ECE 20875:
Python for Data Science

Fall 2021
Course Syllabus

Instructors

Section 001 (CRN: 21503)
Prof. Christopher G. Brinton (cgb@purdue.edu)
Lectures: M/W/F 11:30am-12:20pm in WALC 1018
Office Hours: M/W 3:30pm-4:30pm in MSEE 342

Section 002 (CRN: 23070)
Prof. Qiang Qiu (qqiu@purdue.edu)
Office: MSEE 358
Lectures: Mon/Wed/Fri, 1:30pm - 2:20pm ET in EE 170
Office hours: Tues and Thurs 1-2pm (virtual unless special arrangements made)

Section 004 (CRN: 27067)
Prof. Mahsa Ghasemi (mahsa@purdue.edu)
Office: MSEE 238
Lectures: Mon/Wed/Fri, 1:30pm – 2:20pm in WTHR 320
Office Hours: Wed 2:30pm-4:30pm (virtual unless special arrangements made)

See pinned post on Piazza for video and office hour links.

Graduate TAs

Somosmita Mitra [Lead] (mitra26@purdue.edu)
Nadir Mohamedraf Alawadi (nalawadi@purdue.edu)
Laura M Cruz (lcruzcas@purdue.edu)
Jhanvi Saraswat (jsaraswa@purdue.edu)

Undergraduate TAs

Joseph Bushagour [Lead] (jbushago@purdue.edu)
Julia Taylor [Lead] (taylo773@purdue.edu)
Minjun Zhang [Lead] (zhan3624@purdue.edu)
Jude Adham (jadham@purdue.edu)
Jhen-Ruei Chen (chen2774@purdue.edu)
Noah Criswell (ncriswe@purdue.edu)
Vaishakh Deshpande (deshpav@purdue.edu)
Hsuan-Chen Fang (fang299@purdue.edu)
Yiming Fu (fu287@purdue.edu)
Alex Gieson (agieson@purdue.edu)

Ishaan Jain (jain343@purdue.edu)
Ruibin Jiang (jiang551@purdue.edu)
Pranav Karthikeyan (pkarthik@purdue.edu)
Robert Ketler (rketler@purdue.edu)
Shuihan Liu (liu2244@purdue.edu)
Wang-Ning Lo (lo67@purdue.edu)
Maximilian Manzhosov (mmanzhos@purdue.edu)
Marvin Mui (mui1@purdue.edu)
Jan-Adriel Nacpil (nacpil@purdue.edu)
Kahaan Patel (pate1410@purdue.edu)
Kartik Pattaswamy (kpattasw@purdue.edu)
Ayush Praharaj (aprahar@purdue.edu)
Runjia Shen (shen403@purdue.edu)
Siddharth Srinivasan (sriniv75@purdue.edu)
Muhammad Sulaim (msulaim@purdue.edu)
Param Talwalkar (ptalwalk@purdue.edu)
Avik Wadhwa (wadhwa13@purdue.edu)
Runlin Wang (wang4229@purdue.edu)
Henry Lee Wong (wong313@purdue.edu)
Bo-Yang Wu (wu1563@purdue.edu)]

Communication via Piazza
We will use Piazza as the primary form of announcements and communication between the teaching staff and the students. **If you have any questions about the course, homeworks, or project, you should post them on Piazza**. It is a shared discussion forum, where your question can be answered by the instructors, the TAs or your fellow students!

All students are expected to join the Piazza classroom within the first week of class and preferably as soon as possible.

Find our Piazza page here:
piazza.com/purdue/fall2021/ece20875

Class Website
To maintain consistency across the sections, we will manage a separate website for the course. This website will have links to the course material, assignments, and other information. Brightspace will be used *solely* to enter and track your grades. All other matters for the course will be administered through Piazza and the website.

Find our course website here:
https://www.cbrinton.net/ECE20875-2021-Fall.html

Sections and Lectures
There are three sections of the course: 001, 002, and 004. All three sections of the course will be managed very similarly. They will have the same homeworks, the same discussion forum, the same announcements, and, as much as possible, the same lectures. The primary exception to this is the lecture component described next.

