# ACST890

# Data Analytics for Finance and Insurance

S1 Evening 2019

*Department of Actuarial Studies and Business Analytics*

# 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 & lecturer
Sachi Purcal
sachi.purcal@mq.edu.au
Contact via sachi.purcal@mq.edu.au
4ER 615
See iLearn

Lecturer
Chi Truong
chi.truong@mq.edu.au
Contact via chi.truong@mq.edu.au
4ER 230
See iLearn

Credit points
4

Prerequisites
Admission to MActPrac or MAppStat

Corequisites

Co-badged status

Unit description
This unit focuses on the rapidly evolving area of high-performance commercial computing in finance and insurance and its associated `Big Data' and predictive analytics applications. It equips students with the necessary computing and statistical tools to be become active participants in the field. Students will solve `Big Data' problems at the Linux command line, harness the R language to perform data analytics and tease data out of relational databases using SQL. The unit examines statistical learning, and treats a variety of topics, from linear regression to support vector machines that, coupled with computing knowledge, take commercial supercomputing beyond analysing bounced cheques, discovering fraud or managing Facebook friends. The statistics learned will be applied using the R language, and R will also enable the study of various data visualisation applications, which are important to both understanding data and communicating findings.

## 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:

Construct Linux command-line solutions to data science problems

Devise R programming language code for data analytics and visualisation

Formulate SQL language approaches to relational database problems

Assemble statistical learning techniques to tackle data science problems

Select, describe, critically examine and employ a variety of data visualisation techniques

# General Assessment Information

**Gradebook**

It is the responsibility of students to view their marks for each within session assessment on iLearn within 20 working days of posting. If there are any discrepancies, students must contact the unit convenor immediately. Failure to do so will mean that queries received after the release of final results regarding assessment marks (not including the final exam mark) will not be addressed.

**Assessment criteria**

Assessment criteria for all assessment tasks will be provided on the unit iLearn site.

## Late submissions and extensions

**Tasks 10% or less** – No extensions will be granted. Students who have not submitted the task prior to the deadline will be awarded a mark of 0 for the task, except for cases in which an application for special consideration is made and approved.

**Tasks above 10%** - No extensions will be granted. There will be a deduction of 10% of the total available marks made from the total awarded mark for each 24 hour period or part thereof that the submission is late (for example, 25 hours late in submission – 20% penalty). This penalty does not apply for cases in which an application for special consideration is made and approved. No submission will be accepted after solutions have been posted.

These requirements mean that **no extensions will be granted** and as each task is worth less than 10% (7%, 3%, 7%, 3%, 3%, 2%, 3%, 2%, 3%, 2% and 5%), **missing a submission deadline will result in a mark of zero for the task**, except where an application for special consideration is made and approved.

**Special consideration**

Where a special consideration application is approved, the student may be offered an alternative assessment or may receive a mark based on the percentage mark achieved by the student in one or more other assessment tasks, at the unit convenor's discretion.

# Assessment Tasks

| Name | Weighting | Hurdle | Due |
|------|-----------|--------|-----|
| Tests (take home) | 20% | No | 21/3, 1/4, 16/5, 27/5/19 |
| Quizzes (take home) | 10% | No | 4/4, 15/4, 2/5, 13/5/19 |
| Assignment | 10% | No | 6/5, 20/5 and 27/5/19 |
| Final Examination | 60% | No | Exam period |

## Tests (take home)

Due: **21/3, 1/4, 16/5, 27/5/19**
Weighting: **20%**

You will use iLearn to submit your solutions (in a PDF file) to two (take home) tests. The material will be more theoretical than practical. If you submit your solutions in a file format other than PDF for a test you will be awarded a mark of zero for the test.

You will use iLearn's peer assessment tool to mark the work of your classmates (due 1/4/19 and 275/19).

Each test submission and your accompanying peer marking performance will total 10% of your final assessment. Of the 10%, your submission task is worth 7% and your peer marking task is worth 3%.

Your first task is due in week 4. Solutions and a marking guide will immediately be made available, giving you valuable feedback as to whether to continue in the unit beyond the census date.

On successful completion you will be able to:

- Construct Linux command-line solutions to data science problems
- Devise R programming language code for data analytics and visualisation
- Formulate SQL language approaches to relational database problems
- Assemble statistical learning techniques to tackle data science problems
- Select, describe, critically examine and employ a variety of data visualisation techniques

## Quizzes (take home)

Due: **4/4, 15/4, 2/5, 13/5/19**
Weighting: **10%**

You will use iLearn to submit solutions to two (take home) quizzes. Each will include a number of practical (coding) problems.

You will use iLearn's peer assessment tool to mark the work of your classmates (due 15/4/19 and 13/5/19).

