Automobile Market Analysis Project

Part 1 and Part 2

Deepak Venkat Sai Ram D
Mentor : Mr. Munna Pandey

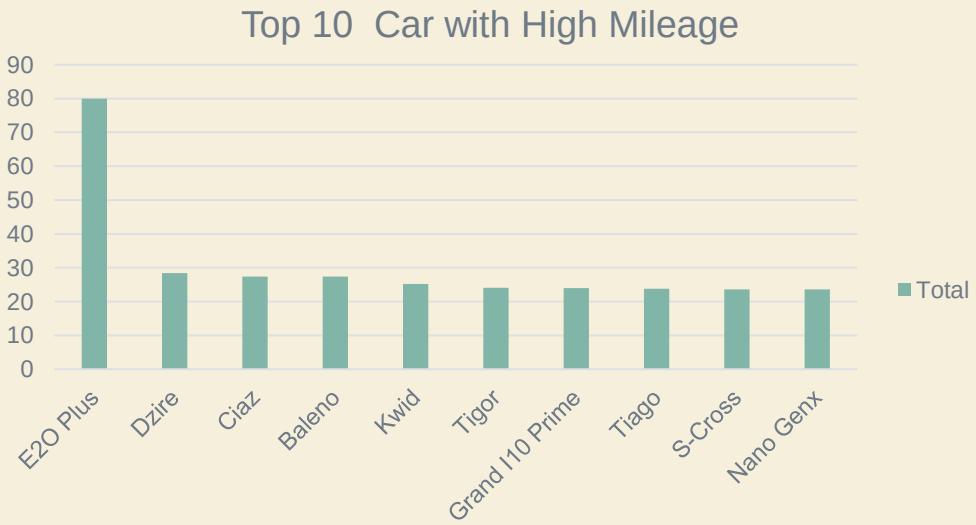
Introduction

1. In the current competitive Automobile market each and every car manufacturers are facing evolving market where the drivers are becoming eco-friendly, Fuel efficiency has become their top priority as well as safety features in cars alongside performance.
2. In order to analyze the various shifts, trends and efficiency in the automobile industry DATA ANALYTICS has been used at every stage of automobile production.
3. In this project, Smith a data analyst with a popular automobile giant need to analyze a car dataset and identify the potential models to launch and present insights about the most diverse car manufacturer in the given country.
4. A data set containing all the details of the automobile production of car manufacturers has been given. Various business questions is been solved to arrive at the conclusion.

Part 1

Question 1 : Identify the top 10 cars with the highest miles per gallon. Also, identify the names of the car manufacturers.

Model	Manufactures
E20 Plus	Mahindra
Dzire	Suzuki
Baleno	Suzuki
Ciaz	Suzuki
Kwid	Renault
Tigor	Tata
Grand i10 Prime	Hyundai
Tiago	Tata
S-Cross	Suzuki
Nano Genx	Tata



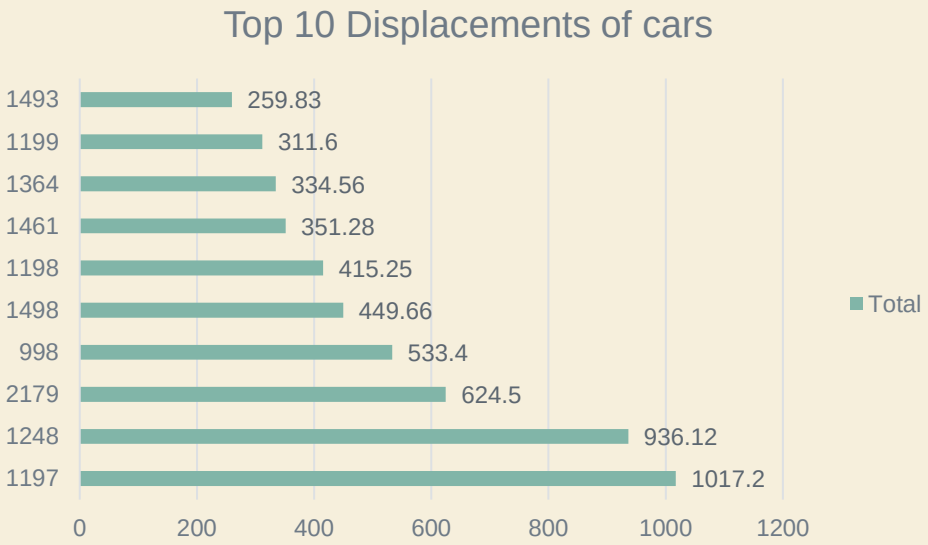
Potential Interpretations:

Fuel Efficiency: Brands like Mahindra, Suzuki, Renault, and Tata are showing high fuel efficiency, which could be appealing for consumers looking to minimize fuel costs, especially in urban settings where fuel consumption can be a significant factor.

Performance vs. Economy: Luxury or performance-oriented brands like BMW may have lower fuel efficiency compared to more economy-focused brands. This might reflect a trade-off between performance features and fuel economy.

Part 1

Question 2: Identify the top 10 cars with the highest miles per gallon based on displacement. Also, identify the names of their car manufacturers.



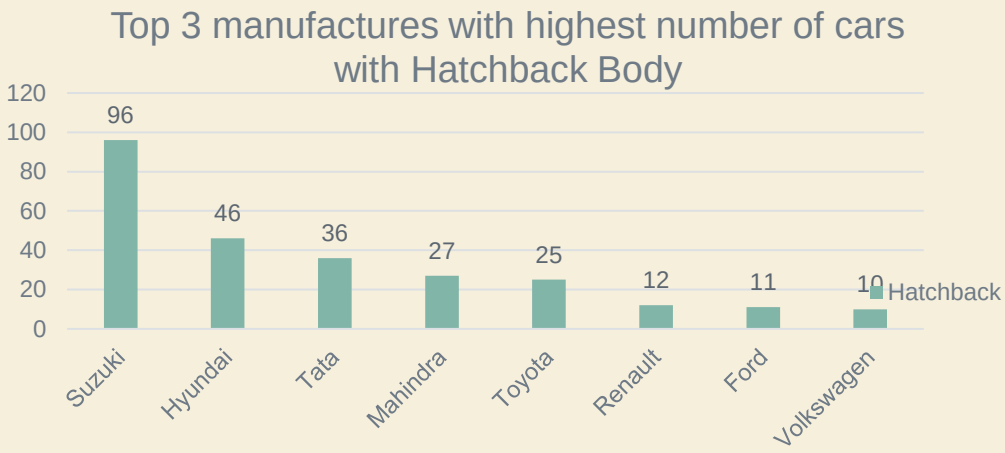
Displacement	Make	model
1197	Suzuki	Ignis
1248	Tata	Bolt
998	Suzuki	Alto K10
1498	Volkswagen	Polo
2179	Mahindra	Tuv300 Plus
1497	Tata	Altroz
1198	Mahindra	Kuv 100 Nxt
1199	Tata	Tiago
1461	Mahindra	Verito Vibe
999	Renault	Kwid

Potential Interpretations:

From above data we have identified the top 10 cars with highest miles per gallon based on displacement and also identified their car manufactures.

Part 1

Question 3 : Identify the top 3 manufacturers with the highest number of cars of the hatchback body type.



Interpretation:

1. Suzuki: 96 entries, indicating it has the largest share in the "Hatchback" segment.
2. Hyundai: 46 entries, suggesting a significant presence.
3. Tata: 36 entries, also a major player in this category.
4. Mahindra: 30 entries, indicating a notable but smaller presence.
5. Toyota: 25 entries, suggesting a moderate presence.
6. Renault: 12 entries, indicating a smaller share.
7. Ford: 11 entries, also a smaller player in the "Hatchback" market.
8. Volkswagen: 10 entries, representing the smallest share among the listed manufacturers.

Key Insights:

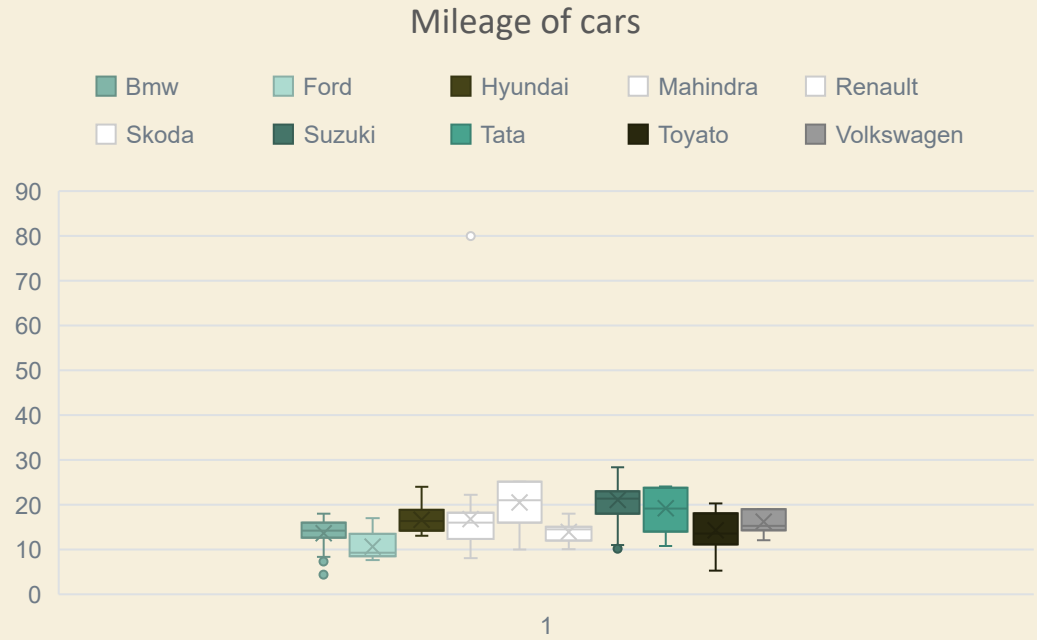
Suzuki dominates the "Hatchback" segment.

Hyundai and Tata also have significant shares, indicating strong competition in the market.

Mahindra, Toyota, Renault, Ford, and Volkswagen have smaller shares, suggesting they are either niche players or have less focus on the "Hatchback" body type compared to others.

Part 1

Question 4 : Compare the mileages of cars produced by all manufacturers and perform outlier analysis.



Interpretation

BMW - Outliers: 4.45, 7.32, 8.4, 28.09, 28.4
Ford - Outliers: None
Hyundai - Outliers: None
Mahindra - Outliers: 27.39, 28.09, 28.4
Renault - Outliers: None
Skoda - Outliers: None
Suzuki - Outliers: 5.3, 28.4
Tata - Outliers: 28.4
Toyota - Outliers: None
Volkswagen - Outliers: None

General Interpretation

1. **Consistency**: Columns like Ford, Hyundai, Renault, Skoda, Toyota, and Volkswagen show no outliers, indicating that these datasets are consistent with no extreme values affecting the overall data range.

2. **Variability**: The presence of outliers in columns such as BMW, Mahindra, Suzuki, and Tata suggests that these datasets have more variability. This could be due to unusual cases, measurement errors, or genuine extreme values.

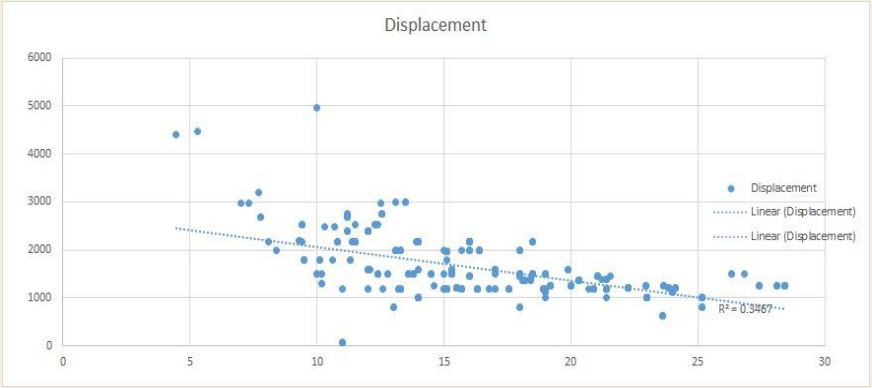
3. **Data Quality**: Outliers might indicate errors or exceptional cases. It's important to review these outliers to ensure they are not due to data entry mistakes and to understand their impact on the analysis.

4. **Further Analysis**: The presence of outliers may warrant further investigation to understand their causes and implications. For instance, are they due to rare but valid events, or are they errors in data collection?

Part 1

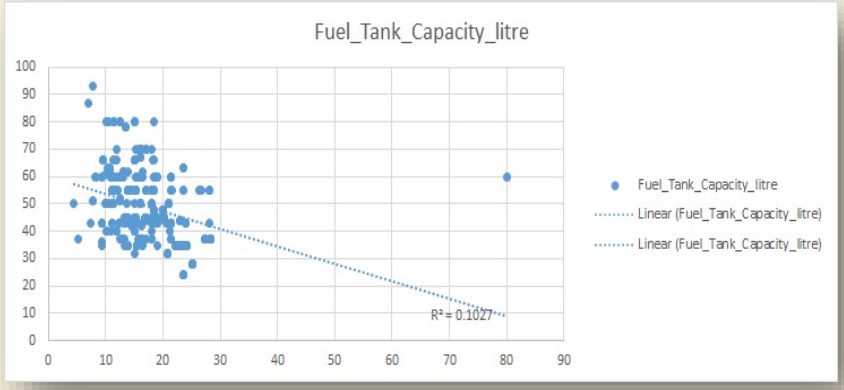
Question 5 : Present the relationship between:

A. Displacement and mileage



Interpretation: Correlation coefficient of displacement and mileage is equal to – 0.34

B. Mileage and fuel tank capacity

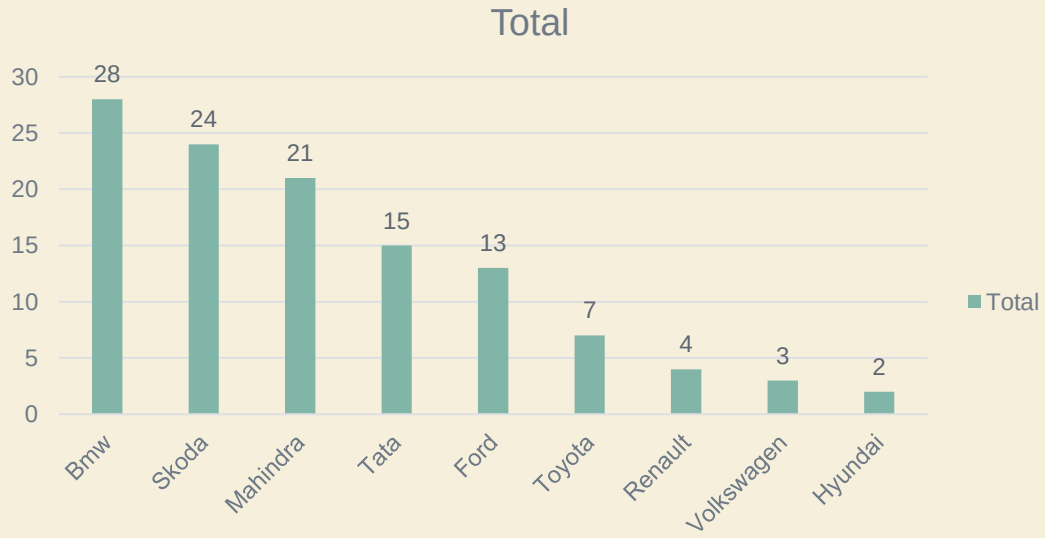


Interpretation: Correlation coefficient of Mileage and fuel tank capacity is –0.38

General Interpretation: Both represents a **negative** correlation

Part 1

Question 6 : Identify a car that would be best for adventurous drives.



BMW is suitable for **Adventures** drivers

Interpretation

BMW: 28 cars (23.93%)

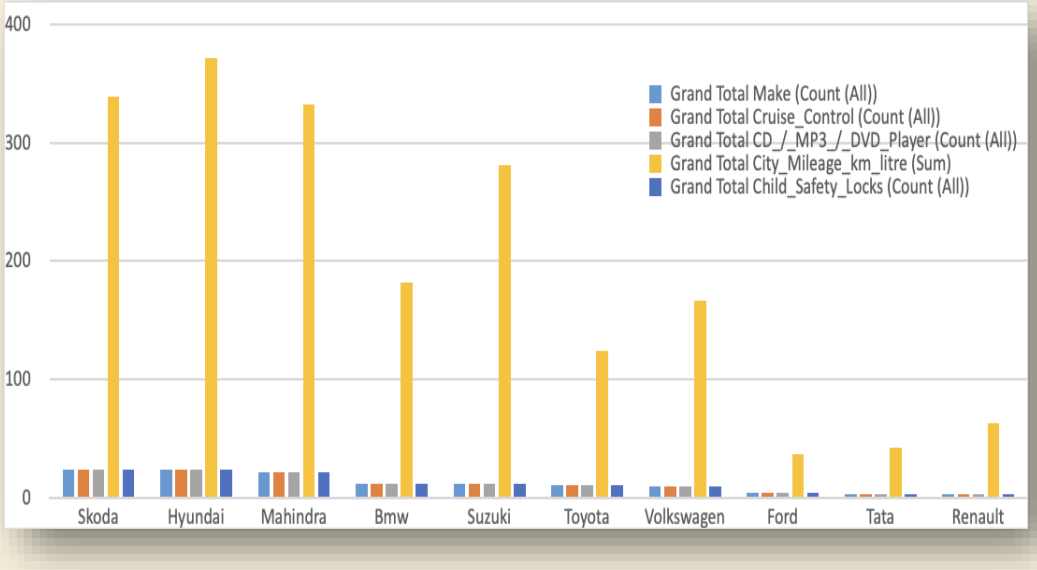
BMW X4 is the most frequently listed make, comprising nearly 24% of the total cars. This suggests a strong preference or higher availability for BMW vehicles in this dataset.

Top Makes: BMW, Skoda, and Mahindra are the most frequent, with BMW leading significantly.

Less Frequent Makes: Brands like Toyota, Renault, Volkswagen, and Hyundai have relatively smaller shares, indicating either less popularity or fewer listings.

Part 1

Question 7 : identify a car suitable for family usage.



Hyundai is the best car for **Family** usage

Recommendation

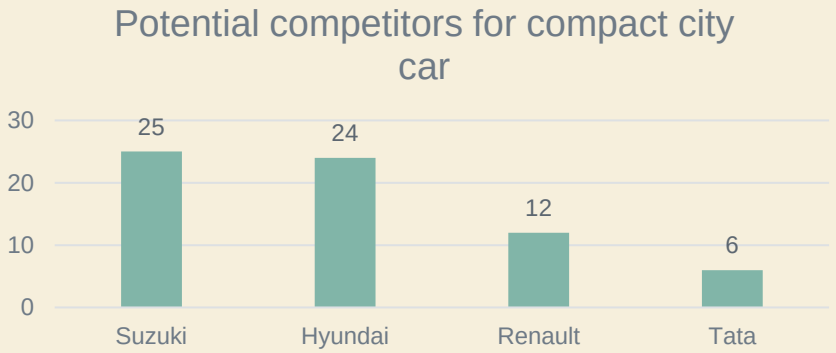
- Hyundai
- Skoda
- Mahindra
- Suzuki & Toyota
- BMW

For a family car, **Hyundai** and **Skoda** are likely to be good options based on fuel efficiency and model variety. **Mahindra** could be ideal if you are looking for larger SUVs or MPVs. Always consider test-driving and evaluating specific models to ensure they meet your family's needs in terms of space, comfort, and overall safety.

Part 1

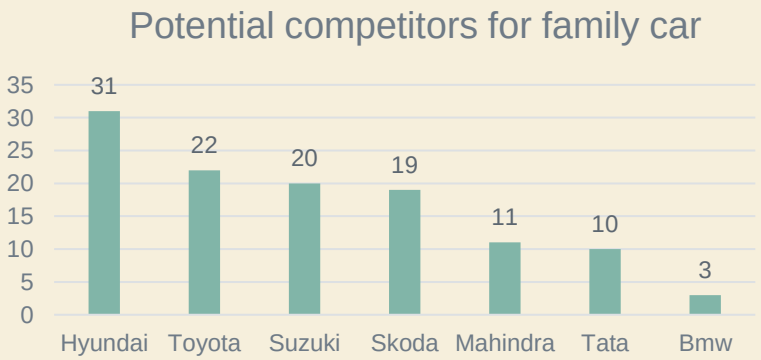
Question 8 : Identify the potential competitor cars for the two types of cars stated above.

- Compact city car



Suzuki and **Hyundai** are the primary competitors in the **compact** city car segment due to their larger number of models or entries.

- Family car

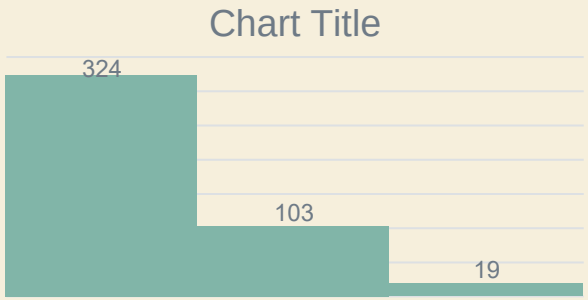


In the **family** car market, **Hyundai** and **Toyota** are the leading competitors due to their high vehicle counts and strong offerings in the family car segment.

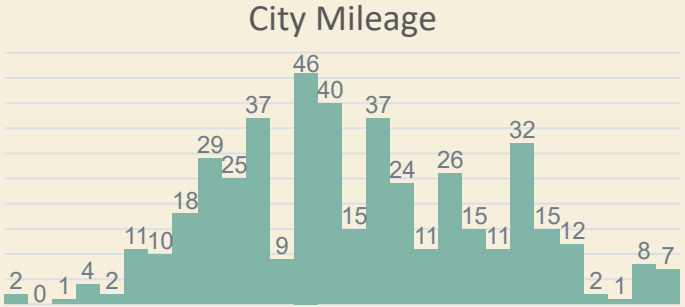
Part 1

Question 9 : Identify the outliers from following columns using Histogram method.

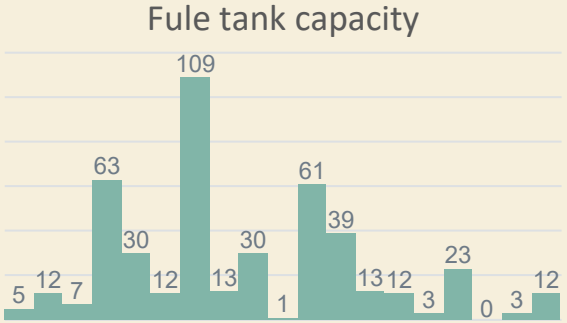
A. Displacement



B. City Mileage(miles/gallon)



C. Fuel Tank Capacity



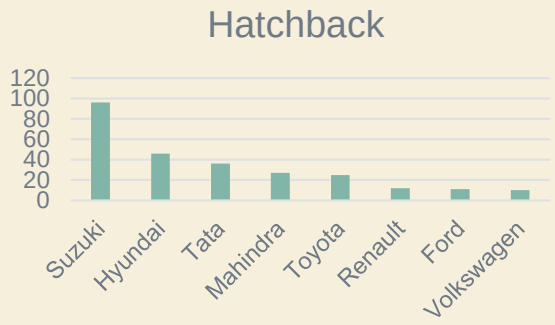
INTERPRETATION:

- Firstly, Outliers are found for City mileage, Displacement and Fuel tank capacity using Histogram and IQR method.
- Displacement: 64 outliers(UF), 3 outliers(LF)
- City mileage : 91 outliers (UF),78outliers (LF)
- Fuel tank capacity: 4 outliers (UF), 0 outliers (LF)
- Then we have removed the outliers and created histogram.

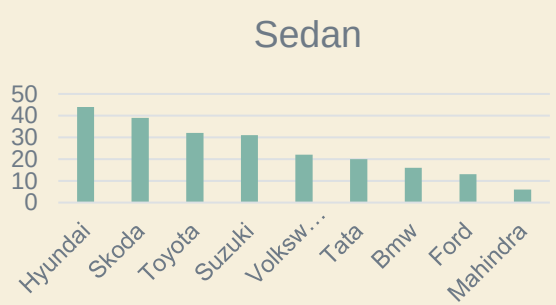
Part 1

Question 10 : Identify the top three car manufacturers with the highest number of variants in following categories

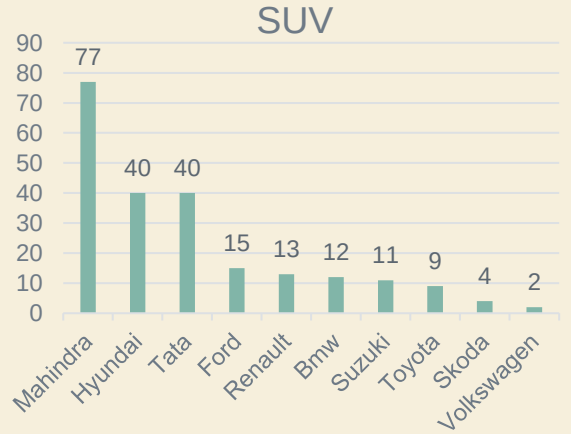
A. Hatchback



B. Sedan



C. SUV



INTERPRETATION:

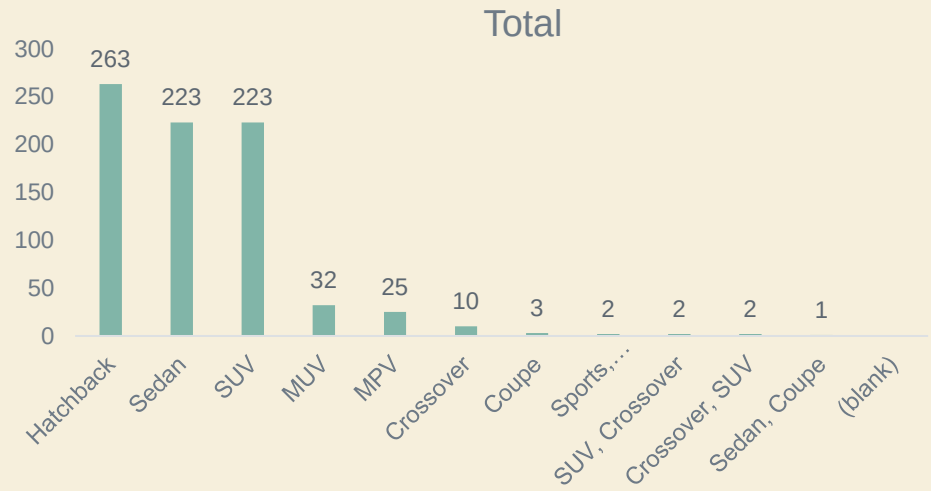
A. Hatchback **: Suzuki, Hyundai, Tata** are the top three manufactures with highest variants in Hatchback category.

B. Sedan : **Hyundai, Skoda, Toyota** are the top three manufactures with highest variants in sedan category.

C. SUV: **Mahindra, Hyundai, Tata** are the top three manufactures with highest variants in SUV category

Part 1

Question 11 : Find the most popular car body type, by count, from the analysis of Task 2.



Hatchback is the most popular body type followed by Sedan and SUV.

General Observations:

Hatchbacks are the most common body type in the dataset, with 263 vehicles, which is the highest count by a significant margin. This indicates a strong preference or availability for hatchbacks in the sample.

Sedans and **SUVs** are tied for the second most common body type, each with 223 vehicles. This suggests a balance between these two popular types of vehicles.

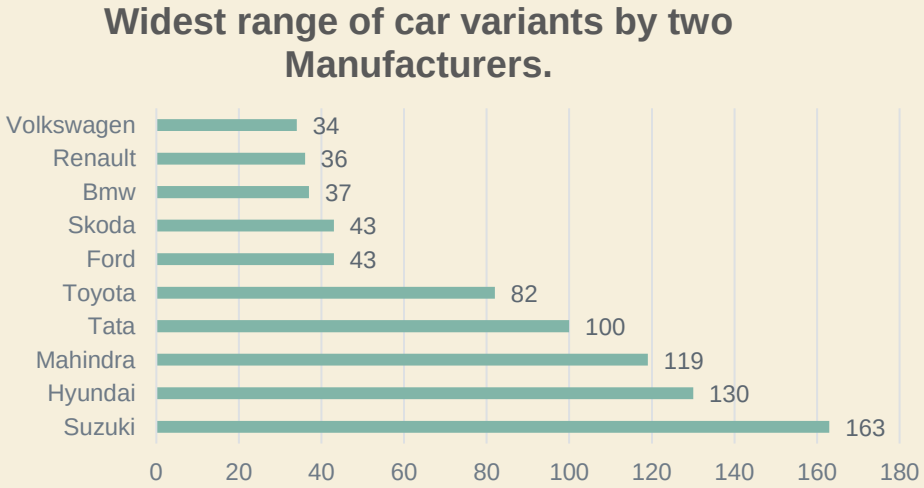
MUVs and **MPVs** are less common, with 32 and 25 vehicles, respectively. These types are not as prevalent in this dataset.

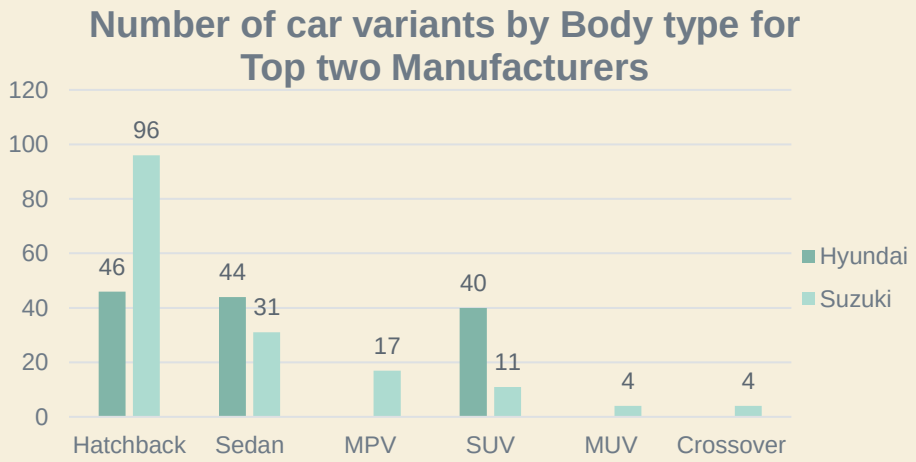
Crossovers have a relatively low count (10 vehicles), and **Coupes** are even less common (3 vehicles).

The combined counts for **Sports, Convertible**, **SUV, Crossover**, **Crossover, SUV**, and **Sedan, Coupe** are very low, indicating these are rare or niche combinations in this dataset.

Part 2

Question 1 : Identify the top two manufacturers who offer the widest range of cars variants. Use appropriate charts/graphs to present the output, capturing the number of car variants under each body type for the two manufacturers identified above.



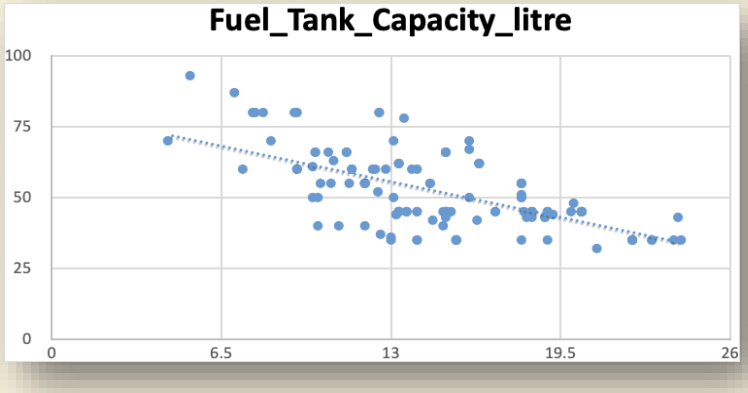


Suzuki and **Hyundai** are the top two with widest range of car variants and number of car variants by body type

Part 2

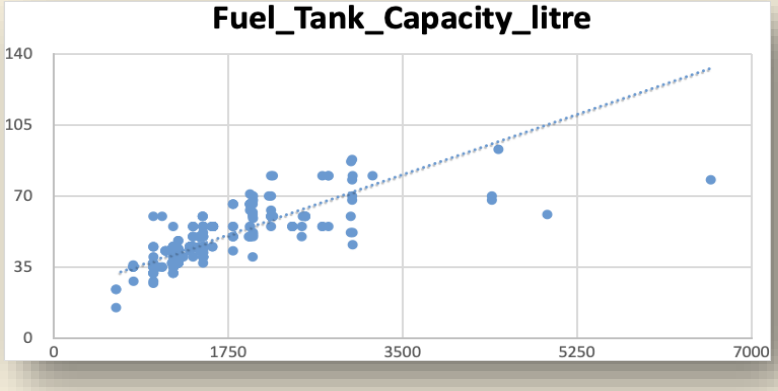
Question 3: Display and check how City Mileage and Displacement depends on Fuel Tank Capacity using scatter plot along with trend line and Pearson's product-moment coefficient(R).

City Mileage



The provided correlation coefficient (R) is -0.666355261, which suggests a moderate to strong negative correlation between City Mileage km litre and Fuel Tank Capacity litre

Displacement

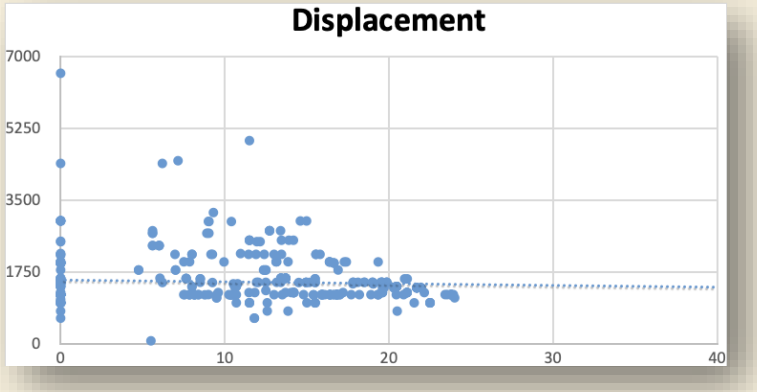


There seems to be a positive correlation: as displacement increases, fuel tank capacity also increases. This is likely due to larger engines requiring more fuel for optimal performance and longer ranges.

Part 2

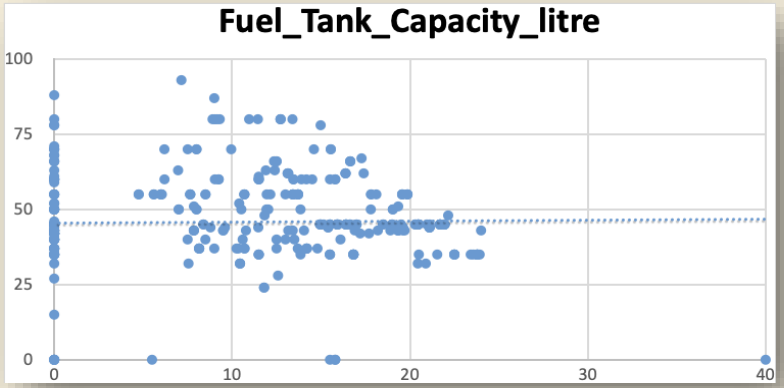
Question 5(B)(b): Use scatterplots and correlation matrices to expiore relationships between fuel efficiency and other variables.

Displacement



Displacement and Mileage: -0.06265422
This indicates a very weak negative correlation between engine displacement and mileage. In practical terms, engine displacement has little to no impact on mileage based on this dataset.

Fuel Tank Capacity



Fuel Tank Capacity and Mileage: 0.019442965
Interpretation: This indicates a very weak positive correlation between fuel tank capacity and mileage. Fuel tank capacity also has little to no impact on mileage based on this dataset.

Summary

1. Next 3 undertakings viz., 6,7 and 8 predominantly relies on investigation the different elements of vehicles to find the best vehicles for explicit purposes like for gutsy rides, where BMW (X4) MODEL is reasonable, ble, for Family utilization its recognized that Suzuki vehicles are appropriate and for reduced vehicle for city use where Goodbye (Nano Genx) is reasonable

2. Task 4 and 5 essentially relies on finding the connection between different elements parched limits between the vehicles delivered by different carmanufacturers. have utilized Dissipate plot diagrams which best depicts the between two factors. Task 3 arrangements with exception examination, where Hyundai being the producer sore number of anomalies which is seriously amazing.

3. Task 6 (Undertaking 1 to 3) predominantly managed distinguishing the Top vehicle producers and their vehicle models in view of wed key elements like City mileage, removal, the bodytype of vehicles whether hatchback, Mov, car and so on. In light of these perceptions Mahindra and Suzuki makers are doing great contrasted with different producers and the two of them are in top 10 records.

4. In the Task 7 (Venture 2) the examination is predominantly founded on noticing eco-friendliness to comprehend the driving examples which further aides in diminishing the nemissions all over the planet. The main 2 producers who has the largest scope of vehicle variations are, SUZUKI-163 Variants Hyundai-130 Variants

5. ask 9 and 10, the histogram is made and we have recognized the adjustment of histogram when the information is with exceptions and without anomalies. The main 3 producers under each bodytype is thought about additionally founded on the count of variations.