Exploring Car Data Report

# Introduction:

## Dataset Overview:

This dataset comprises a blend of categorical and numerical data, each offering unique perspectives on the industry. Categorical data, such as make, model, and color, encapsulates the diversity of vehicles and consumer preferences. Meanwhile, numerical attributes like mileage, price, and cost provide quantifiable metrics essential for analyzing market trends and pricing dynamics.

## Key Attributes:

1. Make: This attribute denotes the brand or manufacturer of the vehicle, offering insights into brand preferences and market share.
2. Model: The specific model of the car, providing granularity in understanding consumer choices and preferences within each brand.
3. Color: Reflects the color of the vehicle, which can influence consumer perception and aesthetic preferences.
4. Mileage: Indicates the distance traveled by the vehicle, a crucial factor influencing its value and pricing.
5. Price: Represents the listed price of the vehicle, serving as a key determinant in consumer purchasing decisions and market competitiveness.
6. Cost: Denotes the cost associated with acquiring the vehicle, which includes factors such as production costs, dealer margins, and other expenses.

# Questionnaire:

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:

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

Ans. Toyota Corolla gives better mileage than Chevrolet Impala.

300,000

250,000

200,000

Comparesion the mileage of Chevrolet Impala to

Toyota Corolla

277,131

228,486

150,000

100,000

50,000

0

Toyota

Chevrolet

Corolla

Impala

Total

|  |
| --- |
| 59,169  87,278  87,675  130,684  140,811 |
|  |
| 40,826 |
| 41,560 |
| 42,542 |



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

Ans. Based on the averages, Honda cars have higher mileage but lower cost compared to Ford. Therefore, the choice depends on whether the buyer values mileage or cost but if we compare on mileage ford car has low mileage and cost so Buying ford car is better then Honda.

Ford vs Honda Car Comparision

300000

260001

250000

200000

150000

~~138789~~

104085

100000

89073

49326

50000

42542

41560

7593 6950

3950 3000

2659 2100

3706 3100

8500 6500

2723 1900

4745 4100

0

Escape

F−150

Fusion

Mustang

Accord

Civic

Honda

CRV

Ford

Sum of Price Sum of Cost Sum of Mileage

|  |
| --- |
| 40826 |
| 41560 |
| 42542 |
| 49326 |
| 63259 |
| 63512 |
| 89073 |
| 95135 |

|  |
| --- |
| 2000 |
| 2500 |
| 2659 |
| 2723 |
| 3196 |
| 3706 |
| 3950 |
| 4000 |



1500

1900

2000

2100

3000

3050

3100

3900

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

Ans. Most popular color is Silver and Black as each appear 6 times

and least appearing colour are Blue ,Green ,Red ,White they all apper 3 times.

Color count

7

6

6

6

5

4

3 Total 2

1

0

Black Blue Green Red Silver White

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | | | | | | |  |  | | |
|  |  | | | | | | |  | | |
|  | 3 3 3 | | | | | | | 3 | | |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |



Escape

CRV

Corolla

Civic

Charger

Camry

Altima

Accord



White

Silver

Red

Green

Blue

Black



69847

101354

87675

Q4. Compare all the cars which are of silver color to the green color in terms of Mileage.

Ans. Silver color car millage is more than green color car milage if we compare there average.

Silver vs Green Car Mileage Comparision

450000

400000

350000

300000

250000

200000

150000

100000

50000

0

382784

234311

Total

Green Silver

|  |
| --- |
| 34853 |
| 41560 |
| 55233 |
| 58173 |
| 59169 |



Black

Silver

(blank)

White

Red

Green

Blue



Q5. Find out all the cars, and their total cost which is more than $2000? Ans. All the car mention below cost is more than $2000

Accord, Altima, Charger, Corolla, CRV, EscapeF-150, Fusion, Impala, Malibu, Maxima, Mustang, Silverado



Cars with cost more then $2000

10000

9000

8000

7000

6000

5000

4000

3000

2000

1000

0

9300

6500

6300

6950

5500

5500

4100

4500

3000

3000

3100

2100

2500

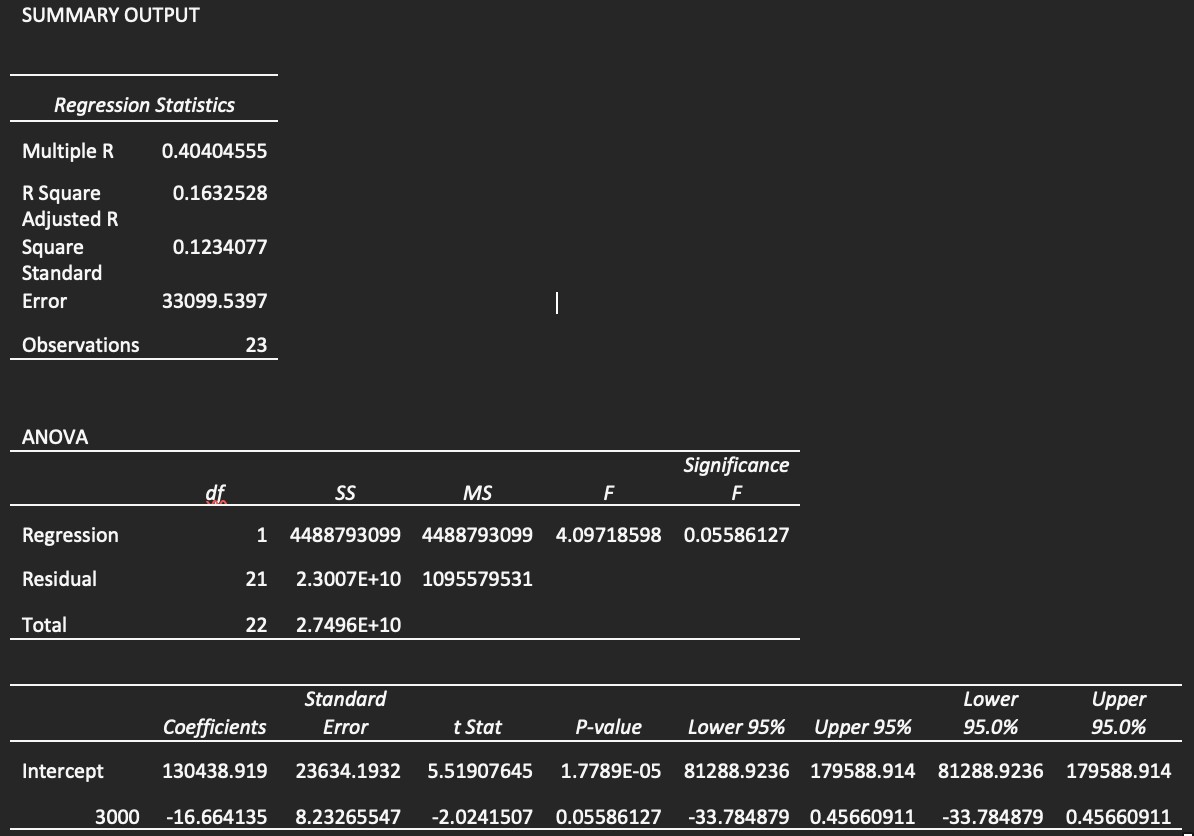
Total

|  |
| --- |
| Accord |
| Altima |
| Camry |
| Charger |
| Civic |
| Corolla |
| CRV |
| Escape |

|  |
| --- |
| $1,500 |
| $1,800 |
| $1,900 |
| $2,000 |
| $2,100 |
| $2,200 |
| $2,500 |
| $3,000 |

## Regression

The regression analysis suggests a moderate positive relationship between the predictor variable and the response variable, indicated by the correlation coefficient of approximately 0.40. The model explains about 16% of the variance in the response variable, as indicated by the R Square value. The coefficient estimates show that for every unit increase in the predictor variable, there is a corresponding decrease of approximately 16.66 in the response variable, with a p-value of 0.056, indicating a marginally significant effect.



