

In [1]: `#importing liberies
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error,r2_score`

In [2]: `#importing the dataset
data=pd.read_csv("Downloads/advertising.csv.xls")`

In [3]: `data`

Out[3]:

	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	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9
...
195	38.2	3.7	13.8	7.6
196	94.2	4.9	8.1	14.0
197	177.0	9.3	6.4	14.8
198	283.6	42.0	66.2	25.5
199	232.1	8.6	8.7	18.4

200 rows × 4 columns

In [4]: `#first 10 rows of data
data.head(10)`

Out[4]:

	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	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9
5	8.7	48.9	75.0	7.2
6	57.5	32.8	23.5	11.8
7	120.2	19.6	11.6	13.2
8	8.6	2.1	1.0	4.8
9	199.8	2.6	21.2	15.6

In [5]: `#last 10 rows
data.tail(10)`

Out[5]:

	TV	Radio	Newspaper	Sales
190	39.5	41.1	5.8	10.8
191	75.5	10.8	6.0	11.9
192	17.2	4.1	31.6	5.9
193	166.8	42.0	3.6	19.6
194	149.7	35.6	6.0	17.3
195	38.2	3.7	13.8	7.6
196	94.2	4.9	8.1	14.0
197	177.0	9.3	6.4	14.8
198	283.6	42.0	66.2	25.5
199	232.1	8.6	8.7	18.4

In [6]: `data.columns`

Out[6]: `Index(['TV', 'Radio', 'Newspaper', 'Sales'], dtype='object')`

In [7]: `data.describe()`

Out[7]:

	TV	Radio	Newspaper	Sales
count	200.000000	200.000000	200.000000	200.000000
mean	147.042500	23.264000	30.554000	15.130500
std	85.854236	14.846809	21.778621	5.283892
min	0.700000	0.000000	0.300000	1.600000
25%	74.375000	9.975000	12.750000	11.000000
50%	149.750000	22.900000	25.750000	16.000000
75%	218.825000	36.525000	45.100000	19.050000
max	296.400000	49.600000	114.000000	27.000000

In [30]: `x=data[['Newspaper']].values
y=data['Sales'].values`

In [31]: `#splitting the data into tranng and testing
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=42)`

In [32]: `#sales prediction model
model=LinearRegression().fit(x_train,y_train)`

In [33]: `#making prediction
y_pred=model.predict(x_test)`

In [34]: `mse,r2=mean_squared_error(y_test,y_pred),r2_score(y_test,y_pred)`

In [35]: `#printing results
print("mean squared error:",mse)
print("r-squared:",r2)`

mean squared error: 30.759376922769615
r-squared: 0.004586344085821592

In [36]: `new_Newspaper=10000
predicted_Sales=model.predict(np.array([[new_Newspaper]]))
print("predicted sales:",predicted_Sales[0])`

predicted sales: 385.96126994791496