In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
```

exploratory data analysis

In [3]:
```
data=pd.read_csv("/home/placement/Downloads/Advertising.csv")
```

In [4]:
```
data.describe()
```

Out[4]:

	Unnamed: 0	TV	radio	newspaper	sales
count	200.000000	200.000000	200.000000	200.000000	200.000000
mean	100.500000	147.042500	23.264000	30.554000	14.022500
std	57.879185	85.854236	14.846809	21.778621	5.217457
min	1.000000	0.700000	0.000000	0.300000	1.600000
25%	50.750000	74.375000	9.975000	12.750000	10.375000
50%	100.500000	149.750000	22.900000	25.750000	12.900000
75%	150.250000	218.825000	36.525000	45.100000	17.400000
max	200.000000	296.400000	49.600000	114.000000	27.000000

In [5]:
```
data.head()
```

Out[5]:

	Unnamed: 0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9

In [6]:
```
data1=data.drop(['Unnamed: 0'],axis=1)
data1
```

Out[6]:

	TV	radio	newspaper	sales
0	230.1	37.8	69.2	22.1
1	44.5	39.3	45.1	10.4
2	17.2	45.9	69.3	9.3
3	151.5	41.3	58.5	18.5
4	180.8	10.8	58.4	12.9
...	...	...	...	...
195	38.2	3.7	13.8	7.6
196	94.2	4.9	8.1	9.7
197	177.0	9.3	6.4	12.8
198	283.6	42.0	66.2	25.5
199	232.1	8.6	8.7	13.4

200 rows × 4 columns

In [7]:
```
y=data1['sales']
x=data1.drop('sales',axis=1)
```

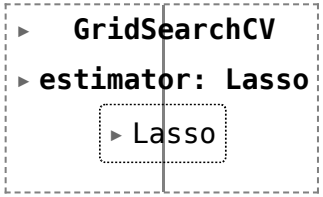
In [8]:
```
y
```

Out[8]:
```
0      22.1
1      10.4
2       9.3
3      18.5
4      12.9
       ...
195     7.6
196     9.7
197    12.8
198    25.5
199    13.4
Name: sales, Length: 200, dtype: float64
```

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]:
```
from sklearn.linear_model import Lasso
from sklearn.model_selection import GridSearchCV
lasso=Lasso()
parameters={'alpha':[1e-15,1e-10,1e-8,1e-4,1e-3,1e-2,1,5,10,20]}
lasso_regressor = GridSearchCV(lasso,parameters)
lasso_regressor.fit(x_train,y_train)
```

Out[10]:



In [11]:
```
lasso_regressor.best_params_
```

Out[11]:
```
{'alpha': 1}
```

In [12]:
```
lasso=Lasso(alpha=0.1)
lasso.fit(x_train,y_train)
y_pred_lasso=lasso.predict(x_test)
```

In [13]:
```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred_lasso)
```

Out[13]:
```
0.8559136390952934
```

In [14]:
```
from sklearn.metrics import mean_squared_error
lasso_Error=mean_squared_error(y_pred_lasso,y_test)
lasso_Error
```

Out[14]:
```
3.718719794627319
```

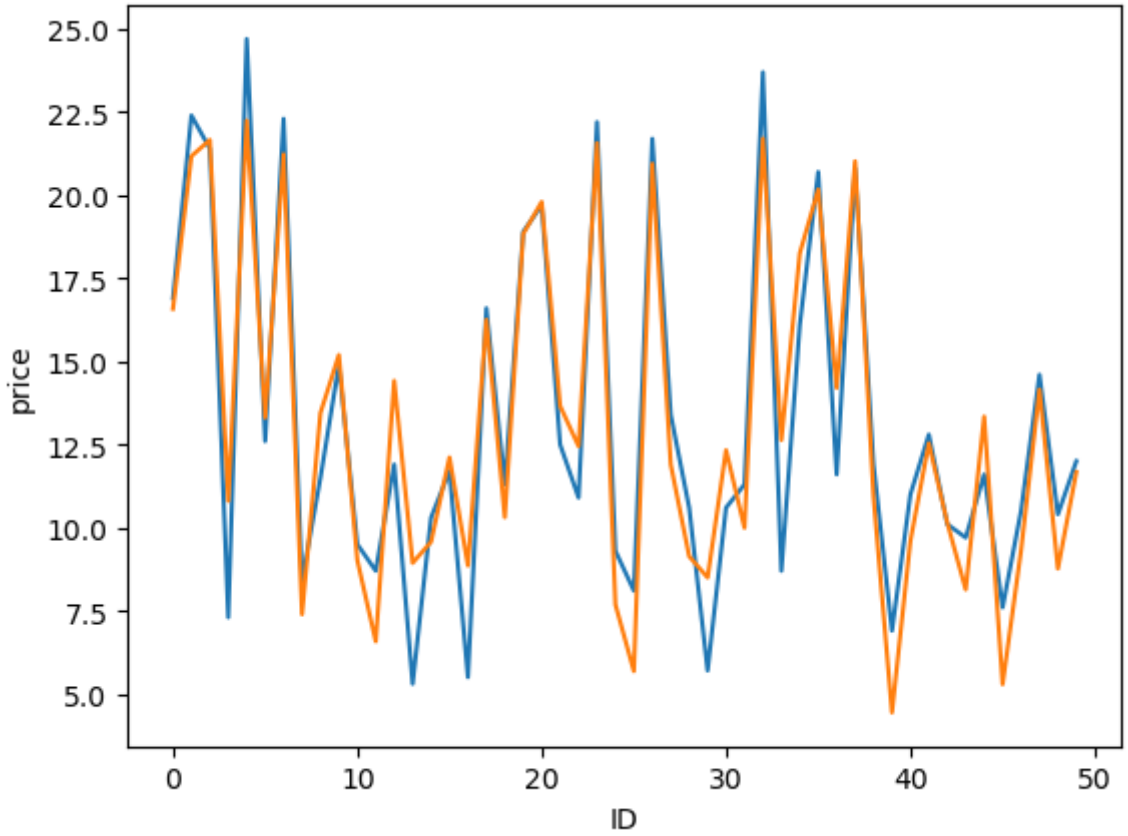
In [15]:
```
Results=pd.DataFrame(columns=['price','predicted'])
Results['price']=y_test
Results['predicted']=y_pred_lasso
Results=Results.reset_index()
Results['ID']=Results.index
Results.head(15)
```

Out[15]:

	index	price	predicted	ID
0	95	16.9	16.580451	0
1	15	22.4	21.173432	1
2	30	21.4	21.663263	2
3	158	7.3	10.804369	3
4	128	24.7	22.245736	4
5	115	12.6	13.307456	5
6	69	22.3	21.231000	6
7	170	8.4	7.391095	7
8	174	11.5	13.449902	8
9	45	14.9	15.194742	9
10	66	9.5	9.019847	10
11	182	8.7	6.577231	11
12	165	11.9	14.415634	12
13	78	5.3	8.936984	13
14	186	10.3	9.572569	14

In [16]:
```
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[16]: []



In [ ]: