

# 國立政治大學課程教學大綱

## Syllabus

課程資訊	
學年學期 Academic Year / Semester	113 學年第 1 學期      Fall Semester, 2024
科目代號 Course Number	354921001
開課單位 Department Name	統計系
課程名稱 Course Name	巨量資料與統計分析 Big Data and Statistical Analysis
授課教師 Instructor	余清祥 (Ching-Syang Jack Yue)
選課人數 Maximum Number of Students	30
學分數 No. of Credits	3
修別 Required / Elective	選修(Elective)
先修科目 Prerequisite(s)	統計學
上課時間 Course Time	每週二 9：10~12：00
點閱核心能力分析圖與授課方式比例圖	
課程簡介 Course Description	
<p>This course is to introduce the concept of data analysis for big data. In addition to the concept and analysis methods, the students will learn how to analyze two kinds of big data, structured data and un-structured data.</p>	
課程目標與學習成效 Goals& Learning Outcomes	
<p>The goal is to develop the skills of a data scientist in the statistical point of view. Emphasized topics include problem definition, data analysis for hard and soft data, data cleaning, and practical aspects of big data analysis. For the first part, we will introduce some basic notions of big data, as well as problem definition. The second and third parts involve case studies of hard and soft data, in addition to the introduction of their analysis methods. Basically, the analysis of hard data includes the data mining techniques (Hastie et al., 2009). On the other hand, there is no standard operating procedure for the soft data yet and the exploratory data analysis skills are suggested.</p> <p>The students are expected to be familiar with the process and methods of big data after finishing this course. This course can be helpful to students for finding jobs in the industries and government.</p>	

## 每週課程進度與作業要求 Course Schedule &amp; Requirements

週次 Week	課程主題 Topic	課程內容與指定閱讀 Content and Reading Assignment	教學活動與作業 Teaching Activities and Homework	學習投入時間 Student workload expectation	
				課堂講授 In-class Hours	課程前後 Outside-of-class Hours
1	課程介紹、SQL 教學	教師自製投影片	教師講述	3 小時	1 小時預習
2	數據科學家	教師自製投影片	教師講述	3 小時	2 小時複習
3	資料蒐集	教師自製投影片	教師講述	3 小時	2 小時複習
4	資料蒐集；結構資料	教師自製投影片	教師講述；第一次作業	3 小時	3 小時複習
5	結構資料	教師自製投影片	教師講述	3 小時	3 小時複習
6	探索性資料分析	教師自製投影片	教師講述	3 小時	2 小時複習
7	探索性及驗證資料分析	教師自製投影片	教師講述；第二次作業	3 小時	2 小時複習
8	驗證資料分析	教師自製投影片	教師講述	3 小時	5 小時預習
9	結構資料案例討論	教師自製投影片	教師講述	3 小時	3 小時複習
10	期中報告；定義問題	參考文獻	教師講述	3 小時	3 小時複習
11	非結構資料	教師自製投影片	教師講述	3 小時	3 小時複習
12	非結構資料	教師自製投影片	教師講述	3 小時	3 小時複習
13	非結構資料案例討論	教師自製投影片	教師講述；第三次作業	3 小時	3 小時複習
14	歸納與演繹	教師自製投影片	教師講述	3 小時	3 小時複習
15	歸納與演繹	教師自製投影片	教師講述	3 小時	5 小時預習及複習
16	道德考量；報告格式	教師自製投影片	教師講述；第四次作業	3 小時	5 小時預習及複習
17	案例討論	教師自製投影片	教師講述	3 小時	3 小時複習
18	期末報告	參考文獻	分組討論	3 小時	3 小時複習

## 授課方式 Teaching Approach

講述 Lecture：60%；討論 Discussion：15%；小組活動 Group Activity：15%； 數位學習 E-learning：10%；其他 Others：0%。
<b>評量工具與策略、評分標準 Evaluation Criteria</b> <b>【明列評量項目與給分標準】</b>
There are regularly based homework, numerical projects, and one final project. The homework counts 45%, numerical projects 25%, and the final project 30%. The final project is due on 1/10/2025, and the one fails to hand in the final project will result “Fail” in the final grade.
<b>課程進行中，可否使用手機等智慧行動裝置 To Use Smart Devices During the Class</b>
<input type="checkbox"/> 是 Yes <input type="checkbox"/> 否 No <input checked="" type="checkbox"/> 需經教師同意始得使用 Approval required <input type="checkbox"/> 其他 Others:
<b>授課教師 Office Hours 及地點 Office Location</b>
Office Hours 在教師研究室（1026 室），授課地點為學校排定教室
<b>教學助理基本資料 Teaching Assistant Information</b>
教學助理在學期開始時，由申請者中挑選適合者
<b>指定／參考書目 Textbook &amp; References</b> <b>【為維護智慧財產權，請務必使用正版書籍】</b>
<ol style="list-style-type: none"> <li>1. 大數據—知識經濟與實務應用(2016)，余清祥、顏貝珊合著</li> <li>2. Problem Solving: A Statistician’s Guide (1995), by C. Chatfield</li> <li>3. The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2009), by T. Hastie, R. Tibshirani, &amp; J. Friedman</li> <li>4. Doing Data Science (2013), by R. Schutt &amp; C. O’Neil</li> <li>5. An Introduction to Statistical Learning, with Applications in R (2013), by James, Witten, Hastie, &amp; Tibshirani</li> </ol>
<b>課程相關連結 Course Related Links</b>
教師個人網站 <a href="http://csyue.nccu.edu.tw">http://csyue.nccu.edu.tw</a>
<b>課程附件 Course Attachments</b>
視課程需要，教師在網站公布相關文獻、連結網站，包括： The software R can be downloaded via <a href="http://www.r-project.org">http://www.r-project.org</a>