

## PRINCIPIOS DE MODELADO Y SIMULACIÓN

### PRIMERA ENTREGA GRUPO 5

Paula Juliana Cañón Ávila<sup>1</sup>, Valentina Pérez Yaya<sup>2</sup>, Laura Lucia Cuellar Morales<sup>3</sup>.

[paula.canona@utadeo.edu.co](mailto:paula.canona@utadeo.edu.co), [valentina.perezy@utadeo.edu.co](mailto:valentina.perezy@utadeo.edu.co), [laural.cuellarm@utadeo.edu.co](mailto:laural.cuellarm@utadeo.edu.co).

Para el proyecto se utilizó una base de datos obtenida de la Red de Monitoreo de Calidad de Aire de Bogotá (RMCAB) con el fin de analizar y visualizar datos recolectados acerca de la calidad del aire en diferentes zonas de la ciudad de Bogotá, la base de datos contiene datos desde el 2019 hasta el 2022 con el propósito de analizar la calidad del aire antes, durante y al finalizar el periodo de la pandemia de COVID-19.

### Diccionario de datos

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
DateTime	PM10_(µg/m3)	CO_(ppm)	OZONO_(ppm)	NO_(ppm)	NO2_(ppm)	NOX_(ppm)	SO2_(ppm)	Vel_Viento_(m/s)	Dir_Viento_(grados)	Temperatura_(C)	Precipitacion_(mm)	PM10_Flow_(L/m)	PM2.5_(µg/m3)	Localidad
1/04/2019	42,3	1,8	4,9	41,5	22,1	63,5	2,1	1,4	301	13,6	0	12,6	10,2	Puente_aranda
1/04/2019	55,9	1,4	12,6	26	24,8	50,7	1,5	1,8	326	17,2	0,3	12,5	28,8	Puente_aranda
1/04/2019	40,8	1,4	7,2	19,7	30,2	49,9	0,7	2,6	308	13,9	7,9	12,6	16,4	Puente_aranda
2/04/2019	28,3	1,5	1	29,8	26,2	56	0,7	1,4	333	12,6	1,8	12,7	13,4	Puente_aranda
2/04/2019	41,6	1,1	11,2	17,4	24,2	41,6	1,5	2,8	292	14,6	10,5	12,6	22,7	Puente_aranda
2/04/2019	53,1	2,6	1,7	76,6	39,2	115,8	1,3	1,5	234	12,2	1,6	12,7	20,8	Puente_aranda
3/04/2019	82,9	1,7	1	62,4	27,7	90,1	2,1	1,6	223	10,7	0	12,8	32,8	Puente_aranda
3/04/2019	76,1	1,3	28,3	22,9	28,1	50,9	2,4	3,5	248	16,8	0	12,5	43,3	Puente_aranda
3/04/2019	40,8	1,1	11,9	6,9	24,1	31	1,1	2,9	291	14,3	0	12,5	13,1	Puente_aranda
4/04/2019	36,4	0,9	11,4	9,1	18,9	28	0,9	1,9	291	12,9	0	12,6	9,6	Puente_aranda
4/04/2019	49,3	1	26,8	10,4	15,3	25,6	1,4	3,8	281	17,7	0	12,4	25,5	Puente_aranda
4/04/2019	34,9	0,7	14,1	4,9	16,6	21,5	0,8	3,3	287	15,3	0	12,5	13,8	Puente_aranda
5/04/2019	36,9	0,8	9,2	9,3	17,4	26,7	0,9	2,2	303	12,6	0,3	12,6	9,4	Puente_aranda
5/04/2019	59,6	0,9	23,2	11,1	17,9	28,9	3,3	3,7	275	17,8	0	12,4	35,9	Puente_aranda
5/04/2019	62,3	1	9,2	8,4	24,5	32,8	1,3	2,9	275	14,6	0	12,5	26,7	Puente_aranda
6/04/2019	34,5	0,8	11,3	8,7	16,5	25,2	0,9	2,3	318	13,2	0	12,6	13,5	Puente_aranda
6/04/2019	37,4	0,8	28,1	6,2	13,2	19,5	1,2	3,8	281	17,6	0	12,4	25,8	Puente_aranda
6/04/2019	43,8	0,9	15,2	4,2	20,9	25,1	0,9	3,1	276	15	0	12,5	13,7	Puente_aranda
7/04/2019	36,9	0,9	12,5	8,6	15,5	24	1,1	1,4	291	13,5	0	12,6	12,9	Puente_aranda
7/04/2019	40,6	0,7	33,9	3,2	8,3	11,5	1,1	4,3	222	18,4	0	12,4	30,7	Puente_aranda
7/04/2019	49,6	0,8	26,4	2,8	14,9	17,8	0,8	3,4	188	14,3	0,7	12,5	14,6	Puente_aranda
8/04/2019	53,4	1,1	18,3	17,3	23,6	40,8	1,4	1,5	221	12,4	0	12,7	22,5	Puente_aranda
8/04/2019		0,8	22,6	7,7	17,6	25,3	1,2	3,9	214	18,7	0	12,4	43	Puente_aranda
8/04/2019	67,1	0,9	14,9	8,3	29,4	37,7	1,2	2,7	250	15,2	0	12,5	24,9	Puente_aranda
9/04/2019	67,8	1,1	6,6	33,4	32,2	65,6	1,6	1,6	255	12,7	2,6	12,6	34,2	Puente_aranda
9/04/2019	76,4	1	17,7	17,4	28,9	46,3	2,5	3,4	267	16,6	1,5	12,4	44,5	Puente_aranda
9/04/2019	67,1	1,4	3,2	36,3	38,4	74,7	1,6	2,1	341	15,1	0	12,5	20,9	Puente_aranda

