

# COMP7990

## Principles and Practices of Data Analytics

Course Instructors:

Dr. ZHANG Lu, Eric

Dr. Wang King Hang, Kevin

Teaching Assistants:

GENG Yu, WANG Ningxia, XU Ke

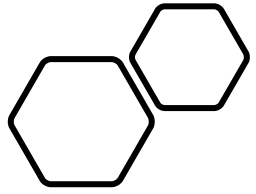
# About Me

- Dr. Zhang Lu Eric
- Office: DLB 641
- Email: [ericluzhang@comp.hkbu.edu.hk](mailto:ericluzhang@comp.hkbu.edu.hk)
- Research Interest: Deep learning in genomics, Complex disease prediction, AI in drug discovery
- Served:
  - Stanford University – Postdoctoral scholar
  - Imperial College of London-Honor research officer



# Contact Information

- Dr. ZHANG Lu, Eric 張璐
  - Office: DLB641
  - Email: [ericluzhang@comp.hkbu.edu.hk](mailto:ericluzhang@comp.hkbu.edu.hk)
- Dr. WANG King Hang, Kevin 王景行
  - Office: RRS712
  - Email: [kevinw@comp.hkbu.edu.hk](mailto:kevinw@comp.hkbu.edu.hk)
- GENG Yu 耿煜
  - Office: DLB625
  - Email: [yugeng@comp.hkbu.edu.hk](mailto:yugeng@comp.hkbu.edu.hk)
- WANG Ningxia 王寧霞
  - Office: DLB625
  - Email: [nxwang@comp.hkbu.edu.hk](mailto:nxwang@comp.hkbu.edu.hk)
- XU Ke 許可
  - Office: DLB626
  - Email: [cskexu@comp.hkbu.edu.hk](mailto:cskexu@comp.hkbu.edu.hk)



# Outline



Timetable



Course Contents



Learning  
Outcomes



Assessment  
Method

# Timetable (Section 1 and 2)

- Instructor: Dr. Eric Lu Zhang & Dr. WANG King Hang, Kevin
- Time of our classes
  - 13 weeks from Sep 10 to Dec 03
  - Time: 18:30~21:20 (**Friday**) (1st Oct - week4 - is a holiday)
  - Make-up class: 2nd Oct (**WLB103, Saturday** 18:30~21:20).  
You can also choose to join the make-up class **by zoom**.
- Quizzes:
  - Quiz1: Oct 16 (afternoon)
  - Quiz2: Nov 20 (afternoon)

[illegible]

# Course Contents

- First Half Semester
  - Instructor: Dr. ZHANG Lu, Eric (Section 1 and 2, Friday)
- Second Half Semester
  - Instructor: Dr. WANG King Hang, Kevin (Section 1 and 2, Friday)

# Course Content

## **First Half Semester:**

- Data analytics – Background mathematics; statistical analysis techniques
- Data Visualization – Concepts of data visualization; charts, maps and infographics

## **Second Half Semester:**

- Data management – Database system concepts; Relational data model; SQL
- Data security and privacy – Concepts of data security and privacy; privacy protection principles
- Data Mining – Data preparation; Data mining algorithms (classification; clustering)

# Tentative Class Schedule

| Week | Topic   | Instructor                |
|------|---|---------------------------|
| 1    | Statistic 1                                       | Dr. ZHANG Lu, Eric        |
| 2    | Statistic 2                                       |                           |
| 3    | Statistic 3                                       |                           |
| 4    | Data Visualization                                |                           |
| 5    | Data Visualization+Lab1: Statistic, Jamovi        |                           |
| 6    | Review+Lab2: Data Visualization Lab + Quiz 1      |                           |
| 7    | Data Management 1                                 | Dr. WANG King Hang, Kevin |
| 8    | Data Management 2                                 |                           |
| 9    | Security and Privacy 1                            |                           |
| 10   | Security and Privacy 2                            |                           |
| 11   | Intro to Data Mining; Data Preprocessing + Quiz 2 |                           |
| 12   | Data Mining: Classification, Clustering           |                           |
| 13   | Lab3,4: SQL, Data Mining using Weka               |                           |



# Course Aims

- This course introduces principal concepts of data management and analysis.
- It covers various topics including database management, data analytics, data mining, data visualization, and data privacy.
- It is expected that students can grasp practical skills about how to collect, store, analyze, and visualize data.

# Expected Learning Outcomes

- Knowledge
  - Describe fundamentals of database management
  - Explain concepts of data analysis techniques and data mining algorithms
  - Describe and explain concepts of data visualization
  - Describe concepts and legal foundations of data security and privacy
- Professional Skill
  - Formulate SQL queries on the database
  - Conduct statistical analysis and design visualization to present analysis results

# Assessment methods

- Continuous Assessment (40%)
  - Labs x 4 (5% each)
  - Quizzes x 2 (10% each)
- Examination (60%)
  - Final examination
- Import Notices
  - Plagiarism: Students who plagiarized and who were plagiarized will be given zero mark.
  - Final Exam: In order to pass this course, students should attain at least 30% of the final examination mark.
  - Cheating in exam: **Students who cheated in the exam/quizzes may receive a failure grade of the course and may defer their study for one year.**
  - A cumulative GPA of at least 2.50 for graduation

# Support

|                    |  |
|--------------------|--|
| <b><u>Post</u></b> | <b>Post your question on Piazza: <a href="https://piazza.com/class/kszcs0ir32u1n8">https://piazza.com/class/kszcs0ir32u1n8</a></b> |
| Email              | Email your instructors or TA   |
| Video              | Zoom lecture video will be posted online   |
| Appointment        | Make appointment for individual consultation with instructor and TA  |

# Student helper (with remuneration) and MSc Research Project 1

- 1. Research Assistant
  - Develop machine learning models for disease prediction by integrating multi-source data (e.g. medical history, genomics, digital device, et al.)
  - Have the opportunity to access ~500k individuals from UK Biobank
- 2. Teaching Assistant
  - I need several students who are familiar with python and data analytics to help prepare some material for an undergraduate course