



CISC367 Introduction to Data Science

Credits: 3, Section(s): 080

Semester: Spring, Year: 2020

Meeting Days, Times, Location and Room: TR 12:30-1:45pm, Online

Syllabus Version: 0.0.1

Table of Contents

[Table of Contents](#)

[1. Instructor Information](#)

[Instructor Contact Information](#)

[Teaching Assistant\(s\) Contact Information](#)

[About the Instructor:](#)

[2. Course Description](#)

[Description](#)

[Prerequisites](#)

[Course Delivery](#)

[3. Learning Outcomes](#)

[4. Learning Resources](#)

[Optional Textbooks](#)

[Technology](#)

[5. Learning Assessment](#)

[Final Grade Breakdown](#)

[Course Component](#)

[Percentage of Total](#)

[Weekly DS Assignments](#)

[Midterm and Final Project](#)

[Participation](#)

[CORGIS Contributions](#)

[Grading Scale](#)

[Grade](#)

[Interval](#)

[6. Course Calendar](#)

[Date](#)

[Theme/Topic](#)

[Assignment](#)

[Technology](#)

[7. Course Policy Document](#)

[Course Specific Policies](#)

[Attendance](#)

[Communication](#)

[Collaboration](#)

[Late Submissions](#)

[Cheating](#)

[Privacy](#)

[Backups and Data Loss](#)

[UD Policies](#)

[Academic Integrity](#)

[Harassment and Discrimination](#)

[Faculty Statement on Disclosures of Instances of Sexual Misconduct](#)

[Accommodations for Students with Disabilities](#)

[Non-Discrimination](#)

[Change Log](#)

1. Instructor Information

Instructor Contact Information

Instructor name: Austin Cory Bart
Instructor preferred pronouns: he/him/his
Email address: acbart@udel.edu
Website: <https://www.acbart.com>
Office location: Smith 411
Office hours: TBA

Teaching Assistant(s) Contact Information

TA name: TBA
E-mail address:
Office location:
Office hours:
Phone number:
Additional information:

About the Instructor:

Dr. Bart is an Assistant Professor of Computer Science. Although his primary research area is Computer Science Education, his work has always had a Data Science component. His dissertation work promulgated the idea of “Pedagogical Datasets” to support introductory Data Science education for non-Computing majors. He has also published in Learning Analytics and Educational Data Mining conferences, helping to advance the field with new data standards and techniques. He has extensive technical experience with the Python Data Science stack (Jupyter, Pandas, Matplotlib, etc.). He is very passionate about teaching Data Science and fostering the critical thinking skills necessary.

2. Course Description

Description

Data Science is an increasingly popular, profitable, and critical discipline. In this course, students will learn essential Data Science fundamentals of data manipulation, analytics, visualization, interpretation, and presentation. Taught in a problem-oriented style, with minimal lecture, the course emphasizes practical techniques and programming pragmatics. Contexts will vary widely across the semester, without a single overall theme, and will occasionally give students a chance to explore their own interests in different application areas.

Prerequisites

- Instructor consent is required
- CISC220 Data Structures
- **Note:** Honors status is appreciated but not required

Course Delivery

This course will be administered online through Canvas with regular meetings in Zoom and Discord.

All of the following is subject to change, but in a typical week,

- On Monday, students will be given a dedicated assignment focused on that week's course content. This will include some short pre-reading or introductory video.
- On Tuesday, students will meet
 - For the first half of lecture time on Zoom, and Dr. Bart will introduce the material.
 - For the second half of lecture time in Zoom breakout rooms, working on the assignment
- On Thursday, students will meet
 - For the first half of lecture time in Zoom breakout rooms, working on the assignment.
 - For the second half of lecture time in the main Zoom room, and Dr. Bart will summarize the lessons to be learned.
- During the rest of the week, students will continue working on the assignment and take advantage of the Discord to ask further questions and collaborate with classmates.

