CMSE 402: Visualization of Scientific Datasets

Instructor: Brian O'Shea, oshea@msu.edu

Course Description

One of the critical steps in data analysis is creating visual representations that facilitate the interpretation of the data, for the purposes of communication, analysis, and decision-making. As datasets and models become ever more complex, and the questions that we ask of them become more sophisticated, it is critical to create graphics that communicate information as clearly and effectively as possible. The overall goal of this course is to give you an introduction to the core principles, methods, and techniques of data visualization that will help you to create meaningful graphics that effectively communicate information. It will also introduce you to a variety of software tools that can be used to visualize a variety of different types of data (e.g., geographic, multivariate, statistical, vector, etc.).

By the end of this course, you will be able to:

1. Identify the key aspects of visualization of datasets and how they affect the interpretation of data
2. Analyze a dataset, determine the questions that can be asked of it, and create one or more visualizations using standard techniques that effectively answer those questions
3. Apply standard techniques to visualize multi-dimensional scalar and vector datasets using pre-built toolkits.

We will work toward the goals expressed above throughout this course using a range of activities – primarily by writing software both individually and in small groups, but also through discussion, presentations, and other types of exercises.

Topics covered

The primary topics covered in this course include:

- The goals of data visualization
- Types of visualizations: 1D, 2D/planar, 3D/volumetric, statistical, temporal, multivariate, hierarchical, network
- The basics of human visual perception
- Evaluating the effectiveness of data visualization
- Visualization tools and toolkits
- Tying data analysis to visualization

Expectations for you (prior to the start of the semester)

The prerequisites for this course are CMSE 201 and 202, as well as multivariate calculus (Calc III; MTH 234 or 254H, LB 220, or the equivalent). In order for you to fully participate in this class, you are expected to do the following prior to the beginning of the semester:

- Be able to use the Unix command line to do basic operations (creating and changing directories, listing their contents, creating, deleting, and moving files, and so forth). (see Software Carpentry Unix command line tutorial)
- Be able to work with the Git version control system to clone repositories, check in changes, resolve merge conflicts, and push changes to a remote repository. (see the Atlassian Git tutorial)
- Be able to program comfortably in the Python 3.x programming language (see the official Python 3 tutorial).
- Have a computer that has Python installed, along with standard numerical packages such as matplotlib, numpy, scipy, h5py, and scikit-learn. I strongly urge you to install the Anaconda Python distribution, available at https://www.continuum.io/downloads . I strongly suggest installing the Python 3.6 version.
- Have a laptop that you can bring to class every week with the required Python distribution, along with a power cord and VGA or HDMI adapter (to plug into the display ports in our classroom). If you do not have a laptop, please let me know as soon as possible.

Required reading materials

This class has two required texts, which will be heavily used starting in the first week of class:

1. *The Visual Display of Quantitative Information, 2nd Ed.* by Edward Tufte (ISBN 978-0961392147; Amazon link)
2. *The Truthful Art: Data, Charts, and Maps for Communication*, by Alberto Cairo (ISBN 978-0321934079; Amazon link)

The total cost of the paper versions of these books is approximately $80 on Amazon.com. These textbooks will be supplemented by a variety of journal articles and web-based materials, which will be provided during the semester.

Other required materials

In-class programming assignments are a critical part of the learning process in this course. To that end, you are expected to bring your laptop, power cord, and VGA or HDMI adapter (to plug into the LCD projector) to class every day. If you do not have a laptop, or if your laptop won't run the software that we need for class, I can arrange to give you a spare machine to use this semester.

Course activities

Class participation: Active class participation (led both by the instructor and by students) is critical to the success of this course. As such, you are expected to attend class every day, bring the required materials (including your laptop and copies of any pre-class assignments and readings) and to actively participate in the in-class activities.

Pre-class assignments: We will often have assignments that are due prior to class. The purpose of these assignments is to introduce new material and give you some practice with it so that we can focus on experimentation and implementation in class. These assignments may consist of watching videos, reading materials, answering questions, or writing software, and will be due at 11:59 p.m. the night before class via the course's GitHub Classroom. We will generally use Python for pre-class assignments that involve programming.

In-class assignments: Class sessions will be held twice a week, and will be broken up into presentations, discussions, and programming/visualization activities that will allow you to implement (and get immediate feedback on) what you have just learned. In-class programming activities will be turned in at the end of the class session via the course's GitHub Classroom.

Homework: You will have periodic homework assignments (every two or three weeks) that will provide a more in-depth exploration of the materials covered in class. These will be pursued either individually or in pairs, and will be turned in by the given deadline using GitHub Classroom.

