

Artículo de Investigación

Datos abiertos para la creación de valor: atención e interés para el marketing

Open data for value creation: attention and interest for marketing

Ricardo Curto Rodríguez¹: Universidad de Oviedo. España.

curtoricardo@uniovi.es

Rafael Marcos Sánchez: Universidad Internacional de La Rioja. España.

rafael.marcos@unir.net

Fecha de Recepción: 20/08/2024

Fecha de Aceptación: 27/09/2024

Fecha de Publicación: 10/10/2024

Cómo citar el artículo:

Curto-Rodríguez, R. y 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>

Resumen:

Introducción: La divulgación de información en manos del sector público mediante datos abiertos se ha consolidado a nivel mundial. Dentro de la diversa tipología de los datos abiertos, la información demográfica presenta un elevado interés para el marketing al constituir una variable útil para realizar segmentaciones de mercados. **Metodología:** Se procede a revisar los conjuntos de datos autonómicos en abril de 2024 observando que, de los casi 40 000 conjuntos de datos existentes, aproximadamente 7 000 presentan la etiqueta "Demografía". **Resultados:** En este apartado muestra una aproximación al contenido de datos demográficos a nivel autonómico. **Discusión:** Se aprecian diferencias inter autonómicas muy importantes tanto en el número de *datasets*, porcentaje respecto al total autonómico, información contenida e incluso en la formulación de los conjuntos de datos. **Conclusiones:** A pesar de que existen datos poblacionales relevantes, la heterogeneidad observada dificulta su tratamiento automatizado y, por tanto, la creación de valor bien sea por empresas infomediarias como por institutos de investigación de mercados.

Palabras clave: España; comunidades autónomas; datos abiertos; marketing; creación de valor; portales de datos abiertos; datos poblacionales; información pública.

¹ Autor Correspondiente: Ricardo Curto Rodríguez. Universidad Internacional de La Rioja (España).

Abstract:

Introduction: The offer of public sector information through open data has been consolidated worldwide. Within the diverse typology of open data, demographic information is of high interest for marketing as it constitutes a useful variable for market segmentation. **Methodology:** The regional datasets are reviewed in April 2024, observing that, of the almost 40,000 existing data sets, approximately 7,000 have the “Demography” label. **Results:** This section offers a summary of the autonomous content of demographic data. **Discussion:** Very important inter-autonomous differences are seen both in the number of datasets, percentage with respect to the autonomous total, information contained and in the formulation of the data sets. **Conclusions:** Although there is relevant data, the observed heterogeneity hinders the automated processing of information and, therefore, the creation of value, either by infomediary companies or by market research institutes.

Keywords: Spain; autonomous communities; open data; marketing; value creation; open data portals; population data; public information.

1. Introducción

Las Administraciones producen y gestionan una gran cantidad de información que hacen pública para facilitar su acceso y su uso denominada datos gubernamentales abiertos (Rangel-Carrillo *et al.*, 2020). Estos datos abiertos u *open data* son parte fundamental de las políticas de gobierno abierto (Díez-Garrido y Melero, 2022) ya que son capaces de facilitar la solución de un amplio conjunto de problemas públicos (Pozen, 2020), de generar valor (Curto-Rodríguez, 2017) y de fomentar el desarrollo de las *smart cities* (Lnenicka *et al.*, 2022).

Vivimos en la era de los datos (Cerrillo-Martínez, 2018) donde los *open data*, de cómoda reutilización por ser fácilmente localizables y no estar sometidos a restricciones técnicas o jurídicas (Cerrillo-Martínez, 2014), tienen gran potencial al poder ser aprovechados por cualquier interesado, bien sea mediante la elaboración de aplicaciones o mediante análisis *database* según los postulados del *crowdsourcing* indicados en Surowiecki (2005).

Estos datos abiertos suelen ofrecerse de forma proactiva y voluntaria a través de plataformas en línea (Mabillard y Keuffer, 2020) y generalmente a través de portales de datos abiertos u *open data portals*. En el mundo existen más de tres mil portales de *open data*, siendo España líder con 279 portales de datos abiertos (Abella *et al.*, 2022).

Esta afirmación puede verse contrastada con las posiciones alcanzadas en varios estudios internacionales. En primer lugar nos referimos al ranking realizado por la Organización para la Cooperación y el Desarrollo Económico denominado OURData (*Open-Useful-Reusable Data Index*) donde España ha obtenido posiciones relevantes en 2017 (sexto puesto), 2019 (octavo lugar), avanzando hasta la quinta posición en 2023. En segundo lugar, hablamos del *Open Data Maturity* realizado por el portal europeo de datos abiertos para el periodo 2015-2023, en donde España ha ocupado la segunda posición en el periodo 2017-2020, retrocediendo levemente en 2021 (tercer puesto), en 2022 (séptimo puesto) y en 2023 (quinto puesto).

Los datos abiertos ofrecen hasta diez utilidades: impulsar el crecimiento económico, favorecer la rendición de cuentas por parte de las instituciones públicas, mejorar los servicios públicos, aumentar los ingresos públicos, mejorar el sector sanitario, fomentar la creación de empleo y reducir el coste de la mano de obra, aumentar la eficiencia energética, mejorar la investigación, crear utilidad en el sector geoespacial y facilitar la justicia social (Verhlust *et al.*, 2020). Con utilidades tan diversas no es de extrañar que la tipología de los datos abiertos sea muy numerosa, de manera que clasificaciones como la del portal español nacional distinga 22

categorías según la propuesta de la 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 Información² (figura 1).

