

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERIA Y ARQUITECTURA
ESCUELA DE INGENIERIA EN SISTEMAS INFORMATICOS



ESPECIALIZACIÓN: INGENIERÍA DE DATOS 2025
DOCENTE: ING. RENÉ FABRICIO QUINTANILLA GOMEZ.

TAREA PRACTICA #1

“DESARROLLO E IMPLEMENTACIÓN DE DATA WAREHOUSE ”

NOMBRE
JOSUE SALVADOR CASTILLO PEREZ

CARNÉ
CP04010

Lunes 11 de Agosto de 2025

Contenido

Objetivos.....	3
Metologia	4
1.Identificacion de proceso de negocio	4
2.Declaracion de Granularidad	4
3.Identificacion de Dimesiones	4
4.Identificacion de Tabla de Hechos.....	4
Desarrollo de Proyecto	5
Conclusiones	6

Objetivos

Objetivo

Desarrollar e implementar un Data Warehouse de ventas para la integración y su análisis histórico de la información

Objetivos Especificos

1. Diseñar un modelo dimensional con esquema estrella
2. Implementar procesos ETL para integrar y depurar datos

Metologia

1. Identificación de proceso de negocio

- Proceso principal: *Ventas*
- Fuentes: tablas transaccionales, archivos de carga, sistemas de ventas.

2. Declaración de Granularidad

- Granularidad definida: *Detalle por línea de venta* (cada transacción individual).

3. Identificación de Dimesiones

- DimFechas
- DimProductos
- DimEstados (o regiones/territorios)
- DimOficinasVentas
- DimPromociones

4. Identificación de Tabla de Hechos

- Fact table: *HechosVentas*
- Métricas: Cantidad, Precio Unitario, Monto de Descuento, Monto Neto.

Desarrollo de Proyecto

1. Diseño Dimensional

- Se elabora un diagrama de entidad y relación en esquema estrella
- Se asegura que haya una integridad referencial con claves sustitutas (SK)
- Se utiliza la técnica scd tipo 2 en consiste poder manejar los cambios de los datos en una tabla de dimensiones, conservando un historial
- Campos históricos como EffectiveFrom, EffectiveTo, IsCurrent en dimensiones SCD Tipo 2.

2. Procesos ETL

Implementación de la herramienta SSIS, elaboración de paquetes para cada proceso de extracción y llenado de tablas

- Se utilizan la técnica de Staging (Tabla Intermedia) Paquetes Staging
- **Paquetes de Dimensiones:** carga controlada de cada dimensión, con lógica de SCD.
- **Paquete de Hechos:** carga de la tabla dw.HechosVentas asegurando integridad de claves.

3. Automatización de Dim Fechas

Para evitar errores por claves de fecha faltantes:

- Se creó el procedimiento dw.usp_EnsureDimFechasFromStaging.
- Se ejecuta automáticamente después de cargar stg.Ventas y antes de poblar dw.HechosVentas.

4. Control de valores nulos

En el proceso de carga de hechos:

- Cualquier valor NULL en claves sustitutas se reemplaza por un registro "desconocido" en la dimensión correspondiente (-1 o "No existe sala").
- Esto asegura que no haya valores nulos en la fact table.

Conclusiones

- Con la aplicación de Metodología Kimball permitio estructurar el trabajo de investigación en fases claras
- La automatizaciende control de fechas y nulos reduce errores y garantiza integridad