TÉCNICO LISBOA

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Seismic risk and real estate prices: an analysis of revealed and stated preferences in Lisbon (Portugal)

João Diogo Fragoso Januário

Supervisor: Doctor Carlos Paulo Novais Oliveira da Silva Cruz
Co-Supervisors: Doctor Humberto Salazar Amorim Varum
Doctor Vítor Faria e Sousa

Thesis approved in public session to obtain the PhD Degree in

Civil Engineering

Jury final classification: **Pass with Distinction and Honour**

2023

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Seismic risk and real estate prices: an analysis of revealed and stated preferences in Lisbon (Portugal)

João Diogo Fragoso Januário

Supervisor: Doctor Carlos Paulo Novais Oliveira da Silva Cruz
Co-Supervisors: Doctor Humberto Salazar Amorim Varum
Doctor Vítor Faria e Sousa

Thesis approved in public session to obtain the PhD Degree in
Civil Engineering

Jury final classification: **Pass with Distinction and Honour**

Jury

Chairperson:
Doctor Jorge Manuel Caliço Lopes de Brito, Instituto Superior Técnico, Universidade de Lisboa
Members of the Committee:
Doctor Rita Maria do Pranto Nogueira Leite Pereira Bento, Instituto Superior Técnico, Universidade de Lisboa
Doctor Carlos Paulo Novais Oliveira da Silva Cruz, Instituto Superior Técnico, Universidade de Lisboa
Doctor Joaquim José Miranda Sarmento, Instituto Superior de Economia e Gestão, Universidade de Lisboa
Doctor João José Lourenço Marques, Universidade de Aveiro
Doctor Maria João Serpa da Lança Falcão da Silva, Laboratório Nacional de Engenharia Civil
Doctor Vítor Emanuel Marta da Silva, Universidade de Aveiro

Funding Institution – FCT: Fundação para a Ciência e a Tecnologia

2023

Resumo

Lisboa tem na sua história diversos eventos sísmicos. Apesar do risco conhecido, os anos recentes demonstraram um crescimento abrupto dos preços do imobiliário em Lisboa, levando investidores e residentes a comprarem propriedades a preços crescentes apesar da vulnerabilidade sísmica do edificado. O mercado imobiliário parece não estar preocupado com as potenciais consequências de um terramoto. Prova disso, é o facto que a maioria das habitações em Portugal não possuem seguro com cobertura de risco sísmico. Este estudo tem como objetivo responder a três questões de investigação: (1) Os valores de mercado refletem uma preferência por imóveis menos vulneráveis a sismos? (2) Como é que os residentes e investidores de Lisboa percecionam o risco? (3) A sua perceção de risco afeta a sua disponibilidade para pagar por um imóvel? Para responder a estas questões foi adotada uma metodologia baseada em métodos implícitos, através da construção de modelos de regressão utilizando dados reais de transações imobiliárias (FGLS, PooledOLS, Fixed-Effects, MARS e XGBoost), e em métodos explícitos baseados em inquéritos à população *(Contingent Valuation Methods)*. Os resultados confirmam as expectativas iniciais do que o mercado é alheio ao risco, mas demonstram que a perceção e ação perante o risco variam entre grupos socioeconómicos, dependendo de fatores como a idade, estado civil e rendimento familiar. Com base nos resultados, são recomendadas várias políticas públicas a adotar para a mitigação do risco sísmico.

Palavras-Chave: Avaliação de Risco, Avaliação Imobiliária, Mercado Imobiliário, Risco Sísmico, Econometria

Abstract

The capital city of Portugal, Lisbon, has a history of suffering from seismic events. Despite the known risk, there has been a recent upsurge in real estate prices in Lisbon, and investors and residents continue to buy properties at increasing prices despite their vulnerability to seismic activity. The market seems to be unconcerned with the potential consequences of an earthquake and most residential properties lack seismic insurance coverage. This study aims to answer three research questions: (1) Do market values reflect a preference for properties less vulnerable to earthquakes? (2) How do residents and investors in Lisbon perceive risk? (3) Does their risk perception affect their willingness to pay for property? The study used both implicit methods, through the construction of regression models on real property transaction data (FGLS, PooledOLS, Fixed-Effects, MARS, and XGBoost), and explicit methods based on surveys (Contingent Valuation Methods). The results confirm the initial expectations that the market is oblivious to risk but show that the perception and action towards risk vary across socioeconomic groups, depending on various factors such as age, marital status, and household income. Based on the results, several public policies to address seismic risk mitigation are recommended.

Keywords: Property Valuation, Real Estate Market, Risk Assessment, Seismic Risk, Econometrics

Acknowledgements

To my parents Fernanda and José, my wife Petra, my family, and everyone who have accompanied me on this journey. A special thank you to Prof. Carlos Oliveira Cruz for all the trust and support. I also thank my co-supervisors, Prof. Vítor Faria e Sousa and Prof. Humberto Varum, for the support given throughout this work, and to Prof. Rita Bento for the guidance throughout the *Infrarisk Programme* that enabled me to explore such fascinating topics. Finally, I would like to thank Fundação para a Ciência e a Tecnologia for the four years of financial support under the PhD grant PD/BD/150402/2019.

Table of Contents

List of Figures

List of Tables

List of Symbols

U(w)	Utility function of wealth
F	Set of feasible alternative investments
I	Investment
I_{opt}	Optimal investment
$X(I)$	Random variable for the ending value of each investment
p_i	Probability of *i*
$V_{EU}(P)$	Expected utility preference function
S_i	Subjective expected utility of outcome
U_i	Utility of *i*
$D(r)$	Level of damage after a certain r measure being implemented
W	Wealth
PR_{it}^{s}	Subjective risk perception
PR_{it}^{o}	Objective risk component
v_{it}	Household's risk measurement error
V_{final}	Final vulnerability
V_i	Initial vulnerability
FM	Modifying factor
$y_{i,t}$	Normalized price for parish i at time t
β	Rgression coefficient
$X_{i,t}$	Matrix of explanatory variables
u_i	Entity-fixed dummy variable
$\varepsilon_{i,t}$	Error term for parish i at time t
α	Constant term
β_{GLS}	Generalized Least Squares regression coefficient
PGA	Peak Ground Acceleration
I_0	Estimated intensity
$V_{u_{req}}$	Required vulnerability
p_{it}	Price for parish *i* at time *t*
BRD	Building Resistance Deficit
u_i	Parish dummy variable
δ	Regression coefficient
ϕ_t	Time-fixed dummy variable
$obj(\theta)$	Objective function
$L(\theta)$	Loss function
$\Omega(\theta)$	Regularizartion term
$RMSE$	Root Mean Squared Error
$\hat{y}$	Predicted value
y_i	Dependent variable
w_i	Weight for i
z	z-score
$\hat{p}$	Expected prevalence (percentage)
ε^2	Confidence interval
WTP	Willingness-to-pay
Y_{300}	Average sale price in a 300m radius

List of Acronyms

(in order of appearance)

IVSC	International Valuation Standards Council
RICS	Royal Institution of Chartered Surveyors
UNDDR	United Nations Office for Disaster Risk Reduction
EM-DAT	International Disaster Database
CRED	Centre for Research on the Epidemionlogy of Disasters
WHO	World Health Organisation
NOAA	US National Oceanic and Atmospheric Administration
GDP	Gross Domestic Product
USD	United States Dollar
ERRMs	Earthquake Risk Reduction Measures
OTM	Out-of-market
USA	United States of America
ISO	International Organization for Standardization
GUT	General Utility Theory
EUT	Expected Utility Theory
SSZ	Special Studies Zones
SEUT	Subjective Expected Utility Model
ART	Alternative Risk Tranfer
CAT	Earthquake Catastrophe Bonds
JERC	Japan Earthquake Reinsurance Company
EQC	New Zealand's Earthquake Commission
TCIP	Turkish Catastrophe Insurance Pool
ICI	Iceland Catastrophe Insurance
CEA	California Earthquake Authority
DRR	Disaster Risk Reduction
CVM	Contingent Valuation Method
WTA	Willingness-to-accept
WTP	Willingness-to-pay
EU	European Union
IMF	International Monetary Fund
OECD	Organisation for Economic Co-operation and Development
LMA	Lisbon Metropolitan Area
ECB	European Central Bank
FED	U.S. Federal Reserve
VAT	Value Added Tax
DECO	Associação Portuguesa para a Defesa do Consumidor
IRS	Personal Income Tax
STR	Short-Term Rental
CML	Lisbon City Council
NRAU	Novo Regime do Arrendamento Urbano
ARI	Autorização de Residªencia para Investimento
SEF	Serfiço de Estrangeiros e Fronteiras
UK	United Kingdom
ATJ	Azores Triple Junction

EC8	Eurocode 8
RSCCS	Regulamento de Segurança das Construções contra os Sismos
PEERS	Plano Especial de Emergência para o Risco Sísmico
GIS	Geographic Information System
LNEC	Laboratório Nacional de Engenharia Civil
DBA	Displacement-Based Seismic Assessment
EC	European Commission
EMS98	European macroseismic scale (1998)
ICIST	Instituto de Engenharia de Estruturas, Território e Construção do Instituto Superior Técnico
RC	Reinforced Concrete
RSEP	Regulamento de Solicitações em Edifícios e Pontes
RSA	Regulamento de Segurança e Ações em Estruturas de Edifícios e Pontes
REBAP	Regulamento de Estruturas de Betão Armado e Pré-Esforçado
SPES	Sociedade Portuguesa de Engenharia Sísmica
APS	Associação Portuguesa de Seguradores
EDA	Exploratory Data Analysis
PCA	Principal Component Analysis
BRD	Building Resistance Deficit
MARS	Multivariate Adaptive Regression Spline
XGBoost	Extreme Gradient Boosting
FM	Modifying Factors
PGA	Peak Ground Acceleration
OLS	Ordinary Least Squares
LISA	Local Indicators of Spatial Association
OMI	Osservatorio del Mercato Immobiliare
MAUP	Modifiable Areal Unit Problem
FE	Fixed-Effects
RE	Random-Effects
GLS	Generalized Least Squares
GMM	Generalized Method of Moments
MLR	Multiple Linear Regression
BLUE	Best Linear Unbiased Estimators
FGLS	Feasible Generalized Least Squares
RAM	Random-access memory
RSS	Residuals Sum of Squares
ANN	Artificial Neural Networks
RMSE	Root Mean Squared Error
EVRI	Environmental Valuation Reference Inventory
VAR	Vector Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
RP	Revealed Preferences
SP	Stated Preferences
BGE	Base de Georreferenciação de Edifícios
AUGI	Urban Areas of Illegal Origin
CBD	Central Business District

1. Introduction

1.1. Context

According to the United Nations (Leilani, 2017), the value of global real estate market is about US$ 217 trillion, nearly 60% of the value of all global assets, with residential real estate comprising 75 per cent of the total. Lisbon, the Portuguese capital city, has seen an upsurge in real estate prices in the period between 2015-2023, leading to a consequent increase in the market value of its housing stock. According to the National Statistics Institute (INE[1]), in 2012, the Lisbon Metropolitan Area (LMA) saw 2.38 billion euros worth of residential properties being transacted. In 2022, that number had increased to 13.26 billion, a 557% increase. Lisbon's seismicity is well known after suffering one of the most devastating earthquakes of 18th-century Europe in 1755. Additionally, 27.3% of residential buildings in the LMA were built before any seismic regulation was in place, and that percentage rises to 63.9% of residential buildings if we only consider Lisbon's municipality (Costa et al., 2008).

Despite the risk, the market seems to be oblivious to it. The city's residents and investors continue to buy properties at increasingly higher prices despite their vulnerability to seismic activity. This is aggravated by the fact that most residential properties do not have seismic insurance coverage and that over 70% of dwellings are occupied by its homeowner. While that number is less significant in Lisbon, around 50% of its dwellings are occupied by their homeowners (CML 2023). This is the case because, like in other Southern European countries, there is a homeownership culture, and acquiring a primary residence is culturally seen as a landmark of reaching adulthood. This leads to most wealth of Portuguese households being held in real estate properties, mainly in the form of primary residences (INE 2020).

According to ISO 31000:2018[2], "risk is usually expressed in terms of risk sources, potential events, their consequences and their likelihood". Regarding the consequences, previous studies have estimated the number of deaths in LMA in more than 12 000, along with several thousand severely damaged or collapse buildings (Costa et al. 2008). On this basis, one should expect the increase risk to be translated into market discounts/premiums for the most/least vulnerable properties. In addition, the added risk makes crucial the study of its perception by residents and investors so measures to mitigate it can be effectively taken. In the end, a clear understanding of the Lisbon's real estate market and of its residents' perception should allow the design of public policies aimed to protect the capital city's inhabitants.

[1] Source: INE
[2] https://www.iso.org/standard/65694.html

1.2. Objective and research questions

This works proposes to address three central research questions:

- R1: Do market values reflect a preference for properties less vulnerable to earthquakes?
- R2: How do residents and investors in Lisbon perceive risk?
- R3: Does their risk perception affect their willingness to pay for property?

The answer to these questions should provide a deeper understanding of Lisbon's reality and serve as a basis for many other cities worldwide where the long-past memory of seismic events puts its housing stock and its households at risk.

1.3. Structure

This work will be structured as follows:

- A comprehensive literature review will present previous studies on natural hazards and their impact on the economy and the real estate market (chapter 2). This section will also cover risk perception, the risk-assessment process made by households and institutions, and their willingness-to-pay. Additionally, an overview of the Portuguese reality and, specifically, Lisbon's case study will be presented, showing its current housing stock, trends, and seismic risk.
- The following section (chapter 3) will be composed of the methodology, which will address the statistical models used to analyze an extensive property transaction database containing over 8 000 transactions in Lisbon from 2008 to 2018 and the analysis of surveys. Building stock, ground motion, spatial and economic considerations will also be addressed in this chapter.
- The results (chapter 4) and the discussion (chapter 5) sections will be divided into subsections according to the research questions they aim to address. One should note that some information serves as the base for multiple research questions.
- The final chapter (chapter 6) will present the conclusions of this work, policy recommendations and future avenues of work.

1.4. Methodology overview

Both revealed and stated preferences were assessed to develop this work and answer the research questions. The former was evaluated through the exploratory data analysis (EDA), containing univariate and bivariate analysis as well as an understanding of the spatial distribution of values. This was achieved through the analysis of central tendency measures for the database attributes, plots using the GeoPandas library in Python, a Principal Component Analysis (PCA), a Feature Importance analysis using the Extreme Gradient Boosting Tree (XGBoost) algorithm, a Best Subset Selection, Moran-I and LISA indicators. After the initial EDA, a set of regression models was constructed to assess the impact of various endogenous attributes, as well as seismic vulnerability, on property prices. Among the methods used, a multiple linear regression model with Feasible Generalized Least Squares estimation, a PooledOLS, a Fixed-Effects model, a Multivariate Adaptive Regression Splines model, and an XGBoost model were implemented.

Then, to assess stated preferences, a survey was conducted on 325 people, aiming to understand their seismic risk perception and their willingness to pay for structural retrofitting and for added seismic resistance. The survey was distributed online to allow for a larger sample and avoid any interviewer bias that could have influenced the results. The results allowed for a deep understanding of the impact of perception on people's awareness and action towards risk. Finally, a joint regressive model was constructed using the results of the stated preferences survey as input variables to estimate property prices in the neighbourhood.

1.5. Main contributions

In conclusion, this thesis aims to explore the relationship between the real estate market, seismic risk, and the perceptions of residents and investors in Lisbon. It will investigate the preference for properties less vulnerable to earthquakes, risk perception, and its influence on the willingness to pay for a property. The research aims to provide a better understanding of the dynamics in the Lisbon real estate market. The findings will contribute to the body of knowledge on natural hazards' impact on the economy and serve as a foundation for developing effective public policies for the city's inhabitants. The results are expected to empower policymakers, residents, and investors to mitigate risks and ensure a safer and more sustainable future for Lisbon and beyond.

2. Literature Review

2.1. Overview

Previous studies have considered the Portuguese capital of Lisbon to have a moderate seismicity risk (Borges et al., 2001). A major earthquake, like the ones of 1531 and 1755, would have significant consequences on the country's economic landscape and the lives of its inhabitants. Many recent studies have focused on the economic assessment of natural hazards and the relationship with the perceived risk (Asgary et al, 2007; Naoi et al., 2009; Modica & Locati, 2016; Peng, 2021). Throughout this chapter, a review of the primary methodologies used for assessing both implicit and explicit preferences involving real estate properties and the perception of risk by investors and homeowners will be carried out. The aim is to understand how the real estate market values natural hazard risks, especially seismic risk, and how they can be quantified. Lastly, some case studies and policy implications will also be revealed.

There are nine main sections in this chapter. After this introductory note, the second section will focus on describing the methodology used for selecting literature. The third section provides a brief description of valuation in the real estate market. The fourth will focus on natural hazards and their impacts on real estate, with a notice on earthquakes. The fifth section will focus on risk perception and the sixth on how this risk is assessed both by households and by institutions. The seventh will present how the willingness-to-pay can be address by explicit and implicit measures. Next, the eight section will present the case study: the economic context, the seismic risk and the vulnerability of Lisbon's building stock (the case study for this research). The ninth and final section will present some concluding remarks from the literature review.

2.2. Review Methodology

The review mainly focuses on the works of the last 20 years, selecting articles published between 2000 and 2020 to build the literature review. Older studies, when considered relevant and helpful to the research using a method of backward snowballing (Jalali & Wohlin, 2012), were also included. The primary databases used during this process were Web of Science, Science Direct, Elsevier and Google Scholar.

For the first part of the review, papers in the English language were considered, using the following keywords: real estate market AND risk, natural hazard risk, property valuation AND risk, seismic risk, post-disaster loss assessment, seismic loss assessment, earthquake insurance, natural disasters AND real estate, risk perception and decision-making process. In this first phase, all geographies were considered. Nevertheless, most papers were found to be related to regions of higher seismic activity, such as the United States (especially California), Iran (e.g., Tehran, Tarom, Zanjan and Rasht), Japan (e.g., Tokyo) and Turkey (e.g., Istanbul). In addition, although the primary focus was on seismic risk, other natural hazards such as floods and typhoons were also considered.

Papers addressing Portugal and Lisbon's Metropolitan Area were analyzed in the second phase. The scope was extended to include works in the Portuguese language using the following keywords: Lisbon real estate market, Lisbon seismic risk, Portugal seismic risk, Lisbon earthquake risk, Lisbon building stock (English); Risco sísmico Lisboa, Mercado imobiliário em Lisboa, vulnerabilidade sísmica, parque habitacional Lisboa (Portuguese).

Summary tables containing the bibliography considered essential from the first two phases of the literature review can be found in Annex A (Table A.0-1 to Table A.0-4).

In the third phase, the biases in the decision-making process and how to address them when delivering questionnaires was analyzed, building up the base for the methodological approach adopted in developing the questionnaire to evaluate risk perception.

The literature review process began at the start of this project and has been updated ever since. Thus, one made an effort to keep up to date with the latest work in the field.

Up to this point, the literature review intended to provide an understanding on how the markets react to natural hazard risk and on how human perception (of risk) is affected by its context, such as the individual's cultural and economic environment. Furthermore, this broader scope would allow to understand the best methodologies to apply to the Portuguese reality. Therefore, the fourth research phase focuses on the Portuguese market and the specificities related to its building stock and geography.

2.3. Real Estate Market

The real estate market is composed by all the real estate property transactions occurring in each location, over a defined timeframe. As mentioned earlier, it constitutes one of the largest asset classes across all capital markets, and is, for most homeowners, the largest purchase made during their entire lifetime. Thus, the dynamics of the real estate have a significant impact on both the government budget (local and/or central government, depending on the location) and on the households' budget. Hence, the study of real estate market cycles is crucial for a deep understanding of larger economic cycles. In order to study market cycles and trends, it is essential to have a clear definition of "valuation".

It is important to distinguish between market value and market price to have a clear understanding of this dissertation. According to the International Valuation Standards Council (IVSC), "market value" is defined (RICS, 2020) as:

"The estimated amount for which an asset or liability should exchange on the valuation date between a willing buyer and a willing seller in an arm's length transaction, after proper marketing and where the parties had each acted knowledgeably, prudently and without compulsion."

As per definition, it is only an estimate of the exchange price if it were to be sold on an open market. The market value is usually attributed by a real estate appraiser, based on a set of assumptions on the property's condition, location, and market conditions (local supply and demand) at the time of the appraisal. It can be understood as a snapshot estimate of its value, based on the application of a valuation method, given it has a temporal validity. The comparable method, the investment/income method and the cost method are some of the most used traditional valuation methods (see Pagourtzi et al., 2003). Note that there are several assumptions taken when valuing a property. In practice, some sales do not meet the assumptions of "proper marketing", and "knowledgeable and prudent" counterparties. Psychological bias also affects both the buyer and the seller. Thus, there might be a gap between the actual market price and the property's valuation.

Market values, such as the ones provided by banking valuations, can nonetheless be used as proxies for the price. Rough estimates of a property's market value are also used to calculate its reconstruction costs after a natural disaster (see Bradshaw (2003)).

Market price ("selling price" or simply "price") is the actual exchange price in the marketplace. In its turn, the market price will then influence the market value of future sales (as in the comparable method). This is the preferred value to be used since it refers to the actual transaction price; hence it translates both the buyers' and sellers' willingness. Authors such as Brookshire et al. (1985), Beron et al. (1997) and Nakagawa et al. (2009) use market prices as variables of interest.

Finally, some authors (see, for instance, Önder et al. (2004)) use listing prices as a reference for modelling. The list price is a property's suggested gross sale price when put on the market. It does not have to equal either the market value or the market price, but it constitutes a proxy for the market price, usually 1%-10% above.

2.4. Natural Hazards

2.4.1. Definition

The United Nations Office for Disaster Risk Reduction (UNDDR) defines hazard as a process, phenomenon or human activity that may cause loss of life, injury or other health impacts, property damage, social and economic disruption, or environmental degradation. When a natural hazard affects a region, causing significant damage to the population affected, it classifies as a natural disaster, that is, a severe disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources. Natural disasters affect millions of people across the world every year. Much scientific research has been devoted to understanding these phenomena, with the goal of increasing the ability to predict and prepare for natural hazards (Cavallo & Noy, 2011).

This chapter uses data from the International Disaster Database (EM-DAT), an extensive resource on natural disaster records created in 1988 by the Centre for Research on the Epidemiology of Disasters (CRED), at the Catholic University of Louvain, with the support of the World Health Organisation (WHO) and the Belgian Government. This database contains records of natural disasters occurring worldwide ranging from 1900 to 2021. It contains more than 22 000 records, and it will serve as base for the analysis.

Before analyzing damage costs, it is essential to distinguish between direct and indirect damages. According to Bradshaw (2003) direct damages occur at the moment of the disaster or within the first few hours. However, depending on the magnitude of the disaster, this time frame may be extended to a period of up to five years. These direct damages account for damage to infrastructure, property, immovable assets on stock (including final goods, goods in process, raw materials, materials and spare parts), livestock, mortality and morbidity that are a direct consequence of the natural phenomenon (i.e., an earthquake, a flood, or a drought). The estimated damage shown in EM-DAT database can be classified as following these criteria. However, one should remember that there are also indirect damages from a natural disaster. Indirect damages refer to economic activity, particularly the production of goods and services, which will not occur after the disaster and because of it (Cavallo & Noy, 2011).

2.4.2. Occurrence

Floods are the most recurring natural disasters, with 34.4% of all natural disaster occurrences in the database, averaging 45 events every year. Storms and earthquakes follow in the number of occurrences with 27.91% (at a rate of approximately 37 events per year, on average) and 9.63% (at a rate of approximately 13 events per year, on average), respectively. Figure 2-1 shows the distribution of natural events by type during the available period.

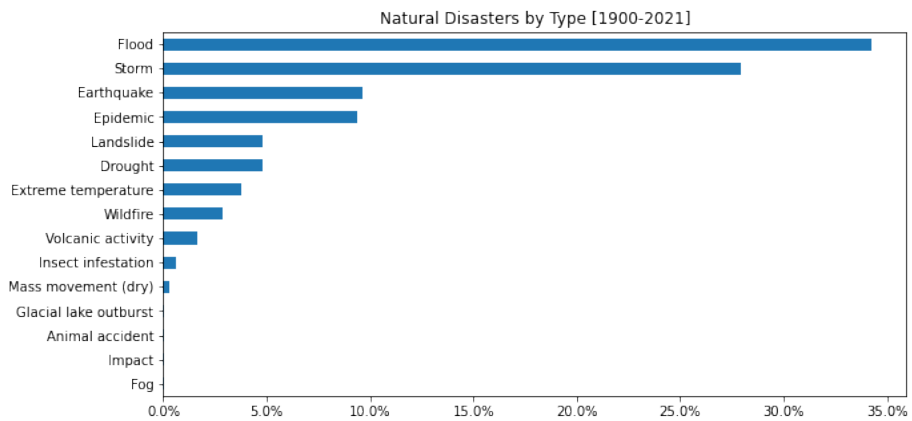


Figure 2-1 - Natural Disasters (1900-2021). Data Source: EM-DAT

As seen in Figure 2-1 floods are the most common natural disaster. However, they are not the one causing the most damage. Storms cause, on average, 13.32 billion dollars in damages each year, killing more than eleven thousand people and affecting the lives of more than ten million. It is worth noting that the value of real estate assets also significantly influences the overall damages' economic assessment. Therefore, the same material damages can have different financial valuations depending on the country of their occurrence. Table 2-1 shows the yearly averages of damage costs, deaths and the number of people affected by the disaster. Note that, while less frequent, at only 4,77% of all occurrences, droughts are the deadliest, having claimed the lives of more than 11 million people over the past 121 years.

Table 2-1 - Natural Disasters summary table (1900-2021). Data Source: EM-DAT

Disaster Type	Total Damages ('000 US$)	Total Deaths	Total Affected	Most Affected Country[3]
Storm	13 323 970.00	11 576	10 284 020	USA
Flood	7 316 628.00	57 784	32 044 550	China
Earthquake	6 938 941.00	19 337	1 658 108	China
Drought	1 505 827.00	96 954	22 867 140	China
Wildfire	984 911.40	37	58 968	USA
Extreme temperature	522 862.30	1593	856 492	India
Landslide	90 982.93	553	120 862	China
Volcanic activity	42 891.83	717	79 128	Indonesia[4]
Insect infestation	1 901.86	0	23159	Niger
Mass movement (dry)	1 727.27	38	228	China
Glacial lake outburst	1 702.48	2	0	India
Impact	272.73	0	2492	Russia
Epidemic	0,06	79 451	408 825	D.R. Congo
Animal accident	-	0	0	Niger
Fog	-	33	0	UK

Geography is one of the most critical factors contributing to a country's exposure to natural hazards. Japan, for instance, is one of the world's most earthquake-prone countries since it lies at the nexus of four tectonic plates (Naoi et al., 2009) in a region called the Ring of Fire[5]. The active faults near Tehran (Iran) (Asgary et al., 2007), in California (USA) (Brookshire et al., 1985), or the ones in the West of Chengdu (China) (Deng et al., 2013) are also responsible for several seismic events. Several hurricanes also hit Florida (USA) due to its location (see Pielke et al. (2008)). Disasters are often described as a combination of hazard, exposure and vulnerability. However, measures to reduce or cope with the potential negative consequences significantly

[3] By number of occurrences
[4] Indonesia is notoriously affected by the Sunda Arc, the region where the Indian-Australian plate meet the Southeast Asian plate (see for instance Curray (1989))
[5] https://www.britannica.com/place/Ring-of-Fire

impact the aftermath. The socioeconomic conditions of a country also play a significant role in the outcome of a natural disaster. According to Sen (1981, as cited in Cavallo & Noy (2011)), the very occurrence of natural disasters are economic events. Therefore, one should expect that countries with higher development and financial stability would be more resilient to economic shocks caused by natural hazards. Toya and Skidmore (2007) argue that countries with higher income, higher educational attainment, greater openness, smaller government and more complete financial systems experience fewer losses due to higher implementation of precautionary measures to reduce the impacts of natural disasters. Higher levels of income also increase private demand for safety. As an example, we may look to the effects of the 2011 Haiti earthquake, a 7.0 in Richer scale earthquake which struck the densely capital of Port-au-Prince, causing between 200 000 and 300 000[6] deaths and a massive impact on the economy (over 100% of country's GDP). In 2009, the country registered a GDP per capita of $1 172,1 USD[7]. One month after the event, in February 2010, Chile suffered an 8.8[8] earthquake near its capital city, Santiago. With more robust financial stability, the country registered a GDP per capita of $12 227,2 USD and suffered less than 1000 deaths (Cavallo & Noy, 2011).

However, a natural disaster's economic impact cannot be measured only by the number of deaths or total damage costs. These events affect the lives of many people and economic sectors: from agriculture and industry to housing and tourism. In this dissertation, the author will focus on one specific economic sector: real estate. Disasters often result in property damage and asset destruction (Bennett,1989; Beron et al., 1997; Badri et al., 2006; Lee et al., 2017). The following chapter will review the existing literature on the impacts of natural disasters on real estate.

2.4.3. Impacts

Many authors have studied the impact of several types of natural disasters on real built assets: floods (Harrison et al., 2001; Bin & Polasky, 2004; Daniel et al., 2009), earthquakes (Brookshire et al., 1985; Beron et al., 1997; Tekeli-Yesil et al., 2011; Deng et al., 2013; Modica & Locati, 2016; Fekrazad, 2019), wildfires (Donovan et al., 2007; Kiel & Matheson, 2018), storms and hurricanes (Hallstrom & Smith, 2005; Ewing et al., 2007; Burrus et al., 2009; Zhang et al., 2010; Asgary & Halim, 2011; Lee et al., 2017).

The damage to property caused by natural disasters depends on the type of natural disaster and the type of construction. Earthquakes typically damage structural elements (beams, joists, panels, load-bearing walls, etc.) and non-structural elements (partition walls, non-structural roofs, furniture, installations, equipment, etc.). Intense storms and hurricanes exert extraordinary pressure on buildings and may lead to structural and non-structural damages (Bradshaw, 2003). Floods may affect a building's structure or bury it in mud, leading to equipment and furniture loss

[6] Estimated number of deaths varies between sources. While the Haitian government as claimed that more than 300.000 people lost their lives (see https://www.britannica.com/event/2010-Haiti-earthquake), other sources claim fewer numbers (see as an example Cavallo (2011))
[7] source: https://data.worldbank.org/indicator/NY.GDP.PCAP.KD?end=2009&start=1960
[8] In Ritcher scale

and rendering it useless. However, across the literature, the general belief is that households' and markets' reaction is more closely linked to the shock rather than to the type of natural disaster.

Bin and Polasky (2004) studied the impact of Hurricane Floyd on the housing market of Pitt County, North Carolina (USA). They found out that properties within the floodplain had a lower market value than others *ceteris paribus* and that the occurrence of this hurricane only aggravated the risk premium. Daniel et al. (2009) reviewed the results of several studies on the effect of floods over multiple locations in the US housing market and concluded that an increase in the probability of flood risk of 0.01 in a year is associated with a difference in transaction price of an otherwise similar house of –0.6%. Donovan et al. (2007) studied the impact of wildfires in Colorado Springs, Colorado (USA) and found a negative effect on housing prices, in the surrounding area, immediately after the event but diminished over time. Hallstrom and Smith (2005) studied the impact of a "near-miss" Hurricane Andrew (1992) on the housing market of Lee County, Florida (USA). They concluded that the "near miss" experience led to a 19% decline in housing prices due to the increased awareness of homeowners of natural hazards. Finally, Zhang et al. (2010) analyzed the impact of several hazards (flood, hurricane and toxic chemicals) and concluded that risk perception is a mediating factor between hazard proximity and property value. Thus, an increased notion of proximity to a hazard will lead to a discount in the property value.

The general belief across the literature is that the initial shock caused by a natural disaster, even if it is a "near-miss" or out-of-market event, causes a change in the households' risk perception leading to a decrease in housing prices, especially in the properties assessed by households as of greater risk. However, some authors have found evidence of slight increases in housing prices due to initial overestimation of damages (see for instance, Beron et al. (1997)) or even negligible effects on housing prices. Several factors should be considered, providing plausible explanations for the variety of results. Previous knowledge of risk, people's attitudes towards risk and their cultural environment can alter individuals' risk perception and their assessment of whether they may be personally affected by a natural disaster (Palm, 1998; Asgary & Willis, 1997; Peng, 2021). Hence, to understand the effect of natural hazards in the real estate market, it is critical first to understand human risk perception and the limitations of human cognition (Pryce et al., 2011) as well as the households' assessment of risk. A deep understanding of these foundational concepts will lead to a better understanding of the actual impact in the market and how governments should effectively communicate and disclose "hazardous information".

A summary table (Table A.0-2) containing the studies mentioned above and their key findings can be found in Annex A.

2.4.3.1. Earthquakes' Impact

According to the International Disaster Database (EM-DAT)[9], yearly, earthquakes have caused an average of 19 336 deaths and almost 7B$ in total damages since 1900. The most affected countries in the world by number of occurrences are China (12,75% of total events), Indonesia (8,69%) and Iran (8,5%) (Figure 2-2).

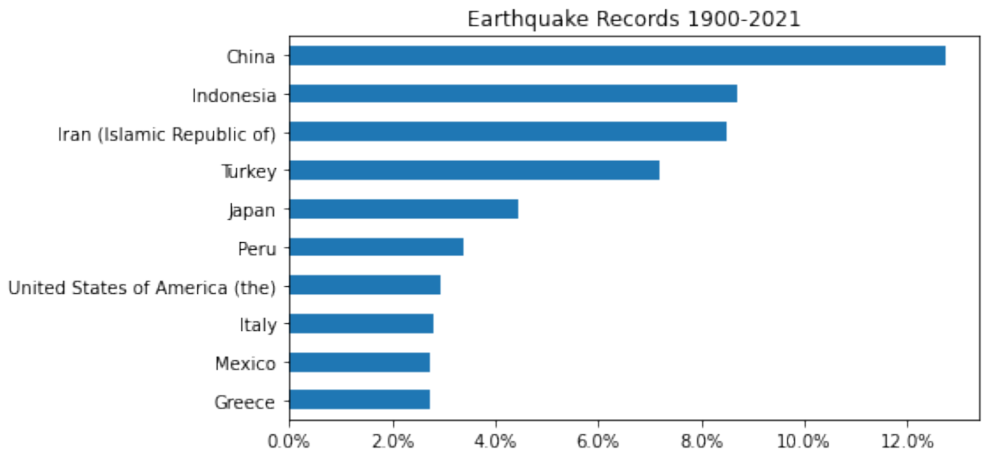


Figure 2-2 - Top 10 countries by percentage of global earthquakes (1900- 2001). Data Source: EM-DAT

China (652 419 deaths), Haiti (222 593 deaths[10]) and Indonesia (203 460 deaths) present the highest death toll, while Japan (384 B$), China (112 B$) and Italy (55 B$) have suffered the most significant total damages. Figure 2-3 and Figure 2-4 show the geographical distribution of deaths and total damage.

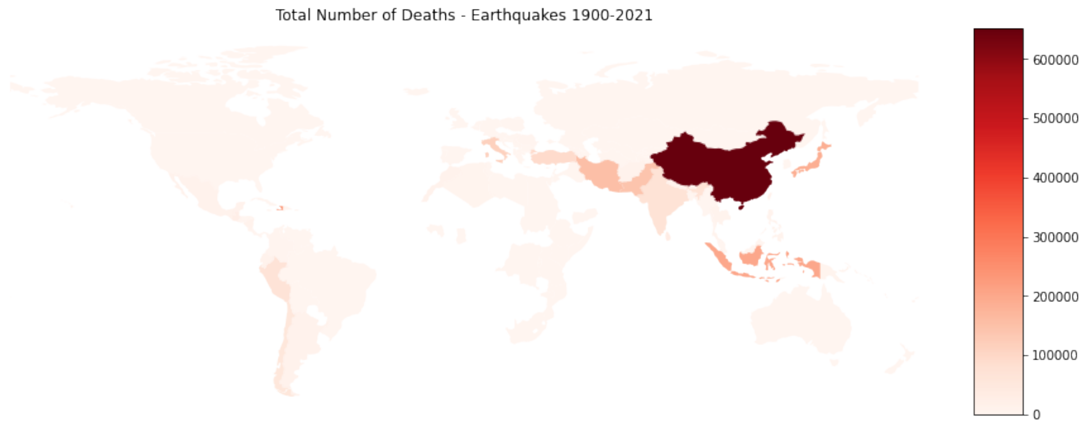


Figure 2-3 – Deaths due to earthquakes (1900-2001). Data Source: EM-DAT

[9] https://www.emdat.be/ ; For an event to be registered in the database it has to fulfil at least one criteria of the following: (a) ten or more people reported killed; (b) Hundred or more people reported affected; (c) Declaration of a state of emergency; (d) Call for international assistance.
[10] In a single earthquake on the 12th of January 2010

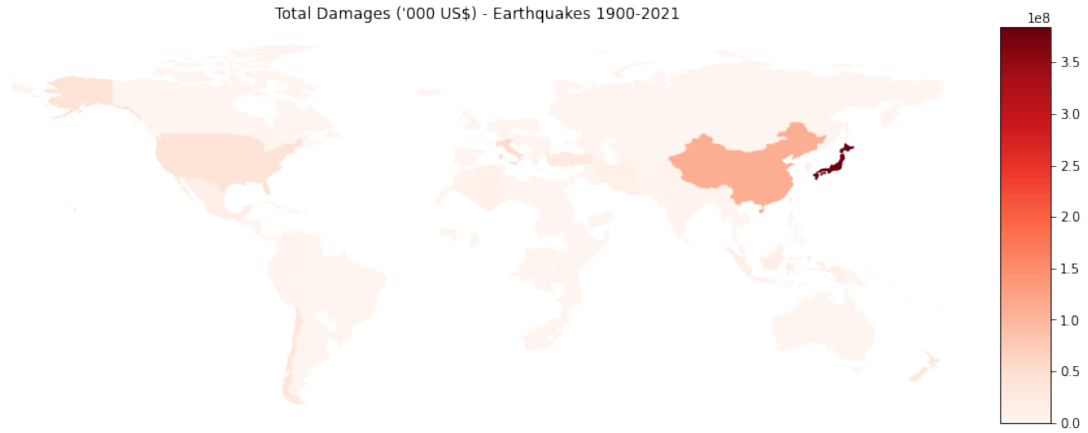


Figure 2-4 – Damages due to earthquakes (1900-2001). Data Source: EM-DAT

The deadliest earthquake[11] was the 1976 Tangshan earthquake, claiming 242.000 lives, and the costliest event was the 2011 Tohoku earthquake and tsunami in Japan, with total damage surpassing 210 B$. It is no surprise that most of the scientific literature concerning the impact of earthquakes on real estate relates to the most affected geographies. Naoi et al. (2009) studied the effect of earthquake risk in the Japanese real estate market, a country with an annual average of 1.14 events equal to or greater than a magnitude of 5.5 on the Ritcher scale between 1980 and 2000. Using nationwide panel data, Naoi et al. (2009) found some modifications to individuals' earthquake risk assessment after a seismic event. Although they found no clear evidence of a negative correlation between the valuation and regional earthquake risk in pre-quake periods, they found a significant correlation in post-quake periods.

Households can hedge against earthquake risk by purchasing insurance policies. Naoi et al. (2009) also stated that earthquake damage is not fully insured in Japan, as in other countries. According to EM-DAT, only 17.8% of the total damage was covered by insurance during the 2011 Tohoku earthquake and tsunami, confirming the findings by Naoi et al., (2009). The author also argued that "*since an earthquake is an exogenous risk factor that is tied to a specific location, its risk should be capitalized into local housing and land prices. Estimating individuals' risk is crucial for evaluating benefits of earthquake damage mitigation policies*". Throughout the literature, there is evidence that increased risk factors should have negative impacts on price and offering a discount on riskier properties is generalized across the literature (Brookshire et al., 1985; MacDonald et al., 1987; Willis & Asgary, 1997; Önder et al., 2004; Bin & Polasky, 2004; Hallstrom & Smith, 2005; Nakagawa et al., 2007; Fekrazad, 2019).

Nakagawa et al. (2007) analyzed the Tokyo real estate market and found evidence that the disclosure of a building-specific earthquake risk index, compiled by the Tokyo Metropolitan Government, led to lower land prices in areas with disclosed substantial exposure to earthquake risk. Two years later, Nakagawa et al. (2009) analyzed land prices in Tokyo during the 1980's and 1990's and found that households had seemed to become increasingly aware of earthquake risk

[11] According to EM-DAT database. Note there is no consensus on the number of deaths on the Port-au-Prince's earthquake (2011)

throughout the last decade. Land prices in riskier locations showed an average 8% discount against that of safest locations. Unlike previous papers, the authors found no evidence of a statistically significant impact of the Great Hanshin-Awaji Earthquake or the disclosure of the hazard map. Naoi et al. (2009) found out that the price discount from locating in a quake-prone area in Tokyo, is significantly more considerable soon after an event than beforehand. Hidano et al. (2015) also analyzed the Tokyo housing market's response to disclosed information on earthquake resistance and found that properties in low-risk zones were priced at a premium of 13 970–17 380 JPY relatively to properties in higher-risk zones. The author also concluded that prices of newly constructed properties with higher structural safety standards were not significantly affected by risk, contrary to their older counterparts.

In China, Deng et al. (2013) analyzed the impact of the 2008 Wenchuan earthquake to analyze the pricing behaviour of units on lower versus upper floors before and after the earthquake. They found that average housing prices decreased after the tremor and that high-floor units were subject to the most significant discounts. Also, the relative prices of low (first and second floor, mainly) to high floor units, considerably increased after the earthquake. This is due to the belief that living on lower levels would allow people to exit the building faster. However, the authors state that there is no evidence of higher survival rates among people living on lower floors. The authors argue that this effect may have resulted from an overreaction of individuals to the dramatic event, which is consistent with the "self-insuring" behaviour of other rational individuals living in hazardous areas. They point out that this overreaction to dramatic events has also been observed across the financial markets, leading to significant corrections in the stock market, as documented by several authors (see for instance, Brooks et al. (2003) and Ederington and Lee (1993, 1995)). According to the authors, the potential fear driving this price differential should dissipate eventually, leading house prices to return to normal levels as people realize that the probability of recurrence of such an event is very small.

In Iran, Willis and Asgary (1997) found a significant price difference between earthquake-resistant houses across all districts in the city of Tehran. The authors argue that this price difference might be further increased with additional disclosure of information about earthquake risk. However, this sensitivity varies between city areas and socioeconomic groups. Despite the awareness in the post-quake period, Earthquake Risk Reduction Measures (ERRMs) had not been widely used due to the increased construction costs. Therefore, legal measures and recent earthquakes only have a short-term impact on the housing market.

Önder et al. (2004) used listing prices from newspapers in 1995 and 2000 to analyze the impact of the 1999 Kocaeli Earthquake on Istambul's housing market. They concluded that the distance from active fault lines is an essential factor explaining house values in the Turkish city and that the event had increased its impact. However, the authors argue that the distance to fault lines, as an explanatory factor, dwindles when controlling for other neighbourhood characteristics. They have also found out that the warnings about the soil type resistance, previously made by geologists, had only been of significance to high-income locations, contributing to the idea that the levels of awareness and tolerance to risk also depend on the socioeconomic environment.

In Italy, Modica and Locati (2016) studied the housing market response to the 2012 Northern Italy earthquake and found evidence of overreaction due to the level of damage produced. There was an average of 4.6% price reduction after the shock, and the regions most affected by the earthquake were also the most affected by this price reduction. It is noteworthy that villas with better structural quality were less affected by the price drop, suggesting that households have factored in the perception of resistance to earthquakes.

Brookshire et al. (1985) found significant impacts on land pricing in the United States after the State Government of California disclosed a hazard map made available by a 1974 state law. The author argued that this disclosure created space for a new market for safe housing in California. Contrary to Brookshire, Beron et al. (1997), in the aftermath of the 1989 Loma Prieta earthquake, found evidence that households had initially overestimated the risks, leading to a downward revision in the market value of earthquake risk in the post-quake period. Palm (1990), as cited in Nakagawa et al. (2007), showed that legislation did not significantly affect the housing market, as most appraisers did not give any higher price on a comparable property in a lower risk area. The author also found that clients had only rarely asked about the seismic hazard, and few purchasers checked for evidence of previous damage from earthquakes.

McGinnis (2004) also analyzed the effect of the 1989 Loma Prieta and 1994 Northridge earthquakes. The author concluded that both events had impacted the consumers' behaviour to earthquake risk over time, thus increasing the risk effect on prices; however, he stated that this impact tends to wear over time.

Fekrazad (2019) studied California's housing market response to the news of out-of-market (OTM) earthquakes. He found a decrease of 3% to 6% in the home value index and median listing prices, respectively, for zip codes with high seismic hazard risk, after high-casualty events (with a death toll of 1000 or more) outside of California. The author also found that higher death tolls would lead to more significant increases in the price differential and that an earthquake in Europe significantly impacts California's market more than a similar event in another region. This "affection reaction" to events happening to people that we consider similar to ourselves, in what might be a manifestation of affinity bias[12] or unconscious bias[13], leads to a higher price differential. In essence, the news of these events reminded households of their risk exposure. Also, such events may receive increased media coverage, contributing to risk awareness. This has resulted in a transfer of bargaining power from the sellers of high-risk properties to the buyers.

[12] Affinity bias is the unconscious tendency to get along with others who are like us. It is easy to socialize and spend time with others who are not different. However, it requires more effort to bridge differences when diversity is present.

[13] Unconscious bias (or implicit bias) is often defined as prejudice or unsupported judgments in favour of or against one thing, person, or group as compared to another in a way that is usually considered unfair.

However, the author also noted that the effects of OTM news dissipate one month after the earthquake occurs.

If the government correctly assesses the earthquake risk and applies policies promoting public disclosure and educational campaigns on the risks, the occurrence of an event should not substantially alter their perception. The increased transparency will benefit the population and the real estate market, leading to a preference for more resistant buildings and less quake-prone areas and, consequently, to lower values of damage and lower price discounts in the real estate market. Since legal and educational measures have short-term effects on the market, educational campaigns should not be a one-time event. They should be repeated across time. If households overestimate the potential damage, the occurrence of a mild seismic event may even lead to a slight increase in prices as demonstrated by Beron et al. (1997).

A summary table (Table A.0-2) containing the studies mentioned above and their key findings can be found in Annex A.

2.5. Risk Perception

ISO 31000 (2009) defines risk as the "effect of uncertainty[14] on objectives" (ISO 3100). It is generally conceived as consisting of two components: the likelihood and the severity of negative outcomes (Van Der Pligt, 1998). Risk is usually assessed from a quantitative point of view, through statistical tools, as in the case of insurance companies, but is also perceived by households in intuitive risk-judgements based on media coverage and social influences transmitted by family, friends and public officials (Slovic et al., 2016). The previous section presents numerous examples of how household risk perception can impact property market valuations.

As human beings, we frequently assess risk empirically in our day-to-day life. However, the risk is usually poorly determined, especially in low-probability events (Kahneman & Knetsch, 1992); this occurs because we are subject to heuristics and biases, rules of thumb and empirical knowledge from past experiences misguiding our quantitative risk estimates. These biases also result from our cultural background, age, gender, educational level, and income levels (Pidgeon et al., 2003; Önder et al., 2004; Tekeli-Yes et al., 2011; Lindell & Perry, 2012). As Wolff et al. (2019) states, it "*does not matter whether someone judges the risk of going to the dentist or going to Mallorca, they will still be influenced by the same heuristics and biases*". Moreover, the perception of hazard is not only limited by the uncertainty of hazard occurrence, but also by the limitations of human cognition (Pryce et al., 2011; Peng, 2021). In sum, many factors may amplify or attenuate our perception of risk. All these affect the purchaser's decision-making process.

[14] According to ISSO 31000 (2009) *"Uncertainty is the state, even partial, of deficiency of information related to, understanding or knowledge of an event, its consequence, or likelihood."*

This chapter will provide an overview of households' perception of risk and the decision-making process, as well as some biases and heuristics that may come into play in assessing of risk, and, ultimately, for the case of this research, in pricing of real estate.

There is a common belief across psychologists that the primary sources of human behaviour result from people's cognitive perceptions, which result from human interactions (social and cultural). Thus, mere exposure to information will not automatically lead to a change in behaviour, and therefore the problem is not limited to a lack of information (Van Der Pligt,1998; Fekrazad, 2019).

Many researchers have identified several ways individuals resist information about their vulnerability to harm (see for reference, Weinstein et al. (2000)). We may hypothesize that this lack of behavioural change upon risk information may also arise from individuals' natural tendency to "bury their head in the sand" and avoid or reject negative details, a phenomenon known in the psychology literature as the "Ostrich Effect" (see, for reference, Karlsson et al. (2009)).

Asgary and Willis (1997) state that the process of risk perception can be modelled as a sequence of hear-perceive-respond. People personalize what they hear in what may be called "perceived vulnerability". Then, based on their assessment, they respond by taking or not taking action. Perception bias refers to the divergence between the objective probability of a given risk and an individual's perception of the risk (Daniel et al., 2009). This bias rising from their subjective perception may be one of the causes why individuals may not be able to perceive actual risks based on expert information or news coverage (Florax et al., 2005). Informing the general public based on scientifically estimated risk does not always promote appropriate risk perception. This leads to distorted perceptions, leading people to strongly oppose some hazards that experts find to be of lower risk while ignoring riskier hazards. The frequency of exposure to information may also influence perceived vulnerability, given that individuals will often violate Bayes' rule, overweighting recent information and underweighting prior data (Zhang et al., 2010; Deng, 2013). This phenomenon is commonly referred to in the psychology literature as "salience bias" (see for reference Tiefenbeck et al. (2018)). Lobb et al. (2012, as cited in Fekrazad (2019)) found that following the 2012 Haiti earthquake, Twitter posts, press releases and charitable donations peaked after one week. The same author argued that individuals might forget the event within one month, thus, the effects on their response to risk may also fade out.

Generally, there is a belief that households tend to underestimate earthquake risk if there has not been a recent occurrence (Naoi et al., 2009). However, natural disasters can dramatically change individuals' perceived vulnerability. For example, Bin and Polasky (2004) found out that the extensive scale damage caused by Hurricane Floyd led to increased awareness and a 24% increase in sales of flood insurance policies in North Carolina. Experiencing events such as floods, earthquakes, or hurricanes may lead to prolonged posttraumatic stress disorders, panic attacks, recurrent nightmares, and survivor-guilt feelings (de la Fuente (1986), as cited in Dooley et al. (1992); Wood et al., 1992). Large-scale disasters have been shown to create a feeling of vulnerability and decrease unrealistic optimism, even among those who were not victims of the event. Even out-of-market or "near-miss" events may change individuals' perceived vulnerability since people tend to learn from comparable circumstances (Weinstein et al., 2000; Hallstrom & Smith, 2005; Fekrazad, 2019). In earthquakes, the distance to the fault and the degree of ground shaking experienced in the same event will also lead to differences in risk perception in the post-quake period (Deng et al., 2013). Different levels of damage during the event have also been found to translate into different levels of perceived risk (Modica & Locati, 2016).

Overall, the literature demonstrates that households are initially unaware of earthquake risk or at least underestimate it in most locations. A seismic event produces an emotional shock that forces the perceived risk to be better aligned with the objective probability and associated consequences. Regions like San Francisco, with a higher public disclosure of seismic risk, may present different initial market conditions, even resulting in an overestimation of the risk (Beron et al., 1997). In addition, a massive earthquake in neighbouring regions or similar socioeconomic markets abroad may alter the household's perception. As people learn from comparable circumstances, this may result in post-quake discounts for property values within quake-prone areas and less resistant buildings. Markets may be distorted by any event changing the perceived importance of previously ignored risks (Lamond et al., 2010).

Despite a high perceived vulnerability, households may decide not to take mitigation measures due to their fatalistic attitude (Asgary & Willis, 1997). They will not take action if they perceive that there is nothing they can do to alter their fate. Some authors argue that, even in the case of increased information, if individuals have low expectations about dealing with risk, they will have a fatalist attitude which will translate into decreased intentions to behave adaptively (see for reference Van Der Pligt (1998)). Conversely, people tend to construct their own theory of why some people or places are at risk. Weinstein et al. (2000) argue that this "loss of perceived behavioural control" may be shattered by the occurrence of a nearby disaster resulting in a decreased bias. Strong emotions have also been proven to lead people to neglect the small probability that the risk will come to fruition (Sunstein, 2001).

Burton et al. (1978) suggest households follow four patterns of response to natural hazards risks:

a. *Denying behaviour* – individuals try to deny risk or convince themselves they will not be harmed by the hazard; This attitude will not lead to preparedness or mitigation measures; This may arise from comparative optimism or unrealistic optimism as people show a tendency to claim they are less likely than their peers to suffer harm (Van Der Pligt, 1998; Weinstein et al., 2000).
b. *Acceptance of loss behaviour* – individuals are aware of the risk but are resigned to accept it (fatalism). They feel there is nothing they can do and therefore respond passively. They may feel that the event is random and uncontrollable or that it is the result of the will of God; therefore, they believe they are unable to reduce risk.

c. *Practical behaviour* – individuals believe their actions can reduce their risk exposure. Consequently, they take preparedness and mitigation measures such as: purchasing earthquake insurance, holding family earthquake drills, choosing a suitable site for the dwelling, fixing high furniture securely and storing basic food and water supplies.

d. *Extreme behaviour* – individuals take extreme measures to avoid hazard risk. As an example, they may build their homes with stricter security levels to seismic events than the ones recommended by building codes.

It is noteworthy that the type of response depends not only on the person (cultural context and socioeconomic condition) but also on the time of the survey. In addition, any individual may change his response to natural hazards based on previous experience.

Education can also be crucial to an adequate response. Hurnen and McClure (1997) found that higher levels of knowledge about the specific causes of damage, and the ways damage can be reduced, are correlated with earthquake preparedness. Additionally, there is evidence that a knowledge-based education strategy is likely to reduce fatalism to earthquakes. Asgary et al. (2007) also found out that the more education and children the respondents have, the more they are willing to pay for earthquake preparedness and mitigation measures. Donovan et al. (2007), while studying wildfires in Colorado Springs, also argues that a considerable portion of the observable changes in households' risk perception can be attributed to educational programs.

Additional to education, age, prior experience with earthquakes, income[15], marital status[16], having children in the household, and length of residence have also been found to positively correlate to the levels of preparedness (Asgary & Willis, 1997; Dooley et al., 1992).

2.6. Risk Assessement

2.6.1. Households

The decision-making process of households has been widely modelled within a Utility Functions framework. These are often used to assess how much risk an investor is willing to take to attain more wealth or, in other words, to measure an investor's relative preference for different levels of total wealth. Utility Functions are part of the General Utility Theory (GUT) framework and have been used to evaluate scenarios of social perception of risk or pure financial risk assessment (e.g. modern portfolio theory) (Asgary & Willis, 1997b; Norstad, 1999).

Utility functions are twice differentiable functions of wealth $U(w)$ (defined for $w>0$) and follow two distinct properties:

[15] According to Asgary and Willis (1997b) safety is a function of income and wealth. Hence higher-income individuals are more willing to pay for additional safety measures.
[16] Dooley et al (1992) found that married individuals are 70% more likely to adopt preventive measures and 20% more likely per each decade of residence in the present home.

a. *Non-satiation* – the investor is never satisfied with his level of wealth; hence he has never so much wealth that getting a little more would not be desirable. Formally, this translates into *U'(w) > 0*.

b. *Risk aversion* – the utility function is concave, therefore the marginal utility of wealth decreases as wealth increases. In practice, a higher income person will attribute less extra (marginal) utility to earn a single euro than a lower income person would. Formally, this translates into *U''(w) > 0.*

It is noteworthy that different types of investors will have different types of utility functions, but they should always comply with the abovementioned properties.

The principle of expected utility maximization (Norstad, 1999) states that, in the face of multiple choices, the rational investor will always choose the one that maximizes his wealth. Formally, this can be formulated as (EQ 1):

$$E\left(U\left(X\left(I_{opt}\right)\right)\right) = \max_{I \in F} E\left(U\left(X(I)\right)\right) \qquad \text{[EQ 1]}$$

where *F* is a set of feasible alternative investments I and $X(I)$ is the random variable providing the ending value of each investment *I* for the period in question. Faced with this choice, the rational investor will choose his optimal investment I_{opt} .

The Expected Utility Theory (EUT), first introduced by Nicolas Bernoulli in the 18th century, is a subset of the GUT and has been one of the most cited frameworks to model the decision-making process. Throughout the 20th century, several authors have built upon this theory[17], most notably Frank Ramsey (1926), von Neumann and Morgenstern (1944), Friedman and Savage (1948), Markowitz (1952), Arrow (1965) and Pratt (1964). According to this theory, individuals act rationally in their decision-making process. It bases on the assumption that individuals would choose a level of risk in which the expected marginal benefits are equal to the expected marginal costs, measured according to the probability of each event. This probability can be objective (based on surveys) or subjective (perceived probability). Formally it can be modelled as a choice between finite-outcome *objective lotteries* (Machina, 2008), yielding a reward x_i with a probability $0 \le p_i \le 1$, such that:

$$P = (x_1 p_1; \dots; x_n p_n)$$

with:

$$\sum_{i=1}^{n} p_i = 1$$

[17] For further information see https://plato.stanford.edu/entries/decision-theory/

The *expected utility preference function* can be defined as (EQ 2):

$$\begin{aligned} V_{EU}(P) &\equiv V_{EU}(x_1 p_1; \ldots; \; x_n p_n) \\ &\equiv U(x_1)p_1 + \cdots + \; U(x_n)p_n \end{aligned} \qquad \text{[EQ 2]}$$

According to the EUT, an individual will prefer option P* over P if and only if $V_{EU}(P^{*}) > V_{EU}(P)$. As it derives from the GUT, EUT also follows the *non-satiation* and *risk aversion* criteria. Von Neumann and Morgenstern (1994) introduced the *independence axiom* to EUT, among others[18], which states that if two options (or gambles) are mixed with a third irrelevant one, the order of preference between the original two options will remain intact.

Brookshire et al. (1985) have argued that the Expected Utility (EU) model is an appropriate descriptor of consumers' behaviour when facing a low-probability high-loss event such as an earthquake. The author analyzed the Californian real estate market and concluded that households process probability information reasonably, specially *"in a market situation with a well-defined institutional mechanism".* Consequently, he found a gradient in housing prices inside and outside the Special Studies Zones[19] (SSZ) delimitation.

However, this formulation has several issues, as modelling human decision-making as a purely rational process can often be inadequate. Von Neumann and Morgenstern (1944) recognize from the start that "*one of the chief difficulties lies in properly describing the assumptions which have to be made out of the motives of the individuals*". Kahneman and Tversky (1979) also argued that the independence axiom is systematically violated in the case of low-probability high-damage events such as earthquakes. Even the linearity assumption, which derives from the *non-satiation*, has been questioned (Lopes, 1994). For instance, one would hardly trade the possibility of a high value option (e.g. winning 1 million euros) for an almost certain chance of winning a larger amount (e.g. 5 million euros). This has been known in the literature as the Allain Paradox (Incekara-Hafalir et al., 2021). Palm (1981) also found evidence of general unwillingness of individuals to insure or engage in preventive expenditure against low-probability high-loss situations such as earthquakes.

Leonard Jimmi Savage (1990) built upon the EUT model to focus on subjective probability assessment. The Subjective Expected Utility Model (SEUT) was innovative by focusing on psychological assessment, bringing a closer approximation to human behaviour during decision-making. According to the SEUT approach, individuals build their assessment not upon objective, quantifiable probabilities but on subjective assessments resulting from their experience, cultural and socioeconomic background and available information. Although innovative, this model follows the same axioms as EUT (Asgary & Willis, 1997b). This model can be expressed a function of

[18] The authors also introduced the *completeness*, *transivity* and *archemedian (continuity)* axioms to the EUT.
[19] Delimitation of high-risk seismic zone in San Francisco after the Alquist-Priolo Special Studies Zone Act (1972)

the subjective expected probability of each outcome P_i and its utility U_i. The subjective expected utility of outcome i is then given by S_i such that (EQ 3):

$$S_i = P_i U_i \qquad \text{[EQ 3]}$$

Asgary and Willis (1997b) analyzed the willingness of household to pay for earthquake mitigation methods. They have formulated the related SEU model as (EQ 4):

$$S_i[U(r)] = PU[W - D(r) - r] + (1 - P)U(W - r) \qquad \text{[EQ 4]}$$

where $S_i[U(r)]$ is the subjective expected utility of risk adverse households associated with investing r units of mitigation, P is the subjective probability of occurrence of an earthquake perceived by the households, W is the level of wealth, $D(r)$ is the level of damage after a certain r measure being implemented (such that $D(0) = D$, and it is a concave function). In this formulation, $PU[W - D(r) - r]$ is the household utility under the occurrence of an earthquake and $(1 + P)U(W - r)$ is the utility perceived in non-earthquake conditions.

The Random Utility Theory (RUT) places the hypothesis that individuals will maximize their expected utility based on objective components along with some degree of randomness expressed by an error component (Snowball & Willis, 2006; Asgary & Halim, 2011). The inclusion of the error component translates the lack of objectivity in the households' assessment, which derives from each individual's perception. Naoi et al. (2009) formulates the subjective risk probability as (EQ 5):

$$PR_{it}^{s} = PR_{it}^{o} + v_{it} \qquad \text{[EQ 5]}$$

where PR_{it}^{s} is the subjective risk perception, PR_{it}^{o} is the objective probability component and v_{it} arises from household's measurement error. For $v_{it} = 0$ it is assumed that households would have a correct assessment of the probability of occurrence of an event.

Other authors have used different methods to assess risk perception. Zhang et al. (2010) used a 1-5 scale of flood, hurricane and general hazard risk and used these variables in a standard hedonic price model regression to study the impact of risk perception against housing features. Tekeli-Yes et al. (2011) used a similar scale method to measure risk awareness and perception among Istanbul residents. The author then applied the survey responses as variables in logistic hedonic regression.

Considering the difficulty of correctly assessing risk perception based on surveys is crucial. Although these represent the most direct methods of evaluation awareness and perception, it has been stated in the literature that various heuristics and biases influence measures of perceived risk via item wording (Wolff et al., 2019). According to Wolff et al. (2019), objective risk is defined as the probability of negative outcomes weighted by severity. Similarly, in the standard economic literature of EUT, subjective probability assessment is often referred to as the perceived probability of a particular hazard weighted by its severity. Although with recognizable flaws, utilitarian economic theories continue to be one of the most utilized ways to assess risk perception by households. Some authors have even argued that worry should be assessed instead of risk perception, as the former has been found to be a better predictor of behaviour than the latter (Wolff et al., 2019).

2.6.2. Institutions

Both governments and insurance companies manage risk on a larger scale. Instead of being dependent on a single investment location and building typology, as is often the case for investors and households alike, governments and insurance companies are exposed to the risk of several properties, whether in a single region or city or across multiple locations. Besides dealing with immediate impacts such as property damage or loss of life, they must also manage the ripple economic effects of such events. For example, major earthquakes will often translate into an abrupt decrease in Gross Domestic Product (GDP), failure of businesses, unemployment, and increased prices due to supply chain shock, putting local governments under tremendous financial stress (Ghesquiere & Mahul, 2010; Goda et al., 2014). In a broader scope of analysis, we can also argue that the length of present global supply chains that stretch across the globe also expose outside regions to supply shocks with far-reaching consequences (Kajitani et al., 2013). Therefore, it is of paramount importance for governments and risk managing institutions to have a correct assessment of natural hazard risks and to enhance their capability of dealing with low-probability high-consequence (LP-HC) events by the adoption of hard (e.g., seismic retrofitting) and soft (e.g. insurance) disaster risk reduction measures (DDR) (Cavallo & Noy, 2011; Michel-Kerjan et al., 2013; Goda et al., 2014; Pothon, 2020; Safiey & Pang, 2021).

Contrary to the case of households, institutions rely heavily on the catastrophes statistical and computational modelling as well as actuarial/financial and economic models in their decision-making process (Pothon, 2020). As an example, the typical risk assessment process from an insurance company will consist of four steps: (i) inventory/exposure database, (ii) hazard characterization, (iii) structural vulnerability assessment, and (iv) loss estimation and insurance portfolio analysis. During this process, they would use Monte Carlo simulations of various scenarios, damage index-based models for structural vulnerability assessment (Safiey & Pang, 2021), compound Poisson processes and Extreme Value Theory to model financial risks, and use expected utility to model the demand decisions by homeowners. There is a clear difference between the risk assessment process of households and institutions. Even their attitude towards risk is modelled differently: homeowners are typically assumed to be risk-averse decision-makers, having a concave utility functions whereas institutions are considered to be risk-neutral decision makers, having a linear utility function. (Goda et al., 2014). Figure 2-5 summarizes the risk modelling overview for institutional decision making.

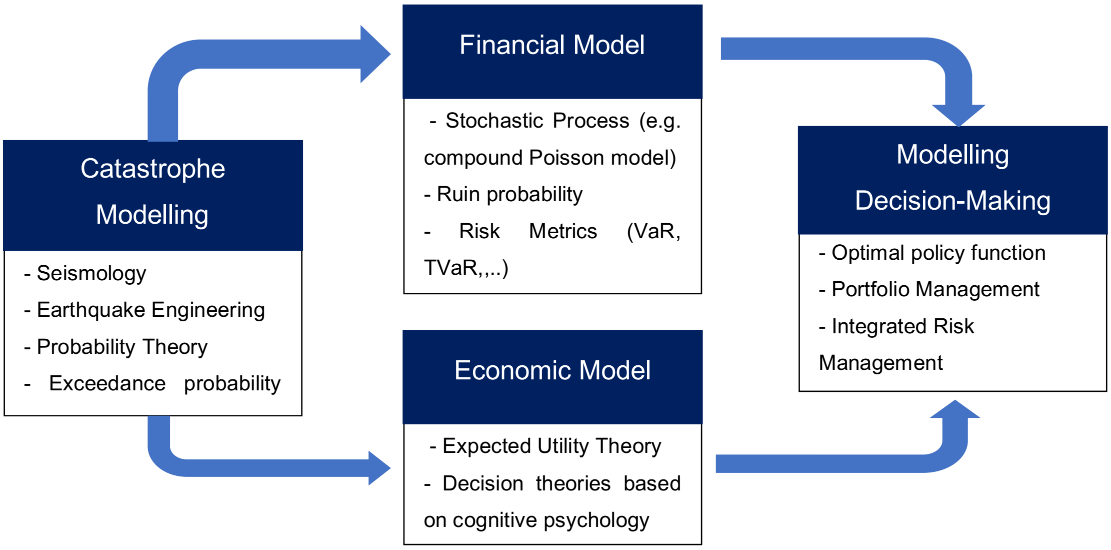


Figure 2-5 – Modelling decision-making. Adapted from Goda (2014)

2.6.2.1. Local Government

One of the main functions of governments is to ensure security and stability for their citizens. Security to natural hazards should therefore come as a priority issue to be addressed. The preparedness and response to natural disasters at an institutional level is shared between local and central governments (e.g., through urban planning and seismic regulations, respectively). In a natural disaster, local governments are responsible for providing first aid to the victims, but they should also play a crucial role in preventing the impacts of natural hazards (Col, 2007; Lee, 2019). There is much evidence in the literature that investing in DRR measures often pays off when nature strikes. Goda et al. (2014) provides several examples that such proactive earthquake risk management is the key to achieving sustainable and resilient communities against earthquake disasters. The implementation of financial incentive schemes to promote structural retrofitting and even urban planning tools, such as the delimitation of riskier zones, can be used to promote a safer housing stock. Accurate information about risk publicly promoted through local media should also play a part in an effective communication strategy to promote safer housing and household decision-making processes (*see previous chapters on Risk Perception and Risk Assessment by Households*). Both disaster promotion and education should be tailored to citizens' needs based on their vulnerabilities and knowledge (Lee, 2019). Local governments should also be liable for this crucial step in pre-disaster preparedness.

However, local governments play an essential role in both pre- and post-disaster management (preparedness/mitigation and response/recovery phases, respectively). They are also responsible for providing the first response in post-disaster situations, including the financial assistance much needed by its citizens. O'Leary (2004,1, as cited in Col, 2007) states, "*Virtually all disasters are experienced at the local level, where many communities can expect to be 'on* *their own' for the first seventy-two hours after disaster impact*". However, they will have to rely on their local governments to take immediate action which requires capital. Ghesquiere and Mahul (2010) listed 11 financial sources available to local governments during post-disaster, which are comprised of 3 basic categories: 1) donations, 2) public debt and 3) insurance. These resources are then used across the three main phases in post-disaster management: Relief (1-3 months), Recovery (3-9 months) and Reconstruction (>9 months). Each with its own resources' requirements (Figure 2-6).

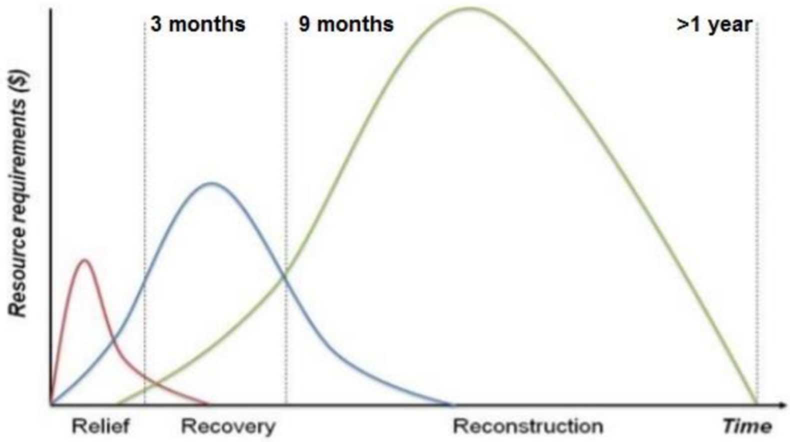


Figure 2-6 - Funding requirements by stage. Source: Ghesquiere and Mahul (2010)

Hence, the financial stability of local governments can broadly impact their handling of the disaster. Lee et al. (2017) have concluded that local governments' independence and financial stability have impacted the damage costs due to natural hazards in South Korea. According to the authors, damage costs can be lowered in a community where the local government can provide enough funding to implement disaster prevention projects. Lee (2019) also highlighted the importance of governments in strengthening risk governance and providing "disaster resilience" to their communities by involving multiple stakeholders. In addition, the author argued that an effective risk government requires a clear understanding of network participants' capabilities and resources.

In sum, local governments play a crucial role in preparing, mitigating, and managing recoveries after natural disasters. Therefore, thorough knowledge of the community and risk assessment based on both physical (catastrophe modelling) and financial/economic modelling is of utmost importance and should be translated into effective preparedness and mitigation measures (based on financial incentive schemes and education) and response/recovery measures (made possible by a strong network of partnerships with multiple stakeholders and strong financials).

2.6.2.2. Insurance

Regarding insurance risk management, earthquake events have an added layer of risk related to their geographical distribution. For instance, car insurance policies can be modelled as independently distributed events, as the occurrence of a car accident leading to an insurance claim is generally randomly distributed in time and space. In contrast, the occurrence of an earthquake disaster will require increased liquidity, given that its risk is spatially and temporally correlated. Hence, to minimize insolvency risk, insurance companies often take their own risk-sharing measures with reinsurers with greater risk-bearing capacities or through alternative risk transfer (ART) tools such as earthquake catastrophe (CAT) bonds. In addition, insurers often have to deal with other general issues in their coverage such as moral hazard[20] or adverse selection[21] (Goda et al., 2014).

It is noteworthy that governments can also act as reinsurers due to their ability to borrow money by issuing bonds. They can either act directly or via publicly-owned companies, as in the case of the Japan Earthquake Reinsurance Company (JERC), New Zealand's Earthquake Commission (EQC), the Turkish Catastrophe Insurance Pool (TCIP)[22], the Iceland Catastrophe Insurance (ICI) or the California Earthquake Authority (CEA).

2.7. Willingness-to-Pay

As mentioned in the previous chapters, risk perception is critical to foster the adoption of DRR measures. Therefore, it also significantly influences decision-making processes regarding adopting earthquake insurance and even home purchase decisions (Kunreuther, 1996). However, having an accurate risk perception or being well informed is not sufficient criteria to act. There are several factors involved in the decision other than the risk perception, such as the price of DRRs or the price difference between safe and unsafe housing, the attitude towards risk (e.g., fatalistic, optimistic attitude), age, gender, income level or cultural environment. We find examples in the literature of this dissonance, as in the case reported by (Pothon, 2020), where results have shown that most California households do not buy earthquake insurance because of the price and not because of risk underestimation. To have a full scope of the relation of households and investors with risk, one must analyze both sides of the spectrum through the analysis of their perception (through explicit models) and their actions, reflected by market dynamics (through implicit methods).

Explicit Methods, such as the Contingent Valuation Method (often referred to in the literature as CVM or CV), estimate the value a person places on a certain good or feature through surveys. It is often applied as a survey design to estimate the willingness-to-pay (WTP) or willingness-to-accept (WTA) for a good or service. For example, in the literature, it has been used to assess the willingness to pay for safer housing, risk insurance or other DRR measures (Palm, 1981; Brookshire et al., 1985; Willis & Asgary, 1997, 1997b; Dooley et al., 1992; Palm, 1998; Asgary et

[20] e.g. a policyholder further damages its property after a disaster to claim a higher insurance payment
[21] e.g. insurer cannot distinguish between high and low earthquake risk potential and charges a flat fee based solely on location
[22] See for instance Yucemen (2005)

al., 2007; Naoi et al., 2009; Tekeli-Yes et al., 2011; Asgary & Halim, 2011; Zhang et al., 2009; Sarmiento et al., 2015).

Implicit Methods, such as hedonic price models, estimate the value that the market puts on a certain good or feature through the construction of statistical models, often price models. In these methods, the feature or good to be assessed is used as an input variable to the model and its impact is measured in the dependent variable of interest. In a classic hedonic price model, the impact of a feature is given by its regression coefficient. As an example, to study the price effect of having a structural reinforcement in a building, it would be possible to use a database of sales prices of several properties with a binary variable as input indicating whether the building has been subject to reinforcement or not (Palm, 1981; Brookshire et al., 1985; Beron et al., 1997; Dooley et al., 1992; Bin & Polasky, 2004; Nakagawa et al., 2007; Toya & Skidmore, 2007; Naoi et al., 2009; Donovan et al., 2007; Deng et al., 2013; Modica and Locati, 2016; Fekrazad, 2019; Peng, 2021).

Hence, several studies have used both methods simultaneously. In this study, the same approach will be followed to comprehensively view the natural hazards in Lisbon's real estate market.

2.8. Real Estate Market

2.8.1. Portuguese Context

2.8.1.1. Macroeconomic context

Portugal is the southwesternmost country in Europe, with a population of 10 286 million[23] and a GDP of 239 billion euros[24]. From being highly dependent on the primary sector in the 1960s and 1970s, Portugal has managed to shift its economy onto the secondary and tertiary sectors, benefitting from the economic assistance of its partners in the EU. As an example, in 1974, the primary sector comprised 35% of the working force, the secondary sector 34% and the tertiary sector 31%; Nowadays, the primary sector accounts for 5% only, while the secondary and tertiary sectors account for 25% and 70%, respectively[25].

The country was one of the hardest hit by the 2007-2008 financial crisis (Serapioni & Hespanha, 2019), with several years of negative GDP growth, which culminated in the 2014 decrease of 4.0%. Moreover, the country had been in a decade-long state of economic stagnation since 2000, with recurring budget deficits, which may have contributed to its frail economic condition at the time of the crisis (Marinheiro, 2013). However, the political measures taken by the government, and the call for IMF assistance in 2011, led to an economic recovery praised by international media and partners (Newel, 2019) (see Figure 2-7).

[23] 2019 population estimates. Source: PORDATA
[24] Source: World Bank (2019). https://data.worldbank.org/country/PT
[25] Source: PORDATA

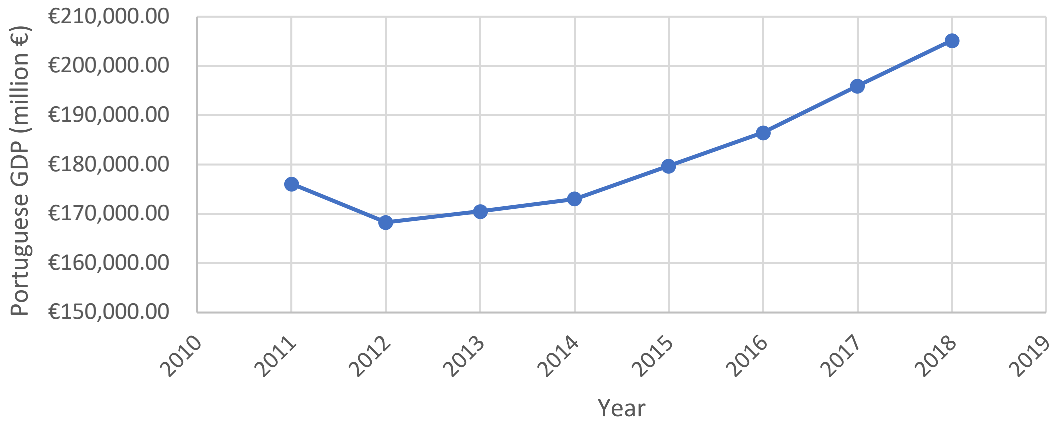


Figure 2-7 - Portuguese GDP (million € - base 2016). Source: PORDATA

In 2019, 76% of the Portuguese GDP came from the tertiary sector. Industries such as tourism, energy and automotive are some of the larger in the country. However, technology is slowly taking its space in the Portuguese economy with the birth of innovative technological startups. Much of this innovation is taking place in Lisbon, the nation's capital city, which has often been considered a thriving tech hub (Tissot, 2016; Hoffower, 2019)

Lisbon's Metropolitan Area is home to 2.8 million people[26] and 24.26% of the country's working population[27], and it is the core location of the Portuguese economy with a GDP per capita[28] of 24 749 €, the highest amongst NUT-II regions. The city of Lisbon has a population of 509 thousand[29], mainly working in the tertiary sector. Despite the decentralization effort made in recent years, Lisbon is still the heart of the Portuguese political and business spheres[30].

2.8.1.2. *Homeownership*

Portugal has traditionally followed the trend of other Southern and East European countries, where people prefer to buy their own homes rather than rent, which is often considered as part of their cultural heritage (Elsinga & Hoekstra, 2005; Antunes & Seixas, 2020). However, this was not always the case. According to the 1970 census, the distribution between renters and owners was roughly the same (49.3% owned their home) (Xerez et al., 2019). The liberalization of the credit market in the 80s and 90s and the economic growth following the Treaty of Accession of Spain and Portugal (1985) provided the conditions for the surge of new construction and access to credit markets needed to increase the homeownership share

[26] 2.863.272 (2019). Source: https://www.pordata.pt/municipios/populacao+residente+total+e+por+grandes+grupos+etarios-390
[27] Source: https://www.pordata.pt/municipios/populacao+ativa+total+e+por+nivel+de+escolaridade+completo-808 (2020)
[28] Source: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0008839&contexto=bd&selTab=tab2&xlang=pt (2017)
[29] 509.515. Source: https://www.pordata.pt/municipios/populacao+residente+total+e+por+grandes+grupos+etarios-390
[30] Source: OECD . https://www.oecd-ilibrary.org/sites/fea62108-en/index.html?itemId=/content/publication/fea62108-en

The 2007-2008 financial crisis brought a new paradigm to the real estate market. Although most people would still prefer to buy their property, for several families, the financial stability needed to be granted a loan was put into jeopardy in the aftermath of the crisis. According to Braga (2013), the constraints imposed by the crisis led to increased demand for renting. The costs related to the purchase of a property and the values required for a down payment would represent an obstacle to the buying option. This has been observed across several European countries (Lennartz et al., 2015).

In 2019, following the economic recovery, the percentage of homeownership stood at 73.9%, lower than the 95.8% observed in Romania, but higher than the European average of 69.3%. It is clear the difference in preference when we compare it with northern economies like Germany (51.1%), Denmark (60,8%) or Sweden (63.6%).

Despite the effects of the COVID-19 pandemic, it is reasonable to assume that the legacy of preference for homeownership will continue in the foreseeable future. In a recent study conducted by real estate brokerage company Century 21 (2020) after the initial shock of the pandemic, 45% of responders considered moving to a new house, of which 89% claimed to prefer buying rather than renting. As for the main reasons for their choice, 59% considered homebuying as an investment in the future and 21% as part of their legacy to their children. According to the same source, 75% of responders were willing to pay between 200€ and 600€ a month for rent/mortgage.

The preference for homeownership has been rooted in the Portuguese tradition and culture, being considered a landmark of adulthood success. Hence, it will likely remain a driving force in demand for property, especially in the purchase and sale market. According to a survey on the financial situation of families, in 2020, only 2,0% of homeowners would prefer to be renters and 63.5% of renters would prefer to be homeowners (INE 2020). According to the same survey, 77.3% of Portuguese households' wealth is held in real estate, whether as a primary residence or as other real estate property. It is also worth mentioning that the recent increase in valuations has not been followed by an increase in salaries which, given the importance of homeownership to the Portuguese society, may lead to social discontent due to lack of housing accessibility with deeper political and social consequences.

2.8.1.3. Credit / Mortgage

The desire to acquire property has long been part of Portuguese culture. However, before the European Communities' access, few families had the financial stability to fulfil the strict conditions needed to access the credit market. Real estate developers were also unfavored by the high credit rates that came with the political instability following the carnation revolution of 1974. A higher cost of capital for developers would translate into higher selling prices, reducing the number of potential buyers.

In 1986, Spain and Portugal were officially integrated into the European Communities, now the European Union. In the years that followed, the country experienced the liberalization of the financial sector, and credit became less expensive[31] and available to a broader market (Lestegás, 2019; Antunes & Seixas, 2020). The increase in accessibility to credit, expanded by the subsidized credit policies, the desire for homeownership, and the improvement of macroeconomic conditions led several Portuguese families to acquire property during the 1990s. Consequently, the debt levels in Portuguese families grew accordingly, a trend that would remain during the beginning of the 21st century (see Figure 2-10), especially in younger families. For example, in 1981, only 7.7% of Portuguese families had a mortgage debt, however, by 2001, that number had increased to 23.6%, reaching 31.2% by 2011. It is worth mentioning that the LMA has been significantly affected by this increase in debt (Xerez et al., 2019). The easiness of access to credit, and the influx of migrants (domestic and foreign[32]) during this period also led to a credit-fueled suburbanization of the LMA (Malheiros & Vala, 2004; Lestegás, 2019).

It is worthy of note that despite the increasing average value of purchase and sale contracts of real estate properties, the actual number of transactions peaked in 1999 at 284 241 contracts (see Figure 2-9 and Figure 2-11).

The increased demand and the lower credit rates led to substantial new construction growth. In 1985, 35 475 dwellings were completed as new buildings for family housing, rising to 125 539 in 2002 (see Figure 2-8). The total number of dwellings grew by over a million (or 24%) in eleven years, from 4.216.540 in 1991 to 5.232.258 in 2002.

