DATA 101: Data Science for the Humanities and Social Sciences

Spring 2026

(Last updated January 19, 2026; syllabus is subject to change)

Instructor: Professor Eren Bilen
Office: Althouse 104
Email: bilene@dickinson.edu
Phone: 717-254-8162
Office Hours: Calendly link
Monday 11:30a-1p
Wednesday 11:30a-1p
or by appointment
QRA: Phuong Hoang, hoangm@dickinson.edu
QRA Office Hours: TBA
QRA Office Hours Location: TBA

Classroom: Althouse 204
Tuesday and Thursday
10:30-11:45a

Optional Course Materials

- Data Science For All by Brennan Davis and Hunter Glanz
- Telling Stories with Data by Rohan Alexander
- Learning Statistics with Python by Ethan Weed. Online version available here.
- Course webpage: Github

Course Overview

What can data science do for you? This application-driven course invites students from the humanities, arts, and social sciences to use data as a lens for asking, and answering, disciplinary questions. Through a task-oriented introduction to statistical software, students will wrangle real-world datasets drawn from literary corpora, maps, music, sports, visual arts, and viral trends. Along the way we introduce core ideas in data science such as descriptive statistics, modeling, visualization, and ethics, while demystifying the field itself. This course equips you with practical, portable skills to address the following types of questions. How has the usage of the word "Queer" evolved? What makes hip hop lyrics different from country lyrics? How do the color palettes of the Renaissance compare to those of the French Impressionists? Did a viral Tik-Tok increase sales of Stanley tumblers? No prior coding or statistics background is assumed.

A student who successfully completes this course should be able to do the following:

- Formulate and/or analyze complex questions that can be addressed with numbers, symbols, or data.
- Evaluate and interpret relevant quantitative information to support an argument.
- Interpret and translate between multiple representations of quantitative information.

- Analyze and interpret data using visualization techniques and statistical summaries.
- Learn how models are built, trained, and used to make sense of real-world systems.

Course Policies

Attendance Policy: This course will be taught in person. Students are expected to attend all in-class meetings. While I will not take formal attendance, it is important for you to attend in class meetings and take notes. If you will be unable to attend a class meeting for any health-related issues or other emergencies, please contact me beforehand so that arrangements can be made.

Use of Laptops, Tablets, and Phones: The college has recently switched to a fully mobile computing plan. Therefore, please make sure to have a laptop computer for this course. If you need one, the Help Desk can provide you one. Here is more information. Laptops and tablets are permitted for note-taking, coding, and activities during this course and you should bring one with you to every lecture. In exchange for trusting you to use these devices, I ask that you not use them as distractions. I maintain the right to change this policy for individual students or for everyone if these tools become a problem during class. Phones are not permitted and should be put away in silent mode.

Grading: Your course grade is based on two closed-book exams, a take-home final exam, homework assignments, and participation in discussions and in-class activities.

Exam 1 (20%):	March 5
Exam 2 (20%):	April 9
Take-home Final (20%):	by May 7, 5pm ET
Assignments (20%):	Due dates TBA
Participation (20%)	

While I will not be giving extra credit in this course, I will drop your lowest assignment score. I expect there to be 10 total assignments (depending on course pacing). Occasionally, an assignment may be weighted to count as two assignments (because of the complexity or length), this will be clearly indicated when it is assigned.

The following scale will be used to determine your final grade:

Score	Letter	GPA	Score	Letter	GPA
$93 \geq x$	A	4.0	$73 \leq x < 77$	C	2.0
$90 \leq x < 93$	A-	3.7	$70 \leq x < 73$	C-	1.7
$87 \leq x < 90$	B+	3.3	$67 \leq x < 70$	D+	1.3
$83 \leq x < 87$	B	3.0	$63 \leq x < 67$	D	1.0
$80 \leq x < 83$	B-	2.7	$60 \leq x < 63$	D-	0.7
$77 \leq x < 80$	C+	2.3	$x < 60$	F	0.0

Programming: We will use Python for basic data exploration and analysis. No prior coding experience is expected. Please download and install the Anaconda distribution from here during the first week of class. Anaconda includes Python, Jupyter Notebook, and standard data science libraries. It works on Windows, macOS, and Linux operating

systems. A tablet or Chromebook is not sufficient for this class. If you need a laptop, you can loan one from the Help Desk. We will go over setting up the environment together in class. I highly recommend going to our QRA for extra support setting Anaconda up on your computer.

