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("C:\\Users\\SRI KAAVYA\\OneDrive\\Desktop\\Internship projects\
df.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]:
```
df.shape
```

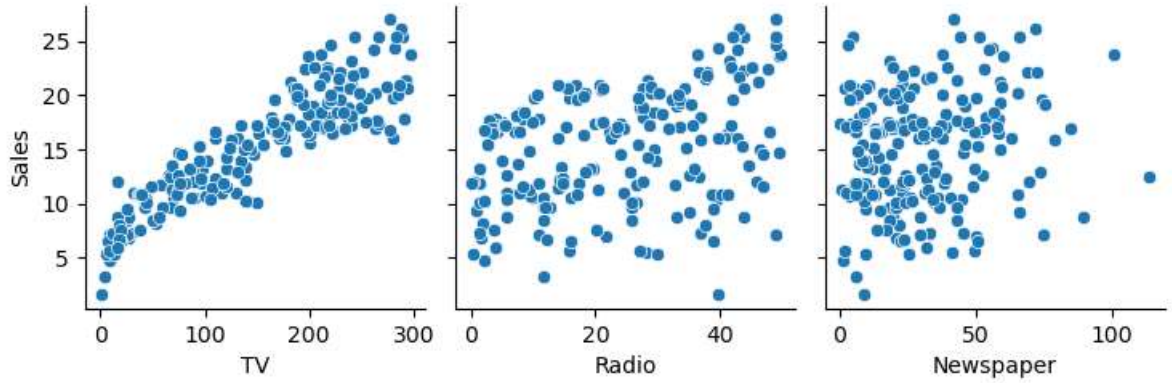
Out[3]:
```
(200, 4)
```

In [4]:
```
df.describe()
```

Out[4]:

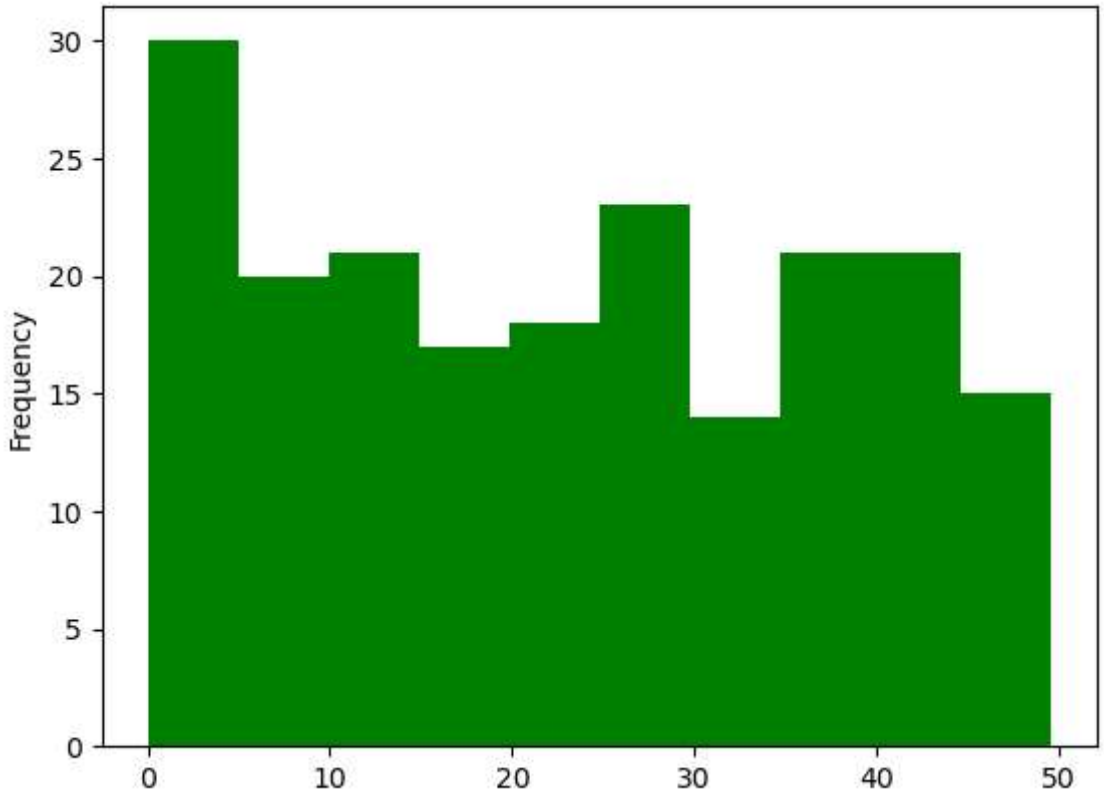
	TV	Radio	Newspaper	Sales
count	200.000000	200.000000	200.000000	200.000000
mean	147.042500	23.264000	30.554000	15.130500
std	85.854236	14.846809	21.778621	5.283892
min	0.700000	0.000000	0.300000	1.600000
25%	74.375000	9.975000	12.750000	11.000000
50%	149.750000	22.900000	25.750000	16.000000
75%	218.825000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

