Python for Data Science by Leo Van - Coding Mentors

Learning outcome:

1. Python programming fundamentals
2. Perform analysis on any dataset and present your findings
3. Build your first Machine Learning model

Tools:

1. Python
2. Jupyter notebooks
3. Visual Studio
4. pandas
5. numpy
6. matplotlib
7. seaborn
8. scikit-learn

Course duration: 5 weeks (4 hours of lecture + XX hours of self-paced study)

Syllabus:

| ID | Week | Duration | Topic | Content |
| --- | --- | --- | --- | --- |
| L00 | W1 | 2 hours | Set up your workstation for data science | * Install anaconda * Install visual studio and useful extensions (can replace with any other IDEs: Pycharm, Sublime text, Atom, etc.) * Customise your terminal |
| L01 | W1 | 2 hours | Data types and operators | * Common types in Python: integers, float, boolean, string, list, tuple, set, dictionary * Perform computations and create logical statements * Declare, assign, and reassign values using Python variables |
| L02 | W2 | 2 hours | Control flow | * Write control flow using if, else, elif statements and boolean expressions * Lists * Use for and while loops to iterate and manipulate lists * Use break and continue to skip iterations in loops * List comprehension |
| L03 | W2 | 2 hours | Functions | * Define custom functions * Function parameters, arguments, keyword arguments * Create and reference variables using the appropriate scope * Write your first docstring to add documentation to the function * Anonymous functions using lambda expressions |
| L04 | W3 | 2 hours | Strings and Dictionaries | * Create, rearrange, reassign, disassemble, and reassemble block of texts * Understand Python dictionary structure and how to use key-value pair |
| L05 | W3 | 2 hours | numpy | * Create, access, modify and sort multidimensional numpy arrays (ndarrays) * Use slice indexing and boolean indexing to access and change subsets of an ndarray * Element-wise operations on ndarrays * Broadcasting |
| L06 | W4 | 2 hours | pandas | * Core objects in pandas: Series and DataFrame * Create, access, and modify core pandas objects * Index, select, and assign * Summary functions and maps * Grouping and sorting * Deal with np.NaN values * Combine 2 or more DataFrames |
| L07 | W4 | 2 hours | Data visualisation | * Use seaborn and matplotlib * Line charts * Bar charts and heatmaps * Scatterplots * Distributions |
| L08 & L09 | W5 | 4 hours | Intro to Machine Learning (ML) | * ML project life-cycle * Basic data exploration * Feature engineering * Feature selection * Build your first ML model * Tune your model |