

Python for Scientists and Engineers

A specialist course in Brisbane



Audience: This is a course for scientists and engineers interested in using Python for solving computational problems that arise in daily work and automating the processing of different kinds of scientific data.

Outcome: By the end of the course, you will have all the knowledge you need to start programming competently in Python for scientific and engineering applications. You will know what's available with Python and how to use its powerful data types and amazing libraries to write clean, efficient code. You will have had experience with using Python for various scientific data manipulation tasks and solving a range of analytical tasks with medium-sized data sets, including easily creating beautiful plots, performing Monte Carlo simulations, constructing statistical models, regression, and analysing images and time-series data.

Duration: 4 days

Dates: 9–12 April 2018

Venue: Training Choice, Mezzanine level, 88 Creek Street, Brisbane CBD

Format: Each topic is a mixture of hands-on exercises and expert instruction.

Instructors: Henry Walshaw, Edward Schofield, Robert Layton, and/or Juan Nunez-Iglesias

Prerequisites: Familiarity with programming concepts (in any language) is assumed.

Optional GIS day: A fifth day on geospatial analysis is available as an extension to the course.

Please see here: <https://pythoncharmners.com/training/python-for-geospatial-analysis>

Course Outline

Day 1: Introduction to Python

Day 1 covers the basics of using the Python language and standard library, with a focus on scientific and engineering applications, including tips and tricks for making this easy:

- The Jupyter notebook and shell for rapid prototyping
- Modules and packages
- Python syntax and concepts
- Essential data types, tips and tricks
- Handling exceptions
- Worked example: fetching and ranking real-time temperature data for global cities from a web API

Day 2: Handling, Analysing, and Presenting Data in Python

Day 2 gives a comprehensive introduction to reading and writing some of the most important data formats in science and engineering and how to analyse and visualise data easily:

- Reading and writing essential data formats: CSV, Excel, SQL databases
- Indexing and selecting data in *Pandas*
- Data fusion: joining & merging datasets
- Summarisation with “group by” operations; pivot tables
- Time-series analysis: parsing dates, resampling
- Visualisation and statistical graphics with *Seaborn*
- Worked example: creating automated reports with *Pandas* and *nbconvert*

Day 3: Essentials of Scientific Computing with Python

Day 3 teaches the use of Python for numerical and scientific computing. It covers array and matrix manipulation and an overview of available scientific routines, including an introduction to statistical modelling:

- Introduction to manipulating vectors and matrices with *NumPy*
- Tour of *SciPy* and related packages for scientific data manipulation, with fancy demos: unit conversions, clustering with *scikit-learn*, interpolation, dense & sparse linear algebra, signal processing, image processing
- Statistics in Python, with scientific applications: modelling, confidence intervals, hypothesis testing, linear regression, Monte Carlo simulation

Day 4: Scaling up

Day 4 focuses on techniques for creating larger codebases in teams and scaling up from small datasets to large ones that are too big for memory or too slow for one computer to process. It also introduces further kinds of machine learning (classification and regression) for automatically inferring models from large datasets:

Morning: best practices

- Defining your own classes: custom data types
- Integrated development environments; Python idioms and style
- Sharing code with others: creating scripts, modules and packages; revision control with Git

Afternoon: scaling up

- Efficiency: profiling, vectorization, and JIT compilation for speeding up code by 4x to 10,000x
- Parallel computing with *Dask*
- Machine learning with *scikit-learn*: classification and nonlinear regression

Topics available on request

Some more specialised topics we can include as supplementary topics on request include:

- Interfacing Python with other languages: Excel, R, C/C++, Matlab, Fortran
- Interfacing with other data formats, e.g. NetCDF, HDF5

Supplemental materials

We will supply you with printed course notes and a USB stick containing kitchen-sink Python installers for multiple platforms, electronic versions of the course notes, solutions to the programming exercises, several written tutorials, and reference documentation on Python and the third-party packages covered in the course.

