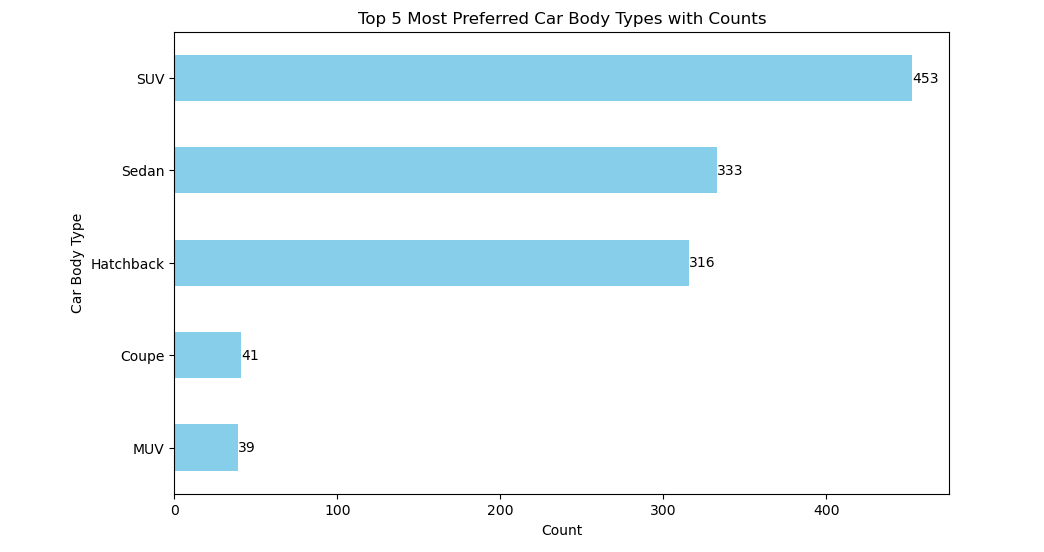
**CAR – PROJECT**

* **PART – 1:**

1. **Market Segmentation Analysis:**

## Task 1 -

**A new car manufacturer wants to introduce a new car and wants information about the top 5 most preferred car body types. Display the result using a horizontal bar chart.**

****

***Interpretation:***

This analysis is to identify the top 5 most preferred car body-types with counts.

* The body-types - **SUV** have the highest count i.e., 453.
* The body-types - **MUV** have the least count i.e., 39.

## Task 2 –

## List all the details of cars that can adapt to various driving conditions such as normal, comfort, eco, sport, and power mode. How many cars have all the above-mentioned features?

## 

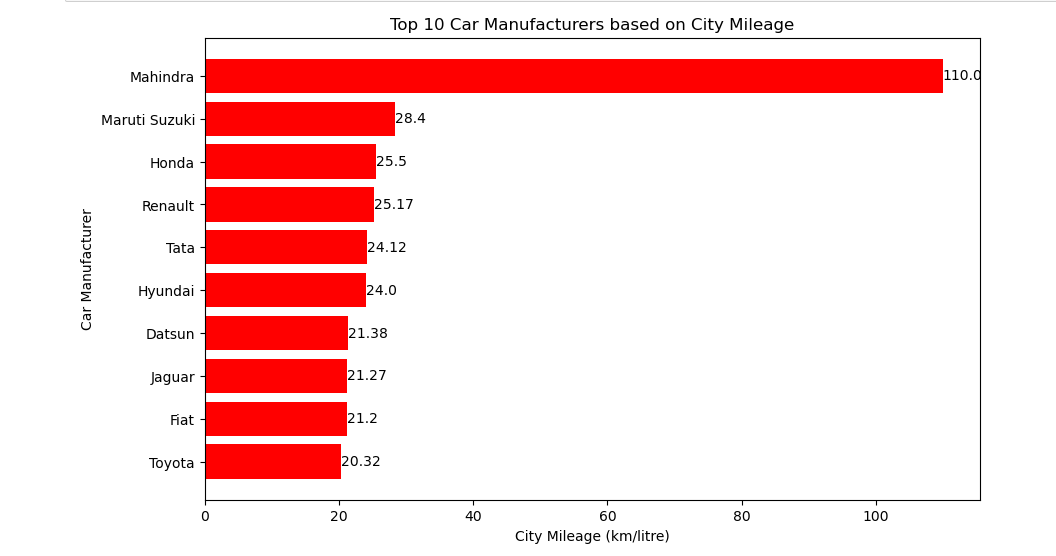
## Interpretation:

Displayed the list of all cars that can adapt to various driving conditions such as **Normal, Comfort, Eco, Sport and Power Mode** and found there were 1 car - **Audi** of 2 models- **Q8 & A8 L**with these specifications.

## Fuel Efficiency Analysis:

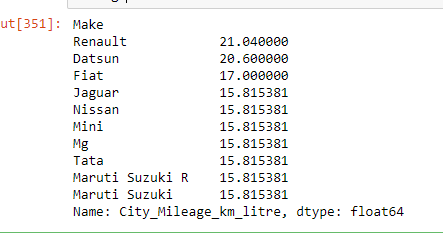
## Task 3 -

**Identify the top 10 car manufacturers based on the city mileage and display the result using a horizontal bar graph with the manufacturer on the y-axis and mileage on the x-axis. Do you notice any outliers in the above output? If yes, what is the reason, and how would you resolve it?**



## Observation:

The **Make - Mahinder** have the high city mileage and the **Make - Toyota** have the low city mileage, which is considered as high and low outliers in horizontal bar graph.

