

TechCore

Nombre del autor: Zambón Enzo Agustín

Email: eazambon97@gmail.com

Cohorte: PT01

Fecha de entrega: 12/11/2025

Institución:

TechoCore es una cadena de tiendas minoristas fundada en 2014 y especializada en la venta y distribución de computadores, equipos electrónicos y accesorios tecnológicos. Con operaciones en Colombia, la empresa se ha consolidado como un referente del sector gracias a su amplio portafolio de productos, su enfoque en la innovación y su capacidad para atender tanto a consumidores individuales como a clientes corporativos. Su trayectoria y compromiso con la tecnología accesible la posicionan como un actor clave en un mercado en constante evolución.

Módulo 4**Introducción**

El presente proyecto tiene como objetivo transformar una base de datos cruda de ventas en un sistema analítico integral que permita comprender el desempeño comercial de una cadena de tiendas minoristas dedicada a la venta de computadores y accesorios tecnológicos. A partir de la limpieza, modelado y análisis de los datos, se busca proporcionar información confiable, estructurada y visualmente accesible para apoyar la toma de decisiones estratégicas de la empresa.

El trabajo se desarrolló en tres etapas principales. En la primera, se realizó la carga, depuración y estandarización de los datos originales, garantizando su calidad y preparándolos para posteriores análisis. En la segunda etapa, se diseñó y construyó un modelo relacional que integra las distintas entidades del negocio —ventas, productos, clientes, ubicaciones y vendedores— asegurando coherencia e integridad en toda la información. Finalmente, en la tercera etapa, se elaboró un dashboard interactivo en Power BI, orientado a responder preguntas clave del negocio, identificar patrones relevantes y exponer hallazgos que permitan evaluar el rendimiento comercial y detectar oportunidades de mejora.

El resultado final es una herramienta analítica robusta y escalable, capaz de transformar datos dispersos en conocimiento accionable, ofreciendo una visión clara del comportamiento de las ventas, las dinámicas de los clientes y las tendencias del mercado.

Módulo 4

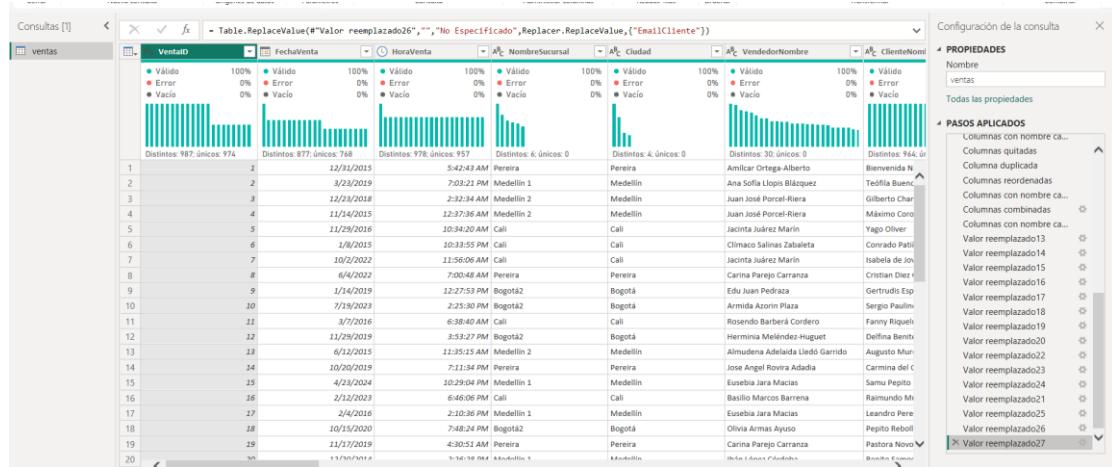
Desarrollo del proyecto:**Avance 1: Carga y transformación de los datos**

El objetivo de esta etapa fue importar, limpiar y preparar la base de datos cruda de facturación en Power BI, dejándola lista para los siguientes análisis y modelados.

