UNIVERSIDAD GALILEO

Postgrado en Análisis y Predicción de Datos

Curso: Diseño y Construcción de Data Warehouses Horario: Lun: 18:00 – 21:00

Tutor: Msc. José Rolando Lucero Morataya

**Diseño y gestión de Data Warahouses**

**Semana 4**

Integrantes:

Andrea María Hernández Marroquín – 240110074

Otto Francisco Chamo Cheley – 19001395

Edgar Geovany Ocaña Orozco – 24010004

Jose Eduardo Calderon Lopez - 14001764

Guatemala, 24 de marzo de 2024

# **INDICE**

Contenido

[**INDICE** 2](#_Toc162183035)

[**INTRODUCCION:** 3](#_Toc162183036)

[**DISEÑO DEL DATA WAREHOUSE** 4](#_Toc162183037)

[**DISEÑO DIMENSIONAL** 5](#_Toc162183038)

[**SQL** 6](#_Toc162183039)

[**CONCLUSIONES** 8](#_Toc162183040)

# **INTRODUCCION:**

El diseño dimensional en un Data Warehouse (DW) es un enfoque de modelado que se utiliza para organizar y estructurar los datos de manera que sean óptimos para el análisis y la generación de informes. A diferencia de los modelos relacionales tradicionales, que están diseñados para transacciones y operaciones de base de datos, el diseño dimensional se centra en facilitar el análisis de datos mediante la creación de estructuras fáciles de entender y consultar.

En un diseño dimensional, los datos se organizan en dos tipos principales de tablas: tablas de hechos y tablas de dimensiones. Las tablas de hechos contienen métricas o medidas cuantitativas que se están analizando, mientras que las tablas de dimensiones contienen descripciones o atributos que proporcionan contexto a las métricas. Este enfoque permite a los usuarios realizar consultas complejas y analizar datos desde diferentes perspectivas.

La presente tarea se presenta un detallado diagrama dimensional para una empresa que busca analizar información de un supermercado para poder tomar decisiones.

## **DISEÑO DEL DATA WAREHOUSE**

Para realizar el diseño dimensional para un Supermercado se propone lo siguiente la creación de las siguientes tablas:

1. Tabla customer\_dimension

Se decidió por construir una dimensión que se lleva registro de los datos del cliente, que contenga información sobre el segmento y código postal de donde hace la compra.

1. Tabla ship\_mode\_dimension

Se construyo una tabla que lleve registro de los tipos modos de entrega para cada uno de los pedidos, para lograr identificar cual es el más usado y, si eventualmente el negocio llega a ofrecer una nueva alternativa.

1. Tabla geography\_dimension

Esta tabla llevara el registro de todos los países y diferentes regiones a las que se ha enviado un cargamento.

1. Tabla producto\_dimension

Una tabla que lleva registro de todos los productos que tiene la empresa y, que ira creciendo según la empresa ofrezca nuevos productos.

1. Tabla sales\_facts

La tabla de hechos que lleva registro de todas las variables que buscamos analizar, lleva registro de cada una de las ventas, cantidad, descuentos y ganancias obtenidas.

## **DISEÑO DIMENSIONAL**

Interfaz de usuario gráfica, Aplicación, Teams

Descripción generada automáticamente

## **SQL**

CREATE DATABASE IF NOT EXISTS retail\_sales;

USE retail\_sales;

CREATE TABLE customer\_dimension (

     sk\_customer   INT NOT NULL auto\_increment PRIMARY KEY,

     customer\_id   VARCHAR (255) NULL,

     customer\_name   VARCHAR (255) NULL,

     segment     VARCHAR (255) NULL,

     postal\_code     int NULL,

     timestamp    DATETIME

);

CREATE TABLE ship\_mode\_dimension

  (

     sk\_ship\_mode   INT NOT NULL auto\_increment PRIMARY KEY,

     ship\_mode   VARCHAR (255) NULL,

     timestamp    DATETIME

  );

CREATE TABLE geography\_dimension

  (

     sk\_geography   INT NOT NULL auto\_increment PRIMARY KEY,

     country   VARCHAR (255) NULL,

     region VARCHAR (255) NULL,

     state  VARCHAR (255) NULL,

     city     VARCHAR (255) NULL,

     postal\_code     int NULL,

     timestamp    DATETIME

  );

CREATE TABLE product\_dimension

  (

     sk\_product   INT NOT NULL auto\_increment PRIMARY KEY,

     product\_id   VARCHAR (255) NULL,

     product\_name VARCHAR (255) NULL,

     subcategory  VARCHAR (255) NULL,

     category     VARCHAR (255) NULL,

     timestamp    DATETIME

  );

CREATE TABLE sales\_facts

  (

     sale\_id          VARCHAR (255) NOT NULL,

     date\_key          INT NOT NULL,

     sk\_customer       INT NOT NULL,

     sk\_geography      INT NOT NULL,

     sk\_ship\_mode      INT NOT NULL,

     sk\_product        INT NOT NULL,

     sales             DECIMAL (18, 2) NULL,

     quantity          INT NULL,

     discount          DECIMAL (18, 2) NULL,

     profit            DECIMAL(18, 2) NULL,

     FOREIGN KEY (sk\_customer) REFERENCES customer\_dimension (sk\_customer),

     FOREIGN KEY (sk\_geography) REFERENCES geography\_dimension(sk\_geography),

     FOREIGN KEY (sk\_ship\_mode) REFERENCES ship\_mode\_dimension(sk\_ship\_mode),

     FOREIGN KEY (sk\_product) REFERENCES product\_dimension(sk\_product)

  );

# **CONCLUSIONES**

Con el diseño del Data Warehouse anterior se desea convertir los datos en información valiosa y accionable para la toma de decisiones empresariales. Al proporcionar una estructura clara y organizada para los datos, se facilita el análisis, se mejora el rendimiento y se impulsa la colaboración en toda la organización.