

## Laboratorio 15

### Sesión # 15 Limpieza con Power BI

**Título del Laboratorio:** Aplicación y uso de la herramienta de Power BI para hacer la limpieza de los datos.

**Duración:** 2 horas

**Objetivos del Laboratorio:**

1. *Afianzar los conocimientos y manejo básico en Power BI para realizar las visualizaciones con ejercicios prácticos planteados.*

**Materiales Necesarios:**

1. *Computador con acceso a internet.*
2. *Colocarlo en el repositorio de Github*
3. *Ampliar el conocimiento con el curso de datos en AWS y Cisco.*
4. *Power BI descargarlo*

**Estructura del Laboratorio:**

#### 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 respectivas capturas de pantalla, esta aplicación es de acuerdo con los escenarios planteados.

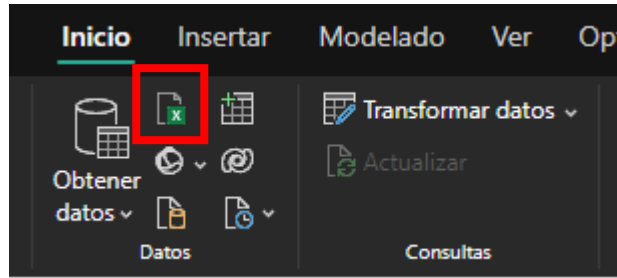
#### 1. Ejercicio de práctica 1.

##### 1. Escenario datos de ventas en línea

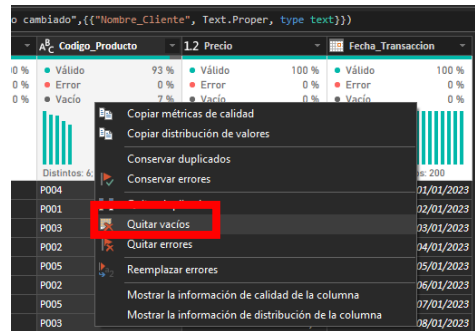
Se tiene un conjunto de datos de una empresa de ventas online que contiene información de clientes, productos vendidos y fechas de transacciones. El archivo tiene problemas comunes como valores faltantes en la columna de códigos de productos, precios con formato incorrecto y nombres de clientes en mayúsculas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

**Data:** data\_ventas\_para\_limpieza 1

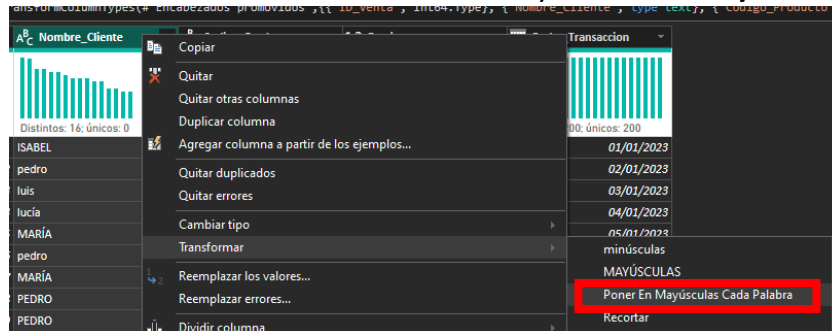
- Primero se deben importar los datos de las tres tablas seleccionando el botón “Importar datos de un libro de Microsoft Excel”.



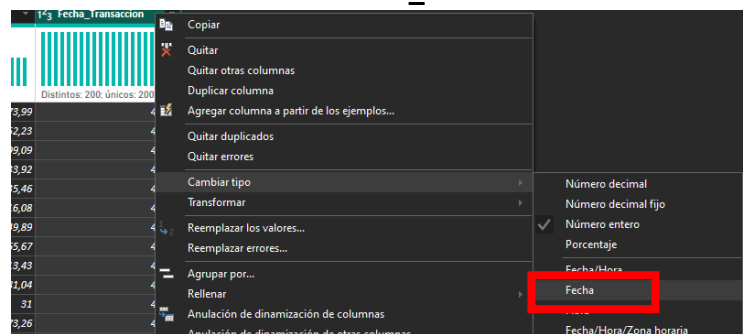
- Quitar los registros de ventas con valores “null” en la columna “Producto”.



