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

In [2]:
```python
df=pd.read_csv(r'C:\Users\mohan\Downloads\advertising.csv')
```

In [3]:
```python
df.head()
```

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

In [4]:
```python
df.tail()
```

Out[4]:

	TV	Radio	Newspaper	Sales
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 [5]:
```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 4 columns):
 #   Column     Non-Null Count  Dtype  
---  ------     --------------  -----  
 0   TV         200 non-null    float64
 1   Radio      200 non-null    float64
 2   Newspaper  200 non-null    float64
 3   Sales      200 non-null    float64
dtypes: float64(4)
memory usage: 6.4 KB
```

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

Out[6]:

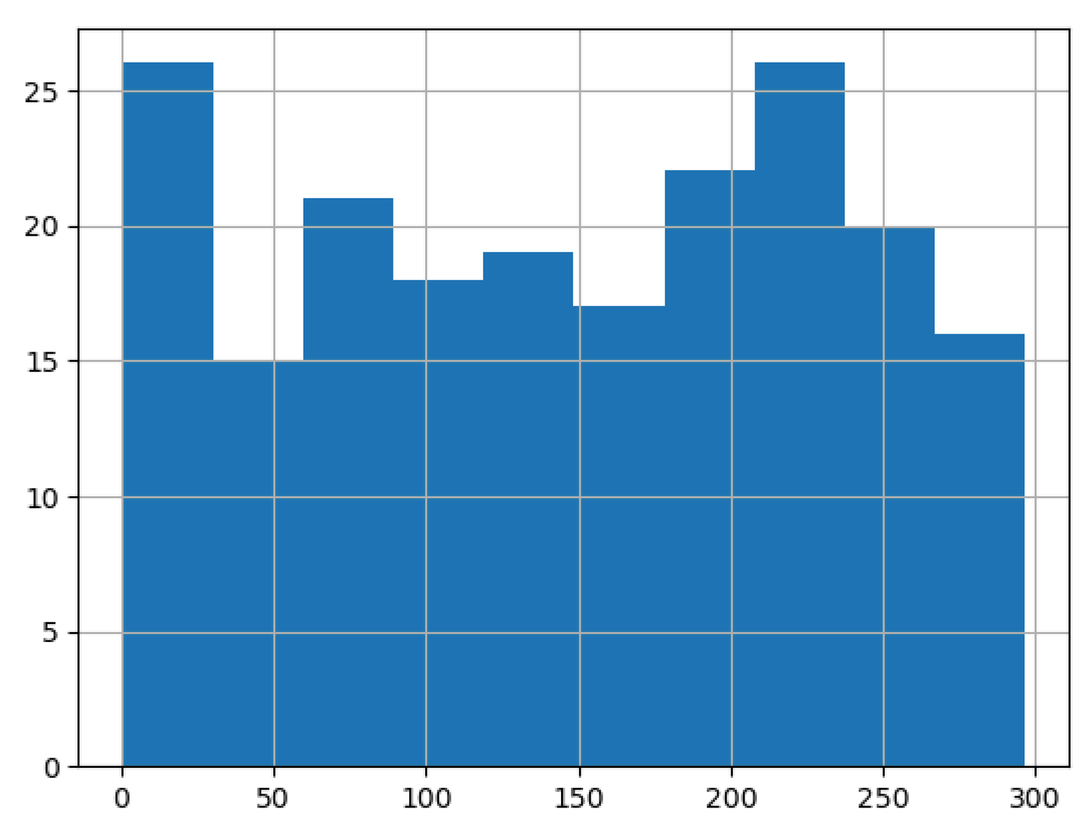
	TV	Radio	Newspaper	Sales
count	200.000000	200.000000	200.000000	200.000000
mean	147.042500	23.264000	30.554000	15.130500
std	85.854236	14.846809	21.778621	5.283892
min	0.700000	0.000000	0.300000	1.600000
25%	74.375000	9.975000	12.750000	11.000000
50%	149.750000	22.900000	25.750000	16.000000
75%	218.825000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

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

Out[7]:
```
TV           0
Radio        0
Newspaper    0
Sales        0
dtype: int64
```

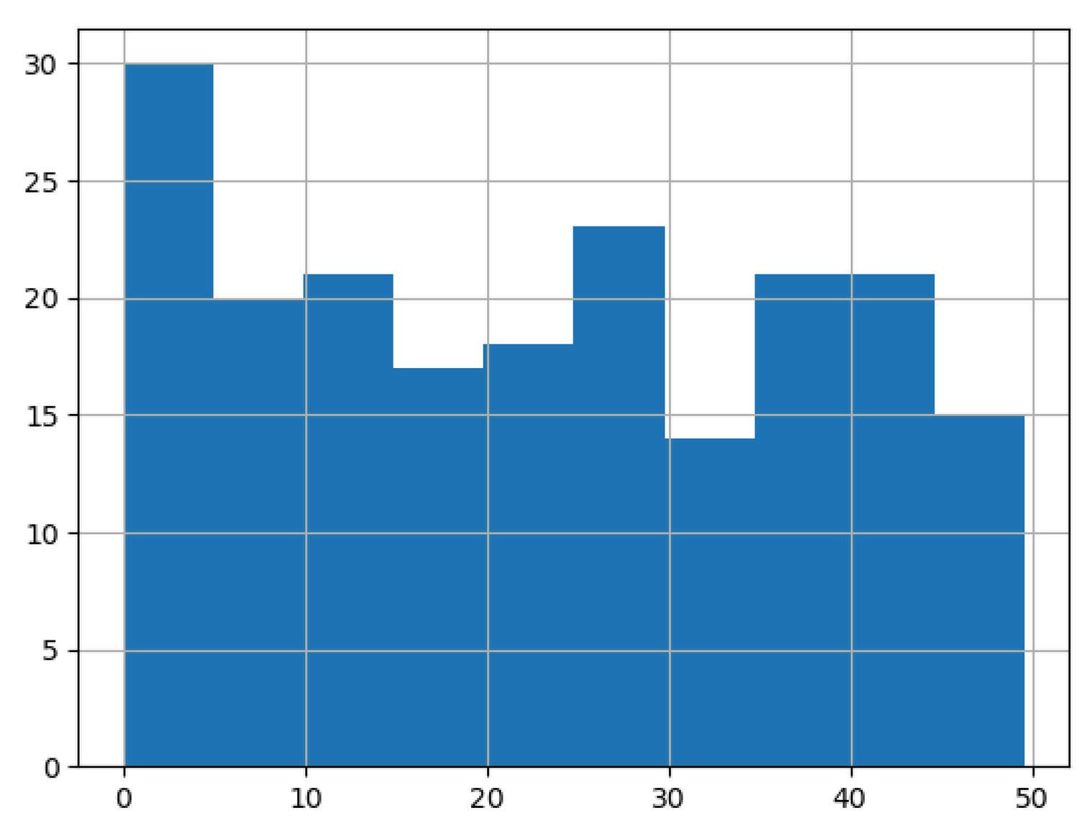
In [8]:
```python
df['TV'].hist()
```

Out[8]: <Axes: >



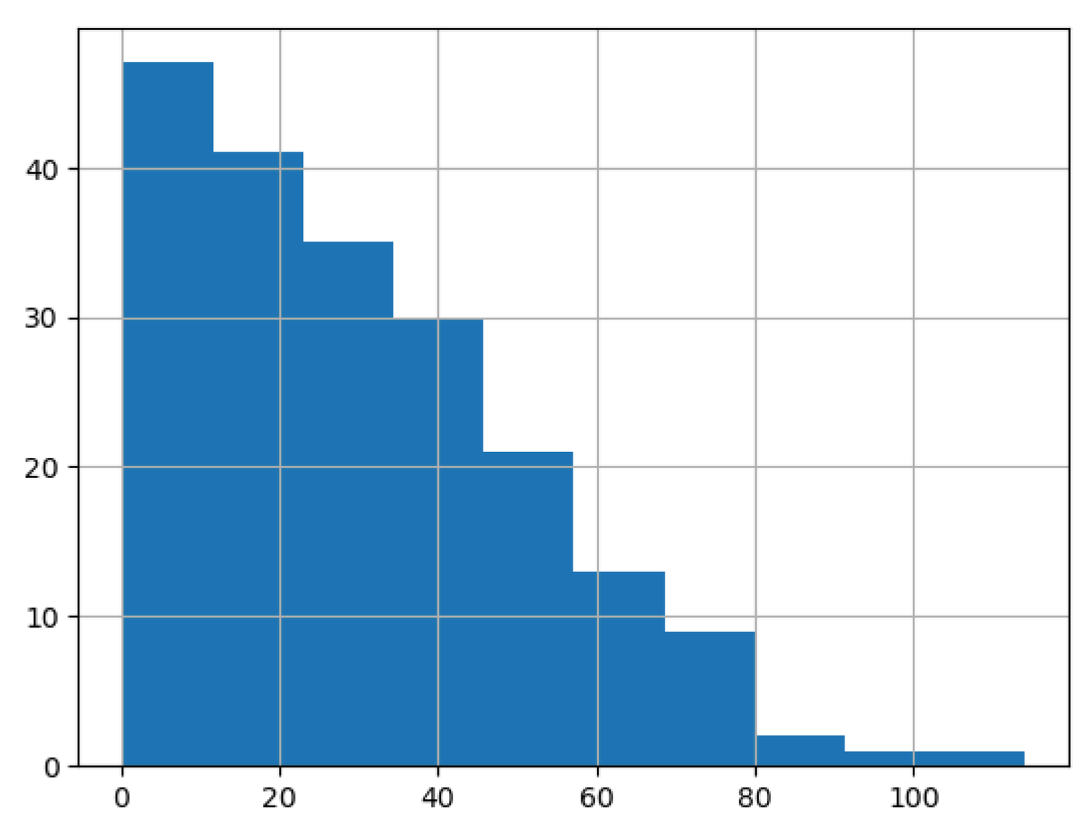
In [9]:
```python
df['Radio'].hist()
```

Out[9]: <Axes: >



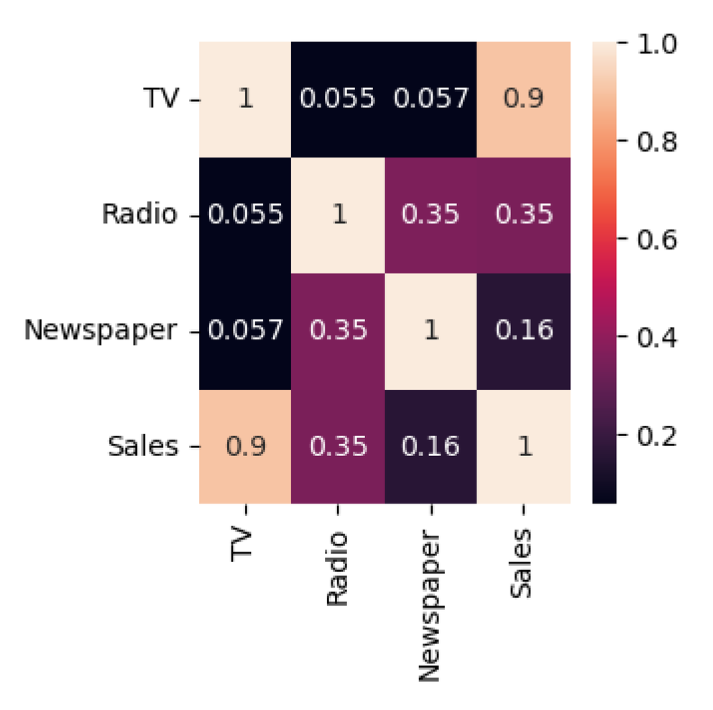
In [10]:
```python
df['Newspaper'].hist()
```

