

DIPLOMADO EN CIENCIA DE DATOS
UNIVERSIDAD PEDAGÓGICA Y TECNOLÓGICA DE COLOMBIA
COHORTE 2024-1
ABRIL 2024

Teniendo en cuenta tu dataset escogido, formula con tu equipo las siguientes preguntas de interés a ser abordadas:

1. ¿Qué contiene mi dataset?
2. ¿Cuántas filas y columnas tiene mi dataset?
3. ¿Qué variables y qué tipos de variables contiene mi dataset?
4. ¿Qué categorías de variables puedo identificar?
5. ¿Cuántos celdas vacías contiene cada una de mis variables?
6. ¿Puedo identificar una variable objetivo?
7. ¿Cuántas variables son numéricas y cuántas categóricas?
8. ¿Qué variables considero relevantes?
9. ¿Qué estadísticas puedo extraer de mis variables?
10. Escoge al menos 5 variables para trabajar, formula preguntas o plantea situaciones que hagan uso de dichas variables para iniciar a trabajar!
11. Realiza gráficos de visualización para las variables de tu interés

Ejemplo:

Si mi conjunto de datos contiene información de las compras de clientes en función del tiempo, unas preguntas que me puedo plantear sería:

- ¿Cuáles son los meses que tienen mayores compradores?
- ¿Cuáles son los productos con mayores ventas?

Situación: Supongamos que nuestra empresa quiere aplicar una campaña publicitaria para promover las ventas de los productos con menores ventas en función del tiempo...

Supongamos que quiero predecir el comportamiento de los clientes al cambiar los precios de los productos en función del tiempo, o que quiero saber cuál es la mejor estrategia para aumentar el precio del producto más vendido (de acuerdo al principio de oferta-demanda)

Discútelos con tu equipo ;)

No olvides consultar a tus tutores... La idea es que puedas hacerlo por tu cuenta y por supuesto, que puedas trabajar con tu equipo ya que los equipos de datos siempre requieren la colaboración!