



# Minería de Datos

Estudio sobre energía renovable en los hogares españoles

## Entregable 1

**Integrantes y porcentaje de trabajo:**

Alumno	Porcentaje de trabajo
Elena Ballesteros Morallón	20 %
Francisco Javier Luna Ortiz	20 %
Antonio Gómez Jimeno	20 %
Sergio Herreros Fernández	20 %
Pedro Sánchez Martín	20 %

## Tabla de contenidos.

1. Introducción.....	4
2. Alcance.....	4
3. Descripción de los datos.....	4
3.1. Análisis exploratorio de datos.....	5
I. Consumo eléctrico.....	5
II. Información sobre las viviendas.....	5
III. Dispositivos de energía renovable.....	6
IV. Capacidad económica de la población.....	6
V. Distribución de edad de la población.....	6
VI. Producción de energía fotovoltaica (PV).....	6
3.2. Análisis detallado.....	7
I. Consumo eléctrico anual.....	7
II. Viviendas según número personas.....	7
III. Dispositivos energía renovable.....	7
IV. Intensidad uso viviendas.....	8
V. Distribución edad población.....	9
VI. Salario medio población.....	9
VII. Producción energía por lugar.....	9
VIII. Tipo núcleo familiar.....	10
IX. Renta Hogar.....	10
4. Trabajos similares y estudio de viabilidad.....	11
5. Objetivos e hipótesis .....	11

5.1.	Objetivo .....	11
5.2.	Hipótesis.....	12
I.	Hipótesis Descriptivas.....	12
II.	Hipótesis Predictiva.....	13
6.	Enriquecimiento de los datos .....	13
7.	Metodología seguida.....	14
I.	Selección de la Temática.....	14
II.	Recopilación de Datos.....	14
III.	Definición de Hipótesis y Objetivos .....	14
8.	Anexo.....	15
I.	Diccionario de datos.....	15
II.	Bibliografía .....	21

## 1. Introducción

El precio de la luz es un gasto importante en todos los hogares y en los últimos años se ha observado un aumento en el consumo y el precio de la electricidad. Tanto es el impacto que mucha gente ha optado por abastecerse ellos mismos por medio de energías renovables, que gracias a los avances tecnológicos y a la nueva legislación son cada vez más accesibles.

Este estudio se plantea desde el punto de vista de una empresa de instalación de sistemas para el autoconsumo energético. El objetivo es hacer un estudio de mercado de los potenciales clientes según las características demográficas y geográficas.

## 2. Alcance

El volumen de datos del número de dispositivos de energía renovable es escaso, en España a nivel de municipios o a nivel europeo, en algunos países. Por ello el estudio se centrará en los datos de España, concretamente a nivel provincial ya que no hay información de todos los municipios, solo los de más de 50.000 habitantes, y los resultados estarían sesgados y no se podrían extrapolar a nivel nacional. Se referirán como provincias, aunque en realidad no lo sean, las ciudades autónomas de Ceuta y Melilla.

El estudio de estas últimas, junto con las islas Baleares y Canarias queda pendiente de ser evaluado en el preprocesamiento y limpieza de los datos (correspondientes al entregable 2). Ya que podría suponer un volumen de datos desparejo a la península, una inversión mayor por desplazamiento, además por su posición geográfica, estas regiones cuentan con un sistema de red eléctrica distinto.

En cuanto a las tecnologías renovables, no nos centraremos en una tecnología en concreto, como podría ser la solar fotovoltaica, ya que los datos están agregados en varias tecnologías. Por ello, se tratarán todas las energías renovables que puedan instalarse en una residencia.

## 3. Descripción de los datos

En esta sección se presentará el análisis exploratorio de los datos necesarios para cumplir con los objetivos de nuestro proyecto. El propósito es definir los conjuntos de

datos que utilizaremos, así como las preguntas clave que estos deben ayudar a responder. Además, se realizará un análisis detallado de cada dataset, describiendo su contenido, relevancia y cómo contribuye a obtener respuestas específicas a nuestras interrogantes.

### **3.1. Análisis exploratorio de datos**

Nos interesa utilizar datos que incluyan información acerca de los siguientes aspectos:

- Consumo eléctrico.
- Información sobre las viviendas.
- Dispositivos de energía renovable.
- Capacidad económica de la población.
- Distribución de edad de la población.
- Producción de energía fotovoltaica (PV).

