STAT302
Graphics, Multivariate Methods and Data Mining
S2 Day 2013

Statistics

Contents

General Information  2
Learning Outcomes  2
Assessment Tasks  3
Delivery and Resources  7
Unit Schedule  9
Policies and Procedures  9
Graduate Capabilities  10
Teaching and Learning Strategy  15
Changes since First Published  15

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.
Unit guide STAT302 Graphics, Multivariate Methods and Data Mining

General Information

Unit convenor and teaching staff
Unit Convenor
Nino Kordzakhia
nino.kordzakhia@mq.edu.au
Contact via nino.kordzakhia@mq.edu.au
E4A 537
Wednesdays 12 - 2pm

Credit points
3

Prerequisites
39cp including (STAT270(P) or STAT271(P) or BIOL235(P) or PSY222(P) or PSY248(P))

Corequisites

Co-badge status

Unit description
This unit is concerned with the structure of multivariate data which is explored graphically, analysed statistically and investigated using data mining methods. Multivariate methods covered include cluster analysis, principal components, and discriminant analysis. Knowledge of simple matrix algebra, although not essential, will be very helpful in understanding and working through these topics. Statistical packages are used extensively to illustrate the concepts in lectures and tutorials.

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

## Assessment Tasks

<table>
<thead>
<tr>
<th>Name</th>
<th>Weighting</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial Participation</td>
<td>10%</td>
<td>Weekly</td>
</tr>
<tr>
<td>Short Class Tests (x2)</td>
<td>10%</td>
<td>Random</td>
</tr>
<tr>
<td>Assignments (x2)</td>
<td>10%</td>
<td>As indicated</td>
</tr>
<tr>
<td>Practical Test</td>
<td>10%</td>
<td>Week 13</td>
</tr>
<tr>
<td>Final Examination</td>
<td>60%</td>
<td>Exam Timetable</td>
</tr>
</tbody>
</table>

### Tutorial Participation

**Due:** Weekly  
**Weighting:** 10%

### Assessment Task

**Weight**

**Due**

Homework and Tutorial Participation  
10%  
Weekly

*Weeks 2 – 12 inclusive.*

Every week tutorial participation will be monitored and most weeks there will be set homework to submit to the lecturer at the start of the following lecture.

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.
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.

Assignments (x2)

Due: As indicated
Weighting: 10%

Assessment Task

Weight

Due

Assignment 1
5%
Assignments There will be two individual assignments due in weeks 5 and 11. The assignment questions will be made available through iLearn.

There is no “group work” assessment in this unit. All work is to be the student’s own.

Marked assignments will be returned to students in the tutorials. On-time submission of assignments is compulsory.

Extensions may be given in exceptional circumstances, students need to apply for special consideration online at ask.mq.edu.au. An extension will not be considered unless the student applies for special consideration.

In circumstances when students cannot complete assignments on time due to unavoidable disruption they will need to consider applying for Special Consideration. Information about the special consideration process is available at http://www.mq.edu.au/policy/docs/special_consideration/policy.html.

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.

**Practical Test**

Due: **Week 13**
Weighting: **10%**

The practical test, using pcs, will be held in the tutorial in Week 13 and covers the first 12 weeks of material in the unit. Permitted materials are a calculator, lecture notes, assignments, assignment solutions and tutorial solutions.

**Assessment Task**

**Weight**

**Due**

Practical Test

10%

W13, 4 Nov

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.

**Final Examination**

Due: **Exam Timetable**

Weighting: **60%**

**Assessment Task**

**Weight**

**Due**

Final Examination

60%

As timetabled

The examination will examine any material covered throughout the course. Students may bring
one A4 sized sheet of hand written notes, formulae, etc., which may be written on both sides and is easily readable. This summary must be submitted with your exam paper and is marked. No other materials such as lecture notes and textbooks are permitted.

Calculators will be needed but must not be of the text/programmable type.

You are expected to present yourself for examination at the time and place designated in the University Examination Timetable. The timetable will be available in draft form approximately eight weeks before the commencement of the examinations and in final form approximately four weeks before the commencement of the examinations. http://www.exams.mq.edu.au/

The only exemptions to sitting an Examination at the designated time are because of documented illness or unavoidable disruption. In these circumstances you may wish to consider applying for Special Consideration. Information about unavoidable disruption and the special consideration process is available at http://www.mq.edu.au/policy/docs/special_consideration/policy.html.

Students need to apply for special consideration online at https://ask.mq.edu.au

If a Supplementary Examination is granted as a result of the Special Consideration process the Examination will be scheduled after the conclusion of the official examination period. Special Consideration will only be granted to students whose performance in all parts of the coursework is satisfactory.

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

In this offering, two short class tests run randomly will replace one of three individual assignments offered in the past.

Lecture notes will be posted on iLearn site of the unit. Students should make sure they login at https://ilearn.mq.edu.au/login/MQ regularly to access the materials and read posts.

There are no prescribed texts for this unit, but the following list provides useful references, which are available in Special Reserve in the Library.
**Recommended texts:**


**Software:**

SPSS, SPSS Modeler, R (an open source, it is available from [http://cran.r-project.org/](http://cran.r-project.org/))

**Computer Labs:**

You may want to bring a memory stick when using the computers in computer labs.

Location: All Computer Labs are located in Building E4B.

