# **2nd Project portfolio**

## **Data set name:** Sample Sales Data

The "Sample Sales Data" dataset on Kaggle provides a comprehensive collection of order details, sales figures, customer information, and shipping data. This makes it an excellent resource for retail analytics tasks such as segmentation, customer analytics, and clustering. The dataset is particularly valuable for gaining insights into customer behavior and enhancing business operations. Its global scope encompasses various aspects of nations worldwide, making it an invaluable tool for developing strategies to boost sales and enhance customer satisfaction in the e-commerce sector.

## **Data Cleaning and uses of Software:**

To start the analysis, I import the data into Python and begin the data cleaning process. This involves removing duplicates, handling null values, standardizing formats, and ensuring accuracy using Python libraries such as Pandas and NumPy. After cleaning the data, I perform data analysis to discover insights and address issues using Python queries. I utilize Python libraries like Matplotlib and Seaborn to create visualizations, and also develop a comprehensive Power BI dashboard to enhance the analysis with compelling visuals.

## Data Analysis:

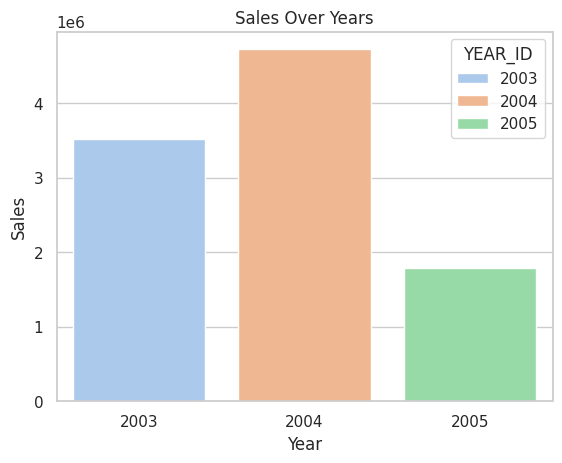
In order to visualize the data, I began by focusing on cleaning and preprocessing it. This involved checking for and removing any null values, identifying and removing duplicate entries, and ensuring that the data types were correctly set for each column. I also eliminated unnecessary columns that did not contribute to the analysis. Specifically, I addressed missing or incomplete values in the Territory column by imputing or updating them as needed. All essential steps were taken to ensure that the data is properly formatted and cleaned, thereby making it ready for effective visualization and analysis.

# **Question and Insights**

**Question No 1identifity the number of sale over past servel year**s

sale\_year= data.groupby('YEAR\_ID')['SALES'].sum().reset\_index()

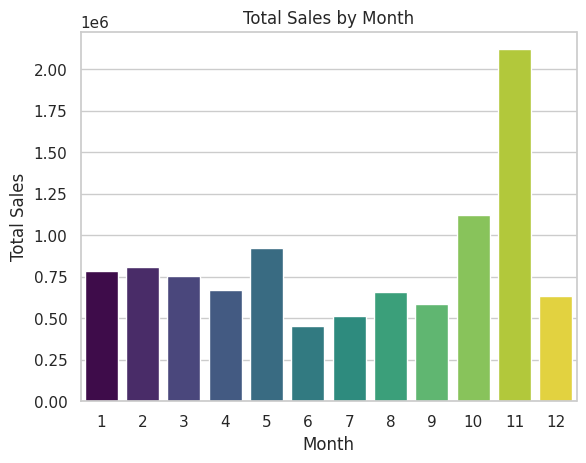
sale\_year



**Insight:**

Sales increased significantly from 2003 to 2004, suggesting a strong year-over-year growth trend. The highest sales figure in 2004 compared to 2003 and 2015 indicates a peak period, possibly due to successful product launches, market expansion, or seasonal factors. The EMEA region consistently had the highest sales, reflecting strong market performance, while NA and APAC regions followed

