

36-721

Statistical Graphics and Visualization

Fall 2015

6.0 units

Homework 1: Legible Graphics

Due Sat 9/5/15, 5pm

You will be assigned a dataset to explore and summarize graphically.

The dataset `GoogleTrendsReport.csv` is a comma-separated value (CSV) file of a Google Trends report. Each row is a weekly record, from the start of 2004 to now. For each week, we have an index of “search interest relative to the highest point on the chart” for the search terms “data scientist,” “statistician,” and “data analyst.”

You can learn more about the Google Trends methodology on their site:

www.google.com/trends/explore#q=data%20scientist%2C%20statistician%2C%20data%20analyst

Plot the data legibly and summarize any interesting trends or features. Feel free to smooth, aggregate, or subset the data if you find it useful.

Please submit

- a **stand-alone image file** of your graph,
- a **one-paragraph writeup** of what you learned about the data, and
- **your code** (or written instructions) for making the graph.

See rubric on next page.

Component	Competent	Not yet competent
Legible	Image is high-resolution bitmap or vector format. Font is large enough to read easily. Data are not hidden or overwhelmed by ticks, axes, or gridlines. Different colors or symbols are easily distinguishable. Aspect ratio shows data clearly.	Image is low-resolution, grainy, or pixelly. Font is too small to read. Data are hidden by other graph elements. Colors or symbols cannot be distinguished. Aspect ratio causes data to be too bunched up or spread out to see patterns easily.
Comprehensible	Graphic has an informative title, axis labels, and legend (if relevant). Axis ticks are labeled with sensible, round numbers. Graphic axes, legend, colors, etc. are consistent across small multiples (if relevant).	Graphic has no (or unclear) title, axis labels, or legend. Axis ticks are unmarked or are marked at arbitrary, unhelpful numbers. Graphic elements are inconsistent across small multiples.
Reproducible	Submission includes minimal, clear, standalone code (or written instruction) for reproducing the graphic.	No code or instructions are included; or they are not standalone or do not run/compile; or contain far more than needed to reproduce this graphic.
Informative	Graph clearly highlights any trend or pattern in the data, which is indicated in the title and described in the writeup. Interesting differences or comparisons are plotted directly.	Graph highlights no interesting or useful pattern. Pattern is not indicated in title or not described in writeup. Readers have to mentally compute differences instead of seeing them directly.
Administrivia	Filenames follow naming convention: 721HW1_andrewID.xxx. Graph is created in R (if MSP).	Filenames do not follow convention. Graph is not in R.