University of Montana

ScholarWorks at University of Montana

University of Montana Course Syllabi

Open Educational Resources (OER)

Fall 9-2-2018

CSCI 444.01: Data Visualization

Jesse V. Johnson
University of Montana, Missoula

Welcome to Data Visualization
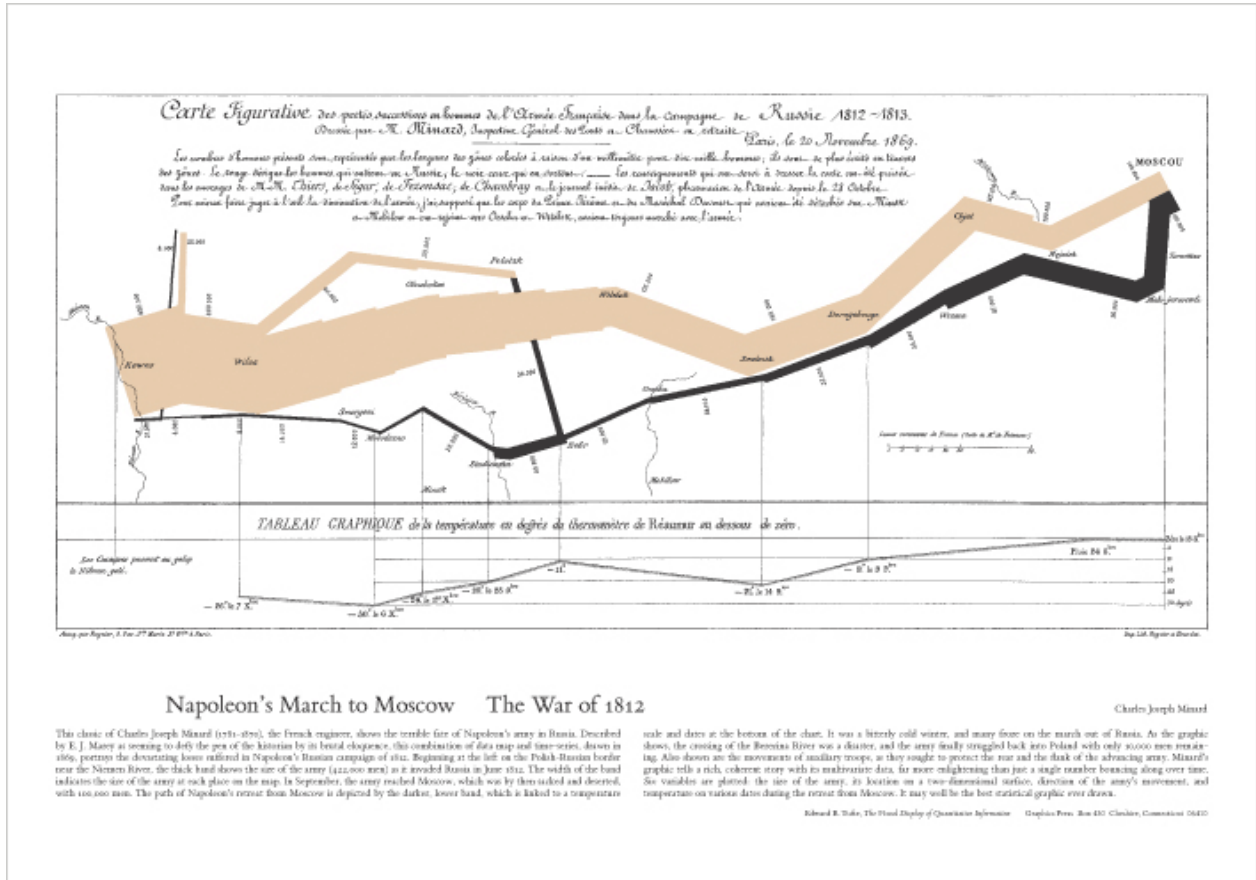

ADVANCED FORUM

- Ability to program in a language of your choice, as demonstrated by successful completion of data structures or a similar course.
- Evidence of mathematical maturity as shown by successful completion of calculus and/or statistics.
- Maturity enough to show up for class, consistently.
- Maturity enough to offer constructive criticism to your peers.

Course Objectives

This course emphasizes the practice of data visualization, compelling students to identify and master tools to produce visualizations of data having different types of relations. Hence, a majority of the student's time is spent in the creation of original visuals. To guide the process, lecture time will initially be spent considering visualizations that represent best, and worst practices. Later lectures will consider the way humans perceive, and how limits of human perception drive decisions about the display of quantitative information. Moreover, students acquire a sophisticated framework for assessing the quality of visualizations, the technical skills required to produce visualizations, and an understanding of the mathematical challenges that underlay many data visualization techniques.

After taking this course, students will be able to:

- quickly adapt to any quantitative visualization programming envioronment,
- write short, modular programs of moderate complexity to visualize quantitative data,
- apply numerical methods to make data more accessible,
- critique visualizations using a framework focused on the accurate display of quantitative information and the limitations of human perception.

Course Logistics

The course is project driven. All projects share the same approach, but differ in data sets used. Given a data set, your approach will always be as follows:

1. Form a hypothesis from the data. State the hypothesis as clearly as possible.
2. Identify two visuals. One should be a good example of how a hypothesis can be supported with similar data (e.g. when dealing with time series, show a time series plot that is conclusive, and especially well done in terms of relating variables). The other visual should be a bad visual that uses similar data, and is difficult or impossible to draw conclusions from. In short provide: one visual for how to do it, and one for how not to do it. Obtain electronic copies of the visuals and submit them with the assignment.
3. Write *one half to one* page about the visuals you found, where you found them, and how you can critique them **based on the lecture material**.
4. Produce three visuals of your own, from the data, that support your hypothesis.
5. Write *one half to one* page detailing how your visuals are consistent with the examples, support the hypothesis, and **are consistent with the material covered in lecture**.

6. Produce *one* visual that refutes or calls into question your hypothesis and provide a paragraph explaining why. Summarize your findings in a few overhead slides and be prepared to defend them to the class.

Given the enrollment this semester, I will come up with a randomized method of determining who presents each project, otherwise too much time would be spent in presentations. Everyone will be prepared to present, but only some of you will. All of you will submit assignments for a grade.

To help differentiate between good and bad visuals, we will be studying the opinions of Edward Tufte. His books are very accessible, and feature a large number of engaging graphics. Those graphics will be displayed on the projector, and we will discuss them in class.

I place no restrictions on the languages/tools used for visualization of data. You may use whatever suites you. However, I personally use Python, and might offer you more insightful advice should you need it, if you use Python. Some of the assignments require the use of Python, at least in the sense that you have to interface your visualization tools to a Python program. D3.js is an especially nice library that is highly amenable to sharing your work with others. Using it makes a lot of sense to me.

Course Topics

Review the weekly material on the Moodle. It is based on how the course was offered last time. Some of the data sets will change, but by and large, we will follow the same trajectory through the course.

Course Materials

We will use the following text:
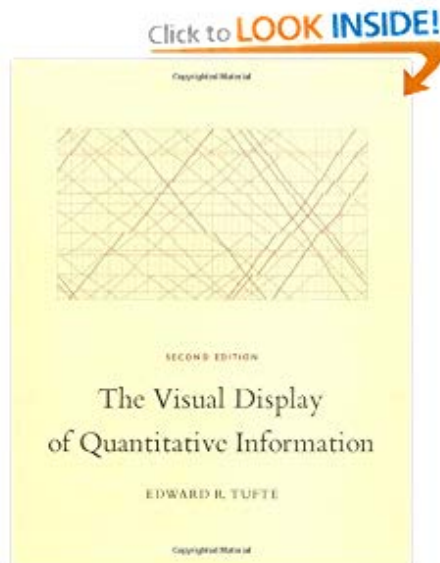
- **Hardcover:** 200 pages
- **Publisher:** Graphics Pr; 2nd edition (May 2001)
- **Language:** English
- **ISBN-10:** 0961392142

- **ISBN-13:** 978-0961392147
- **Product Dimensions:** 9 x 11 inches

At $26, this might be the best book you purchase for college.

Grading

Grading Breakdown

Grades of A-F will be assigned based on a percentage of the total possible points earned. The break points are as follows.

A	94-100
A-	90-93
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	73-76
C-	70-82
D+	67-69
D	63-76
D-	60-62
F	0-59

I reserve the right to make changes to the grade breakdown that will be favorable to students grades.

Assessment

Grades will be based on the following forms of evaluation.

Component	Description of Component	Percent of Grade
Projects	I will try to prepare five of these. Each one will feature a new data set. The requirements for each are described above, under 'course logistics'. The data sets will steadily increase in complexity, dimension, and interrelations.	50%
Final	This is a visualization of your own choosing. You will find your own data and visualize it in whatever manner you like. The evaluation technique is the same.	30%

Engagement	In each class, students will be randomly selected to demonstrate some proficiency in class. This will often be to demonstrate a visual they have seen. Sometimes it could be to show how to perform some task with the program they are using. I will notify students of what is expected before each class. I have been calling these "Batsu" games.	20%

Pass Fail

Students taking the course pass/no pass are required to earn a grade of D or better in order to pass.

Attendance Policy

The design of the course will make attendance necessary. If you know in advance that you will need to miss class, I encourage you to inform me. If you do so, and the reason is fair, I will not call you for engagement activities. If you fail to inform me of an absence, and you are randomly called for some engagement activity, you will receive a zero. Observe that a zero is much more painful than an F.

Late Assignments

Other than in in exceptional circumstances, such as family emergencies, late work will not be accepted. If you do have an emergency that causes you to miss an important classroom activity, please let me know in advance and I will be very accomodating.

Academic Integrity

As a student of the University of Montana, you are responsible for upholding all rules in the student conduct code. There are aspects of that code that are of particular importance in Computer Science courses. The electronic nature of the many assignments facilitates their dissemination. To be clear, from the student conduct code:

1. Plagiarism: Representing another person's words, ideas, data, or materials as one's own.

6. Submitting work previously presented in another course: Knowingly making such submission in violation of stated course requirements.

Of course, all other aspects of the student conduct code will be enforced as well. These are just the two that are commonly violated.

I will interpret these guidelines to the letter. Students found in violation will be penalized with the maximum punishment permitted in the student conduct code. That is to say, the matter will be handed over to the Academic Dean and academic misconduct proceedings will take place.

In order to reconcile encouraged interaction between students and the academic misconduct policies, you must credit other students in your work. If, for example, you worked with others to develop some algorithm, or solve some homework problem, specifically mention those that you have worked with in the assignment that is handed in. Similarly, you must properly document and credit any online resources that you use.

If you collaborate with others, the instructor has the right to question you about the material turned in. If it is evident that your understanding of what you turn in is weak, your grade will be lowered.

Students are to uphold a level of conduct becoming of adults. The use of abusive or demeaning speech is not permitted under the student conduct code, and will not be tolerated in this course.

Disabilities

Students with disabilities are encouraged to meet with me to discuss any accommodations they require.

Other Issues

Turn off your cellphone, or set it to vibrate in class. Take the call outside the classroom.

I expect your full attention during class time. Do not be looking at a cell phone or a computer screen.

Do not talk in the classroom during lecture. Take it outside.

CLOSE