Nos centraremos en España, utilizando datasets divididos por provincias y/o comunidades autónomas como eje central y relacional de nuestra base de datos.

Para identificar los datasets idóneos, es fundamental definir claramente las preguntas que buscamos responder sobre cada aspecto de nuestro proyecto. Al establecer estas preguntas clave, podremos seleccionar conjuntos de datos que proporcionen la información necesaria, asegurando así que nuestro estudio esté basado en datos relevantes y significativos.

#### **I. Consumo eléctrico.**

Preguntas clave:

- ¿Cuál es el consumo eléctrico anual por comunidad autónoma?
- ¿Cómo ha evolucionado el consumo eléctrico en cada comunidad autónoma a lo largo de los años?

#### **II. Información sobre las viviendas.**

Preguntas clave:

- ¿Cómo se distribuyen los hogares según su tamaño (número de integrantes)?
- ¿Cómo se distribuye el consumo eléctrico por vivienda?
- ¿Cuál es la distribución de los hogares según el tipo de núcleo familiar?

### **III. Dispositivos de energía renovable.**

Preguntas clave:

- ¿Qué provincias tienen una mayor proporción de hogares que utilizan dispositivos de energía renovable?
- ¿Qué hogares tienen dispositivos de energía renovable en función de los ingresos familiares en cada provincia?
- ¿En qué zonas, rurales o urbanas se concentra una mayor adopción de dispositivos de energía renovable?
- ¿Cuáles son los tipos de dispositivos de energía renovable disponibles, cuántos hogares utilizan cada uno y si se implementan combinados?

### **IV. Capacidad económica de la población.**

Preguntas clave:

- ¿Cómo varían los salarios medios brutos mensuales entre las distintas comunidades autónomas?
- ¿Existen diferencias entre los salarios de empleados a tiempo completo y a tiempo parcial?
- ¿Cómo ha evolucionado la renta neta media por hogar en España y en cada comunidad autónoma entre 2020 y 2022? (Tener en cuenta también la renta media por hogar con alquiler imputado).

### **V. Distribución de edad de la población.**

Preguntas clave:

- ¿Cómo se distribuye la población por rangos de edad en las diferentes provincias?
- ¿Cómo ha evolucionado la distribución de la población por rangos de edad en España?

### **VI. Producción de energía fotovoltaica (PV).**

Preguntas clave:

- ¿Cuál es la cantidad de energía fotovoltaica producida en un mes específico para cada comunidad autónoma?

- ¿Qué cantidad de energía se generó en una provincia específica durante un mes determinado?

## **3.2. Análisis detallado**

A continuación, se enumerarán los datasets seleccionados para nuestro proyecto y se ofrecerá una descripción detallada de cada uno.

Los datasets cuyos datos se presentan divididos por comunidad autónoma son: Consumo eléctrico anual, Salario medio población y Renta hogar.

El resto de datasets se centran en dividir los datos por provincias.

### **I. Consumo eléctrico anual.**

Este conjunto de datos presenta el consumo eléctrico anual en España, desglosado por comunidad autónoma, para los años 2015, 2017, 2019, 2021 y 2022.

Este dataset es útil para analizar cuáles son las comunidades autónomas con mayor consumo de energía eléctrica, así como para estudiar la relación entre el consumo y la adopción de energías renovables en los hogares.

### **II. Viviendas según número personas.**

Este dataset proporciona información sobre el número de viviendas en cada comunidad autónoma y en cada una de sus provincias, según el número de personas que residen en ellas (1, 2, 3, 4 y 5 personas o más). Los datos abarcan el periodo desde 2020 hasta 2023.

A través de esta información, podemos determinar si en determinadas provincias las familias con un cierto número de integrantes invierten más o menos en energía renovable.

### **III. Dispositivos energía renovable.**

En este dataset se analiza la cantidad de dispositivos de energía renovable por hogar, teniendo en cuenta el tipo de dispositivo de energía renovable (un único tipo de dispositivo: energía solar térmica, un único tipo de dispositivo: energía solar

fotovoltaica, eólica, biomasa, etc., combinación de varios tipos de dispositivos, no dispone de dispositivo de energía renovable y el total).

Además, se tiene en cuenta el tamaño del municipio (entre 50.000 habitantes o menos y más de 50.000 habitantes).

