School of Chemical Engineering

UNSW Engineering

CEIC6789

Data-driven Decision Making in Chemical Engineering and Food Science

Term 2, 2022

Course Overview

Staff Contact Details

Convenors

Name	Email	Availability	Location	Phone
Priyank Vijaya Kumar	priyank.kumar@unsw.edu.au	Via Teams or Email	Science and Engineering Building, E8, 334	

School Contact Information

Enquiries related to the course (e.g. course content, assessment instructions) should be raised during the scheduled classes, office hours, or in Teams channels/Moodle forums designated for that purpose.

Learning and question etiquette:

- Please be prepared for classes and attend the timetabled classes so that you can ask questions during the class time.
- Please respect that demonstrators and tutors have scheduled the class time to help you learn and are likely to be busy with other responsibilities outside those times; questions asked outside of class times will take longer to be answered.
- PhD students and other casuals who are teaching classes are normally only expected to look after the timetabled class and not to provide follow-up one-on-one assistance.
- Please don't ask questions in private that could be reasonably asked in a way that everyone can learn from the discussion.
- As a member of a community of learners, please try answering each other's questions!
- Please limit private messages to staff (via email or Teams) to *confidential* matters related to course administration.

For assistance with enrolment, class registration, progression checks and other administrative matters, please see the Nucleus: Student Hub. They are located inside the Library – first right as you enter the main library entrance. You can also contact them via http://unsw.to/webforms or reserve a place in the face-to-face queue using the UniVerse app.

If circumstances outside your control impact on submitting assessments, Special Consideration may be granted, usually in the form of an extension or a supplementary assessment. Applications for Special Consideration must be submitted online.

Course Details

Units of Credit 6

Summary of the Course

The recent surge in the volume of data collected owing to technological advances provides opportunities to improve processes and take better decisions across various industries. However, in order to turn large data sets into useful insights, combining the knowledge of right data with right analytical tools is important. Data-driven decision making is an industry-oriented course where students learn data management and analytic skills through a major project and real case studies from the School's research strengths/industrial experience in chemical engineering and food science. The course covers advanced methods for obtaining, handling and summarising various categories of data with databases. The course will also focus on how to analyse the collected data efficiently by applying sophisticated analytical techniques including statistical tests, inferences and regression analysis.

Course Aims

CEIC6789 is an undergraduate elective/postgraduate disciplinary knowledge course where the primary aim is for the students to learn data management and analytic skills in chemical engineering and food science through real case studies and hands-on coding exercises. The course builds upon previous mathematics, computing and experimental courses to develop skills in obtaining, handling and summarizing various categories of data through the use of databases. The course will also focus on how to analyze the collected data efficiently by applying relevant statistical analytical techniques including statistical tests, inferences and regression analysis with a focus on the types of data used by chemical engineers and food scientists. Thus, the aim of the course is to provide practical training to equip students with a range of skills necessary for rapidly increasing data science jobs in industries.

Course Learning Outcomes

After successfully completing this course, you should be able to:

Learning Outcome	EA Stage 1 Competencies
1. acquire and verify large data sets from various real-life applications in chemical engineering and food science.	PE1.2, PE2.2
2. describe the nature of data and distinguish between raw, incomplete, noisy and corrupted data for further analysis.	PE1.2, PE2.2
3. choose and apply appropriate statistical tools to analyse data sets across industries in chemical engineering and food science.	PE1.4, PE2.2, PE3.6
4. apply their pre-existing knowledge of food science or chemical engineering to formulate and defend models for selected data sets.	PE1.1, PE3.2

Teaching Strategies

The course consists of a 2 hr/week lecture and a 2+2 hr/week workshop. One of the 2 hr/week workshop slot will be used as office hours, where students can ask questions and discuss topics with the instructor.

The project assessments will require a significant amount of work outside of class time. It is important that students balance their time between team work, individual research, and overall planning in order to meet their assigned objectives. Lecture materials will identify and discuss central topics in the subject but students must identify, find, and study supplementary information. If work is carried out in teams, active participation is required of all team members.

Lectures: Lecture material will be delivered via pre-recorded videos on Moodle; face-to-face lectures will not be held for this course.. The lecture material will be available on Moodle as sets of video lessons. You will be able to watch these videos at any time, not just at the time scheduled in the timetable.

Workshops: We will run workshop classes at the time scheduled in the timetable. We will not use all of the timetabled sessions as we have far more than are needed; we will instead use some of those times for detailed discussion about your progress in the major project and to offer advice and assistance.

Programming: This course involves hand-on coding exercises and projects. Python will be used as a programming language to demonstrate concepts in this course. You will be asked to use Jupyter notebooks (visit jupyter.org) for your work; you can either install Python and Jupyter on your own device or use CoCalc.com/Google Colab as a free Jupyter online host. Although Python is preferred, students are free to choose between Python, Matlab, Octave and R codes to carry out their exercises and projects.

Additional Course Information

Some requisite knowledge is assumed for this course:

- Introductory statistics and data handling from any relevant courses including MATH1231, MATH2089, MATH1041.
- Discipline specific knowledge from relevant courses, for example CEIC3000, CEIC3004, CEIC3005, FOOD2320, FOOD3801, FOOD3220.

Assessment

Assessment task	Weight	Due Date	Course Learning Outcomes Assessed
1. Group selection	N/A	Week 2	1
2. Quiz 1	15%	Tuesday, Week 4	2, 3
3. Quiz 2	25%	Week 9	2, 3
4. Project part 1	20%	Week 7	2, 3, 4
5. Project part 2	40%	Week 10	2, 3, 4

Assessment 1: Group selection

Due date: Week 2

You should find your other group member(s) for the project.

Assessment 2: Quiz 1

Due date: Tuesday, Week 4

Two quizzes will be held covering the understanding of lecture content.

Quizzes will be administered on Moodle, which must be completed individually. Evidence of collaboration will be tracked electronically and academic dishonesty penalties used.

Assessment 3: Quiz 2

Due date: Week 9

Please see above.

Assessment 4: Project part 1

Due date: Week 7

The team project is a core part of this course and is a team-based activity. The project is divided into two parts over the term and the following project break-up is to be expected.

Part 1: Data preprocessing (Report submission via Moodle + Presentation on Teams)

Part 2: Data analytics (Report submission via Moodle + Presentation on Teams)

If any collaboration problems occur, please notify your instructor immediately. Marks for this component will be moderated by team assessment of individual contributions to the team submission.

Only one report and one presentation needs to be made per team. While the reports are graded as a

group, the presentations are assessed on an individual basis. As such, the presentation should be delivered by every team member, not just a single person from the team.

Assessment 5: Project part 2

Due date: Week 10

Please see above.

Attendance Requirements

Students are strongly encouraged to attend all classes and review lecture recordings.

Course Schedule

View class timetable

Timetable

Date	Type	Content
O-Week: 23 May - 27 May	Reading	Revision and light reading of pre-requisites
Week 1: 30 May - 3 June	Online Activity	Introduction to data science
	Workshop	Introduction to Jupyter and programming in python
Week 2: 6 June - 10 June	Online Activity	Data preprocessing
	Workshop	Application of data-preprocessing techniques to data sets
Week 3: 13 June - 17 June	Online Activity	Simple and Multiple linear regression
	Workshop	Performing simple and mulitple linear regression analysis on data sets
Week 4: 20 June - 24 June	Online Activity	Additional concepts related to multiple linear regression
	Workshop	Performing multiple linear regression analysis on data analysis with emphasis on concepts such as dummy variables, test-train split etc.
Week 5: 27 June - 1 July	Online Activity	Simple and multiple logistic regression
	Workshop	Performing logistic regression analysis on data sets
Week 6: 4 July - 8 July	Online Activity	**Flexibility week** Revision and consolidation of Week 1-5; Project discussion
Week 7: 11 July - 15 July	Online Activity	K-means clustering
	Workshop	Project part 1 presentations and discussion
Week 8: 18 July - 22	Online Activity	Case studies; advanced topics

July	Workshop	Performing K-means clustering on data sets
Week 9: 25 July - 29 July	Online Activity	Case studies
	Workshop	Project team meetings and discussion
Week 10: 1 August - 5 August	Online Activity	Case studies
	Workshop	Project part 2 presentations and discussion

Resources

Recommended Resources

Course materials and assessment tasks are delivered through Moodle and students should check regularly for updates. You can also obtain assistance from the UNSW Library. One starting point for assistance is: http://www.library.unsw.edu.au/servicesfor/students.html. Students will also be required to find information to augment lectures and help with their product development projects.

Some useful references are:

1. Jacob T. Vanderplas, Python data science handbook: Essential tools for working with data. O'Reilly media Inc., 2017

2. U. Dinesh Kumar, Business Analytics: The Science of Data-Driven Decision Making, Wiley. ISBN: 9788126568772

3. Andrew Ng, Machine Learning, Coursera (https://www.coursera.org/learn/machine-learning)

4. Kevin Dunn, Process Improvement Using Data (https://learnche.org/pid/preface/index)

5. G.E.P. Box, J.S. Hunter, and W.G. Hunter, Statistics for Experimenters - Design, Innovation and Discovery, 2nd edition, Wiley. ISBN: 978-0471718130.

Course Evaluation and Development

The School of Chemical Engineering evaluates each course each time it is run through (i) myExperience Surveys, and (ii) Focus Group Meetings. As part of the myExperience process, your student evaluations on various aspects of the course are graded; the Course Coordinator prepares a summary report for the Head of School. Any problem areas are identified for remedial action, and ideas for making improvements to the course are noted for action the next time that the course is run. Focus Group Meetings are conducted each term. Student comments on each course are collected and disseminated to the Lecturers concerned, noting any points which can help improve the course.

All of the activities in this course from the online lessons through to the team project have been designed in response to student feedback.

Submission of Assessment Tasks

In the School of Chemical Engineering, all written work will be submitted for assessment via Moodle unless otherwise specified. Attaching cover sheets to uploaded work is generally not required; when you submit work through Moodle for assessment you are agreeing to uphold the Student Code.

Some assessments will require you to complete the work online and it may be difficult for the course coordinator to intervene in the system after the due date. You should ensure that you are familiar with assessment systems well before the due date. If you do this, you will have time to get assistance before the assessment closes.

All submissions are expected to be neat and clearly set out. Your results are the pinnacle of all your hard work and should be treated with due respect. Presenting results clearly gives the marker the best chance of understanding your method; even if the numerical results are incorrect.

Marking guidelines for assignment submissions will be provided at the same time as assignment details to assist with meeting assessable requirements. Submissions will be marked according to the marking guidelines provided.

Late penalties

Unless otherwise specified, submissions received after the due date and time will be penalised at a rate of 5% per day or part thereof (including weekends). For some activities including Moodle quizzes and Team Evaluation surveys, extensions and late submissions are not possible.

Special consideration

If you have experienced an illness or misadventure beyond your control that will interfere with your assessment performance, you are eligible to apply for Special Consideration prior to submitting an assessment or sitting an exam.

UNSW has a Fit to Sit / Submit rule, which means that if you attempt an exam or submit a piece of assessment, you are declaring yourself fit enough to do so and cannot later apply for Special Consideration.

For details of applying for Special Consideration and conditions for the award of supplementary assessment, please see the information on UNSW's Special Consideration page.

Please note that students will need to provide some documentary evidence to support absences from any assessments missed because of COVID-19 public health measures such as isolation. UNSW will **not** be insisting on medical certificates for COVID-related absences of 7 days or less, with the positive PCR or RAT result being sufficient. Longer absences due to self-isolation or COVID-related illness will still need documentation such as a medical certificate.

Applications for special consideration **will still be required** for assessment and participation absences related to COVID-19. Special consideration requests should not be lodged for missing classes if there are no assessment activities in that class.

Academic Honesty and Plagiarism

Academic integrity is fundamental to success at university. Academic integrity can be defined as a commitment to six fundamental values in academic pursuits: honesty, trust, fairness, respect, responsibility and courage (International Center for Academic Integrity, 'The Fundamental Values of Academic Integrity', T. Fishman (ed), Clemson University, 2013). At UNSW, this means that your work must be your own, and others' ideas should be appropriately acknowledged. If you don't follow these rules, plagiarism may be detected in your work.

Further information about academic integrity and plagiarism can be located at:

- The Current Students site
- The ELISE training site

The Conduct and Integrity Unit provides further resources to assist you to understand your conduct obligations as a student: https://student.unsw.edu.au/conduct.

Referencing is a way of acknowledging the sources of information that you use to research your assignments. You need to provide a reference whenever you draw on someone else's words, ideas or research. Not referencing other people's work can constitute plagiarism. Further information about referencing styles can be located at https://student.unsw.edu.au/referencing.

For assessments in the School of Chemical Engineering, we recommend the use of referencing software such as Mendeley or EndNote for managing references and citations. Unless required otherwise specified (i.e. in the assignment instructions) students in the School of Chemical Engineering should use either the APA 7th edition, or the American Chemical Society (ACS) referencing style as canonical author-date and numbered styles respectively.

Academic Information

To help you plan your degree, assistance is available from academic advisors in The Nucleus and also in the School of Chemical Engineering.

