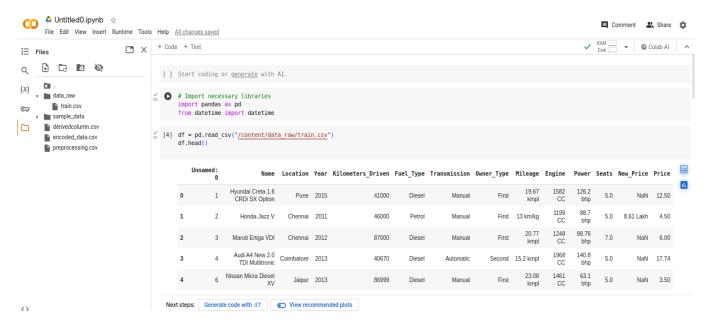
Student Name: Narendra Mamilla

Student ID : 16351892

Assignment - 2

1. Loading the train.csv file into google collab



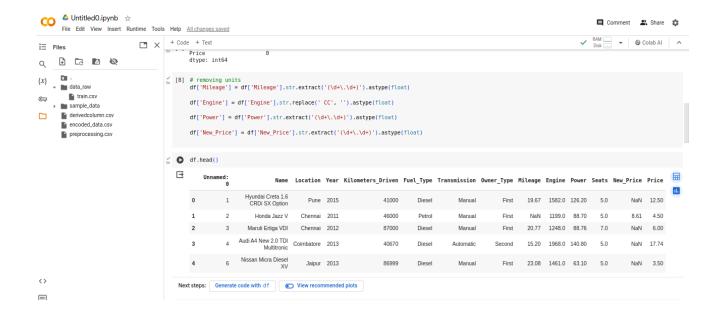
2. Identifying the missing values

```
# Identify missing values
    missing_values = df.isnull().sum()
    print("Missing values in each column:")
    print(missing values)

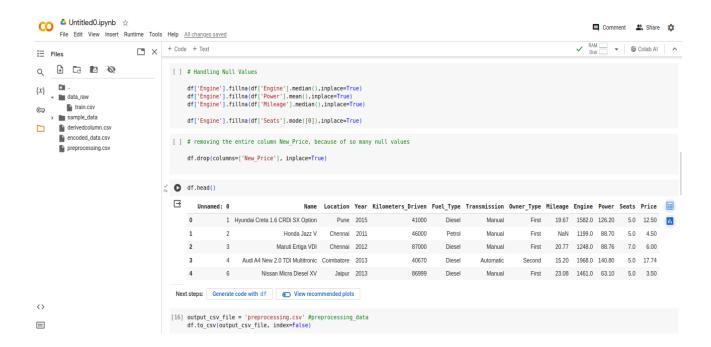
→ Missing values in each column:

    Unnamed: 0
                           0
    Name
                           Θ
    Location
    Year
    Kilometers Driven
    Fuel Type
                           0
    Transmission
                           0
    Owner Type
                           0
    Mileage
    Engine
                          36
    Power
                          36
    Seats
                          38
    New Price
                        5032
    Price
    dtype: int64
```

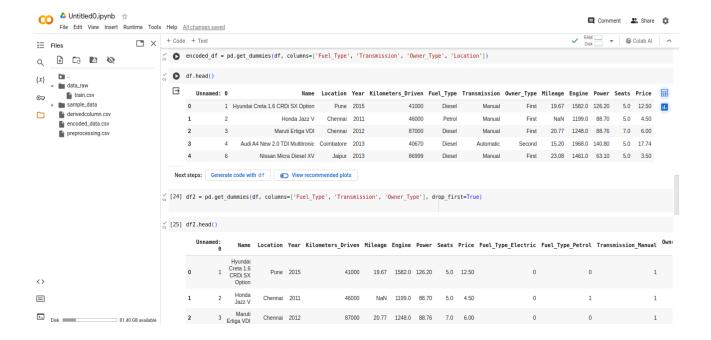
3. Removing the units from Power, Mileage, Engine and New Price



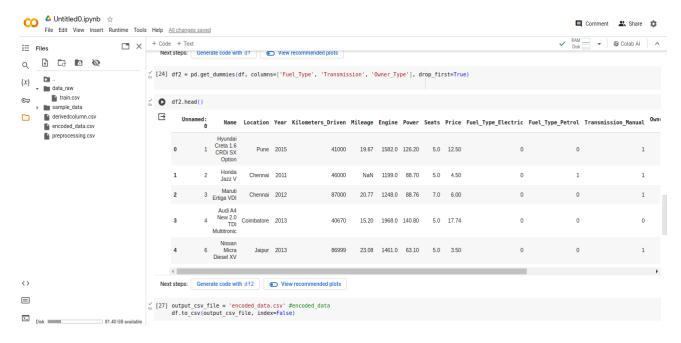
4. Handling the Null values by removing the entire column "New_Price" because it contains so many null values.



5. Encoding the columns Fuel Types, Transmission, Owner Type and Location to numerical by using 1 hot encoding.



6. And then Downloading the encoded csv file.



7. Finding the current age of the vehicle to derive the new column "Mileage per year".



8. Deriving the mileage per year and then downloading the csv file.

