



ESTUDIO CONTAMINACIÓN MEDIO AMBIENTE

Visualización de datos, 2º Ciencia de datos



2 DE MAYO DE 2024

JAVIER BENY RODRÍGUEZ, ADRIÁN BLAT CAMPOS, DIEGO CAMACHO MENGUAL

Índice

1. Introducción.....	2
2. Metodología.....	2
2.1. Análisis exploratorio de los datos.....	2
2.2. Preprocesado de Datos.....	2
2.3. Preprocesado de la información geográfica.....	3
2.4. Elección de gráficas para los distintos tipos de datos.....	9
2.5. Diseño del mapa e interactividad.....	10
2.6. Diseño del cuadro de mandos.....	11
2.7. Implementación.....	11
3. Resultados.....	12
4. Discusión.....	14
5. Conclusiones.....	15
6. Referencias.....	16

1. Introducción

El objetivo de este miniproyecto es el estudio de la contaminación de distintos tipos de gases en los diferentes distritos de Valencia. Vamos a realizar un procesado de capas en Qgis y posteriormente en R realizaremos diferentes gráficas y mapas con las librerías shiny y Leaflet para la realización del estudio.

2. Metodología

2.1. Análisis exploratorio de los datos

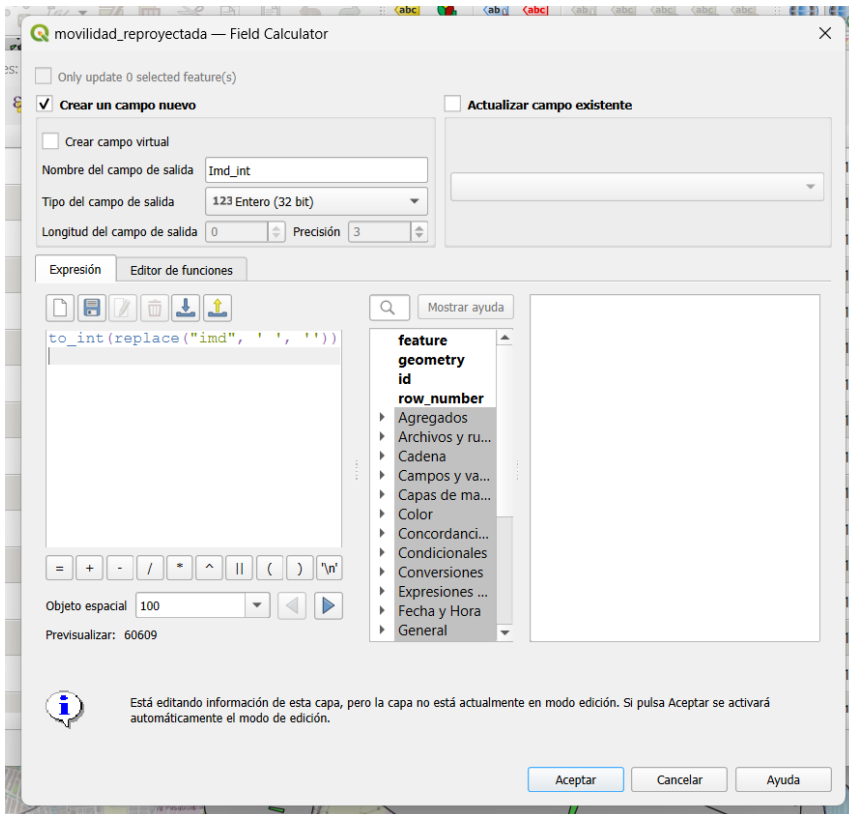
Para este proyecto hemos utilizado los siguientes conjuntos de datos:

- **Datos diarios calidad aire 2004-2022:**
 - Datos diarios de calidad del aire de las estaciones de la red de vigilancia de la ciudad de València. Datos no georreferenciados. Para cada registro, se sabe la estación de medición desde la que se han tomado los datos. Hay un total de 43.388 registros.
- **Estaciones contaminación atmosféricas:**
 - Datos diarios sobre las distintas estaciones de contaminación atmosféricas que hay en Valencia. Datos georreferenciados. La capa vectorial es de tipo Point.
- **Distritos:**
 - Contienen las capas de información administrativa proveniente del servicio de Urbanismo e Infraestructuras de Opendata referentes a distritos. Datos georreferenciados. La capa vectorial es de tipo Polygon.
- **Puntos medida tráfico espiras electromagnéticas:**
 - Ubicación de los puntos de medida de tráfico en Valencia (espiras electromagnéticas). Datos georreferenciados. La capa vectorial es de tipo Point.
- **Datos IMD València desde enero 2016:**
 - Datos de Intensidad Media Diaria mensual de vehículos motorizados de los días laborales de la ciudad de València desde 2016. Datos georreferenciados. La capa vectorial es de tipo Point, y para cada uno de ellos, existen múltiples registros.
- **Zonas Verdes (Planificación):**
 - Localización de lugares planificados para la creación de jardines y espacios verdes en la ciudad de València. Datos georreferenciados. La capa vectorial es de tipo Polygon.
- **Puntos de medida espiras electromagnéticas (bicicletas):**
 - Información geográfica de los puntos de medida de espiras electromagnéticas para bicicletas, con campos como Intensidad/hora de bicicletas y ángulo representación del sentido de la circulación en el punto de medida. Datos georreferenciados. La capa vectorial es de tipo Point.

