

Stats 21 Course Syllabus

Python and Other Technologies for Data Science

Miles Chen, PhD

2023 Fall Quarter

Contents

Course Information	2
Summary	2
Learning Outcomes	2
Prerequisites	2
Lecture Times	2
Office Hours Schedule	2
Exam Date	2
Tentative Outline of topics	3
Websites	3
Office Hours	3
Software	3
Books	3
Grading	4
Grade Breakdown	4
Letter Grade cutoffs	4
Lecture Viewing Quizzes	4
DataCamp Homework	4
Homework	5
Campuswire Grading:	7
Final Exam	7
Make up policy for Exam	7
Course Policies	8
Copyright Policy	8
Academic Integrity	8
Gender Discrimination, Sexual Violence, and Harassment	8
Mental Health and Student Well-Being	8
Accommodations for students with disabilities	9
Course Grade Changes	9
Additional Information	9
Undergraduate Mentoring	9
Requests for Letters of Recommendation	9

Course Information

Summary

This class covers the use of Python and other technologies for data analysis and data science. The course will focus on programming with Python and a selection of its libraries: NumPy, pandas, matplotlib, and scikit-learn for the purpose of data processing, data cleaning, data analysis, and machine learning. Other technologies covered will include Jupyter notebook and Git. Intended for Data Theory majors as an introduction to the Python language and libraries most frequently used in data science.

Learning Outcomes

Upon completion of this course, students

- will have created Jupyter Notebooks and will have published them to a Github repository
- will be able to read and write intermediate code in Python
- will have imported, manipulated, and summarized data in Python for exploratory data analysis
- will have created data visualizations in Python

Prerequisites

Stats 20

Lecture Times

- Lecture 1: MWF 10-10:50AM
- Location: Franz 2258A

Office Hours Schedule

Office Hours schedule subject to change.

- Mondays 11:30AM - 12:30PM (in-person and online)
- Tuesdays 4-5PM (online only)
- Wednesdays 11:30AM - 12:30PM (in-person and online)
- Zoom link:
 - <https://ucla.zoom.us/j/98887074323?pwd=bjFqd2ZCNnRqQmMvU3dnYkVad0RqQT09>
 - Meeting ID: 988 8707 4323
 - Passcode: 916280
 - Please **turn on your cameras**
- Physical Office Location: Math Sciences Building 8105 J (8th floor)

Also general office hours (Open to students in any class and to former students):

- Saturdays 10AM-11AM.
 - <https://ucla.zoom.us/j/96806536111>
 - Meeting ID: 968 0653 6111
 - Passcode: 023194

No office hours on holidays:

- No office hours November 11, 2023 (Veterans Day)
- No office hours November 23 - November 25, 2022 (Thanksgiving)
- No office hours December 10, 2023 - January 6, 2024

Exam Date

- Final Exam will be held during finals week at the times scheduled by registrar's office.
- At the time of this writing, date of Final Exam still needs to be scheduled by registrar's office.
- Final exam will be online / remotely. Final exam will require computer, internet connection, Zoom, camera.

Tentative Outline of topics

- 1) Week 1: Git, GitHub, conda, Jupyter
- 2) Week 2: Python Basics: variables, expressions, statements
- 3) Week 3: Python Functions: encapsulation, recursion, return values, iteration
- 4) Week 4: Python Data types: strings, lists, dictionaries, tuples
- 5) Week 5: Python OOP: Classes and objects, methods, inheritance
- 6) Week 6: Pythonic code: List comprehensions, kwargs; Numpy
- 7) Week 7: Pandas: Importing, reshaping, and cleaning data
- 8) Week 8: Pandas: Wrangling, and Aggregation
- 9) Week 9: Data Visualization
- 10) Week 10: Fitting models with scikit-learn

Websites

Announcements and assignments will be posted via Canvas / Bruin Learn <https://bruinlearn.ucla.edu> Notes will be uploaded to GitHub Repository: <https://github.com/smileschen/2023-fa-stats21>

Homework assignments will be submitted via Gradescope. Students will be invited and will need to create an account. <https://www.gradescope.com>

Course discussions will take place on Campuswire. Students will be invited and will need to create an account.

