Data 504: Network and Text Analytics

Drew University, Fall 2020

<https://drew.zoom.us/j/92435070902>

**Instructor Information:** Ellie Small, HSC 304, x3347, [esmall1@drew.edu](mailto:esmall1@drew.edu)

***Office Hours***: Mo 4:00-5:30, We 12:45-2:15 and by appointment at <https://drew.zoom.us/j/96632237239>

**Meeting Time:** M 6:00 – 8:30 pm. on Zoom: <https://drew.zoom.us/j/92435070902>

**Course Description:** This course will cover networks and their visualization, including weighted and directed networks. It will also discuss important network measures such as centrality, transitivity, and reciprocity. Applications focus on technology, information systems, and social settings, including the internet, the World Wide Web, and social media. In addition, it covers text mining topics including principal phrase mining. You will learn how to do this using the R language in the R-studio environment.

**Course Topics:** Networks and their representation, adjacency matrices, weighted and directed networks, bipartite networks, trees, degree, walks and paths, components, centrality, transitivity, reciprocity, similarity, homophily and assortative mixing, text mining and principal phrases.

**Student Learning Outcomes:** By the end of the course, students will be able to:

1. Create and visualize a network from a dataset consisting of items and connections between them
2. Demonstrate knowledge of network terminology
3. Solve real-world network problems by applying technology to their models
4. Create and demonstrate projects dealing with networks
5. Analyze one, or multiple bodies of texts by obtaining summary information such as frequent words and principal phrases and any interesting aspects of the texts they may reveal
6. Present and explain text mining results to an audience

**Required Text**: *Networks (2nd edition)*, by Mark Newman, Oxford University Press, ISBN-13: 978-0-19-880509-0.

**Course Resources**: Please visit the course Moodle site at: <http://moodle.drew.edu>. The Moodle site contains Power Point presentations, code used in class, homework, assignments, data sets, and other resources.

**Laptop Computer and Statistical Software:** We will be using the software package R and the R-studio environment**.** These are open-source and as such may be downloaded and installed for free. Please visit <https://cran.r-project.org/mirrors.html> to select a mirror from which to download R, and visit <https://www.rstudio.com/products/rstudio/download/#download> to select an installer for R-studio.

You need to complete the installation before class on Wednesday, August 26. Try to install the software as soon as you can, so you have time to get help, if necessary. If you have any trouble, please contact the University Technology Service Center Helpdesk. Contact information and hours may be found here: <http://www.drew.edu/ut/visiting-or-contacting-us/>.

Please log into R and R-studio every day before class; you will be asked to access it and perform small coding assignments during class. Try to use the same computer for homework that you will use for quizzes and exams and regularly practice a variety of problems using R; if you do this, you will be much less likely to experience computer problems during the tests. Should you have a problem with your computer and it was purchased through Drew, take it immediately to the Helpdesk. When you hand in your laptop for repairs, you may request a free loaner. If you purchased your computer elsewhere, you need to arrange for the repair, but you may rent a loaner for $5 per day while your computer is being repaired. Details of the loaner policy may be found at <http://www.drew.edu/ut/policies/loaner-policy/>.

**Attendance Policy**: Class attendance and participation are integral to the academic experience at Drew University, and this remains true for online courses during the Fall 2020 semester. For scheduled online course meetings, attendance and participation require that students join the synchronous Zoom sessions unless prior arrangements have been made with the instructor. Students who face challenges participating at the scheduled time (e.g., issues with time zones or WIFI connectivity) should contact the instructor immediately to make alternate arrangements for class participation.

**University Absence Policy:** In addition to the course attendance policy, students should be aware of their rights and responsibilities regarding absences for legitimate reasons as described in the University’s Absence Policy. This policy states that students can expect reasonable accommodations for (1) the equivalent of one week of class or less missed for legitimate planned absences (religious holidays) so long as they inform the faculty member of their planned absences in the first week of the semester; and (2) the equivalent of one week of class or less for legitimate unplanned absences (illness or a death in the family) so long as they notify the instructor prior to the class and as soon as they are aware of the unplanned absence (or within 24 hours of the absence in extraordinary cases), and submit the [Absence Verification form](http://www.drew.edu/center-academic-excellence/about-us/forms/), including appropriate documentation. For all legitimate absences, students are responsible for all material covered in missed classes, and students should realize that absences can indirectly affect final grades as a result of the impact absences have on learning.

Students experiencing extended illnesses (i.e., greater than one week), due to either the coronavirus or other medical conditions, should contact the Associate Provost at [cae-admin@drew.edu](mailto:cae-admin@drew.edu) to discuss the best course of action.

