Hachibul Islam

Capstone Project: Retail Analysis

**Overview:**

This Retail Analysis project aims to leverage data analytics and insights to enhance the operations, decision-making, and customer experience within a retail business.It focuses on four key areas: Sales Analysis , Product Analysis, Customer Analysis, and Demographic Analysis. In the modern retail landscape, data-driven decision-making is crucial for staying competitive and meeting customer demands. This project will involve the collection, storage, and analysis of various types of data to gain valuable insights that can inform strategies, optimize inventory management, and improve overall business performance.

**The Process:**

The process of gathering data from GitHub, analyzing it in SQL, and creating a Power BI dashboard for retail analysis involves several steps. Here's an overview of the process:

1) Data Collection from GitHub:

Identify and access the relevant GitHub repositories or data sources that contain retail market data and we extract it,

2) Data Transformation:

Prepare the data for analysis. This may involve cleaning, formatting, and restructuring the data to ensure it's suitable for analysis. For instance, we need to convert data formats, handle missing values, and remove duplicates.

3) Data Storage in SQL Database:

Set up a SQL database where you can store and manage our data. Create the necessary tables and schemas to accommodate the data you've extracted and transformed

4) Data Loading:

Load the cleaned and transformed data into the SQL database. We c use SQL commands ETL (Extract, Transform, Load), or scripts to accomplish this task.

5) Data Analysis in SQL:

Utilize SQL queries to perform various analyses on the data stored in our SQL database. The major analysis that we perform are:

a)Sales Analysis.

b)Product Analysis.

c)Customer Analysis.

d)Demographic Analysis.

6) Data Visualization in Power BI:

Power BI is a powerful data visualization and business intelligence tool that can help us present our retail analysis findings in a visually appealing and interactive way. Now we export the data to Power BI to create visualizations . Design and build a Power BI dashboard specifically tailored to our retail analysis project

**Objective:**

The objective of this project is to develop a Power BI dashboard using the Retail Database that provides comprehensive insights into the retail business's performance. The dashboard aims to facilitate data-driven decision-making, optimize sales strategies, and enhance customer experiences. This Retail Analysis project aims to leverage data analytics and insights to enhance the operations, decision-making, and customer experience within a retail business. In the modern retail landscape, data-driven decision-making is crucial for staying competitive and meeting customer demands. This project will involve the collection, storage, and analysis of various types of data to gain valuable insights that can inform strategies, optimize inventory management, and improve overall business performance . The analysis will focus on four key areas: Sales Analysis, Product Analysis, Customer Analysis, and Demographic Analysis. It will cover sales trends, product performance, customer behaviour, and customer demographics to offer a comprehensive understanding of the retail business The primary goal of the Power BI dashboard is to empower retail stakeholders with actionable insights. The dashboard will help identify top-selling products, peak sales periods, customer preferences, and target demographics. It will enable retailers to make informed decisions to boost sales, optimize inventory, and tailor marketing strategies.

**Significance:**

Retail analysis is a critical aspect of the retail industry that involves the systematic examination of various data and factors associated with a retail business. Its significance lies in the valuable insights and benefits it offers, which can greatly impact a retailer's success and sustainability. The Power BI dashboard will provide valuable insights into sales performance, product trends, customer behaviour, and demographic patterns. It will identify high-performing products, low-performing items, customer preferences, and opportunities for customer engagement and retention. Based on the analysis, the dashboard will recommend targeted marketing strategies, inventory optimization, and personalized customer experiences. Retail analysis empowers retailers to make informed and data-driven decisions. By analyzing sales data, customer behaviour, and market trends, retailers can understand what works and what doesn't, enabling them to make strategic choices that maximize profitability. Understanding customer preferences and behaviour through retail analysis allows retailers to personalize marketing, product offerings, and services. This, in turn, enhances the overall customer experience, leading to higher customer satisfaction and loyalty. One of the critical aspects of retail analysis is inventory management. Retailers can use data analysis to optimize inventory levels, reduce carrying costs, and minimize stock-outs or overstock situations. Retail analysis helps retailers determine optimal pricing strategies. By analyzing historical sales data and market competition, retailers can set prices that are competitive yet profitable .Analyzing the success of marketing campaigns and promotional efforts is crucial for retailers. Retail analysis provides insights into which marketing channels and strategies are most effective, allowing retailers to allocate resources more efficiently and achieve a higher return on investment (ROI). Staying competitive in the retail industry requires an understanding of market trends and competitors. Retail analysis can be used to assess the performance of employees, especially in customer-facing roles. For retailers looking to expand or open new locations, retail analysis provides insights into the demographics and buying behaviour of potential customer bases. This information helps in making informed decisions about where to open new stores or branches. The retail industry is constantly evolving due to technological advancements, market shifts, and changing customer expectations. Retail analysis allows retailers to adapt to these changes more effectively, ensuring long-term success in a highly competitive environment.

**Data Dictionary:**

The Entity-Relationship (ER) diagram for the Power BI Dashboard using the above database represents the relationships and interactions between various entities (tables) in the database. Here's an explanation of the ER diagram based on the provided tables:

Entities:

Offices Table

This table stores information about the different offices of the retail company, including the office code, city, phone number, address, state, country, postal code, and territory. Each office is uniquely identified by its office code.

Employees Table

The employees table holds data about the company's employees. It includes fields such as employee number (a unique identifier for each employee), last name, first name, extension, email address, office code (identifying the office where the employee works), reports to (the employee number of the person to whom the employee reports), and job title.

Customers Table

This table contains information about the retail company's customers. It includes fields like customer number (a unique identifier for each customer), customer name, contact last name, contact first name, phone number, address, city, state, postal code, country, sales representative employee number (identifying the employee responsible for the customer), and credit limit.

Products Table

The products table stores details about the various products sold by the retail company. It includes information such as the product code (a unique identifier for each product), product name, product line (categorization of the product), product scale, product vendor, product description, quantity in stock, buy price, and Manufacturer's Suggested Retail Price (MSRP).

Product Lines Table

This table is used to describe the different product lines available in the company's inventory. It includes fields such as the product line name, text description, HTML description (for web-based content), and an image (stored as a BLOB) to represent the product line visually.

Orders Table

The orders table stores data related to customer orders. It includes information like the order number (a unique identifier for each order), order date, required date, shipped date, order status (using an ENUM for predefined statuses), comments, and customer number (identifying the customer who placed the order).

