Division of Economic and Financial studies

STAT 302: Data Mining and Graphics

Second Semester 2005

Unit Outline

Unit Convenor: Kehui Luo
Frederick Wong

Students in this unit should read this unit outline carefully at the start of semester. It contains important information about the unit. If anything in it is unclear, please consult one of the teaching staff in the unit.

ABOUT THIS UNIT

The Graphics section introduces some of the concepts involved in the production of statistical graphics, especially those for multivariate data. This section will be mostly of a practical nature, but will also cover some theoretical material. The tutorial work will be mostly computational in nature, using the statistical package SPSS. Most of the graphical techniques are, however, easily performed by hand (for small data sets), and you may be required to do them by hand in the final examination.

An introduction to Data Mining and Multivariate Analysis follows the Graphics section. This section of the unit provides an introduction to various topics appropriate to the analysis of multivariate data. The emphasis will be on the application of these techniques to multivariate data sets, which we shall accomplish using **SPSS** and **Clementine**. Emphasis will also be placed on the presentation and interpretation of graphics, and results of statistical analyses in report form.

TEACHING STAFF

Lecturer: Dr Frederick Wong
Room C5C 473, phone: 9850 8557
E-mail: fwong@efs.mq.edu.au

Convenor: Dr Kehui Luo
Room C5C 460, phone: 9850 8563
E-mail: kluo@efs.mq.edu.au

Lectures

Stat302 requires students to attend two hours of lecture each week, held on Tuesdays from 3pm – 5pm in E7B 100.

REFERENCES

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

Chambers J M et al (1983) Graphical Methods for Data Analysis
Cleveland W S (1994) The Elements of Graphing Data
Cleveland W S & McGill M E (1988) Dynamic Graphics for Statistics

Du Toit S H C et al (1986) Graphical Exploratory Data Analysis
Ehrenberg A S C (1982) A Primer in Data Reduction
Everitt B S (1978) Graphical Techniques for Multivariate Data
Tufte E R (1983) The Visual Display of Quantitative Information
Manly, B F J (1994) Multivariate Statistical Methods - A Primer
Johnson, R.A. & Wichern, D.W. (2002) Applied Multivariate Statistical Analysis
Johnson, D.E (1998) Applied Multivariate Methods for Data Analysts

ASSESSMENT

The overall assessment for Stat302 is as follows:

Assignments	15%
Homework and Tutorial Participation	10%
Practical Test	15%
Examination	60%

ASSIGNMENTS

Three assignments will be set and on-time submission is compulsory. Students who are unable to submit any assignment on time because of illness or some other cause must report the circumstances in writing to the lecturer, and documentation must also be provided to the Registrar, through the Student Centre.

TUTORIALS

Each student is required to participate in a tutorial session of two hours per week. Times are

Wednesday 12pm – 2pm ***or***
Thursday 1pm – 3pm in

The tutorial sessions are compulsory, and they will be held in **W6B301** laboratory. You will be given an account to use the computers there.

Tutorial exercises will be set each week to be completed during the tutorial session. Some parts of these exercises will need to be written up and submitted at the beginning of the lecture in the following week.

A practical test will be held during the Week 13 tutorials. It will be based on the skills you will learn in Weeks 2 - 12.

EXAMINATION

This will examine any material covered throughout the unit. You may bring one A4 size sheet of notes, formulas, etc, written on one side into the examination. However, any other materials such as lecture notes and text books are not permitted. Scientific calculators should be brought in, but they should not be of the text/programmable type.

Note carefully: In order to pass Stat302 students must satisfy each of the following requirements:

* Participate in tutorials.
* Submit assignments and tutorial exercises (on time).
* Perform satisfactorily (i e achieve pass standard) examination section (i.e. final exam and practical test).
* Perform satisfactorily (i e achieve pass standard) in the overall assessment.

PLAGIARISM

The University defines plagiarism in its rules: "Plagiarism involves using the work of another person and presenting it as one's own." Plagiarism is a serious breach of the University's rules and carries significant penalties. You must read the University's practices and procedures on plagiarism. These can be found in the *Handbook of Undergraduate Studies* or on the web at: http://www.student.mq.edu.au/plagiarism/

The policies and procedures explain what plagiarism is, how to avoid it, the procedures that will be taken in cases of suspected plagiarism, and the penalties if you are found guilty. Penalties may include a deduction of marks, failure in the unit, and/or referral to the University Discipline Committee.

UNIVERSITY POLICY ON GRADING

Academic Senate has a set of guidelines on the distribution of grades across the range from fail to high distinction. Your final result will include one of these grades plus a standardised numerical grade (SNG).

On occasion your raw mark for a unit (i.e., the total of your marks for each assessment item) may not be the same as the SNG which you receive. Under the Senate guidelines, results may be scaled to ensure that there is a degree of comparability across the university, so that units with the same past performances of their students should achieve similar results.

The process of scaling does not change the order of marks among students. A student who receives a higher raw mark than another will also receive a higher final scaled mark.

For an explanation of the policy see
http://www.mq.edu.au/senate/MQUonly/Issues/Guidelines2003.doc or
http://www.mq.edu.au/senate/MQUonly/Issues/detailedguidelines.doc.

Grading in STAT302: Your final grade will be based on your performance in various parts specified in the Assessment section. The grades are awarded according to rules set out in the Bachelor Degree Rules 10 (2) as follows:

HD – High Distinction: Denotes a performance that meets all unit objectives in such an exceptional way and with such marked excellence that it deserves the highest level of recognition.

D – Distinction: Denotes performance that clearly deserves a very high level of recognition as an excellent achievement in the unit.

C – Credit: Denotes performance that is substantially better than would normally be expected of competent students in the unit.

P – Pass: Denotes performance that satisfies unit objectives.

PC – Conceded Pass: Denotes performance that meets unit objectives only marginally.

F – Fail: Denotes performance which does not meet unit objectives.

Once your final grade has been decided, on the basis of your performance in the unit, you are given a *standardized numerical grade* (SNG). SNG is not a mark but a ranking of students.

The SNGs awarded in a particular unit are designed to indicate that the students in each performance band, from HD to F, have satisfied the criteria for inclusion in that band and ranks them by their performance within that band. Since the ranges of SNGs differ from band to band the relationship between raw marks and SNGs may differ from band to band even within the

same unit. The relationship between raw marks and SNGs would almost always differ between units.

STUDENT SUPPORT SERVICES

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

Visit regularly the unit web site, http://www.stat.mq.edu.au/units/stat302/, for possible announcements and updates. Assignment Solution will also be made available here in the due course.

Seek help from your lecturers or tutors sooner by seeing them in their office hours or make an appointment to see a staff at other times.