

# ANÁLISIS ESTADÍSTICO DE LOS INGRESOS DE ANDALUCÍA DURANTE EL AÑO 2021

MANUEL CEBALLOS GONZÁLEZ<sup>1\*</sup>, JOSÉ M. DE LA PUERTA BELZUNEGUI<sup>2</sup>

<sup>1</sup> DEPARTAMENTO DE INGENIERÍA, ESCUELA SUPERIOR DE INGENIERÍA; UNIVERSIDAD LOYOLA ANDALUCÍA; AV. DE LAS UNIVERSIDADES, S/N, 41704 DOS HERMANAS, SEVILLA, ESPAÑA; MCEBALLOS@ULYOYOLA.ES;  [HTTPS://ORCID.ORG/0000-0003-0913-6417](https://orcid.org/0000-0003-0913-6417)

<sup>2</sup> DEPARTAMENTO DE INGENIERÍA, ESCUELA SUPERIOR DE INGENIERÍA; UNIVERSIDAD LOYOLA ANDALUCÍA; AV. DE LAS UNIVERSIDADES, S/N, 41704 DOS HERMANAS, SEVILLA, ESPAÑA; JMPUERTABELZUNEGUI@AL.ULOYOLA.ES;  [HTTPS://ORCID.ORG/0009-0005-3634-9102](https://orcid.org/0009-0005-3634-9102)

\*AUTOR CORRESPONSAL

Citación: Ceballos González, M. y De La Puerta Belzunegui, J. M. (2023). Análisis estadístico de los ingresos de Andalucía durante el año 2021. *Inquietud Empresarial*, 23(2), e16495. <https://doi.org/10.19053/01211048.16495>

Editor: Blanco-Mesa, Fabio

Recibido: 04/09/2023  
Aceptado: 14/11/2023  
Publicado: 22/12/2023

Códigos JEL: C10, C14, L10

Tipo de artículo: Investigación



**Resumen:** Andalucía es una de las comunidades autónomas de España que mayor crecimiento económico presenta. El objetivo del presente artículo es analizar los ingresos de los distintos municipios de Andalucía durante el año 2021. Para ello, se presenta un estudio estadístico sobre la variable ingresos por habitante haciendo uso del programa Stata y una base de datos. Stata es un potente software estadístico cuya primera versión data de 1985 y es muy utilizado hoy en día en el entorno del big data. La base de datos utilizada corresponde a la plataforma del Sistema de Información Multiterritorial de Andalucía (SIMA) de la página web del Instituto de Estadística y Cartografía de la Junta de Andalucía. Los resultados obtenidos permiten concluir que en Andalucía existen diferencias estadísticamente significativas con respecto a los ingresos tanto a nivel municipal como provincial. Esto da lugar a la presencia de datos atípicos. Tras varios tests se concluye que los datos de la variable ingresos por habitante no se ajusta a una distribución normal y que, tras obtener la matriz de correlaciones, existe una correlación lineal con varias variables de la base de datos de SIMA.

**Palabras clave:** análisis estadístico, base de datos, Andalucía, ingresos.

# STATISTICAL ANALYSIS OF THE ANDALUSIAN INCOMES DURING THE YEAR 2021

**Abstract:** Andalusia is one of the autonomous communities of Spain with the highest economic growth. The main goal of this paper is to analyse the budgets in every city and town of Andalusia during the year 2021. In order to do so, we have developed a statistic analysis for the variable income per capita by using Stata software and a database. Stata is a powerful statistical software whose first version is from 1985. It is very common to use this software to deal with problems regarding big data. The database considered corresponds to the Mutiterritorial Information System from the website of the Institute of Statistics and Cartography of the Junta de Andalucía. The results obtained allow us to conclude that in Andalusia there are statistically significant differences with respect to both at the municipal and provincial level. This yields to the existence of atypical and highly atypical data. After various tests, it has been concluded that the data for the incomes per capita does not fit a normal behaviour and that there exists a strong correlation between this variable and other variables from the database.

**keywords:** statistical study, database, Andalusia, incomes.

## I. INTRODUCCIÓN

Andalucía es una de las comunidades autónomas más importantes de España. Destaca por su notable potencial empresarial, una sólida economía y una ubicación estratégica que la convierte en un lugar atractivo para el establecimiento y crecimiento de empresas. El potencial empresarial de Andalucía está respaldado por su amplia infraestructura de transporte, que facilita la conexión con otros mercados nacionales e internacionales. Esta accesibilidad fomenta el comercio y la exportación y crea oportunidades para empresas de diferentes sectores. Además, cuenta con una mano de obra altamente cualificada y diversificada. Por todo ello, ha sido la tercera comunidad que mayor crecimiento empresarial ha experimentado durante el año 2021 (Instituto de Estadística y Cartografía de Andalucía, 2021). Más concretamente, se crearon un total de 15 000 empresas nuevas. La que más destaca es el Grupo Bidadarma, cuyo principal cometido es suministrar medicamentos a más de 12 500 farmacias al día.

No obstante, para comprender el poder adquisitivo de Andalucía y tomar decisiones, resulta fundamental realizar un estudio estadístico. Este tipo de análisis permite conocer aspectos relevantes, como la estructura económica, los sectores emergentes, las tendencias de mercado, etc. Además, permite a los emprendedores, inversores y responsables políticos tomar decisiones estratégicas basadas en evidencia, con el objetivo de aprovechar al máximo las oportunidades empresariales en la región.

Este artículo se propone determinar, analizar y justificar a través de un estudio estadístico las características de los ingresos por habitante durante el año 2021 en todos los municipios de Andalucía. Para ello, se usará el software Stata y una base de datos de la plataforma del Sistema de Información Multiterritorial de Andalucía (SIMA) de la web del Instituto de Estadística y Cartografía de dicha comunidad. El programa Stata es un software estadístico cuya primera versión data de 1985 y ofrece a los usuarios analizar, administrar y generar visualizaciones gráficas de los datos. El Instituto de Estadística y Cartografía de Andalucía está forma-

do por varios organismos, como la Comisión Andaluza de Estadística y Cartografía y la Comisión Técnica de Estadística (Álvaro et al., 2013). La plataforma SIMA se encuentra en la web de dicho instituto y es un banco de datos con información multitemática y multiterritorial de todos los municipios de Andalucía (Sistema de Información Multiterritorial de Andalucía, 2023). Entre estos datos se encuentra el entorno físico y medio ambiente, demografía, población, sociedad, economía, mercado del trabajo y hacienda.

El principal objetivo del presente artículo es el análisis estadístico de los ingresos de todos los municipios de Andalucía durante el año 2021. Para ello, se trabajará con la variable ingresos por habitante. El análisis incluye estadísticas descriptivas, diagramas y estudio de los datos atípicos y extremadamente atípicos. Un segundo objetivo es la clasificación del tipo de ingresos (bajo, medio-bajo, medio-alto o alto) en cada municipio, a través de un criterio basado en los cuartiles para representar mediante mapas la tipología de ingresos en cada provincia. El tercer objetivo es un estudio de los municipios que corresponden a datos atípicos y extremadamente atípicos. El cuarto objetivo es un análisis comparativo de medias de ingresos provinciales estableciendo las parejas de provincias entre las que existen diferencias estadísticamente significativas. El quinto objetivo corresponde a estudiar si los datos de ingresos por habitante se ajustan a una distribución normal, lo que permitiría utilizar los parámetros de dicha distribución para inferir la probabilidad de que se obtenga un determinado valor de ingresos y los mayores o menores que él. Como sexto y último objetivo está el desarrollo de un análisis bivalente con el resto de las variables de la base de datos de SIMA.

Este artículo tiene la siguiente estructura. Tras esta breve introducción, la segunda sección aborda el marco teórico y antecedentes con un estudio de la distribución de empresas por actividad económica. En la tercera sección, se muestra la metodología utilizada. En la cuarta sección, se presenta un estudio estadístico sobre la variable ingresos por habitante con un sumario de estadísticas descriptivas y varios diagramas de representación: histograma, tallo-hojas y cajas-patas; además, se analiza el tipo de ingresos por habitante en cada provincia andaluza, mediante una clasificación basada en los cuartiles; luego se analizan los datos atípicos y extremadamente atípicos a nivel autonómico y provincial y se continúa con un test de hipótesis para comparar las medias provinciales y estudiar si existen diferencias estadísticamente significativas entre ellas; con el resto de variables de la base de datos SIMA se hace un test de normalidad y un análisis bivalente. La quinta sección expone una discusión sobre los resultados obtenidos y limitaciones. Por último, hay una sección de conclusiones y futuras investigaciones, además de la bibliografía.

## 2. MARCO TEÓRICO

En esta sección se hace un resumen de la historia económica de Andalucía, para después analizar la distribución de la tipología de empresas por actividad económica y, por último, revisar estudios estadísticos relacionados con la economía andaluza.

La historia de la economía de Andalucía es muy diversa (Tejedor, 2022). A lo largo de los siglos, ha experimentado varias etapas y ha estado influenciada por factores como la ubicación geográfica, los recursos naturales y los cambios políticos y sociales.

