**ASSIGNMENT 5**

**DATA WAREHOUSE**

**GROUP MEMBERS**

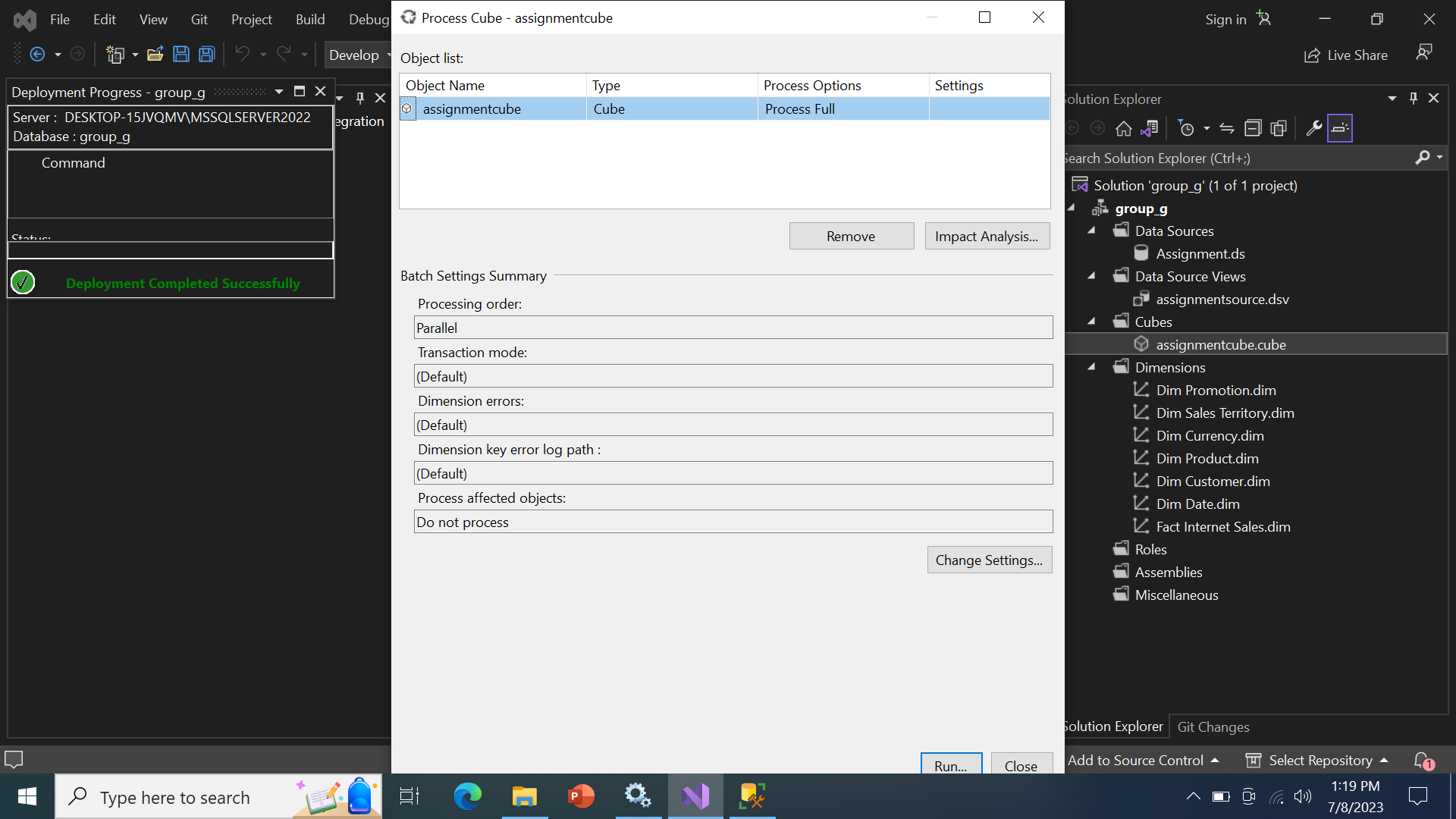
