Final SQL Project: AdventureWorks 2022 Data Analysis

Project Overview

As a newly hired Data Analyst at AdventureWorks, a leading manufacturer and distributor of bicycles and cycling accessories, you are tasked with analyzing various aspects of the company's operations. Your analysis will focus on sales performance, product information, customer behavior, and employee data. The insights you provide will aid in strategic decision-making to enhance business performance.

Project Objectives

- Demonstrate proficiency in writing SQL queries ranging from basic to intermediate complexity.
- Apply SQL concepts such as data retrieval, filtering, sorting, aggregation, joins and subqueries
- Analyze and interpret data to provide meaningful business insights.
- Present findings in a clear and organized manner.

Project Structure

The project is divided into two main sections:

1. Beginner Level Tasks
2. Intermediate Level Tasks

Each section contains specific tasks that progressively build upon each other to assess different SQL competencies.

Section 1: Beginner Level Tasks

Task 1: List All Employees and Their Job Titles

- Objective: Retrieve basic employee information.
- Instructions:
    - Use the HumanResources.Employee and Person.Person tables.
    - Join these tables to get the employees' first and last names along with their job titles.

Task 2: Display All Products and Their List Prices

- Objective: Extract product details.
- Instructions:

- Query the Production.Product table to list all products with their respective list prices.

Task 3: Retrieve Customers Who Placed Orders in 2021

- **Objective:** Identify active customers within a specific timeframe.
- **Instructions:**
  - Use the Sales.SalesOrderHeader, Sales.Customer, and Person.Person tables.
  - Filter orders placed in the year 2021.

Task 4: List the Top 10 Most Expensive Products

- **Objective:** Identify high-value products.
- **Instructions:**
  - Query the Production.Product table.
  - Sort the products by ListPrice in descending order and limit the results to the top 10.

Task 5: Show the Total Number of Orders Placed in 2021

- **Objective:** Aggregate order data for a specific year.
- **Instructions:**
  - Use the Sales.SalesOrderHeader table.
  - Count the number of orders where the OrderDate falls in 2021.

Task 6: List Sales Orders with TotalDue Greater Than $1,500

- **Objective:** Highlight high-value transactions.
- **Instructions:**
  - Query the Sales.SalesOrderHeader table.
  - Filter orders where TotalDue exceeds $1,500.

Task 7: Retrieve Products with ListPrice Between $100 and $500

- **Objective:** Filter products within a specific price range.
- **Instructions:**
  - Use the Production.Product table.
  - Apply a WHERE clause to select products priced between $100 and $500.

Task 8: Retrieve Customers from a Specific Region (e.g., "United States")

- **Objective:** Segment customers based on geographical location.
- **Instructions:**
  - Utilize the Sales.Customer, Person.Person, and Person.Address tables.
  - Filter customers where CountryRegionName is "United States".

Section 2: Intermediate Level Tasks

Task 9: Calculate Total Sales Amount for Each Year from 2020 to 2022

- **Objective:** Analyze sales trends over multiple years.
- **Instructions:**
  - Use the `Sales.SalesOrderHeader` table.
  - Group sales by year and sum the `TotalDue` for each year between 2020 and 2022.

Task 10: Display Number of Orders Placed by Each Customer

- **Objective:** Understand customer ordering behavior.
- **Instructions:**
  - Join `Sales.SalesOrderHeader`, `Sales.Customer`, and `Person.Person` tables.
  - Group by `CustomerID` and count the number of orders per customer.

Task 11: List Products That Have Never Been Sold

- **Objective:** Identify unsold inventory.
- **Instructions:**
  - Use the `Production.Product` and `Sales.SalesOrderDetail` tables.
  - Find products in `Production.Product` that do not have corresponding entries in `SalesOrderDetail`.

Task 12: Find Total Number of Employees with the Title "Sales Representative"

- **Objective:** Assess the size of the sales team.
- **Instructions:**
  - Query the `HumanResources.Employee` table.
  - Count employees where `JobTitle` is "Sales Representative".