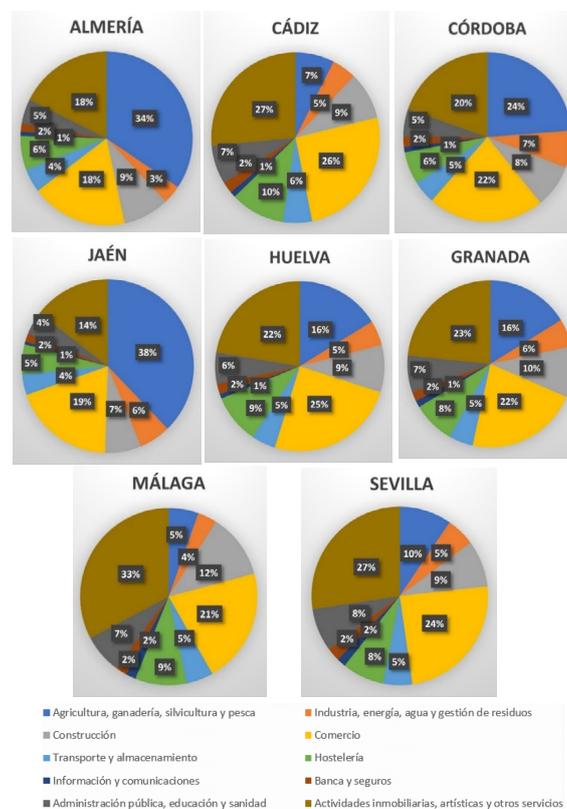
Durante la Antigüedad e Imperio romano, las principales ciudades de Andalucía eran Hispalis (Sevilla), Corduba (Córdoba) y Gades (Cádiz). La región era famosa por su producción de aceite de oliva, vino y productos agrícolas. Los musulmanes en el siglo VIII introdujeron nuevas técnicas agrícolas como el riego por

goteo y la plantación de cultivos de cítricos, arroz y caña de azúcar. Andalucía se convirtió en un importante centro de producción agrícola y comercial. Sus puertos mediterráneos como Almería y Málaga facilitaron el comercio con el norte de África y el Oriente Medio.

Durante la reconquista cristiana del siglo XIII, Andalucía experimentó cambios en su estructura económica. La economía se basaba principalmente en la agricultura y la ganadería, los cultivos de trigo, vid y olivo eran fundamentales. En la era de los descubrimientos, el puerto de Sevilla se convirtió en un centro clave para el comercio con el Nuevo Mundo. La plata, el oro, los productos agrícolas y las especias impulsaron la economía regional. En el siglo XIX, la economía andaluza experimentó una transformación significativa con la industrialización: se fundaron distintas fábricas y minas en Málaga, Sevilla y Huelva. No obstante, la región también sufrió problemas sociales y económicos, como la emigración masiva y la pobreza. Durante el siglo XX, la economía se caracterizó por la agricultura intensiva y la producción de bienes de consumo y las provincias costeras desarrollaron un importante sector turístico, especialmente en la Costa del Sol.

En la actualidad, la economía de Andalucía es diversa y abarca varios sectores. La agricultura sigue siendo relevante con la producción de aceite de oliva, frutas, hortalizas y vino. El turismo es una fuente clave de ingresos con ciudades como Sevilla, Málaga y Granada. Además, la región cuenta con una importante industria.

A continuación, se muestra una distribución de la tipología de empresas por actividad económica durante el año 2021 según la clasificación nacional de actividades económicas (CNAE 09) de Andalucía. En la Figura 1, elaborada con los datos obtenidos a partir del Instituto de Estadística y Cartografía de Andalucía, se observa cómo se destacan los sectores de agricultura, ganadería, silvicultura, pesca y comercio, así como actividades inmobiliarias, profesionales, auxiliares, artísticas y otros servicios por encima de los demás.



**Figura 1.** Distribución de empresas por sector de actividad en Andalucía

En la literatura, se pueden encontrar diversos trabajos previos relacionados con el estudio estadístico de la economía andaluza. Algunos de ellos analizan aspectos como la empleabilidad y los factores económicos que la condicionan. Congregado y García (2002) compararon el mercado laboral andaluz con el español y dedujeron que la probabilidad de que los jóvenes trabajadores andaluces encuentren empleo es mucho menor que la del resto de los españoles. Además, se concluyó que los principales efectos sobre esta probabilidad son el ciclo económico y la educación.

Técnicas como el análisis multivariante y análisis cluster han sido utilizadas para desarrollar estudios sobre la economía andaluza. Estas técnicas han permitido deducir qué variables (físicas, demográficas, sociales, etc.) reflejan la realidad socioeconómica de las localidades de Andalucía (Cruces et al., 2009), lo que ayudó a clasificar los municipios de la comunidad para elaborar políticas de desarrollo. Estas herramientas también se han aplicado con la inferencia estadística para determinar los factores que condicionan la propensión exportadora de las empresas andaluzas (Pérez et al., 2017).

También se utilizaron las Matrices de Contabilidad Social (MCS), para analizar la economía andaluza para el período 1990-2005 (Cardenete y López, 2015) y comparar la estructura económica en 2005-2010 (Campoy-Muñoz et al., 2015). Como resultado, se identificaron los sectores productivos más relevantes de Andalucía y se analizaron los efectos de la crisis económica de 2008.

Por último, cabe también mencionar la teoría de las Cadenas Globales de Valor (GVC, por sus siglas en inglés), que busca comprender la forma en la que las empresas y los países participan en la producción y distribución de bienes y servicios a nivel mundial. Esto se centra en analizar cómo se organiza y coordina la cadena de actividades económicas a lo largo de diferentes etapas de producción, diseño, fabricación, comercialización y entrega al consumidor (Kano et al., 2020).

Se han llevado a cabo estudios sobre la participación de Andalucía en las cadenas globales de valor, lo que ha aportado una visión general de los sectores y actividades económicas en los que esta comunidad tiene una mayor presencia (Instituto de Estadística y Cartografía de Andalucía, 2017). También se ha analizado su relación con la competitividad y el desarrollo económico regional (Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía, 2018).

Un estudio estadístico sobre los ingresos de Andalucía puede ser de gran interés para la teoría de las Cadenas Globales de Valor, pues permite identificar sectores y actividades económicas en los que Andalucía tiene una mayor presencia y participación. Esto ayudaría a comprender las oportunidades para el crecimiento económico en la región. Además, podría ayudar a revelar las interconexiones y los vínculos existentes entre las empresas andaluzas y las empresas a nivel nacional e internacional en las cadenas de valor. Esto proporcionaría información valiosa sobre las relaciones comerciales, la transferencia de conocimientos y la colaboración entre empresas, lo cual es crucial para mejorar la competitividad y la capacidad de innovación y aprovechar las oportunidades de crecimiento económico.

### **3. METODOLOGÍA**

En esta sección, se explica la metodología implementada. En un primer momento, se hace una introducción al análisis de datos, luego se analiza la fuente de datos utilizada y su tratamiento, para finalizar con comentarios acerca de los métodos, técnicas y procedimientos aplicados.

### 3.1 Introducción al análisis de datos

Para llevar a cabo el análisis, se usarán los datos proporcionados por la plataforma SIMA, que se encuentran en la página web del Instituto de Estadística y Cartografía de Andalucía. Los datos recogidos en dicha plataforma son muy diversos y analizan características de cada municipio andaluz: el entorno físico, medioambiente, demografía, población, sociedad, economía, etc. Entre todos ellos, se encuentran los datos de ingresos por habitante y por municipio en el año 2021 (Instituto de Estadística y Cartografía de Andalucía, 2021). Los datos de SIMA se exportan a una hoja de cálculo Excel para su futura importación al programa Stata, el software estadístico utilizado para procesar los datos y llevar a cabo el análisis de los ingresos en cada municipio de Andalucía.

El programa Stata ha tenido tres grandes precursores: Biomedical Program (BMDP), Statistical Analysis System (SAS) y (Statistical Progrms for the Social Sciences (SPSS). De igual manera, son tres las características más destacadas que han logrado que Stata sea una de las aplicaciones más usadas en estadística. En primer lugar, el empleo de instrucciones con un fácil lenguaje de modo interactivo. En segundo lugar, su especialización en el análisis de regresiones, por lo que ofrece una amplia gama de procedimientos, que van desde la regresión simple hasta los modelos de ecuaciones estructurales. Por último, este software además de empaquetar todas las instrucciones en un fichero, un técnico experto también puede elaborar nuevas utilidades diferentes a las implementadas (Universidad Loyola, 2021). Stata se destaca en el análisis de muestras complejas, series temporales, análisis de datos históricos, datos de panel, gráficos de contornos y marginales y el análisis de modelos de ecuaciones estructurales (Bernardi y Escobar Mercado, 2012).

### 3.2 Fuentes y tratamiento de los datos

Para el análisis de datos, se utilizarán los datos proporcionados por la plataforma SIMA, se exportarán a formato Excel y se importarán al editor de variables del programa Stata. Luego se seguirán los pasos que se pueden encontrar en el libro *Análisis de datos de la Universidad Loyola* (2021) para realizar el análisis estadístico y obtener todos los resultados que se presentan en la cuarta sección de este artículo. Todo ello referido a la variable ingresos por habitante, que se define como el cociente entre el total de ingresos en un municipio y el número de habitantes del mismo. Para esta variable, se calculan en la sección 4.1 las estadísticas descriptivas más importantes y diversas representaciones gráficas.

Después, en la sección 4.2, se lleva a cabo una clasificación de la tipología de ingresos en cada municipio según los valores de los cuartiles ( $Q_1$ ,  $Q_2$  y  $Q_3$ ) (a nivel autonómico. Esto permite obtener una clasificación de los datos. Así, un municipio tiene un nivel de ingresos Bajo si se encuentra por debajo del 25 % (valores menores que  $Q_1$ ); ingresos de tipo Medio-Bajo si se encuentra entre el 25 % y el 50 % (valores mayores que  $Q_1$  y menores que  $Q_2$ ); ingresos de tipo Medio-Alto si se sitúa entre el 50 % y el 75 % (valores mayores que  $Q_2$  y menores que  $Q_3$ ); e ingresos de tipo Alto está por encima del 75 % (valores mayores que  $Q_3$ ). A partir del valor de los cuartiles, se obtuvo el valor del rango intercuartílico  $RIC = Q_3 - Q_1$ .

