

Characteristics and Applications of Regional Open Demographic Data

Ricardo Curto-Rodríguez^{1*} , Marta Ortiz-de-Urbina-Criado² ,
Carmen De-Pablos-Heredero² , Rafael Marcos-Sánchez³ 

¹University of Oviedo (Spain)

²University Rey Juan Carlos (Spain)

³International University of La Rioja (Spain)

*Corresponding author: curtoricardo@uniovi.es

marta.ortizdeurbina@urjc.es, carmen.depablos@urjc.es, rafael.marcos@unir.net

Received: February 2026

Accepted: June 2026

Abstract:

Purpose: Open data initiatives offer substantial opportunities for business use, supporting informed decision-making and the creation of economic value. This study analyses the characteristics and applications of regional open demographic data.

Design/methodology/approach: After presenting the overall landscape and its evolution during 2024 and 2025, based on a review of regional portals or the corresponding sections of the national portal, a simple random sampling of demographic datasets is conducted to develop a regional ranking grounded in the availability and quality of data reuse. A detailed assessment of the open demographic data and its usefulness for each region is then provided.

Findings: The findings indicate that only six of the seventeen regions offer high-quality open demographic data.

Originality/value: The study contributes to the literature by improving the understanding of how open data initiatives are designed and implemented in Spain. From a managerial perspective, it identifies recurrent shortcomings and proposes improvements aimed at ensuring that these initiatives genuinely support value creation. Additionally, the identification of exemplary cases—regarding dataset content, information encoding, and data formats, both in scope and level of openness—provides actionable guidance for public managers responsible for open data strategies.

Keywords: open data, reuse, demographic data, regions, Spain

To cite this article:

Curto-Rodríguez, R., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C., & Marcos-Sánchez, R. (2026). Characteristics and applications of regional open demographic data. *Journal of Industrial Engineering and Management*, 19(2), 462–482. <https://doi.org/10.3926/jiem.9387>

1. Introduction

Open government has become one of the most significant paradigms in public administration over the past fifteen years (González-García & García-García, 2022). This governance model has been promoted through transparency legislation, mechanisms for accessing public sector information, and the development of policies and web platforms—such as portals for submitting access-to-information requests and open data portals (Cruz-Meléndez & Valencia-López, 2022). In addition, many countries, particularly developed ones, have adopted strategic plans to implement open government (Othman-Alanaimy, 2023), supported by technological tools that facilitate the use of collaborative intelligence (Criado-Grande, 2013) and, consequently, the multiplication of contributions and perspectives (Barrientos-Matute & Ruvalcava-Gómez, 2023).

It is therefore unsurprising that one of the key components of open government is the open data movement (Ferrer-Sapena et al., 2011), which has economic, political, and social implications (Corrales-Garay et al., 2019), given its potential to generate benefits for businesses, citizens, and governments alike (Putera & Irwanto, 2023). In Spain, the most recent edition of the infomediary sector report (13th edition), titled “*Economía del Dato en su ámbito infomediario (Data Economy in its infomediary sphere)*”, estimates a contribution of around 46 million euros to the national economy (ASEDIE, 2025).

Given this potential, the development of open data portals, particularly governmental ones (Alcaide-Muñoz et al., 2022), has increased substantially in recent years (Nikiforova & McBride, 2021; Gao et al., 2021; Wibowo et al., 2023). However, for data to create value, they must be published and provided appropriately. A single isolated datum has no meaningful value; therefore, as Hofmann (2015) notes, it is more useful to create a conglomerate of data organized around one or more specific themes, referred to as a dataset (Ortiz-de-Urbina-Criado et al., 2023).

Datasets must be accompanied by metadata that provides additional information such as descriptions, categories, publishers, dates, formats, and licenses. This allows datasets to be located and organized in catalogs (Máchová et al., 2018), which, if they also follow the eight open data principles (Open Government Working Group, 2007), are prepared to be freely used, reused, and redistributed for any purpose (Ruvalcaba-Gómez, 2020). In this way, datasets acquire the characteristics needed to enable data processing and value creation (Barrientos-Matute & Ruvalcava-Gómez, 2023), making it essential to have tools that assess the quality of data for reuse (Cerrillo-Martínez & Casadesús-de-Mingo, 2021).

In the business domain, open data can serve as an important secondary source of information, as it is generally provided free of charge and is often promoted through initiatives such as hackathons, forms for requesting new datasets, and application development competitions. In addition, Business Intelligence technologies and market research systems can incorporate available open data, with governmental datasets representing one of the most relevant inputs (Hair et al., 2022) for creating value (Pulido-Barreto, 2021) and supporting business decision-making (Van-den-Berg et al., 2025).

In Spain, information categorization is defined in Annex IV of the Technical Standard for the Reuse of Public Sector Information (Secretaría de Estado de Administraciones Públicas, 2013), which establishes a taxonomy requiring datasets to be classified into one of twenty-two categories. One of these categories —*Demography*— includes information on Family, Childhood, Immigration and Emigration, Older Adults, Women, and Population Registers. This type of information is highly valuable for market segmentation, enabling the identification of groups with distinct characteristics or behaviors (Kotler & Armstrong, 2006), and thus supporting Big Data technologies that are substantially relevant for marketing (Zuñiga et al., 2023).

However, knowledge about open data initiatives—both worldwide and in Spain—remains limited, whether in general or within specific data categories. Early studies, such as that of Ramos-Simón et al. (2012), analyzed ten European open data portals and concluded that despite initial efforts, few initiatives continued expanding their catalogs. Similarly, Jetzek et al. (2014), drawing on the Open Data Barometer, found that open data initiatives are still far from realizing their full potential, with evidence suggesting limited impact.

More concerning than this limited progress are the deficiencies observed in aspects such as licensing (Ramos-Simón, 2017) and data quality, particularly regarding a lack of updating and incompleteness

(Martínez-Gutiérrez, 2023). The main challenge lies in the difficulty of reusing datasets (Ortiz-de-Urbina-Criado et al., 2023), making improvements in standardization, interoperability, and user training necessary to foster greater reuse (Vidal-Cabo et al., 2024).

At the municipal level in Spain, Royo-Montañés and Benítez-Gómez (2019) report that Spanish city councils still have considerable room for improvement, particularly in aspects that promote data use by non-expert audiences (tutorials, demonstrations, or help materials) and in the quality of the data provided (especially metadata definition). At the regional level, the multidimensional analysis conducted by Curto-Rodríguez (2021) shows uneven performance across the four pillars assessed, with some regions achieving very low levels of development. In previous studies of industrial open datasets, a high variability is observed, which limits automated processing and reuse, indicating the need to improve data quality (Curto-Rodríguez & Marcos-Sánchez, 2024; Curto-Rodríguez, Marcos-Sánchez & Ferrández, 2025). Overall, data openness appears to be an evolving process in which Spanish regions still face challenges in promoting data reuse (García-Rodríguez & Mahou-Lago, 2025).

In particular, the demographic data provided by public authorities serve to address many public issues, as well as generating economic and social value and promoting the development of smart cities (Curto-Rodríguez & Marcos-Sánchez, 2024). Accordingly, the objective of this article is to deepen the understanding of open data initiatives in Spain, with a particular focus on the usefulness of demographic datasets at the regional level. From this objective, the study proposes several research questions:

RQ1: What characteristics do regional open demographic datasets present?

RQ2: How and for what purposes can open demographic data be used at the regional level?

To answer the first question, a sample of open demographic datasets published by each Spanish region in 2025 was selected, and both their level of openness and the formats in which they are available were analysed. These analyses provide an overview of the level of heterogeneity between regions and their potential for reuse. To explore the topic further and answer the second research question, an analysis was carried out of the datasets selected for each autonomous community, focusing on the usefulness of regional open demographic data, and several proposals for improvement were put forward.