Instructor bios

Your trainers for the course will be selected from:

Dr Edward Schofield

Ed is well-known in the Python and scientific Python communities as the author of the *future* package and a former release manager of *SciPy*. He has consulted to or trained dozens of organisations in scientific and data analytics applications of Python, including ANU, A*STAR, the Bureau of Meteorology, Cisco, CSIRO, Dolby, Geoscience Australia, Optus, and Toyota Technical Centre. Ed is co-chair of the Python in Science & Data miniconf for PyCon AU and co-organises the Python user group in Melbourne.

Ed holds a PhD in computer science from Imperial College London, where his thesis was in machine learning. He also holds BA and MA degrees in maths and computer science from Cambridge University. He has 20+ years of experience in programming, teaching, and public speaking.

Dr Robert Layton

Robert is the author of the book “Data Mining in Python”, published by O’Reilly. He provides analysis, consultancy, research and development work to businesses primarily using Python. Robert has worked with government, financial and security sectors, in both a consultancy and academic role. He is also a Research Fellow at the Internet Commerce Security Laboratory, Federation University Australia, investigating cybercrime analytics and data-mining algorithms for attribution and profiling.

Robert is a regular contributor to the Python-based scikit-learn open source project for machine learning and writes regularly on data mining for a number of outlets. He has presented regularly at a number of international conferences in Python, data analysis, and its applications.

Henry Walshaw

Henry has over 10 years of experience in GIS, spatial analysis and application development, particularly in the natural resource management field. Henry’s core technical expertise relates to the development and analysis of large scale spatial datasets (primarily using Python), and communicating this understanding to people including subject matter experts and the general public.

Henry has worked in government at federal and state levels, at Geoscience Australia (GA), the Victorian Government Department of Sustainability and Environment (DSE), and the Environmental Protection Agency (EPA). He has also worked in the private sector as Senior Spatial Consultant with Sinclair Knight Merz (SKM) and we-do-IT. He holds a Bachelors in Computational Science.

Dr Juan Nunez-Iglesias Juan Nunez-Iglesias is co-author of the book *Elegant SciPy*, published by O'Reilly Media. Juan is a core developer of the scikit-image Python library, and has contributed to many others in the scientific Python ecosystem, including SciPy, NetworkX, and Matplotlib. He has taught and presented at the SciPy conference in Austin, EuroSciPy, PyCon Australia, the Advanced Scientific Programming in Python summer school, and Software Carpentry workshops. He is currently a research scientist at the University of Melbourne, where he uses image analysis and computer vision to study brain connectivity, malaria parasites, and other biological questions.

Juan has Bachelor's degree in Biomedical Science from the University of Melbourne and both an MSc in Statistics and PhD in Computational Biology and Bioinformatics from the University of Southern California.

Other information

Computer: A computer with an internet connection will be provided for you for the course.

Personal help: Your trainer(s) will be available after the course each day for you to ask any one-on-one questions you like — whether about the course material and exercises or about specific problems you face in your work and how to use Python to solve them.

Food and drink: We will provide lunch, morning and afternoon tea, and drinks.

Timing: The course will run from 9:00 to roughly 17:00 each day, with breaks of an hour for lunch and 15 minutes each for morning and afternoon tea.

Price

\$825 per day per person, including GST.

Booking

To book places on the course, please contact us, or visit:

<https://pythoncharmners.com/training/python-for-scientists-engineers>

Testimonials

Testimonials from past participants of this course are available at pythoncharmners.com/testimonials.

Questions?

You are welcome to contact us if you have any questions before the course. You can reach us at info@pythoncharmners.com.

About Python Charmers

Python Charmers is the leading provider of Python training in the Asia-Pacific region, based in Australia and Singapore. Python Charmers specialises in teaching programming for scientists, engineers, data scientists, and computer scientists in the Python language.

Contact

Phone: +61 1300 963 160
Email: info@pythoncharmners.com
Web: pythoncharmners.com