**ADNAN SAMAD 19B-011-SE A**

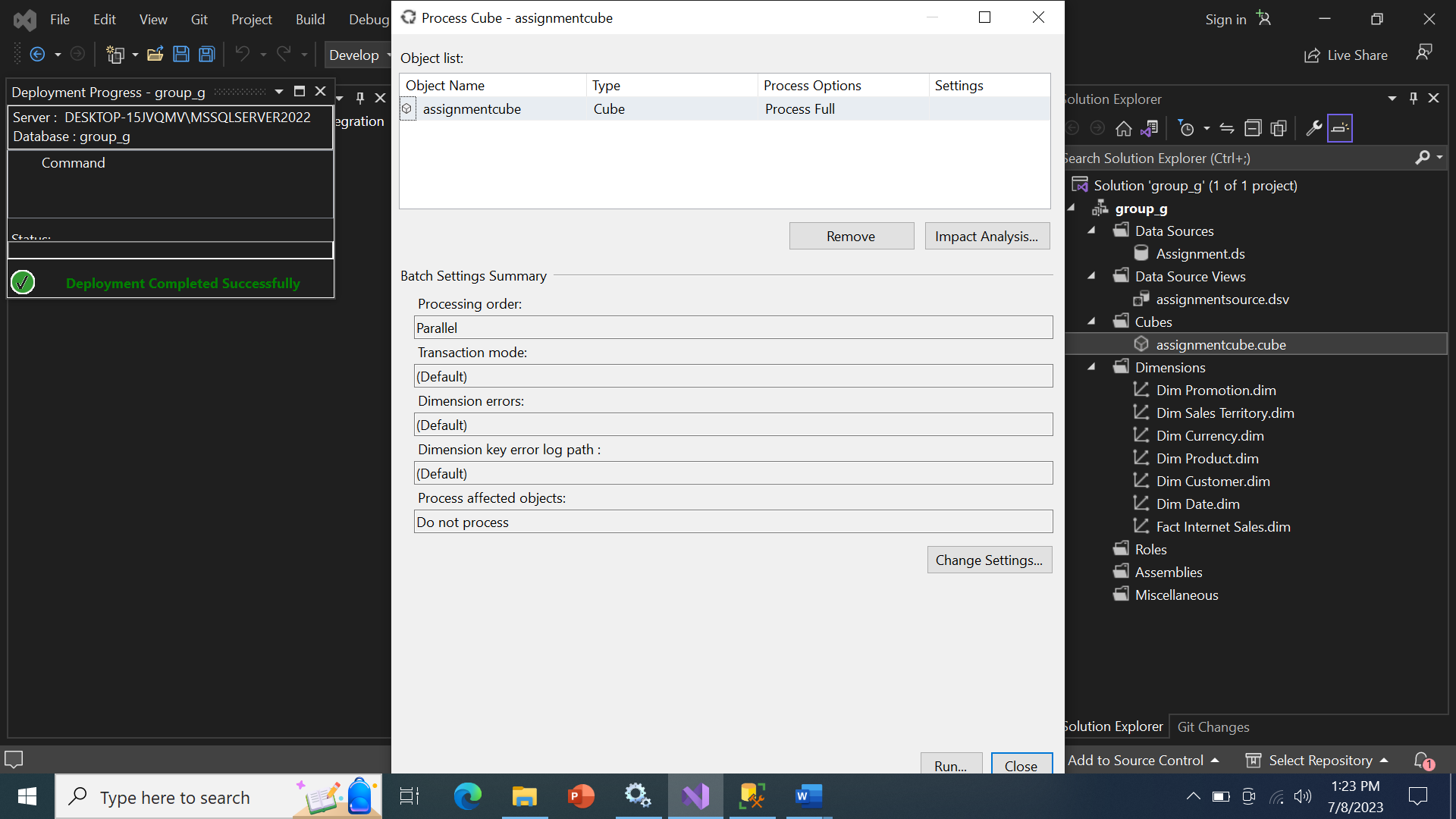
