

「數量方法與金融應用」 ——什麼是統計？

Spring 2026

教師：統計系余清祥

內容：什麼是統計？

下載：<http://csyue.nccu.edu.tw>

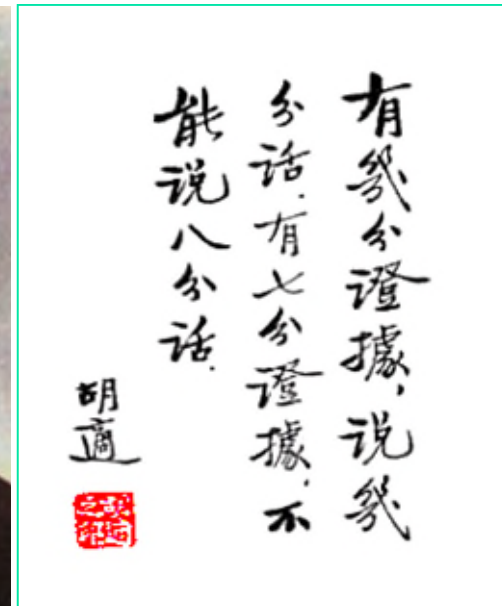
Email: csyue@nccu.edu.tw

2026年3月7日



什麼是統計？

- 統計學是研究定義問題、運用資料蒐集、整理、陳示、分析與推論等科學方法，在不確定(Uncertainty)情況下，做出合理決策的科學。



WHAT IS STATISTICS?

WE MUDDLE THROUGH LIFE MAKING CHOICES
BASED ON INCOMPLETE INFORMATION...

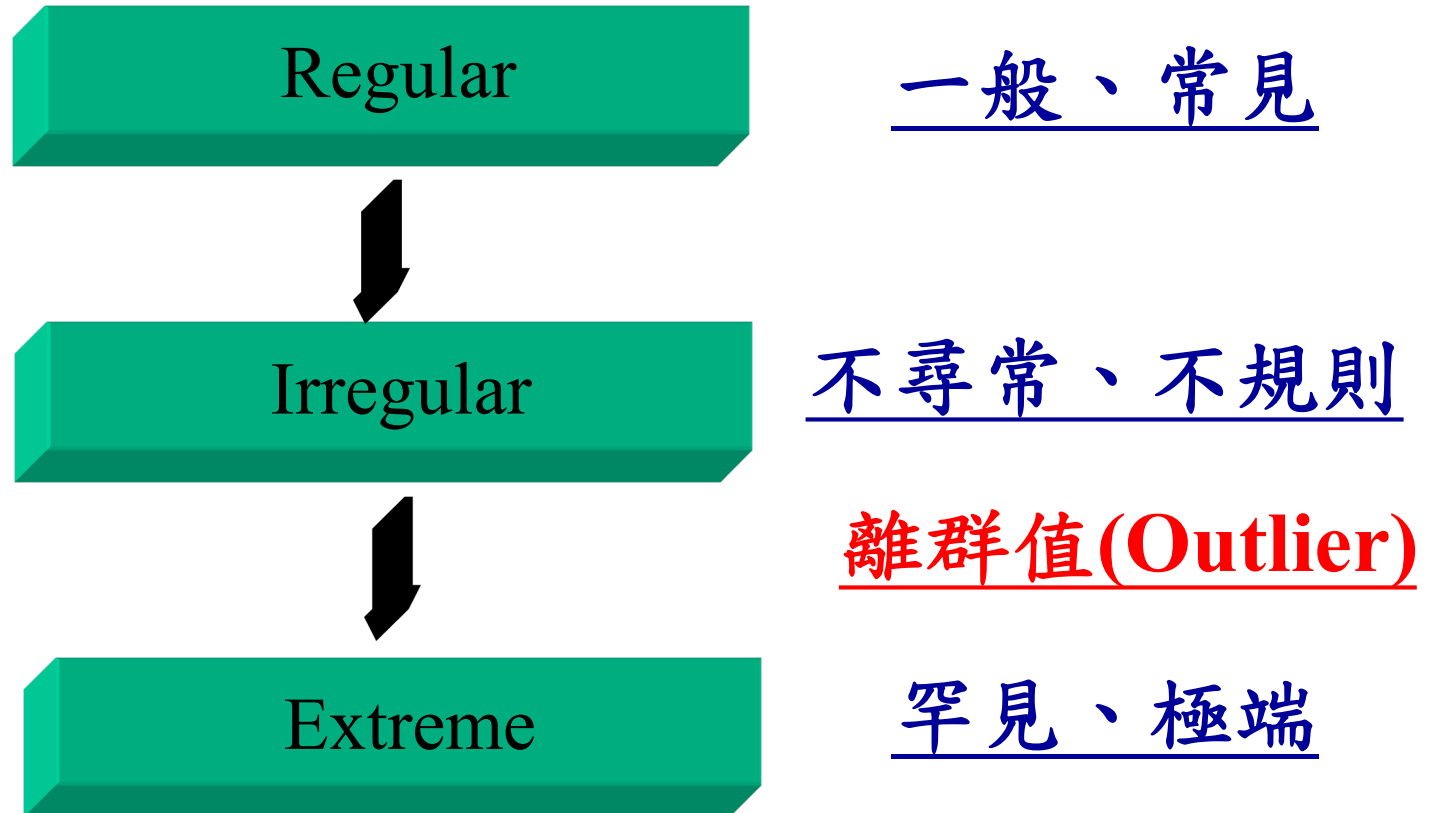
SHOULD I HAVE THE SOUP?
EVERYTHING ELSE IS SO
EXPENSIVE, AND I DON'T
KNOW WHO'S PAYING... ARE
STATISTICIANS STINGY? I'VE
NEVER GONE OUT WITH
ONE BEFORE... THOUGH I
ONCE KNEW A VERY
GENEROUS ACCOUNTANT...

SHOULD I HAVE THE SOUP?
27 OUT OF THE 36 TIMES
I'VE HAD IT, IT WAS PRETTY
GOOD... BUT IS MONDAY THE
REGULAR CHEF'S NIGHT
OFF? AND WHAT IF ALL THE
AIR MOLECULES IN THE
ROOM SUDDENLY FLY UP TO
THE CEILING?



統計與知識

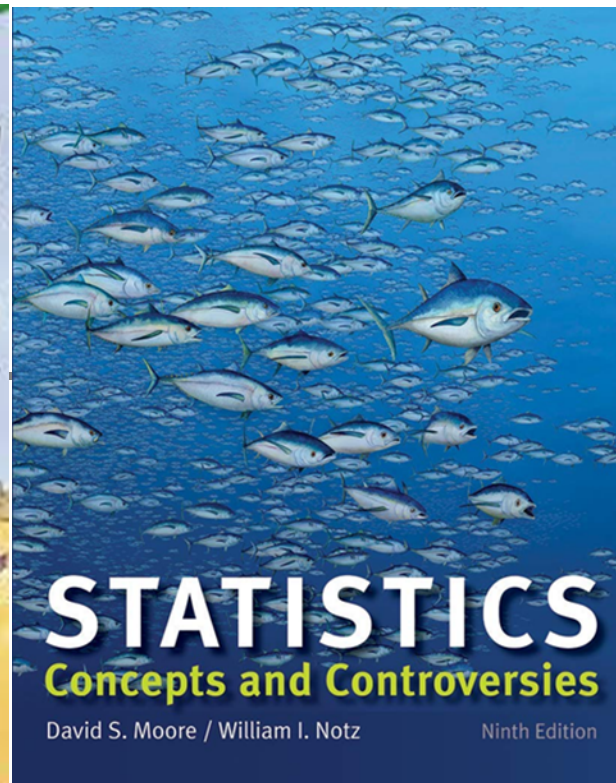
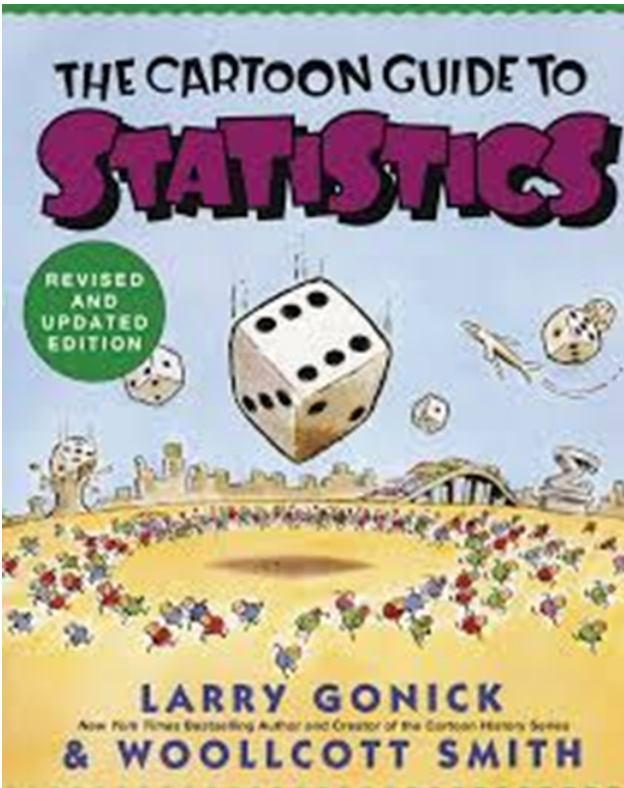
- 資訊整理屬於歸納法(Induction)，從龐雜資料找出共同趨勢，區分資料具有以下哪種特性：



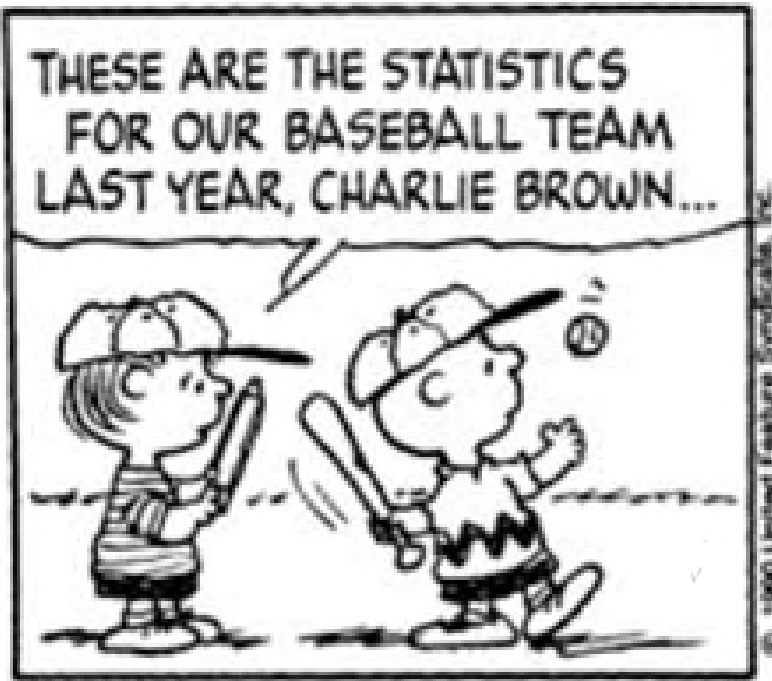


幾本有趣的參考書籍

- 《看漫畫，學統計》(2003)，天下文化。
→ The Cartoon Guide to Statistics (1993)
- 《統計，讓數字說話》(1998)，天下文化。
- 《統計，改變了世界》(2001)，天下文化。
- 《你管別人怎麼想：科學奇才費曼博士》
(2005)，天下文化。
- 《真實的謊言—揭開民調與統計的黑盒子》
(1996)，時報文化。



棒球與統計



棒球與統計高度連結



■ 美國MLB非常熱愛統計

→ 一百多年來創造、紀錄了各類型的統計指標，
用於衡量個人、球隊的表現及特色。

■ 以2025年世界大賽、2024年12強為例

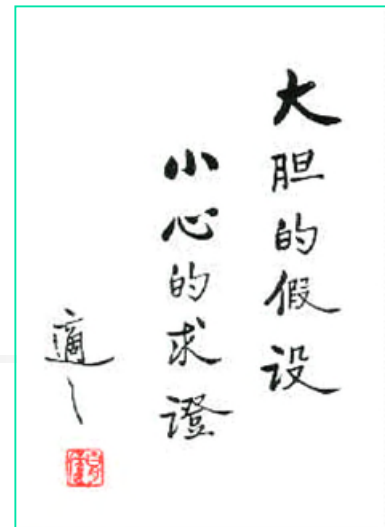
→ 從日籍球員大谷翔平、山本由伸看到什麼？

（註：球探如何物色潛在選手？）

→ 2024年12強中華隊與往年有何不同？



學習統計的幾個要素



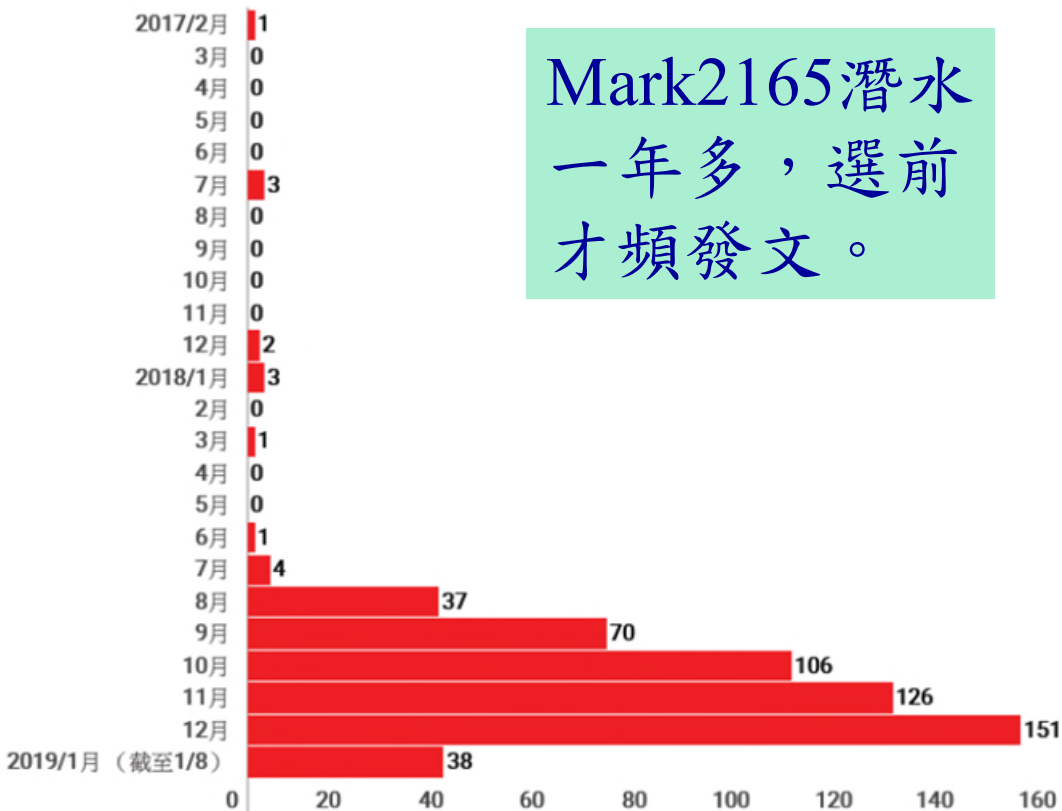
■ 解決問題大致有以下幾個要素：

- 如何定義、測量？(e.g., Variable Format, Data Collection)
- 如何判斷、取捨？(e.g., Estimation, Prediction, Testing)
- 如何詮釋、增加附加價值？(e.g., Utility, Decision)

如何從龐雜訊息中篩選資訊？

- 訊息產生及傳播方式愈趨大量多元，如何篩選、整理、分析變得更加困難。

Mark2165潛水
一年多，選前
才頻發文。



<https://www.wittenberg.edu/sites/default/files/styles/boxlarge2x/public/2020-05/Insta%20Fake%20News.png?itok=16pf5u7H>

