EN 605.42762 Data Visualization

Project #2 – Data Exploration and Design

Announced: First Day of Module #3

Due: First Day of Module #5

I. Purpose:

The purpose of this assignment is to introduce you to the process of exploring and visually analyzing data without even having to develop a visualization tool. You will pick a domain and dataset that you are interested in. The data should have at least 10 variables. The purpose of this assignment is to design different visualizations to illustrate individual aspects of the dataset under consideration.

II. Task:

1. Find a dataset of interest to you
2. Explore and analyze the data using Excel, Google Spreadsheets, R, Python, Matlab, Tableau, or simply by hand (drawing pictures or equations is fine). What are the characteristics of the data? Describe each of the data elements and list their category (nominal, ordinal, quantitative, etc…)
3. List five analytical questions that users examining the data might be wondering
4. By using pen and paper or tools such as Power Point, Excel, Photoshop, Gimp, Paint, Illustrator, Tableau, Qlik, etc… create visualizations that can be used to address some of the analytical questions. Of the five analytical questions, select three of them and design or sketch three different visualizations for each of them. (i.e. 3 analytical questions x 3 visualizations = 9 visualizations).
5. Write one paragraph for each of the drawings explaining the design, the purpose, and the logic behind your design.

III. Useful Links

- Visualization
  - https://github.com/d3/d3/wiki/Gallery
  - https://public.tableau.com/en-us/s/gallery
  - http://christopheviau.com/d3list/gallery.html
  - https://bost.ocks.org/mike/
  - http://textvis.lnu.se/

- Datasets (Tip: Find a dataset that interests you.)
  - www.kaggle.com/datasets
  - https://catalog.data.gov/dataset

- https://github.com/awesomedata/awesome-public-datasets
- Cereal dataset: http://davis.wpi.edu/xmdv/datasets/cereal.html
- http://nwdata.org
- http://www.scribblelive.com/blog/2012/03/30/data-sources/
- https://www.ons.gov.uk
- http://datacatalog.worldbank.org
- http://rhythm-of-food.net
- http://flowingdata.com

IV. What to submit

- Document (2+ pages) describing the data, analytical questions, and the 9 visualizations designed for three of the analytical questions. Document must include screenshots of the drawings, visualizations, or illustrations.

- Submit document through Blackboard. Please use the following file format: your_lastname_project02.docx or your_lastname_project02.pdf

V. Grading

Students will be evaluated based on the quality of the work, logic, clarity, and effort put into the design. Specifically,

- What dataset was used? There are no right or wrong datasets, but put effort in finding something of interest to you.
- How well the data was explored and described? Mean, standard deviation, median, max, etc…? Is there a table listing the elements?
- Quality of the five analytical questions that users examining the data might be wondering.
- Logic of the 9 designs, sketches, pictures, or visualizations proposed to explore the dataset
- Write-up and overall justification.