[31] Partly because of the decrease in inflation rate. As an example, in 1985 the inflation rate stood at 19.5%, while in 2001 it closed at 4.4%. Source: PORDATA

[32] In 1991, foreign residents in LMA represented 1.8% increasing their share to 4.7% in 2001 (Malheiros & Vala 2004)

Interest rates have decreased since the 2007-2008 financial crisis, reaching negative values in 2015[33] (see Figure 2-12). These would also exert pressure on the market, as it is well known that interest rates have a close effect on housing credit, with lower rates leading to increases in demand (Reichert, 1990; Cho & Ma, 2006; Tse et al., 2014)

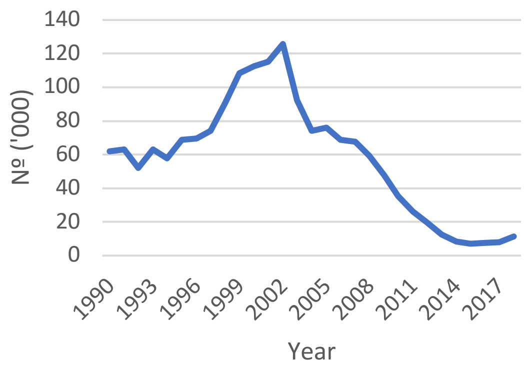


Figure 2-8 - Number of dwellings completed in new buildings (in thousands). Source: INE

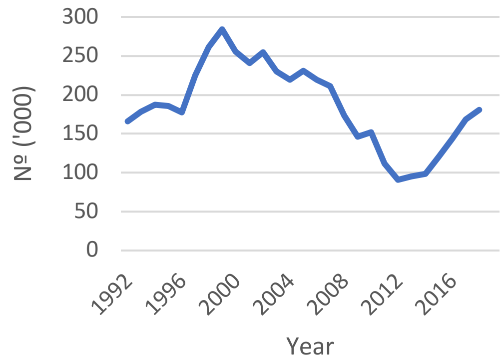


Figure 2-9 - Purchase and sale contracts (in thousands). Source: INE

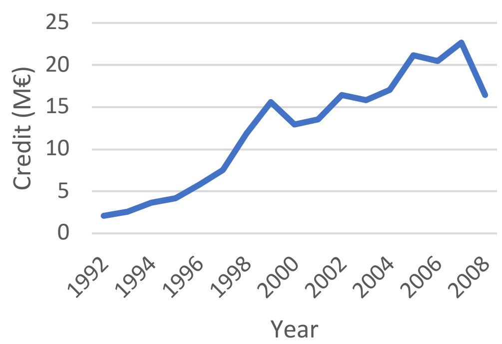


Figure 2-10 - Credit to natural persons (in millions of €). Source: INE

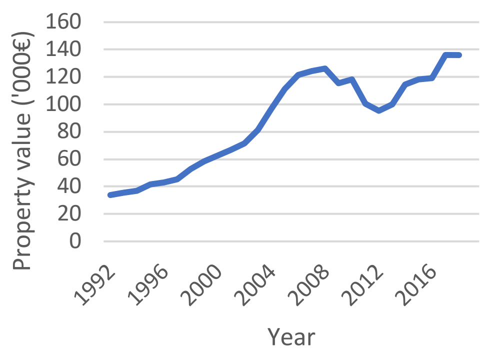


Figure 2-11 – Mean value of property sales (thousands €). Source: INE

Amidst the post-COVID recovery phase, the general sentiment was that credit conditions would remain stable in the near future. In July 2021, the European Central Bank (ECB 2021) European banks reported that the demand for housing credit from households is strong and will continue to increase in the coming month. Hence, one might expect prices will continue to rise, leveraging on the low credit costs (Scope Ratings GmbH 2021) (Figure 2-12). However, due to generalized rise

[33] At the time of writing (20 September 2021)

in prices, since March 2022, the U.S. Federal Reserve (FED), the European Central Bank (ECB) have been drastically increasing rates in an attempt to tame inflation.

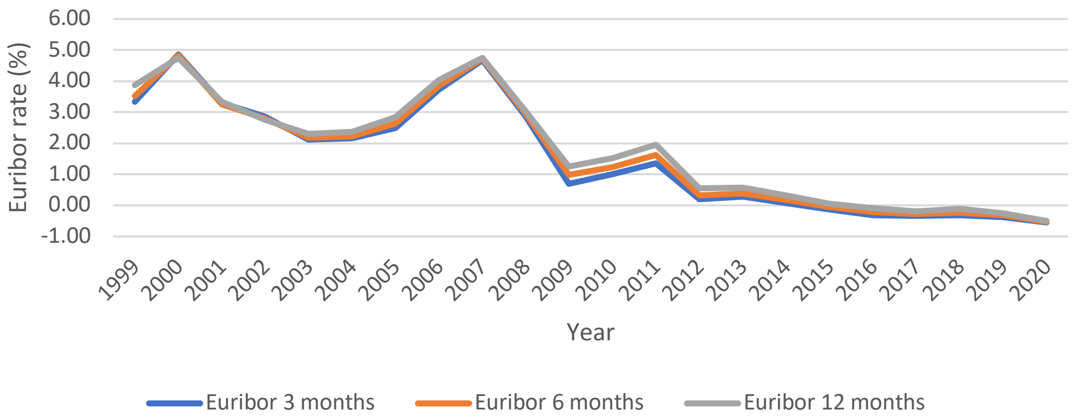


Figure 2-12 – Euribor rates. Source: PORDATA

2.8.1.4. Construction Costs and Taxation

Construction costs have risen, driven mainly by labour costs (see Figure 2-13). The increase in labour costs can be mainly attributed to the rise of minimum wage values and the shortage of qualified workers in construction. Nonetheless, we should note that several immigrants, often in illegal conditions, work in the sector, leading to a notable share of unreported and un-taxable work (Morás, 2018; Lusa, 2021)

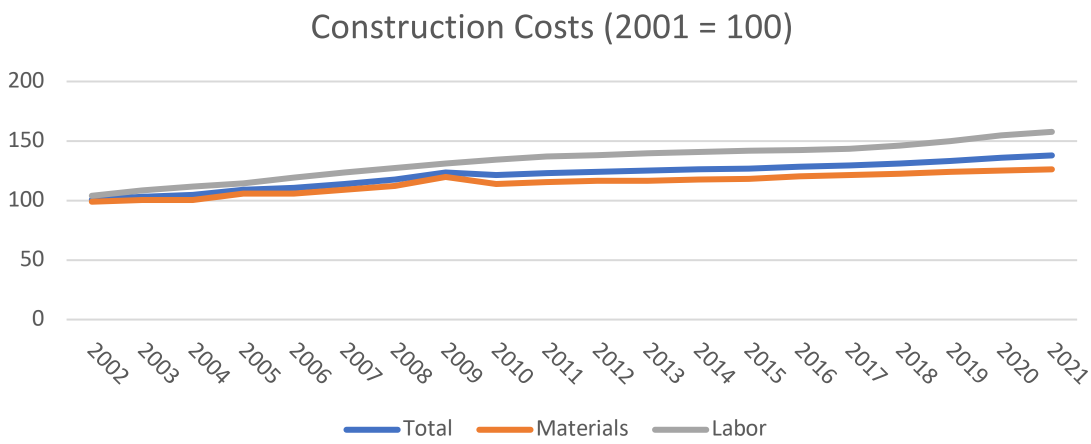


Figure 2-13 – Construction costs index (base 2001 = 100). Source: INE

COVID-19 has also increased construction costs, impacting workers' availability but mainly the cost of materials. For example, the disruption of logistical chains and the increase in oil prices were reflected in the prices of materials, such as steel, in the post-pandemic period. Some companies have reported increases in steel as high as 30% over nine months[34].

New real estate developments, especially in Lisbon and Porto, have primarily focused on luxury properties[35] during the last few years, driven by the demand from foreign and national wealthy investors. Some developers have justified their choice by arguing that the increased construction costs and the high tax rates[36] are some of the main factors contributing to the higher price tags and the preference for the luxury segment's development. However, if we analyze it in the European context, although above an average of approx. 20%, the Value Added Tax (VAT) of 23% applied to new construction is not out of place (see Table 2-2).

Table 2-2 - Real estate property VAT across several European countries. Source: UIPI (2013)

Country	VAT (%)
Hungary	27
Denmark	25
Normay	25
Sweden	25
Finland	24
Greece	23
Portugal	**23**
Belgium	21
Spain	21
Austria	20
Bulgaria	20
Slovakia	20
Slovenia	20
U.K.	20
France	19.6
Germany	19
Romania	19
Cyprus	18
Czech Rep-	18
Ireland	13.5
Italy	10
Switzerland	8

However, it is worth mentioning that the Portuguese VAT has been increasing from 1995 onwards, representing an ever-larger share in price (Table 2-3).

[34] See for reference https://eco.sapo.pt/2021/09/08/escassez-de-materiais-dita-subida-de-66-dos-custos-de-construcao/ , https://www.idealista.pt/news/imobiliario/construcao/2021/06/22/47831-construcao-falta-de-materiais-ameaca-paralisar-o-setor and https://www.dinheirovivo.pt/economia/escassez-de-materiais-de-construcao-leva-precos-a-disparar-acima-de-35--13785630.html

[35] See for instance https://visao.sapo.pt/imobiliario/2021-03-25-a-nova-cara-da-velha-lisboa-os-projetos-que-vao-mudar-a-face-da-capital/ and https://www.idealista.pt/news/imobiliario/habitacao/2021/09/15/48887-casas-de-luxo-em-portugal-quais-os-estrangeiros-mais-interessados

[36] See for instance https://imobiliario.publico.pt/opiniao/30-10/ and https://jornaleconomico.sapo.pt/noticias/aumento-dos-custos-de-construcao-que-impactos-esperar-no-imobiliario-777945

Table 2-3 - VAT rates from 1985 to present day in Portugal. Adapted from Morás (2018)

Year	VAT (normal rate) (%)
1985	30
1986-1991	30
1992-1994	30
1995	17
1996-2001	17
2002-2005	19
2006-2008	21
2009-2010	20
2011- Present	23

Note that the VAT is not the only taxation included in the property price. The value of the corporate tax paid by construction companies (average of 21%[37]), the municipal surtax, and licensing fees are also to be accounted. Additionally, the buyer may be subject to pay a sale tax[38] (up to 6%) and stamp tax[39] (0,8%). During his ownership, the buyer will also have to pay an annual property tax[40] (Morás, 2018).

2.8.1.5. Employment, Demographics and Youth's Housing Struggles

Although the country has experienced a rise in the unemployment rate in the years following the financial crisis of 2007-2008, it has also been steadily decreasing since 2013, contributing to the demand increase as it has been shown in the literature (Agnew & Lyons, 2017). Despite the challenge presented by the COVID-19 pandemic, which increased unemployment momentarily, especially in the tourism sector due to travel and sanitary restrictions, the country has managed to soften its impact, as shown by the rapid decrease shown in Figure 2-14.

Several known factors have been proven to influence the housing demand: births, nest-leaving adults, union formation/marriages, migrations and deaths. Nest-leaving young adults have been considered one of the most critical drivers of the increase in housing demand (Mulder, 2006). However, young adults face increasing challenges in reaching their financial independence, especially concerning housing (Lennartz et al., 2015). In a 2007 study, young Europeans ranging from 15 to 30 years old were asked *"what do you think is the MAIN reason that young adults live in their parents' homes longer than they used to?"*. 62% of responders claimed they couldn't afford to move out (Gallup Organization, 2007). According to an OECD (2018) study, the escalating rents and house prices are increasingly overburdening young people. A survey conducted in the same report showed that around one-third of respondents aged between 20 and 34 picked

[37] Small and Medium Enterprises (SMEs) may benefit from a reduced rate of 17% in the first 15.000€ of revenue. Source: PWC (2021)
[38] As per decree-law n.° 287/2003. Applicable to sales over 92.407€.
[39] As per Law n.°150/99
[40] For more information, please refer to https://info.portaldasfinancas.gov.pt/pt/informacao_fiscal/codigos_tributarios/cimi/Pages/codigo-do-imi-indice.aspx

securing or maintaining adequate housing as one of their top-three short-term concerns, with the share peaking at 40% among 25 to 29 years old (Figure 2-15).

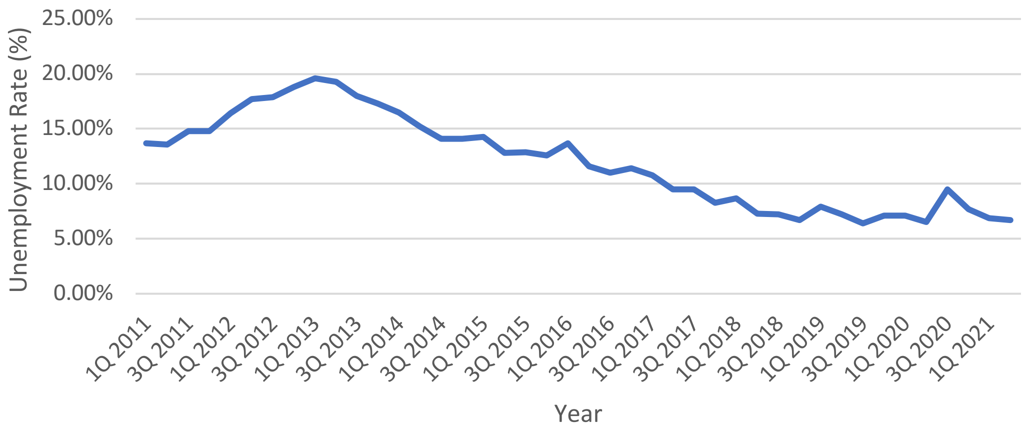


Figure 2-14 – Unemployment Rate in Lisbon's Metropolitan Area. Source: INE

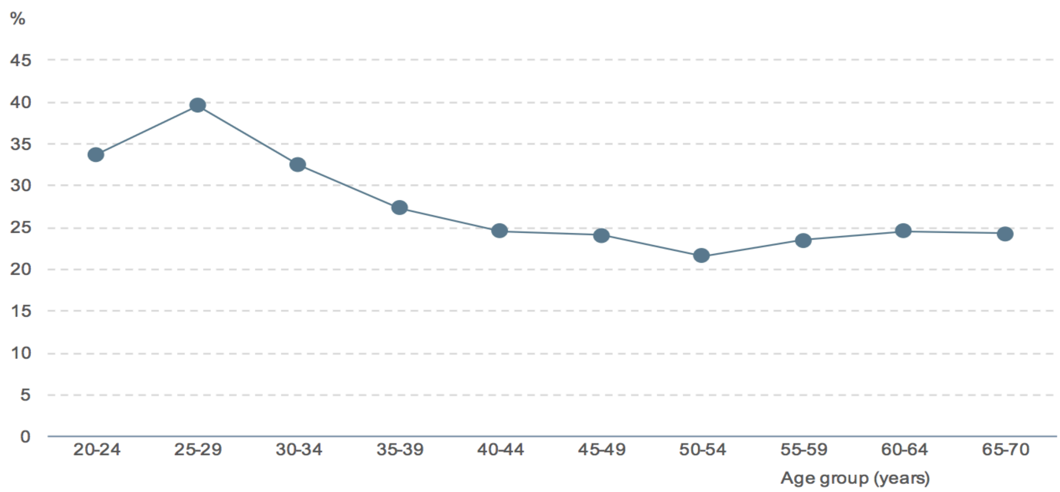


Figure 2-15 – Percentage of by age group. Source: OECD (2018)

Portuguese young adults have also been following this trend. According to Xerez et al. (2019), the rise in sales prices and rents may have been a significant factor in the increased levels of financial stress among renters. The authors found increases in housing deprivation[41] n all age groups between 2011 and 2017. However, renters and owners under 29 years old have shown the highest increases (19.4% to 42.2% and 15.2% to 27%, respectively). This result may be interpreted as the outcome of the inability of young people to keep up with the increase in housing prices, given their job precarity (Tavares et al., 2021)

According to Lennartz et al. (2015), the economic downsizing experienced during the financial crisis, lower real incomes and consequently job precarity, especially of younger adults between 18-34 years old, and higher mortgage deposit requirements, led to sharp drops in homeownership

[41] Housing deprivation is a measure of poor amenities and is calculated by referring to those households with a leaking roof, no bath/shower and no indoor toilet, or a dwelling considered too dark. (source: Eurostat)

in the post-crisis period. Moreover, in southern European countries (Italy, Portugal, Greece and Spain), a very high share of younger adults ended up living with their parents, delaying home-leaving to relative old age. Nevertheless, if younger people left their parental home before age 34, they would be more likely to become homeowners than renters.

Despite the job precarity, the number of one-person households has also been contributing to an increase in housing demand, following the trend of developed countries[[42]. In 2018, one-person households represented 23.8% of the total (see Figure 2-16). According to PORDATA[43], there were 852.800 one-person households in the country, a 20% increase from the levels observed in 2010, of which 42.9% consist of non-senior citizens (younger than 65 years old) a rise from the 36,4% of one-person households in 2010, meaning there have been more young people choosing to live by themselves. One factor contributing to the rise of one-person households is that people are getting married later in life, and divorce rates have also increased from the beginning of the century, doubling from 30% in 2000 to approximately 61% in 2020. In 2000, the average age of a Portuguese man in his first marriage was 27.5 years of age, while in 2020, it was 34,9. It is interesting to notice that, according to the literature, there is evidence that people tend to postpone marriages and union formations when they do not find adequate housing (Clark 2012, as cited in (Garha & Azevedo, 2021)). These changes in demographics also exert pressure in the demand for housing.

Another factor influencing the Portuguese real estate market is the ageing of its population. In 1960, Portugal had an ageing index of 27. In 2019, that number had risen to 161.3, standing 19,74% above the EU27 average[44]. In a 2013 study, the Portuguese Association for Consumer Protection (DECO, 2013) found that 67% of retirees do not have financial conditions to pay for retirement homes and that the families support the majority (53%) of costs. There is also evidence in the literature that newer generations have increasingly desired greater independence (Tinker, 2002). Hence, the number of senior citizens leaving for retirement homes and the number of dwellings available will tend to decline, contributing to the lack of offers and price appreciation of real estate property.

[42] For more information on the rising of single-person households across Europe please see https://ec.europa.eu/eurostat/statistics-explained/index.php?title=People_in_the_EU_%E2%80%93_statistics_on_household_and_family_structures&oldid=375234

[43] https://www.pordata.pt/en/Portugal/Single+person+private+households+total+and+with+a+person+aged+65+and+over-822

[44] Source: https://www.pordata.pt/Europa/%c3%8dndice+de+envelhecimento-1609

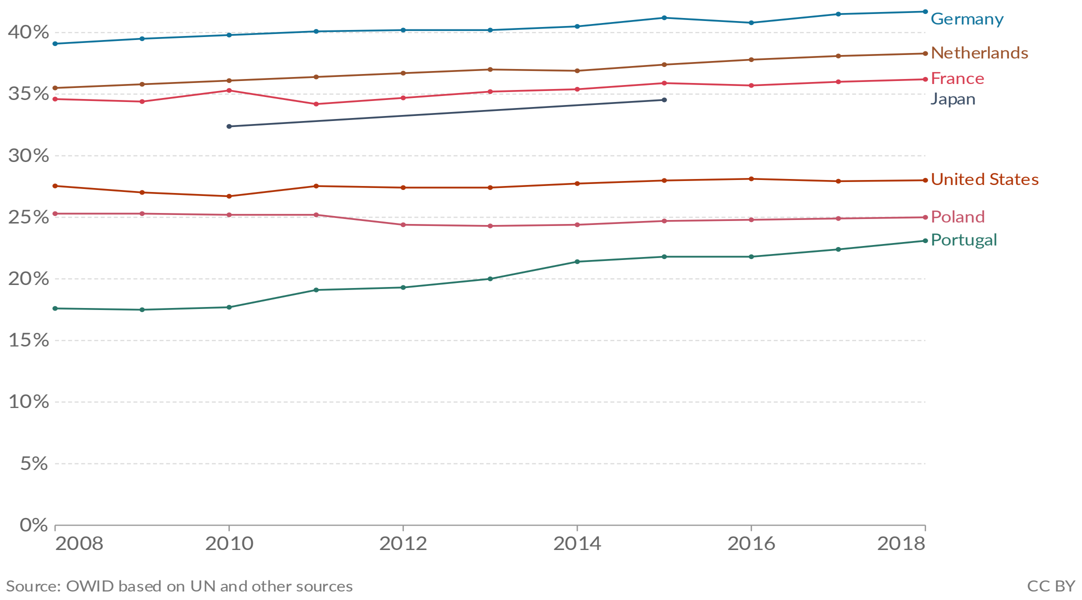


Figure 2-16 – Percentage of Single-Person households, 2008 to 2018. Source: OurWorldInData

2.8.1.6. Tourism

Following the 2007-2008 financial crisis, the austerity policies imposed in Portugal, and similarly in other countries in Southern Europe (Spain, Italy and Greece), reinforced the importance of tourism to the economic recovery (Mendes, 2017, 2021; Lestegás et al., 2019; Amore et al., 2020). This led to several efforts from the central and local governments to promote tourism in the country. As a result, throughout the last decade, the country has been recognized as one of Europe's leading destinations[45] and has also been awarded in the world tourism stage[46] in several categories[47]. Its capital city, Lisbon, is one of the most thriving cities, being recognized as World's Leading City Break Destination 2020[48] and as one the most promising tech hubs worldwide[49].

In 2011, hotel establishments in Lisbon registered 6.419.256 bed nights, increasing to 13.985.262 in 2019, a 117% increase[50]. According to Turismo de Portugal, 16,3 million guests[51] have visited the country, generating €8,7 billion in revenue. In recent years, we have also seen the rise of a new tourism industry, short-term rentals.

The increased numbers in tourism have also resulted in higher purchase and sale contracts of real estate acquired by non-residents (Figure 2-17). In 2021, 6.902 properties were sold to non-

[45] World Travel Awards. Source: https://www.worldtravelawards.com/award-europes-leading-destination-2020
[46] https://econews.pt/2019/11/29/portugal-is-the-best-destination-in-the-world-for-the-third-consecutive-year/
[47] For more information, please refer to https://www.worldtravelawards.com/winners/2020
[48] https://www.worldtravelawards.com/award-worlds-leading-city-break-destination-2020
[49] Source: Business Insider
[50] Source: INE
https://ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0008576&contexto=bd&selTab=tab2
[51] Foreign guests.
http://business.turismodeportugal.pt/pt/Conhecer/Apresentacao/Desempenho_Turistico/Paginas/default.aspx (last visited 8 September 2021)

residents, representing 5,54% of the total number of real estate transactions. In 2018, that number increased to 19.912 or 8,89% of all transactions.

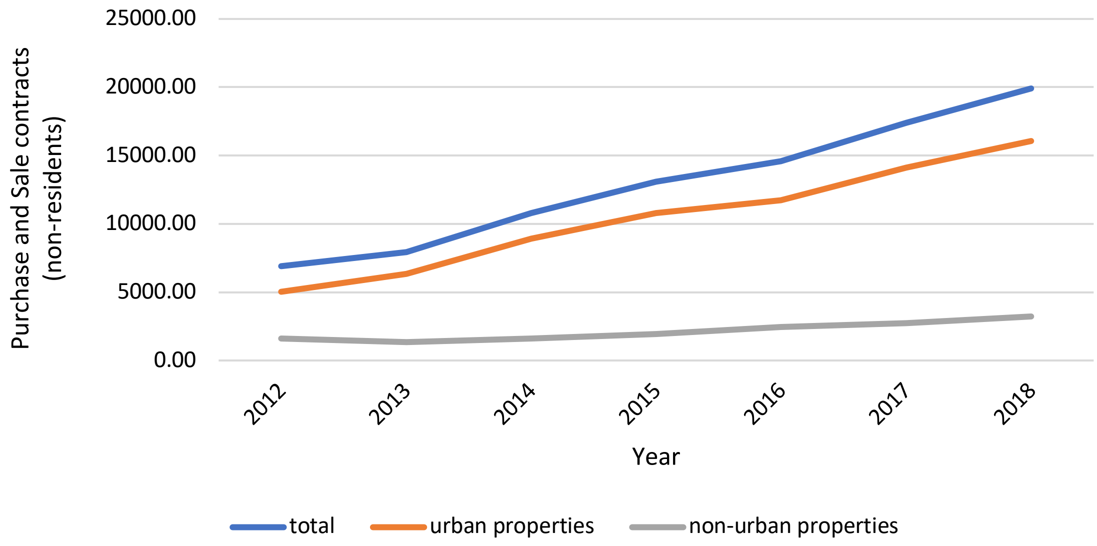


Figure 2-17 -Purchase and sale contracts by non-residents according to nature. Source: INE

The value of properties sold to non-residents have increased from an average value of 115 688.79 € in 2012 to 171 177.53 € in 2018, a 47.96% increase.

It is also noteworthy that in 2009, the Portuguese Government created the tax scheme for Non-Habitual Residents in terms of Personal Income Tax (IRS)[52], which concedes the following benefits for ten years:

- Exemption from taxation in Portugal of foreign source income resulting from dependent work;
- Taxation at the rate of 20% of Portuguese-source income arising from value-added activities[53];

Considering the current distribution of income tax categories[54] and that the average top Personal Income Tax Rate in Europe stands at 37.80%[55], the benefits are clear. Therefore, some consider Portugal one of the friendliest tax regimes in Europe to foreigners[56].

---

[52] As per Decree-law nº 269/2009. More information at https://dre.pt/web/guest/pesquisa/-/search/490420/details/normal?l=1
[53] Architects and engineers, visual artists, actors and musicians, auditors and tax consultants, doctors and dentists, university professors, investors, administrators and managers, senior management, as specified in Ordinance 12/2010, of 7 January.
[54] Annex A
[55] Source: https://tradingeconomics.com/european-union/personal-income-tax-rate
[56] For more information please refer to https://www.acm.gov.pt/documents/10181/753067/IRS_RNH_EN.pdf/3b639655-6f18-4f94-86c2-d3b080bd691d and https://www2.deloitte.com/pt/pt/pages/tax/articles/residentes-nao-habituais.html

2.8.1.7. Valuation

Following the 2007-2008 crisis, the Portuguese purchase and sale market experienced a decrease in valuations until 2013[57]. However, low credit rates, governmental policies promoting tourism, and the surge of short-term rentals, amongst other factors, pushed a market recovery from 2013 onwards, leading the country to one of the highest increases in the valuation of the housing stock across Europe (see Figure 2-18).

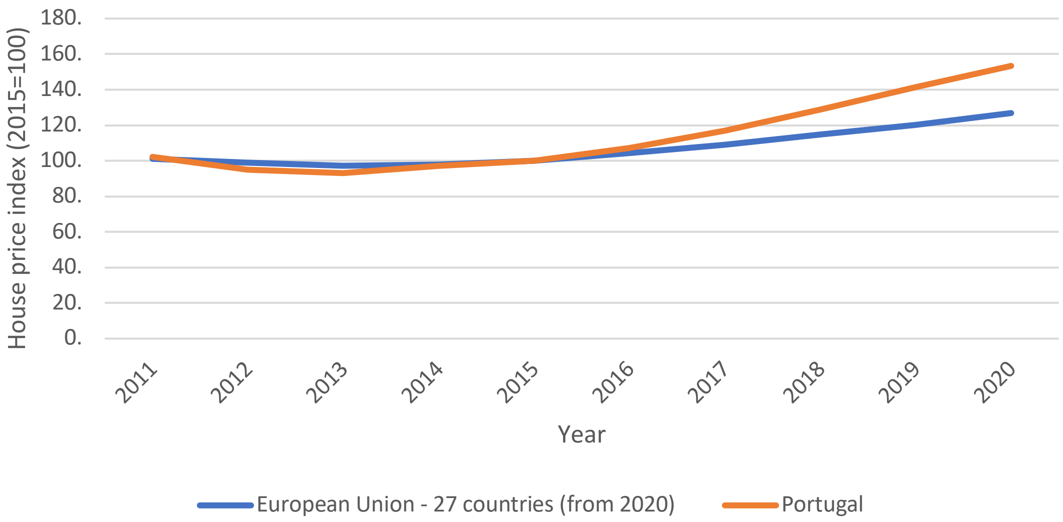


Figure 2-18 - House price index (Portugal vs European average). Source: Eurostat

As mentioned above, the increase in valuations is not a direct consequence of the increase in household income, which has been far more modest[58]. However, others argue that the issue is with the low average income of Portuguese families[59] and that the valuation of properties in the capital city is just following its European counterparts. This is creating a mismatch between the needs of the population for housing and the offer put on the market, contributing to the phenomena of "social filtering"[60], "gentrification", and "commoditization" of housing stock (Mendes, 2013; Lestegás et al., 2019).

The increasing valuations of the Portuguese real estate market, the exposure of households and banks to earthquake risk, and the inexistence of a national coverage fund for seismic activity constitute a delicate situation that should be addressed sooner rather than later.

The following chapters will describe some of the main factors affecting the increase in valuations across the Portuguese territory.

---

[57] Nonetheless, is worth mentioning that this general trend had not been followed by all property segments, namely the luxury market in Lisbon which began to soar during this period.

[58] Figure 2-18 shows an 61,67% increase in the Portuguese house price index between 2011 and 2020. During the same period, households's available income only rose by 13% (source: INE).

[59] See for instance https://www.dinheirovivo.pt/empresas/ceo-da-vanguard-preco-do-imobiliario-nao-e-alto-os-portugueses-e-que-ganham-mal-13803947.html

[60] Process in which less-privileged strata of society find themselves gradually being replaced by people coming from the upper-middle and upper classes who are able to pay for this kind of restored housing. (Mendes 2013)

2.8.2. Lisbon's market (Case Study)

In addition to the macroeconomic conditions, a series of complementary drivers influenced the real estate market valuation, namely: 1) the increasing number of tourists, underpinning the surge of short-term rental (STR) platforms such as AirBnB; 2) the creation of the Golden Visa Program, promoting the development of projects meeting its criteria; and 3) the deregulation of the rental market, in particular the end of the rent freeze by Law 31/2012[61], revitalizing the interest of investors.

These drivers contributed not only to the increase of valuations, but also to an internationalization of the real estate market in Portugal, in general, and in Lisbon, in particular. This reflects an enhanced attractiveness of the market at international level, with a growing percentage of the properties acquired by non-residents (Figure 2-19). These topics will be further detailed in section 3.3.3 *Golden Visa Programme*.

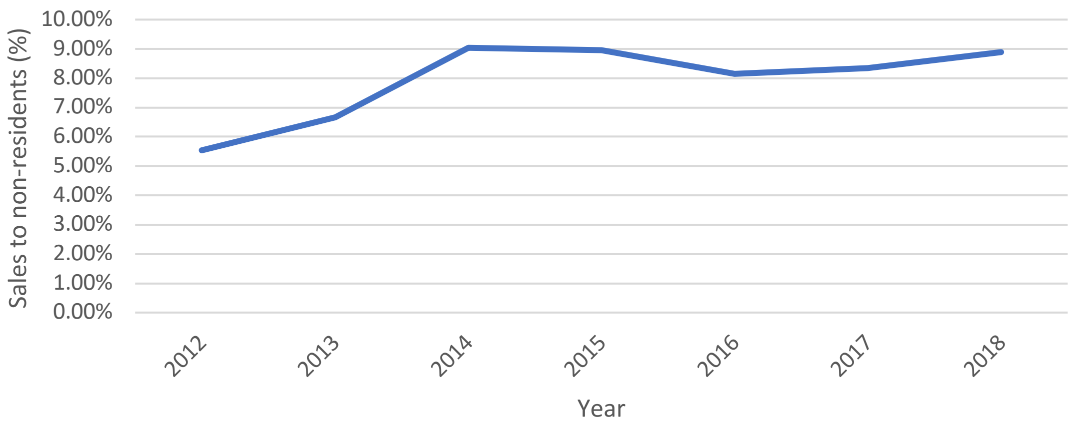


Figure 2-19 - Purchase and sale contracts by non-residents (%). Calculated by the author. Source: INE

It is noteworthy that Lisbon has a lower percentage of homeowners than the national average, at approximately 50% (CML 2023).

2.8.2.1. Lisbon's Valuations

Lisbon's market has seen a rapid increase in valuations, with median banking valuations doubling from 2014 to 2020 (Figure 2-20).

The increase in valuations and the shortage of supply caused by changes in the demand profile and increased construction costs have caused profound changes in the market[62].

[61] See http://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=1779&tabela=leis

[62] Supply shortage is not unique to the Portuguese market. Several countries are currently under a supply shortage in the property market. See for instance https://www.redfin.com/news/were-building-6-homes-for-every-10-new-households-where-will-people-live/

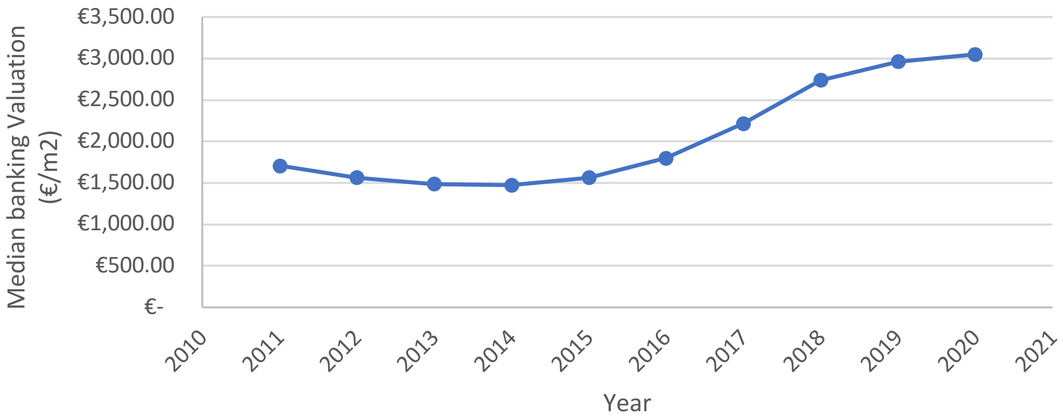


Figure 2-20 - Median banking valuation in Lisbon's county. Source: INE

The increase in prices affects both new construction and the existing housing stock, expanding from Lisbon's historical centre to adjacent locations (Figure 2-21 and Figure 2-22). Given that most of the housing stock in these parishes is comprised of older buildings, this price expansion pattern suggests that investors and homebuyers are oblivious to the differences in the typologies across the city, focusing solely on the financial gains.

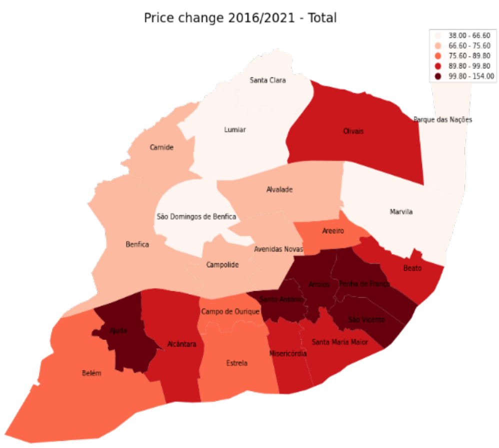


Figure 2-21 - Median sales prices change by parish (%) (Jan 2016 - Jan 2021). Own elaboration with data from INE.

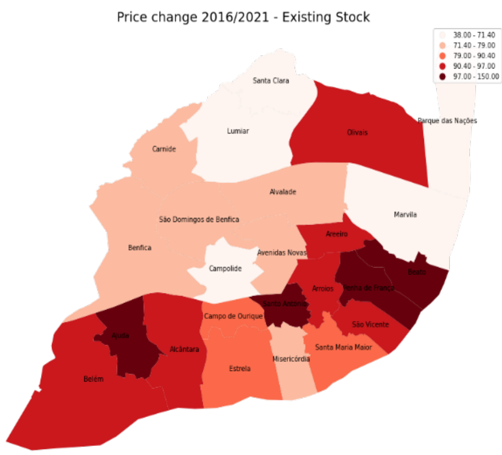


Figure 2-22 - Median sales prices change of existing building stock by parish (%) (Jan 2016 - Jan 2021). Own elaboration with data from INE.

These price increases have also raised some deeper issues regarding affordability and access to housing, even for the Portuguese middle class. This is not exclusive to Lisbon, as this phenomenon has been observed across multiple European cities under the combined influence of digital changes and weak regulatory policies. Terms like 'housing financialization', 'touristification', or 'gentrification' have been increasingly important in the literature (Mendes, 2013, 2017, 2018, 2021; Seixas & Antunes, 2019; Antunes & Seixas 2020). The Portuguese capital has seen a rapid increase in prices in the recovery from the 2008 financial crisis; however, the average income of the families did not follow the same pace (Lestegás et al., 2019).

Some authors suggested that prices should be moderated by introducing new regulations on rehabilitation by the city council, limiting the number of licenses in the city, setting a maximum rent limit[63] to discourage price speculation, or imposing a high property tax on vacant apartments (Mendes, 2013; Garha & Azevedo, 2021).

It is interesting to notice that in 1981, Lisbon was home to almost 9% of the Portuguese population, with 807 thousand people and an average age of 37.7 years; however, in 2019, it was estimated to be home to 509.4 thousand people, with an average age of 45.9 years old. Despite a 36% decrease in population, the city could not house their nest-leaving young adults by providing affordable quality housing (Garha & Azevedo, 2021). Over the same period, we saw the expansion of neighbouring counties, which were able to house these young adults by offering a higher number of dwellings with higher quality and at lower prices (Lestegás et al., 2019).

According to Portuguese Real Estate Professionals and Brokerage Companies (APEMIP)[64], in 2018, there was a shortage of supply of 70.000 properties nationwide, especially in Lisbon and Porto, where is located over 50% of the market demand. On the other hand, some authors have claimed that the number of dwellings in Lisbon currently exceeds households, so there is no mathematical need to build more residential buildings. Additionally, some authors claim that housing affordability should be done by the rehabilitation[65] of poor-quality older buildings rather than by new construction (Garha & Azevedo, 2021). Regardless of the perspective, the increase in valuations is noticeable in the Portuguese capital city.

It is noteworthy that the lack of quality offers in the market and affordability has led households to be less careful when buying a new home, increasing their exposure to financial and physical risks.

2.8.2.2. Deregulation of Rental Market

Since the implantation of the Portuguese Republic in 1910, with the publishing of the decree of November 12th 1910[66], we have seen periods of rent freeze across the country, justified by social and political instability and later by the effects of the first World War[67]. In 1948, under the dictatorship of Salazar, the Portuguese Government approved law 2030 (22nd June 1948), which allowed the increase of rents by a fiscal appraisal except for the cities of Lisbon and Porto. In fact, it was not the rent itself that was frozen but the fiscal appraisal, leaving landlords unable to increase rents (Braga, 2013). Many tenants have benefitted from this policy, and while rents were frozen, inflation was not. In the second half of the 20th century, Portugal experienced several periods of high inflation rates. For example, from 1978 to 1985, annual inflation rates were

[63] similar to the "Rent Price Cap law" in Berlin
[64] Associação dos Profissionais e Empresas de Mediação Imobiliária de Portugal (APEMIP). Source: Idealista last visited 20th September 2020
[65] In restricted terms, the urban rehabilitation process may be defined as an action leading to the significant improvement of the state of housing or a building, where the restoration of the built environment is less extreme than renovation as it consists of rearranging what is already standing and not replacing it with new buildings (Mendes, 2013)
[66] For more information see http://pl.proprietarios.pt/leis/arquivo/historico/G-Decreto-14Nov1910.pdf
[67] For more information see http://pl.proprietarios.pt/leis/arquivo/historico/G-Decreto5411-19Abr1919.pdf

consistently close to 20%[68]. This led to a detachment of landlords and investors from the rental market, in the two largest cities in the country. Consequently, many buildings in the city centre were left without the proper maintenance, resulting in the degradation of a significant portion of the housing stock[69]. According to the 2001 Census (as cited in CML 2011, p. 61), 72% of all rentals[70] in the city had a monthly rent below 100€. That is 34% of all dwellings in Lisbon. It is noteworthy that the limitations imposed by the 1994 Lisbon's Master Plan concerning the area of Baixa-Chiado, which disallowed[71] the approval of new building permits, also contributed to the degradation of the city centre. That limitation was only lifted in 2008 by resolution of the Council of Ministers nº 192/2008[72].

The limitations imposed by the rent freeze, the complex eviction policies[73] and the restrictions imposed by the city's Master Plan led to the unattractiveness of the real estate market to investors. In its turn, the lack of investors led to a shortage in the number of available rental properties and gave rise to illegal practices by unethical landlords who took the opportunity to charge excessive rents, commonly with no rental contracts allowing for tax evasion. In light of this scenario, many people started looking at homebuying.

Following the Carnation Revolution of 1974, which ended a 41-year dictatorial regime, the new central government designed policies to incentivize home buying. In 1976, the first Portuguese Constitution on the newly democratic regime was approved, in which article 65º[74] recognized the right of housing for all its citizens. The Resolution of the Council of Ministers of February 24th, 1976 even goes to the extent of recognizing the right for every family to own their home, despite their income level[75]. The resolution also provided the creation of a subsidized credit regime for the acquisition of real estate. In 1981, Decree-Law nº 340/81 created a housing-savings account[76], with tax exemption, further incentivizing home purchase (Figure 2-23).

Buying a house became increasingly synonymous with adulthood success, working also as a financial safety net to decrease income levels during retirement. On the other hand, renting became synonymous with the lack of financial capability to buy a property, and social housing[77] was associated with low-income families, leading to social stigma. This is not unique to Portugal, as many European countries have traditionally incentivized homeownership above renting, being seen as the "natural presence" and a sign of independence, security and "personal

---

[68] Source: PORDATA.
[69] The decree-law nº157/2006 allowed for coercive rehabilitation works to be enforced by the City Council and repaid by tenants' rents. However, despite the good intentions behind it, this resulted in large financial losses for the City Council (CML, 2011, p70-73).
[70] In 2001, 48% of all dwellings were under a rental contract (CML, 2011, p61)
[71] With the exception of conservation works with no changes in the building (CML, 2011, p21).
[72] https://dre.pt/web/guest/pesquisa/-/search/440824/details/maximized
[73] According to articles 1079º to 1090º of the Civil Code, landlords would only be allowed to dissolve the rental contract after 3 months of unpaid rents. Even after 3 months, the tenant could invoke a whole set of situations provided for in articles 930-A of the Civil Procedure Code (CPC), leading to a complex legal resolution. (Braga, 2013)
[74] See https://dre.pt/web/guest/legislacao-consolidada/-/lc/337/202103262024/73938589/element/diploma#81790
[75] See https://www.igf.gov.pt/leggeraldocs/RCM_024_02_1976.htm (point nº2)
[76] In Portuguese "*conta poupança-habitação*". See https://dre.pt/web/guest/pesquisa/-/search/567419/details/normal?filterEnd=1981-12-31&filterStart=1981-01-01&q=1981&perPage=100&fqs=1981
[77] The percentage of social housing in Portuguese is extremely low. Only 2% of the housing stock is public housing (however in Lisbon and Porto that percentage rises to 10% and 12%) (Xerez et al., 2019; Antunes & Seixas, 2020)

achievement”, leading to higher housing satisfaction. It has also been traditionally seen as a way to create wealth; however, that is not accurate as the type and location of property bought can easily affect its future appreciation. People with lower incomes tend to buy property in less desirable neighbourhoods with fewer benefits. As some authors put it, “*Homeownership may be a game that all can play, but the chances of winning are heavily skewed to certain directions*” (Elsinga & Hoekstra, 2005).

In 1985, with the approval of law 46/85[78], we saw the first steps toward a friendlier environment for investors. The law laid the foundations for a free market, allowing the rent value to be freely agreed upon between landlord and tenants in new contracts. New amendments to the law were gradually made in 1990 and 2006[79], but only in 2011, in the context of the Troika memorandum, were significant changes made leading to the approval of Law 31/2012[80] (Antunes & Seixas, 2020).

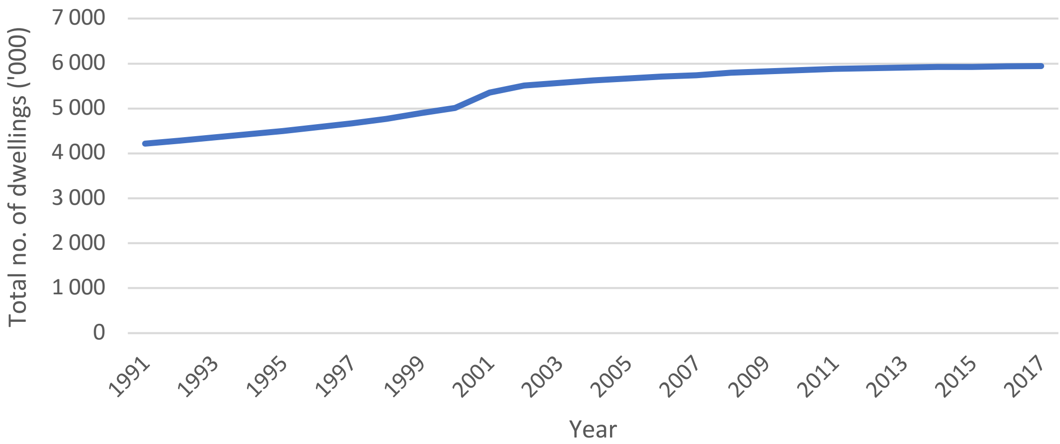


Figure 2-23 - Total number of dwellings for family housing (in thousands). Source: INE

This construction shift led to the rental market's deterioration (Associação Lisbonense de Proprietários, 2011).

According to the 2001 census, 75,4% of families own their home, the highest percentage ever recorded. Lisbon Metropolitan Area (LMA) has seen one of the most significant increases in homeownership, from 34,8% in 1981 to 79,4% in 2001 (Xerez et al., 2019).

In 2012, the government introduced a New Urban Lease Regime update as part of the troika agreement, which deregulated the private rental market by providing landowners and renters more freedom in the stipulation of urban rental agreements (Lei n° 31/2012). This new regime created a dedicated scheme for resolving existing subsidized rentals, thus enabling landowners to quickly relocate properties in the market (Lestegás, 2019; Amore et al., 2020). Previously, the

[78] For more information see https://dre.pt/pesquisa/-/search/177451/details/maximized
[79] Law 6/2006 was still very much protective of the elderly, limiting the increases in rent for tenants over 65 years old to a minimum of 4% of the dwelling's taxable patrimonial value over ten years. For more information see http://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=691&tabela=leis
[80] See http://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=1779&tabela=leis

2006 New Urban Lease Regime (NRAU[81]) tried to tackle this problem; however, economically deprived tenants, the severely disabled and older contracts with elderly tenants (over 65 years old), which comprised the large majority of lower rents, remained protected. Law nº 31/2012 made it possible for landlords to transform older contracts (prior to 1990) into time-limited contracts with rents consistent with free-market prices, leading to a few thousand evictions in downtown areas and a process of gentrification and conversion residential dwellings into tourist accommodation (hotels, hostels, and short-term rentals) (Antunes and Seixas, 2020). Smith (1996:67, as cited in Lestegás (2019)) argued that this gentrification constituted "*a back-to-the-city movement all right, but a back-to-the-city movement by capital rather than people*".

2.8.2.3. Golden Visa Programme

In 2012, the Portuguese Government amended the Portuguese Immigration Law (Lei n.º 23/2007[82]) to create a Residence Permit for Investment Activity (ARI[83]), commonly known as the Golden Visa Programme[84]. The programme is not unique to the country. Several European countries, such as Spain, Greece, Malta, Cyprus and the UK, have similar programmes ongoing (Transparency International, 2018). Under this programme, the beneficiary of a Golden Visa would be allowed to reside in Portugal, work in the country, freely move in the "Shenghen" area[85] and apply for Portuguese citizenship after five years. To apply to this programme, one would have to fulfil one of eight criteria, of which two were directly related to real estate activity:

1. The acquisition of real estate property with a value equal to or greater than 500 000€.
2. The acquisition of real estate property built at least 30 years ago or located in an urban rehabilitation area[86], and carrying out rehabilitation works in a global amount equal to or greater than 350 000€.

According to Foreigners and Frontiers Services (SEF[87]), between October 2012 and June 2021, 9.834 permits were issued with China (50.26%), Brazil (10.41%), Turkey (4.75%), South Africa (4.13%), Russia (3.89%) being the most representative nationalities among beneficiaries. Roughly 94% of the visas were issued based on the abovementioned two criteria, of which 90.5% were based on the acquisition of a property valued over 500 000€ and the remaining 9.5% based on the acquisition of property for rehabilitation. According to a Transparency International (2018) report, Portugal, on average, earns €670 million annually through this programme. Only Spain (976 M€) and Cyprus (914 M€) have reported higher figures. According to the same source, from 2012 to 2018, Portugal managed to attract a total of 4 B€, most of it invested through real estate.

[81] Novo Regime de Arrendamento Urbano (NRAU) – Law nº 06/2006
[82] For more information see http://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=920&tabela=leis
[83] Autorização de Residência para Actividade de Investimento (ARI)
[84] For more information https://www.goldenvisaportugal.pt/pt/golden-visa/beneficios-do-programa/
[85] For more information see https://www.schengenvisainfo.com/schengen-visa-countries-list/
[86] Note that Lisbon comprises a large area dedicated to urban rehabilitation, as seen in Annex 1. This came as a consequence of the state of degradation of a large proportion of its buildgins stock (see for reference CML (2011))
[87] Serviço de estrageiros e fronteiras (SEF). For more stats information, please visit https://www.sef.pt/en/pages/conteudo-detalhe.aspx?nID=51 (last visited 8 September 2021)

In 2014, an article[88] published by real estate broker JLL reported that the government must have largely underestimated the impact of the Golden Visa in the country's real estate market and that the demand had almost completely exhausted the existing high quality housing stock in Lisbon's area. The report pointed out the quality of real estate properties, the low prices and some of the highest *yields* across Europe, and the country's safety, as some of the most important points considered by investors.

The demand for real estate properties, especially in Lisbon and Porto, in the middle to high-end of the housing market fueled the development of this sector, leading to higher prices and a decrease in affordable housing in the city centre. Notably, on the 22nd of December 2020, the Council of Ministers approved an amendment to the Golden Visa programme, restricting investment to counties outside de large metropolitan areas of Lisbon and Porto[89]. The new regime came into effect on July 1st, 2021, with a transition period of 1 year. After it is concluded, new permits will not be issued for housing in the two largest metropolitan areas.

2.8.2.4. Short-Term Rentals

Founded in 2008, Airbnb is one of the dominant organizations in the so-called sharing economy of the housing market. Its impact on the housing market in prime urban tourism areas like Lisbon should, therefore, not be disregarded (Mermet, 2017; Lestegás et al., 2019; Oskam, 2019; Amore et al., 2020). In September 2021, *Inside AirBnB* reports a total of 16 230[90] AirBnB active listings in the city of Lisbon: 74.8% entire homes/apartments, 23,8% entire rooms and 1,4% shared rooms. According to the same source, 70.9% of hosts have multiple active listings, managing more than a single property. A recent study suggests that the top 25 homeowners in AirBnB in Lisbon have at least 60 active listings on the platform, with a personal gain of nearly 25 M€ (Rio Fernandes et al., 2019 as cited in (Amore et al., 2020)). According to Lestegás et al. (2019), the parishes of Misericórdia, Santa Maria Maior, Santo António and São Vicente face the greatest tourist pressure. The historic core, the five civil parishes with more than 1000 listings each, Santa Maria Maior (3101), Misericórdia (2623), Arroios (1893), Santo António (1207), and São Vicente (1159), concentrated, in 2018, 67.8% of the total number of listings in just 10.8% of the municipal area (see Figure 2-24).

[88] See https://www.jll.pt/pt/estudos-e-tendencias/investimento/o-golden-visa-e-uma-arma (last visited 7 September 2021)
[89] https://www.portugal.gov.pt/pt/gc22/governo/comunicado-de-conselho-de-ministros?i=391 (point 11)
[90] Source: Inside AirBnB. http://insideairbnb.com/lisbon/?neighbourhood=neighbourhood_group%7CLisboa&filterEntireHomes=false&filterHighlyAvailable=false&filterRecentReviews=false&filterMultiListings=false (last visited 9 September 2021)

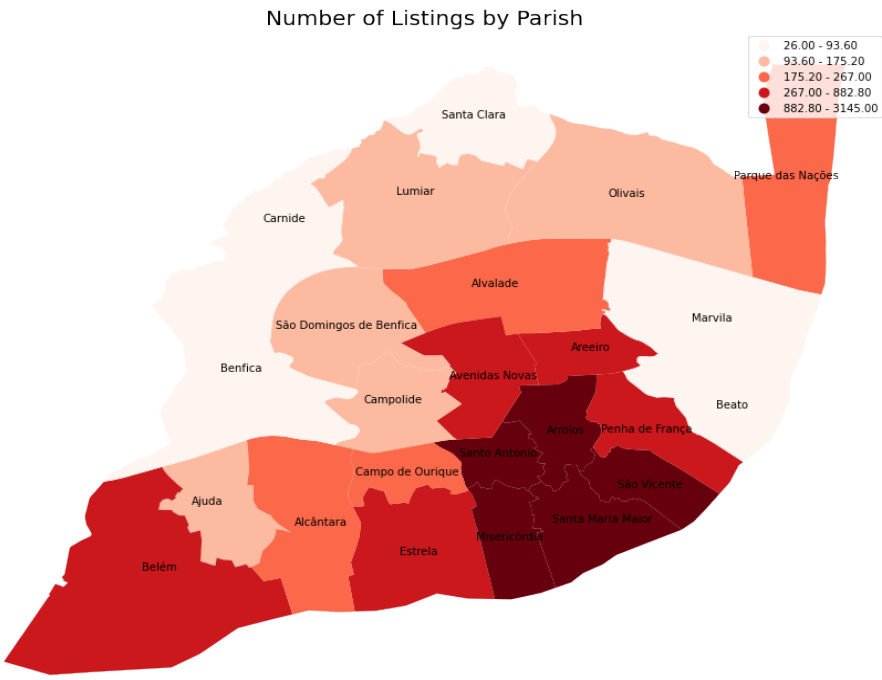


Figure 2-24 - Listings by parish. Own elaboration. Data source: InsideAirbnb (Sep 2021)

In September 2021, the same tests were run under the scope of the present research effort and found that 68.67% of listings were located in these same five parishes (2021 data) (see Table 2-4). Alcântara, Arroios, Belém and Estrela have also been facing significant pressure.

Table 2-4 - Comparison with Lestegás et al. (2019). Own elaboration. Source: InsideAirBnB

Parish	2018	2021	%Change
Arroios	1893	1765	-6.76%
Misericórdia	2623	2331	-11.13%
Santa Maria Maior	3101	3145	1.42%
Santo António	1159	1105	-4.66%
São Vicente	1207	1052	-12.84%

Despite the decrease in the number of listings, which may have resulted from the decrease in tourism during 2020-2021 due to the COVID-19 pandemic, it comes as no surprise to find that these have become some of the most expensive areas in the city, with average prices ranging 3460€ to 5425€ per square meter (Figure 2-25).

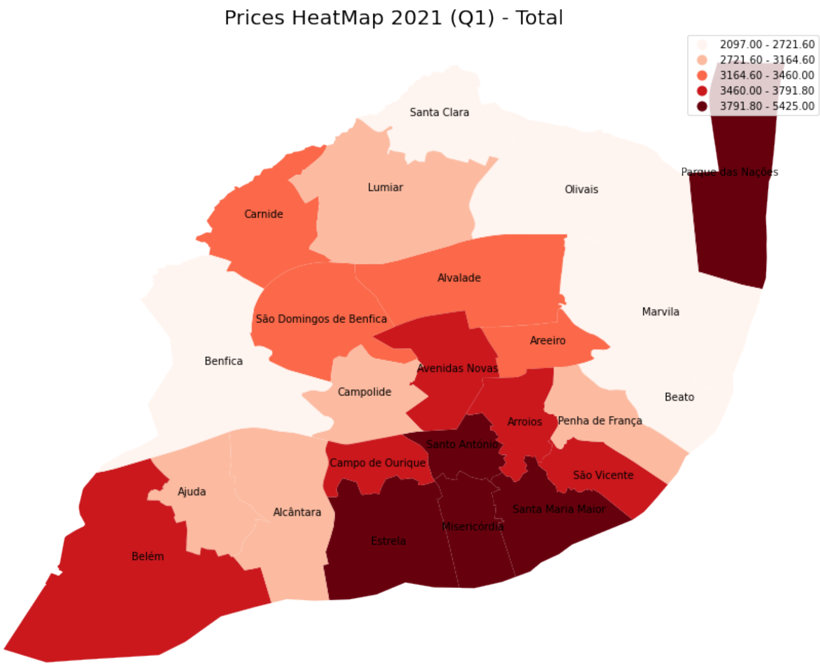


Figure 2-25 - Median Sales Prices (1Q 2021). Own elaboration. Source: INE.

According to InsideAirBnB[91] data, the parishes of São Vicente (89.35%), Santa Maria Maior (87.47%), Misericórdia (86.78%), Parque das Nações (84.84%), Estrela (82.98%) and Ajuda (81.95%) all present more than eighty percent of its listings comprised of entire homes / apartments (Figure 2-26).

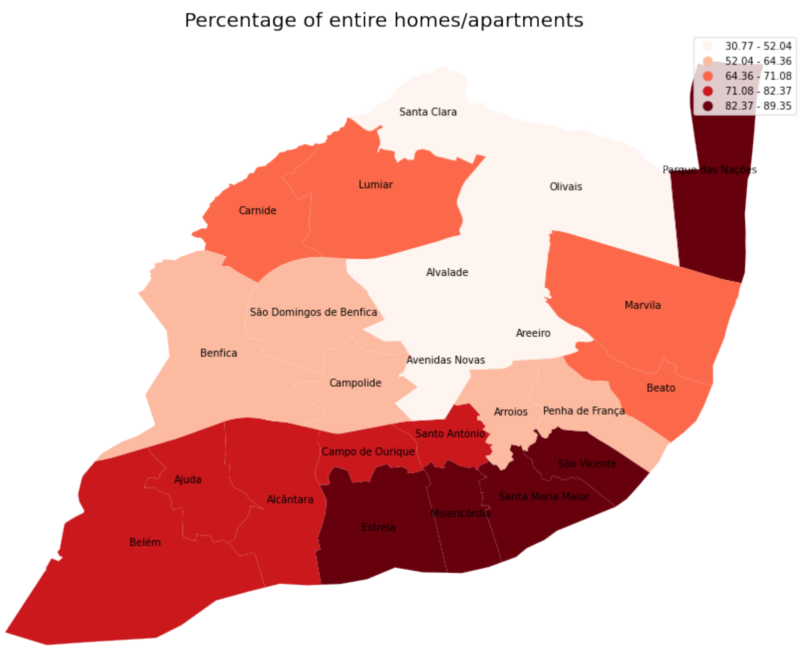


Figure 2-26 - Percentage of entire homes/apartments listings by parish. Own elaboration. Source: InsideAirBnB (Sep 2021)

[91] http://insideairbnb.com/get-the-data.html

In 2015, only three European countries provided nearly half of the total number of Airbnb guests in Lisbon: France (27%), Germany (10%), and the UK (9%) (Lestegás et al. 2019). According to Eurostat (2021), median hourly earnings are much lower in Portugal (5.37€) than in France (15.35€), Germany (17.23€) and the UK (15.2€). This influx of tourists makes it nearly impossible for Portuguese households, with far lower purchasing power, to compete with the rates paid, causing a shortage of available housing for sale and rent.

Cocola-Gant and Gago (2019) analyzed the impact of Airbnb (short-term rentals) in one of the most traditional neighbourhoods in Lisbon, Alfama, between 2015 and 2017. They found strong evidence of tourism-driven displacement caused by a buy-to-let investment strategy from investors (foreign and domestic) using housing as an asset to store capital, taking advantage of the income provided by short-term rentals (STR). In addition, compared with traditional long-term rentals, short-term rentals provide increased liquidity, given by the flexibility to sell the asset at any time. According to their findings, 78% of Airbnb landlords are individual and corporate investors. In the case of Alfama, only 1% of the supply of Airbnb properties is offered by residents sharing their homes. The authors have also identified several techniques used by these landlords to harass renters and force them to move out, which impacts not only the market itself but also its social structure and, ultimately, the uniqueness and tradition that made the neighbourhood attractive in the first place. One of the reported techniques is what one might call "selective rehabilitation", where landlords renovate buildings to open STR while neglecting low-income and elderly tenants under traditional rentals, which are seen as barriers to capital accumulation. Additionally, Cocola-Gant and Gago (2019) identified buy-to-leave strategies in which apartments remain vacant, working purely as a store of capital.

Overall, the phenomena of financialization of rental housing in which multi-storey buildings are converted from its primary function of housing to a pure financial perspective of generating yield and working as a store of value, have driven the gentrification process and contributed to the increase in demand and subsequent increase in prices. This gentrification process affects tenants and homeowners driven by negative externalities of tourism. Amore et al. (2020) compared the phenomena of STR in three major European cities: Athens, Lisbon and Milan. Their findings are aligned with Cocola-Gant and Gago (2019) concerning Lisbon. The authors even state that the "*quick rental turnaround observed in the Alfama-Mouraria is reflection of the surge in property values observed in main urban tourist neighbourhoods of Baixa and Avenida da Liberdade",* building on the idea of reinforcement of price increase driven by STR. According to the study, the policies that regarded STR as a solution for the economy, clearly overlooked socio-economic disparities that emerge between existing residents and tourists and its effects on the market.

Despite the negative consequences of STR, other studies found examples of positive impacts of these platforms, namely the fact that they provide a more convenient accommodation service for customers (Oskam, 2019), create employment opportunities (Romão et al., 2018), create a higher volume of visitors to secondary tourist cities (Ioannides et al., 2018) and even may drive the rehabilitation of the city centre (Lestegás et al., 2019). However, it is undeniable the impact it has on the gentrification process and the increase in prices forced by the reduction of available and affordable housing (Cockburn Association, 2018, p2; Amore et al., 2020).

In 2018, an amendment to the law for local accommodation (Law nº62/2018[92]) allowed City Councils to restrict the number of STR. Under this amendment, the Lisbon City Council has applied restrictions to specific neighbourhoods, disallowing any new requests for STR licenses in Bairro Alto, Madragoa, Castelo, Alfama, and Mouraria, where, at the time, more than 25% of housing was already assigned to STR (Francisco, 2018). However, the clause does not apply to pre-existing listings in these areas (Amore et al., 2020).

Despite the restrictions, from 2016 to 2021, we have seen the continuation of the trend in prices, especially in the city centre (see Figure 2-25), with median sales prices doubling in the parishes o Arroios, Santo António, Penha de França, São Vicente and Ajuda, over just five years

In many ways, during the last decade, the real city has been demonstrating a focus on the *city users* (investors and tourists) rather than the city residents (Martinotti, 1993; Lestegás et al., 2019). This focus has been reflected in the soaring of real estate property prices and increasing the value-at-risk.

2.8.3. Seismic Risk

Previous studies and events have demonstrated that the nation's capital city is at risk (Lopes, 2008; Simões et al., 2012). The Portuguese territory is of moderate seismicity, given the proximity to the Gibraltar-Azores fault. While most activity can be characterized as of low intensity (M<5,0), events of moderate to high intensity (5,0 ≤ M < 7,8) may also occur sporadically (Borges et al., 2001; Sotto-Mayor, 2006; Sousa Oliveira, 2008; Santos et al., 2019). Although the Azores register the highest number of events[93] across the country due to its proximity to the Azores Triple Junction (ATJ) (Caldeira et al., 2018), the mainland has also been severely affected throughout its history.

The earthquakes of 1531 and 1755 have largely been considered the two most devastating to hit the Portuguese capital. Although it would be fair to assume that the 1531 event had a similar extent, the largest reported and most devastating natural disaster to hit Lisbon was the 1755 earthquake, estimated to have reached an 8.70 magnitude (Appleton, 2001; Spence, 2007). The sheer size and impact of the event shocked not only the Portuguese people but the rest of Europe, being one of the most impactful natural disasters in Europe's 18th-century history[94]. At the time,

[92] https://dre.pt/home/-/dre/116152179/details/maximized

[93] Since its discovery in the 15th century, Azores have registered 33 earthquakes with an intensity equal to or greater than VII, causing 6.300 deaths and widespread destruction. The dynamics of the Azores triple junction (ATJ) plate boundary and volcanism are the main causes of these events (for more information, see Caldeira et al. (2018)).

[94] The event had the immediate attention of Europe's great thinkers at the time, having influenced the enlightenment movement (see for reference Voltaire (1756), Rousseau (2005), Yeats (2015), Braga (2017). More recently, it has been described in scientific publishing but also portrayed in documentaries for different occasions (see for instance, "O terramoto de Lisboa de 1755" (2005), " Wild Planet – Lisbon in Jeopardy" (2009) , "God's Wrath" (2013))

Lisbon was the cosmopolitan capital of a wealthy empire stretching out from Brazil to the far east. According to historical records, it was home to between 110.000 to 350.000 living in a total of 30.000 to 50.000 dwellings (Santos, 2008; Sousa et al., 2012). On the morning of the All-Saints' Day, 1st of November 1755, most people were serving their religious duties, attending masses in the several churches and chapels spread across the city (Figure 2-27).

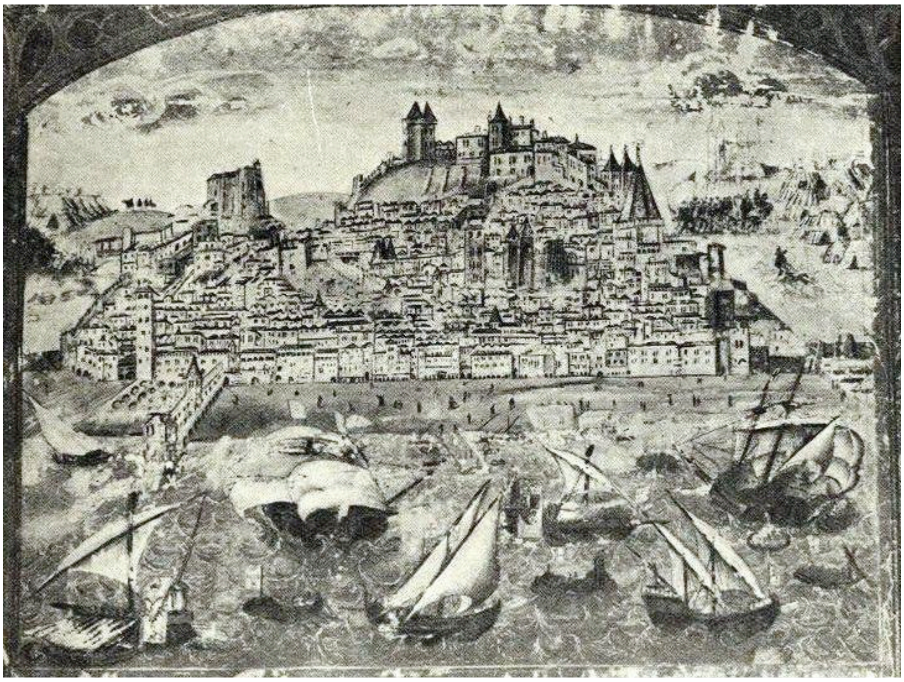


Figure 2-27 - Lisbon before the earthquake (1755). Source: Santos (2008)

In the event's aftermath, thousands of people lost their lives. Depending on the source, the number of deaths ranges from 6.000 to 8.000 (Santos, 2008), to 10.000 (Costa et al., 2012; Santos et al., 2019) to as high as 20.000 (Sousa et al., 2012). Some records reportedly estimate the number of deaths at 100.000, although it may have been exaggerated, especially compared to other sources (Santos, 2008). The event also led to the damage or collapse of 14.000 buildings (Costa et al., 2008; Sousa et al., 2012). The direct economic cost is estimated to have been between 32% and 48% of Portuguese GDP (Pereira, 2009). It is noteworthy that the event changed the economy, urban landscape, and city limits. For example, Santos et al. (2019) showed that the coastline of downtown Lisbon changed significantly since the earthquake, increasing the available land area from 50 m to about 300 m.

Sebastião José de Carvalho e Mello, the first Marquis de Pombal[95], as later came to be known, was a central figure in the awakening of the earthquake event. Mello led the city's reconstruction, based on a plan proposed by General and Royal Engineer [96] Manuel da Maia and put into practice with the help of military engineers Eugénio dos Santos and Carlos Mardel. The reconstruction was initiated right after the event, in 1756, and took approximately one hundred years to be concluded (Santos, 2008). With the seismic event in mind, new construction techniques were

[95] in Portuguese, *Marquês de Pombal*
[96] General e Engenheiro Mor do Reino de Portugal

implemented, giving rise to the “Pombalino” buildings. These buildings were characterized by their wooden structural frames, commonly known as “*gaiola*” (in English “*cage*”), which provided structural robustness and resistance, specially designed to the occurrence of seismic events (Figure 2-28).

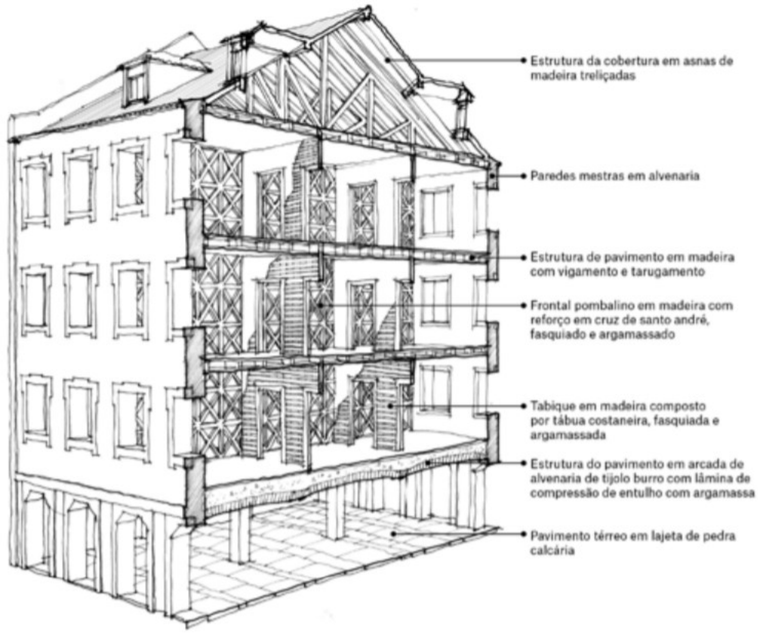


Figure 2-28 - Pombalino building schematic structure. Source: OARSR (2016)

There is a clear distinction in Portuguese housing and urban planning between the pre-and post-1755 earthquake period, with the adoption of suitable building materials and construction techniques. In a broader sense, one may say that the history of building seismic upgrades in Portugal is mainly related to earthquake disasters (such as the 1755 earthquake) or building code enforcement (Costa et al., 2012; Lopes, 2008; Jarimba, 2016).

2.8.4. Risk Assessment

Assessing the building stock’s risk is a crucial step in understanding the impact of a natural disaster. Several authors have focused on the risk of Lisbon’s specific building typologies, allowing for an in-depth analysis of each structural specificity. This section will cover some of the most relevant studies on each building typology. The case study section will present a detailed description of each typology.

Simões et al. (2015) conducted a seismic performance-based assessment analysis of “Pombalino”, “Gaioleiro”[97], and “Placa”[98] buildings and found a significant seismic vulnerability in every typology. The author also found they do not fulfil the requirements for the ultimate limit state

[97] type I - strait size façade walls and one side shaft – see for reference (Appleton, 2005)
[98] “rabos de bacalhau” sub-type

as defined in EC8-1 (Anexo Nacional do EC8, 2009) and Italian Code (NTC, 2008) and urged for improvements in the seismic capacity of these buildings.

Mendes and Lourenço (2010) studied the "gaioleiros" buildings through numerical analysis on a reduced scale 1:3 model, carrying out nonlinear dynamic analysis with time integration and pushover analysis. They concluded that even buildings with appropriate floor-wall connections are at the limit of their loading capacity under seismic action, according to EC8 National Annex (Anexo Nacional do EC8, 2009). During the analysis, the authors also concluded that "gaioleiros" follow the typical collapse of masonry buildings, including (a) cracking around the corners of the openings; (b) out-of-plane collapse (4th floor's piers). Neves (2016) studied the seismic resistance of "gaioleiros" buildings using nonlinear analysis (pushover). The flexible wooden floors cannot provide the stiffness needed to uniformly distribute horizontal loads across the vertical masonry walls. He also confirmed a difference in the stiffness between the two main directions, further contributing to its vulnerability. He concluded there was sufficient evidence to confirm the high vulnerability of "gaioleiros" to seismic events.

Simões (2018) studied the seismic vulnerability of the unreinforced masonry buildings constructed in the transition between the 19th and 20th centuries in Lisbon, based on a group of 3 buildings representative of sub-type of buildings, using nonlinear static and dynamic analysis. Considering the earthquake-resistant code for Lisbon (action type 1 with a return period of 475 years), the results indicated type I buildings[99] to have approximately 50% probability of having hefty damage and more than 30% of probability of collapse. According to the author, the results "*put in evidence the high seismic vulnerability of these buildings and the urgent need for the structural intervention and for the design of retrofitting measures in order to reduce potential losses due to future earthquakes*".

Marques et al. (2017) conducted a cost-benefit analysis of seismic retrofitting for old residential buildings in Lisbon, stone masonry and "Placa" buildings. The results concerning stone masonry buildings are consistent with the studies mentioned above, with the author highlighting the difference in resistance in the two main directions. "Placa" buildings were also found to be unsafe, mainly due to their irregularity in elevation and heterogeneity in the distribution of structural elements. However, the author found that the seismic performance of low height "Placa" building can be increased by 60% using a relatively cheap solution (2.6% of the building value) by combining steel jacketing of RC beams with a partial application of reinforced plaster on a wall.

Caruso and Bento (2019) evaluated the structural seismic behaviour of an old reinforced concrete building (pre-1980). Using a case study building designed under the 1958 old Portuguese Code for Reinforced Concrete and for Earthquake-Resistant Design (RSCCS) [100], they found it only to withstand low seismic forces and had insufficient seismic detailing. Jarimba (2016) have also focused on the seismic vulnerability transition concrete buildings built during 1960-1980 in the

[99] buildings with small size façades and one flat per floor
[100] RSCCS, Regulamento de Segurança das Construções Contra Sismos . DL 41658. (in Portuguese), Lisbon,1958.

neighbourhood of Alvalade. The author states that most of these buildings were built based on unclear regulatory frameworks, which may have led to an inferior seismic building design. She performed a linear and non-linear (pushover) analysis on a sample building model and concluded that the structural behaviour of this typology is conditioned by the type of actions applied to the building: however, it shows an inadequate resistance to seismic action, especially on upper floors.

Several studies have been conducted throughout the last decades on a broader building stock risk assessment. The city council has long shown awareness of the seismic risk through multiple initiatives. One of the earliest examples was the development of an Earthquake Damage Simulator by the Municipal Civil Protection Service (SMPC/CML) [101] in partnership with the scientific community (CGUL/FCUL, DGEOL/FCUL, IST, LNEC, ISEL, among others). This was integrated into the Emergency Plan for Seismic Risk[102] (PERS), leading to the deployment of a GIS-based solution as part of the emergency response and preparedness plan (Pais 2002). This plan was based on previous studies assessing Lisbon's seismic risk (see for reference Sousa et al. (1992); Pais et al., (1996a) and (1996b); Pais (2001)). To develop the simulator, vulnerability and fragility functions were built for each building typology. This allowed the municipality estimate damage percentage and financial losses, the number of damaged buildings and their corresponding damage level, the probability of collapse of each typology, and the percentage of population affected given the structural damages. Infrastructures, such as transportation, gas and electricity networks, were also accounted for in the model. The city's population variation throughout the day was also considered in the model, given the large influx and outflux of people during the business days. By inputting epicentre coordinates, the magnitude of the event, and the hour, the system would produce a spatial estimate of damage to the building stock and infrastructure and estimate the number of casualties, injured, and displaced people (see Figure 2-29). The study concluded that the historic centre and adjacent areas faced the most risk.

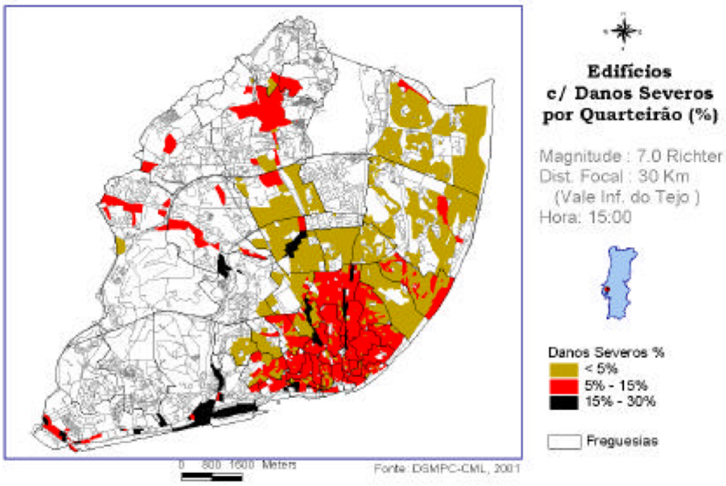


Figure 2-29 - Estimated buildings with severe damage (%) per city block. Source: Pais (2001)

[101] Serviço Municipal de Protecção Civil da Câmara Municipal de Lisboa (SMPC/CML)
[102] Plano de Emergência para o Risco Sísmico (PERS)

LNECLoss consisted of a software simulator for Lisbon and was initially promoted by the National Service of Civil Protection and developed by The National Laboratory in Civil Engineering (LNEC)[103]. It analyzed the earthquake risk not only for a specific building typology but for the overall building stock (Costa et al., 2004; Afonso, 2006; Costa et al., 2008). One of the scenarios simulated a 1755-like event in Lisbon Metropolitan Area (LMA).

The LNEC's contribution to the project allowed to:

1. Characterize geology and seismic action in LMA;
2. Identify and characterize the vulnerability of elements at risk;
3. Assess damages to the building stock and loss of human lives;

In order to assess the vulnerability of the building stock, the authors used information regarding the number of storeys, building year and type of elements used in the construction. By crossing this data, the authors could simulate the building's behaviour based on its building typology[104], using Displacement-Based Seismic Assessment (DBA). The authors have estimated up to 12448 deaths in LMA for a seismic scenario similar to 1755 (Costa et al. 2008). It also states that the low perception of seismic risk among the Portuguese population may result from the low number of casualties caused by earthquakes, in continental Portugal, during the 20th century (Sousa, 2006).

Among its main findings, it is possible to highlight:

1. The economic risk of a seismic event in Lisbon's Metropolitan Area (LMA) was evaluated regarding the replacement cost of the residential building stock. Economic losses vary from 1.3%, for 95 years return period, up to 38%, for 5,000-year return period, of the total replacement cost of LMA residential building stock.
2. After implementing a given strengthening strategy, based on selective retrofitting interventions and applied to typological building classes responsible for the larger seismic economic risk referred to above, it was concluded that economic risk could be mitigated by an 36% for all return periods.
3. The maps of loss estimation after the implementation of the proposed mitigation actions show that the chosen mitigation action has predominant effects on the south margin of Tagus River, where intermediate soils prevail.

At the time of writing, no Web-based simulator for LNECLoss was found, and according to the EC website[105] the LESSLOSS project ended on the 31st of August 2007. Nevertheless, the LNECloss

[103] Laboratório Nacional de Engenharia Civil (LNEC)
[104] Each building was assigned to one of 49 building typologies based on its characteristics
[105] https://cordis.europa.eu/project/id/505448 (last visited 7 September 2021)

tool can be considered a landmark project on the seismic simulation of earthquake scenarios and possible outcomes.

Teves-Costa et al. (2011), analyzed loss scenarios based on the existing housing stock (see Figure 2-30).

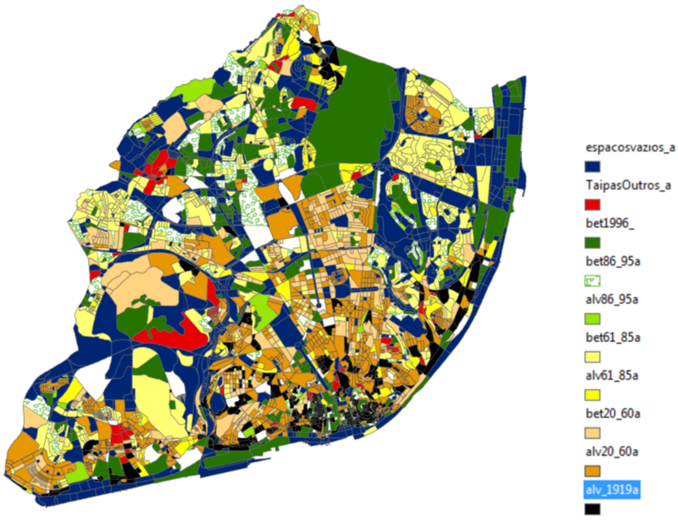


Figure 2-30 – Main building typologies by statistical unit. Source: Teves-Costa et al. (2011)

With a sample of 2500 buildings from the Lisbon City Hall Database, they have applied the EMS-98[106] scale and calculated vulnerability and fragility curves for each typology. Based on this information, the geology and topography of the city, they have created two possible scenarios: (1) one considering a near epicenter, with a macroseismic intensity of VIII (EMS98) and (2) one considering a distant epicenter with a macroseismic intensity of IX (EMS98). Figure 2-31 and Figure 2-32 show the percentage of building with level 4 (very heavy damage) by sub-statistical unit.

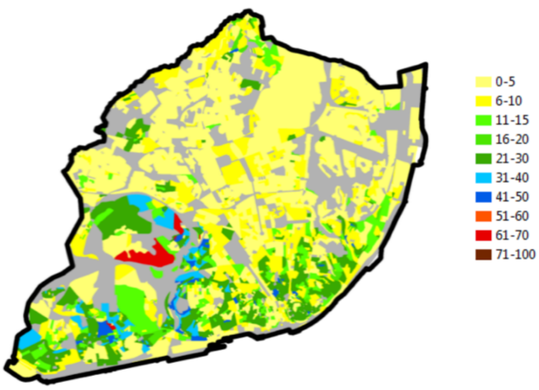


Figure 2-31 - Level 4 damage due to near distance earthquake (%) (level VIII intensity). Source: Teves-Costa et al. (2011)

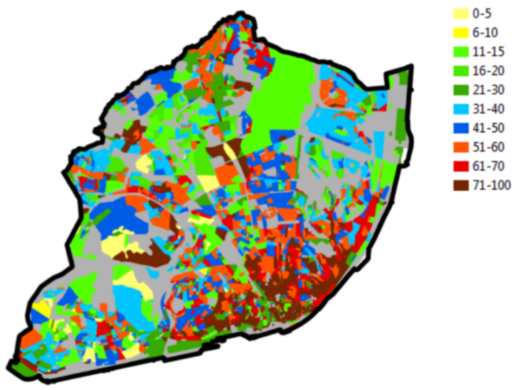


Figure 2-32 - Level 4 damage due to distant earthquake (%) (level IX intensity). Source: Teves-Costa et al. (2011)

[106] The EMS98 scale considers 5 damage levels, from 1 (negligible to slight damage) up to 5 (destruction). Level 4 represents very heavy damages. For more information, please refer to EMS98 or consult the summary table in Annex A.

In 2005, the City Council released their Sectorial Studies on Seismic Risk[107], with the collaboration of The Institute for Structural Engineering, Territory and Construction from Instituto Superior Técnico (ICIST)[108]. The study aimed to produce recommendations to be included in the city's Master Plan such as the definition of critical seismic areas. The document also contains recommendations regarding interventions in existing buildings, such as the disallowance of any intervention that may decrease the global resistance to horizontal loads, such as the construction of new upper floors. Special care is recommended for "pombalinos" and "gaioleiros" buildings. Regarding the rehabilitation of buildings, it is recommended that any rehabilitation project should be required to contain detailed information on the building's seismic resistance explicitly and that maintenance works of more vulnerable buildings should be used as an opportunity for its structural retrofitting. It also recommends a tighter oversight of construction works by the Lisbon City Council. In this study, the authors also recommended the update of the Earthquake Damage Simulator developed in 2001, to reflect the changes in the population dynamics and in the city's urban landscape. Summary tables (Table A.0-3 and Table A.0-4) containing the above-mentioned studies and their key findings can be found in Annex A.

2.8.5. Building Stock

Lisbon's building stock can be divided into seven main building typologies (Lopes, 2008; Jarimba, 2016; Neves, 2016). However, as evolution is a continuous phenomenon across time, this classification should not be interpreted strictly since it may vary slightly across locations. The building typologies are as follows:

1. **Old mansory buiildings (prior to 1755 earthquake)**
   These are buildings composed of thick masonry walls and wooden floors. They present extreme irregularity in the size of its structural elements (namely masonry stone), mortars used and overall quality of construction. Thus, it is difficult to characterize in one single class the structural resistance of these buildings. Some authors even argue that it is difficult to define a specific building typology in pre-1755 buildings, as they emerged across several centuries and with no proper urban planning (Simões et al., 2012). However, some features appear to be common in Lisbon's old masonry buildings. The wall thickness is usually reduced from bottom to top and has also been reduced over time, so newer buildings will likely present thinner walls. It is noteworthy that many of the existing old masonry building stock suffered significant structural changes, with the addition of extra upper floors or even elevators (Simões et al., 2012). Some authors even divide this class into two sub-classes (OASRS, 2016):

[107] For more information see https://www.lisboa.pt/fileadmin/cidade_temas/urbanismo/DELETE_documentos/estudo_sobre_risco_sismico-1.pdf (last visited 8 September 2021)
[108] Instituto de Engenharia de Estruturas, Território e Construção do Instituto Superior Técnico (ICIST)

- **Buildings with "ressalto" floors** – Mainly built during the medieval period (Figure 2-33), prior to the 1499 Royal Letter by D. Manuel, which required the demolition of advanced building facades. It also required the new buildings to use masonry instead of wood in its construction process as a measure of risk reduction to fires and sanitation, with aligned facades and building heights;

Figure 2-33 - Building with "ressalto" floors

- **"Seiscentistas" buildings –** built during the sixteenth century and until 1755 (Figure 2-34);

Figure 2-34 – Casa-Museu Amália Rodrigues ("Seiscentista" building")

Most of these typologies can be found in the old part of the city, mostly in downtown areas such as Castelo, Alfama and Mouraria and Bairro Alto (Appleton, 2001)

2. **“Pombalino” buildings (built during the 1755 reconstruction)**

The 1755 earthquake was a turning point in the city landscape and construction techniques. During the reconstruction, a new building construction typology was created, with a more robust structure and less vulnerability to seismic activity. This type of construction extended for almost a century and was based on prefabricated structural elements (OASRS, 2016; Jarimba, 2016). Wooden vertical and horizontal elements were added to the facades to increase resistance of the masonry wall and to connect them to the three-dimensional wooden frames that composed interior walls, the so-called “gaiola pombalina” (Simões et al., 2012). This wooden truss, commonly known as “Cruz de Santo André” (*in English* “Saint Andrews’ cross”), was then filled with masonry elements (Neves, 2016). The regularity in construction within each compound also provided added structural stability to “pombalino” buildings. Although the added resistance, these buildings still present weaknesses. As an example, one might consider the variation in stiffness across height, caused by the use of different elements across the first floor (built entirely in masonry) and the upper floors (built upon wooden frames), one of its most notable fragilities (Simões et al., 2012). The construction of new upper floors might build upon these fragilities, increasing safety concerns in case of no structural reinforcement (Appleton, 2001). This type of building can be seen in the downtown area, also known as “baixa pombalina”.

3. **“Gaioleiro” buildings (built the last third of the XIX century, corresponding to a phase of large city expansion)**

The city of Lisbon was subject to a significant increase in its population during the last third of the XIX century. Between 1890 and 1911, the city saw its population rise from 301 206 inhabitants to 435 359 inhabitants, a 45% increase over 21 years (Simões, 2018). Hence, there was a need for a rapid expansion of the city's urban landscape. Frederico Ressano Garcia, an engineer and politician, was responsible for designing the city’s expansion to the north through the construction of the *Avenidas Novas* and *Bairro Camões* and consolidating other areas such as *Campo de Ourique*. Due to the pressure to expand the city quickly, the new city plan was flexible in architectural details, with no definition concerning maximum heights, widths, or occupations leading to a heterogeneous building landscape (Appleton, 2005). However, in the absence of a strict and defined plan, real estate developers and builders focused on profits rather than safety, using poor quality materials and changing the construction of structural elements at their will (Lopes, 2008; Neves, 2016). The construction was, in some instances, so poorly conducted that it led to the collapse of several buildings during the construction process (Appleton, 2001). The fading memory of the 1755 events and the approval of the 2030 law, which allowed the replacement of older buildings with poorly constructed new buildings with a higher number of dwellings, further contributed to the deterioration of the “pombalino” construction techniques into what came to be known as the “gaioleiro” buildings (Gomes, 2011). There are several differences between the more robust "pombalino" buildings and the newer "gaioleiros". The timber diagonal elements from the front facade were gradually removed, and brick walls replaced the thick rubble walls on the lower floors and hollow or "tabique" walls on the upper floors, giving little to no structural support to out-of-plain movements (Mendes & Lourenço, 2010). Furthermore, 5 or 6 storeys are common in this typology, contrasting with the "pombalinos" typical three-story design. Hence, "gaioleiro" buildings present significant vulnerability to seismic activity (Simões et al., 2012). This type of construction prevailed in the city from the 1870s until the 1930s. Nowadays, it is estimated that "gaioleiro" buildings represent 24% of all building stock in Lisbon, with a total of 14 000 units (Simões et al., 2014, 2015). In addition, many of these units have not been subject to proper maintenance, presenting signs of degradation of structural elements and increasing safety concerns (Simões et al., 2012).

4. **"Placa" buildings or "transition buildings" (wood and reinforced concrete mixed design)**

   In 1921, a collapse of a building, leading to the death of twelve workers, caused a public demonstration against the "gaioleiros" (Simões, 2018). This event paved the way for the surge of a new typology. In the late thirties and early forties, with a new urbanization plan by engineer Duarte Pacheco, a new building typology using a mix of masonry-reinforced concrete appeared in the city, the "placa" buildings. This typology introduced the use of reinforced concrete in the lower commercial areas needing larger beam spans. On the upper floors, back balconies and service rooms (kitchens and bathrooms) were also built on a lightly reinforced concrete slender slab combined with the use of timber in the remaining floor area (Simões et al., 2012; Jarimba, 2016; OASRS, 2016). It is noteworthy that the use of concrete slabs in the referred areas aimed to diminish the risk of fire by removing inflammable materials and the risk of timber rotting due to humidity. Despite the RC slabs, bearing masonry walls were still in use. This constitutes an evolution not only in the architecture but also in the engineering of new buildings, fostered by increased concern for building safety and the approval of the *Regulamento Geral de Construção Urbana* (RGCU)[109] by Lisbon's City Hall in 1930. The use of concrete slabs introduces the weight of the floor as one of the most significant in the structure. In addition, it provides the structure with increased stiffness at the floor levels. Nevertheless, the difference between the first floor (commercial area) and the remaining upper floor still constitutes a difference in stiffness across the height of the building, which is detrimental to its response to seismic activity.

   The increased regulation of the construction sector, namely by the approval of the *Regulamento de Salubridade das Edificações Urbanas*[110] (1903), *Regulamento Geral da*

[109] The Regulamento Geral de Construção Urbana (*General Law on Urban Construction)* of 1930 imposed stricter resistance and safety measures and discouraged the construction of "gaioleiro" buildings.
[110] Regulation for Health of Urban Buildings

Construção Urbana (1930), *Regulamento do Betão Armado*[111] (1935), *Regulamento Geral das Edificações Urbanas*[112] (1951) and the *Regulamento de Estruturas de Betão Armado* (1967) (OASRS, 2016) have marked this transition phase and is clearly notable in the evolution of mixed masonry-reinforced concrete buildings.

This typology remained in use until the 1960s and constitutes one of the most representative typologies in today's building stock (approx. 22%). It is found to be most expressive in the neighborhoods of *Alvalade*, *Encarnação*, *Serafina*, *Boavista* and *Areeiro* (Appleton, 2001; Jarimba, 2016).

5. **First period of reinforced concrete buildings (before 1961)**

   The first RC buildings were built in the city during the 1940s and 1950s. Initially built in public real estate developments, the market rapidly adopted the new structural approach. Bearing masonry walls were replaced by RC beams and columns in a framed structure. The first construction safety regulation against earthquakes in the country, the *Regulamento de Segurança das Construções Contra Sismos*[113] (RSCCS) was published in 1958. Thus, despite its superior structural stiffness, several of the first RC buildings were not designed considering any seismic activity, using low to moderate resistance concrete and slender concrete slabs (Lopes, 2008; Jarimba, 2016). In 1961, the *Regulamento de Solicitações em Edifícios e Pontes* (RSEP) was published, putting together all actions required for buildings' structural design for the first time. However, despite the RC design, several buildings were built using low-quality concrete with lower resistance, especially noticeable in the mechanical behaviour of columns (Appleton, 2001).

6. **Second period of reinforced concrete buildings (before 1985)**

   After 1961, the building design already accounted for seismic resistance, leading to the second phase of RC construction. However, the publishing of the Regulamento de Segurança e Ações em Estruturas de Edifícios e Pontes (RSA), in 1983, updated the requirements of actions on buildings with stricter regulations. As seen in Figure 2-35 the percentage of buildings built after 1960 is lower in the city centre and more significant in the outer areas, such as Parque das Nações, Olivais and Lumiar.

[111] Regulation for Reinforced Concrete
[112] General Regulation for Urban Buildings
[113] Construction Safety Regulations Against Earthquakes

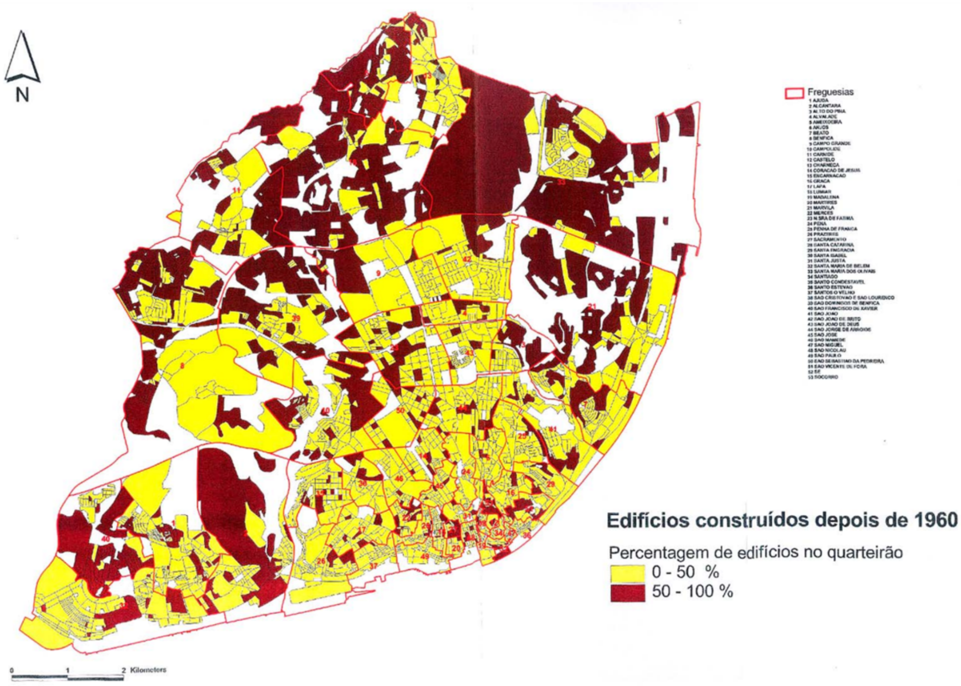


Figure 2-35 - Percentage of buildings (per block) built after 1960. Source: CMLisboa

7. **Modern reinforced and prestressed concrete buildings**

   The adoption of the *Regulamento de Segurança e Ações em Estruturas de Edifícios e Pontes* (RSA) and the *Regulamento de Estruturas de Betão Armado e Pré-Esforçado* (REBAP) layed down the foundations for modern reinforced and prestressed concrete buildings. The Eurocodes published during the last decade of the twentieth century, normalized regulations for buildings structural design across Europe.

Most of Lisbon's housing stock is not well suited to resist a seismic event (CML, 2011, p58). According to Costa et al. (2008), 27.3% of the buildings in Lisbon Metropolitan Area were constructed before the first Portuguese seismic code. That percentage rises to 63,9% if we narrow it to the city of Lisbon. Several studies have analyzed the vulnerability of its housing stock to the occurrence of an earthquake. Appleton (2001) raised awareness by describing potential scenarios caused by 1755-like events in the 21st century, detailing the expected effects in each building typology.

In 2001, the *Sociedade Portuguesa de Engenharia Sísmica* - SPES (Portuguese Society for Earthquake Engineering) and the *Grémio das Empresas de Conservação e Restauro do Património Arquitectónico - GECoRPA* (Portuguese Association of Companies for Preservation and Restoration of the Architectural Heritage), elaborated a reduction program of the seismic vulnerability of the building stock (SPES & GECoRPA, 2001), similar to the ones implemented in other countries (such as USA, New Zealand, Japan, Italy, and Turkey). The programme focused on seismic rehabilitation, conducted an extensive analysis of the vulnerability of the Portuguese building stock and aimed to draw the attention of the civil society to the seismic risk. The authors agreed on 14 best strategies to reduce vulnerability during the debate. The need for a general enforcement of the adequate building codes and inspections, the development of documents intended to supervise rehabilitation, namely the specific creation of a Seismic Rehabilitation Manual devoted to the main typologies of the building stock[114], the guarantee of adequate technical training for professionals and the involvement of various stakeholders (schools, professional associations and business associations) in the process. The *APS – Associação Portuguesa de Seguradores* (Portuguese Insurers Association) has also been pointed out as one of the important stakeholders in controlling the quality of the design and execution of the work. The programme presented a 20-year plan for the rehabilitation of buildings, but by 2007, it had not been acted on (Spence, 2007).

2.9. Final Remarks

After an extensive analysis of the Portuguese real estate markets, its legacy, trends, vulnerability and exposure of the housing stock, we may consider some concluding remarks:

- Lisbon's Metropolitan Area (LMA) contains a significant portion of the Portuguese population, economic and decisional power;
- The Portuguese capital city is considered to be of moderate seismicity, with two significant reported events in its history that changed the city landscape, causing thousands of deaths and financial losses, and shaping construction for the following decades;
- Despite the event, most of Lisbon's housing stock presents high levels of vulnerability to seismic events, especially the "gaioleiros" which constitute approximately one fifth of the city's building stock; Approximately 2/3 of the stock was built before any seismic regulation;
- The increasing population of the LMA has been increasing the exposure to risk during the last decades;
- The traditional preference for homeownership allied with the promotion of governmental policies that favoured home acquisition, the deregulation of the financial market in the 80s and 90s, the deregulation of the rental market in 2012, the surge of short-term rentals (STR), the increased number of tourists and the commoditization of the real estate market led to an extraordinary increase in property valuations and, consequently, value-at-risk;
- The low levels of earthquake risk insurance penetration, and the lack of a regulatory framework requiring mandatory insurance coverage and not allowing for the creation of a seismic emergency fund, increases the exposure of the built environment to financial risk;

[114] The creation of this document has been considered by the authors as an "irrefutable priority".

- Despite the risk, most of the population and investors are unaware of it (Appleton, 2001). Furthermore, the analysis of real estate price changes suggests that investors and homebuyers are oblivious to the risks. However, most of the household's wealth is in real estate property;
- Although there is extensive research on the building stock losses and risk assessment of the housing stock, to the author's best knowledge, there is no comprehensive study on the awareness of real estate stakeholders to the vulnerability of their houses nor their exposure to risk;

Given this context and the lack of research on the awareness of real estate stakeholders about the risks, it is utterly vital to provide an adequate understanding of the impact of seismic risk on the property market, to understand the levels of awareness amongst the population and to provide new information and recommendations. Hopefully, this could assist policy and decision-makers to act accordingly, reducing both the exposure of people and the vulnerability of property stock. Furthermore, a correct assessment and communication of risk will minimize financial losses to all stakeholders and preventable losses of human life.

3. Methodology

3.1. Overall Approach

Several key points are required to analyze the influence of risk in the real estate market, hence the methodology used should be capable of capturing the effects comprehensively. First, one should consider risk itself. Risk is often referred to as a combination of hazard, vulnerability and exposure when dealing with natural events, such as earthquakes, but simplified into the combination of likelihood and consequences in a more general framework. Therefore, one should consider not only the historical seismicity across the city (the epicentre of previous events, magnitude and ground motion effects – i.e., hazard) but also buildings' structural behaviour and response to an earthquake (vulnerability) as well as the population dynamics and demographics (to estimate the number of injured and death – i.e., exposure) (Costa et al., 2008). However, this work will mainly focus on the vulnerability of Lisbon's building stock and its differences across building typologies and neighbourhoods.

Nevertheless, the methodology will start with some considerations that should be taken into account to understand the rationale behind the methodology used. These considerations concern:

1. Building stock - a key point to having a deep understanding of the building typologies across the city and its history.

2. Ground motion - is imperative to understand in order to calculate the requirements that each building should fulfil in order to avoid collapse.

3. Spatial considerations - given that location is known to be one of the main factors in determining real estate values. This includes a number of distinct factors, such as proximity to transport, public facilities, etc.

4. Economic considerations - given that the price of a real estate asset is arguably affected by macroeconomic trends and local effects such as shifting demographics.

Although this might not be an exhaustive list of all the factors affecting the real estate market (McGinnis, 2004), these four large groups were considered to be some of the most important to the understanding of the real estate dynamics.

In the first phase of the study, different techniques will be used to analyze these values from different viewpoints to have a deeper understanding of all the factors impacting the perception of risk and its reflection on real estate values. First, a dataset with over 8.000 real estate transactions will be explored using various techniques often used in Exploratory Data Analysis (EDA), such as univariate analysis, bivariate analysis, multivariate analysis and even spatial analysis techniques. Principal Component Analysis (PCA) and feature selection techniques will also be employed. This will provide an overview of the data and allow building the following phases upon an informed basis.

In the second phase, the author will gather elements from previous works on seismic hazards and soil types in Lisbon will be gathered to estimate the vulnerability of its building stock on a building-by-building automated approach. It is important to highlight that, given the large number of properties under analysis, it would not be feasible to physically analyze each of these buildings; therefore, it was required an implementation of a set of automated techniques and simplifications allowing the attribution of vulnerability rating for each property, based on its construction year, location, the estimated number of floors and building typology. Following the work of Sousa et al. (2012) and Ferreira (2012), this vulnerability will be synthesized into a single index, considering both the actual and the required vulnerability of the building to avoid collapse, under the name of Building Resistance Deficit (BRD). This process will be described in section 4.7.

In the third phase, the impact of risk and its perception on real estate from a revealed preferences point of view will be modelled. Several classic modelling techniques, such as the Multiple Linear Regression will be used to build the hedonic price models, as well as more recent algorithms such as the Multivariate Adaptive Regression Spline (MARS) and Extreme Gradient Boosting (XGBoost). No fragility analysis will be carried out nor estimation of human or economic losses (value per building floor lost), as this work focuses on measuring the impact on actual recorded transactions.

The fourth phase will provide an understanding through the stated preferences viewpoint, using Contingent Valuation Methods (CVM), based on surveys and controlling for some response bias known in the literature. It is noteworthy that respondents' biases always influence data based on surveys. Every human interaction is known to be affected by many biases resulting from the conditions where the survey is delivered, the subject's past experiences and even the wording used. Thus, the survey build and delivery should be thoroughly crafted to minimize the effects of response bias, maximizing the quality of the data obtained.

The fifth and last phase consists of joining all of this data to build a single model encompassing both the revealed and stated preferences, providing an overview of the risk aversion of investors and real estate stakeholders towards earthquake risk (Nakagawa et al., 2009).

3.2. Building Stock Considerations (vulnerability)

One of the core topics of this work is the impact of the building's vulnerability to earthquakes which strongly depends on its structure and, therefore, on its building typology. It is not under the scope of this work to propose capacity (pushover) analysis and fragility curves but to use previous results upon which to build. In section 2.8.4, the city's buildings were classified into seven typologies, according to Lopes (2008):

1. Old masonry buildings (prior to 1755 earthquake).

2. “Pombalino” buildings (built during the 1755 reconstruction).

3. “Gaioleiro” buildings (built in the last third of the XIX century).

4. “Placa” buildings or “transition buildings” (wood and reinforced concrete mixed design, built approximately from 1930 to 1960).

5. First period of reinforced concrete buildings (before 1961).

6. Second period of reinforced concrete buildings (before 1985).

7. Modern reinforced and prestressed concrete buildings (built after 1985).

One can argue that an eighth category could be added, containing the most recent building techniques from the 2000s onward. Sousa et al. (2012), used the classification of the census 1991 tracks (building materials and epoch), and adopted seven typological classes: adobe and rubble stone (all epochs), masonry before 1960, masonry 1961 – 1985, masonry 1986 – 2001, reinforced concrete (RC) before 1960, RC 1961 – 1985 and RC 1986 – 2001. Each typological class was then subdivided into seven classes for the number of floors, obtaining 49 vulnerability classes a similar seismic response. Ferreira (2012) took a similar approach, starting from a total of 27 building typologies. However, it grouped the original twenty-seven categories into five large groups, allowing for a less rigorous but more expedite approach. The final five categories are as follows: masonry (before 1945), RC 1946 – 1960, RC 1961-1985, RC after 1985, up to five floors and RC after 1985 with more than five floors. Each of these five categories serves as the base for the vulnerability values attributed to each building, to which specific factors are added depending on the number of floors, condition of the building, existing damages, among others. In the first approach, the use of this classification to attribute the estimated vulnerability values (Vu_i), will be considered, as per the following table (Table 3-1):

Table 3-1 – Average vulnerability for each typology. Source: Ferreira (2012)

Typology	Vui
Masonry (<1945)	0,70
RC (1946-1960)	0,60
RC (1961-1985)	0,50
RC (>1985 and ≤ 5 floors)	0,40
RC (>1985 and > 5 floors)	0,44

Then, the modifying factors (FM) are added based on the following criteria listed on Table 3-2:

Table 3-2 – Modifying Factors. Source: Ferreira (2012)

Modifying Factors (FM)		**Masonry**	**RC | Regulation**		
			Low (<1960)	**High (>1985)**	**Median (1960-85)**
No. Floors	Low: (M. 1-2) | (RC 1-3)	-0,04	-0,02	-0,02	-0,02
	Median: (M. 3-5) | (RC 4-7)	0	0	0	0
	High: (M. ≥ 6) | (RC ≥ 8)	0,04	0,08	0,06	0,04
Cover	RC	0,04	0	0	0
Existing damages	A, B, D	0,08	0,08	0,08	0,08
	At least 2 structural damages	0,06	0,06	0,06	0,06
	At least 1 structural damage	0,04	0,04	0,04	0,04
	water infiltration	0,04	0	0	0
Building plan	L, T, U	0,02	0,02	0,02	0,02
Atrium/courtyard		0,02	0,02	0,02	0,02
Resistant Walls (elevator shaft)	Centered	0	-0,04	-0,04	-0,04
	Not Centered	0	0,02	0,02	0,02
Mass eccentricity		0,04	0,04	0,04	0,04
Non-uniform height		0,02	0	0	0
Openings	Large openings	0,02	0,02	0,02	0,02
Hollow floors		0,04	0,04	0,04	0,04
Short columns		0	0,02	0,01	0
Condition	Good	-0,04	0	0	0
	Bad	0,04	0,04	0,02	0

The final vulnerability attributed to the building is given by (EQ 5):

$$V_{final} = V_i + FM \qquad \text{[EQ 5]}$$

The higher the level of V_{final} (closer to 1), the less resistance a building has to seismic activity. Thus, the vulnerability value can be translated to a level of expected damage. However, there are some issues with this approach given that many epochs have an overlap of building typologies, which means that for each year, there are several building typologies that could have been attributed to the building if no additional information on its building materials is available. Additionally, the database only provides the condition of the building as a modifying factor, and this scale of “good” and “bad” is not objectively defined. Nevertheless, the Lisbon municipality confirmed that the attribution of this classification was made by the same team across the entire dataset. This implies that the existing biases are consistent throughout the database. No additional information on the building configuration or seismic design has been provided. In the first approach, the number of floors will be used as a modifying factor which was possible by joining multiple data sources, namely the building survey and the topographic survey. The quality of construction and materials used are also determining factors in a structure’s resistant capacity

(Modica et al., 2021). Unfortunately, it was not possible to assess these features in an automated way, given the information available. The entire process will be described in section 3.7.

Given the limitations of this approach, second approach was followed using a probability density function for attributing the building typology and vulnerability, considering not only the building's year but also its location at a geographical statistical unit, as per census tracks. The entire process will be described in section 3.7.4.

Although it is recognized the limitations of the methods to attribute vulnerability to each building based on its estimated building typology and location, given the impracticality of doing a tailor-made building-by-building survey to collect additional features, two different approaches were used to assess the influence on the results. Therefore, using various methods, implementing and programming scripts, several data sources were joined in order to have a more refined and expeditious way of classifying each property's vulnerability to earthquake activity.

3.3. Ground Motion Considerations

According to McGinnis (2004) "*earthquake risk can be broken down into three aspects: distance from the fault, probability of quake on nearby faults, and susceptibility of soil to ground motion produced from a quake*". This work's analysis is based on a citywide scale, thus it is reasonable to assume that soil susceptibility to ground motion will play a crucial part in explaining the differences in vulnerability across buildings, neighbourhoods and even parishes (which are the focus of this work). Even over hundreds of metres, local soil conditions can significantly influence ground shaking values. There are two main approaches to measuring ground motion intensity used in vulnerability and fragility assessment (Rossetto et al., 2014): 1) those based on macroseismic intensity scales (e.g. Modified Mercalli Intensity) and 2) those based on instrumental quantities, such as Peak Ground Acceleration (PGA)[115]. PGA measures the peak acceleration reached during an earthquake and is affected by two main determinants: distance from the fault and soil composition. The greater the PGA, the greater the damage. The approach followed by Ferreira (2012) and Sousa et al. (2012), on which this work's assessment of vulnerability is based, uses PGA as a way to estimate the intensity of ground motion[116] and the building's required vulnerability to avoid structural collapse according to the European Macroseismic Scale (EMS-98). To obtain the PGA value, a soil type was attributed according to the location of each building. It is noteworthy that each property is georeferenced, thus it was possible to join the spatial distribution of the dataset with the Lisbon City Soil Classification Map developed by Oliveira et al. (2019). The process will be described in sections 3.7.2 and 3.7.3.

[115] According to Rossetto et al. (2014), instrumental intensity measures may be less correlated with damage. However, given the absence of a single macro-seismic intensity scale for all locations worldwide, the use of instrumental measures (IM) is preferred.
[116] Following the work of Mota de Sá et al. (2010)

3.4. Spatial Considerations

Every real estate property is unique. The heterogeneity of real estate properties comes from their differences in construction, condition, financing and location (Balchin & Kieve, 1986; Robinson, 1979; Cupal, 2017). Therefore, the market value of a property is intrinsically linked to its location (e.g. proximity to schools, supermarkets, transport and other amenities), and hence it is expected that real estate values show signs of spatial dependency (Bin et al., 2008).

According to Toblers's (1970) "first law of geography": "everything is related to everything else, but near things are more related than distant things". This is the reason why similar properties in nearby locations usually show similar market values (Dubin, 1988; Li & Brown, 1980; LeSage & Pace, 2008), leading to an almost continuous grid of price variance across the city/region. This is a problem because regression models commonly applied to cross-section or panel data assume that observations are independent of one another, neglecting the unobserved dependency[117] (Sarafidis & Wansbeek, 2010; Gow et al., 2010). Hence, in order to have a correct evaluation of real estate values, one should take into account the importance of spatial dependency in the model's results (e.g. spatial autocorrelation of values), especially in cross-section regression models (Cliff & Ord, 1973; Judge et al., 1985; Wyatt, 1996; Anselin, 1998). In the case that spatial autocorrelation is present, estimations using Ordinary Least Squares (OLS)[118] can lead to biased results and, even if unbiased, can lead to inconsistent and inefficient coefficients, with strict limitations to its application and validity of values inferred through the model (LeSage & Pace, 2008; Kunst, 2010). This autocorrelation of values can be detected using a standard Moran-I test, Local Indicators of Spatial Association (LISA)[119],or through an analysis of residuals. One should keep in mind that heteroscedasticity can arise from the spatial variance (and autocorrelation) of values which is not explained by the model (Cuthbertson et al., 1992; Anselin, 1998). Hence, a spatial dependency analysis of values prior to the regression will be carried out, as well as an analysis of residuals afterwards, to ensure the validity of results. Spatial econometrics will also be used by resorting to a fixed-effects regression model. This approach allows to minimize the effects of spatial autocorrelation in the results.

Some authors have defined "homogeneous"[120] real estate market areas as a way to overcome this issue (see, for instance, the housing data from the Osservatorio del Mercato Immobiliare (OMI) used by Modica et al. (2021)). However, this approach has a high level of subjectivity and can lead to biases due to the incorrect location of "frontiers" (Gallimore et al., 1996; González and Formoso, 2000). The aggregation of real estate values also raises a second issue, the Modifiable Areal Unit Problem (MAUP). This occurs when the spatial zoning used to collect/gather

[117] The unobserved dependency of other independent variable(s) is called unobserved heterogeneity and the correlation between the independent variable(s) and the error term (i.e. the unobserved independent variables) is called endogeneity and it is a Endogeneity is a major methodological concern for many areas of business and management research that rely on regression analysis to draw causal inference (Abdallah et al., 2015)
[118] Due to the violation of the independence assumption between the error term and the explanatory variables, a basic assumption behind ordinary least squares (OLS) regression analysis
[119] Refers to Local Moran tests (Anselin, 1995)
[120] With similar properties, demographics and locations;

geographic data is “modifiable” or arbitrary. Thus, the results from this analysis may reflect differences in the partitioning of the data rather than differences in actual behaviour, given that economic phenomena rarely comply with administrative or arbitrary boundaries. In order to minimize its effects, Openshaw and Taylor (1979; 1981) (as cited by Viegas et al. (2009)), recommend that analysts follow three steps:

1. Start from the smallest division available, or the smallest they can process.
2. Aggregate these divisions in a fashion relevant to their investigation.
3. Assess the repetitiveness of their results for several aggregations.

Thus, to minimize the possible effects of MAUP, the analysis will be developed on three different levels:

1. Raw data with no aggregation and no information on location.
2. Aggregation of data using the old parishes of Lisbon (prior to 2013).
3. Aggregation of data using the current parishes of Lisbon.

The author started by running a PooledOLS regression model, which pools sample data across the years and then uses OLS to estimate the regression coefficients. This pooling of data is not as innocuous as it seems at first sight. Indeed, it is based on the assumptions that (1) there are no omitted variables and (2) there is no dynamic and simultaneous endogeneity. Both assumptions seem highly unrealistic because: (1) it is reasonable to assume that not all explanatory variables are included in the model, given that the real estate property is considered to be a complex, composed, and heterogeneous good depending on various features (both endogenous and exogenous) and (2) it is reasonable to assume the existence of spatially-correlated real estate values, leading to the spillover effects often observed in the market (Januário et al., 2021).

An Hausman test was carried out to see whether Fixed-Effects (FE) or Random-Effects (RE) were more appropriate and concluded a FE-model (with entity fixed effects) should be chosen. This fixed-effects or within-group estimation is used to minimize the effects of spatial dependency. However, the coefficients can still show an upward bias, as the number of periods T under analysis is small (10 points per parish, one per year) (Abdallah et al., 2015). The model looks as follows (EQ 6):

$$y_{i,t} = \beta \, X_{i,t} + u_i + \varepsilon_{i,t} \qquad \text{[EQ 6]}$$

where yi is the dependent variable (normalized property price), β is a vector of regression coefficients, $X_{i,t}$ is a matrix containing the yearly values for each variable for each location, u_i is an entity-fixed dummy variable and $\varepsilon_{i,t}$ is the error term.

In the second phase, a Multiple Linear Regression model of all data was chosen, using Generalized Least Squares (GLS[121]) to determine the regression coefficients, to minimize any heteroscedasticity in the residuals (EQ 7).

$$y_i = \alpha + \beta_{GLS} X_i + \varepsilon_i \qquad \text{[EQ 7]}$$

where yi is the dependent variable (normalized property price), α is a constant, β is a $nx1$ vector of GLS-estimated regression coefficients where n is the number of variables used, X is a $n \, x \, m$ matrix of explanatory variables (with m being the total number of observations) and ε_i is the error term.

However, a deep analysis of residuals is required, given that, in the case of dependence, that might be an indicative sign of endogeneity, and in that case, GLS did not improved the validity of the results (relatively to OLS). This is because GLS focuses on the efficiency issue resulting from OLS's failure of classic assumptions, but if OLS is biased, the results obtained by GLS will likely present the same validity failures. In that case, a Generalized Method of Moments (GMM[122]) regressor will be a more suitable solution to deal with endogeneity in the model (Rathnayake, 2019).

3.5. Economic Considerations

It is well documented in the literature the effect of macroeconomic factors in real estate markets (Reichert, 1990; Cho & Ma, 2006; Tse et al., 2014; Agnew & Lyons, 2017). Employment rates, income levels, productivity levels, tax rates (including municipal property taxes), and interest rates (such as Euribor) all impact the decision-making process of buyers and sellers of real estate property and, therefore, on housing pricing. By normalizing real estate prices on a yearly basis, one should be able to control for the effects of macro variables that affect all properties in a region taken into account.

Nonetheless, it is important to acknowledge some limitations to this procedure. Employment rates and income levels should not be uniform across the city, thus, its effects may not be completely controlled for after the normalization. One could expect that neighbourhoods with higher income-

[121] According to Taboga (2021) "*The generalized least squares (GLS) estimator of the coefficients of a linear regression is a generalization of the ordinary least squares (OLS) estimator. It is used to deal with situations in which the OLS estimator is not BLUE (best linear unbiased estimator) because one of the main assumptions of the Gauss-Markov theorem, namely that of homoskedasticity and absence of serial correlation, is violated. In such situations, provided that the other assumptions of the Gauss-Markov theorem are satisfied, the GLS estimator is BLUE.*". GLS allows for heteroskedasticity (the errors can have different variances) and correlations (the covariances between errors can be different from zero)".

[122] One should note that OLS is a specific case of GMM, which is a class of estimators, where the number of moment restrictions equals the number of unknown parameters (Dunn, 2014)

level households should not be as affected pricewise by changing economic conditions as their less wealthy counterparts. However, the lack of detailed information at neighbourhood level makes it nearly impossible to control for these variances. Additionally, one could expect that tax and interest rates will have different levels of impact across the different levels of housing. However, the impact of these factors on affordable housing should not be the same as on luxury housing, as different market segments attract different kinds of investor profiles with different financial capacities. One could even argue that the differences in financial literacy across profiles may also impact the decision-making process and, thus, how changes are perceived and acted upon. Hence, the inclusion of factors into the experimental design should always consider the availability of data and the author's hypothesis associated, while recognizing its limitations.

In this work, the author will follow this strategy, including, whenever possible, some macroeconomic factors that one may consider not adequately controlled by the yearly normalization of property prices.

3.6. Data

The data used in this thesis contains information regarding 8.726 properties in the city sold between 2008 and 2018. The dataset presents 61 variables, such as the normalized sales values, physical characteristics of the property (e.g. number of rooms, year of construction), and distances from points of interest such as schools and public transportation, among other information. Table 3-3 presents a brief description of some of the most important ones:

Table 3-3- Brief description of variables

Variable	Description	Type	Unit	Factors
DATE	Date of sale	Numerical	date	---
QUARTER	Quarter of sale	Numerical	quarter	---
SEMESTER	Semester of sale	Numerical	semester	---
YEAR	year of sale	Numerical	year	---
LOCATION	Address	Categorical	---	Ground Motion, Spatial
PARISH	parish	Categorical	---	Spatial
PARISH_Cod	parish code	Numerical	---	Spatial
ZONE	city zone	Categorical	---	Spatial
ZONE_Code	city zone code	Numerical	---	Spatial
CONSTRUCTI	Construction year	Numerical	year	Building Stock
BUILDING_F	Number of floors (building)	Numerical	floors	Building Stock
TYPOLOGY	Number of Rooms (Portuguese Typology)	Categorical	---	Building Stock
TYPOLOGY_C	Number of Rooms	Numerical	rooms	Building Stock
FLOOR	Apartment floor number	Numerical	---	Building Stock
AREA	dwelling area	Numerical	m2	Building Stock
BUILDING_C	building condition (good, average, bad)	Categorical	---	Building Stock
BUILDING_2	building condition (3,2,1)	Numerical	---	Building Stock
APARTMENT_	dwelling condition (good, average, bad)	Categorical	---	Building Stock

Variable	Description	Type	Unit	Factors
APARTMENT1	dwelling condition (3,2,1)	Numerical	---	Building Stock
ABANDONED	Abandoned (Yes/No)	Categorical	---	Building Stock
ABANDONED_	Abandoned (2/1)	Numerical	---	Building Stock
BUYER	(National / Foreigner and Single/Company)	Categorical	---	Economic
BUYER_Code	(0 to 5)	Numerical	---	Economic
NORMALIZED	Normalized value (year average €/m2)	Numerical	---	Economic
NORMALIZ_1	Normalized value (semestre average €/m2)	Numerical	---	Economic
NORMALIZ_2	Normalized value (month average €/m2)	Numerical	---	Economic
PURCHACE_U	Normalized value (2008 average €/m2)	Numerical	---	Economic
PURCHACE_1	Normalized value (2008 semestre average €/m2)	Numerical	---	Economic
PURCHACE_2	Normalized value (2008 month average €/m2)	Numerical	---	Economic
x	Longitude	Numerical	degrees	Ground motion, Spatial
y	Latitude	Numerical	degrees	Ground motion, Spatial
POI_Metro	Distance to nearest metro station	Numerical	m	Spatial
POI_Mercad	Distance to nearest supermarket	Numerical	m	Spatial
POI_Saude	Distance to nearest health facility	Numerical	m	Spatial
POI_Parque	Distance to nearest green space	Numerical	m	Spatial
POI_Estudo	Distance to nearest library	Numerical	m	Spatial
POI_Emerg	Distance to hospital emergencies	Numerical	m	Spatial
POI_CBD	Distance to Central Business District	Numerical	m	Spatial
POI_Cemit	Distance to nearest cemetery	Numerical	m	Spatial
POI_AcRodo	Distance to nearest access to highway	Numerical	m	Spatial
POI_Univer	Distance to nearest university	Numerical	m	Spatial
POI_Indust	Distance to nearest industrial area	Numerical	m	Spatial
POI_Rio	Distance to Tagus river	Numerical	m	Spatial
POI_Escola	Distance to nearest primary school	Numerical	m	Spatial
POI_ESPriv	Distance to nearest private high school	Numerical	m	Spatial
ESPubl	Distance to nearest public high school	Numerical	m	Spatial
AUGI	Does it belong to an Urban Area of Ilegal Origin	Numerical	---	Building Stock
GEBALIS	Does it belong to social housing	Numerical	---	Building Stock
RUIDO	Level of noise (0 to 5)	Numerical	---	Building Stock
VULInund	Vulnerability to flooding (0 to 3)	Numerical	---	Spatial
VULSismo	Ground motion vulnerability (0 to 4)	Numerical	---	Ground Motion, Spatial
VULMass	Landlslide vulnerability (0 to 3)	Numerical	---	Spatial
Park50M	Number of parking spots wiithin 50 m	Numerical	spots	Spatial
Culture500	Number of cultural points of interest within 500 m	Numerical	points	Spatial

It is noteworthy that Luz (2020) has added the last 23 variables mentioned in Table 3-3 during his work of analyzing the spatial determinants of property prices in Lisbon. This research made use of those variables.

The development of this thesis also required some information to be added as new variables to the database, such as the number of floors (corrected number of floors), the soil type, and peak ground acceleration at the location, amongst others. The following sections will describe the methodology used to do so.

3.7. Attributing Vulnerability

Our dataset analysis started with an Exploratory Data Analysis (EDA), presented in detail in the results chapter. After the initial assessment, a series of seven steps was followed to attribute vulnerability values for each property in the database. In this section, the author presents the description of each of these steps.

3.7.1. Estimating Number of Floors

The number of floors of a building has been known as one of the variables of interest when assessing its vulnerability to seismic risk (Mota de Sá et al., 2010; Simões et al., 2014). Given that the database had a high percentage of missing values for the number of floors in each building, SIG3D model of the city of Lisbon was used, containing the height of the terrain and buildings, which resulted from the work of IST colleagues Arch. Rita Machete and Prof. Ana Paula Falcão. As previously mentioned, the dataset contained the latitude and longitude of every property in the dataset, hence, both datasets were spatially joined to infer the number of floors. This technique was implemented using the *GeoPandas* library available in Python v2.7. It was also possible to cross-check the height records of buildings in the dataset to ensure consistency across all data.

In order to estimate the number of floors, an approximate floor-to-floor height of three meters has been used. In the case of mismatching number, between the existing records and the estimated height a case-by-case analysis was put in place, in order to verify the approximate number of floors. This information was then used as input to the modifying factors used in Ferreira (2012) method.

3.7.2. Estimating Soil Type and Peak Ground Acceleration

The soil type in which a building is located is also a variable of interest when assessing its required vulnerability. Some authors have even studied the direct relationship between housing prices and the type of soil where they are located (see for reference Önder et al. (2004)). Hence, a similar technique to the one used in estimating the number of floors was used, but to spatially join the Lisbon City Soil Classification Map, which resulted from the work of Oliveira et al. (2019), with the property database. This soil classification map was created from a varied set of information, including NSPT values from 8792 boreholes and the thickness of cover formations (see Figure 3-1).

This allowed the attribution of a soil type to each building according to Eurocode 8 (EC8), which classifies it into seven categories: A, B, C, D, E, S1 and S2. Given that the work of Oliveira et al.

(2019) also considers soil intermediate soil types AB and BC, a conservative approach was followed by attributing the worst-case scenario for each of these intermediate types, which means that an AB soil type would be defined in the database as B, while a BC would be defined as C (Figure 3-2).

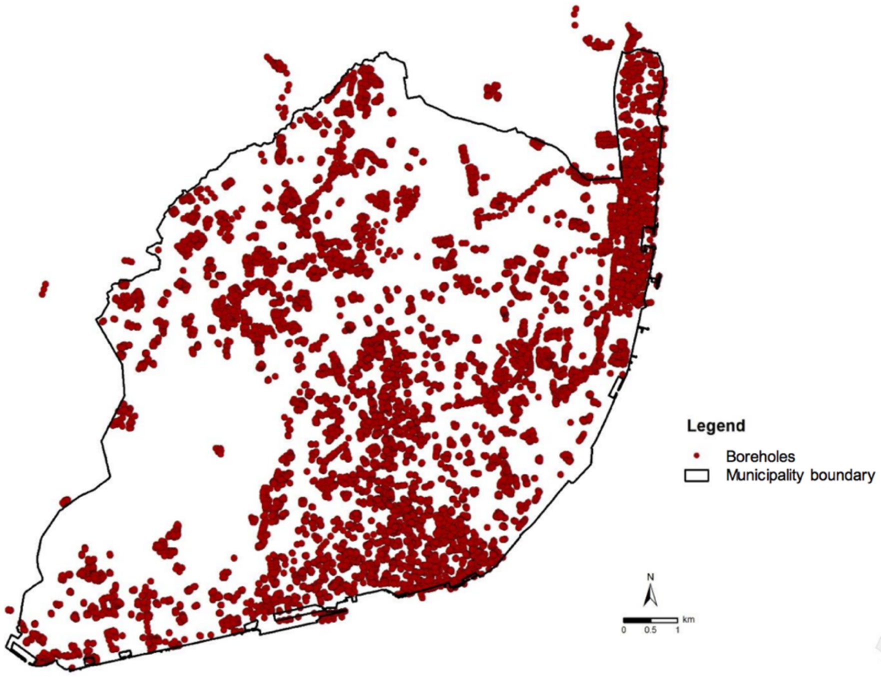


Figure 3-1 – Boreholes for the City Soil Classification Map. Source: Oliveira et al (2019)

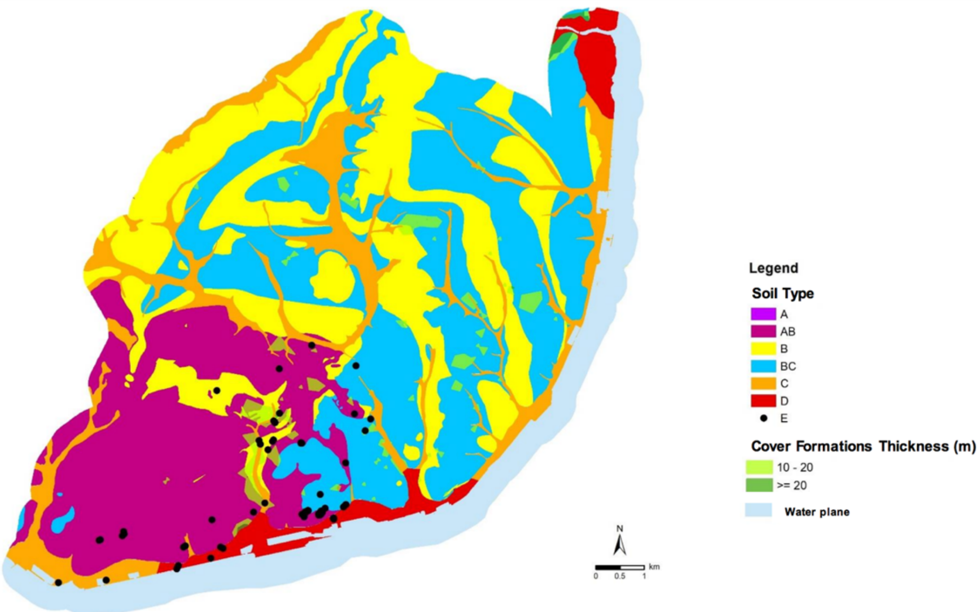


Figure 3-2 - Lisbon City Soil Classification Map. Source: (Oliveira et al 2019)

For each category, a value for peak ground acceleration (PGA) was attributed according to the class its spectral acceleration (close earthquake[123] - type II), and the class of importance of the residential buildings (see Portuguese Norm NP EN 1998 -1[124]). This was an important step as PGA measures the peak acceleration reached during an earthquake. The greater the acceleration, the greater the structural damage (McGinnis, 2004). Having the PGA for each location allowed the author to calculate the intensity I_0, as explained in Mota de Sá et al. (2010). The intensity is calculated as (EQ 8):

$$I_0 = \frac{log\, PGA - 0{,}014}{0{,}3} \qquad \text{[EQ 8]}$$

With the intensity, it is possible to estimate the required vulnerability of a given building in order to avoid structural collapse according to EC8. This required vulnerability (V_{req}) can be calculated as (EQ 9):

$$V_{ureq} = 1{,}82576 - 0{,}178444 I_0 \qquad \text{[EQ 9]}$$

The vulnerability will be used as the denominator in the Building Resistance Deficit ratio (BRD), as explained in section 4.7.5. The large number of boreholes has provided an extensive description of soil types across the city of Lisbon. However, we should note that boreholes are discrete sample points, and given we have fewer boreholes in the northern and western regions of Lisbon, we may assume there might be some uncertainty in the attribution of soil classification, as a continuous distribution of soil is being estimated from a collection of discrete points. Nevertheless, the author has assumed that most attribution should be correct and that results should not be affected by neglecting the uncertainty in the borders between soil types.

3.7.3. Attributing Required Vulnerability Based on Construction Year

The first approach to determining the construction type of a given building followed a simple correspondence between its construction year and a vulnerability class, as Ferreira (2012) suggested. Each vulnerability class represents a group of buildings with expected similar behaviour under seismic action (Mota de Sá et al., 2013). This constitutes a simpler but expeditious approach to the attribution of current vulnerability. It follows a system of five vulnerability classes (Table 3-4):

[123] Translated from "*sismo próximo*".
[124] Available at: http://www2.dec.fct.unl.pt/seccoes/S_Estruturas/Dinamica/mine/EC8_1_Portugues.pdf

Table 3-4 - Average vulnerability for each typology. Source: Ferreira (2012)

Typology	Vui
Masonry (<1945)	0,70
RC (1946-1960)	0,60
RC (1961-1985)	0,50
RC (>1985 and ≤ 5 floors)	0,40
RC (>1985 and > 5 floors)	0,44

Using this approach, the construction typology was estimated and a value for the current vulnerability attributed based only on the available information (construction year and estimated number of floors). This approach was chosen, rather than the one considering only the construction year (see, for instance, Table A.0-8 in Annex 1), as it provided an extra class and thus more detail to the classification system.

This is an adaptation from the 1998 European Macroseismic Scale (EMS98) (Grünthal, 1998), which provides a correspondence between the type of structure and its vulnerability (rated A to F)[125] (Figure 3-3). Herein, *vulnerability* is understood as the capacity of the residential units to suffer a given level of damage according to the intensity of the shock and defines building damage on a scale of 1 to 5 (Figure 3-4) [126].

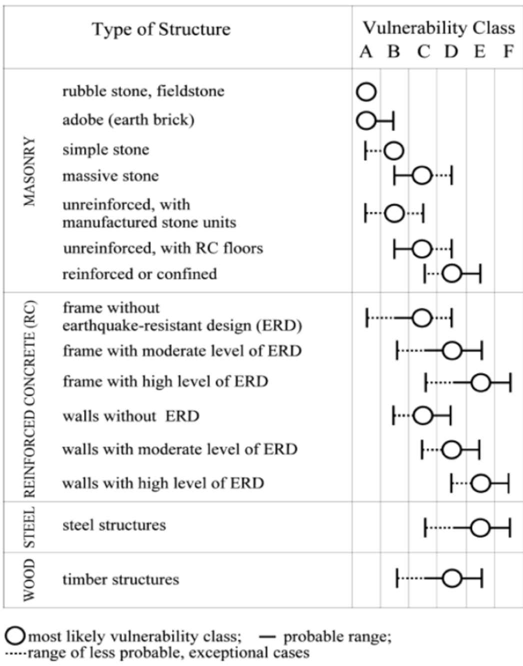
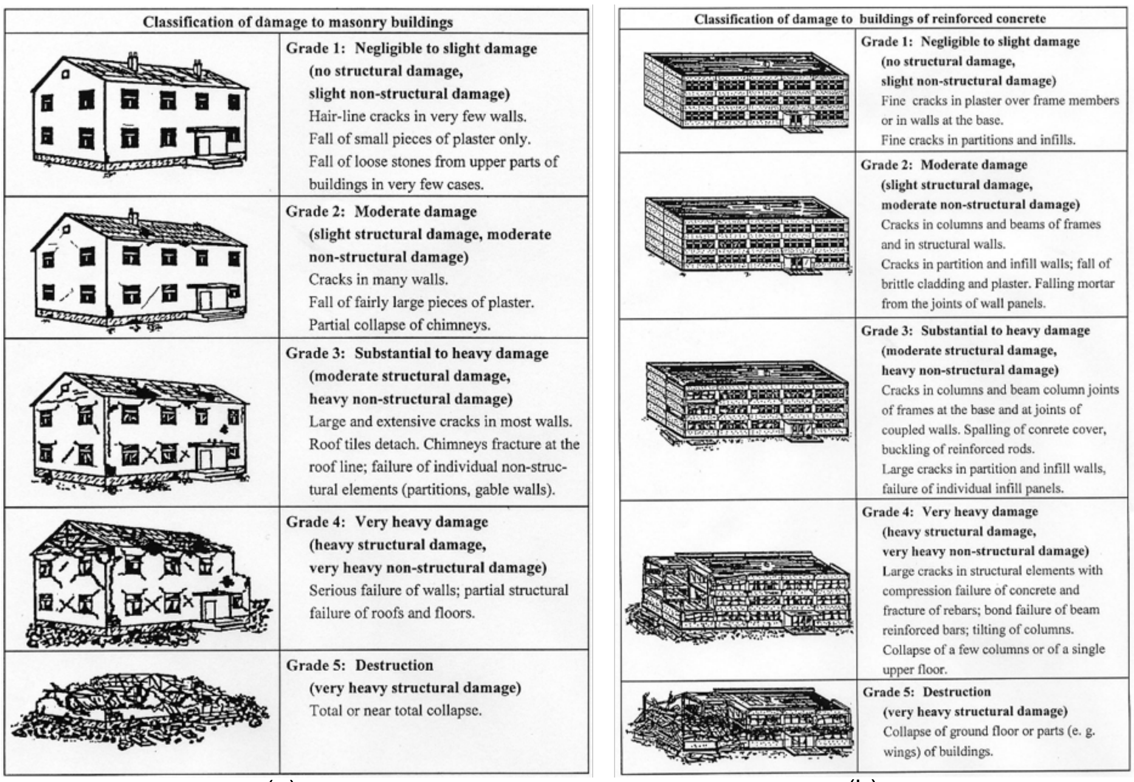


Figure 3-3 - Vulnerability classes of buildings according to the EMS-98;

[125] Some authors have even used the A to F scale of vulnerability as an input for models assessing the effect of earthquake risk in the housing market. See for reference (Modica et al 2021)

[126] This scale was used as part of the bibliography to define vulnerability values on Table A.0-5.

As previously described, to the obtained average vulnerabilities based on construction year, modifying factors (which increase or decrease vulnerability) based on building characteristics were added (see Table 3-2).



(a) (b)

Figure 3-4 - Damage level for masonry (a) and reinforced concrete (b) buildings (EMS-98). Source: Copernicus.eu

3.7.4. Attributing Required Vulnerability Based on Construction Year and Location

In order to have a more refined description of the current vulnerability ($v_{uactual}$), the estimates were based not only on its construction year but also on a building's location. Hence, the author started by defining a set of vulnerability classes with similar seismic behaviour.

The first step was gathering information from existing literature on how to constitute these classes. The LNEC (National Laboratory for Civil Engineering)[127], divides building typologies in the Portuguese territory in 7 distinct classes:

1. Masonry building built before 1755.
2. Post-earthquake masonry buildings (e.g. *pombalino* buildings), built between 1755 and 1880.
3. Gaioleiro masonry buildings, built between 1880 and 1930.
4. Masonry and RC buildings (mixed-structure) built between 1930 and 1940.

[127] Available at: http://www-ext.lnec.pt/LNEC/DE/NESDE/divulgacao/evol_tipol.html

5. Masonry and RC buildings (mixed-structure) built between 1940 and 1960.
6. Modern RC buildings built after 1960.

According to Lopes (2008), Lisbon's housing stock can be classified into seven groups, as previously described in detail in chapter 3.5:

1. Old mansory buildings (prior to 1755 earthquake).
2. "Pombalino" buildings (built during the 1755 reconstruction).
3. "Gaioleiro" buildings (built during the last third of the XIX century, corresponding to a phase of large city expansion).
4. "Placa" buildings or "transition buildings" (wood and reinforced concrete mixed design).
5. First period of reinforced concrete buildings (before 1961).
6. Second period of reinforced concrete buildings (before 1985).
7. Modern reinforced and prestressed concrete buildings.

Carvalho et al. (2002), used the Building Questionnaire of Census 2001 to evaluate the vulnerability of residential buildings in Portugal, and divided them into seven primary classes (Table 3-5):

Table 3-5 - Classes of vulnerability variables (Building survey 2001). Source: Carvalho et al (2002)

Typological Class	Description	Seismic Design Level	
Masonry before 1919	Unreinforced stone and brick masonry	None	Minimum strenght and ductility
Masonry 1919-1960	Unreinforced stone and brick masonry	Low	Low strenght and ductility
Masonry 1960-1985	Unreinforced stone and brick masonry	Low	Low to medium yield strenght; Low ductility
Masonry 1960-2001	Masonry with reinforced concrete confining elements	Low	Low to medium strenght; greater over-strenght after yielding; Low ductility
RC before 1960	Reinforced concrete framed structure	Low	Low to medium strenght; over-strenght; Low ductility
RC 1961-1985	Reinforced concrete framed structure	Medium	Medium strenght and ductility
RC 1986-2001	Reinforced concrete framed structure	Medium	Medium strenght and ductility, higher to previous class

Each class was then divided into seven subclasses according to the number of floors (1,2,3,4,5 to 7, 8 to 15, +15), leading to a total of 49 classes. The same classification was followed by Sousa et al. (2004).

In this thesis, the Census 2011 Building Questionnaire was used, which classifies each building in the city of Lisbon according to its construction year (a total of 10 classes), its type of structure (a total of 6 classes) and some other characteristics such as the type of roofing and cladding. Using the construction year and the type of structure (based on its main resistant elements), a total of 60 different initial categories were found, the total number of possible pairs, each resulting from a combination of the two variables. However, given the similarities between some classes, the information from the literature was used to join the initial groups into 24 different classes. Then, a construction typology (according to EMS-98) was assigned, along with a value for the current vulnerability of each class. Finally, the modifying factors (described in Table 3-2) were added based on the number of floors and condition. Figure 3-5 synthesizes the reasoning followed for this process.

The base values of vulnerability have been based primarily on the works of Lagomarsino and Giovinazzi (2006), Oliveira et al. (2004) and Ferreira (2012). The full table of correspondence and groupings can be found in Annex A (Table A.0-5).

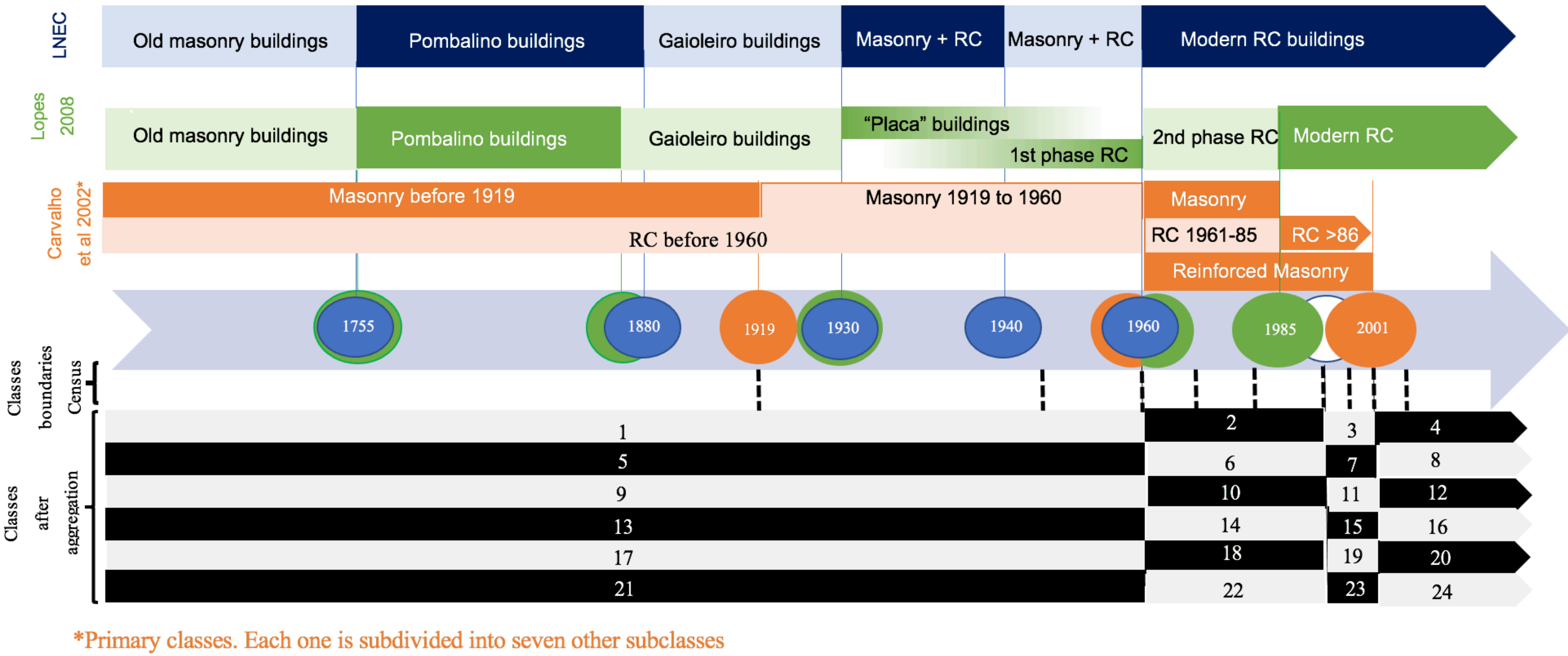


Figure 3-5 - Classes used to model vulnerability.

3.7.5. Attributing Building Resistance Deficit (BRD)

In order to assess the importance of vulnerability in the pricing of Lisbon's properties, the work of Mota de Sá et al. (2013) was followed to define the Building Resistance Deficit index (BRD)[128]. The BRD indicates, for a given level of seismic action, how far is the current resistance of a given building from the required resistance or, in this case, required vulnerability to avoid structural collapse according to EMS-98.

Given the required vulnerability (Vreq) and current vulnerability (Vui) of a building, its BRD is calculated as follows (EQ 10):

$$BRD = \frac{\Delta V}{V_{req}} \quad \text{[EQ 10]}$$

where ΔV is the difference between the current and required vulnerabilities. The BRD value was calculated and added to the database. Despite the practical advantages of using an index value to describe vulnerability, there are limitations associated with it. For instance, it is known that ground motion duration is especially important for degrading structures, such as unreinforced masonry (Bommer et al., 2004; Rossetto et al., 2014), which is not captured by the base value of required vulnerability. The heterogeneity in the construction of older buildings can also pose limitations to using a single index. However, one may argue that the modifying factor concerning the condition of a building, used to calculate its current vulnerability, already accounts for the increased effects of ground motion on a degrading structure.

3.8. Additional Exploratory Analysis

This chapter will describe some additional analyses carried out to deepen the understanding of the data and its variability. A Principal Component Analysis (PCA) and a Feature Importance analysis were conducted to lay down the basis on which the variables should be chosen for the model.

3.8.1. Principles of Principal Component Analysis

PCA is a multivariate technique to extract meaningful information from a dataset by creating a new set of orthogonal vectors[129] from the original inter-correlated data, synthesizing information, by reducing variability within each principal component (Abdi & Williams, 2010). Thus, the PCA was used as part of the exploratory analysis to understand which variables have the most significant contribution to the variability of data in the sample. One should note that this technique will not allow for a direct selection

[128] Translated from Portuguese *Défice de Resistência das Construções*
[129] PCA computes new variables called principal components which are obtained as linear combinations of the original variables. Components have a decreasing level of importance, meaning, the first principal component is the most important one, synthesizing the large portion of variability. The second component will be the one synthesizing the second-largest portion and the same logic applies to the remaining components.

of variables but to provide a deeper understanding of the data's variability. The results will be detailed in section 4.1.2.

3.8.2. Principles of Feature Importance

The second additional analysis conducted as part of the exploratory analysis comes as a subproduct of implementing an extreme gradient boosting tree algorithm or XGBoost[130], which is an ensemble method to combine the outputs of individual decision trees[131] (see for reference Friedman (2001)). Unlike Random Forests, the boosting technique combines decision trees (weak learners with only one split, or decision stump) in series, with each tree set to minimize the errors of its previous. If we take the amount that each attribute (or variable) split point improves the performance of a tree, weighted by the number of observations the node is responsible for, we get a measure of each feature's importance. This enables an understanding of which features may be the most important to be used as inputs in a model and, in this case, which attributes contribute the most to property prices. Section 4.9.3.3 will describe in detail the methodology behind XGBoost algorithms and the parameters used.

This analysis was conducted for the whole city of Lisbon, along with an in-depth analysis for each of its parishes, shedding light on the variability of attributes valued by homebuyers. The results will be detailed in section 4.1.3.

3.9. Revealed Preferences

This chapter will discuss the methodological approach taken to study the revealed preferences of real estate stakeholders regarding property features and their vulnerability to seismic action. This was done by analyzing the data from the property sales dataset. Section 3.9.1 will lay down the basic assumptions made.

3.9.1. General Considerations

Revealed Preferences (RP) methods analyze the decisions made by consumers by looking at their purchase behaviour. This presents an advantage to stated-preferences methods, based on survey responses, because the latter could be subject to behavioural biases (Rogers et al., 2019). Hedonic price models have been one of the most widely used methods to describe housing prices as a function of housing/structural characteristics such as location, typology and condition. Initially developed by Court ((Court, 1939) as cited in (Des Rosiers & Thériault, 2006)) as a framework to estimate prices in the automotive industry, it became a staple for estimating prices in heterogeneous goods. This framework was later applied to real estate valuations by Rosen (1974), which perfectly fits the definition of "market value" as the "most probable price" a property should be sold for in a transparent and

[130] For further information on the Python implementation of XGBoost, please refer to https://xgboost.readthedocs.io/en/stable/tutorials/model.html
[131] Decision trees (DT) are supervised learning models used in machine learning for regression and classification. The algorithm of DTs is optimized to split data into leaves and brunches, selecting the attributes that provide the greatest information gain by reducing the system's (data) entropy. For reference, see Rokach and Maimon (2005)

competitive market setting. Hedonic price models have evolved to consider several factors, such as the influence of spatial relationships between real estate properties.

We should note that location is one of the most critical factors contributing to pricing and that the pricing of any given property is not independent of the surrounding ones. Sampling biases should also be taken into account when modelling data. This derives from the fact that the sample collected may not be representative of an entire population. For instance, if we are modelling the housing prices for the parish of Avenidas Novas, its buildings' age and overall features should be similar to those taken in the sample. For example, if we took a sample where most properties were built after 2001, the results would most likely be biased, given that the average building in this parish is relatively older. Therefore, to properly understand the possible biases in the results, one should compare the sample with some survey results containing information on the entire building stock. In the following sections, the methods applied to reduce any possible biases in the results will be described as well as the regression methods applied in the hedonic price models.

3.9.2. Sampling Bias

Sampling biases can be introduced when a database contains only a subset of properties non-representative of the total housing stock. The smaller the sample size relative to the whole population number, the more likely it is to see its effects. Additionally, the farther the average characteristics of the properties in the sample are from the actual property stock, the larger the biases introduced in the results. In order to control for the effects of sampling bias, the author started by comparing the property sample with the Building Survey from Census 2011. A comparison at three levels will be carried out: 1) at the city level; 2) at the new-parish level; and 3) at the old-parish level. However, one must recognize that, given the nature of the data, there will be few options to deal with a possible mismatch between the sample and the population: 1) take a stratified sampling approach or 2) conduct a higher level of aggregation. Also, a low number of properties in any given unit of analysis can lead to a biased sample; however, in this case, a stratified sampling approach is not a viable option, as any piece of information should be discarded. Nevertheless, it is important to acknowledge any sampling biases and their possible effects. The results of this comparison are shown in section 4.3.

3.9.3. RP Regression

This chapter will describe the regression techniques and algorithms used to assess the impact of several property characteristics in its value.

3.9.3.1. Multiple Linear Regression (MLR)

Multiple linear regression (MLR) techniques have long been used to build hedonic price models (Elsinga and Hoekstra, 2005). Simply put, MLR models allow for the estimation of a target variable dependent on the values of a set of explanatory or independent variables. The influence of each variable on the target is given by its regression coefficients[132], which can be calculated using a variety of estimation techniques, such as Ordinary Least Squares (OLS), which minimize the sum of the residuals (González and Formoso 2000). This is one of the most often used, however, to obtain reliable results, a set of assumptions, called Gauss-Markov assumptions must be followed[133]:

1. **Linearity in parameters**: the parameters of the population model are linear with respect to the outcome (dependent variable); If so, we can write the model as (EQ 11):

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon \qquad \text{[EQ 11]}$$

2. **Random sampling**: the data must have been randomly sampled from the population;

3. **No perfect collinearity**: the regressors being calculated aren't perfectly correlated with each other, and none of the covariates is constant;

4. **Exogeneity** or **Zero Conditional Mean**: the error has an expected value of zero given any values of the independent variables;

5. **Homoscedasticity**: no matter what the values of the regressors might be, the error of the variance is constant.

If all the assumptions are met, then we can assume that the parameters estimated from OLS are BLUE (**B**est **L**inear **U**nbiased **E**stimators). However, as previously mentioned in the Spatial Considerations section (section 4.4), real estate valuations are closely linked to location of a property. Therefore, there usually is a spatial correlation of values, meaning that the error term is closely linked to the location variables and thus leading to non-constant error variance. As Des Rosiers and Thériault (2006), state that the "*hedonic approach raises major methodological issues in relation to the presence of structural heteroskedasticity and spatial autocorrelation in the model residuals. Both are detrimental to the stability of regression coefficients, even more so to the accuracy of their standard errors*". We can make an approximation by: 1) simplifying the model and excluding location variables, but, in that case, the model will be biased as it does not account for one of the main factors influencing the pricing of a property or 2) collecting data from smaller regional areas considered to be roughly homogeneous; under this assumption, we may consider that differences in location are neglectable and that most differences in price will be due to endogenous factors (e.g. the number of rooms, floor area, condition…)[134].

[132] Usually, models are built using a random sample from the population, as a technique for statistical inference. In regression analysis, regression coefficients are estimates of the actual population parameters.
[133] See for reference Perraillon (2019)
[134] For instance, when analyzing the perception of real estate hazard in the Tokyo Metropolitan Area, using seven-digit postcode basis as unit of analysis, Nakagawa pointed out that "Although our estimation procedure does not consider spatial correlation explicitly, given less geographical clustering, standard errors of the ordinary least squares (OLS) estimates may not be seriously biased" (Nakagawa 2007).

Given that the author intends to analyze property values across all Lisbon and considering its differences between parishes, location is an important factor to take into account when building the model. It is noteworthy that OLS is, in fact, a particular case of the Generalized Least Squares (GLS) technique, where the variance of errors is assumed to be constant. In situations where only the homoscedasticity assumption is violated, provided that all other assumptions are satisfied, GLS is BLUE (Taboga, 2021). Hence, in the first approach, given that the true covariance matrix *V* is unknown thus preventing the direct application of GLS, the author has used Feasible Generalized Least Squares (FGLS)[135] to estimate the regression coefficients.

FGLS has been taken as a first approach to a regression model, using a MLR technique, however, as previously mentioned, we acknowledge the limitations of a simple MLR to model real estate values, and the use of FGLS is not a perfect solution. It should be noted that:

1. FGLS is actually biased (especially for small samples)[136].

2. FGLS is consistent (asymptotically unbiased); So, fears go away with larger sample sizes.

3. As n tends to infinity, the variance from FGLS is less than from OLS;

Despite the considerably large sample size, it is not possible to guarantee that estimators presented in this regression are BLUE (as it would have been the case for true GLS estimations). Therefore, more sophisticated models are also used, such as a Fixed-Effects model[137], which allows estimation in the presence of autocorrelation within panels (entities or parishes in this case), cross-sectional correlation and heteroskedasticity across panels[138], and has already been used to model (see for reference Nakagawa (2007), Modica and Locati (2016) and Modica et al (2021)). This is a more appropriate setting for the use of panel data[139]. To use this approach, the properties were grouped according to the administrative boundaries: first using the old parishes (pre-2013) and then using the new, larger parishes. The final fixed-effects model was as follows (EQ 12):

$$\log(p_{it}) = \alpha + \delta X_{it} + \beta BRD_{it} + \mu_i + \phi_t + \varepsilon_{it} \qquad \text{[EQ 12]}$$

Where $\log(p_{it})$ is the mean housing sale price, for parish *i* at time *t;* X_{it} is a set of relevant explanatory variables (including number of rooms, floor area, construction year, distance to the nearest station/bus stop, construction type of dwelling,…) *BRD* is the Building Resistance Deficit (BRD); Parish fixed effects (μ_i) control for all time-invariant characteristics of each parish and time-fixed effects (ϕ_t) control for city-wide shocks in each period; ε_{it} is the error term for parish *i* at time *t.*

---

[135] FGLS is the practical application of GLS when the covariance matrix *V* is actually unknown, being replaced with an estimate, such that $\hat{\beta}_{GLS} = (X^T\hat{V}^{-1}X)^{-1}X^T\hat{V}^{-1}$y (Taboga, 2021).
[136] See for reference Miller and Startz (2018)
[137] Fixe-Effects models have also been found to be useful when assessing the perception of risk and its impacts on the real estate data. See for reference (Modica et al., 2021)
[138] See for reference https://www.stata.com/manuals13/xtxtgls.pdf
[139] Time-stamped data across multiple identifiable locations (entities).

The selection of variables to be used in the model was carried out as a combination of 4 different techniques:

1. PCA to identify the variables which most contributed to data variance (as explained in section 4.8.1)
2. Feature selection derived from XGBoost (as explained in section 4.8.2);
3. Best subset selection and
4. Stepwise regression (both forward stepwise and backward stepwise).

The first two techniques have already been described as part of the additional analysis (section 4.8) and were also featured as part of the EDA.

The Best Subset Selection technique consists of an algorithm that divides the dataset into subsets of variables, runs multiple regression models, and assesses their fitness according to the desired evaluation metric. Then, it selects the models with the best metric values, showing which variables may be the most suited to use as input variables. The problem with this technique is that it requires much computational power, especially in-memory processing (RAM), as it needs to store the results of every model it runs to compare and provide the best results. For example, using this dataset, a selection of 1 explanatory variable would lead to a total of 34 models being built and compared[140]. However, by adding a second variable, the number of models run was increased to 561 models. As a result, a subset of up to 3 variables was run, resulting in a total of 5984 models. The adjusted R^2 was the evaluation metric chosen. Section 4.1.4 will show this technique's results.

Lastly, both forward and backward stepwise regression models[141] were used. The forward stepwise regression starts with a model with no variables (e.g null model), and then starts adding significant variables one after another until a pre-specified criterion (stopping rule) is met or when all available variables are included in the model. In this case, variables providing the highest drop in the model RSS (Residuals Sum of Squares) compared to the other explanatory variables' options were added. The backward stepwise regression is similar to the former. However, it starts with a full model (model with all the available explanatory variables or predictors) and then removes the least significant variables in the model, one after another, until a pre-specified criterion is met. In this case, the author has used the same criteria (RSS), eliminating the variable causing the lowest increase in RSS. According to the literature, backward stepwise regression is preferable and should be applied when possible[142] (Mantel, 1970). The results will be detailed later in this work (section 4.1.5).

3.9.3.2. Multivariate Adaptive Regression Splines (MARS)

Multivariate Adaptive Regression Spline (MARS) is a nonparametric regression technique used to study nonlinear relationships between a target variable and a set of predictors (Kartal & Bozdogan 2015). Linear models present known limitations due to their assumptions of linearity, which can affect their

[140] For more information see https://www.statology.org/best-subset-selection/
[141] Forward and Backward Stepwise have been used in the literature to build hedonic models. See for reference (Elsinga and Hoekstra 2005)
[142] Whenever the number of variables is less than the number of observations.

predictive accuracy. Therefore, the implementation of MARS regression is a middle-ground between the easiness of interpreting the results of linear regression models (often with lower accuracy) and the "black-box", but typically more accurate results from artificial neural networks (ANN). The MARS algorithm was first proposed by Friedman (1991), and it captures nonlinear relationships without the need to explicitly include polynomial terms, as in a polynomial regression model. It captures these nonlinear relationships by assessing cutpoints (knots) in the data. For instance, Figure 3-6 shows an example function such that (EQ 13):

$$Y = f(X) \qquad \text{[EQ 13]}$$

The algorithm starts by identifying the first knot, with two regression lines on each side, intending to minimize RSS. This is defined by a hinge function, which divides the dataset into two parts subject to two different regression lines, in the form of $h(x-a)$ where a is the cutpoint. This procedure is followed to determine the second and following cutpoints, always to minimize RSS.

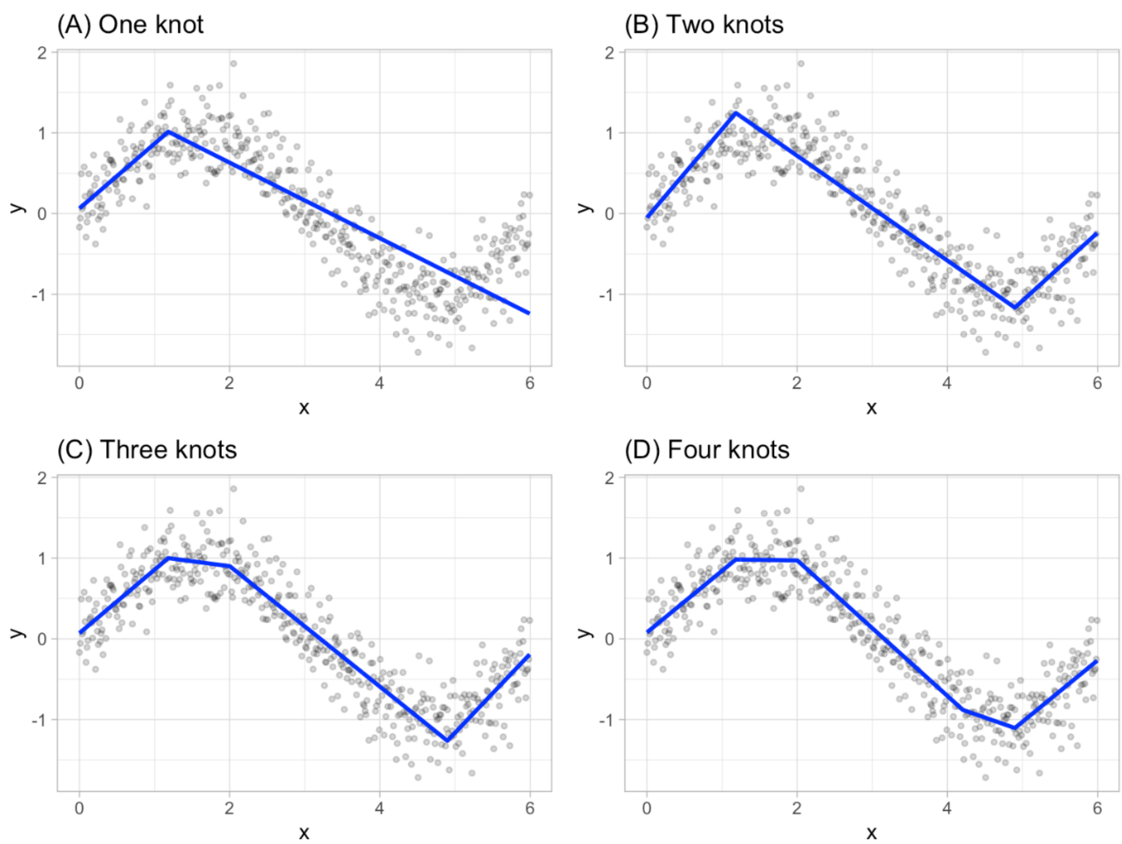


Figure 3-6 - MARS regression procedure. Source: Bradleyboehmke

This procedure can be followed to achieve as many cutpoints as deemed necessary[143]. However, it may produce a highly nonlinear prediction equation, with the risk of *overfitting* and lowering its accuracy in predicting unseen data. Once the complete set of knots is found, knots may be removed in order to

[143] Only limited by the number of data points

generalize the function and avoid overfitting, in a process known as "pruning" (see, for reference, Deichmann et al. (2002)).

This procedure was applied to the dataset, using the full range of explanatory variables. The results can be found in section 4.2.4.

3.9.3.3. Extreme Gradient Boosting (XGBoost)

Lastly, XGBoost was used to predict property values based on a set of explanatory variables. This section will further explain the methodology behind this machine learning technique and describe the parameters used in the regression.

XGBoost is an ensemble learning technique used in supervised learning to predict the values of a target variable y_i from a set of predictors x_i . In essence, XGBoost consists of a set of decision trees[144], ensembled to enhance its predictive power. The model is trained by defining an objective function, containing a loss function $L(\theta)$ and a regularization term (EQ 14):

$$obj(\theta) = L(\theta) + \Omega(\theta) \qquad \text{[EQ 14]}$$

The training loss function measures how far the predictions are from the training set's actual values. In this model, the Root Mean Squared Error (RMSE) was used, which is given by equation 15:

$$RMSE = \sqrt{\frac{\sum_i^n (\hat{y}_i - y_i)^2}{n}} \qquad \text{[EQ 15]}$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value in the training set and n is the total number of observations. This loss function was chosen as it is one of the most commonly used in machine learning[145]. The regularization term is used to modify the learning algorithm to favour "simpler" prediction rules and avoid *overfitting*, which may comprise prediction accuracy with new data. It is important to note that no regularization has been used when running the models to determine feature importance. This is due to the fact that the main goal was to find the most important predictors for the dataset and not to maximize accuracy in predictions using new data (outside the training set). However, in this phase, a L2 regularization term[146] was used, which penalizes weights (w) in the model that are large. It can be defined as (EQ 16):

$$d(w, 0) = \sqrt{\sum_i^k (w_i)^2} = \|w\| \qquad \text{[EQ 16]}$$

[144] This happens because usually a single decision tree is not strong enough to be used in practice. It is also noteworthy that decision trees are optimize for leaf values and not for individual explanatory variables. For more information on decision trees please refer to Rokach and Maimon (2005)

[145] See for reference https://xgboost.readthedocs.io/en/stable/tutorials/model.html

[146] See for reference Paul (2018)

where the norm of w is the vectors length. By minimizing the length of the vector, we are penalizing larger weights in the model and avoiding overfitting. This penalization is usually fine-tuned in the loss function by using a hyperparameter λ, which adjusts the tradeoff between having a low training loss and having low weights, taking the form $L(w) + \lambda\|w\|^2$. Note that the square is only used as a mathematical artefact to eliminate the root making it easier to work with.

The XGBoost model results from the additive training of a set of trees. Therefore, one can take any added tree as a step in the prediction. The prediction of any given tree as $f_i(x_i)$, and the predicted value at each step can be defined as (EQ 17):

$$\begin{aligned}
\hat{y}_i^{(0)} &= 0 \\
\hat{y}_i^{(1)} &= f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\
\hat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\
&\dots \\
\hat{y}_i^{(t)} &= \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i)
\end{aligned} \qquad \text{[EQ 17]}$$

with t being the total number of trees in the model. Note that, at each step, it is added the tree that optimizes the objective function. In this model, an ensemble of 100 trees was used, with a boosting learning rate[147] of 0.08, and a training sample of 75% of the total data set (75/25 split between training and test set). The final results will be presented in section 5.2.5.

