

Informe de Laboratorio 2 - Análisis de Ventas con Power BI

ALUMNA: Gina Grosso.

1. Proceso ETL (Extracción, Transformación y Carga)

Para el desarrollo de este informe se llevó a cabo un proceso integral de ETL utilizando Power Query en Power BI Desktop.

- **Extracción:** Se conectaron cuatro fuentes de datos en formato CSV provenientes de un dataset de ventas regional:
 - *orders* (Tabla de Hechos: Transacciones de venta).
 - *accounts* (Dimensión: Información de clientes).
 - *sales_reps* (Dimensión: Representantes de ventas).
 - *region* (Dimensión: Regiones geográficas).
- **Transformación y Limpieza:** Se aplicaron transformaciones significativas para asegurar la calidad y utilidad de los datos:
 - **Normalización de Texto:** En la tabla *accounts*, se dividió la columna de contacto principal (*primary_poc*) para separar el Nombre y el Apellido en columnas independientes.
 - **Corrección de Formatos Numéricos:** Debido a conflictos de configuración regional, se realizó una corrección matemática en las columnas de coordenadas (*lat*, *long*) y montos monetarios (*total_amt_usd*), aplicando una división estándar para ajustar la escala de los valores decimales.
 - **Limpieza de Fechas:** Se transformó la columna de fecha y hora de la tabla *orders* para conservar únicamente la fecha, facilitando la relación temporal.
 - **Enriquecimiento de Datos:** Se creó una columna condicional llamada "Tamaño Pedido" en la tabla *orders*. Esta columna segmenta las transacciones en "Pequeño", "Mediano" o "Grande" basándose en la cantidad total de unidades vendidas (*total*), permitiendo un filtrado estratégico en el reporte.

2. Modelado de Datos

Se diseñó un Modelo de Datos Relacional siguiendo un esquema de **Copo de Nieve (Snowflake Schema)** para optimizar el rendimiento y la integridad de los filtros.

- **Relaciones:** Se establecieron relaciones de "Uno a Muchos" (1:*) fluyendo desde las tablas dimensionales hacia la tabla de hechos, respetando la jerarquía: *Region* → *Sales Reps* → *Accounts* → *Orders*.
- **Inteligencia de Tiempo:** Se creó una tabla *Calendario* dedicada utilizando fórmulas DAX (*CALENDAR*, *MIN*, *MAX*). Esta tabla se marcó oficialmente como "Tabla de fechas" y se vinculó a la tabla de hechos, habilitando un análisis temporal preciso y escalable.

3. Medidas DAX y Visualización

El análisis cuantitativo se basó en **Medidas Explícitas** creadas con DAX para asegurar cálculos dinámicos según el contexto del filtro:

- **KPIs Principales:**
 - *Total Ventas USD* = SUM(orders[*total_amt_usd*])
 - *Cantidad de Pedidos* = COUNTROWS(*orders*)
 - *Ticket Promedio* = DIVIDE([*Total Ventas USD*], [*Cantidad de Pedidos*])

Diseño del Dashboard:

Se construyó un panel de control interactivo que incluye:

1. **Mapa de Burbujas:** Visualización geoespacial de las ventas según la latitud y longitud de los clientes.
2. **Gráfico de Barras:** Comparativa de rendimiento de ventas por Región.
3. **Gráfico de Líneas:** Análisis de tendencia de ventas a lo largo del tiempo (Año/Mes).
4. **Tarjetas KPI:** Indicadores numéricos de alto nivel para una lectura rápida del estado del negocio.
5. **Segmentación (Filtros):** Se incorporó un panel de segmentadores que permite filtrar el informe por dos dimensiones clave: '**Tamaño del Pedido**' (Pequeño/Mediano/Grande) y '**Región**' (Northeast, West, etc.). Esto habilita una interactividad completa para analizar el desempeño tanto por volumen de venta como por zona geográfica.