# STAT302

## Graphics, Multivariate Methods and Data Mining

S2 Day 2019

*Dept of Mathematics and Statistics*

## Contents

**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 Convenor
Nino Kordzakhia
nino.kordzakhia@mq.edu.au
Contact via nino.kordzakhia@mq.edu.au
Room 610 L6 E7A 12 Wally's Walk
TBA

Lecturer
Balamehala Pasupathy
balamehala.pasupathy@mq.edu.au
Contact via balamehala.pasupathy@mq.edu.au
TBA

Credit points
3

Prerequisites
6cp at 200 level including (STAT270 or STAT271 or BIOL235(P) or PSY222 or PSY248(P))

Corequisites

Co-badged status

Unit description
This unit introduces statistical tools for multivariate data analysis such as statistical graphics, discriminant analysis, principal component analysis, cluster analysis and an introduction to data mining, especially classification. Statistical packages are used extensively to illustrate the concepts in lectures and tutorials. Students are given opportunities to share their learning with their peers in tutorials, by presenting to their peers such as solutions to a problem posed in an earlier class, or summary of a specific weeks' learning in one or two slides, similar to 3 minute thesis presentations.

## 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:

Understand the principles underlying graphics, multivariate methods and data mining;

Choose the appropriate statistical analysis, for a given data set, from a wide range of methods based on multivariate methods and data mining;

Choose appropriate graphical techniques for displaying data;

Use a statistical computer package to carry out chosen analyses and interpret the results with understanding; present the results of analyses in a form which is suitable for publication;

Apply statistical techniques to problems arising from diverse fields of research.

# Assessment Tasks

| Name | Weighting | Hurdle | Due |
|---|---|---|---|
| SGTA Work | 10% | No | Weeks 3, 5, 7 and 10 |
| Mid-Semester Test | 30% | No | Week 8 |
| Practical Test | 60% | No | Week 12 |

## SGTA Work

Due: **Weeks 3, 5, 7 and 10**
Weighting: **10%**

There will be a set of SGTA exercises to submit in Week 3, 5, 7 and 10 worth 2.5% each. SGTA Works are worth 10% in total.

In the case when a student is unable to submit the SGTA Work due to unavoidable circumstances, the student must apply for Special Consideration via https://ask.mq.edu.au/

**Late Submission of Work**

The SGTA Works must be submitted by the official due date and time.

**No marks will be given to late work unless an extension has been granted following a successful application for Special Consideration.**

On successful completion you will be able to:

• Understand the principles underlying graphics, multivariate methods and data mining;

• Choose the appropriate statistical analysis, for a given data set, from a wide range of methods based on multivariate methods and data mining;

• Choose appropriate graphical techniques for displaying data;

• Use a statistical computer package to carry out chosen analyses and interpret the results

with understanding; present the results of analyses in a form which is suitable for publication;

- Apply statistical techniques to problems arising from diverse fields of research.

# Mid-Semester Test

Due: **Week 8**
Weighting: **30%**

The Mid-Semester Test will be 50 minutes long. This is a closed book exam, except you are permitted ONE A4 page of paper containing reference material printed or handwritten on both sides. Calculators will be needed but must not be of the text/programmable type.

If you cannot sit the Mid-Semester Test due to unavoidable circumstances you will need to apply for special consideration

https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policies/special-consideration

On successful completion you will be able to:

- Understand the principles underlying graphics, multivariate methods and data mining;
- Choose the appropriate statistical analysis, for a given data set, from a wide range of methods based on multivariate methods and data mining;
- Choose appropriate graphical techniques for displaying data;
- Apply statistical techniques to problems arising from diverse fields of research.

# Practical Test

Due: **Week 12**
Weighting: **60%**

The Practical Test will be held in Week 12, in a PC lab.

Permitted materials include the lecture notes and other teaching material, which will be distributed via *iLearn*.

Other conditions of the practical test will be specified in the test notification issued closer to the date via *iLearn.*

If you cannot sit the practical test due to unavoidable circumstances you will need to apply for special consideration

https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policies/special-consideration

**Important**

If you receive special consideration for the practical test, the supplementary Practical Test will be scheduled in the interval between the regular exam period and the start of the next session.