Office Hours

Office hours are my preferred method of contact.

Questions and issues are generally resolved much more quickly via office hours.

When you come to office hours, please **introduce yourself**. Say “Hi Miles, I’m Joe Bruin.” Do this **every** time you visit me until I start calling you by your name.

I like when students come to office hours with questions about material. I love to explain things and to help students understand.

I like when students come to office hours to tell me more about themselves and to seek counsel about classes to take or next steps. I am happy to make accommodations for students who face difficult circumstances and may need extensions for assignment deadlines. Please do not hesitate to visit office hours.

I am happy to correct grading mistakes. I do not want to get in arguments with students over points. I do not like arguing whether a particular mistake should be a 5 or 10 point deduction.

Software

Students will need to install git: <https://git-scm.com/downloads>

In this course, we will use Anaconda Python, a popular data science platform. Anaconda can be downloaded and installed for free from: <https://www.anaconda.com/products/distribution>

Books

Think Python, 2nd Edition

- By: Allen B. Downey
- Publisher: Green Tea Press and O’Reilly Media, Inc.
- Print ISBN-13: 978-1-4919-3936-9

The textbook is available free to all at the publisher’s website: <https://greenteapress.com/wp/think-python-2e/>

The print version is also available for purchase at any retailer of your choice. If you decide to purchase the text, you must get the second edition.

This is an excellent text that emphasizes brevity. Please be sure to do your reading. Every sentence is valuable.

Python Data Science Handbook

- By: Jake VanderPlas
- Publisher: O'Reilly Media, Inc.
- Print ISBN-13: 978-1098121211

UCLA students can read the textbook in their web browser at the following site:

<https://learning.oreilly.com/library/view/python-data-science/9781098121211/>

Select “University of California, Los Angeles” and choose “Sign In with my Institution”

The print version is also available for purchase at any retailer of your choice.

Grading

Grade Breakdown

- 15% Lecture Viewing Quizzes
- 20% DataCamp Homework
- 36% Homework Assignments
- 25% Final Exam
- 4% Campuswire Participation

Letter Grade cutoffs

Letter grades are assigned on a straight scale as follows:

- 59.9 and below: F
- 60.0 - 69.9: D
- 70.0 - 76.9: C, 77.0 - 79.9: C+
- 80.0 - 82.9: B-, 83.0 - 86.9: B, 87.0 - 89.9: B+
- 90.0 - 92.9: A-, 93.0 and up A, top 5% of students: A+

I do not curve grades. I do not assign C-, D+ or D-. The top 5% of students will get an A+. An A+ is the only grade assignment that is assigned based on performance relative to peers.

If you are a Stats major or minor it is important that you take the class for letter grade. Taking the grade as pass/no-pass will not fulfill the requirements for completing the major or minor.

Lecture Viewing Quizzes

Each lecture will have a corresponding lecture viewing quiz. The lecture viewing quiz will be multiple choice questions with choices A, B, C, D, E. During each lecture, I will provide the answers to the viewing quiz audibly at arbitrary times in the lecture. Students will not find the answers to the quizzes in the lecture notes.

It is academic dishonesty to share the answers to the viewing quiz or to gain the answers to the viewing quiz by a method that does not involve watching the lecture.

Quizzes close before the next lecture. Students who do not complete the quiz before it closes will receive a 0 on the quiz. Make sure you complete the lecture viewing quiz after each lecture.

I will drop the one lowest viewing quiz grade. (e.g. If you forget to complete the quiz one day and get a zero, it will be dropped.) Do not message me if you forget to take the quiz or if you make a mistake entering your view quiz answers.

DataCamp Homework

Some homework will be assigned via DataCamp.

Students are required to sign up for their free subscription to DataCamp. Students must make sure their DataCamp account has the first and last name as it appears in BruinLearn so that we can transfer the scores from DataCamp to BruinLearn.

