

教育部教學實踐研究計畫成果報告
Project Report for MOE Teaching Practice Research Program

計畫編號/Project Number：PGE1090844

學門專案分類/Division：通識(含體育)

執行期間/Funding Period：2020/08/01 ~ 2021/07/31

運動分析實務問題融入資料科學課程之行動研究

(配合課程：運動資訊管理)

計畫主持人(Principal Investigator)：

許 育 嘉

執行機構及系所(Institution/Department/Program)：

國立臺灣體育運動大學 運動資訊與傳播學系

成果報告公開日期：

立即公開 延後公開(統一於 2023 年 9 月 30 日公開)

繳交報告日期(Report Submission Date)：2021/09/07

運動分析實務問題融入資料科學課程之行動研究

壹、研究動機與目的

「資料科學」(data science)亦稱為「數據科學」,與「大數據分析」、「巨量資料分析」意義相近,是近年來興起的一門新的課程,目標是要學生能從資料中,利用一些所學得的理論與技術,包括包括應用數學、統計、圖型識別、機器學習、資料視覺化、資料倉儲以及高效能計算等,來提取出有價值的知識而加以運用,協助讓非專業人士也能理解問題,幫助正確的處理大量資料,得以進行研究調查,或應用於商業競爭上。而在運動產業中,各式各樣的數據也出現巨量的成長,從過去一直保留至今的各場次比賽紀錄資料庫,到近年來在運動科學上邁向科技化的應用,利用各式感測器(sensor)量測的生心理數據、生物力學數據、空間數據、各式運動測試數據等,這些大量數據後所隱藏的知識,需有足夠的運動領域知識並精通數據分析方法才有機會發掘。然而在主修體育之相關科系中,有關資訊科技應用的課程仍略顯不足(Barneva & Hite, 2017)。

而國內大學「資料科學」的課程,依據教育部「大學課程資源網」¹的資料顯示,最早應始於103學年東吳大學的巨量資料分析學分學程開設「資料科學與巨量資料導論」開始,從103學年全國大學(不含技職)開設課程數1門開始,隔年104學年度共開設27門,而到了107學年更出現爆炸式的成長到有38校共61個系所開設了120門課程。而這些開設課程的系所,大都是資訊相關科系,而少數非資訊科系開課的則有教育、經濟、金融、企管、心理、社會、應數、醫管、土木等系所,可見「資料科學」在各個領域都受到相當程度的重視,且隨著各領域不同的應用特性,課程內容也應隨之調整,如某些系所結合其領域知識開設「健康資料科學概論」、「教育資料科學概論」、「資料科學與社會研究」等課程。

反觀運動產業相關科系,在面臨大數據資訊時代到來之時,運動大數據分析的教學大多分散在碩士班或專題類的課程中以部份時數教授,尚未能有整學期「資料科學」完整課程的開設。但對於運動產業,資料科學並非不是不重要,而是資料科學須結合許多基礎領域的知識,對於數學、統計、程式設計、機器學習等需有一定程度的基礎,而這些基礎領域通常也不在體育相關科系的開課範圍內,因此對體育相關科系形成很高的開課門檻。

在國內體育相關科系中,結合運動產業的特性開設「資料科學」課程實已刻不容緩。而對於這些未能有足夠基礎知識主修體育的學生,一個良好課程的規劃與設計,包含教材內容、教學方法、學習活動、技術實作等,會是影響學習成效最重要的關鍵。

另一方面,「運動分析」這門學科,在Moneyball這類題材近幾年開始受到注目後,以往僅於運動愛好者間普遍討論的棒球統計學(又稱賽伯計量學、Sabermetrics)、籃球計量學(APBRMetrics)等,逐漸擴展至大學課程中,並結合運動情蒐、運動科學、運動感測器等既有運動學術領域之發展,在國外的大學中逐漸演變發展出「運動分析」課程及學位學程,如美國西北大學設有Master's in Sports Administration Specialization in Sports Analytics²,英國羅浮堡大學則設有MSc Sport Analytics and Technologies programs³。

運動分析所探討的實務問題,非常適合發展「資料科學」課程教學之情境以及用於設計課程實作,主要有下列三個理由:

- 一、現有的運動資料集十分完整:各項運動每場比賽的記錄與敘述統計資料幾乎沒有短少。尤其是棒球比賽的數據,歷年來皆維護良好,堪稱是目前世界上最豐富的運動資料庫。而這些資料通常都是免費開放的,可以輕易取得,在進行分析前不太需要

¹ 網址: <http://ucourse-tvc.yuntech.edu.tw/>

² 網址 <https://sps.northwestern.edu/masters/sports-administration/analytics-specialization.php>

³ 網址 <https://www.lboro.ac.uk/study/postgraduate/masters-degrees/a-z/sport-analytics-and-technologies/>

進行資料清理與調整，且每場比賽的規則明確，使得資料的因果關係容易確認。

二、運動分析所要探討的問題定義明確：所有的運動都遵守一項明確的規則：誰贏了？誰輸了？而大多數運動分析主要要解決的問題，就是預測出比賽輸贏。其次的問題，則是要去說明為何會造成輸贏，如探討某位運動員場上的表現，或是因為採用某種策略而導致輸贏。

三、領域層次的知識(domain-level knowledge)容易融入：運動是一項普為被接受的休閒活動，而在運動觀賞行為中，觀眾涉入的程度高且對於球隊或球員會產生忠誠度，而出現愛好者(fans)的行為，除了能對運動競賽規則有清楚的了解外，往往也會對於競賽的結果預測有一些個人的初步分析與看法。而其他領域中的資料分析，分析者較難有相類似的環境，去深入了解這些領域的知識規則。

目前國內的大學，尚未有適合運動相關科系學生修習之「資料科學」課程與教材，因此本計畫將秉持著因材施教適性發展的原則，以「運動分析」之實務問題為基礎來客製化發展適合運動領域學生修習的課程與教材，並實踐於教學課程中進行行動研究。課程的設計需要兼顧學生的學習動機與興趣、符合學生的知識技能與程度，並克服學生對於程式編碼的恐懼，建立學生學習的成就感與自信，並達到一定程度的學習成效。

設計出來的課程，於國立臺灣體育運動大學運動資訊與傳播學系大三(含大四)的「運動資訊管理」課程中實施，並以行動研究來了解學生對於課程的反應，探討課程內容設計、教學教法、技術實作成果，對於學習成就感與自信心、學習成效、學習滿意度之影響，並加以強化與改進課程，期望能達到以下的研究目的：

- 一、探討本研究所提出的課程設計，是否能符合體育相關科系學生之客製化學習需求。
- 二、利用課程中導入「運動分析」之主題與技術實作，探究學生在解析資料與程式實作過程之問題與解決策略。
- 三、探討問題導向式學習、專案導向式學習、合作學習策略、做中學之技術實作，對學生學習投入之影響。
- 四、綜合研究成果提出意見，以做為「資料科學」課程改進之參考。
- 五、透過行動研究歷程，幫助教學者自身的專業成長。

貳、文獻探討

與本計畫研究主題相關之國內外文獻中，篩選出幾個有關「資料科學」課程設計及實務現場之代表性案例進行評析，依所實施的學生主修領域不同，分述如下。

一、於資訊相關科系之案例

Asamoah, Doran 和 Schiller(2015)參酌業界顧問的意見和自身資料分析研究團隊的經驗，發展了一套「資料科學導論」(introduction to data science)的課程，並於美國萊特州立大學(Wright State University)開設，教學的目標主要是培養數據科學家，包括了三個方面：

- 1、數學和統計知識：學生應熟悉一套工具，用於匯總、分類和對（非）結構化數據進行預測分析；能夠構建的基本模型分析數據；通過模式識別和知識發現技術獲得經驗，以建立數據關係；展示數據分析技術和模擬現象。
- 2、資訊技術：學生應該熟悉現代的開源編程語言，以進行數據分析；學習資料視覺化工具的使用，以深入了解數據關係；了解支持基本數據分析任務，包括提取，清洗，分類或預測以及解釋。
- 3、實務上的專業：學生應能提出潛在的分析方法在當今數據豐富的環境中解決實務上的問題；了解如何解釋分析結果；運用科學和商業原則進行分析和使用基於計算機的分析技術和應用程序解釋數據。

課程中共設計了 8 個於課堂中動手實驗的任務，分析的案例非常生活化，分別為股市報

酬、通勤車輛、超市購物、博弈模擬、體適能、運動賽事、電影評論、程式犯罪報告，和一個期末小組專題。而不是個人書面報告和程式作業。課堂動手實驗的任務讓學生以小組合作學習的方式，在協作中學習資料科學工具的操作，並讓學生接觸真實世界中的實務資料以實踐資料分析結果報告和解釋。每一個實驗室作業包括使用 R 或 Python 處理實際數據。而課程主要授課的對象，有各種不同主修學科的學生參與，為了推廣跨領域的學習，小組的成員至少有兩名需來自不同主修領域，藉助學生對於特定領域實務的認知來深度解讀資料，共同完成期末的專題。而實際上在實施時，共有 14 位學生修讀，其中主修資訊管理的學生 4 位，電腦科學 9 位，1 位是非資訊相關科系。

他們所設計發展的課程，其品質最後是以跨領域小組成員間協同合作達成的程度之學習經驗，和個人對於學習成果與滿意度來進行衡量，由修過課的學生進行問卷調查。調查的結果顯示，在小組協同部份，學生對這門課程跨領域的特性、共同學習共同教學的模式感到非常滿意，且在學習經驗方面呈現出正向的態度。而在學習成果部份則反映出學生對於資料科學之知識與技能自認可以達到精通的程度。

二、於非資訊相關科系之案例

Chung(2018)認為應鼓勵非資訊相關科系的學生，主動積極參與學習資料分析。為了達到這個目標，他以醫學臨床學習常用的實證實務 (evidence-based practice, EBP) 來引導，並以學生為導向，開發設計了一個暑期短期共 10 週的課程，以工作坊的形式於美國南伊利諾州大學開課，提供非資訊科系的學生修讀。

課程的實施，包含了 EBP 的研討，以網路建立點對點的支持和讀書會來進行社交活動。而教材則採用主動學習，將教材放置於學習管理系統上讓學生在課前先進行閱讀，並準備了五個測驗題讓學生作答，而課堂上則是依據學生作答的情況，加以強化解決學習的困惑，確保關鍵概念的建立，課堂的最後則讓學生進行數據分析的實驗，以 R 語言進程式編碼進行實作。而課程學習成果的驗證，以學生是否能運用所學得的知識與技能，解答主修領域中的問題。以最後兩週的時間，完成團隊專題。

研究結果顯示，以 EBP 支持科學實證有助於發展應用資料分析的課程。同時採用主動學習法(active learning)使得學生在教師因資源有限而較少介入的情況下，仍可使學習同儕凝聚在一起。而課程也激勵了非資訊科系學生的參與，藉著影響學生去尋找在他所屬的領域中如何應用資料分析。最後課程的專題實作使得學生得以體驗大學階段的研究。

三、於體育相關科系之案例

Barneva & Brimkov(2019)於美國紐約州立大學(State University of New York)中，針對運動管理系所大學部的學生，開設了「資料分析」的課程。修課的學生僅需具備高中數學的程度，且大學修過 Excel、Access 軟體應用及基礎統計學，且不要求學生修過程式設計的課程。他們認為主修體育的學生，通常都對於計量(quantitative)與推理(reasoning)的課程感到困難，且偏好動手實作、服務學習、專題導向、實習的課程，而不是典型的講授課程。與計量有關的課程，大都是在通識教育中修習，且大都僅到達入門的程度。而對於資料科學家不可或缺的 Python 和 R 等程式知識，大都數體育相關科系的學生都沒有修過相關的課程。因此，「資料科學」的課程在體育相關科系中實施的難度，往往比電腦科學或數學相關科系來的高。此外，「資料科學」的課程，須於電腦教室中實施，以提供學生電腦動手實作，而不僅是講授理論。

課程的內容包含了資料的蒐集、資料視覺化、各種決策分析的方法，以及如何使用軟體去衡量運動的表現，競技體育如何透過資料分析來獲益等。並使學生能利用資料科學解決運動管理中常見的問題，如：預測隊伍的表現、制定獲勝策略、球員的分析、球員或隊伍的排名等等。課程的進行，是以專題導向的方式來進行，分階段一步一步地在作業中提出如何應用資料分析於某個運動項目，以提升表現、創造收益、預測趨勢、引導進步，最後期末完成一整個專業的正式報告。教材的部分，因「資料科學」涵蓋範圍很廣，尚未有合適的教科書，因此捨棄傳統的教科書，而採用免費的開放教育資源(Open Education Resource, OER)，包括免

費的線上教學課程、免費開放書籍、短片等。教師花費許多時間整理了適合教學的資源，讓學生在家中事先閱讀並完成指派的進度，而後於課堂中討論。為了克服學生在程式設計方面能力之不足，在軟體部份採用了 Knime 和 Excel 的 Solver 附加套件，使學生能使用軟體解答線性與非線性的問題，求解最佳化問題，且進一步可使用機器學習的模型，來找出致勝策略。

這樣的課程設計，經實施後驗證是非常可行的，最後修課的學生均能在運動的案例探討中，利用資料分析的模型來完成不同的階段任務，並且在獨立個人的專題中，能使用有關的分析模型，以動手做的方式實際進行運動資料分析後提出建議，最後完成專業的書面報告與公開演說發表。

參、研究問題

本計畫主要探討的問題為整合式課程之設計、教材開發，以及學生學習動機、學習成就之提升，而教學模式與策略則參照「主動學習」、「情境式學習」、「做中學」、「合作式學習」、「專題導向式學習」之精神加以融合，於課程教學實踐中進行行動研究。研究問題如下：

一、如何提升學習動機與興趣

藉由運動領域之實務問題來包裝資料科學課程，降低學生對於數理統計和程式設計的恐懼。課程的將參考現有運動領域中「運動分析」(sports analytics)課程常見的分析主題，挑選出適合發展運動產業資料分析的主題，將傳統以統計方法所進行的運動分析，嘗試以資料科學的方法重新進行另一種面向的分析與闡釋。

二、教材是否符合學生知識領域及技能程度

當學生欠缺豐富的高等統計知識及程式設計能力時，設計的教材是否能以高中數理程度和大學通識課程的程度，即能具備修課之基礎，有待進一步研究與設計。

三、如何建立學習成就感與自信心

參酌目前資料分析之競賽平台如 Kaggle，以及 GitHub 上之程式分享來編製課程實作之教材，讓學生親手透過軟體或程式之操作，完成指定的任務，提出真實世界資料解析之報告，讓學生實際感受到「I Can Do It Too」來建立成就感與自信心。

四、學習成效之確保

以團隊合作和專案導向式學習，讓學生完成小組獨立研究，引導學生發掘實務問題，並依循課程中所學的各個實作任務，綜整成為一個完整的專題研究，提出專業的分析報告與決策建議。

肆、研究設計與方法

一、研究架構

本研究採用行動研究(action research)模式來進行，在發展新的課程設計與設計教材的同時，執行經過規劃的教學行動方案、持續改進實務行動、監控並評鑑行動的實施歷程與結果，過程中透過反覆的循環，來改進教學的品質，提升學生學習的成效。

本研究在資料科學之分析工具，選用 Python 程式語言，以接軌學生先前在程式設計課程之基礎。在學生核心能力的部分，所設定的教學目標為：一、具備團隊合作精神，分工進行資料科學技術研究與運動分析問題之探討；二、瞭解運動領域常見的數據分析問題有哪些；三、從案例中觀摩完整之程式應用實例，了解數據分析之流程。針對上述之教學目標，擬訂

教學內容。

整體研究架構設計，如圖 1 所示，共包含四個部分：課程設計、課堂活動、主動學習與回饋反思，如圖七所示。第一部分為教師參照相關文獻之教學策略，以「運動分析」做為知識領域問題的發想，而「資料科學」做為問題的解答方法，兩者跨領域融合在一起規劃教學內容。第二部分課堂活動，為依課程主題所實施之資料分析實務操作，並將分析結果運用運動領域之知識做解析。第三部分主動學習則是為了解決課堂授課時數的限制，希望藉由學生自我主動學習，於課前研讀相關教材，並養成未來可利用網路開放教學資源自我主動學習之能力。第四部份反思與回饋，透過教學反思日誌、學習回饋單、問卷與訪談的方式、歸納分析問題與建議，並對課程進行修正。

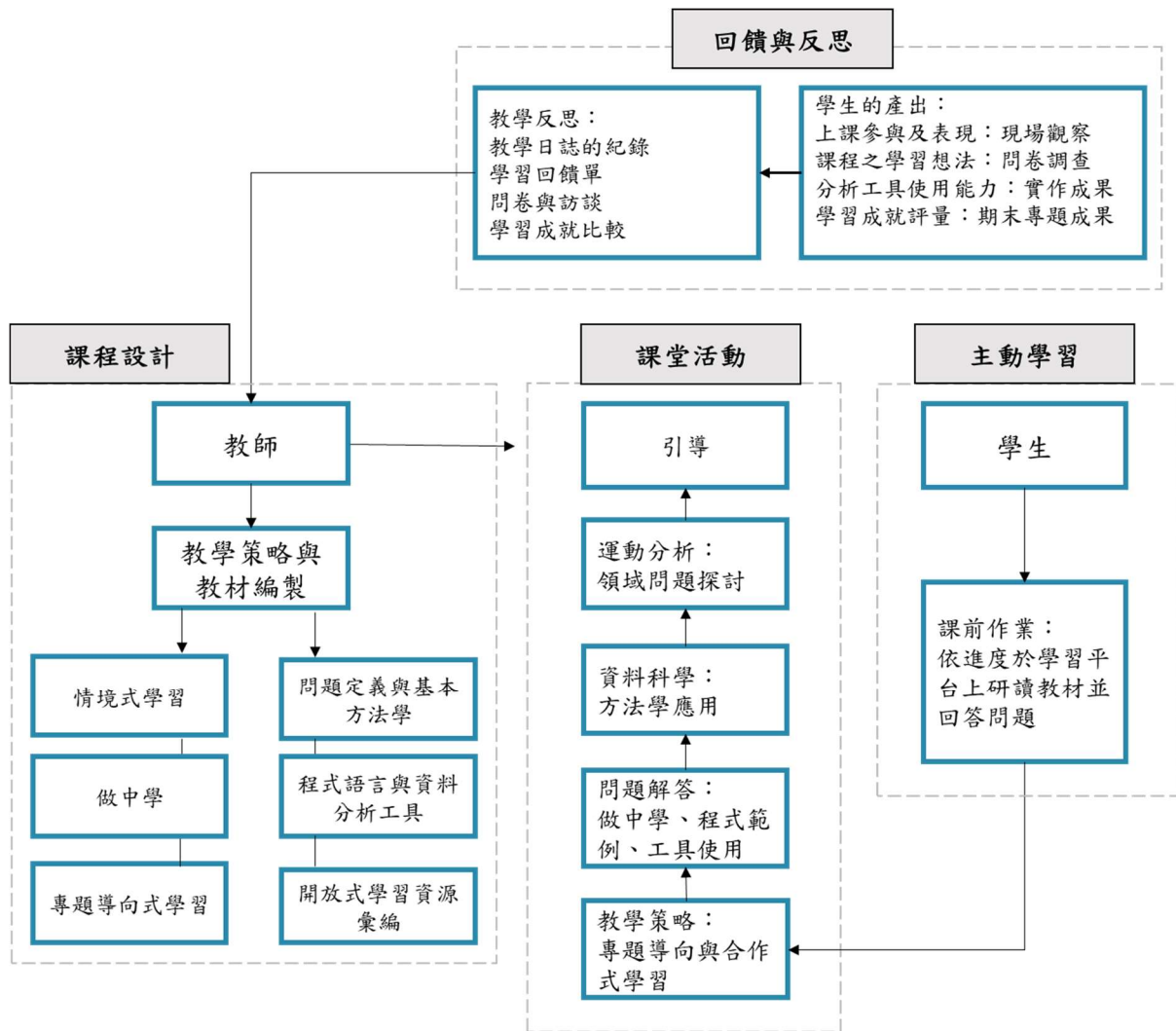


圖 1、研究架構

二、研究流程

本研究之流程如圖 2 所示，分述如下：

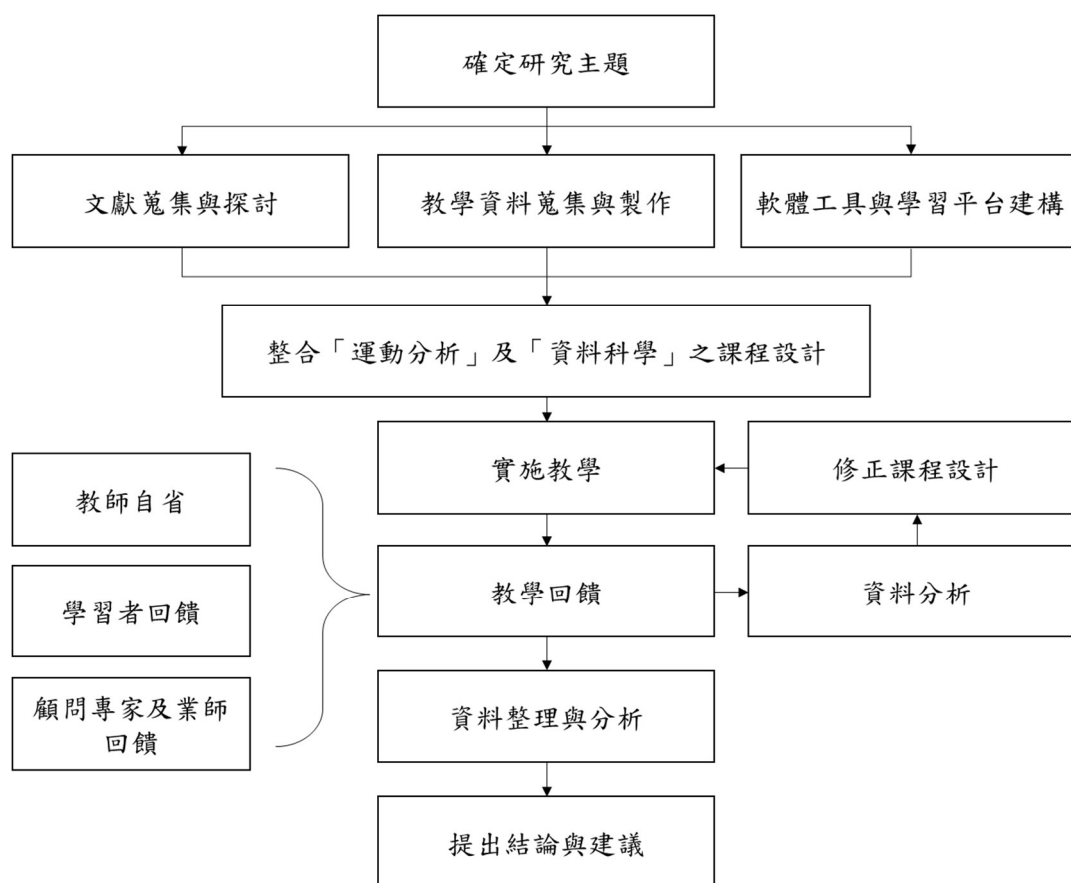


圖 2、研究流程

1. 確定研究主題
研究者由多年的實際程式設計教學經驗中，整理出幾個學習成效上的問題，做為研究的主题。
2. 文獻蒐集與探討
與「運動分析」及「資料科學」教學相關的文獻，在構思研究主题階段已有初步的瞭解。而在執行計畫階段，相關問卷的製作、質性分析的結果、實施成效分析等，需再進一步的參考相關文獻進行探討。
3. 教學資料收集與教材編製
研究者依據教學主题與單元目標，於開放式教育資源中搜尋，廣泛蒐集運動分析領域的各類分析問題與案例，如 Kaggle 平台上與運動有關之資料分析競賽題目，或是 GitHub 上與運動有關之大數據分析程式，經挑選並改編為適合研究對象能自我主動學習的程度，且能於課堂中實作，編製成教材。
4. 網路教學平台之設置
研究中將採用學校現有之 Moodle 平台，做為教材建置、問卷等教學輔助使用。此外，針對課程實務實作部分，則採用免費的開源軟體，如 WEKA、KNIME、Python Scikit-learn 套件等資料分析工具。這些軟體功能十分強大且涵蓋眾多資料科學之方法學，受限於學生程度及授課時數，課程中無法逐一詳述。故需配合教材，一一測試後選擇出最合適的操作的部份，或僅使用單一功能，做為本研究使用。當決定採用之軟體工具後，將相關教材、範例程式碼等，建置於平台上。
5. 整合「運動分析」及「資料科學」之課程設計
依所設計之運動分析主题，如與棒球分析、籃球分析相關之 Sabermetric、APBRmetrics

等領域知識，融入敘述統計、資料視覺化、蒙地卡羅模擬、機器學習演算法等資料科學的教學中。並以運動闖場景發展情境式教學，運動分析之案例實作讓學生能「做中學」，並引導學生以團隊合作學習方式完成期末專題。

6. 實施教學

本研究預計實施教學的時間，為一學期的課程共 18 週。約每 4 到 6 週為一個教學主題，每個主題均包含運動分析之領域知識、資料科學方法學應用、課堂實作練習、業師協同教學講座。課程行動研究以第九週為分隔，分為兩個階段。每階段完成後則進行教學反省與修正，再進入下一階段，如此循環。而在第十八週時，則為期末專題成果報告，做為學習成就評量的一部份。於兩階段教學結束後，再依據相關資料進行研究。

7. 教學回饋

教學過程中，蒐集學生的自我主動學習情形、課程實作之完成度、回饋單、學生訪談記錄、問卷填寫結果、以及研究者之教學省思日誌等，做為教學回饋。

8. 修正教學設計

依據上一階段之教學回饋呈現之缺失或不足處，調整各部分講解重點、實作時間、教材的難易度等，修正後於下一階段的教學活動實施。

9. 資料整理與分析

於每階段完成後之教學回饋資料進行彙整，待學期結束後，針對所蒐集的多方資料進行統計分析、三角檢證與統整，最後進行詮釋與分析。

10. 提出結論與建議

綜整一整個學期課程兩個循環的資料，進行整理並歸納出結論，對結果提出相關建議。

三、研究對象

本研究以研究者所任教之運動資訊與傳播學系的學生為對象，課程為 109 學年下學期大三系定選修之「運動資訊管理」課程，共有 18 位修課學生參與研究。大多數的學生具備入門程度的程式設計與資料庫能力，有大一上下兩個學期共四學分的 Python 程式，以及二年級上學期二學分資料庫管理的學習經驗。

四、研究方法及工具

本研究採行動研究法，在方法論上是採用折衷統合的方法，為質性與量化之綜合研究。其優點是可以運用許多不同的方法，蒐集資料。

在量化的部分，採用李克特式五分量表進行問卷調查，調查參與課程的學生對於教材內容各個知識點之學習困難，以及對教師課程進行方面之滿意度，以進行統計分析。而針對學習成效的部分，則以課堂實務實作任務的完成度，以及期末專題所達到的專業性、完整性來進行評量，並進行統計分析。

