

Advanced Academic Programs
Zanvyl Krieger School of Arts and Sciences
Johns Hopkins University

Course and Instructor Information

Course Number

470.708: Unleashing Open Data with Python

Semester and Course Duration

Spring 2022, 15 weeks

Instructor

Collin Paschall, PhD/JD

301-358-2639

cpaschall@jhu.edu

Office Hours

Virtual by appointment only

Credit Hours

3

Class Times

Online via Blackboard. This course is 15 weeks and follows a weekly cycle of instruction. Each week begins on a Monday and ends on the following Sunday night.

Course Location

Online via Blackboard

Course Prerequisites

470.681 Probability and Statistics and 470.768, Programming and Data Management

Course Description

This course is designed to guide you as you add Python to your statistical programming toolkit. Python is a very popular language in data science circles, and in some sense is viewed as a competitor to R. So, a complete data science toolkit requires familiarity with Python. Many of the concepts in Python are similar to what you have already encountered in Programming and Data Management when you were learning to use R. Python has data structures and libraries that are similar in many ways to what you find in R. The difference is largely in the syntax and useability. R is exceptionally easy to pick up as a scripting language, while Python is a bit more difficult to learn and use.

Program Learning Objectives

When you successfully complete the program requirements, you would achieve the following goals:

P1: Apply theories of social science research to meaningful data challenges.

P2: Analyze quantitative data using an array of methodological approaches, such as descriptive statistics, causal inference, visualization, machine learning, geospatial analysis and predictive modeling.

P4: Develop an original data analysis that answers an empirical research question related to politics, policy or governance; present that analysis in writing using appropriate visuals.

Course Learning Objectives

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

1. Manage types, values, variables, and classes in Python (P2)
2. Import, export, and clean data with Python (P2)
3. Use command structures and mathematical operations in Python (P2)
4. Plot and visualize data with Python (P2)
5. Access open data with webscraping tools using Python (P2)
6. Conduct a statistical analysis using social science data with Python (P1, P2)
7. Present statistical results in an effective and accessible manner (P2, P4)
8. Identify applications of and resources for open data in social science and policy contexts (P1)

Required Texts and Other Materials

Textbooks

Please note that textbook requirements may vary by section. These texts are available for viewing free through the JHU Library.

- *Python for Data Analysis, 2nd edition*, McKinney
- *Practical Statistics for Data Scientists, 2nd Edition*, Bruce, Bruce and Gedeck
- *Mining Social Media: Finding Stories in Internet Data*, Lam Thuy Vo
- *Introducing Python: Modern Computing in Simple Packages, 2nd Edition*, Lubanovic
- *Python Projects for Beginners: A Ten-Week Bootcamp Approach to Python Programming 1st ed. Edition*, Milliken

Specific Technology Requirements & Skills for this Course

You will need a modern Mac OS or Windows PC for this course. We will go over how to install Python. (Of course, you can use Python with Linux, but if you are already doing that, you probably should not be in this class.)

Additional information can be found at the following links concerning Python's [Privacy Policy](#) and [statement on Accessibility](#).

Learning online requires some basic knowledge of computer technology. At a minimum, you need to be able to:

- Navigate in and use Blackboard; the Blackboard Student Orientation course on your “My Institution” page
- Create and save MS Word documents; review [MS Word training and tutorials](#) for PC users (all versions); [Word Help](#) for Mac users
- Find basic resources on the Internet
- Create and organize files & folders on your computer
- Send, receive, and manage email

Evaluation and Grading Policy

Below is a brief description of the course requirements. These may vary by section and are subject to change.

Course Requirements	Evaluation
Discussion Board: There are discussion board activities throughout the semester	15 points
Coding Assignments: There are 11 coding assignments. Each is worth 2 points. You may skip one assignment.	20 points
Project Proposal: This is your proposal for your course project.	10 points
Final Project Presentation: You will record a short presentation summarizing the findings of your project	5 points
Final Project: This is the written report for your final project.	200 points

Grading Scale and Criteria

This course will follow the [Advanced Academic Programs Grading Policies](#). The grading scale for students enrolled for credit is A, A-, B+, B, B-, C and F.

