# Table of Contents

**FIT5145 Introduction to data science - Semester 2, 2015**

- Mode of Delivery ................................................................. 1
- Workload Requirements .......................................................... 1
- Unit Relationships .................................................................. 1
  - Prerequisites .................................................................. 1
- Chief Examiner .................................................................... 1
- Campus Lecturer ................................................................... 2
  - Caulfield ........................................................................ 2
- Your feedback to Us .............................................................. 2

**Academic Overview** ................................................................ 3

- Learning Outcomes .............................................................. 3

**Unit Schedule** .................................................................... 4

- Teaching Approach .............................................................. 4
- Assessment Summary ........................................................... 4

**Assessment Requirements** .................................................. 6

- Assessment Policy ............................................................... 6
- Assessment Tasks ................................................................ 6
  - Participation .................................................................. 6

- Learning resources .............................................................. 8
- Feedback to you .................................................................. 8
- Extensions and penalties ..................................................... 8
- Returning assignments ......................................................... 8
- Assignment submission ....................................................... 8
- Online submission .............................................................. 9
- Technological Requirements ................................................ 9

**Other Information** ............................................................... 10

- Policies ............................................................................. 10
  - Faculty resources and policies ........................................... 10
    - Graduate Attributes Policy ............................................ 10
- Student Charter ................................................................. 10
- Student services ............................................................... 10
- Monash University Library ................................................. 10
- Disability Liaison Unit ....................................................... 10
FIT5145 Introduction to data science - Semester 2, 2015

This unit looks at processes and case studies to understand the many facets of working with data, and the significant effort in Data Science over and above the core task of Data Analysis. Working with data as part of a business model and the lifecycle in an organisation is considered, as well as business processes and case studies. Data and its handling is also introduced: characteristic kinds of data and its collection, data storage and basic kinds of data preparation, data cleaning and data stream processing. Curation and management are reviewed: archival and architectural practice, policy, legal and ethical issues. Styles of data analysis and outcomes of successful data exploration and analysis are reviewed. Standards, tools and resources are also reviewed.

Mode of Delivery

Caulfield (Day)

Workload Requirements

Minimum total expected workload equals 144 hours per semester comprising:

1. Contact hours for on-campus students:
   - Two hours/week lectures
   - Two hours/week laboratories

1. Contact hours for Monash Online students:
   - Two hours/week online group sessions

Online students generally do not attend lecture, tutorial and laboratory sessions, however should plan to spend equivalent time working through resources and participating in discussions.

1. Additional requirements (all students):
   - A minimum of 8 hours per week of personal study (22 hours per week for Monash online students) for completing lab/tutorial activities, assignments, private study and revision, and for online students, participating in discussions.

See also Unit timetable information

Unit Relationships

Prerequisites

(FIT5131 or FIT9131) and (FIT5132 or FIT9132) or equivalent

Chief Examiner

Dr Wray Buntine
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/ and on student evaluations, see:
www.policy.monash.edu/policy-bank/academic/education/quality/student-evaluation-policy.html
Academic Overview

Learning Outcomes

On successful completion of this unit a student should be able to:

1. analyse the role of data in different styles of business;
2. demonstrate the size and scope of data storage and data processing, and classify the basic technologies in use;
3. assess tasks for data curation and management in an organisation;
4. classify participants in a data science project: such as statistician, archivist, analyst, and systems architect;
5. classify the kinds of data analysis and statistical methods available for a data science project;
6. locate and assess resources, software and tools for a data science project.
## Unit Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Activities</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No formal assessment or activities are undertaken in week 0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Overview of data science and what a project looks like.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Roles of a data scientist, and the impact of the field.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Review of Python. Data business models.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Application areas and case studies.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Characterising data and &quot;big&quot; data.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Data sources and case studies.</td>
<td>Test 1.</td>
</tr>
<tr>
<td>7</td>
<td>More Python, regular expressions. Resources and standards.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Resources case studies.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Data analysis theory.</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Data analysis process.</td>
<td>Test 2.</td>
</tr>
<tr>
<td>11</td>
<td>Issues in data management.</td>
<td>Assignment 3 part 1 due.</td>
</tr>
<tr>
<td>12</td>
<td>Data management frameworks.</td>
<td>Assignment 1, Assignment 2, Assignment 3 part 2 due.</td>
</tr>
<tr>
<td>SWOT VAC</td>
<td></td>
<td>No formal assessment is undertaken during SWOT VAC</td>
</tr>
</tbody>
</table>

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

## Teaching Approach

### Other

The teaching approach is flipped classroom. Students will be introduced to topics via online ePub and video material. Lectures will be expert-led sessions where concepts are demonstrated interactively with the class. Students will gain practical experience via lab classes with activities and tutor-led discussion.

