

# 人氣教授養成指南

## — 成為人人推薦的教授全攻略

---

### UrSchool系所教授評價分析

組名：我叫陳伶均請多多指教

組員：杜 昕 陳伶均 廖韋凱 陳柏宇

# 專案流程簡介

## 網路爬蟲

使用 Python 爬取 UrSchool 臺大教授的個人推薦分數、同學評價等資料，並存成 Excel 檔。

## 量化分析

使用 R 語言分析每位教授的推薦分數與所屬系所、口條、課業壓力、給分甜度等因素的關係。

## 語意分析

使用 R 語言分析同學給教授的評價留言，分析私人推薦分數高和私人推薦分數低的教授留言中，留言內容的差異。

# 專案流程簡介

## 網路爬蟲

使用 Python 爬取 UrSchool 臺大教授的個人推薦分數、同學評價等資料，並存成 Excel 檔。

## 量化分析

使用 R 語言分析每位教授的推薦分數與所屬系所、口條、課業壓力、給分甜度等因素的關係。

## 語意分析

使用 R 語言分析同學給教授的評價留言，分析私人推薦分數高和私人推薦分數低的教授留言中，留言內容的差異。

# UrSchool 教授基本資料

國立臺灣大學的教授清單：共收錄3177位教授

國立臺灣大學好嗎？如何選課？高中生、大學生如何選填科系、研究所？來這裡查看系所教授評價

« 1 2 3 4 5 6 7 8 ... 317 318 »

姓名	授課系所	職稱	私心推薦	學到東西	口條好	課業壓力	給分甜度	平均成績	最高學歷	值得一提	綽號	點名方式
林明仁	經濟學系	教授兼系主任	5.0 ★	5.0 ★	5.0 ★	4.0 ★	4.0 ★	A	美國芝加哥大學經濟學博士	芝加哥學派，台灣當代應用個體的代表人物	史上最帥系主任 MJ	加簽同學接受隨機抽點，其他自己選上的不點名
駱明慶	經濟學系	教授	4.0 ★	4.0 ★	4.0 ★	3.0 ★	4.0 ★	A	美國密西根大學經濟學博士	精神不集中會很容易睡著，錯過課程精華.憤世嫉俗		偶爾照點名單念一遍，原則上完全放任
孫維新	物理學系	教授	5.0 ★	5.0 ★	5.0 ★	4.0 ★	4.5 ★		美國加州大學洛杉磯分校（UCLA）天文學博士	2000 金鐘獎 教科文節目主持人獎		
蘇彩足	政治學系	教授	5.0 ★	5.0 ★	5.0 ★	3.0 ★	4.0 ★	A	美國卡納基美隆大學公共政策博士			

爬取基本資料  
進行量化分析

# UrSchool 同學留言評價



最帥系主任 沒有之一 – 4 年前



現在教學：經濟學原理、應用個體經濟學、資料科學、金融實務課程 除了超強個人魅力與讀到教學功力，是個讓人無可挑剔的教授 / 主任；但是想要上主任越深入的課程，就請投入對應的時間與精力把實力建立起來！

– 4 年前



很棒的一堂經濟學原理!!不過難度可以再稍微增加~~ – 3 年前

爬取文字內容  
進行語意分析

# 資料前處理

- 統一授課系所、職稱名稱，並刪去重複的教授資料
- 增加「百分制平均成績」、「學院」欄位

姓名	授課系所	職稱	私心推薦	學到東西	口條好	課業壓力	給分甜度	平均成績	最高學歷	值得一提	綽號	點名方式	調整後平均成績	學院
林明仁	經濟學系	教授	5	5	5	4	4.0	A	美國芝加哥大學經濟學博士	芝加哥學派，台灣當代應用個體的代表人物	史上最帥系主任 MJ	加簽同學接受隨機抽點，其他自己選上的不點名	87	社會科學院
駱明慶	經濟學系	教授	4	4	4	3	4.0	A	美國密西根大學經濟學博士(	精神不集中會很容易睡著，錯過課程精華。憤世嫉俗	NA	偶爾照點名單念一遍，原則上完全放任	87	社會科學院
孫維新	物理學系	教授	5	5	5	4	4.5	NA	美國加州大學洛杉磯分校 (UCLA) 天文學博士	2000 金鐘獎教科文節目主持人獎	NA	NA	NA	理學院

