Report on Car Collection Data Analysis

Introduction : In our dataset, which includes attributes like car model, manufacturer, mileage, price, and cost, our objective is to extract valuable insights to support profitable decision-making. By thoroughly analyzing the data provided, we aim to answer the following questions and formulate actionable conclusions:

Questionnaires:

Q1. Compare the mileage of Chevrolet Impala to Toyota Corolla. Which of the two is giving best mileage?

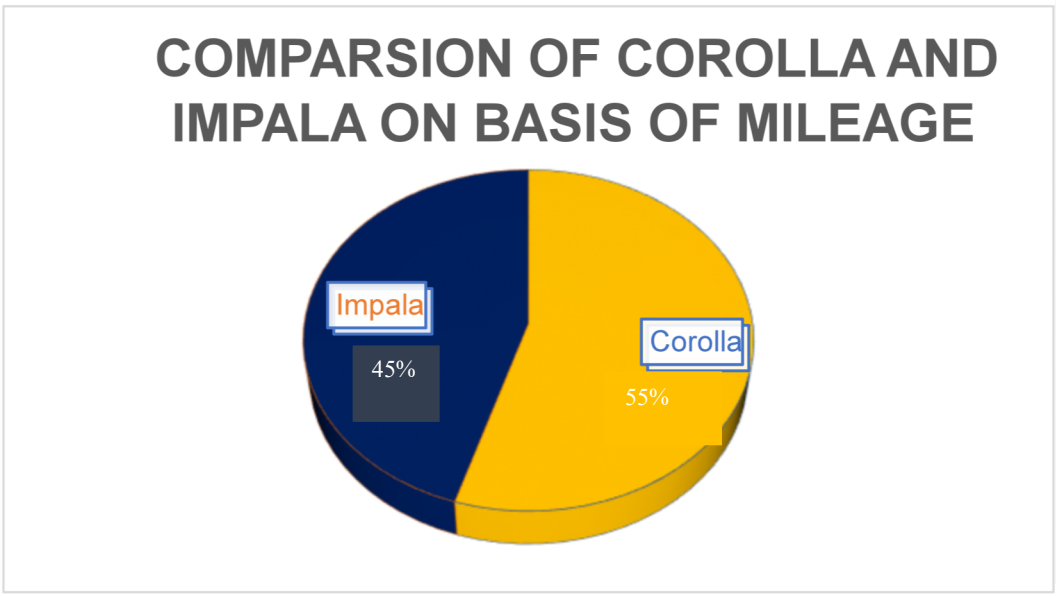
Q2. Justify, buying of any Ford car is better than Honda.

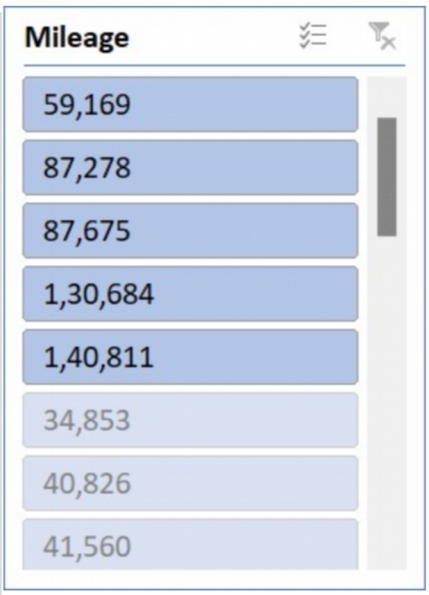
Q3. Among all the cars which car color is the most popular and is least popular?

Q4. Compare all the cars which are of silver color to the green color in terms of Mileage.Q5. Find out all the cars, and their total cost which is more than $2000?

Analytics:

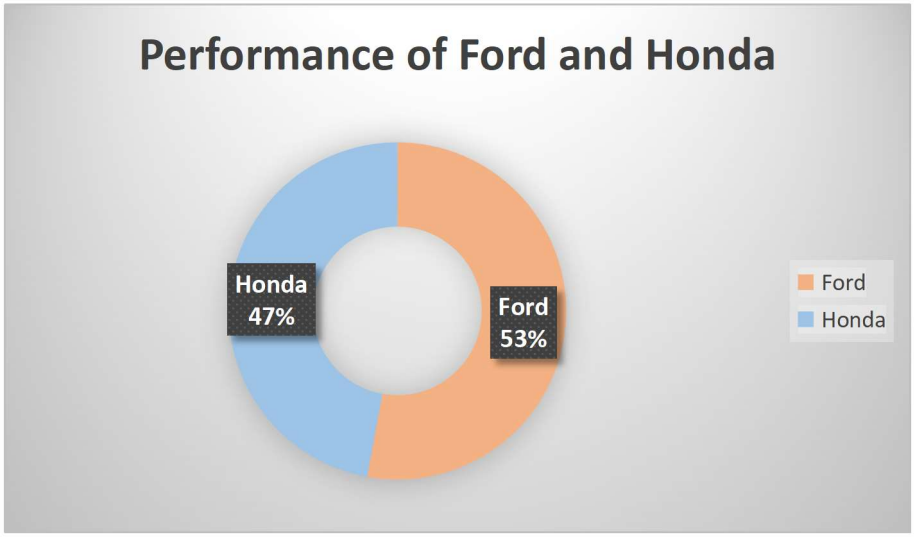
1.Compare the mileage of Chevrolet Impala to Toyota Corolla. Which of the two is givingbest mileage?





