

## Math 1040 StatCrunch Exercise 2

The purpose of this exercise is to introduce you to some more features of the StatCrunch software that is available with your textbook and MyStatLab program. You will learn how to calculate the values that numerically summarize data and how to create a boxplot.

Using the StatCrunch software, load the data for problem number 34 in section 3.2 of your textbook. (If you cannot recall how to accomplish this or any of the previously learned tasks, please refer to StatCrunch Exercise 1 which is posted in MyStatLab.) This data set contains the daily high temperatures recorded in November, 2007 for both Chicago and San Diego. We will be working only with the San Diego data here, so you may wish to delete the data for Chicago.

Begin by clicking on “Stat” on the toolbar at the top of the spreadsheet. Choose “Summary Stats” and click “Columns”. Choose your data by clicking on the title of the column, “San Diego”. To apply the default options, now click on “Calculate”. Copy the table of default summary statistics into the document file that you will submit (Word is a good choice). Save this document. You will be adding to it!

Notice that the table as copied may not look right. In most word processing documents, your text and numbers do not fit inside the cells causing it to wrap. You should never turn in a college level report with such a messy table. One way to make it look better is to highlight the text in your word processing document and change the font to a smaller size. Try it! You can make the text fit properly, but there are still some problems. Can you explain the meaning of standard error (Std. Err.)? If not, then you had better not include it in a report! Also note that we will seldom place as much importance on the variance as we do on the standard deviation, so you may not even want to include the variance. In StatCrunch, close the window with the summary statistics data and let’s try again.

Again click on “Stat”, “Summary Stats”, and select the column of temperatures for San Diego. But now click on “Next>” at the bottom of this window. This menu allows you to select the statistics you want in your output table and remove the statistics that you do not want. You may also change the order in which they appear by toggling all of them off and toggling them back on in the order you prefer. For now, let’s remove “Variance” and “Std. Err.” from the output, but include the interquartile range, IQR. Calculate again and copy the new output table into your document file. Place it below your first table and make sure it is formatted so that the text does not wrap and all values are easy to read.

Notice that the computer does not give you the mode. The easiest way to find the mode of this data set is to sort it and look to see what, if any, value occurs most often. Sort the column of temperatures for San Diego in StatCrunch. Use the arrow key on your keyboard to scroll down the column and determine the mode. In the document file that you will be turning in, type the value of the mode under the two tables of statistics you have already copied.

While your data is sorted, you may also look for outliers. Page 172 of your book explains one method for identifying outliers using the lower and upper fences:

$$\text{Lower fence} = Q_1 - 1.5(\text{IQR})$$

$$\text{Upper fence} = Q_3 + 1.5(\text{IQR})$$

Data below the lower fence or above the upper fence are considered outliers. Calculate the values of the lower and upper fences for the San Diego temperature data (Notice that you have already found all the values you need to make your calculations.) In the document that you will be turning in, type the fence values calculated and list any outliers for this data set.

You should now have a good idea how StatCrunch can help you find your numerical summary statistics. Another tool for summarizing data that you learned in Chapter 3 is the boxplot. To create a boxplot for the San Diego temperature data, click on “Graphics” on the toolbar at the top of the screen and choose “Boxplot”. Select the San Diego data column and create your graph. The default setting is a vertical graph, but since the book uses horizontal graphs, you may prefer to make it horizontal. Click on “Options” and “Edit”. Experiment with menus until you are happy with your boxplot. Use the built-in editing options to give your boxplot a title, and then copy it into the document you will be turning in. Type a sentence under this boxplot that explains the shape of the distribution of the data.

Now you will practice all these skills that you just learned with a new data set! Load the data for problem number 28 in section 3.4 of your textbook. This data set contains the weekly spending on entertainment of a sample of 40 students.

Use StatCrunch to output summary statistics for this data set that include the sample size, mean, standard deviation, 5-number summary values, and interquartile range. Do not include the standard error. Copy the output table into your document file, formatting it so that it is easy to read.

Find the mode, if any, of this data set. Report your findings in your document file.

Calculate the upper and lower fences and list any outliers. Report these values in your document file.

Create a boxplot for this data, give your boxplot a title using the StatCrunch editing options, and then copy it into the document you will be turning in. Type a sentence under this boxplot that explains the shape of the distribution of the data.

Print your document and turn it in by the posted due date.