Order Details Table

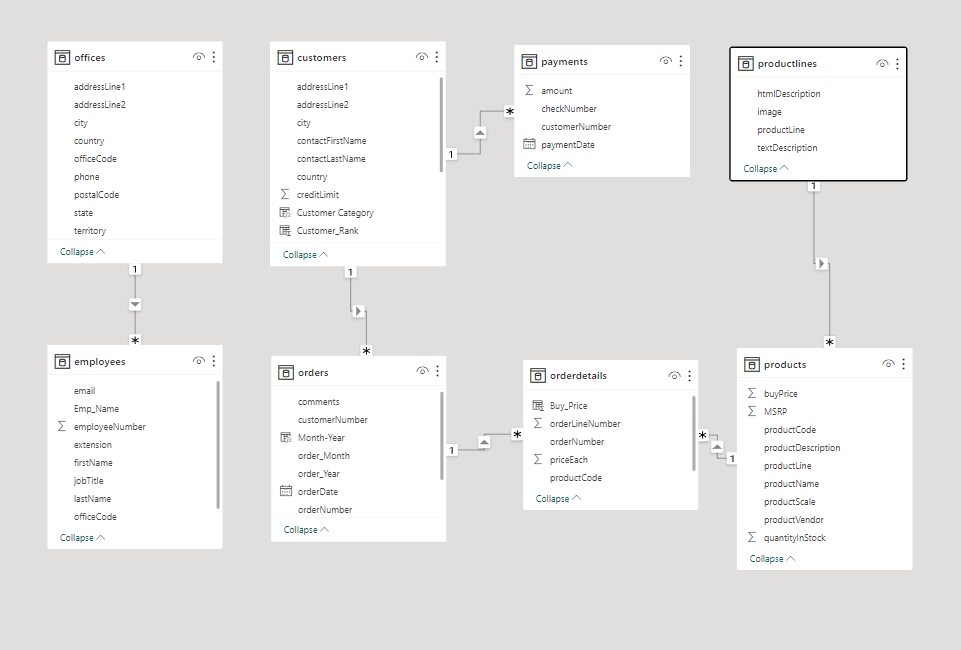
This table contains information about the individual items (line items) included in each order. It includes fields such as the order number (linking to the orders table), product code (linking to the products table), quantity ordered, price per item, and order line number.

Payments Table

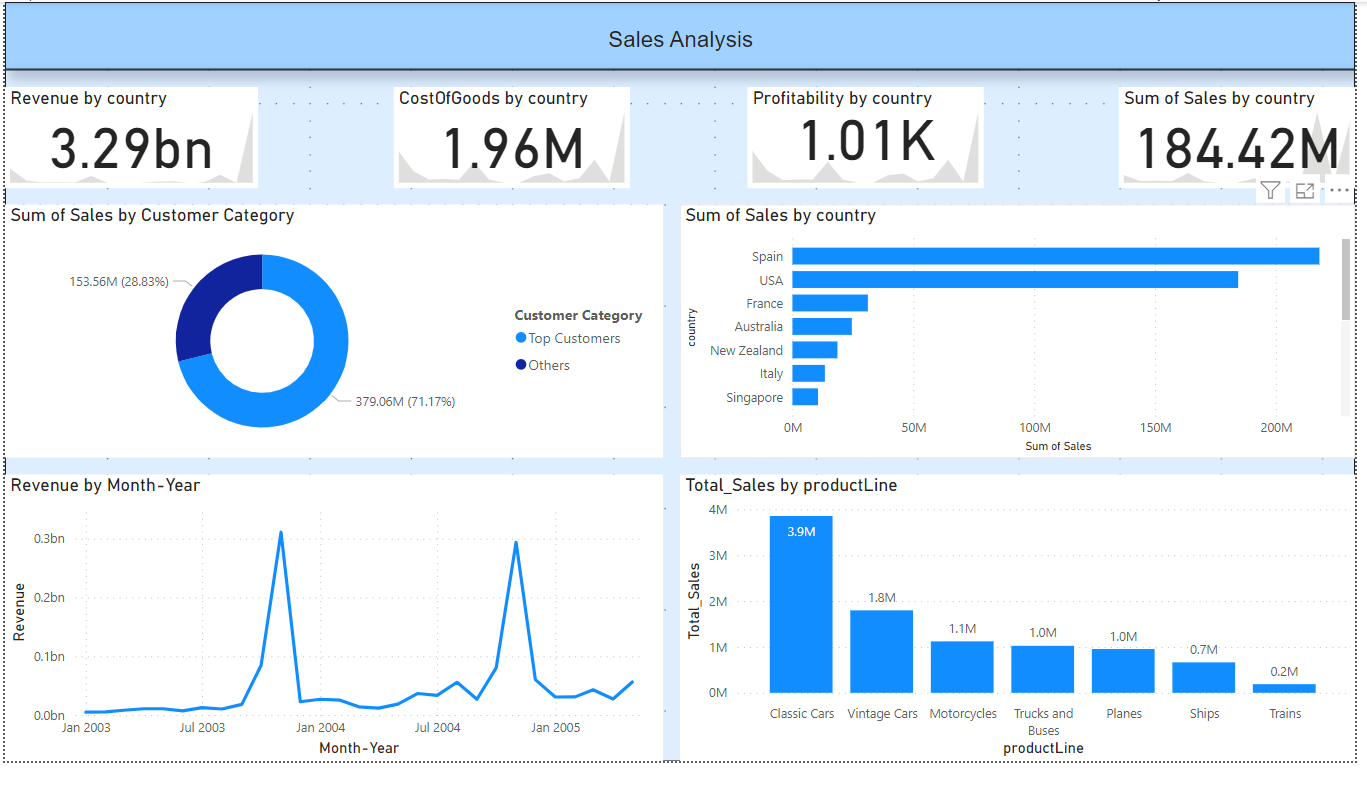
The payments table stores details about payments made by customers. It includes fields such as the customer number (linking to the customers table), check number, payment date, and the payment amount.

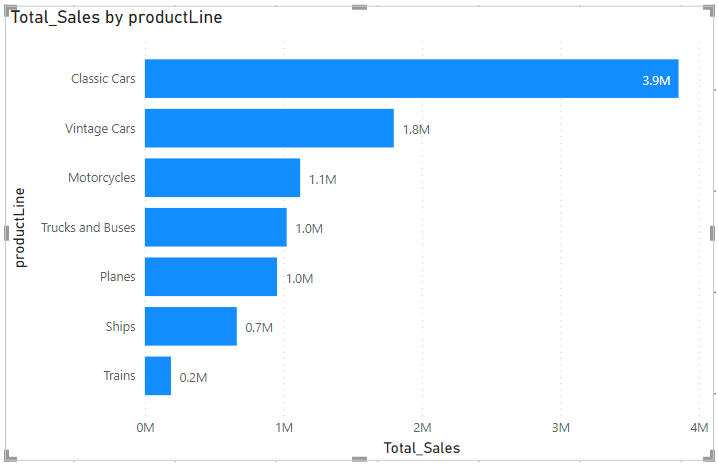
The ER diagram visually presents the structure and relationships of the database, facilitating better understanding and design of the Power BI Dashboard. It provides a clear representation of data entities and their connections, enabling effective data analysis and visualization within the Power BI Dashboard.