Lo anterior se utiliza en la sección 4.3 para estudiar las observaciones cuyos valores difieren del resto de los datos. De esta forma, los datos atípicos vienen determinados por las observaciones por debajo de  $Q_1 - 1.5 \times RIC$  o bien por encima de  $Q_3 + 1.5 \times RIC$ . Los datos extremadamente atípicos están formados por los valores que están por debajo de  $Q_1 - 3 \times RIC$  o bien por encima de  $Q_3 + 3 \times RIC$ . Además de obtener estos valores, se estudió cómo se distribuyen entre las distintas provincias andaluzas.

En la Sección 4.4, se aborda un estudio de comparación de medias provinciales. Para ello, se usó la media, desviación típica y el número de municipios de cada provincia y se aplicó el test de la distribución T de Student para comparar medias y ver si existen diferencias estadísticamente significativas entre las provincias. Este tipo de distribución se utiliza para estimar el valor de la media poblacional en una distribución normal. El test mencionado se aplica cuando el tamaño muestral es demasiado pequeño como para que el estadístico en el que está basada la inferencia esté normalmente distribuido. En cuanto a la formulación del test, comienza considerando dos provincias A y B y realiza el contraste de hipótesis  $H_0: \mu_A = \mu_B$ ,  $H_A: \mu_A \neq \mu_B$

La estadística del contraste viene dada por 
$$T = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{(S_A^2)}{n_A} + \frac{(S_B^2)}{n_B}}}$$

$\bar{X}$ ,  $S^2$ ,  $n$  son la media, desviación típica y número de municipios para cada provincia, respectivamente. Como se puede observar, se define como la diferencia de las medias dividido entre la raíz cuadrada de la suma de los cuadrados de las desviaciones entre el tamaño de su muestra correspondiente.

La sección 4.5 tiene como objetivo realizar los test de normalidad de Kolmogorov-Smirnov y Shapiro-Wilk para contrastar si la variable objeto de estudio se ajusta o no a una distribución normal. Estos tests se basan en un contraste de hipótesis y se realizan mediante el programa Stata con los comandos `swilk` y `ksmirnov`.

Por último, en la sección 4.6, se lleva a cabo un análisis bivalente comparando los ingresos por habitante con el resto de variables de la base de datos de la plataforma SIMA. Mediante Stata se generó una matriz de correlación comparando la variable y todas las demás. También se elaboraron diagramas de dispersión y una tabla de correlación para las variables que tienen mayor correlación con los ingresos por habitante.

## 4. RESULTADOS

En esta sección se desarrolla un estudio estadístico presentado en seis apartados acorde a los objetivos que se indicaron en la introducción. Se trabajará con la variable ingresos por habitante, definida como el cociente entre los ingresos totales de un municipio y el número de habitantes del mismo. Se desarrollará un análisis estadístico utilizando el programa Stata. En primer lugar, se calcularán las estadísticas descriptivas y diagramas representativos más importantes. Luego, se analizará el tipo de ingresos de cada municipio acorde a los criterios que se indicaron en la sección anterior. Después, se estudiarán los datos atípicos y extremadamente atípicos a partir de las estadísticas ya calculadas. En el cuarto apartado, se llevará a cabo una comparación de medias provinciales con respecto a los ingresos por habitante aplicando el test de la distribución T de Student. En el siguiente paso, se aplicarán los Test de Kolmogorov-Smirnov y de Shapiro-Wilk para contrastar si los datos de la variable se ajustan o no a una distribución normal. Por último, se realizará un estudio bivalente analizando si existe correlación entre la variable y las demás recogidas en la base de datos de la plataforma SIMA.

### 4.1 Estadísticas descriptivas y diagramas

Se obtuvo un sumario de estadísticas descriptivas para la variable ingresos por habitante (Tabla 1), a través del análisis de 730 municipios andaluces, y se obtuvo una media de 1519.2 EUR de y una desviación típica de 776.16. El mayor valor obtenido fue 7916.033 EUR, dato que corresponde al municipio de Huévar

del Aljarafe, situado en la provincia de Sevilla. Este dato se debe a que es una población relativamente pequeña, con 3129 censados en el año 2021, y que tiene un presupuesto inusualmente alto para una población de esas características cercano a los 25000000 EUR. El menor valor obtenido fue de 361.9793 EUR, que corresponde al municipio de HuétorVega (Granada). Además, se obtuvieron el valor de los cuartiles para calcular el rango intercuartílico y usar dichos datos en la sección 4.3. Los valores de los cuartiles obtenidos con Stata fueron los siguientes:

$$Q_1=1070.72\text{€}, Q_2=1284.72\text{€}, Q_3=1718.12\text{€}.$$

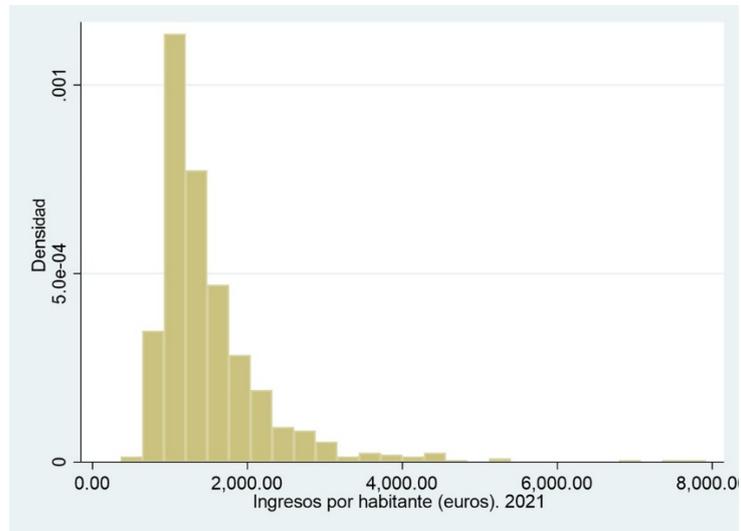
A partir del valor de los cuartiles, se obtuvo el valor del rango intercuartílico:

$$RIC=Q_3-Q_1=647.4 \text{ €}$$

En la Tabla 1, se observa que la media se encuentra ligeramente desplazada a la derecha y que hay una alta dispersión con unos valores de varianza y desviación estándar elevados, además de valores de sesgo y curtosis positivos. A partir de esto, se realizaron los diagramas o representaciones gráficas más importantes para dicha variable: el histograma, el diagrama de tallos y hojas y el de cajas y patas (Figuras 2-4). El histograma muestra una clara asimetría hacia la izquierda, con lo que se evidencia que en el lado opuesto asoman tímidamente algunos de los datos atípicos. En el diagrama de tallos y hojas (Figura 4) los datos atípicos se observan con mayor claridad, mientras que en el diagrama de cajas y patas (Figura 3) se confirma la asimetría y la existencia de numerosos datos atípicos por la derecha; por la izquierda no se observan atípicos, lo que tiene sentido ya, que serían valores cercanos a cero o incluso negativos.

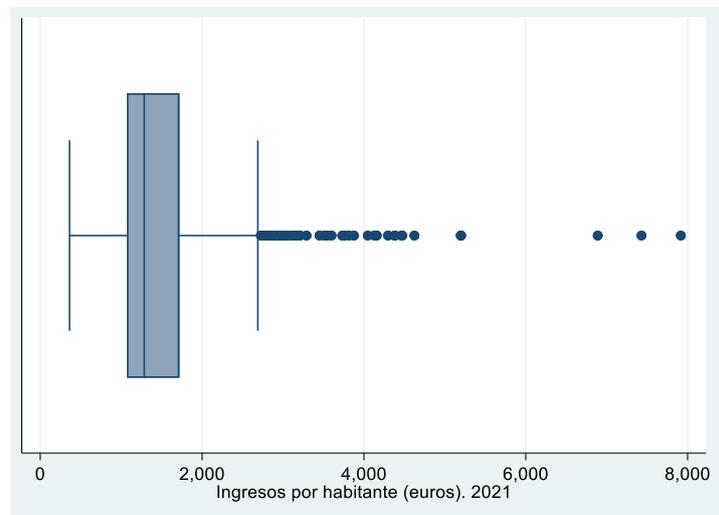
**Tabla 1.** Estadísticas descriptivas

Variable	Observaciones	Media	Media	Mínimo	Máximo
Ingresos por hab.	730	1519.209	776.1599	361.9793	7916.033
<b>Ingresos por habitante (euros). 2021</b>					
	<b>Percentiles</b>	<b>Mínimo</b>			
1%	676.0318	361.9793			
5%	820.6559	610.5836			
10%	921.4852	636.6367	<b>Observaciones</b>		730
25%	1070.716	642.8896			
50%	1284.742		<b>Media</b>		1519.209
		<b>Máximo</b>	<b>Desviación Típica</b>		776.1599
75%	1718.12	5199.882			
90%	2316.041	6889.575	<b>Varianza</b>		60242.1
95%	2888.492	7429.604	<b>Asimetría</b>		3.152467
99%	4464.833	7916.033	<b>Curtosis</b>		19.05513



