Graphing Data and Statistical Analysis with Excel

Instructions: In this exercise, you will apply your basic knowledge of Microsoft Excel to analyze data using Excel graphing tools and its built-in statistical functions. From the data tables, you will create plots, calculate and graph averages and standard deviations, compute other central tendency numbers, and calculate *p*-values using the T-distribution.

Set up a work session:

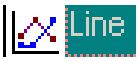
1. Open a new Excel session. Use a full screen window.
2. Use the data set provided below. Alternately, open the workbook posted on Blackboard.

Guided Practice: *Average Faculty Salaries, Males vs. Females*

Instructions. For the next data set, *Average Faculty Salaries, Males vs. Females*, perform Exercises 1 - 6. Correctly label all your formatted graphs and tables with results.
Save your practice in an Excel file named like this: Salaries_YourFullName_Period.xls.

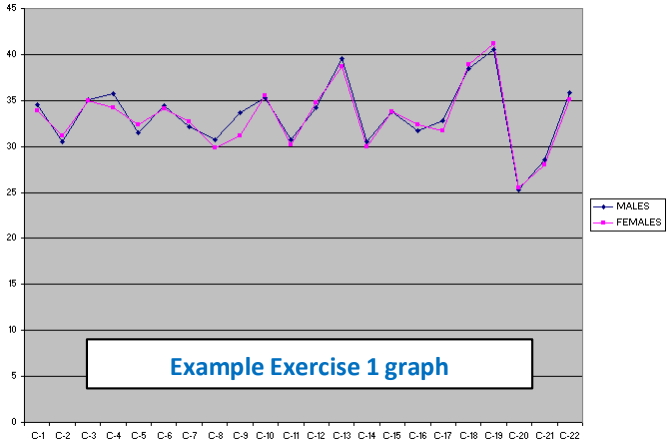
College	Males	Females
C-1	34.5	33.9
C-2	30.5	31.2
C-3	35.1	35.0
C-4	35.7	34.2
C-5	31.5	32.4
C-6	34.4	34.1
C-7	32.1	32.7
C-8	30.7	29.9
C-9	33.7	31.2
C-10	35.3	35.5
C-11	30.7	30.2
C-12	34.2	34.8
C-13	39.6	38.7
C-14	30.5	30.0
C-15	33.8	33.8
C-16	31.7	32.4
C-17	32.8	31.7
C-18	38.5	38.9
C-19	40.5	41.2
C-20	25.3	25.5
C-21	28.6	28.0
C-22	35.8	35.1

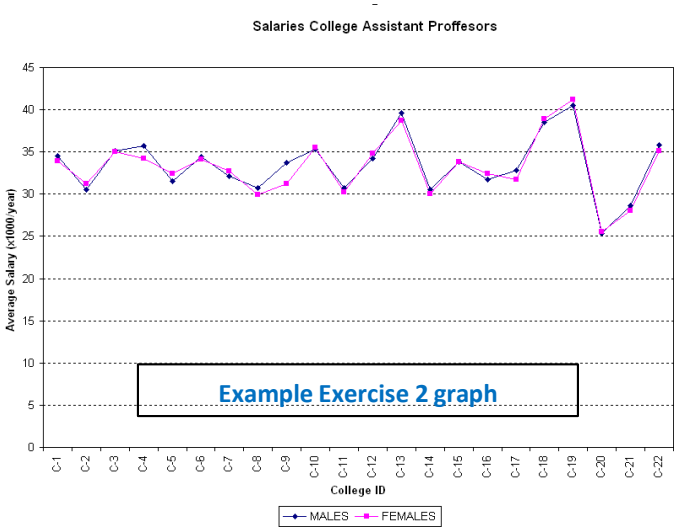
Exercises:

1. *Creating a graph*
   For the paired data set 1, create a line graph. Place this graph *as a new sheet.*
   (Hint: Select data columns *Males – Females* ►  

)
   To make the values in column *College* be the *x*-values in this graph:
   - In the Chart Wizard – Step 2 of 4 – Chart Source Data , select tab Series.
   - Click on box: 

click here
   - Using the mouse, select only the data in column *College* ► press Enter.





