



**MONASH** University  
Information Technology

**FIT3152**  
**Data science**

**Unit Guide**

**Semester 2, 2013**

The information contained in this unit guide is correct at time of publication. The University has the right to change any of the elements contained in this document at any time.

*Last updated: 25 Jul 2013*

# Table of Contents

<b><u>FIT3152 Data science - Semester 2, 2013</u></b> .....	<b>1</b>
<u>Mode of Delivery</u> .....	1
<u>Contact Hours</u> .....	1
<u>Workload requirements</u> .....	1
<u>Unit Relationships</u> .....	1
<u>Prerequisites</u> .....	1
<u>Chief Examiner</u> .....	2
<u>Campus Lecturer</u> .....	2
<u>Clayton</u> .....	2
<u>Tutors</u> .....	2
<u>Clayton</u> .....	2
<b><u>Academic Overview</u></b> .....	<b>3</b>
<u>Learning Outcomes</u> .....	3
<b><u>Unit Schedule</u></b> .....	<b>4</b>
<u>Assessment Summary</u> .....	4
<u>Teaching Approach</u> .....	5
<b><u>Assessment Requirements</u></b> .....	<b>6</b>
<u>Assessment Policy</u> .....	6
<u>Assessment Tasks</u> .....	6
<u>Participation</u> .....	6
<u>Examinations</u> .....	7
<u>Examination 1</u> .....	7
<u>Learning resources</u> .....	7
<u>Feedback to you</u> .....	7
<u>Extensions and penalties</u> .....	8
<u>Returning assignments</u> .....	8
<u>Referencing requirements</u> .....	8
<u>Assignment submission</u> .....	8
<u>Online submission</u> .....	8
<u>Recommended text(s)</u> .....	8
<b><u>Other Information</u></b> .....	<b>9</b>
<u>Policies</u> .....	9
<u>Graduate Attributes Policy</u> .....	9
<u>Student services</u> .....	9
<u>Monash University Library</u> .....	9
<u>Disability Liaison Unit</u> .....	10
<u>Your feedback to Us</u> .....	10
<u>Previous Student Evaluations of this Unit</u> .....	10

# **FIT3152 Data science - Semester 2, 2013**

In recent years the world has seen an explosion in the quantity and variety of data routinely recorded and analysed by research and industry, prompting some social commentators to refer to this phenomenon as the rise of "big data," and the analysts and practitioners who investigate the data as "data scientists."

The data may come from a variety of sources, including scientific experiments and measurements, or may be recorded from human interactions such as browsing data or social networks on the Internet, mobile phone usage or financial transactions. Many companies too, are realising the value of their data for analysing customer behaviour and preferences, recognising patterns of behaviour such as credit card usage or insurance claims to detect fraud, as well as more accurately evaluating risk and increasing profit.

In order to obtain insights from big data new analytical techniques are required by practitioners. These include computationally intensive and interactive approaches such as visualisation, clustering and data mining. The management and processing of large data sets requires the development of enhanced computational resources and new algorithms to work across distributed computers.

This unit will introduce students to the analysis and management of big data using current techniques and open source and proprietary software tools. Data and case studies will be drawn from diverse sources including health and informatics, life sciences, web traffic and social networking, business data including transactions, customer traffic, scientific research and experimental data. The general principles of analysis, investigation and reporting will be covered. Students will be encouraged to critically reflect on the data analysis process within their own domain of interest.

## **Mode of Delivery**

Clayton (Day)

## **Contact Hours**

2 hrs lectures/wk, 2 hrs laboratories/wk

## **Workload requirements**

Lectures: 2 hours per week

Tutorials/Lab Sessions: 2 hours per week per tutorial

and up to an additional 8 hours in some weeks for completing lab and project work, private study and revision.

## **Unit Relationships**

### **Prerequisites**

FIT1006, ETC1000 or equivalent. (For example BUS1100, ETC1010, ETC2010, ETF2211, ETW1000, ETW1010, ETW1102, ETW2111, ETX1100, ETX2111, ETX2121, MAT1097, STA1010)

## **Chief Examiner**

**Dr John Betts**

## **Campus Lecturer**

**Clayton**

**Dr John Betts**

**Dr Sue Bedingfield**

## **Tutors**

**Clayton**

**Mr Rj Chow**

**Dr Kefeng (Jason) Xuan**

# Academic Overview

## Learning Outcomes

At the completion of this unit students will have -A knowledge and understanding of:

- analysing large data sets;
- data cleansing and preparation;
- open source and proprietary software for data analytics;
- techniques and tools for data analytics;
- validation of results.

Developed attitudes that enable them to:

- model business problems by transforming the problem into an analytics problem that can then be solved using data analytics techniques. The insights from the analysis are then related back to the original business problem;
- interpret data within a domain-specific context;
- understand how data analytics may be used within organisations to understand current practice and identify potential opportunities;
- appreciate the value of data analytics over traditional statistical analysis and modelling;
- critically evaluate the limitations and benefits of data analytics.

Gained practical skills to:

- manage large data;
- prepare data for analysis;
- analyse large data sets; in particular textual data sets;
- construct and test the reliability of predictive models;
- techniques and tools for data analytics.

