

**THE UNIVERSITY OF HONG KONG
FACULTY OF BUSINESS AND ECONOMICS**

ECON 3225: Big Data Economics

GENERAL INFORMATION	
<p>Instructor: Dr. Yanhui WU and Dr. Ping YU</p> <p>Email: yanhuiwu@hku.hk; pingyu@hku.hk Office: KKL 931; KKL 1108 Phone: Consultation times: TBD</p> <p>Tutor:</p> <p>Pre-requisite(s): Introductory econometrics (ECON2280) or causal inference (ECON3284) Co-requisite(s): Mutually exclusive:</p> <p>Course website: Other important details:</p>	
COURSE DESCRIPTION	
<p>This course introduces students to fundamental ideas, important methods and popular techniques in big data analysis and machine learning. Combining statistical theory, computational tools, and hands-on experience with real data, this course will provide students with a solid basis for handling big data in economics, finance, and management.</p> <p>The primary focus of this course is on application instead of methodological rigor. Hence, the use of mathematics will be limited to an elementary level. However, students are expected to have a strong background in statistics and/or econometrics. In particular, the course will assume that participants have an understanding of statistical inference using t-tests and have prior experience of interpreting the results of multiple linear regression. We will review these topics briefly during the course.</p> <p>Because of the emphasis on hands-on experience, students are expected not to be scared by data and coding. Previous experience with statistical software and knowledge about computer programming is an advantage but not required. Homework assignments are designed to familiarize students with the necessary programming language. For programming, we will use the statistical package R via a front-end called RStudio. Both R and RStudio are free and open source.</p>	
COURSE OBJECTIVES	
<ol style="list-style-type: none"> 1. Provide students a broad overview of the most popular data science methods in economics and business studies 2. Enhance students' analytical ability to apply appropriate methods in different contexts 3. Equip students with a basic toolkit that can be directly used for their own research 	
FACULTY LEARNING GOALS (FLGs)	
<p>FLG1: Acquisition and internalization of knowledge of the programme discipline FLG2: Application and integration of knowledge FLG3: Inculcating professionalism FLG4: Developing global outlook FLG5: Mastering communication skills FLG6: Cultivating leadership</p>	
COURSE LEARNING OUTCOMES (CLOs)	
Course Learning Outcomes	Aligned Faculty Learning Goals (FLGs)
CLO1. Gain a solid understanding of the principles of applying data science to social sciences	FLG 1, FLG 2

CLO2. Demonstrate a solid grounding in recent developments in big data methods, including state-of-the-art machine learning techniques and their suitability to solve important economic, finance, and business problems.	FLG 1, FLG 2, FLG 5
CLO3. Demonstrate ability to address questions of interest by using applied data science and econometric techniques.	FLG 1, FLG 2, GLG 4, FLG 5
CLO4. Demonstrate facility with implementing the techniques covered in the course using statistical software on real-world datasets	FLG 1, FLG 2, FLG 3

COURSE TEACHING AND LEARNING ACTIVITIES

Course Teaching and Learning Activities	Expected Study Hours	Study Load (% of study)
T&L1. Lectures	36	30
T&L2. Weekly problem sets	36	30
T&L3. Computer programming	24	20
T&L4. Tutorial/self-learning sessions	24	20
Total	120	100

Assessment Methods	Brief Description (Optional)	Weight	Aligned Course Learning Outcomes
A1. Problem sets		50	CLO1, CLO2, CLO3, CLO4
A2. Class participation		10	CLO1, CLO2, CLO3, CLO4
A3. Final exam		40	CLO1, CLO2, CLO3, CLO4
Total		100%	

STANDARDS FOR ASSESSMENT

Course Grade Descriptors	
A	Strong evidence of superb ability to fulfill the intended learning outcomes of the course at all levels of learning: describe, apply, implement, evaluate and synthesis.
B	Strong evidence of ability to fulfill the intended learning outcomes of the course at all levels of learning: describe, apply, implement, evaluate and synthesis.
C	Evidence of adequate ability to fulfill the intended learning outcomes of the course at low levels of learning; such as describe and apply, but not at high levels of learning such as evaluate and synthesis.
D	Evidence of basic familiarity with the subject.
F	Little evidence of basic familiarity with the subject.

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

Please see the attached sheet
COURSE CONTENT AND TENTATIVE TEACHING SCHEDULE
<p>Week 1. Overview of big data economics Weeks 2. Prediction Weeks 3-4. Classification Weeks 5-6. Regularization Weeks 7-8. Support Vector Machine Week 9. Decision Tree Week 10. Clustering Weeks 11-12. Deep learning</p>
REQUIRED/RECOMMENDED READINGS & ONLINE MATERIALS (e.g. journals, textbooks, website addresses etc.)
<p>Required: (ISLR) <i>An introduction to Statistical Learning</i>. James, Witten, Hastie, and Tibshirani. Springer. 2013. (The digital version of this book is freely available on http://faculty.marshall.usc.edu/gareth-james/ISL/)</p> <p>Useful references:</p> <p>Data Science for Business: <i>What you need to know about data mining and data-analytic thinking</i>. Provost and Fawcett. O'Reilly. 2013.</p> <p>Business Data Science: <i>combining machine learning and economics to optimize, automate, and accelerate business decisions</i>. Taddy. McGraw Hill. 2019.</p>
MEANS/PROCESSES FOR STUDENT FEEDBACK ON COURSE
<p><input checked="" type="checkbox"/> conducting mid-term survey in addition to SETL around the end of the semester</p> <p><input checked="" type="checkbox"/> Online response via Moodle site</p> <p><input type="checkbox"/> Others: _____ (please specify)</p>
COURSE POLICY (e.g. plagiarism, academic honesty, attendance, etc.)
<p>1. This is an active learning course, and attendance and participation are extremely important. Please observe appropriate classroom etiquette and be considerate to others. In particular, laptop use should be limited to course-related activities, and cell phones are not allowed in class.</p> <p>2. Students are encouraged to work together in groups to solve the problem sets. However, each student must turn in his or her own homework. Copying another student's answers is not permitted even with consent. All assignments including the term project report must be typewritten.</p> <p>3. Plagiarism and cheating in exams are serious academic offenses.</p>
ADDITIONAL COURSE INFORMATION (e.g. e-learning platforms & materials, penalty for late assignments, etc.)

Assessment Rubrics for Each Assessment

	Grade A	Grade B	Grade C	Grade D	Grade F
Problem sets (50%)	Demonstrate strong evidence of mastering the methods and techniques to solve problems; turn in all problem sets with clear answers in due course	Demonstrate reasonable evidence of mastering the methods and techniques to solve problems; turn in all problem sets with clear answers in due course	Demonstrate some evidence of the capability of applying the methods and techniques to solve problems; turn in most problem sets in due course	Demonstrate limited evidence of the capability of applying the methods and techniques to solve problems; turn in most problem sets in due course	Demonstrate little evidence of the capability of applying the methods and techniques to solve problems; Fail to turn in most problem sets in due course
Class participation (10%)	Attend class regularly; active engagement in class discussion	Attend class regularly; some engagement in class discussion	Attend class regularly; limited engagement in class discussion	Attend most of the classes; limited engagement in class discussion	Fail to attend most of the classes; little engagement in class discussion
Final Exam (40%)	A solid understanding of concepts and super ability of solving problems.	A solid understanding of concepts and good ability of solving problems.	A good understanding of concepts and reasonable ability of solving problems.	A fair understanding of concepts but limited ability of solving problems.	Limited understanding of concepts and little ability of solving problems.