The analyses carried out in this study make a contribution at both national and international levels, with their findings having implications for policy and practice in a range of contexts. In that sense, Spain is consistently ranked highly in global open data maturity rankings. In 2025, for example, it was ranked 5th in the EU Open Data Maturity Ranking and 7th in the Overall Ranking (Page et al., 2025). Furthermore, the latest open data maturity report notes that the Spanish public sector, through the Federation of Municipalities and Provinces, has institutionalised local responsibility for open data sets through various standardisation regulations to maintain a structured inventory of data sets and information sources issued by each local authority (Page et al., 2025, p. 35). However, the reality is that an adequate level of standardisation has not yet been achieved. The results of this study can help raise awareness of the differences between regions. This will help to implement actions and policies aimed at improving the quality and reusability of open data. Improvements in the quality of open data benefit governments, citizens, and businesses alike, as they can access services and applications created from such data and/or use them to launch new businesses.

Following this introduction, Section 2 presents the methodology. Section 3 reports the results addressing Research Question 1, while Section 4 addresses Research Question 2. The article concludes with a summary of the main findings, limitations, and future research directions.

2. Methodology

This section presents, on the one hand, the general context of the study and, on the other, the methodological design.

2.1. General Context of the Research

At the regional level, Spain is divided into 17 regions, known as Autonomous Communities. In a little over forty years, the country has become one of the most decentralized in the world (Martínez-Vázquez, 2014), surpassing

countries with well-established federal systems, such as Germany, in the transfer of competences and resources (Rodríguez-Pose & Gill, 2003). Specifically, these regions manage more than 35% of consolidated public spending, providing them with the capacity and resources necessary to undertake initiatives such as open data portals. In that context, Curto-Rodríguez & Marcos-Sánchez (2024) provide an overview of the main data offered by each autonomous community, with greater detail in the datasets tagged under the 'Demographics' category.

Table 1 shows the open data portals and datasets available in each region in 2024 and 2025. To identify them, the first point of reference was the regional open data portals themselves (except in the cases of the Principality of Asturias and Extremadura), which redirect and deposit their datasets in the national portal.

Autonomous Community	Portal Name	2024	2025	Difference
Andalusia	Datos Abiertos Junta de Andalucía	712	873	161
Aragon	Aragon Open Data	2,713	2,877	164
Balearic Islands	Dades obertes caib	453	15,041	14,588
Basque Country	Opendataeuskadi	11,539	14,877	3,338
Canary Islands	Datos Abiertos Canarias	14,237	22,066	7,829
Cantabria	Icane - datos	290	256	-34
Castilla and Leon	Datos Abiertos de Castilla y León	744	790	46
Castilla-La Mancha	Datos abiertos de Castilla- La Mancha	341	353	12
Catalonia	Datos Abiertos GenCat	1,218	1,443	225
Community Foral of Navarre	Open Data	1,561	1,885	324
Community of Madrid	Comunidad de Madrid datos abiertos	247	5,002	4,755
Extremadura	Gobierno Abierto - Catálogo de datos	12	48	36
Galicia	abert@s	512	554	42
La Rioja	Datos Abiertos Rioja	853	451	-402
Principality of Asturias	Open Data del Principado de Asturias	1,788	733	-1,055
Region of Murcia	Datos abiertos Región de Murcia	961	979	18
Valencian Community	Dades obertes gva	1,453	1,162	-291
TOTALS		39,634	69,390	29,756

Table 1. Regional open data portals and hosted datasets (Updated from Curto-Rodríguez & Marcos-Sánchez (2024) with information from datos.gob.es (Principality of Asturias and Extremadura), on the regional open data portals of each Spanish Autonomous Community)

An examination of Table 1 reveals that, during the period analyzed, the total number of datasets published by Spanish regions increased by nearly 30,000 datasets—an increase of more than 75%. However, the contribution of each region to this growth has not been uniform, and three groups have been identified.

The first group consists of regions that have reduced the number of datasets they publish. These include Cantabria and the Valencian Community, which show a slight or moderate decrease, and La Rioja and the Principality of Asturias, which reduced their dataset offerings by approximately 50%. This reduction in regional catalogs has also been noted by García-García and Curto-Rodríguez (2019), who report that Catalonia consolidated almost 1,000 geographic information system datasets into a single catalog to improve structure, searchability, and data reuse. In the present study, however, both La Rioja and Asturias appear to have removed certain datasets entirely from their portals.

The second group comprises regions with moderate increases, neither reducing their offerings nor showing substantial improvement. This group includes those that added between 12 and 324 datasets, such as Castilla-La

Mancha, the Region of Murcia, Extremadura, Galicia, Castilla and Leon, Andalusia, Aragon, Catalonia, and the Community Foral of Navarre.

Finally, the third group includes four regions that recorded a notable expansion. Two of these were already leading regions in 2024 and continued their upward trend: the Basque Country (with more than 3,000 additional datasets) and the Canary Islands (adding 8,000 datasets). The other two regions, Community of Madrid and the Balearic Islands, completely redesigned their open data initiatives, including the portal layout and functionalities, resulting in substantial increases: nearly 5,000 new datasets in Community of Madrid and 14,500 in the Balearic Islands.

The overall review also shows that most regions use the categories proposed by the Technical Interoperability Standard, except the Basque Country, which includes an additional category (Basque language), and Galicia, which groups certain categories (e.g., economy, business and employment; land, housing, and transport), similarly to Cantabria (e.g., jointly offering industry, energy, construction, and housing). Another interesting observation is that some regions do not assign datasets exclusively to a single category; instead, they assign them to multiple categories simultaneously.

2.2. Methodological Design

To address the first research question, the datasets labeled as demographic in each region were selected. Access was obtained either through the direct-access icons displayed on each portal's homepage—one corresponding to each of the 22 categories recommended by the Technical Interoperability Standard—or, when these were not available, by using the thematic search tool provided on the portal.

3. Results

We begin the results section by providing an overview of the number of datasets available in each category across Spain's regions and Autonomous Communities. The *demography* category contains the largest number of tagged datasets: 15,560 (Table 2). It is followed by employment with 12,257 datasets and the public sector with 9,424. The distribution across categories is highly uneven, as many of them fall well below the average of 4,251 datasets.

Categories/Autonomous Communities	CNR	VAS	ARA	AST	NAV	VAL	CAT	MUR	RIO
Commerce	254	126	68	4	8	0	17	10	0
Culture and Leisure	249	551	145	7	22	10	49	38	0
Demographics	4,884	1,633	162	419	4	41	33	95	2
Economy	1,125	2,071	117	98	156	64	76	34	28
Education	286	215	976	3	8	144	62	33	3
Employment	8,786	1,025	282	61	12	202	56	7	11
Energy	171	69	5	3	4	20	19	7	4
Environment	403	1,019	87	73	604	375	64	189	185
Health	2,772	216	274	35	179	57	70	65	68
Housing	316	590	40	6	20	7	9	0	3
Industry	337	110	110	24	6	4	16	5	6
Legislation and Justice	352	114	66	5	12	42	46	2	1
Public Finance	177	5,930	50	11	112	17	13	73	58
Public sector	1,030	5,367	172	25	678	42	202	286	17
Rural Development	544	689	38	19	70	17	272	1	14

Categories/Autonomous Communities	CNR	VAS	ARA	AST	NAV	VAL	CAT	MUR	RIO
Science and Technology	645	298	69	5	31	0	18	20	1
Security	224	170	23	1	4	0	129	0	0
Society and Well-being	2,216	1,440	49	45	122	12	113	75	10
Sports	174	68	38	2	8	1	9	8	0
Tourism	1,486	443	50	17	23	14	42	35	1
Transport	382	534	33	7	265	22	60	38	10
Urban planning and Infrastructure	281	689	20	8	446	97	67	3	29
TOTAL TAGGED	27,094	23,367	2,874	878	2,794	1,188	1,442	1,024	451

Note: ARA: Aragon, AST: Principality of Asturias, CAT: Catalonia, CNR: Canary Islands, MUR: Region of Murcia, NAV: Community Foral of Navarre, RIO: La Rioja, VAL: Valencian Community, VAS: Basque Country.

