In [1]:
```
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
```

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

In [3]:
```
data.describe()
```

Out[3]:

	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 [4]:
```
data.head()
```

Out[4]:

	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
1	2	pop	51	1186	32500	1	45.666359	12.241890	8800
2	3	sport	74	4658	142228	1	45.503300	11.417840	4200
3	4	lounge	51	2739	160000	1	40.633171	17.634609	6000
4	5	pop	73	3074	106880	1	41.903221	12.495650	5700

In [5]:
```
data1=data.drop(['lat','lon','ID'],axis=1)
data1
```

Out[5]:

	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 [6]:
```
data2=pd.get_dummies(data1)
data2
```

Out[6]:

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

1538 rows × 8 columns

In [7]:
```
data2.shape
```

Out[7]: (1538, 8)

In [8]:
```
y=data2['price']
x=data2.drop('price',axis=1)
```

In [9]:
```
y
```

Out[9]:
```
0       8900
1       8800
2       4200
3       6000
4       5700
        ... 
1533    5200
1534    4600
1535    7500
1536    5990
1537    7900
Name: price, Length: 1538, dtype: int64
```

In [10]:
```
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 [11]:
```
x_test.head(5)
```

Out[11]:

	engine_power	age_in_days	km	previous_owners	model_lounge	model_pop	model_sport
481	51	3197	120000	2	0	1	0
76	62	2101	103000	1	0	1	0
1502	51	670	32473	1	1	0	0
669	51	913	29000	1	1	0	0
1409	51	762	18800	1	1	0	0

In [12]:
```
x_train.head(5)
```

Out[12]:

	engine_power	age_in_days	km	previous_owners	model_lounge	model_pop	model_sport
527	51	425	13111	1	1	0	0
129	51	1127	21400	1	1	0	0
602	51	2039	57039	1	0	1	0
331	51	1155	40700	1	1	0	0
323	51	425	16783	1	1	0	0

In [13]:
```
y_test.head(5)
```

Out[13]:
```
481     7900
76      7900
1502    9400
669     8500
1409    9700
Name: price, dtype: int64
```

In [14]:
```
y_train.head(5)
```

Out[14]:
```
527    9990
129    9500
602    7590
331    8750
323    9100
Name: price, dtype: int64
```

In [15]:
```
from sklearn.linear_model import LinearRegression
reg=LinearRegression()
reg.fit(x_train,y_train)
```

Out[15]:
```
▾ LinearRegression
LinearRegression()
```

In [16]:
```
y_pred=reg.predict(x_test)
y_pred
```