在質性的部分，則採用深度訪談法、學習回饋單，蒐集學生對於教學與自我學習成效之反饋意見，而對於教師則以觀察法，將教學省思記錄到教學日誌中，最後彙整多方資料進行分析，以概念圖、原因和影響分析、解釋性論述，來闡釋資料。

伍、教學暨研究成果

一、教學過程與成果

1. 教材之編撰

本研究在課程設計上，是以開放式教育資源(OER)做為核心。在課程開課前的準備階段，研究者首先設定好授課的主題，再依據主題進行網路教育資源的瀏覽，這些教育資源在經過一系列的評估程序後，將符合條件的教材列入候選，之後配合課程的教學進度，重新組織發展成為正式授課的教材。於課程開始前，建構於學校之 Moodle 學習平台中，包含了許多網頁資料的連結，以及 YouTube 的影音短片，可供學生隨時瀏覽。

本研究採用的開放教育資源教材評估準則，分成五個面向，說明如下：

- 完整性：需有逐步的分析說明，並提供範例之程式碼。程式碼可搭配網頁示範或提供操作介面者為佳。
- 可讀性：能於 30 分鐘內供教師快速瀏覽完畢，了解大意者為佳。適合學生的英語程度或以中文撰寫為佳。
- 易懂性：以入門程度之體驗為主，學生以學習過入門程度之 Python，具備基本運動比賽(籃球、棒球)知識，未曾接觸過資料科學者為對象。
- 權威性：創作者具備相關學識與精力，網站主機具長久性或官方網站者為佳，內容為原創而非抄襲。
- 普遍性：內容為一般常見之主題，具有一定的討論熱度，有一定的瀏覽數、討論度、參與度。

由於網路上的資料在近幾年隨著線上課程的推廣，以及部落客、Youtuber 等內容創作者的興起，充斥著各式各樣良莠不齊的內容，要找到一個有組織化、系統化，內容適合大學生學習，需經過教材開發者一一審視內容，方能列入教材。最終所選用應用於課程中的教材，共分成三個主題，相關資源與說明如附件一所示。在提升學生的學習興趣上，我們也列出了運動分析之問題場景和數據科學教學內容的對照。

課程設計上，第一個主題「運動數據庫的應用」的教學目標，是複習資料分析會使用到的 Python 語法，包含資料型態、以及統計、資料框架、繪圖等套件，以最入門的運動統計為場景。接著，介紹如何取得分析所需的資料，使用了 Python 的網路爬蟲套件或是可直接從運動網站上下載資料的 API 套件。

第二個主題為「NBA 籃球數據分析」，教學目標是要讓初學者了解「資料科學家」這個職業在做甚麼，以及在 NBA 的發展現況。之後則開始導入入門的應用案例，從資料的整理、到視覺化的呈現，進行初步的多變量相關性分析。而在這個主題中，也開始引入 GitHub 和 Kaggle 這類的開放資源，以引導學生於日後能具備自主學習之能力。

第三個主題「MLB 棒球數據分析」的教學目標，則聚焦於機器學習模型之應用，介紹模型建構流程之資料清理與過濾、特徵工程、模型之選擇與誤差測量等程序，以及一些演算法用來解決分類、預測等問題。考量學生的程度，教學中有涉及到機器學習演算的部分，並未詳加介紹，而是將模型當作是黑盒子帶過。

而在作業設計部分，每個主題的結束，均搭配一個以組為單位的作業，共有三個。這三個作業的設計，是隨著對於資料科學的認知，而有不同的目的。第一個作業，主要是讓學生能夠複習 Python 的程式操作，並學會安裝運動分析所需的套件。以課程中所介紹的 sportsipy 套件和範例程式，在眾多數據中找出勝率最高的隊伍、薪資最高的球員、計算某隊某年度之得分等，這類資料過濾與篩選的操作。第二個作業，是讓學生於 GitHub 與 Kaggle 中檢索，報告與 NBA 有關之應用，並要求同學實際依照所提供的程式碼，實際於個人電腦的 Python 環境中，實作過一遍。第三個作業，同樣是讓學生從 GitHub 和 Kaggle 中檢索與 MLB 有關的應用，但檢索的關鍵字需同時加入機器學習有關的關鍵字，如人工智慧、類神經網路、決策樹、叢集分析等。而報告的重點不再要求程式實作，而是讀懂要解決的運動分析問題、解決的方法、使用的資料與欄位、結果呈現、學習心得。

2. 業師協同教學

為了強化學習的動機，課程中也安排了兩次的業師協同教學，邀請實務界的專業人士以自身經驗，說明運動分析在國內的發展。邀請的業師分別為東森媒體的運動主播，以及中華民國運動分析師協會之秘書長。實施的情況如圖 3 所示。而學生對於兩場演講，都給予極高的評價。



(a) 東森媒體熱血體育主播丁元凱



(b) 中華民國運動分析師協會張宏韶秘書長

圖 3、業師協同教學

二、教師教學反思

在實施這套教材教學的同時，教師於課堂上透過與學生的互動，了解到學生的反應及回饋，記錄於教學省思日誌中，同時也在實施問卷調查同學的回饋意見後，對於全體的學生進行深度訪談，以了解學生學習的情況，訪談之問題與紀錄詳見附件二。經過研究者反思後，逐步改進教學，歸納發現如下：

1. 採用 GitHub 與 Kaggle 等開放教育資源，能提升學習動機

許多學生反映在這門課中首次接觸到 GitHub 與 Kaggle，覺得認識這些資源對於學生將來的發展會很有幫助，願意深入去了解它。

2. 了解現今運動領域資料分析師的職業以及資料分析在職業運動中的應用，能提升學習動機

學生對於介紹 NBA 的資料分析師以及他們如何協助 NBA 球隊提高競賽表現的影片，感到新奇。有助於後續各教學單元學習動機之提升。

3. 設計以分組方式從 GitHub 與 Kaggle 檢索後選定內容進行報告的作業，有助於培養學生主動學習的能力

在現今各類網路資料隨手可得的環境中，此類型的作業有助於增加小組協作下之資訊過濾與判讀的能力。而依檢索內容程式碼實作的過程中，所產生之問題與解決，有助於培養主動學習的能力。然而在課程的報告中(作業二、作業三)，發生了有不同組學生擇同一內容報告，主要是 GitHub 與 Kaggle 有完整說明且學生能自行讀懂得資源仍有限。

4. 學生對於運動分析中的運動專業詞彙與計算方式，感到困難及挑戰性。

通常學生僅對單一運動種類有較深入的了解，課程中所介紹的其他運動種類需花較多時間學習。而在棒球統計學、籃球統計學中的表現指標計算，不容易理解，再加上多以英語方式呈現，增加學習的困難度。

