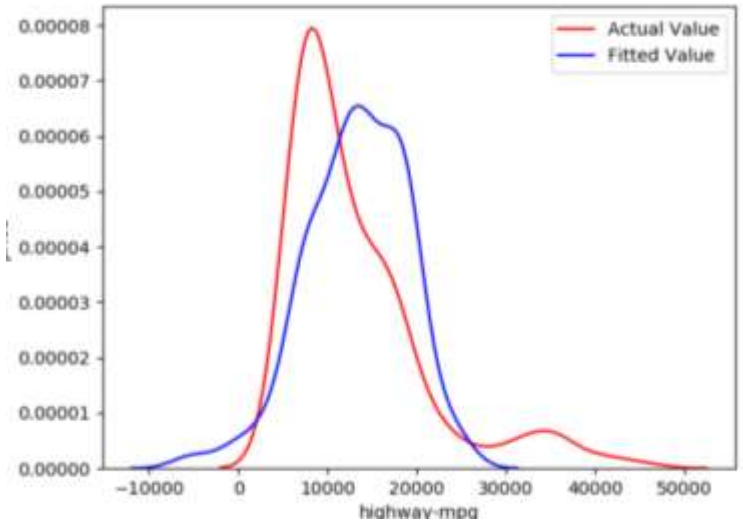



# AUTOMOBILE COMPANY-YEARLY GROWTH ANALYSIS

---

TNP PROJECT

SUBMITTED BY: SHIVA BHADORIYA

ENROLLMENT NUMBER: 0176CS181151

# Data Analysis and visualization to predict car prices based on used car prices data set

---

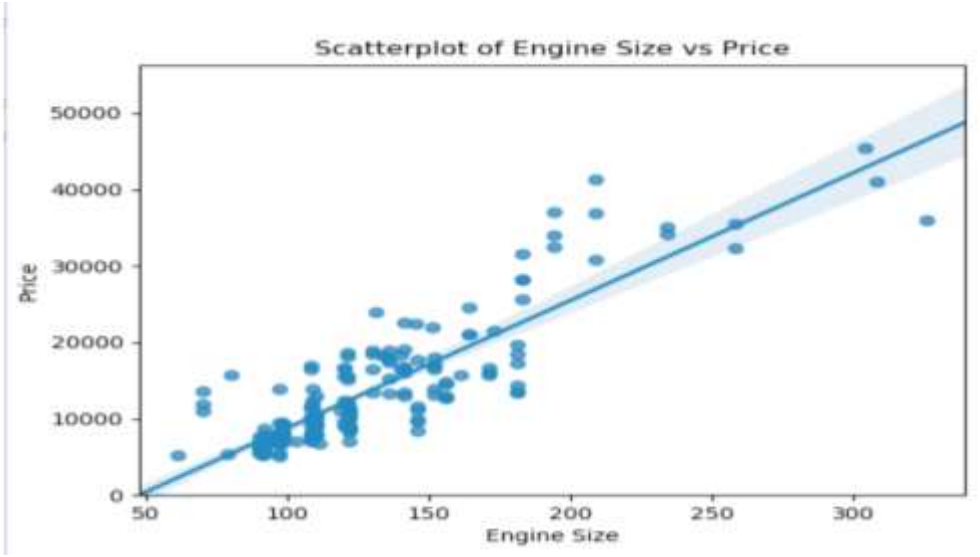
IN THIS PROJECT WE ARE TRYING TO ANALYZE AND VISUALYZE THE USED CAR FROM THE DATA SET AVAILABLE ONLINE IN THE FORM OF CLOUD FROM UCSD UNIVERSITY AVAILBLE AT

https://archive.ics.uci.edu/ml/machine-learning-databases/autos/imports-85.data

IN ORDER TO PREDICT THE MOST PROBABLE CAR PRICES

# LIBRARIES / FRAMEWORK USED:

1. JUPYTER NOTEBOOK
2. NUMPY
3. PANDAS
4. SEABORN
5. MATPLOTLIB
6. REGRESSION AND DISTRIBUTION PLOTS



# Components:

## It is divided into four sections

### 1. DATA OBSERVATION AND WRANGLING:

1. Pre processing data in python
2. Dealing missing values
3. Data formatting
4. Binning

### 2. DATA STANDARDISATION

1. Fitting data in the required format.
2. Labelling it in correct form under correct heading.
3. Standardising correct values for correct comparison.

# Components:

## 3. DATA NORMALISATION

1. Data normalization is the process of structuring a relational database in accordance with a series of so called normal forms in order to reduce data redundancy and improve data integrity.

2.Users can properly utilize the database for further queries and analysis.

## 4. EXPLORATORY DATA ANALYSIS:

1. Thorough analysis and visualization of data obtained by filtering all the anamalies.

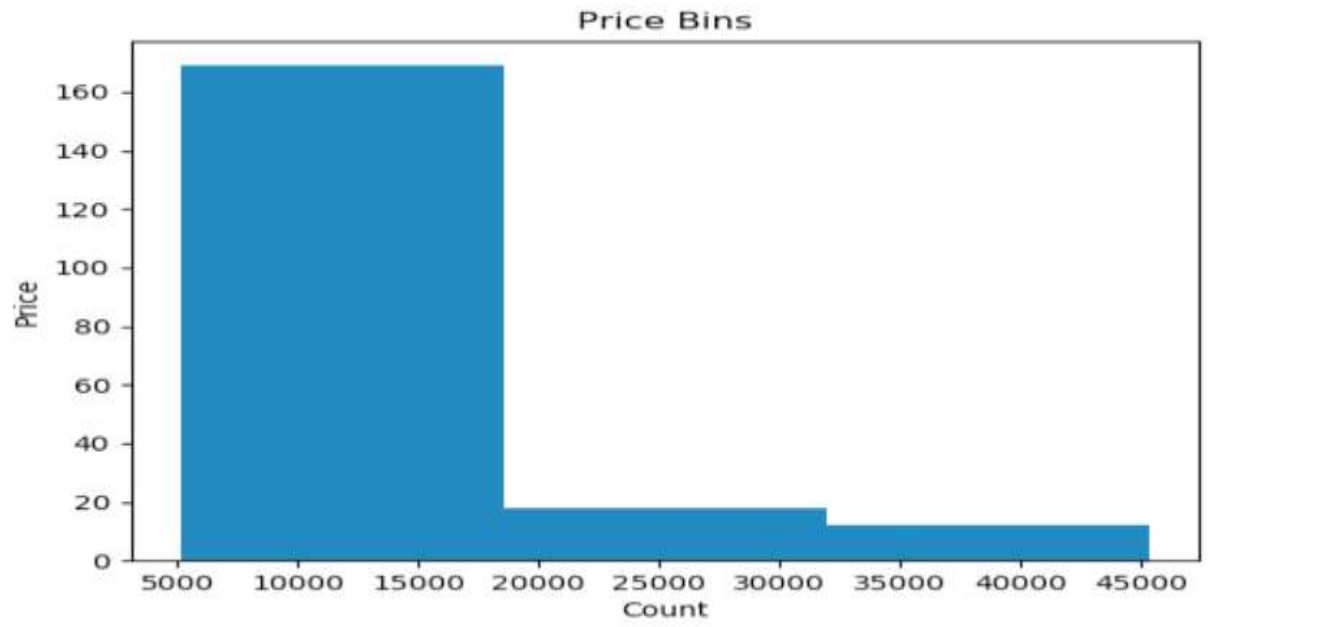
2. Representation of data in interactive way

