In [2]:
```
import pandas as pd
X=pd.read_csv(r"C:\Users\HP\Downloads\data analysis\advertisingML.csv")
X.head()
```

Out[2]:

	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

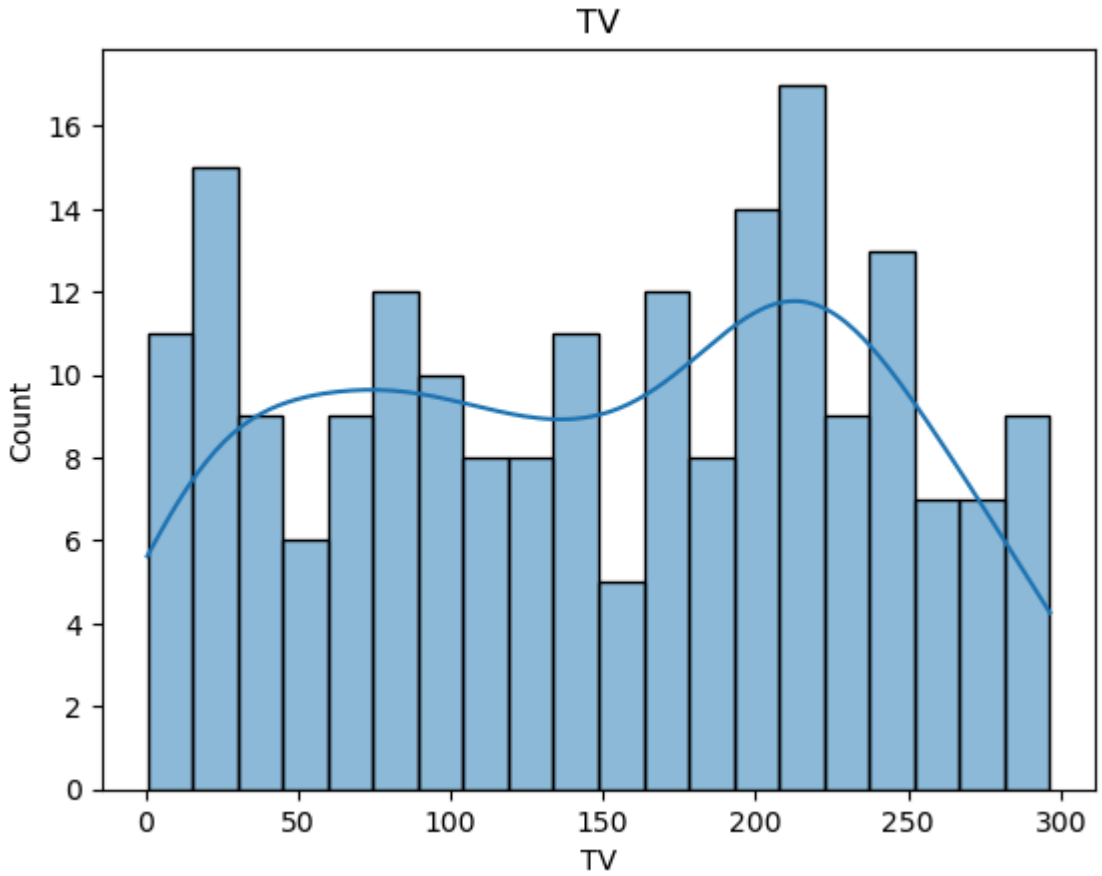
In [3]:
```
X.shape
```

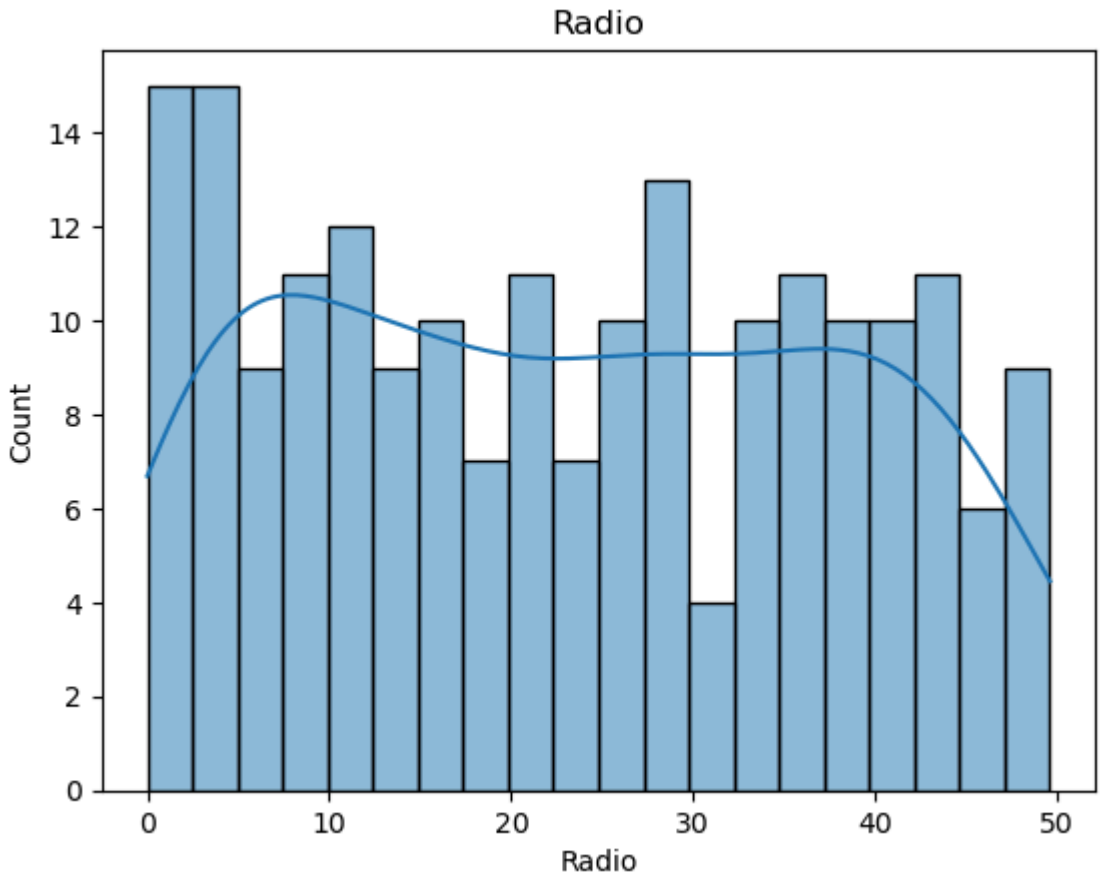
Out[3]:
```
(200, 4)
```

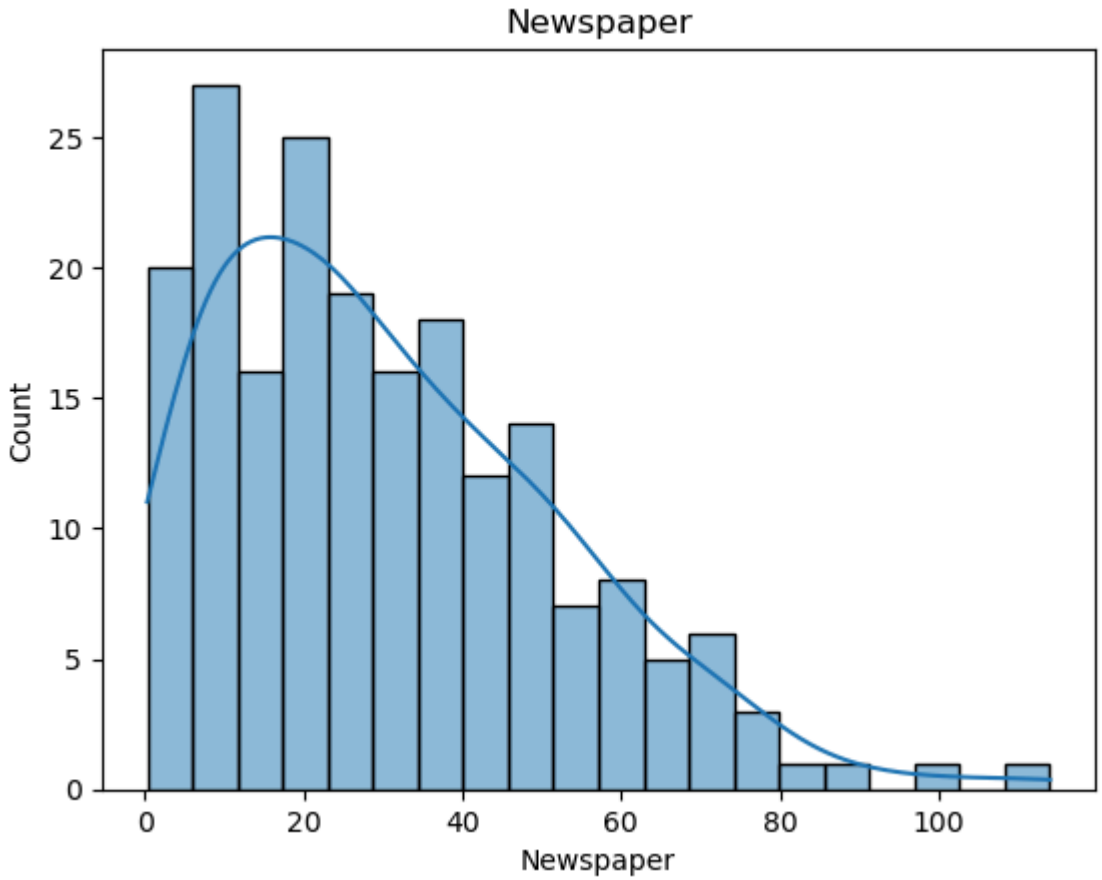