```
       10011.26407146, 10402.02002918,  9945.08219601,  7770.52280413,
        8840.08397206,  9916.27565791, 10287.45603992,  9964.3213269 ,
        8403.51255128,  9345.81907605,  8521.46225147,  9743.68712672,
        9791.34520178,  9779.16293972,  6753.27416058,  7354.16762745,
        8760.24542762,  9923.66596418,  9812.92276721, 10466.90125415,
        8163.46726237,  6659.46839415,  9987.65677522,  8866.7826029 ,
        9952.37340054, 10187.72427693, 10231.39378767, 10091.11325493,
        9365.98570732, 10009.10088406,  9141.00566394, 10099.11667176,
        7803.77049829,  6009.84398185,  8800.33824151, 10237.60733785,
        5609.98366311, 10097.61555355,  9684.99946572,  7644.67379732,
        9276.37891542,  7371.5492091 , 10287.98873148, 10067.26428381,
       10552.64805598,  9966.72383894, 10068.46126756,  6232.53552963,
       10584.55044373,  9965.98687522, 10529.44404458,  9602.67646085,
        9665.77720284,  6186.06948587,  8073.87436253, 10345.58323918,
        6344.74803956,  7361.62678204, 10058.57116223,  6792.219309  ,
        7897.72464823,  5261.45936067,  4540.24137423,  8709.36468047,
        6882.0117409 ,  7406.73353952,  6795.61189392,  7047.27998963,
        9945.33400083,  8856.93910595,  9378.02074127, 10389.561154  ,
       10092.46332921, 10381.52000388,  9723.92466625,  5996.3331428 ,
```

In [17]:
```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred)
```

Out[17]:
```
0.8415526986865394
```

In [18]:
```
from sklearn.metrics import mean_squared_error
mean_squared_error(y_pred,y_test)
```

Out[18]: 581887.727391353

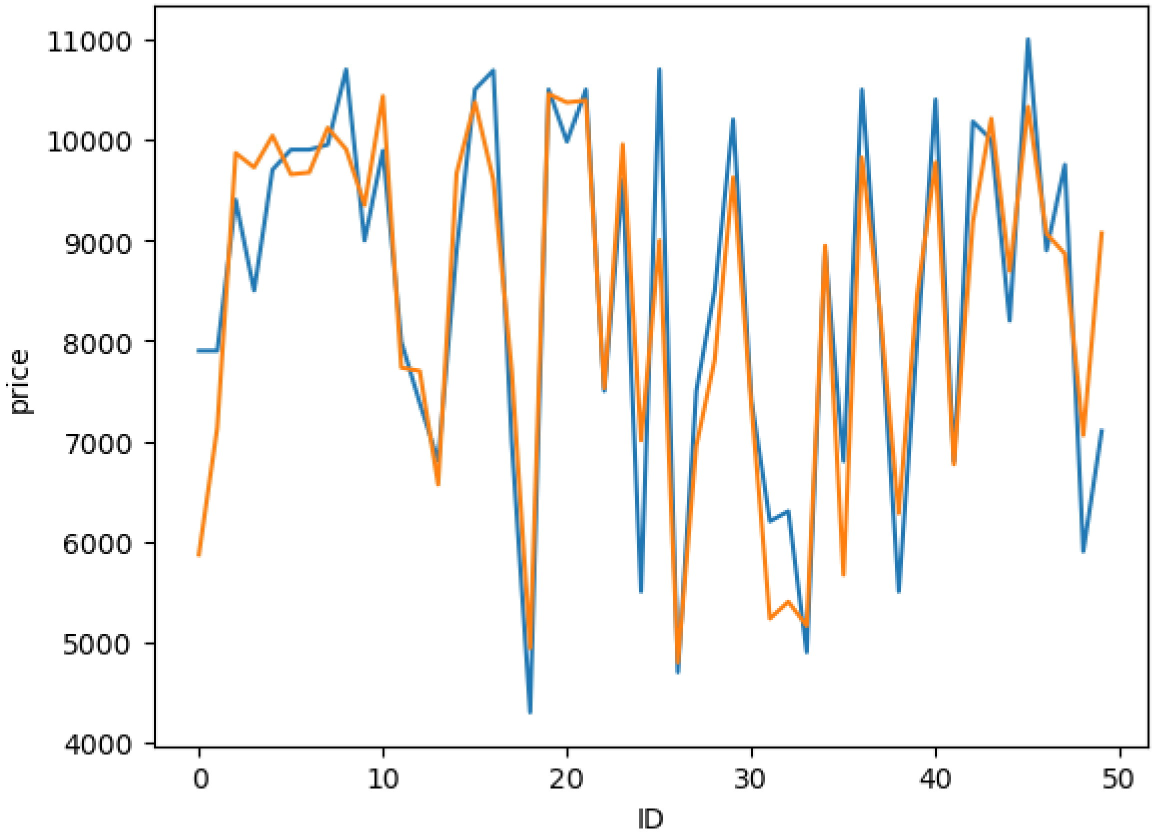
In [19]:
```
Results=pd.DataFrame(columns=['price','predicted'])
Results['price']=y_test
Results['predicted']=y_pred
Results=Results.reset_index()
Results['ID']=Results.index
Results.head(10)
```

Out[19]:

	index	price	predicted	ID
0	481	7900	5867.650338	0
1	76	7900	7133.701423	1
2	1502	9400	9866.357762	2
3	669	8500	9723.288745	3
4	1409	9700	10039.591012	4
5	1414	9900	9654.075826	5
6	1089	9900	9673.145630	6
7	1507	9950	10118.707281	7
8	970	10700	9903.859527	8
9	1198	8999	9351.558284	9

In [20]:
```
import seaborn as sns
import matplotlib.pyplot as plt
sns.lineplot(x='ID',y='price',data=Results.head(50))
sns.lineplot(x='ID',y='predicted',data=Results.head(50))
plt.plot()
```

Out[20]: []



In [ ]:
```
#linear regression end
```

In [21]:
```
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[21]:
```
▸ GridSearchCV
▸ estimator: Ridge
  ▸ Ridge
```

In [22]:
```
ridge_regressor.best_params_
```

Out[22]:
```
{'alpha': 30}
```

In [23]:
```
ridge=Ridge(alpha=30)
ridge.fit(x_train,y_train)
y_pred_ridge=ridge.predict(x_test)
```

In [24]:
```
from sklearn.metrics import mean_squared_error
Ridge_Error=mean_squared_error(y_pred_ridge,y_test)
Ridge_Error
```

