



CALIDAD DEL AIRE EN MADRID

Capstone Bootcamp en Data
Analytics

Ciro Annunziata

Marisa Batalla

Mayelín Dominguez

Borja Muñoz

Roser Ponce



TABLA DE CONTENIDOS

- Introducción: motivación, objetivos y metodología
- Desarrollo: datasets, limpieza y exploración de los datos
- Pruebas y resultados: cómo llegamos al modelo
- Conclusiones: qué aprendimos
- Sigüientes pasos

Calidad del aire

Principales contaminantes:

- Materia particulada (PM10 y PM2.5)
- Dióxido de azufre (SO2)
- Dióxido de nitrógeno (NO2))
- Ozono (O3)

Factores que influyen:

- **Industria**
- **Transporte**
- **Sector energético**
- **Residuos**
- **Topografía**
- **Condiciones metereológicas**

Estado de calidad del aire	SO ₂	NO ₂	O ₃	PM10	PM2,5
MUY BUENO	0-100 µg/m ³	0-40 µg/m ³	0-80 µg/m ³	0-20 µg/m ³	0-10 µg/m ³
BUENO	101-200 µg/m ³	41-100 µg/m ³	81-120 µg/m ³	21-35 µg/m ³	11-20 µg/m ³
REGULAR	201-350 µg/m ³	101-200 µg/m ³	121-180 µg/m ³	36-50 µg/m ³	21-25 µg/m ³
MALO	351-500 µg/m ³	201-400 µg/m ³	181-240µg/m ³	51-100 µg/m ³	26-50 µg/m ³
MUY MALO	501-1250 µg/m ³	401-1000 µg/m ³	241-600 µg/m ³	110-1200 µg/m ³	51-800 µg/m ³

Motivación

Salud pública y medio ambiente:

- Bienestar y calidad de vida
- Mitigación del cambio climático

Cumplimiento normativo :

- Crear políticas para reducir contaminantes
- Tomar acciones para cumplir con estándares internacionales

MEDIO AMBIENTE

Madrid es la ciudad europea con más muertes por aire contaminado de los coches

En el ranking de las ciudades europeas con mayor carga de mortalidad atribuible al NO2 hay tres de España: Madrid, que lidera la tabla; Barcelona, en la sexta posición, y Mollet del Vallés, en la séptima

LA JUSTICIA EUROPEA CONDENA A ESPAÑA POR LA CONTAMINACIÓN DE MADRID Y BARCELONA

El Tribunal de Justicia de la Unión Europea ha dictaminado que España no veló por la calidad del aire en estas dos ciudades entre 2010 y 2018, superándose de manera sistemática y continuada los límites establecidos.

Objetivos

General:

Desarrollar un modelo predictivo de calidad del aire para la ciudad de Madrid, utilizando datos meteorológicos y de tráfico vehicular.

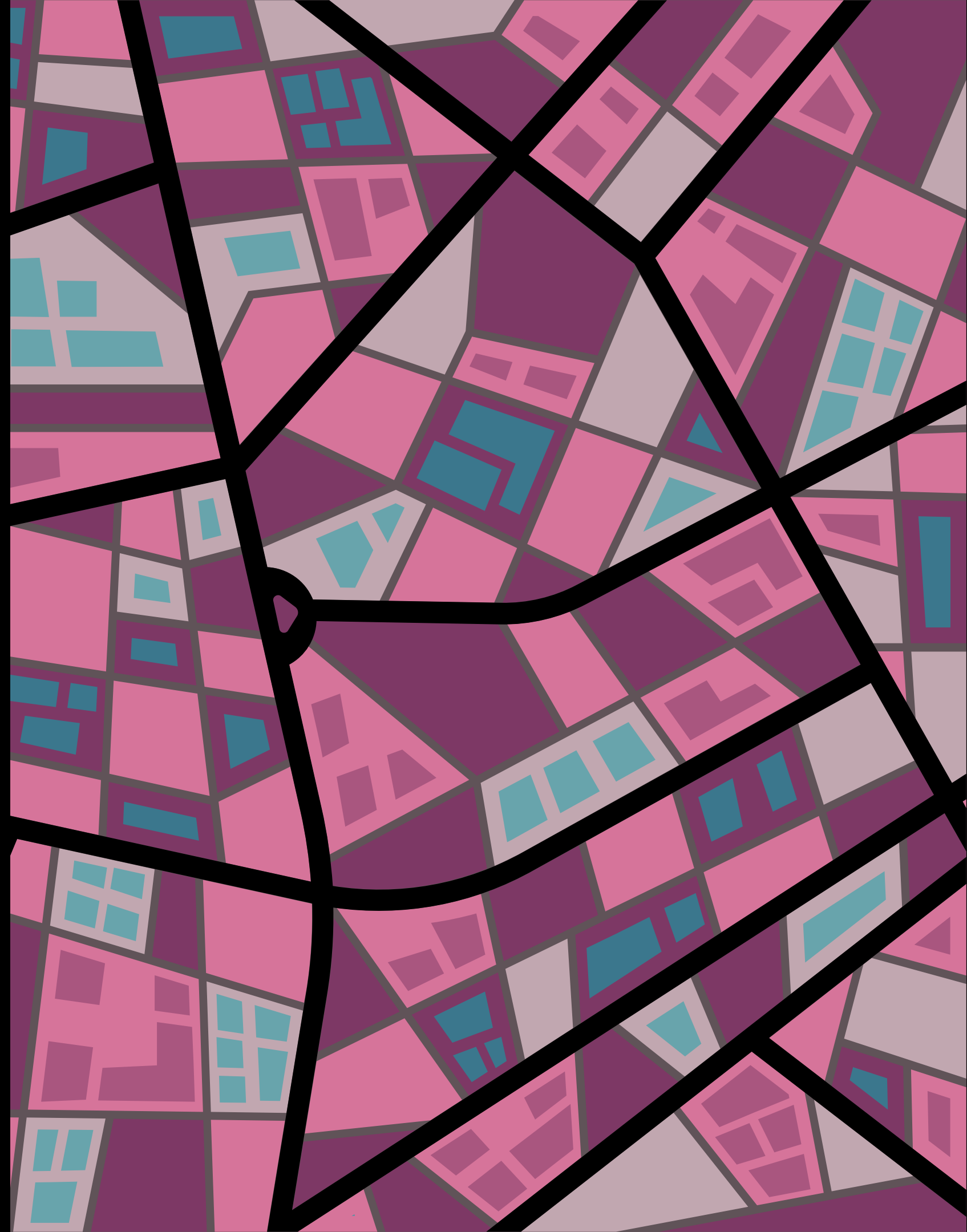
Específicos:

- Recopilar datos históricos de la calidad del aire, el clima y el tráfico.
- Procesar y visualizar los datos para comprender las relaciones entre las variables.
- Implementar y mejorar algoritmos de machine learning para crear un modelo predictivo de contaminación basado en los datos.

Metodología

Análisis cuantitativo

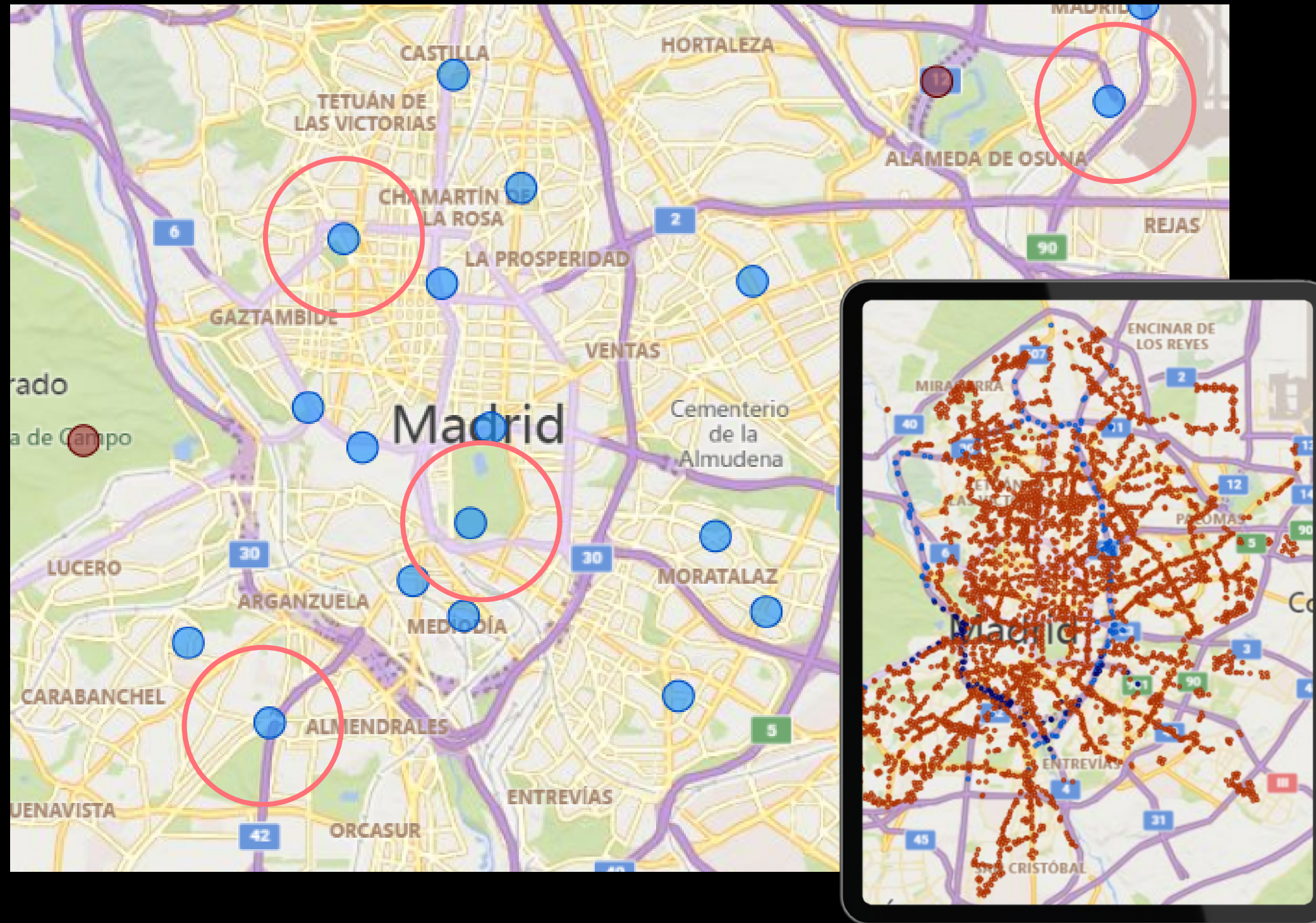
- **Descriptivo:** Resumir datos para proporcionar una descripción general clara y concisa.
- **Exploratorio (EDA):** Profundizar en los datos para descubrir patrones, relaciones y tendencias.
- **Predictivo:** Pronosticar resultados o eventos futuros basándonos en datos históricos.



Datasets

Fuentes:

Agencia Estatal de Meteorología
(AEMET)
Ayuntamiento de Madrid



Zonas analizadas:

Retiro
Cuatro Caminos
Barajas
Plaza Elíptica

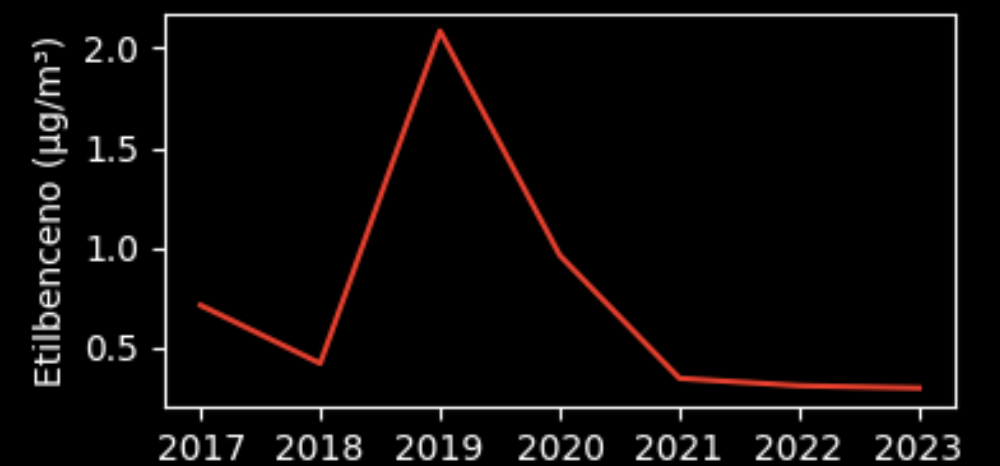
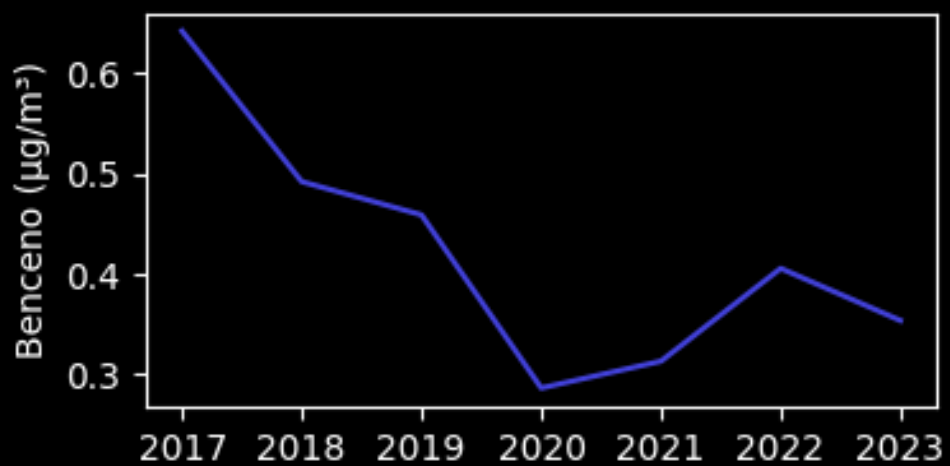
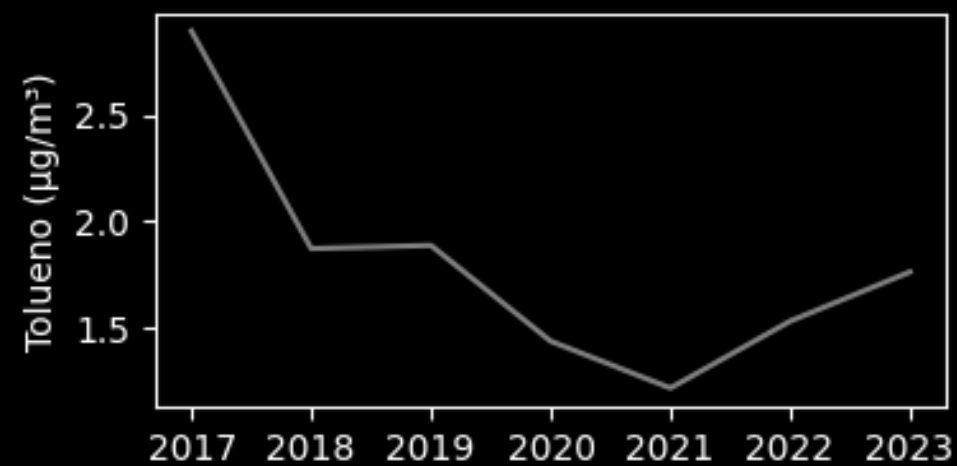
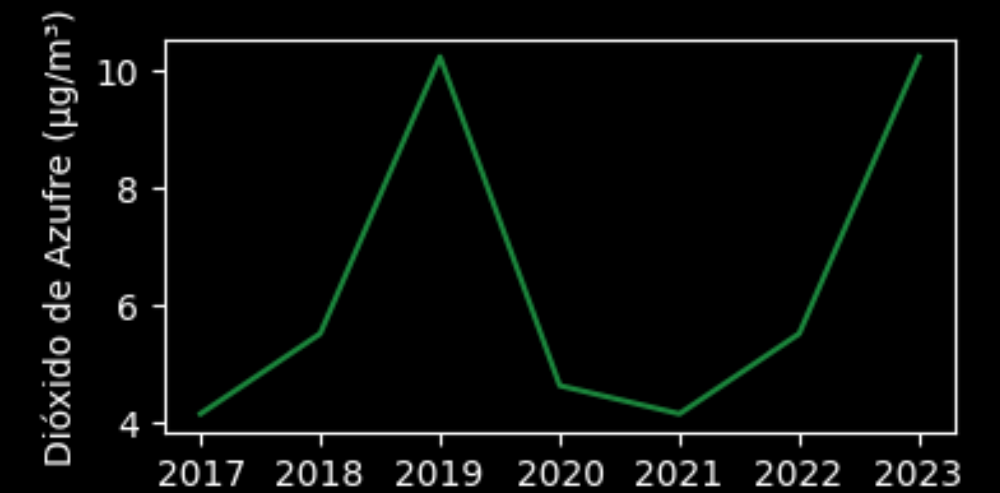
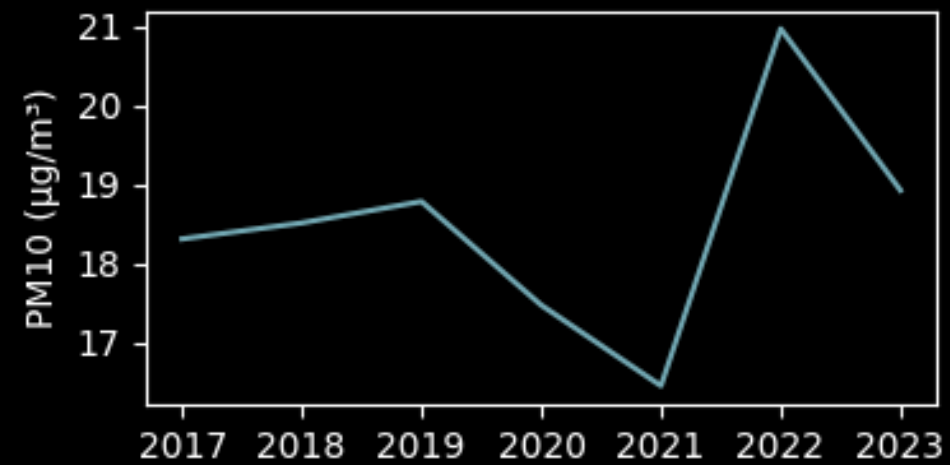
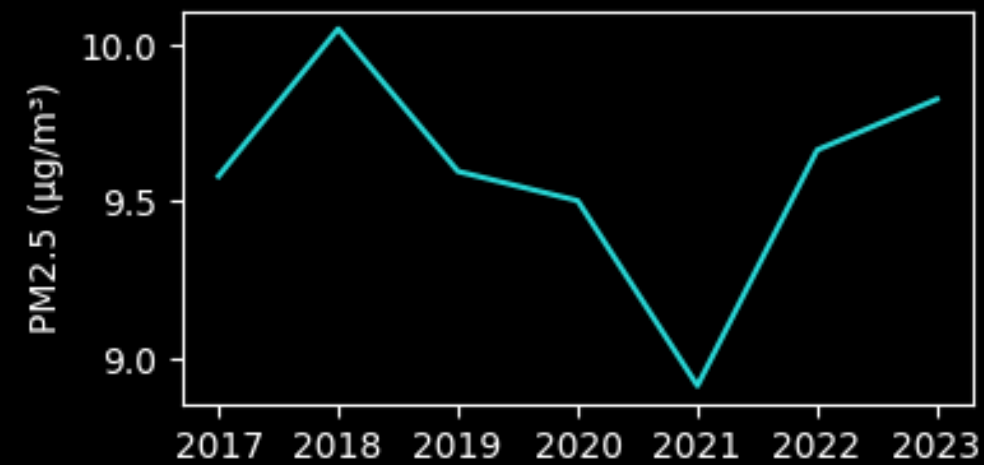
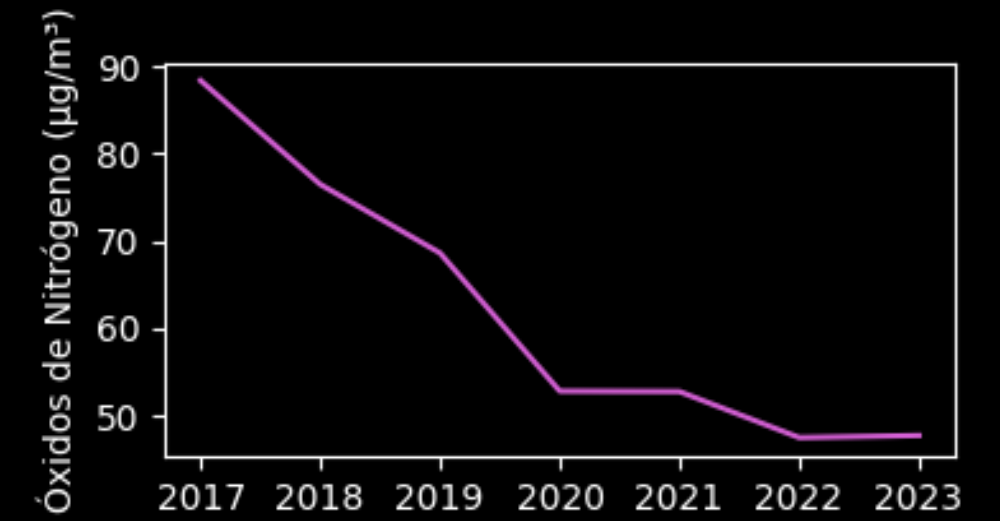
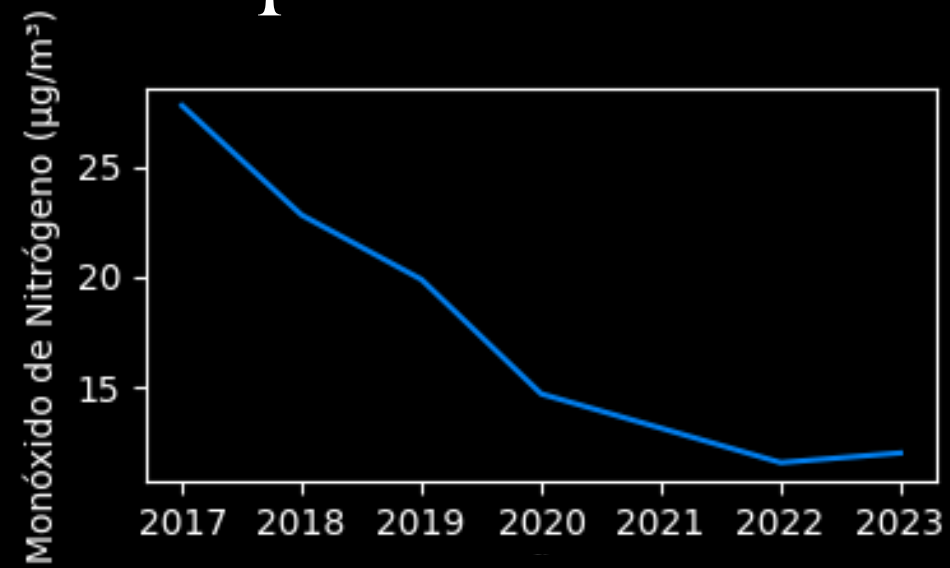
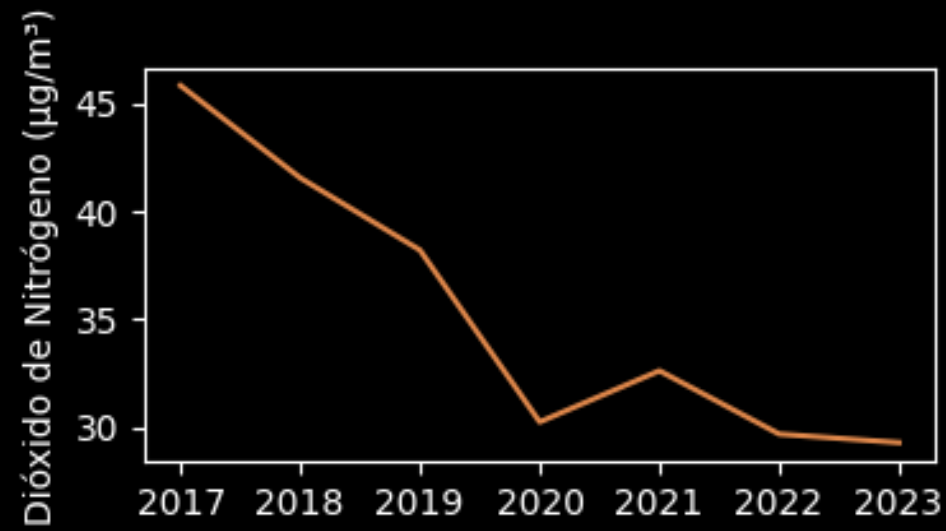
Calidad del dato

Master sheet:

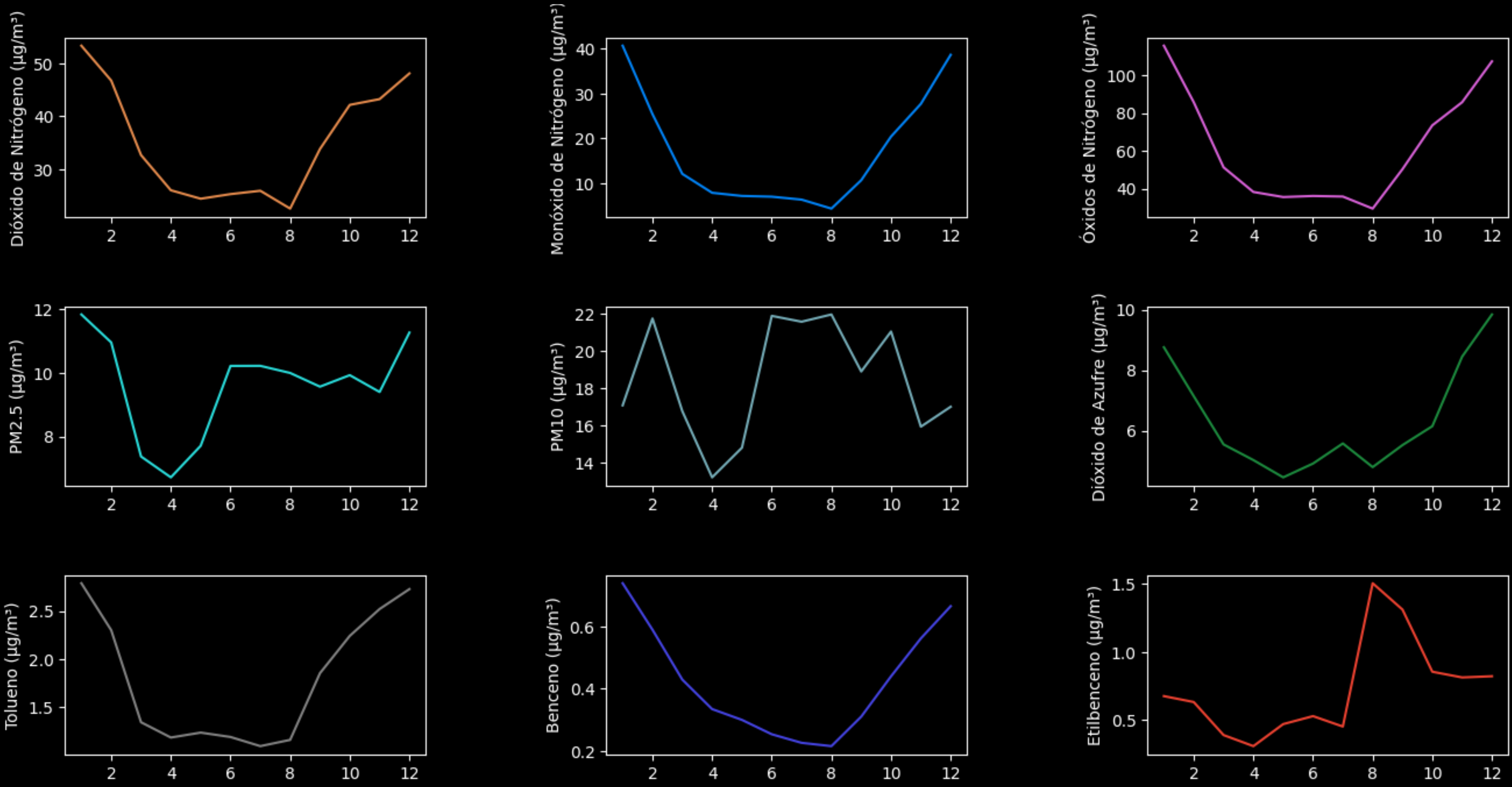
- 74 columnas
- 2555 filas (una por cada día desde el 1 de enero 2017 a 31 de diciembre 2023)

