

Article

Open Data to Promote the Economic and Commercial Development of the Housing Sector: The Case of Spain

Ricardo Curto-Rodríguez ¹, Rafael Marcos-Sánchez ², Alicia Zaragoza-Benzal ³ and Daniel Ferrández ^{3,*}

¹ Facultad de Economía y Empresa, Universidad de Oviedo, 33066 Oviedo, Spain; curtoricardo@uniovi.es

² Facultad de Educación, Universidad Internacional de La Rioja (UNIR), 26006 Logroño, Spain; rafael.marcos@unir.net

³ Departamento de Tecnología de la Edificación, E.T.S. de Edificación de Madrid, Universidad Politécnica de Madrid, Avda. Juan de Herrera, 6, 28040 Madrid, Spain; alicia.zaragoza@upm.es

* Correspondence: daniel.fvega@upm.es

Abstract

Data is the starting point for generating information and knowledge in the decision-making process. Open data, which is information disclosed free of charge through open licenses and reusable formats, has great potential for value creation. Therefore, the objective of this research is to evaluate Spanish autonomous communities' open data initiatives in a category of information of vital importance: housing. The methodology employed was a population analysis of datasets labeled as housing, followed by a necessary data cleansing process due to the identification of various errors, which reduced the number of labeled datasets from 1000 to 599. Only 12 of the 17 autonomous communities provided this type of information. The analysis of the results reveals that autonomous communities' approaches to open data initiatives are highly heterogeneous and that the supply is irregular, with the Basque Country accounting for 70% of the datasets considered in the research. The creation of an indicator that equally assesses the existence of information and file formats (breadth and reusability) continues to identify the Basque Country as the undisputed leader, with Catalonia and Cantabria in second and third place, the only autonomous communities to exceed 50 points out of a possible 100. The study concludes by highlighting that the lack of uniformity in the formulation and implementation of open data policies will limit the use of information and, consequently, its value. Therefore, a series of recommendations is issued in this regard.

Keywords: Spain; autonomous communities; public policy; open data; value creation; reuse; affordable housing; economic and commercial development; public information



Academic Editors: Xin Janet Ge, Jyoti Shukla, Godwin Kavaarpuo, Shuya Yang and Piyush Tiwari

Received: 9 June 2025

Revised: 14 July 2025

Accepted: 15 July 2025

Published: 17 July 2025

Citation: Curto-Rodríguez, R.; Marcos-Sánchez, R.; Zaragoza-Benzal, A.; Ferrández, D. Open Data to Promote the Economic and Commercial Development of the Housing Sector: The Case of Spain. *Urban Sci.* **2025**, *9*, 277. <https://doi.org/10.3390/urbansci9070277>

Copyright: © 2025 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

“Data is the new gold”, Neelie Kroes, former Vice President of the European Commission and head of the Digital Agenda, titled her conference on open data strategy held in Brussels, highlighting the importance of this data in supporting a new model of digital economy [1].

This statement was based on various studies. The first of these was Vickery (2011) [2], “Open data, a driver for innovation, growth, and transparent governance”, which quantifies the value of open data at \$40 billion per year directly, increasing to \$140 billion per year if the indirect impact is also considered. The second, a report entitled “Creating value through open data” [3], estimates the value creation by open data at more than 1% of GDP. Finally, the report entitled “The impact of open data: opportunities for creation in Europe” [4] indicates,

in its conservative forecast, that the turnover of open data in the EU 27+ environment by 2025 would be between 199 and 344.2 billion, generating between 1.12 and 1.97 million jobs.

Therefore, it is not surprising that the interest in open data, a central element of government policies or open government [5], for its ability to generate value, favors the solution of various public problems [6,7] and the development of smart cities [8,9].

In today's digital society [10], public administrations are the organizations that produce and manage the greatest amount of data. They offer open data without technical or legal restrictions and without requirements for use or permissions [11] but with certain characteristics that make it valuable for free reuse [12].

It is worth noting that although many countries have robust data protection legislations in place to preserve individual privacy and respect intellectual property, this does not necessarily conflict with open data initiatives. As mentioned in the first of the eight principles of open data [13], called "complete", all public data that is available and not subject to valid limitations should be made available, thus excluding data that could compromise privacy or security or data subject to legal or commercial restrictions.

For this reason, it is firmly believed that the open data movement is unstoppable. These open data are easily found through online platforms [14] called open data portals.

Thus, given that open data contributes to innovation, it is logical that academic interest has increased [15], despite the fact that its understanding in certain aspects is still scarce [16]. This aspect, together with the undeniable interest in housing for citizens, businesses, and public administrations, has guided this work's one main research question: What is the supply of open data in the housing category being disseminated by Spanish autonomous communities?

Two additional supplementary questions have been added. The first is motivated by the vague definition of housing content contained in the technical standard for interoperability, which invites a typological classification of the results found, differentiated according to their nature: subsidies and grants, prices (rental or sale/purchase), building data, mortgages, statistics, etc. Thus, the question to be answered is: Is there a uniform interpretation by Spanish autonomous communities regarding the content of information to be disclosed?

The second is in line with the obligation contained in Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and reuse of public sector information (amending Directive 2013/37/EU of 26 June 2013) [17,18], which mentions that public sector information must be provided in machine-readable open formats. This aspect raises the following research question: Are the formats used really of high quality? To this end, the formats in which the information is being disseminated have been reviewed (considering both their breadth and their reusability), which, incidentally, allows for the creation of an autonomous community ranking that equally values the number of datasets provided and the formats offered.

To achieve these three tasks, this research has been divided into five sections, with this introduction being the first of them. The second section then deals with the methodology and the justification of the object and approach of the study; the third section shows the main results obtained, which are analyzed and discussed in the fourth section, together with the main conclusions reached and the perspective. The paper concludes with the bibliographical references used in the preparation of the study.

2. Theoretical Framework and Literature Review

Before reviewing academic studies, it is interesting to examine how open data initiatives are organized in Spain. Spain is a leader in open data portals with 279 operational initiatives, although not all of them are identified or classified as efficient initiatives due to

the difficulty in terms of professional use [19]. In this country, information is categorized according to the technical standard for interoperability [20], which proposes a classification of the data provided in 22 categories (see Figure 1).

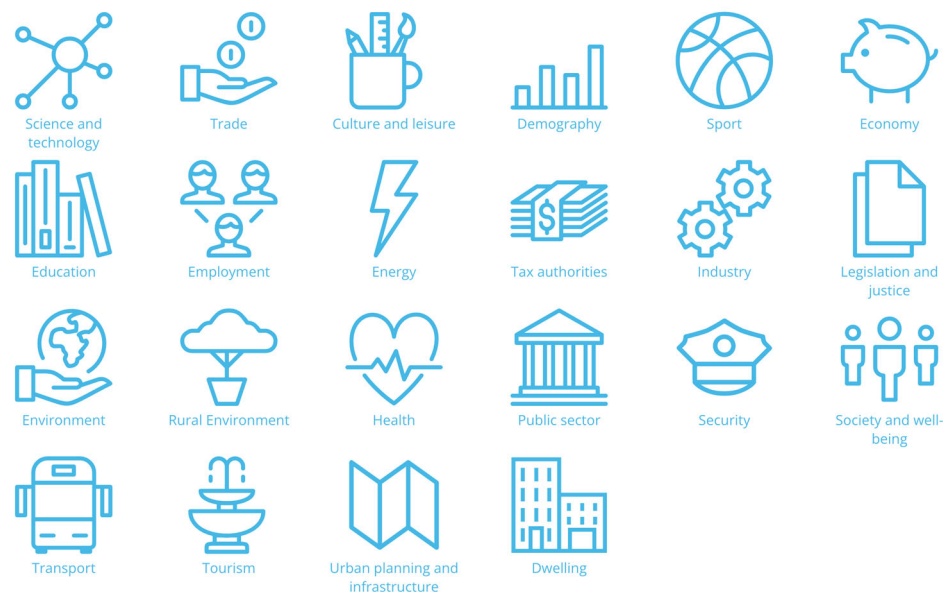


Figure 1. Categories of open data (open data portal of the Castilla-La Mancha).

The technical standard includes a proposal for a common taxonomy made up of the 22 categories mentioned above, indicating, succinctly, the content of each category (for example, it states that “Housing” should include information on the real estate market and housing construction).

There is no doubt that open data related to housing is of unquestionable interest to various actors (urban planners, construction companies, researchers, citizens, or governments), as it allows them to obtain different information such as prices or valuations (rents, purchases), volumes of supply and demand, number of dwellings (free, social or subsidized), and characteristics of the properties (surface area, number of rooms, age, etc.) that will favor informed decision making. Likewise, open housing data is a valuable resource for anticipating price trends; identifying geographic areas with potential growth or in the field of marketing (adapting products to the needs of consumers, creating content, carrying out geolocalized or better-segmented campaigns, useful for optimizing distribution channels, analyzing the competition, etc.); and finally, for creating public policies to promote affordable housing for vulnerable sectors such as young people and low-income families, among others.

Various initiatives featured on the datos.gob.es portal show how to identify and validate key data for the economic and commercial development of the real estate sector. Portals such as “Housing in data” from Barcelona City Council and the data visualization portal from Madrid City Council offer open, up-to-date, high-quality datasets such as consumer and housing price indices, land registry data, information for builders and companies for housing development, and data on energy demand in buildings [21]. Likewise, private platforms such as Idealista and Fotocasa, together with companies specializing in data reuse such as La Gistería, Inspide, and Sociedad de Tasación, integrate cadastral, registry, and statistical data to generate interactive maps and automated valuations [22]. These examples demonstrate that essential data—prices, volumes, building stock attributes, mortgages, and energy information—can be obtained through open portals that guarantee accessibility (csv, xml, json, and jscon formats), interoperability, and open licenses, facilitating their use in market analysis, commercial segmentation, and public policy formulation.

