**Evaluating Investment Potential in the Vehicle Dealership Industry**

Aghamusa Azizov  
Department of Statistics,  
Faculty of arts and sciences,Ankara, Turkey  
aghamusa.azizov@gmail.com

**Abstract**

Many individuals in the investment industry have questioned whether it is worth investing in the vehicle dealership business and what one should look out for while doing so. In this paper, I am trying to answer this and many other questions with dashboards, and sheets in the Tableau Public Version.

My dashboards are going to display many correlations, charts, and plots that have direct answers to the main questions. I believe the users who will be seeing this paper and dashboard most likely are people in the field of business.

Keywords— Car Dealership Investment, Business Analytics, Tableau Dashboards, Market Trends, Data Visualization.

Introduction (*Heading 1*)

In the introduction I will be introducing the data and the variables in it. Firstly, the autosale dataset includes values connected to the auto sales obviously and other factors related to this matter:

**ORDERNUMBER** – it represents unique identifiers for each order in the dataset;

**QUANTITYORDERED** – It shows the number of products ordered;

**PRICEEACH** – tells the price for each product in the dataset autosales;

**ORDERLINENUMBER** – it represents the number of each item in the given order;

**SALES** – the total amount of sales for each order in the dataset;

**ORDERDATE** - the exact date accurate till the day the order was made in our dataset;

**DAYS\_SINCE\_LASTORDER** – this column represents how many days have passed since the particular customer has made an order;

**STATUS** – labels status of the shipment of the order of all of the products in the dataset as either "Shipped," "In Process," "Cancelled," "Disputed," "On Hold," or "Resolved.";

**PRODUCTLINE** – This column tells exactly which product line the product belongs to such as “Classic Cars”, “Motorcycles”, “Planes”, “Ships”, “Trains”, “Trucks and Busses” and “Vintage Cars”;

**MSRP** – an abbreviation that stands for Manufacturer's Suggested Retail Price. It can help people decide if the price is reasonable in the given situations.

**PRODUCTCODE** – represents the code for each product available in the dealership in the dataset;

**CUSTOMERNAME** – it represents the full name of the customer who placed their order at the dealership in our dataset;

**PHONE** – denotes phone numbers of people or organizations who made the order in the dataset;

**ADDRESSLINE1** – gives us the first line of the address of the person who made the order;

**CITY** – this provides us with the city where the customer made their order from in the dataset;

**POSTALCODE** – the postal code also known as ZIP code that is associated with the customer's address;

**COUNTRY** – this denotes the country where the customer in the dataset made their order;

**CONTACTLASTNAME** – stands for the last name of the contact person who is connected to the customer;

**CONTACTFIRSTNAME** - this column provides us with the last name of the contact person who associates themselves with customer;

**DEALSIZE** – this denotes the size of the order given in the dataset by labeling it as follows: "Small," "Medium," or "Large.";

Now I will be explaining the variables of the world-data-2023 dataset:

**Country** – gives us the name of the country;

**Latitude** – the exact Latitude coordinates of the point on the world map where the given country resides;

**Longitude** - the exact Longitude coordinates of the point on the world map where the given country resides;

**Birth Rate** – stands for the count of births per 1000 people in the country’s total population annually;

**CO2-Emission** – the amount of Carbon dioxide emissions in the atmosphere measured in tons;

**CPI** – Consumer Price Index, a variable in the dataset that measures inflation and purchasing power;

**Gasoline Price** – this variable represents prices of petrol in each country priced at the country’s main currency;

**GDP** – an abbreviation stands for Gross Domestic Product, the full value of products produced in each country;

**Life expectancy** – the mean amount of years the person is expected to live when they are born;

**Population** - total number of people residing in the given country;

**Tax Revenue(%)** – The revenue given as the percentage of Gross Domestic Product in each country;

**Total tax rate** – Total tax burden presented as a percentage of commercial revenue;

**Unemployment rate** – the percentage of people from the given country in the working age that are unemployed;

**Urban\_population** – the percentage of the total population of the whole country who are living in urban regions of the country.

# Data Preprocessing

I used operations in Tableau to complete the cleaning data stage. zn(Lookup(SUM([PRICEEACH], 0)) (e.g.) to change the missing values in column **PRICEEACH** to zero.

If zn(SUM([PRICEEACH])) = 0 THEN 80,000 ELSE SUM([PRIEEACH])  END (e.g)

to change the missing value in **PRICEEACH** to the value we want in this situation we want mean value.

