

## 中山管院微學程修習心得

姓名：鄭子婷 系所：資訊管理學系 填寫日期：110/06/17

※請勾選修習之微學程

CM504 數位行銷與電子商務微學程

CM505 商業大數據分析微學程

CM506 金融服務與金融科技微學程

1. 當初為什麼會選擇此微學程來修習。

中山管院致力於推廣商業數據相關課程，在修習此學程前已經有相關數據分析的經驗，因此想藉有完整的微學程課程更加提升在數據分析上的知識與技術。而在三個微學程中對於金融的知識較為薄弱，希望能藉由金融與數據的實作案例，了解大數據運用在金融案例方式，以及加強自己對於金融服務與科技上的知識。

2. 請分別敘述三門課 (R 語言、大數據/巨量、資料案例)的最大收穫。

1. **R 語言**：這門課程的內容是從基礎的 R 語言到模型預測再到最後的商業規劃，雖然老師每周的進度很多，但因為老師完整的對資料分析的整個流程進行教學，因此除了學到技術外，我覺得最大的收穫是**了解了商業上如何進行完整的資料分析與運用**。我曾經也以為資料分析僅僅是將資料做分析、統計、模型預測，只要模型結果好就是好的資料分析，但上完這堂課才了解到資料分析應該要包含資料分析結果的運用，若能真的使用在企業的行銷、成本評估等等的才能算是好的分析，因此我覺得這門課格外重要，除了**學習到基礎的 R 語言程式能力，也能真正了解資料分析，更能加強商業數據的能力**。

2. **巨量**：在巨量的課程中，最大的收穫是**學會許多機器學習的重要觀念**，在修這門課之前我只會對資料進行處理、探索性分析、統計檢定等，對於模型預測比較不熟悉，而上完這門課後，學習了很多機器學習上的概念，例如如何進行特徵工程、如何對模型做參數的調整、衡

量模型的好壞，而除了技術面的觀念，這門課的實作也很多，老師也讓我們自己跑模型，在建模的過程中看到模型的成效一次比一次好也獲得了很多成就感和經驗，透過實作更加了解模型的運作。

3. **資料案例**：這門課雖然老師上課的時間不多，但是作為微學程的最後一門實作課，我認為可以說是收穫滿滿，除了讓我們**了解了真實的公司是如何使用數據進行分析、預測、行銷規劃外，也讓我真切的感受到大數據在現代對於企業的重要性**。在這門課中，我們從一開始的資料探索、特徵工程、模型建置，到最後的行銷方案都是透過小組不斷討論與實作的成果，在過程中雖然遇到種種困難，但是經過實作，不論是 R 語言的程式能力、建置模型的方法與優化、商業的領域知識、或是行銷方案的規劃都獲得了很大的提升。而我選修的是金融服務與金融科技學程，除了把前兩門課程所學的知識與技術實作在真實的資料外，原本對於金融一竅不通的我也獲得了許多相關知識，像是許多金融相關的指標、重要的名詞、銀行業的各種產品。

3. 三門微學程強調同學的跨域核心能力，請舉例說明你如何從課程中獲得跨域能力。

在 R 語言的課程中，我與企管系以及社會系的同學為一組，在討論的過程中，因為我比較擅長程式碼的部分，因此我多半負責複雜的技術以及幫忙除錯，而其他像是資料分析的方向、行銷方案的規劃等等企管系的同學都能發揮他們的專業，提出很好的想法，而社會系的同學也在我們實作台灣地區分析的作業時，提出許多社會指標的計算方式，讓小組作業能夠順利完成。

而在資料案例的課程中，我與一位資管碩的學長以及一位財管系的同學一組，我們依照不同領域進行分工，資管系的同學負責實作模型，而財管系的同學便運用他們的財金領域知識對我們的模型進行解釋與行銷方

案的應用。因為我選擇的是金融科技學程，我們的資料屬於銀行須保密的資料，因此資料都進行了去識別化，這也提升我們在資料分析的過程中對專業領域知識的吸收，例如我們必須了解何謂資產配置、證券戶、薪轉戶等金融相關知識，而在小組討論的時間中，財管系的同學就能夠清楚的為我們解釋。