3. Learning Outcomes

By the end of the course, students will be able to...

1. Find and obtain datasets from a variety of sources
2. Process and transform data from a variety of raw formats into a cleaned representation
3. Analyze cleaned data using both descriptive and inferential statistics
4. Create visualizations from data that accurately represent their reality
5. Interpret visualizations and analyses in order to generate conclusions
6. Present conclusions, results, and methodology in order to tell a clear data story
7. Criticize a data science methodology, results, and conclusions from both a technical and societal context

These objectives are aligned with the following topics from the ACM Body of Knowledge:

- Computational Science: CN4/Interactive Visualization, CN5/Data, Information, and Knowledge, CN6/Numerical Analysis
- Information Management: IM1/Information Management Concepts, IM2/Database Systems, IM3/Data Modeling, IM6/Query Languages, IM10/Data Mining
- Social Issues and Professional Practice: SP2/Analytical Tools, SP3/Professional Ethics, SP6/Professional Communication

4. Learning Resources

Optional Textbooks

No textbooks will be required for this course. Any required reading or viewing assignments will use only free online resources.

However, if you are interested in more resources related to the course material, the following list of texts can be considered:

- VanderPlas, J. (2016). Python data science handbook: Essential tools for working with data. " O'Reilly Media, Inc."
- Wilkinson, L. (2012). The grammar of graphics. In Handbook of Computational Statistics (pp. 375-414). Springer, Berlin, Heidelberg.
- Wickham, H. (2010). A layered grammar of graphics. Journal of Computational and Graphical Statistics, 19(1), 3-28.

- Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons.
- Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python*. O'Reilly Media.
- McIlwain, C. D. (2019). *Black Software: The Internet and Racial Justice, from the AfroNet to Black Lives Matter*. Oxford University Press, USA.
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Technology

- **Canvas:** In this class, Canvas, UD's online learning management system, will be used for all course activities and critical communication channels. All assignments will be posted through the Canvas course site unless otherwise directed.

Information on how to use Canvas is available through the [Canvas Student Guide](#).

Canvas can also be accessed via [MyUD](#).

- **Zoom:** In this class, Zoom will be used for video conferencing. You will receive a meeting invitation with a specific web link and meeting ID number for use in this class.

Zoom is a web-based application that can be used from a computer, tablet, or phone. From a computer, click the class web link and be sure to install the small program that will download to your computer. You only need to install the program once. From a mobile device, download the free Zoom app and type in the class meeting ID number.

A microphone and camera are recommended for online meetings. Information on how to test your computer's audio and video can be found on [Zoom's website](#). Learn more info about [Zoom at UD](#).

- **Discord:** Discord will be used as a real-time chat platform. Verification, participation, and monitoring of the Discord will be required. Although all critical course announcements will be posted to Canvas, the instructor will often use the Discord to provide additional help and resources in the course.

5. Learning Assessment

Final Grade Breakdown

The final course grade will be calculated using the following components:

Course Component	Percentage of Total
Weekly DS Assignment	50%
Midterm Project	10%
Final Project	25%
Participation	15%
CORGIS Contributions	Up to 10%

Weekly DS Assignments

Throughout the semester, students will be tasked with weekly assignments to collect, manipulate, analyze, visualize, interpret, present, and criticize data using Python and associated tools. For many of these assignments, they will focus on a specific Data Science skill or subskill, although some may require a synthesis of these skills in concert.

Often, the work will require coding in Python and/or creating a Jupyter Notebook. Some may be automatically assessed with unit tests, while others will be graded with a rubric.

Contexts will vary greatly across the assignments. Some may allow students to explore their own interests, while others will require students to focus on a particular application area (e.g., health, geography, sales, literature, gaming).

