



# COMP6200

## Data Science

Session 2, In person-scheduled-weekday, North Ryde 2024

*School of Computing*

## Contents

|                                       |    |
|---------------------------------------|----|
| <u>General Information</u>            | 2  |
| <u>Learning Outcomes</u>              | 2  |
| <u>General Assessment Information</u> | 3  |
| <u>Assessment Tasks</u>               | 4  |
| <u>Delivery and Resources</u>         | 7  |
| <u>Unit Schedule</u>                  | 7  |
| <u>Policies and Procedures</u>        | 8  |
| <u>Changes from Previous Offering</u> | 10 |

### Disclaimer

Macquarie University has taken all reasonable measures to ensure the information in this publication is accurate and up-to-date. However, the information may change or become out-dated as a result of change in University policies, procedures or rules. The University reserves the right to make changes to any information in this publication without notice. Users of this publication are advised to check the website version of this publication [or the relevant faculty or department] before acting on any information in this publication.

## General Information

Unit convenor and teaching staff

Unit Convener

Guanfeng Liu

[guanfeng.liu@mq.edu.au](mailto:guanfeng.liu@mq.edu.au)

Lecturer

Xuhui Fan

[xuhui.fan@mq.edu.au](mailto:xuhui.fan@mq.edu.au)

Lecturer

Xuyun Zhang

[xuyun.zhang@mq.edu.au](mailto:xuyun.zhang@mq.edu.au)

Credit points

10

Prerequisites

Corequisites

Co-badged status

Unit description

This unit introduces students to the fundamental techniques and tools of data science, such as the graphical display of data, predictive models, evaluation methodologies, regression, classification and clustering. The unit provides practical experience applying these methods using industry-standard software tools to real-world data sets. Students who have completed this unit will be able to identify which data science methods are most appropriate for a real-world data set, apply these methods to the data set, and interpret the results of the analysis they have performed.

## Important Academic Dates

Information about important academic dates including deadlines for withdrawing from units are available at <https://www.mq.edu.au/study/calendar-of-dates>

## Learning Outcomes

On successful completion of this unit, you will be able to:

**ULO1:** Identify the appropriate Data Science analysis for a problem and apply that method to the problem.

**ULO2:** Interpret Data Science analyses and summarise and identify the most important aspects of a Data Science analysis.

**ULO3:** Present the results of their Data Science analyses both verbally and in written form.

**ULO4:** Discuss the broader implications of Data Science analyses.

## General Assessment Information

### Requirements to Pass this Unit

To pass this unit you must:

- Achieve a total mark equal to or greater than 50%, and
- Participate in, and undertake all hurdle activities for, a minimum of 8 of the 12 weekly workshops

#### Hurdle Assessments:

##### Assessment 1-Practical-based task (10%)

- Development of knowledge and skills requires continual practice at authentic problems in a laboratory-based setting. This unit has weekly laboratory classes and you must demonstrate your progress in developing and communicating knowledge and skills in a minimum of 8 of the 12 workshop practicals. This is a hurdle assessment meaning that failure to meet this requirement may result in a fail grade for the unit. It is allowed to take a second chance at each workshop hurdle assessment until they complete these workshops. **Additional absences will require approval of Special Consideration** (see below)

### Late submission

Unless a Special Consideration request has been submitted and approved, a 5% penalty (of the total possible mark of the task) will be applied for each day a written report or presentation assessment is not submitted, up until the 7th day (including weekends). After the 7th day, a grade of '0' will be awarded even if the assessment is submitted. The submission time for all uploaded assessments is **11:55 pm**. A 1-hour grace period will be provided to students who experience a technical concern. The late submission rule was changed to align with the new Faculty policy.

For any late submission of time-sensitive tasks, such as scheduled tests/exams, performance assessments/presentations, and/or scheduled practical assessments/labs, please apply for [Special Consideration](#). For example, if the assignment is worth 8 marks (of the entire unit) and your submission is late by 19 hours (or 23 hours 59 minutes 59 seconds), 0.4 marks (5% of 8 marks) will be deducted. If your submission is late by 24 hours (or 47 hours 59 minutes 59 seconds), 0.8 marks (10% of 8 marks) will be deducted, and so on.

## Assessments where Late Submissions will be accepted

- Assessment Portfolios – YES, Standard Late Penalty applies
- Assessment Workshops – YES, Standard Late Penalty applies
- Assessment Reflective Report – YES, Standard Late Penalty applies
- Assessment Critical Assessment Task – YES, Standard Late Penalty applies
- Assessment Mid Exam – NO, unless Special Consideration is Granted
- Assessment Final Exam – NO, unless Special Consideration is Granted

## Special Consideration

The Special Consideration Policy aims to support students who have been impacted by short-term circumstances or events that are serious, unavoidable, and significantly disruptive, and which may affect their performance in assessment.

*Written Assessments:* If you experience circumstances or events that affect your ability to complete the written assessments in this unit on time, please inform the convenor and submit a Special Consideration request through [ask.mq.edu.au](mailto:ask.mq.edu.au).

*Weekly practice-based tasks:* To pass the unit you need to demonstrate ongoing development of skills and application of knowledge in 8 out of 12 of the weekly practical classes. If you miss a weekly practical class due to a serious, unavoidable, and significant disruption, contact your convenor ASAP as you may be able to attend another class that week. If it is not possible to attend another class, you should still contact your convenor for access to class material to review in your own time.

## Assessment Tasks

| Name                          | Weighting | Hurdle | Due   |
|-------------------------------|-----------|--------|---|
| <u>Weekly Tasks</u>           | 10%       | Yes    | 11:55pm on Sunday ending Each Week              |
| <u>Examinations</u>           | 40%       | No     | TBA on Week 8 and Exam Week                     |
| <u>Data Science Portfolio</u> | 35%       | No     | 11:55pm on Saturday ending Weeks 5, 7, 9 and 11 |
| <u>Critical Analysis Task</u> | 15%       | No     | 11:55pm on Saturday ending Week 13              |

## Weekly Tasks

Assessment Type <sup>1</sup>: Practice-based task

Indicative Time on Task <sup>2</sup>: 0 hours

Due: **11:55pm on Sunday ending Each Week**

Weighting: **10%**

**This is a hurdle assessment task (see [assessment policy](#) for more information on hurdle assessment tasks)**

Development of knowledge and skills requires continual practice at authentic problems. During weekly workshops you will practice a range of tasks recording your progress on worksheets. To pass this hurdle assessment, you must be able to demonstrate your progress in developing and communicating knowledge and skills in a minimum of 8 of the 12 weekly workshops.

On successful completion you will be able to:

- Present the results of their Data Science analyses both verbally and in written form.

## Examinations

Assessment Type <sup>1</sup>: Examination

Indicative Time on Task <sup>2</sup>: 10 hours

Due: **TBA on Week 8 and Exam Week**

Weighting: **40%**

Examinations will assess your knowledge and understanding of the data analysis and machine learning methods covered in the semester.

On successful completion you will be able to:

- Interpret Data Science analyses and summarise and identify the most important aspects of a Data Science analysis.
- Discuss the broader implications of Data Science analyses.

## Data Science Portfolio

Assessment Type <sup>1</sup>: Portfolio

Indicative Time on Task <sup>2</sup>: 45 hours

Due: **11:55pm on Saturday ending Weeks 5, 7, 9 and 11**

Weighting: **35%**

The portfolio assessment will consist of a number of data analysis problems that you will be given through the semester. These will involve writing code to analyse one or more data sets. These will be marked individually through the semester and then as an overall portfolio at the end of semester.

On successful completion you will be able to:

- Identify the appropriate Data Science analysis for a problem and apply that method to the problem.
- Interpret Data Science analyses and summarise and identify the most important aspects of a Data Science analysis.
- Present the results of their Data Science analyses both verbally and in written form.

## Critical Analysis Task

Assessment Type <sup>1</sup>: Report

Indicative Time on Task <sup>2</sup>: 15 hours

Due: **11:55pm on Saturday ending Week 13**

Weighting: **15%**

You will be given a sample notebook describing the analysis of a dataset. You will provide a critical analysis of this notebook and suggest improvements in the way that data is analysed and results are presented.

On successful completion you will be able to:

- Identify the appropriate Data Science analysis for a problem and apply that method to the problem.
- Interpret Data Science analyses and summarise and identify the most important aspects of a Data Science analysis.
- Present the results of their Data Science analyses both verbally and in written form.
- Discuss the broader implications of Data Science analyses.

---

<sup>1</sup> If you need help with your assignment, please contact:

- the academic teaching staff in your unit for guidance in understanding or completing this type of assessment
- the [Writing Centre](#) for academic skills support.

<sup>2</sup> Indicative time-on-task is an estimate of the time required for completion of the assessment task and is subject to individual variation

## Delivery and Resources

### Classes

There will be one two-hour on-campus lecture each week, and one two-hour workshop in the computing laboratory. The on-campus lecture would be in the form of live streaming or prerecorded lecture videos as well. You are expected to attend both classes as they provide complimentary learning activities each week. In practical classes, you will write code and experiment with various data sets; in lectures, we will discuss the methods you are learning and how the results of your analysis can be interpreted.

*Week 1 Classes* In week 1, there is a two-hour lecture. Workshops will commence in week 1 as well. Every student should attend the workshop session in which they have enrolled. For details of days, times, and rooms consult the timetable system.

### Methods of Communication

We will communicate with you via your university email and through announcements on iLearn. Queries to convenors can either be placed on the iLearn discussion board or sent to the unit convenor via the contact email on iLearn.

### COVID Information

For the latest information on the University's response to COVID-19, please refer to the Coronavirus infection page on the Macquarie website: <https://www.mq.edu.au/about/coronavirus-faqs>. Remember to check this page regularly in case the information and requirements change during the semester. If there are any changes to this unit in relation to COVID, these will be communicated via iLearn.

### Textbooks

We will refer to the following texts during the semester: Introduction to Data Science A Python Approach to Concepts, Techniques and Applications Igual, Laura, Seguí, Santi (electronic edition available via MQ Library) Computational and Inferential Thinking: The Foundations of Data Science By Anil Adhikari and John DeNero (available on GitBooks). You will be given readings from these and other sources each week.

### Technology Used and Required

We will make use of Python 3 for data analysis, including a range of modules such as scikitlearn, pandas, and numpy that provide additional features. These can all be installed via the Anaconda Python distribution. We will discuss this environment and the installation process in the first week of classes. We will use Jupyter Notebook as a way of developing and presenting the analysis results. This is included in the full Anaconda distribution.

### Unit Schedule

- Week 1: Overview of DS, Learning Python, Notebooks

- Week 2: Data formats, Python input and output
- Week 3: Descriptive Statistics, Simple visualisation
- Week 4: Causality and correlation; Visualisation, Data Ethics
- Week 5: Predictive Modelling: Linear and Logistic Regression
- Week 6: Software Engineering for Data Science and Data Privacy
- Week 7: Feature sets and spaces; Unsupervised learning
- Week 8: Supervised Learning: K-Nearest Neighbours
- Week 9: Naive Bayes Classifiers
- Week 10: Artificial Neural Networks
- Week 11: Learning Decision Trees
- Week 12: Summary
- Week 13: Data Science Guest Lecture

## Policies and Procedures

Macquarie University policies and procedures are accessible from [Policy Central](https://policies.mq.edu.au) (<https://policies.mq.edu.au>). Students should be aware of the following policies in particular with regard to Learning and Teaching:

- [Academic Appeals Policy](#)
- [Academic Integrity Policy](#)
- [Academic Progression Policy](#)
- [Assessment Policy](#)
- [Fitness to Practice Procedure](#)
- [Assessment Procedure](#)
- [Complaints Resolution Procedure for Students and Members of the Public](#)
- [Special Consideration Policy](#)

Students seeking more policy resources can visit [Student Policies](https://students.mq.edu.au/support/study/policies) (<https://students.mq.edu.au/support/study/policies>). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

To find other policies relating to Teaching and Learning, visit [Policy Central](https://policies.mq.edu.au) (<https://policies.mq.edu.au>) and use the [search tool](#).

## Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: <https://students.mq.edu.au/admin/other-resources/student-conduct>

## Results

Results published on platform other than [eStudent](#), (eg. iLearn, Coursera etc.) or released directly by your Unit Convenor, are not confirmed as they are subject to final approval by the

University. Once approved, final results will be sent to your student email address and will be made available in [eStudent](#). For more information visit [connect.mq.edu.au](http://connect.mq.edu.au) or if you are a Global MBA student contact [globalmba.support@mq.edu.au](mailto:globalmba.support@mq.edu.au)

## Academic Integrity

At Macquarie, we believe [academic integrity](#) – honesty, respect, trust, responsibility, fairness and courage – is at the core of learning, teaching and research. We recognise that meeting the expectations required to complete your assessments can be challenging. So, we offer you a range of resources and services to help you reach your potential, including free [online writing and maths support](#), [academic skills development](#) and [wellbeing consultations](#).

## Student Support

Macquarie University provides a range of support services for students. For details, visit <http://students.mq.edu.au/support/>

### The Writing Centre

[The Writing Centre](#) provides resources to develop your English language proficiency, academic writing, and communication skills.

- [Workshops](#)
- [Chat with a WriteWISE peer writing leader](#)
- [Access StudyWISE](#)
- [Upload an assignment to Studiosity](#)
- [Complete the Academic Integrity Module](#)

The Library provides online and face to face support to help you find and use relevant information resources.

- [Subject and Research Guides](#)
- [Ask a Librarian](#)

## Student Services and Support

Macquarie University offers a range of [Student Support Services](#) including:

- [IT Support](#)
- [Accessibility and disability support](#) with study
- Mental health [support](#)
- [Safety support](#) to respond to bullying, harassment, sexual harassment and sexual assault
- [Social support](#) including information about finances, tenancy and legal issues
- [Student Advocacy](#) provides independent advice on MQ policies, procedures, and processes

## Student Enquiries

Got a question? Ask us via the [Service Connect Portal](#), or contact [Service Connect](#).

## IT Help

For help with University computer systems and technology, visit [http://www.mq.edu.au/about\\_us/offices\\_and\\_units/information\\_technology/help/](http://www.mq.edu.au/about_us/offices_and_units/information_technology/help/).

When using the University's IT, you must adhere to the [Acceptable Use of IT Resources Policy](#). The policy applies to all who connect to the MQ network including students.

## Changes from Previous Offering

Student feedback from the previous offering of this unit was very positive overall, with students pleased with the clarity around assessment requirements and the level of support from the teaching staff. As such, no change to the delivery of the unit is planned, however, we will continue to strive to improve the level of support and the level of student engagement.

---

Unit information based on version 2024.04 of the [Handbook](#)