Figura 1. Categorías de los datos abiertos.



Fuente: Ministerio para la transformación digital y de la función pública (2021).
<https://datos.gob.es/es/noticia/conoces-los-sectoriales-de-datosgobes>

El contenido de cada categoría viene definido en el anexo IV de la citada Norma Técnica de Interoperabilidad de Reutilización de Recursos de la Información que se refiere a la taxonomía³ de sectores primarios donde se especifican los temas relacionados de cada uno:

Ciencia y tecnología:

Incluye: Innovación, Investigación, I+D+i, Telecomunicaciones, Internet y Sociedad de la Información

Comercio:

Incluye: Consumo.

Cultura y ocio:

Incluye: Tiempo libre.

Demografía:

Incluye: Inmigración y Emigración, Familia, Mujeres, Infancia, Mayores, Padrón.

Deporte:

Incluye: Instalaciones deportivas, Federaciones, Competiciones.

Economía:

Incluye: Deuda, Moneda y Banca y finanzas.

Educación:

² [www.boe.es/eli/es/res/2013/02/19/\(4\)](http://www.boe.es/eli/es/res/2013/02/19/(4))

³ La clasificación, según cita la propia norma, ha sido elaborada en base al documento “propuesta de taxonomía común para los procedimientos y servicios electrónicos en el marco de la ley 11/2007”

Incluye: Formación.

Empleo:

Incluye: Trabajo, Mercado laboral.

Energía:

Incluye: Fuentes renovables.

Hacienda:

Incluye: Impuestos.

Industria:

Incluye: Minería.

Legislación y justicia:

Incluye: Registros.

Medio ambiente:

Incluye: Meteorología, Geografía, Conservación fauna y flora.

Medio Rural:

Incluye: Agricultura, Ganadería, Pesca y Silvicultura.

Salud:

Incluye: Sanidad.

Sector público:

Incluye: Presupuestos, Organigrama institucional, Legislación interna, Función pública.

Seguridad:

Incluye: Protección civil, Defensa.

Sociedad y bienestar:

Incluye: Participación ciudadana, Marginación, Envejecimiento Activo, Autonomía personal y Dependencia, Invalidez, Jubilación, Seguros y Pensiones, Prestaciones y Subvenciones.

Transporte:

Incluye: Comunicaciones y Tráfico.

Turismo:

Incluye: Alojamientos, Hostelería, Gastronomía.

Urbanismo e infraestructuras:

Incluye: Saneamiento público, Construcción (infraestructuras, equipamientos públicos).

Vivienda:

Incluye: Mercado inmobiliario, Construcción (viviendas).

En todo caso, lo normal en la literatura es la reagrupación de las tipologías de los datos abiertos en dos: fomento de la transparencia y creación de valor (Cruz-Rubio, 2005), de manera que unos estudios consideran la contribución de los datos para la rendición de cuentas de las Administraciones públicas, mientras que otros lo hacen para valorar la cuestión económica.

En este último objetivo señalado, es decir, la contribución de los datos abiertos para generar valor y riqueza económica es especialmente relevante el informe Vickery (2011), realizado por encargo de la Comisión Europea, que cuantifica los datos públicos de la Unión Europea para la ciudadanía y para las empresas infomediarias en 40.000 millones al año de manera directa (incrementando la cifra hasta los 140.000 millones/año si se extiende la consideración a la forma indirecta).

Dentro del mundo de la empresa existen distintos tipos de datos capaces de generar valor y en concreto, susceptibles de ser explotados por los sistemas de investigación comercial. Uno de ellos son los datos poblacionales que permiten identificar grupos de consumidores homogéneos y realizar estrategias de segmentación. Esta expresión de segmentación de mercado fue acuñada por primera vez en la obra *Product Differentiation and Market Segmentation as Alternative Marketing Strategies* (Smith, 1956) y es, junto con la personalización efectiva del producto, dos factores de gran influencia para aumentar la competitividad de los productos en los mercados en que la empresa compite (Sudirjo, 2023).

En esta acción de segmentar el mercado, definida por Kotler y Armstrong (2006) como la acción de dividir un mercado en grupos que pudieran presentar diferentes características, deseos o comportamientos, de manera que puedan necesitar diferentes productos o el empleo de campañas de marketing diferenciadas, comprobamos que las tecnologías y, en concreto, el *Big Data*, son de gran interés (Zuñiga *et al.*, 2023) en un escenario donde la Inteligencia artificial se convierte en una herramienta esencial para una segmentación efectiva (Calle García *et al.* 2024).

En base a la importancia que tiene la existencia de datos para la toma de decisiones en el ámbito del marketing, esta investigación tiene el objetivo de valorar la atención que los *open data* a nivel autonómico español prestan a los datos poblacionales, realizando, en este apartado, una primera aproximación su contenido, estructura y tipología.

2. Metodología

Como hemos señalado, el ámbito de este trabajo es el nivel autonómico español, un nivel administrativo de gran importancia ya que España es probablemente el estado más descentralizado de Europa occidental (Rodríguez-Pose y Gill, 2003). Por ello no es de extrañar que sus Comunidades autónomas dispongan de gran autonomía política (Fernández Leiceaga y Lago Peñas, 2013) y gestionen un importante volumen de recursos públicos, lo que le permite implementar iniciativas de datos abiertos.

Y es que estudios como el de Curto-Rodríguez (2020) muestran que el progreso de los datos abiertos a nivel autonómico en España ha sido imparable ya que, a los once portales en funcionamiento en 2013, se han ido añadiendo iniciativas hasta alcanzar los catorce repositorios activos en 2015 y un total de quince portales en 2017. El incremento ha continuado y ya en 2019 todas las comunidades autónomas contaban con portal *open data* operativo. La situación actual (abril de 2024) se ofrece en la tabla 1.

Tabla 1. Portales de datos abiertos autonómicos y datasets albergados en abril de 2024

Comunidad Autónoma	Nombre del Portal	Datasets
Andalucía	Datos Abiertos	712
Aragón	Aragón Open Data	2 713
Canarias	Datos Abiertos Canarias	14 237
Cantabria	Icane - datos	290
Castilla y León	Datos Abiertos de Castilla y León	744
Castilla-La Mancha	Datos abiertos de Castilla- La Mancha	341
Cataluña	Datos Abiertos GenCat	1.218
Comunidad de Madrid	Comunidad de Madrid datos abiertos	247
Com. Foral de Navarra	Open Data	1.561
Comunidad Valenciana	Dades obertes gva	1.453
Extremadura	Gobierno Abierto - Catálogo de datos	12
Galicia	abert@s	512
Islas Baleares	Dades obertes caib	453
La Rioja	Datos Abiertos Rioja	853
País Vasco	Opendataeuskadi	11.539
Principado de Asturias	Open Data del Principado de Asturias	1.788
Región de Murcia	Datos abiertos Región de Murcia	961
CONJUNTOS DE DATOS TOTALES		39 634

Fuente: Elaboración propia (2024).

A continuación, hemos procedido a realizar dos tablas que muestran los datos según las categorías propuestas por cada comunidad autónoma si bien es necesario señalar, como se expondrá posteriormente, que alguna autonomía no se ha adaptado totalmente a lo propuesto por la norma técnica. En primer lugar, se ofrecen la tabla 2 con información sobre las siete autonomías que superan la oferta total de 1.000 conjuntos de datos.

Tabla 2. Portales de datos abiertos con más de 1000 datasets. Resultados por categoría

Categorías / Autonomías	CNR	VAS	ARA	AST	NAV	VAL	CAT
Ciencia y Tecnología	478	287	68	22	27		17
Comercio	84	114	67	207	9		13
Cultura y Ocio	77	648	144	5	17	8	34
Demografía	4.705	1.385	48	803	6	38	33
Deporte	11	40	38	2	8		9
Economía	735	1.778	117	179	154	37	68
Educación	110	192	960	3	8	80	50
Empleo	427	902	280	121	11	160	0
Energía	4	62	3	9	4	16	17
Hacienda	12	4169	40	45	89	3	12
Industria	164	107	109	69	6	3	16
Legislación y Justicia	187	57	60	5	11	40	38
Medio Ambiente	235	858	177	165	553	677	45
Medio Rural	307	412	35	60	67	12	218
Salud	2.664	188	228	54	128	50	67
Sector Público	809	4389	164	81	518	51	158
Seguridad	60	132	23	1	3		112

Sociedad y Bienestar	2.316	1.454	49	99	101	10	106
Transporte	174	454	16	25	238		66
Turismo	1.716	409	48	110	17	10	15
Urbanismo e infraestructuras	140	572	23	28	429	230	62
Vivienda	143	554	16	28	22	28	7

Nota: CNR, Islas Canarias; VAS, País Vasco, ARA, Aragón; NAV, Comunidad Foral de Navarra; VAL, Comunidad Valenciana; CAT, Cataluña.

Fuente: Elaboración propia (2024).

Como muestra la tabla 2, todas las autonomías cuya oferta supera los 1.000 conjuntos de datos ofrecen información demográfica, si bien con diferencias apreciables entre Canarias con más de 4.000 *datasets* y la Comunidad Foral Navarra con tan sólo 6.

La tabla 3, por su parte, ofrece el detalle de las comunidades autónomas que no llegan a suministrar el millar de conjuntos de datos.

Tabla 3. Portales de datos abiertos con menos de 1000 *datasets*. Resultados por categoría

Categorías/ Autonomías	MUR	RIO	CTL	AND	GAL	BAL	CTM	CNT	MAD	EXT
Ciencia y Tecnología	21	6	10	14	39	2	8	3		
Comercio	10	2	6	15			9			
Cultura y Ocio	39	1	24	34	80		31		10	2
Demografía	97	18	25	31		6	9	15	12	
Deporte	8		7	10		1	8			
Economía	42	65	16	48	196	8	44	27	2	6
Educación	33	3	24	31	44		22		4	1
Empleo	7	39	28	84	30	5	24	31	1	
Energía	7	4	17	5		14	5			
Hacienda	51	51	7	28		10	18		1	
Industria	5	6	3	13		3	9			3
Legislación y Justicia	2	5	2	57		10	28		10	
Medio Ambiente	189	166	167	268	54	261	61	18	114	3
Medio Rural	1	14	59	30		11	55		1	
Salud	65	65	92	12	86	28	48		18	
Sector Público	280	22	115	15	77	54	66	15	26	
Seguridad			3	8			5			
Sociedad y Bienestar	75	12	18	18		2	19	60	10	3
Transporte	39	1	25	14	9	3	12		29	
Turismo	35	1	28	19		9	38		2	
Urbanismo e infraestructuras	6	364	14	14	19	0	10		7	
Vivienda		8		5	128	4	8			1

Nota: MUR, Murcia; RIO, La Rioja; CTL, Castilla y León; AND, Andalucía, GAL, Galicia; BAL, Islas Baleares; CTM, Castilla-La Mancha; CNT, Cantabria; MAD, Comunidad de Madrid; EXT; Extremadura.

Fuente: Elaboración propia (2024).

Como indica la tabla 3, la divulgación de datos poblacionales por parte de las autonomías españolas que no llegan a alcanzar los 1.000 conjuntos de datos suministrados es mayoritaria (8 de las 10 comunidades autónomas analizadas poseen al menos un *dataset* de esta categoría). El próximo apartado, destinado a mostrar los resultados de la información demográfica, mostrará un breve resumen de la información existente de cada autonomía.

3. Resultados

Se realiza una descripción del principal contenido ofertado por cada Comunidad autónoma, con mayor detalle en los conjuntos de datos etiquetados en la categoría Demografía.

Andalucía

De los 712 *datasets*, 31 pertenecen a la categoría de Demografía. Destacamos “Movimiento Natural de la Población” con información de nacimientos, matrimonios y defunciones. También señalar “Andalucía Como Hemos Cambiado” que recoge las variaciones en aspectos como el demográfico y cómo no, el Sistema de Información Demográfica de Andalucía (SIDEMA) que presenta una visión global de la evolución demográfica. Otros *datasets* serían “Estadísticas de Población de Andalucía Basadas en Registros Administrativos” y “Censos de Población y Viviendas” que ofrecen información demográfica, cultural, económica y social de todos los habitantes de Andalucía y provincias.

Aragón

El portal ofrece 2.713, de ellos 48 etiquetados como Demografía, siendo el *dataset* más típico “Pirámides de Población” que ofrece un conjunto de datos por año dentro de la serie temporal 1998 a 2020. Otro nutrido subconjunto de datos responde a la tipología PERE (Padrón de Españoles Residentes en el Extranjero) segmentando la información por sexo, edad o titulación. También se suministran “Revisión del Padrón Municipal” y “Número de Extranjeros Empadronados” entre otros.

Canarias

Es el portal autonómico más poblado con 14.237 conjunto de datos, de los cuales 4.705 pertenecen a la categoría Demografía. Si bien el *dataset* más habitual es “Población” por franjas de edad (de 0 a 14, de 15 a 64, etc.) por comarcas, distritos, etc. También existe abundante información como “Altas y Bajas Poblacionales”, “Crecimiento Vegetativo”, “Defunciones”, “Densidad de población”, “Información de Estudiantes”, “Edades Medias”, “Información de Hogares” e “Indicadores demográficos”.

Cantabria

El repositorio cántabro cuenta con 290 conjuntos de datos, pero las categorías existentes no se ajustan 100% a la norma nacional por lo que hemos realizado una clasificación manual. Algunas de estas asignaciones se han realizado de manera sencilla dada la similitud (como población, a Demografía; I+D y TIC, a Ciencia y Tecnología), mientras que otras han sido más complejas (Industria, Energía, Construcción y Vivienda con 53 *datasets*; Servicios, con 23 conjuntos de datos; Sector Primario, con 34, Nivel, Calidad y Condiciones de Vida con 11; Datos regionales, con 200; Datos Municipales, con 41; Datos Históricos, con 77 o Datos de Agrupaciones Municipales con 9 conjuntos de datos).

En todo caso, con la etiqueta Población están clasificados 15 conjuntos de datos cuyo contenido más destacado es “Censo de Población y Viviendas”, “Encuesta Continua (de hogares, inmigrantes, de población, de migraciones)”, “Indicadores Demográficos”, “Movimiento natural de población” y “Población censal 1900-2021”, entre otros.

Castilla y León

Su portal regional contiene un total de 744 conjuntos de datos clasificando a 25 de ellos dentro de la categoría Demografía. Aquí podemos encontrar *datasets* relativos a las defunciones, matrimonios, inmigraciones y emigraciones de/a otras Comunidades autónomas, parejas de hecho, población total por provincia y sexo, diversos indicadores demográficos, así como datos de población total por edades y por sexo.

Castilla-La Mancha

Nos ofrece 341 *datasets* en su portal, si bien únicamente 9 de ellos pertenecen a la categoría población. De ellos, destacamos tan sólo 2 de interés: “Datos estadísticos de población en Castilla-La Mancha”, con información relativa a las cifras oficiales del Padrón de Municipios referentes a Castilla-La Mancha y “Padrones municipales de Castilla-La Mancha”.

Cataluña

Presenta 1.218 conjuntos de datos, 33 de los cuales son poblacionales. A destacar que la información se ofrece de manera muy completa y no desagregando la información en función de distintas variables como realizan otras autonomías. Deben citarse “Censo de Población y Viviendas”, “Estadísticas de defunciones”, “Estadísticas de los Municipios”. También se ofrecen encuestas diversas (usos lingüísticos, uso del tiempo, etc.), así como indicadores demográficos y de movimientos migratorios, nombres (y apellidos) de la población, información del padrón y proyecciones, etc.

Comunidad de Madrid

Se ofrecen 12 conjuntos de datos, dentro de los 247 disponibles con la etiqueta Demografía. Existen 2 a nivel municipal (Madrid capital) y los 10 restantes ofrecen información a nivel autonómico como “Municipios de la Comunidad de Madrid (con información variada como densidad de población)”, varios conjuntos de datos de defunciones y de nacimientos, y diversos *datasets* sobre residentes (por edad, por nacionalidad, etc.).