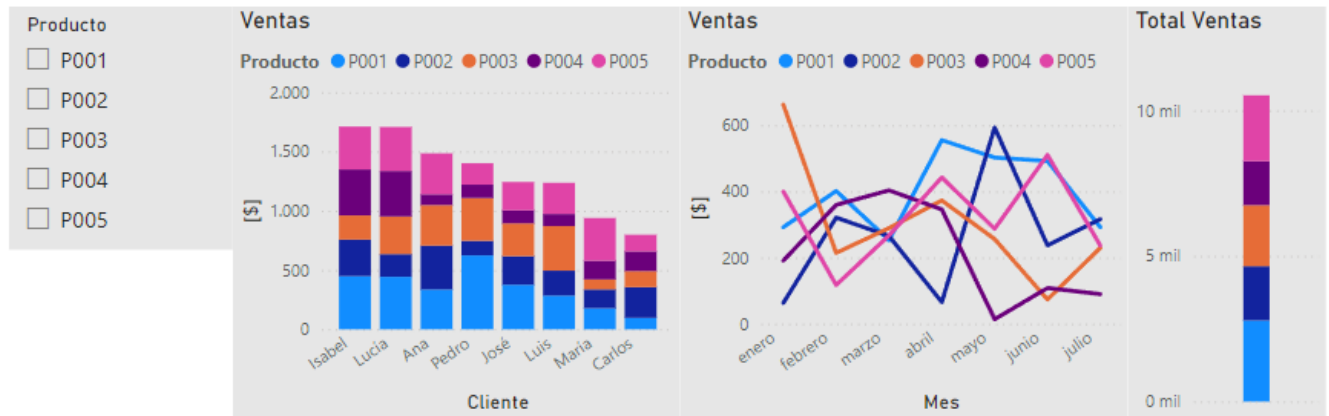
- Capitalizar los nombres mediante “Transformar” y “Poner en Mayúsculas Cada Palabra”.



- Cambiar el tipo de dato de la columna “Fecha\_Transacción” a “Fecha”.



- Después se pueden insertar las visualizaciones que ayuden a mostrar las medidas solicitadas.



## CONCLUSIONES

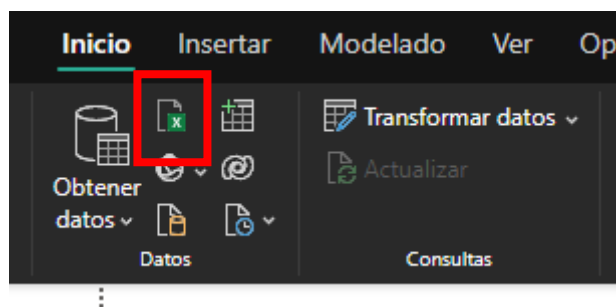
El mejor cliente es “Isabel” y el mejor producto es el “P001” con un total de ventas de **\$2.787,63**.

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

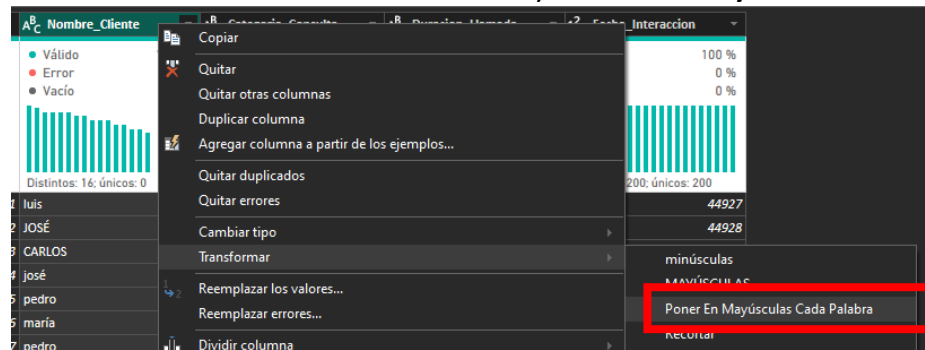
Trabajas en un centro de atención al cliente y tienes un conjunto de datos que contiene información sobre las interacciones con los clientes. Este dataset incluye el ID del cliente, Nombre del cliente, Categoría de consulta, Duración de la llamada y Fecha de la interacción. Algunos registros tienen problemas 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

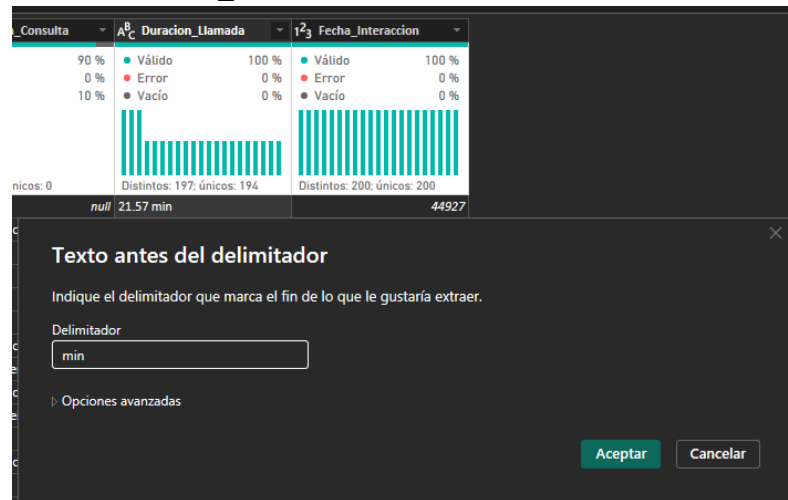
- Primero se deben importar los datos de las tres tablas seleccionando el botón “Importar datos de un libro de Microsoft Excel”.



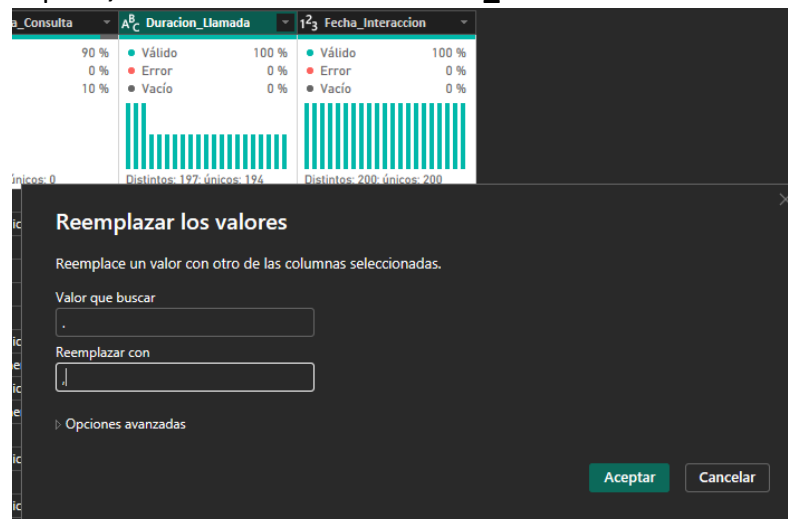
- Capitalizar los nombres mediante **“Transformar”** y **“Poner en Mayúsculas Cada Palabra”**.



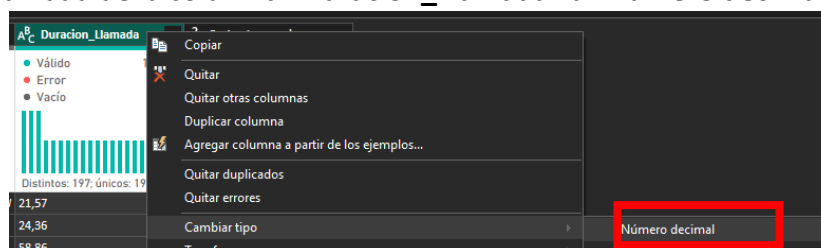
- Extraer el texto de **“Duración\_Llamada”** antes de **“ min”**.