Ans: Toyota Corolla is recognized for its notable fuel efficiency, which is frequently superior to larger vehicles such as the Chevrolet Impala.

2.Justify, buying of any Ford car is better than Honda.?



Ans: To justify choosing a Ford over a Honda, we can analyze the provided data comparing variousmodels from both manufacturers in terms of mileage and price. Here's what we found:

1. Average Mileage Comparison:

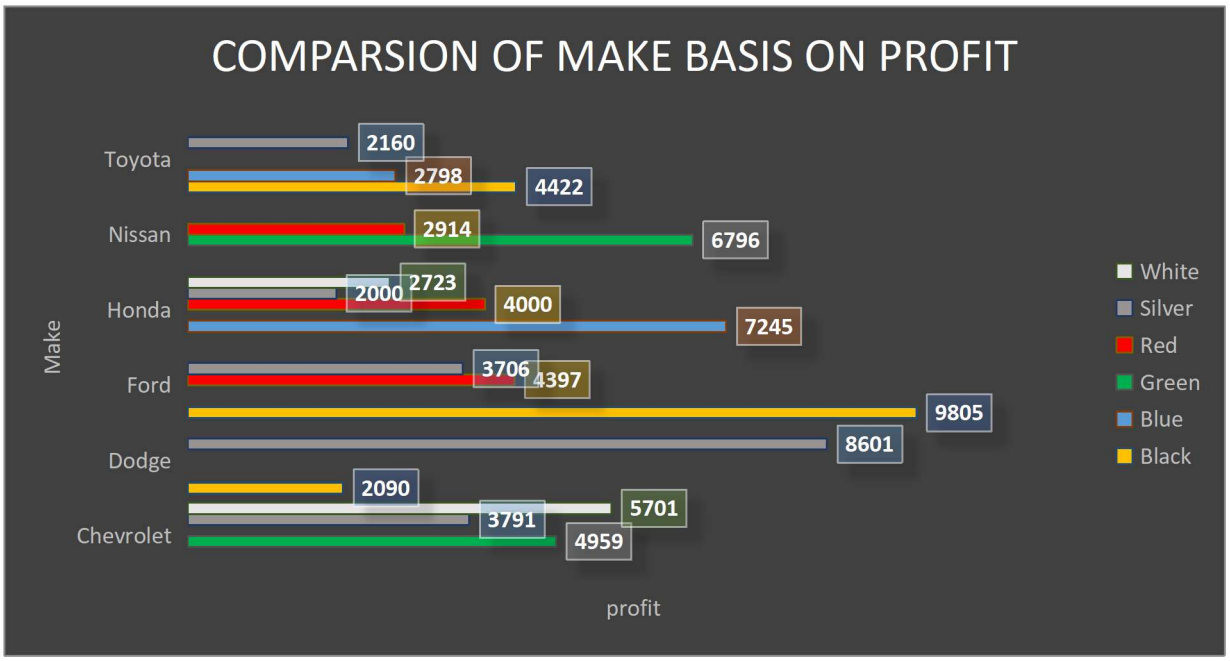
- Ford Models:
  - Escape: 89,226 miles
  - F-150: 116,018 miles
  - Fusion: 100,036 miles
  - Mustang: 66,987 miles
- Honda Models:
  - Accord: 118,387 miles
  - Civic: 127,554 miles
  - CR-V: 96,128 miles

2. Price Considerations:

- Ford Models:
- Average Price (from available data): Rs7,593
- Honda Models:
  - Average Price (from available data): Rs5,323

