PSYC 3803/8803 Syllabus

Welcome to PSYC 3803/8803:
Introduction to Programming in Python and R

Lectures (https://gatech.instructure.com/courses/547270/pages/lectures-slides)	(https:/ (https:/ (https:/

Course Instructor

Eunbee Kim

PhD in Quantitative Psychology
School of Psychology

Lecture Time: MW 9:30-10:45AM

Lecture Location: Coon 248

Email: **eunbee.kim@gatech.edu (mailto:eunbee.kim@gatech.edu)**

Office: Coon G53C

Office Hours: W 11AM-12 PM

Course Description & Objectives

This course introduces students to programming in Python and R, emphasizing their applications in data science, statistical analysis, and data visualizations. Students will gain practical experience through hands-on coding, real-world projects, and comparative analysis of the two languages.

By the end of this course, students will be able to:

- Understand and apply core programming concepts in both Python and R.
- Perform data manipulation and analysis using appropriate libraries in each language.
- Apply statistical models in Python and R.
- Visualize data effectively using Python and R.
- Select the appropriate language for specific data analysis tasks.
- Integrate Python in R and operate both languages interchangeably.

Prerequisites and/or Co-Requisites

A basic understanding of Linear Algebra. No programming experience is required for this class (in fact, it is best if you have none). This course is not recommended for students who have credit for CS 1371 or higher.

- No prior programming experience required.
- Basic understanding of statistics is recommended.

Textbook and Other Materials

- Online resources (tutorials, documentation) for Python and R
  - [main textbook] **Spciy-Lectures-textbook.pdf (https://gatech.instructure.com/courses/547270/files/73941201?wrap=1)**
  - **https://lectures.scientific-python.org/ (free) (https://lectures.scientific-python.org/)**
  - **http://openbookproject.net/thinkcs/python/english3e/index.html (http://openbookproject.net/thinkcs/python/english3e/index.html)** (free)
- *[Optional] Python for Data Analysis* by Wes McKinney ($10-44)
- *[Optional] R for Data Science* by Hadley Wickham and Garrett Grolemund ($12-50)

[Optional] Additional/supplementary materials for statistics

- **https://onlinestatbook.com/index.html (https://onlinestatbook.com/index.html)**
- **https://open.umn.edu/opentextbooks/textbooks/559 (https://open.umn.edu/opentextbooks/textbooks/559)**
- https://ethanweed.github.io/pythonbook/landingpage.html

Assessment of Learning

Lecture Attendance	100 points (14.3%)
In-Class Activity & Quizzes	100 points (14.3%)

Assignments/Homework	100 points (14.3%)
Ed Discussion	100 points (14.3%)
Midterm Project Report	100 points (14.3%)
Final Project Presentation / Final Report	200 points (28.6%)
Total Points	**700 points (100.1%)**
	For each section, the earned points will be proportionally converted based on the rubric's weighting. For instance, if you earn 95 out of 100 possible scores for the final project presentation, this score will be scaled to 190 out of 200 points according to the rubric.

LECTURE ATTENDANCE: From past experience teaching various courses, we have found a strong statistical relationship between class attendance and overall course performance. To that end, attendance will be counted as a small part of the course grade to help students perform better overall.

Beginning on the second week of class, attendance will be taken at all lecture sessions. To receive points, students are expected to arrive on time and stay for the entire class period. Arriving more than five minutes late, leaving class early, or acting in a disruptive manner during class will forfeit the points. Students must attend the lecture for which they are officially registered to earn attendance points.

This attendance policy is designed to promote class participation; therefore, no makeups will be permitted under any circumstances, and no absences will be classified as "excused" without documentations. However, to accommodate valid reasons for missing class, **students can still earn full attendance points even with up to two absences without documentation**.

Attendance will be strictly monitored, and points for a session will be annulled for the entire class if the number of respondents exceeds the total number of students present on that day, so please do not have another student sign in for you, and do not sign in for another student.

Class disruptions of ANY kind will NOT be tolerated and may result in your removal from the classroom and/or loss of participation points for that day. Please show courtesy to your fellow classmates and instructor or teaching assistant by adhering to the following class rules:

• Come to class on time and stay for the entire class period.

• Refrain from conversing with your fellow students.

• Put away any reading materials, cellular phones, and other electronic devices unrelated to the course.

IN-CLASS ACTIVITIES/QUIZZES/PARTICIPATION: You are expected to come prepared and actively participate in the class sessions. **In the event of an absence, you are responsible for all missed materials, quizzes, assignments, and any additional announcements or schedule changes given in class**.

