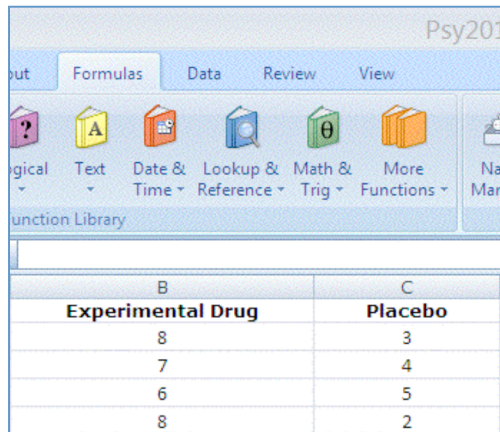# Psy201
Module 3 – Study and Assignment Guide

Using Excel to Calculate Descriptive and Inferential Statistics

What is Excel?

Excel is a spreadsheet program that allows one to enter numerical values or data into the rows or columns of a **spreadsheet**, and to use these numerical entries for such things as calculations, graphs, and statistical analyses.

	Experimental Drug	Placebo
	8	3
	7	4
	6	5
	8	2

What is a spreadsheet?

A spreadsheet is the computer equivalent of a paper ledger sheet. It consists of a grid made from columns and rows. The spreadsheet environment can make number manipulation easy and somewhat painless. The advantage of Excel is that you can experiment with numbers without having to RE-DO all of the calculations. **LET THE COMPUTER DO IT FOR YOU!!**

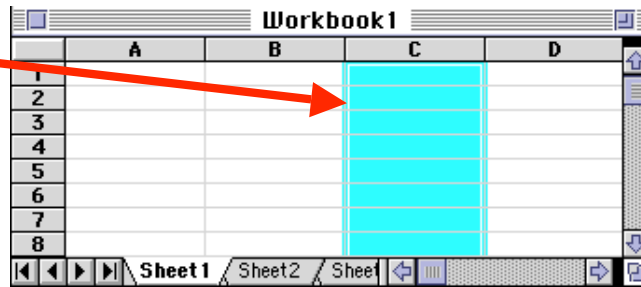|	A	B	C
1		computer ledger	
2			
3		car loan	$12,000.00
4		interest	9.60%
5		# of payments	60
6			
7		Monthly Pmt.	$252.61

Spreadsheet Basics

Spreadsheets are made up of…
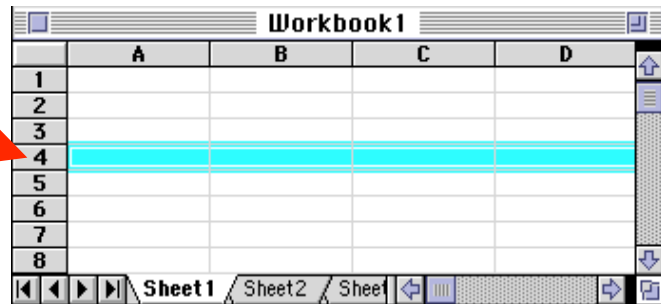
1. **COLUMNS**
2. **ROWS**
3. **CELLS**

COLUMNS:

1. Vertical spaces going up & down.
2. **Letters** are used to designate each **COLUMN'S** location.
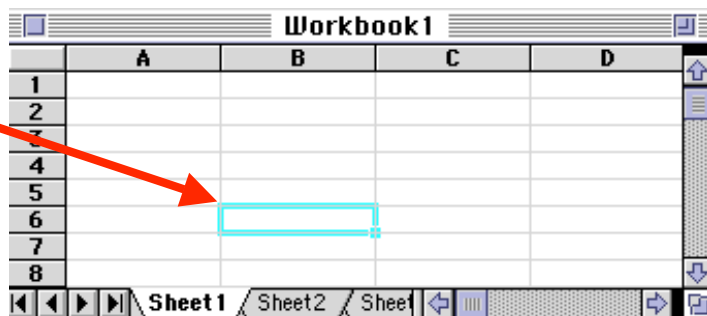3. So here column C is highlighted.

