Module Handbook

Data Driven Organisation

INF11XXX - Session 2016/2017 – Trimester 1

Module leader: Laura Muir

Contents

[About the Module 2](#_Toc449478655)

[What you will learn and what this is about 2](#_Toc449478656)

[Learning Outcomes 2](#_Toc449478657)

[Module Team 2](#_Toc449478658)

[Indicative References and Reading List 2](#_Toc449478659)

[Sample papers 3](#_Toc449478660)

[Lecture and Tutorial Activities 3](#_Toc449478661)

[Assessment Information 3](#_Toc449478662)

[Module plan 3](#_Toc449478663)

[Key dates 3](#_Toc449478664)

[Teaching schedule 5](#_Toc449478665)

# About the Module

## What you will learn and what this is about

On this module you will examine the principles behind the data driven organisation and the basis of analytics to support decision making. This will including the governance of data and the issues that face organisations in using data for decision making. You will also learn about the roles of governance frameworks, the professions, and the consequences of setting and implementing data policies and standards across an organisation. In terms of analytics you will learn how to devise analytic strategies and apply these to large and complex data sets. Focus is on the selection of appropriate methods, evaluating their performance and implementing and communicating solutions.

To achieve the above objectives you will gain an in-depth knowledge of approaches to the management of information, taking into account the human, organisational and technical, and strategic factors required to establish a cohesive data science programme, from a current to future state, supporting the organisation goals.

### Learning Outcomes

On completion of this module you will be able to:

LO1: Critically assess the roles and impact of ethics, governance and professionals in data analysis

LO2: Research and evaluate organisational drivers for advanced analytics

LO3: Develop and apply data analytics strategies and evaluate their organisational impact

LO4: Present and communicate the results of the above to an appropriate professional and academic standard.

## Module Team

The module is delivered by Calum Alexander [xxxx@napier.ac.uk](mailto:xxxx@napier.ac.uk) and Robert Raeside [r.raeside@napier.ac.uk](mailto:r.raeside@napier.ac.uk), extn 4308. Calum will be responsible for the organisation and business aspects of the module and Robert will be responsible for the statistical and R part of the module.

The module leader is Peter Cruickshank   
Room C56, extn 2309, email [p.cruickshank@napier.ac.uk](mailto:p.cruickshank@napier.ac.uk)

The module leader is Laura Muir, [l.muir@napier](mailto:l.muir@napier), Room C60, extn 2381 and non-technical queries about the module should be addressed to Laura

## Indicative References and Reading List

Because of the pace of change and the wide ranging nature of this course, there is no single course textbook. Instead, there is a selection of **directed reading** to accompany each lecture, the majority of which will be made available through Moodle. The following reading is essential:

* Building a Digital Analytics Organization: Create Value by Integrating Analytical Processes, Technology, and People into Business Operations (2013), Pearson, Judah Phillips, ISBN-13: 978-0-13-337278-6
* Data Science for Business: What you need to know about data mining and data-analytic thinking (1st Edition) (2014), O'Reilly, Foster Provost & Tom Fawcett, ISBN-13: 978-1449361327
* Big Data at Work: Dispelling the Myths, Uncovering the Opportunities (2014), Harvard Business Publishing, Thomas H. Davenport (Author) ISBM 987-1-4221-6816-5
* R for Data Science (2014), Packt Publishing Ltd, Birmingham, D. Toomey, ISBN-13: 978-1784390860
* Practical Data Science with R. (2014), N. Zumel and J Mount Manning Publications, NY, N. Zumel and J Mount ISBN-10: 1617291560
* Good Practice Team, Effective tables and graphs in official statistics, Government Statistical Service, <https://gss.civilservice.gov.uk/wp-content/uploads/2014/12/Effective-graphs-and-tables-in-official-statistics-version-1.pdf>, 2014

Mike 2.0 (Method for an Integrated Knowledge Environment) is an additional source throughout the course. Mike resources can be downloaded for free on registration from [mike2.openmethodology.org](http://mike2.openmethodology.org/), Extracts will be provided if necessary.

### 

# Lecture and Tutorial Activities

There is a lot of material to cover – you are expected to engage with the whole course. In particular, please note videos of all lectures are made available online. There are a number of optional online quizzes you can use to test your knowledge

Lectures are intended as gateways to topic areas, and tutorial exercises are essential to consolidate your understanding. It’s vital that you keep up to date with the reading, which is substantial, and which forms part of the independent study element on the module descriptor.

For **Blended Learning** students, face to face teaching takes place on three occasions though the Trimester. Distance students can access remote support using Moodle forums and email and Skype with the lecturer.

**Distance learning students** can access remote support using Moodle forums and email and regular Webex sessions with the lecturer.

Support is also available through using forum discussions and email.

# Assessment Information

The assessment is by two pieces of coursework.

The First Coursework is submitted in Week 5, should be around 3 000 words in length, and is an academic report worth 30% of the overall mark for the module. Detailed feedback will be provided to students in Week 8. This assessment will involve articulation of the business need for data analytics and a report on an analysis of a data set using R.

The Second Coursework is submitted in Week 14, and will be a case study presentation and report of around 5000 words in length, and is worth 70% of the overall module mark. This will be based on a mini project set in the students’ organisation and will involve formulating a business problem, data collection, analysis and an implementation report.

# Module plan

## Key dates

|  |  |
| --- | --- |
| Date 1 | Preparation material made available for day 1 (released). |
| Date 2 | Day 1 Materials. Coursework 1 hand out. |
| Date 3 | Preparation material made available for day 2 (released). |
| Date 4 | Day 2 Materials. Coursework 2 hand out. |
| Date 5 | Preparation material made available for day 3 (released). |
|  |  |
| Date 6 | Module submission: Presentation (Webcast, Slides + Report) |

Teaching days for Blended Mode students start at 9:30am and take place in room C18, Merchiston.

.

## Teaching schedule

*Lecture and other material will be made available in three blocks to allow you pace your study around other work and life commitments*

|  | Topic | Textbook reading. (NB Research papers will also be posted for you to read) |
| --- | --- | --- |
|  | Introduction to the module. Overview of the subject |  |
|  | Topic 1: The Data Scientist:   1. Bring change to an organisation. 2. Interacting with decision makers – working across an organisation leadership structure. |  |
|  | Topic 2: Data Strategy (1):   1. Requirements/ Stages 2. Governance: Policies & Standards 3. Data Architecture 4. Data Quality   Business perspective: threats and risks |  |
|  | Topic 3: Data Science as a process:   1. The scientific method 2. Understanding of how data is collected & analysed, & applied 3. data synthesis: understanding data from multiple perspectives. 4. Tools: people, processes, and technology. 5. Statistical approaches and the generation and use of hypothesis statements. |  |
|  | Topic 4: Describing and communicating data   1. Numerical and graphical descriptive statistics 2. Assessing against a target 3. Comparing between groups   ***Coursework 1 handout*** | R for Data Science (2014), Packt Publishing Ltd, Birmingham, D. Toomey  Good Practice Team, Effective tables and graphs in official statistics |
|  | R workshop   1. Introduction to R 2. Data entry 3. Descriptives and graphics | Introduction to R DataCamp https://www.datacamp.com/courses/free-introdu |
|  | Topic 4: Data Strategy (2):   1. Master Data Management (MDM) 2. Governance, Risk Management & Compliance 3. Governance & Information frameworks: Mike, Togaf, etc. 4. Security & privacy |  |
|  | Topic 5: Deployment & Implementation:   1. Driving policies & Standards: Approaches - Business buy-in 2. Communications and Reporting 3. Measurement – KPIs and Linking data back to business goals 4. Professionals and ethics |  |
|  | Topic 7. Multivariate data analysis   1. Visualisation methods 2. Correlation 3. Data Reduction- Principal components analysis 4. Cluster and segmentation – K means and modes – classification trees | * R for Data Science (2014), Packt Publishing D. Tomey |
|  | Workshop 2: Delivery of analytics capabilities  Coursework 2 handout |  |
|  | Case Studies/Business Application (1)  Changing the culture of the organisation - Data Culture |  |
|  | Case Studies/Business Application (2)  Using techniques from other disciplines: Case Study, DOE. |  |
|  | Case Studies/Business Application (3)  Industry Leaders - Case Study from industry leaders e.g. Davenports Ideas |  |
|  | Statistical modelling  Building predictive models  Diagnosing models  The General Linear Model | R for Data Science (2014), Packt Publishing D. Tomey |
|  | Workshop 3: The effective data & analytics driven organisation |  |