Each quiz submission and your accompanying peer marking performance will total 5% of your final assessment. Of the 5%, your submission task is worth 3% and and your peer marking task is worth 2%.

On successful completion you will be able to:
- Construct Linux command-line solutions to data science problems
- Devise R programming language code for data analytics and visualisation
- Formulate SQL language approaches to relational database problems
- Assemble statistical learning techniques to tackle data science problems
- Select, describe, critically examine and employ a variety of data visualisation techniques

# Assignment

Due: **6/5, 20/5 and 27/5/19**
Weighting: **10%**

There will be three tasks. Your reports must be submitted electronically via iLearn.

For your second task, you will use iLearn's peer assessment tool to mark the work of your classmates (due 20/5/19).

The first task will be worth 3% of your final assessment, the second 2% and the third 5%.

On successful completion you will be able to:
- Construct Linux command-line solutions to data science problems
- Devise R programming language code for data analytics and visualisation
- Formulate SQL language approaches to relational database problems
- Assemble statistical learning techniques to tackle data science problems
- Select, describe, critically examine and employ a variety of data visualisation techniques

# Final Examination

Due: **Exam period**
Weighting: **60%**

A three hour final examination for this unit will be held during the University Examination period.

You are permitted ONE A4 page of paper containing reference material printed on both sides. The material may be handwritten or typed.

The page will not be returned to you at the end of the final examination.

On successful completion you will be able to:

- Construct Linux command-line solutions to data science problems

- Devise R programming language code for data analytics and visualisation

- Formulate SQL language approaches to relational database problems

- Assemble statistical learning techniques to tackle data science problems

- Select, describe, critically examine and employ a variety of data visualisation techniques

# Delivery and Resources

## Textbook

The textbook for this course is

- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013), *An Introduction to Statistical Learning with Applications in R*, Springer

which is available online or may be purchased.

## Technology Used and Required

### Tests

You need to electronically submit your solutions (using iLearn) to the (take home) tests as PDF files. Some possibilities to produce these files are: scan your handwritten solutions to a PDF file; use WORD (or similar software) to type out your solutions and save your work as a PDF file. For each test you can only submit one PDF file; it must contain all the pages of your solution (so make sure, if your scanner produces separate PDF pages, that you can combine them into one file). Your PDF file must be able to read by the current version of Adobe Acrobat Reader—check this is the case before submission (if it can't be read, then it can't be marked and you will be awarded a mark of zero for the assessment task).

# 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

## Supplementary exams

Information regarding supplementary exams, including dates, is available at: http://www.businessandeconomics.mq.edu.au/current_students/undergraduate/how_do_i/special_consideration

# 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

## PG - Discipline Knowledge and Skills

Our postgraduates will be able to demonstrate a significantly enhanced depth and breadth of knowledge, scholarly understanding, and specific subject content knowledge in their chosen fields.

This graduate capability is supported by:

### Learning outcomes

- Construct Linux command-line solutions to data science problems
- Devise R programming language code for data analytics and visualisation
- Formulate SQL language approaches to relational database problems
- Assemble statistical learning techniques to tackle data science problems
- Select, describe, critically examine and employ a variety of data visualisation techniques

### Assessment tasks

- Tests (take home)
- Quizzes (take home)
- Assignment
- Final Examination

## PG - Critical, Analytical and Integrative Thinking

Our postgraduates will be capable of utilising and reflecting on prior knowledge and experience, of applying higher level critical thinking skills, and of integrating and synthesising learning and knowledge from a range of sources and environments. A characteristic of this form of thinking is the generation of new, professionally oriented knowledge through personal or group-based critique of practice and theory.

This graduate capability is supported by:

## Learning outcomes

- Construct Linux command-line solutions to data science problems
- Devise R programming language code for data analytics and visualisation
- Formulate SQL language approaches to relational database problems
- Assemble statistical learning techniques to tackle data science problems
- Select, describe, critically examine and employ a variety of data visualisation techniques

## Assessment tasks

- Tests (take home)
- Quizzes (take home)
- Assignment
- Final Examination

# PG - Research and Problem Solving Capability

Our postgraduates will be capable of systematic enquiry; able to use research skills to create new knowledge that can be applied to real world issues, or contribute to a field of study or practice to enhance society. They will be capable of creative questioning, problem finding and problem solving.

This graduate capability is supported by:

## Learning outcomes

- Construct Linux command-line solutions to data science problems
- Devise R programming language code for data analytics and visualisation
- Formulate SQL language approaches to relational database problems
- Assemble statistical learning techniques to tackle data science problems
- Select, describe, critically examine and employ a variety of data visualisation techniques

## Assessment tasks

- Tests (take home)
- Quizzes (take home)
- Assignment
- Final Examination

# Research and Practice

This unit uses research by Macquarie University researchers and external sources (references will be given in the lectures, tutorials and assignment).