In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
data=pd.read_csv("/home/placement/Downloads/fiat500.csv") #reading the data
data.describe()
```

Out[2]:

| | ID | engine_power | age_in_days | km | previous_owners | lat | lon | price |
|---|---|---|---|---|---|---|---|---|
| **count** | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 |
| **mean** | 769.500000 | 51.904421 | 1650.980494 | 53396.011704 | 1.123537 | 43.541361 | 11.563428 | 8576.003901 |
| **std** | 444.126671 | 3.988023 | 1289.522278 | 40046.830723 | 0.416423 | 2.133518 | 2.328190 | 1939.958641 |
| **min** | 1.000000 | 51.000000 | 366.000000 | 1232.000000 | 1.000000 | 36.855839 | 7.245400 | 2500.000000 |
| **25%** | 385.250000 | 51.000000 | 670.000000 | 20006.250000 | 1.000000 | 41.802990 | 9.505090 | 7122.500000 |
| **50%** | 769.500000 | 51.000000 | 1035.000000 | 39031.000000 | 1.000000 | 44.394096 | 11.869260 | 9000.000000 |
| **75%** | 1153.750000 | 51.000000 | 2616.000000 | 79667.750000 | 1.000000 | 45.467960 | 12.769040 | 10000.000000 |
| **max** | 1538.000000 | 77.000000 | 4658.000000 | 235000.000000 | 4.000000 | 46.795612 | 18.365520 | 11100.000000 |

In [3]:
```
data1=data.drop(['lat','ID'],axis=1)
```

In [4]:
```
data2=data1.drop('lon',axis=1)
```

In [5]:
```
data2['model']=data['model'].map({'lounge':1,'pop':2,'sport':3}) #making model values string to number
```

In [6]:

```
data2
```

Out[6]:

|  | model | engine_power | age_in_days | km | previous_owners | price |
|---|---|---|---|---|---|---|
| **0** | 1 | 51 | 882 | 25000 | 1 | 8900 |
| **1** | 2 | 51 | 1186 | 32500 | 1 | 8800 |
| **2** | 3 | 74 | 4658 | 142228 | 1 | 4200 |
| **3** | 1 | 51 | 2739 | 160000 | 1 | 6000 |
| **4** | 2 | 73 | 3074 | 106880 | 1 | 5700 |
| **...** | ... | ... | ... | ... | ... | ... |
| **1533** | 3 | 51 | 3712 | 115280 | 1 | 5200 |
| **1534** | 1 | 74 | 3835 | 112000 | 1 | 4600 |
| **1535** | 2 | 51 | 2223 | 60457 | 1 | 7500 |
| **1536** | 1 | 51 | 2557 | 80750 | 1 | 5990 |
| **1537** | 2 | 51 | 1766 | 54276 | 1 | 7900 |

1538 rows × 6 columns

In [7]:

```
y=data2['price']
x=data2.drop('price',axis=1) # droping the price column
```

In [8]:

```
x
```

Out[8]:

```
      model  engine_power  age_in_days      km  previous_owners
   0      1            51          882   25000                1
   1      2            51         1186   32500                1
   2      3            74         4658  142228                1
   3      1            51         2739  160000                1
   4      2            73         3074  106880                1
 ...    ...           ...          ...     ...              ...
1533      3            51         3712  115280                1
1534      1            74         3835  112000                1
1535      2            51         2223   60457                1
1536      1            51         2557   80750                1
1537      2            51         1766   54276                1

1538 rows × 5 columns
```

In [9]:

```
from sklearn.model_selection import train_test_split
x_train, x_test,y_train, y_test=train_test_split(x,y,test_size=0.33,random_state=42)
```

In [10]:
```
x_test.head(5)
```

Out[10]:

| | model | engine_power | age_in_days | km | previous_owners |
|---|---|---|---|---|---|
| **481** | 2 | 51 | 3197 | 120000 | 2 |
| **76** | 2 | 62 | 2101 | 103000 | 1 |
| **1502** | 1 | 51 | 670 | 32473 | 1 |
| **669** | 1 | 51 | 913 | 29000 | 1 |
| **1409** | 1 | 51 | 762 | 18800 | 1 |

In [11]:
```
y_test.head(5)
```

Out[11]:
```
481     7900
76      7900
1502    9400
669     8500
1409    9700
Name: price, dtype: int64
```

In [12]:
```
from sklearn.linear_model import LinearRegression #importing linear regression
reg=LinearRegression()
reg.fit(x_train,y_train)
```

Out[12]:
```
▾ LinearRegression
LinearRegression()
```

In [13]:
```
ypred=reg.predict(x_test)
```

In [14]:
```
ypred # array values of ypred
```

Out[14]:
```
array([ 5994.51703157,  7263.58726658,  9841.90754881,  9699.31627673,
       10014.19892635,  9630.58715835,  9649.4499026 , 10092.9819664 ,
        9879.19498711,  9329.19347948, 10407.2964056 ,  7716.91706011,
        7682.89152522,  6673.95810983,  9639.42618839, 10346.53679153,
        9366.53363673,  7707.90063494,  4727.33552438, 10428.17092937,
       10359.87663878, 10364.84674179,  7680.16157493,  9927.58506055,
        7127.7284177 ,  9097.51161986,  4929.31229715,  6940.60225317,
        7794.35120591,  9600.43942019,  7319.85877519,  5224.05298205,
        5559.52039134,  5201.35403287,  8960.11762682,  5659.72968338,
        9915.79926869,  8255.93615893,  6270.40332834,  8556.73835062,
        9749.72882426,  6873.76758364,  8951.72659758, 10301.95669828,
        8674.89268564, 10301.93257222,  9165.73586068,  8846.92420399,
        7044.68964545,  9052.4031418 ,  9390.75738772, 10267.3912561 ,
       10046.90924744,  6855.71260655,  9761.93338967,  9450.05744337,
        9274.98388541, 10416.00474283,  9771.10646661,  7302.96566423,
       10082.61483093,  6996.96553454,  9829.40534825,  7134.21944391,
        6407.26222178,  9971.82132188,  9757.01618446,  8614.84049875,
        8437.92452169,  6489.24658616,  7752.65456507,  6626.60510856,
        8329.88998217, 10412.00324329,  7342.77348105,  8543.63624413,
        9796.44743777, 10010.43582651,  7256.86786062,  9523.1488851 ,
```

In [15]:
```
from sklearn.metrics import r2_score
r2_score(y_test,ypred)
```

Out[15]:
```
0.8383895235218546
```

In [16]:
```
from sklearn.metrics import mean_squared_error
b=mean_squared_error(ypred,y_test)
```

In [17]:
```
srt=b**(1/2)
print(b)# getting mean square value
print(srt) #square root value for mean square
```

```
593504.2888137395
770.3922954013361
```

In [18]:

```
results=pd.DataFrame(columns=['price','predicted'])
results['price']=y_test #getting the result of price
results['predicted']=ypred #getting results of predicted price
results.head(15)
```

Out[18]:

| | price | predicted |
|---:|---:|---:|
| **481** | 7900 | 5994.517032 |
| **76** | 7900 | 7263.587267 |
| **1502** | 9400 | 9841.907549 |
| **669** | 8500 | 9699.316277 |
| **1409** | 9700 | 10014.198926 |
| **1414** | 9900 | 9630.587158 |
| **1089** | 9900 | 9649.449903 |
| **1507** | 9950 | 10092.981966 |
| **970** | 10700 | 9879.194987 |
| **1198** | 8999 | 9329.193479 |
| **1088** | 9890 | 10407.296406 |
| **576** | 7990 | 7716.917060 |
| **965** | 7380 | 7682.891525 |
| **1488** | 6800 | 6673.958110 |
| **1432** | 8900 | 9639.426188 |

In [19]:

```
#results=results.reset_index()
#results['Id']=results.index()
```