CUATRO CAMINOS	líneas vacías	media	mediana	min	max	valores únicos
precipitaciones	0	1	0	0	67,7	156
temperatura media	0	16	15,4	-3,2	33,4	314
temperatura mínima	0	9,9	9,3	-11	26,2	316
temperatura máxima	0	22	21,5	0,9	42,2	357
humedad media	0	52,2	50	16	99	207
humedad mínima	0	36	32	10	98	184
humedad máxima	0	76,2	80	32	100	170
dióxido de nitrógeno	0	35,3	31	3	147	113
monóxido de nitrógeno	0	17,4	9	1	254	131
óxidos de nitrógeno	0	62	44	5	519	251
PM25	0	9,6	9	1	48	44
PM10	0	18,5	16	1	271	81
dióxido de azufre	0	6,3	5	1	24	22
tolueno	0	1,8	1,3	0,1	19	105
benceno	0	0,4	0,3	0,1	3	29
etilbenceno	0	0,7	0,4	0,1	16,6	87
intensidad del tráfico	0	409,5	432	29,8	646,2	2553
ocupación del tráfico	0	5	5,1	0,7	9,9	2553
carga del tráfico	0	18,3	19,4	2	29	2553

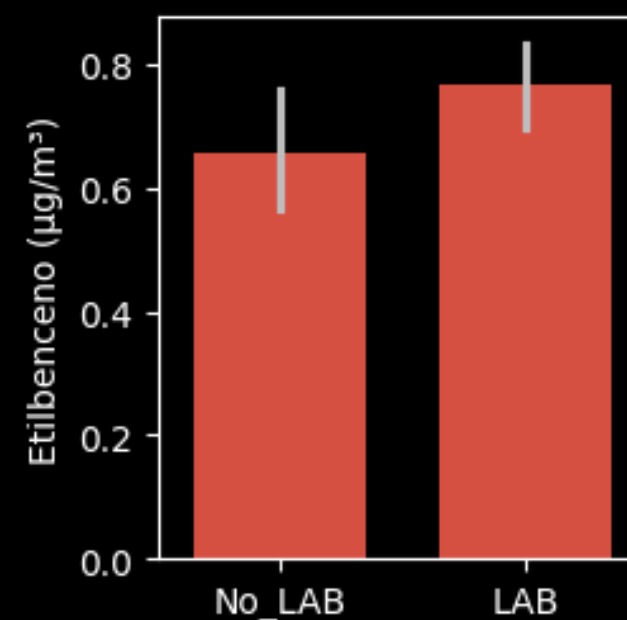
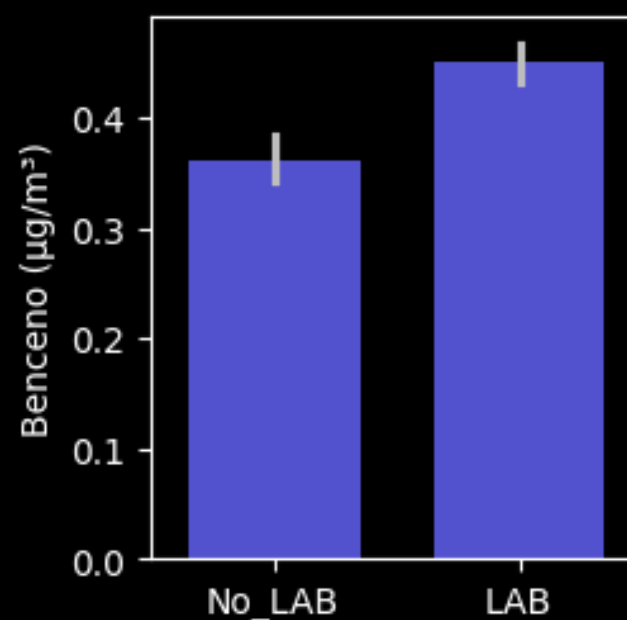
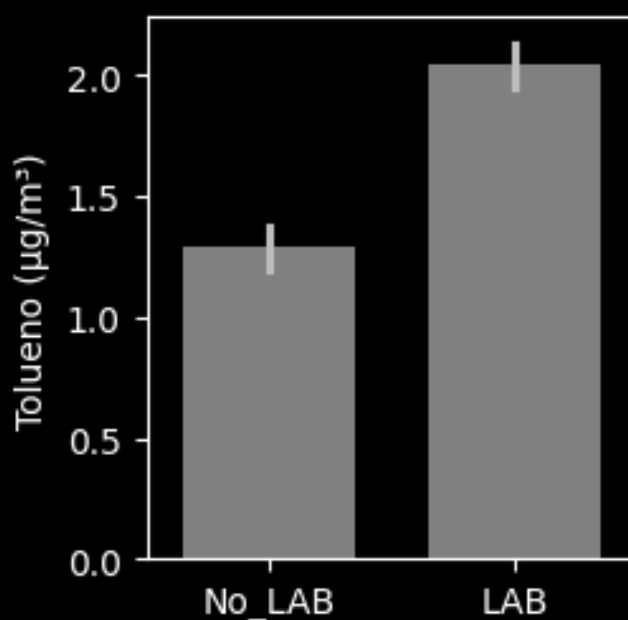
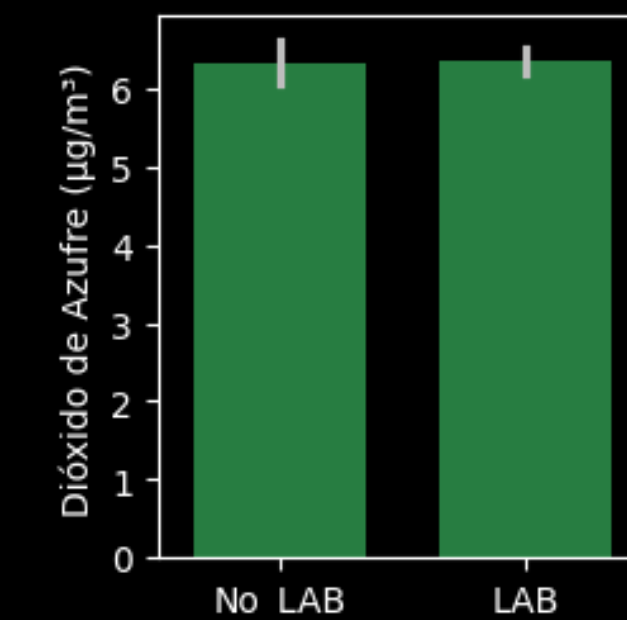
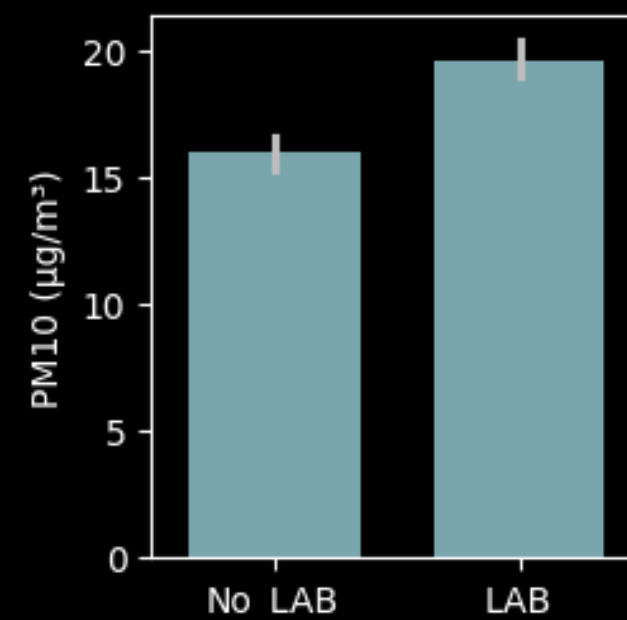
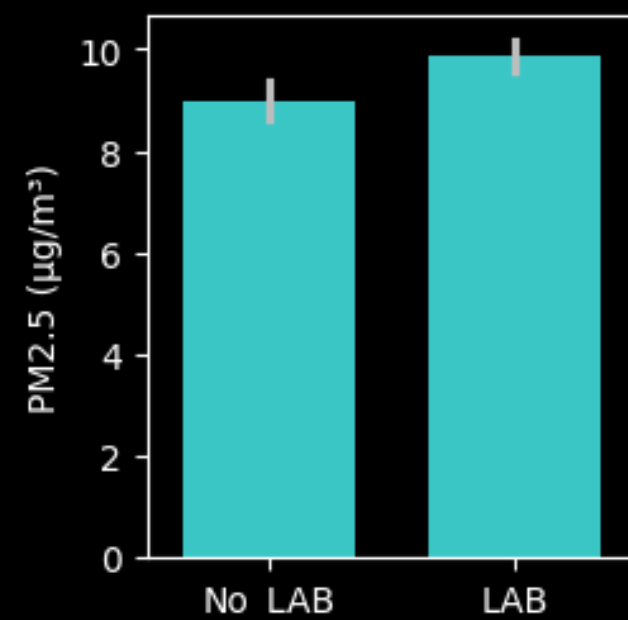
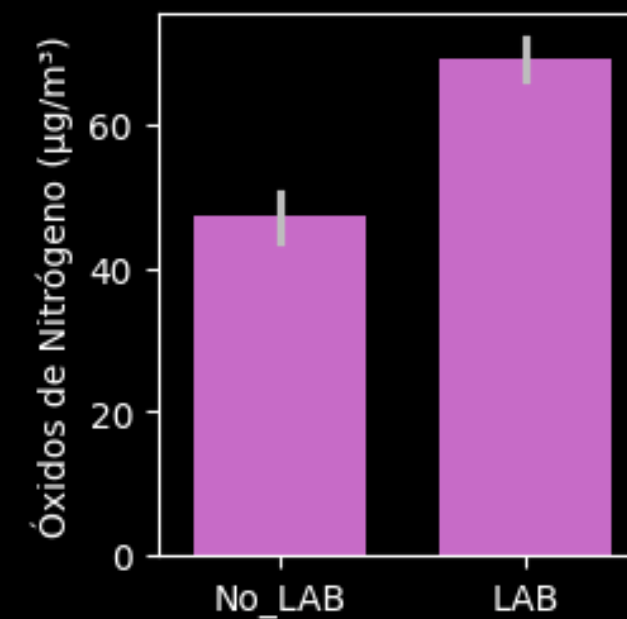
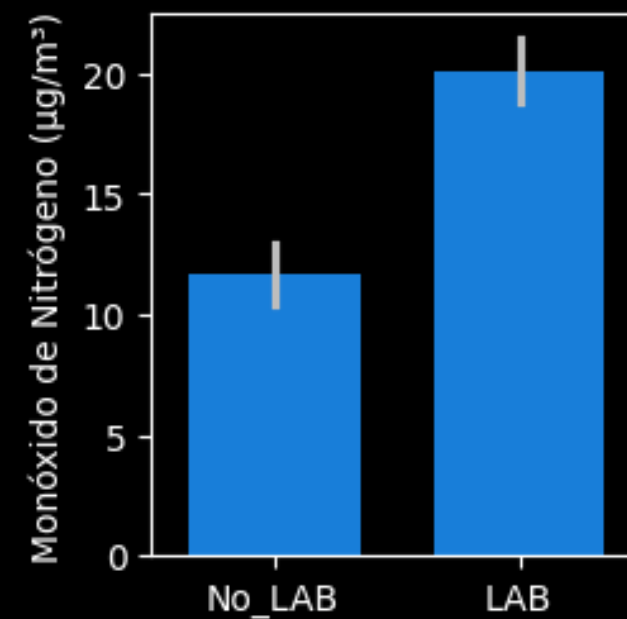
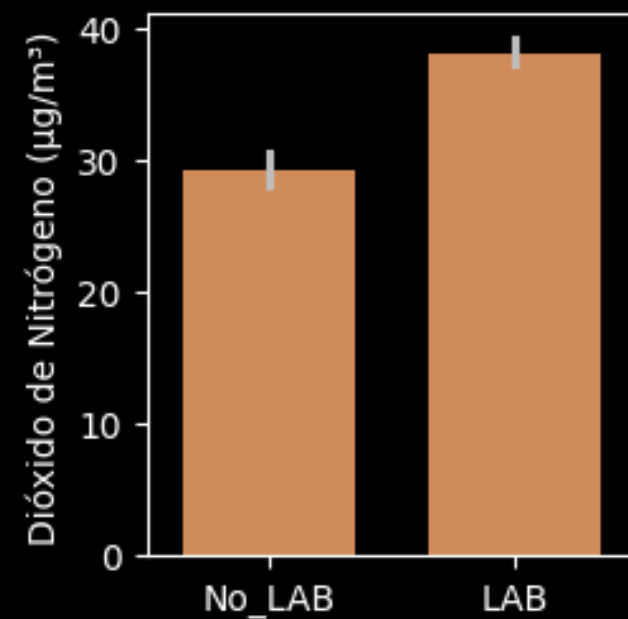
Distribuciones de las concentraciones medias de contaminantes por año



