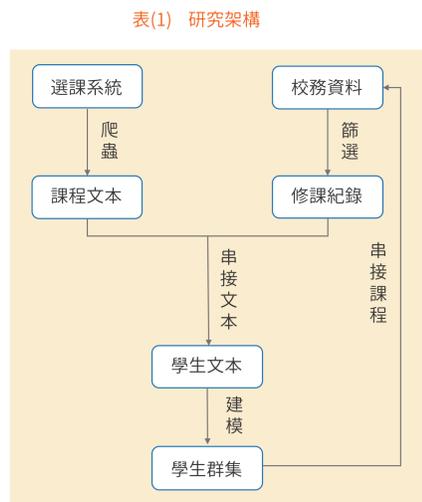


一、前言

近年來，學校越來越重視跨領域人才的培育，但受到系所限制，不易觀察共同重視的領域能力。

假定相似的課程應有類似的課程敘述，則能打破系所的框架，透過文本的角度將「修課文本」類似的學生分群，建構各群之課程地圖，提供學校跨領域課程設計的建議，乃至不分系入學生的課程參考。

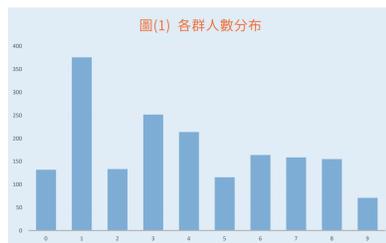
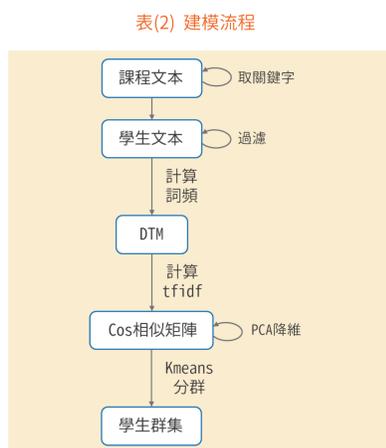


二、不同「領域」的學生樣貌

A、Kmeans分群

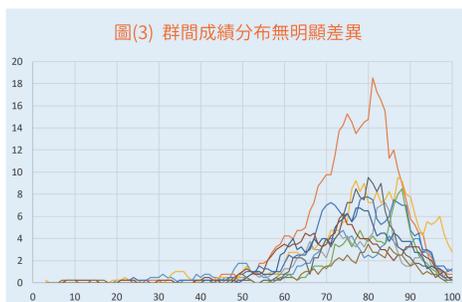
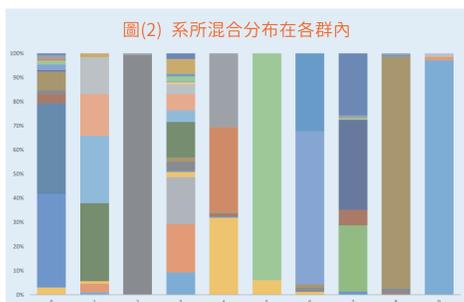
將爬蟲取得的課程文本，以人工方式標記出每堂課程的關鍵字，再串接學生的修課紀錄，合併後得到「學生文本」；即把每個學生都視為單一文本，據此建立文件詞頻矩陣(DTM)。

接著，將DTM採逆向文件頻率(TF-IDF)去噪後，計算兩兩文本間的餘弦相似度，再將得到的相似矩陣當作文本特徵，以主成分分析(PCA)降至8維後，使用k-means將學生分成10群。



B、檢視分群效果

為避免分群結果與類別變數具有相關，故就「系所比例」與「成績分布」兩個方面檢視分群的有效性；各系所大致混合分布在不同群內(圖2)，群集間也具有類似成績分布(圖2)，顯示整體的分群結果尚可。



