Project Dimensional Data Analysis

1. Schema Overview

The dimensional model is designed in a snowflake schema, which enhances normalization by separating out address/location data into a dedicated Dim_Location table, linking stores, resellers, and customers. The model supports analysis across multiple dimensions: products, stores, resellers, customers, sales channels, and dates.

The core fact tables:

- Fact_Sales

- Fact_Target_Sales_Channel

- Fact_Target_Sales_Product

These tables capture actual sales transactions and sales targets, enabling a robust Actual vs. Target analysis at a daily grain.

2. Fact Tables

- **Fact_Sales**

Grain: One row per sales transaction at the line-item level (SalesDetailID).

Purpose: Captures all individual sales transactions, tracking both direct-to-customer and reseller sales. Supports reporting across stores, resellers, products, channels, and dates.

Key Columns: SalesHeaderID (PK), SalesDetailID (PK), Dim_Reseller_ID (FK), Dim_Store_ID (FK), Dim_Customer_ID (FK), Dim_Product_ID (FK), Dim_Channel_ID (FK), Dim_Location_ID (FK), Dim_Date_ID (FK), SalesAmount, SalesQuantity, SalesUnitPrice, SalesTotalProfit

Why modeled this way: It captures detailed transaction info that allows deep slicing and dicing by customer, reseller, store, and location. Including both SalesAmount and SalesProfit allows flexible profitability analysis.

- **Fact_Target_Sales_Channel**

Grain: One row per Channel, Date, Store, Customer, and Reseller.

Purpose: Captures daily sales targets set per channel/store/reseller/customer.

Key Columns: Dim_Customer_ID (FK), Dim_Date_ID (FK), Dim_Store_ID (FK), Dim_Channel_ID (FK), Dim_Reseller_ID (FK), Target_Channel_ID (PK), SalesAmount_Target

Why modeled this way: Although target data is annual, we normalize it down to daily targets to align with the fact sales grain, enabling like-for-like comparisons.

- **Fact_Target_Sales_Product**

Grain: One row per Product, Date.

Purpose: Captures daily targets for product sales quantity.

Key Columns: Dim_Product_ID (FK), Dim_Date_ID (FK), Target_Product_ID (PK), SalesQuantityTarget

Why modeled this way: Supports tracking whether specific products (e.g., Women's Apparel) are meeting sales expectations on a daily basis.

3. Dimension Tables

- **Dim_Product**

Grain: One row per unique product.

Purpose: Stores detailed product attributes: price tiers (retail, wholesale), cost, and category/type, enabling profitability calculations.

Key Columns: Dim_Product_ID (PK), ProductID (Natural key), ProductName, ProductType, ProductCategory, RetailPrice, WholesalePrice, Cost, ProfitMargin

Why modeled this way: Essential for calculating: Sales Profit: (Quantity × Price) - (Quantity × Cost), Profit Margin %: (Price - Cost) ÷ Price

- **Dim_Store**

Grain: One row per store.

Purpose: Captures store-specific details and links to location.

Key Columns: Dim_Store_ID (PK), StoreID (Natural key), Dim_Location_ID (FK), StoreName, StoreManager

Why modeled this way: Allows location-based analysis

- **Dim_Reseller**

Grain: One row per reseller.

Purpose: Tracks resellers like Department Stores and Branded Franchises.

Key Columns: Dim_Reseller_ID (PK), ResellerID (Natural key), Dim_Location_ID (FK), ResellerName, Phone,Email

- **Dim_Customer**

Grain: One row per customer.

Purpose: Captures customer profiles for direct sales.

Key Columns: Dim_Customer_ID (PK), CustomerID (Natural key), Dim_Location_ID (FK), Customer_FirstName, Customer_LastName, Segment, SubSegment

- **Dim_Channel**

Grain: One row per channel.

Purpose: Classifies sales channels by Direct or Indirect categories.

Key Columns: Dim_Channel_ID (PK), ChannelID (Natural key), ChannelCategory

- **Dim_Date**

Grain: One row per date.

Purpose: Standard calendar table supporting daily analysis.

Key Columns: Dim_Date_ID (PK), Full_Date, Month, Year

- **Dim_Location**

Grain: One row per unique address/location.

Purpose: Stores geographical data for customers, resellers, and stores.

Key Columns: Dim_Location_ID (PK), Address, City, PostalCode, Country

4. Special Design Considerations

- **Snowflake schema:** Location and contact details are normalized into Dim_Location.

- **Surrogate keys:** Used throughout for performance and integrity (e.g., Dim_Store_ID vs. StoreID).

- **Natural keys:** Included for business transparency and alignment.

- **Targets:** Annual targets are **decomposed into daily records** to align grains.