**Imagen 1.** Base de datos utilizada. Tomado de *Red de monitoreo de calidad de aire de Bogotá*.

La base de datos es un archivo de Excel donde se encuentran datos recopilados de los años 2019 a 2023 únicamente de los meses de abril y mayo de cada año para 5 zonas de la ciudad de Bogotá las cuales son

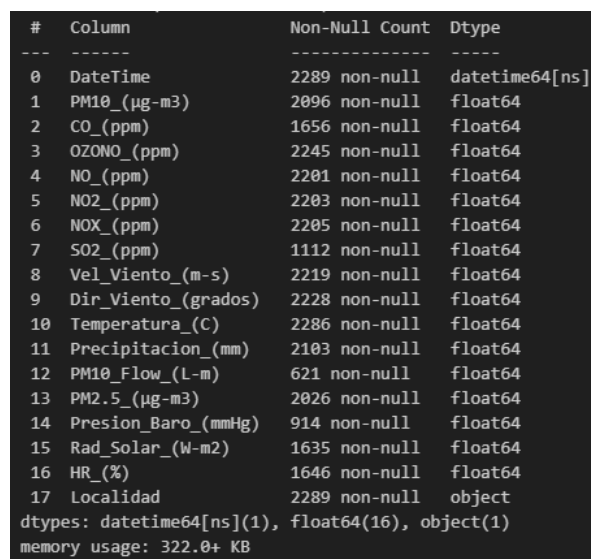
- Puente Aranda
- Centro de alto rendimiento
- Las Ferias
- Guaymaral
- Tunal

Se implemento el método `df.describe(include='all')` el cual proporciona un resumen estadístico de las columnas de nuestra base de datos, incluyendo la cuenta, media, desviación estándar, valores mínimos y máximos, y los percentiles (25%, 50%, 75%).

