SC1015 Car Project (Overview):

Technical Components:

1. 10% for coming up with your own problem definition based on a dataset
2. 20% for the use of machine learning techniques to solve specific problem

Analytical Components:

1. 20% for exploratory data analysis/visualization to gather relevant insights
2. 20% for the presentation of data-driven insights and the recommendations

Supporting Components:

1. 10% for data preparation and cleaning to suit the problem of your choice
2. 10% for the quality of your final team presentation and overall impressions

Note: 10% for learning something new and doing something beyond this course

**Problem Statement:**

Can the features of a car predict it’s MSRP?

Data Set: <https://www.kaggle.com/CooperUnion/cardataset?select=data.csv>

Exploratory Analysis(Used):

* Linear Regression
* Classification tree
  + Task: To explore in scikit-earn existing model to fit the project.
* Explore train-test split for the data set(such as 80-20 split)
* Data Cleaning(Content)

Visualisation:

* Boxplot
* Correlation
* Data Cleaning (Presentation)
* Bar Plot

16th March: Explore other Scikit-learn model(Different models to be used)