## Assessment Summary

In-semester assessment: 100%

<table>
<thead>
<tr>
<th>Assessment Task</th>
<th>Value</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment 1: Data Science and me.</td>
<td>10%</td>
<td>Friday Week 12</td>
</tr>
<tr>
<td>Assignment 2: Data Science Resources</td>
<td>20%</td>
<td>Friday Week 12</td>
</tr>
<tr>
<td>Assignment 3: Business and data case study</td>
<td>40%+10%</td>
<td>Friday Week 11 + Friday</td>
</tr>
</tbody>
</table>
Unit Schedule

Test 1: size and scope of data storage and data processing, and their technologies 10% Friday Week 6
Test 2: kinds of data analysis and statistical methods 10% Friday Week 10
Assessment Requirements

Assessment Policy

Faculty Policy - Unit Assessment Hurdles

Academic Integrity - Please see resources and tutorials at
http://www.monash.edu/library/skills/resources/tutorials/academic-integrity/

Assessment Tasks

Participation

• Assessment task 1

  Title: Assignment 1: Data Science and me.
  Description: Learning outcomes 4 and 6. For this assessment task, students are required to create and maintain a reflective journal that specifies their goals and aspirations for a career in data science. In the journal, students must also specify training and resources required for them to achieve their goal. Approximately 200 words are to be entered 6 times (totalling approx 1200 words) during the teaching period.

  Weighting: 10%
  Criteria for assessment:
  1. Classify roles in data science project such as statistician, archivist, analyst and systems architect
  2. Identify and assess resources, software and tools.

  Due date: Friday Week 12

• Assessment task 2

  Title: Assignment 2: Data Science Resources
  Description: Learning outcomes 5 and 6. For this assessment task, students are required to create and maintain their own detailed annotated list of resources (software, tools, data) required for a proposed data science project as well as a list of the particular standards they may use. Annotation should explain justifications. They are not doing the data science project, they just have to describe it sufficiently to then present resources.

  Weighting: 20%
  Criteria for assessment:
The report will be assessed on demonstration of:

1. Classifying the kinds of data analysis and statistical methods available for a data science project.
2. Identifying and assessing resources, software and tools for a data science project.

Due date:
Friday Week 12

• Assessment task 3

Title:
Assignment 3: Business and data case study

Description:
Learning outcomes 1, 2, 3 and 4. For this assessment task, students are required to prepare and present a detailed report based on a business and data case study. The report must explain:

♦ how the case study fits into the classification and lifecycle models
♦ what sorts of data is required, and its ‘V’ characteristics
♦ the enabling factors behind the success of the project
♦ its business value to the organisation
♦ data curation and management

The report will be in the form of both a 1000 word report and a 3 minute video presentation.

Second, the student is required to do a peer review of another student's report.

Weighting:
40%+10%

Criteria for assessment:
The report and video will be assessed on the demonstration and knowledge of unit outcomes. The peer review will be assessed on the analysis of same.

Due date:
Friday Week 11 + Friday Week 12

• Assessment task 4

Title:
Test 1: size and scope of data storage and data processing, and their technologies

Description:
Learning outcome 2. A multiple choice test of 25 questions will be done.

Weighting:
10%

Criteria for assessment:
Correctness.

Due date:
Friday Week 6
**Assessment Requirements**

- **Assessment task 5**

  **Title:**
  Test 2: kinds of data analysis and statistical methods

  **Description:**
  Learning outcome 5. A multiple choice test of 25 questions will be done.

  **Weighting:**
  10%

  **Criteria for assessment:**
  Correctness.

  **Due date:**
  Friday Week 10

**Learning resources**

Monash Library Unit Reading List (if applicable to the unit)
http://readinglists.lib.monash.edu/index.html

**Feedback to you**

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

- Graded assignments with comments
- Test results and feedback
- Quiz results

**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.

**Assignment submission**

It is a University requirement (http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-managing-plagiarism-collusion-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 electronic submission). 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.

Technological Requirements

Students must regularly check Moodle for announcements. Video, audio, PDF and ePUBs for unit material will be made available through Moodle so access to a laptop or similar to view the material is necessary. Some workk will be done in the language Python so you must have access to a machine running Python with appropriate libraries.
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:

Faculty resources and policies

Important student resources including Faculty policies are located at
http://intranet.monash.edu.au/infotech/resources/students/

Graduate Attributes Policy

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

Student Charter


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 Malaysia 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 or the library tab in my.monash portal for more information. At Malaysia, 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 Malaysia
• Email: 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, Malaysia Campus