

---

# **hydrosoc-2018 Documentation**

*Release 1.0.0*

**Mike Kittridge**

**Dec 02, 2018**



---

## Sections

---

<b>1</b>	<b>Intended audience</b>	<b>3</b>
<b>2</b>	<b>Course summary</b>	<b>5</b>
<b>3</b>	<b>Prerequisites</b>	<b>7</b>
<b>4</b>	<b>Registration</b>	<b>9</b>
<b>5</b>	<b>Instructors</b>	<b>11</b>
5.1	Date and Location . . . . .	11
5.2	Prerequisites . . . . .	12
5.3	Schedule and Modules . . . . .	12
5.4	License and terms of usage . . . . .	13



The volume of data collected and available to scientists has grown in magnitudes over the past decades. It has grown to an extent that cannot be effectively handled by spreadsheets passed from one person to another. Fortunately, the toolsets to handle large amounts of data have grown as well, although it can be quite hard to keep up with the rapid developments of these toolsets.

The Python programming language is one of the most popular languages for both general purposes and scientific. This is due in part by the ease of use, open source code, and the large open source community that has developed a number of professional level toolsets for a wide range of applications.

The general goal of this workshop is to teach Python tools specifically beneficial for natural/environmental scientists in New Zealand handling large amounts of data. This will be accomplished through a combination of practical exercises and presentations.



# CHAPTER 1

---

## Intended audience

---

The intended audience for the workshop are people with very little to some experience with programming (Python or otherwise). Those people with a lot of Python programming experience will not likely get much out of the workshop unless they have not used the [Pandas package](#) in the past.





## CHAPTER 2

---

### Course summary

---

The workshop will cover the fundamental handling of tabular data and the associated processing and analysis tools. We will be primarily using the toolset contained within the [Pandas package](#). This will include reading/writing data, indexing, reshaping, computations, joining tables, time series handling, and visualization.



## CHAPTER 3

---

### Prerequisites

---

The **main prerequisite** for the workshop is the [Introduction to Python](#) course by the Monash University. The links provided by each chapter can be run independently without the installation of Python on your PC. Please go through at least the first 5 chapters before coming to the workshop. The last two are optional but recommended.



## CHAPTER 4

---

### Registration

---

Please contact the instructors via the email addresses below to sign up for the workshop. There will be a maximum of 30 attendees for the workshop. Suggestions for advanced topics or examples are welcome.



- Mike Kittridge
  - Senior Scientist - Hydrologist
  - Environment Canterbury
  - [mike.exner-kittridge@ecan.govt.nz](mailto:mike.exner-kittridge@ecan.govt.nz)
- Wilco Terink
  - Senior Scientist - Hydrologist
  - Environment Canterbury
  - [wilco.terink@ecan.govt.nz](mailto:wilco.terink@ecan.govt.nz)

## 5.1 Date and Location

### 5.1.1 Date

The workshop will take place at 9:00 on Monday Dec 3rd, 2018. It will likely run until 16:00-17:00 with a morning and an afternoon tea break and lunch. Go to the section *Schedule and Modules* for more details.

### 5.1.2 Location

The location of the workshop will be at Environment Canterbury (200 Tuam Street, Christchurch) in the Waimakariri Room. For public transportation options from the airport to ECan, use [this google map link](#).

## 5.2 Prerequisites

### 5.2.1 Python fundamentals

The **main prerequisite** for the workshop is the [Introduction to Python](#) course by the Monash University. The links provided by each chapter can be run independently without the installation of Python on your PC. Please go through at least the first 5 chapters before coming to the workshop. The last two are optional but recommended.

### 5.2.2 Python environments

#### Using the Jupyter Notebooks for the workshop modules

This workshop uses self contained code sets called [Jupyter Notebooks](#). The workshop will not explicitly require you to install python on your PC, but you are welcome to try as described in the next paragraph. Consequently for the workshop, the preferred method to run the notebooks will be through the [binder](#) links that build the correct python environment for the notebooks to be run under. This ensures that no one will have issues properly running the notebooks.

**Please run through** the short notebook [A quick tour of Jupyter/IPython Notebooks](#). It will familiarise you with Jupyter notebooks and some of its capabilities, but don't worry if you don't understand everything.

#### Installing a Python environment

If you would like to install your own Python environment, then please read through and follow the documentation in [Intro to Python for Environmental Scientists](#) especially as it relates to installing Anaconda/Miniconda and creating multiple python environments. If you feel like you'd like to give it a try, download the [environment.yml](#) from this workshop's github repo and run the following line from the Anaconda prompt:

```
conda env create -f environment.yml
```

It might take a couple minutes, but just be patient...

## 5.3 Schedule and Modules

Through much internal debate, we've decided to utilise a couple pre-existing Python courses that are very well made. All are completely self-contained and the Pandas fundamentals include an accompanying Youtube video series associated with the exercises. Pandas fundamentals use the fantastic open-access course provided by the [Data School](#).

### 5.3.1 Schedule

Time	Module
9:00 - 10:30	<a href="#">Introduction to Python recap</a>
10:30 - 10:45	BREAK
10:45 - 12:15	<a href="#">Pandas fundamentals</a>
12:15 - 13:15	LUNCH
13:15 - 14:45	<a href="#">Visualisation and Time series</a>
14:45 - 15:00	BREAK
15:00 - 16:30	To be determined...



### 5.3.2 Post-workshop exercises and/or courses

- More Pandas functionality
  - Best practices with Pandas
  - Official Pandas docs
- Python and GIS spatial analysis
  - Automating GIS Processes (Fantastic course!)

### 5.3.3 Reference material

- Basic Python
  - Python Quick Reference (Jupyter Notebook)
  - Introduction to Python for Environmental Scientists
  - General Introduction to Python
- Pandas
  - Pandas cheat sheet
  - Official Pandas docs

### 5.3.4 Glossary of terms

Python 3 glossary

## 5.4 License and terms of usage

This package is licensed under the terms of the Apache License Version 2.0 and can be found on the [GitHub project page](#).