Topics in Environmental Geography
GEOG-G 440/540

Python Programming

Department of Geography
Indiana University
Spring 2019 Syllabus

Lecture: Tuesday & Thursday 1:00 PM to 2:15 PM
Lecture Location: Student Building 221

Instructor: Dr. Natasha MacBean
Office: Student Building 204
Email: nmacbean@indiana.edu
Office hours: Tuesday & Thursday 2:15 PM to 3:15 PM

Course description

This course introduces students to the Python programming language and the basics of scientific computing. We will learn how to use Python to perform advanced geospatial data analyses and data visualization with large spatiotemporal datasets (for example modeling, remote sensing or GIS data).

In the course we will learn how to install python and how to download the various packages we will use, as well as how to download the remote sensing, climate, and modeling datasets we may analyze. Data handling topics will include, but will not be limited to: fundamental concepts in Python (e.g. data types, data structures, indexing, subsetting, looping over data), reading and manipulating large, multi-layer geospatial datasets, vector and raster data manipulation, map reprojection, area selection based on geographic coordinates), masking data, data visualization (including plotting maps).

Data analyses will include applying basic statistics to large datasets (sub, mean, max, min), interpolation and smoothing, regression, function fitting for extracting information from time series data, image classification and spatial clustering, dimensionality reduction (e.g. EOF analysis), point pattern analysis, and analysis of spatial heterogeneity and spatial autocorrelation.

Students will become familiar with the following many if not all of the following Python packages: NumPy (Numerical Python), GDAL (Geospatial Data Abstraction Library), SciPy (Scientific Python), (Geo)Pandas, StatsModels, scikit-learn, Matplotlib and Seaborn.

We will meet in a computing cluster and all classes will be lab focused where topics will be introduced with practical exercises. Exercises will be mostly taught via DataCamp.com No previous programming experience is required, but it would be helpful (for example GEOG-G 250/577, GEOG-G 488/588).

Optional texts

There are no required texts for this course. We will mostly be using the tutorials and exercises via DataCamp.com, supplemented with extra exercises in certain classes.
Everyone has their own view as to whether textbooks are useful for learning how to program/code, as there are often a lot of useful tutorials and websites with advice online. This is especially true for a freely available programming language such as Python. However, students may find the following textbooks useful. I will bring copies of these textbooks to class in the first few weeks so students can have a look and decide if they are worth buying themselves.

- McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc.
- Scopatz, A., & Huff, K. D. (2015). *Effective computation in physics: Field guide to research with python*. O'Reilly Media, Inc.
- Lawhead, J. (2015). *Learning Geospatial Analysis with Python Second Edition*. Packt Publishing Ltd.
- Grus, J. (2015). *Data science from scratch: first principles with python*. O'Reilly Media, Inc.

The course page will also be updated with useful websites and articles specific to each topic that will help us during each class.

Course Goals and Learning Outcomes

Course goals:
Students who engage in this class will be equipped to use Python to develop their own scientific code to analyses large spatiotemporal data analyses that will be useful in their own research and careers in geospatial data analysis. Students will be given a grounding in the basic principles of algorithm development and the construction of scientific programs for analyzing data that can be transferred and widely applied outside the topics covered in the class.

Learning outcomes:
At the end of this course, students should:

- Understand the basics of scientific computing
- Have a working knowledge of linux/unix operating systems
- Have a foundational knowledge of the Python programming language
- Know how to obtain, compile and install commonly available scientific software packages
- Be able to develop code and construct scientific programs Python to perform data manipulation tasks with large geospatial datasets
- Be able to carry out a variety of analytical tasks on large, multi-layer spatiotemporal datasets
- Understand technical issues with data visualization and geospatial data analysis and approaches to solve those issues.