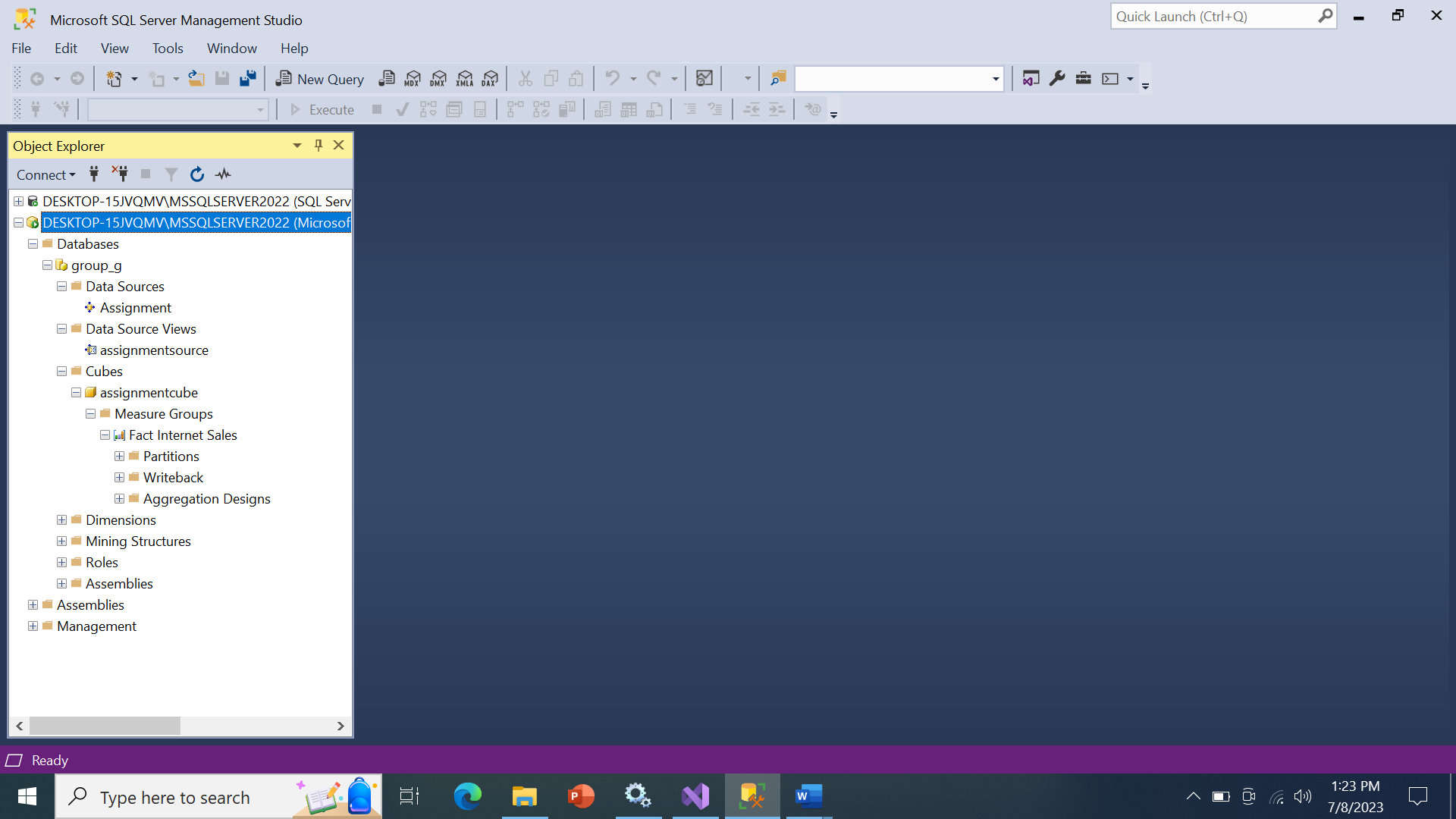
**AIZAH KHAN 19B-016-SE A  
ALISHA HAMEED 19B-017-SE**

