LABORATORIO 15

Parte 1

En la primera parte se aplicarán los temas vistos en la sesión como es la limpieza de los datos, se deberá realizar el paso a paso con las respetivas capturas de pantalla, esta aplicación es de acuerdo con los escenarios planteados.

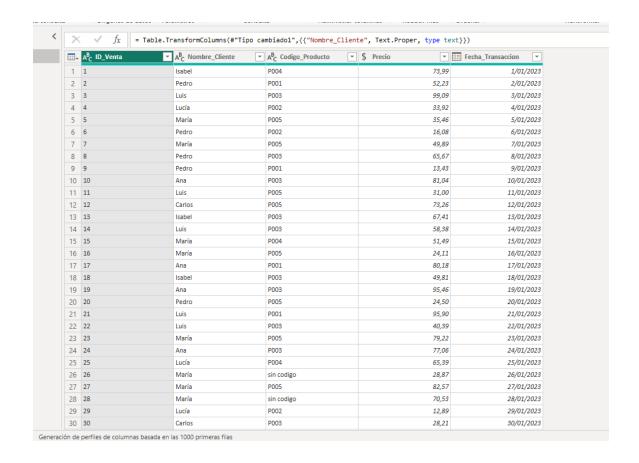
1. Ejercicio de práctica 1.

1. Escenario datos de ventas en línea

Se tiene un conjunto de datos de una empresa de ventas online que contiene información de clientes, productos vendidos y fechas de transacciones. El archivo tiene problemas comunes como valores faltantes en la columna de códigos de productos, precios con formato incorrecto y nombres de clientes en mayúsculas y minúsculas mezcladas.

Data: data_ventas_para_limpieza 1

Paso a paso para la limpieza



CONCLUSIONES

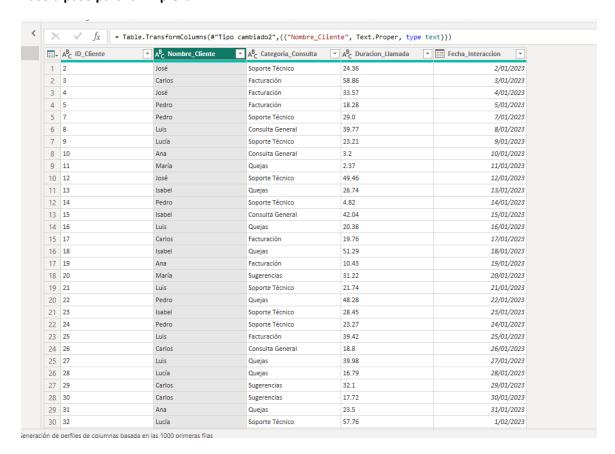
- . Manejamos valores faltantes en la columna
- .Normalizamos el valor con el formato de nombre
- . Normalizamos el formato de fecha

2. Escenario: Limpieza de datos de un servicio de atención al cliente

Trabajas en un centro de atención al cliente y tienes un conjunto de datos que contiene información sobre las interacciones con los clientes. Este dataset incluye el ID del cliente, Nombre del cliente, Categoría de consulta, Duración de la llamada y Fecha de la interacción. Algunos registros tienen problemas como nombres en minúsculas y mayúsculas inconsistentes, valores faltantes en la categoría de la consulta, y duraciones de llamada con errores de formato.

Data: data_atencion_cliente_para_limpieza 2

Paso a paso para la limpieza



CONCLUSIONES

- . En ID_Cliente lo pasamos de numero entero a texto
- . En Nombre_Cliente lo que hicimos fu tranformar los las mayusculas y minusculas que habian por "poner en mayuscula cada palabra"
- . Tambien cambiamos la duracion de llamada a numero decimal fijo
- . y por ultimo la fecha de interaccion lo pasamos de ser numero entero a FECHA

3. Escenario: Tienda minoritaria de tecnología

Trabajas para una tienda minoritaria tecnología, el equipo de inventario tiene dificultades para gestionar adecuadamente los datos de productos debido a la calidad inconsistente de

los registros. El conjunto de datos de inventario incluye información sobre ID de producto,

Nombre del producto, Cantidad disponible, Precio unitario y Fecha de ingreso al inventario.

Sin embargo, se han detectado varios problemas en los datos que impiden su uso eficaz para

análisis de ventas y reposición de stock.

Los problemas para solucionar incluyen:

• Nombres de productos faltantes.

• Precios de algunos productos faltantes.

• Cantidades de productos con texto adicional "unidades".

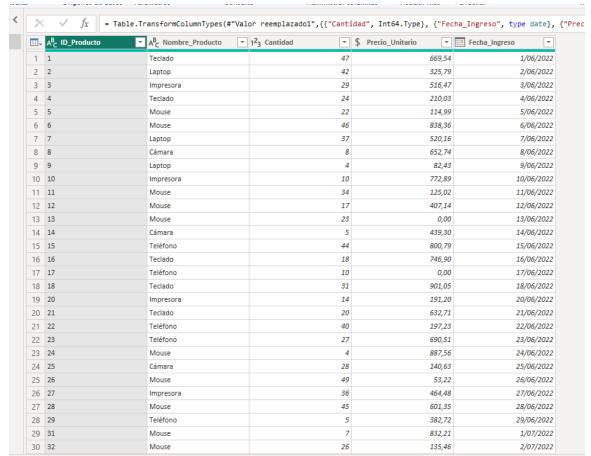
El objetivo es optimizar la calidad de los datos de inventario mediante la corrección de

errores comunes, asegurando que los registros sean consistentes, completos y adecuados

para análisis de inventarios y toma de decisiones estratégicas en la tienda.

Data: data_inventario_para_limpieza 3

Paso a paso para la limpieza



varación da porfilar da calumnar barada an las 1000 primaras filas