Midterm and Final Project

For both the midterm and final assessment periods, students will be tasked with completing a complete Data Science process end-to-end (“Full Stack Data Science”), including a presentation of their findings. This will be assessed with a rubric evaluating their performance in each of the major skills covered. The major projects will be an opportunity for students to tailor to their specific interests. Significant independent research will be expected in exploring the dataset, with the expectation that students will clearly demonstrate the critical thinking skills essential to Data Science.

Participation

Throughout the semester, students are expected to be regularly engaged with the course material and members. This engagement can come in many forms, such as by active interactions during the class lectures, communication on the course Discord, or meeting independently with classmates to collaborate on assignments. The exact nature and extent of the interactions will vary depending on their time and nature. This category will be graded subjectively by the instructor based on the evidence of student participation gathered over the course of the semester.

CORGIS Contributions

Up to 5 times throughout the semester, students will be given the opportunity to contribute a new or improved CORGIS dataset, each effectively worth up to 2% of extra course credit (max of 10% course credit). The [CORGIS Project](#) (Collection of Really Great and Interesting dataSets) is an open-source research project to create high-quality “Pedagogical Datasets” for novices learning Computational Thinking. Each dataset is individually curated to appeal to niche interests, and requires application of Data Science skills learned during the semester. By creating pedagogical datasets, students will actively demonstrate their understanding of the course material while contributing to a popular computing community resource. The quality of student-created datasets will be evaluated by the instructor using a rubric; the instructor reserves the right to reject or request revisions to a dataset.

Grading Scale

Students will be assigned the following letter grade based on the calculation coming from the course assessment section.

Grade	Interval
A	94.0 and above
A-	90.0 to <94.0
B+	87.0 to <90.0
B	84.0 to <87.0

B-	80.0 to <84.0
C+	77.0 to <80.0
C	74.0 to <77.0
C-	70.0 to <74.0
D+	67.0 to <70.0
D	64.0 to <67.0
D-	60.0 to <64.0
F	<60.0

6. Course Calendar

The following calendar should be considered preliminary.

Date	Theme/Topic	Assignment	Technology
1	Basics of Data Text-based Data Formats Basics of Data Science Basics of Git	Interpret the story of some data Leverage a Github repository	JSON, CSV Git
2	Python Crash Course Jupyter Notebooks	Recreate an existing Jupyter Notebook	Python Jupyter
3	Data Transformation Tidy Data	Reshape datasets to match a target format	Pandas
4	Basics of Visualization Grammar of Graphics	Create of visualizations	Matplotlib Seaborn
5	Practical Statistics Descriptive Statistics Inferential Statistics	Run statistical tests and interpret results	Scipy Scikit-Learn
6	Data Storytelling Ethical Reasoning Critical Peer Evaluation	Midterm Presentations Present a story using some data	YouTube
7	Dealing with Missing Data Interpolation Imputation	Process missing/bad data	<i>No new tech</i>
8	Basics of SQL Sqlite	Process data in SQLite format	Sqlite SqlAlchemy

9	Pedagogical Datasets Data Science Pedagogy	Develop/update CORGIS Datasets	<i>No new tech</i>
10	Regular Expressions Parsing	Parse text files	Regex
11	Web Scraping Parsing	Scrape websites	Requests BeautifulSoup Scrapy
12	Binary Data Formats Audio, Video, Graphics	Extract data from non-text file formats	<i>Various</i>
13-14	Final Project	Present a complex analysis by processing complex data from disparate sources	

7. Course Policy Document

Course Specific Policies

Attendance

Due to the ongoing global pandemic, all absences are immediately excused without documentation required. However, assignment due dates and lock dates are still enforced. Further, multiple absences may require makeup activities, or risk a drop in the Participation grade. Students are expected to communicate issues that may affect their participation in the course with the professor in a timely fashion, depending on the nature of their circumstances.

Absences on religious holidays listed in university calendars are also recognized as an excused absence. Nevertheless, students are urged to remind the instructor of their intention to be absent on a particular upcoming holiday. Absences on religious holidays not listed in university calendars, as well as absences due to athletic participation or other extracurricular activities in which students are official representatives of the university, shall be recognized as excused absences when the student informs the instructor in writing during the first two weeks of the semester of these planned absences for the semester.