Grading (440): In-class exercises – 60%
Final project – 40%

Grading (540): In-class exercises – 40%
Final project – 60%

In class exercises: We will be using DataCamp.com's Python tutorials and exercises throughout most of the semester. For some topics, the DataCamp exercises will be supplemented by extra questions to test students' comprehension of the material. Both the DataCamp assignment and any extra questions will be listed on Canvas for the date of the class. If there are extra questions, students must answer the questions in a word document and upload that to Canvas by the deadline. The grade for each assignment will be based partly on timely completion of the DataCamp.com exercise and partly on current answers to any extra questions on Canvas.
Final Project: The final project will be designed and carried out in Python by each student on a topic of their own choosing. This project will be completed in the final 4 weeks of the course. The student will use the data manipulation and analysis methods learned in the class to answer a scientific question pertaining to their own data analysis and research interests. More information about the final project will be provided at the start of the semester.
NOTE: The initial 1-page proposal for your project will be due after Spring Break.
Additional requirements for G540 students: Additional questions may be added to the assignments for G540 students. The final project will be more extensive and based on your research interests. More information will be given in class.

Grading Scale

97% - 100%	A+	77% - < 80%	C+
93% - < 97%	A	73% - < 77%	C
90% - < 93%	A-	70% - < 73%	C-
87% - < 90%	B+	67% - < 70%	D+
83% - < 87%	B	63% - < 67%	D
80% - < 83%	B-	60% - < 63%	D-
		< 60%	F

Grade Dissemination:
All grades will be posted on Canvas. Please make sure to track your own grades, as mistakes can occasionally occur. If you have received a grade by mistake, please see me for a correction in Canvas.

Python
Python is a high-level programming language that is freely available, relatively easy to learn and portable across different computing systems. In Python, you can rapidly develop

code/programs to perform the analyses you may need to do both in your research and future career. Code written in Python is also easy to maintain, is (or should be) self-documented, and can easily be linked to code written in other languages.
Advantages of Python include:

- it is automatically compiled and executed
- code is portable provided you have the appropriate Python modules installed.
- there are many well-documented existing packages you can employ for you data analysis needs
- there is an active user and development community, which means that there is a lot of useful advice and documentation online. Within this community, new capabilities appear over time and there are many existing extensions and enhancements easily available to you.

We suggest you use the anaconda python distribution if you are using your own computer.

Course Policies & Services

Course Policies:

Assignments and Late Work: You may work with others on the exercises, but you may not copy directly from anyone. Your work must be your own. *Late exercises will have 10% deducted per day, but exercises will not be accepted once graded copies are returned to the class.* The only allowable exceptions pertain to the IU policy on religious observances, military duty and family emergencies, illness with a doctor's note, or prior permission from me. If you have an excused absence for any of these reasons, make arrangements in advance or ASAP to makeup missed activities. *Plagiarism and cheating in any form will not be tolerated.*

Academic Integrity: As a student at IU, you are expected to adhere to the standards detailed in the *Code of Student Rights, Responsibilities, and Conduct* (*Code*). Academic misconduct is defined as any activity that tends to undermine the academic integrity of the institution. Violations include: cheating, fabrication, plagiarism, interference, violation of course rules, and facilitating academic dishonesty. When you submit an assignment with your name on it, you are signifying that the work contained therein is yours, unless otherwise cited or referenced. Any ideas or materials taken from another source for either written or oral use must be fully acknowledged. All suspected violations of the *Code* will be reported to the Dean of Students and handled according to University policies. Sanctions for academic misconduct may include a failing grade on the assignment, reduction in your final course grade, and a failing grade in the course, among other possibilities. If you are unsure about the expectations for completing an assignment or taking a test or exam, be sure to seek clarification from your instructor in advance. Please also see this website for the IU College of Arts and Science's policies on Academic Integrity.

Plagiarism and Academic Dishonesty: Plagiarism is the act of taking someone else's work and presenting it as your own. Plagiarism can occur in several forms, but whether the action is intentional or not, it is in violation of the IU Student Code of Conduct. Copying and pasting

text off of the Internet or any other source is NOT acceptable, as this is stealing someone else's work. If you use an author's materials verbatim, you must place these words in quotation marks, and must correctly reference them. Proper references must be placed in the text as well as in the bibliography. Be aware that when citing sources, you must do so not only when a passage is a direct quotation, but also when paraphrasing. Failure to do so will result in a "0" for the project. If you have questions or concerns regarding how to properly cite your resources, please ask me or contact Campus Writing Services. Additionally, if you are new to this university, are unclear about what plagiarism is, or would like a brief review of IU's standards, please look here and here.

