STAT-112

TABLEAU

VISUALIZATION

PROJECT

By ÖZGÜR KILIÇ – 2614733

**ABSTRACT**

This project utilizes a detailed dataset that captures the sales performance of an automobile company across 195 countries, encompassing 14 distinct variables. By incorporating demographic data, economic metrics, and other contextual details, the dataset offers a comprehensive view of global sales patterns. Using Tableau for visualization and analysis, the project seeks to uncover critical insights into the factors shaping automobile sales, identify trends and anomalies, and facilitate informed decision-making. It emphasizes the relationship between regional demographic characteristics and economic conditions, providing valuable guidance for optimizing global sales strategies and enhancing market positioning.

**INTRODUCTION TO DATASET**

**NUMERICAL VARIABLES:**

QUANTITYORDERED: It indicates the number of items ordered in each order

PRICEEACH: This column specifies the price of each item in the order

ORDERDATE: It denotes the date on which the order was placed

DAYSSINCELASTORDER: This column represents the number of days that have passed since the last order for each customer. It can be used to analyze customer purchasing patterns.

MSRP: It stands for Manufacturer’s Suggested Retail Price and represents the suggested selling price for each item

LATITUDE: Latitude coordinate of the country’s location

LONGITUDE: Longitude coordinate of the country’s location

BIRTHDATE: Number of births per 1,000 population per year

CO2-EMISSION: Carbon dioxide emissions in tons

CPI: Consumer Price Index, a measure of inflation and purchasing power

GASOLINE PRICE: Price of gasoline per liter in local currency

GDP: Gross Domestic Product, the total value of goods and services produces in the country

LIFE EXPECTANCY: Average number of years a newborn is expected to live

POPULATION: Total population of the country

TAX REVENUE(%): Tax revenue as a percentage of GDP

TOTALTAXRATE: Overall tax burdan as a percentage of commercial profits

UNEMPLOYMENTRATE: Percentage of the labor force that is unemployed

URBANPOPULATION: Percentage of the population living in urban ares

**CATEGORICAL VARIABLES**:

ORDERNUMBER: This column represents the unique identification number assigned to each order

ORDERLINENUMBER: It represents the line number of each item within an order

STATUS: It indicates the status of the order, such as “Shipped”, “In Process”, “Cancelled”, “Disputed” , “On Hold” or “Resolved”

PRODUCTLINE: This column specifies the product line categories to which each item belongs

PRODUCTCODE: This column represents the unique code assigned to each product

CUSTOMERNAME: It denotes the name of the customer who placed the order

PHONE: This column contains the contact phone number for the customer

ADRESSLINE1: It represents the first line of the customer’s adress

CITY: This column specifies the city where the customer is located

POSTALCODE: It denotes the postal code or ZIP code associated with the customer’s adress

COUNTRY: This column indicates the country where the customer is located

CONTACTLASTNAME: It represents the last name of the contact person associated with the customer

CONTACTFIRSTNAME: This column denotes the first name of the contact person associated with the customer

DEALSIZE: It indicates the size of the deal or order, which are the categories “Small”, “Medium” or “Large”

