Findings from the data file

File Name : used\_cars\_prices.csv Findings:

1. Though there are lot of metrics to consider while choosing the used car, the historical data is not well organized for each brand vehicle
2. Initial phase : Data Cleansing as part of pre-processing stebfor analysis -
   1. Replacing the missing numerical values with 0 using pandas method   
      fillna()
   2. Removing the duplicate rows based on all columns which helps to   
      eliminate redundant data which may skew the results. Method used is   
      drop\_duplicates() with parameter *inplace=True*
   3. Converting the data type for the below columns to int type from   
      floating\decimal  
      1. *Year*2. *Seller Reviews Count* 3. *Car Reviews Count* 4. *Mileage*
   4. Saved the new file with name ‘cleaned\_used\_cars\_prices\_datafile.csv’
3. Used set\_index() method over the columns - *Year & Manufacturer* which helps to access and analyze the data efficiently.
4. Prepared Box Plot from seaborn library to understand the price range of the first 10 cars. As per the outcome, manufacturers - *Toyota & Lexus* looks best pick.
5. Prepared Histogram from seaborn understand the price range of the first 10 cars. Out of 10 cars, *Jeep* is the costliest and *FIAT* is the cheapest car.
6. Prepared Line Plot from seaborn library to find the price trend of the top 3 unique cars - *Hyundai,MINI & Mitsubishi*
7. Using linear regression, predict the car price value based on one feature - [year of manufacture] , 2 features - [year of manufacture & Mileage] and 3 features - [year of manufacture, Mileage & customer review count]
8. The below metrics are evaluated - Mean Squared Error & Mean Absolute Error
9. Mean squared error is low considering single feature instead of multi feature consideration. Might be due to data!
10. Whether customer will purchase an used car using logistic regression? If yes, what all metrics to consider.
11. The below metrics are evaluated - accuracy, precision and recall
12. All the above metrics resulted 1.0