



校務發展與大數據分析

銘傳大學

王金龍





課程內容

簡介

校務研究

資料處理

資料分析

商業智慧

研究議題

研究案例

結語





簡介



科技如何改變我們的生活經驗？

雜貨店

超市

便利商店

網路購物



科技如何改變我們的生活經驗？

紙本書信

電子郵件

手機簡訊

即時通訊





校園資訊部門-業務興隆



資訊安全

電腦教室

校務行政

網際網路

無線網路

CA TV 網路

安全監控

遠距教學

數位電話網路

IoT 圖書資料通行證

E 化教室

教學平台內容

APP 系統

雲端系統

3D 列印

VR/AR

E-portfolio

課程地圖

MOOCs



個資保護

大數據

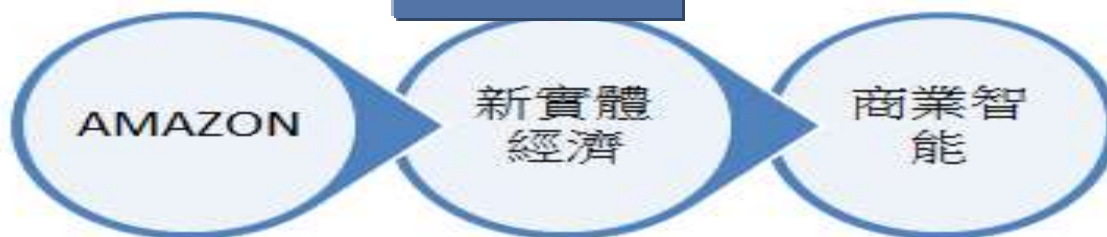




學習成效
數位學習
產學銜接
國際接軌



校務研究
教學卓越
系所評鑑
校務評鑑



資訊系統於教育運用的趨勢

研究

高速計算
科研模擬
實驗網站
開放原始碼

行政

教務系統
學務系統
財務系統
人事系統
決策系統

教學

教學平台
MOOCs
學生E-portfolio
教師E-portfolio
學習地圖

校務研究

大數據分析
學習成效分析
教學平台數據分析
招生研究
畢業生追縱



大學資訊環境

校務行政 線上異動處理

高教行政教學順利推動
Database

校務自動化的進行(MIS)

應用系統各自產生報表

無法提供者整合性資訊

校務研究 線上分析處理

支援學校管理者的決策
Data Warehouse

商業智慧(BI)

資料整合/主題導向

不變動性/時間差異





校務研究

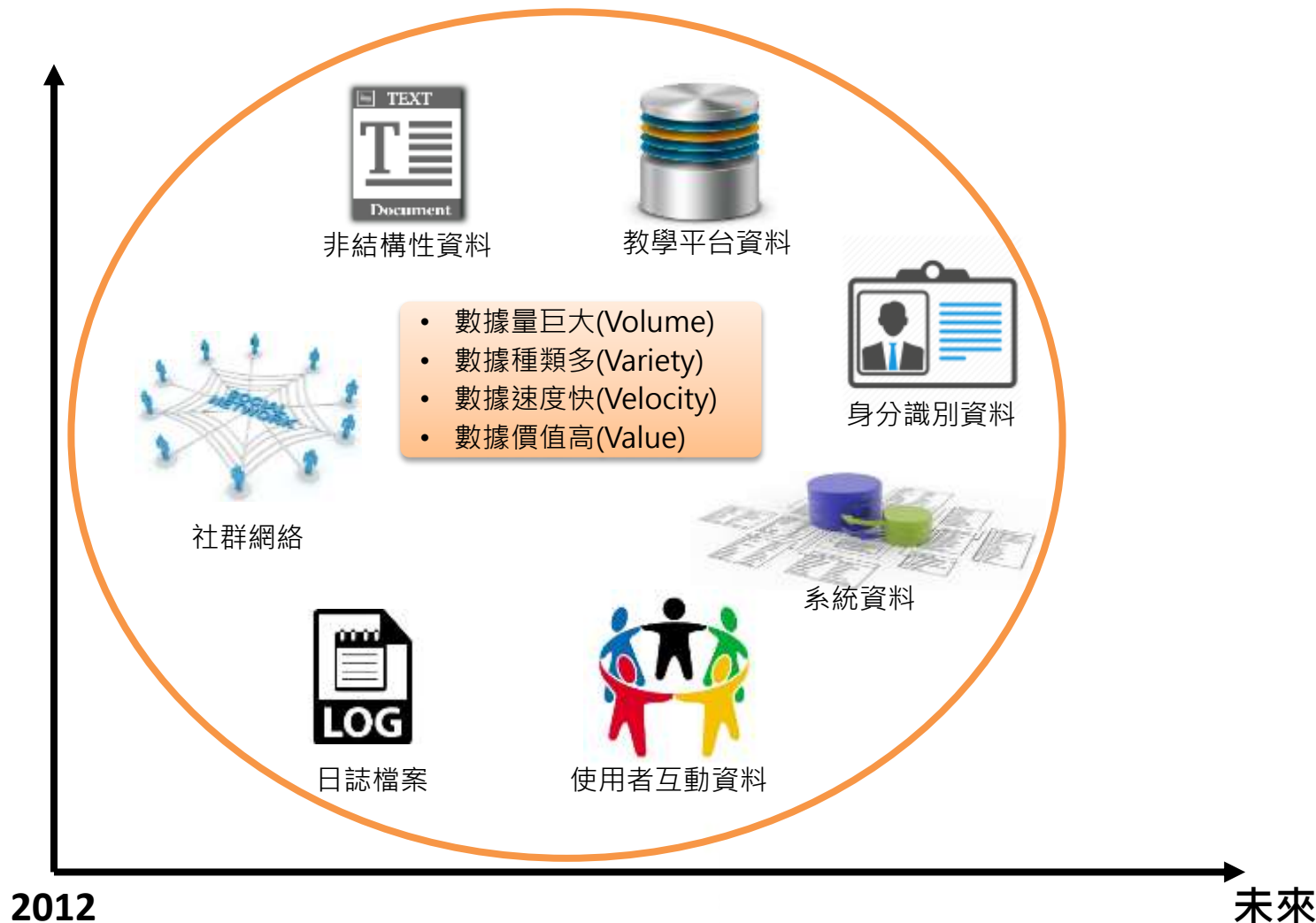


校務研究

- GPS 導航系統：校務發展規劃
- 落實自我改善：學校基礎工程
- 確保品質保證：驅動教與學提升
- 提升校務行政：服務品質、效率與效能
- 數據基礎決策：系統性地收集、決策參考
- 高度相關：中長程計畫、教學卓越計畫、校務評鑑、系所評鑑

全方面教育大數據

資料量





大學資訊系統架構

教育、學校管理人員

院長、系主任、教師

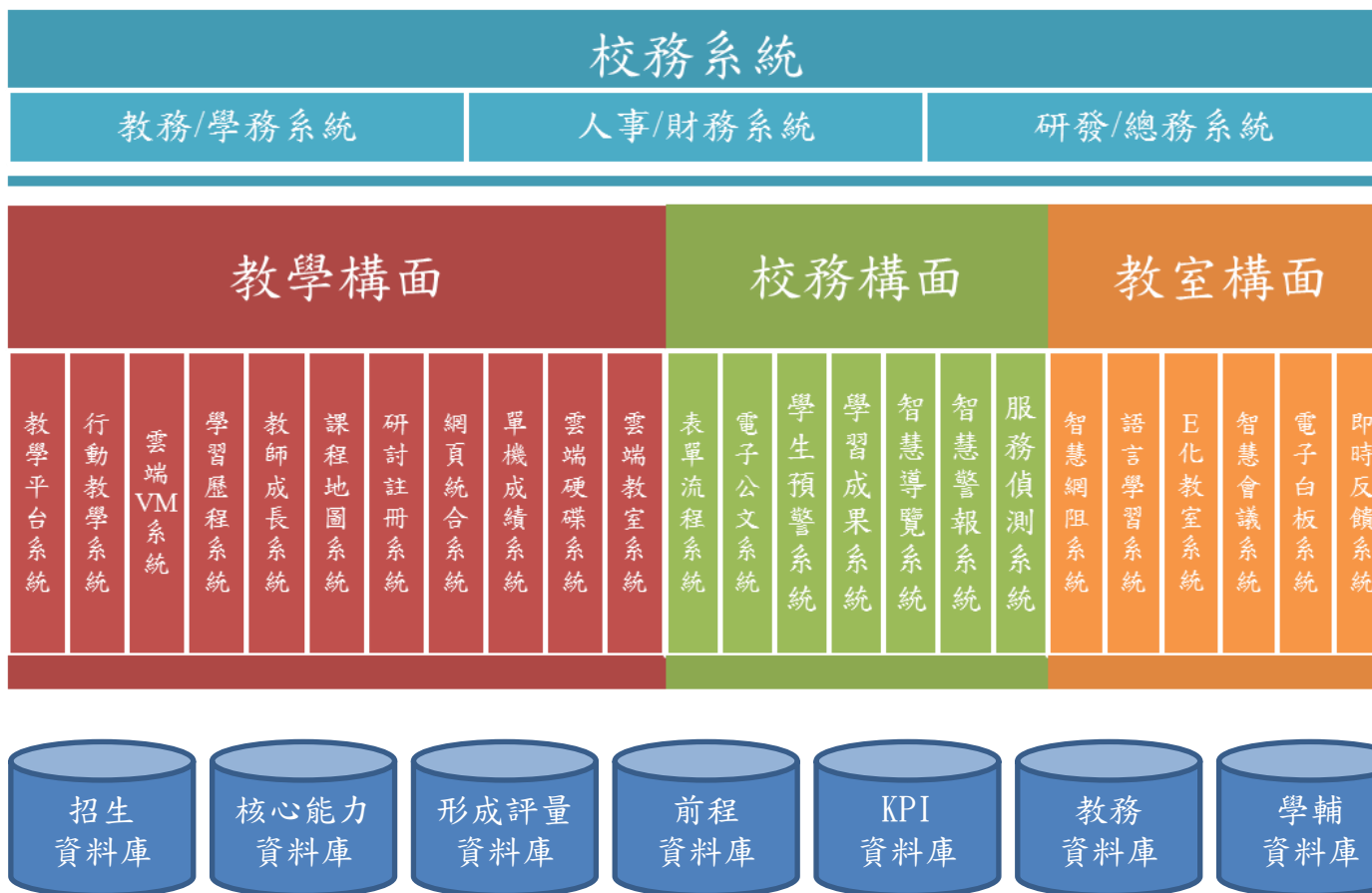
學生、家長、社會人士

技術服務團隊

標準規範體系

持續營運體系

安全管理體系



教學與學習

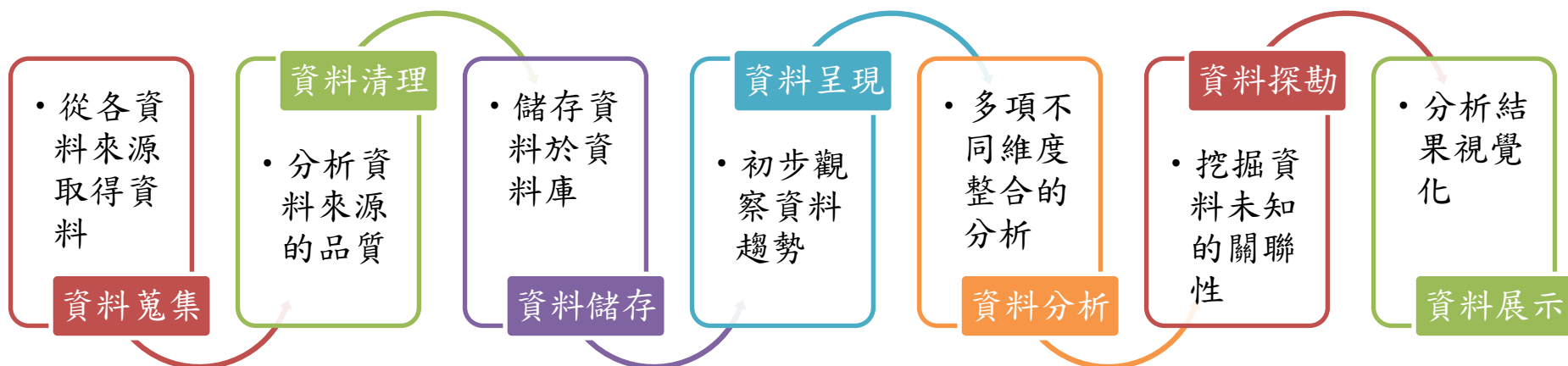




資料處理



資料處理模式



資料匯整：70%

資料分析：20%

結果：10%



資料匯整

- 有效分析的關鍵：從各種運作系統的資料中，提取出有用資料並進行**清理**，保證資料的正確性
- **擷取、轉換和載入**：合併到的資料倉儲中
 - 資料擷取：**從資料來源處擷取所需之數據資料**
 - 資料轉換：針對所擷取出之數據資料，依照邏輯需求，針對數據資料作適當的轉換
 - 資料載入：最後將已作適當轉換過的數據資料載入到目的地



彙總資料須轉換

省市	100學年度				101學年度			
	學士	碩士	博士	合計	學士	碩士	博士	合計
福建	198	38	6	242	381	80	10	471
江蘇	62	25	6	93	91	69	10	170
浙江	187	36	1	224	390	89	5	484
廣東	177	31	4	212	317	81	7	405
上海	55	17	2	74	99	36	5	140
北京	45	34	4	83	100	84	11	195
湖北								
遼寧								
合計	724	181	23	928	1378	439	48	1865



	A	B	C	D
1	學年	省市	學制	人數
2	100	福建	學士	198
3	100	福建	碩士	38
4	100	福建	博士	6
5	100	江蘇	學士	62
6	100	江蘇	碩士	25
7	100	江蘇	博士	6
8	100	浙江	學士	187
9	100	浙江	碩士	36
10	100	浙江	博士	1
11	100	廣東	學士	177
12	100	廣東	碩士	31
13	100	廣東	博士	4
14	100	上海	學士	55
15	100	上海	碩士	17
16	100	上海	博士	2
17	100	北京	學士	45
18	100	北京	碩士	34
19	100	北京	博士	4
20	101	福建	學士	183
21	101	福建	碩士	42

須將資料轉換為詳細資料



資料管理之建議

資料收集

資料字典：規劃資料
參數與欄位的定義

資料輸入：發生端

資料分類統整
定期自我盤點

資料儲存

資料來源：選擇/儲存
資料(代碼)

建立標準schema

資料庫：資料中心
雲端環境

資料使用

總整資料：反正規化
資料共享：維持一致

視覺：BI
分析：SPSS, Excel
預測：Weka

教務處：建立資料
總務處：讀取資料





資料分析





大數據資料分析



Descriptive Analytics

名詞

- 描述性分析、敘述性分析

釋義

- 尋求事件發生的狀況，在歷史數據中挖掘有意義的資訊，反映與剖析過去表現

方法

- 使用BI工具觀察資料變化
- Power BI、Qlik View、Tableau





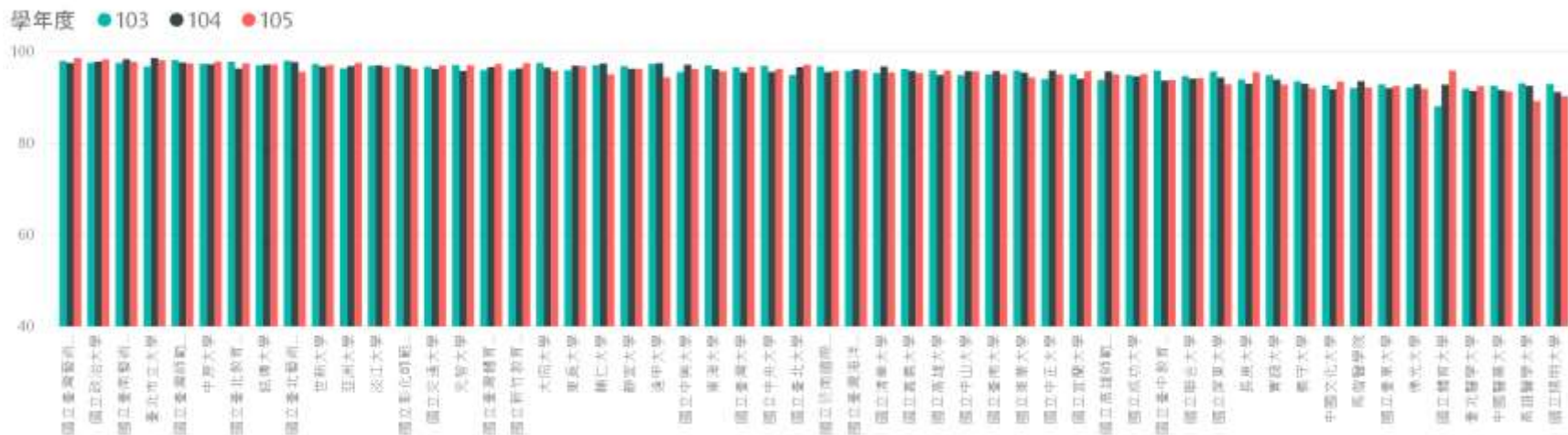
103~105各校學系註冊率

學年度	設立別	學校代碼/名稱	學校代碼	學校名稱	學門代碼	學門名稱	學類代碼	學類名稱	科系代碼	科系名稱	註冊率
105	私立	1004 中原大學	1004	中原大學	14	教育學門	1406	特殊教育學類	140601	特殊教育學系	100
105	私立	1006 中國文化大學	1006	中國文化大學	14	教育學門	1401	綜合教育學類	140101	教育學系	94.83
105	私立	1006 中國文化大學	1006	中國文化大學	14	教育學門	1403	專業科目教育學類	140315	體育學系	94.9
105	私立	1067 台灣首府大學	1067	台灣首府大學	14	教育學門	1404	學前教育學類	140401	幼兒教育學系	88.57
105	私立	1048 亞洲大學	1048	亞洲大學	14	教育學門	1404	學前教育學類	140401	幼兒教育學系	98.85
105	私立	1020 南華大學	1020	南華大學	14	教育學門	1404	學前教育學類	140401	幼兒教育學系	93.33
105	公立	0013 國立中正大學	0013	國立中正大學	14	教育學門	1405	成人教育學類	140501	成人及繼續教育學系	100
105	公立	0020 國立東華大學	0020	國立東華大學	14	教育學門	1401	綜合教育學類	140125	教育與潛能開發學系	95.35
105	公立	0020 國立東華大學	0020	國立東華大學	14	教育學門	1403	專業科目教育學類	140329	體育與運動科學系	92.68
105	公立	0020 國立東華大學	0020	國立東華大學	14	教育學門	1404	學前教育學類	140401	幼兒教育學系	97.37
105	公立	0020 國立東華大學	0020	國立東華大學	14	教育學門	1406	特殊教育學類	140601	特殊教育學系	100
105	公立	0020 國立東華大學	0020	國立東華大學	14	教育學門	1407	教育行政學類	140710	教育行政與管理學系	97.5
105	公立	0052 國立屏東大學	0052	國立屏東大學	14	教育學門	1401	綜合教育學類	140101	教育學系	97.5
105	公立	0052 國立屏東大學	0052	國立屏東大學	14	教育學門	1401	綜合教育學類	140107	教育心理與輔導學系	92.31
105	公立	0052 國立屏東大學	0052	國立屏東大學	14	教育學門	1403	專業科目教育學類	140315	體育學系	88.89
105	公立	0052 國立屏東大學	0052	國立屏東大學	14	教育學門	1404	學前教育學類	140401	幼兒教育學系	95
105	公立	0052 國立屏東大學	0052	國立屏東大學	14	教育學門	1406	特殊教育學類	140601	特殊教育學系	95
105	公立	0052 國立屏東大學	0052	國立屏東大學	14	教育學門	1499	其他教育學類	149911	科普傳播學系	86.67
105	公立	0001 國立政治大學	0001	國立政治大學	14	教育學門	1401	綜合教育學類	140101	教育學系	100
105	公立	0014 國立高雄師範大學	0014	國立高雄師範大學	14	教育學門	1401	綜合教育學類	140101	教育學系	96.88
105	公立	0014 國立高雄師範大學	0014	國立高雄師範大學	14	教育學門	1403	專業科目教育學類	140302	工業科技教育學系	92.5
105	公立	0014 國立高雄師範大學	0014	國立高雄師範大學	14	教育學門	1403	專業科目教育學類	140315	體育學系	87.5
105	公立	0014 國立高雄師範大學	0014	國立高雄師範大學	14	教育學門	1406	特殊教育學類	140601	特殊教育學系	94.44
105	公立	0038 國立新竹教育大學	0038	國立新竹教育大學	14	教育學門	1401	綜合教育學類	140108	教育心理與諮商學系	97.5
105	公立	0038 國立新竹教育大學	0038	國立新竹教育大學	14	教育學門	1401	綜合教育學類	140122	教育與學習科技學系	98.72
105	公立	0038 國立新竹教育大學	0038	國立新竹教育大學	14	教育學門	1402	普通科目教育學類	140212	英語教學系	100



103-105學年度各校系註冊率

各學年度各校系註冊率



設立別

- ☐ 公立
☐ 私立

學年度

- ☐ 103
- ☐ 104
- ☐ 105

領域名稱

- ☐ 人文及藝術
- ☐ 其他
- ☐ 工程、製造及營造
- ☐ 教育
- ☐ 服務
- ☐ 社會科學、商業及法律
- ☐ 科學
- ☐ 農學

學門名稱

- ☐ 人文學門
- ☐ 傳播學門
- ☐ 其他學門
- ☐ 商業及管理學門
- ☐ 工程學門
- ☐ 建築及都市規劃學門
- ☐ 教育學門
- ☐ 數學及統計學門

學類名稱

- ☐ 一般商業學類
- ☐ 一般大眾傳播學類
- ☐ 一般法律學類
- ☐ 一般農業學類
- ☐ 中國語文學類
- ☐ 人類學學類
- ☐ 企業管理學類
- ☐ 兒童保育學類

學校名稱

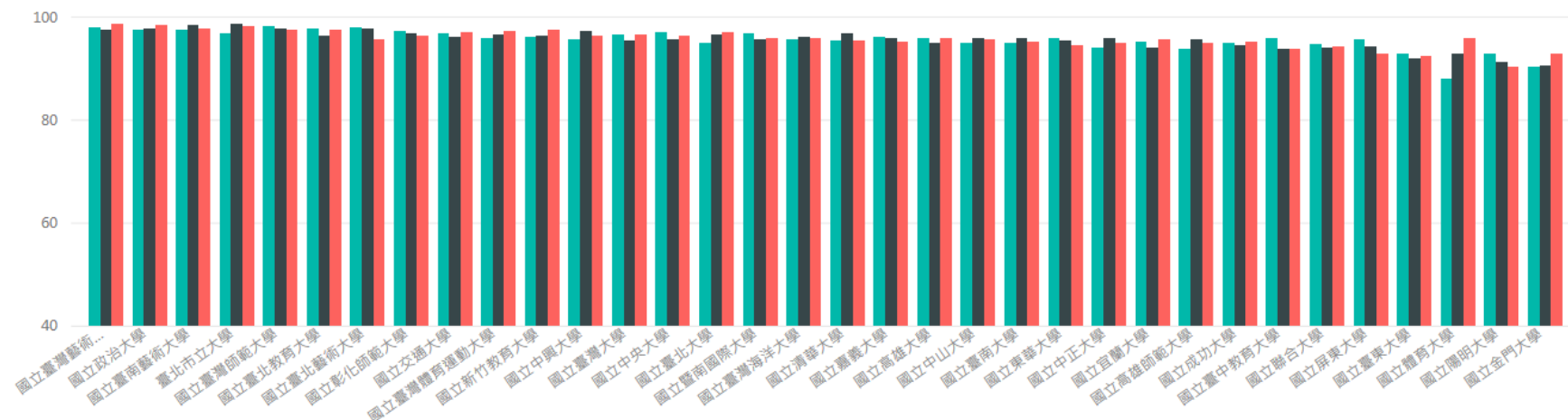
- ☐ 世新大學
- ☐ 中信金融管理學院
- ☐ 中原大學
- ☐ 中國文化大學
- ☐ 中國醫藥大學
- ☐ 中山醫學大學
- ☐ 中華大學
- ☐ 亞洲大學



103-105公立大學註冊率排名

各學年度各校系註冊率

學年度 ● 103 ● 104 ● 105



設立別

- ☒ 公立
☐ 私立

學年度

- ☐ 103
☐ 104
☐ 105

領域名稱

- ☐ 人文及藝術
☐ 其他
☐ 工程、製造及營造
☐ 教育
☐ 服務
☐ 社會科學、商業及法律
☐ 科學
☐ 農學
☐ 醫藥衛生及社福

學門名稱

- ☐ 人文學門
☐ 傳播學門
☐ 其他學門
☐ 商業及管理學門
☐ 工程學門
☐ 建築及都市規劃學門
☐ 教育學門
☐ 數學及統計學門
☐ 生命科學

學類名稱

- ☐ 一般大眾傳播學類
☐ 一般法律學類
☐ 一般農業學類
☐ 中國語文學類
☐ 人類學學類
☐ 企業管理學類
☐ 公共行政學類
☐ 公共衛生學類
☐ 其他學類

學校名稱

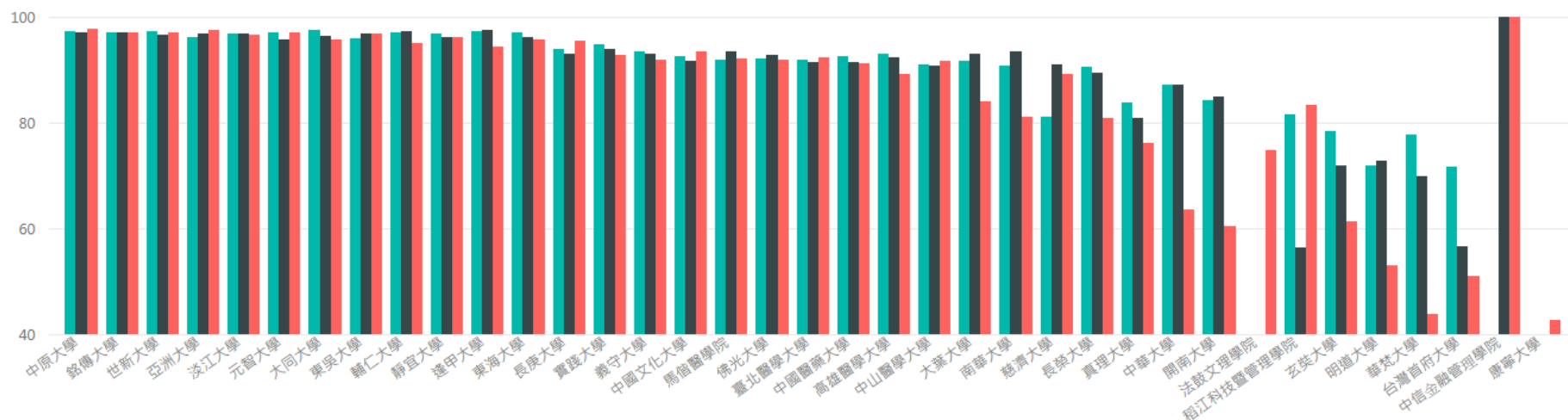
- ☐ 國立中央大學
☐ 國立中山大學
☐ 國立中正大學
☐ 國立中興大學
☐ 國立交通大學
☐ 國立嘉義大學
☐ 國立宜蘭大學
☐ 國立屏東大學
☐ 國立臺南大學
☐ 國立陽明大學
☐ 國立金門大學



103-105私立大學註冊率排名

各學年度各校系註冊率

學年度 ● 103 ● 104 ● 105



設立別

☐ 公立

☒ 私立

學年度

- ☐ 103
- ☐ 104
- ☐ 105

領域名稱

- ☐ 人文及藝術
- ☐ 工程、製造及營造
- ☐ 教育
- ☐ 服務
- ☐ 社會科學、商業及法律
- ☐ 科學
- ☐ 農學
- ☐ 醫藥衛生及社福

學門名稱

- ☐ 人文學門
- ☐ 傳播學門
- ☐ 商業及管理學門
- ☐ 工程學門
- ☐ 建築及都市規劃學門
- ☐ 教育學門
- ☐ 數學及統計學門
- ☐ 民生學門

學類名稱

- ☐ 一般商業學類
- ☐ 一般大眾傳播學類
- ☐ 一般法律學類
- ☐ 一般農業學類
- ☐ 中國語文學類
- ☐ 人類學學類
- ☐ 企業管理學類
- ☐ 兒童保育學類

學校名稱

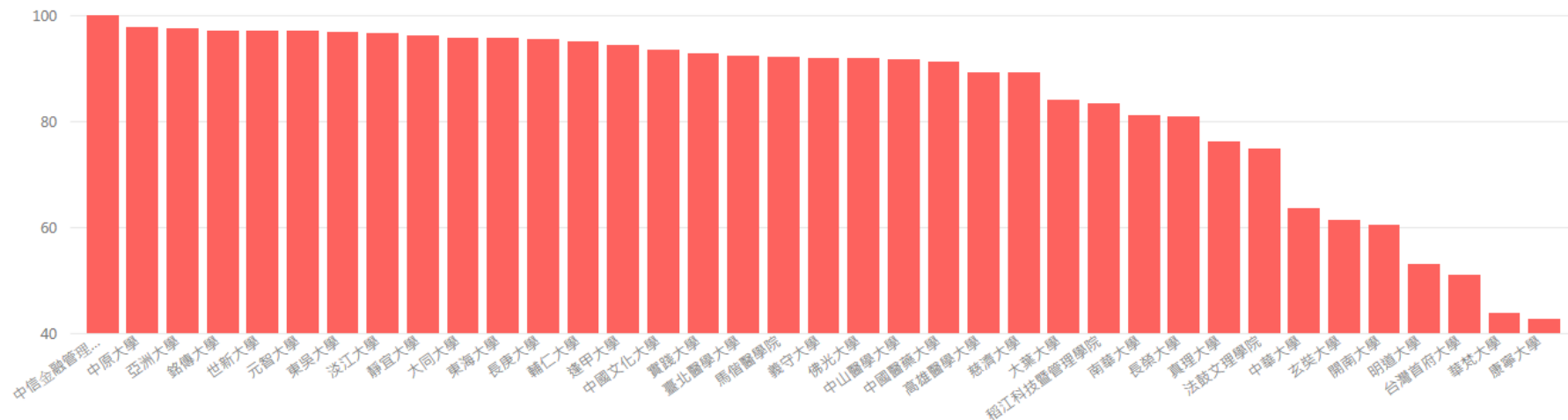
- ☐ 世新大學
- ☐ 中信金融管理學院
- ☐ 中原大學
- ☐ 中國文化大學
- ☐ 中國醫藥大學
- ☐ 中山醫學大學
- ☐ 中華大學
- ☐ 亞洲大學



105私立大學註冊率排名

各學年度各校系註冊率

學年度 ● 105



設立別

- ☐ 公立
☒ 私立

學年度

- ☐ 103
☐ 104
☒ 105

領域名稱

- ☐ 人文及藝術
☐ 工程、製造及營造
☐ 教育
☐ 服務
☐ 社會科學、商業及法律
☐ 科學
☐ 農學
☐ 醫藥衛生及社福

學門名稱

- ☐ 人文學門
☐ 傳播學門
☐ 商業及管理學門
☐ 工程學門
☐ 建築及都市規劃學門
☐ 教育學門
☐ 數學及統計學門
☐ 民生學門

學類名稱

- ☐ 一般商業學類
☐ 一般大眾傳播學類
☐ 一般法律學類
☐ 一般農業學類
☐ 中國語文學類
☐ 企業管理學類
☐ 兒童保育學類
☐ 公共行政學類

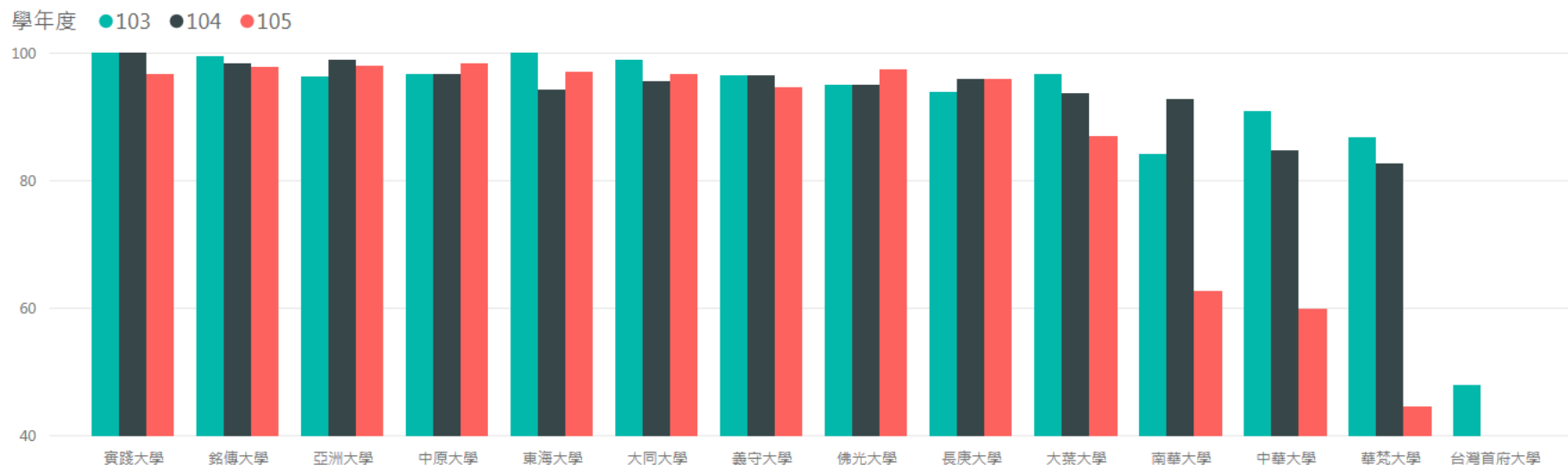
學校名稱

- ☐ 世新大學
☐ 中信金融管理學院
☐ 中原大學
☐ 中國文化大學
☐ 中國醫藥大學
☐ 中山醫學大學
☐ 中華大學
☐ 亞洲大學



103-105私立產品設計學類排名

各學年度各校系註冊率



設立別

- ☐ 公立
☒ 私立

領域名稱

- ☒ 人文及藝術

學門名稱

- ☒ 設計學門

學類名稱

- ☒ 產品設計學類
☐ 空間設計學類
☐ 綜合設計學類
☐ 視覺傳達設計學類

學校名稱

- ☐ 中原大學
☐ 中華大學
☐ 亞洲大學
☐ 佛光大學
☐ 南華大學
☐ 台灣首府大學
☐ 大同大學
☐ 大葉大學

學年度

- ☐ 103
☐ 104
☐ 105

Diagnostic Analytics

名詞

- 診斷分析、特徵分析

釋義

- 發現或確定事情為何發生

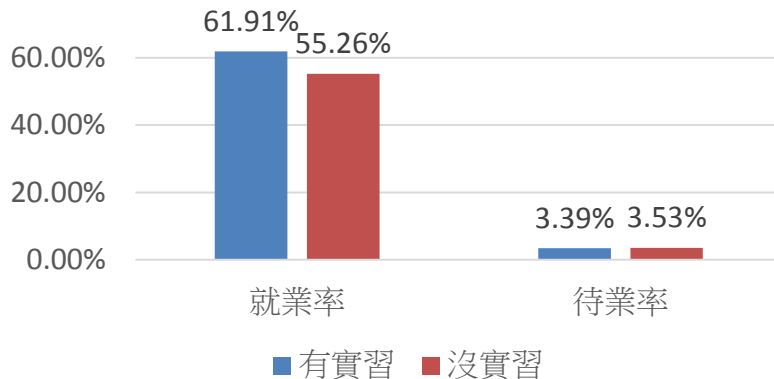
方法

- 使用統計分析工具，了解資料關聯性
- 相關性分析、迴歸分析
- SAS、SPSS、R

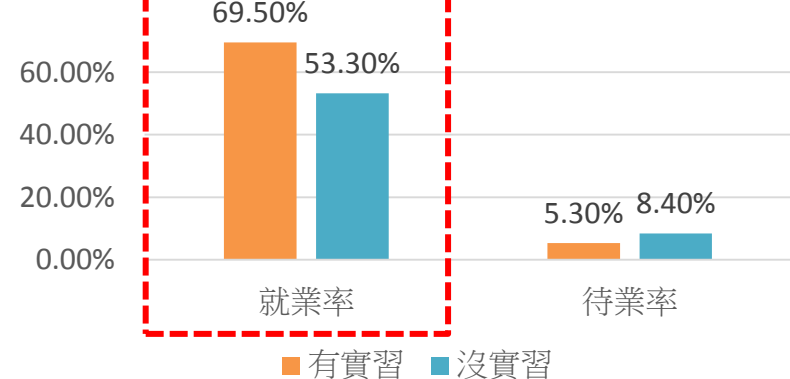


實習與就業率分析

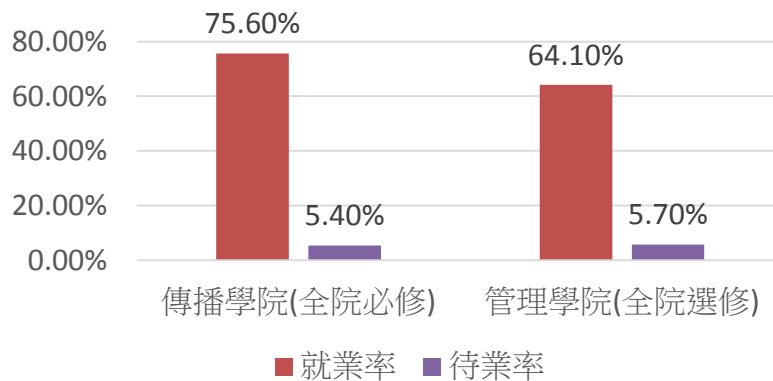
畢業後一年調查與實習關係



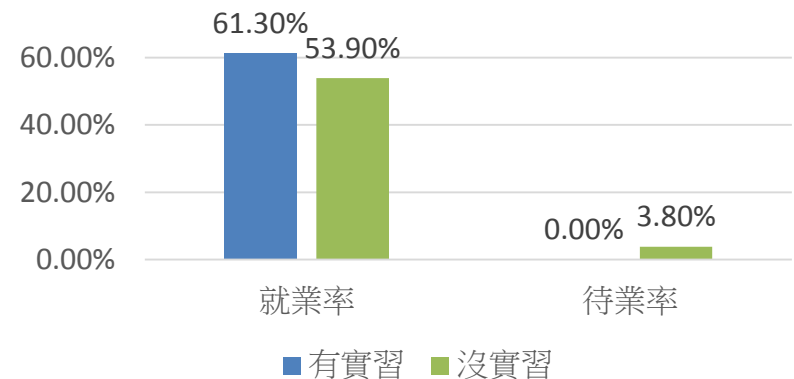
大學部就業率與待業率



學院之就業率與待業率



學系之就業率與待業率-公事



上述就業率及待業率之計算不含升學、服役和其他比率



Predictive Analytics

名詞

- 預測分析

釋義

- 估計將會發生什麼。會根據歷史數據，結合限制條件和因素，評估最大可能發生的結果

方法

- 使用資料探勘工具，預測未來狀況
- Weka、SAS、SPSS



多益考試與校內EPT

資料蒐集

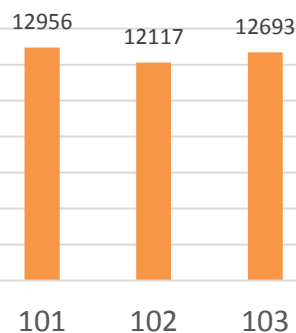
建立模型

驗證模型

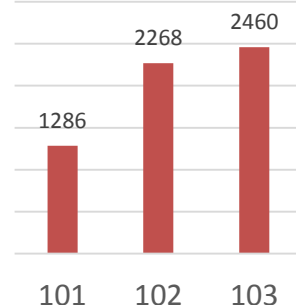
進行預測

結果分析

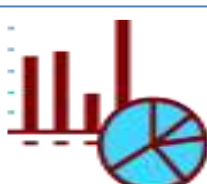
EPT資料筆數



多益資料筆數



複迴歸

二元
複迴歸二元二次
迴歸

參數估計

殘差分析

模型評估

模型修正

TOEIC

人數

500-525

747

525-550

685

550-575

538

575-600

375

600-625

172

625-650

85

650-990

25

推薦
632位TOEIC500 ↑
455位預測成功率
72%

數位教學分析：尋找輔導標的

10301-05141論文研討 (一)

姓名	學號	學分	成績	備註	備註	備註	備註
10301111 陳建宏	261	1	3	31	55	54	1
10301122 謝建宏	189	0	3	24	12	36	3
10301133 謝建宏	171	1	3	26	29	15	1
10301144 謝建宏	203	1	3	23	21	40	1
10301155 謝建宏	126	0	3	14	17	32	3
10301166 謝建宏	85	0	3	18	11	36	1
10301177 謝建宏	149	0	3	9	18	11	1
10301188 謝建宏	199	0	3	21	29	22	1
10301199 謝建宏	162						
10301210 謝建宏	112						

報表

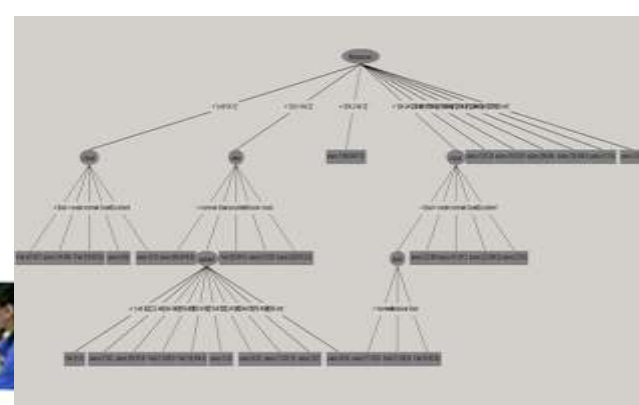
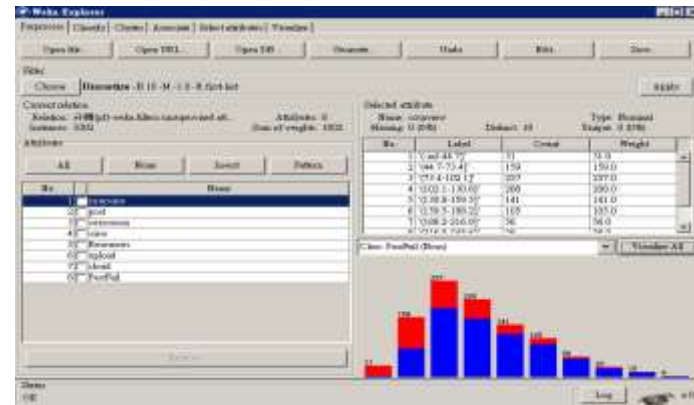
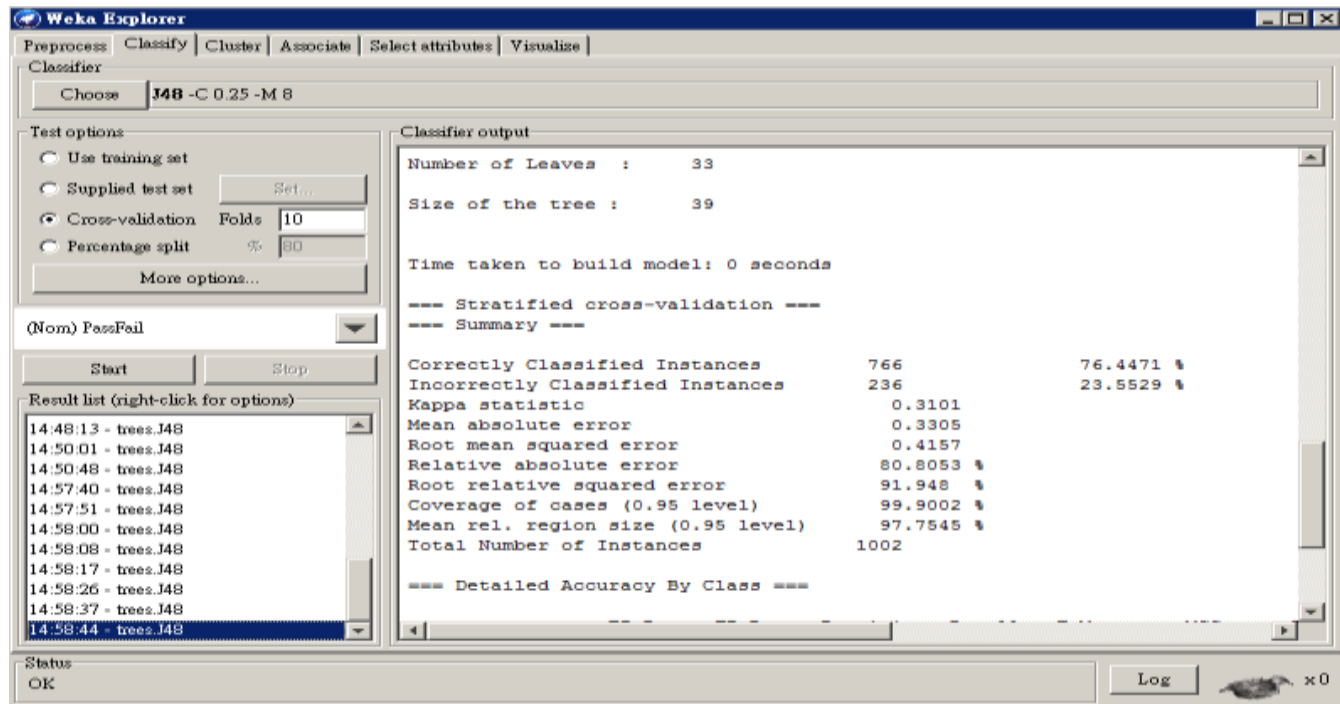


全校統計

全校統計 (比例)	全校統計 (比例)
已修課人數 (比例)	1210
已修課人數 (比例)	45186
已修課人數 (比例)	2389
已修課人數 (比例)	25412
已修課人數 (比例)	2131
已修課人數 (比例)	7714
已修課人數 (比例)	1107

WEB

Classification : J48



Prescriptive Analytics

名詞

- 指導性分析、處方分析、規範性分析、時效性分析

釋義

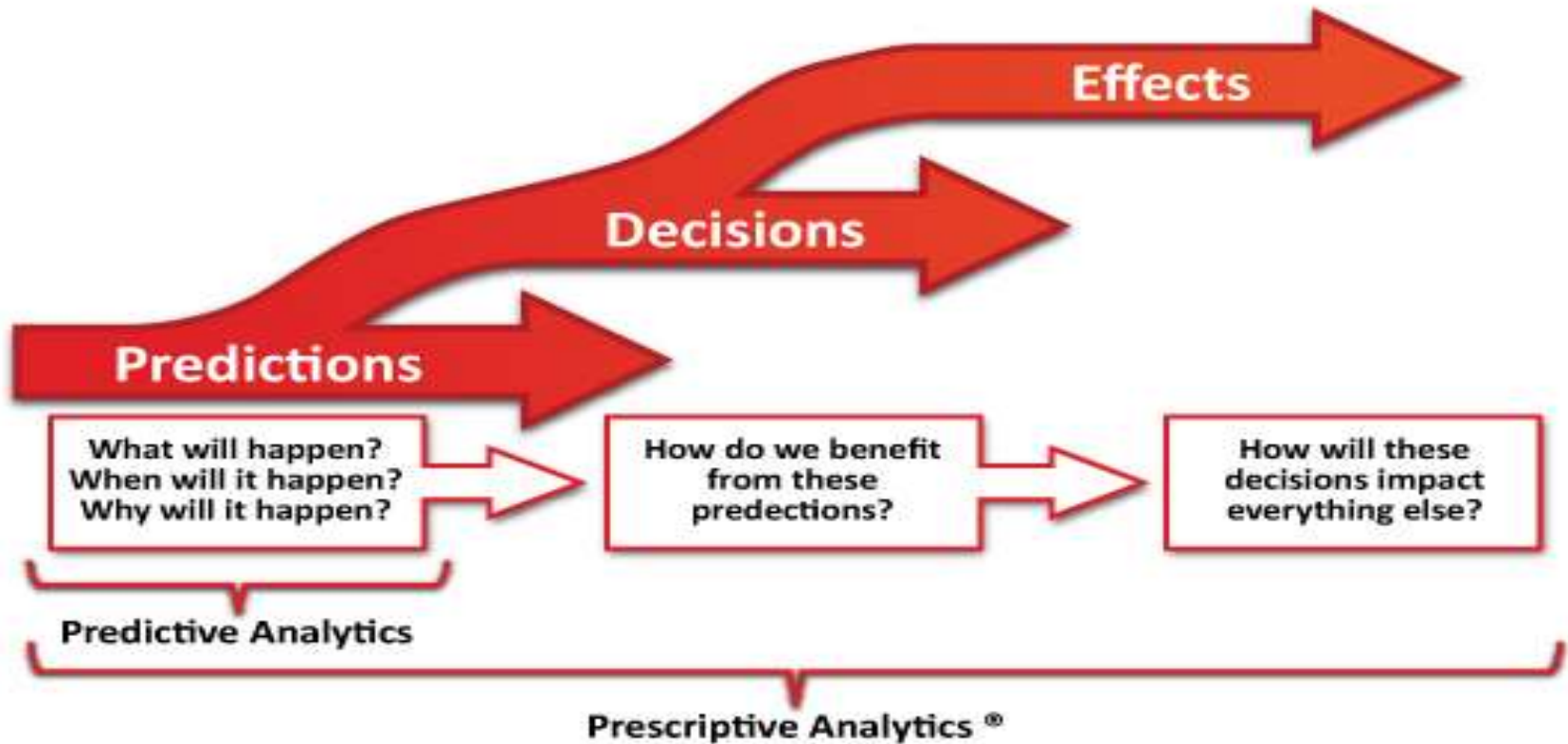
- 根據預測分析的結果，總結及建議不同結果的優化行動

方法

- 透過預測分析結果，進行決策
- 人工智慧、類神經網絡、決策支援



Prescriptive Analytics



Prescriptive Analytics extends beyond predictive analytics by specifying both the actions necessary to achieve predicted outcomes, and the interrelated effects of each decision

指導性分析(範例:交通疏運計畫)





商業智慧

Business Intelligent



校務研究與商業智慧（BI）

- 商業智慧（Business Intelligence，簡稱BI），指用現代數據倉儲技術、線上分析處理技術、數據挖掘和數據展現技術進行數據分析以實現商業價值。
- 商業智慧作為一個工具，用來處理學校中現有數據，並將其轉換成知識、分析和結論，輔助主管做出正確且明智的決定，在推動校務發展時可有憑依據。



商業智慧

定義

用現代數據倉庫技術、
線上分析處理技術、
數據挖掘和數據展現
技術進行數據分析以
實現商業價值

工具種類

資料處理

☐ Power BI☐ Qlik View☐ Tableau

資料分析

☐ SPSS☐ SAS☐ R

資料倉儲

☐ Oracle☐ MS SQL Server☐ Teradata AsterData

資料探勘

☐ WEKA☐ SAS☐ MS SQL

商業智慧的優點

- 加快資料的收集、處理、分析流程：分析企業數字背後的意義
- 使用簡單：不需要懂 IT 的艱深技術
- 方便快捷：匯整能力
- 量化的策略分析及決策支援：有效降低企業風險與營運成本(任意角度進行資料分析、比對、預測)



資料分析

線上即時分析 (OLAP)

- 採用多維度之資料結構將資料載入，以進行多項不同維度整合的分析

統計分析

- 類別分析、多因子變異數、相關係數、迴歸分析

資料探勘

- 利用統計、人工智慧 (AI) 或其他的分析技術，在巨量歷史資料內深度尋找與發掘未知、隱藏性的關係與規則，從而達成分類(classification)、推估(estimation)、預測(prediction)、關聯分組(affinity grouping)、同質分組(clustering) 等結果





結果呈現

報表

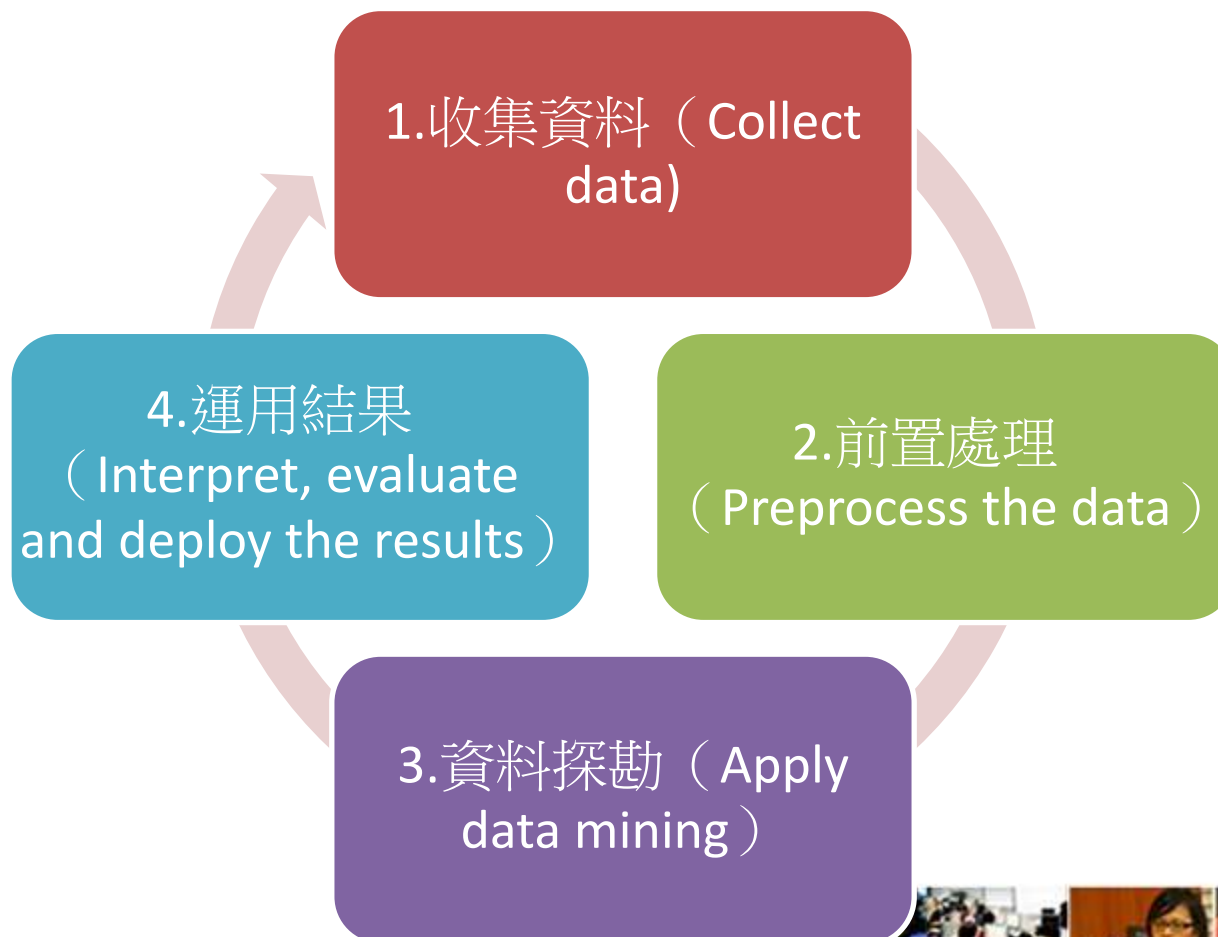


教育的資料探勘(EDM)

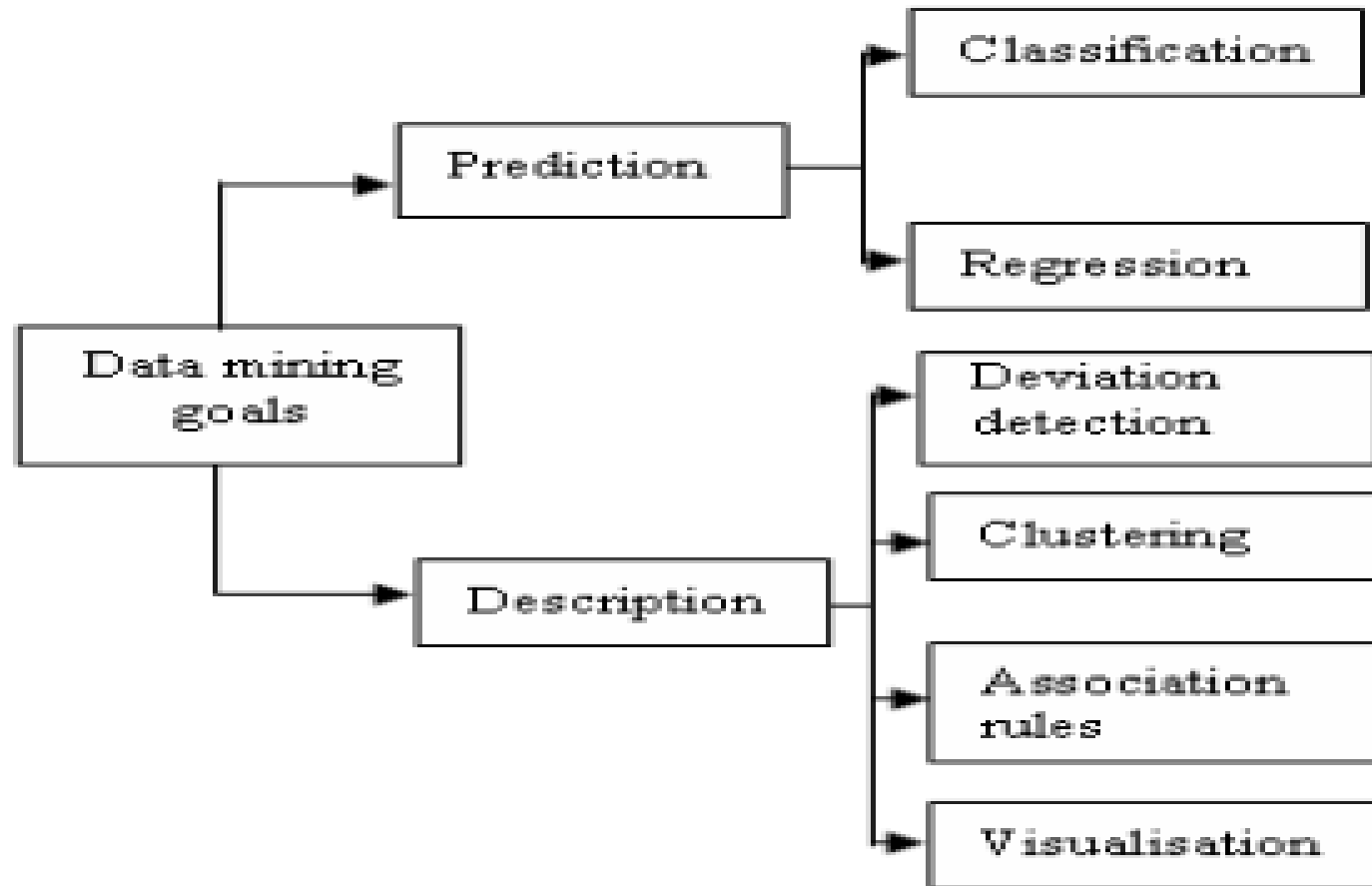
教育大數據分析



資料探勘的四個步驟

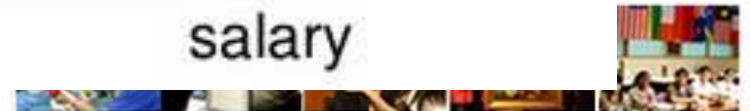
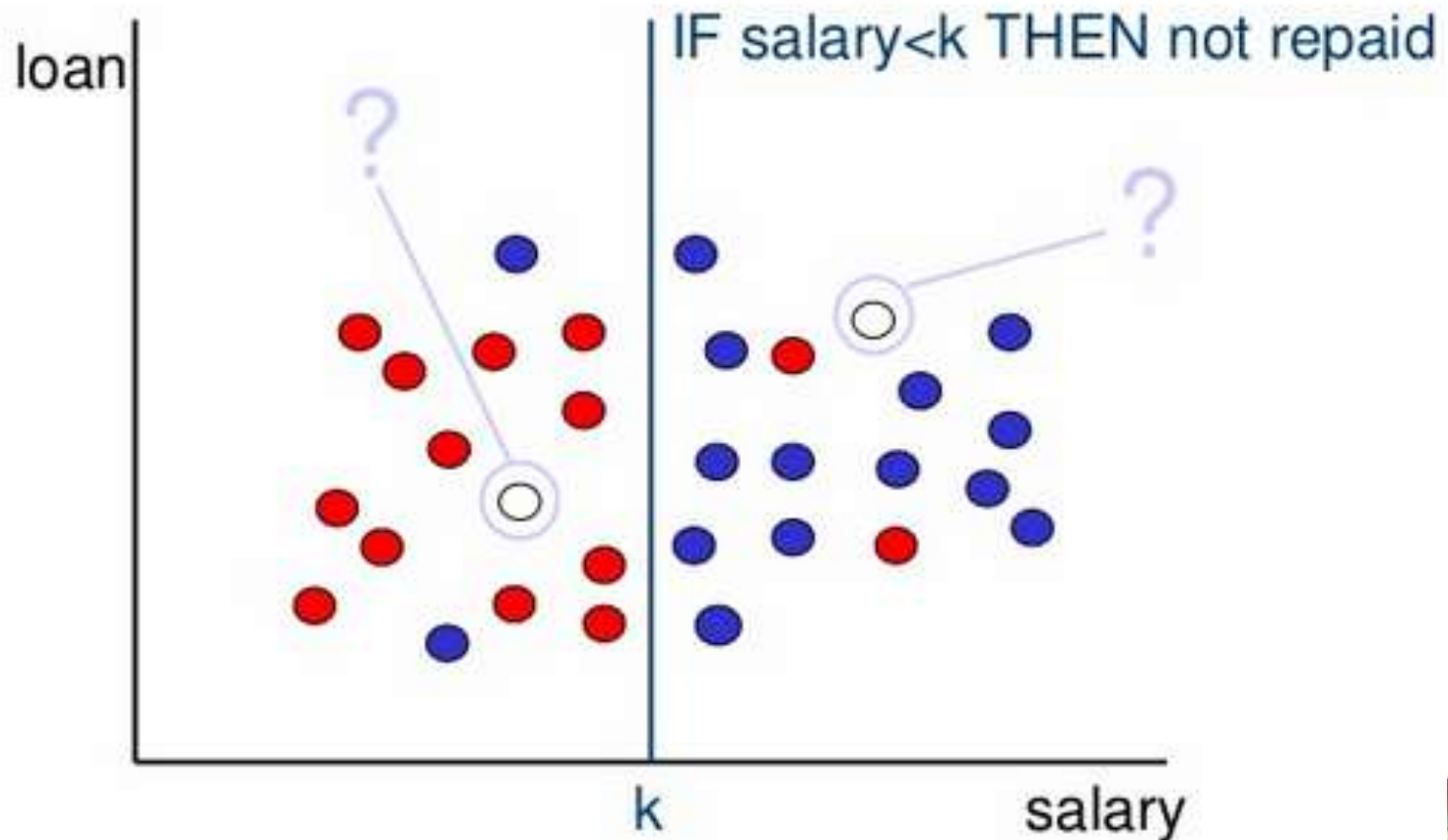