A	=	94-100%
A-	=	90-93.9%
B+	=	88-89.9%
B	=	84-87.9%
B-	=	80-83.9%
C	=	70-79.9%
F	=	below 70%

How the Final Paper Is Evaluated

The final project in this class is a short research note between 2,500 and 3,500 words in length, including all notes, references, tables, and figures. You will post your data and replication code to Github as an online appendix, along with uploading your code file to Blackboard for record keeping purposes.

More details are provided on the course site, but the research note is a short empirical analysis using Python. You will state a research question, provide a short literature review, test an empirical hypothesis, and discuss the implications of your findings.

An excellent research note has the following characteristics:

- Readily identifiable research question that makes an understandable claim.
- Synthetic, objective, efficient review of previous research that provides an overview of key background information and scholarship. A well-crafted theory section/literature review gives the reader the context required to understand how the paper relates to a broader scholarly or policy conversation. It also sets up a hypothesis that you can test with statistical methods. In research notes, aim for a literature review/theory section of about 1,000-1,500 words, drawing on about 10 peer reviewed sources.
- A clearly stated hypothesis and relates independent and dependent variables.
- Appropriately detailed description of data used in the analysis.
- Selection of a statistical methodology that tests the hypothesis.
- Statistical results that are presented in an aesthetically pleasing format, in tables and/or figures.
- A discussion of results that accurately interprets the analysis and discloses any limitations of the analysis.
- Conclusion that summarizes key findings, identifies the original contribution of the paper, and suggests future avenues of inquiry.
- In general, excellent written composition skills, including skillful organization and transitions.
- Appropriate conventions and citation format.
- Carefully written replication file that includes comments and all necessary documentation to generate project results using Python.

Your letter grade is determined, converted to a percentage grade as follows, then multiplied by 200 total possible points to determine the number of points towards your final course grade.

A+ (100%) grade: A stellar project of superior quality. The instructor could see a lightly revised version of this as an academic conference paper. Note that while it is possible to earn an A+ on this individual assignment, the JHU grade scale for a course has a maximum of A.

A (97%) grade: The note displays a high degree of skill and care in preparation. While there may be ways these projects could be improved against some hypothetical “gold” standard, these notes show great effort and/or outstanding creativity in analysis and presentation. Revisions would serve to improve an already strong project, not to remedy any significant deficiency in the document.

A- (92%) grade: The note aligns with assignment guidelines but would benefit from some degree of revision before it would be an exemplar of the assignment objectives. There are some identifiable but readily correctable shortcomings with this work. Minor substantive, structural, methodological, or stylistic revisions would meaningfully strengthen the final product. On the whole, the quality of the work suggests that the student could follow up on these revisions successfully without further guidance.

B+ (89%) grade: The note falls significantly short on at least one guideline of the assignment or represents a marginal or incomplete effort in multiple dimensions. Significant revision is needed. At least one section of the paper (introduction, literature review, analysis, or conclusion) needs

major reworking, or, alternatively, there is a significant flaw that runs throughout all sections of the document. Very roughly, this submission would require an amount of revision corresponding to approximately 25% of the total effort involved in producing the document.

B (86%) or *B-* (82%) grade: This is deficient on multiple dimensions, similar in nature to the problems with a *B+* submission but to a greater or much more extensive degree. Several features of the note could merit this assessment. The quality of writing could be inconsistent with graduate level expectations. The analysis could be superficial, poorly executed, or deceptive. There may be no cognizable central claim of the paper. The review of previous literature could be incomplete or overlook important context. This submission requires major, wholesale revision or reworking.

C (75%) or *F* (<70%) grade: These grades are reserved for submissions that indicate disengagement with the course, frustration of the course purposes, or a need for remediation.

