



QUE ES EDA?

Análisis Exploratorio de Datos

Es un proceso analítico que sirve para comprender los datos en profundidad y conocer las diferentes características de estos, generalmente usando medios visuales.

Se utiliza para descubrir tendencias, patrones, verificar supuestos, obtener insights, dar sentido a los datos, además de eliminar irregularidades y valores innecesarios de los datos.







OBJETIVOS EDA

- Identificar y eliminar valores atípicos (outliers) de los datos.
- Identificar tendencias en tiempo y espacio.
- Descubrir patrones relacionados con el objetivo.
- Crear hipótesis y comprobarlas mediante experimentos.
- Identificar nuevas fuentes de datos.



PASOS INVOLUCRADOS EN EL ANÁLISIS EXPLORATORIO DE DATOS

1. Comprensión inicial de los datos

- Explorar la estructura (forma) del DataFrame.
- Revisar las primeras y últimas filas (Head y Tail).
- Identificar tipos de datos (dtypes).
- Obtener estadísticas descriptivas (describe).
- Evaluar correlaciones entre variables.

2. Preparación y limpieza de datos

- Eliminar información irrelevante o duplicada.
- Renombrar columnas para mayor claridad.
- Crear nuevas variables útiles (Feature Engineering).

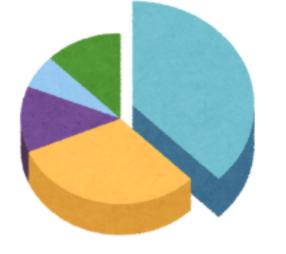
3. Comprensión de variables individuales

 Realizar análisis univariado para entender distribución y características de cada variable por separado.

4. Explorar relaciones entre variables

- Aplicar análisis bivariado para estudiar relaciones entre pares de variables.
- Realizar análisis multivariado para identificar patrones entre múltiples variables.







TIPOS DE ANÁLISIS EXPLORATORIO DE DATOS

Análisis Univariado

En el análisis univariado, el resultado está relacionado con una sola variable, y todos los datos recopilados son para analizarla. No existe ninguna relación causa-efecto.

Análisis Bivariado

En el análisis bivariado, el resultado depende de dos variables, y se analiza la relación entre ellas comparando estas dos variables.

Análisis Multivariado

En el análisis multivariado, el resultado depende de más de dos variables. El análisis se realiza sobre variables que pueden ser numéricas o categóricas. El resultado del análisis se representa en valores numéricos, visualizaciones o gráficos.