2.2. Preprocesado de Datos

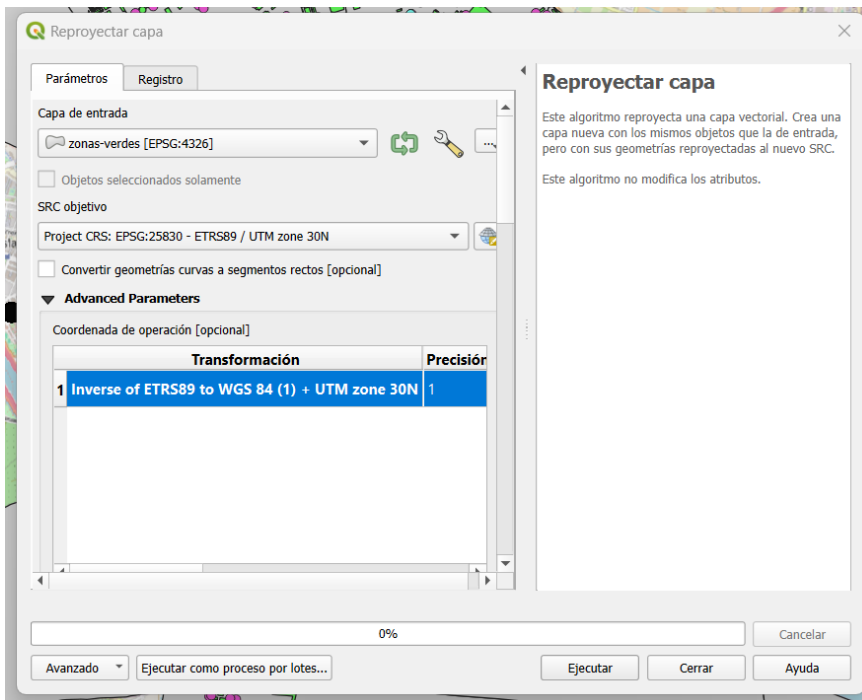
Pasar a int y quitar espacios entre números (en mobilitat, la columna imd):

Para hacer esto hemos hecho uso de la calculadora de campos. Con el siguiente código hemos quitado los espacios a la variable "imd2 y posteriormente lo hemos pasado a tipo entero:



2.3. Preprocesado de la información geográfica

Antes que nada, hemos proyectado todas las capas para tenerlas todas en el mismo sistema de coordenadas de referencia (25830). Aquí un ejemplo con la capa zonas verdes:



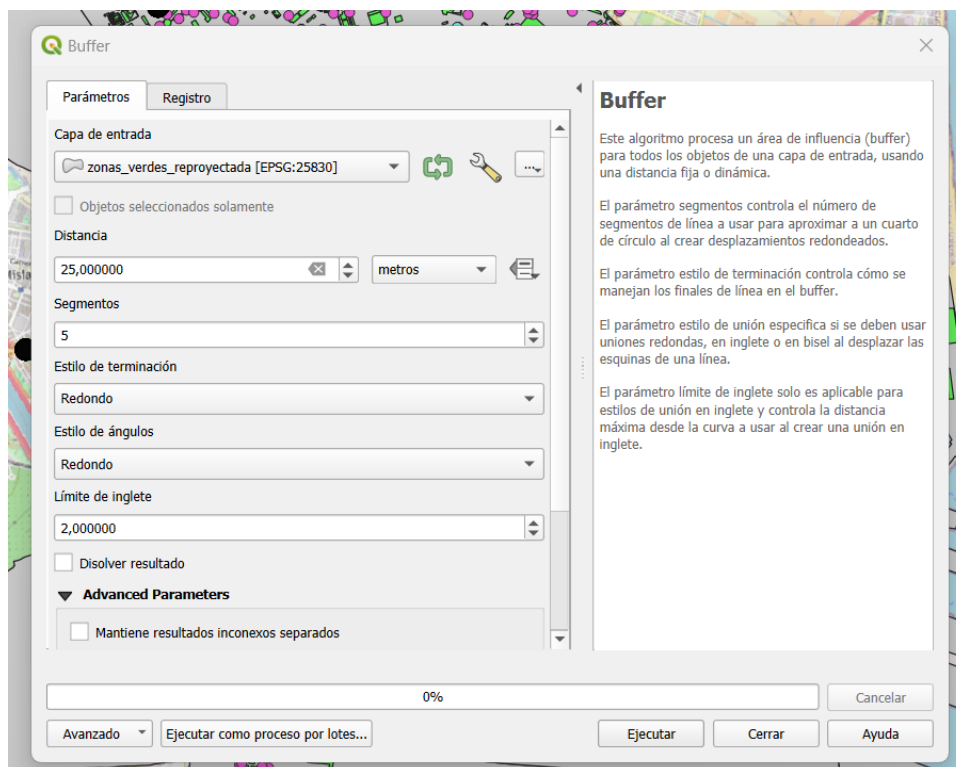
Después de reproyectar las capas hemos utilizado en este proyecto tanto unión, intersección como buffers:

Realizamos 2 Buffer de las capas Zonas verdes y Bicicletas:

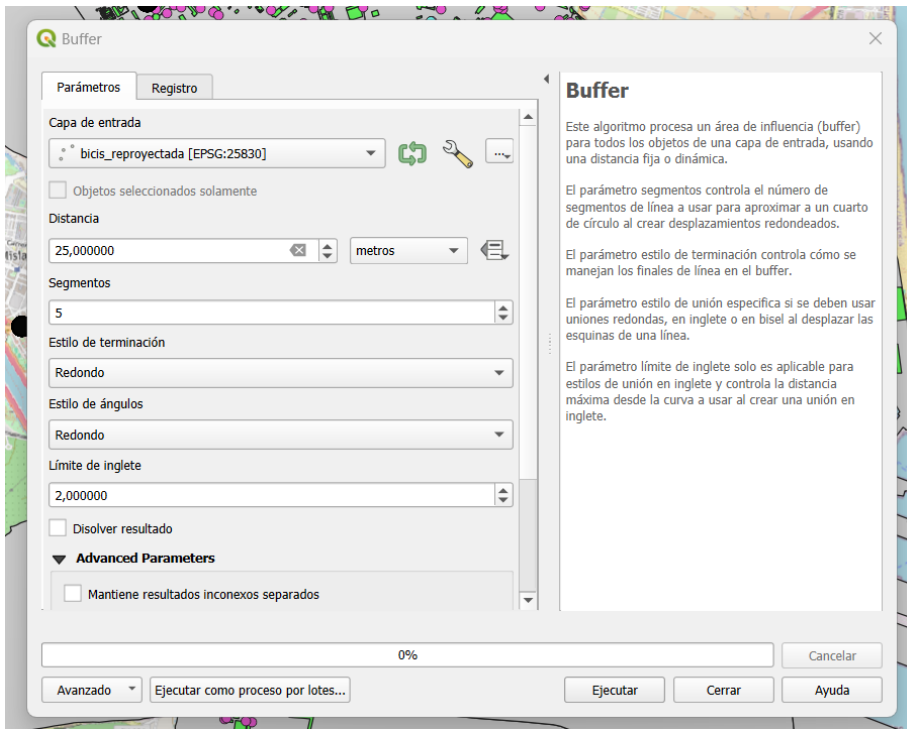
Para hacer un buffer, nos hemos ido al apartado Vectorial > Herramientas de geoprocso > Buffer...



