In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
```

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

Out[2]:

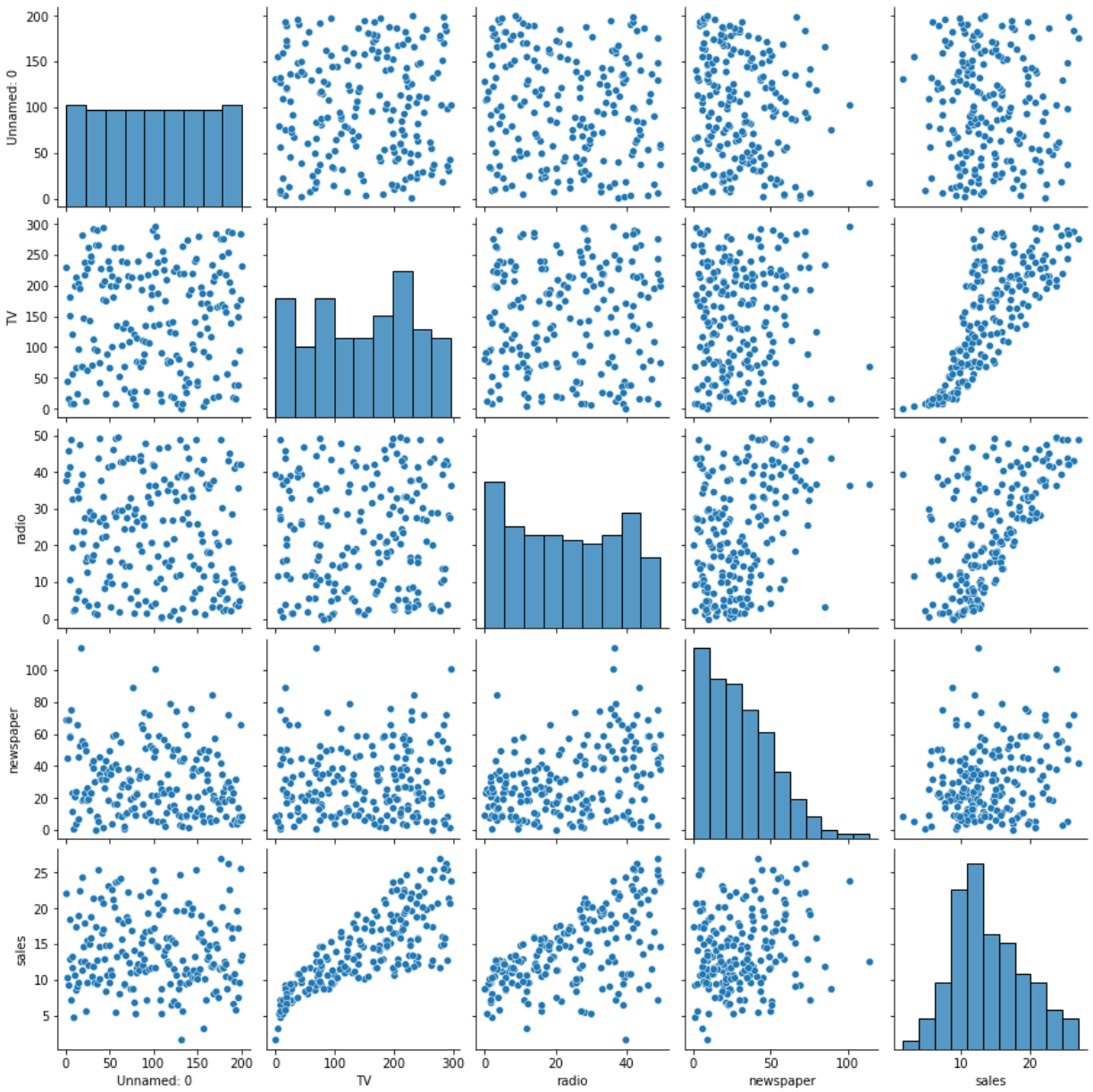
	Unnamed: 0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9

In [3]:
```
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
```

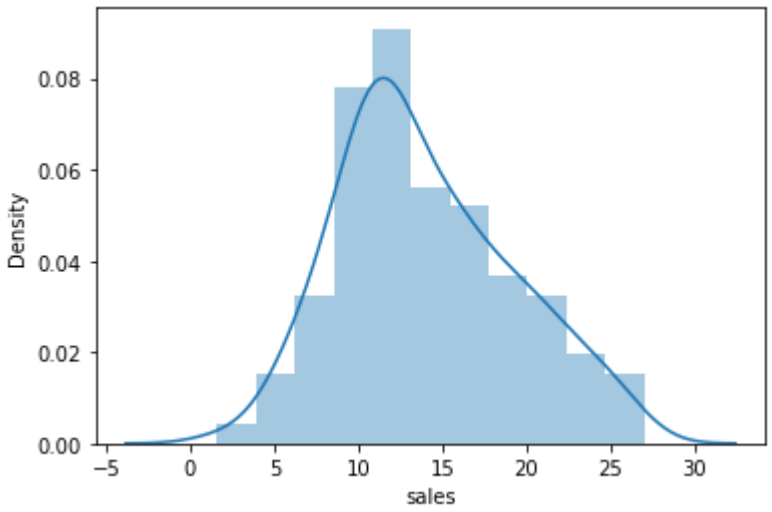
In [4]:
```
sns.pairplot(df)
```

Out[4]:
```
<seaborn.axisgrid.PairGrid at 0x8ac94f6c40>
```



In [5]:
```
sns.distplot(df["sales"])
```

Out[5]:
```
<AxesSubplot:xlabel='sales', ylabel='Density'>
```



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

Out[6]:

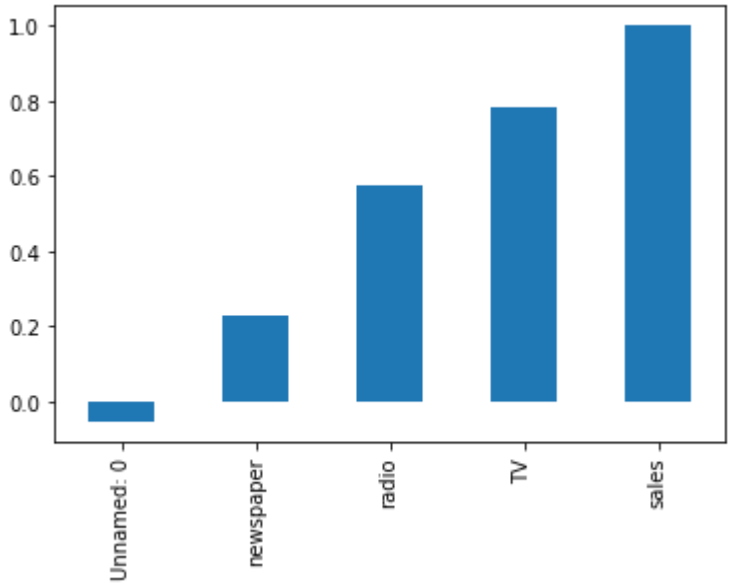
	Unnamed: 0	TV	radio	newspaper	sales
Unnamed: 0	1.000000	0.017715	-0.110680	-0.154944	-0.051616
TV	0.017715	1.000000	0.054809	0.056648	0.782224
radio	-0.110680	0.054809	1.000000	0.354104	0.576223
newspaper	-0.154944	0.056648	0.354104	1.000000	0.228299
sales	-0.051616	0.782224	0.576223	0.228299	1.000000

In [7]:
```
df.corr()["sales"].value_counts()
```

Out[7]:
```
 1.000000    1
 0.576223    1
-0.051616    1
 0.782224    1
 0.228299    1
Name: sales, dtype: int64
```

In [8]:
```
df.corr()["sales"].sort_values().plot(kind="bar")
```

Out[8]: <AxesSubplot:>



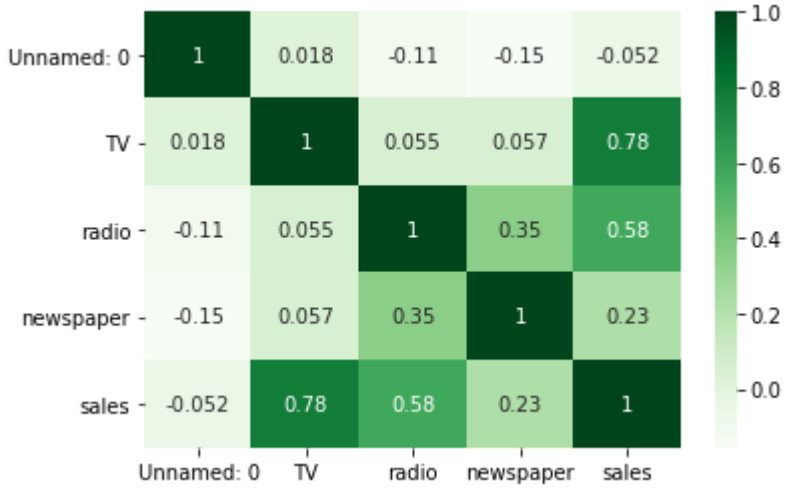
In [9]:
```
df.corr().style.background_gradient()
```

Out[9]:

	Unnamed: 0	TV	radio	newspaper	sales
