EN 605.662 Data Visualization

Project #4 – Interactive Visualization using JavaScript, R, or Python

Announced: First Day of Module #7

Due: First Day of Module #10

I. Purpose:

During the last few years, a number of software tools have been designed to help users explore complex data by creating visualizations and dashboards. Popular off-the-shelf tools include Tableau, Qlikview, Spotfire, Microsoft Power BI, MicroStrategy, Birst, and Logi among many others. Unfortunately, those tools only provide a small set of built-in visualizations that the user must use to visualize the data. Often users are interested in using advanced visualization techniques that are not available within those off-the-shelf applications.

Recently, a number of libraries have been released to help organizations develop new visualizations and illustration tools. Some of the popular libraries include D3, Chart.js, Plot.ly, Highcharts, Bokeh, ggplot, matplotlib, ProtoVIS, R Shiny, NVD3, etc… As data scientists, it is important for us to have a basic understanding of those libraries and their capabilities. The purpose of this assignment is to get familiar with open source libraries by developing three sample visualizations using any libraries for JavaScript, R, or Python.

II. Task:

1. Find 3 datasets that you would like to analyze. The data should have 3 or more variables and 100 or more rows. Datasets included within the packages / libraries cannot be used in the project.

2. Develop three <u>different</u> visualizations by leveraging the library of your choice to illustrate the datasets under consideration.
   - Students can develop their own visualizations or select samples from the corresponding libraries and update those as needed.
   - If using an existing sample from the library of your choice, aside from updating the data source, students should make at least one significant change to each visualization / source code they selected. Changes can be (but not limited to) changing colors, adding tooltips, updating axes, etc…. If students are modifying existing sample code, students must describe in the document the changes that they did to each of the original programs.
   - The visualization must have some sort of user interaction (e.g. either sorting, filtering, tooltip, etc…) to enable data exploration.
   - If using Jupyter Notebooks or R Markdown, please make sure you create three different files. See file structure below.

3. File structure: students should structure their project the following way.
    - your_lastname_project04/
        - Paper: your_lastname_project04.pdf
            - Introduction: What the project is about
            - Dataset: explain and provide links to the sources
            - Approach: explain which libraries you selected
            - Visualization #1: Explain, provide screenshot, and justify
            - Visualization #2: Explain, provide screenshot, and justify
            - Visualization #3: Explain, provide screenshot, and justify
            - Conclusion
            - References
        - src/
            - Sample01/
                - Index1.html or Sample1.py or Sample1.R, etc…
                - Data1.csv
                - Screenshot_sample01.jpg
            - Sample02/
                - Index2.html or Sample2.py or Sample2.R, etc…
                - Data2.csv
                - Screenshot_sample02.jpg
            - Sample03/
                - Index3.html or Sample3.py or Sample3.R, etc…
                - Data3.csv
                - Screenshot_sample03.jpg
        - Requeriments.txt: List dependencies (e.g. R, Shiny R, Python, matplotlib, D3, etc…). If using JavaScript, students must include a directory with the JavaScript files.

III. Useful Links

- https://d3js.org/
- https://github.com/d3/d3/wiki/Tutorials
- http://alignedleft.com/tutorials/d3/
- https://shiny.rstudio.com
- http://nvd3.org
- https://plot.ly/python/

IV. What to submit

- A .zip file with the file structure shown above
- A paper describing the projects, the datasets that were chosen, the thee visualizations or dashboards that were developed (including screenshot), and explanation of what was updated from any sample code that was used.
- Submit document through Blackboard. Please use the following file format: your_lastname_project04.zip