All of this improves the functioning of the housing market and, therefore, greater efficiency, generating value.

Housing has become a constant demand worldwide, and as part of government openness, portals have been promoted with the aim of providing information that encourages decision making and joint interinstitutional collaboration [23]. In Spain, housing is a sector of great relevance due to its contribution to GDP, and although the effects of recent interest rate increases are being felt [24], housing sales and purchases have been at levels not seen since the real estate boom of the early 20th century in 2021 [25].

Consequently, one would expect a wealth of research on open data related to housing in Spain. However, a consultation of a bibliographic database shows that this is far from the case. The database chosen was Scopus because it is considered suitable for demonstrating this aspect, both in terms of quantity and quality of articles: choosing WoS would have selected articles of generally higher quality but would have returned fewer results, while relying on Google Scholar would have reversed the parameters (a greater number of publications, but not all of a scientific/academic nature).

After justifying the choice of database, the search performed is shown below: open data OR open data AND Spain OR Spain, which located 112 works, a figure considered insufficient for a topic as crosscutting as open data.

The narrative review of the aforementioned articles shows that many of them focused their efforts on analyzing a specific topic such as health [26–28], forestry data [29,30], education [31], oceanography [32], research [33,34], legal information [35], libraries [36], archaeology [37], urban planning [38], energy consumption [39], attention to information science [40], or transparency [41,42].

Only one of those studies, by Barrientos Matute and Ruvalcaba-Gómez, dealt with housing and concluded that non-governmental portals offer information of greater value than governmental ones. However, curiously, it did not refer to Spain but to Guadalajara in Mexico [23].

After highlighting the lack of open data studies related to the housing sector in Spain, a brief review of the literature was conducted, analyzing a series of studies that present different approaches to studying open data in this country. Given that most of them aimed to assess whether initiatives are well or poorly designed by issuing a verdict on the matter, this paper outlines how they were approached, allowing for a comparison of the results with those of this study in the discussion section.

One of the pioneering studies is that of Garriga Portolá (2013) [43], who begins by pointing out that, at the time of publication of the study, Spain was the only country in the European Union with more than one million inhabitants that did not have a law regulating the fundamental right of access to public information. The author notes that, despite this, Spain remains a benchmark in data openness and has made significant progress, although there are still notable gaps.

He then refers to a relevant group of researchers, led by Alberto Abella, who are engaged in analyzing open data in that country, conducting a series of periodic studies entitled “The reuse of open data in Spain”, whose conclusions are based on sample analyses.

The first of these [44] points out that open data is an opportunity for Spain in three ways: increasing trust in public management, improving efficiency through easier use of public resources, and creating computer applications that take advantage of this data and generate economic value. However, over the years, there has been a noticeable shift in the researchers’ opinions, with a more critical tone. In the latest study [45], it is noted that the increase in information provided is mostly statistical rather than primary.

Similarly, reference should be made to the research by Curto-Rodríguez, all of which is autonomous community in scope. The author incorporates the role of open data in

generating value into his main line of research, which analyzes the role of open data in promoting transparency. The multidimensional study of the portals [46] assesses the existence of information, existing computer applications, the possibilities for interaction with the portal, and its functionalities, highlighting the different performances of each autonomous community in each of the dimensions analyzed. Other works carried out by other authors focus on analyzing the population or demographic data provided [47] or the role of accounting information in open data for accountability [48].

After conducting a brief review of the literature, which revealed an interesting knowledge gap, this research focused on open data relating to housing in Spain, a largely decentralized country [49] that is undoubtedly among the most decentralized in Western Europe [50]. In Spain, the Spanish public sector is divided into three levels: the state, autonomous communities (comprising 17 autonomous communities, see Figure 2), and local governments (comprising 50 provinces and 8112 municipalities), with a relationship among these governments based on competencies rather than hierarchy [51]. The framework chosen was that of the autonomous communities, which have a great deal of autonomy and important handling of competencies, managing a high volume of public resources [52].



Figure 2. Map of Spain's autonomous communities.

3. Materials and Methods

This section describes the procedure used in this study, with the aim of quantifying the actual provision of information on housing by the Spanish autonomous communities, analyzing the nature or type of information that predominates, and assessing the formats in which the information is being disseminated. All of this answers the three main research questions posed: Is the provision of housing information through open data sufficient?

Are the Spanish autonomous communities uniform in terms of the content of the housing information they provide to users? Moreover, is the training offered comprehensive and of high quality for reuse purposes?

The methodology employed was a population analysis of the datasets categorized by housing type. To this end, and first of all, all the portals of the Spanish autonomous communities' open data were visited throughout the months of October and November 2024. Sixteen of Spain's seventeen autonomous communities have an operational, modern, and functional data portal, which is the primary point of access for information. On the other hand, the portal of the Principality of Asturias (the mainland) experiences problems accessing data, which led to visiting the national portal and consulting the information that the Asturian government sends to the national government, thus accessing the content.

Once the portals were identified, the desired information was consulted, i.e., the datasets relating to housing. Using the information that all open data portals contain, the housing category was selected, resulting in a list of all datasets that had been tagged in this way. Unfortunately, in certain cases, the information was collected manually, as the searches carried out could not be downloaded (it was not even possible to sort the results alphabetically).

Consultation of this first database revealed that not all of them could be selected due to various errors identified. Firstly, datasets that were incorrectly labeled and did not belong to the housing category were eliminated, as were those that do not allow their contents to be downloaded. Secondly, those that were literally offered in duplicate or those that do not add value because they were merely a temporary fragmentation of another previously offered file were not taken into consideration. Thirdly, those datasets that presented autonomous community content were selected, and those provided by lower-level bodies (such as municipal or provincial) were discarded.

This screening was carried out by consensus of the research team. The section on the preparation of results is addressed, with the maximum level of detail possible, in the following section.

4. Results

Fieldwork began with counting the datasets that each autonomous community disclosed (Table 1). Table 1 shows the autonomous communities ranked according to the total number of datasets.

This initial collection of information shows that the total autonomous community offering is quite significant, as it is close to 40,000 datasets. In the analysis of Table 1, the uneven content offered by each respective autonomous community stands out. The Canary Islands stand out positively with 20,856 datasets and the Basque Country with 12,493, contributing 44.26% and 26.51%, respectively (which together means that these two autonomous communities comfortably supply more than two-thirds of the national total). At the opposite end of the spectrum is Extremadura, whose 13 datasets account for less than one-thousandth of the total.

The distribution is particularly uneven, as the expected average of 2772 datasets per autonomous community is only exceeded by three autonomous communities (the Canary Islands, the Basque Country, and Aragon), while fourteen are below the average. Obviously, this has an impact on the dispersion measures of the sample, whose standard deviation is 5482 (approximately double the value of the mean).

However, given that the focus of this research is on the datasets related to housing, Table 2 shows the autonomous community summary by category.

Table 1. Name of the autonomous community portals and their number of datasets in October 2024.

Autonomous Community	Portal Name/Web Address	Datasets
Canary Islands	Datos Abiertos Canarias	20,856
Basque Country	Opendataeuskadi	12,493
Aragon	Aragón Open Data	2849
Principality of Asturias *	Open Data del Principado de Asturias (*)	1586
Autonomus Community of Navarre	Open Data	1561
Valencian Community	Dades obertes gva	1494
Catalonia	Datos Abiertos GenCat	1279
Andalusia	Datos Abiertos	979
Region of Murcia	Datos abiertos Región de Murcia	975
Castilla and Leon	Datos Abiertos de Castilla y León	744
La Rioja	Datos Abiertos Rioja	528
Galicia	abert@s	512
Balearic Islands	Dades obertes caib	363
Castilla-La Mancha	Datos abiertos de Castilla- La Mancha	346
Cantabria	Icane—datos	291
Community of Madrid	Comunidad de Madrid datos abiertos	250
Extremadura	Gobierno Abierto—Catálogo de datos	13

* The Principality of Asturias deposits the datasets on the national portal, so the information can be accessed at <https://datos.gob.es/es/> (accessed 14 July 2025), since their own portal contains less data and is poorly designed. Source: Own elaboration.

Table 2. Autonomous communities with information on the subject of open data related to housing.

Category/Autonomous Communities	VAS	GAL	CNR	ARA	AST	NAV	VAL	CAT	CNT	CTM	RIO	AND	BAL
Commerce	122	0	85	67	136	9	0	15	6	9	1	16	0
Culture and Leisure	654	81	81	144	6	22	8	39	0	31	1	35	0
Demographics	1432	0	4707	65	654	6	38	33	10	9	10	32	5
Economy	1816	196	793	117	176	153	41	72	9	45	65	49	8
Education	195	45	111	969	3	8	78	50	15	22	3	31	2
Employment	909	30	7046	280	115	11	173	55	18	25	39	113	5
Energy	63	0	6	3	3	4	9	18	11	5	4	5	0
Environment	882	54	239	179	136	561	752	52	9	61	179	268	183
Finance	4789	0	14	48	0	89	2	14	5	18	51	30	10
Health	191	86	2691	274	54	129	50	63	14	48	68	12	26
Housing	568	153	145	40	26	22	14	9	8	8	7	5	4
Industry	108	0	166	110	55	6	4	16	26	9	6	13	3
Legislation and Justice	65	0	216	60	5	12	42	40	0	28	5	57	13
Public Sector	4562	78	899	169	43	529	49	167	4	66	17	154	30
Rural Environment	646	0	380	36	57	67	14	222	0	55	14	30	11
Science and Technology	290	39	463	71	18	28	0	17	0	8	1	15	2
Security	139	0	60	23	1	3	0	118	0	5	0	9	0
Society and Welfare	1473	0	2386	49	63	99	11	108	0	20	10	19	1
Sports	44	0	11	38	2	8	1	9	16	8	0	10	1
Tourism	417	0	1385	50	42	23	7	40	9	38	1	19	9
Transportation	458	9	179	32	7	239	31	55	11	12	10	14	3
Urbanism and Infrastructure	573	19	118	24	17	435	170	65	0	10	37	14	29
Sum of Categories	20,396	790	22,181	2848	1619	2463	1494	1277	171	540	529	950	345

Note: VAS, Basque County; GAL, Galicia; CNR, Canary Island; ARA, Aragon; AST, Principality of Asturias; NAV, Autonomous Community of Navarre; VAL, Valencian Community, CAT, Catalonia; CNT, Cantabria; CTM, Castilla-La Mancha; RIO, La Rioja; AND Andalusia; BAL, Balearic Island. Source: Own elaboration (2025).