**ER Diagram:**



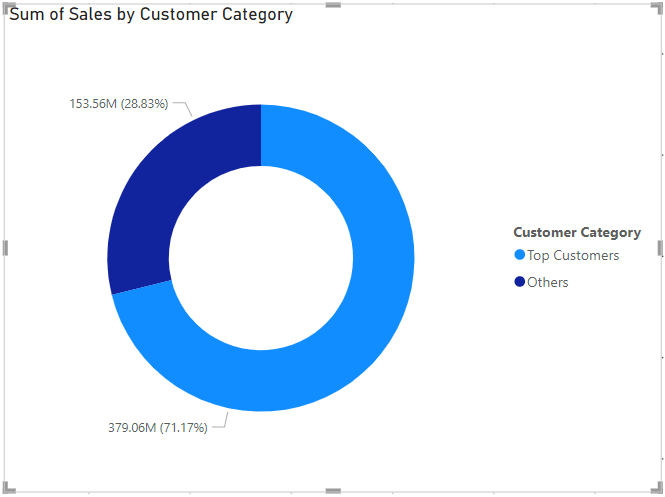
**Power BI Problem Statements**



Which product line has the highest sales:

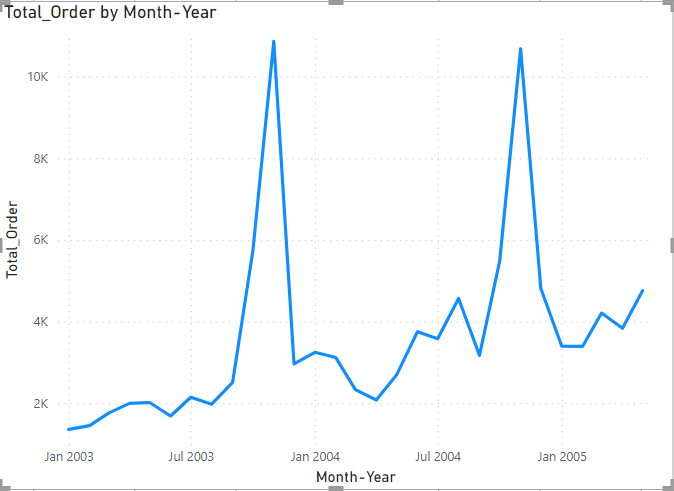
To visualize the highest sales in terms of product line in PowerBI , we create a visual that shows total sales by product line. By creating this visual we easily identify and analyze the highest sale in terms of product line .Here we see which product line contribute the most in terms of sales and we can use this information to make business decisions.

How does the sales performance of top customers compare to the rest:



This Donut Chat shows the sale performance of top customer compare the other customers, making it easy to see how the top customers are performing in relation to the over all customer base. We can also adjust the number of “top-customer” by changing the threshold in the DAX formula used to categorize customers.

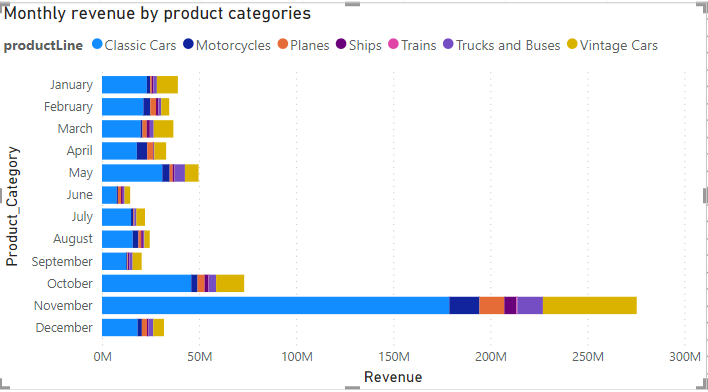
What is the trend in customer order volume over the past year:



To visualize the trend of customers order by month and year in PowerBI we create a visual that shows customer order by month-year.

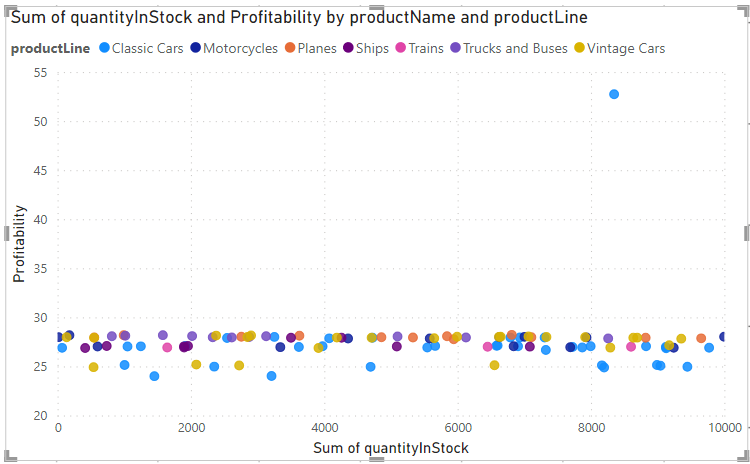
By creating these visuals, we can easily visualize and analyze the trend of customer order in different month and years, helping us gain insights into customer order base and make data-driven decisions related to marketing, product development and any other aspects of business.

How does monthly revenue vary across different product categories:



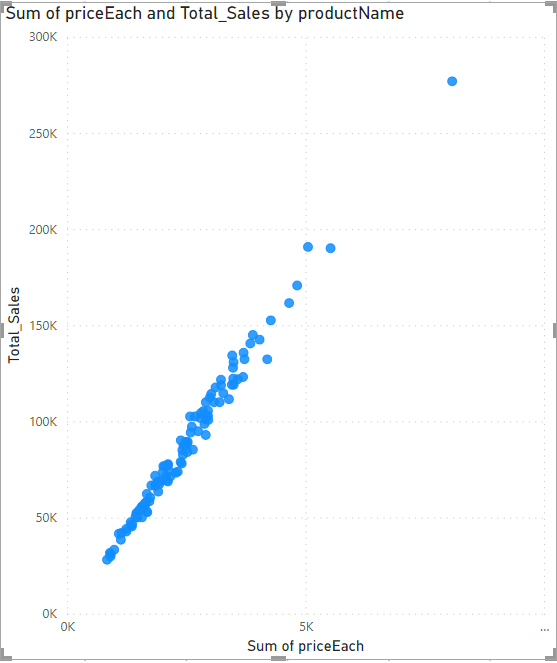
This is how monthly revenue vary according to Different product categories. Here we clearly see that in the month of November ,the monthly revenue is high as compare to other month as well as the revenue generated by classic cars is higher as compare to all the other product line.

How does the profitability of different products compare based on their quantity in stock

 To compare the Profitability of different products based on their quantity on stock in PowerBI .

We create a visual that shows the relationship between product profitability and quantity in stock and this visual allowing us to identify which products are more profitable relative to their stock level.

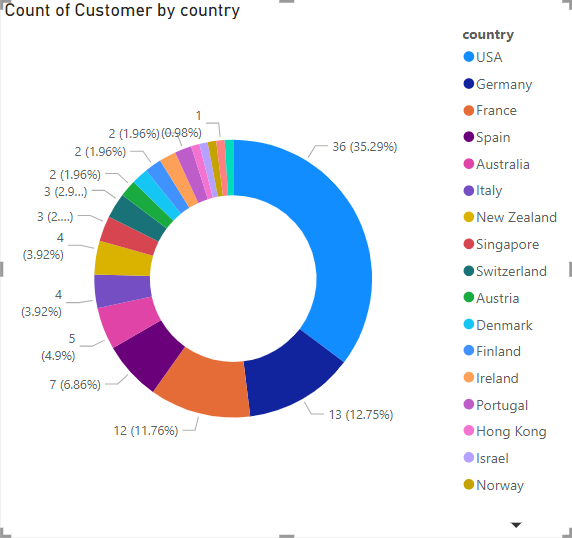
How does product pricing impact sales volume:



To analyze how product pricing impacts sales volume in PowerBI we can create a visual that shows the relationship between product price and sales volume.