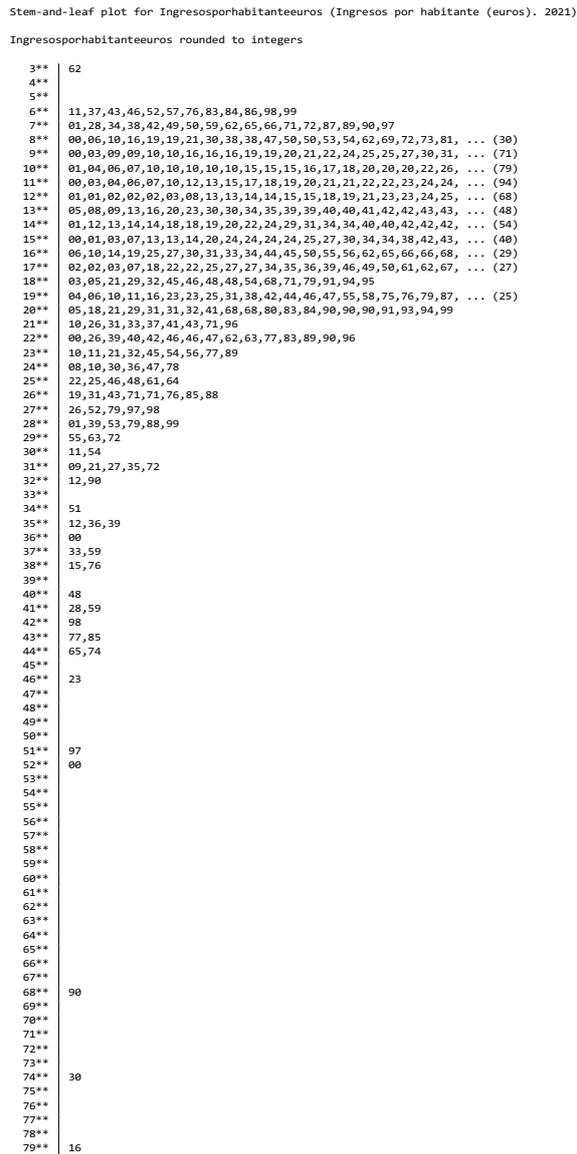
**Figura 2.** Histograma

Nota. Creación propia en Stata.



**Figura 3.** Diagrama de cajas y patas

Nota. Creación propia en Stata.

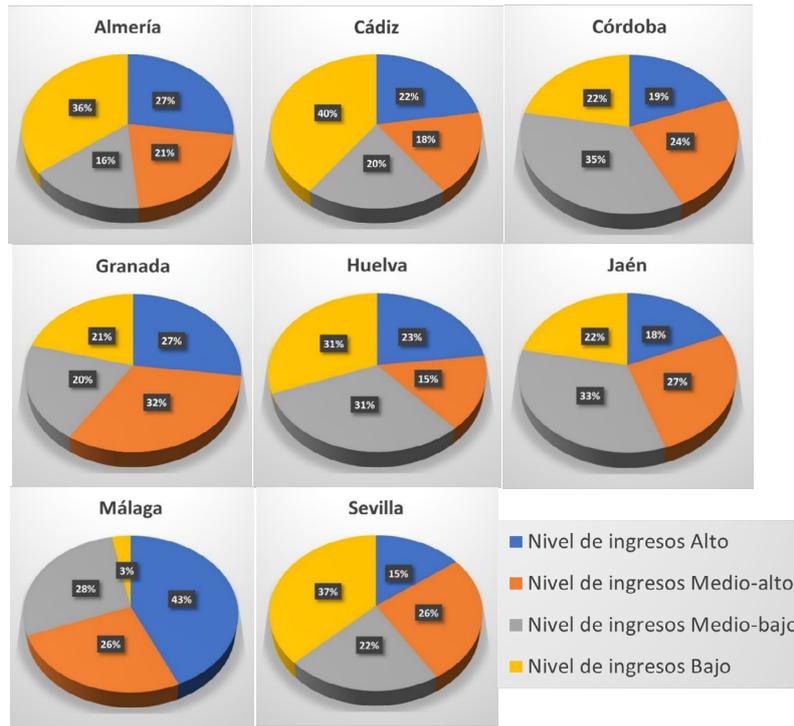


**Figura 4.** Diagrama de tallos y hojas

Nota. Creación propia en Stata.

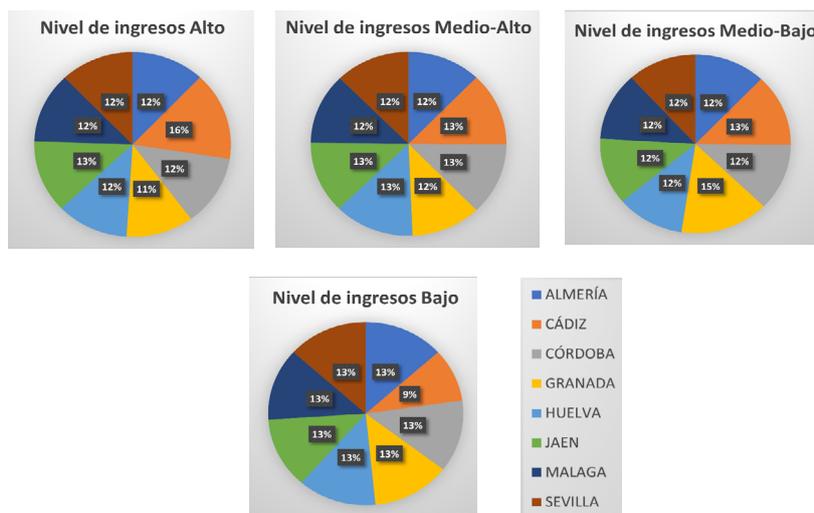
## 4.2 Tipos de potencial

En este apartado se analizará el tipo de ingresos de cada municipio andaluz, a través de un diagrama sectorial (Figura 5) que muestra la distribución de la tipología de ingresos (Bajo, Medio-Bajo, Medio-Alto y Alto). Se observa cómo en la provincia de Málaga predominan los ingresos altos, mientras que en Córdoba, Huelva y Jaén los ingresos son medio-bajos y en las provincias de Almería, Cádiz y Sevilla se destacan los bajos, mientras que los medio-altos son los más comunes en Granada.



**Figura 5.** Tipos de ingresos en cada municipio de Andalucía

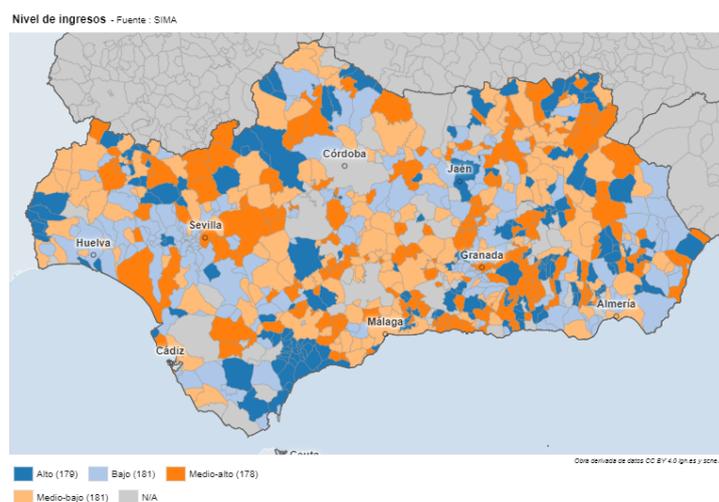
Después, en la Figura 6, se muestra cómo se distribuye el tipo de ingresos en las ocho provincias. Estos gráficos fueron normalizados en proporción al número de municipios que posee cada provincia. Se observa que Cádiz es la provincia que cuenta con mayor proporción de municipios con potencial empresarial alto; en los tipos de ingresos Medio-Alto y Bajo, todas las provincias tienen valores similares; y Granada es la provincia que más se destaca con respecto al tipo Medio-Bajo. Por último, en la Figura 7 se recoge la distribución del tipo de ingresos en toda Andalucía.



**Figura 6.** Distribución de los tipos de ingresos en Andalucía

A continuación, se comentan las conclusiones extraídas de las figuras mencionadas:

- En Almería predominan los ingresos bajos, seguidos de los altos, por lo que se puede concluir que existe mucha desigualdad entre los municipios. La mayoría de las localidades con ingresos bajos se reparten entre el norte y el sur de la provincia, mientras que los del altos y medio-altos, se distribuyen uniformemente por la zona central.
- Al igual que en Almería, en la provincia de Cádiz predominan los ingresos bajos, aunque en este caso con mayor peso y no están distribuidos de una manera determinada. Quizás se podría destacar la presencia de varios municipios con ingresos altos en la zona sureste de la provincia. Es posible que este hecho se deba a la cercanía con la Costa del Sol, una de las zonas más poderosas económicamente de la comunidad autónoma.
- En Córdoba y Jaén, existe una clara predominancia de los ingresos medios, especialmente de los medio-bajos. Se puede afirmar que se encuentran distribuidos de manera más o menos uniforme en ambas provincias.
- En la provincia de Granada, predominan los ingresos medio-altos, seguidos de los ingresos altos, por lo que se concluye que se trata una región con alta actividad económica. Se aprecia que la mayoría de los ingresos altos o medio-altos se encuentran en el centro y al norte, en la zona de la Sierra de Cazorla.
- Un 31 % de los municipios de Huelva tienen tanto ingresos bajos como medio-bajos. Por tanto, más de la mitad de los municipios están por debajo de la media andaluza. La mayoría de las localidades con ingresos bajos se encuentran al sur, cerca de la costa. Al oeste se destaca un grupo de municipios con ingresos altos, todos ellos fronterizos con Portugal.
- En términos económicos, la provincia de Málaga es la más importante de la Comunidad Autónoma de Andalucía. Cuenta con un 69 % de municipios con ingresos por encima de la media (43 % nivel alto) y con un ínfimo 3 % de municipios con ingresos bajos. La mayoría de las localidades con ingresos altos se aglutinan al suroeste de la provincia, muy cerca de la Costa del Sol.
- En Sevilla predominan los ingresos bajos y es la provincia con menor porcentaje de localidades con ingresos altos. Sin embargo, no es la provincia con menor actividad económica. La mayoría de los municipios con ingresos altos y medio-altos se encuentra en la mitad norte de la región.



