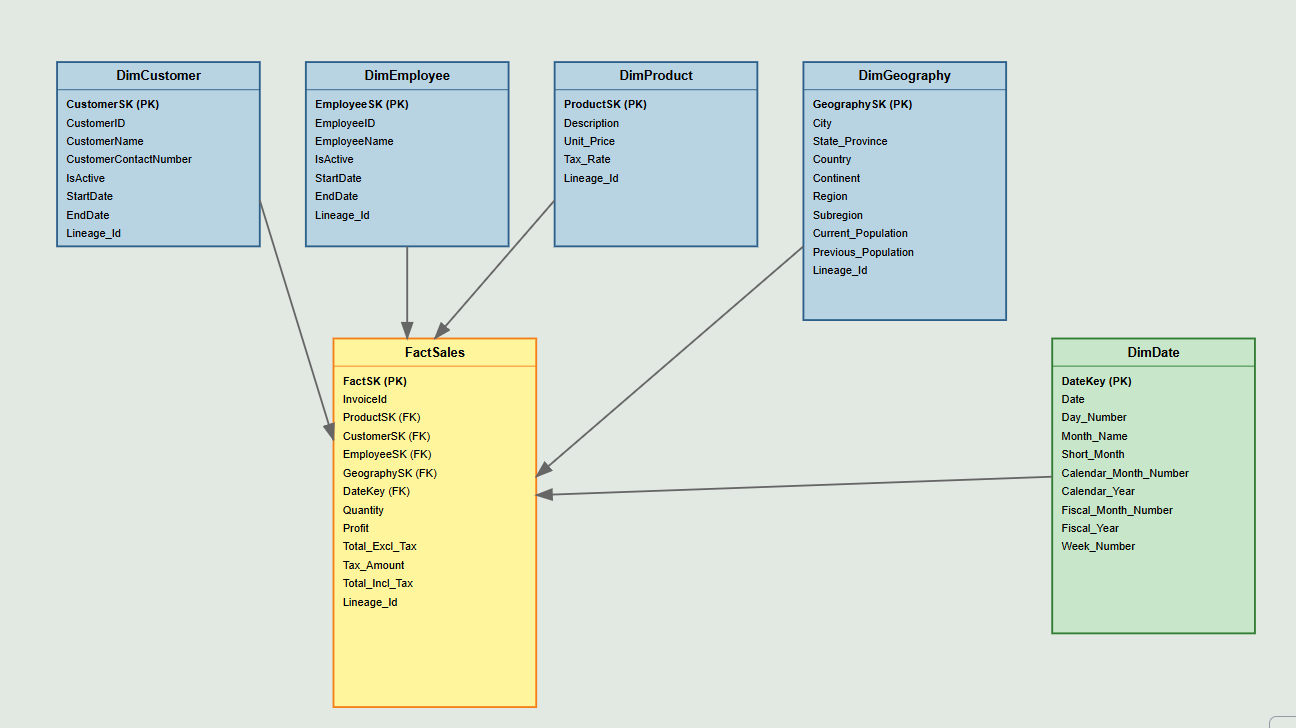
Insignia DW - ETL and Data Modeling Documentation

# 1. Introduction

This document outlines the implementation of a Data Warehouse solution for Insignia Corporation. It includes schema design, ETL procedures, dimension loading strategies (SCD Type 1, 2, 3), fact table loading, and a lineage tracking mechanism to ensure traceability and auditability of data loads.



This diagram shows:

* All Dimension Tables (Customer, Employee, Product, Geography, Date)
* Central Fact Table (FactSales)
* Relationships via surrogate and foreign keys

# 2. Step-by-Step Implementation in SSMS

## 2.1 Create Database

Run the following SQL in SSMS to create and switch to the working database:

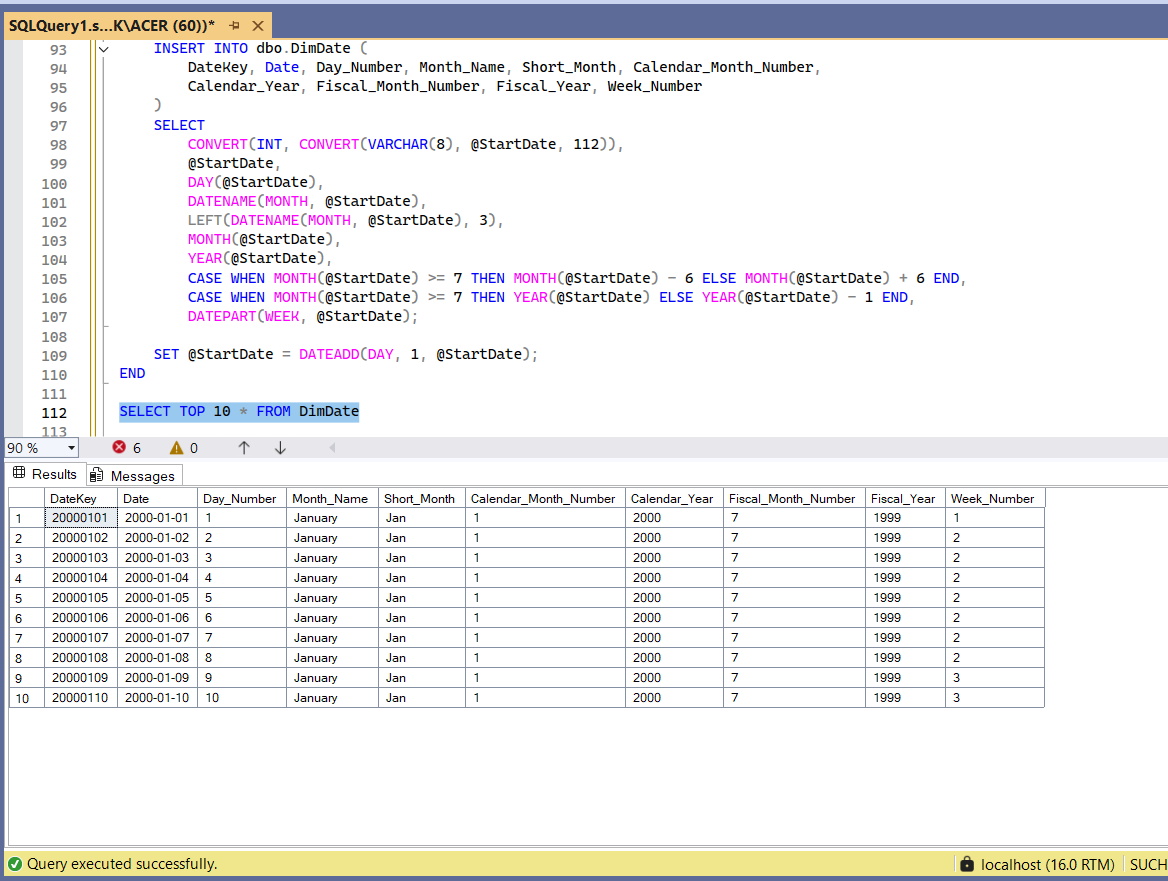
USE master;  
GO  
CREATE DATABASE InsigniaDW;  
GO  
USE InsigniaDW;

## 2.2 Create Tables

Create all required dimension, fact, and lineage tables using the scripts provided.

## 2.3 Populate Date Dimension

Use the date generation script to populate DimDate from the year 2000 to 2023.



## 2.4 Load Staging Copy

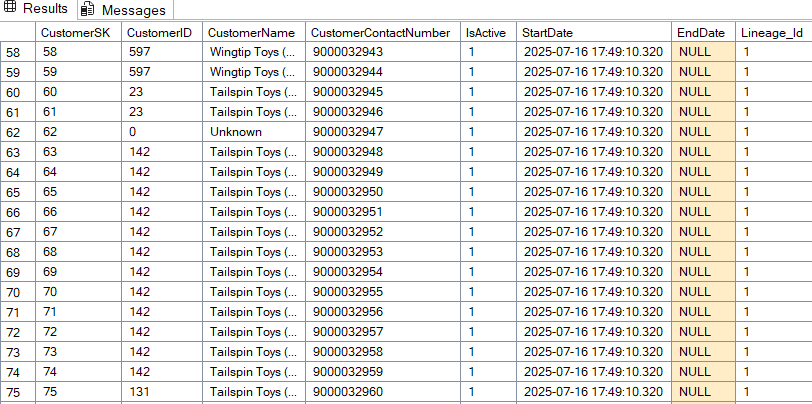
Create `Insignia\_staging\_copy` as a copy of `Insignia\_staging`. Truncate before each load, then insert data from `Insignia\_incremental`.

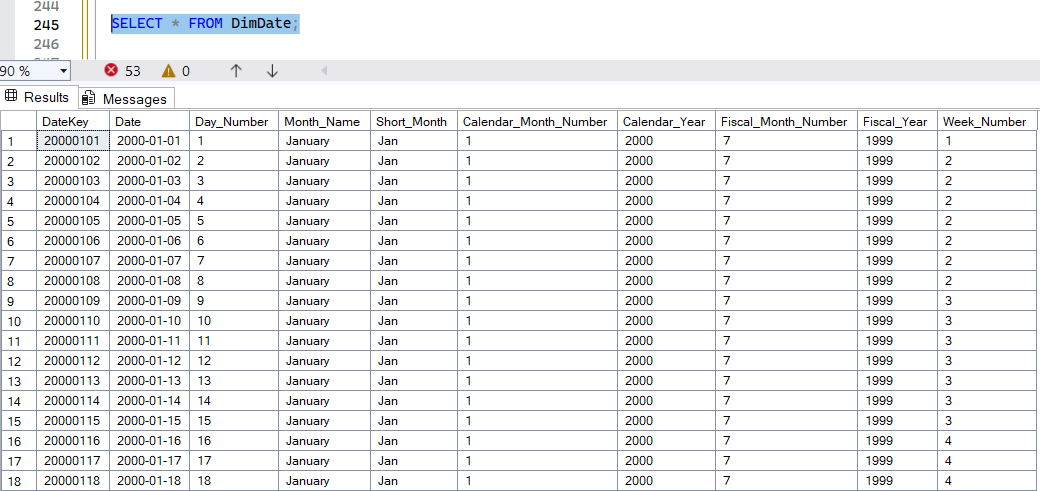
## 2.5 Run ETL Scripts

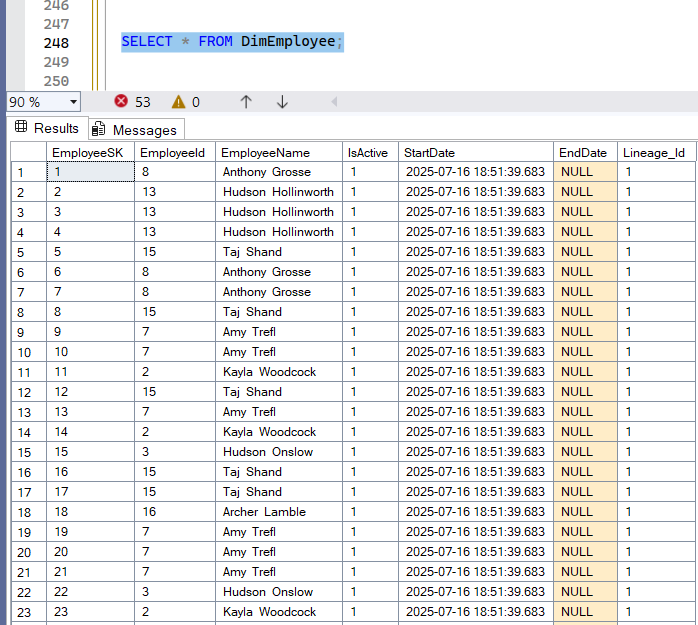
Execute ETL logic in this order:

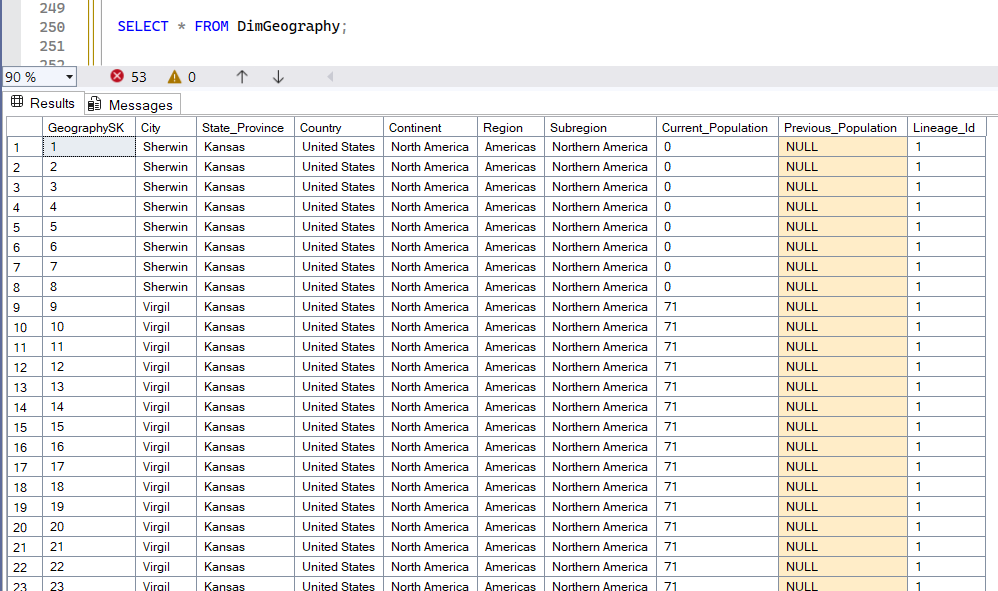
1. DimCustomer (SCD Type 2)  
2. DimEmployee (SCD Type 2)  
3. DimProduct (SCD Type 1)  
4. DimGeography (SCD Type 3)  
5. FactSales (fact loading)