Data Mining Goals



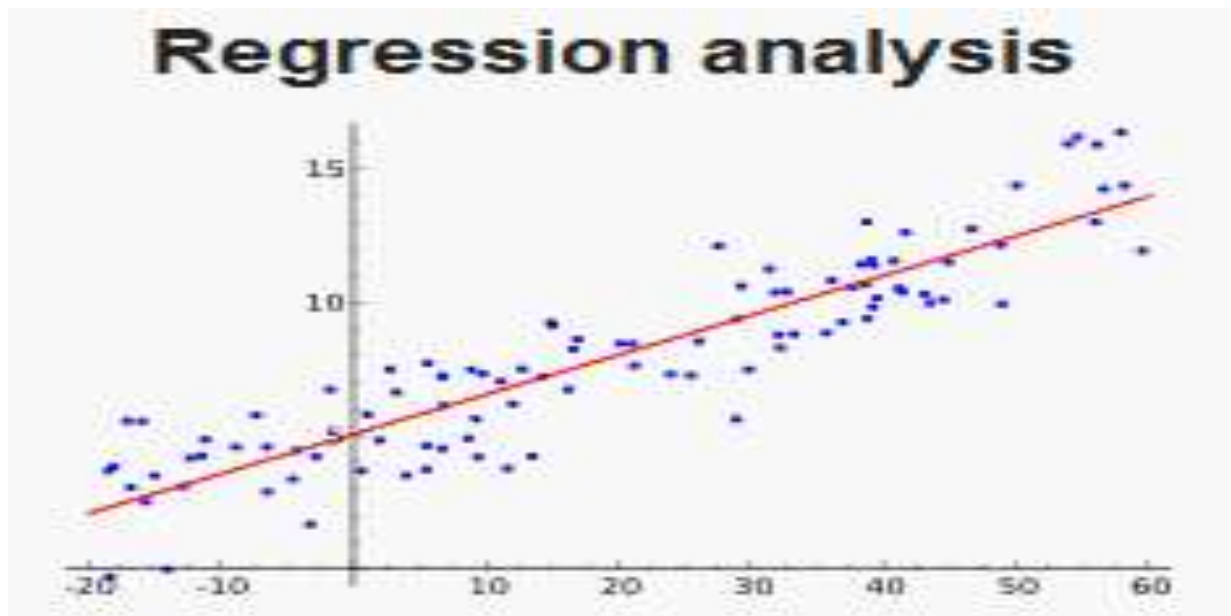
Danubianu M., Pentiuc St. Gh., Socaciu T. (2009). Towards the optimized personalized therapy of speech disorders by data mining techniques, The Fourth International Multi Conference on Computing in the Global Information Technology ICCGI 2009, Vol: CD, 23-29 August, Cannes - La Bocca, Franc

Classification



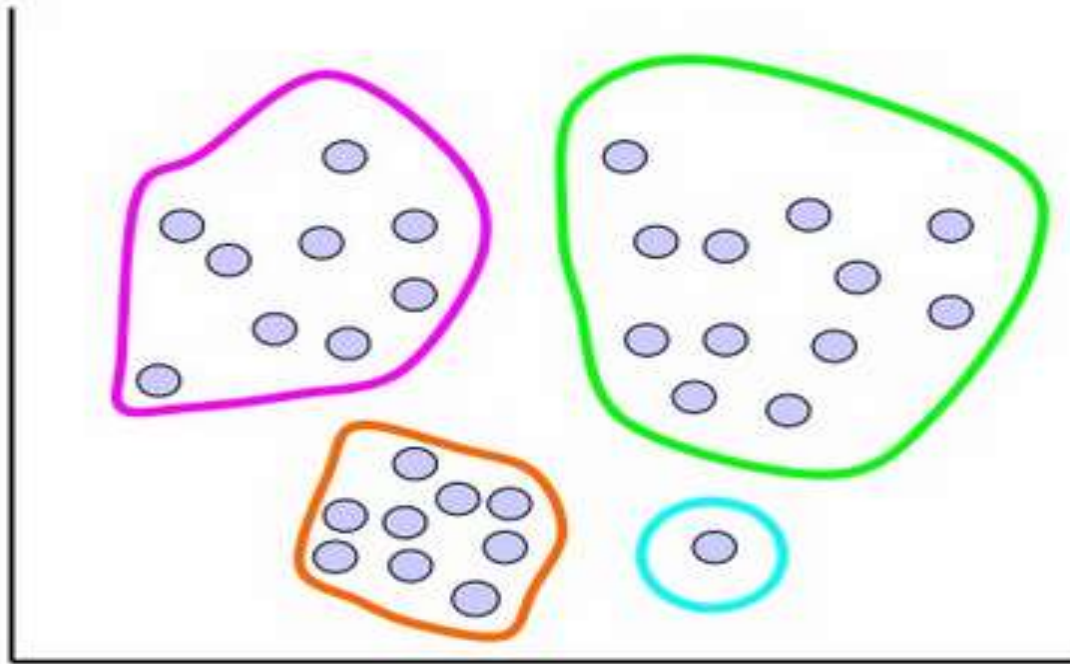
Regression

- Regression is used in predicting values of a dependent variable by estimating the relationship among variables using statistical analysis



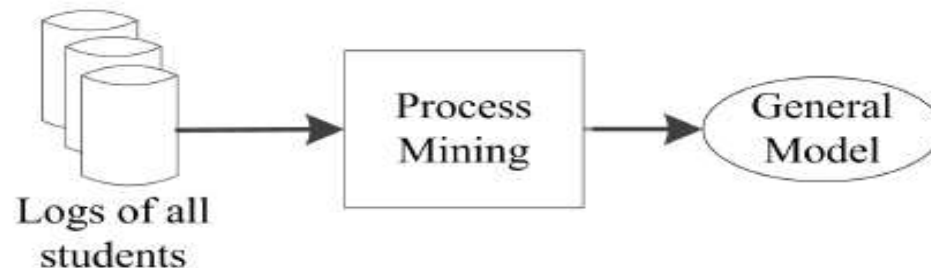
分群 (Clustering)

- 分群分析是資料探勘的方法之一，將比較相似的樣本叢集在一起，形成集群

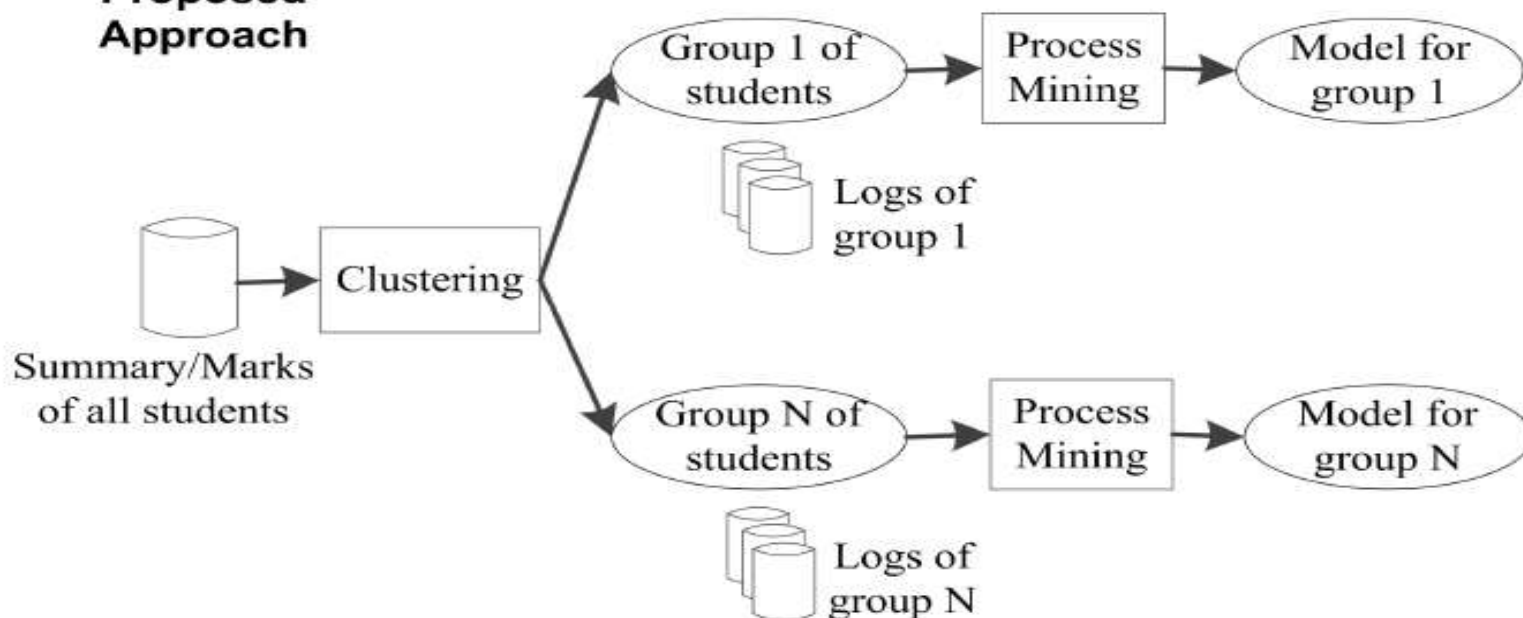


分群->建模

Traditional Approach



Proposed Approach



Association Rules

- Given a set of transactions, find rules that will predict the occurrence of an item based on the occurrences of other items in the transaction

