E-commerce Data Analysis Using SQL Server 2024

Enhancing Performance and Insights

Created by: Lama Ibrahim Al-Marhabi

Introduction:

Project Story

In the era of digital evolution, e-commerce is no longer an optional add-on but has become an essential part of everyday consumer life. As online platforms grow in number and variety, and as product offerings and promotions expand, competition intensifies, bringing new challenges for companies in this field. Among these challenges, understanding **customer behavior** stands out as a crucial factor – knowing what a customer prefers, when they are likely to buy, and their preferences is key to providing a personalized shopping experience that fosters loyalty. Additionally, **improving product quality** is essential for attracting new customers and retaining existing ones, while **evaluating the effectiveness of promotional offers** requires a precise understanding of how discounts and special offers influence purchasing decisions.

This led to the creation of the **E-commerce Data Analysis Project**, which utilizes advanced analytical tools like SQL. This project represents a vital step towards transforming massive amounts of raw data into actionable insights, enabling these insights to support more effective decision-making and achieve a competitive edge. Here, the objective goes beyond static figures to become a practical guide for decision-makers, providing them with an in-depth understanding of customer needs and the dynamic demands of the market.

Project Objectives

The primary goal of this project is to **extract valuable insights** from e-commerce platform data.

The project aims to identify different **purchasing patterns** among customers, analyze **product ratings** to gauge customer satisfaction, and evaluate the **impact of promotional offers** on sales.

By analyzing customer behavior and purchase trends, this project provides a comprehensive view that aids in:

1. **Enhancing customer experience** by personalizing offers and products according to customer interests.
2. **Improving product quality** by monitoring ratings and identifying products that need development.
3. **Forecasting future demand** based on past purchasing patterns and seasonal trends analysis.
4. **Measuring the effectiveness of promotional offers** by understanding factors that lead to increased or decreased sales.

Data Environment and Database Creation

Database Structure

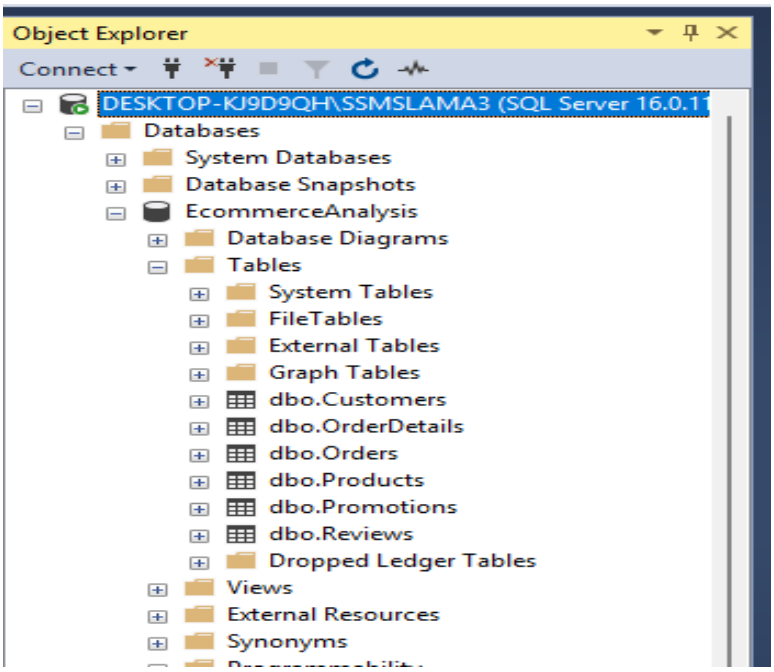
In order to conduct an effective analysis of e-commerce data, the database has been designed to include several foundational tables that encompass essential aspects of customer interactions and product performance. The primary tables in this project are **Customers**, **Products**, **Orders**, **Reviews**, and **Promotions**, each serving a specific role in capturing data critical to the e-commerce platform's performance.

1. **Customers Table**: This table stores customer-specific information, including unique identifiers, registration details, and geographical locations. Each customer has a unique `CustomerID` that links to their respective orders and reviews, allowing for customer behavior analysis across demographics.
2. **Products Table**: Here, details of each product, such as `ProductID`, name, category, price, and stock, are stored. The `ProductID` is a primary key that links to other tables, particularly `OrderDetails` and `Reviews`, enabling tracking of sales performance and customer satisfaction.
3. **Orders Table**: This table records each order placed by customers, including order dates and total amounts. The `OrderID` is a primary key that connects with `OrderDetails` and links to the `Customers` table through a foreign key, ensuring each order can be traced back to its respective customer.
4. **OrderDetails Table**: The OrderDetails table captures the specifics of each item within an order. Through foreign keys `OrderID` and `ProductID`, this table serves as an intermediary linking orders to products and detailing the quantity and unit price for each item purchased.
5. **Reviews Table**: To monitor product satisfaction, this table includes customer reviews, capturing the rating, review text, and review date. Each review is linked to the `Products` and `Customers` tables through `ProductID` and `CustomerID`, allowing for precise analysis of customer feedback per product.
6. **Promotions Table**: This table stores promotional campaigns applied to products, including discount percentages and start and end dates. Each promotion links back to specific products through `ProductID`, providing insights into the effects of discounts on product sales.

ER and EER Models

The database structure is visually represented through **Entity-Relationship (ER) Diagrams** and **Enhanced Entity-Relationship (EER) Diagrams** to provide a clear understanding of table relationships and data flow within the platform. The **ER Diagram** illustrates the fundamental relationships between entities, focusing on primary and foreign keys, while the **EER Diagram** expands on this by incorporating advanced features like inheritance hierarchies and specialization where applicable.

Analysis

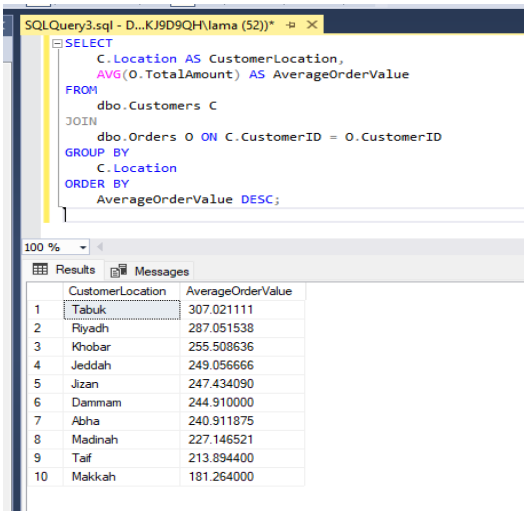


Customer Behavior Analysis

Analyzing customer behavior is critical for understanding purchasing patterns and preferences across different regions. By examining customer data, we can evaluate factors such as geographic location and average order value, which are valuable for optimizing marketing strategies and personalizing offers.

1. **Analysis by Geographic Location and Average Order Value**:

- This analysis provides insight into the purchasing habits of customers based on their location and the average value of their orders. Using SQL, we can group data by geographic region and calculate the average order value for each location.

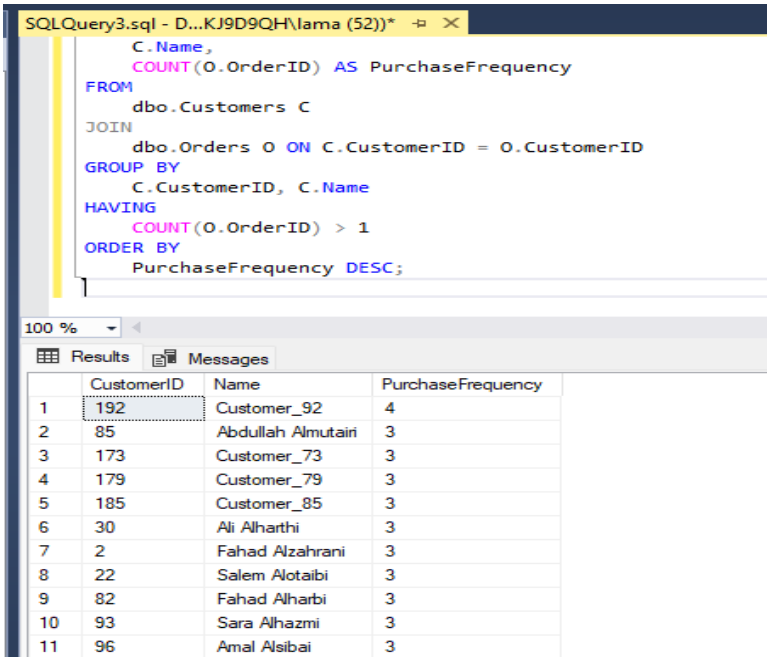


	CustomerLocation	AverageOrderValue
1	Tabuk	307.021111
2	Riyadh	287.051538
3	Khobar	255.508636
4	Jeddah	249.056666
5	Jizan	247.434090
6	Dammam	244.910000
7	Abha	240.911875
8	Madinah	227.146521
9	Taif	213.894400
10	Makkah	181.264000

This query outputs the average order value per location, helping to identify high-value regions and tailor marketing campaigns accordingly.