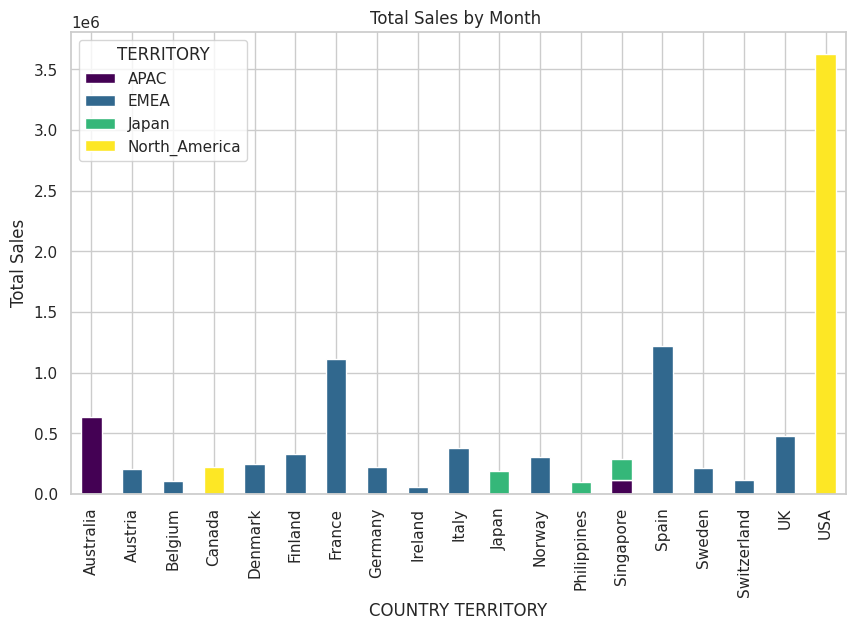
**Question No 2 Which Month Has the Most Sales?**



**Insights**

The significant sales increase in the 11th month suggests a strong seasonal trend, likely year-end promotions, while other months experience lower sales, indicating that November might be a key period for targeted marketing and sales strategies.Other months have lower sales figures, indicating that November is a critical period for maximizing revenue and focusing marketing efforts

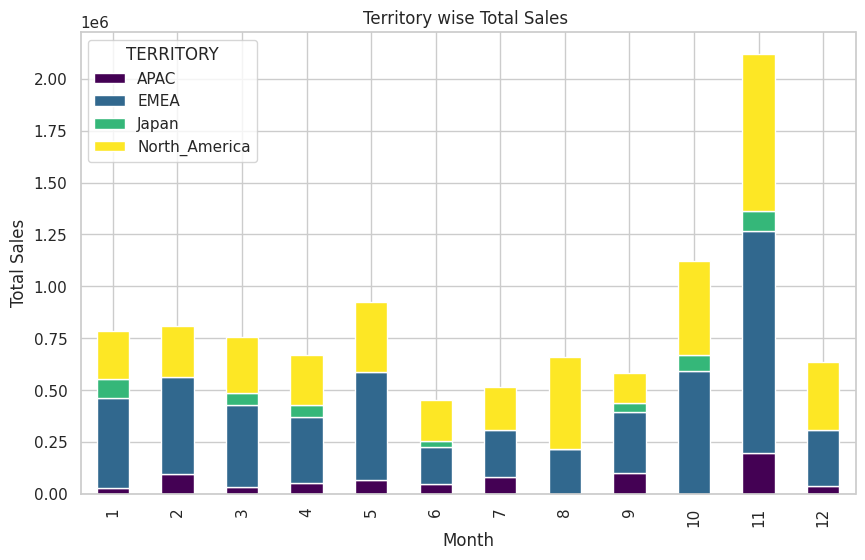
**question:3 Which country have most sale**



**Insights**

The United States dominates sales, contributing over 52% of total global sales, indicating it is the primary market driver while the France and Spain is following by USA.

