The BRITISH UNIVERSITY IN EGYPT

Introduction to Data Science

23CSAI01I

Dr. Nahla Barakat

MOHAMED KHAIRI – 224483 – A3

ABDELRAHMAN GAMAL – 227998 – A3

KARIM YASSER – 206703 – A2

Cars Dataset - Provided by the Dr. Nahla Barakat

Description:

In this dataset we have 9 columns.

1) **Car**: This column represents the name of each car.
2) **MPG**: This column represents cars fuel efficiency. A lower MPG indicates better fuel consumption, and a higher one means that the car consumes a lot of fuel.
3) **Cylinders**: This column represents the number of cylinders for each car engine, and there is a very close relationship between it and the displacement column.
4) **Displacement (engine size)**: This column represents the total size of all of an engine's cylinder. High displacements have high horsepower.
5) **Horsepower**: This column represents the horsepower of all car engines (power output), and a car with a higher horsepower will have a higher number of cylinders.
6) **Weight**: This column represents the weight of the car. And the higher the weight of the car (heavy), the greater the displacement of the car.
7) **Acceleration**: This column represents the car's acceleration and is calculated in seconds. Cars that have high acceleration have less horsepower.
8) **Model**: This column represents the year in which the car was manufactured. The newer model ranges from 79 to 82, and the newer the model, the higher the fuel consumption.
9) **Origin**: This column represents the country that manufactured the car and can be useful for comparing each car in different regions or determining the concentration of each origin.

The BRITISH UNIVERSITY IN EGYPT

Assignment 1

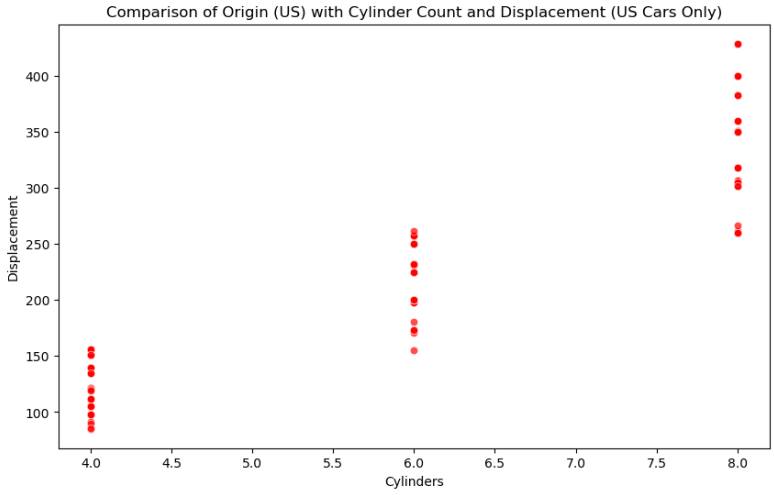
Answers for Questions based on Provided Dataset and Visualization:

1) How (US) cars are always noticed to be V8 (8 cylinders). For what reasons (US) manufactures tend to use large displacements capacities and do that effect the Milage per Gallon (MPG)?

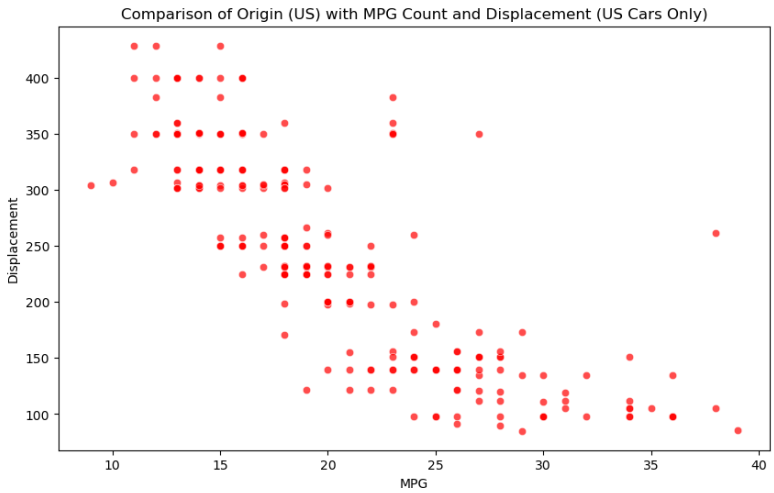
Hint: American citizens prefer large displacements engines, in other words high performance cars.

Answer: Almost all US cars have large engines in another words they are equipped with High Displacement engines with large number of cylinders. Nevertheless, US cars are can obtain covering high mileage per gallon and have high MPG. And this what makes it more preferred by Americans and they prefer cars that could travel for high distances.

Visualization:



Notice:
When displacement increases number of cylinders increases and vice versa.



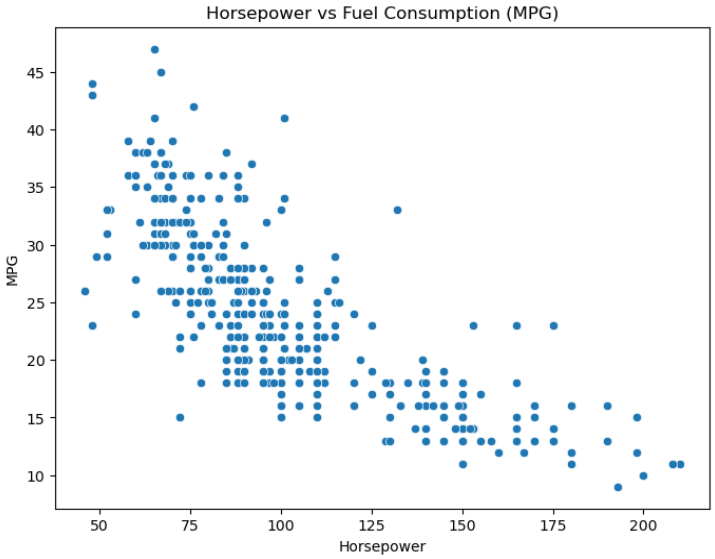
Notice:
When Displacement increases MPG Decreases and vice versa.

2) How high horsepower cars tend to have low fuel consumption?

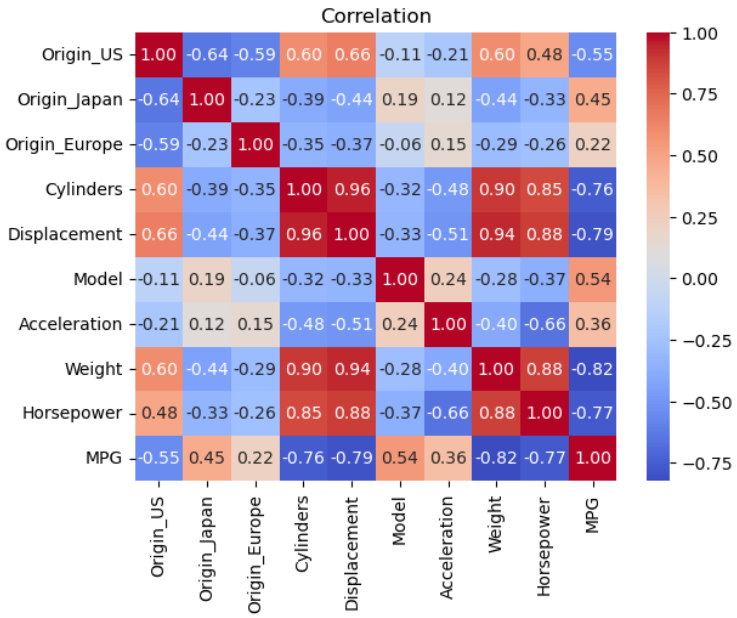
Hint: When we need to increase horsepower fuel consumption (MPG) decreases sequentially.

Answer: in this cars dataset there is a relation between Horsepower and MPG where, when increasing horsepower MPG decreases. In other words, cars that have high horsepower are noticed to be fuel efficient.

Visualization:



Notice:
Horsepower is inversely proportional with MPG.



Notice:
From the correlation visualization Horsepower is related with MPG with around -0.77, this means that is a close relationship.

