**Five Number Summary**

Based on the below five points we can summarize the given distribution list and can discard the outlier.

1. Minimum value in the distribution
2. First Quartile (Q1)
3. Median
4. Third Quartile(Q3)
5. Maximum Value in the distribution

Let’s consider the example:-

Say a distribution list {1, 2, 2, 2, 3, 3, 3, 4, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 9, 21}

Here the size of the distribution list is N= 20

Lower fence >>>>>> Higher fence

**Minimum Value**= 1

**First Quartile(Q1/25 percentile)-**

Q1= (25/100)\*(N+1)

Q1= (25/100)\*21

=5.25 is the 5th index value =3

**Median-**

The size of the distribution list is 20, thus the median would be (10th + 11th position index value)/2

Thus it would be (5+5)/2=5

**Third Quartile(Q3/75 percentile)-**

Q3= (75/100)\*(N+1)

= (0.75\*21)

=15.75 is the average of 15th and 16th index is (7+8)/2=7.5

**Maximum value**= 9

Inter Quartile range (IQR) =Q3-Q1

=7.5-3

=4.5

Lower Fence = Q1 – 1.5 (IQR)

= 3-1.5(4.5)

= -3.65

Higher Fence = Q3+1.5(IQR)

=7.5+1.5(4.5)

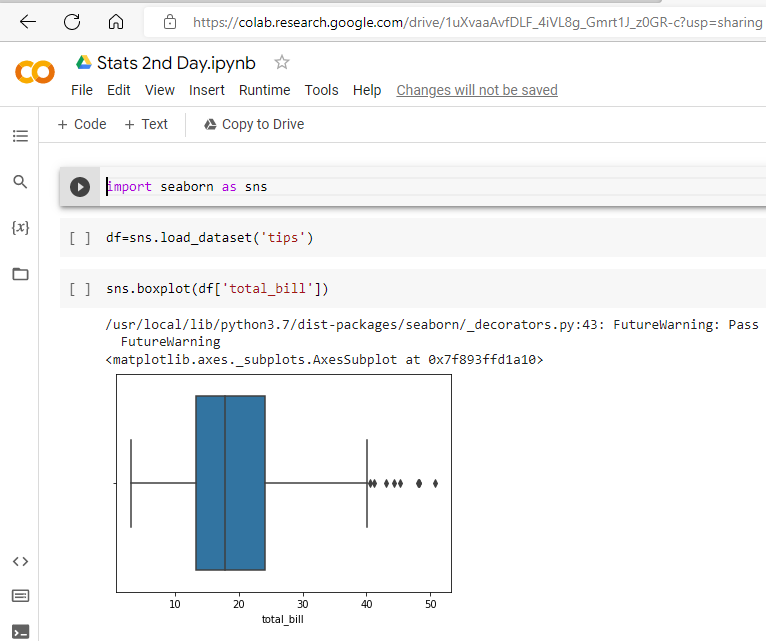
= 14.5

Hence the distribution list should be fall in between -3.65 to 14.5 and any element apart from that range should be an outlier and can be discarded.

In the above distribution list 21 is the outlier and can be discarded which falls after 14.5.

By using the above 5 points we can find the box plot.

**Box plot:-**



The dots at the extreme end are known as the outliers and they can be removed or ignored.