```
In [71]: import pandas as pd
```

```
In [72]: data=pd.read_csv("/home/placement/Downloads/fiat500.csv")
```

```
In [73]: import warnings
         warnings.filterwarnings('ignore')
```

```
In [74]: data.describe()
```

Out[74]:

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

In [76]:
```
data1
```

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

1538 rows × 6 columns

In [77]:
```
data2=data.loc[(data.model=='lounge')]
```

In [78]:
```
data2
```

Out[78]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1	lounge	51	882	25000	1	44.907242	8.611560	8900
3	4	lounge	51	2739	160000	1	40.633171	17.634609	6000
6	7	lounge	51	731	11600	1	44.907242	8.611560	10750
7	8	lounge	51	1521	49076	1	41.903221	12.495650	9190
11	12	lounge	51	366	17500	1	45.069679	7.704920	10990
...	...	...	...	...	...	...	...	...	...
1528	1529	lounge	51	2861	126000	1	43.841980	10.515310	5500
1529	1530	lounge	51	731	22551	1	38.122070	13.361120	9900
1530	1531	lounge	51	670	29000	1	45.764648	8.994500	10800
1534	1535	lounge	74	3835	112000	1	45.845692	8.666870	4600
1536	1537	lounge	51	2557	80750	1	45.000702	7.682270	5990

1094 rows × 9 columns

In [79]:
```
data2=pd.get_dummies(data2)
```

In [80]:

```
data2
```

Out[80]:

	ID	engine_power	age_in_days	km	previous_owners	lat	lon	price	model_lounge
0	1	51	882	25000	1	44.907242	8.611560	8900	1
3	4	51	2739	160000	1	40.633171	17.634609	6000	1
6	7	51	731	11600	1	44.907242	8.611560	10750	1
7	8	51	1521	49076	1	41.903221	12.495650	9190	1
11	12	51	366	17500	1	45.069679	7.704920	10990	1
...	...	...	...	...	...	...	...	...	...
1528	1529	51	2861	126000	1	43.841980	10.515310	5500	1
1529	1530	51	731	22551	1	38.122070	13.361120	9900	1
1530	1531	51	670	29000	1	45.764648	8.994500	10800	1
1534	1535	74	3835	112000	1	45.845692	8.666870	4600	1
1536	1537	51	2557	80750	1	45.000702	7.682270	5990	1

1094 rows × 9 columns

In [81]:

```
data2.shape
```

Out[81]:

```
(1094, 9)
```

In [82]:

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

In [83]:
```
y
```

Out[83]:
```
0        8900
3        6000
6       10750
7        9190
11      10990
        ...  
1528     5500
1529     9900
1530    10800
1534     4600
1536     5990
Name: price, Length: 1094, dtype: int64
```

