

capstone project-retail analysis



The dataset has significant amount of data that is enough to bring important insights like top performing products, customers choice, employees’ performance, and sales trend over the time. This data is helpful to make data driven decisions and future growth of company.

Overview

THE PROCESS

1.Data Acquisition from GitHub:

Obtain the requisite dataset from a GitHub repository.

2.Data Transformation:

Execute data transformation procedures to ensure data quality and consistency. Also add some new problem statements to enrich the analysis potential.

3.Connecting with Tools:

Establish the connections between the dataset and various analytical tools. connect the dataset with Power BI, Excel, and MySQL Workbench.

4.Problem Statement Solution in Power BI:

Utilize Power BI dashboard to use important insights, Analysis sales trend, product performance, customer preference, and employee performance.

5.Exploratory Data Analysis (EDA):

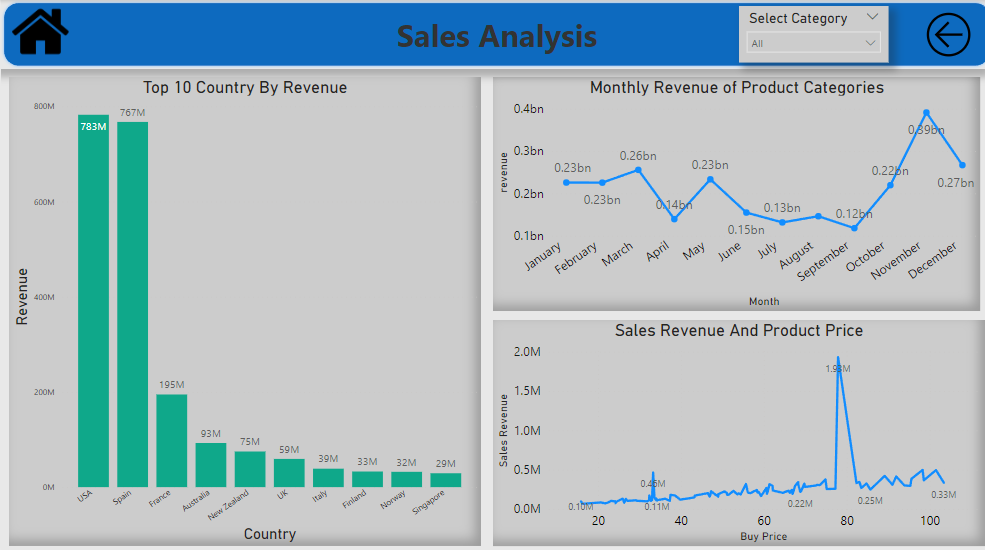
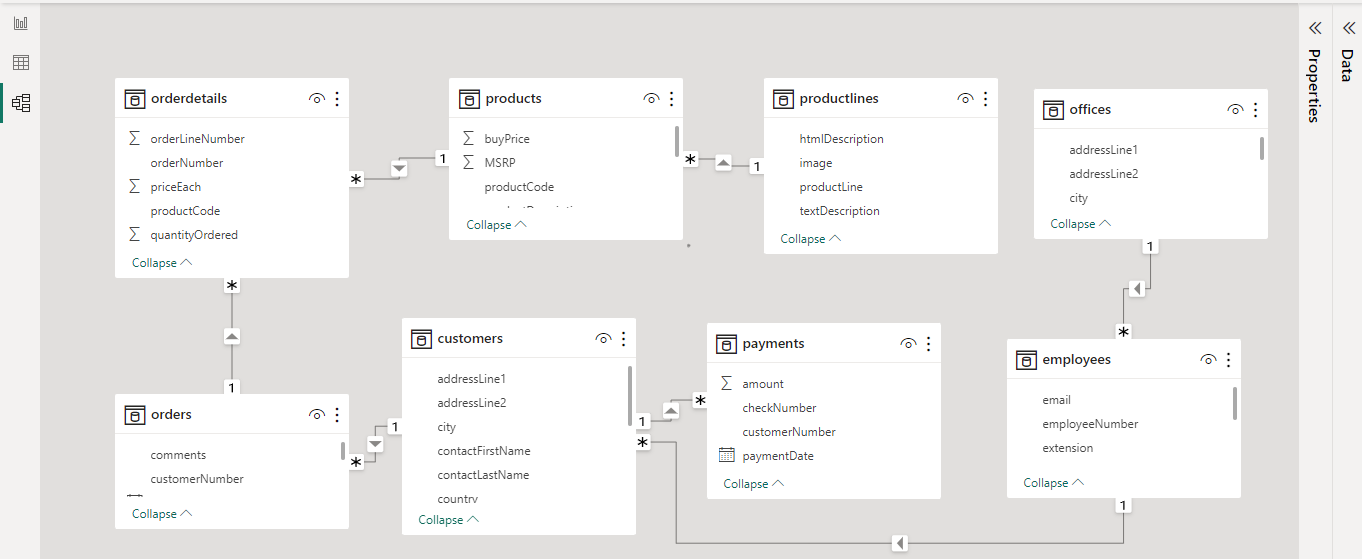
Perform exploratory data analysis using SQL and make visuals in Excel. Extract meaningful patterns and trends from the data to make data driven decision.