Communication

Students should use email for:

- Discussing issues related to course participation

Students should use Discord for:

- Discussing course material

Students should regularly check Canvas Announcements for:

- Important course updates

Students should use Canvas Submission Comments for:

- Discussing the grade of a specific assignment.

In all cases, we will attempt to return communication within 1-2 business days, ideally much less.

Collaboration

Social learning has a huge impact on developing expertise, and Data Science is inherently a collaborative industry. Therefore, the course will leverage multiple opportunities for collaboration. In particular, students will be sorted into smaller learning communities (Cohorts) and additionally paired with a peer, to promote a culture of students working closely on assignments together. In fact, some assignments will explicitly require group interaction and submission.

Because collaboration is so important in computing, communicating respectfully is essential. This includes any interactions between the instructors, teaching assistants, and students. Whether in online discussion forums or in-class peer interactions, the same rules apply. Abusive behavior will not be tolerated in any form, especially: Racism, sexism, homophobia, transphobia. Students are expected to conduct themselves in an ethical, professional manner. Offering criticism should be an opportunity to help others improve, not to take them down.

Please be excellent to each other.

Late Submissions

Due dates and lock dates are strictly enforced. Submission after the lock date is not allowed without explicit instructor permission and exceptional circumstances. Submission after the due date (without prior approval) will incur a linearly increasing penalty (e.g., -10% after the first day, -20% after the second day). The instructor reserves the right to change these policies when needed to maximize student motivation and learning.

Cheating

You can assume that all assignments are collaborative unless they are specifically stated to be otherwise. However, students may not avoid learning by cooperating to split up work. All assignments completed collaboratively should be completed together, with all participants developing their own understanding and answers.

In the cases below, both partners are at fault and could be penalized:

- Splitting up a problem set into two halves - you do the first half and your friend does the second half without talking to each other.
- Your classmate emails you the solution to a quiz and you fill in the answers for your copy.
- You play on your phone while your partner works on a programming problem.

Other examples of external cheating include, but are not limited to:

- Copying the language, structure, code, ideas, and/or thoughts of someone else and presenting it as your own without proper citation
- Making up data and results
- Subverting automatic grading mechanisms or tracking software to achieve apparently incorrect results or falsify logs.
- Submitting work that was previously submitted for credit in another course.
- Submitting an instructor reference solution as your own.
- Submitting work after an assignment's lock date without instructor permission
- Hiring someone to do your work for you.
- Helping others to cheat, either intentionally or unintentionally.
- Lying, either verbally or in writing, about the above.

Privacy

You are requested to share some online identities that you may wish to keep private (e.g., Github account name, Discord username). If you are uncomfortable sharing this information, please get in touch with the instructor to discuss appropriate alternatives.

Students are strongly encouraged to have their web cameras on at all times. However, we understand that not everyone has an environment where they are comfortable having their webcams on, and that at times we simply do not want to be on video. Active Participation does not require a web camera on, but it will be considered strong evidence in favor of engagement (i.e., you can earn participation points by having your web camera on regularly). We do ask that you have your web camera on when meeting in small groups or for 1-1 meetings, although again we understand if this is not possible. However, barring exceptional and temporary technical difficulties, we do expect you to have your microphone on.

Backups and Data Loss

Make sure to keep backups of all your work! Hardware failure or accidentally losing your code is not an allowable excuse to be granted an extension. Use Git, commit often. Keep instructors informed about technological failures in a timely manner.

