

Project 1 – Using Minitab to Compute Descriptive Statistics

In this lab you will use the Minitab software to analyze data about the life length (measured in days) of rats that are fed two different diets. The two diets were the *restricted* diet and *free range* diet (hereafter referred to as the *regular* diet).¹

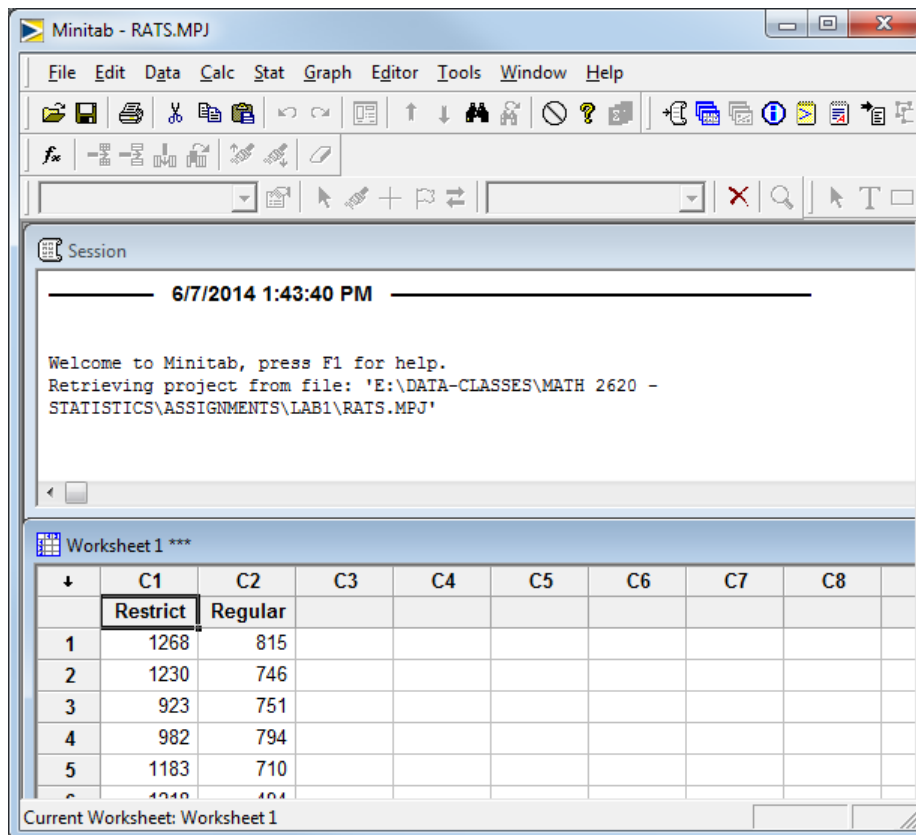
Follow these directions step-by-step. As you compute various statistics you will copy them to the Project 1 report file. The report will be printed and turned in in class on the due date is shown on the Schedule.

This project utilizes the Minitab software which is only available on the lab we use in class.

Step 1 - Preliminary

You will download two files and need to save them in a place that they can be retrieved later, so use a flash drive or network drive to store the files.

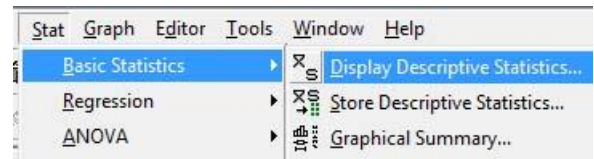
- 1.1. Do the following:
 - a. Download the project Template (p1_template.docx) from the Projects page, save to a flash/network drive and open it in MS Word.
 - a. Download the Data (p1_data.mpj) from the Projects page, save to a flash/network drive and then double-click to open it in Minitab.
 - b. Your Minitab screen should look something like this:



¹ I borrowed this data long ago and can't find the source.

