

Age\_Balance\_散佈圖



Age\_Balance\_散佈圖\_log\_趨線圖



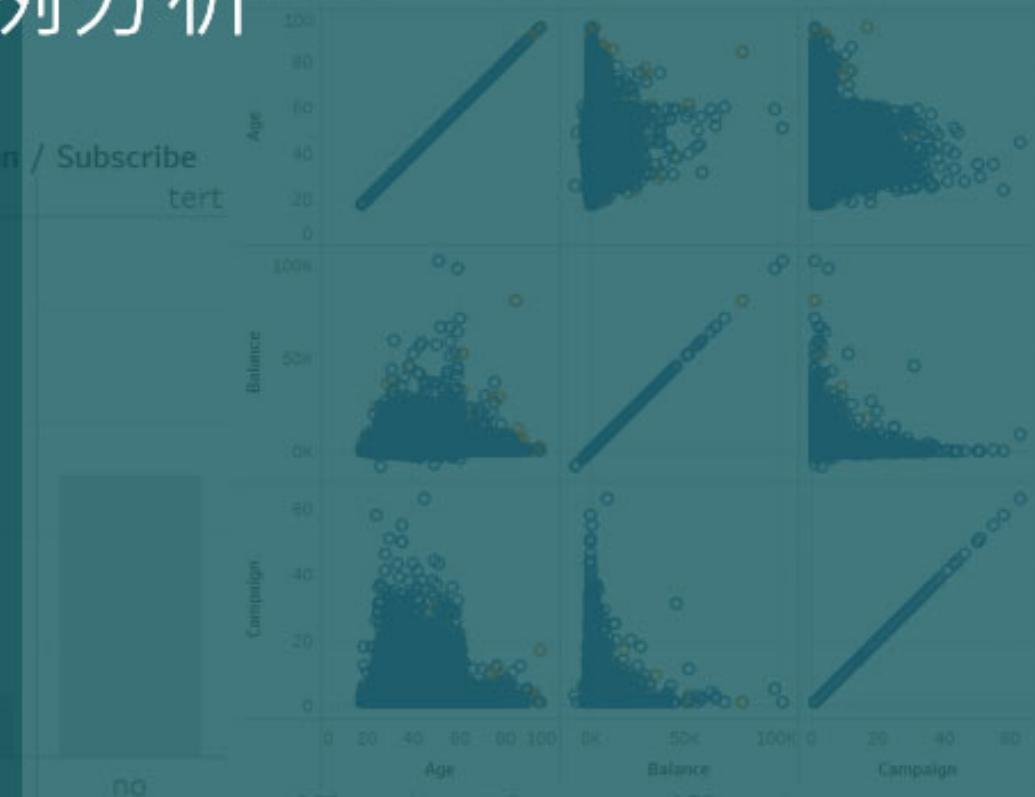
# 國立臺北科技大學資訊與財金管理系(所)

## Tableau 金融商品案例分析

Education\_Subscribe\_併排長條圖

主講人：李明昌 博士

2020.11.16



Subscribe  
no  
yes

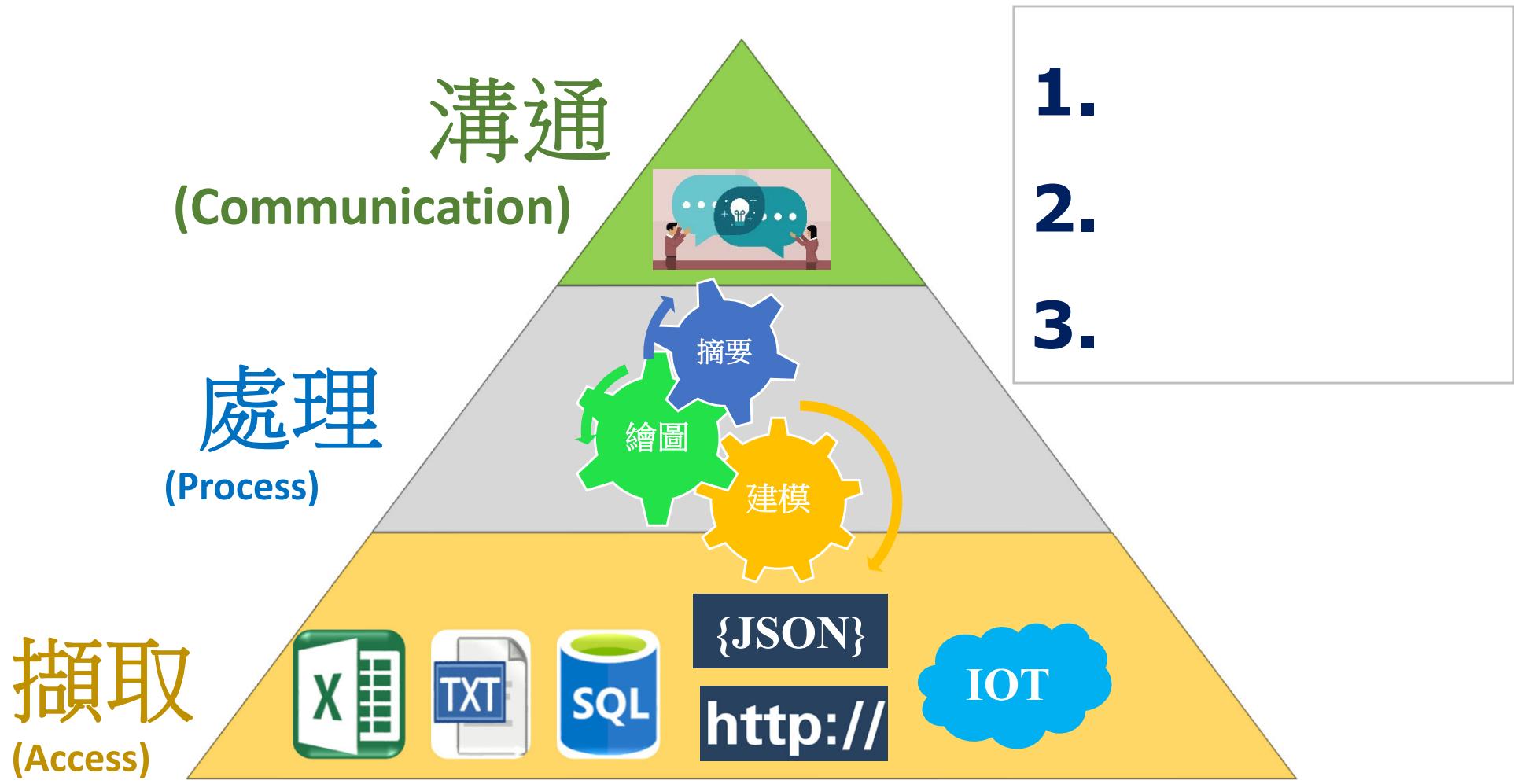
# 大綱

1. 資料視覺化簡介
2. Tableau 簡介
3. Tableau 金融商品案例分析

- 下載檔案: <https://github.com/rwepa/DataDemo>
  - Tableau\_Financial\_Products\_Case\_Study.pdf
  - Tableau\_Financial\_Products\_Case\_Study.twbx

# 1. 資料視覺化簡介

# 資料分析架構→APC方法

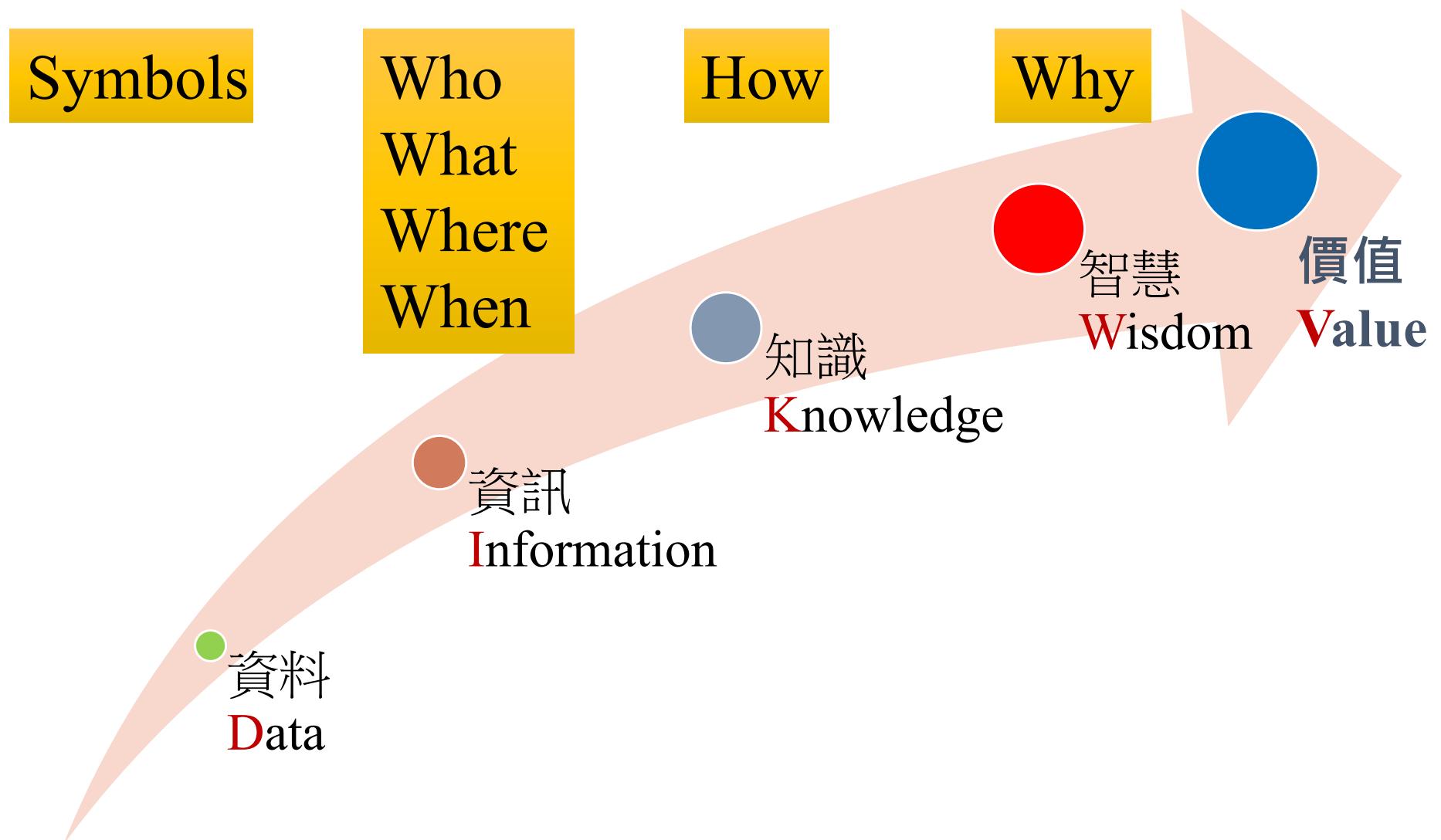


# 資料視覺化 (Data Visualization)

- 資料視覺化不是僅有表面敘述或是畫出圖形
- 資料視覺化使用統計圖形、圖表、資訊圖表和其他工具。可以使用點、線或條對數字資料進行編碼，以便在視覺上傳達定量/定性資訊(參考資料1)。
- 資料視覺化包括下列轉換過程(參考資料2)：
  - 概念 (concepts)
  - 想法 (ideas)
  - 特性 (properties)
- 參考資料1 <https://zh.wikipedia.org/wiki/数据可视化>
- 參考資料2 Tufte, E. R., Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University, maintains that “excellence in statistical graphics consists of complex ideas communicated with **“clarity, precision, and efficiency”**”, 1983.

正確性, 精確性, 效率性

# 現代資料視覺化的智慧需求



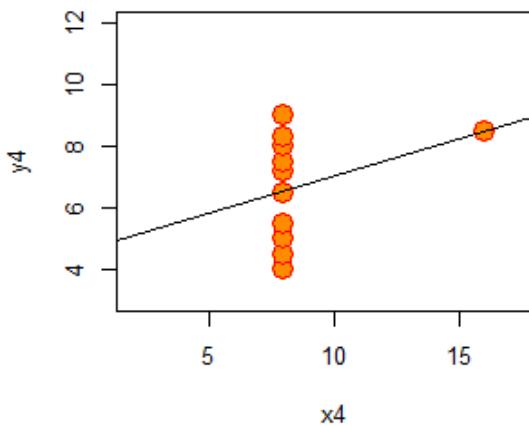
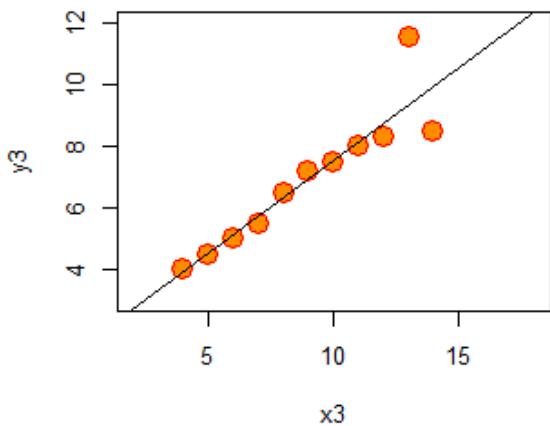
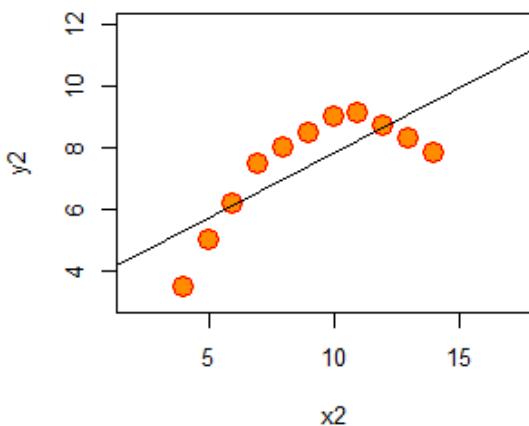
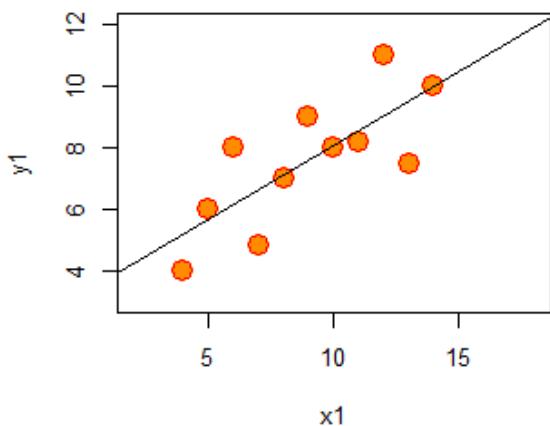
# 為什麼需要視覺化

- 人腦不善於閱讀和分析大量資料

	A	B	C	D	E	F	G
1	Country or Area	Subgroup	Year	Source	Unit	Value	Value Footnotes
2	Afghanistan	Male 5-14 yr	2002-2011	UN_Demographic	a Percent	11	1
3	Afghanistan	Female 5-14 yr	2002-2011	UN_Demographic	a Percent	10	1
4	Afghanistan	Total 5-14 yr	2002-2011	UN_Demographic	a Percent	10	1
5	Albania	Male 5-14 yr	2002-2011	UN_Demographic	a Percent	14	1
6	Albania	Female 5-14 yr	2002-2011	UN_Demographic	a Percent	9	1
7	Albania	Total 5-14 yr	2002-2011	UN_Demographic	a Percent	12	1
8	Algeria	Male 5-14 yr	2002-2011	UN_Demographic	a Percent	6	2
9	Algeria	Female 5-14 yr	2002-2011	UN_Demographic	a Percent	4	2
10	Algeria	Total 5-14 yr	2002-2011	UN_Demographic	a Percent	5	2
11	Angola	Male 5-14 yr	2002-2011	UN_Demographic	a Percent	22	3
12	Angola	Female 5-14 yr	2002-2011	UN_Demographic	a Percent	25	3
13	Angola	Total 5-14 yr	2002-2011	UN_Demographic	a Percent	24	3
14	Argentina	Male 5-14 yr	2002-2011	UN_Demographic	a Percent	8	2
15	Argentina	Female 5-14 yr	2002-2011	UN_Demographic	a Percent	5	2
16	Argentina	Total 5-14 yr	2002-2011	UN_Demographic	a Percent	7	2
17	Armenia	Male 5-14 yr	2002-2011	UN_Demographic	a Percent	5	2
18	Armenia	Female 5-14 yr	2002-2011	UN_Demographic	a Percent	3	2
19	Armenia	Total 5-14 yr	2002-2011	UN_Demographic	a Percent	4	2

# 為什麼需要視覺化 (續)

- 但是人腦很善於閱讀和分析圖形



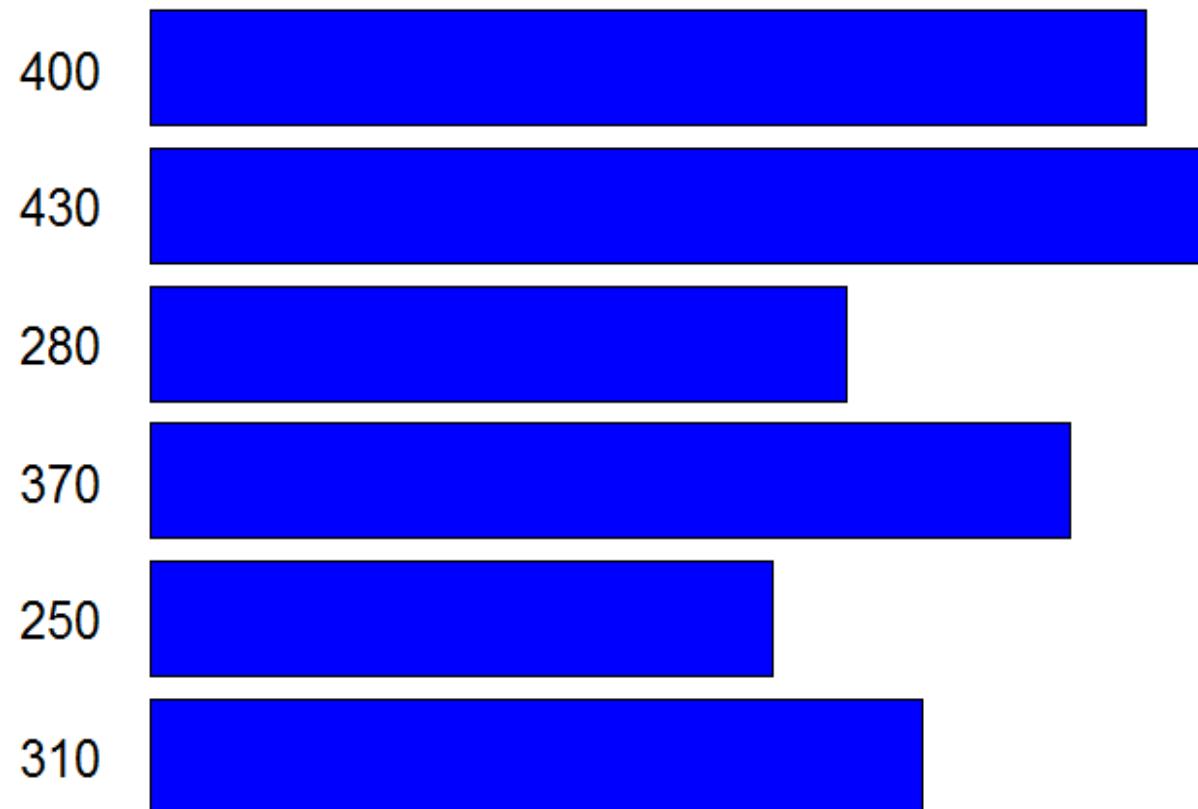
# 為什麼需要視覺化 (續)

- 人類是**視覺動物**，其視覺神經系統有強大的模式識別和分析能力，視覺化是啟動這套系統的途徑
- 視覺化是一種高效的**資訊壓縮和展示方法**，能將大量資料快速傳輸給人的大腦
- 視覺化能**探索並提煉資料**，並促進新的問題的提出和解決
- 探索式資料分析是資料視覺化的重要應用
  - identify properties, relationships, regularities, or patterns