Each professor will give lectures for their section in person during the scheduled class times and locations (see above). We may **optionally** record the lectures and post them online afterwards for asynchronous viewing. As much as possible, you are cautioned not to use the recordings as a replacement for the in-person lectures, for a few reasons:

- There are several obvious advantages of being in-person, e.g., you can ask questions and interact with the instructor in real time, and it provides structure in the day.
- It is possible that minor portions of the material covered live (e.g., using the whiteboard) will not be captured well on the recording.

That being said, due to the ongoing COVID complications, we want to be as sensitive as possible to those who are really uncomfortable attending in person. Therefore, attendance will not factor into your grade (see Attendance Policy).

See pinned post on Piazza for links to the recorded lectures for each section.

Attendance Policy
While attendance will not be a factor in the final grades, students are expected to attend all classes in-person (or watch recorded lectures in a timely fashion) unless they are ill or otherwise unable to attend class. If they feel ill, have any symptoms associated with COVID-19, or suspect they have been exposed to the virus, students should stay home and contact the Protect Purdue Health Center (496-INFO).

In the current context of COVID-19, attendance will not be a factor in the final grades. Students need to inform the instructor of any conflict that can be anticipated and will affect the timely submission of an assignment or the ability to take an exam.

Classroom engagement is extremely important and associated with your overall success in the course. The importance and value of course engagement and ways in which you can engage with the course content even if you are in quarantine or isolation, will be discussed at the beginning of the semester. Student survey data from Fall 2020 emphasized students' views of in-person course opportunities as critical to their learning, engagement with faculty/TAs, and ability to interact with peers.

Only the instructor can excuse a student from a course requirement or responsibility. When conflicts can be anticipated, such as for many University-sponsored activities and religious observations, the student should inform the instructor of the situation as far in advance as possible. For unanticipated or emergency conflicts, when advance notification to an instructor is not possible, the student should contact the instructor/instructional team as soon as possible by email, through Brightspace, or by phone. In cases of bereavement, quarantine, or isolation, the student or the student's representative should contact the Office of the Dean of Students via email or phone at 765-494-1747. Our course Brightspace includes a link to the Dean of Students under 'Campus Resources.'

TA Lab Hours

We are fortunate to have many wonderful graduate and undergraduate TAs on the teaching staff this semester. Many of the TAs have taken the class before and/or previously served as a TA. With their help, we are able to offer a substantial amount of TA lab hours to help answer any questions you have. Please take advantage of these sessions to get one-on-one assistance, particularly for help on the homework assignments.

Due to COVID-19, the primarily delivery of these lab hours will be online using the Discord platform. Additionally, we will have some limited availability for in-person lab hours in EE 206 and 207. The protocols for these are given below. We will offer the following schedule:

- *Virtual lab hours (Discord)*: Monday through Friday 9-11am ET, 1-2pm ET, and 5-8pm ET
- *Limited in-person lab hours (EE 207 and EE 206)*: Tuesday and Thursday 5-8pm ET (in EE 207), Friday 5-8pm ET (in EE 206&207)

***Protocol for using Discord*:**

- During lab hours, each TA will join one of the open voice channels in Discord. Students who have questions can join these voice channels to talk with a TA one-on-one and share screens if needed.
- After discussion, the student should leave the TA voice channel and go back to the general channel to allow other students to talk to the available TAs.
- Further instructions on Discord can be found on the class website.
- **Discord link:** https://discord.gg/E9YTfzA3hx**.**

***Disclaimers for limited in-person office hours*:**

- EE 206 and 207 are small rooms, so there is only room for 10-15 students in person at a time. Due to COVID-19 protocols, these limits must be strictly enforced. This means Tuesdays and Thursdays have a limit of 10-15 students (in EE 207), while Friday it is 20-30 students (in EE 206&207).
- As a result, unless you have a strong preference for in-person office hours, we suggest you take advantage of Discord. Additionally, if you are waiting for a long time to get into EE 206/207, you may be better off trying Discord instead.