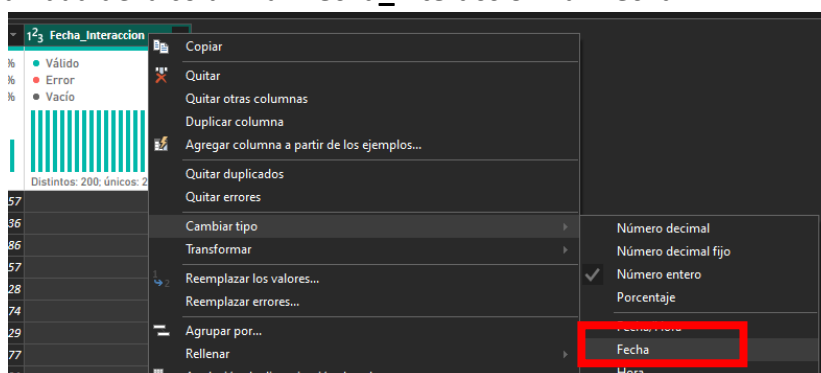
- Reemplazar **“.”** por **“,”** en el texto de la **“Duración\_Llamada”**.



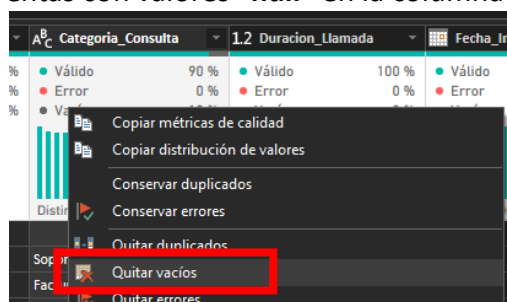
- Cambiar la unidad de la columna “Duración\_Llamada” a “Número decimal”.



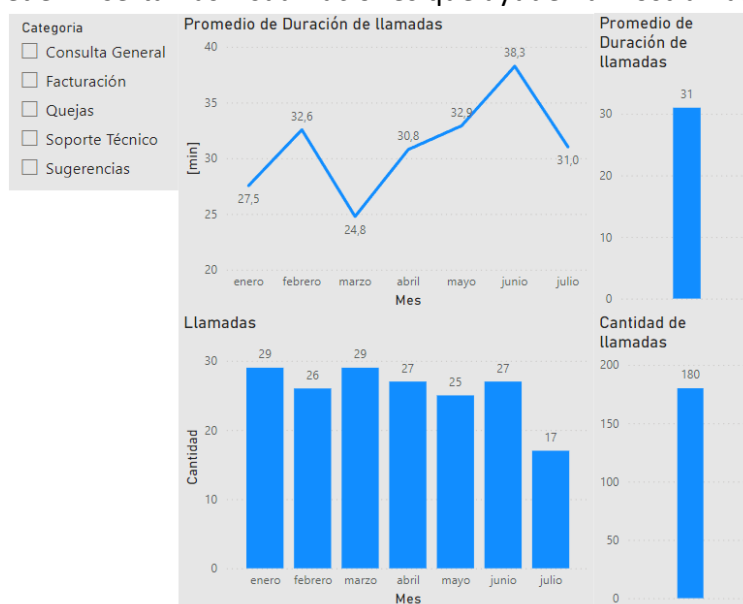
- Cambiar la unidad de la columna “Fecha\_Interacción” a “Fecha”.



- Quitar los registros de ventas con valores “null” en la columna “Categoría\_Consulta”.



- Después se pueden insertar las visualizaciones que ayuden a mostrar las medidas solicitadas.



