

Please complete each problem and submit a PDF with your solutions. If you use code to solve the problems, you should include it in your submission.

Problem 1 (10 points)

For Problem 1, you will need to refer to the dataset we used in Homework 2—storms.csv.

- A. A new file, names.csv, contains the "gender" of the hurricane name (by convention, the World Meteorological Society determines a list of storm names, alternating names traditionally used for males and names traditionally used for females every calendar year—see [Tropical Cyclone Naming History and Retired Names](#)). Combine storms.csv with the new file so that you have a dataframe with the name, "gender", and year of each storm. You should have 512 distinct storms in the new dataframe (remember, some names were used in more than one year). (1 pt)
 - B. With your new dataframe, create a stacked bar chart showing the number of storms in each year. The "stacked" categories should be the gender of the storm name. Use an appropriate color scheme for the gender categories based on the type of data you are displaying, and explain why you selected the scheme. Include a caption discussing your graph. (2 pts)
 - C. Using your original dataframe, create a scatterplot showing the average longitude for each storm on the x -axis and the average latitude for each storm on the y -axis. Use an appropriate color scheme for the average windspeed based on the type of data you are displaying, and explain why you selected the scheme. Include a caption discussing your graph. (2 pts)
 - D. Using the graph from Part C, highlight the hurricane you tracked in Homework 2 with a different color. Use an appropriate color scheme to highlight your hurricane, and explain why you selected that scheme. (1 pt)
 - E. Revisit your hurricane tracking plot from Part E in Homework 2, Problem 1 (longitude on x -axis, latitude on y -axis, circles indicating the radius of tropical storm and hurricane winds). Color the path of the hurricane according to the Saffir-Simpson scale using any color scheme from the set of `viridis` color schemes except `viridis` itself. Use an appropriate color scheme for the Saffir-Simpson scale based on the type of data you are displaying, and explain why you selected that scheme. Include a caption discussing your graph. (2 pts)
- Note:** Make sure that all tropical storm circles have a color consistent with category 0, and all tropical depression circles have a color consistent with category -1.
- F. At some point in your career as a data scientist, you may want to create and use a custom color palette. In your data visualization module/package of choice (`altair`, `ggplot2`, `matplotlib`, `plotnine`, or `seaborn`), create a custom six color palette based on something you enjoy—e.g., a movie, musical album, sports team, TV show, etc. For example, consider this [palette based on colors worn in the 2021 Presidential Inauguration](#). You may want to use a color picker such as <https://imagecolorpicker.com/en>. Apply the new color scheme to your graph in Part C. (2 pts)

Problem 2

Using a dataset of your choosing, create a distinct type of data visualization that you have not made before. This means you may not make a bar chart, histogram, boxplots, or scatterplot—consider instead types such as a directed network graph or choropleth map. Examine the data, find a fact that is contained within the data, and design a visualization that communicates that fact.

As usual, include a figure caption and just one paragraph discussing your findings and the graphical design. Attach any code you used to produce the visualization. The figure caption should describe the origin of the dataset.

You do not have to use any specific tools to produce the visualization (you could even draw it by hand) but you need to find something interesting and display it effectively.