# RECORD THE LECTURE



Australian National University

## REVISION

#### COMP2420/COMP6420 INTRODUCTION TO DATA MANAGEMENT, ANALYSIS AND SECURITY

#### WEEK 12 – LECTURE 2

Wednesday 25 May 2022

Priscilla Kan John School of Computing College of Engineering and Computer Science

Credit: Ramesh Sankaranarayana (previous course convenor)



### HOUSEKEEPING



### **Course items**

- Release of Midsemester results
- Appeals within 14 days through form on Wattle
- Assignment 2 in, some extensions granted.

Cheerful news! Timing competition (lab 2)

- Announcing Daniel Simpson as the winner!
- Daniel will receive a small prize
- Total of 4 submissions before solutions were released

Table 1 – Mean timing	Tutor computer timing (s)	ANU lab computer timing (s)
results for each participant		
Rank		
1	0.002	0.003
2	0.005	0.006
3	0.01	0.042
4	3.241	3.182



Learning Outcomes Upon successful completion of this course, you should be able to:	01	Demonstrate a conceptual understanding of database systems and architecture, data models and declarative query languages.
	02	Define, query and manipulate a relational database
	03	Demonstrate basic knowledge and understanding of descriptive and predictive data analysis methods, optimization and search, and knowledge representation.
	04	Formulate and extract descriptive and predictive statistics from data
	05	Analyse and interpret results from descriptive and predictive data analysis
	06	Apply their knowledge to a given problem domain and articulate potential data analysis problems
	07	Identify potential pitfalls, and social and ethical implications of data science
	08	Explain key security concepts and the use of cryptographic techniques, digital signatures and PKI in security

### **COURSE REVISION**



## Final exam

- Final exam details and sample exams are up <u>on the</u> <u>course website</u>
- Saturday 04 June 2022 starting at 12.20pm
- 5 mins early to show id, decrypt file, sign statement of originality
- 15 mins reading (can start writing)
- 180 mins writing
- Final PUSH before deadline, commits only NOT valid
- Statement of Originality needs to be submitted by deadline
- All examination rules \*strictly\* enforced, no excuses accepted unless extremely exceptional circumstances



# Examinable material

- The examinable material comprises the all content of the course, e.g. lectures, labs, previous assessment, preparation material from Roger Clarke. (The only exception is recorded guest lecture by Prof Stephen Gould).
- There will be a mix of theoretical and practical questions, on a rough 50-50 split. There will be around 5 main questions in the exam.



### Exam environment

The exam environment will be similar to the one that you had for the mid-semester exam. It will be open book and gitlab based. You will have three hours and fifteen minutes in which to do the exam.

You will need to record a self-invigilation video to submit, including a screen recording, sound and live web cam capture, same as you did for the midsem exam. You need to show your ID (student card, passport, driving licence) at the beginning of the exam. Failure to do so is a breach of exam conditions. This will be very strictly enforced.



# Exam advance set-up

- We are investigating use of gitlab exam server for the final exam
- More details to be communicated via Piazza and Wattle
- You need to ensure that you have tested that you can decrypt the sample exam provided on the course website, as well as the 2021 exam provided in the lab repo.

### Exam submission

- A complete exam submission consists of both your completed Jupyter notebook AND your filled-in Statement of Originality. Your submission is not considered valid otherwise. No late submission of the Statement of Originality will be permitted (this will be very strictly enforced, no excuses accepted).
- You need to PUSH your complete exam submission \*before\* the deadline of the exam. Commits are NOT sufficient for time considerations. This will also be very strictly enforced.
- Save, Commit and PUSH regularly throughout the exam, especially after completing each question.
- Check your answers are being saved (remember raw text won't display on a browser)



### Exam selfinvigilation

- You need to submit your self-invigilation on wattle video 3 days after the exam. (ie by Tuesday 07 June 12.20pm)
- If you could not record it for some reason, you can submit justification through wattle. You will likely be called for an oral exam if this happens.
- The deadline for self-invigilation video will be strictly enforced. No submission means a breach of examination conditions.



# Enforcing exam rules

- The final exam is a summative assessment
- It tests your knowledge and skills with respect to the learning outcomes but also management of time and efficiency under given conditions.
- For fairness to everyone, examination rules will be very strictly enforced and no excuses accepted. You need to be prepared and know the rules and conditions and work within those.







### Content

Week 1: Data Science

Week 2: Visualisation and Data Analysis Week 3: Machine Learning, Prediction Week 4: Classification, Linear Classification Week 5: Decision Trees, Clustering Week 6: Ethics

LO 3-7

LO3. Demonstrate basic knowledge and understanding of descriptive and predictive data analysis methods, optimization and search, and knowledge representation.
LO4. Formulate and extract descriptive and predictive statistics from data
LO5. Analyse and interpret results from descriptive and predictive data analysis
LO6. Apply their knowledge to a given problem domain and articulate potential data analysis problems

LO7. Identify potential pitfalls, and social and ethical implications of data science

### Programming

- As covered in the lectures, labs
- Only application of algorithms using Python libraries
- No implementation of algorithms themselves
- Look at the lectures/labs/assignments/sampleexams for practice questions







### Content

Week 6: Data Types, Database Systems Week 7: Relational Model, SQL, ER Model, Normalisation

Week 12: NoSQL, XML

LO1. Demonstrate a conceptual understanding of database systems and architecture, data models and declarative query languages.

LO2. Define, query and manipulate a relational database

LO3. Demonstrate basic knowledge and understanding of descriptive and predictive data analysis methods, optimization and search, and knowledge representation.

# Programming (databases)

- SQL based programming using Python
- Awareness of xquery







### Content

Week 9: Intro, Public/Secret key crypto

Week 10: Internet Security, Data Protection and Privacy (IBM guest lectures AND Roger Clarke's lecture and notes)

Week 11: Digital Signatures, Public Key Infrastructure, Networks

LO8. Explain key security concepts and the use of cryptographic techniques, digital signatures and PKI in security

### Programming

- Hashing
- Encryption/decryption
- Digital signatures





CSSA event or Course organised event on 01 June evening The CSSA session intended for June 01 will be about the exam.

Make sure that you attend with your questions.

Tentative at the moment.

But regardless, we will hold a study event for this course if it does not go ahead.

6pm on 01 June 2022 Online details to be communicated.



#### **EXAM TAKING TIPS**



### Exam taking tips

- Time management is critical
- Skim through the questions and mentally mark out the easy/medium/hard ones for yourself
- Keep track of time and mark allocation when answering questions (e.g. don't spend 20 mins answering a 2 marks questions)
- Prepare your support material in advance just like you would for an in-person exam



### Thank you

- It's been trying times. We have all managed as well as we can.
- Hope the course delivery has gone on relatively well.
- We've had some very good tutors to help.
- feedback would be much appreciated.
- It has been a real pleasure teaching you all.

### Good luck

Good luck with your exams

All the best

Take care and stay safe



# End of lecture