**Figura 7.** Mapa de la distribución de los tipos de ingresos en Andalucía

**Nota.** Creación propia con Atlas Interactivo Nacional de España a partir de los datos de SIMA.

### 4.3 Datos atípicos y extremadamente atípicos

En los diagramas de cajas y patas y tallos y hojas que se obtuvieron en la sección 4.1, se observa que existen datos atípicos y extremadamente atípicos en la distribución. En la Figura 8, se muestra la distribución de datos atípicos y extremadamente atípicos por provincias y se observa que la provincia que más datos atípicos contiene es Málaga seguida de Almería, mientras que Cádiz, Huelva y Jaén son las que albergan menos datos de este tipo. En cuanto a los datos extremadamente atípicos, se destaca la provincia de Málaga con gran diferencia, y Granada y Cádiz son las que registran menor número de esta tipología de datos.

Municipios con datos atípicos: Almócita, Bacares, Benitagla, Cantoria, Castro de Filabres, Cóbdar, Lucainena de las Torres, Nacimiento y Serón (Almería); Chipiona, Villaluenga del Rosario (Cádiz); Alcaracejos, Conquista, El Guijo, Fuente la Lancha, La Graniuela, Valsequillo (Córdoba); Juviles, Lobras, Soportújar, Escúzar (Granada); El Almendro, Cumbres de Enmedio (Huelva); Aldeaquemada, Jaén (Jaén); Alfarnatejo, Alpandeire, Árchez, Atajate, Benadalid, Benahavís, Benalauría, Benarrabá, Cartajima, Faraján, Genalgucil, Jimera de Líbar, Júzcar, Manilva, Parauta, Pujerra, Salares (Málaga); Carrión de los Céspedes, Huevar del Aljarafe, El Madroño, Las Navas de la Concepción (Sevilla).

Municipios con datos extremadamente atípicos: Bacares, Benitagla (Almería); Villaluenga del Rosario (Cádiz); Conquista, El Guijo, Valsequillo (Córdoba); Escúzar (Granada); Cumbres de Enmedio (Huelva); Jaén (Jaén); Alpandeire, Atajate, Benadalid, Cartajima, Faraján, Júzcar, Parauta, Salares (Málaga); Huevar del Aljarafe (Sevilla).

**Figure 8.** Distribución de los datos tipos y atípicos extremos por provincias

### 4.4. Comparación de medias

En esta sección se comparan las medias provinciales aplicando el test de la distribución T de Student. Para ello, se considera la Tabla 2, en la que se calcularon la media y la desviación típica de los ingresos por habitante en cada provincia y se indica además el número de municipios de cada una. Tal y como se puede observar, a partir de los datos de dicha tabla, la provincia de Málaga es la que tiene mayor media seguida de Córdoba; las que menos son Cádiz y Jaén. Estos datos concuerdan con el estudio de los datos atípicos y extremadamente atípicos realizado en la sección anterior.

Para analizar más en profundidad la comparación de las medias provinciales, se examinó si existen diferencias estadísticamente significativas (con un nivel de significación del 95%) entre las distintas parejas de provincias andaluzas. Los resultados obtenidos son que las únicas parejas entre las que existen tales diferencias son Almería y Málaga; Cádiz y Málaga; Córdoba y Málaga; Granada y Málaga; Huelva y Málaga; Jaén y Málaga; Málaga y Sevilla.

En las Tablas 3-9 se puede observar que se obtiene un p-valor menor que 0.05. Por tanto, se rechaza la hipótesis nula y existen diferencias significativas entre las provincias indicadas en cada caso. Los valores que se obtienen en dichas tablas son, respectivamente: 0.0004, 0.0008, 0.0025, 0, 0.0006, 0 y 0.

**Tabla 2.** Media, desviación estándar y número de municipios para la variable ingresos por habitante

Provincia	Media ingresos por habitante (EUR)	Desviación estándar	N.º de municipios
Almería	1492.70	757.22594	99
Cádiz	1370.41	614.59012	40
Córdoba	1517.28	766.45592	68
Granada	1501.02	577.14331	159
Huelva	1456.94	864.69447	78
Jaén	1386.64	558.44998	87
Málaga	1938.18	1037.65219	97
Sevilla	1395.19	814.84804	102

**Tabla 3.** Test t para las provincias de Almería (x)-Málaga (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	99	1492.7	76.10407	757.2259	1341.674	1643.726
Y	97	1938.18	105.3576	1037.652	1729.047	2147.313
Combinada	196	1713.167	66.5516	931.7224	1581.914	1844.42
Dif		-445.48	129.5635		-701.0138	-189.9462
Dif=media(x)-media(y) Ho: Dif=0			t=-3.4383 grados de libertad=194			
Ha: Dif < 0 Pr(T<t)=0.0004		Ha: Dif = 0 Pr( T > t )=0.0007		Ha: Dif > 0 Pr(T>t)=0.9996		

Nota. Creación propia con Stata a partir de los datos de SIMA

**Tabla 4.** Test t para las provincias de Cádiz (x)-Málaga (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	40	1370.41	97.17523	614.5901	1173.855	1566.965
Y	97	1938.18	105.3576	1037.652	1729.047	2147.313
Combinada	137	1772.408	82.63393	967.2054	1608.994	1935.821
Dif		-567.77	175.7502		-915.3499	-220.1901
Dif=media(x)-media(y) Ho: Dif=0			t=-3.2306 grados de libertad=135			
Ha: Dif < 0 Pr(T<t)=0.0008		Ha: Dif = 0 Pr( T > t )=0.0016		Ha: Dif > 0 Pr(T>t)=0.9992		

Nota. Creación propia con Stata a partir de los datos de SIMA,

**Tabla 5.** Test t para las provincias de Córdoba (x)-Málaga (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	68	1517.28	92.94643	766.4559	1331.758	1702.802
Y	97	1938.18	105.3576	1037.652	1729.047	2147.313
Combinada	165	1764.718	74.40495	955.7488	1617.803	1911.633
Dif		-420.9	147.9986		-713.1416	-128.6584
Dif=media(x)-media(y) Ho: Dif=0			t=-2.8439 grados de libertad=163			
Ha: Dif < 0 Pr(T<t)=0.0025		Ha: Dif = 0 Pr( T > t )=0.005		Ha: Dif > 0 Pr(T>t)=0.9975		

Nota. Creación propia con Stata a partir de los datos de SIMA

**Tabla 6.** Test t para las provincias de Granada (x)-Málaga (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	159	1501.02	45.77044	577.1433	1410.619	1591.421
Y	97	1938.18	105.3576	1037.652	1729.047	2147.313
Combinada	256	1666.663	50.6556	810.4896	1566.906	1766.419
Dif		-437.16	100.9655		-635.9961	-238.3239
Dif=media(x)-media(y) Ho: Dif=0			t=-4.3298 grados de libertad=254			
Ha: Dif < 0 Pr(T<t)=0.0000		Ha: Dif = 0 Pr( T > t )=0.0000		Ha: Dif > 0 Pr(T>t)=1.0000		

Nota. Creación propia con Stata a partir de los datos de SIMA

**Tabla 7.** Test t para las provincias de Huelva (x)-Málaga (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	78	1456.94	97.90737	864.6945	1261.982	1651.898
Y	97	1938.18	105.3576	1037.652	1729.047	2147.313
Combinada	175	1723.684	74.92759	991.1988	1575.801	1871.568
Dif		-481.24	146.6871		-770.7669	-191.7131
Dif=media(x)-media(y) Ho: Dif=0			t=-3.2807 grados de libertad=173			
Ha: Dif < 0 Pr(T<t)=0.0006		Ha: Dif = 0 Pr( T > t )=0.0013		Ha: Dif > 0 Pr(T>t)=0.9994		

Nota. Creación propia con Stata a partir de los datos de SIMA

**Tabla 8.** Test t para las provincias de Jaén (x) – Málaga (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	87	1386.64	59.87212	558.45	1267.618	1505.662
Y	97	1938.18	105.3576	1037.652	1729.047	2147.313
Combinada	184	1677.398	65.42656	887.489	1548.31	1806.485
Dif		-551.54	124.8848		-797.9482	-305.1318
Dif=media(x)-media(y) Ho: Dif=0			t=-4.4164 grados de libertad=182			
Ha: Dif < 0 Pr(T<t)=0.0000		Ha: Dif = 0 Pr( T > t )=0.0000		Ha: Dif > 0 Pr(T>t)=1.0000		

Nota. Creación propia con Stata a partir de los datos de SIMA

**Tabla 9.** Test t para las provincias de Málaga (x)-Sevilla (y)

	Observaciones	Media	Error estándar	Desviación típica	[Intervalo de confianza 95 %]	
X	97	1938.18	105.3576	1037.652	1729.047	2147.313
Y	102	1395.19	80.68198	814.848	1235.139	1555.241
Combinada	199	1659.864	68.5374	936.839	1524.707	1785.02
Dif		542.99	131.9097		282.8537	803.1263
Dif=media(x)-media(y) Ho: Dif=0			t=4.1164 grados de libertad=197			
Ha: Dif < 0 Pr(T<t)=1.0000		Ha: Dif = 0 Pr( T > t )=0.0001		Ha: Dif > 0 Pr(T>t)=0.0000		

Nota. Creación propia con Stata a partir de los datos de SIMA

#### 4.5 Test de normalidad

En esta sección se estudia si la variable objeto de estudio se ajusta a una distribución normal. Para ello se usan los tests de Kolmogórov-Smirnov y de Shapiro-Wilk. Ambos se usan para contrastar si un conjunto de datos sigue una distribución normal o no y están basados en un contraste de hipótesis. Se analizaron los resultados obtenidos con Stata en las Tablas 10 y 11 y como se puede comprobar en ambos casos, el p-valor obtenido es 0. Por tanto, se puede afirmar que la variable ingresos por habitante no sigue una distribución normal. Esto no ocurre con muchas otras variables (Akella, 2017).

