Module Descriptor

Title	Advanced Data Science		
Session	2024/25	**Status**	Published
Code	COMP11068	**SCQF Level**	11
Credit Points	20	**ECTS (European Credit Transfer Scheme)**	10
School	**Computing, Engineering and Physical Sciences**		
Module Co-ordinator	Tahir Mahmood		

Summary of Module

The connected world generates large volume of data that need to be understood and analysed to develop insight in particular application areas for instance financial sector, healthcare, market analysis, community behaviours to name a few. With the adoption of smart devices and ubiquitous deployment of sensing capabilities around us in the form of Internet-of-Things (IoT), data is becoming increasing important. Data from various modalities ranging from personal devices to mass deployment of IoT are utilised to provision personalised services commonly known as data driven services. The efficacy of these services is greatly influenced by intelligence gained through data. The availability of high speed internet and connected devices generates data that need to be integrated, analysed and then fed to data driven services.

Data science is an emerging filed which offers innovation solutions catering all 4V's of big data; with its roots in statistical analysis and machine learning it offers powerful algorithms, methodologies, and tools to manage, transform and analyse big data. This module address advance topics of data science yet in incremental manner – starting from fundamentals of data analysis life cycle to data stream processing. This module is designed to cover theoretical knowledge of data processing, management and analysis to hands-on experience on big data frameworks used to crunch massive data sets.

The following topics will be covered in this module:

- Introduction to data science
- Data management for data science services and applications
- Data science at scale
- Data stream processing
- New developments in data science
- This module will work to develop a number of the key 'I am UWS' Graduate Attributes to make those who complete this module: Universal • Critical Thinker • Ethically-minded • Research-minded Work Ready • Problem-Solver • Effective Communicator • Ambitious Successful • Autonomous • Resilient • Driven

Module Delivery Method	On-Campus[1] ☐	Hybrid[2] ☒	Online[3] ☐	Work -Based Learning[4] ☐
Campuses for Module Delivery	☐ Ayr ☐ Dumfries	☐ Lanarkshire ☐ London ☒ Paisley	☐ Online / Distance Learning ☐ Other (specify)	

Terms for Module Delivery	Term 1	☒	Term 2	☐	Term 3	☐
Long-thin Delivery over more than one Term	Term 1 – Term 2	☐	Term 2 – Term 3	☐	Term 3 – Term 1	☐

Learning Outcomes	
L1	Demonstrate an extensive knowledge of the fundamental data science concepts and their usage for varied dataset types and volume.
L2	Demonstrate a comprehensive understanding of data analytics algorithms and libraries to design and develop data driven services.
L3	Analyse and apply widely used big data frameworks for various application domains and expected outcomes.
L4	Develop skills to make use of big data frameworks for data science pipelines supporting data driven services.
L5	

Employability Skills and Personal Development Planning (PDP) Skills	
SCQF Headings	**During completion of this module, there will be an opportunity to achieve core skills in:**
Knowledge and Understanding (K and U)	**SCQF 11** Extensive understanding of data science concepts. Comprehensive knowledge of data analytics life cycle and its application for various domains.
Practice: Applied Knowledge and Understanding	**SCQF 11** In-depth knowledge to various data analysis methodologies and data management platforms to design and develop data driven services.
Generic Cognitive skills	**SCQF 11**

[1] Where contact hours are synchronous/ live and take place fully on campus. Campus-based learning is focused on providing an interactive learning experience supported by a range of digitally-enabled asynchronous learning opportunities including learning materials, resources, and opportunities provided via the virtual learning environment. On-campus contact hours will be clearly articulated to students.

[2] The module includes a combination of synchronous/ live on-campus and online learning events. These will be supported by a range of digitally-enabled asynchronous learning opportunities including learning materials, resources, and opportunities provided via the virtual learning environment. On-campus and online contact hours will be clearly articulated to students.

[3] Where all learning is solely delivered by web-based or internet-based technologies and the participants can engage in all learning activities through these means. All required contact hours will be clearly articulated to students.

[4] Learning activities where the main location for the learning experience is in the workplace. All required contact hours, whether online or on campus, will be clearly articulated to students

	Critical and theoretical analysis of state-of-the-art presented in published technical reports and scholarly articles.
Communication, ICT and Numeracy Skills	**SCQF 11** Assignments and discussions session will assist students to work collaboratively and discuss possible solutions to a problem related to the knowledge they gained in lectures and lab sessions.
Autonomy, Accountability and Working with Others	**SCQF 11** Each student will generate a comprehensive report summarizing his/her finding for a given scenario. For groups based assignments, participants of each group will have to justifying their findings within a group discussion lead by a group leader, and finally submit a collective report – delineating their findings.

Prerequisites	**Module Code**	**Module Title**
	Other	
Co-requisites	**Module Code**	**Module Title**

Learning and Teaching

In line with current learning and teaching principles, a 20-credit module includes 200 learning hours, normally including a minimum of 36 contact hours and maximum of 48 contact hours.

This module comprises of lectures, labs, and assignments. The lectures will deliver fundamental knowledge of data science and application in various domains. Through lectures students will be able to learn various methodologies to prepare, persists and process data from varied sources in large volume. Furthermore these lectures will develop theoretical understanding of algorithms and frameworks used to process big data. Students will also be introduced to selected topics of data science research. Labs will help in developing in-depth understanding of the knowledge delivered in the lectures, and critical evaluation of algorithms and methodologies when applying for a specific problem.