3.10. Stated Preferences

Stated Preferences (SP) methods, also called Contingent Valuation (CV) methods, rely on survey-based data to establish valuations (Kroes & Sheldon, 1988; Mitchel & Carson, 1989; Carson et al., 1996; Carson, 2000; Venkatachalam, 2004; Rahmatian, 2005; Carson & Louviere, 2011). They consist of various techniques to perform valuations based on individual respondents' statements, often to hypothetical scenarios presented, to model the expected utility of amenities. They are also well-suited to perform statistical modelling of subjective concepts such as safety or perception of risk and its impact on valuation (see, for reference, Asgary and Willis (1997)). This section will be describing the methodology used to assess the knowledge and preferences stated by Lisbon's real estate stakeholders, allowing for a different perspective on the impact of real estate risk (and perception) in the market, a subjective risk perception (Naoi et al., 2009). The presentation will start by laying down some general considerations on these methods.

[147] The learning rate refers to how fast the model learns. Each tree added modifies the overall model. The impact (or magnitude) of this modification is given by the learning rate. A lower learning rate means the model is learning more slowly; however, it will make it more robust and efficient, with a better expected performance. However, if the model is learning slower, it will take more time to learn and will need a higher number of trees in order to achieve the best results. It is noteworthy that a high number of trees may also increase the risk of overfitting.

3.10.1. General Considerations

The high flexibility of SP methods allows for an in-depth analysis of real estate stakeholders' decision-making process and risk perception. They allow conducting monetary estimates based on an asset's market and non-market characteristics (e.g. flood risk) and having a first-hand insight into the rationale of people's decisions. It is noteworthy that stated preferences are usually used to determine willingness-to-pay (WTP) for a particular feature or commodity (Rogers et al. 2019). Despite the advantages, Stated Preferences methods are prone to errors primarily from response biases. This can arise from unclear questions, wrong wording choices, types of questions (open vs closed), the delivery method (in-person or telephone interviews, e-mail, online survey) or even respondents' personal experiences and demographics. These methods are also subjected to sampling biases, as previously explained in section 3.9.2. Nevertheless, despite the bias-induced weaknesses of the method, CV methods have been widely used in the scientific community. As a reference, the Environmental Valuation Reference Inventory (EVRI)[148], an extensive online database currently being assembled for policy-making purposes by Environment Canada, with documents produced by the European Union, the U.S. EPA, the World Bank, the environmental agencies of Chile and Mexico, and the Economy and Environment Program for Southeast Asia, as of November 2021, had 2535 studies containing CV methods.

Some authors have even used these methods to analyze the evolution of the perception of seismic risk across a timeframe (see, for reference, McGinnis (2004)). However, given the lack of historical data, it was not possible to replicate this analysis for the city of Lisbon. Thus, the implementation of time-fixed effects models, or even autoregressive models (e.g. Vector Autoregressive (VAR) and Autoregressive Conditional Heteroskedasticity (ARCH) models) is unviable.

Thus, the focus is on the spatially and demographic distribution of risk perception through an EDA of the results, by building new regression models as explained in section 3.10.3, and by incorporating it into joint regression models in section 3.11.

3.10.2. Surveys

The surveys will be designed to capture various features that will allow one to have a deep understanding of the perception of risk by real estate stakeholders. Given that, to the author's best knowledge, it will be one of the first of its kind to be conducted in the city of Lisbon, the author will try to capture a full range of features, which will be divided into four groups:

1. Demographics.
2. Housing.
3. Perception and action towards risk.
4. Willingness to pay for mitigating measures;

[148] For reference, see http://www.evri.ec.gc.ca/evri/

The first group, Demographics, will contain questions regarding the number of household members, marital status, number of children in the household, income level, head of household's age, gender and education level. This will assess the differences in perception across demographics.

The second group, Housing, will focus on housing-specific characteristics such as the ownership status (owned, private rental or public rental), construction type (masonry, reinforced concrete or unknown), age of construction, perceived structural resistance of their dwellings, total surface area, number of rooms[149]. This is critical information to have since it has been shown in previous studies that homeowners and renters have different subjective perspectives towards risk. Also, the anti-seismic quality of a dwelling largely depends on its construction type (see, for reference, Naoi et al. (2009)). The location of their home (neighbourhood and parish) will also be addressed in this section. This information will allow the author to assess the differences in taking risk-mitigating measures across the type of tenure and respondents' perception of structural features.

The third group will address the perception and attitude towards risk. The seismic risk perception will be ranked amongst other types of risks to their properties. Several statements will be presented to the respondents to assess their knowledge of risk-mitigating measures. The survey also attempts to capture the respondents' trust in their local authorities, central government, insurance companies and existing building codes for mitigation and disaster relief. These will be the base layer upon which the models will be built.

The fourth and last group will focus on the WTP for mitigating measures. The author will assess the number of respondents with seismic risk coverage in their insurance policies, their willingness to pay for structural retrofitting of their properties and their preferred payment method (e.g. through municipal taxes, monthly/yearly/one-time payment[150], direct payment with tax benefits) will be assessed.

Given the complexity of this topic, and the significance of having unbiased, accurate information, the next sections will address the methodologies used to select the sample (size, and sampling methods), to ensure an accurate representation of the population and avoid sampling bias and to minimize response bias. Finally, the methodology used to structure the survey will be covered and its final version presented.

The survey totals 40 questions, 39 of which are multiple choices/close-ended questions and only one is open-ended. It can be found in Annex B and was designed to take up to 10 minutes to complete[151].

[149] The total surface area and number of rooms will allow the author to determine the WTP for square metre/number of rooms. This is important information especially for implementing policy measures. See, for reference, Manganelli, Vona and De Paola (2018)

[150] As a premium in the property value or has part of the Property Transfer Tax (IMT)

[151] Following the research guidelines by Revilla and Ochoa (2017)

3.10.2.1. Sampling Strategy

This section will describe the sampling strategy's methodology followed. The author started by comparing some strategies and results obtained in previous studies. Table 3-6 shows some of the methods, sample sizes, and response rates obtained in previous studies.

Table 3-6 - Summary table with literature examples of stated preferences methods

	Sample Size	sampling strategy	Methods	Response Rates	Location	Total Population
Dooley et al (1992)	4000	random digit dialing (RDD)	telephone interview	61.40%	Orange County, Califonia, US	2,410,556 (1990)[(a)]
Hanley et al. (1998)	800	random sample	telephone interview	51.6%	Alberta, Canada	2,701,000 (1994)
Asgary and Willis (1997)	1300	multi-stage cluster sampling	written questionnaire	64%	Tehran, Iran	6,831,000 (1997)[(a)]
Asgary and Willis (1997)	600	multi-stage cluster sampling	written questionnaire	59%	Rasht, Iran	427,000 (1997)[(a)]
Asgary et al. (2007)	1000	cluster sampling method	written questionnaire	51.20%	Tehran, Iran	7,845,000 (2007)[(a)]
Palm (1998)	2600	random sample	written questionnaire	60%	Los Angeles, California, US	3,722,500 (1998)[(c)]
Palm (1998)	2000	random sample	written questionnaire	75%	Yokohama, Japan	3,307,140 (1995)[(d)]
Earnhart (2001)	464	two-stage cluster sampling	written questionnaire	22,.6%	Farifield, Connecticut, US	3,548 (2001)[(e)]
Tekeli-Yeşil et al. (2010; 2011)	1123	two-stage cluster sampling	field survey	93.60%	Istanbul, Turkey	12,585,000 (2010)[(a)]

source: (a) macrotrends.net; (b) StatCan; (c) LAO - Government of Califonia; (d) population.city; (e) U.S.

Given the values observed in previous studies, it is possible to conclude that random and two-stage cluster sampling are the most frequent sample methods. Therefore, initially, it was aimed to pursue a two-stage cluster sampling technique, as it allowed for a better representation of Lisbon's population.

A written questionnaire was chosen as the preferred method of distribution. This comes as a result of the analysis of previous studies and three other main factors: cost, time and bias minimization. Surveys of large samples require a significant amount of time and specialized staff to conduct, which implies a high financial cost. Thus, to maximize the number of respondents and minimize time and cost, person-to-person interviews or field surveys were excluded from the initial options. Telephone interviews would be a viable option, as they require less effort and are considerably less costly. However, as with the case of field surveys and person-to-person interviews, they are prone to interviewer bias[152] (see, for reference, Salazar (1990)). Hence, the data collected using self-completion questionnaires may produce more reliable results. The survey will be constructed using *ArcGIS Survey123* software, allowing georeferenced survey responses. Georeferenced responses can later be used in any level of aggregation, using administrative boundaries with parishes or counties or even terrain meshes (e.g., 500m x 500m grid), giving the flexibility to develop any future works with the data collected.

[152] According to Salazar (1990), "*ties of the interviewer generally are acknowledged to be key determinants of the outcome of an interview. Biases introduced by the interviewer can directly affect the validity and reliability of the ultimate findings of the study*".

According to Manganelli et al. (2018), "*statistics suggest that the minimum sample size does not depend on the size of the population. Therefore, it is logical that the calculation of the sample number does not take into account the size of the population when it is too large or unknown*". Hence, the sample size required will be determined by the variance (confidence interval or error margin), which measures the estimate's accuracy, and the desired width of the confidence interval (confidence level), which measures the security of the estimate.

The minimum sample size for vast or unknown population (n) is given by (EQ 18):

$$n = \frac{z^2 \cdot \hat{p}(1-\hat{p})}{\varepsilon^2} \qquad \text{[EQ 18]}$$

where z is the *z-score*, $\hat{p}$ is the expected prevalence (percentage) and ε is the confidence interval. The most commonly used confidence levels are 90%, 95% and 99%. In this work, the 95% confidence interval is used, which corresponds[153] to a *z-score* of 1.96. For $\hat{p}$, the value of 50% was chosen, which corresponds to the worst scenario possible[154]. The author will aim for a margin of error of 5%, leading to a minimum sample size of 385 responses. The second-best scenario would be to aim for a margin of error of 10%, as used in Manganelli et al. (2018), which leads to a minimum sample size of 97. This would be the minimum sample for a survey across the city of Lisbon. A higher number of responses will lead to a narrower margin of error; however, this number of responses would be enough to achieve statistical significance in the results. According to the provisional 2021 census results[155], Lisbon has a total of 545 923 inhabitants, with the distribution presented in Table 3-7.

Table 3-7 - Population distribution across parishes and required sample sizes

Parish	No. Inhabitants	%Population[1]	Sample Size[2]
Ajuda	14313	2.62%	10
Alcântara	13852	2.54%	10
Alvalade	33313	6.10%	23
Areeiro	21167	3.88%	15
Arroios	33307	6.10%	23
Avenidas Novas	23261	4.26%	16
Beato	12185	2.23%	9
Belém	16549	3.03%	12
Benfica	35367	6.48%	25
Campo de Ourique	22146	4.06%	16
Campolide	14794	2.71%	10
Carnide	18029	3.30%	13
Estrela	20308	3.72%	14
Lumiar	46338	8.49%	33
Marvila	35482	6.50%	25
Misericórdia	9660	1.77%	7
Olivais	32184	5.90%	23

[153] See, for reference, https://www.calculator.net/sample-size-calculator.html
[154] Corresponds to a scenario where nothing is known about the population
[155] See https://www.ine.pt/scripts/db_censos_2021.html (last visited August 25th, 2022)

Parish	No. Inhabitants	%Population[1]	Sample Size[2]
Parque das Nações	22382	4.10%	16
Penha de França	28485	5.22%	20
Santa Clara	23650	4.33%	17
Santa Maria Maior	10052	1.84%	7
Santo António	11062	2.03%	8
São Domingos de Benfica	34081	6.24%	24
São Vicente	13956	2.56%	10

1 – Percentage of Lisbon's total population; 2 – Sample size as percentage of the total number of Lisbon's inhabitants

As seen in Table 3-7, 385 responses would be enough to provide a statistically sound result for the city of Lisbon, but, given that we intend to analyze differences between parishes, a minimum of 30 respondents[156] per survey would be the bare minimum required to have a statistically significant analysis. However, the number of respondents should be significantly higher to guarantee a margin of error of 10% or 370-380 respondents per parish for 5%. This would lead to a significantly higher number of responses (from 2304 to 9024), considering the current 24 parishes. The number of total respondents required would be even larger to apply the same rational to the older parishes. In this work, the aim is a total of 1,032 results, 43 per parish, which would give a larger margin of error of 15%; however, we should note that this sample size is in line with the literature previously shown in Table 3-6.

Regarding the method of distribution, four different alternatives were developed: 1) send the survey by email through parish council internal services; 2) send it by email through some of the largest employers in the city (e.g. CTT, EDP, BNP Paribas, and other large companies); 3) send email to institutions in the city and to neighborhood associations; and 4) distribute flyers to residents with a QR code allowing for the survey to be completed online;

The timeframe for collecting data has been fixed to a maximum of 3 months or until the total desired number of responses is achieved. According to the provisional results of census 2021, most population in the city is aged between 25 and 64 (> 50%, except for Belém, at 48%). The minimum response rate should be at 60.31%, the average obtained for the written questionnaires, as shown in Table 3-6. The survey will also be pre-tested in a small group of 15 respondents to ensure understandability.

3.10.2.2. Comparison With Demographics

In order to guarantee the validity[157] of the results, one should assess the presence of sampling bias. This will be done by comparing the demographics of the respondents with the demographics of their parish and of the city of Lisbon. However, it is noteworthy that, given the distribution methods chosen

[156] As a general rule of thumb, for the Central Limit Theorem (CLT) to be applied, a minimum sample size of n≥30 should be considered. See, for reference, https://guides.fscj.edu/Statistics/centrallimit

[157] In simple terms, validity refers to the 'accuracy' and reliability refers to 'consistency' or 'reproducibility' of the CV results (Kealy et al. (1990) as cited in Venkatachalam (2004))

and that most population is of working age (see Table A.0-9 in Annex A), a significant sampling bias in the results is not expected.

3.10.2.3. Response Bias

It is well documented in the literature that several cognitive and response biases, commonly present in surveys, may jeopardize the accuracy of results (Kahneman & Tversky, 1979; Diamond & Haussman, 1994; Venkatachalam, 2004; Manganelli et al., 2018). It should be noted that the order and type[158] of questions, and even its wording, may also influence the decision-making process of respondents. Therefore, in this section, some of the most common sources of bias will be listed and the methods used to minimize their effects identified.

These biases can arise from respondents' personal experiences, their social context or demographics, or even from the interaction with interviewers, as previously mentioned. According to Manganelli et al. (2018), data collected using self-completion methods may produce more reliable results than face-to-face interviews, due to the reduction of interviewer bias, reducing influence in their responses[159], and because respondents are given more time to think. On the other hand, respondents of self-administered written questionnaires may try to look up online or consult other household members to choose the "right" answer. It can be hypothesized that using simple and clearly stated questions may reduce the need to consult external sources.

Table 3-8 presents a list of some of the biases most likely to affect the results, briefly describing the question groups[160] considered the most likely to be affected and how they are addressed in this work. These were considered so that their effect should be accounted for when designing the survey.

Table 3-8 – List of most probable biases to affect the results.

Bias	Description	Most likely to affect	How it is addressed
Iterviewer bias	The interviewer voice or even facial expressions may bias the results by implying a particular "right" answer.	Groups 3 and 4	Self-administered questionnaires over the internet, exclude human interaction, avoiding facial and voice hints
Social Desirability	Respondents provide a more "socially accepted" response instead of their honest thoughts on the subject.	All groups	The fact that this is an online, anonymous survey may decrease the effects of this bias. Also, we should include both direct and indirect question to minimize this effect.
Non-response	Occurs when people who don't respond to a survey question differ significantly to those who do respond	Groups 1 and 3	Use of clear statements avoiding misunderstanding or confusion and keeping the survey not too long

[158] Closed vs open-ended questions
[159] The interviewer voice or even facial expressions may bias the results by implying a particular "right" answer. Additionally, respondents may change their response, seeking the interviewer's approval, a phenomenon known as social desirability bias (Manganelli, Vona and De Paola 2018)
160 Question groups: 1) Demographics; 2) Housing; 3) Perception and action towards risk; Willingness to pay (WTP) for mitigating measures.

Bias	Description	Most likely to affect	How it is addressed
Demand bias	Participants change their behaviors or views simply because of prior knowledge of the research agenda (e.g. title of the study, prior information on the study subject and procedures)	Groups 3 and 4	Only disclosing as much information as it is needed to clearly understand the questions
Extreme response	Participants respond with an extreme view, even if they don't have an extreme opinion on the matter	Groups 3 and 4	Use of carefully worded questions that are clear and neutral in tone help to provoke a thoughtful response.
Voluntary response	Occurs when the sample is made of people who volunteered to take the part in the survey. These may have a larger than average interest/knowledge in the research topic, biasing the final results.	All groups	A high response rate (60% to 80%) helps to minimize potential problems with extrapolating to the population of interest
Range bias	Given a range of values, individuals may perceive it as the acceptable bid range and therefore they may adjust their offered bids (especially concerning when assessing WTP)	Group 4	Use of both closed and open-ended questions. Inclusion of zero as an acceptable answer.
Importance bias	Occurs when zero is not considered as an accepted value	Group 4	Include zero as an acceptable answer
Income effect	Occurs when the WTP is affected by the respondent's income level	Group 4	Include answers as a percentage of total income or total property value
Loss aversion	The loss of a commodity for an individual is considered to be greater than the gain derived from buying the same commodity.	Group 3 and 4	In this work we did not take measures to minimize this bias, as it should reflect the natural decision process made by households
Sequencing effect (or "question order bias")	The sequence of questions influences responses	All groups	Carefully sequencing questions in a neutral way
Strategic Bias	Occurs when an individual "deliberately misrepresents their preferences in order to influence the decision-making process" to influence, for instance, policymaking	Group 4	Stating that the survey will be addressed to a large number of respondents may disincentivize strategic bias, as respondents may conclude that their opinion will not have a large influence overall result
'Hypothetical bias' (Neill et al., 1994)	The potential divergence between the real and hypothetical payments	Group 4	A high response rate should minimize its effects
Framing Effect	The framing effect is when our decisions are influenced by the way information is presented.	Group 3 and 4	Use of neutral wording, a neutral sequencing of questions. Relates to Sequencing effect
Anchoring Bias	Cognitive bias that causes us to rely too heavily on the first piece of information we are given about a topic.	Group 3	Only disclosing as much information as it is needed to clearly understand the questions
Salience bias	Tendency to focus on items or information that are more noteworthy while ignoring those that do not grab our attention.	Group 3	Avoid highlighting the importance of seismic risk to this study, by asking about other types of risks
Hyperbolic Discounting	Hyperbolic discounting is our inclination to choose immediate rewards over rewards that come later in the future, even when these immediate rewards are smaller.	Group 3 and 4	In this work we did not take measures to minimize this bias, as it should reflect the natural decision process made by households

Bias	Description	Most likely to affect	How it is addressed
Ostrich Effect	Cognitive bias that causes people to avoid negative information.	Group 3	In this work we did not take measures to minimize this bias, as it should reflect the natural decision process made by households
Optimism Bias	Tendency to overestimate our likelihood of experiencing positive events and underestimate our likelihood of experiencing negative events.	Group 3	In this work we did not take measures to minimize this bias, as it should reflect the natural decision process made by households

Responses in groups 3 (perception and action towards risk) and 4 (willingness to pay for mitigation measures) pose the most serious concerns, as they may be more prone to biases.

In response to the range bias, it was chosen to include open-ended questions in the survey. However, the number of this type of questions will be limited, as they usually elicit a high number of “protest zeros” and a small number of extremely high responses (Manganelli et al., 2018). Regarding WTP questions (in group 4), the frequency of payments will also be accounted for. A one-time payment generally produces more conservative estimates (Kahneman & Knetsch, 1992). Finally, control questions should also be included to identify respondents ticking boxes randomly. In these cases, the questionnaires should be excluded from the sample.

There are usually two paths to be taken to check for the validity of CVM results: 1) testing for convergent validity – i.e., comparing the results to other revealed preference methods such as hedonic price models (see, for reference, Brookshire et al. (1982)).); and 2) collecting a second sample from the same population and comparing results. In this work, the results obtained through the survey will be compared with the results from hedonic price models. Still, mismatches in results will be expected due to differences in the stated actions and the actual behaviour of respondents. Due to the relatively large number of respondents aimed for, the author will not consider collecting a second sample of 1,032 respondents will not be considered, as it would require a significant increase in time and effort. However, results in sub-samples may be compared to check their overall consistency.

3.10.3. SP Regression

The following step on the SP method, was to regress the WTP on a set of explanatory variables such as gender, age, whether own or rent, number of children under 18 living at home, whether married or not, education, and earthquake experience. Hence, the WTP is defined as (EQ 19):

$$WTP = f(age, gender, income, ownership, etc) \qquad \text{[EQ 19]}$$

and the regression function can be defined as (EQ 20):

$$WTP = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon \qquad \text{[EQ 20]}$$

here WTP is the willingness-to-pay, β_0 is the intercept, X_n are the selected explanatory variables, β_1 are the regression coefficients and ε is the error term. WTP will be measured as the willingness to pay a premium for a less vulnerable property, a binomial logistic regression will be used. This will enable to model the WTP based on demographic features, which could be helpful in defining policies for the structural retrofitting of the city's housing stock.

3.11. Joint Preferences Model

The join preferences model will be using survey data as the independent variables and property valuation as the dependent variable. To limit the influence of the market, a 300m (approx. 5-minute walk) circular area around each respondent's address is considered. This approach should allow for an accurate representation of each local market pricing dynamics. The join model will take the form of a multivariate regression model, such as (EQ 21):

$$Y_{300} = f(a, b, \dots, n) \quad \text{[EQ 21]}$$

Where Y_{300} is the average sale price in a 300m of each respondent, and $a, b, \dots, n$ are the responses to some perception questions in the survey.

4. Results

4.1. Overview

This chapter will review the results, showing what has been concluded from the Explanatory Data Analysis (EDA) section, the Revealed Preferences models (RP), the remaining Stated Preferences (SP) models, based on the surveys, and the Joint Regression model. However, before proceeding to present the results, the three initial hypotheses that laid the ground for the development of this thesis are recalled:

1. R1: Do market values reflect a preference for properties less vulnerable to earthquakes?

2. R2: How do residents and investors in Lisbon perceive risk?

3. R3: Does their risk perception affect their willingness to pay for property?

These four sections address our three research questions as follows:

1. EDA is the basis for understanding the three research questions. It aims to provide a general understanding of the dataset, the distribution of properties, their features (such as the price and BRD) and spatial distribution and how our sample compares with the actual building stock. It should also provide a first look at the relationship between variables in our dataset and what components or variables should be used in our models to follow;

2. The RP section focuses on our first research question (R1), using the transaction dataset to validate whether there is a market preference for less vulnerable properties. In order to achieve robust results, five different types of models will be presented: Feasible Generalized Least Squares (FGLS), PooledOLS, Fixed-Effects, Multivariate Adaptive Regression Spline (MARS), and Extreme Gradient Boost (XGBoost). One should highlight that these vary in complexity, explainability and fitness to use for spatial modelling of data;

3. The SP section focuses on the second research question, foreshadowing the third research question. It analyses the results of a survey of 325 respondents in Lisbon. It draws some first takeaways from each of the four question groups (Demographics (A), Housing (B), Perception and action towards risk (C) and Willingness to pay for mitigating measures (D)). It also presents a regression of the *"willingness to pay to a premium for a less vulnerable"* using encoded responses from groups A, C and D as input variables;

4. The Joint Regression Model section presents the results from the last model, which combines encoded answers from the survey and observable market values from the transaction database, to focus on whether the risk perception of residents affects market values (R3).

The next section will start by presenting the EDA results.

4.2. R1, R2 and R3: Exploratory Data Analysis

As mentioned in section 3.6, the data used in this thesis was provided by the Lisbon Municipality, containing information regarding the sales of 8.726 properties, ranging between 2008 and 2018. A statistical description of variables, including measures of central tendency and dispersion, can be found in Table A.0-10, Annex A.

The database was first analyzed for some possible errors. The first step was to analyze the location of the properties. Figure 4-1 shows the exact location of each property. Most properties contained their address, latitude and longitude, except for 1029 (11,79% of the dataset).

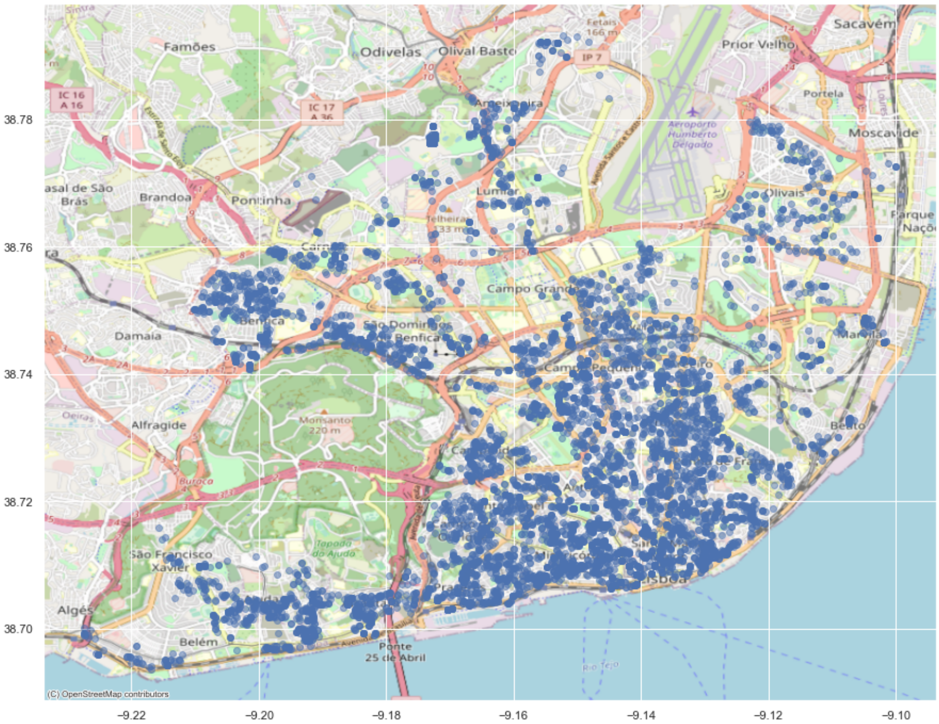


Figure 4-1 - Location of properties. Own elaboration.

We can see that most properties are located in the city centre, especially in nowadays *Arroios*, *Avenidas Novas*, *Santa Maria Maior* (see Figure 4-2).

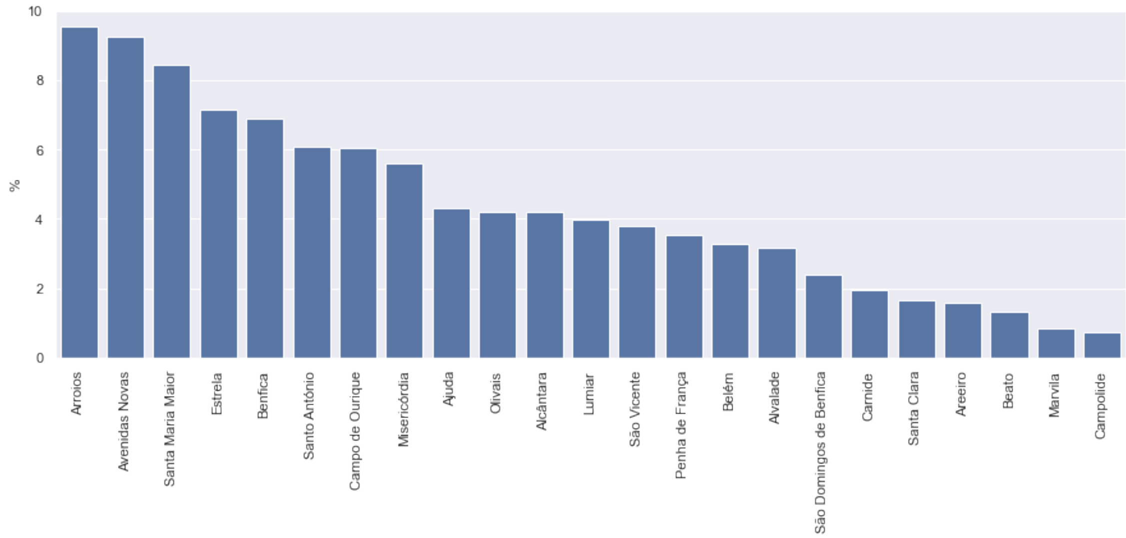


Figure 4-2 - Number of properties by parish as a percentage of total

A comparison between the number of properties in each parish and the total number of floors (NPAV) according to the BGE 2011 was done. It is acknowledged that the total number of floors cannot be simply converted as the number of properties in each parish, as each building type could have a various number of dwellings per floor, however, it was considered to be a good proxy for the number of properties. According to the BGE 2011, there are a total of 197.336 floors spread across 52.696 buildings. By comparing the percentage by parish in the BGE 2011 and in the database, it provides an idea of the magnitude of sampling bias in the dataset (see Table 4-1). At the extremes, there is an overrepresentation of parishes like *Santa Maria Maior* (+4,94%), *Estrela* (+3,72%) and *Avenidas Novas* (+2.34%) and an underrepresentation of *Parque das Nações* (-2.58%), *Olivais* (-2.47%) and *Alvalade* (-1.90%). However, one can argue that, overall, this database is a good representation of Lisbon's housing stock spatial distribution.

Table 4-1 - Comparison of properties by parish

Parish	NPAV	% BGE	% Database	Difference
Ajuda	6393	3.24%	2.77%	-0.47%
Alcântara	5095	2.58%	3.60%	1.02%
Beato	4812	2.44%	1.21%	-1.23%
Benfica	11871	6.02%	5.30%	-0.71%
Campolide	6116	3.10%	2.77%	-0.33%
Carnide	5441	2.76%	1.66%	-1.09%
Lumiar	11363	5.76%	4.19%	-1.57%
Marvila	6455	3.27%	2.21%	-1.06%
Olivais	10132	5.13%	2.67%	-2.47%
São Domingos de Benfica	9776	4.95%	4.80%	-0.16%
Alvalade	10823	5.48%	3.59%	-1.90%
Areeiro	7067	3.58%	3.34%	-0.24%
Arroios	13633	6.91%	8.04%	1.13%
Avenidas Novas	10114	5.13%	7.46%	2.34%
Belém	8547	4.33%	4.38%	0.05%

Parish	NPAV	% BGE	% Database	Difference
Campo de Ourique	8406	4.26%	4.04%	-0.22%
Estrela	9722	4.93%	8.65%	3.72%
Misericórdia	8356	4.23%	5.92%	1.68%
Parque das Nações	5444	2.76%	0.18%	-2.58%
Penha de França	9747	4.94%	3.60%	-1.34%
Santa Clara	5442	2.76%	0.98%	-1.78%
Santa Maria Maior	8753	4.44%	9.37%	4.94%
Santo António	6935	3.51%	5.64%	2.13%
São Vicente	6893	3.49%	3.63%	0.13%

Regarding the year of the sale, most properties were sold between 2010 and 2015 (60.29%) (see Figure 4-3). It is noteworthy that this corresponds to the period of a down market, the consequence of the 2007-2008 financial crisis, which was only felt a few years later in Portugal, as described in section 2.8.1.1.

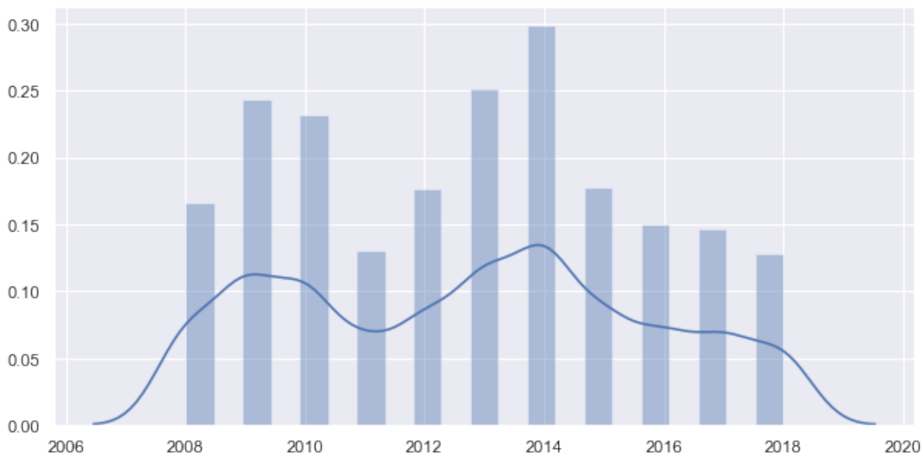


Figure 4-3 - Transaction date [Year of the properties in the database

Most properties in the dataset were located from the ground floor to the third (approximately 71.55% of properties). This was expected since, according to BGE 2011[161], more than two thirds of the buildings in Lisbon have between one and six floors (Figure 4-4).

Regarding the properties area, the author found some errors in the database, reporting apartment areas larger than 500 square meters, which is highly unlikely for a Lisbon apartment. Fortunately, these represented a small fraction of 0.02% of the dataset; hence they were removed from the analysis. Likewise, apartments with less than 30 m^2 (1.58% of total dataset) were also removed. After their removal, the mean area of properties in the dataset was found to be 102.59 m^2, with a standard deviation of 56.23 m^2, indicating a predominance of smaller one-bedroom to three-bedroom apartments. The median area stands at 88 m^2. Figure 4-5 shows the area distribution.

[161] Buildings Georreferenced Database 2011 (Base de Georeferrenciação de Edifícios - BGE), kindly provided by the National Statistics Institute (INE)

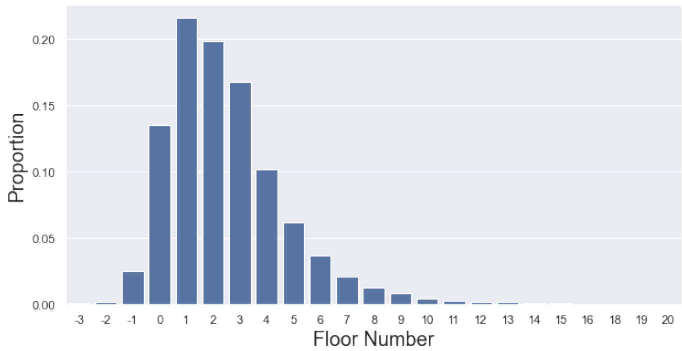


Figure 4-4 - Number of properties by floor, as percentage of total dataset

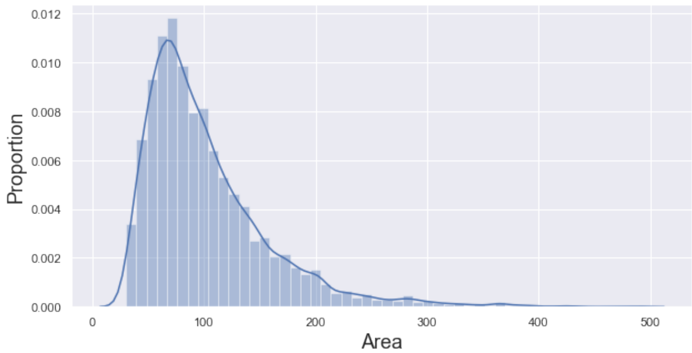


Figure 4-5 - Area distribution after outliers' removal

Lumiar shows the largest median area of the properties, at 122.80 m^2, followed by *Avenidas Novas* (111.80 m^2) and *Estrela* (107 m^2). *Beato* (65 m^2), *São Vicente* (68 m^2) and *Ajuda* (70 m^2) show to be the locations of smaller dwellings.

Regarding the construction year, 98.61% of properties were built during the 20th and 21st centuries[162] (see Figure 4-6).

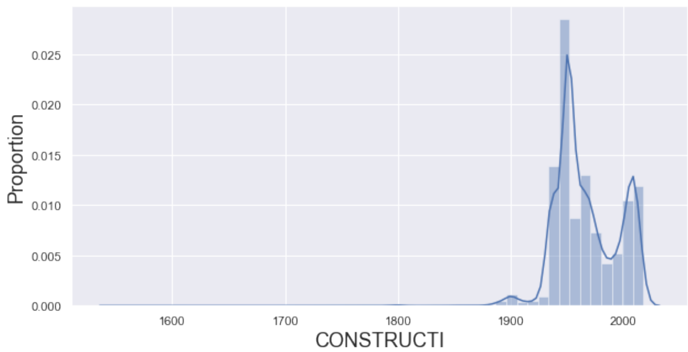


Figure 4-6 - Distribution of construction years

The older properties are mostly located in the city centre, in the parishes of *Misericórdia*, *Arroios* and *Santa Maria Maior*, while the newer are located in *Avenidas Novas*, *Santa Maria Maior*[163] and *Estrela*.

[162] The most recent property was built in 2018
[163] Probably as result of rehabilitation programs

Overall, compared with the construction epochs defined in BGE 2011, we may say the dataset is representative of the city's housing stock; however, it seems to be slightly underrepresented in the older buildings (built before 1919) and overrepresented in the buildings built between 1946 and 1960 (see Table 4-2). Nevertheless, one should note that the data from BGE 2011 refers to buildings, while this dataset refers to unique properties (mostly apartments).

Table 4-2 - Percentage of buildings by construction epoch. Comparison of BGE 2011 and dataset

Epoch	BGE 2011	Dataset	Difference
before 1919	19.51%	2.14%	17.36%
1919 to 1945	18.50%	14.75%	3.75%
1946 to 1960	24.95%	33.41%	-8.45%
1961 to 1970	13.22%	11.28%	1.94%
1971 to 1980	8.23%	6.84%	1.39%
1981 to 1990	4.05%	3.87%	0.18%
1991 to 1995	2.22%	1.93%	0.30%
1996 to 2000	3.49%	2.26%	1.23%
2001 to 2005	3.36%	4.06%	-0.70%
2006 to 2011	2.48%	8.74%	-6.26%
after 2011	N/A	5.16%	---

Regarding the number of rooms, we see that most of the properties are one to three-bedroom apartments (70,20% of total) (Figure 4-7).

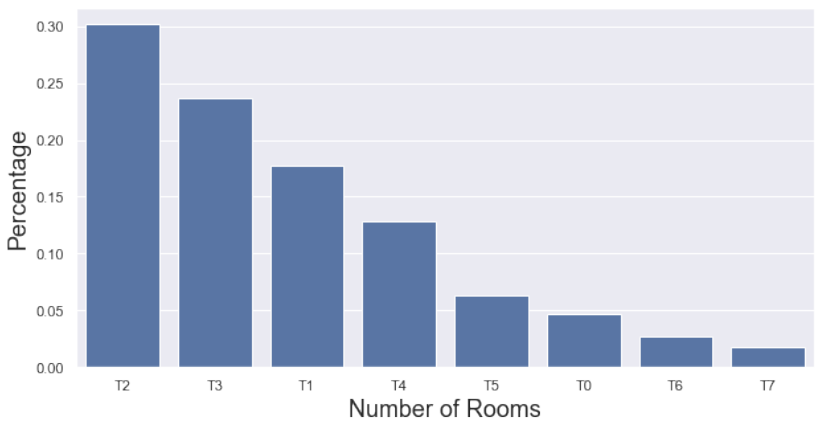


Figure 4-7 - Percentage of properties by number of rooms[164]

The analysis of the yearly normalized prices reveals that the smaller apartments are some of the most expensive, on a price per square metre basis (see Figure 4-8)

[164] In this image we are using the Portuguese designation for the number of rooms. For instance, a T2 apartment has two bedrooms while a single room apartment is defined as T1.

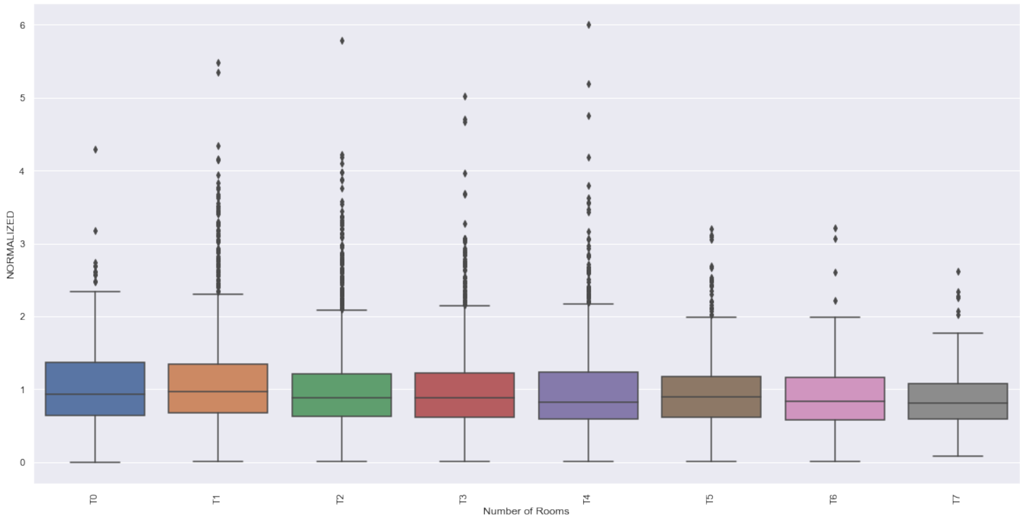


Figure 4-8 - Yearly normalized prices by number of rooms

The most expensive properties are mainly located in the city centre in the parishes of *Avenidas Novas*, *Misericórdia* and *Arroios* (Figure 4-9).

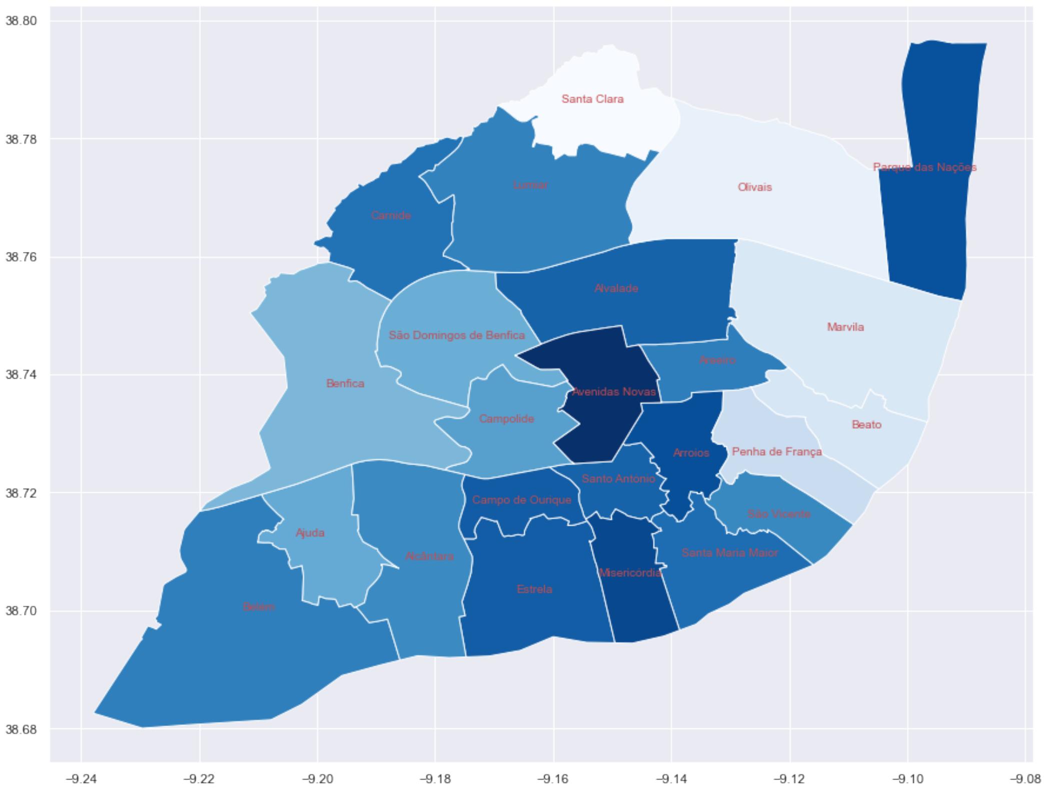


Figure 4-9 - Geo heatmap of normalized sales in the dataset. Higher values with darker shades.

Although they present the highest average prices, the wide range of values that can be observed in these parishes is also noticeable (Figure 4-10). It is possible to hypothesize that this dispersion of values should mainly be related to differences in the buildings' condition and the amenities presented to homebuyers.

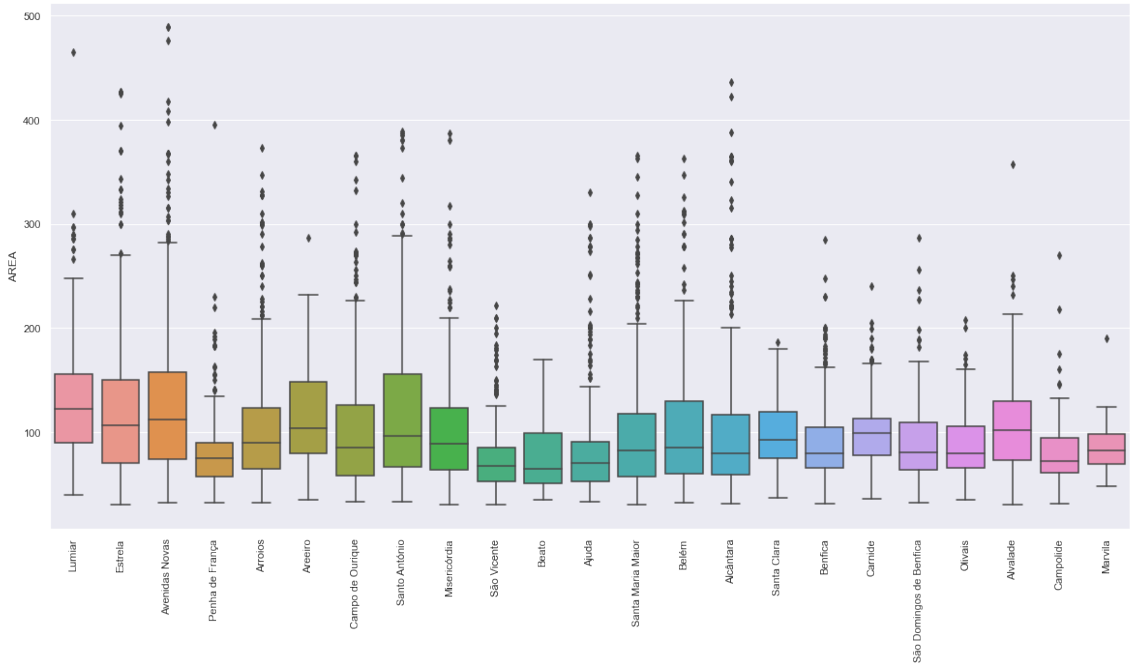


Figure 4-10 - Dispersion of prices across parishes

Buildings and apartments' condition have been classified as "bad", "reasonable", and "good". Most buildings appear to be in reasonable condition according to the dataset, with 41.05% being classified as in "good" condition, 53.11% in "reasonable" condition, and only 5.86% in "bad" condition (see Figure 4-11). The apartments follow a similar distribution, with 37.26% as "good", 52.46% labelled as "reasonable" and only 10.26% as "bad" (see Figure 4-12).

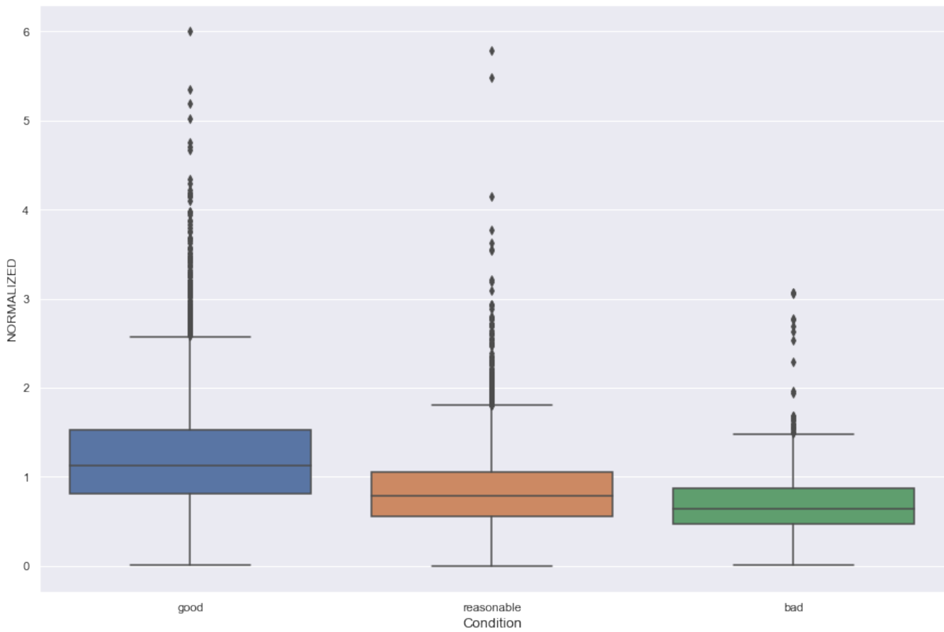


Figure 4-11 - Distribution of prices by building condition

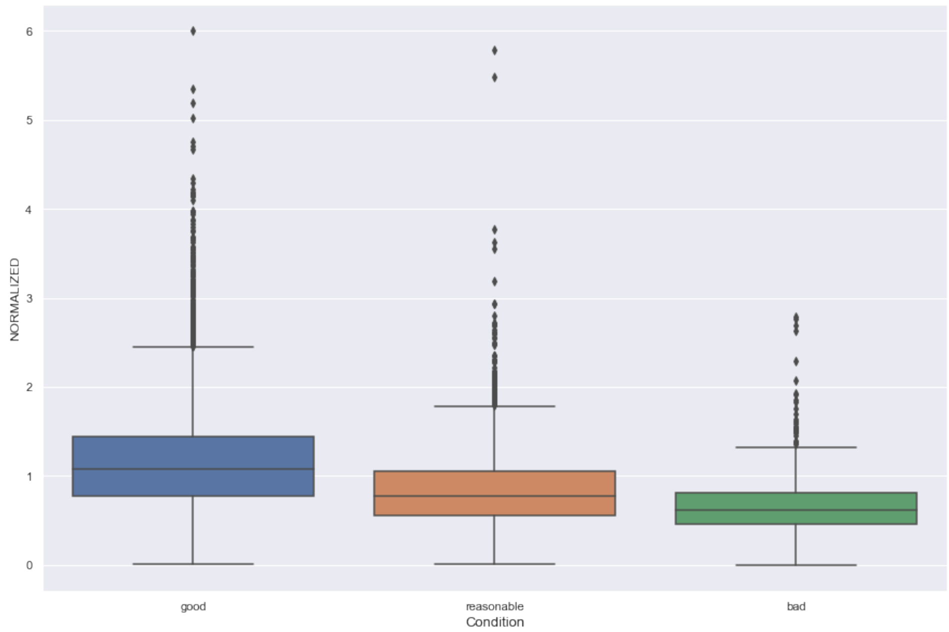


Figure 4-12 - Distribution of prices by apartment condition

Only 0,17% of the properties are located in Urban Areas of Illegal Origin (AUGI[165]), and only 4.27% are managed by Gebalis[166] as part of the city's social housing stock (see Figure 4-13).

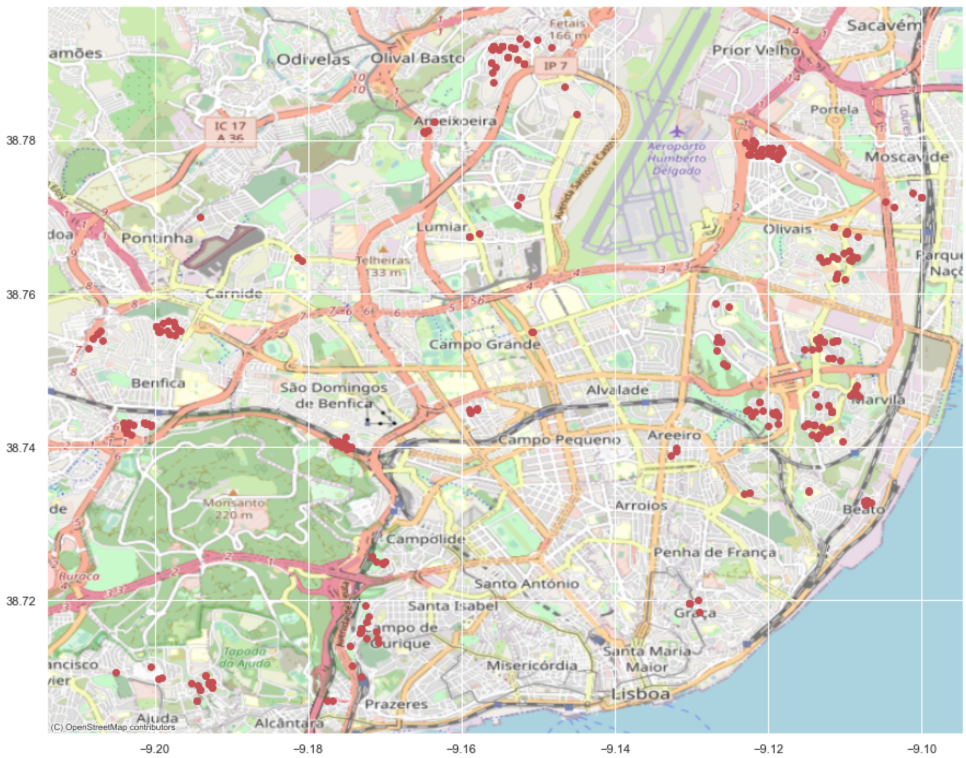


Figure 4-13 - Properties managed by Gebalis

[165] Translated from Portuguese, Áreas Urbanas de Génese Ilegal (AUGI)

[166] Gebalis is a municipally owned public company, responsible for Lisbon's social housing management. More information in https://www.gebalis.pt/Paginas/default.aspx

The dataset also contains information regarding noise levels (RUIDO) (on a scale from 0 – lowest noise levels- to 5 – highest), the area's vulnerability to flooding (VULInund), landslides (VULMass) and earthquakes (VULSismo)[167] (on a scale from 0 to 4). Table 4-3 and Table 4-4 summarize the percentage of properties in each class.

Table 4-3 - Percentage of properties by noise level

Class	RUIDO
0	11.79%
1	54.92%
2	15.13%
3	9.07%
4	7.40%
5	1.66%

Table 4-4 - Percentage of properties by risk class

Class	VULInund	VULSismo	VULMass
0	70.12%	11.79%	97.99%
1	16.53%	16.72%	0.74%
2	3.79%	40.15%	0.83%
3	9.54%	22.14%	0.42%
4	---	9.19%	---

Looking at the spatial distribution of the values (Figure 4-14 to Figure 4-17), some spatial patterns are noticeable: 1) most properties with high noise levels are located near the largest distribution lanes in the city, namely in *Alcântara*, *Campolide* and *Avenidas Novas*; 2) the city centre parishes of *Misericórdia*, *Santa Maria Maior* and *Arroios*, despite showing as one of the location with the highest vulnerability to flooding events and earthquake risk, based on soil type, is also one of the highest values in the city, hinting that investors and homebuyers alike are oblivious to earthquake risk; and 3) the zones with the higher risk to landslides are located near the river Tagus, and especially across Lisbon's downtown.

[167] Solely based on the soil type. We should note that the original Soil Seismic Vulnerability Map, elaborated by Faculdade de Ciências da Universidade de Lisboa, only has 4 classes: very high, high, average and low. The latter 3 types can be associated to the one defined in the Safety Regulations and Actions for Building Structures and Bridges (Regulamento de Segurança e Acções para Estruturas de Edifícios e Pontes – RSA). The full map is available at: https://www.lisboa.pt/fileadmin/cidade_temas/seguranca/documentos/Carta_de_Vulnerabilidade_Sismica_dos_Solos.pdf (Last visited September 1, 2022)

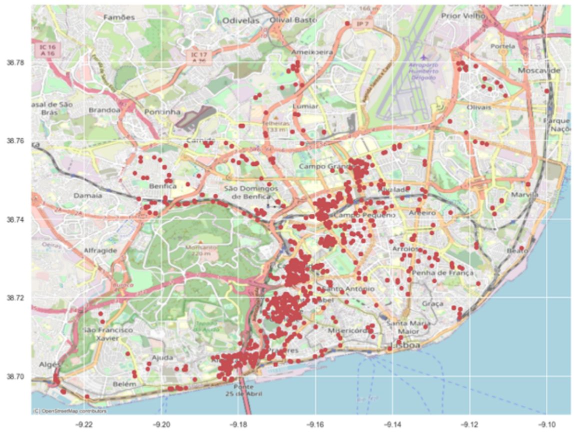


Figure 4-14 - Properties with higher noise levels (RUIDO>=3)

Figure 4-15 - Properties with higher vulnerability to earthquakes, based only on soil type (VULSismo>=3)

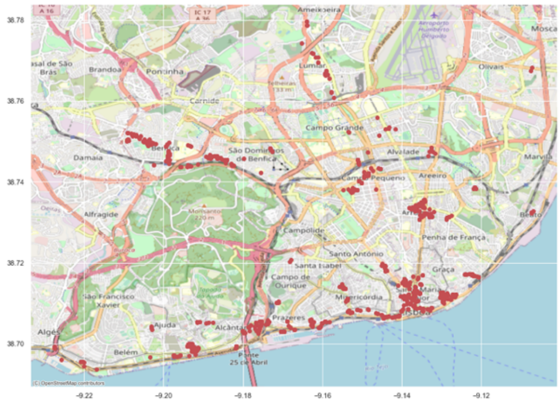


Figure 4-16 - Properties with higher vulnerability to flooding (VULInund>=2)

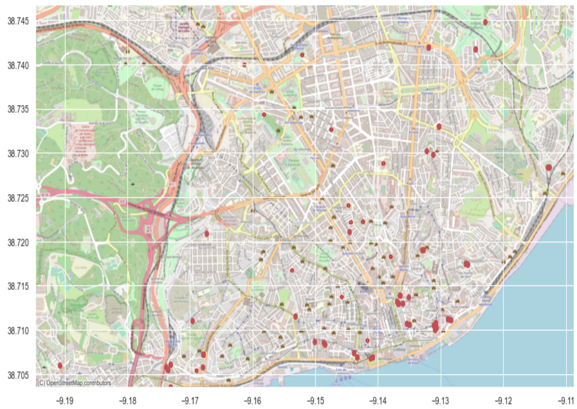


Figure 4-17 - Properties with higher vulnerability to landslides (VULMass>=2)

A correlation analysis of the values in the dataset was also carried out. Figure 4-18 shows the correlation heatmap. Notably, only the NORMALIZED price variable has been chosen as a proxy. This happens because there is a high correlation between all price variables; thus, including them in this analysis would be redundant. Additionally, only numerical variables have been considered in this correlation analysis.

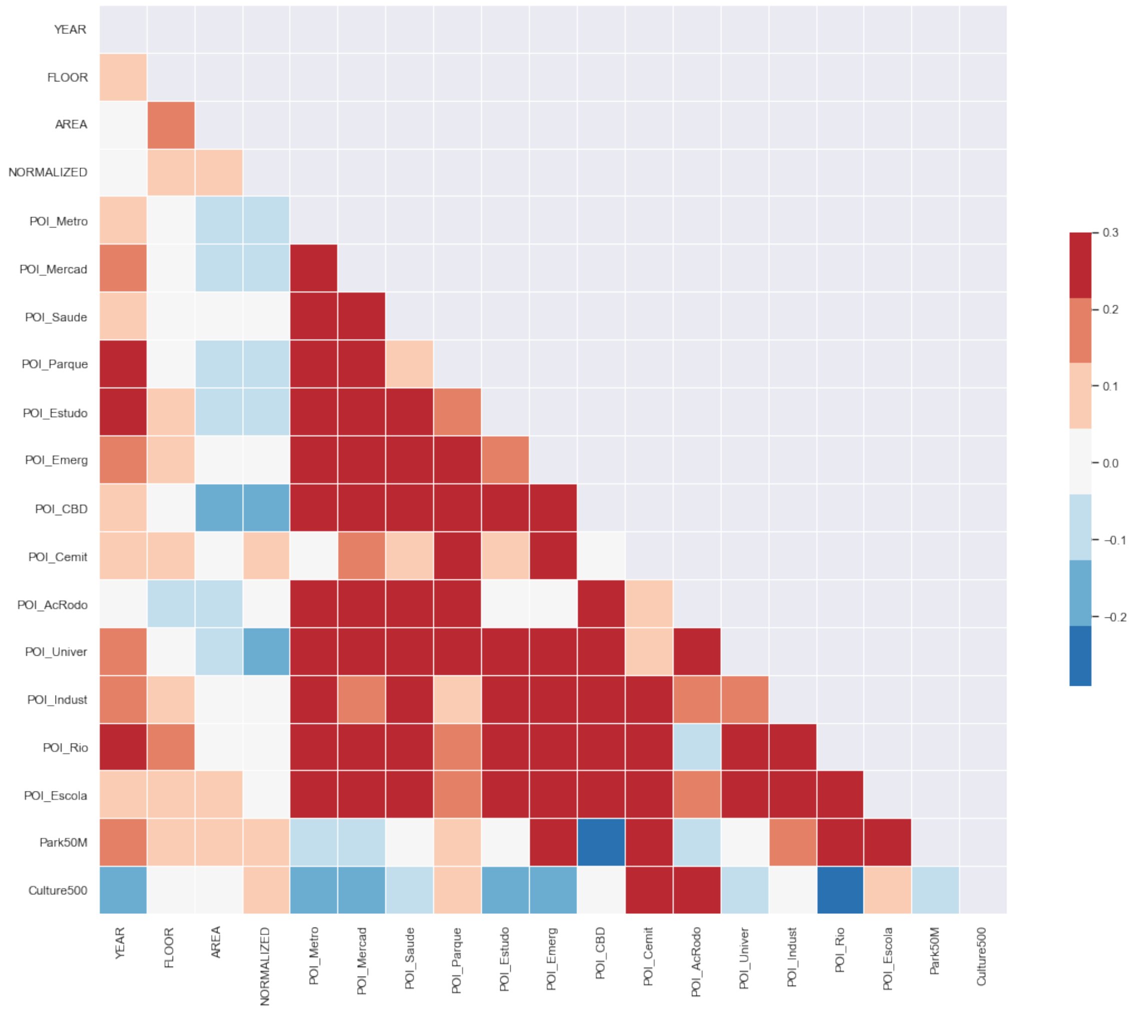


Figure 4-18 - Correlation heatmap

From the correlation analysis, and broader EDA, it is possible to take some early conclusions:

1. Overall, the sample seems to be representative of Lisbon's total housing stock, regarding its spatial distribution and age; however, there is some level of overrepresentation of parishes like *Santa Maria Maior* (+4,94%), *Estrela* (+3,72%) and *Avenidas Novas* (+2.34%) and an underrepresentation of *Parque das Nações* (-2.58%), *Olivais* (-2.47%) and *Alvalade* (-1.90%). Buildings prior to the 1900s also seem to be underrepresented (-17.36%) and buildings from the 1940s to the 1960s seem to be overrepresented (+8.45%).

2. The condition of the building, the condition of the apartment, the construction year, the area and the floor are some of the variables hinting a positive impact with the yearly normalized prices. This is consistent with the initial expectations and with the existing literature. Newer apartments with better condition are expected to be the most valued by the market. Larger apartments are also usually more expensive than smaller ones[168], and higher floors are also most valuable since they enjoy less noise and pollution, and often better views than lower

[168] Although we should note that the price per square metre usually decreases with the size of the apartment, as smaller units usually have a broader market, with higher demand, contrasting to larger ones which are often only preferred by numerous families. Given the current trends in demographics, one would expect a prevailing demand in the small apartments' market.

floors[169]. However, one should note the positive correlation between area and price per square metre. This could happen due to the highest dispersion of values in in some larger typologies.

3. The distance to the CBD (-0.163) and to the Universities (-0.140) seems to be the most negatively correlated variables to the prices. This is consistent with the existing literature, as proximity to business and education facilities is usually highly valued in the market.

4. Investors and homebuyers seem to be oblivious to natural hazards such as earthquakes, floods and landslides. This is shown by the low correlation with prices and by the spatial distribution of properties with the highest vulnerability to natural hazards.

4.2.1. Spatial Analysis of Property Values and Building Resistance Deficit

This section will focus on the relation between the spatial distribution of property values (Figure 4-19) and seismic vulnerability through a Building Resistance Deficit (BRD) spatial analysis and a spatial autocorrelation price analysis.

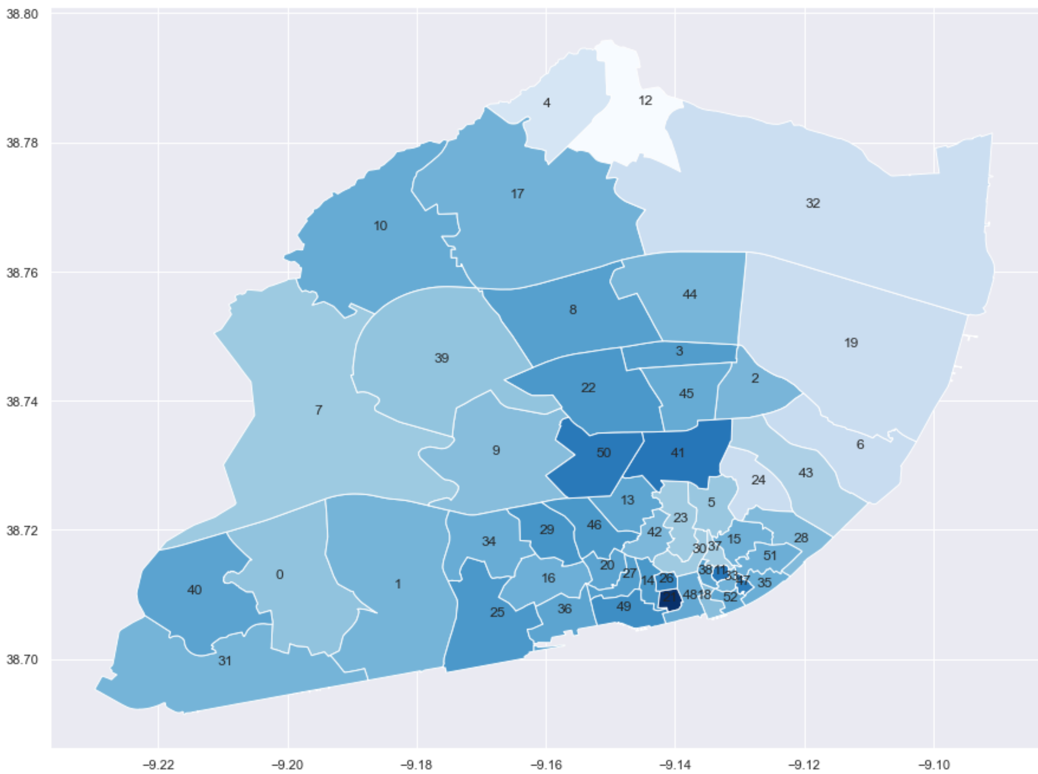


Figure 4-19 - Mean normalized values distribution in old parishes (pre-2013)[170].

[169] Although this dynamic may be different in older buildings with no elevator, in which case the lower levels may be more valuable.
[170] Parishes: 00 - Ajuda; 01 – Alcântara; 02 - Alto do Pina; 03 – Alvalade; 04 – Ameixoeira; 05 – Anjos; 06 – Beato; 07 – Benfica; 08 - Campo Grande; 09 – Campolide; 10 – Carnide; 11 – Castelo; 12 – Charneca; 13 - Coração de Jesus; 14 – Encarnação; 15 – Graça; 16 – Lapa; 17 – Lumiar; 18 – Madalena; 19 - Mártires; 20 – Marvila; 21 – Mercês; 22 - Nossa Senhora de Fátima; 23 – Pena; 24 - Penha de França; 25 – Prazeres; 26 – Sacramento; 27 - Santa Catarina; 28 - Santa Engrácia ; 29 - Santa Isabel; 30 -

In order to analyze the spatial dependency of values, a Global Moran-I test was done and the Local Indicators of Spatial Association (LISA) calculated, using a Queen contiguity matrix. As mentioned in section 4.4, these methods allow for the detection of spatial correlation of values. The old parishes were analyzed first due to their smaller size, which would be more approximate to the ideal homogenous units of analysis. The test statistic for the Moran-I test, for old parishes was 0.355, with a significance level of 0.1%. It is important to remind that Moran-I tests the hypothesis that the values on the map are randomly generated. The 0.1% significance level means that, if we generated a large number of maps with the same values but randomly allocated over space and calculated the Moran's I statistic for each of those maps, only 0.1% of them would display a larger (absolute) value than the one we obtain from the real data. Given the results, we can reject the null hypothesis that the values on the map are randomly generated can be rejected and it is concluded that the location of properties is a defining feature of their price. Figure 4-20 shows the relationship between the normalized values (price) and the spatial lag, which clearly shows signs of positive spatial autocorrelation, given the positive slope of the regression curve.

This is associated with the presence of positive spatial autocorrelation: similar values tend to be located close to each other. This means that high-value locations tend to be surrounded by other high-value locations, and low-value locations tend to be surrounded by low-value locations, hence clustered over space. This is according to the initial expectations, as real estate markets tend to behave in a "spill-over" pattern, leading to clusters of high value and low value locations.

However, this does not mean that this is the only situation in the dataset. There can, of course, be particular cases where low ones surround high values and vice-versa. This is clearly revealed by the LISA indicators, which are similar to a Global Moran-I test but applied to each unit of analysis. It compares its values with its neighbours, resulting in as many statistics as original observations.

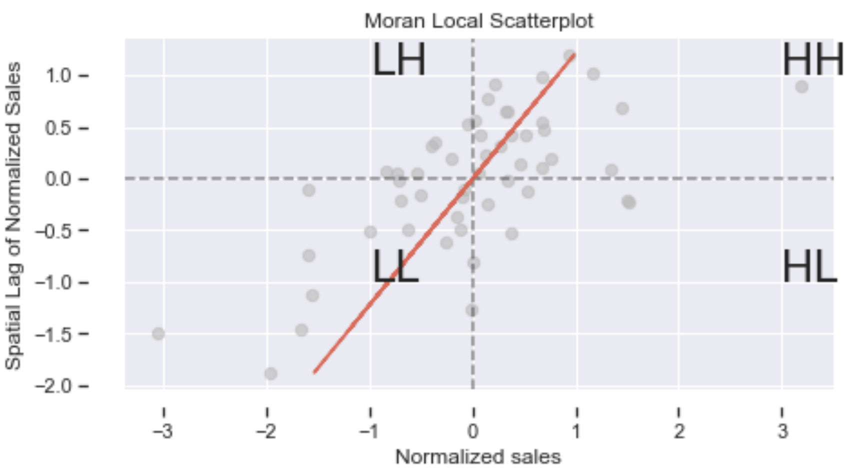


Figure 4-20 - Local Moran Scatter plot (old parishes).

Santa Justa; 31 - Santa Maria de Belém; 32 - Santa Maria dos Olivais; 33 – Santiago; 34 - Santo Condestável; 35 - Santo Estêvão; 36 - Santos-o-Velho; 37 - São Cristóvão e São Lourenço; 38 - São Domingos de Benfica; 39 - São Francisco Xavier; 40 - São João; 41 - São João de Brito; 42 - São João de Deus; 43 - São Jorge de Arroios; 44 - São José; 45 - São Mamede; 46 - São Miguel; 47 - São Nicolau; 48 - São Paulo; 49 - São Sebastião da Pedreira; 50 - São Vicente de Fora; 51 – Sé; 52 - Socorro

The spatial distribution of values indicates that the parishes of *Lumiar*, *Ameixoeira*, *Charneca* and *Santa Maria dos Olivais* are a cluster of low values, and the parishes of *São Paulo, Encarnação, Mártires, Sacramento*, and *São Nicolau* are a cluster of high values. It also shows the parish of *Beato*, surrounded by other low value locations, namely *Marvila*, *Alto do Pina* and *São João* (Figure 4-21). This hints that positive values are spilled over from the higher valuations in the downtown area to the lower valuations on the city's outskirts. However, it should be noted that different results would be expected in the analysis using the new parishes, as the newly created parish of *Parque das Nações* is currently one of the most prime locations for real estate properties.

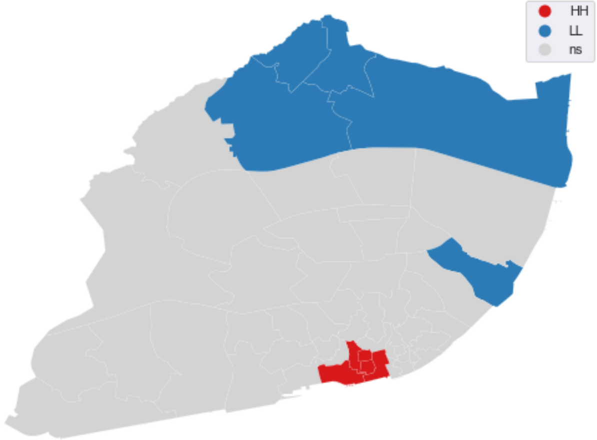


Figure 4-21 - LISA indicators (old parishes). Own elaboration.

After calculating the BRD, as explained in section 3.7.5, it is noticeable the lack of importance attributed by investors and homebuyers to the vulnerability of the housing stock (Figure 4-22). The highest mean values of BRD are located in the parishes of *São Nicolau* (0.838), *Madalena* (0.721) and *São Paulo* (0.627). The full table of mean BRD values can be found in Annex A (Table A.0-11).

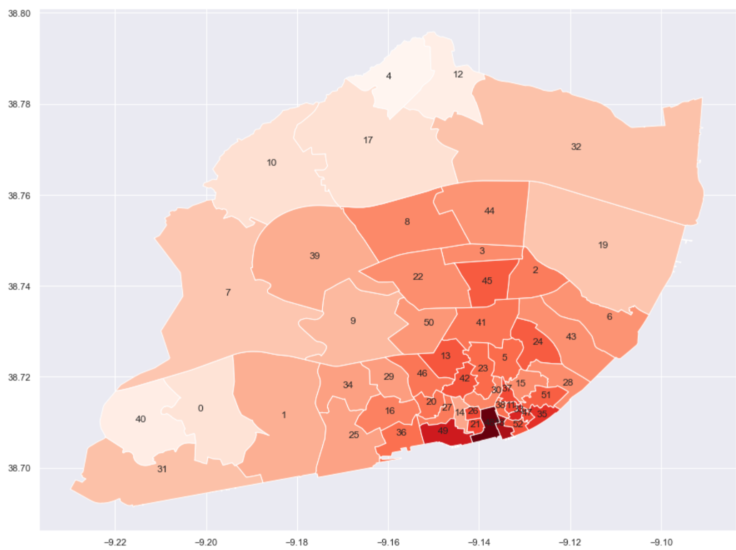


Figure 4-22 - Mean BRD values by parish (old parishes)[171]. Higher values with darker shades.

[171] Parishes: 00 - Ajuda; 01 – Alcântara; 02 - Alto do Pina; 03 – Alvalade; 04 – Ameixoeira; 05 – Anjos; 06 – Beato; 07 – Benfica; 08 - Campo Grande; 09 – Campolide; 10 – Carnide; 11 – Castelo; 12 – Charneca; 13 - Coração de Jesus; 14 – Encarnação; 15

Under the new administrative limits, the results are similar. The Global Moran-I test resulted in a test statistic of 0.220, with a 3,6% significance level (Figure 4-23). The larger area covered by the new parishes leads to less homogenous units of analysis, thus leading to less statistical significance in the spatial correlation test, however, it is still lower than 5%.

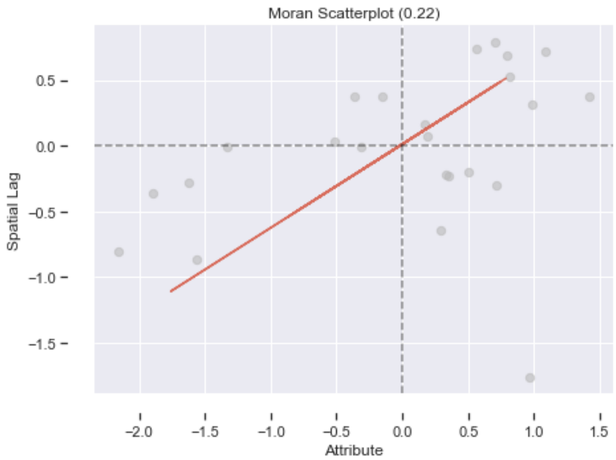


Figure 4-23 - Moran Scatterplot for normalized prices (according to new parishes)

As expected, the LISA indicators show *Parque das Nações* as an outlier, showing higher values than its neighbouring parishes (Figure 4-24)[172]. Nonetheless, it is clear that the high property values are spilled over from the city centre to the city's outer limits.

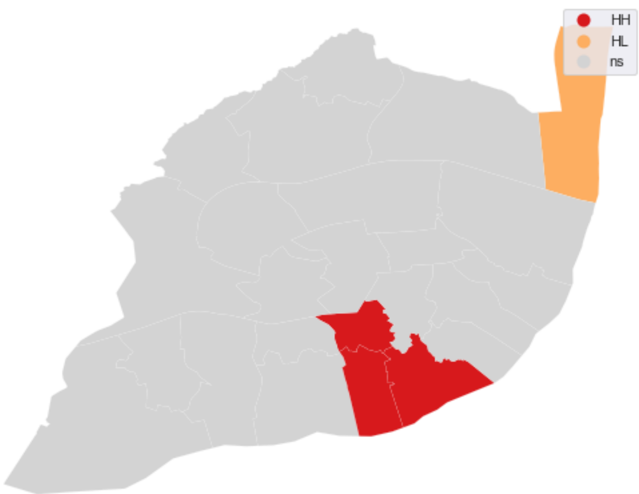


Figure 4-24 - LISA indicators for the new parishes.

– Graça; 16 – Lapa; 17 – Lumiar; 18 – Madalena; 19 - Mártires; 20 – Marvila; 21 – Mercês; 22 - Nossa Senhora de Fátima; 23 – Pena; 24 - Penha de França; 25 – Prazeres; 26 – Sacramento; 27 - Santa Catarina; 28 - Santa Engrácia ; 29 - Santa Isabel; 30 - Santa Justa; 31 - Santa Maria de Belém; 32 - Santa Maria dos Olivais; 33 – Santiago; 34 - Santo Condestável; 35 - Santo Estêvão; 36 - Santos-o-Velho; 37 - São Cristóvão e São Lourenço; 38 - São Domingos de Benfica; 39 - São Francisco Xavier; 40 - São João; 41 - São João de Brito; 42 - São João de Deus; 43 - São Jorge de Arroios; 44 - São José; 45 - São Mamede; 46 - São Miguel; 47 - São Nicolau; 48 - São Paulo; 49 - São Sebastião da Pedreira; 50 - São Vicente de Fora; 51 – Sé; 52 - Socorro

[172] Note that HH and HL correspond to high-high clusters of values and high-low, or high value outliers.

By comparing the BRD results to property prices, the same conclusion regarding the importance given by investors and homebuyers to the vulnerability of the housing stock can be drawn. The parishes of *Santa Maria Maior* (0.545), *Misericórdia* (0.464) and *Santo António* (0.452) show the highest values of BRD (Figure 4-25) while showing some of the highest property values in the municipality (see Figure 4-26). The full table of mean BRD values can be found in Annex A (Table A.0-12).

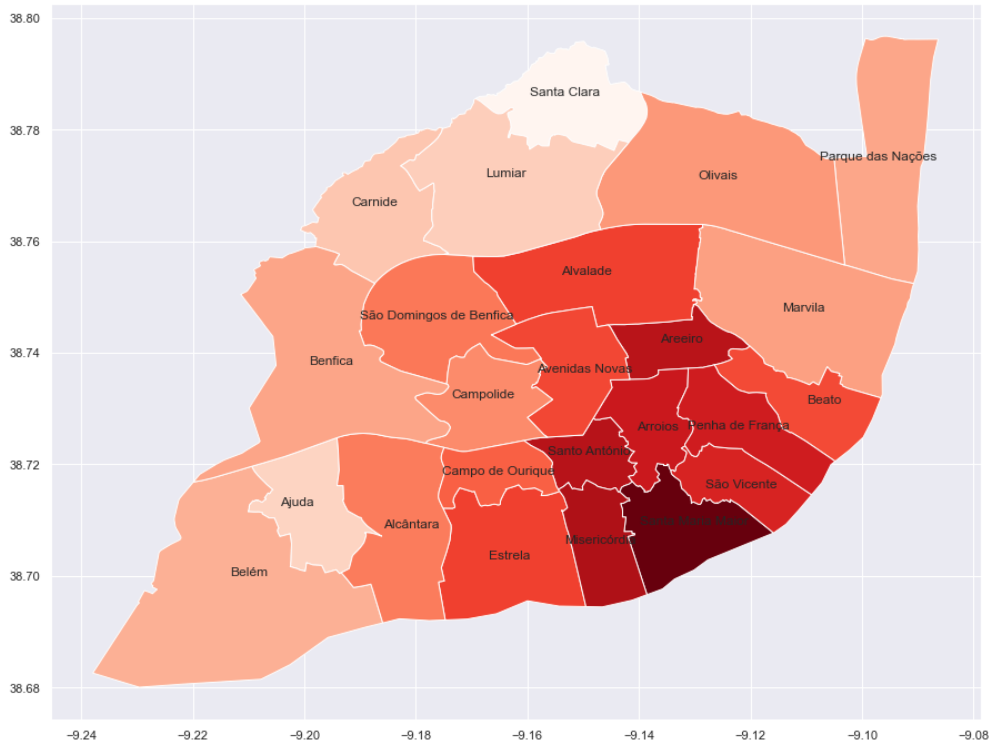


Figure 4-25 - Mean BRD for new parishes. Higher values with darker shades.

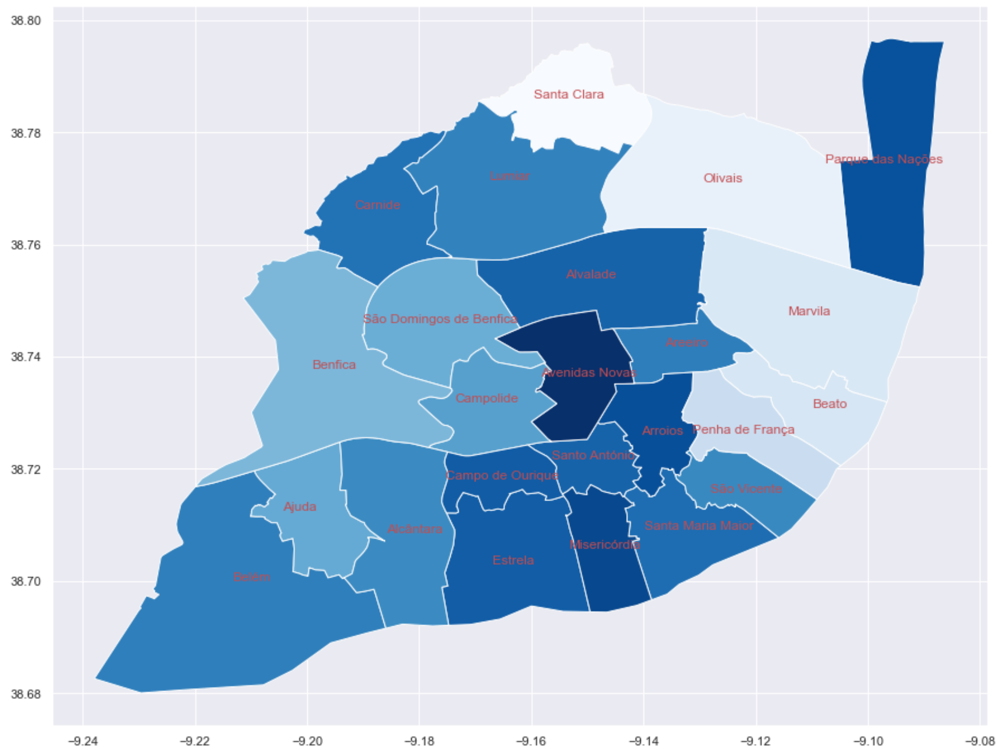


Figure 4-26 - Mean normalized property prices. Higher values with darker shades.

One should note that, while being the sixth parish with the lowest mean BRD value, it was expected *Parque das Nações* to have an even lower BRD, given that most of the buildings in the area were built since the late 90s. However, after a closer look, it is noticeable that the properties in the dataset are actually located near the neighbourhood of *Moscavide* (Figure 4-27), a working-class, older neighbourhood of *Loures* municipality, which explains the higher BRD.

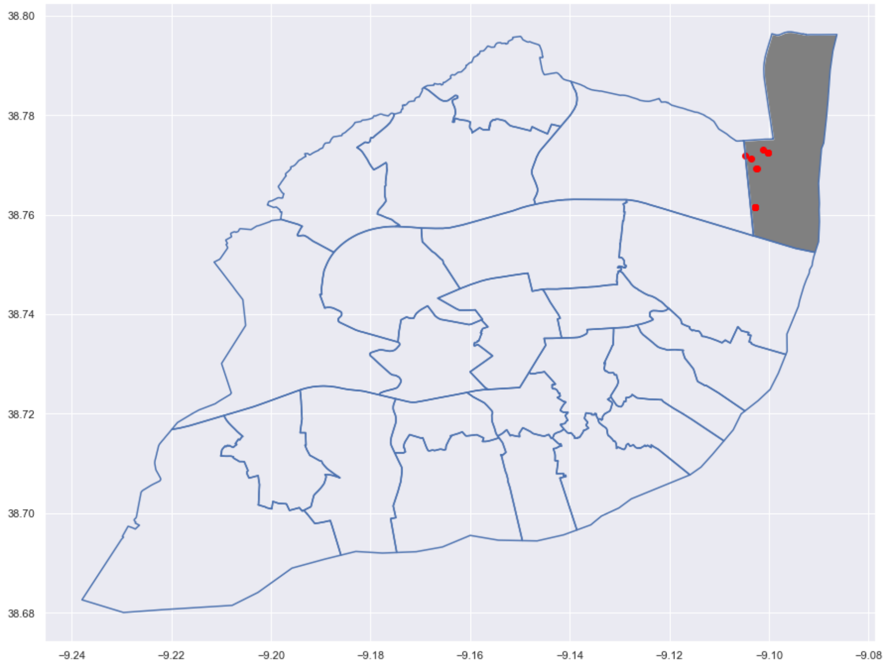


Figure 4-27 - Location of properties in Parque das Nações

4.2.2. Principal Component Analysis (PCA)

As part of the EDA, a PCA analysis was carried out to identify the variables which contributed the most to explaining variance in the dataset. This started with the analysis of the cumulative explained variance as a function of the number of components to determine the recommended number (see Figure 4-28). A total of 23 components have been identified as the required number to achieve 95% of explained variance.

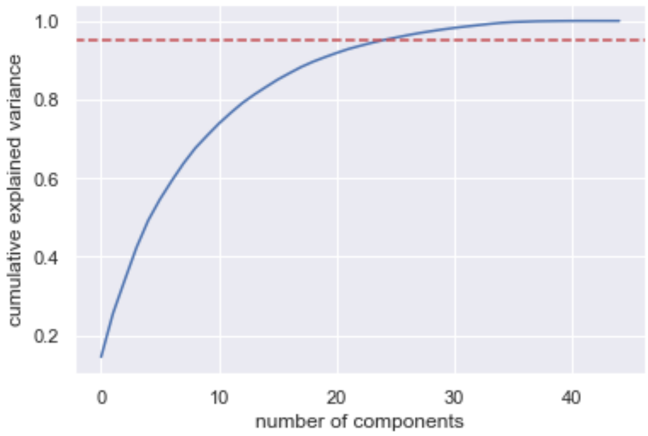


Figure 4-28 - Cumulative explained variance by number of principal components

However, the first 5 components already provided a cumulative explained variance of 72.37%, so it was decided to focus on these. The next step was to analyze the components loading or weights, which represents the elements of the eigenvector, and translate the correlation between the original variables and the principal components (vectors). Table 4-5 and Table 4-6 synthesize the variables with the most weight on each of the first five components. Fig A.0-2, in Annex A, shows the loadings' full correlation matrix plot.

Table 4-5 - Loadings on first three principal components

PC1		**PC2**		**PC3**	
variable	weight	variable	weight	variable	weight
VULInund	0.195	APARTMENT1	0.502	VULSismo	0.460
VUACTUAL_MAF_FINAL	0.192	VULInund	0.456	POI_Rio	0.437
BRD_AFASTADO	0.188	BUILDING_2	0.390	y	0.354
BRD_PROXIMO	0.187	VULSismo	0.333	VULInund	0.330
x	0.181	Culture500	0.204	POI_Indust	0.245
Culture500	0.170	POI_Cemit	0.175	POI_Estudo	0.188
VULSismo	0.132	x	0.142	POI_Cemit	0.185
POI_Cemit	0.079	ABANDONED_	0.092	POI_Emerg	0.183
FM_MAF	0.060	NORMALIZED	0.079	BUILDING_F	0.110
VULMass	0.028	POI_Escola	0.042	BUYER_Code	0.104

Table 4-6 - Loading on fourth and fifth components

PC4		**PC5**	
variable	weight	variable	weight
VULInund	0.488	POI_CBD	0.136
POI_CBD	0.423	VULInund	0.096
ABANDONED_	0.293	POI_AcRodo	0.088
POI_AcRodo	0.213	APARTMENT1	0.086
POI_Estudo	0.211	VULSismo	0.079
POI_ESPriv	0.144	POI_Estudo	0.054
POI_Metro	0.141	Culture500	0.050

PC4		PC5	
POI_Indust	0.125	POI_ESPriv	0.041
POI_Univer	0.096	POI_Metro	0.034
ESPubl	0.091	POI_Univer	0.029

To clearly interpret the loadings, it is necessary to step back from the property values analysis. PCA focuses on the overall variance of data, not on a specific target variable. Thus, it provides us with a better understanding of the commonalities and differences between observations (properties) and not their specific contribution to their value. Nevertheless, as previously mentioned, a property value can be seen as a function of its endogenous and exogenous features, hence one can still extract value from a non-price-targeted analysis.

Looking at the loading in each component, it is difficult to separate them as unique, given the significant number of variables simultaneously being identified as important in several components. Nevertheless, the first component could be identified as being related to the proximity to the river and the centrality, the second factor as related to the condition of the dwelling and building itself, and the remaining three components to the amenities nearby and the centrality of location.

It is noteworthy that some variables are present in several principal components: VULSismo (5), VULInund (4), POI_Estudo (3) are the most frequently referred. By looking at the spatial distribution of these variables, we see that their location is the feature underlying the importance of these components. For example, the city downtown and the riverside locations have high vulnerability to earthquakes (VULSismo) and flooding (VULInund). These hints that there is a similarity between housing stock in these locations. While the distance to libraries (POI_Estudo) reveals, once more, the importance of centrality (see Figure 4-29)[173].

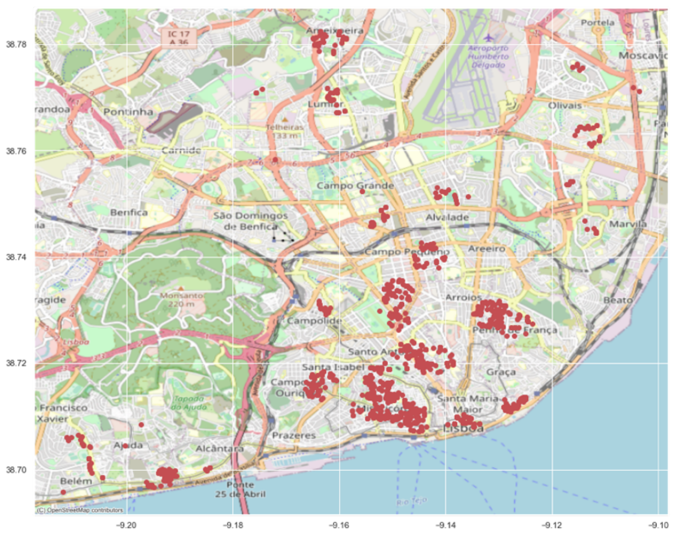


Figure 4-29 - Properties less than 300m from a library (POI_ESTUDO<300m)..

[173] The mean distance to a library in the dataset is 640m, and the median is 485m

4.2.3. Feature Importance

As explained in section 3.8.2, the XGBoost feature importance module was also used to classify the importance of each feature in the dataset as predictors. It is translated by the F-Score which, simply put, means the number of times a feature is used to split the data across all trees. These results are based on a XGBoost regressor with 100 trees and a leaning rate of 0.08. The author started by analyzing the results for Lisbon's whole city, then drilled down to the parish level (Fig. A.0-3 to Fig A.0-25) to understand the differences in preferences. Figure 4-30 shows the result for Lisbon.

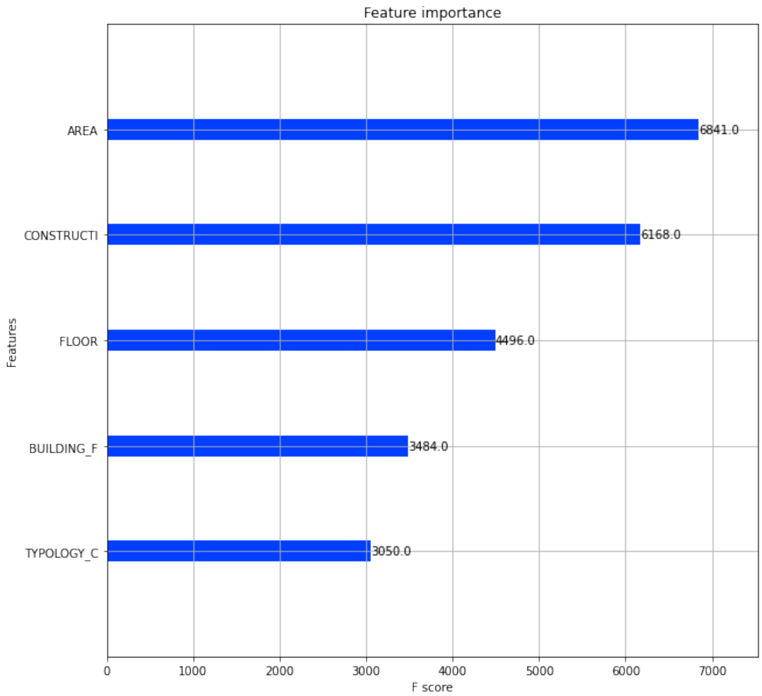


Figure 4-30 - Feature Importance F-Scores for Lisbon municipality

The results are according to the initial expectations and general market knowledge. The area (AREA) is the feature with the most impact on the valuation of a property. The construction year (CONSTRUCTI) is the second most important feature. It is no surprise as newer apartments are often preferred over older ones, due to increased comfort and available amenities. The third feature is the floor number (FLOOR). As previously mentioned, upper floors usually have a higher valuation because they are not subject to the same street level noises as the lower floors, are exposed to less pollution, and often offer better views. However, this rationale is slightly different in buildings with no elevator, given that the effort of climbing an extra flight of stairs is reflected in the price of the upper floors. This feature is followed by the number of floors in the building (BUILDING_F), which is closely related with the previous two variables. Newer buildings are usually taller than older ones, which means that a higher total number of floors is probably synonymous with a newer building, offering more comfort and amenities. The number of rooms (TYPOLOGY_C) is the last feature on the list, which is often referred to as a decisive factor in the home buying process.

Drilling down to the parish level, we notice that the same features keep showing as the most important (area, floor number, construction year and the number of rooms), with a few exceptions. For instance, in *Marvila*, the longitude is the most important feature as it determines the proximity to the wealthy neighbourhood of *Parque das Nações*. In *Campolide*, the longitude is also one of the most important features, defining the proximity to *Amoreiras* (*Santo António*) and *Avenidas Novas*.

4.2.4. Best Subset Selection

This section will be detailing the results for the best subset selection of features. As explained in section 3.9.3.1.1, this method compares multiple models with different predictors, leading to the selection of the best model under the criterion of minimum RSS. The first model consists of a simple example with only one predictor (Figure 4-31). A total of 34 models were tried, and the building condition (BUILDING_2) was selected as the best explanatory variable to explain the normalized price of an apartment. According to the results, the building's condition explains 78% of the variance in the model. The positive sign of the coefficient is also expected, given how the condition scale is defined (1 - bad to 3 - good). It should be noted that, contrarily to the method of PCA used in the previous section, this is a supervised method, which measures the success/importance of a variable by how good of a predictor it is given a target variable.

```
Processed 34 models on 1 predictors in 0.12030887603759766 seconds.
                                 OLS Regression Results
=======================================================================================
Dep. Variable:             NORMALIZED   R-squared (uncentered):                   0.780
Model:                            OLS   Adj. R-squared (uncentered):              0.780
Method:                 Least Squares   F-statistic:                          2.725e+04
Date:                Tue, 06 Sep 2022   Prob (F-statistic):                        0.00
Time:                        13:14:17   Log-Likelihood:                         -6112.3
No. Observations:                7702   AIC:                                  1.223e+04
Df Residuals:                    7701   BIC:                                  1.223e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
BUILDING_2     0.4165      0.003    165.073      0.000       0.412       0.421
==============================================================================
Omnibus:                     2760.173   Durbin-Watson:                   1.606
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            16227.514
Skew:                           1.602   Prob(JB):                         0.00
Kurtosis:                       9.348   Cond. No.                         1.00
==============================================================================
```

Figure 4-31 - Results for 1 predictor

Using two predictors, a total of 561 models were compared (Figure 4-32). The model has increased its adjusted R^2 by 1.8%, and the features selected were the parish code (PARISH_COD) and the apartment condition (APARTMENT1). This means that, for this model, the location of an apartment and its condition are the most important features to be considered for price estimation. It is possible to argue that the condition of the building and the condition of the apartment should be closely linked though not the same. A highly maintained building would suggest care of its inhabitants towards it, which in turn should suggest care for their own apartments. Once again, we see the importance of condition as a predictor of property value.

```
Processed 561 models on 2 predictors in 4.037552118301392 seconds.
                                 OLS Regression Results
=======================================================================================
Dep. Variable:             NORMALIZED   R-squared (uncentered):                   0.798
Model:                            OLS   Adj. R-squared (uncentered):              0.798
Method:                 Least Squares   F-statistic:                          1.524e+04
Date:                Tue, 06 Sep 2022   Prob (F-statistic):                        0.00
Time:                        13:14:25   Log-Likelihood:                         -5771.8
No. Observations:                7702   AIC:                                  1.155e+04
Df Residuals:                    7700   BIC:                                  1.156e+04
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
PARISH_Cod     0.0219      0.001     27.886      0.000       0.020       0.023
APARTMENT1     0.2801      0.005     57.007      0.000       0.270       0.290
==============================================================================
Omnibus:                     2897.880   Durbin-Watson:                   1.570
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            19344.010
Skew:                           1.650   Prob(JB):                         0.00
Kurtosis:                      10.028   Cond. No.                         13.5
==============================================================================
```

Figure 4-32 - Results for 2 predictors

The last set of models tested included three predictors (Figure 4-33). It was possible to compare 5984 models to find that the parish code (PARISH_COD), the building condition (BUILDING_2) and the apartment's condition (APARTMENT1) were the most important features in price estimation. Note that the adjusted R squared only increased by 0.4% from the previous two predictor model.

```
Processed 5984 models on 3 predictors in 39.997554302215576 seconds.
                                 OLS Regression Results
=======================================================================================
Dep. Variable:             NORMALIZED   R-squared (uncentered):                   0.803
Model:                            OLS   Adj. R-squared (uncentered):              0.802
Method:                 Least Squares   F-statistic:                          1.043e+04
Date:                Tue, 05 Apr 2022   Prob (F-statistic):                        0.00
Time:                        12:05:02   Log-Likelihood:                         -5690.2
No. Observations:                7702   AIC:                                  1.139e+04
Df Residuals:                    7699   BIC:                                  1.141e+04
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
PARISH_Cod     0.0190      0.001     23.355      0.000       0.017       0.021
BUILDING_2     0.1450      0.011     12.844      0.000       0.123       0.167
APARTMENT1     0.1602      0.011     15.224      0.000       0.140       0.181
==============================================================================
Omnibus:                     3053.075   Durbin-Watson:                   1.577
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            21649.834
Skew:                           1.734   Prob(JB):                         0.00
Kurtosis:                      10.446   Cond. No.                         41.5
==============================================================================
```

Figure 4-33 - Results for 3 predictors

These results hint that the market values a property's location and condition over its vulnerability to seismic activity or any other natural hazard, which is consistent with the previous findings. Nevertheless, these results should be taken with a grain of salt. Figure 4-34 shows the residuals plot for the last model with three predictors.

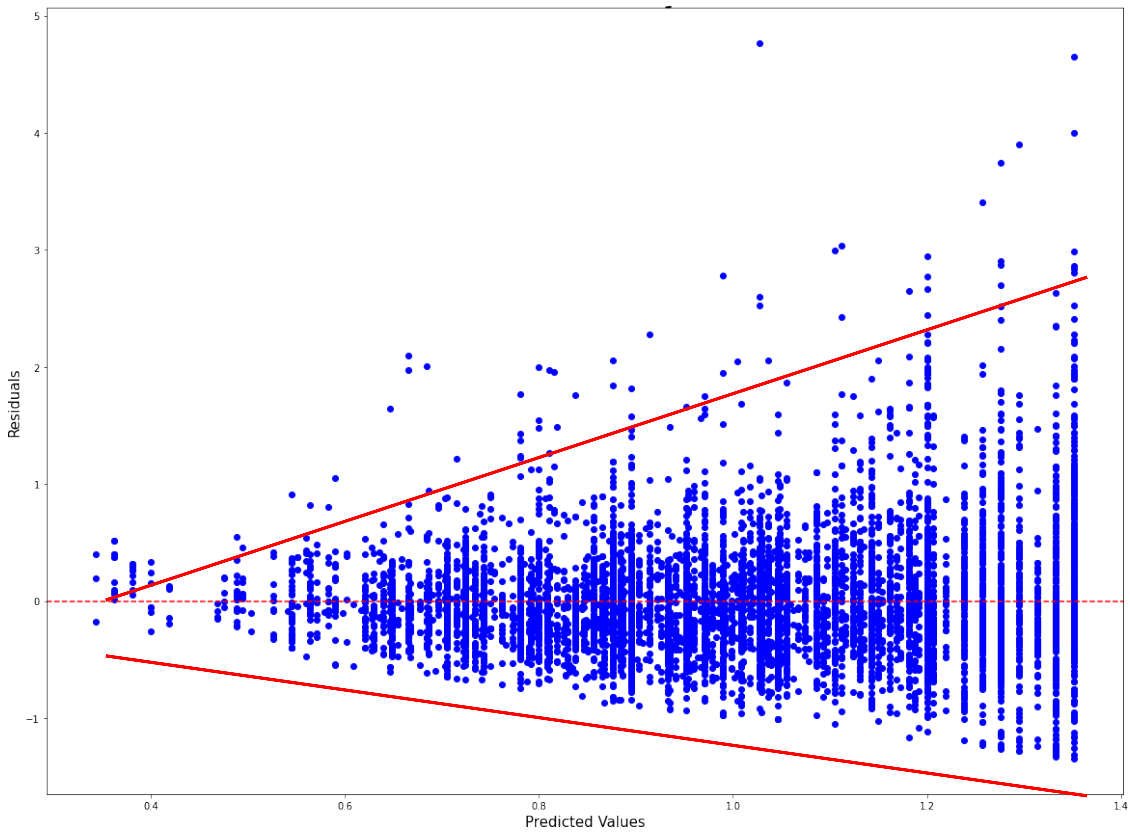


Figure 4-34 - Residuals plot for OLS from Best Subset Selection

Observing Figure 4-34 above[174], it is clear that the variance of residuals is not constant, thus it does not comply with the Markov-Gauss assumption of homoscedasticity of residuals, which means that, even if the estimated regression coefficients are unbiased and consistent, they are not efficient and therefore not BLUE. The same pattern is observed in the regressions using one and two predictors. This is expected given the importance of location in the real estate market price dynamics, as explained in section 3.4.

4.2.5. Stepwise Regression

The stepwise regression consists of a regression algorithm which adds (forward stepwise) or removes (backward stepwise) variables in a model based on their ability to minimize RSS, as explained in section 3.9.3. Nevertheless, one should note that both regressions were used not as final estimation models but as tools to narrow the list of possible predictors to be used in the final models. As so, the residuals will be only analyzed through plots, and no heteroscedasticity test, such as the Breusch-Pagan test, will be applied.

[174] Residuals plot for OLS regression obtained with the Best Subset selection method. Two trend lines are drawn in red.

The forward stepwise model was built first and the number of variables in the model restricted to 7 (Figure 4-35). This restriction is based on the goal of this section to find the most important features contributing to price formation.

```
                                 OLS Regression Results
=======================================================================================
Dep. Variable:             NORMALIZED   R-squared (uncentered):                   0.810
Model:                            OLS   Adj. R-squared (uncentered):              0.810
Method:                 Least Squares   F-statistic:                              4696.
Date:                Tue, 06 Sep 2022   Prob (F-statistic):                        0.00
Time:                        16:11:32   Log-Likelihood:                         -5535.0
No. Observations:                7702   AIC:                                  1.108e+04
Df Residuals:                    7695   BIC:                                  1.113e+04
Df Model:                           7
Covariance Type:            nonrobust
================================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
BUILDING_2     0.1367      0.012     11.612      0.000       0.114       0.160
PARISH_Cod     0.0200      0.001     22.176      0.000       0.018       0.022
APARTMENT1     0.1742      0.010     16.693      0.000       0.154       0.195
POI_ESPriv    -0.0002   1.42e-05    -10.829      0.000      -0.000      -0.000
BRD_AFASTADO  -0.2682      0.023    -11.760      0.000      -0.313      -0.223
ABANDONED_     0.0540      0.008      6.751      0.000       0.038       0.070
Park50M        0.0014      0.000      6.145      0.000       0.001       0.002
==============================================================================
Omnibus:                     2929.337   Durbin-Watson:                   1.592
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            20741.115
Skew:                           1.651   Prob(JB):                         0.00
Kurtosis:                      10.330   Cond. No.                     2.99e+03
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[3] The condition number is large, 2.99e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure 4-35 - Forward stepwise regression model with 7 variables

The results indicate that the building's condition (BUILDING_2), the location given by the parish code (PARISH_COD), the apartment's condition (APARTMENT1), the distance to private high schools (POI_ESPriv), the building resistance deficit calculated to a type 2 earthquake (BRD_AFASTADO), the state of abandonment (ABANDONED_) and the number of parking spots within 50m (Park50M) are the most significant variables, all with a p-value of less than 5%. Once again, the condition of the building and apartment itself are shown to be one of the most important features to be taking into pricing a property. All signs of coefficients are according to expected expect for the state of abandonment. Given that it is classified as (1) abandoned and (2) not abandoned, it was expected that the coefficient should present a positive sign, as non-abandoned properties should be of higher value than the abandoned ones. This happens because occupied ones should, in theory, have better maintenance than those left to abandonment. The negative sign on the coefficient related to the BRD should be interpreted with caution, although in theory it was expected to be negative. Given the preliminary results during the EDA, there is no clear evidence of earthquake risk awareness by real estate stakeholders. Thus, it is possible to hypothesize that the negative value on this variable is caused by the high negative correlation between the BRD and the construction year (-0.77), and the building condition (-0.42). This means that newer and better-maintained buildings are more likely to present a lower BRD. Hence, it is plausible to consider that the appearance of this variable in the model, and its negative sign, translates not a direct preference for mode earthquake resistance buildings, but for newer and well-maintained ones, which is consistent with the results obtained with the previous feature selection techniques.

The residuals reveal heteroscedasticity of errors, as shown in Figure 4-36, although showing an improvement from prior models obtained by Best Feature Selection methods, displaying a lower coefficient of variation.

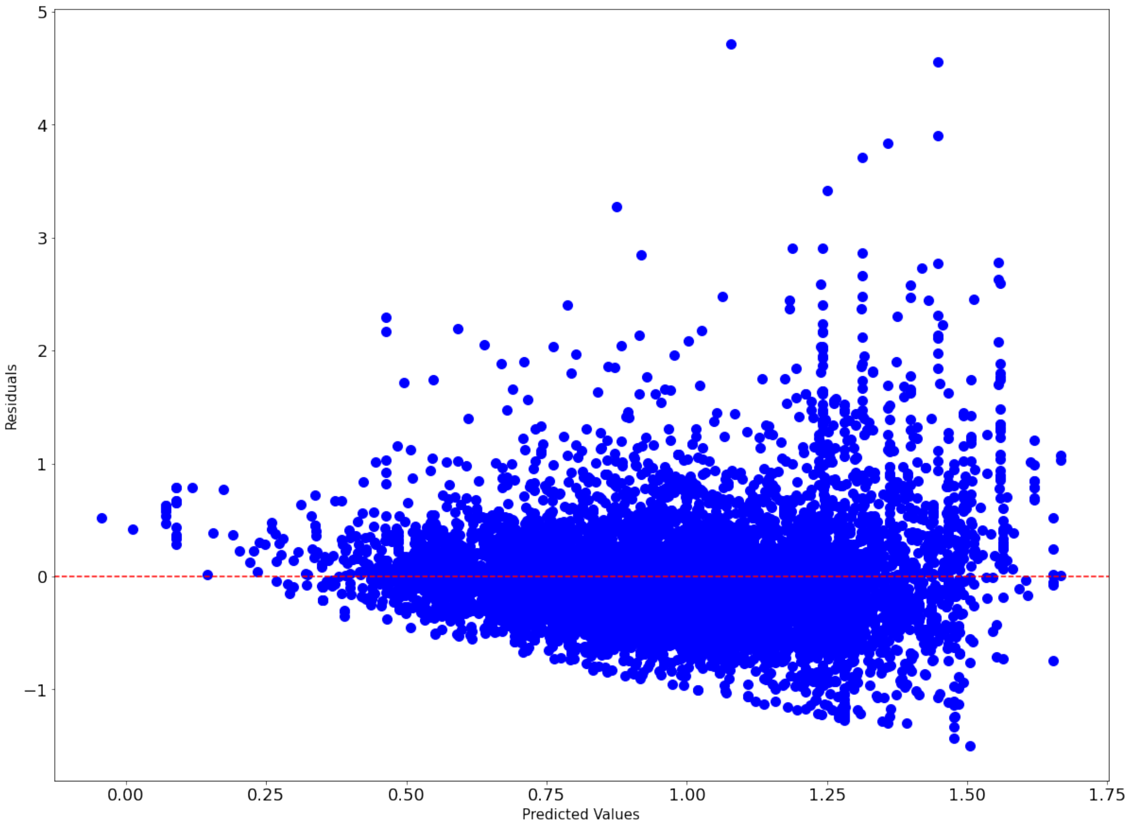


Figure 4-36 - Residuals plot from forward stepwise regression

As mentioned in section 3.9.3.1.1, the backward stepwise regression is preferable over the forward stepwise and should be applied whenever possible. Thus, the analysis was followed by building a backward stepwise regression model. The same rationale of the forward stepwise model was followed, selecting only a reduced number of seven variables to be included in the model. Figure 4-37 shows the final results.

```
                             OLS Regression Results
=======================================================================================
Dep. Variable:            NORMALIZED   R-squared (uncentered):                   0.811
Model:                           OLS   Adj. R-squared (uncentered):              0.810
Method:                Least Squares   F-statistic:                              4706.
Date:               Wed, 07 Sep 2022   Prob (F-statistic):                        0.00
Time:                       10:59:19   Log-Likelihood:                         -5529.0
No. Observations:               7702   AIC:                                 1.107e+04
Df Residuals:                   7695   BIC:                                 1.112e+04
Df Model:                          7
Covariance Type:           nonrobust
===============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------
PARISH_Cod     0.0193      0.001     20.501      0.000       0.017       0.021
BUILDING_2     0.1712      0.012     14.823      0.000       0.149       0.194
APARTMENT1     0.1891      0.010     18.028      0.000       0.169       0.210
POI_CBD    -4.088e-05   5.31e-06     -7.703      0.000   -5.13e-05   -3.05e-05
POI_AcRodo     0.0002   2.23e-05      8.873      0.000       0.000       0.000
POI_ESPriv    -0.0001   1.58e-05     -8.294      0.000      -0.000      -0.000
BRD_PROXIMO   -0.1976      0.020     -9.904      0.000      -0.237      -0.158
==============================================================================
Omnibus:                     2695.610   Durbin-Watson:                   1.591
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            17918.504
Skew:                           1.517   Prob(JB):                         0.00
Kurtosis:                       9.829   Cond. No.                     8.60e+03
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[3] The condition number is large, 8.6e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure 4-37 - Backward stepwise regression model with 7 variables

The results are quite similar to the ones from the forward stepwise regression model. Comparing to the previous model, the R^2 only improved by 0.1%. The location of the property, given by the parish code and the condition of both the building and the apartment remain some of the most important features in predicting property values. The BRD calculated to a type 1 earthquake (BRD_PROXIMO) also appears to be as one of the determining factors in this model, however the same rationale applied in the previous regression can be applied to this one. The BRD for the two different types of seismic hazards are very alike, showing a correlation of 0.99. Thus, as expected, its relationship with the building condition (-0.42) and construction year (-0.76) is very similar. The distance to the CBD (POI_CBD), to the nearest highway (POI_AcRodo) and private high schools (POI_ESPriv) are also selected by the model as predictors.

Although the residuals plot (Figure 4-38) shows signs of heteroskedasticity, it should be noted that the backward stepwise regression presented less variance of residuals than the forward version. This is clear by analyzing the coefficients of variation (CoV) of both regressions, with the former presenting a lower CoV than the latter.

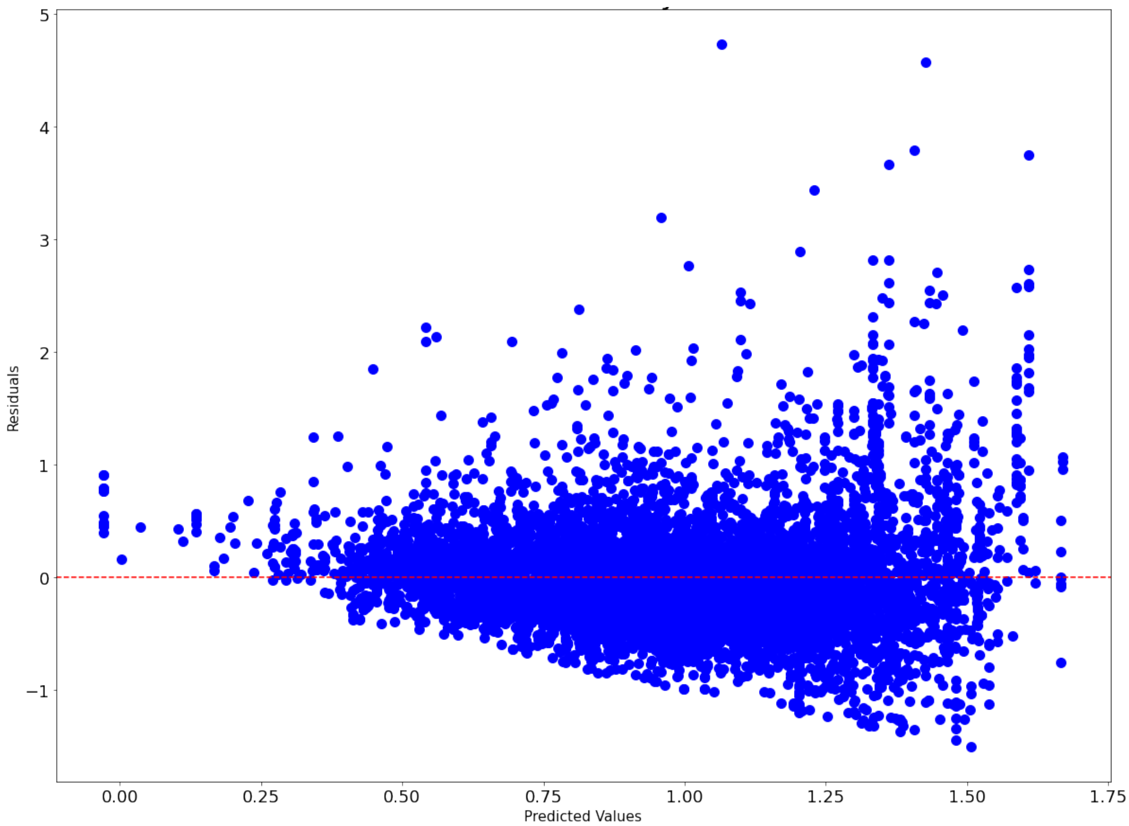


Figure 4-38 - Residuals plot from backward stepwise regression

4.3. R1: Revealed Preferences Regression

This section will detail the final regression models built for the revealed preferences (RP). It will cover PooledOLS regression, fixed-effects regression models, multivariate adaptive regression splines (MARS) and extreme gradient boosting trees (XGBoost).

As mentioned in section 3.4, in order to minimize the possible effects of MAUP, the analysis is developed on three different levels:

1. Raw data with no aggregation and no information on location.
2. Aggregation of data using the old parishes of Lisbon (prior to 2013).
3. Aggregation of data using the current parishes of Lisbon.

Based on the results of the EDA (section 4.1), it was possible to conclude that the following variables should be considered in the model: 1) building's condition (BULDING2); 2) apartment's condition (APARTMENT1); 3) location based on the parish code (PARISH_COD); 4) area (AREA); and 5) the number of floors (FLOOR). Furthermore, given the purpose of this work, the BRD for type II earthquake (BRD_PROXIMO) is also considered.

The analysis starts by modelling the raw data using a FGLS regression and a PooledOLS regression models. Given that information based on location is to be excluded, the parish code will not be taken into account in these two approaches.

4.3.1. Feasible Generalized Least Squares

In this first approach, the author will be using each observation as independent. Heteroscedasticity will be controlled for by using heteroskedasticity consistent estimators (HC0) for covariance. The results are as follows[175] (Figure 4-39):

```
                            GLS Regression Results
==============================================================================
Dep. Variable:             NORMALIZED   R-squared:                       0.142
Model:                            GLS   Adj. R-squared:                  0.142
Method:                 Least Squares   F-statistic:                     258.2
Date:                Wed, 07 Sep 2022   Prob (F-statistic):         6.81e-256
Time:                        16:23:37   Log-Likelihood:                -5925.2
No. Observations:                7702   AIC:                         1.186e+04
Df Residuals:                    7696   BIC:                         1.190e+04
Df Model:                           5
Covariance Type:                  HC0
===============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
-------------------------------------------------------------------------------
const          0.1241      0.035      3.503      0.000       0.055       0.194
BUILDING_2     0.1773      0.013     13.340      0.000       0.151       0.203
APARTMENT1     0.1735      0.009     18.797      0.000       0.155       0.192
AREA           0.0001      0.000      0.866      0.386      -0.000       0.000
FLOOR          0.0143      0.003      5.224      0.000       0.009       0.020
BRD_PROXIMO   -0.0487      0.027     -1.774      0.076      -0.102       0.005
==============================================================================
Omnibus:                     3013.238   Durbin-Watson:                   1.594
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            19782.704
Skew:                           1.734   Prob(JB):                         0.00
Kurtosis:                      10.044   Cond. No.                         775.
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity robust (HC0)
```

Figure 4-39 - Results for Feasible Generalized Least Squares

This model presents an adjusted R^2 of 14.2%. All signs are according to the expected, with the building condition (BUILDING_2), the apartment's condition (APARTMENT1), the area (AREA) and floor number (FLOOR) contributing positively to increase the property value. However, it is noticeable that the AREA variable presents a p-value larger than 5%, which means there is no statistical evidence that this coefficient is truly different from zero (not null). On the opposite side, there is a negative contribution of the BRD, which would be according to the initial expectations, however, we should note that, given its p-value of over 5%, it is not possible to reject the null hypothesis of a null coefficient.

Even after using heteroskedasticity consistent estimators, to control for heteroskedasticity, it was not possible to eliminate it completely from the residuals (see Figure 4-40). But this is not unexpected and it is intrinsic to the data. The most expensive houses per unit area attract a type of clients with lower

[175] Note that, despite the identification of entities and time periods, none were identified during the regression, making it a PooledOLS regression model

budget restrictions, allowing them to make decisions considering a wider range of factors, some very specific and personal (e.g., location close to family members) that are not captured in the database.

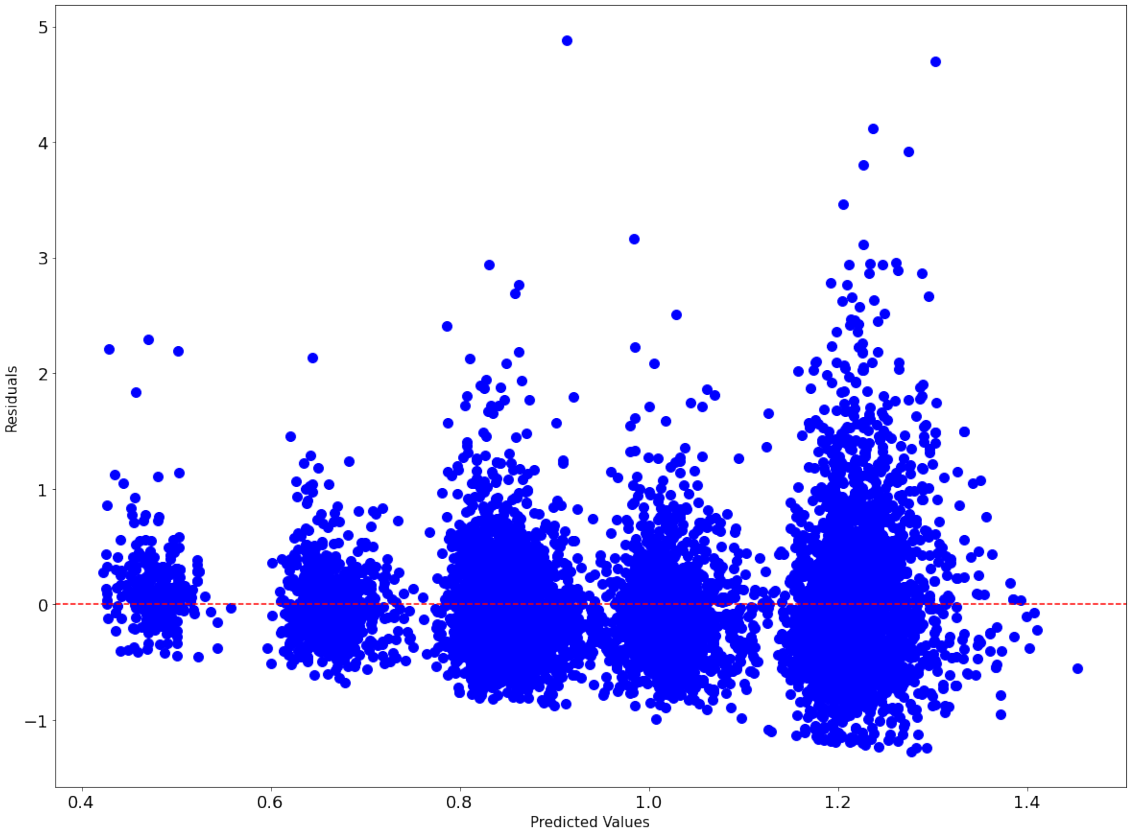


Figure 4-40 - Residuals plot resulting from FGLS regression

It is noticeable that the residuals variance is not constant across the price spectrum (i.e. "funnel errors") which may result from misspecification of the heteroscedasticity type. This means that it may result from factors not being accounted for in this model. Thus, although asymptotically consistent, it is not possible to guarantee the coefficient generated by FGLS to be unbiased, hence not BLUE. According to the literature, the potential for misspecification of the form of heteroskedasticity is pointed as one of the reasons limiting the use of FGLS in applied research (Miller & Startz, 2018). This could be one of these cases.

4.3.2. PooledOLS

This second model controls for heteroscedasticity by using White's covariance estimator[176]. The results are as presented in Figure 4-41[177].

[176] Heteroskedastic robust covariance matrix
[177] Note that, despite the identification of entities and time periods, none were identified during the regression, making it a PooledOLS regression model

PooledOLS Estimation Summary

Dep. Variable:	NORMALIZED	R-squared:	0.1175
Estimator:	PooledOLS	R-squared (Between):	0.1967
No. Observations:	522	R-squared (Within):	0.1027
Date:	Wed, Sep 07 2022	R-squared (Overall):	0.1175
Time:	15:48:32	Log-likelihood	-172.51
Cov. Estimator:	Robust		
		F-statistic:	13.739
Entities:	53	P-value	0.0000
Avg Obs:	9.8491	Distribution:	F(5,516)
Min Obs:	2.0000		
Max Obs:	11.000	F-statistic (robust):	10.887
		P-value	0.0000
Time periods:	11	Distribution:	F(5,516)
Avg Obs:	47.455		
Min Obs:	42.000		
Max Obs:	51.000		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	0.1564	0.1359	1.1508	0.2503	-0.1106	0.4235
BUILDING_2	0.1881	0.0446	4.2187	0.0000	0.1005	0.2756
APARTMENT1	0.1015	0.0331	3.0676	0.0023	0.0365	0.1664
AREA	-0.0001	0.0006	-0.1792	0.8579	-0.0013	0.0011
FLOOR	0.0326	0.0135	2.4083	0.0164	0.0060	0.0592
BRD_PROXIMO	0.1814	0.1263	1.4359	0.1516	-0.0668	0.4295

Figure 4-41 - Results from PooledOLS regression with robust estimation of covariance matrix

Looking at the results, the negative coefficient of the AREA variable is not surprising. It is well known that a larger apartment would be more valuable than a smaller one, all other features remaining equal, on an absolute value basis[178], but less valuable on a square metre basis. However, it is noteworthy that the p-value on this variable is significantly larger than 5%, standing at 85%. In the case of the variable concerning the BRD for a type I seismic action a negative coefficient was expected, signaling that the less vulnerable a property was to earthquakes, the more valuable it would be. However, its p-value also is larger than 5%, standing at 15%, which means there is no statistical significance on this coefficient. In other words, it is not possible to prove that this explanatory variable affects the dependent variable. The residuals analysis presented better results than the previous FGLS model, as per Figure 4-42.

[178] Not necessarily in a price per square metre basis

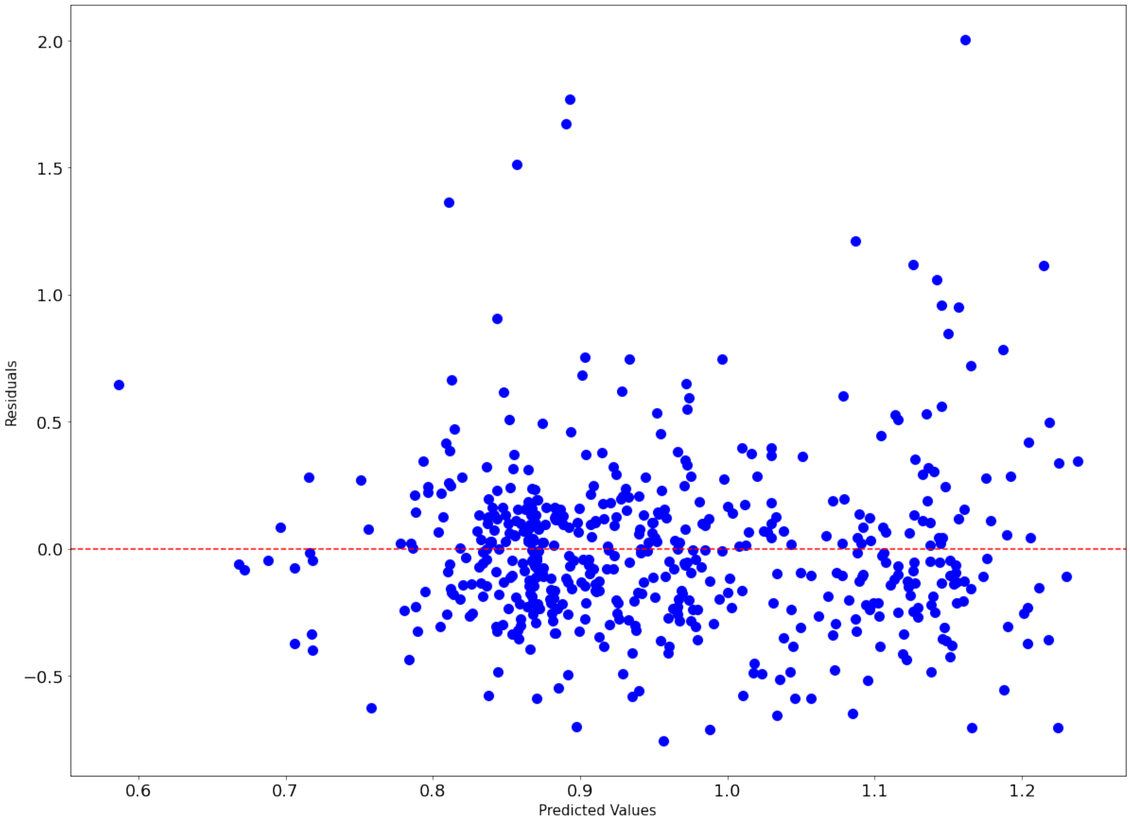


Figure 4-42 - Residuals plot resulting from PooledOLS regression

Despite some visual evidence of “funnel errors”, the author decided to carry out a White- and a Breusch-Pagan-test. Both test the null hypothesis that “homoscedasticity is present” against the alternative hypothesis that “homoscedasticity is not present” (i.e. heteroscedasticity exists). The White-test output was a test statistic of $X^2 = 91.68$ and a p-value of 1.82×10^{-13}, while the Breusch-Pagan-test yielded a test statistic of $X^2 = 22.78$ and a p-value of 1×10^{-4}. Both tests presented a p-value significantly less than 5%, rejecting the null hypothesis of homoscedasticity.

4.3.3. Fixed-Effects

This section will describe the results of applying fixed-effects regression. The analysis starts by the smallest feasible area of real estate data in the database, the old parishes. The results are shown in Figure 4-43.

The results show a R^2 of 18.86%, and all variables present the expected signs, except the area (AREA) variable. However, it should be noted that neither the floor number (FLOOR) nor the area (AREA) present p-values of less than 5%, hence should be discarded from the analysis, given the absence of statistical proof of their impact on the dependent variable.

```
                         PanelOLS Estimation Summary
================================================================================
Dep. Variable:             NORMALIZED   R-squared:                        0.1886
Estimator:                   PanelOLS   R-squared (Between):             -0.4769
No. Observations:                 522   R-squared (Within):               0.1886
Date:                Wed, Sep 07 2022   R-squared (Overall):             -0.0156
Time:                        17:33:44   Log-likelihood                   -65.261
Cov. Estimator:                Robust
                                        F-statistic:                      21.566
Entities:                          53   P-value                           0.0000
Avg Obs:                       9.8491   Distribution:                   F(5,464)
Min Obs:                       2.0000
Max Obs:                       11.000   F-statistic (robust):             11.809
                                        P-value                           0.0000
Time periods:                      11   Distribution:                   F(5,464)
Avg Obs:                       47.455
Min Obs:                       42.000
Max Obs:                       51.000

                             Parameter Estimates
===============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------
const           0.5976     0.1284     4.6542     0.0000      0.3453      0.8500
BUILDING_2      0.1593     0.0425     3.7498     0.0002      0.0758      0.2427
APARTMENT1      0.0780     0.0321     2.4255     0.0157      0.0148      0.1411
AREA           -0.0009     0.0007    -1.2500     0.2119     -0.0022      0.0005
FLOOR           0.0225     0.0125     1.8023     0.0722     -0.0020      0.0471
BRD_PROXIMO    -0.4365     0.1343    -3.2493     0.0012     -0.7005     -0.1725
===============================================================================

F-test for Poolability: 4.5349
P-value: 0.0000
Distribution: F(52,464)

Included effects: Entity
```

Figure 4-43 - Results of the first FE model for old parishes

A second model was built using the number of floors in the building instead of the apartment's area and floor number. The results show R^2 increased by 0.27%, as per Figure 4-44. All signs are according to the expected, and all variables achieved a p-value of less than 5%.

```
                         PanelOLS Estimation Summary
================================================================================
Dep. Variable:             NORMALIZED   R-squared:                        0.1913
Estimator:                   PanelOLS   R-squared (Between):             -0.5156
No. Observations:                 522   R-squared (Within):               0.1913
Date:                Wed, Sep 07 2022   R-squared (Overall):             -0.0307
Time:                        18:43:34   Log-likelihood                   -64.378
Cov. Estimator:                Robust
                                        F-statistic:                      27.502
Entities:                          53   P-value                           0.0000
Avg Obs:                       9.8491   Distribution:                   F(4,465)
Min Obs:                       2.0000
Max Obs:                       11.000   F-statistic (robust):             13.832
                                        P-value                           0.0000
Time periods:                      11   Distribution:                   F(4,465)
Avg Obs:                       47.455
Min Obs:                       42.000
Max Obs:                       51.000

                             Parameter Estimates
===============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------
const           0.5149     0.1215     4.2370     0.0000      0.2761      0.7538
BUILDING_2      0.1559     0.0419     3.7189     0.0002      0.0735      0.2383
APARTMENT1      0.0768     0.0323     2.3757     0.0179      0.0133      0.1404
BRD_PROXIMO    -0.4273     0.1310    -3.2633     0.0012     -0.6847     -0.1700
BUILDING_F      0.0144     0.0048     3.0073     0.0028      0.0050      0.0237
===============================================================================

F-test for Poolability: 4.7010
P-value: 0.0000
Distribution: F(52,465)
```

Figure 4-44 - Results of the second FE model for old parishes

In order to test for spatial autocorrelation of the error term, the author calculated the average error of each parish and plotted its spatial distribution (see Figure 4-45). The parishes of *Santos-o-Velho*, *Pena* and *São Vicente de Fora* presented the larger average errors, while *Sacramento*, *São Nicolau* and *Castelo* reported the smaller error terms.

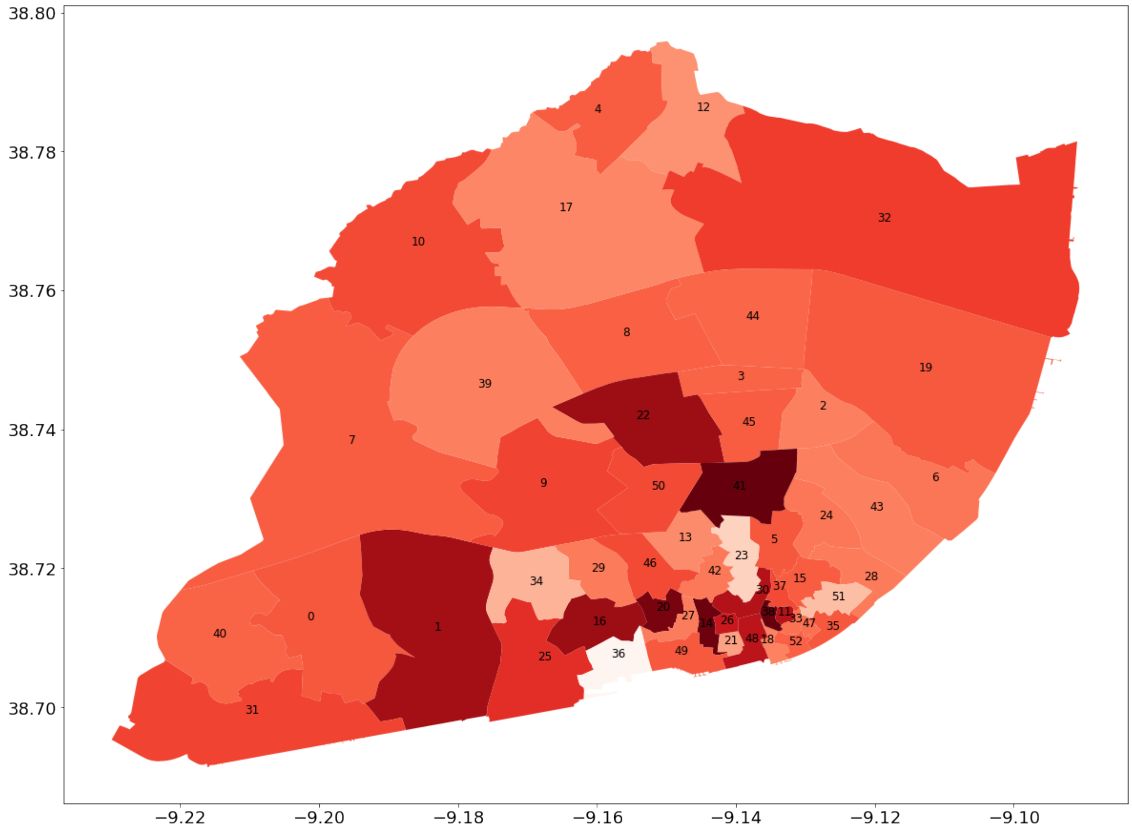


Figure 4-45 - Spatial distribution of average errors (Fixed-Effects model, old parishes)[179]

A Global Moran-I to the regression residuals was carried out and obtained a test statistic of -0.037 and a p-value of 0.463, which means that it failed to reject the null hypothesis at 5% significance level. This means that the spatial distribution of errors does not follow a spatial pattern and might be randomly distributed across space. Thus, one might assume there is no spatial autocorrelation of the error term, proving the validity of the results.

Next, the data was grouped according to the new parishes. The redistribution to 23 parishes led to a larger spatial aggregation in most cases. A model similar to the one firstly built with the old parishes was developed first. The results are shown in Figure 4-46.

[179] Parishes: 00 - Ajuda; 01 – Alcântara; 02 - Alto do Pina; 03 – Alvalade; 04 – Ameixoeira; 05 – Anjos; 06 – Beato; 07 – Benfica; 08 - Campo Grande; 09 – Campolide; 10 – Carnide; 11 – Castelo; 12 – Charneca; 13 - Coração de Jesus; 14 – Encarnação; 15 – Graça; 16 – Lapa; 17 – Lumiar; 18 – Madalena; 19 - Mártires; 20 – Marvila; 21 – Mercês; 22 - Nossa Senhora de Fátima; 23 – Pena; 24 - Penha de França; 25 – Prazeres; 26 – Sacramento; 27 - Santa Catarina; 28 - Santa Engrácia ; 29 - Santa Isabel; 30 - Santa Justa; 31 - Santa Maria de Belém; 32 - Santa Maria dos Olivais; 33 – Santiago; 34 - Santo Condestável; 35 - Santo Estêvão; 36 - Santos-o-Velho; 37 - São Cristóvão e São Lourenço; 38 - São Domingos de Benfica; 39 - São Francisco Xavier; 40 - São João; 41 - São João de Brito; 42 - São João de Deus; 43 - São Jorge de Arroios; 44 - São José; 45 - São Mamede; 46 - São Miguel; 47 - São Nicolau; 48 - São Paulo; 49 - São Sebastião da Pedreira; 50 - São Vicente de Fora; 51 – Sé; 52 - Socorro

```
                        PanelOLS Estimation Summary
================================================================================
Dep. Variable:             NORMALIZED   R-squared:                        0.0849
Estimator:                   PanelOLS   R-squared (Between):             -0.2314
No. Observations:                 238   R-squared (Within):               0.0849
Date:                Wed, Sep 07 2022   R-squared (Overall):             -0.0491
Time:                        22:33:24   Log-likelihood                    73.788
Cov. Estimator:                Robust
                                        F-statistic:                      3.8966
Entities:                          23   P-value                           0.0021
Avg Obs:                       10.348   Distribution:                   F(5,210)
Min Obs:                       7.0000
Max Obs:                       11.000   F-statistic (robust):             3.2844
                                        P-value                           0.0071
Time periods:                      11   Distribution:                   F(5,210)
Avg Obs:                       21.636
Min Obs:                       19.000
Max Obs:                       23.000

                             Parameter Estimates
===============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------
const           0.6622     0.1489     4.4465     0.0000      0.3686      0.9558
BUILDING_2      0.0754     0.0343     2.1998     0.0289      0.0078      0.1430
APARTMENT1      0.0517     0.0272     1.8978     0.0591     -0.0020      0.1053
AREA           -0.0001     0.0009    -0.1197     0.9048     -0.0020      0.0017
FLOOR          -0.0005     0.0124    -0.0416     0.9669     -0.0250      0.0240
BRD_PROXIMO    -0.1771     0.1486    -1.1918     0.2347     -0.4700      0.1158
===============================================================================

F-test for Poolability: 6.4121
P-value: 0.0000
Distribution: F(22,210)

Included effects: Entity
```

Figure 4-46 - Results of the first FE model for new parishes

As expected, the results are worse than the ones for old parish limits. The R^2 dropped from 18.86% to 8.49%, and only one variable was proven to be statistically significant in the model, showing a p-value of less than 5% (BUILDING_2). The larger unit areas encircle vast heterogeneous areas of real estate, with different products and dynamics, leading to worse model performance. A second model was also built, using the same variables as with the old parishes (Figure 4-47).

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:             NORMALIZED   R-squared:                        0.1061
Estimator:                   PanelOLS   R-squared (Between):             -0.3336
No. Observations:                 238   R-squared (Within):               0.1061
Date:                Wed, Sep 07 2022   R-squared (Overall):             -0.0787
Time:                        22:43:35   Log-likelihood                    76.573
Cov. Estimator:                Robust
                                        F-statistic:                      6.2587
Entities:                          23   P-value                           0.0001
Avg Obs:                       10.348   Distribution:                   F(4,211)
Min Obs:                       7.0000
Max Obs:                       11.000   F-statistic (robust):             4.7179
                                        P-value                           0.0011
Time periods:                      11   Distribution:                   F(4,211)
Avg Obs:                       21.636
Min Obs:                       19.000
Max Obs:                       23.000

                             Parameter Estimates
===============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------
const           0.5946     0.1135     5.2382     0.0000      0.3708      0.8183
BUILDING_2      0.0838     0.0345     2.4294     0.0160      0.0158      0.1518
APARTMENT1      0.0498     0.0281     1.7707     0.0781     -0.0056      0.1052
BRD_PROXIMO    -0.2070     0.1440    -1.4374     0.1521     -0.4908      0.0769
BUILDING_F      0.0113     0.0051     2.1901     0.0296      0.0011      0.0214
===============================================================================

F-test for Poolability: 6.9629
P-value: 0.0000
Distribution: F(22,211)

