

CSSS 569 · Visualizing Data and Models

Problem Set 2

Professor: Chris Adolph, Political Science and CSSS

Winter Quarter 2018

Due in class on Tuesday, 6 February 2018

General instructions for homeworks: Homework can be handwritten or typed. For any exercises done with R or other statistical packages, you should attach all code you have written and all (interesting) output. Materials should be stapled together in order by problem. The most readable and elegant format for homework answers incorporates student comments, code, output, and graphics into a seamless narrative, as one would see in a textbook.

Problem: Using Graphics to Explore a Dataset

For this problem, you are encouraged to find (or quickly assemble from existing resources) a dataset from your field, from your own research, or that you simply find intriguing. You will use this dataset to practice creating strong exploratory graphics, and to learn new graphics skills (for most of you, in R).

I encourage you to use your own data, but if you prefer, you may instead use data from gapminder.org. From gapminder.org, click on the Data tab, and start downloading variables you wish to include in your analysis. You may find the site's visualization tools helpful in deciding which variables to select. Note that the data vary over years and countries; you will need to decide how to use or subset data over these ranges.

- a. Create two or three figures exploring either your own data or the gapminder data. These plots should either provide a powerful view of interesting patterns in the data, or drive home a finding from the data.

- You are allowed to transform any variables in any manner you wish. This includes creating categorical variables from continuous ones, or creating interaction terms.
 - You may run any data analysis you know on the data, and then integrate the results into your displays.
 - At least one display should use high-level plotting functions. Be sure to annotate this graphic as well as you can.
 - At least one display should rely only on graphical primitives. In R, these are things like `lines()`, `points()`, `text()`, and `polygon()`. If you don't use R, do your best to find your package's primitive graphics functions and use only those functions for one figure.
 - A figure consisting of small multiples counts as one figure.
 - Think carefully about the use of color. If you use color, explain your color choices.
 - Provide detailed explanatory captions. I should be able to understand your graphic just based on the caption and figure itself.
- b.** For each plot, also provide a brief explanation of why you chose that style of plot, referencing our discussions on aesthetics and cognition.