Teams: It is important that you get practice in team work, as data analysis benefits from having a second pair of eyes. You will be asked to complete your assignments in teams of two. You are free to choose your teammate and may switch if you wish. Note: If you are worried that your teammate is not helpful, fear not! Read about the `rubber duck`!

Datacamp: You will receive an invitation to Datacamp.com, which will grant you access to all their tutorials on data science for the duration of the semester. Use this as an opportunity to pick up the skills that might be necessary for your own projects, or to gain practice before the exams.

Make-up: If you must be away from campus on university business or due to an emergency, or you will miss a presentation day due to sickness, please reach out to me as soon as you can and provide documentation. The weight of the missed assignment will be shifted to other assignments.

Getting Help

Office Hours: Please use this `Calendly link` to schedule an office meeting so that you have a confirmed slot scheduled with me. I will be holding office hours each week. Please see the first of page of the syllabus for my hours. I am also available by appointment. If there is a conflict and you are unable to make it to any of my hours, please feel free to send me an email. My availability outside office hours is not guaranteed, however I devote my attention fully to you during my office hours. Therefore, I highly encourage you to come to my office hours and ask questions.

Quantitative Reasoning Associate: This semester, we are fortunate to have a Quantitative Reasoning Associate (QRA) working with us. A QRA is a fellow student who completed Data Analytics courses in the past and will be helping us as a course facilitator and student mentor. This semester, the QRA for our course is Phuong Hoang. Phuong is a Data senior and has completed all courses in her major which includes upper level Data classes. She will be holding office hours during the hours on the first page of the syllabus. Location is TBA. In addition, Phuong will host a study session before each exam, which will be announced closer to exams. I highly encourage reaching out to Linh or dropping by during her office hours.

Quantitative Reasoning Center & Data and Scientific Reasoning Lab: Dickinson College provides additional support for students taking courses with quantitative content across the curriculum through the Quantitative Reasoning (QR) Center. For the Spring 2026 semester, the QR Center will offer tutoring for our course, in addition to general quantitative support. You can go to dickinson.mywconline.com to make an appointment, or drop in during designated lab hours with information available here.

Other Important Information

Generative Artificial Intelligence (Gen-AI) Statement: Use of Gen-AI (such as ChatGPT) is allowed in this class, but I encourage you to use it responsibly. This means you should understand each step it takes, especially when it comes to coding or writing. Research shows that Gen-AI can help you learn more quickly, *if* you are already engaged with the material. It also shows that students with a solid coding and domain foundation benefit more from using Gen-AI, so it is important for you to invest in your own skills. My recommendation is that if you are going to use Gen-AI, treat it as a tutor, not a solver. While you are learning new concepts, it can be helpful for explanations or clarifications, but you should avoid using it to complete full assignments for you. You will not have access to Gen-AI during our two in-person tests, so you need to be sure that you can apply what you've learned on your own. Finally, while Gen-AI can serve as a useful editor, it is not good at generating original ideas. Its suggestions tend to reflect average or typical content from its training data. For creative or analytical work, your ideas should come from you, not from a statistical average.

Referencing the Work of Others: When submitting your work, you must follow commonsense ground rules. You must cite all your resources in any delivery of your work. Failure to cite your references constitutes plagiarism. If you have any questions about what constitutes plagiarism, you should reach out to me, or you may contact the writing center. Here is a nice resource that you can refer to for any questions regarding plagiarism.

Statement on Disabilities: Dickinson values diverse types of learners and is committed to ensuring that each student is afforded equitable access to participate in all learning experiences. If you have (or think you may have) a learning difference or a disability – including a mental health, medical, or physical impairment – that would hinder your access to learning or demonstrating knowledge in this class, please contact Access and Disability Services (ADS). They will confidentially explain the accommodation request process and the type of documentation that Dean and Director Marni Jones will need to determine your eligibility for reasonable accommodations. To learn more about available supports, go to www.dickinson.edu/ADS, email access@dickinson.edu, call (717) 245-1734, or go to the ADS office in Room 005 of Old West, Lower Level (aka "the OWLL").

If you have already been granted accommodations at Dickinson, please follow the guidance at www.dickinson.edu/AccessPlan for disclosing the accommodations for which you are eligible and scheduling a meeting with me as soon as possible so that we can discuss your accommodations and finalize your Access Plan. If test proctoring will be needed from ADS, remember that we will need to complete your Access Plan in time to give them at least one week's advance notice.