## CONCLUSIONES

La categoría “**Facturación**” tiene menor promedio de duración de llamadas de “**24.72 min**”. “**Soporte Técnico**” es la categoría que más llamadas atiende.

### 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.

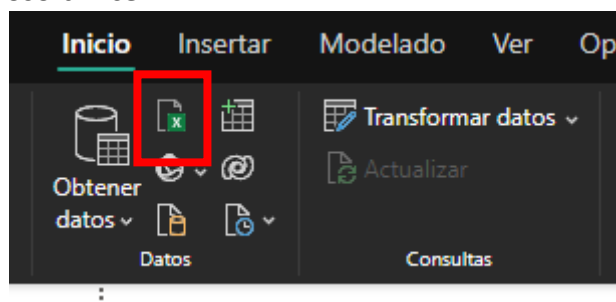
Los problemas para solucionar incluyen:

- Nombres de productos faltantes.
- Precios de algunos productos faltantes.
- Cantidades de productos con texto adicional "unidades".

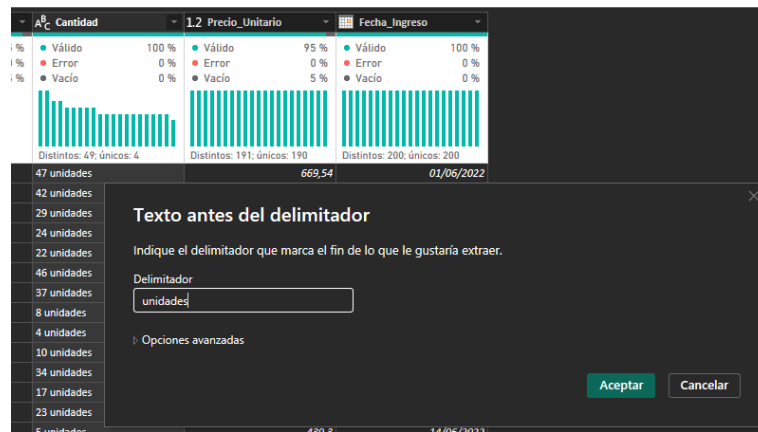
El objetivo es optimizar la calidad de los datos de inventario mediante la corrección de errores comunes, asegurando que los registros sean consistentes, completos y adecuados para análisis de inventarios y toma de decisiones estratégicas en la tienda.

Data: data\_inventario\_para\_limpieza 3

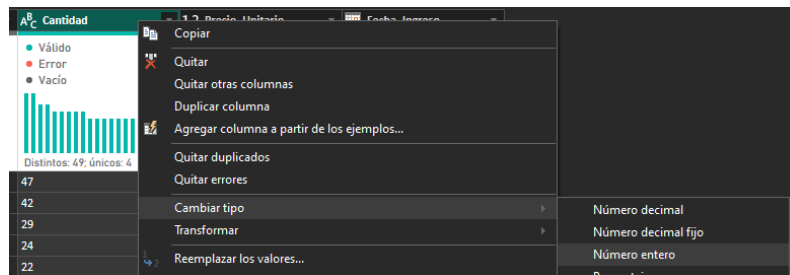
- Primero se deben importar los datos de las tres tablas seleccionando el botón “**Importar datos de un libro de Microsoft Excel**”.



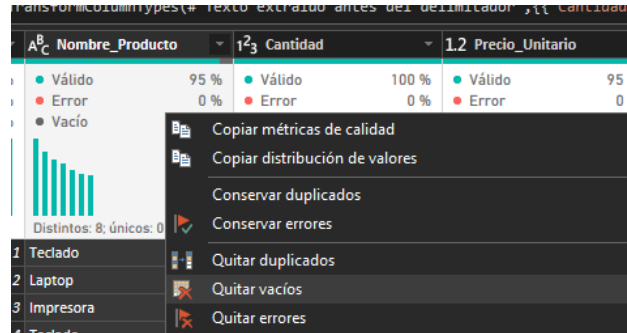
- Extraer el texto de “Cantidad” antes de “ unidades”.