Market-Basket transactions

<i>TID</i>	<i>Items</i>
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

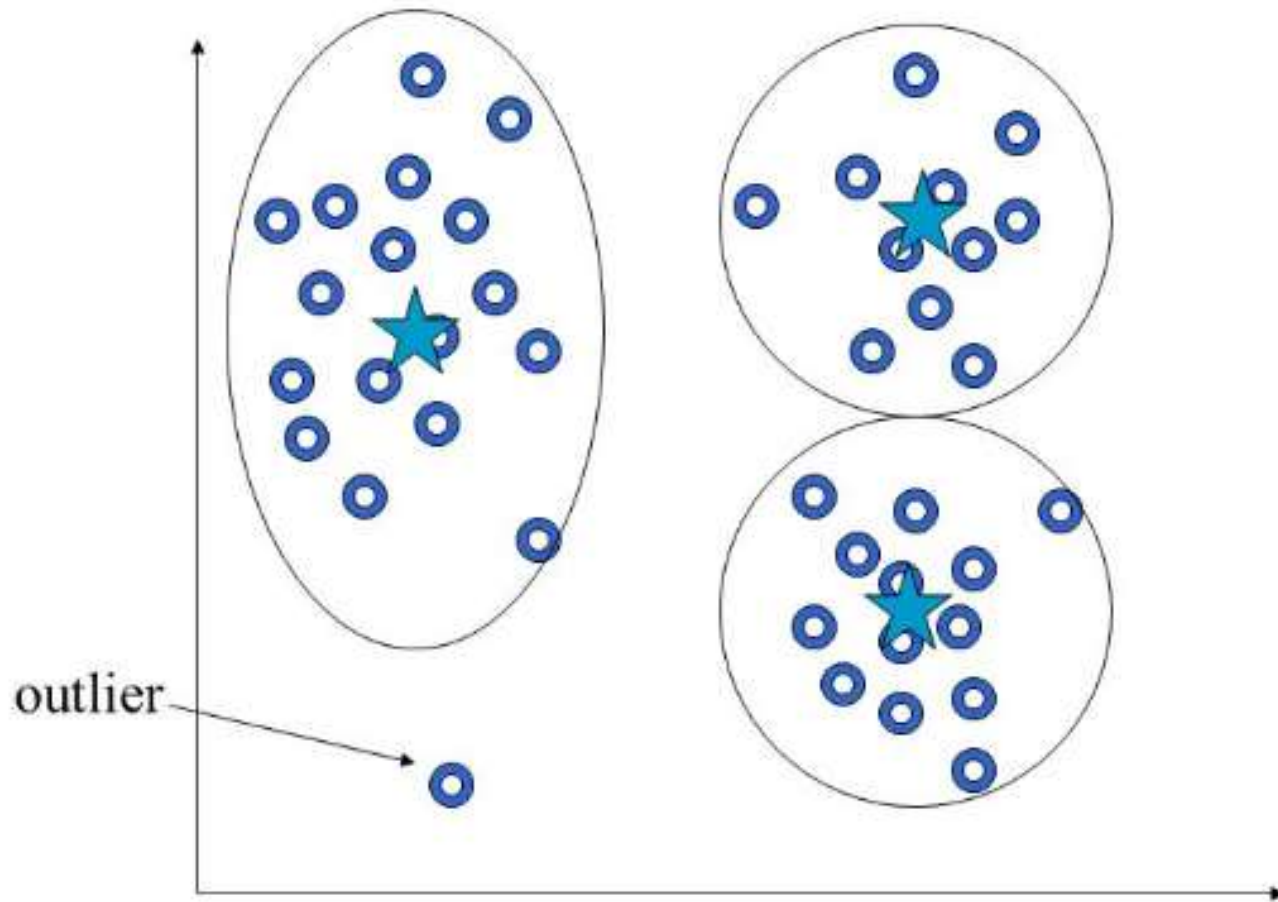
Example of Association Rules

$\{\text{Diaper}\} \rightarrow \{\text{Beer}\},$
 $\{\text{Milk, Bread}\} \rightarrow \{\text{Eggs, Coke}\},$
 $\{\text{Beer, Bread}\} \rightarrow \{\text{Milk}\},$

Implication means co-occurrence,
not causality!

Deviation Detection

- Find unexpected values
- Outlier





研究議題

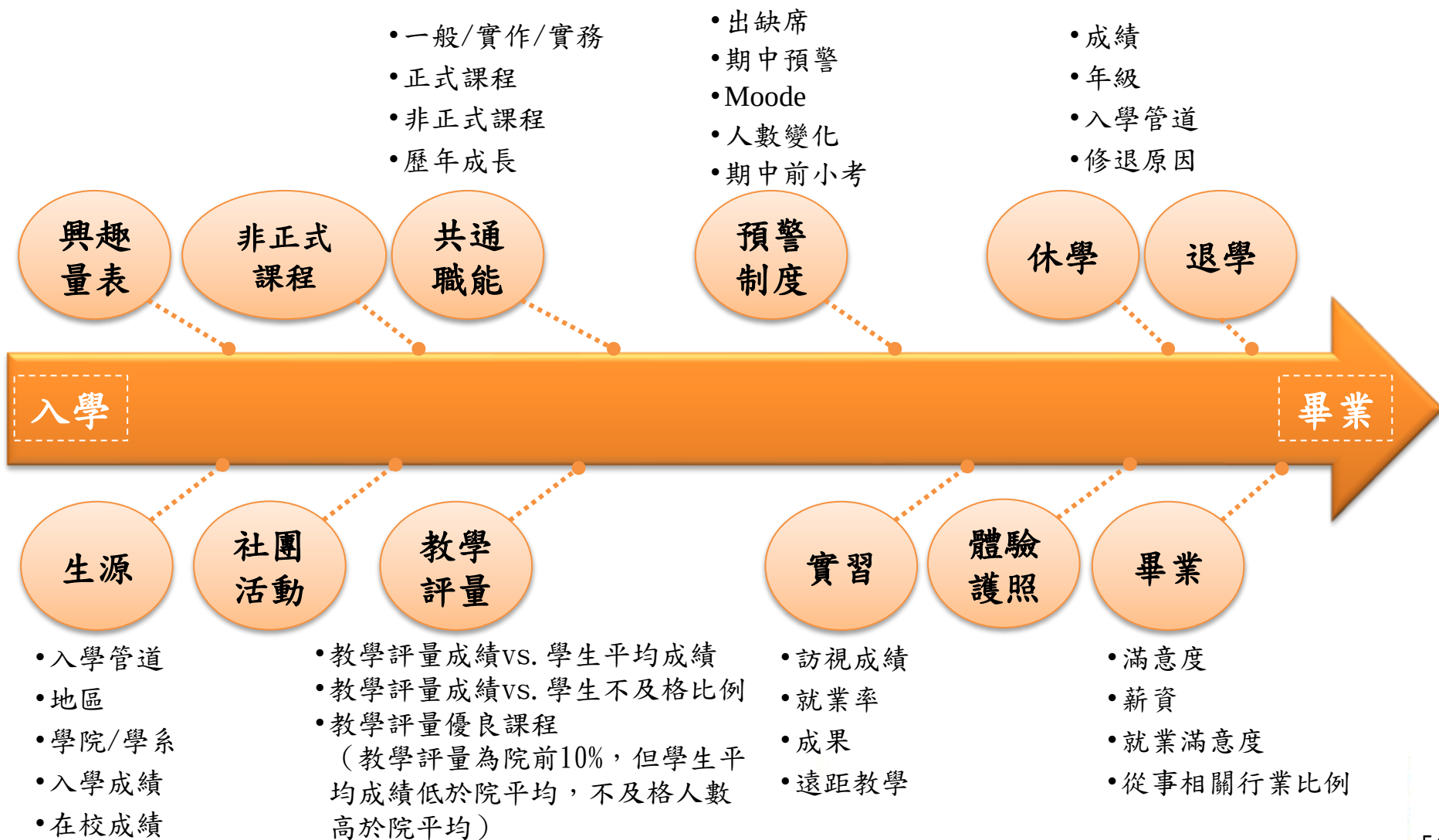


校務研究議題面向

- 1) 學校教學輔導系統與學生學習成效分析系統之整合
- 2) 校務資源分配與學生學習成效分析系統之整合
- 3) 教師升等制度與學生學習成效分析系統之整合
- 4) 依學習成效數據分析提供學生個別化學習經驗及主動輔導之措施
- 5) 大學招生策略與新生特質分析
- 6) 少子女化趨勢對大學永續經營之策略分析
- 7) 增加學生續讀率或降低學生休退學率之分析
- 8) 畢業生在校學習成效與就業職場關聯性分析
- 9) 教師教學成效分析與學生成績



校務研究議題和相關因子





研究案例

教學平台



網路教學平台

- 學習管理系統(LMS) ，以**知識管理的概念原理配合資訊化科技的應用**，將素材、教學單元、主題教學網頁三階段的媒體納入管理的系統，它可以說是一個網路教學和知識管理的混合體。
- 主要的功能需包含**教學參與者管理**、**課程內容管理**、記錄課程**參與者上課記錄**，及**評量學習成效**，並將相關學習過程的資料匯集分成各報表。
- TronClass、Moodle、Blackboard、WebCT、Oracle iLearning



優質的教與學

- 完整的教學管理

教學

- 便利的師生互動

互動

檔案

- 充沛的檔案資源

評量

- 多元的學習評量

系統

穩定的系統運作

行政

整合的行政系統

資料探勘軟體Weka

- [Weka](#)是以Java為基礎的資料探勘(Data mining)與(Machine learning)軟體，也是自由軟體(Open Source Software)。
- 紐西蘭的懷卡託大學(The University of Waikato)
- Weka整合了大量的資料探勘的演算法
- Windows x86、Windows x64、Mac OS X、Linux



「資訊科技：辦公室應用課程」

- 蒐集學生參與課程教學活動的數據
 - 將課程檢視、發表文章、檢視討論區、檢視文章、線上資源、上傳檔案
 - 加上進入銘傳大學雲端教室操作練習題的次數
 - 對1,002位學生參與TQC Word技能檢定的成績通過或不通過 (pass/fail)
 - 利用J48決策樹演算法進行建模預測



目的





Confusion Matrix

Prediction Actual	Positive	Negative
	True Positives (TP)	False Negatives (FN)
Positive		
Negative	False Positives (FP)	True Negatives (TN)

Precision	proportion of instances that are truly of a class divided by the total instances classified as that class, i.e. Precision of Positive = $TP/(TP+FP)$ and Precision of Negative = $TN/(FN+TN)$.
Recall	proportion of instances classified as a given class divided by the actual total in that class (equivalent to TP rate), i.e., Recall of Positive = $TP/(TP+FN)$ and Recall of Negative = $TN/(FP+TN)$.



結果

- 數據的執行結果，預測的準確率可達七成左右，的確可以作為學習預警的目標，期望作為學生核心重要課程的學習監控。
- 活躍的課程：預測的準確度
- 舊學期->新學期：教學方式一致性
- 教師對數位教學平台有更多的投入和使用



感想

- 內容：去蕪存菁--先結構資料，後非結構資料
- 因果：因果非所想像
- 需求：系統需求不確定→先工具，後系統
 - 因素分析、關聯分析、迴歸分析、資料探勘
- 系統：
 - Cloud Computing：DW + Web + Platform
 - 分析：Excel + SPSS + DataMing + Visualization
- 美麗的結果：艱辛的過程



結語

- 領域知識 + 統計能力 + 資訊能力
- 不同學校具備不同的資料完整度
- 商業智慧工具有效剖析校務資訊
- 校務研究是校務決策的重要依據

Data is King





簡報完畢，敬請指教！