Table 2. Regional datasets tagged under each category of the technical standard (first part)

Categories/Autonomous Communities	CTL	AND	GAL	BAL	CTM	CNT	MAD	EXT	TOTAL
Commerce	6	16	0	642	10	0	48	0	1,209
Culture and Leisure	24	32	59	468	37	0	90	10	1,791
Demographics	25	41	0	7,961	10	24	226	0	15,560
Economy	16	52	248	869	48	25	169	4	5,200
Education	26	31	24	515	24	0	531	2	2,883
Employment	28	110	0	1,365	26	34	252	0	12,257
Energy	17	0	0	409	5	0	28	2	763
Environment	168	270	46	571	61	17	1,187	4	5,323
Health	127	9	73	631	48	0	802	17	5,443
Housing	0	0	0	423	7	0	102	1	1,524
Industry	3	13	0	440	9	34	22	0	1,139
Legislation and Justice	2	57	0	421	29	0	233	3	1,385
Public Finance	7	33	0	869	19	0	157	0	7,526
Public sector	121	156	73	1,027	68	14	143	3	9,424
Rural Development	59	28	0	460	56	23	128	2	2,420
Science and Technology	10	16	30	1,844	8	0	30	1	3,026
Security	3	9	0	409	5	0	26	0	1,003
Society and Well-being	18	57	0	3,088	22	54	373	1	7,695
Sports	7	10	0	469	11	0	10	1	816
Tourism	30	27	0	509	39	0	74	11	2,801
Transport	25	13	132	419	13	0	126	5	2,084
Urban planning and Infrastructure	14	15	0	446	12	0	130	1	2,258
TOTAL TAGGED	736	995	685	24,255	567	225	4,887	68	93,530

Note: AND: Andalusia, BAL: Balearic Islands, CNT: Cantabria, CTL: Castilla and Leon, CTM: Castilla-La Mancha, EXT: Extremadura, GAL: Galicia, MAD: Community of Madrid.

Table 2. Regional datasets tagged under each category of the technical standard (second part)

3.1. Demographic Datasets by Region

Table 3 presents the results for demographic datasets in each region for the most recent years (2024 and 2025).

Autonomous Communities	2025	2024	Difference
BAL	7,961	6	7,955
CNR	4,884	4,705	179
VAS	1,633	1,385	248
AST	419	803	-384
MAD	226	12	214
ARA	162	48	114
MUR	95	97	-2
AND	41	31	10
VAL	41	38	3
CAT	33	33	0
CTL	25	25	0
CNT	24	15	9
CTM	10	9	1
NAV	4	6	-2
RIO	2	18	-16
EXT	0	0	0
GAL	0	0	0
TOTAL	15,560	7,231	8,329

Note: AND: Andalusia, ARA: Aragon, AST: Principality of Asturias, BAL: Balearic Islands, CAT: Catalonia, CNR: Canary Islands, CNT: Cantabria, CTL: Castilla and Leon, CTM: Castilla-La Mancha, EXT: Extremadura, GAL: Galicia, MAD: Community of Madrid, MUR: Region of Murcia, NAV: Community Foral of Navarre, RIO: La Rioja, VAL: Valencian Community, VAS: Basque Country.

Table 3. Demographic dataset offering by region and year

Overall, regional datasets have increased, particularly in the Balearic Islands (51.1%), the Canary Islands (31.3%), and the Basque Country (10.4%). However, when examining the evolution in more detail, the Balearic Islands stand out as the main contributor, accounting for more than 95% of the total increase. They are followed by the Basque Country, Community of Madrid, the Canary Islands, and Aragon. A more moderate, almost symbolic increase can be observed in Andalusia, Cantabria, the Valencian Community, and Castilla-La Mancha. Conversely, several regions have not expanded their dataset offerings. The first group maintains their dataset count, including Catalonia and Castilla and Leon (alongside Extremadura and Galicia, which continue to provide no datasets). A second group has fewer tagged datasets than before, such as the Region of Murcia, the Community Foral of Navarre, La Rioja, and the Principality of Asturias, the latter showing a sharp reduction of 384 tagged datasets.

3.2. Regional Ranking of Demographic Information

To develop a regional ranking of demographic information, given the substantial volume and heterogeneity of datasets, a sampling approach was adopted. Ten datasets were selected per region, or all available ones if the total number was lower.

A systematic simple random sample was conducted by ordering all demographic datasets by relevance to ensure the selection of diverse dataset types. When this filter was unavailable, datasets were ordered by number of queries or most recent update. The resulting sample of 136 datasets is provided as supplementary information in an Excel file

that includes: the Autonomous Community, dataset title, dataset URL, exclusion status (with justification), and the formats available for download.

A data-cleaning process was conducted to exclude datasets that could not be considered demographic (i.e., not related to the category), or not regional (i.e., datasets sourced from non-regional providers or non-regional scopes). Datasets not offered as datasets were also excluded. Table 4 shows the initial number of datasets considered in the sampling and the final number retained. Only 57 of the 136 initial datasets passed all filters (42%).

The sampling also accounted for the formats in which datasets were available for download, to evaluate not only the number of datasets retained (10 points for each non-excluded dataset), but also their reusability quality. Reusability quality was measured using two dimensions: openness level and number of available formats, allowing each dataset to obtain up to 10 additional points. Regarding openness, the 5-star data model proposed by Berners-Lee (2010) was adapted, assigning points according to the highest-quality format offered (e.g., html: 1 point; doc: 2 points; csv: 3 points; rdf: 4 points). Regarding format breadth (1 point per format), a maximum of six different formats was established, based on the assumption that this is sufficient to meet the needs of diverse user profiles. Moreover, offerings exceeding six formats were generally variations of a previously listed standard format (e.g., json and JSON-Google or json-stat).

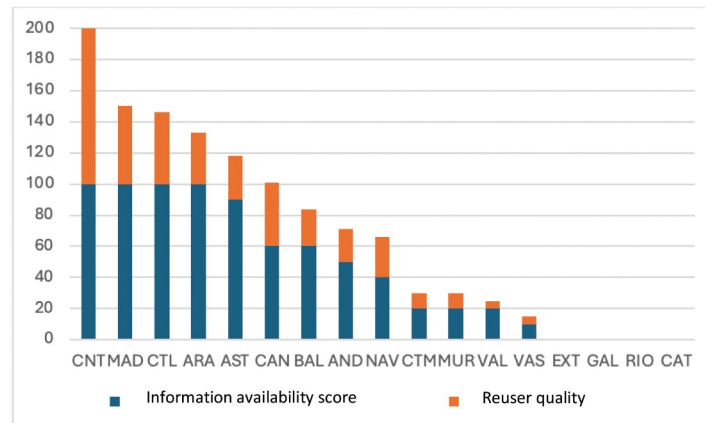
Table 4 shows the resulting scores, which range from 0 points—when sampling indicates that a region does not provide demographic regional datasets—to 200 points, when a region offers all ten sampled datasets without incurring any exclusion criterion, provides at least six formats, and includes at least one dataset reaching the four-star level.

Autonomous Communities	Initial Datasets	Not Excluded	Information availability score	Total Formats	Maximum Opening	Reuser quality	TOTAL
CNT	10	10	100	60	40	100	200
MAD	10	10	100	20	30	50	150
CTL	10	10	100	16	30	46	146
ARA	10	10	100	12	21	33	133
AST	10	9	90	10	18	28	118
CAN	10	6	60	23	18	41	101
BAL	10	6	60	6	18	24	84
AND	10	5	50	7	14	21	71
NAV	4	4	40	14	12	26	66
CTM	10	2	20	4	6	10	30
MUR	10	2	20	4	6	10	30
VAL	10	2	20	2	3	5	25
VAS	10	1	10	2	3	5	15
CAT	10	0	0	0	0	0	0
RIO	2	0	0	0	0	0	0
EXT	0	0	0	0	0	0	0
GAL	0	0	0	0	0	0	0
TOTAL	136	57	770	180	219	399	1,169

Note: AND: Andalusia, ARA: Aragon, AST: Principality of Asturias, BAL: Balearic Islands, CAT: Catalonia, CNR: Canary Islands, CNT: Cantabria, CTL: Castilla and Leon, CTM: Castilla-La Mancha, EXT: Extremadura, GAL: Galicia, MAD: Community of Madrid, MUR: Region of Murcia, NAV: Community Foral of Navarre, RIO: La Rioja, VAL: Valencian Community, VAS: Basque Country.

