**Graphing Data and Statistical Analysis with Excel Practice**

*Instructions*: In this practice, 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 scatter 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 sets provided below.

**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 ID** | **Male AP** | **Female AP** |
| 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 – Female*s ► ► ► )

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.

**Example Exercise 1 graph**

**Example Exercise 2 graph**

1. *Formatting a graph*
2. Place the graph legend at the bottom of the graph.
3. Eliminate the plot area default gray color. (*Hint*: Click on Plot Area ► Format ► Select Plot Area or double click on Plot Area.)
4. Change the major gridlines to a broken line. (*Hint*: *Double click in one of the gridlines.*)
5. Insert the next labels. For *x*-axis: *College ID*; for *y*-axis: *Average Salary* (x*1,000/year*)
(*Hint:* Chart►Chart Options ► Titles)
6. Include in the graph title: *College Assistant Professor Salaries. Males vs. Females*
7. *Calculating statistics*
8. Compute the data differences.
9. Compute samples/differences means.
[*Hint*: use function =average()]
10. Compute sample/differences standard deviations.
[*Hint*: use function =stdev()]
11. Find the sample/differences maximum values.
[*Hint*: use function =max()]
12. Find the sample/differences minimum values.
[*Hint*: use function =min()]
13. Find the sample/differences ranges.
14. Find the sample/differences medians.
[*Hint*: use function =median()]

**Example Exercise 3 results**



1. *Graphing data differences*

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

* 1. Delete the graph legend.
	2. Add a y-axis label: *Average Differences* (x *$*1,000*/year*)
	3. Title the graph: *College Assistant Professor Salary Differences: Males vs. Females*

**Example Exercise 4 results**

1. *Graphing mean and standard deviation for the differences*
	1. 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*].)
	2. Include in the graph horizontal lines representing mean ± 1 standard deviation.
	(Hint: Create list with ± SD, then ► Source Data ► Add [*Select created data].*)
	3. Include in the graph a horizontal line representing mean ± 2 standard deviations.
	4. Format the sample mean line: Change the color to red and select the next thicker line. (*Hint*: Double click on line.)
	5. 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**

1. *Compute the sample differences t-value, p-value and sampling standard deviation*
	1. Compute the sample associated *t*-value or sample test statistic.

*Use equation:* , *where* *is difference mean, n is sample size*, *and Sd is difference standard deviation. Using the values in the table:*

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

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

* 1. Compute the sampling standard deviation for this difference.
	*Use equation*: :

$$S\_{\overbar{d}}=\frac{0.84594}{\sqrt{22}}=0.18035$$

* 1. Do your results support the claim that no significant salary difference exists between male and female college professors…
	…at the 5% level of significance?

**Example Exercise 6 results**

…at the 10% level of significance?

(*Write your conclusions in a textbox on the results spreadsheet. Include an explanation. Example:*)

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 salary as the male assistant professors.

**Exercises:**

1. Create a graph
2. Format a graph (a-e)
3. Calculate statistics (a-g)
4. Graph data differences (a-c)
5. Graph mean and standard deviation for the differences (a-e)
6. Compute the sample differences t-value, p-value and sampling standard deviation (a-d)

**Independent Practice: *Unemployment: College vs. High School Graduates***

*Instructions:* For the next data set, *Unemployment: College vs. High School Graduates*, perform Exercises 1 - 6.
Correctly label all your formatted graphs and tables with results.
Save your practice in an Excel file named like this: Unemployment\_YourFullName\_Period.xls.

|  |  |  |
| --- | --- | --- |
| **Year** | **College** | **High School** |
| 1999 | 2.8 | 5.9 |
| 2000 | 2.2 | 4.9 |
| 2001 | 2.2 | 4.8 |
| 2002 | 1.7 | 5.4 |
| 2003 | 2.3 | 6.3 |
| 2004 | 2.3 | 6.9 |
| 2005 | 2.4 | 6.9 |
| 2006 | 2.7 | 7.2 |
| 2007 | 3.5 | 10.0 |
| 2008 | 3 | 8.5 |
| 2009 | 1.9 | 5.1 |
| 2010 | 2.5 | 6.9 |

**Independent Practice: *Birth Rates vs. Death Rates***

*Instructions*. For the data set, *Birth Rates vs. Death Rates*, perform Exercises 1 - 6.
Correctly label all your formatted graphs and tables with results.
Save your practice in an Excel file named like this: BDRates\_YourFullName\_Period.xls.

|  |  |  |
| --- | --- | --- |
| **County ID** | **Birth** | **Death** |
| CO-01 | 12.7 | 9.8 |
| CO-02 | 13.4 | 14.5 |
| CO-03 | 12.8 | 10.7 |
| CO-04 | 12.1 | 14.2 |
| CO-05 | 11.6 | 13.0 |
| CO-06 | 11.1 | 12.9 |
| CO-07 | 14.2 | 10.9 |
| CO-08 | 12.5 | 14.1 |
| CO-09 | 12.3 | 13.6 |
| CO-10 | 13.1 | 9.1 |
| CO-11 | 15.8 | 10.2 |
| CO-12 | 10.3 | 17.9 |
| CO-13 | 12.7 | 11.8 |
| CO-14 | 11.1 | 7.0 |