

A8DP: Data Programming

Module Code:	A8DP
Long Title	Data Programming APPROVED
Title	Data Programming
Module Level:	LEVEL 8
EQF Level:	6
EHEA Level:	First Cycle
Credits:	5
Module Coordinator:	MICHAEL BRADFORD
Module Author:	Madita Feldberger
Departments:	School of Computing
Specifications of the qualifications and experience required of staff	
Learning Outcomes	
<i>On successful completion of this module the learner will be able to:</i>	
#	Learning Outcome Description
LO1	Design algorithms and implement key programming patterns and constructs for data analytics
LO2	Apply practical skills to process data and access databases using a professional tool/language of data analytics (e.g., R)
LO3	Assess the challenges associated with processing big data datasets, and compare and contrast programming for big data vis-à-vis programming for conventional datasets
LO4	Assess methods and practices for software development in order to design and implement data programming requirements
Dependencies	
Module Recommendations	
No recommendations listed	
Co-requisite Modules	
No Co-requisite modules listed	
Entry requirements	

A8DP: Data Programming

Module Content & Assessment			
Indicative Content			
Data Programming and Software Development Algorithm design Program I/O, Data types and data structures, Program control and process models, Programming constructs, Programming types (imperative, declarative, functional, logic) Programming languages for data analytics (e.g., R, Python), Developing programs for data processing activities (e.g., data extraction, cleaning, merging, aggregation, analysis, reporting)			
Data Structures for Analytics Data structures facilitating data analysis, Data structures facilitating data analysis Use of support libraries (e.g., R packages, Pandas),			
Database Programming Database system technologies Programmatically connecting to databases Create/Read/Update/Delete (CRUD) Operations			
Big Data Programming Challenges associated with programming for big data, Parallelism for computational processes, Distributed computing. Utilisation of cloud computing platforms for big data processing			
Assessment Breakdown			%
Coursework			100.00%
Assessments			
Full Time			
Coursework			
Assessment Type:	Formative Assessment	% of total:	Non-Marked
Assessment Date:	n/a	Outcome addressed:	
Non-Marked:	Yes		
Assessment Description: Formative assessment will be included by the provision of class case studies and short answer questions. Feedback will be provided individually or as a group in written and oral format, or on-line through Moodle. In addition, in class discussions will be undertaken as part of the practical approach to learning.			
Assessment Type:	CA 1 (0380)	% of total:	15
Assessment Date:	n/a	Outcome addressed:	1,2,4
Non-Marked:	No		
Assessment Description: The first test will assess apprentices' knowledge of data types and programming I/O.			
Assessment Type:	CA 2 (0390)	% of total:	15
Assessment Date:	n/a	Outcome addressed:	1,2,4
Non-Marked:	No		
Assessment Description: The second test will assess apprentices' knowledge of programming control constructs.			
Assessment Type:	CA 3 (0420)	% of total:	15
Assessment Date:	n/a	Outcome addressed:	1,2,4
Non-Marked:	No		
Assessment Description: The third test will assess apprentices' knowledge of programming data processing activities and software design.			
Assessment Type:	Continuous Assessment (0200)	% of total:	15
Assessment Date:	n/a	Outcome addressed:	1,2,4
Non-Marked:	No		
Assessment Description: The fourth test will assess apprentices' knowledge and programming data operations with database connectivity			
Assessment Type:	Project (0050)	% of total:	40
Assessment Date:	n/a	Outcome addressed:	3
Non-Marked:	No		
Assessment Description: The project will entail the apprentice conducting research to review the current state of programming technologies and computing environments that facilitate data processing activities for datasets of varying types and sizes.			
No End of Module Assessment			
No Workplace Assessment			
Reassessment Requirement			
Coursework Only <i>This module is reassessed solely on the basis of re-submitted coursework. There is no repeat written examination.</i>			

A8DP: Data Programming

Module Workload

Module Target Workload Hours 0 Hours

Module Resources	
<i>Recommended Book Resources</i>	
<p>Marz N. and Warren J.. (2013), BIG DATA : PRINCIPLES AND BEST PRACTICES OF SCALABLE REALTIME DATA SYSTEMS., Manning Publications, [ISBN: 9781617290343].</p> <p>Paul Teetor. R Cookbook, O'Reilly Media, p.436, [ISBN: 0596809158].</p> <p>Wes McKinney. (2017), Python for Data Analysis: Data Wrangling with Pandas, O'Reilly Media.</p> <p>Edward Capriolo, Dean Wampler, Jason Rutherglen. (2012), Programming Hive, O'Reilly Media, p.200, [ISBN: 9781449319335].</p>	
<i>Supplementary Book Resources</i>	
<p>Thomas A. Runkler. (2012), Data Analytics, Vieweg+Teubner Verlag, p.147, [ISBN: 9783834825889].</p> <p>Tom White. (2012), Hadoop: The Definitive Guide, O'Reilly Media, p.625, [ISBN: 1449311520].</p> <p>Jimmy Lin, Chris Dyer, Graeme Hirst (Series Editor). (2010), Data-Intensive Text Processing with MapReduce, Morgan and Claypool Publishers, p.178, [ISBN: 9781608453429].</p> <p>Boris Lublinsky, Kevin T. Smith, Alexey Yakubovich.. (2013), Professional hadoop solutions, Chichester; John Wiley and Sons, p.1 online resource (505) :, [ISBN: 9781118611937].</p> <p>Alex Holmes. (2012), Hadoop in Practice, Manning Publications, p.536, [ISBN: 9781617290237].</p> <p>Alan Gates. (2011), Programming Pig, O'Reilly Media, p.400, [ISBN: 9781449302641].</p>	
<i>This module does not have any article/paper resources</i>	
<i>Other Resources</i>	
<p>[Website], MIT Open Courseware videolectures.net. (2005), Introduction to Algorithms, http://videolectures.net/mit6046jf05_introduction_algorithms/</p> <p>[Website], MIT Open Courseware. Introduction to Computer Science and Programming, http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-00sc-introduction-to-computer-science-and-programming-spring-2011/index.htm</p>	
Discussion Note:	