



## Article

# Analyzing Economic and Social Inequalities in Housing: A Visual Storytelling Case Study in Portugal

Afonso Crespo<sup>1</sup>, José Barateiro<sup>2,\*</sup>  and Elsa Cardoso<sup>1,3</sup> 

<sup>1</sup> Instituto Universitário de Lisboa (ISCTE-IUL), Av. das Forças Armadas, 1649-026 Lisboa, Portugal; afonso\_crespo@iscte-iul.pt (A.C.); elsa.cardoso@iscte-iul.pt (E.C.)

<sup>2</sup> NOVA Laboratory for Computer Science and Informatics (NOVA LINCS), Centro de Investigação Tecnológica do Algarve (CINTAL), Universidade do Algarve (UALg), Campus de Gambelas, 8005-139 Faro, Portugal

<sup>3</sup> Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura (ISTAR-Iscte), Av. das Forças Armadas, Edifício II, 1649-026 Lisboa, Portugal

\* Correspondence: jebarateiro@ualg.pt

## Abstract

Housing inequalities remain a major challenge for contemporary urban governance, as they combine economic, social, spatial, and demographic dynamics that are difficult to capture through single indicators. This paper develops a data-driven assessment of housing inequalities in Portugal between 2015 and 2025, drawing on official national and European statistics and applying a Business Intelligence (BI) and urban analytics framework oriented towards policy monitoring. Official data from Statistics Portugal and Eurostat are integrated through an analytical pipeline including automated extraction via public APIs, data enrichment, and visual analytics. The workflow follows a CRISP-DM-inspired structure, creating a set of normalized indicators to capture different dimensions of housing conditions. The results point to a structurally polarized housing market. Housing valuations increased across all regions, but at uneven rates, reinforcing territorial disparities rather than convergence. Metropolitan and tourism-oriented regions experienced faster appreciation and indirect effects, while year-over-year growth in completed dwellings slowed after 2021–2022, indicating an uneven supply response. Beyond its empirical findings, the primary contribution of this study lies in demonstrating how BI and data science methodologies can be operationalized to monitor housing inequalities using official statistics. The proposed framework is replicable and can be adapted to other territorial and policy contexts.

**Keywords:** housing inequalities; territorial disparities; urban analytics; business intelligence; visual analytics; official statistics



Academic Editors: Jungho Baek and Manfred Max Bergman

Received: 10 January 2026

Revised: 30 March 2026

Accepted: 30 April 2026

Published: 15 May 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

In the current socioeconomic landscape, housing is established as a fundamental right in the Constitution of the Portuguese Republic, playing a key role in promoting social cohesion and addressing structural inequalities. However, in recent decades, complex economic, demographic, and urban dynamics have intensified disparities in access to adequate housing conditions. Municipalities and central governments face rising pressures to mitigate the effects of soaring rental prices, a scarcity of affordable options, and gentrification phenomena, which have contributed to the intensification of urban polarization and the vulnerability of large population segments. These challenges are further amplified in metropolitan areas where the cost of living increasingly outpaces income growth. As the

housing market continues to fluctuate, the ability to systematically identify and monitor these disparities has become essential for effective governance. In this context, monitoring social and economic inequalities becomes crucial, to raise public awareness and as pathway to a more inclusive society.

Housing inequalities constitute one of the most significant obstacles to social and economic progress. They affect not only access to housing but also the quality of living conditions and their consequences for health, education, and community integration. The lack of adequate housing restricts both individual and collective opportunities, accentuating cycles of poverty and social exclusion while undermining territorial cohesion. In a context characterized by significant shifts in housing markets and demographic pressure in urban centers, a systematic approach is crucial to monitor these inequalities. Existing official data sources offer valuable insights into these dynamics but require careful integration. The resulting analysis can be used by public authorities to gain a deeper understanding of the structural challenges of this sector and create effective public policies aimed at mitigating social and economic inequalities.

PLANAPP—the Centre for Planning and Public Policy Evaluation (<https://planapp.gov.pt/>)—is a state body that aims to support the design and implementation of public policies and prospective analysis. Its main objective is to support and strengthen the different phases of the public policy cycle (planning, monitoring, evaluation, and impact analysis) with the aim of improving public policy design and implementation. This study is a result of a collaboration between academia and PLANAPP, focused on developing an analytical framework to monitor social and economic inequalities in housing. The study uses public data from official and internationally recognized sources, such as Statistics Portugal (<https://www.ine.pt/>) (INE, from Portuguese National Statistics Institute); Eurostat (<https://ec.europa.eu/eurostat>); the Organization for Economic Co-operation and Development (OECD) (<https://www.oecd.org/>); and the Strategy and Planning Office (<https://www.gep.mtsss.gov.pt/>) (GEP, from the Portuguese name) from the Ministry of Labor, Solidarity, and Social Security. The use and integration of multiple official sources enables a richer analysis and a comparison of indicators with different scales, granularity, and context.

This study provides a data preparation and visualization approach that can be applied to other social and economic inequality dimensions or adapted to other contexts to analyze official statistical data for better governance. This paper presents an analysis of housing inequalities in Portugal, examining the evolution of bank valuation values, the structure of the housing stock, and the dynamics of new construction between 2015 and 2025. Through this data-driven approach, this study contributes to ongoing efforts to enhance public policies and provides a replicable framework for understanding the reality of a national housing market.

Despite the growing research on housing inequalities, relatively few studies translate official statistical data into integrated and reproducible analytical workflows supporting continuous monitoring across territorial scales. In practice, many analyses remain descriptive or confined to a single spatial level, limiting their usefulness for tracking indirect and regional behavior over time. This paper addresses this gap by combining a multi-scale empirical analysis with a Business Intelligence (BI) and visual analytics framework designed for policy monitoring.

The remainder of this paper is structured as follows: Section 2 presents related work and a literature review on social and economic inequalities in housing. Section 3 describes the study area, data, and method adopted. Section 4 presents the results, discussing their significance regarding territorial polarization and housing supply dynamics. Finally, Section 5 outlines the main conclusions and contributions of this study.

## 2. Related Work

Housing is widely recognized in the academic literature not merely as a physical shelter but also as a determinant of health, well-being, and social inclusion [1]. Access to adequate housing is intrinsic to the realization of the United Nations Sustainable Development Goals (SDGs), particularly SDG 11 on sustainable cities and communities and SDG 10 on reducing inequalities. Despite its formal recognition as a basic human right, housing remains a major source of inequality worldwide.

Housing is one of the key domains where socioeconomic inequalities are reproduced and reinforced. Housing inequality research documents persistent disparities in access, affordability, and housing quality, alongside the need for more inclusive housing policies and sustainable urban planning frameworks [1]. Empirical evidence indicates that vulnerable social groups face elevated risks of housing discrimination and deprivation even after accounting for socioeconomic characteristics, pointing to structural barriers within housing markets [2]. Evidence from Northern Europe also suggests that urban densification strategies do not automatically promote social inclusion, as they may entail socio-environmental trade-offs that require complementary policy measures to mitigate exclusionary outcomes [3].

Housing inequality is characterized by the unequal availability of safe, affordable, accessible, and dignified living spaces [4]. James et al. conceptualize housing inequality as an uneven distribution of housing opportunities and resources, manifested through inadequate living conditions, unaffordable rents or mortgages, and structural barriers limiting access to desirable residential environments [5]. These disparities are deeply intertwined with broader socioeconomic and demographic factors. Characteristics such as age, gender, ethnicity, education, and income are key determinants influencing both the probability of homeownership and the housing quality [6–9]. Furthermore, the rapid urbanization and market dynamics observed in recent decades have intensified these issues. The increasing demand in urban areas has driven housing prices beyond the reach of lower-income populations, increasingly relegating them to substandard living arrangements [10].

A critical dimension of housing inequality is discrimination, which creates barriers that extend beyond financial capacity. The literature distinguishes between taste-based discrimination, where housing providers exhibit preferences for members of their own social group, and statistical discrimination, whereby stereotypes are used to infer the reliability of potential tenants [11]. These mechanisms are particularly evident in the rental market.

Experimental studies across Europe, such as the work of [12] in Ireland, have revealed evidence of ethnic and gender bias, demonstrating that minority groups face disproportionate difficulties in securing dignified housing compared to the majority population. Closely related is the concept of housing deprivation, defined as a form of material deprivation associated with physical structural deficiencies (such as dampness or inadequate heating) and a lack of basic facilities [13]. This form of deprivation is not evenly distributed for all societal groups. Recent analyses also in Ireland, based on the Survey on Income and Living Conditions (SILC), indicate that vulnerable groups, including young adults aged 18 to 34, persons with disabilities, and single-parent households, are significantly more likely to experience these precarious conditions, even when controlling for other socioeconomic variables [2].

Parallel to social exclusion, the economic dimension of housing has become more prominent in European policy debates, often framed as a persistent “housing crisis.” Le Goix et al. show that housing costs in major European cities have increased at a faster pace than household incomes since the 1990s, generating a widening gap in housing accessibility [14]. The concept of housing regimes, defined as the systematic interaction

between housing systems, tenure structures, and welfare-state arrangements, has been used to explain these dynamics [15]. These regimes have undergone substantial transformation over the past decades, particularly in relation to tenure structures and market regulations. There has been a notable decline in homeownership in certain countries, accompanied by an unprecedented expansion of the private rental sector and the emergence of corporate landlords [16]. In contrast, countries like the Netherlands and France have attempted to stabilize ownership and affordability through regulatory reforms and renewed investments in social housing [17,18]. To assess these trends quantitatively, the Price-to-Income Ratio (PIR) remains a widely used indicator of housing affordability, relating housing prices to income levels, offering a clear metric for assessing financial pressure on households and informing housing policy intervention [19,20].

Recent European (EU) research shows that socioeconomic and housing inequalities are shaped by interactions between income distribution, household characteristics, territorial contexts, and institutional frameworks. Comparative studies reveal substantial EU cross-national variation in inequality of opportunity, particularly in labor markets, with estimates ranging from relatively low to markedly high levels, and little evidence that urban areas systematically function as “social elevators” relative to less urbanized regions [21]. These findings challenge simplified narratives regarding the equalizing effects of urbanization and highlight the importance of national and regional policy contexts.

Beyond aggregate income measures, the literature also focuses on household-level mechanisms that shape inequality. Unequal intra-household resource allocation has been shown to bias conventional poverty and inequality estimates downward, masking higher levels of deprivation and disproportionately disadvantaging women, particularly in Central and Eastern Europe [22]. Household composition, occupational sector, and degree of urbanization also emerge as key determinants of income inequality, with rural areas and low-wage sectors consistently associated with lower income levels across European countries [23]. Educational similarity bias in partner selection (i.e., mating) further amplifies household income inequality, especially in Eastern European contexts, while its effects appear more limited in Nordic countries due to stronger gender equality and more redistributive welfare regimes [24]. Territorial inequalities and measurement challenges also feature prominently in the European literature. Studies on regional poverty and inequality indicators demonstrate that standard measures often suffer from high uncertainty at sub-national scales, limiting their usefulness for policy design and evaluation [25].

Comparative and intergenerational perspectives highlight the growing role of housing wealth in deepening socioeconomic inequalities. Cross-city analyses reveal substantial variation in housing affordability across Europe, reflecting dynamics between housing prices and household incomes [14]. The increasing concentration of housing wealth reinforces both intergenerational and territorial inequalities, strengthening calls for fiscal and housing policies aimed at a more equitable redistribution of capital gains [26]. Multidimensional and extreme poverty approaches further identify specific household profiles (such as unemployed individuals approaching retirement age and living alone) as being at heightened risk of deprivation, underscoring the importance of targeted housing and social policy interventions [27].

Analyzing these inequalities requires a granular understanding of income distribution that goes beyond national averages, which often mask significant territorial heterogeneity. Ref. [25] demonstrated that regional disparities in poverty and inequality are profound, with regions such as Southern Italy and Eastern Germany exhibiting much higher vulnerability than their respective national averages. Their findings suggest that standard indicators may suffer from high uncertainty at the sub-national level, advocating for the use of small-area estimation models, such as the Fay–Herriot model, to produce more

stable estimates that can support targeted policy interventions. Equally important is the intra-household distribution of resources. Using small-area estimation techniques has improved the reliability of regional indicators, strengthening the empirical basis for targeted poverty-reduction strategies. Complementary approaches propose composite poverty indices to identify particularly vulnerable regions within the European Union, thereby supporting more effective spatial targeting of cohesion policies [28].

While this literature provides important insights into the drivers and consequences of housing inequalities, fewer studies focus on how these dynamics can be operationalized through integrated analytical frameworks using official statistics. In particular, challenges related to data fragmentation, scale mismatches, and indicator comparability remain under-explored. This gap is especially relevant for policy-oriented analysis, where reproducible and scalable monitoring tools are required.

### 3. Data and Method

#### 3.1. Study Area

This research focuses on the housing market dynamics in Portugal, which is currently experiencing a severe housing crisis, marked by shifts in real estate valuation and social accessibility. The analysis covers the entire national territory, focusing on the asymmetries between the coastal urban areas and the interior regions. According to the 2021 Census by Statistics Portugal (INE), the country is characterized by a “two-speed” demographic and economic reality. The metropolitan areas, specifically in Lisbon, Porto, and Algarve, concentrate most of the population, economic activity, and tourism pressure. These regions have witnessed the most significant accelerations in bank valuation values, creating a widening gap compared to the interior.

Identifying and understanding territorial disparities is critical as housing is not only a market asset but also a determinant of social cohesion. The study area is analyzed through a multidimensional approach, utilizing the NUTS 2024 classification (Nomenclature of Territorial Units for Statistics). The analysis drills down from the broad NUTS II regions (e.g., *Norte*, *Centro*, *Grande Lisboa*, *Península de Setúbal*, *Alentejo*, and *Algarve*) to the NUTS III level and, where granular data permits, to specific municipalities. This granular approach is essential to unmask local inequalities, such as the indirect effects spreading from Lisbon to the adjacent area of *Península de Setúbal*, which are hidden in national averages. Figure 1 displays the different Portuguese regions at the NUTS II level.



**Figure 1.** Portuguese NUTS II regions (figure adapted from [https://app.escolavirtual.pt/fs/c/6/8/d/configuracao\\_nuts\\_2024.pdf](https://app.escolavirtual.pt/fs/c/6/8/d/configuracao_nuts_2024.pdf), accessed on 2 March 2026).

### 3.2. Data Sources

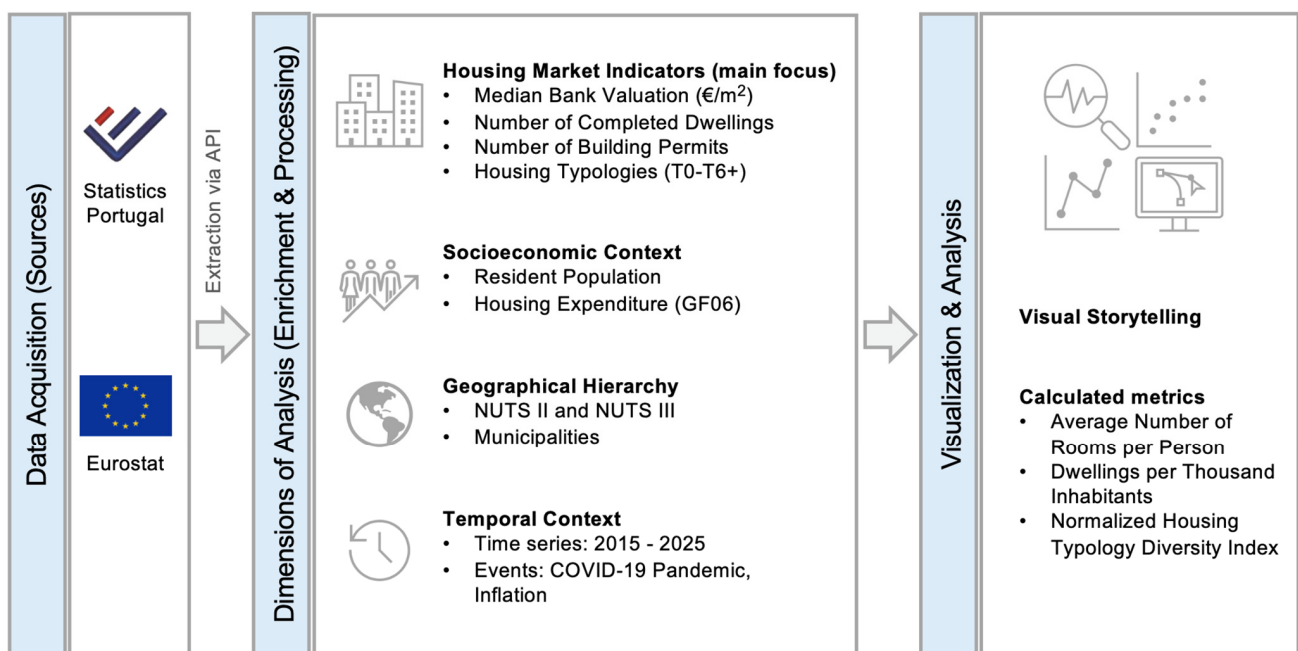
This study relies exclusively on official data sources, integrating two primary data streams:

- **INE (Statistics Portugal):** This source provides detailed national data on housing characteristics, construction activity, and bank valuations. Key datasets include indicators on the number of completed dwellings, building permits, median bank appraisal values per square meter, and housing stock characteristics (e.g., typology and area). These data are used to characterize the supply side of the market and the financial valuation of properties across Portuguese municipalities;
- **Eurostat:** This source contextualizes the Portuguese reality within the European landscape, specifically regarding public expenditure on housing and harmonized social indicators. This allows for a comparative analysis of how Portugal positions itself in terms of public investment and housing affordability relative to other EU Member States.

By combining national micro-data with EU-level indicators, this study builds a multi-dimensional dataset that captures local housing dynamics, while positioning them within broader EU trends.

### 3.3. Method

Figure 2 illustrates the methodological workflow developed for this study, transforming raw data into actionable insights. The approach is grounded in the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology, adapted for social data analysis and public policy monitoring.



**Figure 2.** Overview of this study's methodological workflow.

As depicted in Figure 2, the workflow is organized into three sequential stages:

1. **Data Acquisition (Sources):** The process begins with the automated ingestion of data from INE and Eurostat via their respective APIs. This stage involves handling query parameters to retrieve specific indicators related to banking valuations, construction statistics, and government expenditure;

2. Dimensions of Analysis (Enrichment and Processing): This is the core transformation phase. Raw data is cleaned, normalized, and enriched across four dimensions:
  - a. Housing Market Indicators: The primary focus, containing the hard metrics of value and supply;
  - b. Socioeconomic Context: Adding layers of population density and income to understand the “demand” side;
  - c. Geographical Hierarchy: Mapping data to NUTS II, NUTS III, and Municipalities (Concelhos, in Portuguese) to allow for spatial analysis;
  - d. Temporal Context: Organizing data into consistent time series (2015–2025) to identify trends and shocks (e.g., COVID-19 and inflation);
3. Visualization and Analysis: The final output involves the calculation of additional indicators and the generation of visual storytelling elements to detect inequality trends and patterns.

### 3.4. Data Understanding

The Data Understanding phase focused on exploring the structure, granularity, and quality of the available data. An initial assessment of the INE API revealed that while data is extensive, it is often fragmented across different levels of detail that require careful mapping and aggregation. For instance, the indicator for Bank Valuation allows for segmentation by “Type of Building” (Apartments vs. Houses) and “Geography” (NUTS/Municipality).

An important outcome of this phase was the identification of temporal and spatial gaps. Unlike typical “big data” scenarios where missing values might be imputed, in the context of territorial inequality, a missing value (e.g., no bank valuations in a rural municipality for a specific month) is a finding that indicates a dormant market or a lack of liquidity. Therefore, the strategy adopted was to preserve these nulls to highlight regional exclusion rather than masking them with artificial averages.

Data retrieval was performed through direct service invocations to the INE public API, where each request explicitly specifies the indicator code (varcd) and the relevant analytical dimensions. Figure 3 shows the JSON response of calling the API (the parametrized service endpoint used is [https://www.ine.pt/ine/json\\_indicador/pindica.jsp?op=2&varcd=0012503&Dim1=S7A2021&Dim2=111&Dim3=3&Dim4=2&lang=PT](https://www.ine.pt/ine/json_indicador/pindica.jsp?op=2&varcd=0012503&Dim1=S7A2021&Dim2=111&Dim3=3&Dim4=2&lang=PT), accessed on 18 February 2026 where parameter varcd=0012503 identifies the housing indicator under analysis, Dim1=S7A2021 defines the period of reference, and Dim2=111 constrains the extraction to a specific territorial unit at the NUTS III level. The remaining parameters (Dim3 and Dim4) correspond to structural housing characteristics, namely the number of divisions and the dwelling area class, while the parameter lang=PT ensures consistency in metadata labelling.

This parameterized structure was consistently applied across all INE indicators used in the study, with variations only in the number and type of dimensions, depending on the statistical definition of each indicator. Such an approach guarantees full control over temporal and spatial disaggregation while maintaining comparability across datasets.

A similar approach was adopted for the extraction of EU-level data from Eurostat. Data was accessed through the Eurostat Statistics API, which allows programmatic queries with explicit control over temporal, spatial, and thematic dimensions. In this study, the dataset gov\_10a\_exp (the dataset is available at [https://ec.europa.eu/eurostat/databrowser/product/page/GOV\\_10A\\_EXP](https://ec.europa.eu/eurostat/databrowser/product/page/GOV_10A_EXP), with a persistent identifier [https://doi.org/10.2908/GOV\\_10A\\_EXP](https://doi.org/10.2908/GOV_10A_EXP), accessed on 18 February 2026, was used, focusing on the COFOG (COFOG is an international classification of general government expenditure by function; GF06 refers to *Housing and community amenities*) 1999 function *Housing and community amenities* (GF06).

```

{
  "IndicadorCod": "0012503",
  "IndicadorDsg": "Alojamentos familiares clássicos de residência habitual (N.º) por
Localização geográfica à data dos Censos [2021] (NUTS - 2024), Escalão de divisões e Escalão de área
útil; Decenal - INE, Recenseamento da população e habitação - Censos 2021",
  "MetaInfUrl": "https://www.ine.pt/bddXplorer/htdocs/minfo.jsp?var_cd=0012503&lingua=PT",
  "DataExtracao": "2026-02-19T11:00:08.299Z",
  "DataUltimoAtualizacao": "2024-03-19",
  "UltimoPref": "2021",
  "Dados": {
    "2021": [
      {
        "geocod": "111",
        "geodsg": "Alto Minho",
        "dim_3": "3",
        "dim_3_t": "Com 3 divisões",
        "dim_4": "2",
        "dim_4_t": "30 m² - 39 m²",
        "ind_string": "657",
        "valor": "657"
      }
    ]
  }
},

```

**Figure 3.** Excerpt of the INE API service response used for data extraction. ([https://www.ine.pt/ine/json\\_indicador/pindica.jsp?op=2&varcd=0012503&Dim1=S7A2021&Dim2=111&Dim3=3&Dim4=2&lang=PT](https://www.ine.pt/ine/json_indicador/pindica.jsp?op=2&varcd=0012503&Dim1=S7A2021&Dim2=111&Dim3=3&Dim4=2&lang=PT), accessed on 22 February 2026).

Figure 4 shows the JSON response of calling the API (the parametrized service endpoint used is [https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/gov\\_10a\\_exp?format=JSON&lang=en&cofog99=GF06&geo=PT&time=2020](https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/gov_10a_exp?format=JSON&lang=en&cofog99=GF06&geo=PT&time=2020), accessed on 18 February 2026, where the parameter `gov_10a_exp` identifies the statistical dataset under analysis. The parameter `cofog99=GF06` restricts the extraction to public expenditure associated with housing and community amenities, ensuring thematic alignment with the objectives of this study. The parameter `time=2020` defines the temporal reference, while `geo=PT` limits the extraction to a specific territorial unit at the national level. The parameters `format=JSON` and `lang=en` specify the response format and the language used for metadata labelling, respectively. This request retrieves public expenditure on housing and community amenities for a specific country and year, following the same principle of explicit dimensional control applied to the INE data. The same parameterized structure was used consistently across all Eurostat indicators included in the analysis, with adjustments limited to the temporal range (2019–2025); geographical scope; and, where applicable, additional analytical dimensions such as the unit of measurement or institutional sector.

```

{
  "version": "2.0",
  "class": "dataset",
  "label": "General government expenditure by function (COFOG)",
  "source": "ESTAT",
  "updated": "2026-02-02T23:00:00+0100",
  "value": {
    "1": 287.2,
    "4": 2.1,
    "2": 21.9,
    "5": 0.7,
    "13": 14.8,
    "18": 43.9,
    "19": 41.0,
    "27": 4.3,
    "24": 261.3,
    "0": 413.5,
    "3": 415.6,
    "29": 570.2,
    "31": 570.2,
  }
}

```

**Figure 4.** Excerpt of the Eurostat API service response used for data extraction. ([https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/gov\\_10a\\_exp?format=JSON&lang=en&cofog99=GF06&geo=PT&time=2020](https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/gov_10a_exp?format=JSON&lang=en&cofog99=GF06&geo=PT&time=2020), accessed on 22 February 2026).

Table 1 presents the key variables included in the final dataset, detailing their source, measurement unit, and role in the analysis. This study adopts a hierarchical territorial classification (NUTS II, NUTS III, and municipality) for spatial aggregation. The empirical analysis is restricted to Mainland Portugal (NUTS I) due to data limitations affecting the remaining NUTS I regions, namely the autonomous regions of the Azores and Madeira.

**Table 1.** Housing inequality dataset description.

Variable Name	Source	Measurement	Description
Median Bank Valuation (Cod: 0012248)	INE	€/m <sup>2</sup>	The median value per square meter of dwellings is valued for mortgage purposes. Serves as the primary proxy for market prices. Detailed level: Geography (Municipality) and Time (2019–2025).
Number of Completed Dwellings (Cod: 0013255)	INE	Number	Number of residential units completed in new constructions or rehabilitations during the reference year. Detailed level: Geography (Municipality) and Time (2015–2025).
Number of Building Permits (Cod: 0012778)	INE	Number	Leading indicator of future supply, representing the number of licenses issued for new work. Detailed level: Geography (Municipality) and Time (2019–2025).
Number of Classic Dwellings of Habitual Residence (Cod: 0012503)	INE	Number	Number of conventional family dwellings used as usual residence, by geographic location. Also named housing stock. Detailed level: Geography (NUTS 2024), Time (Census 2021), Typology of Number of Rooms (T0–T6+), and useful floor area.
Number of Habitable Rooms on Occupied Dwellings (Cod: 0012853)	INE	Number	Total number of rooms intended for habitation in occupied dwellings within a given geographic area, reflecting housing capacity and living conditions of the resident population. Detailed level: Geography (Municipality) and Time (Census 2021).
Resident Population (Cod. 13267)	INE	Number	Total number of individuals usually residing in each geographic area at the reference date, serving as a key demographic indicator for territorial analysis and socio-economic planning. Detailed level: Geography (Municipality) and Time (Census 2021).
Housing Expenditure (GF06)	Eurostat	%GDP	General government expenditure on “Housing and community amenities” (COFOG), allowing EU-wide comparison.

### 3.5. Data Enrichment

In the Data Enrichment phase, the objective was to transform isolated administrative figures into meaningful socioeconomic indicators. This process was implemented in Python 3.13, including data cleaning, standardization, spatial–temporal integration, and feature engineering through the calculation of synthetic indices.

The initial step of standardization entailed the harmonization of disparate coding systems from the primary data sources. Administrative codes returned by the APIs were mapped to human-readable designations using metadata dictionaries, ensuring that the final output was intelligible for decision-makers. Regarding missing data, unlike standard big data approaches that often impute missing values, this study explicitly preserved null values. In the context of housing inequality, the absence of transactional data in rural municipality is not a system error but a significant finding, signaling a dormant market, a lack of liquidity, or deep territorial exclusion.

Following standardization, the integration of spatial and temporal dimensions was necessary to visualize the “two-speed” reality of the Portuguese housing market. Geo-

graphically, all micro-data from INE was already aligned with the current NUTS 2024 classification system (the previous NUTS classification system was valid since 2013). This hierarchical mapping enabled the analysis to seamlessly drill down from broad national trends to specific NUTS II regions and, where data quality permitted, to the municipal level. Temporally, the disparity between monthly bank valuations and annual construction statistics was resolved by aggregating data into consistent quarterly and annual averages. This synchronization created a time series spanning from 2015 to 2025, enabling the correlation of housing market behaviors with macro-economic disruptions, such as the COVID-19 pandemic or fluctuations in interest rates.

To move the analysis beyond simple descriptive statistics, feature engineering was applied to create three metrics to capture the structural dimensions of inequality, based on the source variables displayed in Table 1.

First, the “Average Number of Rooms per Person” metric [29] was computed to assess the availability of residential space at the individual level and to capture potential housing pressure and crowding conditions within the housing stock (Equation (1)). This metric, used by Eurostat and the OECD in the analysis of housing conditions, relates the total number of habitable rooms in occupied dwellings to the total number of residents. Higher values indicate a greater availability of living space per person and are generally associated with higher housing standards and lower levels of overcrowding. Conversely, lower values suggest more constrained living conditions, reflecting higher residential density and increased pressure on the housing stock, typically observed in high-demand urban or metropolitan areas. By normalizing dwelling size by household composition, this indicator allows for consistent comparisons across regions with different demographic structures and settlement patterns.

$$\begin{aligned} & \text{Average Number of Rooms per Person} \\ & = \frac{\text{Number of Habitable Rooms on Occupied Dwellings}}{\text{Resident Population}} \end{aligned} \quad (1)$$

Second, the “Dwellings per Thousand Inhabitants” indicator [30] was computed to assess the availability of housing stock relative to population size and to capture structural pressures in housing markets (Equation (2)). This indicator is widely used by the OECD in the analysis of housing supply conditions, housing adequacy, and long-term developments in residential construction. The indicator relates the total number of classic dwellings of habitual residence to the total resident population, expressing the result per 1000 inhabitants. By measuring the housing stock relative to population, it provides a standardized indicator of housing supply intensity that allows for meaningful comparisons across countries and regions with different population sizes and demographic dynamics. Higher values indicate a relatively larger housing stock compared to the size of the population and are generally associated with lower structural housing shortages and greater availability of dwellings. Conversely, lower values suggest tighter housing markets, potentially reflecting insufficient housing supply, higher demographic pressure, or rapidly growing populations. By normalizing the number of dwellings by the number of inhabitants, this indicator enables consistent cross-country and cross-regional comparisons, independently of absolute population size.

$$\begin{aligned} & \text{Dwellings per Thousand Inhabitants} \\ & = \frac{\text{Number of Classic Dwellings of Habitual Residence}}{\text{Resident Population}} * 1000 \end{aligned} \quad (2)$$

Finally, the “Normalized Housing Typology Diversity Index” [31], based on Shannon entropy, was computed using official housing data from the INE to assess the degree of heterogeneity in housing typologies across regions (Equation (3)). Specifically, using the indicator Number of Classic Dwellings of Habitual Residence (Cod: 0012503), dwellings

were then classified into six mutually exclusive typological categories (T0–T1, T2, T3, T4, T5, and T6+), constructed by combining information on useful floor area and number of rooms. This typological aggregation allows for a consistent representation of the structural composition of the housing stock while ensuring comparability across territorial units. The indicator captures the evenness of the distribution of housing units across typologies, rather than their absolute quantity. Higher values reflect a more diversified and balanced typological structure, indicating the coexistence of multiple housing typologies within the same region. Conversely, lower values signal a concentration of the housing stock in a limited number of typologies, suggesting a more specialized or homogeneous housing structure. To ensure comparability across regions and time, Shannon entropy is normalized by its theoretical maximum, in this case,  $\ln(6)$ , yielding a dimensionless index bounded between 0 and 1.

$$\text{Normalized Housing Typology Diversity Index} = \frac{\sum_{i=1}^6 p_i \ln(p_i)}{\ln(6)} \quad (3)$$

where

- $p_i$  denotes the proportion of housing units belonging to typology  $i$ , defined as

$$p_i = \frac{\text{Number of Dwellings of Typology } i}{\text{Total Number of Dwellings}} \quad (4)$$

- $i \in \{T0 - T1, T2, T3, T4, T5, T6+\}$

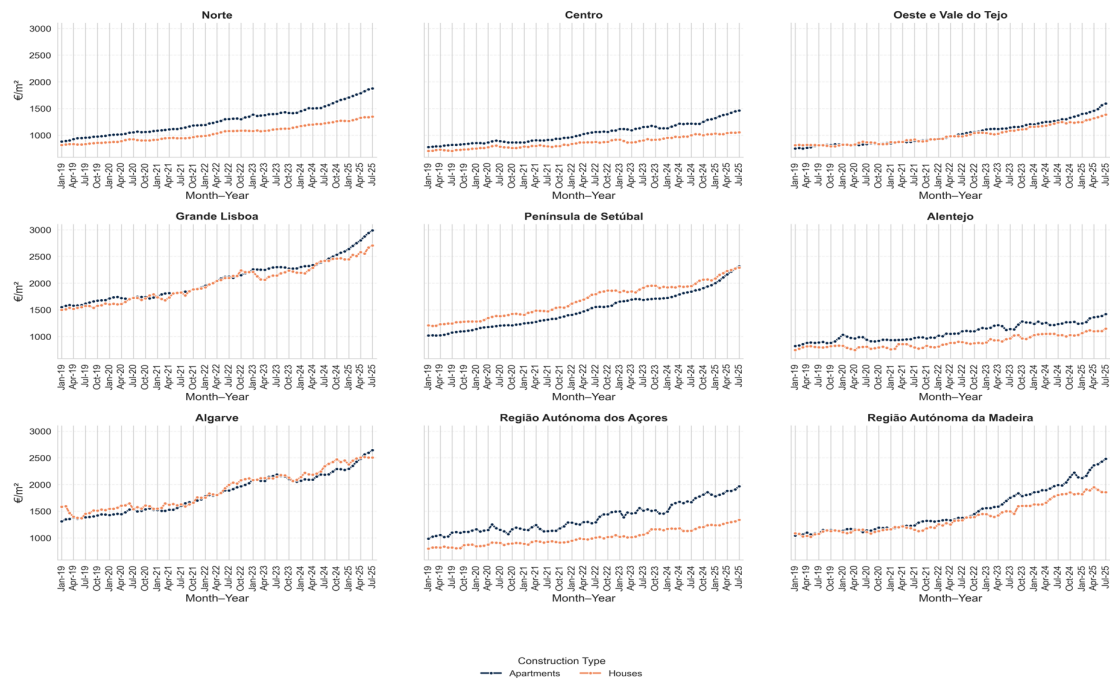
#### 4. Results and Analysis

This section presents the results of the analyses of housing inequalities in Portugal, organized into four dimensions: (1) evolution of housing market valuation; (2) spatial analysis at the regional level; (3) characterization of the housing market; and (4) dynamics of new supply and construction.

##### 4.1. Evolution of Housing Market Valuation

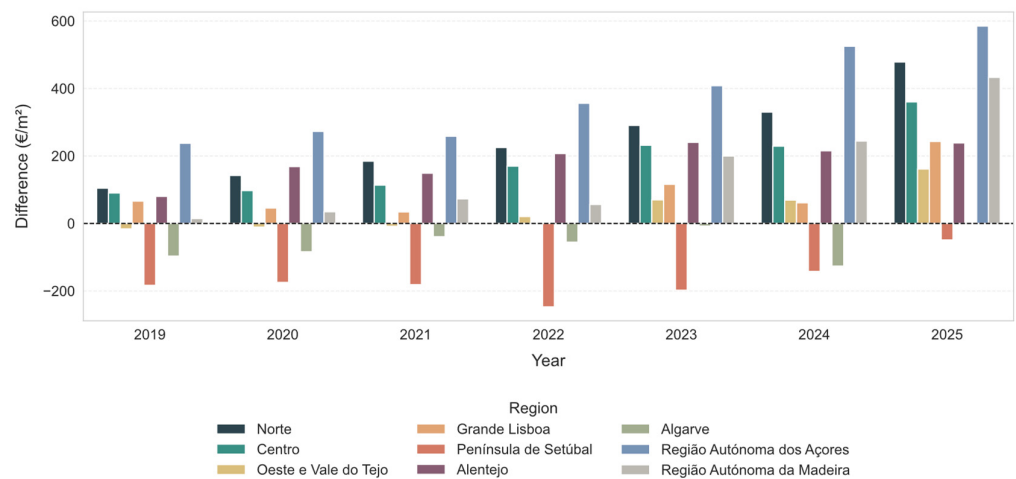
The “Median Bank Appraisal Value” (€/m<sup>2</sup>) is used as a proxy for monitoring housing market dynamics, reflecting both the financial sector’s risk perception and credit demand. This indicator is publicly available in official datasets, and although it does not represent the actual market values, it is considered a good estimate by region of the real housing market value. The analysis of the period between January 2019 and July 2025 reveals a structural transformation in the Portuguese housing market, characterized by a generalized increase in values and a deepening of inequalities between typologies and territories.

As illustrated in Figure 5 (Figure 5 uses the small multiples visualization technique that breaks a dataset into multiple small frames with consistent scales and axes to analyze general trends [32,33]; in this article, we use this technique to visualize trends by NUTS II regions and municipalities), there is an upward trend across all regions, with a notable acceleration starting in 2022. However, this appreciation is not homogeneous. Apartments consistently present higher valuations per square meter when compared to houses. This disparity suggests that market pressure is concentrated on typologies associated with apartments, which are often the target of investment and demand. In regions where these differences were previously less pronounced, particularly Lisbon (*Grande Lisboa*), the *Algarve*, and the *Oeste e Vale do Tejo* regions, there has been a clear increase in apartment prices relative to house prices since 2024. The *Península de Setúbal* is the only region displaying a different pattern, as houses market values have consistently been higher than those of apartments; however, since 2025, this region has begun to exhibit a reversal of this trend.



**Figure 5.** Monthly evolution of median bank appraisal value (2019–2025) by region: NUTS II.

The widening of this gap is quantified in Figure 6, which displays the price differential between apartments and houses. In almost all regions, living in an apartment has become progressively more expensive relative to houses, reinforcing barriers for young families and middle-income households seeking residence in consolidated urban areas. These trends suggest that recent housing policies have not been sufficient to mitigate pressure on the most sought-after market segments.



**Figure 6.** Price gap (€/m<sup>2</sup>) between apartments and houses by region (NUTS II) and year.

Beyond the typology gap, territorial inequality has intensified significantly. Figure 7 illustrates the continuous increasing in the differential between the most expensive and the cheapest regions in the country. In 2019, this territorial gap was approximately 800–900 €/m<sup>2</sup>; by 2025, it approached 1500 €/m<sup>2</sup>. This trend indicates that the market did not converge; instead, regions that were already expensive (such as *Grande Lisboa* and *Algarve*) experienced much faster price growth than the rest of the country, intensifying socio-territorial inequalities.

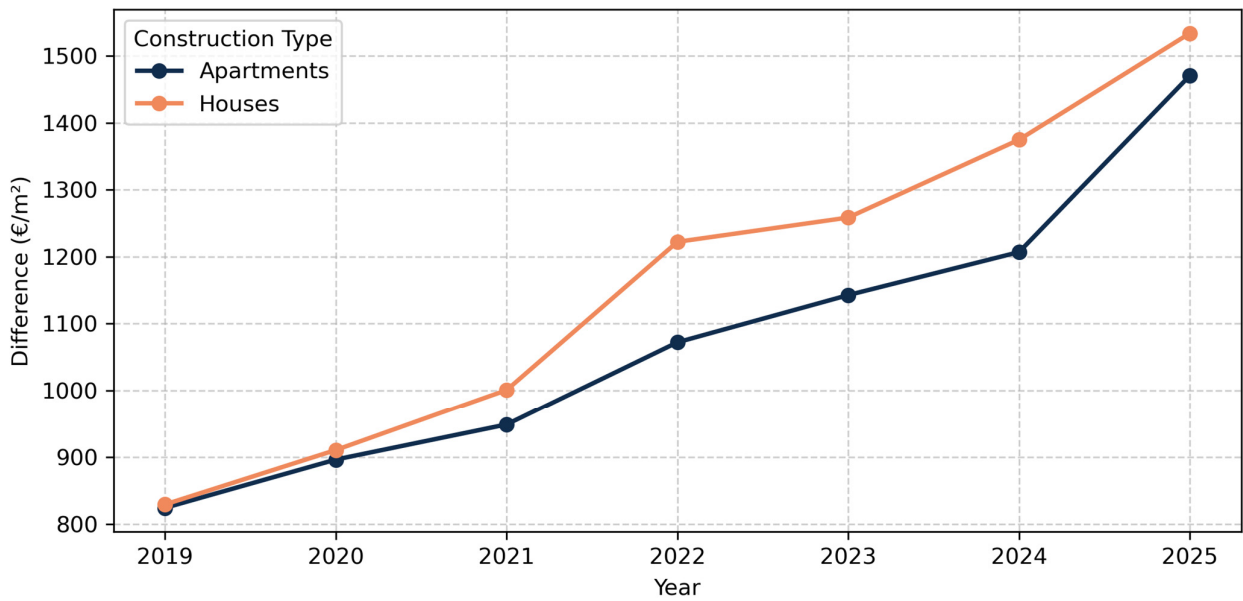


Figure 7. Evolution of the gap between the most expensive and cheapest regions (€/m<sup>2</sup>).

To visualize the magnitude of this territorial gap in (€/m<sup>2</sup>), the heatmap in Figure 8 details the absolute values. The regions of *Grande Lisboa*, the *Algarve*, and the *Península de Setúbal* stand out as the territories with the highest absolute variations. Notably, the *Península de Setúbal* recorded growth exceeding 100% in six years, suggesting the indirect effects spreading from Lisbon to the southern bank of the Tagus River, impacting the regional housing affordability.

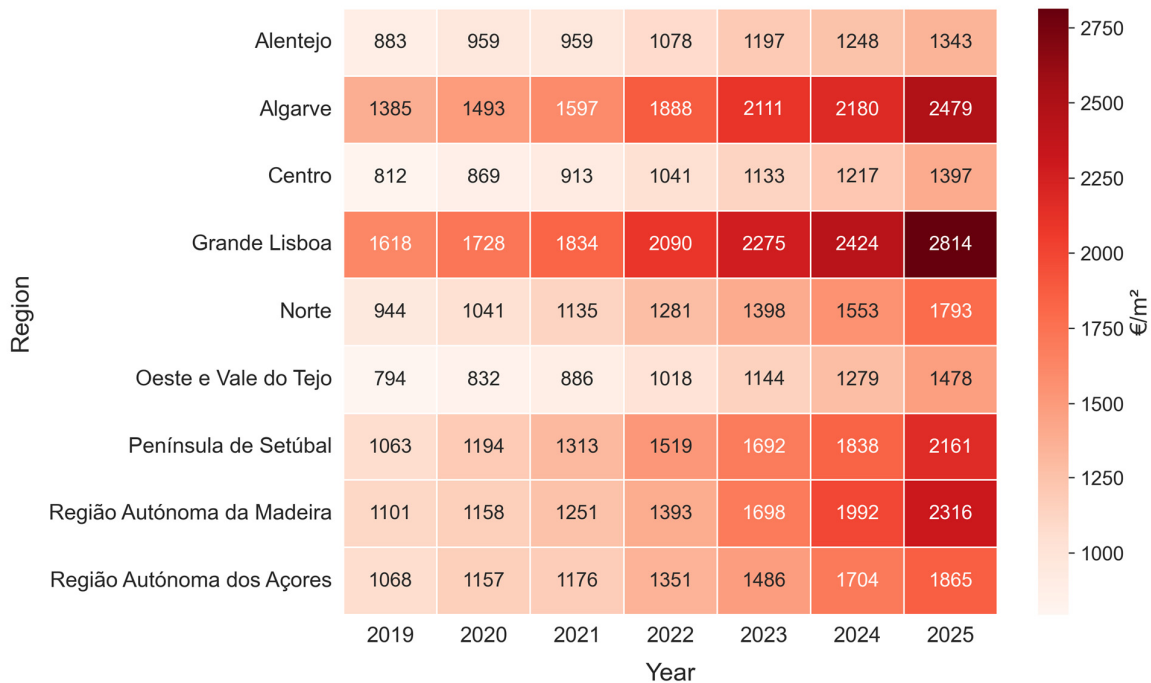
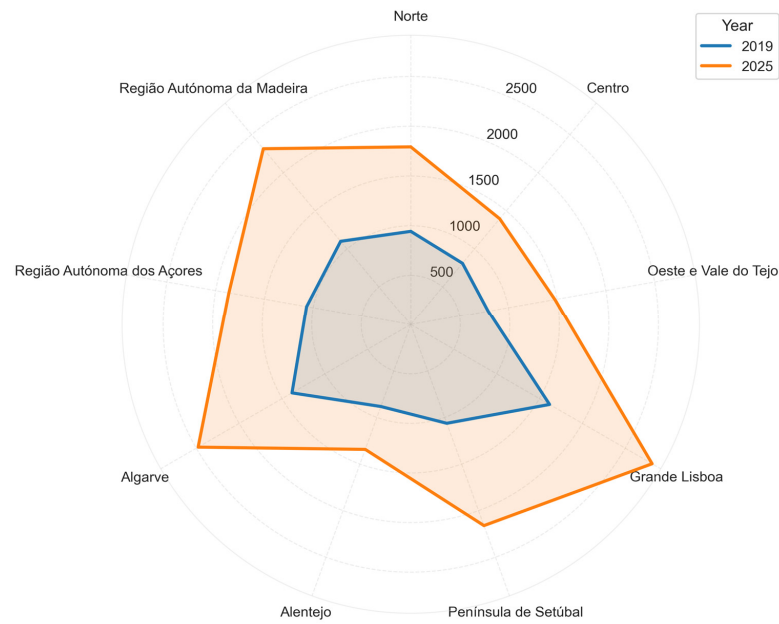


Figure 8. Heatmap of median value (€/m<sup>2</sup>) for apartments by region: NUTS II.

This transformation is succinctly captured in the radar chart in Figure 9, which compares the regional profiles of 2019 versus 2025. The chart highlights with a bubble mark the four specific regions that have detached from the national average, creating a polarized market structure in 2025 (namely, *Grande Lisboa*, *Algarve*, *Região Autónoma da Madeira*, and *Península de Setúbal*).

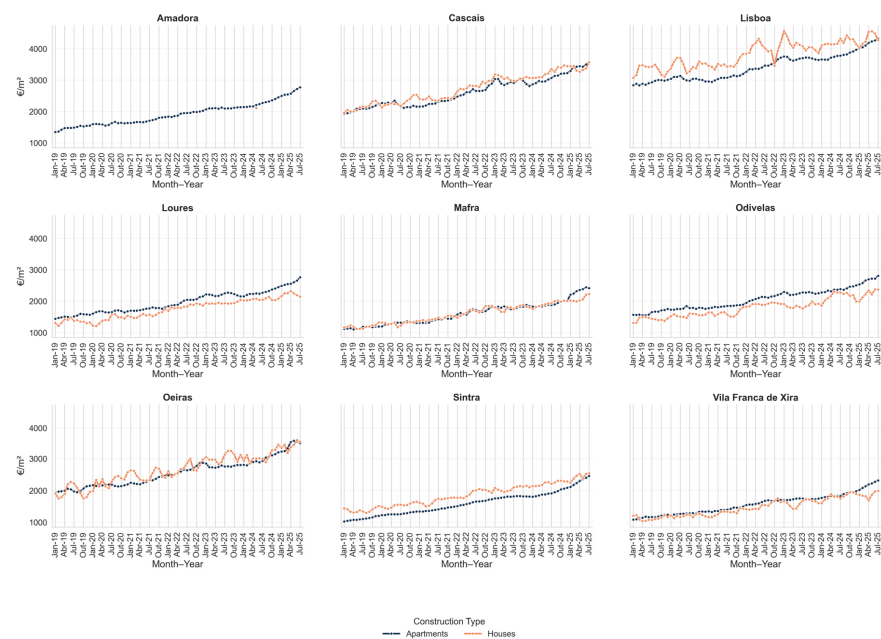


**Figure 9.** Regional comparison of apartment valuations (€/m<sup>2</sup>): 2019 vs. 2025.

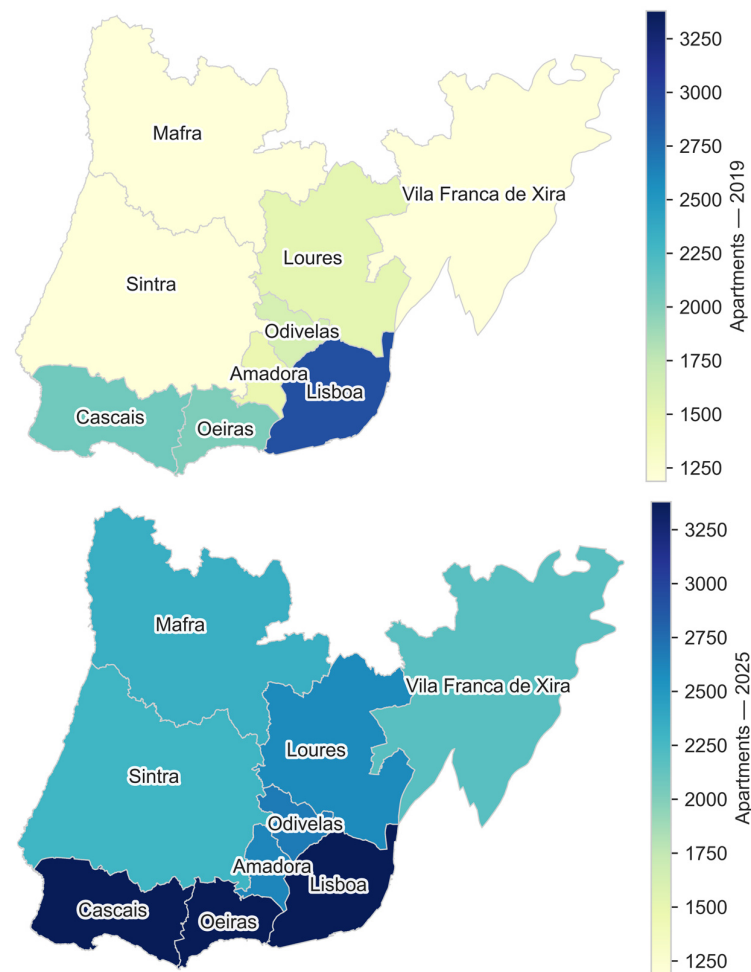
**4.2. Spatial Analysis at Regional Level**

At the national level, housing valuations increased, with pronounced market polarization in the regions of *Grande Lisboa*, *Península de Setúbal*, *Algarve*, and the *Região Autónoma da Madeira*. The following sections will focus on a municipal-level analysis of *Grande Lisboa* and *Península de Setúbal*, as these regions provide the most comprehensive data at this level of spatial detail (from the official source).

In *Grande Lisboa*, the ranking of municipalities in 2025 (Figure 10) confirms a consolidated hierarchy. Lisbon (at the municipality level, Lisbon is confined to the city of Lisbon, while the region of Grande Lisbon includes nine municipalities (as detailed in Figure 11)), *Cascais*, and *Oeiras* remain detached at the top, with values significantly higher than the regional average. However, all municipalities show a steady and continuous increase in housing valuation. Additionally, the dataset lacks information on houses in *Amadora*.



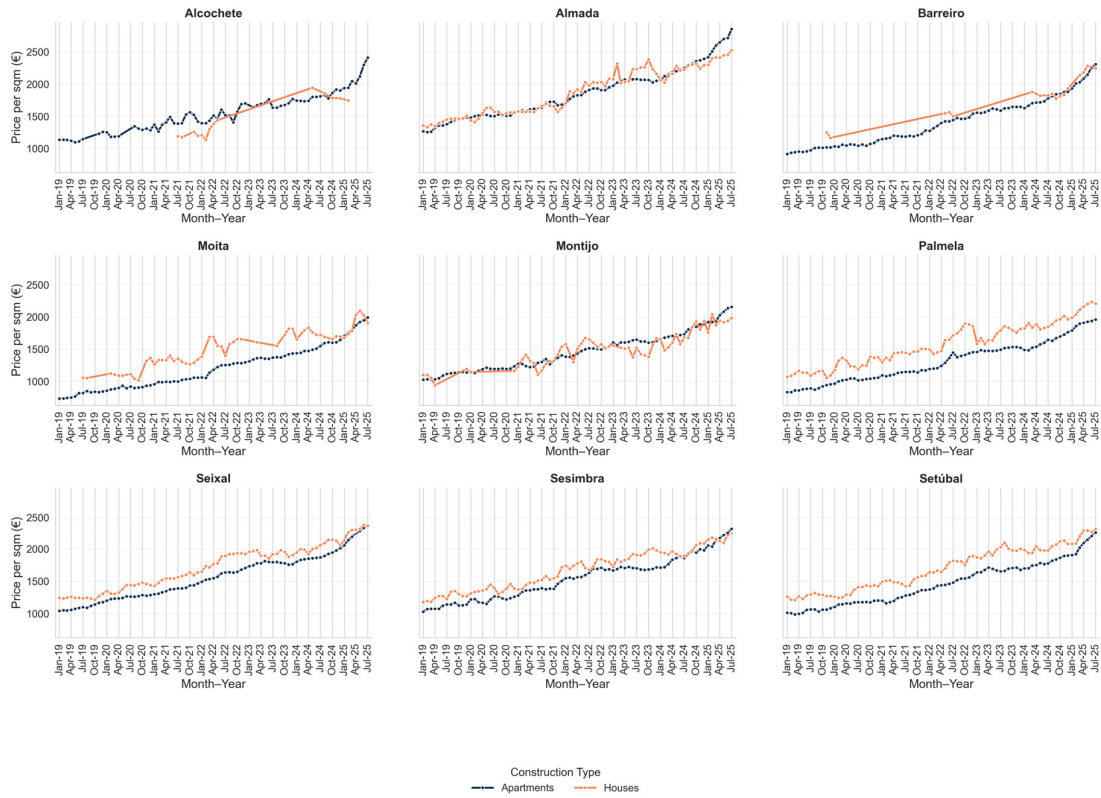
**Figure 10.** Evolution of median bank appraisal value (2019–2025) by municipalities in Grande Lisboa (€/m<sup>2</sup>).