SELECT \* FROM DimCustomer;



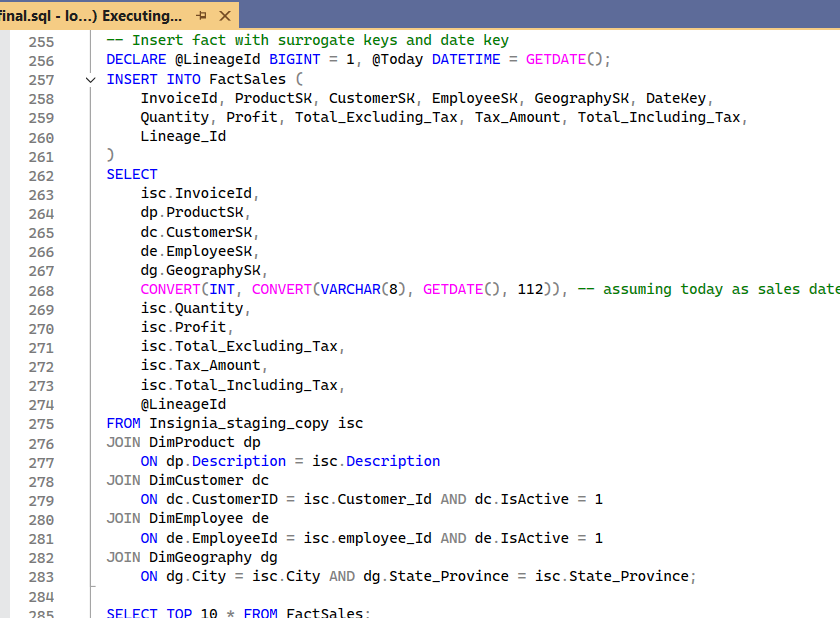






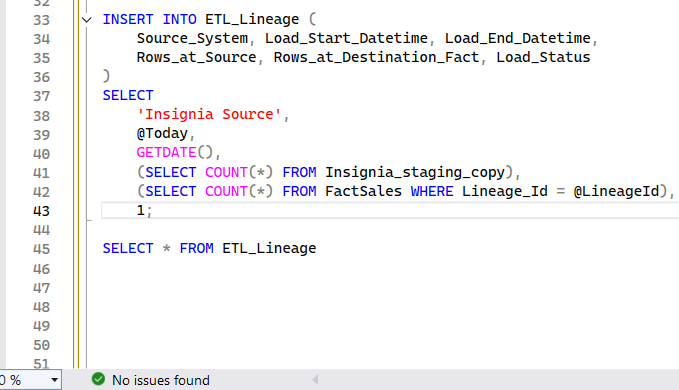
## 2.6 Insert into FactSales

After all dimensions are loaded, insert into `FactSales` with surrogate keys and DateKey.



## 2.7 Insert into ETL\_Lineage

Insert ETL audit log into the `ETL\_Lineage` table to track load status.



**Conclusion**

The implementation of the Insignia Data Warehouse solution enables the organization to streamline its data processing pipeline and ensures robust, scalable analytics capabilities. The data model was carefully designed to capture key business entities and their relationships through dimension and fact tables, following data warehousing best practices.

The ETL process incorporates support for Slowly Changing Dimensions (SCD Type 1, 2, and 3), lineage tracking for audit purposes, and logic to handle late-arriving dimension records. With the inclusion of a detailed Date Dimension and a Lineage table, the solution not only supports historical tracking but also offers insights into the quality and completeness of each data load.

The clear separation of dimensions and fact tables ensures a normalized and performance-optimized structure, which is crucial for analytical reporting and BI tools integration. With this foundation in place, Insignia Corporation is well-positioned to scale their data operations and drive better decision-making through reliable and timely insights

-Sucharita Gorai