

Homework #3

DS 3001, D-term 2018

100 points total [9% of your final grade]

Due: April 12, 2018 by 11:59pm

[no submission will be accepted after April 15, 2018 at 11:59pm]

Delivery: Submit via canvas

Overview

In this homework, you will create a data visualization based on an exploration of dataset of your choice. Your visualization should capture Tufte's principles of graphical excellence: "complex ideas communicated with clarity, precision, and efficiency."

Task: Visualization

To begin, you should identify a dataset of your choice and ask a compelling question of the data. As a starting point, here a few datasets (and lists of datasets):

- [AWS Public Data Sets](#)
- [Quandl](#), a search engine for datasets
- <http://www.kdnuggets.com/datasets/index.html>
- <https://github.com/caesar0301/awesome-public-datasets>
- <https://www.kaggle.com/datasets>

Feel free to use your own choice of dataset, so long as the data is non-trivial.

Based on the data of your choice, you should ask and answer a compelling question via a data visualization of your choice. You may clean, manipulate, and transform the data as you like. You may integrate multiple datasets, or none at all.

Ultimately, you should produce:

- A single image (JPG, PNG, or GIF) that effectively communicates the data. You may use any data visualization toolkit you like.
- A brief explanation describing your design. (a few paragraphs, 1/2 page or so).

Your explanation should justify the design choices you have made. Why did you choose the particular visualization type? What motivated your choice of size, color, and scale? Why are they appropriate for what you are trying to communicate?

Your explanation should document what you are trying to communicate, the benefits of your choice of visualization, as well as any downsides to your viz (e.g., are certain connections downplayed due to your choices?).

What to turn in:

- You should turn in One file: a PDF report containing your image and brief explanation about it. (hw3_yourname.pdf)
- This is an individual assignment, but you may discuss general strategies and approaches with other members of the class (refer to the syllabus for details of the homework collaboration policy). At the top of your report, please write the names of classmates you consulted and the nature of your discussion.