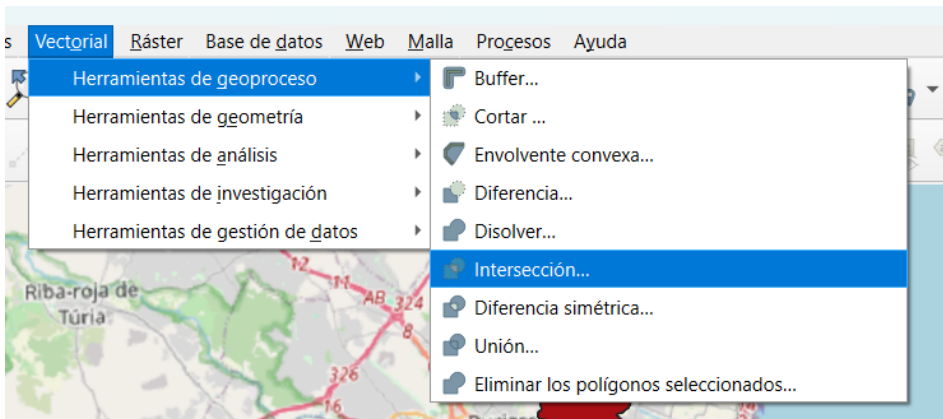
Y posteriormente hacemos un buffer de de la capa de zonas verdes reprojectada con una distancia de 25 metros:



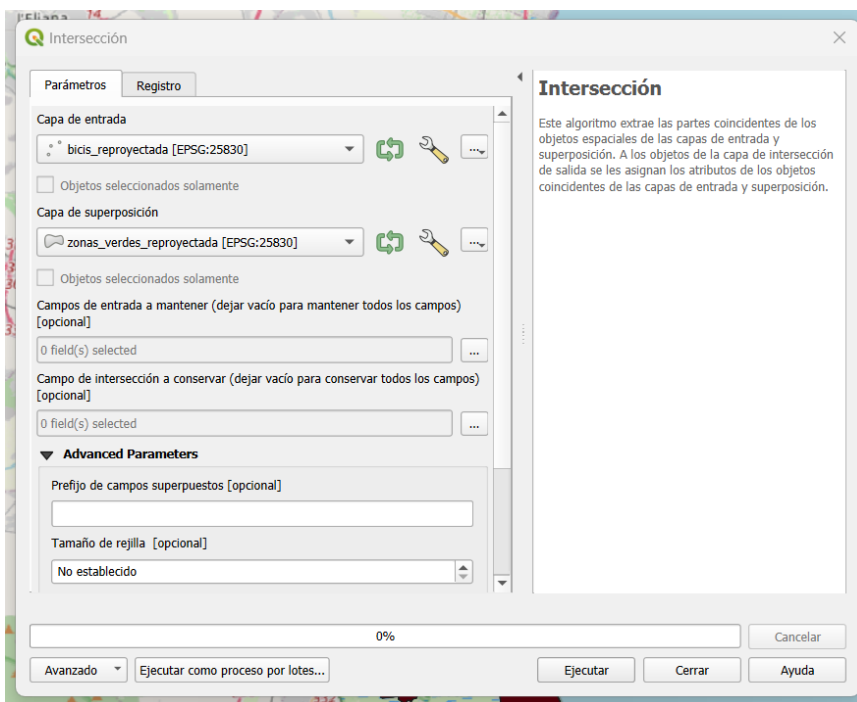
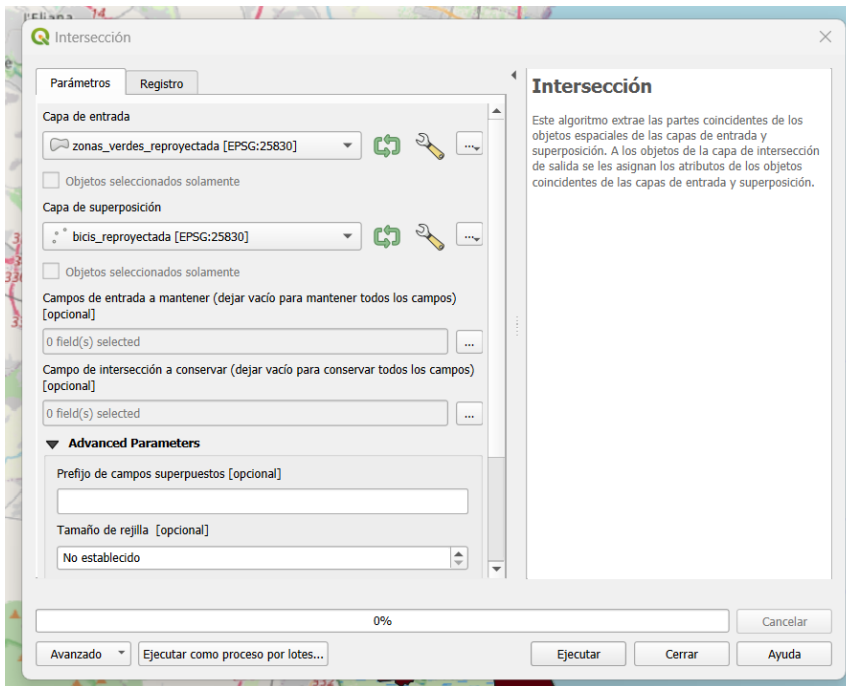
Nuevamente hacemos otro buffer de la capa bicicletas re proyectada:



Una vez tenemos los buffer hechos, realizamos una intersección entre las 2 capas, para ello tenemos que irnos a Vectorial > Herramientas de geoprocso > Intersección...

