

Tarea: Investigación y Desarrollo de un Análisis Exploratorio de Datos (EDA)

Los coders deben investigar cómo se realiza un proceso de **EDA (Exploratory Data Analysis)** aplicando Python y librerías comunes (**pandas**, **matplotlib**, **seaborn**, etc.).

Parte 1: Análisis Exploratorio de Datos

Cada estudiante debe investigar y responder las siguientes preguntas:

1. ¿Qué es el EDA y cuál es su propósito en el análisis de datos?
Añade un ejemplo práctico de un caso en el que el EDA sea fundamental antes de aplicar modelos de Machine Learning.
2. ¿Qué tipos de datos existen (categóricos, numéricos, ordinales, etc.) y cómo se tratan en un EDA?
3. ¿Qué técnicas básicas se usan en un EDA?
(Ejemplo: estadísticas descriptivas, histogramas, boxplots, correlaciones, detección de valores atípicos).
4. ¿Cuál es la diferencia entre análisis univariado, bivariado y multivariado?
5. ¿Qué es la limpieza de datos y qué tareas suelen incluirse en este paso?
(Ejemplo: manejo de valores nulos, duplicados, outliers, normalización de datos).
6. ¿Qué papel juegan las librerías pandas, matplotlib y seaborn en un EDA?
7. Explica brevemente el flujo típico de un EDA desde la carga del dataset hasta las conclusiones.
8. ¿Qué es una matriz de correlación y para qué sirve en el EDA?
9. ¿Qué son los outliers y qué métodos existen para detectarlos y tratarlos en un análisis exploratorio?
10. ¿Qué es hipótesis testing y para qué sirve en el EDA?

Entrega

1. Subir un PDF al classroom con tus respuestas.