

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [2]: df=pd.read_csv("Advertising.csv")
df
```

```
Out[2]:
```

	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
...	...	...	...	...	...
195	196	38.2	3.7	13.8	7.6
196	197	94.2	4.9	8.1	9.7
197	198	177.0	9.3	6.4	12.8
198	199	283.6	42.0	66.2	25.5
199	200	232.1	8.6	8.7	13.4

200 rows × 5 columns

```
In [3]: df.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]: df.tail()
```

```
Out[4]:
```

	Unnamed: 0	TV	Radio	Newspaper	Sales
195	196	38.2	3.7	13.8	7.6
196	197	94.2	4.9	8.1	9.7
197	198	177.0	9.3	6.4	12.8
198	199	283.6	42.0	66.2	25.5
199	200	232.1	8.6	8.7	13.4

```
In [5]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 5 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Unnamed: 0   200 non-null    int64
1   TV           200 non-null    float64
2   Radio        200 non-null    float64
3   Newspaper    200 non-null    float64
4   Sales        200 non-null    float64
dtypes: float64(4), int64(1)
memory usage: 7.9 KB
```

In [6]: `df.describe()`

Out[6]:

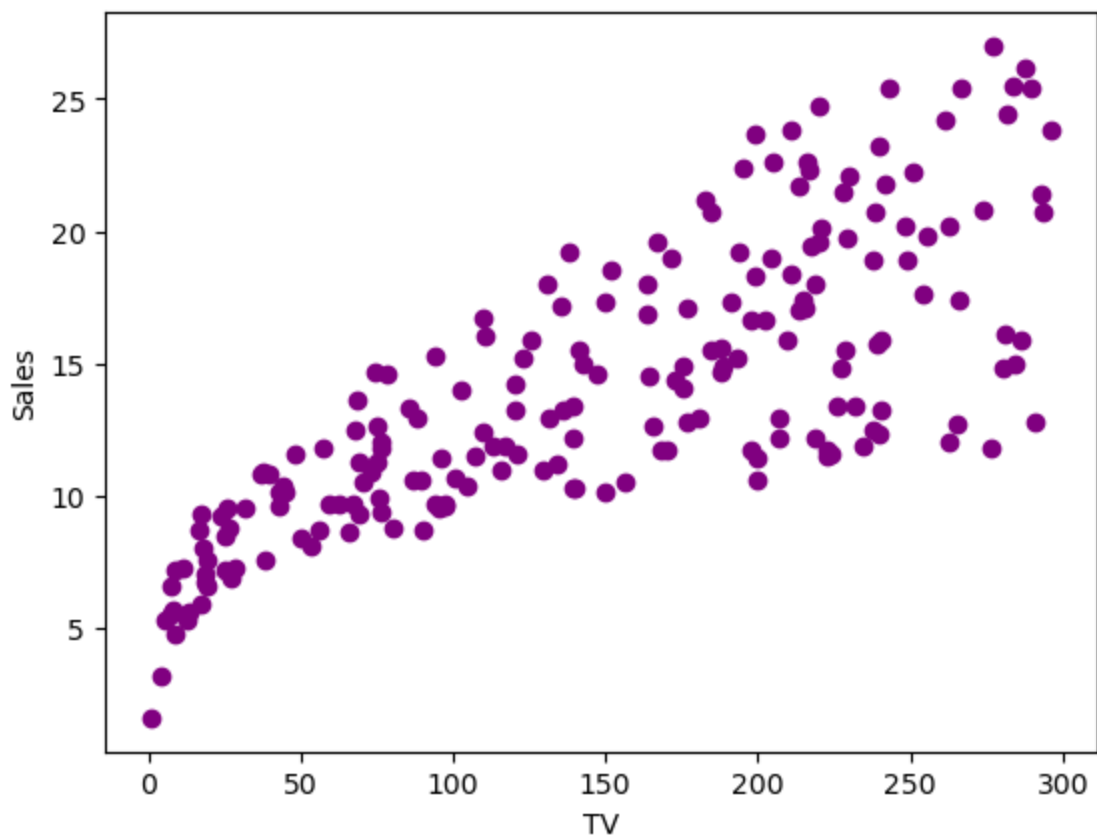
	Unnamed: 0	TV	Radio	Newspaper	Sales
<b>count</b>	200.000000	200.000000	200.000000	200.000000	200.000000
<b>mean</b>	100.500000	147.042500	23.264000	30.554000	14.022500
<b>std</b>	57.879185	85.854236	14.846809	21.778621	5.217457
<b>min</b>	1.000000	0.700000	0.000000	0.300000	1.600000
<b>25%</b>	50.750000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

In [7]: `df.isnull().sum()`

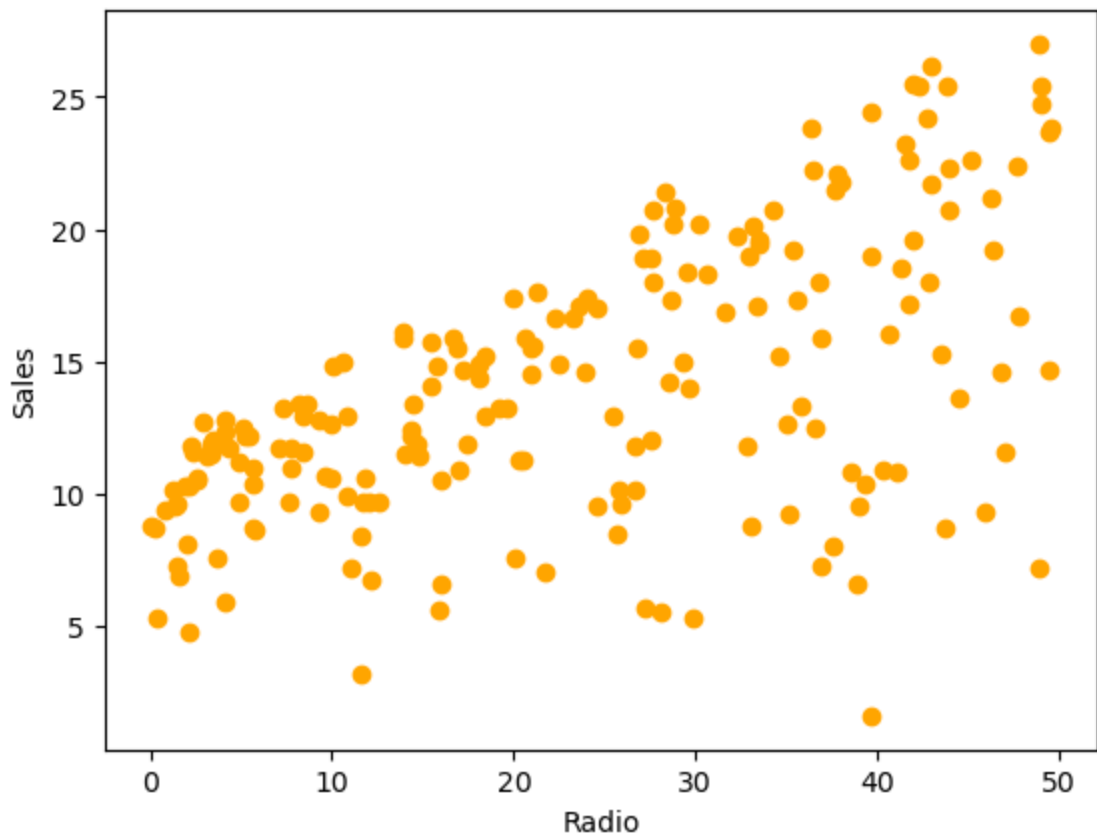
Out[7]:

```
Unnamed: 0    0
TV            0
Radio         0
Newspaper     0
Sales         0
dtype: int64
```

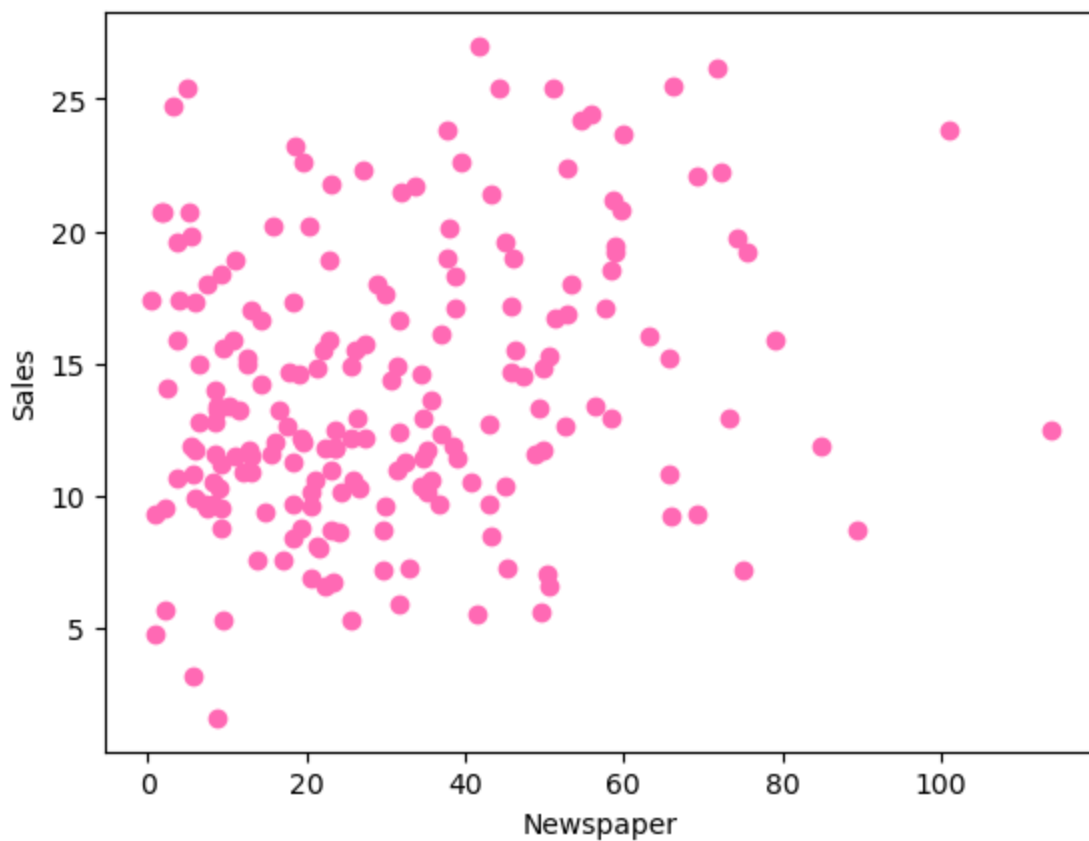
In [10]: `plt.scatter(df['TV'],df['Sales'],color='purple')
plt.xlabel('TV')
plt.ylabel('Sales')
plt.show()`



```
In [11]: plt.scatter(df['Radio'],df['Sales'],color='orange')
plt.xlabel('Radio')
plt.ylabel('Sales')
plt.show()
```



```
In [13]: plt.scatter(df['Newspaper'],df['Sales'],color='hotpink')
plt.xlabel('Newspaper')
plt.ylabel('Sales')
plt.show()
```



```
In [14]: from sklearn.model_selection import train_test_split
X=df.drop(columns=['Sales'])
Y=df['Sales']
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.30)
```

```
In [15]: from sklearn.linear_model import LinearRegression
model=LinearRegression()
```

```
In [16]: model.fit(x_train, y_train)
```

```
Out[16]: ▼ LinearRegression
LinearRegression()
```

```
In [17]: model.score(x_train,y_train)*100
```

```
Out[17]: 87.12169659684093
```

```
In [ ]:
```