By creating this visual we gain insights that help us to make informed decision to optimize our sales and revenue

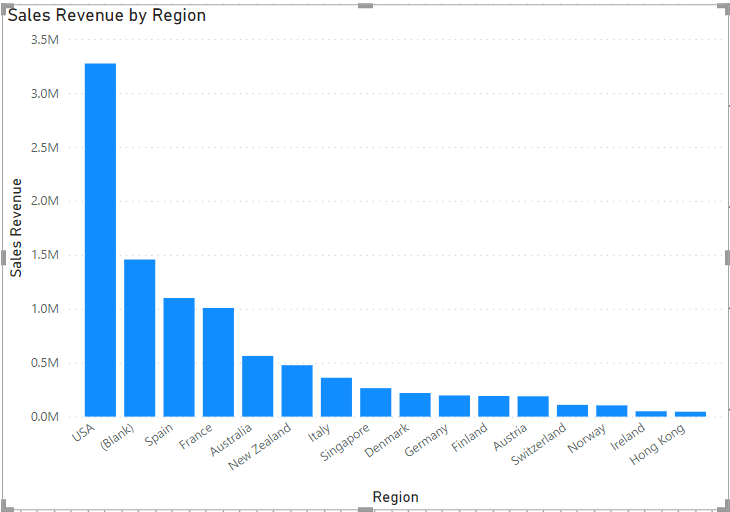
What is the distribution of customers across different demographic segments:



To visualize the distribution of customers across different demographic segments in PowerBI we create a visual that shows customer segmentation by demographics.

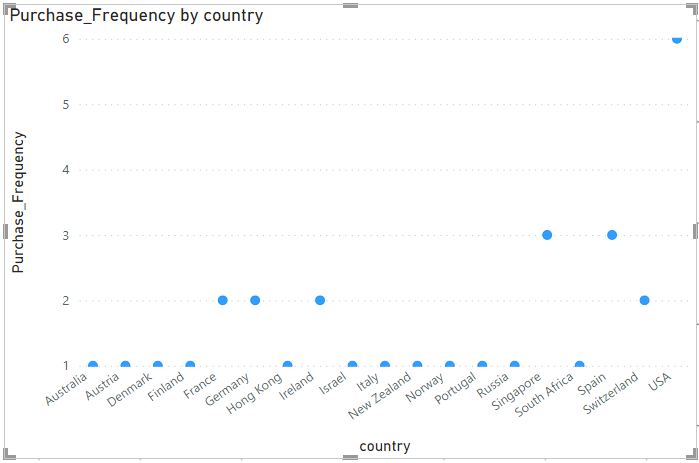
By creating these visuals, we can easily visualize and analyze the distribution of customer across different demographic segments, helping us gain insights into customer base and make data-driven decisions related to marketing, product development and any other aspects of business.

What are the top regions in terms of sales revenue:



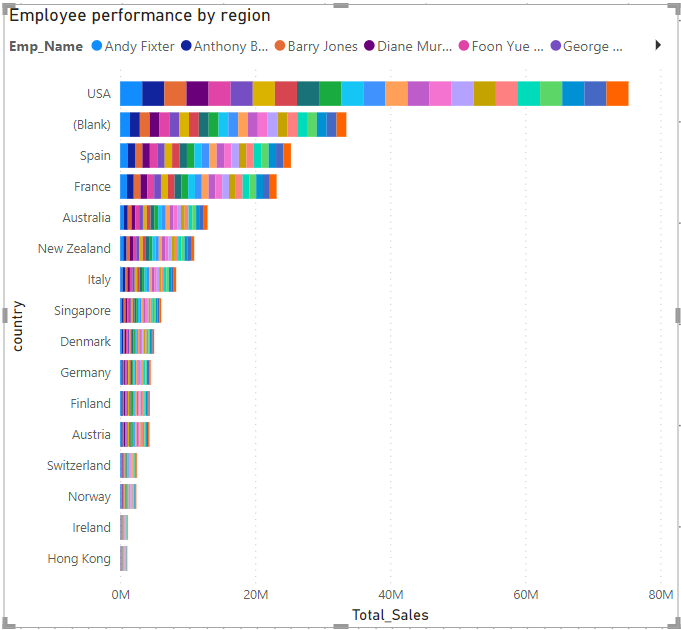
To visualize the top regions in terms of sales revenue in PowerBI, we create a visual that shows sales revenue by region.

By creating this visual we easily identify and analyze the top regions in terms of sales revenue.Here we see which regions contribute the most in terms of sales ,we can use this information to make business decisions.

What is the correlation between customer demographics and purchase frequency : To visualize the correlation between customer demographics and purchase frequency in PowerBI .

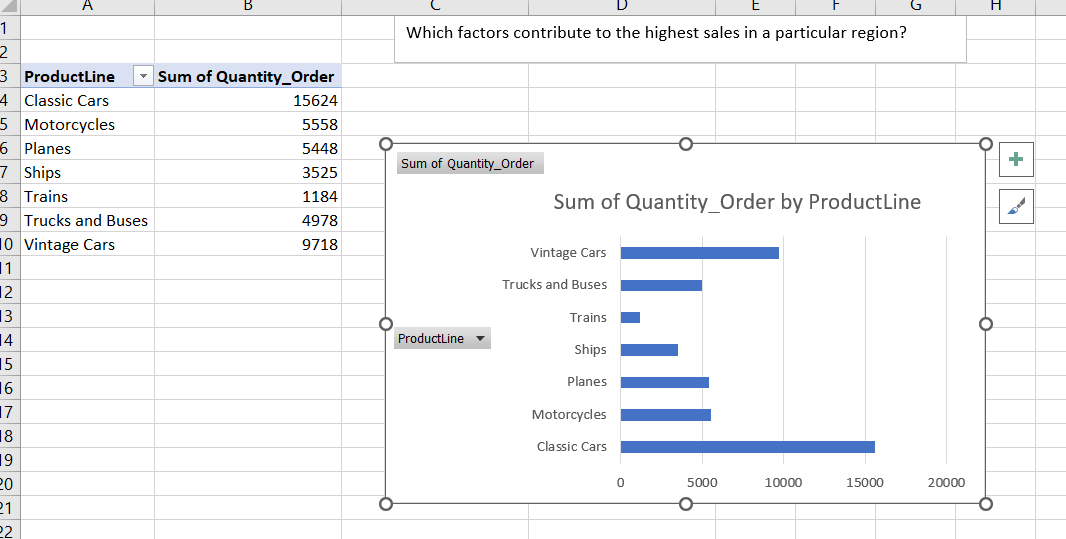
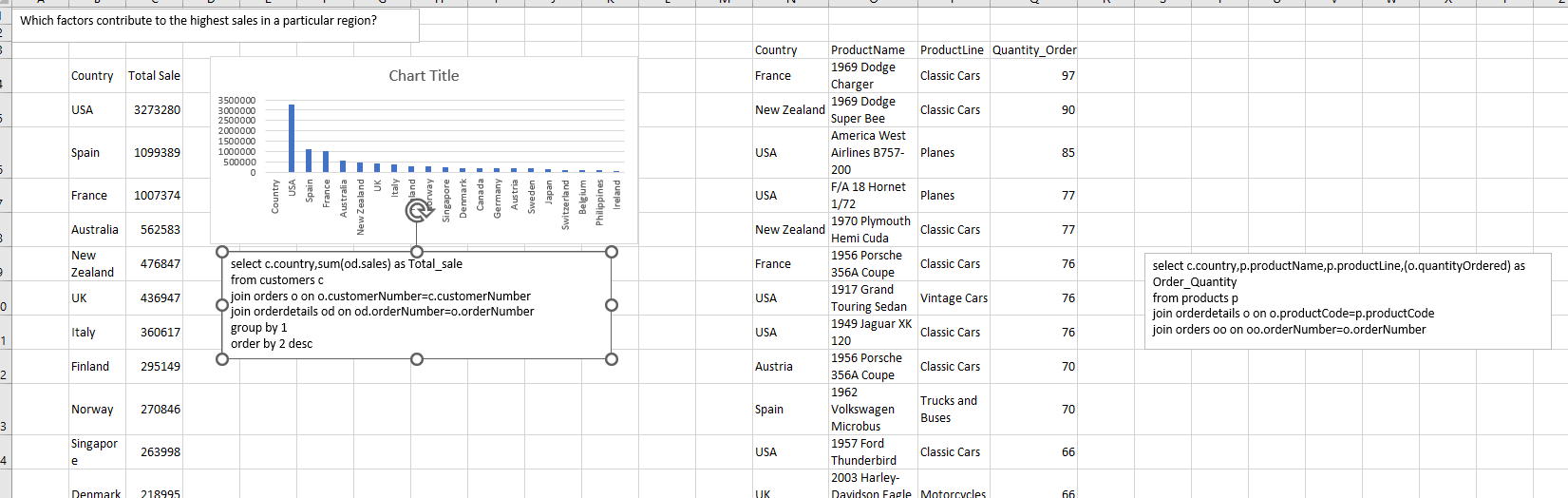
We create a visual that shows the relationship between customer demographics and purchase frequency and this visual allowing us to identify in which country the purchase frequency is higher as compare to the rest.

How does the performance of sales employees vary across different regions:

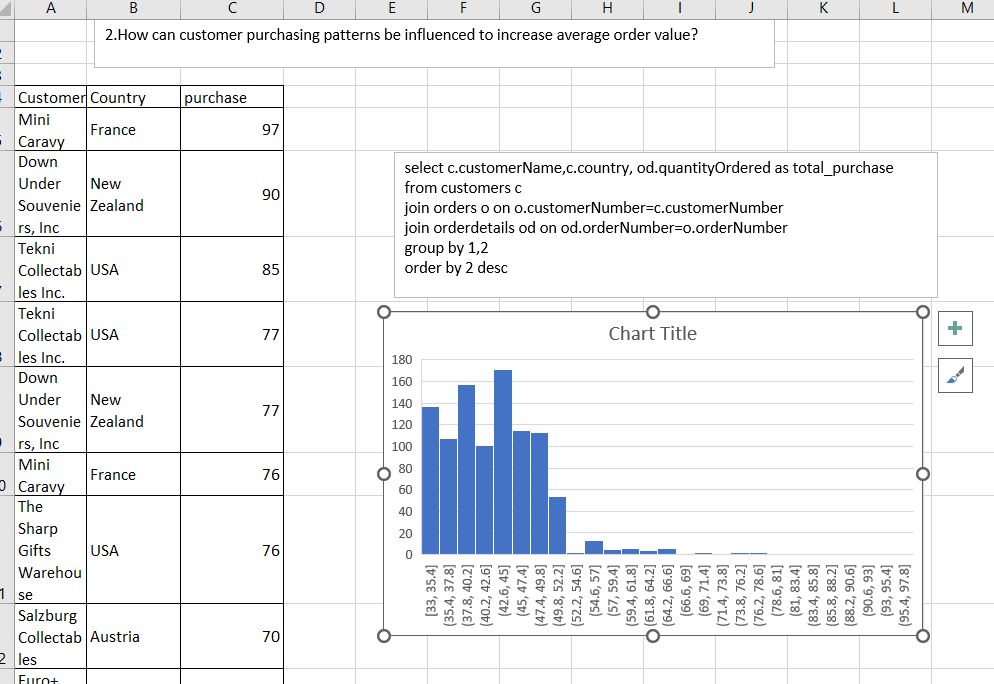
 To analyze how the performance of sales employees varies across diferrent regions in PowerBI, we create a visual that shows employee performance in terms of sales by region. By creating this visual we can easily identify top-performing employees in each regions and assess which regions may need additional support or improvement

**EDA Problem Statements:**

**Which factors contribute to the highest sales in a particular region:**

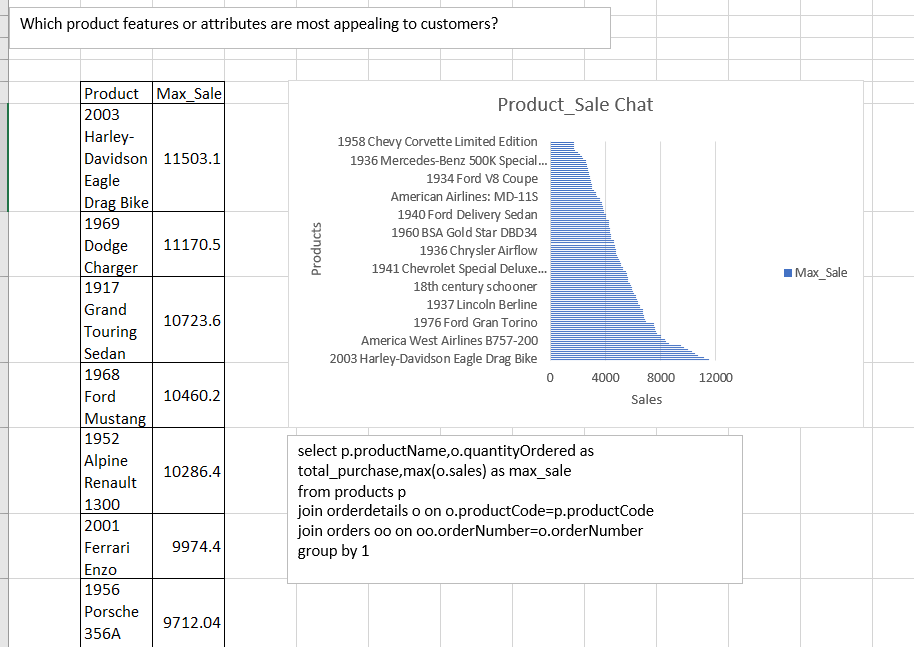
  The factors that contribute to the highest sales in a particular region can vary widely depending on the specific industry, business, and market conditions. However, here we consider some common factors that can influence sales in a particular region. Understanding the demographics of the region's population, such as age, gender, income level, and preferences, can help tailor products or services to better meet the needs of the local market.