**Question 4: How does the total sales distribution vary across different territories on a month-by-month basis?**

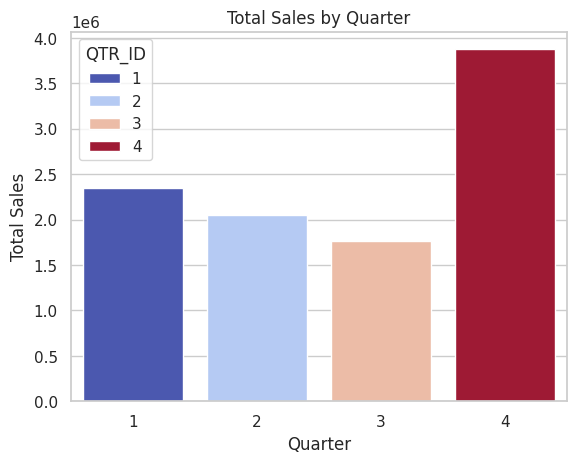
****

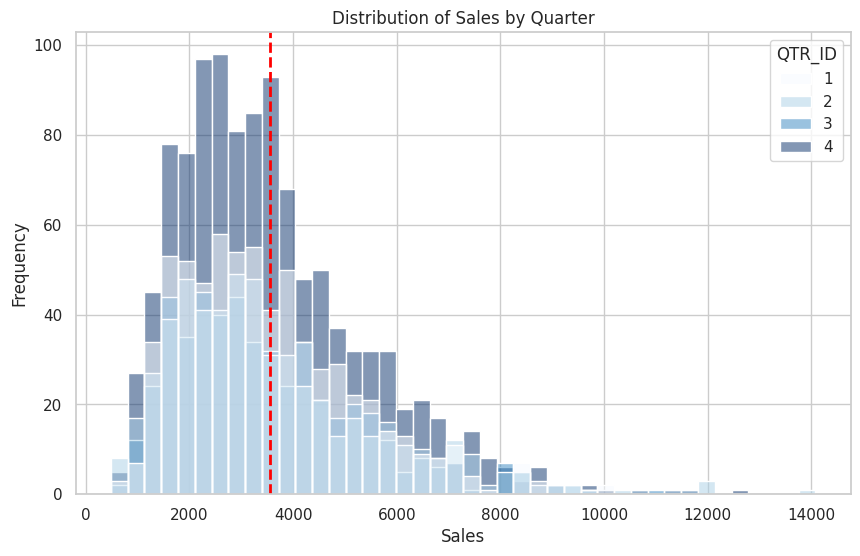
**Insights:**

Sales peak significantly in November across all territories, with EMEA and North America showing the highest figures. EMEA maintains strong sales throughout the year, while North America has notable peaks in May and November. APAC and Japan exhibit more variable sales patterns.

The month-by-month sales distribution reveals that November consistently sees the highest sales across all territories, likely due to holiday shopping. North America leads in sales, reflecting its strong consumer market, while Japan shows the lowest sales, possibly due to different cultural shopping habits or market saturation.

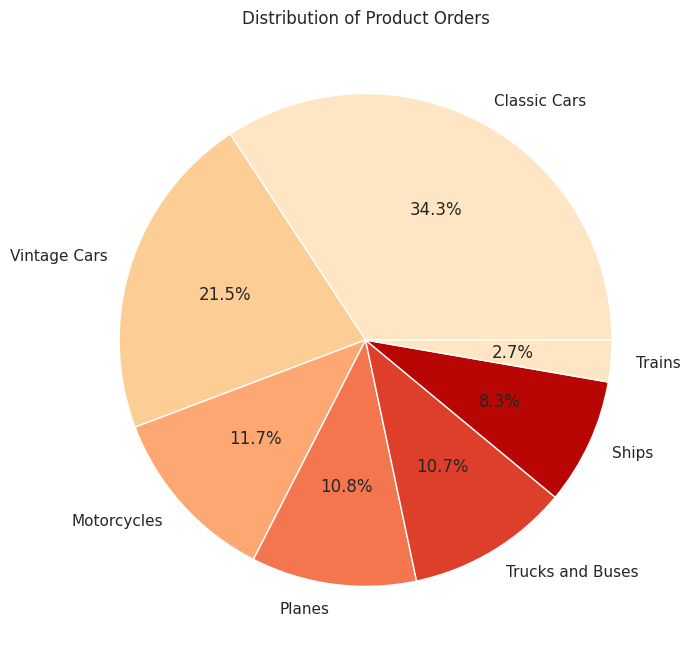
**question:5** **What are the total sales by quarter?**

