



BUSINESS SCHOOL
Te Kura Pakihi

INFO 204: Introduction to Data Science

COURSE OUTLINE

Semester Two, 2021

Contents

Paper Description and Aims	1
Learning Outcomes	1
Teaching Staff	2
Course Delivery.....	2
Locations and Times	3
Expectations and Workload	3
Course Learning Resources	4
Blackboard.....	4
Assessment.....	4
Course Requirements.....	5
Special Consideration	5
Course Calendar.....	6
Disclaimer	7

Paper Description and Aims

INFO 204 covers introductory theory and methods for performing data-driven decision making. Data statistics, data manipulation and integration, learning algorithms, model evaluation, and ethical considerations of data-driven decision making, are discussed at an introductory level, providing the knowledge base and skill set as required for further exploration of data science.

The importance of data science, and data analytic thinking, is becoming increasingly evident in modern business environments. Businesses are relying upon data-driven decision making at an ever-increasing rate, so individuals with a mind towards data science thinking have a competitive advantage in industry. The role of data scientist has been reported by Forbes as “the top job” in recent years, and there are constant “data scientists” vacancies both within New Zealand and overseas.

In addition to being a core topic of Information Science, the concepts discussed in this paper would be of interest to a wide range of specialties, including: computer science, marketing, management, statistics and finance.

INFO 204 operates under a fairly traditional model – lectures, tutorials and labs will be used to present and discuss material relevant to the paper. Within these sessions, there will be opportunities for discussion, and where appropriate, classes may draw on more interactive techniques to support the content, or refer to external online content for context.

Learning Outcomes

Upon successful completion of this paper, you should be able to:

1. Define data science as a field that integrates concepts from information technology and statistical/machine learning and combines with organisational context
2. Describe the basic strengths and weaknesses of decision making based upon data science methodologies
3. Explain the ethical and behavioural impacts and opportunities for innovation that data science methods can introduce within small and large businesses
4. Perform basic data engineering for various data domains
5. Apply basic data-driven modelling techniques to solve classification and regression problems
6. Use appropriate visualisation and reporting techniques to convey knowledge acquired through data science processes.

Teaching Staff

Paper Coordinator and Lecturer

Name: Grant Dick
Office: 9.08, 9th Floor, Otago Business School
Email: grant.dick@otago.ac.nz
Office Hours: TBC – check Blackboard

Labs

Name: Chris Edwards
Office: 8.01, 8th Floor, Otago Business School
Email: chris.edwards@otago.ac.nz

Name: Abira Sengupta
Contact: Available during lab times

Other staff will be present in labs on a more ad hoc basis for support and general discussion. You will be introduced to these staff during your lab times.

Course Delivery

Every week students must attend *two 50 minute lectures* and a *single 110 minute lab* (see '[Locations and Times](#)' for details of days, times, and rooms):

- **Lectures** present the key conceptual material through discussion and interaction between teaching staff and students. Multimedia and class exercises may be used to supplement the presentation. Lectures are occasionally supported by readings as indicated on Blackboard.
- **Labs** are interactive, collaborative sessions in which students attempt to cement concepts presented at lectures with their peers in a supportive environment. Labs are also used to discuss and work on assignment tasks.
You are expected to attend the lab in your assigned lab stream. You are also welcome to attend more than one lab stream if you feel that you need more time to focus on the concepts discussed in a given lab as long as spare seats are available in the additional lab stream you are attending.
- **Tutorials** are interactive, collaborative sessions in which students attempt to cement concepts presented at lectures and labs through conceptual tasks with their peers in a supportive environment.

The [Course Calendar](#) (page 6) details semester dates, class topics, and assessment related scheduling information. **Note that this calendar may change as the course proceeds.** These will be announced at lectures and detailed on Blackboard.

Students are required to prepare for and attend all classes to gain full benefit from the course.

These activities should be prepared for by reviewing information detailed on Blackboard and completing any assigned readings. Students unable to attend a lecture are expected to catch up on missed material. Unless stated otherwise, all aspects of the course are examinable (see below).

Locations and Times

Your personal timetable is in [eVision](#) – please check that for your streamed lab time and location. More information about the locations of rooms can be found at the University's [List of Lecture Theatres](#) page.

- Monday Lectures are in Burns 7 Lecture Theatre
- Tuesday Lectures are in Archway 1
- Labs run in either [Arts CAL](#) or OBS 3.27 (check your [eVision](#) timetable!)
- Tutorials run in ME2.02 in the Mellor Labs

	Monday	Tuesday	Wednesday	Thursday	Friday
9am					
10am				Lab CA3	
11am			Lab CA1	OBS3.27	Lab CA4
12pm			Arts CAL	Tutorial TA1	OBS3.27
1pm			Lab CA2	Tutorial TA2	
2pm	Lecture LA1	Lecture LA2	OBS3.27		
3pm					
4pm					
5pm					

Figure 1 Lecture and lab timetable. Attend one class of each colour each week.

