

課程大綱

Program Syllabus

(English course)

系所 Department	企業管理研究所 / Graduate Institute of Business Administration	必選修 compulsory/elective	選修 elective
課程名稱 (含英文名稱) Course title	大數據程式設計(含 Python)(一) Big Data Analysis with Python (I)	全英文授課 EMI	■ Yes
學年/學期 academic year/Semester	114-1 學期 Autumn semester 2025	學分數 Credit(s)	3
講授教師 Instructor	宋豪漳 Hao-Chang Sung	上課地點 Classroom	TBA
教師辦公室&諮詢時間 Instructor office number & office hour	hcsung@ccu.edu.tw 管院: 421 室	上課時間 Time	Thursday 13:10~16:00
助教 Teaching Assistant	洪晨芮 chanrui93@gmail.com 陳委岑 show05732@gmail.com	教師聯絡資訊 Instructor Contact	Email: hcsung@ccu.edu.tw
先修課程 Pre-requisite courses	統計學、微積分、基礎編程能力 Statistics, Calculus, Basic Programming Skills		
課程目標 Course Objective	This course consists of three parts: - This course will focus on using Python for data analysis: - Introducing essential Python libraries: Numpy, Pandas, Scikit-learn, and Matplotlib. - Introducing manipulating, processing, cleaning, and crunching data in Pandas. We will also discuss using Pandas for time series data analysis, such as financial data. - We will introduce using Scikit-learn and TensorFlow for supervised learning and unsupervised learning. - supervised learning: linear regression, logistic regression, decision trees, unsupervised learning: K-means. - We will introduce the concept of data in the real world that arrives incrementally		

	and real-time data handling (for financial data) i. vectorized backtesting ii. working with real-time data	
AACSB 學習品質保證學習目標 Assurance of Learning (AOL) Learning goals *請先選填為主要或次要學習目標(Major or minor learning goal)，再選擇對應之學習目標		
主要學習目標 Major learning goal 目標 1：創新思考 LG1:Creative Thinking	主要學習目標 Major learning goal 目標 4：全球視野 LG4:Global Perspectives	次要學習目標 Minor learning goal 目標 2：溝通能力 LG2:Communication Skills
教材 Teaching materials	Teaching materials are selected chapters of texts and corresponding handouts used in class. These handouts are made available only for the personal use of the students.	
網址 Course website	TBA	
教科書/參考書 Textbooks/Reference	Textbooks (Required book): 1. Wes McKinney*, (2022). <i>Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter</i> . 3 rd edition, O'Reilly Media. (Ch.4-8; Ch.11) *: Creator of Panadas module 2. Hilpisch, Y. (2020). <i>Python for Algorithmic Trading</i> . O'Reilly Media. (https://www.books.com.tw/products/0010897570) 3. Yves, H. (2019). <i>Python for finance: mastering data-driven finance</i> . O'Reilly Media.(https://www.books.com.tw/products/F014247464) References: 1. Resources: • https://github.com/TA-Lib/ta-lib-python • https://docs.data.nasdaq.com/ 2. 陳強(編著), 2021, 機器學習及 Python 應用, 高等教育出版社(簡體書). 3. Hull, J. <i>Machine Learning in Business: An Introduction to the World of Data Science</i> . Amazon Distribution. 2020. (Ch.2-4) 「請尊重智慧財產權,不得非法影印教師指定之教科書籍」	
評量方式(請填百分比) Assessment	■課堂參與 Participation 15% ■作業 Homework 20% ■期中考 Midterm Exam25% ■期末報告 Final project 40%	
其他說明 Other description	1. The course will be offered for graduate students and undergraduates (junior and senior). 2. Teaching Approach(es):	

	Lecture: 60% Class Discussion: 10% Group Activity: 30% 3. Course Contents: i. Class Participation/Attendance (15%) Class attendance and participation are essential. Students need to send an email in advance with their excuse for absences. (In case of an emergency or illness, they are allowed to send me an email after their absence) ii. Homework Assignments (20%) Homework assignments will be provided once a week or once every two weeks. Students are required to hand it in before the deadline. iii. Open-book Midterm Exam (25%) There will be one midterm exam for this course. The exam will consist of 2 to 3 essay questions and 3 to 5 Python code problems with the help of a PC. Each exam is graded on a 100-point basis. iv. Final-term Project (40%) a. Students need to do a final data project (on a topic related to financial data) designed to engage in data analysis using Python. b. Students will form a team (1-2 classmates) and present the final projects in the last two weeks. Each group must analyze a chosen topic and present the findings. c. By using Python, each group will need to acquire and clean the data and use tools from the course to explore, describe, and analyze the data. d. To polish your project's findings, each group should evaluate the project's findings by making predictions (via the machine learning method) or trading based on automated algorithmic trading with real-time or streaming data. Each group should be confident in their knowledge of the course material by engaging in the final project.	

課程規劃表 Course Schedule

週次 week	日期 Date	內容 Description	教材章節 Textbook	其他說明 Remark
1.	9/11	1. Course Overview 2. Introduction: a. Why Python for data analysis b. Essential Python Libraries a.	Ch.1 of McKinney (2022)	
2.	9/18	NumPy Basics: Arrays and Vectorized Computation-(1)	Ch.4 of McKinney (2022)	
3.	9/25	NumPy Basics: Arrays and Vectorized Computation-(2)	Ch.4 of McKinney (2022)	
4.	10/02	Getting Started with Pandas-(1)	Ch.5 of McKinney (2022)	
5.	10/09	Getting Started with Pandas-(2)	Ch.5 of McKinney (2022)	
6.	10/16	(停課一周, 另討論補課時間)Data Loading, Storage, and File Formats	Ch.6 of McKinney (2022)	
7.	10/23	Data Cleaning and Preparation	Ch.7 of McKinney (2022)	
8.	10/30	Time Series Data-(1)	Ch.11 of McKinney (2022) & Ch.3 of Hilpisch, Y. (2020).	
9.	11/06	Midterm Exam (Open book)		
10.	11/13	Time Series Data-(2)	Ch.11 of McKinney (2022) & Ch.4 of Hilpisch, Y. (2020).	
11.	11/20	Predicting Market Movements with Machine Learning (I)	Ch. 3 of Hull (2020) & Ch.. 5 of Hilpisch, Y. (2020).	

12.	11/27	Predicting Market Movements with Machine Learning (II)	Ch. 4 of Hull (2020) & Ch . 5 of Hilpisch, Y. (2020).	
13	12/04	Backtesting	Ch.6 of Hilpisch, Y. (2020).	
14.	12/11	Working with Real-Time Data and Sockets	Ch 7 of Hilpisch, Y. (2020).	
15.	12/18	Final Project Presentation-(I)		
16.	12/25	Holiday		
17.	01/01	Holiday		
18.	01/08	Self-studying		