STAT 112: Introduction to Data Processing and Visualization

TABLEAU VISUALIZATION PROJECT:

WHO ARE OUR CUSTOMERS AND WHAT DO THEY PREFER?

ALI ALTUNTAŞ

2646529

Tableau Public URL: [LINK](https://public.tableau.com/app/profile/ali.altunta./viz/CustomerAnalysis_17336746504880/Dashboard)

Github URL: [LINK](https://github.com/ForxDeven/customer_analysis)

INTRODUCTION:

This report presents an analysis of the auto sales dataset, discovering different dimensions such as customer behavior, sales trends, and product performance. The objective of this analysis is to come up with insights that can help with strategic decision-making on increasing customer loyalty, acting according to market trends, and revenue optimization using computer applications Tableau and Microsoft Excel. The dataset includes detailed information about orders, customers, and product categories across multiple countries and regions. The insights which stakeholders can get from this report can help them with having a good understanding of the customer preferences and different products across multiple product lines.

DATA DESCRIPTION:

In this analysis, 2 datasets have been used. The first one being the company’s “autosales” dataset and the second one being a dataset that includes information about all the countries in the world called “world-data2023”. Although the world-data2023 dataset gives valuable insights about different countries, it has not been used in this analysis except for location data it includes.

DATA CLEANING AND TIDYING:

To ensure that the data is reliable, several data cleaning methods have been used on this dataset:

1. MISSING VALUES:
   1. 48 whitespaces have been found in the ADDRESS column of the autosale dataset and they have been trimmed.
   2. Some missing values in the world\_data\_2023 table have been detected but since those values are outside of the values of AutoSales table and therefore they are irrelevant to this analysis, they have been neglected.
2. STANDARDIZING DATA FORMATS:
   1. AutoSales PRICEEACH and MSRP column’s data type has been changed as currency.
   2. AutoSales PHONE column’s data type has been changed as string.
   3. world\_data\_2023 GDP and Gasoline Price columns’ data type has been changed as currency.
   4. AutoSales Order\_Number column’s data type has been changed as string on Tableau.
   5. All column names in both tables have been changed in a single format such as “This\_Is\_A\_Title”.
3. REMOVING DUPLICATES:

There are no duplicates found in the dataset.

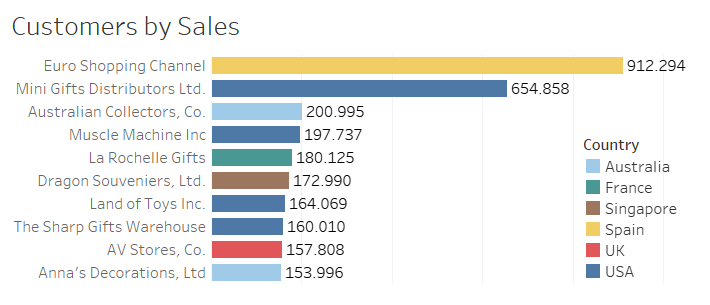
1. JOINING:

The two datasets autosales and world\_data2023 have been joined on country names.

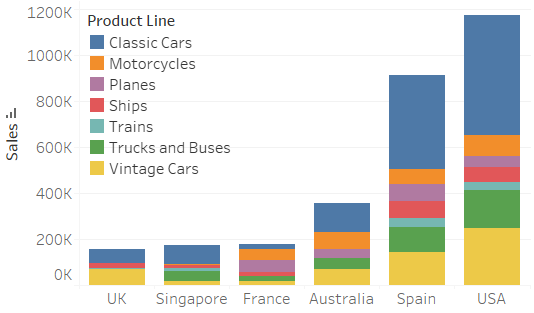
EXPLORATORY DATA ANALYSIS:

In this analysis, 6 research questions has been conducted:

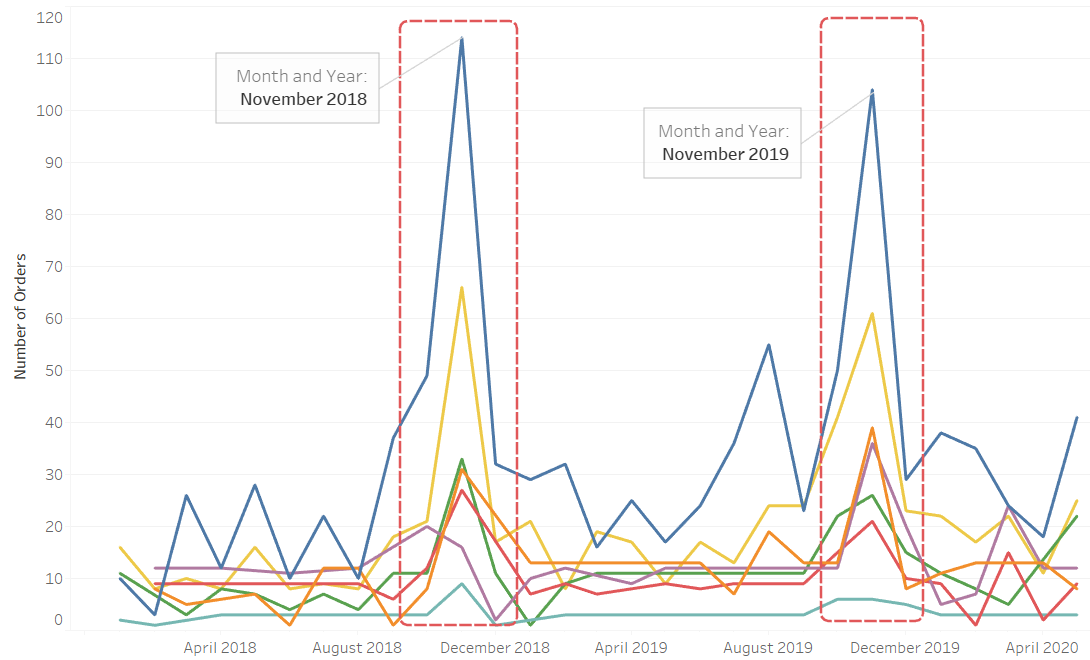
1. Who are the top 10 performing customers based on sales numbers and which product categories are the most preferred by these companies?