## Información General de la Tabla

- **Nombre de la Tabla:** Red de monitoreo de calidad de aire de Bogotá
- **Descripción:** Base de datos de la secretaría de ambiente de Bogotá
- **Fecha de Creación:** 2019-2023
- **Autor:** Secretaría de ambiente de Bogotá

La base de datos tiene en total 18 columnas con las siguientes variables



#	Column	Non-Null Count	Dtype
0	DateTime	2289 non-null	datetime64[ns]
1	PM10_(µg-m3)	2096 non-null	float64
2	CO_(ppm)	1656 non-null	float64
3	OZONO_(ppm)	2245 non-null	float64
4	NO_(ppm)	2201 non-null	float64
5	NO2_(ppm)	2203 non-null	float64
6	NOX_(ppm)	2205 non-null	float64
7	SO2_(ppm)	1112 non-null	float64
8	Vel_Viento_(m-s)	2219 non-null	float64
9	Dir_Viento_(grados)	2228 non-null	float64
10	Temperatura_(C)	2286 non-null	float64
11	Precipitacion_(mm)	2103 non-null	float64
12	PM10_Flow_(L-m)	621 non-null	float64
13	PM2.5_(µg-m3)	2026 non-null	float64
14	Presion_Baro_(mmHg)	914 non-null	float64
15	Rad_Solar_(W-m2)	1635 non-null	float64
16	HR_(%)	1646 non-null	float64
17	Localidad	2289 non-null	object

dtypes: datetime64[ns](1), float64(16), object(1)  
memory usage: 322.0+ KB

**Imagen 2.** Inspección de la base de datos usando Python. Elaboración propia.

## Descripción de las variables

Todas las variables cuentan con las unidades en las que fueron medidas.

**Columna 1.** Datetime: representa la fecha de la toma del dato en el formato día/mes/año

**Columna 2.** PM10\_(µg-m3) : partículas en suspensión atmosférica con un diámetro menor o igual a 10 µm.

**Columna 3.** CO\_(ppm): concentración de CO medida en partes por millón 50 ppm no genera efectos en la salud durante 8 horas de exposición.

**Columna 4.** OZONO\_(ppm): gas contaminante medido en partes por millón. Se recomienda que la cantidad máxima de exposición al ozono debe ser de 0,1 ppm diarios.

**Columna 5.** NO\_(ppm): Óxido nítrico medido en partes por millón. La máxima exposición es de 20 ppm al día.

**Columna 6.** NO2\_(ppm): Dióxido de nitrógeno medido en partes por millón. La máxima exposición permitida es de 0.053 ppm en un día.

**Columna 7.** NOX\_(ppm) Mezcla de gases muy reactivos que contienen oxígeno y nitrógeno, tiene que ver con los gases de combustión.

**Columna 8.** SO2\_(ppm): Dióxido de azufre. La máxima exposición permitida al día es de 200 ppm.

**Columna 9.** vel\_viento(m-s): Velocidad del viento medida en metros por segundo.

**Columna 10.** Dir\_Viento\_(grados): Direccion del viento medida en grados.

**Columna 11.** Temperatura\_(C): Temperatura ambiente medida en grados Celsius.

**Columna 12.** Precipitacion\_(mm): Precipitación de lluvia medida en mm

**Columna 13.** PM10\_Flow\_(L-m): flujo de partículas PM10 medida en L/m

**Columna 14.** PM2.5\_(µg-m3): partículas de suspensión atmosférica que tienen un diámetro de 2.5 micrómetros, este contaminante fue medido en micrómetros por metro cúbico.

**Columna 15.** Presion\_Baro\_(mmHg): Presión barométrica medida en milímetros de mercurio.

**Columna 16.** Rad\_Solar\_(W-m2): Radiación solar medida en watts por metro cuadrado.

**Columna 17.** HR\_(%): porcentaje de humedad relativa en el aire, es la relación entre la presión parcial del vapor de agua y la presión de vapor de equilibrio del agua a una temperatura dada.