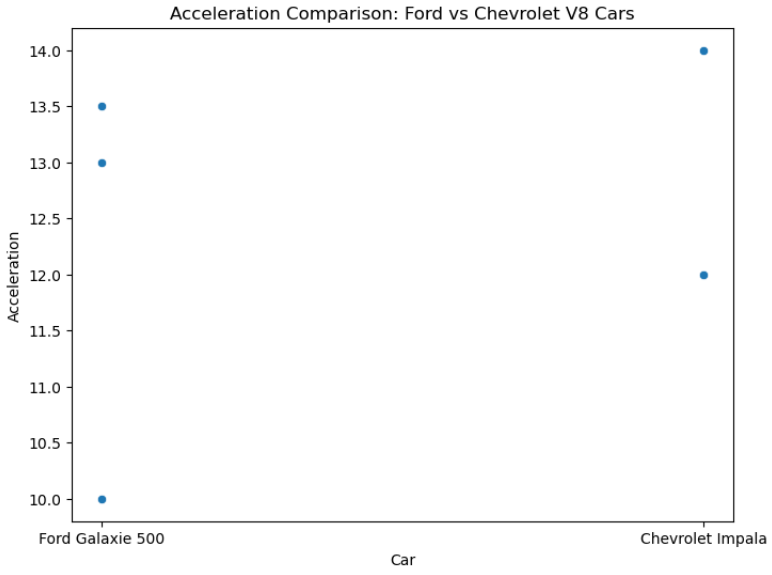
3) How drivers who are interested in (US) cars claim that Ford V8 (8 cylinders) cars performs better than Chevrolet's V8 (8 cylinders) cars interims of weight and Milage Per Gallon (MPG) and acceleration.

Hint: Ford engineers have efficient engineering designs than Chevrolet's.

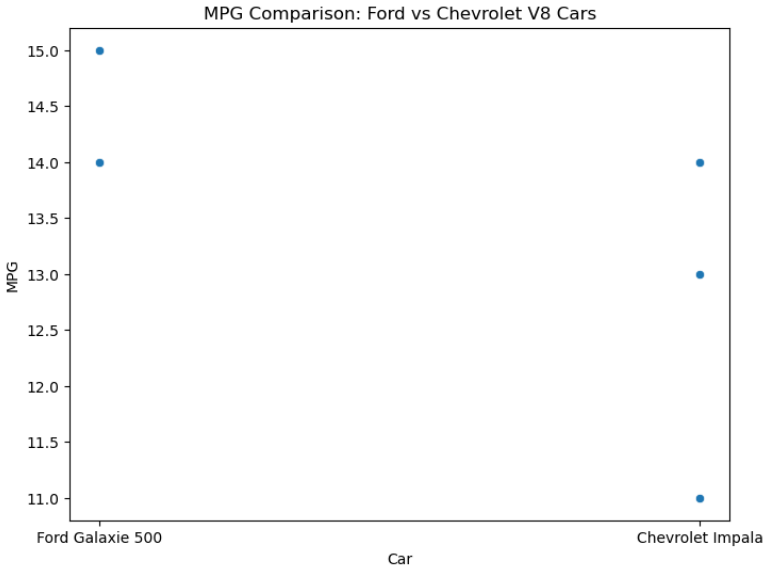
Answer: When talking about both brands in general Ford might be better than Chevrolet. However, according to the visualization, Chevrolet outperforms ford in many fields. Nevertheless, we could easily discover that each car maker has a point to shine in.

Visualization:

This visualization inculdeds copmapring 4 american V8 cars from both ford and Chevrolet inorder to have a clearer knowlage of which car maker outperformes the other.

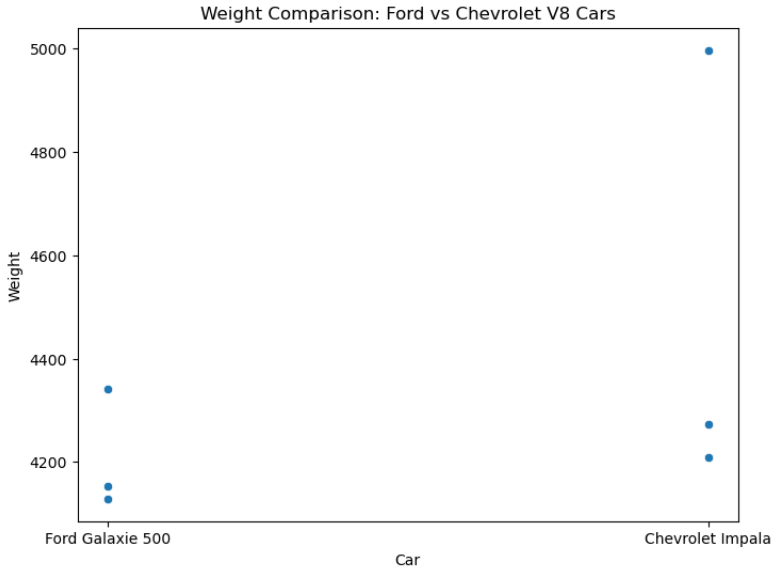


Notice:
(Ford Galaxie 500 vs Chevrolet Impala) Chevrolet Impala has higher acceleration most of the time in other words Ford Galaxie 500 reaches higher speeds faster than Chevrolet Impala.



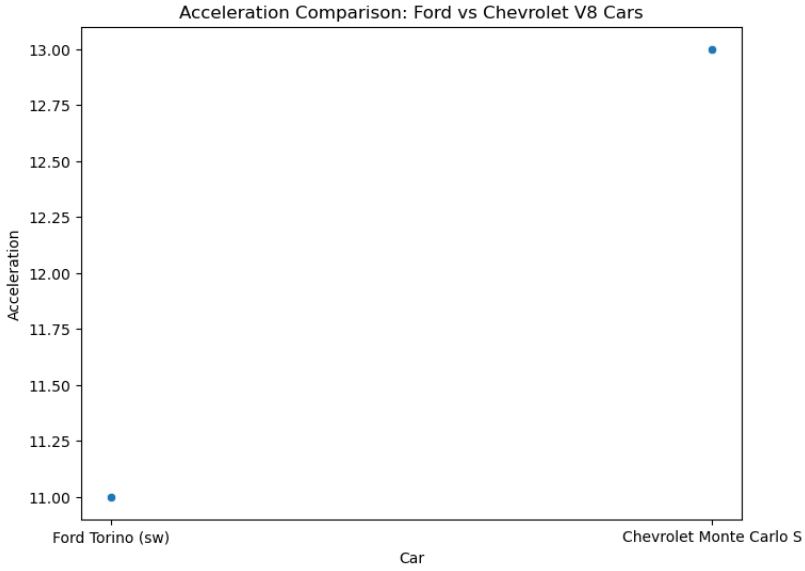
Notice:

(Ford Galaxie 500 vs Chevrolet Impala) Chevrolet Impala has better fuel consumption than Ford Galaxie 500



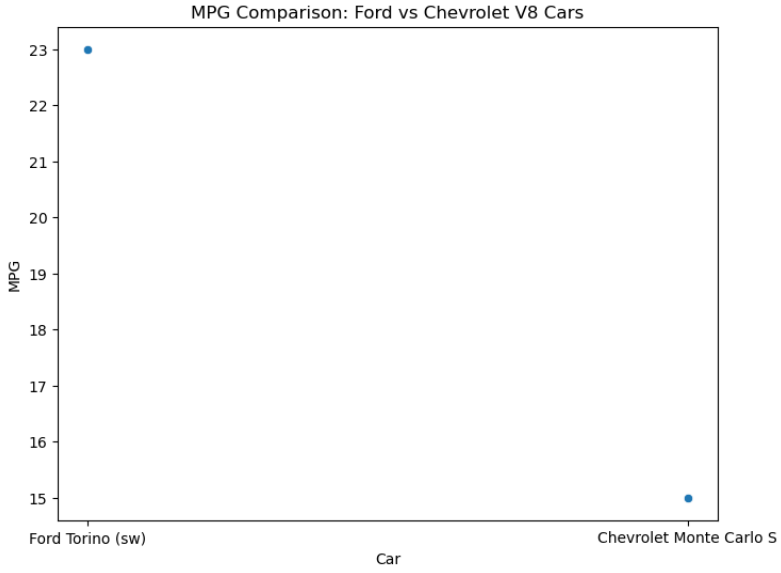
Notice:

(Ford Galaxie 500 vs Chevrolet Impala) Ford Galaxie 500 has lower weight than the Chevrolet Impala



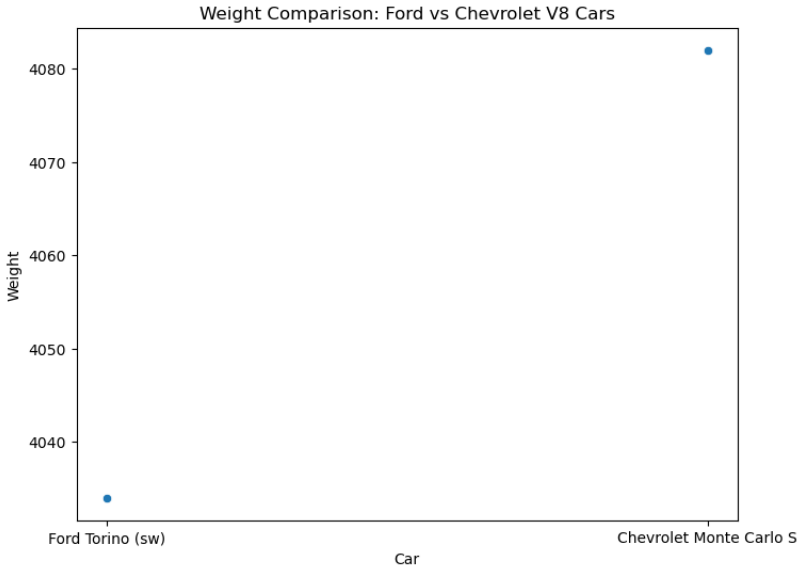
Notice:

(Ford Torino (SW) VS Chevrolet Monte Carlo S) Chevrolet Monte Carlo S has a lower acceleration time with 2 seconds after the Ford Torino (SW).