DataCamp only has three possible statuses for homework assignments:

- completed on time
- completed late
- not completed

We will record grades for the homework assignment 7 days after its due date. Assignments completed on time will receive 100 points. Assignments completed late will receive 80 points. Assignments not completed by the time we record the score in the gradebook will receive 0 points.

Please make sure you complete the assignment in DataCamp well ahead of its deadline to ensure you receive full credit for the assignment.

Unfortunately, the rigidity of the DataCamp assignment interface does not allow me to grant extensions.

Homework

There will be several homework assignments (5 or 6). Each homework assignment is worth the same amount. Each homework assignment will be scored out of a total of 100 points. None of the homework scores will be dropped. It is your responsibility to verify that your homework assignment successfully uploaded by the deadline.

All homework assignments will be posted as an ipynb file.

You will download the ipynb file and modify it with your answers.

After entering all of your answers, select “Kernel” > “Restart and Run All”. Check to make sure all requested answers are clearly visible.

Then students will submit the PDF file to Gradescope. Do not submit the ipynb file.

Homework submission requirements

The graders have to go through many students’ submissions. The grader will not spend more than 5-10 minutes on each submission. The grader does not have the luxury of trying to run code or debug code, or even see what individual lines of code can do.

The grader will look for the printed results, so it is **very important** that students properly print the requested results in their documents.

It is your responsibility to read through the rendered output and to make sure that every problem and task has the required output before submitting your assignment.

Example of bad submission

Problem 1: Use a loop to produce a list of the first 10 numbers in the Fibonacci sequence

```
x = [1, 1]
for i in range(1, 9):
    x.append(x[i-1] + x[i])
```

The grader has no idea what was produced. You will receive 0 points for this submission.

Example of good submission

Problem 1: Use a loop to produce a list of the first 10 numbers in the Fibonacci sequence

```
x = [1, 1]
for i in range(1, 9):
    x.append(x[i-1] + x[i])
print(x)

[1, 1, 2, 3, 5, 8, 13, 21, 34, 55]
```

This is the exact same code, but adds the line `print(x)`. This small change makes a big difference.

It is your responsibility to read through the rendered output and to make sure that every problem and task has the required output before submitting your assignment.

Gradescope submission requirements Submissions to Gradescope must be a PDF with letter-sized pages.

Improperly formatted submissions will lose points.

Students must properly tag the pages of their homework assignments on Gradescope. If the output for a particular problem spans multiple pages, all relevant pages must be tagged for that portion. Improperly tagged submissions will lose points.

Late Policy for Homework

Homework deadlines are based on local time in Los Angeles, California. If you are located in a different time zone, you are still expected to submit the assignments by the deadline in California. Please be aware of what time these deadlines are for you.

There is a 10 minute grace period. Submissions up to 10 minutes and 59 seconds late will be accepted with no penalty.

Homework assignments submitted 11 minutes late or more will be accepted with penalty. There is a deduction of 5 points for being less than one hour late. An additional 5-point deduction will be taken for each additional hour it is late.

Deductions are suspended between the hours of 12 am and 5:59 pm the following day. The penalties resume at 6 pm.

The maximum late penalty is 60 points.

So if an assignment is due at 6pm, a homework assignment submitted on Gradescope at

- 6:10:59 pm will receive no penalty.
- 6:11:00 pm will receive 5-point penalty.
- 6:59:59 pm will receive 5-point penalty.
- 7:00:00 pm will receive 10-point penalty.
- 7:59:59 pm will receive 10-point penalty.
- 8:00:00 pm will receive 15-point penalty.
- 11:59:59 pm will receive 30-point penalty.
- 12:00:01 am (following day) will receive 30-point penalty.
- 5:59:59 pm (following day) will receive 30-point penalty.
- 6:00:00 pm (following day) will receive 35-point penalty.
- 10:59:59 pm (following day) will receive 55-point penalty.
- 11:00:00 pm (following day) will receive 60-point penalty.
- Anything later (but less than 7 days late) will be accepted with a 60-point penalty.

I strongly advise uploading the homework to Gradescope well in advance of the deadline in case there are connectivity problems or server issues.

After 7 calendar days, late homework assignments will not be accepted. Exceptions beyond the 7-day submission window will be granted in only extreme circumstances and will only be granted via an office hours meeting (Zoom or in-person) with the professor.

Homework extensions

I will grant a 72-hour reprieve from the late penalty for homework submissions if the homework is accompanied with documentation. If you need a homework extension but do not have documentation, you can contact the professor to request an extension. The professor's reply granting approval of an extension can then be used as documentation for an extension.

The documentation must be included as part of the homework submission PDF.

Gradescope accepts only one file per submission for homework assignments. Your homework PDF should include the documentation showing the **date** and reason for your inability to submit the homework on time.

Do not email the professor or your TA with your documentation.

Extensions beyond 72 hours will be granted in only extreme circumstances and will only be granted via an office hours meeting (Zoom or in-person) with the professor.

Homework Regrades

Homework assignments are graded by designated graders (graduate students in the Statistics Department). Occasionally a grader may make a mistake in grading your submission.

Both the TA and the professor have the authority to change the grade of a homework assignment.

All requests for regrading a homework assignment must be submitted via Gradescope.

Do not email the TA or the professor for a request to regrade an assignment.

Please only submit a regrade request if you are certain that a mistake was made in the grading of your homework.

During a regrade request, the TA or Professor may discover and deduct points for mistakes that were not found during the initial grading. It is possible your regrade request results in a lower grade for the assignment.

Campuswire Grading:

Participation on Campuswire accounts for 4 percentage points of your final grade. I will download Campuswire's reputation report on Week 10 Friday and will enter grades as follows:

- Students who remain at noob level (Level 0) will get 0 points.
- Students who achieve starter level (Level 1) will get 1 point.
- Students who achieve starter level (Level 1) AND earn 50 reputation points will get 2 points.
- Students who achieve intermediate level (Level 2 - Red bird) AND 100 reputation points will get 3 points.
- Students who achieve intermediate level (Level 2 - Red bird) AND 150 reputation points will get 4 points (full credit).
- Students who achieve advanced level (Level 3 - Eagle) will get 5 points (full credit plus 1 extra credit point)

Students can check their reputation level in the "Grades" section on Campuswire.

Low quality posts may be removed at the instructor's discretion and will not count towards a student's reputation score.

It is important that when you create your Campuswire account, you enter your first and last name as they appear on Canvas / Bruin Learn.

Final Exam

There will be a final exam covering programming concepts in Python.

Make up policy for Exam

If you know in advance that you will miss an exam, contact me and I can schedule you for a make-up exam.

If an emergency arises and you feel you will not be able to perform well on the exam, let me know right away with documentation and I can schedule you for a make-up. Please do not take the exam, perform poorly, and then request a make-up exam. Once you take the exam, I cannot allow you to retake it. If you have an emergency, I will be flexible and will schedule a make-up exam.

If you miss an exam, you will need to provide documentation for your reason for missing the exam.

Make-ups of the exam must be taken within 7 days of the actual exam or a zero will be recorded, unless an exception is made by the professor. (For example, a missed exam originally given on Wednesday must be made up by the following Wednesday by 5pm.)

Course Policies

Copyright Policy

All course materials posted on the course website or distributed in class (including but not limited to lecture slides, homework assignments, lecture videos, quizzes, exams) are intended for personal use only by students enrolled in Stats 21. **It is a violation of course policy to post, share, or distribute any course material electronically or physically without permission from the instructor, even after the course is over.**

Academic Integrity

As a student and member of the University community, you are here to get an education and are, therefore, expected to demonstrate integrity in your academic endeavors. All students must uphold University of California Standards of Student Conduct as administered by the Office of the Dean of Students. Students are subject to disciplinary action for several types of misconduct, including but not limited to: cheating, multiple submissions, plagiarism, prohibited collaboration, facilitating academic dishonesty, or knowingly furnishing false information.

