**Excel Examples and Useful Tricks**

**Things to keep in mind**

* Google is your friend! If you have a question, someone has probably answered it. Try this first.
* If you get stuck, try to understand the logic of an example (for instance, where are the parentheses and commas placed in an IF statement? Does order matter?) and apply the same rules to the task you are trying to do.
* If you are uncertain of what to do, you don’t have to do everything in Excel: to make things less challenging for yourself, try and identify what you have to do within the software and what you might be able to calculate by hand.

**Steps**

* Clean and transform data
* Calculate quantities and summaries (e.g. crosstabs)
* Visualize and summaries

**Cleaning/transforming the data**

1. To make this easier in Excel, you probably want to delete the second and third row of the dataset. You can also copy and paste columns of data into a new sheet for ease. For this exercise, copy columns “GPA” and “social gatherings”
2. *Missing values:* Remove the rows with NAs / empty cells
   1. Choose “Edit” -> “Go to” -> “Special” (You may need to do “Edit” -> “Find” -> “Go to” -> “Special”)
   2. Select “Blanks”
   3. Excel will select the blank cells
   4. On the home tab choose “Delete” -> “Delete sheet rows”
3. *Recoding/creating new variables:* We know from the questionnaire we designed that the variable named “GPA” refers to the question: “How satisfied are you with you GPA at Harvard?”, with responses falling on a Likert scale. For our purposes, we want to simplify this variable to identify responses as generally satisfied or generally not satisfied.
   1. To do so, let’s make a new variable. Write “GPA\_satisfaction” on the same row as the other variable names to create a new column
   2. Use the “if statement” to create values under this column. Here’s how you do that:
      1. In the cell in row 2 in the same column write “=IF(OR(B2="Very Satisfied", B2="Somewhat Satisfied"), "Satisfied", IF(OR(B2="Very Dissatisfied", B2="Somewhat Dissatisfied"), "Unsatisfied", "Neutral"))” where S2is the first cell in the regime type column (check if S is also GPA in your spreadsheet and adjust accordingly).
      2. Let’s also calculate a separate variable for each category. For satisfied, called the variable “GPA\_satisfied”. Following the steps in (i), the IF statement would be: “=IF(OR(B2="Very Satisfied", B2="Somewhat Satisfied"), 1,0)”. This will assign a value of 1 for respondents who are satisfied with their GPA and 0 to those who are neutral or unsatisfied.

*TIP: You can use if statements like this for numbers, not just characters, for example “99” etc.*

* 1. Apply the formula to the rest of the column dragging the bottom right corner of the cell downward

*TIP: You can do this faster by double-clicking on the bottom right hand corner of the cell (i.e. on the box that you would otherwise drag). Make sure, though, that the formula didn’t go past the last row of the dataset*

**Creating summaries**

We can start by calculating proportions:

1. Select any new blank cell somewhere on your sheet and enter formula =AVERAGE(D:D) where D is the “GPA\_satisfied” column with values 0 and 1. This will calculate the proportion of people overall who are satisfied with their GPA.

We can also create conditional proportions:

1. Select any new blank cell somewhere on your sheet and enter formula =AVERAGEIF(A:A, "Always",D:D) where A is the “social gathering” column and D is thee “GPA\_satisfied” column with values 0 and 1. This will calculate the proportion of people who are satisfied with their GPA among those who always attend social gatherings.

We can also create cross-tabs from categories. Cross-tabs group variables to illustrate correlations, patterns and trends in data. First, for purposes of cross-tabulation, your data has to be a table, so highlight the data (e.g. “ctrl + A”), and click on “tables” menu on the top and select “new table” to convert it. (You may instead need to go to “Insert” then “Table.”)

1. Select the columns in your worksheet
2. From the Tables tab, select “Summarize with Pivot Table.”
3. When a new window pops up, click “OK”. The PivotTable will be placed in a new worksheet.
4. Where it says “PivotTable Fields” or “PitvotTable Builder” on the righthand side of the worksheet, tick the variables you’re interested in under “Field Name.” In this case, let’s choose “social gathering” and “GPA\_satisfaction”.
5. “social gathering” should be displayed in the rows and “GPA\_satisfaction” in values. Click and drag accordingly.
6. In the value field, it will then say “Count of GPA\_satisfaction”. This will tell you the number of people in each cell.
7. We can also calculate these counts as percentages. Click on the “i” again, select the “Show Value As” tab, then select “% of Grand Total” from the first drop down list and click Okay.