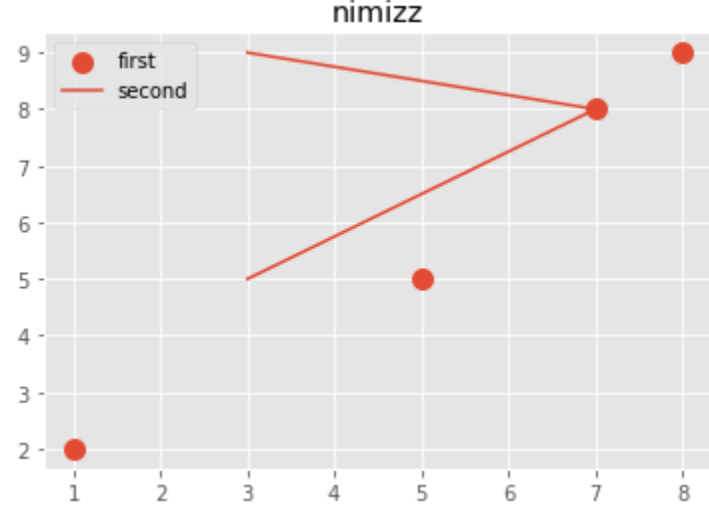


```
In [1]: pip install matplotlib

Requirement already satisfied: matplotlib in c:\users\mujjj\desktop\anaconda\new folder\new folder\pythonanac
onda\lib\site-packages (3.5.1)
Requirement already satisfied: pyparsing>=2.2.1 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pyth
onanaconda\lib\site-packages (from matplotlib) (3.0.4)
Requirement already satisfied: packaging>=20.0 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pytho
nanaconda\lib\site-packages (from matplotlib) (21.3)
Requirement already satisfied: pillow>=6.2.0 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pytho
nanaconda\lib\site-packages (from matplotlib) (9.0.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pyt
honanaconda\lib\site-packages (from matplotlib) (4.25.0)
Requirement already satisfied: numpy>=1.17 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pythonana
conda\lib\site-packages (from matplotlib) (1.21.5)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\mujjj\desktop\anaconda\new folder\new folder
\pythonanaconda\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: cycler>=0.10 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pythonan
aconda\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pyt
honanaconda\lib\site-packages (from matplotlib) (1.3.2)
Requirement already satisfied: six>=1.5 in c:\users\mujjj\desktop\anaconda\new folder\new folder\pythonanacon
da\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

```
In [1]: %matplotlib inline
from matplotlib import pyplot as plt
```

```
In [2]: from matplotlib import pyplot as plt
from matplotlib import style
style.use("ggplot")
x=[1,5,7,8]
y=[2,5,8,9]
x2=[3,7,3]
y2=[5,8,9]
plt.scatter(x,y, label="first", linewidth=5)
plt.plot(x2,y2, label="second")
plt.title("nimizz")
plt.legend()
plt.show()
```



statistics for data analyst

```
In [3]: import seaborn as sns
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
%matplotlib inline
```

```
In [4]: import statistics
```

```
In [5]: sns.__version__
```

```
Out[5]: '0.11.2'
```

```
In [6]: sns.get_dataset_names()
```

```
Out[6]: ['anagrams',
'anscombe',
'attention',
'brain_networks',
'car_crashes',
'diamonds',
'dots',
'dowjones',
'exercise',
'flights',
'fmri',
'geyser',
'glue',
'healthexp',
'iris',
'mpg',
'penguins',
'planets',
'seaice',
'taxis',
'tips',
'titanic']
```

```
In [7]: df=sns.load_dataset('tips')
df.head()
```

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