How can customer purchasing patterns be influenced to increase average order value:



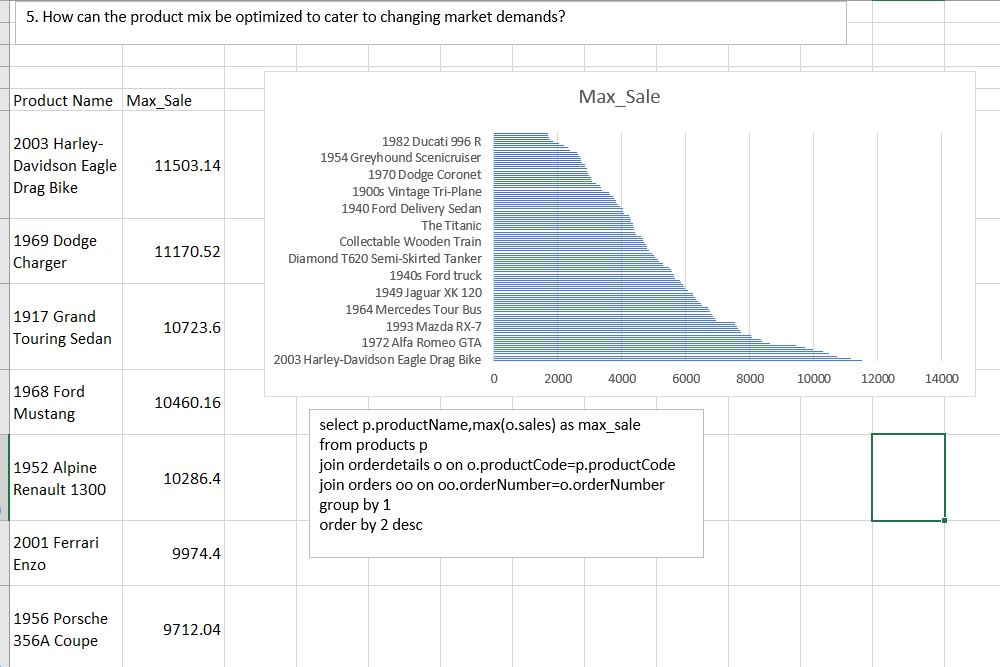
**Influencing customer purchasing patterns to increase the average order value (AOV) involves strategic planning and the implementation of various techniques. By combining strategies and tailoring them to our specific industry and target audience, you can effectively influence customer purchasing patterns and increase the average order value. Regularly analyze data and customer feedback to refine your approach and stay responsive to evolving market dynamics.**

Which product features or attributes are most appealing to customers:



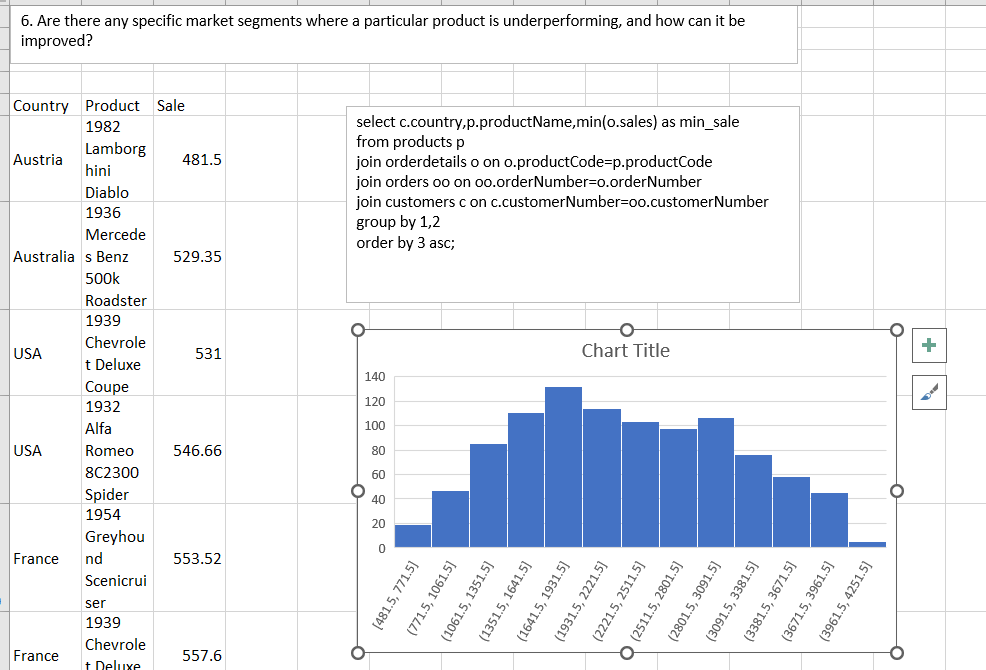
**The appeal of product features or attributes to customers can vary based on factors such as industry, target audience, and market trends. However, several common product features are often considered appealing to customers across different sectors. Understanding the preferences of the target audience, staying attuned to market trends, and incorporating relevant features and attributes into product development are key strategies for enhancing the appeal of products to customers.**

How can the product mix be optimized to cater to changing market demands:

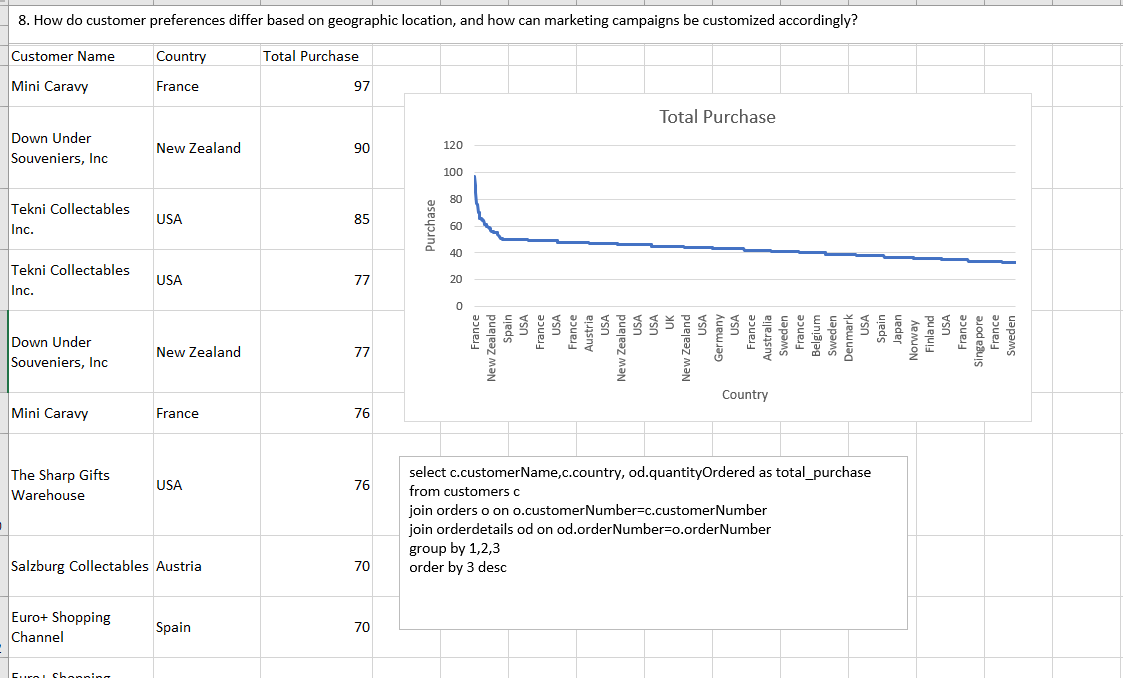


**Optimizing the product mix to cater to changing market demands is essential for staying competitive and meeting the evolving needs of consumers. By adopting a strategic and proactive approach to adjusting the product mix, businesses can not only meet changing market demands but also position themselves for sustained growth and competitiveness. Regularly reassessing the product mix in response to market dynamics is a fundamental aspect of successful business management.**

