DSC 223 Syllabus
Introduction to Data Science

Fall 2021 Block 4 (11/15 to 12/15)
West Science Hall 201

Class Times
M/W/Tr/F: 9am-11am, T: 9:30am – 11am
M-F: 1:00pm-3:00pm

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!

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.

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-**

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 using version control. 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 patters or trends and exploring them to in learn more. Your reasoning skills are built and tested when making decisions based on the data using your own programmed visualizations and numerical summaries. In your group work in class and group project presentations 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.

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.

Moodle

Moodle is where you will find course announcements, technology instructions, links to other course components. Check here many times each day to stay caught up in our course.

Software – No need to install

We will use a combination of technologies in this course including a program called 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!

IT Help Desk: HERE

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. 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.

DISABILITIES AND ACCOMODATIONS 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, see HERE.

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."

Grade Breakdown

Homework (200pts):
These 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.

Labs and Application Exercises (100pts)
Your 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.

Quiz's (100pts)
The quizzes will be over both the assigned readings and any content in class we have recently covered and/or practiced.

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 (200 pts each – Total 400pts)
There will be a Midterm (morning of 12/3) exam and a final exam (morning of 12/15).

Total Points: 1000

Current Covid-19 Conditions
The Delta variant of COVID-19, coupled with vaccination rates too low to achieve herd immunity, have resulted in an increase in COVID-19 cases in our area. The following course policies are a response to this fluid, evolving situation, and may be adjusted as needed during the block.

Mask Requirement
You are required to wear a mask in class. Mainly this is because during our class we have fall break. Since we will sometimes be handing each other's computers (rare but possible) please also clean and sanitize your hands often (it is flu season too).

Positive Covid Test Policy
If you test positive for Covid during the block, need to quarantine, or need to isolate, you need to inform me directly; the Health Center cannot inform me on your behalf. If someone in the class tests positive, all close contacts in the classroom will be required to take a COVID-19 test at the Health Center. To prepare for this possibility, we will have consistent seating. Non-immune close contacts (individuals who are neither vaccinated nor have a documented case of COVID-19 in the immediately preceding 10 months) will need to quarantine. Immune contacts will be exempt from quarantine unless they become symptomatic. If there is a positive COVID-19 case in the class, everyone in the class may take a COVID-19 test at the Health Center if they wish.

Illness Policy
If you are experiencing COVID-19 symptoms, do not attend class. Notify me (that you will be gone) and contact Director of Student Health Services Lynn O'Brien at lobrien@cornellcollege.edu to arrange a COVID-19 test at the Health Center.
If you need to isolate or quarantine due to COVID-19, or if you become unable to attend class for any other health reason, you will need to meet with me (virtually) as soon as possible to determine if you are able to continue in the class. A Withdrawal for Health Reasons may be required.

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.