Included effects: Entity
```

Figure 4-47 - Results of the second FE model for new parishes

Although it showed a slight improvement from the previous model, increasing the R^2 from 8.49% to 10.61%, it delivers a substantially poorer performance than the one developed for old parishes. Only the building's condition (BUILDING_2) and the total number of floors (BUILDING_F) are statistically sound as positively impacting the normalized property prices. Figure 4-48 shows the spatial distribution of error.

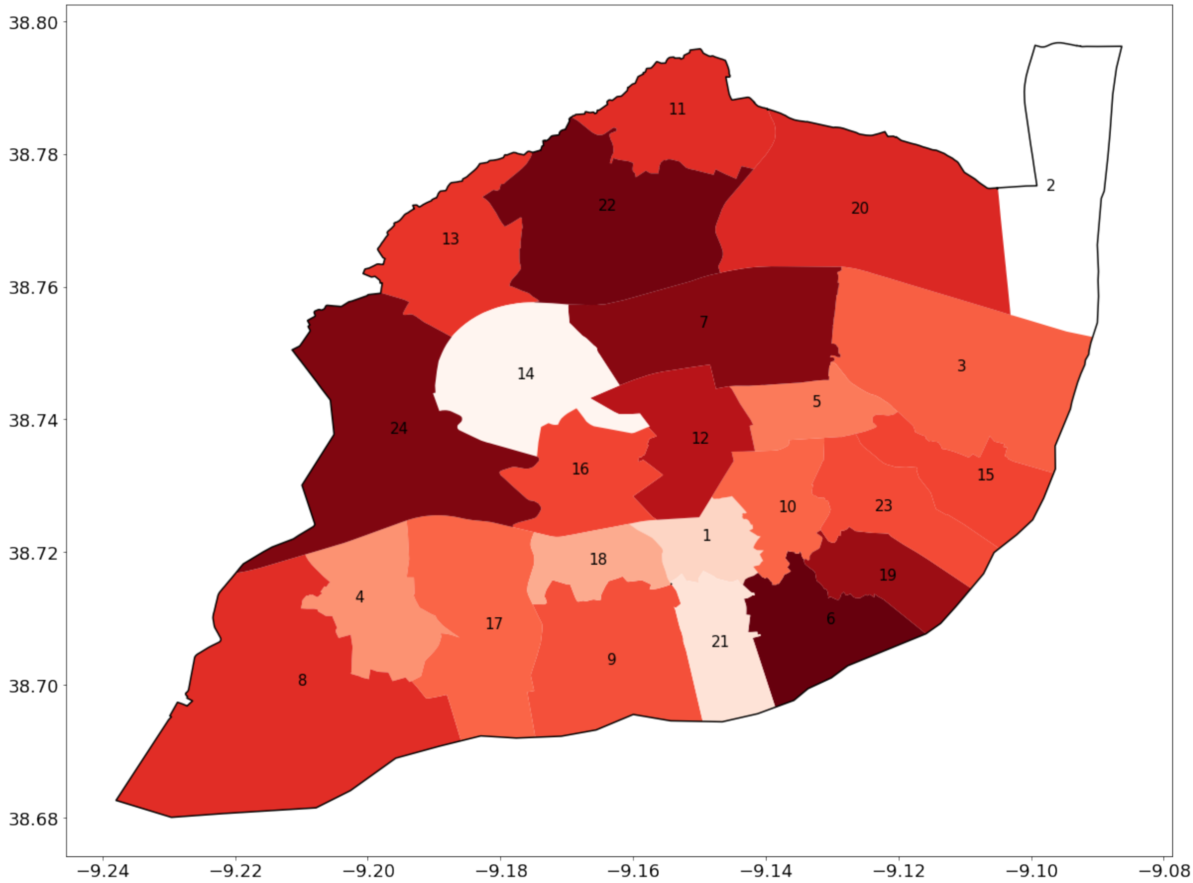


Figure 4-48 - Spatial distribution of average errors (second Fixed-Effects model , new parishes)[180]

In this case, the Global Moran-I test rejected the null hypothesis at 5%, suggesting the presence of spatial autocorrelation of errors.

The next step was to develop models based on the vulnerability values given by the year of construction and the location of the building. A probability function at the level of the statistical subsection was used to estimate the vulnerability given by the class, as described in section 3.7.5. In essence, the models' previous build was replicated while replacing the BRD variable with an updated one with the new vulnerability values (BRD_PROXIMO_POND). It is important to note that it was only possible to estimate vulnerabilities with the second approach for approximately half of the properties in the database. This could result from the misplacement of buildings in the database, errors in the insertion of coordinates or in the attribution of construction years. Hence, in order to have a larger pool of properties to model with, it was assumed the other half (50.66%) to have the same vulnerability levels as calculated using the methodology of Ferreira (2012).

[180] 1 - Santo António; 2 - Parque das Nações; 3 - Marvila; 4 - Ajuda; 5 - Areeiro; 6 - Santa Maria Maior; 7 - Alvalade; 8 - Belém; 9 - Estrela; 10 - Arroios; 11 - Santa Clara; 12 - Avenidas Novas; 13 - Carnide; 14 - São Domingos de Benfica; 15 - Beato; 16 - Campolide; 17 - Alcântara; 18 - Campo de Ourique; 19 - São Vicente; 20 - Olivais; 21 - Misericórdia; 22 - Lumiar; 23 - Penha de França; 24 - Benfica;

Table 4-7 summarizes the overall results. Each column represents a different set (two different sets) of variables included in the model.

Table 4-7 - Regression results for FE models

	BRD - First approach				BRD - second approach			
	Old Parishes		New Parishes		Old Parishes		New Parishes	
const	0.597***	0.519***	0.662***	0.594***	0.475***	0.403***	0.578***	0.484***
BUILDING_2	0.159***	0.155***	0.075**	0.083**	0.151***	0.153***	0.095***	0.108***
APARTMENT1	0.078**	0.076**	0.051*	0.049*	0.112***	0.112***	0.052*	0.052*
AREA	-0.001	--	-0.0001	--	-0.0005	--	9.18e-5	--
FLOOR	0.022*	--	-0.0005	--	0.020	--	-0.012	--
BRD_PROXIMO	-0.436***	-0.427***	-0.177	-0.207	-	--	--	--
BRD_PROXIMO_POND	--	--	--	--	-0.223**	-0.221**	-0.0002	0.002
BUILDING_F	--	0.014***	--	0.011***	--	0.014***	--	0.009*
Entities	53	53	23	23	52	52	23	23
Periods	11	11	11	11	11	11	11	11
R2	18.86%	19.13%	8.49%	10.61%	20.48%	21.39%	9.54%	10.57%
Spatial Autocorrelation of Residuals	No	No	Yes	Yes	No	No	Yes	Yes

***p-value<1%; **p-value<5%; *p-value<10%

The analysis of the results allows concluding that the old parishes units are a better overall approach to model data. This is consistent with the initial expectations as smaller spatial units allow for more homogenous housing stock and real estate market dynamics. It is also possible to conclude that the building's and apartments' condition, as well as the number of floors in the building, positively impact the property value. With a more detailed BRD approach, the second approach seems to yield better results in modelling property prices overall. Despite the significant impact of the BRD variables in the old parishes' models, this should be taken with due care, given the high correlation with the age of the building, which may reveal a preference for newer buildings rather than a direct preference for a lower vulnerability to seismic activity. Additionally, the area appears insignificant in the model, which would be counterintuitive. However, one should note that the price per square metre is being used. Thus, this normalization would already consider the area as it only shows that there is no significant difference in prices per square metre across the number of rooms.

4.3.4. MARS

The next model to be developed was the Multivariate Adaptive Regression Spline (MARS). As explained in section 3.9.1, this model is a good compromise between the explainability of linear regression models and the adaptability to non-linear relationships. Following the previous fixed-effects model, the input variables chosen were: BUILDING_2, APARTMENT1, AREA, FLOOR, BRD_PROXIMO, and

BUILDING_F (Table 4-8). However, BRD_PROXIMO_POND was not used in this case due to the more significant number of missing values. Thus, it was opted out of this model.

Table 4-8 - Results of MARS model

Basis Function	Pruned	Coefficient
(Intercept)	No	-3.255
APARTMENT1	No	0.175
BUILDING_2	No	0.175
h(AREA-49.31)	No	0.005
h(49.31-AREA)	Yes	None
FLOOR	No	0.017
h(AREA-388.12)	No	-0.011
h(388.12-AREA)	No	0.010
h(AREA-74.16)	No	0.006
h(74.16-AREA)	Yes	None
h(AREA-163.22)	No	-0.001
h(163.22-AREA)	Yes	None
BRD_PROXIMO	No	-0.061

All the variables but the AREA were considered significant to the model (not pruned[181]). The AREA variable was excluded from the model for values between 49.31m^2 and 74.16m^2 and between 163.22m^2 and 388.12m^2. This could be the case as it is already accounted for in the normalized variable. For interpretation, note that, in the case of the variables APARTMENT1, its coefficient means that a 1 unit increase in the variables is expected to lead to a 0.175 increase in the response variables (the normalized sale price, in this case). The results show that a better condition of the building and apartment are the most critical variables considered by the market, contributing positively to the price increase (0.175 each). The other variables proved to have a much lower impact on the target variable. This model achieved a Mean Squared Error (MSE) of 0.2669, a R-squared of 16.09% and a Generalized R-Squared[182] (GRSQ) of 15.59%.

[181] Note that "pruning" means to change the model by deleting the child nodes of a branch node

[182] GRSQ is an adjusted version of the R-squared that accounts for the complexity of the model. It penalizes models with more basis functions and provides a more conservative estimate of the model's goodness of fit. Like the R-squared, GRSQ values range from 0 to 100%, where 100% represents a perfect fit.

4.3.5. XGBoost

The last type of model applied to the revealed preferences data was the Extreme Gradient Boosting (XGBoost). Several tunings to the hyperparameters were performed, such as the number of trees and the learning rate, to assess the model's sensitivity and the difference in results, creating four different models. The maximum depth considered was 2.

All four models were built using a sub-sample of 75% as a training set and ten folds[183] for cross-validation. The results are shown in Table 4-9.

Table 4-9 - Results of XGBoost models

Parameters	Regression 1	Regression 2	Regression 3	Regression 4
n_estimators	1000	500	100	100
learning_rate	0.08	0.08	0.08	0.04
sub_sample	75%	75%	75%	75%

Metrics	Regression 1	Regression 2	Regression 3	Regression 4
RMSE	0.398	0.403	0.423	0.443
R2 Score	71.29%	70.38%	66.74%	62.55%

Cross Validation	Regression 1	Regression 2	Regression 3	Regression 4
#Folds	10	10	10	10
Mean cross-validation score	0.48	0.47	0.41	0.36

The R squared and the RMSE were the metrics selected to evaluate the accuracy of the models. The difference in RMSE and R squared across the four models is significant, varying 11.30% and 8.74%, respectively. However, the results obtained hint some takeaways:

1. From model 1 to model 2, the number of estimators was decreased leading to a lower R squared and a higher RMSE, which may hint that the 1000 trees in the first model probably led to overfitting. A lower number of estimators led to poorer results.
2. By reducing the learning rate from model 3 to model 4, both the R-squared value and the RMSE were decreased, which means that the number of trees was probably not enough to improve the results, given that a lower learning rate would require a higher number of steps (or estimators) to converge to an optimal valued.

It is worth mentioning that the parameters defined for XGBoost Regression 1 were used to extract the feature importance, previously mentioned.

[183] A value of k=10 is very common in the field of applied machine learning for K-Fold Cross Validation

Figure 4-49 shows a plot of the predicted and test values. The results are satisfactory, nonetheless the model seems to present a larger error in the higher price segment, leading to a higher dispersion of values.

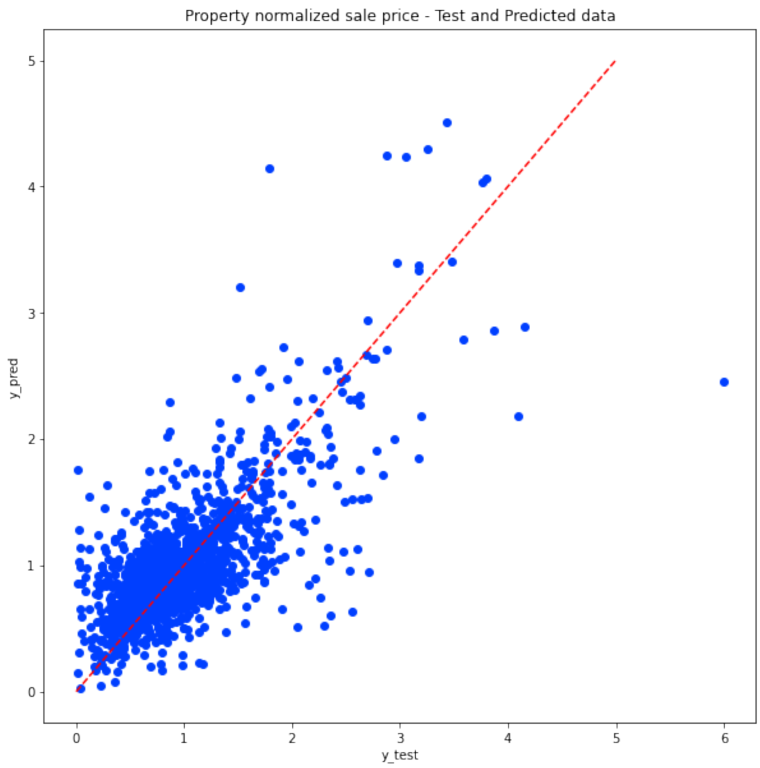


Figure 4-49 - Predictions model

4.4. R2 and R3: Stated Preferences Results

4.4.1. Survey Results

The first phase consisted of a validation stage with 15 respondents for one week, intending to test the survey's understanding. After the test period, it was possible to conclude that no significant issues were found; only minor adjustments were needed (for example, one of the building periods was set as "1775 to 1880" instead of "1755 to 1880", by mistake).

The survey reached 325 respondents for approximately one month (from March 7th to April 11th), of which 75% identified as Lisbon residents. Considering the total respondents, it achieved a margin of error of 5.44% for a 95% confidence interval.

A. <u>Demographics</u>

Most respondents identify as Lisbon municipality residents (74,7%), with the remaining living in its neighboring municipalities within Lisbon's Metropolitan Area, as per Figure 4-50. Approximately 50%

are 2-to-3-person households (Figure 4-51). Thus, the sample's average household size[184] is 3.18, slightly above Lisbon's 2021 census value of 2.2.

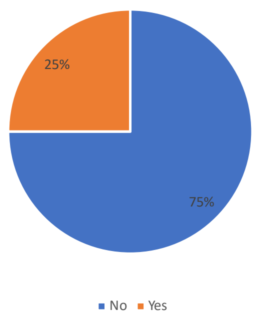


Figure 4-50 - Lisbon residents in the sample (%).

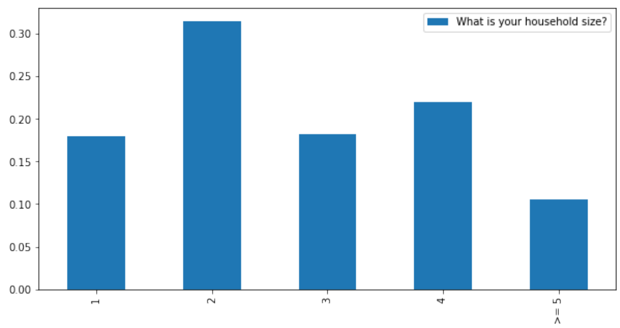


Figure 4-51 - Households by household size (%)

According to the INE (INE, 2023), the gross reported income less personal income paid tax per inhabitant in Lisbon for 2020 was 13.890€. Thus, for a family of 3.18 people, one could expect an annual income of approximately 44.170€ or a monthly income of 3.155€. The sample's average household income is 3.123€ (Figure 4-53), which aligns with initial expectations. However, it should be noted that this corresponds to average values, not to median ones, and that income distributions can be biased. According to the report on the *average monthly gross wage per worker*, published by INE (2023), the

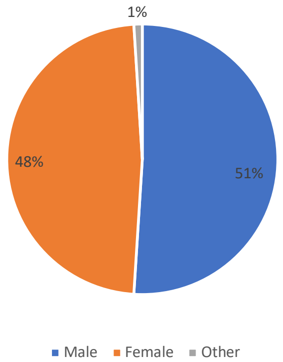


Figure 4-52 – Respondents by gender (%)

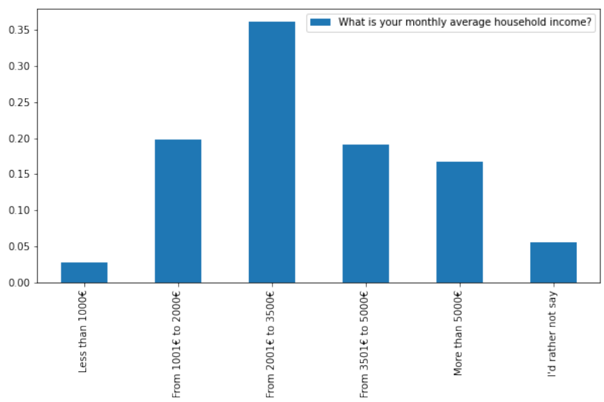


Figure 4-53 – Households by income class (%)

income distribution in the country is positively skewed (or right-skewed), with a high concentration of low wages. Within the respondents, 48% of our respondents identify as female, 51% as male and only 1% as other (Figure 4-52).

Regarding the age of respondents, most are younger than 55 (58%) (Figure 4-54). Compared with the 2021 census, the sample as a slight underrepresentation of the older population and an overrepresentation of the younger class (Figure 4-55). This may result from the chosen survey

[184] The author considered 5 as the mark in the "5 or more" class.

distribution: an online survey reduces the biases in our sample (such as Interviewer and Social Desirability biases), but also excludes older and “info excluded” individuals.

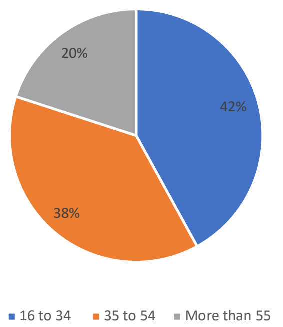


Figure 4-54 – Respondents by age (%)

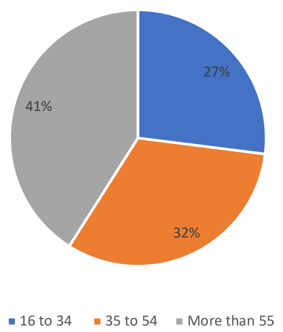


Figure 4-55 –Lisbon residents by age, according to the 2021 census (%). Source: PORDATA

Regarding marital status, 47% of respondents are married, 40% are single, 7% are divorced and 5% claim other statuses (Figure 4-56). In addition, 37% of respondents have children (under 18 years old) living in their dwelling (Figure 4-57).

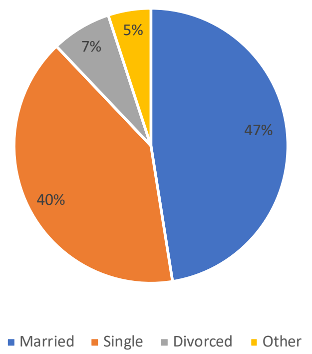


Figure 4-56 – Marital status of respondents (%)

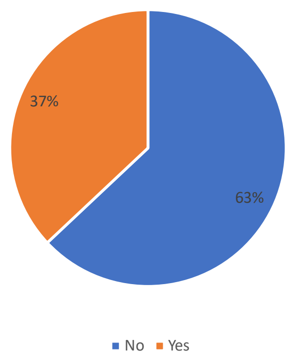


Figure 4-57 – Respondents with children (under 18 years old) living in their dwelling (%)

Most respondents are highly educated: 39,81% hold a bachelor’s degree, 44,13% a master’s degree, 5% hold a PhD, and 0,3% only concluded elementary school (Figure 4-58).

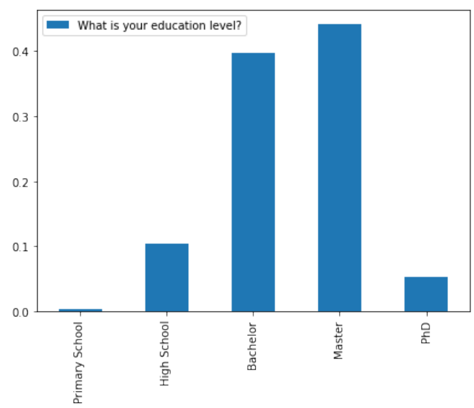


Figure 4-58 - Education level of respondents (%)

Figures Figure 4-59 and Figure 4-60 show the spatial distribution of respondents.

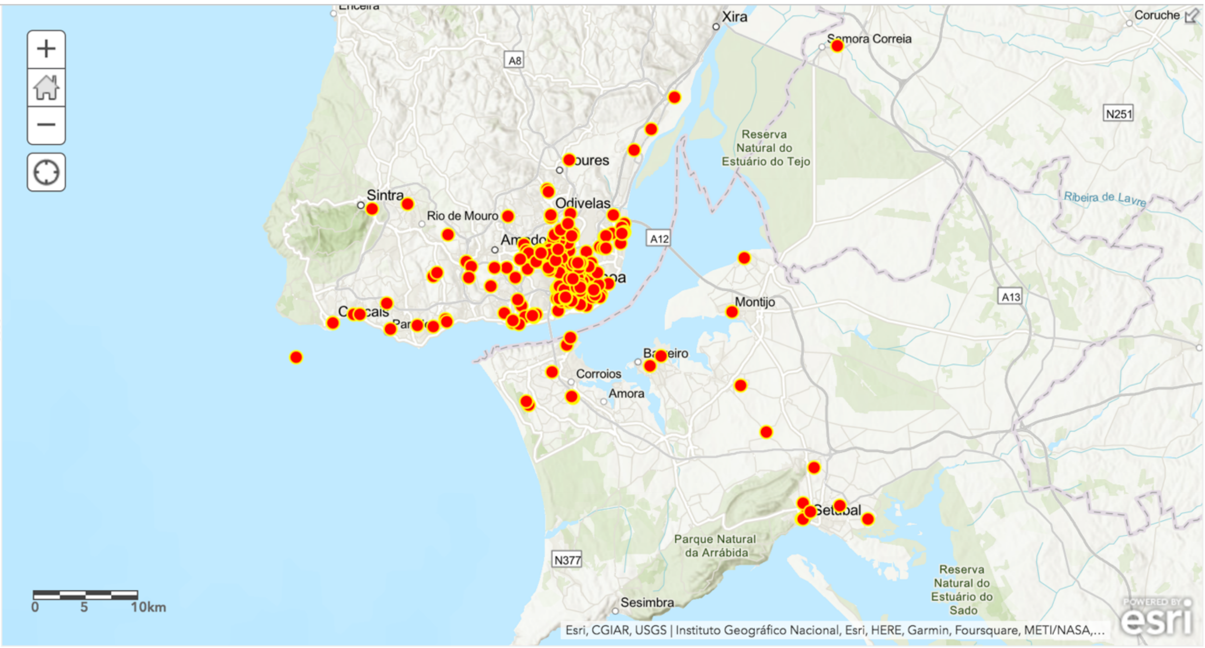


Figure 4-59 - Spatial distribution of respondents across Lisbon's Metropolitan Area.

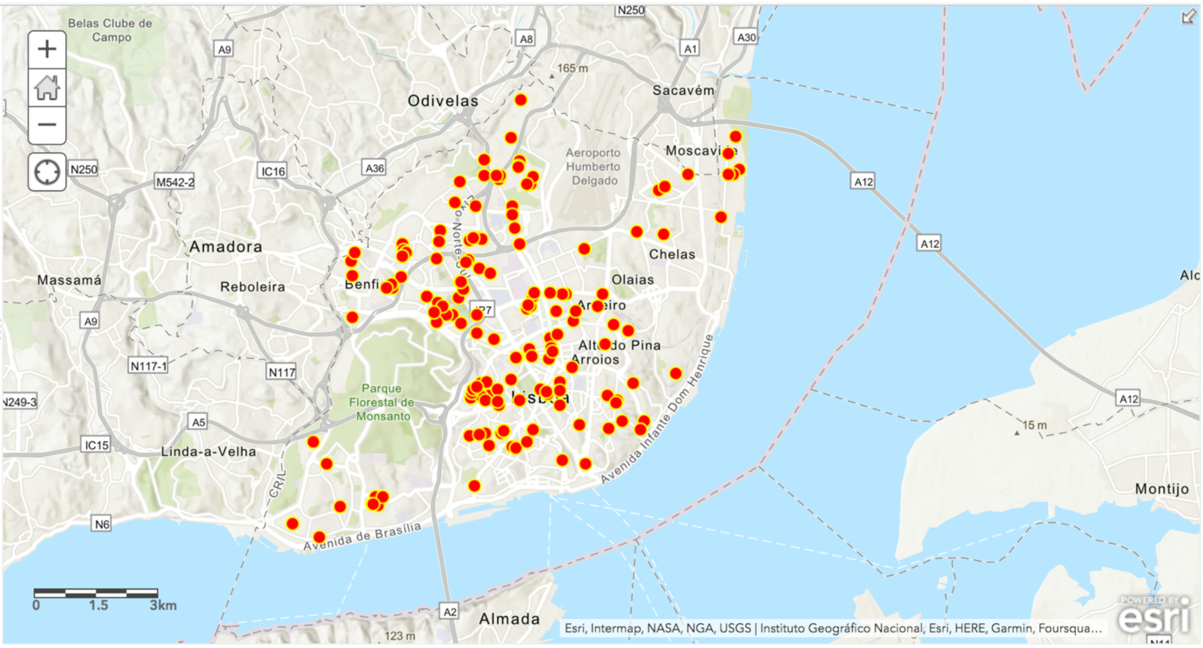


Figure 4-60 - Spatial distribution of respondents in Lisbon's municipality.

Table 4-10 shows the number of inhabitants and population percentage per the 2021 census and compares it with the survey's sample. Notably, it is achieved a slight overrepresentation of parishes such as *Avenidas Novas* and *Campolide* and an underrepresentation of others like *Marvila* and *Arroios*.

Table 4-10 - Analysis of respondents by parish (census vs survey sample)

Parish	No. Inhabitants	%Population	Corresponding Sample Size	Responses	%Sample
Ajuda	14313	2.62%	10	6	2.97%
Alcântara	13852	2.54%	10	1	0.49%
Alvalade	33313	6.10%	23	16	7.92%
Areeiro	21167	3.88%	15	7	3.46%
Arroios	33307	6.10%	23	6	2.97%
Avenidas Novas	23261	4.26%	16	20	9.90%
Beato	12185	2.23%	9	2	0.99%
Belém	16549	3.03%	12	10	4.95%
Benfica	35367	6.48%	25	15	7.42%
Campo de Ourique	22146	4.06%	16	6	2.97%
Campolide	14794	2.71%	10	21	10.39%
Carnide	18029	3.30%	13	9	4.45%
Estrela	20308	3.72%	14	8	3.96%
Lumiar	46338	8.49%	33	24	11.88%
Marvila	35482	6.50%	25	2	0.99%
Misericórdia	9660	1.77%	7	1	0.49%
Olivais	32184	5.90%	23	6	2.97%
Parque das Nações	22382	4.10%	16	6	2.97%
Penha de França	28485	5.22%	20	5	2.47%
Santa Clara	23650	4.33%	17	8	3.96%
Santa Maria Maior	10052	1.84%	7	1	0.49%
Santo António	11062	2.03%	8	5	2.47%
São Domingos de Benfica	34081	6.24%	24	14	6.93%
São Vicente	13956	2.56%	10	3	1.48%

In sum, a significant sample's portion comprises highly educated young adults living in Lisbon and with an average household monthly income of around two to three thousand euros.

B. Housing

The second section of the survey regards housing. Most respondents are homeowners (70%), 26% are private market renters, and only 1% are public market renters (Figure 4-61). According to the Lisbon Municipality (CML 2023), the city has roughly the same proportion of renters and homeowners (120.188 renters and 121.884 resident homeowners). Thus, one can conclude that our sample is slightly biased towards homeownership status. 76% of renters claim to pay a monthly rent below 1000€, and only 7% pay more than 1500€ (Figure 4-62).

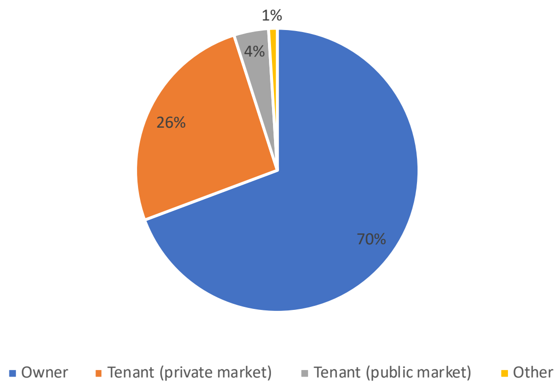


Figure 4-61 - Respondents by housing tenure (%)

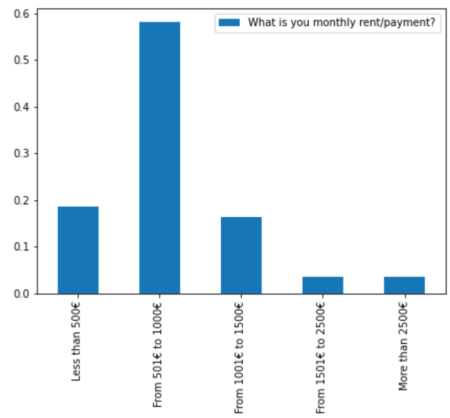


Figure 4-62 - Respondents by monthly rent class (%)

The average estimated property value was 544 690€: 43% of respondents estimate their apartment value to be between 250 000€ and 500 000€, followed by 100 001€ to 250 000€ (20%) and 500 001€ to 750 000€ (18%); 15% estimate it to be over 750 000€ and only 2% estimate it to be less than 100 000€ (Figure 4-63)[185]. Regarding the number of rooms, most respondents live in 2 and 3-bedroom apartments (Figure 4-64).

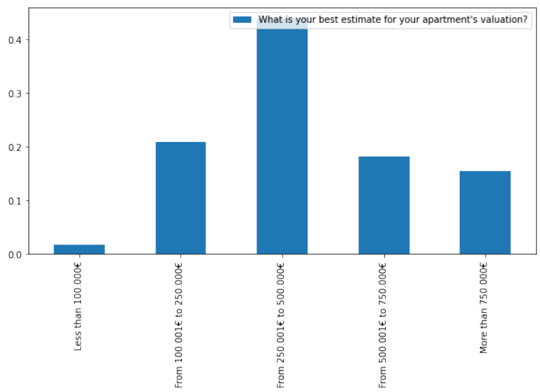


Figure 4-63 - Estimates of property valuation (%)

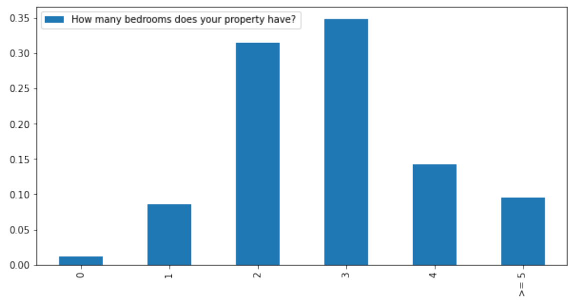


Figure 4-64 - Number of rooms (%)

Regarding the property's area, 71% of respondents estimate it to be between 51 and 150 m^2, which is in line with the median area of 95 m^2 of family accommodation sold in Portugal (INE 2019, p. 252) (Figure 4-65), which gives us an estimated average area of 116,94 m^2. Dividing the average estimated property valuation by the average estimated area, we get 4655€/m^2, which aligns with expectations. For reference, as of March 2023, according to the real estate intelligence company Alfredo AI, the average price for a property in Lisbon is 4469€/m^2. Furthermore, half the respondents live in a building less than six stories high (Figure 4-66).

[185] For comparison, as of March 2023, according to the real estate intelligence company Alfredo AI, the average sale price for Lisbon Municipality stands at 410.000€ or 4469€/m^2 (see https://alfredo.pt/price-index). Note that, since 2014, INE only provides median banking valuation values at the municipal level.

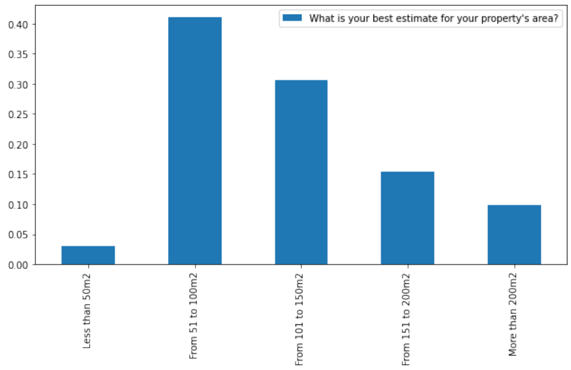


Figure 4-65 - Estimated useful area by class (%)

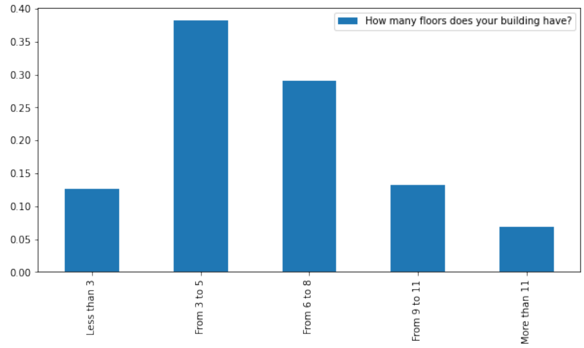


Figure 4-66 - Number of stories in the respondent's building by class (%)

Most respondents (64%) knew the construction year of their buildings. As expected, this was the case with newer buildings. Based on the respondents who knew the construction year of their building, the mean corresponds to 1986 and a the median to 1991 (Figure 4-68). To those who did not, it was asked an estimated year range, which corresponded to older buildings, with most respondents (63%) pointing to an estimated construction year between 1930 and 1985 (Figure 4-67).

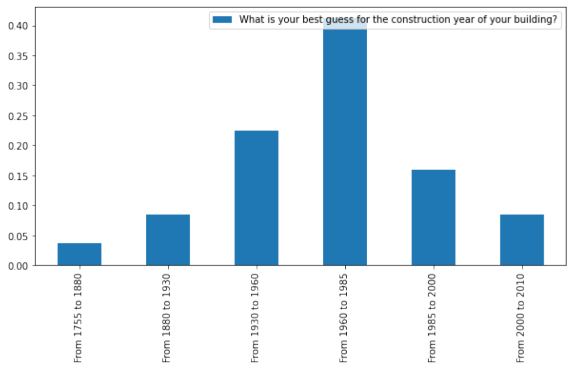


Figure 4-67 - Estimated construction year by class (%)

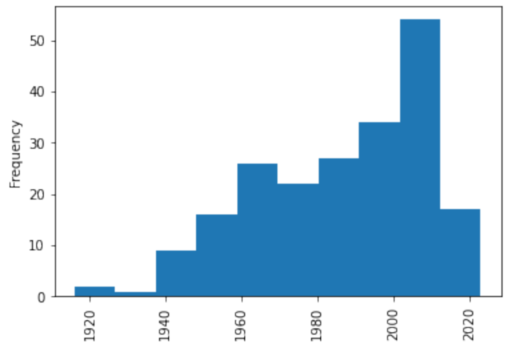


Figure 4-68 - Histogram of construction years

C. <u>Perception and action towards risk</u>

In this section, the perception and action towards risk are under analysis. The first question asked the respondents to order five risks to their dwellings by the level of importance: earthquakes, home robberies, fires, floods and others. A value between 1 and 5 was attributed to each risk, corresponding to scale from a low-concern to a high-concern risk. Earthquakes achieved an average of 3.74, followed by fires (3.64), home robberies (3.21), floods (2.55) and others (1.87) (Figure 4-69). Notably, no significant differences were seen in results across marital status, size and income of the household, education level, age, gender, or the existence of children in the dwelling. However, there is a difference between Lisbon residents and others. The former's top concern is earthquakes, while the latter mentioned home robberies as their top concern. These results should be interpreted in light of the context lived in when the survey was carried out. On February 6th, 2023, a 5.2 magnitude earthquake caused thousands of fatalities in Turkey and Syria, having significant news coverage, which led to a nationwide discussion on the seismic risk of the city of Lisbon. This coverage, however diminishing over time, was still present during March and April as the survey was being carried out[186] and may have led to "salience bias" caused by an out-of-market event.

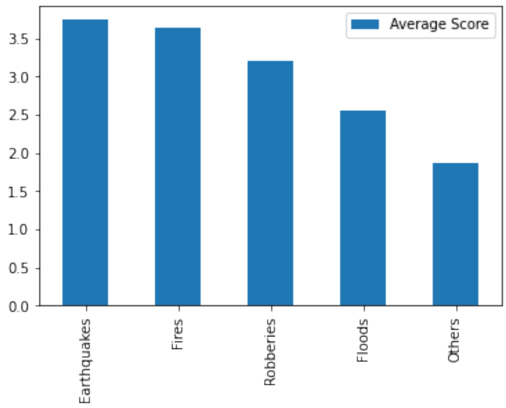


Figure 4-69 - Average importance score by risk type

Figure 4-70 shows the prevailing risk perceived by each parish[187].

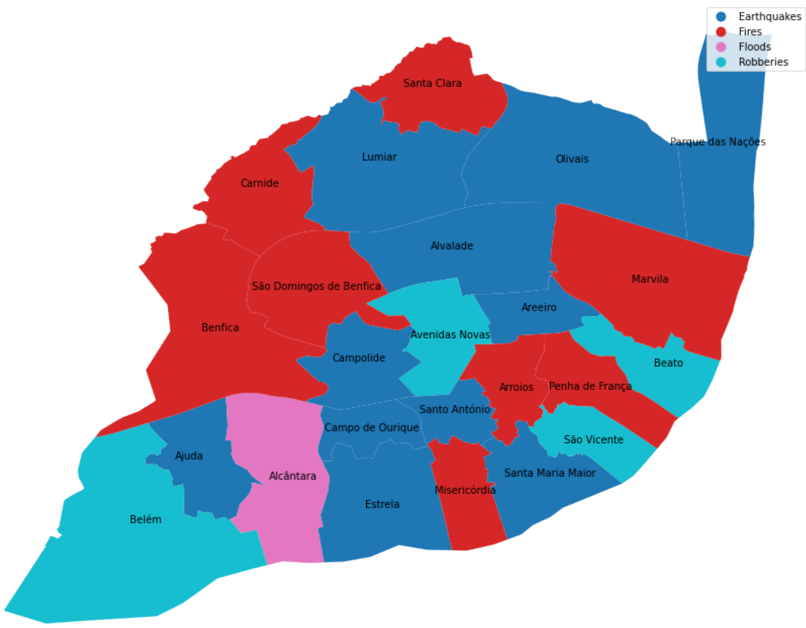


Figure 4-70 - Spatial distribution of the mode of top risk perceived by respondents.[188]

[186] For reference, please see https://observador.pt/2023/03/06/sismo-na-turquia-e-siria-e-o-pior-desastre-na-europa-num-seculo/ and https://expresso.pt/sociedade/2023-04-15-Turquia-dia-1-do-diario-do-epicentro-resiliencia-sismica-d94aac2e

[187] Note that the sample size may not be statistically significant for some parishes. Nevertheless, it is interesting to note that Alcântara, the only parish to have floods as their top concern, was also one of the most affected by the floods of December 2022. Therefore, it may hint at the existence of salience bias. For reference, please visit https://sicnoticias.pt/meteorologia/2022-12-08-Alcantara-e-uma-das-zonas-mais-afetadas-pelas-inundacoes-06458768

[188] Home robberies (dark blue), fires (red), floods (pink) and earthquakes (light blue). Note that, in the face of the small sample size, the results may not be statistically significant. For the entire sample size used for each parish, please refer to Table A.7 12.

Table 4-11 shows a set of 9 statements aiming to evaluate the perception and action towards risk. A colour grade has been used to enhance the visual interpretation of results, highlighting from red to green the percentage of respondents.

The first statement intends to evaluate the necessity to protect their home. The results align with initial expectations, revealing that 83% of respondents agree that their home is their priority after their families' safety. Next, the fatalistic attitude towards different risks was assessed: home robberies and earthquakes. This tried to evaluate their perception of how their attitudes may change their vulnerability to risk. Notably, while fatalism is not significant in home-robberies, opinions are divided regarding earthquakes, with a slight inclination for a fatalistic attitude. This reveals that people may feel it is out of their hands to take any action that significantly decreases their vulnerability. The next phase measured people's expectations towards the city council/civil parish in risk mitigations. Interestingly, most respondents believe that local government institutions can lessen the effects of floods and earthquakes by taking action to minimize risk. The following two statements address the central government and insurance companies. The results hint at a lack of trust towards both types of institutions, revealing people's disbelief in the ability of the government to take care of the response to disaster situations and a distrust of insurance institutions to cover their needs in the aftermath of an earthquake. The last two statements evaluate the trust in building codes and professional practices. In both cases, we can see an unfamiliarity with building codes and practices, with most respondents not agreeing or disagreeing.

Table 4-11 - Statements on risk perception in their original order

Statement	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Following the safety of my family, the safety of my home is my number one priority	0%	4%	13%	56%	27%
If my home is going to be robbed, there isn't much I can do about it	12%	55%	18%	12%	3%
If floods are going to occur, there is not much my city council/civil parish can do to lessen its effects	38%	42%	10%	6%	3%
If an earthquake is going to harm me, it will, and there isn't much that I can do about it	9%	27%	18%	35%	11%
If an earthquake is going to occur, there is not much my city council/civil parish can do to lessen its effects	26%	44%	17%	10%	3%
If an earthquake is going to occur, the government will take care of me	31%	39%	19%	9%	2%
If an earthquake is going to occur, my insurance policy will take care of me	15%	35%	27%	19%	3%
My building is safe against earthquakes	11%	24%	40%	22%	2%
All recent buildings are safe to earthquakes due to restrictive building codes	10%	29%	38%	21%	2%

Crossing this information with the demographic data provides some insights on the differences of perception. As shown in Figure 4-71, people living in older buildings tend to disagree with the statement "My building is safe against earthquakes" showing that, on average, people have a good understanding that older building may have a higher degree of vulnerability to earthquakes. This is especially significant on people who estimate their building construction year to be between 1755 and 1880 and between 1930 and 1985.

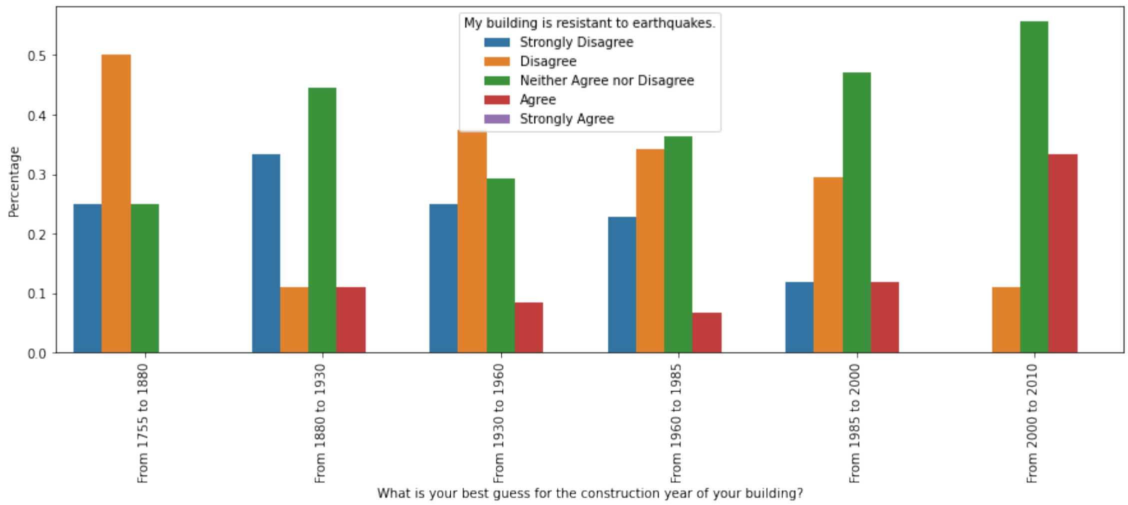


Figure 4-71 - Perception of building vulnerability to earthquake risk by estimate of construction year.[189]

The analysis also reveals a tendency to believe that a more expensive apartment should be less vulnerable to earthquakes. However, this is not always the case. Newer buildings may be subjected to more restrictive codes and have a higher market valuation, but the value (by itself) should translate into less vulnerability (Figure 4-72).

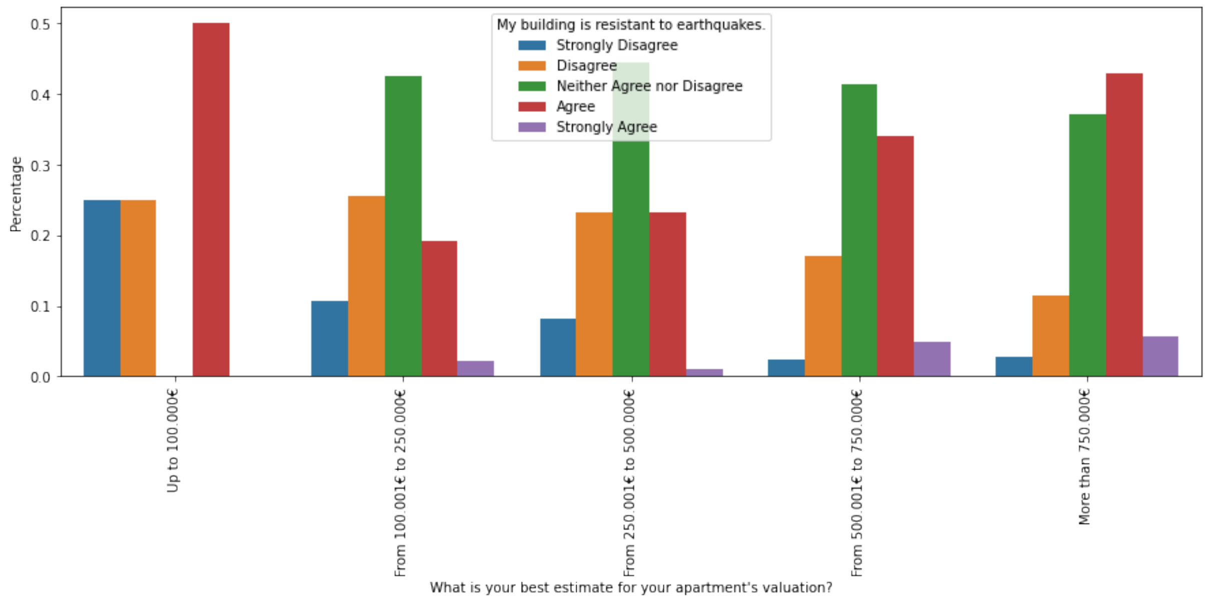


Figure 4-72 - Perception of building vulnerability to earthquake risk grouped by estimate of property market value.[190]

Evaluating the perception of risk with the level of education, reveals that higher educated people tend to be more aware of their buildings' vulnerability. This is especially noticeable from bachelors to master's degree, and from master's to PhD (Figure 4-73).

[189] Percentage of responses: Strongly disagree (dark blue), disagree (orange), neither agree or disagree (green), agree (red) and strongly agree (purple)
[190] Percentage of responses: Strongly disagree (dark blue), disagree (orange), neither agree or disagree (green), agree (red) and strongly agree (purple)

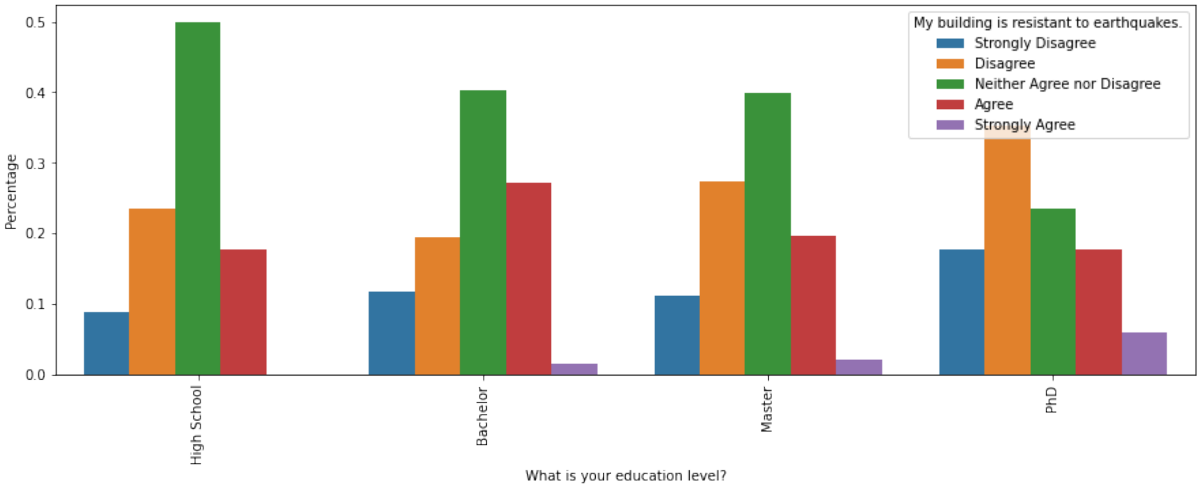


Figure 4-73 - Perception of building vulnerability to earthquake risk grouped by education level.[191]

The analysis of the perceived vulnerability by gender, age, income and marital status revealed no significant differences across groups.

Most respondents have experienced an earthquake (67%), and most homeowners have earthquake insurance (55%) (Figure 4-74 and Figure 4-75, respectively). This significantly exceeds the national average of 16% of earthquake insurance coverage[192]. Notably, respondents who experienced an earthquake are not significantly more willing to buy earthquake insurance (only 55,5% of respondents who experienced an earthquake have insurance against it, a 0,5% increase from people who did not experience such an event). This might be the case due to Lisbon's lower magnitude of earthquakes in recent years. However, some respondents also provided feedback, saying that the extra cost could have been more affordable, leaving informed homeowners without insurance coverage. Others said that the newer restrictive building codes made them feel secure and with no need to pay the extra cost.

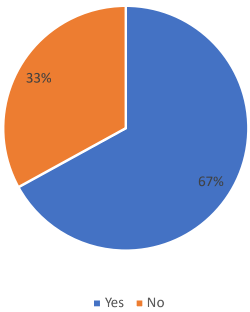


Figure 4-74 - People who experienced an earthquake (%)

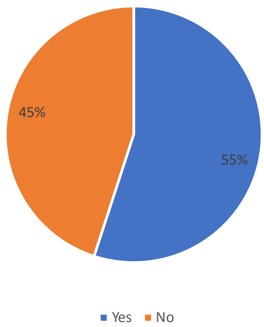


Figure 4-75 - Homeowners with earthquake insurance (%)

[191] Elementary school level was excluded from this analysis due to the small number of respondents. Percentage of responses: Strongly disagree (dark blue), disagree (orange), neither agree or disagree (green), agree (red) and strongly

[192] For reference, see the interview to the President of the Portuguese Association of Insurers (APS – Associação Portuguesa de Seguradores), José Galamba de Oliveira, available at the following link: https://www.sabado.pt/ultima-hora/detalhe/so-16-das-habitacoes-tem-cobertura-de-risco-sismico-que-deve-ser-obrigatoria

Of those paying for earthquake insurance, 46% pay up to 236€ annually, 27% pay between 236€ and 452€, 13% pay between 452€ and 668€ and only 14% pay more than 668€.

Table 4-12 shows some example of proposed mitigation measures, some of which should be followed while others should be avoided. Respondents were asked to choose from the list, those which they considered to be valid mitigation measures. Remarkably, most respondents identified "shelter against an inside wall, or under sturdy furniture such as heavy desks or tables" and "Secure heavy objects to walls and floors, such as shelves, bookcases, cabinets, and water heaters" as valid mitigation measures.

Table 4-12 - Proposed mitigation measures and percentage of respondents

Proposed Mitigation Measures	Percentage of respondents (%)
Secure heavy objects to walls and floors, such as shelves, bookcases, cabinets, and water heaters	63,53%
In case of a seismic event, get into the elevator as soon as possible to exit the building	0,3%
In case of a seismic event, shelter against an inside wall, or under sturdy furniture such as heavy desks or tables	90,58%
In case of a seismic event, shelter against an outside wall, or near a window, so you can see what is going outside	3,04%

D. Willingness to pay for mitigating measures

In this section the willingness to pay of respondents is evaluated. The first two questions evaluated the willingness to pay for higher seismic resistance. The results show that 80% of respondents were willing to pay more for a more resistant building to seismic events (Figure 4-76).

Of those who responded positively, 53% stated they would be willing to pay up to 5% more for a property with a higher seismic resistance, 36% were willing to pay between 5% to 10% more and only 6% would pay more than 15% (Figure 4-77).

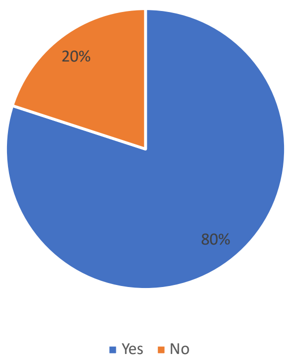


Figure 4-76 - Respondents willing to pay more for building's resilience (%)

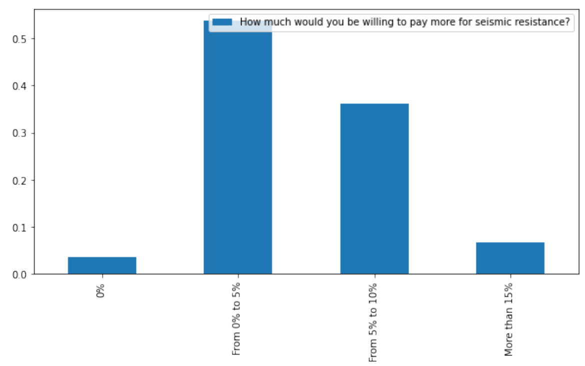


Figure 4-77 - Increase in property price (%)

Of those who were not willing to pay for a more resilient property, 21% claimed they do not worry about seismic risk, 20% say their building is already safe, 3% claimed that it would be a waste of money and the remaining 54% claimed "other" reason. From the feedback received from respondents, some argued that prices were already high enough and that they would not be willing to pay more. Some of the "other" reasons enunciated by respondents were:

- "The decision to buy a more resistant property shouldn't be on the buyers' side; It should be required by the municipality during the licensing process and it should not be a sole responsibility of the design engineer"
- "Prices are already too high"
- "Buildings should already be safe"
- "Resistance should already be granted by the developer"
- "Seismic resistance should be already be accounted for in all new developments and rehabilitation projects"

It is noteworthy that the high prices, and the financial struggles to access housing, was the most common justification. This is consistent with the existing literature (see, for reference, Pothon (2020)). Financial struggles are also reflected on the willingness to accept a discount for a less resistant building (48% agreed to have a discount).

When asked if they would be willing to pay for structural retrofitting of their building, 73% agreed (Figure 4-78). Some of those who disagreed did it due to the fact that their buildings are recent. Regarding the value of this structural retrofitting, when asked as a percentage of the total property price, 57% said they would be willing to pay up to 5% of the property value (for structural retrofitting), 30% would pay between 5% and 10%, and only 10% would be willing to pay more than 10%. When asked for an absolute value they would be willing to pay, respondents provided a median value of 5.000€ and a mean of 10.304€ (Figure 4-79).

As mentioned in the literature review, some authors found differences in risk perception and willingness to pay for mitigation measures across gender, age or educational and income levels. Thus, a cross-analysis was carried out to assess the willingness to pay for structural retrofitting. No significant differences were found across gender, with both male and female respondents showing a 75% willingness to pay for structural retrofitting. Regarding marital status, married individuals were more likely to pay (81%), followed by divorced (75%). Single individuals were the least likely to pay (63%). Households with children were also more likely to pay (79%) than those without children living at home (70%).

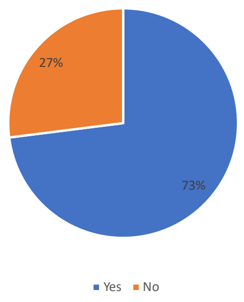


Figure 4-78 - Respondents willing to pay for structural retrofitting (%)

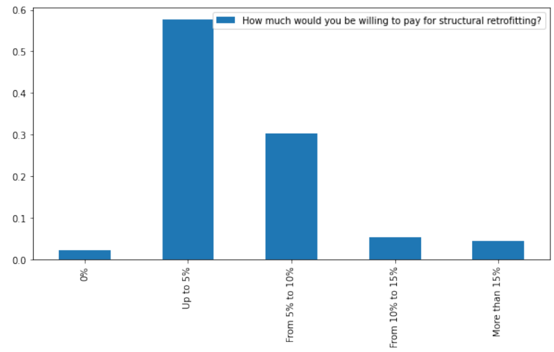


Figure 4-79 – Percentage of total property price that respondents would be willing to pay for structural retrofitting

It is also possible to see a slight increase in the willingness to pay for structural retrofitting in wealthier households (Figure 4-80). For example, households with a monthly income higher than 5000€ were 6% more willing to pay for structural retrofitting than those earning between 2001€ and 3500€.

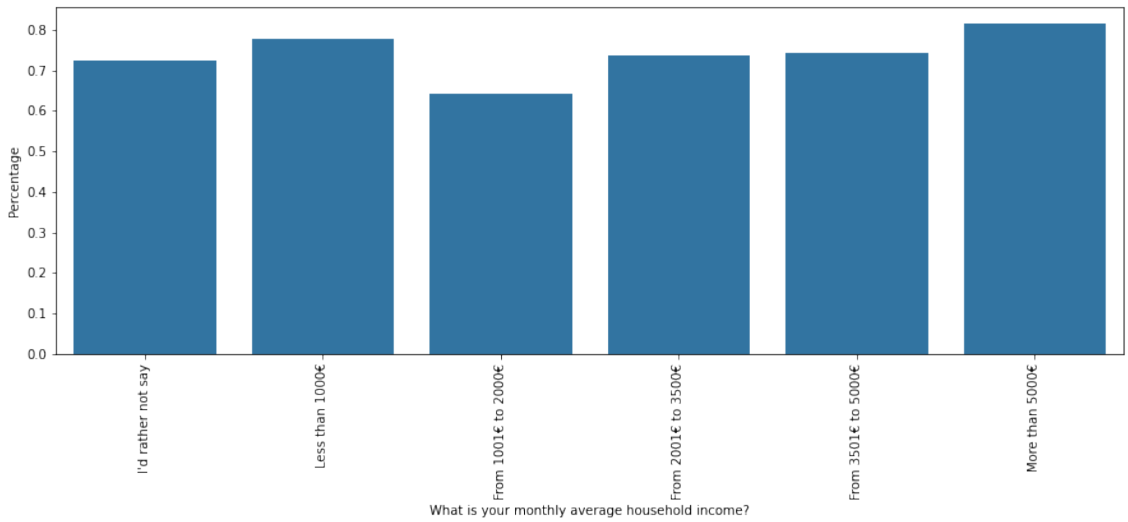


Figure 4-80 – Respondents agreeing to pay for structural retrofitting (%), grouped by income level

Regarding the values they would be willing to pay, the data also shows that wealthier households are willing to pay more for retrofitting (Figure 4-81). These observations align with the existing literature on the relationship between willingness to accept risk and household wealth. As previously mentioned, Asgary and Willis (1997b) showed that safety is a function of income and wealth. Hence higher-income individuals are more willing to pay for additional safety measures. Thus, this relationship suggests the existence of a utility function between safety and income.

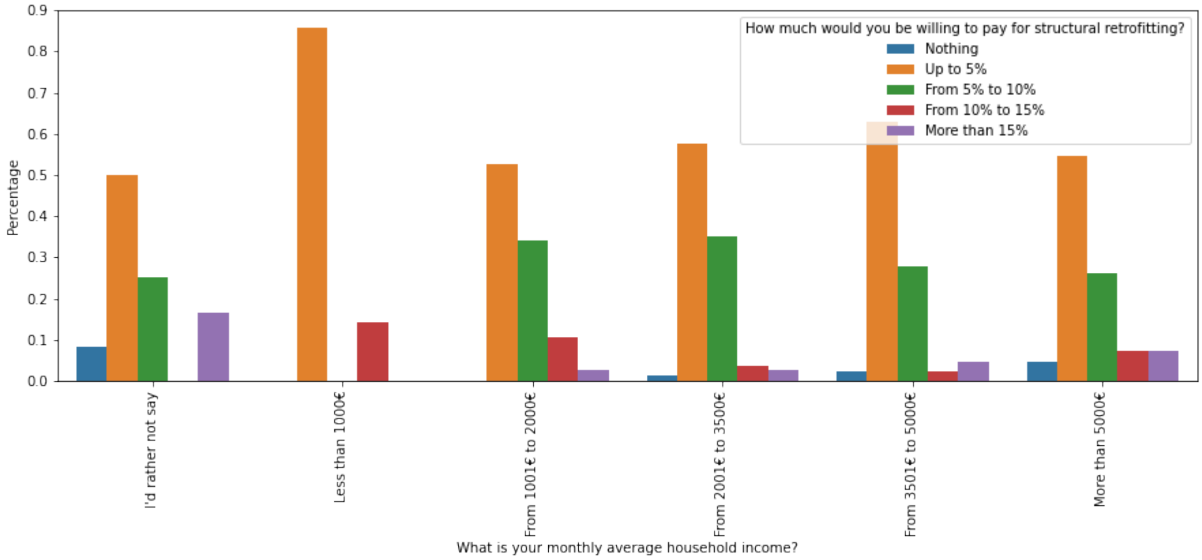


Figure 4-81 -Respondents willing to pay for structural retrofitting of their building (%), grouped by income level.

The willingness to pay also varies across age groups, with older individuals being more willing to pay for retrofitting. For example, 84% of individuals older than 55 agreed to pay, decreasing to 77% for ages between 35 and 54 and to 64% for ages between 16 and 34. Regarding the age of the building, we see a clear willingness to pay for households living in buildings pre-1960, with acceptance levels over 70% (Figure 4-82).

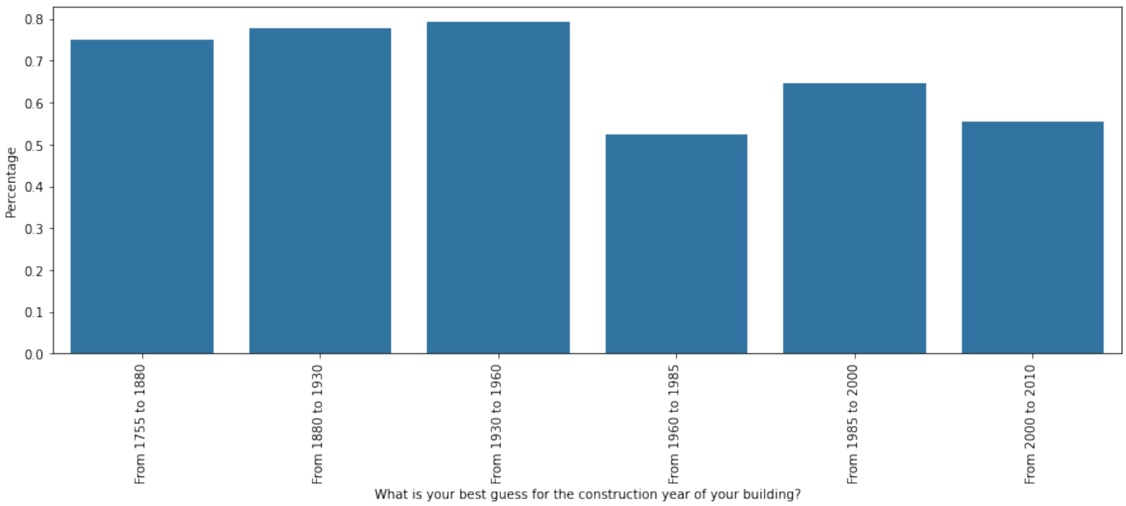


Figure 4-82 - Respondents willing to pay for structural retrofitting of their building (%), grouped by income level.

Regarding the spatial distribution of willingness to pay for structural retrofitting, parishes such as *Carnide*, *Alvalade* and *Benfica* seem to be some of the most receptive showing high percentages of acceptance (Figure 4-83).

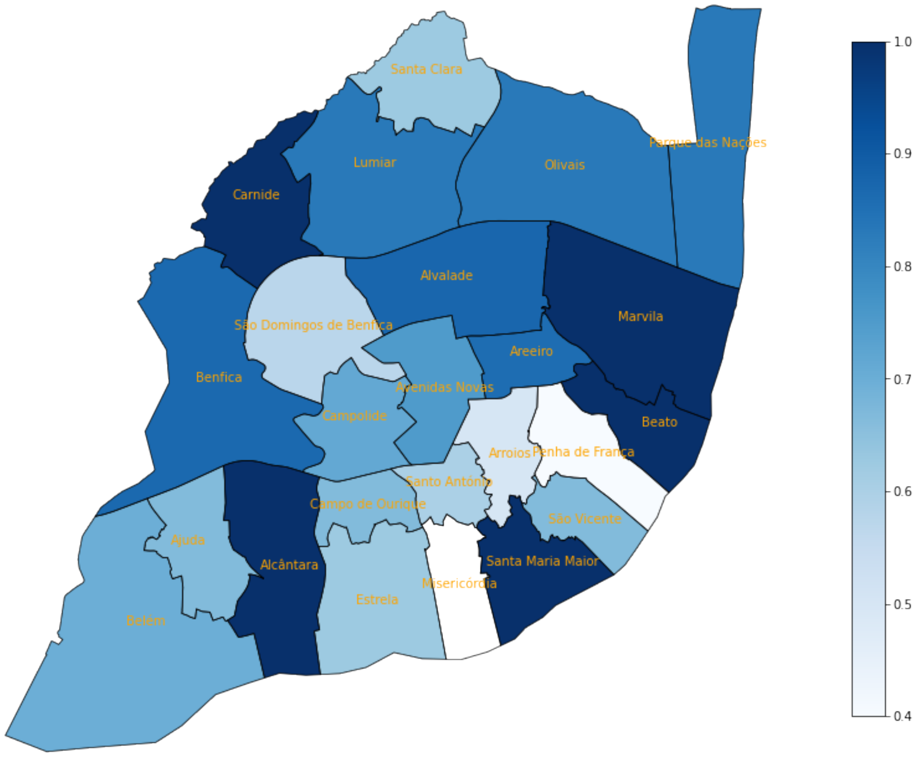


Figure 4-83 - Spatial distribution WTP for structural retrofitting (%; 1.0 = 100)[193]

The last question focused on the preferred way to pay for structural retrofitting. Of those who agreed to pay for it, 64% would prefer to pay it as a direct payment to contractors, with tax benefits for the property owner, 20% would pay it as a percentage of the municipal property tax (IMI), and only 16% would choose a one-time payment to the municipality.

4.4.2. Regression Results

Before proceeding with building the regression model, the dataset was screened for outliers and extreme responses. In total 9 variables were selected from the dataset: household size (categorical variable), the existence of children under 18 in the household (binary), household income (categorical), gender (binary[194]), age (categorical), education level (categorical), marital status (categorical), experience of a previous seismic event (binary), the willingness to pay a premium for a less vulnerability building (binary) and the willingness to pay for structural retrofitting (binary). The choice of the variables aimed to characterize demographically each respondent (Table 4-13).

193 Note that, in the face of the small sample size, the results may not be statistically significant. For the entire sample size used for each parish, please refer to Table A.0 12.

194 Given the low percentage (1%) of respondents not identifying with male nor female genders, the author chose to remove the records from the dataset, allowing this variable to be treated as binary, and eliminating the necessity to encode the variable. It is should be mentioned that this decision was purely based on the practicality of including gender as a binary variable in the model, despite the acknowledgment of non-binary identities.

Table 4-13 - Set of chosen variables

Variable	Description	Type	Obtained from (Question)
tam_agg	Household size	Categorical	A.1
criancas	Existence of children under 18 in the household	Binary	A.2
rendimento	Household income	Categorical	A.3
masculino	Gender	Binary	A.4
idade	Age	Categorical	A.5
estadoCivil	Marital status	Categorical	A.6
educacao	Education level	Categorical	A.7
sismo_exp	Experience of a previous seismic event	Binary	C.11
pag_reforco	Willingness to pay for structural retrofitting	Binary	D.5
pag_resil	Willingness to pay a premium for a less vulnerability building	Binary	D.1

After the initial screening, household size, income and age and education level were converted into ordinal numerical variables. As for the marital status variable, it was encoded into *dummy variables* using python's *OneHotEncoder* module (from *scikit-learn* library).

As mentioned in the methodology section, a binomial logistic regression was used. The variable *pag_resil*, which concerns the willingness to pay to a premium for a less vulnerable, was chosen as the dependent variable, and the remaining variables were selected as independent variables, with the expection of *pag_reforco,* which was excluded from the model due to the high correlation with our target variable. Only the "married" dummy variable was included to avoid redundancy. The results are shown in Figure 4-84.

Logit Regression Results

Dep. Variable:	pag_resil	No. Observations:	301
Model:	Logit	Df Residuals:	292
Method:	MLE	Df Model:	8
Date:	Tue, 02 May 2023	Pseudo R-squ.:	0.07166
Time:	18:07:23	Log-Likelihood:	-136.95
converged:	True	LL-Null:	-147.52
Covariance Type:	nonrobust	LLR p-value:	0.006776

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.8970	0.715	-1.254	0.210	-2.299	0.505
tam_agg	-0.0742	0.195	-0.381	0.703	-0.456	0.308
criancas	0.0219	0.465	0.047	0.962	-0.889	0.932
rendimento	0.4822	0.178	2.713	0.007	0.134	0.831
masculino	-0.0349	0.312	-0.112	0.911	-0.647	0.577
idade	0.5808	0.246	2.363	0.018	0.099	1.063
educacao	0.2293	0.216	1.062	0.288	-0.194	0.652
sismo_exp	0.3466	0.319	1.086	0.277	-0.279	0.972
estadoCivil_Casado	-0.2664	0.324	-0.822	0.411	-0.902	0.369

Figure 4-84 - Results of first binomial logistic regression model

In this first model, only two variables proved to be statistically significant, showing a p-value lower than 5%: household income (*rendimento*) and age (*idade*). Thus, in the second model, only these two variables were included. The results can be seen in Figure 4-85.

Logit Regression Results

Dep. Variable:	pag_resil	No. Observations:	301
Model:	Logit	Df Residuals:	299
Method:	MLE	Df Model:	1
Date:	Tue, 02 May 2023	Pseudo R-squ.:	0.05818
Time:	18:08:23	Log-Likelihood:	-138.94
converged:	True	LL-Null:	-147.52
Covariance Type:	nonrobust	LLR p-value:	3.425e-05

	coef	std err	z	P>|z|	[0.025	0.975]
rendimento	0.3645	0.061	5.987	0.000	0.245	0.484
idade	0.4733	0.214	2.215	0.027	0.055	0.892

Figure 4-85 - Results of second logistic regression model

The results show that both the household income and the respondents age have a positive effect on the willingness to pay a premium for a less vulnerable property, being statistically significant in the model for a significance level of 5%.

4.5. R3: Joint Regression Model

The joint regression model uses the data obtained from the survey to assess possible impacts on the property values, as measured through the normalized transaction prices. Figure 4-86 shows the spatial distribution of properties (in blue) and respondents (in red) living in Lisbon municipality. The green rings represent the limit of the area surrounding the respondents' residence, considered to calculate the average sale price. It represents a circle of 300 m in radius, or nearly 5 minutes walking. This distance was chosen in order to capture the price dynamics of each neighborhood / local market.

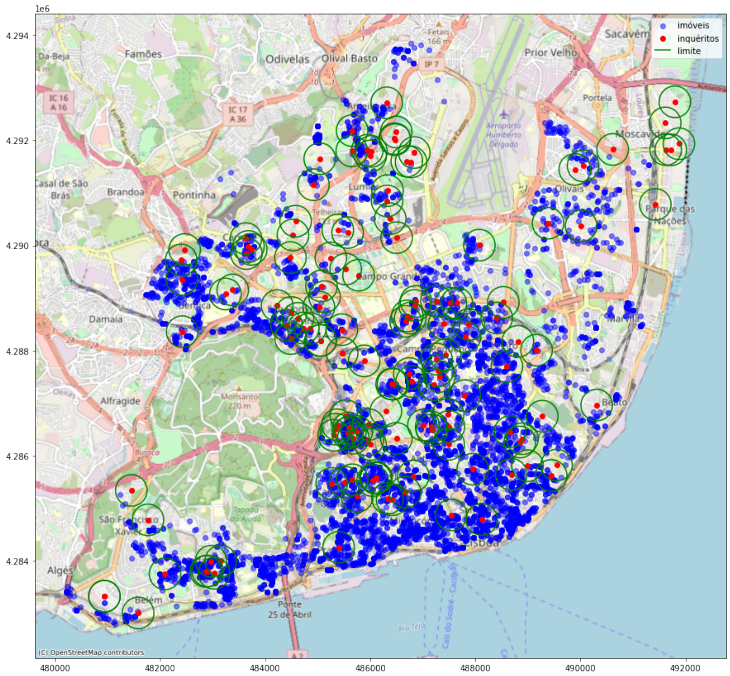


Figure 4-86 - Properties in the dataset (blue), survey respondents (red) and area limit (green)

A total of 11 respondents had no properties referenced in the database within a 300m radius (Figure 4-87).

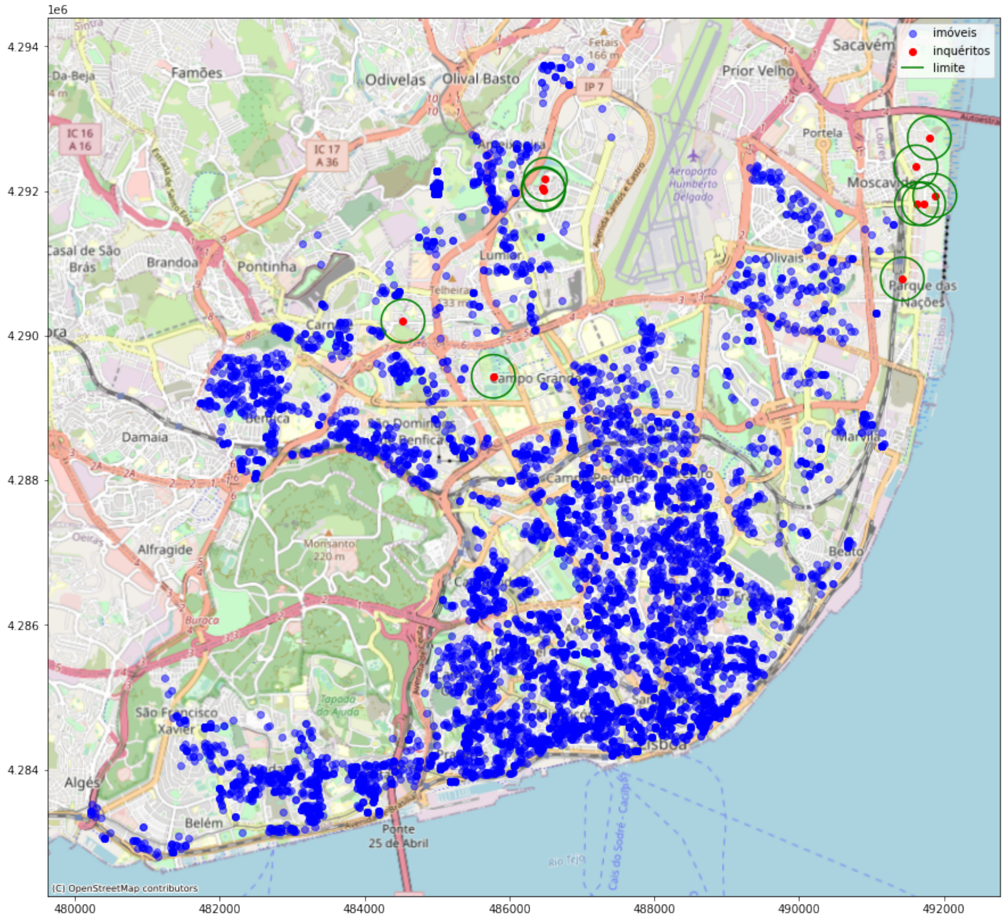


Figure 4-87 - Respondents with no properties within a 300 metres radius

Some of the responses to the statement in section C of the survey, regarding the respondents' perception and action towards risk, were encoded into variables suitable to be used as inputs. Responses were scaled from 1 to 5, where 1 corresponds to "completely disagree" and 5 to "completely agree". Table 4-14 shows their correspondence to the set of variables used.

Table 4-14 - Statements encoded as variables

Statement	Corresponding variable
"If an earthquake is going to harm me, it will, and there isn't much that I can do about it"	stat1
"If an earthquake is going to occur, there is not much my city/community can do to lessen its effects"	stat2
"If an earthquake is going to occur, the government will take care of me."	stat3
"If an earthquake is going to occur, my insurance policy will take care of me."	stat4
"My building is safe against earthquakes."	stat5
"All recent buildings are safe to earthquakes due to restrictive building codes."	stat6

The heatmap in Figure 4-88 reveals little correlation between the responses, with the exception of three cases:

- **Case 1** - A correlation of 0.41 between *stat5* and *stat6*. This may hint that the knowledge of the construction activity can raise the awareness of vulnerability, leading to a strong correlation between respondents who feel their building is secure and those who trust in enforcement and restrictiveness of building codes to provide structural resilience;
- **Case 2 -** A correlation of 0.38 between *stat3* and *stat4*. This may be used as a proxy for the trust (or distrust) shown by respondents in public and private institutions and in their ability to handle a catastrophe, especially in the ability of municipalities and insurance companies to take care of them and their families;
- **Case 3 -** A correlation of 0.33 between *stat1* and *stat2*. This can be associated to fatalism, showing a positive correlation between those who feel helpless to mitigate earthquake consequences and those who believe there is not much local governments can do to mitigate risks.

In addition, a positive correlation exists between another set of variables: household size (*tam_agg*), the existence of children under 18 in the household (*criancas*), household income (*rendimento*) and marital status (*estadoCivil*). This is expected as married couples are more likely to have a higher (combined) household income as well as to have children. Having children will, naturally, lead to having children living with them and to increase their household size.

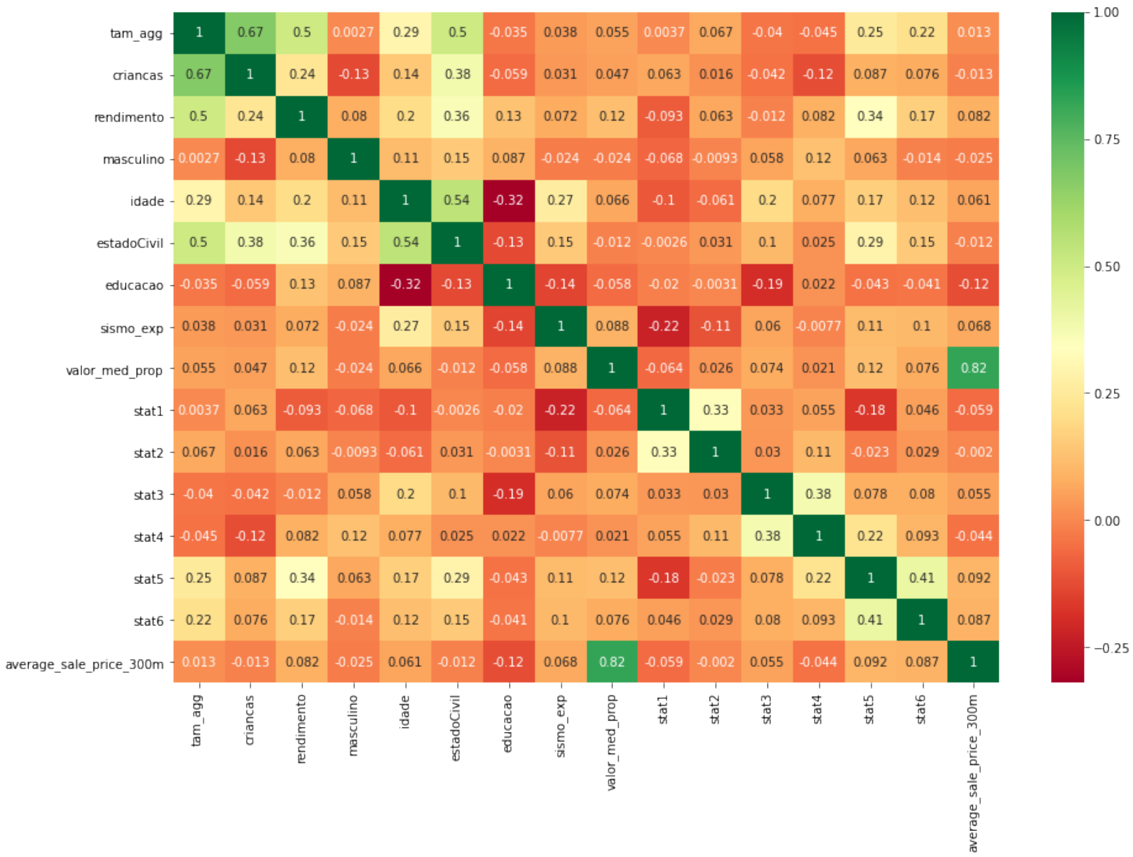


Figure 4-88 - Correlation heatmap

Based on the correlation analysis, a set of 4 independent variables was chosen to be included in the regression: one variable for each case (1, 2 and 3) and one to serve as proxy for the socioeconomic situation. This was to minimize the collinearity between independent variables. In each case, within each pair of variables, the author chose the one with the highest absolute value of correlation to the dependent variable, the average sale price within a 300m radius. *Stat1*, *stat3*, *stat5* and *tam_agg* were the chosen input variables. The results were as follows (Figure 4-89):

