

H9PDA: Programming for Data Analytics

| | |
|--|---|
| Module Code: | H9PDA |
| Long Title | Programming for Data Analytics SUPERSEDED |
| Title | Programming for Data Analytics |
| Module Level: | LEVEL 9 |
| EQF Level: | 7 |
| EHEA Level: | Second Cycle |
| Credits: | 10 |
| Module Coordinator: | MICHAEL BRADFORD |
| Module Author: | MICHAEL BRADFORD |
| 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 | Analyse, compare, contrast and critically evaluate the characteristics of programming languages and programming environments commonly utilised for data analytics solution implementation |
| LO2 | Critically assess the challenges associated with processing big data datasets and compare and contrast programming for big data vis-à-vis programming for conventional datasets |
| LO3 | Determine algorithm complexity and develop cost functions associated with data intensive problem solutions |
| LO4 | Evaluate, develop and implement solutions for processing datasets and solving complex problems in distributed computing and cloud computing environments using relevant programming paradigms (e.g., MapReduce, parallelism), relevant programming languages (e.g., Pig, Hive), and associated tools and techniques (e.g., data compression). |
| Dependencies | |
| Module Recommendations | |
| No recommendations listed | |
| Co-requisite Modules | |
| No Co-requisite modules listed | |
| Entry requirements | |

H9PDA: Programming for Data Analytics

| Module Content & Assessment | | | |
|--|------------------------------|---------------------------|----------|
| Indicative Content | | | |
| No indicative content | | | |
| Assessment Breakdown | | | % |
| Coursework | | | 100.00% |
| Assessments | | | |
| Full Time | | | |
| Coursework | | | |
| Assessment Type: | Continuous Assessment (0200) | % of total: | 20 |
| Assessment Date: | n/a | Outcome addressed: | 1,2,3 |
| Non-Marked: | No | | |
| Assessment Description: May be assessed through continuous assessment in which learners will be required to conduct research and provide reviews regarding the characteristics of programming languages, environments, and technologies utilised in the field of data analytics. Learners may also be assessed during practical sessions in which particular problems are set as a challenge for which learners will be required to develop and present solutions. | | | |
| Assessment Type: | Practical (0260) | % of total: | 20 |
| Assessment Date: | n/a | Outcome addressed: | 1,2,3,4 |
| Non-Marked: | No | | |
| Assessment Description: n/a | | | |
| Assessment Type: | Project | % of total: | 60 |
| Assessment Date: | n/a | Outcome addressed: | 1,2,3,4 |
| Non-Marked: | No | | |
| Assessment Description: May be assessed through a project in which learners must identify and source a large set of raw data design, develop, implement, and document a process for efficiently processing and analysing the data to answer a novel question utilising a distributed computing environment and appropriate programming languages present project work | | | |
| No End of Module Assessment | | | |
| No Workplace Assessment | | | |
| Reassessment Requirement | | | |
| Repeat examination <i>Reassessment of this module will consist of a repeat examination. It is possible that there will also be a requirement to be reassessed in a coursework element.</i> | | | |

H9PDA: Programming for Data Analytics

| Module Workload | | | | |
|---|-----------------------------|--------------|-------------------|--|
| Module Target Workload Hours 0 Hours | | | | |
| Workload: Full Time | | | | |
| <i>Workload Type</i> | <i>Workload Description</i> | <i>Hours</i> | <i>Frequency</i> | <i>Average Weekly Learner Workload</i> |
| Lecture | No Description | 2 | Every Week | 2.00 |
| Practical | No Description | 2 | Every Week | 2.00 |
| Assignment | No Description | 17 | Once per semester | 1.42 |
| 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 | No Description | 2 | Every Week | 2.00 |
| Practical | No Description | 2 | Every Week | 2.00 |
| Assignment | No Description | 17 | Once per semester | 1.42 |
| Total Weekly Contact Hours | | | | 4.00 |

Module Resources

Recommended Book Resources

- Marz N. and Warren J.. (2013), Big Data: Principles and best practices of scalable realtime data systems, Manning Publications, [ISBN: 13:978-16172].
- Lublinsky B., Smith K. T. and Yakubovich A. (2013), Professional Hadoop Solutions, Wrox, [ISBN: 13:978-11186].
- Holmes A. (2012), Hadoop in Practice, Manning Publications, [ISBN: 13:978-16172].
- McKinney W.. (2012), Python for Data Analysis, O'Reilly Media, [ISBN: 13: 978-14493].

Supplementary Book Resources

- Runkler T.A.. (2012), Data Analytics: Models and Algorithms for Intelligent Data Analysis, Vieweg+Teubner Verlag, [ISBN: 13:978-38348].
- Tom White. (2012), Hadoop: The Definitive Guide, O'Reilly Media/Yahoo Press, p.625, [ISBN: 1449311520].
- Lin J. and Dyer C.. (2010), Data-Intensive Text Processing with MapReduce, Morgan and Claypool Publishers, [ISBN: 1397816084].
- Gates A. (2011), Programming Pig, O'Reilly Media, [ISBN: 13: 978-14493].
- Capriolo E. and Wampler D.. (2012), Programming Hive, O'Reilly Media, [ISBN: 13: 978-14493].

This module does not have any article/paper resources

Other Resources

- [website], MIT Open Courseware videolectures.net. (2005), Introduction to Algorithms, http://videolectures.net/mit6046jf05_introduction_algorithms/
- [website], Cloudera University. The Apache Hadoop Ecosystem, <http://university.cloudera.com/onlineresources/hadoopecosystem.html>

Discussion Note: