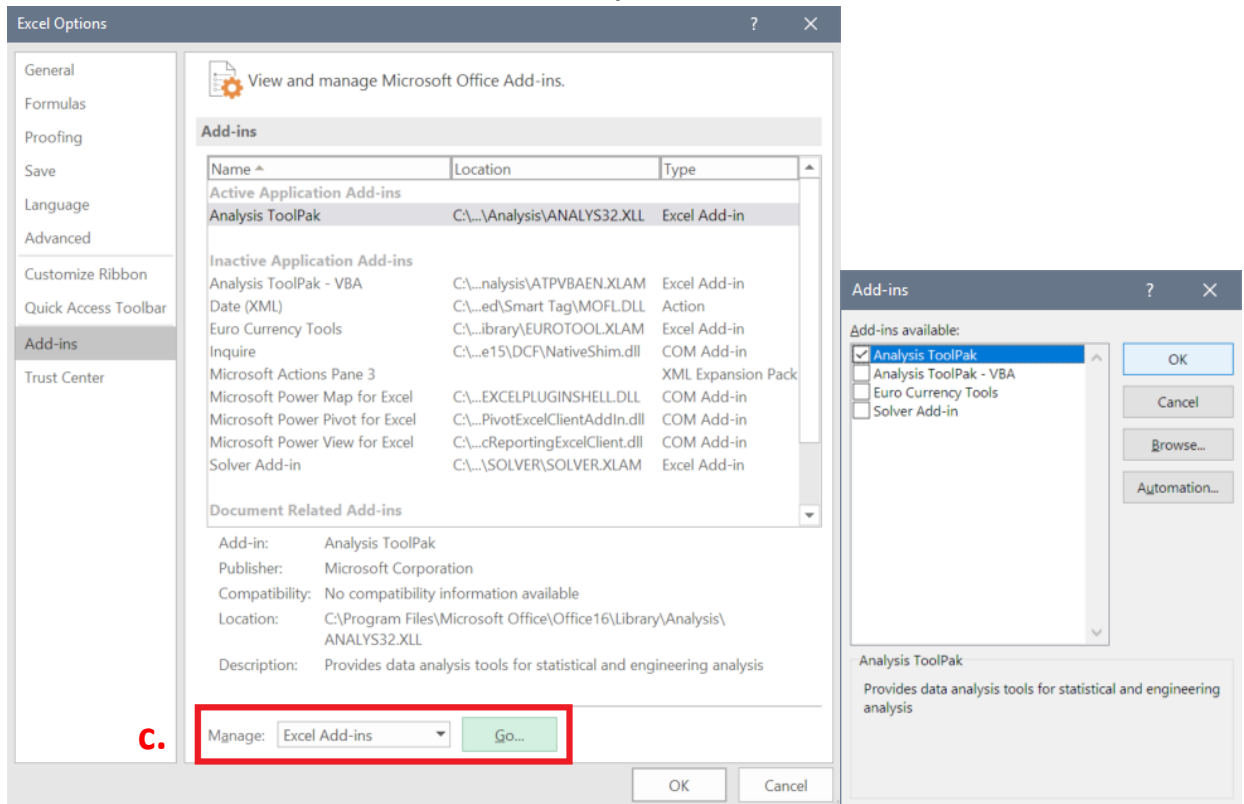


## Agenda:

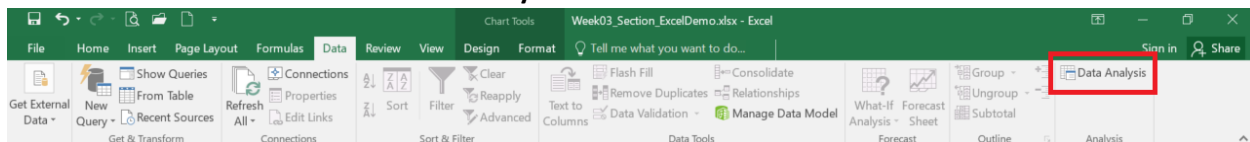
1. Calculating means and SDs
2. Graphing
3. T-tests

## Set-up:

1. Make sure you have some data to work with. Your options are:
  - a. The duration data from Assignment 1. Use this if you have already put your duration measurements into the spreadsheet template Prof. Keating provided
  - b. Fake duration data that I made using the same template. You can download this from CCLE.
2. WINDOWS Users: Activate the Analysis ToolPak add-in on Excel
  - a. **File>Options**
  - b. In the Excel Options window, select **Add-ins**
  - c. In the Manage box, select **Excel Add-ins** and click **Go...**
  - d. In the Add-ins window, select **Analysis ToolPak** and click **OK**



- e. If you get prompted that the Analysis ToolPak is not currently installed on your computer, click **Yes**
- f. You will find the **Data Analysis** command under the **Data** tab



