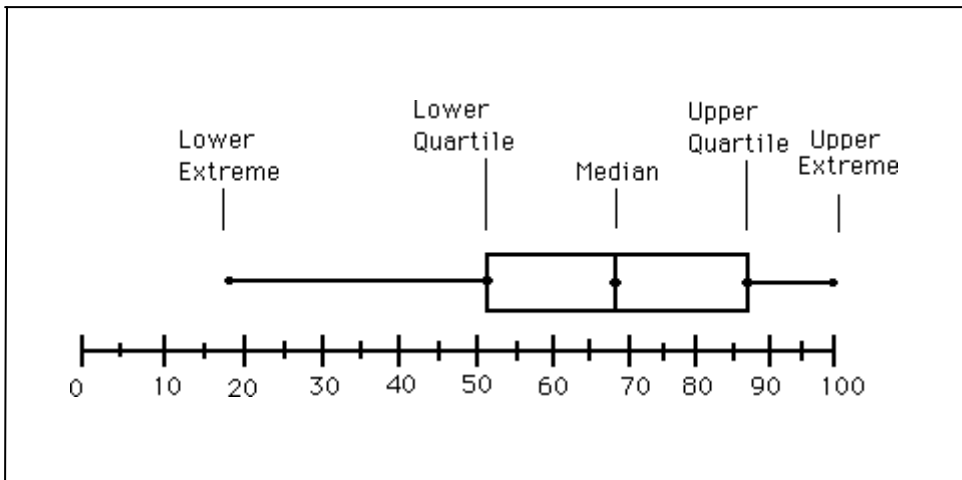


Box and Whisker Plots- Five Number Summary

Data: A Box-and-Whisker Plot shows the distribution of a set of data along a number line, dividing the data into four parts using the median and quartiles.



John Wilder Tukey

"An appropriate answer to the right problem is worth a good deal more than an exact answer to an approximate problem."

Tukey invented the box-and-whisker plot in 1977 to display certain groups of data. He picked $1.5 \times \text{IQR}$ (interquartile range) as the demarcation line for outliers. Typically, five values from a set of data are used; the extremes, the upper and lower quartiles (hinges) and the median.

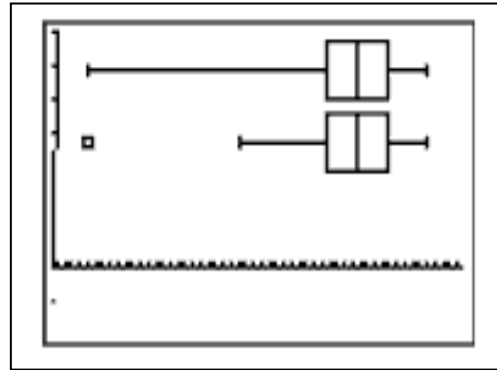
- It's called the **FIVE NUMBER SUMMARY or FIVE STATISTICAL SUMMARY** because it consists of all five numbers for the box-and-whisker plot including the minimum value or lower extreme, Q_1 (the bottom of the box), Q_2 (the median of the set), Q_3 (the top of the box), and the maximum value or upper extreme (which is also Q_4). So I need to order the set from least to greatest, find the median and the sub-medians, and then list the required values in order.
- *The whiskers are the lower and upper extremes (endpoints).

Box and Whisker Plots

Given the data set
{85, 100, 97, 84, 73, 89, 73, 65, 50, 83, 79, 92, 78, 10},
create a box and whisker plot to represent this data.

1. CLEAR out the graphs under **y =** (or turn them off).

2. Enter the data into the calculator lists.
Choose **STAT, #1 EDIT** and type in entries.
(See [Basic Commands](#) for entering data.)



3. Two icons for Box-and-Whisker Plots:

Choose the **second icon** for beginning level work.

Press **2nd STATPLOT** and choose **#1 PLOT 1**.
You should see the screen at the right. Be sure the plot is **ON**, the second box-and-whisker icon is highlighted, and that the list you will be using is indicated next to **Xlist**. **Freq: 1** means that each piece of data will be counted one time.



What about that other icon?

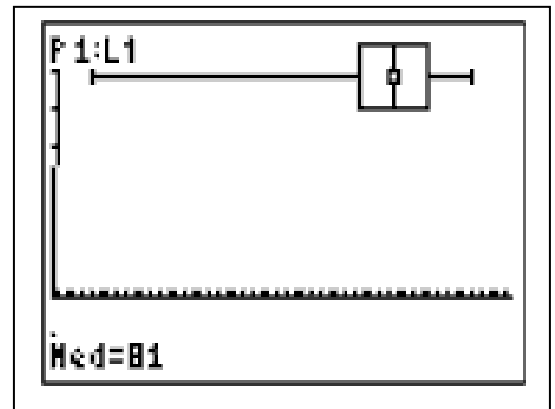
The first box-and-whisker icon is the modified box plot dealing with **outliers**. This modified version will not plot points that are $1.5 \cdot \text{IQR}$ beyond the quartiles. These points, called **outliers**, are plotted as individual points beyond the whisker in an attempt to give a more accurate picture of the dispersion of the data. Notice the two plots displayed at the top of this page representing the same set of data. **NOTE:** IQR stands for the Interquartile Range which is $Q3 - Q1$.



Used in more advanced statistics.

4. Seeing the graph:

To see the box-and-whisker plot, press **ZOOM** and **#9 ZoomStat**. Press the **TRACE** key to see on-screen data about the box-and-whisker plot. The whiskers extend from the minimum data point in the set to the first quartile, and from the third quartile to the maximum point. The box itself is defined by $Q1$, the median and $Q3$. The spider will jump from the minimum value to $Q1$, to median, to $Q3$ and to the maximum value.



Example # 2:

The following set of numbers is the amount of dolls fifteen different girls own in order from least to greatest.

18 27 34 52 54 59 61 68 78 82 85 87 91 93 100

L1	L2	L3
82		
85		
87		
91		
93		
100		

1. First find the *median*. The median is the value exactly in the middle of an ordered set of numbers. *

68 is the median

2. Consider only the values to the left of the median: 18 27 34 52 54 59 61 and find the median of this set of numbers. Remember, the median is the value exactly in the middle of an ordered set of numbers. Therefore, 52 is the median of the scores less than the median of all scores, and is the *lower quartile*.

52 is the lower quartile

3. Now consider only the values to the right of the median: 78 82 85 87 91 93 100. We now find the median of this set of numbers. The median 87 is therefore called the *upper quartile*.

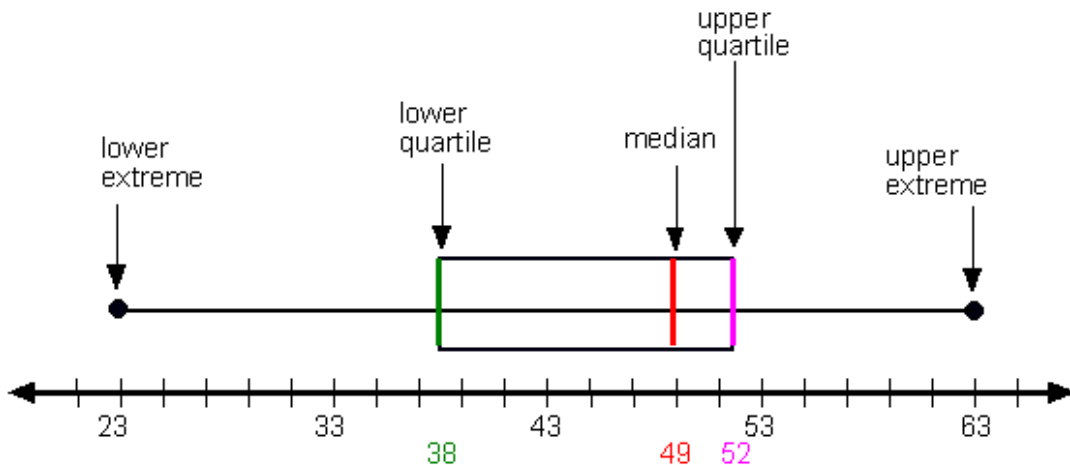
87 is the upper quartile

(*If you're finding the median in an ordered set with an even number of values, you must take the average of the two middle numbers. For example: 3, 5, 7, and 10. Add the two middle numbers. $5 + 7 = 12$. Then, divide 12 by 2 to get the average. $12 / 2 = 6$. Therefore, 6 is the median for the ordered set of 3, 5, 7, and 10.)

4. You are now ready to find the *interquartile range (IQR)*. The interquartile range is the difference between the upper quartile (Q3) and the lower quartile (Q1). In our case the $IQR = 87 - 52 = 35$. The IQR is a very useful measurement because it is less influenced by extreme values, it limits the range to the middle 50% of the values.

35 is the interquartile range

Now that you have all of the information, you can draw your graph or put it in your calculator.



Example # 3:

Construct a box plot for the following data:

12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25

Solution:

L1	L2	L3	1
5			
22			
30			
7			
36			
14			
42			
L1(2)=5			

Step 1: Arrange the data in ascending order.

Step 2: Find the median, lower quartile and upper quartile

5, 7, 12, 14, 15, 22, 25, 30, 36, 42, 53
 ↑ ↑ ↑
lower quartile median upper quartile

(If there is an even number of data items, then we need to get the average of the middle numbers.)

What is the median? _____

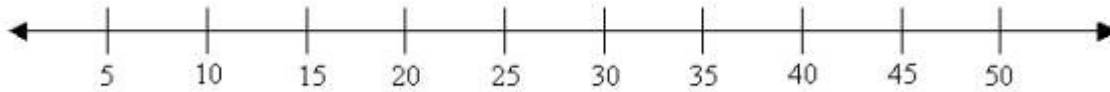
Upper quartile? _____

Lower Quartile? _____

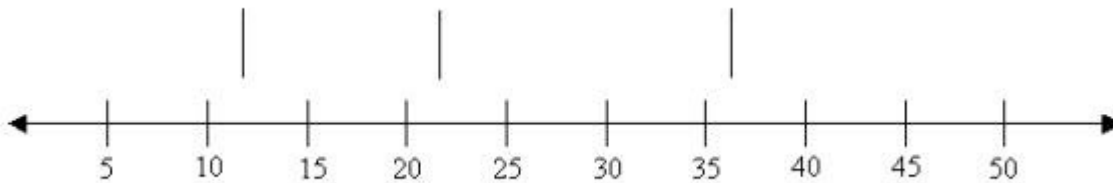
Lower extreme? _____

Upper extreme? _____

Step 3: Draw a number line that will include the smallest and the largest data.



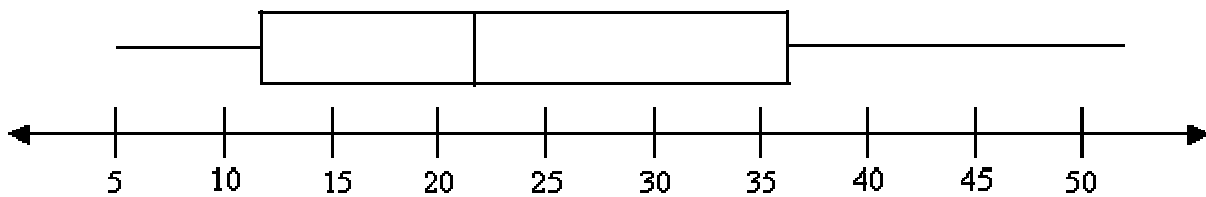
Step 4: Draw three vertical lines at the lower quartile (12), median (22) and the upper quartile (36), just above the number line.



Step 5: Join the lines for the lower quartile and the upper quartile to form a box.



Step 6: Draw a line from the smallest value (5) to the left side of the box and draw a line from the right side of the box to the biggest value (53).

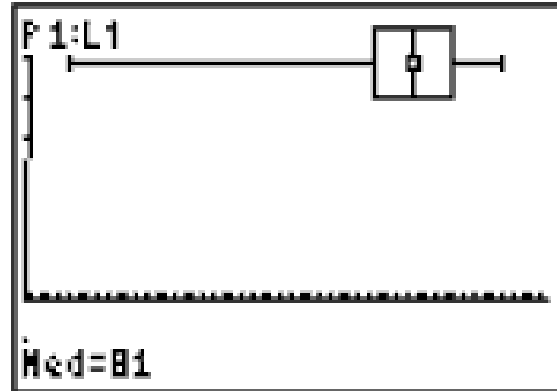


***When you put this information into your calculator following the correct steps, your box and whisker plot should look like this and five number summary should have included the values 22 for the median, 12 for the lower quartile, 36 for the upper quartile, 5 for the lower extreme and 53 for the upper extreme.**

Five Number Summary

STEP 1. A box and whisker plot will clearly show a five number summary.

Pressing the **TRACE** key will display the values. The whiskers extend from the minimum data point in the set to the first quartile and from the third quartile to the maximum point. The box itself is defined by Q1, the median and Q3. The spider will jump from the minimum value to Q1, to median, to Q3 and to the maximum value.



STEP 2. A complete five number summary is displayed on the lower portion of the **1-Var Stats** screen.

L1	L2	L3	1
50			
83			
96			
92			
78			
10			
L1(15) =			

Enter the data in a list.

```

EDIT [2ND] [MODE] TESTS
1: 1-Var Stats
2: 2-Var Stats
3: Med-Med
4: LinReg(ax+b)
5: QuadReg
6: CubicReg
7: QuartReg
    
```

Go to **STAT - CALC** and choose **1-Var Stats**

```

1-Var Stats L1
    
```

On the **HOME** screen, when **1-Var Stats** appears, type the list containing the data.

```
1-Var Stats
 $\bar{x}$ =75.57142857
 $\Sigma x$ =1058
 $\Sigma x^2$ =86772
 $S_x$ =22.90016555
 $\sigma_x$ =22.06715169
↓n=14
```

Hit **ENTER**.

When the **1-Var Stats** information appears, notice that there is an arrow pointing downward at the bottom of the screen. Arrow down.

```
1-Var Stats
↑n=14
minX=10
Q1=73
Med=81
Q3=89
maxX=100
```

The **five number summary** is the listed as the last 5 items on this screen

Resources:

<http://mathbits.com/mathbits/tisection/Statistics1/BoxWhisker.htm>

<http://www5555.morris.umn.edu/~sungurea/introstat/history/w98/Tukey.html>

<http://www.purplemath.com/modules/boxwhisk.htm>

<http://www.onlinemathlearning.com/box-plot.html>

Cassandra Gucciardo

Assignment- Technology/Graphing Calculator- Option 1