2. **Purchase Frequency and Loyalty Analysis**:

- To assess customer loyalty, we can analyze purchase frequency, identifying customers with recurring purchases. Loyal customers are valuable assets, so tracking their behavior helps refine retention strategies.



	CustomerID	Name	PurchaseFrequency
1	192	Customer_92	4
2	85	Abdullah Almutairi	3
3	173	Customer_73	3
4	179	Customer_79	3
5	185	Customer_85	3
6	30	Ali Alharthi	3
7	2	Fahad Alzahrani	3
8	22	Salem Alotaibi	3
9	82	Fahad Alharbi	3
10	93	Sara Alhazmi	3
11	96	Amal Alsibai	3

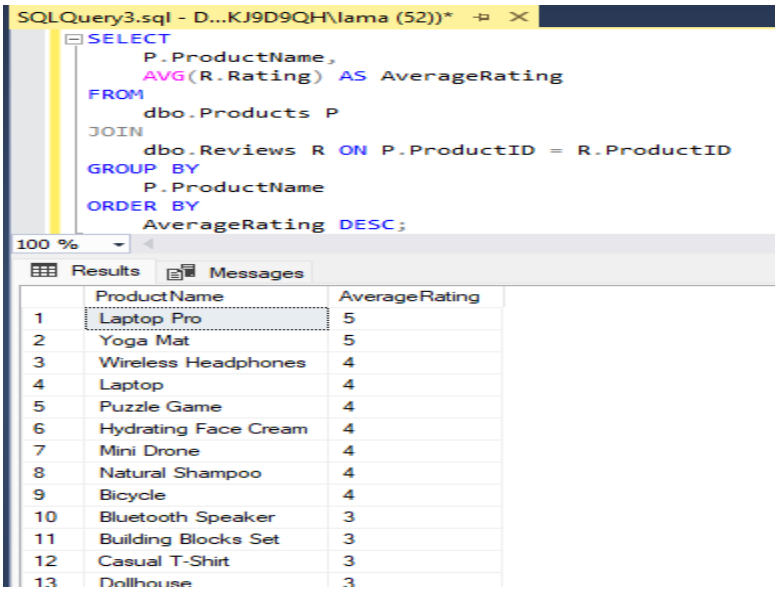
This query lists customers with more than one order, indicating repeat buyers. High purchase frequency can be used to identify loyal customers and develop targeted reward programs.

Product Ratings Analysis

To maintain product quality and improve customer satisfaction, it's essential to analyze product ratings and customer feedback.

1. **Average Product Rating**:

- Calculating the average rating for each product provides insight into overall customer satisfaction and highlights products that might require quality improvements.



	ProductName	AverageRating
1	Laptop Pro	5
2	Yoga Mat	5
3	Wireless Headphones	4
4	Laptop	4
5	Puzzle Game	4
6	Hydrating Face Cream	4
7	Mini Drone	4
8	Natural Shampoo	4
9	Bicycle	4
10	Bluetooth Speaker	3
11	Building Blocks Set	3
12	Casual T-Shirt	3
13	Dollhouse	3

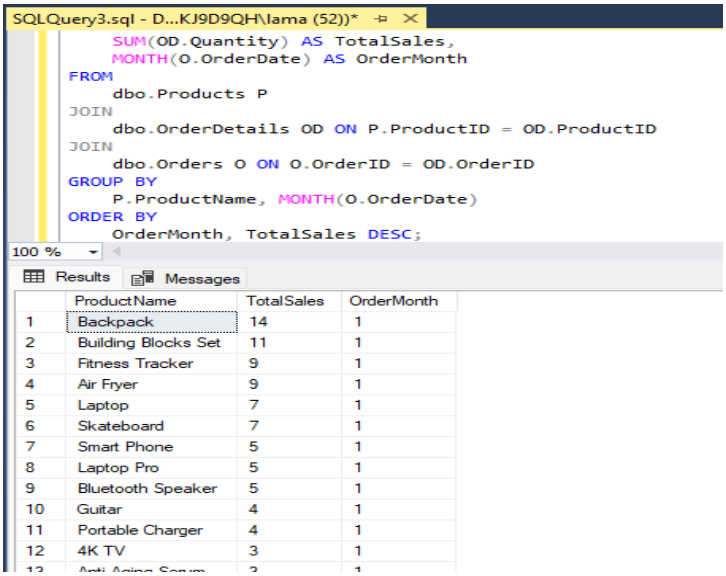
This query ranks products based on average customer ratings, allowing the identification of high-performing products and those in need of improvement.

Seasonal Demand Analysis and Future Demand Prediction

Seasonal demand analysis helps identify high-demand periods, allowing the company to optimize stock levels and marketing efforts during peak seasons.

1. **Monthly Sales Trends**:

- Analyzing monthly sales trends provides insight into product popularity across different times of the year. Grouping sales by month helps identify the most popular products for each period.



	ProductName	TotalSales	OrderMonth
1	Backpack	14	1
2	Building Blocks Set	11	1
3	Fitness Tracker	9	1
4	Air Fryer	9	1
5	Laptop	7	1
6	Skateboard	7	1
7	Smart Phone	5	1
8	Laptop Pro	5	1
9	Bluetooth Speaker	5	1
10	Guitar	4	1
11	Portable Charger	4	1
12	4K TV	3	1

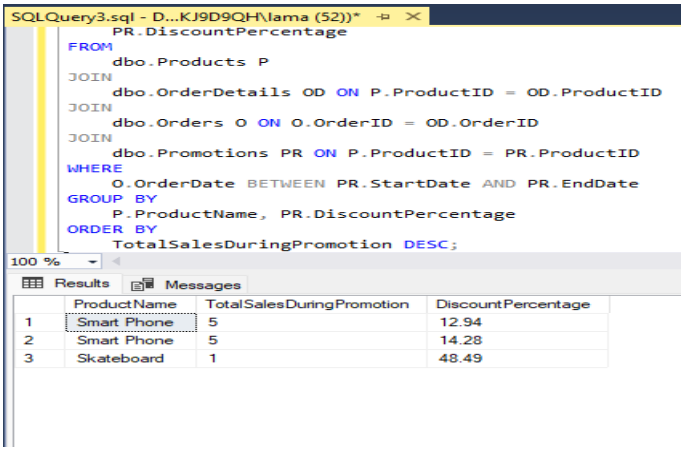
This query lists total sales per product by month, providing insights into seasonal demand trends that support inventory planning and promotional efforts.

Promotion Effectiveness Analysis

To maximize the impact of promotions, analyzing how discounts influence sales is essential. Tracking sales during promotional periods helps determine if promotions effectively increase demand.

1. **Sales During Promotion Periods**:

- This query calculates total sales for products under promotion, providing insights into the effectiveness of different discount levels.