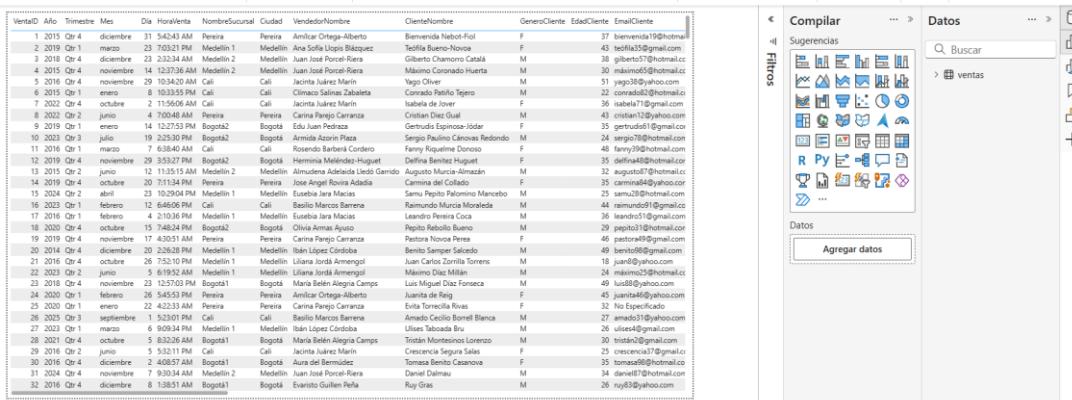
Entre los pasos realizados se destaca:

1. Encabezados promovidos
2. Tipos de datos de columnas cambiados
3. Se reemplazaron los valores nulos por cero
4. Se uso un único criterio para los nombres de las columnas
5. Se reemplazaron valores en varias de las columnas para unificar los valores de las clasificaciones, entre otros.

En total fueron aplicados 43 pasos en el proceso de limpieza de la base de origen “ventas.csv”



Posteriormente, se procedió a crear un objeto visual con forma de tabla para importar los datos en un archivo nuevo: “VentasTransformed”



Módulo 4**Avance 2: Diseño del Modelo Relacional**

En esta nueva instancia, el objetivo fue diseñar e implementar el modelo relacional que permitirá integrar la información de ventas con las demás entidades del negocio —productos, clientes, sucursales y vendedores— asegurando coherencia, integridad y trazabilidad de los datos. Para ello, se cargo en python el dataset limpio VentasTransformed.csv y se construyeron las tablas que conforman la estructura del modelo: Facturas, DetalleFacturas, Productos, Clientes, Sucursales, Ciudades y Vendedores.

```

> <ipython>
    import pandas as pd
    # Cargar el archivo
    df = pd.read_csv('C:/Users/Enzo/OneDrive/Escritorio/Data Science Henry/Módulo 3/Avance 1/VentasTransformed.csv')
[1]

```

	# VentaID	# Año	Δ Trimestre	Δ Mes	# Dia	Δ HoraVenta	Δ NombreSucursal
0	1	2015	Otr 4	diciembre	31	1899-12-30 05:42:43	Pereira
1	2	2019	Otr 1	marzo	23	1899-12-30 19:02:21	Medellin 1
2	3	2018	Otr 4	diciembre	23	1899-12-30 02:32:34	Medellin 2
3	4	2015	Otr 4	noviembre	14	1899-12-30 00:37:36	Medellin 2
4	5	2016	Otr 4	noviembre	29	1899-12-30 10:42:40	Cali
5	6	2015	Otr 1	enero	8	1899-12-30 22:35:55	Cali
6	7	2022	Otr 4	octubre	2	1899-12-30 11:56:06	Cali
7	8	2022	Otr 2	junio	4	1899-12-30 07:00:46	Pereira
8	9	2019	Otr 1	enero	14	1899-12-30 12:27:53	Bogota2
9	10	2023	Otr 3	julio	19	1899-12-30 14:25:30	Bogota2