6.Create Informative PowerPoint:

Make a comprehensive PowerPoint presentation that contain project objective, methodologies, problem statement solutions and visualizations.

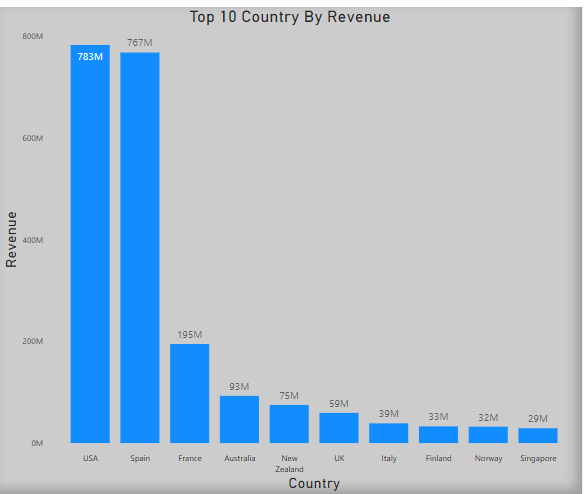
7. Detailed Documentation:

Create a detailed report of project that cover complete analysis process.



ER Diagram

Power BI Problem Statements and Solutions

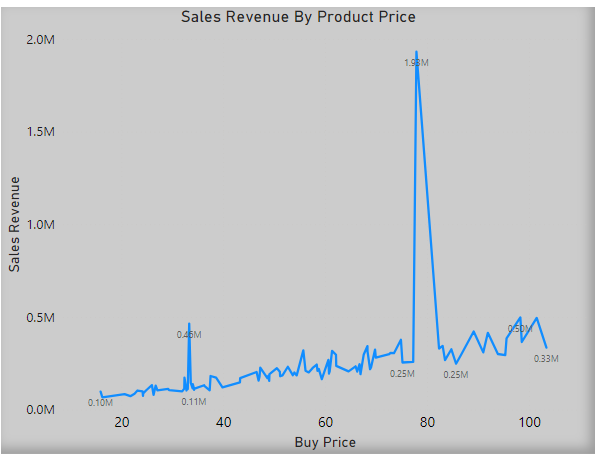


**What are the top regions in terms of sales revenue?**

**In terms of sales revenue USA and Spain generate highest revenue 783M and 767M. In visual we can see top 10 country by sales revenue.**

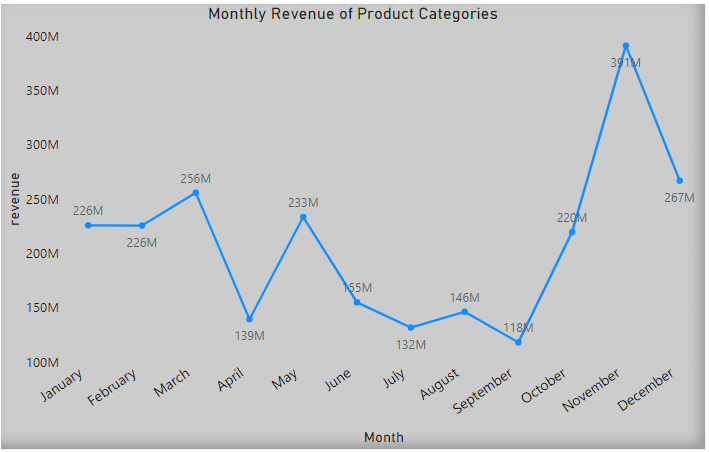
**How does product pricing impact sales volume?**

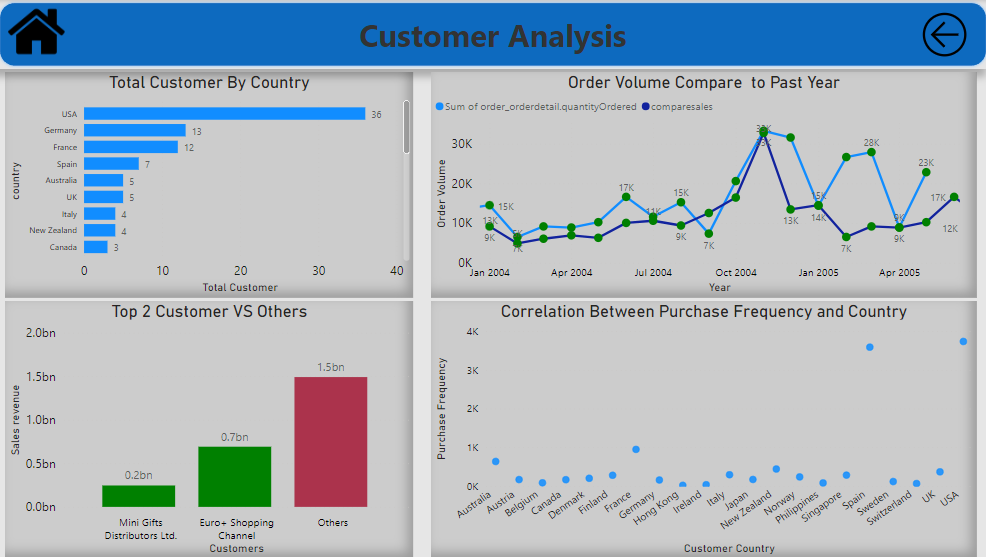
Product price makes a significant impact on sales revenue. In this visual we can see when product price in 10-50 then sales revenue is not much and when price is 85-100 then revenue is greater then10-50 price range, but when product price in midrange like 60-85 then revenue is significantly high, it indicates that customer preference is midrange price.



**How does monthly revenue vary across different product categories?**

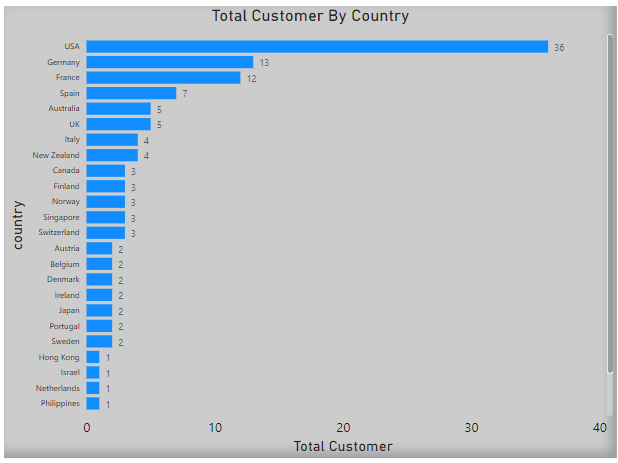
As we can see monthly revenue of product categories is start decreasing in May and in September it hit lowest of the year 118M but in next two month it increases significantly and in November it hit highest of the year 391M.





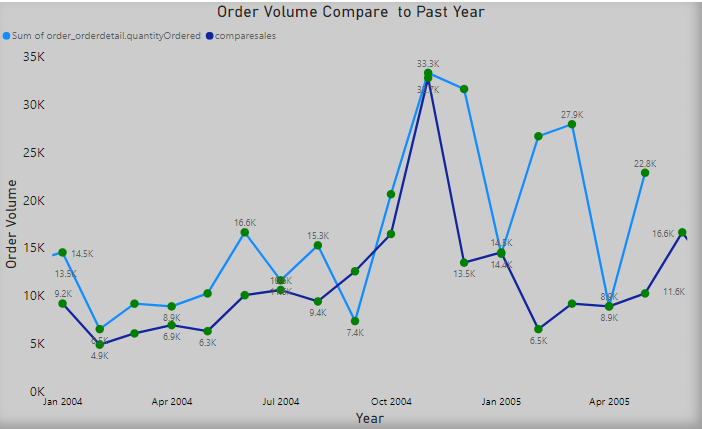
**What is the distribution of customers across different countries?**

As we can see USA has highest number of customer 36 and after Germany, France, Spain, Australia, and UK. We can least number of customers are in Hong Kong, Israel, Netherlands, and Philippines**.**



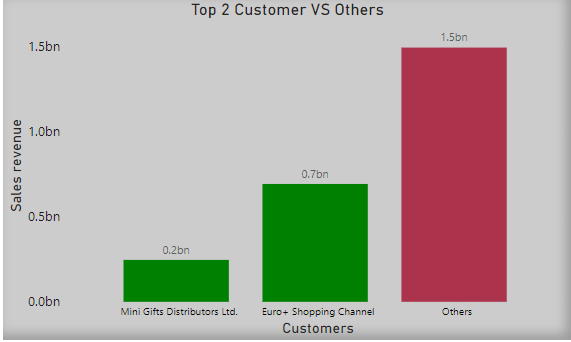
**What is the trend in customer order volume over the past year?**

In this visual we can see order volume in year 2004 compared to 2003 is less but in July to October its cross last year order volume and in November it almost equals to last year highest. Order volume in Jan 2005 and April 2005 is equal to year 2004 but less in February and March.



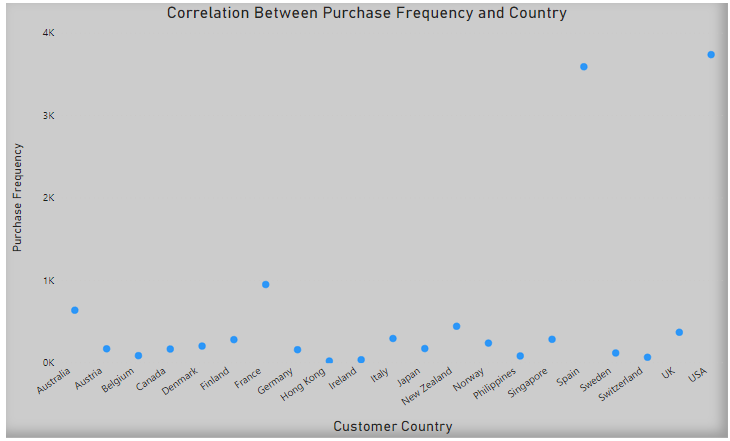
**How does the sales performance of top customers compare to the rest?**

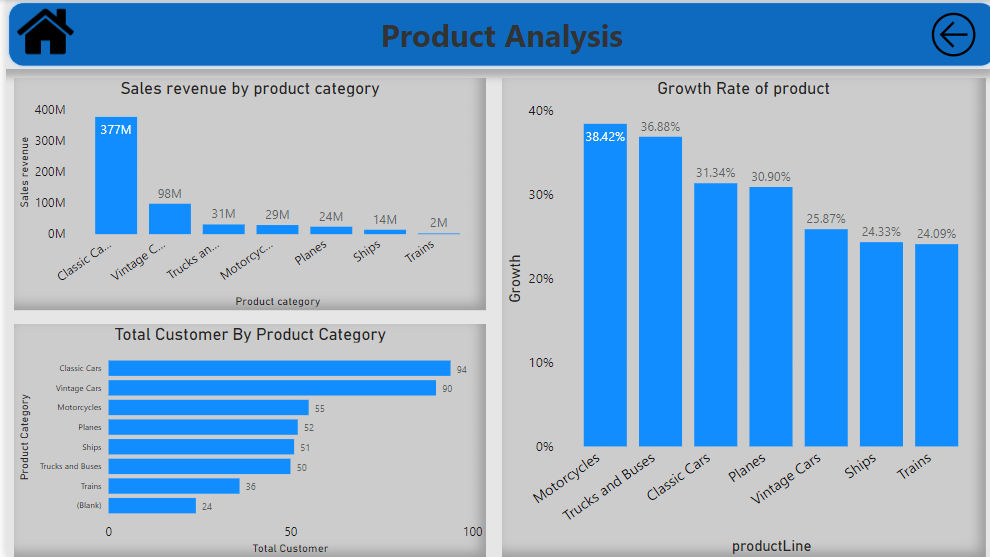
We can see top two customer Mini gifts distribution ltd and Euro+ Shopping channel contributing 0.9bn that is 37.5% of our total sales revenue.



**What is the correlation between customer demographics and purchase frequency?**

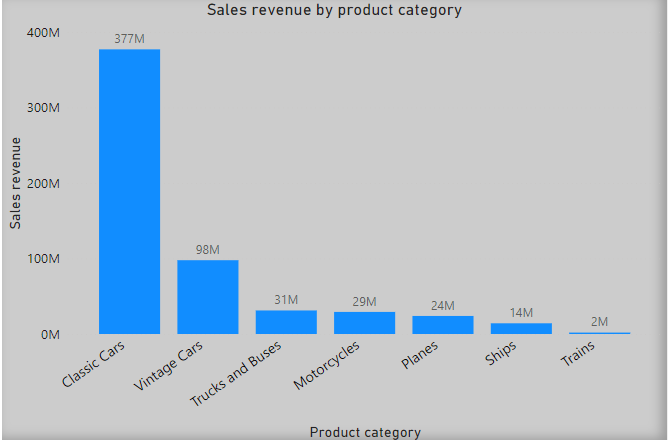
As we can see correlation between customer country and their purchase frequency, USA and Spain has highest purchase frequency and Hong Kong has lowest.





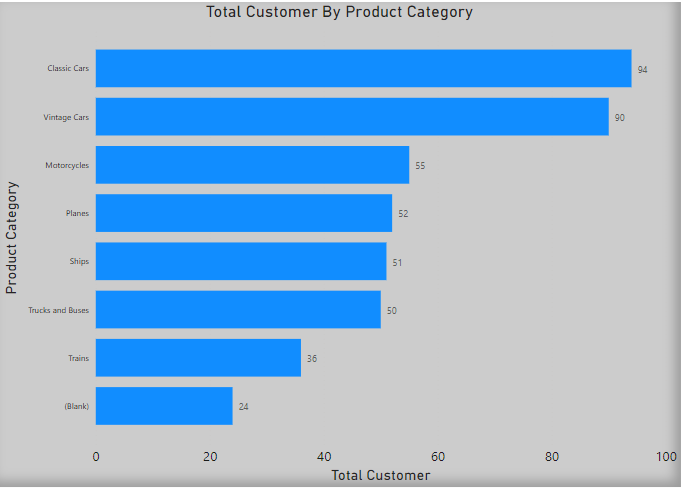
**What is the distribution of product sales across different product line?**

Here we can clearly see that classic cars generate highest revenue 377M after that vintage car (98M). Revenue from Trains is lowest (2M).



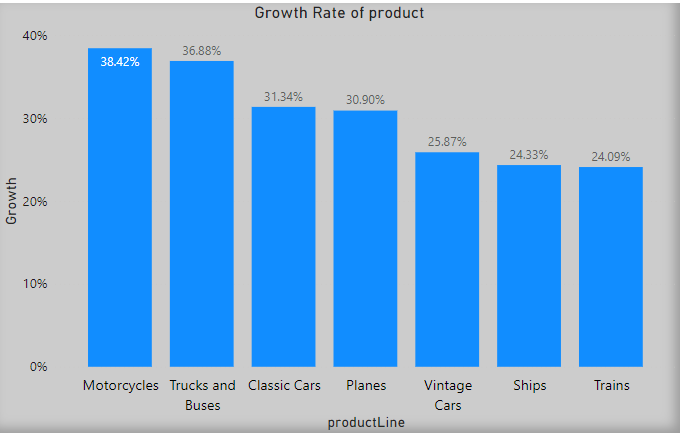
**What is the distribution of customers across different Product categories?**

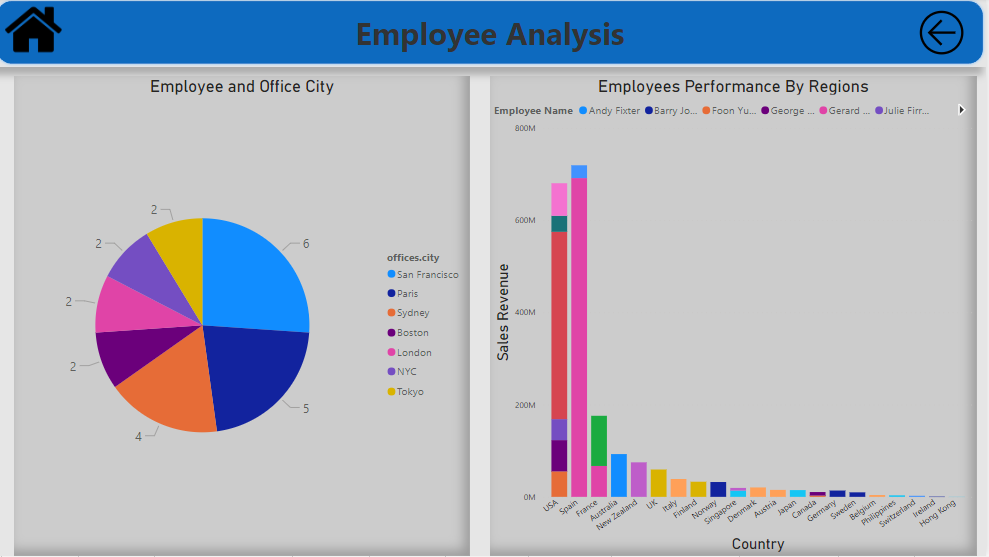
Classic cars and Vintage cars are on top with 94 and 90 customers. After that Motorcycles, Planes, Ships, Trucks and Buses, and last is Train with 36 customers.