Distribuciones de las concentraciones medias de contaminantes por mes



Distribuciones de las concentraciones medias de contaminantes por tipo de día



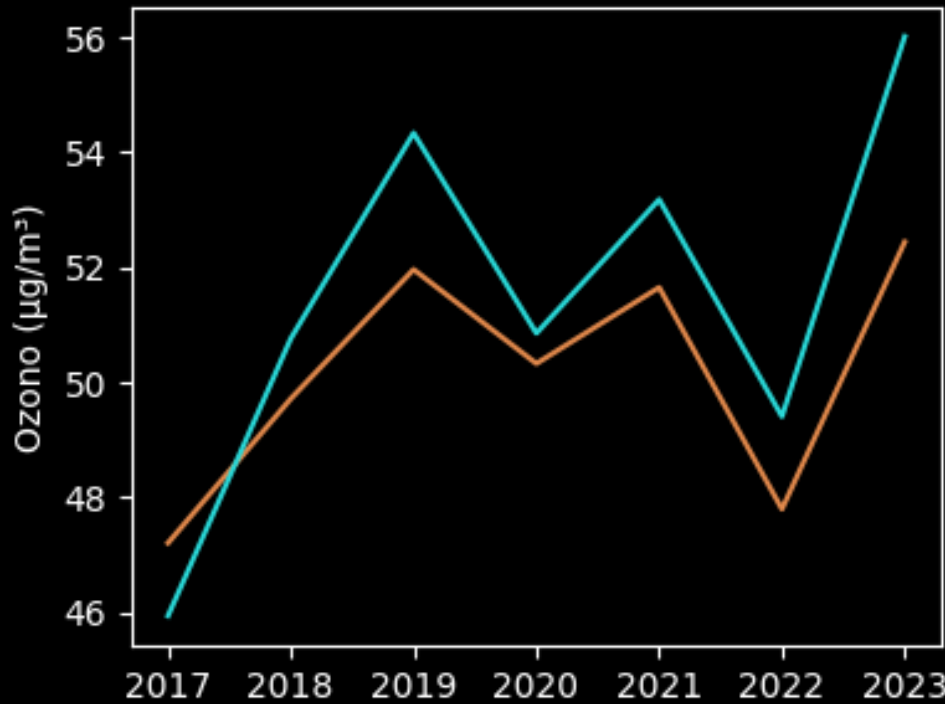
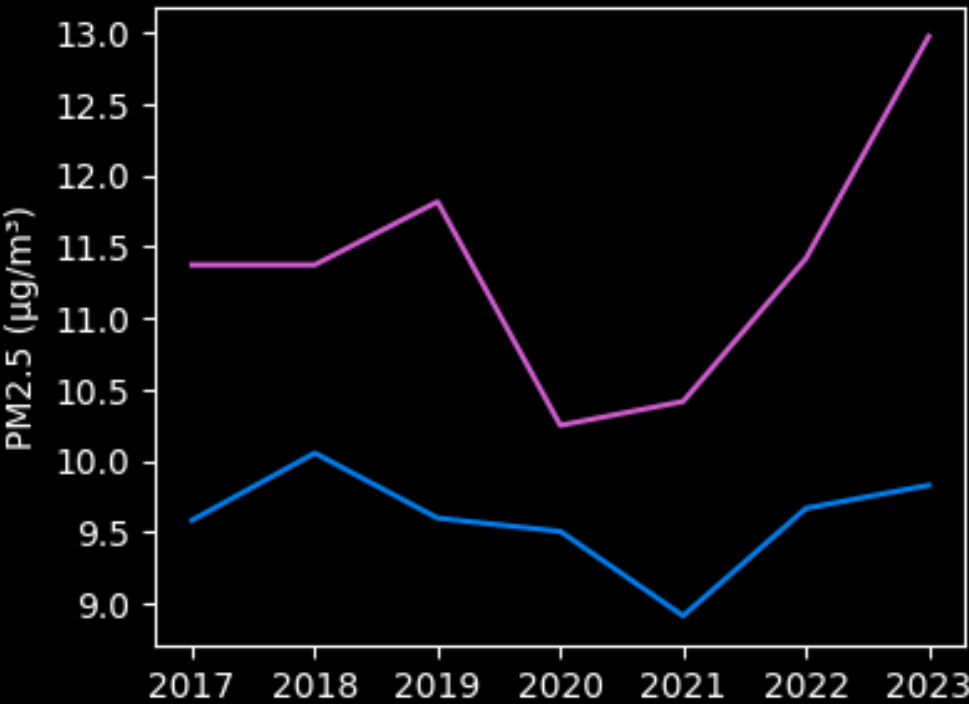
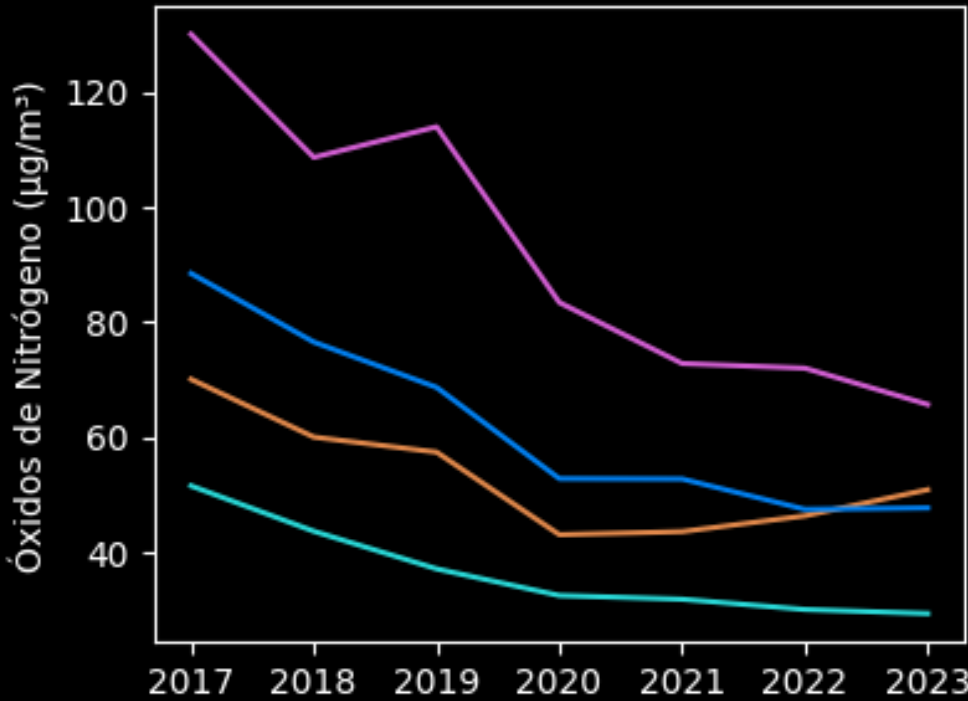
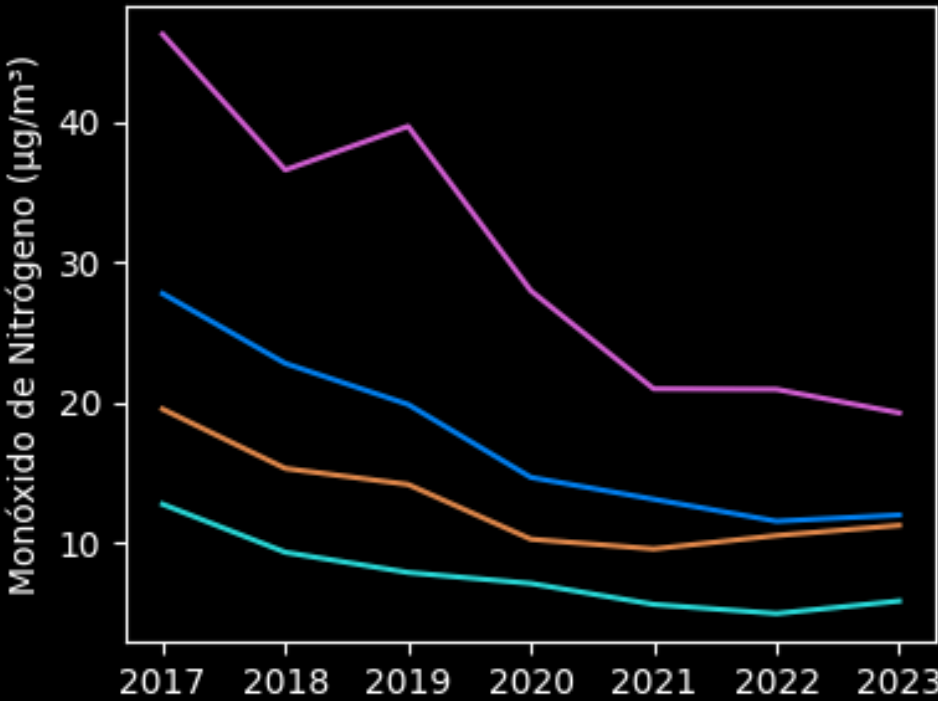
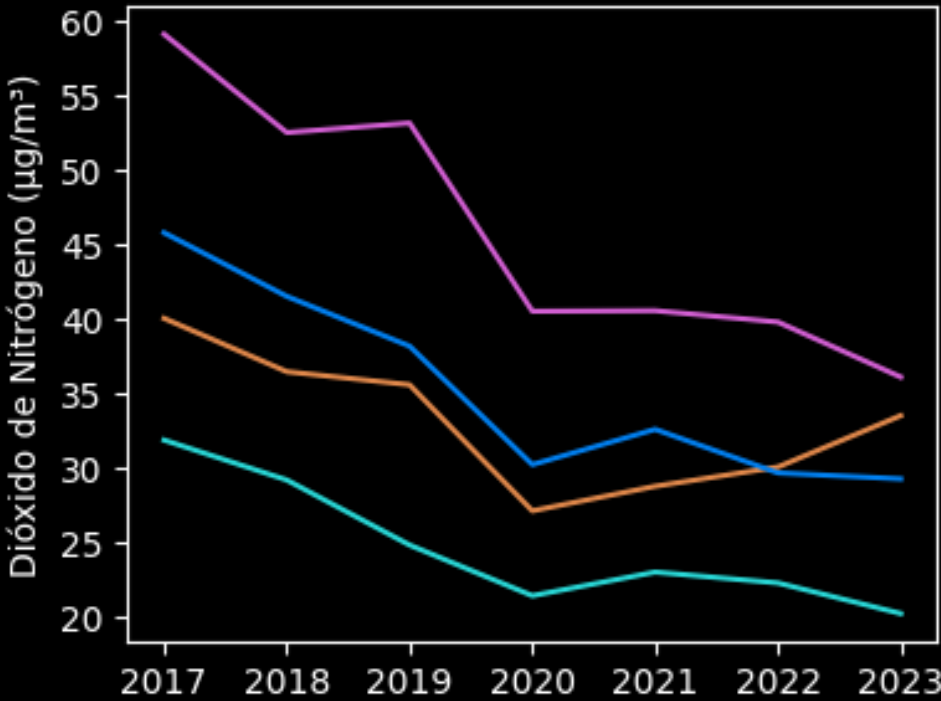
Concentraciones medias de contaminantes por año