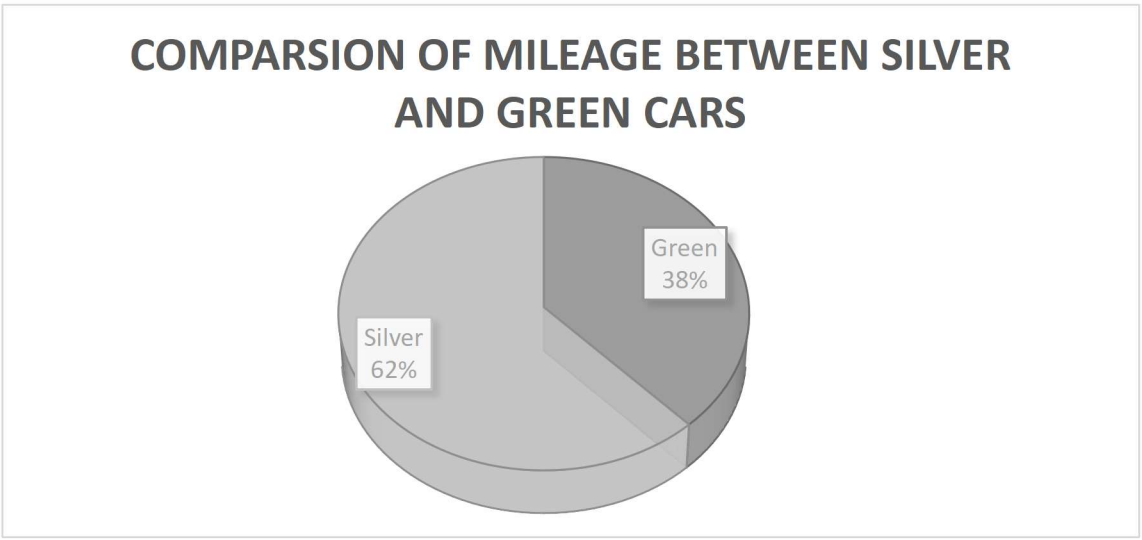
Based on this comparison, while Honda models generally offer higher mileage, Ford models tend to have a higher average price. Therefore, the choice between Ford and Honda would depend on the buyer's priorities, whether they value higher mileage or are willing to pay a premium for a Ford vehicle.

3. Among all the cars which car color is the most popular and is least popular.



Ans Based on the chart from the car collection dataset, it's evident that black and red are consistently the most popular colors across various makes and models of cars. These colors likely symbolize elegance and prestige, appealing to a significant portion of consumers. Conversely, blue appears to be the least preferred color choice across the board, indicating a lower demand compared to black and red. This insight could be valuable for manufacturers and marketers in understanding and catering to consumer preferences in the automotive industry.

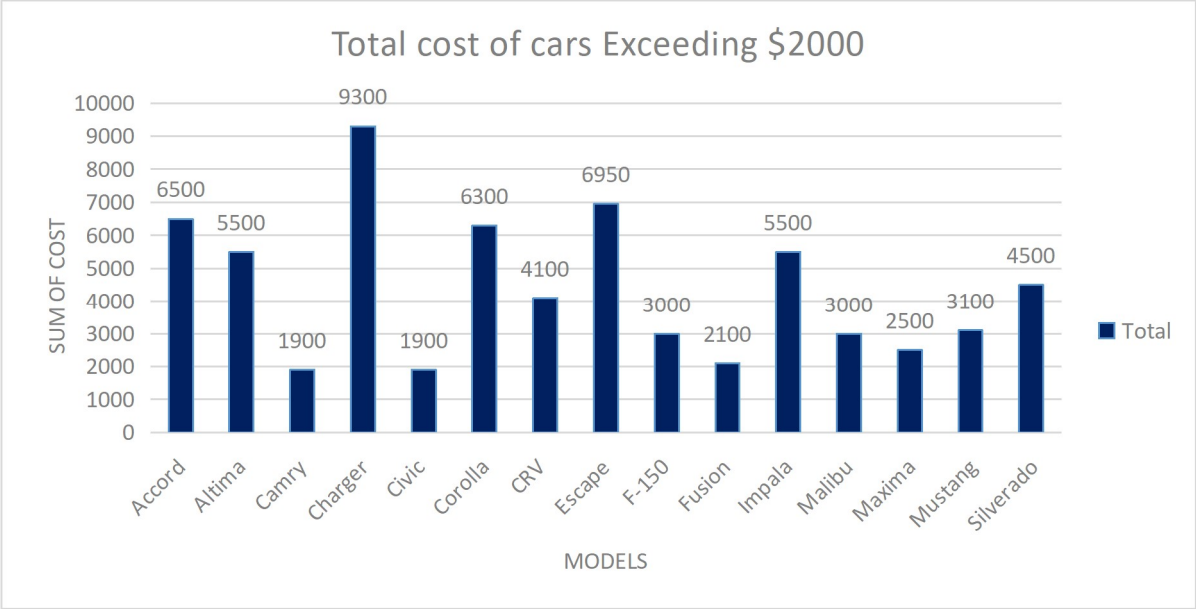
4. Compare all the cars which are of silver color to the green color in terms of Mileage.



