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

Problem 1 (5 points)

Freedom House is an organization that tracks changes in political rights and civil liberties in the world. The Excel spreadsheet, All_data_FIW_2013-2022.xlsx, contains data for the survey over the last ten years.

A. Suppose we are interested in seeing how civil liberties have changed in the pandemic era (for our purposes, defined as 2019-2022). First, read in the data and calculate the differences in total civil liberties scores from 2019 to 2022 (1 pt).

B. Using the calculated differences, create a chloropleth map showing the change for every country/territory. Make sure to use a meaningful color scheme using guidelines from the last few weeks–give a brief description of the scheme you used and a justification for using it. According to your chloropleth map, what are the global trends in changes in civil liberties in the pandemic era? (3 pts)

C. Zoom into Europe, and change the Coordinate Reference System (CRS) of your graph. (1 pt)

Problem 2 (5 points)

The United States Census Bureau's American Community Survey replaced the long form census in 2000. It collects data describing the American population, housing, habits, etc. The csv file, ACSSPP1Y2021.S0201-Data.csv, contains data from the 2021 survey.

A. In class, we viewed a scatterplot describing the relationship between the relationship between median age and median income for Pennsylvania. Recreate this plot, making sure to include error bars for both the x and y variables, for all of the counties in the spreadsheet. Write a suitable caption discussing the relationship for the plot. (2 pts)

You will need the ACSSPP1Y2021.S0201-Column-Metadata.csv file to identify the relevant columns.

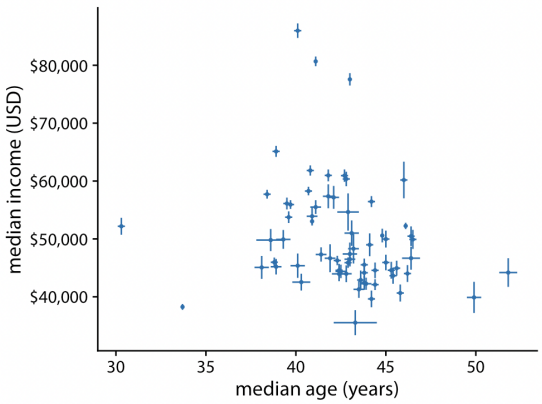

Figure 16.11: Median income versus median age for 67 counties in Pennsylvania. Error bars represent 90% confidence intervals.
Data source: 2015 Five-Year American Community Survey

Wilke, Claus O. Fundamentals of data visualization: a primer on making informative and compelling figures, Chapter 16. O'Reilly Media, 2019.

B. While discussing this plot, we mentioned that there may be a relationship between the size of the sample in a given county and the length of the error bars. Using your graph from part A, encode information about the county size in the plot. How does this information change the relationship discussed in part A? (2 pts)

C. Add a linear regression line to your plot, include standard error ribbons. Does this help you spot any outliers? Is there a trend among the outliers? (2 pts)

Problem 3 (10 points)

Using a dataset and corresponding graph previously created for this class, examine the data, find an additional fact that is contained within the data, and design a visualization with an interactive component 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 will need code for the interactive component) but you need to find something interesting and display it effectively.