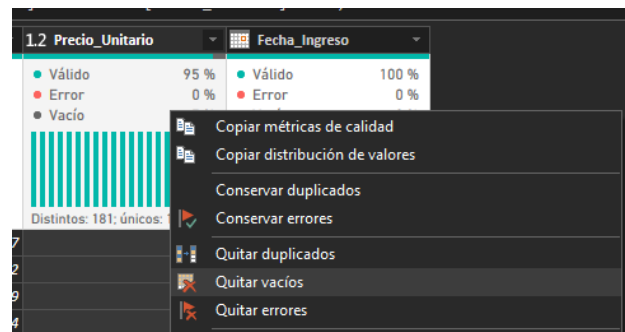
- Cambiar la unidad de la columna “Cantidad” a “Número entero”.



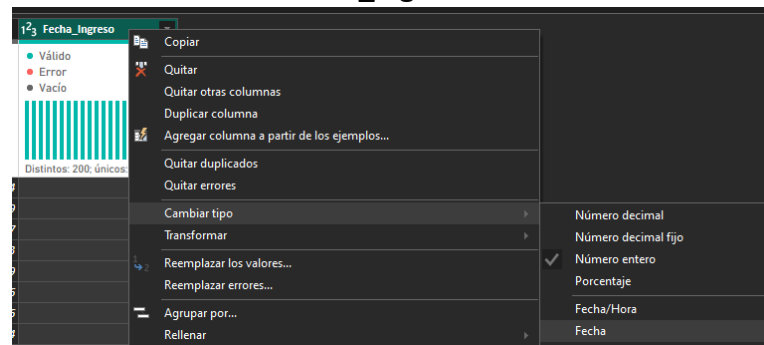
- Quitar los registros de ventas con valores “null” en la columna “Nombre\_Producto”.



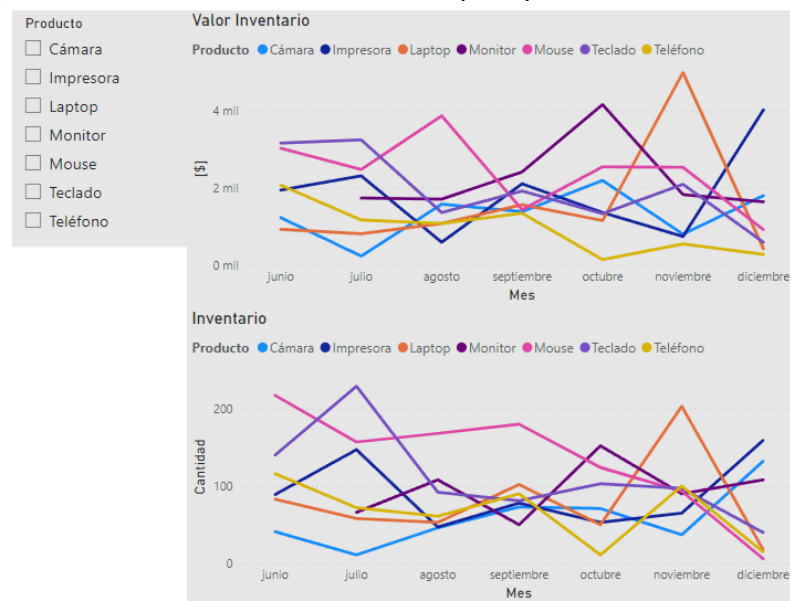
- Quitar los registros de ventas con valores “null” en la columna “Precio\_Unitario”.



- Cambiar la unidad de la columna “Fecha\_Ingreso” a “Fecha”.



- Después se pueden insertar las visualizaciones que ayuden a mostrar las medidas solicitadas.



## CONCLUSIONES

En las gráficas se puede observar que el producto “**Laptop**” tuvo un incremento muy alto del valor del inventario en el mes de “**noviembre**”, se debe evaluar porqué están por fuera de su valor normal de compra.