

**ANÁLISIS DE DATOS
DE REDES
INALAMBRICAS COMO
PROXY DE DATOS
SOCIECONÓMICOS**

Juan Lobato Camacho

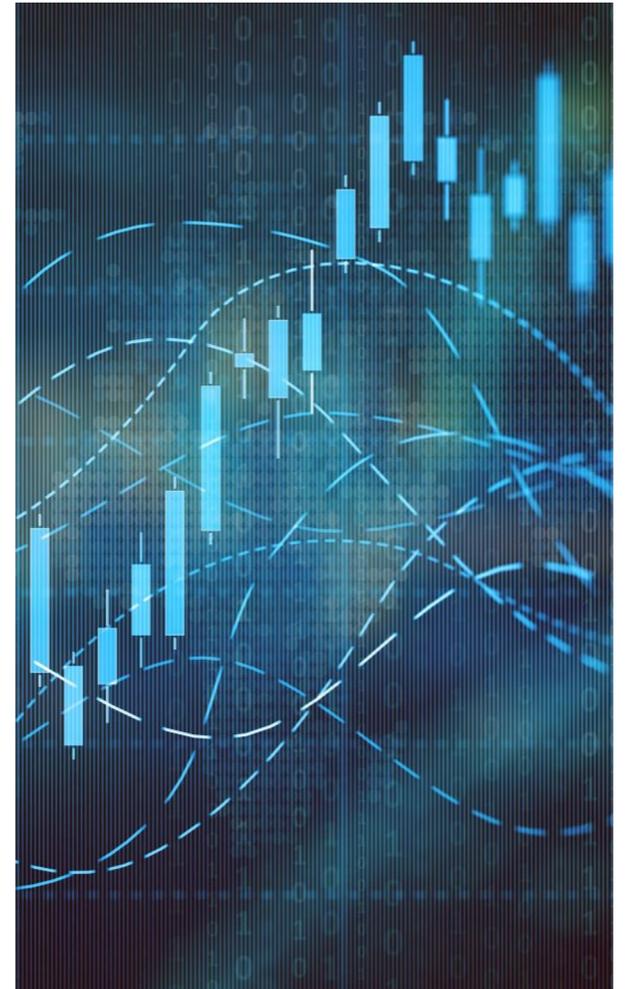


TABLA DE CONTENIDOS



1. INTRODUCCIÓN



2. ESTADO DEL ARTE



3. METODOLOGÍA



4.1 RESULTADOS ESPAÑA



4.2 RESULTADOS EUROPA



5. CONCLUSIONES

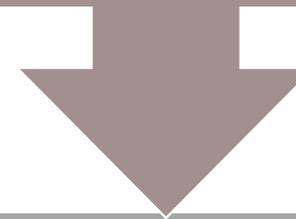
I. INTRODUCCIÓN

I.1 INTRODUCCIÓN

I.2 CONTEXTO HISTÓRICO

I.I INTRODUCCIÓN

El proyecto consiste en poder determinar si únicamente el **número de redes wifi** que hay en una zona determinada, sirve como **indicador** de otros **datos socioeconómicos**.



Datos Socioeconómicos

Número de habitantes	Superficie en Km2	Renta per Cápita Bruta	Habitantes mayores de 65 años	Número de viviendas familiares
----------------------	-------------------	------------------------	-------------------------------	--------------------------------

1.2 CONTEXTO HISTÓRICO Y SOCIAL



- A finales de los **años 90** aparecen las primeras redes wifi, que son redes que permiten la **transmisión de datos** de forma **inalámbrica** a **grandes velocidades** .
- Aparecen principalmente en **grandes empresas** y **universidades**.
- En los años 2000 comienza a extenderse el uso a **negocios** como cafeterías, restaurantes, hoteles y **zonas públicas** como parques, ayuntamientos etc.
- Desde la década de los 2010 se hace un **indispensable** llegando a muchos **hogares**, escuelas y oficinas gracias a la evolución de la misma debido a protocolos como 802.11ac o 802.11ax.

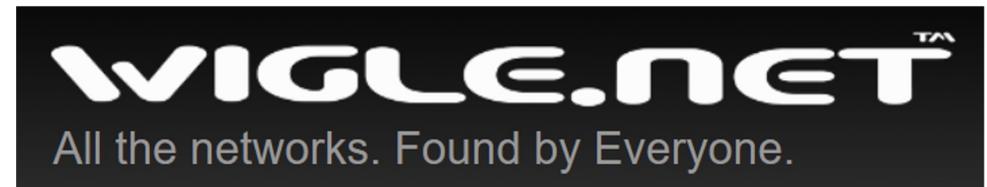
2. ESTADO DEL ARTE

2.1 BASES DE DATOS DE REDES WIFI

2.2 POLÍGONOS DE SUPERFICIE

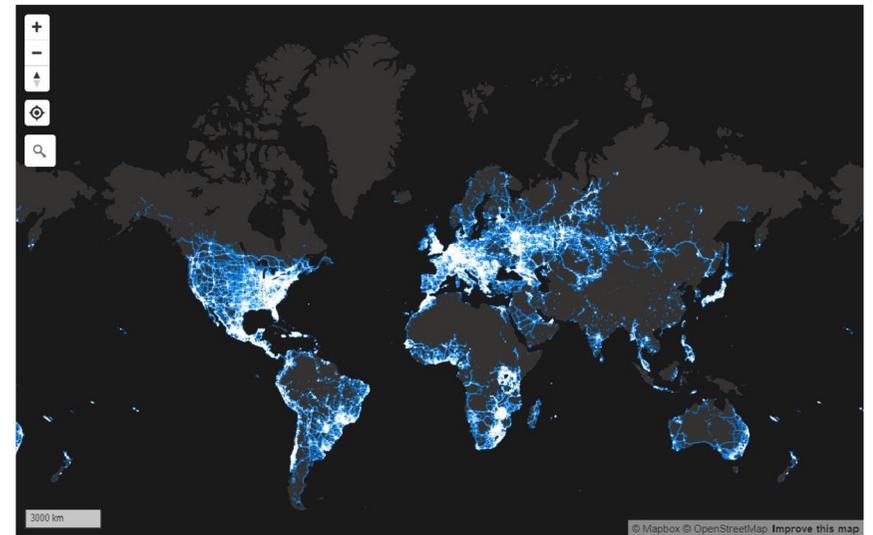
2.1 BASES DE DATOS DE REDES WIFI

Encontramos 2 posibles bases de datos:



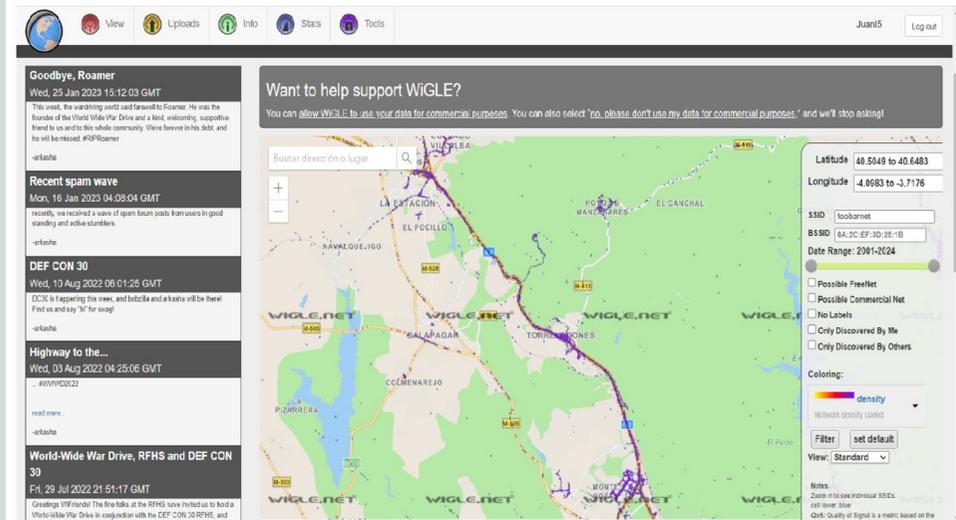
SERVICIO DE LOCALIZACIÓN MOZILLA

- Recopila datos de redes de telefonía móvil, bluetooth y **wifis**, tanto **desarrolladores** como **usuarios** recopilan la información.
- Únicamente permite acceso a personas que estén desarrollando aplicaciones de código abierto que ellos **autoricen**.