How Course Engagement Is Calculated

The most important component of your grade in this course is the final submission. However, the course is designed to foster classroom engagement. Recognizing this, excellent or poor course engagement can affect your grade, to a degree, as specified below.

In an online asynchronous course, engagement consists of active participation in course discussions, presentations, and peer review. In this coding class, engagement will include short coding assignments designed to make sure you are keeping up with the course skills.

For course engagement activities, your submissions are assessed using the following procedure, with point values as specified in individual assignments posted on the discussion board.

- Full credit = 100%
- Submitted but below expectations = 92%
- Submitted but *de minimis* effort = 82%
- Not submitted = 0%

Assignment Submission

The instructor uses the SafeAssign tool for written assignments. Please review the JHU Ethics Statement below prior to submission.

Assignment Feedback

The instructor will aim to return assignments to you within 5-7 days following the due date, depending on the length of the assignment.

Late Policy

The instructor will accept late work without prior arrangement in the case of extenuating circumstances (such as hospitalization, childbirth, major accident, injury or bereavement).

Students who suffer such a circumstance must notify the instructor as soon as possible of the extenuating circumstance that prevented them from submitting work on time and determine a deadline with the instructor for submitting the work.

If you know you are going to miss a deadline, please contact the instructor before the due date. Late work submitted more than seven days after the due date, without explanation, will not earn better than an A- or 92%.

Additional Course Logistics

Synchronous Sessions

The instructor may hold live, synchronous sessions in Zoom. Attendance for synchronous sessions, while not required, is highly recommended. If you cannot attend a synchronous session, you will be responsible for watching the recording at a later time.

Time Management Expectations

It is expected that you look ahead to schedule your time. Plan to complete coursework across several days of the week rather than all in one day. Be sure to consider how group activities impact your schedule as well.

Some assignments require that you work on them for multiple weeks. Be sure to review the assignment directions at the beginning of the course so that you can plan your time accordingly. Please seek help before becoming frustrated and spending a significant amount of time to resolve an issue.

Directions for Students

Next Steps: Carefully review the remaining sections of the syllabus before beginning the first week's activities, which are located in the **Lessons** area of the navigation menu in your online course.

Once you feel that you are ready to dive into the first week's activities, select **Lessons** on the navigation menu. Then, select **Week 1** to begin.

Course Schedule

This course is 15 weeks in length and includes individual and group activities in a weekly cycle of instruction. Each week begins on a Monday and ends on the following Sunday. Please review the course syllabus thoroughly to learn about specific course outcomes and requirements. Be sure to refer to the Checklist each week, which provides a week-at-a-glance and shows targeted dates for the completion of activities.

Course Participation

You are expected to log into Blackboard regularly throughout the week - a daily check-in is recommended. It is your responsibility to read all announcements and discussion postings within your assigned forums. You should revisit the discussion multiple times over the week to contribute to the dialogue.

Online Etiquette

In this course, online discussion will primarily take place in our online Voice Thread discussion board. In all Voice Thread online communication, it is important to follow proper rules of online

etiquette - communicating with others in a proper and respectful way. For helpful tips, please see these [Ground Rules for Online Discussions](#).

Course Protocols and Getting Help

Amendments to the Course

Changes to the course will be posted in the Announcements section of your course. Please check announcements every time that you log into your online course.

Course Communication

You should communicate often with your classmates and the instructor. The majority of communication will take place within the Discussion forums. When you have a question about an assignment or a question about the course, please contact your instructor, or post your question in the course's "Syllabus & Assignment Question" forum.

Email Communication

For questions regarding course activities and assignments that would be general interest to other students, please post those in the Discussion forum. If you have a question regarding course activities and assignments of a personal nature, please send an email message to the instructor and observe the following guidelines:

- Include the title of the course in the subject field (e.g., JHU Insert Name of Course).
- Keep messages concise, and check spelling and grammar.
- Sign your full name (the sender's email is not always obvious).

Feel free to contact your instructor with comments, questions, and concerns. All email messages will be sent to you via your JHU email account, so you should be in the habit of checking that account every day or you should ensure that your JHU email account forwards messages to another account of your choice.

Email messages will be responded to within 24-48 hours.

University Policies

General

This course adheres to all University policies described in the academic catalog. Please pay close attention to the following policies.

Academic Conduct

All JHU students assume an obligation to conduct themselves in a manner appropriate to the Johns Hopkins University's mission as an institution of higher education and with accepted standards of ethical and professional conduct. Students must demonstrate personal integrity and honesty at all times in completing classroom assignments and examinations, in carrying out their fieldwork or other applied learning activities, and in their interactions with others. Students are obligated to refrain from acts they know or, under the circumstances, have reason to know will impair their integrity or the integrity of the University. Students and faculty in Advanced Academic Programs are required to adhere to the academic integrity guidelines and process laid out in the Graduate Academic Misconduct Policy. Refer to the website for more information regarding the academic misconduct policy.

If you use Safe Assign or Turnitin or any other plagiarism detection tool add the following: Please note that student work may be submitted to an online plagiarism detection tool at the discretion of the course instructor. If student work is deemed plagiarized, the course instructor will follow the policy and procedures governing academic misconduct.

Ethics & Plagiarism

JHU Ethics Statement: The strength of the university depends on academic and personal integrity. In this course, you must be honest and truthful. Ethical violations include cheating on exams, plagiarism, reuse of assignments, improper use of the Internet and electronic devices, unauthorized collaboration, alteration of graded assignments, forgery and falsification, lying, facilitating academic dishonesty, and unfair competition. Report any violations you witness to the instructor.

Read and adhere to JHU's Notice on Plagiarism.

Copyright Policy

All course materials are the property of JHU and are to be used for the students' individual academic purpose only. Any dissemination, copying, reproducing, modification, displaying, or transmitting of any course material content for any other purpose is prohibited, will be considered misconduct under the JHU Copyright Compliance Policy, and may be cause for disciplinary action. In addition, encouraging academic dishonesty or cheating by distributing information about course materials or assignments which would give an unfair advantage to others may violate AAP's Code of Conduct and the University's Student Conduct Code. Specifically, recordings, course materials, and lecture notes may not be exchanged or distributed

for commercial purposes, for compensation, or for any purpose other than use by students enrolled in the class. Other distributions of such materials by students may be deemed to violate the above University policies and be subject to disciplinary action.

Students with Disabilities

Johns Hopkins University is committed to providing reasonable and appropriate accommodations to students with disabilities. Students with documented disabilities should contact the coordinator listed on the [Disability Accommodations](#) page. Further information and a link to the Student Request for Accommodation form can also be found on the [Disability Accommodations](#) page.

Dropping the Course

You are responsible for understanding the university's policies and procedures regarding withdrawing from courses found in the current catalog. You should be aware of the current deadlines according to the [Academic Calendar](#).

Getting Help

You have a variety of methods to get help on Blackboard. Please consult the resource listed in the "Blackboard Help" link for important information. If you encounter technical difficulty in completing or submitting any online assessment, please immediately contact the designated help desk listed on the [AAP online support page](#). Also, contact your instructor at the email address listed in the syllabus.

Title IX Confidentiality and Mandatory Reporting

As an instructor, one of my responsibilities is to help create a safe and inclusive learning environment on our campus. I also have mandatory reporting responsibilities related to my role as a Responsible Employee under the Sexual Misconduct Policy & Procedures (which prohibits sexual harassment, sexual assault, relationship violence and stalking), as well as the General Anti-Harassment Policy (which prohibits all types of protected status based discrimination and harassment). It is my goal that you feel able to share information related to your life experiences in classroom discussions, in your written work, and in our one-on-one meetings. I will seek to keep information you share private to the greatest extent possible. However, I am required to share information that I learn of regarding sexual misconduct, as well as protected status based harassment and discrimination, with the Office of Institutional Equity (OIE). For a list of individuals/offices who can speak with you confidentially, please see Appendix B of the [JHU Sexual Misconduct Policies and Laws](#).

