



Title:	Scientific Prog in Python APPROVED
Long Title:	Scientific Prog in Python
Module Code:	COMP8060
Credits:	5
NFQ Level:	Advanced
Field of Study:	Computer Science
Valid From:	Semester 1 - 2018/19 (September 2018)
Module Delivered in	2 programme(s)
Module Coordinator:	Donna OShea
Module Author:	Ted Scully
Module Description:	In this module, the learner will use the Python programming language to manipulate, manage and process data. More specifically, statistical and numerical libraries will be applied to analyse and manipulate complex data sets. The learner will also use Linux commands to perform basic system and file operations.
Learning Outcomes	
<i>On successful completion of this module the learner will be able to:</i>	
LO1	Apply Linux commands to perform basic system and file operations.
LO2	Implement a program to solve data-driven problems by applying standard programming concepts.
LO3	Utilise standard programming libraries and their associated functionality to perform analysis of datasets.
LO4	Apply programming techniques to clean, transform and query data.
LO5	Illustrate data sets using appropriate visualisation techniques.
Pre-requisite learning	
Module Recommendations	
<i>This is prior learning (or a practical skill) that is strongly recommended before enrolment in this module. You may enrol in this module if you have not acquired the recommended learning but you will have considerable difficulty in passing (i.e. achieving the learning outcomes of) the module. While the prior learning is expressed as named CIT module(s) it also allows for learning (in another module or modules) which is equivalent to the learning specified in the named module(s).</i>	
No recommendations listed	
Incompatible Modules	
<i>These are modules which have learning outcomes that are too similar to the learning outcomes of this module. You may not earn additional credit for the same learning and therefore you may not enrol in this module if you have successfully completed any modules in the incompatible list.</i>	
No incompatible modules listed	
Co-requisite Modules	
No Co-requisite modules listed	
Requirements	
<i>This is prior learning (or a practical skill) that is mandatory before enrolment in this module is allowed. You may not enrol on this module if you have not acquired the learning specified in this section.</i>	
No requirements listed	
Co-requisites	
No Co Requisites listed	

Module Content & Assessment

Indicative Content

Linux

Introduction to Linux file system and commands. Command line overview. Directory navigation, creation of directories and files, manipulating files, permissions, regular expression, process management, editors.

Programming Concepts:

Categories of programming language, their typical application, programming in an analytical and scientific context, Language, Syntax, error checking and debugging, variables and basic data types, lists, dictionaries and sets. Processing data structures: Conditionals and Loops. Efficient code structure: functions, modules, packages and files. Engineering code: objects, classes, and Object Oriented Programming (OOP). Introduction to version control, repositories, branching, commits and merging.

Data Array Manipulation

Overview of standard programming libraries for numerical computation. Creating multi-dimensional arrays. Performing operations such as indexing, slicing, boolean indexing, fancy indexing, building queries, transposing and applying conditional logic to arrays.

Numerical and Scientific Computing:

Use of open source numerical and scientific libraries for: performing mathematical and statistical routines, filtering, sorting, ranking, function application and mapping. Data loading, storage and different file formats. Data wrangling in the form of cleaning, transforming, merging and reshaping data. Data aggregation and grouping operations.

Visualisation

Visualisation and interpretation of data using 2D and 3D plots such as libraries such as Matplotlib and Seaborn. Line graphs, bar plots, violin plots, KDE plots, heatmaps, etc.

Python Packages

NumPy, SciPy, Matplotlib, Seaborn, Pandas

Assessment Breakdown

%

Course Work

100.00%

Course Work

Assessment Type	Assessment Description	Outcome addressed	% of total	Assessment Date
Practical/Skills Evaluation	Laboratory exam that assesses proficiency in applying Linux commands.	1	20.0	Week 3
Practical/Skills Evaluation	Practical lab assessment focused on building a program that applies a range of programming concepts and libraries to solve data driven problems.	2,3,4	40.0	Week 11
Project	Develop a program to solve specific data-driven problems by using numerical libraries and techniques and applying appropriate visualisations.	3,4,5	40.0	Sem End

No End of Module Formal Examination

Reassessment Requirement

Coursework Only

This module is reassessed solely on the basis of re-submitted coursework. There is no repeat written examination.

The institute reserves the right to alter the nature and timings of assessment

Module Workload

Workload: Full Time				
<i>Workload Type</i>	<i>Workload Description</i>	<i>Hours</i>	<i>Frequency</i>	<i>Average Weekly Learner Workload</i>
Lecture	Module Content delivery	2.0	Every Week	2.00
Lab	Laboratory Practical	2.0	Every Week	2.00
Independent Learning	Student undertakes independent study and develops programming skills	3.0	Every Week	3.00
Total Hours				7.00
Total Weekly Learner Workload				7.00
Total Weekly Contact Hours				4.00

Workload: Part Time				
<i>Workload Type</i>	<i>Workload Description</i>	<i>Hours</i>	<i>Frequency</i>	<i>Average Weekly Learner Workload</i>
Lecture	Module Content delivery	2.0	Every Week	2.00
Lab	Laboratory practical	2.0	Every Week	2.00
Independent Learning	Student undertakes independent study and develops programming skills	3.0	Every Week	3.00
Total Hours				7.00
Total Weekly Learner Workload				7.00
Total Weekly Contact Hours				4.00

Module Resources

Recommended Book Resources

- Wes McKinney 2017, *Python for Data Analysis*, 2nd Ed., O' Reilly Media [ISBN: 9781491957660]
- Joel Grus 2015, *Data Science from Scratch: First Principles with Python*, 1st Ed. [ISBN: 9781491901427]

Supplementary Book Resources

- Philipp K. Janert, *A hands-on guide for programmers and data scientists: Data Analysis with Open Source Tools*, O' Reilly [ISBN: 9780596802356]

This module does not have any article/paper resources

Other Resources

- Website: *Python Documentation*
<http://www.python.org/doc/>
- Website: *Python Data Analysis Library*
<http://pandas.pydata.org/>
- Website: *Python Scientific Library*
<http://www.scipy.org/>
- Website: *Python Numerical Library*
<http://www.numpy.org/>
- Website: *Python Visualisation Library*
<http://matplotlib.org/>
- Website: *Python Interactive Computing*
<http://ipython.org/>

Module Delivered in

Programme Code	Programme	Semester	Delivery
CR_SDAAN_8	<u>Higher Diploma in Science in Data Science & Analytics</u>	1	Mandatory
CR_SDAAN_9	<u>Master of Science in Data Science & Analytics</u>	1	Mandatory