**Box-and-Whisker Plots**

A **box-and-whisker plot** is a convenient way to display the five-number summary. To draw a box-and-whisker plot:

1. Mark theminimum, maximum, median, Q1, and Q3 above the numbers on your number line.
2. Draw a box that represents the middle 50% of the data by drawing a box from Q1 to Q3. The length of the box represents the interquartile range (IQR).
3. Draw a vertical line segment inside the box to show the median.
4. Draw “whiskers” to represent the lowest 25% (by connecting Q1 to the minimum value) and highest 25% (by connecting Q3 to the maximum value) of the data.

![MP900315438[1]]()

|  |  |  |  |
| --- | --- | --- | --- |
| **September** | **High** | **Average** | **Low** |
| 8 | 78 | 67 | 57 |
| 9 | 77 | 68 | 55 |
| 10 | 71 | 62 | 52 |
| 11 | 63 | 59 | 55 |
| 12 | 69 | 64 | 60 |
| 13 | 82 | 72 | 64 |
| 14 | 78 | 68 | 57 |
| 15 | 81 | 68 | 57 |
| 16 | 70 | 63 | 55 |
| 17 | 63 | 56 | 50 |
| 18 | 75 | 60 | 48 |
| 19 | 71 | 60 | 51 |
| 20 | 73 | 58 | 43 |
| 21 | 77 | 60 | 46 |
| 22 | 75 | 64 | 53 |
| 23 | 81 | 74 | 66 |

1. The data set on the right lists the high, average, and low temperatures in Farmington, CT from September 8 to September 23, 2011. Complete the table, using your calculator. (Suggestion: enter the data in L1, L2, and L3)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Min | Q1 | Med | Q3 | Max | IQR | Sx |
| High |  |  |  |  |  |  |  |
| Avg. |  |  |  |  |  |  |  |
| Low |  |  |  |  |  |  |  |

The statistics for the high temperatures are displayed in a box-and-whisker plot below.

1. Draw box-and-whisker plots below for the average and low temperatures.

**Box-and-Whisker Plots on the Calculator**



1. Use the calculator to make the three box & whisker plots for the temperature data from
problem 1. You may display them side by side if you use L1 in StatPlot 1, L2 in StatPlot 2, and L3 in StatPlot 3 and turn all plots on. Select either of the two icons for Box & Whiskers shown

to the right.

 Indicate what values you used in the Window menu:

 Xmin = \_\_\_\_\_\_\_\_\_\_\_\_\_ Xmax = \_\_\_\_\_\_\_\_\_\_\_\_\_ Xscl = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 Ymin = \_\_\_\_\_\_\_\_\_\_\_\_\_ Ymax = \_\_\_\_\_\_\_\_\_\_\_\_\_ Yscl = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Now change the values of Ymin and Ymax. Does this affect the way the box-and-whisker plots are displayed? Explain.



1. Match the box-and-whisker plots on the right to the descriptions they match.
2. WWhich plot has the greatest range?
3. WWhich plot has the greatest IQR?
4. WWhich plot has the greatest Q1?
5. WWhich plot has the greatest median?
6. WWhich plot has the greatest Q3?



1. The following set of data lists the top 18 New York Yankees Salaries for 2011. Find the five-number summary, the range, and the interquartile range. Then make a box-and-whisker plot.

|  |  |  |
| --- | --- | --- |
| **RK** | **Player** | **Salary (millions)** |
| 1 | Alex Rodriguez | 32.0 |
| 2 | CC Sabathia | 24.3 |
| 3 | Mark Teixeira | 23.1 |
| 4 | A.J. Burnett | 16.5 |
| 5 | Mariano Rivera | 14.9 |
| 6 | Derek Jeter | 14.7 |
| 7 | Jorge Posada | 13.1 |
| 8 | Robinson Cano | 10.0 |
| 9 | Nick Swisher | 9.1 |

|  |  |  |
| --- | --- | --- |
| **RK** | **Player** | **Salary (millions)** |
| 10 | Rafael Soriano | 9.0 |
| 11 | Curtis Granderson | 8.3 |
| 12 | Russell Martin | 4.0 |
| 13 | Phil Hughes | 2.7 |
| 14 | Eric Chavez | 1.5 |
|  | Andruw Jones | 1.5 |
|  | Freddy Garcia | 1.5 |
| 17 | Boone Logan | 1.2 |
| 18 | Bartolo Colon | 0.9 |

