In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
%matplotlib inline
```

Exploratory Data Analysis

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

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

Out[3]:

	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 [4]:
```
y=data['sales']
x=data.drop('sales',axis=1)
```

In [5]:
```
x
```

Out[5]:

	Unnamed: 0	TV	radio	newspaper
0	1	230.1	37.8	69.2
1	2	44.5	39.3	45.1
2	3	17.2	45.9	69.3
3	4	151.5	41.3	58.5
4	5	180.8	10.8	58.4
...	...	...	...	...
195	196	38.2	3.7	13.8
196	197	94.2	4.9	8.1
197	198	177.0	9.3	6.4
198	199	283.6	42.0	66.2
199	200	232.1	8.6	8.7

200 rows × 4 columns

In [6]:
```
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 [7]:
```
from sklearn.linear_model import Lasso
reg=Lasso()
reg.fit(x_train,y_train)
```

Out[7]:

▾ Lasso

Lasso()

In [8]:
```
y_pred=reg.predict(x_test)
```

In [9]:
```
y_pred
```

Out[9]:
```
array([16.52356688, 21.07183096, 21.63653184, 10.73467623, 22.18465223,
       13.23922119, 21.16664756,  7.4431686 , 13.52971188, 15.20686921,
        9.06623028,  6.63289843, 14.40121642,  8.95402791,  9.6407048 ,
       12.14981955,  8.84314265, 16.26648664, 10.32853903, 18.85137298,
       19.70493689, 13.5213463 , 12.41160087, 21.46562312,  7.78549068,
        5.78265455, 20.87376833, 11.96631497,  9.18377863,  8.5245863 ,
       12.43342356, 10.04074515, 21.58122748, 12.49489552, 18.26811877,
       20.16474367, 14.14870494, 20.95277886, 10.97979616,  4.53957393,
        9.67385982, 12.60613272, 10.14475068,  8.20968861, 13.42902186,
        5.36899823,  9.33781109, 14.18378647,  8.82071962, 11.68378102,
       15.56752612, 11.83221277, 13.07693412, 10.97649377,  6.53276607,
        9.92514199,  9.50755099, 24.21620593,  7.71785174, 12.42302645,
       17.64869313, 15.30224537, 11.48429434, 11.04853032, 16.71492188,
        6.9526154 ])
```

In [10]:
```
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import Lasso

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]:

▸ **GridSearchCV**

▸ **estimator: Lasso**

▸ Lasso

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

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

In [12]:
```
lasso=Lasso(alpha=0.01)
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)
```

Out[13]:
```
0.8589177083282906
```

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

Out[14]:
```
3.72733659282993
```

In [ ]:
```

```

In [15]:
```
Results=pd.DataFrame(columns=['Actual','predicted'])
Results['Actual']=y_test
Results['predicted']=y_pred_lasso
#Result['km']=x_test['km']
Results=Results.reset_index()
Results['price']=Results.index
Results.head(10)
```

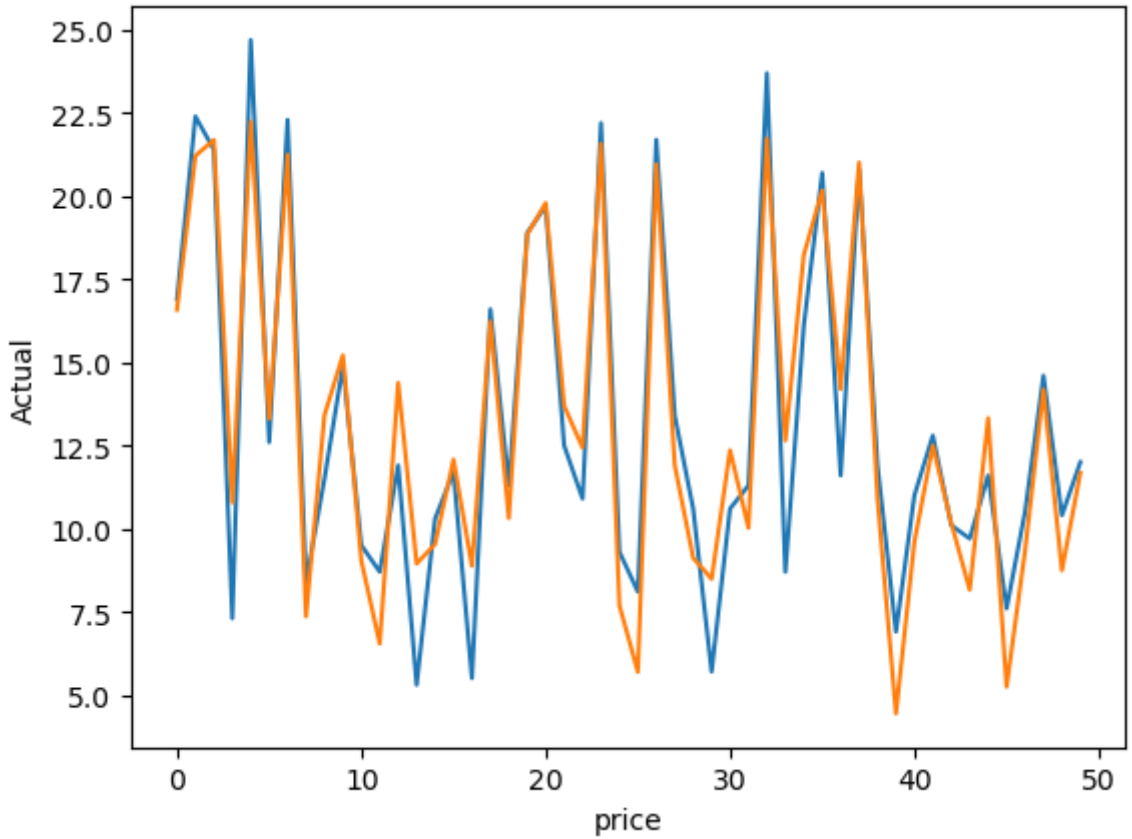
Out[15]:

	index	Actual	predicted	price
0	95	16.9	16.585351	0
1	15	22.4	21.213889	1
2	30	21.4	21.691715	2
3	158	7.3	10.786696	3
4	128	24.7	22.243765	4
5	115	12.6	13.305613	5
6	69	22.3	21.249652	6
7	170	8.4	7.359812	7
8	174	11.5	13.415111	8
9	45	14.9	15.214711	9

In [16]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

In [17]:
```
sns.lineplot(x='price',y='Actual',data=Results.head(50))#bule
sns.lineplot(x='price',y='predicted',data=Results.head(50))#orange
plt.plot()
```

Out[17]: []



In [ ]: