Excel Lab, April 11, 2023

Intro

1. Our current dataset can be found **here**. (It was downloaded from the Webtools survey in csv format. Imported into Excel using Unicode (UTF-8) format, and saved as an Excel workbook.)

2. Spend a few minutes exploring the raw data. For reference, our survey can be found **here**. The survey numbers correspond to the spreadsheet column headings.

3. While not required, I recommend freezing the first row of the Excel Spreadsheet. The top row labels serve as landmarks to aid in remembering what is in the columns.
   i. On the **View tab**, click **Freeze Top Row**.

4. I know many of you are familiar with Excel functions. A recap of some of the basic functions
   i. Summation of a column:
      a. Let's look at question 4: "To your knowledge, does UIUC have a data privacy policy in place?"
      b. Let's count how many people think there is a privacy policy in place, but have not read it (Column R, labeled Q4:Yes, but I haven't read it).
      c. A few rows underneath the end of Column R, type **=SUM(R2:R68)**
         This adds all the elements in the column.
      d. This sum tells us how many people believe there is a policy in place, but have not read it.
      e. To experiment, sum the column to the left of it, Column Q by entering:
         =SUM(Q2:Q68).
      f. Now, sum columns S and T.
      g. Let's graph this.
         1. First, lets add labels for each of these sums above each value. (I used Yes, and read it, Yes, haven't read, No, Unsure)
         2. Drag a rectangle across the numbers and their corresponding labels.
         3. Go to the **Insert tab**. Click on the bar chart or **Recommended Charts**.
         4. Enter a title. I used "To your knowledge, does UIUC have a data privacy policy in place?"
         5. Now try creating sums of the gender and geography columns to create new visualizations. Experiment with bar charts, pie charts, and more.

            (Note: This approach for creating visualizations is great for your presentations.)

   ii. We can also sum values with the COUNTIF function.
      a. Pick a cell under Column R and enter **=COUNTIF(R2:R68, 1)**

   iii. IF Statement: Takes 3 arguments: IF (A, B, C)
      a. IF(A=1,B ,C) says IF(A = 1 (or Yes/True), then return B, otherwise return C).
      b. Let's test this. Find a cell under the end of Column EW.
      c. Enter **=IF(EW2, "Rural", "no")**
      d. This will return the word "Rural" if the participant in that row selected it. And will return the word "no" if they did not.

iv. Excel can also help you with word counts. For example, if we want to see in Column Z how many people think UIUC might collect location data, we can count the number of times the word "location" appears in every row. One way to do this is by entering:
=SUM(LEN(Z2:Z68)-LEN(SUBSTITUTE(LOWER(Z2:Z68),"location","")))/ LEN("location")

(Using formulas like this is not the norm. It is here in case it comes in handy in text exploration.)

5. You'll notice that each survey multiple choice option has its own column. For legibility, we will aggregate those into one column.
   i. Let's create a column that contains the gender entered by each survey participant.
   ii. Go to column ES. Click at the top of it, so that the entire column is selected.
   iii. Right click, and select **Insert**.
   iv. Label the top of the column **Q29: Gender**.
   v. Click on the cell below this new label and enter:
   =IFERROR(IF(EQ2, ER2, INDEX($EM$1:$EP$1,1,MATCH(1, EM2:EP2,0))), "N/A")

   (This column will be helpful later when we explore the data with Pivot Tables.)

Pivot Tables

1. **Pivot Tables** can help you explore your data.

2. We can replicate some of the above operations using Pivot Tables.

3. Let's go to Question 4 (Q4) in Columns Q-T.

4. Select a rectangle with the diagonal corners Q1 to T68.

5. Go to the **Insert tab**. Select **PivotTable**.
   i. In the box that appears, keep the range as is. Make sure **New worksheet** is selected. And click **OK**.
   ii. Select all of the field names that appear in the PivotTable Fields window that popped up on the right.
   iii. On the bottom right of that window, you will see a space with a summation label (Σ). This will add entries you put in this rectangle. Move the four Q4 fields here.
   iv. Click on **PivotChart** in the upper right. And your chart appears. You can rearrange the the order of the bars by rearranging the order of the values in your rectangle.
   v. To add a title, click on the **Design tab** -> **Add Chart Element** -> **Chart Title** and select where you want to place your title.

6. Spend some time exploring this data with Pivot Tables.

7. Now let's explore some characteristics of our survey participants.
   i. Create a selection rectangle to select all of the survey data, including labels.
   ii. Go to the **Insert tab**. Select **PivotTable,** and create it in a new worksheet.
   iii. In the new worksheet, go to the **PivotTable Fields** window.

iv. From the entries in the PivotChart Fields, select: Q29:Female, Q29:Male, Q29:Non-Binary, Q29:Prefer not to disclose, Q29:Prefer to self-describe
v. Move the entries into the rectangles as in the image to the right (Figure 1).
vi. Explore the different chart options.

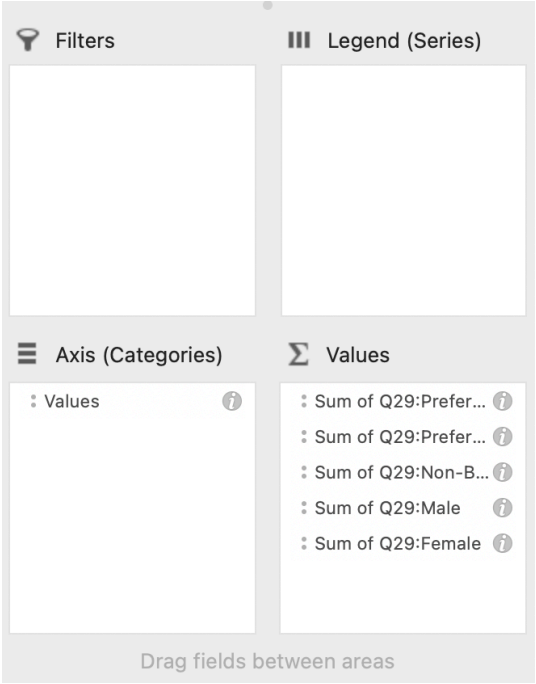

Figure 1

8. Let's also look at data data with some survey participant characteristics.
   i. Create a selection rectangle from cell B1 to the bottom right hand edge of the data in your spreadsheet. (Do not include the extra cells we added at the bottom to explore functions and formulas).
   ii. Go to the **Insert tab**. Select **PivotTable,** and create it in a new worksheet.
   iii. In the new worksheet, go to the **PivotTable Fields** window.
   iv. In the **FIELD NAME** window, select all the entries for Q1.
   v. Also select the Q29 aggregated column we created in step 5 of the *Intro* section of this document.. Match the entries with the image to the right (Figure 2).
   vi. Explore the different chart options. And explore the changes you see as you drag the fields across different areas in the **FIELD NAME** window.

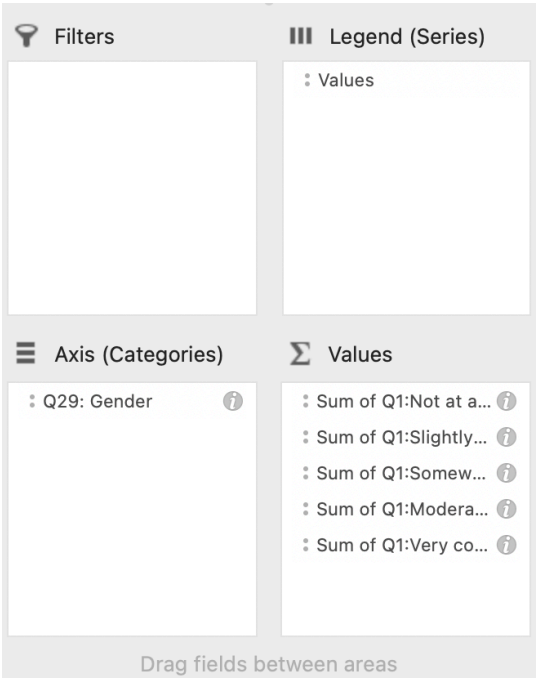

Figure 2

9. You can find these examples **here**.

Other Visualization Resources

1. Many sites that create word clouds (e.g., https://www.wordclouds.com/)
   i. Note that if data is sensitive, you may not want to put it on external servers.
2. Tableau. As UIUC students, you have free Tableau access.
3. Some visualization resources from the UIUC library here.

We showed you a spectrum of techniques to explore, arrange, and present data in Excel. Your goal is not to use all of them, but to let your story guide you. Explore the data for patterns, insights and inconsistencies.