

Análisis de Ventas de una Tienda Online Minorista

Estructura del Proyecto

Contexto

Este dataset contiene los registros de transacciones de una tienda online minorista con sede en el Reino Unido, que opera a nivel internacional. La tienda vende productos variados, como "White Hanging Heart T-Light Holder", "White Metal Lantern", "Cream Cupid Hearts Coat Hanger" y "Knitted Union Flag Hot Water Bottle". Las transacciones, iniciadas al desde diciembre de 2010 , incluyen compras y devoluciones realizadas por clientes identificados (`CustomerID`) y anónimos, en diversos países, con un ejemplo inicial de "United Kingdom".

La gerencia de esta tienda busca analizar estos datos para entender su rendimiento comercial, optimizar el inventario, mejorar la experiencia del cliente y aumentar los ingresos en un entorno donde las tendencias temporales y las dinámicas de mercado son críticas.

Premisa

La premisa del proyecto es que los datos de ventas registrados en este dataset proporcionan información concreta y accionable para mejorar la toma de decisiones estratégicas de la tienda. Con 541,909 transacciones que incluyen detalles como cantidad (`Quantity`), precio unitario (`UnitPrice`), y fechas específicas (`InvoiceDate`), el análisis puede revelar patrones de compra, productos clave, comportamientos de clientes y problemas operativos como devoluciones o inconsistencias en los datos. Este enfoque basado en datos permitirá a la empresa responder a las condiciones actuales y planificar estrategias futuras con precisión.

Importancia de Analizar esta Data

El análisis de este dataset es fundamental por las siguientes razones:

- **Optimización del inventario:** Los datos de `Quantity` y `StockCode` permiten identificar qué productos tienen mayor demanda o generan más devoluciones, ayudando a gestionar el stock de manera eficiente y reducir costos operativos.
- **Fidelización de clientes:** Con 406,829 transacciones vinculadas a `CustomerID`, analizar el comportamiento de los clientes recurrentes puede impulsar estrategias de retención y personalización.
- **Expansión internacional:** La columna `Country` registra las ubicaciones de los compradores, lo que permite evaluar el desempeño por región y detectar oportunidades de crecimiento en mercados específicos.
- **Gestión de riesgos:** Los valores negativos en `Quantity` (mínimo -80,995) y `UnitPrice` (mínimo -11,062) indican devoluciones o ajustes que afectan los ingresos. Cuantificar y mitigar estos casos es clave para la estabilidad financiera.
- **Planificación temporal:** Las fechas detalladas en `InvoiceDate` ofrecen la capacidad de identificar tendencias estacionales o picos de ventas, mejorando la preparación para períodos de alta demanda.

Este análisis tiene un impacto directo en la operación diaria y establece una base sólida para decisiones estratégicas a largo plazo, utilizando herramientas como `pandas` y `matplotlib` para transformar los datos en resultados prácticos.

Estructura del proyecto

- familiarizarse con el dataset
- limpieza de datos
- tratamiento de datos faltantes
- tratamiento de datos duplicados
- asignar fecha como índice
- evaluar valores atípicos, Quantity

- categorizar las compras hechas con una nueva columna, con las categorías alta para venta total mayor a 50, media para venta total entre 20 y 50 y baja para venta total menor a 20 (pista: función apply)
- obtener columnas de año, y mes para un análisis temporal
- obtener ranking de los 3 países con mejores ventas
- obtener las ventas por semestre en el último año y graficar
- obtener las ventas por trimestre en el último año y graficar
- obtener las proporciones de categorías de gastos y graficar
- obtener el top 10 productos más vendidos
- que otra información consideras relevante para el propósito de este proyecto, considera al menos 2