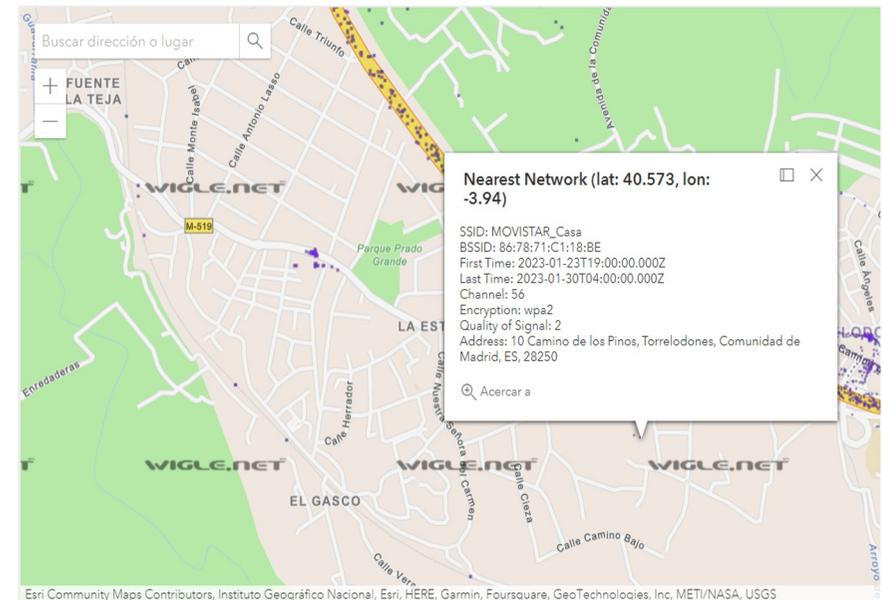
WIGLE

- Contiene información de redes de telefonía, bluetooth y redes wifi recopilados únicamente por usuarios.
- Acceso gratuito para todo el mundo.
- Permite crear cuenta que da más posibilidades.
- Cuenta con una **API** muy útil y completa



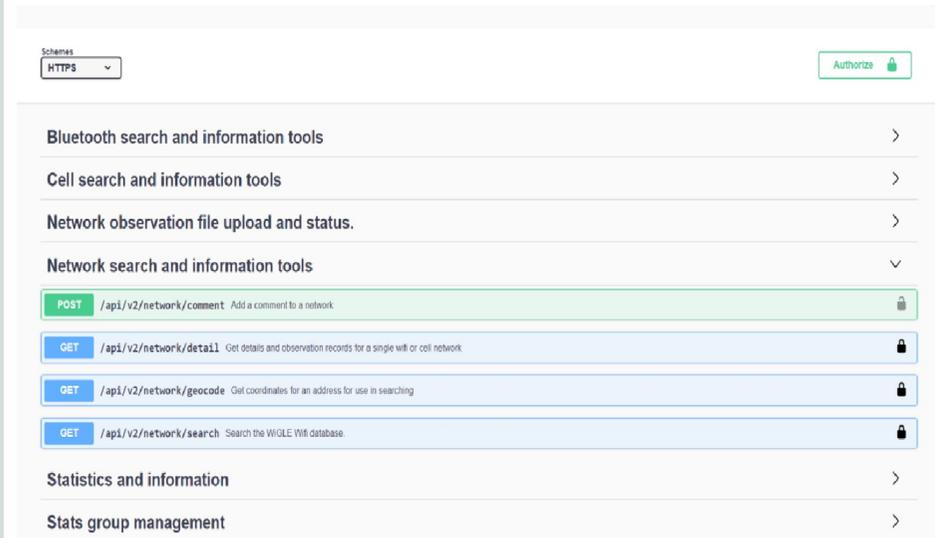
DATOS WIGLE

- **SSID:** (Service Set Identifier) Nombre público que identifica una red.
- **BSSID:** Identificador único de cada dispositivo inalámbrico
- **First Time:** La primera vez que se vio la red Wifi
- **Last Time:** la última vez que se vio la red Wifi
- **Channel:** división específica de frecuencias dentro de una banda inalámbrica en cuestión
- **Encryption:** Nivel de encriptación de la red
- **Quality of Signal:** Calidad de la señal
- **Address:** Dirección exacta del Wifi



API WIGLE

- La API de Wigle permite realizar peticiones **GET** y **POST**.
- Permite **100** peticiones al día.
- Necesario estar **registrado**.
- Métodos útiles:
 - **Network.geocode**: a partir del nombre de un municipio te devuelve las coordenadas.
 - **Network.search**: a través de unos criterios de búsqueda te devuelve los **datos** que le indiques de las redes wifi.

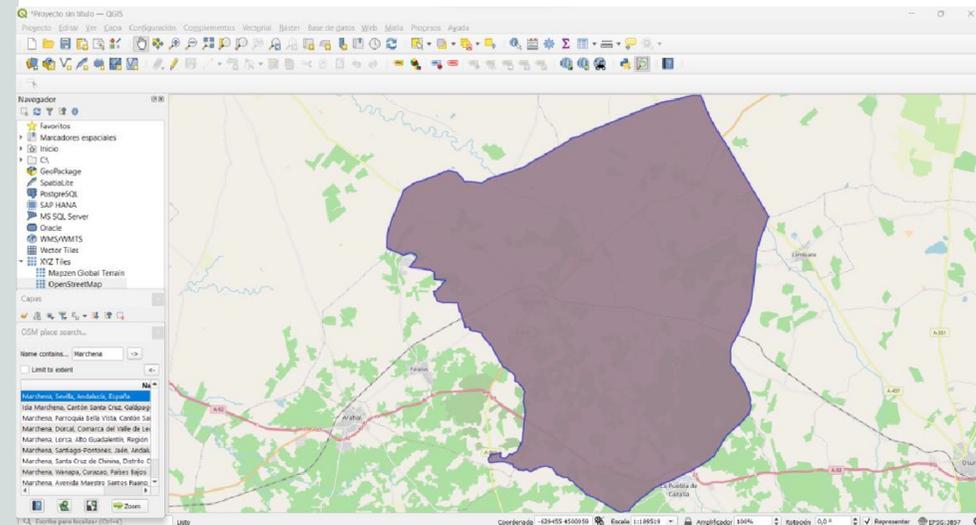


2.2 POLÍGONOS DE SUPERFICIES



QGIS

- QGIS es una aplicación de escritorio para ordenador que permite obtener **información geográfica**.
- Las **funcionalidades principales** que he utilizado en QGIS, han sido:
 - **OSM Place Search:** permite buscar por nombre un determinado territorio.
 - **Mapa de OpenStreetMaps:** mapa que representa el mundo entero.
 - **Generación de polígono de coordenadas**



REPOSITORIO GITHUB

- Repositorio de GitHub que generaba los polígonos de coordenadas de cualquier zona geográfica que se quisiera.
- Además, contaba con un **documento .xml** el cual contenía los polígonos de todos los municipios españoles.
- <https://github.com/hegdehog/PolylineGenerator>



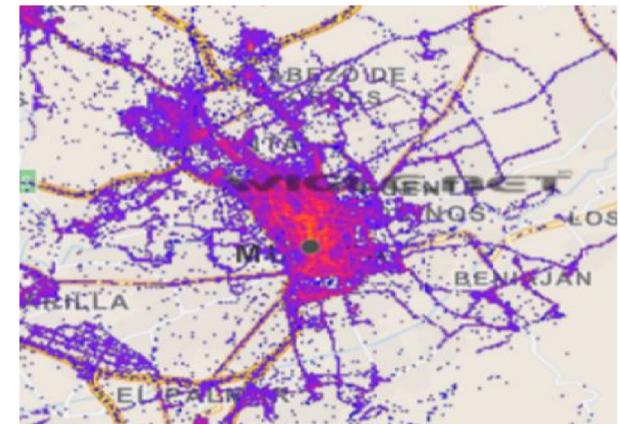
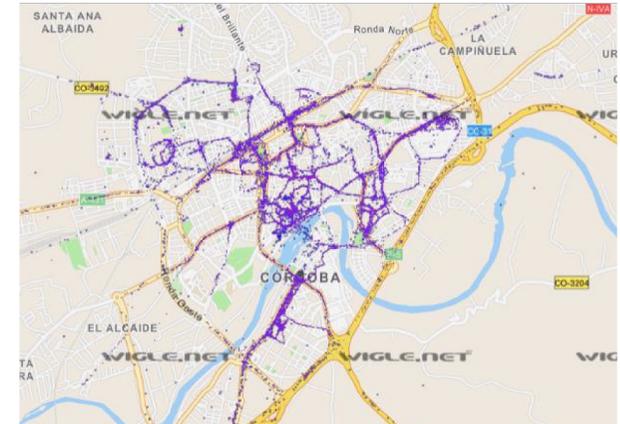
3.METODOLOGÍA

