**Stat 401 A Lab 2**

# Goals:

Demonstrate how to read space delimited files

Demonstrate the importance of the Modeling type attribute.

Learn how to save JMP graphs or other JMP output in a word processor.

Demonstrate how to compute p-values from randomization tests.

# Lab Activities

We will analyze the creativity study, one of the Chapter 1 case studies.

**Reading space delimited files:**

1. Last week we saw how to read comma delimited (.csv) files. We will also have to read space delimited (.txt) files. This week, there is a space-delimited version of the creativity data on the web site, creativity.txt. Download creativity.txt into a convenient folder and use File / Open to open the filename dialog. Change the file type from “All JMP files” to Text files, select the file, and click Open.

The data window will show a messed up version of the data file. We want one column labelled treatment and a second column labelled score. Instead, we get one column with a nonsense variable called "treatment score".

The problem is that the default text import does not care about spaces.

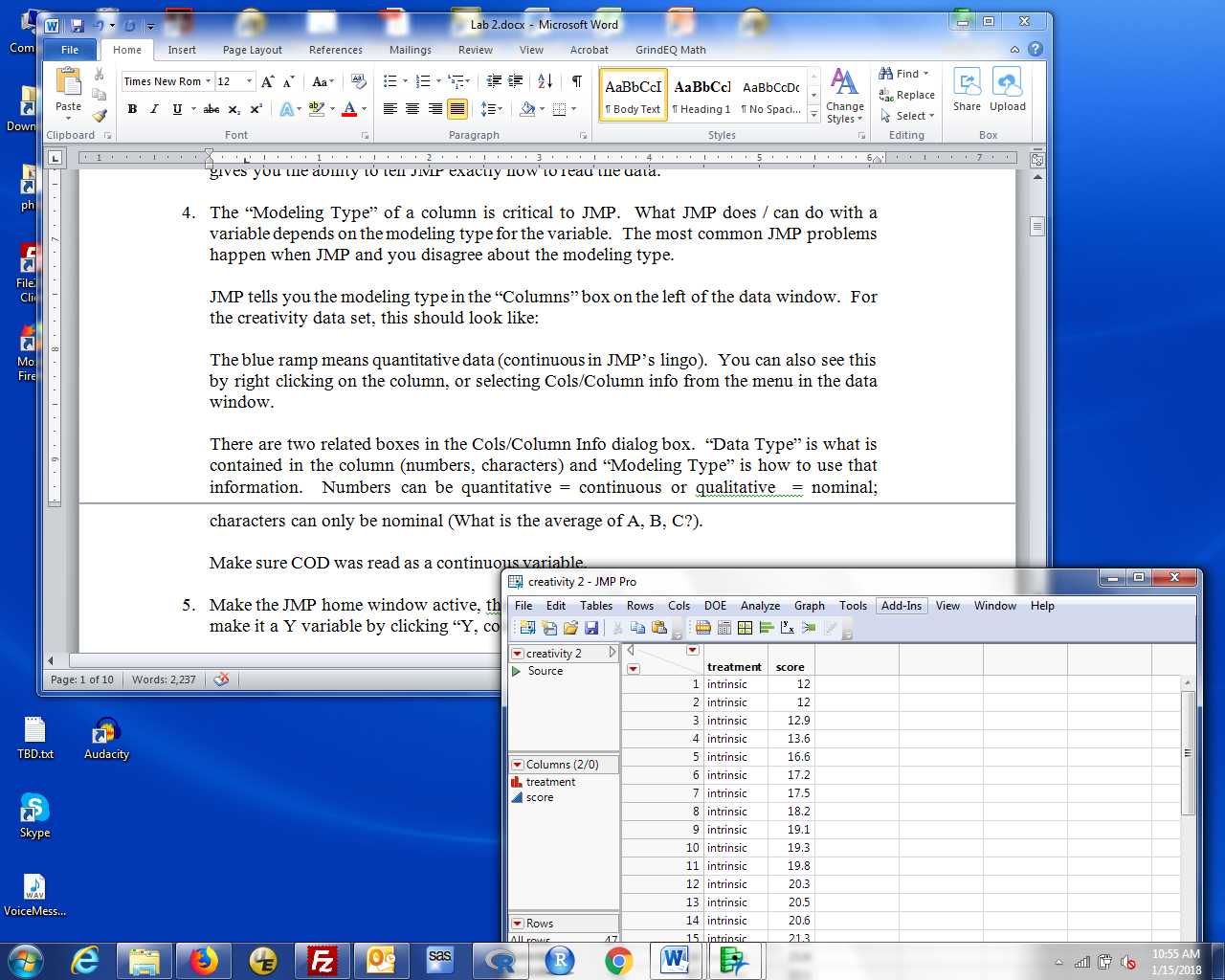
1. To read the data file correctly, choose File / Open again and look below the list of file names. You see Data, using Text Import Preferences and 3 other options. Select Data, using best guess, then left click the open button.

If the data was not read properly (the data in the window doesn’t look like what you expected), reread the using “Data, with preview” the third option for “Open As”. This gives you the ability to tell JMP exactly how to read the data.

**Changing the modeling type:**

1. The “Modeling Type” of a column is critical to JMP. What JMP does / can do with a variable depends on the modeling type for the variable. The most common JMP problems happen when JMP and you disagree about the modeling type.

JMP tells you the modeling type in the “Columns” box on the left of the data window. For the creativity data set, this should look like:



The blue ramp icon means quantitative data (continuous in JMP’s lingo). The red bars icon means categorical data (nominal in JMP's lingo). You can also see this by right clicking on the column name and selecting Column Info, or selecting the column (click anywhere to highlight the column name) and choosing Cols/Column info from the menu in the data window.

There are two related boxes in the Cols/Column Info dialog box. “Data Type” is what is contained in the column (numbers, characters) and “Modeling Type” is how to use that information. Numbers can be quantitative = continuous or qualitative = nominal; characters can only be nominal (What is the average of A, B, C?).

1. To see why the modeling type matters, use Analyze / Distribution to calculate summary statistics for all 47 scores (covered in creativity1 information last week). This should give you estimates of the mean, standard deviation, and other results appropriate for quantitative data.

Now, change the modeling type for score to nominal. Bring up the Column Info dialog (see step 3), select Nominal in the Data Type box, and click OK. Score now changes to a red bar variable. use Analyze / Distribution, choose Y as score, and look at the results. You get something very different and certainly don't see any quantitative summaries.

If JMP doesn’t behave as expected or doesn’t give you the options you expected, check that the modeling type is correct.

**Copying graphs:**

1. When I produce a report that presents results from a statistical analysis, I find it easier to retype numeric results so I include only the results I need in my report. However, it is convenient to be able to copy graphs or plots from JMP without printing them.

The easiest way is to copy and paste using keyboard shortcut keys. Left click on the graph, to make it active, then type ctrl-c (hold down the Ctrl key while typing once on the c key). That copies the graph to the clipboard. Navigate to your Word (or other word processing) document and type ctrl-v to paste the graph into your report. If you left click on the graph in Word, you activate the picture menu that allows you to resize or edit the graph. Here’s a copy and paste of the bar chart of COD values, using Graph/Chart.

If you click on the window with the Analyze/Distribution results, you can copy the entire window into Word, then delete the parts you don’t want.

Later, we will see how to use a JMP Journal to collect results. That also makes it easy to copy and paste selected results to Word. If you know about Journals or other ways to copy graphs, feel free to use what works for you.

**Randomization tests**:

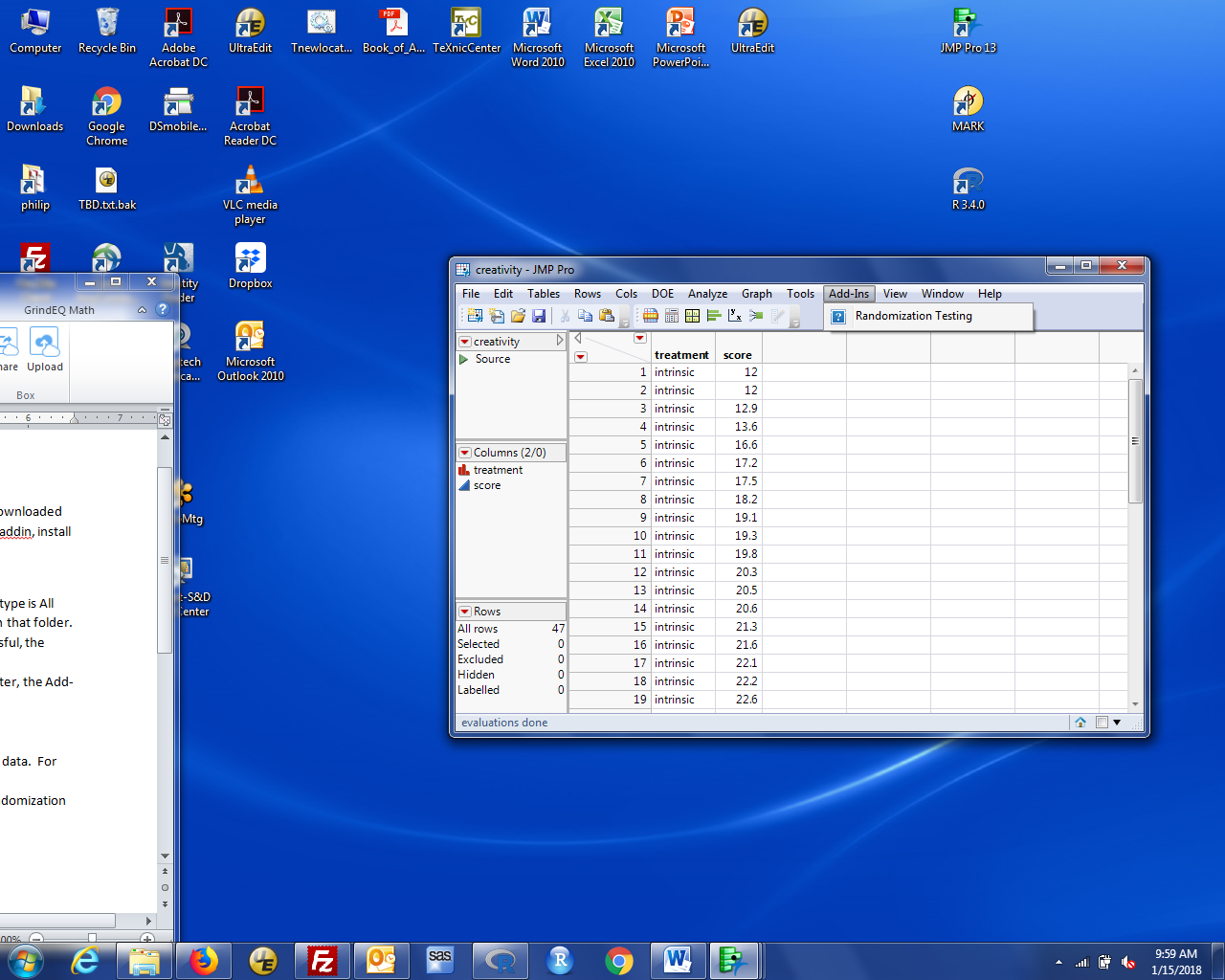
This requires an addin - Code that implements an analysis not part of regular JMP. I have downloaded the addin and put it on the class web site. To run this code, you need to download and the addin, install it (only necessary once), then run it.

1. Download the add in and save it in your 401 folder.
2. Start JMP 13 Pro, then File / Open and navigate to your 401 folder. The default file type is All JMP files. That includes addin files. You should see Randomization Testing Beta 3 in that folder. Select that file, Click Open, then Choose Install in the confirmation dialog. If successful, the main JMP menu should include a new item: Add-Ins.