In [4]:
```
X.isnull().sum()
```

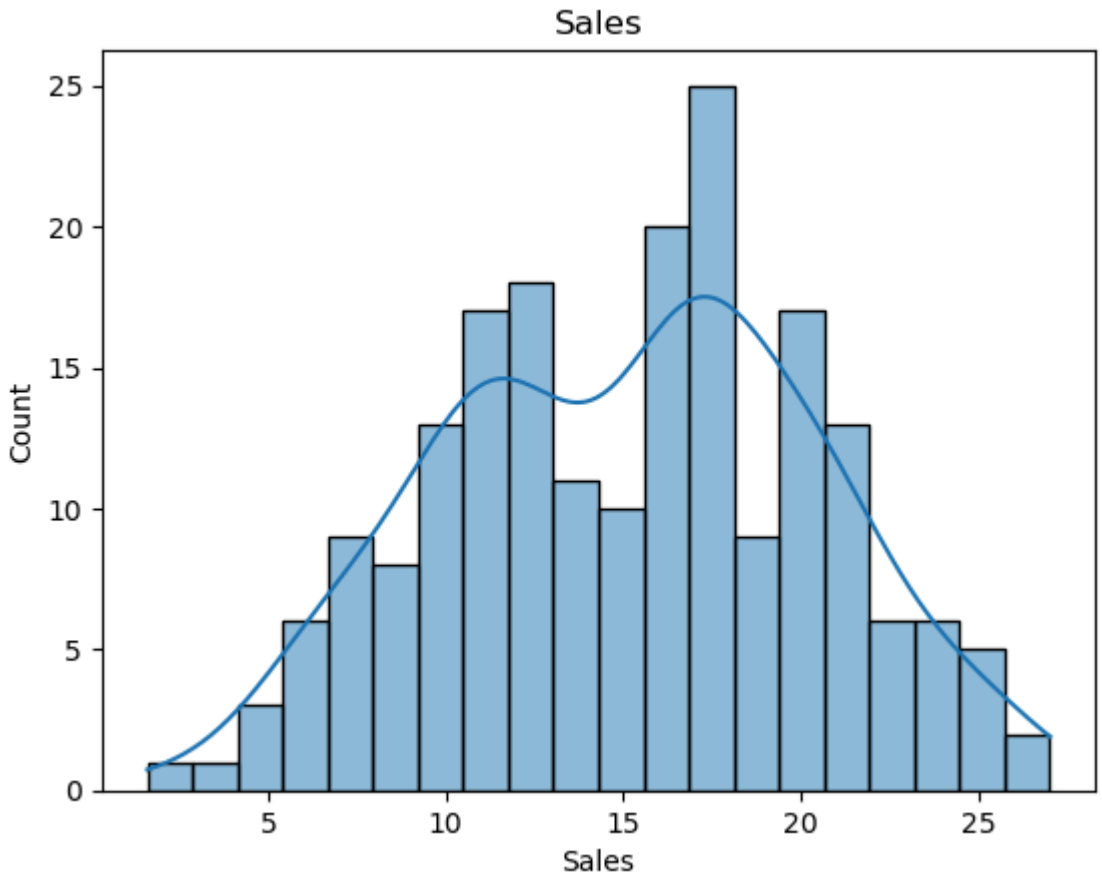
Out[4]:
```
TV           0
Radio        0
Newspaper    0
Sales        0
dtype: int64
```

In [10]:
```
import matplotlib.pyplot as plt
import seaborn as sns
for i in X.columns:
    sns.histplot(X[i],bins=20,kde=True)
    plt.title(i)
    plt.show()
```




Radio
Count
Radio
0
2
4
6
8
10
12
14
0
10
20
30
40
50


Newspaper
Count
25
20
15
10
5
0
0
20
40
60
80
100
Newspaper


Sales
Count
Sales
0
5
10
15
20
25
5
10
15
20
25

In [11]:
```
X
```

Out[11]:

	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

SIMPLE LINEAR REGRESSION

In [90]:
```
F=X[['TV']] #2D
T=X[['Sales']]
```

In [91]:
```
F
```

Out[91]:

	TV
0	230.1
1	44.5
2	17.2
3	151.5
4	180.8
...	...
195	38.2
196	94.2
197	177.0
198	283.6
199	232.1

200 rows × 1 columns

In [92]:
```
T
```

Out[92]:

	Sales
0	22.1
1	10.4
2	12.0
3	16.5
4	17.9
...	...
195	7.6
196	14.0
197	14.8
198	25.5
199	18.4

200 rows × 1 columns

In [93]:
```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(F,T,test_size=0.10)
```

In [94]:
```
X_train.shape
```

Out[94]:
```
(180, 1)
```

In [95]:
```
X_test.shape
```

Out[95]:
```
(20, 1)
```

In [96]:
```
y_train.shape
```

Out[96]:
```
(180, 1)
```

In [97]:
```
y_test.shape
```

Out[97]:
```
(20, 1)
```

In [98]:
```
from sklearn.linear_model import LinearRegression
A=LinearRegression()
```

In [99]:
```
A.fit(X_train,y_train)
```

Out[99]:
```
LinearRegression
LinearRegression()
```

In [100]:
```
pred=A.predict(X_test)
pred[1:5]
```

Out[100]:
```
array([[ 7.70138691],
       [20.33780937],
       [11.16846003],
       [11.22977483]])
```

In [101]:
```
y_test[1:5]
```

Out[101]:

	Sales
108	5.3
81	17.3
82	11.3
80	11.8

In [102]:
```
from sklearn.metrics import mean_squared_error  ,mean_absolute_error  ,r2_score
```

In [103]:
```
mean_squared_error(y_test,pred) #MSE
```

Out[103]: 8.783669429800064

In [104]:
```
mean_absolute_error(y_test,pred) #MAE
```

Out[104]: 2.390229793788078

In [105]:
```
R=r2_score(y_test,pred) #R SQUARED
R
```

Out[105]: 0.7346114972837725

In [106]:
```
import numpy as np #RMSE
np.sqrt(mean_squared_error(y_test,pred))
```

Out[106]: 2.9637255996127685

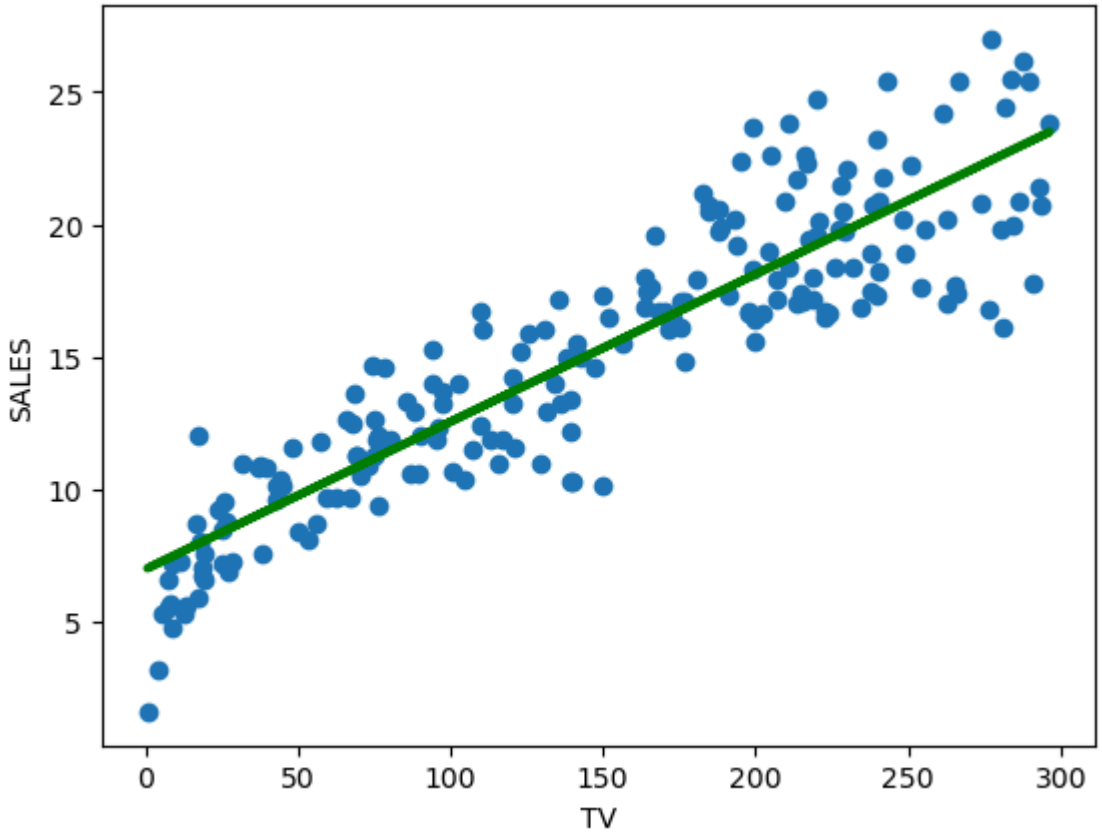
In [107]:
```
#ADJ R SQUARED
1-(((1-R)*(X.shape[0]-1))/(X.shape[0]-X.shape[1]-1))
```

Out[107]: 0.7291676305613883

In [ ]:
```

```

In [112]:
```
plt.scatter(F,T)
plt.plot(F,A.predict(F),color='g',lw=3)
plt.xlabel('TV')
plt.ylabel('SALES')
plt.show()
```



In [113]:
```
A.intercept_#C
```

Out[113]:
```
array([6.9711834])
```

In [114]:
```
A.coef_#M
```

Out[114]:
```
array([[0.05574073]])
```

In [ ]:
```

```