**HAFIZ TAHA SHARIF 19B-035-SE A  
MUHAMMAD OMER SIDDIQUI 19B-004-SE A**

**PROFESSOR**

**DR LUBAID AHMED**





Scenario: Creating a Data Warehouse with AdventureWorks

In this scenario, we are creating a data warehouse using the AdventureWorks database. AdventureWorks is a sample database provided by Microsoft that simulates an e-commerce system. The goal of the data warehouse is to consolidate and integrate data from multiple sources within the AdventureWorks database, enabling efficient reporting, analysis, and decision-making.

Definition of a Data Warehouse:

A data warehouse is a central repository that stores structured, historical data from various sources. It is designed to support business intelligence and reporting activities by providing a consolidated and consistent view of data for analysis and decision-making purposes. Data warehouses typically involve the extraction, transformation, and loading (ETL) of data from source systems into a dimensional model that optimizes querying and reporting performance.

Explanation of the Scenario:

In this scenario, we are using the AdventureWorks database as the source for our data warehouse. AdventureWorks contains tables representing different aspects of an e-commerce system, such as customers, orders, products, and sales. Our goal is to create a data warehouse that consolidates and integrates these tables to provide a unified view of the data.

To achieve this, we perform the following steps:

1. Data Extraction: We extract the relevant data from the AdventureWorks tables using SQL queries. This involves selecting the necessary columns and applying any necessary transformations or aggregations.

2. Data Transformation: We transform the extracted data to conform to the dimensional model of the data warehouse. This includes cleaning the data, resolving inconsistencies, and structuring it in a way that optimizes querying and reporting.

3. Data Loading: We load the transformed data into the appropriate tables of the data warehouse. These tables are designed using a dimensional model, typically consisting of fact tables and dimension tables. Fact tables contain the numerical and measurable data, while dimension tables provide the descriptive attributes.

4. Dimensional Modeling: We design and create the dimension and fact tables based on the business requirements. Dimension tables capture the descriptive attributes of the data, such as customer details, product information, and date dimensions. Fact tables store the numerical measures, such as sales quantities and amounts.

5. Reporting and Analysis: With the data warehouse in place, business users can perform various reporting and analysis tasks. They can generate reports, run ad-hoc queries, perform trend analysis, identify patterns, and gain insights to support decision-making processes. The data warehouse provides a unified and consistent view of the data, enabling efficient and accurate analysis.

By creating a data warehouse with AdventureWorks, we are able to consolidate and integrate data from various tables within the database, providing a solid foundation for reporting and analysis. The dimensional model and optimized structure of the data warehouse allow for efficient querying and reporting, empowering business users to gain valuable insights and make informed decisions based on the data.

SELECT p.ProductLine, SUM(f.SalesAmount) AS TotalSalesAmount, f.ProductKey, p.ProductKey AS Expr1, f.SalesAmount, p.ListPrice, p.ProductLine AS Expr2

FROM FactInternetSales AS f INNER JOIN

DimProduct AS p ON f.ProductKey = p.ProductKey

GROUP BY p.ProductLine, f.ProductKey, p.ProductKey, f.SalesAmount, p.ListPrice, p.ProductLine