Notice:

(Ford Torino (SW) VS Chevrolet Monte Carlo S) Chevrolet Monte Carlo S has low MPG when compared to the Ford Torino (SW), which means that Ford Torino (SW) could cover higher distances per gallon than Chevrolet Monte Carlo S.



Notice:

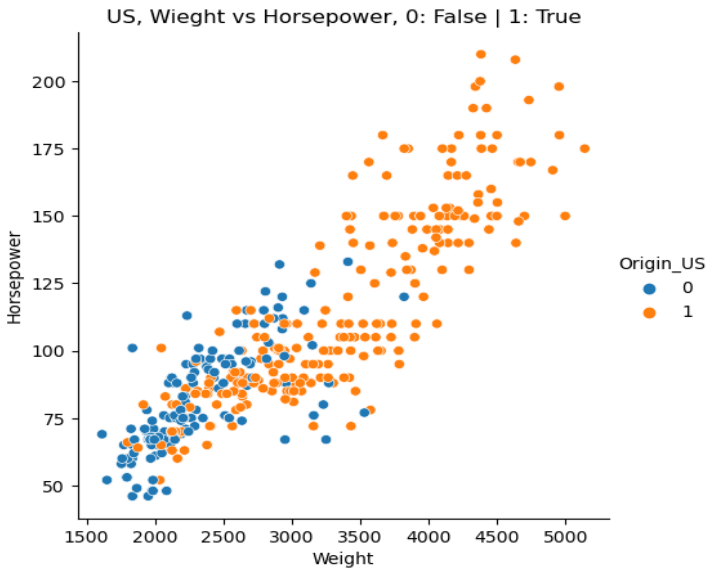
(Ford Torino (SW) VS Chevrolet Monte Carlo S) *when it comes to wight the Chevrolet Monte Carlo S is heavier (have high wight) when compared to the Ford Torino (SW).*

4) Why identical cars in weight and model from different origin have various horsepower's and consumptions in (MPG), in other words does the country of origin affect the engineering quality?

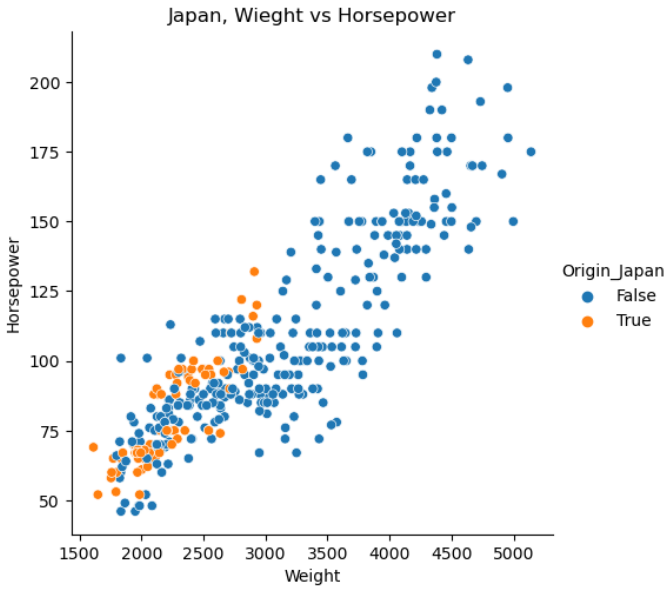
Hint: Environment Eco-friendly rules in Europe are so strict so that car manufactures tend to scarify performance (displacements and fuel consumption (MPG)) in favor of being eco-friendly.

Answer: Manufactures around the 3 car origins of manufacturing have different strategies when it comes to cars and engineering as for instance when comparing the 3 origins of manufacturing (US, Japan, Europe) we could find that each country has its approach to focus on.

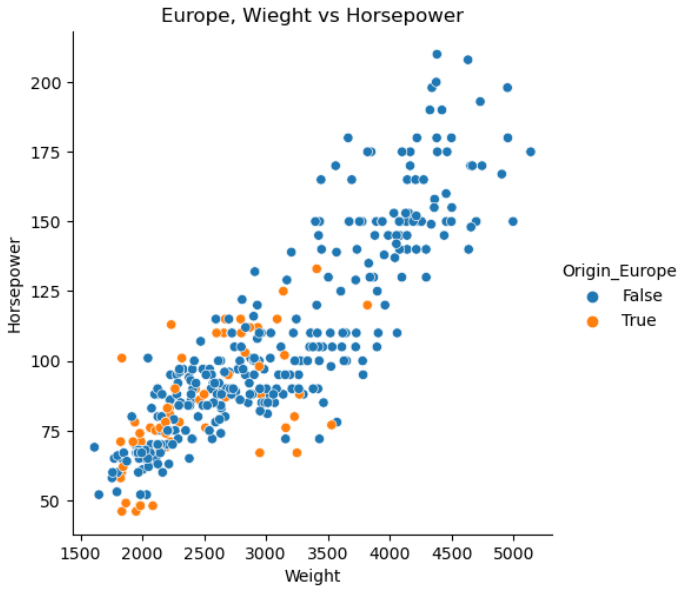
Visualization:



Notice:
US Manufactures always tend to have e high horsepower and large weight.

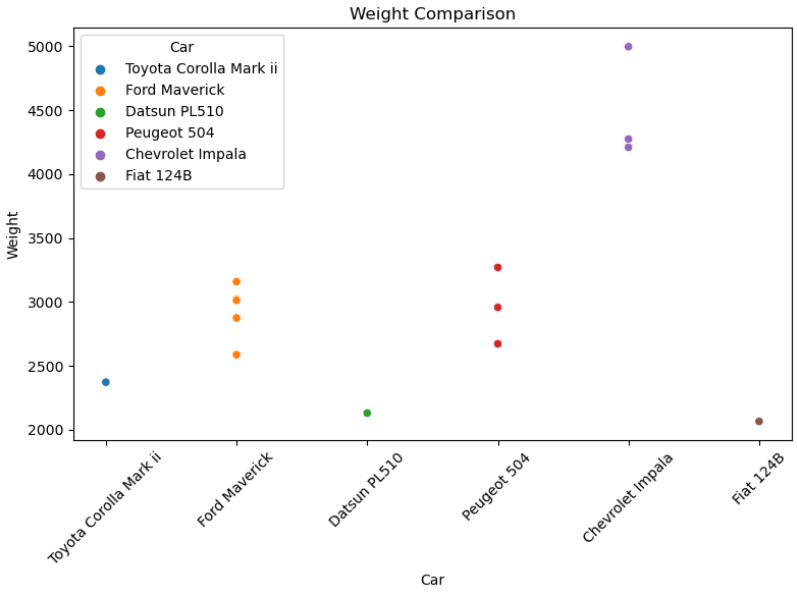


Notice:
Japanese cars manufacture prefers making Cars that are low on Horsepower and efficient in wight as well.

