

與國立臺北教育大學通識課程校際選課 課程大綱

開課學校	國立臺灣大學				
開課單位/系所	共同教育中心				
授課教師姓名	蔡芸琇	職級	專案計畫助理教授	專/兼任	專任
中英文課程名稱	(中文) 資料科學程式設計 (英文) Data Science Programming				
全英語授課	☐ 是 ☒ 否				
課程領域別	基本能力課程				
開放外校修課人數	共 5 人				
學分數	3 學分				
課程概述	拜寬頻網路普及化之賜，大數據熱潮來襲，許多資料科學相關的需求因應而出，《哈佛商業評論》將會寫程式、又有專業領域知識的資料科學家，封為「21世紀最性感行業」。本課程結合 R 語言與其他電腦資訊技術，運用在各專業領域上。例如：以虛擬社群做社會行為的量化分析、以量化研究方法做人口統計學上的社會變遷、以質性研究方法針對虛擬社群做網誌調查、以線上問卷系統取代傳統紙本問卷、資料視覺化與結合地圖資訊系統等。				
課程目標	想透過這門課，和學生們一同製作出屬於台大的資料開放平台，就暫且稱呼為 NTU GOV 吧！進行資料分析不是為了要批評或鬥爭，而是想要讓我們的學校可以透過資訊技能，看到問題點在哪裡，進而得到更好的決策參考。批評很容易，放炮也很簡單。最難的是要如何在舊有的體制內做軟性的創新，讓大家都願意一同進步，期待熱血的你們加入。 以 http://gov.tw/zh-TW/ 的概念為基礎，對台大財務報告進行討論與分析 http://www.ntuacc.ntu.edu.tw/chief/chief8.jsp，或其他台大相關開放資料。				

教學內容及進度 (每週進度及教學內容簡述、課程若邀請學者專家演講，請敘明其姓名、單位、職稱及演講主題)	課程內容		
	週次		單元主題
	第 1 週		上課規則、計分規則。介紹 R 語言、上手 R 語言。作業說明。Rmarkdown 使用教學。本週課程投影片。
	第 2 週		R 語言的資料型態。本週課程投影片與範例程式。資料科學修課心得分享。https://www.datacamp.com/courses/free-introduction-to-r
	第 3 週		dplyr 與 Data Frame。本週投影片。
	第 4 週		ggplot2 繪圖教學
	第 5 週		透過 r apply family 執行文本蒐集。本週課程投影片。

		ion/d/e/2PACX-1vQj5V83I9CDw9soje3k3tx4oZBbTNHvfFWWh0UIOSdT5FQC10tNs6Qe08k08gpf7ube6_ZwAvvsQsfCX/pub?start=false&loop=false&delayms=3000>本週投影片。
第 6 週		臉書爬蟲與文字雲實作。
第 7 週		Term to Document Matrix 本周投影片
第 8 週		以 TF-IDF 進行 K-Means 分群。 本週投影片
第 9 週		1. 機器學習簡介 2. PCA 3. SVM 本週投影片
第 10 週		雙變數分析。
第 11 週		利用 R 和 Shiny 製作網頁作品。 https://featherwind.shinyapps.io/final/

	第 12 週		利用 R 和 Shiny 製作網頁作品。 https://feeatherwind.shinyapps.io/final/	
	第 13 週		https://docs.google.com/presentation/d/e/2PACX-1vQBwdoVGj1Zut4nHNiFWoLpXJ5dvmPWLP-Pof-rx1WDYJKwpKr9zJbmHsGYFezjb9RKyhSmJpEpsLI7/pub?start=false&loop=false&delayms=3000 自創 R 套件包	
	第 14 週		https://ithelp.ithome.com.tw/articles/10187683 使用 Deep Learning 於 R 中	
			https://docs.google.com/presentation/d/e/2PACX-1vTuV2TnskjNMY00_DKH3YBFoVV3LL1EVmysS4DnRiYQYd1IVzL6t0Qdz7KnMIixeCQzXM4YLfy_a9I7/pub?start=false&loop=false&delayms=3000 本週投影片	
	第 15 週		小組討論。	
	第 16 週		五次作業 Demo	
	第 17 週		期末專案 Demo	
	第 18 週		期末考停課。	
指定閱讀資料及 延伸閱讀 (請詳列每週學生應配合閱讀之篇章)	http://datascienceandr.org/ http://r4ds.had.co.nz/			
成績評量方式 (請說明各項評量項目、比例)	No.	項目	百分比	說明
	1.	期末專	40%	共 1 次，佔個人總分的 30%，於期中考後，進行分組，五個人一組，最後有參加期末發表的組別，期末報告成績為 A，於期末發表會得獎的組別，期末報告成績為 A+，其餘依照

及標準)		題		完成度斟酌給分。
	2.	個人作業	60%	共 5 次，每次佔個人總分的 12%，於作業公布後兩周內於上課前繳交，每次作業分數依照指示給分，於公布作業當天公告，遲交斟酌扣分。
課程網址	(網址：http://，資料若無則免填)			
修課注意事項 (若無則免填)	優先錄取具備以下條件者，上限 50 人，收滿為止： 1. 已安裝好 R 和 RStudio 者。 2. 已使用過 GitHub 者。 3. 能提出台大開放資料相關構想者(想做的題目、資料來源)。 作業繳交規範： https://joel88032.github.io/pecuclass1/regulation.html 課程投影片清單：https://ntu-csplusx.github.io/2017-DS-Programming/			