Table 4. Score obtained in the sampling of demographic data

Table 4 indicates that Cantabria is the leading region, achieving the maximum score. It is followed by Community of Madrid (75% of the maximum score), Castilla and Leon, Aragon, the Principality of Asturias, and the Canary Islands, all surpassing 50%. On the other hand, several regions do not reach 100 points. Notably, four regions receive zero points: Extremadura and Galicia, which have no tagged datasets, and La Rioja and Catalonia, which do not provide the information as datasets but instead redirect users to external links. Figure 1 graphically displays these results.



Note: AND: Andalusia, ARA: Aragon, AST: Principality of Asturias, BAL: Balearic Islands, CAT: Catalonia, CNR; Canary Islands, CNT: Cantabria, CTL: Castilla and Leon, CTM: Castilla-La Mancha, EXT: Extremadura, GAL: Galicia, MAD: Community of Madrid, MUR: Region of Murcia, NAV: Community Foral of Navarre, RIO: La Rioja, VAL: Valencian Community, VAS: Basque Country.

Figure 1. Ranking by regions

Figure 1 shows that, except for Cantabria, which achieves 100% of the maximum score in both dataset availability and reusability quality, the remaining regions display an imbalanced performance. Following Community of Madrid, where reusability represents 50% of dataset availability, no region meets that threshold, indicating that the number of formats and openness levels fall well short of Cantabria's offering.

Figure 2 displays the formats in which the datasets are made available. The most common format is CSV (Comma-Separated Values), a plain-text file format used to store and exchange tabular data where values are separated by a delimiter. This format is similar to XLS (Microsoft Excel spreadsheet), the third most common format, which—although also accessible to non-technical users—offers lower interoperability and is proprietary, placing it at a lower level in the five-star data hierarchy.

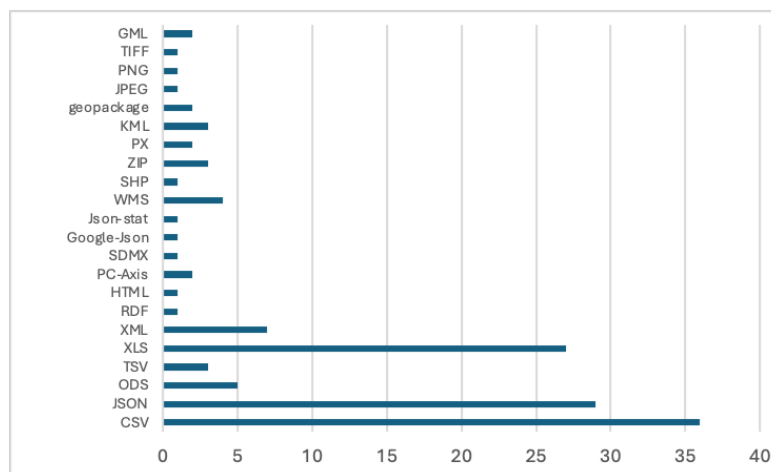


Figure 2. Formats of non-excluded demographic datasets in the sample

JSON (JavaScript Object Notation) files occupy an intermediate position between CSV and XLS. JSON is a lightweight text format for data interchange that uses simple key–value pairs and list structures, making it readable for humans and highly processable for machines.

Use of the remaining formats is significantly less frequent, particularly noteworthy in the cases of XML and RDF, two high-openness formats that facilitate automated processing and advanced reuse. Their limited use may hinder information processing by advanced users who require automated data handling and integration.

The results section concludes by examining correlations among several variables derived from the study (Table 5).

		Total Datasets	Demographics datasets	Information availability score	Reuser quality	Total Score
Total Datasets	Pearson Correlation	1	0.829**	0.085	0.082	0.087
	Sig. (two-tailed)		0	0.746	0.753	0.738
	N	17	17	17	17	17
Demographics datasets	Pearson Correlation	0.829**	1	0.121	0.076	0.107
	Sig. (two-tailed)	0		0.645	0.773	0.682
	N	17	17	17	17	17
Information availability score	Pearson Correlation	0.085	0.121	1	0.834**	0.974**
	Sig. (two-tailed)	0.746	0.645		0	0
	N	17	17	17	17	17
Reuser quality	Pearson Correlation	0.082	0.076	0.834**	1	0.937**
	Sig. (two-tailed)	0.753	0.773	0		0
	N	17	17	17	17	17
Total Score	Pearson Correlation	0.087	0.107	0.974**	0.937**	1
	Sig. (two-tailed)	0.738	0.682	0	0	
	N	17	17	17	17	17

Table 5. Correlations between variables

Regarding positive associations, the most notable correlation (0.829) is between the total number of datasets offered by each region and the number of datasets tagged as demographic. This confirms the pivotal role of this category, the largest among the 22 defined by the Technical Interoperability Standard. A positive correlation also exists between total score and dataset availability and reusability quality, since both are components of the overall score.

However, Table 5 also reveals one of the study's key findings: the absence of a relationship (correlation values near zero) between each region's sampling score and the total amount of information available on its portal—whether overall or demographic. This suggests that dataset volume is not a determining factor, and that much of the information supplied suffers from at least one of the exclusion criteria (non-demographic information, non-regional scope, or information not provided as datasets).

4. Proposed Improvements and Usefulness of Regional Open Demographic Data

This section provides a summary of the information tagged under the demography category in Spain's Autonomous Communities (from highest to lowest number of tagged datasets, as shown in Table 3).

4.1. Analysis of Open Demographic Data by Region

4.1.1. Balearic Islands

As noted earlier, the increase in the Balearic Islands has been exponential, rising from 6 datasets in 2024 (some of them referring to specific years, such as *Census Sections 2017* or *Census Districts 2017*) to nearly 8,000 datasets tagged as demographic.

Summarizing such an enormous volume of information presents a major challenge. This difficulty is intensified by the fact that many datasets are tagged in up to nine categories simultaneously. As a result, most of them do not relate to the expected content. Examples include *Measurement Concepts of Collective Agreements* or *Measurement Concept Scheme for the Publication of Quality Wine Data of the Balearic Islands*.

Although some datasets are accurately tagged—such as population reports by marital status or sex, register data, censuses, immigration data, migratory balances, and related indicators—these represent only a small fraction of the total. Even using the portal's search engine, locating them is difficult.

4.1.2. Canary Islands

The Canary Islands occupy the second position in the supply of demographic information. Although the region has approximately 3,000 datasets behind the Balearic Islands, it also surpasses the third region by a similar margin.

Between 2024 and 2025, its dataset offering increased by 179 datasets (3.80%), reaching 4,884 datasets. The most common information refers to population data, heavily segmented by age groups, districts, counties, etc., and complemented by birth and death data, natural increase, population density, household information, and demographic indicators.

Given the large dataset offering, a considerable number of non-demographic datasets were expected—and indeed found—such as *Percentage of Schools with Internet Access by School Ownership and Type of Connection* or *Classification of Housing and/or Environment Problems*, among others.

4.1.3. Basque Country

The Basque Country, the third region in the dissemination of population data, increased its offering by 248 datasets (18% more than in 2024), reaching a total of 1,633 datasets in 2025.

