

# DATA 170A Syllabus

**Instructor:** Dr. Durell Bouchard  
**Office Hours:** M-F: 1:00 PM - 2:00 PM  
**Office:** Trexler 365-C  
**E-Mail:** [bouchard@roanoke.edu](mailto:bouchard@roanoke.edu)  
**Phone:** 375-4901

## Course Objectives

---

The sheer volume of data we generate in our interactions with the digital world is staggering. [In just one second](#), users create 347,222 posts on Instagram, upload 500 hours of video to YouTube, and place orders for 6,659 packages on Amazon. To comprehend the immense quantities of information generated, we require sophisticated algorithms to simplify, analyze, and visualize the data. Additionally, the companies collecting this digital information need data scientists who possess a deep understanding of these algorithms. This course aims to introduce you to the tools and techniques essential for becoming a data scientist. You will gain proficiency in applying machine learning to vast datasets to explore data and make predictions.

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

1. Develop programs utilizing machine learning to make predictions.
2. Properly format and manage data.
3. Report and interpret the performance of prediction models.

## Course Content

---

**Prerequisites:** CPSC120

**Project:** The course will culminate in a project utilizing machine learning to create a model capable of making predictions from data. This project is designed to allow you to apply all the skills and techniques you've learned throughout the semester to explore a dataset that piques your interest.

**Assignments:** We will have regular small programming assignments that reinforce class concepts. These assignments prepare you to succeed on the project.

**Activities:** During class, programming activities will provide you with a structured experience in data cleaning and analysis. activities connect the reading and lectures to the practice of data science and prepare you for the assignments.

**Co-curricular:** The Department of Mathematics, Computer Science, and Physics is offering a series of lectures designed to engage the campus community in discussions of ongoing research, novel applications, and other issues within these disciplines. You may submit up to two papers reflecting on a talk you attend for extra credit.

**Grading:** Course grades are assigned based on the following weights and scale:

Grade Weights			
Category	Weight		
Project	20%		
Assignments	40%		
Activities	30%		

  

Grade Scale			
Grade	Range	Grade	Range
A	93-100	C	73-76
A-	90-92	C-	70-72
B+	87-89	D+	67-69
B	83-86	D	63-66
B-	80-82	D-	60-62
C+	77-79	F	0-59

## Course Policies

---

**Attendance:** If you have a temperature of 100.4 or higher or other COVID symptoms, do not come to class. Call Health Services IMMEDIATELY. For your absence to be excused, you must allow Health Services to notify me that you have consulted them about COVID symptoms. If Health Services instructs you to isolate and not attend class for multiple days, inform me so that we can devise a plan to keep you up-to-date with the course material. All absences caused by consultation with Health Services regarding coronavirus symptoms or isolation ordered by Health Services will be excused. You will still need to complete the work and graded assignments, even if we extend your deadline.

Class attendance is crucial for success in this course; the material covered during missed sessions remains the responsibility of the student. Conversations in class clarify the published class materials and are subject to evaluation in subsequent tests and quizzes. If

you anticipate being unable to attend class, please email me before class to request an excusal.

**Late Work:** Unless specified otherwise, assignments must be submitted before the start of class on the due date. If you anticipate being unable to meet a deadline, email me before the deadline to request an extension. Unexcused late work will receive no credit.

**Academic Integrity:** Collaboration is a fundamental part of learning. You are encouraged to discuss and learn from one another while working on the activities. However, collaboration on the group project requires proper attribution. Copying someone else's work or submitting someone else's work as your own is NEVER allowed. Using someone else's work or ideas without appropriate attribution constitutes plagiarism and is considered an academic integrity offense. Please familiarize yourself with Roanoke College's standards for academic integrity. If you are unsure about how the policy applies to any assignments in this course, seek clarification from me.

**Subject Tutoring:** located on the lower level of Fintel Library (Room 5), is open 4-9 PM, Sunday-Thursday. Subject Tutors are highly trained, current students who offer free, one-on-one (and small group) tutorials in over 80 courses taught at Roanoke College, including: Business, Economics, Mathematics, INQ 240, Modern Languages, Lab Sciences, and Social Sciences. Check out all available subjects and schedule 30- or 60-minute appointments at <[www.roanoke.edu/tutoring](http://www.roanoke.edu/tutoring)>. If you have a question, feel free to stop by, or contact us at [subject\\_tutoring@roanoke.edu](mailto:subject_tutoring@roanoke.edu) or 540-375-2590. See you soon!

**Accessible Education Services:** Accessible Education Services (AES) is located in the Goode-Pasfield Center for Learning and Teaching in Fintel Library. AES provides reasonable accommodations to students with documented disabilities. To register for services, students must self-identify to AES, complete the registration process, and provide current documentation of a disability along with recommendations from the qualified specialist. Please contact Dustin Persinger, Assistant Director of Academic Services for Accessible Education, at 540-375-2247 or by e-mail at [aes@roanoke.edu](mailto:aes@roanoke.edu) to schedule an appointment. If you have registered with AES in the past and would like to receive academic accommodations for this semester, please contact Dustin Persinger at your earliest convenience to schedule an appointment and/or obtain your accommodation letter for the current semester.

**Student Health and Counseling Services:** Student Health & Counseling Services supports students through in-person health appointments, in-person counseling, 24/7 telehealth (TimelyCare), Therapy Assistance Online, as well as resources related to general wellness, LGBTQ+, sexual assault, substance abuse, and suicide prevention. Unmet health needs can negatively impact your performance in this course. Student Health & Counseling Services can help. Please see <https://www.roanoke.edu/shcs> for more information and to access services.

**Diversity:** This classroom is a place where all individuals will be treated with respect, and students of all backgrounds, beliefs, ethnicities, genders, gender identities, gender

expressions, national origins, religious affiliations, sexual orientations, abilities, and other visible and non-visible differences are welcomed. All members of this class are expected to contribute to creating a respectful, welcoming, and inclusive environment for each other.

**Preferred Name/Pronoun:** I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester to make appropriate changes to my records.

## Course Schedule

---

This course expects you to spend at least 12 hours of work each week inside and outside of class.

<b>Date</b>	<b>Topic</b>
Tuesday, January 16	Python - Introduction
Thursday, January 18	Python - Testing and Functions
Tuesday, January 23	Python - Conditionals
Thursday, January 25	Python - Loops
Tuesday, January 30	Python - Strings and Arrays
Thursday, February 1	Python - Dictionaries and Objects
Tuesday, February 6	Classification - File I/O
Thursday, February 8	Classification - Perceptron
Tuesday, February 13	Classification - Normalization
Thursday, February 15	Classification - Training
Tuesday, February 20	Classification - Precision and Recall
Thursday, February 22	Vectorization - NumPy
Tuesday, February 27	Vectorization - Array Operations
Thursday, February 29	Vectorization - Array Indexing
<b>Spring Break</b>	
Tuesday, March 12	Vectorization - Array Shaping
Thursday, March 14	Vectorization - Fancy Indexing
Tuesday, March 19	Vectorization - CSV
Thursday, March 21	Regression - Pandas
Tuesday, March 26	Regression - Matplotlib
Thursday, March 28	Regression - Cleaning
Tuesday, April 2	Regression - Time Series
Thursday, April 4	Regression - Scikit-Learn
Tuesday, April 9	Regression - Cross Validation

<b>Date</b>	<b>Topic</b>
Thursday, April 11	Regression - Feature Engineering
Tuesday, April 16	Regression - Feature Selection
Thursday, April 18	Project
Monday, May 1 (10:00 AM)	Project Presentations