Toda esta información se encuentra desglosada por provincias y en función de los ingresos netos mensuales por hogar (menos de 1000 euros, de 1000 euros a menos de 1500 euros, de 1500 euros a menos de 2000 euros, de 2000 euros a menos de 3000 euros y 3000 euros o más).

Este dataset presenta información del año 2021.

#### **IV.Intensidad uso viviendas.**

En este conjunto de datos se determina el número de viviendas en cada provincia según varios factores.

- Viviendas totales.
- Viviendas vacías.
- Viviendas con bajo consumo.
- Viviendas de uso esporádico (total y desglosado en rangos).
  - 251 a 500 kwh.
  - 501 a 750 kwh.
- Viviendas según niveles de consumo (desglosado en rangos).
  - 751 a 1.000 kwh.
  - 1.001 a 2.000 kwh.
  - 2.001 a 3.000 kwh.
  - 3.001 a 4.000 kwh.
  - 4.001 a 5.000 kwh.
  - 5.001 a 6.000 kwh.
  - 6.001 a 7.000 kwh.
  - 7.001 a 8.000 kwh.
  - 8.001 a 9.000 kwh.
  - 9.001 a 10.000 kwh.
  - Consumo mayor de 10.000 kwh.

Además, se incluye como dato adicional la mediana del consumo anual por provincia.



Este dataset presenta información de un censo de población y viviendas de 2021.

## **V.Distribución edad población.**

Este dataset abarca información desde el año 2015 hasta 2022 y proporciona un análisis tanto a nivel general como por provincias. Se establece el total de habitantes, así como la distribución de estos por grupos de edad de cinco años, abarcando desde 0 hasta 100 años y más.

Este dataset es una herramienta valiosa para estudiar la distribución de la población en España y su edad. Al analizar los datos, podemos identificar las provincias con mayor población, así como aquellas con una población más joven. Esta información permite comprender las dinámicas demográficas del país.

## **VI.Salario medio población.**

Este conjunto de datos aborda los salarios medios brutos mensuales del empleo totales y en cada comunidad autónoma. La información se presenta desglosada en tres categorías: el total de salarios, los salarios correspondientes a jornadas a tiempo completo y los salarios de aquellos empleados a tiempo parcial.

Además de proporcionar estos totales, los datos se dividen en deciles. Los deciles son medidas estadísticas que dividen un conjunto de datos en diez partes iguales, permitiendo así analizar la distribución de los salarios. Cada decil representa el umbral bajo el cual se encuentra un porcentaje específico de la población. Esta desagregación permite comprender mejor la desigualdad salarial y las variaciones en los ingresos en las diferentes comunidades autónomas.

Los datos recogidos corresponden a los años 2021 y 2022.

## **VII.Producción energía por lugar.**

Este dataset proporciona información sobre la producción de energía fotovoltaica (PV) en diferentes localizaciones de España y está compuesto por varias columnas clave: la posición geográfica, el mes (indica el periodo en el que se registró la producción), la cantidad de energía PV producida en kilovatios hora (kWh), la comunidad autónoma y la provincia.

Estos datos fueron obtenidos a través de web scraping, una técnica que permite extraer automáticamente información de sitios web.

Este dataset presenta información actualizada.

### **VIII. Tipo núcleo familiar.**

Este dataset muestra la cantidad de hogares según el tipo de núcleo familiar, tanto a nivel nacional como por provincias, en 2021. La información se organiza considerando varios aspectos, dividiendo los datos en dos dimensiones: el tipo de núcleo familiar y la cantidad de hijos que viven en el núcleo familiar.

En cuanto al tipo de núcleo familiar, se incluyen las siguientes categorías (total de núcleos familiares, pareja casada con o sin hijos, pareja no casada con o sin hijos, madre con hijos y padre con hijos).

Por otro lado, en relación con la cantidad de hijos que viven en el núcleo familiar, se desglosa la información en: total de hijos conviviendo, 0 hijos, 1 hijo y 2 o más hijos. Además, se presenta la misma categorización, pero especificando que sean menores de 25 años.

### **IX. Renta Hogar.**

El siguiente dataset presenta la renta por hogar, tanto a nivel total en España como desglosada por comunidades autónomas, correspondiente a los años 2020, 2021 y 2022. Este conjunto de datos incluye dos variables clave: la renta neta media por hogar y la renta media por hogar, teniendo en cuenta el alquiler imputado.