**Columna 18.** Localidad donde se tomo la medición en Bogotá.

Para la visualización de datos se utilizó la plataforma Power BI, empleando la base de datos previamente mencionada (Imagen 1). Se realizó una versión resumida de esta tabla (Imagen 3), enfocada en el análisis de dos contaminantes: monóxido de carbono (CO) y ozono (O3). Estos contaminantes fueron seleccionados debido a que la Resolución 2254 de 2017 (Imagen 4), emitida por el Ministerio de Ambiente y Desarrollo Sostenible, establece los contaminantes criterio, sus niveles máximos permisibles según el tiempo de exposición, y los niveles de prevención y alerta, además del Índice de Calidad del Aire (ICA). Según esta norma, el CO y el O3 son los contaminantes que se miden cada 8 horas, por lo que se analizó este periodo de tiempo durante los meses de abril y mayo de 2019 a 2021. En la base de datos anterior se incluía la estación de monitoreo Guaymaral, pero se decidió eliminarla en la nueva tabla subida a Power BI. Esto se debe a que, en varias fechas, esta estación no registraba datos de concentración de monóxido de carbono, lo que dificultaba calcular un promedio confiable para un periodo tan extenso. Además, no era viable reemplazar estos datos faltantes con el valor de 0, ya que una concentración de 0 representaría un valor real, y no es preciso asumir que en ciertos días la contaminación fue inexistente.

	A	B	C	D
1	Fecha	CO µg/m3	O3 µg/m3	Estacion
2	1/04/2019	2083,38843	12,8394229	puente aranda
3	1/04/2019	1620,41322	33,0156589	puente aranda
4	1/04/2019	1620,41322	18,8660908	puente aranda
5	2/04/2019	1736,15702	2,62029039	puente aranda
6	2/04/2019	1273,18182	29,3472524	puente aranda
7	2/04/2019	3009,33884	4,45449366	puente aranda
8	3/04/2019	1967,64463	2,62029039	puente aranda
9	3/04/2019	1504,66942	74,154218	puente aranda
10	3/04/2019	1273,18182	31,1814556	puente aranda
11	4/04/2019	1041,69421	29,8713104	puente aranda
12	4/04/2019	1157,43802	70,2237824	puente aranda
13	4/04/2019	810,206612	36,9460945	puente aranda
14	5/04/2019	925,950413	24,1066716	puente aranda
15	5/04/2019	1041,69421	60,790737	puente aranda
16	5/04/2019	1157,43802	24,1066716	puente aranda
17	6/04/2019	925,950413	29,6092814	puente aranda
18	6/04/2019	925,950413	73,6301599	puente aranda
19	6/04/2019	1041,69421	39,8284139	puente aranda

**Imagen 3.** Base de datos resumida utilizada en la visualización de datos. Tomada y modificada de *Red de monitoreo de calidad de aire de Bogotá*.

Cabe mencionar que en la Resolución 2245 se mide los índices de calidad en unidades de µg/m3, por lo que fue necesario realizar la conversión de unidades que nos proporciona la pagina de RMCAB las cuales en el caso del monóxido de carbono y el ozono fueron en ppm.

Índice de Calidad del Aire			Puntos de corte del ICA						
ICA	Color	Categoría	PM <sub>10</sub> µg/m <sup>3</sup> 24 horas	PM <sub>2.5</sub> µg/m <sup>3</sup> 24 horas	CO µg/m <sup>3</sup> 8 horas	SO <sub>2</sub> µg/m <sup>3</sup> 1 hora	NO <sub>2</sub> µg/m <sup>3</sup> 1 hora	O <sub>3</sub> µg/m <sup>3</sup> 8 horas	O <sub>3</sub> µg/m <sup>3</sup> 1 hora <sup>(1)</sup>
0-50	Verde	Buena	0 – 54	0 – 12	0 – 5094	0 – 93	0 – 100	0 – 106	.....
51-100	Amarillo	Aceptable	55 – 154	13 – 37	5095 – 10819	94 – 197	101 – 189	107 – 138	.....
101 - 150	Naranja	Dañina a la salud de Grupos Sensibles	155 – 254	38 – 55	10820 – 14254	198 - 486	190 – 677	139 - 167	245 – 323
151 - 200	Rojo	Dañina a la salud	255 – 354	56 – 150	14255 – 17688	487 – 797	678 - 1221	168 - 207	324 – 401
201 - 300	Púrpura	Muy dañina a la salud	355 – 424	151 – 250	17689 – 34862	798 – 1583	1222 – 2349	208 - 393	402 – 794
301 - 500	Marrón	Peligrosa	425 - 604	251 – 500	34863 - 57703	1584 - 2629	2350 - 3853	394 <sup>(2)</sup>	795 – 1185

Imagen 4. Tabla de Resolución 2245 de 2017. Tomado de: Ministerio de Ambiente y Desarrollo Sostenible.

$$\text{Microgramos por metro cúbico } \left(\frac{\mu\text{g}}{\text{m}^3}\right) = \text{ppm} * PM_x * \left(\frac{10^3}{24.5}\right)$$

Peso molecular del CO (PM): 28,1 g/mol

Peso molecular del O3 (PM): 48 g/mol

Para la visualización de datos, se cargó la tabla con las concentraciones de los dos contaminantes y de las cuatro estaciones a analizar. Se decidió transformar y organizar los datos de la siguiente manera:

Fecha	Estacion	Atributo	1.2 Valor	ABC 123 Personalizado
1/04/2019	puente aranda	CO µg/m3	2083,38843	Antes
1/04/2019	puente aranda	O3 µg/m3	12,8394229	Antes
1/04/2019	puente aranda	CO µg/m3	1620,413223	Antes
1/04/2019	puente aranda	O3 µg/m3	33,0156589	Antes
1/04/2019	puente aranda	CO µg/m3	1620,413223	Antes
1/04/2019	puente aranda	O3 µg/m3	18,8660908	Antes
2/04/2019	puente aranda	CO µg/m3	1736,157025	Antes
2/04/2019	puente aranda	O3 µg/m3	2,620290389	Antes
2/04/2019	puente aranda	CO µg/m3	1273,181818	Antes
2/04/2019	puente aranda	O3 µg/m3	29,34725235	Antes
2/04/2019	puente aranda	CO µg/m3	3009,338843	Antes
2/04/2019	puente aranda	O3 µg/m3	4,454493661	Antes
3/04/2019	puente aranda	CO µg/m3	1967,644628	Antes
3/04/2019	puente aranda	O3 µg/m3	2,620290389	Antes
3/04/2019	puente aranda	CO µg/m3	1504,669421	Antes
3/04/2019	puente aranda	O3 µg/m3	74,154218	Antes
3/04/2019	puente aranda	CO µg/m3	1273,181818	Antes
3/04/2019	puente aranda	O3 µg/m3	31,18145562	Antes
4/04/2019	puente aranda	CO µg/m3	1041,694215	Antes
4/04/2019	puente aranda	O3 µg/m3	29,87131043	Antes