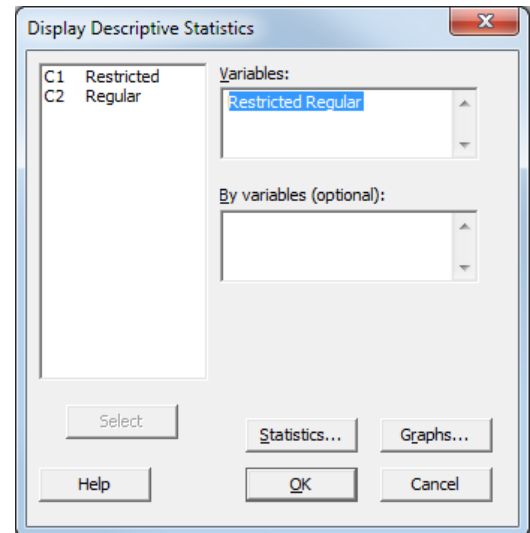
Step 2 - Numerical Descriptive Statistics

2.1. Choose: Stat, Basic Statistics, Display Descriptive Statistics:



2.2. Do the following:

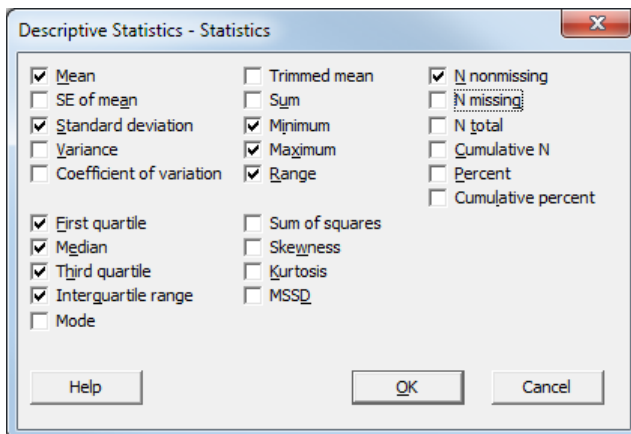
- Double-click each of the data sets names you wish to analyze. The name of each should appear under the *Variables* heading (see Figure on right)
- Choose the *Statistics* button.



2.3. On the dialog that appears, do the following:

- Uncheck: *SE of mean*
- Uncheck: *N missing*
- Check: *Range*
- Check: *Interquartile range*

The result will look like this:



e. Choose *OK*

2.4. Choose *OK* (again). The descriptive statistics will be shown in the Session window:

Descriptive Statistics: Restricted, Regular

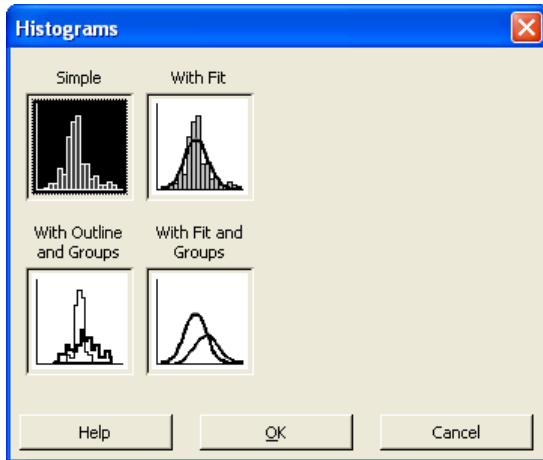
Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Range
Restricted	106	968.7	284.6	105.0	859.3	1035.5	1170.8	1435.0	1330.0
Regular	89	684.0	134.1	89.0	643.5	710.0	764.5	963.0	874.0