I noticed that **ORDERNUMBER** was a dimension so I converted it to string since it’s unique identification.

I countered the duplicates by finding a column that should not be duplicating which was **ORDERNUMBER** and made a calculated field called Unique PRICEEACH and inserted such a command

{FIXED [ORDERNUMBER] : MIN([PRICEEACH])  this command would group all duplicating **ORDERNUMBER** as one and show the **PRICEEACH** as a new variable called unique PRICEEACH.

# Exploratory Data Analysis

To get the most out of the data given I made some questions to answer and discuss.

1. First of all, the question that I think is the most crucial since it will decide the future of the business is: Which year in recent times was the best for the car marketplace, and are sales expected to grow in the future?

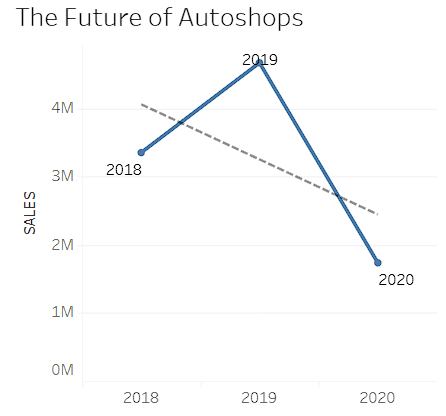


Figure 1.1

As you can see in the figure I made a line chart that indicates the total sales for the corresponding year. Even though three years are not quite enough to make a prediction and the data is outdated due to the fact that it is missing 2021, 2022, 2023, and 2024, I would like to foretell that the sales are going down due to the trendline having a negative slope. I wish I had a wider timeline to compare and predict more accurately. Seeing this unfortunate outcome I would suggest waiting a little longer until the trendline gets a positive slope coefficient.

Overall, the best year for the vehicle market was 2019 since it had 4,669,925 sales while others had 3,353,014 and 1,737,283.

2. Secondly, the second question I would like to state that this question will help the reader decide the location of the business.

In which countries does MSRP align with an actual price?

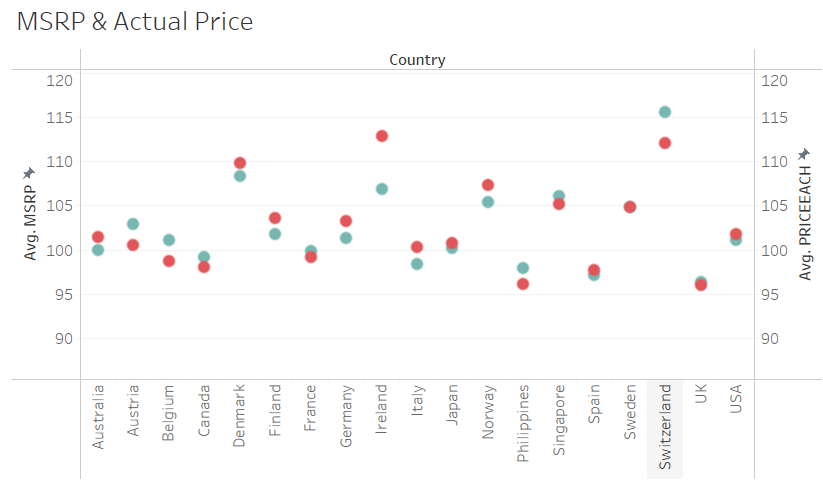


Figure 2.1

In my scatterplot, you can see how close or how far the are dots referring to MSRP and PRICEEACH. Red for PRICEEACH and gray for MSRP which as I mentioned before stands for Manufacturer's Suggested Retail Price.

