Industrial Data Systems

Course Description: The course will provide an introduction to Python and databases in industry practice. Four aspects will be included: (a) Basic Python programming knowledge (b) Practical data analysis techniques with Python and Pandas (c) Database design principles; (d) SQL and Python interfaced database manipulations.

Prerequisites: Basic knowledge in programming language is recommended but not required

Schedule:

Lecture: MoWe 12:00PM - 01:20PM, Rm 4620, Lift 31-32 (126)

Lab 1: We 09:00AM - 10:50AM Rm 3207, Lift 21

Lab 2: Fr 09:30AM - 11:20AM

Instructor:

CHEN Yi. Email: yichen@ust.hk, Office: Room 5559C

TAs:

- Tianzhuang Xu. Email: txuay@connect.ust.hk,

- HUANG Zhenhong. Email: zhuangdr@connect.ust.hk,

- YANG Rushuai. Email: ryangap@connect.ust.hk,

- OUYANG Boer. Email: bouyangaa@connect.ust.hk,

Course webpage: https://canvas.ust.hk

Lecture/lab notes, homework assignments, solutions, and other announcements will be posted on this site. Please check this site regularly.

Course Learning Outcomes: This course focuses on practical programming techniques in data analysis and relational database system. You will learn how to use Python, Pandas, and SQL to do data manipulation and analysis in this course. You will also learn the fundamental principle and basic theory of database system design. This course is heavy in coding. Upon the successful completion, you will learn:

- Basic Python programming techniques

- Use Pandas to do data analysis and visualization

- Principle of relational database, including: Entity-Relationship models, functional dependencies and normalization, relational algebra, etc.

- Write and use SQL to manipulate database

- Use Python to manipulate database and conduct advanced data analysis.

- Advanced topics, e.g., non-relational database system, differential privacy.

Textbook: *Fundamentals of Database Systems.* R. Elmasri and S. B. Navathe. Addison-Wesley; 6 edition. 2010. (I will provide comprehensive and systematic class.You do not need to purchase the book.)

Grading: homework (20%), midterm (30% each), and final (50%).

- Exams:

    - The midterm is scheduled in the mid of semester (in lecture). The specific date will be announced later.

    - There will be a final test as scheduled by the university.

    - The exams are entirely closed. Memory and proficiency are what to be tested. Thus, no smart devices, no books, no notes, **limited cheat sheets will be allowed**.

- Homework:

    - Two types: coding homework and hand-writing homework.

    - Late submission can be graded but can earn at most the HALF points.

Course Code

- Be responsible to yourself: skip the class and copying homework at your own cost! Cheating in exams at your own risk!

- Respect others: Do not be late, no cell-phone and computer noise, no talking in class.

- Checking on grading must be done within one week after handing out the graded homework or exam.

- Make-up exams will only be given if you have a compelling reason, such as being hospitalized. This requires an official written document.