```
In [8]: df.shape
```

```
Out[8]: (244, 7)
```

```
In [9]: np.mean(df['total_bill'])
```

```
Out[9]: 19.785942622950824
```

```
In [10]: np.median(df['total_bill'])
```

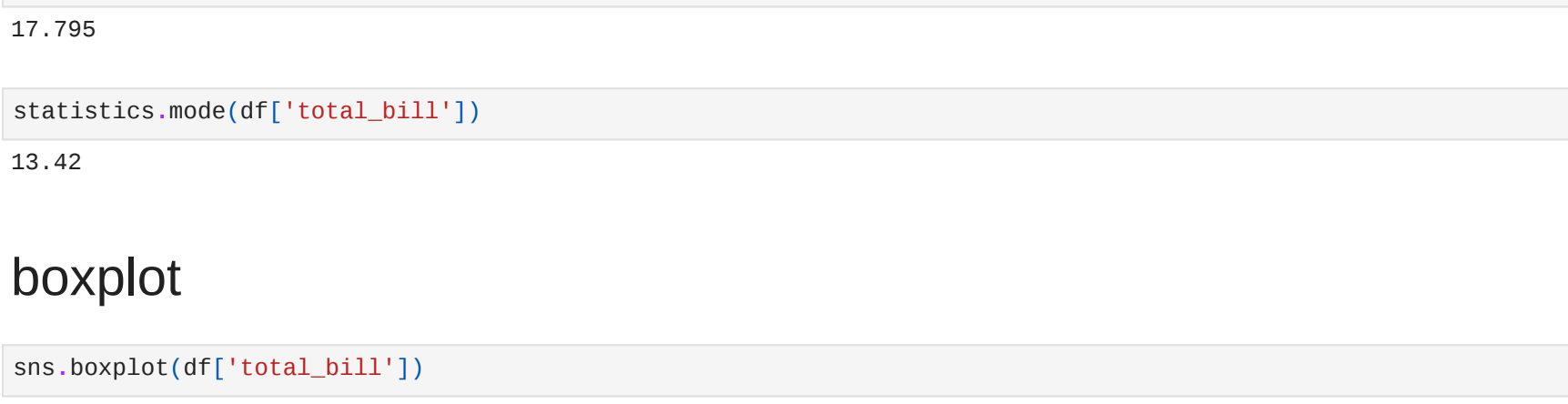
```
Out[10]: 17.795
```

```
In [11]: statistics.mode(df['total_bill'])
```

```
Out[11]: 13.42
```

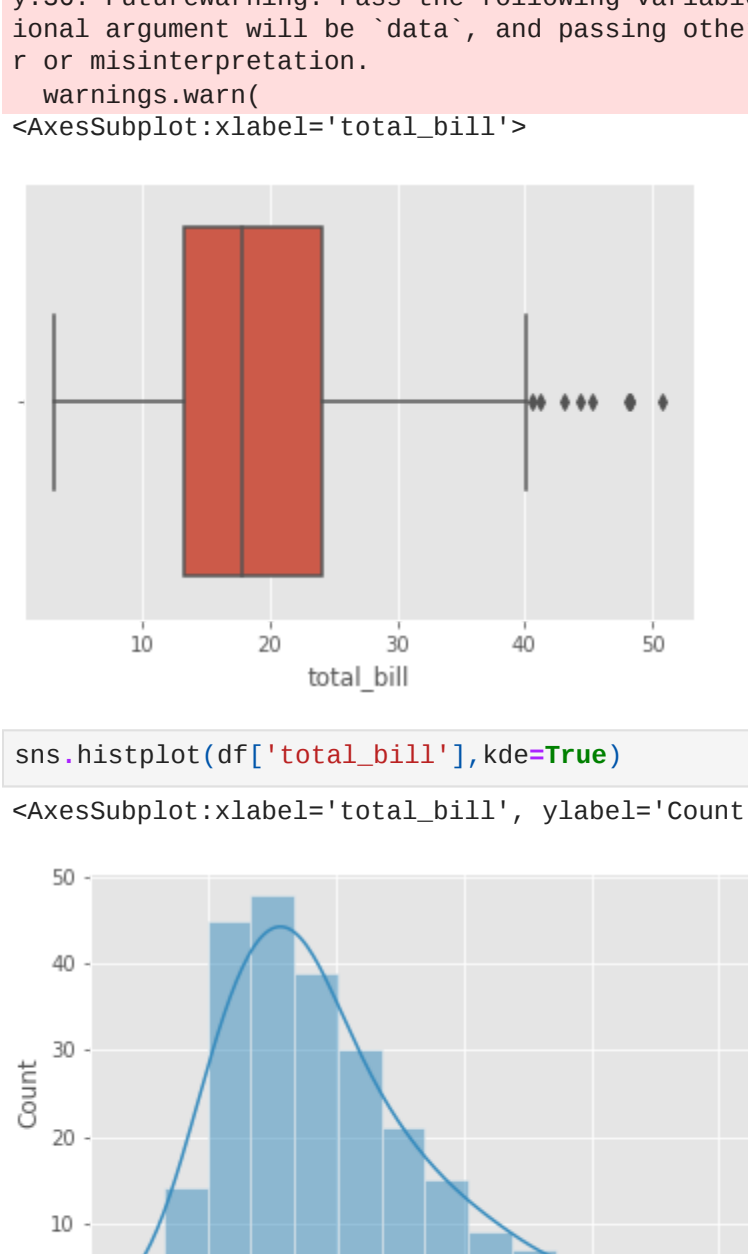
boxplot

```
In [12]: sns.boxplot(df['total_bill'])
```

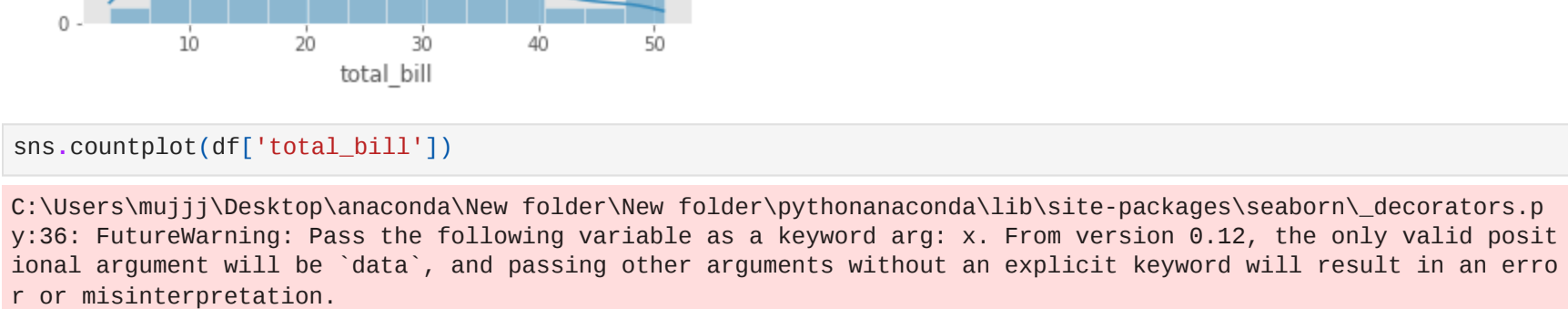


```
In [13]: sns.histplot(df['total_bill'],kde=True)
```

```
Out[13]: <AxesSubplot:xlabel='total_bill', ylabel='Count'>
```



```
In [14]: sns.countplot(df['total_bill'])
```



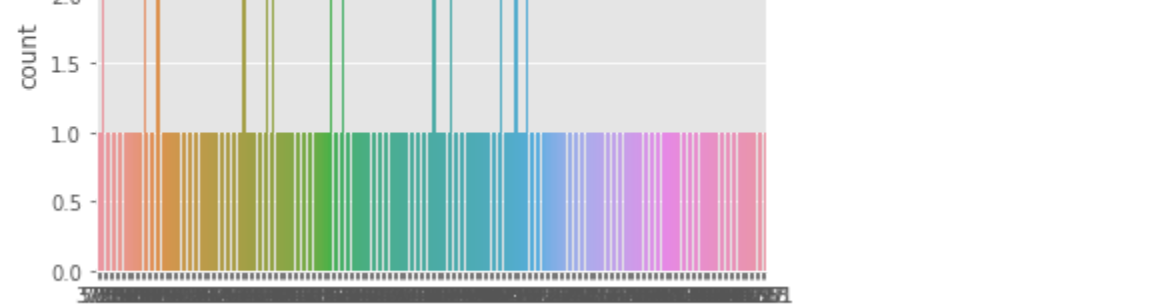
```
In [15]: np.percentile(df['size'],[25,75])
```