- Discord will also have a substantially larger number of TAs, so it is likely you can get assistance faster.

Instructor Virtual Office Hours

The instructors will host their office hours virtually by default (times given on page 1). If you want to meet with your instructor in-person, please schedule a time via email.

During office hours, we will expect that you have been watching the lectures regularly and reviewing the notes before asking questions, i.e., these venues must be used as a supplement, not an alternate, to the lectures. Additionally, we will expect that you have put some effort into understanding the material before coming to office hours. We will give preference to conceptual or clarification questions over debugging questions because debugging questions could be answered during the daily TA lab hours.

Email

Technical questions should be posted to Piazza, raised during TA lab hours, or raised during instructor office hours (primarily questions on the lecture). This allows other students who might have similar questions to benefit from the answers. As a result, the professor and TAs will not typically answer programming questions via email. Of course, if you have questions of a personal or confidential nature, or want to set up a time to meet with your instructor in-person, we welcome your email.

Course Outcomes

A student who successfully fulfills the course requirements will have demonstrated:

1. An understanding of regular expressions [1, 2]
2. An ability to use Python to write data analyses [1, 2, 6]
3. An ability to explain when particular data analyses are appropriate [1, 2, 3, 6]
4. An ability to explain the results of data analyses [2, 3, 5]
5. An ability to incorporate classes in their Python programs [1, 2, 6]
6. An ability to incorporate associative arrays in their programs [1, 2, 6]
7. An ability to work with a partner to choose appropriate analyses to solve a problem, perform those analyses, and interpret the results of those analyses [1, 2, 3, 5, 6]

These outcomes are extremely high level. In more detail, after taking this course you will be able to:

- Write programs in Python that incorporate: basic control structures; functions; data structures such as lists, tuples and associative arrays; classes and objects; higher order functions; and regular expressions.
- Perform basic scripting tasks in Bash and Python (e.g., redirecting the output of one program to the input of another, setting and reading environment variables).
- Understand the basics of sampling, estimation, and hypothesis testing.

- Write data analyses that perform: visualization; textual analysis; regression; classification; and clustering.
- Perform more complex analyses using approaches like neural networks.

Course Assessment: The achievement of the course objectives will be through programming assignments (covering outcomes 2, 4, 5, and 6), exams (covering outcomes 1, 3, and 4), and a project (covering outcome 7).

Course Grading

Grades will be assigned as follows:

50% — Three exams (all equally weighted)
40% — 10-12 programming assignments (all equally weighted)
10% — Mini-project

The students who are the most active (and helpful) in answering questions on Piazza may receive bonus points.

Programming Assignments

Programming assignments will be due every 1-2 weeks, except for the weeks of the exams and during the final two weeks (which is set aside for the mini-project). They will test the concepts covered in class, both programming and statistical. The following rules apply to *all* programming assignments:

1. All programs should run correctly in the versions of Python available on the Scholar cluster. You will have a Scholar account created for you at the beginning of the semester.
2. Some assignments will include a separate writeup as well as a code submission that will produce the required output when run.
3. Unless otherwise specified, assignments are due at 11:59 PM on the deadline.

Programming assignments will be submitted via GitHub Classroom (https://classroom.github.com). As such, *you are required to have a GitHub account.* These can be obtained for free at https://github.com.

Please fill out the following form *as soon as the semester starts* to provide your GitHub account information:
https://forms.gle/d1boSm5ufCmn9pNx9

Without your GitHub username, we cannot put your grades on Brightspace! We use the grades released on Brightspace to calculate your final scores.

Assignment late submission policy: Except for medical and family emergencies (accompanied by verification), there will be *no individual extensions* granted for programming assignments or the mini-project. Late submissions will be scaled according to lateness, docking 10% from your score per day late, up to a maximum of 50%. Submissions more than 5 days late will be assigned a score of 0.

