

Grado de Ciencia de Datos, ETSE-UV  
Visualización de Datos

# MEMORIA DE MINIPROYECTO: Estudio de la calidad del aire - Valencia

Raúl Rodríguez,  
Feiyi Yu,  
Israel Carbonell,  
Joyce Kemeni,  
Yoana Boyanova

# Indice

1.	Introducción	
1.1.	Presentación del contexto y la importancia del problema .....	p.2
1.2.	Revisión de la literatura relevante .....	p.2
1.3.	Objetivos del estudio .....	p.2
2.	Metodología	
2.1.	Análisis exploratorio de los datos .....	p.3
2.2.	Preprocesado de Datos .....	p.3
2.3.	Preprocesado de la información geográfica .....	p.6
2.4.	Elección de gráficas para los distintos tipos de datos.....	p.9
2.5.	Diseño del mapa e interactividad	
2.5.1.	Mapa de coropletas.....	p.11
2.5.2.	Mapa de tráfico y contaminación en tiempo real.....	p.12
3.	Resultados	
3.1.	Presentación de los datos obtenidos de manera clara y organizada.....	p.14
3.2.	Utilización de tablas, gráficos u otros recursos visuales para facilitar la comprensión.....	p.14
4.	Discusión	
4.1.	Interpretación de los resultados en relación con los objetivos del estudio y la literatura existente.....	p.17
4.2.	Análisis de las implicaciones de los hallazgos.	
4.3.	Discusión sobre las limitaciones del estudio y posibles áreas para futuros trabajos.	
4.4.	Dificultades encontradas.....	p.19
5.	Conclusiones	
5.1.	Recapitulación de los hallazgos principales.....	p.20
5.2.	Respuestas a los objetivos planteados en la introducción.....	p.21
5.3.	Implicaciones prácticas y científicas de los resultados.....	p.21
6.	Referencias.....	p.22

# 1.Introducción

## 1.1 Presentación del contexto y la importancia del problema.

La composición del aire es de gran importancia para nosotros como raza humana, pues el aire es un elemento fundamental para tanto nuestra supervivencia, como la de los demás seres vivos. Es por ello, que este proyecto con enfoque al estudio de la contaminación atmosférica se ajusta con una de las mayores problemáticas a las que nos enfrentamos hoy en día: el impacto de la contaminación y el cambio climático en la salud humana y el bienestar del planeta. En el estudio nos centraremos en los datos de las estaciones medidoras de calidad ambiental de Valencia establecidas por la Red Valenciana de Vigilancia y Control de la Contaminación Atmosférica (RVVCCA), principalmente en los datos monitorizados de los principales contaminantes gaseosos: SO<sub>2</sub>, NO<sub>2</sub> y O<sub>3</sub><sup>[1]</sup>. Además hemos incluido en nuestro estudio las PM<sub>2.5</sub> y PM<sub>10</sub><sup>[2]</sup>, partículas finas de tamaño inferior a 2.5 µm e inferior a 10 µm respectivamente.

## 1.2 Revisión de la literatura relevante

Dado el enfoque de nuestro estudio, nos hemos basado principalmente en otros estudios del aire realizados, así como en la documentación ofrecida por la OMS (Organización Mundial de la Salud)<sup>[1]</sup>, por la Generalitat Valenciana<sup>[3]</sup> y el Ministerio para la Transición Ecológica y el Reto Demográfico del gobierno español<sup>[4]</sup>.

## 1.3 Objetivos del estudio

La problemática de la calidad del aire es causa de múltiples problemas de salud y nos afecta a nivel personal. Sin embargo, también es la humanidad la que favorece el aumento de la contaminación atmosférica y, por tanto, se convierte en nuestra responsabilidad como sociedad disminuirla.

El principal objetivo de este estudio es ofrecer al lector y/o usuario la información clave para hacer frente a la polución, para así contribuir a la posibilidad de alcanzar las metas del ODS 3 (Salud y Bienestar), así como informarle de las medidas adecuadas que puede tomar para disminuirla, contribuyendo a los ODS 13 y 15 (Acción por el clima y Vida de ecosistemas terrestres, respectivamente). Para ello, el proyecto incluye un componente interactivo, que le permite llegar a los ciudadanos de forma accesible y eficaz, es decir, se establece una relación proveedor-usuario.

Otra de nuestras metas es la concienciación sobre la importancia del efecto que tiene la contaminación atmosférica sobre la salud. Según los estudios del año 2019 de la OMS<sup>[1]</sup>, la contaminación atmosférica provoca aproximadamente 4.2 millones de muertes prematuras a nivel mundial cada año. Además en el caso del territorio español, los niveles de NO<sub>2</sub>, PM<sub>10</sub> y PM<sub>2.5</sub> superan los niveles establecidos como aceptables por la OMS en la mayor parte del territorio español, especialmente este último, el de las PM<sub>2.5</sub> que de hecho tienen efectos más severos sobre la salud que las de un tamaño mayor, ya que son 100% respirables y se

componen de compuestos orgánicos y metales pesados<sup>[5]</sup>, que supera el valor recomendado en el 93% del territorio español, según un estudio de la Moncloa de 2022<sup>[6]</sup>.

## 2. Metodología

### 2.1 Análisis exploratorio de los datos

Para la realización del proyecto hemos utilizado los datasets obligatorios, la contaminación, los distritos y la calidad del aire. También algunos de los opcionales.

El conjunto de datos de distritos, es un dataset georeferenciado que contiene el nombre del distrito junto a su id y el área.

El dataset de la contaminación contiene los principales contaminantes ambientales, a tiempo real (los datos son tomados cada 3 min), para diferentes localizaciones georeferenciadas. Además, contiene el tipo de zona y de emisión, que más adelante podrá ser utilizada en los mapas.

En cuanto al dataset de calidad de aire es el más importante, aquí se encuentran todos los contaminantes ambientales para cada estación con su histórico, empezando en 2004. Este conjunto de datos, no está georeferenciado y el nombre de las estaciones difiere de los nombres del dataset de contaminación, más adelante en la preparación de los datos se explicará como se ha solucionado.

Para los datos opcionales hemos usado los conjuntos de datos de zonas verdes y manzanas para analizarlos en QGIS. Además, hemos considerado importante el efecto que tiene el estado del tráfico sobre la medición de la contaminación, por lo que hemos incluido el conjunto de datos del tránsito en tiempo real para implementarlo junto al de contaminación, también en tiempo real.

### 2.2 Preprocesado de Datos

Esta es una de las partes más importantes porque dependiendo de lo que se haga habrá que manejar más adelante los datos de una forma u otra. Además, hay que asegurarse de que estén los datos en el formato correcto y no haya errores.

Para el dataset de distritos, hemos eliminado la columna de 'objectid' que no la vamos a usar y también la de 'gis\_gis\_dis' que contiene el área de cada distritos pero contiene NAs. Después, hemos transformado 'coddistrit' a integer (estaba en chr) y hemos calculado el área haciendo uso de la librería de sf.

Para la contaminación hemos transformado las columnas 'tipozona', 'tipoemision', 'calidad\_amb' y 'nombre' a factor por comodidad. Además, hemos cambiado las unidades de PM10 y PM2.5 dividiéndolas entre 10.



En el de calidad del aire también hemos convertido muchas de las columnas a numérico, el nombre de la estación lo hemos pasado a factor y hemos borrado otras columnas que, o son redundantes (Dia de la semana o del mes, que se puede sacar desde la propia fecha) o no lo vamos a utilizar para nada ('Id', 'fecha baja / creacion').

El primer problema al que nos hemos enfrentado es el de georeferenciar las estaciones de calidad aire, para esto había que utilizar el dataset de contaminación que contenía las ubicaciones, con el problema de que los nombres de las estaciones eran diferentes en ambos datasets y había estaciones que no compartían. Para resolverlo hicimos un vector de correspondencias con el nombre de las estaciones de calidad de aire que corresponden a los nombres de contaminación y unimos ambos datasets obteniendo los nombres correctos de las estaciones y la georreferencia. Las estaciones que no estaban en ambos las filtramos, eliminando así para el dataset con datos históricos (calidad\_aire) "Nazaret Meteo", "Conselleria Meteo", "Puerto Valencia", "Patraix", "Cabanyal" y "Dr. Lluch".

Después de hacer esto, ya tenemos georeferenciadas todas las estaciones que están en el dataset de contaminación y de calidad aire, no obstante quedan dos estaciones de calidad aire que no se han georeferenciado al unir los datasets pero tienen muchos registros como para eliminarlas, por tanto hemos tenido que hacerlo manualmente. Para ello, hemos usado la API de google cloud, Geocoding, y usando la librería de R ggmap hemos podido obtener las coordenadas de las estaciones "Puerto Iltic Turia" y "Puerto Moll Trans. Ponent".

Para los NAs podíamos realizarlo en QGIS o en R, decidimos hacerlo en R ya que teníamos experiencia previa. Lo primero es visualizar los NAs en cada una de las variables.

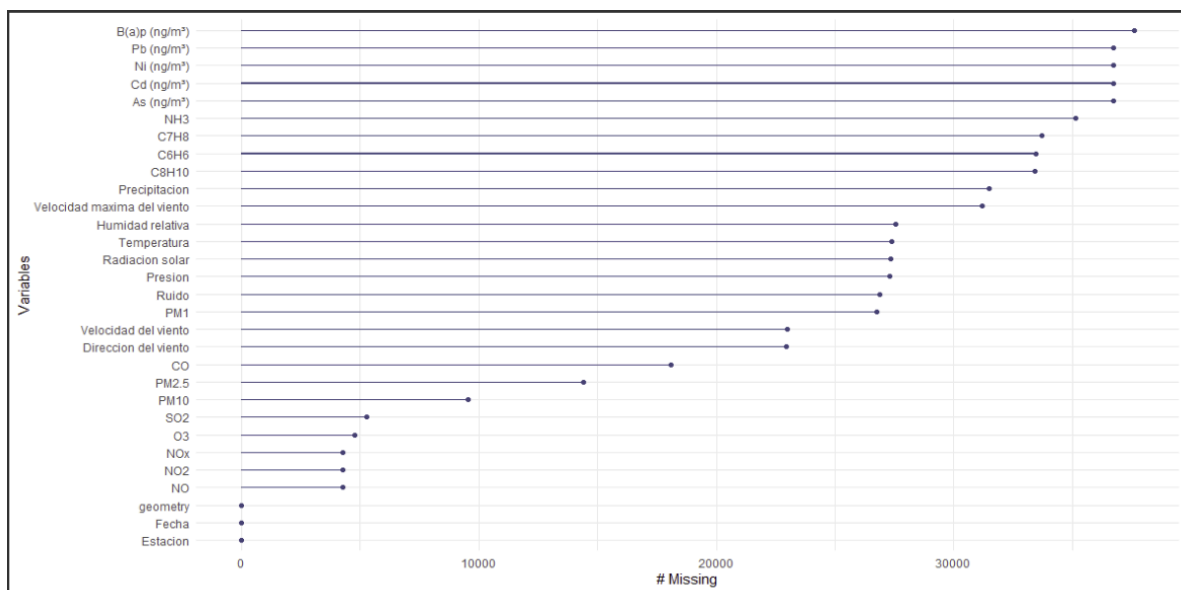


Figura 1. Representación de los datos faltantes

Tal y como vemos en la Figura 1, los principales contaminantes son los que menos NAs tienen, ya que tienen mayor importancia. Nos interesan los contaminantes NO<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> y en este caso el NO<sub>x</sub> y el NO también porque está relacionado con los elevados niveles de PM<sub>2.5</sub>. Como habían muchos registros con NAs como para eliminarlos decidimos imputarlos usando el paquete mice de R.

Lo primero fue ver la distribución de los datos, vimos que los datos estaban sesgados por la derecha, así que para evitar fallos en la imputación decidimos transformarlos a logaritmos, sumándole una constante muy pequeña a los ceros absolutos para poder transformarlos. Para aumentar la precisión había que tener en cuenta los datos temporales, por lo que hicimos para cada contaminante una variable que tuviera en cuenta la temporalidad, ej. SO<sub>2</sub> -> SO<sub>2</sub>ayer, esta variable transpone el valor original al registro de abajo relacionando cada valor con el valor que tenía esa variable el día anterior. Por último, hemos tenido en cuenta la localidad, para ello hemos transformado a numérico la estación.

Para comprobar qué método era mejor y si las pruebas que hacíamos eran mejores o peores que las anteriores, utilizamos un conjunto de datos de entrenamiento que consistía en 10000 datos de calidad del aire a los cuales le habíamos quitado los NAs y habíamos creado NAs artificiales de manera que teníamos los valores reales y los imputados, y podíamos comparar mediante el RMSE y la mean difference si lo que hacíamos era beneficioso y cuál era el mejor método para cada variable.

Este es el mejor resultado que hemos obtenido (están en escala logarítmica)

name <chr>	RMSE_PM2.5 <dbl>	MediaSumas_PM2.5 <dbl>	RMSE_PM10 <dbl>	MediaSumas_PM10 <dbl>	RMSE_SO2 <dbl>	MediaSumas_SO2 <dbl>	RMSE_NOx <dbl>
norm.boot	0.1128731	0.02274669	0.08800812	0.01585712	0.11815953	0.02429593	0.02403720
norm.nob	0.1133340	0.02414696	0.08792685	0.01653093	0.11759629	0.02393762	0.02444965
lasso.norm	0.1037898	0.02328124	0.08368427	0.01504113	0.11972134	0.02462741	0.02588130
pmm	0.1115618	0.02140851	0.08522016	0.01502377	0.12155144	0.02185150	0.02427622
midastouch	0.1105830	0.02148698	0.08528437	0.01533472	0.10739226	0.01846124	0.02308632
cart	0.1077188	0.02082699	0.08203520	0.01394747	0.08459369	0.01470743	0.02642253

Tabla 1a). Comparación del rendimiento y ajuste de métodos de imputación

MediaSumas_NOx <dbl>	RMSE_O3 <dbl>	MediaSumas_O3 <dbl>	RMSE_NO2 <dbl>	MediaSumas_NO2 <dbl>	RMSE_NO <dbl>	MediaSumas_NO <dbl>	RMSE_PM2.5ayer <dbl>
0.004792480	0.07933512	0.01623810	0.03302303	0.006810777	0.07333113	0.014619268	0.09683489
0.004696191	0.07033548	0.01387856	0.03270217	0.006214248	0.07532167	0.014976587	0.09573421
0.005223823	0.07389689	0.01458215	0.03099188	0.006338644	0.07734748	0.014881429	0.09685453
0.004135550	0.08107470	0.01552768	0.02682939	0.004923179	0.06687415	0.009477295	0.09022512
0.004075297	0.06764875	0.01279098	0.02730481	0.004955352	0.06773915	0.009102176	0.09309887
0.005288436	0.06599749	0.01288672	0.03547147	0.006922323	0.07475453	0.011121355	0.07238496

Tabla 1b). Comparación del rendimiento y ajuste de métodos de imputación

Viendo los resultados hemos seleccionado el método que mejor se ajusta a cada parámetro y hemos realizado la imputación para el dataset completo, reconvirtiendo las variables a la escala original y borrando las columnas temporales.

Por último, una vez imputados los datos calculamos la calidad ambiental, para ello hay que seguir los umbrales dados por las siguientes categorías:

SO <sub>2</sub>		PM <sub>2,5</sub>		PM <sub>10</sub>		O <sub>3</sub>		NO <sub>2</sub>		CATEGORÍA DEL ÍNDICE
0	100	0	10	0	20	0	50	0	40	BUENA
101	200	11	20	21	40	51	100	41	90	RAZONABLEMENTE BUENA
201	350	21	25	41	50	101	130	91	120	REGULAR
351	500	26	50	51	100	131	240	121	230	DESFAVORABLE
501	750	51	75	101	150	241	380	231	340	MUY DESFAVORABLE
751-1250		76-800		151-1200		381-800		341-1000		EXTREMADAMENTE DESFAVORABLE

Tabla 2. Tabla de colores de definición de la calidad del aire establecida por el Gobierno

### 2.3 Preprocesado de información geográfica

Un aspecto importante es saber el alcance que tenemos con nuestros datos. Como en este caso queríamos trabajar con los datos de estaciones hemos estimado conveniente realizar un buffer alrededor de las estaciones de contaminación atmosférica, dándonos así una idea aproximada de su zona de influencia.



Figura 2. Zonas de influencia de las estaciones

Aplicamos una estética sobre el campo NO<sub>2</sub> para que visualmente seamos capaces de discernir entre zonas con peor calidad ambiental y las mejores. Esta estética no se corresponde con el color RGB dado por el gobierno ya que para los datos que teníamos se

hubiesen agrupado en solo 2 categorías y no hubiese representado esa variación visualmente y de esta forma se representa mejor esa variación en los niveles de NO<sub>2</sub>

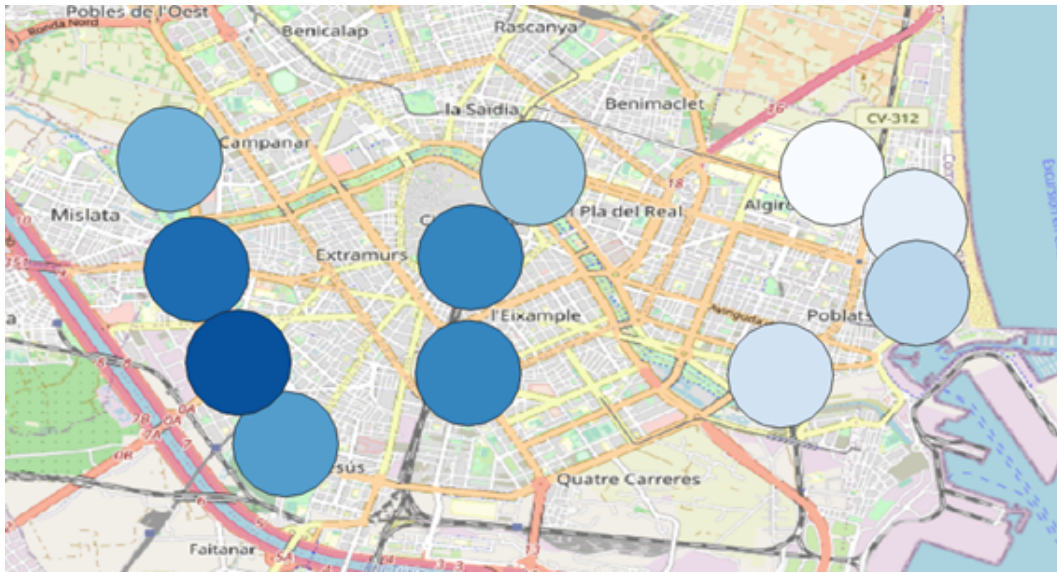


Figura 3. Zonas de influencia de las estaciones coloreadas según su nivel de NO<sub>2</sub>

Una de nuestras hipótesis era que aquellas áreas con menor número de zonas verdes, eran potencialmente las que peor calidad de aire de NO<sub>2</sub> tenían.

Para comprobarlo hemos hecho una intersección entre nuestro buffer anterior y el dataset de zonas verdes. Visualmente vemos que efectivamente áreas con aparentemente menos zonas verdes como la de las estaciones de Patraix, Olivereta y Centro son también aquellas con mayor concentración de NO<sub>2</sub>.

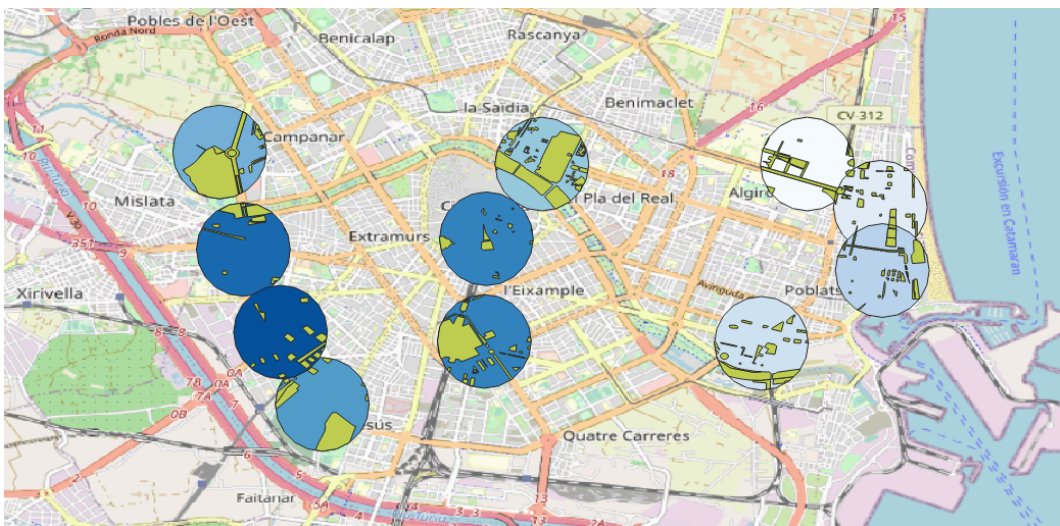


Figura 4. Zonas verdes contenidas en las zonas de influencia de las estaciones



Como no solo queríamos comprobarlo de manera visual, hemos utilizado la calculadora de campos sobre nuestra capa de intersección para que nos cree un nuevo campo donde esté calculada el área de cada zona verde correspondiente al buffer de su estación. De esta manera podemos calcular la proporción de área verde respecto al buffer y así comprobar si esas 3 estaciones son las que menor cantidad de zonas verdes contienen en su zona de influencia.

Inte\_verd\_buff—Objetos Totales: 312, Filtrados: 312, Seleccionados: 0

objectid	nombre	direccion	tipozona
226	Pista de Silla	C/ Filipinas, s/n	Urbana
227	Pista de Silla	C/ Filipinas, s/n	Urbana
228	Pista de Silla	C/ Filipinas, s/n	Urbana
229	Pista de Silla	C/ Filipinas, s/n	Urbana

area	tipoemision	fecha_carg
1966,366631887...	Tráfico	2024-04-08 14:2...
1894,565534576...	Tráfico	2024-04-08 14:2...
1822,653859052...	Tráfico	2024-04-08 14:2...
1704,820392921...	Tráfico	2024-04-08 14:2...

Tablas 3 a) y b). Tabla de atributos de la capa de zonas verdes una vez calculada su área

Aprovechamos el buffer anterior de las estaciones para realizar una intersección con un dataset de Manzanas que contiene información sobre la población de las mismas.

La finalidad de esto es saber, no solo la zona de influencia de cada estación, sino a cuánta gente afectará la calidad del aire correspondiente a dicha estación.

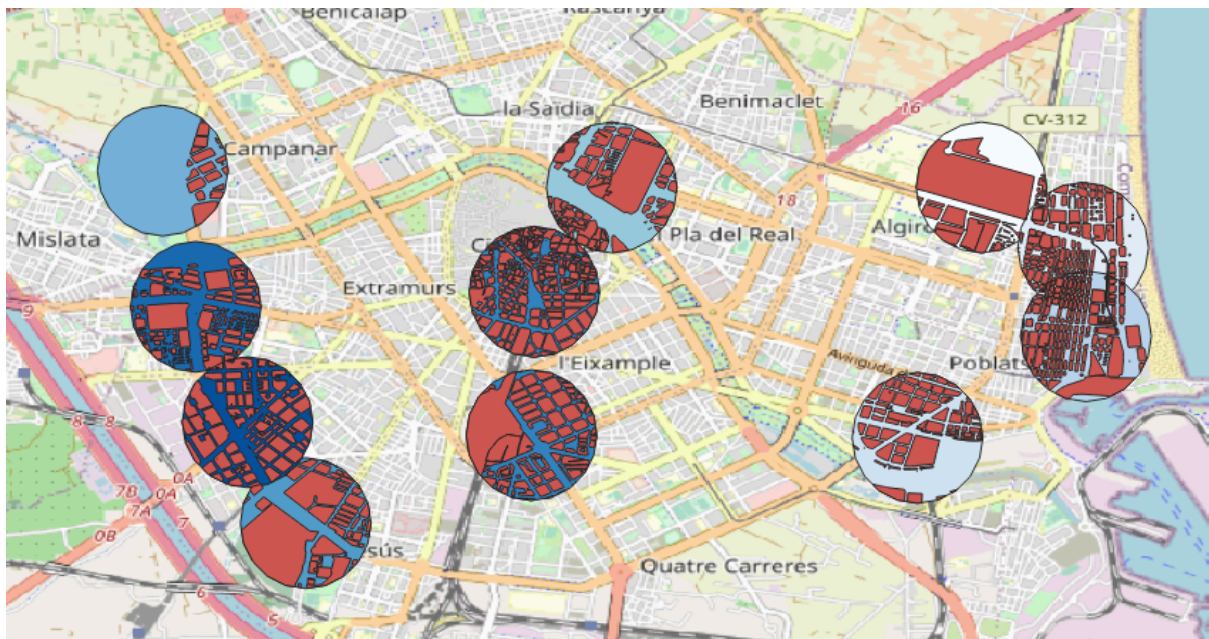


Figura 5. Manzanas contenidas en las zonas de influencia de las estaciones

## 2.4 Elección de las gráficas para los distintos tipos de datos

Para el primer gráfico hemos escogido un mapa de calor ya que tenemos muchos datos que de otra forma sería difícil representar sin que quede saturado. Hemos representado la calidad ambiental para cada día, representando en el eje 'x' la semana y en el eje 'y' el día de la semana, además hemos separado por meses en cada columna y por año en cada fila.

