

Tarea: Exploración de los datos y análisis descriptivo

Objetivo de la tarea

El propósito de esta tarea es que los estudiantes apliquen herramientas y técnicas de análisis descriptivo de datos, comprendiendo las características principales de los datos, realizando un análisis exploratorio y utilizando sistemas para crear visualizaciones y cuadros de mando.

Instrucciones

1. Descargar un conjunto de datos

Descarga un archivo de datos (puede ser en formato CSV o Excel) con información ficticia de ventas, clientes o cualquier otra temática relevante.

- **Ejemplo de dataset:** Contiene columnas como Fecha, Región, Producto, Ventas, Costos y Satisfacción del cliente.

2. Dividir la tarea en los siguientes apartados:

1. Caracterización de los datos

- **Objetivo:** Identificar las principales características del conjunto de datos.
- **Ejercicio:**
 1. Cargar el conjunto de datos en una herramienta como Excel, Python (pandas), R o Power BI.
 2. Determinar:
 - El tamaño del dataset (número de filas y columnas).
 - Los tipos de datos de cada variable.
 - Los valores faltantes, si existen, y cómo se podrían manejar.
 - La descripción estadística de las variables numéricas (mínimo, máximo, media, mediana, desviación estándar).
 3. Responder las preguntas:
 - ¿Cuáles son las variables categóricas y cuáles las numéricas?
 - ¿Existen valores atípicos en las variables? ¿Cómo los identificaste?

2. Análisis exploratorio de datos

A) Análisis univariante

- **Objetivo:** Analizar una variable a la vez para identificar patrones básicos.
- **Ejercicio:**
 1. Crear gráficos para las variables principales:
 - Histogramas para las variables numéricas (Ventas, Costos).
 - Diagramas de barras para variables categóricas (Región, Producto).
 2. Interpretar los gráficos:
 - ¿Qué distribución tienen las ventas?
 - ¿Qué producto es el más vendido?

B) Análisis multivariante

- **Objetivo:** Explorar relaciones entre dos o más variables.
- **Ejercicio:**
 1. Generar gráficos como:
 - Diagramas de dispersión para analizar la relación entre Ventas y Satisfacción del cliente.
 - Tablas cruzadas para comparar Región y Producto.
 2. Calcular las correlaciones entre las variables numéricas.
 3. Interpretar:
 - ¿Hay una relación entre el costo y las ventas?
 - ¿En qué región se venden más productos?

3. Análisis multidimensional

- **Objetivo:** Construir cuadros de mando que resuman los indicadores clave.
- **Ejercicio:**
 1. Seleccionar indicadores como:
 - Ventas totales por región.

- Producto más vendido.
 - Promedio de satisfacción del cliente por región.
2. Crear un cuadro de mando o dashboard utilizando:
- Excel (tablas dinámicas y gráficos).
 - Power BI o Tableau (si tienen acceso).
3. Responder:
- ¿Qué regiones tienen mejor desempeño?
 - ¿Qué áreas necesitan más atención según los indicadores?
-

4. Sistemas para análisis descriptivo

- **Objetivo:** Aplicar herramientas específicas para análisis descriptivo.
 - **Ejercicio:**
 1. Usar una de las siguientes herramientas para realizar análisis descriptivo y visualización:
 - Python (librerías como pandas, matplotlib y seaborn).
 - R (paquetes como ggplot2).
 - Power BI.
 - Tableau.
 2. Realizar:
 - Una comparación visual de los resultados obtenidos con diferentes sistemas.
 - Una reflexión sobre qué herramienta prefieres y por qué.
-

Entrega

- **Formato:** Un informe en Word o PDF con capturas de los gráficos y las interpretaciones.
- **Fecha límite:** [Definir fecha].
- **Evaluación:** Se calificará la claridad de los análisis, la interpretación de resultados y la creatividad en las visualizaciones.