By making a special consideration application for the practical test you are declaring yourself available for a resit during the designated time period and will not be eligible for a second special consideration approval based on pre-existing commitments.

On successful completion you will be able to:

- Understand the principles underlying graphics, multivariate methods and data mining;
- Choose the appropriate statistical analysis, for a given data set, from a wide range of methods based on multivariate methods and data mining;
- Choose appropriate graphical techniques for displaying data;
- Use a statistical computer package to carry out chosen analyses and interpret the results with understanding; present the results of analyses in a form which is suitable for publication;
- Apply statistical techniques to problems arising from diverse fields of research.

# Delivery and Resources

**Classes**

**Lectures begin in Week 1.**

**SGTA classes begin in Week 2.**

Students must attend 2 hours of lectures and 2 hours of SGTA classes per week.

Lecture notes will be posted on *iLearn* site of the unit before the lecture.

Students should make sure they login at

https://ilearn.mq.edu.au/login/MQ/

regularly to access the teaching material.

 **Software:**

**SPSS** and **R**

There are no prescribed texts for this unit, but the following list provides useful references.

**Recommended texts:**

Chambers J M et al (1983) Graphical Methods for Data Analysis.

Cleveland W S (1994) Elements of Graphing Data.

Tufte E R (2001) The Visual Display of Quantitative Information.

Everitt B S et al (2001) Applied multivariate data analysis.

Johnson, R.A. & Wichern, D.W. (2002) Applied Multivariate Statistical Analysis.

Manly, B F J (2004) Multivariate Statistical Methods - A Primer.

# Unit Schedule

| WEEK | TOPICS | WORK DUE |
|------|--------|----------|
| 1 | Introduction | |
| 2 | Different graphical displays | |
| 3 | Displaying multivariate data | **SGTA Work** |
| **4** | Similarities and distances | |
| 5 | Hierarchical cluster analysis | **SGTA Work** |
| 6 | K-means clustering | |
| 7 | Eigenvalues and eigenvectors | **SGTA Work** |
| *Midsession Break – Two Weeks* | | |
| **8** | Principal component analysis | **Mid-Semester Test** |
| 9 | Principal component analysis cont. | |
| **10** | Discriminant analysis | **SGTA Work** |
| 11 | Classification Trees  Revision | |
| 12 | **Practical Test** | **Practical Test** |
| 13 | No classes | |

# Policies and Procedures

Macquarie University policies and procedures are accessible from Policy Central (https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central). 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 *(**Note:** The Special Consideration Policy is effective from 4 December 2017 and replaces the Disruption to Studies Policy.)*

Undergraduate students seeking more policy resources can visit the Student Policy Gateway (https://students.mq.edu.au/support/study/student-policy-gateway). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

If you would like to see all the policies relevant to Learning and Teaching visit Policy Central (https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central).

## Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: https://students.mq.edu.au/study/getting-started/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 improve your marks and take control of your study.

- Workshops

- StudyWise

- Academic Integrity Module for Students

- Ask a Learning Adviser

## 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.

# Graduate Capabilities

## Creative and Innovative

Our graduates will also be capable of creative thinking and of creating knowledge. They will be imaginative and open to experience and capable of innovation at work and in the community. We want them to be engaged in applying their critical, creative thinking.

This graduate capability is supported by:

### Learning outcomes

- Choose the appropriate statistical analysis, for a given data set, from a wide range of methods based on multivariate methods and data mining;
- Apply statistical techniques to problems arising from diverse fields of research.

### Assessment task

- SGTA Work

## Capable of Professional and Personal Judgement and Initiative

We want our graduates to have emotional intelligence and sound interpersonal skills and to demonstrate discernment and common sense in their professional and personal judgement. They will exercise initiative as needed. They will be capable of risk assessment, and be able to handle ambiguity and complexity, enabling them to be adaptable in diverse and changing environments.