5. 學生對於自行依範例進程式碼實作，感到困難及挑戰性。

雖然學生已具備 Python 語言的入門基礎，但對於資料科學之分析程式實作，這類資料導向的程式設計，仍感到困難。此外，學生自行依 GitHub 上之程式碼實作時(作業一)，常會遇到 Python 和套件版本之環境問題不知如何自行解決，以致無法完成實作。建議未來可於這門課之前，另外安排資料導向程式設計之課程。且再補充 Open source 之軟體版本解決方案之操作。

三、學生學習回饋

研究中，我們分別在期中(第九週前測)和期末(第十八週後測)，以學習體驗的自我評量問卷來蒐集學生的學習意見回饋。學習體驗的自我評估(self-assessment)有三項：(一)較以往更深的理解(以下簡稱為「更理解」)。(二)處理信心提升(以下簡稱為「有信心」)。(三)對設計課的造形發展有幫助(以下簡稱為「有益設計課」)。同一群修課的學生以同一份問卷進行前後測，因樣本數 N=18 屬小樣本，因此不進行信度、效度檢定，僅統計平均分數與標準差，以及前後測之成對樣本 t 檢定，來驗證教學前後之差異，問卷之內容及結果如表一。

表一、學習體驗自我評量問卷統計結果

問題	前測		後測		顯著性 (雙尾)
	平均值	標準差	平均值	標準差	
1. 我認為本課程的學習目標是清楚的。	3.89	0.471	4.17	0.618	0.172
2. 我很想瞭解如何運用“資料科學”來進行“運動分析”的探討。	3.44	0.618	3.61	0.922	0.592
3. 我認為在“資料科學”的教學中，使用“運動分析”做為案例，可提升學習的興趣。	3.78	0.922	4.06	0.778	0.263
4. 目前，我對“資料科學”了解的程度是清楚的。	3.06	0.778	3.72	0.732	0.055
5. 藉由實作練習作業，我對“資料科學”的理解比以前更深入。	3.83	0.732	4.06	0.725	0.386
6. 在完成實作練習作業之後，我有信心將“資料科學”處理得比以前更好。	3.61	0.725	3.78	1.056	0.564
7. 練習處理 GIT/Kaggle 上之開放程式資源，對我的“資料科學”能力發展有幫助。	3.61	1.056	3.78	0.752	0.130
8. 目前，我對“運動分析”了解的程度是清楚的。	3.11	0.752	3.89	0.618	0.054
9. 藉由實作練習作業，我對“運動分析”的理解比以前更深入。	3.89	0.618	4.06	0.639	0.507
10. 在完成實作練習作業之後，我有信心將“運動分析”處理得比以前更好。	3.50	0.639	3.94	0.698	0.134
11. 練習處理 GIT/Kaggle 上之開放程式資源，對我的“運動分析”能力發展有幫助。	3.50	0.698	3.94	0.878	0.163
12. 現在我有信心將“資料科學”應用於“運動分析”上，同時處理完善。	2.72	0.878	3.61	0.778	0.011 *

表一顯示，學生對於課程的內容與目標(問題 1)是清楚的，而且非常認同課程中的程式實作活動對學習有幫助(問題 5、問題 9)。然而對於信心程度(問題 12)在前後測都同樣相對略低，但前後測間的提升，卻是顯著的。而在前後測的比較，對於資料科學(問題 4)與運動分析(問題 8)的認識皆有所提升。

陸、建議與省思

本研究在實際實施一學期課程後，提出之結論如下：

1. 運動分析實務問題與資料科學兩者融合為一，可提升學生之學習動機。
2. 利用開放教育資源的學習經驗，有助於提升學習成就上的自信心。
3. 學生之間所具備之程式語言基礎及運動知識存在著顯著差異。
4. 對於利用資料科學進行運動實務問題的分析，學生感到最大的學習困難為程式的實作，以及有關運動統計及其計算公式之英文詞彙。
5. 利用開放教育資源的學習經驗，有助於提升學習成就上的自信心。

而對於未來教學上，提出之建議與省思如下：

1. 在有限教學時數下，融合了兩種學科進行教學，課程須設定為入門程度，無法探討到進階的應用。在課程推廣上，也可考慮定位在通識課程中實施。
2. 利用開放教育資源來開發教材，未來可考慮利用更嚴謹的決策方法(如：德菲法)加入專家和學生的意見，來挑選更合適的教材。
3. 課程的設計應搭配更完整的課程地圖之配套措施，在先備課程先建立好學生之基礎能力。

參考文獻

- Asamoah, D. A., Doran, D., & Schiller, S. (2015). Teaching the Foundations of Data Science: An Interdisciplinary Approach. 2015 Pre-ICIS SIGDSA Workshop, Fort Worth, Texas.
- Barneva, R.P., Hite, P. (2017). Information technology in sport management curricula. *J. Educ. Technol. Syst.* 45(3), 326–342.
- Barneva, R. P., & Brimkov, V. E. (2019). Teaching Data Analytics to Sport Management Majors: A Practical Approach. In *International Workshop on Learning Technology for Education in Cloud* (pp. 408-417). Springer, Cham.
- Chung, S. (2018). Data Analytics Workshop Series for Non-Computing Major First-Generation-College-Bound Students. *Information Systems Education Journal*, 16(5), 37.
- Schumaker, R. P., Solieman, O. K., & Chen, H. (2010). *Sports Data Mining*. ISBN 978-1-4419-6729-9. Springer Science+ Business Media, LLC.

附件一：教材內容與相關開放教學資源

單元名稱	運動分析之問題與目的	資料科學所用到的技能	OER 網址	OER 說明
運動數據庫之應用	1. 認識運動科學中會用到的統計計算 2. 認識運動數據分析的資料來源，以及可用的數據類型 3. 查詢並擷取 reference 網站的運動數據	1. 資料型態 List、Dictionary 2. 資料框架 Pandas 3. 網站爬取 Beautiful soap 套件 4. 套件的使用: sportsPI	Python for Sport Scientists: Descriptive Statistics https://towardsdatascience.com/python-for-sport-scientists-descriptive-statistics-96ed7e66ab3c	這是 Towards Data Science Inc，發布於 Medium 線上部落格發布平台的一系列文章。可提供讀者免費存取。在運動科學家的 python 文章中，教導了常被使用到敘述統計之平均值、中位數、標準差、變異數，以及相關係數之 python 程式案例，使用了虛構的跑步距離與速度當資料集。
			sportsipy https://pypi.org/project/sportsipy/ https://sportsreference.readthedocs.io/en/stable/	sportsipy 是一個免費的 python API，可直接從 www.sports-reference.com 擷取所需要的資料，以便利 Python 應用程式之開發，特別是資料分析、機器學習有關之應用。
籃球數據分析	1. 數據分析時代的 NBA 2. APBRMetrics 簡介 3. 分析 Mapping Kobe Bryant's shots. 4. 籃球數據分析的程序 5. 球員 RPM 分析、進攻與防守、薪水的分析	1. 資料科學家的職業 2. 資料地圖的視覺化分析 3. NBA Stats API 4. 資料分析流程 5. Git 與 Kaggle 資源與介面的介紹 6. seaborn 中的 heatmap 函數於相關性分析 7. seaborn 的多變量分析及圖表	The NBA Data Scientist https://youtu.be/MpLHMKTolVw	介紹一位女性數學博士，於 NBA 費城 76 人隊擔任資料科學家。說明工作的性質與日常工作
			SportsVU NBA https://youtu.be/jOQE1_tkEwE	介紹 SportsVU 以攝影機為基礎之球員追蹤技術
			The math behind basketball's wildest moves https://www.ted.com/talks/rajiv_mahe_swaran_the_math_behind_basketball_s_wildest_moves	介紹可運用機器來計算籃球比賽中，距離、位置、命中率、投籃熱區等數值
			How data transformed the NBA The Economist https://youtu.be/oUvvhkXyOA	NBA 在數據分析發展下的轉變
			Every shot Kobe Bryant ever took. https://wikieducator.org/Sport_Informatics_and_Analytics/Pattern_Recognition/Python , https://graphics.latimes.com/kobe-every-shot-ever/ , https://github.com/datadesk/kobe-every-shot-ever	Kobe Bryant 曾經於球場上所投過的球，以於球場地圖上以視覺化的方式呈現。提供了可供互動查詢的網站，以及背後運作的程式碼。而這個資源也被 Wikieducator 的課程引用做為案例
			NBA Stats API. http://danielwelch.github.io/documenting-the-nba-stats-api.html , https://danielwelch.github.io/little-pynny/	供 python 讀取 stats.nba.com 網站數據之免費 API，開發應用程式使用
			NBA Analytics With Python. https://princetonportsanalytics.wordpress.com/2017/09/30/nba-analytics-with-python-a-tutorial/	Princeton Sports Analytics club 所提供之 NBA 分析簡單入門範例，涵蓋了資料的取得、處理、視覺化、學習模型之應用。

			NBA Player RPM Prediction Defense vs Offense. https://www.kaggle.com/dhamlett/nba-player-rpm-prediction-defense-vs-offense/notebook	分析球員的效能
			球員效率與薪資的分析. https://zhuanlan.zhihu.com/p/42625197	中文分析球員效率與薪水
棒球分析	1. Sabmetrics 及常用統計數值介紹 2. 以歷史對戰紀錄進行蒙地卡羅計算勝率(非OER) 3. 不同年代之每場比賽得分、勝場數 4. 球員入選棒球名人堂之預測 5. 打擊角度、速度與結果之分類	1. Excel 之數據整理、統計工具箱(非'OER) 2. 蒙地卡羅模擬範例(非'OER) 3. 視覺化之資料探索(Exploring and Visualizing The Data) 4. 機器學習模型之流程: 資料清洗與前處理、特徵工程、模型選擇與誤差度量、 5. sklearn 套件之使用-K mean、logistic regression、random forest 6. 機器學習分類模型之介紹:KNN、SVC、gBDT	GLOSSARY https://www.mlb.com/glossary	MLB 官網上所提供與棒球相關之專有詞彙說明, 包含許多賽博計量學所使用之表現統計指標
			Scikit-Learn Tutorial: Baseball Analytics Pt 1. https://www.datacamp.com/community/tutorials/scikit-learn-tutorial-baseball-1	datacamp 主要是提供付費的資料科學課程, 而有部分為免費的入門教學。此資源為一步一步以 Python 程式碼介紹如何利用機器學習模型來進行資料分析與預測
			體育分析與機器學習 https://www.zhihu.com/column/c_1005779975458443264	知乎上的專欄, 約有 20 多篇利用機器學習在運動分析上應用的短篇文字, 主要在說明運動分析的步驟, 專有名詞的解釋, 部分有提供程式碼。
			Classifying MLB Hit Outcomes. http://tylerjamesburch.com/blog/baseball/hit-classifier-1 https://github.com/tjburch/mlb-hit-classifier/blob/master/notebooks/1-model-selection.ipynb	波士頓紅襪隊分析師 Tyler James Burch 個人的部落格, 說明打擊結果之分類, 並提供 github 之程式碼

附件二：學生訪談問題與紀錄

1. 你覺得課程中的哪個部分，是最具挑戰性的？請具體說明一到兩項。

- 我覺得因為是有 project 要做，我對那個 NBA 或者是棒球他們的專業像 era 這種指標數據原本是沒有什麼概念，所以我覺得如果要做報告的話就要特別去查這些指標代表的意義是什麼，會比較陌生。
- 我覺得要把資料全部讀懂，就是你要看得懂所有的程式碼它在幹嘛，然後要做什麼，然後跟所有的結果，就是要把它全部讀懂我覺得比較難的。
- 我覺得是 MLB，就是最後一個作業，因為我自己平常是對 NBA 籃球比較有興趣，數據也都比較看得懂，那如果是 MLB 的話就是要再去了解一些數據方面的東西，規則都了解，但是像我今天報告 war 是我第一次知道的事情，然後很多數據的方面我都是沒有什麼概念，就是比如說像籃球好了，我知道這個數據跟這個數據結合在一起可以做什麼用，但是如果像棒球的話我就必須要去更了解這項運動才可以去，就是怎麼講，就是可以讓人家，就我自己懂的我講出來大家也會懂這樣，就可以理解這個東西。
- 最具挑戰性就是這個期末報告的前一個報告，就是要做實作的那個部分，一開始總是會失敗，就是有時候會找不到原因。
- 因為像我們有時候就不一定知道說一些算數學公式嗎？我們知道說可能可以切入的點有哪些，比如說籃球可以從一些投籃命中率或者是一些其他東西下手，但有時候有些數學公式就不知道到底要怎麼去運用出來。
- 我覺得是作業 3 個裡面可能我自己覺得籃球吧，因為那時候好像還要改那個程式碼的作業，就是要自己再去嘗試看哪個方式會成功，然後因為我自己覺得我對程式碼沒有很熟悉到可以一下就可以完成，所以還要上網查資料之類的，就是試比較多次，所以這個可能會花一點時間這樣。
- 因為我個人沒有看一些運動比賽的習慣，所以對於這個運動資料大概要整理哪一個部分的敏銳度比較低，就是不知道哪一個數據加減乘除會比較影響到那個比賽的重要性，當然什麼攻擊能力、防守能力這種對於比賽的這個成績有比較顯著的影響這個當然知道，但是一些什麼像是籃板、三分球什麼的，甚至是一些比較比較複雜的學術名詞因為他們都會用縮寫，所以其實全英文的話直接放翻譯也不太會知道他的那個縮寫到底在講什麼，就是可能這個入門會比較困難一點。
- 我覺得如果不要講到程式相關我都覺得還好，如果是單純數據分析那些我覺得蠻有興趣，但是程式的話我可能就沒辦法那麼去了解。
- 應該是就是要閱讀英文，因為我的閱讀真的不是很行。如果他今天是中文的話其實就很好懂。
- 應該是比較後面就是籃球跟棒球這些數據的分析，就是我對他那個數據分析比較不擅長，然後又加上英文比較差，所以看那些數據就很吃力。
- 課程的部分，我覺得應該是最後這一個機器學習的部分，因為就很多專業術語要去了解，不然文章會看不懂他到底是在說什麼。

