Fall term 2016/2017

CNSC6012: DATA AND NETWORK VISUALIZATION

Course coordinator: Professor Roberta Sinatra, sinatrar@ceu.edu
No. of Credits: 2
Prerequisites: Basic programming skills in any programming language (python, matlab and R preferred, but not necessary), basic statistics
Course Level: Master and PhD
Office (temporary): 609 Nador 11, Budapest
Office hours: TBA or by appointment

Brief introduction to the course:
The Internet and modern computers have given us vast amounts of data, so it is more important than ever to understand and analyze these data. A picture is worth a thousand words, so visualizations, from scientific plots and infographics to interactive data explorers, are crucial to summarize and communicate new discoveries.

The goals of the course:
The major goals of this course are to understand how visual representations can help in the analysis and understanding of complex data, how to design effective visualizations, and how to create your own visualizations using programming skills and visualization software. We will achieve these goals by evaluating existing visualizations, teaching general theoretical principles of information design, and by walking students through the analysis and visualization of real world datasets during hands-on classes.

The learning outcomes of the course:
By the end of the course, students will have acquired the following skills:
- Basic data harvesting, storage and cleaning with automated computer programs
- Apply methods for visualization of data from a variety of fields
- Use basic principles of human perception and cognition in visualization
- Create some basic web-based interactive visualizations
- Learning to layout and visualize network data

Course Organization:
Lectures: 12 classes of 100min. For most classes, we will spend the first 30 minutes introducing the day's concepts, and spend the rest of the class doing the lab exercises or evaluating visualizations. Therefore, use of a computer will be required during some lectures. Students can form groups and use their own laptops.
Topics
- Basics of human perception and cognition of visualization
- Meaningful visualization for research
- Data “munging” or cleaning to process and visualize data
- Visualization of static data with Matplotlib, using Python
- Principles of Network visualizations (layouts, information reduction)
- Network visualization using Gephi
- Basic interactive visualizations using D3

Tentative calendar
1st class: Is a picture worth 1000 words? Principles of information Design
2nd class: Getting you from a dataset to several plots
3rd class: Lots of visualization examples, and practice going from data to chart
4th class: Line charts, Pie-charts, Box plots, Violin plots
5th class: Visualizing multivariate Data
6th class: More practice with multivariate data
7th class: Discussion about good and bad visualization practices, taken from research papers and from the web (as part of the course assignment)
8th class: Visualizing networks with Gephi
9th class: More visualization with Gephi. Visualization with Networkx
10th class: Interactive visualization
11th class: Interactive visualization
12th class: Project Presentation

Course Requirements/Assessment
Students are expected to attend lectures and hands-on sessions, to hand in one assignment during the course and to develop a project during the entire term.
Grading:
- Attendance of the classes and hands-on sessions: 40% of the final grade
- Assignment: 20% of the final grade
- Final project: 40% of the final grade

Final Project
For the final project, students will have to apply and show proficiency with the principles and tools used during the course. A number of options for projects will be suggested in class.
Textbooks
- The Visual Display of Quantitative Information
- Interactive Data Visualization for the Web
- Bill Mark Lutz, Learning Python, O’Reilly (2013) – Also available for free online

Further information, such as the course website, assessment deadlines, office hours, contact details etc. will be given during the course.

Important note
This is the first time CEU offers this course. Hence it is possible that some parts of the syllabus will change during the course, based on the feedback during the first weeks and to tailor the course on the students’ skills. Any modification to the syllabus will be discussed with the students. Students are responsible for information given in class.

Cheating
In short: don’t do it! You may work with friends to help guide problem solving or consult stack overflow or similar to work out a solution, but copying—from friends, previous students, or the Internet—is strictly prohibited. If caught cheating, you will fail this course. Ask questions in recitation and at office hours.