Comunidad Foral de Navarra

A pesar de que Navarra es una de las autonomías con el portal más poblado (con 1.561 conjuntos de datos), es una de las Comunidades autónomas que menor atención presta a los datos demográficos con tan sólo seis *datasets*. Destacamos información sobre el censo de solicitantes de vivienda (uno por lugar de empadronamiento y otro por sus características), datos de población por zonas básicas de salud y tres conjuntos de datos con indicadores de género en población (un *dataset* relativo a solteros, otro a viudos y otro a población total).

Comunidad Valenciana

El portal contiene 1.453 *datasets*, 38 de ellos son categorizados como Demografía. Algunos de ellos relativos al censo de viviendas (un *dataset*, para el ejercicio 2001 y otro para el 2011), otros recogen información sobre inmigrantes, mientras que otros ofrecen estimaciones (de hogares

y de población). De todas formas, el conjunto de datos más frecuente es el padrón municipal continuo por distritos y secciones a razón de un *dataset* por año (del 1999 al 2022).

Extremadura

El portal tiene 12 conjuntos de datos, si bien no posee ningún *dataset* poblacional de interés.

Galicia

El portal señala la existencia de 512 conjuntos de datos. Su oferta por categorías agrupa varios de los epígrafes propuestos por la normativa nacional lo que ha obligado a realizar una asignación manual para reclasificar estos conjuntos de datos según categorías estándar. Ninguno de los conjuntos de datos estaba asociado con información de carácter demográfico.

Islas Baleares

Este repositorio presenta la singularidad de que, aunque se puede elegir la lengua del portal entre catalán y español, el idioma sólo afecta al continente (menús, opciones, etc.) pero no se aplica a los conjuntos de datos, que sólo pueden ser consultados en catalán.

En todo caso, tan sólo encontramos 6 conjuntos de datos (de los 453 existentes en el portal balear) que pertenezcan a la categoría Población: “Datos Estadísticos de la Población de las Islas Baleares”, “Secciones censales 2017”, “Municipios”, “Portal” (infraestructuras y servicios asociados a los transportes), “Distritos Censales 2017” y “Núcleos de población según clasificación del Instituto Nacional de Estadística”.

La Rioja

Tiene 853 conjuntos de datos, 18 de los cuales son relativos a la Demografía. De esta naturaleza observamos que se ofrecen varios conjuntos de datos de flujos de inmigración y emigración, así como varios *datasets* con información del padrón municipal de habitantes (por distritos y secciones censales, por comarcas agrarias, grupos de edad y sexo, por municipios). También ofrece información sobre matrimonios, nacimientos y defunciones.

País Vasco

Es el segundo portal más poblado conteniendo una oferta total de 11.539 conjuntos de datos que están identificados con las etiquetas propuestas por la norma nacional, si bien incorpora una categoría adicional denominada Euskera que contiene 190 *datasets*.

Dentro de la categoría población existen 1.385 conjuntos de datos. Algunos de ellos no son de contenido demográfico como “calles de Bilbao” y otros se alejan bastante de un contenido “estándar” de lo que puede ser considerado como conjunto de datos con información poblacional. Si nos circunscribimos a este tipo de temática, observamos *datasets* diversos como “Estadísticas de Personas Beneficiarias de Prestaciones”, aunque la información que predomina es la del tipo habitantes (según edad, lugar de nacimiento, nivel de estudios y las variaciones en el padrón municipal), todos ellos a nivel de entidades locales -municipios-.

Principado de Asturias

La peculiaridad de esta comunidad autónoma reside en que, al acceder al apartado descarga de datos de su portal (alojado en <https://transparencia.asturias.es/open-data>), se redirige al

portal nacional (www.datos.gob.es⁴) ofreciendo un total de 1.788 conjunto de datos que muestran la etiqueta de publicación “Principado de Asturias”.

De los 803 *datasets* poblacionales destacan “Ficheros Adquirientes” (personas que han adquirido la nacionalidad española) por sexos, tramos de edad, lugar de nacimiento, “Censos de Población y Vivienda”, “Encuesta de Población Activa”, “Estructura de Población”, “Datos de Evolución” (por familias, hogares, mercado laboral, mortalidad, movimientos de población o nivel de estudios. La tipología más nutrida es “Datos de Población” que son ofrecidos por sexo y edad a nivel de concejos, comarcas o a nivel agregado de autonomía (con información por periodos anuales). Análogamente existe una serie de conjuntos de datos denominados “Principales Indicadores Demográficos”.

Región de Murcia

El portal autonómico ofrece 961 conjuntos de datos de los cuales 97 están categorizados como Demografía. La mayor parte de ellos son presentados a nivel local, como los del Ayuntamiento de Lorca o los del Ayuntamiento de Molina de Segura, entre otros. A nivel autonómico se presentan varios conjuntos denominados “Evolución” (de nacidos, de partos, de defunciones, de matrimonios, de inmigraciones, de emigraciones, etc.). También se ofrece un fichero de “Indicadores Demográficos de Nacimientos” y otro de “Evolución de los SalDOS Migratorios con otras Provincias”.

4. Discusión

Las Autonomías españolas son muy distintas entre sí a nivel de superficie, población, riqueza, lengua propia, etc., por lo que no debe extrañar que existan diferencias en la implantación de las iniciativas de sus datos abiertos.

Estas diferencias en sus iniciativas de datos abiertos, ya señaladas a nivel multidimensional por Curto-Rodríguez (2021), se analizan en este texto desde la propia observación de lo propuesto siguiendo la norma técnica nacional en cuanto a la clasificación de sus conjuntos de datos. Como hemos puesto de manifiesto, existen comunidades autónomas como País Vasco que añaden categorías nuevas mientras que otras como Galicia realizan una agrupación que engloba varias de las tipologías indicadas.

Además, hemos podido constatar que el propio etiquetado se realiza de forma distinta, de manera que tan sólo Aragón, Comunidad Valenciana, La Rioja y Comunidad de Madrid realizan una asignación única (es decir, conjunto de datos que sólo pertenecen a una categoría) mientras que para Galicia, Principado de Asturias, País Vasco, Navarra, Castilla-La Mancha, Murcia, Andalucía, Islas Baleares, Región de Murcia y Canarias se observa el fenómeno opuesto, esto es, una asignación no unívoca de cada *dataset* a la categoría (se hace a más de una ya que la suma de los conjuntos de datos por categoría supera holgadamente a la oferta total del portal). Por otro lado, hemos comprobado que en Castilla y León, Cataluña y Cantabria se han dejado conjuntos de datos sin etiquetar puesto que la cifra total indicada su portal es inferior a la suma de las categorías (según se desprende de la observación de las tablas 2 y 3).

⁴ El portal nacional reúne todos los conjuntos de datos de los organismos públicos españoles que hayan realizado el proceso de sindicación o federación en el portal. Es, por tanto, habitual encontrar allí los mismos datos que las Comunidades autónomas o entidades locales muestran en su portal propio, siendo singular el caso del Principado de Asturias ya que no ofrece los datos en su repositorio, remitiendo al nacional.

Si consideramos las iniciativas desde una perspectiva global, las diferencias también son evidentes ya que el rango, calculado como las diferencias entre la Comunidad autónoma con portal de datos abiertos más poblado (Canarias con 14.237) y el que menos *datasets* ofrece (Extremadura con 12) supera los 14.000 conjuntos de datos (véase la tabla 4).

Obviamente estas diferencias a nivel agregado tienen su impacto en el número de conjuntos de datos de la categoría “Demografía” disponibles por cada una de las Comunidades autónomas. Nuevamente es Canarias la autonomía que más información demográfica está ofreciendo con 4.705 conjuntos de datos, seguida por la segunda comunidad autónoma en divulgación de información, el País Vasco donde, del total de 11.539 *datasets* ofertados, 1.385 pertenecen a la categoría “Demografía”, ocupando el tercer lugar el Principado de Asturias con 803 conjuntos de datos. Si dejamos a un lado a estas tres Autonomías, podemos observar que ninguna de las restantes alcanza los 100 conjuntos de datos con información de carácter demográfico (véase tabla 4). Finalmente estarían Galicia y Extremadura que son las únicas Autonomías españolas que no están divulgando ninguna información de la tipología “Demografía”. El análisis porcentual realizado, que mide que tanto por ciento ocupan los datos de la categoría “Demografía” dentro del total autonómico, se ofrece en la tabla 4.

Tabla 4. Total de conjuntos de datos abiertos, *datasets* demográficos y porcentaje sobre el total.

Datos/ Autonomía	CNR	VAS	ARA	AST	NAV	VAL	CAT	MUR	RIO	CTL	AND	GAL	BAL	CTM	CNT	MAD	EXT
<u>Demografía</u>	<u>4.705</u>	<u>1.385</u>	<u>48</u>	<u>803</u>	<u>6</u>	<u>38</u>	<u>33</u>	<u>97</u>	<u>18</u>	<u>25</u>	<u>31</u>		<u>6</u>	<u>9</u>	<u>15</u>	<u>12</u>	
<u>Totales</u>	14.237	11.539	2.713	1.788	1.561	1.453	1.218	912	853	744	712	512	453	341	290	247	12
<u>Porcentaje</u>	33%	12%	2%	45%	0%	3%	3%	11%	2%	3%	4%	0%	1%	3%	5%	5%	0%

Nota: CNR, Islas Canarias; VAS, País Vasco, ARA, Aragón; NAV, Comunidad Foral Navarra; VAL, Comunidad Valenciana; CAT, Cataluña; MUR, Murcia; RIO, La Rioja; CTL, Castilla y León; AND, Andalucía, GAL, Galicia; BAL, Islas Baleares; CTM, Castilla-La Mancha; CNT, Cantabria; MAD, Comunidad de Madrid; EXT; Extremadura.

Fuente: Elaboración propia (2024).

La observación de la tabla 4 permite apreciar que los líderes en cuanto al total de conjunto de datos demográficos son los mismos que a nivel porcentual, si bien el orden difiere ligeramente (Principado de Asturias, Canarias y País Vasco). La tabla 4 también permite observar que, salvo Región de Murcia, con un 11% de información demográfica, el resto de Comunidades autónomas no supera el 5%.

Previo a la finalización de este apartado hemos pretendido averiguar si existe asociación significativa a nivel estadístico entre en total de conjuntos de datos autonómicos y la existencia de datos demográficos, por lo que hemos decidido calcular el coeficiente de correlación de Pearson que arroja un valor de +0,9, lo que indica una correlación positiva directa muy marcada entre los datos demográficos y el total de datos de cada autonomía con una significación estadística al nivel 0,01.

Debemos, antes de pasar al apartado de las conclusiones, señalar como limitaciones del estudio su carácter descriptivo justificado por el hecho de tratarse de una investigación exploratoria. Por otro lado, la ausencia de aportaciones que emplearan una temática similar ha impedido la comparativa con la literatura existente que podría servir de contrapunto a las afirmaciones realizadas, aspecto éste de innegable interés. Futuras líneas de investigación podrían estar asociadas a la valoración cuantitativa de cada Autonomía, por ejemplo mediante el empleo de un indicador que se basara en una lista de comprobación, lo que podría incorporar el empleo de variables adicionales como los formatos en que se suministra la información -que permiten su uso por distintos perfiles de reutilizadores-, así como las licencias asociadas a los conjuntos de datos -que van a determinar los posibles usos de la información suministrada-. De todas

formas, la tarea no parece sencilla dadas las importantes diferencias por Autonomías en todos y cada uno de los aspectos señalados.

5. Conclusiones

Administraciones públicas de todos los niveles y países están implementando políticas de datos abiertos debido a su gran valor, tanto económico como social. En España todas las Autonomías disponen de portal de datos abiertos, si bien existen importantes diferencias entre ellos, tanto a nivel general (diseño del portal, iniciativas de fomento de reutilización de los conjuntos de datos, aplicaciones desarrolladas o número de *datasets* ofertados) como a nivel detallado de la categoría “Demografía”, objeto de este estudio.

La atención a los datos demográficos, de gran interés para el marketing por su usabilidad para la aplicación de técnicas de segmentación, muestra tres autonomías líderes en valores absolutos en esta categoría en: Canarias, País Vasco y Principado de Asturias. Si analizamos la importancia que estos datos tienen respecto al total suministrado, calculado como el porcentaje de *datasets* de contenido demográfico respecto al valor total, las Autonomías serían las mismas pero en un orden distinto: Principado de Asturias, Canarias y País Vasco. En todo caso, con la salvedad mencionada de estas tres autonomías, es necesario señalar que el resto parece mostrar un menor desempeño tanto el valor absoluto como en porcentaje respecto al acumulado de conjuntos de datos suministrados.

De todas formas, y en base al resumen del contenido de la categoría “Demografía” efectuada, hemos observado que el conjunto de información demográfica suministrada por las Autonomías es muy heterogéneo. En primer lugar se observa que no existe un contenido uniforme en la información que se está divulgando y que la formulación concreta de cada *dataset* varía mucho de una Autonomía a otra. Así, se observa que los conjuntos de datos de Cataluña agregan la información mientras que otras Autonomías como Canarias codifican y publican la información de manera desagregada en función de distintas variables como (sexo, franjas de edad, lugar de residencia, etc.). Concluimos que éste es uno de los hallazgos más relevantes de esta investigación, puesto que este aspecto de ausencia de estandarización (además de dificultar una comparativa real entre autonomías) entorpece el aprovechamiento de la información, al requerir un tratamiento individualizado de la información que se está suministrando. Nuestra recomendación al respecto sería exigir a las Administraciones una mayor homogeneización pues esto permitiría la reutilización y el aprovechamiento de estos datos de una manera más sencilla y facilitaría el tratamiento automatizado.

6. Referencias

- Abella, A., Ortiz-de-Urbina-Criado, M. y De-Pablos-Heredero, C. (2022). Criteria for the identification of ineffective open data portals: pretender open data portals. *Profesional de la Información*, 31(1), 1-12. <https://doi.org/10.3145/epi.2022.ene.11>
- Calle García, A. J., Quimis Vera, M. C., Piguave Vargas, M. T. y Zambrano Luzardo, J. S. (2024). La inteligencia artificial como herramienta en la segmentación de mercado. *Ciencia y Desarrollo*, 27(1), 193-202. <http://dx.doi.org/10.21503/cyd.v27i1.2556>
- Cerrillo Martínez, A. (2018). Datos masivos y datos abiertos para una gobernanza inteligente. *Profesional de la Información*, 27(5). <http://dx.doi.org/10.3145/epi2018.sep.16>
- Cerrillo-Martínez, A. (2014). Los principios de datos abiertos en la legislación española. *Revista*

- de Internet, Derecho y Política, 19, 62-77. <http://dx.doi.org/10.7238/idp.v0i19.2418>
- Cruz-Rubio, C. N. (2015). ¿Qué es (y qué no es) gobierno abierto? Una discusión conceptual. *EUNOMÍA. Revista en Cultura de la Legalidad*, 8, 37-53. <https://e-revistas.uc3m.es/index.php/EUNOM/article/view/2475>
- Curto-Rodríguez, R. (2017). Los portales autonómicos de datos abiertos y la información relacionada con la rendición de cuentas: punto de partida y situación tras la entrada en vigor de la ley 19/2013 de transparencia, acceso a la información pública y buen gobierno. *Revista Española de la Transparencia*, 5, 80-93. <https://revistatransparencia.com/ojs/index.php/ret/article/view/152/230>
- Curto-Rodríguez, R. (2020). Gobiernos autonómicos españoles y transparencia activa de sus repositorios digitales. Una revisión taxonómica. *Política y Sociedad*, 57(2), 543-565. <https://doi.org/10.5209/poso.68119>
- Curto-Rodríguez, R. (2021). Análisis multidimensional de los portales de datos abiertos autonómicos españoles. *Revista Española de Documentación Científica*, 44(1). <https://doi.org/10.3989/redc.2021.1.1745>
- Díez-Garrido, M. y Lázaro, M. M. (2022). La transparencia como medicina. La comunicación de la pandemia a través de los portales de datos abiertos de las autonomías. *Estudios sobre el Mensaje Periodístico*, 28(1), 65-80. <https://doi.org/10.5209/esmp.77172>
- Fernández Leiceaga, X. y Lago Peñas, S. (2013). Sobre el reparto de los derechos de déficit entre las Comunidades Autónomas: Una propuesta alternativa. *Revista de economía aplicada*, 21(63), 117-136. <https://www.redalyc.org/articulo.oa?id=96929566005>
- Kotler, P. y Armstrong, G. (2006). *Principles of Marketing* (11^a ed.). Prentice-Hall, Inc.
- Lnenicka, M., Nikiforova, A., Luterek, M., Azeroual, O., Ukpabi, D., Valtenbergs, V. y Machova, R. (2022). Transparency of open data ecosystems in smart cities: definition and assessment of the maturity of transparency in 22 smart cities. *Sustainable Cities and Society*, 82, 103906. <https://doi.org/10.1016/j.scs.2022.103906>
- Mabillard, V. y Keuffer, N. (2020). What drives online transparency practices at the local level? Evidence from 2222 municipalities in Switzerland. *Jahrbuch der Schweizerischen Verwaltungswissenschaften*, 11(1), 158-170. <https://doi.org/10.5334/ssas.143>
- Pozen, D. E. (2020). Seeing transparency more clearly. *Public Administration Review*, 80(2), 326-331. <https://doi.org/10.1111/puar.13137>
- Rangel-Carrillo, A. M., Maestre-Góngora, G. P. y Osorio-Sanabria, M. A. (2020). Principios, lineamientos, dimensiones y atributos para la evaluación de calidad de datos abiertos de gobierno. *Aibi revista de investigación, administración e ingeniería*, 8(S1), 54-65. <https://doi.org/10.15649/2346030X.950>
- Rodríguez-Pose, A. y 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>