The list of indicative lectures, labs and tutorials:

Lectures

• Introduction to data science – motivation, real world applications, use case studies

• Data analytics life cycle – discovery, preparation, planning, building, results and operationalization

• Data structures and data management for data science

• Programming for data science

• Programming libraries for data science

• Introduction to big data frameworks - abstraction, joins, implementation

• Data science as scale

• Data stream processing – concepts, limitations

• Data stream processing frameworks

• Research developments in data science

Labs and Tutorials

• Setting up NoSQL database; data modelling and querying

• Programming for data science

• Dataset manipulation

• Data analytics

• Big Data frameworks

• Data analytics with big data framework

Group / Individual assignments

• Data pipeline with real world dataset

Learning Activities During completion of this module, the learning activities undertaken to achieve the module learning outcomes are stated below:	**Student Learning Hours** (Note: Learning hours include both contact hours and hours spent on other learning activities)
Lecture / Core Content Delivery	20
Laboratory / Practical Demonstration / Workshop	16
Personal Development Plan	10
Independent Study	154
Please select	
Please select	
TOTAL	200

Indicative Resources

The following materials form essential underpinning for the module content and ultimately for the learning outcomes:

The following materials form essential underpinning for the module content and ultimately for the learning outcomes:

Lectures notes and hand-outs will be provided through moodle.

Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data by EMC Education Services; John Wiley & Sons.

Data Science from Scratch: First Principles with Python Paperback by Joel Grus

Python Programming for the Absolute Beginner 3rd Edition by Mike Dawson

Programming Python Paperback by Mark Lutz

Software Packages:

- PyCharm - IDE for Python, - Python 3.x or latest release - MongoDB 3.x or latest release - VirtualBox 5.x or latest release - Libraries NumPy, SciPy, Matplotlib, Pandas, scikit-learn
(N.B. Although reading lists should include current publications, students are advised (particularly for material marked with an asterisk*) to wait until the start of session for confirmation of the most up-to-date material)

Attendance and Engagement Requirements
In line with the Student Attendance and Engagement Procedure, Students are academically engaged if they are regularly attending and participating in timetabled on-campus and online teaching sessions, asynchronous online learning activities, course-related learning resources, and complete assessments and submit these on time. **For the purposes of this module, academic engagement equates to the following:** In line with the Academic Engagement Procedure, Students are defined as academically engaged if they are regularly engaged with timetabled teaching sessions, course-related learning resources including those in the Library and on the relevant learning platform, and complete assessments and submit these on time. Please refer to the Academic Engagement Procedure at the following link: Academic engagement procedure

Equality and Diversity
The University's Equality, Diversity and Human Rights Procedure can be accessed at the following link: UWS Equality, Diversity and Human Rights Code. The University policies on equality and diversity will apply to this module. In relation to students with special needs, when a student discloses a disability the individual module tutor, in consultation with the special needs coordinator will agree any appropriate adjustments to be made. Students should note that the language of instruction is English and that they will need to have a reasonable grasp of the language in order to understand the teaching materials. UWS Equality and Diversity Policy
(N.B. Every effort will be made by the University to accommodate any equality and diversity issues brought to the attention of the School)

Supplemental Information

Divisional Programme Board	**Computing**
Overall Assessment Results	☐ **Pass / Fail** ☒ **Graded**
Module Eligible for Compensation	☐ **Yes** ☐ **No** **If this module is eligible for compensation, there may be cases where compensation is not permitted due to programme accreditation requirements. Please check the associated programme specification for details.**
School Assessment Board	Business & Applied Computing
Moderator	Naeem Ramzan
External Examiner	R Menzies

Accreditation Details	
Module Appears in CPD catalogue	☐ **Yes** ☒ **No**
Changes / Version Number	1.06

Assessment (also refer to Assessment Outcomes Grids below)
Assessment 1
During the laboratory sessions each student will be required to successfully complete the tasks mentioned in the lab sheet (weighted 30%), consequently assessing the achievement of L1 and L2.
Assessment 2
Each student will select a topic for term project (weighted 70%) to analysis, design, and develop a data science pipeline using skills developed through lectures and lab sessions. Term project will evaluate L3 and L4. Topic of the term project must be agreed between the student and module coordinator/module tutor to ensure its relevance with the learning outcomes.
Assessment 3
(N.B. (i) Assessment Outcomes Grids for the module (one for each component) can be found below which clearly demonstrate how the learning outcomes of the module will be assessed. (ii) An indicative schedule listing approximate times within the academic calendar when assessment is likely to feature will be provided within the Student Module Handbook.)

Component 1							
Assessment Type	**LO1**	**LO2**	**LO3**	**LO4**	**LO5**	**Weighting of Assessment Element (%)**	**Timetabled Contact Hours**
Report of practical/ field/ clinical work	☒	☒	☐	☐	☐	30	0

Component 2							
Assessment Type	**LO1**	**LO2**	**LO3**	**LO4**	**LO5**	**Weighting of Assessment Element (%)**	**Timetabled Contact Hours**
Dissertation/ Project report/ Thesis	☐	☐	☒	☒	☐	70	0

Component 3							
Assessment Type	**LO1**	**LO2**	**LO3**	**LO4**	**LO5**	**Weighting of Assessment Element (%)**	**Timetabled Contact Hours**
	☐	☐	☐	☐	☐		
Combined total for all components						100%	hours

Change Control

What	When	Who