

Primer Avance

Tema

Análisis de factores que influyen en el aumento de las ventas en un ecommerce.

Link Dataset

<https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store>

Formato:

CSV

Tamaño:

Nº de filas: 2,753,091

Nº de columnas: 7

Fuente:

Kaggle

Objetivo de análisis:

Identificar los factores que más influyen en la conversión de visitas a compras dentro de una tienda online.

Pregunta de Investigación:

¿Qué variables del comportamiento de los usuarios en el sitio están asociadas con una mayor probabilidad de compra?

Hipótesis:

1. Los usuarios que visualizan más productos tienen mayor probabilidad de comprar.
2. La categoría del producto influye significativamente en la tasa de conversión.
3. Las compras ocurren con mayor frecuencia en ciertos días u horas específicas.

Métricas de Éxito:

- Tasa de conversión: compras / visitas
- Precisión del modelo predictivo si se construye uno (ej. accuracy, ROC-AUC)
- Insights accionables derivados del análisis para estrategias comerciales

Qué datos tengo para lograr el objetivo

- Eventos generados por usuarios (ver producto, agregar al carrito, compra)
- Categoría del producto
- Hora y fecha del evento
- ID del usuario
- Precio del producto
- ID del producto

¿Cómo se llaman las columnas y qué significan? (Diccionario de datos)

Columna	Significado
event_time	Fecha y hora del evento
event_type	Tipo de evento: view, cart, purchase, remove_from_cart
product_id	Identificador único del producto
category_id	Identificador de la categoría del producto
category_code	Código que representa la categoría del producto
brand	Marca del producto
price	Precio del producto
user_id	Identificador del usuario

Granularidad de la tabla:

Cada fila representa un evento individual realizado por un usuario en un instante determinado (por ejemplo, una visualización de producto o una compra).

¿Estos datos me sirven para lograr el objetivo?

Sí, permiten analizar patrones de comportamiento, tasa de conversión y segmentar por producto, usuario, categoría o tiempo.

Tipos de Datos:

Columna	Tipo de dato
event_time	Fecha y hora
event_type	Categorico
product_id	Categorico

category_id	Categórico
category_code	Categórico
brand	Texto
price	Numérico
user_id	Categórico