3.1
LIMITACIONES

3.2
METODOLOGÍA

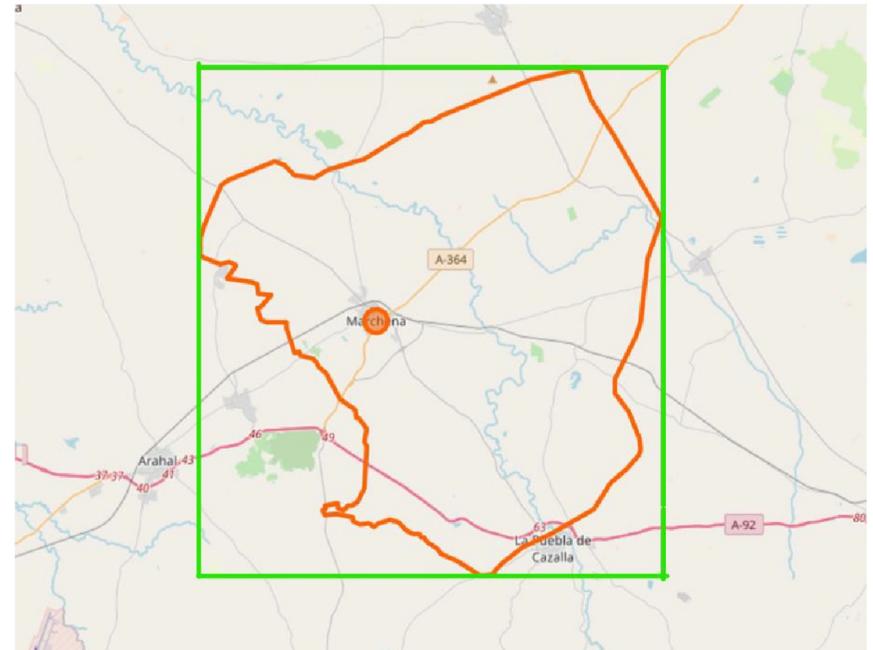
3.1 LIMITACIONES

- **Disponibilidad de datos en Wigle:** los datos que recoge Wigle han sido **subidos** por parte de los **usuarios**, lo que hace que haya zonas que no cuentan con datos completos.
- **Solución:** realizar capturas de pantalla de todas las zonas de estudio, y determinar cuáles de ellas tienen un nivel esperado de redes wifi en la base de datos de Wigle



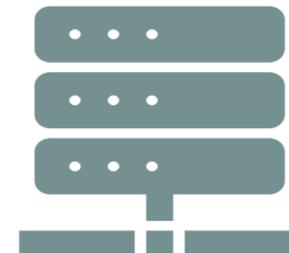
3.1 LIMITACIONES

- **Cuadrículas de Coordenadas (Bbox):** la única forma de buscar datos que se corresponden con una zona determinada en Wingle es a través de una cuadrícula de coordenadas o **Bbox**, que consiste en el conjunto de las **dos latitudes y longitudes máximas y mínimas** de un municipio.
- **Solución:** no hay solución posible, pero se la mayor parte de las redes se encuentran en el núcleo urbano, por lo que el número de wifis que cogemos de otros municipios será muy reducido



3.1 LIMITACIONES

- **Límite de peticiones de Wigle:** la API de wigle permite hacer 100 peticiones al día,
- **Solución:** me aumentaron el número de peticiones a 250.
- **Redes duplicadas:** algunos de los datos recogidos, son versiones de la **misma red**, en versiones de **2.4 GHz** y **5 GHz** que se corresponden con el mismo usuario.
- **Solución:** no hay solución posible, pero nuevamente no afecta demasiado al estudio puesto que el **porcentaje de redes repetidas** es muy pequeño.



3.2 METODOLOGÍA



1. Decidir **muestra** a estudiar.



2. Programa en Python que busca en el **documento .xml** para obtener los puntos de coordenadas, y **genera un Bbox** para cada ciudad.



3. Método en Python que interactúa con la **API** de Wigle mediante la librería Pygle, que recibe un **Bbox** y devuelve el número de redes **wifi** que hay en esas coordenadas.



4. Descartar **datos incompletos** mediante la realización de capturas en Wigle.



5. Obtener el resto de los datos **socioeconómicos**.



6. Realizar **gráficas** en Python para poder observar los resultados.



7. Obtener **conclusiones**.

4. RESULTADOS

3.1 RESULTADOS ESPAÑA

3.2 RESULTADOS EUROPA

4.1 RESULTADOS ESPAÑA



La muestra seleccionada para este estudio son los 64 municipios españoles que cuentan con más de 100.000 habitantes en el año 2023.



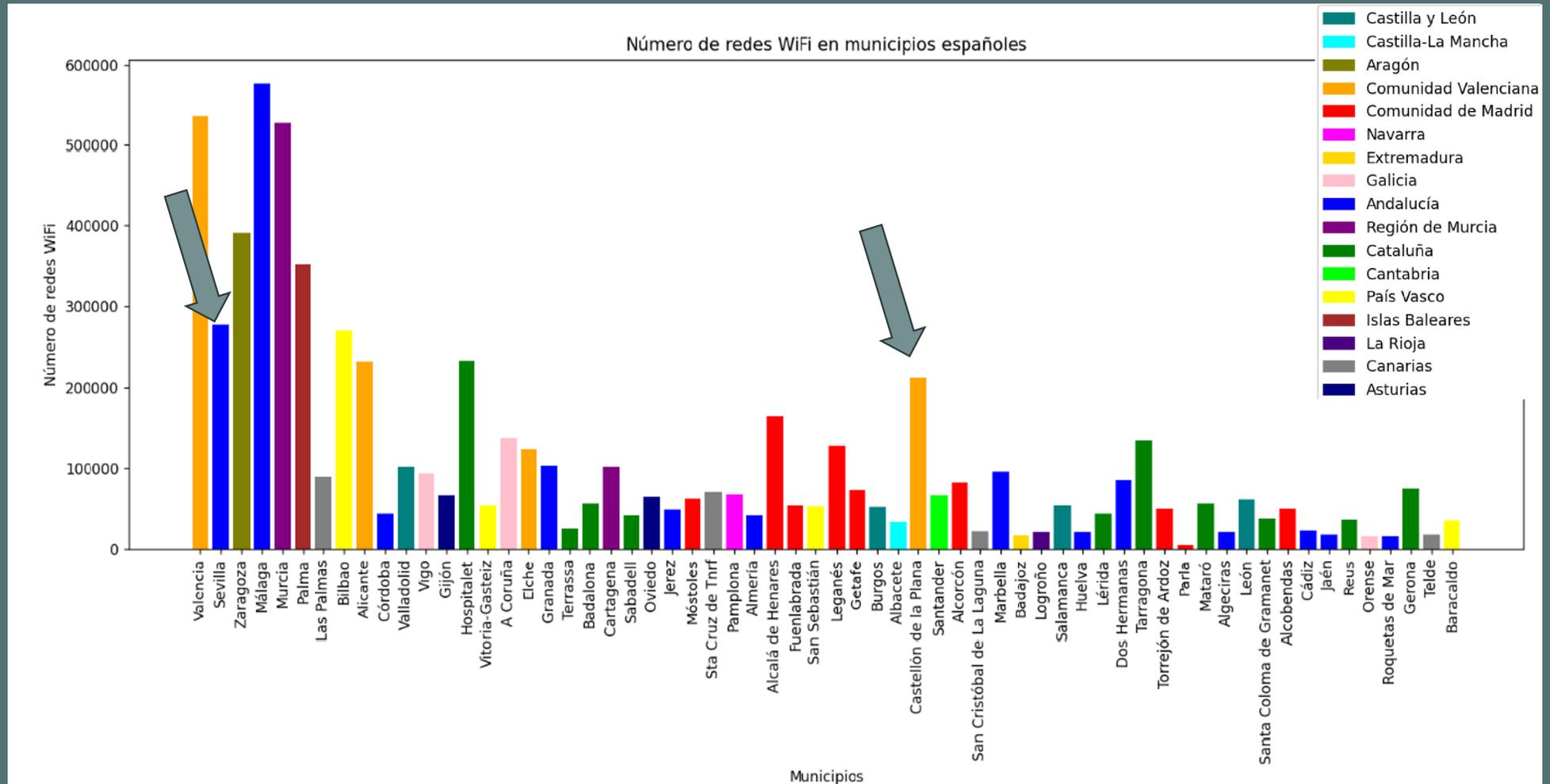
En primera posición encontramos Madrid con 3.280.782 habitantes y en último Baracaldo con 100.535.