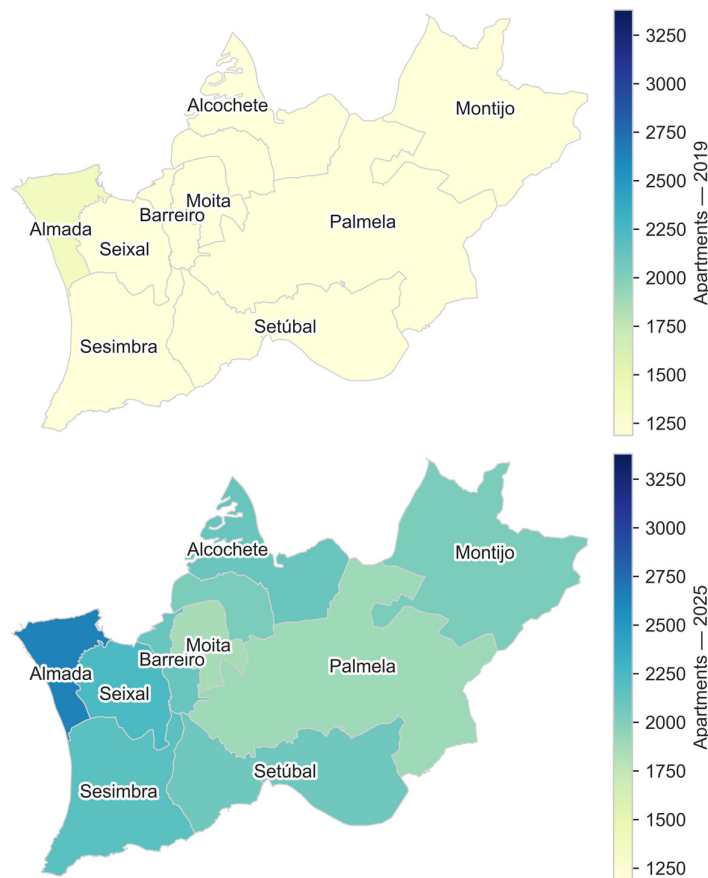
**Figure 11.** Spatial distribution of median bank appraisal value in Grande Lisboa: 2019 (**top**) vs. 2025 (**bottom**).

A comparison between the 2019 and 2025 snapshots, displayed in the heatmaps of Figure 11, shows that valuation averages in all municipalities are converging to the valuation of the Lisbon municipality (City of Lisbon). This expansion forces a trade-off for families between price, commuting time, and location quality, pushing demand further to the periphery. Even the peripheral regions such as *Vila Franca de Xira*, *Sintra*, and *Mafra* have reached a significant increase in valuation levels since 2019.

The *Península de Setúbal* exhibits a rapid increase in housing valuations across all municipalities, as illustrated in Figure 12. *Almada* emerges as a particularly salient case, with valuation levels approaching those observed in *Grande Lisboa*. A comparison between 2019 and 2025 (Figure 13) reveals dynamics similar to the indirect effects previously identified in Lisbon and other municipalities within *Grande Lisboa*. Specifically, housing valuations in the *Península de Setúbal* increased by more than 100% over the six-year period, indicating a diffusion of price pressures from Lisbon to the southern bank of the Tagus River, with significant implications for regional housing affordability. Moreover, municipalities such as *Seixal* and *Barreiro* have substantially narrowed their valuation gap relative to *Almada*. This trajectory suggests that the southern bank of the Tagus River has transitioned from a predominantly low-cost, commuter-oriented area to an integral component of the high-valuation metropolitan core, thereby introducing new affordability challenges at the local level.



**Figure 12.** Evolution of median bank appraisal value (2019–2025) by municipalities in Península de Setúbal (€/m<sup>2</sup>).



**Figure 13.** Spatial distribution of median bank appraisal value in the Península de Setúbal: 2019 (top) vs. 2025 (bottom).

### 4.3. Characterization of the Housing Stock

Data on the characterization of the housing stock are drawn from the most recent population and housing census (2021). At the national level, as shown in Figure 14, the Portuguese housing stock is predominantly composed of dwellings with four to five divisions (corresponding to T3 and T4 typologies) and six divisions (T5). In the Portuguese housing market, the “T” classification refers to the number of bedrooms. Figure 15 displays the same information using aggregated T typologies.

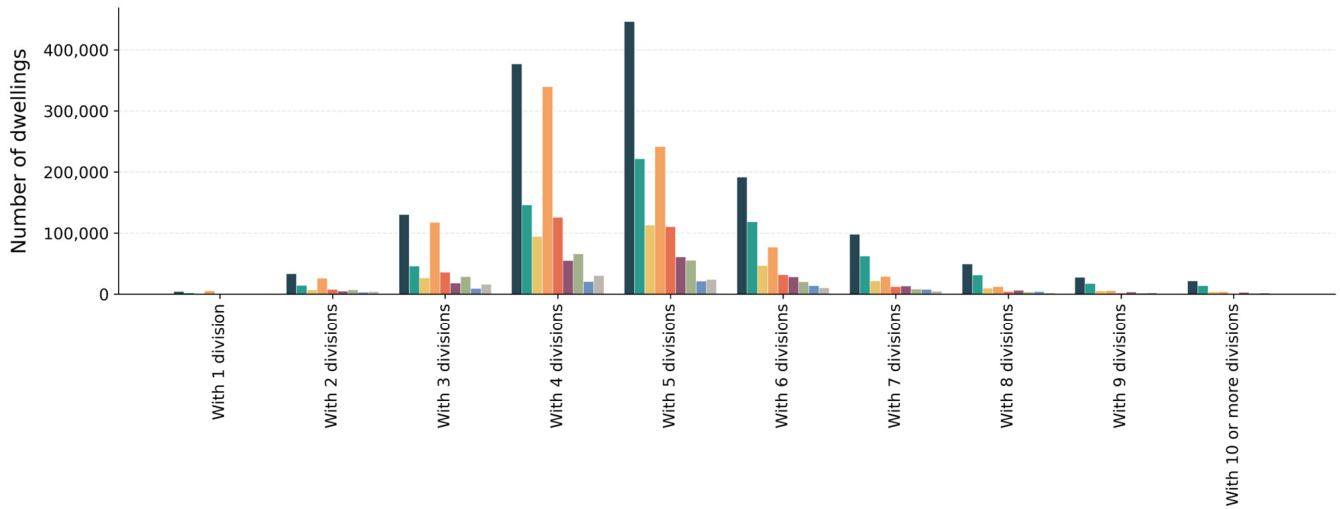


Figure 14. Distribution of dwellings by number of divisions and region (NUTS II) (2021).

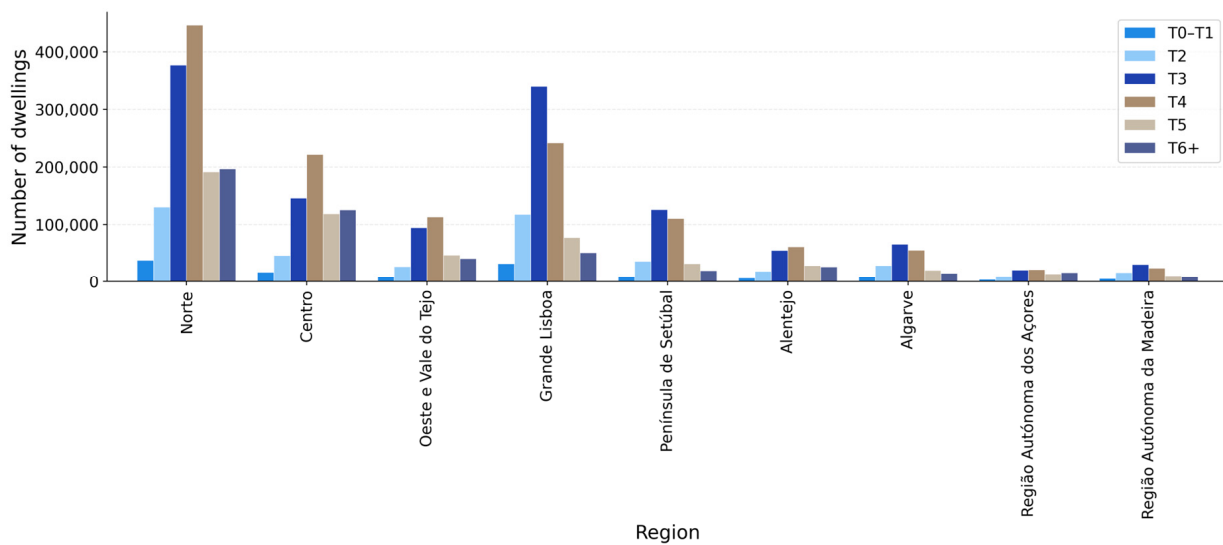


Figure 15. Distribution of dwellings by housing typology and region (NUTS II) (2021).

Figures 15 and 16 show clear differences in housing typologies across regions in Portugal. The *Norte*, *Centro*, and *Oeste e Vale do Tejo* regions exhibit a housing stock dominated by T4 typologies. *Grande Lisboa*, *Península de Setúbal*, *Algarve*, and *Madeira* display a distribution curve characteristic of high-density urban environments, where T3 typologies prevail, and the supply of large properties (T5–T6+) contracts sharply. As previously discussed, these are the most expensive regions (see Figure 9). Notably, despite these regional asymmetries, T0–T1 units remain marginally represented across the entire territory, signaling a generalized undersupply of compact housing options suitable for smaller households.

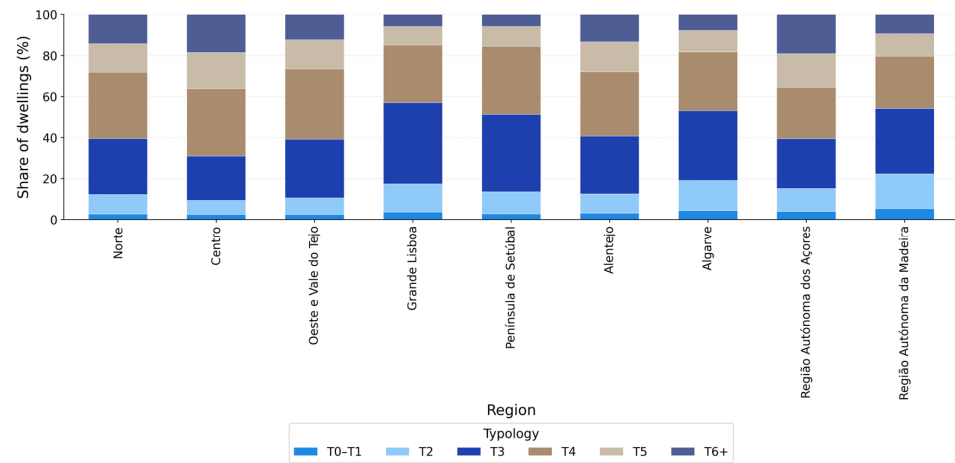


Figure 16. Housing typology share by region: NUTS II (2021).

The Normalized Housing Typology Diversity Index indicates consistently high levels of typological diversity across Portuguese regions in 2021, with values clustered between approximately 0.82 and 0.90. This pattern suggests a broadly balanced distribution of housing typologies nationwide and the absence of pronounced regional specialization. Slightly higher diversity levels are observed in the autonomous regions and in *Alentejo*, pointing to more heterogeneous housing structures, whereas metropolitan areas such as *Grande Lisboa* and the *Península de Setúbal* display marginally lower values, reflecting a more standardized typological composition. Despite these differences, the overall dispersion of the index remains limited, underscoring the strong structural similarity of regional housing stocks. These regional patterns are illustrated in Figure 17.

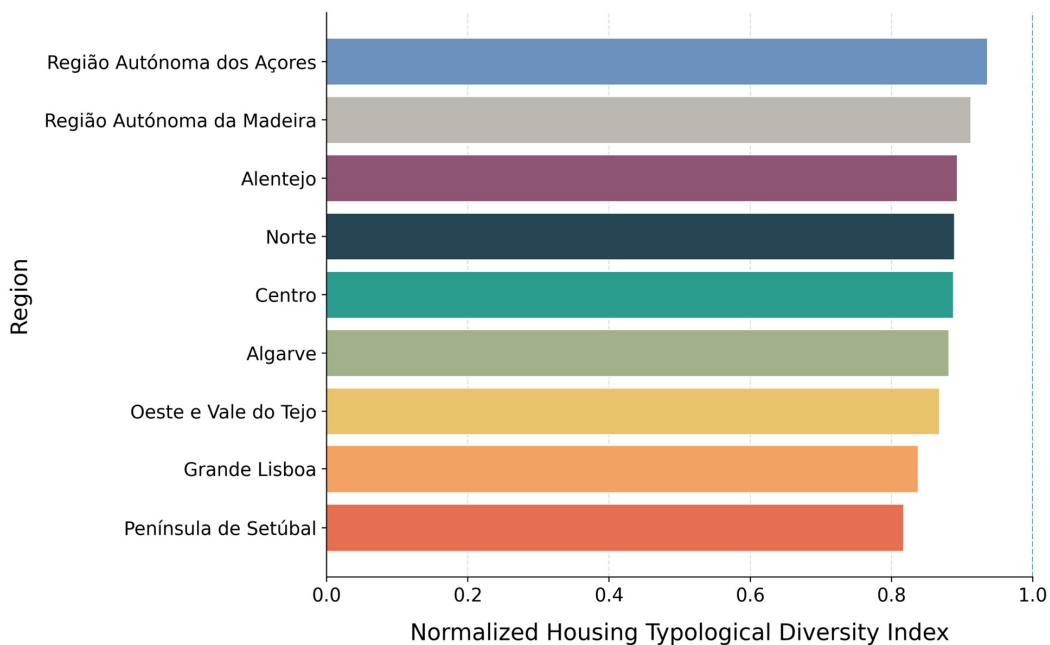


Figure 17. Normalized Housing Typology Diversity Index by region: NUTS II (2021).

Figure 18 displays the distribution of useful floor area, with a predominance of dwellings between 80 and 119 m<sup>2</sup>. Although dwellings with four to six divisions (T3–T5), as previously discussed, dominate the housing stock, their useful floor area is comparatively small for typologies of this size, suggesting a mismatch between dwelling classification and spatial adequacy.

