

GRAPHING IN EXCEL

EXCEL LAB #2

BUSN/ECON/FIN 130: Applied Statistics
Department of Economics and Business
Lake Forest College
Lake Forest, IL 60045
Copyright, 2013

Overview

This lab is written for Excel 2010, which is available to students in the library. The notation => can be read as "go to" or "click on." This notation will most often be used when navigating the menu or toolbars in Excel. To indicate a command or icon that you might click on or search for in Excel, **bold** will be used. Likewise, anything that you are to type into Excel will be **bolded** in the instructions. Do not enter such text as bolded text unless the instructions ask you to do so.

Tutorial

1. Open **2002 International Transparency Data.xlsx**, containing 5 variables and 179 observations on most of the world's countries in 2002.
2. Rename the **Sheet 1** that contains the data to **Original Data**. Rename **Sheet 2** to **Tutorial**. Rename **Sheet 3** to **Exercises** (hint: right click on the sheet name => rename).
3. Copy the five columns of data, and paste them into the **Tutorial** worksheet: in **Original Data**, **left click on column A and drag through column E => right click on the shaded area => Copy => left click on the Tutorial worksheet tab => right click on cell A1 => Paste**.
4. In the **Tutorial** worksheet, set all columns to be right justified with Arial 9 point font: **left click on the blank box to the left of column A and above row 1** (there is a small shaded triangle in the box, and when you click on it the entire worksheet will be highlighted) => **Home** tab => **Alignment** box => **click on right-justify icon** (third from the left in the second row) => **Font** box => **select Arial and 9 point font**.
5. Set the width of column F and G to 15 and 12 respectively (hint: right click on a column and select column width).
6. In cell F1 enter **Income Level** (in bold). In cell G1 enter **Income Value** (in bold).

7. We are going to create two new variables that assign countries to one of three income levels:

Low income: per capita GDP \leq \$5,000
Middle income: \$5,000 < per capita GDP \leq \$15,000
High income: \$15,000 < per capita GDP.

We will do this with a couple of **IF** commands. The **IF** command in Excel requires a binary logic statement, then a value to assign to the cell if the statement is true, and then a value to assign the statement if the statement is false.

Enter in cell G2: **=1+IF(E2>5000,1,0)+IF(E2>15000,1,0)**.

1. The first part of this function, the 1, assigns a value of 1 to the cell regardless of what comes later.
 2. The first IF statement adds either 1 or 0 to the cell based on whether the expression in the first IF statement is true. Specifically, if the value in cell E2 is greater than \$5,000, then the formula adds 1 more to the original 1 for a value of 2. This happens for all per capita GDP levels above \$5,000 (in particular, this part of the formula adds one for middle income and high income GDP levels).
 3. The second IF statement adds either 1 or 0 to the cell based on whether the expression in the second IF statement is true. Specifically, if the value in cell E2 is greater than \$15,000, then the formula adds 1 more to the cell. Thus, for any country with per capita GDP greater than \$15,000, the total value of 3.
 4. Be sure you understand why the formula entered into cell G2 assigns a value of:
1 to all countries with per capita GDP \leq \$5,000
2 to all countries with \$5,000 < per capita GDP \leq \$15,000
3 to all countries with per capita GDP > \$15,000 as desired.
8. Copy and paste cell G2 to cells G3 through G180: **right click on cell G2 => Copy => left click on cell G3 and drag through G180 => right click on the shaded area => Paste.**
 9. We now want to assign the income levels of "Low Income," "Middle Income," and "High Income" to countries appropriately. To do this, we are going to use nested **IF** functions. You will also notice that when we want to assign words to the values associated with **IF** functions, the words need to be in quotes.

In cell F2 enter:

=IF(G2=1, "Low Income",IF(G2=2, "Middle Income",IF(G2=3, "High Income", "Error")))

The formula gives the cell the following options:

1. If cell G2 = 1, the function returns a value of "Low Income."
 2. Given cell G2 \neq 1, if G2 = 2, the function returns a value of "Middle Income."
 3. Given cell G2 \neq 1 or 2, if G2 = 3, the function returns a value of "High Income."
 4. Given cell G2 \neq 1, 2, or 3, the function returns a value of "Error."
10. Copy and paste cell F2 to cells F3 through F180.

11. Copying and Pasting Values from Formulas.

As we will be moving some columns around eventually, we want to change columns F and G from being formula-based to reporting the actual values of the formulas.

To copy and paste columns F and G to where they already are, but to “paste” the values into the cells rather than having the cell be formula based: left click on column F => drag to highlight column G => right click in the shaded area => **Copy** => right click in the shaded area again => **Paste Special** => **Values** (paste) => **OK**. If you click on cell F2 or G2 and look in the textbox, you will see that the formulas that derived these values have now been replaced with the actual values.

12. Creating a Frequency Table.

The first table we are going to make in Excel is a frequency table of the **Transparency** variable.

1. Set column I's width to 12. Set the column J through O's width to 10.
2. Merge cells I1 through K1: **left click on cell I1 and drag through cell K1** => **right click on the shaded area** => **Format Cells** => **Alignment** (tab) => **Merge Cells** => **OK**.
3. In cell I1 (which is now I1 through K1), enter **Transparency Frequency Table** in bold, underlined, and horizontally centered. In cells I2 through K2 respectively enter **Tranparency**, **Count**, **Freq (%)** (all underlined but not bolded).
4. As the **Transparency** variable takes on the values of Open (if the county is transparent to the United Nations regarding its conventional arms trades) and Closed (if the country is not transparent to the United Nations), in cell I3 enter **Open** and in cell I4 enter **Closed** (neither bolded).
5. To have Excel provide the country counts of both values of **Transparency**, we use the **COUNTIF** command.

In cell J3 enter: **=COUNTIF(\$C\$2:\$C\$180,I3)**.

The **COUNTIF** formula counts the number of cells from C2 through C180 that have the same value as cell I3. You should see that there are 112 “open” countries in the sample. The dollar signs before the C's and before the 2 and the 180 allow us to copy and paste the formula without having Excel change the range of the cells being considered in the **COUNTIF** command. Without the dollar signs, Excel automatically “rolls” the columns and rows of formulas when they are copied and pasted elsewhere.

6. Copy and paste cell J3 to cell J4. You should see that Excel is counting 67 countries as being “closed”.

Click on cell J4, and look at the textbox at the top of the screen. You will see the formula **=COUNTIF(\$C\$2:\$C\$180,I4)**. In particular, notice that the data range of C2 to C180 has not changed (thanks to the dollar signs), but the qualifying cell I4 has changed from I3. So in cell J4 Excel is counting the number of cells in C2 through C180 that have the same value as cell I4, which is “closed”.

7. At this point, we have 112 countries that are open and 67 countries are closed. The table does not yet, however, report the distribution in percentages. The following steps will produce a percentage column.
 1. In cell I7 enter **Total** (not in bold).
 2. In cell J7 enter: **=SUM(J3:J4)**. The value 179 should appear in cell J7.
 3. To calculate the percentages, in cell K3 enter: **=J3/\$J\$7**. Then copy and paste cell K3 to cell K4. Notice that you can copy and paste cell K3 into cell K4 to calculate the percent of observations that are closed because of the dollar signs in **\$J\$7**. Without the dollar signs, the pasting would not only have rolled the **J3** to **J4** (which it did, and which we needed to have happen), but Excel would have also rolled the **J7** to **J8** which would not be correct.
 4. Set the format of cell K3 and K4 to report in percentages to one decimal place: **left click on cell K3 and drag through cell K4 => right click in the shaded area => Format Cells => Number (tab) => Percentage (category) => 1 (decimal places) => OK.**

Cells K3 and K4 should now be showing that the 112 open countries represent 62.6% of the sample of 179 countries and that the 67 closed countries represent 37.4% of the sample of 179 countries. The entire table should look like the following:

<u>Transparency Frequency Table</u>		
<u>Transparency</u>	<u>Count</u>	<u>Freq %</u>
Open	112	62.6%
Closed	67	37.4%