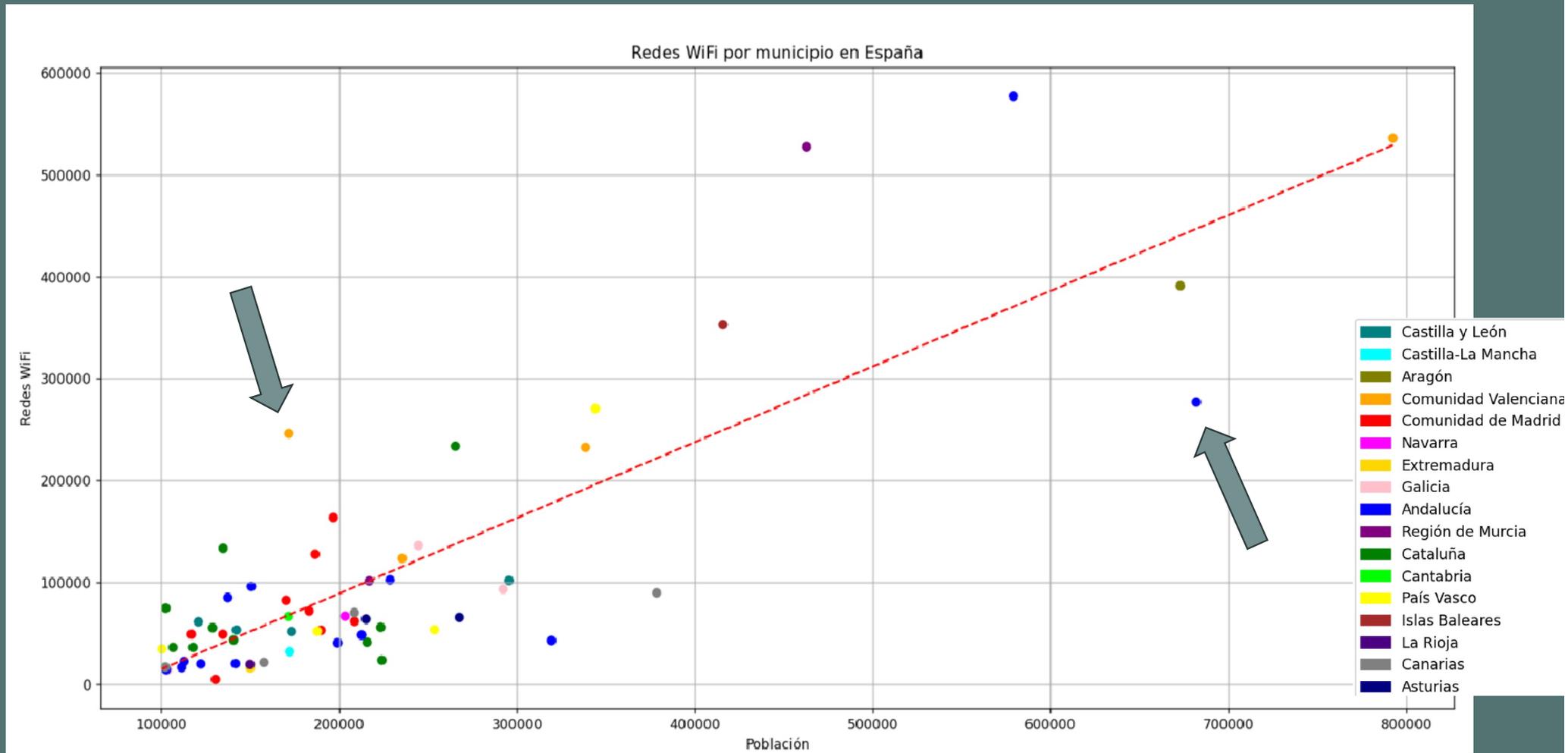
Los datos socioeconómicos a estudiar con respecto al número de wifi son:

Número de habitantes
Superficie en Km2.
Renta per Cápita Bruta.
Personas mayores de 65 años.
Número de viviendas familiares.

4.1.1 NÚMERO DE REDES WIFI

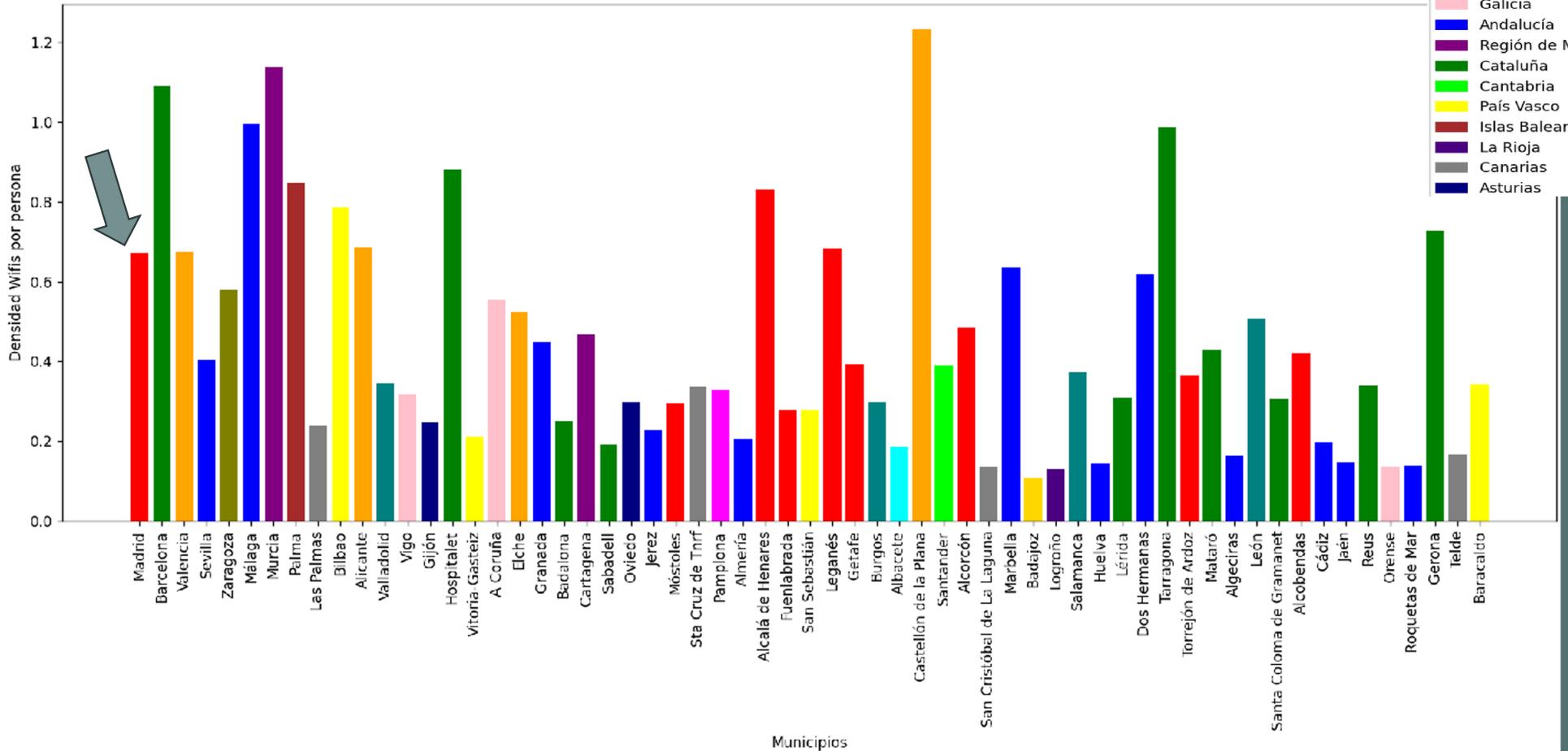


4.1.2 NÚMERO DE REDES WIFI VS POBLACIÓN



4.1.3 DENSIDAD DE REDES WIFI POR HABITANTE

Densidad de redes WiFi por habitante en municipios españoles

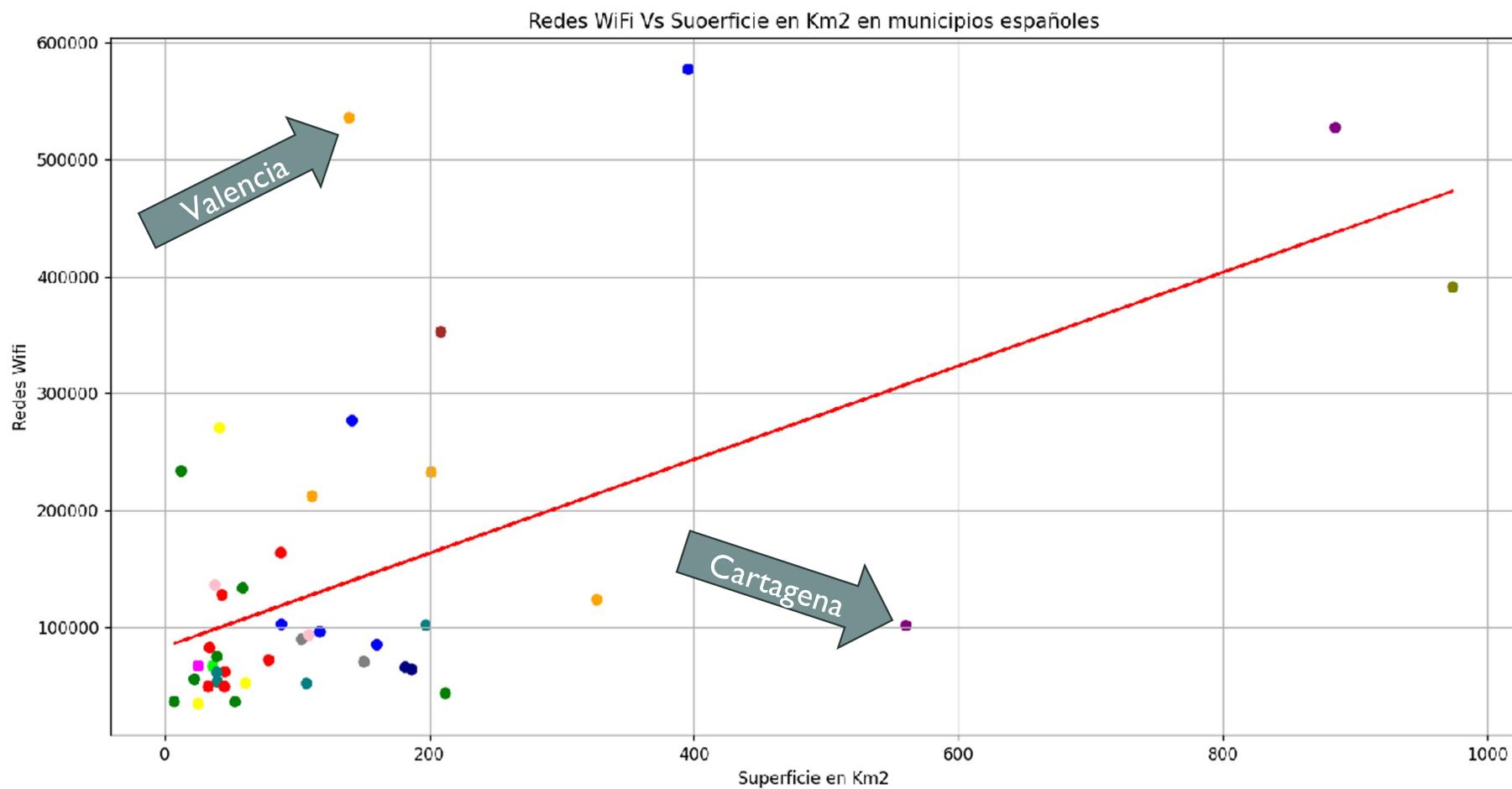


DESCARTE DE MUNICIPIOS

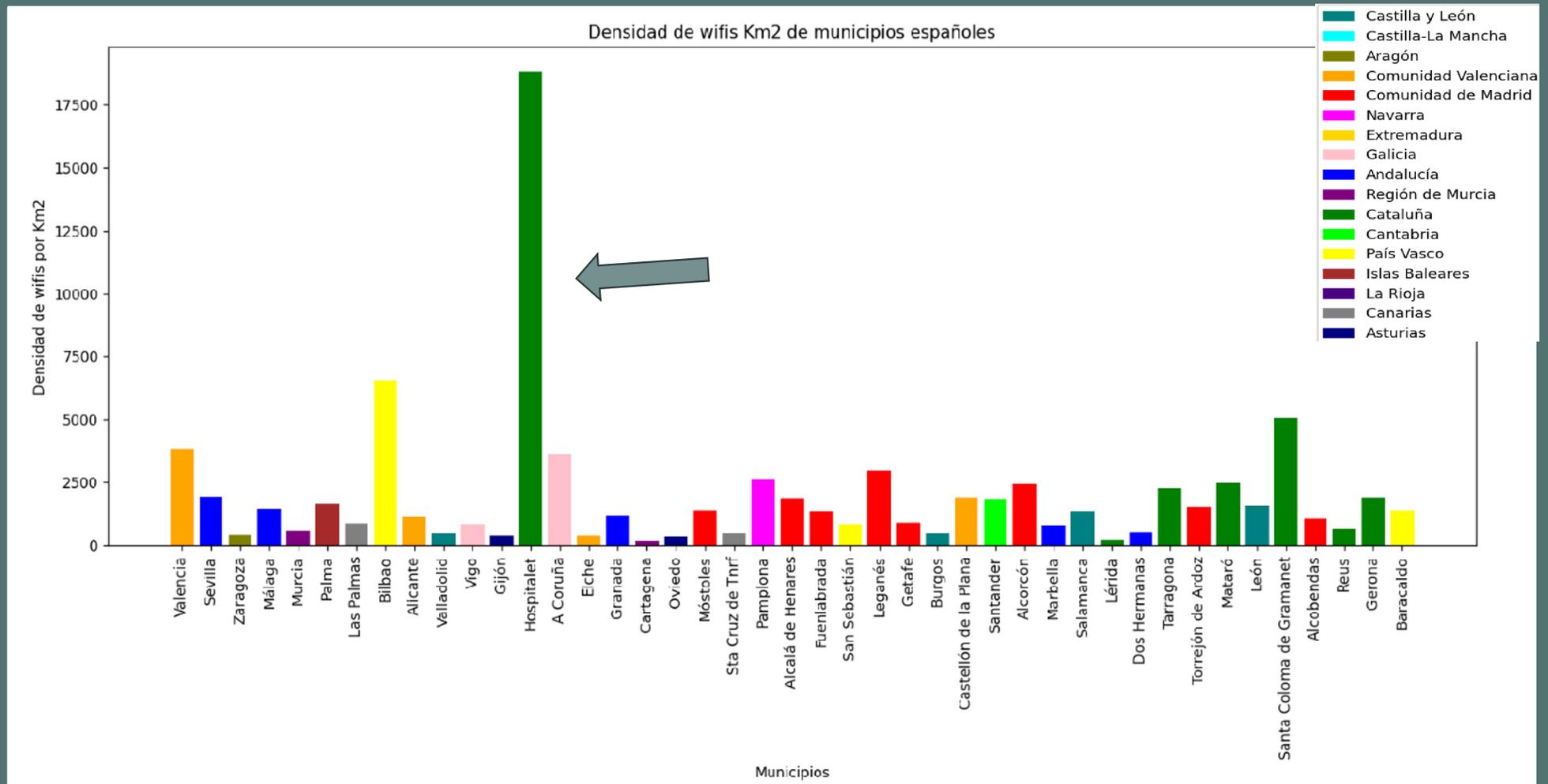
- La media de **habitantes por hogar** en España es de **2,5 personas** (INE), debería de haber entre **0,5 y 0,333 wifis por habitante**, si **únicamente** se tuviesen en cuenta los wifis que hay en los **hogares** españoles, la media debería superar estas estimaciones.
- **Descartamos** los municipios cuya densidad de wifis sea **menor de 0,3 wifis por habitante**, deberíamos dejar fuera del estudio los siguientes municipios: Vitoria-Gasteiz, San Cristóbal de La Laguna, Badajoz...



