



# PYTHON FOR MACHINE LEARNING

COURSE GUIDE (V1): FEBRUARY – DECEMBER 2024

© 2024 PYTHON CHARMERS®

# Python for Machine Learning

A specialist course

**Audience:** This is a course for data scientists, quants, financial analysts, researchers, statisticians, and software developers interested in learning the fundamentals of data science and machine learning in Python.

**Context:** Trawling through large volumes of data and making accurate inferences and predictions are now skillsets in fierce demand. Python offers some of the best tools for this in any language.

**Overview:** This intensive, hands-on, practical training course will teach you how to apply some of Python's most powerful and productive tools processing, analyzing, modelling, and visualizing various kinds of data. It will then teach you in-depth how to apply the three primary modelling strategies of "classical" machine learning: clustering, classification, and regression.

**Format:** Live instructor-led training (online). Each topic is a mixture of expert instruction, worked examples, and hands-on exercises with help from the instructor(s).

**Expert instructors:** See bios below.

**Duration:** 4 days

## Modules:

Days 1–4: Python for Machine Learning  
Days 1–2: Introduction to Python  
Days 3–4: for graduates of other Python Charmers courses  
or people with 3–6 months Python experience

## Price:

Regular course (4 days): AUD \$3,600 (excl GST)  
Modular / partial courses: AUD \$900 (excl GST) per day

## Dates (February – December 2024):

12 - 15 February 2024  
18 - 19 March 2024 (days 1-2); 25-26 March 2024 (days 3-4)  
29 April - 2 May 2024  
3 - 6 June 2024  
24 - 27 June 2024  
12 - 15 August 2024  
21 - 24 October 2024  
9 - 12 December 2024

**PYTHON**  
C H A R M E R S

# Skills & Activities

## Skills

**Overall:** You will gain a broad understanding of methods and tools in machine learning and come to appreciate the elegance and power of the Python language and its powerful ecosystem of packages for data science.

**Days 1–2:** You will gain a solid understanding of the Python language and experience using Python for a wide range of scripting and data-manipulation tasks with data in a variety of useful formats; and creating beautiful statistical graphics and simple dashboards.

**Day 3:** You will learn about time-series analysis; manipulating vector/matrix data; performing Monte Carlo simulations; clustering; and outlier / anomaly detection.

**Day 4:** You will learn how to use machine learning (ML) to construct powerful predictive models using regression and classification; and how to evaluate, refine, and deploy ML models.

## Activities

**Worked examples:** To present the concepts and prepare you for the exercises, the trainer will present worked examples and demos and help you to follow along with these on your own computer.

**Exercises:** There will be practical exercises throughout the training course. These will be challenging and fun, and the solutions will be discussed after each exercise and provided as source code. During the exercises, the trainer will offer you help and suggestions.



# Topic outline

## Day 1: Python basics

Day 1 covers how to use Python for basic scripting and automation tasks, including tips and tricks for making this easy:

- Why Python? What's possible?
- The *Jupyter* notebook for rapid prototyping
- Modules and packages
- Python concepts: an introduction through examples
- Essential data types: strings, tuples, lists, dicts
- Worked example: retrieving real-time data from a REST web API
- Raising and handling exceptions

# Topic outline

## Day 2: Handling, analyzing, and presenting data in Python

Python offers amazingly productive tools like Pandas for working with different kinds of data. Day 2 gives a thorough introduction to analyzing and visualizing data easily:

- Reading and writing essential data formats:  
CSV, Excel, SQL, time-series (others on request)
- Indexing and selecting data in *Pandas*
- Data fusion: joining & merging datasets
- Data summarization with pivot tables
- Visualization and statistical graphics with *Plotly Express*
- Worked example: creating automated reports
- Creating interactive dashboards with *Streamlit*

# Topic outline

## Day 3: Time-series, simulation, clustering, outlier detection

Day 3 shows you how to manipulate time-series and matrix/vector data. It then gives examples of Monte Carlo simulation and shows you how to apply powerful unsupervised learning techniques to applications in clustering and outlier / anomaly detection:

- Time-series analysis: parsing dates; resampling; interpolation
- Introduction to *NumPy* for manipulating vector and matrix data: data types, powerful indexing, reshaping, *ufuncs*
- Monte Carlo simulation and applications
- Clustering with *scikit-learn*; applications to customer profiling, text correction, quantization
- Outlier and anomaly detection with *pyod*; applications to time-series

# Topic outline

## Day 4: Machine learning

Day 4 gives you a practical and comprehensive introduction to supervised methods of machine learning for powerfully inferring complex models from data. It uses a range of datasets, including time-series and spatial data, and gives examples from a range of industries:

- Intuition behind ML; overview of the ML package ecosystem in Python
- Classification; application to diagnosis, AI systems, satellite imagery, ...
- Validation and model selection; diagnostic tools; *yellowbrick*
- Linear and nonlinear regression; application to time-series forecasting
- Feature engineering and selection; pipelines
- Deploying machine learning models in production
- Intuition behind core ML algorithms in *scikit-learn*:  
Naive Bayes, logistic regression, SVMs, random forests



## Personal help

We are happy to offer on-the-spot problem-solving after each day of the training for you to ask one-on-one questions — whether about the course content and exercises or about specific problems you face in your work and how to solve them. If you would like us to prepare for this in advance, you are welcome to send us background info before the course.



