**Used Cars dataset**

**Problem Statement**: To build a machine learning model that will predict selling price of the cars.

**Data Set Description**: This dataset contains information about used cars. Use the independent variables to predict selling price of the cars. Variable names are self-explanatory:

Below is the list of variables in the data:

* name
* year
* **selling\_price**
* km\_driven
* fuel
* seller\_type
* transmission
* owner
* mileage
* engine
* max\_power
* torque
* seats

**Expectations**

1. Perform Data cleaning – remove redundant variables, treat missing value and outliers.
2. Perform a detailed exploratory data analysis (EDA) on the dataset to understand variables and their interactions (relationship with target variable)
3. Do feature engineering and bring external features/variables to make the model more robust.
4. Build regression model using different algorithms to predict selling\_price of the used cars.
5. Compare performance of different models using best model evaluation metrics (**rmse**) and choose the best regression model for this problem.