

```
import numpy as np
from scipy import linalg

("mcycle.csv")

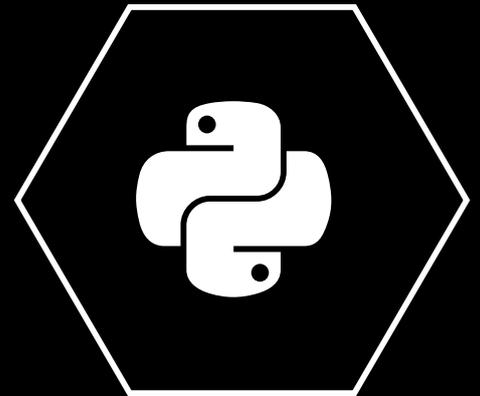
, w):
w)

X = np.stack((np.ones(x.shape), x), 1)
return linalg.inv(X.transpose()@W@X)@X.transpose()@y

def predict_wls(x, y, w, x0):
return np.sum(weighted_ls(x, y, w) * np.array([1, x0]))
```

# Data Programming in Python

Course for non-student learners  
(Full course, 11 weeks)  
Course information sheet 2019-20



The course introduces learners to object-oriented programming, the programming language Python and its use for data programming and analytics.

## Prerequisite Knowledge

The course requires no prerequisite knowledge. However, some of the examples require a basic understanding of matrix algebra and statistics.

The course is suitable for learners with no prior experi-

ence of programming, however the course advances at a brisk pace. Learners with no prior experience in programming should expect a larger time commitment in order to fully benefit from the course.

## Intended Learning Outcomes

By the end of this course learners will be able to:

- design and implement functions and classes in Python;
- make efficient use of the data structures built into Python, such as lists;
- describe and exploit features of object-oriented design such as polymorphism and inheritance
- implement data management and visualisation tasks in Python; and
- implement data-analytic tasks in Python using external libraries such as scikit-learn, NumPy/SciPy and pandas

## Syllabus

### Week 1

- Installing Anaconda Python
- Overview over front ends
- Overview of distinctive features of Python

### Week 2

- Data types in Python
- Strings
- Control structures: if, for and while

### Week 3

- Data frames
- Transforming, subsetting and merging data frames
- Reading/writing data from/to files

### Week 4

- List, tuples and sets
- Dictionaries
- Comprehensions

### Week 5

- Object-oriented programming and duck typing
- Creating classes

### Week 6

- Working with vectors and matrices in NumPy
- Linear algebra in NumPy and SciPy

### Week 7

- Pandas Series
- Pandas DataFrames
- Data manipulation in pandas

### Week 8

- Efficient methods for data management in pandas
- Merging, grouping and summarising data in pandas

### Week 9

- Plotting using matplotlib
- Data visualisation using seaborn and the plotting functions in pandas.

### Week 10

- Simple statistical inference using SciPy.
- Fitting regression models using statsmodels.

### Week 11

- Fitting machine learning models using scikit-learn
- Pre-processing data for machine learning models
- Creating pipelines

## Online Learning

The course is delivered by on-line distance learning and consists of

- Weekly live sessions with tutor(s)
- Weekly learning material (reading material, videos, exercises with model answers)
- Bookable one-to-one sessions with tutor(s)

## Textbooks

M. Lutz. Learning Python. O'Reilly.

A. B. Downey. Think Python. O'Reilly.

<https://greenteapress.com/wp/think-python-2e>

J. Vanderplas. Python Data Science Handbook. O'Reilly.

<https://jakevdp.github.io/>

[PythonDataScienceHandbook/](#)  
W. McKinney. Python for Data Analysis. O'Reilly.

## Hardware and Software Requirements

To take our courses all you need is a computer with an internet connection, an up-to-date version of a standard browser (such as Google Chrome, Firefox, Safari, Internet Explorer or Microsoft Edge) and a PDF reader (such as Acrobat Reader).

We encourage learners to install Anaconda Python and provide detailed installation instructions, but learners can also use free cloud-based services (Google Colab) if they prefer not to install software on their computer.

Learners need to install Zoom for participating in videoconferencing sessions. We recommend the use of a headset for videoconferencing sessions.

## Assessment

Courses for non-student learners have quizzes which allow you to check your progress. You need to attempt at least one quiz to obtain a certificate of attendance.

DATA ANALYTICS   
GLASGOW

School of Mathematics and Statistics

University of Glasgow

<http://gla.ac.uk/mdatagov>

<http://gla.ai>

Email:

[maths-stats-analyticscpd@glasgow.ac.uk](mailto:maths-stats-analyticscpd@glasgow.ac.uk)