SQL Data Analytics Project - Notes

Overview

This project involves building a SQL-based analytics solution using a structured data warehouse approach. It focuses on data modeling, ETL, and various analytical techniques essential for business intelligence and data storytelling.

Key Modules

1. SQL Data Warehouse Project

- Designing fact and dimension tables
- Understanding star schema structure
- Relationships and foreign keys
- Importance of surrogate keys

2. ETL Process

- Extract data from CSV files
- Transform data for consistency and structure
- Load into SQL tables using BULK INSERT
- Use schemas (e.g., gold) to organize layers

Core Analytical Techniques

Change Over Time

Analyze trends over daily, monthly, or yearly periods using GROUP BY, DATEPART, ORDER BY, AVG, SUM.

Cumulative Analysis

Use SUM(...) OVER(ORDER BY ...) for rolling totals. Useful for financial tracking, growth metrics.

Performance Measurement

SQL Data Analytics Project - Notes

Calculating key metrics like sales, quantity, average price. Evaluate trends and compare KPIs.

Data Segmentation

Grouping by customer attributes (e.g., age, country, marital status). Enables targeted insights and cohort analysis.

Part-to-Whole Analysis

Understand proportions using percentage calculations. Common for pie charts, stacked bars, contribution metrics.

Dimensions vs Measures

Dimensions: Qualitative attributes (e.g., country, gender, product)

Measures: Quantitative values (e.g., sales amount, quantity)

Tips

Use views or CTEs for modular and readable queries

Validate data before and after loading

Always GROUP BY non-aggregated fields when using SUM, COUNT, etc.