Variable	IQR
Restricted	311.5
Regular	121.0

2.5. Copy these descriptive statistics and paste them into the project template in the proper place.

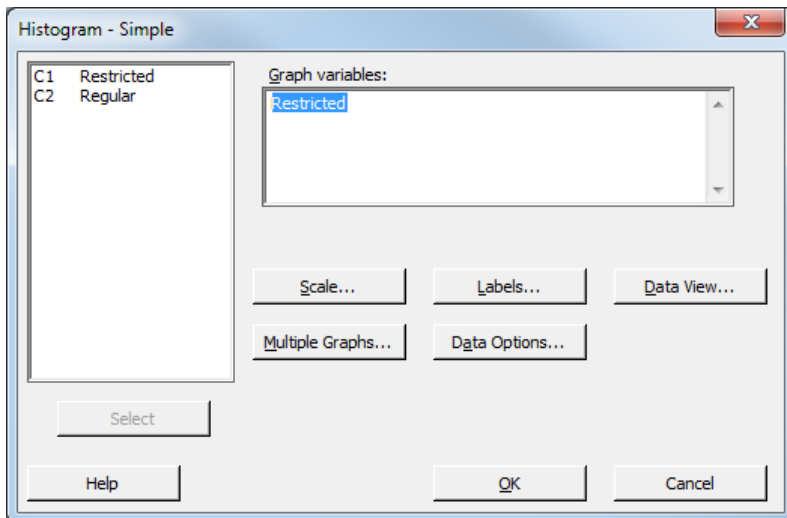
Step 3 - Histogram for Restricted Diet

3.1. Choose: *Graph, Histogram, Simple, OK*.

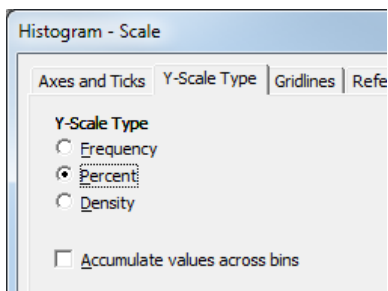


3.2. On the dialog that appears, do the following:

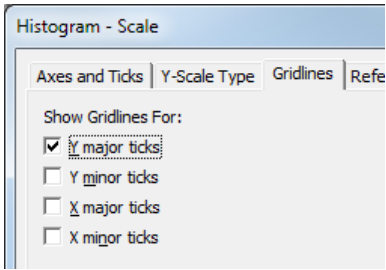
- a. Double-click "C1 Restricted" from the left pane and it will appear on the right pane under the *Graph Variables* heading.



- b. Choose: *Scale*. On the dialog that appears, choose the *Y-Scale Type* tab. Select: *Percent*.

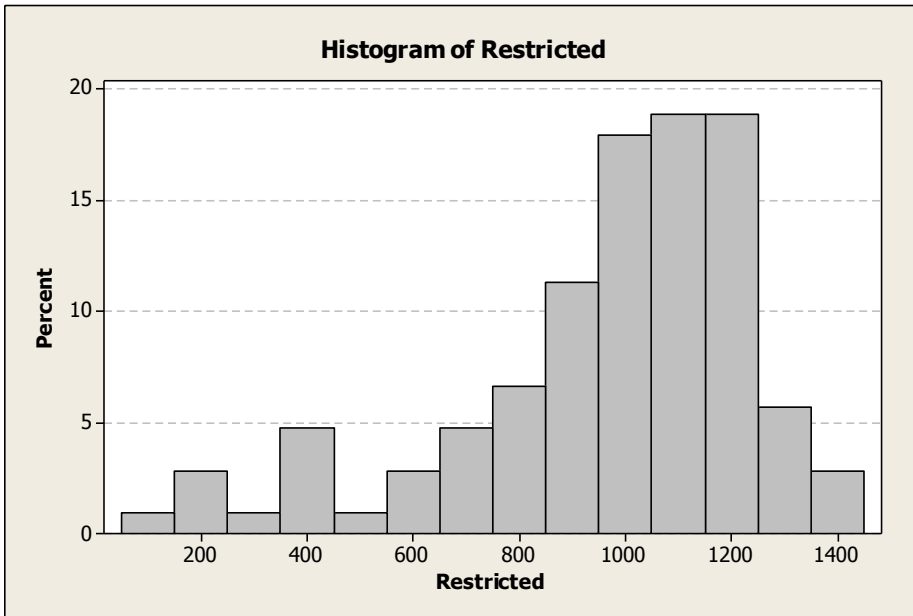


c. Choose the *Gridlines* tab and select *Y major ticks*



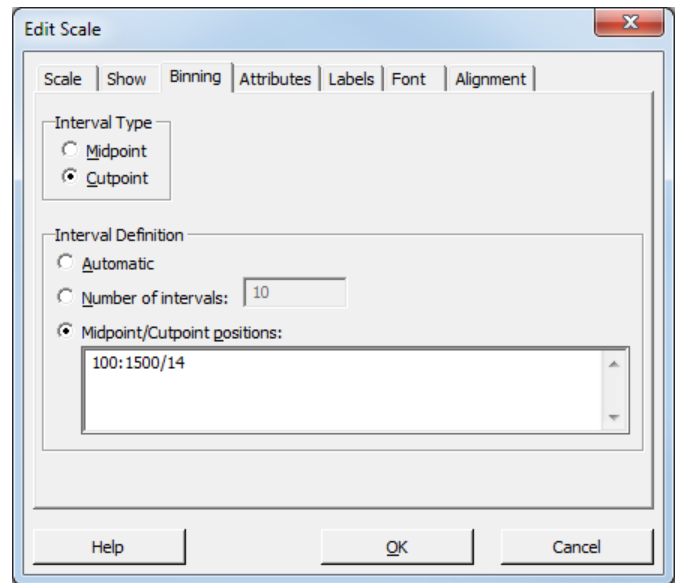
d. Choose: *OK*

e. Choose: *OK* (again). The graph is shown in a separate window:



3.3. Next, we will modify the x-axis. Do the following:

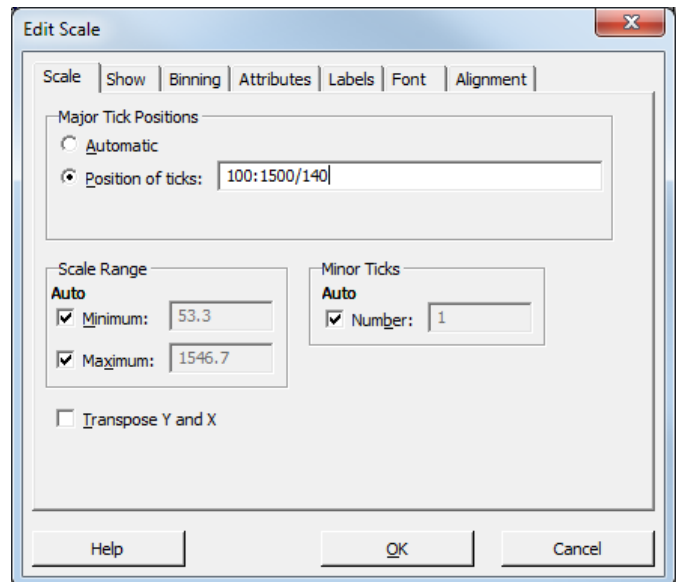
- a. Double-click a value on the x-axis.
- b. Choose the *Binning* tab.
- c. Select *Cutpoint* under *Interval Type* heading.
- d. **(Read: no action required)** Determine the appropriate number of bars for the histogram. A rule-of-thumb we use is to take the square-root of the sample size: $\sqrt{106} = 10.3$. Thus, we should use either 9, 10, or 11 bars. Here, we will use 10.
- e. **(Read: no action required)** Determine an appropriate range for the x-axis. Usually, we start the histogram a little below the smallest value and end it a little above the largest value. For this example, I chose to start the histogram at 100 since the minimum data value was 105 days. For the end, I chose 1500 since the maximum data value was 1435 days.
- f. **(Read: no action required)** Compute the bar width: $\frac{1500-100}{10 \text{ bars}} = 140 \text{ days}$.
- g. Select: *Midpoint/Cutpoint positions* under the *Interval Definition* heading. Add this text in the textbox (as shown in the figure above): **100:1500/140**



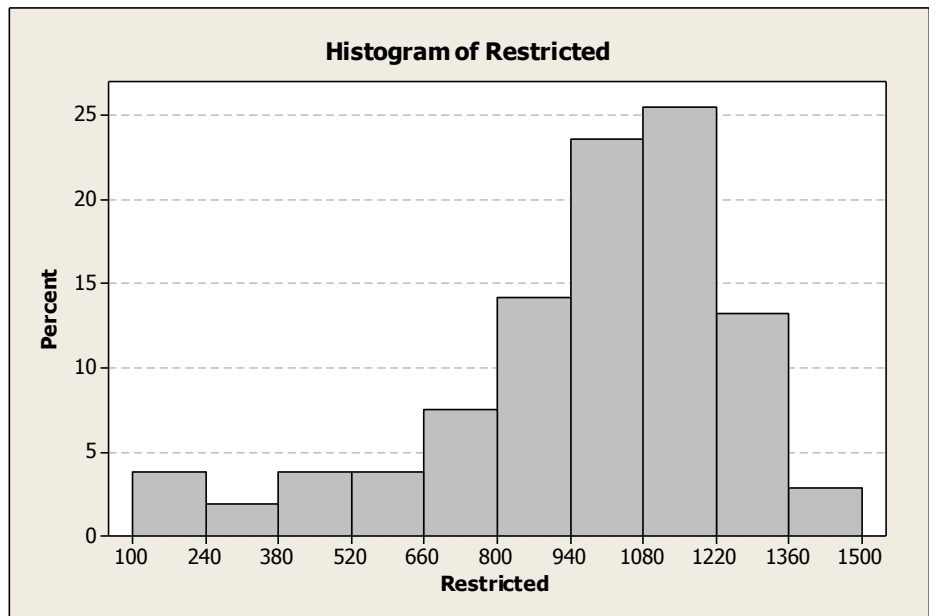
What does 100:1500/140 mean? It means that we are starting the histogram at 100 days, ending it at 1400 days, and each bar has width 140 days.

In general, you will want to play with these values so that you get values that are easy to interpret.

- h. Choose the *Scale* tab. Select *Position of ticks* and specify the same value as above: **100:1500/140**. (Ignore "Scale Range", you may have different values there-don't change anything).



- i. Choose *OK* and the resulting histogram will look like this:



- 3.4. Sometimes the values on the y-axis are spaced too widely. In the example above, the spacing seems fine. It goes from 0 to 25 with increments of 5. We could tighten this up if we want by having it go from 0 to 25 by 2.5 increments. To do this:

- a. Double-click the y-axis
- b. Select *Position of ticks*
- c. Supply the value: 0:25/2.5
- d. Choose *OK* and the result will look like this:



- 3.5. Make the gridlines black. When you print the histogram with the thin gray gridlines as shown above, they tend to not show up very well. Let's make them black by doing the following:

- a. Double-click a gridline
- b. Select *Custom*
- c. Choose the color black
- d. Choose *OK*

- 3.6. Copy the graph and paste it into the project template in the proper place.

Step 4 - Histogram for Regular Diet

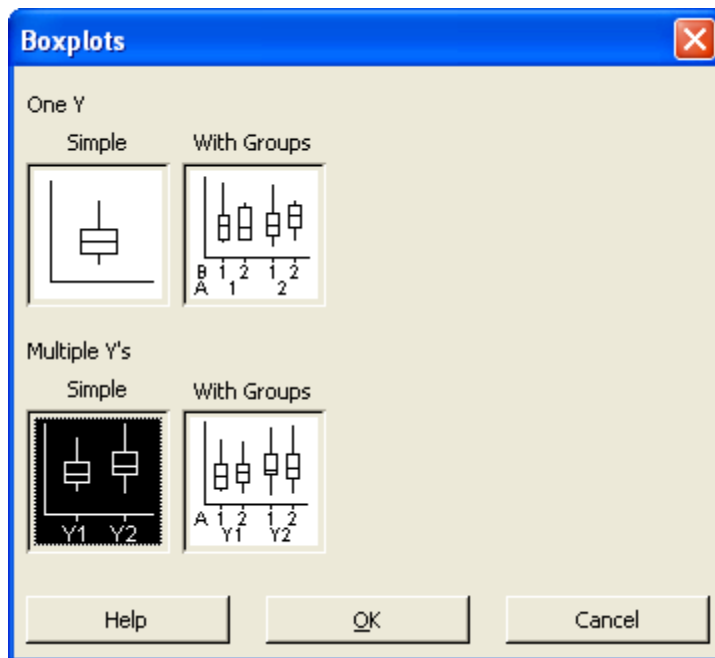
4.1. Repeat Steps 3.1-3.6 with the *Regular* diet.

Notes:

- For steps 3.3 *g* and *h*, we will run the histogram from 0 to 1000 with bars every 100 days. Thus, you will supply the value: **0:1000/100**.
- Ignore step 3.4