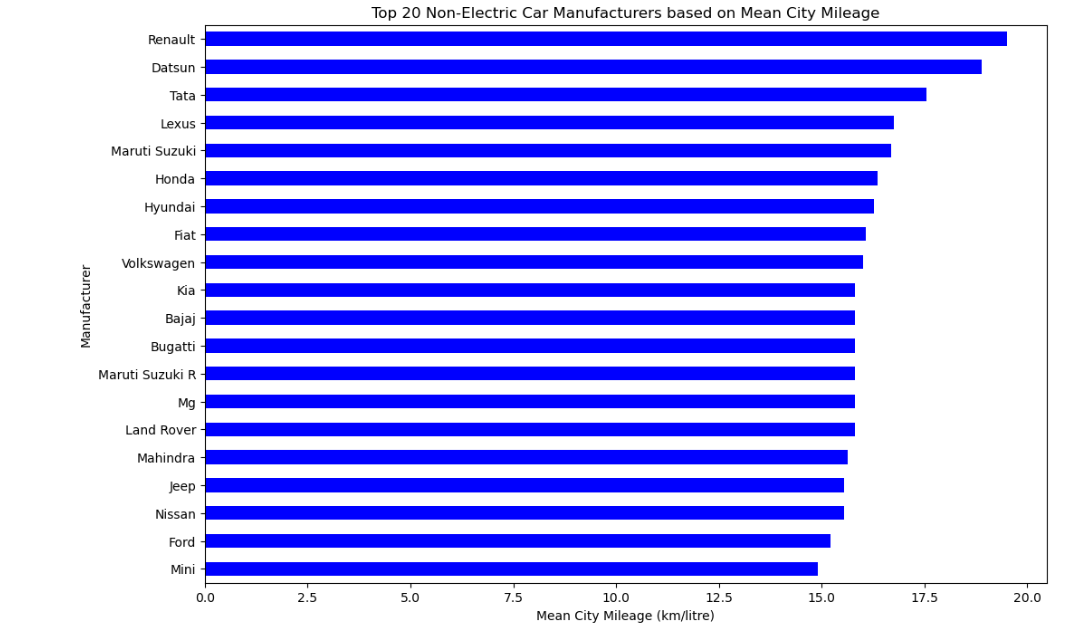
****

## Interpretation:[¶](http://localhost:8888/notebooks/Car_Project.ipynb#Interpretation:)

Performed Inter Quartile Range(IQR) to resolve the outliers that were found as high & low city mileage.

## Task 4 -

**Identify the top 20 non-electric car manufacturers based on city mileage and display the results using a bar graph similar to that in task 1.**

****

***Interpretation:***

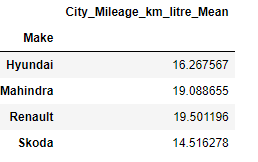
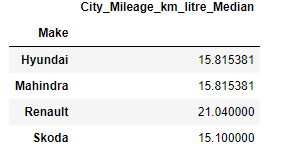
This analysis is to identifying top 20 car manufacturers based on mean city mileage.

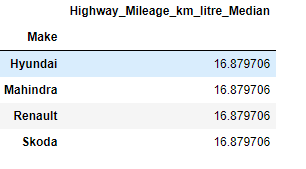
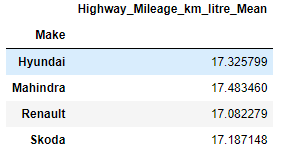
* The car manufacturer - **Renault** got the highest mean city mileage among all 20 car manufacturers.
* The car manufacturer - **Mini** got the least mean city mileage.
* **Task 5 -**

**Display the car mileages for the following four companies using a pivot table: Hyundai, Mahindra, Renault, and Skoda**

**Create 4 separate pivot tables capturing the following details:**

1. **Mean of city mileage**
2. **Median of city mileage**
3. **Mean of highway mileage**
4. **Median of highway mileage**

