

JOSE DAVID CARDONA MAZO - Avance 1

Proyecto Integrado III - Analítica de Datos (Sharon Karin Camacho) -
PREICA2502B020071

UNIVERSIDAD DIGITAL DE ANTIOQUIA - 2025

Primer Avance	
Tema	Ambiente y Desarrollo Sostenible - Calidad Del Aire En Colombia (Promedio Anual)
Link Dataset	https://www.datos.gov.co/resource/kekcd-7v7h.json?
Columnas	28733
Fuente	Portal de Datos Abiertos de Colombia
Objetivo de análisis:	Evaluar la influencia de las variables meteorológicas (temperatura, precipitación, viento, radiación solar, humedad) en la concentración y variabilidad de los contaminantes atmosféricos (PM10, PM2.5, NO2, SO2, CO, O3, entre otros) en las estaciones de monitoreo de calidad del aire de Colombia, para identificar patrones y predecir posibles excedencias de los límites establecidos por las normativas ambientales.
Pregunta de Investigación:	¿Cómo afectan las variables meteorológicas (temperatura, precipitación, viento, radiación solar, humedad) a la concentración y variabilidad de los contaminantes atmosféricos (PM10, PM2.5, NO2, SO2, CO, O3, entre otros) en las estaciones de monitoreo de calidad del aire en Colombia?
Hipótesis	Las condiciones meteorológicas (temperatura, precipitación, viento, radiación solar y humedad) tienen un impacto significativo en las concentraciones y la variabilidad de los contaminantes atmosféricos medidos, donde cambios en las condiciones climáticas, como temperaturas extremas o baja humedad, se correlacionan con niveles más altos de contaminación.
Métricas de Éxito:	Correlación entre las condiciones meteorológicas y los contaminantes (R^2).
	Número de días de excedencia de los límites de calidad del aire.
	Desviación estándar de los contaminantes y su relación con las condiciones meteorológicas.
	Precisión del modelo predictivo (evaluado con RMSE o R^2).
Qué datos tengo para lograr el objetivo	Sí, estos datos son útiles para lograr el objetivo del análisis, Las variables incluyen tanto contaminantes como factores meteorológicos, lo que nos permite analizar cómo afectan las condiciones ambientales a la calidad del aire

DICCIONARIO DE DATOS

Columna	Descripción	Tipo de Dato	Ejemplo
id_estacion	Identificador único asignado a la estación de monitoreo.	Numérico	20383
autoridad_ambiental	Autoridad ambiental encargada de la estación.	Texto	CORANTIOQUIA
estacion	Nombre de la estación de monitoreo.	Texto	ALTAVISTA
latitud	Coordenada geográfica que indica la latitud de la estación.	Numérico	6.222.584
longitud	Coordenada geográfica que indica la longitud de la estación.	Numérico	-75.628.207
variable	Contaminante o variable meteorológica medida en la estación (PM10, PM2.5, temperatura, etc.).	Texto	PM10
unidades	Unidad de medida para la variable.	Texto	ug/m3
tiempo_de_exposicion_horas	Horas de exposición durante el monitoreo.	Numérico	1
a_o	Año en el que se realizaron las mediciones.	Numérico	2011
promedio	Promedio anual de la medición de la variable (contaminante o meteorológica).	Numérico	75.5
suma	Total acumulado anual de la medición de la variable.	Numérico	329454.0
no_de_datos	Número total de datos válidos recopilados en ese año de medición.	Numérico	4363
representatividad_temporal	Porcentaje del tiempo que tiene datos válidos registrados.	Categórico	50%
excedencias_límite_actual	Número de veces que el contaminante ha superado el límite permitido en ese año.	Numérico	0
porcentaje_excedencias_límite	Porcentaje de veces que se superó el límite permitido durante el año.	Numérico	0.0%
mediana	Mediana de la medición del contaminante durante el año.	Numérico	64.0
percentil_98	Percentil 98 de la medición del contaminante durante el año (valor por debajo del cual está el 98% de los datos).	Numérico	208.0
maximo	Valor máximo de la medición durante el año.	Numérico	524.0
fechas_horas_del_maximo	Fecha y hora del valor máximo registrado durante el año.	Fecha/Tiempo	27/10/2011 8:00:00 p. m.
minimo	Valor mínimo de la medición durante el año.	Numérico	7.0

fechas_horas_del_minimo	Fecha y hora del valor mínimo registrado durante el año.	Fecha/Tiempo	27/10/2011 8:00:00 p. m.
dias_de_excedencias	Número de días en que se superó el límite permitido.	Numérico	0
codigo_del_departamento	Código DANE del departamento donde se encuentra la estación.	Numérico	5
nombre_del_departamento	Nombre del departamento.	Texto	ANTIOQUIA
codigo_del_municipio	Código DANE del municipio donde se encuentra la estación.	Numérico	5001
nombre_del_municipio	Nombre del municipio.	Texto	MEDELLÍN
tipo_de_estacion	Tipo de estación de monitoreo (fija o móvil).	Texto	Fija
ubicacion	Información geoespacial de la estación (coordenadas en formato JSON).	Geoespacial	{'type': 'Point', 'coordinates': [-75.628207, 6.222584]}