Smith, W. R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing*, 21(1), 3-8.
<https://doi.org/10.1177/002224295602100102>

Sudirjo, F. (2023). Marketing strategy in improving product competitiveness in the global market. *Journal of Contemporary Administration and Management (ADMAN)*, 1(2), 63-69.
<https://doi.org/10.61100/adman.v1i2.24>

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor Books.

Verhulst, S., Winowatan, M., Young, A., Zahuranec, A. J. y Shaw, A. (2020). *Open Data Index: 10 insights on the value of open data*. <https://acortar.link/Xlhxkw>

Vickery, J. (2011). *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.
<https://acortar.link/2wLmFr>

Zúñiga, F., Mora Poveda, D. A. y 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>

CONTRIBUCIONES DE AUTORES/AS, FINANCIACIÓN Y AGRADECIMIENTOS

Conceptualización: Curto-Rodríguez, R. y Marcos Sánchez, R. **Análisis formal:** Curto-Rodríguez, R. y Marcos Sánchez, R. **Curación de datos:** Curto-Rodríguez, R. y Marcos Sánchez, R. **Redacción-Preparación del borrador original:** Curto-Rodríguez, R. y Marcos Sánchez, R. **Redacción-Revisión y Edición:** Curto-Rodríguez, R. y Marcos Sánchez, R. **Todos los/as autores/as han leído y aceptado la versión publicada del manuscrito:** Curto-Rodríguez, R. y Marcos Sánchez, R.

Financiación: Esta investigación no posee financiamiento externo.

AUTOR/ES:**Ricardo Curto-Rodríguez**

Universidad de Oviedo.

Diplomado en Ciencias Empresariales, Licenciado en Administración de Empresas y Doctor *Cum Laude* en Economía y Empresa. Acreditado ANECA como Contratado Doctor. Profesor en el departamento de Administración de Empresas de la Universidad de Oviedo y en el Máster de Formación del Profesorado de la Universidad Internacional de la Rioja. Cuenta con un sexenio de investigación reconocido. Autor de una quincena de artículos publicados en revistas indexadas de alto impacto y de siete capítulos de libro. Su línea de investigación son los datos abiertos, la transparencia y la rendición de cuentas. Es miembro del equipo editorial de la Revista Española de la Transparencia, revisor en revistas indexadas en WoS, y ponente y coordinador en congresos (coordinador general del CIT 2021).

curtoricardo@uniovi.es**Índice H: 9****Orcid ID:** <https://orcid.org/0000-0002-6628-5278>**Scopus ID:** <https://www.scopus.com/authid/detail.uri?authorId=57204037374>**Google Scholar:** <https://scholar.google.com/citations?user=x9Bw22YAAAAJ&hl=es>**ResearchGate:** <https://www.researchgate.net/profile/Ricardo-Curto-Rodriguez>**Rafael Marcos Sánchez**

Universidad Internacional de La Rioja.

Doctor en Economía y Licenciado en Ciencias Económicas, Empresariales y Actuariales por la Universidad Complutense de Madrid. Desde 2011, es profesor en la Universidad Internacional de La Rioja (UNIR) y docente en CES Valdecás (Madrid) desde 2001. Cuenta con un sexenio de investigación reconocido. Con experiencia como Director Técnico del CES Valdecás (2007-2011), ha impartido asignaturas de economía, gestión educativa, didáctica y metacognición. Sus líneas de investigación incluyen la metacognición, gestión de centros y *coaching* educativos. Ha dirigido Trabajos de Fin de Máster y enseña en programas de formación del profesorado y dirección de centros educativos. Además, ha desarrollado cursos sobre competencias básicas y liderazgo educativo.

rafael.marcos@unir.net**Índice H: 5****Orcid ID:** <https://orcid.org/0000-0002-3179-6211>**Google Scholar:** <https://scholar.google.com/citations?user=8sl099AAAAAJ&hl=es>**ResearchGate:** <https://www.researchgate.net/profile/Rafael-Marcos>