30,000 rows x 33 cols 10 per page << Page 1 of 3000 >> ...

```

# --- TABLA: CIUDADES ---
dim_ciudad = df[['Ciudad']].drop_duplicates().reset_index(drop=True)
dim_ciudad['CiudadID'] = dim_ciudad.index + 1

# --- TABLA: SUCURSALES ---
dim_sucursal = df[['NombreSucursal', 'Ciudad']].drop_duplicates().reset_index(drop=True)
dim_sucursal = dim_sucursal.merge(dim_ciudad, on='Ciudad', how='left')
dim_sucursal['SucursalID'] = dim_sucursal.index + 1

# --- TABLA: VENDEDORES ---
dim_vendedor = df[['VendedorNombre']].drop_duplicates().reset_index(drop=True)
dim_vendedor['VendedorID'] = dim_vendedor.index + 1

# --- TABLA: CLIENTES ---
dim_cliente = df[['ClienteNombre', 'GeneroCliente', 'EdadCliente', 'EmailCliente',
                  'TelefonoCliente', 'DireccionCliente']].drop_duplicates().reset_index(drop=True)
dim_cliente['ClienteID'] = dim_cliente.index + 1

# --- TABLA: MÉTODO DE PAGO ---
dim_metodo = df[['MetodoPago']].drop_duplicates().reset_index(drop=True)
dim_metodo['MetodoPagoID'] = dim_metodo.index + 1

# --- TABLA: PRODUCTOS ---
# Reorganizamos los productos (1, 2 y 3)
productos = []
for i in range(1, 4):
    temp = df[[f'NombreProducto{i}', f'MarcaProducto{i}', f'PrecioUnitarioProducto{i}']].drop_duplicates()
    temp.columns = ['NombreProducto', 'MarcaProducto', 'PrecioUnitario']
    productos.append(temp)

dim_producto = pd.concat(productos).drop_duplicates().reset_index(drop=True)
dim_producto['ProductoID'] = dim_producto.index + 1

```

Módulo 4

```

# --- TABLA: FACTURAS ---
facturas_cols = [
    'VentaID', 'Año', 'Trimestre', 'Mes', 'Dia', 'HoraVenta',
    'NombreSucursal', 'VendedorNombre', 'ClienteNombre', 'MetodoPago',
    'DescuentoVenta', 'TotalVenta'
]

facturas = df[facturas_cols].drop_duplicates().reset_index(drop=True)

# --- Vinculamos claves foráneas ---
facturas = (
    facturas
        .merge(dim_sucursal[['NombreSucursal', 'SucursalID']], on='NombreSucursal', how='left')
        .merge(dim_vendedor[['VendedorNombre', 'VendedorID']], on='VendedorNombre', how='left')
        .merge(dim_cliente[['ClienteNombre', 'ClienteID']], on='ClienteNombre', how='left')
        .merge(dim_metodo[['MetodoPago', 'MetodoPagoID']], on='MetodoPago', how='left')
)

# --- Eliminamos las columnas originales de texto ---
facturas = facturas.drop(columns=['NombreSucursal', 'VendedorNombre', 'ClienteNombre', 'MetodoPago'])

# --- Generamos la clave primaria ---
facturas['FacturaID'] = facturas.index + 1

# --- Reordenamos columnas ---
facturas = facturas[[[
    'FacturaID', 'VentaID', 'Año', 'Trimestre', 'Mes', 'Dia', 'HoraVenta',
    'SucursalID', 'VendedorID', 'ClienteID', 'MetodoPagoID',
    'DescuentoVenta', 'TotalVenta'
]]]

# --- TABLA: DETALLE FACTURAS ---
detalles = []
for i in range(1, 4):
    temp = df[['VentaID', f'NombreProducto{i}', f'CantidadProducto{i}'],
               f'PrecioUnitarioProducto{i}', f'SubtotalProducto{i}']].copy()
    temp.columns = ['VentaID', 'NombreProducto', 'Cantidad', 'PrecioUnitario', 'Subtotal']
    detalles.append(temp)

detalle_facturas = pd.concat(detalles).dropna(subset=['NombreProducto'])

detalle_facturas = detalle_facturas.merge(facturas[['VentaID', 'FacturaID']], on='VentaID', how='left')
detalle_facturas = detalle_facturas.merge(dim_producto[['NombreProducto', 'ProductoID']], on='NombreProducto', how='left')
detalle_facturas = detalle_facturas[['FacturaID', 'ProductoID', 'Cantidad', 'PrecioUnitario', 'Subtotal']]