Table 2 only includes thirteen of Spain's seventeen autonomous communities. The reason for this is that four of them—Extremadura, Madrid, Murcia, and Castile and Leon—did not have datasets labeled in the housing category, so they were obviously not included in the rest of the study, which focused on analyzing this category of information.

On the other hand, Table 2 reveals an interesting observation about the labeling of datasets, namely the heterogeneity in the interpretation of the standard by authorities of the autonomous communities when categorizing datasets for digital portals. In this sense, some autonomous communities applied multiple labels to their datasets (for example, the total offer of the Basque Country is 12,493 datasets, while the sum of the content of each category exceeds 20,000), such as the Valencian Community that performs it in a unique way (the total coincides with the sum of the categories) and others that have uncategorized datasets (such as Cantabria whose total offer is 291, while the sum of categories is 171).

Having made this comment, the content provided by the autonomous communities that offer datasets labeled under the category of housing is detailed below, in order of the total number of datasets labeled as housing (Table 2). Although this review is lengthy and tedious (and could well have been placed in an appendix to the research), it is considered particularly important in order to illustrate the reality and peculiarities of the offer from the autonomous communities.

4.1. Basque Country

Its autonomous portal contains 568 datasets within the housing category, making it the autonomous community with the most data available, almost quadrupling the labeled data of the second-ranked autonomous community, Galicia. However, many of these datasets do not contain information at the autonomous community level, such as the 114 datasets of the cadastral parcel at the municipality level or the urban planning approach (all offered by the Provincial Council of Vizcaya); others containing information on buildings (seven from the city council of Bilbao and one from Alava); and finally, a small group of datasets unrelated to housing such as “Basque society facing the future” and “GED: Demographic Survey Manager.”

Once this information was excluded, the remaining 423 datasets were provided by two data suppliers, EUSTAT (the Department of Statistics) and the Basque government, with no apparent difference in the content of each provider. In both cases, recurrent information, i.e., that which is updated every year or quarter in a new dataset, was therefore considered appropriate to proceed with the review by content blocks.

The first block collects the characteristics of both persons living in rental households (xls and csv), and needy dwellings (a dataset for 2005 offered in xls and another for 2006 offered in pdf) or non-mainstream dwellings (xls and csv).

A second block is related to housing. Specifically, a dataset is offered on the quantification of the need for access to housing in the autonomous community (xls and csv), together with seven other varied datasets that measure the demand for housing at different periods (1, 2, and 4 years) for the years 2005, 2006, and 2007, formulated in xls or pdf. Also related to buildings, “Building, residential and housing” is available in ods, and nine others relating to the buildings of the autonomous community by territorial areas that segment the information (by the state of the building, by its destination, by year of construction, etc.) are available in csv and xls. The following refers to “emancipation and housing” available in a wide range of formats (xml, xls, json, jsonp, and csv) and to “distribution of municipally owned housing in the autonomous community of Euskadi by historical territory and use”. Finally, there are twenty-five datasets that begin with the word housing: three on used housing for sale (years 2005–2007); two on planned housing; ten on family housing; three on free rentals; and six that offer information by territorial areas, neighborhoods, or municipalities.

Surveys occupy the third place with thirty datasets. The first dataset concerned rental households (xls), five others on housing needs (period 2008–2013), and twenty-four on real estate supply offered by quarters within the period of 2008–2013.

The fourth grouping, the most numerous, is made up of statistics. Sixteen datasets, mostly available in CSV and XLS, are related to building or housing, offering the information by quarters within the period of 2011–2015. Thirty-seven others concern housing starts and completions (one dataset per quarter for 2016–2024), accompanied by two summaries: “housing starts and completions statistics 1990–1999” and “housing starts and completions statistics 2000–2009”. Finally, twelve datasets on real estate registration statistics are provided (one dataset per year within the period of 2012–2024), mostly in csv and xml. One hundred and eighteen statistical tables are also found, such as twenty-four on building and housing (quarters in csv and xls or years up to 2022 in xls), six on housing needs and demand survey, another six on housing use (one dataset biannually in the 2015–2023 series), and twenty-four on the real estate supply survey (csv and xls) with quarterly or annual information for the period 2013–2021.

In fifth place, eighteen datasets called evolution deal with both housing supply and price, with information relating to the years 2005, 2006, and 2007.

The sixth block is also very numerous as it is made up of 102 datasets relating to financing: housing and land protection actions (one dataset per quarter for the period 2014–2024 available mostly in csv and xls); the loans approved in this regard; the loans formalized; and the subsidies recognized (all three offer a dataset on a quarterly basis for the period 2007–2011 in pdf).

Seventh, twelve datasets stand out as municipal sustainability indicators: average age of family homes, comfort index, etc. (lod, api rest, xls, pdf, and csv); the technical inspection of buildings (same formats); two licenses for major works (xls and csv) houses need or reasons for such need (years 2005–2006); three of renters (xls and csv); three of average housing price; and three on the average monthly rent. There is also a registry of public housing and the monitoring of housing policies (both in zip).

The review of the Basque Country ends by pointing out four datasets on households, all in XLS and CSV, and three on renting (socioeconomic characteristics, useful area, and monthly rent), accompanied by households by year according to the type of housing need. It ends with “Hollows in normal buildings”, which offers information on the census of buildings and their number of floors and hollows; Bizilagun Appointment, an entity that advises on the rights and duties of owners, landlords, or neighbors (zip); and “comparative of loans formalized by financial institutions” (pdf, xls, and csv).

4.2. Galicia

This autonomous community groups information about territory, housing, and transport into a single category, but it is curious that some information is not related to any of the three categories, such as “Galician blue flag beaches” (offering data from several years), “deaths”, etc.

It was necessary, therefore, to conduct a manual review of the 153 datasets to extract those with content regarding dwellings. The review begins with two datasets of statistical information (which have not been recently updated): “Survey of household living conditions: housing register”, which in txt format offers one dataset per year within the series 2007–2012, and “Structural household survey. General module”, which offers a dataset for the 2014–2021 series in xls format.

Referring to expenditure, the following are available—“Average monthly household expenditure on rent of main dwelling” and “Average monthly household expenditure on mortgage of main dwelling”—accompanied by four lists: “Households that have expenditure on mortgage of main dwelling”, “Households that have expenditure on rent of main dwelling”, “Households according to monthly income interval” and “Households according to household typology broken down” (all six offered in csv and xls). With the

same purpose, there is “Number and surface area of buildings and dwellings by type of construction” or “Construction statistics of buildings or dwellings that have certain”, both available in csv and xls.

Finally, and with the aim of providing information on housing, two sets of data are offered: “News of the Galician Institute of Housing and Health” (rss) and “Permanent advertising of second transmission housing” (html and pdf).

4.3. Canary Islands

This island is an autonomous community that offers the largest number of datasets at the national level; however, for the subset of data on housing, it is in third place after the Basque Country and Galicia.

Of the 145 datasets labeled in this category, a large number of datasets were found that did not seem to fit there, highlighting the 69 datasets relating to active social security registrations, four on registered employment, and other more general ones such as consumer price indices or a certain economic macro-magnitude. Despite the above, there is abundant information on housing, all offered in json, xml, tsv, csv, and xls formats, which are summarized below.

There are four sets of data relating to buildings: “Characteristics of buildings with partial demolition works by Canary Islands provinces and years”, “Characteristics of buildings with total demolition works by Canary Islands provinces and years”, “Characteristics of residential buildings to be constructed by Canary Islands provinces and years”, and “Characteristics of non-residential buildings to be constructed by Canary Islands provinces and years.”

This information is provided together with nine other datasets that quantify buildings or dwellings. As for the buildings, the following are indicated: “Number of non-residential buildings and surface area to be constructed according to destinations by provinces of the Canary Islands and years”, “Number of buildings according to types of work by municipalities of the Canary Islands and years”, “Number of buildings according to types of work and destinations by islands of the Canary Islands and years”, “Number of buildings according to types of work and destinations by provinces of the Canary Islands and years”, and “Number of buildings according to types of work and destinations-Spain and autonomous communities by periods”. Regarding housing, the following were found: “Number of dwellings of new construction, extension or refurbishment -Spain and autonomous communities by periods”, “Number of dwellings by type of construction by Canary Islands and years”, “Number of dwellings by type of construction by municipalities of the Canary Islands and years”, “Number of dwellings by type of construction and destination by provinces of the Canary Islands and years”, and “Number of dwellings by type of construction -Spain and autonomous communities by periods.” Finally, a dataset containing both terms was found: “Number of buildings and dwellings according to types of developer, -Spain and autonomous communities by periods”.

Next, four other datasets regarding operational concepts of statistics are highlighted: of appraised value of housing, of construction of buildings, of official bidding in construction, and of construction management visas and housing price index.