```
                                 OLS Regression Results
==============================================================================================
Dep. Variable:     average_sale_price_300m   R-squared (uncentered):                   0.805
Model:                                 OLS   Adj. R-squared (uncentered):              0.801
Method:                      Least Squares   F-statistic:                              211.7
Date:                     Wed, 24 May 2023   Prob (F-statistic):                    1.36e-71
Time:                             12:18:59   Log-Likelihood:                         -137.92
No. Observations:                      209   AIC:                                      283.8
Df Residuals:                          205   BIC:                                      297.2
Df Model:                                4
Covariance Type:                 nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
tam_agg        0.2964      0.011     26.662      0.000       0.274       0.318
stat1         -0.0069      0.028     -0.245      0.807      -0.063       0.049
stat4         -0.0254      0.032     -0.784      0.434      -0.089       0.039
stat5         -0.1076      0.034     -3.127      0.002      -0.175      -0.040
==============================================================================
Omnibus:                        0.434   Durbin-Watson:                   1.408
Prob(Omnibus):                  0.805   Jarque-Bera (JB):                0.206
Skew:                          -0.045   Prob(JB):                        0.902
Kurtosis:                       3.124   Cond. No.                         3.76
==============================================================================
```

Figure 4-89 - Results of the joint regression

The results show that only the socioeconomic proxy (*tam_agg*) and the knowledge of the construction sector (*stat5*) have significance in the model, showing a p-value of less than 5%. This can hint that neither fatalism nor the trust in the government and insurance companies has significant effect on market values. It also hints that the demand in the market, and that property prices, are mostly defined by the socioeconomic conditions of the household and not by their perception of risk. However, we may note that the knowledge of the construction sector, and the belief in the correct enforcement of the best building practices and inspection may also influence the decision-making process of households.

5. Discussion

5.1. Overview

This section will further interpret and discuss the results presented in the previous section in light of the three initial research sections:

4. R1: Do market values reflect a preference for properties less vulnerable to earthquakes?
5. R2: How do residents and investors in Lisbon perceive risk?
6. R3: Does their risk perception affect their willingness to pay for property?

The content will be divided into three sections, each addressing a research question. Each section will present a summary of the findings, and a comparison to previous research. Throughout the discussion, the focus will be on answering the research questions by considering their meaning, significance, and limitations. Additionally, the chapter will conclude with suggestions for future research or analysis.

5.2. R1: Do market values reflect a preference for properties less vulnerable to earthquakes?

The initial EDA has shown that the spatial distribution of property values overlays with the highest building vulnerability and flooding risk, namely in the *Misericórdia*, *Santa Maria Maior*, *Campolide* and *Avenidas Novas* parishes. These are parishes central to the city, thus hinting that their location would be one of the major factors contributing to their market value. A correlation analysis also showed that the condition of the building (0.338), the condition of the apartment (0.336), the construction year (0.277), the area (0.100) and the floor (0.099) are some of the variables showing the highest positive correlations with the yearly normalized prices. This is consistent with the initial expectations as newer and better-conditioned properties are more appealing to potential buyers, therefore, more valuable to the market. Larger apartments are usually more expensive than smaller ones, and higher floors are also most valuable since they enjoy less noise and pollution and often better views than lower floors. A negative correlation of prices to the distance to the CBD (-0.163) also showed that location was a primary factor being considered by the market. In order to avoid the Modifiable Unit Areal Problem (MAUP), the raw values were used for the analysis with aggregation into old parishes (with smaller sizes, thus more approximate to the ideal homogeneous units of analysis) and aggregation into new parishes (as defined by Law n.º 56/2012). The results were consistent across all aggregation levels. A Building Resistance Deficit (BRD) was calculated as a proxy of the building's vulnerability based on its age, typology, location and number of floors. The highest mean values of BRD are located in the parishes of *São Nicolau* (0.838), *Madalena* (0.721) and *São Paulo* (0.627), in the city centre, in some of the most expensive downtown locations, next to the Tagus river. In sum, a first analysis hinted that the market is oblivious to vulnerability, factoring in the location (especially the distance to the city centre), the condition of the building and of the apartment itself. When aggregating in the new parishes, *Santa Maria Maior* (0.545), *Misericórdia* (0.464) and *Santo António* (0.452), also located downtown, show the highest values of BRD while showing some of the highest property values in the municipality. Prices generally demonstrated a spill-over pattern from the city centre to its outskirts, with no regard to seismic risk. *Parque das Nações*, a newer wealthy neighbourhood, also demonstrated higher prices despite the distance to the city centre. A PCA analysis revealed that a significant part of the data variance (25.46%) could be explained by the first component which could be identified as being related to the proximity to the river and the centrality, and by the second factor (19.26%) related to the condition of the building and the dwelling itself. The following three components were also related to amenities nearby. The feature importance analysis, carried out using an XGBoost algorithm, showed that the area, the construction year, the floor number and the number of rooms were also the most important features considered by the market. In this analysis the importance of the location was, again, highlighted with the proximity of *Parque das Nações* being one of the most important features in the parish of *Marvila*, and the proximity of *Amoreiras* (one of the business centres in Lisbon) and *Avenidas Novas* being one of the decisive factors, pricewise, for properties located in *Campolide*. Regarding the selection of variables for the regression analysis, a *best subset selection* and a *stepwise regression* were carried out. They showed consistency with the previous findings, highlighting the importance of location and that newer and better-maintained buildings are in higher demand presenting higher market values. These first regression methods, here included in the EDA, also showed the heterocedasticity of errors, usually found when modelling highly location-dependent values as property prices.

Next, four types of regression models, with different aggregation units and levels of complexity, were built in order to explore the robustness of results: (1) Feasible Generalized Least Squares regression; (2) PooledOLS; (3) Fixed-Effects; (4) Multivariate Adaptive Spline Regression (MARS); and (5) Extreme Boosting Tree (XGBoost).

The first set of models (FGLS) showed that the building's condition (BUILDING_2), the apartment's condition (APARTMENT1), the area (AREA) and the floor number (FLOOR) were the most important variables, contributing positively to the property's price. On the opposite side, the building resistance deficit has shown to be negative, which would be according to our initial expectations; however, it was not possible to prove its significance and reject the null hypothesis of a null coefficient given that its p-value was lower than 5%. The results obtained from the PooledOLS model were aligned with the previous findings; nonetheless, the area (AREA) showed a p-value larger than the 5% significance level. The Fixed-Effects model, built for different aggregation levels, showed a negative coefficient for the building resistance deficit (BRD_PROXIMO), and positive coefficients for the area (AREA) and floor number (FLOOR). The best results were obtained for the old parishes aggregation level, showing no spatial autocorrelation of the error term, thus showing the validity of the results. The best results were achieved for the second model for old parishes (based on construction year and location), with an r-squared of 21.39%, followed by the first model of new parishes, with the first approach of BRD attribution, which is based on the construction year, with an r-squared of 10.61%. The MARS model showed that the apartments' condition and the floor number where the most important features to the property price. Finally, the XGBoost model showed the highest r-squared value at 71.29%.

The overall results were measured using the R-squared (RSQ), which measures the proportion of variance in the dependent variable that the model can explain. In other words, it measures the goodness of fit of our model. The comparison between our best models, for each type, can be found in Table 5-1.

Table 5-1- Best results obtained in each type of model

Metrics	FGLS	PooledOLS	Fixed-Effects	MARS	XGBoost
R^2	14.20%	11.75%	21.39%	16.09%	71.29%

In light of the results, it is possible to conclude that the market does not reflect a preference for properties less vulnerable to earthquakes. However, it is worth notice that it presents a preference for newer, better buildings, which may present less vulnerability, driven not by the concern of safety itself but by the amenities offered the property, namely its condition and the proximity to high-value locations (e.g. downtown city centre, proximity to the river or proximity to the CBD).

These results do not contrast with the existing literature. It is known that earthquakes typically damage structural elements (beams, joints, panels, load-bearing walls, etc.) and non-structural elements (partition walls, non-structural roofs, furniture, installations, equipment, etc.) and therefore, theoretically, the increased vulnerability of a property should be accounted for in its pricing. As mentioned by Naoi et al. (2009) "*since an earthquake is an exogenous risk factor that is tied to a specific location, its risk should be capitalized into local housing and land prices".* This is observable in various locations subject to earthquake risk, such as Tokyo (Nakagawa et al. 2007), Iran (Willis and Asgary 2007) and California (Brookshire et al. 1985). However, the literature also shows that most of these price-reduction effects occur after the disclosure of risk information in high-seismicity locations (see, for instance, Naoi et al. (2009)), and especially after in post-earthquake situations (as shown by Modica and Locati (2016)). Traumatic events, such as earthquakes, have a profound effect not only on a person's mental health (leading to prolonged posttraumatic stress disorders, panic attacks, recurrent nightmares, and survivor-guilt feelings (de la Fuente (1986), as cited in Dooley (1992); Wood et al., 1992)) but also in the market as a whole, which is subject to the judgement of all stakeholders which, ultimately, culminates in its price dynamics. Nevertheless, other authors have shown that the overestimation of seismic risk may even lead to price reductions, or even slight increases, in post-earthquake periods if damages are less than initially expected. It is primarily a question of individuals' perceived vulnerability. Additionally, the literature has shown that legislation did not significantly affect the housing market, as most appraisers did not give any higher price on a comparable property in a lower-risk area (Palm (1990), as cited in Nakagawa et al. (2007)).

There are several implications to these findings:

- Markets seem to be oblivious to the building's vulnerability to seismic events. However, despite its possibly low probability, municipalities must take a serious and structured approach to risk communication. This should be done in proximity to the population, by disclosing hazard maps or risk indicators, and through education. Even if an event does not occur in the addressed location, the effect of Out-Of-Market events may add the disclosed information, increasing awareness and which should be translated into an increase in the perceived vulnerability of its inhabitants, which, in its turn, should be reflected into the price.

- The results have shown that the building and apartment's condition, as well as its area and location, are some of the most important features taken into consideration by prospective homebuyers during their decision process. Therefore, it should be mandatory for rehabilitation projects to account for structural retrofitting, especially in a city where most residential buildings were built before any seismic regulations. It should also translate into a higher degree of inspection by the municipality during licensing and construction phases to ensure compliance with the highest standards of buildings' seismic performance. It is well known that Portuguese municipalities struggle to respond timely to licensing requests, let alone the availability to have tight surveillance over construction projects. However, it should be noted that people's lives are at stake and that any underestimation of seismic effects or lack of detail during the construction process may lead to severe consequences.
- Buildings being built or rehabilitated in high-risk areas should be closely monitored by inspectors due to their added layer of risk, given that miscalculations or lack of detail during the construction process can be aggravated in these locations. Additional mitigation measures could also be required for such locations, which should be explicitly defined in land-use municipal plans and regulations.
- Municipalities should collaborate with insurance companies to define hazard maps that should be reflected in the insurance premiums and coverage policies. These should also reflect the structural vulnerability of the building, which insurers presenting lower premiums to less vulnerable buildings, thus incentivizing private owners to invest in structural retrofitting.
- There should also be a collaboration between the municipality and the remaining real estate stakeholders (such as real estate developers and brokers, appraisers, construction companies, neighbours' associations, academia and other public institutions). A seismic resilience committee should be gathered to create and monitor policy proposals that align all the incentives from various parts and foster the adaptation of mitigation measures. The results of each measure should be thoroughly analyzed and iterated over to ensure the effectiveness of such policies.
- Appraisers should also be encouraged to incorporate risk assessment into property valuation processes to improve industry practices. Most residential properties are appraised using the comparable method, where several features from comparable properties are taken into account. A publicly available index for the buildings' structural vulnerability could also be partly incorporated into this valuation method.

This set of recommendations should be applicable not only to Lisbon but to other moderate seismicity locations, with similar cultural backgrounds, with no record of severe seismic action in recent years.

Notably some of these recommendations are already be addressed by Lisbon's municipality. The ReSist[195] initiative has already taken significant steps to increase the city's resilience to seismic events, defining a set of 47 measures based on six goals, namely:

- The standardization of technical standards and methodologies for assessing the city's seismic vulnerability.
- The development of operational actions set to effectively promote resilience, implemented through inspection campaigns, projects, and structural reinforcement works that focus on the built stock and vulnerable infrastructures.
- Carrying out awareness and dissemination campaigns to involve society, to train the population in general.
- The development of information management systems that streamline the tasks of knowledge sharing and program execution among the various municipal structures;
- The definition and implementation of alert criteria and intervention prioritization that guide the action of municipal services.
- Strategic partnerships must be defined with external entities to optimize actions.

The initiative's leadership has both municipal and academic representation, and its technical-scientific committee also counts on private construction companies and entities such as the National Laboratory for Civil Engineering (LNEC) and the engineers' and architects' national associations.

Additionally, under new regulation[196], which took effect on November 15th, 2019, expansion, alteration or reconstruction works may be subject to the preparation of a seismic vulnerability assessment report and even require the preparation of a seismic reinforcement project.

These constitute necessary steps to mitigate seismic risk and the vulnerability of Lisbon's housing stock. However, as previously mentioned, several additional steps could be taken in order to address risk and which could ultimately be translated into the market pricing dynamics.

Lastly, some methodological implications should be taken. As the results have shown, smaller aggregation units should be chosen over larger ones, as local supply and demand factors especially condition real estate markets. Therefore, smaller aggregation units allow the researcher to capture local market dynamics in ways that larger aggregation areas cannot. Also, given that the real estate market is subject to phenomena that are not limited to administrative boundaries, studies should be carried out considering various aggregation levels to further increase the results' robustness. Spatial effects should also be considered whenever possible, as spill-over dynamics are frequent in this market.

The choice of regression models to apply should also depend on the research goal. For example, as shown by the results, the Extreme Boosting Tree algorithm's deployment significantly improved the data

[195] For more information, please refer to https://informacoeseservicos.lisboa.pt/prevencao/resiliencia-urbana/projetos/resist
[196] Portaria n.º 302/2019, de 12 de setembro. For more information, please refer to https://dre.pt/dre/detalhe/portaria/302-2019-124642991?_ts=1683796107224

goodness of fit but at the expense of explainability. Thus, a more generalized approach, using less complex statistical models, could be a better fit for research where the goal is to analyze the impacts of each variable, while advanced models, using complex decision tree assembly methods, or even the employment of neural networks, should be reserved to contexts where the models' performance is the primary focus.

5.3. R2: How do residents and investors in Lisbon perceive risk?

Several factors may contribute to how individuals perceive risk. Age, marital status, presence of children in the dwelling, previous experiences, cultural factors and media exposure all can affect our perception of reality and, hence, our perception of risk. Therefore, in order to analyze how risk is perceived by Lisbon's residents and investors, one should have a clear understanding of its demographics. The sample collected for this study comprised 325 surveys answered over approximately one month, achieving a margin of error of 5.44% for a 95% confidence interval. The sample demographics showed that 47% of respondents are married, 40% are single, 7% are divorced and 5% claim other statuses. In addition, 37% of households had children under 18 living in their dwellings. Most respondents were highly educated, with nearly half holding a master's or PhD degree.

A significant portion of the sample comprises young adults living in Lisbon with an average monthly household income of around two to three thousand euros. Most of these respondents are homeowners (70%) or private market renters (26%), of which 76% pay rent below 1.000€. Respondents have shown proper knowledge of their housing conditions and features, such as the estimated area and market value. Most had an adequate notion of their building's age. The assessment of their perception of risk has shown that earthquakes are on top of their concerns, followed by fires, home robberies and floods. However, these results have to be taken with due care.

The survey was conducted in the aftermath of Turkey's February 6th, 2023, which had extensive media coverage and raised a nationwide debate on the preparedness of the Portuguese state and infrastructures to deal with a similar event, given that the southernmost part of the country is especially vulnerable to seismic events. This was especially the case for the city of Lisbon. This coverage, however diminishing over time, was still present during March and April as the survey was being carried out and may have led to "salience bias" (Tiefenbeck et al., 2018) caused by an out-of-market event. Nevertheless, the results surprisingly revealed no significant differences in the perception towards risk across different marital statuses, size and income of the household, age, gender, or existence of children in the dwelling. However, there is a difference between Lisbon residents and others. The former's top concern is earthquakes, while the latter mentioned home robberies as their top concern. This probably surged due to differences in the media coverage, which mainly focused on the vulnerability of Lisbon's municipality housing stock. Most neighbouring municipalities received little coverage regarding their preparedness to respond to seismic events.

Respondents also showed a difference in their response towards risk-mitigation measures across different risks, showing a much less fatalistic attitude towards home robberies than towards earthquakes. In fact, despite a clear division across attitudes towards earthquakes, respondents showed a slight inclination for a fatalist attitude, showing feelings of helplessness to change their fate in the case of an event, translating in a less proactive attitude to taking mitigation measures. They also showed distrust in government and insurers' ability to handle an adequate response to earthquakes.

Regarding the effectiveness of building codes and professional practices, respondents diverge in their responses, with most respondents not agreeing or disagreeing. However, respondents seem to favour newer buildings given that their likeliness to classify them as unsafe significantly correlated with their construction year: older buildings were generally regarded as more vulnerable to seismic events, especially pre-1985 buildings. Additionally, there seems to be a correlation between the property value and the perception of its safeness, with pricier properties being regarded as less vulnerable. In practice, this is not always the case. Newer and more expensive buildings do not necessarily translate to safer ones, as there are several empirical examples in the industry of deceiving practices in the construction process.

Nevertheless, the perception of risk differs across education levels, with higher education levels revealing increased awareness of their building's vulnerability, especially from bachelor to master's degree and from master's to PhD. The results also showed a high percentage of respondents who had already felt an earthquake (67%), and a high percentage of earthquake insurance (55%), well above the national average coverage of 16%. Most respondents showed a good knowledge of adequate mitigation measures that could be taken before or how to respond during such an event.

There are several similarities between the results obtained and the ones from the literature. For example, Önder et al. (2004) found out that the warnings about the soil type resistance, previously made by geologists, had only been of significance to high-income locations; therefore, there would be a dependence of the tolerance to risk on the household's income and socioeconomic environment. Furthermore, we find in the literature evidence that higher levels of knowledge about the specific causes of damage and how damage can be reduced correlate with earthquake preparedness (Hurnen & McClure, 1997), consistent with this survey's results. But the perception problem is not limited to lack of information. Hence mere exposure to information will not immediately change the behaviour and the willingness to take mitigation measures (Van Der Pligt,1998).

Turkey's event may have triggered an emotional shock and reminded households of their risk exposure, translating into the added media coverage of earthquake engineers warning of seismic risk. We can argue that the subjective probability of risk perceived by households has momentarily come closer to the objective probability (Lamond et al., 2010). Previous experiences with natural hazards may also alter an individual's risk perception. In this case, the results showed that respondents who experienced an earthquake are not significantly more willing to buy earthquake insurance (only 55,5% of respondents who experienced an earthquake have insurance against it, a 0,5% increase from people who did not experience such an event). This may result from a fatalistic attitude or an "acceptance of loss behaviour" (Burton et al. 1978). In fact, the analysis showed a slightly fatalistic attitude towards earthquakes, which was not observed in the case of home robberies. This lack of significance in the willingness to pay for earthquake insurance may also result from Lisbon's lower magnitude of earthquakes in recent years. Another possible justification may concern the extra cost. With the extra housing costs, which have increased in recent years, households may still opt not to have earthquake insurance, which may constitute a barrier to wider-scale adoption. The literature has also addressed this phenomenon by modelling the decision-making process, namely through the Utility Functions framework. The increment in wealth provided by not spending money on earthquake insurance is justified by the disparity between subjective and objective probability (often neglected). The marginal utility of an extra euro attributed by a wealthier household will be less than the one attributed by a more modest one. Thus, even if the subjective probability is the same for both households, the wealthier one should be more likely to minimize the risk by buying insurance. Thus, the behaviour observed in the survey may hint at the existence of a utility function for earthquake insurance and, broadly, to the investment in mitigation measures.

Another topic worth discussing is the disbelief of households in the governments' and insurance companies' ability to deal with a disaster. It is known that, at the moment following a disaster, most will have to rely on their local governments to take immediate action, as they play a crucial role in preparing, mitigating, and managing recoveries after natural disasters. However, 70% of respondents don't trust the government to be able to take care of them and their families in such an event. This is concerning and should be addressed by both the central and local governments, as one of their core objectives is to provide their citizens with safety, health, and overall good quality of life.

Based on the results, some implications should be taken:

- Since the survey indicates that earthquakes are perceived as a high-concern risk, it is crucial to continue raising awareness among the population about the potential impact of earthquakes and the importance of preparedness. Public education campaigns, community workshops, and informational materials can disseminate knowledge about seismic risk, mitigation measures, and the role of individuals in reducing vulnerability. Continual educational efforts can help dispel misconceptions and promote a more comprehensive understanding of seismic risk, building vulnerability, and the role of building codes and professional practices.
- Disclosure requirements should be established. Implementing regulations that mandate sellers and landlords to disclose the seismic vulnerability of a property can increase awareness among potential buyers and tenants. This can be achieved through a standardized system of seismic risk ratings or certificates, similar to energy efficiency certificates.
- Robust emergency preparedness and response plans should be invested in, including early warning systems, evacuation routes, and public shelters, to minimize seismic events' potential impact and enhance community resilience. Media coverage should then address these investments, strengthening the government's image next to the population. Insurance companies should also be part of such campaigns, showing concern with the matter and a proactive attitude to protect their clients' families.

- Additionally, insurance companies should provide affordable coverage solutions and post-earthquake assistance policies, ensuring that homeowners have trust in the insurance industry's ability to support their needs.

Nevertheless, some limitations of this work should be acknowledged. The first regards the distribution channels chosen to deliver the survey. During the design phase of this study, the author planned four possible approaches to be taken:

1. Send the survey by email through parish council internal services.
2. Send it by email through some of the largest employers in the city (e.g. CTT, EDP, BNP Paribas, and other large companies).
3. Send emails to institutions in the city and to neighborhood associations.
4. Distribute flyers to residents with a QR code allowing the survey to be completed online.

By March 14th, 2023, 85 contacts by email were made to professional associations, neighbours' associations, the municipality and all its civil parishes individually, and some of the largest companies in Lisbon. The complete list of entities is shown in Table A.0-14, in Annex A.

However, the results were different than expected. Most parishes did not respond to the call, while others claimed they were not authorized to distribute surveys to inhabitants. Most professional associations from the real estate sector did not provide any response. The same happened with the initially planned large firms, which were contacted through their Data Protection Officer. Thus, most responses obtained were part of a "friends and family" effort to spread the survey through their contacts and LinkedIn and Facebook. It is also noteworthy that the author contacted Instituto Superior Técnico, through the Area of Studies, Planning and Quality (AEPQ), however, after two and a half months of contacts, as of May 22nd, the survey was yet to be distributed to its alum network. By the end of May, two short stories were published on Instagram, providing 15 additional responses. In order to reach a sizable sample, results were closely monitored to see which parishes could be misrepresented in the sample.

After some follow-up emails were sent, the following approach was twofold: (1) to send the survey to real estate agencies (190 in total) and (2) to print flyers with a QR code to the online survey and distribute it across the city, in the most underrepresented neighbourhoods. A flyer copy can be found in Annex A (Figure A.0-27). A thousand flyers were printed and distributed by hand or through mailboxes. Despite the effort to achieve a high survey dissemination and response rate, our sample has shown a slight overrepresentation of younger classes and a slight underrepresentation of older individuals compared to the 2021 census. This could have also resulted from the distribution channel chosen for this survey, leading to less tech-savvy individuals being excluded. Geographically, the sample also achieves a slight overrepresentation of the *Avenidas Novas* and *Campolide* parishes and an underrepresentation of *Marvila* and *Arroios*. Ideally, the respondents to this survey would have been the sellers or buyers of the properties in the transaction database. Only then would we have a complete correspondence between their perception and actual buying/selling decision. Lastly, it should be pointed out that the author contacted the National Insurance Association (APS) but, unfortunately, no information regarding the spatial distribution of earthquake insurance premiums was provided. This would have been an interesting addition to our database, as it would provide an insurer's perspective of risk and a deep understanding of residents' and investors' insurance purchase decision-making.

5.4. R3: Does their risk perception affect their willingness to pay for property?

As previously mentioned, an increase in the perceived vulnerability does not automatically translate into action towards risk or an increase in the willingness to pay more for a less vulnerable property. The survey results have shown that 80% of respondents were willing to pay more for a more resistant building to seismic events, and most were willing to pay up to 5% more. However, the analysis of the transaction database also shows that that is not the case. One possible explanation is the sharp increase in property prices, especially compared to 2014 values, and the significant decrease in housing affordability. Some respondents claimed they were unwilling to pay more for a more resilient property and claimed that this decision should be on the buyer's side but should be required by the municipality during the licensing process. Regarding structural retrofitting of existing buildings, 73% of survey respondents agreed to pay for it, of which 87% would pay no more than 10% of their estimated property value. When asked for an absolute value 50% of respondents were only willing to pay up to 5.000€.

While no significant difference in willingness to pay was noted across the respondent gender, the results showed that married individuals were the most likely to pay for structural retrofitting. In contrast, single individuals were the least likely to pay. Households with children living in their dwellings were also 9% more likely to pay than those without. Wealth and age were also shown to be part of the household's decision-making process, with wealthier and older respondents showing a higher disposition to pay for retrofitting. In regard to the spatial distribution of responses, the parishes of *Carnide*, *Alvalade* and *Benfica* seem to be some of the most receptive to pay, showing high percentages of acceptance.

Households living in pre-1960 buildings also show a higher willingness to pay. Additionally, the preferred way to pay for structural retrofitting would be as a direct payment to contractors, with tax benefits for the property owner (64%), 20% would pay it as a percentage of the municipal property tax (IMI), and only 16% would choose a one-time payment to the municipality.

The regression results have shown that market price is mainly defined by the socioeconomic conditions of households and not by their perception of risk or attitude towards it, with the possible expectation of their knowledge and belief in the proper enforcement of building codes and construction practices, leading the more skeptic to have a more thoughtful decision-making process.

The consistency of these results with the existing literature is not straightforward, as the literature shows mixed evidence concerning the willingness to pay for risk premiums. Various examples have shown that people are more willing to pay after experiencing a natural hazard. That has been shown for floods (Bin & Polasky, 2004; Daniel et al., 2006), wildfires (Donovan et al., 2007), hurricanes (Ewing et al., 2007; Burrus et al., 2009; Zhang et al., 2010; Asgary & Halim, 2011; Lee et al., 2017) and earthquakes (Önder et al., 2004; McGinnis, 2004; Naoi et al., 2009; Deng et al., 2013). Other authors have even shown the effect of “out-of-market” events (Hallstrom & Smith, 2005; Modica & Locati, 2016; Fekrazad, 2019) which negatively impacted the price of the more vulnerable properties. Others have found that if people initially overestimate the risk, a mismatch between the expected damage and actual damages may even backslash, leading to a price increase in riskier areas (Beron et al., 1997).

We can argue that the defining topic in this discussion is the occurrence of an event. As referenced in the literature, people tend to overweight recent information and underweight prior data (Zhang et al., 2010; Deng, 2013). As stated by Fekrazad (2019), the higher the death tolls in the “out-of-market” (OTM) event, the higher the price differential in the aftermath (a decrease in 3%-6% in prices depending on the zip code’s seismic risk). Thus, we could hypothesize that the large Turkey earthquake, which led to the death of 40.000 people, could have impacted the prices in Lisbon. One aspect remarkably consistent with the literature and the survey results is that the willingness to pay and the effect of OTM events are not constant across socioeconomic classes. Toya and Skidmore (2007) state that “higher levels of income also increase private demand for safety”. In other words, “safety is a function of wealth and income” (Asgary & Willis, 1997b). Therefore, the high prices and financial struggles felt by some of the most vulnerable Portuguese families, especially younger and single-parent families, can limit the ability to choose a safer dwelling, which is consistent with the existing literature (see, for reference, Pothon (2020)).

One should also highlight that most respondents account the municipality liable, during the licensing process, in addition to the contractor, for verifying and certifying the structural resilience of the building. To that extent, we may argue that some recent political measures have taken an opposite direction. With the goal to address housing affordability issues, by increasing private construction, the Portuguese government launched, in February 2023, a set of measures in which architectural and engineering processes are no longer subject to municipal licensing, being solely subject to a term of responsibility for the designers, with municipal licensing being limited to urban requirements[197]. Some specialists argue that design projects may be subject to error, which may lead to disastrous consequences. At the same time, others go as far as to acknowledge that design engineers may even deceive the municipality to get the project's approval[198]. There is to say that the requirement to meet and verify the construction standards concerning the seismic vulnerability of new builds is solely attributed to design engineers. However, we should note that recent regulations (Portaria n.º 302/2019) require the delivery of seismic vulnerability assessment report elaborated by a specialized entity in the case of rehabilitation or reconstruction.

Based on the results, some implications should be taken:

- Households place most of the responsibility on municipalities and contractors to comply with structural regulations and ensure the building stock's structural resilience. Therefore, given that one of governments’ primary goals should be to provide safety to its citizens, tighter surveillance

[197] See, for reference, https://www.portugal.gov.pt/pt/gc23/comunicacao/noticia?i=governo-aprova-pacote-mais-habitacao
[198] See, for reference, https://eco.sapo.pt/2023/02/09/edificios-reabilitados-em-portugal-falham-inspecao-de-risco-sismico/

of the construction and reconstruction processes should be taken by municipalities. Given that most municipalities struggle with limitations in human resources, they should, at least, increase the monitoring and the number of inspections of buildings under construction or reconstruction in the most vulnerable locations of the municipality, which could be easily done through soil vulnerability maps. In other words, the prioritization of the inspections should be directly proportional to the soil's vulnerability.

- The Lisbon municipality should continue their efforts to improve risk communication. Increased transparency in the construction practices and realities may increase the awareness of potential homebuyers leading to more thoughtful decision processes, which, in turn, may translate into their willingness to pay a premium for less vulnerable properties.
- In order to promote structural retrofit across the city, one possible measure to be taken by the Lisbon municipality would be to take a pilot case study by addressing families living with children in their apartment, in pre-1960 buildings, in the parishes of *Carnide*, *Alvalade* and *Benfica*, and providing them tax breaks for structural retrofitting of their buildings.
- More broadly, financial incentives and tax breaks given to property owners who voluntarily retrofit their buildings to meet higher seismic standards can encourage proactive action and investment in improving the safety of structures.
- Given that wealth and income are limiting factors to adopting structural retrofitting, local and central governments should facilitate access to financing and grants. Providing financial assistance, low-interest loans, or grants to property owners, especially those with limited resources, can help overcome financial barriers to retrofitting or implementing seismic resilience measures. Collaborating with financial institutions and seeking external funding sources can expand the availability of such assistance programs.
- Governments should partner with academia to provide specialization courses in rehabilitation and seismic resilience to private construction and design companies. Certified companies should benefit from tax breaks, for instance, paying fewer taxes for each certified professional. This would incentivize both professionals and companies to deepen their knowledge and, in the long term, increase the pool of companies with specific skills in rehabilitation, which would increase the market offerings and therefore decrease its price.

Nevertheless, one should note a limitation to these results. In an ideal condition, the surveys should have been answered by the buyers or sellers of such properties. Only in that case one would have a clear perception of the impact of knowledge on their willingness to pay. However, one may argue that the spatial limitation to the neighborhood may compensate for this issue and be an adequate solution to capture and relate the willingness to pay with perception and socioeconomic conditions of residents and investors.

6. Conclusion

6.1. Final Remarks

Earthquakes are responsible for thousands of lives lost and economic damages worldwide. Its mark is visible across all continents and cultures, and it is a certainty to several locations across our globe. Lisbon is not different. Being of moderate seismicity, the city will undoubtedly be subject to seismic events, as it was in the past. The only reasonable action is to mitigate its risk, reducing its consequences. Risk is a function of natural hazard, vulnerability, and exposure. In practice, vulnerability is the only of these points we can reasonably improve and depends not only on our collective but our individual actions and choices. Real estate properties are crucial for human well-being, playing a twofold function of shelter and store of value. It shelters millions of households and is the storage of most of the wealth in the world. Therefore, reducing the vulnerability of the housing stock is of critical importance. Failing to do that is to set ourselves and our communities to failure in the case of an event. In light of that, it is crucial to understand how the housing stock vulnerability is accounted for in the real estate market. Additionally, it is vital to understand how we, as humans and citizens, perceive risk so that political action and communication can be effectively addressed and foster action towards mitigating it.

This thesis was set to contribute to this body of knowledge to provide a deep understanding of how earthquake risk is quantified in the real estate market, how its stakeholders perceive it, and what impact it has. To do that, three research questions were presented:

1. Do market values reflect a preference for properties less vulnerable to earthquakes?
2. How do residents and investors in Lisbon perceive risk?
3. Does their risk perception affect their willingness to pay for property?

The results have shown that the answers to these questions are not straightforward, as markets present different segments and different behaviours depending on the information presented, the news coverage, the levels of education, the wealth of individuals, and the number of children, amongst many other variables. Markets are the sum of all our individual actions. Real estate markets are, simply put, the sum of all our preferences as consumers and our willingness to pay for them, which constitutes the demand side, and the availability of properties in the market, or the offer side, which depends on public and private initiative to build and to develop, and of availability of sellers in the market. During this work, we came to answer our initial questions. Some remarks we can take from this work are the following:

1. **Do market values reflect a preference for properties less vulnerable to earthquakes?**

   Based on the analysis of the literature and the results, it can be concluded that the market does not necessarily reflect a preference for properties less vulnerable to earthquakes. Although the vulnerability of a property to seismic activity is theoretically expected to be accounted for in its pricing, the research shows that the market values newer and better-conditioned buildings, driven not by the concern of safety itself but by the amenities offered by the property, such as its location or proximity to high-value locations. The analysis also reveals that investors and homebuyers seem oblivious to natural hazards such as earthquakes, as demonstrated by the low correlation with prices and the spatial distribution of properties with the highest vulnerability to natural hazards. Nevertheless, it is essential to note that the market can be subject to the judgement of all stakeholders, which, ultimately, culminates in its price dynamics. Traumatic events such as earthquakes can have a profound effect not only on a person's mental health but also on the market as a whole. Additionally, the literature has shown that legislation did not significantly affect the housing market, as most appraisers did not give any higher price on a comparable property in a lower-risk area.

2. How do residents and investors in Lisbon perceive risk?

Risk perception and awareness is a complex matter influenced by several factors, such as age, marital status, cultural environment, media exposure, and previous experiences. Although natural disasters such as earthquakes can lead to traumatic shocks and, in some cases, to prolonged post-traumatic stress disorders, panic attacks, and recurrent nightmares, their shock effect usually fades away with time. That means that the longer it has passed from an event, the less memory most residents have and the less effect it has on the market. Even the initial shock may not necessarily lead to price differentials, as it is also limited by the market itself. In the case of Lisbon, the author found differences in perception and action towards risk across different socioeconomic groups, mainly depending on their marital status, age, education level and number of children. It was also observed a slightly fatalistic attitude towards earthquake risk, which has not been observable in the case of home robberies, for instance. The results also hint at some unfamiliarity with building codes and construction practices. Additionally, it is concerning that a high percentage of respondents do not trust the government to be able to take care of them and their families in the case of a natural disaster. Overall, the study suggests a need for greater awareness and education on the subject of natural hazards and risk perception.

3. Does their risk perception affect their willingness to pay for property?

Based on the literature review, the survey, and the analysis of the transaction database, it is apparent that investors and homebuyers are not currently factoring in seismic risk when determining the market value of a property. The research also shows that the market price is mostly defined by the socioeconomic conditions of households, not by their perception and attitude towards risk. The analysis of sale prices shows that some of the most expensive locations are also the ones with the highest vulnerability to natural hazards, mainly those in the city centre, indicating that investors and homebuyers alike are oblivious to earthquake risk. The study also found that the proximity to business and education facilities is highly valued. Although the survey results have shown that most respondents were willing to pay more for a more resistant building to seismic events, the regression analysis did not find a clear correlation between market price and seismic risk. Furthermore, it is crucial to note that most respondents accounted the municipality liable for, during the licensing process, and in addition to the contractor, verifying and certifying the structural resilience of the building.

In light of these answers, some policy recommendations can be made, amongst which we may highlight:

- Cities must treat risk communication seriously and systematically, even though earthquake events may have a low likelihood. This needs to be done close to the population by disclosing risk indicators or hazard maps and through education. The effect of Out-Of-Market events may add the disclosed information, increasing awareness, which should translate into an increase in the perceived vulnerability of its inhabitants, which, in turn, should be reflected in the price, even if an event does not occur in the location addressed;
- Due to the increased risk, inspectors should constantly monitor buildings constructed or renovated in high-risk zones. This is because calculation mistakes or a lack of attention to detail throughout the construction process can be made worse in these situations. Such areas can additionally necessitate additional mitigation measures, which should be expressly included in land-use municipal plans and laws;
- The municipality and the other real estate stakeholders (such as developers and brokers, appraisers, construction firms, neighbours' associations, academic institutions, and other governmental agencies) should also work together. A seismic resilience committee should be established to develop and oversee policy proposals that coordinate all the incentives from multiple sources and encourage the adaption of mitigation measures. To ensure the efficacy of such programs, the outcomes of each measure should be carefully examined and revised;
- Establishing disclosure standards and implementing laws requiring sellers and landlords to disclose a property's seismic vulnerability can raise awareness among prospective purchasers and tenants. A standardized system of seismic risk ratings or certificates, similar to energy efficiency certificates, can accomplish this.
- Wealth and income are deterrents to adopting structural retrofitting, so local and national governments should make grants and finance more accessible. Giving grants, low-interest loans, or financial assistance to property owners, particularly those with limited resources, can help them overcome financial obstacles to retrofitting or putting seismic resilience measures in place. The availability of these support programs can be increased by working with financial institutions and looking for outside funding sources;
- Governments should collaborate with academic institutions to offer specialized training in earthquake resilience and rehabilitation for commercial construction and design firms. Companies that are accredited should receive tax incentives, such as paying less taxes for each certified expert they hire. Professionals and businesses would be encouraged to expand their knowledge, which would eventually lead to a growth in the number of businesses with specialized knowledge in rehabilitation, which would raise the market's supply and lower the price;

The study of Lisbon's reality has contributed to deepening the study of the effects of earthquake perception on the Portuguese capital city's property market. However, these six highlighted recommendations can be applied not only in Lisbon but in all locations with moderate seismicity activity. As we have seen, the perception of risk varies depending on the frequency of events. However, we can argue that households should respond similarly in risky locations with no recent memory of significant earthquake events. Similar locations could benefit from the findings presented in this work across Spain, Italy, the Balkans, and broadly, across the world. One can even argue that these conclusions may contribute to the overall knowledge of the household's perception and action towards all high-risk low-probability natural hazards. Ultimately, this study contributes to the broader understanding of household perception and action towards seismic risk and serves as a valuable resource for future research and decision-making in disaster-prone areas globally.

6.2. Main Contributions

This research has significantly contributed to understanding the relationship between real estate valuation and seismic risk.

First, various econometric models were adopted, including FGLS (Feasible Generalized Least Squares), PooledOLS (Pooled Ordinary Least Squares), Fixed-Effects, MARS (Multivariate Adaptive Regression Splines), and XGBoost. These models were applied to real transactional data, allowing for an implicit and comprehensive analysis of the relationship between seismic risk and property values. Leveraging large-scale datasets permitted to capture the intricate dynamics and complexities of real estate markets in the presence of seismic risk.

Furthermore, the employment of Contingent Valuation Methods (CVM) as an explicit approach allowed to assess the willingness-to-pay of property owners to reduce seismic risk. By conducting surveys and analyzing the responses, the author was able to quantify the economic value assigned to risk reduction measures.

Applying these methods provided valuable insights into quantifying the impact of seismic risk on real estate values. Through rigorous analysis and statistical techniques, this study identified key variables and factors significantly influencing property values in seismic-prone areas.

6.3. Future Developments

The progress made in this study has provided an opportunity to formulate public policies to mitigate earthquake risk. However, it also raised several questions that should be further explored in future developments:

- What specific socioeconomic factors contribute to the differences in risk perception and action observed across different groups in Lisbon? How can these factors be addressed to promote a more comprehensive understanding and response to seismic risk?

- What are the specific reasons for the fatalistic attitude observed towards earthquake risk, as opposed to other risks like home robberies? Are there cultural or psychological factors at play that can be further explored?
- What are the potential implications of a lack of trust in the government's ability to respond to natural disasters? How can this trust be rebuilt or strengthened to ensure effective disaster response and risk reduction?
- What are the potential implications of the market's indifference to seismic risk on long-term urban planning, infrastructure development, and disaster resilience efforts? How can stakeholders and policymakers address this gap to ensure a more resilient built environment?
- Are there any potential policy or regulatory interventions that could incentivize or require the consideration of seismic risk in property pricing? How can market mechanisms be aligned with risk mitigation goals?

The answers to these questions could build on this work and increase understanding of households' perceptions of risk. This understanding could then be used to design public policies aligned with stakeholders' incentives. Specifically, one could explore the factors that shape households' perceptions of risk and how these perceptions impact their behaviour. One could also examine how different social groups perceive risk and how this varies across different contexts. Additionally, one could investigate the relationship between risk perception and trust in institutions and the role that communication plays in shaping risk perceptions or conduct a contrafactual analysis to isolate vulnerability's effect. By gaining a deeper understanding of these issues, we could develop more effective public policies that better serve the needs of all stakeholders.

References

Abdallah, W., Goergen, M., & O'Sullivan, N. (2015). Endogeneity: how failure to correct for it can cause wrong inferences and some remedies (Version 1). Loughborough University. https://hdl.handle.net/2134/17490

Abdi, H., & Williams, L.J. (2010). Principal component analysis. WIREs Comp Stat, 2: 433-459. https://doi.org/10.1002/wics.101

Afonso, N. (2006) SIMULADOR DE CENÁRIOS SÍSMICOS EM AMBIENTE WEBSIG. Instituto Superior Técnico.

Agnew, K., & Ronan L. (2017). The Impact of Employment on Housing Prices: Detailed Evidence from FDI in Ireland. Trinity Economics Papers tep0417, Trinity College Dublin, Department of Economics.

Amore, A., Bernardi, C., & Arvanitis, P. (2020): The impacts of Airbnb in Athens, Lisbon and Milan: a rent gap theory perspective. Current Issues in Tourism. DOI: 10.1080/13683500.2020.1742674

Anexo Nacional do EC8 (2009). Documento de trabalho da CT115, LNEC, Lisboa, Fevereiro.

Antunes, G., & Seixas, J. (2020). Housing Market Access in the Lisbon Metropolitan Area Between the Financial and the Pandemic Crises. Critical Housing Analysis. Volume 7. Pp 58-72. DOI: 10.13060/23362839.2020.7.2.515.

Anselin, L. (1995) Local Indicators of Spatial Association—LISA. Geographical Analysis. Volume 27, Issue 2, pp 93-115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Anselin, L. (1988) Spatial econometrics: Methods and models. Dordrecht: Kluwer Academic.

Appleton J. (2001). O Megasismo de Lisboa no século XXI ou Vulnerabilidade sísmica do Parque Edificado de Lisboa. In: Redução da Vulnerabilidade Sísmica do Edificado, Eds. SPES e GECoRPA, Lisboa, pp. 95-104.

Appleton, J. (2005) Rehabilitation of Gaioleiro buildings, 1st ed. Orion Editors, Lisbon

Asgary, A., & Willis, K. (1997). Household Behaviour in Response to Earthquake Risk: An Assessment of Alternative Theories. Disasters, Volume 2, pp 354-365. DOI: 10.1111/1467-7717.00067

Asgary, A., & Willis, K. G. (1997b). Estimating the benefits of construction measures to mitigate earthquake risks in Iran. Environment and Planning B: Planning and Design, Volume 24(4), pp. 613-624.

Asgary, A., Jason, L., & Nader, M. (2007). Estimating Willingness to Pay for a Hypothethical Earthquake Early Warning System. Environmental Hazards Volume 7, pp. 312-320, DOI: 10.1016/j.envhaz.2007.09.003

Asgary, A., & Halim, M. (2011). Measuring people's preferences for cyclone vulnerability reduction measures in Bangladesh. Disaster Prevention and Management. Volume 20, pp.186-198. DOI: 10.1108/09653561111126111.

Arrow, K.J. (1965) Aspects of the Theory of Risk Bearing. Yrjo Jahnssonin Saatio, Helsinki.

Associação Lisbonense de Proprietários (2011). Retrato da habitação em Portugal. Características e Recomendações. Available at: https://observatorio-lisboa.eapn.pt/ficheiro/habitacaoportugal.pdf. (last visited 15 September 2021)

Autoridade Tributária e Aduaneira (2020). Código do Imposto sobre o Rendimento das Pessoas Singulares. Available at: https://info.portaldasfinancas.gov.pt/pt/informacao_fiscal/codigos_tributarios/Cod_download/Documents/CIRS.pdf Last visited October 18, 2021

Badri, S. A., Asgary, A., Eftekhari, A. R., & Levy, J. (2006). Post-disaster resettlement, development and change: a case study of the 1990 Manjil earthquake in Iran. Disasters, Volume 30(4), pp. 451-468. DOI: 10.1111/j.0361-3666.2006.00332.x

Balchin, P. N., & Kieve, J.L. (1986). Urban land economics (3ed) London: McMillan, 1986.

Bennett, M. J. (1989). Lessons Learned from the Loma Prieta, California, Earthquake of October 17, 1989. Volume 1045. US Government Printing Office.

Beron, K. J., Murdoch, J. C., Thayer, M. A., & Vijverberg. W. (1997). An Analysis of the Housing Market before and after the 1989 Loma Prieta Earthquake. Land Economics, Volume 73(1), pp. 101–113. DOI: 10.2307/3147080

Bin, O., & Polasky, S. (2004). Effects of flood hazards on property values: Evidence before and after hurricane Floyd. Land Economics, 80(4), 490-500. DOI: 10.2307/3655805

Bin, O., Kruse, J. B., & Landry, C. E. (2008). Flood hazards, insurance rates, and amenities: Evidence from the coastal housing market. Journal of Risk and Insurance, Volume 75(1), pp. 63-82. DOI: 10.1111/j.1539-6975.2007.00248.x

Bommer, J. J., Magenes, G., Hancock, J., & Penazzo P. (2004) The influence of strong-motion duration on the seismic response of masonry structures. Bulletin of Earthquake Engineering. Volume 2, No. 1, pp. 1-26. DOI: 10.1023/B:BEEE.0000038948.95616.bf

Borges, J., Fitas, A., Bezzeghoud, M., & Teves-Costa, P. (2001). Seismotectonics of Portugal and its adjacent Atlantic area. Tectonophysics. Volume 331(4), pp. 373-387. DOI: 10.1016/S0040-1951(00)00291-2

Bradshaw, S. (2003). Handbook for estimating the socio-economic and environmental effects of disasters. United Nations. Available at: https://www.cepal.org/en/publications/2782-handbook-estimating-socio-economic-and-environmental-effects-disasters (last visited May 25, 2023)

Braga. F.M. (2013). Mercado Imobiliário em Portugal. JURISMAT, Portimão, n.º 2, 2013, pp. 199-216.

Braga, Lucas. (2017). Immanuel Kant and his writings on the Lisbon earthquake of 1755. Conference: VI Congresso Internacional UFES/Paris-EstAt: Vitória

Brooks, R. M., Patel, A., & Su, T. (2003). How the Equity Market Responds to Unanticipated Events. The Journal of Business, Volume 76(1), pp. 109–133. DOI: 10.1086/344115

Brookshire, D. S., Thayer, M. A., Schulze, W. D., & d'Arge, R. C. (1982). Valuing Public Goods: A Comparison of Survey and Hedonic Approaches. The American Economic Review. Volume 72(1), pp. 165–177. http://www.jstor.org/stable/1808583

Brookshire, D. S., Thayer, M. A., Tschirhart, J., & Schulze, W. D. (1985). A Test of the Expected Utility Model: Evidence from Earthquake Risks. Journal of Political Economy, Volume 93(2), pp. 369–389. DOI: 10.1086/261304

Burton, I., Kates, R.W., & White, G.F. (1978). The Environment as Hazard. Oxford University Press. Oxford, 240pp.

Burrus, R., Graham Jr.,T., Hall, E., William, W., & Schuhmann, P. (2009). Home-buyer Sentiment and Hurricane Landfalls. Hurricane Floyd Symposium, East Carolina University, September 2009. Available at: https://thescholarship.ecu.edu/handle/10342/1734 (last visited October 21, 2021)

Caldeira, B., Fontiela, J., Borges, J., & Bezzeghoud, M. (2018). Large earthquakes in the Azores - Grandes terremotos en Azores. Fisica de la Tierra. Volume 29, pp. 29-45. DOI: 10.5209/FITE.57601

Câmara Municipal de Lisboa (CML) (2011). Estratégia de Reabilitação urbana de Lisboa 2011/2024. Available at: https://www.lisboa.pt/fileadmin/cidade_temas/urbanismo/reabilitacao_Urbana/documentos/estrategia_reabilitacao.pdf (last visited November 14, 2022)

Câmara Municipal de Lisboa (CML) (2023) Apresentação da Carta Municipal de Habitação 2022-2023. Available at: https://www.lisboa.pt/fileadmin/atualidade/noticias/user_upload/CMH_REUNIAO_6_MARCO_PRIORIDADES_DA_CARTA.pdf (last visited April 14, 2023)

Carson, R.T., Flores, N.E., Martin, K.M. & Wright, J.L. (1996). Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. Land Economics. Volume 72(1), pp. 80-99. DOI: 10.2307/3147159

Carson, R. T. (2000). Contingent Valuation: A User's Guide. Environmental Science & Technology, Volume 34, pp. 1413-1418. DOI: 10.1021/es990728j

Carson, R.T., & Louviere, J.J. (2011). A Common Nomenclature for Stated Preference Elicitation Approaches. Environmental and Resource Economics.Volume 49, pp. 539-559. DOI: 10.1007/s10640-010-9450-x

Caruso, C., & Bento, R. (2019). Seismic assessment and strengthening of wall-frame RC building through a case study in Lisbon. 7th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering Methods in Structural Dynamics and Earthquake Engineering. pp. 2983-2992. DOI: 10.7712/120119.7126.18525.

Carvalho, A, Coelho, E., Campos Costa, A., Sousa, M.L., & Candeias, P. (2002). Vulnerability Evaluation of Residential Buildings in Portugal. Proceedings Londres. Proceedings da 12ª European Conference on Earthquake Engineering. Referência artigo 696, Londres, Reino Unido, Elsevier Science Ltd.

Cavallo, E., & Noy, I. (2011). Natural disasters and the economy—a survey. International Review of Environmental and Resource Economics. Volume 5(1), pp. 63-102. DOI: 10.1561/101.00000039

Century21 (2020). ESTUDO SOBRE O IMPACTO DA COVID-19 NOS CRITÉRIOS DA HABITAÇÃO EM PORTUGAL. Available at: https://www.century21.pt/observatorio-imobiliario/ . (last visited 15 September 2021)

Cockburn Association (2018). The Airbnb phenomenon: Impact and opportunities of the collaborative economy and disruptive technologies: How should Edinburgh respond to short-term letting?. Available at: https://greens.scot/sites/default/files/201803114_CockburnAssoc_AirbnbPhenomenonEvent.pdf (last visited May 25, 2023)

Cócola-Gant, A., & Gago, A. (2019). O alojamento local e a gentrificação turística em Alfama, Lisboa. In: Santos, Ana Cordeiro (org.). A nova questão da habitação em Portugal: uma abordagem de economia política. Lisboa: Actual, pp. 143-170. ISBN: 9789896943912

Cho, D., & Ma, S. (2006). Dynamic Relationship between Housing Values and Interest Rates in the Korean Housing Market. Journal of Real Estate Finance and Economics. 32, 169–184. DOI: 10.1007/s11146-006-6013-6

Cliff, A. C., & Ord, J. K. (1973). Spatial autocorrelation. Pion Limited, London.

Col, J. M. (2007). Managing disasters: The role of local government. Public administration review, Volume 67, pp. 114-124. DOI: 10.1111/j.1540-6210.2007.00820.x

Costa, A., Sousa, M., Carvalho, A., Bilé, J.B, & Martins, A. (2004). Simulador de cenários sísmicos integrado num sistema de informação geográfica. SÍSMICA 2004 - 6º Congresso Nacional de Sismologia e Engenharia Sísmica.

Costa, A.C., Sousa, M.L., Carvalho, A., & Coelho, E. (2008). Seismic Loss Scenarios Based on Hazard Disaggregation. Application to the Metropolitan Region of Lisbon, Portugal. In: Oliveira, C.S., Roca, A., Goula, X. (eds) Assessing and Managing Earthquake Risk. Geotechnical, Geological And Earthquake Engineering. vol 2. Springer, Dordrecht. DOI: 10.1007/978-1-4020-3608-8_21

Cupal, M. (2017). Sales Comparison Approach Indicating Heterogeneity of Particular Type of Real Estate and Corresponding Valuation Accuracy. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis. Volume 65, pp.977-985. DOI: 10.11118/actaun201765030977.

Cuthbertson, K., Hall, S. G., & Taylor, M. P (1992). Applied econometric techniques. New York: Harvester Wheatsheaf, 1992.

Daniel, V. E., Florax, R. J., & Rietveld, P. (2009). Floods and Residential Property Values: A Hedonic Price Analysis for the Netherlands. Built Environment (1978-), Volume 35(4), pp. 563–576. DOI: 10.2148/benv.35.4.563

DECO (2013). Lar, Caro Lar. Dinheiro&Direitos 116. Available at: https://tinyurl.com/4frrfz7p (last visited October 20, 2021)

Deichmann, J., Eshghi, A., Haughton, D., Sayek, S., & Teebagy, N. (2002). Application of multiple adaptive regression splines (MARS) in direct response modeling. Journal of Interactive Marketing. Volume 16, pp. 15-27. DOI: 10.1002/dir.10040.

Deng, G., Gan, L., & Hernandez, M. A. (2013). Do People Overreact? Evidence from the Housing Market after the Wenchuan Earthquake. NBER Working Paper No. w19515, DOI: 10.3386/w19515

Des Rosiers, F., & Thériault, M. (2006). Mass Appraisal, Hedonic Price Modelling and Urban Externalities: Understanding Property Value Shaping Processes. Discussion Paper presented at the Advances in Mass Appraisal Methods Seminar. Delft University of Technology, October 30-31, 2006. DOI: 10.1002/9781444301021.ch6

Diamond, P. A., & Hausman, J. A. (1994). Contingent Valuation: Is Some Number better than No Number? The Journal of Economic Perspectives. Volume 8(4), pp. 45–64. DOI: 10.1257/jep.8.4.45

Donovan, G.H., Champ, P.A., & Butry, D.T. (2007). Wildfire risk and housing prices: a case study from Colorado Springs. Land Economics. Volume 83(2), pp. 217-233. DOI: 10.3368/le.83.2.217

Dooley, D., Catalano, R., Mishra, S., & Serxner, S. (2006). Earthquake Preparedness: Predictors in a Community Survey1. Journal of Applied Social Psychology. Volume 22, pp. 451-470. DOI: 10.1111/j.1559-1816.1992.tb00984.x

Dubin, R. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. Review of Economics and Statistics. Volume 70, pp. 466-474. DOI: 10.2307/1926785

Dunn, R. (2014). Re: Why do we often use a GMM approach?. Retrieved from: https://www.researchgate.net/post/Why_do_we_often_use_a_GMM_approach/53231e72d11b8bf7698b4575/citation/download (last visited May 29, 2023)

Earnhart, D. (2001). Combining revealed and stated preference methods to value environmental amenities at residential locations. Land economics, Volume 77(1), pp. 12-29. DOI: 10.2307/3146977

Ederington, L. H., & Lee, J. H. (1993). How Markets Process Information: News Releases and Volatility. The Journal of Finance. Volume 48(4), pp. 1161–1191. DOI: 10.2307/2329034

Ederington, L. H., & Lee, J. H. (1995). The Short-Run Dynamics of the Price Adjustment to New Information. The Journal of Financial and Quantitative Analysis, Volume 30(1), pp. 117–134. DOI: 10.2307/2331256

Elsinga, M., & Hoestra, J. (2005) Homeownership and housing satisfaction. Journal of Housing and the Built Environment. Volume 20, pp. 401–424. DOI 10.1007/s10901-005-9023-4

European Central Bank (ECB) (2021). Bank lending survey (BLS). Available at: https://www.ecb.europa.eu/stats/ecb_surveys/bank_lending_survey/html/ecb.blssurvey2021q2~b868c78ada.en.html (last visited October 18th 2021)

European Committee for Standardization (CEN) (2004) Eurocode 8: design of structures for earthquake resistance — part 1: general rules, seismic actions and rules for buildings (EC8-1)

Ewing, B.T., Kruse, J.B. & Wang, Y. (2007). Local housing price index analysis in wind-disaster-prone areas. Natural Hazards, Volume 40, pp. 463–483. DOI: 10.1007/s11069-006-9005-1

Fekrazad, A. (2019). Earthquake-risk salience and housing prices: Evidence from California. Journal of Behavioral and Experimental Economics (formerly The Journal of Socio-Economics), Elsevier, Volume 78(C), pp. 104-113. DOI: 10.1016/j.socec.2019.01.001

Ferreira, M. A. (2012). Risco Sísmico em Sistemas Urbanos. Instituto Superior Técnico. Available at: http://www.civil.ist.utl.pt/~monicaf/Tese_MAF.pdf (last visited November 5, 2021)

Florax, D. J., Travisi, A. M., & Nijkamp, P.(2005) A Meta-Analysis of the Willingness to Pay for Reductions in Pesticide Risk Exposure. European Review of Agricultural Economics, Volume 32(4), pp. 441-467. DOI: 10.1093/erae/jbi025

Francisco, S. (2018, October 18). Veja os bairros históricos de Lisboa onde já não pode haver mais alojamentos locais. Diário de Notícias. Available at: https://www.dn.pt/poder/veja-os-bairros-historicos-de-lisboa-onde-ja-nao-pode-haver-mais-alojamentos-locais-10022441.html (last visited May 26, 2023)

Friedman, M., & Savage, L. J. (1948). The Utility Analysis of Choices Involving Risk. Journal of Political Economy, Volume 56(4), pp. 279–304. DOI: 10.1086/256692

Friedman, J. H. (1991) Multivariate Adaptive Regression Splines. Ann. Statist. Volume19(1), pp. 1-67, March. DOI: 10.1214/aos/1176347963

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics, Volume 29(5), pp. 1189–1232. DOI: 10.1214/aos/1013203451

Gallimore, P., Fletcher, M., & Carter, M. (1996) Modelling the influence of location on value. Journal of Property Valuation and Investment, Volume 14(1), pp.6-19. DOI: 10.1108/14635789610107444

Gallup Organization (2007) Flash Eurobarometer Series #202: Youth Survey among People Aged between 15-30, in the European Union. Available at: https://www.jugendpolitikineuropa.de/downloads/22-177-510/eurobaro%20youth.pdf (Last visited April 27th 2022)

Garha, N. S., & Azevedo, A.B. (2021). Population and Housing (Mis)match in Lisbon, 1981–2018. A Challenge for an Aging Society. Social Sciences, Volume 10, pp. 102. DOI: 10.3390/socsci10030102

Ghesquiere, F., & Mahul, O. (2010). Financial Protection of the State Against Natural Disasters: A Primer World Bank Policy Research Working Paper No. 5429. DOI: 10.1596/1813-9450-5429

Goda K., Wenzel F., & Daniell, J. (2014) Insurance and Reinsurance Models for Earthquake. In: Beer M., Kougioumtzoglou I., Patelli E., Au IK. (eds) Encyclopedia of Earthquake Engineering. Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-36197-5_261-1

Gomes, R., (2011). Sistema Estrutural de Edifícios Antigos de Lisboa - Os Edifícios “Pombalinos” e os Edificíos “Gaioleiros.”. Instituto Superior Técnico.

González, M. A. S., & Formoso, C. T. (2000). Análise conceitual das dificuldades na determinação de modelos de formação de preços através de análise de regressão. Engenharia Civil. Um, Volume 8, pp. 65-75.

Gow, I. D., Ormazabal, G., & Taylor , D. J. (2010). Correcting for cross-sectional and time-series dependence in accounting research, Accounting Review, Volume 85, pp. 483– 512. DOI: 10.2308/accr.2010.85.2.483

Grünthal, G. (1998) European Macroseismic Scale 1998, Cahiers du Centre Européen de géodynamique et de séismologie, Conseil de L’Europe, Report Luxembourg.

Hallstrom, D. G., & Smith, K. (2005) Market responses to hurricanes, Journal of Environmental Economics and Management, Volume 50(3), pp. 541-561. DOI: 10.1016/j.jeem.2005.05.002

Hanley, N., Wright, R.E., & Adamowicz, V. (1998). Using Choice Experiments to Value the Environment. Environ Resource Econ. Volume 11, pp. 413–428. DOI: 10.1023/A:1008287310583

Harrison, D., Smersh, G. & Schwartz Jr., A. (2001). Environmental Determinants of Housing Prices: The Impact of Flood Zone Status. Journal of Real Estate Research. Volume 21, pp. 3-20. DOI: 10.1080/10835547.2001.12091045.

Hidano, N., Hoshino, T., & Sugiura, A. (2015). The effect of seismic hazard risk information on property prices: Evidence from a spatial regression discontinuity design, Regional Science and Urban Economics, Volume 53, pp. 113-122, ISSN 0166-0462, DOI: 10.1016/j.regsciurbeco.2015.05.005.

Hoffower, H. (2019). Forget San Francisco — these 6 global cities have thriving tech hubs that could make them the next Silicon Valley. Available at: https://www.businessinsider.com/silicon-valley-around-the-world-tel-aviv-berlin-stockholm-2019-6 (last visited October 14, 2021)

Hurnen, F., & McClure, John (1997) The Effect of Increased Earthquake Knowledge on Perceived Preventability of Earthquake Damage. Department of Psychology, Victoria University of Wellington. Available at: https://www.massey.ac.nz/~trauma/issues/1997-3/mcclure1.htm (last visited September 12, 2022)

Incekara-Hafalir, E., Kim, E. & Stecher, J.D. (2021). Is the Allais paradox due to appeal of certainty or aversion to zero?. Exp Econ, Volume24, pp. 751–771. DOI: 10.1007/s10683-020-09678-4