Attendance is expected for this course. To further encourage active learning and ensure your understanding of the material, a series of unannounced in-class activities will be administered throughout the semester. Some activities will be completed individually, whereas others will be completed in a group. Activities will be given at different points during class time (beginning, middle, or end), so it is important that you arrive to class on time and stay for the entire duration of the class. This should encourage you to attend class on time, keep up with your readings and assignments, and engage in class discussion.

All the Quizzes/Exams are closed-notes and closed-books; cheat sheets are NOT allowed.The work on all exams must be your OWN work, that is, they are not to be completed with th e help or aid of others or outside materials. See **http://www.catalog.gatech.edu/policies/honor-code/Links to an external site. (http://www.catalog.gatech.edu/policies/honor-code/%22%20%5Ct%20%22_blank)** or **http://www.catalog.gatech.edu/rules/18/Links to an external site. (http://www.catalog.gatech.edu/rules/18/%22%20%5Ct%20%22_blank)** for information on Georgia Tech's Academic Honor Code.

Make-Up Assessments (Quizzes/Exams/Presentations): Make-up assessments may be arranged only with an Institute-Approved Absence or an official letter from the Dean of Students. Documentation must be provided prior to the scheduled quiz/exam date and before arranging the make-up exam.

ASSIGNMENTS: Credit for assignments will be based on your individual contribution to the in-class activity and the accuracy of your responses.

Missed assignments can only be made up with an Institute Approved Absence or an official letter from the Dean of Students and must be completed within 7 days of the original in-class activity day. No exceptions will be made to this policy. It is your responsibility to contact the instructor via email to make up missed activities.

ED DISCUSSION: While in-class participation will be encouraged, this course also includes an asynchronous discussion component via Ed Discussion. This platform is designed to provide fast and efficient support from classmates, the TA, and myself.

To earn full credit for the Ed Discussion portion of your grade, you must average at least one post per week, either by asking a question or answering a classmate's question. A total of 12 posts is required by Nov 30 to receive full credit. Posts can include programming questions, psychological concepts, problem-solving inquiries, approaches to assignments, or any other course-related topics. Simple posts that ask about logistics (e.g., "When is the exam?", "Do we have a class tomorrow?") rather than course content generally wouldn't count towards the requirement unless they stimulate further discussion on course-related topics.

You are encouraged to actively engage by answering each other's questions; however, course instructors will also participate, and particularly insightful or notable posts may be highlighted during lectures or labs for further discussion.

Most of your Ed Discussion posts should be made to the full class. The exception is if you have a question about a particular assignment, quiz, or anything else that could spoil answers for classmates. In these cases, you should post to instructors only. Please post these messages rather than email or DM them, as we may determine they are OK to make public (with your permission) if we think it would be helpful to others.

MIDTERM PROJECT REPORT: requires students to conduct a reproducible data analysis using an existing dataset (no new data collection). Students may work individually or in groups of up to three. The primary goal is to draw clear, well-supported conclusions using techniques covered in class. Analyses should be simple, focused, and interpretable, and students may use additional methods only if they fully understand and can clearly explain them. The final submission must be an HTML file exported from Jupyter Notebook (approximately 500–1500 words, excluding code and figures) and include: a clearly stated research question/hypothesis, dataset justification, literature review, analysis strategy, results write-up, and an appendix containing code, dataset link, and reproducibility steps.

The project is graded out of 100 points based on clarity of hypothesis, dataset justification, literature engagement, exploratory data analysis (EDA), presentation of results, code documentation, and reproducibility. Undergraduates must cite at least three scholarly sources in APA format and demonstrate clear understanding of prior research. Graduate students must cite at least five sources

and provide deeper synthesis, critical evaluation, and stronger methodological reasoning. The detailed grading rubric is available in the Projects/Assignments tab on Canvas.

FINAL PRESENTATION / FINAL PROJET REPORT : The Final Project is a continuation of the midterm project and must be fully reproducible. Students are expected to apply techniques covered in class and draw clear, well-supported conclusions from the data. Analyses should remain focused, interpretable, and logically structured. All work must be submitted as a knitted HTML document generated from an R Markdown (.Rmd) file, with Python integrated via the reticulate package. The primary emphasis of grading will be on code functionality, clarity, organization, and reproducibility.