As for licenses, these datasets were “Number of municipal licenses for major construction work granted according to types of work by Canary Islands and years”, “Number of municipal licenses for major construction work granted according to types of work by Canary Islands municipalities and years”, and “Number of municipal licenses for major construction work granted according to Canary Islands provinces and years”, together with another five sets of data on tenders: “official tender in construction according to contracting agents and extended classification of work typologies in the Canary Islands by years”,

“Official tender in construction according to contracting agents and work typologies in the provinces of the Canary Islands by years”, “Official tender in construction according to contracting agents and work typologies-Spain and autonomous communities by periods”, “Official tender in construction according to forms of contracting in the Canary Islands by years”, and “Official tender in construction according to work typologies and extended classification of contracting agents in the Canary Islands by year”.

Next, there are four others with information about mortgages: “Mortgages canceled according to the nature of the property and the lender—autonomous communities and provinces by period”; “Mortgages cancelled according to the nature of the property—autonomous communities and provinces by period”; “Mortgages modified according to the nature of the property—autonomous communities and provinces by period”; and finally, “Mortgages modified according to the type of change in mortgage conditions—autonomous communities and provinces by period”. The following information should be added to this: “New mortgages constituted according to the nature of the property and the lender—autonomous communities and provinces by period” and “New mortgages constituted according to the nature of the property—autonomous communities and provinces by period.”

Subsequently, there is another block related to surface area consisting of six datasets: “Surface area to be extended or rebuilt according to type of work by provinces of the Canary Islands and years”, “Surface area to be built according to destination by Canary Islands and years”, “Surface area to be built according to destination by municipalities of the Canary Islands and years”, “Surface area to be built according to destination by provinces of the Canary Islands and years”, “Surface area to be built according to destination—Spain and autonomous communities by period”, and “Average surface area per dwelling of new construction or extension—Spain and autonomous communities by period.”

Finally, various information is mentioned, such as two sets of data called “Average appraised value of free housing by age of the dwelling—Spain, autonomous communities and provinces by quarters” and “Average appraised value of subsidized housing—Spain, autonomous communities and provinces by quarters”, together with “News from the Department of Public Works, Transport and Housing of the Canary Islands”, “Free and subsidized housing starts and completions—Spain, autonomous communities and provinces by quarters”, “Periods and free and subsidized housing starts and completions—Spain, autonomous communities and provinces by periods”, and “Number of certificates and settlement value -Spain and autonomous communities by periods.”

4.4. Aragon

It has 40 datasets related to housing, which will be discussed below. First, there is a series of datasets on the age of buildings, on real estate according to use, on rural land, on urban land, on the value of real estate, and on cadastral surfaces, which offer information differentiated at the level of municipalities, counties, provinces, or at the autonomous community level (all in csv format).

Secondly, there are datasets called “Buildings, surface area, and housing”, “Municipal building licenses”, and “Building management permits”, all in xls format, which are offered in a segmented manner, either by municipalities, by counties, or by provinces (this last formulation also includes the autonomous community level as a whole).

To conclude the analysis of this autonomous community, various information is reviewed, such as that relating to mortgages: “Distribution of the volume of new mortgage credit and its average duration”, “Average mortgage and average interest rate of new mortgage loans”, and “Average amount of mortgage credit per housing transaction” (all in csv and px formats). Finally, it mentions “Number of housing transactions” and “Average

surface area of transferred housing” (csv and px), accompanied by “Housing rental bonds deposited with the government of Aragon” (zip).

4.5. *Principality of Asturias*

This autonomous community has its own open data repository with an inadequate design, both in terms of the container (portal design) and content (number of datasets), which seem to be the reason for syndicating its datasets directly to the national portal. Therefore, the national portal was accessed at data.gob, and the Principality of Asturias was searched for in order to access the 26 datasets of the housing category, identifying two predominant blocks.

On the one hand, there are six datasets called population and housing censuses in Asturias. The first one offers detailed information at the council level (html), while the remaining five, available in xls, ods, json, and csv and which only contain information up to 2011, mention their characteristics: number of floors, parking spaces, etc.

On the other hand, there is information about the real estate market, with fifteen datasets offered in html (which redirects to a search engine that allows for downloading the information in xls). Four of them contain valuations: the average value (and total value) of free housing transactions, the appraised value of dwellings, and the value of free housing in municipalities with more than 25,000 inhabitants. There are another seven on transfers of properties, land (according to the size of the municipality or according to the nature of the purchaser), and dwellings (according to the nationality of the purchaser, according to the type of dwelling, or according to regime and state), completing the information in the dataset “Real estate market in Asturias: properties and dwellings transferred according to nature and title of acquisition”. Two others deal with prices (average price of urban land according to the size of the municipality) and the housing price index (according to type), completing this block on the stock of unsold new housing and the evolution of the average area of the transferred dwellings.

The remaining datasets to complete the 26 datasets offered in the housing category are those relating to energy efficiency certificates of buildings and housing (xls), indicators of the sustainable development objectives of the Principality of Asturias (xls and zip), and municipal building permits in Asturias with monthly data on new buildings (html) and by type of developer (format not specified).

4.6. *Autonomous Community of Navarre*

The Autonomous Community of Navarre offers twenty-two datasets within the housing category, although three of them present general information such as companies with economic activity in the Autonomous Community of Navarre and two directories (of companies and establishments) that carry out their activities there (csv, json, ods, tsv, xls, and xml).

More relevant to the topic are, first, two datasets containing a census of housing applicants (by characteristics and by place of census registration) available in xml, while the rest refer to housing itself.

Of the latter, the first block to be highlighted (downloadable in csv, json, ods, tsv, xls, and xml) is composed of the total number of housing starts and the total number of housing completions in the Autonomous Community of Navarre. In addition to these two datasets showing the total by years, housing—both started and completed—is offered for the period 1994–2012 (which does not seem to add any usefulness to the previous ones), and two other datasets offer that information broken down by locality accompanied by information on developments (kml).

In terms of subsidized housing, there are datasets on subsidies for subsidized housing (csv, json, ods, tsv, xls, and xml) and on surface area and visas for such housing (both in ods and xls). In addition, there are four datasets related to subsidized housing rehabilitation: started, started by locality, finished, and finished by locality (csv, json, ods, tsv, xls, and xml).

Finally, there is the general building assessment register for the Autonomous Community of Navarre (curiously, there are two sets of data with the same content), one in kmz format and the other in csv, json, ods, tsv, xls, and xml, and the evaluation reports carried out (xls).

4.7. Valencian Community

This autonomous community offers fourteen datasets. Five of them are called “viewers of the Valencian Community”, with information on sensitive urban spaces, and there are three datasets on housing rental prices (price table 2018, average price for 2018 and period 2016–2018) and on accessibility in buildings and organized public spaces. Only the first of the five datasets specifies format (wms, wfs, and pdf) since the rest redirect to a map with the information but does not allow for downloading the results.

There are also two other datasets that do not indicate format: “Integrated sustainable urban development strategy” and “Urban renewal and regeneration areas of the Autonomous Community”. The datasets that complete the offer are “Housing of the public patrimony of the Catalonia: offered and awarded” (csv), “Buildings with more than 50 years old and residential use” (wms and wfs), and “Evaluation report of the building of the housing of the Valencian Community” (wms and wfs). Information on rehabilitation areas is also available (pdf) and programmed rehabilitation residential environments (wms and wfs) as well as the study of universal accessibility in public buildings (wms, wfs, and pdf) and a detailed study of such accessibility conditions (wfs and pdf).

4.8. Catalonia

Catalonia has nine datasets hosted in its portal within the housing category, all of them, except “Census of buildings”, offered in html and offered in a wide formative offer (geojson, kml, kmz, rdf, rss, shapefile, and xml).

Three datasets show historical information on housing: “Housing starts and completions, annual historical series 1990-present”, “Quarterly historical series 2000-present”, and “Official protection annual historical series from 2002-present”. Two datasets are on the average rental price of housing (by the municipality and by supra-municipal area), and the remaining datasets are “Certificates of habitability”, “Actions of the neighborhood program 2004–2010”, and “Areas and zones of reference in terms of housing policy in Catalonia”.

4.9. Cantabria

It offers eight datasets with the label housing, all of them with a rather brief and generic title, which are available for download in html, rdf, xls, pc-axis, sdmx, and json.

The datasets are “Urban land prices”, “Estimated housing stock”, “Real estate transactions”, “Subsidized housing statistics”, “Free housing statistics”, “Housing price index”, “Appraised value of housing”, and “Real estate registry statistics.”

4.10. Castilla-La Mancha

It has a total of eight datasets related to housing. The first two contain information on mortgages: “Mortgage intermediation offices” (ods, csv, json, and pdf) and “Assistance, advice and mortgage intermediation in Castilla-La Mancha” (xls and csv). The following three could favor access to housing: “Housing aid and subsidies” (ods, csv, and json), “Aid for urban and rural regeneration and renovation in Castilla-La Mancha” (ods and csv) and “Supply of subsidized housing” (html).

There are also two other related datasets: “Entities for quality control” (xls and csv) and “List of testing laboratories for building quality control declared in Castilla-La Mancha” (ods and csv). Finally, there is the “Statistics Service Database” of general content as it is labeled in the 22 categories included in the technical standard for interoperability.

4.11. La Rioja

The seven datasets assigned to La Rioja within the housing category are available for download in four different formats: csv, json, xls, and xml.

Firstly, reference is made to the datasets “Certifications of completion-Number of dwellings by developer” and “Certifications of completion-Number of buildings by developer”, which show the evolution of construction activity and the building stock based on the information provided by the building permits and certificates of completion.

Another block of information is related to permits, such as “Visas and new construction management-Number of dwellings by type of work” and “Visas and new construction management-Number of buildings by main destination” or “Municipal licenses-Number of dwellings by type of work” and “Municipal licenses-Number of buildings by type of work.” The last dataset available in this category contains information on housing prices under the heading “New and second-hand housing price index.”

4.12. Andalusia

The main page of the open data portal of this autonomous community only shows twenty direct access icons, corresponding to each of the categories included in the technical standard, with only two missing (energy and housing). However, upon accessing the portal, it was discovered that they do exist as sectors of activity and that they offer a total of five datasets each.

