

Introduction to Python for Data Science (Part 2)

This planning document is intended to support teachers who are delivering the NPA/PDA Data Science or for students who are learning independently. It also aligns with the Data Skills for Work framework.

Contents

Lesson Description	1
Lesson Contents	2
Learning Intentions	2
Success Criteria	2
Knowledge Prerequisites	2
Lesson Requirements.....	4
Jupyter Notebook.....	4
Python Learning Resources.....	6
How you can use this lesson	6

Lesson Description

Lesson Overview	An introduction to the use of Python for Data Science projects.
Topic	Tools and Languages
Book Chapter(s)	“Tools and Languages”

NPA level	5, 6
PDA level	7, 8
Data skills for work level	Analysis

Lesson Contents

This lesson consists of:

- A lesson plan (this document)
- A PowerPoint presentation, 'introduction to Python for data science'
- 2 Jupyter notebooks:
 - 'intro_to_python_for_data_science_part_2.ipynb' (for learners)
 - 'intro_to_python_for_data_science_with_answers_part_2.ipynb' (for teachers)

The Jupyter notebook for teachers contains answers to the tasks set for learners.

Learning Intentions

We are learning about Python, specifically to:

- understand Python data types and data structures that are important for data science
- manipulate strings
- create and call Python functions
- call a Python object's methods and access its properties
- perform a sequence of operations using method chaining

Success Criteria

I can *describe* the key Python data types and data structures that are important for data science.

I can *manipulate* strings.

I can *create* and *call* Python functions.

I can *call* a Python object's methods and *access* its properties.

I can *perform* a sequence of operations using method chaining.

Knowledge Prerequisites

- Python programming to at least the level defined in SQA Computer Programming Level 5 (HY2C 45)
- How to use a Jupyter notebook to write, edit and run Python code
- Introduction to Python for Data Science (Part 1)

For learners who do not have the prerequisite Python skills, a list of external learning resources that could be used to develop these skills is provided in the Python Learning Resources section below.



Lesson Requirements

	PDA	NPA	Data Skills for work
Qualification	Yes	Yes	Yes
Outcome ID(s)	WD7.3a, WD8.1j	DS5.1d, DC6.2b	A2.1, A5.1, A5.2
Outcome description(s)	WD7.3a Types of software for data analysis WD8.1j Tools for data analysis	DS5.1d Describe the tools that can be used at each stage in the life cycle DC6.2b Explain how data can be analysed and the tools that can be used to perform analysis	A2.1 Use of tools to analyse data A5.1 Use of programming languages A5.2 Programming for analysis
Level	7, 8	5, 6	Analysis
Software language	Python	Python	Python
Required equipment /software for student	Lesson: PowerPoint Python notebook: Jupyter notebook environment	Lesson: PowerPoint Python notebook: Jupyter notebook environment	Lesson: PowerPoint Python notebook: Jupyter notebook environment

Jupyter Notebook

There is a Jupyter notebook for this lesson that provides examples and programming tasks for learners, drawn from the examples in the lesson PowerPoint.

The notebook uses Python 3.x and the following packages:

- [platform](#) – provides access to underlying platform’s identifying data
- [pandas](#) - for data manipulation
- [s3fs](#) - an API to AWS S3 (Simple Storage Service), used to import datasets

The tasks are described in the table below.

Notebook section	Task	Description

Data Types	Task 1 - What's My Type?	Use the type() function to determine the data type of a value.
	Task 2 - Just My Type	Use the type() function to determine the data type of a variable.
Strings	Task 3 - Just a Letter	Extract a character from a string using an index.
	Task 4 - Another Letter	Extract a character from a string using an index.
	Task 5 - Slice	Extract a sequence of characters using the slice operator, ..
Lists	Task 6 - These are No Ordinary Foods...	<ul style="list-style-type: none"> • create a list • access the value at a specific index in the list • append an item to the list • calculate the length of the list • use online documentation
Dictionaries	Task 7 - What Ever Happened to Those (Super) Heroes?	<ul style="list-style-type: none"> • create a dictionary • access a value in the dictionary • hypothesise about what will happen if you try to access a dictionary with a key that does not exist • debug a simple error message • look at built-in and online sources of information that are useful when debugging • use online documentation
Functions	Task 8 - Python Rocks	Write a simple function which takes no arguments and returns a string.
	Task 9 - Numbers Numbers Everywhere	Write a simple function which takes 2 numeric arguments and returns a number.
	Task 10 - Show Me the Number	Call the function written in Task 9 and display its return value.
	Task 11 – System Information	<ul style="list-style-type: none"> • use online reference documentation to undertake simple research on a Python module • call a function in a module

		<ul style="list-style-type: none"> display the return value of the function.
Use Objects	Task 12 - Animals Sorted	<ul style="list-style-type: none"> use online reference documentation to undertake simple research on a Python data type (lists) sort a list by calling a list method
	Task 13 - Is it Empty?	<ul style="list-style-type: none"> use online reference documentation to undertake simple research on a pandas data type (DataFrame) access a pandas DataFrame property
	Task 14 - Tail	<ul style="list-style-type: none"> use online reference documentation to undertake simple research on a pandas method (tail()) call the pandas method
	Task 15 - Pipe Dreams	<ul style="list-style-type: none"> use string methods to gradually build a 'method chain' use the type() function to determine the data type of variables

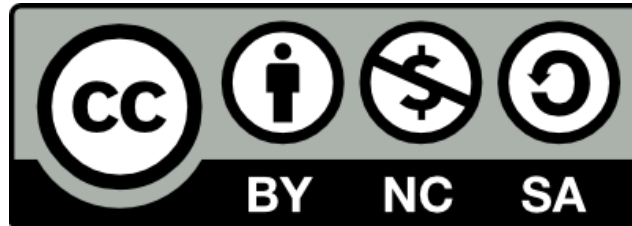
Python Learning Resources

- Python For Everybody (<https://www.py4e.com/html3/>)
- W3 Schools Python Tutorial (<https://www.w3schools.com/python/>)
- Data Carpentry Introduction to Python (<https://swcarpentry.github.io/python-novice-gapminder/>)
- Python.Land (<https://python.land/>)

How you can use this lesson

This lesson has been created by Effini in partnership with Data Education in Schools, The Data Lab and Data Skills for Work, with funding from the Scottish Government.

© 2021. This work is licensed under a [CC BY-NC-SA 4.0 license](https://creativecommons.org/licenses/by-nc-sa/4.0/).



You are free to:

- **Share** – copy and redistribute the material in any medium or format
- **Adapt** – remix, transform and build upon the material

Under the following terms:

- **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **NonCommercial** — You may not use the material for [commercial purposes](#).
- **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.