**Tabla 10.** Test de Kolmogorov-Smirnov

Grupo	D	p-valor
Ingresos por habitante	0.1671	0.000
Acumulado	-0.1384	0.000
Test K-S	0.1671	0.000

Nota. Creación propia en Stata.

**Tabla 11.** Test de Shapiro-Wilk

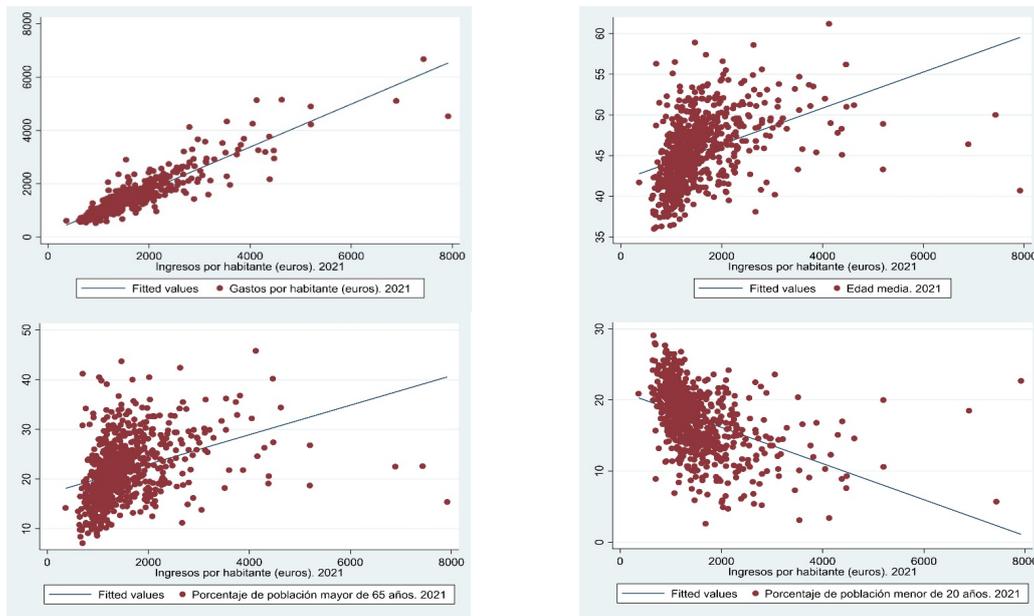
Variable	Observaciones	W	V	z	Prob > z
Ingresos por habitante	730	0.73243	126.753	11.835	0.00000

Nota. Creación propia en Stata.

### 4.6 Análisis bivariante

Esta sección se ocupa de la correlación entre la variable ingresos por habitante y las demás variables recogidas en la base de datos de SIMA. Para este fin, se generó una matriz de correlación (Figura 9) en la que se observa que las variables con las que los ingresos tienen mayor correlación son gastos por habitante, edad media, población menor de 20 años y población mayor de 65 años. Por su parte, la Figura 10 presenta los diagramas de dispersión relativos a estos análisis bivariantes, y la Tabla 12 muestra el cálculo del coeficiente de correlación para cada pareja de variables. Cabe agregar que estas correlaciones pueden ser útiles para predecir valores de otras variables.

