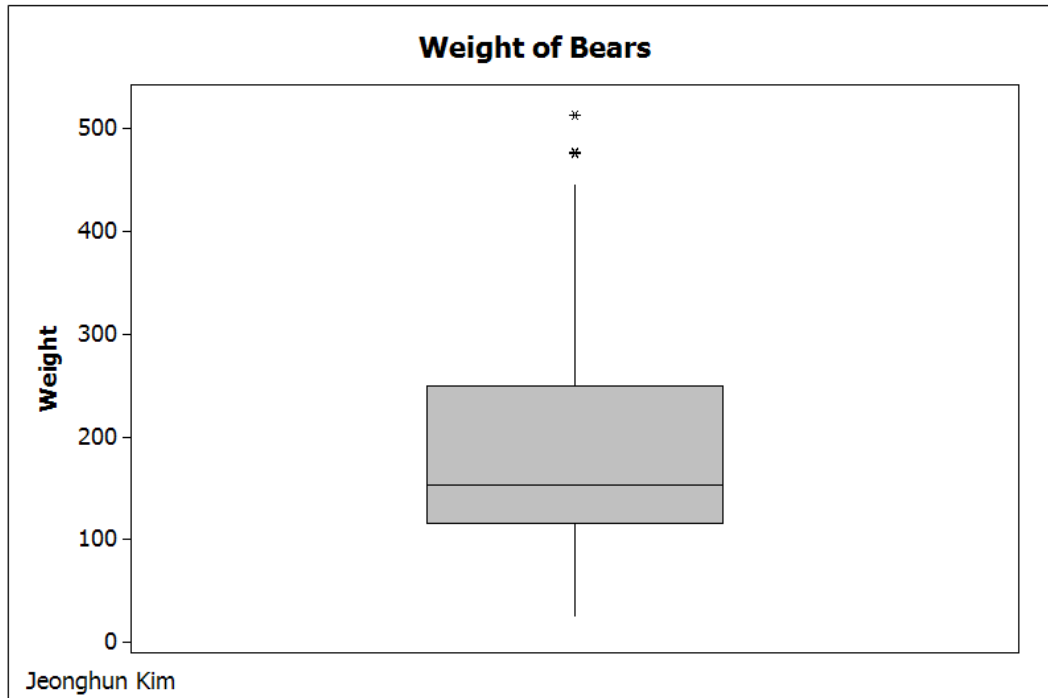


[Return to Cover Page](#)

## LESSON 7 – BOX-AND-WHISKER PLOTS

Before starting this section, be sure that your toolbars are set up as indicated at the beginning of Lesson 3 (see page 12).

Retrieve the worksheet K:\minitab15\MINITAB\MiniTab\Bears.MTW. Delete everything below the date\time stamp then type your name, Lesson 7 and Examples on separate lines. Now let's make a box-and-whisker plot of the bears' weight. Click on Graph > Boxplot, click "OK" on the first dialog box. In the second dialog box select C10 Weight into the "Graph variables:", use "Labels" to enter a title (Weight of Bears) and your name as a footnote, then click "OK". After making the tan area white, you should see a graph like the figure below.

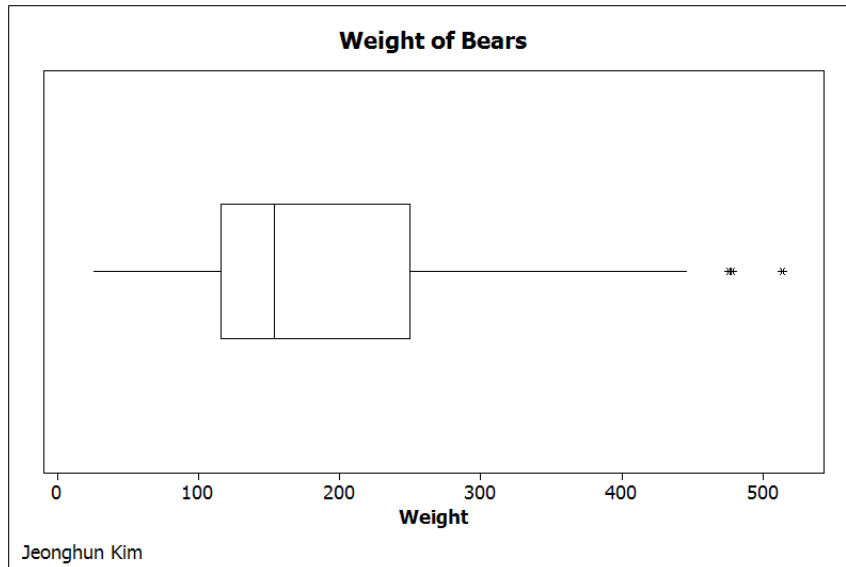


Notice the box-and-whisker plot looks different from graphs we have seen in our text. The three data points at the top of the box are values that are so much larger than the rest. Minitab considers them suspected outliers. Box-and-whisker plots with suspected outliers shown in this manner are called *modified boxplots*. Minitab program always draws modified boxplots instead of standard box-and-whisker plots. Obviously if there are no outliers, there is no difference between those two graphs.

To make the box white, edit the "Interquartile Range Box" in the same way that we edited the bars in the histograms in Lesson 3.

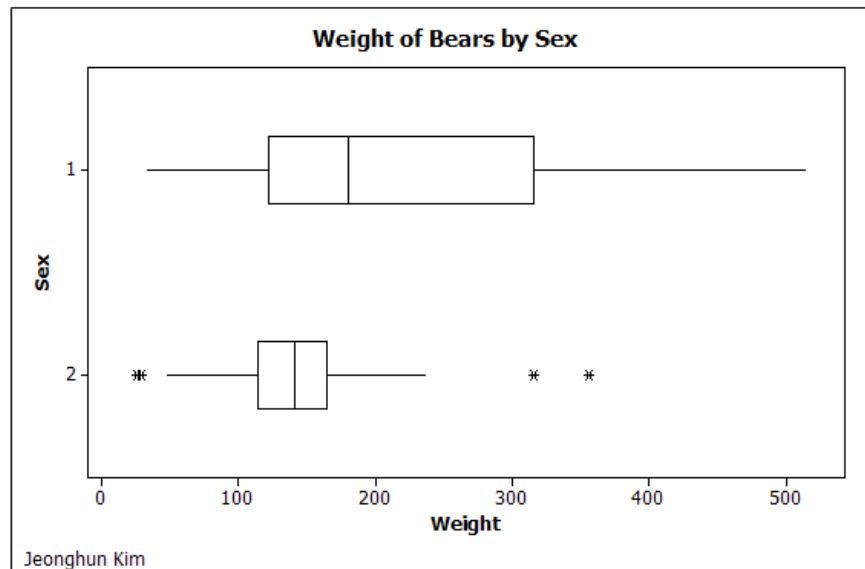
Notice that this modified boxplot runs vertically instead of horizontally as we have drawn

them in class. To change this select either "X Scale" or "Y Scale" as the item to edit, then click on the "Edit" tool. Now check the box in front of "Transpose value and category scales" then click "OK". Your graph should now look like the image below.



Although boxplots are sometimes drawn vertically in the literature, you should draw all your boxplots horizontally for this course.

To see the boxplots of the weight by sex, click on Graph > Boxplot, select "With groups" in the dialog box that opens, then click "OK". Select C10 Weight into the "Graph Variables:" box and C4 Sex into the "Categorical variables" box. Use "Labels" to enter a title and your name. Now click on "Scale" and check the box for "Transpose value and category scales" Click "OK" and "OK". Your graph should look like the figure below after you remove the color.



Notice that the first graph shows that when all the bears are together, the distribution of weight is significantly skewed to the right. After separation, however, only the males are skewed to the right. The distribution of the weight for the females is almost perfectly symmetric except for a few outliers.

Now let us find the mean, standard deviation, median, and interquartile range for the weight of the bears, first all together, then by sex. The results are shown below.

<b>Descriptive Statistics: Weight</b>						
Variable	Total Count	Mean	SE Mean	StDev	Median	IQR
Weight	143	192.16	9.24	110.54	154.00	134.00

<b>Descriptive Statistics: Weight</b>							
Variable	Sex	Total Count	Mean	SE Mean	StDev	Median	IQR
Weight	1	99	214.0	12.0	119.7	180.0	194.0
	2	44	143.05	9.72	64.48	141.00	50.50

Notice that for the bears all together and for the male bears, the standard deviation is smaller than the interquartile range, but for the female bears, the interquartile range is larger than the standard deviation. A look at the boxplots tells us why. Remember that the interquartile range measures the range of the middle 50% of the data, it is not influenced by outliers. The standard deviation, on the other hand, takes all of the data into consideration. A few outliers that are very far from the mean compared to most of the data, will result in the standard deviation as a measure of variation indicating greater dispersion among the data than is warranted. The bears all together have only three outliers out of 143 data points, and those are not that far out. The male bears have no outliers. The females, however, have four outliers out of only 44 data points. Almost 10% of the data are outliers, and the two on the right are very far out compared the rest of the data. This has greatly exaggerated the standard deviation. Thus, in the case of the females, the interquartile range is a more appropriate measure of variation.

## MINITAB ASSIGNMENT 7

**See instructions on page 8.**

1. Using the data from Minitab Assignment 6, construct a boxplot for the GPA of all students together and boxplots for GPA by Class. How do you explain the strange behavior observed in Minitab Assignment 6 for the standard deviation and interquartile range for sophomores? Type the answer to this question in the session window.
2. Do Problem 23 on Page 110. Display the data after entering it.

[Return to Cover Page](#)