

Python for Data Science

Applied Technology Academy

Five days that supply the Python tooling data science runs on. The course starts with Python fundamentals, moves to iPython and Jupyter notebooks as a platform for exploring data, then covers pandas extensively - loading, exploratory analysis, transforming, reshaping and exporting - before building useful visualizations with Matplotlib and Seaborn. NumPy, SciPy, Excel files, object-oriented basics and error handling are covered along the way. No prior programming experience is required.

LEVEL	DURATION	DELIVERY
Introductory	5 Days	Instructor-led

Course Overview

- Five days, hands-on, with scripts and interactive notebooks used side by side.
- This is not a data science course in itself - it provides the tools to do data science.
- pandas is covered extensively: loading, exploratory analysis, cleaning, reshaping and exporting.
- Materials: a 400+ page course manual and a Python quick reference.

Course Outline

- **Day One**
 - The Python Environment - starting Python, using the interpreter, running a script.
 - Variables and Values - variables, strings and string operations, numbers, converting types.
 - Basic I/O - writing to the screen, string formatting, command-line parameters, the keyboard.
 - Flow Control - conditional expressions, Boolean values, relational operators, while loops, loop exits.
- **Day Two**
 - Array types - lists and tuples, indexing and slicing, iteration, sequence functions and operators, list comprehensions, generator expressions, nested sequences.
 - Working with files - opening, reading and writing text files, raw binary data.
 - Dictionaries and Sets - creating and iterating dictionaries, creating and working with sets.
 - Functions, modules and packages - return values, parameter types, scoping, documentation, creating and importing modules, organizing packages.
- **Day Three**
 - Errors and Exception Handling - syntax errors, exceptions, try/except/else/finally.
 - Introduction to Python Classes - defining classes, constructors, instance methods and data, attributes, inheritance.
 - Excel spreadsheets - the openpyxl module, reading, creating and modifying workbooks.

- iPython - features, magic commands, configuration.
- **Day Four**
 - Jupyter Notebooks - using notebooks, Jupyter Lab, markdown, managing output, exporting.
 - Brief intro to SciPy - what it provides, useful functions, subpackages.
 - Intro to NumPy - creating arrays, indexing and slicing, large number sets, transforming data.
 - pandas I, basics - Series and DataFrames, loading data, data summaries, data types, basic plotting.
- **Day Five**
 - pandas II, exploring and cleaning - selecting and querying, indexes, dropping and modifying data,
 - missing values, groupby, multi-indexes.
 - pandas III, reshaping - pivot tables, merging, melting, stacking and unstacking.
 - Plotting - getting started with Matplotlib, common plots, Seaborn, tweaking plots.

Prerequisites

- No prior programming experience is required.
- Participants should be comfortable with science and math concepts.
- Comfort working with files and folders is expected.
- Familiarity with the command line in Linux, Windows or macOS is useful but not crucial.

Who Should Attend

- Analysts and scientists moving into Python.
- Users of R, Julia, SAS or MatLab who want to apply that experience in Python.
- Engineers and technical staff who need to explore and visualise real data.

What You'll Learn

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

- create and run Python programs, and understand the fundamentals behind them
- design and code modules and classes
- process CSV and Excel files
- manipulate arrays with NumPy
- work with the breadth of subpackages that make up SciPy
- use Series and DataFrames in pandas to load, query, clean and reshape data
- create plots with Matplotlib and Seaborn
- use Jupyter notebooks for ad hoc calculation, plotting and what-if analysis

Environment and Tools

- Latest Python 3 from Anaconda, plus the additional modules the labs use.
- An IDE or editor: Visual Studio Code, PyCharm or Eclipse.
- Lab files are supplied, with detailed setup instructions before class.

Ready to enroll? Call 800.674.3550 or email Info@AppliedTechAc.com — private team cohorts, GSA MAS purchasing, military credentialing funding, VA GI Bill and VR&E and student financing available.