However, many datasets are unrelated to demography—such as *Bilbao Street Names*—or diverge significantly from the theme, such as the more than 250 datasets on municipal sustainability indicators (urban planning, energy, facilities, water, and air). This appears to stem from the extensive use of multiple classifications (up to five categories in some sampled datasets). Large amounts of non-regional information, particularly from provincial councils or municipalities, were also observed.

These issues hinder the identification of regional demographic datasets containing the expected content—population and inhabitants, marriages, separations, divorces, deaths, births, and similar indicators.

4.1.4. Principality of Asturias

The Principality of Asturias has reduced its offering by almost half (from 803 to 419 datasets) but still ranks as the fourth-largest regional supplier of demographic datasets.

Population data are the most frequent content (300 datasets), segmented by sex, age, and place of birth, usually referring to specific years. Additional datasets include those on population structure (28 datasets), key demographic indicators (23 datasets), and data on inactive, employed, or unemployed populations (18 datasets), along with migration flows, nomenclature, and related information.

Some datasets appear to be incorrectly tagged, such as registers of cooperatives, administrative units, buildings in the Principality of Asturias, or registers of labor societies—also tagged under industry or employment.

4.1.5. Community of Madrid

In 2024, the Community of Madrid offered 12 datasets (2 municipal and 10 regional). In 2025, it became the third region in the growth of demographic datasets, reaching 226. This improvement is associated with a complete redesign of the initiative and a modern, user-friendly portal that provides information in table, text, JSON, analytical, dictionary, API, or downloadable formats (CSV added to the existing JSON option).

The Community of Madrid offers a diverse set of demographic datasets, including: 15 on nationality acquisition; 13 on deaths; 10 on foreign residents; 15 on marriages; 27 on women; 10 on births; 13 on population registers; 38 on population; 22 on population projections; and 10 on various rates (fertility, marriage, mortality, etc.).

4.1.6. Aragon

Aragon ranks sixth in the dissemination of population data in 2025 (and fifth in dataset growth), increasing from 48 datasets in 2024 to 162 in 2025.

Its portal offers two large information series: Population Pyramids (21 datasets, 1998–2020, except for 1999) and Municipal Population Register as of January 1 (7 datasets, 2004–2010).

Other information includes datasets on births and deaths (3 each), age and rates at the regional and provincial levels (20 datasets), county and provincial levels (17 datasets), municipality level (10 datasets), the register of Spanish residents abroad (6 datasets), foreign residents (16 datasets), immigration and emigration flows (13 each), population indicators (16 datasets), various migration balances (11 datasets), and several additional datasets.

4.1.7. Region of Murcia

The region of Murcia shows a slight decrease, from 97 to 95 datasets. Most datasets do not present regional aggregation but rather come from specific municipalities, such as Lorca (43) or Molina de Segura (11). Only a few are provided by the regional government.

Examples of region-level datasets include *Evolution of Migrations*, *Evolution of Births*, *Evolution of Deaths*, *Evolution of Marriages*, *Demographic Birth Indicators*, and *Evolution of Migration Balances with Other Provinces*.

4.1.8. Andalusia

Andalusia increases its offering from 31 datasets in 2024 to 41 in 2025. It provides comprehensive, non-disaggregated datasets on fertility rates, the population register, population by province, sex, and age, as well as natural population changes, annual reports, statistical censuses, and nomenclature.

However, several non-demographic datasets are also found, such as the *Multiterritorial Information System of Andalusia (SIMA)*, *Sustainable Development Indicators for the 2030 Agenda*, *Employment Agenda Monitoring Indicators*, and *Quarterly Indicators of Economic Activity of the Andalusian Population (ITAEPA)*.

4.1.9. Valencian Community

The Valencian Community holds a mid-level national position with 41 datasets in 2025 (3 more than in 2024).

The most recurrent dataset is the *Continuous Municipal Register by Districts and Sections*, with one dataset per year from 1999 to 2022 (23 datasets). Additional datasets include housing censuses (2001, 2011, 2021), immigration data, population and household estimates, projections, residential variations, and related statistics.

However, most datasets are not true datasets but links to the Valencian Institute of Statistics, where data can be accessed through multi-field search tools.

4.1.10. Catalonia

Catalonia provides 33 datasets in both 2024 and 2025, with comprehensive but non-disaggregated information.

Content includes the census of population and housing, household surveys, death, marriage, and birth surveys, population estimates, migration data, population registers, and projection datasets.

Although most datasets fit the category well, occasional unrelated datasets appear (e.g., *Elections to the National Congress*). A major limitation is that all sampled datasets are redirected to an external link (Catalonia's statistical office), which uses configurable extractors rather than actual datasets.

4.1.11. Castilla and Leon

Castilla and Leon and Catalonia are the regions that offer the same number of demographic datasets in both years: 25.

Two information types are identified: classic population data (population by province and sex, births, deaths, marriages, migrations, demographic indicators, etc.) and non-population content related to category-appropriate

themes (e.g., childcare center registers, family meeting point programs), all aligned with the Technical Interoperability Standard categories related to demography.

4.1.12. Cantabria

Cantabria increases its range of offers from 15 to 24 datasets. These include municipal census changes, annual population and housing censuses, population, household, and immigrant surveys, migration statistics, demographic indicators, natural population movement data, population registers, and census-level population.

All datasets are offered in nine different formats, with an identical set of formats available for each dataset.

4.1.13. Castilla-La Mancha

Castilla-La Mancha increases its range slightly, from 9 to 10 datasets. However, extensive multiple categorization results in many datasets unrelated to demography, such as territorial integration areas, municipal entities, geographical data portals, or territorial extensions.

Only three datasets contain demographic content, *Municipal Population Registers*, *Population Projections*, and *Population Statistics*, but none are actual datasets, as they redirect to external links.

4.1.14. Community Foral of Navarre

The Community Foral of Navarre's portal contains nearly 2,000 datasets in total, but demographic offerings are extremely limited, decreasing from 6 datasets in 2024 to 4 in 2025.

It retains datasets on housing applicant registers and health-area population data but replaces three gender-indicator datasets with a *Nomenclator* dataset listing municipalities and population entities.

4.1.15. La Rioja

La Rioja offered 18 datasets in 2024 (emigration, immigration, population register data, etc.), but this number drops to 2 in 2025: *Births, Deaths, and Marriages* and *Population Censuses*. Neither is a true dataset, as both redirect to external links at the regional statistics institute.

With La Rioja, the review of regional demographic datasets concludes, as Extremadura and Galicia do not provide any datasets in this category.

4.2. Proposal for Improvement

Previous studies, such as that by Curto-Rodríguez and Marcos-Sánchez (2024), found that open data in general, and demographic information available in 2024, varied across regions. In some cases, information on variables such as gender, age groups or place of residence was disaggregated, whilst in others it was aggregated, highlighting the lack of standardisation. Although the number of datasets has doubled during the period analyzed (2024-2025), both overall and within the demography category, the errors detected highlight the need for improvements to ensure that these initiatives can generate social and economic value.

The analysis shows a significant asymmetry among regional open data initiatives: the Balearic Islands, the Canary Islands, and the Basque Country account for 92% of all demographic information tagged at the regional level. Heterogeneity is also evident in terms of where information is hosted (fifteen regions use their own portals, while Principality of Asturias and Extremadura rely exclusively on the national portal); in the partial adoption of the twenty-two categories defined by the Technical Interoperability Standard (followed consistently by only fourteen regions); and in the widespread use of multiple category assignments for individual datasets.

This issue is particularly concerning because much of the information supplied does not relate to demography, as confirmed by the correlation analyses. Additionally, the segmentation of datasets results in information overload in highly populated portals. For example, entering the term “population” in the Canary Islands search tool returns 4,416 datasets, including entries such as *Employed population by sex and type of working day. Canary Islands. Third quarter 2025*. This, combined with the absence of a uniform standard for structuring information, makes dataset integration difficult.

Based on the study's findings, several improvement proposals are outlined below.

