

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

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

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

Out[3]:

	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 [4]: list(data)
```

Out[4]: ['Unnamed: 0', 'TV', 'radio', 'newspaper', 'sales']

```
In [5]: data1=data.drop(['Unnamed: 0'],axis=1)
```

```
In [6]: 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]:

x

Out[8]:

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

200 rows × 3 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
```

```
Out[10]:
```

	TV	radio	newspaper
95	163.3	31.6	52.9
15	195.4	47.7	52.9
30	292.9	28.3	43.2
158	11.7	36.9	45.2
128	220.3	49.0	3.2
...
97	184.9	21.0	22.0
31	112.9	17.4	38.6
12	23.8	35.1	65.9
35	290.7	4.1	8.5
119	19.4	16.0	22.3

66 rows × 3 columns

```
In [11]: y_test
```

```
Out[11]: 95      16.9
          15      22.4
          30      21.4
          158      7.3
          128      24.7
          ...
          97      15.5
          31      11.9
          12       9.2
          35      12.8
          119      6.6
          Name: sales, Length: 66, dtype: float64
```

```
In [12]: x_train
```

```
Out[12]:
```

	TV	radio	newspaper
42	293.6	27.7	1.8
189	18.7	12.1	23.4
90	134.3	4.9	9.3
136	25.6	39.0	9.3
51	100.4	9.6	3.6
...
106	25.0	11.0	29.7
14	204.1	32.9	46.0
92	217.7	33.5	59.0
179	165.6	10.0	17.6
102	280.2	10.1	21.4

134 rows × 3 columns

```
In [13]: y_train
```

```
Out[13]: 42      20.7
         189      6.7
         90     11.2
         136      9.5
         51     10.7
         ...
        106      7.2
         14     19.0
         92     19.4
        179     12.6
        102     14.8
        Name: sales, Length: 134, dtype: float64
```

```
In [14]: 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[14]: ▸ GridSearchCV
         ▸ estimator: Lasso
           ▸ Lasso
```

```
In [15]: lasso_regressor.best_params_
```

```
Out[15]: {'alpha': 1}
```

```
In [16]: lasso = Lasso(alpha=30)
         lasso.fit(x_train, y_train)
         y_pred_lasso = lasso.predict(x_test)
```

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

```
Out[17]: 0.7659327162791038
```

```
In [ ]:
```