

Data 227 - Autumn 2023 HW 7
Data Visualization and Communication - Trimble
Due: Monday Dec 4, 2023 11:59pm (after final project!)

1. Draw some error bars

The American Community Survey is a major sampling campaign by the U. S. Census Bureau which collects a wide variety of statistics about people who live in the United States. For some statistics, the Census publishes uncertainty estimates (conventionally 90% error bands) for their estimates.

You can get data from the firehose at: <https://data.census.gov/table>

Make a scatterplot of two statistics (mean income vs. mean age, or mean years education vs. mean ages) from the ACS where each point is a county. Indicate the error bands in both dimensions on the graph. Use color to encode county population.

2. Network diagram

Make a visualization (static or interactive) that communicates something about a network dataset of your choice. Include a full caption that explains the origin of the dataset.

3. Interaction

Make a graph using a dataset of your choice that includes at least 50 data points and is interactive. Use a mode of interaction other than hover-over text (you are welcome and encouraged to use hover-over, but make sure you use something else too), such as brushing, subset-selection, or animation of some sort. If appropriate, you can use hyperlinks in your graph.

To share an interactive graph, you will need to export your graph as html, and should use a software package that embeds javascript containing the data into html. Altair has a limit of 5000 rows per datastructure; you probably shouldn't use more than 5000 rows for plotly either; it is best to choose a dataset (or subset a dataset) until you have fewer than 5000 things to display.

Since this is a dataset of your choice, please write a full caption that indicates where the dataset is from and what the dataset shows.