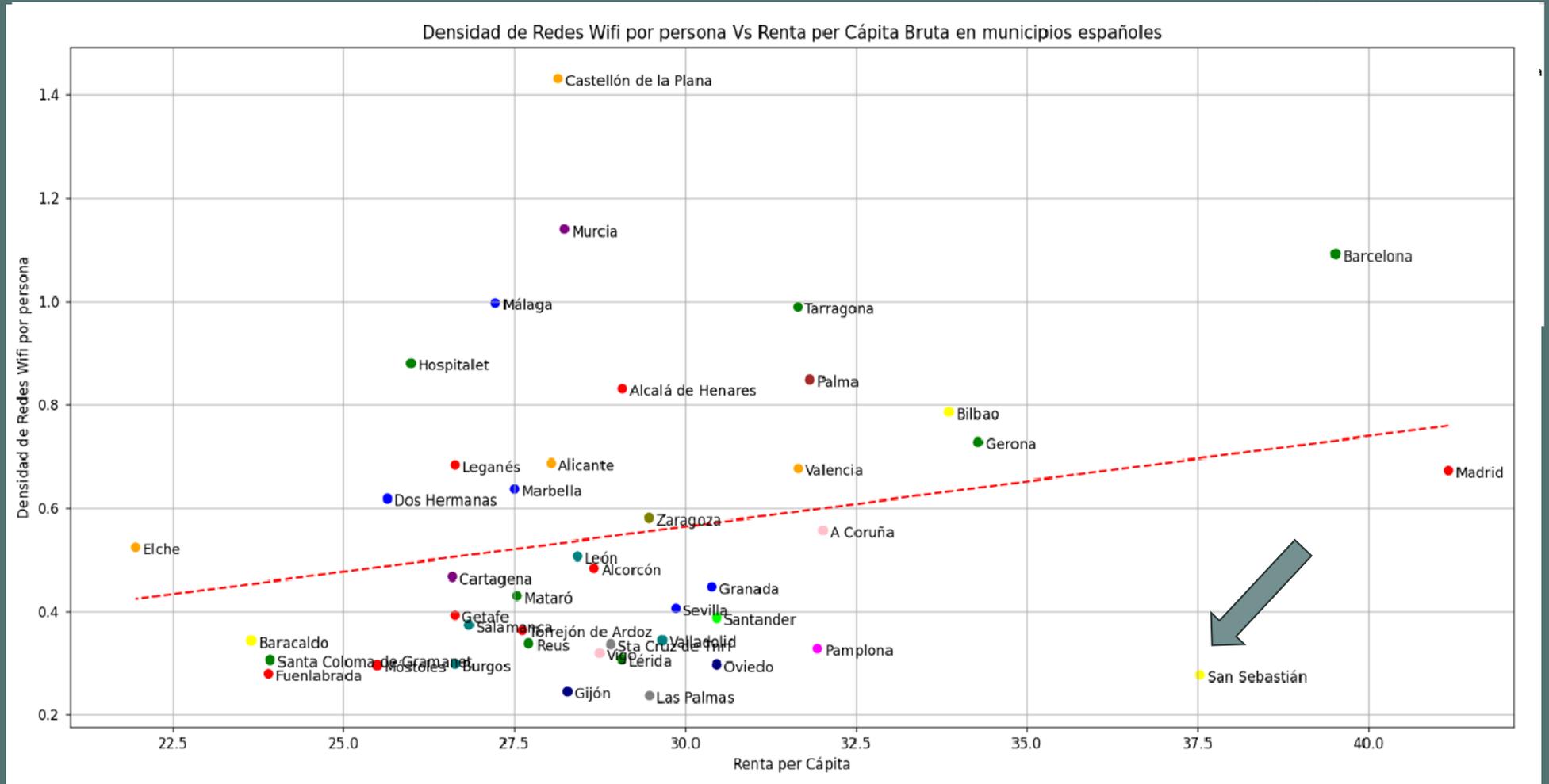
4.1.5 SUPERFICIE EN KM₂ VS WIFIS



4.I.6 DENSIDAD DE WIFIS POR KM₂



4.1.8 RENTA PER CÁPITA BRUTA VS DENSIDAD DE WIFIS



4.2 RESULTADOS EUROPA



La muestra seleccionada para este estudio son los 17 municipios europeos que cuentan con más de 1.000.000 habitantes en el año 2023.



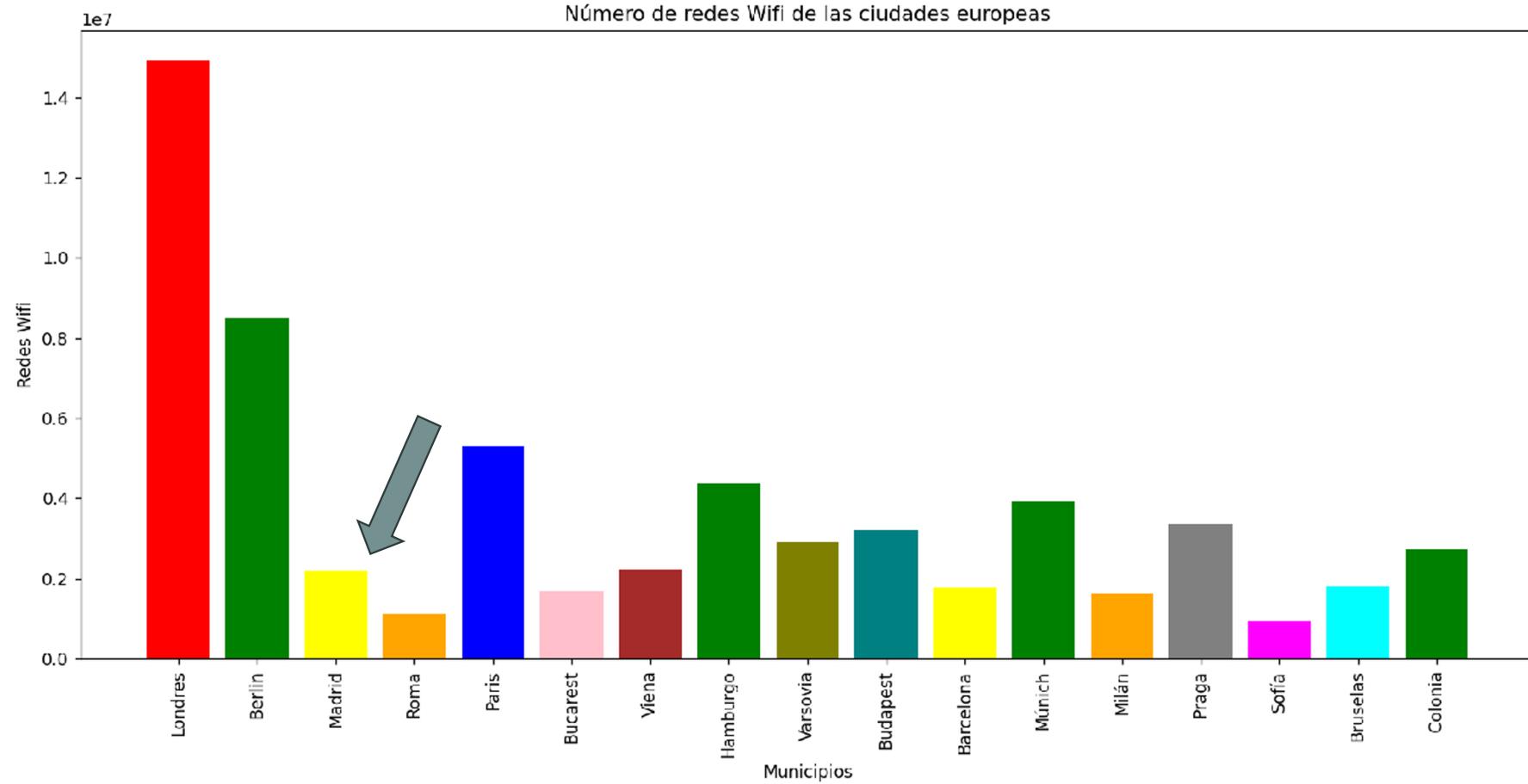
En primera posición encontramos Londres con 8.982.000 habitantes y en último Colonia con 1.089.984.