El alquiler imputado se aplica a los hogares que no pagan un alquiler completo por ser propietarios o por ocupar una vivienda alquilada a un precio inferior al de mercado o a título gratuito. El valor que se imputa es el equivalente al alquiler que se pagaría en el mercado por una vivienda similar a la ocupada, menos cualquier alquiler realmente abonado. Asimismo, se deducen de los ingresos totales del hogar los intereses de los préstamos solicitados para la compra de la vivienda principal.

## 4. Trabajos similares y estudio de viabilidad

Como ya se ha mencionado anteriormente, la electricidad representa un gran porcentaje del gasto mensual de los hogares. El coste de generar esta electricidad está influenciado por el origen de las fuentes de energía: combustibles fósiles, gas natural o fuentes renovables. De esta forma, el precio también influye en la demanda, que será diferente según las características de la vivienda, las particularidades de sus inquilinos (el número, salario, edad o género) y el clima de la zona.

Según (Bueno et al., 2020) los precios de la electricidad han tenido un efecto negativo, es decir, cuando suben los precios, la demanda baja. También influye si es un hogar monoparental o con residentes extranjeros. En cambio, la demanda aumenta con mayores ingresos, más horas de sol y aumentos de la temperatura, aunque estas variables tienen menos influencia.

Los usos de esa electricidad son muy variados, desde el uso de dispositivos electrónicos, electrodomésticos hasta sistemas de calefacción y refrigeración. En una encuesta desarrollada en 2015, (Ortega-Izquierdo et al., 2019) identificaron que, en la elección de estos sistemas, pocos usuarios preferían las energías renovables. También, que factores sociales como la edad, el género o la educación de los encuestados no estaban relacionados con el resultado, mientras que el tipo de edificio o los ingresos sí.

El propósito de este estudio es tener una visión más actualizada de la realidad de la implantación de las energías renovables en los hogares españoles. A continuación, se presenta una explicación más desarrollada de los objetivos.

## 5. Objetivos e hipótesis

### 5.1. Objetivo

El objetivo principal es analizar el uso y la rentabilidad de las energías renovables en cada provincia de España.

A través de la aplicación del proceso *KDD*, se busca identificar los perfiles demográficos y socioeconómicos de las personas que utilizan dispositivos de energía renovable, considerando variables como la edad, la situación económica y familiar y el lugar de residencia. Además, se tendrán en cuenta las necesidades eléctricas de los hogares, lo que permitirá comprender mejor cómo estos factores influyen en la adopción de dispositivos de energía renovable.

También, se pretende evaluar las regiones de España dónde existe una mayor tendencia a la adopción de este tipo de tecnologías, así como las condiciones climáticas que pueden influir en su uso.

## 5.2. Hipótesis

A continuación, se plantean una serie de hipótesis de distinta naturaleza sobre el estudio realizado.

### I. Hipótesis Descriptivas.

Las hipótesis descriptivas que se plantean en este estudio buscan explorar las relaciones entre diferentes variables demográficas, socioeconómicas y geográficas y la adopción de energías renovables y su rendimiento en las distintas provincias de España. Son las siguientes:

1. *Las provincias con una edad media menor y una renta media por hogar mayor a la nacional, suelen estar más concienciadas con el uso de energías renovables y utilizan más dispositivos que aprovechan este tipo de energía.*

La justificación de esta hipótesis es que las personas jóvenes suelen estar más concienciadas con el medio ambiente, además de que, tienen más tiempo para amortizar la inversión de este tipo de tecnología a lo largo de su vida.

2. *Las provincias con mayor proporción de viviendas con bajo consumo eléctrico (por debajo de 3.000 kWh) son más propensas a adoptar energías renovables que aquellas con mayor proporción de consumo alto (8.000 kWh).*

Las viviendas con menor consumo eléctrico necesitan menos infraestructura para generar la electricidad que demandan. Es por ello que invertir en este tipo de tecnologías para viviendas pequeñas puede ser una opción a tener en cuenta. Además, las personas con viviendas de alta demanda energética suelen tener un poder adquisitivo mayor, y, generalmente, prefieren contratar la electricidad a lidiar con los problemas técnicos de este tipo de dispositivos.