Barajas

Cuatro Caminos

Plaza Elíptica

Retiro



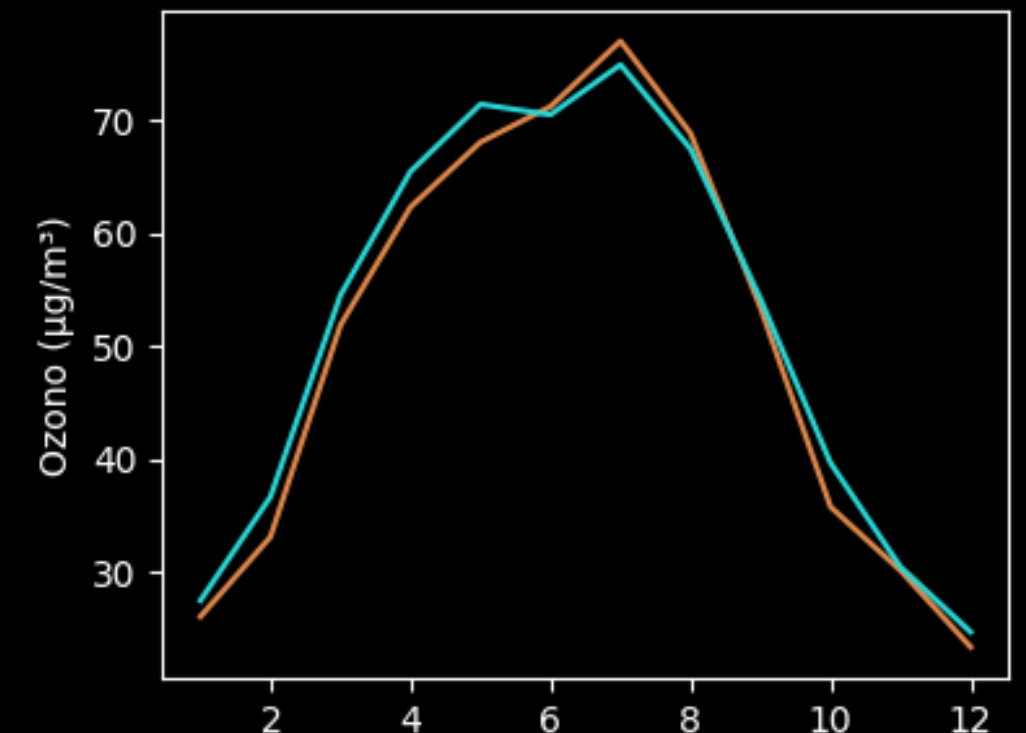
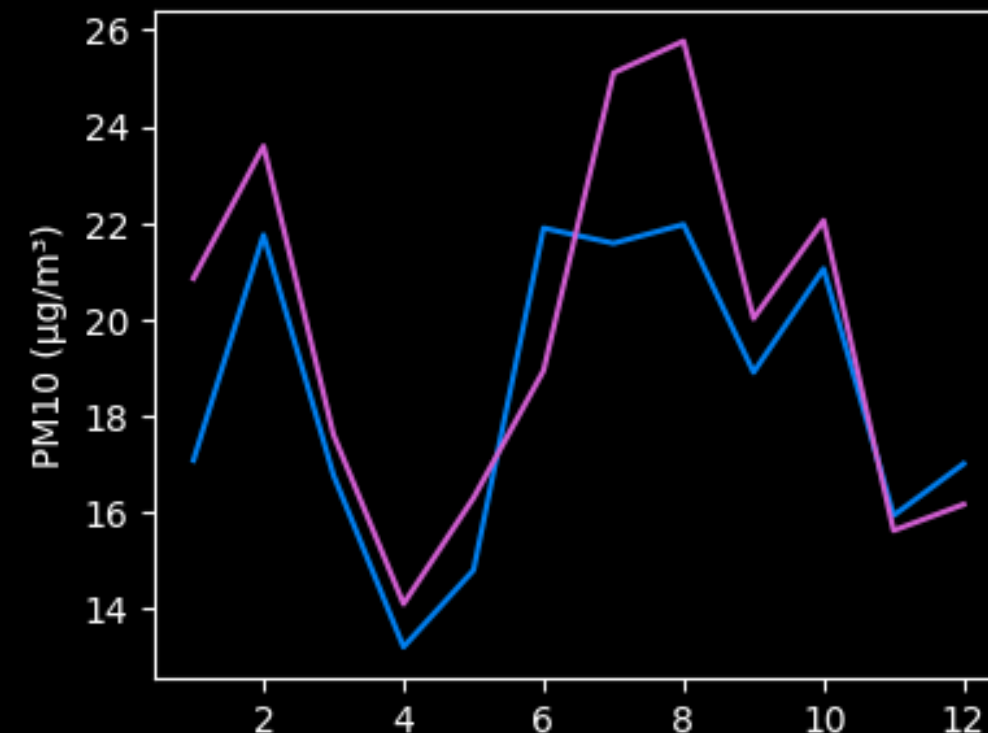
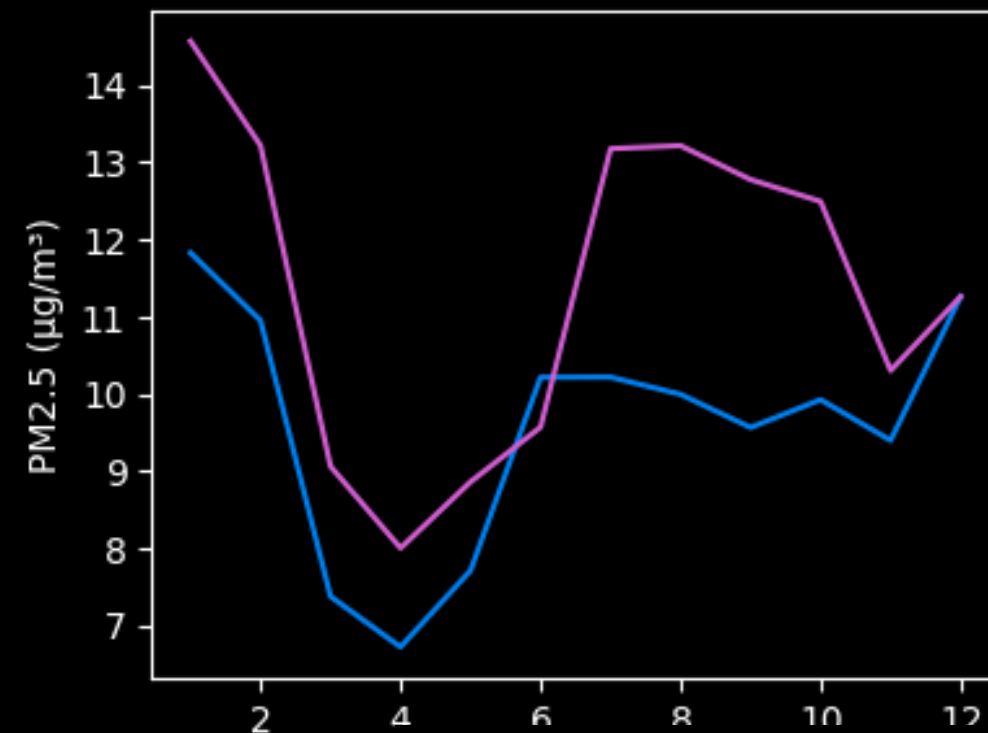
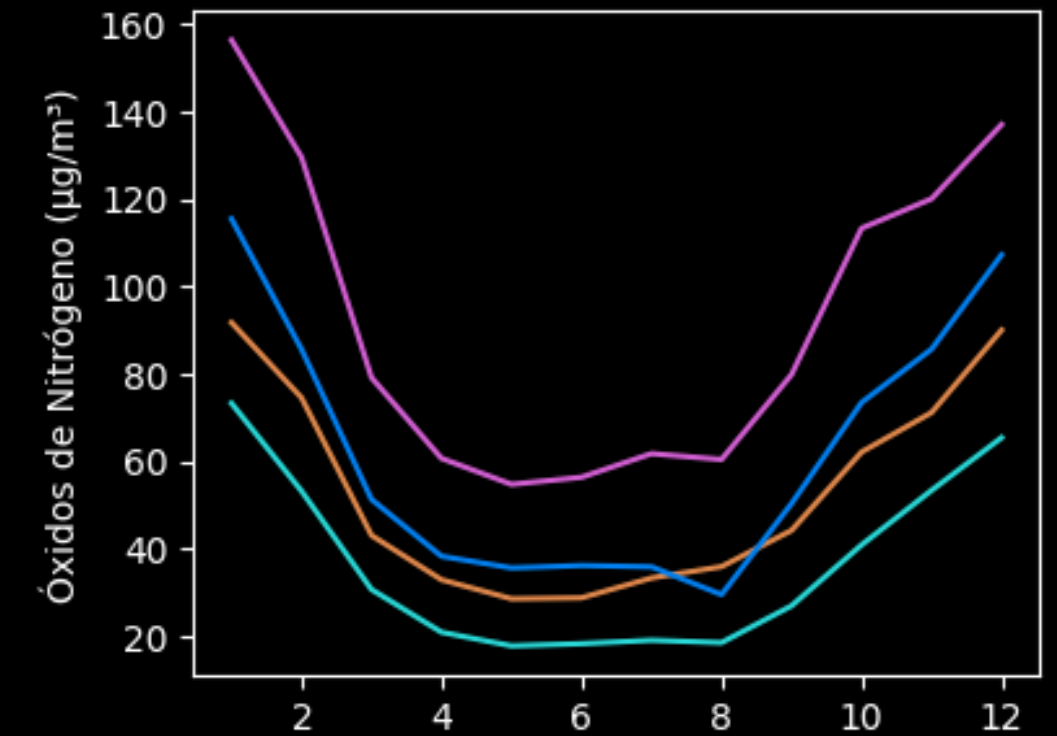
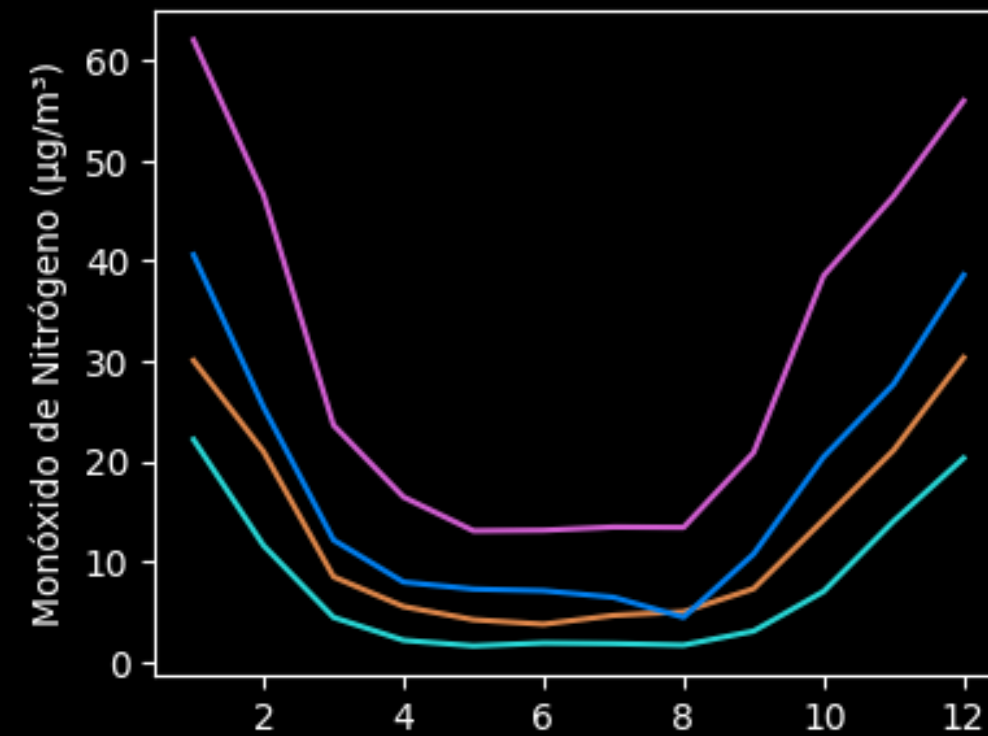
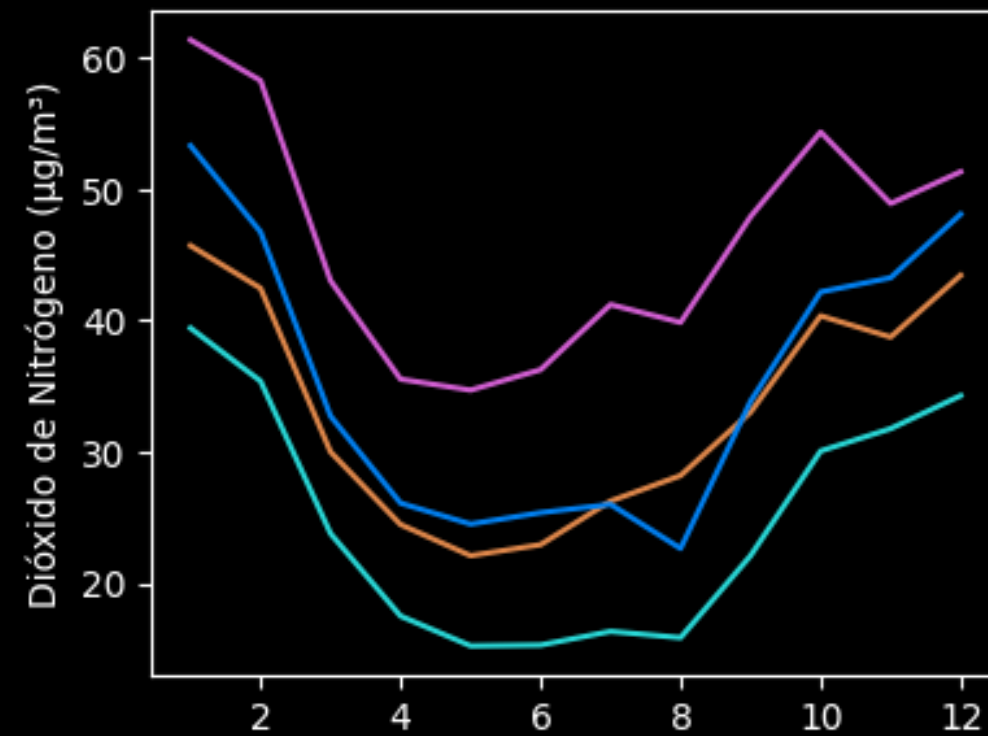
Concentraciones medias de contaminantes por mes