Thanks to this, information on housing can be found, such as “Spatial data (they are cartographic bases) of reference of Andalusia” (wms, wfs, html, shp, and gpkg), “Statistical Yearbook of Andalusia”, “System of indicators for the monitoring and evaluation of the agenda for employment” (csv), “Multi-territorial information system of Andalusia” (csv), and “System of indicators of sustainable development of Andalusia for the 2030 agenda” (csv, html, and json). Unfortunately, it should be noted that none of them have a close relationship with housing.

4.13. Balearic Islands

Although the repository allowed us to consult the general content of the portal in Catalan or Spanish, the detailed information of the datasets is only available in Catalan, so it had to be translated.

The Balearic Islands contains four sets tagged in the housing category: “Residential vacant land 2015”, “Current habitability certificates of Menorca”, “Tourist accommodations of Menorca” and ‘Stays and tourist holiday homes of Menorca’, all of them exportable in csv, geojson, kml, kmz, rdf, rss, tsv, and xml.

With the Balearic Islands, the review of tagged information is complete in the category for the thirteen Spanish autonomous communities that have information available. The following section proceeds to analyze the results.

5. Results Analysis

After the laborious process of reviewing the more than 1000 datasets labeled in the housing category supplied by the thirteen Spanish autonomous communities, the analysis of results begins by showing, first of all, an overview of the general situation, pointing out that these 1004 labeled datasets only represent half of those expected in the face of a uniform distribution (47,119 datasets/22 categories = 2141 datasets). This shows that

attention to housing-related information is not a priority, a statement that is evident when, upon totaling the autonomous community supply by category, it is observed that it ranks fifteenth, only ahead of the public sector at 983; security at 361; society and welfare at 553; transport at 482; tourism at 361; urbanism and infrastructure at 163; and housing with 157 datasets (the most populated is Science and Technology with 8855 datasets).

In addition, and as the detailed review of the labeled datasets has shown, purification was essential. For this reason, those that were still labeled as housing, but which had no relation to that category (for example, Galician beaches with blue flag), were eliminated. Those that were duplicated or contained redundant information (housing starts and housing completions in the Autonomous Community of Navarre, accompanied by another dataset relating to the period 1994–2012) or those that did not allow for downloading the information (accessibility in building and in public spaces organized by the Valencian Community) were also removed. The results, after applying this filtering process, are shown in Table 3.

Table 3. Datasets tagged and selected in the housing category. Percentages.

Autonomous Community	VAS	ARA	CNR	AST	NAV	GAL	CAT	VAL	CNT	CTM	RIO	BAL	AND	Totals
Total offer autonomous community	12,493	2849	20,856	1586	1561	512	1279	1,494	291	346	528	363	979	45,137
Tagged as housing (%)	568	40	145	26	22	153	9	14	8	8	7	4	5	1009
Selected housing datasets	423	40	38	25	18	12	9	8	8	7	7	4	0	599
% Retained	74	100	26	96	82	8	100	57	100	88	100	100	0	59
% Selected out of total	3.39	1.40	0.18	1.58	1.15	2.34	0.70	0.54	2.75	2.02	1.33	1.10	0.00	1.33

Note: VAS, Basque County; GAL, Galicia; CNR, Canary Island; ARA, Aragon; AST, Principality of Asturias; NAV, Autonomous Community of Navarre; VAL, Valencian Community, CAT, Catalonia; CNT, Cantabria; CTM, Castilla-La Mancha; RIO, La Rioja; AND, Andalusia; BAL, Balearic Island.

Table 3 reveals some interesting findings, the first being that Andalusia, despite having five datasets labeled in the housing category, did not pass the exclusion filters because, as already pointed out in the detailed analysis of each autonomous community, none of them had a close relationship with housing. Secondly, Table 3 shows that after the datasets were cleaned up, the selected ones did not reach 600 (approximately 60% of those labeled). It is also noted that the percentages of dataset conservation were very disparate, with five autonomous communities maintaining 100%, while Galicia reduced its percentage to 8%. On the other hand, there was greater homogeneity in the percentage of preserved datasets with respect to the total, with a range of just over 3% (3.39% in the Basque Country compared to 0.18% in the Canary Islands).

In any case, and with the aim of clearly presenting the contribution of each autonomous community in terms of the selected datasets, Figure 3 is shown below.

As Figure 3 shows, the Basque Country provided more than 70% of the selected housing datasets. Aragón was in second place with 6.7%, followed by the Canary Islands with 6.3%. Notably, no other autonomous communities reached a contribution of 5%.

In relation to the second objective of the analysis of results, which involves categorizing the information provided, Table 4 is presented, ordered by autonomous community from highest to lowest number of items addressed to identify the type of information provided clearly.

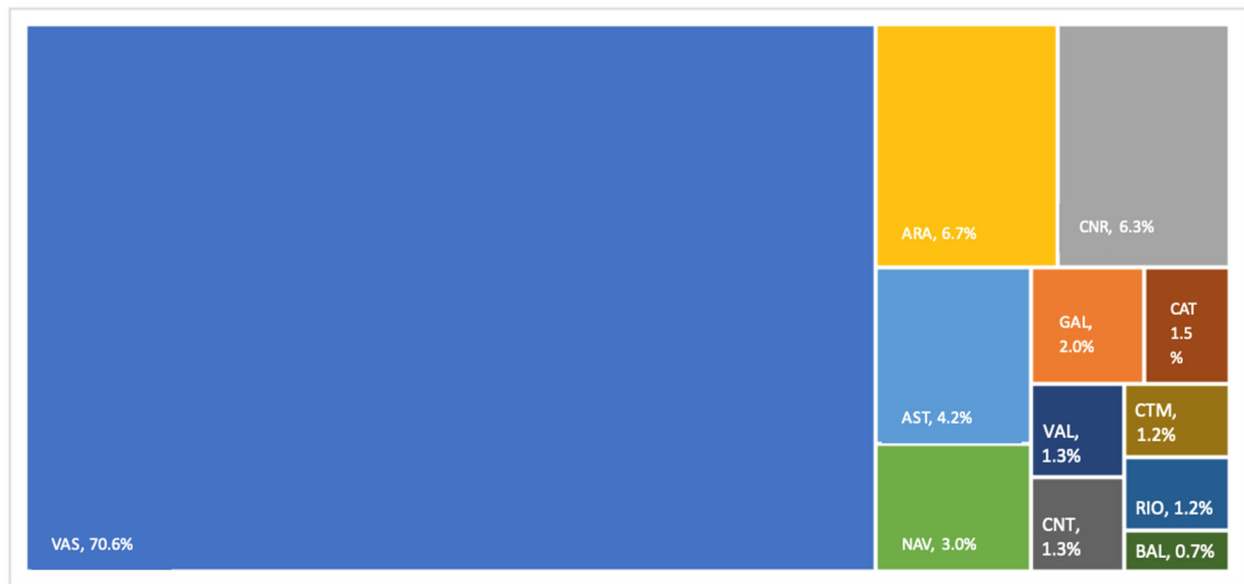


Figure 3. Role of each autonomous community in housing datasets (selected datasets). Note: VAS, Basque County; GAL, Galicia; CNR, Canary Island; ARA, Aragon; AST, Principality of Asturias; NAV, Autonomous Community of Navarre; VAL, Valencian Community, CAT, Catalonia; CNT, Cantabria; CTM, Castilla-La Mancha; RIO, La Rioja; BAL, Balearic Island.

Table 4. Housing information provided.

Category	Subcategory	Autonomous Community											Totals	
		VAS	ARA	CNR	AST	CNT	CTM	NAV	GAL	BAL	RIO	CAT		VAL
Aid and subsidies		1	0	0	0	0	1	1	0	0	0	0	0	3
Buildings	Census/historical information	1	0	1	0	0	0	1	1	0	0	1	0	5
	Certificates of occupancy—assessment reports	0	0	0	0	0	0	1	0	1	0	0	0	2
	Features-age-use	1	1	0	1	0	0	0	0	0	0	0	0	3
	Land/vacant land/value	0	1	0	0	1	0	0	0	1	0	0	0	3
	Prices/valuation	1	1	1	1	1	0	0	0	0	1	0	0	6
	Surfaces	1	1	1	0	0	0	0	0	0	0	0	0	3
Licenses or visas		1	1	1	1	0	0	0	0	0	1	0	0	5
Mortgages	Average mortgage cost	0	1	0	0	0	0	0	1	0	0	0	0	2
	Duration	0	1	0	0	0	0	0	0	0	0	0	0	1
	Interest rate	0	1	0	0	0	0	0	0	0	0	0	0	1
	Other	1	0	1	0	0	1	0	0	0	0	0	0	3
Rentals		1	0	0	0	0	0	0	0	0	1	1	3	
Statistics surveys		1	0	1	1	1	1	0	1	0	1	0	0	7
Subsidized housing		1	0	0	0	0	1	1	0	0	0	1	0	4
Tourist homes		0	0	0	0	0	0	0	0	1	0	0	0	1
Transactions		1	1	0	0	1	0	0	0	0	0	0	0	3
Items handled by autonomous community		11	9	6	4	4	4	4	3	3	3	3	1	55

Note: VAS, Basque County; GAL, Galicia; CNR, Canary Island; ARA, Aragon; AST, Principality of Asturias; NAV, Autonomous Community of Navarre; VAL, Valencian Community, CAT, Catalonia; CNT, Cantabria; CTM, Castilla-La Mancha; RIO, La Rioja; BAL, Balearic Island. Source: Own elaboration (2025).