3. *Las provincias que tienen mayor porcentaje de familias de 4 o más integrantes invierten más en energía renovable.*

Aunque en principio esta hipótesis parece entrar en contradicción con la anterior, ya que a mayor número de personas viviendo en un hogar se supone que habrá mayor consumo.

Es interesante comprobar si la anterior es falsa, si el número de personas viviendo es la justificación o en caso contrario, si es verdadera, qué relación hay entre ambas.

4. *Las provincias situadas en el sur de España (aquellas al sur de Madrid) tienden a utilizar menos dispositivos de aprovechamiento de energías renovables que las del norte.*

Aunque las regiones del sur de España cuentan con un mayor potencial solar, su menor consumo eléctrico asociado al calentamiento de los hogares influye en la menor adopción de dispositivos de energía renovable. Por otra parte, las provincias del norte, que suelen tener climas más fríos, suelen gastar más en calefacción, lo que genera una mayor necesidad de invertir en esta tecnología para reducir sus costos.

## II. Hipótesis Predictiva.

La hipótesis predictiva que se plantea en este estudio busca anticipar el consumo eléctrico de una de las Comunidades Autónomas de España en el año 2028. La hipótesis es la siguiente:

5. *¿Qué cantidad de placas solares fotovoltaicas se necesitarían instalar para abastecer el consumo eléctrico de Castilla-La Mancha en el año 2028?*

Esta hipótesis es interesante para conocer cuál es el verdadero potencial y la rentabilidad de uno de los dispositivos de aprovechamiento de energías renovables más comunes y utilizados actualmente, las placas solares fotovoltaicas.

## 6. Enriquecimiento de los datos

El enriquecimiento de datos es el proceso de incorporar información adicional a los datos existentes, complementándolos para hacerlos más descriptivos y útiles. En este caso, al no partir de un único dataset con suficiente información para generar las hipótesis, si no que a partir de un objetivo se han buscado datasets para unirlos. Por lo que, es difícil diferenciar entre información básica y aquella que la complementa.

Aun así, se podría incluir como información complementaria la agrupación de provincias según la comunidad autónoma a la que pertenecen en caso de que se amplie alguna hipótesis a nivel regional.

## 7. Metodología seguida

En este apartado se detalla la metodología seguida para llevar a cabo este proyecto. El proceso se desarrolló en varias fases, que son las siguientes:

### I. Selección de la Temática

Inicialmente, se llevó a cabo una búsqueda entre todos los integrantes del grupo para seleccionar la temática del estudio. Se analizaron distintas propuestas, aunque optamos por las energías renovables.

### II. Recopilación de Datos

Una vez definida la temática, se procedió a recopilar datos de distintas fuentes oficiales. Algunas de estas fuentes son el *Instituto Nacional de Estadística* (INE) y el *Centro Común de Investigación* (JRC) de la *Comisión Europea*.

Aunque la gran mayoría de datos se han obtenido de datasets preexistentes, otros se han obtenido mediante técnicas como *web scraping*.

Además, todos los datos recopilados para este documento y la explicación de estos se han incluido en este documento (Ver apartados 3 y 8.I).

### III. Definición de Hipótesis y Objetivos

Por último, teniendo en cuenta los datos disponibles, se formularon las hipótesis y los objetivos del estudio, ya que, debido a la dificultad de encontrar datos útiles y suficientemente detallados, fue necesario ajustar las hipótesis iniciales que se plantearon en la fase de *Selección de la Temática*.

## 8. Anexo

### I. Diccionario de datos

#### Consumo eléctrico anual

Nombre del campo	Tipo de dato	Descripción
Comunidades Autónomas y Ciudad Autónomas	String	Nombre de la comunidad o ciudad autónoma de España donde se registra cierto consumo eléctrico anual.
Total (Consumo eléctrico anual)	Float	Consumo eléctrico en miles de euros registrado en la comunidad autónoma para un año específico.
año	Int	Año del registro del consumo eléctrico (2015-2022)

#### Viviendas según el número de personas

Nombre del campo	Tipo de dato	Descripción
Comunidades y Ciudades Autónomas	String	Nombre de la comunidad o ciudad autónoma donde se registra el número de viviendas según el número de personas.
Provincias	String	Nombre de la provincia donde se registra el número de viviendas según el número de personas.
Tamaño del hogar	String	Clasificación de viviendas por número de personas residiendo. Por ejemplo: "1 persona", "5 personas o más", "Total".
Año (2020-2023)	Int	Año del registro del número de vivienda en cada comunidad autónoma.

