PYTHON
CHARMERS®

PYTHON FOR BEGINNERS

TRAINING FOR IMC: 2024

Python for Beginners

An interactive training course at IMC

Audience: People with no or little prior experience in programming. A mixed group of traders, developers, network & performance engineers, and/or business analysts.

Overview: This intensive, hands-on, practical training course will give you all the knowledge you need to start using Python competently for automatically processing and analysing financial data with a focus on financial time-series.

Skills: You will learn the fundamentals of Python's powerful data types, the advantages and disadvantages of each, and when to use which data type. You will also learn when and when not to use the *Pandas* package and how to slice, dice, merge, aggregate, pivot, clean, munge, resample, and visualize time-series and other data with ease.

Overall, you will appreciate the elegance and power of the Python language and the breadth of its amazing ecosystem of powerful packages for solving many kinds of problems. You will be well-placed to continue learning more as you use Python day-to-day.

Duration: 4 sessions x ½ day (mornings), followed by an optional internal IMC session on retrieving IMC data

Instructor: Dr Edward Schofield and/or Dr Robert Layton

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

Location: Online (live)

PYTHON
CHARMERS

Topic outline

Session 1: Basics

Session 1 will teach you about *Jupyter* and introduce core *Python* concepts through examples, with a focus on *Python*'s powerful data types:

- The *Jupyter* notebook for rapid prototyping

- Functions; docstrings; importing modules and packages

- Essential data types and their uses: *str*, *list*, *dict*

- Worked example: fetching real-time data from a REST web API

- Raising and handling exceptions

PYTHON CHARMERS

Topic outline

Session 2: *Python* in practice

Session 2 will teach you about further data types, *Python* language topics, how to structure code for re-use, and tooling:

- Worked example: finding duplicate files (by contents)

- Further data types: *set*, *tuple*, *namedtuple*

- Further worked examples

- Creating your own modules and packages

- Code re-use; brief intro to the *VS Code* IDE

- Creating and managing *conda* environments

PYTHON
CHARMERS

Topic outline

Session 3: Handling, analyzing, and presenting data

Python offers amazingly productive tools like *Pandas* for working with different kinds of data. Session 3 will give you an introduction to slicing, dicing, and analyzing tabular data in powerful ways:

- Reading and writing essential data formats:
    CSV, Excel, SQL, time-series, JSON

- Indexing and selecting data in *Pandas*

- Data fusion: joining & merging datasets

- Summarization with "group by" operations; pivot tables

PYTHON CHARMERS

Topic outline

Session 4: Time-series; visualizing and presenting data

Session 4 focuses on analyzing time-series data and visualizing various kinds
of tabular data:

- Handling time-series data:
        parsing dates; resampling; rolling-window operations

- Intermediate *Pandas*: dtypes, reshaping data; hierarchical indices; styles

- Introduction to visualizing data in Python

- Worked examples: aggregating and summarizing market data

PYTHON
CHARMERS

Topic outline

Session 5 (Optional, IMC internal): Retrieving IMC data

This session describes how to access data within IMC using:

- Impala

- Dremio (Parquet file format)

- Spark / PySpark

- Authentication with *kerberos*

PYTHON
CHARMERS

Personal help

We are happy to offer on-the-spot problem-solving after each day of the training for participants to ask one-on-one questions — whether about the course content and exercises or about specific problems they face in their work and how to solve them.

If you would like us to prepare for this in advance, you are welcome to send us sample datasets and background info before the course.

Other information

Custom topics: We would be happy to customise the above program to include other desired topics and/or sample datasets.

Materials: We will provide participants downloadable course notes (PDF and *Jupyter* notebooks), cheat sheets, solutions to the programming exercises, several written tutorials, and reference documentation on Python and the third-party packages covered in the course.

Exercises: There will be practical programming exercises throughout the 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(s) will offer individual help and suggestions.

Live online delivery: We have cloud infrastructure (using *JupyterHub*) that allows participants to code online via their web browser. Our trainer(s) can see in real-time how the participants are faring with the exercises and offer help. The cloud coding environment also contain our materials, interactive notebooks and sample datasets.

PYTHON
C H A R M E R S

Instructor bio

Dr Edward Schofield

Ed has consulted to or trained over 3000 people from dozens of organizations in data science and Python, including ANZ, Atlassian, Barclays, Cisco, CSIRO, Deloitte, Dolby, Harvard University, IMC, Oracle, Shell, Singtel Optus, 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. In normal times he regularly presents at conferences in data science and Python in Australia and internationally.

Ed holds a PhD in machine learning from Imperial College London. 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.

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 was the author of the website "LearningTensorflow.com", sold 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 6,000 delighted people from organizations such as AGL, Atlassian, Barclays, CSIRO, Cisco, Deloitte, Dolby, IMC, pwc, Singtel Optus, Shell, Sportsbet, Telstra, Toyota, Verizon, Westpac, and Woolworths. Python Charmers specializes in teaching programming and data science to scientists, engineers, data analysts, quants, and computer scientists.

Python Charmers' trainers boast years of experience with data science, data analytics, statistical modelling, and programming, 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*, *Scikit-Image*, *NetworkX*, and *Python-Future*.

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

https://pythoncharmers.com/testimonials/

Questions: We would be happy to hear from you. Please let us know if you have any questions.

Contact:
Phone: +61 1300 963 160
Email: info@pythoncharmers.com
Web: http://pythoncharmers.com

PYTHON
C H A R M E R S