Exams

The first two exams will take place roughly during the 6th and 11th weeks of the semester, and the third exam will take place during finals period. Exams will be administered in-person, in the evening. We will be cancelling lectures to compensate for exam periods. The exact date, time, and location of each exam will be provided on the course website and Piazza when it is available.

Exams are equally weighted and are not cumulative.

Missing an exam will result in a grade of 0 for that exam. Exam reschedules will only be granted if you demonstrate that it is absolutely necessary (accompanied by verification) and obtain written approval from one of the instructors in advance of the exam.

Regrade Requests

If you believe that we have made an error in grading on a homework assignment or exam, please submit a regrade request to ece20875.purdue.university@gmail.com that includes:

1. Relevant identifying information. For homeworks, a link to your Github repository is needed.
2. Your reasoning or explanation for a regrade, with specific details and justification.

All such requests must be made within one week of when the grade was posted on Brightspace. Regrade requests arriving later than this will not be considered.

Please note that during the regrading process, we reserve the right to regrade all portions of your submission.

Tentative Course Schedule

Below is a rough schedule of when and what will be covered in class:

Week	Topics covered and assignments/exams
1 8/23	Introduction Python basics (variables, control structures, functions) Git basics

Week	Topics covered and assignments/exams
2 8/30	Python data structures (lists, tuples, associative arrays) Histograms Probability distributions **Homework 1 Due**
3 9/6	**9/6: Labor Day (No classes)** Probability distributions (continued) Random variables **Homework 2 Due**
4 9/13	Filters and map/reduce Higher-order functions Sampling and estimation **Homework 3 Due**
5 9/20	Sampling and estimation (continued) Hypothesis testing Confidence intervals **Homework 4 Due**
6 9/27	Hypothesis testing (continued) Confidence intervals (continued) **Exam 1 (date and location TBD)**
7 10/4	Regular expressions Basic text processing Bash scripting **Homework 5 Due**
8 10/11	**10/11-12: Fall Break (No classes)** Introduction to supervised learning Matrix algebra **Homework 6 Due**
9 10/18	Regression and least squares Regularization Cross validation
10 10/25	Regression (continued) n-grams Intro to natural language processing **Homework 7 Due**
11 11/1	Objects and Classes Introduction to unsupervised learning **Homework 8 Due**
12 11/8	Clustering: K-Means **Exam 2 (date and location TBD)** Clustering: GMMs
13 11/15	Inheritance Classification: Naive Bayes Classification: kNN

Week	Topics covered and assignments/exams
	Homework 9 Due
14 11/22	**No class: 11/22 makeup for Exam #2, 11/24&26 Thanksgiving Break**
15 11/29	Classification: Logistic regression Intro to deep learning Neural network architectures **Homework 10 Due**
16 12/6	Backpropagation Deep learning for visual recognition Conclusion **Mini-Project Due**
Finals 12/13-12/18	**Exam 3 (date and location TBD)**

Academic Honesty

Unless expressly allowed, you are expected to complete all assignments by yourself. However, you are allowed to discuss general issues with other students (programming techniques, clearing up confusion about requirements, etc.). You may discuss particular algorithmic issues on Piazza (but do not post or copy code!). If there is any doubt, please email the instructors.

We will be using software designed to catch plagiarism in programming assignments and copying on exams. **A student is considered in violation of the academic honesty policy regardless of whether they are the one "copying" or the one "being copied from."**

Punishments for academic dishonesty are severe, including receiving an F in the course or being expelled from the University. By departmental rules, all instances of cheating will be reported to the Dean of Students. On the first instance of cheating, students will receive a 0 on the assignment or exam; the second instance of cheating will result in a failure of the course.

