**Course- PGP in Business**

**Analytics**

**Name - Prabir Das**

**Topic - Identifying Key Factors**

**Influencing Automobile**

**Pricing and Performance**

**Across Different**

**Manufacturers**

**Date - 15/06/2024**

**Executive Summary:**

The primary purpose of this study is to identify and analyse the key factors that influence automobile prices and to examine the variations in builds manufactured by different companies. By understanding these factors, we aim to provide insights that can help consumers make informed purchasing decisions, assist manufacturers in optimizing their pricing strategies, and contribute to the overall understanding of market dynamics in the automotive industry. Also, contribute to the existing body of knowledge in the automotive industry by offering detailed insights into the relationship between automobile features, manufacturer attributes, and pricing.

The study reveals that most cars in the dataset are relatively inexpensive. However, there is a considerable range in prices, indicating significant variability in the market. When looking at factors that influence the weight of a car, the analysis shows that the length and width of a vehicle are particularly important, with longer and wider cars generally weighing more. This is supported by a model that explains a substantial 91% of the variation in curb weight, suggesting a very strong relationship between these dimensions and the car's overall weight.

Additionally, the analysis identifies key factors that significantly impact car prices. Engine size stands out as having the most substantial positive effect, meaning cars with larger engines are typically priced higher. Conversely, stroke has a negative impact on price, indicating that cars with higher stroke values tend to be cheaper. Other factors such as curb weight also play a notable role, with heavier cars generally being more expensive. Overall, the models used in the study provide a strong explanation of the factors influencing both car weight and price.

The study highlights several key insights for manufacturers. Understanding that factors like engine size and performance metrics significantly impact car prices allows manufacturers to tailor their models to consumer preferences, potentially boosting sales. Gas cars with higher horsepower suggest a design focus on performance, while diesel cars may prioritize fuel efficiency. Brands like Toyota and Honda, with their diverse range of models, cater to various market segments, enhancing their presence. In contrast, luxury brands like Jaguar, Mercedes-Benz, and Porsche focus on high-end offerings that attract affluent customers willing to pay premium prices. This reveals a strategic approach to balancing market coverage and brand prestige.

The study also emphasizes the trade-offs between performance and fuel efficiency. Manufacturers face challenges in balancing engine size, horsepower, and fuel economy, prompting innovation in powerful yet efficient engines. The negative impact of vehicle weight on fuel efficiency suggests a trend towards using lightweight materials to reduce weight without compromising safety or performance, potentially lowering manufacturing costs. Additionally, the strong correlations between vehicle dimensions indicate that design changes in one aspect necessitate adjustments in others, influencing overall design and manufacturing processes. Ultimately, aligning vehicle specifications with consumer expectations can enhance customer satisfaction and loyalty, driving long-term success.

**INDEX**

**TOPIC Page Number**

**1. Executive Summary 1**

**2. Introduction 3**

**3. Understanding Individual Features 4**

**4. Understanding The Relationship Between**

**Different Features 11**

* **Corelation Between Numerical Variables 11**
* **Chi-square Test Between Categorical Variables 16**
* **T-test Between Categorical and Numerical**

**Variables 21**

**5. Identifying significant Features that Influence Price 29**

**6. Identifying significant Features that Influence**

**Curb-weight 31**

**7. Identifying significant Features that Influence**

**Horsepower 33**

**8. Conclusion 35**

**INTRODUCTION:**

The automotive industry stands as a pillar of the global economy, undergoing continuous evolution and transformation. This report embarks on a comprehensive analysis of the automobile dataset, delving deep into market dynamics to identify key insights shaping the industry's growth trajectory.

The analysis not only helps in understanding market dynamics but also extends to predicting prices and identifying significant features driving consumer preferences and market trends. By leveraging advanced analytical techniques, we aim to provide a holistic view of the industry landscape.

Our examination of the dataset begins with a meticulous exploration of customer preferences and purchasing patterns, unravelling intricate trends that dictate demand for various vehicle types and features. Using regression analysis, we further predict prices based on key variables such as vehicle horsepower, mileage, brand, and features.

Moreover, we delve into sales data to understand market segmentation, pricing strategies, and the performance of different automotive categories. This exploration enables us to identify significant features that influence price variations and consumer choices, guiding strategic decision-making for industry stakeholders.

In addition to consumer-driven aspects, our analysis extends to the evaluation of industry benchmarks, regulatory influences, and technological innovations. Factors such as fuel efficiency, safety standards, and performance profoundly impact market dynamics, shaping consumer choices and industry competitiveness.

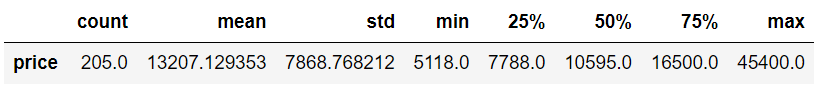
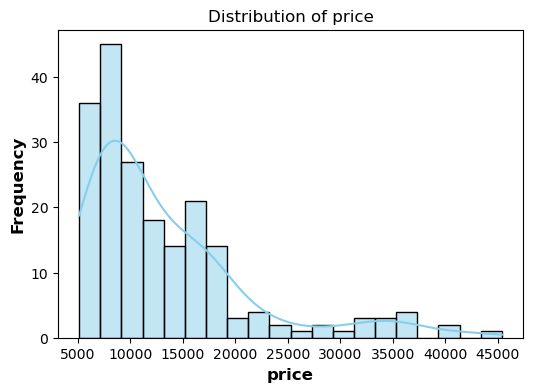
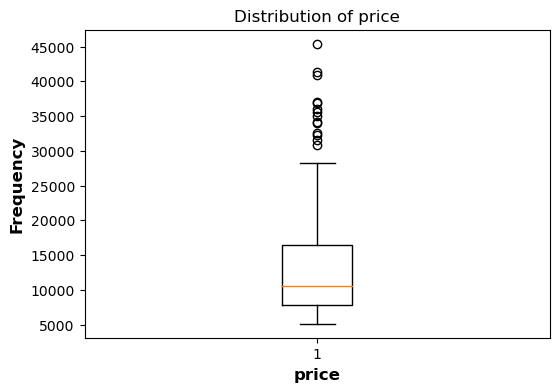
Furthermore, a comprehensive review of the competitive landscape and industry trends provides valuable insights into market positioning, competitive advantages, and areas for improvement for automotive companies.

This report culminates with strategic recommendations tailored to help industry players capitalize on emerging trends, optimize operational efficiencies, and foster sustainable growth. By leveraging data-driven insights, predictive analytics, and strategic foresight, stakeholders can navigate the complex automotive landscape with confidence and agility.

# **References**

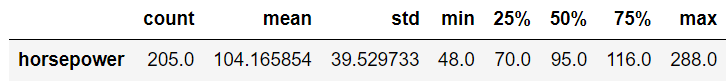
Automobile Dataset. Retrieved from Kaggle: https://www.kaggle.com/datasets/toramky/automobile-dataset

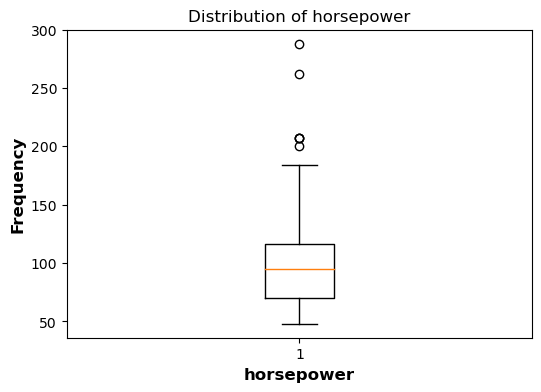
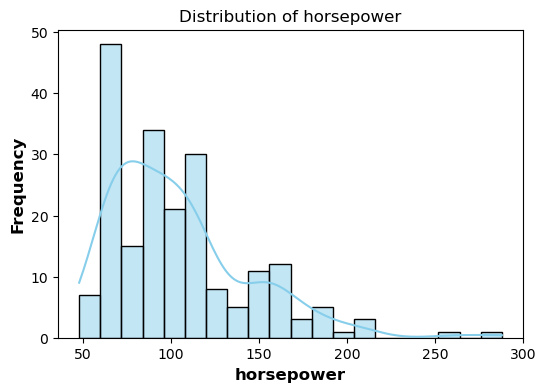
**Understanding Individual Features:**

**Distribution of Price:**  

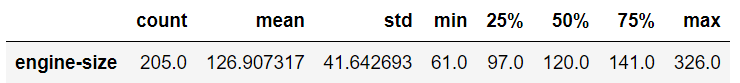
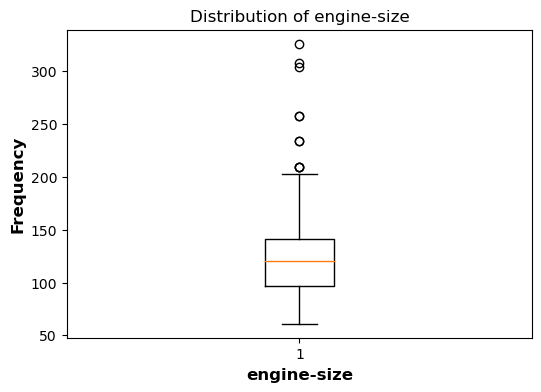
1. The skewness value is 1.83, which indicates Price has a positively skewed distribution. This means that there are more cars priced lower.
2. The average car price in the dataset is $13,207.13 and standard deviation is $7,868.77. This high value suggests there is considerable variability in car prices, indicating a wide range of car prices.

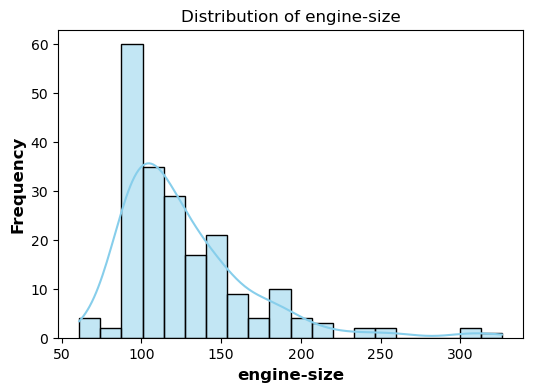
**Distribution of Horsepower:**



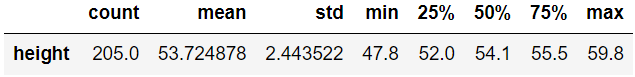
 

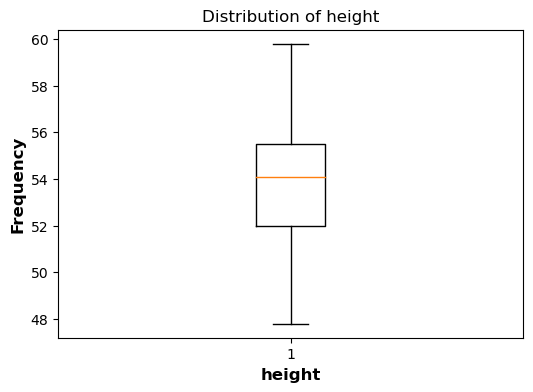
1. The average horsepower of the cars in the dataset is 104.17. This indicates that, on average, cars in the dataset have a moderate level of engine power.
2. The standard deviation of horsepower is 39.53. This relatively high value suggests significant variability in the horsepower of cars.
3. The skewness value is 1.40, which is positive, indicating that the distribution of horsepower is right-skewed. This means that there are more cars with lower horsepower, with a few high-horsepower cars.

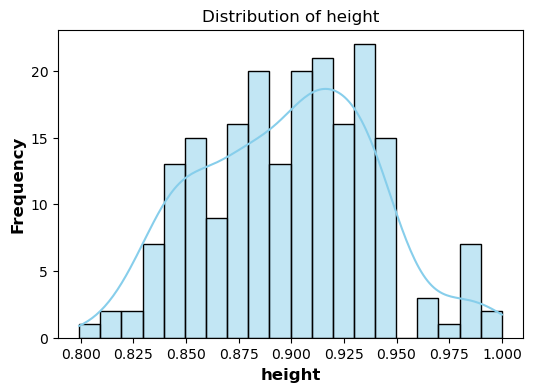
**Distribution of Engine-size: ** 



1. The data reveals that the average engine size among the cars in the dataset is 126.91 cubic inches.
2. With a standard deviation of 41.64 cubic inches, there is a notable degree of variation in engine sizes.
3. The skewness value of 1.95 indicates Engine-size has positively skewed distribution which means there are more cars with smaller engine.
4. The kurtosis value of 5.31, which is well above 3, points to a distribution with heavy tails and a sharp peak.

**Distribution of Height:** ****

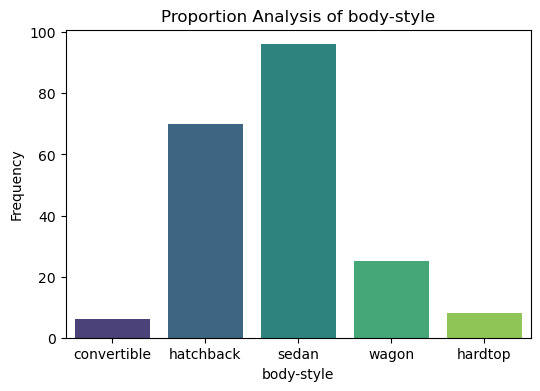




1. The data reveals that the average height among the cars in the dataset is 53.72.
2. The standard deviation and variance indicate that there is some variability in height, but the values are relatively close to the mean.
3. The height data is approximately symmetric with a skewness close to zero.

**Distribution of Body-style:**

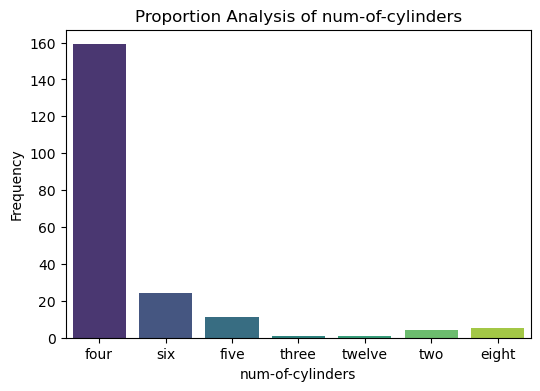
| **body\_style** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **Frequency** | **Percent** | **Valid Percent** | **Cumulative Percent** |
| **Valid** | **convertible** | **6** | **2.9** | **2.9** | **2.9** |
| **hardtop** | **8** | **3.9** | **3.9** | **6.8** |
| **hatchback** | **70** | **34.1** | **34.1** | **41.0** |
| **sedan** | **96** | **46.8** | **46.8** | **87.8** |
| **wagon** | **25** | **12.2** | **12.2** | **100.0** |
| **Total** | **205** | **100.0** | **100.0** |  |



1. we see that most of the cars are either sedans or hatchbacks. This indicates that these two body styles dominate the market.
2. Sedans are the most common body style in the dataset. This indicates that sedans are a highly popular choice among car buyers.
3. Convertibles and hardtops are the least common body styles, with only 6 (2.9%) and 8 (3.9%) cars, respectively. This low frequency indicates that these body styles are less preferred.