Patterns 樣式

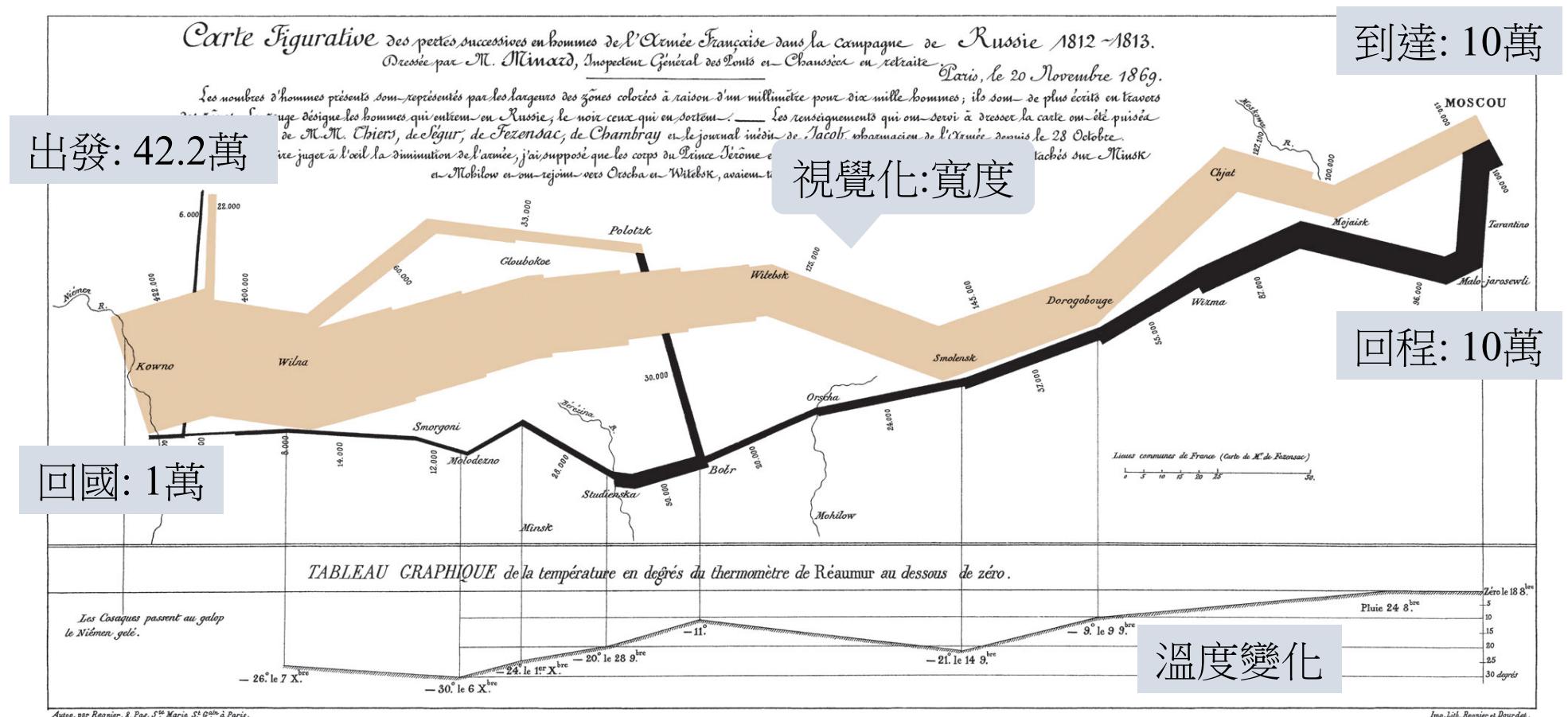
# 為什麼需要視覺化 (續)

- 人們接收視覺屬性(顏色, 大小, 形狀)非常好
- 圖形認知優於心智思考



## 俄法戰爭

## 第一個著名的視覺化

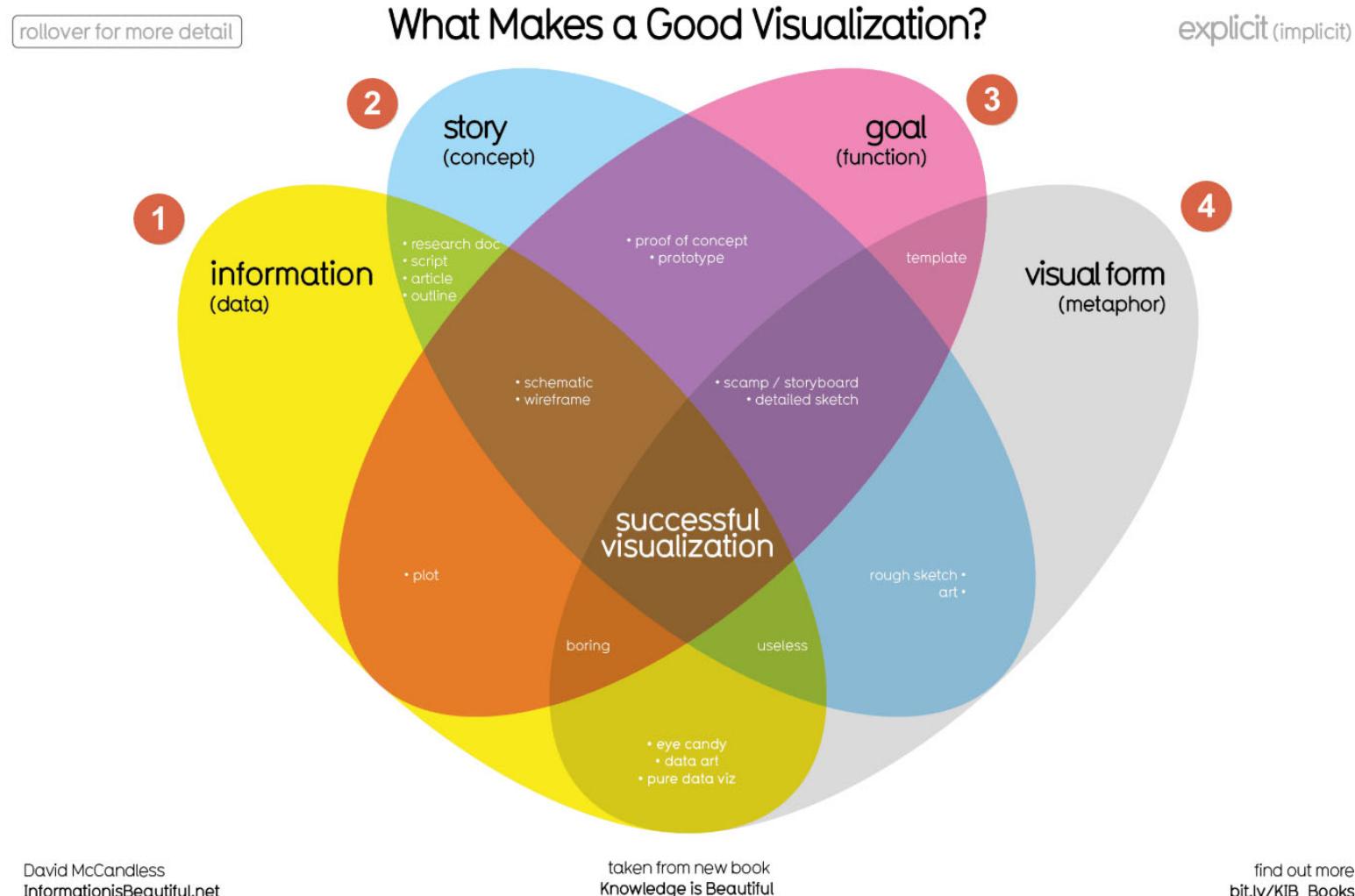


- Charles Joseph Minard, a French engineer, in 1869.
- Illustrate the number of losses suffered by Napoleon's army during the disastrous march toward Moscow in 1812.6 - 1812.11

# 資料視覺化基本概念

- 明確視覺化的**具體目標**
  - 可視化的類型？
  - 要傳送什麼樣的資訊？哪些最重要？使用什麼資料？
- 考慮**觀眾/聽眾**之不同
  - 閱讀者的角色和知識背景？
  - 需要何類資訊？細節程度？
- 風格簡約，凸顯主題
  - 閱讀者的注意力是有限的，提煉重點
  - 最大化 Data-ink ratio
- 選擇合適的視覺編碼方法
  - 位置、長度、尺寸、角度、顏色、形狀...

# 成功的資料視覺化



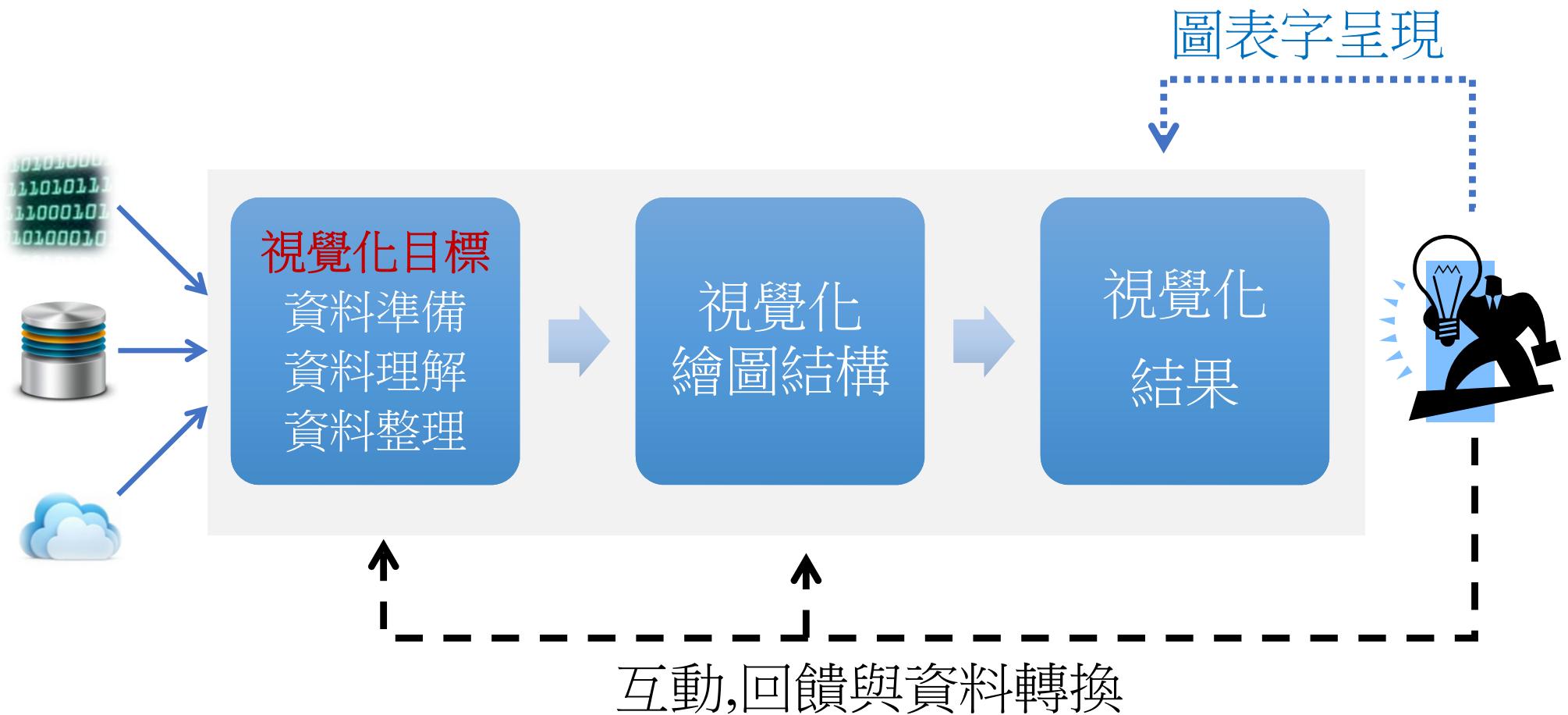
參考資料 <https://informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/>

# 資料視覺化4大成功關鍵

- 訊息 Information: 準確性, 真實性, 一致性
- 故事 Story: 資料視覺化的認同性
- 目標 Goal: 找到重要結論
- 視覺表現 Visual Form: 藝術的呈現

參考資料 David McCandless , Knowledge Is Beautiful: Impossible Ideas, Invisible Patterns, Hidden Connections - Visualized, Oct 21, 2014.

# 資料視覺化流程



# 視覺化目標

- 資料來源 (內部/外部, 免費/付費)
- 聽眾 vs. 觀眾
- 期望結果: 高層主管 vs. 基層員工
- 確認視範化目標
  - 產品銷售資料分析(產品別, 部門別, 時間別)
  - 教務研究-學生休退學預警分析(院別, 系別, 學習, 休學, 退學)
  - 醫學研究-疾病/藥品預測分析 (類別預測, 數值預測)
  - 製造分析-生產最佳化分析
  - 金融商品分析

# 資料準備 - Open Data 下載

- 政府資料開放平台 <https://data.gov.tw/>
- UCI Machine Learning Repository  
<https://archive.ics.uci.edu/ml/datasets.php>
- Google Dataset Search  
<https://toolbox.google.com/datasetsearch>

New !
- Kaggle Dataset <https://www.kaggle.com/datasets>

參考資料 <http://rwepa.blogspot.com/2019/09/dataset.html>

# 資料理解

- 包括描述資料、探索資料、核驗資料品質
- 敘述統計分析
  - 六力分析(R: summary函數): 最小值, 25百分位數, 中位數, 平均值, 75百分位數, 最大值
- 繪圖
  - 依**群組**特性
  - 依**時間**特性
  - 趨勢
  - 離群值 (Outliers)
  - 散佈圖、散佈圖矩陣
  - 盒鬚圖
  - **樣式 (Patterns)**

# 視覺化繪圖結構

- 資料分析-用統計和挖掘方法對資料進行模式提煉，通常是解釋性視覺化的前提
- 設計特徵 (features) 映射，將資料特徵映射到圖形元素，可手繪草圖構思

# 視覺化繪圖結構(續)

Points



Size



Lines



Orientation



Surfaces



Color



Volumes



Texture



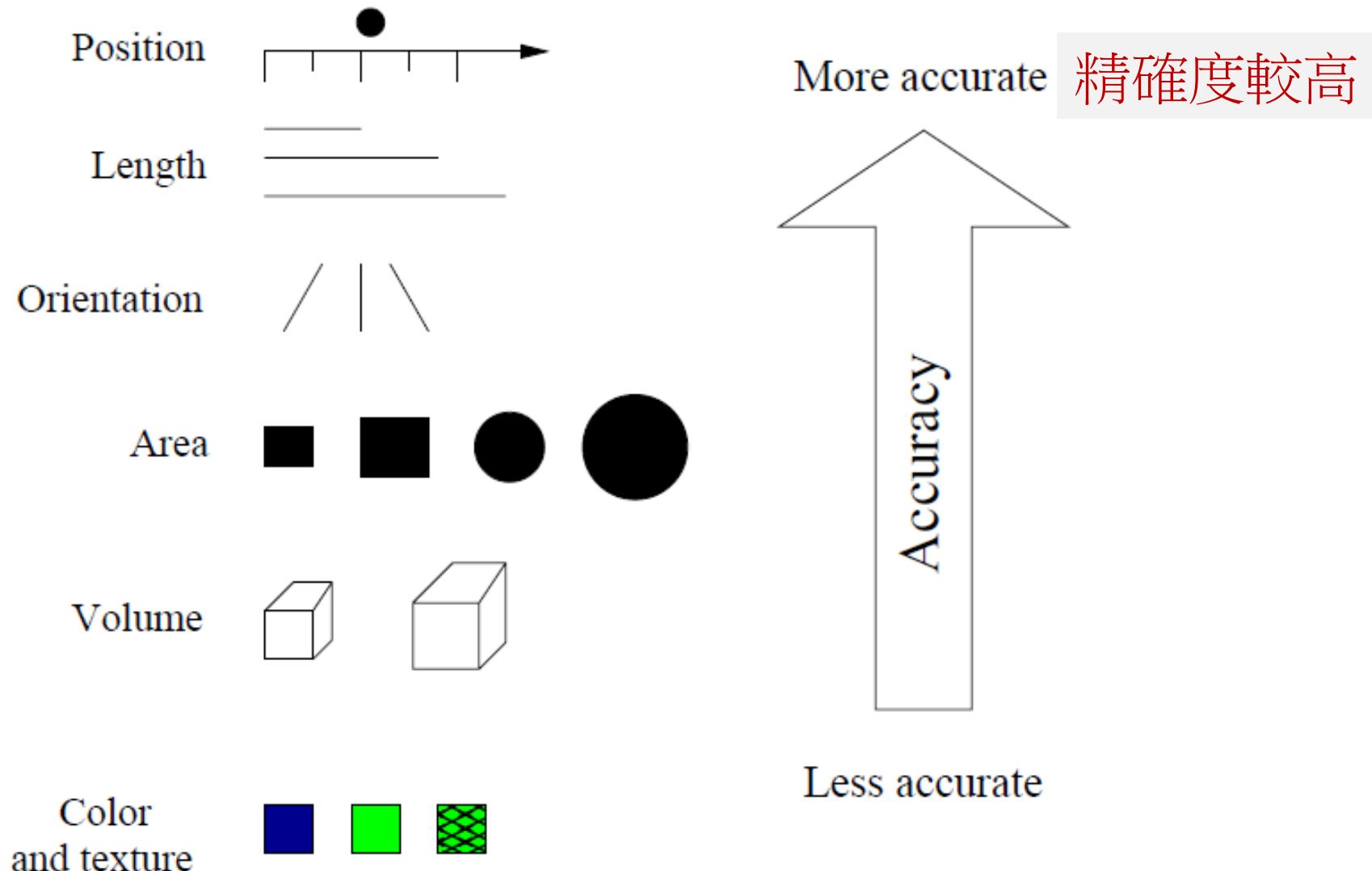
Shape



**Graphical elements**

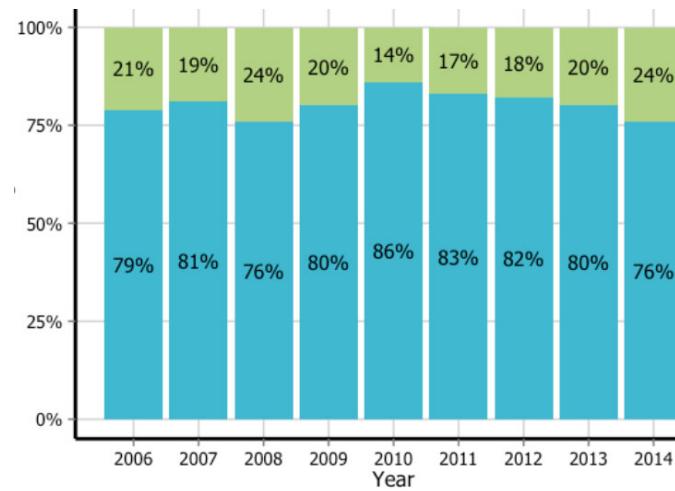
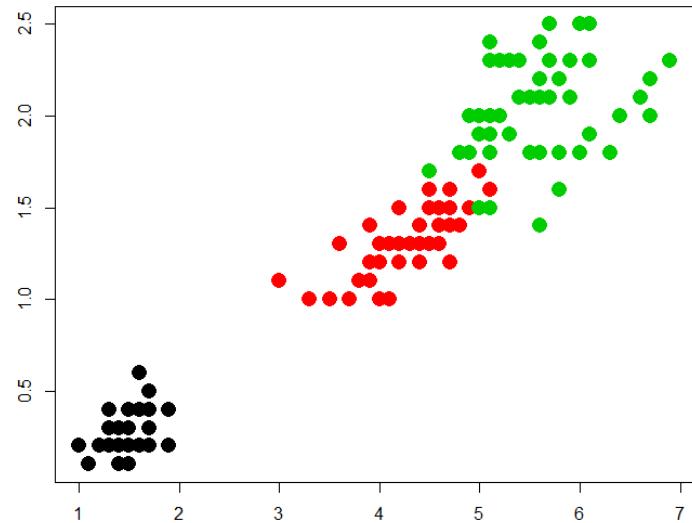
**Graphical properties**