UD Policies

Academic Integrity

Please familiarize yourself with UD policies regarding academic dishonesty. To falsify the results of one's research, to steal the words or ideas of another, to cheat on an assignment, to re-submit the same assignment for different classes, or to allow or assist another to commit these acts corrupts the educational process. Students are expected to do their own work and neither give nor receive unauthorized assistance. Complete details of the university's academic integrity policies and procedures can be found at sites.udel.edu/studentconduct/sgup/ Office of Student Conduct, 218 Hullahen Hall, (302) 831-2117. E-mail: student-conduct@udel.edu

Harassment and Discrimination

The University of Delaware works to promote an academic and work environment that is free from all forms of discrimination, including harassment. As a member of the community, your rights, resource and responsibilities are reflected in the non-discrimination and sexual misconduct policies. Please familiarize yourself with these policies at www.udel.edu/oei . You can report any concerns to the University's Office of Equity & Inclusion, at 305 Hullahen Hall, (302) 831-8063 or you can report anonymously through UD Police (302) 831-2222 or the EthicsPoint Compliance Hotline at www1.udel.edu/compliance. You can also report any violation of UD policy on harassment, discrimination, or abuse of any person at this site: sites.udel.edu/sexualmisconduct/how-to-report/

Faculty Statement on Disclosures of Instances of Sexual Misconduct

If, at any time during this course, I happen to be made aware that a student may have been the victim of sexual misconduct (including sexual harassment, sexual violence, domestic/dating violence, or stalking), I am obligated by federal law to inform the university's Title IX Coordinator. The university needs to know information about such incidents to, not only offer resources, but to ensure a safe campus environment. The Title IX Coordinator will decide if the incident should be examined further. If such a situation is disclosed to me in class, in a paper assignment, or in office hours, I promise to protect your privacy--I will not disclose the incident to anyone but the Title IX Coordinator.

For more information on Sexual Misconduct policies, where to get help, and reporting information, please refer to www.udel.edu/sexualmisconduct. At UD, we provide 24/7/365 crisis assistance and victim advocacy and counseling. Contact 302-831-1001 to get in touch with a sexual offense support advocate, as well as confidential and anonymous counseling services for other concerns.

Accommodations for Students with Disabilities

Any student who thinks he/she may need an accommodation based on a disability should contact the Office of Disability Support Services (DSS) office as soon as possible. Students who have documentation of their need for accommodation should register via the SAM platform: andes.accessiblelearning.com/UDEL/. Reach DSS in the following ways: Phone: 302-831-4643, fax: 302-831-3261, [DSS website](#). Email: dssoffice@udel.edu or visit at 240 Academy Street, Alison Hall Suite 130.

During COVID-19, Disability Support Services staff are available remotely. Please call 302-831-4643 during business hours (8-5 M-F) or email dssoffice@udel.edu for assistance.

Non-Discrimination

The University of Delaware does not discriminate against any person on the basis of race, color, national origin, sex, gender identity or expression, sexual orientation, genetic information, marital status, disability, religion, age, veteran status or any other characteristic protected by applicable law in its employment, educational programs and activities, admissions policies, and scholarship and loan programs as required by Title IX of the Educational Amendments of 1972, the Americans with Disabilities Act of 1990, Section 504 of the Rehabilitation Act of 1973, Title VII of the Civil Rights Act of 1964, and other applicable statutes and University policies. The University of Delaware also prohibits unlawful harassment including sexual harassment and sexual violence.

For inquiries or complaints related to non-discrimination policies, please contact: Office of Equity & Inclusion-oei@udel.edu, 305 Hullahen Hall Newark, DE 19716 (302) 831-8063

For complaints related to Section 504 of the Rehabilitation Act of 1973 and/or the Americans with Disabilities Act, please contact: Office of Disability Support Services, dssoffice@udel.edu, Alison Hall, Suite 130, Newark, DE 19716 (302) 831-4643 OR contact the [U.S. Department of Education - Office for Civil Rights](#)

Change Log

The changelog presented below reflects the dynamic nature of the course. The instructor reserves the right to make changes as needed. When possible, this log will be updated in sync.

- Version 0.0.1 (10:58am at 12/23/2020): Initial version created