SOAR: Academic Success Support: Students can find a wealth of strategic guidance by going to www.dickinson.edu/SOAR. This website for SOAR (Strategies, Organization, and Achievement Resources) includes apps, tips, and other resources related to time management, study skills, memory strategies, note-taking, test-taking, and more. You will also find information aimed to helps students "SOAR Through Academic Challenges," as well as a schedule of academic success workshops offered through Academic Advising. If you would like to request one-on-one assistance with developing a strategy for a manageable and academically successful semester, email SOAR@dickinson.edu.

Other talks, events: I highly encourage you to actively attend campus talks organized by the departments of Data Analytics, Math & CS, Economics, among others. You should network with presenters, and chat about how they approached tackling on their project. Your attendance on these talks will also count towards your participation.

Course Outline: Below is a schedule for what will be covered in this course including application topics. There may be adjustments on the list during the semester depending on progress made in class. On Tuesdays, we will have an in-class discussion/lecture on the week's topic. Thursdays will consist of a application session during which we will do live-coding and you will get to have a head start on your assignment in class which will be due the following Tuesday. The exact due dates for each assignment will be announced.

- **Week 1**
  Introduction to the course. Overview of the course structure and assignments.
  Lecture topic: What is Data? Who uses it, and why?
  Application topic: Explore domains that make use of data.
  Resources: Digital Humanities, NYT: YouTube, the Great Radicalizer
- **Week 2**
  Primer on statistics. Introduction to Explarotary Data Analysis (EDA), descriptive statistics, distributions, types of data, where do data come from?
  Lecture topic: Primer on Statistics.
  Application topic: Exploring and describing real-world data
  Resources: Part IV. Statistical Theory, tidy tuesday
- **Week 3**
  Visualization of data. Histograms, bar charts, and scatterplots.
  Lecture topic: How do you visualize data and why?
  Application topic: Data visualization exercises.
  Resources: reddit.com/r/dataisbeautiful, datavizproject
- **Week 4**
  Introduction to experiments. How is evidence obtained in the social sciences? Randomized trials, natural experiments, research ethics.
  Lecture topic: Experiments.
  Application topic: Design your own experiment.
  Resources: Seeing Theory, 1, 2, 3
- **Week 5**
  Introduction to Python.
  Lecture topic: Introduction to Python: basics of coding
  Application topic: Getting practice with Python
  Resources: Part II. Introduction to Python, Part III. Working with Data
- **Week 6**
  Maps and locations.
  Lecture topic: Introduction to spatial data.
  Application topic: Creating a map of Australia
  Resources: GeoPandas
- **Week 7**
  Review for Exam I.
  Resources: Course notes, assignments, study guide posted on Github.

EXAM 1: MARCH 5

- **Week 8**
  Spring Break: no class on Tuesday and Thursday.
  Lecture topic: What is a model? What is prediction?
  Application topic: Predict the value of a home in Carlisle
  Resources: Interactive simple linear regression
- **Week 9**
  Text analysis I. Bag of words. Sentiment analysis.
  Lecture topic: What is text data? How do you analyze it?
  Application topic: Sentiment analysis on presidential speeches.
  Resources: Presidential speeches
- **Week 10**
  Text analysis II. Authorship classification.
  Lecture topic: Authorship classification: How can we predict the author of a text?
  Application topic: Determine the author of a given piece of text.
  Resources: Poetry database, Poetry assessor
- **Week 11**
  Text analysis III. Changes in language and word usage. Trends in word frequency.
  Lecture topic: How do words evolve over time? How can we quantify language trends?
  Application topic: Track and visualize the usage of words or phrases across decades using historical corpora.
  Resources: Google Books Ngram Viewer
- **Week 12**
  Review for Exam II
  Resources: Course notes, assignments, study guide posted on Github.
  EXAM 2: APRIL 9
- **Week 13**
  Data Science and music
  Lecture topic: What is data in music? What can you do with it?
  Application topic: Find two songs that are "similar" and claim why they are so from a data point of view.
- **Week 14**
  Visual data: Data science and paintings.
  Lecture topic: What is data in the context of paintings? How do we extract data from images?
  Application topic: Track an artist's use of colors across their career.
  Resources: Artworks dataset
- **Week 15**
  Lecture topic: NA- tie loose ends & summary of the semester
  FINAL PROJECT: BY MAY 7 AT 5PM