Superstore Sales & Profitability Analysis – SQL Portfolio Project

Objective: Analyze a retail superstore dataset using SQL to understand sales trends, profit drivers, customer value, and discount impact. This project demonstrates foundational data analytics skills.

Data Pipeline

• RAW layer: Import CSV into **superstore_raw** with all fields as text for safe loading.

• CLEAN layer: Convert raw text into typed columns (DATE, NUMERIC, INT). Stored in **superstore_clean**.

• ANALYTICS layer: Summary tables for insights (sales by region, customer value, product profitability).

Pipeline Tables

Layer	Table	Purpose
RAW	superstore_raw	Direct import, no data types, ensures no load failures.
CLEAN	superstore_clean	Typed, validated data ready for analysis.
ANALYTICS	superstore_analytics_region	Summary of sales & profit by region.

Business Questions Answered

• Which regions generate highest sales and profit?

• Which products are most profitable or loss-making?

• How do discount levels impact profit margins?

• Who are the top customers by revenue and profit?

Key SQL Techniques Used

• GROUP BY, ORDER BY, LIMIT

• Aggregate functions (SUM, AVG, COUNT)

• CASE statements for discount banding

• Date conversion using TO_DATE()

• Creating analytic summary tables