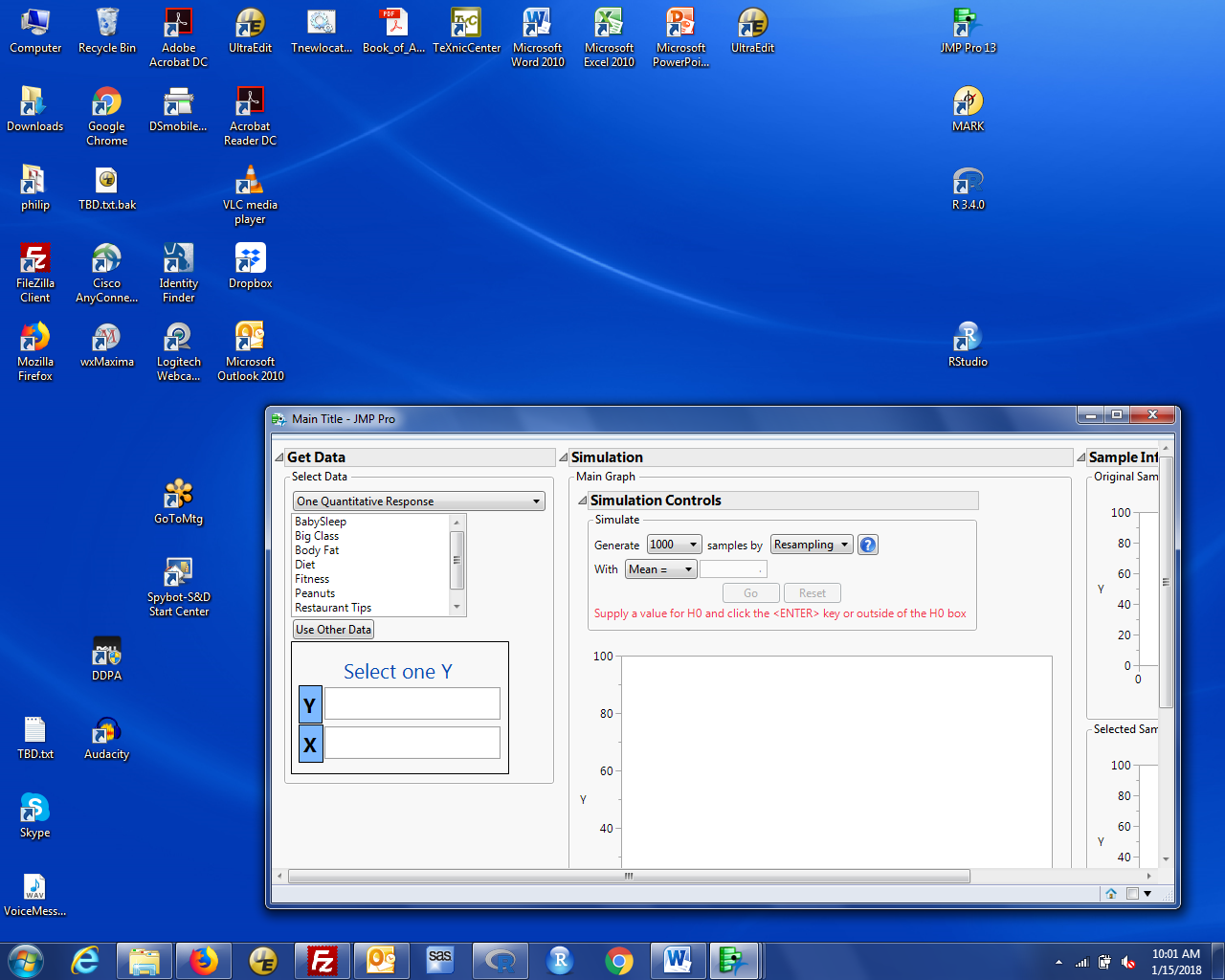
This installation is permanent (until uninstalled), so if you close JMP and restart it later, the Add-In is still available.

To run a randomization test:

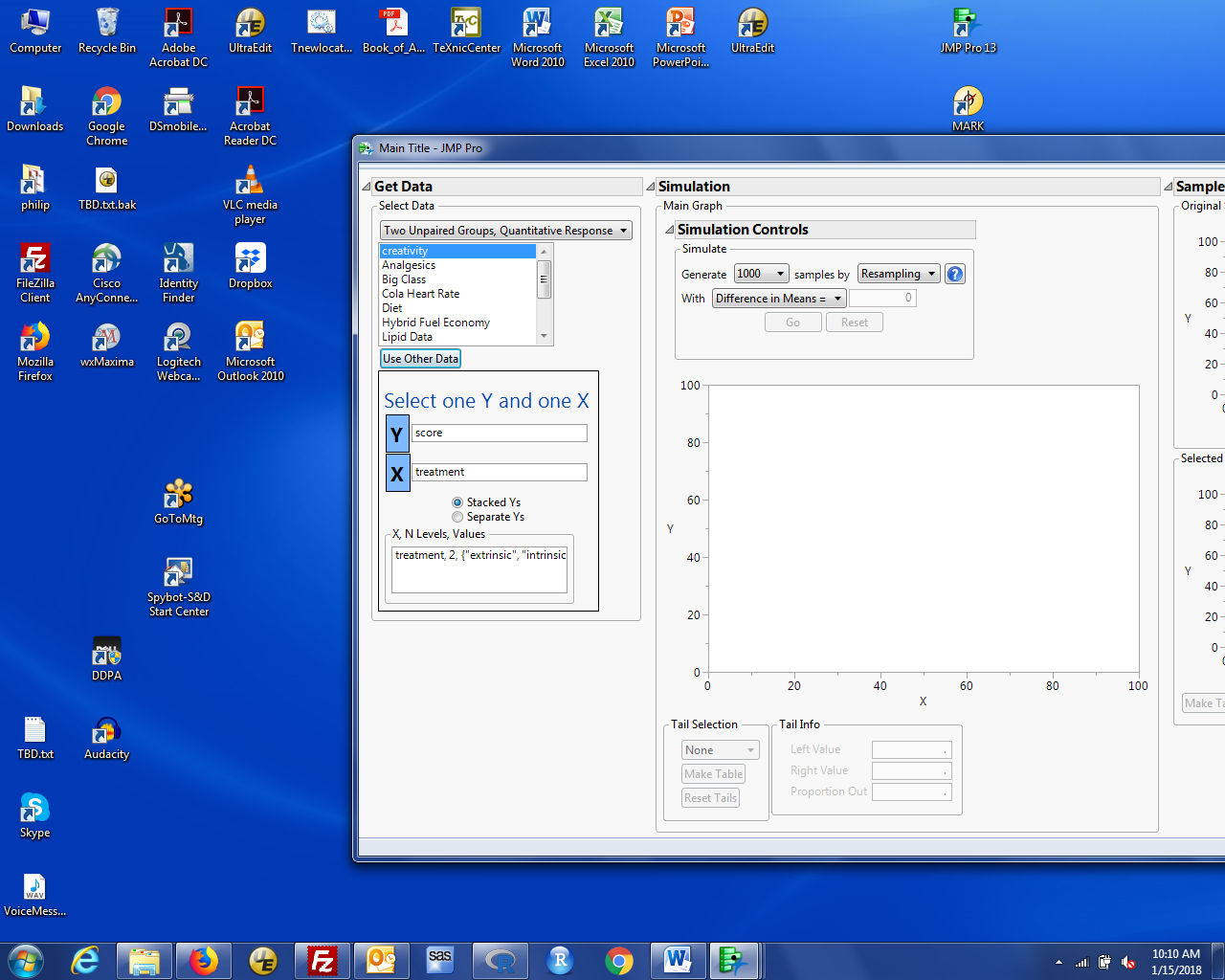
1. The add in requires a saved JMP data set (.jmp file extension). Read in the data, so there is a JMP Data window containing the data. For the creativity data set, this window will be named creativity. Then either File / Save or ctrl-s (the shortcut for save). The dialog will suggest the name of the data set (creativity here) and (important) by default save it as a JMP Data Table (.jmp file). Remember the folder where the data set will be saved (probably your Stat 401 folder) and left click Save.
2. Left click Add-Ins (in either the data window or the JMP window), then left click Randomization Testing.



That will open the following window:



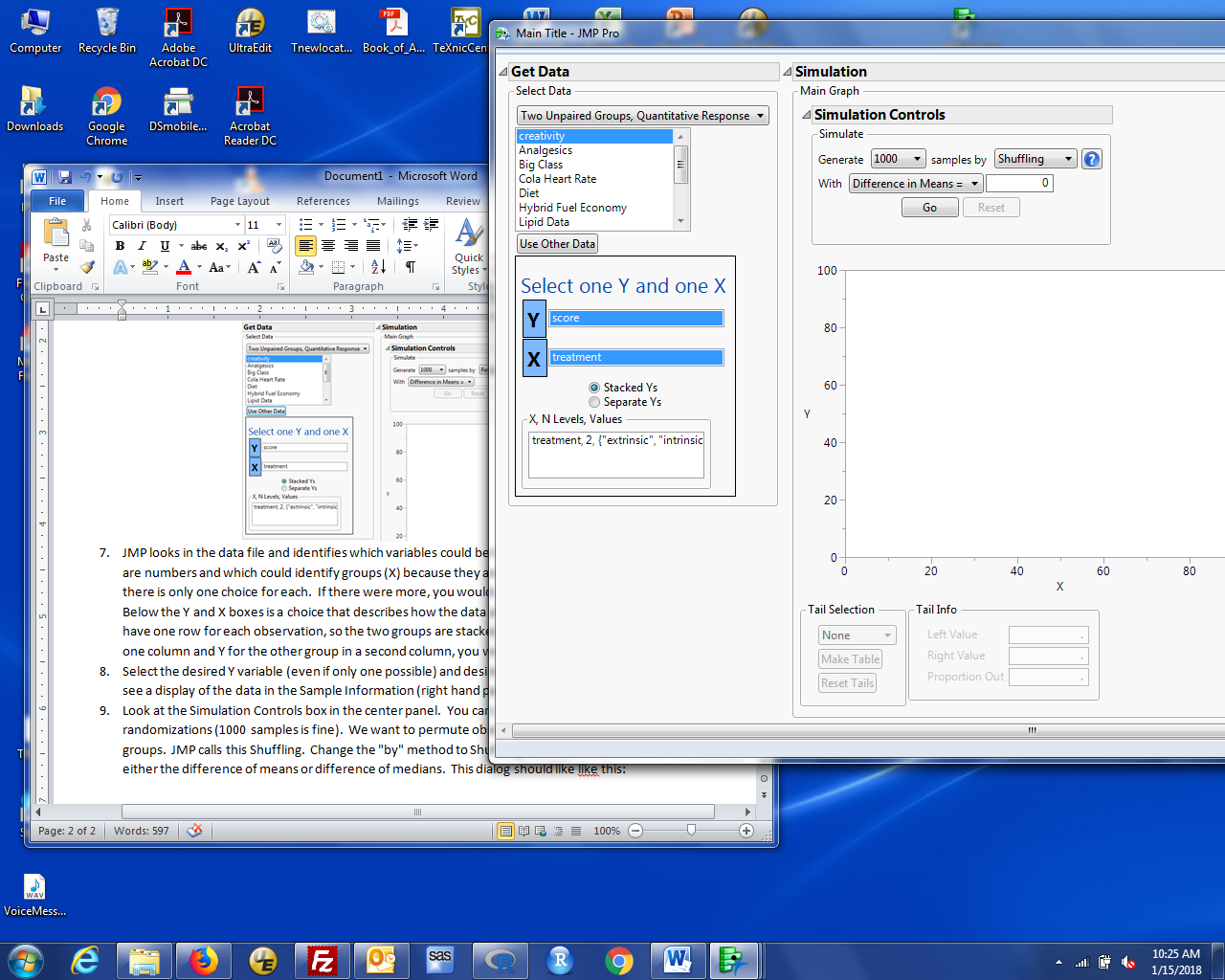
1. Select the appropriate type of test in the Get Data / Select Data box. We have two unpaired groups with a quantitative response. (We'll talk in a couple of weeks about paired and unpaired data). Below that choice is a list of example data sets provided with the addin. We want our own data set.
2. Left click Use Other Data below the list of example data sets. That will open a file open dialog. If step 3 above was successful, you should see creativity in that dialog with a JMP table icon next to it. Select that file and click open. You should see variable names in the Y and X boxes. For the creativity analysis, the left side of the window should look like this:



1. JMP looks in the data file and identifies which variables could be the response (Y) because they are numbers and which could identify groups (X) because they are not numbers. In this case, there is only one choice for each. If there were more, you would see more names in each box.

Below the Y and X boxes is a choice that describes how the data are stored. Our data sets will have one row for each observation, so the two groups are stacked. If you had Y for one group in one column and Y for the other group in a second column, you would have separate Y's.

1. Select the desired Y variable (even if only one possible) and desired X variable. You should now see a display of the data in the Sample Information (right hand part of the window).
2. Look at the Simulation Controls box in the center panel. You can specify the number of randomizations (1000 samples is fine). We want to permute observations between the two groups. JMP calls this Shuffling. Change the "by" method to Shuffling. JMP allows you to use either the difference of means or difference of medians. This dialog should like like this:



Click Go

1. You see a histogram of the difference in means found in the randomized samples. The observed difference (-4.1) is marked.
2. Locate tail selection box below the histogram. Select two-sided. JMP now marks the more extreme values and counts them.
3. The p-value is in the Proportion out box.

Note: JMP calculates the p-value as (# as or more extreme) / (# samples), which will be ever so slightly different than the suggested calculation (# more extreme + 1 ) / (# samples + 1)

If you want more samples (so less Monte-Carlo variation in the estimated p-value), just click Go again (or again and again).