**Distribution of Number of Cylinders:**

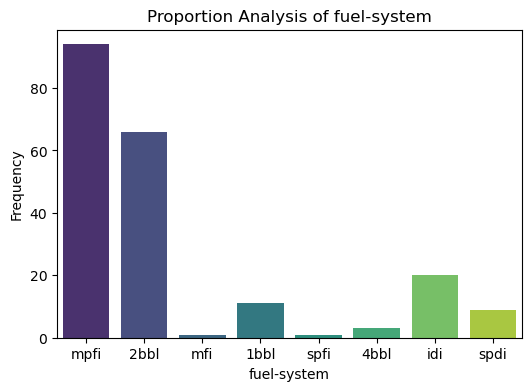
| **num\_of\_cylinders** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **Frequency** | **Percent** | **Valid Percent** | **Cumulative Percent** |
| **Valid** | **eight** | **5** | **2.4** | **2.4** | **2.4** |
| **five** | **11** | **5.4** | **5.4** | **7.8** |
| **four** | **159** | **77.6** | **77.6** | **85.4** |
| **six** | **24** | **11.7** | **11.7** | **97.1** |
| **three** | **1** | **.5** | **.5** | **97.6** |
| **twelve** | **1** | **.5** | **.5** | **98.0** |
| **two** | **4** | **2.0** | **2.0** | **100.0** |
| **Total** | **205** | **100.0** | **100.0** |  |



1. Four-cylinder engines dominate the dataset, with 159 cars, accounting for 77.6% of the total. This indicates a strong preference for four-cylinder engines.
2. Engines with three, twelve, and two cylinders are extremely rare. This highlights that engines with these cylinder counts are not popular choices.
3. This distribution emphasizes a clear trend towards smaller, more efficient engines, with only a small fraction of cars using engines with fewer or more than these common configurations.

**Distribution of Fuel-System:**

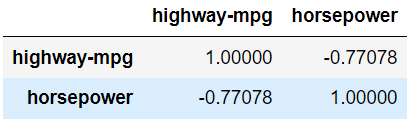
| **fuel\_system** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **Frequency** | **Percent** | **Valid Percent** | **Cumulative Percent** |
| **Valid** | **1bbl** | **11** | **5.4** | **5.4** | **5.4** |
| **2bbl** | **66** | **32.2** | **32.2** | **37.6** |
| **4bbl** | **3** | **1.5** | **1.5** | **39.0** |
| **idi** | **20** | **9.8** | **9.8** | **48.8** |
| **mfi** | **1** | **.5** | **.5** | **49.3** |
| **mpfi** | **94** | **45.9** | **45.9** | **95.1** |
| **spdi** | **9** | **4.4** | **4.4** | **99.5** |
| **spfi** | **1** | **.5** | **.5** | **100.0** |
| **Total** | **205** | **100.0** | **100.0** |  |



1. The most common fuel system is mpfi, used by 94 cars (45.9%).
2. The second most common is 2bbl, used by 66 cars (32.2%).
3. Including the previous category 78.1% of the cars use either a 2-barrel or mpfi system. This indicates a strong preference for 2-barrel or mpfi fuel-system.
4. Other fuel systems (1bbl, 4bbl, idi, mfi, spdi, spfi) have lower frequencies, with mfi and spfi being the least common, each used by only 1 car (0.5%).

**Understanding The Relationship Between Different Features:**

**Correlation of Highway-mpg and Horsepower:**

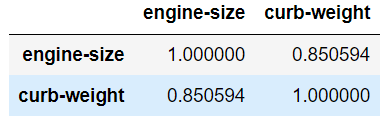


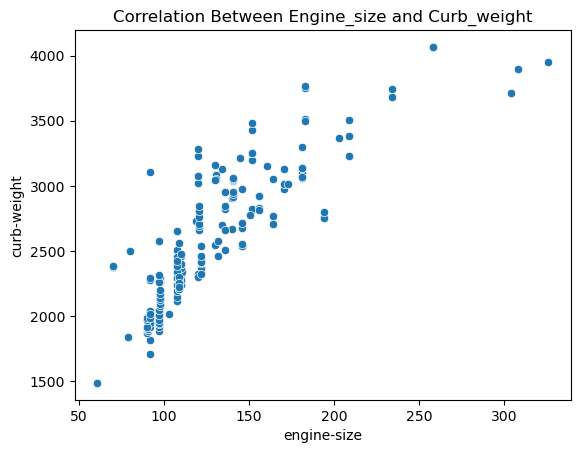
A graph with blue dots

Description automatically generated

1. There is a strong negative correlation (-0.77) between highway miles per gallon (mpg) and horsepower. This suggests that as highway mpg increases, horsepower tends to decrease, or vice versa.
2. The negative correlation between highway mpg and horsepower implies that vehicles with higher horsepower may have lower fuel efficiency on highways.
3. Understanding this correlation can be valuable for predictive modelling in the automotive industry.

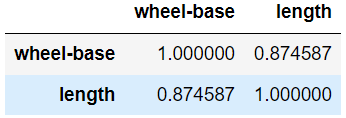
**Correlation of Engine-size and Curb-weight:**

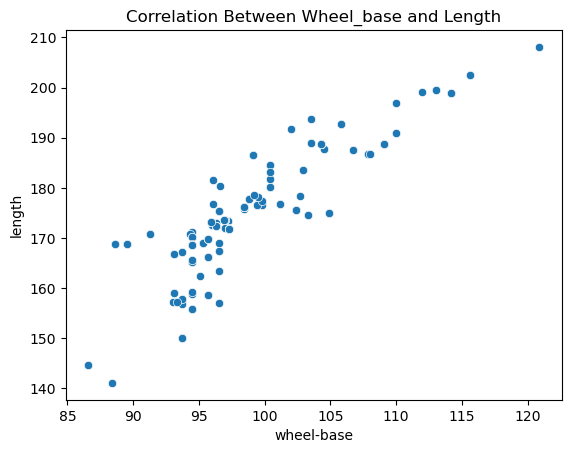




1. The Pearson correlation coefficient between engine size and curb weight is 0.85. This indicates a strong positive correlation.
2. When engine size increases, curb weight tends to increase as well.
3. The scatterplot shows a clear upward trend, reinforcing the strong positive correlation.
4. The strong positive correlation between engine size and curb weight suggests that vehicles with larger engines tend to be heavier.

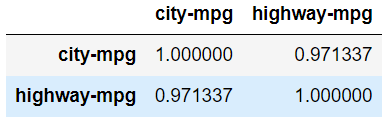
**Correlation of Length and Wheelbase:**

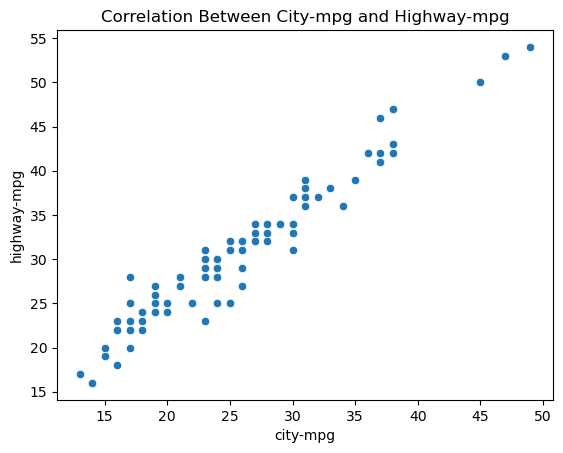




1. There is a strong positive correlation between the length of the vehicle and its wheelbase, with a correlation coefficient of 0.874. This suggests that as the length of a vehicle increases, its wheelbase tends to increase as well.
2. The correlation value of 0.874 is very high, close to 1, indicating a very strong linear relationship between the two variables.
3. The strong correlation suggests that knowing the length of a vehicle can be a good predictor of its wheelbase, and vice versa.

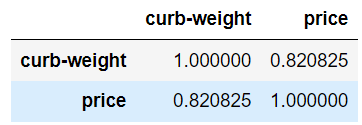
**Correlation of City-mpg and Highway-mpg:**





1. There is an extremely high positive correlation between city miles per gallon (city-mpg) and highway miles per gallon (highway-mpg), with a Pearson correlation coefficient of 0.971.
2. Vehicles that are fuel-efficient in city driving conditions are almost always fuel-efficient on the highway as well.
3. The high correlation suggests that knowing the city-mpg of a vehicle can be a very good predictor of its highway-mpg, and vice versa.

**Correlation of Curb-weight and Price:**

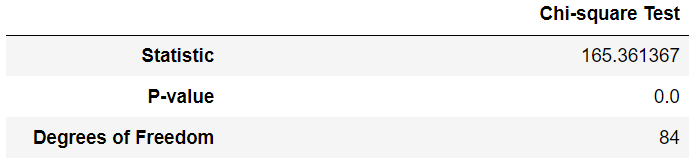
****

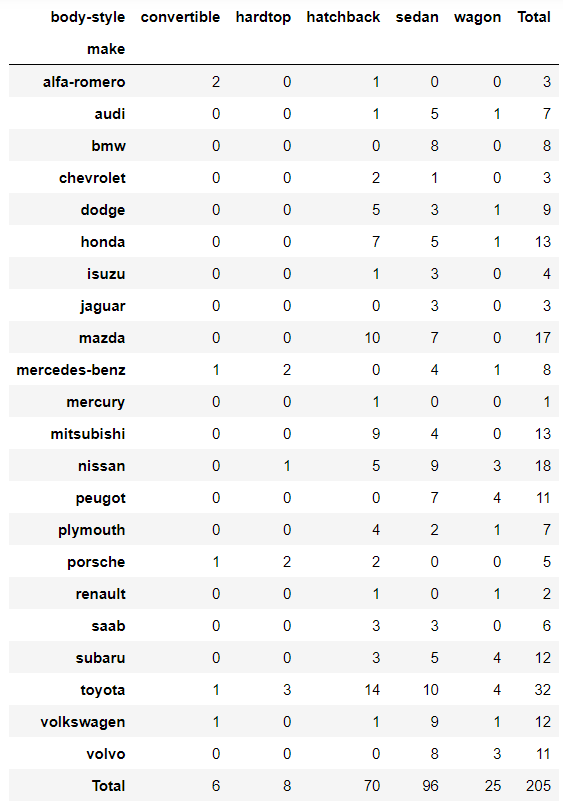
A graph with blue dots

Description automatically generated

1. There is a strong positive correlation (r = 0.821) between curb weight and price, indicating that as the curb weight of a car increases, its price tends to increase as well.
2. The prices of the cars vary widely, ranging from under $10,000 to over $45,000, as shown in the scatterplot.
3. There are several outliers with high prices and high curb weights, suggesting the presence of luxury or high-performance vehicles that are heavier and more expensive.

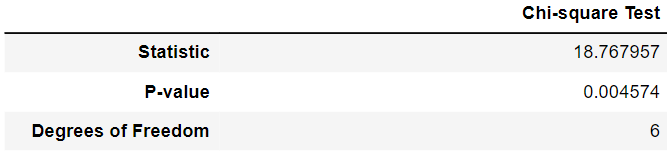
**Distribution of Body-style based on Make:**

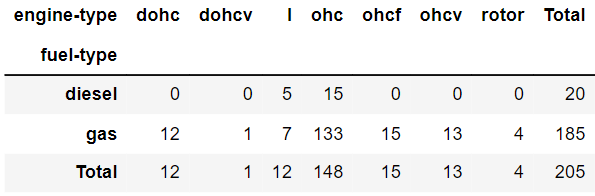
vblllllbivdhssssssssvdsbikv 

****

1. Chi-square P-value is less than level of significance so we can say there is dependency between make and body-style.
2. Toyota makes more cars compared to others.
3. BMW and jaguar only make cars with sedan body style.
4. Chevrolet and Isuzu only make cars with sedan or hatchback body-style.
5. There are very less cars with convertible or hardtop body-style

**Distribution of Engine-type based on Fuel-type:**



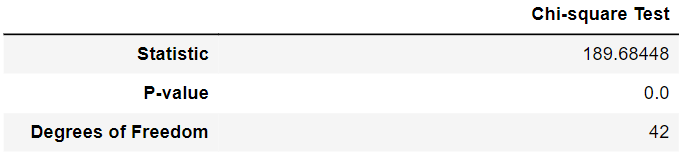


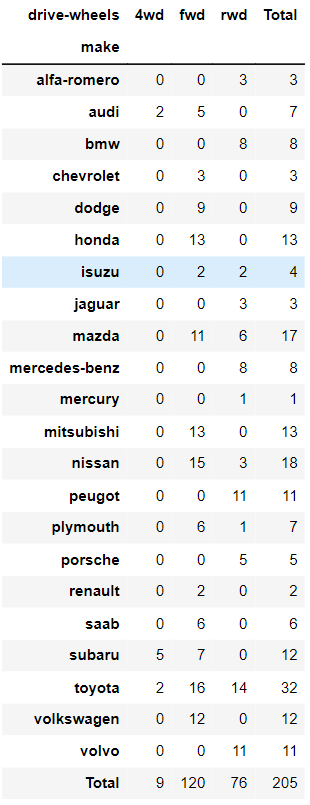
A graph of a number of different types of fuel type

Description automatically generated

1. The chi-square test indicates a statistically significant relationship between engine type and fuel type (Pearson Chi-Square value = 18.76, p = 0.004), suggesting that the distribution of fuel types varies significantly among different engine types.
2. The OHC (overhead camshaft) engine type is the most prevalent, with a total of 148 vehicles, the majority of which use gas (133 vehicles).
3. Some engine types, such as DOHC, DOHCV, OHCF, OHCV, and rotor, exclusively use gas.
4. The rotor engine type is the least common, with only 4 vehicles, all using gas.
5. Across all engine types, gas is the predominant fuel, with 185 out of 205 vehicles using it.