* Total Sales Amount by Product Line: The query calculates the total sales amount for each product line in the AdventureWorks data warehouse. Executives can analyze this information to understand which product lines generate the highest sales revenue.
* Product Key and List Price: The query includes the product key and list price columns from the DimProduct table. Executives can use this information to identify specific products and their corresponding list prices in relation to the sales performance.
* Grouping and Aggregation: The GROUP BY clause groups the results by product line, product key, sales amount, list price, and product line again. Executives can analyze the aggregated values to gain insights into sales trends within specific product lines and identify the products contributing most to the overall sales.
* Comparison of Sales Amount and List Price: By comparing the sales amount and list price columns, executives can assess the effectiveness of pricing strategies. They can identify products that generate higher sales despite having lower list prices, indicating potential opportunities for pricing adjustments or promotions.
* Product Line Performance: Executives can analyze the total sales amount for each product line to assess the performance of different product categories. This analysis can help them identify product lines that are driving revenue growth and those that may require strategic adjustments or further investment.
* Identifying Top-Selling Products: By examining the product keys and their corresponding sales amounts, executives can identify the top-selling products within each product line. This information can guide marketing and sales efforts, inventory management, and resource allocation.

Overall, the provided query enables business executives to analyze sales performance, product line profitability, and identify opportunities for growth and improvement within the AdventureWorks data warehouse.

SELECT c.CurrencyName, SUM(s.SalesAmount) AS TotalSalesAmount, c.CurrencyName AS Expr1

FROM FactInternetSales AS s INNER JOIN

DimCurrency AS c ON s.CurrencyKey = c.CurrencyKey

GROUP BY c.CurrencyName

* Total Sales Amount by Currency: The query calculates the total sales amount for each currency used in the AdventureWorks data warehouse. Executives can analyze this information to understand the contribution of different currencies to the overall sales revenue. They can identify which currencies are generating the highest sales amounts, indicating potential markets or customer segments with strong purchasing power.
* Currency Performance Comparison: By grouping the results by currency name, the query allows executives to compare the total sales amounts across different currencies. This analysis helps identify currencies that are driving significant sales volume and those that may require further evaluation or strategic consideration. Executives can focus their attention on currencies with higher sales amounts to identify factors contributing to their success, such as favorable exchange rates or targeted marketing efforts.
* Identifying Currency Preferences: The query joins the FactInternetSales and DimCurrency tables based on the currency key. This enables executives to understand customer preferences for specific currencies when making purchases. By examining the total sales amount associated with each currency, they can identify which currencies are commonly used by customers, potentially indicating target markets or regions where specific currencies hold prominence.
* Currency Impact on Sales Revenue: By analyzing the correlation between currency and sales amounts, executives can assess the impact of currency fluctuations on sales revenue. They can identify currencies that have a strong influence on sales performance and monitor changes in exchange rates or economic conditions that may affect customer buying behavior. This analysis can help in developing strategies to mitigate risks associated with currency fluctuations and optimize revenue generation.

Overall, the provided query enables business executives to analyze sales performance across different currencies in the AdventureWorks data warehouse. They can gain insights into currency preferences, identify high-performing currencies, and evaluate the impact of currency fluctuations on sales revenue. This information can guide decision-making related to market expansion, pricing strategies, and risk management in international markets.

SELECT p.ProductLine, SUM(s.OrderQuantity) AS TotalQuantitySold

FROM FactInternetSales AS s INNER JOIN

DimProduct AS p ON s.ProductKey = p.ProductKey

GROUP BY p.ProductLine

ORDER BY TotalQuantitySold DESC

* Product Line Performance: The query calculates the total quantity of products sold for each product line in the AdventureWorks data warehouse and presents the results in descending order based on the total quantity sold. Executives can use this information to assess the performance of different product lines and identify the top-selling product lines. They can focus their attention on product lines with the highest total quantity sold, indicating strong customer demand and potential areas for further investment or marketing efforts.
* Trend Analysis: Executives can track the quantity of products sold across different product lines over time by comparing the results of this query with historical data. By monitoring the trends in total quantity sold for each product line, they can identify patterns or shifts in customer preferences. This analysis can inform strategic decisions such as inventory management, product line expansions or adjustments, and targeted marketing campaigns to maximize sales and profitability.

Overall, the provided query allows business executives to analyze the quantity of products sold by product line in the AdventureWorks data warehouse. They can evaluate product line performance, identify top-selling product lines, and track trends to make data-driven decisions that optimize sales and drive business growth.