# 視覺化繪圖結構(續)



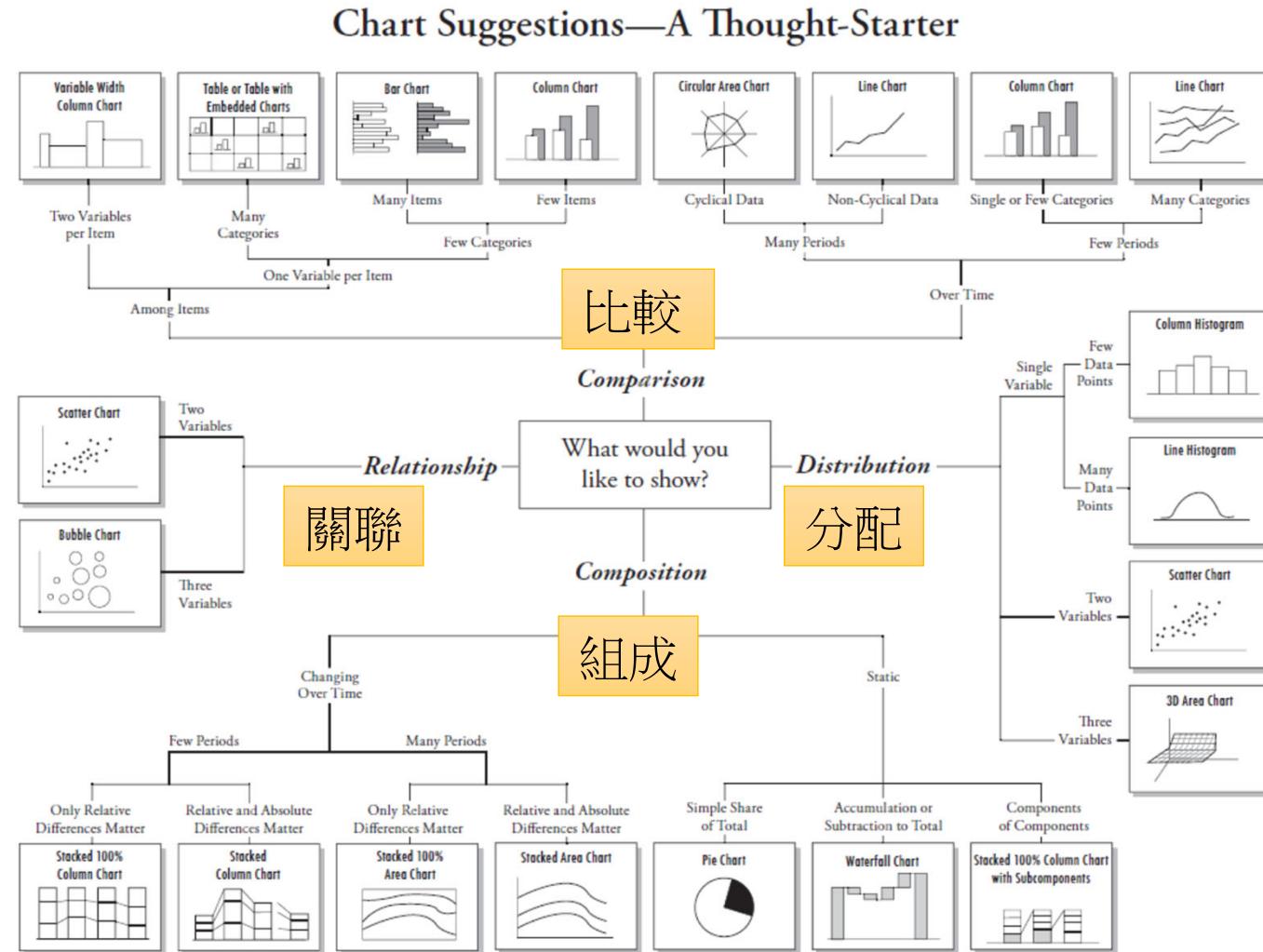
# 視覺化結果 - 四大繪圖類型

- 比較 Comparison
- 組成 Composition
- 分配 Distribution
- 關聯 Relationship



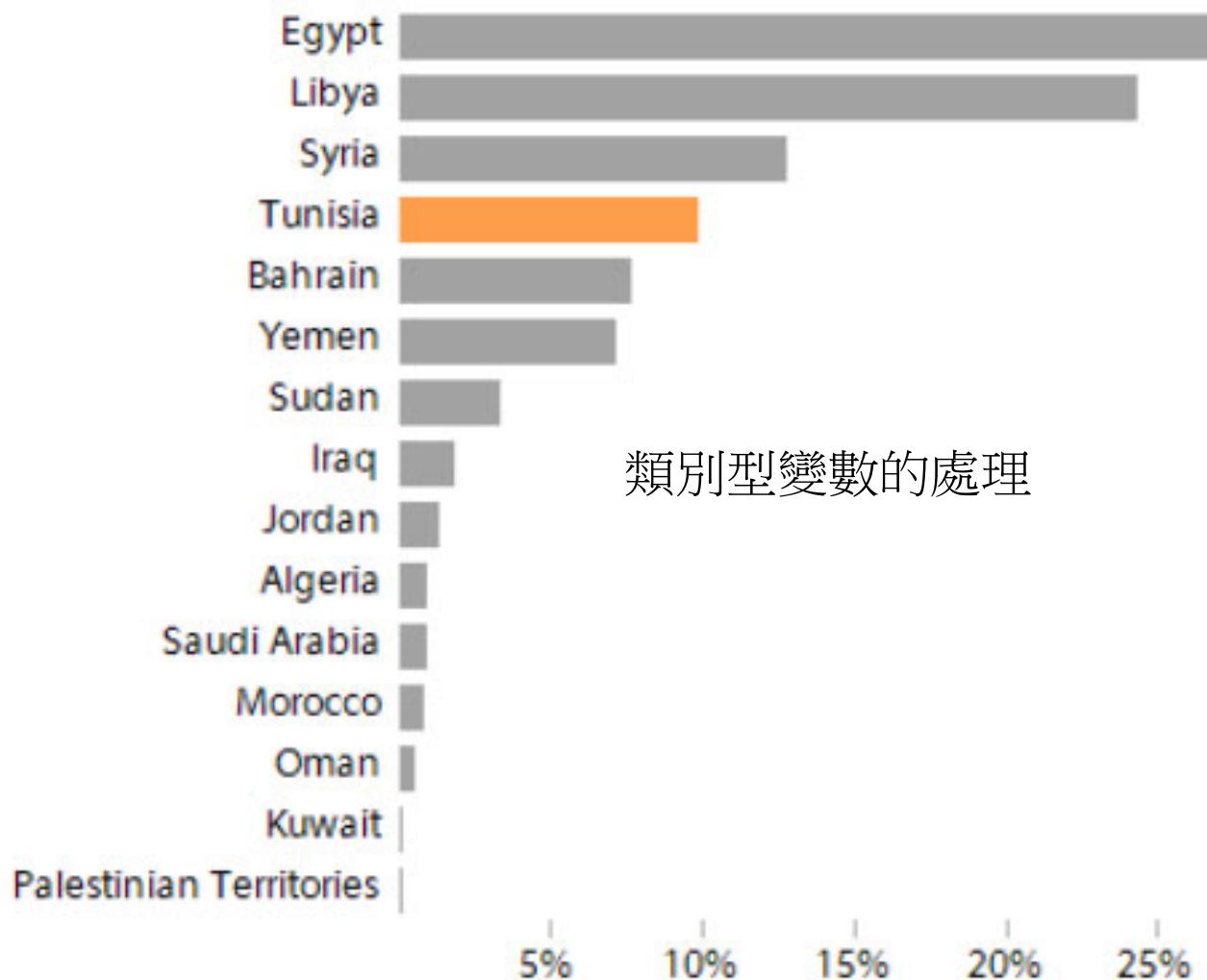
參考資料 <http://www.tatvic.com/blog/7-visualizations-learn-r/>

# Chart Suggestions—A Thought-Starter

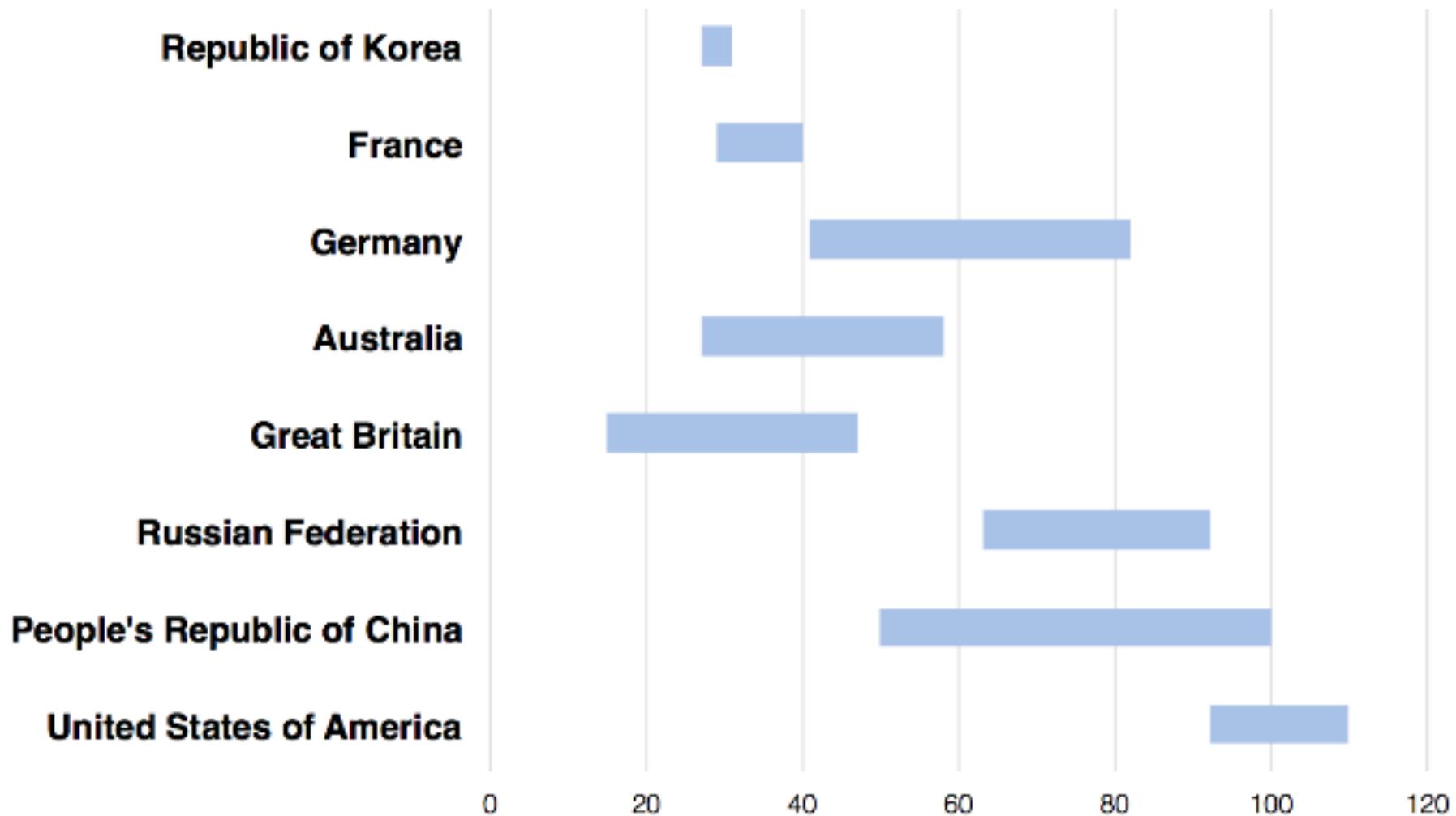


參考資料 <https://extremepresentation.typepad.com/files/choosing-a-good-chart-09.pdf>

# 資料視覺化的各類應用



# 類別型變數的處理

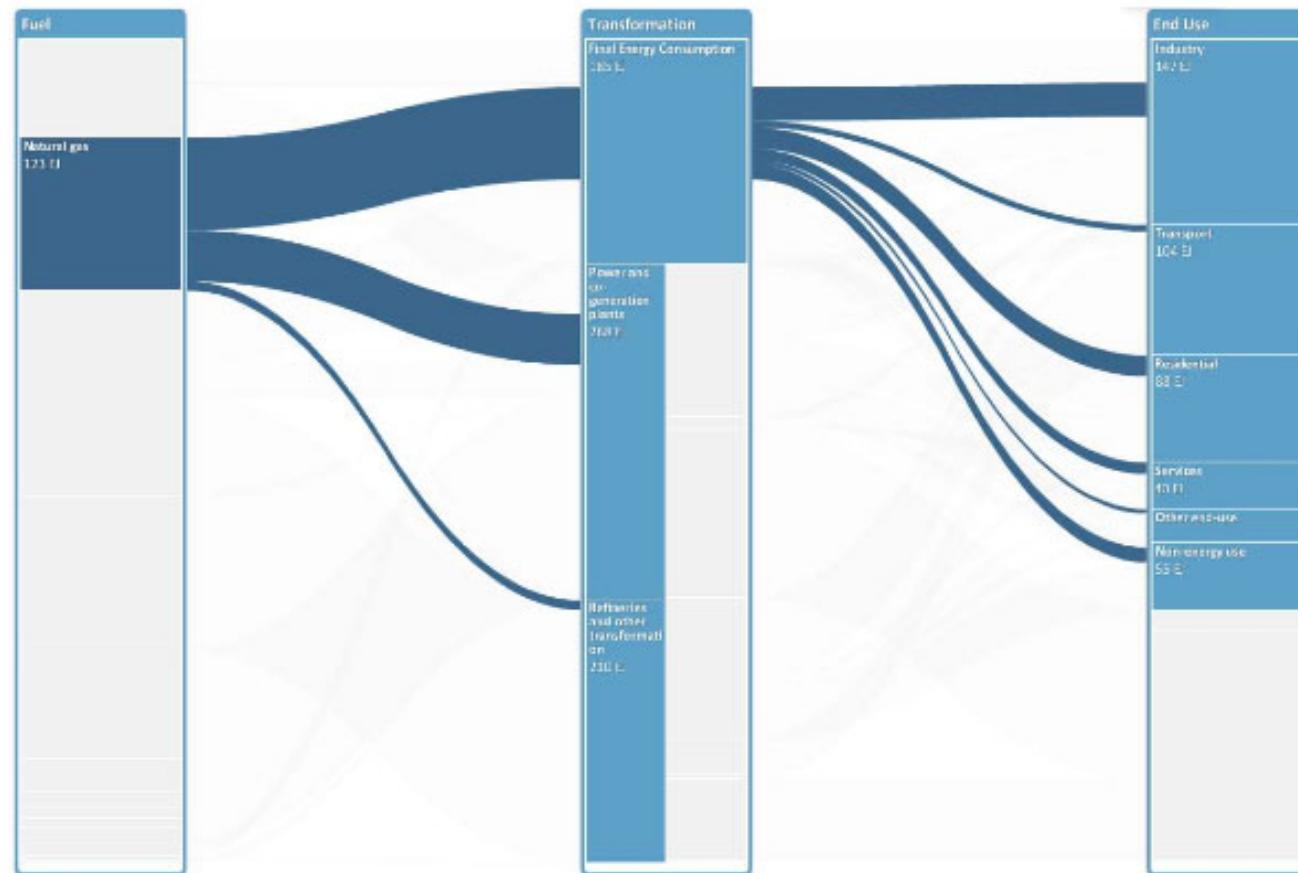


# 類別型變數的處理

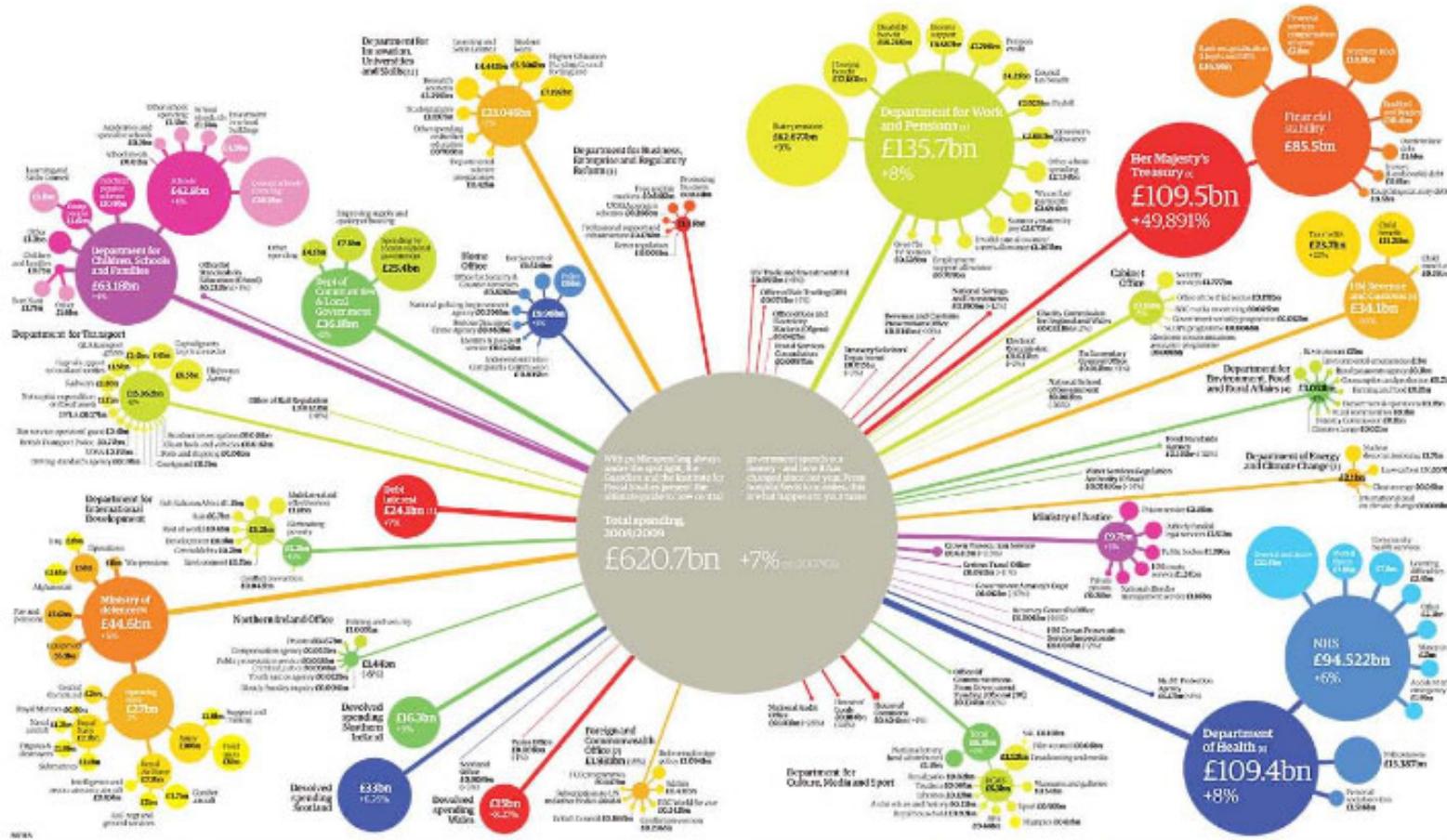
# 層次結構的展示

Fast Start for Anthony and 8-1 Start for Knicks	Toronto Blue Jays hire John Gibbons as manager   Baseball	San Francisco approves public nudity ban; protesters strip down	Driver of Midland parade float described as 'in shock'	Wen Jiabao adamant on South China Sea claim	Mumbai attacker executed in India	Humor in the Burst Tel Aviv Bubble
NHL awaits economic proposals	Boxer Macho Camacho shot in Puerto Rico	Kevin Durant, NBA players shamed by D-1 player Jack Taylor's 138 points	South Carolina tax chief resigns after report reveals 4 million taxpayers hacked	Panetta says US must press fight against A-Codes	EU mayor to honor Twitter food stamp challenge	Three people killed in central Kabul suicide bombing
Barclays Center Brooklyn, NY	Saint Louis can't overcome 10s in loss to Kansas	Hiroki Kuroda returns to Yankees for \$15 million	Wireless Shooting victim pleaded, Please, not Please!	San Diego media figure dies after car crash	One Thanksgiving Day Parade, They're Not Having That!	John McAfee, Unhinged: His Bizarre Breaks From Reality
Luck by chance	Erin Brockovich and Apple Askin On a Collision Course With the Devil	The 'Voice' recap: Crazy Eight	Toys safer than ever, but a few dangerous	Rip current has already arrived in Alaska, state says	Oliver Holt column Worldwide Becks appeal: The global opportunities available ...	Gatland: I could have lost a leg
With 'Breaking Dawn,' I Remember My Love At First Bite	Guardians' rise to the occasion	NBC Wins November Ratings for First Two Weeks of 2012, Tele視視	Doctors: Birth control pills should be sold over the counter	Karzai Gives His Approval for Execution of Prisoners	Method behind Ronaldo's misery	Method behind Ronaldo's misery
Black Friday Protesters to Take on Wal-Mart	Japan trade deficit jumps as exports to China fall	Best Buy Romin: Dismal as Holiday Season Begins	Court approves FTC's \$22.5 million penalty for Google over Safari cookies	New Zealand's Prime Minister Declares a Hobbit State	FOREX-Euro slumps after no Greece deal, yen falls broadly	Eurozone finance ministers fail to agree Greece bailout
	Hostess Judge to Weigh Shutdown After Mediation Fails	US Treasury Allows Major Asian markets	AU jury awards W man in first trial of March 2012	Energy proposals may push up bills	Ronaldo scores over 600 goals in his career	Ronaldo scores over 600 goals in his career
		Japan says off-shore trade related			Tolkien Estate sues over Lord of the Rings slot machines	Tolkien Estate sues over Lord of the Rings slot machines

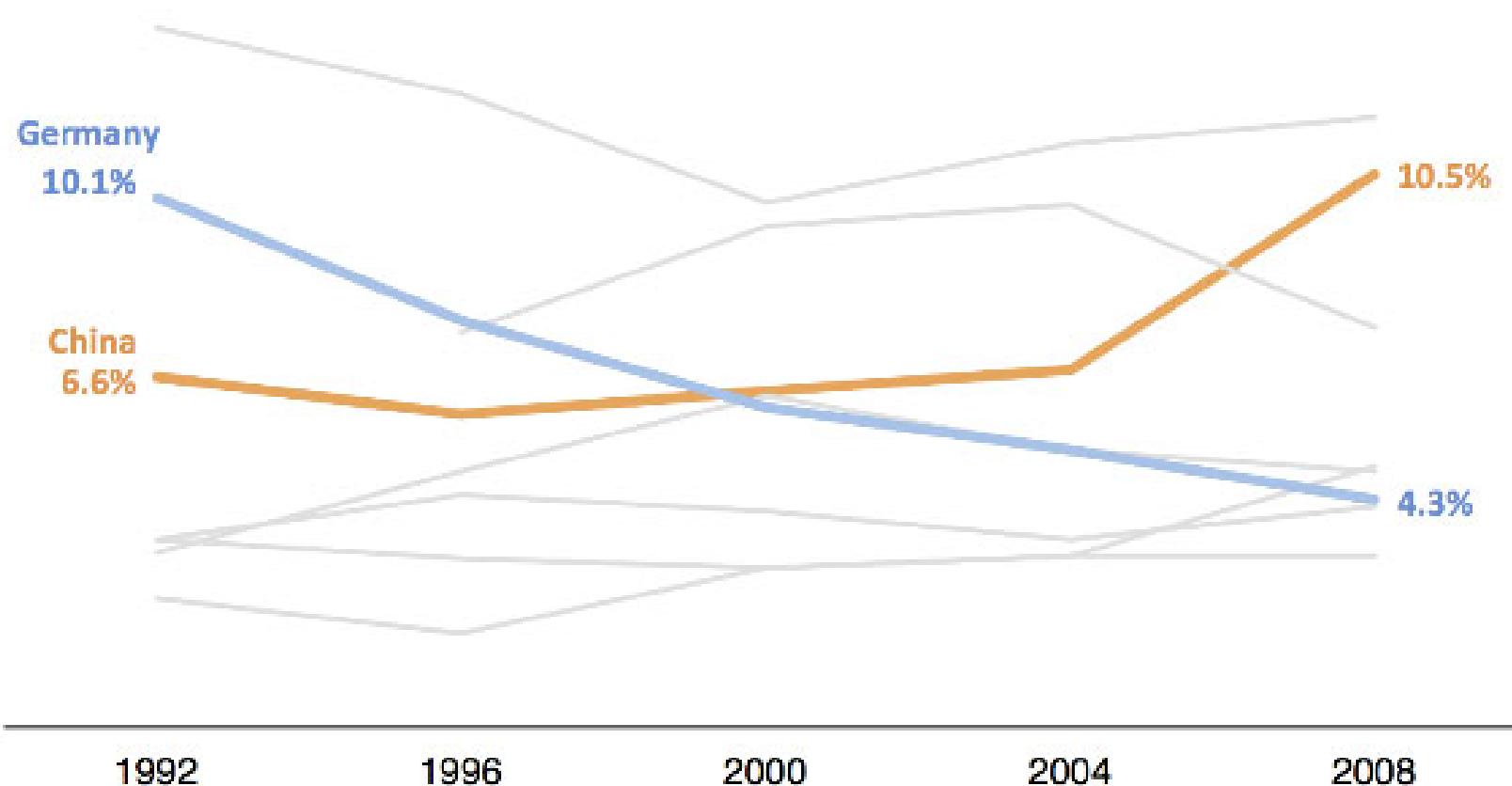
# 層次結構的展示(續)



# 層次結構的展示(續)



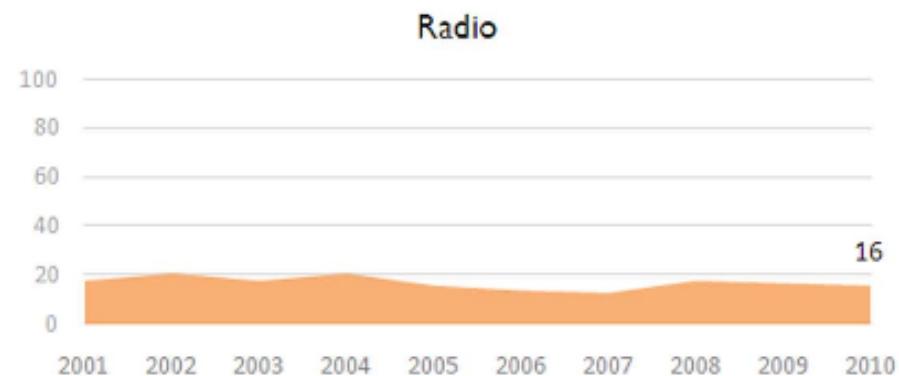
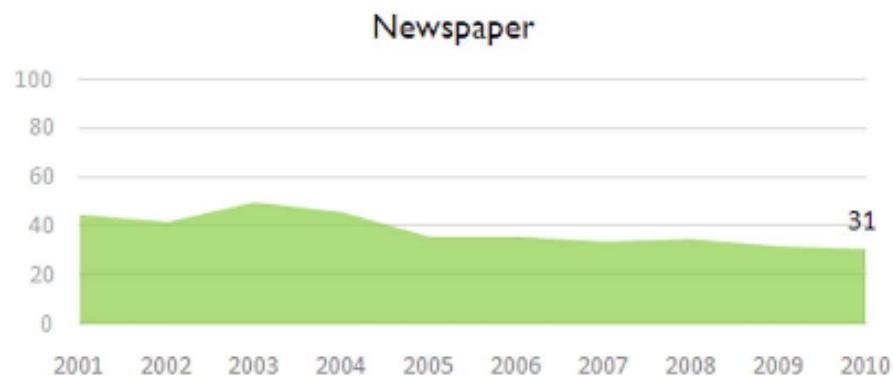
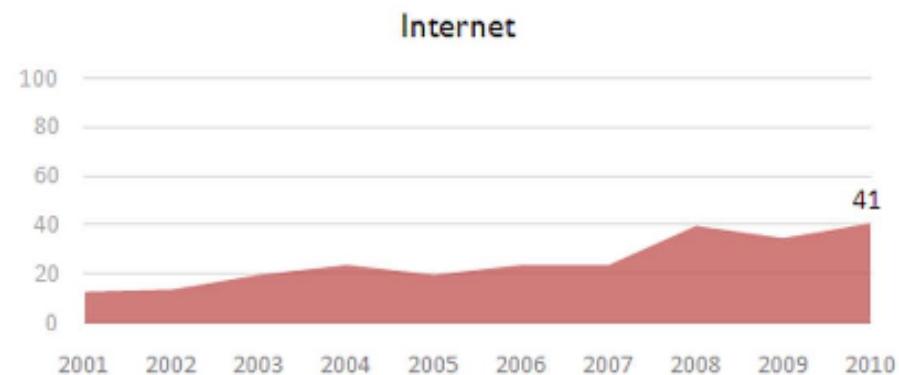
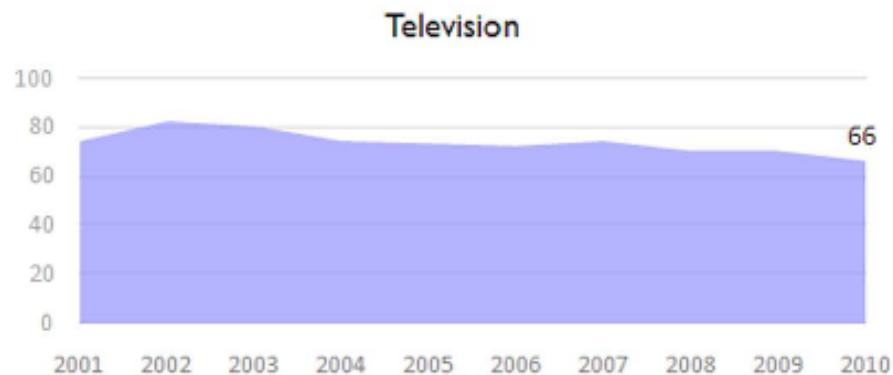
# 視覺化案例：顯示時間的流動



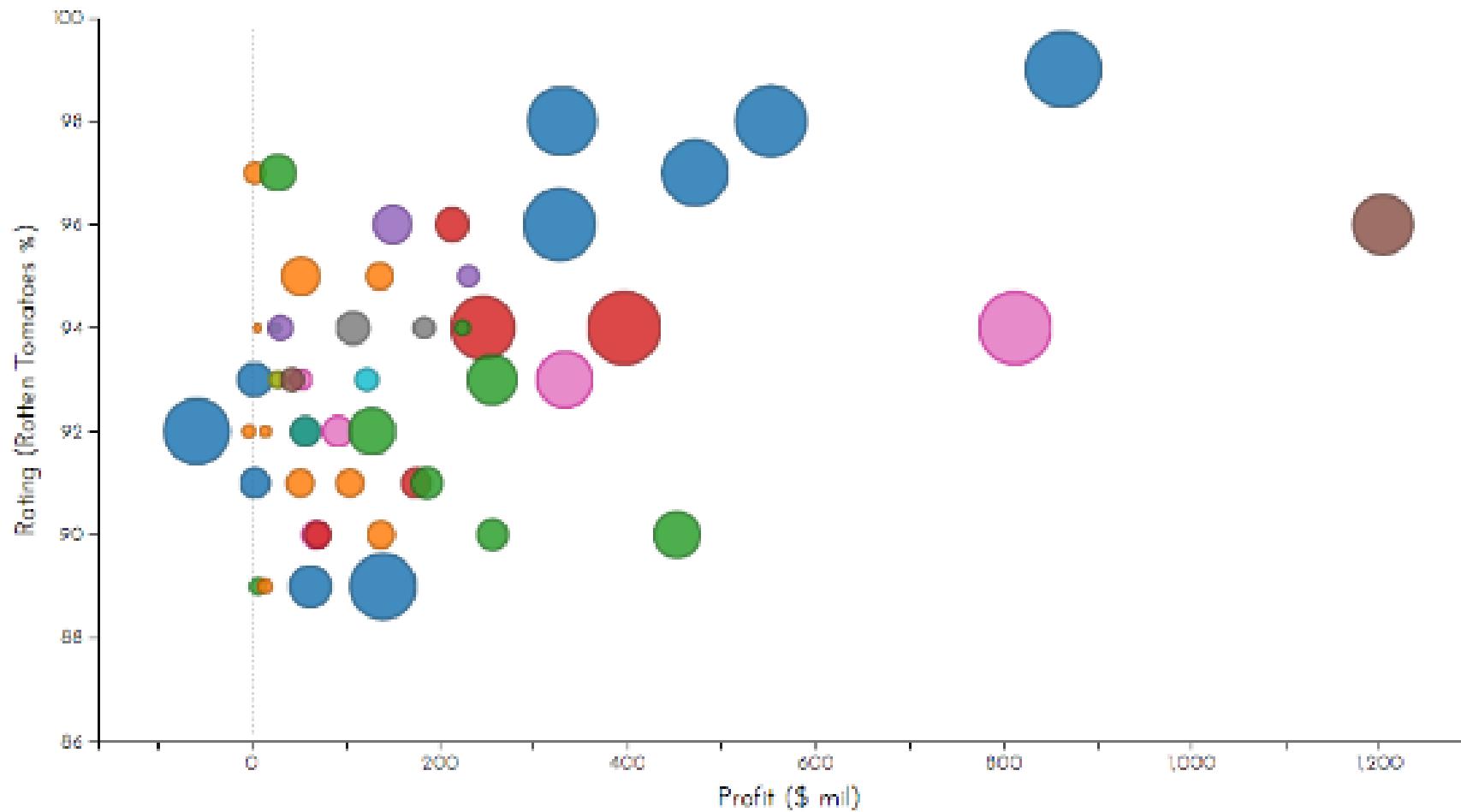
# 視覺化案例: 顯示時間的流動(續)



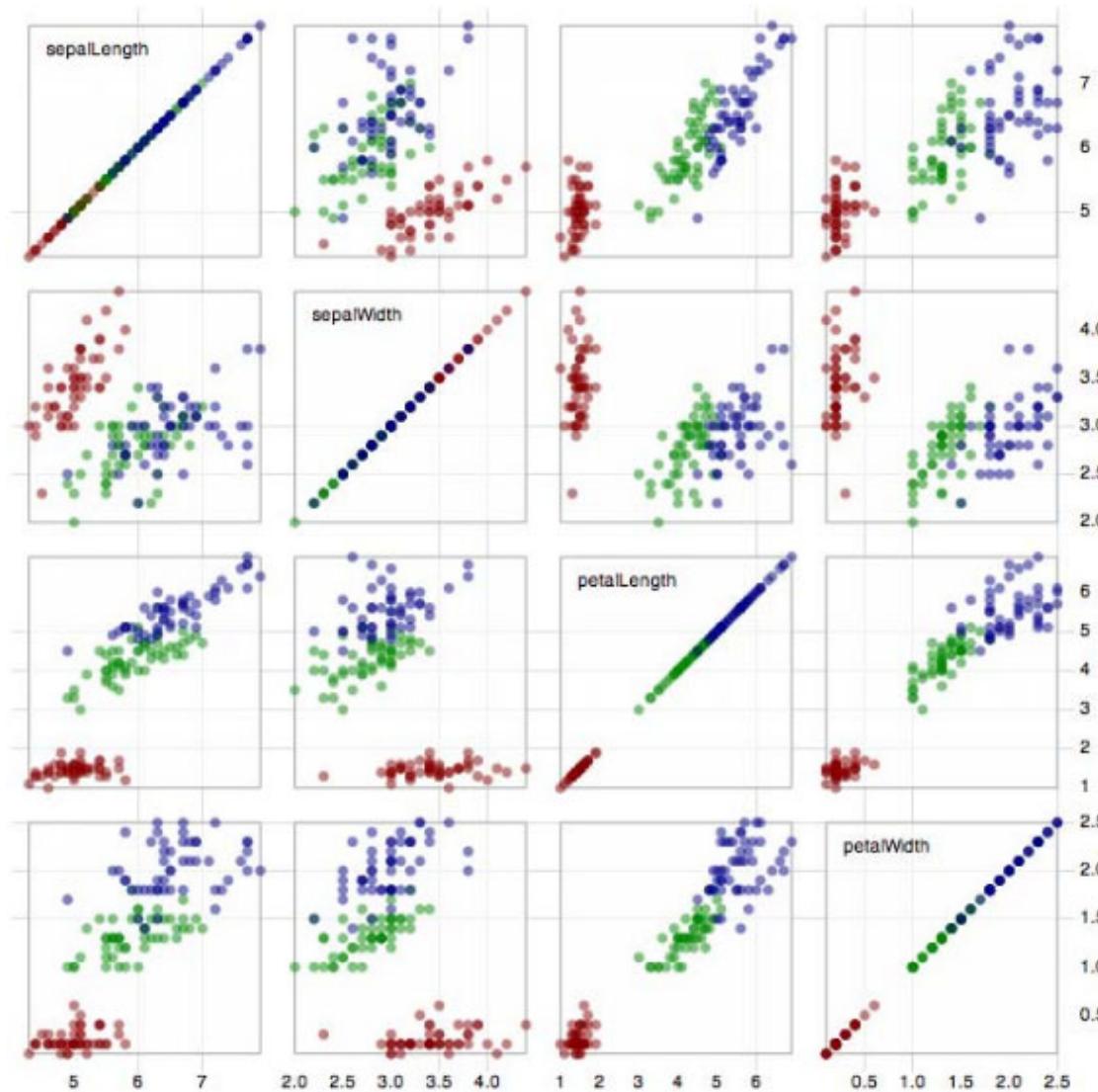
# 視覺化案例：顯示時間的流動(續)