Total	Int	Número total de viviendas en las que viven cierta cantidad de personas dependiendo del tamaño del hogar.
-------	-----	--

### Dispositivos de energía renovable

Nombre del campo	Tipo de dato	Descripción
Provincias	String	Nombre de la provincia española en la que se registra cierto total de dispositivos de energía renovable.
Nivel de ingresos mensuales netos del hogar	String	Intervalo de ingresos netos mensuales de los hogares, clasificados en tramos (por ejemplo, "Menos de 1.000 euros", "De 1.500 euros a menos de 2.000 euros", etc.).
Tamaño del municipio	String	Clasificación de datos para diferenciar municipios de 50.000 habitantes o menos y aquellos que tienen más.
Tipo de dispositivo de energía renovable	String	Categorización para distinguir qué dispositivo de energía renovable se utiliza. Así como: "No dispone de dispositivo de energía renovable", "Un único tipo de dispositivo: Energía solar térmica", etc.
Total	Int	Número total de tipo de dispositivos renovables utilizados.

### Intensidad de uso en viviendas

Nombre del campo	Tipo de dato	Descripción
Comunidades y Ciudades Autónomas	String	Nombre de la comunidad o ciudad autónoma en España a la que pertenece la provincia.



Provincias	String	Nombre de la provincia española con cierto consumo eléctrico y el número de viviendas con tal consumo.
Consumo eléctrico	String	<p>Clasificación de viviendas según su consumo para cada comunidad autónoma. Puede tomar los siguientes valores:</p> <ul style="list-style-type: none"> <li>- Viviendas totales</li> <li>- Viviendas vacías</li> <li>- Mediana consumo anual (kWh)</li> <li>- Viviendas con bajo consumo</li> <li>- Viviendas de uso esporádico</li> <li>- Viviendas de uso esporádico (entre 251 y 500 kWh)</li> <li>- Viviendas de uso esporádico (entre 501 y 750 kWh)</li> <li>- Viviendas con consumo entre 751 y 1.000 kWh</li> <li>- Viviendas con consumo entre 751 y 1.000 kWh</li> <li>- Viviendas con consumo entre 2.001 y 3.000 kWh</li> <li>- Viviendas con consumo entre 3.001 y 4.000 kWh</li> <li>- Viviendas con consumo entre 4.001 y 5.000 kWh</li> <li>- Viviendas con consumo entre 5.001 y 6.000 kWh</li> <li>- Viviendas con consumo entre 6.001 y 7.000 kWh</li> <li>- Viviendas con consumo entre 7.001 y 8.000 kWh</li> <li>- Viviendas con consumo entre 8.001 y 9.000 kWh</li> <li>- Viviendas con consumo entre 9.001 y 10.000 kWh</li> <li>- Viviendas con consumo mayor de 10.000 kWh</li> </ul>
Total	Int	Número total de viviendas según su intensidad de uso y provincia.

**Distribución de edad de la población española**

Nombre del campo	Tipo de dato	Descripción
Comunidades y Ciudades Autónomas	String	Nombre de la comunidad o ciudad autónoma de España a la que pertenecen los datos de edades.
Provincias	String	Nombre de la provincia española a la que pertenecen los datos de edades.
Edad (grupos quinquenales)	String	Edad distribuida de en grupos quinquenales (para la distribución de la población). Ejemplos: “0-4 años”, “55-59 años”, etc.
Periodo	String	Año del registro de edades de la población (2015-2022). Ejemplo: “1 de enero de 2016”
Total	Int	Número total de personas por comunidad autónoma (o en España) que su edad se encuentra en un grupo quinquenal específico, en cierto año.

**Salario medio de la población española**

Nombre del campo	Tipo de dato	Descripción
Tipo de jornada	String	Si el salario medio es un dato de “Jornada a tiempo completo”, “Jornada a tiempo parcial” o “Total”
Comunidades y Ciudades Autónomas	String	Nombre de la comunidad o ciudad autónoma en España con los datos salariales.
Decil	String	La población se divide en diez subsecciones de igual tamaño. El primer decil separa los salarios medios entre el 10% de los valores inferiores y el resto de los datos. El segundo decil separa los salarios medios entre el 20%

		de valores inferiores y el 80% de los valores superiores. Así consecutivamente.
Total	Float	Valor del salario medio según el decil de la población censada.
Periodo	Int	Año en el que se ha deducido el salario medio, ya sea en cada decil o en total (2021-2022).

