



STAJ1379

Statistical Technologies for Data Science

NMJI Term 1 (Sept), In person-scheduled-weekday, China 2026

School of Mathematical and Physical Sciences

Contents

<u>General Information</u>	2
<u>Learning Outcomes</u>	3
<u>General Assessment Information</u>	3
<u>Assessment Tasks</u>	4
<u>Delivery and Resources</u>	6
<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

Jun Ma

jun.ma@mq.edu.au

Greg Teuss

gregory.teuss@mq.edu.au

Credit points

10

Prerequisites

STAJ1371

Corequisites

Co-badged status

Unit description

Professional statistical work often involves a blend of statistical modelling, coding, and effective communication of results. This unit introduces R, an industry-standard coding language for statistical computing and graphics, to address a wide spectrum of data science challenges, encompassing tasks like data importation, data wrangling, and visualisation. Through hands-on projects and real-world case studies, students will gain expertise in the end-to-end data science workflow. Students will also acquire proficiency in utilising relational databases and executing SQL queries to leverage the full potential of data for informed decision-making. The unit places a strong emphasis on practicality, with most classes and assessments conducted in a computer lab. Students who have completed this unit will emerge with a versatile skill set that bridges the gap between statistical analysis and data technology, positioning them for success in a data-driven world.

Learning in this unit enhances student understanding of global challenges identified by the United Nations Sustainable Development Goals ([UNSDGs](#)) Quality Education; Decent Work and Economic Growth; Industry, Innovation and Infrastructure

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: Translate statistical thinking into coding syntax and appropriate data structures

ULO2: Compose SQL queries to manage and retrieve data from relational databases

ULO3: Import, clean, and transform real-world datasets using R

ULO4: Develop and implement computational strategies to tackle data challenges

ULO5: Generate meaningful visualisations using R and communicate results to diverse audiences

ULO6: Apply good practice of statistical project workflow

General Assessment Information

Attendance and participation

We strongly encourage all students to participate actively in all learning activities. Regular engagement is crucial for your success in this unit, as these activities provide opportunities to deepen your understanding of the material, collaborate with peers, and receive valuable feedback from instructors, to assist in completing the unit assessments. Your active participation not only enhances your own learning experience but also contributes to a vibrant and dynamic learning environment for everyone.

Requirements to Pass this Unit

To pass this unit you must:

- Achieve a total mark equal to or greater than 50%.

Hurdle Assessments

There is no Hurdle Assessment in this unit.

Late Submission Policy

- **5% penalty per day:** If you submit your assessment late, **5% of the total possible marks** will be deducted for each day (including weekends), up to 7 days.
 - *Example 1 (out of 100):* If you score 85/100 but submit 20 hours late, you will lose 5 marks and receive **80/100**.
 - *Example 2 (out of 30):* If you score 27/30 but submit 1 day late, you will lose 1.5 marks and receive **25.5/30**.
- **After 7 days:** Submissions more than 7 days late will receive a **mark of 0**.
- **Extensions:**

- **Automatic short extension:** Some assessments are eligible for automatic short extension. You can only apply for an automatic short extension before the due date.
- **Special Consideration:** If you need more time due to serious issues and for any assessments that are not eligible for Short Extension, you must apply for Special Consideration.

Need help? Review the Special Consideration page [HERE](#)

Assessments where Late Submissions will be accepted.

- Problem Set 1 – YES, Standard Late Penalty applies;
- Problem Set 2 – YES, Standard Late Penalty applies;
- Individual Project – YES, Standard Late Penalty applies.

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. If you experience circumstances or events that affect your ability to complete the assessments in this unit on time, please inform the convenor or tutor.

Written Assessments/Quizzes/Tests: If you experience circumstances or events that affect your ability to complete the written assessments in this unit on time, please inform the convenor or tutor.

Assessment Tasks

Name	Weighting	Hurdle	Due	Groupwork/Individual	Short Extension	AI Approach
Problem Set 1	20%	No	23/10/2026	Individual	Yes	Open
Problem Set 2	30%	No	13/11/2026	Individual	Yes	Open
Individual Project	50%	No	04/12/2026	Individual	Yes	Open

Problem Set 1

Assessment Type ¹: Problem-based task

Indicative Time on Task ²: 10 hours

Due: **23/10/2026**

Weighting: **20%**

Groupwork/Individual: Individual

Short extension ³: Yes

AI Approach: Open

This task will test your ability to use data technologies to analyse provided problems.

On successful completion you will be able to:

- Translate statistical thinking into coding syntax and appropriate data structures
- Import, clean, and transform real-world datasets using R
- Develop and implement computational strategies to tackle data challenges
- Generate meaningful visualisations using R and communicate results to diverse audiences

Problem Set 2

Assessment Type ¹: Problem-based task

Indicative Time on Task ²: 10 hours

Due: **13/11/2026**

Weighting: **30%**

Groupwork/Individual: Individual

Short extension ³: Yes

AI Approach: Open

This task will test your ability to use data technologies to analyse provided problems.

