

DSC 365/465 Syllabus

Fall 2020

Basic Info:

Dr. Eli T Brown

Office: CDM 711 (<https://depaul.zoom.us/j/92704658191>)

Office Hours:

Mondays from 3:00-4:00pm

Mondays after class

Wednesdays from 12:00-1:00pm

Website: <https://www.cdm.depaul.edu/Faculty-and-Staff/Pages/faculty-info.aspx?fid=1311>

e-mail: eli.t.brown@depaul.edu

will respond by night of next business day

include CSC 365/465 in the subject!

Location:

Zoom Link: <https://depaul.zoom.us/j/94516042159?pwd=Z1pmNU5aVmdSNnFHQXp5VTJjUETjZz09>

Password: 465VIS

Dates:

Midterm: week of Oct 19

Final Presentations: week of Nov 16

Final Project Reports Due: Nov 23

Course Summary:

This course will be an introduction to data visualization techniques for exploration and analysis of large data sets with a wide variety of data types, such as might be discovered from a wide range of fields including commercial, financial, medical, scientific and engineering. Topics will include visual encoding of numeric data, human visual perception, effective visualization design, graphical integrity, visualizing distributions, color theory, basic network graph visualization, geospatial visualization and some additional topics.

Prerequisites: PREREQUISITE(S): IT 403 and (CSC 401 or IT 411)

This course assumes that you have had a basic course in statistics along with an introductory programming course (e.g. Python).

Learning Goals:

By the end of the quarter, students will be able to use visualization to discover insights in data and craft visualizations to communicate such insights. You will be able to choose appropriate visualizations to perform exploratory analysis and investigate data, and you will be able to create explanatory visualizations to communicate your findings. To do this you will leverage your knowledge of a collection

of different visualization techniques along with understanding of how to appropriately apply them and optimize their appearance for clutter, clarity, graphical integrity and human perception. Specific goals for each module are included to help you track and reinforce your learning.

Expectations:

Contacting Me

Please get in touch if you have questions or would like to schedule a meeting outside office hours. **Email is the best option.** NOTE: I respond quickly when possible, but my policy is that you should receive a response by the night of the next business day (i.e. an email Tuesday gets a response by Wednesday night).

Additionally, please do not email me questions to which the answer is right here on the syllabus (e.g. when is the midterm?) or posted in a note on the D2L page. I may not answer such emails, so if you do not receive a response, double-check that the answer is not easy to find. When you email me, **include DSC 465 online** in the subject.

My policy below about not debugging student code is especially true by email. It is rarely productive as I cannot see it run or test a small change for you. Do not send me a snippet of code and ask 'what is wrong?'. You will likely not get any reply. See Expectations/Coding for what you should do.

Finally - come to my office hours or call me during office hours. I encourage you to get help if you need it. I will not judge you for needing help. I will help you.

Workload

Visualization is a wide topic and we have a lot to cover, so the course is a significant amount of work. Part of the work is learning to create visualizations, and like any skill this takes practice. I expect you to put in the time making mistakes and figuring things out, and in return you can expect me not to waste your time with unhelpful exercises. Much of that practice will be on the homework assignments and the project. Once we get to the second half of the class, there is less new material and less individual homework because we shift focus to group project work. Expect that the content and homework assignments are a bit front-loaded. The dividing line between halves of the class is the midterm in week 6, which will test your core visualization knowledge. Expect it to be fair but not easy.

Finally, there are additional weekly exercises that will aid your learning and build community, including **quizzes** to reinforce key points, **tutorials** to give you practice with R implementation, and **discussions** to develop the class' higher level skills like critique. These are individually small parts of your grade, but I expect you to complete all of them because they will make a big difference in your outcomes for the course. The grades section below explains the components of the course in more detail.

Coding

We will use Tableau and R in this course, but Tableau has a very shallow learning curve so we will focus efforts on R. Though you may use Tableau for most assignments, you will likely find yourselves drawn toward coding in R because you get more control. If this sounds unlikely to you, at least give it a shot though the tutorials. The prerequisites require some familiarity with coding, so I will not be covering the very basics, e.g. concepts of writing instructions to the computer, variables and loops. Learning R and its

data toolkits is well worth the time, but you should be able to get through by adapting examples from the ggplot book (see below) and elsewhere, including my demonstrations. This will require experimenting.

Finally, note that as a matter of policy, I do not debug student code. You cannot learn coding without practicing and making mistakes (banging your head against the wall, frankly). Work with each other, use the recommended books and provided tutorials. Start by making an established example run and then start adding and changing. Note that I am an expert in data visualization and not an expert in R anyway. That said, I have significant software experience and if you hit a wall, I'm happy to try to help. Come to office hours (or screenshare) and always start by telling me what you have already tried (this is always a good idea when asking anyone a technical question, including a boss or IT support). Treat me as a last resort for debugging.

Discussions Boards and Homework Discussion

The discussion boards are for you. I generally stay off them but browse to look for issues from time to time. Specifically, what I must check for and cannot tolerate on the boards is academic integrity violations. You may not post answers or answer code to the forums. It is okay to discuss homework, give suggestions for plan of attack or discuss strategies, but everyone must do and submit their own work. The forums are also used for forming project groups and weekly discussion posts. In all cases I expect you all to be courteous to each other and respectful of others' time and effort. If there are issues, I will have to shut the forums down for the whole class.

Tutorials, Quizzes

Like the Discussion boards, Tutorials and Quizzes are there *for you*. Quizzes help you remember key points of information and Tutorials give you examples to work from for your assignments.

Grades

See the Grades section for details on the proportional components of your final grade and policies, e.g. for late assignments.

Textbooks and Software:

Books

The following books are **necessary** (they are 'required', but they are also free to read online, and 'required' makes it sound like you have to buy them, so I use a different word. It's worth noting that it's not expensive, as textbooks go, to get a hardcopy.

- Covers ggplot with examples that will be critical for doing the homework:
R Graphics Cookbook: Practical Recipes for Visualizing Data, 2nd Edition. Winston Chang. O'Reilly Media, November 2018. <https://r-graphics.org/>

The following books are **strongly** recommended for the course. You need them to get the most out of the course.

- Covers the world of R libraries for data manipulation and more on ggplot:
R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. Hadley Wickham and Garrett Grolemund. O'Reilly Media, January 2017. <https://r4ds.had.co.nz/>

- Nearly a textbook for the course, provides more detail on most course topics:
Fundamentals of Data Visualization. Claus O. Wilke. O'Reilly Media, April 2019.

The following books are **recommended**, with explanations of what they cover:

- Excellent coverage of the graphing parts of the course, including clutter, clarity, graphical integrity and guidelines. This is a very influential book on visualization design and is beautifully written with excellent explanations and illustrations (some used for the course).
The Visual Display of Quantitative Information, 2nd Edition, Edward Tufte. ISBN-13 978-0961392147
- An actual visualization textbook and written by one of the top scientists and most influential people in the field. It is a little too geared for future visualization researchers to be our textbook, e.g. including how to perform design studies to make interactive visualization software, but I definitely recommend it for its overall coverage of the science and practice of visualization.
Visualization Analysis and Design. Tamara Munzner. A K Peters/CRC Press. December 2014.
- Much more detail on visual perception for a visualization audience.
Visual Thinking for Design. Colin Ware. ISBN-13: 978-0123708960
- Introduction to visualization from someone who has built a career out of doing it in industry, starting at Google and now consulting. Implementation info is available in Excel.
Storytelling with Data. Cole Nussbaumer Knaflic. ISBN-13: 978-1119002253.
- More design-oriented view of visualization, beautiful book including case studies.
The Functional Art: An Introduction to Information Graphics and Visualization. Alberto Cairo. ISBN-13: 978-0321834737
- Covering D3, a JavaScript-based library for building interactive tools that we discuss at the end of the course.
- *Interactive Data Visualization for the Web*. Scott Murray. <https://alignedleft.com/work/d3-book-2e>
- Using MATLAB for visualization
MATLAB Data Analysis and Visualization. Antonio Siciliano. ISBN-13 978-9812837516

Software

- **Tableau** kindly provides student licenses!
 - You can go to their website and sign up as a student and they will verify your student address at DePaul: <https://www.tableau.com/academic/students>
 - You have to renew every year but it is free while you are a student
- **R Studio** is a convenient and free way to use R (the language). You can download it here, but note in the installation instructions that if you don't have R itself, you'll need to install that too by following their link: <https://rstudio.com/products/rstudio/download/> (Get the *Desktop* version. Server is for running your own server for interactive R visualizations.)

Grading:

Work in this course will be evaluated principally through homework assignments (4), a midterm exam and a final project. The breakdown for the grading will be

- Quizzes: 10%
- Discussions and Participation: 10%

- Homework: 30%
- Midterm: 20%
- Final Project 30%

Quizzes

Weekly quizzes will be administered through D2L and will test the material in the lectures. These quizzes are open book/notes/friends. I ask that you consider the questions yourself in order to test yourself on the material as this 'recall practice' helps you remember the material. I will review each quiz the following week, but you can still retake to improve your score.

Discussions and Participation

The discussion assignments will help you engage with the material and the community by finding discussing examples and questions. You must make your own post and then respond to others. This is worth a portion of your grade. Additional credit can be granted for strong participation in the forums, e.g. helping fellow students.

Homework

Homework assignments will be submitted online on the D2L website. No email submissions will be accepted. Submissions after the due date will come with an automatic 20% penalty. No homework will be accepted after the lecture following the due date because I may review assignment material in a posted note or video. The same type of discussion with peers that is allowed on the discussion board is allowed with respect to homework assignments, however, each student must turn in original work, i.e. your own write-up. Please ask if you need an extension on assignments.

Midterm

The midterm will be held on the sixth week of class (week of **Oct 19**). Makeups for the midterm will only be given in extreme circumstances and requests for a makeup must be made as soon as possible. Documentation must be supplied of the relevant circumstances. You must take the midterm to pass the course.

Final project

The final project in the course will be a group effort to build an original collection of visualizations for a dataset of your choice. Most of the work in the latter half of the course is focused on the project. The deliverables consist of check-ins, a presentation and a written final report. You will present your final project on the 10th week of class, and then you will have until the final exam date (one week later) to submit your final report.

The project description document is your friend. There are ideas for data sources, an effective outline for the report, and the grading rubric. Most importantly, refer to that document for project milestones, group composition rules, and policies, e.g. for removing a group member if absolutely necessary.

Extra Credit

There are experiments going on with human computer interaction (HCI) and psychology researchers at DePaul. They need participants. You can get extra credit in this class for participating. Usually, they will start to be available around the fifth week of the quarter. Go to this link to check it out:

<http://www.cdm.depaul.edu/academics/research/Pages/Instructions-for-Participants.aspx>

Final Grading Note

I aim to make grading fair and focused on feedback. Email me if you think this has not been the case.

Feedback for Me

Online Course Evaluations

Evaluations are a way for students to provide valuable feedback regarding their instructor and the course. Detailed feedback will enable the instructor to continuously tailor teaching methods and course content to meet the learning goals of the course and the academic needs of the students. They are a requirement of the course and are key to continue to provide you with the highest quality of teaching. The evaluations are anonymous; the instructor and administration do not track who entered what responses. A program is used to check if the student completed the evaluations, but the evaluation is completely separate from the student's identity. Since 100% participation is our goal, students are sent periodic reminders over three weeks. Students do not receive reminders once they complete the evaluation. Students complete the evaluation online in [CampusConnect](#). I personally read these after each quarter, so while Administration will review them for my job performance, I will be looking carefully for useful constructive criticism that I can use to improve the course. Please help me out.

Other Feedback

Please feel free to email me feedback about the course, during or after. I work hard to make this effective at interesting. If you let me know what worked for you and didn't, I may be able to improve the course. Even if I cannot do that in time to help you, future students will appreciate your time. I will too.

Policies

Changes to Syllabus

This syllabus is subject to change as necessary during the quarter. If a change occurs, it will be thoroughly addressed during class, posted under Announcements in D2L.

Academic Integrity and Plagiarism

This course will be subject to the university's academic integrity policy. More information can be found at <http://academicintegrity.depaul.edu/>. If you have any questions be sure to consult with your professor.

Academic Policies

All students are required to manage their class schedules each term in accordance with the deadlines for enrolling and withdrawing as indicated in the [University Academic Calendar](#). Information on enrollment, withdrawal, grading and incompletes can be found at: cdm.depaul.edu/enrollment.

Preferred Name & Gender Pronouns

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance,

and nationalities. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the quarter so that I may make appropriate changes to my records. Please also note that students may choose to identify within the University community with a preferred first name that differs from their legal name and may also update their gender. The preferred first name will appear in University related systems and documents except where the use of the legal name is necessitated or required by University business or legal need. For more information and instructions on how to do so, please see the Student Preferred Name and Gender Policy at <http://policies.depaul.edu/policy/policy.aspx?pid=332>

Students with Disabilities

Students who feel they may need an accommodation based on the impact of a disability should contact the instructor privately to discuss their specific needs. All discussions will remain confidential.

To ensure that you receive the most appropriate accommodation based on your needs, contact the instructor as early as possible in the quarter (preferably within the first week of class), and make sure that you have contacted the Center for Students with Disabilities (CSD) at: csd@depaul.edu.

Lewis Center 1420, 25 East Jackson Blvd.

Phone number: (312)362-8002

Fax: (312)362-6544

TTY: (773)325.7296