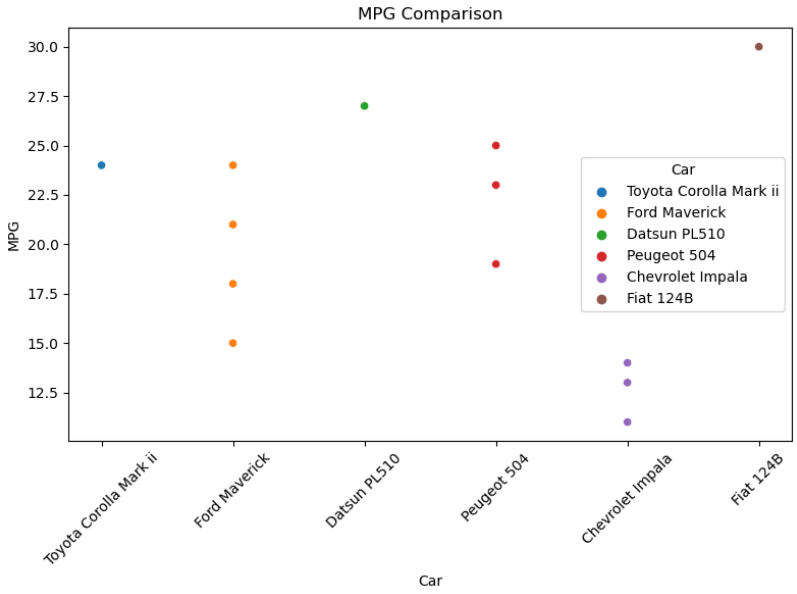


Notice:
Cars that are made in Europe are almost similar to Japanese when it comes to efficient low wight and Low horsepower.

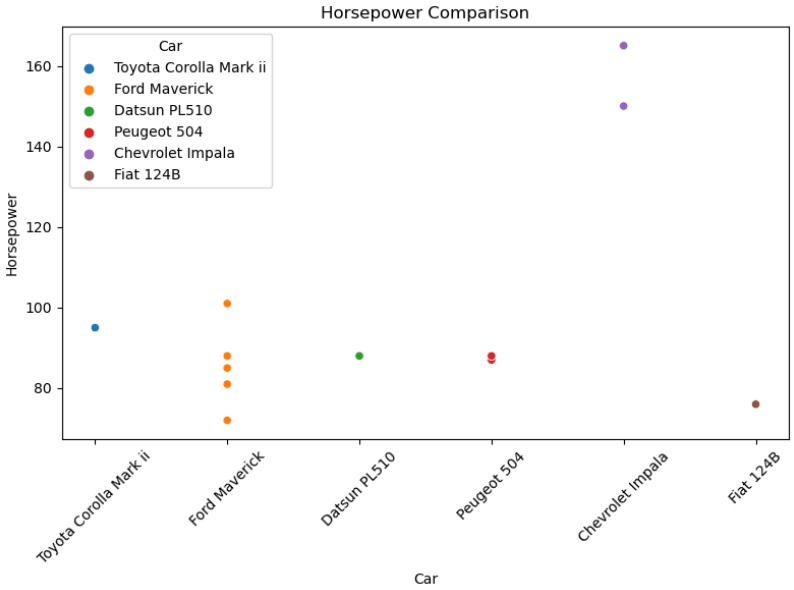
Visualization *for justifying the answer (comparing 6 cars from the 3 origins of manufacturing)*



Notice:
Chevrolet Impala and Ford Maverick (US cars) are having large weights while European manufactures tend to keep weight in range. On the other hand, Japanese manufactures tend to make their cars low in weight.



Notice:
As mentioned in the first question, US cars tend to be fuel efficient than both Japanese and European cars, coming in the second place are Japanese cars makers with midrange MPG not to Hight but at the same time not considered fuel efficient. on the other hand, European cars have high fuel consumption.



Notice:
When it comes to horsepower US cars shine, while other origins of car manufacturers like (European and Japanese cars) are limited in horsepower when compared with the US car maker.

Overall:
US cars have high weight but good in performance and good MPG, while Japanese cars European cars shine interims of wight.

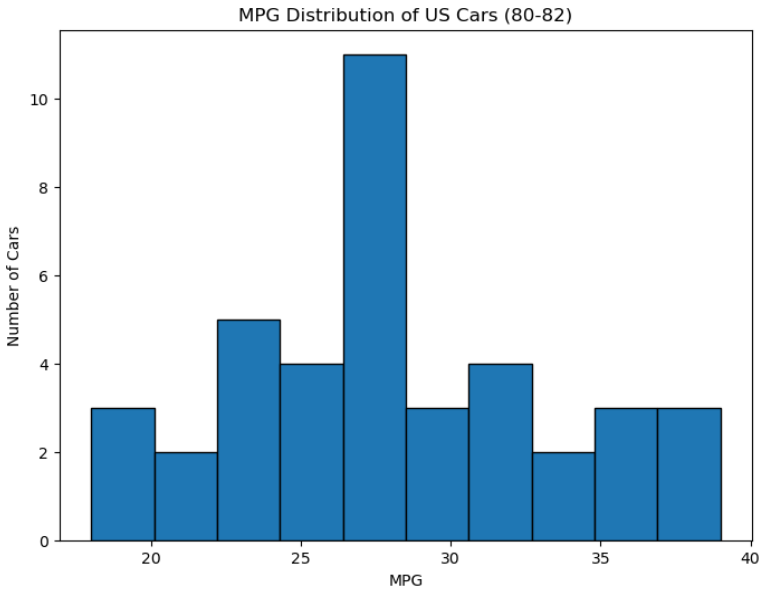
Conclusion:
Each origin has an Idle design in favor of manufacturing, when it comes to engineering cars this might be to weather conditions or preferences of citizens of origin.

5) How 70s (US) cars used to have high displacement in their cars, and they are maintaining to use this large displacement nevertheless, these cars have high (MPG) with their newer models, in other words (During 80s models). On the other hand, European manufacturers tend to use less displacement in their models across the period (70s to 80s) while maintaining higher or the same acceleration as (US)?

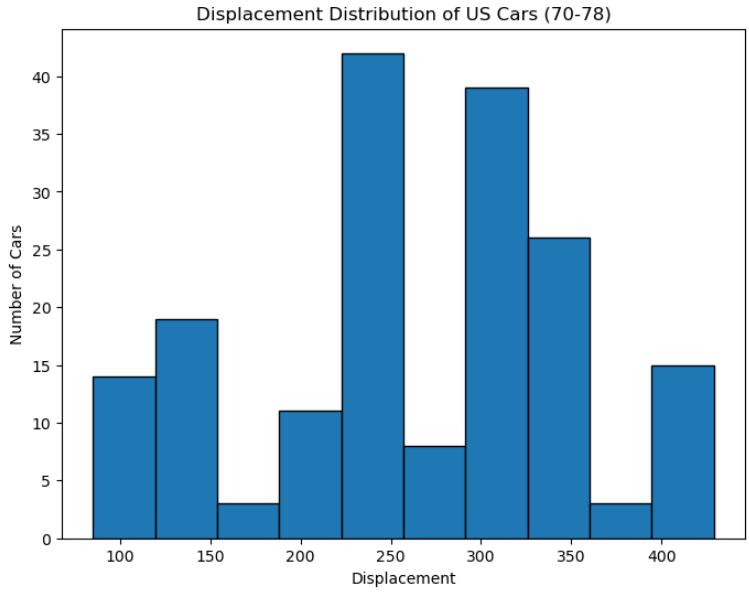
Hint: Some car manufactures believe in using efficient materials as well as displacement per cylinder which decreases (MPG) to achieve their goals regarding efficiency.

Answer: From 70s to 80s cars manufacturing has changed a lot as cars started to consume more fuel (MPG) and to have lower displacement compared to older cars like US cars which their newer models (80s) have lower displacement compared to older models (70s). Let's break this down, first the US cars which are considered the most fuel efficient (lower MPG) with consumption ranging from US cars 25 to 30. Moreover, US cars have high acceleration time range between 14 and 19 which is high compared with the European mid-range acceleration which is about 15. Nevertheless, European car makers across 70s and 80s maintained to have quite lower displacement.

Visualization:

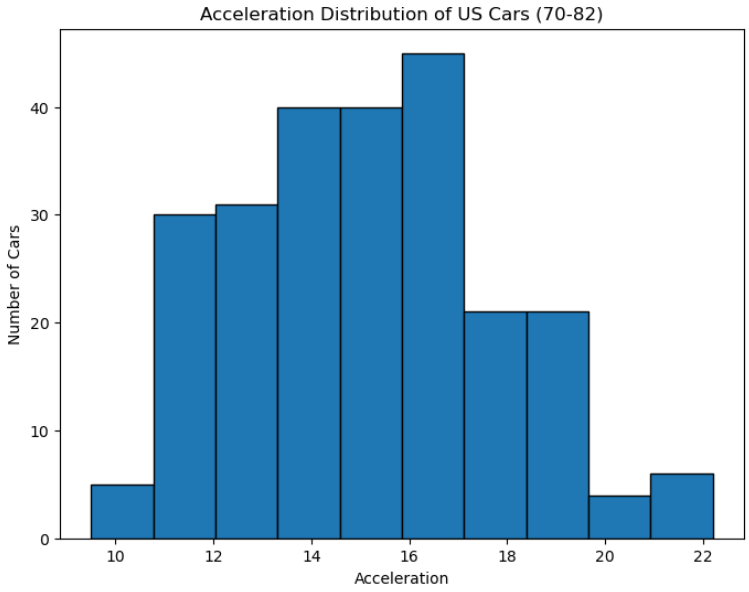


Notice:
US cars have an MPG consumption ranging from 25 to 30 in other words US cars could obtain high milage (distance) per gallon.



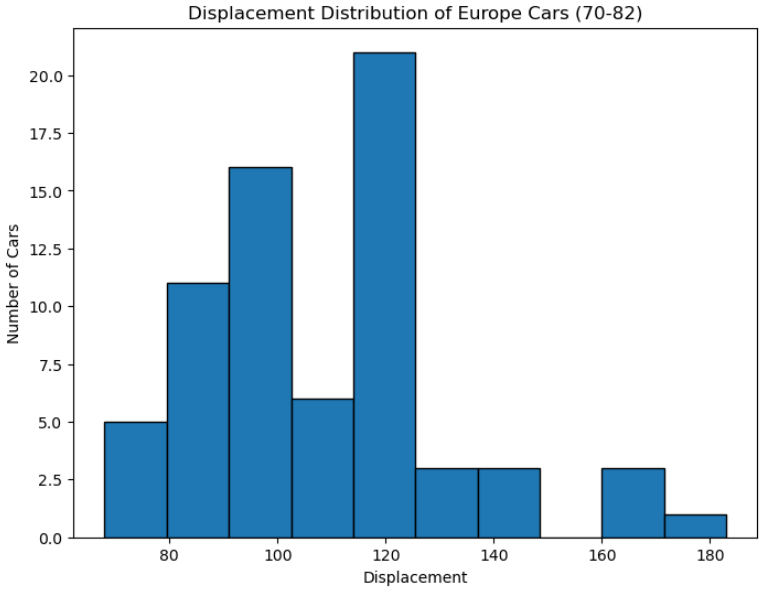
Notice:

Almost all US Cars During 70s used to have high displacement.



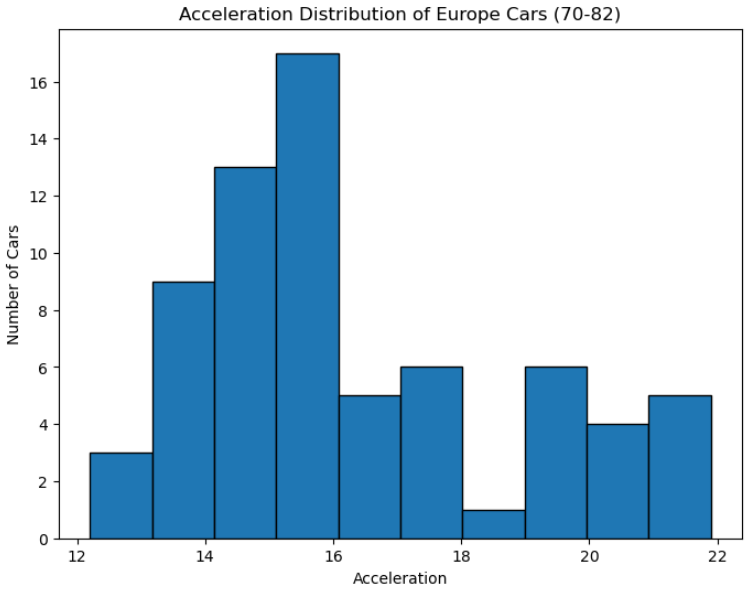
Notice:

American cars tend to have high Acceleration in range between 14 and 19.



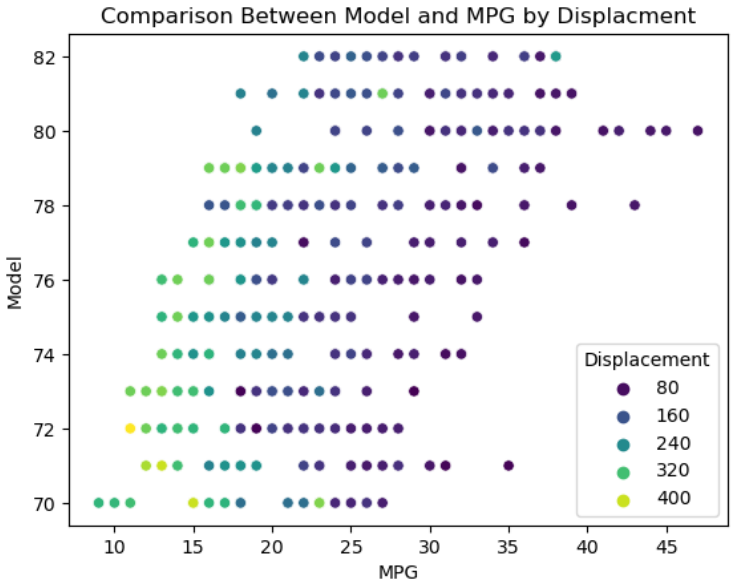
Notice:

European cars tent to have low displacement engines as many Europe cars have displacement from 80 to 120.



Notice:

European cars have quite low acceleration time about 15.



Notice:

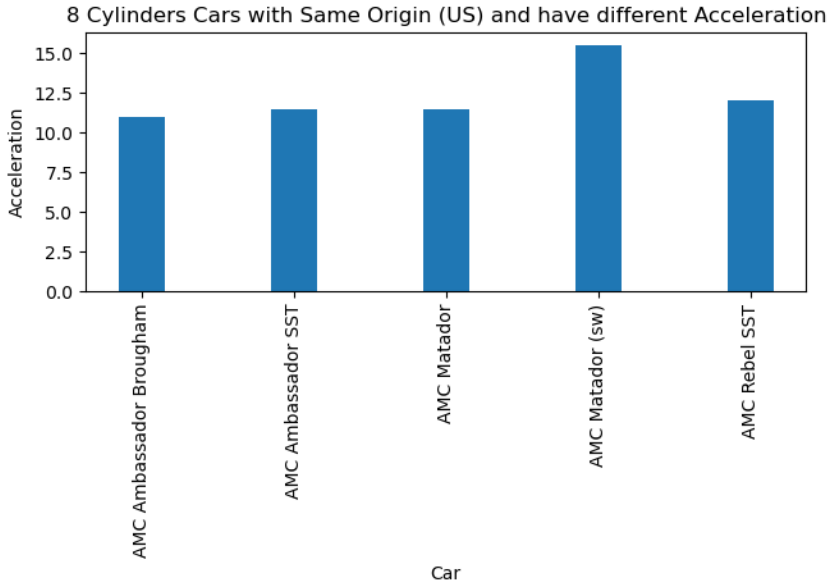
Cars during 80s have higher consumption than cars during 70s. However older cars have higher displacements than newer ones.

6) Does cars that have the same origin and the same number of cylinders have different acceleration?

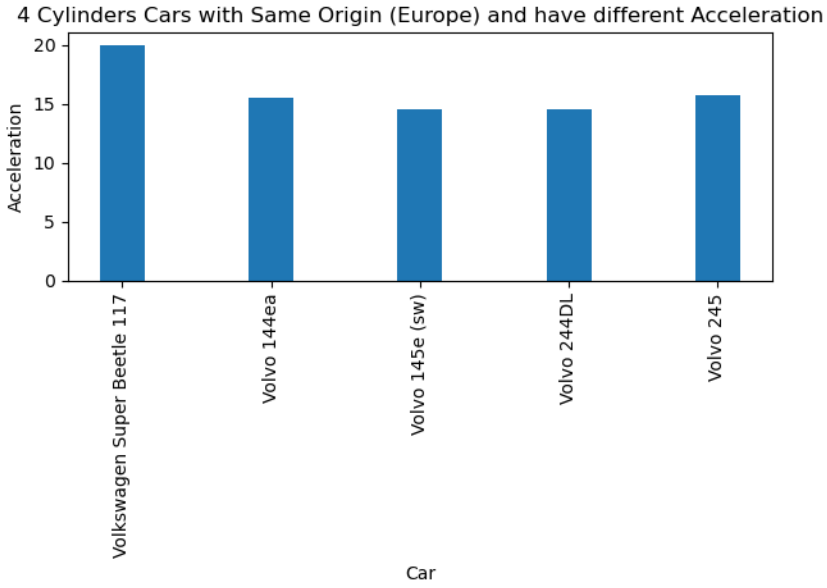
Hint: Some manufactures believe in having low power while achieving a high life span while others believe in providing the driver with the best performance out of the engine while not caring how long will the car last.

Answer: car makers in the same region tend to have various acceleration for cars from the same model and the same origin these is because of different manufacturing techniques and different engineering from one company to other.

Visualization:



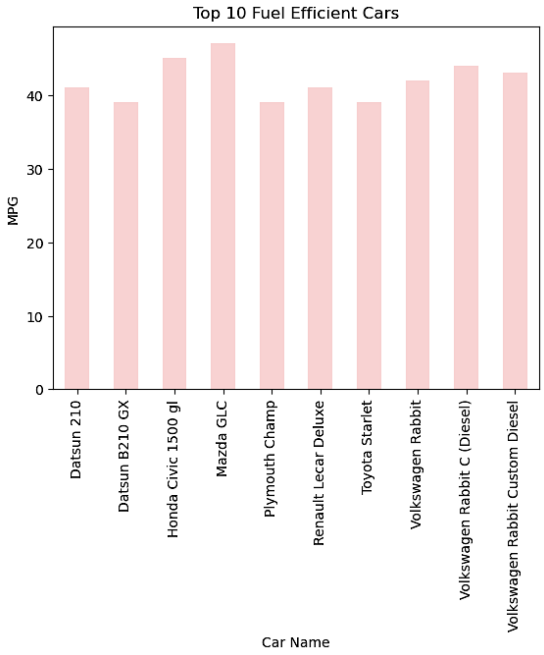
Notice:
All US car makers are in range with maintaining a mid-range (10 to 15) acceleration, Leaded by AMC Matador sw.



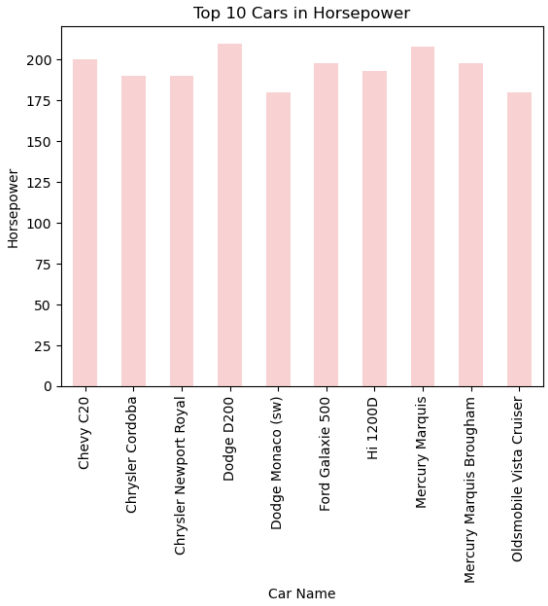
Notice:

Most Europe car makers are maintaining a high acceleration specially, Volkswagen Super Beetle 117, Nevertheless some have a mid-range like the Volvo cars 144ea and 244DL.

Insights:



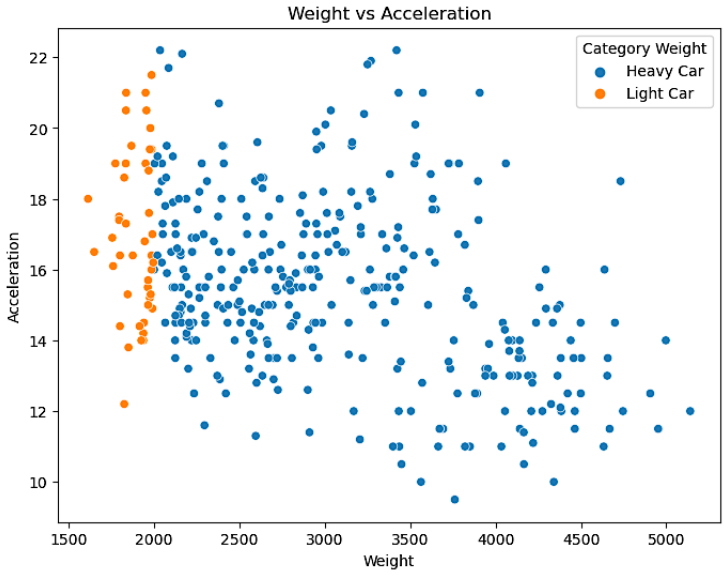
According to the analysis and visualization, these are the top 10 cars in terms of fuel efficiency, and this will be ideal for those who want cars that do not consume a lot of gasoline.



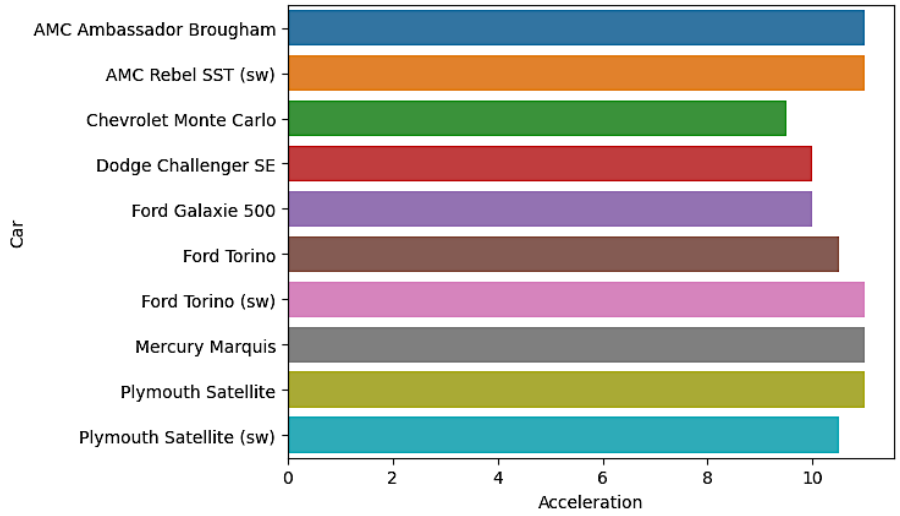
According to analysis and visualization, these are the 10 best cars in terms of horsepower, and this will be ideal for those who passionate and want powerful cars.

This is a table containing the name of each car, the number of cylinders, displacement, and horsepower.

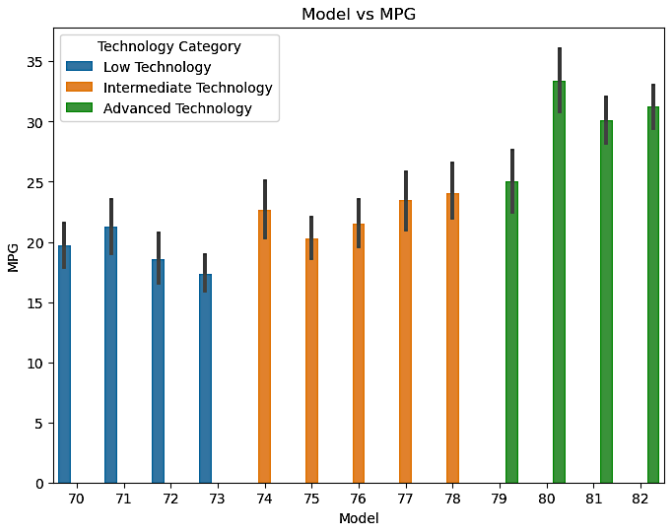
	Car	Cylinders	Displacement	Horsepower
0	Chevy C20	8	307.0	200.0
1	Chrysler Cordoba	8	400.0	190.0
2	Chrysler Newport Royal	8	400.0	190.0
3	Dodge D200	8	318.0	210.0
4	Dodge Monaco (sw)	8	383.0	180.0
5	Ford Galaxie 500	8	429.0	198.0
6	Hi 1200D	8	304.0	193.0
7	Mercury Marquis	8	429.0	208.0
8	Mercury Marquis Brougham	8	429.0	198.0
9	Oldsmobile Vista Cruiser	8	350.0	180.0



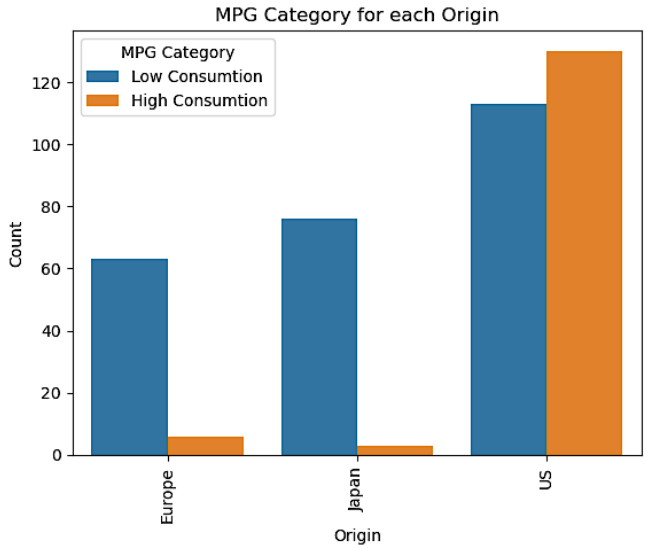
According to analysis and visualization, most light cars have higher acceleration than heavy cars, so this would be ideal for people who are interested in low acceleration to turn to heavy cars.