SELECT c.CustomerKey, c.FirstName, c.LastName, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimCustomer AS c ON s.CustomerKey = c.CustomerKey

GROUP BY c.CustomerKey, c.FirstName, c.LastName

ORDER BY TotalSalesAmount DESC

* Top Performing Customers: The query calculates the total sales amount for each customer in the AdventureWorks data warehouse and presents the results in descending order based on the total sales amount. Executives can identify the top-performing customers by analyzing the results. They can focus their attention on customers with the highest total sales amounts, indicating valuable and profitable customer relationships. This information can help in understanding customer preferences, loyalty, and identifying opportunities for targeted marketing campaigns or personalized customer experiences.
* Customer Segmentation: Executives can use the query results to segment customers based on their sales performance. By grouping customers based on their total sales amounts, they can identify different customer segments, such as high-value customers, regular customers, or potential customers who have shown significant sales potential. This segmentation can inform customer retention strategies, customer relationship management initiatives, and targeted sales and marketing approaches tailored to each segment's needs and preferences.

Overall, the provided query allows business executives to analyze customer sales performance in the AdventureWorks data warehouse. They can identify top-performing customers, understand customer segmentation, and leverage these insights to drive customer-focused strategies and maximize sales and customer satisfaction.

SELECT d.CalendarYear, d.EnglishMonthName, SUM(s.SalesAmount) AS MonthlySalesAmount, d.EnglishMonthName AS Expr1

FROM FactInternetSales AS s INNER JOIN

DimDate AS d ON s.OrderDateKey = d.DateKey

GROUP BY d.CalendarYear, d.EnglishMonthName, d.EnglishMonthName

ORDER BY d.CalendarYear, d.EnglishMonthName

* Monthly Sales Analysis: The query calculates the total sales amount for each month in each calendar year in the AdventureWorks data warehouse. Executives can use this information to analyze sales patterns and trends on a monthly basis. By examining the monthly sales amounts over different calendar years, they can identify seasonal variations, peak sales periods, or any noticeable changes in customer buying behavior. This analysis can help in planning marketing campaigns, optimizing inventory management, and making strategic decisions to maximize sales during specific months or periods.
* Yearly Performance Comparison: Executives can compare the total sales amounts across different calendar years using the results of this query. By examining the aggregated sales data for each calendar year, they can assess the overall performance and growth trends over time. This analysis helps in identifying years with significant sales growth, identifying areas for improvement in underperforming years, and gaining insights into the effectiveness of business strategies implemented in specific calendar years. It can guide decision-making related to resource allocation, goal setting, and performance evaluation.

Overall, the provided query allows business executives to analyze monthly sales amounts across different calendar years in the AdventureWorks data warehouse. They can identify sales patterns, seasonal variations, and make informed decisions to optimize sales performance and drive business growth.

SELECT st.SalesTerritoryGroup, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimSalesTerritory AS st ON s.SalesTerritoryKey = st.SalesTerritoryKey

GROUP BY st.SalesTerritoryGroup

* Sales Performance by Territory Group: The query calculates the total sales amount for each sales territory group in the AdventureWorks data warehouse. Executives can use this information to assess the sales performance of different territory groups and identify which groups contribute the most to the overall sales revenue. They can analyze the total sales amounts for each group and compare them to set targets or benchmarks to evaluate the effectiveness of sales strategies and identify areas for improvement or investment.
* Territory Group Comparison: By grouping the results by sales territory group, the query enables executives to compare the sales performance among different groups. This analysis helps identify the top-performing territory groups that generate significant sales amounts. Executives can focus their attention on these high-performing groups to understand the factors contributing to their success, such as effective sales strategies, strong customer relationships, or favorable market conditions. They can also identify underperforming territory groups that may require additional support or adjustments to drive sales growth.