ROWS:

1. Horizontal space going across.
2. **Numbers** are used to designate each **ROW'S** location.
3. Do here row 4 is highlighted.

CELLS:

1. Space where row & column intersect.
2. Named according to **COLUMN letter** & **ROW number**.
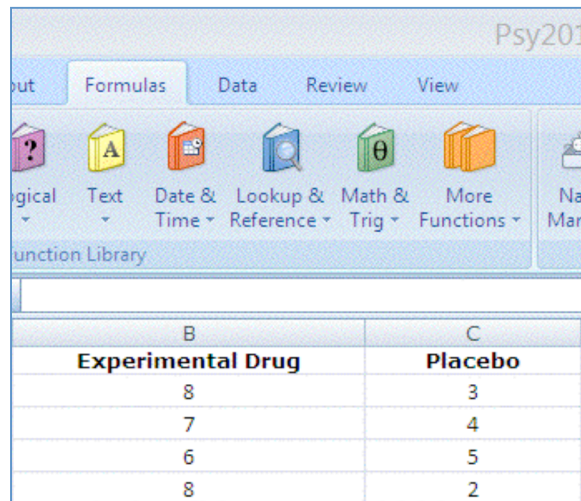3. So here cell B6 is highlighted, where B = column and 6 = row.

How to use Excel to Calculate Descriptive Statistics

Psychologists can use Excel to organize, describe, present, and analyze data.

Suppose you conducted an experiment testing the effects of a new memory drug (compared to a placebo) on the ability to remember a list of words, and you collected data and used Excel to organize, describe, and calculate some descriptive statistics. The independent variable (IV) is what the experimenter manipulates. In this case the IV is experimental drug vs. placebo. The dependent variable (DV) is what the experimenter measures as a result of the manipulation of the IV. In this case, the DV is the number of words recalled.

You might enter the data into Excel like this:

B	C
Experimental Drug	**Placebo**
8	3
7	4
6	5
8	2

Types of Data

It is important to note that there are **3 basic types of data** that can be entered into Excel:
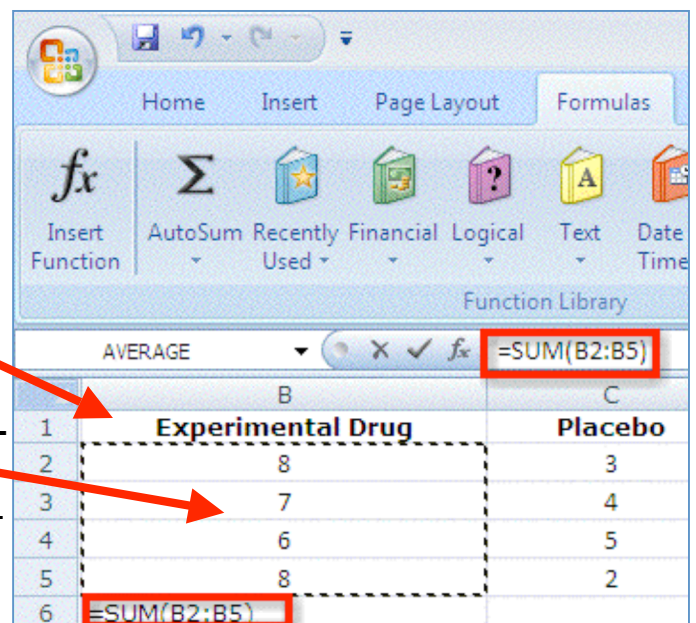
1. **Labels**
   - Text entries with no numerical value
   - In this example, the Labels include the two levels of the **Independent Variable**: Experimental Drug and Placebo.

2. **Constants**
   - Just a number (i.e. constant value)
   - In this example, the constants are the numerical data (i.e. **Dependent Variable -** number words recalled - 8, 7, etc.) located in the columns beneath the labels.

3. **Formulas***
   - A mathematical equation used to calculate a value to display
   - *ALL formulas MUST begin with an equal sign (**=**).
   - In this example, the following formula was entered into Cell B6, = sum(B2:B5). This formula calculates the total of the four memory scores in the Experimental Drug Group

Using Excel to Calculate Descriptive Statistics

Formulas or Functions

Remember that when creating formulas or functions, it is important that you BEGIN it with an equals sign (=).

