

DSC 223 Syllabus

Introduction to Data Science

Fall 2022 Block 2 (9/26 to 10/19)

Professor:

Dr. Tyler George
West Science 311
tgeorge@cornellcollege.edu

Office Hours (time for you):

Monday – Thursday
3:05pm – 4:05pm
West Science 311

Other Office Hours:

I am available far beyond those times listed. Please email me and we can set up a time to chat about class material or whatever you prefer!

Class Meeting Times and Locations

| Room | Monday | Tuesday | Wednesday | Thursday | Friday |
|----------|------------------|------------------|------------------|----------|-----------------|
| West 200 | 9:00am - 11:00am | 9:30am - 11:00am | 9:00am - 11:00am | | |
| West 320 | 1:00pm - 3:00pm | | | | 1:00pm - 2:30pm |

You Are A Priority

My goal this block is to help you learn the material. I want to first and foremost recognize that you are an individual and thus are unique and may learn uniquely. Additionally, your health and well being are priority one. Learning cannot happen effectively if you don't meet your other personal needs. That all being said, I have structured the class in a way that I, from experience teaching and learning myself, think will be most beneficial for the majority of students. I promise you that I will do my best to create an inclusive and engaging learning environment. I ask that you keep an open line of communication between us for when you may need help and/or flexibly. You and your learning are why I am here.

Course Description

Managing and interpreting an overwhelming amount of raw data is part of the foundation of our information society and economy. People use computers and statistics to translate, process, and visualize raw data, enabling new understandings that in turn contribute new knowledge to the world. Data Science is a newly developing field that merges ideas from both statistics and computer science to address these issues. In this course statistics will inform the discussion about what appropriate goals are for learning from the data and how the data will answer the questions raised. The computer science perspective will help us figure out which goals are actually feasible computationally, and how to achieve them.

Learning Objectives

At the end of this course I would like you to be to use software's including RStudio and GitHub to respect, explore, understand, and utilize data in a way that is replicable. This course supports the Educational Priorities and Outcomes of Cornell College with emphasis on knowledge, inquiry, reasoning, and communication, ethical behavior, citizenship, and vocation. Your emphasis on knowledge is in the skills you will learn and apply in various interdisciplinary fields. You will inquire when investigating data – seeing patterns or trends and exploring them to. Your reasoning skills are built and tested when making decisions based on the data and your own programmed visualizations and numerical summaries. Your group work in class and group project presentations will help you practice your communication of statistical analysis. When you make decisions about what data to work with, how to treat the data, and how to talk about your results in an ethical way you practice good ethical behavior. Some of our analysis' will be with data from institutions such as governments or organizations that have an influence on the public – these types of analysis' can inform public policies and are our way, as data scientists, to practice citizenship. Lastly, you will learn about the field of data science and the types of knowledge and training that would be required to support your vocation as a data scientist.

Tentative Grade Scale

| | |
|--------|----|
| 93-100 | A |
| 90-92 | A- |
| 87-89 | B+ |
| 83-86 | B |
| 80-82 | B- |
| 77-79 | C+ |
| 73-76 | C |
| 70-72 | C- |
| 67-69 | D+ |
| 63-66 | D |
| 60-62 | D- |

Prerequisite:

To be successful in this class, you should have completed either Foundations of Computer Science (CSC 140) or Statistical Methods I (STA 201).

Open Access Books – Free!

All of materials for this class are free. Our primary book is [Data Science in a Box](#) by Mine Çetinkaya-Rundel. This book is a fabulous book for both R and version control (our major topics). Our secondary books are [R for Data Science](#) by Wickham and Grolemund and [Introduction to Modern Statistics](#) by Mine Çetinkaya-Rundel and Johanna Hardin.

Course Site and Moodle

Moodle is where you will find course announcements, technology instructions, links to some course components. My course website, linked on Moodle, is where our notes, assignments, and other materials will be hosted.

Software – No need to install

We will use a combination of technologies in this course including Git, GitHub, R, and RStudio (server). Luckily for you I have put lots of effort into setting all of this on a machine we have on campus that we will all access with a web browser! You don't need to install any – in fact for a while I prefer you don't. More on this in class.

If you have any technical problems you should contact IT as soon as possible. [Submit a Work Order!](#)

Group Work

In this class, I would like you to work in groups for a variety of reasons. A large part of this class is communicating analysis – not just completing analysis. At the beginning of the block, groups will be formed. You should expect to work with this group every day. When we work in groups in class we will decide on roles, specifically who is controlling the one screen will rotate). Group members will rotate roles between tasks to help make sure everybody is sharing work load, feels included, and learns equally. Groups will be randomized and change twice during the block. You won't be working in a group for everything; your quizzes, and exams may all be individual. I may also have some activities where some, if not all of it will be completed individually. This is to encourage creative thinking.

