Special Topic: Introduction to Big Data Analysis - Course Outline
COMP 473: 2016 Trimester 1

This document sets out the workload and assessment requirements for COMP 473. It also provides contact information for staff involved in the course. If the contents of this document are altered during the course, you will be advised of the change by an announcement in lectures and/or on the course web site. A printed copy of this document is held in the School Office.

Objectives

By the end of the course, students should be able to:

- Understand the basic concepts of how and where Big Data arises

- Describe data structures and mechanisms for the capture, storage, processing, summary and visualisation of Big Data

- Implement practical methods for data acquisition and management using appropriate software (SQL, R, Python, Perl; Hadoop, Mapreduce; Weka)

- Understand basic methods of analysis of Big Data, including methods from machine learning for high dimensional data.

Textbook

There will be no textbook for COMP 473. Instead, a number of books can be used as a reference, which will be listed on the course website.

Lectures, Tutorials, Laboratories, and Practical work

A schedule of lecture topics, readings, and assignment due dates is available online

Lectures for COMP 473 are: *Wednesday 4-6pm, in RB 901 (library building)*

Tutorials and/or labs, and help desk will be announced when the course starts.

The trimester starts on Monday 29 Feb and ends on Friday 3 June. The exam will be held in 10 June - 29 June.

Assignments and Projects

There will be four assignments, each for a big part/aspect of the course from one university. They are designed for meeting the course objectives.

Item	Due
Assignment 1	10:00am Monday 4 April 2016
Assignment 2	10:00am Monday 2 May 2016
Assignment 3	10:00am Monday 30 May 2016
Assignment 4	18:00pm Friday 3 June 2016

Late assignments may not be accepted. You may apply for an aegrotat consideration in cases of illness or crises around the due date.

Workload

In order to maintain satisfactory progress in COMP 473, you should plan to spend an average of at least 10 hours per week on this paper. A plausible and approximate breakdown for these hours would be:
- Lectures and tutorials: 3
- Readings: 2
- Assignments: 4

School of Engineering and Computer Science

The School office is located on level three of the Cotton Building (Cotton 358).

Staff

The course is taught between four NZ Universities (UC, Otago, VUW and Massey) via videoconferencing. The course will be taught in four sections, one from each university. Each module will have a coordinator from that university. You should contact this person for any questions about their section.

- UC: Jennifer Brown Jennifer.brown@canterbury.ac.nz
- Otago: Matt Parry mparry@maths.otago.ac.nz
- VUW: Bing Xue Bing.Xue@ecs.vuw.ac.nz
- Massey: Mathieu Vignes M.Vignes@massey.ac.nz

At VUW, the course organiser for COMP 473 is Mengjie Zhang. The lecturers for the course are Bing Xue and Mengjie Zhang. Their contact details are:

- *Mengjie Zhang*
- Cotton 355
- +64 4 463 5464
- Mengjie.Zhang@ecs.vuw.ac.nz

- *Bing Xue*
- Cotton 352
- +64 4 463 5542
- Bing.Xue@ecs.vuw.ac.nz

Tutor details will be announced when the course starts.

Announcements and Communication

The main means of communication outside of lectures will be the COMP 473 web area at http://ecs.victoria.ac.nz/Courses/COMP473_2016T1/. There you will find, among other things, this document, the lecture schedule and assignment handouts, and the COMP 473 Forum. The forum is a web-based bulletin board system. Questions and comments can be posted to the forum, and staff will read these posts and frequently respond to them.

Assessment

There are four assignments, each worth 25%. There is no final exam. The assignments will involve considerable time and effort, but balance that with the fact there is no final exam. Assignments must be submitted in an electronic form through the Learn site. Your grade for COMP 473 will be determined based on the following assessment weightings:

Item	Weight
Assignment 1	25%
Assignment 2	25%
Assignment 3	25%
Assignment 4	25%

Tests and Exams

There are no term tests and the final exams for this course.

Practical Work

There will be four assignments, handed out on 9 March (week 2), 13 April (week 5), 11 May (week 8), and 25 May (week11), and due three weeks later except that the last one is relatively small [4 April (ass1), 2 May (ass2), 30 May (ass3) and 3 June (ass4)]. Each of these assignments is worth 25%.

All assignments must be handed in on time unless you have made a prior arrangement with the lecturer or have a valid medical excuse (for minor illnesses it is sufficient to discuss this with the lecturer). The penalty for assignments that are handed in late without prior arrangement is one grade reduction per day. Assignments that are more than one week late will not be marked.

Plagiarism

Working Together and Plagiarism

We encourage you to discuss the principles of the course and assignments with other students, to help and seek help with programming details, problems involving the lab machines. However, any work you hand in must be your own work.

The School policy on Plagiarism (claiming other people's work as your own) is available from the course home page. Please read it. We will penalise anyone we find plagiarising, whether from students currently doing the course, or from other sources. Students who knowingly allow other students to copy their work may also be penalised. If you have had help from someone else (other than a tutor), it is always safe to state the help that you got. For example, if you had help from someone else in writing a component of your code, it is not plagiarism as long as you state (eg, as a comment in the code) who helped you in writing the method.

Mandatory Course Requirements

This course has a mandatory requirement of submitting reasonable attempts for at least three out of the four assignments.

Any student who is concerned that they have been (or might be) unable to meet any of the MCRs because of exceptional personal circumstances, should contact the course coordinator as soon as possible.

Passing COMP 473

To pass COMP 473, a student must satisfy mandatory requirements and gain at least a **C-** grade overall.

Withdrawal

The last date for withdrawal from COMP 473 with entitlement to a refund of tuition fees is Friday 11 March 2016. The last date for withdrawal without being regarded as having failed the course is Friday 13 May 2016 -- though later withdrawals may be approved by the Dean in special circumstances.

Rules & Policies

Find key dates, explanations of grades and other useful information at http://www.victoria.ac.nz/home/study.

Find out about academic progress and restricted enrolment at http://www.victoria.ac.nz/home/study/academic-progress.

The University's statutes and policies are available at http://www.victoria.ac.nz/home/about/policy, except qualification statutes, which are available via the Calendar webpage at http://www.victoria.ac.nz/home/study/calendar (See Section C).

Further information about the University's academic processes can be found on the website of the Assistant Vice-Chancellor (Academic) at http://www.victoria.ac.nz/home/about/avcacademic

All students are expected to be familiar with the following regulations and policies, which are available from the school web site:

Grievances
Student and Staff Conduct
Meeting the Needs of Students with Disabilities
Student Support
Academic Integrity and Plagiarism
Dates and Deadlines including Withdrawal dates
School Laboratory Hours and Rules
Printing Allocations
Expectations of Students in ECS courses

The School of Engineering and Computer Science strives to anticipate all problems associated with its courses, laboratories and equipment. We hope you will find that your courses meet your expectations of a quality learning experience.

If you think we have overlooked something or would like to make a suggestion feel free to talk to your course organiser or lecturer.

Course Outline as PDF