

**Name:** Syllabus for DATA 608 - Knowledge and Visual Analytics

**Description:** In this course students will learn non-statistical aspects of elucidating from data its information content which leads to knowledge. Several differing visual techniques will be examined to gain this knowledge through exploratory use of visualizations as well as visualization techniques for presenting data to a variety of stakeholders. Exploratory techniques look to find patterns in the data. Finding patterns that underlie the system's characteristics when the data sets are very large or have many dimensions by reducing the dimensionality in intelligent ways is a complex task that often includes user direction. Presentation visualizations provide the viewer with useful information and knowledge since the visualizations are created with context in mind. This course will be completed completely online: assignments will be issued most weeks and due before the next class begins. One larger visualization project will be completed at the end of the course and will be displayed on a public website.

**Pre-requisites:** DATA 602 and DATA 607

**Course objectives:** At the completion of this course, students should be able to:

- Understand the principles of data visualization
- Apply these principles in both interactive and status graphics
- Demonstrate familiarity with several modern data visualization tools and libraries
- Efficiently create visualizations for diagnostics and analysis

**Grading:** You will be graded on your understanding of the theory of data visualization, your presentation of information, and how you apply the tools we will be using the course (code quality, etc.). Each homework assignment will be worth 10% of your overall class grade, with your lowest score being dropped. Late projects will be docked 1% of your overall class grade for each day late. Additionally, there will be a small quiz at the beginning of some classes which will be worth 2% of your final grade, with your two lowest scores being dropped. The breakdown of points available over the course will be as follows:

Homework Assignment: 60%

Discussion / In-class 10%

Final Project: 30%

Students are encouraged to discuss approaches to homework assignments together, and I expect many of you to take similar approaches to visualization, though you each must turn in your own work, and it should have been completed solely by you. However, the final project will be done individually. Participation is part of the grade, but those who engage above and beyond the minimum discussion questions and replies will be awarded up to 10% of additional credit at the instructor's discretion.

The only reason students have failed this course in the past is that they did not put enough thought into their Final Projects and their rights to the data they were working with and did not communicate issues to me in a timely manner.

**Textbooks:** There is no textbook for this course: readings will be linked to or distributed by your instructor.

**Hardware, Software, and Network Requirements:** You should have network access for lectures (every other week). You should also have a computer on which you can save software and different files. A Github account will be mandatory for submitting projects.

**Academic Integrity & Netiquette:** I expect that you will honor the principle of academic integrity. This means that all academic work will be done by the individual to whom it is assigned, without unauthorized aid of any kind. I strongly encourage you to look for help on coding problems from the internet (StackOverflow, etc.) If you use code found online in your academic work, provide a link to the original code as a reference so it is clear what was added. We all stand on the shoulders of giants, and you will be graded on how you creatively apply borrowed code to a novel problem.

Constructive criticism of your fellow students (and of me and the class) is encouraged. It should go without saying that bullying and intolerant speech will not be tolerated.

**Projects & Schedule:** Project due dates are listed in the class-by-class schedule below:

Module	Meeting	Project Due Date	Module Notes
Exploratory Data Analysis	Wednesday, 02/09/2020	Sunday, 02/09/2020	Exploratory data visualization with R.  Skill focus: static visualizations, R Markdown
Big Data Visualization	Wednesday, 02/19/2020	Sunday, 02/23/2020	Big Data Visualization with R or Python  Skill focus: Data cleansing, “Bring your code to the data”, Bigvis, Datashader
Interactive visualizations in R	Wednesday, 03/04/2020	Sunday, 03/08/2020	Interactive data visualizations in Shiny  Skill focus: web development concepts, Shiny
Interactive Visualizations in Python	Wednesday, 03/18/2020	Sunday, 03/22/2020	Interactive data visualizations with Plotly and Dash  Skill focus: web development concepts, Plotly, Dash
Introduction to Javascript	Wednesday, 04/01/2020	Sunday, 04/05/2020	Build your first true web app.  Skill focus: web development concepts (frontend), HTML, basic JavaScript
Visualizing data using JavaScript	Wednesday, 04/15/2020	Sunday, 04/19/2020	Visualize data on the web.  Skill focus: d3.js, Flask (optional)
Final Projects	Wednesday, 05/13/2020	Sunday, 05/17/2020	Present your final projects