Evidencia de aprendizaje 4. Data Mart

Shamir Tano Morales – Grupo: 76

Kelly Vega Morales – Grupo: 94

Yampier Ruiz Patiño – Grupo: 94

Daniela Vallejo Bustamante – Grupo: 94

Facultad de Ingeniería y Ciencias Agropecuarias

Institución Universitaria Digital de Antioquia IU DIGITAL

Base de Datos II

Ing. Víctor Hugo Mercado

mayo de 2024

Tabla de contenido

[Introducción 3](#_Toc167614940)

[Objetivos 4](#_Toc167614941)

[Planteamiento del Problema 5](#_Toc167614942)

[Análisis del Problema 6](#_Toc167614943)

[Propuesta de la Solución 7](#_Toc167614944)

[Modelo Estrella 7](#_Toc167614945)

[Descripción del Modelo Estrella Propuesto 7](#_Toc167614946)

[Diseño del Modelo Estrella 7](#_Toc167614947)

[DimProducto: 7](#_Toc167614948)

[DimCliente: 7](#_Toc167614949)

[DimTiempo: 8](#_Toc167614950)

[DimPedido: 8](#_Toc167614951)

[DimDetallePedido: 8](#_Toc167614952)

[FACTVentas: 8](#_Toc167614953)

[Implementación del Proceso ETL (Extracción, Transformación y Carga) 9](#_Toc167614954)

[Extracción: 9](#_Toc167614955)

[Transformación: 9](#_Toc167614956)

[Carga: 9](#_Toc167614957)

[Conclusiones 10](#_Toc167614958)

# Introducción

En el contexto actual del análisis de datos, la migración de datos desempeña un papel fundamental para garantizar la integridad, calidad y eficiencia en el procesamiento de la información. En este documento, se abordará el proceso de migración de datos desde la base de datos Jardinería a una base de datos Staging, que servirá como un entorno intermedio para la preparación y análisis de los datos antes de su procesamiento final en un Data Mart estructurado según el modelo estrella. Este enfoque asegura que los datos estén limpios, transformados y preparados para un análisis eficiente y detallado.

# Objetivos

* Realizar una migración efectiva de los datos de la base de datos Jardinería.
* Diseñar la estructura de las tablas en la base de datos Staging basada en el modelo del Data Mart.
* Desarrollar consultas SQL para la extracción, transformación y carga de datos desde la base de datos Jardinería.
* Validar la integridad de los datos migrados en la base de datos Staging.
* Crear un Data Mart utilizando el modelo estrella para optimizar el análisis de datos.
* Proveer una documentación clara y concisa del proceso de migración para futuras referencias.

# Planteamiento del Problema

El procesamiento y análisis de grandes volúmenes de datos requiere un cuidadoso proceso de preparación para garantizar la calidad y consistencia de la información. La base de datos de Jardinería, aunque contiene datos relevantes para su análisis, puede no estar optimizada para este propósito. Por lo tanto, surge la necesidad de crear una base de datos Staging donde los datos puedan ser limpiados, transformados y preparados antes de su carga en un Data Mart diseñado para análisis detallado.

# Análisis del Problema

La base de datos Jardinería contiene múltiples tablas que almacenan información sobre oficinas, empleados, clientes, pedidos, productos y más. Sin embargo, esta estructura puede no ser la más adecuada para el análisis de datos. Por lo tanto, se propone la creación de una base de datos Staging que refleja la estructura relacional de la base de datos Jardinería pero que esté optimizada para el análisis de datos. En la base de datos Staging, se diseñarán tablas que reflejen las tablas originales de la base de datos Jardinería, con ajustes según sea necesario para mejorar la eficiencia y facilitar el análisis de los datos.

# Propuesta de la Solución

## Modelo Estrella

El modelo estrella propuesto consta de una tabla de hechos central que registra las ventas, y varias dimensiones principales: Pedido, DetallePedido, Producto, Cliente y Tiempo.

## Descripción del Modelo Estrella Propuesto

El modelo estrella se diseña con una tabla de hechos (FACTVentas) que contiene las métricas de ventas y está conectada a varias tablas de dimensiones que describen los contextos de las ventas (Producto, Cliente, Tiempo, Pedido, DetallePedido).

## Diseño del Modelo Estrella

Basado en el modelo ER mostrado, el modelo estrella propuesto incluye las siguientes tablas:

### DimProducto:

* + ID\_Producto (INT)
  + Nombre (VARCHAR)
  + Categoría (VARCHAR)
  + Dimensiones (VARCHAR)
  + Proveedor (VARCHAR)
  + Descripción (TEXT)
  + Cantidad\_en\_Stock (SMALLINT)
  + Precio\_Venta (NUMERIC)
  + Precio\_Proveedor (NUMERIC)

### DimCliente:

* + ID\_Cliente (INT)
  + Cliente (VARCHAR)
  + Contacto (VARCHAR)
  + Teléfono (VARCHAR)
  + Fax (VARCHAR)
  + Línea\_Dirección1 (VARCHAR)
  + Línea\_Dirección2 (VARCHAR)
  + Ciudad (VARCHAR)
  + Región (VARCHAR)
  + País (VARCHAR)
  + Código\_Postal (VARCHAR)
  + Empleado (INT)

### DimTiempo:

* + ID\_Tiempo (INT)
  + Fecha (DATE)
  + Año (INT)
  + Mes (INT)
  + Día (INT)
  + DíaSemanaNum (INT)
  + Trimestre (INT)
  + DíaAño (INT)

### DimPedido:

* + ID\_Pedido (INT)
  + ID\_Cliente (INT)

### DimDetallePedido:

* + ID\_Detalle\_Pedido (INT)
  + Nombre (VARCHAR)
  + Cantidad (INT)
  + Estado (VARCHAR)
  + Descripción (TEXT)
  + Precio\_Venta (NUMERIC)
  + Dimensiones (VARCHAR)
  + Categoría (VARCHAR)

### FACTVentas:

* + ID\_Venta (INT)
  + ID\_Pedido (INT)
  + ID\_Detalle\_Pedido (INT)
  + ID\_Producto (INT)
  + ID\_Cliente (INT)
  + Cantidad\_Vendida (INT)
  + Monto\_Total (NUMERIC)
  + ID\_Fecha (INT)

## Implementación del Proceso ETL (Extracción, Transformación y Carga)

### Extracción:

* + Consultas SQL para extraer datos de las tablas de la base de datos Jardinería.

### Transformación:

* + Limpieza de datos para eliminar inconsistencias y duplicados.
  + Transformación de datos para adaptarse a la estructura de la base de datos Staging y posteriormente al Data Mart.

### Carga:

* + Inserción de los datos transformados en la base de datos Staging.
  + Validación de los datos para asegurar su integridad.
  + Carga final de los datos en el Data Mart basado en el modelo estrella.

# Conclusiones

El modelo estrella propuesto proporciona una estructura organizada y eficiente para analizar las ventas de la empresa de jardinería. Al integrar datos de productos, clientes, tiempo, pedidos y detalles de pedidos en una estructura centralizada, facilita la generación de informes y análisis detallados. La migración de datos hacia la base de datos Staging y su posterior transformación y carga en el Data Mart garantizan una consulta más óptima y confiable, respondiendo efectivamente a los requerimientos del cliente. Este proceso asegura que los datos estén preparados de manera adecuada para su análisis, mejorando así la toma de decisiones basada en datos.