Overall, the provided query allows business executives to analyze sales performance across different sales territory groups in the AdventureWorks data warehouse. They can assess the contribution of each group to the total sales revenue, compare performance among groups, and make informed decisions to optimize sales strategies and allocate resources effectively.

SELECT p.EnglishPromotionName, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimPromotion AS p ON s.PromotionKey = p.PromotionKey

GROUP BY p.EnglishPromotionName

ORDER BY TotalSalesAmount DESC

* Promotion Effectiveness: The query calculates the total sales amount for each promotion in the AdventureWorks data warehouse and presents the results in descending order based on the total sales amount. Executives can use this information to evaluate the effectiveness of different promotions in driving sales. They can identify the top-performing promotions based on the highest total sales amount and assess the return on investment for each promotion. This analysis helps in determining which promotions are most successful in attracting customers and generating revenue.
* Promotion Strategy and Planning: Executives can use the query results to inform their promotion strategy and planning. By analyzing the total sales amount for each promotion, they can identify trends and patterns in customer behavior. This information can guide decision-making related to future promotion planning, such as allocating resources to the most successful promotions, identifying areas for improvement in underperforming promotions, and designing targeted marketing campaigns to maximize sales impact.

Overall, the provided query enables business executives to analyze the sales performance of different promotions in the AdventureWorks data warehouse. They can evaluate promotion effectiveness, identify top-performing promotions, and use the insights gained to refine their promotion strategy and drive business growth.

SELECT d.CalendarYear, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimDate AS d ON s.OrderDateKey = d.DateKey

GROUP BY d.CalendarYear

* Sales Performance by Calendar Year: The query calculates the total sales amount for each calendar year in the AdventureWorks data warehouse. Executives can analyze this information to understand the revenue generated in different years. They can identify which calendar years had the highest total sales amounts, indicating periods of strong sales performance. This analysis can help in evaluating the growth trajectory of the business over time and identifying factors that contributed to successful sales years.
* Yearly Sales Trends: By grouping the results by calendar year, the query allows executives to examine sales trends over time. They can observe the variations in total sales amounts across different years and identify patterns or fluctuations in sales performance. This analysis can provide insights into the impact of external factors such as economic conditions, marketing strategies, or product launches on sales performance. Executives can use this information to make informed decisions, such as adjusting marketing campaigns or allocating resources based on the performance trends observed in different calendar years.

Overall, the provided query enables business executives to analyze sales performance based on calendar years in the AdventureWorks data warehouse. They can assess the revenue generated in different years, identify top-performing years, and observe sales trends over time. This information can guide strategic decision-making, sales forecasting, and resource allocation to optimize business performance.

SELECT d.DayNumberOfWeek, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimDate AS d ON s.OrderDateKey = d.DateKey

GROUP BY d.DayNumberOfWeek

ORDER BY d.DayNumberOfWeek

* Day of the Week Sales Analysis: The query calculates the total sales amount for each day of the week in the AdventureWorks data warehouse and presents the results in ascending order based on the day number of the week. Executives can use this information to analyze sales patterns and trends based on specific days of the week. They can identify which days consistently generate higher sales and determine if there are any notable variations in sales performance across different days. This analysis can help in optimizing staffing, inventory management, and marketing efforts to align with peak sales days and capture potential opportunities for increasing sales.
* Weekly Sales Comparison: Executives can compare the total sales amounts across different days of the week to identify patterns or trends in weekly sales. By examining the sales amounts in ascending order based on the day number of the week, they can determine which days have the highest and lowest sales volumes. This analysis can guide decision-making related to sales promotions, pricing strategies, and resource allocation. For example, if certain days consistently have lower sales, executives may consider implementing targeted marketing campaigns or incentives to boost sales on those specific days.

Overall, the provided query enables business executives to analyze sales performance based on the day of the week in the AdventureWorks data warehouse. They can identify sales patterns, optimize operations for peak sales days, and make data-driven decisions to enhance overall sales effectiveness and drive business growth.