As can be seen in Table 4, the most common information is related to buildings, as all the autonomous communities except Castilla-La Mancha cover one (or more) of the six subsections that comprise it, with data on prices standing out. This is followed by information on mortgages and information of a statistical nature, with clearly less attention being paid to the rest of the items considered. The analysis by autonomous communities shows that those with the most items are Basque Country with 11; Aragon with 9; and Canary Islands with 6. The fact that these three autonomous communities, in that order, are the autonomous communities with the most selected housing datasets requires calculating the Pearson correlation coefficient between these two variables, whose value is +0.775, indicating a strong positive association (and significant at $p < 0.01\%$).

The third task of the study was to analyze the licenses in which datasets are being offered as a preliminary step to the elaboration of an autonomous community ranking of the supply of housing datasets. To this end, five datasets from each autonomous community were sampled (four in the case of the Balearic Islands, since it did not reach this figure), for a total of 60 datasets (10% of the total). The formative breadth dimension (1 point per format—maximum 6 points) and the quality dimension (1 point per Tim Berners-Lee level—maximum 4 points) were evaluated. The formulation of these dimensions can be consulted in Berners-Lee [53], with the only adaptation of grouping the four- and five-star levels (both belonging to the semantic web level). The results are shown in Table 5, where the autonomous communities have been ranked according to the total score achieved by summing the existence of information and formats.

Table 5. Ranking of the autonomous communities to their housing datasets.

Scoring	VAS	CAT	CNT	CNR	NAV	RIO	BAL	CTM	ARA	GAL	VAL	AST
Autonomies number of housing datasets	423	9	8	38	18	7	4	7	40	12	8	25
Existence of information	50.00	1.06	0.95	4.49	2.13	0.83	0.47	0.83	4.73	1.42	0.95	2.96
Amplitude	12	30	30	25	21	20	24	13	7	9	9	6
Aperture	15	20	20	20	19	20	15	12	14	14	10	8
Formats	27	50	50	45	40	40	39	25	21	23	19	14
Total Score	77	51.06	50.95	49.49	42.13	40.83	39.47	25.83	25.73	24.42	19.95	16.96

Note: VAS, Basque County; GAL, Galicia; CNR, Canary Island; ARA, Aragon; AST, Principality of Asturias; NAV, Autonomous Community of Navarre; VAL, Valencian Community, CAT, Catalonia; CNT, Cantabria; CTM, Castilla-La Mancha; RIO, La Rioja; BAL, Balearic Island. Source: Own elaboration (2025).

As shown in Table 5, the Basque Country is the leading autonomous community in Spain thanks to being the largest disseminator of information, although its evaluation of the formats is intermediate. The second and third places are occupied by Catalonia and Cantabria, which achieve the highest scores in terms of the formats offered. As can be seen (with the exception of the Basque Country), the difference in the number of datasets selected is not very great, so the formats have played a decisive role in the final positions. To this effect, and in a negative sense, the performance of the Valencian Community and the Principality of Asturias should be highlighted, as they do not even reach a third of the score in formats, which condemns them to the last positions.

6. Discussion and Conclusions

Article 47 of the Spanish Constitution, a central element of the country's legal system, establishes the right to enjoy decent and adequate housing, attributing to the public authorities the obligation to promote the necessary conditions to make this right effective [54] and facilitating access to affordable housing through public housing policies. In Spain, where the weight of real estate activities in the GDP has not fallen below 10% in the last fifteen years, it can be seen that housing accounts for an increasing proportion of the disposable income of families, which causes unease and even gives rise to citizen demonstrations.

All of the above is proof of the interest in housing for companies, citizens, and public administrations alike. Therefore, in order to favor efficient decision making in this sector, quality data should be available that can be transformed into information and subsequently into knowledge [55,56]. This quality data would favor a commercial boost in infrastructure and the supply of affordable housing, especially for vulnerable sectors such as young people and low-income families, among others.

However, it seems that the problem of open data, which has advanced to be established as the second most advanced pillar of Spanish autonomous open government initiatives [57], is not with its quantity, since its supply has grown exponentially, but its quality [58]. This is partly due to the fact that many of the open data portals have not been designed for professional use [59]. This team of researchers has been developing reports measuring data disclosure in Spain for years, from the first work [45] to its fourth edition [44]. In the reports, they identify problems such as a lack of maintenance or updating of data, formats not suitable for reuse, and lack of standardization, which makes it difficult to consolidate data common to different sources.

This research also identifies certain gaps. Firstly, there is deficient attention given to the category “housing” given that only thirteen of the autonomous communities have some dataset labeled in this category of information, which is reduced to twelve after the filtering since the datasets labeled by Andalusia do not contain information in this regard. This statement is aggravated when, of the approximately 1009 datasets identified, a first review and debugging of non-downloadable and poorly labeled datasets, erroneous content, duplicate information, etc., reduced this to 599, leaving the housing category as one of the worst of the twenty-two proposed by the technical standard, with only 1% of the total data offered by the autonomous communities.

Secondly, there is a lack of relevant information on both general issues (such as prices, supply and demand volumes, etc.) and specific issues (such as the dissemination by all the Spanish autonomous communities of the energy certificates for buildings). In this sense, and in line with goal number 11 of the 2030 agenda for sustainable development, which aims to make cities and human settlements inclusive, safe, resilient, and sustainable, it is vital since, although cities barely occupy 3% of the earth’s surface, they account for 75% of carbon emissions and between 60% and 80% of energy consumption [60]. Smart cities require portals with a large amount of data that can be used by public and private entities to create new services and/or enable the improvement of internal processes [61]. The push for open data, combined with strategic planning, will drive public and, by extension, corporate policies that incorporate social justice objectives [62] through initiatives such as affordable housing.

Thirdly, the small number of “common” or identically formulated datasets will complicate the aggregation or combination of information from different suppliers, making automated data processing difficult. Therefore, it is necessary to reduce heterogeneity in the publication of open data [63]. This “Frankenstein” of Spanish open data initiatives [43], also highlighted by Curto-Rodríguez in his multidimensional analysis, is also evident here by not being able to identify many datasets with the same content [46]. On the other hand, the dissemination of housing datasets could be qualified as deficient since only one leading autonomy with good performance is identified (the Basque Country, with a score of 77 out of 100), and three other autonomous communities are around 50%. The difference between autonomous communities is, in general, very high since the average score of 38.65 points is accompanied by high dispersion values: range, 60.04; standard deviation, 17.22; and coefficient of variation, 28.69%.

These findings are in line with other studies [47,48], which point out a limited attention and interest in demographic information for marketing (given the high heterogeneity

observed) and an unequal attention to transparency through open data, with scores ranging from 0 points in Madrid and Extremadura to 53 in Catalonia [64].

Based on the conclusions reached in this research, which have been limited to being issued (due to the high heterogeneity observed) as a comparative approximation, similar to the work of Criado et al., (2023) [65] on the maturity of open public data portals in the public sector, a series of recommendations are made.

Firstly, there is a need for the unification of the criteria for the interpretation of what is indicated by the technical standard of interoperability for the reuse of information resources regarding the fact that each dataset must be classified into ONE of the proposed categories (it cannot be that some autonomous communities label univocally and others multiple or, what is worse, that some do not label their datasets, which makes it difficult to locate the desired information through the built-in search engines). Secondly, there is a lack of standardization in the content, which hinders its aggregation, giving as an example the roadmap developed by local governments (provinces and local entities) for the opening of their data, which shows a minimum set of 80 datasets to be disclosed along with a series of recommendations [66]. Thirdly and finally, there should be a requirement for an adequate range of formats to meet the requirements of different types of re-users as well as a sufficient formative quality typical of open data initiatives (a file offered in html or pdf should not be considered open data).

The paper concludes by mentioning that there is still a wide margin for improvement. Despite the fact that many hackathons and competitions with interesting prizes are being held for the development of open data if the raw material is not of high quality, this poses an additional challenge in terms of reuse by the various stakeholders involved [67]. It is essential to recall that these datasets serve as inputs for the development of artificial intelligence tools that can enhance strategic decisions in an increasingly complex business environment [68] and revolutionize the public sector [69]. The more information you have and the more data you have, the easier it will be to develop appropriate housing policies.

As future lines of research, the extensive and detailed database generated can be utilized to explore different approaches to the autonomous community reality of open data on housing. It can also focus efforts on one of the subcategories of information proposed in this work. It would also be interesting to develop this exploratory study longitudinally, which would show the evolution experienced. In any case, it is hoped that these studies will show that this is not a missed opportunity, or at least not one that has been sufficiently exploited.

Author Contributions: Conceptualization, R.C.-R. and R.M.-S. methodology, R.C.-R. and R.M.-S.; software, R.C.-R.; validation, R.C.-R., R.M.-S., A.Z.-B., and D.F.; formal analysis, R.C.-R. and R.M.-S.; investigation, R.C.-R. and R.M.-S.; resources, A.Z.-B. and D.F.; data curation, R.C.-R. and R.M.-S.; writing—original draft preparation, R.C.-R. and R.M.-S.; writing—review and editing, R.C.-R., R.M.-S., A.Z.-B., and D.F.; visualization, R.C.-R., R.M.-S., A.Z.-B., and D.F.; supervision, R.M.-S.; project administration, R.M.-S., A.Z.-B., and D.F.; funding acquisition, D.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are included in the manuscript.

Acknowledgments: The authors would like to thank to the Escuela Técnica Superior de Edificación de Madrid for the support shown during the author's research stay (2024–2025).