3. MAC Users with Excel 2016 (maybe 2013 too?): Activate the Analysis ToolPak add-in on Excel
  - a. **Tools>Add-ins**
  - b. Select the **Data Analysis ToolPak** and click **OK**
  - c. You will find the **Data Analysis** command under the **Data** tab
4. MAC Users with older versions of Excel: Download StatPlus Mac LE
  - a. Download it here: [www.analystsoft.com/en/products/statplusmacle/](http://www.analystsoft.com/en/products/statplusmacle/)

### Calculating Means and Standard Deviations in Excel

1. The Functions:
  - a. **Mean** =AVERAGE(*data range*)                      e.g. =AVERAGE(B4:B11,D4:D11)
  - b. **SD**     =STDEV(*data range*)                         e.g. =STDEV(B4:B11,D4:D11)
2. Things to think about:
  - a. How to arrange the data so that the values you want to take the mean and SD from are grouped together.
  - b. What is the range of data you want to enter into your function
  - c. Where you want to enter your functions in the spreadsheet so that the means and SD that you are calculating correspond with the data they are taken from.
3. EXERCISE: For the (fake) duration data from Assignment 1, calculate the mean and standard deviation across speakers for each interval in each word you measured (e.g. the mean and SD of the VOT in 'tuppence').

### Graphing in Excel

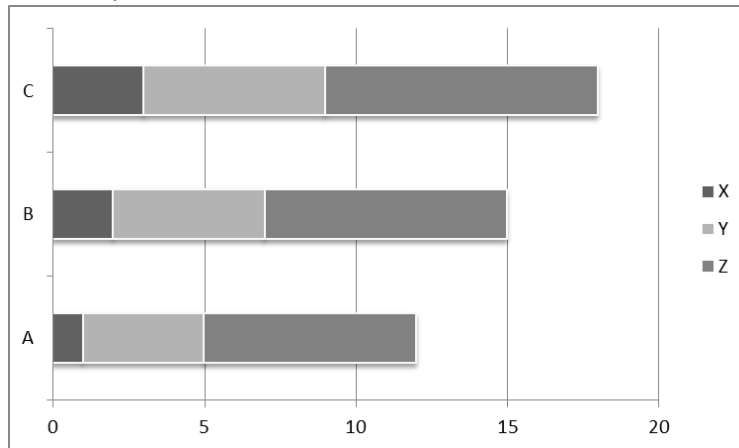
1. It's good practice to graph your data BEFORE you run statistical tests. You can get a feel for the data and make predictions about whether or not your statistical tests will be significant.
2. How to create a graph:
  - a. Arrange your data (e.g. mean values) in continuous data range (i.e. in a row, in a column, or in a matrix). If you are copy-pasting data from a formula to a new location in your spreadsheet, remember to paste it as a VALUE (see handout from Week 0).
  - b. Select the data range you want to graph
  - c. Go to **Insert>Chart** and select the type of graph you want to produce
  - d. CHECK THE GRAPH. Make sure it is representing your data the way you want it to. You will likely have to make some adjustments, such as:
    1. Change the range of your x- or y-axis
    2. Add axis titles
    3. Add a chart title
    4. Change the legend labels

These options are available to you under the **Chart Tools** tab when the graph is selected.

Some useful graphs:

3. Stacked Bar
  - a. Useful for showing total measurements with component parts.

b. An example:



c. The data range that produced this example:

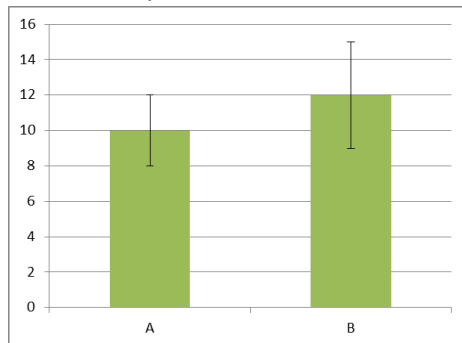
	A	B	C
X	1	2	3
Y	4	5	6
Z	7	8	9

d. Q: What, in our stop duration data, can we represent with this type of graph?

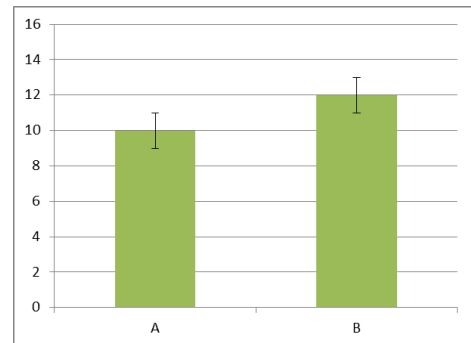
4. (Clustered) Column with error bars

a. Useful for group comparisons. Unlike the stacked bar chart, this type of chart allows us to plot both the mean and standard deviation (as error bars) of any given data group.

b. Some Examples:



	A	B
Mean	10	12
SD	2	3



	A	B
Mean	10	12
SD	1	1

c. Extra steps to generating this type of graph:

1. Create a column chart with ONLY the mean values. This will produce a simple column chart with no error bars.

2. Add error bars

a. Select the chart

- b. Under **Chart Tools>Design** (or **Chart Design** for Macs), select **Add Chart Element>Error Bars>More Error Bars Options...**
- c. In the pop-up window or panel, select **Both** for Direction and **Cap** for End Style.
- d. Under Error Amount, select **Custom** and click **Specify Value**
- e. In the pop-up window, you will be asked to enter a data range for the Positive Error Value and the Negative Error Value. For each of these, delete what is in the field, and select the data range for standard deviations corresponding to the mean values

#### Exercises:

5. Using the (fake) duration data, create a stacked bar graph showing the total duration of [p] in each word (i.e. a pennant, appendix, a penultimate, tuppence) and the component duration measurements within the stop consonant (i.e. voiced portion of the closure, voiceless portion of the closure, VOT)
6. Create a column graph comparing the duration of the voiced portion of the closure to the voiceless portion of the closure in 'a pennant'.

#### T-tests in Excel

WINDOWS (and MAC users who were able to activate the Analysis ToolPak add-in)

1. Under the **Data** tab, click **Data Analysis**
2. In the Data Analysis window, select the kind of test you want to perform. Click **OK**
3. In the new pop-up window, enter in the following:
  - a. **Variable 1 Range:** *Click in the field to activate it, then select a data range in your spreadsheet. Do not include the mean or SD.*
  - b. **Variable 2 Range:** *Click in the field to activate it, then select another data range in your spreadsheet. Do not include the mean or SD.*
  - c. **Hypothesized Mean Difference:** 0 (unless you have reason to change this)
  - d. **Labels:** *Check this if the data ranges in (a) and (b) include a row of column labels.*
  - e. **Alpha:** 0.05 (unless you have reason to change this)
  - f. **Output options:** *Select where you want to put the results of the t-test*

Click **OK**

4. Get the relevant numbers from the output table of results

- a. **degrees of freedom**
- b. **t-value**
- c. **p-value** (take the value from the two-tailed test unless you have reason to choose otherwise)

	Variable 1	Variable 2
Mean	19.81508	60.76994
Variance	110.2335	93.95193
Observations	8	8
Pearson Correlation	-0.49106	
Hypothesized Mean Difference	0	
df	7	<b>a.</b>
t Stat	-6.64231	<b>b.</b>
P(T<=t) one-tail	0.000146	
t Critical one-tail	1.894579	
P(T<=t) two-tail	0.000293	<b>c.</b>
t Critical two-tail	2.364624	

MAC (if you don't have the add-in and had to download StatPlus)

1. Open StatPlus.
2. In StatPlus, open up the Excel spreadsheet containing your duration data
3. Select **Basic>Compare Means [T-Test]...**
4. In the new pop-up window, enter in the following:
  - a. **Variable #1:** Click the button to the right that has a tiny image of a spreadsheet on it. This will bring your spreadsheet forward. Select a set of data in your spreadsheet. Do not include the mean or SD.
  - b. **Variable #2:** Click the button to the right that has a tiny image of a spreadsheet on it. This will bring your spreadsheet forward. Select another set of data in your spreadsheet. Do not include the mean or SD.
  - c. **Labels In First Row:** Check this if your data ranges in (a) and (b) include a row of column labels at the top.
  - d. **T-Test Type:** Select the type of t-test you want to perform
  - e. **Hypothesized Mean Difference:** 0 (unless you have reason to change this)
  - f. The alpha value is set by default at 5%
  - g. The output option is set by default to produce the results table in a new workbook.

Click **OK**

5. Get the relevant numbers from the output table of results
  - a. **degrees of freedom (df)**
  - b. **t-value**
  - c. **p-value** (take the value from the two-tailed test unless you have reason to choose otherwise)

Compare Means				
Descriptive Statistics				
VAR	Sample size	Mean	Standard Deviation	Variance
Column B	8	19.81508	10.49922	110.23354
Column C	8	60.76994	9.69288	93.95193
Paired two-sample t-test				
Degrees of Freedom	7	a.		
Hypothesized Mean Difference	0			
Pooled Variance	102.09274			
Test Statistics	6.64231	b.		
Pearson R	-0.49106			
Two-tailed distribution				
p-level	c.	0.00029	Critical Value (5%)	2.36462
One-tailed distribution				
p-level		0.00015	Critical Value (5%)	1.89458
G-criterion				
Test Statistics	1.42262	Critical Value (5%)		0.30600
p-level	0			
Pagurova criterion				
Ratio of variances parameter	0.53987			
Test Statistics	8.10659	Critical Value (5%)		0.06385
p-level	1.00000			