CIS 299-2 K1 TR: Special Topics - Introduction to Data Science II
Truman College

One of the City Colleges of Chicago
Spring 2024

Tuesdays and Thursdays, 2:00 PM – 3:45 PM
Synchronous Online

February 13, 2024 – May 10, 2024

Course Prefix and Number: 67800 (CIS 299)
Course Title: Introduction to Data Science II
Length of Course: 12 weeks
Semester Credit Hours: 3.0
Contact Hours (minutes): 3.0 (2400 minutes)
Lab Hours (minutes): 0.0 (0 minutes)
Prerequisites: Eligibility for English 101, Math 205, Department Chairperson
Method of Delivery: Face to Face _____ Online ___X___ Hybrid __

Instructor Information
Dr. Amanda R. Kube Jotte
Email: akube1@ccc.edu
Office: 1345
Office Phone: 773.907.4794
Student Office Hours: T 1:00 PM – 2:00 PM

"Never be afraid to try something new. Remember amateurs built the ark, but professionals built the Titanic." – Unknown

Course Description

Data has become increasingly available in recent years, not just to scientists but to the average citizen as well. With this information at our fingertips, there is increased need for training in how to use and understand this data. In this course, we will grapple with how to use data to study the world around us. We will learn how to collect and store data, use programming languages to explore data, and study methods for making inferences about that data. In addition to learning to manipulate and test data, together we will discuss when and how to use data and whether data is always neutral. How might data exacerbate existing biases? When is the use of data appropriate? How can we be more critical consumers of the data that is constantly being presented to us? We will learn to see data in our everyday lives and work with datasets that are exciting and relevant to ourselves and our community.

The main purpose of this course is to:

1. Introduce students to the data science pipeline.
2. Develop students' abilities to be informed and critical readers of quantitative, data-based arguments.
3. Enable students to perform data analysis using Python and Jupyter Notebooks.
4. Develop students' abilities to integrate code, background information, results and interpretation to

communicate data analyses effectively.
5. Help students gain flexible problem-solving and programming skills applicable to a large variety of problems independently.
6. Understand foundational concepts of probability, statistics, and machine learning.
7. Develop skills to apply modeling techniques for data analysis.

Upon satisfactory completion of the course…

1. Students will be able to code and plot in a programming language, using relevant data science packages and software.
2. Students will be able to think critically about the use and collection of data.
3. Students will be able to clean, filter, group, and visualize datasets and use these methods to better understand and explain their data.
4. Students will gain an understanding of statistical methods including hypothesis testing, confidence intervals, and bootstrapping.
5. Students will be able to model data using basic methods of machine learning.
6. Students will be able to use concepts of statistical inference to engage with research questions.
7. Students will be informed and critical readers of quantitative data-based arguments.

Required Books and Materials

You do not need to purchase a printed textbook as I will be assigning readings from free online textbooks such as:

- "Python Data Science Handbook" by Jake VanderPlas available at https://jakevdp.github.io/PythonDataScienceHandbook/
- "Computational and Inferential Thinking" available at https://www.inferentialthinking.com/chapters/intro
- "Learning Data Science" available at https://learningds.org/intro.html

For a more advanced introduction to Python, I suggest Wes McKinney. (2017) "Python for Data Analysis", 2nd edition. O'Reilly Media.

Additional materials and required skills include the following:

- access to a personal computer and reliable internet access
- basic computer and software skills: typing, web browsing, Zoom, Brightspace
- Python and Jupyter Notebook installed on your computer and/or access to Google Colab

Links to readings, handouts, other course documents, and additional information/resources for students are available under our course on Brightspace.

Course Policies

Additional information may be found in the *Academic and Student Policy* manual: https://www.ccc.edu/menu/Documents/Academic_Student_Policy/AcademicStudentPolicy.pdf.

NSW – No-Show Withdrawal: Students enrolled in this course will be issued a no-show withdrawal (NSW) if they fail to show in face-to-face class meeting sessions, and the Brightspace website and enter the course content areas (in each course in which they are registered) at least once on two different days

within the first week. *Per CCC policy: In person classes: a student who is absent from the first two class sessions and has not contacted his/her instructor with intent to pursue the course will have his/her registration canceled by the college and will be issued a no-show withdrawal (NSW). For classes meeting only once a week, an NSW will be recorded if the student misses the first class session.*

ADW – Administrative Withdrawal: Students enrolled in this course who are not actively pursuing the course at midterm will be withdrawn from the course and issued a grade of ADW. Active pursuit should be measured by the Brightspace Course Statistics such as log in frequency, Gradebook, Discussion Board, electronic submission of assignments, and online assessments. Please refer to Student Policy Manual for more details.

Late Work & Assignment Makeup Policy
In general, the material taught in this class builds on past material. For this reason, it is extremely important that students keep up with the material so that they do not fall behind on future material as well. Any late assignment will incur at 10% penalty per day late. For example, a homework assignment that was submitted one day late could earn a maximum of 90%. If you are having an emergency that is making it difficult for you to complete you work on time, please contact me as soon as possible so we can work out a way forward together.

Classroom and Course Discussion Policies
Students who are disrespectful or offensive to the instructor or any member of the class will first be addressed by the instructor. If there are no improvements, students will be referred promptly for disciplinary action. Please consult your student policy manual on page 66 for additional details.

Academic Dishonesty:
For the first plagiarism offense, the student receives a warning and a zero (0) or F for the assignment with a follow-up discussion with the instructor. On the second plagiarism offense, the student receives an F in the course.
Please consult your student policy manual on page 66 for additional details.

Assessment Methods (Formative & Summative)

In addition to the required readings, the course will consist of lectures, class discussion, in class activities, labs, homework, projects, quizzes, and exams.

- ***In-class activities*** – coding worksheets for practice applying concepts, to be completed in class
- ***Lab activities*** – longer activities for practice thinking through a data science problem, to be completed in class
- ***Homework*** – more challenging questions to be completed outside of class, students will have a week to complete
- ***Quizzes*** – knowledge checks at the end of sections of class, to be completed on Brightspace during class
- ***Midterm Project*** – practice working from start to end on a data science question of your own, completed outside of class, students will have several weeks to complete
- ***Final Exam*** – take-home exam covering all material learned over the semester

Grade Distribution

Category	Percentage of Grade
Participation	10%
Labs	10%
Homework	20%
Quizzes	10%
Midterm Project	25%
Final Exam	25%
Total	**100%**

Grading Scale

Grade	Percentage Range	Description
A	90-100%	Excellent
B	80-89%	Good
C	70-79%	Average
D	60-69%	Minimum Passing
F	<60%	Failure

Schedule of Readings and Assignments

This schedule will likely change to suit the evolving needs of our course. A continuously updated schedule will be provided on Brightspace and announcements will be made in class to keep students aware of any changes.

Date	Topic
February 13 ❤️	Proficiency Testing
February 15	Probability Review
February 20	Data Wrangling
February 22	Exploratory Data Analysis Review
February 27	Statistical Inference
February 29	Hypothesis Testing
March 5	Hypothesis Testing
March 7	Hypothesis Testing
March 12	Confidence Intervals and Bootstrapping

March 14	Confidence Intervals and Bootstrapping
March 19	Confidence Intervals and Bootstrapping
March 21	Confidence Intervals and Bootstrapping
March 26	Spring Break 🌴
March 28	Spring Break 🌴
April 2	Linear Dependence and Correlation
April 4	Simple Linear Regression
April 9	Simple Linear Regression
April 11	Multiple Linear Regression
April 16	Multiple Linear Regression
April 18	Feature Engineering and Model Selection
April 23	Feature Engineering and Model Selection
April 25	Basics of Classification
April 30	Logistic Regression
May 2	Logistic Regression
May 7	Other Classification and Regression Methods
May 9	Final Exam 💯

Student Resources

You may also need to seek help from our support centers or technology help desk for additional support. Here is a partial list of some of Truman's other support services:

Navigate
Navigate is designed to help students take full advantage of the resources Truman has to offer in order to succeed in their coursework. It records attendance and allows teachers to issue progress reports for students. After a few weeks, you may be contacted by an advisor as a result of something I have posted in the system. As always, please see me if you have any questions or concerns.

Disability Access Center
The Disability Access Center is located in room 1435 of the main building. This center is responsible for verifying that students have a disability and for providing disability-related needs for academic accommodations, in cooperation with the students themselves and their instructors. Students who need academic accommodations should request them from the DAC Center. The director, Lauren Daley, may be reached at 773-907-4725 or ldaley@ccc.edu; office hours are from 9:00 am- 5:00 p.m. Monday through Wednesday; 9:00 am – 6:00 pm Thursday; and 9:00 am – 1:00 pm Friday.

Tutoring Services at Truman College
Great news! Truman has a variety of free tutoring and support services to help you succeed. Truman has free tutoring for all Credit and Adult Education students. All centers offer appointments that can be booked through Navigate, on Zoom, by phone, in person, or via email. For Fall 2022, all centers are planning to offer both in-person and online tutoring. Please call, email or consult our websites for the most up-to-date information.

Undocumented Liaison
Assists students looking for additional campus and/or community resources including access to financial aid and academic support to successfully matriculate to degree completion. This includes support for undocumented and mixed status students -family member is undocumented. Main Building room 2143. Josiel Marrufo / jmarrufo9@ccc.edu.

Homeless Liaison:

Provides assistance to students who lack or in risk of lacking a fixed, regular, and adequate nighttime residence or whose parent or legal guardian is unable or unwilling to provide shelter and care. Housing resources and counseling are offered in the Wellness Center. The Financial Aid Office offers assistance with financial aid eligibility including eligibility as an independent student. Main Building, room 1946· Karen Caldwell-Littleton / kcaldwell-littleton@ccc.edu.

Financial Aid assistance:
Contact Executive Director, Sean Heraty at sheraty@ccc.edu in the McKeon Building, Financial Aid Office.

Benefits Navigator:
Supports students in determining eligibility for benefit programs as well as identifies campus-wide and community resources. Main Building, room 1435. Suha Jarad / sjarad2@ccc.edu.

Other resources:

Tutoring Center	2230A	907-4785
TRiO	*162	907-4797
Wellness Center	1946	907-4045
Security	1428	907-4801
Transfer Resources Center	*118	907-4724
Admissions and Advising	*118	907-4000
Library	L625	907-4865
Reflection Room	173	

*located in the Larry McKeon building.