Data Visualizations with Flourish

Created by Dr. Jen Andrella
Center for Digital Humanities, USM

CDH Hours: Monday-Friday 8am-4:30pm
Email: dh@usm.edu
For additional resources and tutorials, visit www.usmcdh.org

Overview

Data visualizations are a useful way of modeling and interpreting large quantities of information. Flourish is a powerful, free online software that utilizes quantitative information to create static and interactive charts, graphs, projection maps, diagrams, and other unique visualizations. Data visualizations are more than illustrations; use them to engage your audience, to think about the information in question, and to empower your sources through critical thinking and analysis.

This guide covers the basics of finding data, preparing the data for Flourish, and generating two distinctive visualizations– a simple graph and a projection map. All visualizations may be publicly published, downloaded, and embedded into other websites.

Flourish is free to use, but your data and visualizations will be publicly accessible. Please consider the sensitivity of your data and confirm the rights and permissions prior to using Flourish.

Where to Find Data

You can import different types of tabular data (structured information organized into a table with clear labels) into Flourish. This may be census information, geographic information, or even a hand-created Excel spreadsheet. If you are not creating your own spreadsheet, you may look to the following resources for open and free datasets.

Dataset Resources:

- **Pew Research Center Datasets:** www.pewresearch.org/tools-and-resources
- **World Bank Open Data**: data.worldbank.org
- **USDA Economic Research Service** (county level data): www.ers.usda.gov/data-products/county-level-data-sets
- **IPUMS** US and international census and survey data: www.ipums.org
    a. NOTE: The **IPUMS NHGIS** option includes census + geographic boundary files that can be used to create a data map (see details in projection map section): www.nhgis.org
- **U.S. Census Bureau**: data.census.gov/cedsci
- **World Health Organization (WHO)**: www.who.int/data/gho
- **Google Dataset Search**: datasetsearch.research.google.com
    a. **Google Public Data Directory**: www.google.com/publicdata/directory#
- **Kaggle** (variety of datasets and formats)**:**www.kaggle.com/datasets
- **Data.gov** diverse topics of data in the United States: data.gov
- Data.census.gov, the main platform to access US Census Bureau data. The Decennial Census is a full count of the population every ten years, while the American Community Survey (ACS) is an annual sample count that produces one-year and five-year estimates for different census geographies.
- Global Open Data Index, by the Open Knowledge Foundation.
- Harvard Dataverse
- openAfrica, by Code for Africa.
- Open Data Inception, a map-oriented global directory.
- Open Data Network, a directory by Socrata, primarily of US state and municipal open data platforms.
- World Inequality Database, global data on income and wealth inequality.Sample Dataset for Tutorial:

Sample Dataset: 1840 Free/Enslaved Data by Age and Sex (auto-download CSV) (IPUMS)

