Nastasia Pollas

Abstract
This Census project will offer viewers a chance at an exploratory analysis of the population, occupation type by sex in the US over a 5-year estimate, from 2017 to 2021.
Using https://www.socialexplorer.com/, I generated a report called occupation by stats by sex, using the following variables: total population by sex and age, educational attainment by sex, occupation categories, employment status and employment sector by sex, and occupation by total and sex. Viewers will get a basic overview of which states are the most densely populated, what sectors most people are employed in, and what the most common occupation types. The data has a few important gaps–(1) notably sex is categorized as a binary and may not accurately reflect the full population; (2) occupation is coded based on written responses, and, assumed by the Census, if not descriptive enough; (3) occupation excludes military persons, but employment sector includes them, so the employment and occupation may not be comparable to viewers.

Proposed site architecture
The website will contain 2 visualizations. The first visualization will be three bar charts connected by filter action. The first should have 51 marks (1 per state, including DC). The viewer can then select a state, and the corresponding two bar charts will redraw and display the population per sex.

The second visualization will be the scatterplot, filtered by state, showing the difference between occupation types. The dots will be colored based on which percentage of occupation by sex is higher in the state. The viewer can then hover over the marks to see the values.

If time permits, the last visualization will be a pie chart showing the different employment sectors per sex, since there are only 5 categories.

Citation of data source
Social Explorer Tables: ACS 2021 (5-Year Estimates) (SE), ACS 2021 (5-Year Estimates), Social Explorer; US Census Bureau

Plan for data analysis
1. Transpose the original data so that all the states are in 1 column, the variables are the header columns for the spreadsheet (complete)
    a. Removed Puerto Rico and re-summed the columns to change the total states values
    b. Removed the blank columns during the transpose
2. Import the data via csv file
3. For horizontal bar charts
    a. Set the width, height, and margin of the SVGs
    b. Width/x-scale will correspond to population

    c. Height/y-scale will be the age categories
    d. Change the geographical state using state management in D3
    e. Find some way to connect the 3 bar charts by filter so that filtering a state will change the male/female bar charts
    f. Find some way to display the charts in one row (side-by-side)
4. For scatterplot
    a. Set height, width, margin, and radius
    b. Create SVG
    c. Append the circles
        i. Cx = ≤ Female occupation category
        ii. Cy = ≥ corresponding male occupation category
    d. Label the plots
5. For pie chart
    a. Use the arc function and set it to the employment sector path