In addition, each student is the sole owner of their own code and work and must NOT:

- Submit work that is not original.
- Publish code or solutions online.
- Post the course questions on forums other than the designated course discussion forum (Campuswire). This means students cannot post questions on places like Stack Overflow or other similar places.
- Submit someone else's work or a modification of that work, with or without that person's knowledge.
- Allow someone else to submit their work or a modification of their work.
- Contract course work out to others.
- Plan or execute with another student some form of cheating during an exam.
- Make use of unauthorized material during an exam.

For more information about academic integrity and student conduct code, please go to <http://www.deanofstudents.ucla.edu/Student-Conduct>

Gender Discrimination, Sexual Violence, and Harassment

Title IX prohibits gender discrimination, including sexual harassment, domestic and dating violence, sexual assault, and stalking.

<http://www.sexualharassment.ucla.edu/>

<http://www.sexualviolence.ucla.edu/>

The University of California is committed to creating and maintaining a community dedicated to the advancement, application and transmission of knowledge and creative endeavors through academic excellence, where all individuals who participate in University programs and activities can work and learn together in an atmosphere free of harassment, exploitation, or intimidation.

If you, or someone you know, have been a victim of sexual harassment or sexual violence, please seek help.

Any type of harassment of any other student will not be tolerated. Please be aware that your TA and Lecturer are legally compelled to report any information in regards to sexual harassment or sexual violence to the Title IX officer.

Confidential support and advocacy is available for students who have experienced sexual harassment or sexual violence at the CARE Advocacy Office for Sexual and Gender-Based Violence, John Wooden Center West 1st Floor, (310) 206-2465, CAREadvocate@careprogram.ucla.edu. Sexual violence or sexual harassment can also be reported directly to the University's Title IX Coordinator, 2241 Murphy Hall, titleix@conet.ucla.edu, (310) 206-3417.

Reach a confidential UCLA advocate, 24 hours a day, 7 days a week by calling the CAPS 24/7 Crisis Hotline (310) 825-0768

Mental Health and Student Well-Being

Counseling and Psychological Services (CAPS) is a multidisciplinary student mental health center for the UCLA campus. CAPS supports the academic and student development missions of the University

and the Division of Student Affairs. We're here to support your mental health needs as you pursue your academic goals. Our services are designed to foster the development of healthy well-being necessary for success in a complex global environment.

<http://www.counseling.ucla.edu>

Attending and completing university coursework is difficult and can become burdensome both physically and psychologically. Take care of your mental well-being and seek support from some of the many excellent resources offered at CAPS.

Accommodations for students with disabilities

Students needing academic accommodations based on a disability must contact the Center for Accessible Education (CAE) at (310) 825-1501 or present in person at Murphy Hall A255. As the professionals delegated authority from the campus to determine reasonable disability accommodations, CAE will assess all requested accommodations and communicate appropriately with faculty. In the event that a student has approval for proctoring arrangements during exams, please inform your respective professors and/or Teaching Assistant before date of exams. When possible, students should contact the CAE within the first two weeks of the term as reasonable notice is needed to coordinate accommodations.

For more information visit <http://www.cae.ucla.edu>

Course Grade Changes

After course grades have been submitted to the registrar, grades are final. Grade changes will only be considered if there has been a clerical or procedural mistake. Students have one quarter to make requests for a grade change. Graded exams and other materials will be kept for one quarter. After one quarter, course grade changes will not be made.

Additional Information

Undergraduate Mentoring

I am a mentor for undergraduates for the UCLA Department of Statistics. This means that you may visit me during office hours (or by appointment) for an informal meeting where you can ask questions unrelated to course content, such as research opportunities, graduate studies, career paths, or any other topic pertinent to your education. You are welcome to visit and meet even after the course ends.

Requests for Letters of Recommendation

I am generally only willing to write letters of recommendation for students who I know and recognize. If you think you might request a letter of recommendation in the future, please make an effort to make yourself known to me.

I am more likely to recognize students who keep their cameras turned on and who actively participate.

I am also more likely to recognize students who make several visits to office hours during the quarter.

For a full guideline, please see this page on my personal website: <https://smileschen.com/lor.html>

Copyright Miles Chen. Do not post or distribute without permission.