Expectations and Workload

The teaching team (see [Teaching Staff](#)) are committed to creating the best possible environment to facilitate student learning. However, learning is a joint activity that requires active participation from the learner. Students are therefore expected to attend and participate in all facets of the course. This includes activities such as revising material, completing assigned work, spending extra time researching difficult concepts, in addition to participating in lectures, labs, and assessments.

INFO 204 is worth 18 points, which equates to spending 12 hours per week on the course (in accordance with University guidelines). This calculation includes the mandatory contact hours of lectures, labs, and tutorials totalling 5 hours per week. The remaining 7 hours should be used for reading (assigned and personal research), lab preparation and completion, and course revision.

Course Learning Resources

Blackboard

Blackboard <https://blackboard.otago.ac.nz/> provides you with access to course materials, class notices, and resources. Blackboard is used to email the whole class, so it is important that you check your student email and *Blackboard* regularly.

More general-purpose information related to your studies can be found on Blackboard under the "Study-Related Information" section.

Assessment

All material presented is examinable (except where stated otherwise) by internal assessment and the final examination. All relevant assessment information such as due dates and times, content, guidelines and so on will be discussed at lectures and, where appropriate, detailed on Blackboard. *Students are responsible for ensuring that they are aware of this information, keeping track of their own progress, and catching up on any missed classes.*

Unless otherwise stated, assessment deadlines are 5pm on Friday – please check on Blackboard to confirm deadlines.

Assessment	Due date	% of final grade	Requirements to pass this paper
Lab assessments	Posted on Blackboard for each lab	(best 10 out of 11 labs) 10 * 2 = 20	Terms (see Course Requirements below)
Assignment 1	Week 34	10	
Assignment 2	Week 41	15	
Practical Test	Week 36	15	
Final Exam	TBA	40	Exam Hurdle (see Course Requirements below)

Course Requirements

INFO 204 has the following *terms requirements*.

1. You must at least 7 lab assessments (out of 11). Students *failing to participate in at least seven lab assessments will not be allowed to sit the final exam.*

INFO 204 has the following *exam hurdle*.

1. You must *achieve a mark of 40% or better in the final examination to pass the paper, regardless of your internal assessment performance.*

Special Consideration

On certain grounds (e.g., illness or bereavement), special consideration on internal assessment may be given. Special consideration is given on a case-by-case basis and *must be requested from the course coordinator prior to the assessment deadline.*

Course Calendar

Week	Lecture	Lab	Tut.	Assessment Details
28 12-Jul 16-Jul	1. Introduction 2. Data Science Fundamentals	Getting Started		
29 19-Jul 23-Jul	3. Data Wrangling 4. Statistical and Machine Learning	Statistics and Plotting in Python	T ₁	
30 26-Jul 30-Jul	5. Supervised Learning I 6. Supervised Learning II	Data Wrangling	T ₂	
31 02-Aug 06-Aug	7. Supervised Learning III 8. Supervised Learning IV	Regression	T ₃	
32 09-Aug 13-Aug	9. Neural Networks and Deep Learning I 10. Neural Networks and Deep Learning II	Classification	T ₄	
33 16-Aug 20-Aug	11. Neural Networks and Deep Learning III 12. Model Selection	Neural Networks	T ₅	
34 23-Aug 27-Aug	13. Feature Engineering 14. Part I revision	Model Selection and Feature Engineering	T ₆	<i>Assignment 1 due by 5pm Fri</i>
35 30-Aug 03-Sep	Mid-Semester Break			
36 06-Sep 10-Sep	15. Unsupervised Learning I 16. Unsupervised Learning II	Catch-Up (not assessed)	T ₇	<i>Practical Test</i>
37 13-Sep 17-Sep	17. Unsupervised Learning III 18. Unsupervised Learning IV	Clustering and Dimension Reduction	T ₈	
38 20-Sep 24-Sep	19. Ensemble Learning 20. Non-Tabular Data	Ensemble Learning	T ₉	
39 27-Sep 01-Oct	21. Text Mining I 22. Text Mining II	Spam Detection	T ₁₀	
40 04-Oct 08-Oct	23. Pipelines for Data Science and Machine Learning 24. Optimisation	Optimisation	T ₁₁	
41 11-Oct 15-Oct	25. Data Science Ethics 26. Review		T ₁₂	<i>Assignment 2 due by 4pm Monday</i>

Disclaimer

While every effort is made to ensure that the information contained in this document is accurate, it is subject to change. Changes will be notified in class and via Blackboard. Students are encouraged to check Blackboard regularly. It is the student's responsibility to be informed.