Additional support for students

- Current Student Gateway
- Engineering Current Student Resources
- Student Support and Success
- Academic Skills
- Student Wellbeing, Health and Safety
- Equitable Learning Services
- IT Service Centre

Course workload

Course workload is calculated using the Units-Of-Credit (UOC). The normal workload expectation for one UOC is approximately 25 hours per term. This includes class contact hours, private study, other learning activities, preparation and time spent on all assessable work.

Most coursework courses at UNSW are 6 UOC and involve an estimated 150 hours to complete, for both regular and intensive terms. Each course includes a prescribed number of hours per week (h/w) of scheduled face-to-face and/or online contact. Any additional time beyond the prescribed contact hours should be spent in making sure that you understand the lecture material, completing the set assignments, further reading, and revising for any examinations.

On-campus class attendance

Physical distancing recommendations must be followed for all face-to-face classes. To ensure this, only students enrolled in those classes will be allowed in the room. Class rosters will be attached to corresponding rooms and circulated among lab demonstrators and tutors. No over-enrolment is allowed in face-to-face class. Students enrolled in online classes can swap their enrolment from online to a **limited** number of on-campus classes by Sunday, Week 1.

In certain classroom and laboratory situations where physical distancing cannot be maintained or the staff running the session believe that it will not be maintained, face masks will be designated by the course coordinator as **mandatory PPE** for students and staff. Students are required to bring and use their own face mask. Mask can be purchased from IGA Supermarket (Map B8, Lower Campus), campus pharmacy (Map F14, Middle Campus), the post office (Map F22, Upper Campus) and a vending machine in the foyer of the Biological Sciences Building (Map E26, Upper Campus).

Your health and the health of those in your class is critically important. You must stay at home if you are sick or have been advised to self-isolate by NSW health or government authorities. Current alerts and a list of hotspots can be found here. Do not come to campus if you have any of the following symptoms: fever (37.5 °C or higher), cough, sore throat, shortness of breath (difficulty breathing), runny nose, loss of taste, or loss of smell. If you need to have a COVID-19 test, you must not come to campus and remain in self-isolation until you receive the results of your test.

You will not be penalised for missing a face-to-face activity due to illness or a requirement to self-

isolate. We will work with you to ensure continuity of learning during your isolation and have plans in place for you to catch up on any content or learning activities you may miss. Where this might not be possible, an application for fee remission may be discussed. Further information is available on any course Moodle or Teams site.
For more information, please refer to the FAQs: https://www.covid-19.unsw.edu.au/safe-return-campus-faqs

Note: This course outline sets out description of classes at the date the Course Outline is published. The nature of classes may change during the Term after the Course Outline is published. Moodle should be consulted for the up to date class descriptions. If there is any inconsistency in the description of activities between the University timetable and the Course Outline (as updated in Moodle), the description in the Course Outline/Moodle applies.

Image Credit

Dr Peter Wich

CRICOS

CRICOS Provider Code: 00098G

Acknowledgement of Country

We acknowledge the Bedegal people who are the traditional custodians of the lands on which UNSW Kensington campus is located.

Appendix: Engineers Australia (EA) Professional Engineer Competency Standard

Program Intended Learning Outcomes	
Knowledge and skill base	
PE1.1 Comprehensive, theory based understanding of the underpinning natural and physical sciences and the engineering fundamentals applicable to the engineering discipline	✓
PE1.2 Conceptual understanding of the mathematics, numerical analysis, statistics, and computer and information sciences which underpin the engineering discipline	✓
PE1.3 In-depth understanding of specialist bodies of knowledge within the engineering discipline	
PE1.4 Discernment of knowledge development and research directions within the engineering discipline	✓
PE1.5 Knowledge of engineering design practice and contextual factors impacting the engineering discipline	
PE1.6 Understanding of the scope, principles, norms, accountabilities and bounds of sustainable engineering practice in the specific discipline	
Engineering application ability	
PE2.1 Application of established engineering methods to complex engineering problem solving	
PE2.2 Fluent application of engineering techniques, tools and resources	✓
PE2.3 Application of systematic engineering synthesis and design processes	
PE2.4 Application of systematic approaches to the conduct and management of engineering projects	
Professional and personal attributes	
PE3.1 Ethical conduct and professional accountability	
PE3.2 Effective oral and written communication in professional and lay domains	✓
PE3.3 Creative, innovative and pro-active demeanour	
PE3.4 Professional use and management of information	
PE3.5 Orderly management of self, and professional conduct	
PE3.6 Effective team membership and team leadership	✓