13. Next we will repeat what we have just done to create a frequency table of the **Freedom** variable.
 1. Merge cells M1 through O1: **left click on cell M1 and drag through cell O1 => right click on the shaded area => Format Cells => Alignment (tab) => Merge Cells => OK.**
 2. In cell M1 (which is now M1 through O1), enter **Freedom Frequency Table** (both in bold and underline). In cells M2 through O2 respectively enter **Freedom**, **Count**, **Freq (%)** (all underlined but not bolded).
 3. Center columns M through O horizontally and vertically: **left click on column M and drag through column O => Home tab => Alignment box => horizontally-centered and vertically-centered icons.**
 4. From column D we know that the **Freedom** variable takes on the values of Free (if the county is Free according to Freedom House), Partially Free (if the country is partially free according to Freedom House), and Not Free (if the country is not free in its personal freedoms or civil liberties according to Freedom House). Therefore, in cell M3 enter **Free**, in cell M4 enter **Partially Free**, and in cell M5 enter **Not Free** (none in bold).
 5. Have Excel provide the country counts of all three values of **Freedom** in cells N3 through N5: in cell N3 enter **=COUNTIF(\$D\$2:\$D\$180,M3)** => **right click on cell N3 => Copy => left click on cell N4 and drag through cell N5 => right**

click on the shaded area => Paste. Cell N3 should now have a value of 81. Cells N4 and N5 should now have values of 52 and 46 respectively.

6. Add up all of the Freedom counts: in cell M7 enter **Total** => in cell N7 enter **=SUM(N3:N5)**. Once again, as every country is classified as being free, partially free, or not free, the total number of countries is 179.
7. Calculate the percentages of each class in cells O3 through O5: **in cell O3 enter =N3/\$N\$7 => right click on cell O3 => Copy => left click on cell O4 and drag through cell O5 => right click in the shaded area => Paste.**
8. Set the format of cells O3 through O5 to report in percentages to one decimal place: **left click on cell O3 drag through cell O5 => right click in the shaded area => Format Cells => Number (tab) => Percentage (category) => 1 (decimal places) => OK.**

The Freedom Table should look like the following:

<u>Freedom</u>	<u>Count</u>	<u>Freq %</u>
Free	81	45.3%
Partially Free	52	29.1%
Not Free	46	25.7%

14. **Creating a Histogram.**

Next we will use Excel to create histogram charts of both frequency tables.

1. Create a rough histogram of the Transparency counts: **left click on cell I3 and drag through cell J4 => Insert tab => Chart box => Column => 2-D Column** (left icon in the top row).
2. Your graph will appear somewhere on the worksheet.
 1. Left click on its boundary and drag it so that the top left corner begins at cell I9.
 2. Look at your toolbar. At this point you should see that **Insert** is selected, and to the far right is an options box at the very top of the screen called **Chart Tools**. This is a very important box when creating graphs in Excel. Notice that within **Chart Tools** are three options – **Design**, **Layout**, and **Format**.
 3. We want our histogram to have 2 particular features and not to have 2 other particular features. Specifically, we want a title and a y-axis title (**Number of Countries**) but we don't want the graph to have an x-axis title (because the histogram will already label the x-axis categories) nor do we want a legend as a legend is superfluous in a histogram that only has one variable). To accomplish all of this:
 1. Under **Chart Tools** => **Layout** => **Labels** box => **Chart Title** => **Above Chart** (third from top).
 2. Under **Chart Tools** => **Layout** => **Labels** box => **Axis Titles** => **Primary Vertical Axis Title** => **Rotated Title** (second from top).
 3. There should not be an x-axis title in your histogram right now, but if there is, you can get rid of it in the following steps: **Chart Tools**

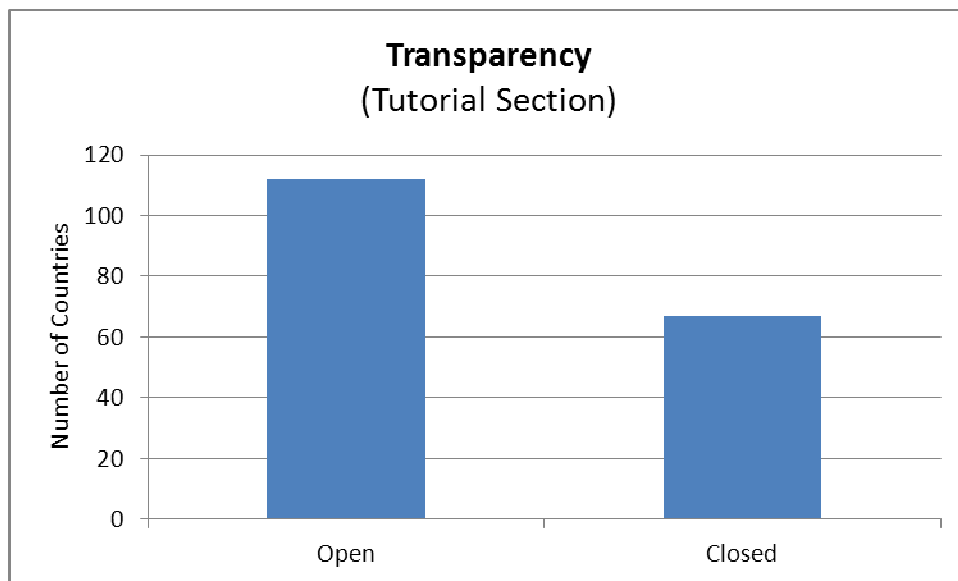
=> **Layout => Labels box => Axis Titles => Primary Horizontal Axis Title => None** (first from top).