**Lab opening hours:**

<table>
<thead>
<tr>
<th></th>
<th>Term Time (Teaching &amp; Exam)</th>
<th>Outside Term Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday-Friday</td>
<td>8 am-10 pm</td>
<td>9 am-7 pm</td>
</tr>
<tr>
<td>Weekend</td>
<td>9 am-5 pm</td>
<td>9 am-5 pm</td>
</tr>
</tbody>
</table>

Students should be aware that classes will be scheduled intermittently in these rooms. During these times, students not involved in the relevant classes must immediately vacate the rooms, and refrain from any action that may disrupt a class in progress.

Students may occupy rooms for individual coursework when no classes are scheduled.

**WARNING:** students are strongly advised not to remain alone in the labs after normal office hours. You should seek out a lab that has other students working in it and/or has a lab monitor.
## Unit Schedule

<table>
<thead>
<tr>
<th>WEEK</th>
<th>TOPICS</th>
<th>WORK DUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (29 Jul)</td>
<td>Introduction &amp; presenting data numerically</td>
<td></td>
</tr>
<tr>
<td>2 (5 Aug)</td>
<td>Different graphical displays</td>
<td></td>
</tr>
<tr>
<td>3 (12 Aug)</td>
<td>Displaying multivariate data</td>
<td></td>
</tr>
<tr>
<td>4 (19 Aug)</td>
<td>Similarities and distances</td>
<td></td>
</tr>
<tr>
<td>5 (26 Aug)</td>
<td>Hierarchical cluster analysis</td>
<td>10:05 Assignment 1</td>
</tr>
<tr>
<td>6 (2 Sept)</td>
<td>K-means clustering</td>
<td></td>
</tr>
<tr>
<td>7 (9 Sept)</td>
<td>Eigenvalues and eigenvectors</td>
<td></td>
</tr>
</tbody>
</table>

*Midsemester Break – Two Weeks*

<table>
<thead>
<tr>
<th>WEEK</th>
<th>TOPICS</th>
<th>WORK DUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (1 Oct)</td>
<td>Principal component analysis</td>
<td></td>
</tr>
<tr>
<td>9 (7 Oct)</td>
<td>Labor Day Public Holiday</td>
<td></td>
</tr>
<tr>
<td>10 (14 Oct)</td>
<td>Discriminant analysis</td>
<td></td>
</tr>
<tr>
<td>11 (21 Oct)</td>
<td>Multiple discriminant analysis</td>
<td>10:05 Assignment 2</td>
</tr>
<tr>
<td>12 (28 Oct)</td>
<td>Classification and regression trees</td>
<td></td>
</tr>
<tr>
<td>13 (4 Nov)</td>
<td>Review</td>
<td>Practical Test</td>
</tr>
</tbody>
</table>

## Policies and Procedures

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


Graduation Capabilities

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:

Student Support

Macquarie University provides a range of Academic Student Support Services. Details of these services can be accessed at: http://students.mq.edu.au/support/

UniWISE provides:

- Online learning resources and academic skills workshops http://www.students.mq.edu.au/support/learning_skills/
- Personal assistance with your learning & study related questions.
- The Learning Help Desk is located in the Library foyer (level 2).
- Online and on-campus orientation events run by Mentors@Macquarie.

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

Details of these services can be accessed at http://www.student.mq.edu.au/ses/.

IT Help

If you wish to receive IT help, we would be glad to assist you at http://informatics.mq.edu.au/help/.

When using the university's IT, you must adhere to the Acceptable Use Policy. The policy applies to all who connect to the MQ network including students and it outlines what can be done.

Graduate Capabilities

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

• 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

• Tutorial Participation
• Short Class Tests (x2)
• Assignments (x2)
• Practical Test
• Final Examination

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

• Tutorial Participation
• Assignments (x2)
• Practical Test
• Final Examination

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

- Tutorial Participation
- Final Examination

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 outcomes

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

- Tutorial Participation
- Short Class Tests (x2)
- Practical Test
- Final Examination

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 outcomes**

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

- Short Class Tests (x2)
- Assignments (x2)
- Practical Test
- Final Examination

**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**

- 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**

- Tutorial Participation
- Assignments (x2)
- 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:

**Learning outcome**

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

**Assessment tasks**

- Tutorial Participation
- Short Class Tests (x2)
- Assignments (x2)
- Practical Test

Engaged and Ethical Local and Global citizens

As local citizens our graduates will be aware of indigenous perspectives and of the nation's historical context. They will be engaged with the challenges of contemporary society and with knowledge and ideas. We want our graduates to have respect for diversity, to be open-minded, sensitive to others and inclusive, and to be open to other cultures and perspectives: they should have a level of cultural literacy. Our graduates should be aware of disadvantage and social justice, and be willing to participate to help create a wiser and better society.

This graduate capability is supported by:

**Learning outcomes**

- 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.

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:

**Learning outcome**

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

• Assignments (x2)
• Practical Test
• Final Examination

Teaching and Learning Strategy

All unit related queries should be directed to the unit convenor Dr Nino Kordzakhia using the Macquarie University e-mail system.

Lectures begin in Week 1.

Tutorials (1 x 2 hour tutorial) will start in the second week. In weeks 2 to 13 you will be required to submit homework and this work will also count towards your assessment.

The timetable for classes can be found on the University web site at https://timetables.mq.edu.au/2013/

Students are expected to:

• attend all the lectures (beginning in Week 1) and tutorials (beginning in Week 2);
• submit assignment solutions by due dates;
• submit homeworks by due dates;
• contact the unit convenor in advance if, for any reason, they cannot hand in their assessment tasks on time.

Changes since First Published

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/08/2013</td>
<td>Removed blank lines.</td>
</tr>
</tbody>
</table>