SELECT p.ProductLine, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimProduct AS p ON s.ProductKey = p.ProductKey

GROUP BY p.ProductLine

ORDER BY TotalSalesAmount DESC

* Product Line Revenue: The query calculates the total sales amount for each product line in the AdventureWorks data warehouse and presents the results in descending order based on the total sales amount. Executives can use this information to identify the product lines that generate the highest revenue. By focusing on the product lines with the highest total sales amounts, executives can allocate resources and develop strategies to further maximize sales and profitability in those specific product lines.
* Performance Comparison: Executives can compare the total sales amounts across different product lines to gain insights into their relative performance. By analyzing the revenue contribution of each product line, they can identify areas of strength and areas that may require improvement. This analysis can help executives allocate resources, prioritize marketing efforts, and make strategic decisions regarding product line expansion, product positioning, or pricing strategies.

Overall, the provided query allows business executives to analyze the sales performance of different product lines in the AdventureWorks data warehouse. They can identify top-selling product lines, assess their revenue contribution, and make data-driven decisions to optimize sales and drive business growth.

SELECT st.SalesTerritoryRegion, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimSalesTerritory AS st ON s.SalesTerritoryKey = st.SalesTerritoryKey

GROUP BY st.SalesTerritoryRegion

ORDER BY TotalSalesAmount DESC

* Sales Territory Performance: The query calculates the total sales amount for each sales territory in the AdventureWorks data warehouse and presents the results in descending order based on the total sales amount. Executives can use this information to assess the performance of different sales territories and identify the top-performing territories. They can focus their attention on territories with the highest total sales amount, indicating strong revenue generation and potential areas for further investment or expansion.
* Comparative Analysis: By comparing the total sales amounts across different sales territories, executives can gain insights into the relative performance of each territory. They can identify territories that are driving significant sales revenue and those that may require further evaluation or strategic attention. Executives can analyze the factors contributing to the success of high-performing territories, such as customer demographics, market dynamics, or effective sales strategies, and replicate those practices in other territories to drive overall sales growth.

Overall, the provided query allows business executives to analyze sales performance across different sales territories in the AdventureWorks data warehouse. They can evaluate territory performance, identify top-performing territories, and conduct comparative analysis to inform decision-making related to resource allocation, sales strategies, and target market focus. This information enables executives to optimize sales efforts, maximize revenue, and drive business success.

SELECT d.FullDateAlternateKey, SUM(s.SalesAmount) AS TotalSalesAmount

FROM FactInternetSales AS s INNER JOIN

DimDate AS d ON s.OrderDateKey = d.DateKey

GROUP BY d.FullDateAlternateKey

ORDER BY d.FullDateAlternateKey

* Sales Trend Analysis: By grouping the sales data by the full date alternate key, the query allows executives to analyze the total sales amount over different dates. They can observe the trends in sales over time, identifying periods of high or low sales activity. This analysis enables executives to understand the seasonality of sales, identify peak sales periods, and detect any notable patterns or fluctuations in sales performance.
* Performance Evaluation: Executives can use the query results to evaluate the overall sales performance based on the full date alternate key. By ordering the results in ascending or descending order, they can identify the dates that contribute the most to the total sales amount. This information helps executives to focus their attention on specific dates or periods that generate significant sales revenue, enabling them to make informed decisions regarding resource allocation, promotional activities, or strategic planning.

Overall, the provided query allows business executives to analyze the total sales amount over time in the AdventureWorks data warehouse. They can gain insights into sales trends, evaluate performance based on specific dates, and utilize this information to make data-driven decisions that optimize sales strategies, allocate resources effectively, and drive business growth.

CONCLUSION

By creating a data warehouse with AdventureWorks, we are able to consolidate and integrate data from various tables within the database, providing a solid foundation for reporting and analysis. The dimensional model and optimized structure of the data warehouse allow for efficient querying and reporting, empowering business users to gain valuable insights and make informed decisions based on the data.