



COMP2200

Data Science

Session 2, Special circumstances 2021

School of Computing

Contents

<u>General Information</u>	2
<u>Learning Outcomes</u>	3
<u>Assessment Tasks</u>	3
<u>Delivery and Resources</u>	6
<u>Unit Schedule</u>	7
<u>Policies and Procedures</u>	7

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.

Session 2 Learning and Teaching Update

The decision has been made to conduct study online for the remainder of Session 2 for all units WITHOUT mandatory on-campus learning activities. Exams for Session 2 will also be online where possible to do so.

This is due to the extension of the lockdown orders and to provide certainty around arrangements for the remainder of Session 2. We hope to return to campus beyond Session 2 as soon as it is safe and appropriate to do so.

Some classes/teaching activities cannot be moved online and must be taught on campus. You should already know if you are in one of these classes/teaching activities and your unit convenor will provide you with more information via iLearn. If you want to confirm, see the list of [units with mandatory on-campus classes/teaching activities](#).

Visit the [MQ COVID-19 information page](#) for more detail.

General Information

Unit convenor and teaching staff

Lecturer and Convener

Yipeng Zhou

yipeng.zhou@mq.edu.au

Tutor

Asim Adnan Eijaz

asimadnan.eijaz@mq.edu.au

Contact via Email

Tutor

Bayzid Hossain

bayzid-ashik.hossain@mq.edu.au

Tutor

Subhash Sagar

subhash.sagar@mq.edu.au

Tutor

David Warren

david.warren@students.mq.edu.au

Tutor

Yao Deng

yao.deng@mq.edu.au

Tutor

Jiwei Guan

jiwei.guan@mq.edu.au

Tutor

Jianchao Lu

jianchao.lu@hdr.mq.edu.au

Steve Cassidy

steve.cassidy@mq.edu.au

Credit points

10

Prerequisites

(COMP1000 or COMP115 or COMP1010 or COMP125) and (STAT1170 or STAT170 or STAT1371 or STAT171 or STAT1250 or STAT150)

Corequisites**Co-badged status**

comp6200

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.

Assessment Tasks

Name	Weighting	Hurdle	Due
Final Exam	40%	No	Final Exam Period
Weekly Submissions	10%	Yes	Weekly
Data Science Portfolio	20%	No	Weeks 4, 6 & 8 for feedback. Week 10 final.

Name	Weighting	Hurdle	Due
Data Science Project	30%	No	Week 7, Week 13

Final Exam

Assessment Type ¹: Examination

Indicative Time on Task ²: 10 hours

Due: **Final Exam Period**

Weighting: **40%**

The exam 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.

Weekly Submissions

Assessment Type ¹: Participatory task

Indicative Time on Task ²: 0 hours

Due: **Weekly**

Weighting: **10%**

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

A submission of a small task based on the workshop each week. This may be a short quiz or the result of a practical task.

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

Data Science Portfolio

Assessment Type ¹: Portfolio

Indicative Time on Task ²: 30 hours

Due: **Weeks 4, 6 & 8 for feedback. Week 10 final.**

Weighting: **20%**

The portfolio assessment will consist of **three** small data analysis problems that you will be given through the semester. These will involve writing code to analyse one or more data sets. You will show the versions in the workshops for feedback and then submit a final version towards 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.
- Discuss the broader implications of Data Science analyses.

Data Science Project

Assessment Type ¹: Report

Indicative Time on Task ²: 40 hours

Due: **Week 7, Week 13**

Weighting: **30%**

In groups of 3-4, students will be given or will find one or more datasets and are asked to develop an analysis of this data and present a report. This project should include using more than one dataset, cleaning and analysing the data, training at least two different predictive models and using the model to make some conclusions. The report should be reproducible, all methods not only documented but available as an executable archive along with the data.

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.

¹ 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.

² 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 online lecture each week, and one two hour workshop in the computing laboratory or online. The online lecture would be in the form of live streaming or pre-recorded lecture videos. 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.

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 Ani 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 *scikit-learn*, *pandas*, *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.

Project Work

A major part of the assessment in this unit is based on a project that you will complete in groups. This will allow you to explore the techniques you are learning in class in a real-world data analysis exercise.

Unit Schedule

Unit Schedule

The indicative list of topics is shown here, this is subject to change based on feedback from the class.

1	Overview of DS, Learning Python, Notebooks	SS
2	Data formats, Python input and output	SS
3	Descriptive Statistics, simple visualisation	SS
4	Causality and correlation; Visualisation	SS
5	Predictive Modelling: Linear and Logistic Regression	SS
6	Software Engineering for Data Science	SS
7	Feature Engineering; Unsupervised Learning	SS/XZ
8	K-Nearest Neighbours Classifiers	XZ
9	Naive Bayes Classifiers	XZ
10	Artificial Neural Networks	XZ
11	Decision Tree Models	XZ
12	Advanced Topics / Guest Lecture	Guest
13	Summary	All

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](#)
- [Grade Appeal Policy](#)
- [Complaint Management 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 ask.mq.edu.au or if you are a Global MBA student contact globalmba.support@mq.edu.au

Student Support

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

Learning Skills

Learning Skills (mq.edu.au/learningskills) provides academic writing resources and study strategies to help you improve your marks and take control of your study.

- [Getting help with your assignment](#)
- [Workshops](#)
- [StudyWise](#)
- [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

Students with a disability are encouraged to contact the [Disability Service](#) who can provide appropriate help with any issues that arise during their studies.

Student Enquiries

For all student enquiries, visit Student Connect at ask.mq.edu.au

If you are a Global MBA student contact globalmba.support@mq.edu.au

IT Help

For help with University computer systems and technology, visit 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.