

Universidad Rafael Landívar
Inteligencia Artificial
Primer Semestre 2025
Ing. Max Cerna

Laboratorio 1

Análisis de Datos Climáticos

Un equipo de meteorología del INSIVUMEH ha recolectado datos sobre la temperatura diaria de la ciudad de Guatemala durante el año 2024. El objetivo es analizar las tendencias, identificar valores atípicos y visualizar la información para tomar decisiones sobre futuras estrategias climáticas.

Tarea

Dado un archivo CSV llamado *clima.csv*, que contiene los siguientes datos:

Fecha: Fecha en formato YYYY-MM-DD.

Temperatura: Temperatura diaria en grados Celsius.

Precipitación: Cantidad de lluvia diaria en milímetros.

Realiza las siguientes tareas utilizando *Pandas*, *NumPy* y *Matplotlib*

1. Carga el archivo CSV en un DataFrame de Pandas.
2. Verifica si hay datos faltantes y, si los encuentras, rellénalos correspondiente.
3. Calcula:
 - a. Temperatura media, máxima y mínima del año.
 - b. Precipitación total del año.
4. Identifica valores atípicos en la columna temperatura utilizando el rango intercuartílico (IQR = $Q3 - Q1$):
 - a. Calcula el primer cuartil (Q1) y el tercer cuartil (Q3).
 - i. Determina los límites inferior y superior para identificar valores atípicos.
 - ii. Filtra los días con temperaturas fuera de estos límites y muestra el resultado.
5. Genera las siguientes gráficas:
 - a. Una gráfica de línea que muestre la evolución de la temperatura a lo largo del año.
 - b. Un histograma que muestre la distribución de las temperaturas.
 - c. Una gráfica de barras que muestre la precipitación total por mes.



Outliers Formula

$$\text{Lower Outlier} = Q1 - (1.5 \times IQR)$$

$$\text{Higher Outlier} = Q3 + (1.5 \times IQR)$$



Entregables

El estudiante deberá:

1. Subir el código Python utilizado en un archivo **.ipynb**.
2. Asegurarse de que todas las gráficas generadas sean claras y estén debidamente etiquetadas.
3. Proporcionar un resumen breve de los hallazgos en formato Markdown dentro del Jupyter Notebook.

ANEXO

Glosario

Un valor atípico o outlier es un dato que se encuentra significativamente alejado del resto de los datos en un conjunto.

Ejemplo de Detección de Outliers con el IQR

Supongamos un conjunto de datos ordenado:

[1, 3, 4, 5, 7, 8, 9, 10, 12]

1. **Determinar los cuartiles:**
 - $n = 9$
 - $Q1 = 4$ (25% de los datos).
 - $Q3 = 9$ (75% de los datos).
2. **Calcular el IQR**
 - $IQR = Q3 - Q1$
 - $9 - 4 = 5$
3. **Detectar valores atípicos:**
 - Límites:
 - Límite inferior = $Q1 - 1.5 \times IQR = 4 - 1.5 \times 5 = -3.5$
 - Límite superior = $Q3 + 1.5 \times IQR = 9 + 1.5 \times 5 = 16.5$
 - Cualquier valor fuera del rango $[-3.5, 16.5]$ sería un outlier.



Outliers Formula

$$\text{Lower Outlier} = Q1 - (1.5 \times IQR)$$

$$\text{Higher Outlier} = Q3 + (1.5 \times IQR)$$