Psychology students will find the **Statistics functions** very useful!

Common functions that you will use in Psychology statistics or research methods classes include descriptive statistics such as:

1. The **Sum** function
2. The **Average** function
3. The **Standard Deviation** function

A list of formulas or functions is available within Excel under the menu *Formula*, down to *Insert Function.*

You can also manually insert formulas by typing them directly into the cell where you want the result of the formula to appear.

1. The SUM Function (SUM)

The **SUM FUNCTION** takes all of the values in each of the specified cells and totals their values.

For example, in order to get the sum of all the words remembered in the Experimental Drug group, we would type in the following formula into cell B6, =SUM(B2:B5)
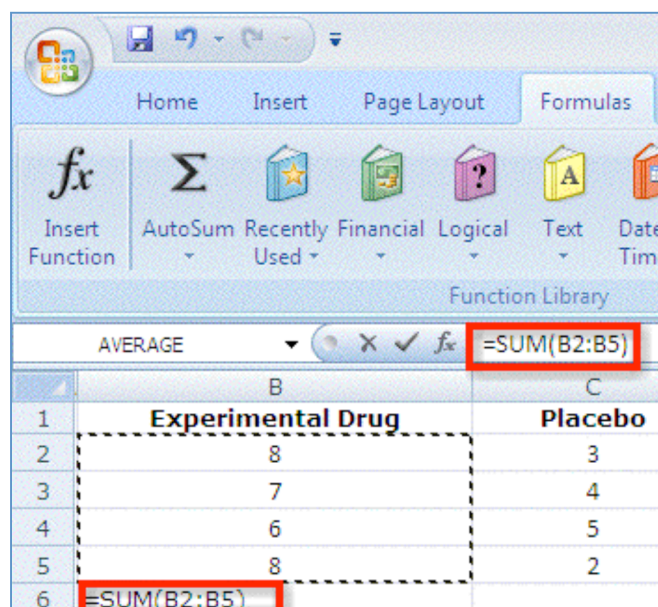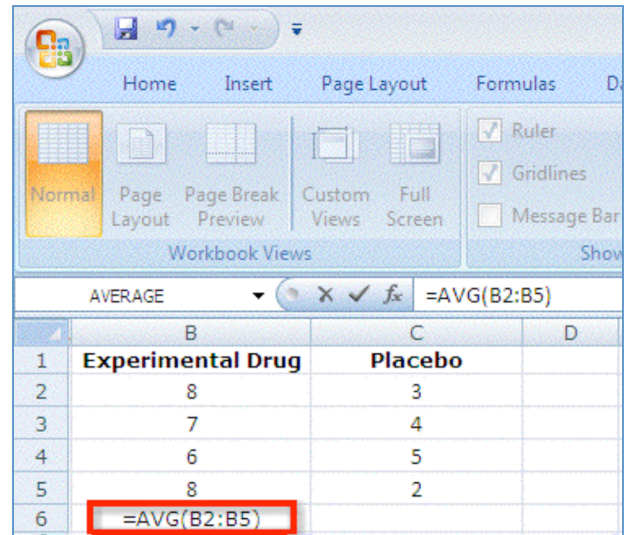
Formulas or Functions, cont'd

2. The Average Function (AVG)

The **AVERAGE** is a measure of central tendency and gives you an idea of the typical score.

The **AVERAGE FUNCTION** finds the average of the specified data.

For example, in order to get the average of all the words remembered in the Experimental Drug group, we would type in the following formula into cell B6, **=AVG(B2:B5).**
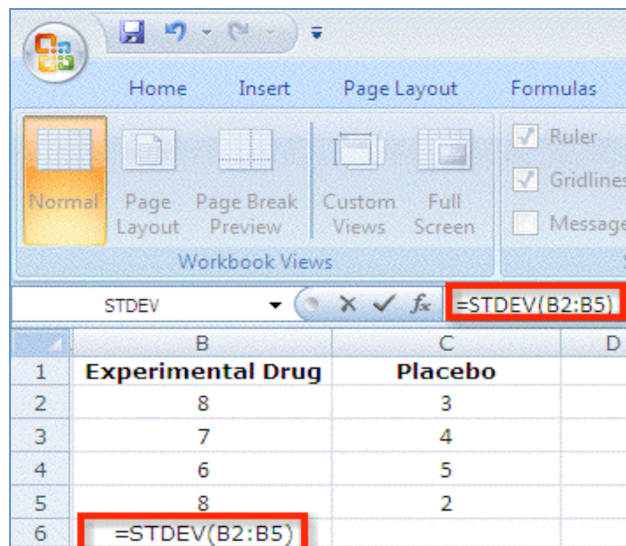

3. Standard Deviation Function (STDEV)

The **STANDARD DEVIATION** is a measure of the spread of variance of the data and gives you an idea of how different each score is from the average.

The **STANDARD DEVIATION FUNCTION** finds the standard deviation of the specified data. For example, in order to get the standard deviation of all the words remembered in the Experimental Drug group, we would type in the following formula into cell B6, **=STDEV(B2:B5).**
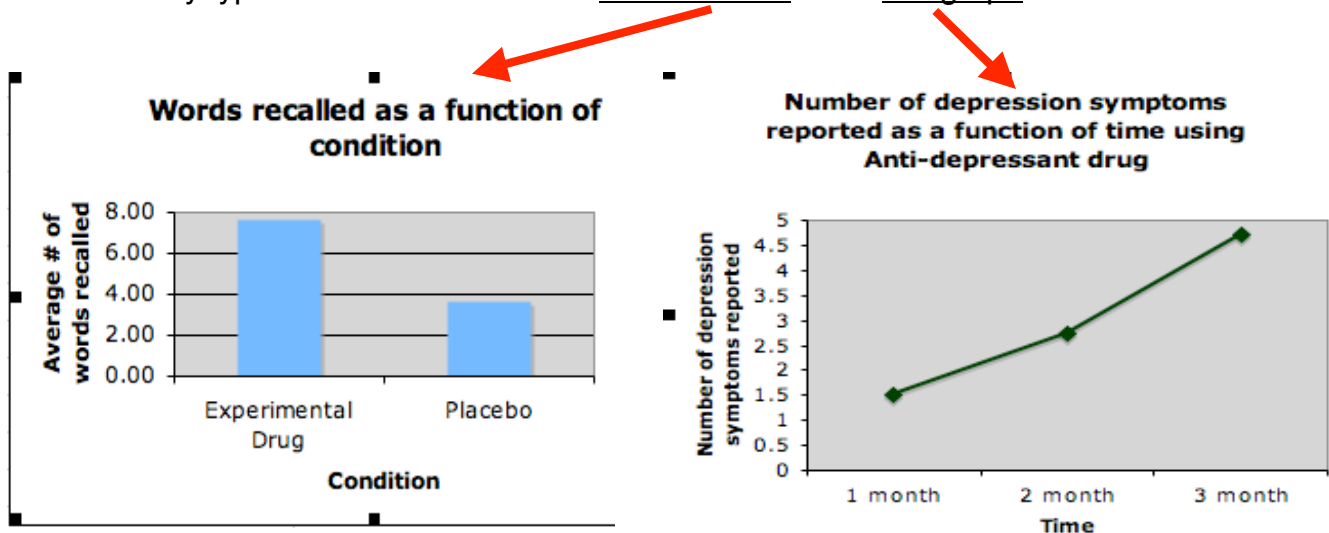
Making a Chart using Excel 2007 or Newer Version

Another important feature of Excel that psychologists often utilize is the chart feature. A chart is also called a graph. Excel has a chart program built into its main program.

There are many types of charts such as the <u>column chart</u> or the <u>line graph</u>

Words recalled as a function of condition

Number of depression symptoms reported as a function of time using Anti-depressant drug

1. First, enter the data to be graphed. Returning to our earlier example, enter the data labels (Experimental Drug and Placebo) and the average memory score for each as shown below:

	A	B	C	
1				
2		Number of Words Recalled		
3		Experimental Drug	Placebo	
4		8	3	
5		7	4	
6		6	5	
7		8	2	
8	Average =	7.25	3.50	
9				
10				
11		Average # of Words Recalled		
12		Experimental Drug	Placebo	
13		7.25	3.50	
14				

2. Next highlight the data to be graphed, including the data labels (in other words highlight Experimental Drug, Placebo, and the averages 7.25 and 3.50)

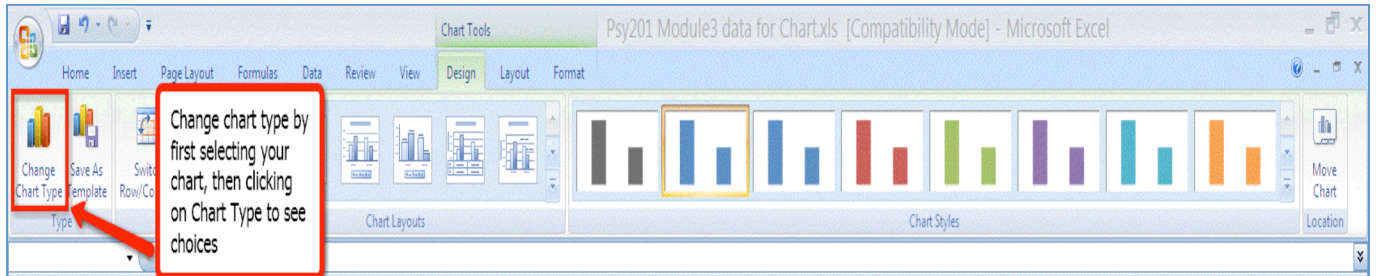3. When labels and averages have been entered and your cursor is immediately below the entered data, press the F11 key on the top row of your keyboard. A chart will appear.
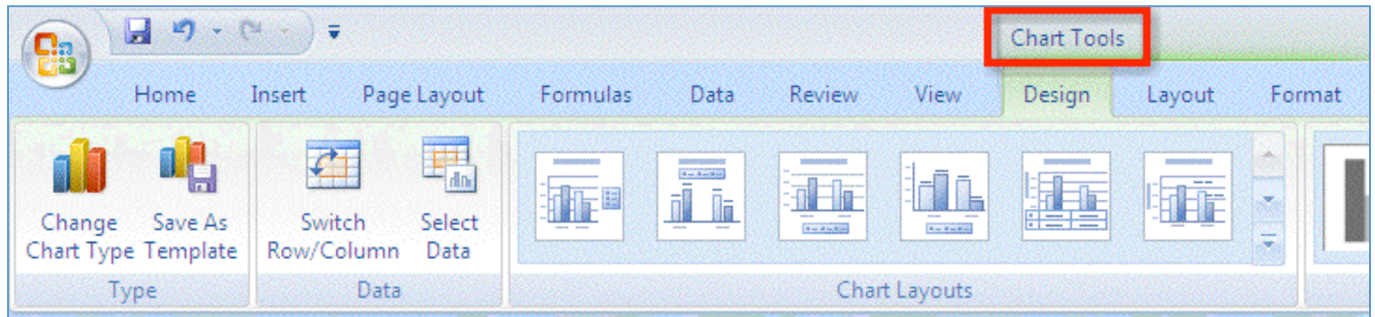
4. To change the **Chart Type:**
    a. Select your chart
    b. Click on the Change Chart Type button on the left to see all of the available chart types, click on Column, then click OK