In [5]:
```
sns.pairplot(df,x_vars=['TV','Radio','Newspaper'],y_vars='Sales',kind='scatter
plt.show()
```



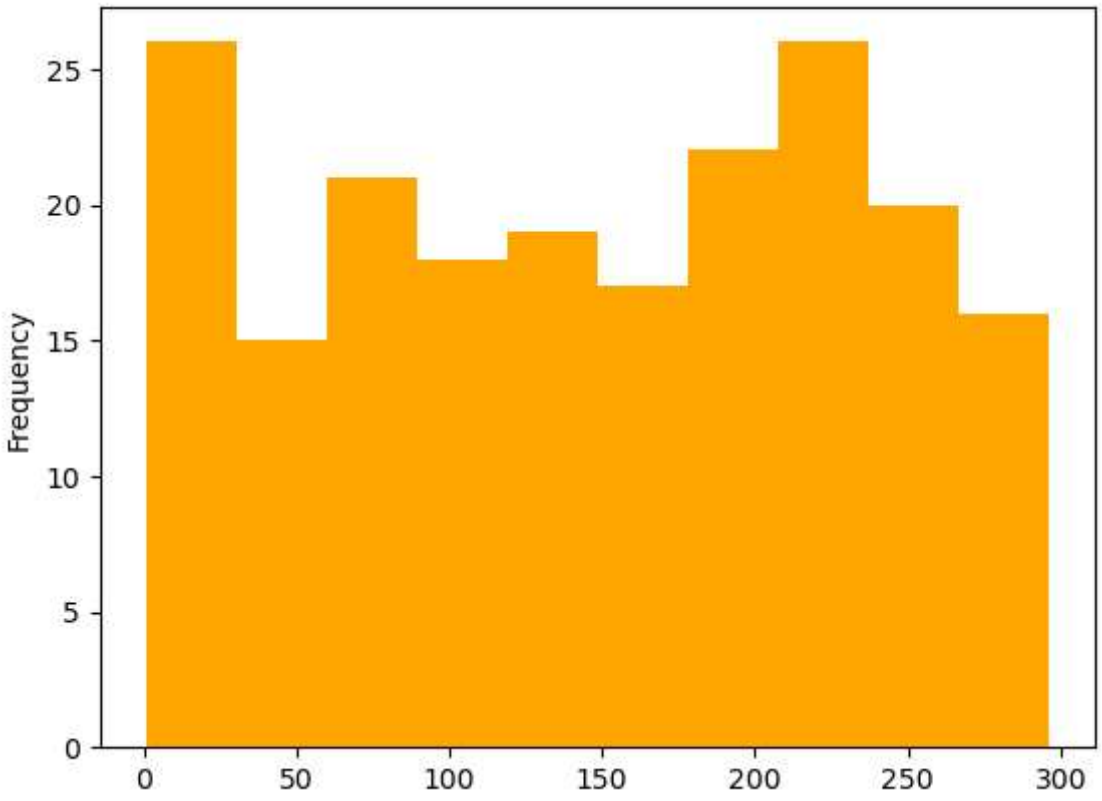
In [6]:
```
df['Radio'].plot.hist(bins=10, color='green',xlabel='Radio')
```