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

References

1. Kroes, N. Data Is the New Gold. 2021. Available online: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/speech_11_872/SPEECH_11_872_EN.pdf (accessed on 25 May 2025).
2. Vickery. Datos Abiertos un Motor para la Innovación, el Crecimiento y la Gobernanza Transparente. Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. 2011. Available online: https://datos.gob.es/sites/default/files/files/.opendata_es.pdf (accessed on 25 May 2025).
3. Capgemini. Creating Value Through Open Data. 2015. Available online: https://data.europa.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf (accessed on 25 May 2025).
4. Capgemini. El Impacto de los Datos Abiertos: Oportunidades de Creación en Europa. 2020. Available online: https://www.capgemini.com/es-es/wp-content/uploads/sites/16/2022/12/NP_Open-datav.pdf (accessed on 25 May 2025).
5. Díez-Garrido, M.; Melero Lázaro, M. La transparencia como medicina. La comunicación de la pandemia a través de los portales de datos abiertos de las autonomías. *Estud. Sobre Mensaje Periodístico* **2022**, *28*, 65–80. [CrossRef]
6. Pozen, D.E. Seeing transparency more clearly. *Public Adm. Rev.* **2020**, *80*, 326–331. [CrossRef]
7. Lnenicka, M.; Nikiforova, A.; Luterek, M.; Milic, P.; Rudmark, D.; Neumaier, S.; Santoro, C.; Flores, C.C.; Janssen, M.; Rodríguez Bolívar, M.P. Identifying patterns and recommendations of and for sustainable open data initiatives: A benchmarking-driven analysis of open government data initiatives among European countries. *Gov. Inf. Q.* **2024**, *41*, 101898. [CrossRef]
8. Lnenicka, M.; Nikiforova, A.; Luterek, M.; Azeroual, O.; Ukpabi, D.; Valtenbergs, V.; Machova, R. Transparency of open data ecosystems in smart cities: Definition and assessment of the maturity of transparency in 22 smart cities. *Sustain. Cities Soc.* **2022**, *82*, 103906. [CrossRef]
9. Lnenicka, M.; Nikiforova, A.; Clarinval, A.; Luterek, M.; Rudmark, D.; Neumaier, S.; Kevic, K.; Rodríguez-Bolívar, M.P. Sustainable open data ecosystems in smart cities: A platform theory-based analysis of 19 European cities. *Cities* **2024**, *148*, 104851. [CrossRef]
10. Hrustek, L.; Kević, K.; Varga, F. Open Data for Transparency of Government Tenders: A State Analysis in Croatian Agriculture Land Lease. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 401. [CrossRef]
11. López Domínguez, P.M.; Ramírez Amador, B.D. El concepto de Gobierno Abierto en la Administración Pública. *Rev. Univ. Potos.* **2023**, *23*, 22–27. Available online: <https://leka.uaslp.mx/index.php/universitarios-potosinos/article/view/382> (accessed on 25 May 2025).
12. Ruvalcaba-Gómez, E.A. Datos abiertos. In *EUNOMÍA, Revista en Cultura de la Legalidad*, 18th ed.; Universidad Carlos III de Madrid: Madrid, Spain, 2020; pp. 327–334. [CrossRef]
13. Sunlight Foundation. The 9 Principles of Open Government. 2017. Available online: <https://opengovdata.org> (accessed on 27 June 2025).
14. Mabillard, V.; Keuffer, N. What Drives Online Transparency Practices at the Local Level? Evidence from 2222 Municipalities in Switzerland. *Yearb. Swiss Adm. Sci.* **2020**, *11*, 158–170. [CrossRef]
15. Alexopoulos, C.; Saxena, S. Moderating effects of gender, perceived importance and usage experience towards Open Government Data (OGD) adoption and usage. In Proceedings of the 24th Annual International Conference on Digital Government Research, Gdańsk, Poland, 11–14 July 2023. [CrossRef]
16. Tang, Y.; Li, S. Research of open government data in China: A systematic literature study in public administration (2013–2023). *Chin. Public Adm. Rev.* **2024**, *15*, 15396754241237634. [CrossRef]
17. Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de Junio de 2013, por la que se Modifica la Directiva 2003/98/CE Relativa a la Reutilización de la Información del Sector Público. 2013. Available online: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2013-81251> (accessed on 25 May 2025).
18. Directiva(UE) 2019/1024 del Parlamento PARLAMENTO EUROPEO Y DEL CONSEJO de 20 de junio de 2019 Relativa a los datos Abiertos y la Reutilización de la Información del Sector Público. 2019. Available online: <https://www.boe.es/doue/2019/172/L00056-00079.pdf> (accessed on 25 May 2025).
19. Abella, A.; Ortiz-de-Urbina-Criado, M.; De-Pablos-Heredero, C. Criteria for the Identification of Ineffective Open Data Portals: Pretender Open Data Portals. *Prof. Inf. Prof.* **2022**, *31*, 1. [CrossRef]
20. Ministerio de Hacienda y Administraciones Públicas. Resolución de 19 de Febrero de 2013, de la Secretaría de Estado de Administraciones Públicas, por la que se aprueba la Norma Técnica de Interoperabilidad de Reutilización de recursos de la Información. Boletín Oficial del Estado num. 54, de 4 de Marzo de 2013. 2013. Available online: [https://www.boe.es/eli/es/res/2013/02/19/\(4\)/con](https://www.boe.es/eli/es/res/2013/02/19/(4)/con) (accessed on 25 May 2025).
21. Ministerio Para la Transformación Digital y de la Función Pública. Datos Abiertos para Conocer Mejor la Situación de la Vivienda en España. 24 February 2025. Available online: <https://datos.gob.es/es/blog/datos-abiertos-para-conocer-mejor-la-situacion-de-la-vivienda-en-espana> (accessed on 14 July 2025).

22. Ministerio Para la Transformación digital y de la Función Pública. 9 Ejemplos de cómo los Datos Abiertos Pueden Facilitar el Acceso a una Vivienda. 17 August 2025. Available online: <https://datos.gob.es/es/blog/9-ejemplos-de-como-los-datos-abiertos-pueden-facilitar-el-acceso-una-vivienda> (accessed on 14 July 2025).
23. Barrientos Matute, J.P.; Ruvalcaba-Gómez, E.A. Metodología para explorar el valor público de datos abiertos ofertados en portales web de vivienda. *Econ. Soc. Y Territ.* **2023**, *23*, 405–431. [[CrossRef](#)]
24. Rodríguez-López, J. El mercado de vivienda en 2023: Se advierten los efectos de los aumentos de los tipos de interés. *Ciudad Y Territ. Estud. Territ.* **2024**, *56*, 307–320. [[CrossRef](#)]
25. Lajer Baron, A.; López-Rodríguez, D.; San Juan, L. El Mercado de la Vivienda Residencial en España: Evolución Reciente y Comparación Internacional. Documentos Ocasionales/Banco de España, 2433. 2024. Available online: <https://www.bde.es/f/webbe/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosOcasionales/24/Fich/do2433.pdf> (accessed on 25 May 2025).
26. Bertos Quesada, M.A.; Marco Cuenca, G. Análisis del estado actual de apertura de datos en salud a nivel autonómico a través del estudio de los portales de datos abiertos. *Ibersid Rev. Sist. Inf. Y Doc.* **2020**, *14*, 87–94. Available online: <https://www.ibernid.eu/ojs/index.php/ibernid/article/view/4680/4265> (accessed on 25 May 2025).
27. Muñoz-Lezcano, S.; de la Fuente, C.; Estella, A.; Amaya-Villar, R.; Puppo-Moreno, A.M.; Armengol-de la Hoz, M.A.; Garnacho-Montero, J. The open data revolution: Enhancing healthcare in intensive care units. *Med. Intensiv.* **2024**, *48*, 421–423. [[CrossRef](#)]
28. Ritoré, A.; Jiménez, C.M.; González, J.L.; Rejón-Parrilla, J.C.; Hervás, P.; Toro, E.; Parra-Calderón, C.L.; Celi, L.A.; Túnez, I.; Armengol de la Hoz, M.A. The role of Open Access Data in democratizing healthcare AI: A pathway to research enhancement, patient well-being and treatment equity in Andalusia, Spain. *PLOS Digit. Health* **2024**, *3*, e0000599. [[CrossRef](#)] [[PubMed](#)]
29. Valle-Basto, D.F.; Espinosa-Quiñones, T.; Limache-de-la-Fuente, D. Evaluación de la deforestación (2000-2020) en concesiones forestales peruanas en la provincia de Tambopata (Madre de Dios) usando plataformas de datos abiertos. *Rev. Kawsaypacha: Soc. Y Medio Ambiente* **2023**, *12*, A-010. [[CrossRef](#)]
30. Vega-Gorgojo, G.; Ordóñez, C.; Giménez-García, J.M.; Bravo, F. Explorando datos abiertos forestales masivos con un navegador web. *Ecosistemas* **2022**, *31*, 2452. [[CrossRef](#)]
31. Rivas-Rebaque, B.; Gertrudix-Barrio, F.; de Cisneros, J.C. The perception of the University Professor regarding the use and value of Open Data. *Educ. XXI Rev. Fac. Educ.* **2019**, *22*, 141–163. [[CrossRef](#)]
32. Wulff, E. Open Standards used in Oceanography Research Spatial Data Repositories in Spain. *DESIDOC J. Libr. Inf. Technol.* **2020**, *40*, 306–312. [[CrossRef](#)]
33. Peset, F.; Aleixandre-Benavent, R.; Blasco-Gil, Y.; Ferrer-Sapena, A. Datos abiertos de investigación. Camino recorrido y cuestiones pendientes. *An. Doc.* **2017**, *20*, 1–12. [[CrossRef](#)]
34. Aleixandre-Benavent, R.; Ferrer, A.; Peset, F. Compartir los recursos útiles para la investigación: Datos abiertos (open data). *Educ. Médica* **2021**, *22*, 208–215. [[CrossRef](#)]
35. Baquero Pérez, P.J. Datos abiertos entrelazados para una mejor información jurídica: Estado actual del Identificador Europeo de Legislación en España. *Rev. Española Transparencia* **2021**, *12*, 87–123. [[CrossRef](#)]
36. Obregón Sierra, A. Inserción de metadatos de las bibliotecas españolas en Wikidata: Un modelo de datos abiertos enlazados. *Rev. Española Doc. Científica* **2022**, *45*, a330. [[CrossRef](#)]
37. Previtali, M.; Valente, R. Archaeological documentation and data sharing: Digital surveying and open data approach applied to archaeological fieldworks. *Virtual Archaeol. Rev.* **2019**, *10*, 17–27. [[CrossRef](#)]
38. Patiño Quinchía, J.E.; Prieto-Curiel, R.; Anderson, B. Análisis de problemáticas urbanas a escala continental basado en datos abiertos: Espacios verdes, forma urbana y sostenibilidad futura de las ciudades en África. *Territorios* **2023**, *49*, 1–24. [[CrossRef](#)]
39. García-López, J.; Domínguez-Amarillo, S.; Sendra, J.J. Clustering Open Data for Predictive Modeling of Residential Energy Consumption across Variable Scales: A Case Study in Andalusia, Spain. *Buildings* **2024**, *14*, 2335. [[CrossRef](#)]
40. De Sena, L.G.; Ferreira, M.R.; de Castro, T.L. Application of Open Government Data in the Light of Information Science. *Cienc. Da Inf.* **2020**, *49*, 69–81. [[CrossRef](#)]
41. García García, J.; Curto-Rodríguez, R. El ejercicio de la rendición de cuentas mediante portales de datos abiertos en las comunidades autónomas españolas. *IDP. Rev. Internet Derecho Y Política* **2019**, *29*, 1–15. [[CrossRef](#)]
42. Curto Rodríguez, R. Operational transparency of the spanish autonomous communities through their open data portals= Transparencia operativa de las comunidades autónomas españolas mediante sus portales de datos abiertos. *Prof. Inf.* **2020**, *29*, 1–13. [[CrossRef](#)]
43. Garriga Portolá, M. El Frankenstein español del open data: Avances importantes, lagunas clamorosas. *Telos Cuad. Comun. E Innovación* **2013**, *94*, 68–73. Available online: <https://telos.fundaciontelefonica.com/archivo/numero094/el-frankensteinspanol-del-open-data/> (accessed on 8 June 2025).
44. Abella, A.; Ortiz-de-Urbina-Criado, M.; De-Pablos-Heredero, C. La Reutilización de Datos Abiertos: Una Oportunidad Para España. Cotec. Universidad Rey Juan Carlos. 2017. Available online: <https://burjcdigital.urjc.es/server/api/core/bitstreams/d389e82c-5c9b-491c-97b0-1c3adcf70fa0/content> (accessed on 25 May 2025).