Creating a Flourish Account

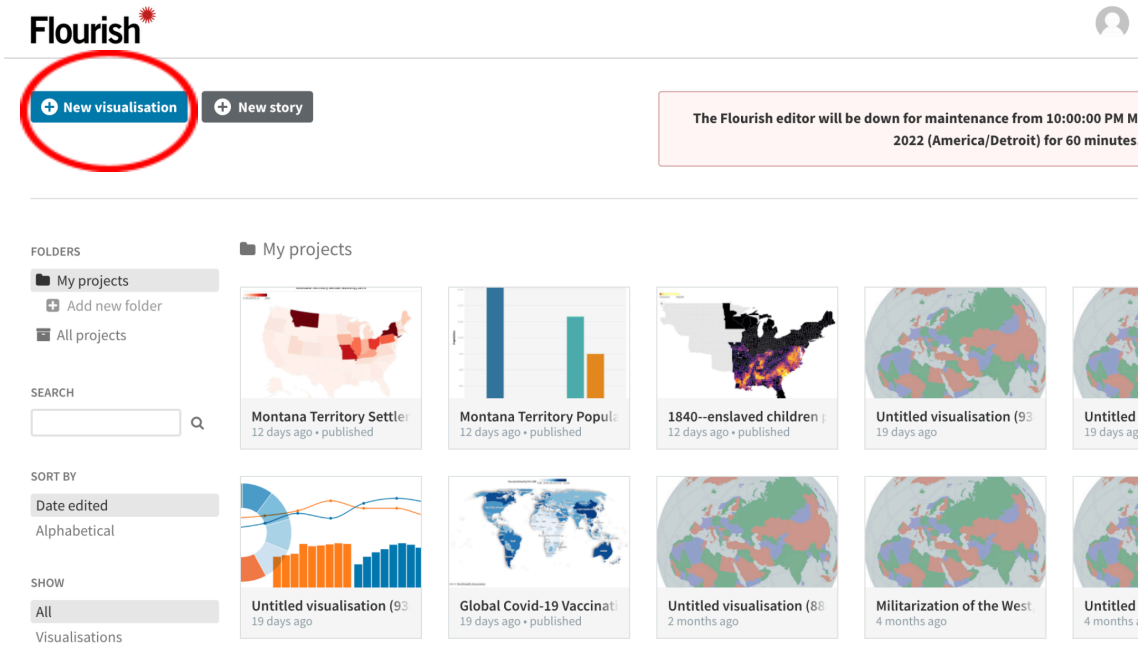
1. Go to: flourish.studio in a web browser.
2. Click on the **Get started for free** button in the center of the page.
3. Enter your preferred email, username, and password in the fields, or create an account using Google.
4. Click **Sign up,** and you will be logged in to the dashboard
5. After creating your account, you will be able to login in the future by clicking the **sign in** option in the top right corner of the screen.

Building a Basic Data Visualization

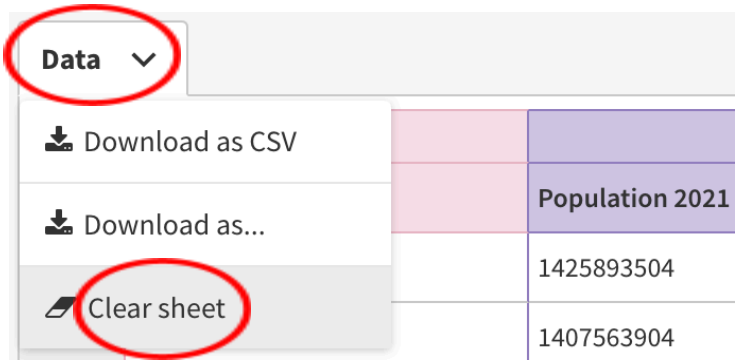
For this section, you may use the sample dataset linked above, or use your own data.

1. Review your data using Excel or Google Sheets and consider research questions. How might visualizing this information address those questions? It might help to develop a list of questions and note which aspects of the data you want to visualize and compare.

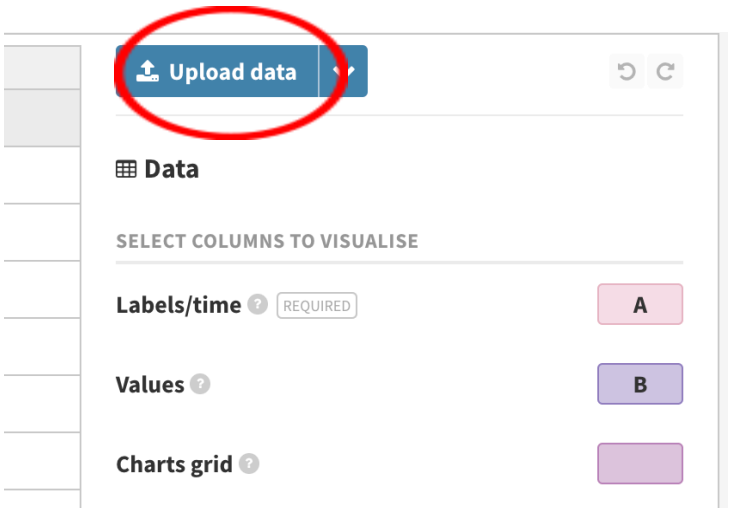
2. Select **New Visualization** in the upper left corner of the screen.