	Ingresos	Gasto-20	Ext-2201	Per-2019	Altitud	Número-e	Dist-201	Po-12021	Poblac..	Poblac..	-eos2021
Ingresos-p	1.0000										
Gastospor-20	0.9076	1.0000									
Extensi-2201	-0.1123	-0.1197	1.0000								
Permet-2019	-0.1233	-0.1300	0.8026	1.0000							
Altitudsob-r	0.2142	0.2821	-0.0634	-0.0503	1.0000						
Númerod-e	-0.1091	-0.1083	0.4153	0.3839	-0.1541	1.0000					
Distanci-201	0.2442	0.2490	0.1472	0.1663	0.2844	0.0128	1.0000				
Poblacion-m	-0.0760	-0.0823	0.1414	0.1701	-0.2354	-0.2039	-0.2093	1.0000			
Poblacion-h	-0.0791	-0.0856	0.1449	0.1822	-0.2410	0.2113	-0.2121	0.9998	1.0000		
Poblacion-w	-0.0732	-0.0792	0.1382	0.1761	-0.2301	0.1971	-0.2067	0.9998	0.9993	1.0000	
Pobl-eos2021	-0.0738	-0.0794	0.1372	0.1757	-0.2302	0.1942	-0.2066	0.9996	0.9992	0.9997	1.0000
Pobl-dos2021	-0.1045	-0.1262	0.1942	0.1783	-0.2592	0.4195	-0.1647	0.3424	0.3525	0.3331	0.3172
Edadmed-2021	0.4166	0.4891	-0.0942	-0.0843	0.5034	-0.1058	0.3951	-0.1890	-0.1960	-0.1831	-0.1834
Porcent-norde	-0.4376	-0.5089	0.1991	0.1863	-0.5015	0.1337	-0.3834	0.1934	0.2000	0.1872	0.1879
Porcent-surde	0.3084	0.0100	-0.0100	-0.0104	0.0953	0.0277	0.1544	-0.1342	-0.1348	-0.1336	-0.1330
Variacion-r	0.2440	0.2904	-0.1254	-0.1308	0.3299	-0.0207	0.1625	-0.1281	-0.1299	-0.1264	-0.1273
Númerod-2021	-0.0544	-0.0581	0.1026	0.1277	-0.2405	0.2002	-0.1577	0.0862	0.8111	0.8014	0.8032
Porcent-j-c	0.0384	0.0100	-0.0100	-0.0104	0.0953	0.0277	0.1544	-0.1342	-0.1348	-0.1336	-0.1330
Emigra-2021	-0.0732	-0.0788	0.1019	0.1384	-0.2298	0.2181	-0.2182	0.9621	0.9623	0.9617	0.9617
Imigra-2021	-0.0710	-0.0781	0.1051	0.1393	-0.2487	0.2459	-0.2190	0.9376	0.9395	0.9356	0.9358
Nacimie-2021	-0.0800	-0.0803	0.1404	0.1838	-0.2439	0.2196	-0.2127	0.9953	0.9959	0.9945	0.9942
Defunci-2021	-0.0678	-0.0728	0.1460	0.1842	-0.2069	0.1764	-0.1788	0.9918	0.9908	0.9913	0.9920
Matrimo-2021	-0.0800	-0.0873	0.1363	0.1741	-0.2390	0.1921	-0.2159	0.9949	0.9943	0.9950	0.9947
Centrosde-I-e	-0.0725	-0.0790	0.1447	0.1811	-0.2274	0.1992	-0.2137	0.9915	0.9909	0.9917	0.9913
Centrosde-I-o	-0.0721	-0.0848	0.1434	0.2010	-0.2304	0.2159	-0.2040	0.9908	0.9903	0.9900	0.9902
Centrosde-I-a	-0.0741	-0.0782	0.1512	0.1950	-0.2081	0.1911	-0.1867	0.9848	0.9832	0.9859	0.9851
Centrosde-I-o	-0.0661	-0.0695	0.1400	0.1867	-0.2019	0.1747	-0.1936	0.9822	0.9800	0.9839	0.9833
Centrosde-I-u	-0.0673	-0.0702	0.1639	0.2011	-0.1921	0.1725	-0.1818	0.9626	0.9594	0.9651	0.9641
Centrosde-I-r	-0.0408	-0.0422	0.1228	0.1614	-0.1528	0.1337	-0.1643	0.9901	0.9855	0.9841	0.9827
Centrosde-I-s	-0.1376	-0.1433	0.2043	0.2446	-0.2304	0.1738	-0.1533	0.8847	0.8830	0.8860	0.8854
Bibliot-2020	-0.1273	-0.1428	0.2079	0.2284	-0.2097	0.2099	-0.1818	0.8497	0.8514	0.8479	0.8482
Centros-2021	-0.0721	-0.0750	0.1482	0.2116	-0.1966	0.1925	-0.1957	0.9608	0.9592	0.9630	0.9611
Consult-2021	-0.0286	-0.0137	0.2789	0.2710	0.0330	0.4981	0.0785	0.1531	0.1574	0.1490	0.1485
Vivriendas-f-s	-0.0698	-0.0750	0.1377	0.1764	-0.2176	0.1861	-0.1989	0.9977	0.9964	0.9985	0.9982
Transaccio-v	-0.0289	-0.0392	0.0772	0.1283	-0.1388	0.2008	-0.1491	0.7992	0.8020	0.7955	0.7977
Númerode-202	-0.0545	-0.0631	0.1064	0.1524	-0.2137	0.1643	-0.1642	0.9186	0.9111	0.9007	0.9105
Elecciones-c	-0.0741	-0.0801	0.1415	0.1798	-0.2227	0.1838	-0.2055	0.9969	0.9957	0.9976	0.9972
Elecciones-en	-0.0795	-0.0866	0.1509	0.1899	-0.2407	0.2082	-0.2037	0.9950	0.9952	0.9945	0.9941
Elecciones-ao	-0.0754	-0.0815	0.1424	0.1806	-0.2267	0.1887	-0.2082	0.9977	0.9968	0.9962	0.9977
Elecciones-en	-0.0769	-0.0840	0.1553	0.1946	-0.2452	0.2160	-0.2096	0.9933	0.9939	0.9924	0.9921
Elecciones-to	-0.0780	-0.0859	0.1290	0.1633	-0.2246	0.1845	-0.1882	0.9647	0.9638	0.9653	0.9650
Elecciones-r	-0.0684	-0.0743	0.1340	0.1724	-0.2239	0.1923	-0.2010	0.9969	0.9965	0.9969	0.9965
Superficie-h	-0.0911	-0.1138	0.1571	0.1517	-0.4537	0.0313	0.8134	0.8552	0.8552	0.8552	0.8552
Principal-re	-0.0773	-0.0901	0.2713	0.2747	-0.2067	0.1929	-0.1800	0.1012	0.1056	0.0972	0.0974
Principal-se	-0.0656	-0.0666	0.5108	0.3895	-0.1200	0.3846	0.0159	0.0573	0.0597	0.0550	0.0548
Superficie-l	-0.1303	-0.1295	0.4957	0.4968	0.0341	0.1678	0.0230	0.0672	0.0689	0.0657	0.0664
Principal-ga	-0.0772	-0.0776	0.2773	0.3302	0.0204	0.0989	-0.0470	0.0372	0.0382	0.0362	0.0373
Principal-ca	-0.1048	-0.0974	0.3920	0.3855	0.0858	0.0926	0.0651	0.0315	0.0323	0.0307	0.0310
Sinial-2021	-0.0638	-0.0672	0.1381	0.1738	-0.2121	0.2078	-0.1800	0.9817	0.9809	0.9811	0.9814
Hastasa-2021	-0.0662	-0.0705	0.1548	0.1935	-0.2133	0.1915	-0.1930	0.9844	0.9832	0.9833	0.9831
Entrefe-2021	-0.0585	-0.0622	0.1561	0.2042	-0.1849	0.1789	-0.1900	0.9810	0.9795	0.9821	0.9819
De20má-2021	-0.0593	-0.0552	0.1365	0.1730	-0.1946	0.1718	-0.1800	0.9852	0.9834	0.9805	0.9805
Totales-2021	-0.0634	-0.0677	0.1456	0.1834	-0.2110	0.2076	-0.1906	0.9854	0.9846	0.9857	0.9854
Vehiculost-l	-0.0770	-0.0840	0.1474	0.1845	-0.2376	0.2157	-0.2076	0.9983	0.9984	0.9979	0.9978
Autorizaci-a	-0.0237	-0.0242	0.0800	0.1164	-0.1295	0.0923	-0.1346	0.9560	0.9514	0.9600	0.9600
Autorizaci-e	-0.1189	-0.1251	0.2069	0.2323	-0.2560	0.3126	-0.1927	0.0637	0.0682	0.0593	0.0600
Autorizaci-l	-0.0278	-0.0285	0.0877	0.1215	-0.1284	0.0903	-0.1363	0.9450	0.9413	0.9408	0.9407
Vehiculos-l	-0.0488	-0.0606	0.1111	0.1482	-0.2332	0.2066	-0.1968	0.9559	0.9563	0.9551	0.9547
Vehiculos-t-s	-0.0407	-0.0544	0.1002	0.1358	-0.2229	0.1894	-0.1896	0.9269	0.9269	0.9266	0.9264
Oficinasde-o	-0.0623	-0.0650	0.1307	0.1517	-0.1916	0.1796	-0.1865	0.9832	0.9800	0.9840	0.9842
Consumodee-w	-0.0335	-0.0520	0.1527	0.1979	-0.2481	0.2362	-0.1800	0.8724	0.8732	0.8714	0.8721
Consumodee-s	-0.0595	-0.0669	0.1332	0.1702	-0.2244	0.2170	-0.2022	0.9792	0.9787	0.9794	0.9791
Hotels-2020	-0.0234	-0.0297	0.1242	0.1601	-0.1603	0.1752	-0.1301	0.8420	0.8389	0.8445	0.8418
Hostale-2020	-0.0222	-0.0300	0.1183	0.1395	-0.1703	0.1700	-0.1201	0.8323	0.8282	0.8350	0.8317
Plazasse-2020	-0.0109	-0.0150	0.0425	0.0686	-0.1953	0.1364	-0.1451	0.6657	0.6669	0.6644	0.6642
Plazassenho-s	-0.0246	-0.0306	0.1234	0.1414	-0.1672	0.1671	-0.1318	0.8804	0.8768	0.8835	0.8803
Par-bres2021	-0.0761	-0.0838	0.1404	0.1768	-0.2345	0.1976	-0.2062	0.9960	0.9959	0.9958	0.9954
Par-bres2021	-0.0732	-0.0803	0.1334	0.1703	-0.2289	0.1896	-0.2023	0.9959	0.9954	0.9961	0.9957
Paroregls-20	-0.0593	-0.0616	0.1415	0.1586	-0.2310	0.2382	-0.1768	0.8510	0.8556	0.8464	0.8490
Parotit-2021	-0.0745	-0.0816	0.1398	0.1751	-0.2354	0.2011	-0.2040	0.9967	0.9960	0.9962	0.9961
Tasamanl-202	-0.0117	-0.0660	0.0720	0.0566	-0.1571	-0.0439	0.0154	0.0904	0.0908	0.0900	0.8892
Contrat-eres	-0.0538	-0.0591	0.1743	0.2013	-0.2179	0.1737	-0.1919	0.9726	0.9718	0.9738	0.9731
Contrat-bres	-0.0745	-0.0802	0.1968	0.2312	-0.2303	0.1857	-0.1886	0.9578	0.9574	0.9578	0.9581
Contrat-sur-d	-0.0439	-0.0483	0.1290	0.1616	-0.1890	0.1786	-0.1749	0.9798	0.9781	0.9800	0.9807
Contrat-sur-e	-0.0676	-0.0733	0.1931	0.2241	-0.2294	0.1813	-0.1927	0.9665	0.9656	0.9670	0.9669
Contrat-sur-r	-0.0744	-0.0774	0.2274	0.2295	-0.2412	0.2307	-0.1775	0.6356	0.6401	0.6312	0.6316
Trabajador-s	-0.2021	-0.2185	0.3155	0.2919	-0.2715	0.3381	-0.0650	0.1102	0.1131	0.1076	0.1080
Presupuest-o	0.0236	0.0167	0.1573	0.1890	-0.2082	0.1958	-0.1926	0.9417	0.9408	0.9422	0.9418
Presupuest-s	-0.0439	-0.0479	0.1254	0.1654	-0.2047	0.1792	-0.1817	0.9856	0.9833	0.9874	0.9867
Númerod-2020	-0.0696	-0.0744	0.1383	0.1764	-0.2195	0.1887	-0.2082				



**Figura 10.** Diagramas de dispersión para las variables de mayor correlación con los ingresos por habitante

Nota. Creación propia en Stata.

**Tabla 12.** Coeficiente de correlación con la variable ingresos por habitante

Variable	Coefficiente de correlación
Gastos por habitante	0.9076
Edad media	0.4166
Población menor de 20 años	-0.4376
Población mayor de 65 años	0.3676

## 5. DISCUSIÓN

Se debe resaltar la gran oportunidad que brinda el Instituto de Estadística y Cartografía de la Junta de Andalucía al contar con la plataforma del Sistema de Información Multiterritorial de Andalucía, pues tiene una amplia recolección de datos de todos los municipios andaluces. Esto permitió llevar a cabo el estudio de los ingresos durante el año 2021 en esta comunidad autónoma, lo que se suma a la relevancia de otros autores le han dado a la gran labor de este instituto (Egea y Luque, 2018). Además, se han desarrollado otros análisis estadísticos basados en la plataforma SIMA (Cruces et al., 2009; Pita y Mateos, 2011).

Además de las referencias mencionadas en la sección 2, se han desarrollado otros estudios sobre la comunidad andaluza donde se plantean modelos que permitan una reducción del consumo (Román-Collado y Carmona, 2021) y del consumo de energía eléctrica (Jadraque et al., 2011), así como un análisis de la influencia de la temperatura en la demanda de electricidad (Gutiérrez y Arévalo-Quijada, 2006).

Tras el análisis realizado en la sección 2, se observa que, a lo largo de la historia, Andalucía ha experimentado cambios significativos con respecto a su economía. El sector industrial se ha diversificado abarcando muchas actividades económicas. La construcción ha crecido durante los últimos 30 años pasando a suponer un 12 %, a pesar de que este sector atravesó un declive durante la crisis de 2008. Estos cambios reflejan la dinámica económica y los desafíos a los que se ha enfrentado Andalucía en su búsqueda por fortalecer su potencial empresarial y promover un desarrollo sostenible. El sector de agricultura, ganadería y pesca se destaca en las provincias de Almería, Córdoba y Jaén. Las actividades inmobiliarias, artísticas y otros servicios destacan en las provincias de Cádiz, Granada, Málaga y Sevilla y el comercio es relevante en Huelva, Cádiz, Granada y Sevilla.