45. Abella, A.; Ortiz-de-Urbina-Criado, M.; De-Pablos-Heredero, C.; García Luna, D. La Reutilización de los Datos Abiertos en España IV. Cotec. Universidad Rey Juan Carlos. 2024. Available online: <https://burjcdigital.urjc.es/server/api/core/bitstreams/1e628bc5-3439-4465-9f11-be5b05d09b6c/content> (accessed on 25 May 2025).
46. Curto-Rodríguez, R. Análisis multidimensional de los portales de datos abiertos autonómicos españoles. *Rev. Española Doc. Científica* **2021**, *44*, e284. [[CrossRef](#)]
47. Curto-Rodríguez, R.; Marcos-Sánchez, R. Datos abiertos para la creación de valor: Atención e interés para el marketing. *Eur. Public Soc. Innov. Rev.* **2024**, *9*, 1–16. [[CrossRef](#)]
48. Curto-Rodríguez, R.; Marcos-Sánchez, R.; Ferrández, D. Open Active Transparency in Spain: Regional Conglomerates and the Role of Accounting Information. *Adm. Sci.* **2025**, *15*, 80. [[CrossRef](#)]
49. Martínez-Vázquez, J. La descentralización tributaria a las comunidades autónomas en España: Desafíos y soluciones. In *Por Una Verdadera Reforma Fiscal*; Duran, J., Esteller-Moré, A., Eds.; Editorial Ariel: Barcelona, Spain, 2014.
50. Rodríguez-Pose, A.; Gill, N. The global trend towards devolution and its implications. *Environ. Plan. C Gov. Policy* **2003**, *21*, 333–351. [[CrossRef](#)]
51. Curto-Rodríguez, R. La información relacionada con la rendición de cuentas suministrada por las Comunidades Autónomas españolas. *Audit. Pública* **2016**, *67*, 61–68. Available online: <https://asocex.es/la-informacion-relacionada-con-la-rendicion-de-cuentas-suministrada-por-las-comunidades-autonomas-espanolas/> (accessed on 25 May 2025).
52. Carbonell, J.R.H.; Domínguez, M.B. Las Comunidades Autónomas ante la crisis económica: Impacto territorial de la recesión, políticas autonómicas de reactivación y tensiones en las cuentas públicas. In *Informe Comunidades Autónomas 2009*; Tornos, J., Moya, D., Eds.; Instituto de Derecho Público: Barcelona, Spain, 2010; Available online: https://idpbarcelona.net/docs/public/iccaa/2009/informe_2009.pdf (accessed on 8 June 2025).
53. Berners-Lee, T. Linked Data. The Year Open Data Went Worldwide. [Video Podcast]. 2010. Available online: https://www.ted.com/embed/tim_berniers_lee_the_year_open_data_went_worldwide (accessed on 25 May 2025).
54. Gobierno de España. Constitución Española. Boletín Oficial del Estado, 29 de diciembre de 1978, núm. 311. 1978. Available online: [https://www.boe.es/eli/es/c/1978/12/27/\(1\)/con](https://www.boe.es/eli/es/c/1978/12/27/(1)/con) (accessed on 25 May 2025).
55. Trespalacios-Gutiérrez, J.A.; Vázquez-Casielles, R.; Bello-Acebrón, L. *Investigación de Mercados: Métodos de Recogida y Análisis de la Información para la Toma de Decisiones en Marketing*; Editorial Paraninfo: Madrid, Spain, 2005.
56. Trespalacios-Gutiérrez, J.A.; Vázquez-Casielles, R.; De-la-Ballina-Ballina, F.J.; Suárez-Vázquez, A. *Investigación de Mercados: El Valor de los Estudios de Mercado en la era del Marketing Digital*; Editorial Paraninfo: Madrid, Spain, 2016.
57. Curto-Rodríguez, R.; Marcos-Sánchez, R.; Ferrández, D. Open government in Spain: An introspective analysis. *Adm. Sci.* **2024**, *14*, 89. [[CrossRef](#)]
58. Gao, Y.; Janssen, M.; Zhang, C. Understanding the evolution of open government data research: towards open data sustainability and smartness. *Int. Rev. Adm. Sci.* **2023**, *89*, 59–75. [[CrossRef](#)]
59. Ortiz-de-Urbina-Criado, M.; Abella, A.; De-Pablos-Heredero, C. Proposal for an index measuring the reputation of open data portals: The Odapri. *Prof. La Inf.* **2023**, *32*, e320312. [[CrossRef](#)]
60. Naciones Unidas. Informe de los Objetivos de Desarrollo Sostenible 2023: Edición Especial. 2023. Available online: https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023_Spanish.pdf?_gl=1*li1ez2*_ga*NjA4MjU5NDM2LjE3NDI2NTIwMDg.*_ga_TK9BQL5X7Z*MTc0MjY1MjAwNy4xLjEuMTc0MjY1MjA4NC4wLjAuMA (accessed on 25 May 2025).
61. Abella, A.; Ortiz-de-Urbina-Criado, M.; De-Pablos-Heredero, C. A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems. *Cities* **2017**, *64*, 47–53. [[CrossRef](#)]
62. Grigg, N.S. Local Public Works Management for Sustainable Cities: The United States Experience. *Urban Sci.* **2025**, *9*, 96. [[CrossRef](#)]
63. Ortiz-de-Urbina-Criado, M.; Abella, A.; García-Luna, D. Open data-set identifier for open innovation and knowledge management. *J. Knowl. Manag.* **2023**, *27*, 2779–2796. [[CrossRef](#)]
64. Curto-Rodríguez, R.; Marcos-Sánchez, R.; Del Canto-García, S. Diez años de transparencia activa autonómica en España mediante datos abiertos: Periodo 2013-2023. *Rev. Lat. Comun. Soc.* **2025**, *14*, 89. [[CrossRef](#)]
65. Criado, J.I.; Jiménez, C.; Alcaide Muñoz, L. Hacia la madurez de los portales de datos públicos abiertos en el sector público. Un análisis comparado del nivel local de gobierno en España. *Rev. Del CLAD Reforma Y Democr.* **2023**, *85*, 117–160. [[CrossRef](#)]
66. FEMP. Datos Abiertos FEMP. 80 Conjuntos de Datos a Publicar por las Entidades. 2024. Available online: <https://redtransparenciayparticipacion.es/wp-content/uploads/2025/03/Datos-abiertos-80-conjuntos-de-datos-vf.pdf> (accessed on 25 May 2025).
67. Begany, G.M. Open data in the health sector: Exploring questions of use and engagement. In *Research Handbook on Open Government*; Edward Elgar Publishing: Cheltenham, UK, 2025; pp. 106–122. [[CrossRef](#)]

68. López-Solís, O.; Luzuriaga-Jaramillo, A.; Bedoya-Jara, M.; Naranjo-Santamaría, J.; Bonilla-Jurado, D.; Acosta-Vargas, P. Effect of Generative Artificial Intelligence on Strategic Decision-Making in Entrepreneurial Business Initiatives: A Systematic Literature Review. *Adm. Sci.* **2025**, *15*, 66. [[CrossRef](#)]
69. Brandão, F.J.S.; Correia, R.M.; Paio, A. Measuring Urban Renewal: A Dual Kernel Density Estimation to Assess the Intensity of Building Renovation—Case Study in Lisbon. *Urban Sci.* **2018**, *2*, 91. [[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.