### Producción de energía por lugar

Nombre del campo	Tipo de dato	Descripción
Latitud	Float	Localización de la provincia en dirección Norte o Sur desde el ecuador.
Longitud	Float	Localización de la provincia en dirección Este u Oeste desde el meridiano de Greenwich.
Mes	String	Mes del año 2022 en el que se registra la cantidad de producción de energía PV.
Energía PV producida en kWh	Float	Energía fotovoltaica (PV) producida en kilovatio hora (kWh) en diferentes provincias de España y en cierto mes.
Comunidad Autónoma	String	Comunidad autónoma de la provincia de la cual se tiene el dato de la energía PV producida.
Provincia	String	Provincia de España del cual se tiene el dato de la energía PV producida.

### Tipo de núcleo familiar

Nombre del campo	Tipo de dato	Descripción
------------------	--------------	-------------

Provincias	String	Nombre de la provincia en la que se registran los datos poblacionales.
Tipo de núcleo familiar	String	Tipo de núcleo familiar (pareja casada, pareja no casada, padre/madre solo; con o sin hijos convivientes). Ejemplos: “Pareja casada con o sin hijos, con o sin otras personas”, “Padre con hijos, con o sin otras personas”, etc.
Número de hijos	String	Clasificación del número de hijos convivientes (0, 1, 2, 3 o más) y si son menores de 25 años. Ejemplos: “3 o más hijos conviviendo menores de 25 años”, “0 hijos conviviendo”, etc.
Total	Float	Número total de familias con el número de hijos correspondiente para el tipo de núcleo familiar especificado.

### Renta hogar

Nombre del campo	Tipo de dato	Descripción
Periodo	Int	Clasificación por año en el que se registran las rentas netas medias por hogar en cada comunidad autónoma. (2020-2022).
Comunidades y Ciudades Autónomas	String	Nombre de la comunidad autónoma en la que se registran los datos de renta neta media por hogar.
Renta anual neta media por hogar	String	Clasifica por un lado la renta neta media por hogar y por otro incluye la renta neta media por hogar con alquiler imputado. Por ejemplo: “Renta media por hogar (con alquiler imputado)”
Total	Float	Renta neta media total según el año, la comunidad autónoma, y si se el alquiler está imputado o no.

## II. Bibliografía

### Fuentes datasets:

Consumo eléctrico anual: <https://www.ine.es/jaxi/Tabla.htm?tpx=31411>

Viviendas según número personas: <https://www.ine.es/jaxi/Tabla.htm?path=/t20/p276/2020-2035/l0/&file=01001.px&L=0>

Dispositivos energía renovable: <https://www.ine.es/jaxi/Tabla.htm?tpx=56927&L=0>

Intensidad uso viviendas: <https://www.ine.es/jaxi/Tabla.htm?tpx=59531>

Distribución edad población: <https://www.ine.es/jaxiT3/Tabla.htm?t=36781&L=0>

Salario medio población: <https://www.ine.es/jaxiT3/Tabla.htm?t=13930>

Producción de energía por lugar: [https://re.jrc.ec.europa.eu/pvg\\_tools/es/](https://re.jrc.ec.europa.eu/pvg_tools/es/)

Tipo núcleo familiar: <https://www.ine.es/dynt3/inebase/es/index.htm?padre=9544&capsel=9548>

Renta Hogar: <https://www.ine.es/jaxiT3/Tabla.htm?t=9949>

### Referencias:

Bueno, J., Romero-Jordán, D., & Del Río, P. (2020). Analysing the Drivers of Electricity Demand in Spain after the Economic Crisis. *Energies*, 13(20), 5336.  
<https://doi.org/10.3390/en13205336>

Ortega-Izquierdo, M., Paredes-Salvador, A., & Montoya-Rasero, C. (2019). Analysis of the decision making factors for heating and cooling systems in Spanish households. *Renewable and Sustainable Energy Reviews*, 100, 175-185.  
<https://doi.org/10.1016/j.rser.2018.10.013>