CIS 299-2 K1 TR: Special Topics – Introduction to Data Science
Truman College
One of the City Colleges of Chicago
Spring 2023

Tuesdays and Thursdays, 2:00 PM – 3:45 PM   February 14, 2023 – May 11, 2023
Truman College Building 1, Room L983

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

Instructor Information
Dr. Amanda Kube
Email: akube1@ccc.edu
Office: 1345
Office Phone: 773.907.4794
Student Office Hours: T and Th 3:45 PM – 4:45 PM

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 and statistics.
7. Develop skills to apply probability and statistics 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 bootstrap.
5. Students will be able to use concepts of statistical inference to engage with research questions.
6. Students will be informed and critical readers of quantitative data-based argument

Required Books and Materials

Intro to Data Science for 11800 and 11900 freely available at:
https://ds1.datascience.uchicago.edu/intro.html

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 into 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, late assignments will only be accepted due to emergency situations. Students must notify me before a deadline (where possible) that work will be late and provide an acceptable explanation. Make-

Ups for exams will only be allowed for documented emergency reasons. If you become ill, experience a family emergency, or have difficulty accessing technology, it is your responsibility to contact me as soon as possible, in writing, so that we can address them together and figure out a way forward. Documentation will be required.

Classroom and Course Discussion Board 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 chapter 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* – in class written exam covering all material learned over the semester

Grade Distribution

Category	Percentage of Grade
In-class activities	5%
Lab activities	15%
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

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.

Additional Student Support Services (New Fall 2022)

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.

Schedule of Readings and Assignments

Course Calendar (Subject to Change)			
Date	**Topic**	**Reading**	**Due**
February 14 ❤️	What is Data Science?		
February 16	Introduction to Python	Chapter 3	
February 21	Lists and Arrays	Chapter 4	Lab 1
February 23	DataFrames	Chapter 6	Homework 1
February 28	Group, Pivot, and Merge	Chapter 7	
March 2	Pandas Lab		Lab 2
March 7	Functions, Conditionals, and Iteration	Chapter 5	
March 9	Functions, Conditionals, and Iteration Lab		Lab 3 Homework 2
March 14	Visualization	Chapter 9	
March 16	Data Collection	Chapter 10	
March 21	Rules of Probability	Chapter 11	
March 23	Probability and Empirical Distributions	Chapter 12	Homework 3
March 28	Probability Lab		Lab 4
March 30	Statistical Inference		Midterm Project
April 4	SPRING BREAK 🌴		
April 6	SPRING BREAK 🌴		
April 11	Hypothesis Testing	Chapter 13	Homework 4
April 13	Hypothesis Testing Cont		
April 18	Hypothesis Testing Lab		Lab 5

April 20	Confidence Intervals and Bootstrap	Chapter 14	
April 25	Confidence Intervals and Bootstrap Cont		
April 27	Confidence Intervals and Bootstrap Lab		Lab 6
May 2	Data Ethics and Society	Chapter 15	Homework 5
May 4	Review		
May 9	Review		
May 11	Final Exam 💯		

This schedule may change to suit the evolving needs of our course. If that happens, students will be provided with a revised schedule.