First, the criteria should be standardized:

- Adopt a single tagging scheme.
- Establish a minimum dataset aligned with FEMP 80 or with the EU's list of high-value datasets.
- Harmonize dataset specifications and designs to facilitate interoperability and combination.

Second, information should be provided as actual datasets, since, by definition, an open data portal must contain open datasets (not external links to search engines). Open data are essential for artificial intelligence applications, and appropriate metadata (title, description, license, standards, etc.) must also be supplied.

Third, the quality of formats offered should be strengthened, with greater availability of XML or RDF files. CSV files are useful and widely accepted, but they should be accompanied by formats with higher levels of openness and reusability.

Finally, although the dissemination of open datasets from certain municipalities is positive, their natural location is the respective local portals. Regional open data portals should provide information at the regional level of aggregation. If a local entity does not have its own portal, two separate sections should be created within the regional portal.

Table 6 summarizes the most relevant recommendations for each region based on the study's findings.

Andalusia	Apply a single tagging scheme
Aragon	Avoid excessive segmentation of information
Balearic Islands	Apply a single tagging scheme – Avoid excessive segmentation of information
Basque Country	Apply a single tagging scheme – Avoid excessive segmentation of information – Provide local information separately
Canary Islands	Apply a single tagging scheme – Avoid excessive segmentation of information
Cantabria	
Castilla and Leon	
Castilla-La Mancha	Apply a single tagging scheme – Increase the amount of information offered
Catalonia	Provide datasets rather than external links
Community Foral of Navarre	Increase the amount of information offered
Community of Madrid	
Extremadura	Increase the amount of information offered (none currently provided)
Galicia	Increase the amount of information offered (none currently provided)
La Rioja	Increase the amount of information offered – Provide datasets rather external links
Principality of Asturias	Apply a single tagging scheme
Region of Murcia	Provide local information separately
Valencian Community	Provide datasets rather than external links

Table 6. Summary of recommendations for the Autonomous Communities

4.3. Usefulness of Regional Open Demographic Data

Regional open demographic data can support multiple practical applications. From a business and marketing perspective, such data enable market segmentation, customer profiling, and location-based decision-making, as well as enriching business intelligence systems with external, authoritative information not available in internal databases. From a public and social perspective, demographic open data support evidence-based policymaking, territorial planning, and citizens' decisions such as choosing a place of residence.

However, for open data to generate real practical value, its use must be actively facilitated. Based on the empirical findings of this study, the following recommendations are proposed. They are organised into four complementary categories, each specifying concrete actions and their contribution to enhancing the usability and reuse of regional open demographic data.

4.3.1. Standardisation of Data Structure and Level of Disaggregation

A first critical step toward practical usability is agreement on the appropriate level of data disaggregation within the demography category. Concrete actions include:

- Defining a common baseline of disaggregation criteria (e.g., sex, age groups, municipality/province) applicable across regions.
- Avoiding excessive dataset fragmentation that produces thousands of narrowly segmented datasets, while also preventing overly aggregated data that restricts analytical use.
- Aligning dataset structures with widely adopted standards, such as FEMP recommendations or EU high-value dataset guidelines.

These measures reduce user uncertainty about expected content, facilitate comparison across regions, and improve interoperability for analytical and business uses.

4.3.2. Enrichment and Harmonisation of Metadata

Metadata quality is a key determinant of whether datasets can be understood and reused effectively. Recommended actions include:

- Enriching metadata beyond basic elements (title, description, license) to clearly specify territorial coverage, temporal coverage, update frequency, and level of aggregation.
- Harmonising metadata structures across regions to enable cross-regional comparison and dataset combination.
- Introducing quality classification mechanisms (e.g., basic, medium, good, excellent), inspired by initiatives such as Open Data Certificates (ODI), based on criteria such as completeness, clarity of documentation, update regularity, and formats offered.

Clear and harmonised metadata lower entry barriers for reusers and increase trust in the datasets, thereby enhancing their practical value.

4.3.3. Enhancement of Data Accessibility Through Visualisation and Interaction Tools

To support non-technical users and broaden reuse, open demographic data should be accompanied by accessible exploration tools. Concrete actions include:

- Integrating interactive dashboards, dynamic cartographic maps, automated population pyramids, and visual time-series analysis tools into regional portals.
- Supporting exploratory analysis through interfaces that do not require advanced technical skills.
- Learning from existing initiatives such as Open Data Barcelona or international platforms like DataUSA.

These tools transform datasets from static files into usable resources, facilitating understanding and decision-making for businesses, citizens, and public managers.

4.3.4. Technical Enablement for Advanced Reuse and Integration

For advanced analytical and professional reuse, technical access mechanisms must be strengthened. Recommended actions include:

- Developing homogeneous and well-documented APIs across regions, particularly for high-demand use cases such as population projections and mobility analysis.

- Documenting APIs using OpenAPI standards, including practical usage examples in R, Python, or Excel.
- Enabling parametric queries (e.g., by age, year, municipality, sex) to support flexible data extraction and automation.

5. Conclusions

This study has addressed the two research questions posed. First, it analyzed regional demographic datasets and their characteristics. The findings indicate a predominance of datasets categorized as demographic (more than 15,000), although contributions from the regions are highly unbalanced. The Balearic Islands, the Canary Islands, and the Basque Country together account for more than 90% of all demographic datasets (51%, 31%, and 10%, respectively). The recent study by García-Rodríguez and Mahou-Lago (2025), which examines open data initiatives across Spain's Autonomous Communities, also highlights uneven development, reporting an average maturity level of 65% in a context where some regions are clearly advanced, and others face substantial shortcomings. Frequent weaknesses include limited effective reuse, insufficient metadata quality, and the lack of advanced usability tools. The results of the present study are consistent with prior research, such as Curto-Rodríguez and Marcos-Sánchez (2024), both in the heterogeneity observed and in the conclusion that regional open data initiatives still have substantial room for improvement.

It is also important to note that many regions—particularly those with the highest number of demographic datasets—exhibit errors in their publication practices. An analysis of a sample of dataset shows that Cantabria achieves 100% in both dataset availability and reusability. Community of Madrid, Castilla and Leon, and Aragon make all information available, but with lower reusability quality, although all exceed 120 out of the 200 possible points. In contrast, eleven regions do not surpass 100 points, and four obtain a score of zero. The main issues identified include deficient tagging, excessive dataset fragmentation, a lack of homogeneity and interoperability across regional initiatives, and the need to enforce formats that enable greater reuse.

These findings regarding the persistent heterogeneity in regional initiatives are particularly relevant for policymakers. These disparities could be reduced through more targeted interventions that enhance the overall impact of open data ecosystems by making information easier to combine. Therefore, the suggestions put forward in Curto-Rodríguez, Marcos-Sánchez et al. (2025) regarding the necessary unification of interpretation criteria for the technical interoperability standard on tagging, the lack of standardization in content, and the provision of an adequate range of formats (both in quality and quantity) to meet the needs of different types of users are considered appropriate. Meetings such as those held within the Spanish Federation of Municipalities and Provinces, which have already established a minimum of 80 datasets that must be disclosed, along with a series of recommendations (FEMP, 2024), need to take place at the regional level. On the other hand, the systematic review conducted by Curto-Rodríguez et al. (2026) shows that the literature identifies another set of obstacles that go beyond the lack of standardization, such as poor information quality and the vague definition of reuse conditions, which should also be addressed by regional authorities.

To address the second research question, a detailed regional assessment was conducted concerning the availability and usefulness of open demographic data. The Balearic Islands, the Canary Islands, the Basque Country, the Principality of Asturias, Andalusia, and Castilla-La Mancha have substantial numbers of datasets that are not genuinely demographic. Other regions, such as Region of Murcia, include datasets that are not regional, with more than 50% coming from two municipalities. Furthermore, Castilla-La Mancha, the Valencian Community, and Catalonia do not provide actual datasets but instead redirect users to external sites to generate the data.