Demonstrated the communication skills necessary to:

- frame a business problem in terms of a formulation suitable for the application of data analytics tools;
- communicate and report analysis and findings.

## Unit Schedule

Week	Activities	Assessment
0		No formal assessment or activities are undertaken in week 0
1	Introduction to Data Science. Introduction to R. Review of basic statistics using R	Tutorial Participation assessed Weekly
2	Exploring data using graphics in R	
3	Analytics and modelling in R	
4	Data cleansing, consulting, case studies. (Guest Lecture)	
5	Programming in R	Group Assignment (Initial report) due 30 August 2013
6	Classification using decision trees	
7	Comparing classification models, ensemble techniques	
8	K-Means and hierarchical clustering	
9	Text analysis	
10	Scalable algorithms. Map Reduce	Individual Assignment due 11 October 2013
11	Student Presentations	Students will give a brief presentation of their group project results. Group Assignment (Final report) due 18 October 2013
12	Review of the course and exam preparation	
	SWOT VAC	No formal assessment is undertaken in SWOT VAC
	Examination period	LINK to Assessment Policy: <a href="http://policy.monash.edu.au/policy-bank/academic/education/assessment/assessment-in-coursework-policy.html">http://policy.monash.edu.au/policy-bank/academic/education/assessment/assessment-in-coursework-policy.html</a>

\*Unit Schedule details will be maintained and communicated to you via your learning system.

## Assessment Summary

Examination (2 hours): 60%; In-semester assessment: 40%

Assessment Task	Value	Due Date
Group Assignment	20%	Initial report due 30 August 2013, Final report due 18 October 2013
Individual Assignment	10%	11 October 2013
Tutorial Participation	10%	Weekly
Examination 1	60%	To be advised

## **Teaching Approach**

### **Lecture and tutorials or problem classes**

This teaching and learning approach helps students to initially encounter information at lectures, discuss and explore the information during tutorials, and practice in a hands-on lab environment.

# Assessment Requirements

## Assessment Policy

Faculty Policy - Unit Assessment Hurdles

(<http://www.infotech.monash.edu.au/resources/staff/edgov/policies/assessment-examinations/unit-assessment-hu>)

Academic Integrity - Please see the Demystifying Citing and Referencing tutorial at

<http://lib.monash.edu/tutorials/citing/>

## Assessment Tasks

### Participation

#### • Assessment task 1

**Title:**

Group Assignment

**Description:**

Students will work in groups to analyse a large data set and report their findings.

**Weighting:**

20%

**Criteria for assessment:**

- ◆ Understanding of the real-world problem, and how the data might be used to solve the problem.
- ◆ Cleansing and pre-processing the data.
- ◆ Visual representation of the data, and initial insights into the data. (Initial report at this milestone)
- ◆ Accuracy and reliability of the model.
- ◆ Reporting and communication of results.

As this is a group project, students in each group will allocate a weighting of the final results to each member of the group based on a consensus estimate of each member's contribution.

**Due date:**

Initial report due 30 August 2013, Final report due 18 October 2013

#### • Assessment task 2

**Title:**

Individual Assignment

**Description:**

Students will individually analyse a data set and report their findings.

**Weighting:**

10%

**Criteria for assessment:**

- ◆ Understanding of the problem, and how the data might be used to solve the problem.
- ◆ Cleansing and pre-processing the data.
- ◆ Visual representation of the data, and initial insights into the data.



## Assessment Requirements

- ◆ Accuracy and reliability of the model.
- ◆ Reporting and communication of results.

**Due date:**

11 October 2013

• **Assessment task 3**

**Title:**

Tutorial Participation

**Description:**

Students will be assessed on their participation during tutorials.

**Weighting:**

10%

**Criteria for assessment:**

- ◆ Participation in tutorials
- ◆ Completion of class exercises
- ◆ Contribution to class discussions

**Due date:**

Weekly

## Examinations

• **Examination 1**

**Weighting:**

60%

**Length:**

2 hours

**Type (open/closed book):**

Closed book

**Electronic devices allowed in the exam:**

Electronic calculators permitted in the exam.

## Learning resources

Monash Library Unit Reading List

<http://readinglists.lib.monash.edu/index.html>

## Feedback to you

Types of feedback you can expect to receive in this unit are:

- Informal feedback on progress in labs/tutes
- Graded assignments with comments
- Solutions to tutes, labs and assignments

## Extensions and penalties

Submission must be made by the due date otherwise penalties will be enforced.

You must negotiate any extensions formally with your campus unit leader via the in-semester special consideration process: <http://www.monash.edu.au/exams/special-consideration.html>

## Returning assignments

Students can expect assignments to be returned within two weeks of the submission date or after receipt, whichever is later.

## Referencing requirements

As per Faculty policy (referencing for Master coursework and undergraduate - see <http://intranet.monash.edu.au/infotech/resources/staff/edgov/policies/units/style-masters-ug-degrees.html>), the Unit Guide will include links to the relevant referencing requirements for the unit.