Out[6]:
```
<Axes: ylabel='Frequency'>
```



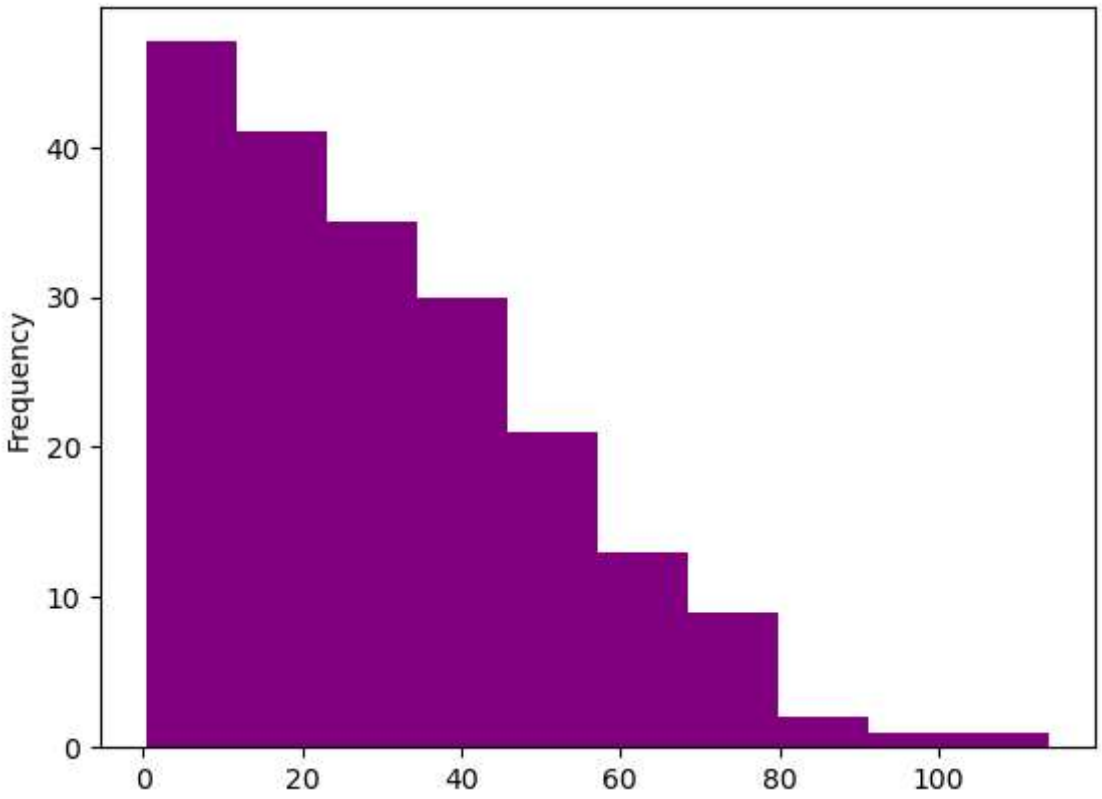
In [7]:
```
df['TV'].plot.hist(bins=10, color='orange',xlabel='TV')
```

Out[7]:
```
<Axes: ylabel='Frequency'>
```

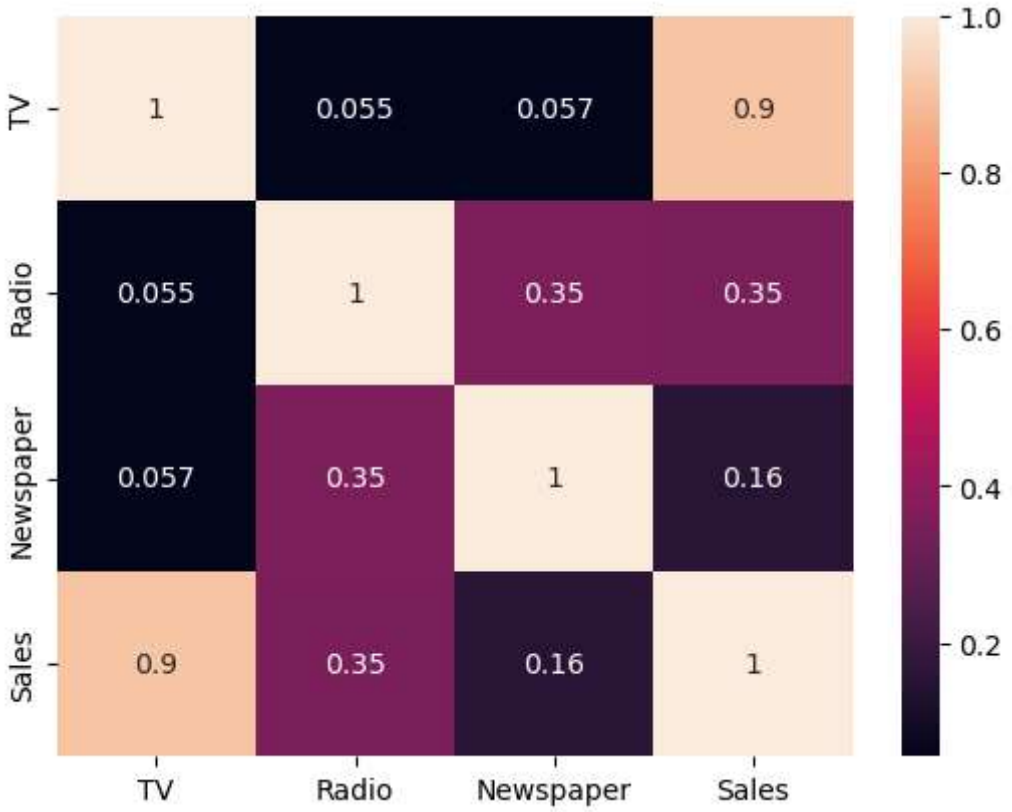


In [8]:
```
df['Newspaper'].plot.hist(bins=10, color='purple',xlabel='Newspaper')
```

Out[8]:
```
<Axes: ylabel='Frequency'>
```



In [9]:
```
sns.heatmap(df.corr(),annot=True)
plt.show()
```



In [10]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train,y_test=train_test_split(df[['TV']],df[['Sales']],test
```

In [11]:
```
print(X_train)
```

```
        TV
187  191.1
46    89.7
138   43.0
195   38.2
108   13.1
..     ...
67   139.3
192   17.2
117   76.4
47   239.9
172   19.6

[120 rows x 1 columns]
```

In [12]:
```
print(y_train)
```

```
     Sales
187   17.3
46    10.6
138    9.6
195    7.6
108    5.3
..     ...
67    13.4
192    5.9
117    9.4
47    23.2
172    7.6

[120 rows x 1 columns]
```

In [13]:
```
print(X_test)
```

```
        TV
18    69.2
170   50.0
107   90.4
98   289.7
177  170.2
..     ...
89   109.8
109  255.4
14   204.1
27   240.1
141  193.7

[80 rows x 1 columns]
```

In [14]:
```
print(y_test)
```

```
     Sales
18    11.3
170    8.4
107   12.0
98    25.4
177   16.7
..     ...
89    16.7
109   19.8
14    19.0
27    20.9
141   19.2

[80 rows x 1 columns]
```

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

Out[15]:
```
LinearRegression()
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

In [16]:
```
res=model.predict(X_test)
print(res)
```

```
[[10.87270287]
 [ 9.80413715]
 [12.05257753]
 [23.14451236]
 [16.49380382]
 [10.14919483]
 [ 7.50560775]
 [20.38405091]
 [ 8.34599017]
 [18.01873615]
 [21.56392556]
 [11.87448324]
 [15.7369031 ]
 [17.4733224 ]
 [11.27341502]
 [13.71108058]
 [21.80323976]
 [ 7.06037204]
 [11.17880243]
 [18.8980767 ]
 [23.02763798]
 [14.82973532]
 [16.7665107 ]
 [14.35110692]
 [ 9.99892778]
 [13.87247853]
 [16.23779328]
 [18.43057919]
 [19.48801402]
 [ 8.41834098]
 [10.79478662]
 [18.09108696]
 [22.64362218]
 [20.47309805]
 [ 7.75605285]
 [ 8.0621524 ]
 [10.33842001]
 [17.08374114]
 [10.8281793 ]
 [ 8.41277553]
 [ 9.07506366]
 [ 8.77452955]
 [14.97443693]
 [18.68658973]
 [19.00938563]
 [12.73712744]
 [ 7.50004231]
 [ 7.96197436]
 [14.01718013]
 [12.84287093]
 [13.13227414]
 [19.79411358]
 [21.14651707]
 [17.31192445]
 [ 9.50916849]
 [10.48868707]
 [23.32260665]
```

```
[16.62180909]
[18.29144303]
[ 7.4276915 ]
[21.78097797]
[18.01873615]
[19.18191447]
[15.21931657]
[16.55502373]
[19.1373909 ]
[21.64184181]
[16.1209189 ]
[12.60912218]
[11.26784957]
[17.31192445]
[14.49580853]
[22.25404092]
[23.51739728]
[12.37537342]
[13.13227414]
[21.23556422]
[18.38049017]
[20.38405091]
[17.80168374]]
```