5. On the **Chart Tools** tab, the third section from the left is named **Chart Layouts**
    a. Near the bottom right portion of that area you will see a small button that will allow you to see all available layouts.
    b. Click one time on the button to see the layouts.
    c. Select the layout that will allow you to add a title to the top of the chart, and label the horizontal X-axis and the vertical Y-axis.
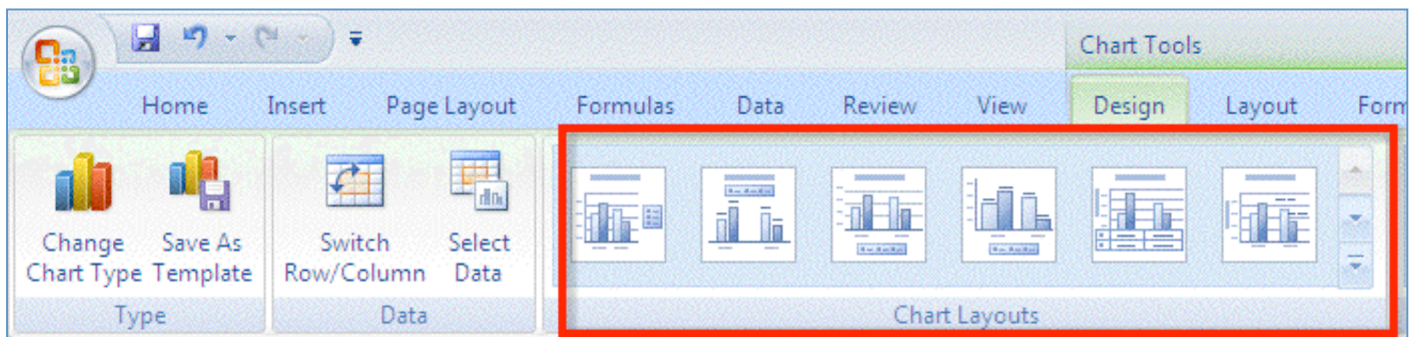
Sample Chart of Example Data

Average number of words recall as a function of condition (experimental drug vs. placebo)
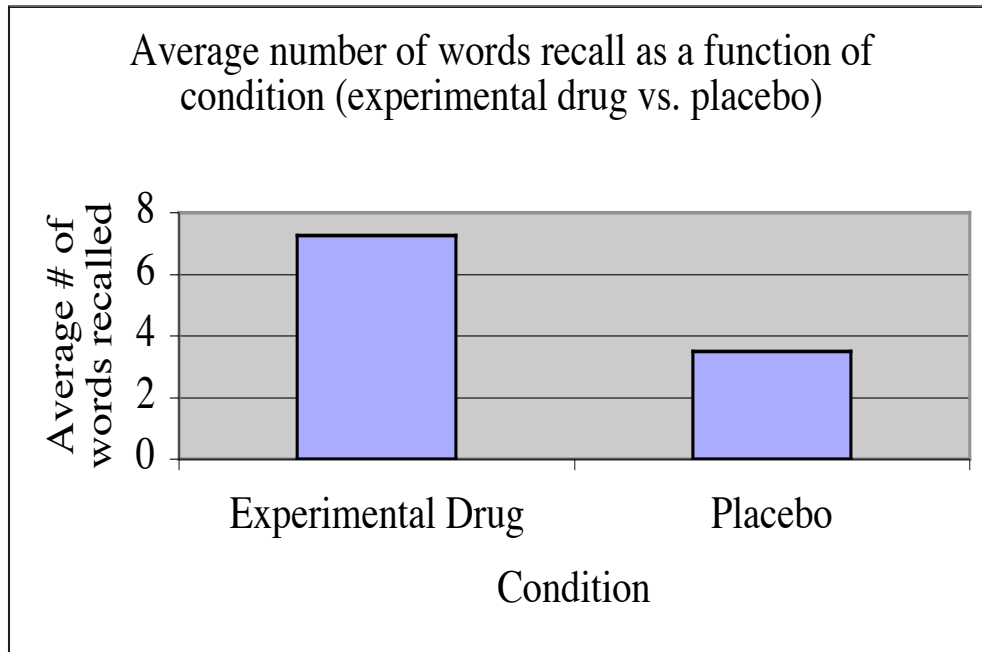
Using Excel 2007 or Newer to Analyze Data

Let's return to our example to make this easier to understand.

Suppose that you, the researcher, are interested in the effects of an experimental drug (vs. placebo) on memory for a list of words. You hypothesize that those taking the experimental drug will exhibit better memory after 8 weeks than those taking the placebo. So you randomly assign people to one of two groups:

1. <u>Experimental Drug Group</u>: this group of 4 people takes the experimental drug for 8 weeks.