<https://www.cw.com.tw/article/5093610>

解決問題(Problem Solving)流程



<https://www.oneeducation.org.uk/course/creativity-and-problem-solving-skills/>

定義問題



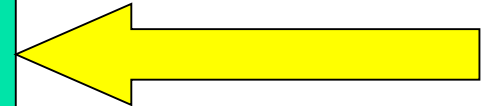
蒐集資料



分析資料

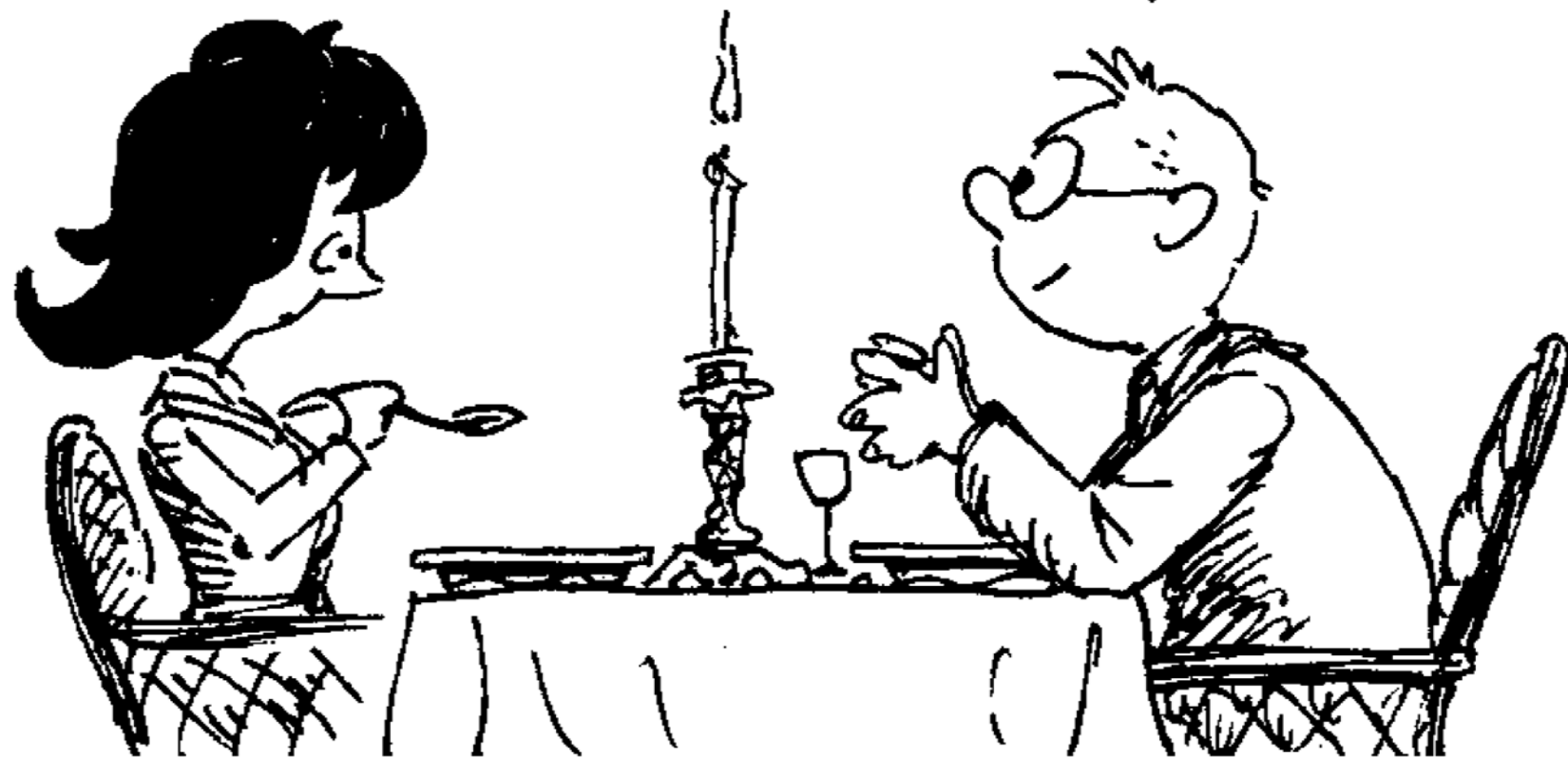


詮釋結果



絕大多數的
資料分析重心

GOOD CHOICE! I'M 95%
CONFIDENT THAT TONIGHT'S
SOUP HAS PROBABILITY
BETWEEN 73% AND 77% OF
BEING REALLY DELICIOUS!



馬克吐溫對統計的想法

There are three kinds of lies:

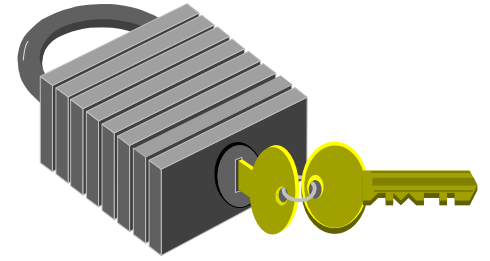
Lies,

Damned lies,

and **Statistics!!**

by Benjamin Disraeli (英國首相)

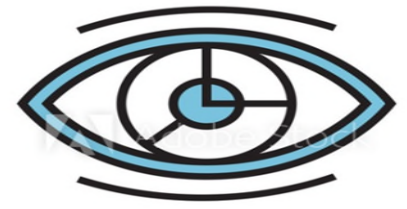




什麼是統計？

- 統計學是研究定義問題、運用資料蒐集、整理、陳示、分析與推論等科學方法，在不確定(Uncertainty)情況下，做出合理決策的科學。

註：通常由部分資料（「樣本」 Sample）推敲出整體（「母體」 Population）的特性。



Visualization

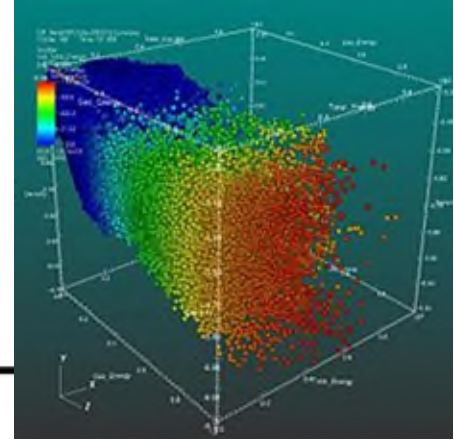
#212488495

文字分析：觀念的視覺化

為什麼需要文學？了解文學、接近文學，對我們形成價值判斷有什麼關係？如果說，文學有一百種所謂「功能」而我必須選擇一種最重要的，我的答案是：德文有一個很精確的說法「*macht sichtbar*」，意思是「**使看不見的東西被看見**」。在我自己的體認中，這就是文學跟藝術最重要、最實質、最核心的一個作用。

——龍應台

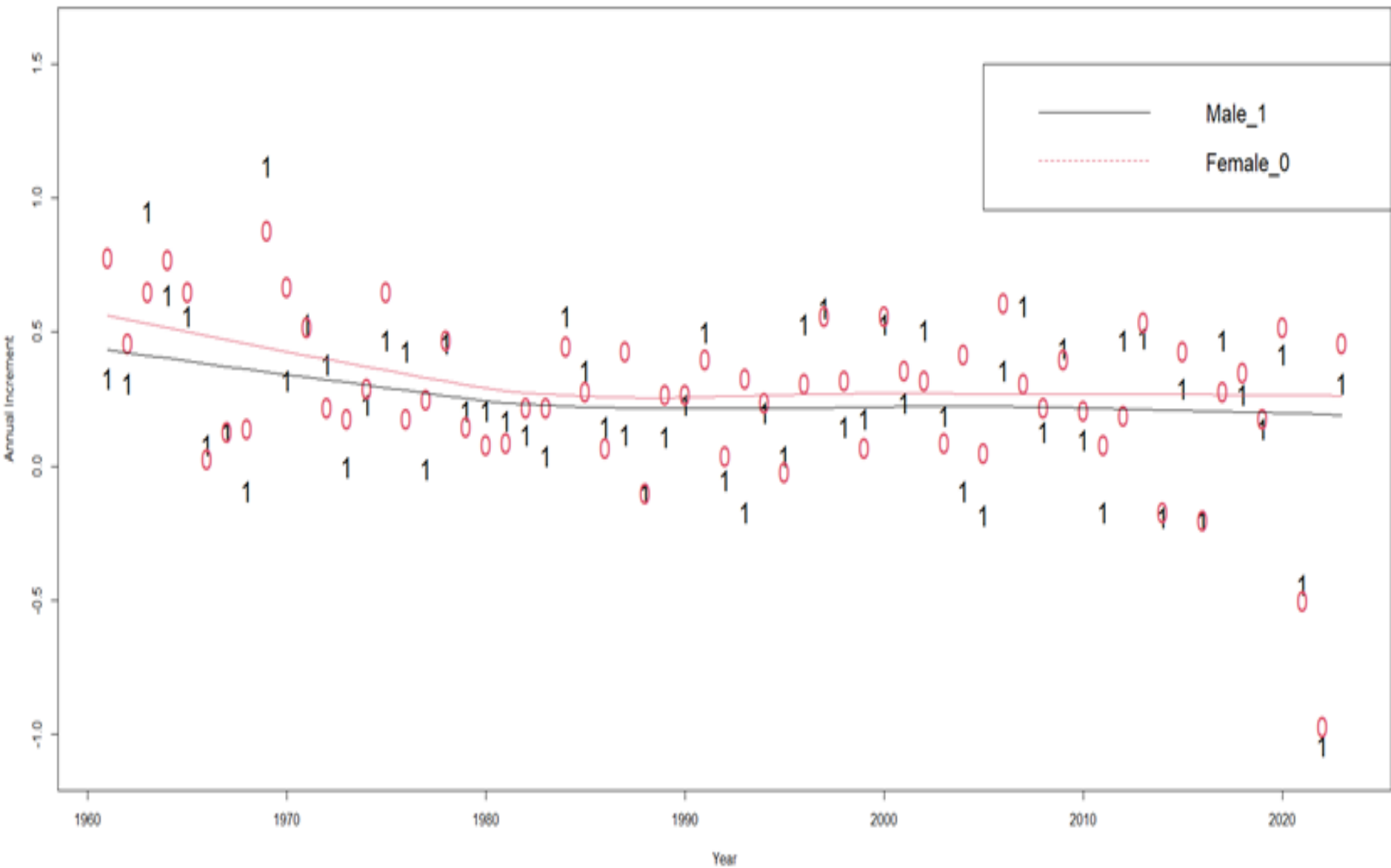
Data visualization



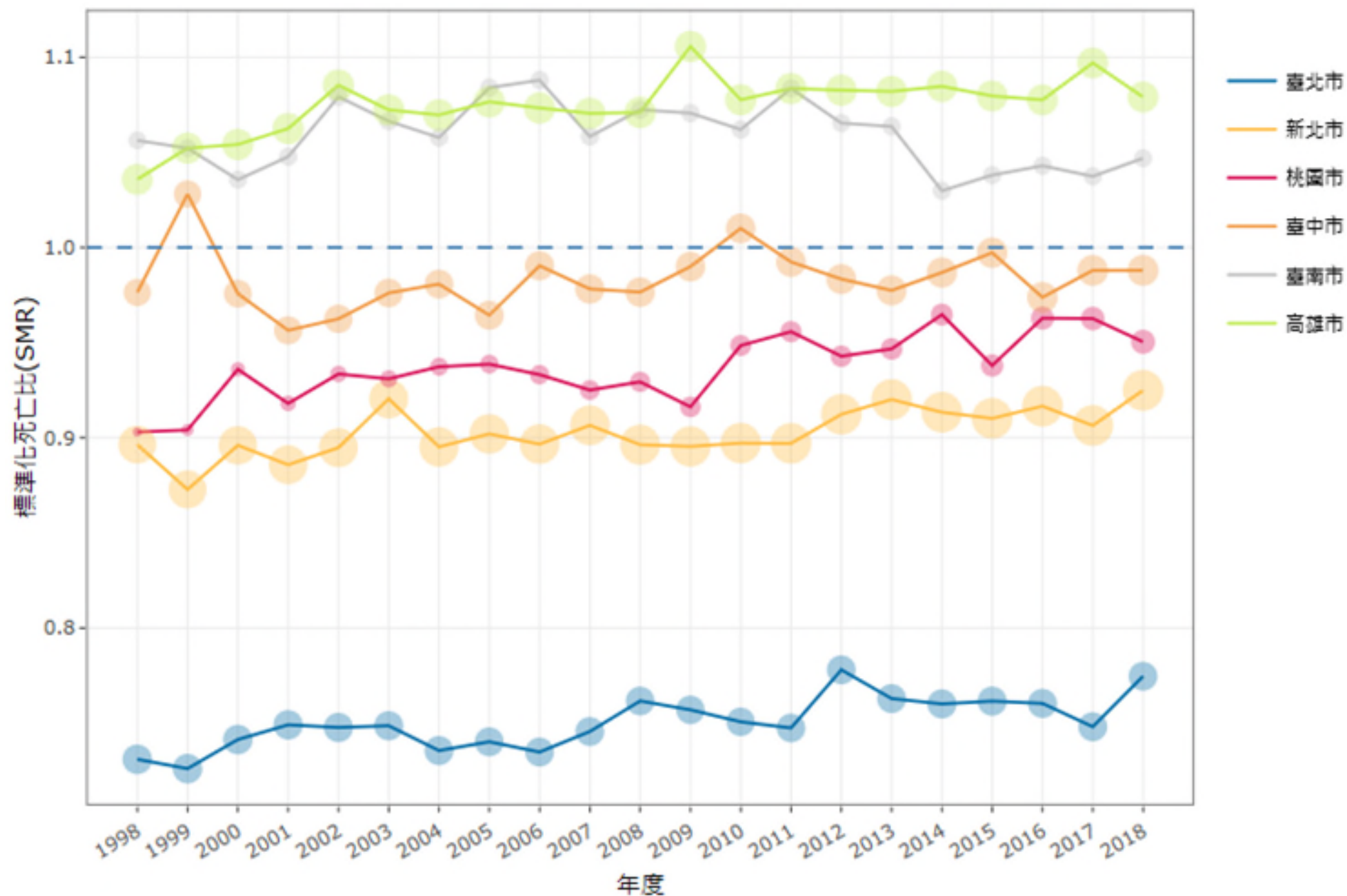
Data visualization is the graphic representation of data. It involves producing images that communicate relationships among the represented data to viewers of the images. This communication is achieved through the use of a systematic mapping between graphic marks and data values in the creation of the visualization. This mapping establishes how data values will be represented visually, determining how and to what extent a property of a graphic mark, such as size or color, will change to reflect changes in the value of a datum.

To communicate information clearly and efficiently, data visualization uses statistical graphics, plots, information graphics and other tools. Numerical data may be encoded using dots, lines, or bars, to visually communicate a quantitative message.^[1] Effective visualization helps users analyze and reason about data and evidence. It makes complex data more accessible, understandable and usable. Users may have particular analytical tasks, such as making comparisons or understanding causality, and the design principle of the graphic (i.e., showing comparisons or showing causality) follows the task. Tables are generally used where users will look up a specific measurement, while charts of various types are used to show patterns or relationships in the data for one or more variables.

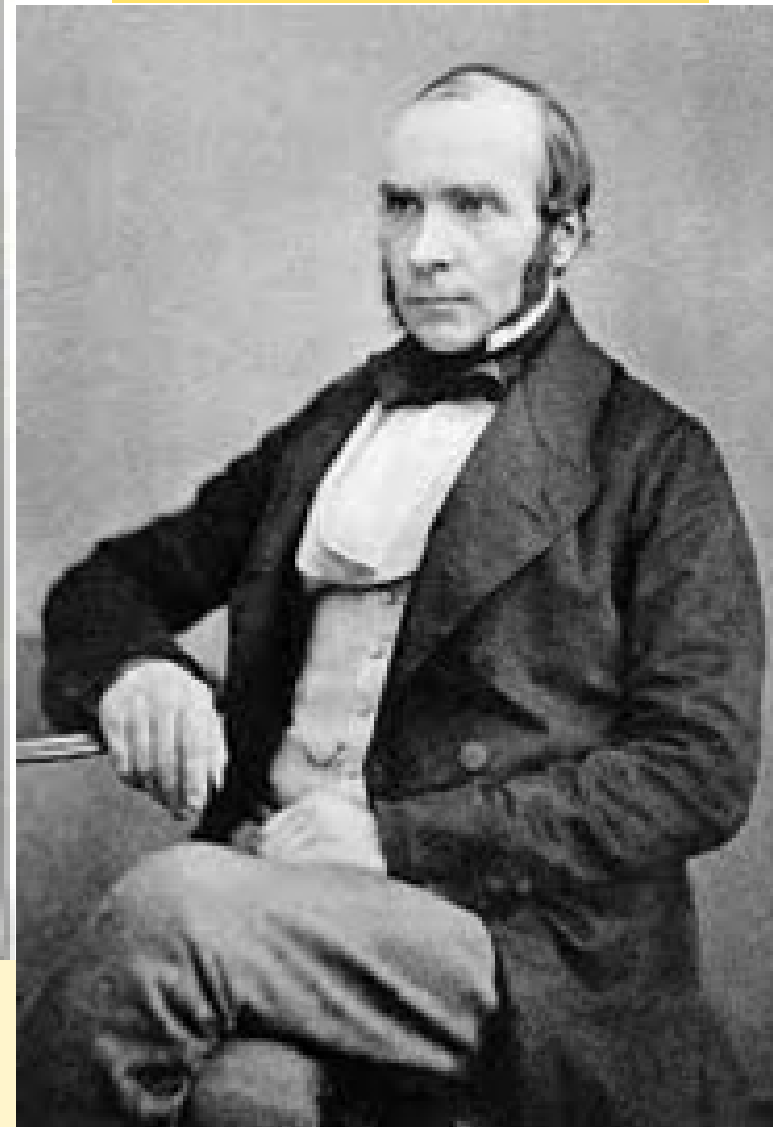
臺灣男女性平均壽命延長趨勢



六都1998-2018年男性標準死亡比



John Snow & Spatial Analysis



巴黎48個行政區的霍亂熱區圖

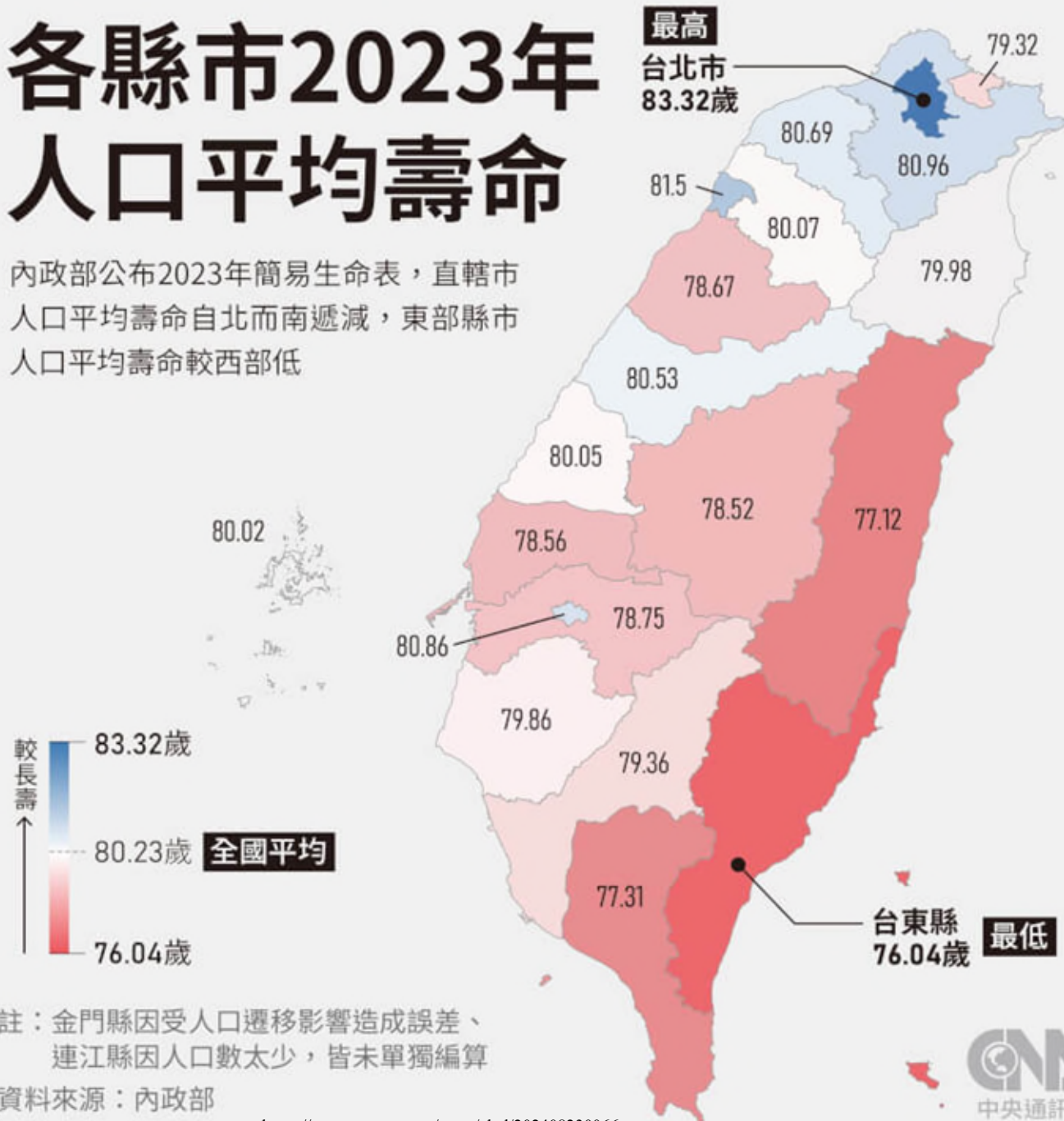
Charles Picquet

https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSh_go8YVCyyZvQsr7UBFWtJ66b1GW2cN0Rlg&usqp=CAU

<https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcTUhpPXin7G9FYiZDkwcle2ZV1YIoqMCEm2g&usqp=CAU>

各縣市2023年人口平均壽命

內政部公布2023年簡易生命表，直轄市人口平均壽命自北而南遞減，東部縣市人口平均壽命較西部低



112年我國10大死因

1 癌症	6 COVID-19
2 心臟疾病	7 高血壓性疾病
3 肺炎	8 事故傷害
4 腦血管疾病	9 慢性下呼吸道疾病
5 糖尿病	10 腎炎、腎病症候群及腎病變

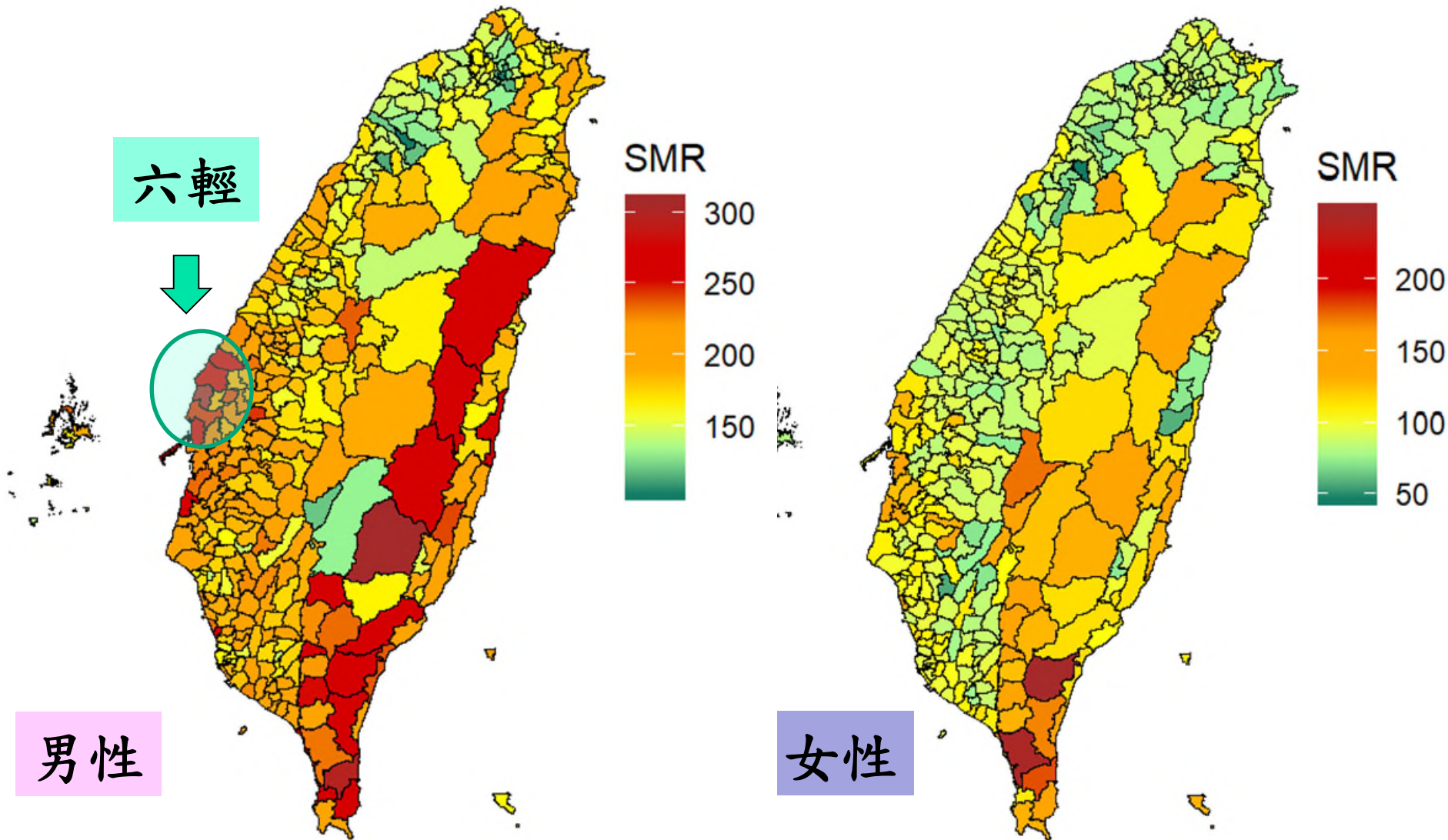
<https://news.pts.org.tw/article/700413>

註：金門縣因受人口遷移影響造成誤差、連江縣因人口數太少，皆未單獨編算

資料來源：內政部

<https://www.cna.com.tw/news/ahel/202408230066.aspx>

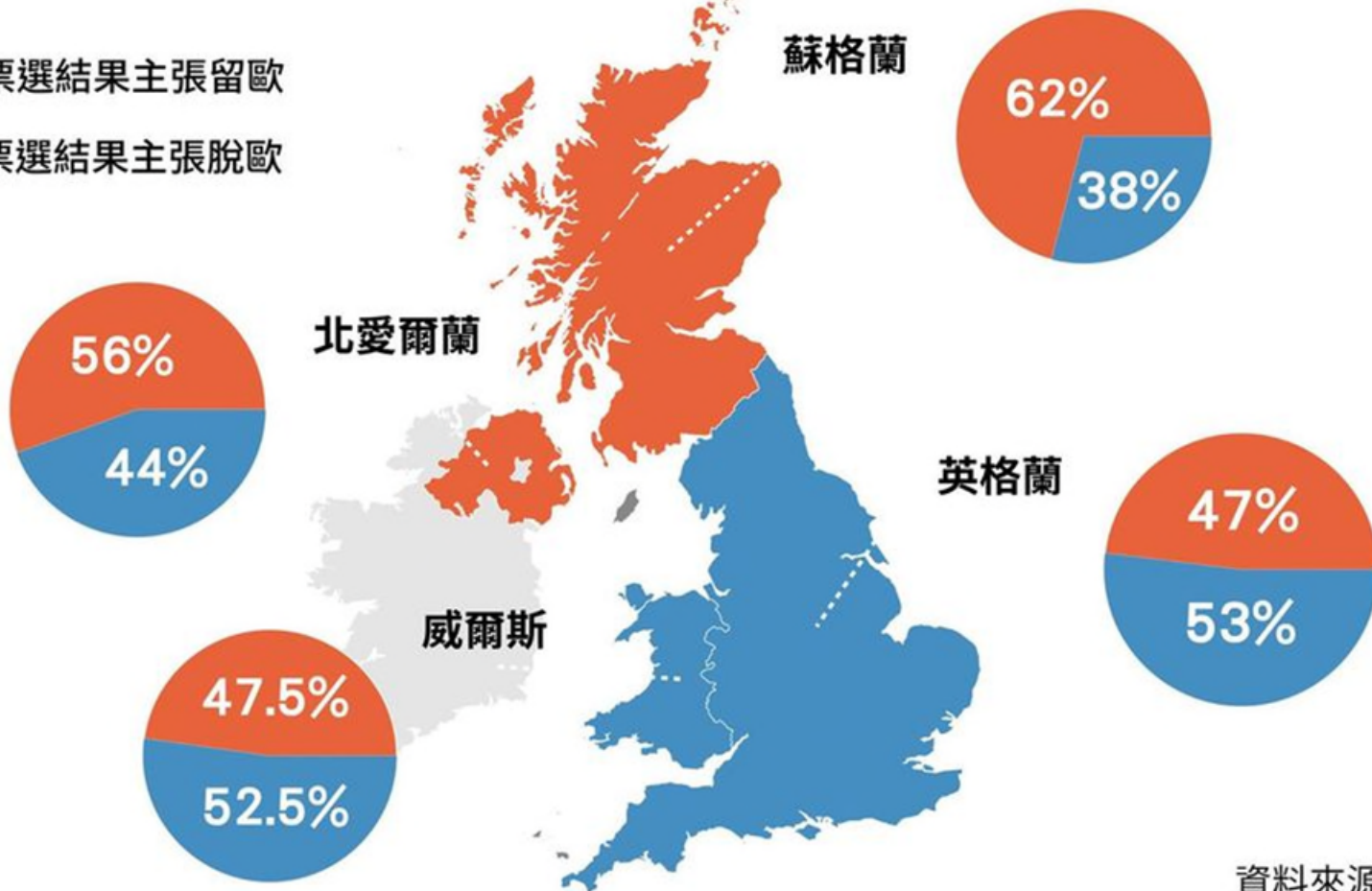
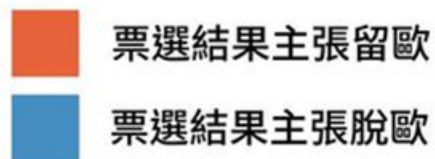
台灣2014~18年鄉鎮市區癌症死亡SMR