**How to get the most out of your Drew online experience:** Drew faculty have been working diligently to design interactive online courses that highlight student-faculty and peer-to-peer interactions that are hallmarks of the Drew experience. Many class sessions will use Zoom class meetings to facilitate seminar discussions and collaborative work. Here are several suggestions for making the most of your synchronous Drew Zoom learning experience:

* Join the Zoom class from a computer (with a video camera), one participant per computer.  As you will need to open documents, participate in chat, and join breakout groups, it is not possible for you to participate in class fully from a phone or tablet alone. If joining via phone or tablet, you will need an additional device to simultaneously access class documents.
* All participants are encouraged to have their cameras turned on so that the class is a face-to-face experience for everyone.
* If possible, you should join Zoom classes from a quiet, adequately-lit room without distractions. If you are uncomfortable Zooming into class from your home, please find a safe, private space from which to join the Zoom classroom. You may also hang a sheet to provide a neutral background when Zooming into class with the camera turned on.
* When in the Zoom room, students will need to learn and utilize features that manage classroom participation. For example, you may need to mute your microphone when you are not speaking. Use the chat and raise your hand features of Zoom to join the conversation when others are speaking.
* Consider using headphones that have a microphone built into the cord as this will maximize your listening and speaking participation.
* For further information on Zoom usage, consult the [Learning Remotely](http://www.drew.edu/university-technology/teaching-learning-and-working-remotely/learning-remotely/) website.

**Canceled Class:** If class is canceled for any reason on an exam date, the exam will be held on the next class period.

**Extra Credit:** There is none. You are all expected to do the work that is assigned to the best of your ability.

**Email:** Check your email regularly, since assignments and important announcements (such as cancellation of class) are delivered via email. It is every Drew student’s responsibility to regularly check his/her email and to not allow the email box to completely fill up (once filled, no new incoming messages can be received). Claiming to have missed an announcement or an assignment because you did not receive the email due to a full mailbox is not an acceptable reason to be excused or extended more time to complete the work.  
  
**Office Hours:** Please feel free to see me during office hours about any difficulties with the course before it gets too close to an exam date. If the office hours conflict with your course schedule, meetings can be arranged by appointment. Note that problems faced early are much easier to solve! Often a new idea learned takes time for its meaning and implications to settle in; allow yourself enough time to digest it properly. NOTE, however, that poor attendance without a valid reason is not an excuse to misuse office hours. Office hours are intended as an opportunity to ask questions about the material and homework after time has been spent on your own with the material in question; they are not make-up lectures or tutoring sessions. Students who routinely miss lectures and expect to make up the time during office hours will be disappointed.

**Homework Assignments**: There will be a homework assignment for this course most weeks. The homework assignments are mandatory; it is imperative that you attempt to do all of them, as they are instrumental to your understanding of the course material. Homework assignments are not graded per se. However, you may be called upon to show your work when homework assignments are discussed in class and your performance at that time will count towards your participation grade.

**Policy on Late Work and Missed Assignments:** *There are no make-ups in this class, even for excused absences*. *Missed or late assignments will earn a zero.* If you know in advance that you will miss a class, please let me know as soon as possible. If you know in advance that you will miss a quiz or exam and you let me know ahead of time, you may be able to take it early by making appropriate arrangements with me.Students unable to submit work on time as a result of a documented extended illness, due to either the coronavirus or other medical condition, should contact their instructor and [Dana Giroux](mailto:dgiroux@drew.edu) in the Office of Accessibility Resources (OAR) to coordinate appropriate accommodations.

**Components of Your Grade:**

|  |  |  |
| --- | --- | --- |
| **Component** | **Weight in Final Grade** | **Description** |
| **Internship application** | **5%** | Every student in the MS Data Analytics program must apply for a summer internship, which is managed through the career center. This part of the grade is for successfully completing your resume and cover letter and applying for at least one internship. Please note that many internships have September application deadlines! For those students who do not have internship requirements, please talk to me about alternative ways to earn this grade, such as applying for a job or graduate school. |
| **Attendance and Participation** | **15%** | You are expected to attend class as much as you can, and actively participate during class. Attending the class does not simply mean showing up and extremely poor participation during class will also cause your attendance grade to be lowered. Examples of poor participation include arriving late, leaving early, checking your phone during class and using your computer inappropriately. |
| **Quizzes and assignments** | **20%** | There will be a quiz or assignment most weeks; quizzes will be approximately 15 minutes during class. Due dates for assignments will be made known when the assignment is set. |
| **Midterm** | **25%** |  |
| **Final Exam** | **35%** | The final exam will consist of a project that will be presented during the final week. Grade is for the quality of the written project as well as the quality of the presentation. |

**NOTE: The material covered and format of the exams are subject to change.**

**Grade Cutoffs:** We will use the following scale in determining your course grade from your course average. The course grades are not curved.