Out[24]:
```
579521.7970897449
```

In [25]:
```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred_ridge)
```

Out[25]: 0.8421969385523054

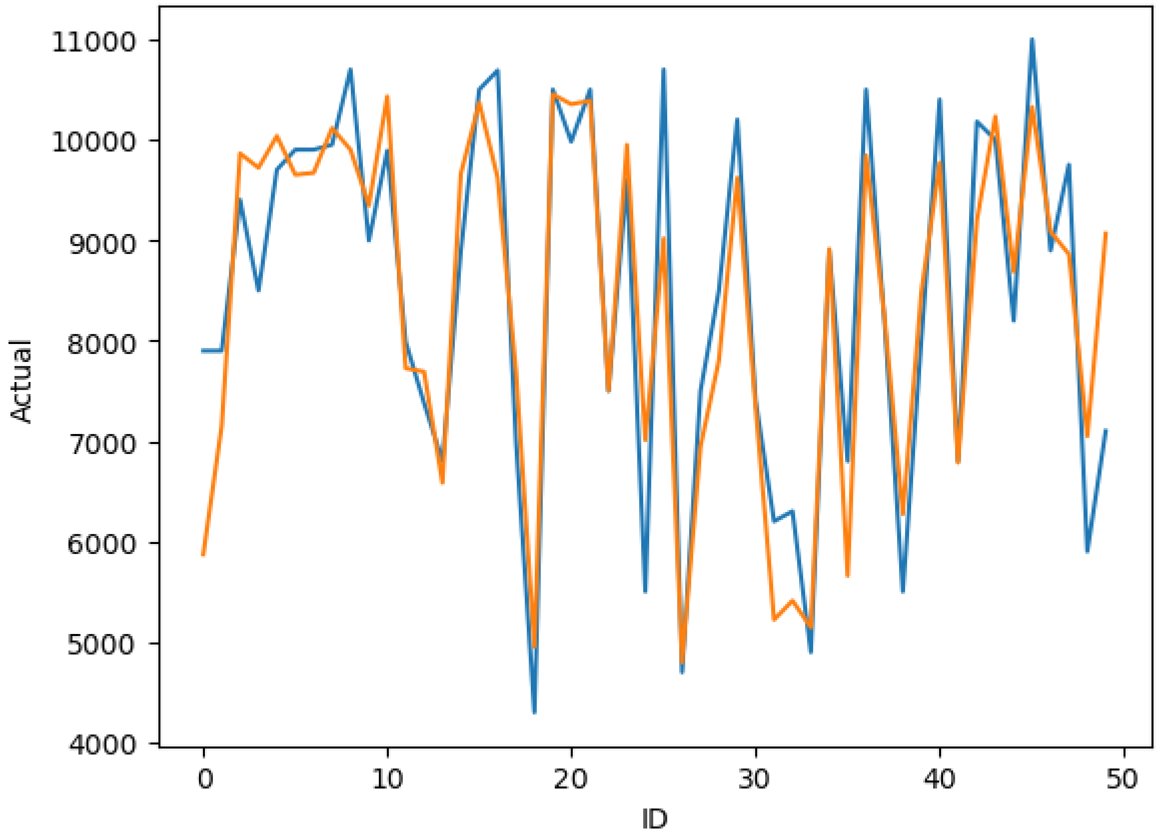
In [26]:
```
Results=pd.DataFrame(columns=['Actual','predicted'])
Results['Actual']=y_test
Results['predicted']=y_pred_ridge
Results=Results.reset_index()
Results['ID']=Results.index
Results.head(10)
```

Out[26]:

	index	Actual	predicted	ID
0	481	7900	5869.741155	0
1	76	7900	7149.563327	1
2	1502	9400	9862.785355	2
3	669	8500	9719.283532	3
4	1409	9700	10035.895686	4
5	1414	9900	9650.311090	5
6	1089	9900	9669.183317	6
7	1507	9950	10115.128380	7
8	970	10700	9900.241944	8
9	1198	8999	9347.080772	9

In [28]:
```
import seaborn as sns
import matplotlib.pyplot as plt
sns.lineplot(x='ID',y='Actual',data=Results.head(50))
sns.lineplot(x='ID',y='predicted',data=Results.head(50))
plt.plot()
```

Out[28]: []



In [ ]:
```
#ridge regression ends
```

In [29]:
```
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import ElasticNet

alpha = [1e-15, 1e-10, 1e-8, 1e-4, 1e-3,1e-2, 1, 5, 10, 20,30]

elastic = ElasticNet()

parameters = {'alpha': alpha}

elastic_regressor = GridSearchCV(elastic, parameters)

elastic_regressor.fit(x_train, y_train)
```

Out[29]:

GridSearchCV

```
GridSearchCV(estimator=ElasticNet(),
             param_grid={'alpha': [1e-15, 1e-10, 1e-08, 0.0001, 0.001, 0.01, 1,
                                   5, 10, 20, 30]})
```

estimator: ElasticNet

ElasticNet

In [30]:
```
elastic_regressor.best_params_
```

Out[30]:
```
{'alpha': 0.01}
```

In [31]:
```
elastic=ElasticNet(alpha=.01)
elastic.fit(x_train,y_train)
y_pred_elastic=elastic.predict(x_test)
```