2. *Formatting a graph*
   a. Place the graph legend at the bottom of the graph.
   b. (If using Excel 2007 or earlier) Eliminate the plot area default gray color. (*Hint*: Click on Plot Area ► Format ► Select Plot Area ordouble click on Plot Area.)
   c. Change the major gridlines to a broken line. (*Hint*: *Double click in one of the gridlines.*)
   d. Insert the next labels. For *x*-axis: *College ID*; for *y*-axis: *Average Salary* (x*1,000/year*) (*Hint:* Chart ► Chart Options ► Titles)
   e. Include in the graph title: *College Assistant Professor Salaries. Males vs. Females*

3. *Calculating statistics*
   a. Compute the data differences.
   b. Compute samples/differences means. [*Hint*: use function =average()]
   c. Compute sample/differences standard deviations. [*Hint*: use function =stdev()]
   d. Find the sample/differences maximum values. [*Hint*: use function =max()]
   e. Find the sample/differences minimum values. [*Hint*: use function =min()]
   f. Find the sample/differences ranges.
   g. Find the sample/differences medians. [*Hint*: use function =median()]

Example Exercise 3 results →

	A	B	C	D
8	College	MALES	FEMALES	Difference
9	C-1	34.5	33.9	0.6
10	C-2	30.5	31.2	-0.7
11	C-3	35.1	35	0.1
12	C-4	35.7	34.2	1.5
13	C-5	31.5	32.4	-0.9
14	C-6	34.4	34.1	0.3
15	C-7	32.1	32.7	-0.6
16	C-8	30.7	29.9	0.8
17	C-9	33.7	31.2	2.5
18	C-10	35.3	35.5	-0.2
19	C-11	30.7	30.2	0.5
20	C-12	34.2	34.8	-0.6
21	C-13	39.6	38.7	0.9
22	C-14	30.5	30	0.5
23	C-15	33.8	33.8	0
24	C-16	31.7	32.4	-0.7
25	C-17	32.8	31.7	1.1
26	C-18	38.5	38.9	-0.4
27	C-19	40.5	41.2	-0.7
28	C-20	25.3	25.5	-0.2
29	C-21	28.6	28	0.6
30	C-22	35.8	35.1	0.7
31		Males	Females	Difference
32	Mean	33.4318	33.2	0.23182
33	St Dev	3.54621	3.60119	0.84594
34	Max	40.5	41.2	2.5
35	Min	25.3	25.5	-0.9
36	Range	15.2	15.7	3.4
37	Median	33.75	33.25	0.2

4. *Graphing data differences*

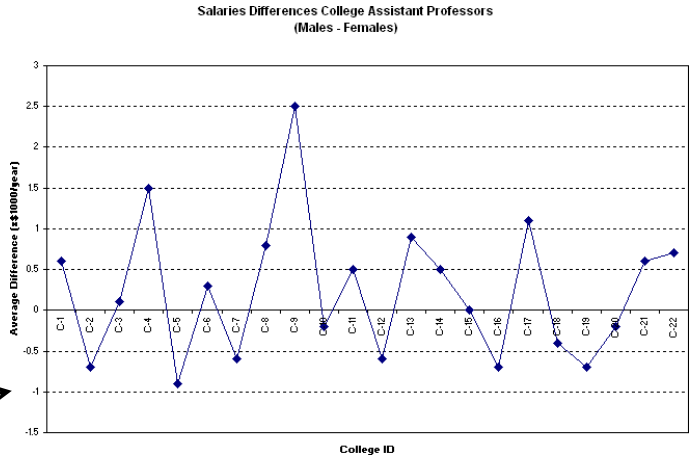
   Repeat Exercises 1 and 2 for the data differences obtained in Exercise 3, with the next changes:

   a. Delete the graph legend.

   b. Add a y-axis label: *Average Differences* (x *$1,000/year*)

   c. Title the graph: *College Assistant Professor Salary Differences: Males vs. Females*

Example Exercise 4 results



5. *Graphing mean and standard deviation for the differences*

   a. Include in the graph a horizontal line representing the sample mean. (*Hint:* Create a list with *mean values,* then Chart ► Source Data ► Add [*Select created data*].)

