

```
import numpy as np
from scipy import linalg

("mcycle.csv")

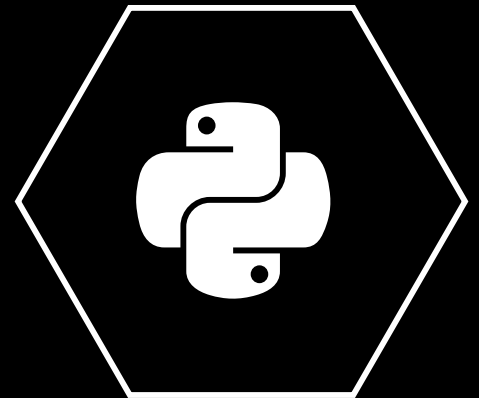
, w):
w)

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 (Part 2/2)

Data Science in Python
Course for non-student learners
(Half course, 6 weeks)
Course information sheet 2019-20



The course trains learners in the use of Python for programming and analytics.

Prerequisite Knowledge

The course requires learners to have basic Python programming skills. Learners with no prior experience of programming in Python should take part 1 of this course first.

Intended Learning Outcomes

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

- 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

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

Week 2

- Pandas Series
- Pandas DataFrames
- Data manipulation in pandas

Week 3

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

Week 4

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

Week 5

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

Week 6

- 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

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

[https://jakevdp.](https://jakevdp.github.io/)

[github.io/](https://jakevdp.github.io/)

[PythonDataScienceHandbook/](https://jakevdp.github.io/PythonDataScienceHandbook/)

W. McKinney. Python for Data Analysis. O'Reilly.

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.

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.

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