**Name: Shubham Dhadage**

**Subject : Linear Models**

**Problem Statement**: The aim of this study is to develop a model that can accurately predict the price of a car based on its features such as model, year of manufacture, fuel type, and transmission type. Linear regression will be used as the primary statistical method to build this model. Because there are only two variables which have high correlation.

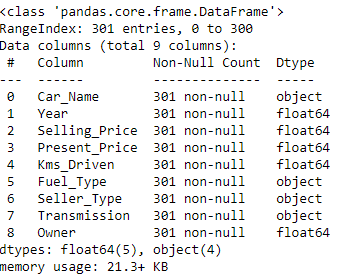
The goal of this study is to help car buyers and sellers make informed decisions by providing them with an accurate prediction of the fair market value of a car. This model will be trained on a dataset of historical car sales data, and the accuracy of the model will be evaluated using statistical metrics such as R-squared and mean squared error.

The data taken for this analysis is secondary data type. Which is collected for internet.

Following are some steps we done while analysis.

**Step 1:**

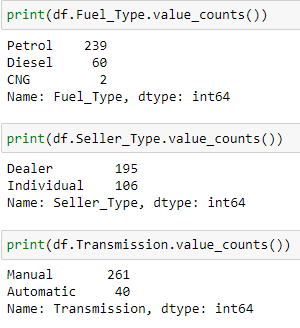
Importing libraries , loading the data set using df=pd.excel\_excel

Checking the information in data set using df.info()

Checking for the missing values and describing dataset 

**Step2:**

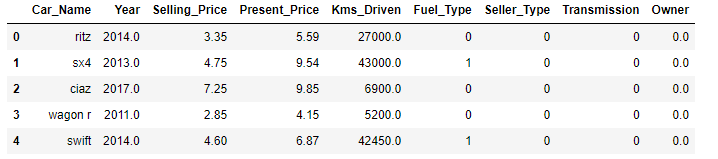
Taking count of data points in data set



In data set the the Fuel type is written as “Petrol”,”Diesel”,”CNG”

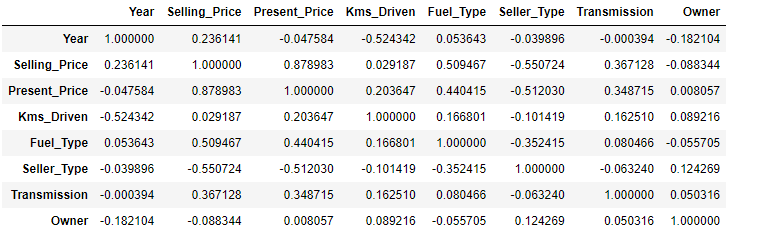
So we replace the variable into numeric form such as petrol =0,diesel=1 and CNG=2

Same with seller type and Transmission.



**Step:3**

Finding the correlation



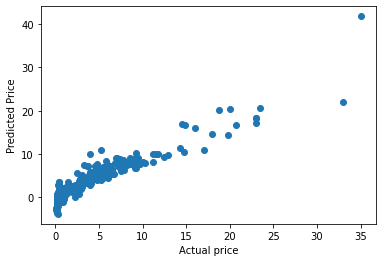
The correlation between the two variable i.e Selling\_Price and Present\_price have high correlation

And the correlation is equal to 0.878983.

Here two variable and high correlation means we have to do linear regression

Step 4:

Fitting the linear regression model



As we see that points does not show a pattern ,we can say that the linear model is appropriate for modelling the data

