Introduction to Python Programming for Life Sciences

Course description:

This course is designed for students with no programming background who want to learn professional programming in Python, a computer language popular across disciplines. In this course, the examples and exercises are drawn from life science applications. Students learn how to do number calculations, use multi-dimensional arrays, and basically think in vectors. Students learn how to operate on tables and time series and manipulate spreadsheets that can be larger than the ones most spreadsheet programs deal with. Topics covered include functions, recursion, data structures, reading and writing files, object-oriented design, debugging, and databases using industry best practices. Students are introduced to the basic Python libraries, such as the numerical performance library NumPy, the data structuring library Pandas, the scientific library SciPy, as well as data algorithms in scikit-learn and visualization using matplotlib.

Tentative Topics:

Week 1. Introduction to Python and Python Notebooks, variables, expressions, operators, conditionals, and iteration.

Week 2. File input and output, strings, regular expressions, lists, sets, tuples, dictionaries, modules

Week 3. Numerical performance with NumPy and data tabulation with Pandas (series, dataframes, descriptive statistics, missing data)

Week 4. Data Modeling with probability and statistics

Week 5. Object oriented programming, inheritance, objects, classes, methods

Week 6. Graphs with NetworkX and Linear Algebra with the Power Method

Week 7. Midterm Exam

Week 8. Python databases with MongoDB and a touch of Web programming with Flask

Week 9. Decision Trees for bioinformatics

Week 10. Visualization, graphics and animation

Week 11. Shrinking data with k-means and Principal Component Analysis

Week 12. Simple Linear Regressions for Data Modeling

Week 13. Classification and clustering in scikit-learn

Week 14. Final Exam

A “Project Zero”: an initial diagnostic assessment will be given at the end of the first day of class. Completion of this assignment is required for acceptance to the class. It consists of a **brief, ungraded assignment that will allow you and your students to evaluate if they currently have the skills and knowledge necessary to succeed in this course.**

The Extension School is committed to providing an accessible academic community. The Accessibility Office offers a variety of accommodations and services to students with documented disabilities. Please visit https://www.extension.harvard.edu/resources-policies/resources/disability-services-accessibility for more information.

You are responsible for understanding Harvard Extension School policies on academic integrity (https://www.extension.harvard.edu/resources-policies/student-conduct/academic-integrity) and how to use sources responsibly. Not knowing the rules, misunderstanding the rules, running out of time, submitting the wrong draft, or being overwhelmed with multiple demands are not acceptable excuses. There are no excuses for failure to uphold academic integrity. To support your learning about academic citation rules, please visit the Harvard Extension School Tips to Avoid Plagiarism (https://www.extension.harvard.edu/resources-policies/resources/tips-avoid-plagiarism), where you'll find links to the Harvard Guide to Using Sources and two free online 15-minute tutorials to test your knowledge of academic citation policy. The tutorials are anonymous open-learning tools.