TLT Google Sheets Directions

Setting up

1. Sign in to your Google Account.
2. Go to the dataset at https://goo.gl/ooSkqj (or https://tinyurl.com/loynobaroqueart)
3. Go to "File," "Make a copy" to add the folder to your personal Drive so you can edit it
4. Open "baroqueart.dump.20140122.csv" from your copy in Drive

Understanding your data

This dataset was downloaded from http://baroqueart.cultureplex.ca/

1. Using the website and the Google Sheet, try and answer the following questions:
    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?

Best practices

You have already made a copy of your dataset; this is a good practice in general in case you accidentally delete or move something. Some other good practices are:

1. Use **Data—Protected Sheets and Ranges…** to prevent collaborators from editing individuals cells, rows, columns, or entire sheets
2. Use formatting (bold, italics, colors, etc) sparingly. If you save your spreadsheet as a CSV, Tab-delimited text file, or other format, these changes will be lost. You can, however, use **Format—Conditional Formatting…** to add visual cues to your sheet. You can even sort by color using a script or an add-on.
3. Avoid using multiple tables within one spreadsheet.
4. Avoid spreading data across multiple tabs (these will again get lost if you change formats).
5. DO use new tabs to record data cleaning, manipulations, graphs, tables, etc.
6. Include only one piece of information in a cell.
7. Avoid special characters
8. Record metadata and actions taken in a separate txt file

Cleaning and viewing

Sometimes your data is messy. Cleaning it up before you start analyzing it will save you a lot of time.

1. Delete any data that isn't clearly attached to a data point.
2. 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**.
3. Expand any columns where you can't view your data by double-clicking in between it and the column to its right, or by manually resizing
4. To sort, highlight the column you want to sort by. Go to **Data—Sort Sheet**. Using "Sort Sheet" makes sure all of your rows stay together; if you use "Sort Range," only that column will sort and your data will no longer be valid. ***How could you sort this sheet?***

5. You can also filter your data to only see certain categories or entries in a single column. This will also help you identify where you might need to clean some of your data. For example, create a filter for **Column F current_place** by clicking the filter icon on the toolbar, and search for New York. ***Can you combine or rename any current_places?***
6. Use **Data—Data validation…** to make sure the correct information is in a row or column. You can specify a list certain values that should be used (named, places, etc) or that the values use certain criteria (such as a column that should only contain 9-digit numbers). For example, highlight **Column C creation_year_end** and select **Data—Data validation…**Enter the following under Criteria: **Number** -- **between** -- **1550** and **1850** to see if all the paintings were actually completed between 1550 and 1850. ***What did you find?***
7. Use formulas to translate foreign languages! Use the formula =GOOGLETRANSLATE("cell") to automatically translate text. Try it for cell A2, "La defensa de la eucaristía con santa Rosa de Lima" (hint: the formula will be **=GOOGLETRANSLATE(A2)**). To repeat the formula for each row, just click and drag down the little blue box in the right bottom corner of the cell (or double-click it to fill out the entire column). Look at how the formula has changed for each row.
8. If you want to split the data in a column into multiple columns, use **Data—Split Text to Columns**, and select where you want to split the data (using punctuation or something else). Note that this will create new columns and overwrite any columns in the way, so you may want to copy the data to a column all the way on the right of your spreadsheet before trying. For example, try splitting the data in **Column F current_place**, using the "." as a delimiter (and remember that "undo" is your friend).

Pivot Tables

Pivot tables let you combine data and move your categories around without having to run formulas manually.

1. Highlight the entire spreadsheet. Click "Data" and then "Pivot table..."
    a. Your pivot table will open in a new sheet. The table will probably be blank.
2. First, let's look at which museum currently has the most Hispanic Baroque paintings (**Column F current_place** on our original sheet)
3. Next to "**Row**," click "**Add field**"
    a. Select "**current_place**"
4. Next to "**Value**," click "**Add field**"
    a. Select "**current_place**"
    b. Next to the "**Summarize by**" drop-down menu, select "**COUNTA**"
5. You should now have a table listing how many Hispanic Baroque paintings each museum currently has. ***Do you see any data that needs more cleaning?***
6. This table is very large. You can filter within a Pivot Table. To only look at the number of paintings per museum in the United States, highlight **Column A current_place** and select the **Filter** icon. Click the filter icon now on **Column A current_place.** Select **Filter by condition…** and **Text contains**. Enter "United States" and click **OK**. Notice how your Pivot Table changes.
7. To view the table in descending order, copy it to a new sheet by right-clicking your mouse, selecting "Paste special" and "Paste values only." This will paste your table data without recreating your Pivot table. You can also paste as values only by selecting Ctrl/Shift/V on your keyboard (PC) or Command/Shift/V (Mac). Do not include blank rows or "Grand Total." Make sure to paste your table at the top-most left-most cell on your sheet.

8. Right-click the sheet's tab and rename it **Tables**.
9. Sort by the **COUNTA of current_place** column on your table

Graphs and Charts

It's finally time to visualize our data! For this chart, we want to show which museums in the United States have the most Hispanic Baroque paintings.

1. Our table is too long to make a good data visualization. Let's only use museums that have at least 10 paintings. Select just the data you need and click the **Chart** icon on your toolbar. ***Are there any rows we should exclude from our chart?***
2. Google Sheets automatically recommends some appropriate chart types to you (or click the "Chart types" tab for more options). Spend a few minutes selecting different chart types. ***Which ones make the most sense for this type of data?***

Customization

Once you've settled on a graph type, you can customize it.

1. First, copy it over to a new sheet. Click the dropdown arrow on your graph and select "Move to own sheet."
2. On the new sheet, right click the sheet tab and rename it "Current US Museums."
3. Click the "Advanced Edit" button at the top to customize your pie chart's display text, color, legend, and more. ***What is a good title for this chart?***
4. Follow best practices by using the Data Visualization Checklist (https://tinyurl.com/datavizcheck)

Saving and Exporting

Finally, you'll want to reuse your graphs outside of your Google sheet.

1. Publish to the web
    a. To create a link to your chart: Click "Publish chart" from your "Current US Museums" sheet. In the first drop-down menu, select the name of your chart. In the second drop-down select whether you want to publish the sheet with your pie chart on it (which would be interactive) or a PDF.
    b. To embed your chart on a website: Click "Publish chart" from your "Current Museum" sheet. Click the "Embed" tab. Click "Publish." Copy the embed code and paste it into your website.
2. Save as a png
    a. Click "Save image."
    b. Save to your computer
    c. Use the image to insert your graph into text documents, presentation slides, and more
3. Copy (for Google products only)
    a. Click "Copy Chart"
    b. Paste chart into Google Docs, Sheets, Slides, etc.
    c. You will be given the option to "Link" or "Don't Link" to your original chart. Linking will automatically update your chart if you change it; however, if you then delete the original chart, it will also be deleted from its linked version.

Questions?

Explore on Your Own

1. Select another data source from the Data Sources box on the Research Guide at http://researchguides.loyno.edu/data_workshop
2. Spend some time on your own interviewing, cleaning, analyzing, and visualizing your data
3. Share:
    a. **What's something interesting you learned from your data?**
    b. **What's something you want to learn from your data but don't know how?**