One of the best methods to compare several categorical variables such as Customer\_Name in our dataset is bar charts. The bar chart above highlights the top 10 performing customers ranked by their total sales. These customers represent a significant portion of overall revenue, with clear leaders emerging from different countries. It can be clearly seen that the top 2 customers Euro Shopping Channel and Mini Gifts Distributors Ltd. have significantly higher sales numbers then the other 8 customers and exceed $500.000 in terms of total sales while other 8 companies’ sales are between $150.000 and $200.000.

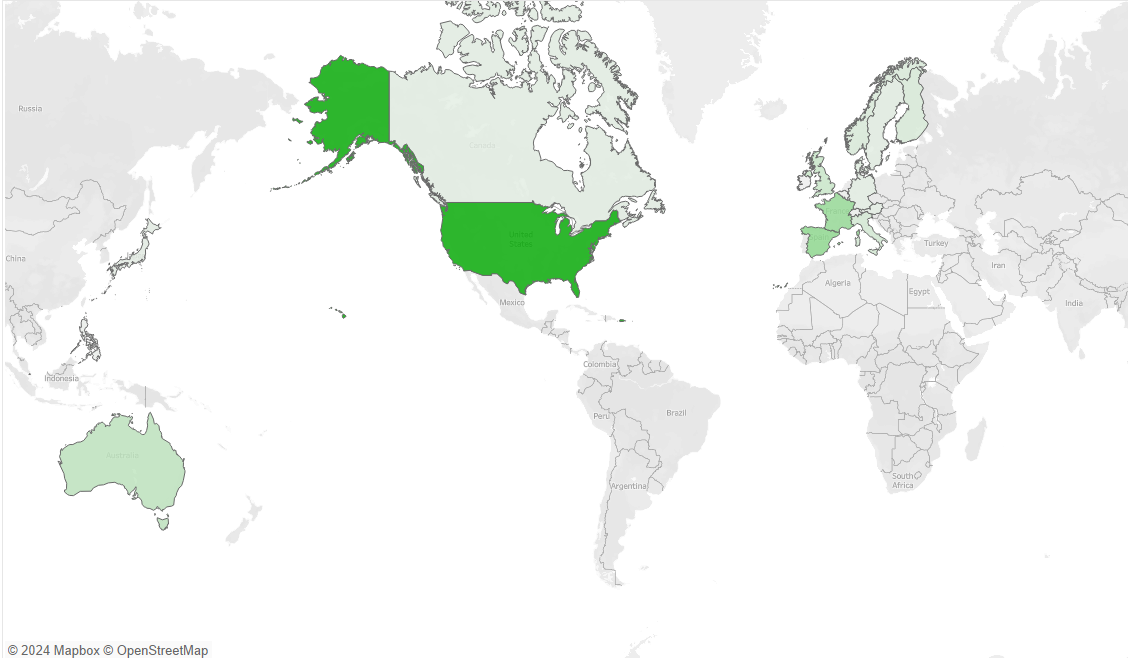
In order to have a better understanding of these customers’ purchasing patterns it is crucial for us to know their preferred product types. To make a comparison such as this, a stacked bar chart is one of the best visualization techniques. These 10 companies are from 6 different countries and we can clearly see that except for the customers from France and the UK the most preferred product type is Classic Cars and for France and the UK it is Planes and Vintage Cars respectively. This gives us valuable information to use it in the future to develop sale strategies aiming these 10 companies.