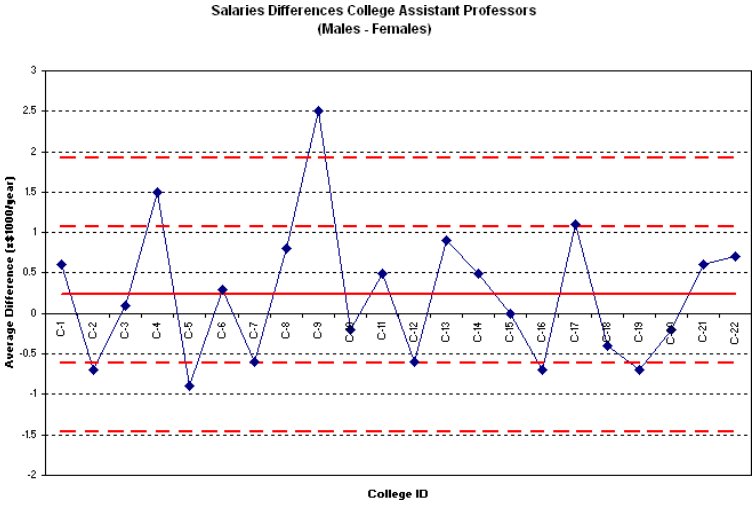
   b. Include in the graph horizontal lines representing mean ± 1 standard deviation. (Hint: Create list with ± SD, then ► Source Data ► Add [*Select created data].*)

   c. Include in the graph a horizontal line representing mean ± 2 standard deviations.

   d. Format the sample mean line: Change the color to red and select the next thicker line. (*Hint*: Double click on line.)

   e. Format the standard deviation lines: Change the color to red and select a broken thicker line (*Hint*: Double click on line.)

Example Exercise 5 graph



6. *Compute the sample differences t-value, p-value and sampling standard deviation*

Test whether the difference in salaries is statistically significant i.e. test

H_o: (Mean (Males) - Mean(Females)) = 0

Vs

H_a: Mean (Males) > Mean(Females)

a. Compute the sample associated *t*-value or sample test statistic.

Use equation: $t = (\bar{d} - 0) \cdot \sqrt{n} / s_d$, *where* $\bar{d}$ *is difference mean, n is sample size, and* S_d *is difference standard deviation. Using the values in the table:*

$$t = \frac{0.23182 \cdot \sqrt{22}}{0.84594} = 1.28535$$

b. Compute *p*-value using T-distribution.

We can use the function

=tdist(D32,21,1)

where "1" indicates one-tail test.

Alternately, use function ttest() with the values in table shown in step 2

=ttest(B8:B30,C8:C30,1,1)

where the first "1" indicates one-tail test, and the second "1" indicates a paired test.

c. Compute the sampling standard deviation for this difference.

Use equation: $s_{\bar{d}} = s_d / \sqrt{n}$:

$$s_d = \frac{0.84594}{\sqrt{22}} = 0.18035$$

	A	B	C	D
20	C-12	34.2	34.8	-0.6
21	C-13	39.6	38.7	0.9
22	C-14	30.5	30	0.5
23	C-15	33.8	33.8	0
24	C-16	31.7	32.4	-0.7
25	C-17	32.8	31.7	1.1
26	C-18	38.5	38.9	-0.4
27	C-19	40.5	41.2	-0.7
28	C-20	25.3	25.5	-0.2
29	C-21	28.6	28	0.6
30	C-22	35.8	35.1	0.7
31		Males	Females	Difference
32	Mean	33.4318	33.2	0.23182
33	St Dev	3.54621	3.60119	0.84594
34	Max	40.5	41.2	2.5
35	Min	25.3	25.5	-0.9
36	Range	15.2	15.7	3.4
37	Median	33.75	33.25	0.2
38				
39			t-value	1.28535
40			p-value	0.10633
41			S-StDev	0.18035

Example Exercise 6 results

d. Do your results support the claim that no significant salary difference exists between male and female college professors…

…at the 5% level of significance?

…at the 10% level of significance?

Because *p*-value = 0.10633 is greater than 0.05 or 0.10, we have no evidence at the 5% or 10% level of significance to reject the original assumption (H_0) that female assistant professors receive, on average, the same as the male assistant professors.