El análisis estadístico descriptivo reveló que el municipio con mayores ingresos por habitante es Huévar del Aljarafe (Sevilla). Se trata de un municipio de poco más de 3.200 habitantes situado en el Aljarafe sevillano. Al tratarse de una variable tan dependiente de las políticas territoriales, es difícil sacar conclusiones certeras de por qué hay municipios a los que se les asigna o disponen de más presupuesto que otros, independientemente de su tamaño y población. Otros de los municipios con datos inesperados para su tamaño y población son Cumbres de Enmedio (Huelva), Salares (Málaga), Atajate (Málaga), El Guijo (Córdoba), o Villaluenga del Rosario (Cádiz).

Se encontraron algunas limitaciones para el estudio estadístico derivadas de la base de datos SIMA. La más importante es el hecho de que para algunos municipios no había datos registrados para ciertas variables, razón por la cual no se contó con algunos datos a la hora de realizar determinados análisis bivariantes. Dichas ausencias de datos han sido muy pocas y, por tanto, despreciables. Otra limitación es la ausencia registros similares de otras comunidades autónomas para hacer una comparativa con distintas regiones de España o incluso dar una visión a nivel nacional de los ingresos.

## 6. CONCLUSIONES

En este artículo se analizó una base de datos para conocer los datos de ingresos por habitantes de cada uno de los municipios de Andalucía durante el año 2021. Los resultados obtenidos permiten concluir que en Andalucía existen desequilibrios con respecto al dinero destinado a cada habitante, tanto a nivel municipal como provincial. Se analizaron 730 municipios andaluces y se obtuvo una media de 1519.2 EUR y una desviación típica de 776.16. El mayor valor obtenido fue 7916.033 EUR, dato corresponde al municipio de Huévar del Aljarafe (Sevilla) y el menor 361.9793 EUR, que corresponde al municipio de Huétor Vega (Granada).

El estudio económico de empresas por sectores concluyó que, en Andalucía Oriental, como Almería y Jaén, hay una clara predominancia del sector agrícola-ganadero, sin embargo, en las provincias con más afluencia turística (Málaga, Sevilla y Cádiz), predomina el sector comercial. Si en el caso de Málaga se le suma que posee un potente sector de actividades inmobiliarias, hacen que sin duda sea la provincia con más actividad económica de Comunidad Autónoma. Provincias más interiores como Córdoba, tienen un sector industrial algo más desarrollado. Huelva y Granada, al contar tanto con zonas costeras como con zonas interiores de sierra, cuentan con sectores distribuidos de manera más uniforme. Los sectores más comunes en términos generales son el agrícola-ganadero, el comercial y el de actividades inmobiliarias.

Con respecto a los datos atípicos, Málaga es la provincia que más datos registra, mientras que Cádiz y Huelva son las que menos tienen. Es muy importante controlar los ingresos de los municipios correspondi-

entes a los datos atípicos, debido a que este tipo de datos influyen en el modelo de predicción que se utiliza para poder conjeturar o predecir valores futuros.

Los tests de normalidad permitieron afirmar que los datos correspondientes a la variable ingresos por habitante no se ajustan a una distribución normal.

Como perspectiva al futuro, queda pendiente un análisis similar sobre la renta declarada por habitante posterior al año 2020. Esto puede resultar de gran interés, pues durante la primavera del año 2020 se atravesó por una fase de confinamiento derivada de la pandemia del COVID-19 y, desde entonces, ha habido numerosos cambios a nivel social, empresarial y económico. Se pretende realizar una comparativa posteriormente con los datos esperados y los reales que aparezcan recogidos en SIMA. Aunque la variable renta media declarada no tenga una correlación elevada con la de este estudio, puede resultar interesante analizar los ingresos de las regiones desde otra perspectiva. Otra variable de índole social y económica que podría ser estudiada es la tasa municipal de desempleo.

Se espera que el análisis estadístico llevado a cabo en el presente artículo ayude a comprender mejor cómo están distribuidos los ingresos en Andalucía y cuáles son los sectores de actividad predominantes. Quizás pueda ser de utilidad a personas que estén dirimiendo dónde invertir su dinero o a personas que quieran emprender un nuevo negocio.

#### **CONTRIBUCIONES DE LOS AUTORES**

Ambos autores contribuyeron equitativamente a la realización de todas las secciones del artículo y coordinaron en todo momento la conceptualización, metodología, validación, análisis, redacción-borrador original, redacción-revisión y edición. Todos los autores han leído y aceptado la versión publicada del manuscrito.

#### **FINANCIACIÓN**

Esta investigación no recibió financiación externa.

#### **DECLARACIÓN DE DISPONIBILIDAD DE DATOS**

El estudio realizado en el presente artículo utilizó datos recogidos en la plataforma del Sistema de Información Mutiterritorial de Andalucía del Instituto de Estadística y Cartografía de Andalucía. Dichos datos se pueden consultar en la web <https://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/info.htm?f=k03>

#### **AGRADECIMIENTOS**

Los autores querrían resaltar la gran oportunidad brindada por el Instituto de Estadística y Cartografía de la Junta de Andalucía al tener en su página web la plataforma del Sistema de Información Multiterritorial de Andalucía con una amplia recolección de datos de todos los municipios de Andalucía. Gracias a dichos datos fue posible llevar a cabo este estudio.

#### **CONFLICTOS DE INTERESES**

Los autores declaran no tener ningún conflicto de intereses.

## **7. REFERENCIAS**

- Álvaro, C.F., Daniel, D. y Espinosa, I. (2013). El Instituto de Estadística y Cartografía de Andalucía. XIV Congreso de Enseñanza y Aprendizaje de las Matemáticas.
- Akella, S., Díaz, P. & Baluguri, S. (2017). Exam Results Fit to a Normal Distribution. *DJ Journal of Engineering and Applied Mathematics*, 4(1), 20-28. <https://doi.org/10.18831/djmaths.org/2018011002>
- Bernardi, F. y Escobar Mercado, M. (2012). *Análisis de datos con Stata* (2a. ed.). CIS.
- Campoy-Muñoz, P., Cardenete, M. y Delgado, M. (2015). Análisis estructural a través de Matrices de Contabilidad Social: una aplicación a la economía andaluza para el período 2005-2010. *Revista Perspectiva Socioeconómica*, (1), 7-28. <https://doi.org/10.21892/24627593.211>
- Cardenete, M. A. y López, J. M. (2015). Análisis de sectores clave a través de Matrices de Contabilidad Social. El caso de Andalucía. *Estudios de Economía Aplicada* 33(1), 203-222.
- Congregado, E. y García, J. I. (2002). El problema de desempleo en la economía andaluza (1990-2001): análisis de la transición desde la educación al

- mercado laboral. Economic Working Papers at Centro de Estudios Andaluces, E2002/17.
- Conserjería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía. (2018). Análisis de las Cadenas Globales de Valor en Andalucía. Junta de Andalucía.
- Cruces, E.M., de Haro, J. y Sarrión, M.D. (2009). Análisis estadístico de la realidad socioeconómica en Andalucía. Una aproximación a escala municipal. *Investigaciones Regionales*, (18), 107-138.
- Egea, R. y Luque, P.L. (2018). Instituto de Estadística y Cartografía de Andalucía (IECA): fuentes de datos [Trabajo fin de grado]. Universidad de Sevilla. <https://idus.us.es/handle/11441/77593>
- Gutiérrez, E. y Arévalo-Quijada, M. (2006). Análisis de la influencia de la temperatura en la demanda residencial de energía eléctrica en Andalucía. *Revista de Estudios Andaluces*, 26, 93-122. <https://doi.org/10.12795/rea.2006.i26.04>
- Instituto de Estadística y Cartografía de Andalucía. (2017). Análisis de las Cadenas de Valor en Andalucía: Propuesta Metodológica y Estimación de la Participación Regional.
- Instituto de Estadística y Cartografía de Andalucía. (2021). Informe Económico de Andalucía 2021 y Andalucía en cifras. Instituto de Estadística y Cartografía de Andalucía.
- Instituto de Estadística y Cartografía de Andalucía. (2021). Andalucía Pueblo a pueblo. Fichas municipales: Ingresos por habitante (euros). Junta de Andalucía. <https://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/info.htm?f=k03>
- Jadraque, E., García, J. & Estrella, A. (2011). Development of an energy model for the residential sector: Electricity consumption in Andalusia, Spain. *Lancet* 43,(6) 1315-1321. <https://doi.org/10.1016/j.enbuild.2011.01.016>
- Junta de Andalucía (2023). Sistema de Información Multiterritorial de Andalucía. <https://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/index2.htm>
- Kano, L., Tsang, E.W.K., & Yeung, H.Wc. (2020). Global value chains: A review of the multi-disciplinary literature. *Journal of International Studies*, (51), 577-622.
- Pérez, M., Sánchez, I. y Espasandín, F. (2017). La propensión exportadora de la economía social andaluza: análisis clúster. *REVESCO. Revista de Estudios Cooperativos*, (125), 159-186.
- Pita, M. y Mateos, B. (2011). Tercer Informe de Desarrollo Territorial de Andalucía. Secretariado de Publicaciones de la Universidad de Sevilla.
- Román-Collado, R. & Carmona, M. (2021). Energy efficiency's key role in explaining the performance of energy consumption in Andalusia (Spain). *Environmental Science and Pollution Research*, (28), 1-21. <https://doi.org/10.1007/s11356-020-11829-2>.
- Tejedor, M. (2022). Más allá del acento: la otra historia de la economía andaluza [Trabajo fin de grado inédito]. Universidad de Sevilla.
- Universidad Loyola (2021). Guía de Stata.