



**Universidad de Oriente**

**Facultad de Ingeniería y Arquitectura**

**Carrera:**

INGENIERIA EN DESARROLLO DE SOFTWARE

**Cátedra:**

DESARROLLO DE FUNCIONES AVANZADAS DE BASES DE DATOS

**Docente:**

Ing. CARLOS ISAIAS RIOS LUNA

**Actividad:**

Avance del Proyecto - Lab III

**LINK's:**

<https://drive.google.com/drive/folders/1e2nvAZ0wV9Ug-mM0NRsMQ7GHCjyRuhs6?usp=sharing>

**Estudiante:**

MEJIA MOLINA, RODRIGO ALEXANDER

U20220069

## **1. Introducción**

En el contexto del análisis de datos empresariales, el presente proyecto tiene como objetivo principal aplicar técnicas de minería de datos sobre la base de datos ContosoRetailDW, desarrollando un dashboard en Power BI Desktop que brinde información útil para la toma de decisiones. A lo largo de este documento, se definirá el enfoque del proyecto, partiendo de procesos ETL, modelado de cubos OLAP, diseño de reportes y prototipado de un dashboard funcional.

## **2. Lógica del problema**

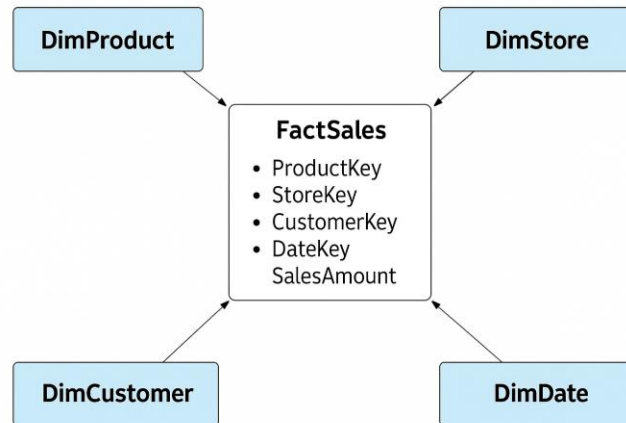
La base de datos ContosoRetailDW representa un escenario empresarial del sector retail, almacenando información sobre productos, ventas, territorios, clientes y más. El objetivo del proyecto es identificar los productos más vendidos por categoría y región, permitiendo a los responsables de ventas y marketing tomar decisiones informadas sobre qué líneas de productos potenciar según la demanda en diferentes zonas geográficas.

En este contexto, se puede aplicar minería de datos para detectar patrones de compra, segmentar mercados y optimizar la estrategia comercial.

## **3. Diagrama relacional de la base de datos**

La base de datos ContosoRetailDW está organizada bajo un modelo de esquema en estrella (star schema), una estructura típica en almacenes de datos que facilita el análisis multidimensional. En este modelo, la tabla de hechos FactSales ocupa el centro y se relaciona directamente con diversas tablas de dimensiones que describen los diferentes contextos del hecho de venta.

Figura 1. Diagrama relacional de la base de datos ContosoRetailDW



#### **Lógica del diagrama:**

- **FactSales:** Es la tabla principal de hechos. Contiene los registros detallados de cada transacción de venta, como el monto, el producto, la tienda y la fecha.
- **DimProduct:** Se vincula con FactSales a través del campo ProductKey, proporcionando información del producto vendido.
- **DimProductSubcategory** y **DimProductCategory:** Amplían la jerarquía de los productos permitiendo agruparlos por subcategorías y categorías generales.
- **DimSalesTerritory:** Se conecta con FactSales mediante SalesTerritoryKey, permitiendo segmentar las ventas por territorio comercial.
- **DimGeography:** Está relacionada con DimSalesTerritory a través de GeographyKey, lo que permite analizar las ventas según regiones geográficas como países o ciudades (RegionCountryName).
- **DimDate:** Proporciona el componente temporal del análisis, facilitando segmentaciones por año, trimestre, mes o día.

Este diseño permite realizar consultas rápidas y eficientes sobre grandes volúmenes de datos, y es ideal para la generación de cubos OLAP y dashboards interactivos en herramientas como Power BI.

#### **4. Tablas, vistas y procesos a utilizar**

A continuación se detallan las tablas y campos clave seleccionados para el desarrollo del proyecto de minería de datos, los cuales permiten realizar análisis de ventas por categoría de producto, región y período de tiempo:

### **Tablas utilizadas:**

- FactSales: Tabla de hechos que almacena información detallada de cada transacción de venta.
- DimProduct: Contiene los detalles de cada producto.
- DimProductCategory: Define las categorías generales de productos.
- DimProductSubcategory: Clasifica los productos en subcategorías dentro de una categoría.
- DimSalesTerritory: Proporciona información sobre las zonas de venta asignadas.
- DimGeography: Contiene datos geográficos relevantes, como país, ciudad o región.
- DimDate: Incluye la dimensión de tiempo asociada a cada venta.