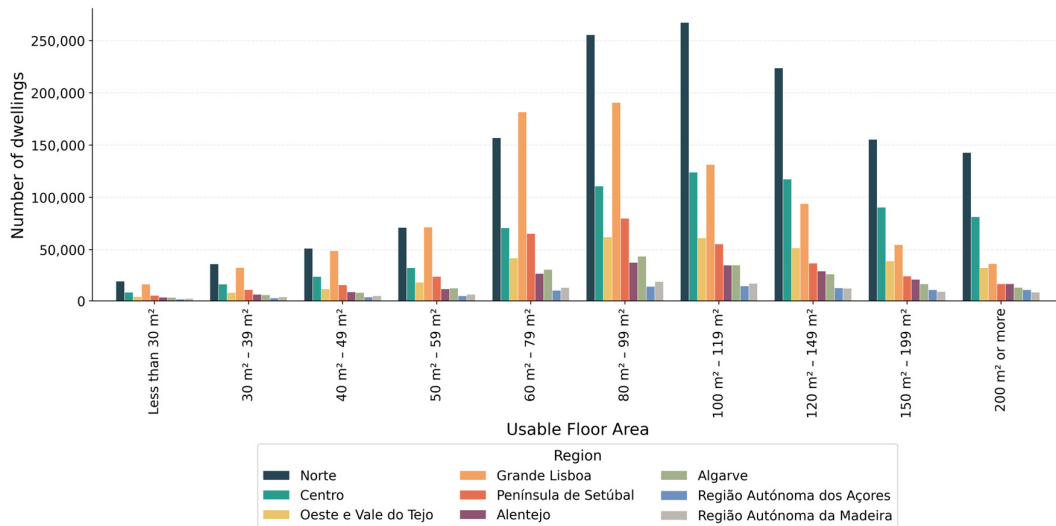


Figure 18. Distribution of dwellings by useful floor area by region: NUTS II (2021).

The normalized average number of rooms per person reveals clear regional disparities in residential space availability across Portugal, as shown in Figure 19. Regions such as *Centro* and *Alentejo* display ratios above the national average, indicating a greater availability of housing space per resident, which is consistent with lower population density and more dispersed settlement patterns. In contrast, metropolitan areas, including *Grande Lisboa* and the *Península de Setúbal*, record systematically lower ratios, reflecting higher residential density and increased pressure on the housing stock. The *Norte* region remains close to the national benchmark, suggesting a relatively balanced housing structure. The autonomous regions, particularly *Madeira*, present the lowest normalized values, pointing to structural limitations in dwelling size and housing composition. Overall, the ratio highlights persistent territorial inequalities closely linked to urbanization and demographic concentration.

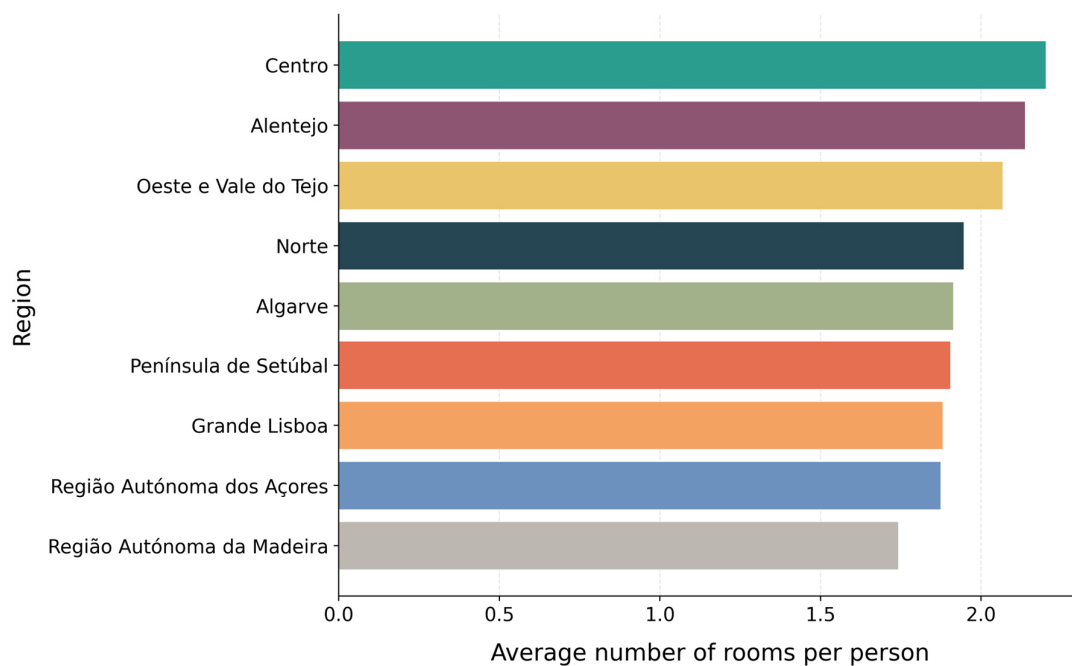
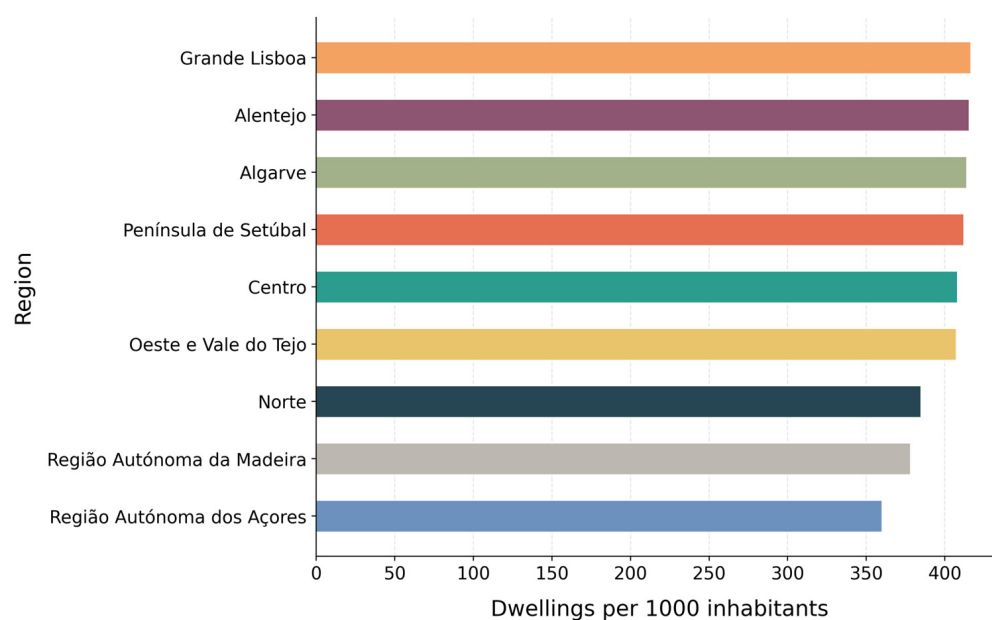


Figure 19. Average number of rooms per person by region: NUTS II (2021).

Complementing the previous evidence on internal dwelling space (Figure 19), Figure 20 reports dwellings per capita, capturing the intensity of the housing stock relative to the resident population. The indicator displays a comparatively narrower regional spread than the rooms-per-person ratio, suggesting that differences in housing space availability are more pronounced than differences in the number of housing units per inhabitant. Still, a clear ranking emerges: *Alentejo* and *Algarve* appear among the highest values, consistent with less compact settlement structures and a relatively larger housing stock per resident. Notably, *Grande Lisboa* also records a high dwellings-per-capita level, which should be interpreted cautiously, as this metric reflects total stock and may incorporate vacant or secondary dwellings rather than effective access or affordability. At the lower end, the autonomous regions—especially the *Região Autónoma dos Açores*—exhibit reduced values, pointing to tighter housing stock relative to population. Overall, Figure 20 suggests that regional disparities in housing conditions are not solely driven by the quantity of dwellings but also by differences in dwelling size and occupancy patterns, reinforcing the need to interpret housing availability through multiple, complementary ratios.



**Figure 20.** Dwellings per thousand inhabitants by region: NUTS II (2021).

This section showed that the housing typology in Portugal has asymmetries by region. On one hand, less populated regions, such as *Centro* and *Alentejo*, have a higher number of rooms per person. On the other hand, the number of dwellings per thousand of inhabitants is higher mainly in *Grande Lisboa*, *Alentejo*, and *Algarve*. This indicator is affected by houses occupied for tourism activities, particularly in the regions of *Grande Lisboa* and *Algarve*, where the tourism pressure is higher.

#### 4.4. Dynamics of New Supply and Construction Patterns

The evolution of completed buildings and dwellings provides insight into the responsiveness of the supply side. Between 2015 and 2024, there was a consistent expansion in the number of completed buildings. However, as Figure 21 demonstrates, this growth relies on new construction, with rehabilitation (including the alteration, extension, and reconstruction types) playing a residual role. This pattern suggests a model of growth based on urban expansion rather than the renewal of the existing housing stock.

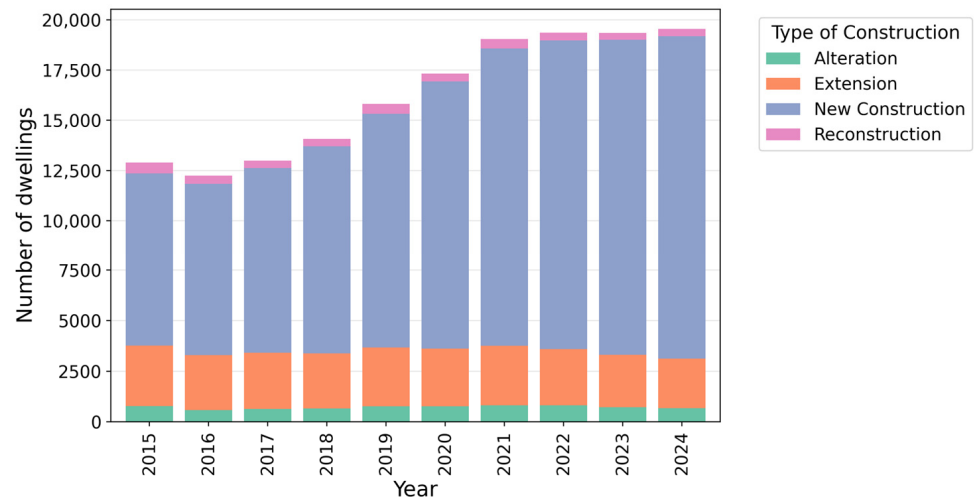


Figure 21. Distribution of building works by construction type (2015–2024).

Analyzing the distribution of the number of completed dwellings (including new constructions and rehabilitations) in absolute values by region accumulated from 2019 and 2024, the *Norte* region dominates. The Pareto chart, also in Figure 22, shows that the regions of *Norte*, *Centro*, and *Grande Lisboa* account for approximately 70% of all completed dwellings. When adding the *Península de Setúbal* and the *Oeste e Vale do Tejo* regions, this figure exceeds 80%.

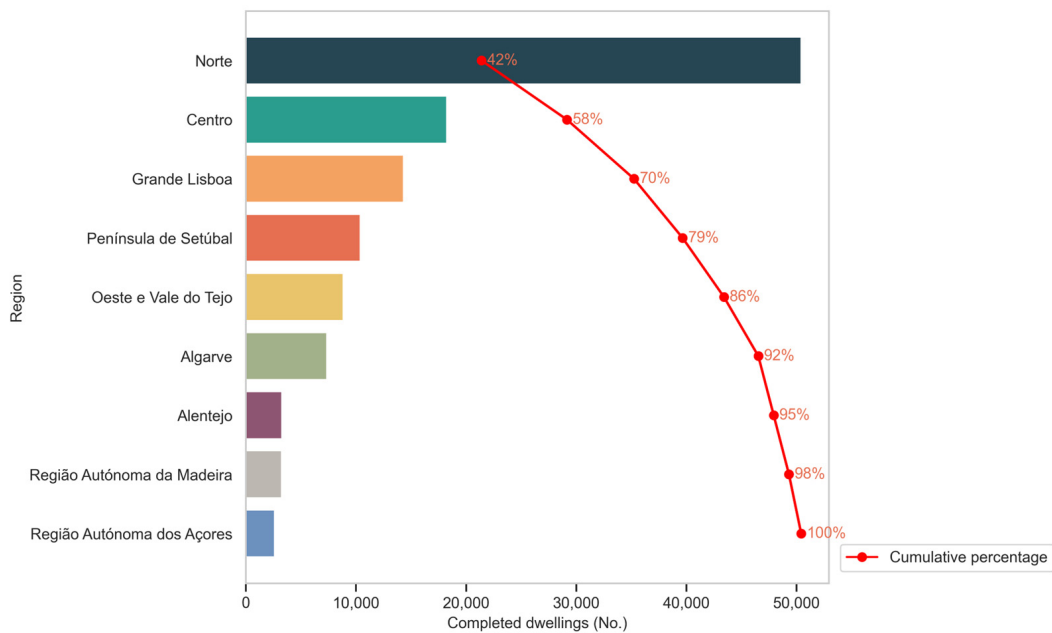
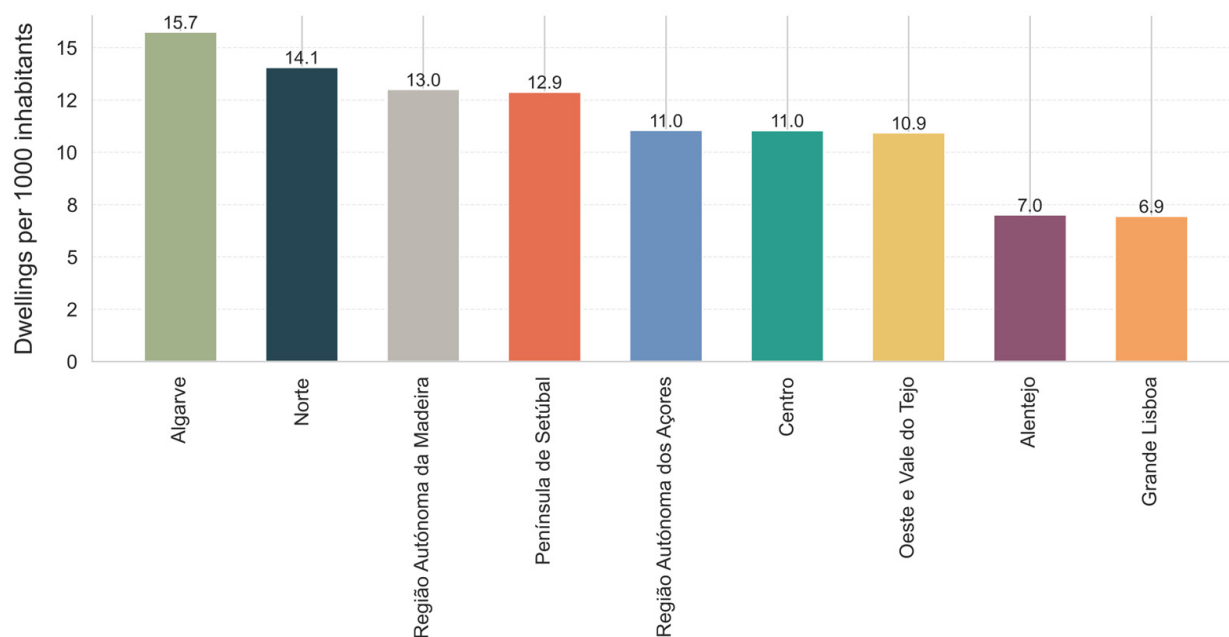


Figure 22. Pareto analysis of completed dwellings by region: NUTS II (2019–2024).

