

DSC 465 Syllabus: DRAFT - this may change but will be thoroughly discussed if so.

Data Visualization (Formerly CSC 465)

Fall 2019

Classroom: LEWIS 01516 at Loop Campus

Terriell (Doug) Scrimager, MSPA

Location: LEWIS 01516 at Loop Campus

Office Hours: Th from 4:00 - 5:00 pm

Th from 9:00 - 9:30pm

I can schedule other sessions as needed - typically outside of regular business hours with some limited weekend availability.

Website: <http://d2l.depaul.edu>

Office Phone: I don't have an assigned Office Phone ATM. Please reach out by email.

e-mail: tscrimag@depaul.edu

will respond by night of next business day

Include DSC465 in the subject!

While email is generally reliable - if you do not hear back within next business day please resend.

Key Dates

Midterm: October 10th (through 12th for online students)

Project Proposal Due: October 17th (the earlier the better)

Final Presentations: November 21st

Final Project Reports Due: November 25th

Grades Due: December 5th

Drop Dates: <https://academics.depaul.edu/calendar/Pages/default.aspx>

Course Summary:

This course will be an introduction to data visualization techniques for exploration and analysis of large data sets from a wide range of fields including commercial, financial, medical, scientific and engineering applications. Topics will include visual encoding of numeric data, effective visualization design, graphical integrity, visualizing distributions and correlation, false-color techniques for feature extraction and enhancement, basic network graph visualization, geospatial visualization and some additional topics.

Student/Instructor responsibilities:

First and foremost, this is a graduate course so you are expected to be a part of a learning community and to take primary responsibility for your overall experience in this class. This class provides a structured learning experience, guidance, resources as well as the instructor which are a significant part of your success in learning so please take advantage of all parts of the course. As an instructor I am here to facilitate learning and am dedicated to the success of each individual in it. The online tools are there to help you manage your course-load, due dates, homework and class project activities. All work is expected to be completely reproducible with all needed code, data, citations and documentation included in any submission whether homework, online activities or final project. I should be able to follow your process and achieve/find the same results. For every quarter-hour expect to spend a similar

amount outside of class each week so up to 4 hours per week over and above the lecture. Final projects may require some extra effort.

Special Consideration for Distance Learning:

There are some special considerations for distance-learning. Please refer to this page:

<https://www.cdm.depaul.edu/onlinelearning/Pages/Exams.aspx>. If you are a distance learner it is critical that you keep up with class recordings and participate in the online forums. The recordings are available soon after the classroom period is over - my advice is not to wait until Sunday night to review the online updates and watch the lectures or you may find yourself running out of time to complete the upcoming assignments and to get feedback if you have questions about the material, assignments or online activities. If you are new to distance learning or still unsure of anything please let me know ASAP.

Amount of Coding:

The course will explore both visualization software packages and code interfaces for data visualization. We will use Tableau and R, and you will likely find yourselves drawn toward coding in R because you get more control. I may also do examples in Python if there is interest. At the end of the quarter I will discuss options for interactive visualization. Not all projects will need a heavy amount of custom coding, but do expect that you will need to condition your data and I will need to be able to understand how you did so. For instance, if you performed any of your conditioning without code (such as in Excel) you will need to thoroughly document what you did to be completely reproducible.

Learning Goals:

By the end of the quarter, students will be able to use visualization to discover insights in data. 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 clarity, target audience and integrity.

Textbooks and Printed Resources:

The following texts are **strongly** recommended for the course.

- William Cleveland, The Elements of Graphing Data, ISBN-13 978-0963488411
- Edward Tufte, The Visual Display of Quantitative Information, 2nd Edition, ISBN-13 978-0961392147
- (on R) Winston Chang, R Graphics Cookbook: Practical Recipes for Visualizing Data, Publisher: O'Reilly Media, December 2012

Also, the following books are recommended depending on your interests

- (on D3) Scott Murray, Interactive Data Visualization for the Web, see <http://chimera.labs.oreilly.com/books/1230000000345/index.html>
- (more on perception) Colin Ware, Visual Thinking for Design, ISBN-13: 978-0123708960
- (intro to vis, but not practical implementations) Cole Nussbaumer Knaflic, Storytelling with Data, ISBN-13: 978-1119002253.
- (more design-based view of vis) Alberto Cairo, The Functional Art: An introduction to information graphics and visualization, ISBN-13: 978-0321834737

- (on MATLAB) Antonio Siciliano, MATLAB Data Analysis and Visualization, ISBN-13 978-9812837516

Software Resources:

- **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://www.rstudio.com/products/rstudio/download3/>
- **Python** programming language - PyCharm and Spider (From Anaconda) are available IDE's.

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, R, C, C++, etc.) including sufficient knowledge of programming to be capable of data preparation and ability to implement algorithmic approaches to data transformation. For both Python and R, the typical experience is to use existing libraries and not to have you implement detailed algorithms that are already available.

Grading:

Work in this course will be evaluated principally through a mixture of in-class and online participation in the community (not limited to graph examples, reflections, peer support, deconstruction/construction exercises and media searches), homework assignments (4), a midterm exam and a final project. The breakdown for the grading will be

- Weekly Online Activities 20%
- Homework 20%
- Midterm 20%
- Final Project 40%

Points Scale:

90 - 100	A
80 - 89	B
70 - 79	C
60 - 69	D
0 - 59	F

Participation

Participation in class and on the discussion board will help you engage with the material and the community, and improve your understanding of the material. Online activities will be administered

through D2L and will test the material in the lectures. These activities are open book/notes/friends. I ask that you consider the questions yourself in order to test yourself on the material. The online activities will be graded for content, effort and contribution to material. While effective writing is a key part of communication, the intent will not be to grade heavily for grammar - although today we all have tools to help with the mechanics such as spelling, grammar, etc. Asking questions, providing additional (positive) insights on peer work and identifying material that enriches peer understanding are all expectations for your participation. All students will need to complete the online activities to receive full credit for participation. The weekly online activities should be possible to complete in about an hour but sometimes may take a little more. Please do review DePaul's policy 'Guiding Principles on Speech and Expression' (http://policies.depaul.edu/documents/other/Guiding_Principles_20170530.pdf) so we stay focused on discussions that will move forward our comprehension of the material.

Homework

Homework will be submitted online on the D2L website. No email submissions will be accepted. Homework submitted 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 will review the assignment. The same type of discussion with peers that is allowed on the discussion board is allowed with respect to homework assignments. Each student must turn in original work, i.e. your own write-up. Please ask if you need an extension on assignments.

Midterm

The in-class midterm will be held on the sixth week of class. If you are in-class student the expectation is that you will take the midterm in class - only under extraordinary circumstances will you be allowed to schedule a proctored exam and this must be discussed with me prior if at all possible. 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. The open period for the midterm is from the Oct 10th through Oct 12th in order to schedule proctored exams for online students. Please note: proctored and makeup students do NOT receive the same exam as the in-class students although all material on any test will have been discussed and reviewed.

Final project

The final project in the course will be a group effort to build an original collection of visualizations for some data set. It will be up to your group to find an appropriate dataset. I will be providing a list of sites that have good datasets, but you are more than welcome to use data of your own. Also, keep in mind the data set you select should be free of personally identifiable information and protected health information. If there are any concerns with data privacy I will work with you to find another dataset. The group assignments should be identified by the midterm so you can start with group activities at that point. Please start putting together ideas and forming groups online as early as the second week. As part of your project proposal you will need to outline areas such as identified dataset, participant work breakdown and final deliverables. I will provide a format for your project proposal submission.

The final project consists of both a presentation and a write-up. You will present your final project on the 10th week of class, November 21st and then you will have until the end of Monday, November 25th to submit your final report.

Group Composition

Groups are generally a mixture of in-class and online students, but these rules are crucial:

- There can be in-class-only and online-only groups.
- Online-only groups **MUST** have someone available to present for the last class and **MUST** check in with me **WEEKLY** by email in the last few weeks of class.
- Groups that have a mix of online and in-class students **MUST** have at least 2 online students.
- In order to enforce these rules so this process runs smoothly, I reserve the right to edit group membership after you have formed them.

It can be difficult to keep online and in-class students working smoothly together. The skills practiced in doing this will be useful to your careers. I expect you all to handle this professionally, courteously and respectfully. Deviations from this will adversely affect your grade in two ways: your participation score may be reduced, and your teammates may report their dissatisfaction on the peer evaluations.

A note about final project groups and non-performance as a team-member on a project:

The final project in this course is very broad in its scope allowing your group to focus on a wide range of dataset types for visualization, and on a wide range of techniques for visualizing the data. Group members are expected to participate fully and equitably in the group, and part of the final project grading rubric will be a peer evaluation that will form part of the final project grade.

Usually, the peer evaluation and documentation, including the meeting minutes, in addition to an overall desire for excellence, is sufficient motivation for individuals to contribute a fair share to the team project. However, in extreme cases, individuals have been known to completely cease contributing to a team project. If this is the case, a team has the right to notify the instructor **unanimously (other than the individual being sanctioned)** that the individual is no longer contributing and the team no longer wants the individual on the team.

It is expected that a team will be able to show significant effort towards reconciling the issue prior to such an extreme action. Note also that this is not a decision to be made lightly, as expulsion from a team will result in **the loss of 40% of the class grade**, i.e. the group portion of the class grade, for the person expelled. Because this is such a serious decision, any team that makes this decision will also experience a deduction of **10% of the final project grade**.

Attendance and Participation:

Attendance (watching the lectures for online students) will count towards your participation score, as will participation in the online forum. You are expected to participate in online discussions with your peers. One of the most important parts of a class is building a learning community. I will be expecting you to be available to help each other with constructive suggestions and to complete the online activities as discussed previously.

Please ask questions (in class or online forum or both) if something is not clear - if you have a question, chances are that someone else has the same question.

Inappropriate behavior, such as rude responses will be noted and will count against your participation score. Note also that posting of solutions is forbidden and will result in the forum being shut down, and also may constitute an academic integrity violation. You may give suggestions for how to approach a problem, and may discuss strategies, but full code solutions are not to be posted. All online activities should be the product of your effort and not solely the result of peers.

What to Expect:

In this class you should expect to spend a significant amount of time outside of class reviewing course materials, topics, exploring external sources, sharing through online activities and working on homework problems. Throughout the course, I may distribute handouts of notes on various topics, and certain in-class materials including sample programs will be available on the class website. Nevertheless you are expected to take notes during class and/or review lectures online to make sure you understand each week's material.

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 DSC465** in the subject. I am generally not available during the business day except for posted office hours, so while I will watch email and respond ASAP it will usually be after/before work hours.

At this point I have scheduled office hours in our classroom on class days. This is subject to change. I encourage you to get help if you need it - I am here to help. For online students we will make arrangements by email for a tele/zoom conference during regular office hours and some limited availability in other times. Anyone can post to me in the online forum for help as well.

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.

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, I will be looking carefully for useful constructive criticism that I can use to improve the course. **Please help me out as I will consider all input - student input is what drives constant improvement in our courses, our delivery of content and our service to the student community.**

Extra Credit for Experiment Participation

There are experiments going on with human computer interaction (HCI) 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:

<https://www.cdm.depaul.edu/academics/research/Pages/ResearchLabs.aspx>

Capstone

If you are interested in data visualization capstone please contact your advisor - as described above there are many opportunities with various labs to find an interesting and fruitful project.

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.

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