According to analysis and visualization, these are the 10 best cars in terms of acceleration, and this would be ideal for anyone who wants to buy a fast and responsive car.



According to visualization, cars with advanced technology could cover high distance, so anyone who wants to save fuel will resort to cars with advanced technology, because cars with low or intermediate technology consume more gasoline.



According to analysis and visualization, American (US) cars have the best fuel consumption (low consumption) in high distance, while Japan comes next and the worst in fuel consumption is Europe. This tells us that we can improve on Japanese and European cars and is ideal for people who want to buy cars that do not consume a lot of fuel. They should turn to American (US) cars.

	Origin	MPG Category	count
0	Europe	Low Consumtion	63
1	Europe	High Consumtion	6
2	Japan	Low Consumtion	76
3	Japan	High Consumtion	3
4	US	High Consumtion	130
5	US	Low Consumtion	113

According to analysis, this would be Ideal for people those who want to buy cars that consume less in long distance travel.

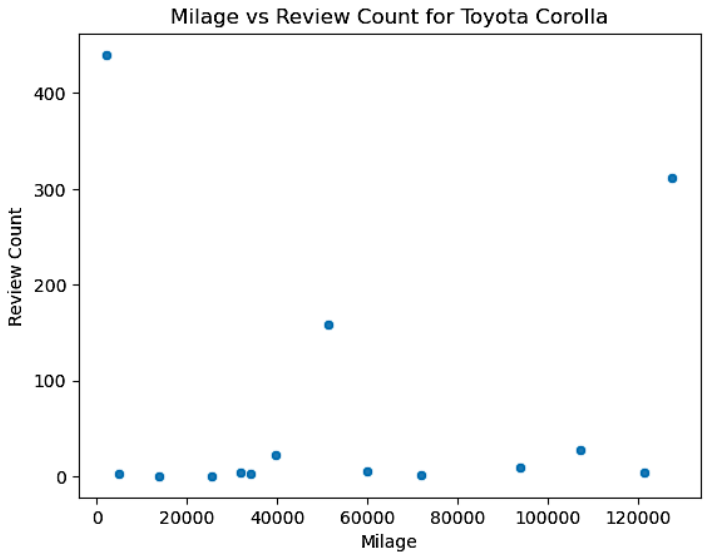
The BRITISH UNIVERSITY IN EGYPT

Assignment 2

1) Does Milage of Toyota Corolla have relationship with the Review Count ?

Answer: There no relationship between Milage and Review Count of Toyota Corolla. Because there is a discrepancy and it is not that the higher the Milage, the greater the Review Count.

Visualization:



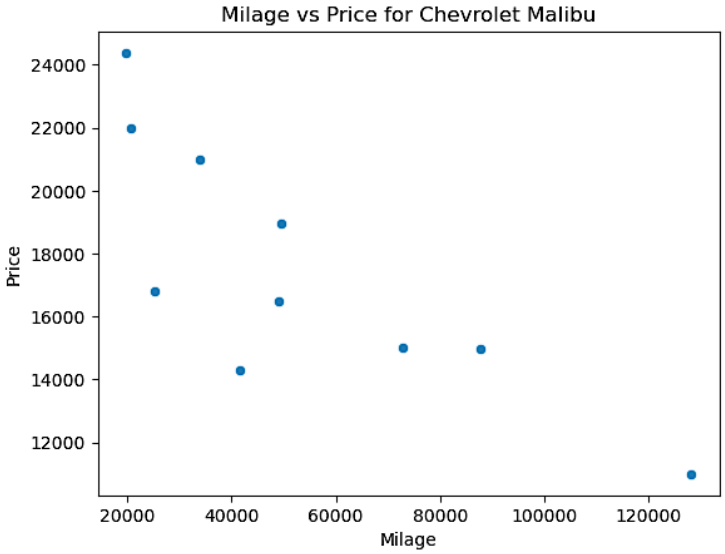
Notice:

There is Toyota Corolla has a low Milage and it has the highest Review Count.
Milage is not indicator of the Review Count there something else that affects the number of reviews.

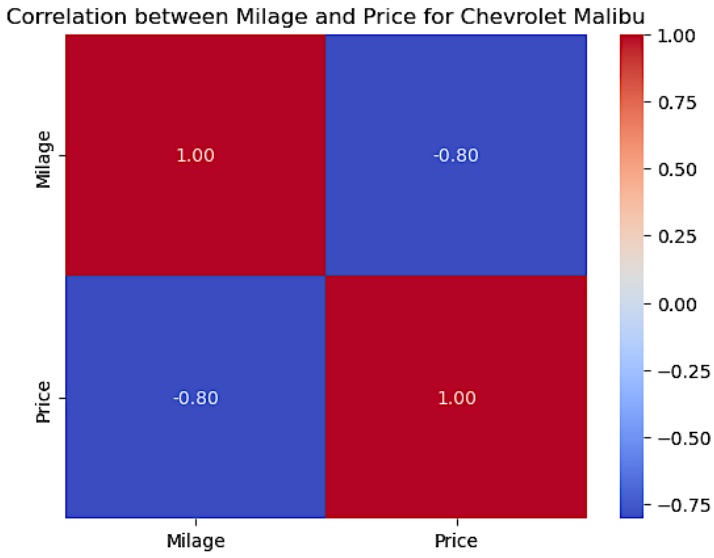
2) Is there any correlation/relationship between Milage of Chevrolet Malibu and the Price ?

Answer: There is relationship between Milage and Price of Chevrolet Malibu. Because when the Milage increase, the price tends to decrease.

Visualization:



Notice:
Milage is the influence on the Price. Car with high milage will have a lower price.

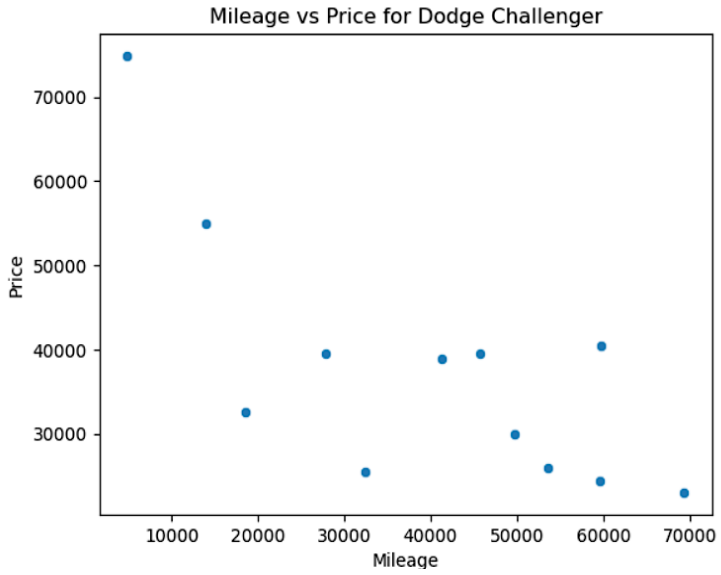


Notice:
Negative correlation between the Milage of Chevrolet Malibu and its Price

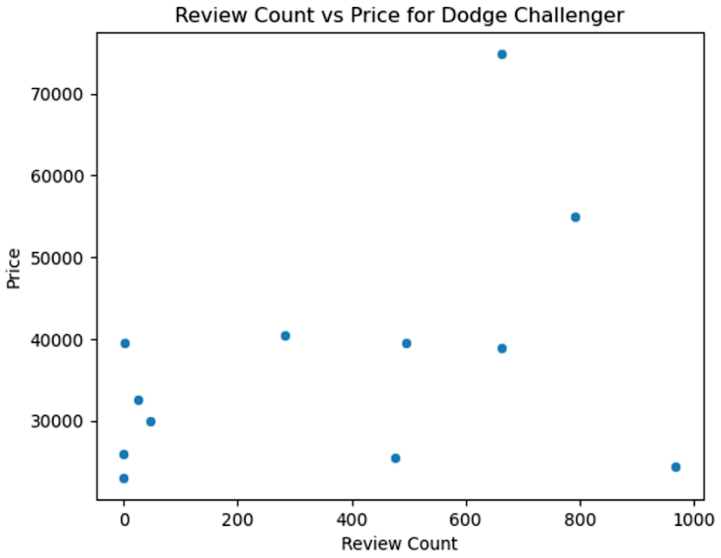
3) Dose the price of Dodge Challenger vary with their mileage and review count ?