For more information on both policies mentioned above, please see: [JHU Relevant Policies, Codes, Statements and Principles](#). Please also note that certain faculty and other University community members also have a duty as a designated Campus Safety Authority under the Clery Act to notify campus security of certain crimes, as well as a duty under State law and University policy to report suspected child abuse and/or neglect.

Diversity

Johns Hopkins is a community committed to sharing values of diversity and inclusion in order to achieve and sustain excellence. We firmly believe that we can best promote excellence by recruiting and retaining a diverse group of students, faculty, and staff and by creating a climate of respect that is supportive of their success. This climate for diversity, inclusion, and excellence is critical to attaining the best research, scholarship, teaching, health care, and other strategic goals of the Health System and the University. Taken together these values are recognized and supported fully by the Johns Hopkins Institutions leadership at all levels. Further, we recognize that the responsibility for excellence, diversity, and inclusion lies with all of us at the Institutions: leadership, administration, faculty, staff, and students.

For more information on JHU's commitment to diversity, please visit the [Diversity at JHU](#) website.

Student Conduct Code

The fundamental purpose of the Johns Hopkins University's (the "University" or "JHU") regulation of student conduct is to promote and to protect the health, safety, welfare, property, and rights of all members of the University community as well as to promote the orderly operation of the University and to safeguard its property and facilities. As members of the University community, students accept certain responsibilities which support the educational mission and create an environment in which all students are afforded the same opportunity to succeed academically.

For a full description of the code please visit the [Student Conduct Code](#) website.

Course Evaluation

At the end of the semester, please complete the online course evaluation survey. These are an important tool in our work to improve the quality of our instruction and programs. The results of the course evaluations are kept anonymous – I will only receive aggregated data and comments for the entire class. An email with a link to the online course evaluation form will be sent to your JHU email address close to the end of the semester.

Appendix A

Tentative Course Schedule

Activity and assignment details will be explained in detail within each week's corresponding learning module (Lessons in Blackboard). If you have any questions, please contact your instructor.

This schedule is subject to change with fair notice. Any changes will be posted via Announcements in Blackboard.

Week	Objectives	Topics	Readings	Activities & Assessments
1	C1, C3	Getting Started with Python	McKinney 1-3	Code: Hello World DB: Intro
2	C1	Object Oriented Programming	Milliken, 7 Lubanovic, 10-11	Code: Objected Oriented Programming DB: Data Science Application/Resource I
3	C1, C3	NumPy	McKinney 4	Code: Numpy DB: Response to Week 2
4	C1, C2, C3	Pandas Part I	McKinney 5	Code: Pandas Part 1 DB: Project Initial Ideas
5	C1, C2	Pandas Part II	McKinney 6-8	Code: Pandas Part 2 DB: Project Initial Ideas Response
6	C2	Pandas Part III	McKinney 10-12	Code: Pandas Part 3 DB: Data Science Application/Resource II
7	C5	APIs	Lam 2-3	Code: API DB: Response to Week 6
8	C5, C6	Webscraping Intro	Lam 5	Code: NONE DB: Proposal Presentation Project: Written Proposal Submission
9	C6, C7	Jupyter NBs and Statistics	BBG 3	Code: Jupyter NB DB: Proposal Responses
10	C4	Visualization Part I	McKinney 9	Code: Matplotlib & Seaborn DB: Visualization Critique
11	C4	Visualization Part II	Online documentation	Code 10: Bokeh and Plotly DB: Response to Visualization Critique
12	C6	Modeling and ML	BBG 4-6	Code 11: Modeling DB: Data Science Application/Resource III
13		Work Week		DB: Response to Week 12
14	C6, C7	Submit Final Projects		DB: Presentation Response Final Project Submission
15	C6, C7	Presentation		DB: Presentation