

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

%matplotlib inline
```

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

```
[3]: df
```

```
[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

```
[5]: df.drop("Unnamed: 0", axis=1, inplace=True, errors='ignore')
```

```
[6]: df
```

```
[6]:
```

	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

```
[7]: df.describe()
```

```
[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

```
[8]: df.isnull()
```

```
[8]:
```

	TV	Radio	Newspaper	Sales
--	----	-------	-----------	-------

0	False	False	False	False
----------	-------	-------	-------	-------

1	False	False	False	False
----------	-------	-------	-------	-------

2	False	False	False	False
----------	-------	-------	-------	-------

3	False	False	False	False
----------	-------	-------	-------	-------

4	False	False	False	False
----------	-------	-------	-------	-------

...
------------	-----	-----	-----	-----

195	False	False	False	False
------------	-------	-------	-------	-------

196	False	False	False	False
------------	-------	-------	-------	-------

197	False	False	False	False
------------	-------	-------	-------	-------

```
[9]: df.duplicated()
```

```
[9]: 0      False
      1      False
      2      False
      3      False
      4      False
      ...
     195     False
     196     False
     197     False
     198     False
     199     False
      Length: 200, dtype: bool
```

```
[10]: df.isnull().sum()
```

```
[10]: TV          0
      Radio       0
      Newspaper   0
      Sales       0
      dtype: int64
```

```
[11]: df.info()
```

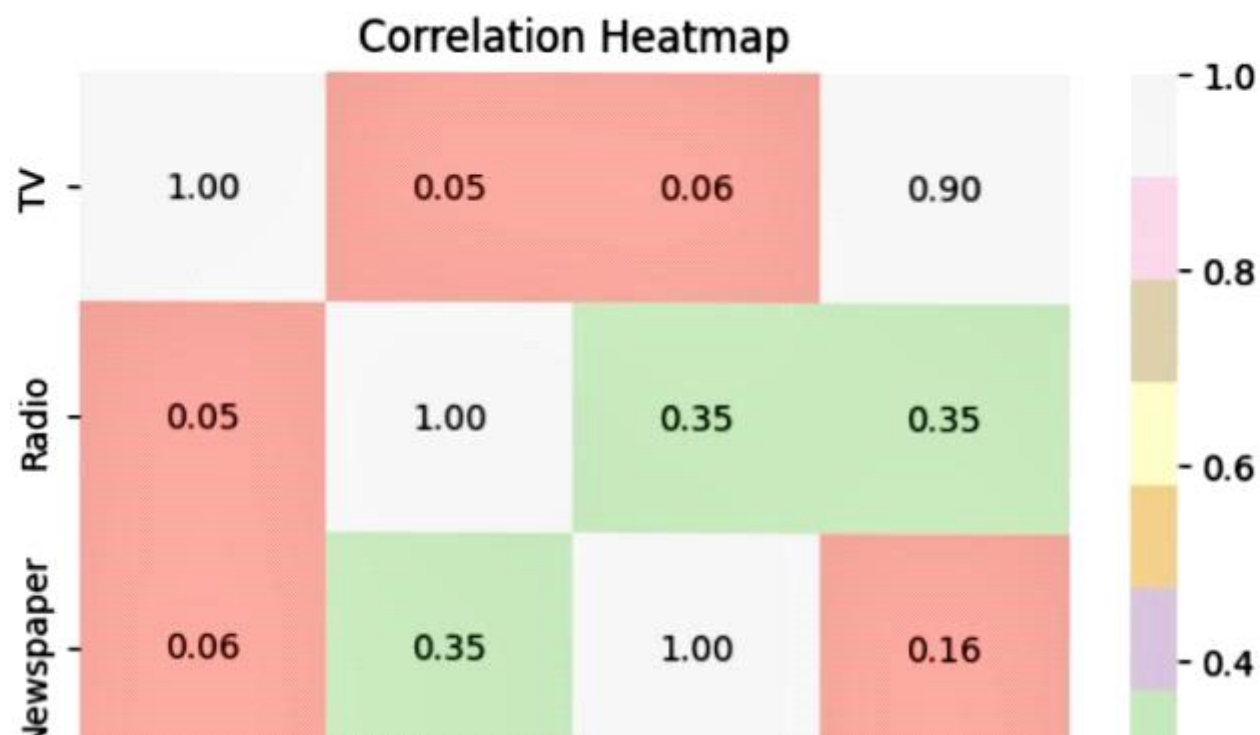
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   TV          200 non-null    float64
1   Radio       200 non-null    float64
2   Newspaper   200 non-null    float64
3   Sales       200 non-null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

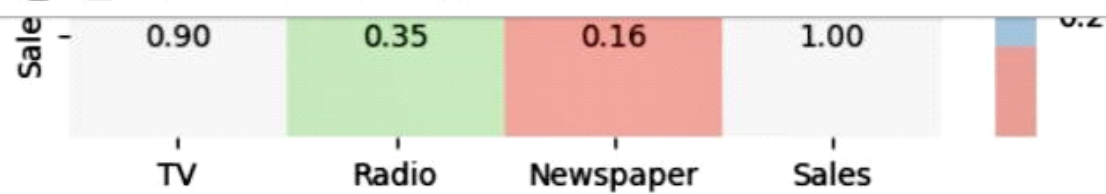
```
[12]: df.corr()
```

