Starting out with the initial stage of [HNG 11 internship program](https://hng.tech/internship), I analysed a retail vehicle sales dataset obtained from Kaggle. This analysis aims to provide a preliminary review of the dataset to identify key patterns, trends, and provide insights for potential further analysis.

**Data Overview**  
The dataset comprises 2,823 entries and includes a variety of data fields such *as order number, quantity ordered, price, order line number, sales, order date, status etc.* Based on the provided data set from Kaggle, the dataset contains information about a company involved in the retail sale of various vehicles, including classic cars, motorcycles, planes, ships, trains, trucks and buses, and vintage cars which spans between the year 2003 to 2005.

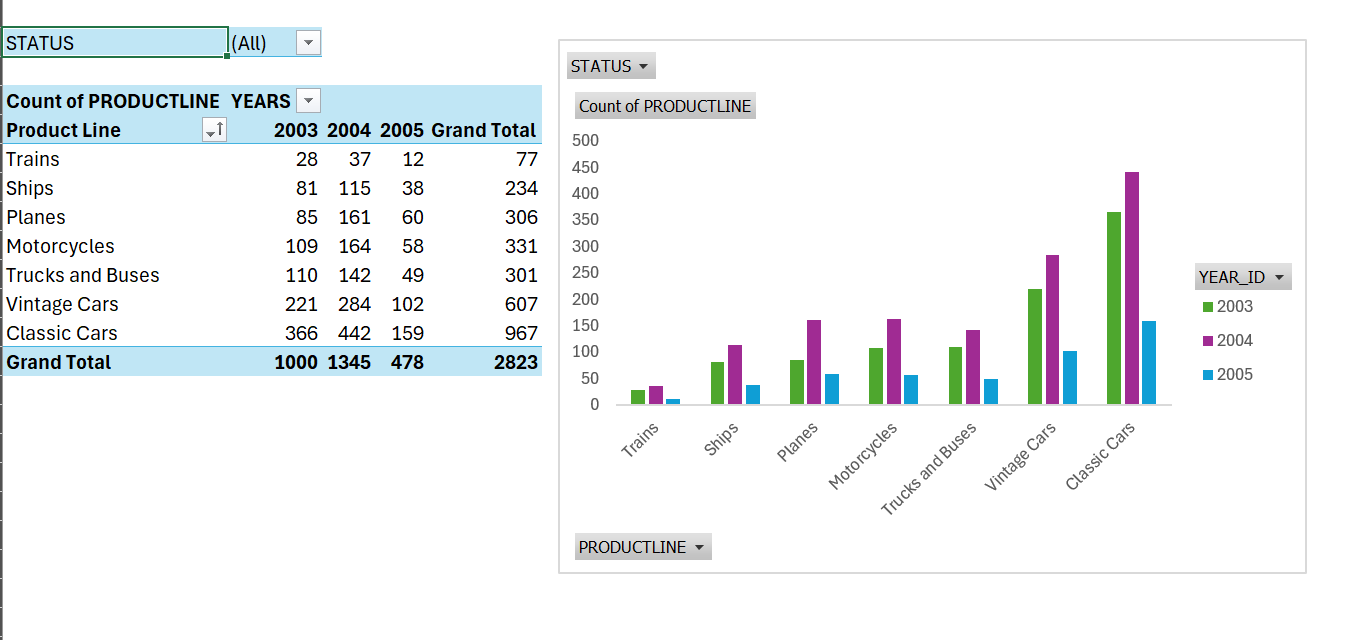
**Data Integrity and Anomalies:** Review shows that there are no duplicate data set.

**Product Analysis**

The year 2004 has the highest number of vehicles sold, with a total of 1,345 units.

1. Among the vehicle categories, classic cars are the most shipped, with a total of 425 units in 2004.
2. There is a noticeable increase in the number of vehicles sold from 2003 to 2004, indicating a significant rise in demand or sales volume during that period. However,

the year 2005 shows strike decline in volume of sales.



**Order Status Analysis** below highlights the differences between each status, suggesting most orders were processed and “shipped” successfully, with minimal issues in other statuses.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated**Geographical Sales Analysis**

A screenshot of a computer

Description automatically generated  
The analysis examines sales performance across territories over a three-year period, focusing on deal sizes. EMEA territory stands out with the highest sales levels, especially in the "medium" deal category. This trend is echoed in other territories, though with lower sales volumes. Additionally, the analysis reveals that the North America territory achieves the highest sales per country (USA), with "San Rafael" in California showing the strongest sales performance at the city level. The chart clearly illustrates that the USA has the highest vehicle prices among the countries listed. Furthermore, the data indicates a consistent upward trend in pricing across all regions, likely influenced by varying levels of demand and purchasing power among customers.

**Summary of Insight Identified**

* Over the years, classic cars have been the most processed and shipped vehicles. This might be due to the product's distinctive quality and what consumers value most, even though it still costs more to buy than some other products.
* In comparison to other states, California maintains the highest level of total deal size and the highest level of vehicle sales across all regions.

**Conclusion**After identifying the sales trends and patterns of these products across specific regions, we should aim to understand the purchasing power of the customers. This involves determining what stands out for each country and analysing how price and order quantity influence demand over the years.