Task 13: Retrieve Average ListPrice for All Products in the "Bikes" Category

- **Objective:** Determine pricing strategy for the Bikes category.
- **Instructions:**
  - Join `Production.Product`, `Production.ProductSubcategory`, and `Production.ProductCategory` tables.
  - Filter products belonging to the "Bikes" category and calculate the average `ListPrice`.

Task 14: List Top 5 Customers Based on Total Order Amount

- **Objective:** Identify top-performing customers.
- **Instructions:**
  - Join `Sales.SalesOrderHeader`, `Sales.Customer`, and `Person.Person` tables.
  - Sum `TotalDue` per customer and select the top 5 customers with the highest total order amounts.

Task 15: Display All Products Sold More Than 50 Times in 2023

- **Objective:** Highlight high-demand products.
- **Instructions:**
  - Join `Production.Product` and `Sales.SalesOrderDetail` tables.
  - Filter sales from 2023 and group by `ProductID` to count sales exceeding 50.

Section 3: Submission Guidelines

Deliverables

1. **SQL Script File:**
   - Compile all SQL queries into a single `.sql` file named `FinalProject_AdventureWorks_YourName.sql`.
   - Each query should be preceded by a comment indicating the task number and description. For example:

     ```sql
     -- Task 1: List All Employees and Their Job Titles
     SELECT ...
     ```

2. **Report Document:**
   - Provide a brief report (in PDF or Word format) summarizing your findings for each task.
   - Include interpretations of the results and any insights or recommendations based on the data analysis.
3. **Presentation (Optional for Extra Credit):**
   - Create a short presentation (slides) showcasing key findings from your analysis.
   - Highlight significant insights and suggest potential business strategies based on the data.

Grading Criteria

Your project will be evaluated based on the following criteria:

1. **Correctness (50%)**

    - SQL queries produce accurate and expected results.
    - Proper use of SQL syntax and functions.
2. SQL Best Practices (30%)
    - Clean and readable code with proper indentation.
    - Meaningful use of table aliases.
    - Efficient use of joins and conditions.
    - Appropriate commenting for clarity.
3. Analysis and Insights (15%)
    - Ability to interpret query results.
    - Providing meaningful business insights based on data.
4. Presentation and Documentation (5%)
    - Organized and well-structured submission.
    - Clear and concise explanations in the report.

Additional Resources

- AdventureWorks Database Documentation:
    - Familiarize yourself with the database schema and relationships.
    - Download Documentation
- Tools for SQL Development:
    - SQL Server Management Studio (SSMS): A comprehensive tool for managing SQL Server databases.
    - Azure Data Studio: A cross-platform tool for data professionals using the Microsoft family of on-premises and cloud data platforms.

Tips for Success

- Understand the Schema: Before writing queries, take time to understand the tables and their relationships within the AdventureWorks database.
- Start Simple: Begin with straightforward queries to ensure you understand how to retrieve basic data before moving on to more complex tasks.
- Use Aliases: Implement table aliases to make your queries more readable and manageable, especially when dealing with multiple joins.
- Validate Results: After writing each query, verify the results to ensure accuracy. Check for edge cases and ensure that your filters and conditions are correctly applied.
- Optimize Your Queries: Aim for efficiency by selecting only necessary columns and avoiding unnecessary computations or joins.
- Comment Your Code: Clearly comment each section of your SQL script to indicate which task it addresses. This aids in readability and assessment.
- Seek Help When Needed: Utilize available resources such as documentation, tutorials, and peer discussions if you encounter challenges.

Final Thoughts

This project is designed to challenge and showcase your SQL skills in a practical, real-world context. By analyzing the AdventureWorks 2022 Database, you will gain valuable experience in data manipulation, querying, and interpretation—essential skills for any aspiring data analyst. Approach each task methodically, ensure clarity in your queries, and strive to provide insightful analyses that can drive business decisions.

Good luck!