```
[12]:
```

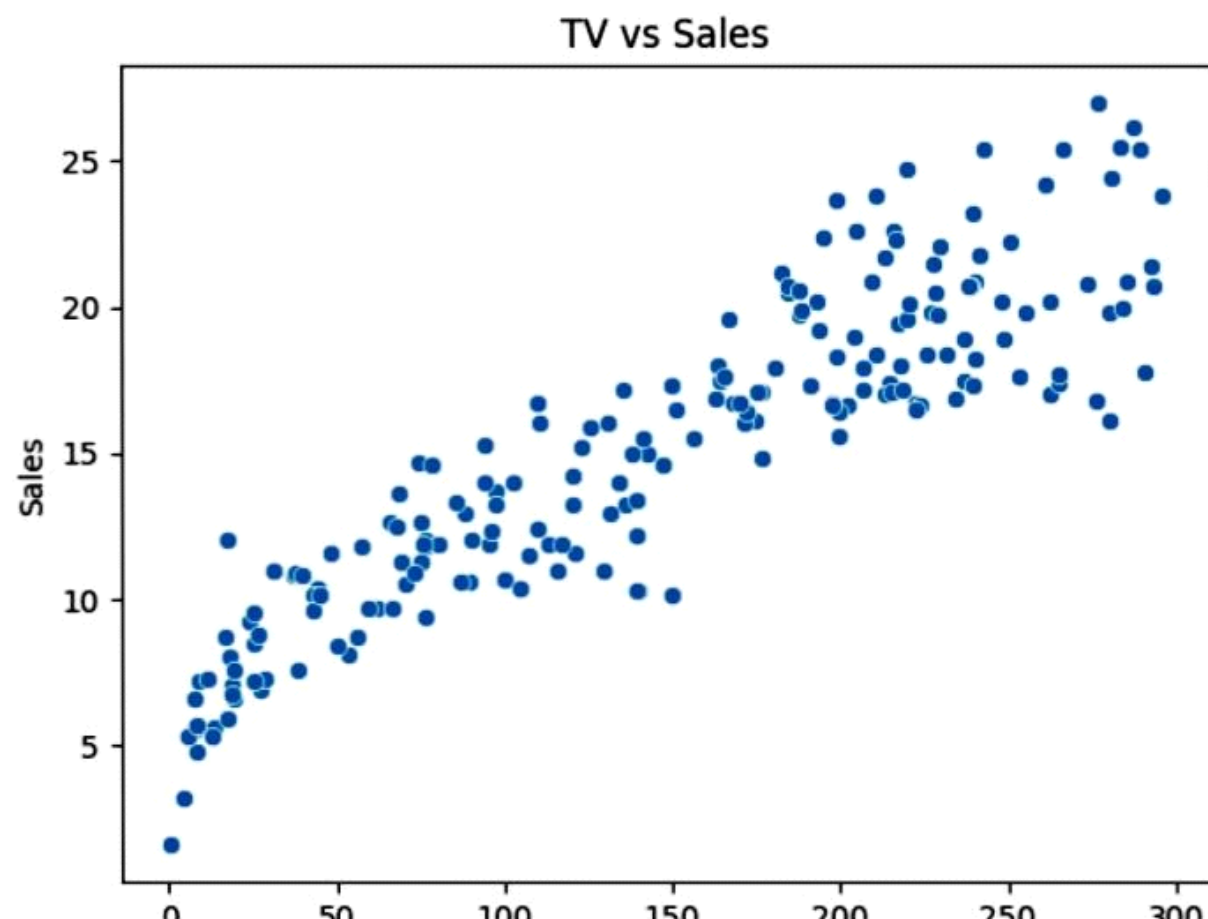
	TV	Radio	Newspaper	Sales
TV	1.000000	0.054809	0.056648	0.901208
Radio	0.054809	1.000000	0.354104	0.349631
Newspaper	0.056648	0.354104	1.000000	0.157960
Sales	0.901208	0.349631	0.157960	1.000000

```
[13]: correlation_matrix = df.corr()  
sns.heatmap(correlation_matrix, annot=True, cmap='Pastel1', fmt='.2f')  
plt.title('Correlation Heatmap')  
plt.show()
```

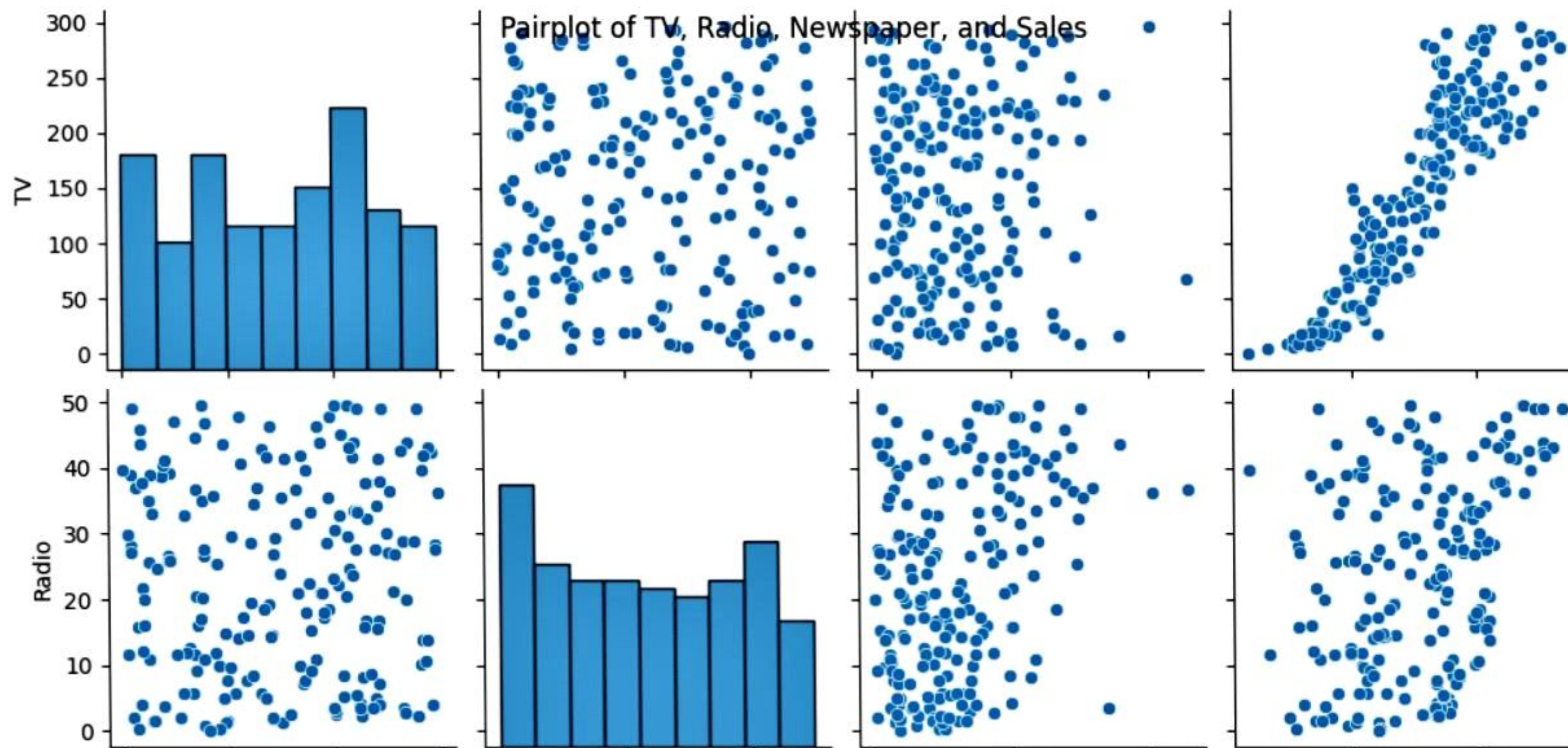


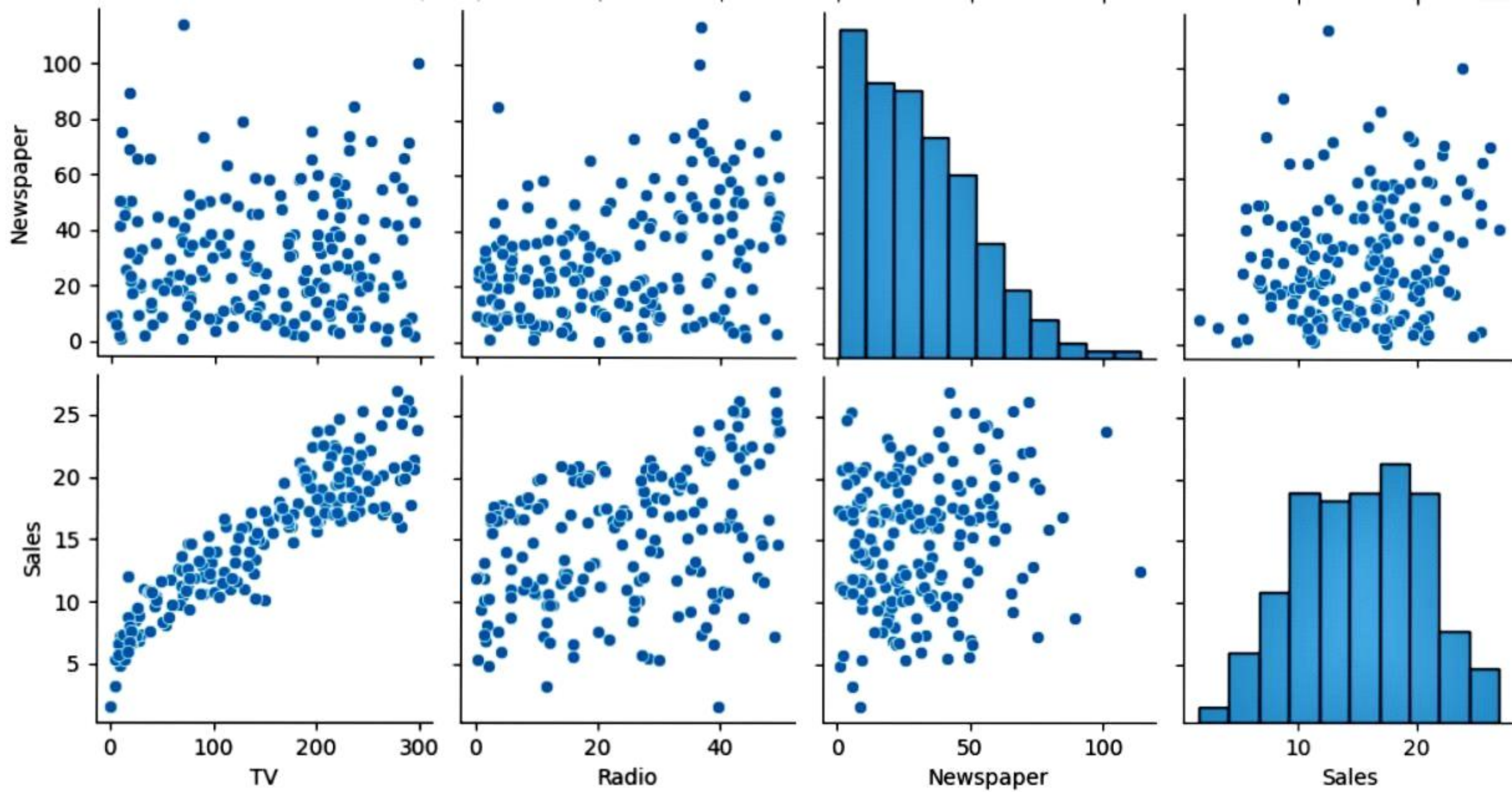


```
[14]: sns.scatterplot(x='TV', y='Sales', data=df)
plt.title('TV vs Sales')
plt.xlabel('TV Advertising Budget')
plt.ylabel('Sales')
plt.show()
```

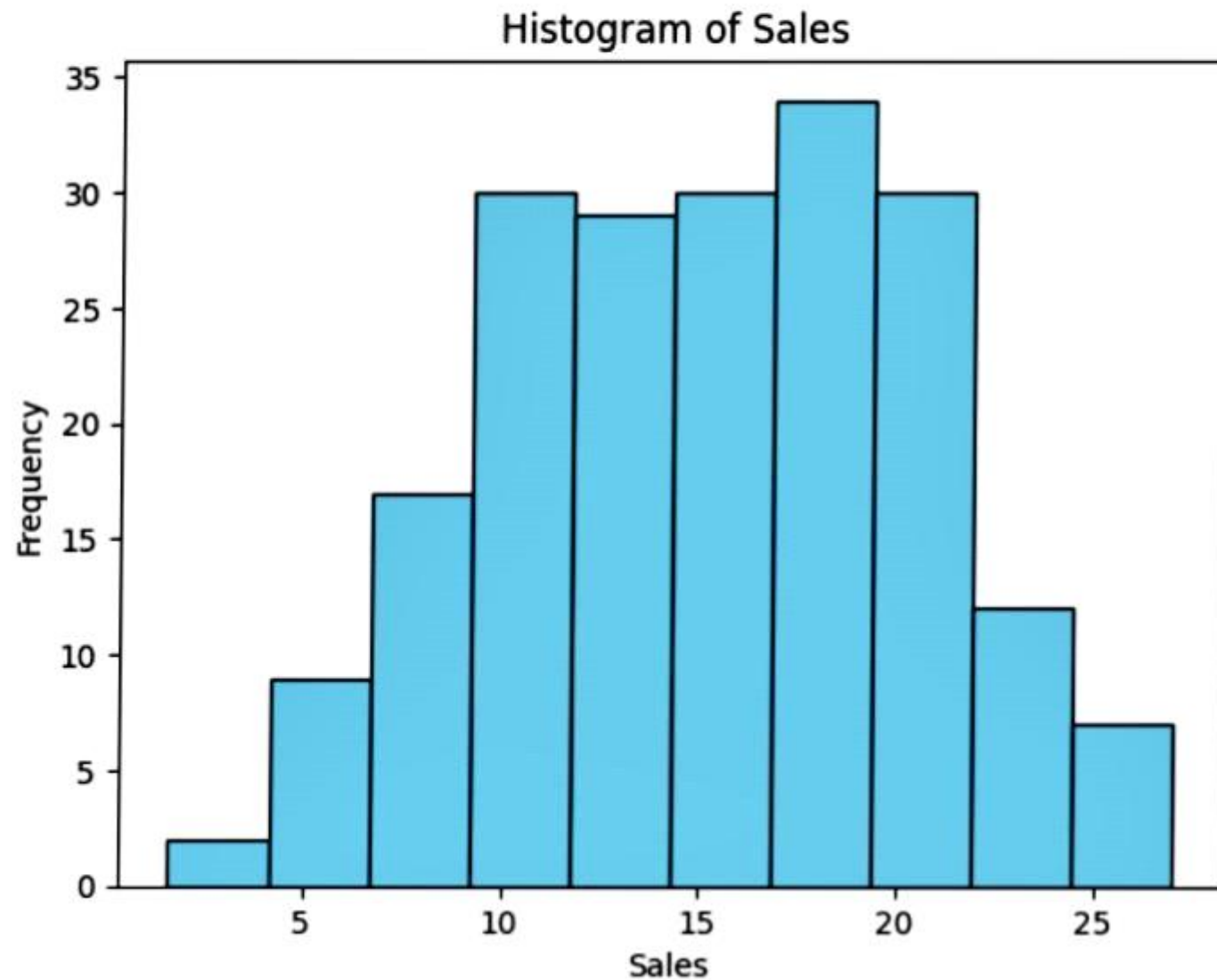