```

Finalmente, se exportaron las tablas en un único Excel:

```

# --- EXPORTAR TODO A UN SOLO ARCHIVO EXCEL ---
with pd.ExcelWriter('Modelo_Ventas.xlsx', engine='openpyxl') as writer:
    dim_producto.to_excel(writer, sheet_name='Productos', index=False)
    detalle_facturas.to_excel(writer, sheet_name='DetalleFacturas', index=False)
    dim_cliente.to_excel(writer, sheet_name='Clientes', index=False)
    dim_ciudad.to_excel(writer, sheet_name='Ciudades', index=False)
    dim_sucursal.to_excel(writer, sheet_name='Sucursales', index=False)
    facturas.to_excel(writer, sheet_name='Facturas', index=False)
    dim_vendedor.to_excel(writer, sheet_name='Vendedores', index=False)
    dim_metodo.to_excel(writer, sheet_name='MetodoPago', index=False)

print("✅ Archivo 'Modelo_Ventas.xlsx' generado con éxito.")

[30]
...
...   ✅ Archivo 'Modelo_Ventas.xlsx' generado con éxito.

```

Tambien se realizó una verificación de consistencia y se efectuaron consultas para verificar la calidad de los datos.

Módulo 4

```
# --- TOTAL DE VENTAS POR MARCA ---

ventas_por_marca = (
    detalle_facturas
    .merge(dim_producto[['ProductoID', 'MarcaProducto']], on='ProductoID', how='left')
    .groupby('MarcaProducto', as_index=False)[['Subtotal']]
    .sum()
    .sort_values(by='Subtotal', ascending=False)
)

print("💡 Total de ventas por marca:")
print(ventas_por_marca.head(10))

]

💡 Total de ventas por marca:
   MarcaProducto      Subtotal
5       Lenovo  117915200000
4        HP    96843000000
3       Dell   86758400000
1      Apple   78021700000
2      Asus   33192000000
0      Acer   22011000000
9     Samsung  12760800000
6       MSI   8525200000
7  Microsoft  8223400000
8      Razer  3936400000
```



```
# --- TOP 10 PRODUCTOS MÁS VENDIDOS (por cantidad) ---

top_productos = (
    detalle_facturas
    .merge(dim_producto[['ProductoID', 'NombreProducto']], on='ProductoID', how='left')
    .groupby('NombreProducto', as_index=False)[['Cantidad']]
    .sum()
    .sort_values(by='Cantidad', ascending=False)
    .head(10)
)

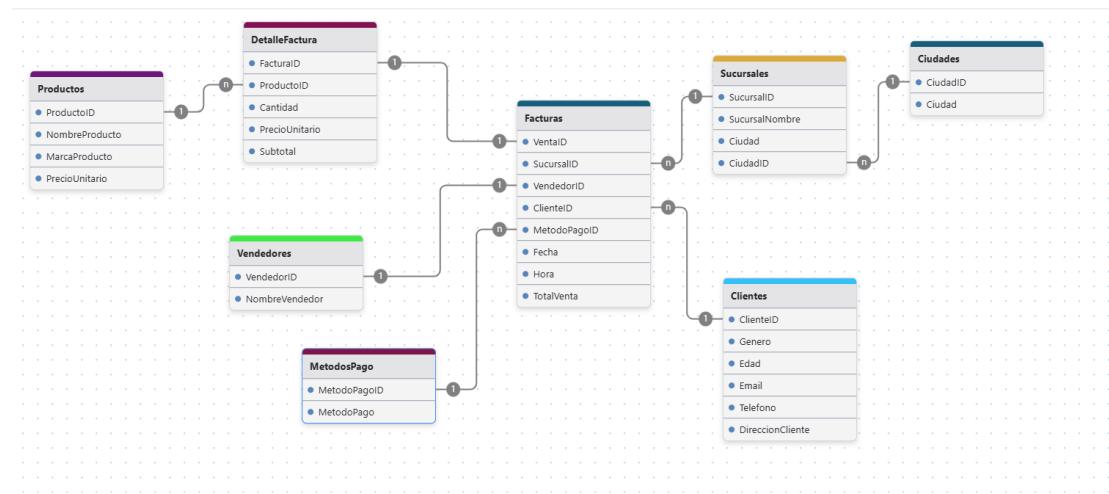
print("🏆 Top 10 productos más vendidos:")
print(top_productos)

]

🏆 Top 10 productos más vendidos:
   NombreProducto  Cantidad
19      HP Spectre x360    7730
21      Lenovo Legion 5 Pro   5864
22  Lenovo ThinkPad X1 Carbon   5814
23      Lenovo Yoga 7i    5054
17      HP Omen 16    5015
20      Lenovo IdeaPad 5    4972
15      Dell XPS 13    4530
14      Dell Latitude 7420   4509
18      HP Pavilion 15    4030
12      Dell Alienware m15   3658
```

El modelo relacional construido es el siguiente:

Módulo 4

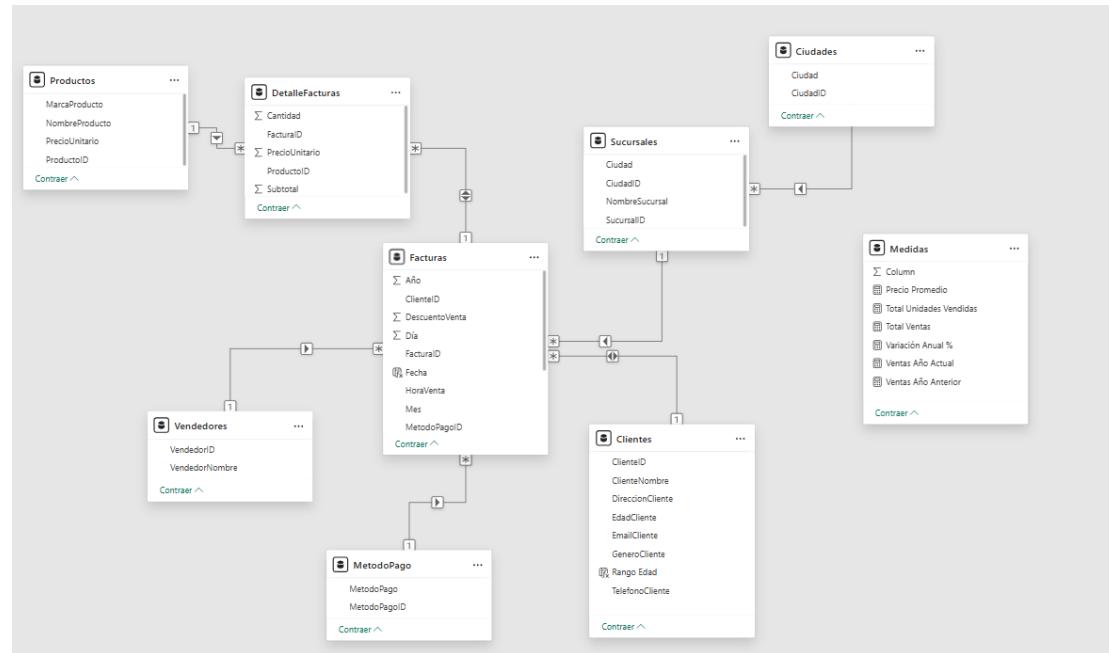


Módulo 4

Avance 3: Desarrollo del Dashboard Interactivo

La ultima etapa consistió en importar el modelo relacional generado en Python dentro de Power BI y diseñar el Dashboard permitiendo transformar los datos en información estratégica para la toma de decisiones.