Semester Projects: This class will have a semester project that will involve synthesizing some subset of the techniques that you have learned about for a project that relates to your personal or research interests. There will be substantial programming involved, as well as a writeup and a presentation at the end of the semester. More details will be available near the middle of the semester.

Course meeting time and location

This class will meet on Mondays and Wednesdays from 10:20-11:40 a.m. in 1220 Engineering Building. The final exam session is scheduled for 7:45-9:45 a.m. on Thursday May 3, 2018, in the same room. We may not use the final exam session, depending on how long presentations take during the last week of class.

Other important information

Course Website, Calendar, and discussion channel: This course uses a GitHub repository to distribute course readings, found at https://github.com/bwoshea/CMSE_402_Spring_2018.

Accompanying course information, including the official course calendar, can be found in this repository. All assignments will be handed out and turned in via a GitHub Classroom. To access both of these things, you will need to create a GitHub account and send it to me. We will also have a course discussion channel (#sciviz-spring18) on the MSU CMSE Slack group, http://cmse-courses.slack.com. If you need an invitation to the Slack group, please contact me and I will issue one (if you have a msu.edu email address, you should be able to create your own). **Most course communication will take place via this Slack channel!** Please note that this course nominally also has a Desire2Learn page (at http://d2l.msu.edu), but at most it will be used for keeping track of grades.

Class attendance: This class is heavily based on material presented and worked on in class, and it is critical that you attend and participate fully every week! Lack of attendance or habitual lateness will not be tolerated, and repeated violations (3 or more unexcused absences or late arrivals) may result in a reduced grade.

Classroom behavior: Respectful and responsible behavior is expected at all times, which includes not interrupting other students, turning off your cell phone ringer, refraining from non-course-related use of electronic devices, and not using offensive or demeaning language in our discussions. Flagrant or repeated violations of this expectation may result in ejection from the classroom, grade-related penalties, and/or involvement of the university Ombudsperson. In particular, behaviors that could be considered discriminatory or harassing, or unwanted sexual attention, will not be tolerated and will be immediately reported to the appropriate MSU office (which may include the MSU Police Department).

Academic Honesty: Intellectual integrity is the foundation of the scientific enterprise. In all instances, you must do your own work and give proper credit to all sources that you use in your papers and oral presentations – any instance of submitting another person's work, ideas, or wording as your own counts as plagiarism. This includes failing to cite any direct quotations in your essays, research paper, class debate, or written presentation. The MSU College of Natural Science adheres to the policies of academic honesty as specified in the General Student Regulations 1.0, Protection of Scholarship and Grades, and in the all-University statement on Integrity of Scholarship and Grades, which are included in Spartan Life: Student Handbook and Resource Guide. Students who plagiarize will receive a 0.0 in the course. In addition, University policy requires that any cheating offense, regardless of the magnitude of the infraction or punishment decided upon by the professor, be reported immediately to the dean of the student's college.

It is important to note that **plagiarism in the context of this course includes, but is not limited to,** directly copying another student's solutions to pre-class, in-class, or homework assignments; copying materials from online sources, textbooks, or other reference materials *without citing those references in your source code or documentation*, or having somebody else do your pre-class work, in-class work, or homework on your behalf. Any work that is done in collaboration with other students should state this explicitly, and have their names as well as yours listed clearly.

More broadly, we ask that students adhere to the Spartan Code of Honor academic pledge, as written by the Associated Students of Michigan State University (ASMSU): “As a Spartan, I will strive to uphold values of the highest ethical standard. I will practice honesty in my work, foster honesty in my peers, and take pride in knowing that honor is worth more than grades. I will carry these values beyond my time as a student at Michigan State University, continuing the endeavor to build personal integrity in all that I do.”

Instructor information

Contact information:

Brian O’Shea
Associate Professor of Computational Mathematics, Science and Engineering
and Physics and Astronomy
Engineering office: 1508F Engineering Building, 517-432-0331
BPS office: 3258 BPS, 518-884-5638
Email: oshea@msu.edu
Web: www.msu.edu/~oshea/

Office hours: Fridays from 1:30-3:30 p.m. in 1508F EB and by appointment.

Scheduled travel: As of the writing of this syllabus, I am scheduled to be out of town on Wednesday January 17th to give a talk, and will likely have a substitute instructor.

Grading information

Activity	Grade percentage
Participation/attendance/in-class assignments	25
Pre-class assignments:	20
Homework assignments	30
Semester project (all components together)	25
Total:	100%

Grading scale

4.0	≥ 90%
3.5	≥ 85%
3.0	≥ 80%
2.5	≥ 75%
2.0	≥ 70%
1.5	≥ 65%
1.0	≥ 60%
0.0	< 60%