

# Informe de Análisis de la Concentración de PM2.5

## Análisis de los Gráficos y Datos

### 1. Gráfico de Dispersión

El gráfico de dispersión muestra la concentración diaria de PM2.5 ( $\mu\text{g}/\text{m}^3$ ) en función del tiempo (días). A simple vista, la distribución de los puntos no sigue una línea clara, sino que forma una nube dispersa, lo que indica una **ausencia de correlación lineal**. Aunque se ha trazado una línea de regresión, su pendiente es casi plana, lo que visualmente confirma la falta de relación.

### 2. Resultados de la Regresión Lineal

La aplicación de un modelo de regresión lineal sobre los datos produjo los siguientes resultados:

- **Ecuación de la recta de regresión:**  $y=0.0005 \times \text{Tiempo} + 7.9152$ 
  - **Pendiente (0.0005):** Este valor, es muy cercano a cero, indica que por cada día que pasa, la concentración de PM2.5 solo aumenta en  $0.0005 \mu\text{g}/\text{m}^3$ . Este aumento es tan insignificante que se considera **estadísticamente irrelevante**.
  - **Intercepto (7.9152):** Representa la concentración de PM2.5 estimada en el día 0 del conjunto de datos.
- **Valor de  $R^2$  (R cuadrado):** 0.0002
  - El valor de  $R^2$  mide la proporción de la varianza en la variable dependiente (concentración de PM2.5) que puede ser predicha por la variable independiente (tiempo). Un valor de  **$R^2$  de 0.0002** es demasiado bajo. Esto significa que **el tiempo explica solo el 0.02% de la variabilidad total** en la concentración de PM2.5. El 99.98% restante es explicado por otros factores.

## Conclusiones

Basado en el análisis de los gráficos y los datos, se concluye que:

- No existe una relación lineal significativa entre el tiempo y la concentración de PM2.5.
- El modelo de regresión lineal utilizado es un **predictor ineficaz** de la concentración de PM2.5.
- La variabilidad observada en los niveles de PM2.5 se debe principalmente a **factores externos** que no se incluyeron en este análisis, como las condiciones climáticas (viento, lluvia, temperatura) o las emisiones de fuentes locales (tráfico, industria).

## Recomendaciones

Para futuros análisis, se sugiere la inclusión de otras variables explicativas. Un **modelo de regresión múltiple** que incorpore datos meteorológicos, información sobre días festivos, o registros de actividad industrial podría ofrecer una comprensión más completa de los factores que impulsan la contaminación por PM2.5. Esto permitiría crear un modelo predictivo más preciso y útil.