****

****

**Insights:** Total sales by quarter reveal that Q4 consistently has the highest sales, likely driven by increased order quantities due to end-of-year corporate purchasing, high demanding of cars, and strategic business spending, making it the most profitable quarter. Q1 shows a strong start to the year, while Q2 and Q3 experience lower sales.

**\*Question:6 define the most order product type \***

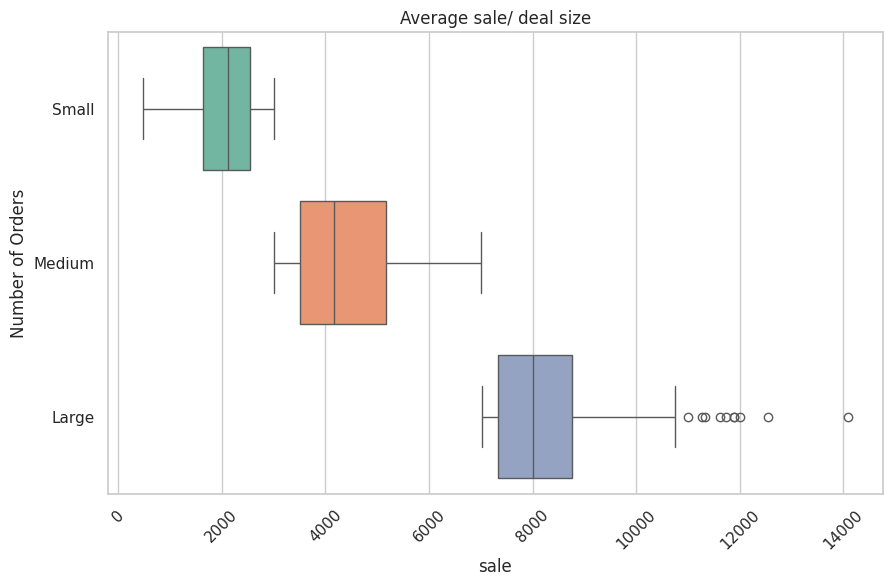


**Insights:**

The above graph shows that the products with the lowest price were sold the most, whereas the products with a high price were the least popular.

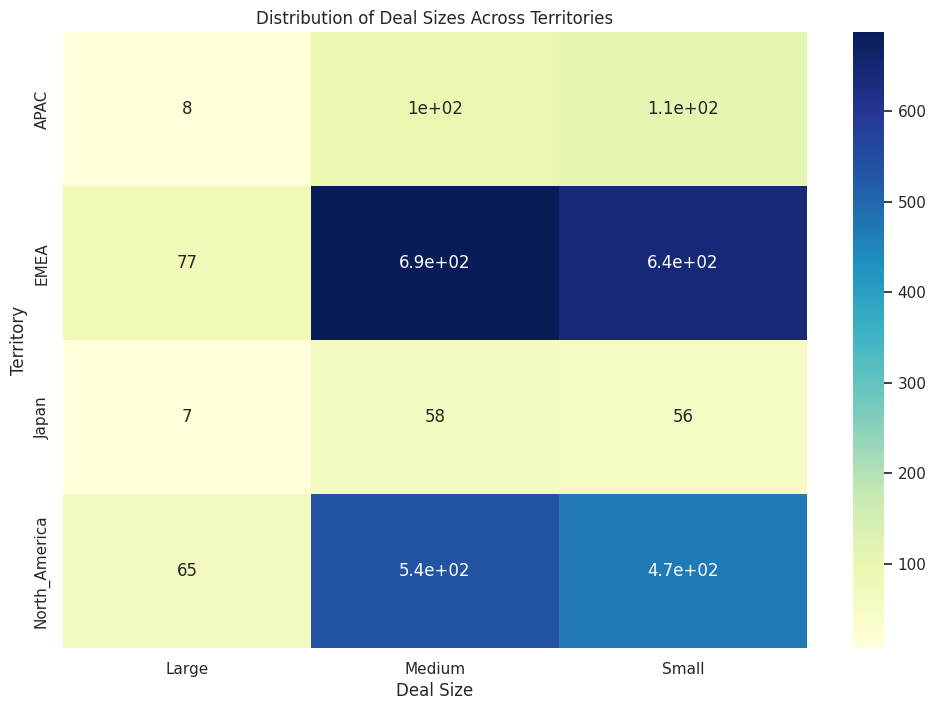
From the plot, 'Classic Cars" & "Vintage Cars" are the most demanded products, 'Trains" was the least demanded product