1. What are the monthly and yearly sales trends, and how do they vary by each product line?



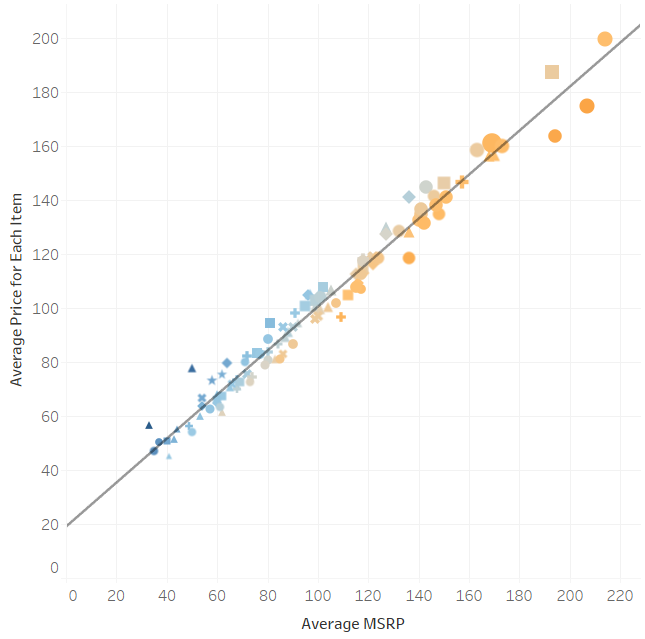
To show how the trends change in time, line graphs are one of the best options. In this graph, 7 different colored lines indicate 7 different product lines with the same color codes used in the stacked bar chart at Question 1. If we look at the sales data, it can be seen that November 2018 and November 2019 have instant spikes which indicates a higher number of orders for those months. Therefore we can clearly say that customers mostly place orders in the month of November.

1. Which countries have the highest and lowest sales performances?



In order to compare different countries, it is best to use a map and color it by the variable of the comparison. In our map above, the darker color indicates a higher sales number. The map shows that the United States with an outstanding sales number exceeding 3 million USD has the highest sales and there are other important countries with high sales numbers such as France, Spain, and Australia but the lowest performing country is Ireland with only $57.756 followed by Philippines with $94.016.

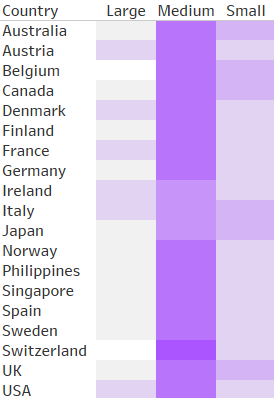
1. Are there significant differences between the MSRP (Manufacturer's Suggested Retail Price) and prices of each product across product lines, and how does this impact profit margins?



One of the best methods to show the relationship between two numerical variables, scatter plots are one best options. They let people see each data point and show the relationship between two numerical variables. The scatter plot on the left shows the relationship between average MSRP and average price for each item where each data point indicate a product, shape of the data points indicate the product line of the given product, the size indicates the number of sales, and the color indicates the percentage difference between average MSRP and average price for the items where the blue color means higher prices than MSRP and orange color means lower prices than MSRP. Although some products have been sold cheaper than its MSRP, generally products with lower prices than MSRP have higher sales numbers.

There are no significant differences between the MSRP and prices of each product. As the MSRP raises prices also increase with it and a linear relationship can be observed between these two variables.

5. What is the most common deal size for each country?



For a comparison for deal sizes, a heat map is an ideal visualization type. It can help users to clearly see how the counts of different types of deals change by each country. In this heat map, the user can clearly see that most of the deals are medium sized for almost all countries. Belgium and Switzerland have never done a large deal while Switzerland has the highest percentage of medium size deals. And last of all Austria, Denmark, France, Ireland, Italy, and the USA have done large deals more often than the other countries.

CONCLUSION:

This analysis of the AutoSales dataset provides actionable insights into customer behavior, sales trends, and product performance across multiple dimensions. By leveraging tools such as Tableau and Microsoft Excel, we uncovered critical patterns that can inform strategic decision-making and drive business growth.

Key findings include:

1. Top Performing Customers:

A significant portion of revenue is driven by a small group of high-value customers, with Euro Shopping Channel and Mini Gifts Distributors Ltd. leading in total sales. Understanding their purchasing preferences, particularly their interest in Classic Cars, offers opportunities for targeted sales strategies.

1. Seasonal Sales Trends:

Sales peak notably in November, indicating the importance of preparing for this period with inventory and marketing efforts.

1. Country-Level Performance:

The United States leads in total sales, while Ireland and the Philippines show the lowest performance, highlighting potential areas for market expansion or realignment.

1. Price-MSRP Relationship:

While no significant discrepancies between MSRP and product prices were observed, a linear relationship suggests pricing strategies should carefully align with MSRP to maintain profitability.

1. Deal Sizes:

Medium-sized deals dominate across most countries, but large deals are prevalent in key markets like the USA and France, suggesting opportunities to optimize sales strategies for different regions.

In summary, this analysis equips stakeholders with a clear understanding of customer preferences, regional trends, and product performance. These insights can be harnessed to refine marketing strategies, optimize pricing, and focus efforts on high-performing markets and customers. Future analyses could explore customer segmentation, profitability, and the impact of specific marketing campaigns to further improve decision-making.