Para cada categoría hemos seleccionado un código de color rgb dado por [la página del gobierno](#):

- Buena (azul): 56, 162, 206
- Razonablemente buena (verde): 50, 161, 94
- Regular (amarillo): 241, 229, 73
- Desfavorable (rojo): 200, 52, 65
- Muy desfavorable (granate): 110, 22, 29
- Extremadamente desfavorable (morado): 162, 91, 164

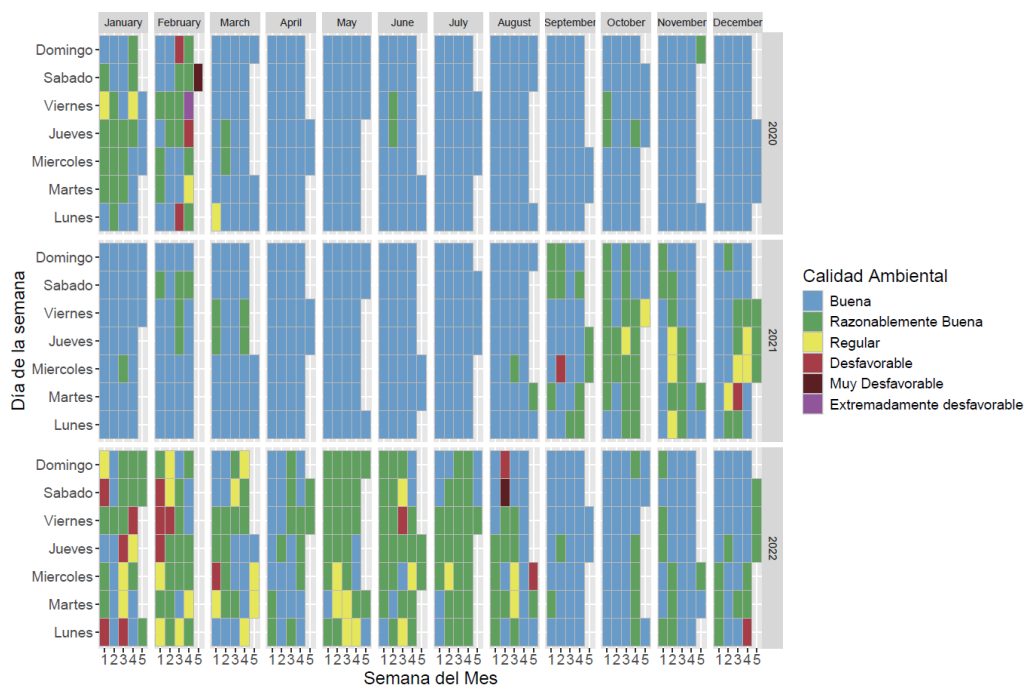


Figura 6. Captura del mapa de calor referente a la calidad del aire según el tiempo

El segundo gráfico es un histórico en el que se representa la evolución de un parámetro, el gráfico está dividido por filas, cada fila corresponde a un parámetro. El usuario puede elegir en este gráfico el intervalo de fechas, los parámetros y la estación.

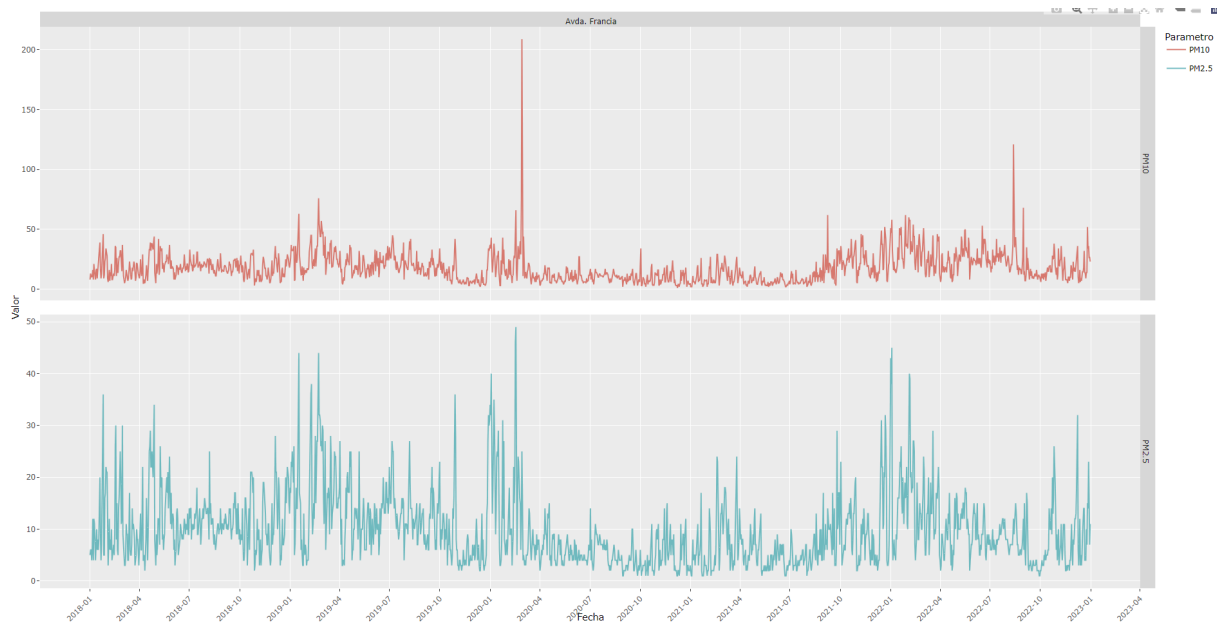


Figura 7. Captura del gráfico histórico de representación de la evolución temporal elegido por el usuario de uno o varios parámetros y una estación seleccionados

Para el tercer gráfico hemos hecho otro histórico de áreas en el cual se puede representar un parámetro en función de una variable temporal, el usuario puede elegir un único parámetro a representar, el intervalo de tiempo y las diferentes estaciones que quiera el usuario incorporar al gráfico.

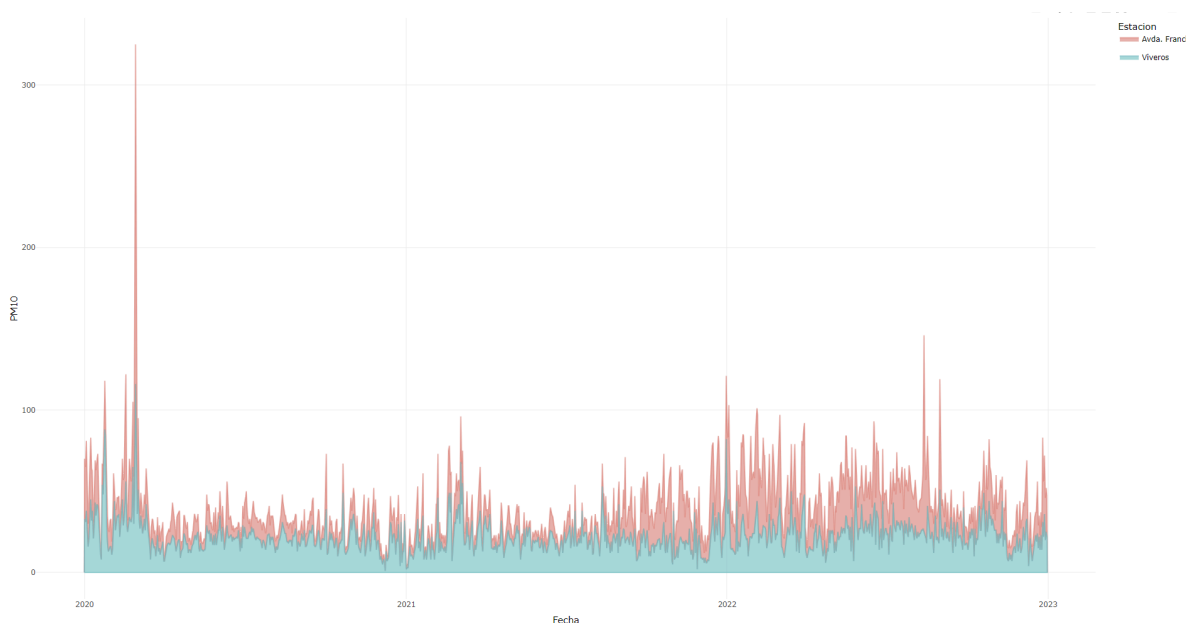


Figura 8. Captura del gráfico histórico de representación de la evolución temporal elegido por el usuario de un parámetro y una o varias estaciones de medición seleccionadas

Para el último gráfico hemos comparado un parámetro con otro, para que el usuario pueda ver tendencias y relaciones. En shiny se puede seleccionar la estación, la fecha y las variables de cada eje.