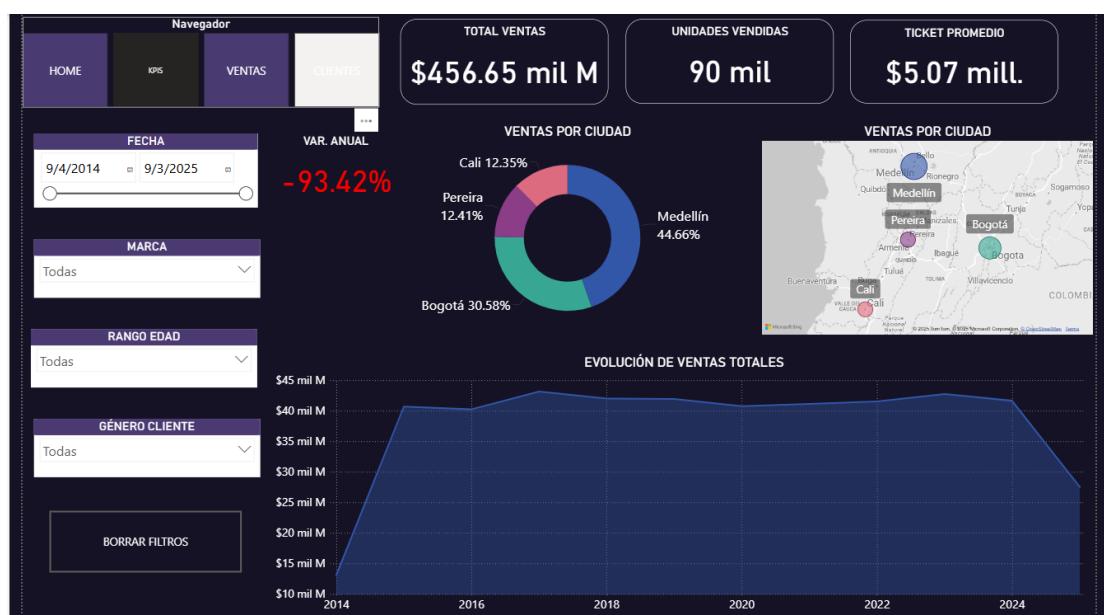
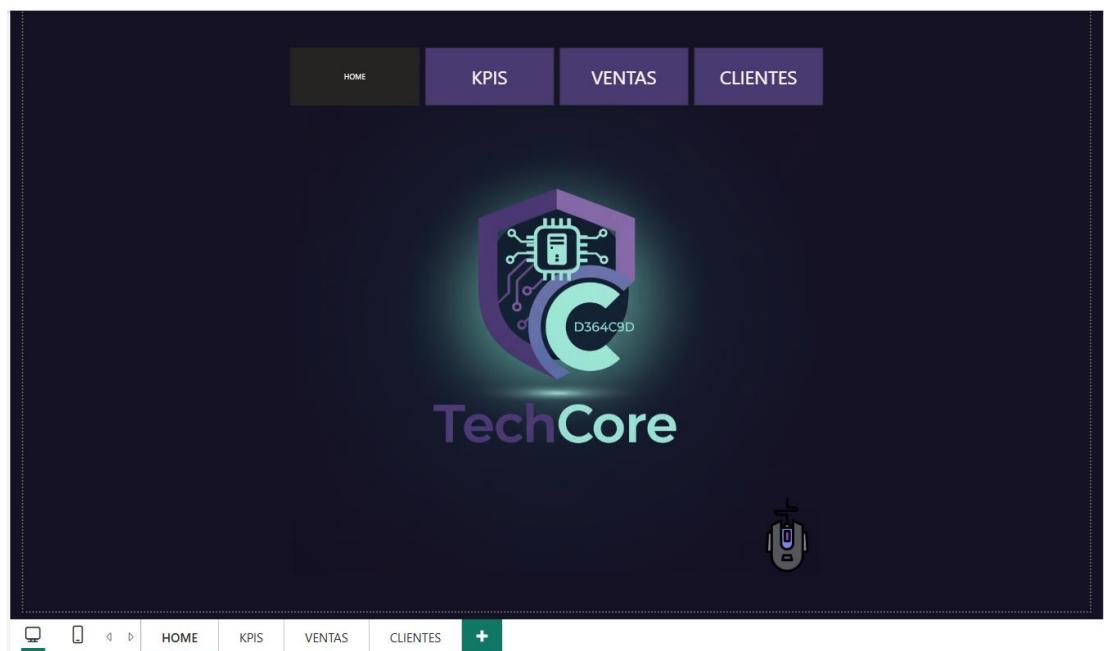
El primer paso fue el de importar el archivo ModeloVentas.xlsx y vincular las tablas para crear el modelo relacional imaginado