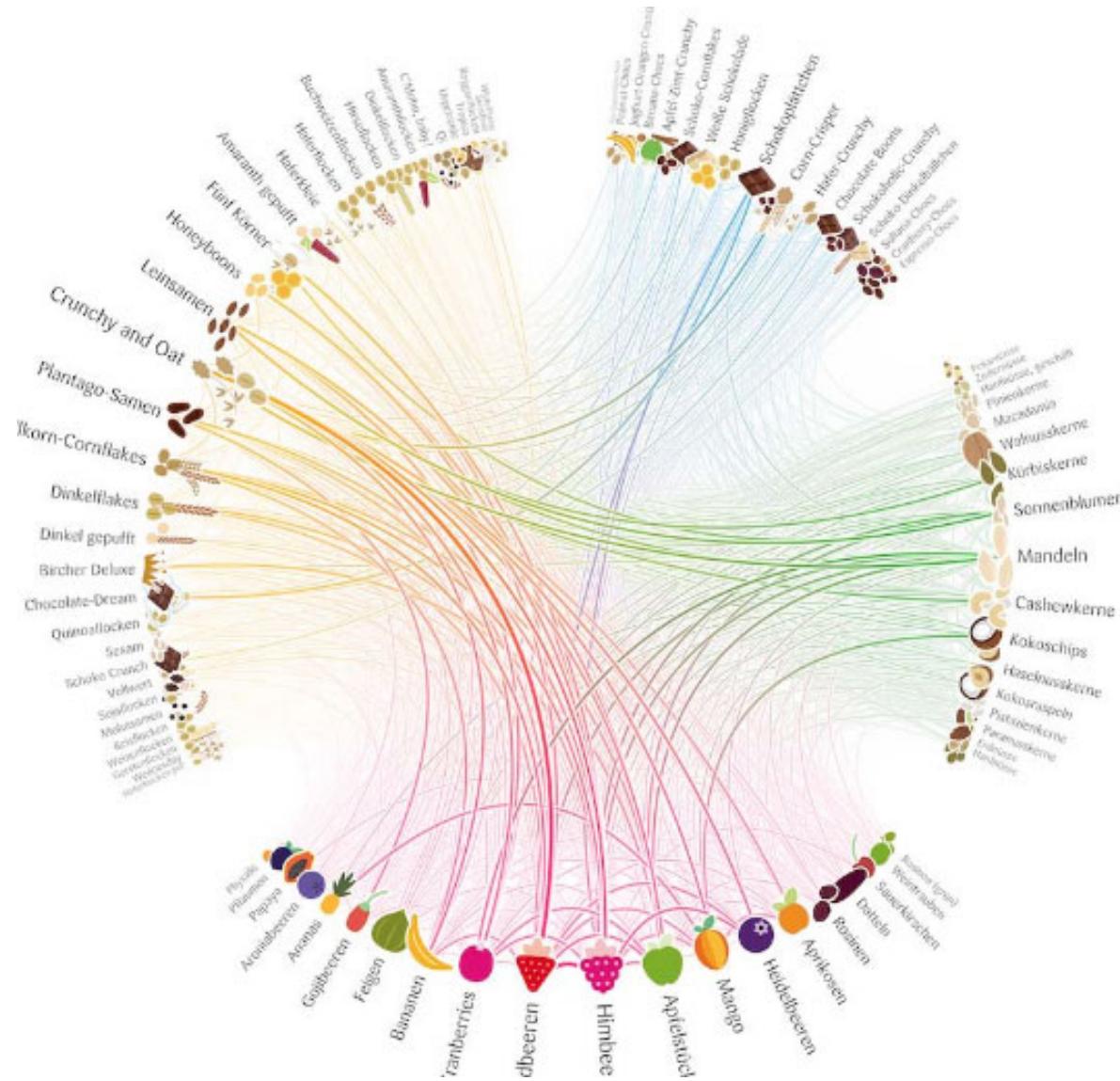
# 視覺化案例：顯示關係的強弱



# 視覺化案例: 顯示關係的強弱(續)

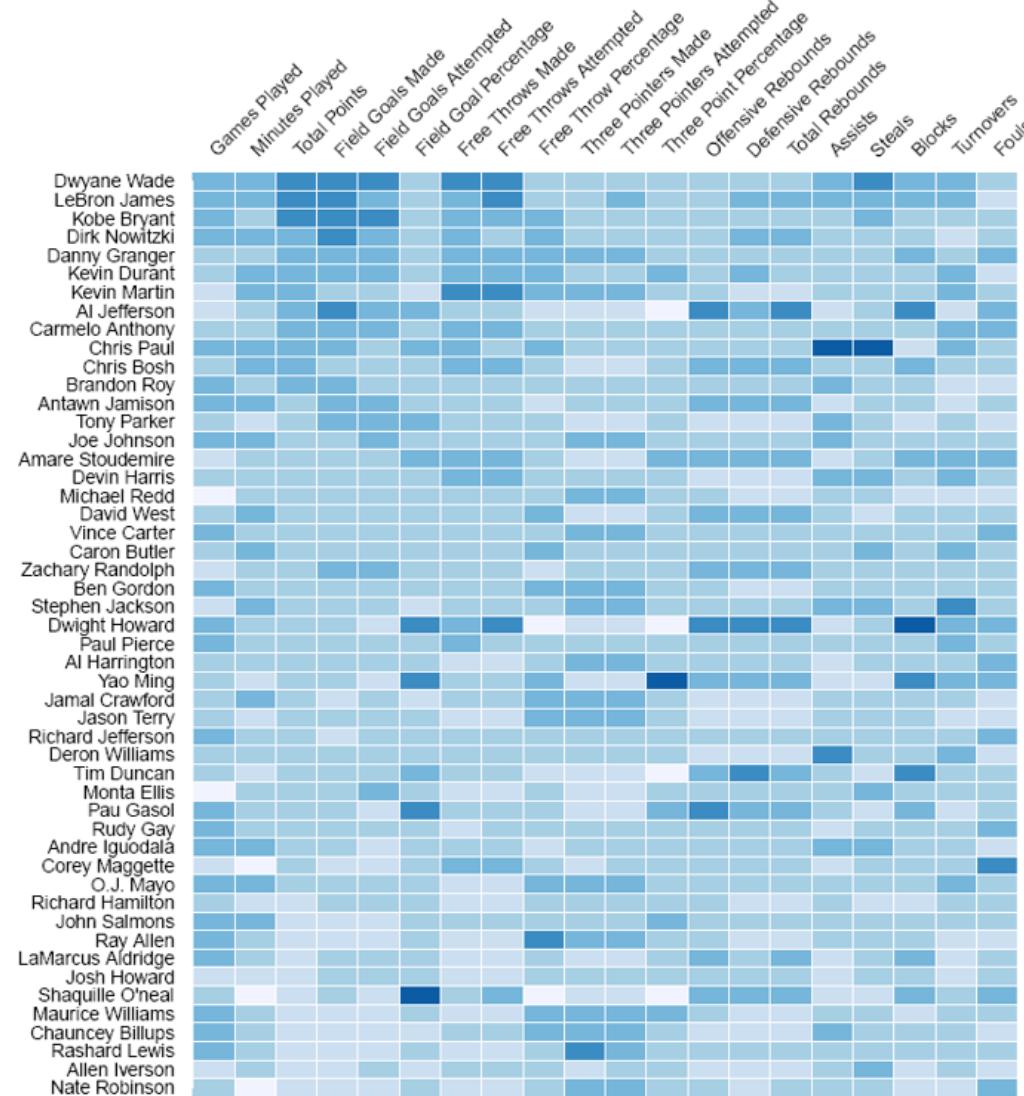


# 視覺化案例：顯示關係的強弱(續)

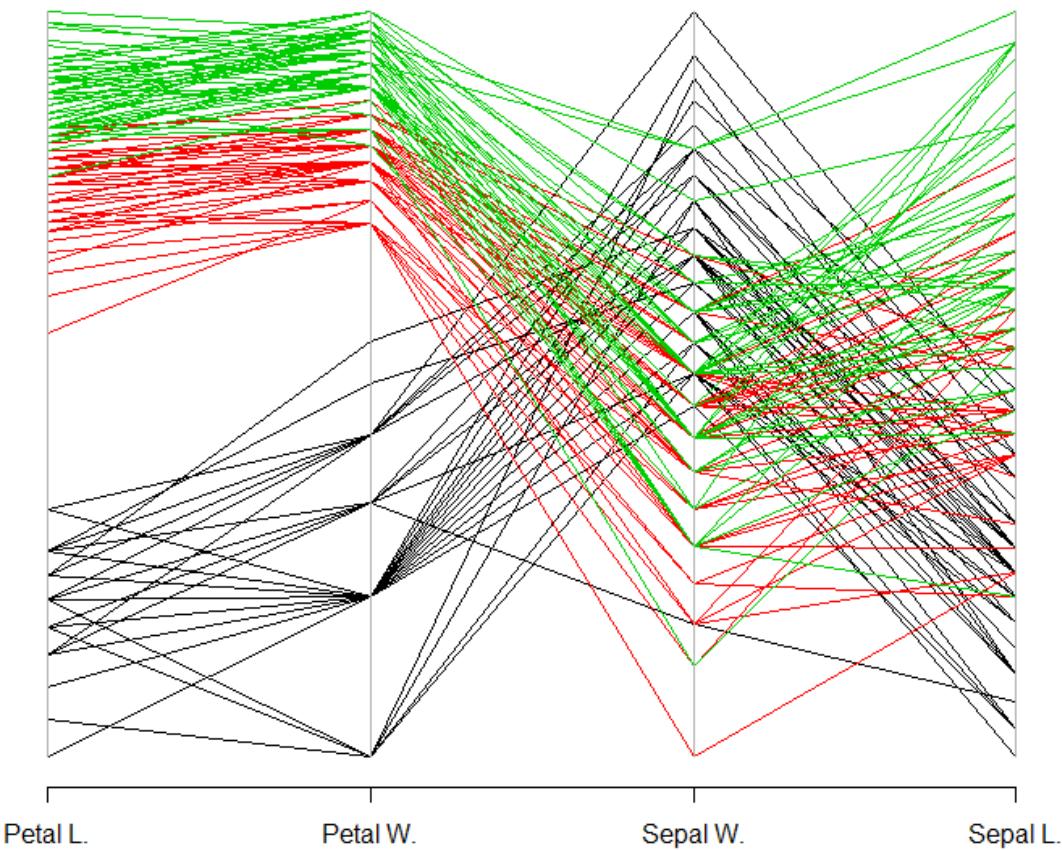


## 熱繪圖

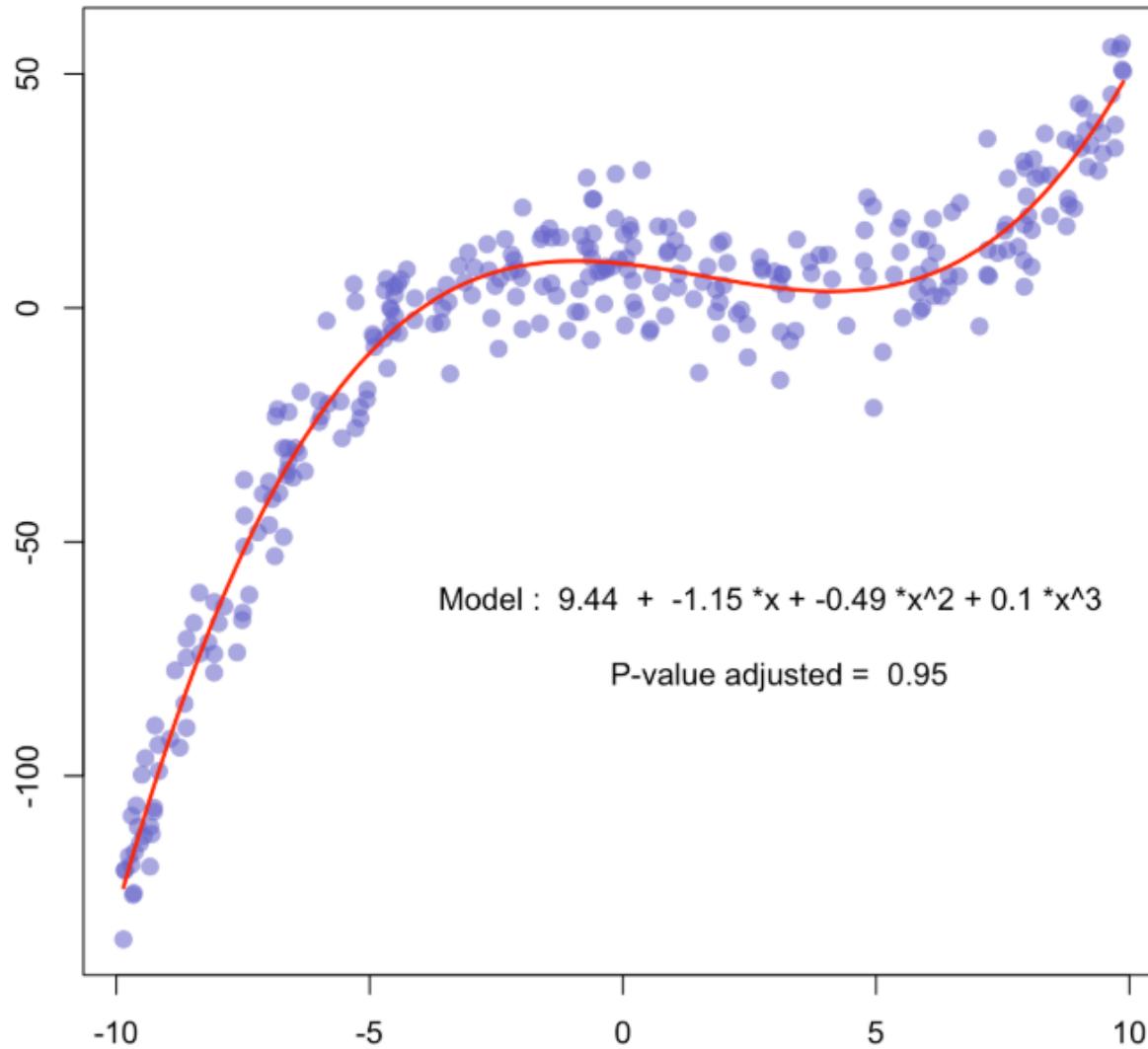
## Heatmap



# 平行座標軸 Parallel coordinates

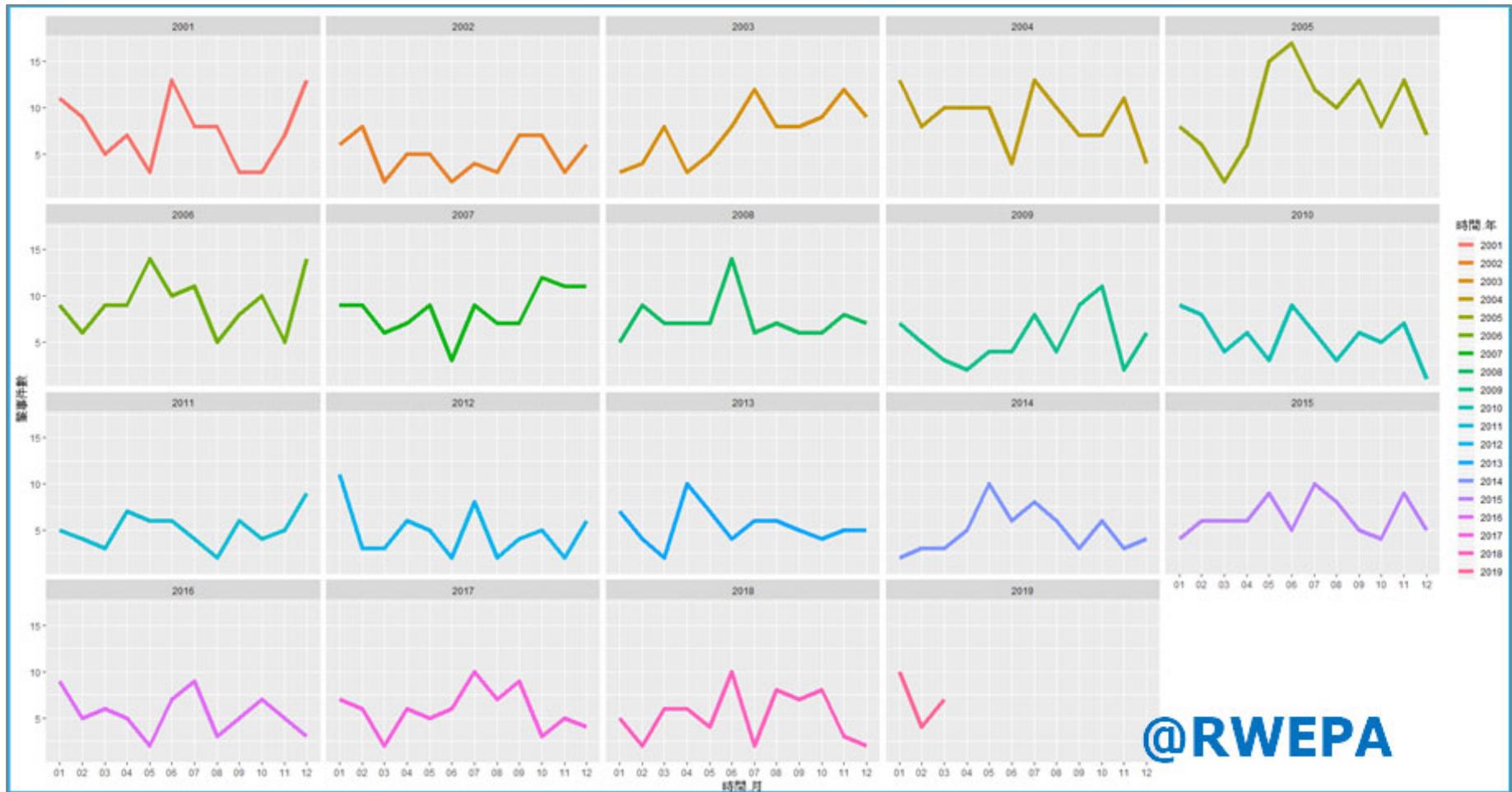


# Polynomial curve fitting



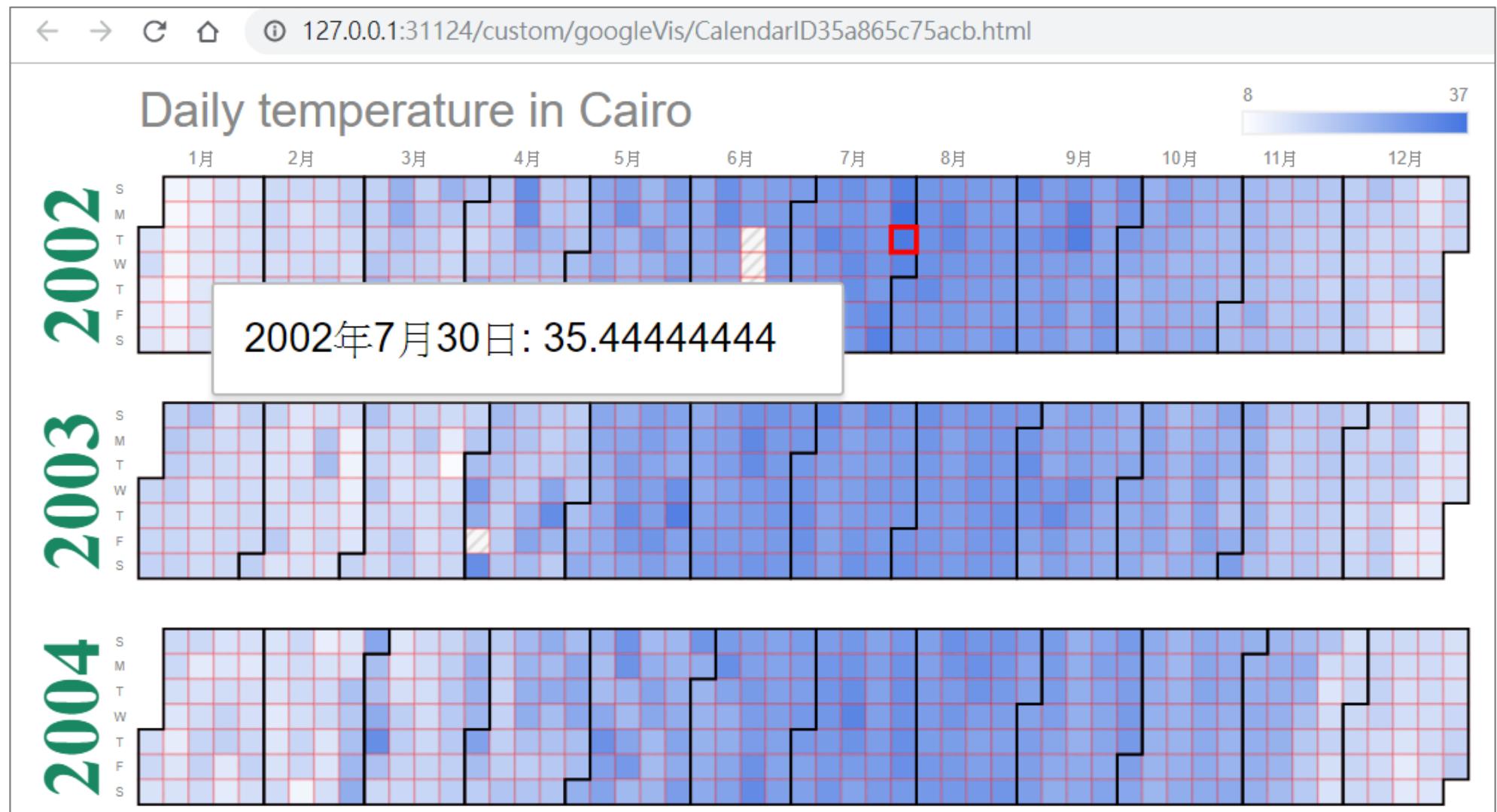
參考資料 <https://www.r-graph-gallery.com/all-graphs>

# 時間線圖 2001~2019 高速公路資料集



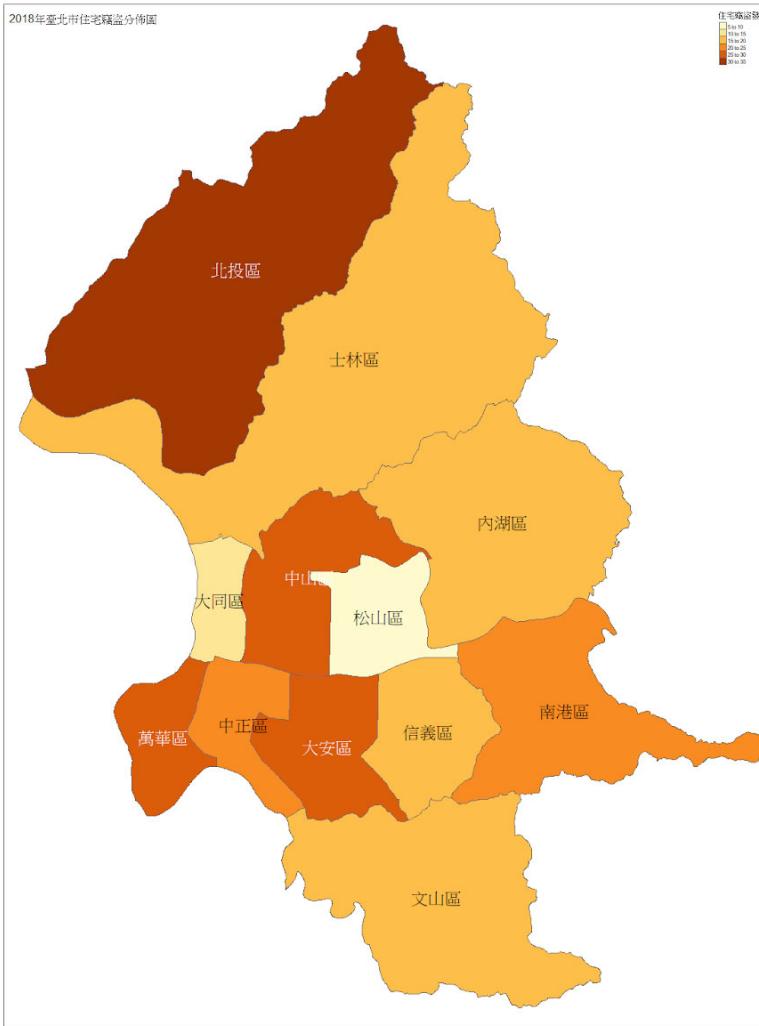
參考資料 <http://rwepa.blogspot.com/2019/05/highway.html>

# 月曆圖 Calendar



# 主題式地圖(Thematic map)

## 2018年臺北市住宅竊盜分佈圖



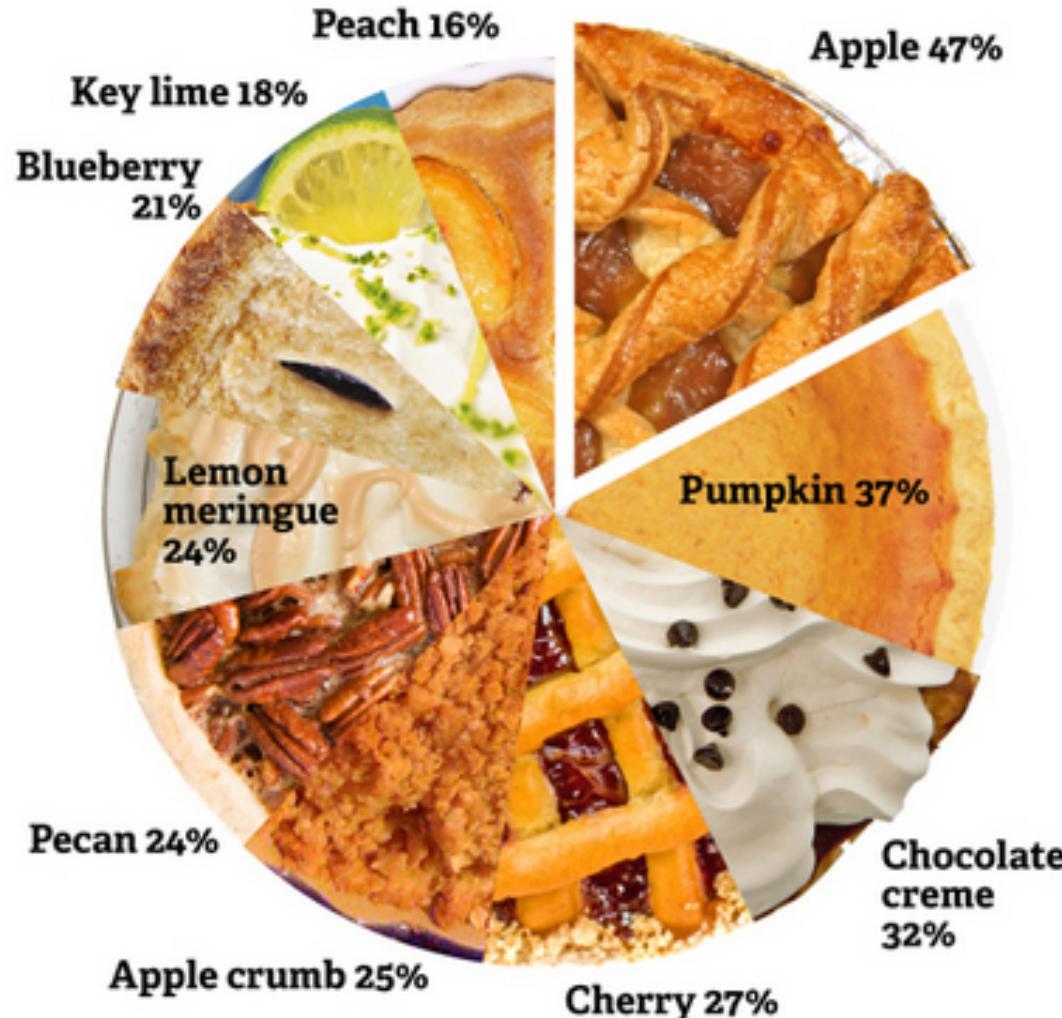
參考資料 <http://rwepa.blogspot.com/2018/10/thematicmap.html>

# 視覺化的陷阱

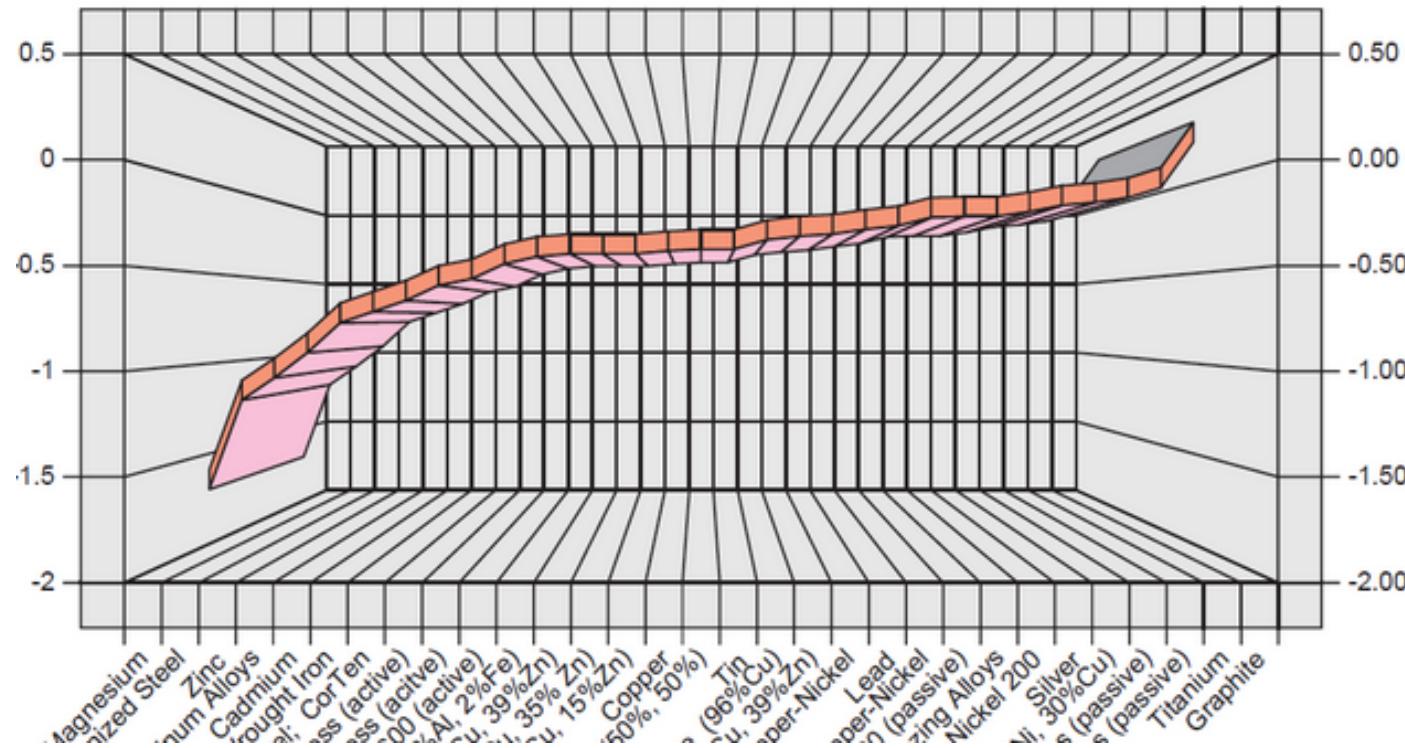
- 避免使用太複雜三維圖形
- 圓形圖處理的類別數目不可過多
- 避免使用漸變色
- 避免使用陰影