On successful completion you will be able to:

- Translate statistical thinking into coding syntax and appropriate data structures
- Compose SQL queries to manage and retrieve data from relational databases
- Import, clean, and transform real-world datasets using R
- Develop and implement computational strategies to tackle data challenges
- Generate meaningful visualisations using R and communicate results to diverse audiences
- Apply good practice of statistical project workflow

Individual Project

Assessment Type ¹: Portfolio

Indicative Time on Task ²: 40 hours

Due: **04/12/2026**

Weighting: **50%**

Groupwork/Individual: Individual

Short extension ³: Yes

AI Approach: Open

You will be assigned a few statistical challenges. You will be required to study these problems using appropriate statistical and computational techniques implemented with statistical software

and other data technologies.

On successful completion you will be able to:

- Translate statistical thinking into coding syntax and appropriate data structures
- Compose SQL queries to manage and retrieve data from relational databases
- Import, clean, and transform real-world datasets using R
- Develop and implement computational strategies to tackle data challenges
- Generate meaningful visualisations using R and communicate results to diverse audiences
- Apply good practice of statistical project workflow

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
- [Academic Success](#) 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.

³ An automatic short extension is available for some assessments. Apply through the [Service Connect Portal](#).

Delivery and Resources

Classes

Workshops (beginning in Week 1): There is one 1-hour workshop each week. During this hour, we will provide a brief overview of the week's content, run a Q&A session, and engage in some interesting activities together.

Pre-recorded Lectures (released weekly from Week 1): There are no formal live lectures scheduled for this unit. Each week we will have some video recordings covering the unit materials.

SGTA classes (beginning in Week 2): Students must register in and attend one 2-hour class per week.

Suggested textbooks

The following textbooks are useful as supplementary resources, for additional questions and explanations. They are available from the Macquarie University library:

- Garrett Grolemund, Hadley Wickham & Mine Çetinkaya-Rundel (2023) R for Data Science, 2nd Edition. O'Reilly Media, Inc. <https://r4ds.hadley.nz>

- Golemund, G. (2014) Hands-on programming with R. 1st edition. Sebastopol, CA, O'Reilly Media. <https://rstudio-education.github.io/hopr/>
- Wickham, H., Navarro, D. & Pedersen T.L. (2025) ggplot2 Elegant Graphics for Data Analysis. 3rd ed (work in progress). Springer. <https://ggplot2-book.org>
- Wickham, H. & Bryan, J. (2023) R packages. 2nd ed. Sebastopol, CA, O'Reilly. <https://r-pkgs.org>

Technology Used and Required

This subject requires the use of the following computer software:

- **R:** R is a free statistical software package. Access and installation instructions may be found at: <https://www.r-project.org/>
- **RStudio:** RStudio is an open-source tool that is used to manage and present work performed using R. Access and installation instructions may be found at <https://rstudio.com/products/rstudio/download/>
- **LaTeX:** LaTeX is a free mathematical typesetting program. Access and installation instructions may be found at: <https://www.latex-project.org/get/>
- **Git:** Git is a [free and open source](#) distributed version control system designed to handle everything from small to very large projects with speed and efficiency.
- **Quarto:** An open-source scientific and technical publishing system, installed by default with the latest release of RStudio.

Students are invited to bring their own devices (BYOD) and a laptop is recommended. Acceptable platforms are Windows, Linux and Mac.

Communication

We will communicate with you via your university email or through announcements on iLearn. Queries to convenors can either be placed on the iLearn discussion forum or sent to your lecturers from your university email address.

Unit Schedule

The following schedule is subject to change.

Week	Topics	Assessment
1	Meet your toolkits: R, RStudio, Git, Quarto	
2	The Basics + Flow Control	
3	Summarise, tidy and transform data with tidyverse	
4	Create elegant data visualisation with ggplot2	
5	Iteration with purrr , a functional programming toolkit	Problem Set 1 due

Week	Topics	Assesement
6	Retrive data from a database with DBI and dbplyr	
7	Writing R packages	
8	Introduction to Markdown	Problem Set 2 due
9	Reproducible report with Quarto	
10	Quarto Presentation in beamer and revealjs	
11	Storytelling with data	
12	Web scrapping with rvest	Project due
13	Looking beyond ST4DS	

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 or if you are a Global MBA student contact 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/>

Academic Success

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

- [Ask a Study Coach](#)
- [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 <https://students.mq.edu.au/support/technology/service-desk>.

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

This is the first time offering of this unit at NMJI

Unit information based on version 2026.04 of the [Handbook](#)