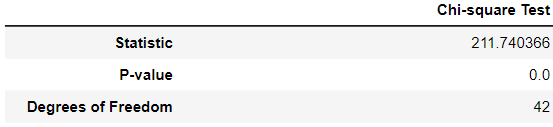
**Distribution of Make based on Driving wheels:**

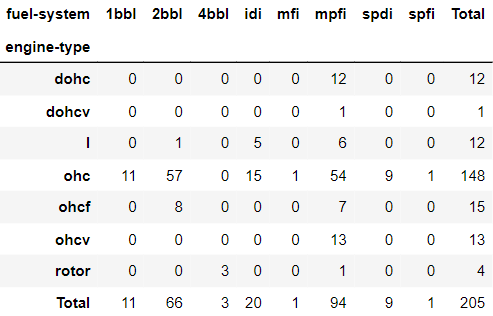




1. The Chi-Square tests indicate a highly significant relationship between car make and drive wheels.
2. FWD is the most common drive type, with 120 out of 205 vehicles having this configuration.
3. Honda (13) and Mitsubishi (13) exclusively use FWD.
4. Only a few brands like Audi (2) and Subaru (5) use 4WD, indicating its specialized application.

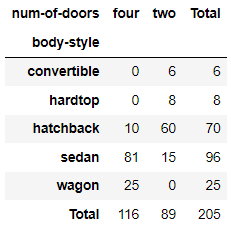
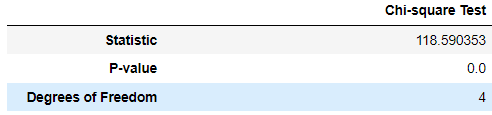
**Distribution of Engine-type based on Fuel-system:**

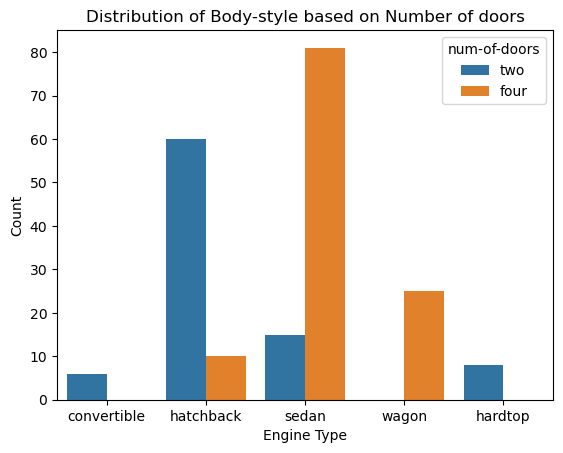




1. The Pearson Chi-Square value of 211.74 with a p-value of 0.000 indicates a statistically significant relationship between engine type and fuel system.
2. The most common combination is the ohc (overhead cam) engine type with the mpfi (multi-point fuel injection) fuel system.
3. The 4BBL (four-barrel carburettor) fuel system is very rare, with only 3 instances, all associated with the rotor engine type.
4. The ohcf (overhead camshaft with forced induction) engine type is primarily paired with 2bbl and mpfi systems, indicating specific fuel system preferences for this engine type.
5. The mpfi fuel system is the most frequently used across different engine types, with 94 instances.

**Distribution of Body-style based on Number of doors:**

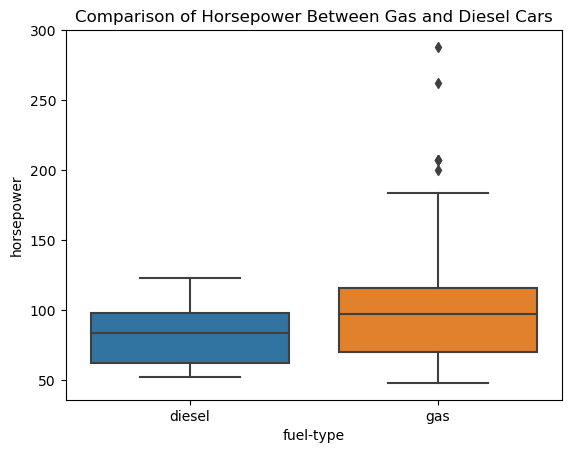
f



1. There are more cars with four doors (116) compared to those with two doors (89).
2. Most sedans have four doors (81 out of 96), indicating that sedans are typically designed with four doors to accommodate more passengers comfortably.
3. There are no convertibles or hardtops with four doors.
4. All wagons in the dataset have four doors (25 out of 25).

|  |  |  |
| --- | --- | --- |
| **t-Test: Two-Sample Assuming Equal Variances** | | |
|  |
|  |  |  |  |
|  | A screenshot of a graph  Description automatically generated |  |  |

**Comparison of Horsepower Between Gas and Diesel Cars:**



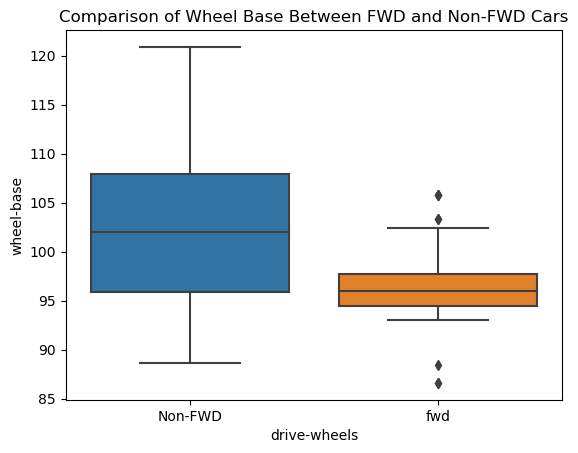
A white background with black text

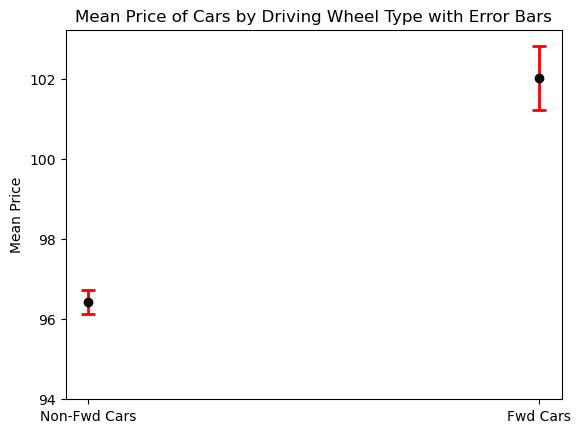
Description automatically generated

1. The p-value for the t-test is less than 0.05, indicating a significant difference in horsepower in the specified direction (gas > diesel).
2. The critical value for the t- test at the 0.05 significance level is 1.97. The t-statistic of 2.37 exceeds both critical values.
3. Based on the p-values and the t-statistic, we conclude that there is a statistically significant difference in the mean horsepower between gas and diesel cars.