The study offers several theoretical contributions. It strengthens the understanding of the open data phenomenon, an area where considerable analysis remains to be done. In decentralized regional systems such as Spain's, development is uneven, an observation also reflected in recent research by Curto-Rodríguez and Marcos-Sánchez (2024), Curto-Rodríguez (2025) and Curto-Rodríguez, Marcos-Sánchez et al. (2025). The study additionally provides a simple tool that quantifies the information supplied, cleans it, and evaluates its reusability quality (in terms of breadth and openness level).

The study's practical implications are also noteworthy. For public managers, it offers a concrete diagnosis of the main issues detected and identifies good practices that could be transferred across regions. For data reusers, it provides a clear summary of the demographic information available and its potential usefulness.

The study is subject to limitations, particularly those related to the sampling performed. The fact is that determining the sample size is key to obtaining information that is fully reliable and representative of the target population (Trespacios-Gutiérrez et al., 2016). Therefore, although the study has sampled a significant volume of data by distinguishing between two groups of autonomous communities (with more or fewer than 10 population datasets), it remains an approximation that could have been combined with stratified random sampling with proportional allocation. Future research could expand the sample size, enabling more robust extrapolations.

Another limitation is that, although the conclusions are drawn from an analysis of the most extensive category of information within a country that is a leader in open data provision, it remains a case study. That is why further work must be conducted on other categories of information or levels of government to allow for the comparison of results, to help generate a broader view of the open data phenomenon in that country, and to conduct comparative assessments with the performance levels of other countries.

A possibility the study could have explored is the use of other indicators to compare the results. Several proposals exist for evaluating the usability of open data portals based on discoverability, accessibility, and data reusability, such as that by Máchová et al. (2018). Other tools, such as Meloda, aim to assess the reusability potential of open data (Abella et al., 2014; Abella, Ortiz-de-Urbina-Criado & De-Pablos-Heredero, 2019), and have been used in academic studies and reports on open data in Spain (Abella et al., 2017, 2021, 2024; Abella, Ortiz-de-Urbina-Criado, De-Pablos-Heredero et al., 2019).

A further line of research would involve studying the measurement of reuse impact as a proxy for the value of information. These measurements include the number of downloads, the volume of API requests, the number of significant public or private reuse projects developed, and academic or institutional references to the datasets. Increases in these metrics reflect greater value generated by the information provided.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This paper has been supported by Projects 'OpenInnova', High-Performance Research Group (V1717, Rey Juan Carlos University, Spain), and PHS-2024/PH-HUM-294 (Madrid Regional Government, Spain).

Authors' contributions

Ricardo Curto-Rodríguez: conceptualization, data processing, research, formal analysis, methodology, writing original draft, review & editing.

Marta Ortiz-de-Urbina-Criado: data processing, formal analysis, methodology, writing original draft, review & editing.

Carmen-de-Pablos-Heredero: formal analysis, methodology, writing original draft, review & editing, acquisition of funds.

Rafael Marcos-Sánchez: data processing, research, formal analysis, methodology, writing original draft, review & editing.

Data availability

Data available upon request.

Use of Artificial Intelligence

The authors declare that they wrote this article without using artificial intelligence for the writing of the text, data analysis, the creation of graphs, or any of the other activities carried out.

References

- Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2014). Meloda, métrica para evaluar la reutilización de datos abiertos. *El Profesional de la Información*, 23(6), 582–588. <https://doi.org/10.3145/epi.2014.nov.04>
- Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2017). *La reutilización de datos abiertos: Una oportunidad para España*. Fundación Cotec. <https://hdl.handle.net/10115/41109>
- Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2019). Meloda 5: A metric to assess open data reusability. *El Profesional de la Información*, 28(6), e280620. <https://doi.org/10.3145/epi.2019.nov.20>
- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C., Vidal-Cabo, C., & Ferrer-Sapena, A. (2019). *La reutilización de datos abiertos en España II*. DesideDatum. <https://hdl.handle.net/10115/41108>
- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C., & García-Luna, D. (2021). *La reutilización de datos abiertos III. Octubre 2021*. ESIC Editorial. <https://hdl.handle.net/10115/41106>
- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C., & García-Luna, D. (2024). *La reutilización de los datos abiertos en España. Actualización 2024. Informe IV octubre 2023*. Fundación Cotec. <https://hdl.handle.net/10115/41110>
- Alcaide-Muñoz, L., Rodríguez-Bolívar, M. P., & Villamayor-Arellano, C. L. (2022). Factors in the adoption of open government initiatives in Spanish local governments. *Government Information Quarterly*, 39(4), 101743. <https://doi.org/10.1016/j.giq.2022.101743>
- ASEDIE (2025). *Economía del dato en su ámbito infomediario* (13.^a ed.) (Informe). <https://www.asedie.es/es/informes-anuales>
- Barrientos-Matute, J. P., & Ruvalcava-Gómez, E. A. (2023). Metodología para explorar el valor público de datos abiertos ofertados en portales web de vivienda. *Economía, Sociedad y Territorio*, 23(72), 405–431. <https://doi.org/10.22136/est20232002>
- Berners-Lee, T. (2010). *5-star Open Data*. <http://5stardata.info/en/>
- Cerrillo-Martínez, A., & Casadesús-de-Mingo, A. (2021). Data governance for public transparency. *El Profesional de la Información*, 30(4). <https://doi.org/10.3145/epi.2021.jul.02>
- Corrales-Garay, D., Ortiz-de-Urbina-Criado, M., & Mora-Valentín, E. M. (2019). Knowledge areas, themes and future research on open data: A co-word analysis. *Government Information Quarterly*, 36(1), 77–87. <https://doi.org/10.1016/j.giq.2018.10.008>
- Criado-Grande, I. (2013). Redes sociales para unas administraciones y gobiernos abiertos. Desafíos para la innovación y la creación de valor público. *Revista Buen Gobierno*, 15, 8–25. <https://www.redalyc.org/articulo.oa?id=569660536001>
- Cruz-Meléndez, C., & Valencia-López, O. D. (2022). Active transparency during the Covid-19 pandemic. *Revista Española de la Transparencia*, 15, 279–311. <https://doi.org/10.51915/ret.205>
- Curto-Rodríguez, R. (2021). Análisis multidimensional de los portales de datos abiertos autonómicos españoles. *Revista Española de Documentación Científica*, 44(1), e284. <https://doi.org/10.3989/redc.2021.1.1745>
- Curto-Rodríguez, R. (2025). Datos abiertos con información de comercio y consumo: un análisis exploratorio a nivel autonómico español. *Revista Española de Documentación Científica*, 48(2), 1675. <https://doi.org/10.3989/redc.2025.2.1675>