** **



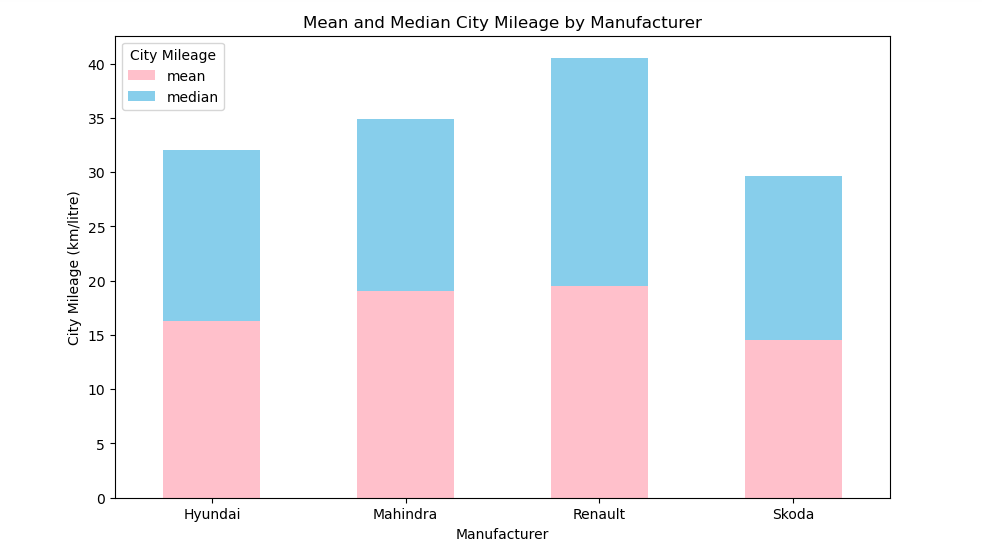
## Interpretation:

Creating pivot tables for 4 car manufacturers i.e., **Hyundai, Mahindra, Renault, Skoda**.

## Task 6 -

**Display the mean and median values of city mileages of the following manufacturers, using a stacked column chart:**

**Hyundai, Renault, Mahindra, and Skoda.**

****

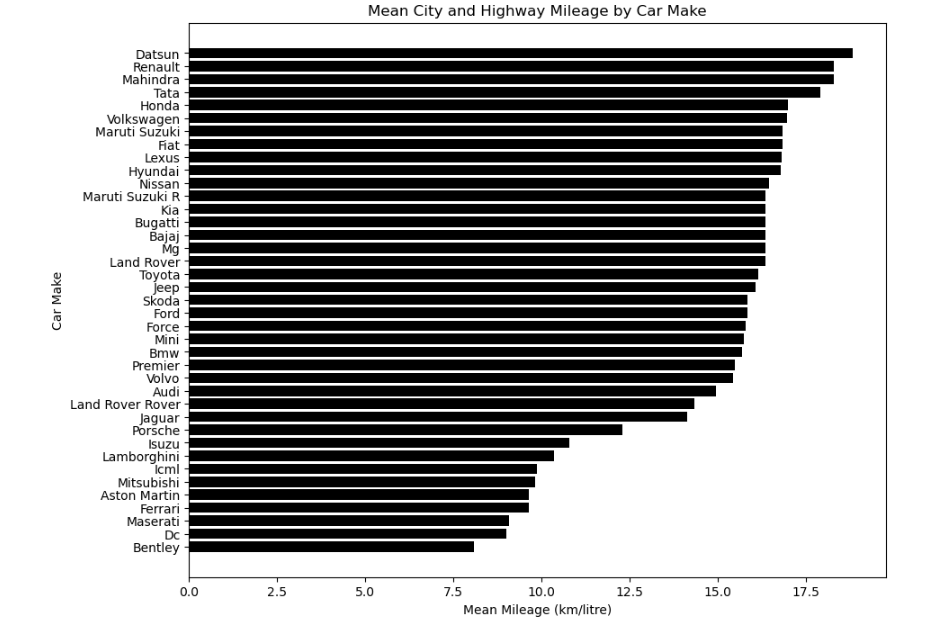
***Interpretation:***

This analysis provides a visual comparison of the mean and median city mileage for different manufacturers.

* The car manufacturer - **Renault** have the highest mean and median city mileage compared to other car manufacturers.
* The car manufacturer - **Skoda** have the least mean and median city mileage.

## Task 7 -

**Display the average mileage (city and highway) of all manufactures using a horizonal bar chart, in ascending order of mileage.**

****

***Interpretation:***

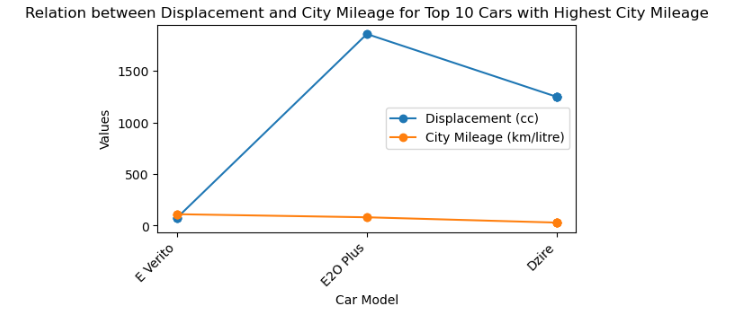
This analysis provides a clear comparison of the average mileage among different car makes.

* The car manufacturer - **Datsun** have the highest mean city mileage compared to all other car manufacturers.
* The car manufacturer - **Bentley** have the lowest mean city mileage.

## Task 8 -

**Plot a line chart to understand the relation between displacement and city mileage by choosing the top 10 cars with the highest city mileage.**

**What do you infer from this graph? Are there any unusual observations? If yes, how would you resolve these?**

****

## Interpretation:

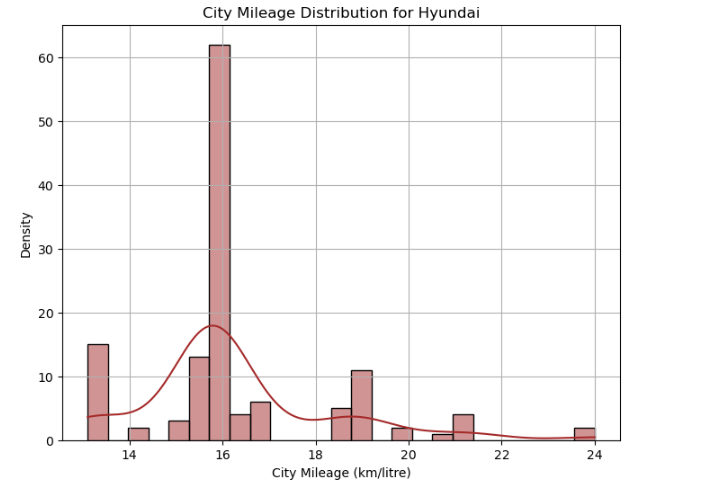
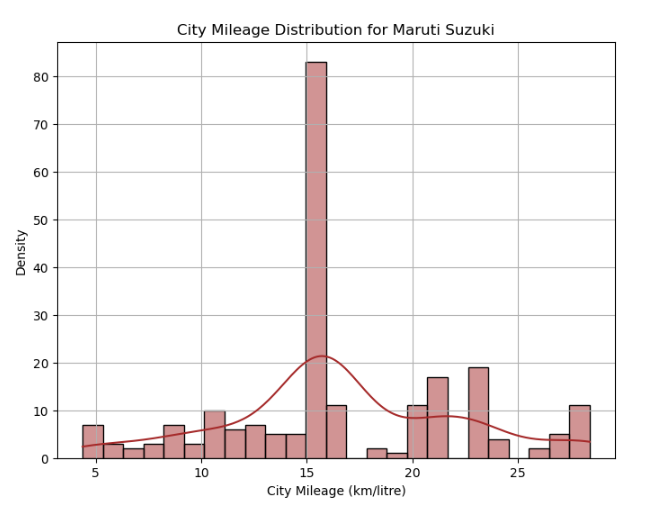
No unusual trend/observation is found, it displays the relation between displacement and city mileage based on highest mileage for top 10 car models.

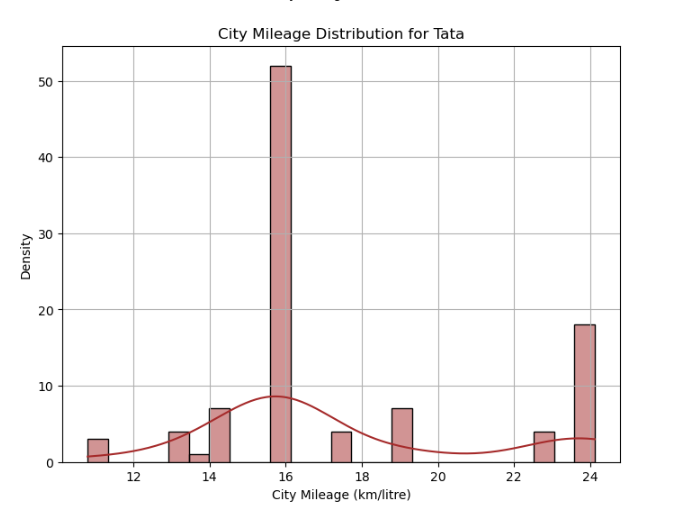
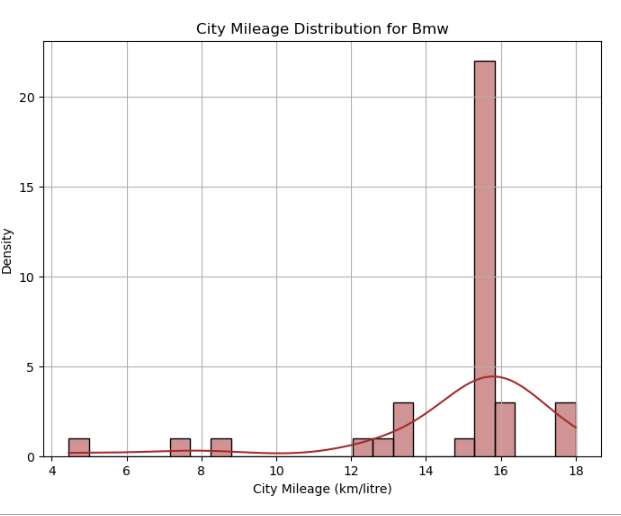
## Task 9 -

**Analyze the spread of fuel economy using a histogram for all car models of the following manufacturers:**

**Hyundai, Suzuki, Tata, and BMW**

**Which of these manufacturers have the most skewed mileage distribution?**

** **

** **

***Interpretation:***[**¶**](http://localhost:8888/notebooks/Car_Project.ipynb#Interpretation:)

The histograms allow comparison of city mileage distributions among different manufacturers **Hyundai, Maruti Suzuki, Tata, Bmw**.

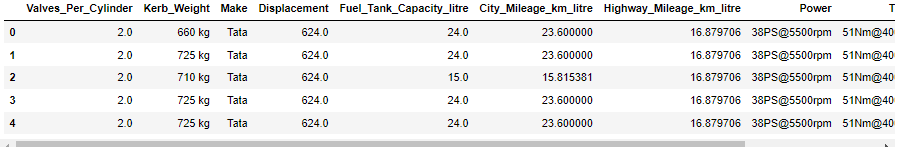
* The City Mileage distribution for car - **Hyundai** showing the **Strong Postive Right Skewness**.
* The City Mileage distribution for car - **Maruti Suzuki** showing the **Weak Postive Right Skewness**.
* The City Mileage distribution for car - **Tata** showing the **Strong Postive Right Skewness**.
* The City Mileage distribution for car - **Bmw** showing the **Strong Negative Left Skewness**.

# **PART – 2:**

* **Task 1 -**[**¶**](http://localhost:8888/notebooks/Car_Project.ipynb#Task-1--)

**Create another data frame that includes only the following columns from the data. Perform missing value treatment for these columns, if required.**

* **Valves\_Per\_Cylinder**
* **Kerb\_Weight**
* **Make**
* **Displacement**
* **Fuel\_Tank\_Capacity\_litre**
* **City\_Mileage\_km\_litre**
* **Highway\_Mileage\_km\_litre**
* **Power**
* **Torque**
* **Gross\_Vehicle\_Weight< font/>**

****

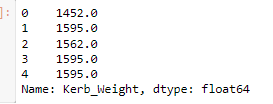
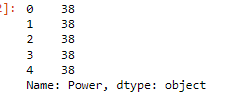
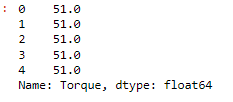
## Interpretation:

Created a new dataframe based on the columns - **Valves\_Per\_Cylinder, Kerb\_Weight, Make, Displacement, Fuel\_Tank\_Capacity\_litre,City\_Mileage\_km\_litre, Highway\_Mileage\_km\_litre, Power, Torque, Gross\_Vehicle\_Weight.**

* **Task 2 -**

**Preprocess the data in the following columns as specified.**

* **Kerb\_weight: Remove "kg" and ensure the column is numerical. Convert the weight in kilograms to pounds, by multiplying each value in the column with the number 2.2.**
* **Power: Sample data = 38PS @ 5500rpm. All entries of the column "Power" should only contain the number 38 after cleaning.**
* **Torque: Sample data = 51Nm@4000rpm. The column "Torque" must only contain the number 51 after cleaning. Perform this operation to all entries of the column and ensure this column is converted to numerical type.**

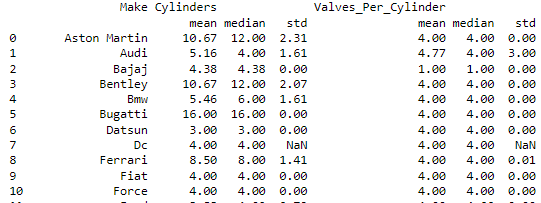
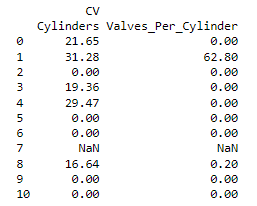
**  **

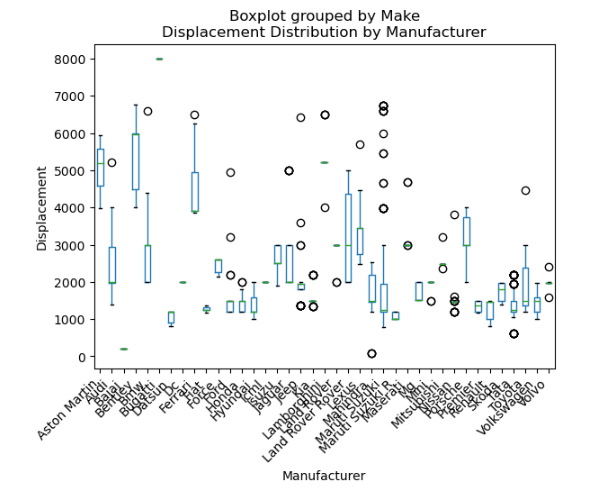
## Interpretation:

This analysis allows to understand the data preprocessing which aims to clean and standardize the format of three columns: **Kerb\_weight, Power, and Torque**.

* In the column - **Kerb\_Weight**, tried to convert the kilograms weight to pounds by multiplying each value by 2.2.
* In the column - **Power**, replaced all the occurrences of '38PS @ 5500rpm' with '38' regardless of case or surrounding spaces.
* In the column - **Torque**, contains values in a specific format, such as "51Nm@4000rpm" and by extracting it converted it to 51.

## Performance Comparison:

** **



## Interpretation:

* The above code analyzes the data by calculating mean, median, standard deviation and coefficient of variation for the number of **cylinders & valves per cylinder** across different car manufacturers.
* Additionally, there is a box-plot which represents the distribution of engine **displacements** among manufacturers.
* In the box-plot, the presence of multiple circles at certain points indicates the **highest density or frequency of data points** at those values.

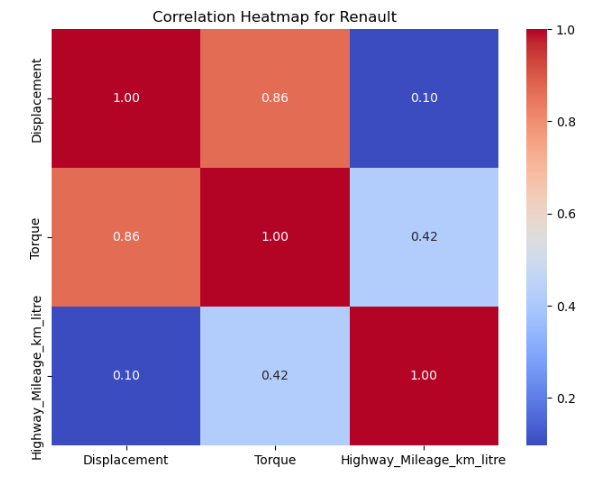
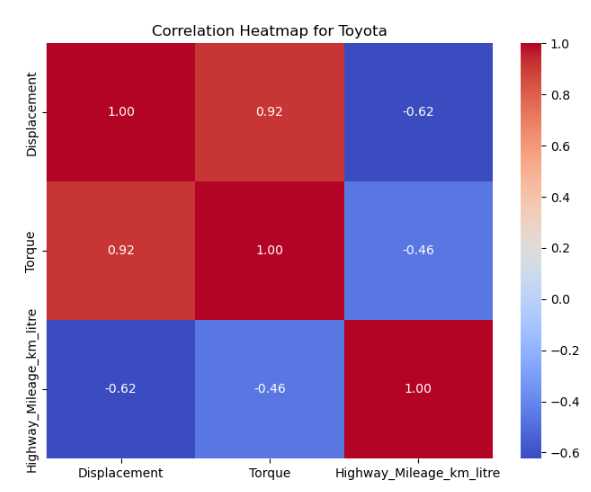
## Task 4 -

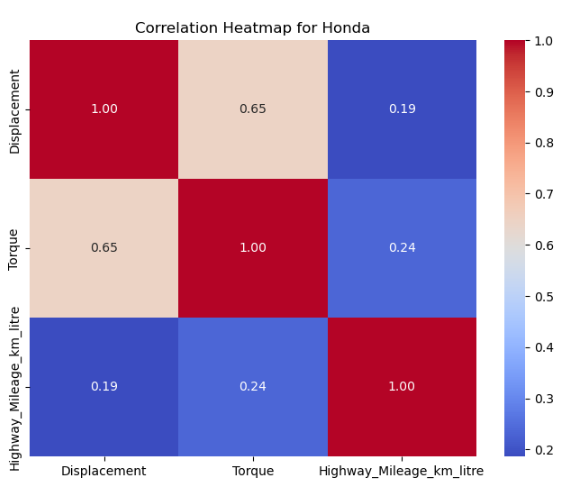
Identify at least three features that affect the mileages (using correlation analysis) of cars from the following manufactures:

Renault, Toyota, and Honda

Represent them visually using appropriate types of charts/graphs in Python.

Illustrate with justification whether these factors affect the mileage of a car positively or negatively.

****

***Interpretation:***

1. The heatmap visualizes the correlation matrix for **Renault(manufacturer)**.

* The relationships between engine displacement, torque, and highway mileage per liter, where it represents the **Strongly Positevely Correlated**.

1. The heatmap visualizes the correlation matrix for **Toyota(manufacturer)**.

* The relationships between engine displacement, torque, and highway mileage per liter, where it represents the **Strongly Negatively Correlated**.

1. The heatmap visualizes the correlation matrix for **Honda(manufacturer)**.

* The relationships between engine displacement, torque, and highway mileage per liter, where it represents the **Strongly Positively Correlated**.

## Weight Distribution Study:

## Task 5 -

**Compare the performance of cars based on factors like kerb weight and gross vehicle weight.**

**Hint: Calculate and display the mean, median, standard deviation, and coefficient of variation for each manufacturer.**

## 

## Interpretation:

* This analysis helps to understand the weight distribution of cars among different manufacturers, for which calculated the average(mean), meadian and variability(standard deviation & coefficient of variation) of **Kerb\_Weight and Gross\_Vehicle\_Weight**.
* The car manufacturer - **Audi** had the highest weight compared to all other manufacturers related to vehicle weight.

# **PART – 3:**

## Safety Feature Assessment:

## Task 1 -

**Examine the presence of safety features such as ABS, airbags, and hill assist, to assess the safety standards of various car models.**

## 

## Interpretation:

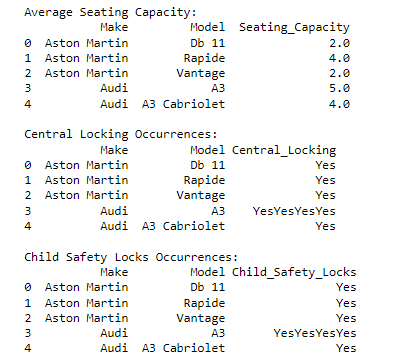
* This analysis defines the safety features based on **ABS\_(Anti-lock\_Braking\_System), Airbags, and Hill\_Assist**, in order to calculate the statistics related to the presence of safety features,converted values to binary i.e.,'Yes' and 'No' values to binary values (1 for 'Yes' and 0 for 'No').
* Calculated the statistics for safety features which include count and presence for each manufacturer.
* The car manufacturer - **Maruti Suzuki** has the presence highest count of ABS\_(Anti-lock\_Braking\_System), Airbags, and Hill\_Assist compared to other car manufacturers.