Out[10]: <Axes: >



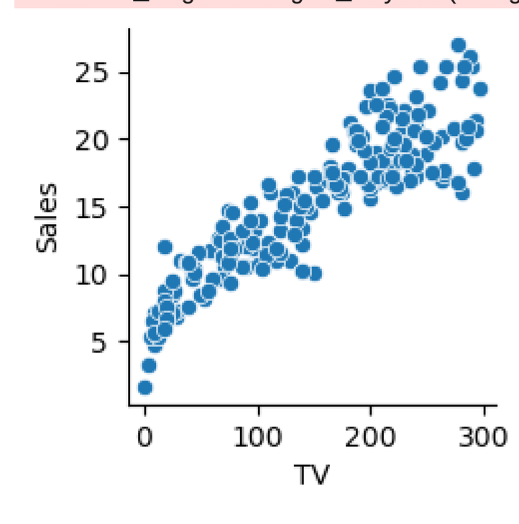
In [11]:
```python
corr=df.corr()
fig,ax=plt.subplots(figsize=(3,3))
sns.heatmap(corr,annot=True)
```

Out[11]: <Axes: >



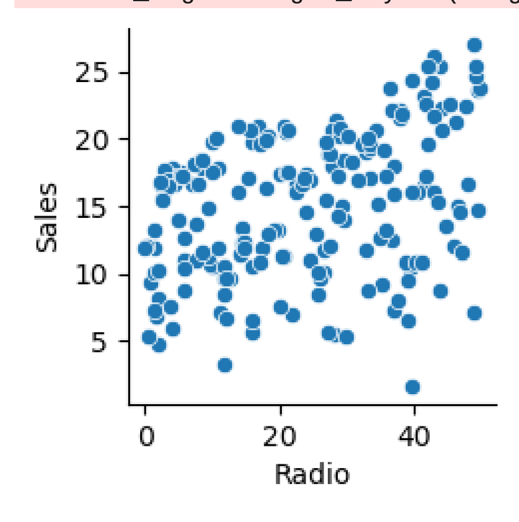
In [12]:
```python
sns.pairplot(df,x_vars=['TV'],y_vars='Sales',kind='scatter')
plt.show()
```

```
C:\Users\mohan\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```



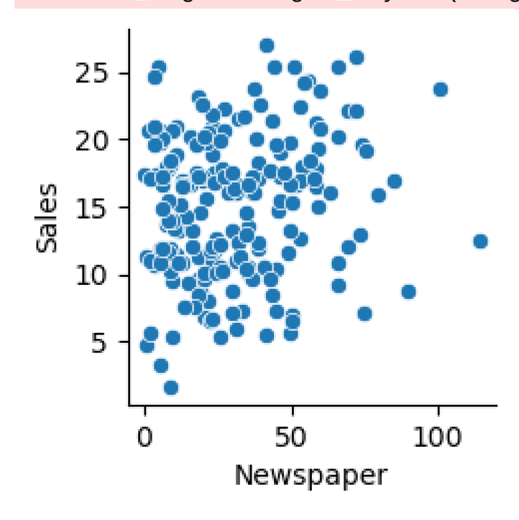
In [13]:
```python
sns.pairplot(df,x_vars='Radio',y_vars='Sales',kind='scatter')
plt.show()
```