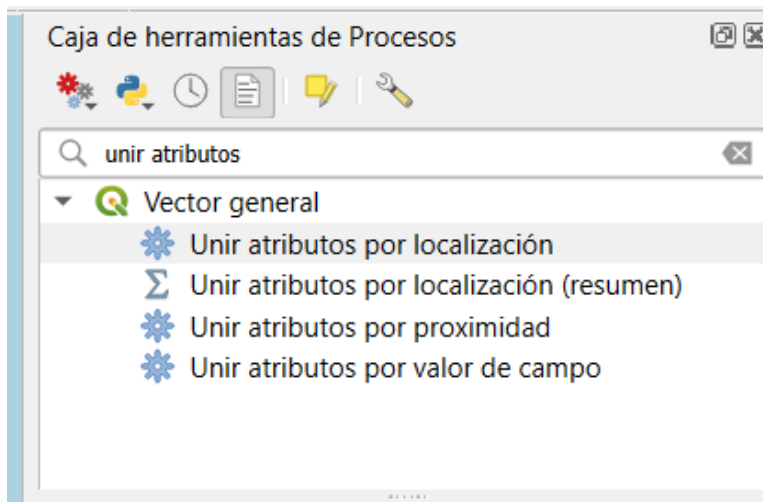


Una vez marcada la opción, tenemos que poner como capa de entrada la de zonas verdes y como segunda capa la de bicis:

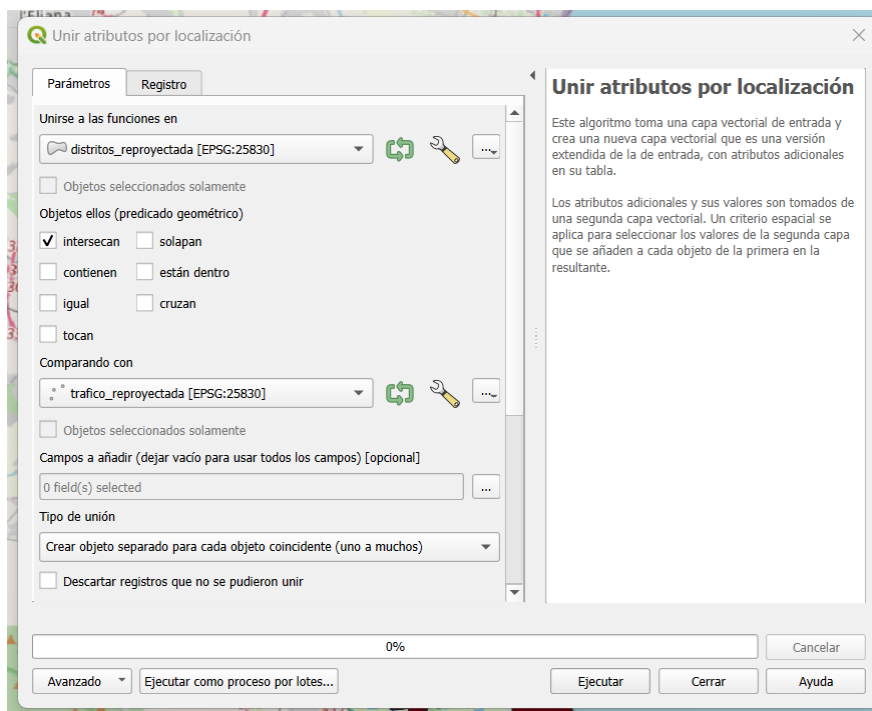


Una vez creada la capa de los buffers hacemos una unión de la capa de distritos reproyectada y la capa de tráfico reproyectada.

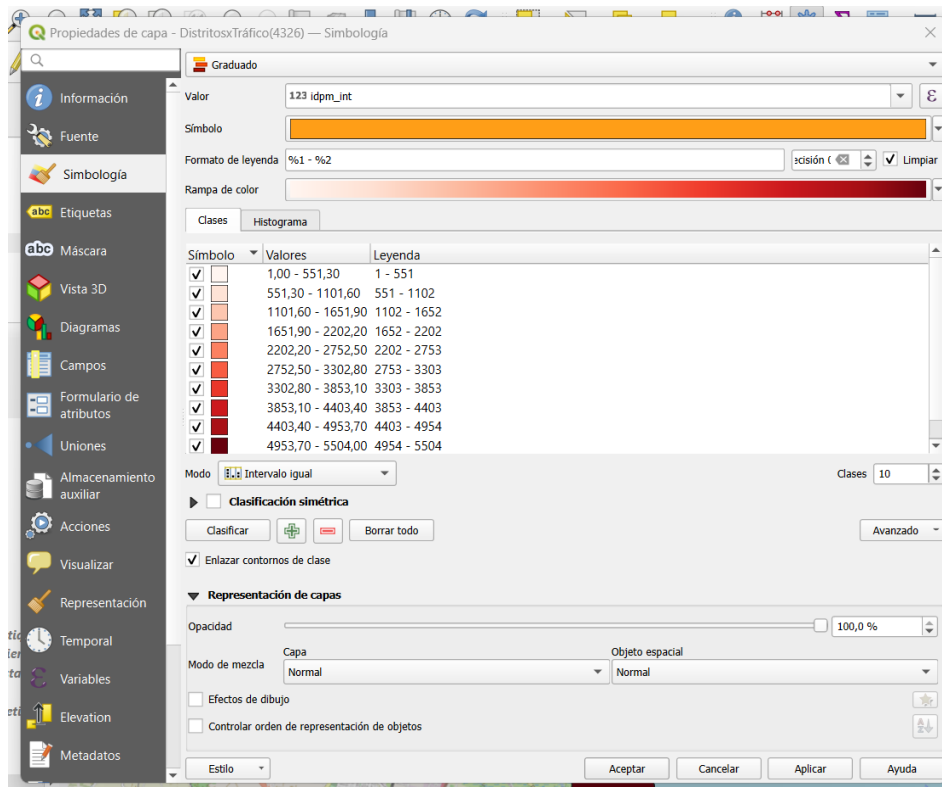
Para ello nos tenemos que ir al apartado herramientas de proceso y buscar Unir atributos por localización:



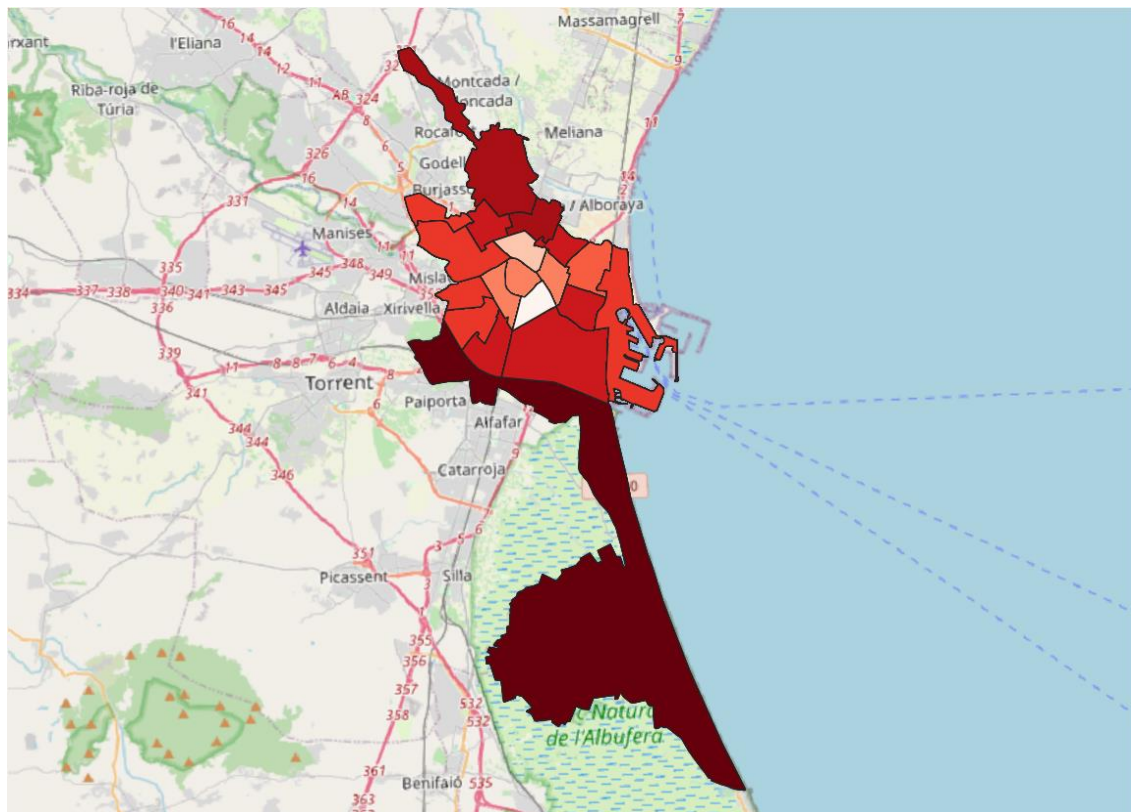
Una vez seleccionado, ponemos como capa de entrada la de distritos reprojectada y como segunda capa la de tráfico reprojectada:



Ya tenemos la capa creada. Ahora solo queda poner cada distrito de color según el tráfico. Para ello tenemos que ir a propiedades de la capa, simbología. Tenemos que poner el color de capa en tipo graduado, seleccionar como variable la de idpm_int y ponemos una rampa de color. El modo lo ponemos en intervalo igual y marcamos 10 clases:



La capa nos quedaría así:



2.4. Elección de gráficas para los distintos tipos de datos

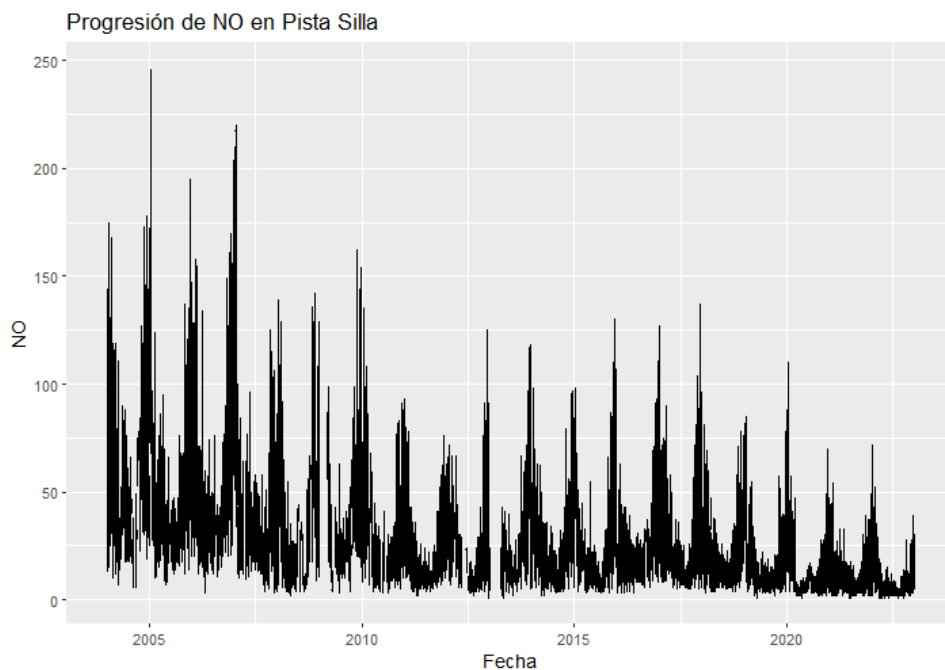
Para este proyecto hemos elegido 2 tipos de gráficos, ambos gráficos de líneas dado que estamos ante el desarrollo de una variable cuantitativa a lo largo del tiempo en ambas gráficas

La primera gráfica elegida es una sobre las estaciones y los gases. Para ello utilizamos los datos de calidad de aire (rvvcca.csv).

