



Benemérita Universidad Autónoma de Puebla

"Ciudad Universitaria 2"

Introducción a la Ciencia de Datos

PROFESOR: JAIME ALEJANDRO ROMERO

Reporte Fase 1: "Descubrimiento del Proyecto"

Alumno:

JULIAN SÁNCHEZ OSORIO

21 de octubre de 2024

"Optimización del Cumplimiento de Pedidos en Amazon"

Objetivo del proyecto

"Reducir la tasa de cancelación de pedidos en un 20% en los próximos 12 meses."

Descripción del Problema

La tasa de cancelación de pedidos es un problema creciente que afecta a la satisfacción del cliente y la rentabilidad de la empresa.

Este problema no solo afecta la satisfacción del cliente, sino que también tienen un impacto negativo en los ingresos y la reputación de la marca. Los motivos detrás de las cancelaciones son diversos y pueden incluir demoras en el envío, falta de disponibilidad de productos y problemas de comunicación durante el proceso de compra.

Por lo tanto, es esencial identificar y analizar los motivos detrás de estas cancelaciones para desarrollar estrategias efectivas que mejoren la experiencia de compra.

Este proyecto busca identificar los factores que contribuyen a las cancelaciones de pedidos mediante un análisis detallado de los datos disponibles, con el fin de implementar estrategias que optimicen el cumplimiento y mejoren la satisfacción del cliente para un aumento en las ventas.

Recursos Disponibles

Tecnología y Herramientas

Python: Para la manipulación y análisis de datos.

Pandas: Para el manejo de estructuras de datos.

Matplotlib y Seaborn: Para la visualización de datos.

Jupyter Notebooks: Para la presentación y desarrollo de análisis interactivos.

Datos-Descripción de las columnas:

index: Identificador secuencial único para cada fila en el conjunto de datos.

Order ID: Identificador único para cada pedido realizado.

Date: Fecha en la que se realizó el pedido.

Status: Estado actual del pedido (Ej. Cancelado, Enviado).

Fulfilment: Método de cumplimiento del pedido (Ej. Fulfilled by Amazon, Merchant).

Sales Channel: Canal a través del cual se realizó la venta (Ej. Amazon.in).

ship-service-level: Nivel de servicio de envío seleccionado (Ej. Standard, Expedited).

Category: Categoría del producto (Ej. Ropa, Electrónica).

Size: Tamaño del producto (Ej. S, M, L).

Courier Status: Estado del envío con el transportista (Ej. En tránsito, Entregado).

Qty: Cantidad de productos en la orden.

currency: Tipo de moneda utilizada en la transacción.

Amount: Monto total de la transacción.

ship-city: Ciudad de envío.

ship-state: Estado de envío.

ship-postal-code: Código postal de envío.

ship-country: País de envío.

B2B: Indica si la venta es de negocio a negocio.

Hipótesis Iniciales

Hipótesis 1: "Los pedidos cumplidos por 'Merchant' tienen una mayor tasa de cancelación en comparación con los pedidos cumplidos por 'Amazon'."

Hipótesis 2: "Los pedidos con ciertos tipos de productos (como prendas de vestir) tienen una mayor tasa de cancelación en comparación con otras categorías."

Hipótesis 3: "Los pedidos con una mayor cantidad de productos tienen una tasa de cancelación más alta"

Definición de Stakeholders Clave

Clientes Actuales: Su satisfacción es fundamental, ya que su experiencia de compra afectará las tasas de cancelación y la lealtad a la marca.

Equipo de Logística: Utilizará los hallazgos para mejorar el proceso de cumplimiento de pedidos.

Equipo de Desarrollo de Producto: Puede adaptar la oferta según las preferencias del cliente y las tendencias de cancelación.

Equipo de Marketing: Utilizará los resultados del análisis para diseñar campañas específicas que aborden la satisfacción del cliente y promuevan productos menos propensos a ser cancelados.

Preguntas Clave

- 1.- ¿Cuáles son las razones más comunes para la cancelación de pedidos?
- 2.- ¿Qué categorías de productos tienen la mayor tasa de cancelaciones?
- 3.- ¿Cómo influyen los tiempos de envío en la decisión del cliente de cancelar un pedido?
- 4.- ¿Existen patrones temporales en las cancelaciones de pedidos?
- 5.- ¿Cómo varían las tasas de cancelación según el canal de ventas?
- 6.- ¿Qué impacto tiene el tamaño del producto en la tasa de cancelación?

7.- ¿Cómo se relacionan las características demográficas de los clientes con la tasa de cancelación de pedidos?

8.- ¿Cuáles son los métodos de envío con menor tasa de cancelaciones?

9.- ¿Qué papel juegan las promociones y descuentos en la tasa de cancelación?

10.- ¿Cómo afecta la ubicación geográfica a la tasa de cancelaciones?

Fuentes de Datos Identificadas

Datos de Pedidos: Información detallada sobre cada pedido, incluyendo el estado de cumplimiento, la cantidad de productos, y el método de cumplimiento (Merchant o Amazon).

Datos de Cancelaciones: Registros específicos sobre pedidos cancelados, permitiendo un análisis profundo de los factores que influyen en este proceso.

Datos de Envío: Información sobre el estado de las entregas, lo que permite evaluar el impacto del cumplimiento en la satisfacción del cliente.

Datos Demográficos de Clientes: Información sobre las características de los clientes (ubicación, edad, etc.) que puede ayudar a segmentar los análisis y comprender mejor las tendencias de cancelación.

Justificación del Proyecto

La optimización del cumplimiento de pedidos es esencial para mejorar la satisfacción del cliente y reducir las tasas de cancelación, lo que tiene un impacto directo en la rentabilidad de la empresa. Las cancelaciones no solo generan pérdidas inmediatas en ventas, sino que también pueden dañar la reputación de la marca y afectar la lealtad del cliente a largo plazo.

Al comprender mejor los factores que contribuyen a las cancelaciones y la experiencia de cumplimiento, se pueden implementar estrategias efectivas que no solo reduzcan la tasa de cancelación, sino que también mejoren la eficiencia operativa y la satisfacción del cliente.

Esto resultará en una mejor experiencia de compra, mayor retención de clientes y, en última instancia, un aumento en las ventas.

¿Cuántos datos y que tipo son?

El conjunto de datos contiene un total de **128,976** registros, que se reparten en 18 columnas con: valores booleanos, números enteros de 64 bits y de cadenas de texto, bool(1), int64(2), object(15).