A = 92 – 100; A- = 90 – 92; B+ = 88 – 90; B = 82 – 88; B- = 80 – 82; C+ = 78 – 80; C = 72 – 78;

C- = 70 – 72; F = 0 – 70

**Academic Integrity**: All students are required to uphold the highest academic standards. Any case of academic dishonesty will be dealt with according to the guidelines and procedures outlined in Drew University’s “Standards of Academic Integrity: Guidelines and Procedures,” which is located in the academic policies section of the catalog. A copy of this document can be accessed on the CLA Dean’s U-Know space by clicking on “Academic Integrity Standards.”

No collaboration is permitted on exams and quizzes and the only programs you are allowed to run during exams and quizzes are R and RStudio. You may discuss the homework assignments with other students, but you should make sure that you understand the responses. If you simply type in the answer that someone else tells you, you lose a valuable opportunity to test your understanding. You are not allowed to have any cell phone out during quizzes and exams. You are not allowed to reference any papers, websites, or other people during exams and quizzes except for the “cheat sheets” you are allowed to have for exams.

**Requesting Accommodations at Drew:**

*Students requesting accommodations for the first time* are instructed to contact Accessibility Resources. Although a disclosure may take place at any time during the semester, students are encouraged to do so early in the semester, because, in general, accommodations are not implemented retroactively. For additional information, visit <http://www.drew.edu/academic-services/disabilityservices>

*Returning Students with Approved Accommodations* are instructed to submit their request to Accessibility Resources, ideally within the first two weeks of class. This allows the office sufficient lead time to process the request. Please complete the accommodations request at: <http://www.drew.edu/academic-services/disabilityservices/request-for-accommodations>  
  
*Office of Accessibility Resources contact information:*  
Director: Dana Giroux  
Location: Brothers College, Room 119-B  
Contact information: Phone 973-408-3962, Email: [dgiroux@drew.edu](mailto:dgiroux@drew.edu) , [disabilityserv@drew.edu](mailto:disabilityserv@drew.edu)

**Study Tips:**Read through the material covered in class shortly after attending to reinforce the material. Make sure you understand everything. Make sure you can do all the R examples and try practicing changing the code in various ways to see how the results are affected.

You should do all assigned homework. Note that assigned homework represents the *minimum* amount of work that you should do between classes. In addition, you should review material covered in earlier classes. If an assignment is difficult, you should practice extra problems of that type. It is expected that a typical student will spend two hours outside of class for every hour inside of class. Thus, to do well, the typical student will be working a minimum of 6-8 hours per week outside of class.

It cannot be stressed enough how essential communication is to succeed in this course. After identifying topics that may be giving you trouble, please communicate this information to your instructor or contact the Center for Academic Excellence. There is no such thing as a bad or unwelcome question. Additionally, please communicate with each other. In a course going this fast, collaboration is almost a necessity to keep up with the material. Explaining concepts and examples to each other is a great way to learn. It is my goal to create a comfortable environment conducive to learning.

Because working with other students is a major part of this course, it is important that everyone participate. You can participate only if you come to class prepared.

**Getting Additional Help:**

The material in this course, as in most technological courses, builds on itself. It is critically important that you do not fall behind. This is not usually a problem for students with good attendance who work consistently outside of class, but if you find yourself having difficulty with the material, you should seek additional help immediately. First and foremost, ask questions in class in the time allotted for this purpose. If you still have questions, resources available to you include instructor office hours. If you choose to utilize this option, it is your responsibility to come prepared with specific questions, as office hours are not meant to be tutoring sessions.

**Keys to Success:**

* Prior to quizzes and exams, be able to solve every homework assignment.
* Review the material from class the same day it is given.
* Study a bit every day rather than in bursts just before an exam.
* Seek help appropriately.
* Be an active learner. When working on problems, try to complete them on your own before asking anyone for help. Pay attention in class and ask questions as they come to you. Do not let small confusions mount up. If you have a question, there are always others in the class with the same question, so please ask!
* Work all the problems assigned (and more if you feel you need extra practice on a given concept). This will expose you to the breadth of ideas not covered in lecture.

**Tentative Schedule:**

|  |  |
| --- | --- |
| Class | Notes: |
| 8/24 |  |
| 8/31 | **Last day to drop without a “W”: 9/4** |
| 9/7 |  |
| 9/14 |  |
| 9/21 |  |
| 9/28 | **No class** |
| 10/5 |  |
| 10/12 |  |
| 10/19 | **Last day to drop with a “W”: 10/23** |
| 10/26 |  |
| 11/2 |  |
| 11/9 |  |
| 11/16 |  |
| 11/23 |  |
| 11/30 |  |
| 12/7 | **Final: 7:30 to 10:30 p.m.** |

***Note that all exam dates are tentative and may change in format as well as time!***

If you need to reschedule the final exam, you need to fill out the form at: <https://www.drew.edu/academic-services/forms/rescheduling-an-exam-final>