Para poder hacer la gráfica hemos tenido que transformar la columna de fecha a tipo date con su formato ya que lo que queremos es representar la progresión de un gas a lo largo del tiempo:

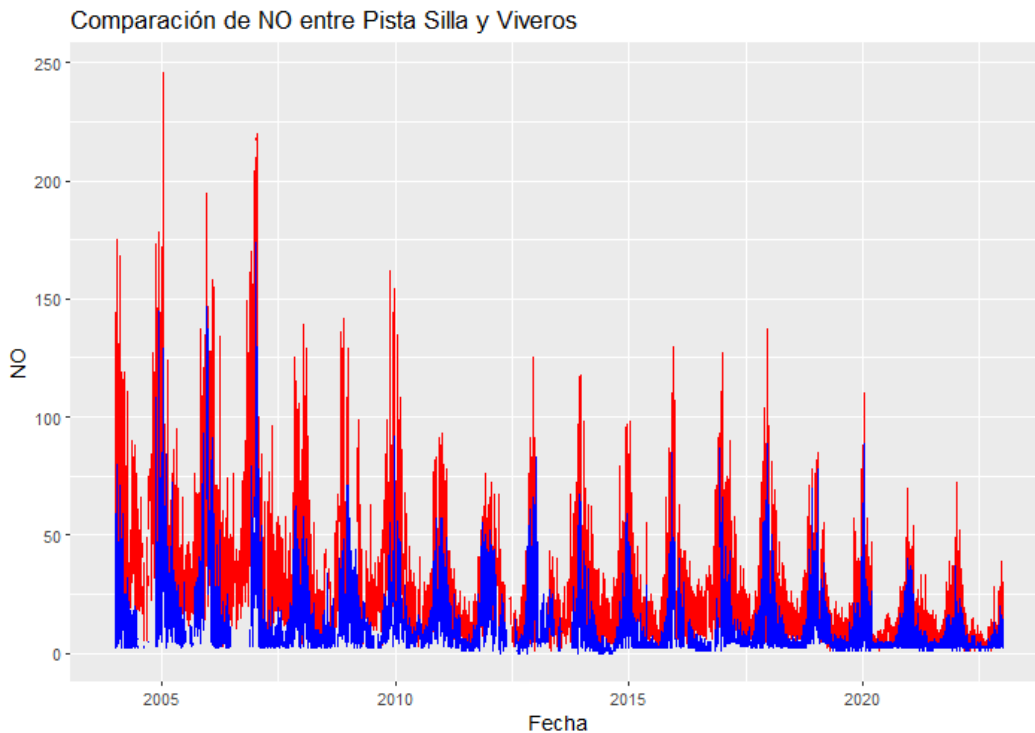
```
calidad_aire$Fecha <- as.Date(calidad_aire$Fecha, format = "%Y-%m-%d")
```

Una vez tenemos los datos como queremos procedemos a realizar el gráfico nos queda de esta manera:



Para el segundo gráfico hemos querido ver la comparación de los gases en dos estaciones a la vez por lo que un gráfico de líneas era la mejor opción.

Al tener los datos como queremos solo tenemos que superponer una gráfica sobre otra, quedaría de esta manera:



2.5. Diseño del mapa e interactividad

En el primer mapa se ha integrado una capa base con el mapa de Valencia. Encima hemos colocado la nueva capa creada en Qgis de distritos y tráfico, la cual está graduada por colores con su respectiva leyenda situada abajo a la izquierda. Además, hemos puesto sobre los distritos unos puntos que representan el IMD mensual de vehículos motorizados con una escala de colores que indica si el IMD es alto o bajo.

En el segundo mapa final también se ha integrado una capa base con el mapa de Valencia donde hemos superpuesto una capa con las zonas verdes de Valencia. También hemos puesto una capa de puntos negros que indican las bicicletas de Valencia y otra capa de puntos que representa la intersección de los buffers de zonas verdes y bicicletas. Además tenemos otra capa de las estaciones que aparecen representados por puntos que varían su color según la calidad ambiental con su respectiva leyenda abajo a la derecha.

Hemos dispuesto así las capas (entre los dos gráficos) para centrar el primer gráfico en el tráfico y todo lo relacionado con él y el segundo mapa con el medioambiente y la contaminación.

Se ha utilizado un ancho del 100%, una altura de 650 y, un zoom de 12 en el primero y 13 en el segundo. Hemos centrado los mapas sobre la estación llamada Centro. Los mapas son interactivos de manera que al clicar en cualquier distrito o estación podemos observar el nombre de estos.

2.6. Diseño del cuadro de mandos

Interfaz de Usuario (UI)

1. titlePanel: Define el título principal del cuadro de mandos.
2. tabsetPanel: Crea un conjunto de pestañas para organizar el contenido en diferentes secciones. Cada pestaña tiene su propio panel.
 - tabPanel: Cada pestaña contiene un panel con contenido específico. En este caso, hay cuatro pestañas: "Tráfico por distritos", "Contaminación y medioambiente", "Información sobre gases por distrito" y "Comparación de gases de estaciones".
 - sidebarLayout: Define un diseño con una barra lateral y un panel principal para algunas de las pestañas. Esto permite incluir controles de selección en la barra lateral y mostrar resultados en el panel principal.
 - sidebarPanel: Contiene controles de selección (selectInput) para filtrar datos o configurar parámetros.
 - mainPanel: Muestra resultados como mapas o gráficos.

Nosotros hemos necesitado cuatro pestañas, dos para los mapas interactivos y otras dos dedicadas a gráficos

Server

El servidor (server) está definido por la función server(), que toma tres argumentos: input, output, y session. Este es el "backend" de la aplicación, donde se define cómo reaccionar a las interacciones del usuario y cómo generar resultados dinámicos.

1. output\$...: Define las salidas que serán mostradas en la interfaz de usuario.
 - renderLeaflet(): Renderiza un mapa interactivo utilizando la biblioteca Leaflet.
 - renderPlot(): Renderiza gráficos utilizando ggplot2.
2. input\$...: Permite acceder a los valores seleccionados por el usuario en la interfaz de usuario. Se utiliza para filtrar datos y personalizar visualizaciones.

En las dos primeras pestañas no hemos usado inputs pero en la tercera hemos usado dos, uno para elegir el distrito y otro para elegir el gas. En la cuarta pestaña el usuario debe escoger dos distritos y un gas por lo que hemos usado 3 inputs

2.7. Implementación

Mapas: Se utilizan datos geoespaciales para representar información sobre distritos, movilidad y estaciones contaminantes. Se utilizan funciones de Leaflet para agregar capas de polígonos, círculos y marcadores al mapa cómo addPolygons, addCircles y addCircleMarkers. También se añaden leyendas para facilitar la interpretación de los datos con la función addLegend.

Gráficas: Se utilizan datos de calidad del aire para generar gráficos de líneas que muestran la progresión de un gas específico a lo largo del tiempo en una estación

seleccionada. Además, se proporciona una explicación del gas seleccionado. Para crear las gráficas se han utilizado las funciones renderPlot, ggplot y geom_line.

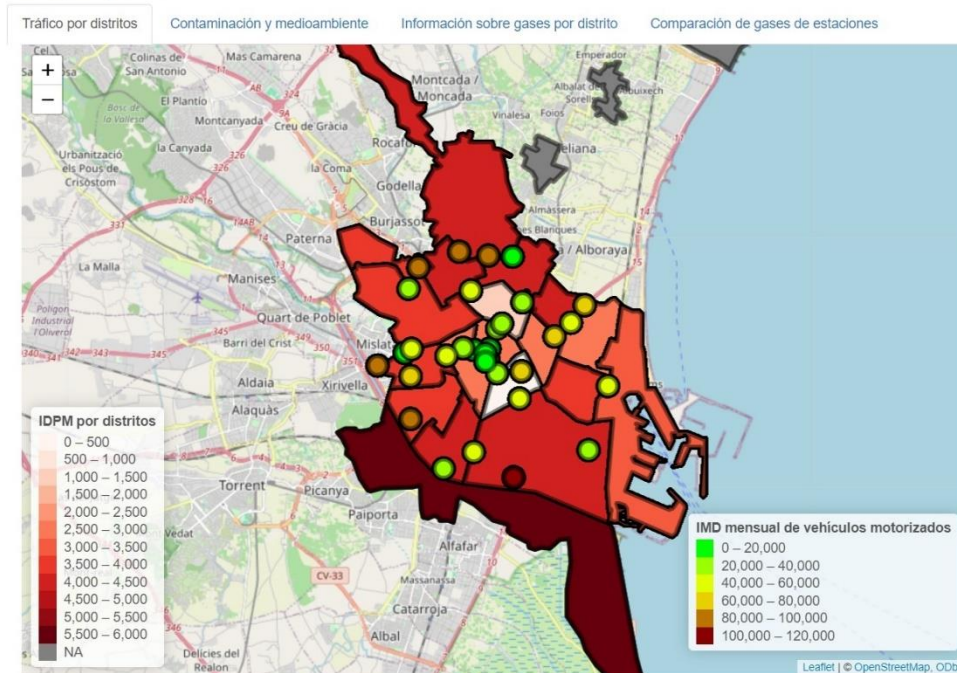
Comparación de estaciones: Se comparan dos estaciones seleccionadas en términos de un gas específico a lo largo del tiempo. Se utilizan dos líneas en el gráfico para representar los datos de las dos estaciones. Se utilizan las mismas funciones que en las gráficas normales.

3. Resultados

Después de realizar los codigos hemos tenido estos resultados:

Tráfico por distritos

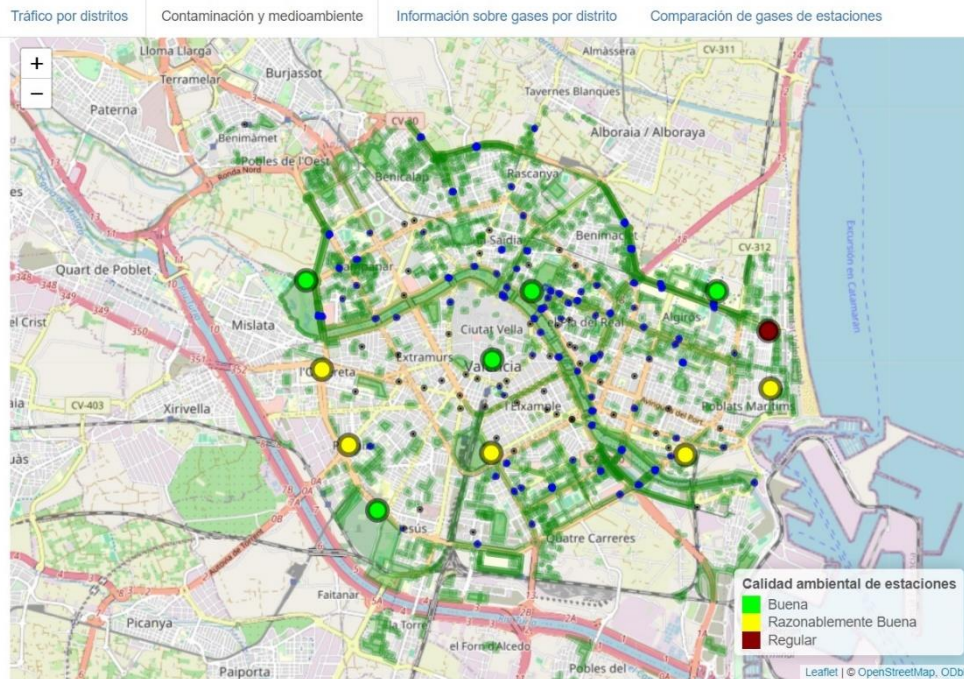
Miniproyecto



En este mapa observamos que los distritos del centro son los de menor IPMD, esto concuerda con los datos de IMD que nos proporcionan los puntos, donde comprobamos que los de menor IMD se encuentran también en el centro. Es decir, los puntos más verdes se encuentran sobre distritos más claros, distritos que se encuentran en el centro de la ciudad.

Contaminación y medioambiente

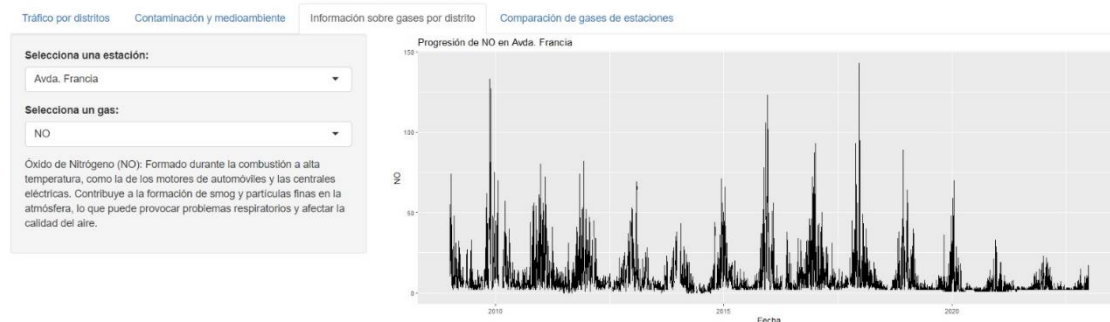
Miniproyecto



En este mapa podemos comprobar que las estaciones con calidad ambiental 'Buena' se encuentran en la parte de arriba de la ciudad que es a su vez la que más zonas verdes y bicicletas tiene. Los puntos azules muestran las bicicletas que están cerca de zonas verdes, podemos comprobar que hay muchos más puntos azules del antiguo cauce del río Turia hacia arriba, en cambio hacia abajo no encontramos muchos puntos azules y hay muchas estaciones con calidad ambiental 'Razonablemente buena'.

Información sobre gases por distrito

Miniproyecto



Comprobamos que para casi cualquier estación y casi cualquier gas hay una estación del año donde la medida sube drásticamente y otra que baja.

En el caso del **NO** los picos se consiguen, para cualquier estación, en invierno y en verano baja drásticamente.

Pero ocurre, al contrario, por ejemplo, con el **NO2** que sube en verano y baja en invierno.

Comparación de gases de estaciones

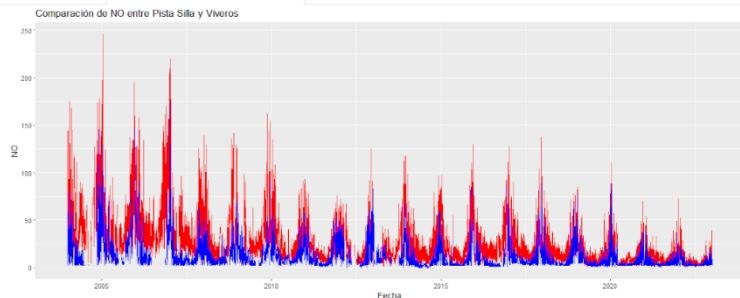
Miniproyecto

Tráfico por distritos Contaminación y medioambiente Información sobre gases por distrito Comparación de gases de estaciones

Selección de la estación 1:
Pista Silla

Selección de la estación 2:
Viveros

Selección de un gas:
NO



En esta gráfica podemos comparar la medida de los gases en diferentes estaciones, esto permite sacar patrones y comparaciones para cada combinación, por lo que pondremos un ejemplo, usaremos las estaciones Pista de Silla y Viveros para el **NO**.

Por norma general, para cualquier estación y gas hay un patrón parecido, es decir, los valores de este gas suben y bajan en las mismas fechas, por lo que son directamente proporcionales

En este ejemplo concreto podemos observar que en la estación de Pista de Silla se mide en casi cualquier fecha más cantidad de **NO** que la de Viveros.

4. Discusión

Observamos que, de manera general, los distritos céntricos de Valencia, como Ciutat Vella son los menos afectados por la contaminación. Mientras que las zonas más exteriores, como Poblats del Sud son las más afectadas por la misma.

Llegamos a la conclusión de que este suceso es debido a las diferencias que hay entre distritos, ya que el centro de Valencia tiene las carreteras más saturadas, pero están formadas por uno o dos carriles y suelen estar saturados, por lo que no pueden circular una cantidad alta de coches.

Por otra parte, si nos centramos en el distrito de Poblats del Sud, aunque puede que parezca una de las zonas más limpias de Valencia por la Albufera, hay que tener en cuenta que cuenta con carreteras en la que pasan coches y autobuses constantemente, por lo que al final acaban pasando una cantidad mayor de vehículos a motor en comparación con los distritos que se encuentran en el centro de Valencia.

Las limitaciones del estudio pueden ser la falta de datos o de recursos, tanto de material como personal.

Hemos encontrado dificultades a la hora de cargar datos de algunos ficheros y capas, ya que aparecían con la clase de variable incorrecta, como por ejemplo puede ser el idpm como cadena (str), en lugar de entero (int).

Lo que hemos hecho para solucionar este tipo de problemas ha sido crear una nueva columna, en la que hemos reconvertido la columna entera, quitando los espacios existentes y pasando correctamente al tipo de dato entero (int).

De esta forma, hemos podido separar por diferentes clases y pintar por colores, mediante una leyenda monocromática roja al ser ya una variable numérica.

Por otra parte, hemos encontrado muchos valores faltantes en gran parte de columnas de algunos archivos que hemos cargado, como por ejemplo el amoníaco (NH₃) en el fichero de calidad del aire. Apenas hay estaciones de medida de la contaminación atmosférica con datos de esta variable, por lo que algunas gráficas aparecen vacías para determinados distritos.

Además, hemos tenido dificultades en cuanto al uso del sistema de referencias de coordenadas, ya que hemos trabajado mediante el 25830 y el 4326 y hemos tenido que reproyectar algunas capas para pasar de QGIS a R.

5. Conclusiones

Hemos observado que conforme vamos saliendo del centro de Valencia a los distritos más exteriores, la contaminación debido a los vehículos motorizados es mayor, mientras que las zonas verdes y los puntos de bicicletas son menores.

Además, como hemos observado anteriormente, las tendencias que siguen los gases de la calidad del aire se corresponde a todos los distritos de Valencia.

Hemos logrado todos los objetivos planteados, ya que hemos utilizado todos los datos disponibles en base a los distritos. De esta forma, hemos realizado el estudio de la contaminación en base a los distritos de Valencia.

Los distritos del centro de Valencia tienen una cantidad de contaminación menor que los distritos exteriores. Esto lo podemos ver mediante los mapas y gráficas creadas en el miniproyecto.

6. Referencias

<https://valencia.opendatasoft.com/explore/dataset/rvvcca/information/>

<https://valencia.opendatasoft.com/explore/dataset/estacions-contaminacio-atmosferiques-estacionescontaminacion-atmosfericas/table/>

<https://valencia.opendatasoft.com/explore/dataset/districtesdistritos/information/?location=10,39.42291,-0.35395&basemap=e4bf90>

<https://valencia.opendatasoft.com/explore/dataset/punts-mesura-traffic-espires-electromagnetiquespuntos-medida-traffic-espiras-ele/table/>

https://valencia.opendatasoft.com/explore/dataset/qlik_datos_imd_mobilitat_enero_2016-diciembre-2022_coord/information/

<https://valencia.opendatasoft.com/explore/dataset/zonasverdes/table/?disjunctive.nivel3>

<https://valencia.opendatasoft.com/explore/dataset/punts-mesura-bicis-espires-electromagnetiquespuntos-medida-bicis-espiras-electr/map/?location=4,21.04195,-3.90988&basemap=e4bf90>