請尊重智慧財產權，合法影印資料並使用正版教科書
Please consult Intellectual Property Rights before making a photocopy. Please use the textbook of copyrighted edition.

國立東華大學

課　　綱　Course Outline

資訊管理學系學士班

中文課程名稱 Course Name in Chinese	巨量資料分析				
英文課程名稱 Course Name in English	Big Data Analysis				
科目代碼 Course Code	IM__33670	班 別 Degree	學士班 Bachelor’s		
修別 Type	學程 Program	學分數 Credit(s)	3.0	時 數 Hour(s)	3.0
先修課程 Prerequisite	統計學(一)				

課程目標
Course Objectives

巨量資料是大量、快速、多樣及真實性可質疑的新興資訊領域。它藉由蒐集並分析大量且多樣的資料集，尋找資料所隱含的各種可能樣式(Patterns)或趨勢(Trends)以利組織管理者進行分析及決策。另外，經由大量資料的計算及分析，用以推論及估計某些事件發生的機率並以此去做出決策。本課程藉由相關課程的講授，以幫助學生能夠瞭解巨量資料，並運用相關技術去描繪資料的輪廓、分析資料間可能存在之差異或相關原因、以及預測資料可能之結果。本課程屬於進階巨量資料分析，除了主要著重相關?述統計、統計檢定、資料探勘理論外，也會輔助以相關實作進行教學。
註：本課程屬於「巨量資料分析」進階課程，不得抵免通識課程之「巨量資料分析」之基礎課程。

系教育目標
Dept.’s Education Objectives

1	培育具備團隊合作、創造與思考以及國際視野能力之人才。 Cultivate the personnel with the capability of teamwork, creative thinking and international vision
2	培育具備資訊專業知識與技術能力之人才。 Cultivate the personnel with the capability of professional knowledge and technical skill
3	培育具備商業管理知識與應用能力之人才。 Cultivate the personnel with the capability of business management and application
4	培育具備整合資訊科技與管理知識之專業人才。 Cultivate the personnel with the capability of integration of information technology and management

系專業能力 Basic Learning Outcomes	課程目標與系專業能力相關性 Correlation between Course Objectives and Dept.’s Education Objectives

A	具備資訊管理基礎與跨學域的應用能力。 Cultivate the personnel with the capability of basic and interdiscipline information management	●
B	具備以資訊科技為核心，擁有高度專業技術與國際視野之能力。 Cultivate professional personnel with international perspective and the capability of using information technology	●
C	具備資訊管理創新、研發、企劃之資訊管理人才之整合能力。 Cultivate senior personnel with the capability of innovation, research and development and integrated planning	●
D	具備企業資訊化的能力。 Cultivate the personnel with the capability of computerization of enterprises	●
E	具備業界多媒體應用、網站經營以及資訊行銷所需之能力。 Cultivate the personnel with the capability of multimedia applications, website operations, and information marketing	○
F	具備認知新興資訊產業發展所需之能力。 Cultivate the personnel with the capability for the emerging information industry	●

圖示說明Illustration ：● 高度相關 Highly correlated ○中度相關 Moderately correlated

課程大綱
Course Outline

巨量資料(Big Data)又稱大數據、海量數據，指的是資料數量的規模龐大到無法透過人工，在合理時間內達到擷取、管理、處理、並整理成為人類所能解讀的資訊。巨量資料幾乎無法使用大多數的資料庫管理系統處理，而必須使用數十、數百甚至數千台伺服器上同時運行。巨量資料的定義與企業資料之多寡，以及其平常用來處理分析資料的軟體之能力息息相關。對某些組織來說，第一次面對每日數百GB的資料量可能讓他們需要重新思考是否需採用不同的資料管理技術。然而，對於跨國(或大型)組織，資料量可能需要達到每日數十或數百TB時，才會對他們造成困擾。巨量資料中有著珍貴的訊息，像是相關性(Unknown Correlation)、未顯露的模式(Hidden Patterns)、市場趨勢(Market Trend)，可能埋藏著前所未見的知識跟應用等著被我們挖掘發現，透過相關分析技術，我們可以找出資料中所存在之隱性模式或趨勢，並以此去預測未來可能發生之事件。巨量資料與傳統資料差異在於下列四個特性(4Vs)：資料量(Volume)非常龐大(大)、資料輸入輸出速度(Velocity)非常快速(快)、資料類型(Variety)非常繁雜(雜)、資料必須確保真實性(Veracity)並過濾有疑慮(疑)資料。

本課程主要包含四個階段：「課程導論與資料處理」、「資料敘述級」、「資料檢定級」、「資料預測級」，每個階段都有1到3個章節的內容。針對第一階段(課程導論與資料處理)，我們除了介紹這門課程(第一章)，而且我們還會讓學生瞭解不同類型資料(第二章)。針對第二階段(資料敘述級)，我們除了讓學生畫圖看出資料的概況(第三章)外，我們也會找出資料的中心與離度(第四章)、分群(第五章)，以及應用規劃求解來找到問題的解答(第六章)。針對第三個階段(資料檢定級)，我們除了介紹資料檢定的意義(第七章)，還會針對不同資料間進行比較，如是否有差異(第八章)、是否有相關(第九章、第十章)、是否明顯比較多(第十一章)。針對第四個階段(資料預測級)，我們除了介紹相關預測處理流程(第十二章)，還會使用貝氏網路的結果預測與原因推理(第十三章)，以及自動文本分類之實作(第十四章)。

資源需求評估（師資專長之聘任、儀器設備的配合・・・等）
Resources Required (e.g. qualifications and expertise, instrument and equipment, etc.)

本課程以熟悉巨量資料分析之師資為主，並且以理論課程為主，實作操作為輔。原則上，針對每個章節，我們除了講授相關理論基礎外，我們也設計數個與該章節相關內容相關之實作學習單，以供學生操作及回答相關問題，因此本課程以電腦教室操作為上課場域。本課程操作相關軟體或工具包含：Google Sheet、Weka以及其它相關巨量資料分析工具。

課程要求和教學方式之建議
Course Requirements and Suggested Teaching Methods

本課程使用教師整理網路相關教材，總計包含14章，教學方式主要包含下列幾項：

1.課程講授 2.上機實作操作實作學習單 3.期未個案報告
其他 Miscellaneous
Grading Criteria 1.出席率、學習態度、實作學習單（20%） 2.期未個案報告（20%） 3.期中評量（30%） 4.期未評量（30%）