```
C:\Users\mohan\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```



In [14]:
```python
sns.pairplot(df,x_vars='Newspaper',y_vars='Sales',kind='scatter')
plt.show()
```

```
C:\Users\mohan\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```



In [15]:
```python
x=df[['TV']]
y=df['Sales']
```

In [16]:
```python
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=0)
```

In [17]:
```python
from sklearn.linear_model import LinearRegression
linear=LinearRegression()
linear.fit(x_train,y_train)
```

Out[17]:
```
▾ LinearRegression
LinearRegression()
```

In [18]:
```python
res=linear.predict(x_test)
```

In [19]:
```python
print(res)
```

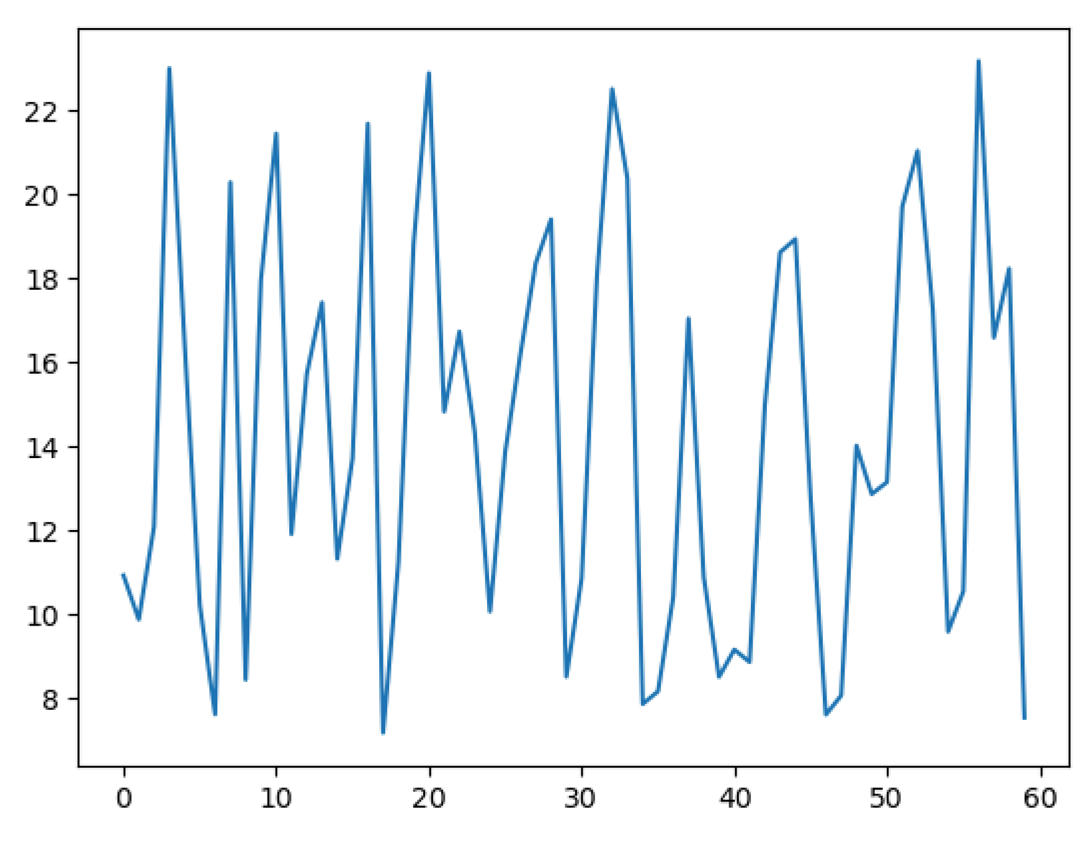
```
[10.93127621  9.88042193 12.09159447 22.99968079 16.45920756 10.21976029
  7.6199906  20.28497391  8.4464437  17.95886418 21.44529217 11.91645209
 15.71485245 17.42249065 11.32534656 13.72260788 21.68063975  7.18213465
 11.23230217 18.82362968 22.88474361 14.82272095 16.72739433 14.35202581
 10.07198391 13.88133066 16.20744039 18.36388094 19.40378881  8.51759529
 10.85465142 18.03001578 22.50709285 20.3725451   7.86628457  8.16731053
 10.40584907 17.03936669 10.88749061  8.51212209  9.16343282  8.86788005
 14.96502414 18.61564811 18.93309367 12.76479799  7.6145174   8.06879294
 14.02363385 12.86878878 13.15339515 19.70481478 21.03480222 17.26376787
  9.59034237 10.55362545 23.17482317 16.58509115 18.22705095  7.54336581]
```

In [20]:
```python
print(y_test)
```