**Comparison of Wheelbase Between Cars with Fwd wheel and Non Fwd Wheel:**

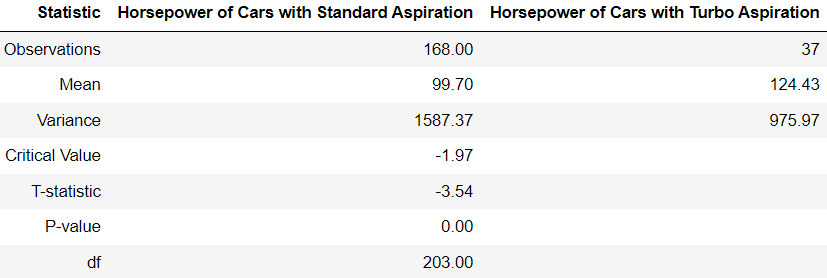
|  |  |  |
| --- | --- | --- |
| **t-Test: Two-Sample Assuming Unequal Variances** | | |
|  |  |  |

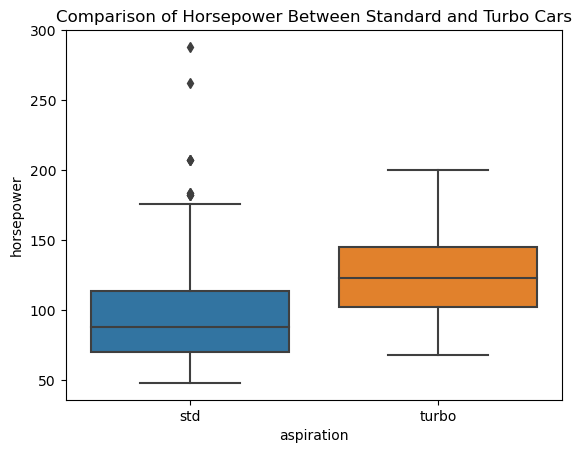


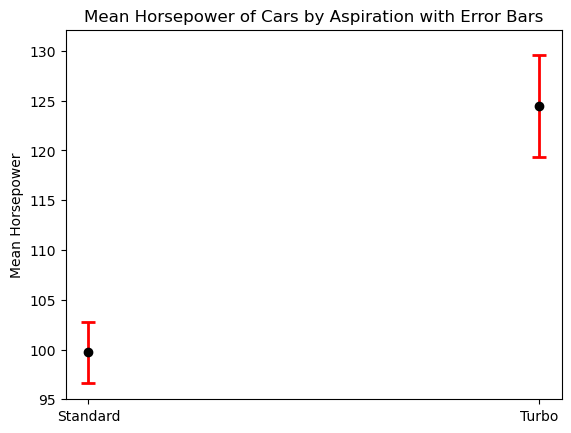


1. The two-tail p-value indicates that the probability of observing such a difference in means by chance is very low, suggesting a significant difference regardless of direction. This suggests a significant difference between the means of the two groups.
2. The significant difference in wheelbase suggests that vehicle manufacturers might design fwd and non-fwd vehicles with different wheel-base lengths due to functional or design considerations.

**Comparison of Horsepower Between Cars with Standard and Turbo Aspiration:**





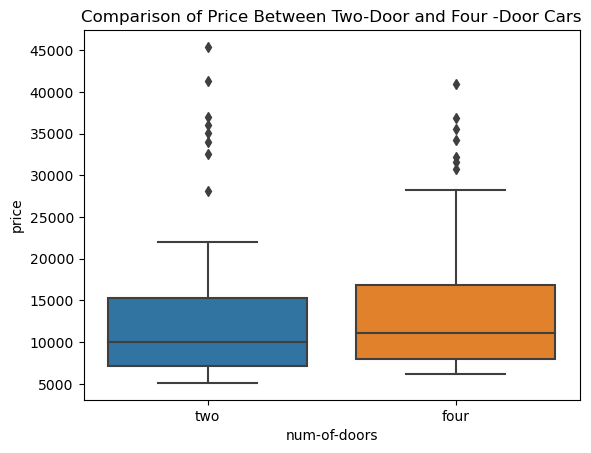


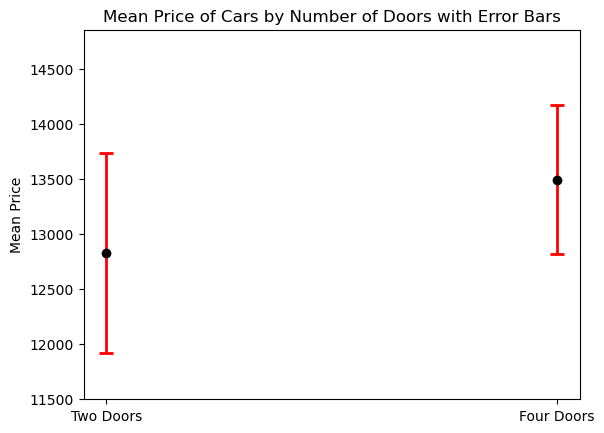
1. Cars with turbo aspiration have a higher mean horsepower (124.43) compared to cars with standard aspiration (99.70).
2. The t-statistic of -3.54 and p-value of 0.00 indicates a significant difference in mean horsepower between the two groups, with the negative value suggesting that the mean horsepower for standard aspiration is significantly lower than that for turbo aspiration.
3. The error bars indicate that there isn’t any overlap in the confidence intervals for the means of the two groups, reinforcing the significance of the difference in means.
4. The analysis shows that cars with turbo aspiration tend to have higher and more consistent horsepower compared to cars with standard aspiration, which have more variation in horsepower values.

|  |  |  |
| --- | --- | --- |
|  | | |
|  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

|  |  |  |
| --- | --- | --- |
| **t-Test: Two-Sample Assuming Unequal Variances** | | |
|  |  |  |

**Comparison of Price Between Two-Door and Four -Door Cars:**

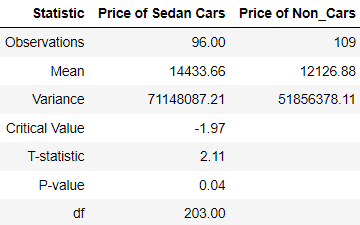




1. The variance in prices is higher for two-door cars compared to four-door cars, indicating more variability in prices among two-door cars.
2. The p-values is 0.55, respectively. These values indicate that there is no statistically significant difference in prices between four-door and two-door cars.
3. The small t statistic and non-significant p-values suggest that any difference in prices between the two types of cars, if present, is not substantial.
4. The sample size for two-door cars is 89 observations, while for four-door cars, it is 116 observations. The difference in sample sizes may influence the reliability of the results.

**Comparison of Price Between Sedan and Non-Sedan Cars:**

**t-Test: Two-Sample Assuming Unequal Variances**

****

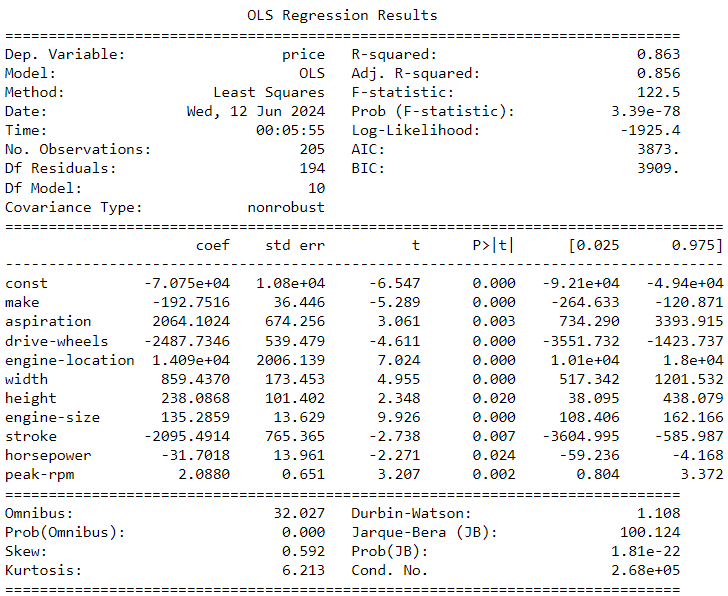


A screenshot of a computer

Description automatically generated

1. The mean price of sedan cars is significantly higher than that of non-sedan cars, indicating a notable price difference between these two categories.
2. The higher mean price and greater variability in sedan car prices suggest a more diverse market segmentation within the sedan car category, catering to different consumer preferences and budget ranges.
3. The t-statistic (2.11) and p-value (0.04) indicate that the difference in mean prices between sedan cars and non-sedan cars is statistically significant, implying that this price difference even though close is likely not due to random chance.
4. The boxplot shows that sedan cars have a higher median price and a wider interquartile range compared to non-sedan cars, indicating more price dispersion among sedans.
5. The error bars show the standard error around the mean prices, indicating no overlap reinforcing the significant difference in mean prices between the two categories.
6. The analysis shows that cars with sedan cars tend to have higher and more consistent Price compared to cars with other body styles.

**Identifying significant Features that Influence Price:**



A graph of a distribution plot

Description automatically generated with medium confidence

1. The model explains 86% of the variance in car prices, indicating a good fit for the data.
2. The ANOVA results are highly significant probability of F-statistic is very low, suggesting that the predictors collectively have a strong impact on car prices.
3. Stroke and Horsepower have negative coefficient, suggesting that higher values for this variable is associated with lower prices.
4. Width has the largest positive coefficient of 859, indicating a strong positive impact on car prices.
5. Height and Engine size have a substantial impact on price, with a coefficient of 238 and 135, indicating that larger cars with powerful engines tend to be priced higher.
6. Make and Drive wheels have negative coefficient of -192 and -2487, indicating non-Fwd cars made by Audi, BMW, Dodge, Honda, Jaguar, Mercedes-Benzes are expensive.
7. The regression equation can be expressed as:

Price = -7.075 + (-192.7516) \*Make + 2064.1024\*Aspiration + (-2487.7346)\*Drive-wheels + 1.409\*Engine-location + 859.4370\*Width + 238.0868\*Height + 135.2859\*Engine size + (-2095.4914)\*Stroke + (-31.7018)\*Horsepower + 2.0880\*Peak-rpm.

A screenshot of a phone

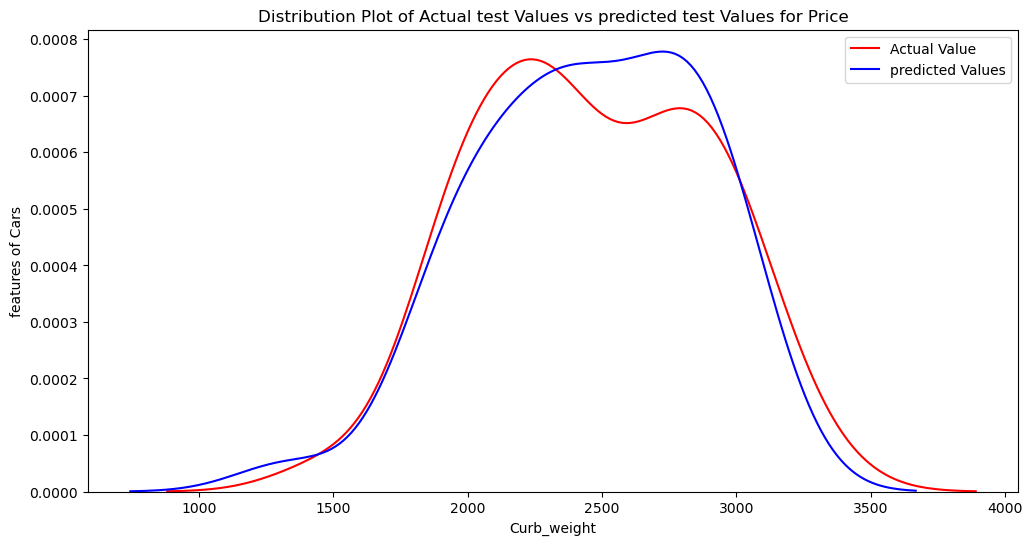
Description automatically generated

1. From the VIF table we cannot see any predictors with VIF more than 10, suggesting this model does not suffer from multicollinearity problem.
2. In the line plot we can see the difference between Actual and predicted values, the closeness of both lines indicating this a strong and reliable model. Also, the values are not perfectly matched, suggesting this model don’t have overfitting problem.

**Identifying significant Features that Influence Curb-weight:**

**A screenshot of a computer

Description automatically generated**



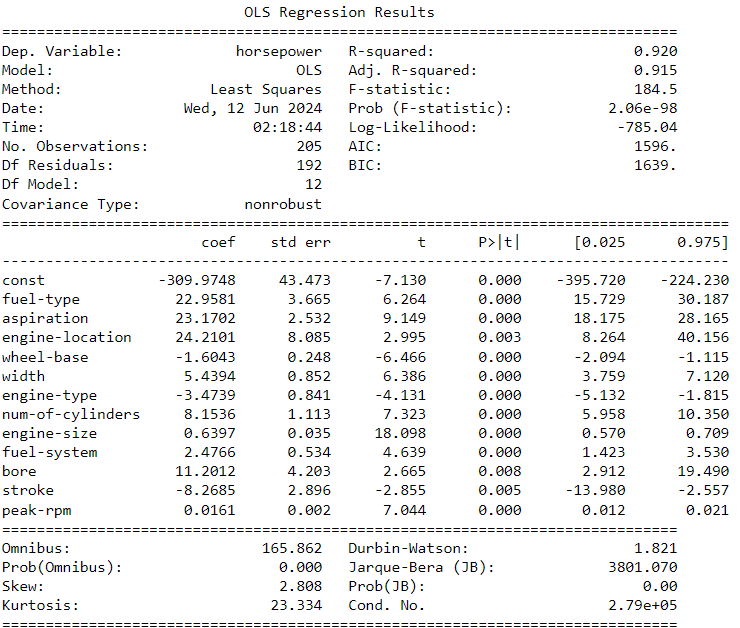
1. The model explains approximately 94% of the variance in curb weight, indicating a strong fit.
2. The ANOVA shows significant results probability of F-statistic is very low, suggesting that the predictors collectively have a strong impact on curb weight.
3. The intercept (constant) has a negative value, indicating that when all predictors are zero, the curb weight is expected to be lower.
4. Highway-mpg have negative coefficient, suggesting that higher values for this variable is associated with lower curb-weight.
5. Among these predictors, Width and Height have standardized coefficient of 53 and 27, indicating their relatively stronger impact on curb weight.
6. Fuel-type and Engine-location have negative coefficient of -283 and -297, indicating diesel cars with front engine are much heavier than gas cars with rear engine.
7. The regression equation based on the provided summary is:

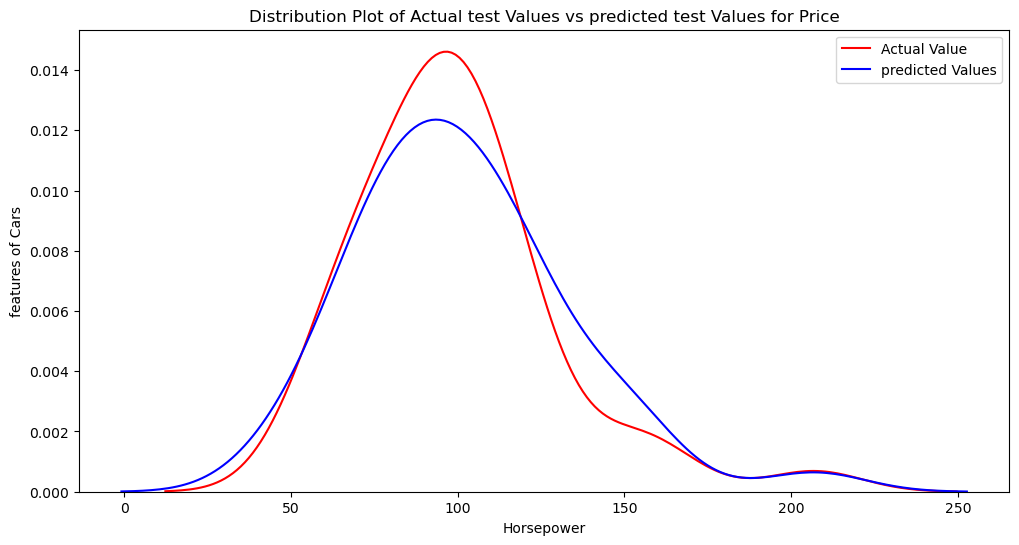
Curb Weight = -2140.1196 + (-283.3436)\*Fuel-type + (-156.2688)\*Drive-wheel + (-297.3058)\*Engine-location + 53.7393\*Width + 27.1596\*Height + 4.0624\*Engine-size + 1.5566\*Horsepower + (-20.5343)\*Highway-mpg.

1. Fuel-type, Engine-location and width appear to be particularly important predictors of curb weight, based on their coefficients and significance levels.
2. A screenshot of a phone

   Description automatically generatedFrom the VIF table we cannot see any predictors with VIF more than 10 with only horsepower having a VIF of 5, indicating this model does not suffer from multicollinearity problem.
3. In the line plot we can see the difference between Actual and predicted values, the closeness of both lines indicating this a strong and reliable model. Also, the values are not perfectly matched, suggesting this model don’t have overfitting problem.

**Identifying significant Features that Influence Horsepower:**

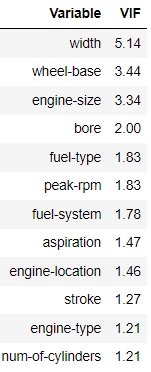
****



1. The model explains approximately 92% of the variance in curb weight, indicating a strong fit.
2. The ANOVA shows significant results probability of F-statistic is very low, suggesting that the predictors collectively have a strong impact on curb weight.
3. The intercept (constant) has a negative value, indicating that when all predictors are zero, the curb weight is expected to be lower.
4. Stroke have negative coefficient, suggesting that higher values for this variable is associated with lower Horsepower.
5. Among these predictors, bore has standardized coefficient of 11, indicating their relatively stronger impact on Horsepower.
6. Wheelbase and Engine-type have negative coefficient of -1.60 and -3.47, indicating non-Fwd cars with ‘Dohc’,’Dohcv’ or ’L’ type engines have much more Horsepower than Fwd cars with ‘Rotor’,’Ohcv’ or ’Ohcf’ type engines.
7. The regression equation based on the provided summary is:

Curb Weight = -309.9748 + 22.9581\*Fuel-type + 23.1702\*Aspiration + 24.2101\*Engine-location + (-1.6043)\*Wheel-base + 5.4394\*Width + (-3.4739)\*Engine-type + 8.1536\*Num-of-Cylinders + 0.6397\*Engine-size + 2.4766\*Fuel-system + 11.2012\*Bore + -8.2685\*Stroke + 0.0161\*Peak-rpm.

1. Fuel-type, Engine-location and Aspiration appear to be particularly important predictors of Horsepower, based on their coefficients and significance levels. They have highest standardized coefficient of 22.9,24.2 and 23.1, reassuring their relatively stronger impact on curb weight and indicating rear engine gas cars with turbo aspiration have more horsepower than front engine diesel cars with standard aspiration.
2. From the VIF table we cannot see any predictors with VIF more than 10 with only width having a VIF of 5, indicating this model does not suffer from multicollinearity problem.
3. In the line plot we can see the difference between Actual and predicted values, the closeness of both lines indicating this a strong and reliable model. Also, the values are not perfectly matched, suggesting this model don’t have overfitting problem.



**Conclusion:**

Understanding that engine size and performance metrics (like peak RPM) significantly impact car prices helps manufacturers tailor their models to consumer preferences, potentially boosting sales.

The higher horsepower in gas cars suggests they may be designed for performance compared to diesel cars, which might focus on fuel efficiency. Brands like Toyota and Honda's broad range of models show a strategy to cover multiple market segments, enhancing their market presence and catering to diverse customer bases. Brands like Jaguar, Mercedes-Benz, and Porsche with limited but high-end offerings focus on brand prestige and performance, attracting affluent customers willing to pay premium prices.

The trade-offs between engine size, horsepower, and fuel efficiency highlight the challenges in balancing performance with fuel economy. Manufacturers must innovate to provide powerful yet efficient engines. Given the negative impact of curb weight on fuel efficiency, the industry may increasingly adopt lightweight materials to reduce weight without compromising safety or performance. Lighter vehicles can be cheaper to produce and transport, impacting overall manufacturing costs.

The strong correlations between length, width, and wheelbase suggest that changes in one dimension will necessitate adjustments in others, influencing overall vehicle design and manufacturing processes. Vehicles with larger dimensions and more powerful engines might be more appealing to consumers who value space and performance.

High horsepower and larger engines contribute to curb weight but are crucial for performance. Manufacturers can use these findings to strike a balance between performance and weight. Insights from the correlations between engine types and fuel systems can guide manufacturers improving fuel efficiency and performance.

Our analysis reveals that fuel type, engine location, and aspiration are particularly significant predictors of horsepower, indicating that rear-engine gas cars with turbo aspiration have significantly higher horsepower compared to front-engine diesel cars with standard aspiration. Furthermore, wheelbase and engine type influence horsepower, with non-FWD cars featuring 'Dohc', 'Dohcv', or 'L' type engines outperforming FWD cars with 'Rotor', 'Ohcv', or 'Ohcf' type engines. This analysis guide manufacturers in optimizing engine configurations and vehicle designs to enhance performance, catering to market demands for higher horsepower vehicles.

In summary, this analysis provides valuable insights into pricing dynamics, vehicle dimensions, vehicle specifications, and market trends within the automobile industry. Ultimately, aligning pricing and vehicle specifications with consumer expectations and preferences can enhance customer satisfaction and loyalty, driving long-term success.