```
[15]: sns.pairplot(df)
plt.suptitle('Pairplot of TV, Radio, Newspaper, and Sales')
plt.show()
```



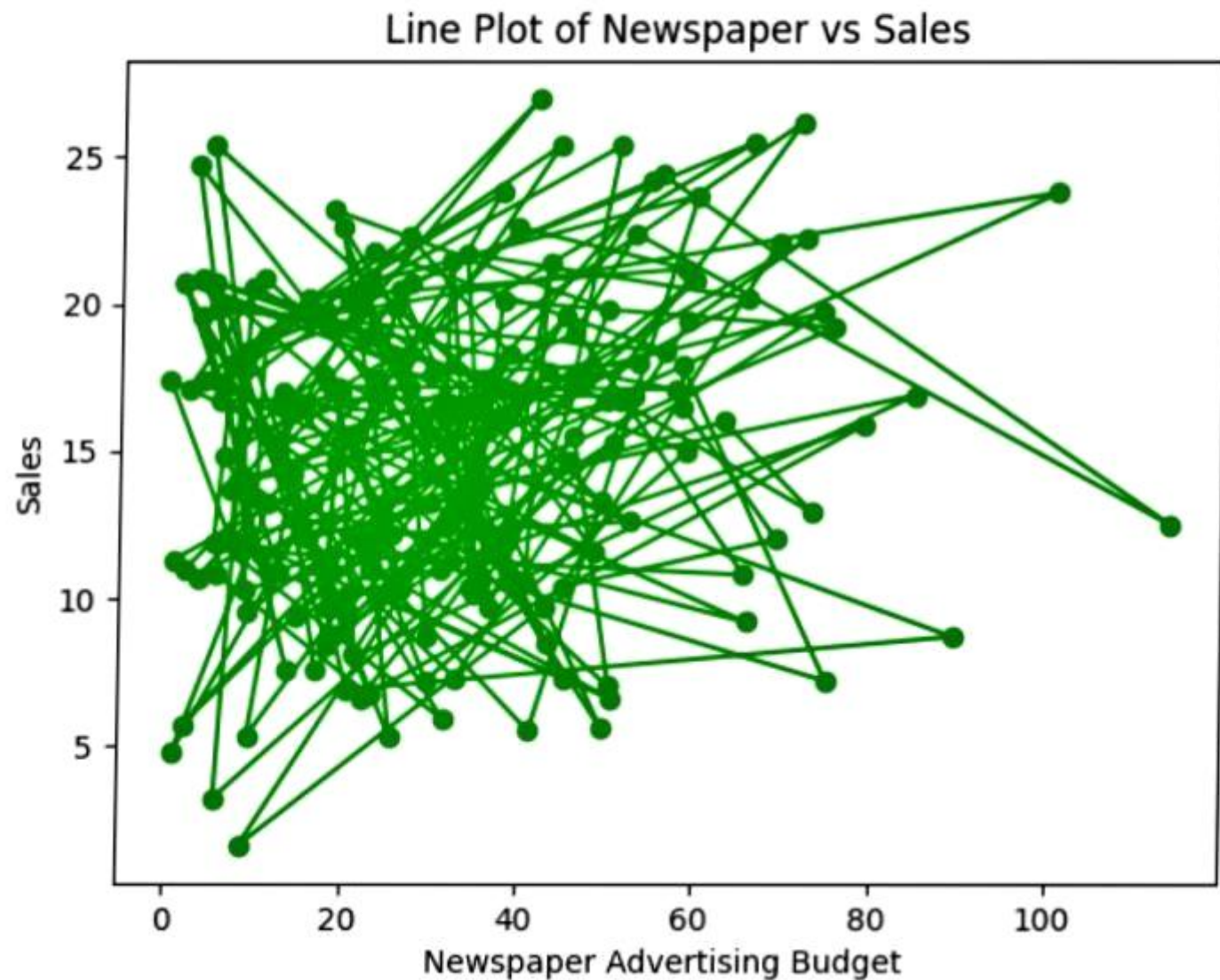



```
[16]: plt.hist(df['Sales'], bins=10, color='skyblue', edgecolor='black')
plt.title('Histogram of Sales')
plt.xlabel('Sales')
plt.ylabel('Frequency')
plt.show()
```



Sales

```
[17]: plt.plot(df['Newspaper'], df['Sales'], marker='o', linestyle='-', color='green')
plt.title('Line Plot of Newspaper vs Sales')
plt.xlabel('Newspaper Advertising Budget')
plt.ylabel('Sales')
plt.show()
```



```
[18]: plt.hexbin(df['TV'], df['Radio'], gridsize=15, cmap='Blues')
plt.title('Hexbin Plot of TV and Radio Advertising Budgets')
plt.xlabel('TV Advertising Budget')
plt.ylabel('Radio Advertising Budget')
plt.colorbar(label='Count')
plt.show()
```

