GENERAL INFORMATION

Instructor: Zigan Wang  
Email: wangzg@hku.hk  
Office: KK902  
Phone: 2859 1045  
Consultation times: By appointment

Tutor: Frankie Ho  
Email: cfhoad@hku.hk  
Office: KK1026  
Consultation times: TBC

Semester: 1  
Lecture: Monday 14:30-17:20, KKLG105

Pre-requisites: ACCT1101 Introduction to financial accounting; or ECON1210 Introductory microeconomics; or FINA1310 Corporate finance  
Co-requisites:  
Mutually exclusive:

Course Website: MOODLE via HKU portal  
Other important details: Please bring your laptop to the lectures and tutorial classes. You must install Anaconda 3 distribution and possibly other packages on your laptop.

COURSE DESCRIPTION

This course provides students a foundation in managing and analyzing financial datasets as well as other datasets. The first part of the course focuses on building skills – data manipulation using Python. The second part introduces various financial databases. Through practice on real-world financial datasets, students will learn methods used to warehouse and retrieve data for statistical computing. The course then turns to analytical methods with a focus on demonstrating these methods on real-data from various contexts in finance. Methods covered include statistical modeling and inference, introduction of machine learning, textual analysis, classification and alternative datasets. Practice on projects will be the primary mode of learning.

COURSE OBJECTIVES

1. Develop skills in database design, management, and processing.  
2. Gain proficiency in programming and performing basic data cleaning, custodianship and data manipulation.  
3. Gain a working understanding of different analytical methods used in finance and where the methods would be appropriate.  
4. Gain fluency for at least one analytical method of the student’s choosing through course projects.

COURSE LEARNING OUTCOMES

<table>
<thead>
<tr>
<th>Course Learning Outcomes</th>
<th>Aligned Programme Learning Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLO1 – Students will learn to store and access data efficiently using modern database storage methods.</td>
<td>CLO 1 – PLO 1,2</td>
</tr>
<tr>
<td>CLO2 – Students will gain an overview of analytical methods used in finance and their typical application and demonstrate understanding of how to apply the methods through highly-supervised programming assignments.</td>
<td>CLO 2 – PLO 1,2,3,5</td>
</tr>
<tr>
<td></td>
<td>CLO 3 – PLO 1,2,3</td>
</tr>
<tr>
<td></td>
<td>CLO 4 – PLO3</td>
</tr>
<tr>
<td></td>
<td>CLO 5 – PLO5</td>
</tr>
</tbody>
</table>
CLO3 – Students will demonstrate strong fluency in one analytical method of their own choice through course projects.

CLO4 – Students will be encouraged to creatively apply methods or data to solve specific industry problems.

CLO5 – Students will be encouraged to communicate ideas.

COURSE TEACHING AND LEARNING ACTIVITIES

<table>
<thead>
<tr>
<th>Course Teaching and Learning Activities</th>
<th>Expected contact hour</th>
<th>Study Load (% of study)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T&amp;L1. - Lectures</td>
<td>36</td>
<td>25%</td>
</tr>
<tr>
<td>T&amp;L2. – Problem sets.</td>
<td>36</td>
<td>65%</td>
</tr>
<tr>
<td>T&amp;L3. - Tutorials</td>
<td>10</td>
<td>10%</td>
</tr>
<tr>
<td>Total</td>
<td>82</td>
<td>100%</td>
</tr>
</tbody>
</table>

Assessment Methods

<table>
<thead>
<tr>
<th>Assessment Methods</th>
<th>Brief Description (Optional)</th>
<th>Weight</th>
<th>Aligned Course Learning Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1. 3 Projects on introductory Python (Codecademy), data transformation and Gephi (data visualization)</td>
<td>Students will have multiple projects involving programming problems. Each step will enforce a particular skill. Students will be then given a choice of an open-ended problem using example data and be asked to solve the problem. Some projects are individual-based and some allow teamwork.</td>
<td>35%</td>
<td>CLO 1-5</td>
</tr>
<tr>
<td>A2. Project on Web scraping</td>
<td></td>
<td>20%</td>
<td>CLO 1-5</td>
</tr>
<tr>
<td>A3. Presentation on Python code to test understanding</td>
<td></td>
<td>20%</td>
<td>CLO 1-5</td>
</tr>
<tr>
<td>A4. In-class test</td>
<td></td>
<td>20%</td>
<td>CLO 1-5</td>
</tr>
<tr>
<td>A5. General engagement and participation</td>
<td></td>
<td>5%</td>
<td>CLO 1-5</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