<div align="center">

****** OR ******

</div>

2. <u>Placebo Group</u>: this group of four people takes a placebo drug for 8 weeks.

At the end of the 8 weeks, you give them a list of words to study and then measure how many words they recall in 90 seconds and enter the data into Excel (as shown below):

	A	B	C
1			
2		Number of Words Recalled	
3		Experimental Drug	Placebo
4		8	3
5		7	4
6		6	5
7		8	2
8	Average =	7.25	3.50
9			
10			
11		Average # of Words Recalled	
12		Experimental Drug	Placebo
13		7.25	3.50
14			

1. First, determine the Independent Variable and the Dependent Variable:

A. The Independent Variable (IV) is the experimental condition (experimental drug vs. placebo).

This IV is manipulated between-subjects (or independent samples), because 1 group <u>only</u> received the experimental drug & the other group <u>only</u> received the placebo.

B. The Dependent Variable (DV) is the number of words recalled.

2. Determine what statistical test you need to use:

When we used Excel to compute the descriptive statistics, we discovered that the experimental drug group recalled more words on average (M = 7.25) than the Placebo Group (M = 3.50). Thus it would appear that our drug was effective at improving memory.

BUT WAIT!! You **CANNOT** just look at the averages and say, "there is a difference between 7.25 and 3.50!"

Instead, you must conduct a statistical analysis to determine if there is a **significant difference.**

In psychology, we accept a **Significance Level** (or p-value) of .05 or less in order to conclude that the means are statistically, significantly different. If the statistical test reveals that the significant level (p-value) is less than .05, this means that there is **less than a 5% chance that the differences occurred by chance.**

Because the IV is manipulated between-subjects, you will need to calculate an inferential statistic called the **independent samples t-test (i.e. two-samples t-test)** in order to determine if the significance level is .05 or less. If the significance level is .05 or less, then we can conclude that there is a *significant difference* between the experimental drug group average word recall of 7.25 and the placebo group average word recall of 3.50.

3. Conduct the analysis:

a. Click in a cell where you want the results of the statistical analysis to appear. In this example, the cell **D7** was chosen.
b. Click on the **Formula Tab** at the top of the Excel window.
c. Select **More Functions**
d. Select **Statistical, TTEST**
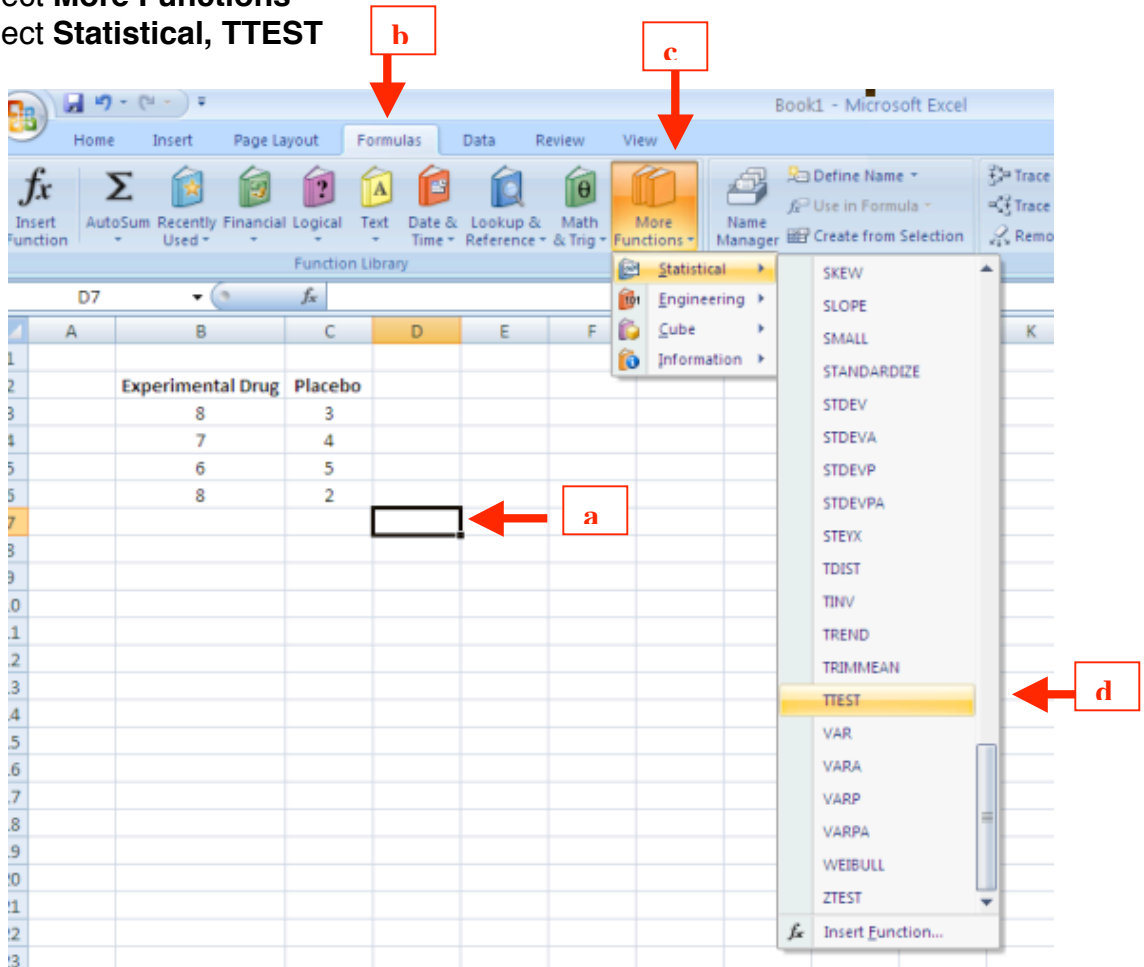
e. In the Functions Arguments Window, enter the data for the first level of the IV (i.e. data for the Experimental Drug group) located in cells B3, B4, B5, and B6 into the textbox called **Array1.** Entering **"B3:B6"** into the Array1 textbox tells Excel to use the constant values found in cells B3, B4, B5, and B6.