Grade Breakdown

Homework (200pts):

There will be a combination of homework that requires coding, book readings, and ethics reflections. Homework requiring coding will be distributed via your own private Git repositories. I will also collect them back from your repositories sometime after the due date. So long as you are *committing* and *pushing* your changes I will have the most current versions. Other homework format will not be uniform but I will post them on Moodle and/or my website for you to get access to them.

Labs and Application Exercises (100pts)

Your effort and completion of these is sufficient for full credit. We will typically complete these in class with leftovers being homework. Application exercise are very directed (to a single concept) while labs are longer and more involved. These will be collected in the same manner as homework's.

Group Project (200 pts)

This will entail multiple stages and some class time will be given for discussing projects with me and your group. Not enough to complete the project. More details on this on its own document. Part of this score is your attendance to all group meetings. Each missed meeting (without justification) will result in a 5% loss in your own final project score. The required number of meetings will be based on the project and group.

Exams (250 pts each – Total 500pts)

There will be a Midterm exam (9/30) and a final exam (morning of 10/13) in West 200. They will be a mix of in class and take home components.

Total Points: 1000

DISABILITIES AND ACCOMMODATIONS POLICY:

Cornell College makes reasonable accommodations for persons with disabilities. Students should notify the Office of Academic Support and Advising and their course instructor of any disability related accommodations within the first three days of the term for which the accommodations are required, due to the fast pace of the block format. For more [information on the documentation required to establish the need for accommodations and the process of requesting the accommodations](#).

ACADEMIC HONESTY POLICY:

Cornell College expects all members of the Cornell community to act with academic integrity. An important aspect of academic integrity is respecting the work of others. A student is expected to explicitly acknowledge ideas, claims, observations, or data of others, unless generally known. When a piece of work is submitted for credit, a student is asserting that the submission is her or his work unless there is a citation of a specific source. If there is no appropriate acknowledgment of sources, whether intended or not, this may constitute a violation of the College's requirement for honesty in academic work and may be treated as a case of academic dishonesty. The procedures regarding how the College deals with cases of academic dishonesty appear in The Catalogue, under the heading "Academic Honesty."

Current Campus Mask Requirement

The following campus mask requirements are currently in effect:

1. As a mask friendly campus, mask-wearing is accepted in all situations without question, and as a community we strive to accommodate the requests of others for each of us to also mask in meetings, offices, and at close-quarter events. All community members are expected to keep a mask with them to wear when requested by others.
2. Masks are not required in any buildings on campus. Faculty and staff retain the ability to require masks in their offices, classrooms, and other instructional spaces at their discretion, and all such requirements will be communicated through course syllabi, signage, or both.
3. Service providers on campus such as IT staff, Facilities Services staff, and athletic trainers can require that service recipients mask while in the same space with you.
4. Signage is available to communicate your masking policy for your office, room, or classroom. You can [print your own](#) or stop by the Mail and Service Center for a pre-printed sign.
5. Individuals in [isolation](#) or [close contacts](#) will mask according to CDC recommendations, which may be up to 10 days. Student Health can help you navigate this.

Mask Optional

At this time, masks are not required during in-person instructional time in this class. However, there are many good reasons to wear a mask, and many of us will choose to do so. Not all members of our community are able to be vaccinated, and masking is an effective way to protect these individuals. Regardless of any individual's decision to wear a mask, I expect that we will all be considerate of each other.

If evolving COVID-19 conditions warrant, I reserve the right to require masks during all in-person instructional time in this class, regardless of the individual immunity status of members of the class. If we do need to shift to a required mask policy, I expect that all students will comply. Non-compliance would result in a conduct report to the Dean of Students, which could result in removal from the course.

Illness Policy

If you are experiencing COVID-19 symptoms, do not attend class. Perform a home test or contact Director of Student Health Services Lynn O'Brien at student_health@cornellcollege.edu immediately to arrange a COVID-19 test at the Health Center.

If you need to isolate due to COVID-19, or if you become unable to attend class for any other health reason, contact me as soon as possible to determine if you are able to continue in the class. A [Withdrawal for Health Reasons](#) may be required.

Positive COVID Test Policy

If you test positive: Please isolate yourself in your room, contact Director of Health Services Lynn O'Brien at student_health@cornellcollege.edu and wait for instructions. A representative from Student Health will contact you to discuss your test and provide you with instructions for isolating and begin contact tracing of on-campus interactions. If you test positive for COVID-19 during the block and need to isolate, you need to inform me directly; the Health Center cannot inform me on your behalf.

Close Contact Policy

If you are named as a close contact of someone who has tested positive on campus, you will be notified by Student Health. (If you learn that you have been in contact with someone who has tested positive from an off campus exposure, you must inform Director of Health Services Lynn O'Brien at student_health@cornellcollege.edu. We will determine your status and next steps using the [CDC guidelines](#). To prepare for the possibility of needing to conduct contact tracing, we will have assigned seating and/or I will take a daily photo of the classroom to document seating proximity.

Going Online

If I must quarantine or isolate during the block, I will need to shift portions of the course online. If this happens, I will communicate with you as soon as possible via email and provide a Zoom link to our next class meeting.