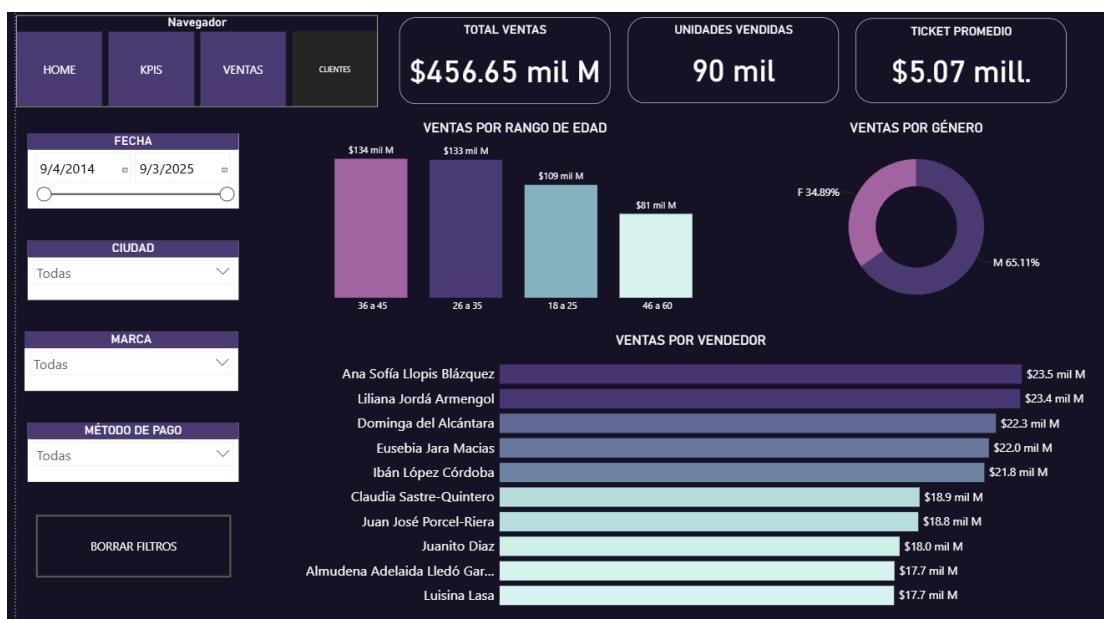
También se crearon las siguientes medidas DAX: Total Ventas, Total Unidades Vendidas, Precio Promedio, Ventas Año Actual, Ventas Año Anterior y Variación % Anual, para visualizar información relevante del negocio.

En cuanto al diseño del Dashboard, este cuenta con una portada de inicio o "Home" con el logo de la empresa, y tres hojas con información

Módulo 4



Módulo 4



Principales Hallazgos:

- La empresa acumula un total de ventas de \$456.65 mil millones de pesos, y 90 mil unidades vendidas.
- El precio promedio de venta es de \$5.07 millones
- En el ultimo año, la empresa registra una caída del 93% en las ventas
- El año con mayor nivel de ventas fue el 2017 con un total de \$43 mil millones
- Medellín (45%) y Bogotá (30%) son las ciudades que mas volumen de ventas abarcan.
- El 40% de los pagos se realizó con tarjeta de crédito, el 20% con débito y el 15% con transferencia.
- Las marcas mas vendidas son Lenovo, HP, Dell y Apple

Módulo 4

- El rango de edad de entre 26 y 45 años representa el 60% de las ventas.
- La proporción de hombres en las ventas es del 65%

Recomendaciones:

1. Revertir la caída del 93% en el último año analizando posibles causas internas: ruptura en la cadena de suministros, aumento de precios, rotación del personal de ventas, baja disponibilidad de stock, etc.
2. Incrementar inversión en Medellín y Bogotá
3. Alianza con marcas de alta rotación (Lenovo, HP, Dell, Apple)
4. Promociones temporales para captar segmentos sensibles al precio.
5. Estrategia de financiamiento: cuotas sin interés para mejorar la conversión, especialmente con tarjeta de crédito (40% del total).
6. Incentivar tarjeta de débito y transferencia con descuentos (bajos costos para la empresa).
7. Ofrecer bonificaciones con tarjeta de crédito en temporadas altas (back to school, día del padre, navidad).
8. Ofrecer beneficios exclusivos para profesionales y estudiantes universitarios (ej.: garantía extendida, financiamiento preferencial).
9. Promociones corporativas orientadas a PyMEs, donde el rango etario de 26 a 45 tiene alta presencia.
10. Diseñar campañas orientadas a mujeres, especialmente enfocadas en diseño, peso del equipo y facilidad de uso.
11. Ofrecer talleres gratuitos (online o presenciales) sobre productividad y tecnología para atraer un público más diverso.

Conclusión:

El análisis permitió transformar la información de ventas en insights clave para comprender el desempeño del negocio y detectar oportunidades de mejora. A través del proceso de limpieza, modelado y visualización en Power BI, se identificaron patrones relevantes en las ventas, el comportamiento de los clientes y la participación por ciudades y marcas. Estos hallazgos sirven como base para orientar decisiones estratégicas que permitan recuperar el nivel de ventas, optimizar recursos y enfocar esfuerzos en los segmentos y regiones con mayor potencial.