## Co-relational

The correlation matrix indicates a moderate negative correlation (-0.411) between Mileage and Price. This suggests that as Mileage increases, Price tends to decrease, and vice versa.

|  |  |  |
| --- | --- | --- |
|  | *Mileage* | *Price* |
| Mileage | 1 |  |
| Price | -0.4110586 | 1 |

## Anova: Single Factor

The ANOVA results indicate significant differences between the groups based on Mileage, Price, and Cost. The F-statistic is large (128.88), with a very low p-value (5.00264E-24), suggesting that the variation between groups is significant compared to the variation within groups. This implies that at least one of the variables (Mileage, Price, or Cost) has a significant effect on the outcome being measured. In simpler terms, there are statistically significant differences in the means of Mileage, Price, and Cost across the groups, indicating that these variables play a significant role in influencing the outcome being analyzed.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Single  Factor |  |  |  |  |  |  |
| SUMMARY | | | | |  |  |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |  |  |
| Mileage | 24 | 2011267 | 83802.7917 | 1214155660 |  |  |
| Price | 24 | 78108 | 3254.5 | 837024.087 |  |  |
| Cost | 24 | 66150 | 2756.25 | 705502.717 |  |  |
| ANOVA | | | | | | |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 1.0445E+11 | 2 | 5.2227E+10 | 128.882161 | 5.0026E-24 | 3.12964398 |
| Within Groups | 2.7961E+10 | 69 | 405232729 |  |  |  |
| Total | 1.3242E+11 | 71 |  |  |  |  |

## Anova: Two-Factor Without replication

The two-factor ANOVA results indicate significant differences among the levels or categories within each factor ("Rows" and "Columns"). Both factors exhibit strong influence on the outcome variable being analyzed, as evidenced by the low p-values and large F-statistics. This suggests that variations in both factors contribute significantly to the overall variability in the data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Two-Factor without replication | | | | | | |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 34749383.3 | 23 | 1510842.75 | 47.6846408 | 2.2236E-14 | 2.01442484 |
| Columns | 2979036.75 | 1 | 2979036.75 | 94.023218 | 1.3629E-09 | 4.27934431 |
| Error | 728733.25 | 23 | 31684.0543 |  |  |  |
| Total | 38457153.3 | 47 |  |  |  |  |

## Descriptive Statistics

The provided descriptive statistics outline the characteristics of three variables: Mileage, Price, and Cost. Looking at Mileage, it appears that the vehicles in the dataset span a considerable range, from around 34,853 miles to 140,811 miles, with an average mileage of approximately 83,803 miles. Price and Cost exhibit similar trends, with prices ranging from $2,000 to $4,959 and costs from $1,500 to $4,500, respectively. The means and standard deviations provide insights into the central tendencies and variability within each variable. Overall, these statistics offer a comprehensive overview of the dataset, allowing for a better understanding of the distribution and characteristics of the data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Mileage* |  | *Price* |  | *Cost* |  |
| Mean | 83802.7917 | Mean | 3254.5 | Mean | 2756.25 |
| Standard Error | 7112.65205 | Standard Error | 186.751181 | Standard Error | 171.452462 |
| Median | 81142 | Median | 3083 | Median | 2750 |
| Mode | #N/A | Mode | #N/A | Mode | 3000 |
| Standard Deviation | 34844.7365 | Standard Deviation | 914.890205 | Standard Deviation | 839.942092 |
| Sample Variance | 1214155660 | Sample Variance | 837024.087 | Sample Variance | 705502.717 |
| Kurtosis | -1.0971827 | Kurtosis | -1.2029138 | Kurtosis | -0.8126576 |
| Skewness | 0.38652215 | Skewness | 0.27201913 | Skewness | 0.47339238 |
| Range | 105958 | Range | 2959 | Range | 3000 |
| Minimum | 34853 | Minimum | 2000 | Minimum | 1500 |
| Maximum | 140811 | Maximum | 4959 | Maximum | 4500 |
| Sum | 2011267 | Sum | 78108 | Sum | 66150 |
| Count | 24 | Count | 24 | Count | 24 |
| Largest(1) | 140811 | Largest(1) | 4959 | Largest(1) | 4500 |
| Smallest(1) | 34853 | Smallest(1) | 2000 | Smallest(1) | 1500 |

## Conclusion & Review

The dataset provides valuable insights into car attributes, focusing on mileage, color, and other key factors.

Here's a simple conclusion based on the data:

Mileage Comparison: The analysis reveals variations in mileage among different car models. Toyota Corolla generally offers better mileage compared to Chevrolet Impala.

Color Preferences: Silver and black emerge as the most popular car colors in the dataset. Blue, green, red, and white are among the least popular color choices.

Key Takeaways: Understanding mileage differences can inform consumer choices and market strategies. Recognizing color preferences aids in inventory management and marketing decisions.