AdventureWorks 2019 SQL Analysis Project Brief

Project Overview

This project uses the AdventureWorks 2019 sample database to provide hands-on experience with SQL and data warehousing concepts. You will begin by preparing the data in a star schema format, creating fact and dimension tables to organize the AdventureWorks sales and customer data. Once the data is structured, you'll conduct various analyses to derive insights about AdventureWorks' sales performance, customer trends, and inventory status.

Learning Objectives

1. **Data Structuring**: Understand and apply concepts of fact and dimension tables for a more efficient, analytical database schema.
2. **Data Retrieval & Aggregation**: Enhance SQL skills for querying, joining, and summarizing data.
3. **Insight Generation**: Perform analyses to extract business insights that can guide AdventureWorks' business strategies.

Project Scope and Tasks

This project consists of two main phases: **Data Preparation** and **Data Analysis**.

Phase 1: Data Preparation - Building Fact and Dimension Tables

In this phase, you will organize raw data tables into a **star schema** by creating fact and dimension tables, optimizing the data structure for analysis. Here is a guide for the tables to create:

1. Define Fact and Dimension Tables

- **Fact Table**: Contains transactional data, generally capturing measurable, quantitative data points.
- **Dimension Tables**: Hold descriptive attributes, such as product names, customer details, and dates.

2. Suggested Fact and Dimension Tables

Table Type	Description
FactSales	Core sales data with transaction-level details for analyzing sales by products, customers, and regions.
DimCustomer	Customer details, including location and demographics for customer segmentation analysis.
DimProduct	Product details, including category and pricing for product performance analysis.
DimEmployee	Employee information, helping analyze sales performance by sales representatives.
DimDate	Date information, enabling time-based trend analysis.

3. Steps for Creating Fact and Dimension Tables

- **Fact Table (FactSales)**: Aggregate transactional sales data from SalesOrderHeader and SalesOrderDetail tables, including fields like order date, product details, customer ID, and total sales amount.
- **Dimension Tables**:
    - **DimCustomer**: Combine data from Customer, Person, and Address tables to get customer names, locations, and other descriptive fields.
    - **DimProduct**: Organize product details, including category and subcategory, from Product, ProductCategory, and ProductSubcategory tables.
    - **DimEmployee**: Capture employee data from Employee, Department, and EmployeeDepartmentHistory tables, focusing on sales representatives.
    - **DimDate**: Create a date table to simplify time-based analysis, breaking dates into year, month, quarter, and weekday attributes.

Phase 2: Data Analysis

Once the data is structured, use SQL to perform analysis tasks across various business areas. Here are suggested analyses and insights to derive from the data:

1. **Sales Analysis**:
    - Examine sales performance by product, region, and sales representative. Determine top-selling products, seasonal sales trends, and regional performance.
    - Use FactSales and DimProduct for sales volume by category, product, and sales territory.
2. **Customer Insights**:
    - Identify the most valuable customers, common purchase patterns, and customer demographics.
    - Link FactSales with DimCustomer to segment customers by region, total purchase amount, and order frequency.
3. **Product Inventory & Performance**:

    - Track inventory levels and assess product performance. Analyze sales data alongside inventory to identify products with low stock or high turnover.
    - Use DimProduct to evaluate category performance, pricing impact, and sales volume for strategic stocking and pricing adjustments.
4. **Employee Performance**:
    - Analyze individual and department-level sales performance. Track sales made by each representative and calculate key metrics like average sales per employee.
    - Link FactSales with DimEmployee to identify top performers and areas for sales training or support.
5. **Financial and Time-Based Analysis**:
    - Perform revenue and expense analysis by linking FactSales and DimDate. Analyze by year, quarter, or month to reveal time-based trends in revenue.
    - Use DimDate to create trend reports on quarterly sales growth, identify peak sales seasons, and project future revenue based on historical patterns.

Deliverables

1. **Data Preparation Documentation**: Outline the process used to build fact and dimension tables, including brief explanations of the fields chosen for each table.
2. **SQL Analysis Queries**: Document the SQL queries used for each analysis task, ensuring clarity and consistency in query structure.
3. **Insights Report**: Summarize your findings from each analysis area, providing insights into AdventureWorks' business trends.
4. **Presentation**: Create a visual presentation that highlights key insights and recommendations for improving sales, inventory, and customer retention.

Grading Criteria

1. **Data Structure**: Adhere to the star schema approach for optimal query performance.
2. **Data Accuracy**: Ensure all joins and aggregations align correctly with primary keys in fact tables and dimension tables.
3. **Query Efficiency**: Aim for clear, efficient SQL queries to handle large data volumes and complex joins effectively.
4. **Insight Quality**: Focus on generating actionable insights that could realistically impact AdventureWorks' business strategy.