Course Code
Big Data Analysis
Module 2, 2021-22

Course Information

Instructor: *郑海洋*
Office: 649
Phone: 0755-2603-4026
Email: hyzheng@phbs.pku.edu.cn
Office Hour:

Teaching Assistant:
Phone:
Email:

Classes:
Lectures:
Venue:

Course Website:
TBA

1. Course Description

1.1 Context

Course overview:
This course introduces the history and background of big data analysis, and teaches some commonly and widely used big data analysis techniques and their applications in various fields. We will cover data processing, data storage, visualization, typical pattern recognition and machine learning methods, and forecasting.

Prerequisites:

Fair Knowledge of programming and statistics

1.2 Textbooks and Reading Materials

- No official textbooks, we use the following two books as References.
- Your Complete Guide to Factor-Based Investing: The Way Smart Money Invests Today, 2016, by Berklin and Swedroe (**CGFBI**)
- Pattern Recognition and Machine Learning, by Bishop (**PRML**)
https://www.microsoft.com/en-us/research/publication/pattern-recognition-machine-learning/

2. Learning Outcomes

2.1 Intended Learning Outcomes

Learning Goals	Objectives	Assessment (YES with details or NO)
1. Our graduates will be effective communicators.	1.1. Our students will produce quality business and research-oriented documents.	YES
	1.2. Students are able to professionally present their ideas and also logically explain and defend their argument.	YES
2. Our graduates will be skilled in team work and leadership.	2.1. Students will be able to lead and participate in group for projects, discussion, and presentation.	YES
	2.2. Students will be able to apply leadership theories and related skills.	
3. Our graduates will be trained in ethics.	3.1. In a case setting, students will use appropriate techniques to analyze business problems and identify the ethical aspects, provide a solution and defend it.	
	3.2. Our students will practice ethics in the duration of the program.	
4. Our graduates will have a global perspective.	4.1. Students will have an international exposure.	YES
5. Our graduates will be skilled in problem-solving and critical thinking.	5.1. Our students will have a good understanding of fundamental theories in their fields.	YES
	5.2. Our students will be prepared to face problems in various business settings and find solutions.	YES
	5.3. Our students will demonstrate competency in critical thinking.	YES

2.2 Course specific objectives

See section 1.1 Context.

2.3 Assessment/Grading Details

Attendance 5%, Assignments 20%, Exams 35%, Final Project 40%

The level of background knowledge may vary among students, but it will be ignored in grading. Grading will be strictly based on outcome, not on effort or progress.

2.4 Academic Honesty and Plagiarism

It is important for a student's effort and credit to be recognized through class assessment. Credits earned for a student work due to efforts done by others are clearly unfair. Deliberate dishonesty is considered academic misconducts, which include plagiarism; cheating on assignments or examinations; engaging in unauthorized collaboration on academic work; taking, acquiring, or using test materials without faculty permission; submitting false or incomplete records of academic achievement; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement; or altering, forging, or misusing a University academic record; or fabricating or falsifying of data, research procedures, or data analysis.

All assessments are subject to academic misconduct check. Misconduct check may include reproducing the assessment, providing a copy to another member of faculty, and/or communicate a copy of this assignment to the PHBS Discipline Committee. A suspected plagiarized document/assignment submitted to a plagiarism checking service may be kept in its database for future reference purpose.

Where violation is suspected, penalties will be implemented. The penalties for academic misconduct may include: deduction of honour points, a mark of zero on the assessment, a fail grade for the whole course, and reference of the matter to the Peking University Registrar.

For more information of plagiarism, please refer to *PHBS Student Handbook*.

3. Topics, Teaching and Assessment Schedule (Tentative)

Week	Dates	Topics
1		An overview of big data analysis, history and development
2		Data processing and databases
3		Hadoop and MapReduce
4		Regressions Techniques for Big Data
5		Logistic Regressions and Machine learning techniques
6		Midterm exam and final project proposals
7		Factors Analysis
8		Factors applications on high frequency data
9		Guest lectures about the latest industry trends

4. Miscellaneous