COUNTRY: Name of the country

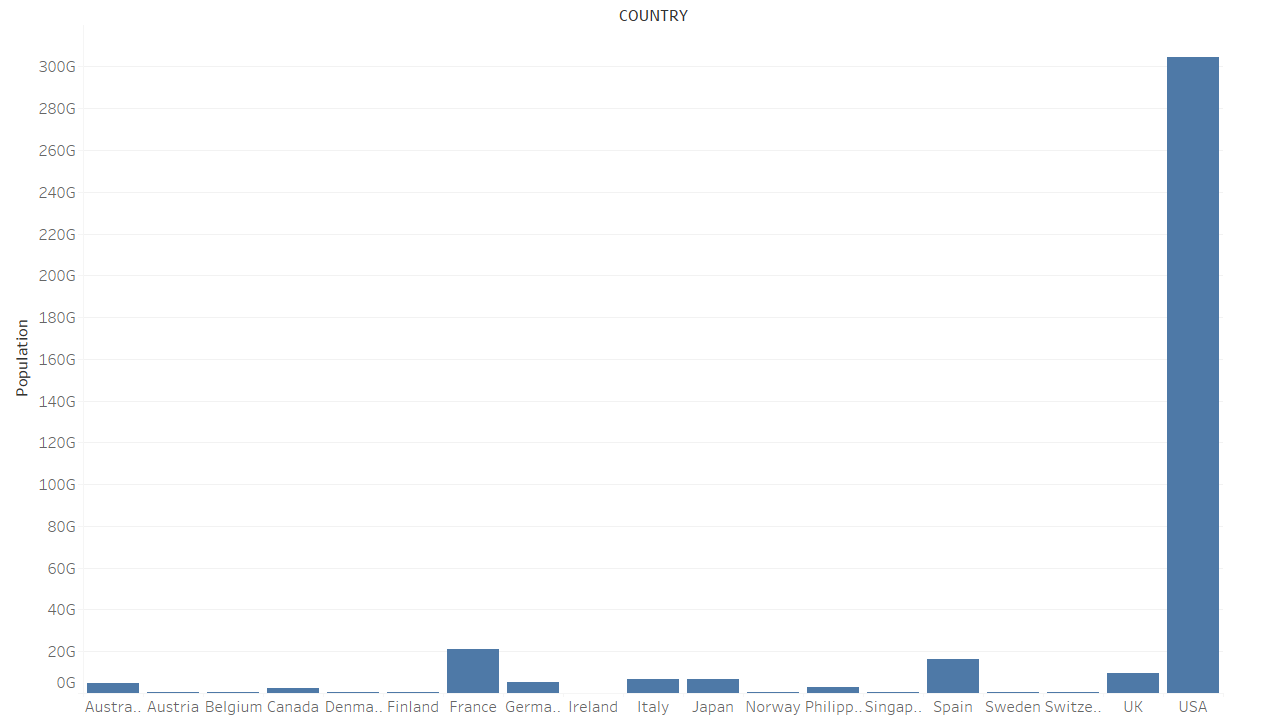
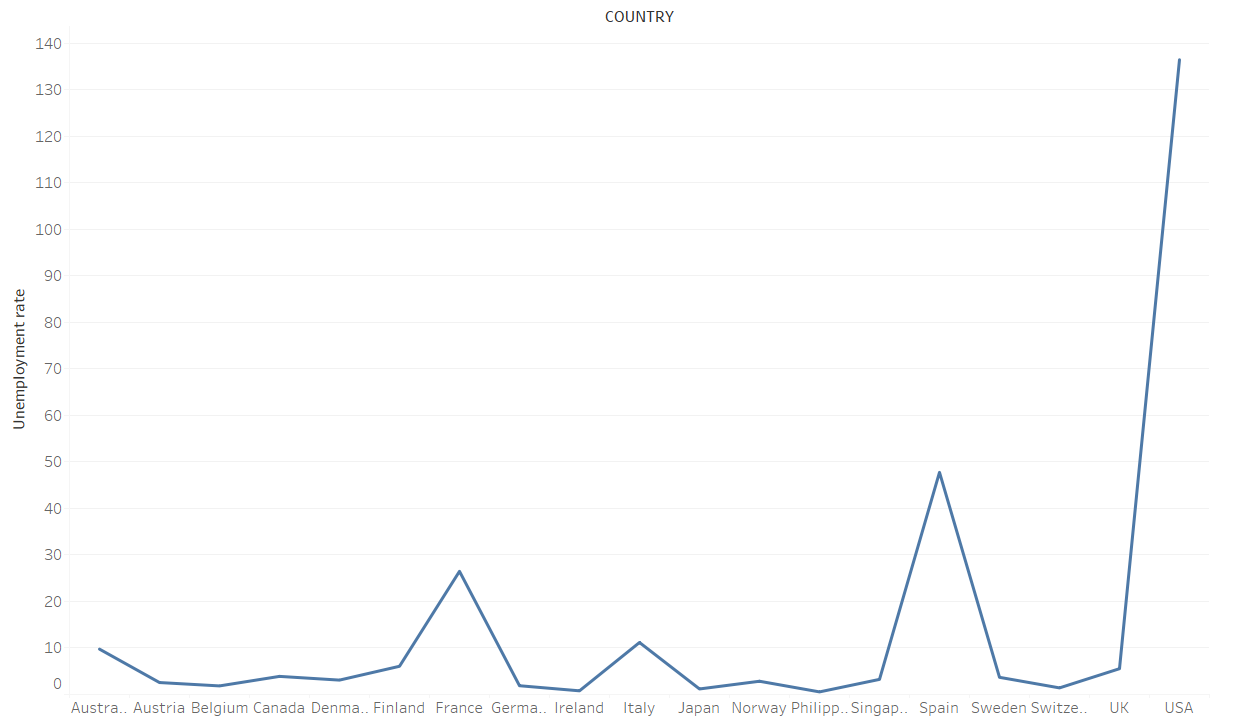
**DATA TIDYING AND CLEANING**