Classroom Guidance Regarding Protect Purdue

The Protect Purdue Plan, which includes the Protect Purdue Pledge, is campus policy and as such all members of the Purdue community must comply with the required health and safety guidelines. Required behaviors in this class include: staying home and contacting the Protect Purdue Health Center (496-INFO) if you feel ill or know you have been exposed to the virus, properly wearing a mask in classrooms and campus building,

at all times (e.g., mask covers nose and mouth, no eating/drinking in the classroom), disinfecting desk/workspace prior to and after use, maintaining appropriate social distancing with peers and instructors (including when entering/exiting classrooms), refraining from moving furniture, avoiding shared use of personal items, maintaining robust hygiene (e.g., handwashing, disposal of tissues) prior to, during and after class, and following all safety directions from the instructor.

Students who are not engaging in these behaviors (e.g., wearing a mask) will be offered the opportunity to comply. If non-compliance continues, possible results include instructors asking the student to leave class and instructors dismissing the whole class. Students who do not comply with the required health behaviors are violating the University Code of Conduct and will be reported to the Dean of Students Office with sanctions ranging from educational requirements to dismissal from the university.

Any student who has substantial reason to believe that another person in a campus room (e.g., classroom) is threatening the safety of others by not complying (e.g., not wearing a mask) may leave the room without consequence. The student is encouraged to report the behavior to and discuss next steps with their instructor. Students also have the option of reporting the behavior to the Office of the Student Rights and Responsibilities. See also Purdue University Bill of Student Rights.

If you feel ill, have any symptoms associated with COVID-19, or suspect you have been exposed to the virus, you should stay home and contact the Protect Purdue Health Center (496-INFO). In the current context of COVID-19, in-person attendance will not be a factor in the final grades, but the student still needs to inform the instructor of any conflict that can be anticipated and will affect the submission of an assignment or the ability to take an exam. When conflicts can be anticipated, such as for many University-sponsored activities and religious observations, the student should inform the instructor of the situation as far in advance as possible. For unanticipated or emergency conflict, when advance notification to an instructor is not possible, the student should contact the instructor as soon as possible by email. When the student is unable to make direct contact with the instructor and is unable to leave word with the instructor's department because of circumstances beyond the student's control, and in cases of bereavement, quarantine, or isolation, the student or the student's representative should contact the Office of the Dean of Students via email or phone at 765-494-1747. Our course Brightspace includes a link on Attendance and Grief Absence policies under the University Policies menu.

Campus Interruptions

In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances beyond the instructor's control. In such an event, information will be provided through the course website, Piazza, and email.

Nondiscrimination Statement
Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life. More details are available on our course Brightspace table of contents, under University Policies.

Accessibility
Purdue University strives to make learning experiences as accessible as possible. If you anticipate or experience physical or academic barriers based on disability, you are welcome to let me know so that we can discuss options. You are also encouraged to contact the Disability Resource Center at: drc@purdue.edu or by phone: 765-494-1247. More details are available on our course Brightspace under Accessibility Information.

Mental Health Statement
If you find yourself beginning to feel some stress, anxiety and/or feeling slightly overwhelmed, try WellTrack. Sign in and find information and tools at your fingertips, available to you at any time.

If you need support and information about options and resources, please contact or see the Office of the Dean of Students. Call 765-494-1747. Hours of operation are M-F, 8 am-5 pm.

If you find yourself struggling to find a healthy balance between academics, social life, stress, etc. sign up for free one-on-one virtual or in-person sessions with a Purdue Wellness Coach at RecWell. Student coaches can help you navigate through barriers and challenges toward your goals throughout the semester. Sign up is completely free and can be done on BoilerConnect. If you have any questions, please contact Purdue Wellness at evans240@purdue.edu.

If you're struggling and need mental health services: Purdue University is committed to advancing the mental health and well-being of its students. If you or someone you know is feeling overwhelmed, depressed, and/or in need of mental health support, services are available. For help, such individuals should contact Counseling and Psychological Services (CAPS) at 765-494-6995 during and after hours, on weekends and holidays, or by going to the CAPS office of the second floor of the Purdue University Student Health Center (PUSH) during business hours.

Disclaimer

This syllabus is tentative and is subject to change. Any substantive changes will be announced in class and on Piazza.