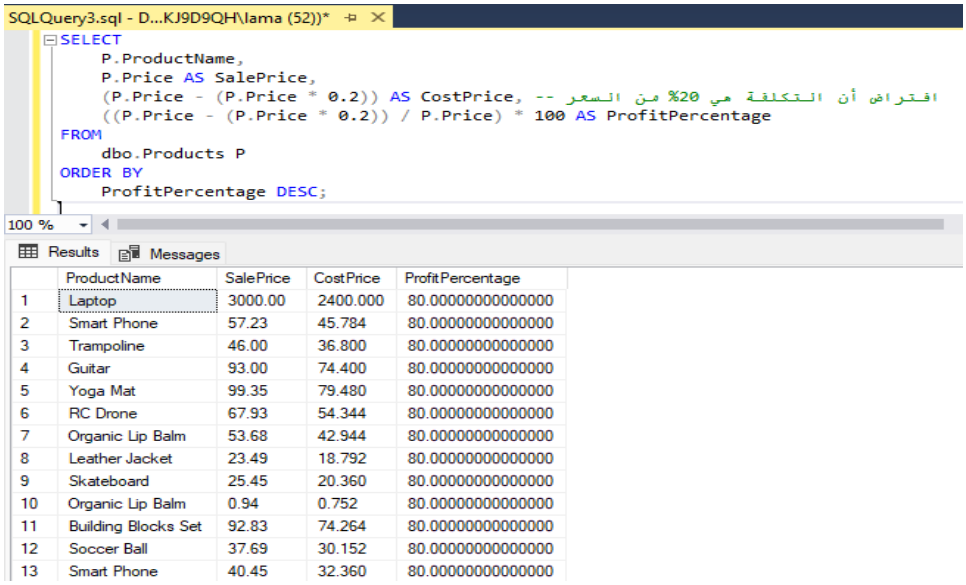


	ProductName	TotalSalesDuringPromotion	DiscountPercentage
1	Smart Phone	5	12.94
2	Smart Phone	5	14.28
3	Skateboard	1	48.49

This query helps in assessing the promotional impact on sales volume, allowing management to refine discount strategies based on demand.

These SQL queries support in-depth analysis across multiple dimensions, from customer behavior to product quality, seasonal demand, and promotion impact, equipping the business with actionable insights for data-driven decision-making.

The query provided aims to calculate the **profit margin** for each product in the database based on the assumed sale price and an estimated cost for the product. In this query, it is assumed that the cost of the product is **20% of the sale price**.



	ProductName	SalePrice	CostPrice	ProfitPercentage
1	Laptop	3000.00	2400.000	80.00000000000000
2	Smart Phone	57.23	45.784	80.00000000000000
3	Trampoline	46.00	36.800	80.00000000000000
4	Guitar	93.00	74.400	80.00000000000000
5	Yoga Mat	99.35	79.480	80.00000000000000
6	RC Drone	67.93	54.344	80.00000000000000
7	Organic Lip Balm	53.68	42.944	80.00000000000000
8	Leather Jacket	23.49	18.792	80.00000000000000
9	Skateboard	25.45	20.360	80.00000000000000
10	Organic Lip Balm	0.94	0.752	80.00000000000000
11	Building Blocks Set	92.83	74.264	80.00000000000000
12	Soccer Ball	37.69	30.152	80.00000000000000
13	Smart Phone	40.45	32.360	80.00000000000000

Through this query, you can identify the products that generate the highest profit margin, which helps make business decisions such as focusing on promoting high-margin products or adjusting the prices of products with lower profit margins

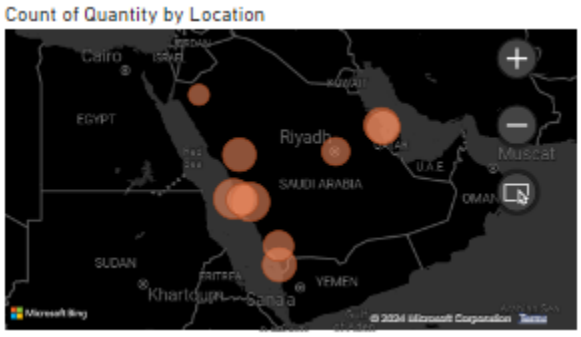
Data Visualization Using Power BI

In this analysis, Power BI was utilized to visualize data from an e-commerce database, which includes information about products, customers, orders, revenue, and payment methods. The visualization aims to provide clear insights that facilitate understanding business performance, identifying purchasing trends, and analyzing customer behavior to support strategic decision-making.

Analytical Objective:

This analysis aims to identify **geographical patterns of demand and distribution**, which helps in:

- Identifying high-demand areas to improve logistical planning.
- Exploring low-demand regions and developing strategies to boost sales.
- Supporting strategic decision-making based on accurate data regarding geographical distribution.

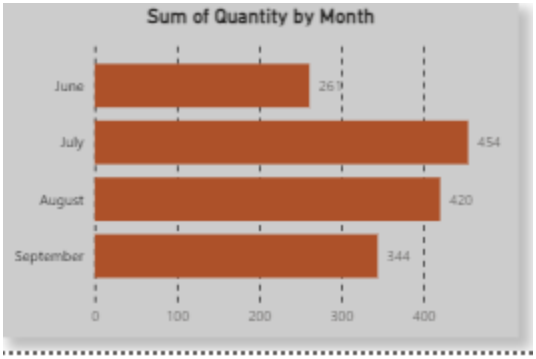


Analytical Objective for the Chart:

This analysis aims to **understand the monthly distribution of quantities** by comparing the total quantities across different months. The bar chart provides clear insights into the fluctuations in demand or distribution over time.