In continuation of the Pareto analysis (Figure 22), the ratio of the number of completed dwellings/resident population (Figure 23) provides a relative measure of housing supply adjusted for demographic size. While *Norte*, *Centro*, and *Grande Lisboa* dominate in absolute numbers, the ratio reveals contrasting regional dynamics. Regions such as *Algarve* and *Madeira* show higher values considering the lower resident population, largely driven by tourism-related and non-permanent housing demand. Conversely, *Grande Lisboa* records the lowest ratio, indicating that housing supply has not kept pace with population pressures.



**Figure 23.** Number of completed dwellings per 1000 inhabitants by region: NUTS II (2019–2024).

The normalized analysis presented in Figure 23 is essential given the highly heterogeneous spatial distribution of the Portuguese population. Population density is relatively low in the inland regions and highly concentrated in areas such as the *Norte* and *Grande Lisboa*. Three contrasting situations emerge from this complementary assessment. In absolute terms, the *Norte* appears to exhibit a substantially higher level of construction activity compared to other regions (Figure 22). However, when the data are normalized by the resident population (Figure 23), it becomes clear that it is not the leading region, and the differences relative to the remaining regions are not particularly significant. *Grande Lisboa* also seems to display a high volume of construction in aggregate terms. Yet, once adjusted for the number of resident inhabitants, it becomes the region with the lowest level of construction per capita. Finally, the *Algarve* presents an asymmetric pattern, largely explained by its relatively small resident population. Although the aggregate figures suggest a low level of new construction, the population-normalized analysis reveals that it stands out as the region with the highest level of new construction relative to its number of residents.

The year-over-year (YoY) growth rates of completed dwellings between 2019 and 2024 reveal substantial regional heterogeneity and volatility (Figure 24). Most regions experienced expansions followed by significant contractions, particularly around the 2021–2022 period. Despite continued increases in completed dwellings across all regions, YoY growth rates have decelerated, particularly from 2021 onwards (the exception is *Oeste e Vale do Tejo*). This behavior may be linked with the COVID-19 pandemic. In contrast, regions such as *Algarve* and the autonomous regions display more cyclical patterns, consistent with investment- and tourism-driven construction dynamics. The widespread deceleration observed in 2022 highlights a generalized slowdown in housing supply growth. Overall, these patterns indicate an unstable and uneven supply response, reinforcing persistent regional disparities in housing availability.

The heatmap in Figure 25 confirms the persistence of this pattern over time. The *Norte* region maintains a high intensity of construction, while *Grande Lisboa* and *Centro* show gradual growth, particularly post-2021.

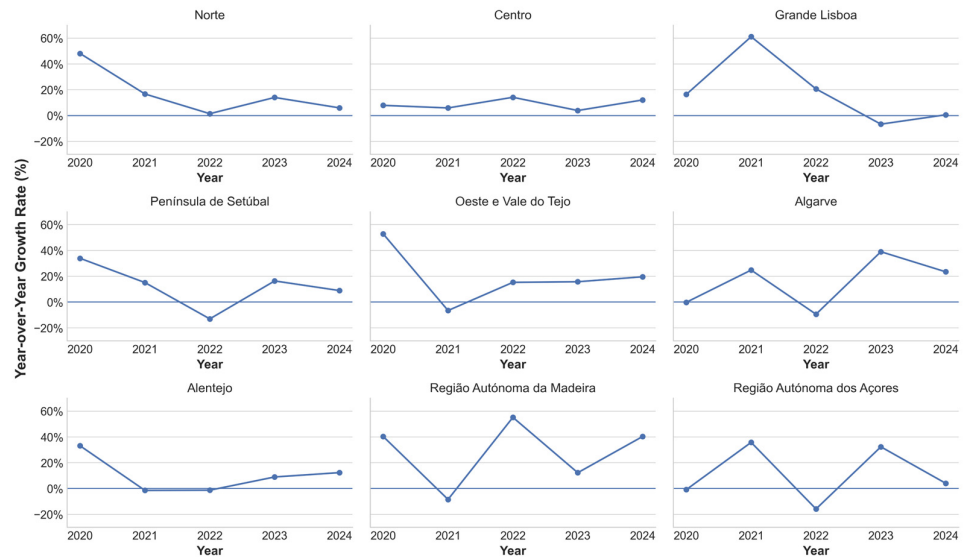


Figure 24. Year-over-year growth rate of completed dwellings by region: NUTS II (2019–2024).



Figure 25. Heatmap of completed dwellings by year and region.

Zooming into *Grande Lisboa* (Figure 26), the temporal evolution of construction between 2019 and 2024 shows a shift from the center to the periphery. Municipalities like *Mafra*, *Sintra*, and *Cascais* have taken the lead in completed dwellings in 2024, surpassing the municipality of Lisbon. This decentralization reflects the scarcity of space and regulatory/cost constraints in the capital.

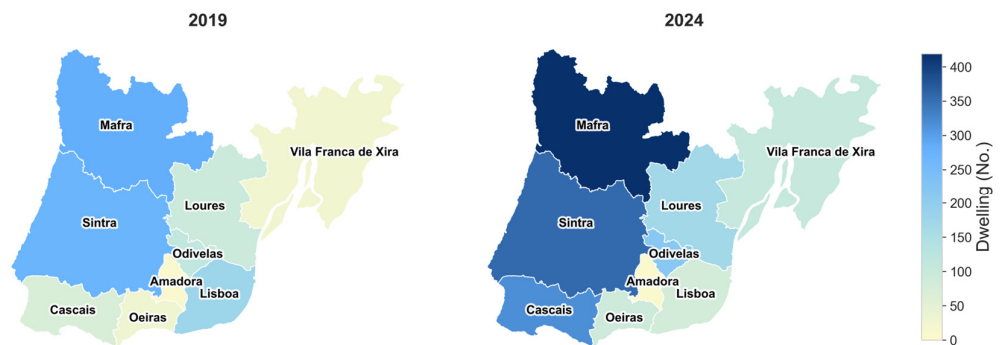
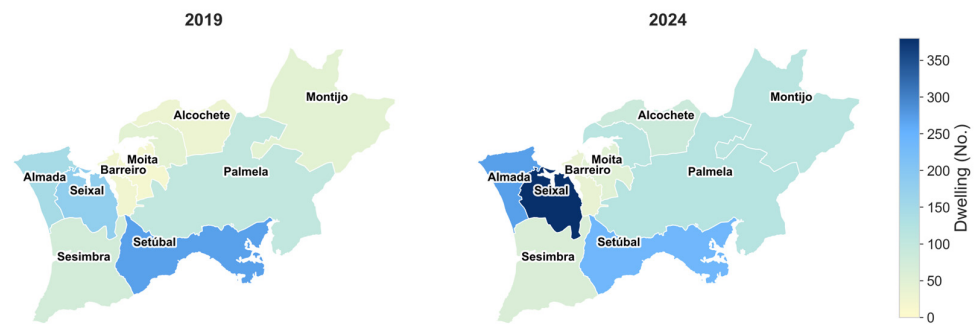


Figure 26. Spatial and temporal evolution of completed dwellings in Greater Lisbon municipalities.

Finally, Figure 27 illustrates the spatial evolution of construction intensity in the *Península de Setúbal*. A comparison between 2019 and 2024 reveals a pronounced increase in construction activity in the regions closer to the two bridges connecting Lisbon to the south bank of the Tagus River: on one hand, *Almada* and *Seixal* (served by *Ponte 25 de Abril*) and, on the other, *Alcochete*, *Montijo*, and *Moita* (served by *Ponte Vasco da Gama*). This reflects the indirect effects from Lisbon and confirms the consolidation of the *Península de Setúbal* as a strategic expansion front of the metropolitan housing market, with direct implications for regional affordability and land-use planning.



**Figure 27.** Spatial and temporal evolution of completed dwellings in *Península de Setúbal* municipalities.

Taken together, the results point to three central findings: (i) persistent territorial polarization in housing valuations; (ii) an uneven and decelerating supply response after 2021–2022; and (iii) the need for multi-scalar monitoring tools capable of relating market dynamics to spatial and social inequalities.

## 5. Discussion

The findings of this study can be considered in the context of a broader European and international discussion on housing inequality, affordability pressures, and uneven housing-market adjustment. Even though the analysis is focused on Portugal, the patterns observed are not particularly unusual. Similar developments have been identified in other settings, especially the widening gap between housing costs and household purchasing power, the concentration of pressure in metropolitan and other high-demand areas, and the limited capacity of housing supply to respond at the same speed as demand [34–36]. Recent comparative research has shown that affordability has worsened across European countries and urban regions, as prices, rents, and related housing costs have increased faster than household incomes, often with significant territorial differences within the same national context [34,35]. Evidence from the Lisbon area points in the same direction, suggesting that the Portuguese case should be understood as part of a wider pattern of spatially uneven housing-market adjustment rather than as a singular national trajectory [36].

Portugal does not appear to be an exception so much as a clear example of broader structural pressures currently affecting housing markets. The valuation growth identified in *Grande Lisboa*, *Península de Setúbal*, *Algarve*, and other high-pressure territories is consistent with what has been described in the literature as housing-market polarization, where metropolitan concentration, tourism-related demand, and uneven investment dynamics tend to reinforce differences in access to housing [36,37]. More recent work has also argued that housing inequality is increasingly shaped by the combined effect of territorial concentration, market pressure, and the growing role of housing as an asset, rather than only as a basic social good [1,37]. In that sense, the Portuguese case helps to illustrate wider dynamics that can also be found elsewhere in Europe [36].

The results also suggest that housing inequality cannot be understood through prices alone. When valuation dynamics are read together with housing stock characteristics, completed dwellings, and the territorial differences identified in the analysis, what emerges is a more complex process in which economic, spatial, and demographic pressures overlap and reinforce one another. This reading is in line with recent work showing that housing inequalities are produced through several dimensions, including affordability constraints, territorial disparities, and unequal access to adequate housing conditions, rather than through a single market indicator [38–40].

A further contribution of this study lies in the way the analysis is operationalized. A large part of the literature on housing inequality is either conceptual or focused on specific empirical outcomes, but fewer studies focus on the construction of reproducible monitoring workflows based on official statistics and scalable analytical procedures. The workflow followed in this manuscript adds not only empirical evidence on Portugal but also a transferable way of integrating public data and communicating territorial patterns through visual analytics [39,40].

In terms of implications for policy makers, the results reinforce the need to move beyond national aggregates. The concentration of housing pressure in specific territories, together with the uneven response of new construction, suggests that housing policy needs to be more sensitive to place and scale. Metropolitan areas and territories shaped by strong tourism demand are unlikely to respond to the same instruments in the same way as lower-pressure regions. The Portuguese case reflects a wider governance problem that is not unique to one country: how to turn fragmented official data into usable and comparable intelligence for policy design, monitoring, and evaluation [39,40].

The results are broadly consistent with what has been discussed in the international literature. First, housing pressures are becoming increasingly territorialized, rather than evenly distributed across space. Second, supply does not always adjust at the same speed as demand and price escalation. Third, the policy value of housing analysis depends heavily on the availability of harmonized and interpretable monitoring tools that can support comparison over time and across territories [34,39,40].

## 6. Conclusions

This paper contributes to the study of housing inequalities by coupling empirical evidence with a reproducible, data science-oriented analytical architecture designed to monitor territorial dynamics. Using Portugal as a case study, the analysis shows how Business Intelligence (BI) and urban analytics methodologies can be used to transform fragmented official statistics into policy-relevant insights.

From an empirical perspective, the results reveal a structurally polarized housing market between 2015 and 2025. Housing valuations increased across all regions, yet at highly heterogeneous rates, reinforcing territorial inequalities rather than promoting convergence. Metropolitan and tourism-intensive regions—namely *Grande Lisboa*, *Península de Setúbal*, *Algarve*, and *Região Autónoma da Madeira*—experienced accelerated appreciation, while spatial spillover effects propagated housing pressure to adjacent territories. On the supply side, although completed dwellings continued to increase nationwide, year-over-year growth rates decelerated after 2021–2022, suggesting a weaker supply response to sustained demand pressures and growing demographic concentration.

From a methodological standpoint, the main contribution of this study is the design and implementation of a workflow, encompassing automated data ingestion from official statistical APIs, schema harmonization, spatial–temporal alignment, feature engineering, and visual analytics. This workflow follows a CRISP-DM-inspired process adapted to

official public-sector data. This pipeline ensures adaptability and replicability, allowing the proposed framework to be used in other contexts and data sources.

One important methodological choice concerns data governance and the treatment of missing values. Rather than applying conventional imputation techniques, the framework explicitly preserves null observations when they reflect territorial realities, such as market inactivity or structural exclusion in low-liquidity regions. This decision avoids the artificial smoothing of spatial inequalities and aligns with best practices in small-area and territorial data analysis.

The feature engineering stage transforms raw administrative data into indicators that capture housing space availability, housing stock intensity, and typological diversity. These indicators enable multidimensional assessments of housing inequality that extend beyond price-based measures, capturing structural and spatial dimensions of housing adequacy. Their normalization allows consistent comparison across regions with heterogeneous population sizes, settlement patterns, and housing structures, reinforcing the robustness of cross-territorial analysis.