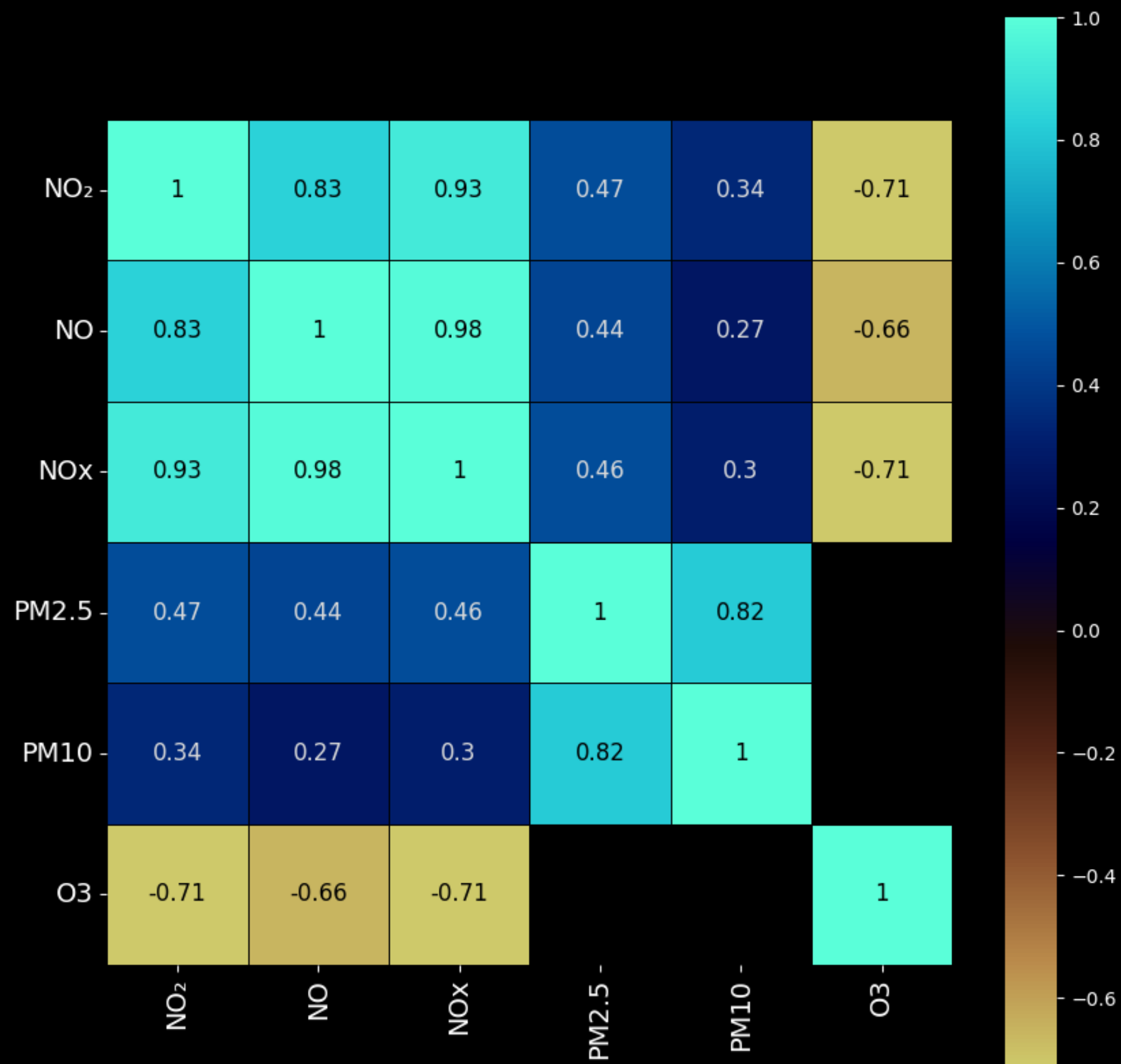
Barajas

Cuatro Caminos

Plaza Elíptica

Retiro





Correlaciones
entre
contaminantes

Modelado y predicción: NO₂

Algoritmos utilizados:

Logistic/linear regression

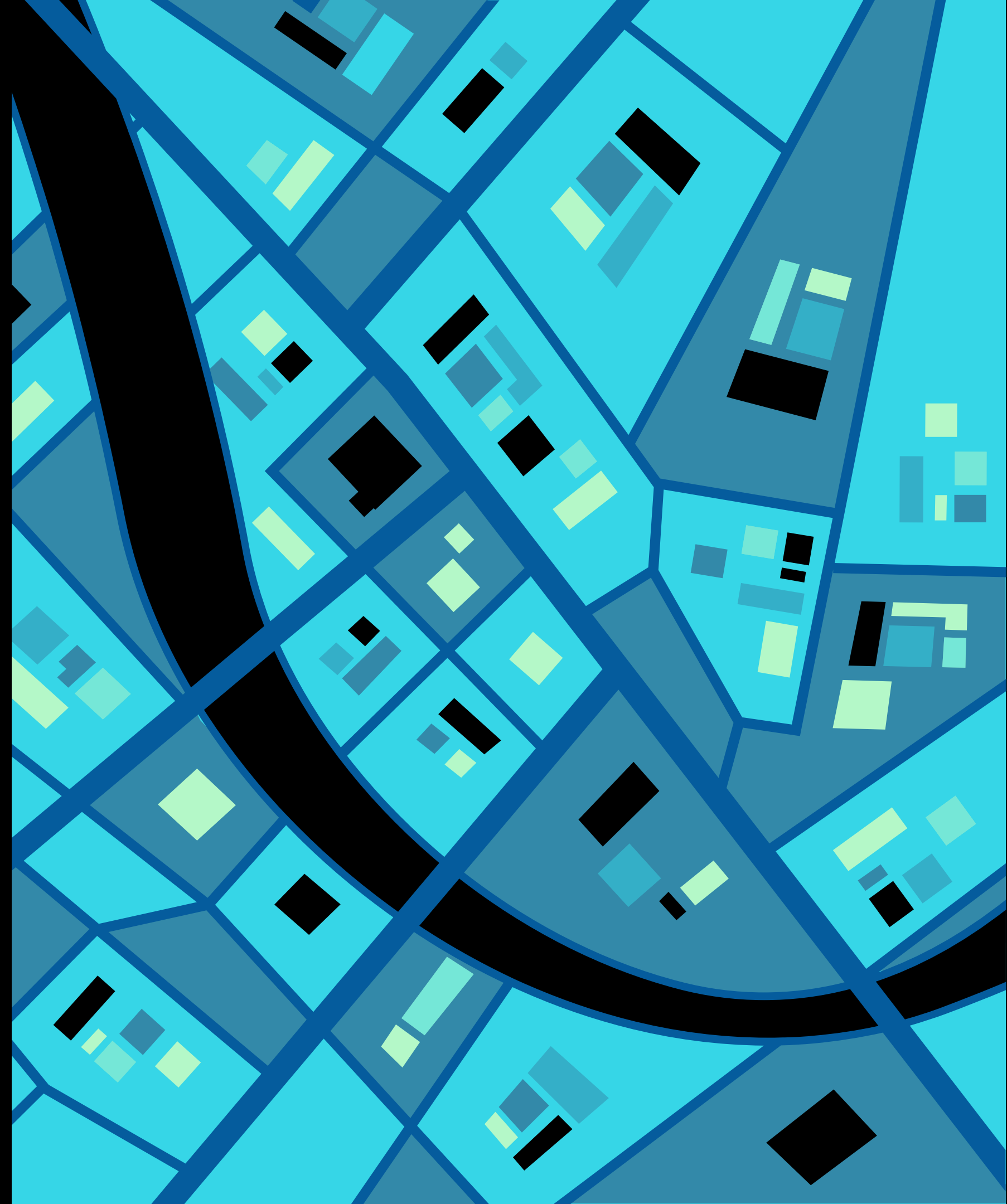
Decision trees

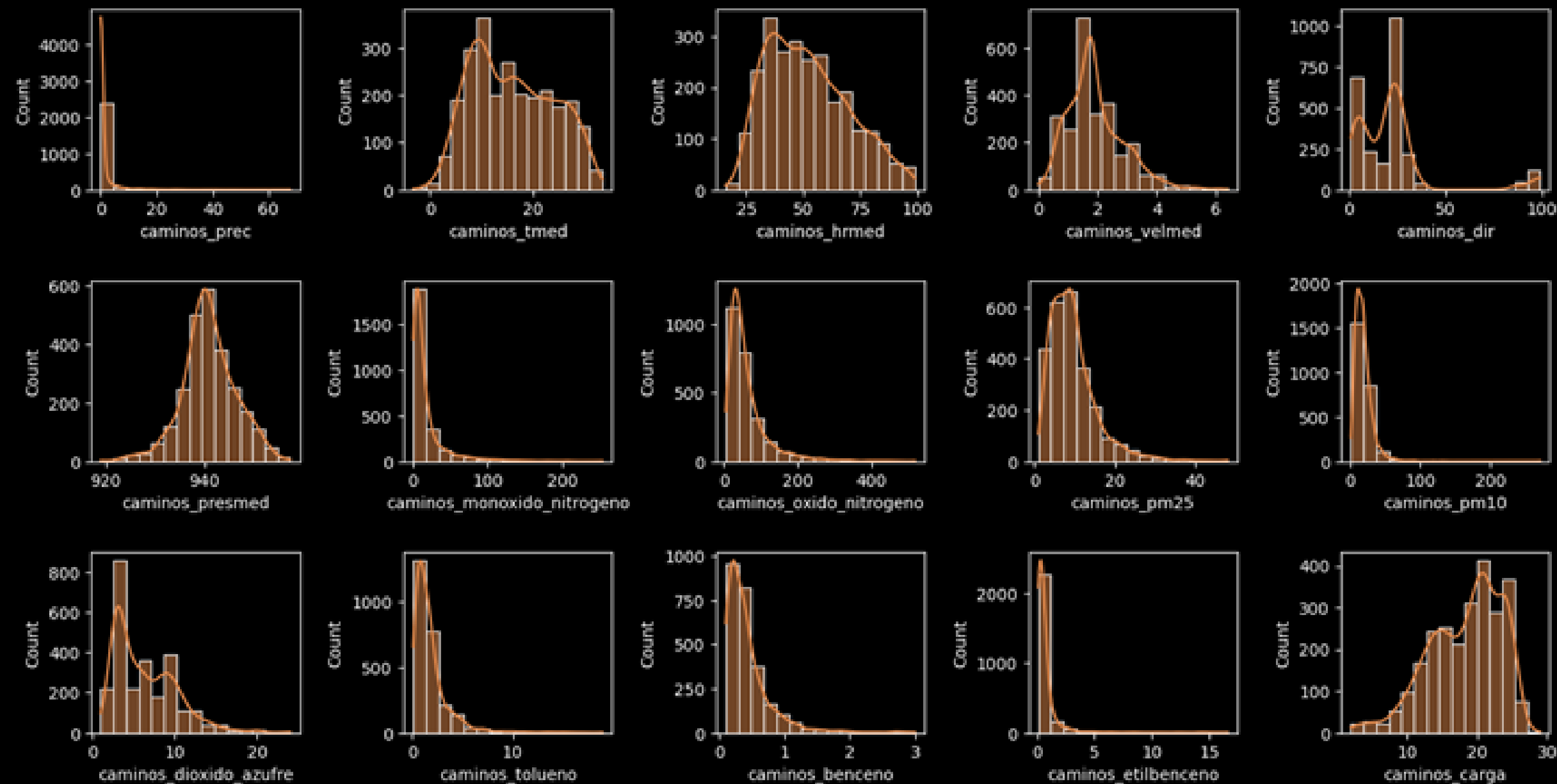
SVMs

Random forest

ADABOOST

XGBoost





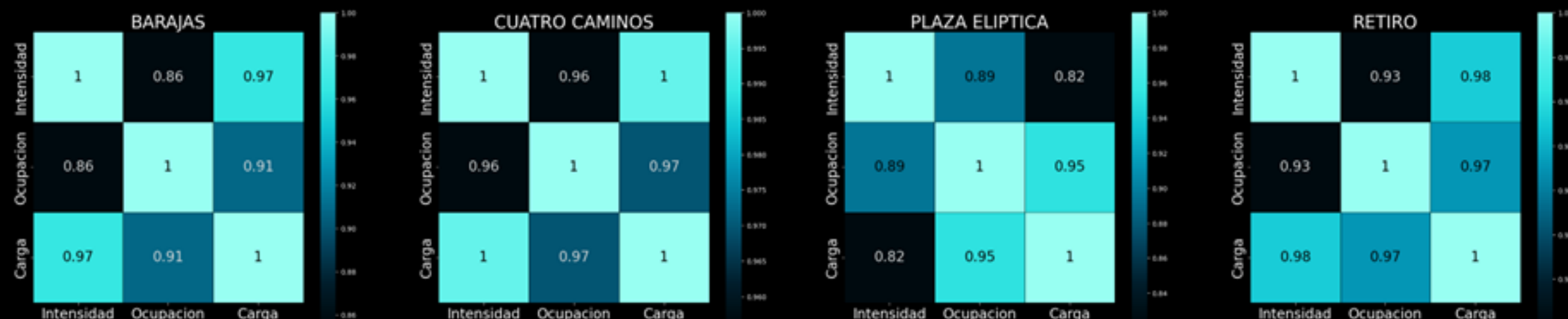
Contaminantes

Meteorología:

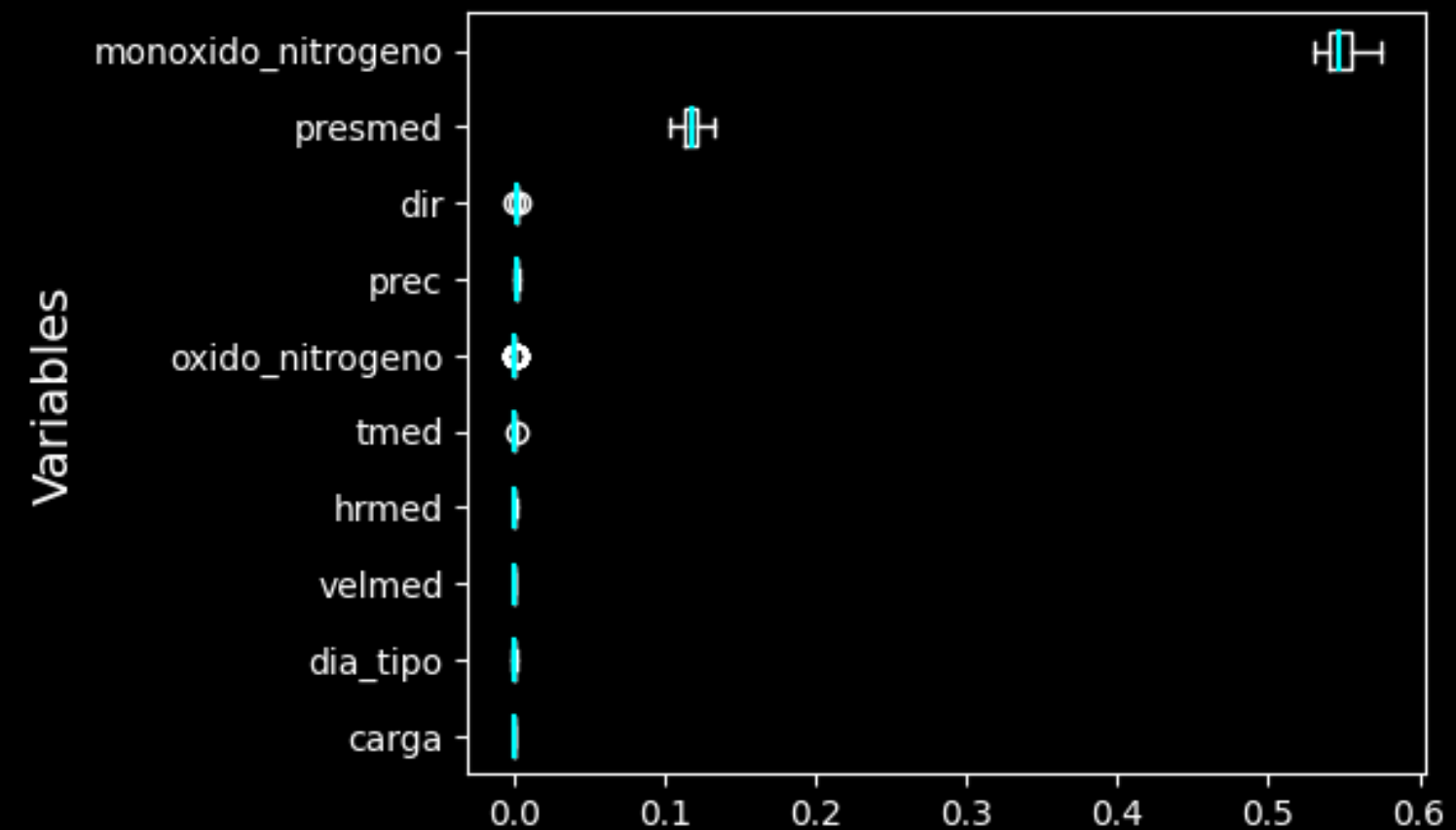
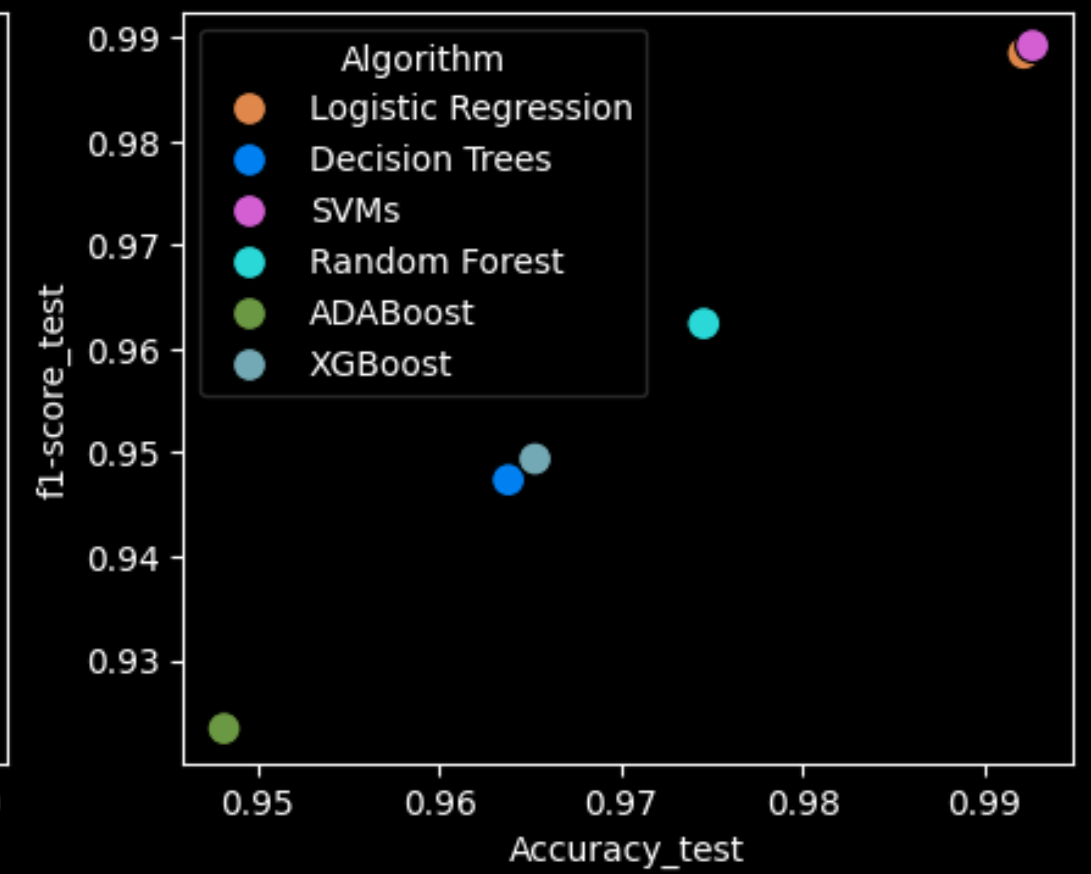
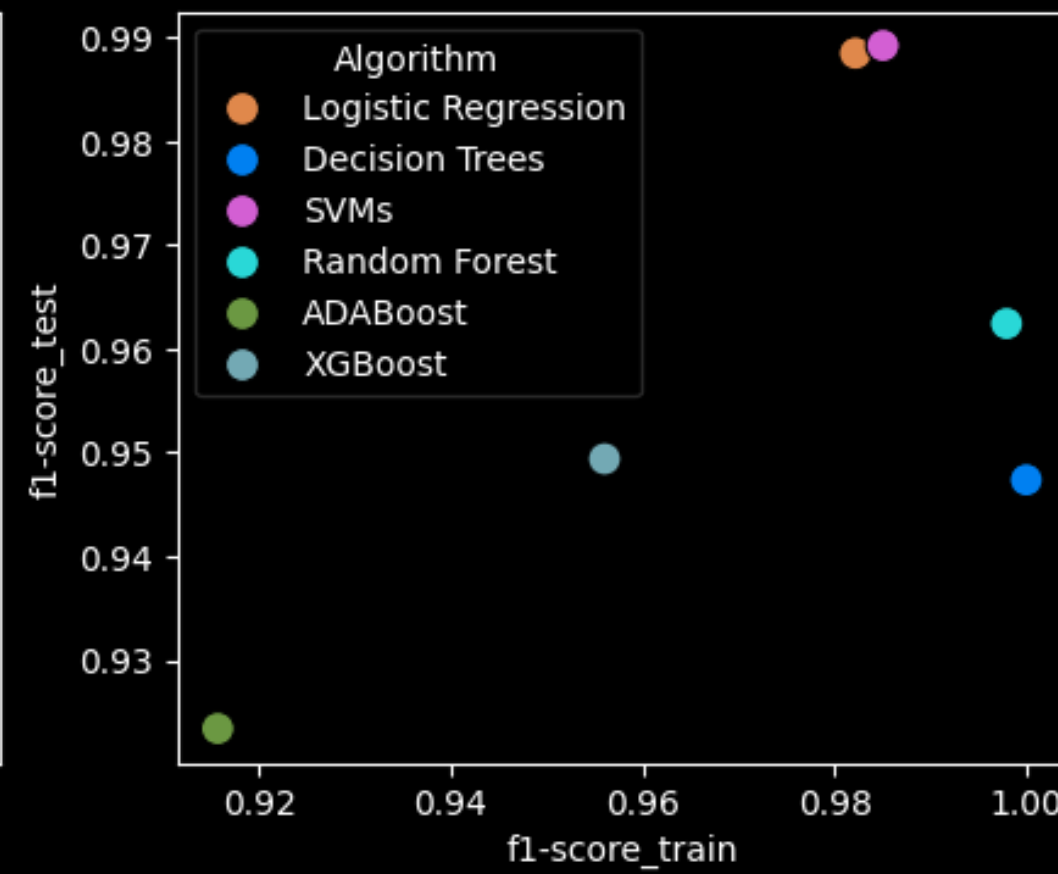
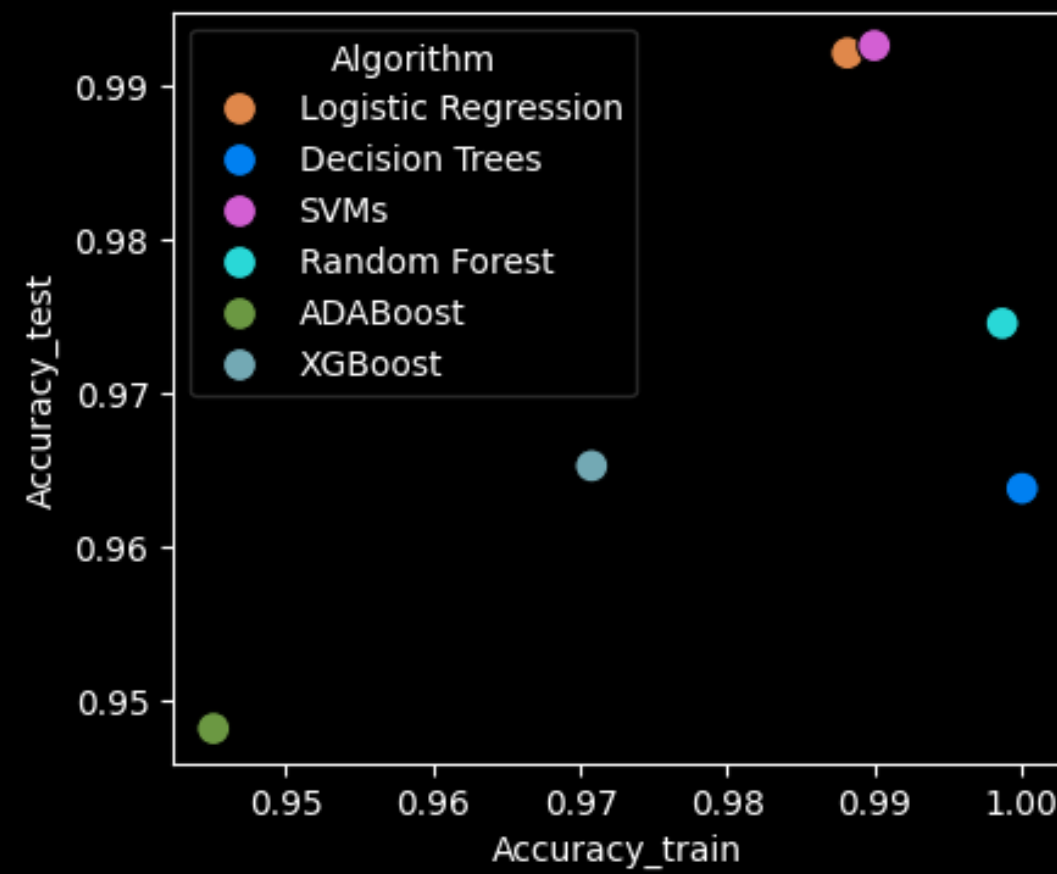
- Valores medios
- Precipitaciones
- Dirección del viento