```
Out[15]: array([2., 3.])
```

outliers & outliers remove

```
In [16]: dataset=[2,3,5,7,8,9,23,34,65,78,95,34,67,20,65,11,22,56,75,98,25,74,67,34,75]
```

```
In [17]: plt.hist(dataset)
```



```
In [ ]:
```

```
In [18]: outliers=[]
def detect_outliers(data):
    threshold=3
    mean=np.mean(data)
    std=np.std(data)
    for i in dataset:
        z_score=(i-mean)/std
        if np.abs(z_score)>threshold:
            outliers.append
    return outliers

In [19]: detect_outliers(dataset)
```

iqr

```
In [20]: dataset= sorted(dataset)
```

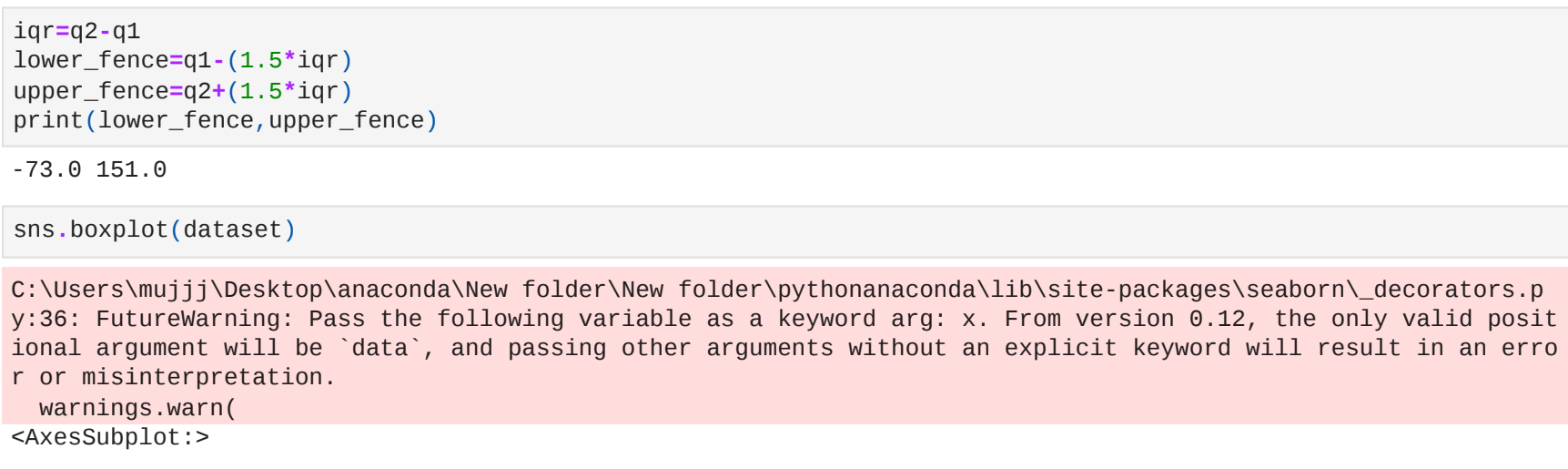
```
In [21]: q1,q2=np.percentile(dataset,[25,75])
print(q1,q2)

11.0 67.0
```

```
In [22]: iqr=q2-q1
lower_fence=q1-(1.5*iqr)
upper_fence=q2+(1.5*iqr)
print(lower_fence,upper_fence)

-73.0 151.0
```

```
In [23]: sns.boxplot(dataset)
```



z_test,p_value

```
In [24]: from statsmodels.stats.weightstats import ztest
data=[88,92,94,94,97,97,105,109,109,109,108,110,112,114,115]
ztest(data,value=110)
```

```
Out[24]: (-2.815610073618764, 0.004868471598155648)
```

```
In [25]: import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [ ]: d
```