Assignment- Data Visualisation

Data Analysis: Derive Insights

Use dataset “CarPrice\_Assignment\_csv” to perform below mentioned tasks:

1. Import following libraries:

a. Numpy and Pandas

b. Matplotlib.pyplot

c. Seaborn

2. Import the dataset using the following command in your Jupyter Notebook:

pd.read\_csv(‘Filename.csv’)

3. Find out unique columns present in the dataset. After you have listed down the unique columns, review the same and highlight any anomaly

4. Check for duplicate values in the dataset

5. Create a Histogram, KDE plot and box plot for variable ‘price’

6. Create Histograms for Car Company, Fuel Type, Car Type, Symboling, Engine Type and put your inference (insights) for the same

7. Create bar plots for the following:

a. Company name vs Average Price

b. Fuel type vs Average Price

c. Car Type vs Average Price

8. Create Histograms for rest of the variables (excluding the ones that have been mentioned above)

9. Create scatter plot and provide insights for the following:

a. Car length vs price

b. Car width vs price

c. Car height vs price

d. Car weight vs price

10. Create scatter plot for the following variables w.r.t. price as provide insights:

✔ Engine size

✔ bore ratio

✔ stroke

✔ compression ratio

✔ horse power

✔ peak rpm

✔ wheel base

✔ city mpg

✔ highway mpg

11. Create scatter plot using the following code (please note the code is incomplete and you will have to complete the same for the entire plot to be visualized) and write down your insights derived

sns.scatterplot(x=cars['fueleconomy'],y=cars['price'],hue=cars['drivewheel'])

12. After above exercise for visualization, can you list down important variable which will impact price of the car.