Instituto Nacional de Estatística (INE) (2019) - Retrato Territorial de Portugal. Lisboa. Available at: https://www.ine.pt/xurl/pub/358634995 . ISSN 1646-0448. ISBN 978-989-25-0490-2 (last visited May 29, 2023)

Instituto Nacional de Estatística (INE) (2020) - Inquérito à Situação Financeira das Famílias 2020. Available at: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=569549494&DESTAQUESmodo=2 (last visited May 10, 2023)

Instituto Nacional de Estatística (INE) (2023) Estatísticas do Emprego - Remuneração bruta mensal média por trabalhador. INE. Abril 2023. Available at: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=415288436&DESTAQUESmodo=2&xlang=pt (last visited April 14, 2023)

International Organization for Standardization (ISO) 31000 (2009). Risk management — Principles and guidelines

(ISO/DIS Standard No. 3001). Available at: https://www.iso.org/iso-31000-risk-management.html

Ioannides, D., Röslmaier, M., & van der Zee, E. (2018). Airbnb as an instigator of 'tourism bubble' expansion in Utrecht's Lombok neighbourhood. Tourism Geographies, Volume 21(5), pp. 1–19. DOI: 10.1080/14616688.2017.1409261

Jalali, S., & Wohlin, C. (2012, September). Systematic literature studies: database searches vs. backward snowballing. In Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement (pp. 29-38).

Januário, J.F., Costa, A., Cruz, C.O., Sarmento, J.M., & Sousa, V.F. (2021) Transport infrastructure, accessibility, and spillover effects: An empirical analysis of the Portuguese real estate market from 2000 to 2018. Research in Transportation Economics, Volume 90, 101130, ISSN 0739-8859. DOI: 10.1016/j.retrec.2021.101130.

Jarimba, A. (2016) Vulnerabilidade Sísmica dos Edifícios porticados de Betão Armado da cidade de Lisboa- décadas de 1960-80. Instituto Superior Técnico.

Judge, G.; Hill, R. C., Griffiths, W., Lütkepohl, H., & Lee, T. (1985). Introduction to the theory and practice of econometrics. 2ed. New York: John Wiley.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, Volume 47(2), pp. 263–291. DOI: 10.2307/1914185

Kahneman, D., & Knetsch, J. L (1992) Valuing public goods: The purchase of moral satisfaction, Journal of Environmental Economics and Management, Volume 22(1), pp. 57-70, ISSN 0095-0696, DOI: 10.1016/0095-0696(92)90019-S

Kajitani, Y., Chang, S. E., & Tatano, H. (2013). Economic impacts of the 2011 Tohoku-Oki earthquake and tsunami. Earthquake Spectra, Volume 29(1), pp. 457-478. DOI: 10.1193/1.400010

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. Journal of Risk and Uncertainty, Volume 38(2), pp. 95–115. DOI: 10.1007/s11166-009-9060-6

Kartal Koc, E., & Bozdogan, H. (2015). Model selection in multivariate adaptive regression splines (MARS) using information complexity as the fitness function. Machine Learning, Volume 101, pp. 35–58. DOI: 10.1007/s10994-014-5440-5

Kiel, K.A., & Matheson, V.A. (2018). The Effect of Natural Disasters on Housing Prices: An Examination of the Fourmile Canyon Fire. Journal of Forest Economics, Volume 33, pp. 1-7. DOI: 10.1016/j.jfe.2018.09.002

Kroes, E. P., & Sheldon, R. J. (1988). Stated Preference Methods: An Introduction. Journal of Transport Economics and Policy, Volume 22(1), pp. 11–25. http://www.jstor.org/stable/20052832

Kunreuther, H. (1996). Mitigating disaster losses through insurance. Journal of risk and Uncertainty, Volume 12(2), pp.171-187. DOI: 10.1007/bf00055792

Kunst, R. M. (2010). Econometric Methods for Panel Data. University of Vienna and Institute for Advanced Studies Vienna. Available Online at: https://homepage.univie.ac.at/robert.kunst/panpres.pdf , (last visited March 29, 2022)

Lagomarsino, S., & Giovinazzi, S. (2006). Macroseismic and mechanical models for the vulnerability assessment of current buildings. Bulletin of Earthquake Engineering. Volume 4, pp. 415-443. DOI:10.1007/s10518-006-9024-z.

Lamond, J., Proverbs, D., & Hammond, F. (2010) The Impact of Flooding on the Price of Residential Property: A Transactional Analysis of the UK Market, Housing Studies, Volume 25(3), pp. 335-356, DOI: 10.1080/02673031003711543

Lee, M. & Hong, J. H, & Kim, K. (2017). Estimating Damage Costs from Natural Disasters in Korea. Natural Hazards Review. Volume 18. DOI: 10.1061/(ASCE)NH.1527-6996.0000259.

Lee, D. W. (2019) Local government's disaster management capacity and disaster resilience, Local Government Studies, Volume 45(6), pp. 803-826, DOI: 10.1080/03003930.2019.1653284

Leilani, F. (2017). Report of the Special Rapporteur on adequate housing as a component of the right to an adequate standard of living, and on the right to non-discrimination in this context, New York, New York, USA: United Nations. Available at: https://digitallibrary.un.org/record/861179 (last visited May 26, 2023)

Lennartz, C., Arundel, R., & Ronald, R. (2016). Younger adults and homeownership in Europe through the global financial crisis. Population, Space and Place, Volume 22(8), pp. 823-835. DOI: 10.1002/psp.1961

Li, M., Brown, M., & James., H. (1980) Micro-neighborhood externalities and hedonic housing prices. Land Economics, Volume 56(2), pp.125-141.

Lindell, M. K., & Perry, R. W. (2012). The protective action decision model: Theoretical modifications and additional evidence. Risk Analysis: An International Journal, Volume 32(4), pp. 616-632. DOI: 10.1111/j.1539-6924.2011.01647.x

LeSage, J., & Pace, R.K., (2008). Introduction to Spatial Econometrics. Chapman and Hall/CRC. ISBN 9781420064247

Lestegás, I. (2019). Lisbon After the Crisis: From Credit-fuelled Suburbanization to Tourist-driven Gentrification. International Journal of Urban and Regional Research. Volume 43, pp. 705-723. DOI: 10.1111/1468-2427.12826.

Lestegás, I., Seixas, J., & González, R.C.L. (2019). Commodifying Lisbon: A Study on the Spatial Concentration of Short-Term Rentals. Social Sciences. Volume 8, pp.33. DOI: 10.3390/socsci8020033.

Lopes, L. L. (1994). Psychology and economics: Perspectives on risk, cooperation, and the marketplace. Annual Review of Psychology, Volume 45, pp. 197. DOI: 10.1146/annurev.ps.45.020194.001213

Lopes, M. (2008) Sismos e Edifícios. Portugal, Edição Orion. ISBN: 9789728620110

Lusa (2021, August 10th). Custos de construção de habitação nova cresceram 6,5% em junho. Available at: https://www.tsf.pt/portugal/economia/custos-de-construcao-de-habitacao-nova-cresceram-65-em-junho-14021008.html (last visited October 18, 2021)

Luz, R. (2020) Análise geoespacial dos determinantes de preço do mercado imobiliário em Lisboa. Instituto Superior Técnico. Universidade de Lisboa. Available at: https://fenix.tecnico.ulisboa.pt/cursos/mec/dissertacao/1128253548922544

MacDonald, D. N., Murdoch, J. C., & White, H. L. (1987). Uncertain Hazards, Insurance, and Consumer Choice: Evidence from Housing Markets. Land Economics, University of Wisconsin Press, Volume 63(4), pp 361-371. DOI: 10.2307/3146293

Machina, M.J. (2008). Non-Expected Utility Theory. In: Durlauf, S.N., Blume, L.E. (eds) The New Palgrave Dictionary of Economics. Palgrave Macmillan, London. DOI: 10.1007/978-1-349-58802-2_1189

Malheiros, J. M., & Vala, F. (2004) Immigration and city change: The Lisbon metropolis at the turn of the twentieth century, Journal of Ethnic and Migration Studies, Volume 30:6, pp. 1065-1086, DOI: 10.1080/1369183042000286250

Manganelli, B., Vona, M., & De Paola, P. (2018). Evaluating the cost and benefits of earthquake protection of buildings. Journal of European Real Estate Research, Volume 11(2), pp.263-278, DOI: 10.1108/JERER-09-2017-0029

Mantel, N. (1970) Why Stepdown Procedures in Variable Selection, Technometrics, Volume 12(3), pp. 621-625, DOI: 10.1080/00401706.1970.10488701

Marinheiro, C. F. (2013). A Política Orçamental em Portugal entre Duas Intervenções do FMI: 1986-2010. Grupo de Estudos Monetários e Financeiros (GEMF). Faculdade de Economia da Universidade de Coimbra. Available at: http://www4.fe.uc.pt/gemf/workingpapers/pdf/2013/gemf_2013-25.pdf (last visited October 14, 2021)

Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, Volume 7(1), pp. 77–91. DOI: 10.2307/2975974

Martinotti, G. (1993). Metropoli: La nuova morfologia sociale della città. Metropolis: The New Social Morphology of the City. Bologna: Il Mulino.

Marques, R., Lamego, P., Lourenço, P.B., & Sousa, M.L. (2018) Efficiency and Cost-Benefit Analysis of Seismic Strengthening Techniques for Old Residential Buildings in Lisbon, Journal of Earthquake Engineering, Volume 22(9), pp. 1590-1625, DOI: 10.1080/13632469.2017.1286616

McGinnis, A. (2004). The Effect of Earthquake Risk on the San Francisco Bay Area Housing Market: How Has it Changed Since the Loma Prieta and Northridge Earthquakes? Available at: https://api.semanticscholar.org/CorpusID:1581314 (last visited October 22, 2021)

Mendes, L. (2013). Public policies on urban rehabilitation and their effects on gentrification in Lisbon. AGIR. Revista Interdisciplinar de Ciências Sociais e Humanas, vol. 1, n.º 5, pp.200-218.

Mendes, L. (2017). Gentrificação turística em Lisboa: neoliberalismo, financeirização e urbanismo austeritário em tempos de pós-crise capitalista 2008-2009. Cad. Metrop., São Paulo, Volume 19(39), pp. 479-512, maio/ago 2017 DOI: 10.1590/2236-9996.2017-3906

Mendes, L. (2018). 'Gentrification and the New Urban Social Movements in Times of Post-capitalist Crisis and Austerity Urbanism in Portugal.' Arizona Journal of Hispanic Cultural Studies Volume 22, pp. 199-215. DOI: 10.1353/hcs.2018.0013

Mendes, Luis (2021): Transnational gentrification and the housing market during pandemic times, Lisbon style, Urban Geography, Volume 42(7), pp. 1003-1010. DOI: 10.1080/02723638.2021.1949832

Mendes, N., & Lourenço, P. (2010) Seismic Assessment of Masonry "Gaioleiro" Buildings in Lisbon, Portugal. Journal of Earthquake Engineering. Volume 14(1), pp. 80-101. DOI: 10.1080/13632460902977474

Mermet, A. C. (2017). Airbnb and tourism gentrification: critical insights from the exploratory analysis of the 'Airbnb syndrome'in Reykjavik. In Tourism and gentrification in contemporary metropolises, pp. 52-74. Routledge. ISBN: 9781315629759

Michel-Kerjan, E., Hochrainer-Stigler, S., Kunreuther, H., Linnerooth-Bayer, J., Mechler, R., Muir-Wood, R., & Young, M. (2013). Catastrophe risk models for evaluating disaster risk reduction investments in developing countries. Risk Analysis, Volume 33(6), pp. 984-999. DOI: 10.1111/j.1539-6924.2012.01928.x

Miller, S., & Startz, R. (2018) Feasible Generalized Least Squares Using Machine Learning. Available at https://ssrn.com/abstract=2966194 or DOI: 10.2139/ssrn.2966194 (last visited August 17, 2022)

Mitchell, R.C., & Carson, R.T. (1989) Using Surveys to Value Public Goods: The Contingent Valuation Method. Resources for the Future, Washington DC. ISBN 9780915707324

Modica, M., & Locati, M. (2016). Housing Market Response to 2012 Northern Italy Earthquake: The role of house quality and changing risk perception. SEEDS Working Papers. 4/2016. Available at: http://www.sustainability-seeds.org/papers/RePec/srt/wpaper/0416.pdf (last visited June 26th, 2023)

Modica, M., Zoboli, R., Meroni, F., Pessina, V., Squarcina, T., & Locati, M. (2021). 'Near miss' housing market response to the 2012 northern Italy earthquake: The role of housing quality and risk perception. Urban Studies, Volume 58(11), pp. 2293–2309. DOI: 10.1177/0042098020943490

Morás, B. (2018). Os custos da construção residencial em Portugal e no Brasil: Caso de estudo. Escola Superior de Tecnologia e Gestão - Instituto Politécnico de Bragança. Available at: https://bibliotecadigital.ipb.pt/bitstream/10198/19302/1/pauta-relatorio-11.pdf (last Visited October 18, 2021)

Mota de Sá, F., Oliveira, C. S., & Ferreira, M. A. (2010). SIRIUS - INDICADOR DE RISCO SÍSMICO NO ESPAÇO URBANO. SÍSMICA 2010 – 8º Congresso de Sismologia e Engenharia Sísmica. Conference Paper. Available at: http://www.civil.ist.utl.pt/~monicaf/SISMICA_104.pdf (last visited November 5th 2021)

Mota de Sá, F., Oliveira, C. S., & Ferreira, M. A. (2013). SIRIUS, Seismic Risk Indicator in Urban Space. Earthquake Spectra, Volume 29(2), pp. 573–595. DOI: 10.1193/1.4000149

Mulder, C. H. (2006). Population and housing: A two-sided relationship. Demographic Research Volume 15, pp. 401–12. DOI: 10.4054/DemRes.2006.15.13

Nakagawa, M., Saito, M., & Yamaga, H., (2007). Earthquake risk and housing rents: Evidence from the Tokyo Metropolitan Area. Regional Science and Urban Economics, Elsevier, Volume 37(1), pp. 87-99. DOI: 10.1016/j.regsciurbeco.2006.06.009

Nakagawa, M., Saito, M., & Yamaga, H. (2009). Earthquake Risks And Land Prices: Evidence From The Tokyo Metropolitan Area. The Japanese Economic Review, Japanese Economic Association, Volume 60(2), pp. 208-222. DOI: 10.1111/j.1468-5876.2008.00438.x

Naoi, M., Seko, M., & Sumita, K. (2009) Earthquake risk and housing prices in Japan: Evidence before and after massive earthquakes. Regional Science and Urban Economics, Volume 39(6), pp. 658-669, ISSN 0166-0462, DOI: 10.1016/j.regsciurbeco.2009.08.002.

Newel (2019). How Portugal engineered a remarkable recovery. Available at: https://www.worldfinance.com/markets/how-portugal-engineered-a-remarkable-recovery (last visited Ocotober 14 2021)

Neves, V. (2016) Edifícios Gaioleiros da cidade de Lisboa. Avaliação Sísmica de um edifício tipo. Instituto Superior Técnico. Available at: https://fenix.tecnico.ulisboa.pt/downloadFile/281870113703470/Dissertacao%20VN.pdf . (last visited September 15th, 2021)

Norstad, J. (1999). An introduction to utility theory. Unpublished manuscript at http://homepage. mac. com/j. norstad.

Norme Tecniche per le Costruzioni (NTC) (2008). Decreto del Ministero delle infrastrutture, Supplemento Ordinario n.30 alla Gazzetta Ufficiale della Repubblica italiana, 29, 4/02/2008, Italy.

Oliveira, C.S., Ferreira, M.A., & Mota de Sá, F. (2004) Seismic Vulnerability and Impact Analysis: Elements for Mitigation Policies. XI Congresso Nazionale "L'ingegneria Sismica in Italia", Genova 25-29 gennaio 2004.

Oliveira, L., Teves-Costa, P., Gomes, R., Pinto, C., Almeida, I., Pereira, T., & Sotto-Mayor, M. (2019). Microzonagem sísmica baseada em dados de sondagens geotécnicas: aplicação a Lisboa. Conference: 11º Congresso Nacional de Sismologia e Engenharia Sísmica. Instituto Superior Técnico

Önder, Z., Dökmeci, V., & Keskin, B. (2004). The Impact of Public Perception of Earthquake Risk on Istanbul's Housing Market. Journal of Real Estate Literature, Volume 12(2), pp.181–194. DOI: 10.1080/10835547.2004.12090136

Ordem dos Arquitectos Secção Regional Sul (OASRS)(2016) Reabilitação e Conservação do Património Edificado - cadernos técnicos. Ordem dos Arquitectos – Secção Regional do Sul. Dezembro 2016. Available at: https://oasrs.org/media/uploads/4_CT_Conservacao.pdf (Last visited November 15, 2022)

Organisation for Economic Co-operation and Development (OECD) (2018) Risks that Matter Survey. Available at: https://www.oecd.org/els/soc/Risks-That-Matter-2018-Main-Findings.pdf (Last visited November 14, 2022)

Oskam, J. A. (2019). The future of Airbnb and the sharing economy: The collaborative consumption of our cities. Channel View Publications. ISBN-10: 1845416724

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: A review of valuation methods. Journal of Property Investment & Finance. Volume 21, pp. 383-401. DOI: 10.1108/14635780310483656.

Pais, I., Teves-Costa, P. & Cabral, J. (1996a) Emergency Management of Urban Systems under Earthquake Damage Scenarios. Proceedings da 11th World Conference on Earthquake Engineering, Junho 1996, edição Elsevier Science LTD, Acapulco, México.

Pais, I., Ribeiro, M.J., Oliveira, C.S., Teves-Costa, P., & Cabral, J. (1996b) Planeamento e Gestâo da Emergência Sísmica na Cidade de Lisboa. Revista do SNPC VIII, II Série, nº 9, Fevereiro/Maio 1996, pp. 28-35, Lisboa.

Pais, I., Oliveira, C.S., Mota de Sá, F., & Teves-Costa P. (2001) O Uso do Simulador de Danos Sísmicos no Planeamento e Gestão da Emergência. A Experiência da Cidade de Lisboa. Proceedings do 5ª Encontro de Sismologia e Engenharia Sísmica – Sísmica 2001, Laboratório Regional de Engenharia Civil dos Açores / SPES, Açores

Pais, I. (2002) O Simulador de Danos Sísmicos do Serviço Municipal de Protecção Civil de Lisboa. O Contributo das Tecnologias SIG para a sua Construção. SMPCL. Lisboa.

Palm, R. (1998). Urban earthquake hazards: The impacts of culture on perceived risk and response in the USA and Japan, Applied Geography, Volume 18(1), pp. 35-46, ISSN 0143-6228. DOI: 10.1016/S0143-6228(97)00044-1.

Paul, M. (2018). Regularization. Applied Machine Learning. University of Colorado Boulder. Available at: https://cmci.colorado.edu/classes/INFO-4604/files/slides-6_regularization.pdf (Last visited August 19, 2022)

Peng, T. C. (2021). The effect of hazard shock and disclosure information on property and land prices: a machine-learning assessment in the case of Japan. Review of Regional Research. Volume 41, pp. 1–32. DOI: 10.1007/s10037-020-00148-1

Perraillon, M. C. (2019). Linear Model Assumptions and Diagnostics. Health Services Research Methods I. HSMP 7607. Anschutz Medical Campus. University of Colorado. Available at: https://clas.ucdenver.edu/marcelo-perraillon/sites/default/files/attached-files/week_7_diagnostics_0.pdf (last visited August 17, 2022)

Pereira, A. S. (2009). The Opportunity of a Disaster: The Economic Impact of the 1755 Lisbon Earthquake. The Journal of Economic History, Volume 69(2), pp. 466–499. DOI: 10.1017/S0022050709000850

Pielke Jr, R. A., Gratz, J., Landsea, C. W., Collins, D., Saunders, M. A., & Musulin, R. (2008). Normalized hurricane damage in the United States: 1900–2005. Natural hazards review, Volume 9(1), pp. 29-42. DOI: 10.1061/(ASCE)1527-6988(2008)9:1(29)

Pothon, A. (2020) Seismic loss modelling in insurance industry: towards a new model for better claims management. Applied geology. Université Grenoble Alpes. Available at: https://tel.archives-ouvertes.fr/tel-02862122 (last visited November 14, 2022)

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. Econometrica,Volume 32(1/2), pp. 122–136. DOI: 10.2307/1913738

Pryce, G., Chen, Y., & Galster, G. C. (2011). The impact of floods on house prices : an imperfect information approach with myopia and amnesia. Housing Studies, Volume 26(2), pp. 259-279. DOI: 10.1080/02673037.2011.542086

PWC (2021). Guia Fiscal 2021. Available at: https://www.pwc.pt/pt/pwcinforfisco/guia-fiscal/2021/irc.html Last visited October 18, 2021

Ramsey, F. B. (1926) TRUTH AND PROBABILITY & "Further Considerations" and "Probability and Partial Belief". Foundations of Mathematics and other Logical Essays, Volume VII, pp. 156-198. Available at: http://fitelson.org/probability/ramsey.pdf (last visited November 4th, 2022)

Rahmatian, M. (2005). Contingent Valuation Method. Available at: https://iwlearn.net/resolveuid/68f0313f42a531088300e467dca9d00c (last visited November 5, 2021)

Rathnayake, D. (2019). How to decide to use panel data analysis GLS and GMM?. Retrieved from: https://www.researchgate.net/post/How-to-decide-to-use-panel-data-analysis-GLS-and-GMM/5c6baf2a11ec737f365e59e3/citation/download. (last visited May 29, 2023)

Reichert, A. K. (1990). The impact of interest rates, income, and employment upon regional housing prices. The Journal of Real Estate Finance and Economics, Volume 3, pp. 373–391. DOI: 10.1007/BF00178859

Robinson, R. (1979). Housing economics and public policy. Palgrave Macmillan London. DOI: 10.1007/978-1-349-16069-3

Rogers, A.A., Dempster, F.L., & Hawkins, J.I. (2019). Valuing non-market economic impacts from natural hazards. Natural Hazards, Volume 99, pp. 1131–1161. DOI: 10.1007/s11069-019-03761-7

Rokach, L., & Maimon, O. (2005). Decision Trees. In: Maimon, O., Rokach, L. (eds) Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA. DOI: /10.1007/0-387-25465-X_9

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy, Volume 82(1), pp. 34–55. DOI: 10.1086/260169

Romão, J., Kourtit, K., Neuts, B., & Nijkamp, P. (2018). The smart city as a common place for tourists and residents: A structural analysis of the determinants of urban attractiveness. Cities, Volume 78, pp. 67–75. DOI: 10.1016/j.cities.2017.11.007

Rossetto, T., Ioannou, I., Grant, D.N. & Maqsood, T. (2014), Guidelines for empirical vulnerability assessment, GEM Technical Report 2014-08 V1.0.0, 140 pp., GEM Foundation, Pavia, Italy, doi: 10.13117/GEM.VULN-MOD.TR2014.11. Available at: https://core.ac.uk/download/pdf/21620143.pdf (last visited 15 November 2021

Rousseau, J. (2005). Lettre à Monsieur de Voltaire sur ses deux poèmes sur « la Loi naturelle » et sur « le Désastre de Lisbonne ». Écologie & politique, Volume 30, pp. 141-154. DOI: 10.3917/ecopo.030.0141

Royal Institution of Chartered Surveyors (RICS)(2020) International Valuation Standards (IVS). Available at: https://www.rics.org/globalassets/rics-website/media/upholding-professional-standards/sector-standards/valuation/international-valuation-standards-rics2.pdf (last visited November 14, 2022)

Safiey, A., & Pang, W. (2021). A new approach to assessing reparability for seismic risk assessment of buildings. Earthquake Spectra, Volume 37(1), pp. 284–303. DOI: 10.1177/8755293020944176

Santos, M. M. (2008). Estudo dos danos ocorridos em Lisboa causados pelo terramoto de 1755—Quantificação e Discussão (Dissertation Submitted for the Master Degree in Civil Engineering. IST/UTL).Available at: https://fenix.tecnico.ulisboa.pt/downloadFile/395137857471/Tese_C%C3%A9u_IMPRIMIR_ENTREGA_FINAL.pdf (Last visited November 14, 2022)

Santos, A., Correia, M., Loureiro, C., Fernandes, P., & Costa, N. (2019). The historical reconstruction of the 1755 earthquake and tsunami in downtown Lisbon, Portugal. Journal of Marine Science and Engineering. Volume 7(7). pp. 208. DOI: 10.3390/jmse7070208.

Savage, I.R. (1990). Leonard J. Savage. In: Eatwell, J., Milgate, M., Newman, P. (eds) Utility and Probability. The New Palgrave. Palgrave Macmillan, London. DOI: 10.1007/978-1-349-20568-4_31

Salazar, M. K. (1990) Interviewer Bias - How It Affects Survey Research. AAOHN JOURNAL, DECEMBER 1990, VOL. 38, NO. 12. pp: 567-572. Available at: https://journals.sagepub.com/doi/pdf/10.1177/216507999003801203 (last visited August 25th, 2022)

Sarafidis, V., & Wansbeek, T. (2010) Cross-sectional Dependence in Panel Data Analysis. University of Sidney. Available at: https://mpra.ub.uni-muenchen.de/20815/ (last visited February 8, 2022)

Sarmiento, J.P., Hoberman, G.,Ilcheva, M., Asgary, A., Majano, A.M., Paggione, S., & Duran, L. R., (2015) Private Sector and Disaster Risk Reduction: The Cases of Bogota, Miami, Kingston, San Jose, Santiago, and Vancouver. DRR Faculty Publications. 13.https://digitalcommons.fiu.edu/drr_fac/13

Scope Ratings GmbH. (2021). European real estate credit outlook 2021: overall resilience; pandemic's uneven impact hurts retail. Available at: https://www.scoperatings.com/#!search/research/detail/166399EN (last visited October 18, 2021)

Seixas, J. & Antunes, G. (2019). Tendências recentes de segregação habitacional na Área Metropolitana de Lisboa. Cidades, Comunidades e Território. Volume 39, pp. 55-82. DOI: 10.15847/citiescommunitiesterritories.dec2019.039.art01.

Serapioni, M., & Hespanha, P. (2019) Crisis and Austerity in Southern Europe: Impact on Economies and Societies. e-cadernos CES [Online], 31 | 2019, Online since 15 June 2019, connection on 21 September 2021. URL: http://journals.openedition.org/eces/4068 ; DOI: https://doi.org/10.4000/eces. 4068

Simões, A., Lopes, M., Bento, R., & Gago, A. (2012) Characterization and Classification of Lisbon Old Masonry Buildings. Available at: https://www.iitk.ac.in/nicee/wcee/article/WCEE2012_2434.pdf (Last visited November 14, 2022)

Simões, A., Bento, R., Cattari, S. & Lagomarsino, S. (2014) Seismic performance-based assessment of "Gaioleiro" buildings. Engineering Structures, Volume 80, Pages 486-500, ISSN 0141-0296. DOI: 10.1016/j.engstruct.2014.09.025

Simões, A. & Milosevic, J., Meireles, H., Bento, R., Cattari, S., & Lagomarsino, S. (2015). Fragility curves for old masonry building types in Lisbon. Bulletin of Earthquake Engineering. Volume 13, pp. 3083–3105. DOI: 10.1007/s10518-015-9750-1

Simões, A. (2018) Evaluation of the seismic vulnerability of the unreinforced masonry buildings constructed in the transition between the 19th and 20th centuries in Lisbon, Portugal. Instituto Superior Técnico.

Sotto-Mayor, M. (2006) Risco Sísmico em Portugal Continental. Instituto Superior Técnico. Available at: http://www-ext.lnec.pt/projects/ASAz/sousa06.pdf (last visited November 15, 2022)

Sousa, M.L., Martins, A. & Oliveira, C.S. (1992). Compilação de Catálogos Sísmicos da Região Ibérica. Relatório 36/92 – NDA, LNEC, Lisboa,.

Sousa, M.L, Costa, A.C., Carvalho, A., & Coelho, E., (2004). An automatic seismic scneraio loss methodology integrated on a geographic information system. 13th World Conference on Earthquake Engineering Vancouver, B.C., Canada. August 1-6, 2004. Paper No. 2526. Available at: http://www-ext.lnec.pt/projects/SPARC/Sousa_et_al_2004.pdf

Sousa, M. L. (2006). Risco sísmico em Portugal continental. Colecção teses e programas de investigação LNEC, Laboratório Nacional de Engenharia Civil (LNEC), Lisboa.

Sousa, ML, Ferreira, M.A., Sá, F. Mota de, Oliveira, C.S., Raposo, S., Sigbjornson, R., Rupakhety, R., Zonno, G., Meroni, F., Bianco, F., & D'Amico, S. (2012) Quantitative Sesimic Evaluation and Mapping: cases of schools and residential facilities in Lisbon and Algarve. 15th WCEE, Lisbon. Conference Paper

Sousa Oliveira, C. (2008) Efeitos Naturais, impacte e mitigação. Chapter of the book "Sismos e Edifícios" (in Portuguese), "Earthquakes and Buildings" (in English); Lopes, M. (Coordinator); Publisher: Orion

Slovic, P., Kunreuther, H., & White, G. (2016). Decision processes, rationality and adjustment to natural hazards. In The perception of risk (pp. 39-69). Routledge.

Snowball, J.d. & Willis, K.g. (2006). Building Cultural Capital: Transforming The South African National Arts Festival, South African Journal of Economics, Economic Society of South Africa, Volume 74(1), pp. 20-33, March. DOI: 10.1111/j.1813-6982.2006.00046.x

Spence, R., So, E., Ameri, G., Akinci, A., Cocco, M., Cultrera, G., Franceschina, G., Pacor, F., Pessina, V., Lombardi, A., Zonno, G., Carvalho, A., Costa, A., Coelho, E., Pitilakis, K., Anastasiadis, A., Kakderi, K., Alexoudi, M., Ansal, A., Erdic, M., Tönük, G., & Demircioglu, M. (2007). Earthquake Disaster Scenario Prediction and Loss Modelling for Urban Areas. Available at: https://www.earth-prints.org/bitstream/2122/3421/1/Report7.pdf (last visited May 26, 2023)

SPES, GECoRP (2001) - Redução da vulnerabilidade sísmica do edificado, Lisboa (também disponível em http://www.spes-sismica.org).

Sunstein, C. R. (2002). Probability Neglect: Emotions, Worst Cases, and Law. The Yale Law Journal, Volume 112(1), pp. 61–107. DOI: 10.2307/1562234

Taboga, M. (2021). "Generalized least squares", Lectures on probability theory and mathematical statistics. Kindle Direct Publishing. Online appendix. https://www.statlect.com/fundamentals-of-statistics/generalized-least-squares.

Tavares, I., Cândido, A.F., & Carmo, R.M. (2021), Desemprego e Precariedade Laboral na População Jovem: Tendências Recentes em Portugal e na Europa, Lisboa, Observatório das Desigualdades, CIES-Iscte Available at https://www.observatorio-das-desigualdades.com/2021/07/15/desempregoe-precariedade-laboral-na-populacao-jovem-tendencias-recentes-em-portugal-e-na-europade-ines-tavares-ana-filipa-candido-e-renato-miguel-do-carmo/ (last visited September 20th, 2021)

Tekeli-Yeşil S, Dedeoğlu N, Braun-Fahrlaender C, & Tanner, M. (2010). Factors motivating individuals to take precautionary action for an expected earthquake in Istanbul. Risk Analysis. 2010 Aug. Volume 30(8), pp.1181-95. DOI: 10.1111/j.1539-6924.2010.01424.x

Tekeli-Yeşil, S., Dedeoğlu, N., Braun-Fahrlaender, C., & Tanner, M. (2011). Earthquake awareness and perception of risk among the residents of Istanbul. Natural Hazards, Volume 59, pp. 427–446. DOI: 10.1007/s11069-011-9764-1

Teves-Costa, P., Barreira, E., & Omira, R. (2011) Estimativa de cenários de danos para a cidade de Lisboa. RISK-EU. Lisboa. Available at: http://lxrisk.cm-lisboa.pt/risco_sismico.html . (last visited 15 September 2021)

Tinker, A. (2002) The social implications of an ageing population, Mechanisms of Ageing and Development, Volume 123(7), pp 729-735. ISSN 0047-6374, DOI: 10.1016/S0047-6374(01)00418-3.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., & Staake, T. (2018) Overcoming Salience Bias: How Real-Time Feedback Fosters Resource Conservation. Management Science, Volume 64(3), pp. 1458-1476. DOI: 10.1287/mnsc.2016.2646

Tissot, C. (2016) Is Lisbon the next Silicon Valley?. Available at: https://id-norway.com/is-lisbon-the-next-silicon-valley/ (last visited October 14th, 2021)

Tobler W. (1970) A computer movie simulating urban growth in the Detroit region. Economic Geography, Volume 46(Supplement), pp. 234–240. DOI: 10.2307/143141

Toya, H., & Skidmore, M. (2007). Economic development and the impacts of natural disasters. Economics letters, Volume 94(1), pp. 20-25. DOI: 10.1016/j.econlet.2006.06.020

Transparency International (2018) European Getaway: Inside the Murky World of Golden Visas. Available at: https://www.transparency.org/en/publications/golden-visas (Last visited November 14, 2022)

Tse, C.B., Rodgers, T., & Niklewski, J. (2014) The 2007 financial crisis and the UK residential housing market: Did the relationship between interest rates and house prices change?, Economic Modelling, Volume 37, pp. 518-530, ISSN 0264-9993, DOI: 10.1016/j.econmod.2013.08.013.

Union Internationale de la Propriété Immobiliére (UIPI) (2013). A Comparative Study on Real Estate Property Taxation in Europe. Available at: https://www.uipi.com/wp-content/uploads/2018/09/Comparative-Analysis-on-Real-Estate-Taxation-2013.pdf (last visited October 18, 2021)

Van Der Pligt, J. (1998). Perceived risk and vulnerability as predictors of precautionary behaviour. British journal of health psychology, Volume 3(1), pp. 1-14. DOI: 10.1111/j.2044-8287.1998.tb00551.x

Venkatachalam, L. (2004) The contingent valuation method: a review. Environmental Impact Assessment Review, Volume 24(1), pp. 89-124, ISSN 0195-9255, DOI: 10.1016/S0195-9255(03)00138-0

Viegas, J. M., Martinez, L. M., & Silva, E. A. (2009). Effects of the Modifiable Areal Unit Problem on the Delineation of Traffic Analysis Zones. Environment and Planning B: Planning and Design, Volume 36(4), pp. 625–643. DOI: 10.1068/b34033

Voltaire, M. (1756) Poème sur le désastre de Lisbonne. Available at: https://gallica.bnf.fr/essentiels/anthologie/poeme-desastre-lisbonne (last visited May 26, 2023)

von Neumann, J., & Morgenstern, O., (1944). Theory of Games and Economic Behavior. Princeton: Princeton University Press. Section 3, chapter I. reprinted in Alfred N. Page. (1968). Utility Theory: A Book of Readings. New York: Wiley, pp. 215–233. ISBN: 9780691130613

Xerez, R., Pereira, E., & Cardoso, F. D. (2019). Habitação própria em Portugal numa perspetiva intergeracional. Relatório. Fundação Calouste Gulbenkian. Available at: https://content.gulbenkian.pt/wp-content/uploads/sites/46/2020/05/06150255/Habita %C3 %A7 %C3 %A3o-Pr %C3 %B3pria-Relat %C3 %B3rio-Final.pdf (last visited November 14, 2022)

Willis, K. G., & Asgary, A. (1997). The Impact of Earthquake Risk on Housing Markets: Evidence from Tehran Real Estate Agents. Journal of Housing Research, Volume 8(1), pp. 125–136. http://www.jstor.org/stable/24833635

Willis, K. G., & Asgary, A. (1997). Estimating the Benefits of Construction Measures to Mitigate Earthquake Risks in Iran. Environment and Planning B: Planning and Design, Volume 24(4), pp. 613–624. DOI: 10.1068/b240613

Weinstein, N. D., Lyon, J. E., Rothman, A. J., & Cuite, C. L. (2000). Changes in perceived vulnerability following natural disaster. Journal of Social and Clinical Psychology, Volume 19(3), pp. 372. DOI: 10.1521/jscp.2000.19.3.372

Wolff, K., Larsen, S., & Øgaard, T. (2019). How to define and measure risk perceptions. Annals of Tourism Research, Volume 79, Article 102759. DOI: 10.1016/j.annals.2019.102759

Wood, J. M., Bootzin, R. R., Rosenhan, D., Nolen-Hoeksema, S., & Jourden, F. (1992). Effects of the 1989 San Francisco earthquake on frequency and content of nightmares. Journal of Abnormal Psychology, Volume 101(2), pp. 219. DOI: 10.1037/0021-843X.101.2.219

Wyatt, P. (1996). The development of a property information system for valuation using a geographical information system (GIS). Journal of Property Research, Volume 13, pp.317-336. DOI: 10.1080/095999196368826

Yeats, R. (2015). Age of Enlightenment and the 1755 Lisbon earthquake. In Earthquake Time Bombs (pp. 158-163). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781316048184.015

Zhang, Y., Hwang, S. N., & Lindell, M. K. (2010). Hazard Proximity or Risk Perception? Evaluating Effects of Natural and Technological Hazards on Housing Values. Environment and Behavior, Volume 42(5), pp. 597–624. DOI: 10.1177/0013916509334564

ANNEX A

Table A.0-1 - Impacts of natural disasters on real estate (literature review summary table)

Study	Subject	Location	Key Findings
(Harrison et al., 2001)	Floods	Alachua County, Florida, USA	(1) Homes located within a flood zone sell, on average, for less than homes located outside flood zone; Pattern of price differential reinforced after the National Flood Insurance Reform Act of 1994;
(Bin and Polasky, 2004)	Floods	Pitt County, North Carolina, USA	(1) House located within a floodplain has a lower market value than an equivalent house located outside the floodplain (2) Price differential increased after Hurricane Floyd
(Daniel et al., 2009)	Floods	Netherlands	(1) Decrease of 9% of property values in the floodplain after the 1993 and 1995 Meuse river floods; However, there is an upward valuation of 3% on properties closer to water;
(Brookshire et al., 1985)	Earthquakes	California, USA	(1) People do not behave as if they are maximizing expected utility for low-probability, high-loss events such as natural disasters. (2) The Alquist-Priolo Act has caused a structural change in the market between 1972 and 1978 through the disclosure of risks.
(Beron et al., 1997)	Earthquakes	San Francisco Bay, California, USA	(1) The results indicate that individuals initially overestimate the probability of damage from unlikely events (2) The results indicate that the hedonic price of earthquake hazard fell after the event. The authors place the hypothesis that it happended because perhaps the damage pattern from earthquake was unrelated to Special Studies Zones (SSZ) designation. The information gained from the earthquake was relatively more "productive" in changing subjective risk than in changing the stigma of an SSZ.
(Tekeli-Yeşil et al., 2011)	Earthquakes	Istanbul, Turkey	(1) preparedness programmes should focus more on information about mitigation and preparedness measures regarding earthquakes; (2) the media can be the main channel for reaching the population in general, but if people are to understand, believe and personalize the information, various methods of risk communication should be used—risk communication is a two-way process, in which the sender and the audience are active participants; (3) the information provided by the media might be more useful if it helped people to question and understand the risk and their ability to cope with it, as well encourage them to take precautions; (4) the Internet also seems to be a promising option for disseminating information. programmes should target people with lower educational and socio-economic levels.
(Asgary and Halim, 2011)	Storms and Hurricanes	Bangladesh	(1) Access to training and education and cyclone warning systems are also found to have significant impacts on households' choices of cyclone vulnerability reduction (2) Structural mitigation measures and access to power and decision making, though significant, were found to have the least impact.

Study	Subject	Location	Key Findings
(Deng et al., 2013)	Earthquakes	Wenchuan, China	(1) Property prices decreased after the 2008 Wenchuan earthquake in the affected areas. (2) elative price of low to high floor units, particularly units located in the first and second floor, considerably increased for several months after the earthquake due to the perception of risk of living in upper floors. This shows an overreaction of buyers and investors after the tremor.
(Modica and Locati, 2016)	Earthquakes	Northern Italy, Italy	(1) There was a difference of 4.6% in the price reduction of houses after the earthquake. The most affected region by the earthquake was also the most affected by the reduction of prices (2) average price of villas in the treated area do not show significant differences with respect to that of the control group (3) houses with better structural quality were less affected by the price drop. the quality of the houses in term of perception of resistance to the tremor plays an important role
(Fekrazad, 2019)	Earthquakes	California, USA	(1) both home value index and median listing price decrease by approximately 6% and 3% respectively in California zip codes with high seismic hazard (relative to those with low seismic hazard) after high-casualty earthquakes occur outside of California.(2) the higher an earthquake's death toll, the larger the increase in the price differential; (3) an earthquake in Europe has a significantly larger impact on California's price differential than a similarly deadly earthquake in another region; (4) the effects are transient and dissipate one month after the earthquake occurs.
(Donovan et al., 2007)	Wildfires	Colorado Springs, Colorado, USA	(1) Negative effect of risk of wildfire in house prices immediately after the event, which diminishes over time (consistent with the literature on other natural disasters). (2) the authors believe that a considerable portion of the observable changes in household risk perception can be attributed to an educational program on wildfires.
(Kiel and Matheson, 2018)	Wildfires	Boulder, Colorado	(2) buyers in the highest risk area are most likely to change their perceptions in response to a fire with houses experiencing a statistically significant 21.9% decline in sale price
(Hallstrom and Smith, 2005)	Storms and Hurricanes	Lee County, Florida, USA	(1) Experiencing a "near miss" Hurricane led to a 19% decline in housing prices (2) Natural hazards are noticed by homeowners. People learn from comparable circumstances.
(Ewing et al., 2007)	Storms and Hurricanes	USA	(1) Results suggest that windstorms (hurricanes and/or tornados) result in an immediate one-half to two percent reduction in total Metropolitan Statistical Area housing value.
(Burrus et al., 2009)	Storms and Hurricanes	North Carolina, USA	(1) The researchers concluded that the market suffers after successive hurricane landfalls, but that sentiment recovers a year or more after the hurricane
(Zhang et al., 2010)	Storms and Hurricanes	Texas, USA	(1) Risk perception is a mediating factor between hazard proximity and property value, but there is some evidence that the mediation is partial rather than complete. Hazard proximity can be perceived as a potential risk and an environmental amenity at the same time for certain types of hazards. These two perceptions operate in opposite directions when affecting housing value.
(Lee et al., 2017)	Storms and Hurricanes	South Korea	(1) The result shows that a 1% increase in annual precipitation and impervious surfaces increases damage costs from hurricanes by 4.52 and 1.74%, respectively

Study	Subject	Location	Key Findings
(Palm, 1998)	Earthquakes	USA and Japan	(1) two populations were indistinguishable in their general beliefs that individuals can not do much to prevent an earthquake from harming them, but that cities and communities can take actions to lessen the effects of earthquakes. (2) In Yokohama, family income, age of the homeowner and perceived vulnerability of the home were the best predictors to insurance adoptance. In the western San Fernando valley, income and perceived vulnerability were also important, but age was not related. Instead, belief about the ability of the household or community to lessen damage.
(Asgary and Willis, 1997)	Earthquakes	Tehran and Rasht, Iran	(1) Response to risk information does not automatically. Response to earthquake risk is not significantly influenced by attitudes toward earthquakes as a threat to life. This means that households who consider earthquakes as a threat to life and are worried about the risk, do not necessarily exhibit mitigatory behaviour as response to receiving information
(Peng, 2021)	Earthquakes	Japan	(2) housing/land attributes are still the key features to real estate values in Hokkaido. In comparison with short-term unexpected hazard shock (i.e. 2018 Hokkaido Eastern Iburi Earthquake), the long-term hazard threats are more influential to prices of properties and lands
(Pryce et al., 2011)	Floods	UK	(1) home price adjustment to ever-increasing levels of risk associated with global climate change is likely to evince an uneven pattern of inertia followed by rapid, tipping-point declines (2) individuals' risk perception of hazards is determined by the uncertainty of hazard occurrence, and the limitations of human cognition

Table A.0-2 - Impacts from Earthquakes (literature review summary table)

Study	Subject	Location	Key Findings
(Naoi et al., 2009)	Earthquakes	Japan	(1) While there is no clear evidence about pre-quake responses to earthquake risk, both housing rents and owner-occupied home values are significantly and negatively correlated with regional earthquake risk in post-quake periods. Themost plausible interpretation for these results is that both renters and homeowners are initially unaware of, or at least underestimate, earthquake risk. (2) Households can hedge against earthquake risk through the purchase of insurance policies (....) However, earthquake risks are partly reflected in the differentiated premiums in the insurance market.
(Brookshire et al., 1985)	Earthquakes	California, USA	(1) People do not behave as if they are maximizing expected utility for low-probability, high-loss events such as natural disasters. (2) The Alquist-Priolo Act has caused a structural change in the market between 1972 and 1978 through the disclosure of risks.

Study	Subject	Location	Key Findings
(MacDonald et al., 1987)	Floods	Monroe, Lousiana, USA	(1)Without insurance, the consumer's willingness to pay is captured in the sales price differential between houses located in areas with varying probabilities of flooding. With insurance, the sales price differential captures the change in the insurance cost and the willingness to pay to move to an area with a lower probability of flooding
(Willis and Asgary, 1997)	Earthquakes	Tehran, Iran	(1) People show little concern for earthquake risk; (2) Legal measures and recent earthquakes only have short term impact on the housing market;
(Önder et al., 2004)	Earthquakes	Instambul, Turkey	(1) Distance from fault lines is an important factor in explaining house values and its impact on house values increased after the 1999 Kocaeli earthquake. (2) no evidence was found of the linear impact of soil type in property prices, however it is possible there is a non-linear relationship between soil and property type; (3) None of the earthquake risk measures significantly affects the change in house values before and after the Kocaeli earthquake. So, the impact of geologists' warnings about soil types was limited to only a few high-income locations, thus not thus influencing the regression resul
(Bin and Polasky, 2004)	Floods	Pitt County, North Carolina, USA	(1) House located within a floodplain has a lower market value than an equivalent house located outside the floodplain (2) Price differential increased after Hurricane Floyd
(Hallstrom and Smith, 2005)	Storms and Hurricanes	Lee County, Florida, USA	(1) Experiencing a "near miss" Hurricane led to a 19% decline in housing prices (2) Natural hazards are noticed by homeowners. People learn from comparable circumstances.
(Nakagawa et al., 2007)	Earthquakes	Tokyo, Japan	(1) housing rents are substantially lower in risky areas than in safer areas, even after controlling for other possible effects, and that the rent of an apartment built prior to the Building Standard Law (1981) being amended is discounted more substantially in risky areas than those built after this date (2) Earthquake risk is reflected in housing rents to a large extent even after controlling for other possible effects; that is, housing rents are substantially lower in the areas with exposure to earthquake risk
(Fekrazad, 2019)	Earthquakes	California, USA	(1) both home value index and median listing price decrease by approximately 6% and 3% respectively in California zip codes with high seismic hazard (relative to those with low seismic hazard) after high-casualty earthquakes occur outside of California.(2) the higher an earthquake's death toll, the larger the increase in the price differential; (3) an earthquake in Europe has a significantly larger impact on California's price differential than a similarly deadly earthquake in another region; (4) the effects are transient and dissipate one month after the earthquake occurs.
(Nakagawa et al., 2009)	Earthquakes	Tokyo, Japan	(1) The study finds strong evidence for the impact of earthquake risks on land pricing; land prices have been substantially lower in risky areas than in safe areas.

Study	Subject	Location	Key Findings
(Hidano et al., 2015)	Earthquakes	Tokyo, Japan	(1) The prices of apartments in safe zones are significantly higher than those in unsafe zones on average. (2) The impact of risk information is not significant for newly built apartments.
(Deng et al., 2013)	Earthquakes	Wenchuan, China	(1) Property prices decreased after the 2008 Wenchuan earthquake in the affected areas. (2) elative price of low to high floor units, particularly units located in the first and second floor, considerably increased for several months after the earthquake due to the perception of risk of living in upper floors. This shows an overreaction of buyers and investors after the tremor.
(Brooks et al., 2003)	Financial Markets	US	(1) Selling pressure and higher volume (for over an hour) when firms experience events that are totally unanticipated by the equity market
(Ederington and Lee, 1993)	Financial Markets	US	(1) The study finds huge volatility in prices after huge macroeconomic news annoucements. This volatility then fades over time.
(Ederington and Lee, 1995)	Financial Markets	US	(1) Overreaction from investors immediately after the release of huge macroeconomic news. Volatility is corrected in the following moments.
(Modica and Locati, 2016)	Earthquakes	Northern Italy, Italy	(1) There was a difference of 4.6% in the price reduction of houses after the earthquake. The most affected region by the earthquake was also the most affected by the reduction of prices (2) average price of villas in the treated area do not show significant differences with respect to that of the control group (3) houses with better structural quality were less affected by the price drop. the quality of the houses in term of perception of resistance to the tremor plays an important role
(Beron et al., 1997)	Earthquakes	San Francisco Bay, California, USA	(1) The results indicate that individuals initially overestimate the probability of damage from unlikely events (2) The results indicate that the hedonic price of earthquake hazard fell after the event. The authors place the hypothesis that it happended because perhaps the damage pattern from earthquake was unrelated to Special Studies Zones (SSZ) designation. The information gained from the earthquake was relatively more "productive" in changing subjective risk than in changing the stigma of an SSZ.
(McGinnis, 2004)	Earthquakes	San Francisco Bay, California, USA	(1) Conclusions moderately support a theory in which the psychological effect of large earthquakes increases the effect of earthquake risk on prices but this impact tends to wear off over time
(Lamond et al., 2010)	Floods	UK	(1) the results suggest that the reaction of property markets to risk can be highly subjective in the absence of enforced disclosure. (2) Markets may become distorted by any event that changes the perceived importance of previously ignored risks.

Table A.0-3 - Building Stock's Risk Assessment - Specific building typologies (summary table)

Study	Building Typology	Key Findings
(Simões et al., 2015)	"Pombalino", "Gaioleiro" and "Placa"	(1) All buildings have very high seismic vulnerability and do not fulfil the requirements for the ultimate limit state as defined in EC8-1 and Italian Code (NTC 2008)
(Mendes and Lourenço, 2010)	"Gaioleiro"	(1) Even buildings with appropriate floor-wall connections are in the limit of their loading capacity under seismic action, according to EC8 National Annex
(Neves, 2016)	"Gaioleiro"	(1) The flexible wooden floors are not able to provide the stiffness needed to evenly distribute horizontal loads across the vertical masonry walls. (2) The difference in the stiffness between the x and y directions, further contributing to its vulnerability.
(Simões, 2018)	Unreinforced masonry buildings	(1) type I buildings to have approx. 50% probability of having very heavy damage and more than 30% of probability of collapse
(Marques et al., 2017)	"Placa"	(1) "Placa" buildings were also found to be unsafe, mainly due to its irregularity in elevation and heterogeneity in distribution of structural elements
(Caruso and Bento, 2019)	Old reinforced concrete building (pre-1980)	(1) The study found old reinforced concrete building (pre-1980) to only withstand low seismic forces and to have insufficient seismic detailing
(Jarimba, 2016)	Transition concrete buildings (1960-1980)	(1) the structural behavior of this typology is conditioned by the type of actions applied to the building however it shows an inadequate resistance to seismic action, especially in upper floors

Table A.0-4 – Building Stock's Risk Assessment – Overall building stock (summary table)

Study	Key Findings
(Pais 2002)	(1) Lisbon's downtown area is the most vulnerable to severe damage
LNECLoss (Costa et al., 2004; Afonso, 2006; Costa et al., 2008)	(1) Economic losses derived from an seismic event vary from 1.3%, for 95 years return period, up to 38%, for 5,000 years return period, of the total replacement cost of Lisbon Metropolitan Area's residential building stock. (2) Structural retroffiting could lead to a reduction of uo to 36% of economic losses (3) the chosen mitigation action has predominant effects on the South margin of Tagus River, where intermediate soils prevail
(Teves-Costa et al., 2011)	(1) Lisbon's downtown area is the most vulnerable to severe damage

Figure A.0-1 - EMS98 damage levels. [source]

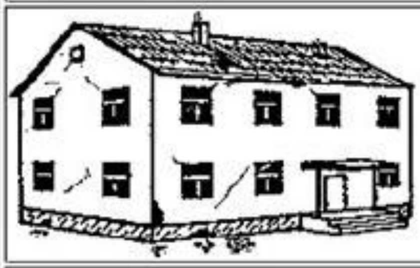
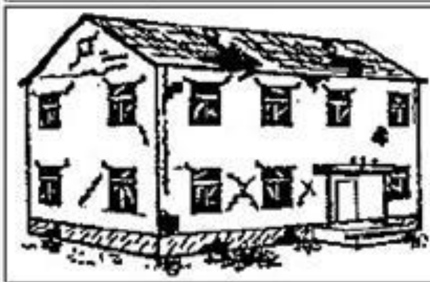
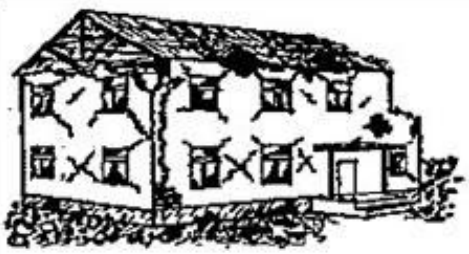
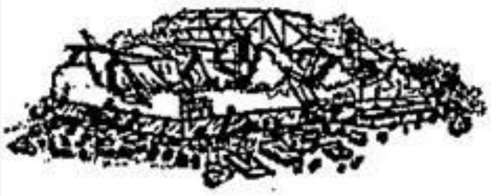
Classification of damage to masonry buildings	
	Grade 1: Negligible to slight damage (no structural damage, slight non-structural damage) Hair-line cracks in very few walls. Fall of small pieces of plaster only. Fall of loose stones from upper parts of buildings in very few cases.
	Grade 2: Moderate damage (slight structural damage, moderate non-structural damage) Cracks in many walls. Fall of fairly large pieces of plaster. Partial collapse of chimneys.
	Grade 3: Substantial to heavy damage (moderate structural damage, heavy non-structural damage) Large and extensive cracks in most walls. Roof tiles detach. Chimneys fracture at the roof line; failure of individual non-structural elements (partitions, gable walls).
	Grade 4: Very heavy damage (heavy structural damage, very heavy non-structural damage) Serious failure of walls; partial structural failure of roofs and floors.
	Grade 5: Destruction (very heavy structural damage) Total or near total collapse.

Figure A.0-2 - Vulnerability classes of buildings (EMS-98). Source: Modica and Locati (2016)

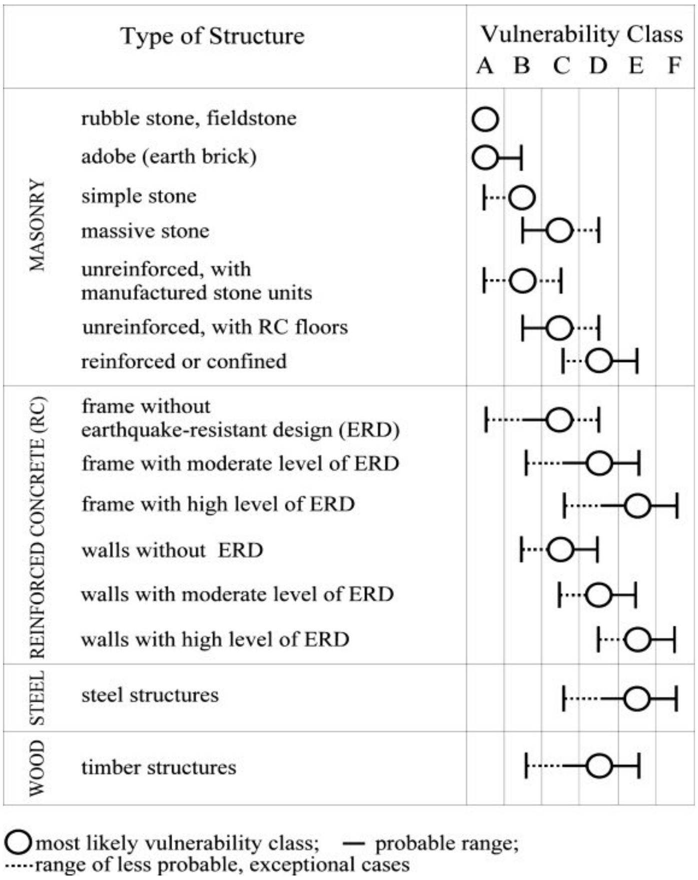


Table A.0-5 - Classification into classes and corresponding vulnerability

Tipologia	Descrição	Novo Grupo Tipológico	Tipologia (EMS-98)	Vu
T1	Edifícios construídos antes de 1919 em Paredes de adobe ou alvenaria de pedra solta	1	M3+	0,830
T2	Edifícios construídos de 1919 a 1945 em Paredes de adobe ou alvenaria de pedra solta			
T3	Edifícios construídos de 1946 a 1960 em Paredes de adobe ou alvenaria de pedra solta			
T4	Edifícios construídos de 1961 a 1970 em Paredes de adobe ou alvenaria de pedra solta	2	M3	0,767
T5	Edifícios construídos de 1971 a 1980 em Paredes de adobe ou alvenaria de pedra solta			
T6	Edifícios construídos de 1981 a 1990 em Paredes de adobe ou alvenaria de pedra solta			
T7	Edifícios construídos de 1991 a 1995 em Paredes de adobe ou alvenaria de pedra solta	3	M3-	0,660
T8	Edifícios construídos de 1996 a 2000 em Paredes de adobe ou alvenaria de pedra solta			
T9	Edifícios construídos de 2001 a 2005 em Paredes de adobe ou alvenaria de pedra solta	4	M3-	0,660
T10	Edifícios construídos de 2006 a 2011 em Paredes de adobe ou alvenaria de pedra solta			
T11	Edifícios construídos antes de 1919 em Paredes de alvenaria, sem placa	5	M5+	0,830
T12	Edifícios construídos de 1919 a 1945 em Paredes de alvenaria, sem placa			

T13	Edifícios construídos de 1946 a 1960 em Paredes de alvenaria, sem placa			
T14	Edifícios construídos de 1961 a 1970 em Paredes de alvenaria, sem placa			
T15	Edifícios construídos de 1971 a 1980 em Paredes de alvenaria, sem placa	6	M5	0,740
T16	Edifícios construídos de 1981 a 1990 em Paredes de alvenaria, sem placa			
T17	Edifícios construídos de 1991 a 1995 em Paredes de alvenaria, sem placa	7	M5-	0,650
T18	Edifícios construídos de 1996 a 2000 em Paredes de alvenaria, sem placa			
T19	Edifícios construídos de 2001 a 2005 em Paredes de alvenaria, sem placa	8	M5-	0,650
T20	Edifícios construídos de 2006 a 2011 em Paredes de alvenaria, sem placa			
T21	Edifícios construídos antes de 1919 em Paredes de alvenaria, com placa			
T22	Edifícios construídos de 1919 a 1945 em Paredes de alvenaria, com placa	9	M6+	0,793
T23	Edifícios construídos de 1946 a 1960 em Paredes de alvenaria, com placa			
T24	Edifícios construídos de 1961 a 1970 em Paredes de alvenaria, com placa			
T25	Edifícios construídos de 1971 a 1980 em Paredes de alvenaria, com placa	10	M6-	0,490
T26	Edifícios construídos de 1981 a 1990 em Paredes de alvenaria, com placa			
T27	Edifícios construídos de 1991 a 1995 em Paredes de alvenaria, com placa	11	M7	0,384
T28	Edifícios construídos de 1996 a 2000 em Paredes de alvenaria, com placa			
T29	Edifícios construídos de 2001 a 2005 em Paredes de alvenaria, com placa	12	M7	0,384
T30	Edifícios construídos de 2006 a 2011 em Paredes de alvenaria, com placa			
T31	Edifícios construídos antes de 1919 em Betão armado			
T32	Edifícios construídos de 1919 a 1945 em Betão armado	13	RC1	0,644
T33	Edifícios construídos de 1946 a 1960 em Betão armado			
T34	Edifícios construídos de 1961 a 1970 em Betão armado			
T35	Edifícios construídos de 1971 a 1980 em Betão armado	14	RC4	0,544
T36	Edifícios construídos de 1981 a 1990 em Betão armado			
T37	Edifícios construídos de 1991 a 1995 em Betão armado	15	RC5	0,384
T38	Edifícios construídos de 1996 a 2000 em Betão armado			
T39	Edifícios construídos de 2001 a 2005 em Betão armado	16	RC6	0,224
T40	Edifícios construídos de 2006 a 2011 em Betão armado			
T41	Edifícios construídos antes de 1919 em Outra(Madeira, metálica, ...)			
T42	Edifícios construídos de 1919 a 1945 em Outra(Madeira, metálica, ...)	17	W+	0,640
T43	Edifícios construídos de 1946 a 1960 em Outra(Madeira, metálica, ...)			
T44	Edifícios construídos de 1961 a 1970 em Outra(Madeira, metálica, ...)	18	W	0,447
T45	Edifícios construídos de 1971 a 1980 em Outra(Madeira, metálica, ...)			

T46	Edifícios construídos de 1981 a 1990 em Outra(Madeira, metálica, ...)			
T47	Edifícios construídos de 1991 a 1995 em Outra(Madeira, metálica, ...)	19	W	0,447
T48	Edifícios construídos de 1996 a 2000 em Outra(Madeira, metálica, ...)			
T49	Edifícios construídos de 2001 a 2005 em Outra(Madeira, metálica, ...)	20	W	0,447
T50	Edifícios construídos de 2006 a 2011 em Outra(Madeira, metálica, ...)			
T51	Edifícios construídos antes de 1919 sem classificação			
T52	Edifícios construídos de 1919 a 1945 sem classificação	21	M2	0,645
T53	Edifícios construídos de 1946 a 1960 sem classificação			
T54	Edifícios construídos de 1961 a 1970 sem classificação	22	M2	0,633
T55	Edifícios construídos de 1971 a 1980 sem classificação			
T56	Edifícios construídos de 1981 a 1990 sem classificação	23	M2	0,534
T57	Edifícios construídos de 1991 a 1995 sem classificação			
T58	Edifícios construídos de 1996 a 2000 sem classificação			
T59	Edifícios construídos de 2001 a 2005 sem classificação	24	M2	0,466
T60	Edifícios construídos de 2006 a 2011 sem classificação			

Table A.0-6 - Number of Portuguese households by annual gross income

Annual Gross Income	Nº Households	% Total
0-5.000€	677673	12.5%
5.001€-10.000€	1446100	26.7%
10.001€-13.500€	841952	15.6%
13.501€-19.000€	783469	14.5%
19.001€ -27.500€	689398	12.7%
27.501€ -32.500€	237623	4.4%
32.501€ - 40.000€	231445	4.3%
40.001€ - 50.000€	187013	3.5%
50.001€ - 100.000€	261496	4.8%
100.001€ -250.000€	47939	0.9%
250.000€+	4180	0.1%

Table A.0-7 - Income tax rate by category (2021). Source: AT

income tax category (by Annual Gross income)	Marginal Tax Rate (%)	Average Tax Rate(%)
0 - 7.112€	14.50%	14.50%
7112€ - 10.732€	23%	17.37%
10.732€ - 20.322€	28.50%	22.62%
20.322€ - 25.075€	35%	24.96%
25.075€ - 36.967€	37%	28.83%
36.967€ - 80.882€	45%	37.61%
80.822€+	48%	-

Table A.0-8 - Current vulnerability based on construction year. Source: Ferreira (2012)

Unknown Structure	Vuactual
< 1960	0,645
1960 - 1969	0,633
1970 - 1985	0,536
>1985	0,466

Table A.0-9 - Distribution of residents by age groups. Aging index is Pop(65+)/Pop(<14). Source: INE (census 2021)

	Total	0 - 14		15 - 24		25 - 64		65+		Aging Index
Location	N.º	N.º	%	N.º	%	N.º	%	N.º	%	
Lisboa	545923	71245	13%	55002	10%	291881	53%	127795	23%	1.79
Ajuda	14313	1796	13%	1267	9%	7326	51%	3924	27%	2.18
Alcântara	13852	1774	13%	1196	9%	7540	54%	3342	24%	1.88
Alvalade	33313	4639	14%	3335	10%	17100	51%	8239	25%	1.78
Areeiro	21167	2781	13%	2218	10%	11105	52%	5063	24%	1.82
Arroios	33307	3456	10%	3135	9%	20039	60%	6677	20%	1.93
Avenidas Novas	23261	3047	13%	2460	11%	12296	53%	5458	23%	1.79
Beato	12185	1351	11%	1144	9%	6547	54%	3143	26%	2.33
Belém	16549	2468	15%	1727	10%	7961	48%	4393	27%	1.78
Benfica	35367	4102	12%	3194	9%	17595	50%	10476	30%	2.55
Campo de Ourique	22146	3156	14%	2071	9%	11489	52%	5430	25%	1.72
Campolide	14794	1702	12%	1387	9%	8108	55%	3597	24%	2.11
Carnide	18029	2462	14%	2016	11%	9282	51%	4269	24%	1.73
Estrela	20308	3028	15%	2107	10%	10982	54%	4191	21%	1.38
Lumiar	46338	6790	15%	5182	11%	24677	53%	9689	21%	1.43
Marvila	35482	4699	13%	3721	10%	18521	52%	8541	24%	1.82
Misericórdia	9660	953	10%	845	9%	5627	58%	2235	23%	2.35
Olivais	32184	3867	12%	3260	10%	16504	51%	8553	27%	2.21
Parque das Nações	22382	3961	18%	2707	12%	12315	55%	3399	15%	0.86
Penha de França	28485	3077	11%	2579	9%	16243	57%	6586	23%	2.14
Santa Clara	23650	4464	19%	2884	12%	12427	53%	3875	16%	0.87
Santa Maria Maior	10052	875	9%	852	8%	6298	63%	2027	20%	2.32
Santo António	11062	1306	12%	1067	10%	6308	57%	2381	22%	1.82
São Domingos de Benfica	34081	4058	12%	3469	10%	17557	52%	8997	26%	2.22
São Vicente	13956	1433	10%	1179	8%	8034	58%	3310	24%	2.31

Table A.0-10 - Variables description of original dataset

variable	count	mean	std	min	25%	50% (median)	75%	max
FID	8726	4362.50	2519.12	0.00	2181.25	4362.50	6543.75	8725.00
CODE	8726	21691.25	21765.90	1.00	4679.75	13695.50	30675.75	79710.00
QUARTER	8726	2.55	1.12	1.00	2.00	3.00	4.00	4.00
SEMESTER	8726	1.51	0.50	1.00	1.00	2.00	2.00	2.00
YEAR	8726	2012.71	2.97	2008.00	2010.00	2013.00	2015.00	2018.00
PARISH_Cod	8726	14.32	6.42	1.00	9.00	15.00	20.00	23.00
ZONE_Code	8726	4.23	1.92	1.00	3.00	4.00	6.00	7.00

variable	count	mean	std	min	25%	50% (median)	75%	max
CONSTRUCTI	8726	1967.20	27.94	1550.00	1951.00	1960.00	1992.00	2018.00
BUILDING_F	8726	4.58	3.55	0.00	2.00	4.00	7.00	26.00
TYPOLOGY_C	8726	3.75	1.78	1.00	3.00	3.00	5.00	17.00
FLOOR	8726	2.50	2.30	-3.00	1.00	2.00	4.00	20.00
FLOOR_Extr	8726	0.01	0.13	0.00	0.00	0.00	0.00	3.00
AREA	8726	102.55	67.21	10.00	62.80	87.00	124.50	2355.00
BUILDING_2	8726	2.35	0.59	1.00	2.00	2.00	3.00	3.00
APARTMENT1	8726	2.42	0.67	1.00	2.00	3.00	3.00	3.00
ABANDONED_	8726	1.65	0.67	0.00	2.00	2.00	2.00	2.00
ONUS_Code	8726	3.89	1.70	0.00	3.00	5.00	5.00	5.00
BUYER_Code	8726	0.55	1.11	0.00	0.00	0.00	0.00	5.00
NORMALIZED	8726	1.00	0.58	0.00	0.62	0.89	1.24	6.00
NORMALIZ_1	8726	1.00	0.58	0.00	0.62	0.89	1.23	6.02
NORMALIZ_2	8726	1.00	0.58	0.00	0.62	0.89	1.24	6.14
PURCHACE_U	8726	0.94	0.55	0.00	0.56	0.82	1.17	5.39
PURCHACE_1	8726	0.94	0.56	0.00	0.56	0.82	1.17	5.40
PURCHACE_2	8726	0.94	0.56	0.00	0.56	0.82	1.17	5.40
x	8726	-8.08	2.95	-9.23	-9.17	-9.15	-9.13	0.00
y	8726	34.16	12.49	0.00	38.71	38.72	38.74	38.79
POI_Metro	8726	443.27	365.40	-1.00	194.62	373.17	607.49	2723.38
POI_Mercad	8726	208.52	193.99	-1.00	69.94	158.55	294.18	1256.31
POI_Saude	8726	135.31	115.27	-1.00	60.24	116.58	187.79	838.02
POI_Parque	8726	182.59	152.74	-1.00	63.87	158.79	269.15	1181.44
POI_Estudo	8726	640.01	652.37	-1.00	274.86	485.78	730.58	3623.18
POI_Emerg	8726	586.14	458.65	-1.00	252.46	482.61	872.79	2191.54
POI_CBD	8726	1682.61	1381.01	-1.00	562.26	1387.68	2530.87	6568.03
POI_Cemit	8726	1378.01	840.92	-1.00	743.81	1388.38	1993.21	3386.58
POI_AcRodo	8726	349.54	288.95	-1.00	138.67	297.24	496.16	1698.42
POI_Univer	8726	633.04	518.14	-1.00	275.76	537.83	850.27	3491.78
POI_Indust	8726	1372.25	1010.38	-1.00	611.34	1283.74	1941.66	4580.90
POI_Rio	8726	2071.29	1969.26	-1.00	549.46	1401.07	3020.91	8051.90
POI_Escola	8726	266.81	170.10	-1.00	144.27	264.62	379.82	873.58
POI_ESPriv	8726	552.58	414.92	-1.00	251.04	494.93	782.09	2205.28
ESPubl	8726	600.71	400.80	-1.00	339.87	581.83	828.72	2944.08
AUGI	8726	0.00	0.04	0.00	0.00	0.00	0.00	1.00
GEBALIS	8726	0.04	0.20	0.00	0.00	0.00	0.00	1.00
RUIDO	8726	1.50	1.15	0.00	1.00	1.00	2.00	5.00
VULInund	8726	0.53	0.95	0.00	0.00	0.00	1.00	3.00
VULSismo	8726	2.00	1.11	0.00	1.00	2.00	3.00	4.00
VULMass	8726	0.04	0.28	0.00	0.00	0.00	0.00	3.00
Park50M	8726	22.05	25.16	0.00	0.00	13.00	40.00	160.00
Culture500	8726	2.19	4.34	0.00	0.00	0.00	2.00	26.00

Table A.0-11 - BRD mean values for old parishes.

Parish Code	OLD_PARISH	BRD_PROXIMO
48	São Nicolau	0.839
18	Madalena	0.721
49	São Paulo	0.628
33	Santiago	0.597
35	Santo Estêvão	0.565
21	Mártires	0.522
47	São Miguel	0.509
26	Sacramento	0.509
11	Castelo	0.507
37	Socorro	0.502
52	Sé	0.487
42	São José	0.487
13	Coração de Jesus	0.478
45	São João de Deus	0.469
24	Penha de França	0.466
51	São Vicente de Fora	0.459
23	Pena	0.434
5	Anjos	0.431
36	Santos-o-Velho	0.420
38	São Cristóvão e São Lourenço	0.419
20	Mercês	0.410
41	São Jorge de Arroios	0.407
46	São Mamede	0.404
27	Santa Catarina	0.393
2	Alto do Pina	0.390
16	Lapa	0.375
30	Santa Justa	0.367
8	Campo Grande	0.358
3	Alvalade	0.352
15	Graça	0.349
28	Santa Engrácia	0.348
22	Nossa Senhora de Fátima	0.345
43	São João	0.338
6	Beato	0.333
44	São João de Brito	0.329
50	São Sebastião da Pedreira	0.324
29	Santa Isabel	0.307
34	Santo Condestável	0.301
25	Prazeres	0.294
14	Encarnação	0.281
39	São Domingos de Benfica	0.268
1	Alcântara	0.262

Parish Code	OLD_PARISH	BRD_PROXIMO
9	Campolide	0.239
31	Santa Maria de Belém	0.218
32	Santa Maria dos Olivais	0.216
19	Marvila	0.205
7	Benfica	0.202
10	Carnide	0.134
17	Lumiar	0.130
0	Ajuda	0.120
40	São Francisco Xavier	0.070
12	Charneca	0.067
4	Ameixoeira	0.031

Table A.0-12 - BRD mean values for new parishes

Parish Code	NEW_PARISH1	BRD_PROXIMO
20	Santa Maria Maior	0.546
15	Misericórdia	0.465
21	Santo António	0.451
3	Areeiro	0.446
4	Arroios	0.419
18	Penha de França	0.410
23	São Vicente	0.397
2	Alvalade	0.348
12	Estrela	0.347
5	Avenidas Novas	0.337
6	Beato	0.333
9	Campo de Ourique	0.304
22	São Domingos de Benfica	0.268
1	Alcântara	0.262
10	Campolide	0.238
16	Olivais	0.217
14	Marvila	0.205
8	Benfica	0.197
17	Parque das Nações	0.196
7	Belém	0.179
11	Carnide	0.144
13	Lumiar	0.129
0	Ajuda	0.120
19	Santa Clara	0.035

Figure A.0-3 - Correlation matrix of variables and components

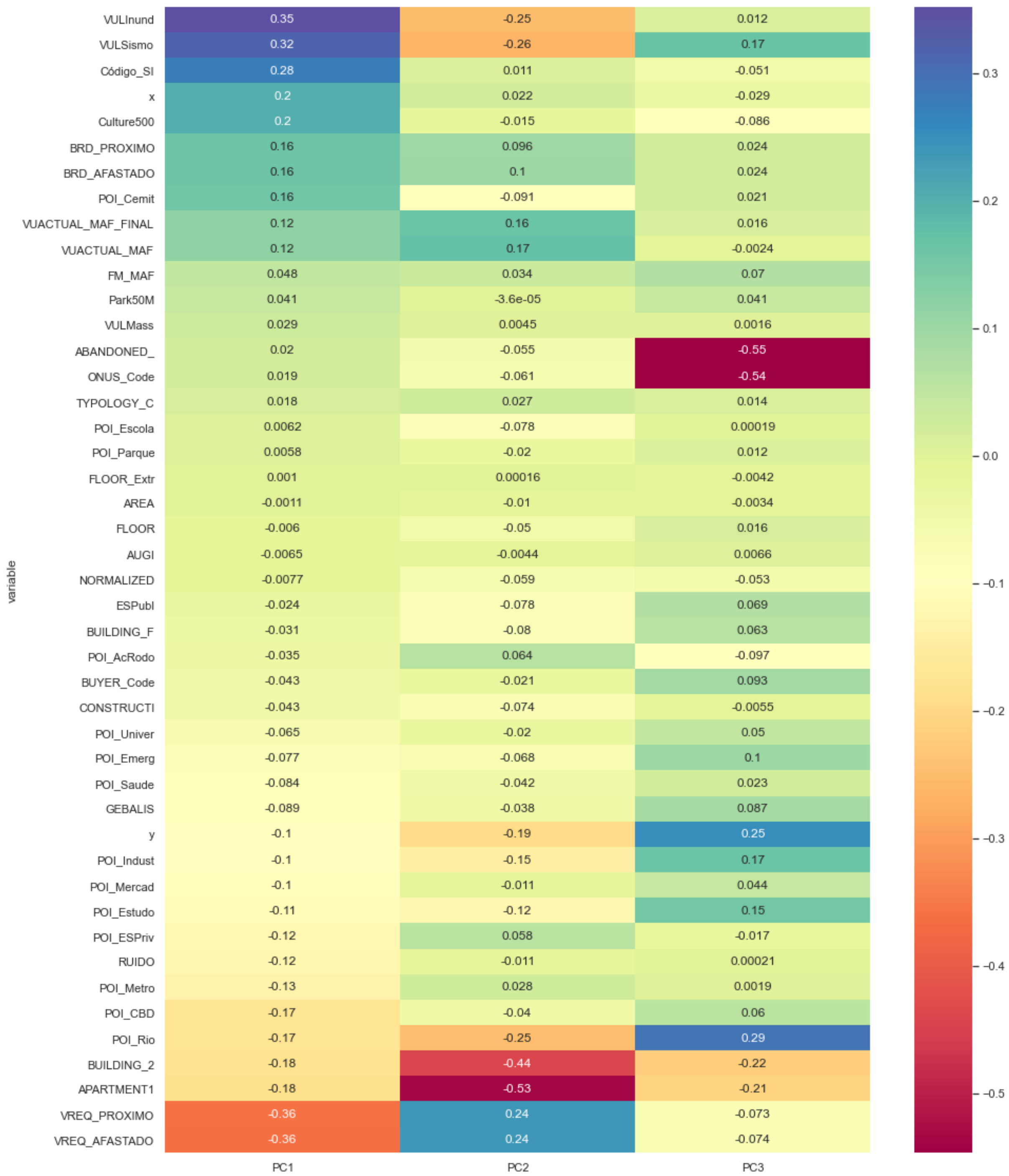


Figure A.0-4 - Most important features for Lumiar (RMSE: 0.388666)

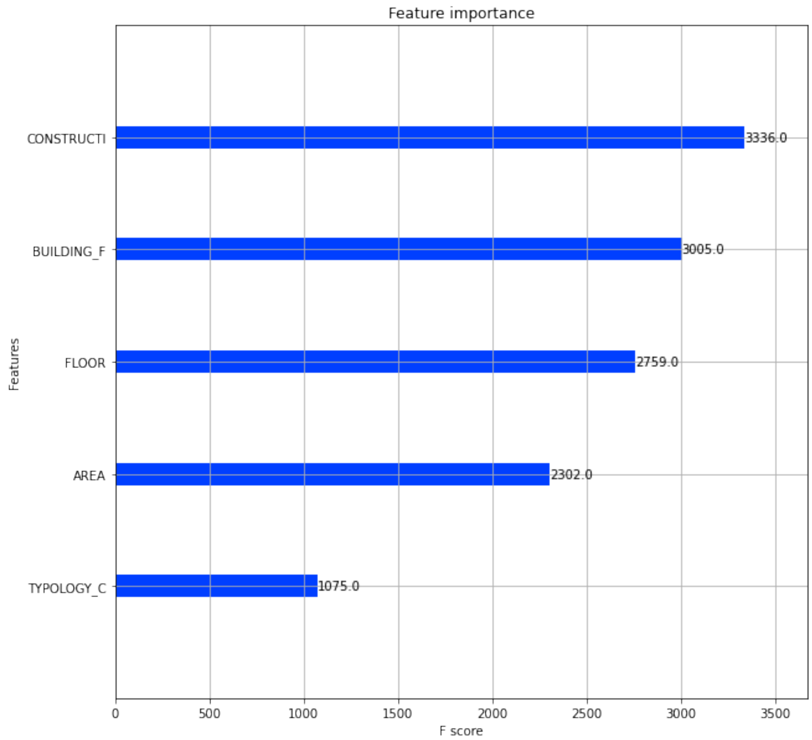


Figure A.0-5 - Most important features for Estrela (RMSE: 0.727539)

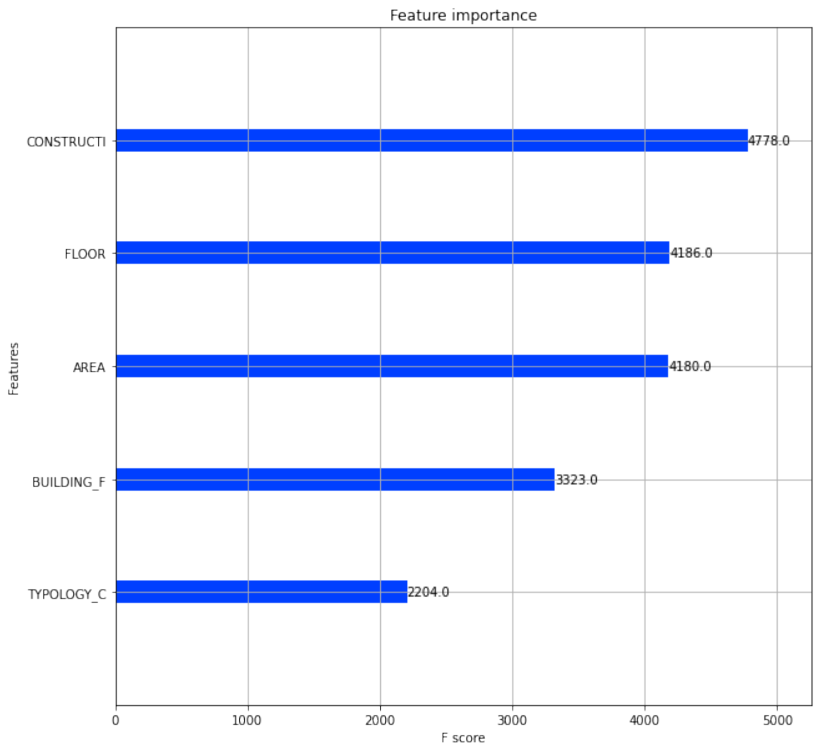


Figure A.0-6 - Most important features for Avenidas Novas (RMSE: 0.636906)

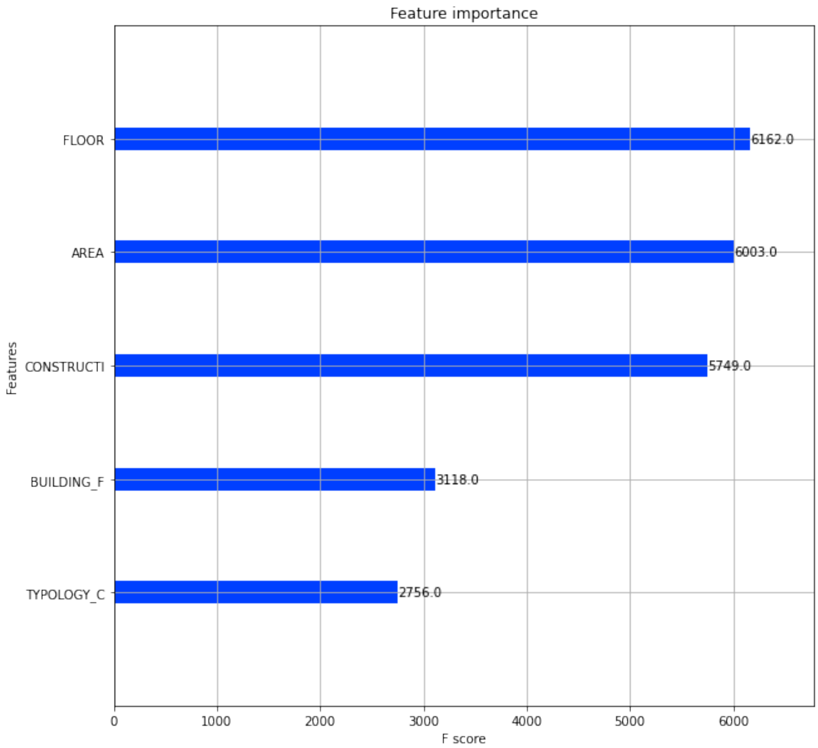


Figure A.0-7 - Most important features for Penha de França (RMSE: 0.271429)

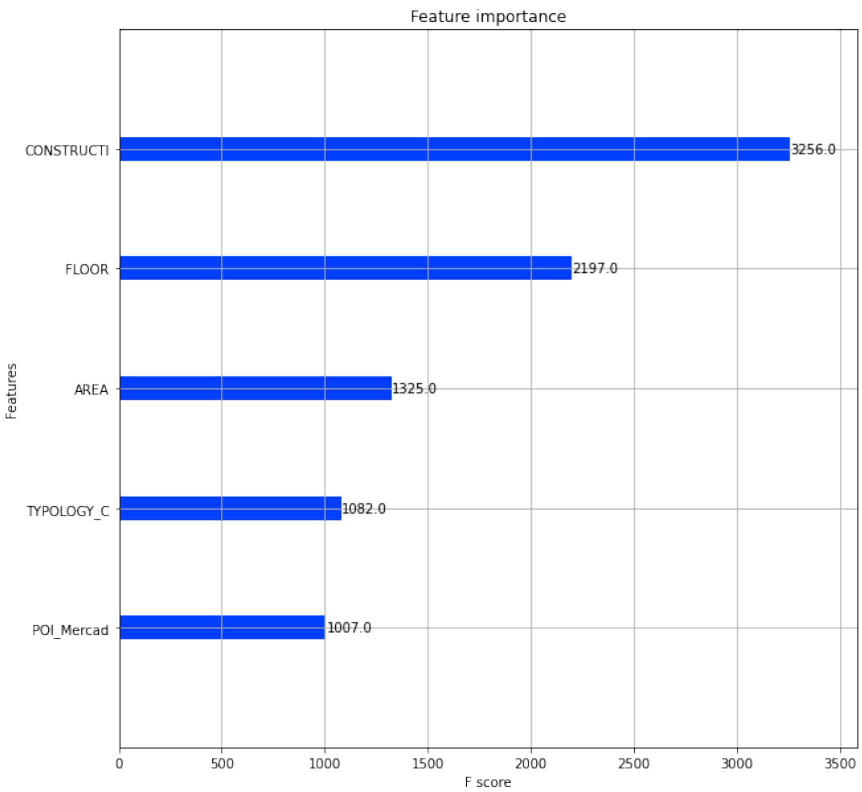


Figure A.0-8 - Most important features for Arroios (RMSE: 0.804470)

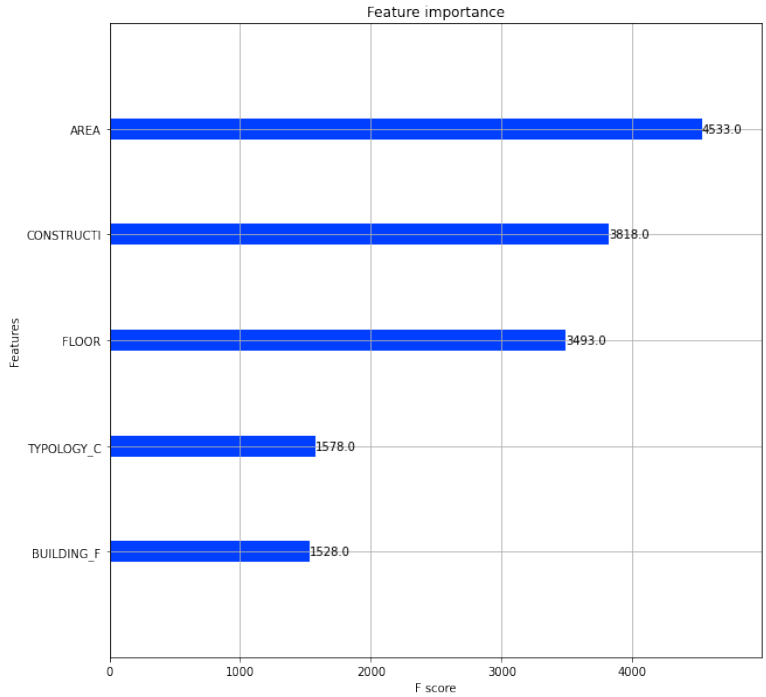


Figure A.0-9 - Most important features for Areeiro (RMSE: 0.452114)

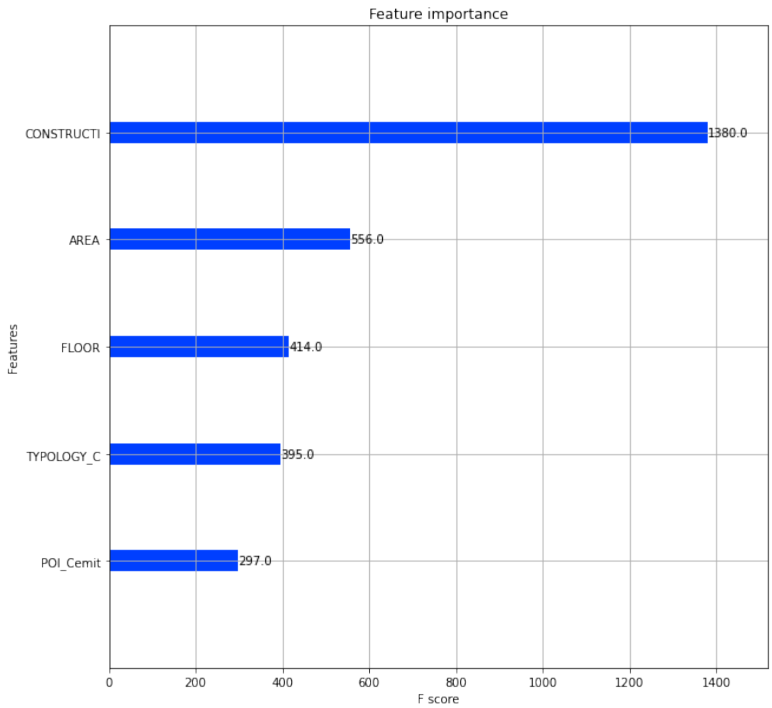


Figure A.0-10 - Most important features for Campo de Ourique (RMSE: 0.393831)

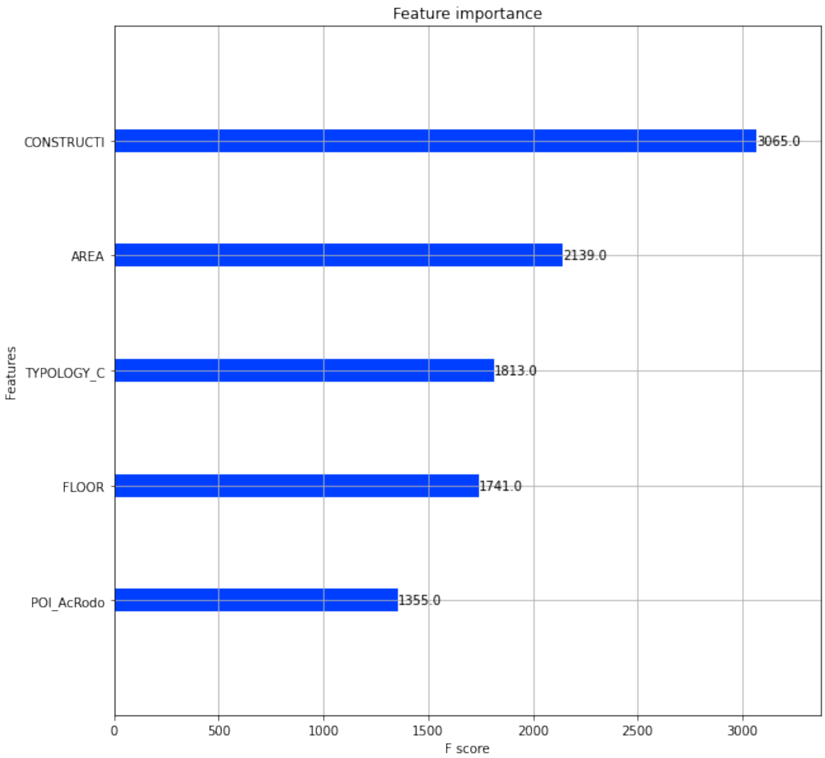


Figure A.0-11 - Most important features for Santo António (RMSE: 0.462867)

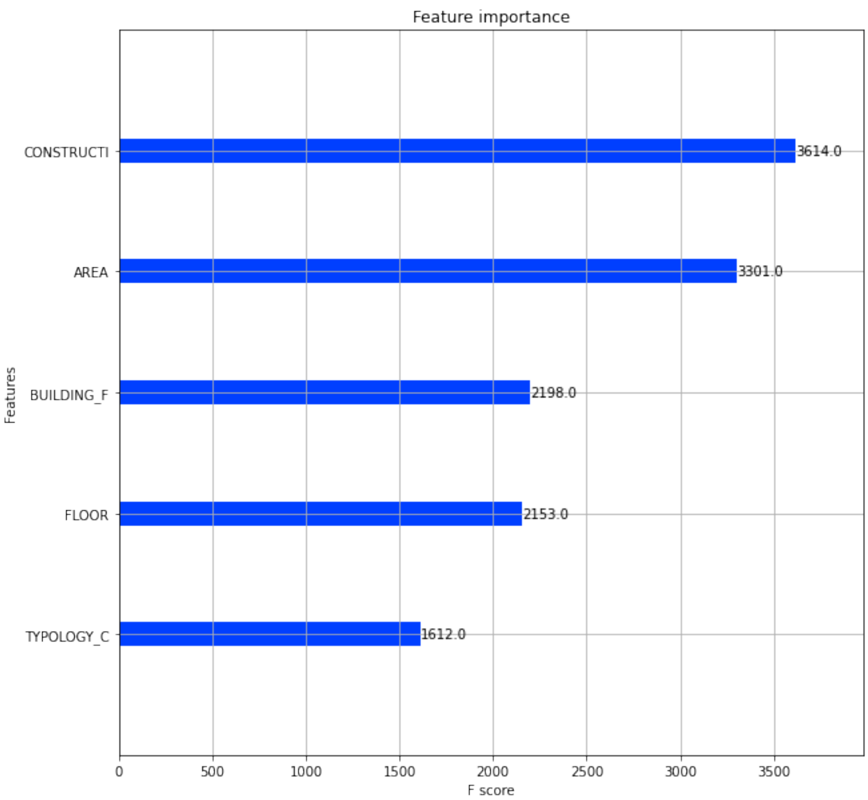


Figure A.0-12 - Most important features for Misericórdia (RMSE: 0.471758)

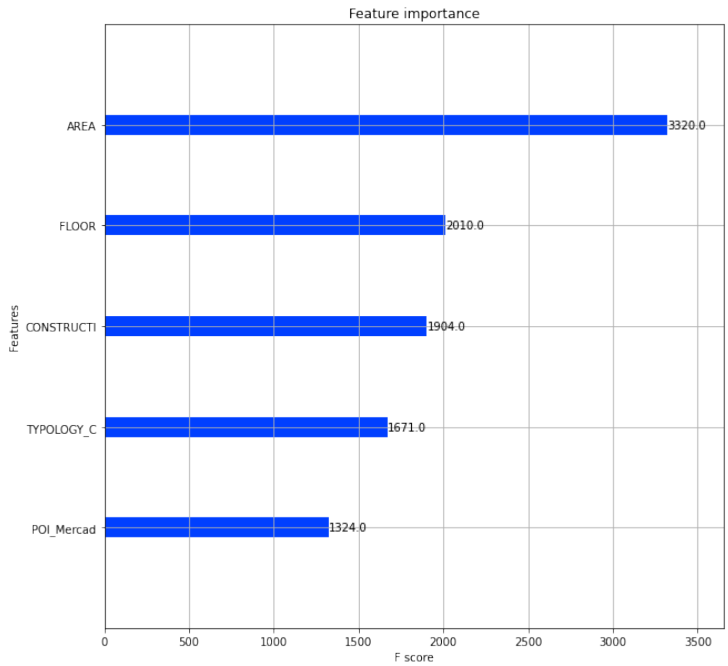


Figure A.0-13 - Most important features for Santa Maria Maior (RMSE: 0.479415)

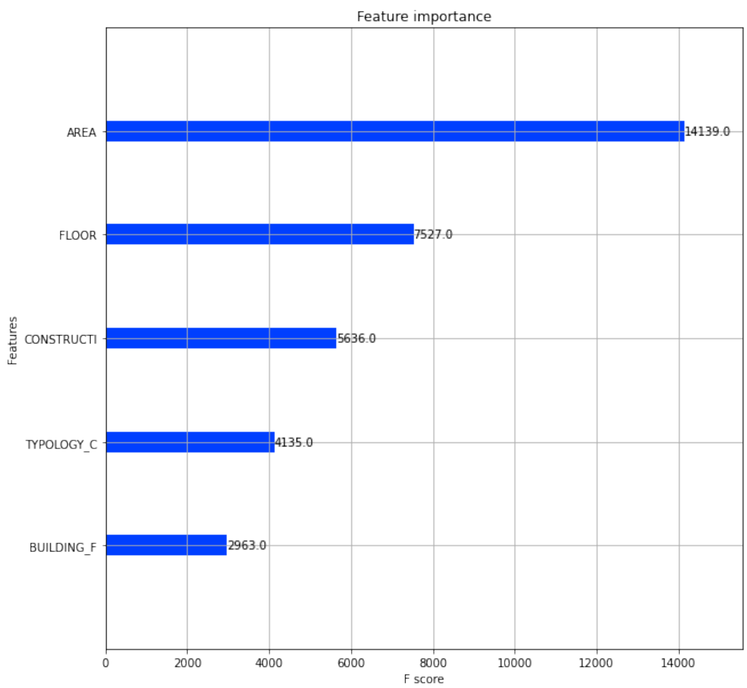


Figure A.0-14 - Most important features for São Vicente (RMSE: 0.495672)

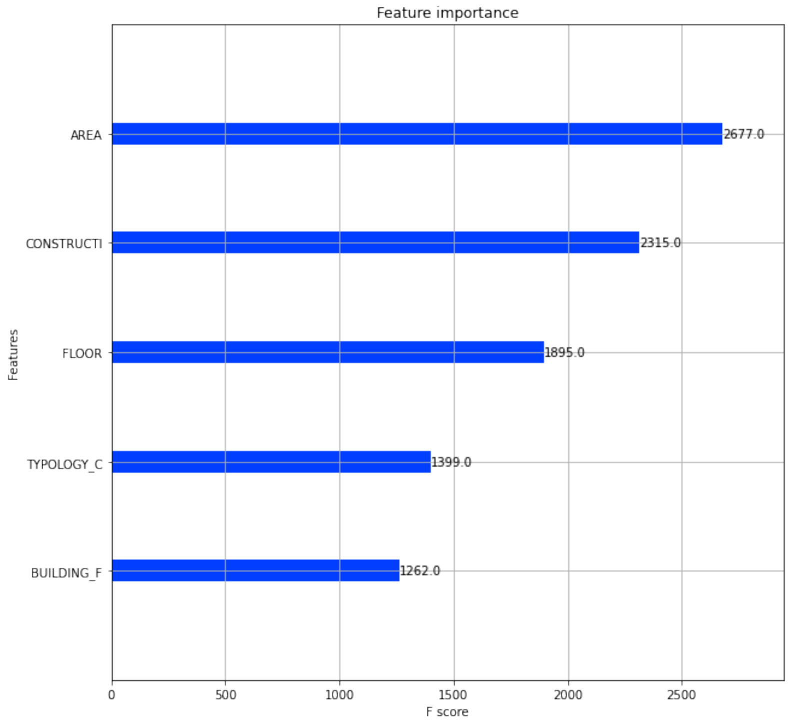


Figure A.0-15 - Most important features for Beato (RMSE: 0.301045)

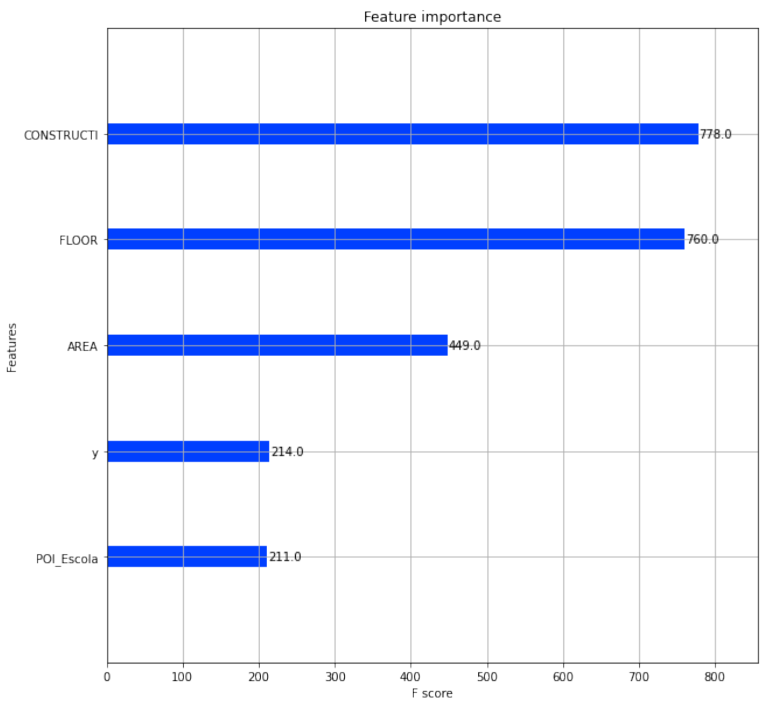


Figure A.0-16 - Most important features for Ajuda (RMSE: 0.363067)

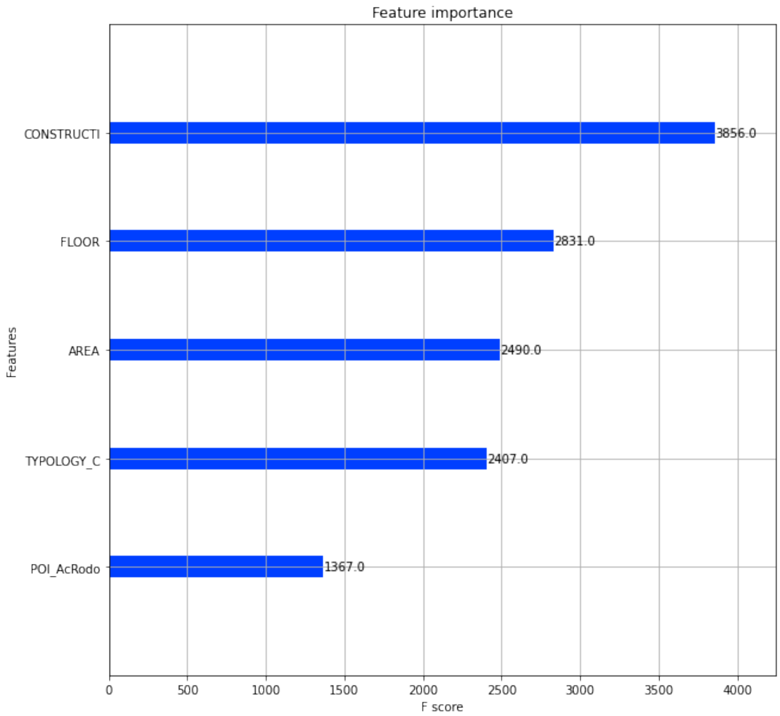


Figure A.0-17 - Most important features for Belém (RMSE: 0.537783)

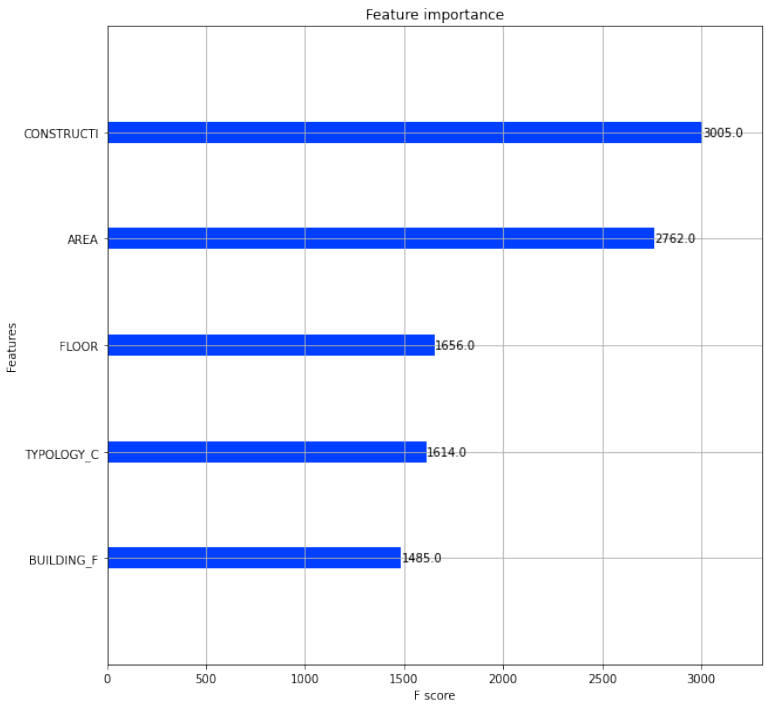


Figure A.0-18 - Most important features for Alcântara (RMSE: 0.397019)

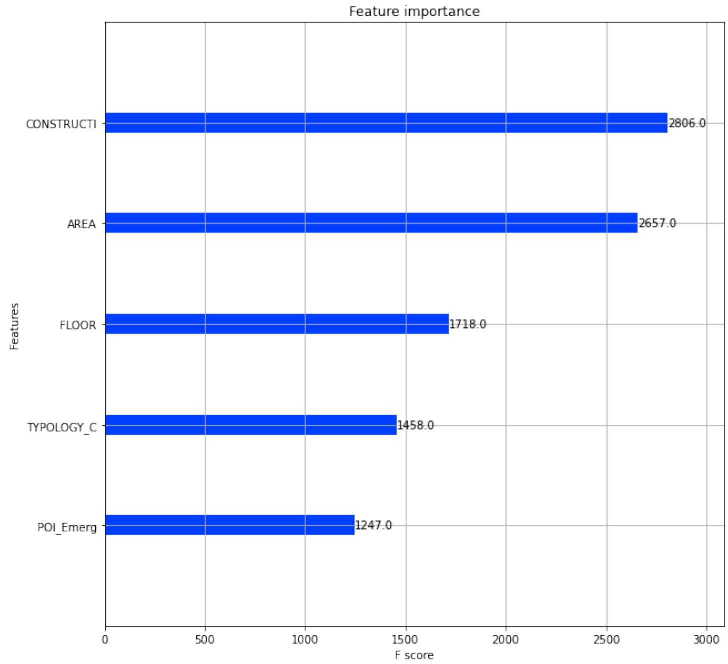


Figure A.0-19 - Most important features for Benfica (RMSE: 0.288159)

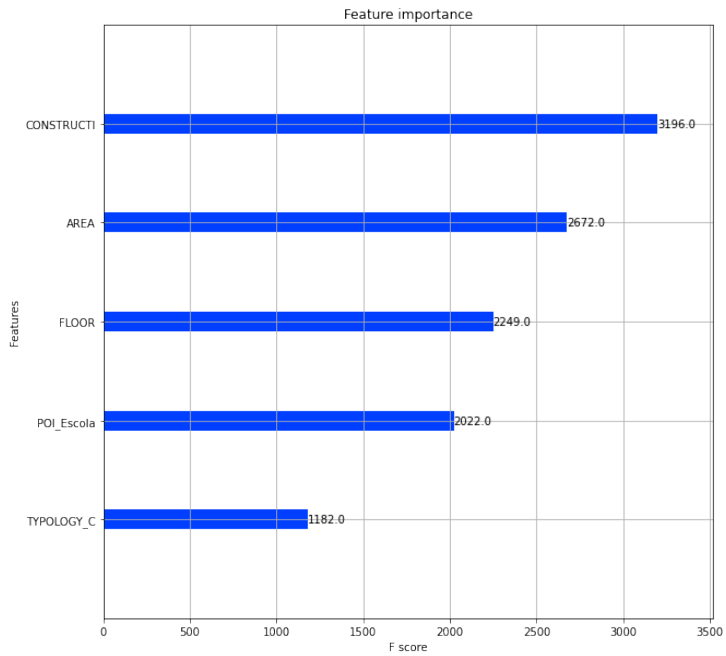


Figure A.0-20 - Most important features for Santa Clara (RMSE: 0.175795)

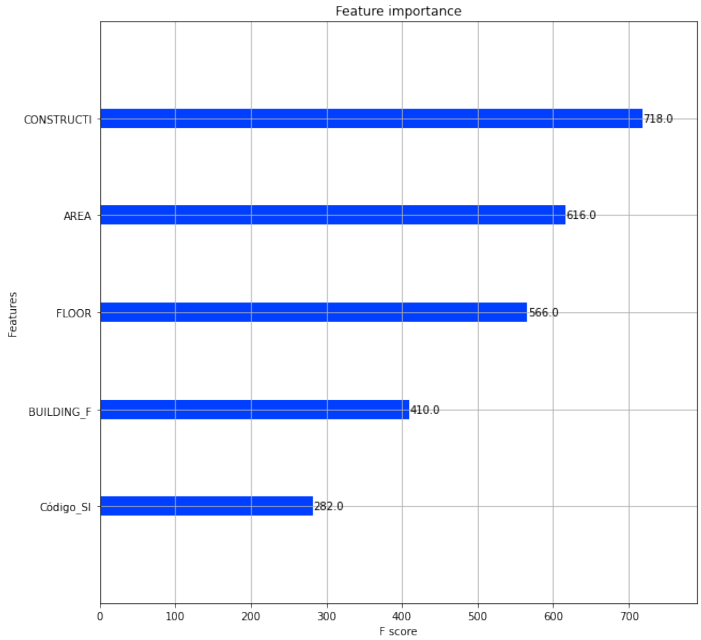


Figure A.0-21 - Most important features for Carnide (RMSE: 0.747987)

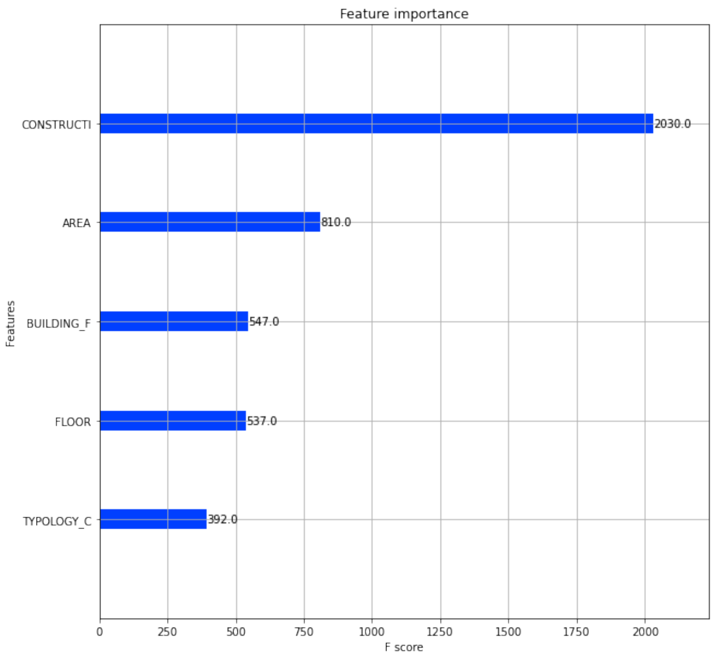


Figure A.0-22 - Most important features for São Domingos de Benfica (RMSE: 0.305182)

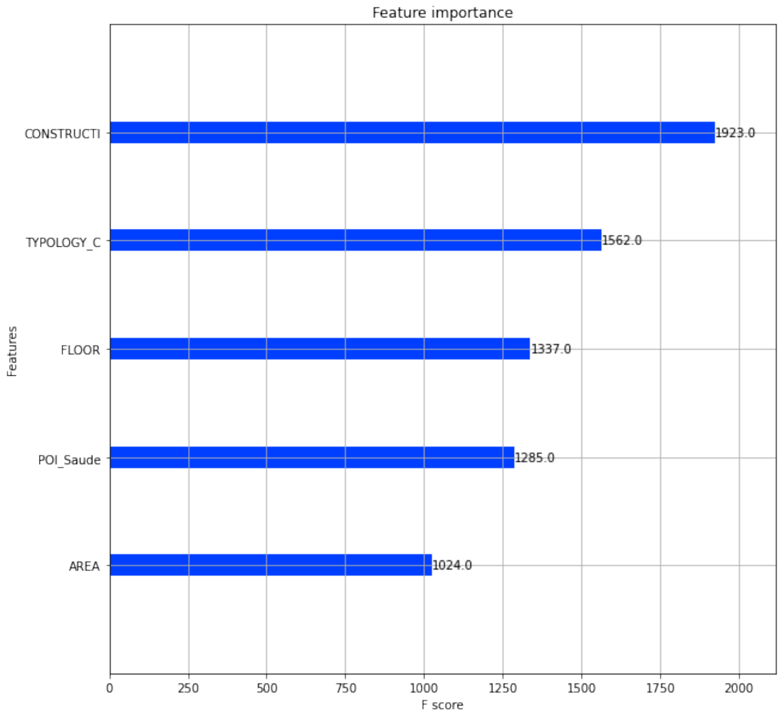


Figure A.0-23 -Most important features for Olivais (RMSE: 0.230601)

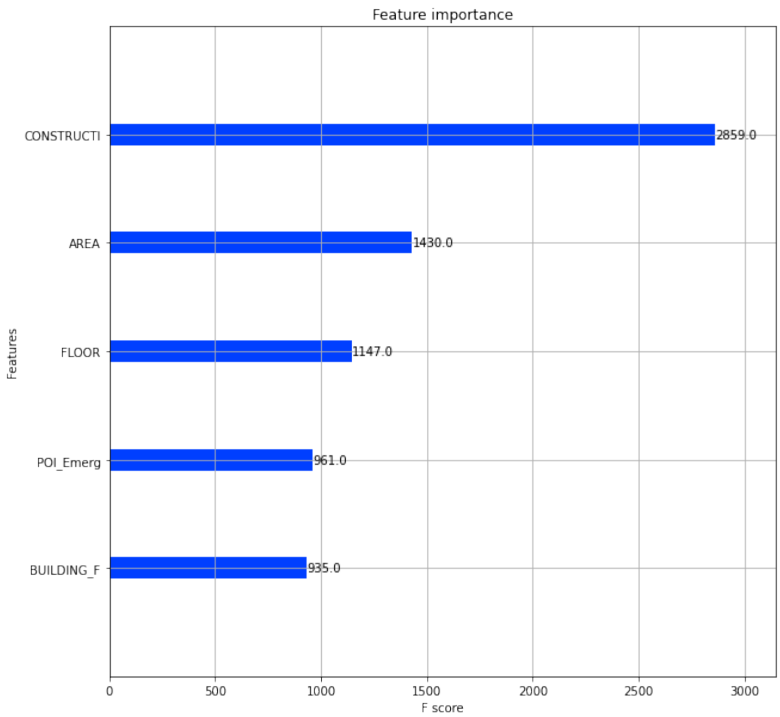


Figure A.0-24 - Most important features for Alvalade (RMSE: 0.425449)

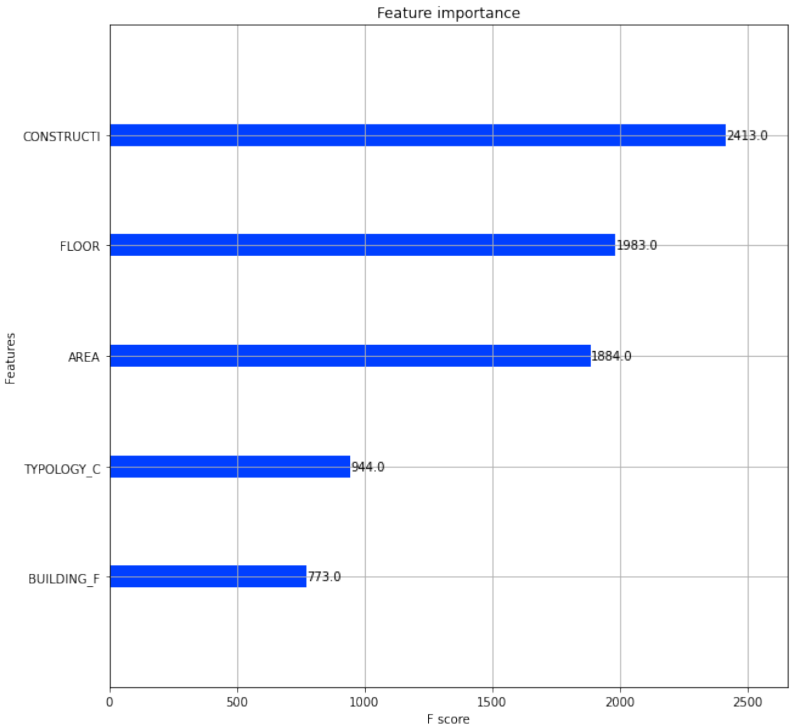


Figure A.0-25 -Most important features for Campolide (RMSE: 0.414885)

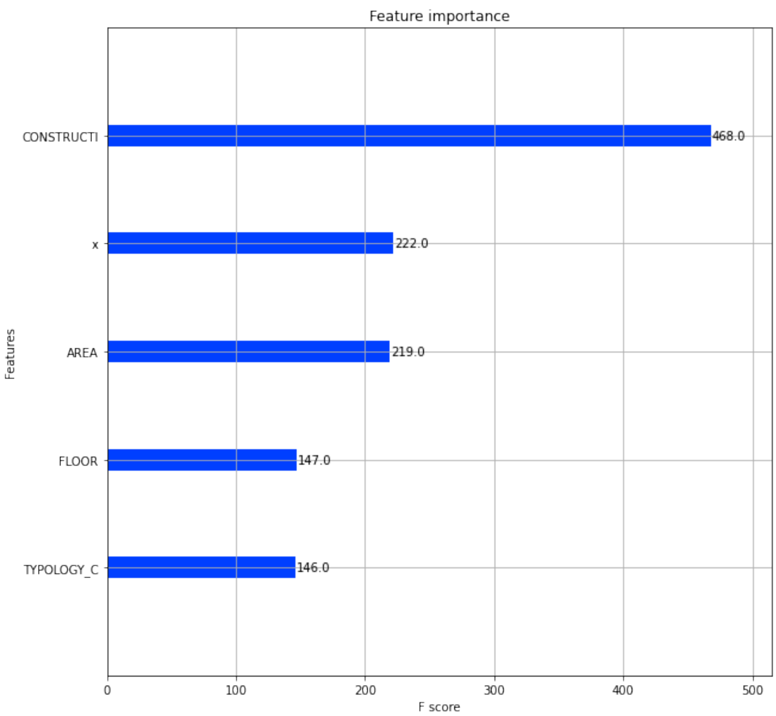


Figure A.0-26 - Most important features for Marvila (RMSE: 0.120865)

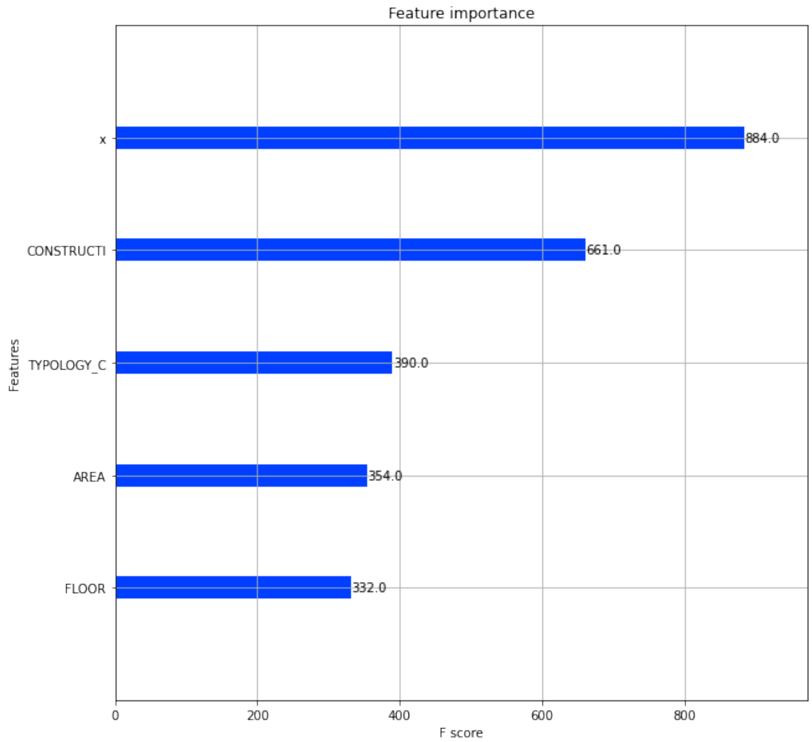


Table A.0-13 - Mode of top mentioned risks in the survey, grouped by parish

Parish	Mode	Sample Size
Ajuda	Sismos	6
Alcântara	Inundações	1
Alvalade	Sismos	16
Areeiro	Sismos	7
Arroios	['Incêndios', 'Sismos']	6
Avenidas Novas	['Assaltos', 'Sismos']	20
Beato	['Assaltos', 'Inundações']	2
Belém	Assaltos	10
Benfica	['Incêndios', 'Sismos']	15
Campo de Ourique	Sismos	6
Campolide	Sismos	21
Carnide	Incêndios	9
Estrela	Sismos	8
Lumiar	Sismos	24
Marvila	['Incêndios', 'Sismos']	2
Misericórdia	Incêndios	1
Olivais	Sismos	6
Parque das Nações	Sismos	6
Penha de França	Incêndios	5

Parish	Mode	Sample Size
Santa Clara	Incêndios	8
Santa Maria Maior	Sismos	1
Santo António	Sismos	5
São Domingos de Benfica	['Incêndios', 'Sismos']	14
São Vicente	['Assaltos', 'Incêndios', 'Sismos']	3

Table A.0-14 - Entities contacted by email

Professional Associations
Ordem dos Engenheiros - Região Sul
Ordem dos Engenheiros Técnicos
Ordem dos Arquitectos - Região Sul
Real Estate Associations
APPII
APPC
AICCOPN
AECOPS
Associação Lisbonense de Proprietários
Associação Nacional de Proprietários
habita! - Associação pelo direito à habitação e à cidade
APEMIP - Associação dos Profissionais e Empresas de Mediação Imobiliária de Portugal
Grupo Habitar - Associação Portuguesa para a Promoção da Qualidade Habitacional
ASSOCIAÇÃO PORTUGUESA DE FUNDOS DE INVESTIMENTO, PENSÕES E PATRIMÓNIOS
FENACHE
COOHABITA - Cooperativa Nacional de Habitação
Civil Parishes
Junta de Freguesia de Alcântara
Junta de Freguesia do Areeiro
Junta de Freguesia de Arroios
Junta de Freguesia das Avenidas Novas
Junta de Freguesia do Beato
Junta de Freguesia de Belém
Junta de Freguesia de Benfica
Junta de Freguesia de Campo de Ourique
Junta de Freguesia de Campolide
Junta de Freguesia de Carnide
Junta de Freguesia da Estrela
Junta de Freguesia do Lumiar
Junta de Freguesia de Marvila

Civil Parishes
Junta de Freguesia da Misericórdia
Junta de Freguesia dos Olivais
Junta de Freguesia do Parque das Nações
Junta de Freguesia da Penha de França
Junta de freguesia de Santa Clara
Junta de Freguesia de Santa Maria Maior
Junta de Freguesia de Santo António
Junta de Freguesia de São Domingos de Benfica
Junta de Freguesia de São Vicente
Junta de Freguesia de Alvalade
Junta de Freguesia da Ajuda
Neighbors' Associations
Associação de Moradores do Bairro de S. João de Brito - Alvalade - Lisboa
Associação de Moradores do Casal Ventoso
Vizinhos de Arroios – Associação de Moradores
Associação de Moradores do Bairro Novo de Carnide e Quinta do Bom Nome
ASS. MORADORES BAIRRO ALTO
ARAL - Associação de Residentes do Alto do Lumiar
Avenidas Novas - Associação de Moradores das Avenidas Novas de Lisboa
ASSOCIAÇÃO PRÍNCIPE + REAL
Associação dos Moradores de Santos
Associação de Moradores do Alto da Serafina
Comissão De Moradores Do Bairro Da Calçada Dos Mestres
Associação Recreativa de Moradores e Amigos do Bairro da Boavista
Associação de Moradores do Bairro do Charquinho
Associação de Moradores do Bairro de Santa Cruz de Benfica e Zonas Contíguas
Associação de Moradores do Bairro das Pedralvas
Associação de Moradores do Bairro do Calhariz de Benfica
Associação de Moradores da Quinta da Granja e Colégio Militar
Associação de Moradores da Monsanto Benfica
Associação de Pais do Jardim de Infância Nº1
Associação de Pais da E.B. 2.3 Pedro de Santarém
Associação de Pais EB1 Jorge Barradas
Associação de Pais e Técnicos do Infantário Pedrita
Associação de Educadores da Escola Básica Quinta de Marrocos
Associação de Pais e EE EB1 n.º 124 e JI n.º2 da Escola Parque Silva Porto
Vizinhos.org
ATENEU DA MADRE DE DEUS
ASSOCIAÇÃO DE MORADORES DAS VILAS OPERÁRIAS DO BEATO
Associação de Moradores das Torres do Alto da Eira
ASSOCIAÇÃO DE MORADORES VIVER MELHOR NO BEATO

Neighbors' Associations
ASSOCIAÇÃO MORADORES E EMPREENDEDORES DO BEATO
Associação Passeio Público
AMBCVL – Associação Moradores Bº Cruz Vermelha Lumiar
Comissão de Moradores do Bairro Azul
Associação de Actividades Sociais do Bairro 2 de Maio
CASES - Cooperativa António Sérgio para a Economia Social
Large companies operating in Lisbon
EDP
CTT
TAP
Other Public Companies and Institutions
CIUL
SRU
CERIS
IRHU

Figure A.0-27 - Copy of the flyers distributed by hand and through mailboxes

ANNEX B

This section will detail the survey questions asked in each group.

A. <u>Demographics</u>

1) Please indicate the total number of household members. *(Options: 1; 2; 3; 4; 5+)*
2) Are any children (<18 years old) currently living in this dwelling? *(Binary: yes/no)*
3) How much is your household's average monthly income?[199] *(Options: <1000€; 1001-2000€; 2001-3500€; 3501-5000€; >5000€)*

<u>The following questions concern the head of household:</u>

4) Please indicate your gender. *(Options: Male; Female; Other)*
5) Please indicate your age group. *(Options: 16–34 years; 35–54 years; >55 years)*
6) Please indicate your marital status. *(Options: Single; Married; Divorced; Widowed; Other)*
7) Education level *(Options: No schooling completed; Elementary school; High school; Bachelor's degree; Master's degree; PhD)*

B. <u>Housing</u>

1) Do you live in Lisbon municipality? *(Binary: yes/no)*
2) Please indicate your address or zip-code. *(Open question)*
3) Please indicate your parish. *(Options: selected from a list of Lisbon's parishes)*
4) What is your house's current ownership status? *(Options: homeowner; private market renter; public housing renter)*
5) <u>If you are a renter</u>, how much do you pay per month? *(Options: <500€; 501-1000€; 1001€-1500€; 1501-2500€; >2500€)*
6) <u>If you are a homeowner</u>, what is your best guess for the house's market value? *(Options: <100k; 100k-250k; 250k-500k; 500-750k; >750k)*
7) Do you know the exact construction year of your building? *(Open: indicate year)*
8) When was your building built? (If you don't know the answer, please provide your best guess) *(Options: <1755; 1775-1880; 1880-1930; 1930-1960; 1960-1985; 1985-2000; 2000-2010; >2010)*
9) What is your dwelling's total surface area? *(Options: < 50 m^2; 51-100m^2; 101-150m^2; 151-200m^2; >200m^2)*
10) How many rooms does it have? *(Options: 0; 1; 2; 3; 4; 5+)*
11) What is your building's construction type? *(Options: masonry, reinforced concrete, steel frame, other, unknown)*
12) How many floors does it have? *(Options: <3; 3-5; 6-8; 9-11; >11)*

C. <u>Perception and action towards risk</u>

1) Rank the following risks to your property from the most concerning (1) to the least concerning (5). *(Options: home robbery; floods; earthquakes; fire; others)*

<u>Classify the following statements from (1) Completely disagree to (5) completely agree</u>

2) "Following the safety of my family, the safety of my home is my number one priority".
3) "If my home is going to be robbed, there isn't much I can do about it"
4) "If floods are going to occur, there is not much my city/community can do to lessen its effects"
5) "If an earthquake is going to harm me, it will, and there isn't much that I can do about it"

[199] As reported by Manganelli et al. (2018), some individuals may be reluctant to declare personal income. The formulation in income brackets is an attempt to overcome it

6) “If an earthquake is going to occur, there is not much my city/community can do to lessen its effects”
7) “If an earthquake is going to occur, the government will take care of me.”
8) “If an earthquake is going to occur, my insurance policy will take care of me.”
9) “My building is safe against earthquakes.”
10) “All recent buildings are safe to earthquakes due to restrictive building codes.”

11) Have you experienced an earthquake before? *(Binary: yes/no)*
12) Do you have earthquake insurance? *(Binary: yes/no)*
13) If yes, how much do you pay per year? *(Open question)*
14) Select which of the following are valid earthquake risk mitigating measures. *(Options: (a) Secure heavy objects to walls and floors, such as shelves, bookcases, cabinets, and water heaters; (b) In case of a seismic event, get into the elevator as soon as possible to exit the building (c) In case of a seismic event, shelter against an inside wall, or under sturdy furniture such as heavy desks or tables; (d) In case of a seismic event, shelter against an outside wall, or near a window, so you can see* what *is going outside.)*

D. Willingness to pay for mitigating measures.

1) Would you be willing to pay more for a house safer to earthquake risk? (Binary: yes/no)
2) If yes, how much would you be willing to pay for it? As a percentage of house market price/rent *(Options: 0%; 1-5%; 5-10%; 10-15%; >15%)*
3) If no, why is that? *(Options: It is a waste of money; I am not worried about earthquake risk; My building is already safe; Other)*
4) Would you be willing to have a discount on price, for a house less resistant to earthquakes? *(Binary: yes/no)*
5) Would you be willing to pay for the structural retrofitting of your building? *(Binary: yes/no)*
6) If yes, how much would you be willing to pay for it? As a percentage of house market price/rent *(Options: 0%; 1-5%; 5-10%; 10-15%; >15%)*
7) If yes, how much would you be willing to pay it? Please indicate in euros *(Open-ended question)*
8) If yes, how would you prefer to do it? (Options: as a percentage of the municipal property tax (IMI); as a one-time payment to the municipality; direct payment to contractors with tax benefits)