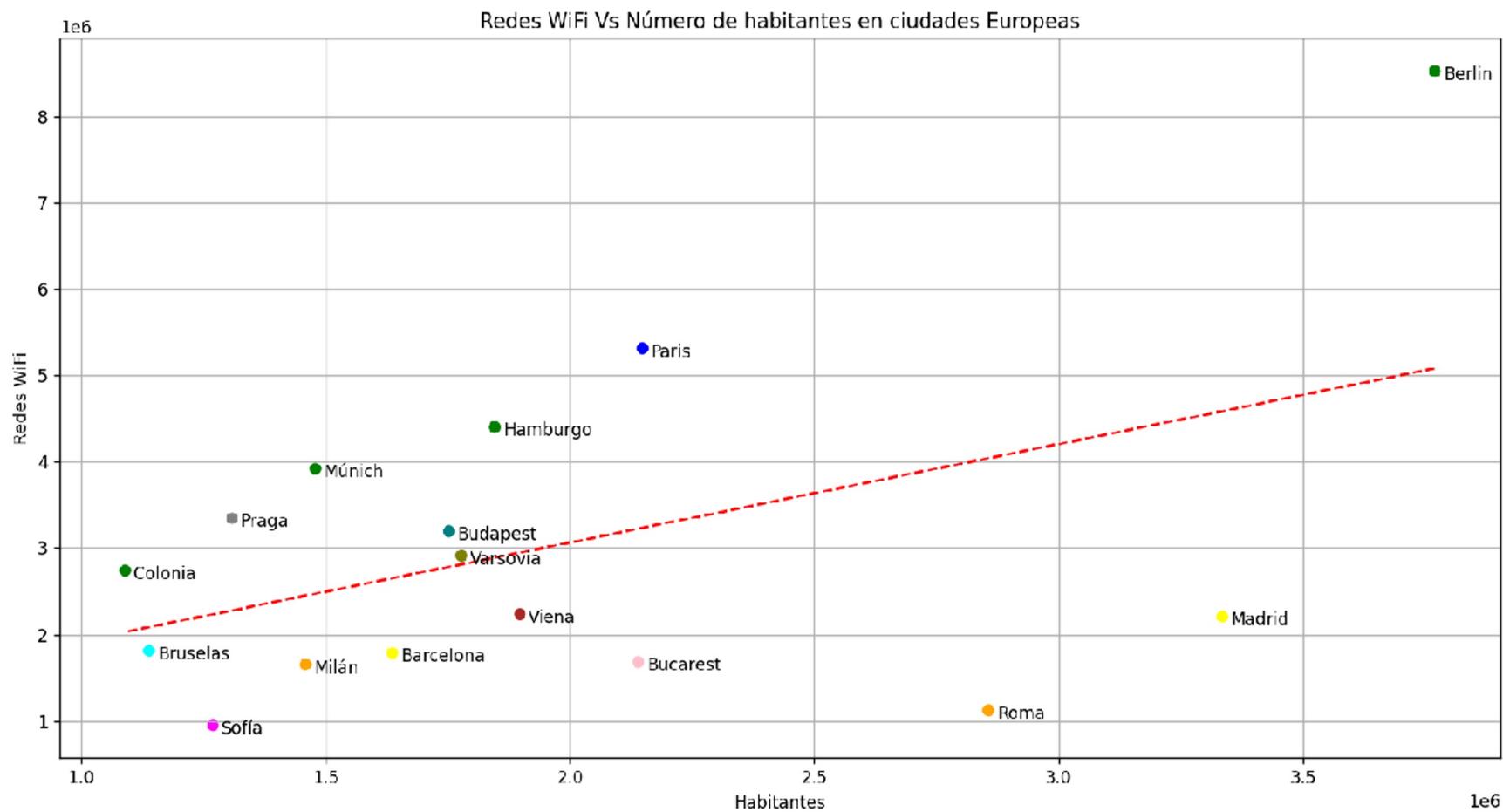
Los datos socioeconómicos a estudiar con respecto al número de wifi son:

Número de habitantes
Superficie en Km2.
PIB per cápita.

