Assignment 2: Information Visualization

SDS235: Visual Analytics

Due: March 1, 2017 by 11:55pm

Now that you’ve mastered the basics of data scraping and analysis, it’s time to dig in to the visual half of Visual Analytics. In this assignment, you’ll apply the techniques and tools we’ve discussed in class to make visual representations of a multidimensional dataset of your choosing. You may use the dataset you scraped in A1, something from Kaggle, or something completely different. You are welcome to use D3, Tableau, or the visualization packages in R or pandas to complete this assignment. As before, you’re welcome to work together on this assignment, as well as reference any material (books, online sources, etc.) that you find useful. However, you should cite all resources you utilize and you must submit your write-ups individually.

1 Visualization

Many real world datasets have a lot of richness in terms of the variety of data types: we may have have temporal (time-based) data, geographic data, categorical data, and quantitative financial data all rolled into one. Because of this, there are many ways one could choose to visualize a dataset to either explore, prove a point, or tell a story. Here are a few (low-hanging fruit) examples:

- Use a scatterplot to compare sales and profits
  Useful examples: D3 — Tableau — R — pandas

- Use line graphs to compare trends over time
  Useful examples: D3 — Tableau — R — pandas

- Use a parallel coordinates plot to show relationships between all columns
  Useful examples: D3 — Tableau — R — pandas

Or perhaps you can come up with something more... compelling? You may want to try out more than one method. The choice is up to you!

2 Reflection

To receive credit for this assignment, please post your visualization(s) along with a brief reflection to Moodle by 11:55pm on Wednesday March 1st. Specifically, please address the following questions:

- Which tool(s) did you use?
• Did you preprocess the data? Why or why not?
• What visual mappings / representations did you choose? Why those?
• Are there any patterns/trends in the data that your visualization highlights?
• Is your visualization useful? If yes, how so? If no, why not?
• Was anything harder than you expected? Easier?