C、「領域」樣貌 — 以管院為例

第一群

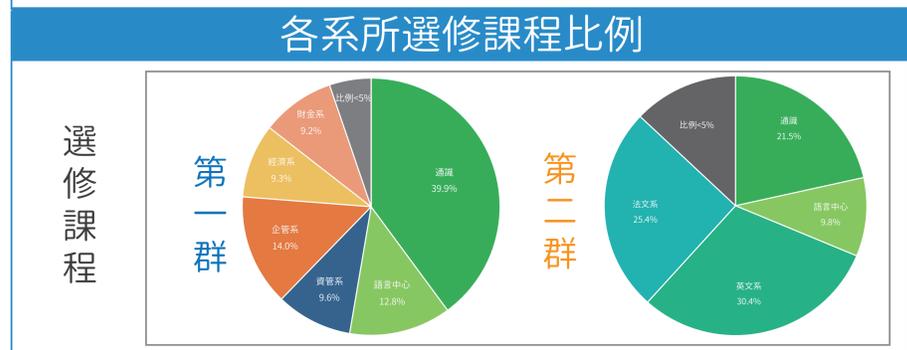
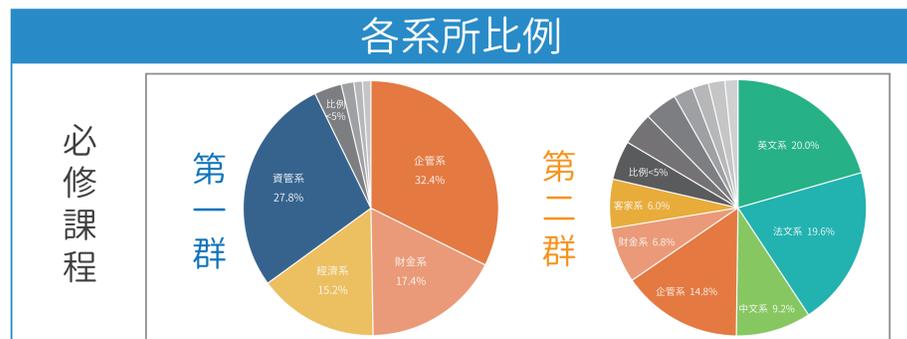
第二群



三、不分系入學的課程地圖 — 以管院學生為例

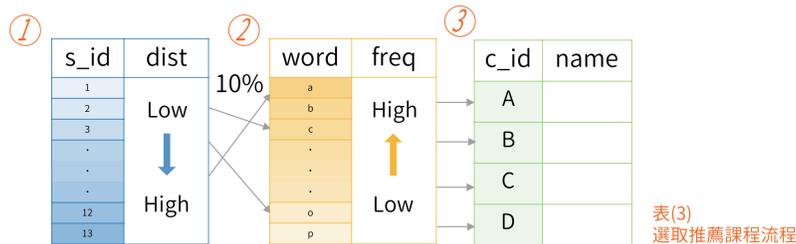
A、領域能力

由於院訂必修仍為基礎必要之核心能力，故用這兩群學生內，占比較高的系所，作為必修課程的依據，至於選修課程則選出詞頻高者作為推薦。



兩群的學生佔比皆以管理、文、客家學院為主，其中以第二群的學生組成比例更多元。以詞彙頻率及開課單位來看，第一群著重管理決策、金融市場、經濟分析等商業行為；第二群則重外文、溝通表達、文化、國際觀及貿易等面向。

B、課程地圖



- ① 先找出此群距離中心最小的前10%學生
- ② 取那些學生中出現次數最多的前10個詞彙
- ③ 再將詞彙對應到課程，依出現次數做排序

第一群以管院為主，推薦課程著重管理決策、金融市場、經濟分析等；第二群以文院為主，側重外文、溝通表達、文化、國際觀及貿易等課程。

推薦的選修課程	第一群	第二群
	企業資源規劃	法語文化概論
	初階程式設計	新聞法文
	管理會計	法文視聽會話
	行銷學	國際視野 I
	中級會計學	法國藝術與文化

四、結語

為確保文本內容充足，分群一定程度上會受到必修課程影響，使得推薦的課程內容，可能與使用院來區分的結果相似。另外，學生修課會受到系所的限制，換句話說，學生能自由修課、透過修課資料反映興趣動機的效果有限。因此，這些課程所代表的關鍵詞聲量，將會壓過興趣選修或是「跨領域」的課程，就結果來論恐有循環論證之虞。但回到系所混合的學生群集，確實能夠跨越系所限制，一定程度上反映系所間的相似性，提供不分系入學發展的領域參考。