```
18     11.3
170     8.4
107    12.0
98     25.4
177    16.7
182     8.7
5       7.2
146    18.2
12      9.2
152    16.6
61     24.2
125    10.6
180    15.5
154    20.6
80     11.8
7      13.2
33     17.4
130     1.6
37     14.7
74     17.0
183    26.2
145    10.3
45     16.1
159    12.9
60      8.1
123    15.2
179    17.6
185    22.6
122    16.6
44      8.5
16     12.5
55     23.7
150    16.1
111    21.8
22      5.6
189     6.7
129     9.7
4      17.9
83     13.6
106     7.2
134    10.8
66     11.0
26     15.0
113    20.9
168    17.1
63     14.0
8       4.8
75      8.7
118    15.9
143    10.4
71     12.4
124    19.7
184    17.6
97     20.5
149    10.1
24      9.7
30     21.4
160    16.4
40     16.6
56      5.5
Name: Sales, dtype: float64
```

In [21]:
```python
plt.plot(res)
```

Out[21]: [<matplotlib.lines.Line2D at 0x230ea63f8d0>]



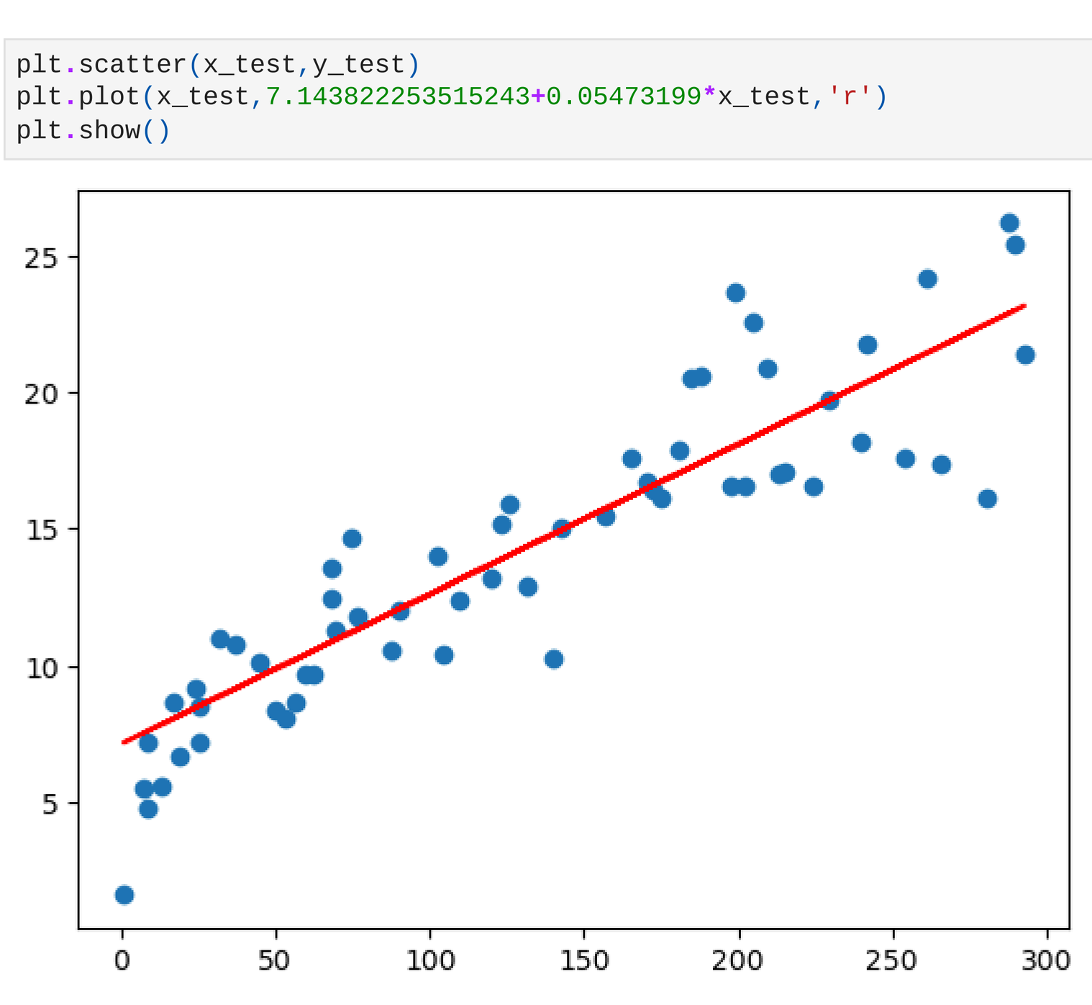
In [22]:
```python
linear.intercept_
```

Out[22]: 7.143822253515243

In [23]:
```python
linear.coef_
```

Out[23]: array([0.05473199])

In [24]:
```python
plt.scatter(x_test,y_test)
plt.plot(x_test,7.143822253515243+0.05473199*x_test,'r')
plt.show()
```



In [ ]:

In [ ]: