

1. **情境一**：國內媒體 readr 經常會將半自動化蒐集來的人工驗證資料當作開放資料釋出。例如他們與政治大學選舉研究中心都曾經在選舉期間蒐集街頭的選舉看板，來估計不同的候選人可能花費多少錢在選舉看板上，或做後續資料分析利用。所以他們會設計一套網頁服務，邀請公眾或自家媒體從業人員參與，當在路上看到看板時，就把他拍照拍下來上傳，然後由媒體內部人員來做後續的核對，並對圖像內容做編碼。這種 Crowd Sourcing 的任務是誰拍下來、誰編碼、怎麼編碼都非常重要。

資料庫規劃 (25%)：作為一個有計算機相關能力的媒體從業人員，你被指派要設計一個資料庫來提供這個服務。因為要邀請公眾參與，所以必須要紀錄參與者的資訊；所紀錄的看板資訊，也應該有核對人員做檢核。請用實體關係圖 (Entity-Relationship-Diagram) 設計一個資料庫來支援上述情境的資料。請仔細考慮在這樣的情境下，你的資料庫可能要有哪些變項欄位？應拆分出哪些資料表？並提供一筆假想的範例資料作為參考。

2. **前後端實作 (25%)**：承上題**情境一**，從資料庫設計、網頁設計，你可能會用到哪些服務、資料庫、軟體、程式套件（也可以都採用雲端服務）。請畫一個架構圖說明，你會怎麼設計這套系統，甚至要如何將這樣的工具以 API 的方式提供有興趣的大眾存取？請盡可能地描繪軟硬體需求。

3. **Crowd-sourcing (10%)**：上述這種情境的做法稱為 Crowd-sourcing，也是資料新聞蒐集資料的方式之一。就過去所涉獵的資料新聞而言，你還有看過哪些是屬於這種類型的新聞？請嘗試描述其資料取得和結果。如果你沒有印象的話，就請構思一個透過 Crowd-sourcing 來產製資料新聞的點子，包含新聞的主題，資料蒐集方式，可能的解釋變項等。

4. **下標偵測 (20%)**：「假新聞」或所謂的「假消息」事實上意義非常模糊，有新聞媒體報錯新聞這種假新聞，也有政黨的政治宣傳 (Propaganda)、也有 LINE 上面傳的長輩貼文、小道消息，也有中共網軍、大外宣、也有誘導你觀看，斷章取義、標題和內文不一樣的「誘餌標題」。針對誘餌標題或所謂的標題黨，現在甚至有媒體會聘請工讀生來調整標題，以獲得比較好的點閱。據此情境請回答下列兩個問題：

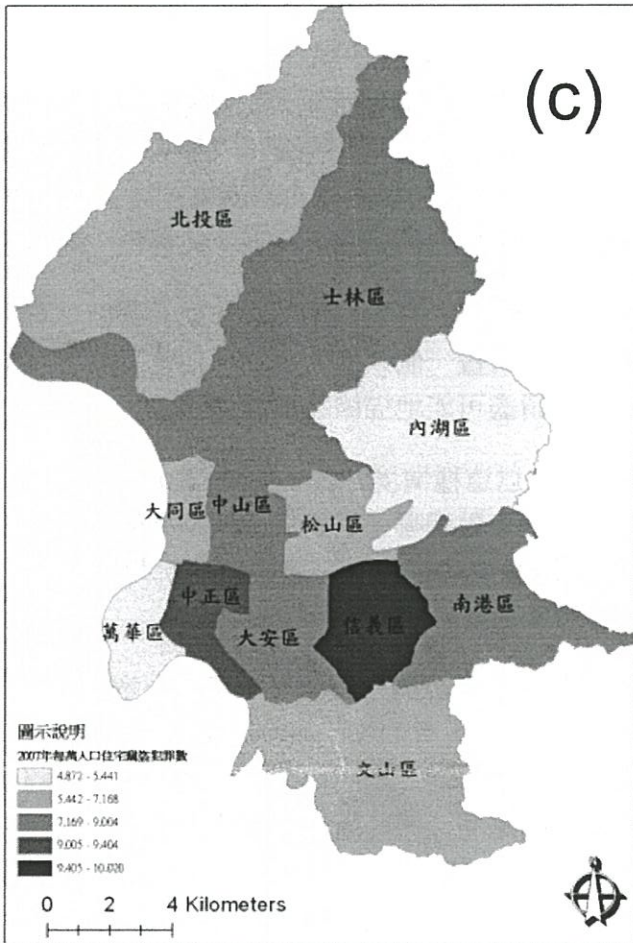
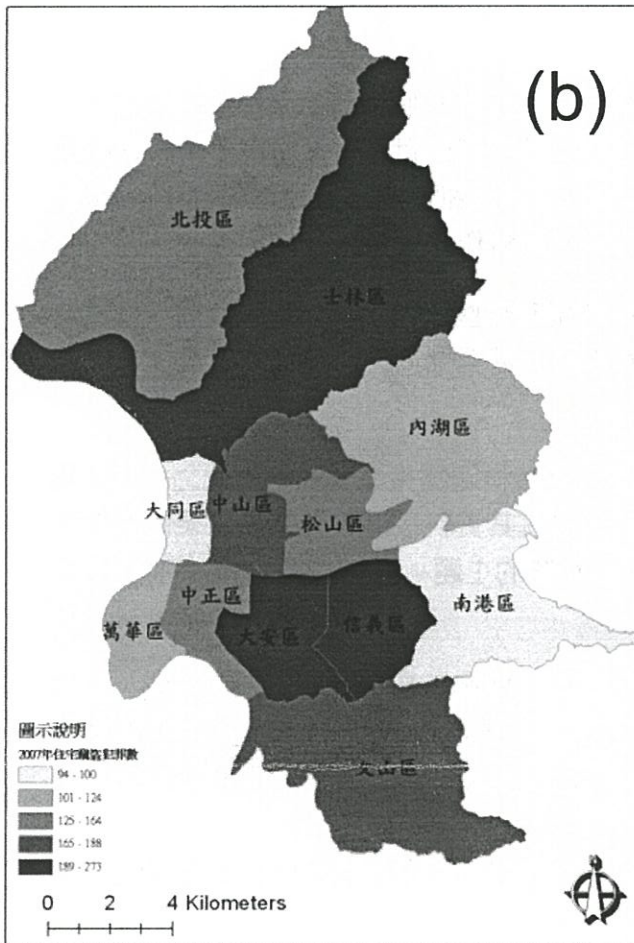
- A. 因為您要報考新聞所，可假設您對新聞的觀察有一定的程度。那麼，如果要去調新聞標題和內文的差異性，你認為現在線上標題的書寫可能會有哪些與新聞內文不一致或一致的情形？
- B. 如果要偵測新聞標題和內文的主題是否一致，從技術上來說，請繪製流程圖來說明，你要如何做比對？除此之外，由於標題和內文長度相差很多，請也在流程圖上說明，你可以怎麼解決這樣的問題。

5. **資料視覺化 (20%)**：下圖為「住宅竊盜點位資訊」匯聚成各行政區案件量的兩種視覺化方法，左側以案件量來繪圖 (方法一)、右側以案件量除以各行政區人口數來繪圖 (方法二)，而也可考慮將案件量除以行政區面積 (方法三)。請問就住宅竊盜點位資訊而言，哪一種方法相較而言最適合？為什麼？請問就近期 COVID-19 初期傳染而言，你認為哪一種方法較適合？為什麼？請問有哪一種開放資料案例可能會適合方法三？為什麼？

見背面

Counts

Risk = counts / pop-at-risk



試題隨卷繳回