# 專案流程簡介

## 網路爬蟲

使用 Python 爬取 UrSchool 臺大教授的個人推薦分數、同學評價等資料，並存成 Excel 檔。

## 量化分析

使用 R 語言分析每位教授的推薦分數與所屬系所、口條、課業壓力、給分甜度等因素的關係。

## 語意分析

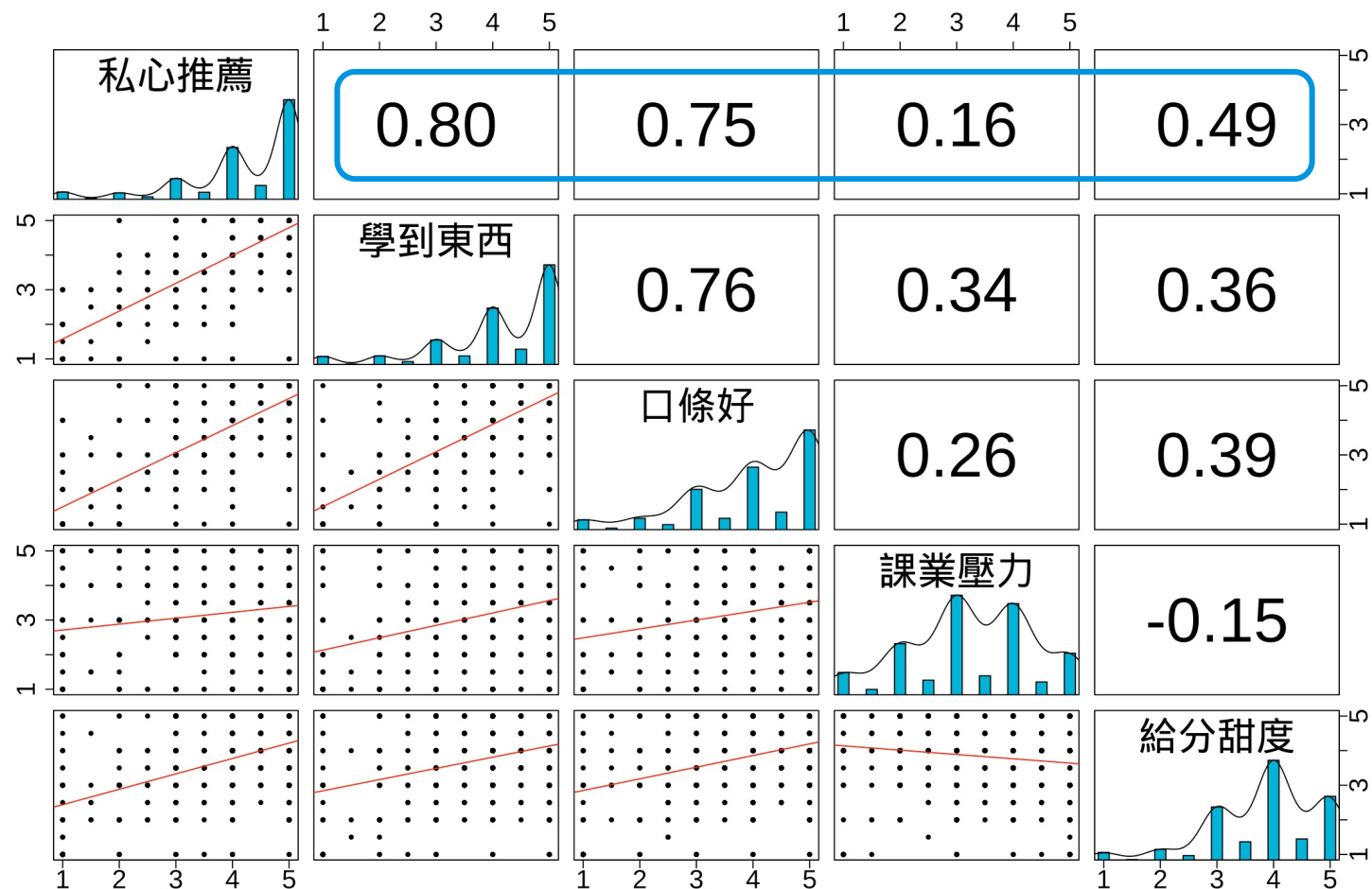
使用 R 語言分析同學給教授的評價留言，分析私人推薦分數高和私人推薦分數低的教授留言中，留言內容的差異。

