

PRINCIPIOS DE MODELADO Y SIMULACIÓN

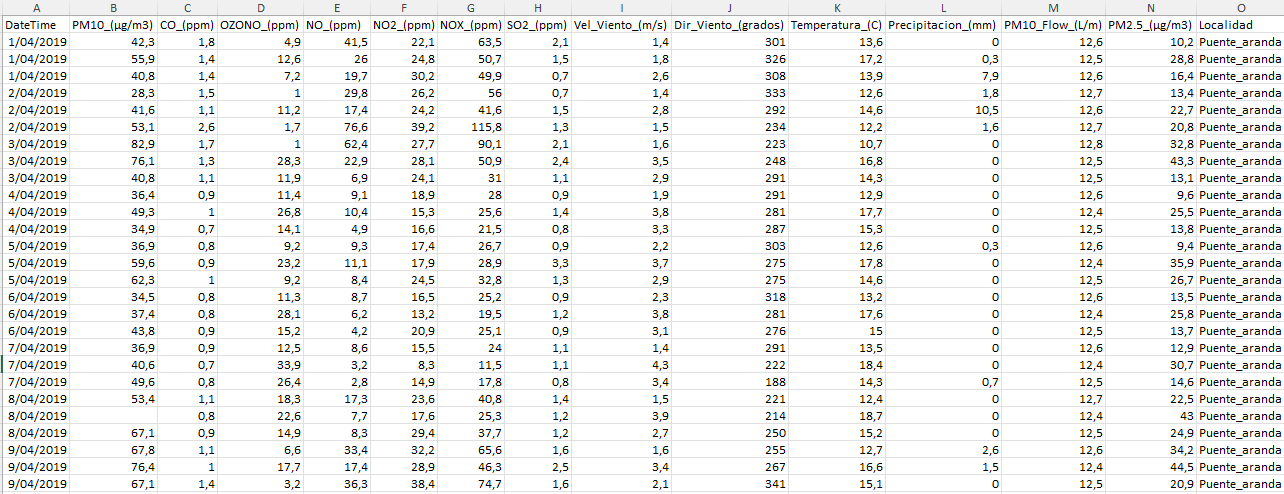
PRIMERA ENTREGA GRUPO 5

Paula Juliana Cañón Ávila1, Valentina Pérez Yaya2, Laura Lucia Cuellar Morales3.

[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 cuidad 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**

****

**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

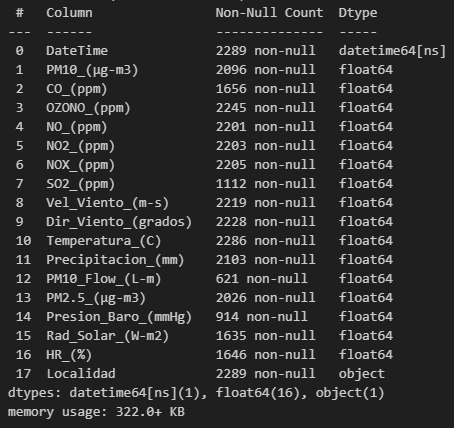
* Puente Aranda
* Centro de alto rendimiento
* Las Ferias
* Guaymaral
* Tunal

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



**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.

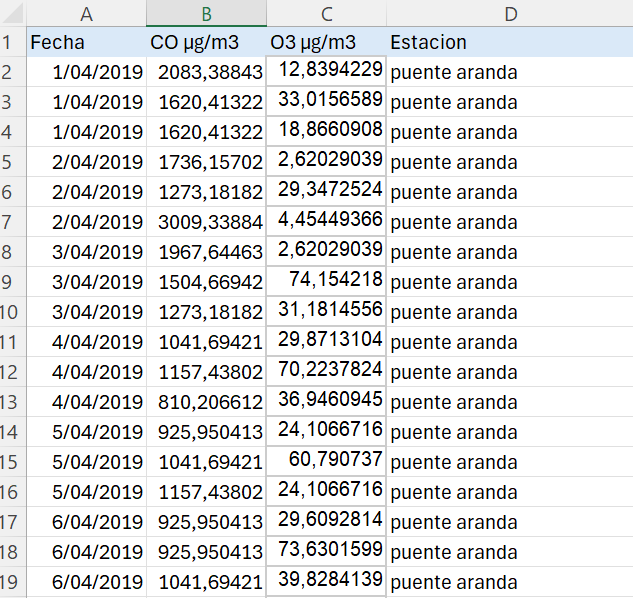
**Columna 15**. Presion\_Baro\_(mmHg): Presión barométrica medida en milímetros de mercurio.

**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.

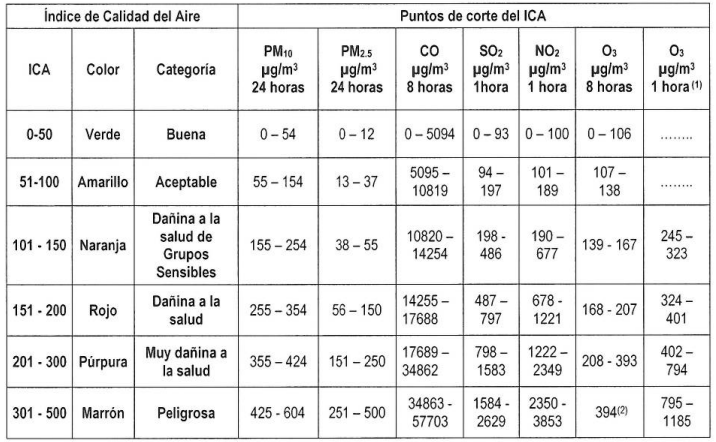
**Columna 18.** Localidad donde se tomo la medición en Bogotá.

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.



**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.



**Imagen 4.** Tabla de Resolución 2245 de 2017. Tomado de: Ministerio de Ambiente

y Desarrollo Sostenible.

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:

Tabla

Descripción generada automáticamente

**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".

Imagen que contiene Texto

Descripción generada automáticamente

**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).

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

**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.

Gráfico, Gráfico de burbujas

Descripción generada automáticamente

**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.

Escala de tiempo, Gráfico de líneas

Descripción generada automáticamenteGráfico, Gráfico de líneas

Descripción generada automáticamente

**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 O3 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 O3 en todas las zonas?