4. 總結性資料案例課程主要以專案方式結合廠商內部資料，讓學生發揮實作與分析能力，請闡述如何實踐「做中學」。

我所實作的廠商是台新銀行，總共的資料是兩百多萬筆的資料，首先是在探索性資料分析的部分，在這個實作過程中我學到的是如何挑選欄位進行資料探索，一開始我對於要如何進行資料探索無從下手，因為資料的欄位數太多，不知道應該挑選甚麼欄位進行分析，但在詢問助教後，助教教導我們先利用模型找出變數的重要性，並將重要的變數挑出來進行分析，如此就可以避免對一些較不重要的資料進行探索而沒有實質的幫助結果。

接著是模型預測的部分，在實作模型時，才發現真實的資料應該要進行許多處理，在過去的課程中往往會使用網路上已經整理好的資料集進行模型訓練，但我們所拿到的資料是沒有經過特別處理的資料，在過程中我學會對資料進行資料型態的轉換、根據模型使用目的切分訓練集與測試集、還有對類別變數的 one-hot encoding、對連續變數的正規化。另外是資料集的樣本不平衡問題，因為從來沒有實作過有這樣問題的資料集，因此透過這次實作讓我了解真實世界中有許多像是這樣的不平衡問題，也讓我學會處理這樣問題的方式，例如 undersampling、oversampling 等作法，以及這些作法的好處與壞處。而最後在評估模型的階段，我也學會應該依照不同的目的對模型進行評估，選擇不同的指標和評估方式對模型結果的解釋也會有所不同。

接著是對客戶做分群行銷，在這個過程中我學到如何結合跨領域知識與我的程式能力，例如我利用爬蟲與資料分析的技巧分析 PTT 的相關知識，以及利用演算法進行分群，而財管系的同學從分析的結果中提出各種行銷方案，讓我學會如何運用分析的結果，以及藉由這些運用使分析的結果為企業帶來更多的效益。

5. 對於微學程整體設計或課程上之建議。

我認為第一門的 R 語言課程，學生的比例應該規劃為至少一半的同學為有資訊背景的或寫過程式的經驗，因為這門課雖然屬於基礎的 R 語言課程，但有許多實作性的作業，在期中與期末也有報告競賽，但因為有資訊背景的同學太少，因此導致每組只有一位較會寫程式的同學，這樣會造成技術分配不均，在討論的過程中雖然老師規劃是會的人教導不會的同學，但 1 對 4 的家教教學還是令人頗有負擔。

巨量分析的課程，我認為應該設立門檻，例如需要先修 R 語言課程，或是其他資料分析和機器學習等相關課程，因為在分組實作的時候發現有許多同學根本不會寫 R 語言，對機器學習也沒有很熟悉，導致幾乎無法進行小組討論。

另外我認為兩門共同課都有線上資源和老師上課的影片是非常好的做法，因為有許多非資訊背景的同學，或許在第一次聽課會無法馬上理解，但如果這些資源可以有很大的幫助。

6. 對於非資訊背景學弟妹修習第一門共同課-R 語言，請給予學習上之建議。

我認為在修這一門課之前，如果能夠有基礎的程式能力會比較能跟上老師的進度，因為寫程式並非只有了解程式語言，邏輯也是很重要的，因此如果能先寫過簡單的程式，能有比較好的邏輯學習新的程式語言。另外，也需要具備基礎的統計能力，因為這門課的進度很快，許多統計概念老師並不會一一講解，若能有基礎的知識，在老師講解時不必花費太

多時間理解，否則在一堂課內除了要聽懂統計的運算外還要看懂程式碼需要花費很多時間，也有許多同學因為無法在一堂課，甚至一周內吸收這些知識所以棄選，因此我**建議非資訊背景的同学，在修這門課前可以到管院商業數據的網站查看歷年的教學大綱和教材，若能事先有基礎能力在學習上會有較好的效果。**