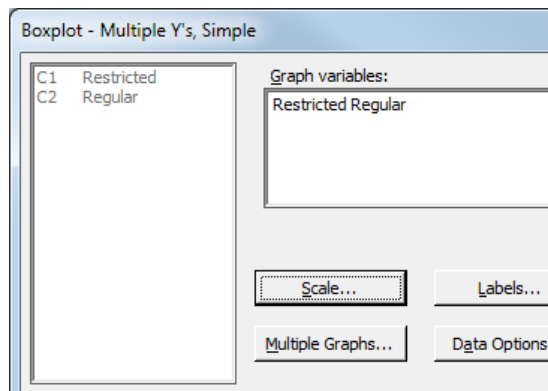
Step 5 - Boxplots

5.1. Choose: *Graph, Boxplot* and then select: *Multiple Y's – Simple*, then choose *OK*.

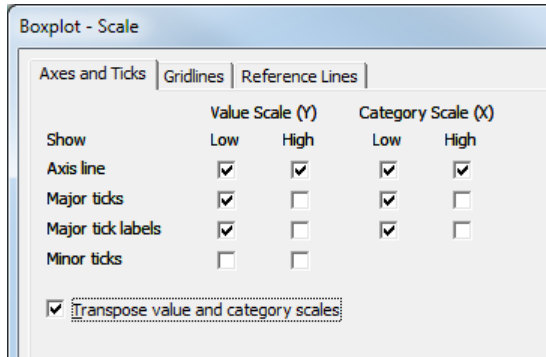


5.2. On the dialog that appears, do the following:

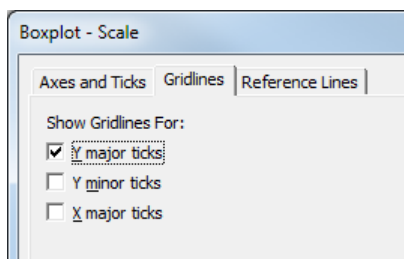
- Select the two data sets



b. Press the *Scale* button. On the dialog that appears, check: *Transpose value and category scales*

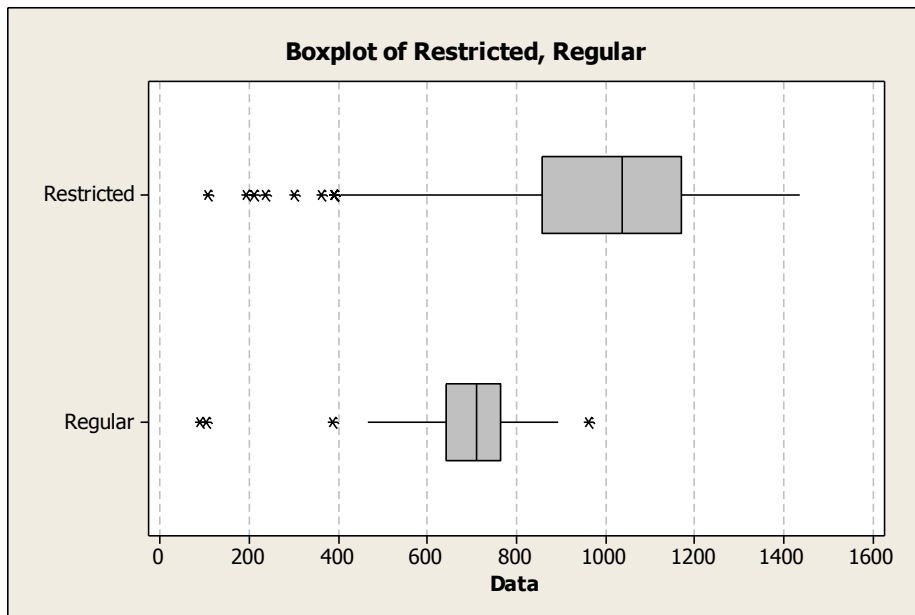


c. Select the *Gridlines* tab and select *Y major ticks*:



d. Choose *OK*

e. Choose *OK* (again). The boxplot is displayed



f. Modify the gridlines so that they are black.

Note (optional): You can modify the x-axis scale as we did before by double-clicking a value on the x-axis and specifying the position of the ticks. For example, you could use: 0:1500/100

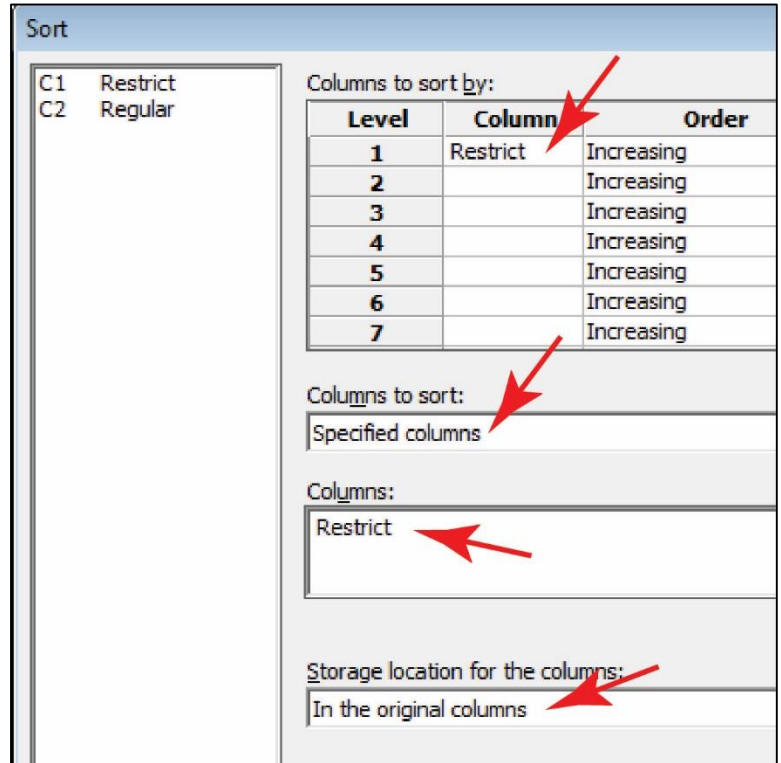
5.3. Copy the graph and paste it into the project template in the proper place.

Step 6 - Z-Scores

6.1. Choose: *Data, Sort*

- Specify the four items indicated with red arrows in the figure on the right.
- Choose *OK*

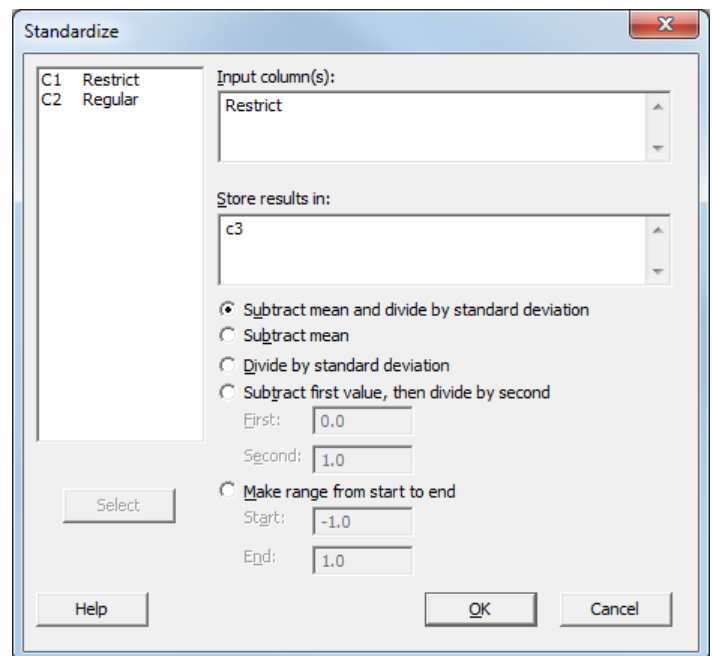
The data in column 1 (Restricted) should now be sorted.



6.2. Choose: *Calc, Standardize*

- Specify the *Input Column* (I've used *Restricted* below)
- Specify the *Store results in*. Below, I've used *C3* to indicate that I want the z-scores to be put in Column 3
- Choose *OK*. The z-scores will be shown in the *Worksheet*.
- Add a title to column 3 (e.g. Z-rest)

	C1	C2	C3
	Restrict	Regular	Z-Restricted
1	105	815	-3.03516
2	193	746	-2.72593
3	211	751	-2.66268
4	236	794	-2.57483
5	302	710	-2.34291



- 6.3. Repeat Steps 6.1 and 6.2 above using the other *Regular* data set. Be careful in Step 6.2 to put the z-scores in Column 4 (C4).
- 6.4. Copy the data and z-scores and paste them into the project template in the proper place. To copy, drag your mouse along the column headings: C1, C2, C3, C4. Then, right-click and “copy”. Finally, paste into Word.
- 6.5. Save your Minitab file and close
- 6.6. Move to the report file.
 - a. Put the section number (b or c), your name, and date at the top.
 - b. Format as necessary to make sure things appear nicely formatted and page breaks are at appropriate places.
 - c. Save and close.
- 6.7. Print your report, staple, and turn in in class.

You are done!

- ~~6.8. (ignore, summer 2016 classes) Submit the completed project template and minitab file on Blazeview (due date is on Schedule). Go to the Dropbox and find Project 1 and submit.~~