3. Use of tables, Graphs and animation to visualize and understand the data

4. This is what we are looking forward to do.

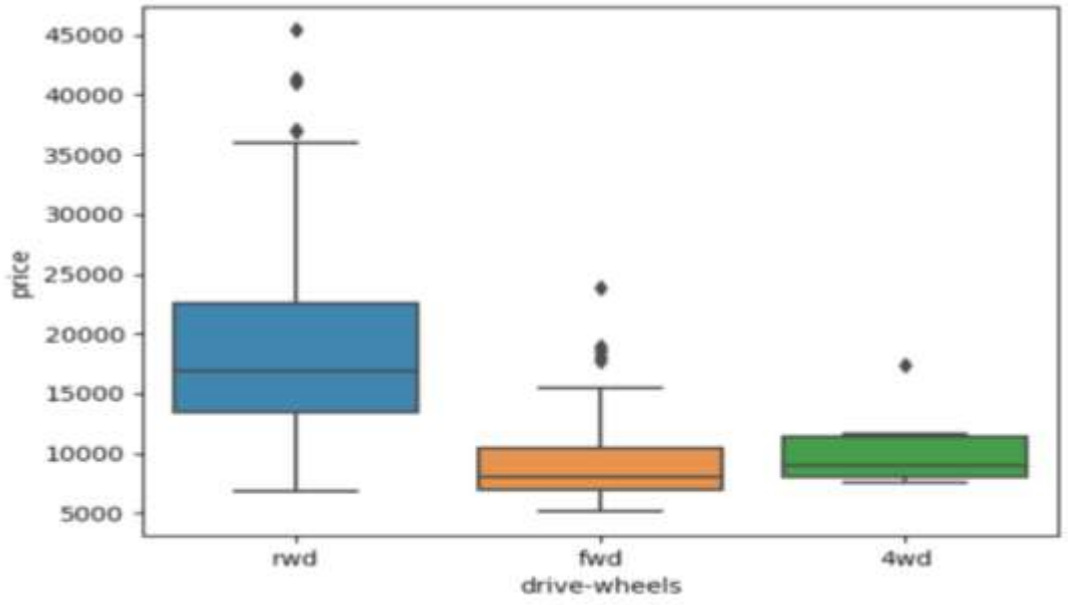
# Analysis and Visualization:

1. Histograms representing Binned prices in Low, Medium, High

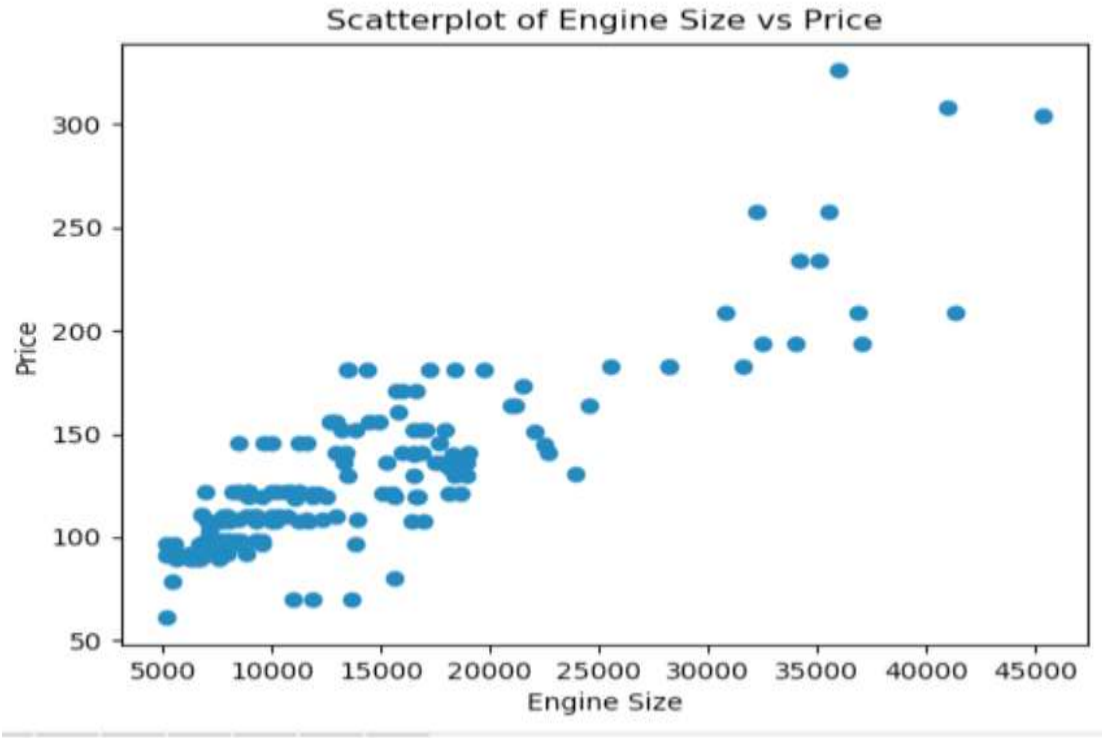


# Analysis and Visualization:

Boxplots representing effect of wheel frive with prices:

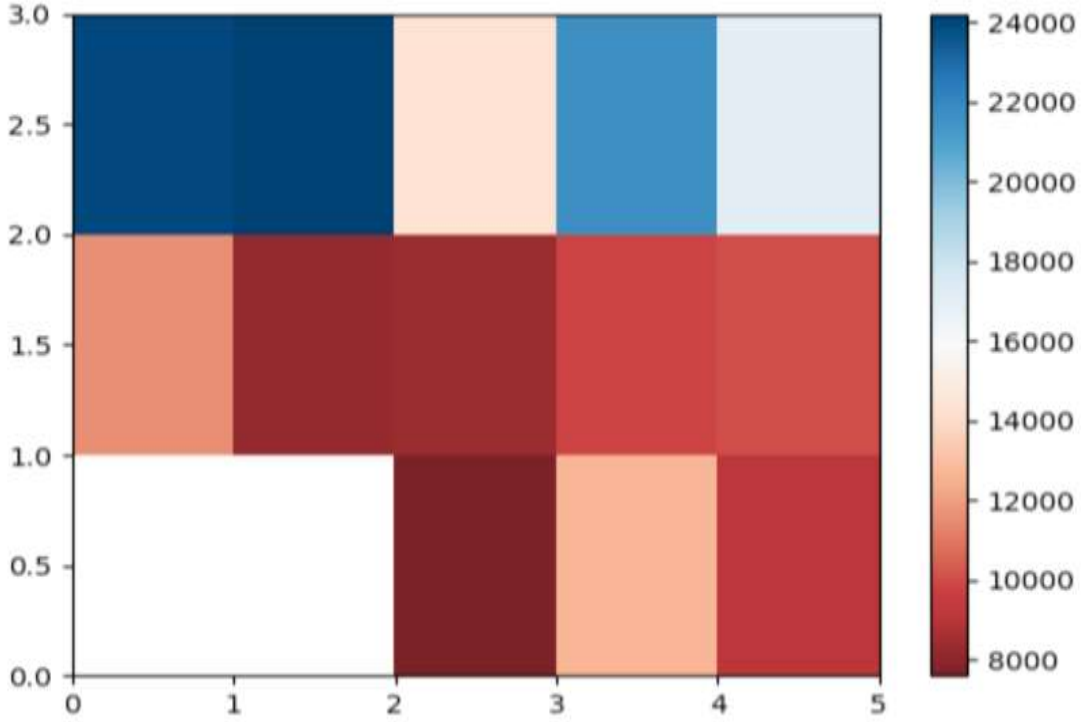



Scatter plot of price vs engine size:

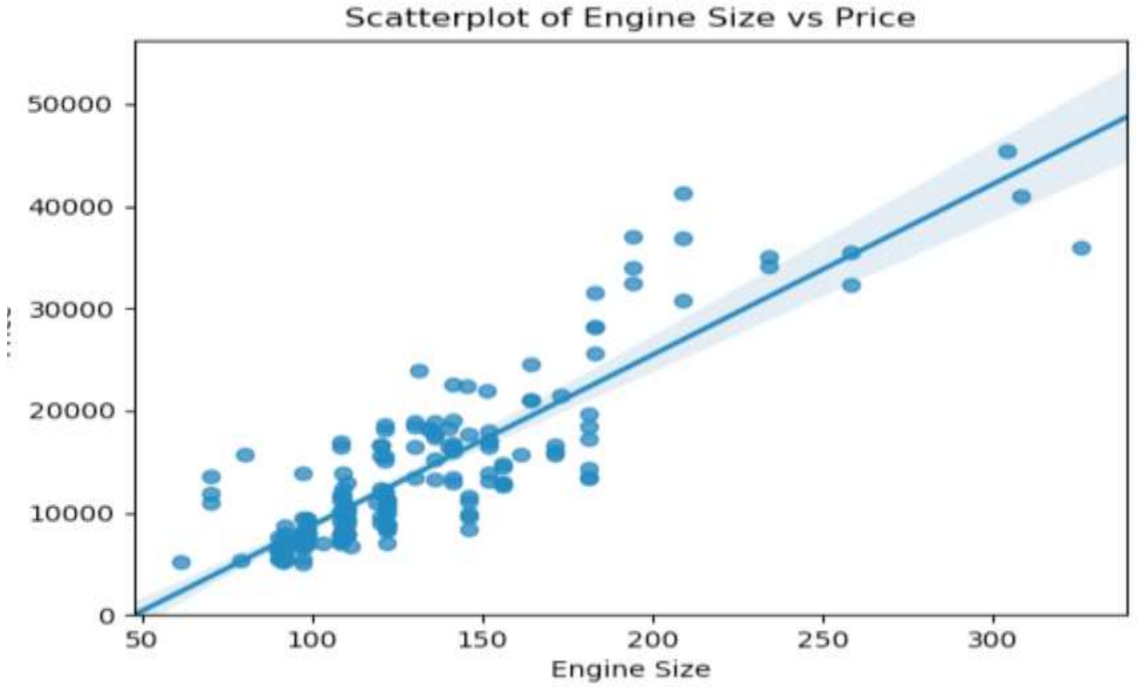


# Analysis and Visualization:

HeatMap with wheel drive in y axis and body style in x axis

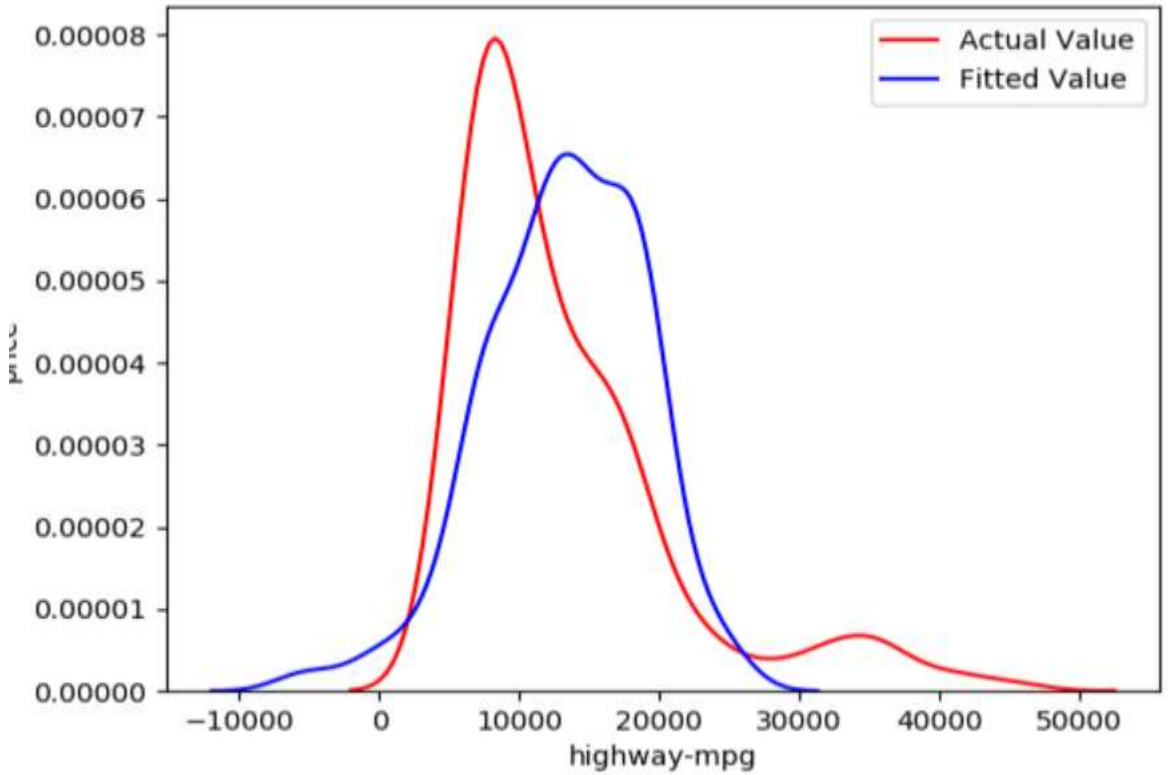



Positive Linear Relationship between engine size and price

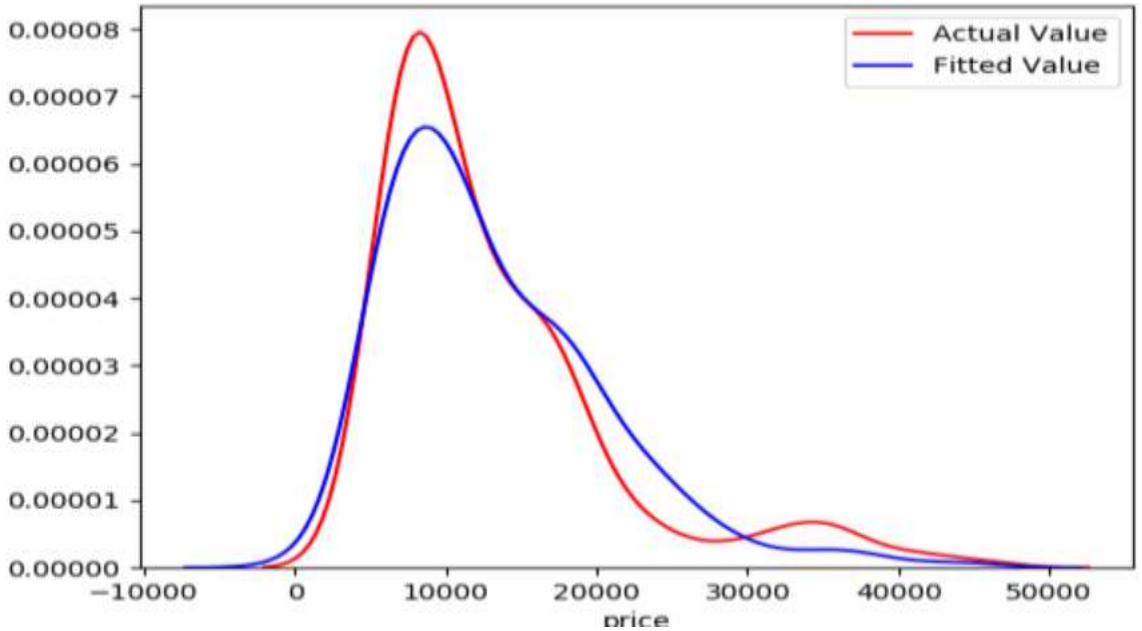


# Analysis and Visualization:

SIMPLE LINEAR REGRESSION PLOT:




MULTIPLE LINEAR REGRESSION PLOT:



# Conclusion:

The distribution plot of Linear Regression and Multiple Regression technique shows how the model predicts the prices of automobiles based on “Horsepower”, “Curb-weight”, “Engine-size”, and “Highway-mpg”.

Comparing these three models, we conclude that the MLR model is the best model to be able to predict price from our dataset. This result makes sense, since we have 27 variables in total, and we know that more than one of those variables are potential predictors of the final car price.

# THANK YOU

TITLE: AUTOMOBILE COMPANY YEARLY- GROWTH ANALYSIS

SUBMITTED BY: SHIVA BHADORIYA

ENROLLMENT NUMBER: 0176CS181151

CLASS ROLL NUMBER: 147

SECTION: C