In [84]:
```
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 [85]:
```
x_test.head(5)
```

Out[85]:

	ID	engine_power	age_in_days	km	previous_owners	lat	lon	model_lounge
676	677	51	762	18609	1	41.572239	13.33369	1
215	216	51	701	25000	1	44.988739	9.01050	1
146	147	51	4018	152900	1	43.067532	12.55155	1
1319	1320	51	731	20025	1	41.689281	13.25494	1
1041	1042	51	640	38231	1	41.107880	14.20881	1

In [86]:
```
x_train.head(5)
```

Out[86]:

	ID	engine_power	age_in_days	km	previous_owners	lat	lon	model_lounge
441	442	51	762	36448	1	45.571220	9.15914	1
701	702	51	701	27100	1	41.903221	12.49565	1
695	696	51	3197	51083	1	45.571220	9.15914	1
1415	1416	51	670	33000	1	42.287029	12.40754	1
404	405	51	456	14000	1	40.840141	14.25226	1

In [87]:
```
y_test.head(5)
```

Out[87]:
```
676     10250
215      9790
146      5500
1319     9900
1041     8900
Name: price, dtype: int64
```

In [88]:
```
y_train.head(5)
```

Out[88]:
```
441      8980
701     10300
695      5880
1415    10490
404      9499
Name: price, dtype: int64
```

In [89]:
```
x_train.shape
```

Out[89]:
```
(732, 8)
```

In [90]:
```
y_train
```

Out[90]:
```
441      8980
701     10300
695      5880
1415    10490
404      9499
        ...  
459     10850
654      5900
189     10000
1455     9400
1218     8900
Name: price, Length: 732, dtype: int64
```

In [91]:
```
x_train
```

Out[91]:

	ID	engine_power	age_in_days	km	previous_owners	lat	lon	model_lounge
441	442	51	762	36448	1	45.571220	9.15914	1
701	702	51	701	27100	1	41.903221	12.49565	1
695	696	51	3197	51083	1	45.571220	9.15914	1
1415	1416	51	670	33000	1	42.287029	12.40754	1
404	405	51	456	14000	1	40.840141	14.25226	1
...	...	...	...	...	...	...	...	...
459	460	51	397	15628	1	45.512569	10.32901	1
654	655	51	3227	95554	1	45.069679	7.70492	1
189	190	51	1431	81900	1	44.138371	12.23882	1
1455	1456	51	701	33942	1	41.107880	14.20881	1
1218	1219	51	882	25000	1	44.907242	8.61156	1

732 rows × 8 columns

In [92]:
```
#ridge regression
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import Ridge
alpha = [1e-15, 1e-10, 1e-8, 1e-4, 1e-3,1e-2, 1, 5, 10, 20,30]
ridge = Ridge()
parameters = {'alpha': alpha}
ridge_regressor = GridSearchCV(ridge, parameters)
ridge_regressor.fit(x_train, y_train)
```

Out[92]:
```
▸ GridSearchCV
▸ estimator: Ridge
    ▸ Ridge
```

In [93]:
```
ridge_regressor.best_params_
```

Out[93]:
```
{'alpha': 30}
```

In [94]:
```
ridge=Ridge(alpha=30)
ridge.fit(x_train,y_train)
y_pred_ridge=ridge.predict(x_test)
```

In [95]:
```
from sklearn.metrics import mean_squared_error
Ridge_Error=mean_squared_error(y_pred_ridge,y_test)
Ridge_Error
```

Out[95]:
```
529111.0455362241
```

In [96]:
```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred_ridge)
```

Out[96]:
```
0.8343797517106646
```

In [104]:
```
Results=pd.DataFrame(columns=['Actual','Pedicted'])
Results['Actual']=y_test
Results['Predicted']=y_pred_ridge
Results=Results.reset_index()
Results['ID']=Results.index
Results.head(10)
```

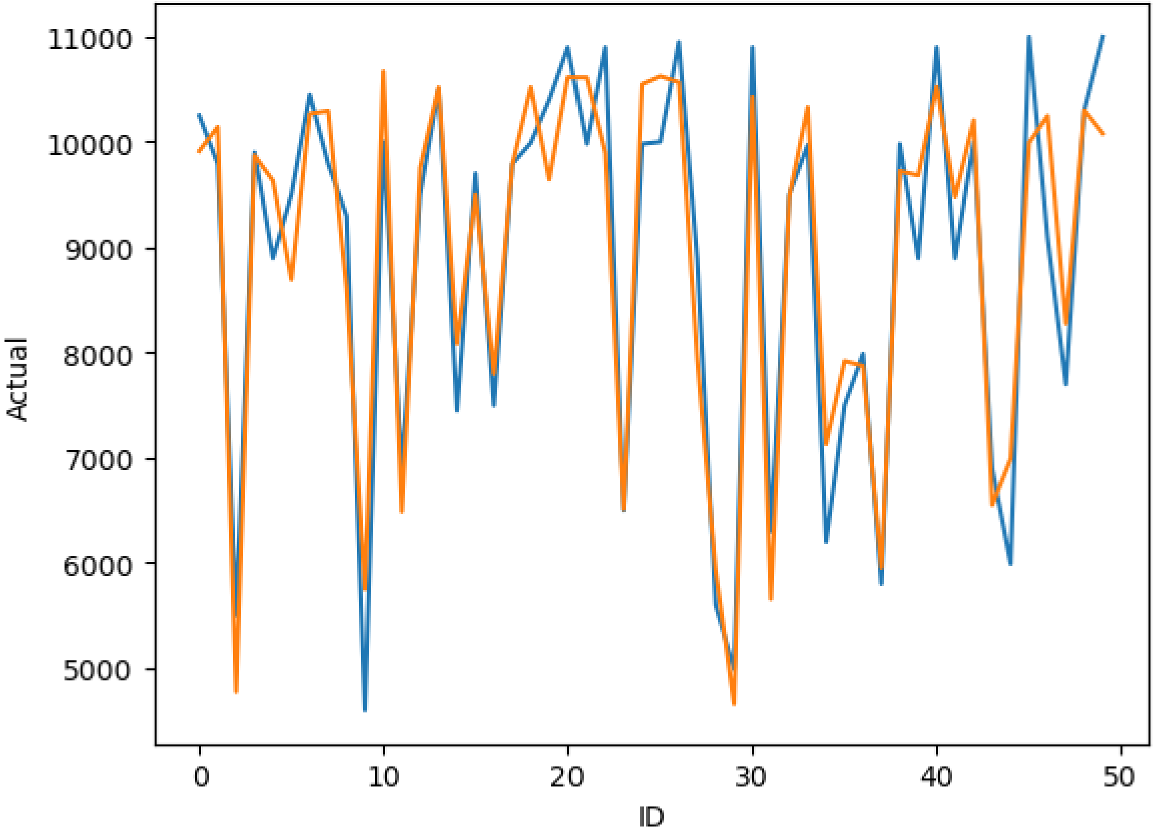
Out[104]:

	index	Actual	Pedicted	Predicted	ID
0	676	10250	NaN	9912.601754	0
1	215	9790	NaN	10141.748493	1
2	146	5500	NaN	4775.235521	2
3	1319	9900	NaN	9870.926966	3
4	1041	8900	NaN	9630.417885	4
5	1425	9500	NaN	8697.092014	5
6	409	10450	NaN	10265.822884	6
7	617	9790	NaN	10293.851867	7
8	1526	9300	NaN	8614.349738	8
9	1010	4600	NaN	5749.673567	9

In [105]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [107]: sns.lineplot(x='ID',y='Actual',data=Results.head(50))
          sns.lineplot(x='ID',y='Predicted',data=Results.head(50))
          plt.plot()
```

```
Out[107]: []
```



In [ ]: