

R 語言 期末報告說明

王貿

國立臺灣大學行為與資料科學研究中心助理研究員

國立臺灣大學政治學系博士、兼任講師

maowang01@gmail.com

課程主題重點

- 期末報告內容說明
- 可使用的語料介紹
- 請填寫「分組表」與「初步構想」（如果有的話）
（ <https://reurl.cc/V5OGkA> ）



期末報告內容

- 「口頭」(10%) 以及「書面」(25%) 報告。
- **主題**：自選 (需有能力自行取得文字資料)
- **分組**：1-3人
- **內容**：**明確的研究問題**、文字資料蒐集方式、分析方法、研究發現、研究結論與建議等部分。(文獻部分可以只呈現在書面報告)
- 最後兩週 (1/18, 1/25) 上臺報告。
- 書面報告請於**1/22中午前**繳交。

高階程度者

- 自行爬蟲。
- 手邊已有自行整理的其他文字資料檔。



中階程度者

- 使用API抓取所需的資料。



可使用的語料

- [manifestoR](https://manifesto-project.wzb.eu/) package (<https://manifesto-project.wzb.eu/>)
- [newsanchor](#): Client for the News API package (cannot download more than 1,000 results in a given search.)

初階使用者

- 使用他人已經蒐集好的資料，只是尚須經過資料清理。

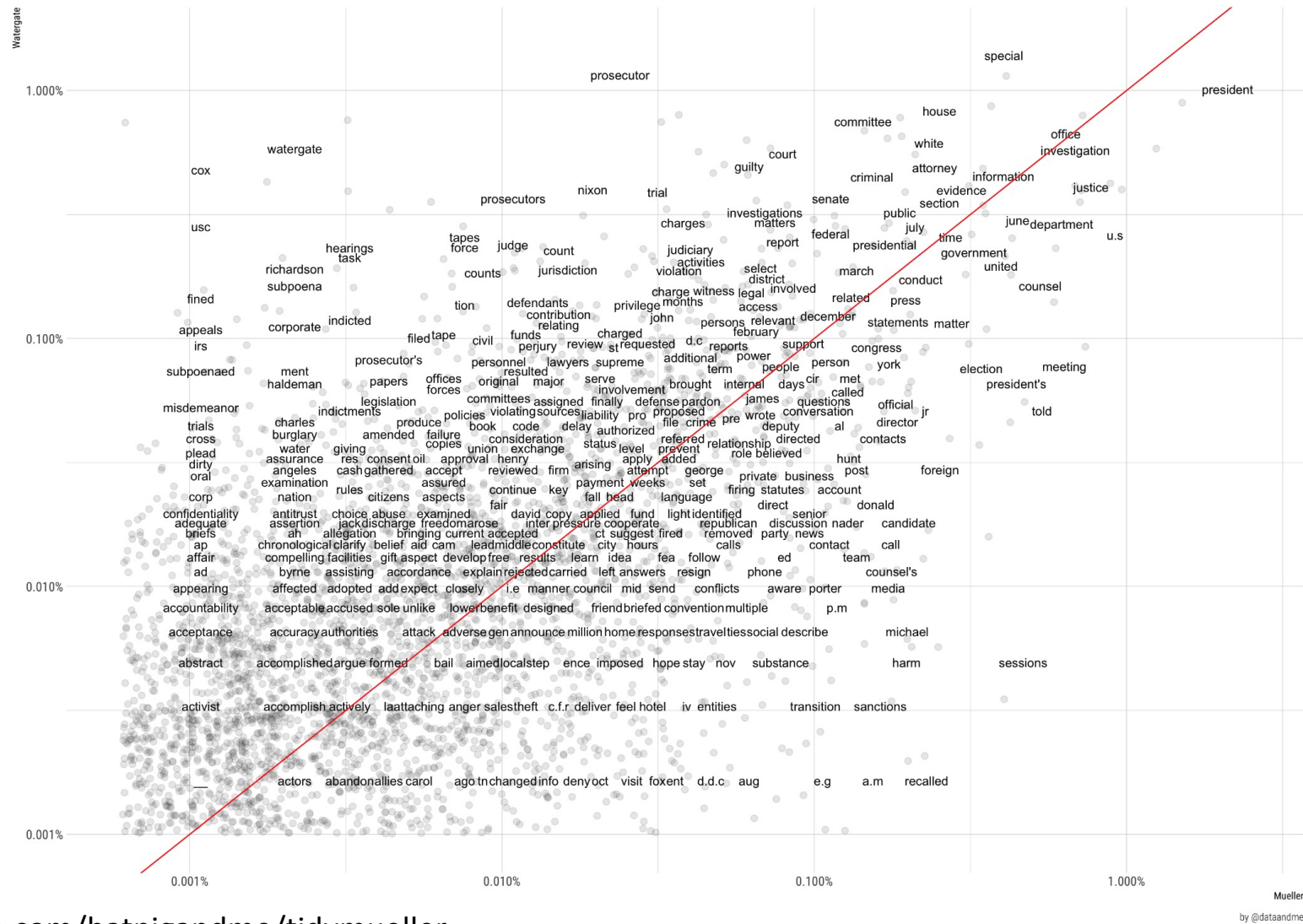


可使用的語料

- Mueller Report
(<https://github.com/gadenbuie/mueller-report>)
- Canadian Hansard dataset (<https://www.lipad.ca/>)
- Harvard Dataverse (<https://dataverse.harvard.edu/>)
 - [UNSC-NE](#): A Named Entity Extension to the UN Security Council Debates Corpus
 - [MAPLE Parliamentary Datasets](#): Full-text and annotated corpora of parliamentary speeches in the legislatures of six European democracies

Word frequencies in Mueller vs. Watergate Reports

Text from Mueller Report (2019), Watergate Special Prosecution Force Report (1975)



<https://github.com/batpigandme/tidymueller>

以前同學做過的主題

- ptt政黑版當成民意探勘的管道
- ptt八卦板看鄉民對同性戀的態度
- 總統經濟言論與經濟表現的關聯
- 不同報紙對於中國相關報導的內容差異
- 不同政治立場媒體對房市報導的差異
- Freedom House人權報導重點與偏誤
- 微博用戶對臺灣政治態度的探討
- 臺北市議員關注議題與政黨區隔
- 行政院新聞澄清機制與選舉之關係
- 公務出國報告內容品質
- 巴黎協定的氣候承諾
- 柯文哲與民粹主義語言：以兩次臺北市長選戰為例
- 輿論對於航空公司罷工之看法
- 總統候選人韓國瑜是否符合民粹主義者定義
- 分析1998到2017年各年度教育年報的總論
- 美國地理距離對移民政策態度的影響
- 普悠瑪新聞資料分析

期末報告分組表

- 填寫連結 (<https://reurl.cc/V5OGkA>)



專案分享