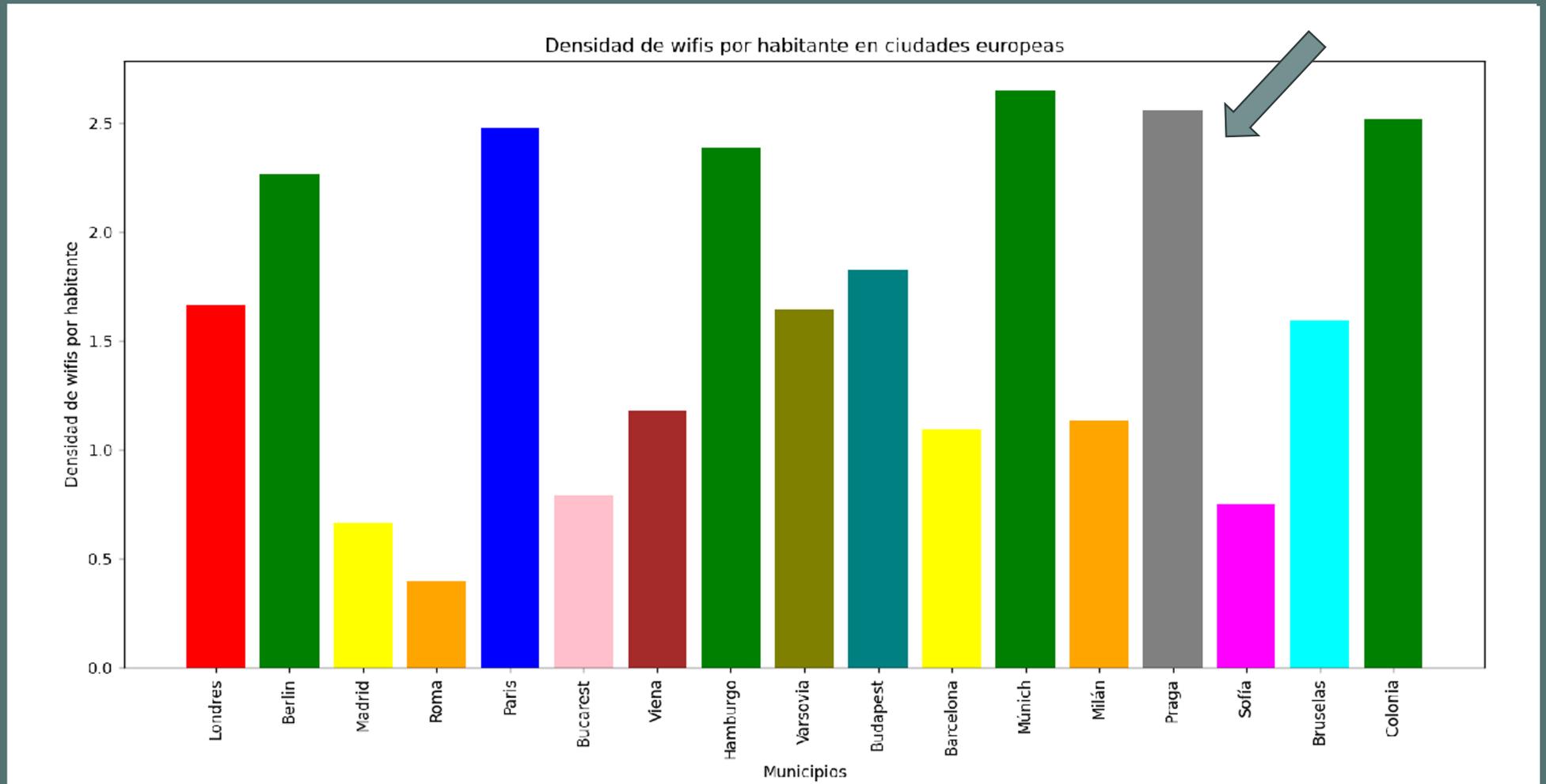
4.2.1 REDES WIFI



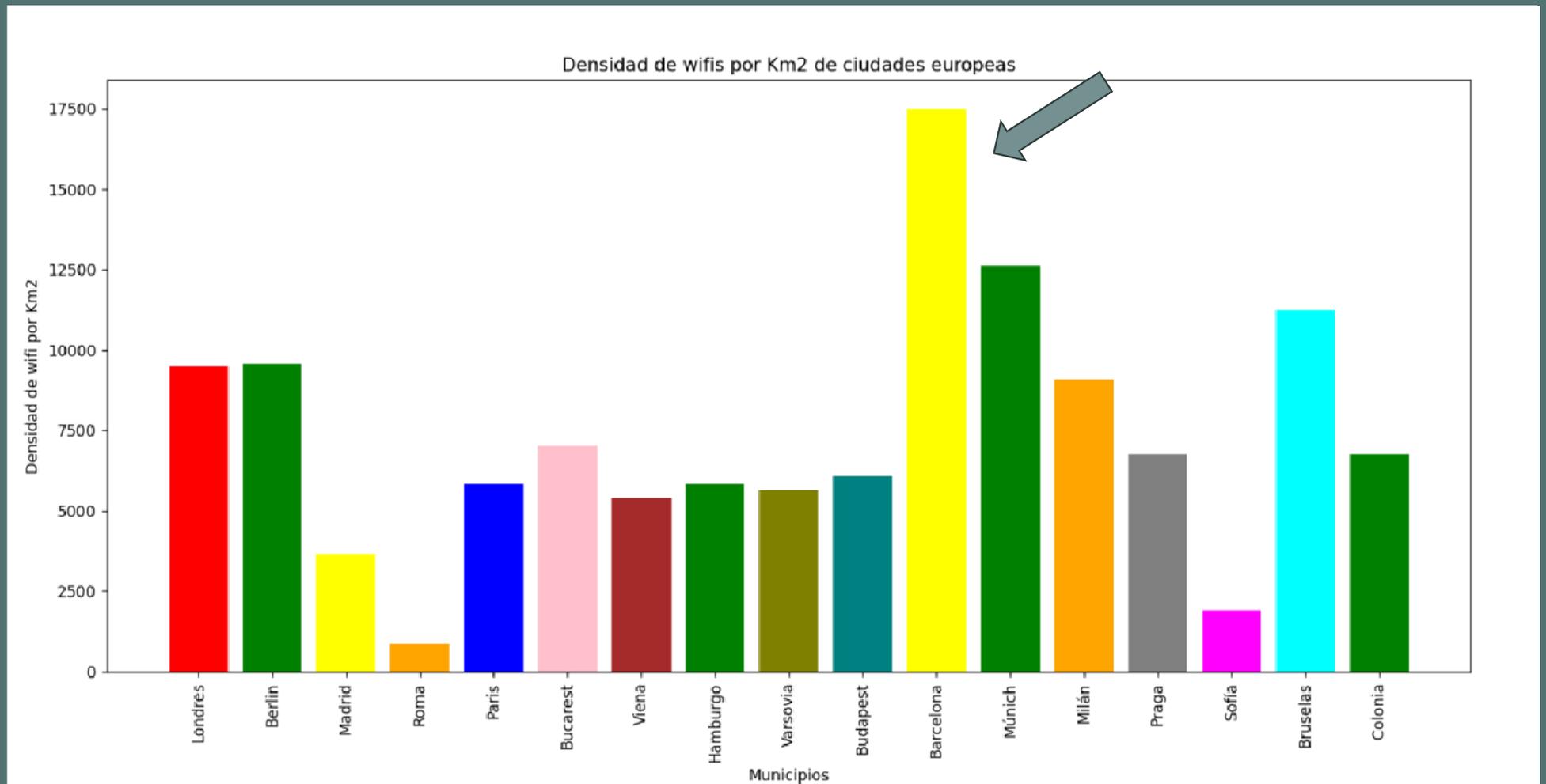
4.2.2 NÚMERO DE HABITANTES VS REDES WIFI



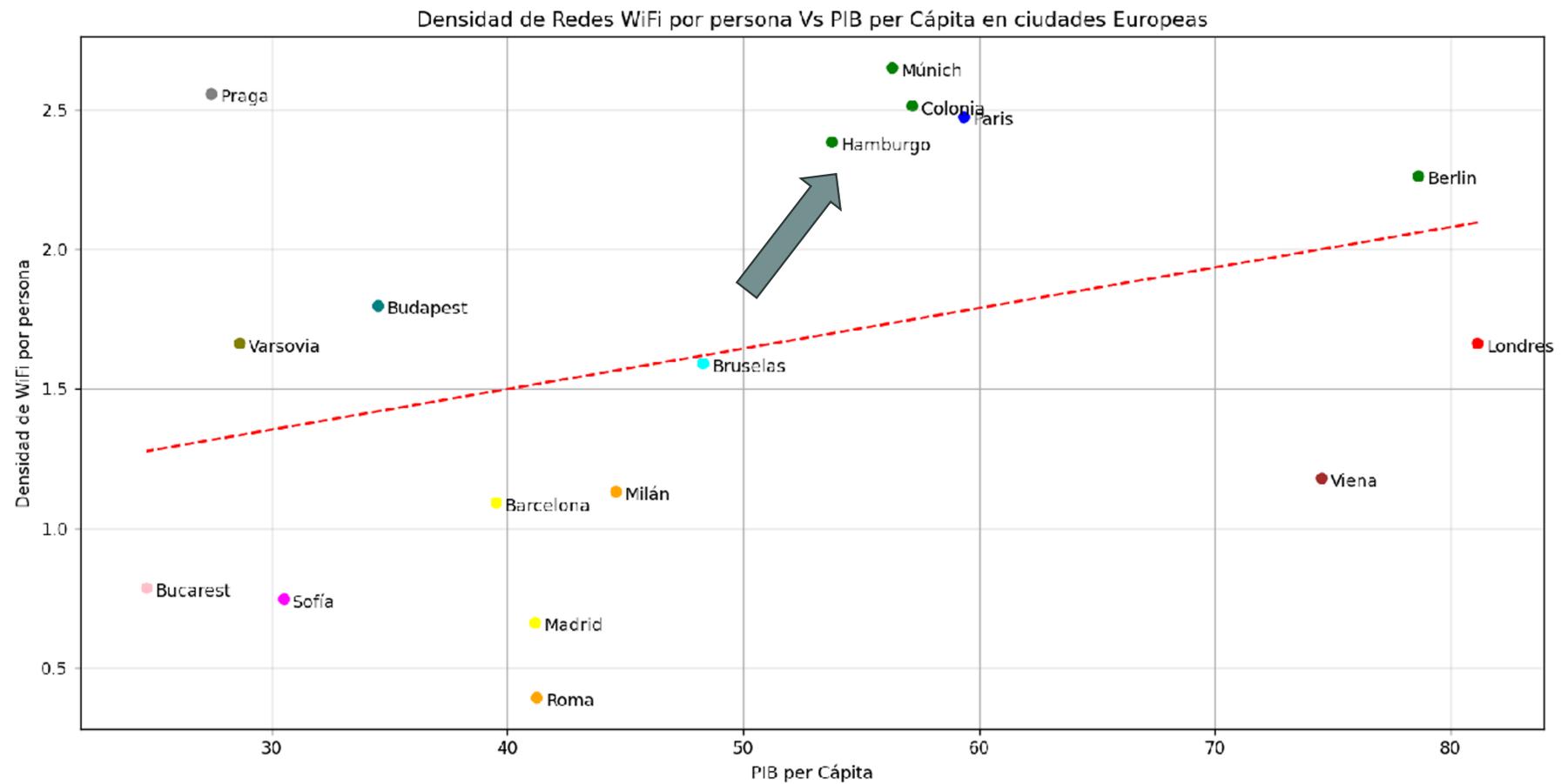
4.2.3 DENSIDAD DE WIFI POR HABITANTE



4.2.4 DENSIDAD DE WIFI POR KM₂



4.2.6 DENSIDAD DE WIFIS POR PERSONA VS PIB PER CÁPITA



5. CONCLUSIÓN



5. CONCLUSIÓN

- La importancia del trabajo consiste en la **dificultad** para **obtener** los datos y la **metodología** utilizada, permitiendo resolver un auténtico problema de ingeniería y **no** tanto en los **resultados** obtenidos.
- Los datos utilizados **no son** del todo **completos**, por lo que debemos de ser **cautos** a la hora de sacar conclusiones, podemos intuir que efectivamente el **número de redes wifi** puede ser un **buen indicador** del número de **habitantes** de un municipio y del **nivel de renta** de dicho municipio, mientras que **no de la superficie** del municipio.
- Este estudio puede servir como **precedente** y **punto de partida** para **otros trabajos** o **investigaciones**.