# 各學院各項指標平均分數

學院	教授資料數	平均 私心推薦	平均 學到東西	平均 口條好	平均 課業壓力	平均 給分甜度
文學院	271	4.47	4.32	4.21	3.21	3.96
理學院	169	4.10	4.03	3.90	3.41	3.78
社會科學院	130	4.24	4.12	3.94	3.05	3.72
醫學院	112	4.00	4.21	3.96	3.41	3.67
工學院	154	4.19	4.14	3.94	3.29	3.96
生農學院	177	3.80	3.75	3.68	3.11	3.72
管理學院	118	4.19	4.15	3.99	3.12	3.94
公衛學院	16	4.25	3.91	3.78	3.25	3.69
電資學院	111	4.29	4.25	4.06	<b>3.59</b>	3.93
法律學院	48	<b>4.53</b>	<b>4.56</b>	<b>4.55</b>	3.47	<b>4.23</b>
生命科學院	53	3.79	3.95	3.73	3.11	3.58
共同教育中心	1	4.00	4.00	4.00	4.00	3.00
其他	3	4.50	4.67	4.67	4.50	3.67



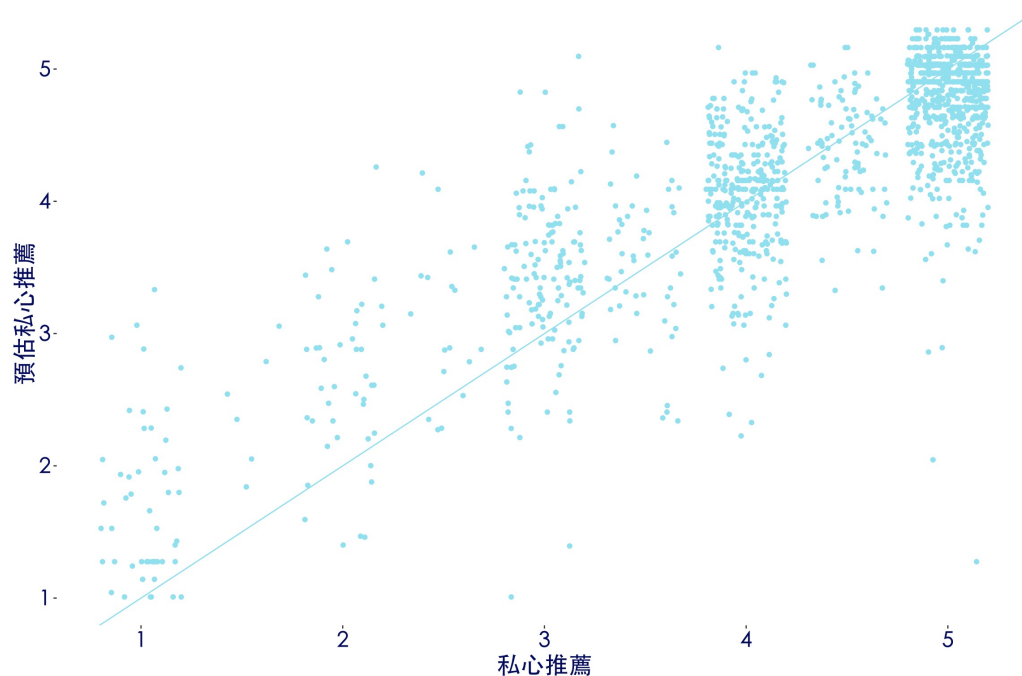
# 各項指標與私心推薦分數的相關性



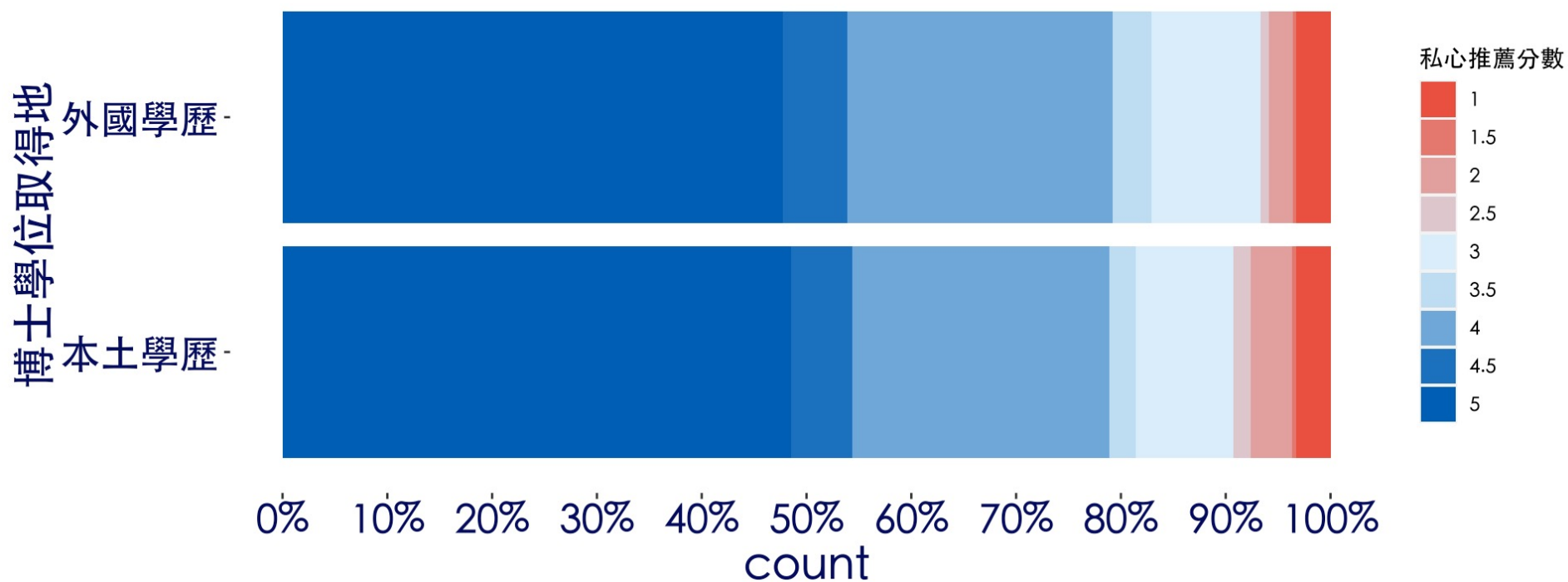
# 使用迴歸模型預測推薦分數

lm ( 私心推薦 ~ 學到東西 + 口條好 + 課業壓力 + 給分甜度 )

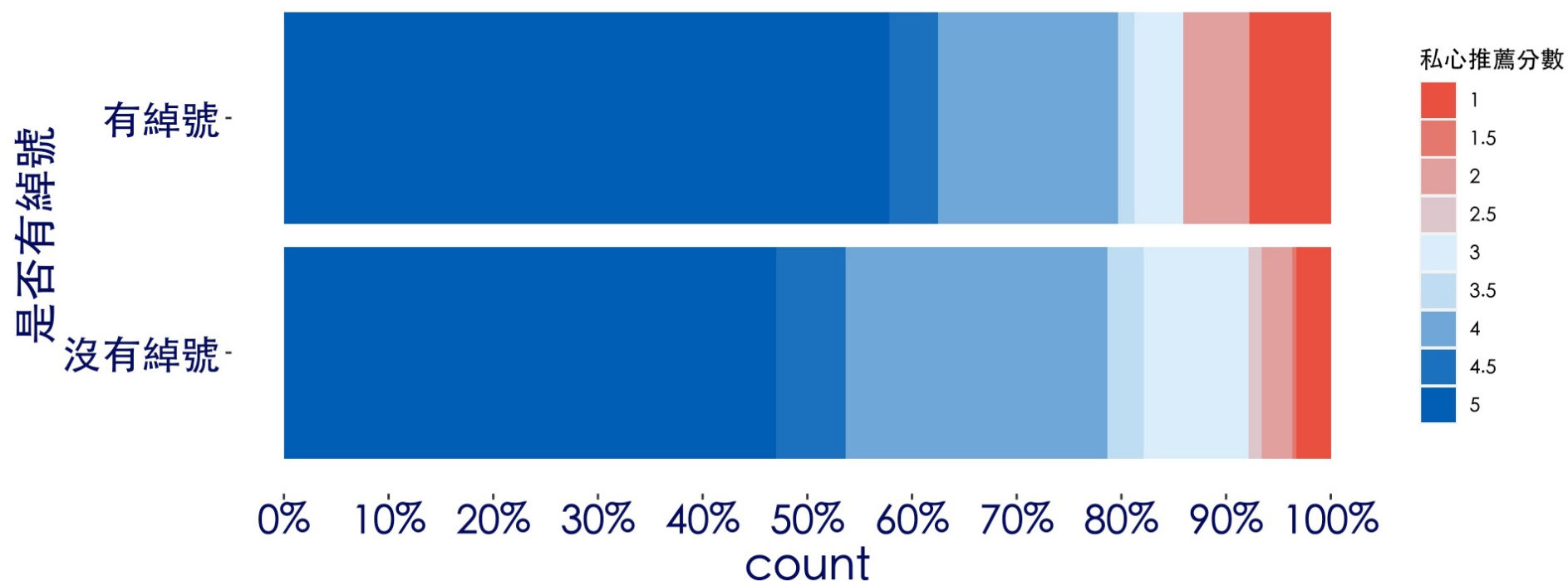
→ 預測私心推薦分數 =  $0.33 + 0.54 \text{學到東西} + 0.27 \text{口條好}$   
 $- 0.07 \text{課業壓力} + 0.19 \text{給分甜度}$



# 博士學位取得地與私心推薦分數的關聯性小



# 有綽號是否與私心推薦分數有些許關聯



# 專案流程簡介

## 網路爬蟲

使用 Python 爬取 UrSchool 臺大教授的個人推薦分數、同學評價等資料，並存成 Excel 檔。

## 量化分析

使用 R 語言分析每位教授的推薦分數與所屬系所、口條、課業壓力、給分甜度等因素的關係。

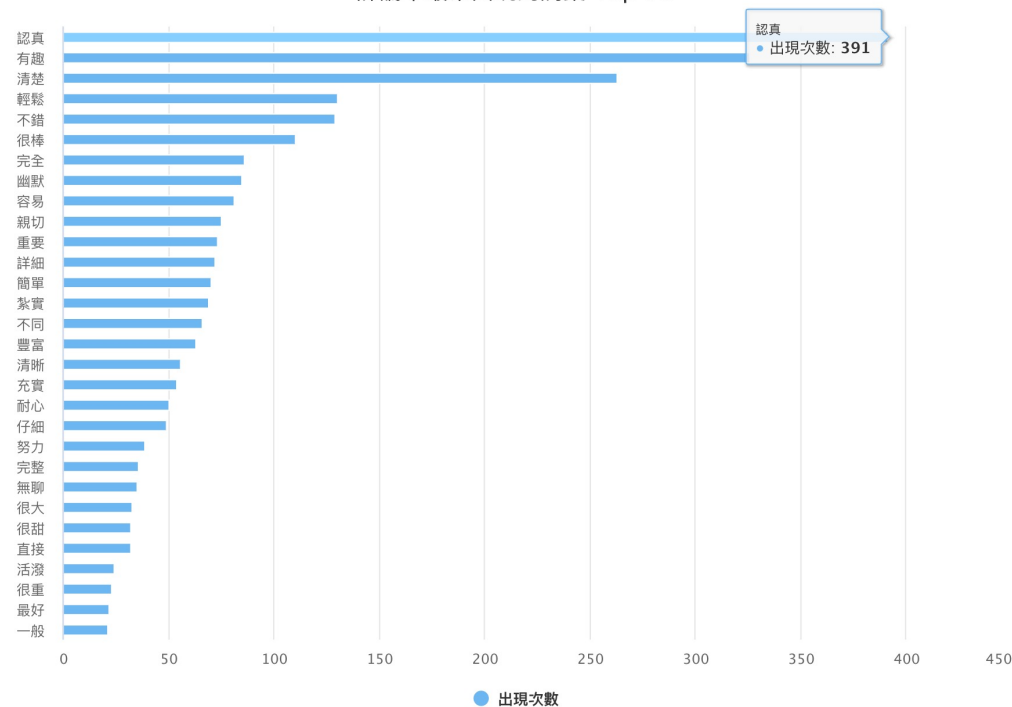
## 語意分析

使用 R 語言分析同學給教授的評價留言，分析私人推薦分數高和私人推薦分數低的教授留言中，留言內容的差異。

# 使用 jieba 找出評論中最常出現的形容詞



評論中最常出現的詞彙 Top 30

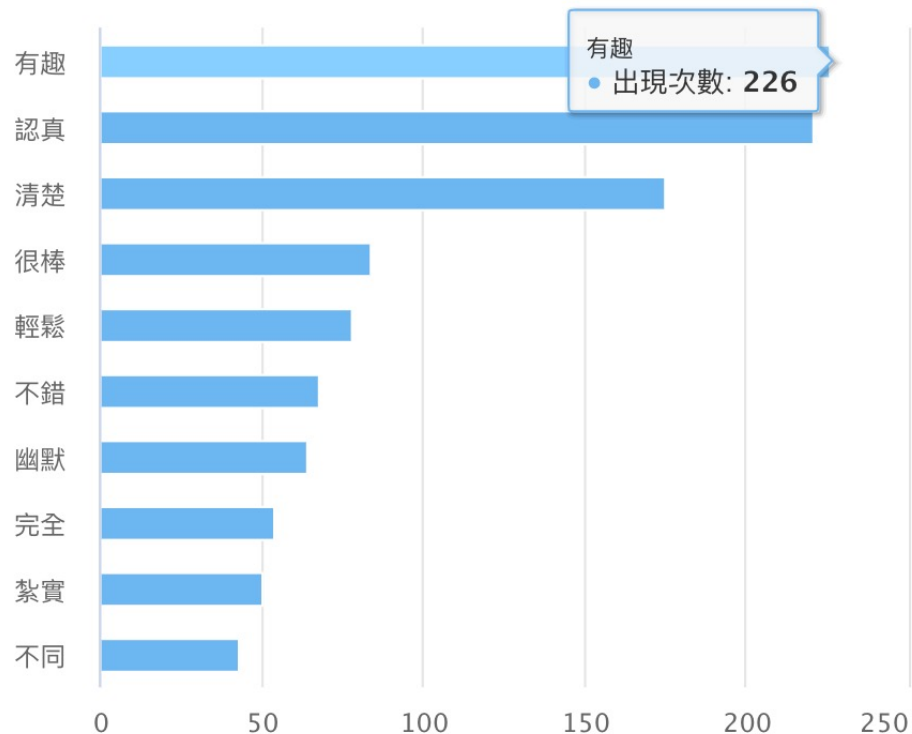


# 各個學院教授最常形容詞：認真、清楚、有趣

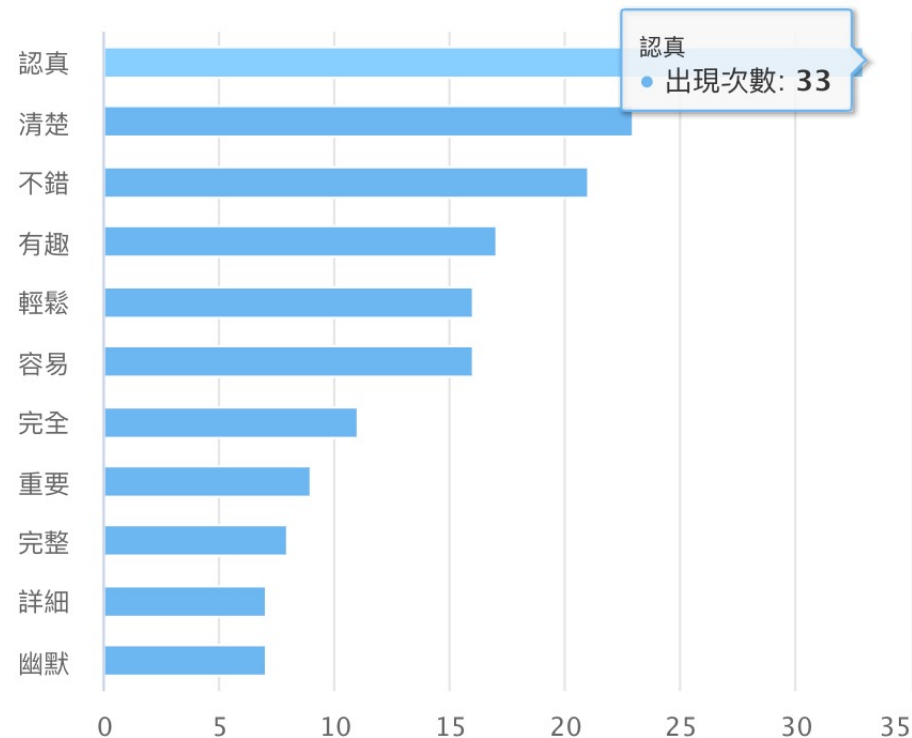
學院	最常形容	出現次數	學院	最常形容	出現次數
文學院	有趣	120	管理學院	認真	33
理學院	認真	75	公衛學院	認真	33
社會科學院	清楚	51	電資學院	無聊	2
醫學院	清楚	18	法律學院	清楚	30
工學院	認真	43	生命科學院	清楚	13
生農學院	認真	42	共同教育中心	有趣	8

# 高評價、低評價教授常出現形容詞無顯著差異

高評價教授評價最常出現的詞彙 Top 10



低評價教授評價最常出現的詞彙 Top 10



前10名共同詞彙：有趣、認真、清楚、輕鬆、不錯、完全



# 使用RMarkdown進行靜態資料呈現

## 1 專案背景與資料簡介

### 1.1 專案背景

### 1.2 專案流程

### 1.3 資料簡介

#### 1.3.1 讀入資料

#### 1.3.2 資料前處理

## 2 量化分析

### 2.1 各系所各項指標分析

### 2.2 各項指標與私心推薦指數的相關性

### 2.3 使用迴歸模型預測推薦指數

### 2.4 其他可能影響私心推薦指數的指標

#### 2.4.1 土博士 v.s. 洋墨水

#### 2.4.2 教授有綽號是否代表與學生比較喜歡？

## 3 語意分析

### 3.1 學生最重視什麼？

### 3.2 學生最重視什麼？（各個學院）

### 3.3 高評價教授 v.s. 低評價教授

## 4 總結

## 人氣教授養成指南 — 成為人人推薦的教授全攻略

### UrSchool系所教授評價分析

組名：我叫陳伶均請多多指教

組員：杜昕 陳伶均 廖韋凱 陳柏宇

## 1 專案背景與資料簡介

### 1.1 專案背景

還記得幾年前剛升上大學，拿到一張空白的課表，人生第一次要由自己來填上。焦急地查閱各種網路文章、評論留言、評價分數……一位位素昧平生的教授就這樣被化成一項項量化的分數，我們或多或少參考了這些分數，默默地過完了好一段大學歲月。

想想自己有沒有可能因此而與某些知識領域擦身而過，或是不小心「誤入歧途」，得要逼自己習慣不適合自己的授課方式？

從這個身為大學生揮之不去的煩惱出發，我們想知道這些對於教授、課堂的評論與分數到底是怎麼一回事。為此我們選了「Urschool大學科系教授評價網」作為研究對象，目標是找到一系列對於教授的評分與文字評價當中的端倪。

21世紀是個「線上時代」，每個人都是一個巨量數據的集合。

不論大至攸關國家社稷，小至謹謹生活瑣事，當我們面對一個陌生的問題時，第一個反應往往是上網查資料。然而透過區區幾十吋的電腦螢幕，我們能在瞬間獲得千百筆資料，新聞媒體、社群平台、電商官網……承載資料的網站五花八門，而在其間四處流竄的資料更是琳瑯滿目，在資料科學與視覺化技術的加持下，在我們眼前以各種形式呈現。

# 感謝觀看

---

杜 昕 陳伶均 廖韋凱 陳柏宇

# 附錄：分工表

組員	負責內容	投入程度
杜昕	資料前處理、量化分析、網頁架設、投影片製作	6
陳伶均	網路爬蟲、語意分析	6
廖韋凱	網路爬蟲、語意分析	6
陳柏宇	量化分析、網頁文字撰寫	6