Are there any specific market segments where a particular product is underperforming, and how can it be improved:

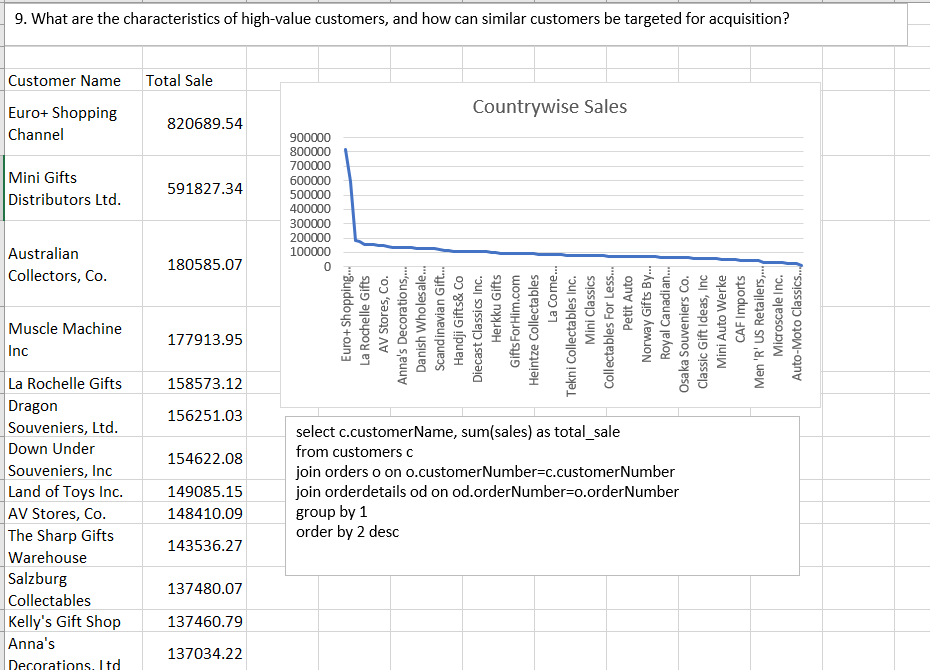


Identifying specific market segments where a particular product is underperforming and determining how to improve its performance often involves a combination of data analysis, market segmentation, and strategic planning.How do customer preferences differ based on geographic location, and how can marketing campaigns be customized accordingly:



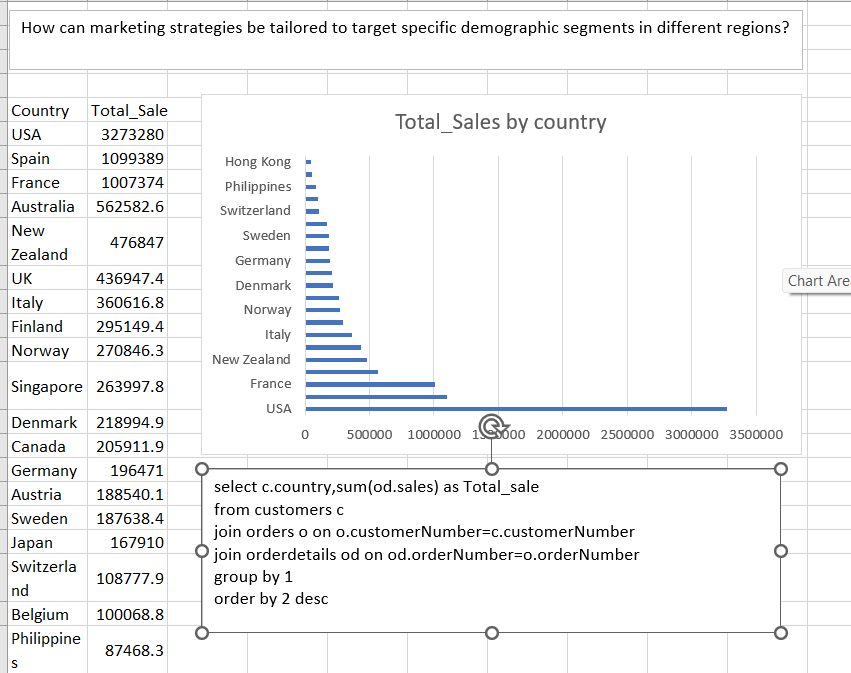
**Customer preferences can indeed vary based on geographic location due to cultural, economic, and environmental differences. Understanding and adapting to these variations is crucial for successful marketing campaigns By incorporating these considerations and strategies, businesses can create more effective and culturally resonant marketing campaigns that connect with customers in different geographic locations.**

What are the characteristics of high-value customers, and how can similar customers be targeted for acquisition:



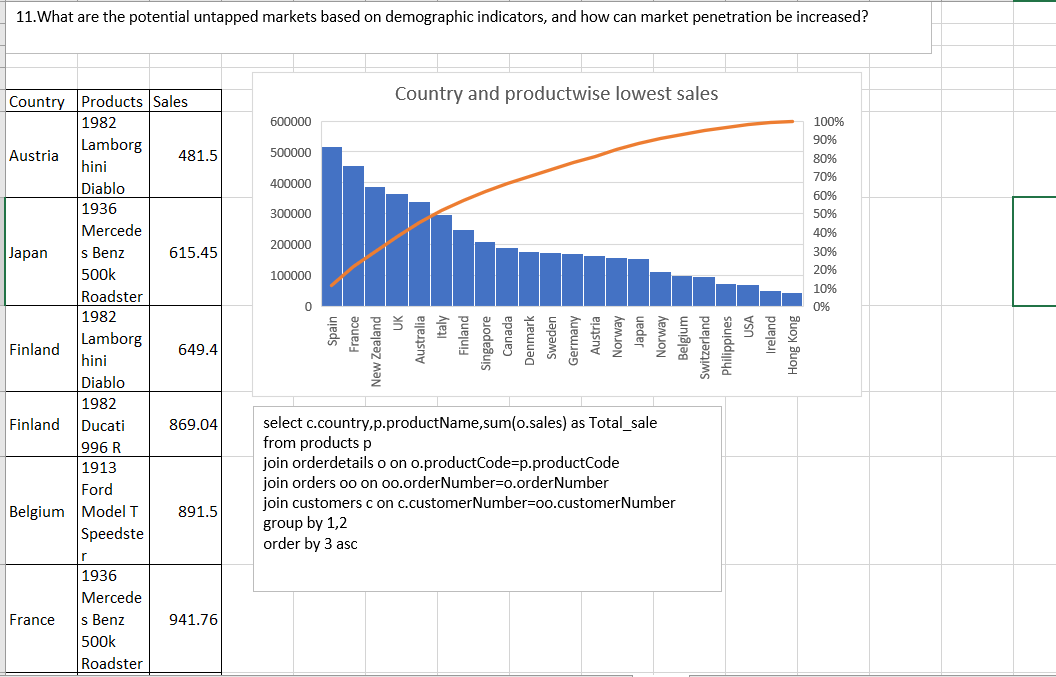
**High-value customers are those who contribute significantly to a business's revenue and profitability. Characteristics of high-value customers can vary depending on the industry and specific business, but there are common traits that are often associated with this segment. Once these characteristics are identified, businesses can develop targeted strategies to acquire similar customers.**