Visual analytics play an important role within this framework. Heatmaps, Pareto distributions, radar plots, and temporal growth visualizations are employed as exploratory and diagnostic instruments, supporting pattern recognition, anomaly detection, and the identification of non-linear dynamics in housing markets. This visual storytelling approach improves interpretability for decision-makers while preserving analytical depth, bridging the gap between advanced data processing and evidence-based urban governance.

Despite the contributions of this study, limitations should also be acknowledged. The reliance on aggregated data and the lack of granular household-level microdata prevented a direct causal analysis between socioeconomic characteristics and housing outcomes. Moreover, the limited availability and consistency of data at the municipal level restricted a more detailed spatial analysis, constraining the assessment of intra-regional disparities. Additional data gaps in specific NUTS II regions—most notably in the *Algarve*—further hindered a fully homogeneous territorial reading. These constraints constituted a major limitation of this study and ultimately led to a focus exclusively on the NUTS II of Mainland Portugal, as the available data for the *Região Autónoma da Madeira* and the *Região Autónoma dos Açores* were particularly scarce and fragmented. Additionally, the bank appraisal value indicator used as a proxy for the housing market prices, introduces a bias in the analyses. On one hand, it may under- or overestimate the real market value, depending on the region. On the other, as this indicator is an estimation made by banks, there may be a time gap between the actual market values and the estimations.

These limitations point to several directions for future research. One relevant extension would be the integration of additional data sources, such as administrative microdata, geospatial information, or alternative indicators such as affordability (price-to-income ratios and housing overburden rates) and short-term rentals, to better capture emerging forms of housing pressure. Future work could also expand the analytical framework to a comparative, cross-country setting, enabling systematic assessments of housing inequalities. In this respect, the reliance on harmonized official statistics offers a clear advantage for international replication. Finally, the framework could be further developed to support prospective and scenario-based analysis, linking housing indicators to demographic projections, policy interventions, or macroeconomic conditions.

This work constitutes the foundation for the development of an online interactive visual storytelling tool developed for PLANAPP to support policy monitoring and communication of housing inequalities to policymakers and other stakeholders.

In conclusion, this study shows that integrating BI pipelines, data science methodologies, and urban analytics techniques into the analysis of official housing statistics substan-

tially improves the capacity to diagnose, monitor, and communicate housing inequalities. Although applied to a single country in this manuscript, the proposed framework is transferable across territorial contexts, offering a scalable template for data-driven housing observatories and public-sector analytics.

More broadly, this study shows that the joint analysis of valuation dynamics, housing stock, and construction trends can help identify economic, social, and spatial inequalities in housing in a systematic and scalable way. The use of official data sources and BI-oriented analytical pipelines creates the basis for a comparable monitoring architecture that may support housing observatories, territorial diagnostics, and evidence-based policy design across different institutional contexts.

Finally, the Portuguese patterns identified in this paper, namely, price polarization, territorial mismatch between pressure and supply, and stronger metropolitan concentration, should be interpreted as part of wider international housing trends rather than as an isolated national phenomenon. For that reason, the results are relevant to a global audience concerned with affordability pressures, uneven territorial development, and the need for more robust public-sector tools to monitor housing policy and inequality.

**Author Contributions:** Conceptualization, J.B. and E.C.; methodology, J.B. and E.C.; data curation, A.C., J.B. and E.C.; formal analysis, A.C.; investigation, A.C.; visualization, A.C.; writing—original draft preparation, A.C., J.B. and E.C.; writing—review and editing, J.B. and E.C.; supervision, J.B. and E.C.; project administration, J.B. and E.C.; funding acquisition, J.B. and E.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by FCT—Fundação para a Ciência e Tecnologia, grant number 2024.07395.IACDC, project ISDAPPP—Sistemas de Informação e Análise de Dados para o Planeamento de Políticas Pública (in Portuguese).

**Data Availability Statement:** All data used in this study is publicly available from official sources. Details to access the data are provided under Section 3 of this article.

**Acknowledgments:** This work is supported by UIDB/04466/2023, UIDP/04466/2023, and UID/04516/2025 with the financial support of FCT—Fundação para a Ciência e Tecnologia. During the preparation of this manuscript, the authors used AI-assisted tools, namely ChatGPT (GPT-5), for language editing and clarification. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

AML	Lisbon Metropolitan Area
AMP	Porto Metropolitan Area
API	Application Programming Interface
BI	Business Intelligence
CRISP-DM	Cross-Industry Standard Process for Data Mining
EU	European
GEP	Strategy and Planning Office
INE	Statistics Portugal (from Portuguese, National Statistics Institute)
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organization for Economic Co-operation and Development
PLANAPP	Centre for Planning and Public Policy Evaluation
PIR	Price-to-Income Ratio
SGD	Sustainable Development Goals

## References

1. Tayefi Nasrabadi, M.; Larimian, T.; Timmis, A.; Yigitcanlar, T. Mapping four decades of housing inequality research: Trends, insights, knowledge gaps, and research directions. *Sustain. Cities Soc.* **2024**, *113*, 105693. [[CrossRef](#)]
2. Grotti, R.; Russell, H.; Maitre, B.; Gritti, D. The Experience of Housing Discrimination and Housing Deprivation Across Social Groups in Ireland. *Soc. Indic. Res.* **2024**, *175*, 195–215. [[CrossRef](#)]
3. Cavicchia, R. Are Green, dense cities more inclusive? Densification and housing accessibility in Oslo. *Local Environ.* **2021**, *26*, 1250–1266. [[CrossRef](#)]
4. Dorling, D. *All That Is Solid: How the Great Housing Disaster Defines Our Times, and What We Can Do about It*; Penguin: London, UK, 2014.
5. James, L.; Daniel, L.; Bentley, R.; Baker, E. Housing inequality: A systematic scoping review. *Hous. Stud.* **2024**, *39*, 1264–1285. [[CrossRef](#)]
6. Gupta, A.; Das, P.; Coulson, N.E.; Das, A. Salience of social identities in explaining homeownership patterns in India. *Urban Stud.* **2024**, *62*, 1435–1463. [[CrossRef](#)]
7. Sharma, M. Race/Ethnicity and Homeownership in an Emerging Immigrant Gateway of the US Southeast: A Neighborhood Scale Analysis. *Soc. Sci.* **2024**, *13*, 624. [[CrossRef](#)]
8. Farahi, A.; Jiao, J. Analyzing racial disparities in the United States homeownership: A socio-demographic study using machine learning. *Cities* **2024**, *152*, 105181. [[CrossRef](#)]
9. Vangeel, W.; Defau, L.; De Moor, L. Young Households' Diminishing Access to Homeownership Attainment in Europe. *Sustainability* **2023**, *15*, 6906. [[CrossRef](#)]
10. Zhang, L.; Ye, Y.; Chen, J. Urbanization, informality and housing inequality in indigenous villages: A case study of Guangzhou. *Land Use Policy* **2016**, *58*, 32–42. [[CrossRef](#)]
11. Flage, A. Ethnic and gender discrimination in the rental housing market: Evidence from a meta-analysis of correspondence tests, 2006–2017. *J. Hous. Econ.* **2018**, *41*, 251–273. [[CrossRef](#)]
12. Gusciute, E.; Mühlau, P.; Layte, R. Discrimination in the rental housing market: A field experiment in Ireland. *J. Ethn. Migr. Stud.* **2022**, *48*, 613–634. [[CrossRef](#)]
13. Nolan, B.; Winston, N. Dimensions of Housing Deprivation for Older People in Ireland. *Soc. Indic. Res.* **2011**, *104*, 369–385. [[CrossRef](#)]
14. Le Goix, R.; Ysebaert, R.; Giraud, T.; Lieury, M.; Boulay, G.; Coulon, M.; Rey-Coyrehourcq, S.; Lemoy, R.; Ramasco, J.J.; Mazzoli, M.; et al. Unequal housing affordability across European cities. The ESPON Housing Database, Insights on Affordability in Selected Cities in Europe. *Cybergeo Eur. J. Geogr.* **2021**, 974. [[CrossRef](#)]
15. Arbaci, S. Ethnic Segregation, Housing Systems and Welfare Regimes in Europe. *Eur. J. Hous. Policy* **2007**, *7*, 401–433. [[CrossRef](#)]
16. Byrne, M. Generation rent and the financialization of housing: A comparative exploration of the growth of the private rental sector in Ireland, the UK and Spain. *Hous. Stud.* **2020**, *35*, 743–765. [[CrossRef](#)]
17. Hochstenbach, C.; Ronald, R. The unlikely revival of private renting in Amsterdam: Re-regulating a regulated housing market. *Environ. Plan. A Econ. Space* **2020**, *52*, 1622–1642. [[CrossRef](#)]
18. Wijburg, G. Privatised Keynesianism and the state-enhanced diversification of credit: The case of the French housing market. *Int. J. Hous. Policy* **2019**, *19*, 143–164. [[CrossRef](#)]
19. Kim, K.-H.; Renaud, B. The Global House Price Boom and its Unwinding: An Analysis and a Commentary. *Hous. Stud.* **2009**, *24*, 7–24. [[CrossRef](#)]
20. Mueller, E.; Tighe, J.R. *The Affordable Housing Reader*, 2nd ed.; Routledge, Taylor & Francis Group: Abingdon, UK, 2022.
21. Mogila, Z.; Melo, P.C.; Gaspar, J.M. Estimating labour-income-based inequality of opportunity for a selection of EU-SILC countries: National and urban–rural perspectives. *Reg. Sci. Policy Pract.* **2022**, *14*, 1158–1188. [[CrossRef](#)]
22. Fialová, K.; Mysíková, M. Intra-household distribution of resources and income poverty and inequality in Visegrád countries. *Int. J. Soc. Econ.* **2021**, *48*, 914–930. [[CrossRef](#)]
23. Baláková, I.; Stávková, J.; Hudec, P. Determinants of individual income in EU countries: Implications for social policy targeting. *Front. Sociol.* **2023**, *8*, 1205094. [[CrossRef](#)]
24. Eeckhaut, M.C.; Stanfors, M.A. Educational assortative mating, gender equality, and income differentiation across Europe: A simulation study. *Acta Sociol.* **2021**, *64*, 48–69. [[CrossRef](#)]
25. Benedetti, I.; Crescenzi, F. The role of income poverty and inequality indicators at regional level: An evaluation for Italy and Germany. *Socio-Econ. Plan. Sci.* **2023**, *87*, 101540. [[CrossRef](#)]
26. Soaita, A.; MacLennan, D.; Gibb, K. The social and economic consequences of housing wealth inequalities. *Hous. Financ. Int.* **2020**, *2020*, 27–35.
27. Voronov, V.V.; Ruza, O.P.; Stasane, J. Multidimensional Poverty in the Baltic States in the EU Context: Theoretical and Practical Aspects. *Balt. J.* **2020**, *12*, 72–88. [[CrossRef](#)]

28. Annoni, P.; Weziak-Bialowolska, D.A. Measure to Target Antipoverty Policies in the European Union Regions. *Appl. Res. Qual. Life* **2016**, *11*, 181–207. [[CrossRef](#)] [[PubMed](#)]
29. EU Statistics on Income and Living Conditions (EU-SILC) Methodology—Housing Conditions. Available online: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU\\_statistics\\_on\\_income\\_and\\_living\\_conditions\\_\(EU-SILC\)\\_methodology\\_-\\_housing\\_conditions](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_housing_conditions) (accessed on 10 January 2026).
30. OECD Affordable Housing Database—HM1.1. Housing Stock and Construction. Available online: [https://webfs.oecd.org/Els-com/Affordable\\_Housing\\_Database/HM1-1-Housing-stock-and-construction.pdf](https://webfs.oecd.org/Els-com/Affordable_Housing_Database/HM1-1-Housing-stock-and-construction.pdf) (accessed on 10 January 2026).
31. Glossary: Shannon Evenness Index (SEI). Available online: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Shannon\\_evenness\\_index\\_\(SEI\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Shannon_evenness_index_(SEI)) (accessed on 10 January 2026).
32. Tufte, E. *The Visual Display of Quantitative Information*; Graphics Press: Cheshire, CT, USA, 1983.
33. Hosseinpour, H.; Matzen, L.E.; Divis, K.M.; Castro, S.C.; Padilla, L. Examining limits of small multiples: Frame quantity impacts judgments with line graphs. *IEEE Trans. Vis. Comput. Graph.* **2024**, *31*, 1875–1887. [[CrossRef](#)]
34. Breidenbach, P.; Schaffner, S. Inequality of earnings and housing costs—Developments between and within regions. *J. Labour Mark. Res.* **2025**, *59*, 4. [[CrossRef](#)]
35. Matějková, J.; Tichá, A. Housing Market Trends and Affordability in Central Europe: Insights from the Czech Republic, Slovakia, Austria, and Poland. *Buildings* **2025**, *15*, 1729. [[CrossRef](#)]
36. Leal, M.; Carreiras, M.; Alves, S. Decoding the Spatial Dynamics of Sales and Rental Prices in a High-Pressure Portuguese Housing Market: A Random Forest Approach for the Lisbon Metropolitan Area. *Cities* **2025**, *156*, 105631. [[CrossRef](#)]
37. European Commission. *Market Developments, Underlying Drivers, and Policies*; Discussion Paper 228; European Commission: Brussels, Belgium, 2024.
38. Kraff, N.J.; Wurm, M.; Taubenböck, H. Housing Forms of Poverty in Europe—A Categorization Based on Literature Research and Satellite Imagery. *Appl. Geogr.* **2022**, *149*, 102820. [[CrossRef](#)]
39. Oleńczuk-Paszal, A.; Sompolska-Rzechuła, A. Housing Conditions and the Quality of Life of the Populations of the European Union Countries. *Sustainability* **2025**, *17*, 1550. [[CrossRef](#)]
40. Ulbrich, P.; Porto de Albuquerque, J.; Coaffee, J. The Impact of Urban Inequalities on Monitoring Progress towards the Sustainable Development Goals: Methodological Considerations. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 6. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.