This graduate capability is supported by:

### Learning outcomes

- Choose the appropriate statistical analysis, for a given data set, from a wide range of

methods based on multivariate methods and data mining;

- Choose appropriate graphical techniques for displaying data;
- Use a statistical computer package to carry out chosen analyses and interpret the results with understanding; present the results of analyses in a form which is suitable for publication;
- Apply statistical techniques to problems arising from diverse fields of research.

## Assessment tasks

- SGTA Work
- Mid-Semester Test
- Practical Test

# Commitment to Continuous Learning

Our graduates will have enquiring minds and a literate curiosity which will lead them to pursue knowledge for its own sake. They will continue to pursue learning in their careers and as they participate in the world. They will be capable of reflecting on their experiences and relationships with others and the environment, learning from them, and growing - personally, professionally and socially.

This graduate capability is supported by:

## Learning outcome

- Apply statistical techniques to problems arising from diverse fields of research.

## Assessment tasks

- SGTA Work
- Practical Test

# Discipline Specific Knowledge and Skills

Our graduates will take with them the intellectual development, depth and breadth of knowledge, scholarly understanding, and specific subject content in their chosen fields to make them competent and confident in their subject or profession. They will be able to demonstrate, where relevant, professional technical competence and meet professional standards. They will be able to articulate the structure of knowledge of their discipline, be able to adapt discipline-specific knowledge to novel situations, and be able to contribute from their discipline to inter-disciplinary solutions to problems.

This graduate capability is supported by:

## Learning outcomes

- Understand the principles underlying graphics, multivariate methods and data mining;
- Choose the appropriate statistical analysis, for a given data set, from a wide range of

methods based on multivariate methods and data mining;

- Choose appropriate graphical techniques for displaying data;

- Use a statistical computer package to carry out chosen analyses and interpret the results with understanding; present the results of analyses in a form which is suitable for publication;

- Apply statistical techniques to problems arising from diverse fields of research.

## Assessment tasks

- SGTA Work
- Mid-Semester Test

# Critical, Analytical and Integrative Thinking

We want our graduates to be capable of reasoning, questioning and analysing, and to integrate and synthesise learning and knowledge from a range of sources and environments; to be able to critique constraints, assumptions and limitations; to be able to think independently and systemically in relation to scholarly activity, in the workplace, and in the world. We want them to have a level of scientific and information technology literacy.

This graduate capability is supported by:

## Learning outcome

- Apply statistical techniques to problems arising from diverse fields of research.

## Assessment tasks

- SGTA Work
- Mid-Semester Test
- Practical Test

# Problem Solving and Research Capability

Our graduates should be capable of researching; of analysing, and interpreting and assessing data and information in various forms; of drawing connections across fields of knowledge; and they should be able to relate their knowledge to complex situations at work or in the world, in order to diagnose and solve problems. We want them to have the confidence to take the initiative in doing so, within an awareness of their own limitations.

This graduate capability is supported by:

## Learning outcome

- Apply statistical techniques to problems arising from diverse fields of research.

## Assessment tasks

- Mid-Semester Test

- Practical Test

# Effective Communication

We want to develop in our students the ability to communicate and convey their views in forms effective with different audiences. We want our graduates to take with them the capability to read, listen, question, gather and evaluate information resources in a variety of formats, assess, write clearly, speak effectively, and to use visual communication and communication technologies as appropriate.

This graduate capability is supported by:

## Assessment tasks

- SGTA Work
- Mid-Semester Test
- Practical Test

# Socially and Environmentally Active and Responsible

We want our graduates to be aware of and have respect for self and others; to be able to work with others as a leader and a team player; to have a sense of connectedness with others and country; and to have a sense of mutual obligation. Our graduates should be informed and active participants in moving society towards sustainability.

This graduate capability is supported by:

## Assessment task

- Practical Test

# Changes from Previous Offering

# Teaching and Learning Strategy

**Students are expected to**

- attend two-hour lectures (beginning in Week 1) and two-hour SGTA (beginning in Week 2). Attendance of SGTA classes is strongly recommended.
- complete the Assessment Tasks according to schedule.