Unnamed: 0	1.000000	0.017715	-0.110680	-0.154944	-0.051616
TV	0.017715	1.000000	0.054809	0.056648	0.782224
radio	-0.110680	0.054809	1.000000	0.354104	0.576223
newspaper	-0.154944	0.056648	0.354104	1.000000	0.228299
sales	-0.051616	0.782224	0.576223	0.228299	1.000000

In [10]:
```
sns.heatmap(df.corr(), annot=True, cmap="Greens")
```

Out[10]: <AxesSubplot:>



In [11]:
```
x = df.iloc[:, 1:-1].values
y = df.iloc[:, -1].values
```

In [12]:
```
x
```

Out[12]:
```
array([[230.1,  37.8,  69.2],
       [ 44.5,  39.3,  45.1],
       [ 17.2,  45.9,  69.3],
       [151.5,  41.3,  58.5],
       [180.8,  10.8,  58.4],
       [  8.7,  48.9,  75. ],
       [ 57.5,  32.8,  23.5],
       [120.2,  19.6,  11.6],
       [  8.6,   2.1,   1. ],
       [199.8,   2.6,  21.2],
       [ 66.1,   5.8,  24.2],
       [214.7,  24. ,   4. ],
       [ 23.8,  35.1,  65.9],
       [ 97.5,   7.6,   7.2],
       [204.1,  32.9,  46. ],
       [195.4,  47.7,  52.9],
       [ 67.8,  36.6, 114. ],
       [281.4,  39.6,  55.8],
       [ 69.2,  20.5,  18.3],
```

In [13]:
```
y
```

Out[13]:
```
array([22.1, 10.4,  9.3, 18.5, 12.9,  7.2, 11.8, 13.2,  4.8, 10.6,  8.6,
       17.4,  9.2,  9.7, 19. , 22.4, 12.5, 24.4, 11.3, 14.6, 18. , 12.5,
        5.6, 15.5,  9.7, 12. , 15. , 15.9, 18.9, 10.5, 21.4, 11.9,  9.6,
       17.4,  9.5, 12.8, 25.4, 14.7, 10.1, 21.5, 16.6, 17.1, 20.7, 12.9,
        8.5, 14.9, 10.6, 23.2, 14.8,  9.7, 11.4, 10.7, 22.6, 21.2, 20.2,
       23.7,  5.5, 13.2, 23.8, 18.4,  8.1, 24.2, 15.7, 14. , 18. ,  9.3,
        9.5, 13.4, 18.9, 22.3, 18.3, 12.4,  8.8, 11. , 17. ,  8.7,  6.9,
       14.2,  5.3, 11. , 11.8, 12.3, 11.3, 13.6, 21.7, 15.2, 12. , 16. ,
       12.9, 16.7, 11.2,  7.3, 19.4, 22.2, 11.5, 16.9, 11.7, 15.5, 25.4,
       17.2, 11.7, 23.8, 14.8, 14.7, 20.7, 19.2,  7.2,  8.7,  5.3, 19.8,
       13.4, 21.8, 14.1, 15.9, 14.6, 12.6, 12.2,  9.4, 15.9,  6.6, 15.5,
        7. , 11.6, 15.2, 19.7, 10.6,  6.6,  8.8, 24.7,  9.7,  1.6, 12.7,
        5.7, 19.6, 10.8, 11.6,  9.5, 20.8,  9.6, 20.7, 10.9, 19.2, 20.1,
       10.4, 11.4, 10.3, 13.2, 25.4, 10.9, 10.1, 16.1, 11.6, 16.6, 19. ,
       15.6,  3.2, 15.3, 10.1,  7.3, 12.9, 14.4, 13.3, 14.9, 18. , 11.9,
       11.9,  8. , 12.2, 17.1, 15. ,  8.4, 14.5,  7.6, 11.7, 11.5, 27. ,
       20.2, 11.7, 11.8, 12.6, 10.5, 12.2,  8.7, 26.2, 17.6, 22.6, 10.3,
       17.3, 15.9,  6.7, 10.8,  9.9,  5.9, 19.6, 17.3,  7.6,  9.7, 12.8,
       25.5, 13.4])
```

In [14]:
```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error as mse, r2_score
```

In [15]:
```
xtrain, xtest, ytrain, ytest = train_test_split(x,y, test_size=0.25, random_state
linreg = LinearRegression()
linreg.fit(xtrain, ytrain)
ypred = linreg.predict(xtest)
```

In [16]:
```
linreg.coef_
```

Out[16]:
```
array([0.04656457, 0.17915812, 0.00345046])
```

In [17]:
```
linreg.intercept_
```

Out[17]:
```
2.876966622317923
```

In [18]:
```
x = df.iloc[:, 1:-1]
y = df.iloc[:, -1]
```

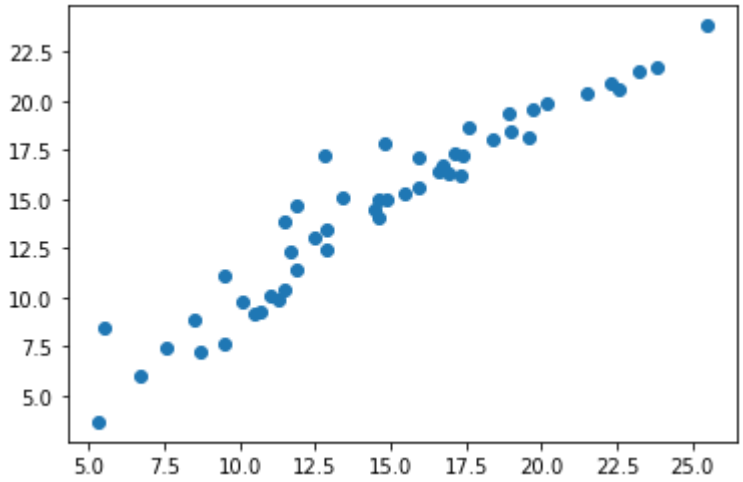
In [19]:
```
coef_df = pd.DataFrame(linreg.coef_, x.columns, columns=["Coefficiet"])
coef_df
```

Out[19]:

	Coefficiet
TV	0.046565
radio	0.179158
newspaper	0.003450

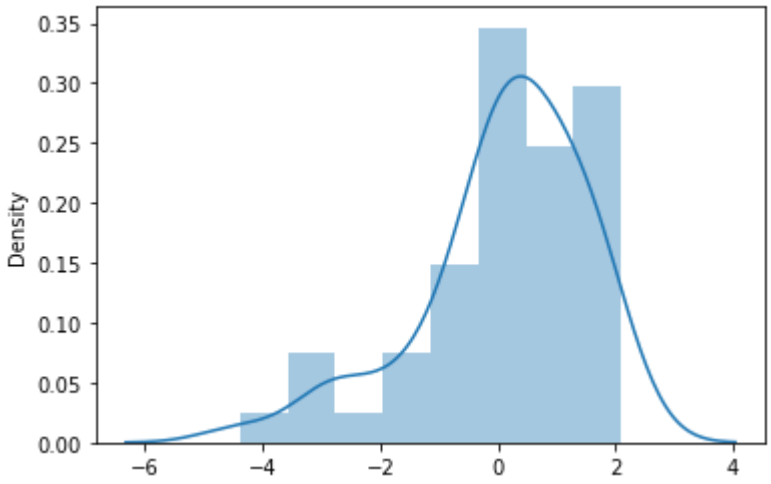
In [20]:
```
plt.scatter(ytest, ypred)
```

Out[20]:
```
<matplotlib.collections.PathCollection at 0x8ace220ee0>
```



In [21]:
```
sns.distplot((ytest-ypred))
```

Out[21]:
```
<AxesSubplot:ylabel='Density'>
```



In [22]:
```
print(mse(ytest, ypred))
print(np.sqrt(mse(ytest, ypred)))
print(r2_score(ytest, ypred))
```

```
1.973045620228338
1.4046514230328953
0.9156213613792232
```

In [23]:
```
df.head()
```

Out[23]:

	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 [24]:
```
x = df.iloc[:,1:-2]
y = df.iloc[:, -1]

xtrain, xtest, ytrain, ytest =train_test_split(x,y, test_size=0.25, random_state=

linreg = LinearRegression()

linreg.fit(xtrain, ytrain)

ypred = linreg.predict(xtest)

print(mse(ytest, ypred))
print(np.sqrt(mse(ytest, ypred)))
print(r2_score(ytest, ypred))
```

```
1.9262760418667417
1.3879034699382886
0.9176214942248908
```

In [25]:
```
coef_df = pd.DataFrame(linreg.coef_, x.columns, columns=["Coefficiet"])
coef_df
```

Out[25]:

	Coefficiet
TV	0.046602
radio	0.181180

In [ ]:
```

```