**Assignment -2**

**Name: Sowmya Patlolla**

**Id: 16338256**

**Step 1**: Collected the data from the following link.

Data link: <https://app.box.com/s/jm6pw202asu4xd3uypwtry2rqk691y1i>

**Step 2**: Loading the csv data file into the jupyter notebook and fetched the data from the file.

A screenshot of a computer

Description automatically generated

**Step 3**: Checking or finding the missing values from all the columns in the dataset.

A screenshot of a computer

Description automatically generated

**Step 4**: Filling the missing values in the columns with the mean values of the numerical column data also dropping the unnecessary column which has more no. of missing values.

A screenshot of a computer

Description automatically generated

**Step 5**: Removing the units and keeping only the numerical values in the columns.

A screenshot of a computer

Description automatically generated

**Step 6**: Converting the categorical variables into numerical one - hot encoded values.

A screenshot of a computer

Description automatically generated

**Step 7**: Preprocessed the dataset and removed all the missing values from the dataset.

A screenshot of a computer

Description automatically generated

**Step 8**: Created a new column named “Current Age” it provides the current age of the car.

A screenshot of a computer

Description automatically generated

**Step 9**: Created another new column named “Mileage\_per\_Horsepower” which provides the mileage of the car per horsepower.

A screenshot of a computer

Description automatically generated