4. There is most likely a legend box in your histogram right now (and most likely labeled Series 1. This needs to be removed: **Chart Tools => Layout => Labels box => Legend => Legend Off** (first from top).
3. If you want to change a label in a graph that already exists, you can click on the label and directly edit the text. For example, click on the title that current reads **Chart Title**, and edit it to read

Transparency (Tutorial Section)

Make **Transparency** bolded but (Tutorial Section) not bolded. And make the entire title 14 point font (under the **Home** tab).

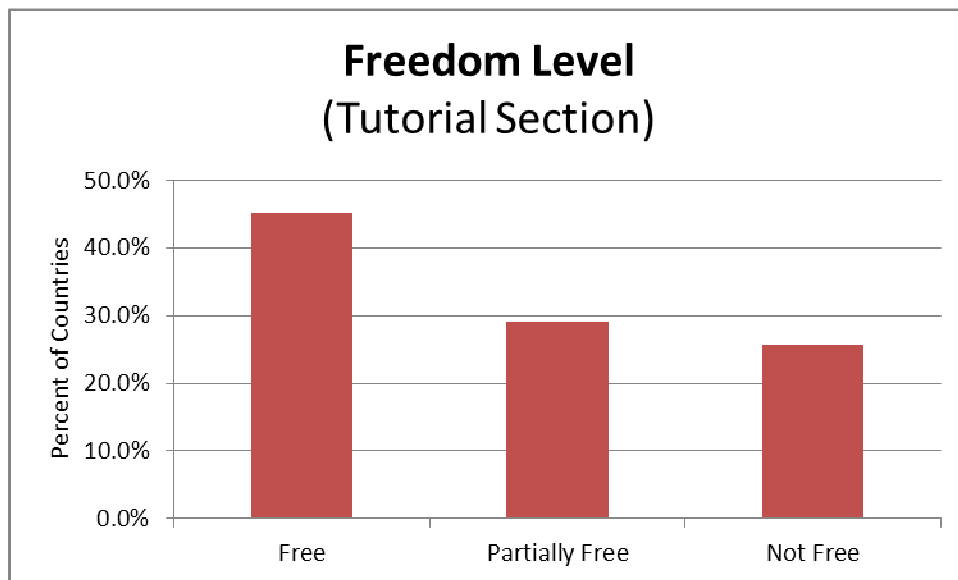
4. If you want to change an axis label in a graph that already exists, you can click on the label and directly edit the text. For example, click on the y-axis label and enter **Number of Countries** (not in bold).
5. Your histogram of **Transparency** now looks like the following:



15. Next we will create the histogram for **Freedom**, but instead of reporting country counts we will report the percent frequency.

1. **Left click on cell M3 and drag through cell M5 => press and hold the Control Key => left click on cell O3 and drag through cell O5 => Insert tab => Charts box => Column => 2-D Column** (left icon in the top row).

2. Your graph will appear somewhere.
 1. Click on its border and drag it so that the top left corner begins at cell I34.
 2. Give the histogram a title and y-axis label but not an x-axis label or a legend:
 1. Under **Chart Tools** => **Layout** => **Labels** box => **Chart Title** => **Above Chart** (third from top).
 2. Under **Chart Tools** => **Layout** => **Labels** box => **Axis Titles** => **Primary Vertical Axis Title** => **Rotated Title** (second from top).
 3. There should not be an x-axis title in your histogram right now, but if there is, you can get rid of it in the following steps: **Chart Tools** => **Layout** => **Labels** box => **Axis Titles** => **Primary Horizontal Axis Title** => **None** (first from top).
 4. There is most likely a legend box in your histogram right now (and most likely labeled Series 1. This needs to be removed: **Chart Tools** => **Layout** => **Labels** box => **Legend** => **Legend Off** (first from top).
 3. Click on the title of the graph, and make it be **Freedom Level** (in bold).
 4. Click on the y-axis label and make it **Percent of Countries** (no bold).
3. The graph seems a bit busy as there are horizontal lines drawn at every 5% intervals. For just three categories (Free, Partially Free, Not Free), this seems a bit excessive. Right click on any of the numbers on the y-axis => **Format Axis** => **Major Unit** (Axis Options) => **Fixed** => **0.10** => **Close**.
4. You may want to change the color of the bars in the histogram: **Insert** tab => **Chart Tools** (toolbar) => **Design** box => **select the color of your choice**. Your Freedom histogram should now look like the following:



The next graph we want to create in Excel is a histogram of freedom level (just like above), except we want to have freedom level separated out by transparency status. The first step is to have Excel create the cross-tabulation of freedom and transparency. In Excel, cross-tabulations are called pivot tables.

16. **Creating Pivot Tables (Cross-Tabulations).**

1. On the toolbar: **Insert** tab => **Tables** box => **Pivot Table** => **Pivot Table** => enter **C1:D180** (range) => select **Existing Worksheet** => enter **T4** under location => **OK**.
2. Several boxes will appear on the worksheet. Under **PivotTable Field List** you will see two variables – **Transparency** and **Freedom**.
 1. Left click and drag **Freedom** to the **Row Labels** box.
 2. Left click and drag **Transparency** to the **Column Labels** box.
 3. Left click on **Freedom** again, and drag it to the **Values** box.

A table with counts (freedom by transparency) will appear. For example, there are 14 closed and free countries, and there are 33 open and partially free countries.

If any other Excel pivot table boxes are still open, such as **Pivot Table Field List**, you can close them by left-clicking on the x in the upper right.

17. Because this cross-tabulation (pivot table) was created in Excel, Excel is a bit sensitive about the graphs made from the data. To trick Excel into letting us do what we want, we will copy the table: **left click on cell V8 and drag through cell T5** => **right click on the shaded area** => **Copy** => **right click on cell T12** => **Home** tab => **Clipboard** box => **Paste** => **Paste Values** (first on left). You should now have the exact same information from the Pivot Table listed in cells T12 through V15.

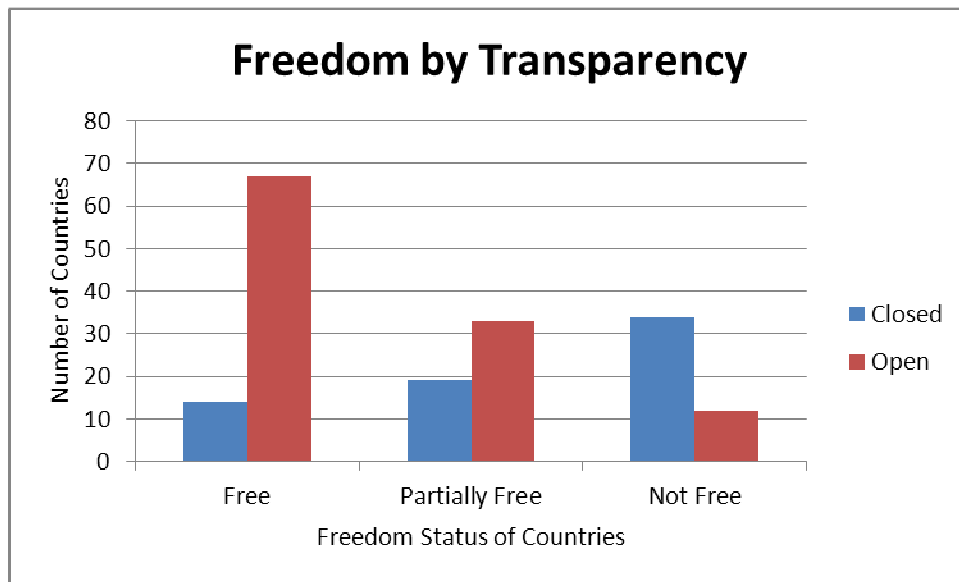
18. Before making the histogram, we want to change the order of the freedom level, which Excel organizes alphabetically: **left click on cell T14 drag through cell V14** => **right click on the shaded area** => **Cut** => **right click on cell T16** => **Insert Cut Cells**. This moves the Not Free row to be below the Partially Free row, and it moves the Partially Free row up one so there are no space rows. (You could have also completed this task with a couple copy and pastes.)

19. Now we are ready to make a histogram from the cross-tab data: **left click on cell T12 and drag through cell V15** => **Insert** tab => **Chart** box => **Column** => **2-D Column** (left icon in the top row).

20. Your graph will appear somewhere.

1. Click on it and drag it so that the top left corner begins at cell T18.
2. For this histogram, because it is plotting information from a cross-tab (i.e., information on freedom based on transparency), we will want the graph to have a title, y-axis label, x-axis label, and a legend:
 1. Under **Chart Tools** => **Layout** => **Labels** box => **Chart Title** => **Above Chart** (third from top). Enter a title of **Freedom by Transparency** (in bold).
 2. Under **Chart Tools** => **Layout** => **Labels** box => **Axis Titles** => **Primary Vertical Axis Title** => **Rotated Title** (second from top). Enter a y-axis label of **Number of Countries** (not in bold).
 3. Under **Chart Tools** => **Layout** => **Labels** box => **Axis Titles** => **Primary Horizontal Axis Title** => **Title Below Axis** (second from top). Enter a x-axis label of **Freedom Status of Countries** (not in bold).
 4. There is most likely a legend box in your histogram right now with the classes of **Open** and **Closed** defined. If so, you are good to go. If not, edit the current legend or include a legend: **Chart Tools** => **Layout** => **Labels** box => **Legend** => **Show Legend At Right** (second from top).

3. Your histogram should now look like the following:



The above graph shows the (somewhat intuitive) negative relationship between a country's transparency with the international community and the level of freedom it offers its citizens. In particular, there is a clear negative relationship between these two country measures: the two bars on the far left show that most free countries are also open (i.e., transparent) with the international community while the two bars on the far right show that most not free countries are also closed (i.e., not transparent) to the international community.

Exercises

1. Copy the data from the **Original Data** worksheet, and paste it into the **Exercises** worksheet.
2. In (roughly) cells G1 through I9, create a frequency distribution of the region in which each country is located. Include in the table the count of countries and a percent distribution.
3. Create a histogram of the region counts. Put **Number of Countries** on the y-axis. Make the title of the graph:

Countries by Region
(Your Name)

Print this graph.

4. Create a second histogram of the region percentages. Put **Percent of Countries** on the y-axis. Make the title of the graph:

Countries by Region (%)
(Your Name)

Print this graph.

5. Make a cross tab (roughly in cells P3 through S11) that counts the number of countries by **Transparency** (columns) and **Region** (rows).
6. Save this file as YourName_Lab2.xlsx.

Turning in your work

Email YourName_Lab2.xlsx to your professor as a file attachment to a single email with the subject heading Excel Lab 2: Your Name. In class turn in your hard copies of the two graphs you printed from the **Exercises** section.

Codebook for 2002 International Transparency Data.xlsx.

This data set contains information on 179 countries in 2002.

Country (Column A) lists each country's name.

Region (Column B) lists for each country one of six global regions in which the country is located – Asia, Europe, North Africa & the Middle East, North America, South America, and Sub Saharan Africa.

Transparency (Column C) associates with each country a label of either "open" or "closed." To be open, the country must have filed reports in 2002 to the United Nations on both its exports and imports of conventional arms as stipulated by the United Nation's Register of Conventional Arms. Failure to file either report is associated with the closed label. (You can learn more about the register of conventional arms at www.un.org/Depts/ddar/Register/Register.html.)

Freedom (Column D) is Freedom Houses classification of the country's freedom level based on the extent of personal freedoms and civil liberties offered to the residents of the country. (You can learn more about Freedom House at www.freedomhouse.org.)

Per Capita GDP (Column E) is each country's per capita GDP as measured by the United Nations in terms of U.S. dollars.