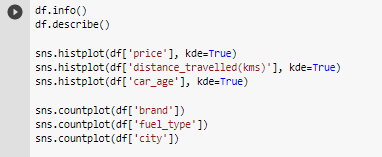
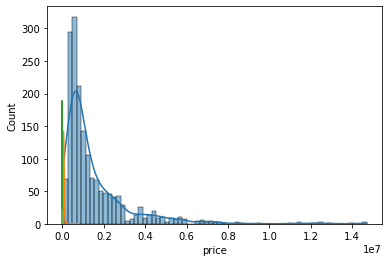
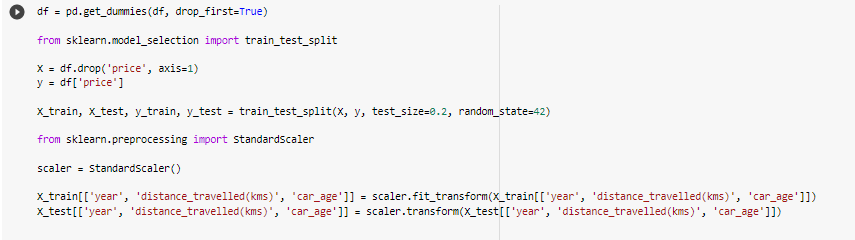
**USED CARS PRICE PREDICTION**

In today's world, the used car market is becoming more and more popular due to its affordability and accessibility. The pandemic has further accelerated this trend as more people opt for personal mobility to avoid public transportation. With the growth of the used car market, there is an increasing demand for predictive models that can accurately determine the selling price of a used car.

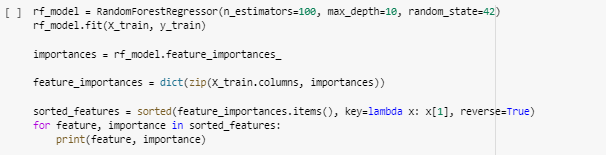


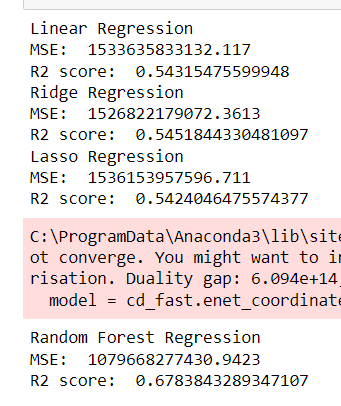


In the EDA phase, univariate analysis is conducted to understand the distribution of each feature in isolation. This is done using histograms and box plots. Multivariate analysis is then conducted to determine correlations between features. This is done using scatter plots and correlation matrices. For example, we can plot a scatter plot of the distance\_travelled(kms) feature against the price feature to determine if there is any correlation between the distance travelled by the car and its selling price. Correlation matrices are used to visualize the correlation between all pairs of features in the dataset. This helps to identify any strong correlations between features and determine which features have a significant impact on the selling price of the car.



Once the EDA is complete, the next step is to layout the regression experiment space. This involves selecting the appropriate regression algorithm and setting the hyperparameters. The next step is to train the model using the training data and evaluate its performance using the test data. The performance of the model can be evaluated using metrics such as mean squared error, root mean squared error, and R-squared. These metrics provide an indication of how well the model is performing in predicting the selling price of the used car.

Finally, the most important features determining the price of used cars can be listed out. This is done by analyzing the coefficients of the linear regression model. The coefficients represent the magnitude and direction of the impact of each feature on the selling price of the car. Features with higher coefficients have a greater impact on the selling price of the car.



It has been observed from the above figure that the all the accuracy are medium; among all the accuracy the accuracy regarding the random forest regression is high. Thus, based on the technical work, it can be stated that the random forest regression is the best method for conducting the task. Lastly, it can be concluded that building a predictive model to determine the selling price of a used car is an important task in today's digital age. It can also be concluded the among all the applied methods, random forest regression is one of the most effective methods. By leveraging data science and machine learning techniques, we can accurately predict the selling price of a used car and identify the most important features that determine the price. This provides valuable insights to car resellers who are looking to optimize their revenue and tap into the growing market for used cars.