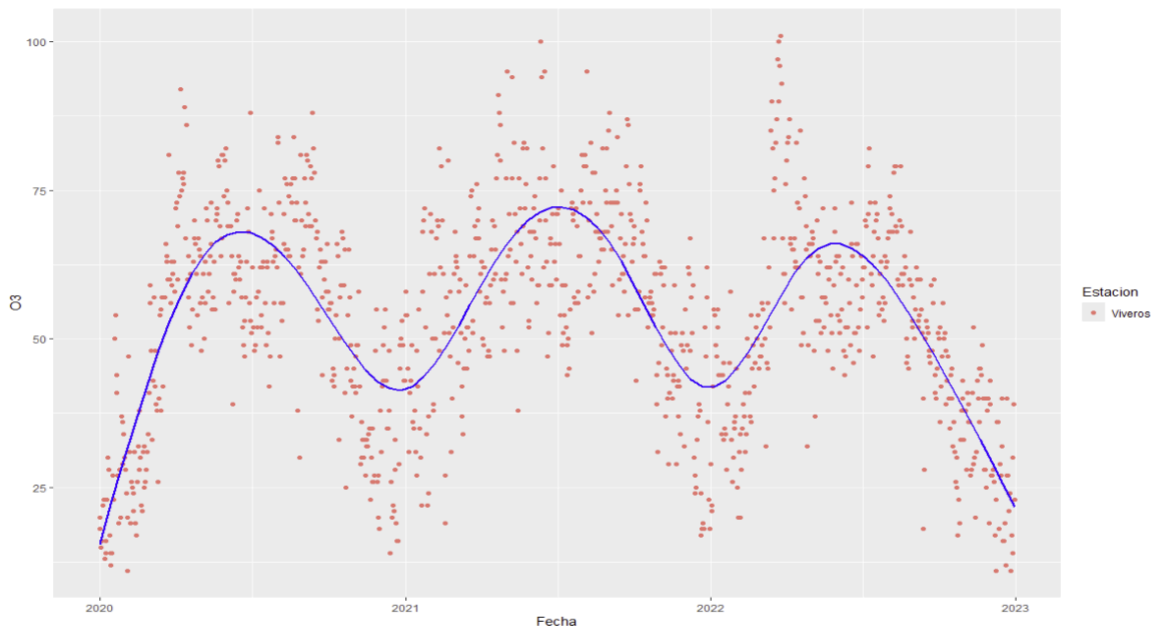


Figura 9. Captura del gráfico histórico de estudio de la relación entre dos parámetros seleccionados en una estación

## 2.5 Diseño del mapa e interactividad

### 2.5.1 Mapa de coropletas

En el caso de la representación de los gases, hemos considerado que sería interesante para el usuario poder visualizar la medición de cada uno de los gases disponibles según el distrito de la ciudad de Valencia en el que se encuentre cada estación de medición. Gracias al cuadro de mandos en shiny que hemos añadido, el usuario puede seleccionar el gas de interés y la fecha en la que lo quiere visualizar.

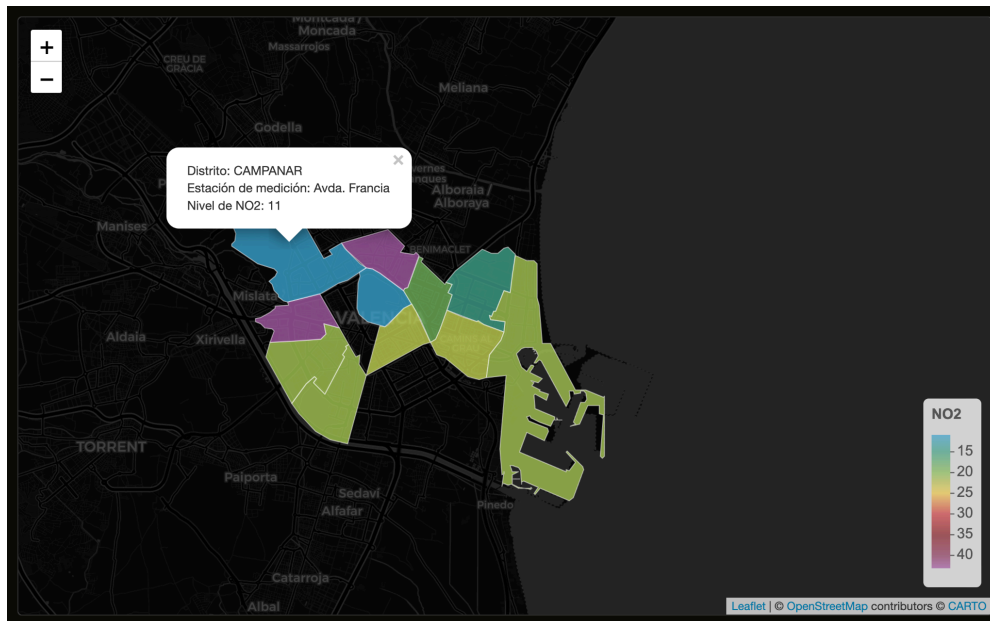
Para ello hemos representado la capa de los distritos que contienen una estación, una capa que hemos obtenido previamente seleccionando dichos distritos desde la tabla de atributos. Además hemos cargado en el entorno la capa del histórico de los datos, para poder rellenar los popups de cada distrito con el nombre de la estación de medición en ese distrito, el nombre del distrito y el nivel del gas seleccionado por el usuario en dicho distrito. Esta capa se ha utilizado además para la coloración y la leyenda del mapa.

Los elementos interactivos del mapa, a parte de los widgets de shiny para la personalización por parte del usuario, son el zoom (puesto originariamente a nivel de 12.3 en las coordenadas  $39.47^\circ$  de latitud y  $-0.335^\circ$  de longitud).

En cuanto a la estética de la cartografía base, nos hemos decantado por el proveedor CartoDB, en especial por su mapa base Dark Matter, debido a que queríamos resaltar los tonos de colores de los gases. Dado que las variables de los datos son numéricas continuas,



hemos implementado una escala de colores graduada de los colores mostrados en la Tabla de colores anteriormente señalada.



*Figura 10. Captura del mapa de coropletas implementado para visualizar los niveles de un parámetro y fecha seleccionados por el usuario*

## 2.5.2 Mapa de tráfico y contaminación en tiempo real

Nuestra decisión de crear este mapa se basa en el hecho de la necesidad de mejorar la interactividad con el usuario. Dado que el tráfico tiene un gran impacto en la contaminación atmosférica, pues es el causante del 10.4% de  $PM_{2.5}$  y del 43.4% de óxidos de nitrógeno<sup>[7]</sup>. Los datos de contaminación se están actualizando constantemente, así como los del estado de tráfico, por lo que hemos creado un mapa que se actualiza cada 3 minutos con la descarga de los mismos desde la página de datos abiertos del Ayuntamiento de Valencia.

Las capas que hemos utilizado en este caso son las de actualización periódica de la contaminación y del tráfico, así como la calculada a partir del buffer de las estaciones que representa la zona de influencia de las mismas, es decir, el área para el que se aplican las mediciones de cada estación. Además, hemos relacionado la contaminación con el área de las zonas verdes presentes dentro de cada zona de influencia, añadiendo una capa adicional con éstas.

Respecto a la estética, las líneas de la capa de tráfico han sido coloreadas según su estado (cuya descripción hemos obtenido a partir de los metadatos de la página de datos abiertos del Ayuntamiento de Valencia). Mientras tanto, para los colores de los marcadores de las estaciones hemos utilizado los colores señalados en la Tabla anteriormente mostrada según la calidad del aire en ese instante. Las zonas de influencia han sido añadidas en gris con un nivel de opacidad bajo para evitar la saturación, al igual que las zonas verdes solo que añadidas en color verde. Como cartografía base hemos elegido DarkMatter del proveedor CartoDB para resaltar los elementos del mapa. Cada capa también se corresponde con una leyenda, en la que también se incluye la posibilidad de no disponer de todos los datos en ese instante.

El mapa incluye varios elementos interactivos destinados a ofrecer un mejor servicio al usuario, como el zoom (puesto originariamente a nivel de 12.5 en las coordenadas 39.47° de latitud y -0.341° de longitud) que permite al usuario ampliar y reducir el mapa, así como diversos labels que se activan cuando el usuario posa el ratón sobre los elementos del mapa y popups para cada estación al pulsar sobre éstas, que indican el nombre de la estación, los niveles de cada gas y de las partículas finas actualizadas. Debido a que el mapa podría parecer sobresaturado al incluir todo el tráfico, el usuario tiene la opción de seleccionar y deseleccionar la capa del tráfico (la leyenda de esta capa se incluye o no según si la capa misma está o no seleccionada).

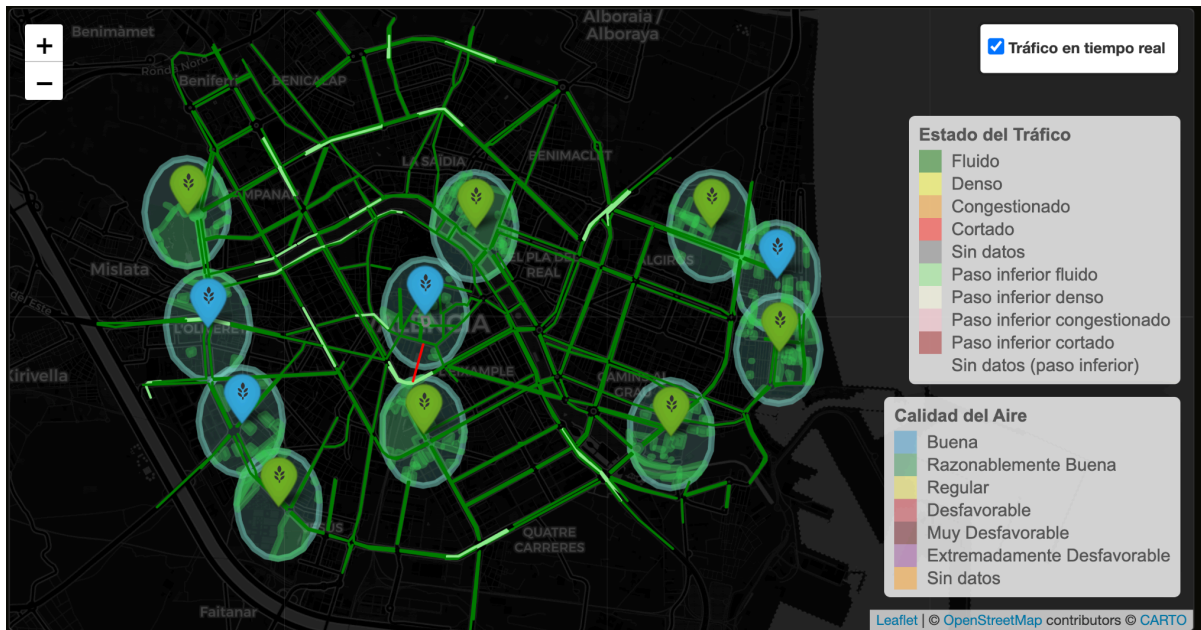


Figura 11. Mapa de contaminación y de tráfico a tiempo real

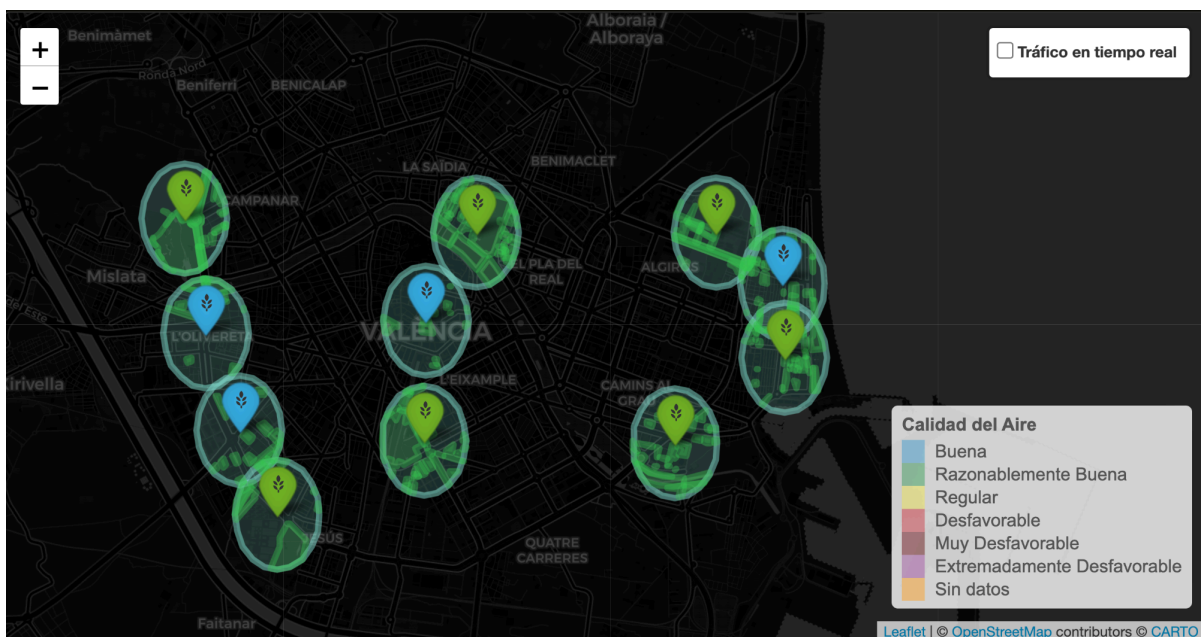


Figura 12. Mapa de contaminación y de tráfico a tiempo real, estando la capa de tráfico sin seleccionar

### 3. Resultados

#### 3.1. Presentación de los datos obtenidos de manera clara y organizada.

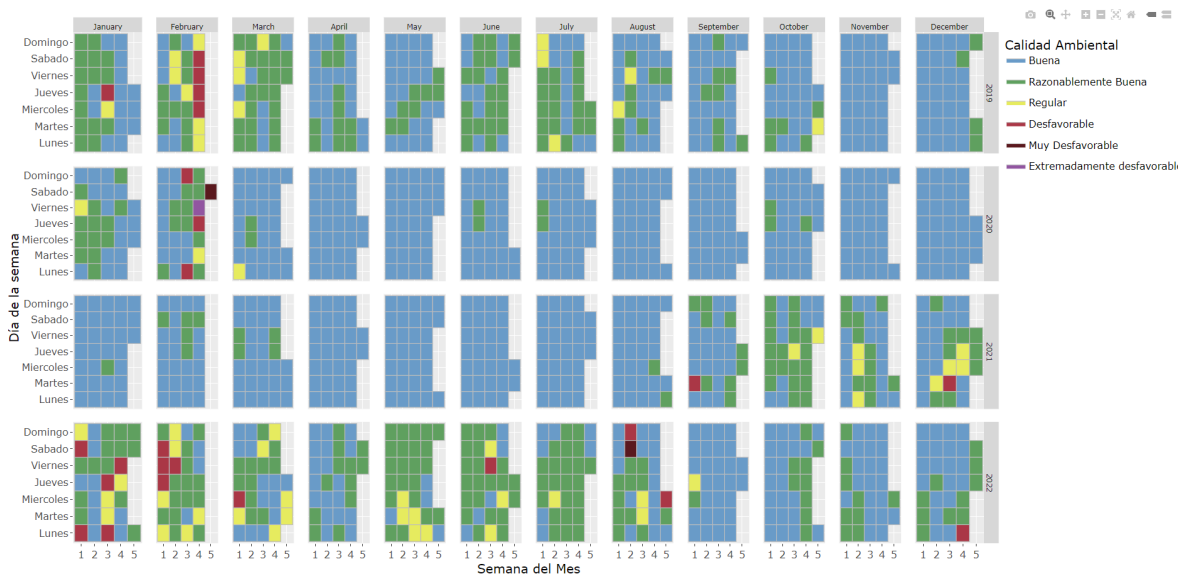
Hemos utilizado distintos métodos para visualizar la calidad ambiental y la concentración de los diferentes contaminantes en distintos lugares y momentos de la ciudad. Para tener una idea clara de en qué lugares la concentración de cierto parámetro es mayor y que zonas están más afectadas hemos utilizado un mapa de coropletas, de esta manera podemos representar espacialmente cada uno de los distritos y comparar fácilmente.

Para ver la evolución de cada parámetro a lo largo del tiempo hemos utilizado tres gráficos diferentes, en cada uno se puede observar diferentes eventos y tendencias. En el mapa de calor podemos observar en qué meses o etapas la calidad ambiental para un parámetro y una estación concreta es peor o mejor. En el primer gráfico histórico se puede ver para una estación la evolución de los diferentes parámetros a lo largo del tiempo, pudiendo observar en qué meses hay picos y cuando decrece el valor del parámetro. El segundo gráfico histórico es lo contrario, se puede ver la evolución de un parámetro para cada estación, permitiendo la clara visualización de qué estación está más afectada. Por último, hemos creado un gráfico en el que relacionamos un parámetro con otro para ver la tendencia y si las diferentes variables están relacionadas y así poder interpretar y sacar conclusiones.