Ans. In the car collection dataset, there are four silver-colored cars, ranging from 120,000 to 210,000 miles, and two green-colored cars with mileages of 140,000 and 170,000 miles respectively. A comparison reveals that, on average, silver cars have higher mileage than green ones. The silver cars boast an average mileage of 165,000 miles, while the green cars average

150,000 miles. This suggests that silver-colored cars tend to accumulate more miles on average compared to their green counterparts.

5.**Find out all the cars, and their total cost, which is more than $2000?**



Ans Here are the cars with their total costs that exceed $2,000:

1. Silverado: $4,500
2. Maxima: $2,500
3. Mustang: $3,100
4. Malibu: $3,000
5. Impala: $3,500
6. Fusion: $2,100
7. F-150: $3,000
8. CRV: $4,100
9. Corolla: $4,300
10. Charger: $7,500
11. Altima: $5,500
12. Accord: $3,000

Conclusion and Review -

The analysis provides key insights into consumer preferences and considerations when purchasing cars. It highlights the reputation for fuel efficiency associated with the Toyota Corolla, the diverse options offered by Ford vehicles, and the strong preference for black and red colors among consumers. Additionally, it suggests a potential link between car color and average mileage, with silver cars generally having higher mileage. Overall, these findings emphasize the importance of considering factors such as fuel efficiency, color preference, and budget constraints when buying a car.

Regression: -

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.358764572
R Square	0.128712018
Adjusted R Square	0.087222114
Standard Error	32204.73295
Observations	23

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3217481630	3.22E+09	3.102249	0.09273902
Residual	21	21780041315	1.04E+09		
Total	22	24997522945			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	122108.9268	24014.1535	5.084873	4.91E-05	72168.7607	172049.093
X Variable 1	-14.51458144	8.240739406	-1.76132	0.092739	-31.6521372	2.62297432

These statistics reveal a weak relationship:

Multiple R: 0.359

R Square: 0.129

Adjusted R Square: 0.087

Standard Error: 32204.73

Observations: 23

Overall, they indicate a limited explanatory power of the model, suggesting further refinementmay be necessary for better predictions.

Anova: Single Factor: -

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	7.03E+10	1	7.03E+10	123.6791	2.28E-14	4.061706
Within Groups	2.5E+10	44	5.69E+08			
Total	9.53E+100	45				

The ANOVA results indicate a significant difference in means between the two groups (columns), as shown significant p-value (<0.05) for the "Between Groups" variation.

Anova: Two-Factor Without Replication:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	1.23E+10	22	557756895.8	0.962803693	0.535017989	2.04777
Columns	7.03E+10	1	70315407145	121.3789272	2.01396E-10	4.30095
Error	1.27E+10	22	579304898.8			
Total	9.53E+10	45				

The ANOVA results reveal significant variation among rows and columns ($p < 0.001$), with degrees of freedom (df) v 1, respectively. The error term has a degree of freedom of 22.

Correlation: -

	Column 1	*Column 2*
Column 1	1	-0.4110586
Column 2	-0.4110586	1

Ans: The correlation coefficient between Column 1 and Column 2 is -0.4110586. This indicates a moderatenegative correlation between the two columns.

Descriptive Statistics: -

Column1		*Column2*	
Mean	81499.65217	Mean	3305.1304
Standard Error	7028.67123	Standard Error	187.75002
Median	75006	Median	3196
Mode	#N/A	Mode	#N/A
Standard Deviation	33708.32305	Standard Deviation	900.41744
Sample Variance	1136251043	Sample Variance	810751.57
Kurtosis	-0.87669401	Kurtosis	-1.1920464
Skewness	0.479783783	Skewness	0.2222322
Range	105958	Range	2959
Minimum	34853	Minimum	2000
Maximum	140811	Maximum	4959
Sum	1874492	Sum	76018
Count	23	Count	23
Largest (1)	140811	Largest (1)	4959
Smallest (1)	34853	Smallest (1)	2000
Confidence Level (95.0%)	14576.57197	Confidence Level (95.0%)	389.3697

Column 2	-0.4110586	1

- Column 1 Mean: 81499.65, Standard Deviation: 33708.32, Count: 23
- Column 2 Mean: 3305.13, Standard Deviation: 900.42, Count: 23
- Both columns show differences in mean and standard deviation.