**PYTHON**  
CHARMERS



## Other information

**Format:** Courses are conducted online via video meeting using Python Charmers' cloud notebook server for sharing code with the trainer(s).

**Computer:**

- **Hardware:** we recommend  $\geq 8$  GB of RAM and a webcam. Preferably also multiple screens and a quiet room (or headset mic).
- **Software:** a modern browser: Chrome, Firefox, or Safari (not IE or Edge); and Zoom.
- **Coding:** we have a cloud-based coding server that supports running code and sharing code with the trainer(s).

**Timing:** Most courses will run from 9:00 to roughly 17:00 (AEST) each day, with breaks of 50 minutes for lunch and 20 minutes each for morning and afternoon tea.

**Certificate of completion:** We will provide you a certificate if you complete the course and successfully answer the majority of the exercise questions.

**Materials:** You will have access to all the course materials via the cloud server. Upon request, a copy of the cheat sheets and a USB stick containing the materials, exercise solutions, and further resources will be sent to your nominated mailing address.

```
data, cmap='winter')
```

```
python3.7/site-packages/matpl  
cnis-1- elementw comparison
```

```
but the future will perfo
```

```
str('ce')
```

**PYTHON**  
CHARMERS



# Instructor bio



**Dr Edward Schofield**

Ed has consulted to or trained over 3000 people from scores of organisations in Python and data science, including Atlassian, Barclays, Cisco, CSIRO, Dolby, Harvard University, IMC, Singtel Optus, Oracle, Shell, Telstra, Toyota, Verizon, and Westpac. He is well-known in the Python community as a former release manager of *SciPy* and the author of the widely used *future* package. He runs leads the Python user group in Melbourne and regularly presents at conferences in data analytics and Python in Australia and internationally.

Ed holds a PhD in machine learning from Imperial College London, with a focus on language models and other generative models for sequence prediction. He also holds BA and MA (Hons) degrees in mathematics and computer science from Trinity College, University of Cambridge. He has 20+ years of experience in programming, teaching, and public speaking.



**PYTHON**  
C H A R M E R S



# Instructor bio



**Dr Robert Layton**

Robert is the author of the book *Data Mining in Python*, published by Packt. 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, investigating cybercrime analytics and data-mining algorithms for attribution and profiling.

Robert is a 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 also created the website *LearningTensorFlow.com* and sold it to DataBricks in 2017. He has presented at a number of international conferences in Python, data analysis, and its applications.



**PYTHON**  
CHARMERS



## About Python Charmers®

Python Charmers is a leading global provider of training in data science and software development, based in Australia and Singapore. Since 2010, Python Charmers has given over 600 training courses and bootcamps to over 6000 delighted people from organizations such as Atlassian, BAE, Barclays, Cisco, CSIRO, Deloitte, Dolby, GDF Engie, IMC, pwc, Qinetiq, Singtel Optus, Shell, Sportsbet, Toyota, Verizon, and Westpac. Python Charmers specializes in teaching data science and programming to scientists, engineers, data analysts, quants, and computer scientists in the Python language.

Python Charmers' trainers have years of experience with machine learning, data analytics, statistical modelling, programming, and teaching, and deep roots in the open source community, as both speakers at events and contributors to well-known open source projects for data science, including *NumPy*, *SciPy*, *Scikit-Learn*, *Pandas*, *Matplotlib*, *NetworkX*, *Dash*, and *Future*.

**Testimonials:** Testimonials from past participants of similar bootcamps and training courses are available at

<https://pythoncharmners.com/testimonials/>

**Questions:** We are happy to customise this program further on request. Please let us know if you would like to discuss this or have any other questions.

**Contact:**

Phone: +61 1300 963 160

Email: [info@pythoncharmners.com](mailto:info@pythoncharmners.com)

Web: [pythoncharmners.com](https://pythoncharmners.com)

The logo for Python Charmers, featuring the word "PYTHON" in a large, bold, sans-serif font, with a small green Python snake icon integrated into the letter "O". Below "PYTHON" is the word "CHARMERS" in a smaller, all-caps, sans-serif font.

PYTHON  
CHARMERS





PYTHON  
CHARMERS®