How can marketing strategies be tailored to target specific demographic segments in different regions:



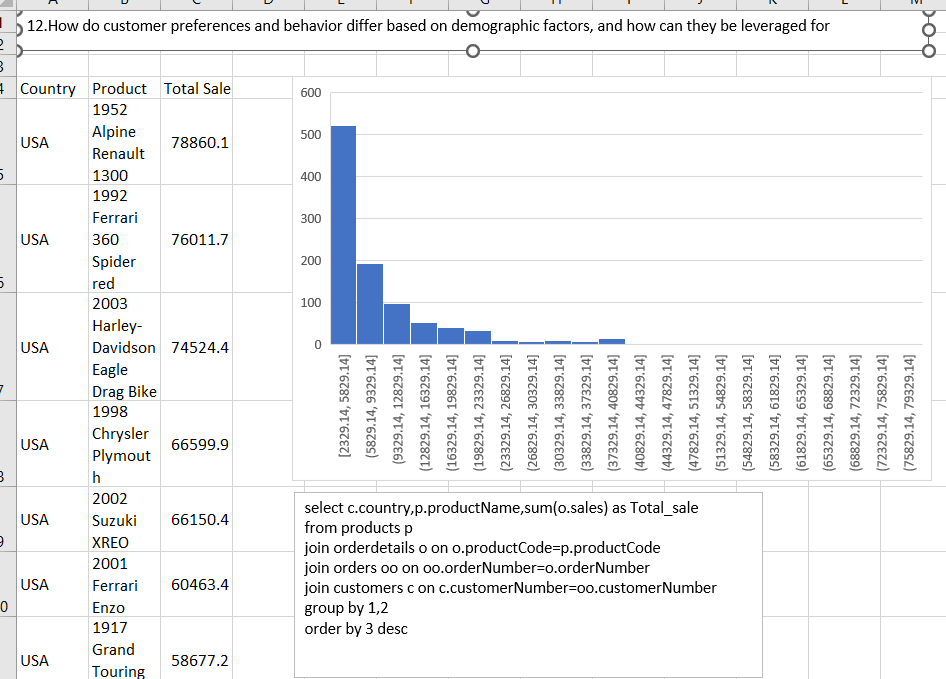
**Tailoring marketing strategies to target specific demographic segments in different regions requires a deep understanding of the local culture, preferences, and behaviors. By incorporating these strategies, businesses can create targeted and culturally relevant marketing campaigns that resonate with specific demographic segments in different regions. Flexibility, adaptability, and a commitment to understanding local nuances are key to success in diverse markets.**

What are the potential untapped markets based on demographic indicators, and how can market penetration be increased:



**Identifying potential untapped markets based on demographic indicators involves recognizing segments of the population that may not have been fully targeted or adequately served by existing products or services. By identifying and tapping into these untapped markets, businesses can expand their customer base and increase market share. It's important to conduct thorough market research, stay attuned to demographic trends, and be responsive to the evolving needs and preferences of different consumer segments.**

How do customer preferences and behavior differ based on demographic factors, and how can they be leveraged for personalized marketing campaigns:



**Customer preferences and behaviour can vary significantly based on demographic factor. Understanding differences is crucial for developing effective personalized marketing campaigns. By understanding and leveraging demographic factors, businesses can create more personalized and resonant marketing campaigns, leading to increased customer engagement and loyalty.**

Conclusion:

The retail analysis process, from data collection on GitHub to data analysis in SQL and the creation of a Power BI dashboard, has yielded valuable insights that are poised to drive significant improvements in our retail operations. This journey of data exploration and interpretation has underscored the critical role that data plays in the success and evolution of our retail business. This project represents a pivotal strategic initiative for retail businesses in today's dynamic and data-centric marketplace. It serves as the backbone for informed decision-making, data-driven insights, and customer-centric strategies that are crucial for the success and sustainability of the retail industry. The project's significance extends far beyond mere data analysis; it fundamentally transforms how retailers operate and interact with their customers.

Through this comprehensive retail analysis, we have unearthed a wealth of insights that hold the potential to shape our business strategies and enhance customer experiences. Some of the key findings include:

a) Sales Trends and Revenue Optimization: We have gained a deep understanding of our sales trends, identifying high-performing products and seasons. This insight allows us to optimize our revenue generation strategies and better forecast sales.

b) Customer Segmentation: The analysis has enabled us to segment our customer base, identifying distinct customer groups with varying preferences and behaviours. This will enable us to tailor marketing and customer engagement strategies for improved personalization.

c) Inventory Management: We have developed more efficient inventory management strategies, reducing carrying costs and mitigating stock-outs. This will result in cost savings and ensure that products are consistently available to meet customer demand.

d) Pricing Strategies: The data analysis has shed light on optimal pricing strategies, helping us set competitive yet profitable prices. We can now implement dynamic pricing mechanisms to remain agile and competitive in the market.

Cost Reduction: We have identified inefficiencies in various operations and marketing areas, which can be addressed to reduce costs and allocate resources more judiciously.

Customer Experience Enhancement: With the insights gained, we are well-equipped to enhance the overall customer experience through personalized recommendations and improved services.

The insights derived from our retail analysis will have a profound impact on our business strategy. They will guide us in making informed decisions, optimizing operations, and enhancing customer interactions. The strategic implications of this analysis include:

Data-Driven Decision-Making: We are committed to a culture of data-driven decision-making, where insights gained from this analysis will be central to our strategic planning and execution.

Enhanced Customer Engagement: Through customer segmentation and personalization, we will enhance our customers' engagement and loyalty, thereby fostering long-term relationships.

Improved Inventory Management: We will implement more efficient inventory management practices to reduce costs and ensure a consistent supply of products.

Competitive Advantage: The ability to respond swiftly to market trends and changing consumer behaviour will give us a competitive edge in a dynamic retail landscape.

Efficient Resource Allocation: By addressing inefficiencies and optimizing resource allocation, we will maximize the value we derive from our investments.

Continuous Monitoring: This is not the end, but the beginning of a data-driven journey. We will continually monitor our operations, update our strategies, and refine our approach to adapt to evolving market conditions.

In the closing, this project has armed us with the tools and insights necessary to navigate the complexities of the retail industry successfully. We look forward to implementing the recommended strategies and capitalizing on the newfound opportunities as we embark on a path of sustainable growth and excellence in the retail sector.