

Data Model Documentation Template

Section	Details
Tables Used	<p>Fact Table:</p> <ul style="list-style-type: none">• Fact_sales (main transactional table) <p>Dimension Tables:</p> <ul style="list-style-type: none">• Dim_country – Country information• Dim_city – City, latitude, longitude• Dim_customer – Customer name, city• Dim_product – Product name, product class• Dim_distributor – Distributor name and ID• Dim_sales_rep – Sales representative details• Dim_sales_team – Sales team and manager info• Dim_subchannel – Subchannel and related Channel_ID• Dim_channel – Channel type• Dim_month – Month details <p>Measure Table:</p> <ul style="list-style-type: none">• Measure Table – Contains calculated measures (e.g., Avg Sales)
Relationships	<ul style="list-style-type: none">- Fact_sales → Dim_city: linked via City_ID- Dim_city → Dim_country: linked via Country_ID- Fact_sales → Dim_customer: linked via Customer_ID- Fact_sales → Dim_product: linked via Product_ID- Fact_sales → Dim_distributor: linked via Distributor_ID- Fact_sales → Dim_sales_rep: linked via Sales_rep_ID- Dim_sales_rep → Dim_sales_team: linked via Sales_team_ID- Fact_sales → Dim_subchannel: linked via Subchannel_ID- Dim_subchannel → Dim_channel: linked via Channel_ID- Fact_sales → Dim_month: linked via Month_ID
Key Calculations / Measures	<ul style="list-style-type: none">- Total Sales = SUM(Fact_sales[Sales])- Average Sales = AVERAGE(Fact_sales[Sales]) (already in Measure Table)- Total Quantity = SUM(Fact_sales[Quantity])- Avg Price per Unit = DIVIDE(SUM(Fact_sales[Sales]), SUM(Fact_sales[Quantity]))- Sales by Channel / Subchannel = Using Dim_subchannel and Dim_channel relationships

- Sales by Region = Using **Dim_city** → **Dim_country** hierarchy

Notes

- The model is a hybrid schema: mostly Star Schema with some Snowflaked dimensions (City → Country, Sales Rep → Sales Team, Subchannel → Channel).
- Ensure one-to-many relationships are single-directional for performance. - All dimension tables have unique keys (e.g., **Customer_ID**, **Product_ID**, etc.).
- Consider flattening snowflaked dimensions if dataset grows large.
- Time-based calculations rely on **Dim_month** (use for YTD, QoQ comparisons).

Diagram:

DATA MODEL