2. 經過這門課的學習，你覺得在哪方面的學習與訓練，獲得了成長，影響你最多？請具體說明一到兩項。

- 我覺得是老師要求的報告模式，因為之前課堂比較少會有這種去研讀一個項目然後去跟大家分享的這種型態，所以我覺得之後也可以把這種形態應用在更多的科目上面。
- 我了解就是在任何體育數據，不管籃球、棒球還是其他運動，運動科學對體育帶來多大的貢獻，就是比如說我可以看到一些數據，我可以在運動比賽的競技過程中可以知道，看到這個數據我可以去為下一場比賽去做準備，就可以去做一個統計，就是大數據統計出來，比如說像剛剛 war 值，這個球員對球隊的貢獻，因為這是在比賽過程中的看不出來的東西，那經過這個大數據分析，這些東西資料庫啊這些跑出來的數據圖模擬圖啊相關圖可以讓一些分析師或者是在找球員的人可以更方便再，怎麼講，可能對球團有比較大的貢獻，怎麼講...
- 可以得知那麼多數據，就是一開始我們其實沒有觸碰過個東西，就是什麼數據然後可以去看這個球員怎麼樣，但是可以透過這些數據然後用程式的方式顯現出來。
- 我覺得比較有印象深刻是老師放一些籃球的研究的影片，比如說可能透過戰術的一些分析或者是某一個球員他很注重在哪一個點進行投射的那種。
- 找 NBA 跟 MLB 這個，用 kaggle 去找。對，因為之前有聽過這些網站，但是也沒有實際過來查，就是有實際查過就發現其實這邊蠻多資訊可以看的。
- 最多應該就是那個程式實際去使用，就是學以致用的感覺，就是學的比較多，因為之前比較說是老師做一個我們做一個，這堂課就是比較多自己去發揮的空間，學比較多。
- 我自己是覺得可能是發現一些新網站，因為老師都會貼像 kaggle 那種網站，然後如果我沒有上這堂課其實我也不知道有這個網站可以去找，用關鍵字去找一些你要的資料的內容，然後之前沒有接觸到，透過作業就可以去了解一些可能資料管理的方式跟做法，所以我覺得這應該是比較特別的地方。
- 學習應該是就是看那個英文吧，因為都是英文的網站，也剛好就是上這堂課就知道說有這個網站可以去查跟程式跟數據有關的東西，就是多一個管道可以去查資料，因為以前不知道，是上這堂課才知道的。
- 對，然後對於那個運動的專有名詞理解程度是增長最多的吧，因為要去找資料，然後資料的話每個都有很多名詞，所以每個都要找，直到找到可以用的那個資料，就是可能會認識了非常多不同的東西，或是那個程式編寫的一些方法。
- 資料的整理吧，雖然程式碼我不會，但是資料整理就是還蠻讓我懂得怎麼收集資料的，像是用在畢製之類的都蠻有用的。
- 有辦法去使用別人寫好的東西，然後甚至是去進一步的修改這樣，我覺得這部份成長比較多。
- 分析這個東西之前比較沒有那麼常碰到程式，就是在程式做統計分析這部分之前比較少做，所以這堂課做了比較多了練習。
- 我自己是覺得像後面兩次的籃球數據跟棒球數據分析，就是都蠻有趣的，然後也利用

kaggle 跟 github 有找到蠻多資料，然後分析的內容也蠻感興趣的這樣。

- 應該就是因為也在摸索當中，所以算是有稍微對於這些數據分析概念比較深一點的感覺，就是有因為接觸這些實作，就是在做作業這幾次這樣慢慢一次一次比較更了解，所以我也覺得說好像要讓大家多做一點，才知道說到底現在上課在上什麼。
- 我覺得應該就是能夠看得懂那些分析出來圖表，然後可以去進行分析，然後了解他，然後就是之後也可以有自主學習的能力這樣。

3. 對於下一個年度已選修這門課的學生，你可以給他們學習上的建議是？請具體說明一到兩項。

- 我覺得他們如果有他們自己本身喜歡看的運動，不管是棒球或是 NBA，可以建議他們就是，因為網路上其實會有很多作家他們都會特別去寫說這些指標代表的意義是什麼，其實都蠻有趣的，如果特別喜歡這些運動的話可以從這些運動下去看說他們自己的指標的意思是什麼。
- 英文要看得懂，還有一些可能數據上面它們的縮寫吧。
- 就是真的要認真上課，沒有認真上課單純自己看是看不懂的，然後就是也不要前一天做報告，前一天基本上是做不了，因為真的會卡很久，就是有挑戰性的那種，不是一般你隨便搞出來的，就是值得學習的這樣子。
- 就是要認真學，因為其實如果將來要走那個情蒐或是數據分析是蠻有用的。
- 就是可能之前的程式課要認真一點，然後課後自己要花多一點的時間去研究，比較不像之前一年級和二年級的靠課堂上就可以。
- 我覺得整體來說這堂課的作業量其實滿足夠的，就是就是大家在上課的過程可以算接觸蠻多蠻多次的，然後如果有興趣的人其實可以學到蠻多東西，然後我覺得如果說可以把作業次數分少一點，但是可以再更難一點嗎？就可以更深入，因為像籃球跟棒球就是他們好像主要是去 kaggle 上網找資料然後去分析人家做的東西，看可不可以再更深入一點這樣，可是作業量可能就相較來說會比較少。
- 可能就是老師這禮拜講的東西要稍微跟上進度，不然可能之後會又要再補上次的進度，就是每次的進度可能都要跟上吧我覺得。
- 主要還是上課要認真聽，老師上課都會講，如果有問題的話可能就是跟同學討論或是再找時間問老師這樣，因為這種東西比較難一點，如果是程式類的話。
- 使用好的關鍵字很重要嗎？因為 kaggle 跟那個其實資料不太好找啦，就是都會有一些資料，應該說資料很多，但是一個完整能夠跑得動或是有一些資料比較少一點，因為很多作者都會把一些運作時候會要跑的一些東西移除嗎？因為他可能什麼著作權或是他自己提出來的，然後這樣的話就會缺少的部分這樣就沒辦法運作，比較像這樣子的？
- 我覺得可能要英文要好一點，因為我們很多網站都是就是是英文的網站，就是你要看得懂英文才比較好理解他的內容。
- 我覺得要對運動可能要有點了解，就是不管是做棒球和籃球，其實最基本就是那些數據的了解，你要對他有一些基本認識才看得懂他到底在講什麼，有一些像英文，因為

都是原文的話就是要看懂話基本上就是要了解他到底是那個數據在表達什麼，他想要呈現的是甚麼，所以我覺得比較大的建議就是要先對他們有基本認識這樣。