f. Enter the data for the second level of the IV (i.e. the data for the placebo group) located in cells C3, C4, C5, and C6 into the textbox called **Array2.** Entering **"C3:C6"** into the Array2 textbox tells Excel to use the constant values found in cells C3, C4, C5, and C6.

g. Enter "**1**" into the textbox called **Tails**. This tells Excel to conduct a **one-tailed test**. You use a one-tailed test when you have a **directional hypothesis**. In this case, you have a directional hypothesis because you expect the results to go in a specific direction - that the experimental drug group will remember more words than the placebo.

***For future reference: You use a **two-tailed test** when you have a **non-directional hypothesis**. If you expect memory differences between the groups, but make no predictions about which group would show enhanced memory performance, then you have a non-directional hypothesis. To conduct a two-tailed test, enter "2" into the textbox called Tails.

h. Enter "**2**" into the textbox called **Type**. This tells Excel what kind of t-test to conduct. In this case we have a between-subjects manipulation of our IV, so we need to conduct an independent samples (also called two-sample t-test). We are going to assume that we have equal variances.

***For future reference: You use a **paired samples t-test** when you have a **within-subjects manipulation of the IV**. If you had a group of participants take the placebo for 8 weeks (then test their memory), and then had the same group take the experimental drug for 8 weeks (and then test their memory again), you would have two memory scores per participants (one score after taking the placebo and one score after taking the experimental drug). This is a within-subjects manipulation because all of the participants are exposed to all levels of the IV. To conduct a paired samples t-test, enter "1" into the textbox called Type.

i. The formula result (= .001722369) appears in the window. This is the significant level.
j. Click OK.

k. The results of the formula will then appear in the Excel spreadsheet.

4. Interpret the Result:

The one-tailed, two samples t-test revealed a significance level of **.0017.**

Remember the significance cut-off or alpha level that we use in the field of psychology is .05. Because .0017 is less than .05, there is less than a 5% chance that the differences in memory performance that we observed between the two groups occurred by chance.

Therefore we can conclude that there is a significant difference between the word recall average by the experimental group of 7.25 and the word recall average by the placebo group of 3.50.