Problem Statement:

There is a huge demand of refurbished cars in the Indian Market today. As sales of new cars have slowed down in the recent past, the refurbished car market has continued to grow over the past year and is larger than the new car market now.

Consider this: In 2018-19, the sales of refurbished cars which were previously owned by someone has increased steadily, and it is currently estimated to be approximately 1.3 times the sales of new cars. There are multiple reasons for this shift. However, the key reason is the increase in total cost of ownership of new cars which includes taxes, insurances.

There is a slowdown in new car sales and that could mean that the demand is shifting towards the pre-owned market.

**The goal here is to predict the Price of a refurbished car based on the variables provided in the data set.**

Working with Data

Data has been split into two groups and provided in the module:

**Training set :** The training set is used to build your machine learning model. For the training set, we provide the price of a car (also known as the variable Price) for each instance.

**Test set :** The test set should be used to see how well your model performs on unseen data. For the test set, it is your job to predict the price of the car (Price) for each instance.

**Data description:**

|  |  |
| --- | --- |
| **Id** | Car ID |
| **Maker** | Maker of the car |
| **model** | Model of the car |
| **Location** | Location of the Car owner |
| **Distance** | Distance travelled |
| **Owner Type** | First hand, second hand ... |
| **manufacture\_year** | Year of manufacture |
| **Age of car** | Age of car |
| **engine\_displacement** | Displacement in cc |
| **engine\_power** | Engine power in bhp |
| **body\_type** | Type of body |
| **Vroom Audit Rating** | Vroom rating |
| **transmission** | Manual or Automatic |
| **door\_count** | Door count |
| **seat\_count** | Seat count |
| **fuel\_type** | type of fuel |
| **Price** | Price in rupees |