Key Analytical Objectives:

1. **Identifying seasonal patterns:** Helps in understanding time periods with higher or lower demand, such as the noticeable peak in July and the decline in June and September.
2. **Logistical planning:** Assists in optimizing inventory and resource management based on months with high or low demand.
3. **Strategic decision-making:** Supports identifying months that require marketing campaigns or discounts to boost sales during less active periods.
4. **Future forecasting:** Contributes to developing accurate demand forecasting strategies based on historical data.

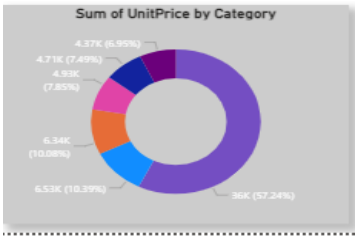


Analytical Objective for the Chart:

This chart aims to **analyze the total unit price by category**, highlighting the categories that contribute the most to overall revenue.

Key Analytical Objectives:

1. **Identifying top-performing categories:** The largest segment of the chart (57.24%) indicates the category that significantly contributes to revenue, helping focus on improving offerings in this category.
2. **Analyzing diversity:** Helps understand the distribution across different categories, showing whether revenue is dependent on specific categories or diversified.
3. **Making marketing decisions:** Guides marketing efforts toward less-performing categories to enhance their market presence.
4. **Resource allocation:** Supports better planning for resource allocation based on high and low-performing categories.

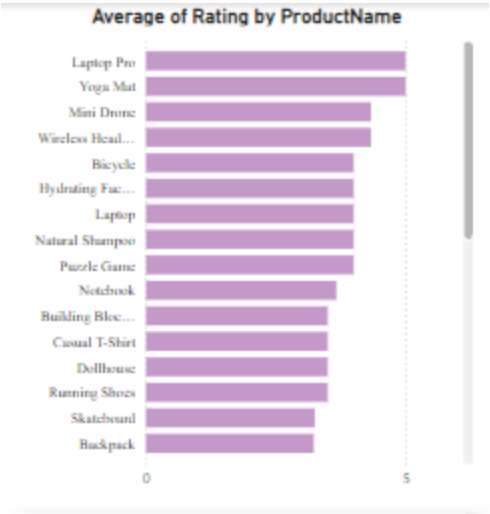


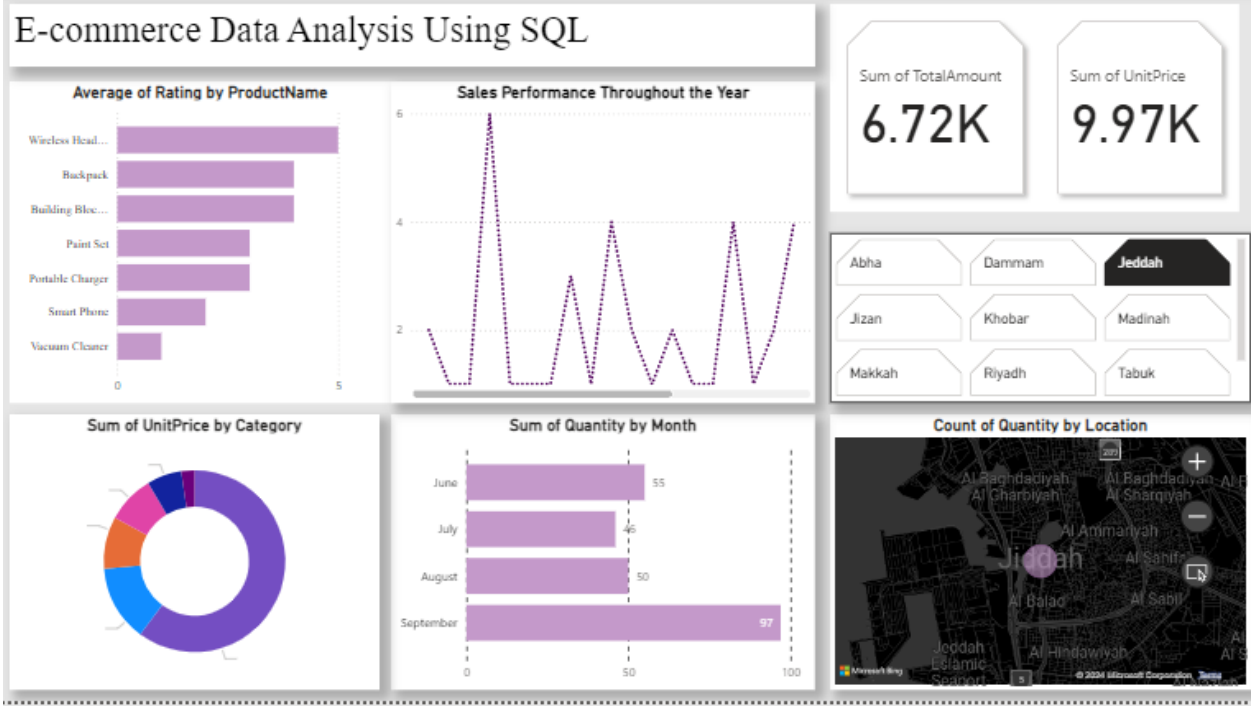
Analytical Objective for the Chart:

This chart aims to **analyze the average rating for each product**, helping assess customer satisfaction with various products.

Key Analytical Objectives:

1. **Identifying high-rated products:** Notable products such as "Laptop Pro" and "Yoga Mat" have the highest average ratings, indicating high customer satisfaction and product quality.
2. **Analyzing low performance:** Helps identify products with lower ratings, providing opportunities to improve their quality or associated services.
3. **Promoting successful products:** Guides businesses to focus on high-rated products in their marketing campaigns to boost sales.
4. **Enhancing customer experience:** Helps understand customer needs and preferences based on ratings to improve the overall experience.





Imagine you're managing an e-commerce business in Saudi Arabia and need a comprehensive view of your performance to make informed decisions. This dashboard tells the story of your business performance, starting with customer feedback, moving through sales trends, and diving into category and location-specific insights.

1. **Customer Satisfaction (Top Left)**:
   The "Average of Rating by ProductName" bar chart reveals customer sentiment toward various products. Items like "Wireless Headphones" and "Backpack" stand out as favorites with high ratings, signaling strong customer satisfaction and potential for further promotion. However, products like "Vacuum Cleaner" show room for improvement, inviting a closer look into customer needs.
2. **Sales Over Time (Center Top)**:
   The "Sales Performance Throughout the Year" line graph highlights peaks and valleys in monthly sales. For example, a noticeable spike suggests seasonal demand, while dips might indicate opportunities for targeted marketing campaigns or special offers to drive revenue.

3. **Category Breakdown (Bottom Left)**:
   The "Sum of UnitPrice by Category" donut chart shows that one category dominates revenue generation, contributing over 57%. This insight enables you to allocate resources effectively while exploring strategies to boost the performance of less successful categories.
4. **Monthly Sales Trends (Bottom Center)**:
   The "Sum of Quantity by Month" bar chart reveals trends in product quantities sold. Notably, September leads in sales volume, while June lags behind, suggesting a potential opportunity to boost sales during slower months through promotional activities.
5. **Geographical Insights (Right Side)**:
   The "Count of Quantity by Location" map and location filters show Jeddah as a primary hub for sales activity. This insight helps refine distribution strategies and allocate marketing budgets more efficiently across regions.
6. **Key Performance Metrics (Top Right)**:
   Summaries of "Total Amount" and "Unit Price" provide quick snapshots of financial performance, offering a clear starting point for evaluating overall business success.

This dashboard not only provides a detailed view of your e-commerce operations but also empowers you to take targeted actions. By leveraging these insights, you can enhance customer satisfaction, optimize sales strategies, and drive growth in key areas, making data-driven decisions your competitive advantage.

Conclusion:

As this analytical journey comes to an end, we find ourselves with a comprehensive dashboard that tells the story of e-commerce performance with clarity and precision. From customer satisfaction reflecting product success, to understanding seasonal patterns that unveil new growth opportunities, from category analysis highlighting impactful resources to deep geographical insights guiding expansion strategies.

This analysis is not just about numbers and visuals; it's a testament to the power of data in shaping decisions. It provides the tools needed to steer your business toward greater success—whether by enhancing performance in key areas, improving underperforming products, or planning new strategies to capitalize on seasonal trends.

The story doesn't end here; it begins. With every step guided by these insights, new opportunities for growth and excellence unfold. Now, you have a clear roadmap inviting you to make bold, informed decisions rooted in data, paving the way for the next chapter of success. Data isn't just a tool; it's your strategic partner in the journey toward distinction.