In [17]:
```
model.coef_
```

Out[17]:
```
array([[0.05565446]])
```

In [18]:
```
model.intercept_
```

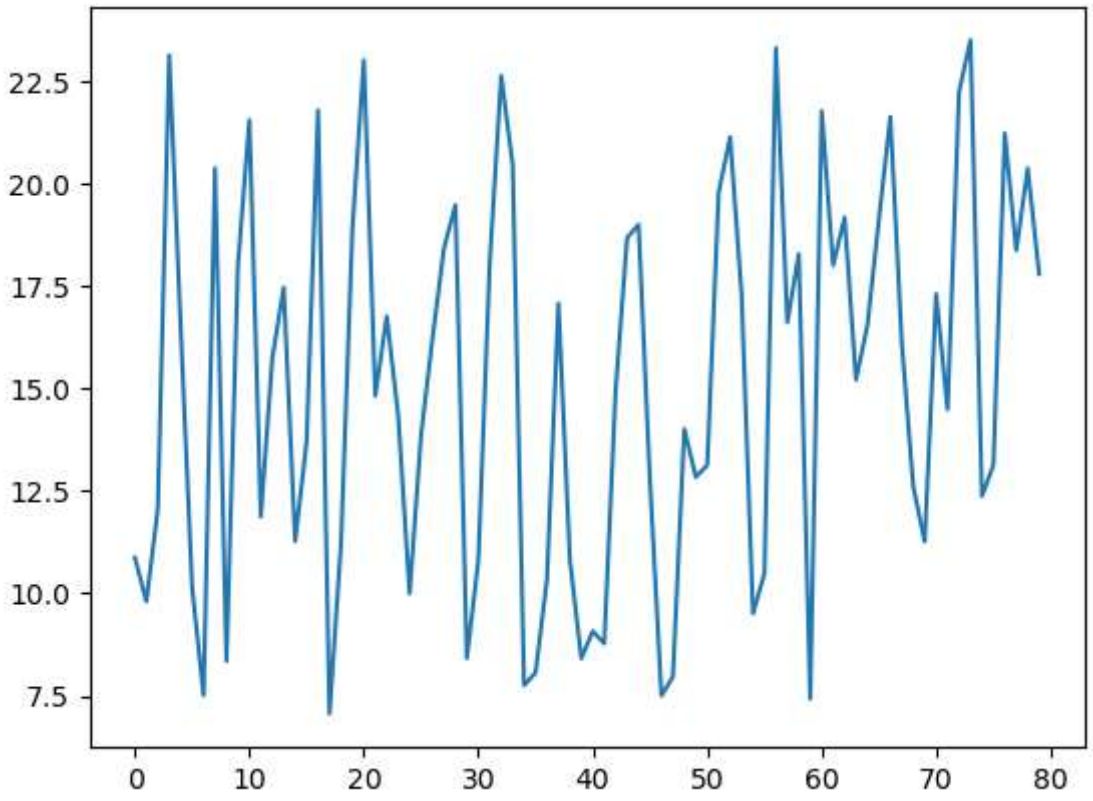
Out[18]:
```
array([7.02141391])
```

In [19]:
```
0.05473199*69.3+7.4573254
```

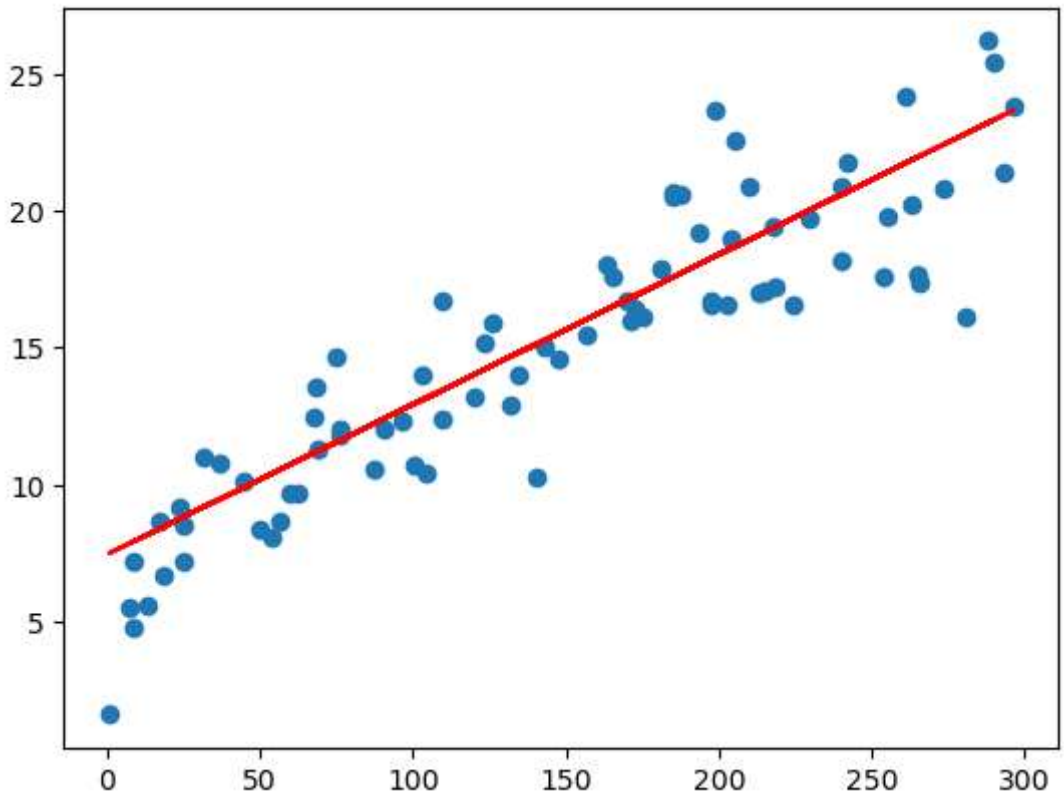
Out[19]:
```
11.250252307
```

In [20]:
```
plt.plot(res)
```

Out[20]:
```
[<matplotlib.lines.Line2D at 0x22a0f577150>]
```



In [26]:
```
plt.scatter(X_test,y_test)
plt.plot(X_test,7.4573254+0.05473199*X_test,'r')
plt.show()
```



In [ ]:
```
```