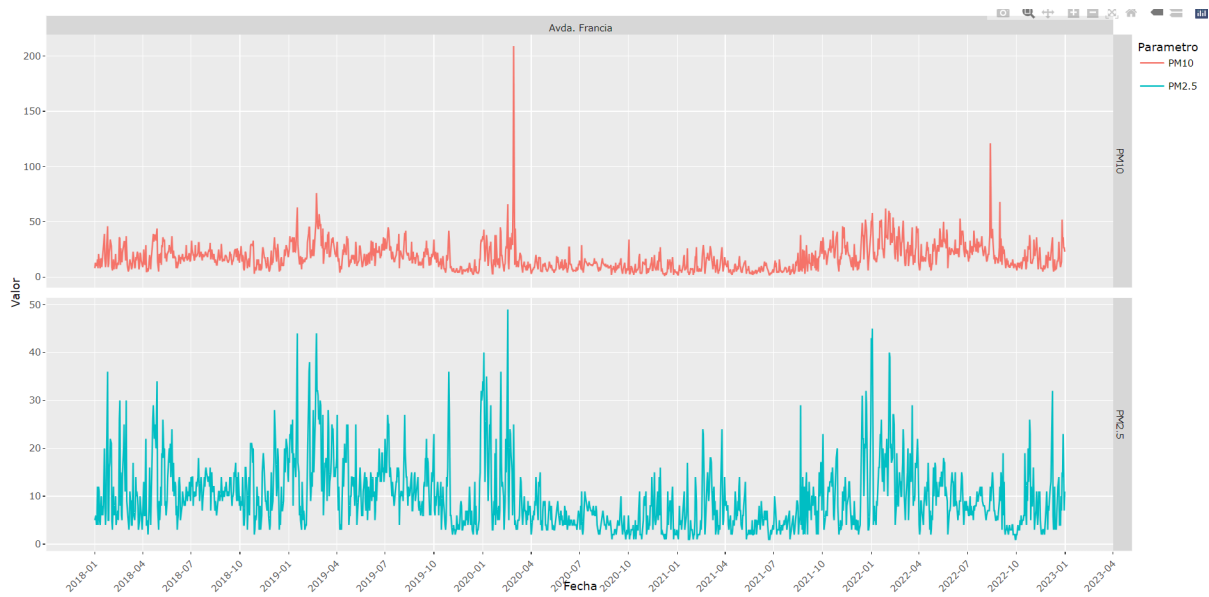
Hemos hecho un segundo mapa en el que se representa en tiempo real la calidad ambiental de las diferentes estaciones y zonas de Valencia, el estado del tráfico en las diferentes carreteras, además de las zonas verdes y la cantidad de personas afectadas por la mala contaminación de determinadas estaciones. Al pulsar sobre el marcador también podemos ver los valores de cada parámetro. Este mapa es muy útil para saber el estado de Valencia en cada momento, pudiendo saber donde el tráfico es denso y si hay zonas en las que la calidad ambiental ha descendido. En el caso de que haya fiestas que contaminan como las fallas se puede ver la influencia fácilmente.

#### 3.2 Utilización de tablas, gráficos u otros recursos visuales para facilitar la comprensión y análisis de los resultados, resaltando las tendencias o patrones observados.

En el siguiente mapa de calor podemos ver dos cosas. Lo primero es que en el periodo en el que duró la pandemia y hubo confinamiento los valores de la calidad ambiental en su mayoría fueron buenos, mientras que de normal la calidad ambiental suele ser razonablemente buena. Lo segundo es que los meses en los que la calidad ambiental disminuye considerablemente es en invierno, esto es debido al aumento en el uso de calefacción, lo que conlleva el aumento de SO<sub>2</sub>, NO<sub>x</sub> y PM entre otros contaminantes. También, otros factores que provocan un aumento de la contaminación son las inversiones térmicas, comunes en Valencia o el aumento del tráfico debido a las temporadas festivas y condiciones climáticas.



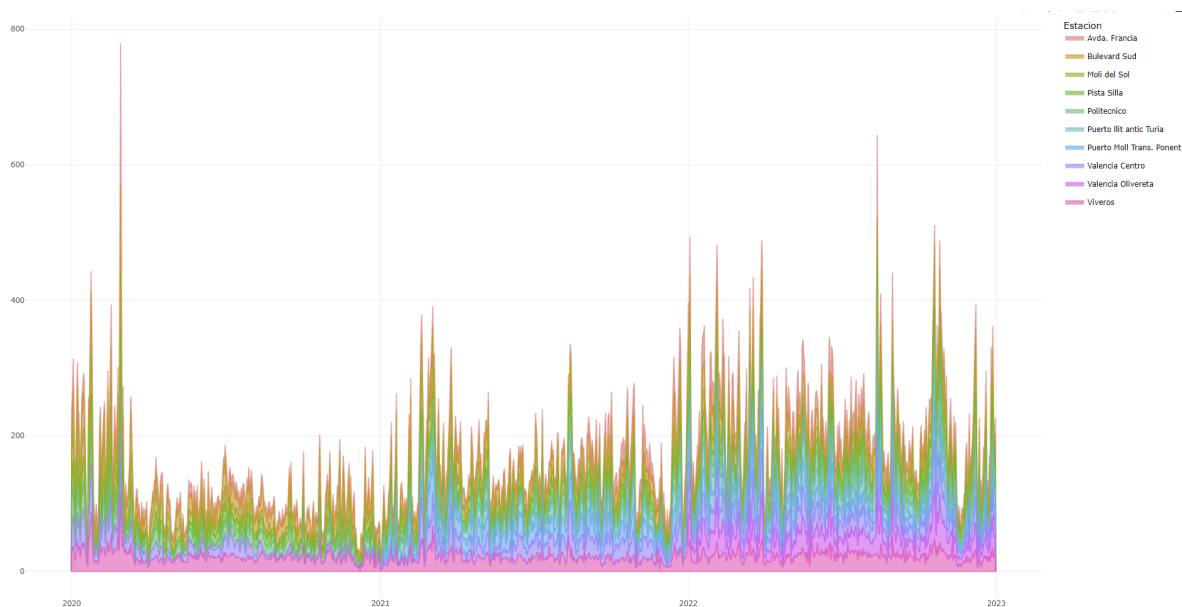
Tal y como se observa en la Figura 6 del mapa de calor, otra de las tendencias que podemos ver es la del PM10 y el PM2.5, durante el periodo del COVID-19 disminuyó drásticamente, obviando esta etapa, cuando más aumenta es en los meses de invierno, desde enero hasta marzo, teniendo picos periódicos entre febrero y marzo. Estos picos pueden deberse a las fallas y las mascletás o quizás a algún fenómeno meteorológico como las inversiones térmicas.



En los gráficos de la Figura 7 podemos observar que el valor del NO2 y el O3 depende de la temporada del año. Para verano el O3 aumenta y el NO2 disminuye mientras que en invierno el O3 disminuye y el NO2 aumenta. Esto se debe a que la intensidad de la radiación solar juega un papel muy importante en la formación del O3, mientras que los niveles de NO2 tienden a aumentar en invierno debido a las condiciones de inversión térmica, llevando a una acumulación de NO2.



En este gráfico interactivo, aunque está saturado nos sirve para ver cual es la estación que alcanza valores más altos y por tanto la que está más contaminada. En este caso para el PM10 y para la mayoría de parámetros la estación más afectada es Avda. Francia. Esto podría deberse a que hay diferentes centros comerciales y tiendas, por lo que es un sitio concurrido en el que hay mucho tráfico.



## 4. Discusión

### 4.1 Interpretación de los resultados en relación con los objetivos del estudio y la literatura existente.

Como uno de los principales objetivos del estudio era la concienciación, vamos a remarcar ciertas partes donde se concluye como los resultados obtenidos coinciden con la literatura existente en la materia.



En nuestros resultados, observamos que durante el periodo entre marzo del 2020 y principios del 2021 hubo una reducción significativa de los datos históricos de partículas PM10 y PM2.5. Esta conclusión adquiere mayor sentido cuando nos paramos a examinar cuáles son las fuentes de emisión de estas mismas partículas, más en concreto las PM2.5, ya que son precisamente estas las que pueden tener mayores impactos negativos en la salud respiratoria cuando se inhalan debido a que por su menor tamaño son capaces de penetrar de forma más profunda en nuestro sistema respiratorio.

que pueden ser inhaladas y penetrar así en el sistema respiratorio; las de menor tamaño (de 2,5 micras de diámetro, o inferior –PM2,5-) pueden incluso alcanzar los alveolos pulmonares, lo que les permite de este modo llevar sustancias nocivas a zonas muy sensibles y agravar patologías que pueden conducir incluso a una muerte prematura. De este modo, las partículas (en especial las PM2,5) pueden estar implicadas en el incremento de la mortalidad y de la morbilidad por causas respiratorias y cardiovasculares. Además de PM10 y PM2,5, las partículas ultrafinas (inferiores a 0,1

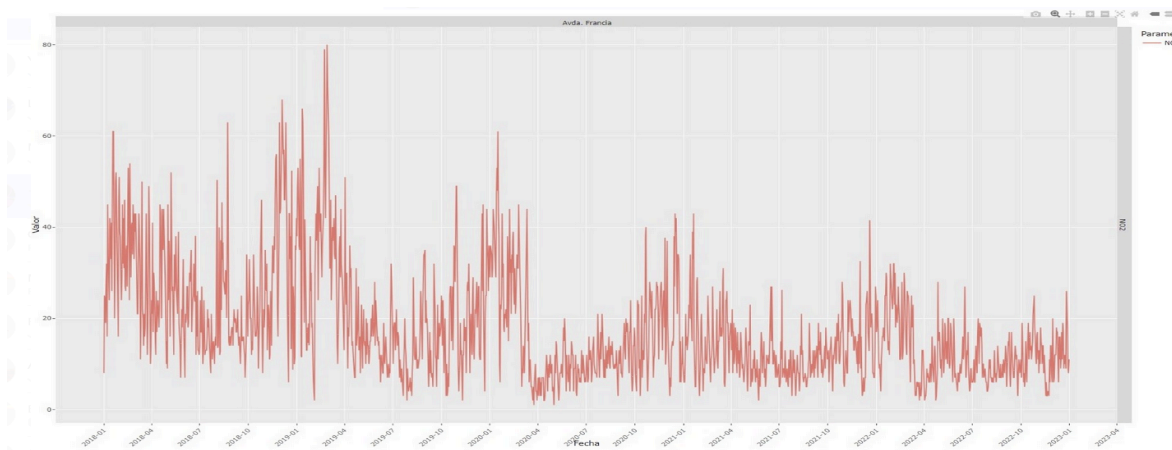
(Ministerio para la transición ecológica y el reto demográfico)

Tal y como se dice en la siguiente cita, las principales fuentes de emisión de PM2.5 serían la combustión en sectores no industriales seguida del tráfico rodado, ambas como resultado de la quema de combustibles fósiles como carbón, petróleo o licuado de petróleo.

Según el Inventario Nacional de Emisiones Contaminantes a la Atmósfera, el foco más influyente es la combustión en sectores no industriales, tanto para las PM10 como para las PM2,5, seguida por el tráfico rodado (en el caso de las PM10) o por la agricultura (en el de las PM2,5).

(Ministerio para la transición ecológica y el reto demográfico)

A su vez, el principal derivado de estas mismas actividades sería el NO2. Si seleccionamos NO2 en nuestro mapa interactivo veremos el siguiente gráfico:



Este mismo nos permitirá compararlo con un gráfico sacado de Investing.com del precio del Brent Oil, el cual es uno de los de referencia mundial para fijar el precio del petróleo, y en él observamos cómo ambos se correlacionan en ese periodo temporal, siendo precisamente esta bajada fruto del cese de la actividad industrial y vehicular de la zona, lo cual nos permitirá sacar conclusiones que detallaremos en su correspondiente apartado.



#### 4.2 Análisis de las implicaciones de los hallazgos.

Como añadido a nuestros objetivos principales, queremos resaltar también que las implicaciones de nuestros resultados pueden ser de mucha utilidad tanto a nivel práctico como científico si se pusiese el foco en ellas y les garantizamos una serie de recursos económicos.

Más en concreto, ponen el foco en la falta de zonas verdes e invitan a la creación de nuevas zonas, lo cual impactaría de forma muy significativa tanto en la calidad aérea como en la salud de las personas.

Un aspecto importante también sería que nuestros resultados nos indican que la calidad aérea de Valencia no es perfecta, y aunque en ningún caso son resultados alarmantes, ponen de manifiesto que existe una posibilidad de mejora.

#### 4.3 Discusión sobre las limitaciones del estudio y posibles áreas para futuros trabajos.

Es necesario resaltar, que nuestro estudio se centra en los datos que hemos mencionado anteriormente, pero estos mismos datos podría ser mucho mayores en número y calidad, por lo que sería interesante una ampliación del contenido con fuentes de datos provenientes de otro tipo de mediciones que ampliasen el abanico de posibilidades que se le ofrece al usuario final, ayudando así a sensibilizar en otras áreas relacionadas.

Un área en la que convendría enfocarse podría ser en evaluar los efectos de las políticas ambientales que se implementan, y así transmitir una mayor sensación de transparencia y utilidad del uso de los recursos públicos empleados en este área.

#### 4.4 Dificultades encontradas.

Algunas de las dificultades que fuimos encontrando fueron por ejemplo la aparición de muchos valores NA en los datos, así como ciertas variables de las que nos deshicimos que no tenían relevancia para nuestro análisis.

También nos aparecieron algunos valores imputados predichos en negativo que no tenían sentido fruto de la distribución y que conseguimos solventar cambiando la misma.

Otro aspecto fue el cómo relacionar la variable temporal con el resto de datos y la localización espacial de cada estación.

### 5. Conclusiones:

#### 5.1 Recapitulación de los hallazgos principales

Los principales hallazgos que nos hemos encontrado han sido:

Ciertas tendencias anteriormente detalladas como, que la **bajada de PM10 y PM2.5** en el periodo mencionado de 2020 estuvo relacionada con el confinamiento y se vió beneficiada del cese de actividad del sector industrial y la ausencia de tráfico vehicular, que observamos que **el valor del NO2 y el O3 depende de la temporada del año** o que **la calidad ambiental disminuye en invierno.**

Que la **emisión de partículas de NO2** está **relacionada con** las actividades de **quema de combustibles fósiles**, viendo como guarda cierta correlación con el precio del petróleo y siendo esto un reflejo de su demanda.

Que por lo visto en el apartado de QGIS, **cuanto más nos alejamos de la zona del puerto peor es la calidad del aire**, y esto bien podría ser por el aumento del flujo de vehículos o por la acumulación de centros industriales.

Sumado a esto, vemos que dentro de estas mismas zonas mencionadas, **aquellas con menor número de zonas verdes son las que más se ven afectadas** en cuanto a niveles de NO2, siendo este el principal hallazgo y en el cual nos gustaría que se pusiera el foco.



## 5.2 Respuestas a los objetivos planteados en la introducción.

En la introducción planteamos 3 objetivos principalmente y estos mismos así como sus respuestas se detallan a continuación:

- 1.- Se ha conseguido visualizar de una forma accesible y eficaz la información recogida por las estaciones, apoyándose en los diferentes elementos interactivos(calidad aérea, estado del tráfico...) de nuestros mapa, así como los labels existentes.
- 2.-Se consigue también nuestra otra prioridad, que es la sensibilización ambiental, la cual se espera que ayude a reducir los niveles de los elementos más perjudiciales a través de la concienciación de los ciudadanos mediante el uso del mapa interactivo.
- 3.-Se informa de medidas adecuadas para disminuir la polución, las cuales serán especificadas en el siguiente apartado.

## 5.3 Implicaciones prácticas y científicas de los resultados:

Lo anteriormente explicado en los hallazgos principales implica varias cosas:

Una es que las zonas céntricas de la ciudad y aquellas masivamente pobladas deberían poner en marcha medidas para mejorar sus niveles de NO<sub>2</sub>, esto como hemos dicho anteriormente se relaciona con el tránsito rodado por lo que una posible actuación podría ser la de fomentar la compra de vehículos eléctricos así como el acondicionamiento de la ciudad a ellos con la instalación de puntos de recarga.

Otra es que la falta de zonas verdes en ciertas áreas como L'olivereta, Patraix o Ciutat Vella empeora la calidad aérea y esto implica que también deberán actuar en consecuencia. Si desean remediarlo, una posible medida podría ser la conversión de áreas abandonadas y degradadas así como solares vacíos en zonas verdes.

Para acabar, queremos resaltar que mediante el uso de nuestro mapa interactivo el usuario medio podría llegar a otras implicaciones, combinando las diferentes capas para llegar a concluir por ejemplo que en Patraix hay una combinación de casuísticas, las 2 anteriores, ya que tiene pocas zonas verdes y es una zona muy transitada, y además dependiendo de la estación del año podría visualizarse si mejora o empeora todavía más su calidad aerea.

## 6. Referencias

1. [https://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
2. <https://www.redalyc.org/journal/104/10463384004/html/>
3. <https://mediambient.gva.es/es/web/calidad-ambiental/calidad-del-aire>
4. <https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/ica.html>
5. <https://www.ecologistasenaccion.org/17842/que-son-las-pm25-y-como-afectan-a-nuestra-salud/>
6. <https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/transicion-ecologica/Paginas/2023/130723-calidad-aire-2022-muestra-mejoria.aspx#:~:text=Para%20el%20PM2%2C5%20el%20totalidad%20de%20las%20zonas.>
7. <https://openroom.fundacionrepsol.com/es/contenidos/trafico-rodado-en-la-calidad-del-aire-urbano-en-europa/#:~:text=El%20tr%C3%A1fico%20rodado%2C%20omnipresente%20en,mejorar%20la%20calidad%20del%20aire.>
8. <https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/salud/particulas.html>
9. <https://www.investing.com/commodities/brent-oil>