Digital Communications Spreadsheet Assignment

Scenario: You're writing an article on employment trends in New Orleans. You need to download some data, understand it, and manipulate it. You'll be using Google Drive for this assignment.

Intro

Hi, I'm Elizabeth Kelly and I am your librarian liaison for Mass Communication. So I am here to help you with any of your research needs or instructional technology needs for this class, or for the other courses that you're taking in Mass Communication. Today what I'm going to be working with you on is an assignment where you work with data. Being able to work with data is such an important skill now in so many different industries, and the reason I first started working with Digital Communication is because the school of Mass Comm did a survey of recent graduates a couple of years ago and asked what they wished they had learned more about when they were here at Loyola. One of the top answers was they wished they knew more about working with data. So what I'm going to do is walk you through an assignment where we find some open data and we clean it up a little and we analyze it and then we create a quick data visualization of it at the end. These are all skills that you could use whether you're interested in going into advertising, or PR, or journalism, so hopefully by the end you will at least feel like you're in a starting place to be able to do some analysis of data. The data set we're going to be working with today comes from the city of New Orleans open data portal. This is something a lot of cities are doing now, where they are publicly releasing the data that they're collecting related to all sorts of aspects of city governance. That can include where the traffic cameras are in town, what companies have applied for new building permits, or even crime data. So there's a lot that we can find from this Open Data Portal and then we can use that data--it's openly licensed, so if we wanted to publish this data in a paper that we were writing or in a newspaper article we would be free to do so. For this assignment we are going to use Google sheets, but everything that we do today could also be done using Microsoft Excel or OpenOffice Calc. The functionality of most spreadsheet software is very similar. There may be some slight changes and differences. The menus may have slightly different titles. But overall once you know how to use one of these programs you'll pretty easily be able to move and use another one. And another advantage is that Google Sheets lets you save your file as a Google Sheet so you can keep working on it in the cloud, you can collaborate with other people, you can share it. But you can also download it as an Excel file and then open it on a desktop using Excel; you can save it as something called a CSV file, a comma separated value file, as well as some other different types of spreadsheet software which then can be opened in other spreadsheet programs or even other data programs that don't use spreadsheets but prefer other types of files. So we'll use Google Sheets for today but just know that hopefully these skills are all translatable to other programs that you might find yourself using.

Setting Up
1. Go to the guide at http://researchguides.loyno.edu/cmmna201 and open "Industry Employment, 2001-2016"
2. Go to File--Make A Copy and copy the dataset to your Google Drive

This data set was downloaded from
https://daisi.datacenterresearch.org/Workforce-Development/Industry-Employment/6aj2-892n

3. Read through the article "How To 'Interview' A Big Pile of Data" (also linked on the Research Guide) from NPR. Then, using the Data Center website and the Google Sheet, try and answer the following questions for yourself:
   a. Who created this dataset?
   b. Who is responsible for maintaining it?

      c.   What point in time or time-range does it apply to?
      d.   How was the data collected?
      e.   How was the data processed after being collected?

Regardless of if you've developed a dataset yourself to work with, someone gave you data to work with, or if, like today, you've downloaded publicly available data, you always want to know how many other hands have "touched" the data before you have, and what exactly is in the data. That way if you're presenting your findings, or writing a paper, or publishing an article on them, that you're representing the data accurately.

Cleaning and viewing

Sometimes your data is messy. This dataset is pretty clean, but in general, we want to delete any extra data and tell Google Sheets how to read certain data.

4. Delete any data that isn't clearly attached to a data point. Today, our data set doesn't have any data that needs to be deleted.
5. Notice that there are letters going across the top of your sheet, and numbers going down the left. The letters are for column names, and the numbers are for your rows. These letters and numbers are important once we start using formulas. Each of the boxes on your sheet is a cell, and each cell has a name based on its column letter and row number. We'll discuss this more later.
6. Next, freeze your top row so you can see what column you're working with even when you're down at the bottom of your data set by going to View--Freeze--1 Row. Now when you scroll up and down, your top row should stay in place.

Sort and filter

Sometimes you want to put your data in a specific order.

7. To sort, highlight the column you want to sort by, then go to Data—Sort Sheet By Column and select either A-Z or Z-A. This will sort the entire sheet based on the column you selected. The other option under Data—Sort is Sort Range by Column. This will only sort the data in that one column and will mess up our spreadsheet. You will almost always want to use Sort Sheet by Column. Undo when you're finished to return to the spreadsheet's original organization.
8. To filter data so you're only looking at some of it at a time, highlight the column you want to filter by, select the Filter icon on the toolbar, click the hamburger menu at the top of your column, and choose which categories you want to look at. Once you have a filter on a column, look at the row numbers on the left of your screen; they may no longer be in consecutive order. Don't worry, your data hasn't disappeared, the filter is just hiding some of it. To get rid of your filter, click the filter icon again. Your row numbers should all be in consecutive order now.

Formulas

Google Sheets will do math for you!

9. Let's find out the average number of people employed in the Performing Arts from 2001-2016.
10. Filter Column C to only include Performing Arts.
11. Create a new sheet in your workbook by clicking the "+" sign at the bottom left of your screen. Right click (PC) or double-click (Mac) the tab for your new sheet on the bottom of your screen and rename it "Performing Arts."
12. Highlight the entire table of Performing Arts data from your original and copy over to a new sheet.

      a.   To highlight: Ctrl+A (PC) or Command+A (Mac)

      b.   To copy: Ctrl+C (PC) or Command+C (Mac)

      c.   To paste: Ctrl+v (PC) or Command+v (Mac)

      d.   OR, in Google Chrome, use the Edit menu to Copy and Paste

13. On your new sheet, highlight E2 through E17, select the Function button and select "Average" from the dropdown menu. You should now see the formula =AVERAGE(E2:E17) in E18. This means that we are calculating the average of the all the data in cells E2 THROUGH E17 (that's what the colon in the formula means). Label your Total by typing "Performing Arts Average" in D18.

      a.   Note that there are many other functions you can use, including Sum and CountNumbers.

Pivot Tables

Another way to calculate sums, averages, and more in Google Sheets is by creating a pivot table. Pivot tables let you combine data and move your categories around without having to run a bunch of formulas manually. Specifically, they let you count, sum, and calculate based on the categories in your dataset. In this case, let's look at the number of workers employed for each industry.

14. Go back to your original sheet and remove the filter. Highlight all of your data and go to Data—Pivot Table. Select "New Sheet" as the location and click Create.

15. Your pivot table will be blank to start. We will need to tell it which parts of our data to look at.

16. Click "Add" next to Rows and select "Industry."

17. Click "Add" next to Values and select "Number Employed." Make sure the "Summarize By" box reads "Sum;" if it says something else (like "Count,") click the drop-down menu and select "Sum."

18. We want to do more with this info, so copy and paste all of your Pivot Table data onto a new sheet (use "paste as values only" from the Edit menu, or Ctrl+Shift+V on PC and Ctrl+Command+V on Mac). Do not copy Blanks or Grand Totals. If you do have a blank row or a row with the Grant Total in it, click the row number on the left and delete it by going to Edit--Delete Row, or right-clicking (PC) and selecting Delete Row.

19. Add a column header in column C that says "Percent of total."

20. In cell C2, create a formula that will add all of the average salaries together and see what percentage of that the entry in your row is. In this case, the formula will be =B2/SUM($B$2:$B$21) (because Google Sheets is taking the sum of B2 through B21 and dividing that into the data in B2). Click enter once you've entered the formula.

21. Click and drag the box at the bottom right corner of C2 all the way down to C21 to make the same calculation for each row. Note that the formula changes for each row, except for the parts of the formula with a $ in front of them. Because we always want the average salaries to be calculated using B2:B21, we need to use the $ to keep that part of the formula the same. The rest of the formula changes depending on which row we're on.

22. Change Column C into a percentage by clicking the Percent (%) symbol on the toolbar.

Visualize

Finally, let's create a data visualization about how employment in each industry has changed over time. Visualizations can help readers and viewers digest your data a little easier.

23. Go back to your Pivot Table sheet. Leave the Rows and Values boxes the same, but add "Year" to the Columns box.

24. Copy and paste all of your Pivot Table data onto a new sheet (paste as values only). Do not copy Blank rows or Grand Totals. If you did, delete them.

25. Highlight the entire table you just created. Select "Insert" and "Charts."

26. Under "Chart Type," change to Line Chart.
27. At the bottom of the chart editor, click all three check boxes (Switch rows/columns, Use Column A as Headers, and Use row 1 as labels)
28. Click the Customize tab on the Chart Editor to make further customizations.
29. Once you like the looks of your visualization, take a few minutes to think about it. What questions does it raise for you about how employment in different industries has changed over time in New Orleans?
30. Finally, click the three dot menu at the top right of your chart to see some options for saving and exporting your chart. You can export it as an image file to include in presentations and papers, or embed in other Google products like Google Slides.