The project includes both a class presentation (100 points) and a final written report (100 points). Presentations will be 15 minutes (individual) or 20 minutes (group), followed by a 5-minute Q&A session. Students must clearly present research context, hypotheses, methodology, exploratory data analysis (EDA), hypothesis testing results, conclusions, and broader discussion. The final report must include a literature review (minimum three peer-reviewed sources for undergraduates; five for graduate students), clearly stated hypotheses, documented data preparation steps, EDA with numerical summaries and visualizations, hypothesis testing with appropriate statistical justification, conclusions, and a general discussion addressing implications and limitations.

Expectations differ by academic level. Undergraduate students must present well-organized results with appropriate numerical summaries, visualizations, labels, and clear interpretation aligned with the research question. Graduate students are held to higher methodological standards, including deeper synthesis of literature, more comprehensive numerical summaries (e.g., effect sizes, diagnostics), explicit justification of analytic decisions, discussion of assumptions and robustness, and publication-quality visualizations with technically precise labeling. Graduate work should demonstrate stronger conceptual integration, critical evaluation, and methodological rigor throughout. The detailed grading rubric is available in the Projects/Assignments tab on Canvas.

Grading Scale:

Grade	%	Score Range
A	90% or above	630 – 700 points

B	80 – 90%	560 – 629.999 points
C	70 – 80%	490 – 559.999 points
D	60 – 70%	420 – 489.999 points
F	Below 60%	419.999 points or fewer

This grading scale **already accounts for rounding** and any applicable curve. No further rounding or adjustments will be made on an individual basis.

**Please refrain from contacting your instructors to request additional rounding or adjustments to the curve.

Course Policies

All work for this class is to be done individually. You are strongly urged to familiarize yourselves with the **GT Student Honor Code (http://osi.gatech.edu/content/honor-code)** rules. Specifically, the following is not allowed:

- Copying, with or without modification, someone else's work when this work is not meant to be publicly accessible *(e.g., a classmate's program or solution).*
- Submission of material that is wholly or substantially identical to that created or published by another person or persons, without adequate credit notations indicating authorship *(plagiarism).*
- Putting your projects on public Github. Otherwise, if a student *(in the future)* copies your codes/projects, the student obviously violates the honor code but you will also be implicated.

Academic Integrity

Students are expected to uphold the highest standards of academic integrity. Any form of cheating, plagiarism, or dishonesty will not be tolerated and may result in disciplinary action.

Zero Tolerance Policy on Cheating and AI Assistance (e.g., Chat GPT, Gemini)

We maintain a strict zero-tolerance policy regarding academic dishonesty, including the use of ChatGPT and other AI tools. Any student found using AI to complete assignments/quizzes/exams will be reported immediately, receive a grade of zero for the submission, and risk a final grade of F.

Disability Accommodations

If you require any accommodation due to a disability, please inform the instructor at the beginning of the course to ensure that appropriate arrangements can be made.

Preliminary Course Schedule (Subject to Change)

Week 1: Introduction to Programming and Course Overview

- Overview of course objectives and expectations
- Introduction to programming concepts (variables, data types, control structures)
- Setting up Python and R environments

Week 2: Python Basics

- Python syntax and data structures (lists, tuples, dictionaries)
- Control flow: loops and conditionals

Week 3 - 5: Data Manipulation in Python

- Introduction to NumPy Arrays
- Data cleaning and transformation techniques

Week 6 - 7: Statistical Analysis and Data Visualization in Python

- Introduction to Pandas and Seaborn
- Creating plots and visualizations

Week 8: Break

Week 9: Introduction to R

- R syntax and basic data structures (vectors, data frames, lists)
- Control flow in R: if statements and loops

Week 10: Data Manipulation in R

- Introduction to dplyr and tidyr
- Data cleaning and transformation techniques

Week 11 - 12: Statistical Analysis & Data Visualization in R

- Introduction to statistical concepts (hypothesis testing, regression)
- Performing statistical analyses in R
- Introduction to ggplot2
- Creating plots and visualizations

Week 13 - 14: Integrating Python and R

- Using R and Python together in data analysis
- Tools for interoperability (e.g., reticulate)
- Strengths and weaknesses of each language
- Use cases for Python and R in data science

Week 15 - 16: Final Project Presentation & Report

- Guidelines for the final project
- Student presentations of final projects
- Peer feedback and discussion
- Work time for project development and instructor feedback

Catch-up week (if applicable): selection between Simulation and Advanced Data Visualization Techniques

- Introduction to simulations for data analyses
- Advanced plotting techniques in R (interactive visualizations)