Answer: There is a change between Milage and review count of Dodge Challenger. Because when the Milage decrease, the price tends to increase, *and when also when review count* decrease, the price also tends to decrease.

Visualization:



Notice:
Milage decrease, the price tends to increase

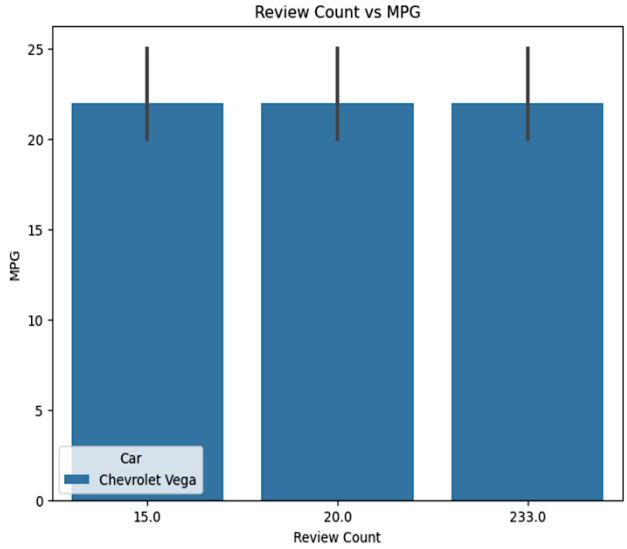


Notice:
review count decrease, the price also tends to decrease.

4) There any relationship between MPG and Car Review Count for Chevrolet Vega ?

Answer: In the range of 0 to 15 review count, MPG tends to be high and also in the range of 15 to 20 review count, MPG also tends to increase. We conclude from this that the more mpg, the more review count.

Visualization:

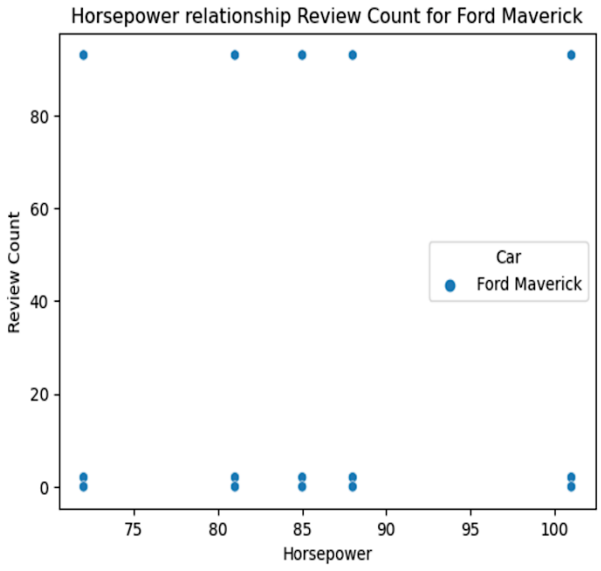


5) Is there any relation between horsepower and review count for Ford Maverick ?

Answer: The reviews of the Ford Maverick tend to be high or low with no direct relation with Horsepower. For instance, Mavericks with low Review count some of them have low horsepower (less than 75) and others have high horsepower above 80 and same for Mavericks with high reviews (more than 100).

Visualization:

Notice:
There is no direct relation between Ford Maverick Horsepower and Review Count.

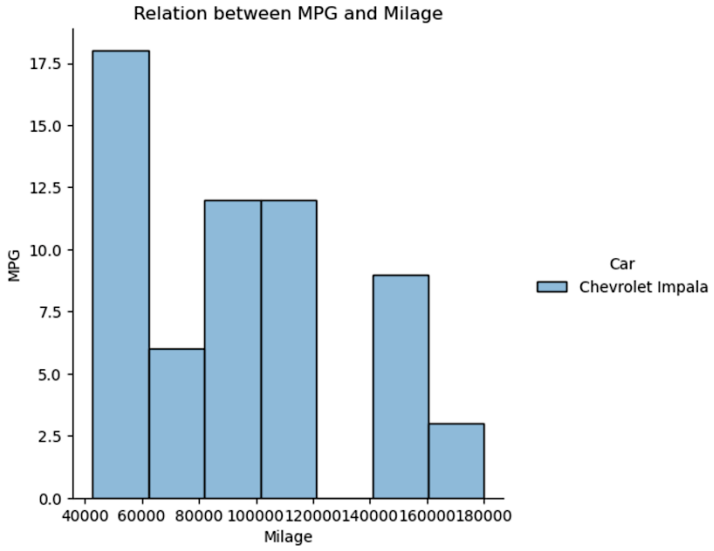


6) Does the Chevrolet Impala Milage have an impact on it MPG ?

Answer:

Chevrolet Impala Milage is inversely proportional with its MPG. For instance, Impala with 17.5 MPG has Millage range between (40000 to 60000). One the other hand Impala with MPG about 2.5 has millage between (160000 to 180000).

Visualization:



Notice:
Millage is inversely proportional with MPG for Chevrolet Impala.

From gathering data to producing valuable analysis:

First, what we did in this data set is that we collected data, then we did data processing, then we did exploratory analysis, then we created visualization, then we did hypothesis testing, and finally we extracted useful and valuable insights for those interested in cars.
The problem that we encountered difficulty with, but we solved it, was to clean the dataset and prepare it first, because it was all in one column. What we did is that we divided this column into more than one column, and we made a drop for the one column that consisted of all the data, and we made a rename for all the columns, and there was a row. It contains the names of the data types for each column (*and they were incorrect, they were just names with no real meaning*). They were all columns (*object*). There were outliers that we dealt with and filled the missing values and created visualizations in order to understand and discover the data better.
We made comparisons between the features or columns of different cars and their trends over time, as well as their distribution in terms of origins and models. Moreover, we obtained outcomes for such aspects as the car origins and model. To add on to that, we also focused on diverse visualization techniques which can be expressed in the form of scatter plots, box plots, histograms and bar plots in order to make our information easily understandable to readers or people from the audience. Because of it (*visualizations*), we extracted very useful insights.

Conclusions drawn from the answers to questions:

Almost all American cars have large engines, which makes them among the cars that can travel long distances
There is a relationship between horsepower and MPG. Cars with high horsepower are more fuel efficient.
In general, Ford is better than Chevrolet, but through analysis, Chevrolet outperforms Ford in many areas (every car manufacturer has a point in which it shines).
Each country has its own approach to focus on.
From the 70s to the 80s, car manufacturing developed significantly, as cars began to consume more fuel and have less displacement compared to old cars, such as American cars. American cars are considered the most fuel efficient (they travel long distances). European car manufacturers maintained this in the 70s and 80s at a much smaller displacement.
Car manufacturers in the same origin have different acceleration speeds for cars of the same origin and model, and this may be due to different manufacturing techniques.
There is a discrepancy between the number of kilometers and the number of inspections for a Toyota Corolla and The higher the mileage of the Toyota Corolla, the lower its price.
The mileage of the Dodge Challenger decreases, the price of the car increases, and as the review count decreases, its price increases.
There is a relationship between the MPG and Review Count of the Chevrolet Vega.
There is no direct relationship between horsepower and the review count of the Ford Maverick.
The Chevrolet Impala's Milage is inversely proportional with MPG.

Table of contributions:

	Contributions
MOHAMED KHAIRI	Participated in all the questions (in the assignment 1), and in the assignment 2 he was responsible for everything related to the Toyota Corolla and Chevrolet Malibu.
ABDELRAHMAN GAMAL	Participated in all the questions (in the assignment 1), and in the assignment 2 he was responsible for everything related to the Dodge Challenger and Chevrolet Vega.
KAREEM YASSER	He participated in all the questions (in the assignment 1), and in the assignment 2 he was responsible for everything related to the Ford Maverick and Chevrolet Impala.