Note Selling: Several commercial services have approached students regarding selling class notes/study guides to their classmates. Selling the instructor's notes/study guides in this course is not permitted. Violations of this policy will be reported to the Dean of Students as academic misconduct (violation of course rules). Sanctions for academic misconduct may include a failing grade on the assignment for which the notes/study guides are being sold, a reduction in your final course grade, or a failing grade in the course, among other possibilities. Additionally, you should know that selling a faculty member's notes/study guides individually or on behalf of one of these services using IU email, or via Canvas may also constitute a violation of IU information technology and IU intellectual property policies; additional consequences may result.

Students with Disabilities: The Americans with Disabilities Act (ADA), the Indiana Civil Rights Act, and Indiana University policy prohibit discrimination in educational programs against students with disabilities. Disabilities may include medical, auditory, visual, learning, psychological, mobility, or neurological problems. It is the policy of Indiana University to provide reasonable accommodations in a timely manner and on an individualized basis while maintaining institutional standards of performance. These accommodations are designed to counter the effects of disabilities where they may pose a barrier to the education process; they will not give the student an easy grade or an advantage over other students. See the Office of Disability Services for Students for accommodation and documentation.

Religious Observances: See here from more information on religious accommodation.

Sexual Harassment: As your instructor, one of my responsibilities is to help create a safe learning environment on our campus. Title IX and our own Sexual Misconduct policy prohibit sexual misconduct. If you have experienced sexual misconduct, or know someone who has, the University can help.

If you are seeking help and would like to speak to someone confidentially, you can make an appointment with:

The Sexual Assault Crisis Service (SACS) at 812-855-8900

Counseling and Psychological Services (CAPS) at 812-855-5711

Confidential Victim Advocates (CVA) at 812-856-2469 IU Health Center at 812-855-4011.

More information about available resources can be found here: http://stopsexualviolence.iu.edu/help/index.html.

It is also important that you know that federal regulations and University policy require me to promptly convey any information about potential sexual misconduct known to me to our campus' Deputy Title IX Coordinator or IU's Title IX Coordinator. In that event, they will work

with a small number of others on campus to ensure that appropriate measures are taken and resources are made available to the student who may have been harmed. Protecting a student's privacy is of utmost concern, and all involved will only share information with those that need to know to ensure the University can respond and assist.
I encourage you to visit http://stopsexualviolence.iu.edu to learn more.

Support services for Students

Knowledge base and UITS support center:
For any technical support, see the Knowledge Base or go to the UITS Support Center website.

Schedule

Please see the following page.

Please note that this tentative syllabus/schedule may change without notice in order to reflect the needs of our classroom. See the course webpage on Canvas for updates.

Week	Date	Topic
1	1/8/19	Course Introduction – What is programming? Python Basics
1	1/10/19	Python lists, functions and packages
2	1/15/19	Introduction to NumPy (arrays and vectorized computation)
2	1/17/19	Introduction to Matplotlib (data visualization)
3	1/22/19	Dictionaries and Introduction to Pandas
3	1/24/19	Flow control, logic, and filtering
4	1/29/19	Loops (and case study)
4	1/31/19	More on writing your own functions
5	2/5/19	Importing data into Python (flat files and other file formats - Part 1 in DataCamp)
5	2/7/19	Cleaning and tidying data in Python
6	2/12/19	Manipulating data in Python
6	2/14/19	*More on data types in Python (Data Types for Data Science)/Pandas Foundations*
7	2/19/19	In-depth Pandas Part 1
7	2/21/19	In-depth Pandas Part 2
8	2/26/19	In-depth Data Visualization with Matplotlib
8	2/28/19	Image (2D array) plotting with Matplotlib
9	3/5/19	Statistical analyses of time series and image data with Seaborn Part 1
9	3/7/19	Statistical analyses of time series and image data with Seaborn Part 2
10	3/12/19	**Spring Break**
10	3/14/19	**Spring Break**
11	3/19/19	Visualizing and manipulating geospatial data in python
11	3/21/19	Visualizing and manipulating geospatial data in python
12	3/26/19	Pattern recognition: classification, clustering, dimension reduction
12	3/28/19	Pattern recognition: classification, clustering, dimension reduction
13	4/2/19	Final Project
13	4/4/19	Final Project
14	4/9/19	Final Project
14	4/11/19	Final Project
15	4/16/19	Final Project

15	4/18/19	Final Project
16	4/23/19	Final Project
16	4/25/19	Final Project