

Taller de Introducción a Python con Pandas y Matplotlib

Dataset: EVSales.csv (Ventas de vehículos eléctricos)

Librerías permitidas: pandas, matplotlib, pyplot

Parte 1: Preguntas Teóricas

Analizar en detalle el dataset y dar respuesta a las siguientes preguntas:

1.1 Describirlo con sus propias palabras.

1.2 Definir 5 potenciales casos de uso que se puedan resolver con ML haciendo uso del dataset.

Parte 2: Preguntas Prácticas

Instrucciones:

- ✓ Descarga el dataset.
- ✓ Responde cada pregunta en una celda de código diferente.
- ✓ Usa comentarios (#) para explicar tu proceso si es necesario.

1. Exploración inicial

- ¿Cuántas filas y columnas tiene el dataset? Muestra las primeras 5 filas.

2. Limpieza básica

- ¿Hay valores nulos en el dataset? Si es así, ¿cuántos por columna?

3. Filtrado simple

- ¿Cuántos vehículos vendidos son de la marca "Tesla" y tienen carga rápida?

4. Agrupación y agregación

- ¿Cuál es el promedio de ingresos (Revenue) por región? Muestra los resultados en un gráfico de barras.

5. Visualización básica

- Genera un histograma de la columna Battery_Capacity_kWh. ¿Cómo se distribuye esta variable?

6. Top marcas más vendidas

- ¿Cuáles son las 5 marcas con más unidades vendidas? Muestra los resultados en un gráfico de barras horizontales.

7. Segmentación por tipo de vehículo

- ¿Cuál es el tipo de vehículo (Vehicle_Type) más vendido en "North America"? Usa un gráfico de pastel para mostrar la distribución.

8. Análisis de descuentos

- ¿Cuál es el porcentaje de descuento promedio aplicado por cada segmento de cliente?

9. Imputar Datos Ausentes

- Seleccionar un feature del dataset, borrar intencionalmente 40 instancias (Pueden seleccionar un rango entre 40 – 100 instancias). Tratar de hallar estos valores haciendo uso de la media y la mediana. Analizar los resultados y llegar a 5 conclusiones del ejercicio.

Los resultados serán revisados durante la sesión.

LET'S CODING ; BE HAPPY



Cata L.