註：標準死亡比(SMR, Standard Mortality Ratio)可視為死亡率比值。

英國四大構成 脫歐留歐各有所好

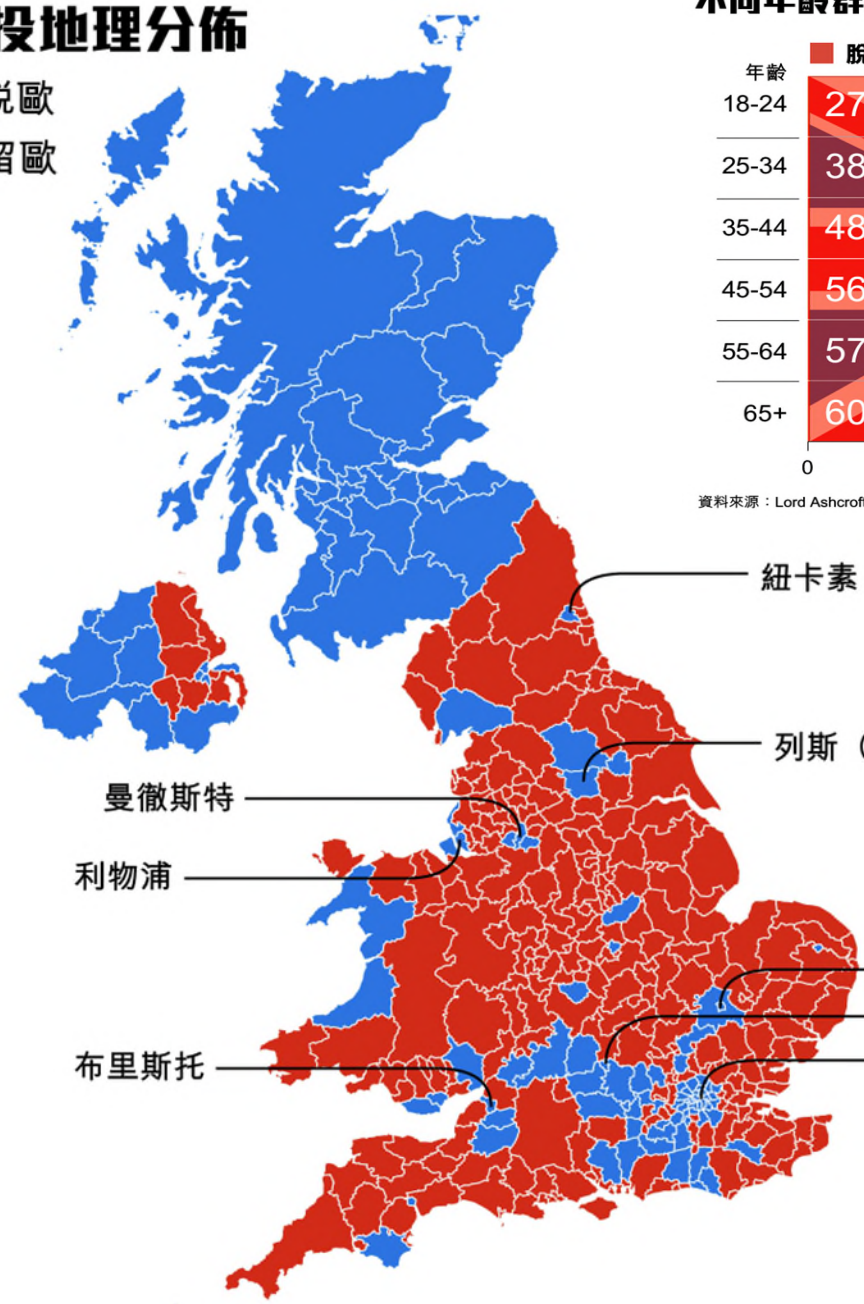
各構成國投票結果



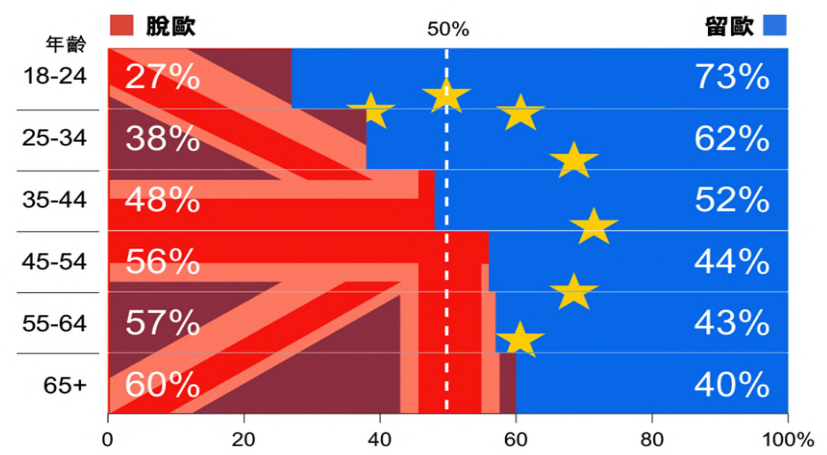
資料來源：BBC

脫歐選舉公投地理分佈

■ 大多數支持脫歐
■ 大多數支持留歐



不同年齡群投票結果



資料來源：Lord Ashcroft Polls、BBC

<https://theinitium.com/article/20160627-opinion-britishandeu/>



資料來源：BBC



<https://theinitium.com/article/20160627-opinion-britishandeu/>

什麼是大數據？



什麼是大數據？

■ 大數據2010年由IBM所提出，涵蓋四個V：

→ 大量化(Volume)：至少TB及PB以上

→ 快速化(Velocity)：即時處理

→ 多樣化(Variety)：視頻、GIS等多樣性

→ 真實性(Veracity)：資料品質（2014年提出）

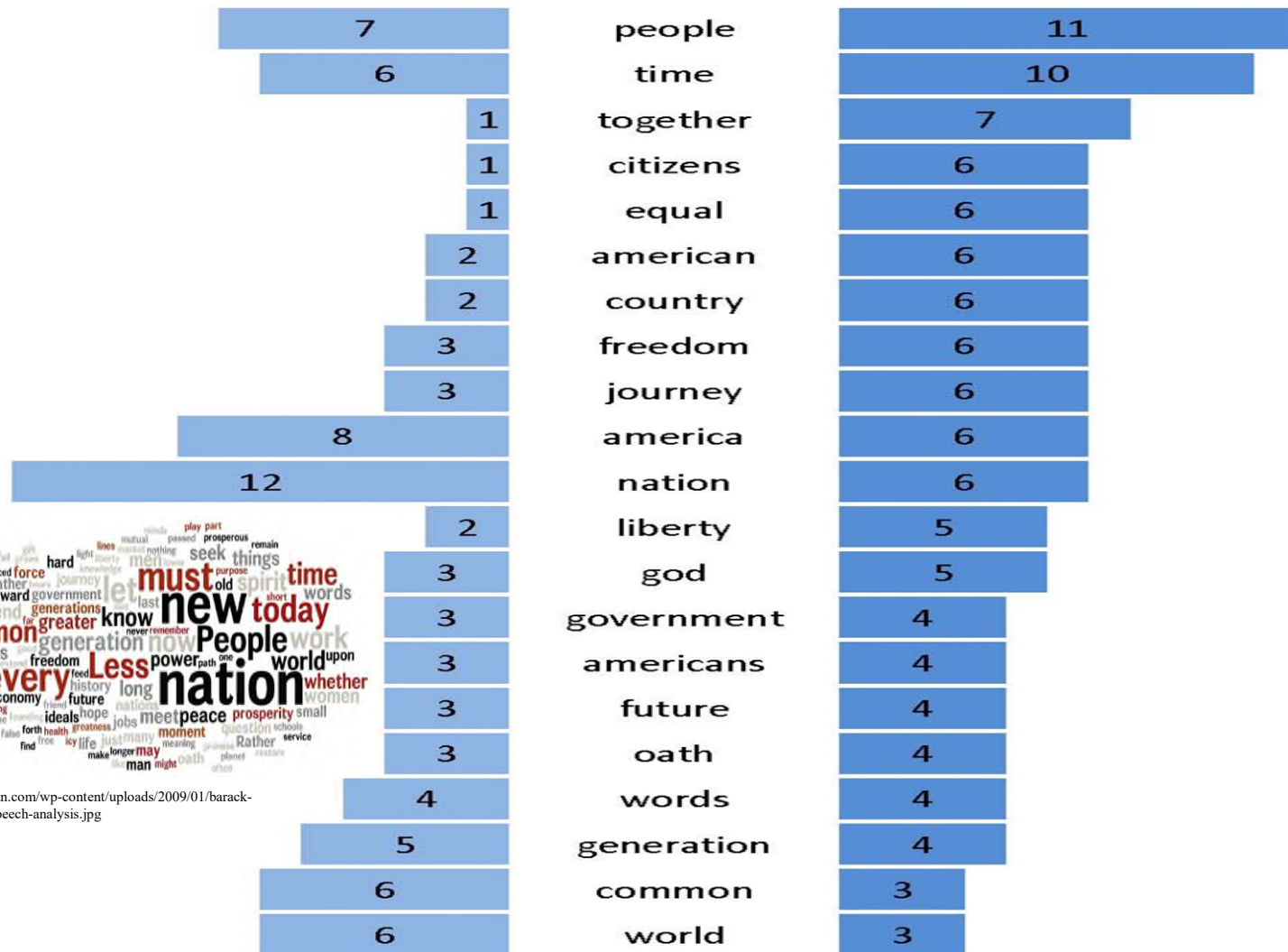
註：加入資料分享與傳遞(Visible)後成為「大、快、雜、疑、轉」；也有人加入價值(Value)成為新的5V。

電腦容量單位的演變（資料爆炸！）

單位	縮寫	意義
Bit	b	1 or 0
Byte	B	8 Bits
Kilobyte	KB	1,024 Bytes
Megabyte	MB	1,024 KB
Gigabyte	GB	1,024 MB
Terabyte	TB	1,024 GB
Petabyte	PB	1,024 TB
Exabyte	EB	1,024 PB
Zettabyte	ZB	1,024 EB
Yottabyte	YB	1,024 ZB

Comparing Inaugural Addresses

■ 2009 ■ 2013



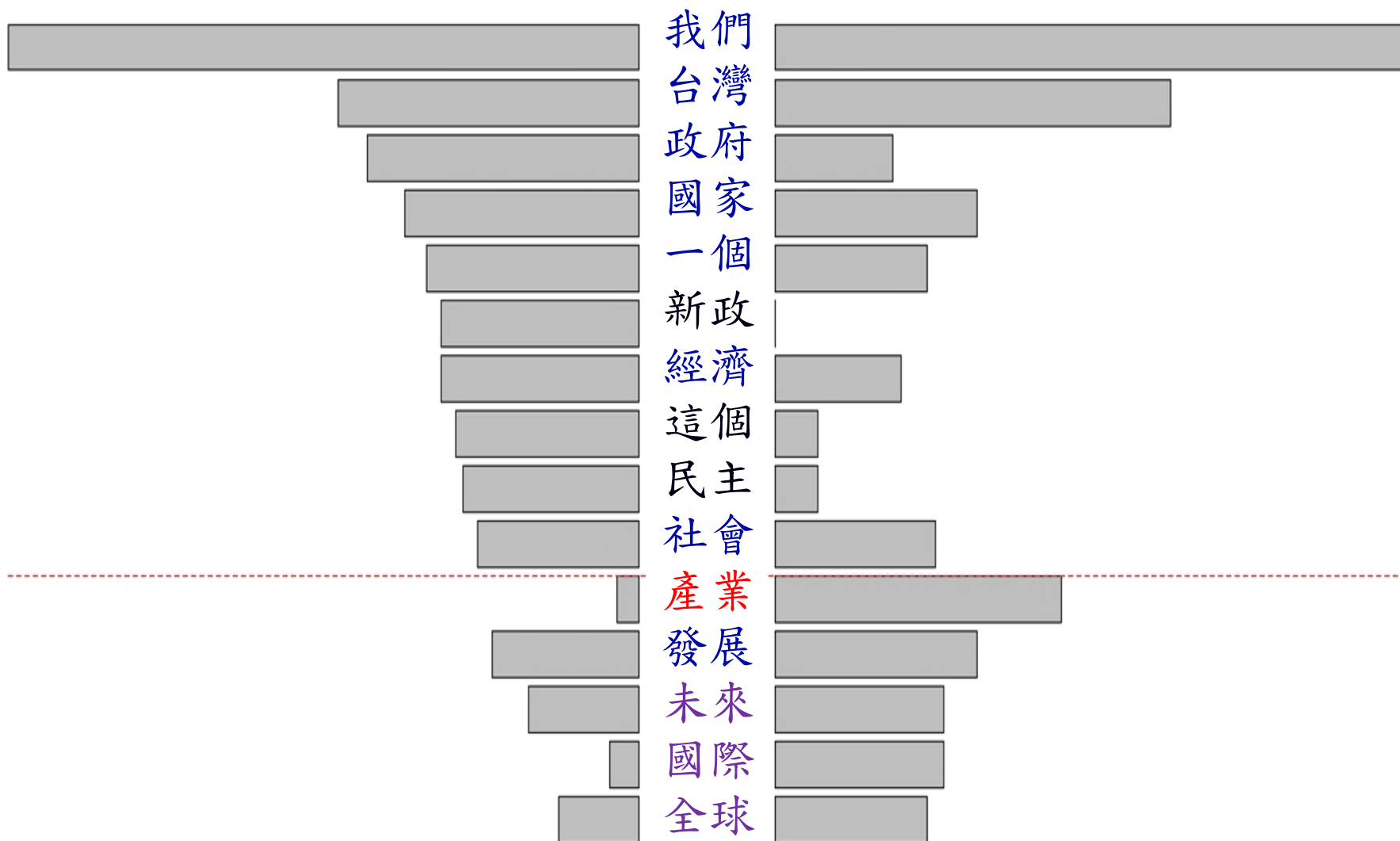
<http://sixminutes.dlugan.com/wp-content/uploads/2009/01/barack-obama-inauguration-speech-analysis.jpg>

Analyzing the speech of President Obama (**Textmining**)

蔡英文總統就職演講稿常見雙字詞

第14任

第15任



臺灣報紙頭條新聞的用字比較(2012~2019年)