### **Campos principales utilizados para el análisis:**

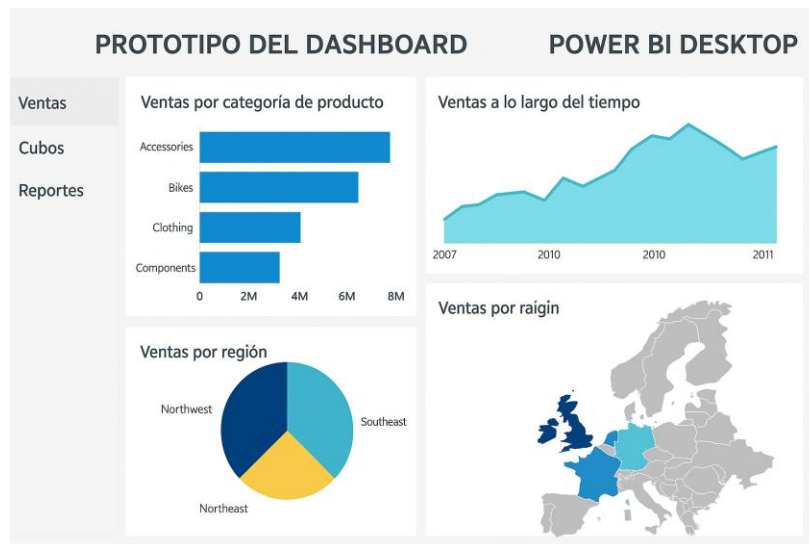
- SalesAmount (FactSales): Monto total de la venta realizada.
- ProductCategoryName (DimProductCategory): Nombre de la categoría del producto vendido.
- RegionCountryName (DimGeography): País o región donde ocurrió la venta.
- SalesTerritoryKey (FactSales, DimSalesTerritory): Clave que relaciona una venta con una zona de ventas.
- GeographyKey (DimSalesTerritory, DimGeography): Clave que vincula una zona de ventas con su información geográfica.

## **5. Prototipo del Dashboard**

El dashboard final muestra:

- Gráfico de columnas: Total de ventas por categoría de producto.
- Mapa geográfico: Ventas totales por país o región (RegionCountryName).
- Segmentadores: Por año, categoría, subcategoría y región.
- Tarjetas de KPIs: Venta total, categoría con mayor venta, región con mayor venta.

Figura 2. Prototipo de dashboard



## 6. Enfoque para desarrollo del proyecto

### a. Proceso ETL (Extracción, Transformación y Carga)

- **Extracción:** Los datos son obtenidos desde SQL Server Management Studio conectando la base de datos.
- **Transformación:** En esta etapa se realiza la depuración de los datos. Se renombraron campos para facilitar la lectura, se eliminaron registros nulos o inconsistentes y se estandarizaron formatos como fechas o valores monetarios.
- **Carga:** Finalmente, los datos limpios y preparados se importarán en Power BI Desktop, seleccionando únicamente las tablas relevantes. Se establecerán las relaciones necesarias entre tablas para conformar un modelo coherente.

### b. Cubo OLAP

- Se construyó un modelo analítico en Power BI, siguiendo la lógica de un esquema en estrella.
- Las relaciones entre la tabla de hechos *FactSales* y sus respectivas dimensiones se representaron mediante claves primarias y foráneas.
- Se utilizaron funciones DAX para definir medidas como “Ventas Totales”, lo que permitió obtener indicadores

### c. Reportes y Visualizaciones

- Se desarrollaron reportes interactivos que permitieron segmentar y comparar datos por categorías de producto, regiones geográficas y periodos de tiempo.
- Se aplicaron formatos condicionales y filtros visuales para facilitar la interpretación de tendencias y comportamientos del mercado.
- Los gráficos incluyen columnas, mapas, tarjetas de KPI y líneas de tiempo.

#### **d. Dashboard Final**

- El producto final fue un dashboard integrado en Power BI, que combino visualizaciones clave para la toma de decisiones.
- El dashboard incluye elementos como indicadores de rendimiento (KPI), mapas de ventas por región, gráficos dinámicos y controles de segmentación.
- Este dashboard podrá ser compartido mediante Power BI Service para facilitar el acceso y colaboración entre usuarios.

### **7. Conclusión**

Este segundo avance consolida el desarrollo completo del dashboard en Power BI, basado en la base de datos ContosoRetailDW. Se aplicaron procesos ETL, se definió un modelo relacional tipo estrella y se construyeron visualizaciones dinámicas e interactivas. Gracias a este enfoque, se logró representar de manera clara las ventas por categoría y región, apoyando decisiones estratégicas. El resultado es un dashboard funcional, filtrable por año y región, que sintetiza los hallazgos más importantes para el negocio.