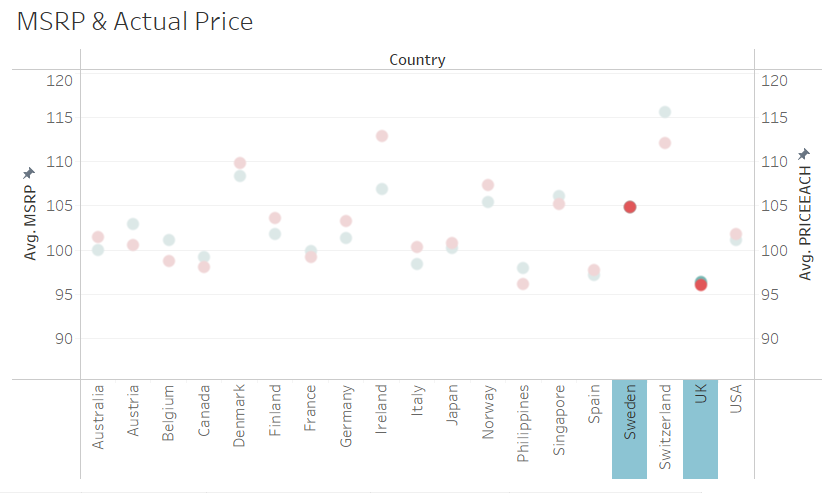


Figure 2.2

As you can see the countries in the figure with almost equal MSRP and PRICEEACH are Sweden and the United Kingdom. This implies that people in these countries are more flexible when discussing prices. I would like to suggest that potential investors focus on these two countries since they have more chances to get their desired profit from a sale.

3. My next question is: Which deal size has the most amount of sales? From this question, we will get the products that have the most interest in the public internationally. This will ensure that we don’t market the wrong product and not be left in a bad financial situation.

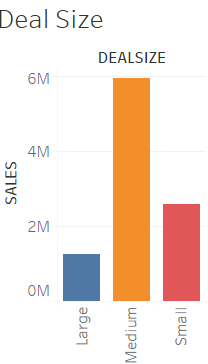


Figure 3.1

In the figure, it is obvious that medium has much more sales than it's more popular in the general public. So it would be more logical to focus on investing in Medium deal size products.

4. The following question is more of a moral dispute that will decide whether is it possible to get into this field without any consequences worsening our atmosphere and environment globally: Would investing in the car dealership business hurt the environment by increasing CO-2 levels?

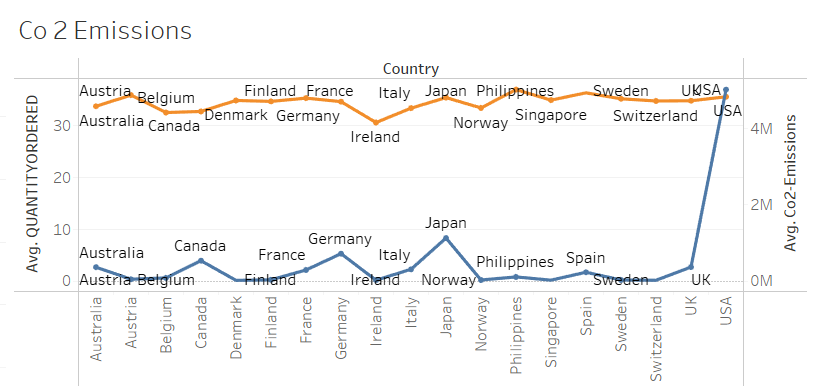


Figure 4.1

In the provided line chart you can witness how the lines that state the average QUANTITYORDERED and average Co-2 Emissions share some spikes as you can see in Figure 4.2 but they are still very different. So I would like to state that they have almost close to zero correlation. This gives us the green light to start investing in this field.

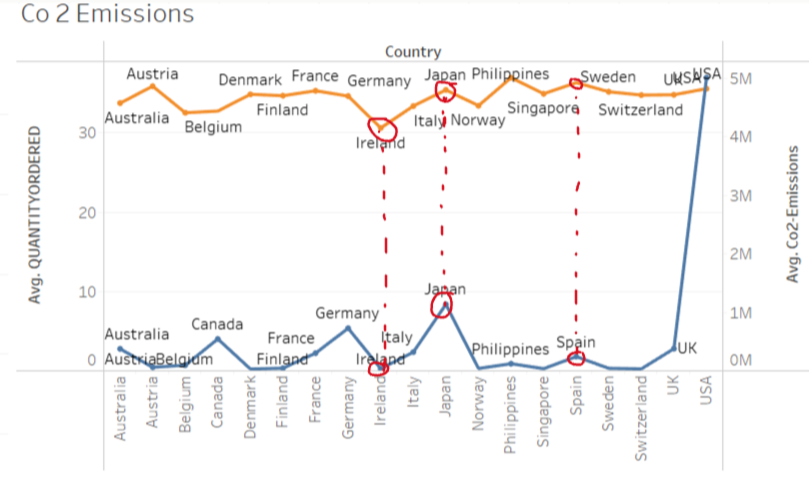


Figure 4.2

5. With this question we are coming back to the topic of choosing the location and narrowing down the choices between potential investments: Are the car dealership businesses more prominent among the urban residents?