《蘋果日報》

卡方指標

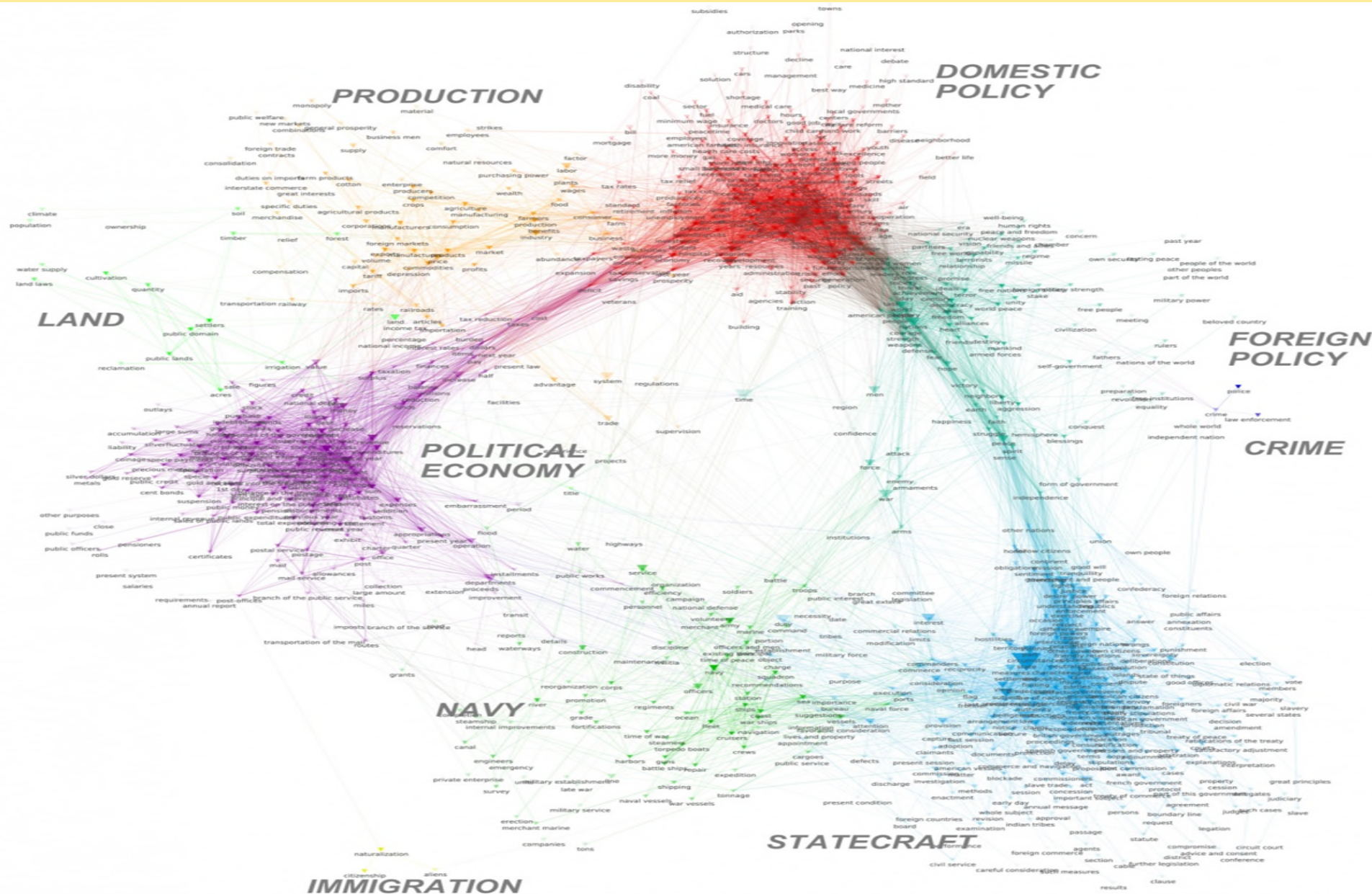


《中國時報》



《自由時報》

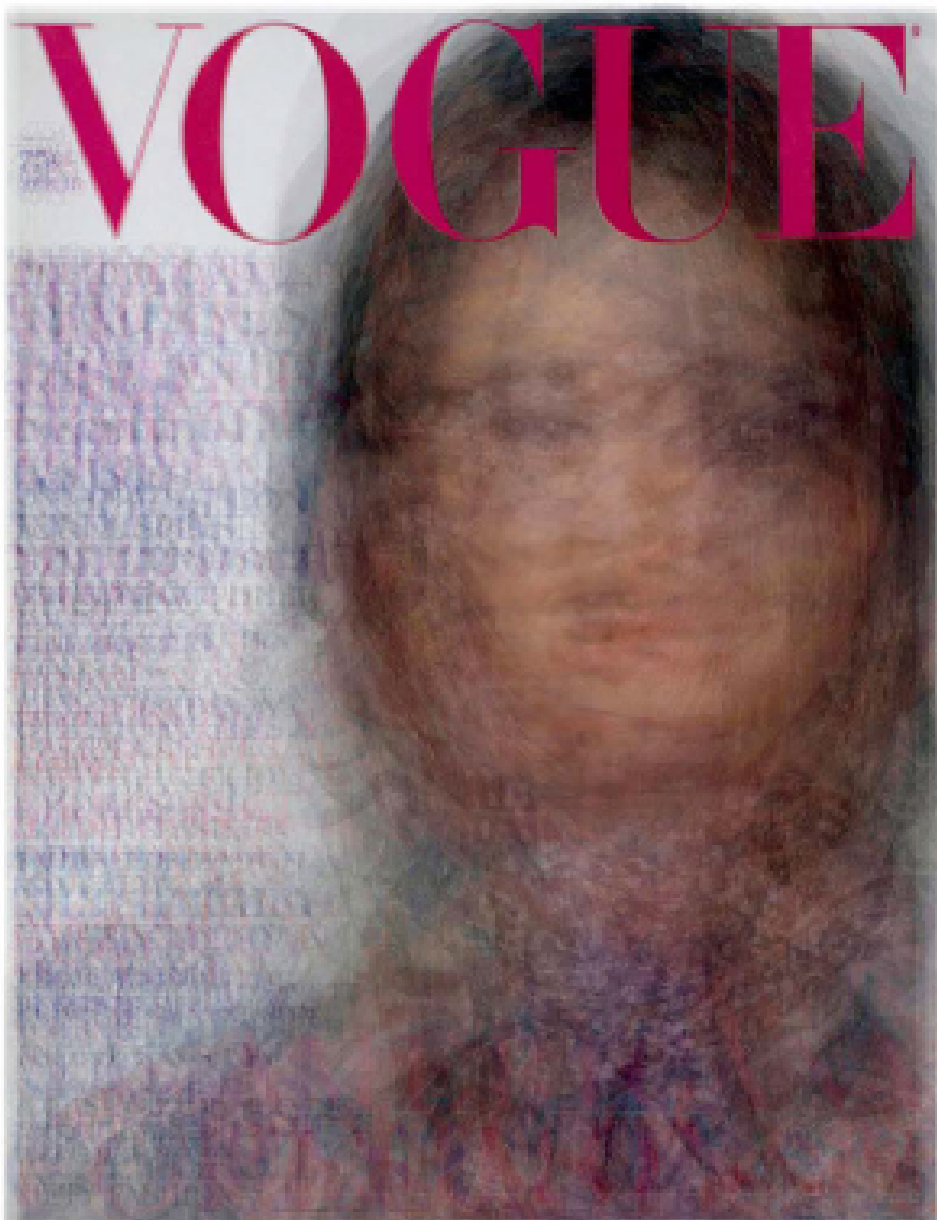
美國總統國情咨文(State of Union;1790年至今)



來源：<https://sciencenode.org/feature/text-mining-strikes-gold-in-political-discourse.php>

耶魯大學數位人文實驗室「Robots Reading Vogue」

1970



1980





https://miro.medium.com/max/3778/1*zdoQ-oKnWAPBKbUMYYL--w.jpeg



資訊解讀的價值



- Data Mining
- Big Data
- Predictive Analytics
- Artificial Intelligence
- Machine Learning

Walmart 賣場——尿布與啤酒

- 沃馬特蒐集與分析顧客資料，並以既有資訊提高銷售業績，成為全球最大的連鎖零售業者。其中發現消費者在週末購物時，許多人會同時購買尿布及啤酒。

→ 為什麼這兩種商品會一起購買？



沃爾瑪於1982年發跡於美國西南部，剛開始只是一家小鎮上的雜貨店，因擅於資料分析，現今是全球最大連鎖零售業者，擁有八千多家分店與超過200萬名的員工。

「尿布與啤酒」的延伸價值

- 尿布與啤酒屬於關連性(Association)的關係，與常見的因果關係(Causality)不同。
- 關連性的價值未必低於因果關係，像是尿布與啤酒的關連，可用於：
 - 商品定價與促銷活動；
 - 商品擺設與賣場動線；
 - 物流倉儲與人員管理。



寶可夢旋風也能帶來商機嗎？



參考資料：<https://tw.news.yahoo.com/%E5%8F%B0%E5%8D%97%E5%A1%9E%E7%88%86-%E5%B0%8F%E9%BB%83%E7%B4%99%E6%A2%9D%E6%9B%9D-%E9%BE%9C%E9%80%9F%E5%8E%9F%E5%9B%A0-%E7%B6%B2%E7%AC%91-%E4%B8%8D%E5%8F%AD%E4%BD%A0%E4%BA%86-143342818.html>

Pokémon GO Safari Zone in Tainan (寶可夢台南狩獵區)：估計主場都會公園奇美博物館有8萬人，大台南全區16萬人，連續五天活動總計主場有56萬人次，台南全區100萬人次。 六億商機！！

今日新聞NOWnews 記者陳聖璋 2018年11月5日





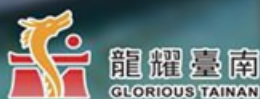
「Pokémon GO Safari Zone in Tainan」今（1）日早上10點半開始將連續5天展開，而這場寶可夢盛會預計帶來的人潮預估將有20萬，各大飯店民宿住房率超過8成，甚至有業者開心表示，「比陸客來台還有用！」

三立新聞網 2018年11月1日 上午11:06

廣告



2024 年 3 月 9-10 日
City Safari 台南



15萬人湧入抓寶!

台南寶可夢大爆發

週六高鐵票已賣光



北門區域

北門遊客中心
北門鹽田
井仔腳鹽田

鹽水月津港區域

鹽水車站
鹽水老街
月津親水公園

安平古堡
安平老街

安平燈區

中西區

國立臺灣文學館
台南林百貨
臺南孔廟

北區及東區

國立成功大學
台南車站
臺南知事官邸

台南高鐵站
沙崙車站
大臺南會展中心

高鐵燈區

原街役場
新化老街
新化武德殿

新化區



戴著狩獵帽的伊布

2024年3月9日及3月10日
Pokémon GO
City Safari：台南

祝大家有個
愉快的學期！