**\*Question 7 :Calculate the average sale per deal size \***

****

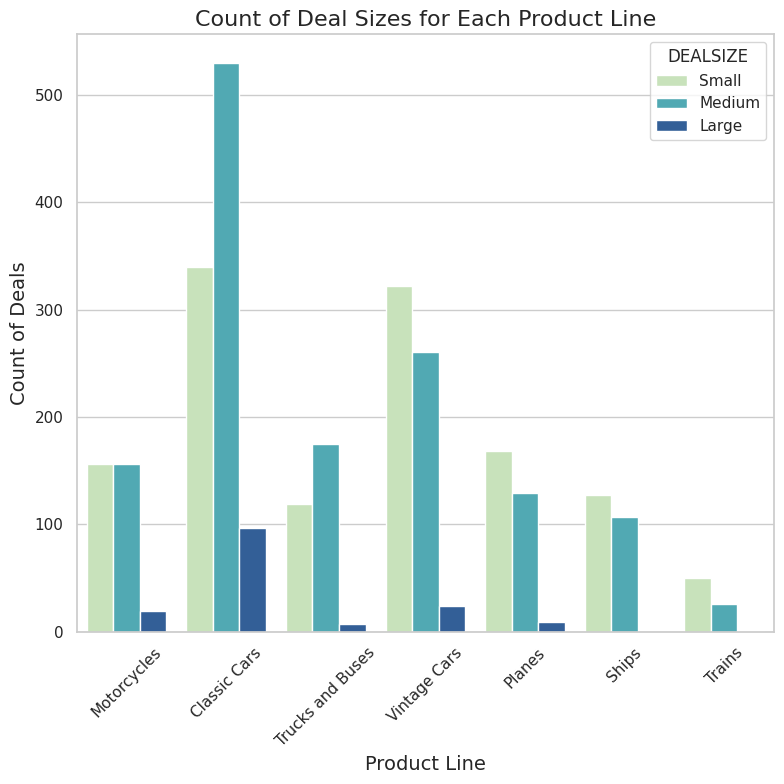
**Insights:** The average sale per deal size shows a clear trend: Large deals generate significantly higher sales compared to Medium and Small deals. This indicates that focusing on larger transactions can substantially boost overall revenue. The pattern suggests that regions with a higher frequency of large deals, such as EMEA and USA, are more profitable, while medium and small deal sizes contribute less to total sales.

**Question \*\*8: find the Distribution of Deal Sizes Across Different Territories**?

****

**Insights:** EMEA leads with the highest counts of large deals, indicating a strong market for high-value transactions, and also shows significant numbers of medium and small deals. North America has a robust market with many medium and small deals, though fewer large ones compared to EMEA. APAC and Japan show lower counts for large deals, with Japan having the smallest overall deal sizes, suggesting potential areas for market growth or strategic adjustments.

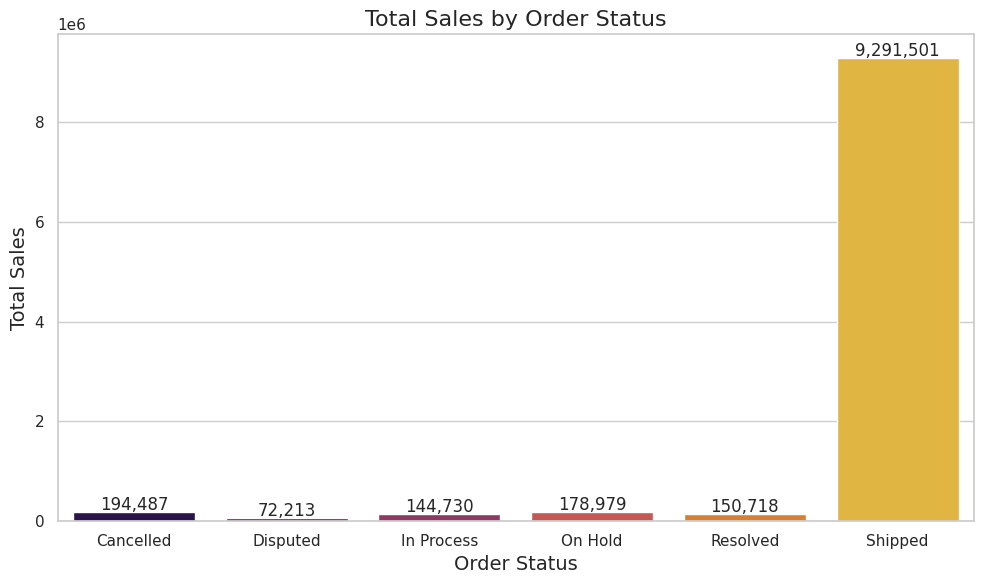
Q**uestion 9: What are the counts of deals for each product line and deal size, and how do these counts vary across different product lines?** **bold text**

****

**Insights:**

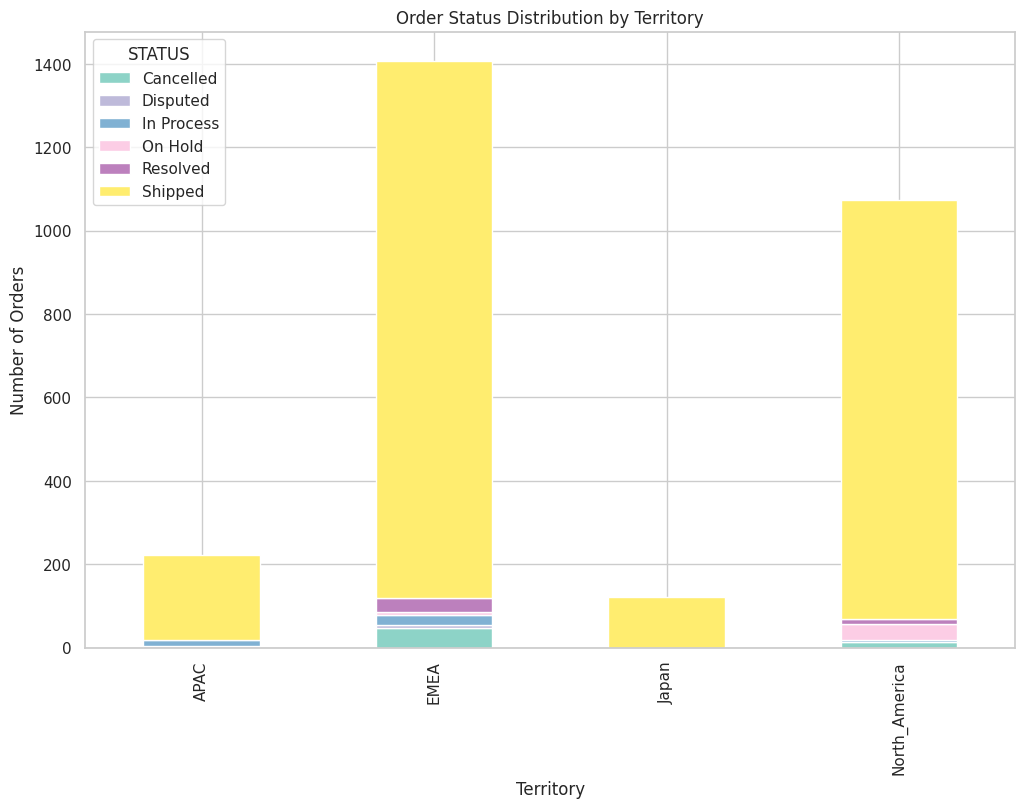
The data reveals that Classic Cars and Vintage Cars dominate in medium-sized deals, with Classic Cars also having a high count of small deals. Classic car has high rate in USA a market. Motorcycles and Planes show fewer large deals, indicating these products are less frequently involved in high-value transactions. Trains and Trucks and Buses have a lower overall deal count, with Trains particularly underrepresented in large deals, suggesting potential areas for growth or reevaluation in high-value sales strategies.

**Question 10: How do total sales differ across various order statuses**

****

**Insights:** The total sales differ significantly across order statuses, with Shipped orders generating the highest sales, indicating successful fulfillment and completion of transactions. In contrast, Cancelled and Disputed statuses show much lower sales, reflecting lost revenue or unresolved issues. In Process, On Hold, and Resolved statuses fall in between, suggesting that while they contribute to sales, they are less impactful compared to completed shipments. This highlights the importance of focusing on order fulfillment to maximize sales revenue.

**Question 11: How are order statuses distributed across different territories, and what does the distribution reveal about the status of orders in each territory?**

****

**Insights:** EMEA and North America lead in shipped orders, indicating strong fulfillment in these regions. EMEA also shows higher counts of cancelled and disputed orders, suggesting potential issues with order processing. North America has a notable number of orders on hold, which could reflect logistical challenges. Japan has a high count of shipped orders with minimal issues, indicating efficient processing

# Recommendations

1. Sales trend analysis: Inspect several years' sales data to evaluate how sales have performed over a specific period. Compare trends from different years to identify long-term sales changes caused by economic conditions, the market environment, and sales promotion strategies.
2. Monthly Sales Analysis: Monthly sales report is meant to assist the business to know the monthly sales of the business. And will assist in analyzing more and to determine more the factors which contributed to the increase or decrease of sales in some months. They can also be used to organize future activities like having promotion or marketing activities that are relevant to the month.
3. Focusing on customers in high-sales locations: Concentrating on the customers in the high-sale areas is also likely to boost sales in a better way. Each of such market area will require a set of marketing strategies and promotional activities.
4. The best-selling products are the cars. The EMEA Territory has the biggest sales, with the USA contributing the highest sales per country.
5. Improving Products or Services: Studying the products or services usually required by the clients can enable the organization identify changes in customer needs. For enhancing or creating the products or services more effectively to satisfy the needs of the customers.
6. Quarter 4, especially November and October, showed higher sales, likely due to anticipation of the December festive period. Increasing my marketing efforts in quarter 4 will likely result in higher sales.
7. Best-selling product rankings: Rankings of the Best-Selling product may be used to identify and allocate resources to those products with the greatest promise on the market. It also aids in enhancing the visibility of such goods and to get those goods to the market faster.
8. Implement improvements in order processing to reduce high cancellation and dispute rates in EMEA. Streamline procedures and resolve logistical challenges in North America to enhance overall fulfillment efficiency and customer satisfaction.
9. Well the bestselling products are the cars and the most lucrative territory is the EMEA territory considering the number of countries to target, it will be best marketing strength is targeted towards the US which us far more lucrative. So targeting The USA market with our cars will be a good marketing strategy.