## User Comfort Exploration:

## Task 2 -

Explore seating capacity, central locking, and child safety locks to understand user comfort and family-friendly features.

Hint: Calculate the average seating capacity and count the occurrences of central locking and child safety locks for each model.



***Interpretation:***

* The average seating capacity for car - **Audi-A3** offers insights into the typical passenger accommodation compared to various car models
* The counts of central locking and child safety locks occurrences for car - **Audi-A3** reveal the safety features are highly integrated compared to other car models.

## Alert Systems Analysis:

## Task 3 -

**Investigate the presence of high-speed alerts, seat belt reminders, and door ajar warnings to assess the integration of safety and convenience features.**

## 

## Interpretation:

This analysis helps understand the distribution of safety and convenience features among different car models.

* The safety feature for car with **High\_Speed\_Alert\_System** has the least percentage for the car model - **718**.
* The safety feature for car with **Fasten\_Seat\_Belt\_Warning** has the second highest percentage for car model - **911**.
* The safety feature for car with **Door\_Ajar\_Warning** has the highest percentage for car model - **3 -Series**.

## Dimensional Analysis:

## Task 4 -

**Analyze the dimensions of cars (height, length, width) to understand size preferences and market demands, aiding in product planning.**

## 

## Interpretation:

The frequency and descriptive statistics help in understanding the distributions and central tendencies of car dimensions, providing insights into the variability and spread of height, length, and width among the cars.

* The **Height** of car displayes the **Right Psoitive Skewness**.
* The **Length** of car displayes the **Left Negative Skewness**.
* The **Width** of car displayes the **Left Negative Skewness**.

## PART – 5:

## ****Summary****

### **Part 1**

### **1. Market Segmentation Analysis**

1. This analysis is to identify the top 5 most prefered car body-types with counts.

* The body-types - SUV have the highest count i.e., 453.
* The body-types - MUV have the least count i.e., 39.

1. Displayed the list of all cars that can adapt to various driving conditions such as Normal, Comfort, Eco, Sport and Power Mode and found there were 1 car - Audi of 2 models- Q8 & A8 Lwith these specifications.

### **2. Fuel Efficiency Analysis**

1. The Make - Mahinder have the high city mileage and the Make - Toyota have the low city mileage, which is considered as high and low outliers in horizontal bar graph.Performed Inter Quartile Range(IQR) to resolve the outliers that were found as high & low city mileage.
2. This analysis is to identifying top 20 car manufacturers based on mean city mileage.

* The car manufacturer - Renault got the highest mean city mileage among all 20 car manfacturers.
* The car manufacturer - Mini got the least mean city mileage.

1. Creating pivot tables for 4 car manufacturers i.e., Hyundai, Mahindra, Renault, Skoda .
2. This analysis provides a visual comparison of the mean and median city mileage for different manufacturers.

* The car manufacturer - Renault have the highest mean and median city mileage compared to other car manufacturers.
* The car manufacturer - Skoda have the least mean and median city mileage.

1. This analysis provides a clear comparison of the average mileage among different car makes.

* The car manufacturer - Datsun have the highest mean city mileage compared to all other car manufacturers.
* The car manufacturer - Bentley have the lowest mean city mileage.

1. No unusal trend/observation is found,it displayes the relation between displacement and city mileage based on highest mileage for top 10 car models.
2. The histograms allow comparison of city mileage distributions among different manufacturers Hyundai, Maruti Suzuki, Tata, Bmw.

* The City Mileage distribution for car - Hyundai showing the Strong Postive Right Skewness.
* The City Mileage distribution for car - Maruti Suzuki showing the Weak Postive Right Skewness.
* The City Mileage distribution for car - Tata showing the Strong Postive Right Skewness.
* The City Mileage distribution for car - Bmw showing the Strong Negative Left Skewness.

### **Part 2**

1. Created a new dataframe based on the columns - Valves\_Per\_Cylinder, Kerb\_Weight, Make, Displacement, Fuel\_Tank\_Capacity\_litre,City\_Mileage\_km\_litre, Highway\_Mileage\_km\_litre, Power, Torque, Gross\_Vehicle\_Weight.
2. This analysis allows to understand the data preprocessing which aims to clean and standardize the format of three columns: Kerb\_weight, Power, and Torque.

* In the column - Kerb\_Weight, tried to convert the kilograms weight to pounds by multiplying each value by 2.2.
* In the column - Power, replaced all the occurrences of '38PS @ 5500rpm' with '38' regardless of case or surrounding spaces.
* In the column - Torque, contains values in a specific format, such as "51Nm@4000rpm" and by extracting it converted it to 51.

### **1. Performance Comparison**

1. The above code analyzes the data by calculating mean, median, standard deviation and coefficient of variation for the number of cylinders & valves per cylinder across different car manufacturers.

* Additionally, there is a box-plot which represents the distribution of engine displacements among manufacturers.
* In the box-plot, the presence of mutliple circles at certain points indicates the highest density or frequency of data points at those values.

1. The heatmap visualizes the correlation matrix for Renault(manufacturer).

* The relationships between engine displacement, torque, and highway mileage per liter, where it represents the Strongly Positevely Correlated.
* The heatmap visualizes the correlation matrix for Toyota(manufacturer).
* The relationships between engine displacement, torque, and highway mileage per liter, where it represents the Strongly Negatively Correlated.
* The heatmap visualizes the correlation matrix for Honda(manufacturer).
* The relationships between engine displacement, torque, and highway mileage per liter, where it represents the Strongly Positively Correlated.

### **2. Weight Distribution Study**

1. This analysis helps to understand the weight distribution of cars among different manufacturers, for which calculated the average(mean), meadian and variability(standard deviation & coefficient of variation) of Kerb\_Weight and Gross\_Vehicle\_Weight.

* The car manufacturer - Audi had the highest weight compared to all other manufacturers related to vehicle weight.

### **Part 3**

### **1. Safety Feature Assessment**

1. This analysis defines the safety features based on ABS\_(Anti-lock\_Braking\_System), Airbags, and Hill\_Assist, in order to calculate the statistics related to the presence of safety features,converted values to binary i.e.,'Yes' and 'No' values to binary values (1 for 'Yes' and 0 for 'No').

* Calculated the statistics for safety features which include count and presence for each manufacturer.
* The car manufacturer - Maruti Suzuki has the presence highest count of ABS\_(Anti-lock\_Braking\_System), Airbags, and Hill\_Assist compared to other car manufacturers.

### **2. User Comfort Exploration**

1. The average seating capacity for car - Audi-A3 offers insights into the typical passenger accommodation compared to various car models

* The counts of central locking and child safety locks occurrences for car - Audi-A3 reveal the safety features are highly integrated compared to other car models.

### **3. Alert Systems Analysis**

1. This analysis helps understand the distribution of safety and convenience features among different car models.

* The safety feature for car with High\_Speed\_Alert\_System has the least percentage for the car model - 718.
* The safety feature for car with Fasten\_Seat\_Belt\_Warning has the second highest percentage for car model - 911.
* The safety feature for car with Door\_Ajar\_Warning has the highest percentage for car model - 3 -Series.

### **4. Dimensional Analysis**

1. The frequency and descriptive statistics help in understanding the distributions and central tendencies of car dimensions, providing insights into the variability and spread of height, length, and width among the cars.

* The Height of car displayes the Right Psoitive Skewness.
* The Length of car displayes the Left Negative Skewness.
* The Width of car displayes the Left Negative Skewness.

**Recommendations**

* SUVs are the most preferred car body-type, followed by others.
* Audi's Q8 and A8 L models offer versatility for various driving conditions.
* Mahindra cars demonstrate high city mileage, while Toyota cars show lower mileage.
* Consider using Inter Quartile Range (IQR) to address outliers for precise analysis of fuel efficiency.
* Analyze mean, median, standard deviation, and coefficient of variation for cylinders and valves per cylinder.
* Utilize box-plots to comprehend engine displacement distributions among manufacturers for performance comparison.
* Audi cars tend to have higher weights compared to other manufacturers.
* Explore mean, median, and variability in kerb weight and gross vehicle weight for weight distribution study.
* Maruti Suzuki prioritizes integrating ABS, Airbags, and Hill Assist features across its models for safety.
* Evaluate Audi-A3 for seating capacity and safety features integration, such as central locking and child safety locks, for user comfort.
* Assess the distribution of safety and convenience features across different car models for alert systems analysis.
* Understand the distribution and central tendencies of car dimensions, noting skewness in height, length, and width for - dimensional analysis.

**Conclusion**

* Market preferences lean heavily towards SUVs, indicating a strong consumer demand for this body-type.
* Audi's Q8 and A8 L models stand out for their adaptability to diverse driving conditions, offering versatility to potential buyers.
* Mahindra cars demonstrate superior city mileage performance compared to Toyota, highlighting potential areas for improvement in fuel efficiency.
* Utilizing Inter Quartile Range (IQR) can help address outliers and ensure more accurate analysis of fuel efficiency trends.
* Detailed performance comparison metrics, such as mean, median, and standard deviation of cylinders and valves per cylinder, provide insights into engine performance across manufacturers.
* Engine displacement distributions, analyzed through box-plots, offer further understanding of performance variations among different car brands.
* Audi cars tend to be heavier than those of other manufacturers, impacting factors like fuel efficiency and handling.
* Safety features are prominently integrated into Maruti Suzuki models, enhancing their appeal to safety-conscious consumers.
* Audi-A3's seating capacity and safety features integration make it a compelling option for those prioritizing comfort and safety.
* Assessing safety and convenience feature distributions across models helps in understanding market trends and consumer preferences.
* Dimensional analysis reveals insights into the variability and distribution of car dimensions, which can inform design and manufacturing decisions.

**THANK YOU**