STANDARDS FOR ASSESSMENT

Course Grade Descriptors

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+, A, A-</td>
<td>The student would can competitively apply analytical methods used in the course independently.</td>
</tr>
<tr>
<td>B+, B, B-</td>
<td>The student appears to be able to apply analytical methods, but requires guidance.</td>
</tr>
<tr>
<td>C+, C, C-</td>
<td>The student has a conceptual understanding of methods applied in the course, but could not be expected to apply all methods used in the course.</td>
</tr>
<tr>
<td>D+, D</td>
<td>The student has shown effort, but a limited understanding of course content.</td>
</tr>
<tr>
<td>F</td>
<td>The student has not demonstrated effort to understand course content.</td>
</tr>
</tbody>
</table>

Assessment Rubrics for Each Assessment (Please provide us the details in a separate file if the space here is not enough)

For projects, the main driver of assessment will be accuracy with respect to the answers on which the assignments are based. An “A” quality course assignment looks professional and any discrepancies can be explained carefully. A “B” course assignment is mostly correct, but gaps in understanding remain. A “C” course assignment shows obvious gaps in understanding.
For the final course project, assessment will be based on quality of execution and originality of the investment idea. An A course project will demonstrate thorough understanding of course methods, careful consideration of pitfalls to analysis, and some element of originality. The work will be well communicated and easy to understand.

### COURSE CONTENT AND TENTATIVE TEACHING SCHEDULE

I assume a twelve-week course schedule. I am going to provide assignments according to the course policy described in the section on **Course Policy**.

I also plan to have perhaps one or two course speakers. I know some Hong Kong-area financial professionals and I will see about their availability – attendance and a short writeup will be equivalent to an assignment.

Programming will be kept on Python.

1. **Topic 1: Introduction to Python**
   - Lecture 1: Introduction of the course, Python distributions, IDE and I/O
   - Lecture 2: Immutable and mutable data structures: string, list, dictionary, tuple, set
   - Lecture 3: Loop, exception handling
   - Lecture 4: Functions: lambda, apply, filter, reduce, map

2. **Topic 2: Introduction to numPy and Pandas**
   - Lecture 5: Introduction of packages: numpy
   - Lecture 5-6: Pandas and dataframe

3. **Topic 3: Applications of Python programming**
   - Lecture 7: OOP introduction, class, geolocation, multiprocessing
   - Lecture 8: Web scraping (beautifulsoup + requests, selenium)

4. **Topic 4: Introduction to Databases**
   - Lecture 9-10: Introduction to WRDS financial database, Compustat, CUSIP/GVKEY

5. **Topic 5: Assessment**
   - Lecture 10: In-class test
   - Lecture 11-12: Presentation on Python code
   - Deadline of all projects during revision week

### REQUIRED/RECOMMENDED READINGS & ONLINE MATERIALS (e.g. journals, textbooks, website addresses etc.)

Coursera, Codecademy

### MEANS/PROCESSES FOR STUDENT FEEDBACK ON COURSE

- conducting mid-term survey in additional to SETL around the end of the semester
- Online response via Moodle site
- Others: ________________________ (please specify)

### COURSE POLICY (e.g. plagiarism, academic honesty, attendance, etc.)

The code of ethics and attendance policy will be applied.

### ADDITIONAL COURSE INFORMATION (e.g. e-learning platforms & materials, penalty for late assignments, etc.)

I may prescribe self-learning materials to supplement students’ learning. Most of this will be optional. Late assignment policy will be to accommodate students who provide reasonable notice about competing obligations. Given that some assignments will be optional, students may also simply do other assignments given later in the course. I want to encourage self-motivated students to outperform, but do not want to discourage students excessively for a new module.