## Assignment submission

It is a University requirement (<http://www.policy.monash.edu/policy-bank/academic/education/conduct/plagiarism-procedures.html>) for students to submit an assignment coversheet for each assessment item. Faculty Assignment coversheets can be found at <http://www.infotech.monash.edu.au/resources/student/forms/>. Please check with your Lecturer on the submission method for your assignment coversheet (e.g. attach a file to the online assignment submission, hand-in a hard copy, or use an online quiz). Please note that it is your responsibility to retain copies of your assessments.

## Online submission

If Electronic Submission has been approved for your unit, please submit your work via the learning system for this unit, which you can access via links in the my.monash portal.

## Recommended text(s)

W. N. Venables, D. M. Smith. (2013). *An Introduction to R*. () Available from: <http://www.cran.r-project.org/doc/manuals/R-intro.pdf>.

M. Allerhand. (2011). *A tiny handbook of R*. () SpringerLink (Online service), Online access via Library.

Pang-Ning Tan, Michael Steinbach, Vipin Kumar. (2006). *Introduction to data mining*. () Addison-Wesley.

Luis Torgo. (2011). *Data mining with R: learning with case studies*. () Chapman & Hall CRC.

## Other Information

### Policies

Monash has educational policies, procedures and guidelines, which are designed to ensure that staff and students are aware of the University's academic standards, and to provide advice on how they might uphold them. You can find Monash's Education Policies at:

[www.policy.monash.edu.au/policy-bank/academic/education/index.html](http://www.policy.monash.edu.au/policy-bank/academic/education/index.html)

Key educational policies include:

- Academic integrity;  
<http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-policy.html>
- Assessment in Coursework Programs;  
<http://www.policy.monash.edu/policy-bank/academic/education/assessment/assessment-in-coursework-policy.html>
- Special Consideration;  
<http://www.policy.monash.edu/policy-bank/academic/education/assessment/special-consideration-policy.html>
- Grading Scale;  
<http://www.policy.monash.edu/policy-bank/academic/education/assessment/grading-scale-policy.html>
- Discipline: Student Policy;  
<http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-discipline-policy.html>
- Academic Calendar and Semesters; <http://www.monash.edu.au/students/dates/>
- Orientation and Transition; <http://intranet.monash.edu.au/infotech/resources/students/orientation/>
- Academic and Administrative Complaints and Grievances Policy;  
<http://www.policy.monash.edu/policy-bank/academic/education/management/complaints-grievance-policy.html>
- Code of Practice for Teaching and Learning;  
<http://www.policy.monash.edu.au/policy-bank/academic/education/conduct/suppdocs/code-of-practice-teaching-and-learning.html>

### Graduate Attributes Policy

<http://www.policy.monash.edu/policy-bank/academic/education/management/monash-graduate-attributes-policy.html>

### Student services

The University provides many different kinds of support services for you. Contact your tutor if you need advice and see the range of services available at <http://www.monash.edu.au/students>. For Sunway see <http://www.monash.edu.my/Student-services>, and for South Africa see <http://www.monash.ac.za/current/>.

### Monash University Library

The Monash University Library provides a range of services, resources and programs that enable you to save time and be more effective in your learning and research. Go to [www.lib.monash.edu.au](http://www.lib.monash.edu.au) or the library tab in [my.monash](#) portal for more information. At Sunway, visit the Library and Learning Commons at <http://www.lib.monash.edu.my/>. At South Africa visit <http://www.lib.monash.ac.za/>.

## Disability Liaison Unit

Students who have a disability or medical condition are welcome to contact the Disability Liaison Unit to discuss academic support services. Disability Liaison Officers (DLOs) visit all Victorian campuses on a regular basis.

Website: <http://www.monash.edu/equity-diversity/disability/index.html> Telephone: 03 9905 5704 to book an appointment with a DLO; or contact the Student Advisor, Student Community Services at 03 55146018 at Sunway Email: [dlu@monash.edu](mailto:dlu@monash.edu) Drop In: Equity and Diversity Centre, Level 1, Building 55, Clayton Campus, or Student Community Services Department, Level 2, Building 2, Monash University, Sunway Campus

## Your feedback to Us

Monash is committed to excellence in education and regularly seeks feedback from students, employers and staff. One of the key formal ways students have to provide feedback is through the Student Evaluation of Teaching and Units (SETU) survey. The University's student evaluation policy requires that every unit is evaluated each year. Students are strongly encouraged to complete the surveys. The feedback is anonymous and provides the Faculty with evidence of aspects that students are satisfied and areas for improvement.

For more information on Monash's educational strategy, see:

[www.monash.edu.au/about/monash-directions](http://www.monash.edu.au/about/monash-directions) and on student evaluations, see:  
[www.policy.monash.edu/policy-bank/academic/education/quality/student-evaluation-policy.html](http://www.policy.monash.edu/policy-bank/academic/education/quality/student-evaluation-policy.html)

## Previous Student Evaluations of this Unit

This is a new unit.

If you wish to view how previous students rated this unit, please go to <https://emuapps.monash.edu.au/unitevaluations/index.jsp>