Ensuring data is well-organized and clean is essential for obtaining accurate and reliable results in any analysis. This typically involves addressing issues such as outliers, missing values, or inconsistencies in the dataset. However, during our initial assessment of the dataset used in this project, we found it to be exceptionally well-prepared, with no such problems identified. A thorough review of the columns and their corresponding data revealed the absence of missing values or outliers that could compromise the quality of the analysis.

As a result, we were able to use the dataset in its original form without the need for additional preprocessing. This allowed us to allocate more time and effort to exploring and analyzing the data rather than addressing data quality concerns. The completeness and accuracy of the dataset provided a strong basis for generating meaningful insights and reliable outcomes from our analysis.

EXPLORATORY DATA ANALYSIS

In this section, we will try to interpret and answer research questions about the datasets we have by using graphing methods in tableau.

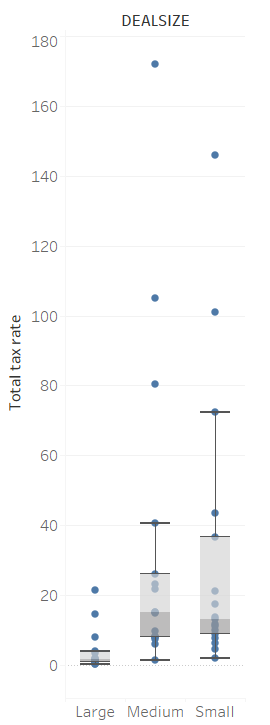
1-) Is there a relationship between countries' population density and unemployment rates?

Looking at the graphs above, we see that the line chart shows the percentage of unemployment by country and the bar chart shows the population by country. In the USA the population is much higher than in other countries and at the same time unemployment is also high. We can say that a similar situation applies to France and Spain. But if we compare, for example, Japan and Norway, Japan has a higher population than Norway, but a lower unemployment rate. As a result, we cannot say that there is a clear direct proportional link between population and unemployment rates.

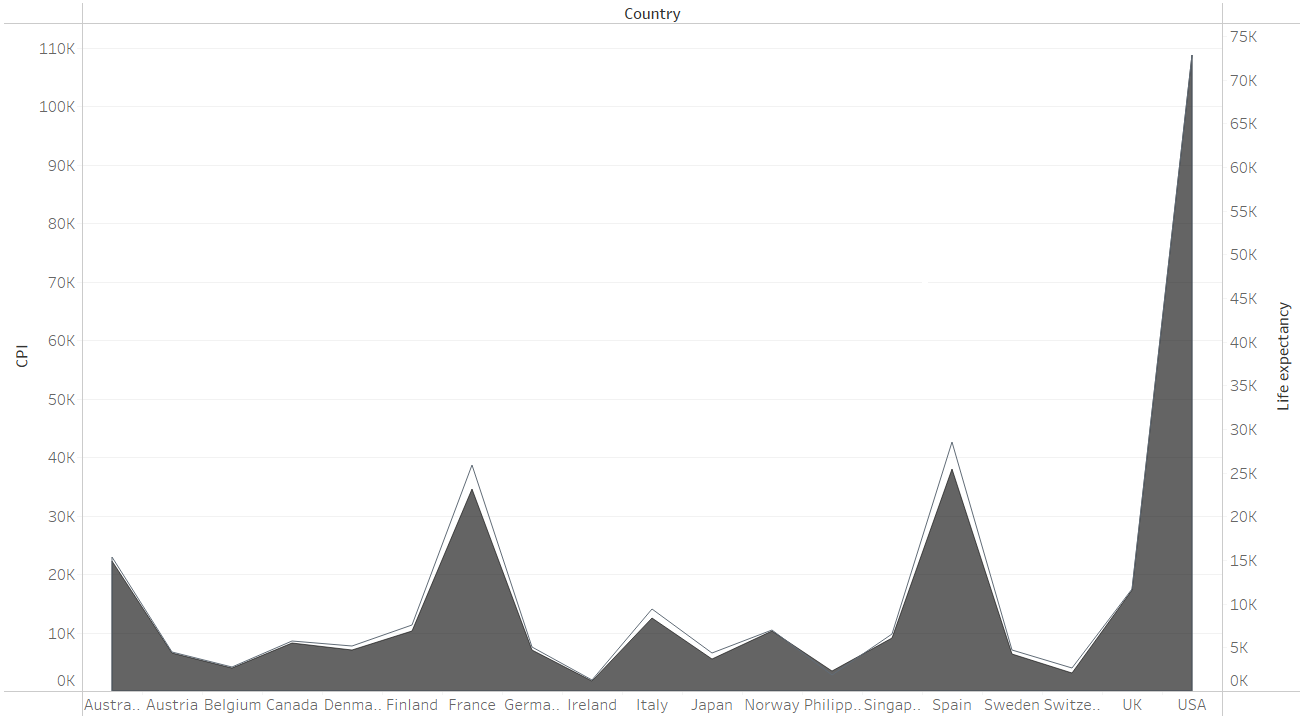
**2-)** Can a link be established between the size of products and tax rates?

When we look at the graphs below, we see that in the order quantities according to the sizes of the products, the most medium and small size products are ordered considerably more than the large size products. When we look at the tax rates according to the sizes of the products in the boxplot graph on the right, we see that the median tax rate in large size products is much less than the others. Based on these inferences, we can say that the answer to our size-tax rate question is partially directly proportional.

**metin, ekran görüntüsü, sayı, numara içeren bir resim

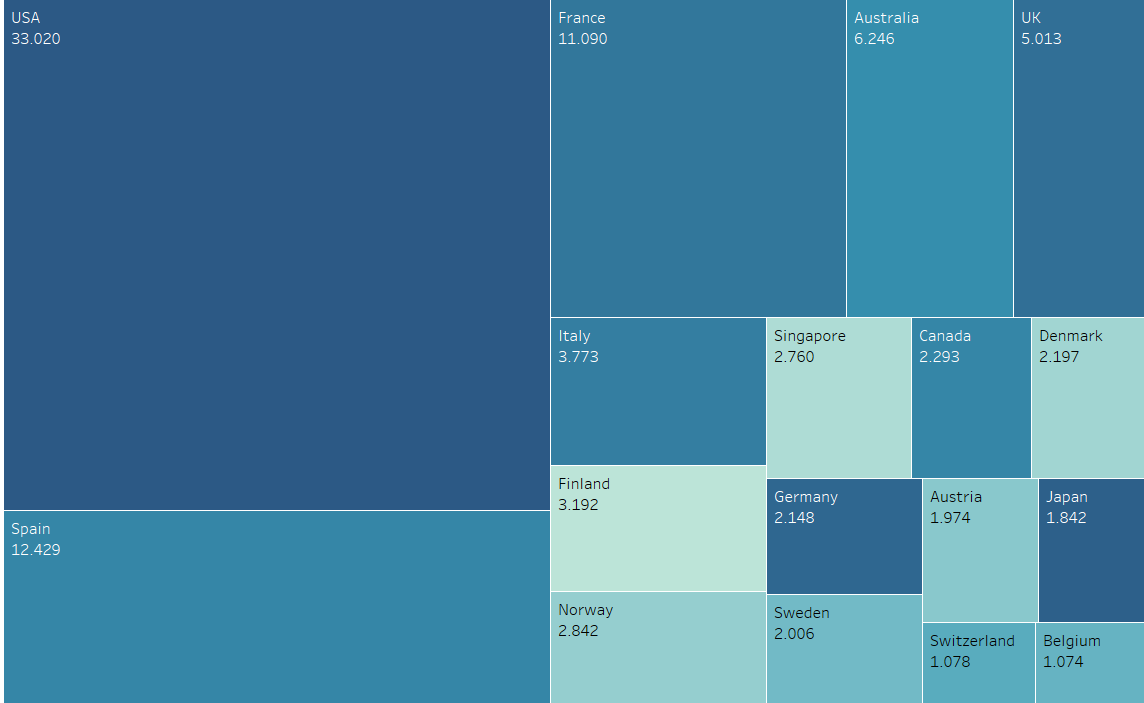
Açıklama otomatik olarak oluşturuldu**

**3-)** Is there a direct correlation between countries' annual CPI and Life Expectancy?

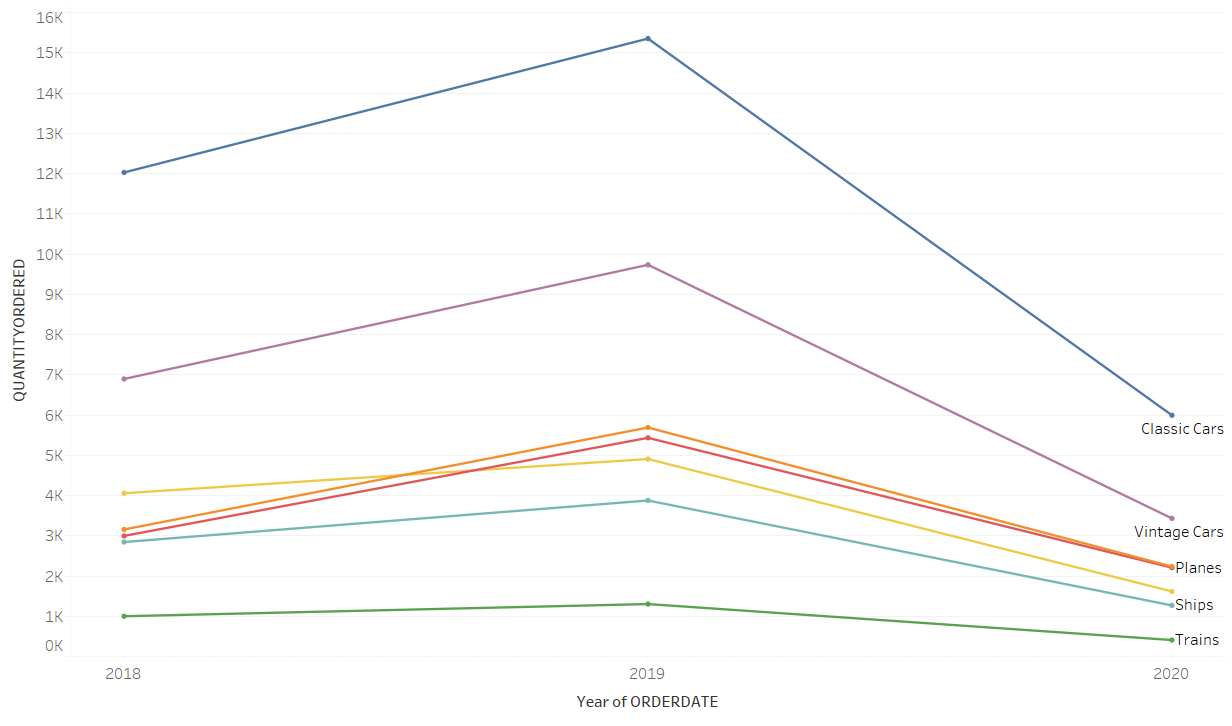
****

Looking at the graph above, the CPI and life expectancy of the countries are almost at the same rate. As a result, we see that there is clearly a direct proportional link.

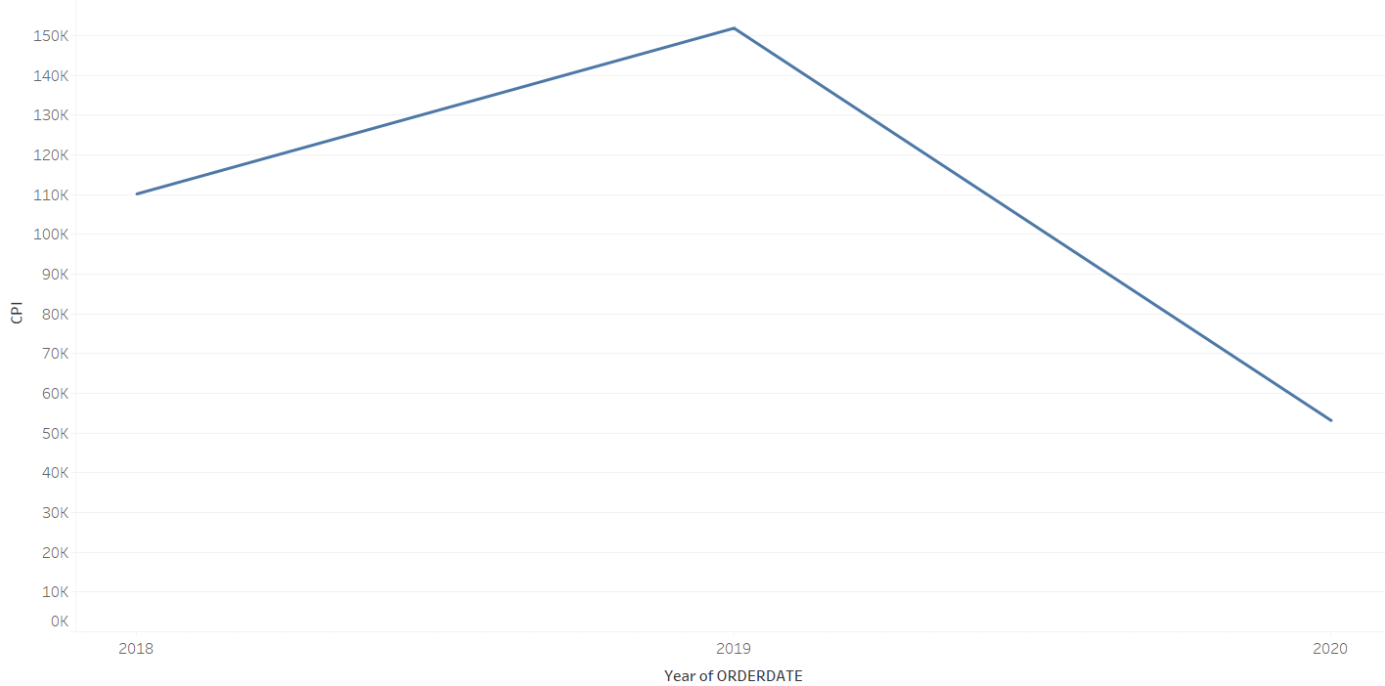
4-) Is there a direct correlation between annual GDP and quantity of orders?



This is a treemap table and in this table, the GDP amounts of the countries are compared and the highest amount is indicated in the boxes in the darkest way. Country names and Quantity of orders are also written in the boxes. When we examine the table, the country of germany draws attention because although it is one of the leading countries in terms of GDP, it is behind most countries in terms of quantity of orders. A similar situation is seen in the country of japan. Although when we examine countries such as USA Spain and France, it seems as if there is a direct proportion, we cannot say that there is actually a definite connection.

5-) How have sales trends and customer habits changed over years?

When the line chart is analyzed, there is a certain upward trend in all products from 2018 to 2019. After 2019, on the contrary, there is a decline.



Looking at the graph of CPI and years given above, there is an increase between 2018-2019 and a similar decline afterwards. Based on these two graphs, we can say that the decrease in the CPI amount of customers, i.e. the weakening of their purchasing power, directly affects the amount of products they order.

CONCLUSION

The results offer a detailed perspective on the data and highlight the need to evaluate trends within a broader context, including economic policies and cultural influences. Future research could delve into the root causes of these patterns to generate more practical and impactful insights.

* The lack of a strong correlation between population density and unemployment, as well as between GDP and order quantities, was unexpected and underscores the complexity of these relationships.
* The inverse relationship between product size and tax rates for large products was an intriguing observation that warrants further investigation.

LINK OF PROJECT AT TABLEAUPUBLIC: [VIZUALIZATION PROJECT - STAT112 | Tableau Public](https://public.tableau.com/app/profile/.zg.r.k.l./viz/VIZUALIZATIONPROJECT-STAT112/Story1?publish=yes)