- Curto-Rodríguez, R., & Marcos-Sánchez, R. (2024). Datos abiertos para la creación de valor: atención e interés para el marketing [Open data for value creation: attention and interest for marketing]. *European Public & Social Innovation Review*, 9, 1–16. <https://doi.org/10.31637/epsir-2024-1195>
- Curto-Rodríguez, R., Marcos-Sánchez, R., & Ferrández, D. (2025). Content, formats and licensing of datasets from autonomous communities: Value generation, sustainability and proposals for improvement. *International Journal of Information Management Data Insights*, 5(2), 100369. <https://doi.org/10.1016/j.jime.2025.100369>
- Curto-Rodríguez, R., Marcos-Sánchez, R., Zaragoza-Benzal, A., & Ferrández, D. (2025). Open Data to Promote the Economic and Commercial Development of the Housing Sector: The Case of Spain. *Urban Science*, 9(7), 277. <https://doi.org/10.3390/urbansci9070277>
- Curto-Rodríguez, R., Leal-Matilla, A., Ferrández, D., & Marcos-Sánchez, R. (2026). Open Data Research in Spain Published via the Diamond Route: A Systematic Review. *Publications*, 14(1), 16. <https://doi.org/10.3390/publications14010016>
- FEMP (2024). *80 conjuntos de datos a publicar*. https://redtransparenciayparticipacion.es/?get_group_doc=14/1687500714-80conjuntosdedatosapublicar.pdf
- Ferrer-Sapena, A., Peset, F., & Aleixandre-Benavent, R. (2011). Acceso a los datos públicos y su reutilización: open data y open government. *El Profesional de la Información*, 20(3), 260–269. <https://doi.org/10.3145/epi.2011.may.03>
- Gao, Y., Janssen, M., & Zhang, C. (2021). Understanding the evolution of open government data research: towards open data sustainability and smartness. *International Review of Administrative Sciences*, 89(1), 59–75. <https://doi.org/10.1177/00208523211009955>
- García-García, J., & Curto-Rodríguez, R. (2019). El ejercicio de la rendición de cuentas mediante portales de datos abiertos en las comunidades autónomas españolas. *IDP. Revista de Internet, Derecho y Política*, 29. <https://doi.org/10.7238/idp.v0i29.3182>
- García-Rodríguez, C. A., & Mahou-Lago, X. M. (2025). Analysis of Data Openness by regions in Spain. *Gestión y Análisis de Políticas Públicas*, 38, 63–84. <https://doi.org/10.24965/gapp.11495>
- González-García, L., & García-García, J. (2022). Meaning of the concept Open Government in Spanish public administrations. *Revista Española de la Transparencia*, 15, 157–180. <https://doi.org/10.51915/ret.224>
- Hair, J. F., Harrison, D. E., & Ajjan, H. (2022). *Fundamentos de analítica de marketing*. McGraw-Hill Interamericana de España.
- Hofmann, A. (2015). Datos abiertos gubernamentales y gobierno abierto ¿Qué tan abierto es abierto? *Buen Gobierno*, 19, 65–84.
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014). Generating sustainable value from open data in a sharing society. In B. Bergvall-Kåreborn & P. A. Nielsen (Eds.), *Creating Value for All Through IT. TDIT 2014. IFIP Advances in Information and Communication Technology* (vol. 429, pp. 62–82). Springer. https://doi.org/10.1007/978-3-662-43459-8_5
- Kotler, P., & Armstrong, G. (2006). *Principles of Marketing* (11^a ed.). Prentice-Hall, Inc.
- Máchová, R., Hub, M., & Lněnička, M. (2018). Usability evaluation of open data portals: Evaluating data discoverability, accessibility, and reusability from a stakeholders' perspective. *Aslib Journal of Information Management*, 70(3), 252–268. <https://doi.org/10.1108/AJIM-02-2018-0026>
- Martínez-Gutiérrez, R. (2023). Datos abiertos, reutilización, Smart Cities y medición de la eficiencia en la contratación pública. *Boletín del Observatorio de Contratación Pública*. <http://hdl.handle.net/10045/141256>

- Martínez-Vázquez, J. (2014). La descentralización tributaria a las comunidades autónomas en España: Desafíos y soluciones. In J. Durán & A. Esteller-Moré (Eds.), *Por una verdadera reforma fiscal*. Editorial Ariel.
- Nikiforova, A., & McBride, K. (2021). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, 58, 101539. <https://doi.org/10.1016/j.tele.2020.101539>
- Open Government Working Group (2007). *Principles for open government data*. <https://doi.org/10.63485/6vfpf-yw667>
- Ortiz-de-Urbina-Criado, M., Abella, A., & García-Luna, D. (2023). Open data-set identifier for open innovation and knowledge management. *Journal of Knowledge Management*, 27(10), 2779–2796. <https://doi.org/10.1108/jkm-07-2022-0514>
- Othman-Alanaimy, A.A. (2023). The Legal System of Open Government. *Regional Studies Journal*, 17(57), 277–304. <https://doi.org/10.33899/regis.2023.179741>
- Page, M., Behrooz, A., & Presse, B.Y. (2025). *2025 Open Data Maturity Report*. Luxembourg: Publications Office of the European Union. <https://doi.org/10.2830/9978861>
- Pulido-Barreto, A. M. (2021). *Plan de apertura y estructuración de los datos abiertos de la alcaldía del municipio de Fusagasugá. [Magister en Gerencia de Sistemas de Información y Proyectos Tecnológicos]*. Universidad EAN.
- Putera, R. E., & Irwanto, E. L. (2023). The Urgency of Embodiment of Open Government as an Effort to Realize the Fulfillment of the Social, Political and Economic Rights of the Indonesian Society. *Publica: Jurnal Pemikiran Administrasi Negara*, 15(1), 56–63.
- Ramos-Simón, L., Arquero-Avilés, R., Botezan, I., Cobo-Serrano, S., Sala-Jiménez, A., Sánchez-Jiménez, R., & Valle-Gastaminza, F. (2012). De la reutilización de la información del sector público a los portales de datos abiertos en Europa. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, 29. <https://bid.ub.edu/29/ramos2.htm>
- Ramos-Simón, L. F. (2017). El uso de las licencias libres en los datos públicos abiertos. *Revista Española de Documentación Científica*, 40(3), e179–e179. <https://doi.org/10.3989/redc.2017.3.1376>
- Rodríguez-Pose, A., & Gill, N. (2003). The Global Trend towards Devolution and its Implications. *Environment and Planning C: Government and Policy*, 21(3), 333–351. <https://doi.org/10.1068/c0235>
- Royo-Montañés, S., & Benítez-Gómez, A. (2019). Portales de datos abiertos. Metodología de análisis y aplicación a municipios españoles. *El Profesional de la Información*, 28(6). <https://doi.org/10.3145/epi.2019.nov.09>
- Ruvalcaba-Gómez, E. A. (2020). Datos abiertos = Open Data. *EUNOMÍA. Revista en Cultura de la Legalidad*, 18, 327–334. <https://doi.org/10.20318/eunomia.2020.5280>
- Secretaría de Estado de Administraciones Públicas (2013). 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*, 54, pp 17045–17071. <https://www.boe.es/eli/es/res/2013/02/19/4>
- Trespalcios-Gutiérrez, J.A., Vázquez-Casielles, R., De-la-Ballina-Ballina, F.J., & Suárez-Vázquez, A. (2016). *Investigación de mercados: El valor de los estudios de mercado en la era del marketing digital*. Ediciones Paraninfo S.A.
- Van-den-Berg, A., Hoekstra, M., & van-Veenstra, A. F. (2025). Government data ecosystems for addressing societal challenges: Lessons from two cases in the Netherlands. *Data & Policy*, 7. <https://doi.org/10.1017/dap.2025.10018>
- Vidal-Cabo, C., Sánchez-Pérez, E. A., & Ferrer-Sapena, A. (2024). Perception and Reuse of Open Data in the Spanish University Teaching and Research Community. *Data*, 9(10), 117. <https://doi.org/10.3390/data9100117>

Wibowo, W. S., Fadhil, A., Sensuse, D. I., Lusa, S., Putro, P. A. W., & Yulfitri, A. (2023). Pinpointing Factors in the Success of Integrated Information System Toward Open Government Data Initiative: A Perspective from Employees. *International Journal of Advanced Computer Science and Applications*, 14(1).
<https://doi.org/10.14569/IJACSA.2023.0140111>

Zúñiga, F., Mora-Poveda, D. A., & Llerena-Llerena, W. V. (2023). El big data y su implicación en el marketing. *Revista de Comunicación de la SEECI*, 56, 302–321. <https://doi.org/10.15198/seeci.2023.56.e832>

Journal of Industrial Engineering and Management, 2026 (www.jiem.org)



Article's contents are provided on an Attribution-Non Commercial 4.0 Creative commons International License. Readers are allowed to copy, distribute and communicate article's contents, provided the author's and Journal of Industrial Engineering and Management's names are included. It must not be used for commercial purposes. To see the complete license contents, please visit <https://creativecommons.org/licenses/by-nc/4.0/>.