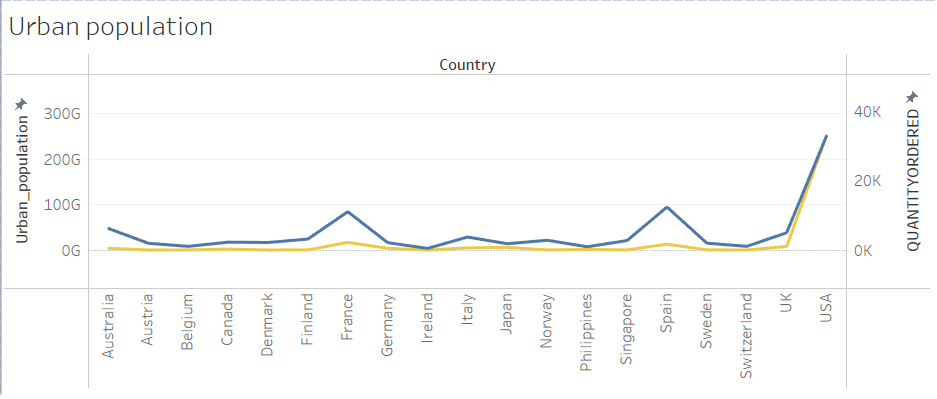


Figure 5.1

Both lines have most of their spikes aligned as you can see I highlighted in the figure 5.2, so I

would like to state that they correlate them. With correlation, I can make the claim we as investors should focus on the urban population, and make them our target audience.

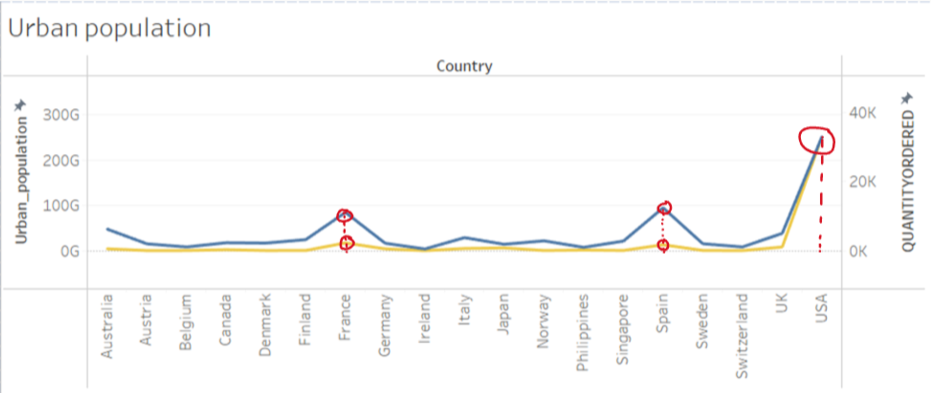


Figure 5.2

6. In the next question we will be coming back to the scatterplot correlation between MSRP and PRICEEACH: How do MSRP and actual prices differ across countries?

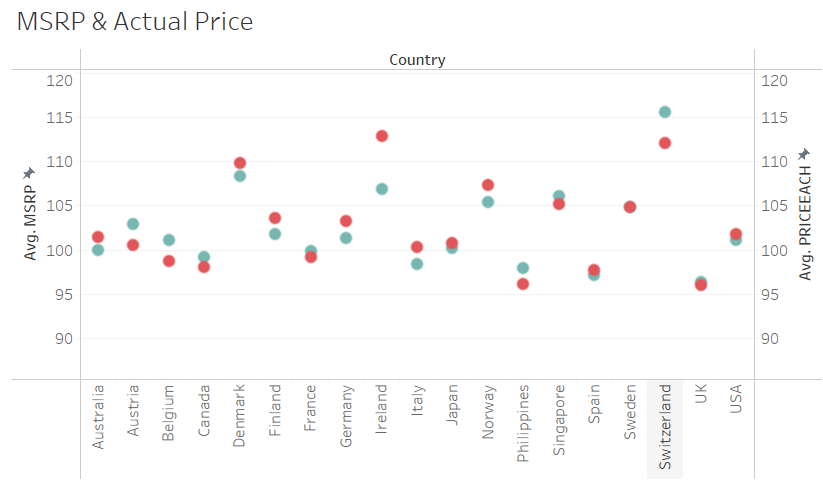


Figure 6.1

As we can see in Figure 6.1 the differences are actually drastically varied due to many factors most of which we are not aware of. So this information will provide some factors to filter the potential investments.

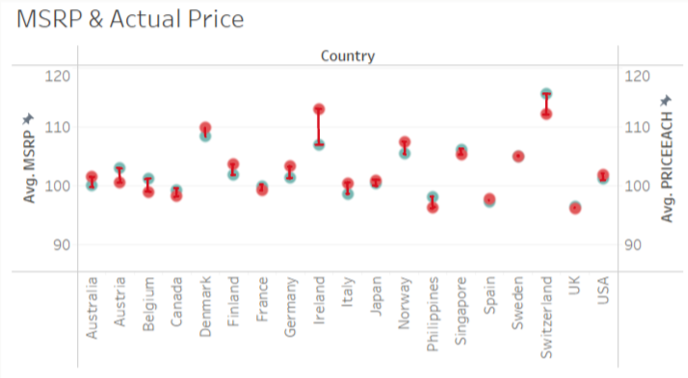


Figure 6.2

7. Is there decline in sales after 2019, what can explain it? As we seen before there is a decline in sales after 2019. I have made a research and concluded that there are two main events that lead to this decline.

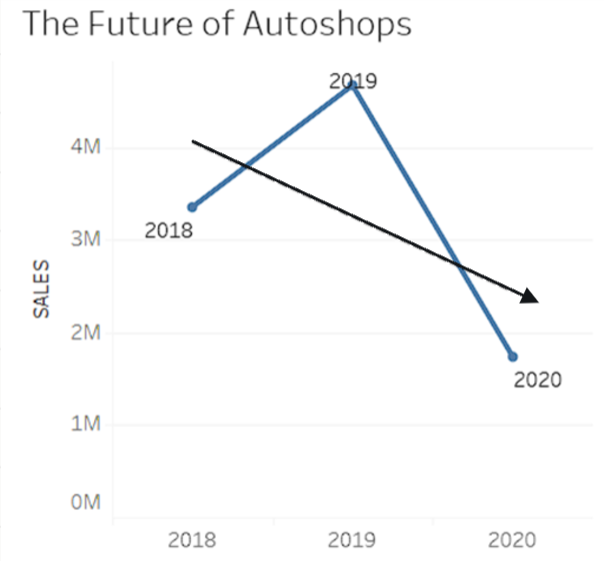


Figure 7.1

(1). I believe the first one has to be the COVID-19 epidemic since it had a major effect on a lot of markets and I believe that the auto field is no exception.

(2). The second factor is the rise of electric cars. Since as we saw the PRODUCTLINE variable has no “Electric cars” subcategory. So it would be logical to assume that as the number of other lines decreased the number of electric cars were increasing.

# Conclusion and discussion

##### All in all, for the final chapter, I would like to discuss the chances and the possibilities and conclude all my findings.

##### From the questions I answered using my dashboard, we can see that selling, reselling, and trading autos is not quite profitable nowadays. The sales are decreasing as we can be witnessed in Figure 1.1 in the first question. This gives us an idea about the worth of auto sales in the present and in the future.

##### Let's say you still want to invest in this business and desire the most out of this unprofitable field, I would suggest investing in companies that are based or base your own company in either Sweden or the United Kingdom since it will guarantee that you will get the price you want and there will be almost no bargain about the price according to Figure 2.2 ( Sweden’s and United Kingdom’s MSRP and actual price are practically the same.)

##### The product you should be selling should be middle-sized and have electricity charging because of the recent rise in the market of electricity cars and descend of sales of other types of cars as seen in Figure 1.1 and as stated in the answer to the 7th question.

##### Your target audience should consist of urban residents and maintain their interest by again focusing on the middle-sized electrical vehicles. References

1. Michael J. Coren, Quartz, “2019 was the year electric cars grew up,”

https://qz.com/1762465/2019-was-the-year-electric-cars-grew-up

1. Joint statement by ILO, FAO, IFAD and WHO, World Health Organization,” Impact of COVID-19 on people's livelihoods, their health and our food systems**”** https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems

My dashboard: https://public.tableau.com/app/profile/aghamusa.azizov6115/viz/therealrealdeal/Dashboard1#1