(Source: [*http://espn.go.com/mlb/team/salaries/\_/name/nyy/new-york-yankees*](http://espn.go.com/mlb/team/salaries/_/name/nyy/new-york-yankees))

Five Number Summary:



1. Michael Jackson’s *Thriller* is the top-selling album of all time, with 110 million albums sold world-wide. The next 24 top-selling albums are below.

|  |  |
| --- | --- |
| **Album** | **Albums sold****(millions)** |
| *Thriller,* Michael Jackson | 110 |
| *Back in Black,* AC/DC | 49 |
| *The Dark Side of the Moon,* Pink Floyd | 45 |
| *The Bodyguard,* Whitney Houston | 44 |
| *Bat Out of Hell,* Meatloaf | 43 |
| *Eagles: Their Greatest Hits, 1971–1975* | 42 |
| *Dirty Dancing,* Various Artists | 42 |
| *Millennium,* Back Street Boys | 40 |
| *Saturday Night Fever* , Bee Gees | 40 |
| *Rumours,* Fleetwood Mac | 40 |
| *Come On Over,* Shania Twain | 40 |
| *Led Zeppelin IV* | 37 |
| *Jagged Little Pill,* Alanis Morisette | 33 |



|  |  |
| --- | --- |
| **Album** | **Albums sold (millions)** |
| *Sgt. Pepper,* The Beatles | 32 |
| *Falling Into You,* Celine Dion | 32 |
| *Music Box,* Mariah Carey | 32 |
| *Dangerous,* Michael Jackson | 32 |
| *1,* The Beatles | 31 |
| *Let’s Talk About Love,* Celine Dion | 31 |
| *Goodbye Yellow Brick Road,* Elton John | 31 |
| *Spirits Having Flown,* Bee Gees | 30 |
| *Born in the U.S.A.,* Bruce Springsteen | 30 |
| *Brothers in Arms,* Dire Straits | 30 |
| *Immaculate Conception*, Madonna | 30 |
| *Bad,* Michael Jackson | 30 |

1. Enter **all of the data** into L1. Using the 1-Var Stats command, find the five-number summary, the range, and the interquartile range, mean and standard deviation of the albums sold.

**With *Thriller***

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | min | Q1 | median | Q3 | max | range | IQR | mean | Sx |
| **Albums sold (millions)** |  |  |  |  |  |  |  |  |  |

1. Apply the 1.5 times IQR rule to find the fences for this set of data.

Upper fence = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Lower fence = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. According to the 1.5 times IQR rule, can *Thriller* be considered an outlier? Explain.
2. Enter the data **without** ***Thriller*** in L2. Using the 1-Var Stats command, find the five-number summary, the range, and the interquartile range, mean and standard deviation of the albums sold for the data **without** ***Thriller***.

**Without *Thriller***

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | min | Q1 | median | Q3 | max | range | IQR | mean | Sx |
| **Albums sold (millions)** |  |  |  |  |  |  |  |  |  |

1. Compare the charts made in parts (a) and (d). Which statistics were most affected when the *Thriller’s* albums sold was removed from the list?
2. Which statistics were least affected when *Thriller’s* albums sold was removed from the list?
3. Statisticians often prefer a statistic that is **not affected by outliers**. Which **measure of** **spread** do you think they prefer? The **range** or the **IQR**? Explain.
4. Compare the values for the mean and standard deviation (Sx) for the data with and without the *Thriller* album. Describe how the outlier affected the **mean** and **standard deviation** of the data set.
5. Now make a box & whiskers plot on your calculator using the album data from problem 7. Turn off all plots except for one and use the data in L1 (**all albums including *Thriller***). Adjust the values in the Window menu so that the entire plot shows.
6. First select the icon “Box and Whiskers Showing Outliers.” Make a sketch of what you see in the space below. Remember to scale your graph.
7. Then select the icon “Box and Whiskers without Outliers.” Make a sketch of what you see in the space below. Remember to scale your graph.
8. When there is an outlier, what is the effect of changing the box-and-whisker display on the length of the whiskers?
9. Now have the calculator make a histogram for the albums sold data. Make a sketch of what you see in the space below.



1. How can you tell from the histogram that there is an outlier in this set of data?
2. Make up a data set with ten values, listed in order, which could have the box-and-whisker plot shown below.



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |

1. Write a story (context) to go along with your data in (9).