# 不佳的視覺化

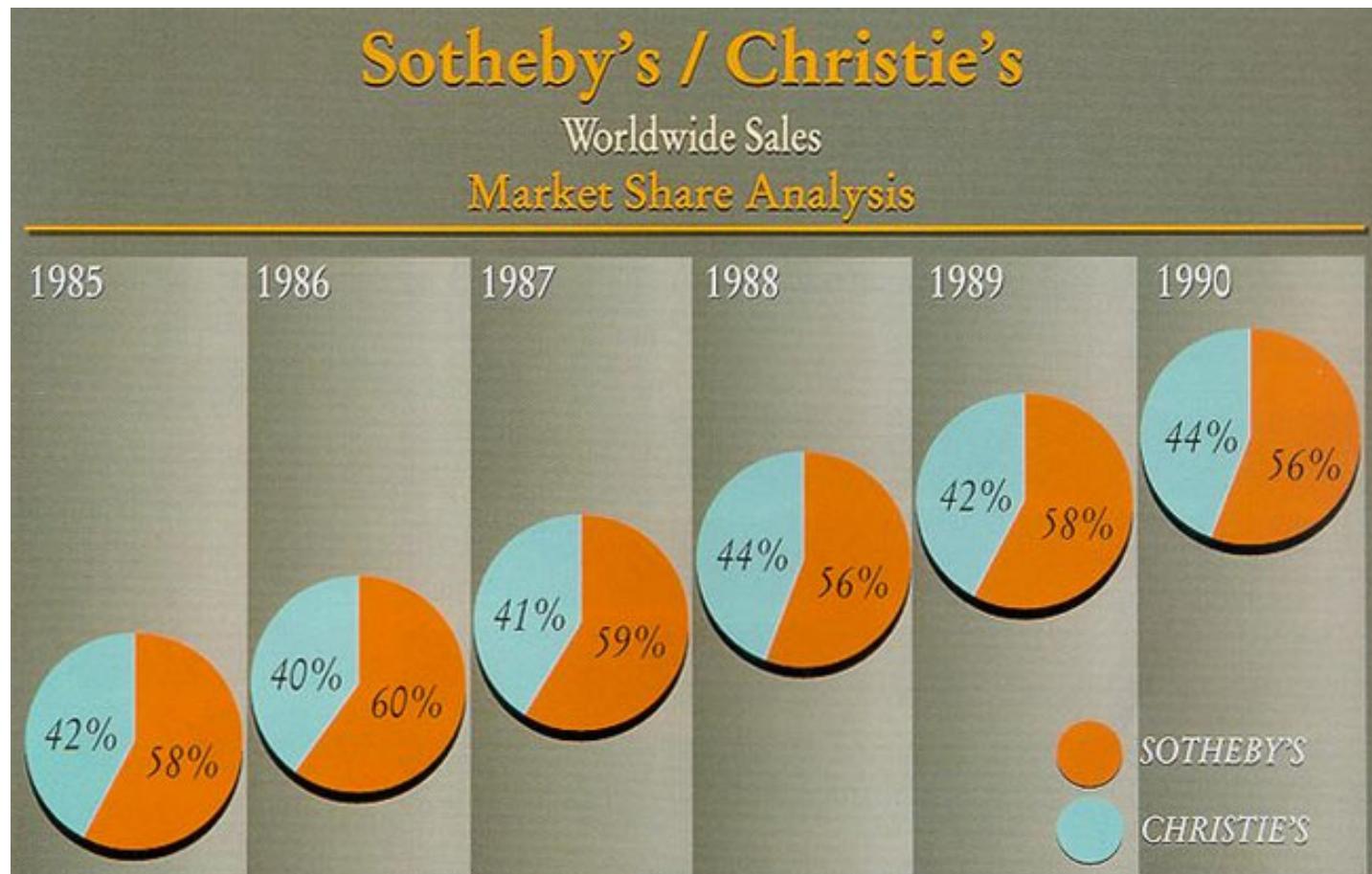
**What are your three most favorite types of pie?**



# 不佳的視覺化(續)

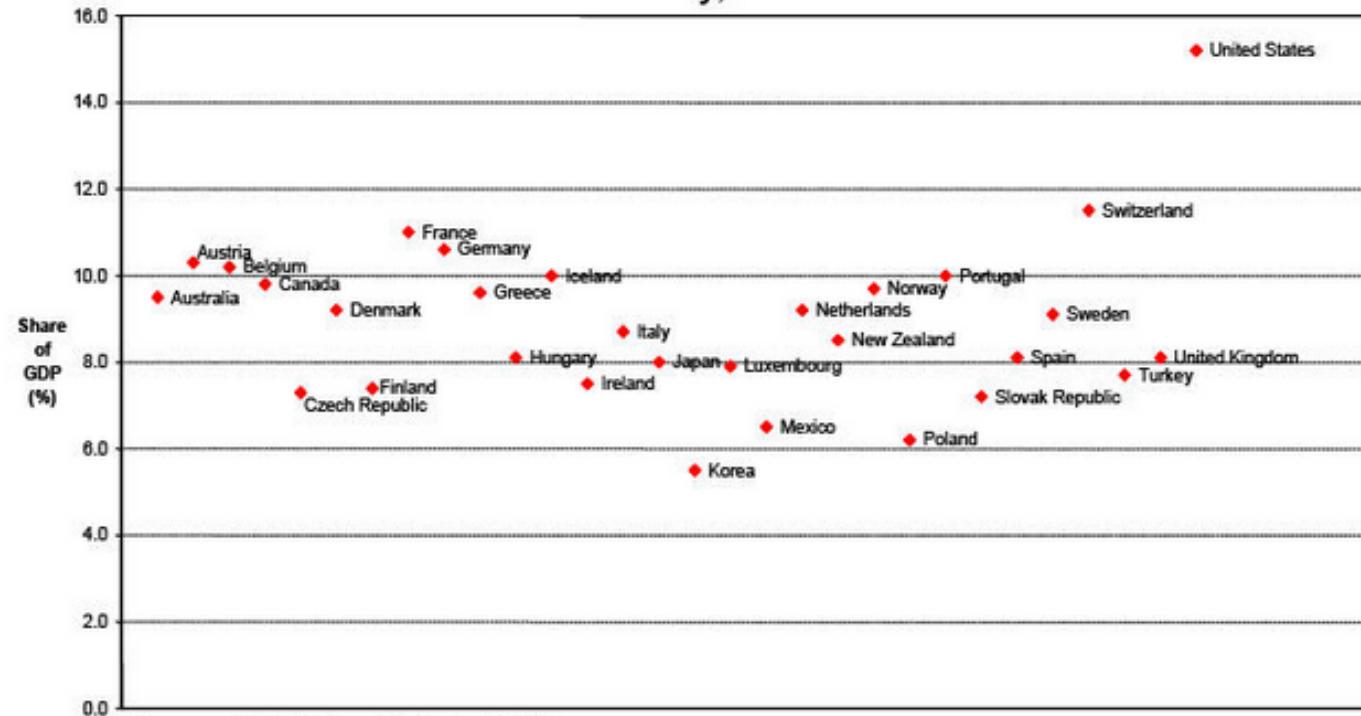


# 不佳的視覺化(續)



# 不佳的視覺化(續)

Chart 2 - Total Expenditures on Health as a Percentage Share of GDP, by OECD Country, 2004

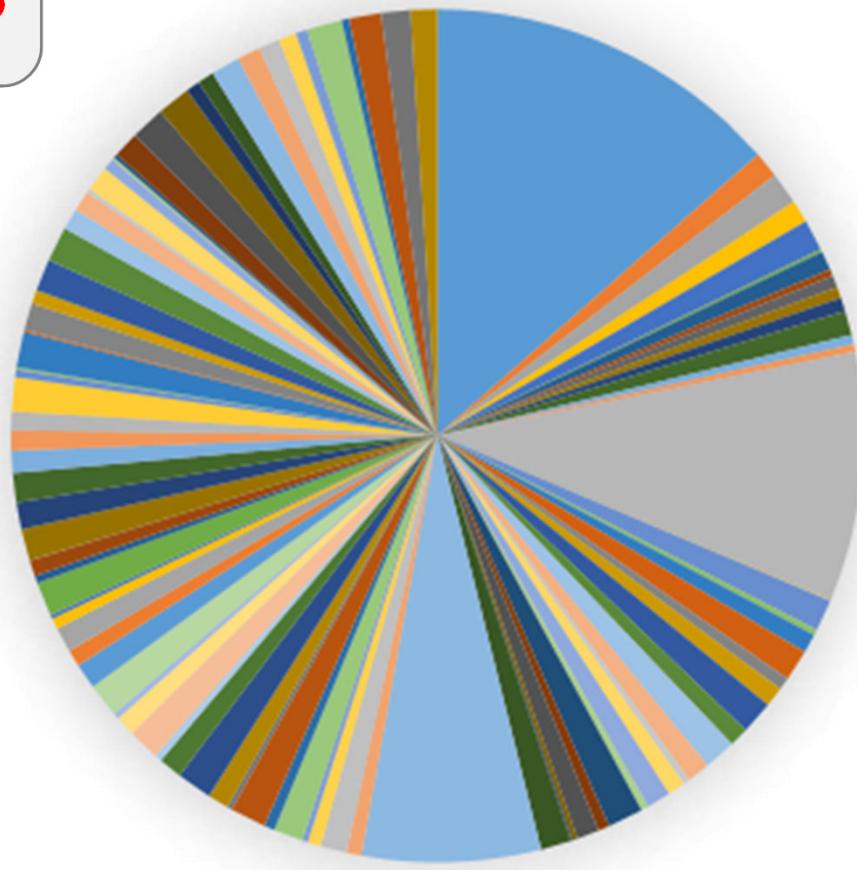


Source: OECD Health Data 2007.

Note: For the United States the 2004 data reported here do not match the 2004 data point for the United States in Chart 1 since the OECD uses a slightly different definition of "total expenditures on health" than that used in the National Health Expenditure Accounts.

# 不佳的視範化(續)

不知在  
視覺化什麼?



## 2.Tableau簡介

# Tableau 新時代視覺應用

- Tableau Software 美國互動式資料視覺化軟體公司，由史丹佛大學電腦科學系 **Christian Chabot** (VizQL技術), **Chris Stolte**(資料庫), **Pat Hanrahan** (算圖引擎) 於2003年1月在加利福尼亞州山景城創立。該公司目前總部位於美國華盛頓州西雅圖市，專注於商業智慧。
- 2019年8月1日，Salesforce.com收購了Tableau。



Christian Chabot



Chris Stolte



Pat Hanrahan

# Tableau - 視覺化創新應用

- 圖形化操作介面
- 多樣化的資料(資料庫)連結
- 輕鬆建立視覺化圖表
- 資料過濾與資料層級剖析
- 快速分享研究結果
- API應用 (C, C++, Java, R, Python, JavaScript,...)
  - 使用R/Python 統計分析功能
  - 使用R/Python建立進階模型 (機器學習, 深度學習)
  - 將R/Python模型回傳至 Tableau , 進行進階視覺化應用

# Tableau安裝

- 產品版本 <https://www.tableau.com/zh-tw/products/desktop>



透過 Tableau Desktop，以思維的速度回答問題

免費試用

查看其運作原理

Tableau Desktop

Tableau Prep

Tableau Server

Tableau Online

Tableau Data Management

Tableau Server Management

Tableau Mobile

內嵌式分析

開發人員工具

最新版本

定價

立即購買

@RWEPA

直接下載 Tableau Desktop 14天適用版:  
<https://www.tableau.com/zh-tw/support/releases>

# 工作表 (Worksheet)

Tableau - Superstore [Read-Only]

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

功能表列

工具列

智慧圖形

資料/分析

分頁

過濾

標記

選取頁面

工作表視圖區

Show Me

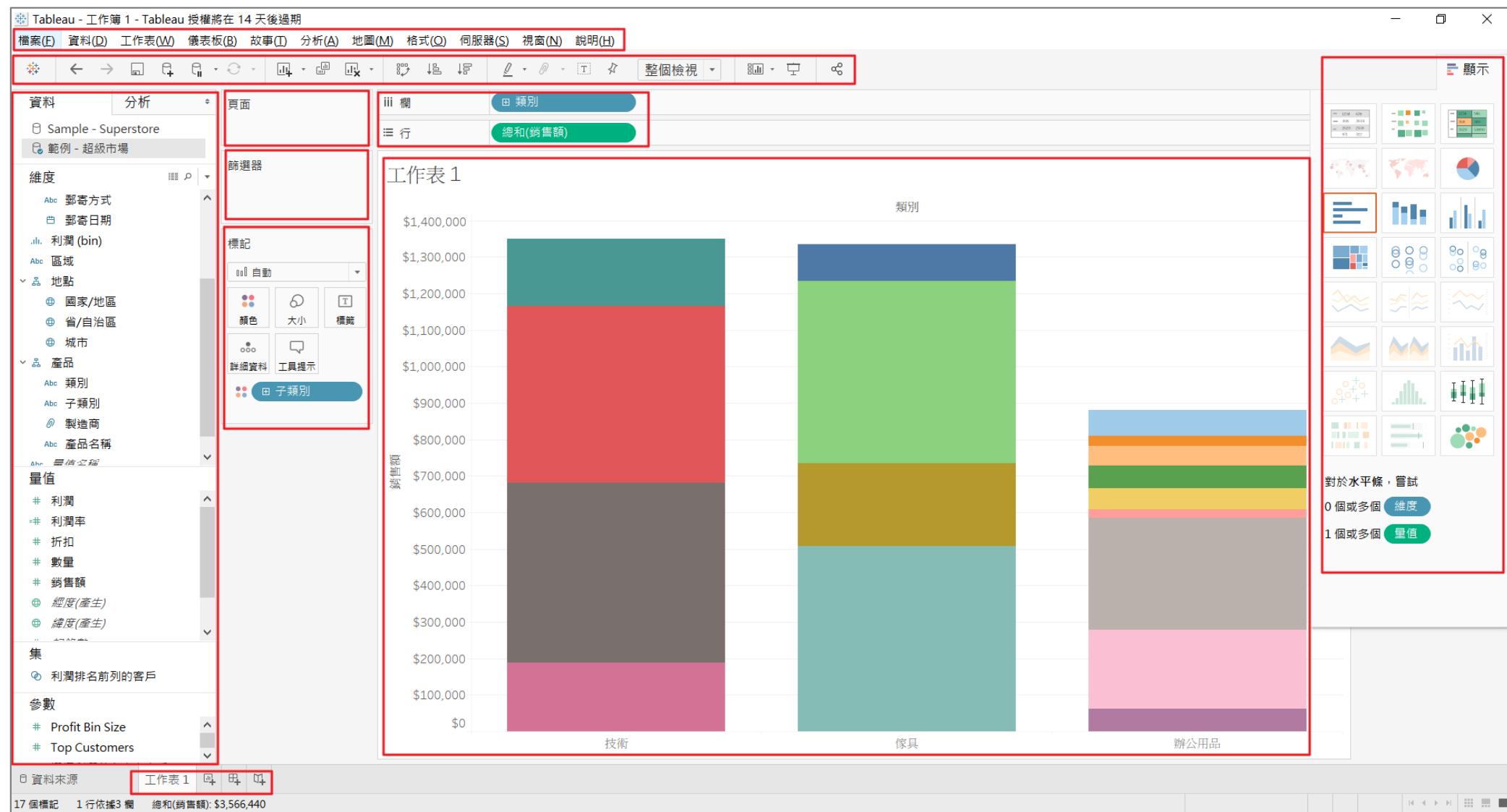
1. Worksheet 工作表  
2. Dashboard 儀表板  
3. Story 故事

City	Profit Ratio (%)
Danville	-1.0%
Missouri City	-5.5%
Deer Park	-35.0%
Littleton	-160.0%
Tyler	-150.0%
Champaign	-150.0%
Waco	-136.7%
Lancaster	-127.8%
Ormond Beach	-120.0%
Hendersonville	-85.8%
Pensacola	-79.3%
Bethlehem	-70.0%
Yuma	-67.6%
Medina	-66.7%
Pueblo	-59.4%
Park Ridge	-55.4%
Clarksville	-54.2%
Cuyahoga Falls	-53.4%
Skokie	-53.0%
Tinley Park	-47.6%
Rockford	-43.2%
Boca Raton	-42.3%
San Antonio	-40.0%
Frisco	-38.6%
Bryan	-35.9%
Orain	-34.4%
Reading	-32.7%
	-32.5%
	-31.4%
	-30.8%

Data Source Overview Tooltip: Profit Ratio by City Product Customers Shipping Performance Commission Model Order Details Forecast What If Forecast

511 marks 511 rows by 1 column SUM of AGG(Profit Ratio): 5972.0%

# 工作表 (Worksheet) - 中文版



# 即時, 摳取



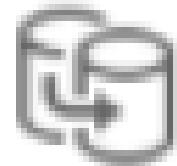
- 即時 Live:

- 即時連線會將查詢傳送至資料庫，並根據結果更新檢視結果。
- 但在最初建立連線時，會定義查詢的特定欄位。
- 即時一般不會將資料儲存於 Tableau server, 而是直接連結資料庫，一般採用此項目。



- 摳取 Extract:

- 主要用於連結靜態資料集。
- 將小量資料讀取至本地端處理。
- 資料將複製一份至 **Tableau server**。
- 重新整理擷取將查詢從其建立擷取的資料來源並重新組建擷取。根據擷取大小，此過程可能需要一些時間。



# 資料欄位管理, 排序

排序欄位 資料來源順序

#	Abc	Abc
Age	Job	Marital
58	management	married
44	technician	single
33	entrepreneur	married
47	blue-collar	married

Abc 類別型資料

日期資料

# 數值資料

排序

#	Abc	Abc
Age	Job	Marital
44	technician	single
33	unknown	single
28	management	single
43	technician	single

# 資料調整 (X 經度, Y 緯度)

#	Dengue_Daily_last12m.csv	#	Dengue_Daily_last12m.c
最小統計區中心點X	最小統計區中心點Y		
121.45269			
120.29430			

修正為地理角色: 地球符號

The screenshot shows the Tableau Data Source pane for the 'Dengue\_Daily\_last12m.csv' file. A context menu is open over the '最小統計區中心點X' column header, specifically over the first data value '121.45269'. The menu is titled '#'. The 'Geographic Role' option is highlighted with a red box. The menu also includes other options like 'Number (Decimal)', 'Number (Integer)', 'Date and Time', 'Date', 'String', 'Boolean', 'Default', and 'None'.

#	Dengue_Daily_last12m.csv		ABC
最小統計區中心點X	最小統計區中心點Y		一級統計區
121.45269			25.01142 A6501-C1-004
120.29430			22.63232 A6407-05-004
121.48052			25.01783 A6501-83-009
120.39661			22.60242 A6414-25-220
120.33063			
121.52055			
120.68748			
121.29346			
121.29693			
120.31614			
121.34725			
120.32670			



Dengue\_Daily\_Last12m.csv

最小統計區中心點Y

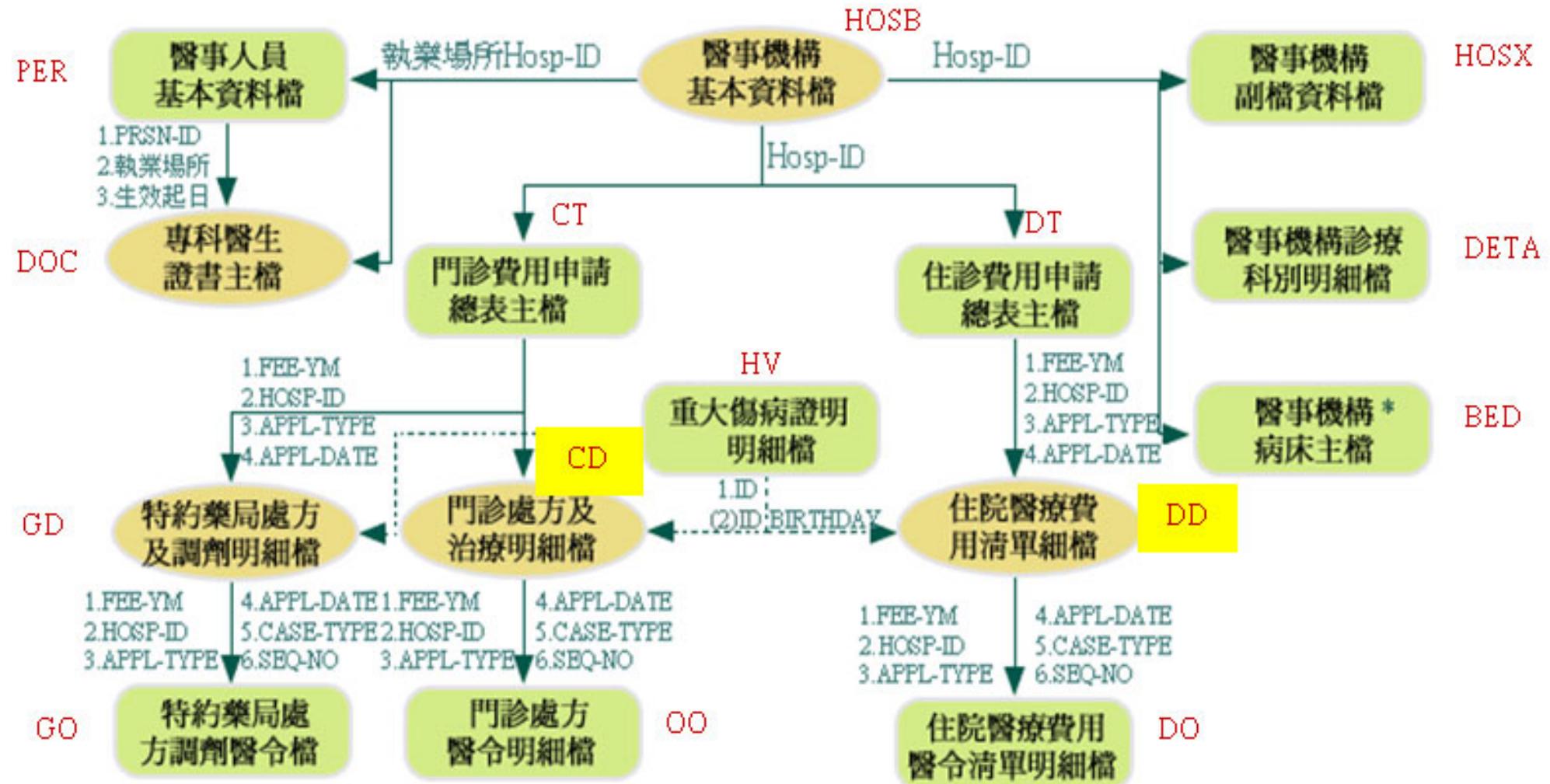
# 資料聯合 (JOIN)

- 二個以上資料表, 依共同欄位合併成為新的資料表.
- 類似資料庫 SQL JOIN, 一般是內部, 即 **INNER JOIN**, 即左右值相同.
- 類似 Excel 使用 =VLOOKUP() 函數

The screenshot shows the Tableau Data Source dialog for a connection named "Dengue\_Daily\_last12m". It displays two data sources: "Dengue\_Daily\_last12m.csv" and "ATM00625\_2019092212154...". A join dialog is open, titled "聯接" (Join). The "内部" (Inner) option is selected. The "資料來源" (Source) dropdown is set to "ATM00625\_2019092212154...". The "左側" (Left) and "右側" (Right) dropdowns both show "Dengue\_Daily..." under the "發病日" (Report Date) column. The "Site" and "count" columns from the left source are mapped to the "Site" and "count" columns in the right source respectively. The "排序欄位" (Sort Fields) section is visible on the left.

# 資料聯合應用案例

## 各檔案間串檔變項說明



### 3. Tableau 金融商品案例分析



# bank\_deposit.csv

- 資料下載

- RWEPA → 右側
  - [GitHub DataDemo](#) →  bank\_deposit.csv
- [https://github.com/rwepa/DataDemo/blob/master/bank\\_deposit.csv](https://github.com/rwepa/DataDemo/blob/master/bank_deposit.csv)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	subscribe
2	58	management	married	tertiary	no	2143	yes	no	unknown	5	may	261	1	-1	0	unknown	no
3	44	technician	single	secondary	no	29	yes	no	unknown	5	may	151	1	-1	0	unknown	no
4	33	entrepreneur	married	secondary	no	2	yes	yes	unknown	5	may	76	1	-1	0	unknown	no
5	47	blue-collar	married	unknown	no	1506	yes	no	unknown	5	may	92	1	-1	0	unknown	no
6	33	unknown	single	unknown	no	1	no	no	unknown	5	may	198	1	-1	0	unknown	no
7	35	management	married	tertiary	no	231	yes	no	unknown	5	may	139	1	-1	0	unknown	no
8	28	management	single	tertiary	no	447	yes	yes	unknown	5	may	217	1	-1	0	unknown	no
9	42	entrepreneur	divorced	tertiary	yes	2	yes	no	unknown	5	may	380	1	-1	0	unknown	no

- 資料筆數: 45212
- 欄位個數: 17

# 銀行定期存款商品銷售

- 1. age : 年齡
  - 2. job : 工作
  - 3. marital : 婚姻 {divorced, married, single}
  - 4. education: 教育程度
  - 5. default : 是否有個人違約記錄
  - 6. balance : 帳戶餘額
  - 7. housing : 是否有房屋
  - 8. loan : 是否有個人貸款
  - 9. contact : 連絡方式
  - 10. day : 最近連絡日
  - 11. month : 最近連絡月
  - 12. duration : 最近連絡通話時間  
秒數. duration愈小時,  
subscribed 較可能為no.
  - 13. campaign: 連絡次數
  - 14. pdays : 最近與客戶聯繫  
之後經過的天數, -1表示以  
前未與客戶聯繫
  - 15. previous : 連絡次數
  - 16. poutcome : 最近行銷結果  
{failure, other, success,  
unknown}
17. subscribe : {yes 有購買定期存款商品, no 沒有購買定期存款商品} --> 反應變數

# Tableau 汇入資料

# 檔案 (匯入資料)

1. 連線至 Tableau Server
2. 到檔案 (CSV, Excel...)
3. 到伺服器
4. 已儲存資料來源



# 選取[文字檔] → 選取bank\_deposit.csv

The screenshot shows the Tableau Data Source interface for selecting a CSV file. The interface includes:

- Step 1:** A red box highlights the "資料(D)" (Data) tab in the top navigation bar.
- Step 2:** A red box highlights the "新增" (New) button in the top right corner of the main pane.
- Step 3:** A red box highlights the "bank\_deposit.csv" file listed under the "bank\_deposit" folder.
- Step 4:** A red box highlights the connection options: "連線" (Connection), "即時" (Real-time), and "擷取" (Extract). "即時" is selected.
- Step 5:** A red box highlights the "Marital" column header in the preview table, which is being sorted by the user.
- Step 6:** A red box highlights the "工作表 1" (Sheet 1) tab at the bottom left.

**1. 資料**

**2. 新增(資料)**

**3. 關聯(2個以上資料表)**

**4. 即時, 擷取**

**5. 排序**

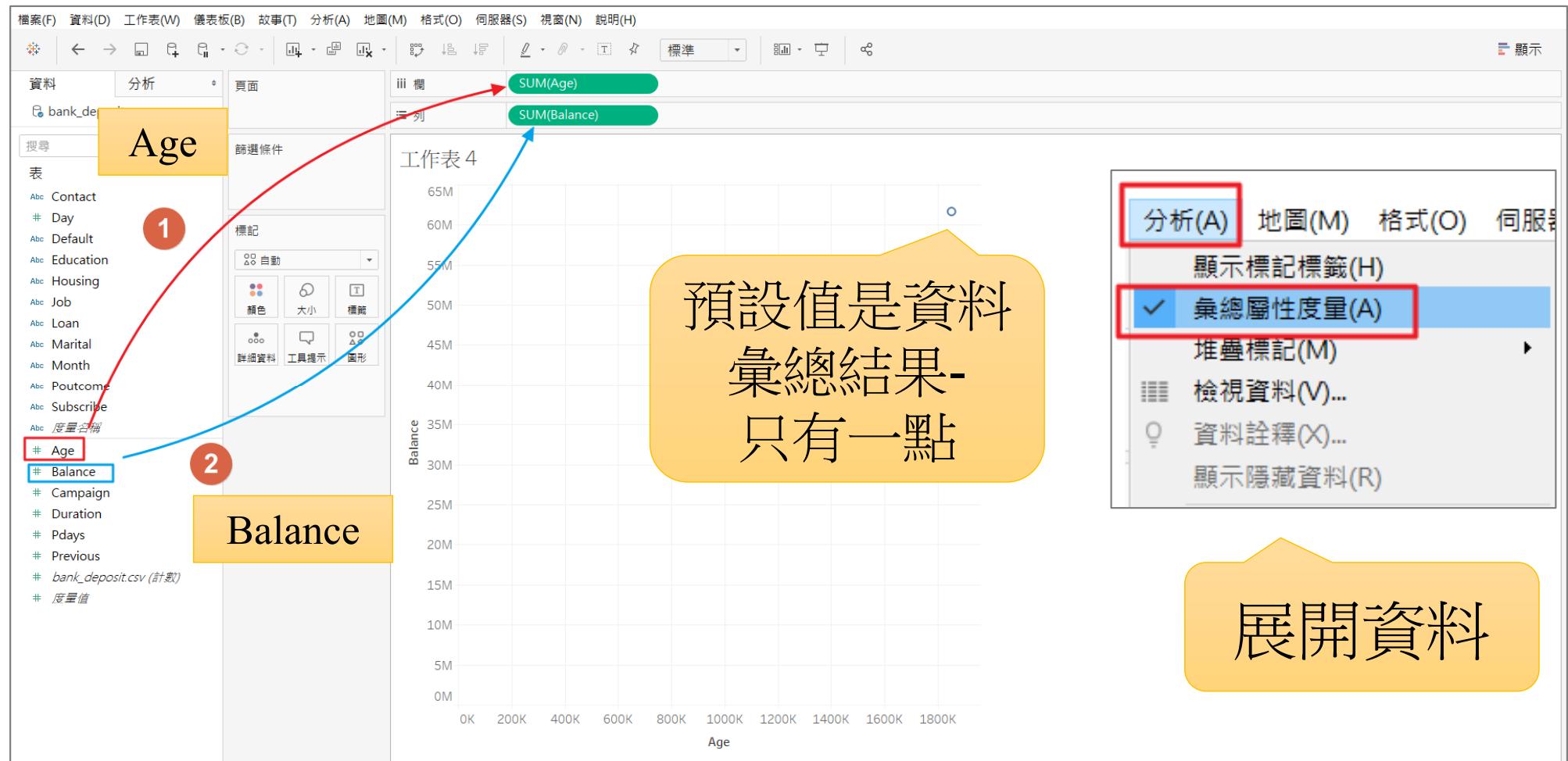
**6. 工作表 (視覺化)**

#	Age	Job	Marital	Education	Default
58	management	bank_deposit.csv.Job	bank_deposit.csv.Marital	bank_deposit.csv.Education	bank_deposit.csv.Default
44	technician	single	secondary	no	no
33	entrepreneur	married	secondary	no	no
47	blue-collar	married	unknown	no	no
33	unknown	single	unknown	no	no
35	management	married	tertiary	no	no
28	management	single	tertiary	no	no
42	entrepreneur	divorced	tertiary	yes	no
58	retired	married	primary	no	no
43	technician	single	secondary	no	no

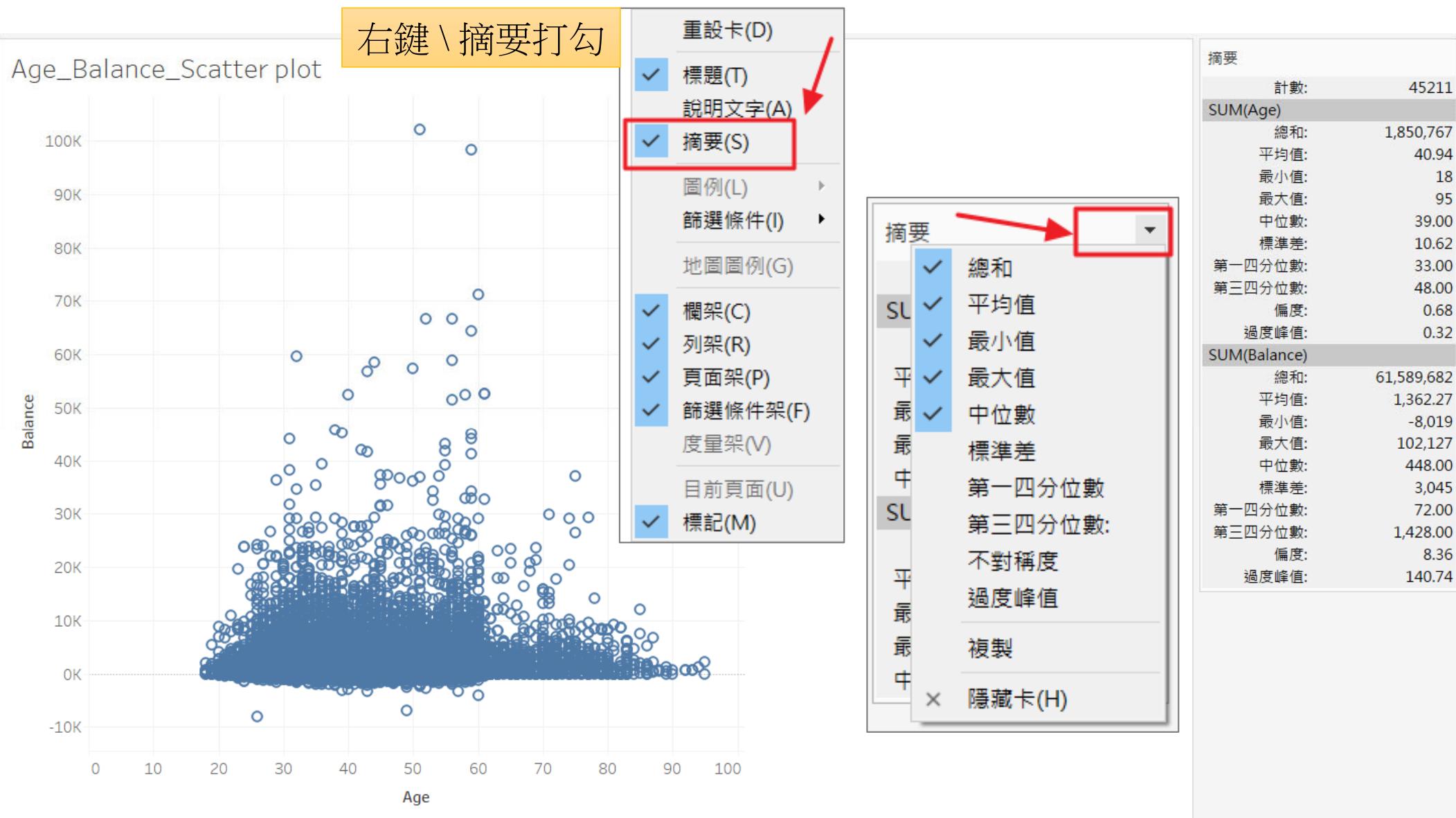
# Scatter plot 散佈圖

適用於X, Y數值資料

# Age\_Balance-Scatter plot 散佈圖

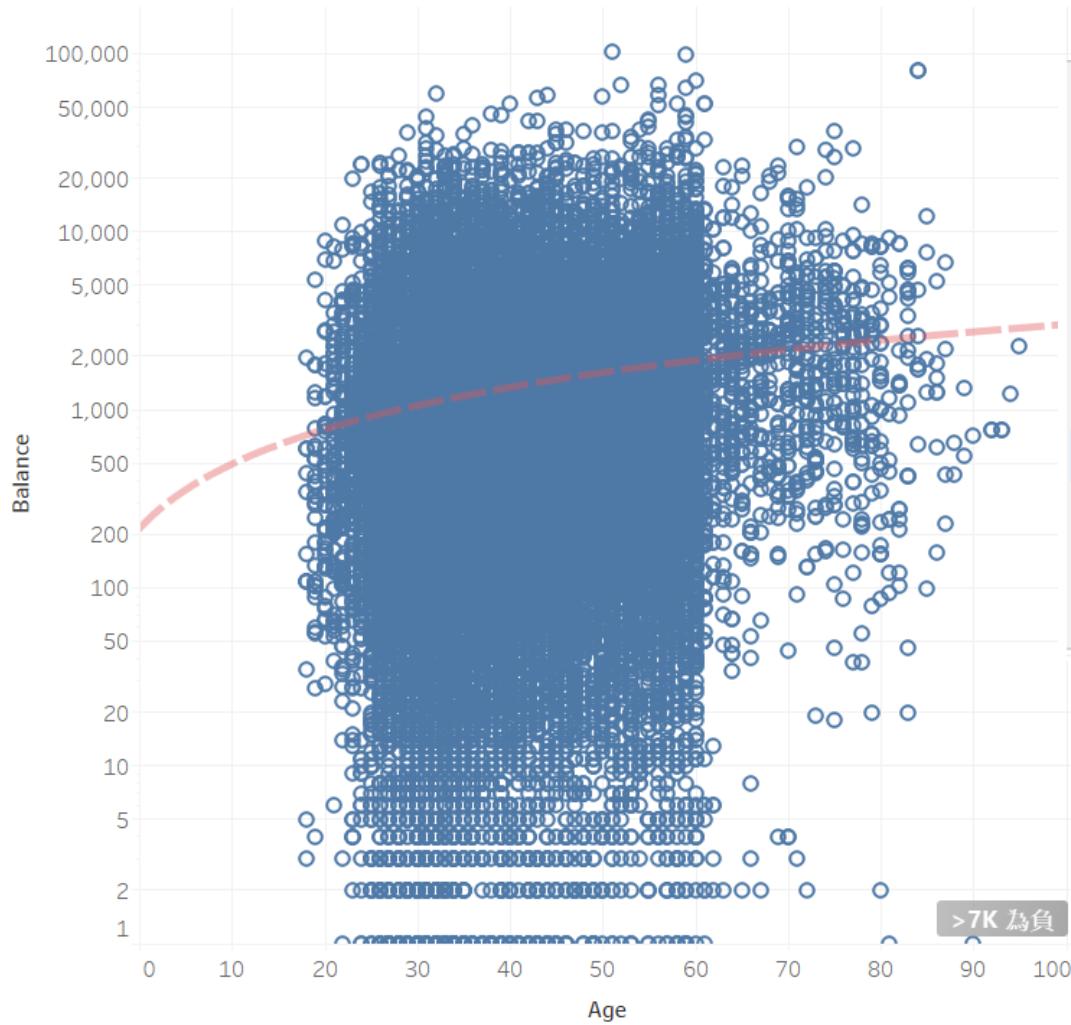


# Age\_Balance\_Scatter plot

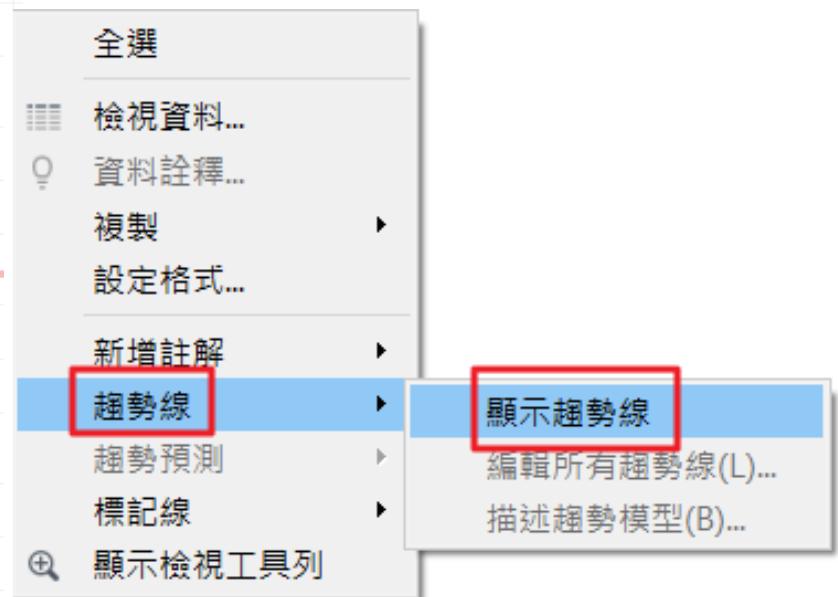


# Age\_Balance\_Scatter plot\_log\_trend

Age\_Balance\_Scatter plot\_log\_trend



- Y 軸 \ 右鍵 \ 編輯軸 \ 對數 log
- 右鍵 \ 趨勢線 \ 顯示趨勢線



# NULL 處理

# NULL (NA值處理方法)

- 忽略方法, 即直接刪除法 → 少用
- 使用出現次數最多者或是中位數等統計量填滿
- 使用演算法填滿, 例: KNN (K近鄰法)
  - [https://en.wikipedia.org/wiki/K-nearest\\_neighbors\\_algorithm](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm)
- 使用預測演算法填滿, 例: 迴歸分析

# NULL 空白值處理

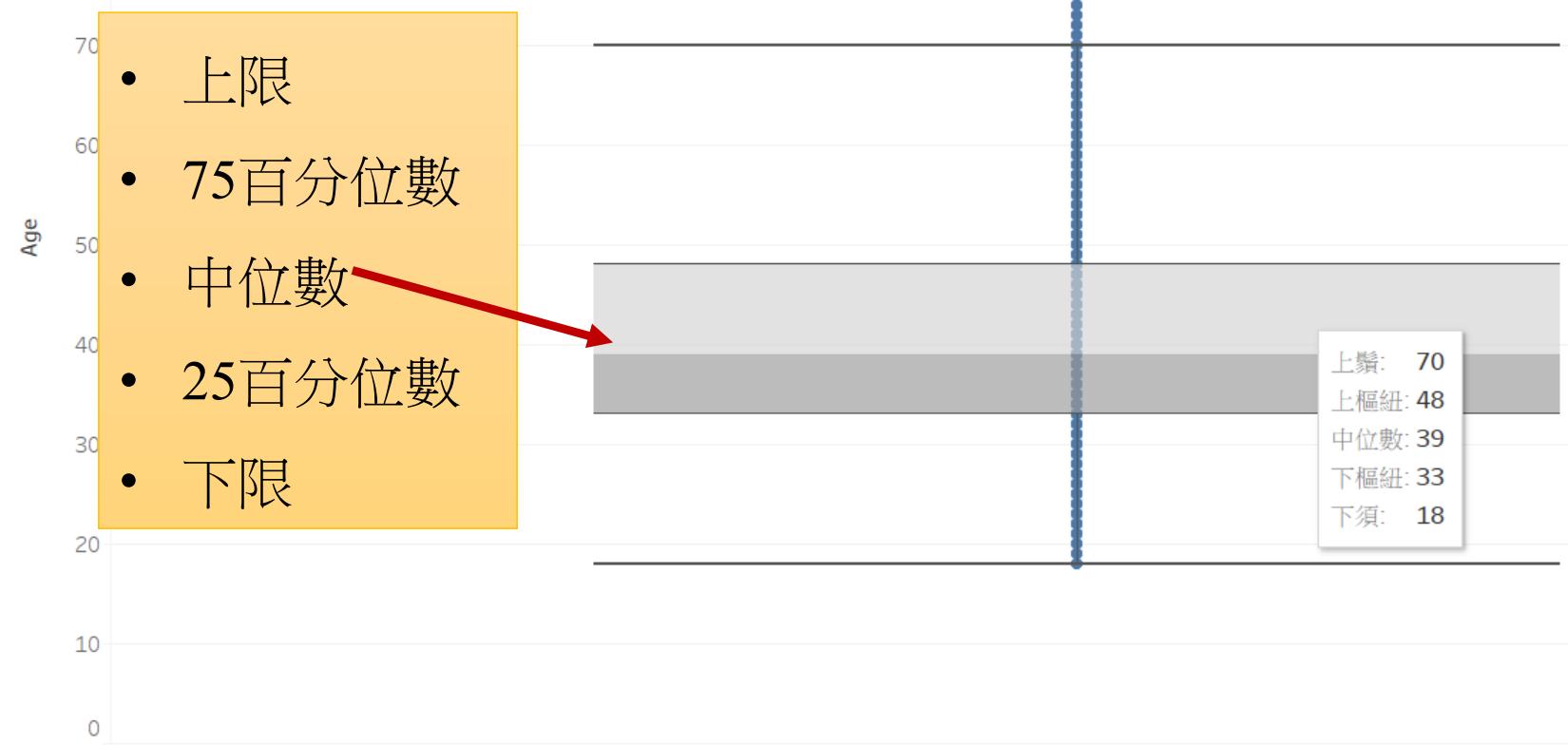
- Tableau 汇入空白資料(遺漏值)會顯示 null
  - 範例數值變數沒有 NULL
  - 範例類別型變數為遺漏值以 “unknown” 表示
- 將 unknown 改為出現最多次類別填補:
  - job
  - education
  - contact
- poutcome 變數不可使用類別填補

- R: NA
- Python: np.nan

# Age\_盒鬚圖

Age\_盒鬚圖

```
> age_plot <- boxplot(bank$age)
> age_plot
$stats
 [,1]
[1,] 18
[2,] 33
[3,] 39
[4,] 48
[5,] 70
```



Upper Whisker:	70
Upper Hinge:	48
Median:	39
Lower Hinge:	33
Lower Whisker:	18

離群值 outliers

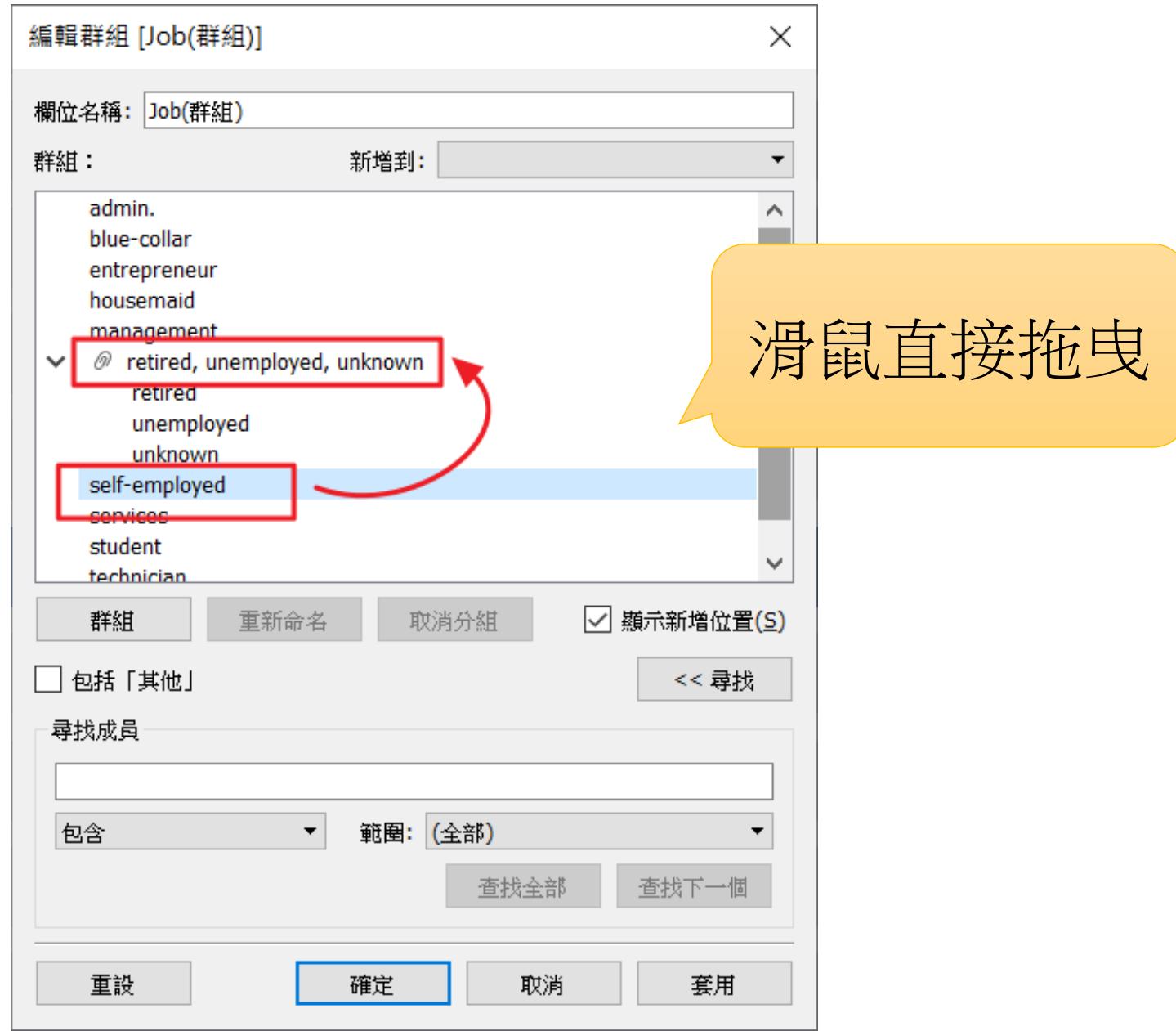
# 群組

The screenshot shows a Tableau interface with a data source containing fields like Age, Balance, Contact, Date, Day, Default, Duration, Education, Housing, Job, Job withNull, Loan, Marital, Month, Poutcome, and S. A specific 'Job' field is selected, displaying categories: retired (1), unknown (2), management, self-employed, unemployed (3), entrepreneur, and housemaid.

A context menu is open over the 'Job' field, with the '群組' (Group) option highlighted and selected. A yellow callout bubble says: 選取3個 Job \ 右鍵 \ 群組 (Select 3 jobs \ Right-click \ Group).

The '群組' option leads to a secondary context menu for the 'Job(群組)' group, which includes options like 新增到工作表 (Add to Worksheet), 顯示篩選條件 (Show Filter Conditions), 剪下 (Cut), 複製 (Copy), 編輯群組... (Edit Group...), 重複 (Duplicate), and 重新命名 (Rename). A red arrow points from the '群組' selection in the first menu to the '編輯群組...' option in the second menu.

# 群組 \ 編輯群組



# Tableau\_Financial\_Products\_Case\_Study.twbx

檔案(E) 資料(D) 工作表(W) 儀表板(B) 故事(I) 地圖(M) 格式(Q) 伺服器(S) 視窗(N) 說明(H)

故事 版面配置 主題 大綱 金融商品資料集 雜誌圖 < >

新增本文點 空白 重複

title 大綱 變數說明 Age\_Balance\_... Age\_Balance\_... Education\_Su... Age\_直方圖 (練習)balance\_... 群組 Job\_table Age\_盒鬚圖 balance\_群組 散佈圖矩陣\_sc... 集群\_初使化 集群\_完成 儀表板 儀表板\_聯動操作

A 拖動以新增文字

顯示標題

大小 故事 (1016 x 964)

資料來源 Tableau 金融商品案例分析 title 大綱 變數說明 Age\_Balance\_散佈圖 Age\_Balance\_散佈圖\_log\_趨線圖 Education\_Subscribe\_併排長條圖 Age\_直方圖 (練習)balance\_直方圖 群組 Job\_table Age\_盒鬚圖 balance\_群組

Tableau demo

# 線上發佈

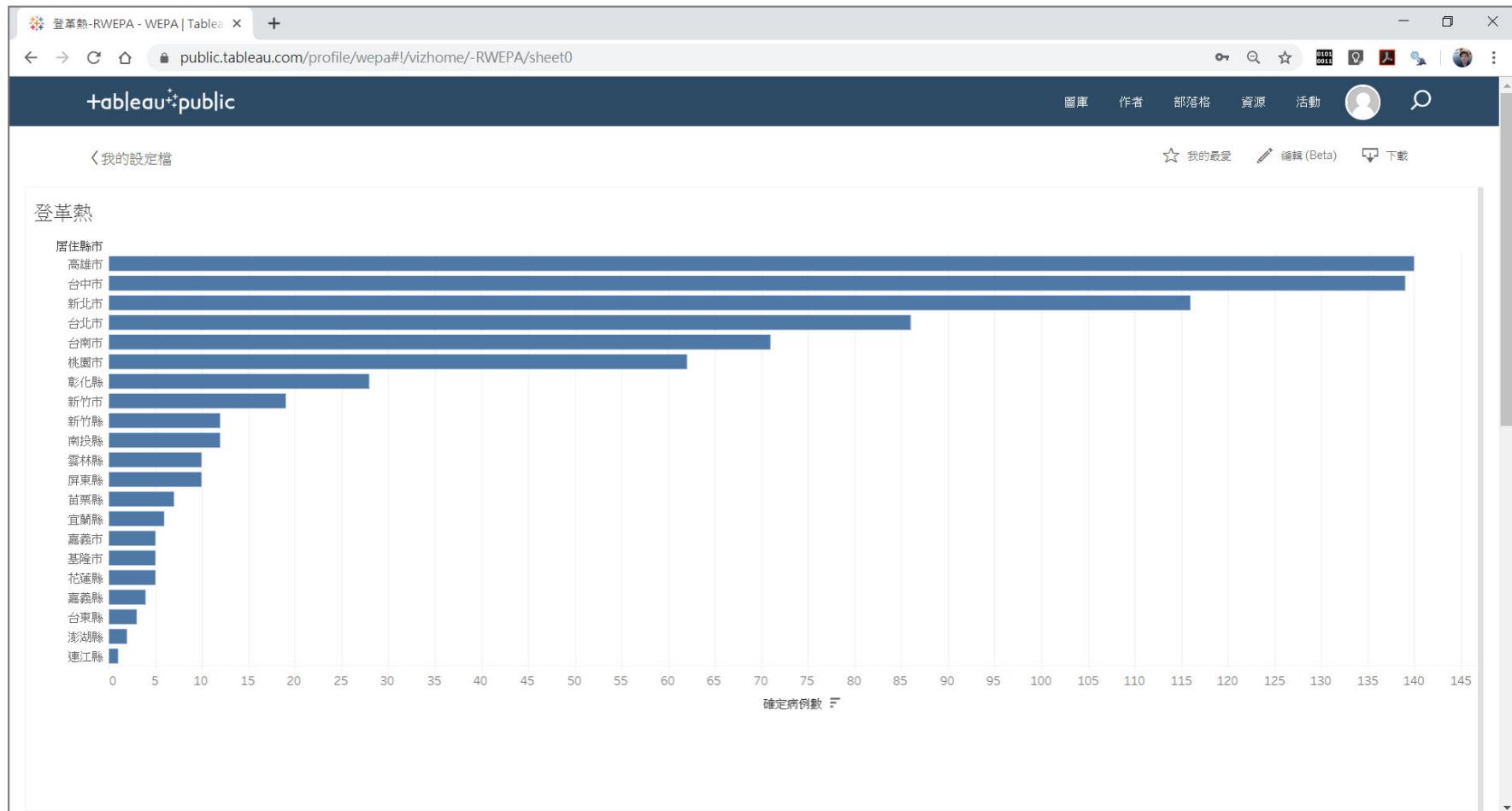
- Tableau server



- Tableau Public

The screenshot shows the 'Tableau Public Sign In' page. It features a large watermark of the Tableau logo ('tableau public') in the center. Below the watermark are two input fields: '電子郵件' (Email) and '密碼' (Password). A large orange '登入' (Log In) button is positioned below the password field. At the bottom of the page, there are links for '忘記密碼?' (Forgot Password?), '是否還沒有設定檔?' (Don't have a profile yet?), and '立即免費建立一個' (Create a free account).

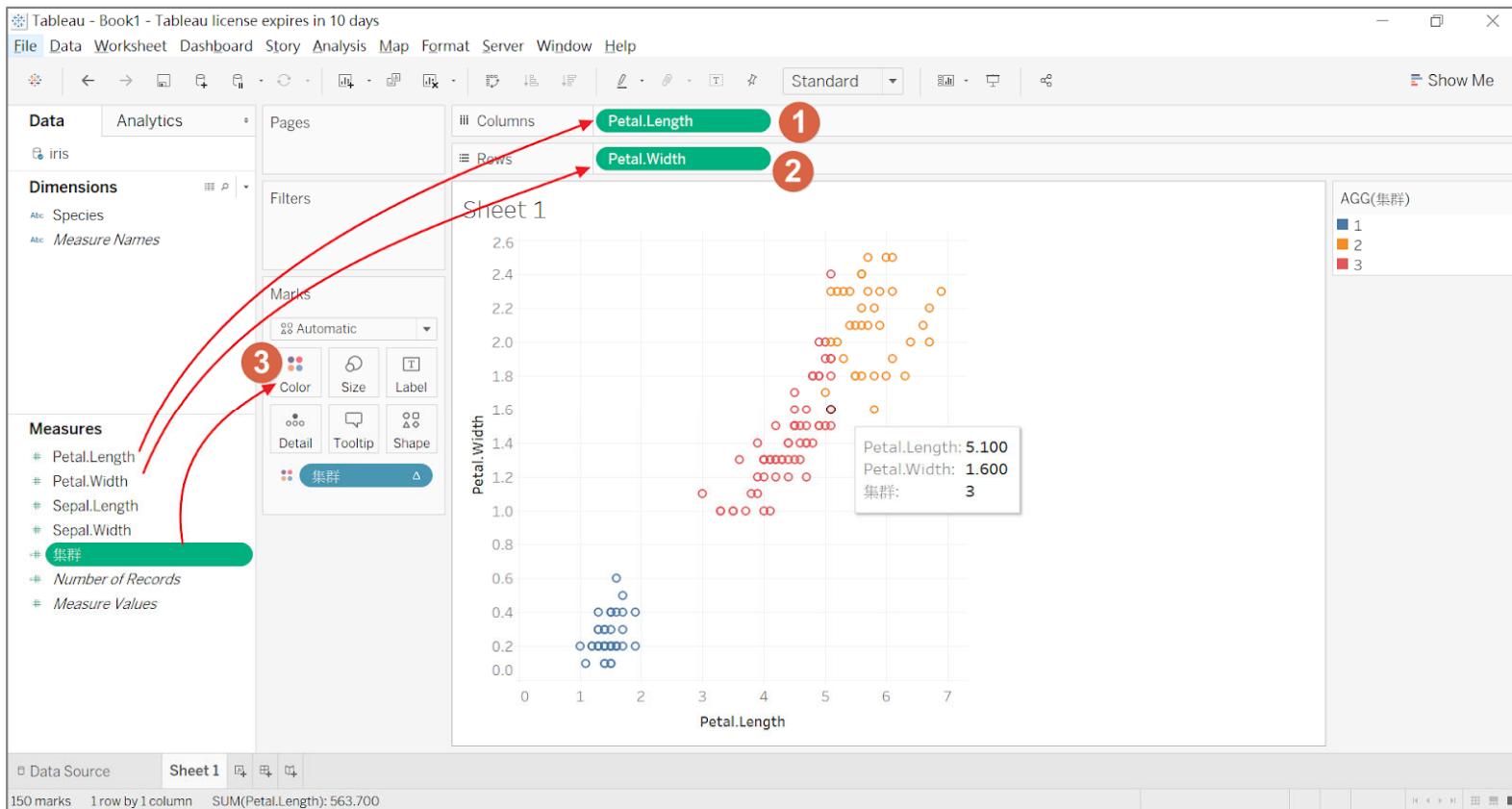
# Tableau Public



WEPA: <https://public.tableau.com/profile/wepa#!/vizhome/-RWEPA/sheet0>

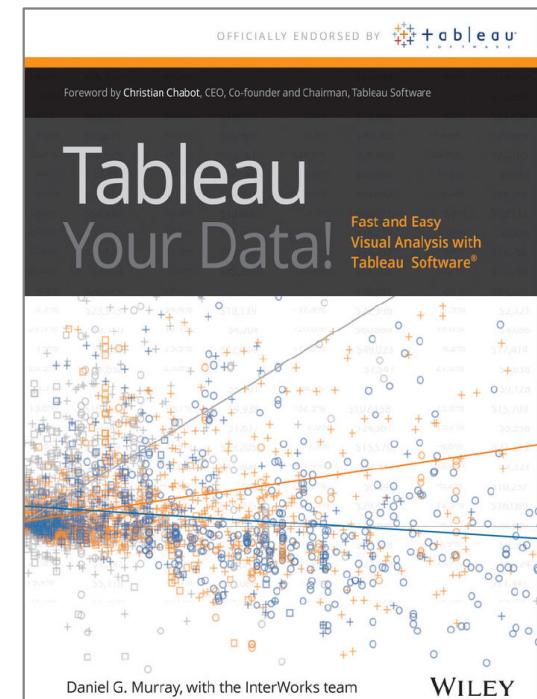
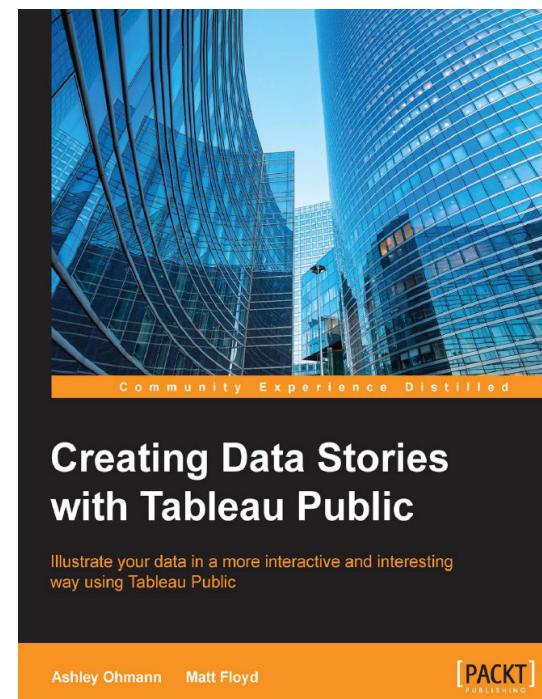
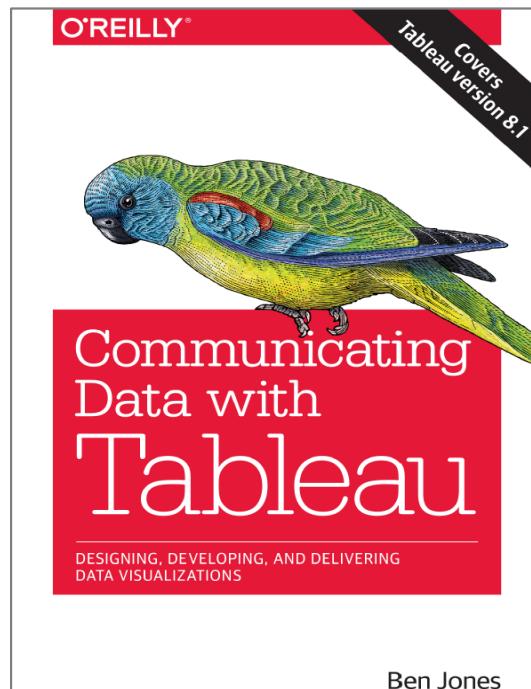
# 補充篇 Tableau + R

- <http://rwepa.blogspot.com/2019/03/tableau-r.html>



# 參考資料

- RWEPA: <http://rwepa.blogspot.com/>
- Youtube-Tableau: <https://www.youtube.com/user/tableausoftware>



# 謝謝您的聆聽



## Q & A

李明昌

*alan9956@gmail.com*

***<http://rwepa.blogspot.tw/>***