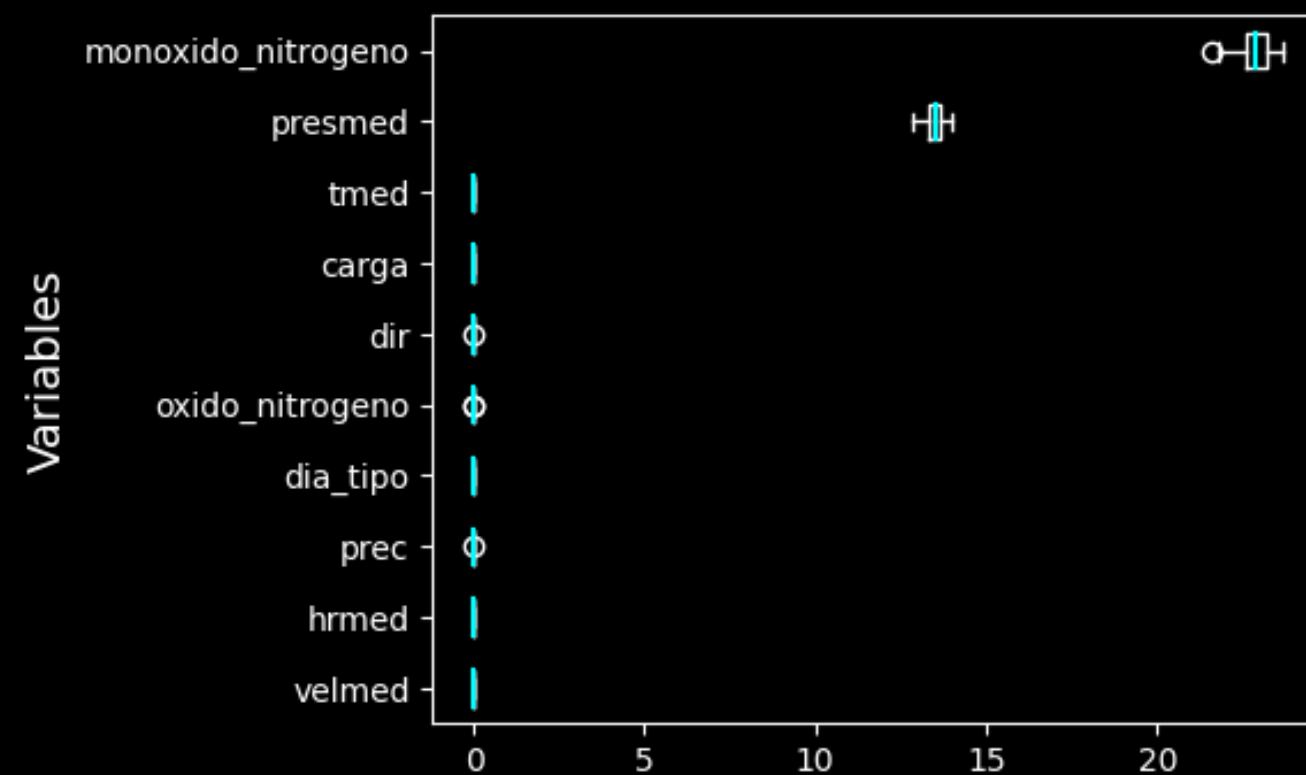
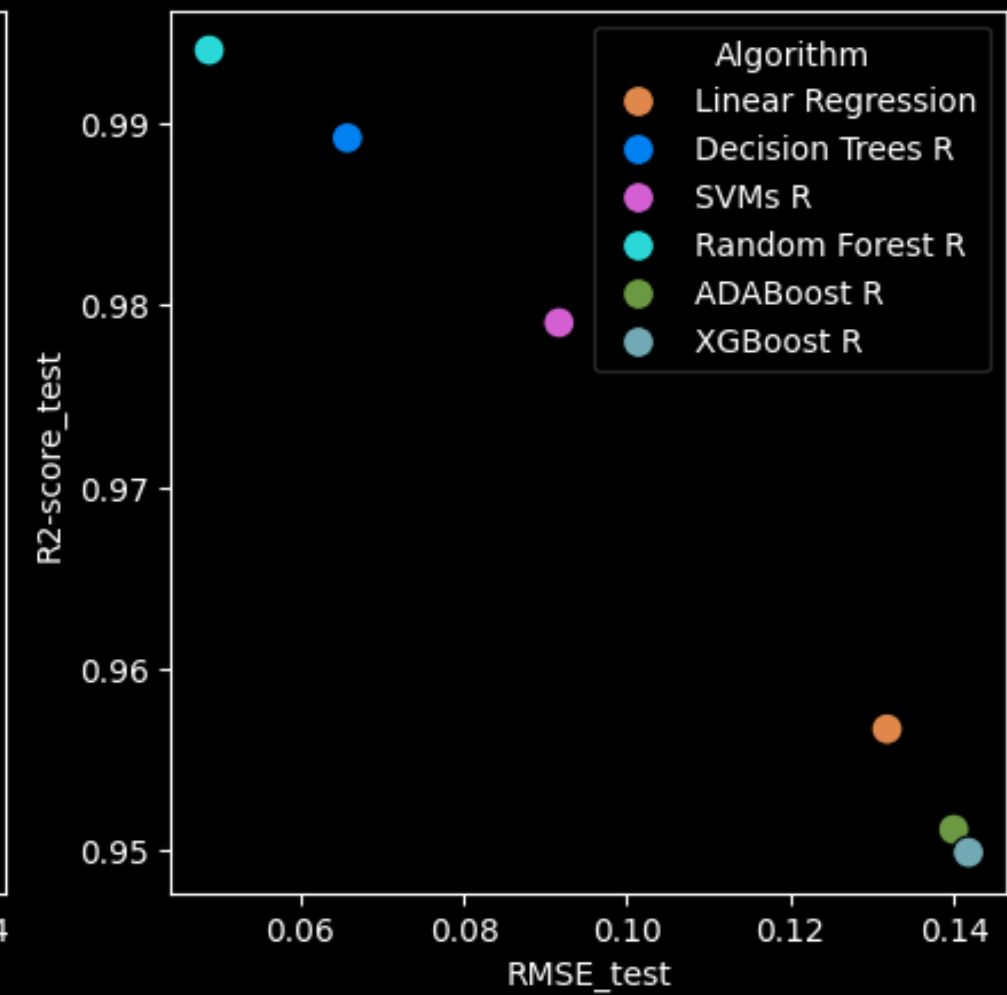
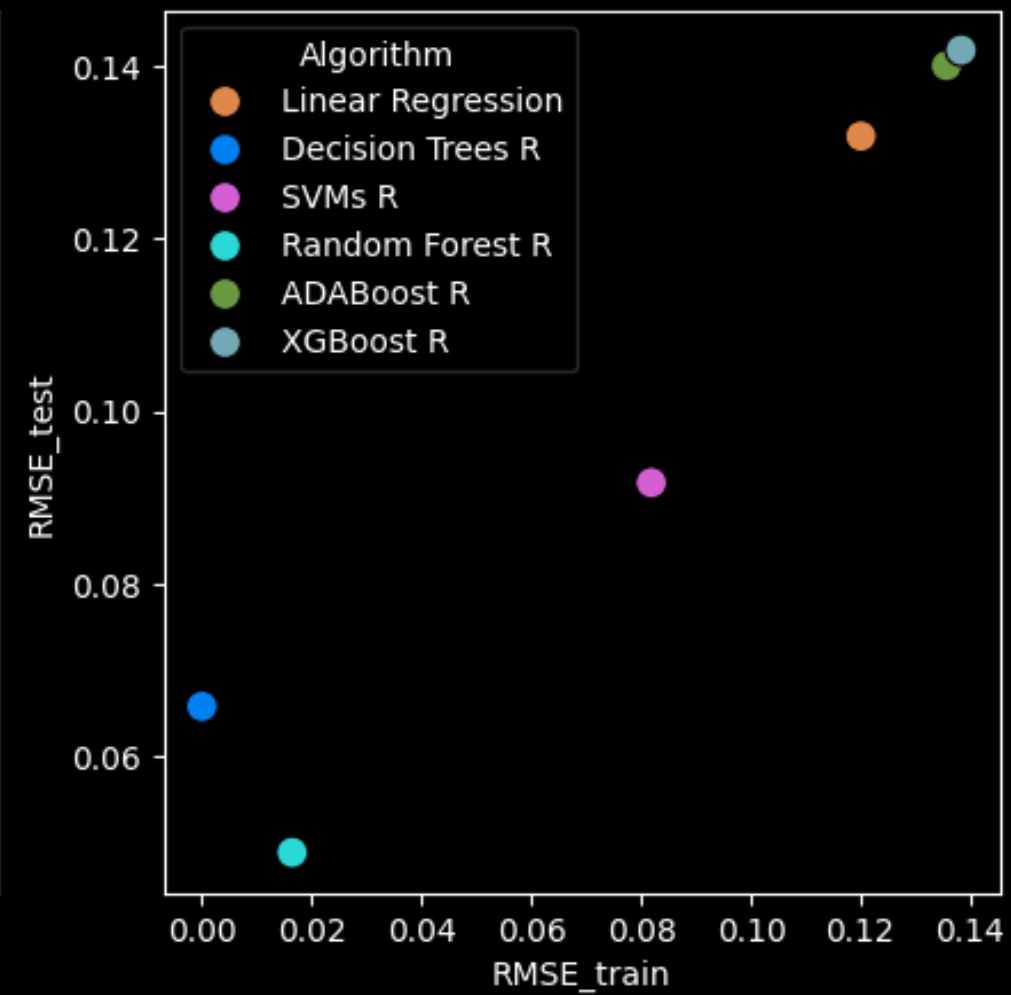
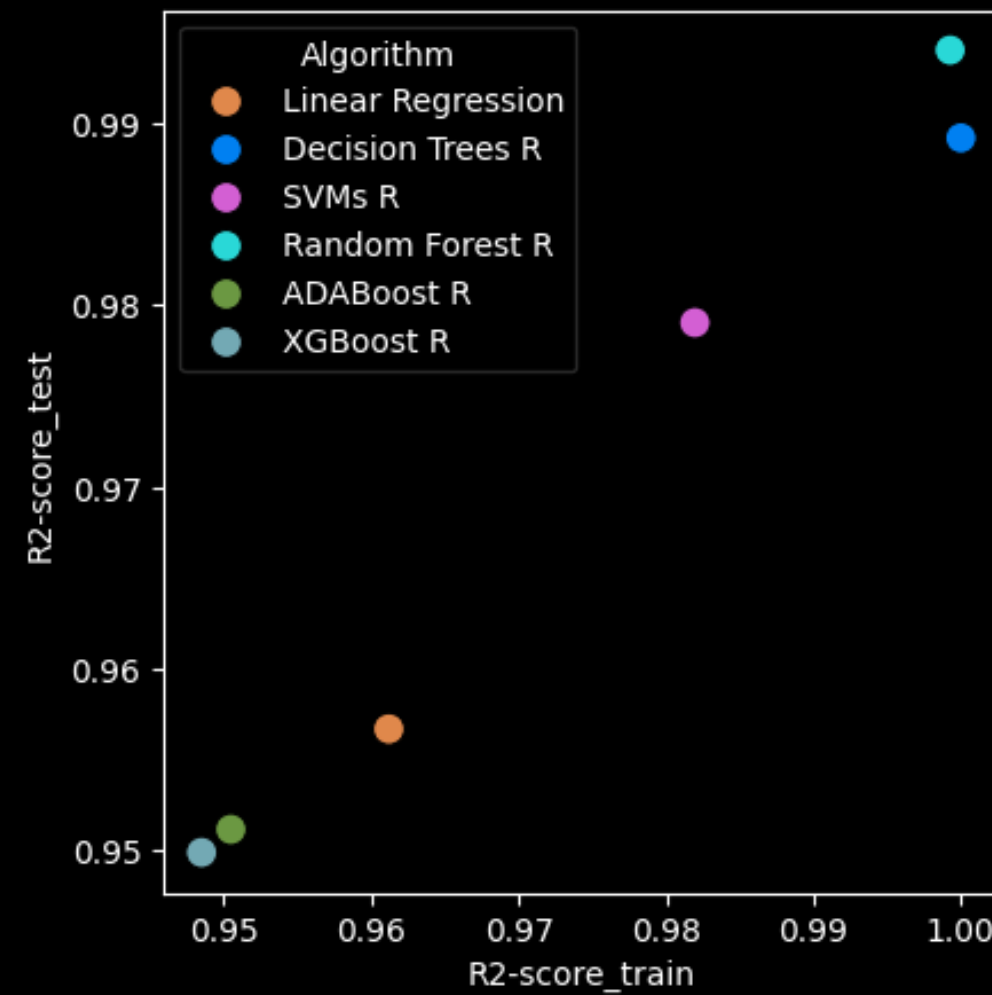
Tráfico:

- Carga
- Tipo de día



Clasificación





Regresión

Conclusiones

El análisis exploratorio de los datos reveló picos de presencia de contaminantes relacionados al nitrógeno:

- meses de invierno
- días laborables

La calidad del aire en Madrid (con respecto al nitrógeno) se ve más afectada por el tráfico vehicular que por los factores meteorológicos.

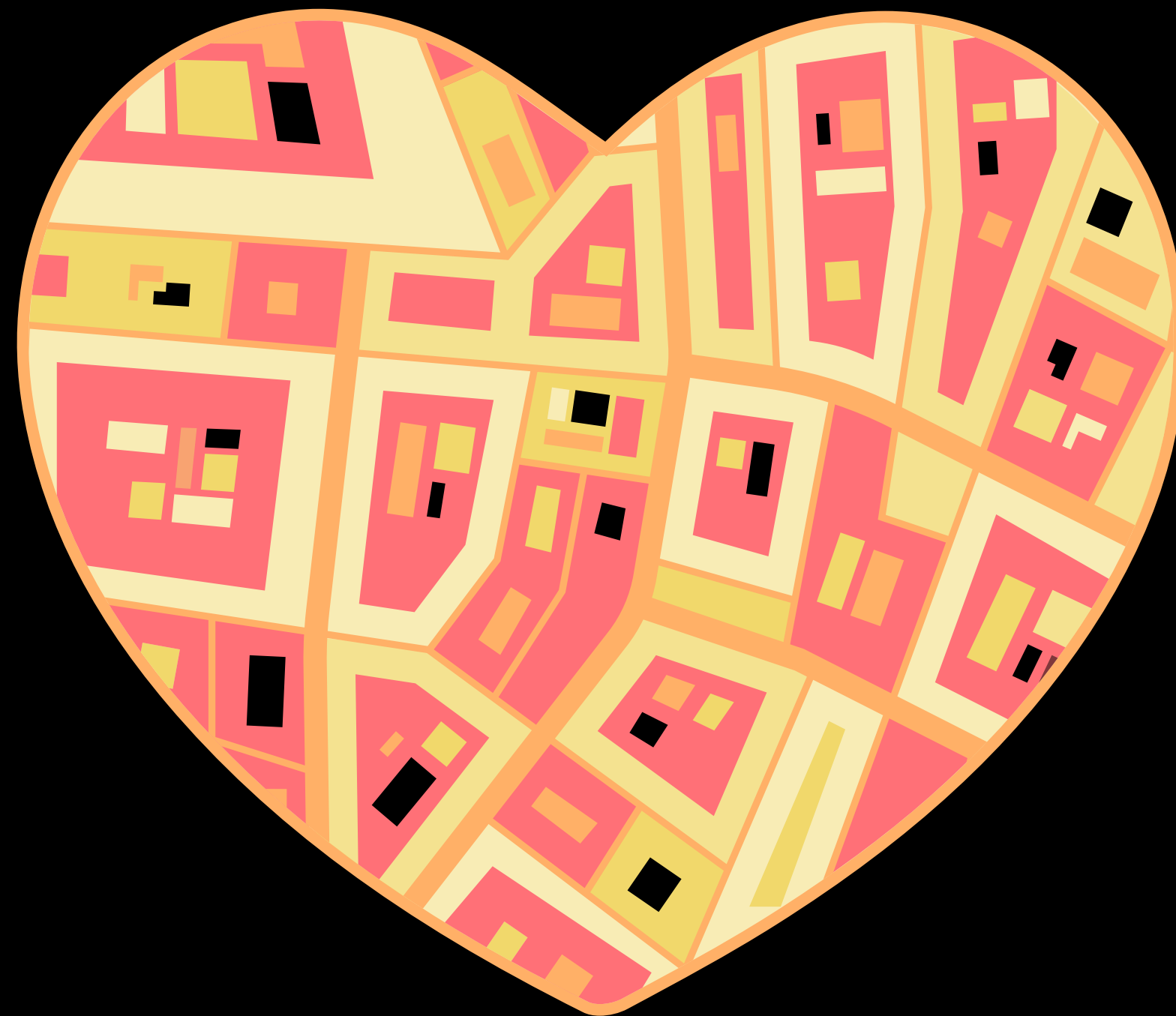
Para mejorar la calidad del aire, se deben implementar políticas y acciones dirigidas a reducir las emisiones de estos contaminantes y tener en cuenta las variaciones estacionales.

AYUNTAMIENTO

Madrid bate en 2023 un nuevo mínimo histórico de contaminación

La capital cumple por segundo año con la directiva europea de calidad del aire

Los modelos de aprendizaje automático predijeron satisfactoriamente los niveles de dióxido de nitrógeno (NO₂) con un buen desempeño.



Siguientes pasos