In [32]:
```
from sklearn.metrics import mean_squared_error
elastic_Error=mean_squared_error(y_pred_elastic,y_test)
elastic_Error
```

Out[32]: 581390.7642825295

In [33]:
```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred_elastic)
```

Out[33]: 0.841688021120299

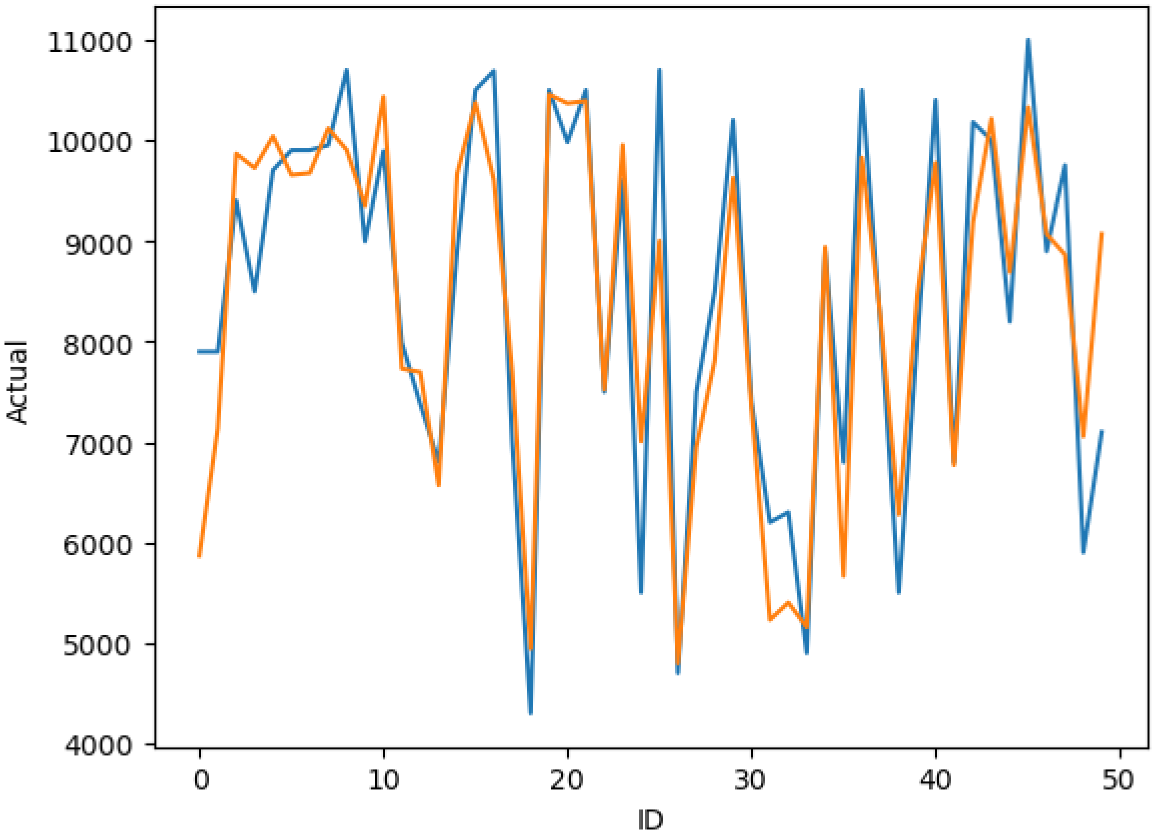
In [34]:
```
Results=pd.DataFrame(columns=['Actual','predicted'])
Results['Actual']=y_test
Results['predicted']=y_pred_elastic
Results=Results.reset_index()
Results['ID']=Results.index
Results.head(10)
```

Out[34]:

	index	Actual	predicted	ID
0	481	7900	5867.742075	0
1	76	7900	7136.527402	1
2	1502	9400	9865.726723	2
3	669	8500	9722.573593	3
4	1409	9700	10038.936496	4
5	1414	9900	9653.407122	5
6	1089	9900	9672.438692	6
7	1507	9950	10118.075470	7
8	970	10700	9903.219809	8
9	1198	8999	9350.750929	9

In [35]:
```
import seaborn as sns
import matplotlib.pyplot as plt
sns.lineplot(x='ID',y='Actual',data=Results.head(50))
sns.lineplot(x='ID',y='predicted',data=Results.head(50))
plt.plot()
```

Out[35]: []



In [ ]:
```
#elastic regression ends
```