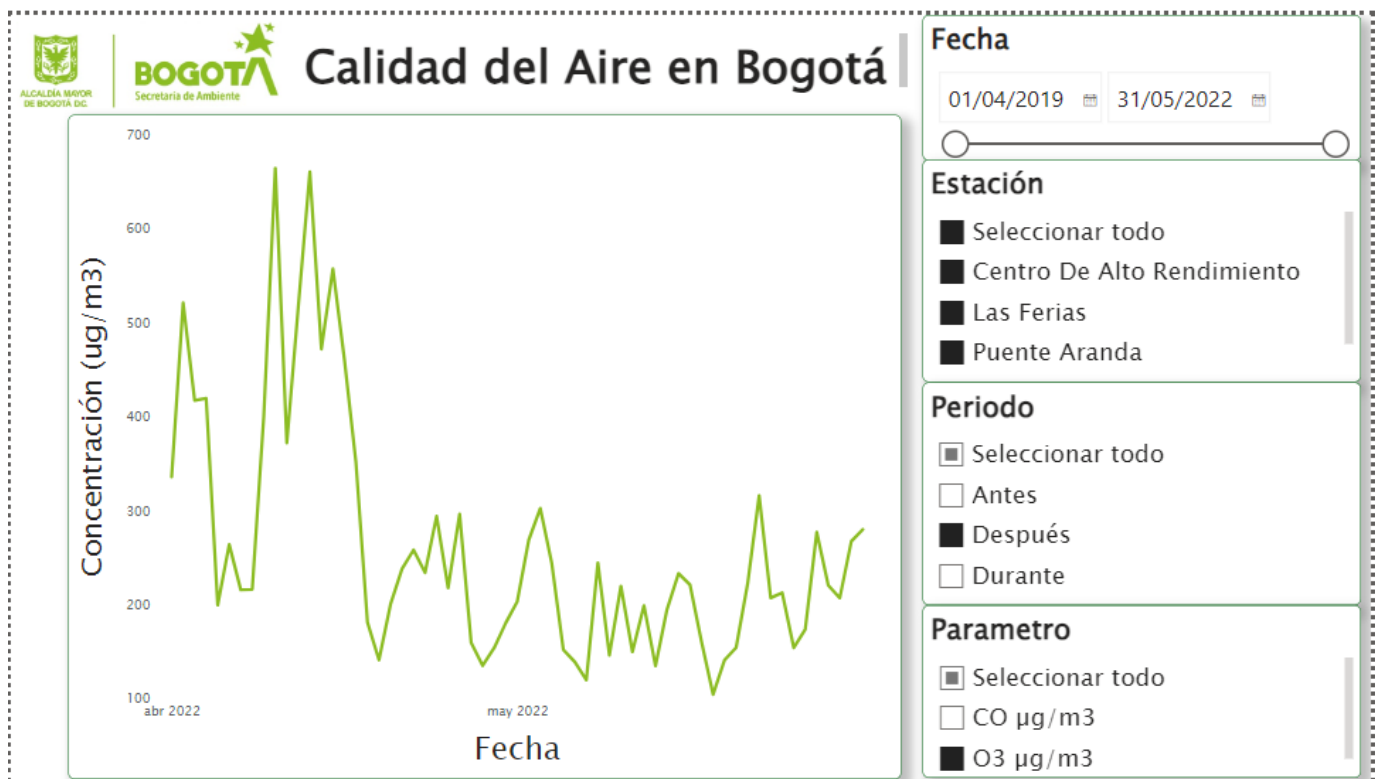
Imagen 5. Datos subidos en Power BI. Elaboración propia.

Se decidió agregar una columna que especifica las etapas "antes", "durante" y "después" de la pandemia según los años. El año 2019 se clasificó como "antes", los años 2020 y 2021 como "durante", y el año 2022 como "después". Esta clasificación se implementó mediante una función condicional "if-then".

```
= Table.AddColumn("#Columnas quitadas", "Personalizado", each if [#"Fecha "] < #date(2020, 3, 20)
then "Antes" else if [#"Fecha "] < #date(2022, 1, 1) then "Durante" else if [#"Fecha "] > #date
(2022, 1, 2) then "Después" else null)
```

**Imagen 6.** Función condicional en Power BI. Elaboración propia.

Finalmente, se construyó la visualización de datos utilizando diversos filtros: 1) por fecha, 2) por estación, 3) por periodo, y 4) por parámetro. Se optó por emplear un gráfico de líneas, ya que es el más adecuado para representar la evolución de los datos a lo largo del tiempo (fechas).

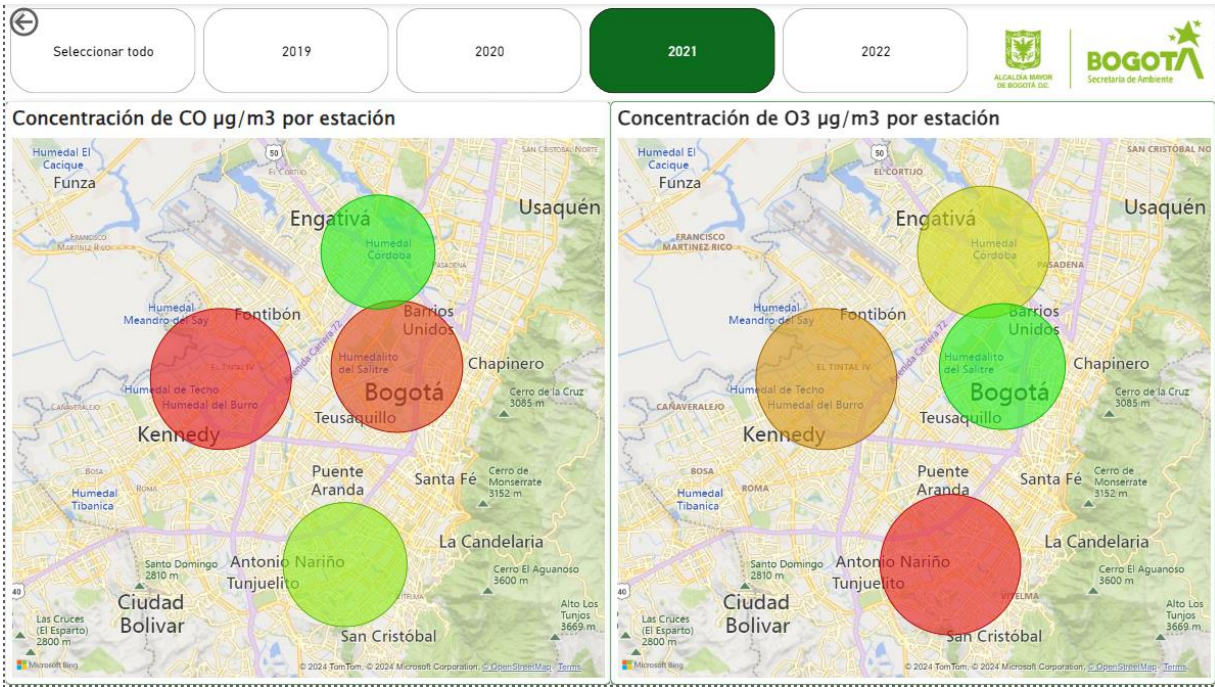


**Imagen 7.** Visualización de datos por gráfico de líneas. Elaboración propia.

Se llevó a cabo un análisis visual mediante mapas aéreos de la ciudad de Bogotá, en el cual se relacionaron los cuatro puntos de monitoreo con cada uno de los contaminantes estudiados, lo cual requirió una modificación de la base de datos inicial al añadir la dirección exacta del punto de monitoreo de cada localidad y adicionándola al Power BI.

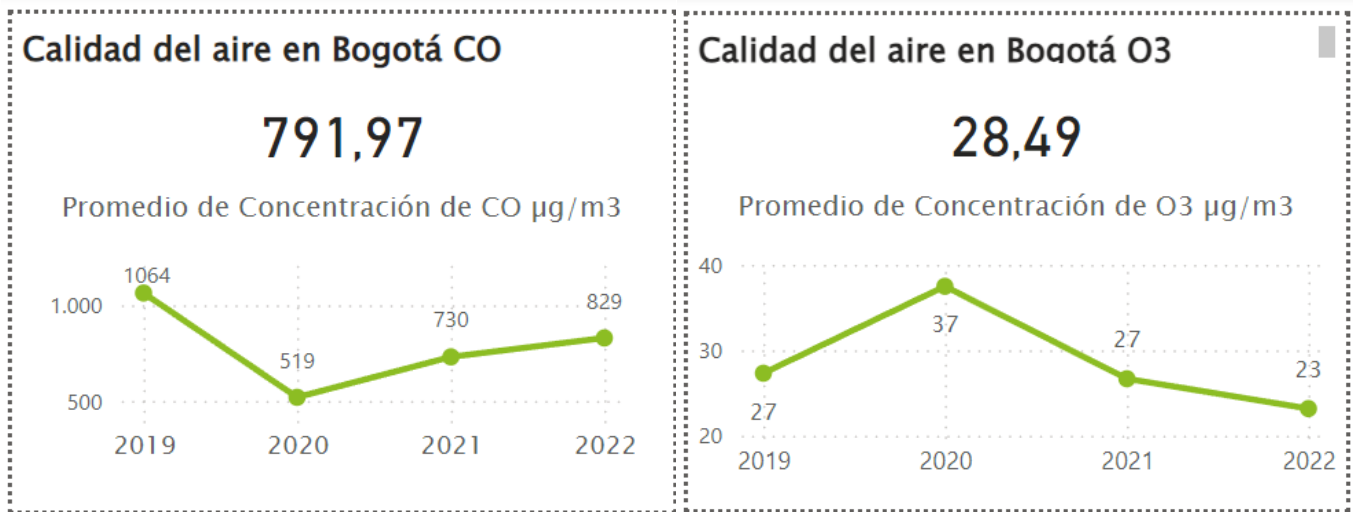
Estos contaminantes se compararon utilizando una escala de colores, donde el color rojo indica los valores más altos del contaminante y el color verde representa los valores más bajos en comparación con el promedio. Para modificar la visualización de los campos visuales, es posible ajustar el año a analizar o seleccionar todos los años, como se muestra en la imagen 7.





**Imagen 7.** Visualización de datos por mapa aéreo. Elaboración propia.

Se añadió una “leyenda” a cada estación en la visualización de mapas aéreos, donde se puede observar la gráfica promedio que se presenta en la imagen 8, así como el promedio de dicho contaminante a lo largo de los cuatro años.



**Imagen 8.** Leyenda de promedio para la visualización de mapa aéreo. Elaboración propia.

## Preguntas:

1. ¿Cuál de las estaciones monitoreadas, Puente Aranda, Tunal, Las Ferias o Centro de Alto Rendimiento, ¿registró los niveles más altos de contaminantes durante el período de la pandemia?
2. Observando las variaciones significativas en las concentraciones de monóxido de carbono y ozono a lo largo de los periodos marcados como "antes", "durante" y "después" de la pandemia, ¿qué factores específicos podrían haber contribuido a los picos y caídas notorios, especialmente durante los años 2020 y 2021, en comparación con los periodos "antes" y "después" de la pandemia?
3. Dado que las estaciones de monitoreo incluyen áreas con distintos puntos industriales y urbanísticos, como Puente Aranda y el Centro de Alto Rendimiento, ¿cómo cambian las tendencias de contaminación entre estas estaciones?
4. ¿Qué diferencias se observan en los niveles de monóxido de carbono (CO) entre las distintas estaciones, y cómo podría estar relacionado esto con la ubicación geográfica y las características de tráfico vehicular o industrial de cada área monitoreada?
5. Teniendo en cuenta la visualización del comportamiento de los contaminantes en los mapas aéreos en las diferentes zonas de Bogotá ¿se evidencia un impacto positivo o negativo después del año 2020?
6. Según la visualización de los mapas aéreos, ¿en qué año se presentó una disminución significativa de ambos contaminantes?
7. ¿Cuál es la zona que tiene una menor contaminación de CO y O<sub>3</sub> en promedio de los 4 años?
8. ¿Considera que la ubicación geográfica tiene un papel importante en cuanto a la disminución significativa del contaminante CO para la zona de Centro de alto rendimiento en los años 2019 a 2021?, ¿Cuál otra razón provocaría esta disminución?
9. ¿Considera que los niveles en cada zona son óptimos?
10. ¿Qué factores cree que son los responsables de que en el año 2020 hubiera un aumento significativo de O<sub>3</sub> en todas las zonas?