****

**What is the growth rate of product category?**

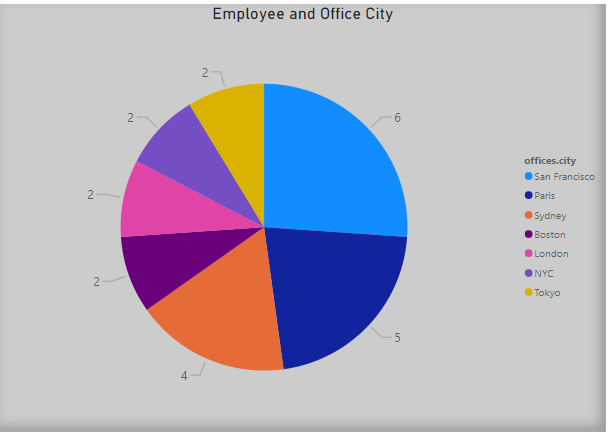
In this visual we can see Motorcycle has the highest growth rate 38.42%. Trucks and buses have Second highest growth rate 36.88%. Trains has the lowest 24.09%.





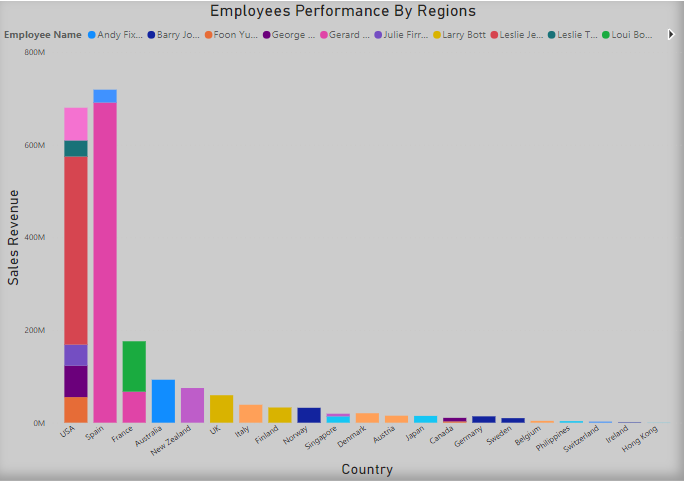
**What is the distribution of employees in various office cites?**

As we can see San Francisco has the highest number of employees. Second and third is Paris and Sydney.

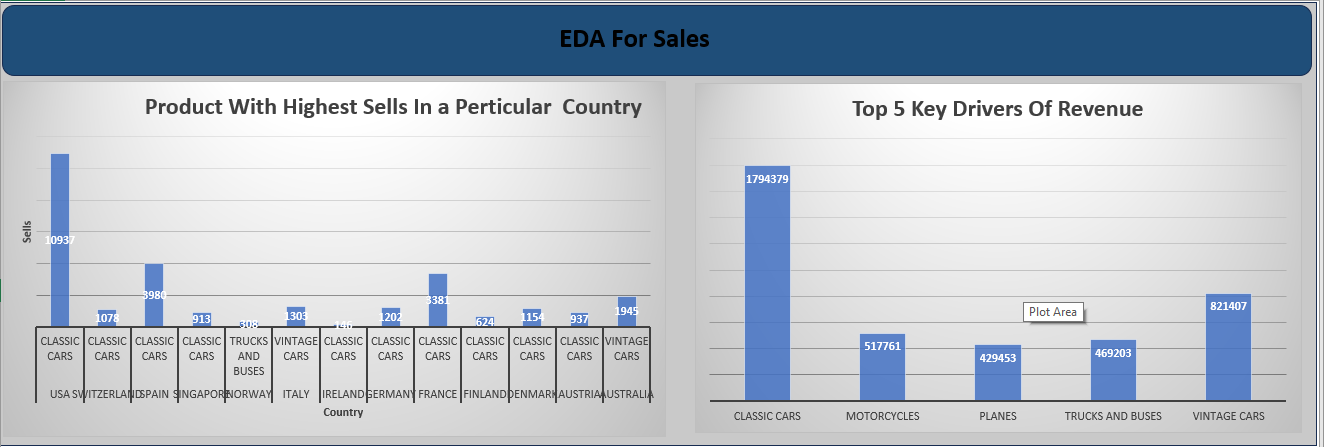


**How does the performance of sales employees vary across different region?**

Performance of sales employees is varied across different region, here we can see Employee Gerard Hernandez sales revenue is above 600M in Spain but in France it is 66M**.**

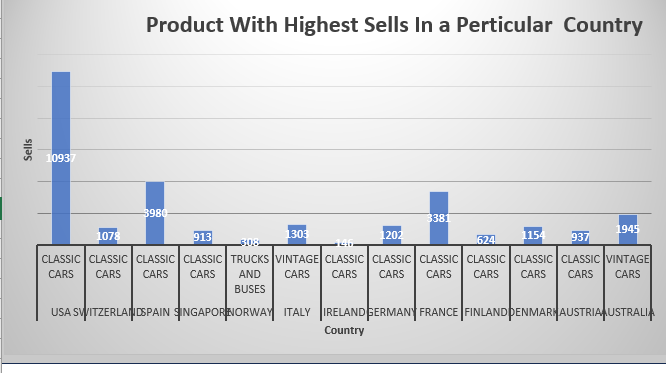


**EDA Problem statements and solutions**



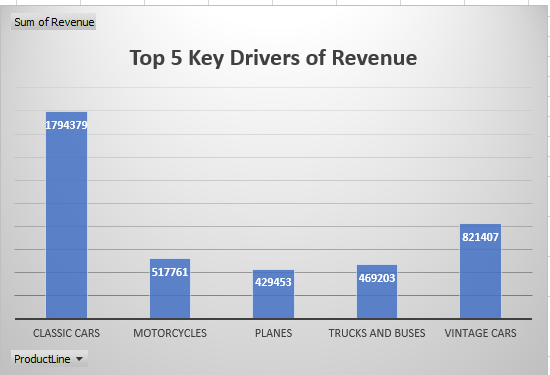
**Which Factors contribute to the highest sales in a particular region?**

**According to the EDA we can see most of highest sales is from classic cars and after it vintage cars but in Norway trucks and Buses contribute highest sales.**



**What are the key drivers of sales growth, and how they be leveraged for future success?**

**Top 5 key driver of sales growth are classic cars, vintage cars, motorcycles, Trucks and buses, and Planes. For future success we can invest more in these product categories, introduce new features, and take marketing decisions.**



A screenshot of a computer

Description automatically generated

**How can customer purchasing pattern be influenced to increase average order value?**

**Here we can see total order quantity of classic cars is high and the average price for it also high. Trend in total order and average price is proportional but Trucks and Buses does not follow the trend.**

A graph on a computer screen

Description automatically generated

**What are the main factors that influence customer loyalty and repeat purchase?**

**Customer loyalty and repeat purchase is dependent on product quality and customer satisfaction. Here we can see 597 products are disputed, 1059 products are cancelled, 1777 are on hold, 1660 resolved and 1379 are in process, these things can influence customer loyalty and repeat purchase.**

**A pie chart with numbers and numbers

Description automatically generated**

**What are the characteristics of high value customer and how can similar customers be targeted for acquisition?**

**Characteristics of high value customers are they have Credit limit more than average credit limit and total orders are more than average orders, so according to these criteria we have 14 high value customers. We can target these customers by giving some discount and special offers.**

A graph of sales and sales

Description automatically generated with medium confidence

A screenshot of a graph

Description automatically generated

**How can Marketing strategies be tailored to target specific demographics segments in different region?**

**We can make our marketing strategies on the basics of top performing category in a specific region. Like in USA classic cars is top performer and in Australia vintage cars is top performer!**

A map of the world

Description automatically generated

**How do customer preference and behaviour differ based on geographics factors and how can they have leveraged for personalized marketing campaigns?**

**Customer preference is differ based on geographic location as we can see in Australia classic cars and vintage cars are preferred by customers but in Italy, we can see planes are ordered more as compared to classic cars and vintage cars. So, we can customise our marketing strategies according to customer preference.**

A graph of blue rectangular bars

Description automatically generated with medium confidence

A graph of a customer

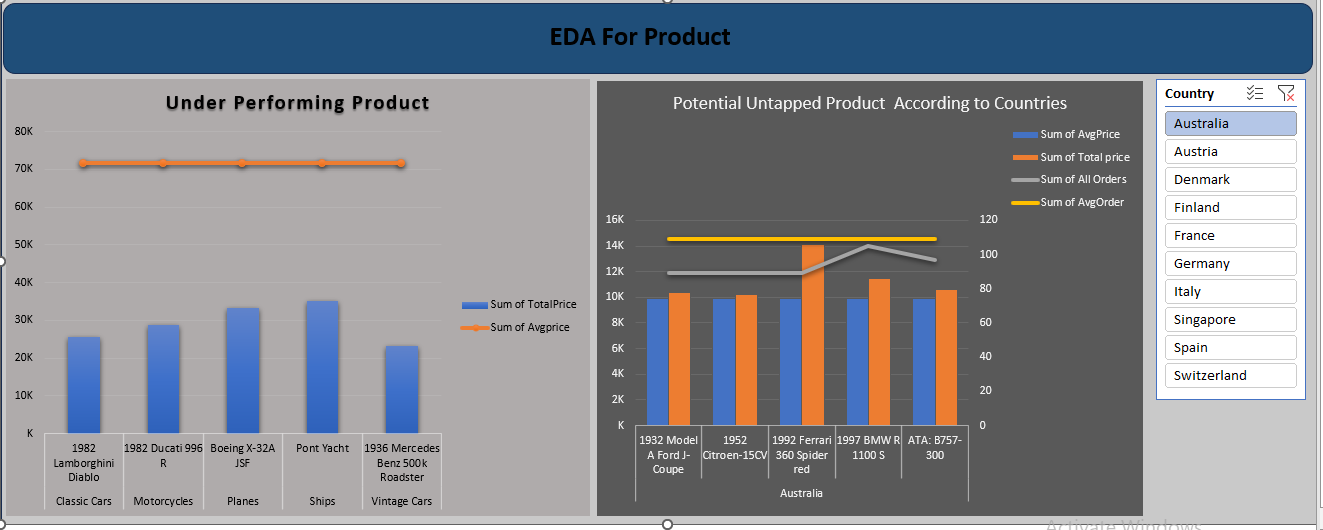
Description automatically generated

**How does customers preference differ based on geographic location, and how can marketing campaigns be customized accordingly?**

**Customer preference is differ based on geographic location as we can see in Australia 1913 Ford Model T Speedster preferred by customers but in Italy, we can see American airlines: MD-115 is ordered more. So, we can customise our marketing strategies according to customer preference.**

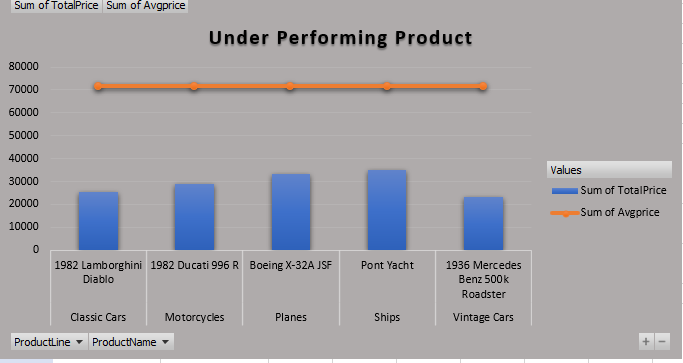
A map of the world with different colored countries/regions

Description automatically generated



**Are there any specific market segments where a particular Product is underperforming and how can it be improved?**

**Yes, as we can see here 5 products in different market segment are under Performing. We can improve them by introducing some new and popular features, giving discount and marketing that can catch pulse of people.**



**What are the potential untapped markets based on demographics indicators, and how can market penetration be increased?**

**Many regions have potential untapped markets. In this visual we can see Australia has 5 products that has enough potential. We can use these products to increase market penetration. We can make some new marketing and pricing strategies for these products.**

