

Module 8: Applied Data Science

Stage	1						
Semester	2						
Module Title	Applied Data Science						
Module Number	8						
Module Status	Mandatory						
Module ECTS Credits	5						
Module NFQ level	9						
Pre-Requisite Module Titles	None						
Co-Requisite Module Titles	None						
Capstone Module	No						
List of Module Teaching Personnel	Mr Barry Denby Dr Waseem Akhtar Guest lecturers						
Contact Hours	Non-contact Hours					Total Effort (hours)	
36				64			100
Lecture	Practical	Tutorial	Seminar	Assignment	Placement	Independent Work	
24			12	30		34	
Allocation of Marks (Within the Module)							
	Continuous Assessment	Project	Practical	Final Examination	Total		
Percentage Contribution	100				100		

Intended Module Learning Outcomes

On successful completion of this module the learner will be able to:

1. Demonstrate “data-analytic thinking” using large datasets from the perspective of real world decision making and problem solving
2. Critically analyse existing big data gathering, storing and sharing practices taking ethical, privacy and their purpose into consideration
3. Employ communication and presentation skills, and tools and techniques, to present insights derived from the analysis of a large dataset that is appropriate for decision making and problem solving purposes
4. Implement a data visualisation system to present information inferred from a large dataset as part of a reporting mechanism.

Module Objectives

This module aims to introduce the learner to the fundamental principles of data science and equips them with “data-analytic thinking” necessary for extracting useful knowledge and business value from the relevant datasets. The module introduces the learner to the principles underpinning the processes and strategies necessary to solve real world problems through data science techniques. The module focuses on data science concepts as applied to practical real world problems and aims to teach learner the underlying concepts behind data science and most importantly how to approach and be successful at problem solving. Problem solving and information discovery strategies will be developed via in-depth analysis of existing Big Data implementations and case studies. As most of the information discovered from large datasets is of direct use to business decisions, both reporting and visualisation are an important element of this module.

Module Curriculum

- **Data for Competitive advantage**
How data science fits in the organisation and the competitive landscape, and in the field of social and scientific research, role of data analyst, how data science may lead to competitive advantage, how to sustain that competitive advantage, how to achieve scientific, tactical and strategic advantage, organisational strategy to gather clean, valid and complete data, what why and how to capture
- **Thinking data-analytically**
Gather data with right volume, variety and velocity, identify appropriate methods and strategies to store and analyse large datasets appropriate for pre-processing, processing, post-processing and visualising; considering data as asset, how to get best leverage form this asset, developing data analytics capabilities to harness the power of big data analytics,
- **Discovering and Presenting Knowledge**
Recognising opportunities and problems from a data perspective, formulating real world problems in terms of data and discovering solutions through data science, identifying appropriate modelling, knowledge discovery and presentation methods, tools and techniques, derive insights and actionable recommendations using appropriate communication and reporting techniques; e.g. the Weka Data Mining Workbench, R, iPython, panda
- **Data Visualisation,**
Identifying and evaluating suitable visualisation methods and data products; visual data analytic; information dashboards; business intelligence systems; Javascript, Tableau
- **Ethical Aspects**
Provenance, privacy, ethics, governance

- **Case Studies & what is out there**
In the area of business intelligence, social and scientific research, social networks & media data, census data and other large datasets and their use, twitter API, Google bigquery, Google Prediction API, cloud analytics
- **Guest Lectures**
Guest lectures from Big Data researchers

Reading Lists and other learning materials

Recommended Reading

Provost F, Fawcett T., 2013, *Data Science for Business: What you need to know about data mining and data-analytic thinking*, O Reilly Media

Granville V., 2014, *Developing Analytic Talent: Becoming a Data Scientist*, Wiley

Minelli M, Chambers M, Bhiraj A., 2013, *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business*, Wiley CIO

Secondary Reading

Lander J, 2013, *R For Everyone: Advanced Analytics and Graphics*, Addison-Wesley Data & Analytic Series

Davenport T, 2014, *Big Data at Work*, Harvard Business Review

Davenport T, Kim J H, 2013, *Keeping Up with the Quants: Your Guide to Understanding and Using Analytics*

Manoochehri M, 2013, *Data Just Right: Introduction to Large-Scale Data & Analytics*, Addison-Wesley Data & Analytic Series

Zumel N, Mount J, 2014, *Practical Data Science with R*, Manning Publication

Additional reading material will include peer reviewed research papers and most recent and relevant case studies.

Module Learning Environment

Accommodation

Lectures are carried out in class rooms / lecture halls in the College. Lab tutorials are carried out in computer labs throughout the Campus. All have the software required to deliver the programme.

Library

All learners have access to an extensive range of physical and electronic (remotely accessible) library resources. The library monitors and updates its resources on an on-going basis, in line with the College's Library Acquisition Policy. Lecturers update reading lists for this course on an annual basis as is the norm with all courses run by Griffith College.

Module Teaching and Learning Strategy

Each week involves both classes and practical laboratory sessions

Classes are used to deliver theoretical content and may be supported by online delivery of notes, examples, and web resources. Introduction of statistical analysis techniques are provided in-class theory and reinforced via worked example. All work is implemented on real-world data.

Laboratory Practical's are used to provide continuous progression of theory presented in lectures with each session building upon ideas of the previous lectures and laboratory sessions.

Module Assessment Strategy

Element	Weighting	Type	Description	Learning Outcomes Assessed
1	25%	Report	For this report learners are required to research a specific topic in Data Science (e.g. a business opportunity or problem, a social or scientific study) identifying and applying best practices, tools and techniques to address that topic. The report must clearly describe the overall objectives of the study, and relevant ethical and privacy issues.	1,2
2	25%	Programming Assignment	Implement informative visualisation of attributes of interest in a large dataset (e.g. a business intelligence dashboard, a scientific data visualisation)	3-4
3	50%	Programming, report writing assignment	Implementation of a complete data science project from start to finish; this will involve selecting a real world scientific or business problem; justifying how concepts of data science can be applied to address this problem; identifying the right dataset and analytics / mining techniques on this dataset (preferably use a dataset analysed as part of Data Mining Algorithms and Techniques module assignment) describe in details the insights gained through data analytics process; use appropriate visualisation techniques to visually present your findings; outline your recommendations useful for decision making; make a presentation	1, 2,3,4