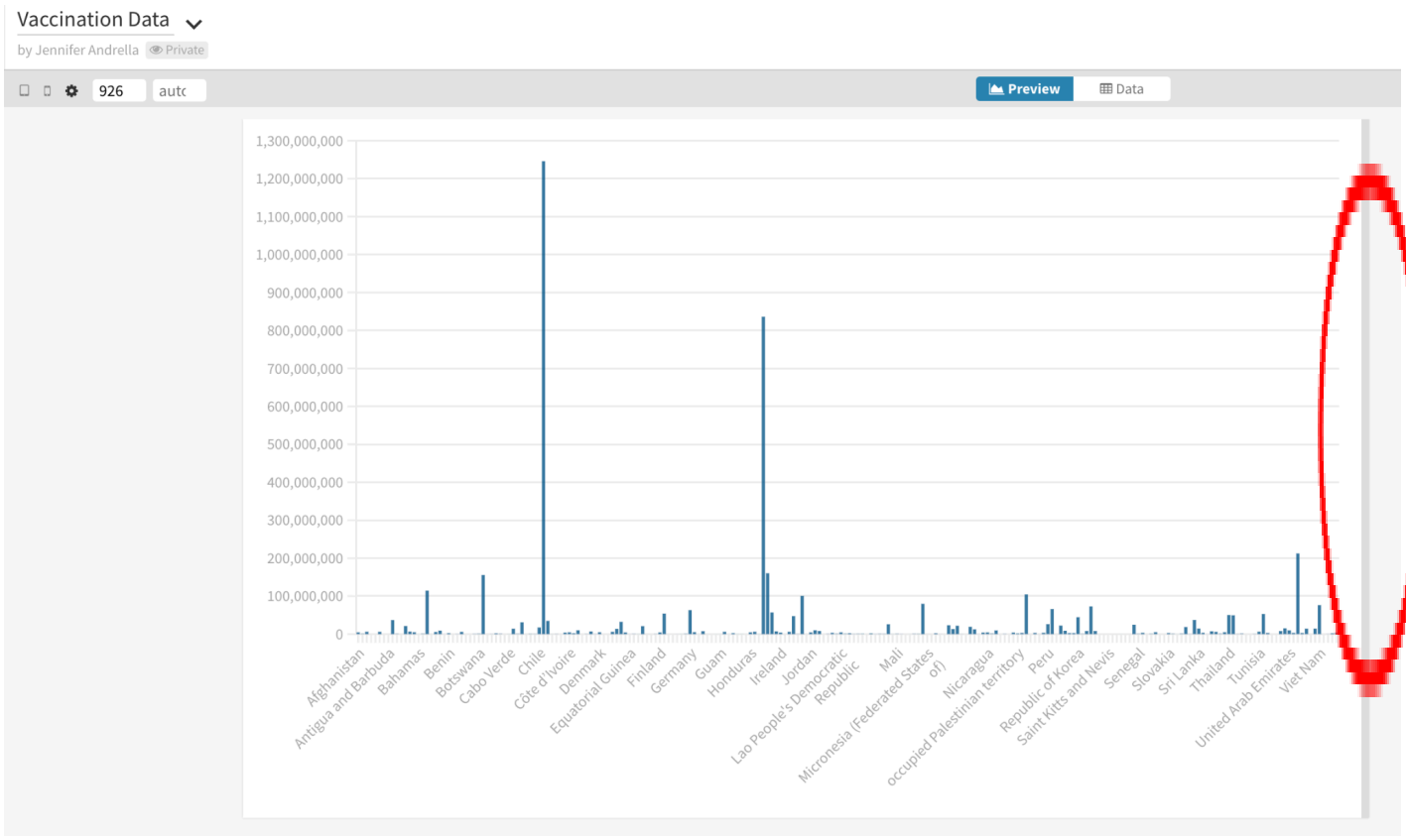
3. Choose a template that complements your data. For example, if you wanted to compare data across categories, a bar graph may be better than a pie chart. For this tutorial, we will use a standard **column chart.** Select this option from the first row of the section, “Line, Bar, and Pie Charts”
4. Give your column chart a **title** in the upper left corner. You will see a preview screen of a sample graph from Flourish.
5. In the center of the screen, select the **Data** tab next to Preview to view and change the data. Again, Flourish provided sample data that we can delete and replace with our own. Select the **drop-down menu** from the Data tab and select **Clear sheet.**



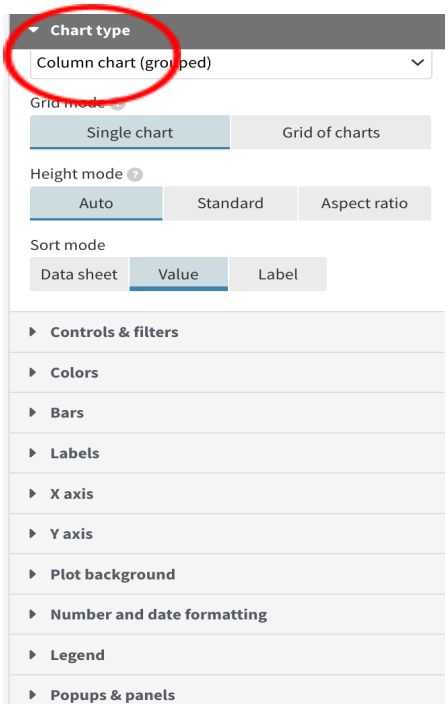
6. Click the blue **Upload data** button.
   Note: The drop down option allows you to merge multiple data sets without deleting the previous one.



7. Select the vaccination data file saved to your computer and **import** it into Flourish
8. Flourish will automatically select columns, but these are just a suggestion. Select a column to represent **Labels/time** (x-axis) and **Values** for those labels (y-axis). The additional fields are optional.
9. Keep **Labels/time** set to **Column A,** and set the **Values** to whichever column(s) you wish to visualize. For example, select **Column L, "Persons fully vaccinated."** You may also select multiple columns by separating them with a comma, or a range of columns by using a hyphen.
   Note: You may add other information in the **Popups** option. This is useful if you want to add other information for comparison or context, such as vaccine types.
10. Select the **Preview tab** to see your graph.
11. Drag the right edge of the graph to expand it and see more countries.



12. Explore the options in the **template menu** to change the appearance of your graph. You can also change the type of graph/chart using the dropdown menu at the top of the menu.



13. When you are ready to publish your visualization, select **Export and Publish** from the upper right corner. This will allow your visualization to be shared by link or by embedding. You may also **Download** the HTML or a still image.

    Note: Flourish **automatically saves** your work and you can always access it from the dashboard for future editing.

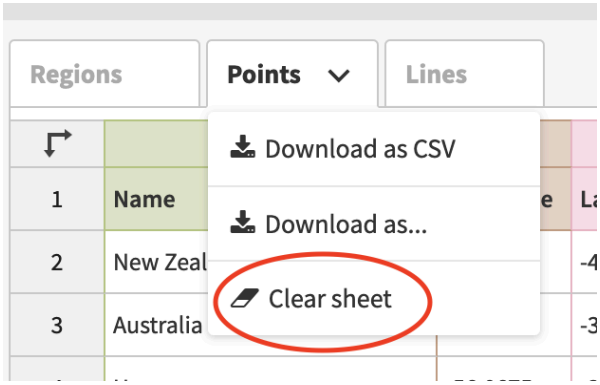
14. After selecting **Publish,** you will be given the options to allow public access to duplicate your data, you may leave this unchecked. You will then be able to **share** your visualization as a link or by using the embed code, if you want to embed it into a website.

    Note: the Link and Embed options will keep your visualization fully interactive for users.

Building a Data Map

Note: We will use the same sample dataset for building a data map.

1. Log in to Flourish and select **New Visualization** from the upper left corner.
2. Scroll down to the Projection Map section and select the **U.S.- County** map option. If using another dataset, select the geography that best matches your data.
3. Give your map a **title**.
4. Flourish provides sample data and a sample map projection. You can edit or delete these in the next step.
5. In the center of the page, select the **data tab** and you will see four sheets: **Regions, Regions Geometry, Points**, and **Lines**. Each of these sheets is pre-populated with its own sample data.
    a. If using the sample data, clear the **Regions, Regions Geometry, and Points** sheets by selecting the drop-down arrow and then **clear sheet.**



6. Under the **Regions** tab, select the **Upload data file** option from the right menu. Select the sample data file (or your own data file) from your computer to import. The **Regions** tab will now include all of your data.
7. **Repeat this step for the Regions Geometry** sheet.

Setting Values for the Data Map

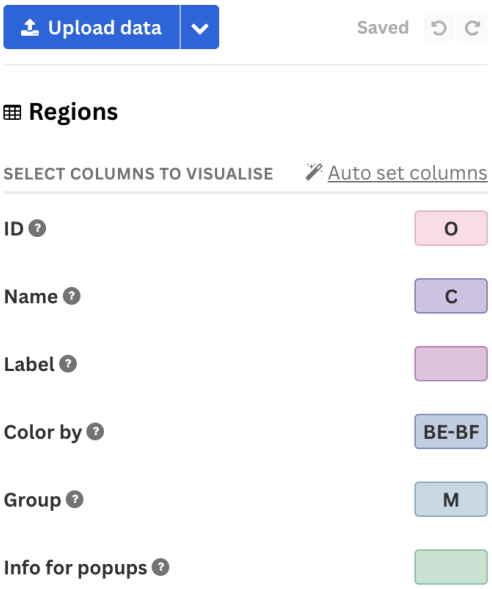
Regions:

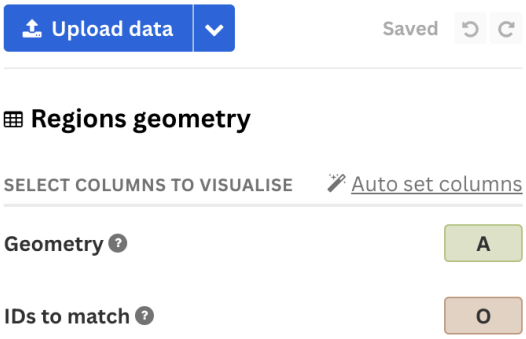
- **ID:** O (PID - County Code)
- **Name:** C (County Name)
- **Label:** C (optional– displays county names on map)

- **Color by:** whichever column(s) of data you want to represent. You can **add multiple columns separated with a comma** or a **range of columns separated by a hyphen**. This will create multiple views of your map to toggle in the preview tab.
- **Group:** M (optional - creates state lines)
- **Info for Popups:** optional, additional data to include in the popups

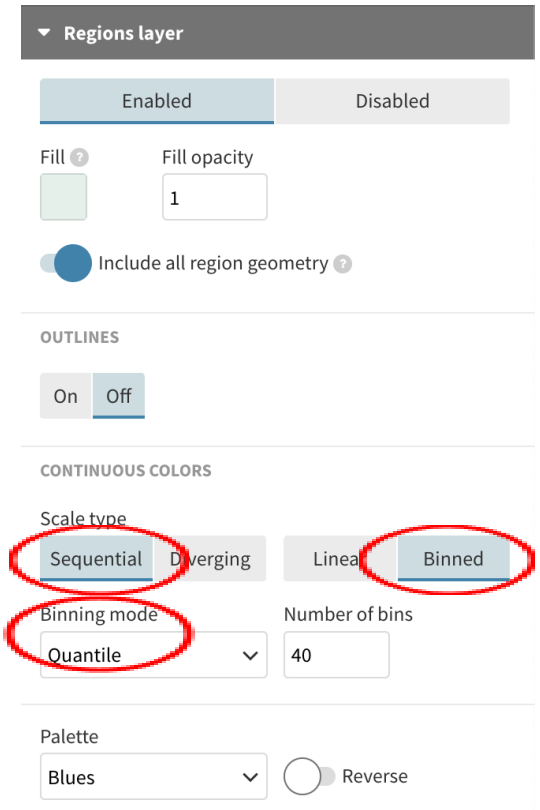
Regions Geometry

- **Geometry**: A (this is a GeoJSON, a special data format for geographic data)
- **IDs to Match**: O (PID for county – matches the ID in the Regions tab)





8. Go to the **Preview** tab to view your map. It should appear in various shades of blue.
9. In the menu on the right side of the screen, find the **Regions Layer** and scroll to the **Continuous Colors** section. Set the scale to **Binned**. Set the Binning Mode to **Quantile** and select your preferred **Palette color.** You can also increase the **Number of bins** to increase the color gradation across the map (at least 5 is recommended, but experiment with this number).



Note: If you are visualizing more than one column, navigate to the **Controls** settings in the right column where you can find options to create a slider, timed slider, toggle buttons, etc. A timed slider creates a looped animation with a play/pause button.

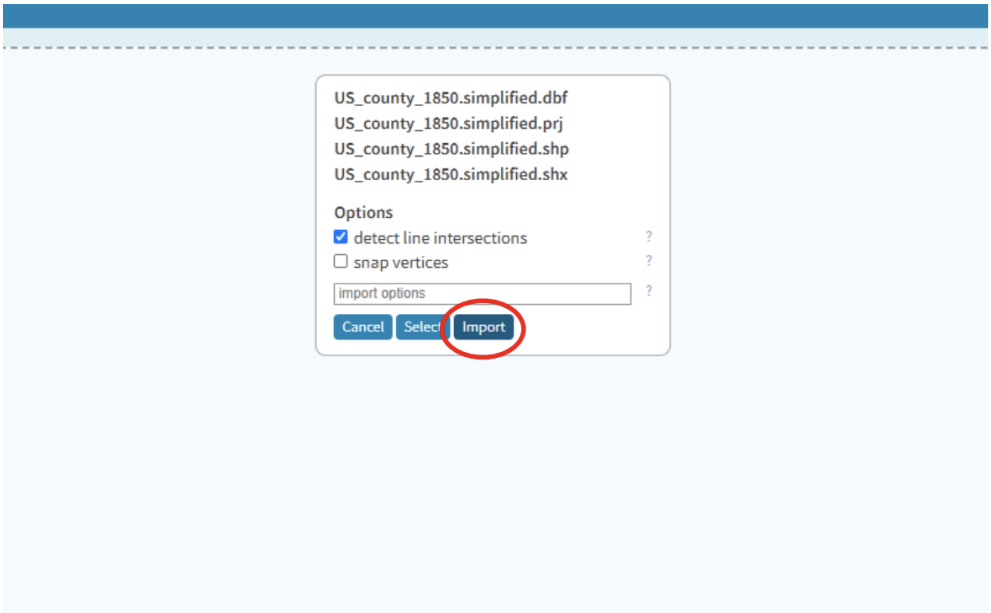
10. Flourish **automatically saves** your work and it can be accessed from the dashboard for future editing.
11. When you are ready to publish your visualization, select **Export and Publish** from the upper right corner. This will allow your visualization to be shared by link or by embedding. You may also **Download** the HTML or a still image.
12. After selecting **Publish,** you will be given the options to allow public access to duplicate your data, you may leave this unchecked. You will then be able to share your visualization as a link or by using the embed code, if you want to embed it into a website.

    Note: the Link and Embed options will keep your visualization fully interactive for users.

Optional: Creating your own map boundaries using Mapshaper

If you are using historical data, you will likely want to create your own map boundaries. To do this, you will need the **GIS Shapefiles** of those boundaries and convert them into a **GeoJson** file. The NHGIS database will offer a download option for shapefiles separate from the CSV census data download.The zipped folder should include **.dbf**, **.prj**, **.shp**, and **.shx** files.

1. After downloading the zipped folder of files, unzip/extract them on your computer.
2. Go to: https://mapshaper.org/
3. Select and import the **.dbf**, **.prj**, **.shp**, and **.shx** files into Mapshaper.



4. Mapshaper will then display your map. There may be red dots, showing uncompromised intersections. Select the X Button in the upper right corner to resolve those intersections.
5. Select **Export** in the upper right and choose the **GeoJSON** file option.
6. Upload the **GeoJSON file** into a **blank projection map** under the **Data** tab.
7. In addition to the map file, **upload and merge** your other data file. If this data is also from NHGIS (recommended), you will need to first open this data in Excel and delete the first row (this ensures a proper match with the map file).
8. When uploading the second file in Flourish, you will be prompted to combine the two files using a matching column name. Select **GISJOIN** from the dropdown menus, which appears in both data files.
9. Now your map should appear in the Preview tab, and you can select the columns you want to visualize on the map. **Return to step 28** in the data map instructions to adjust the Numeric Binning settings.