Information Visualization Introduction

Ming-Te Chi

moodle

Slide: https://goolink.cc/robkC

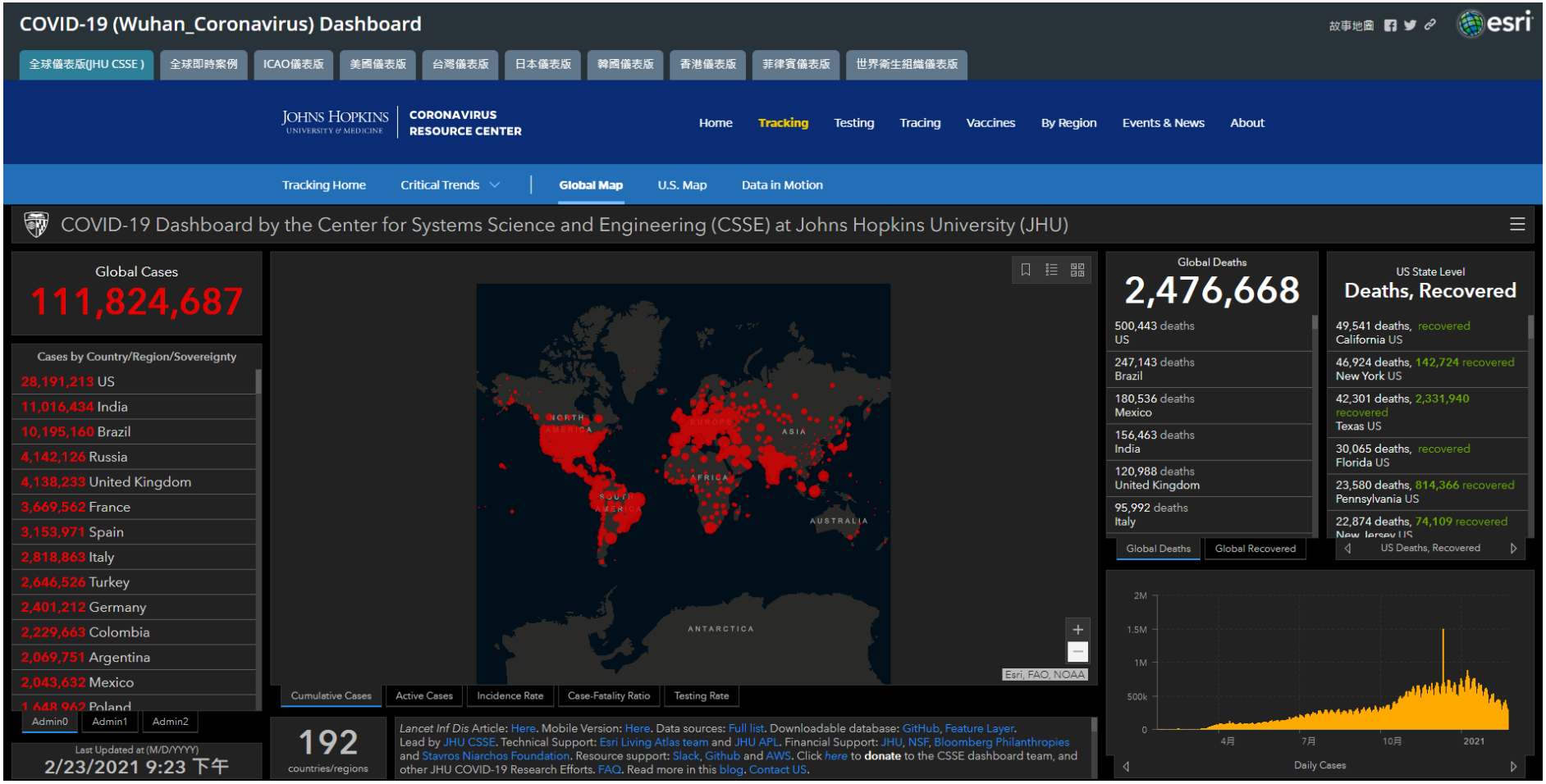
Outline

- Vis Examples
- Syllabus
- News Vis
- Chapter 1

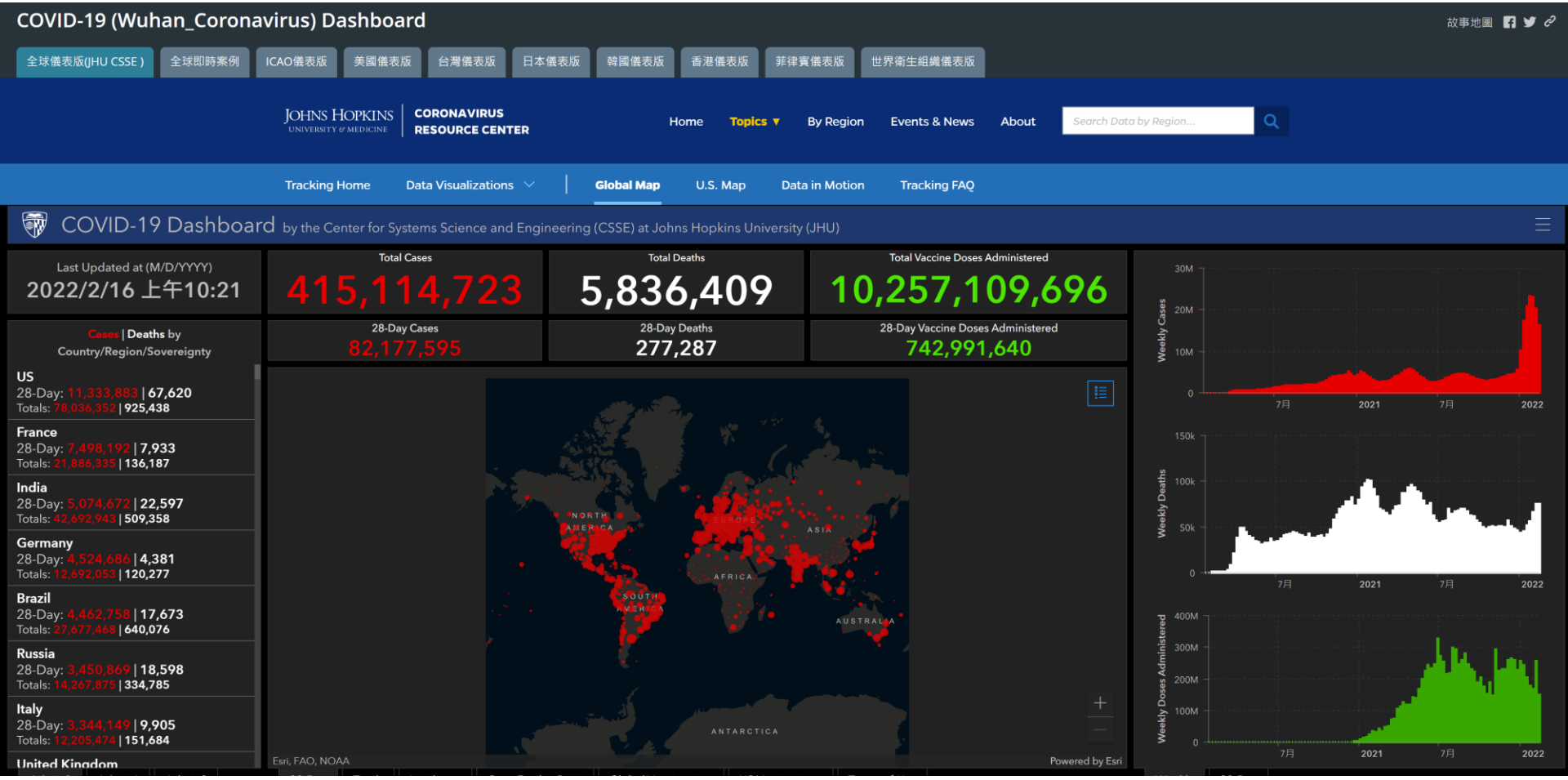
Facebook Friendship in 2010

ten million pairs of friends

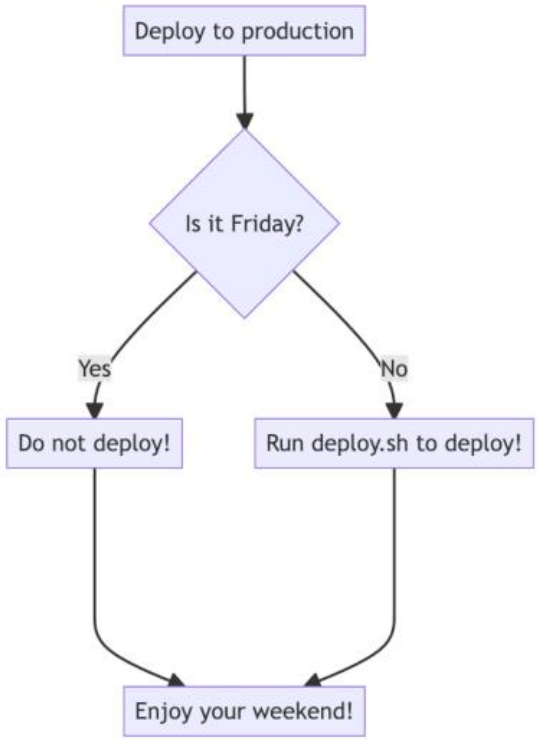
Covid 19 Dashboard



Covid 19 Dashboard



Diagrams



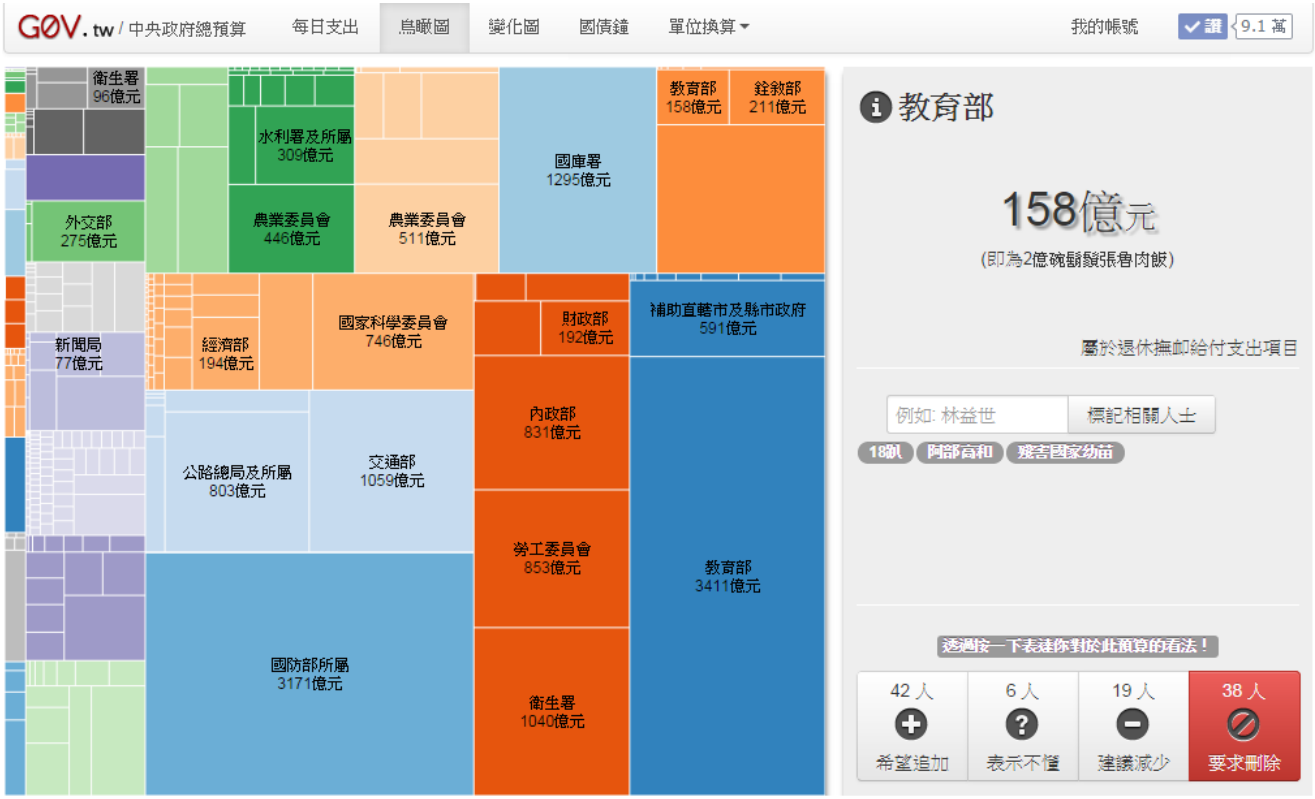
Edit file | Preview

```
## How to deploy

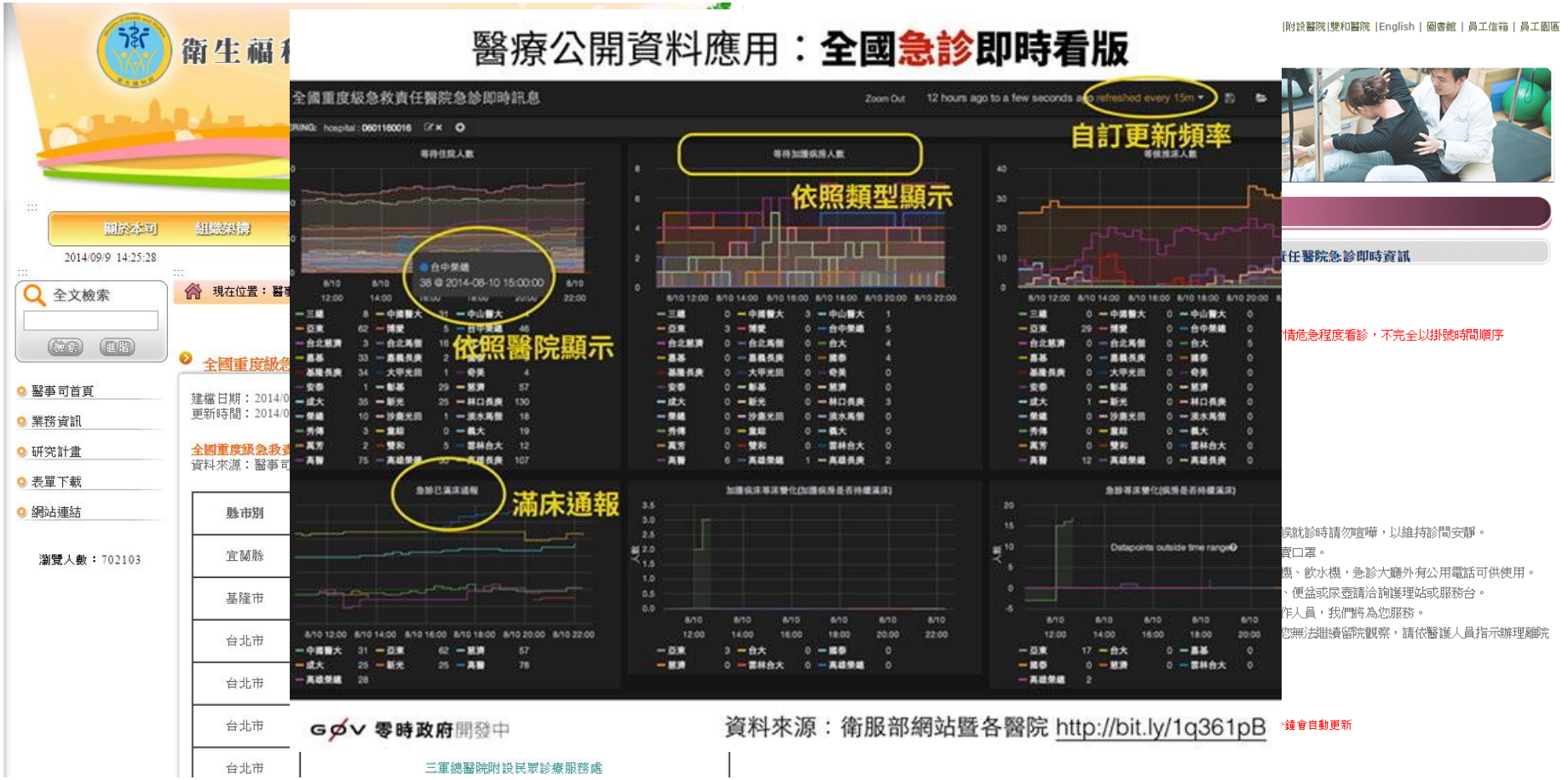
```mermaid
flowchart TD
    A[Deploy to production] --> B{Is it Friday?};
    B -- Yes --> C[Do not deploy!];
    B -- No --> D[Run deploy.sh to deploy!];
    C ----> E[Enjoy your weekend!];
    D ----> E[Enjoy your weekend!];
```
```

https://github.blog/developer-skills/github/include-diagrams-markdown-files-mermaid/

g0v 中央政府總預算


G0V.tw / 中央政府總預算
每日支出
鳥瞰圖
變化圖
國債鐘
單位換算
我的帳號
讚 9.1 萬
衛生署 96億元
水利署及所屬 309億元
國庫署 1295億元
教育部 158億元
銓敘部 211億元
外交部 275億元
農業委員會 446億元
農業委員會 511億元
新聞局 77億元
經濟部 194億元
國家科學委員會 746億元
財政部 192億元
補助直轄市及縣市政府 591億元
內政部 831億元
公路總局及所屬 803億元
交通部 1059億元
勞工委員會 853億元
教育部 3411億元
國防部所屬 3171億元
衛生署 1040億元
教育部
158億元
(即為2億碗鹹鬚張魯肉飯)
屬於退休撫卹給付支出項目
例如: 林益世
標記相關人士
18趴
阿部高和
殘害國家幼苗
透過按一下表達你對於此預算的看法！
42 人
希望追加
6 人
表示不懂
19 人
建議減少
38 人
要求刪除

g0v全國重度級急救責任醫院急診即時訊息



g0v 政治獻金

政治獻金會計報告書查閱制度電腦系統查詢結果列印清單

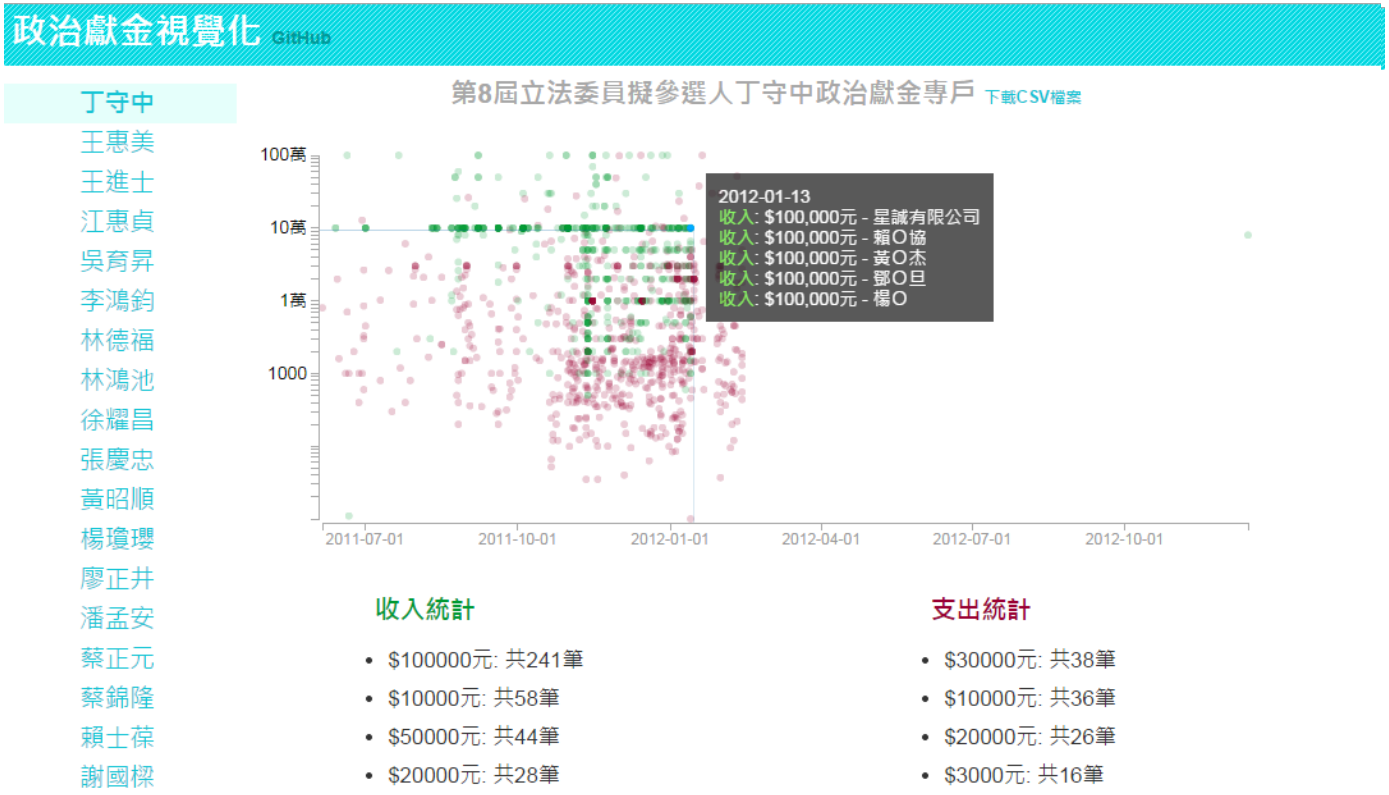
專戶名稱：第1屆新北市市長候選人朱立倫政治獻金專戶

99/11/27 為直轄市長暨市議員選舉日　　首次申報

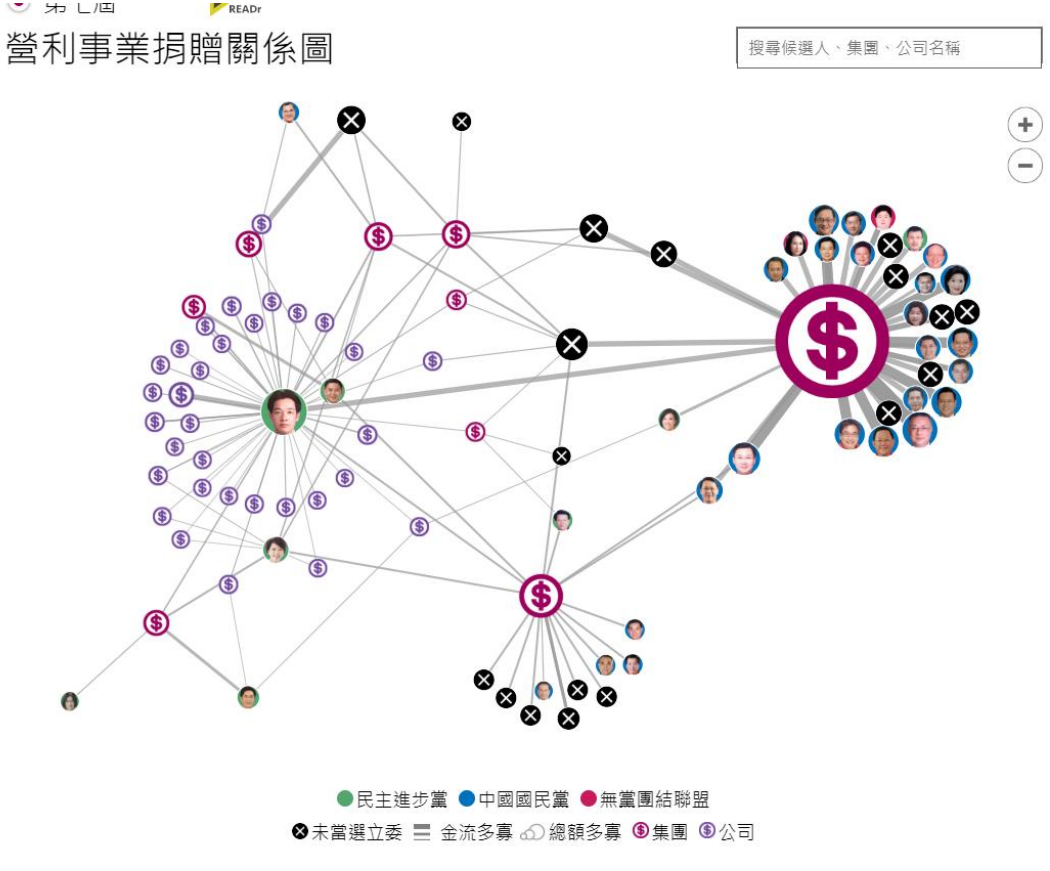
序號	交易日期	收支科目	捐贈者/支出對象	身分證/統一編號	收入金額	支出金額	金融類	地址
1	099/11/30	人事費用支出	邱●能			500	是	****
2	099/11/30	人事費用支出	張●玉			500	是	****
3	099/11/30	人事費用支出	陳●枝			500	是	****
4	099/11/30	人事費用支出	黃●智			500	是	****
5	099/11/30	人事費用支出	陳●冕玉			500	是	****
6	099/11/30	人事費用支出	吳●金鶯			500	是	****
7	099/11/30	人事費用支出	陳●琪			500	是	****
8	099/11/30	人事費用支出	黃●樹			500	是	****
9	099/11/30	人事費用支出	林●郎			500	是	****
10	099/11/30	人事費用支出	王●錦			500	是	****
11	099/11/30	人事費用支出	陳●琴			500	是	****
12	099/11/30	人事費用支出	張●望			500	是	****
13	099/11/30	人事費用支出	吳●城			500	是	****
14	099/11/30	人事費用支出	李●智			500	是	****
15	099/11/30	人事費用支出	張●明			500	是	****
16	099/11/30	人事費用支出	張●宏			500	是	****
17	099/11/30	人事費用支出	李●玉			500	是	****
18	099/11/30	人事費用支出	黃●和			500	是	****
19	099/11/30	人事費用支出	王●月			500	是	****
20	099/11/30	人事費用支出	石●星			500	是	****

本政治獻金資料僅供學術交流使用，嚴禁營利、徵信，或其他不正當目的之使用

g0v 政治獻金視覺化 (2014)



數讀政治獻金2.0 (2018)



賴清德的營利事業捐贈排名

1.	亞利通運	$1,000,000
2.	遠東集團 ▶	$1,000,000
3.	威京集團 ▶	$300,000
4.	大億企業集團 ▶	$300,000
5.	台南企業集團 ▶	$300,000
6.	丞寬企業	$200,000
7.	台南幫集團 ▶	$200,000
8.	東陽事業集團 ▶	$200,000
9.	錦慶鋁業	$200,000
10.	北海道超商	$200,000
11.	寶隆建設	$100,000
12.	儀大 ▶	$100,000

誰跟賴清德的政商關係最接近

1. 陳亭妃 重複 6 家子公司
2. 王昱婷 重複 4 家子公司
3. 李俊毅 重複 4 家子公司
4. 黃偉哲 重複 2 家子公司
5. 高思博 重複 2 家子公司
6. 郭玟成 重複 2 家子公司
7. 林益世 重複 1 家子公司
8. 葉宜津 重複 1 家子公司
9. 余政憲 重複 1 家子公司
10. 林郁方 重複 1 家子公司
11. 鍾紹和 重複 1 家子公司
12. 吳光訓 重複 1 家子公司

賴清德 (61 歲)

民主進步黨 連任

總收入 3,205.5 萬元

營利事業收入 571.1 萬元

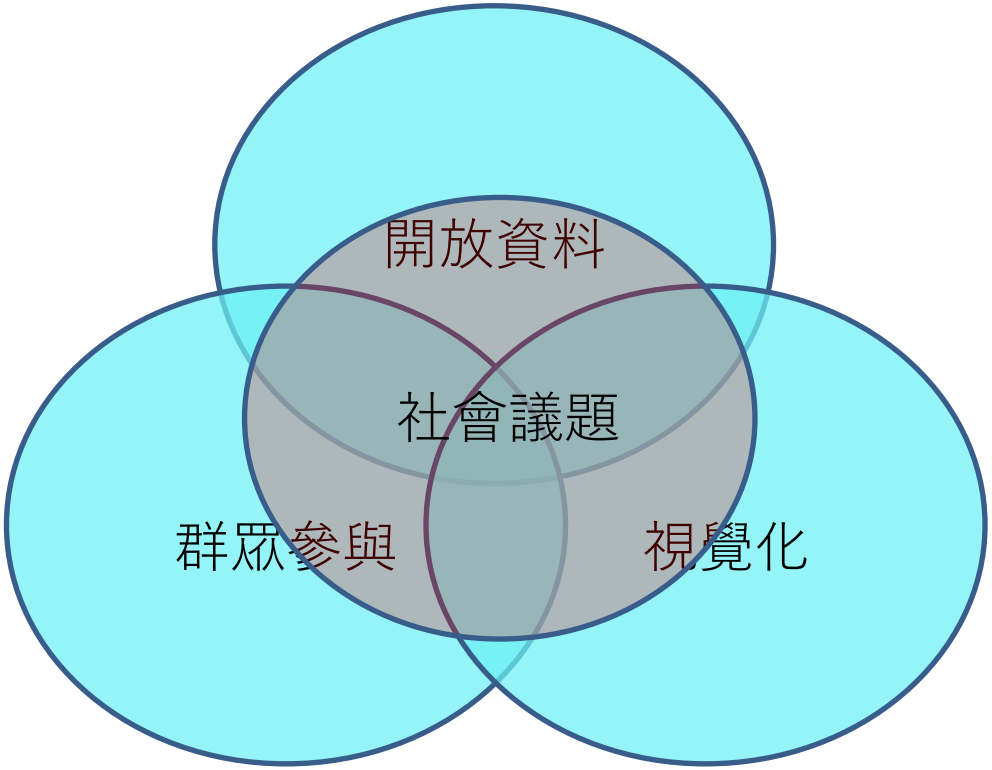
捐贈公司數 41 家

政治獻金收入組成

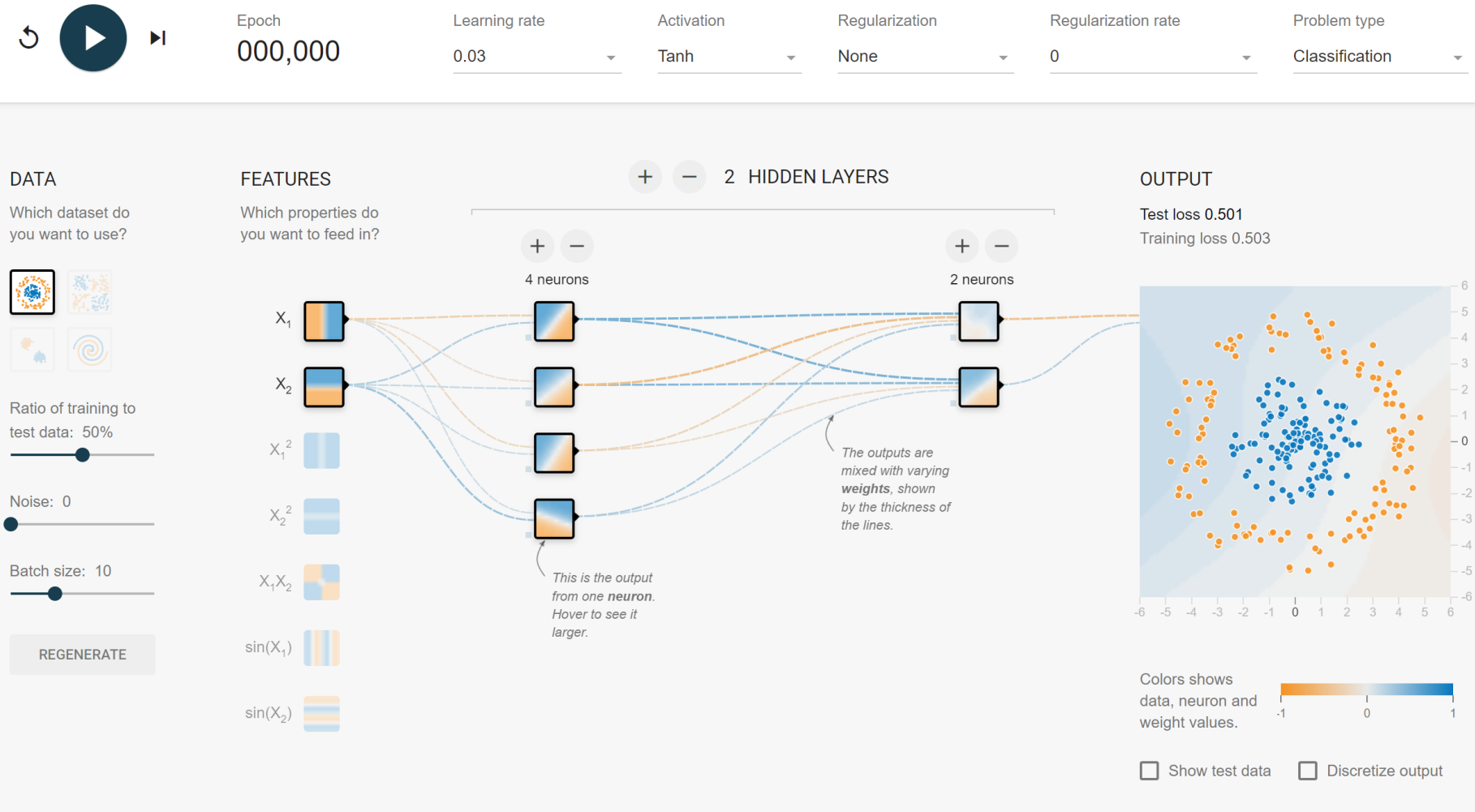
營利事業產業組成 (註)

營利事業選區內 / 選區外

營利事業選區縣市 / 外縣市

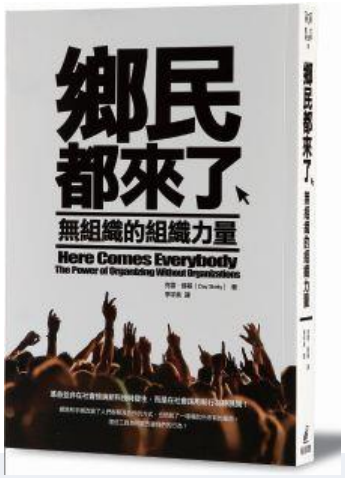

開放資料
社會議題
群眾參與
視覺化

Neural Network Visualization



Trend

- Big data and open data



開放資料與資訊科技的群眾參與

- html5 web platform for interactive display

SYLLABUS

Definition

- The use of computer-generated, interactive, visual representations of data to amplify cognition.

  [Card, Mackinlay, & Shneiderman 1999]

- Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

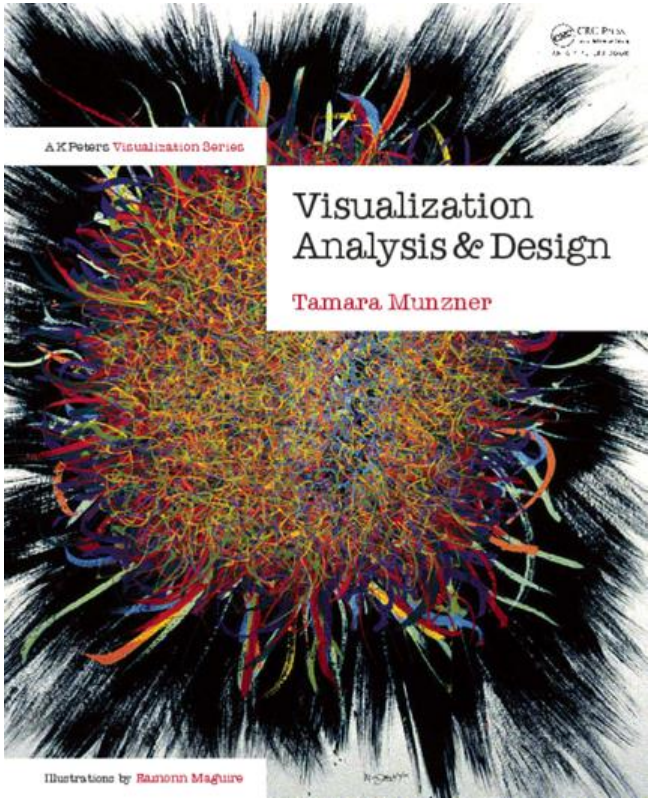
  Tamara Munzner

Text book

Visualization Analysis and Design. Munzner.

A K Peters Visualization Series, CRC Press, Visualization Series,

2014.



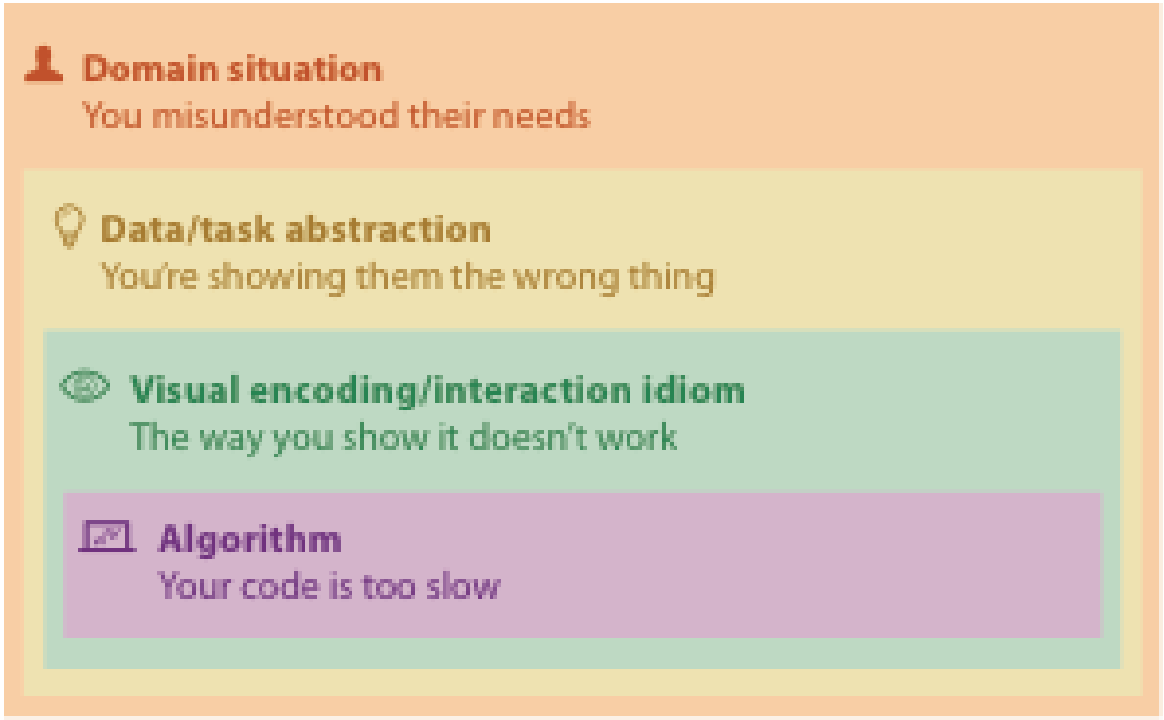
Content

- Visualization Analysis Framework
  - Analysis: What, Why, How
- Spatial Layout
- Color & Interaction
- Guidelines and Examples

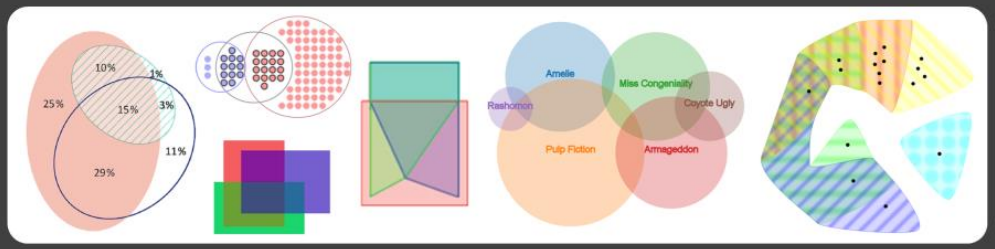
- Tooling mechanics / programming
  - D3, JavaScript, CSS, HTML5

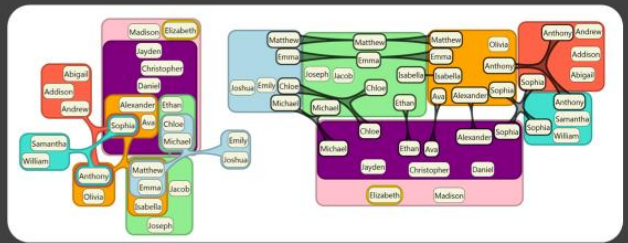
Why is validation difficult?

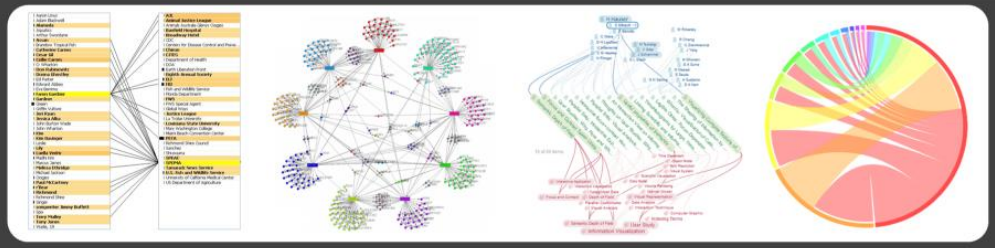
- different ways to get it wrong at each level



Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014.

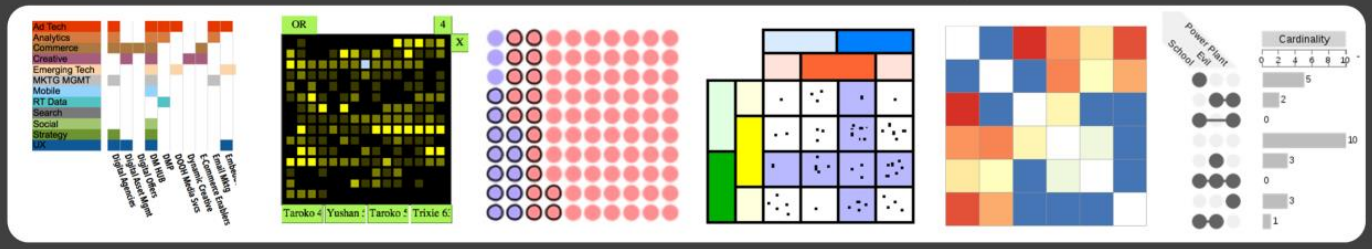


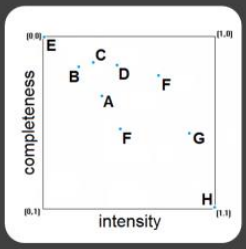


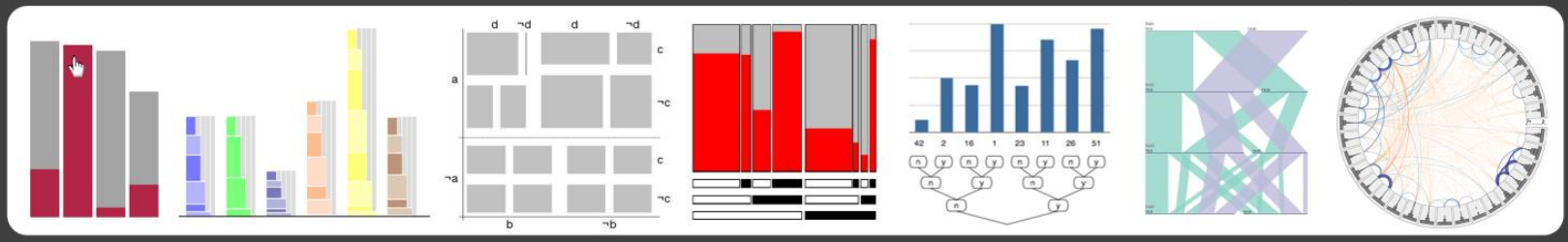


SetViz.net

Visualizing Sets and Set-typed Data







Grading

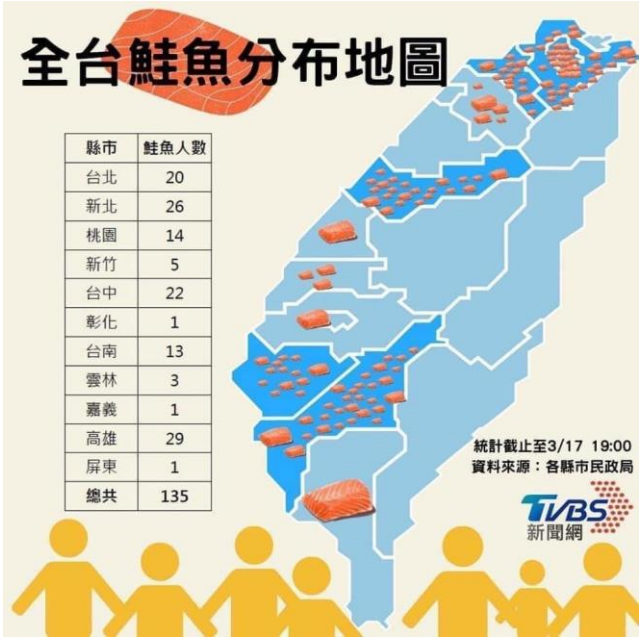
- Homework (40%)
- Paper Presentation/report (20%)
- Final Projects (40%)

Homework (40%)

- HW1: Basic visualization using D3.js
- HW2: Advance visualization using D3.js

Homework (40%)

- HW3: News visualization (team work 1~3 people)
  - Find a news/article with visualization
  - Find the weakness
  - Improve or recreate the visualization



數據看台灣

用數據看台灣

台灣開放即時資料 | 台灣開放統計資料 | 世界即時資訊 | 世界統計資訊 | 開放資料分析部落格 | 關於我們

下雨了，你知道台灣各地的即時降雨情況嗎？

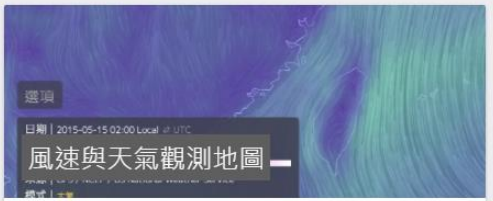
聽說最近缺水，用數據看台灣 的夥伴開發了一個視覺化的網站
可以讓你對現在缺水現況一目了然嗎！

我們每天呼吸的台灣空氣品質到底如何呢？台灣即時空氣品質
指標 PM2.5 版本提供給你台灣各地即時的 PM2.5 指數

北中南的用電到底目前是如何？各電廠所發電出來的用電資訊
為何？讓我們告訴你！

我們每天呼吸的台灣空氣品質到底如何呢？台灣即時空氣品質
指標 PM2.5 版本提供給你台灣各地即時的 PM2.5 指數

豔陽高照的中午，你知道台灣各地現在的紫外線指數嗎？



合作提案、客製化圖表製作、意見回饋 歡迎來信: service@taiwanstat.com

Paper Presentation/report (20%)

Week13~week16

before one of the classes: you each read paper on topic of your choice

- during that class: you present it to everybody else (~10 min)

Final Projects (40%)

- group of 3 ~ 5
- Stages
  - pitches (oral, in class):
  - proposals (written):
  - Progress Presentations:
  - Final presentation(oral and report):

Final Projects

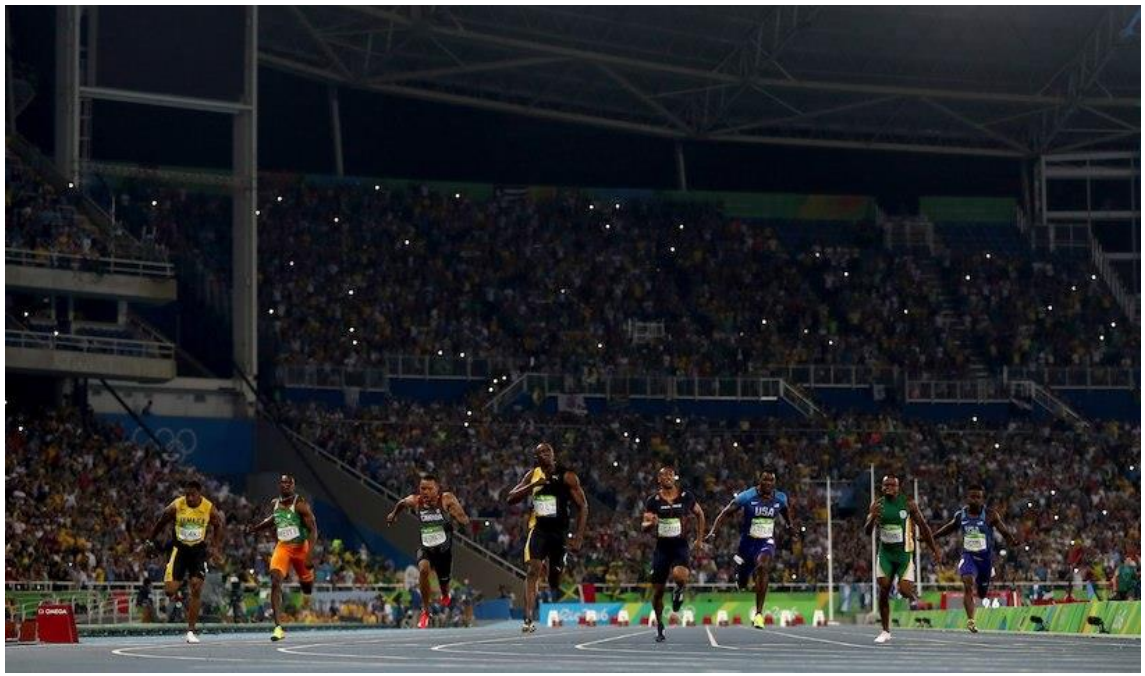
- **Programming**
  - common case
  - three types
    - problem-driven design studies (target specific task/data)
    - technique-driven (explore design choice space for encoding or interaction idiom)
    - algorithm implementation (as described in previous paper)
- **Analysis**
  - analyze a dataset/problem using a variety of existing visualization tools of your choosing, as well as analyze the strengths and weaknesses of those tools and discuss in detail
  - this option is suitable for non-CS students

NEWS VISUALIZATION

NY Times - **Men's 100-Meter Sprint**

- Olympic 2012

	Gold	Silver	Bronze
2012	9.63	9.75	9.79
2008	9.69	9.89	9.91
2004	9.85	9.86	9.87
…			

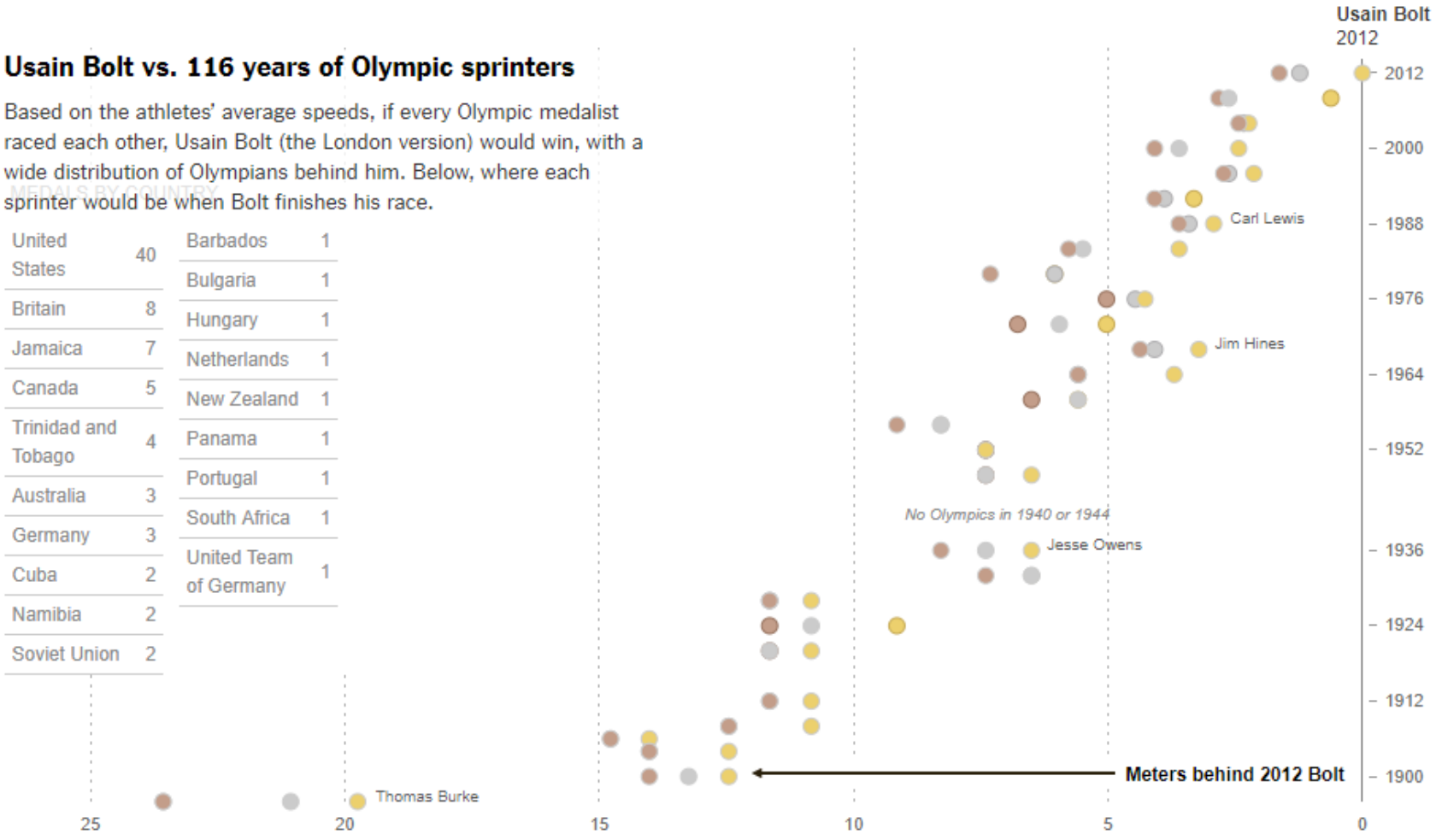


Usain Bolt vs. 116 years of Olympic sprinters

Based on the athletes' average speeds, if every Olympic medalist raced each other, Usain Bolt (the London version) would win, with a wide distribution of Olympians behind him. Below, where each sprinter would be when Bolt finishes his race.

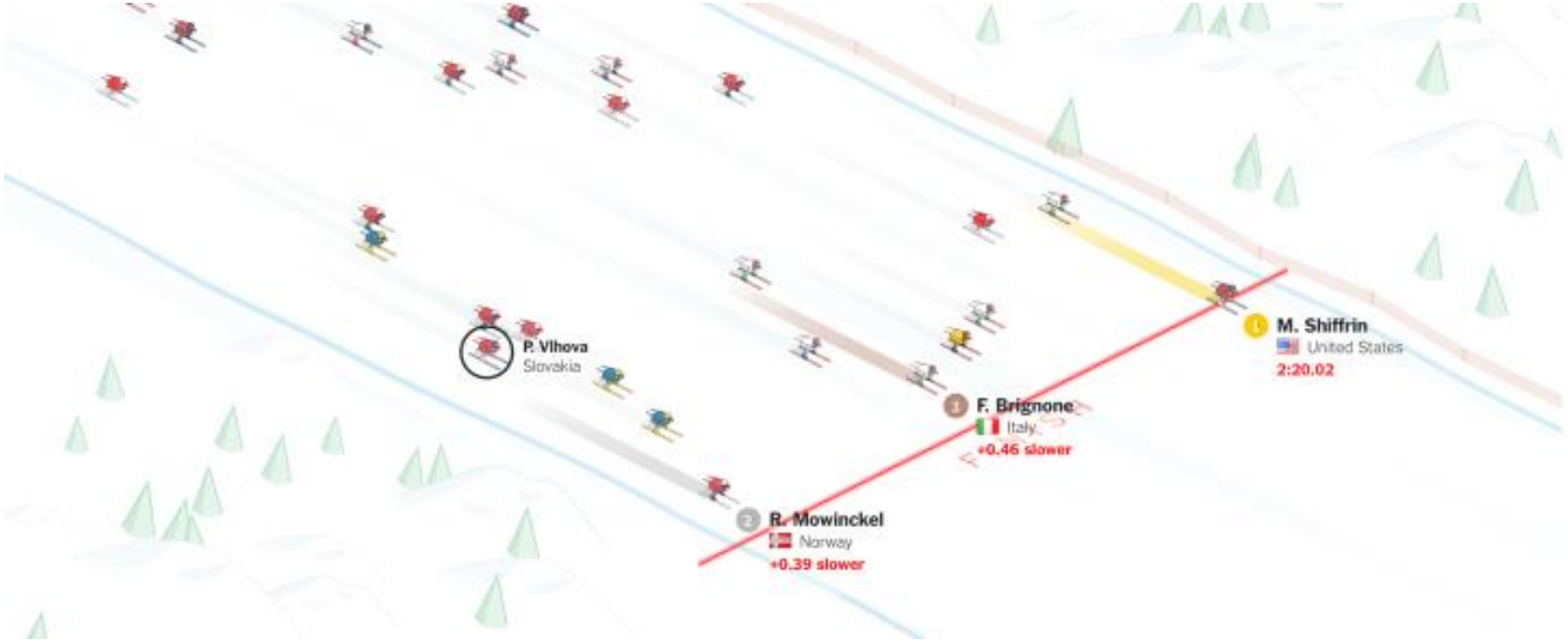
MEDALS BY COUNTRY

United States	40	Barbados	1
Britain	8	Bulgaria	1
Jamaica	7	Hungary	1
Canada	5	Netherlands	1
Trinidad and Tobago	4	New Zealand	1
Australia	3	Panama	1
Germany	3	Portugal	1
Cuba	2	South Africa	1
Namibia	2	United Team of Germany	1
Soviet Union	2		



This chart includes medals for the United States and Australia in the "Intermediary" Games of 1906, which the I.O.C. does not formally recognize.

Mikaela Shiffrin pulling away for gold



數位敘事

互動

動畫

How Uber Uses Psychological Tricks to Push Its Drivers' Buttons

互動

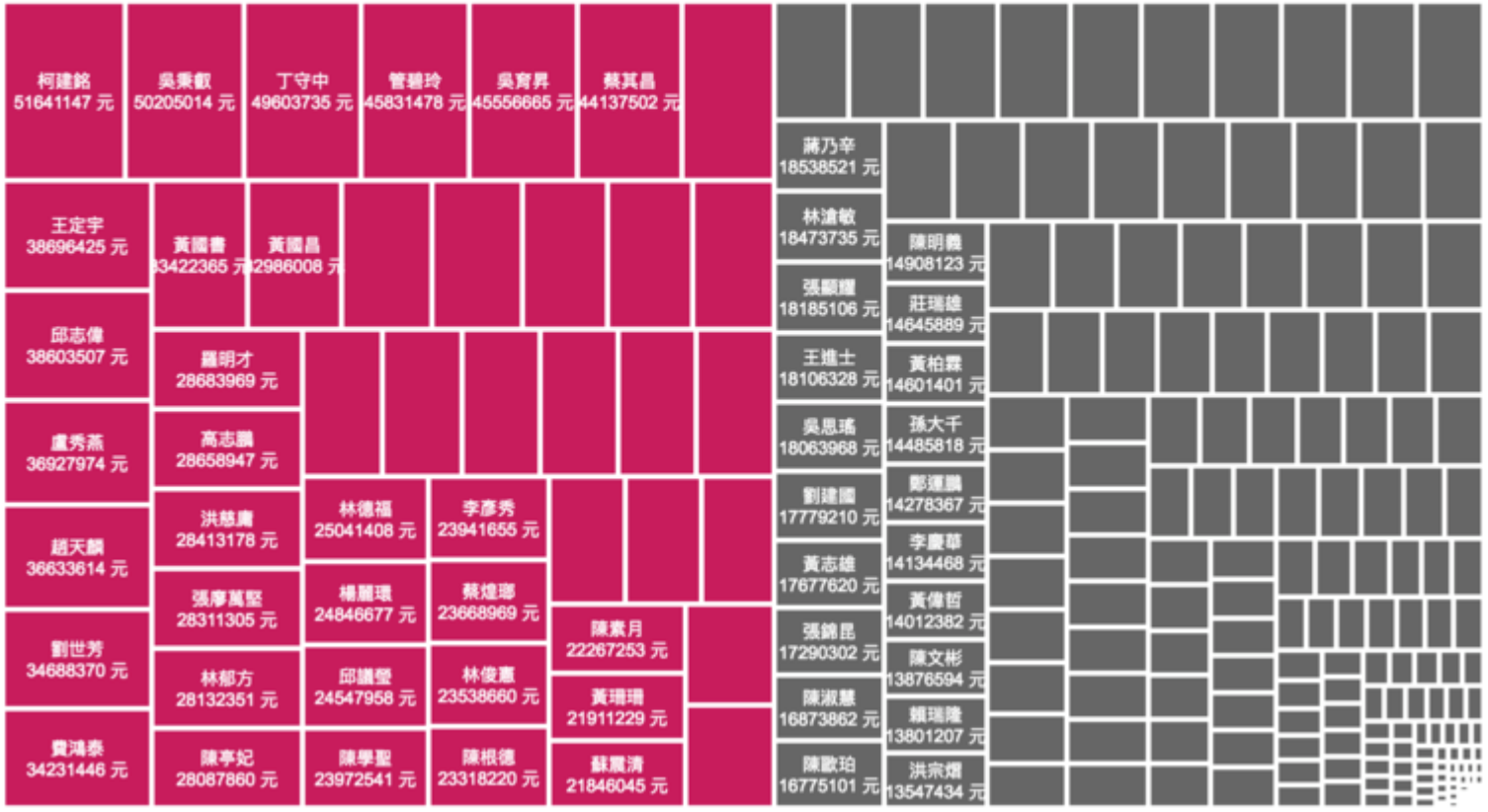
模擬

遊戲

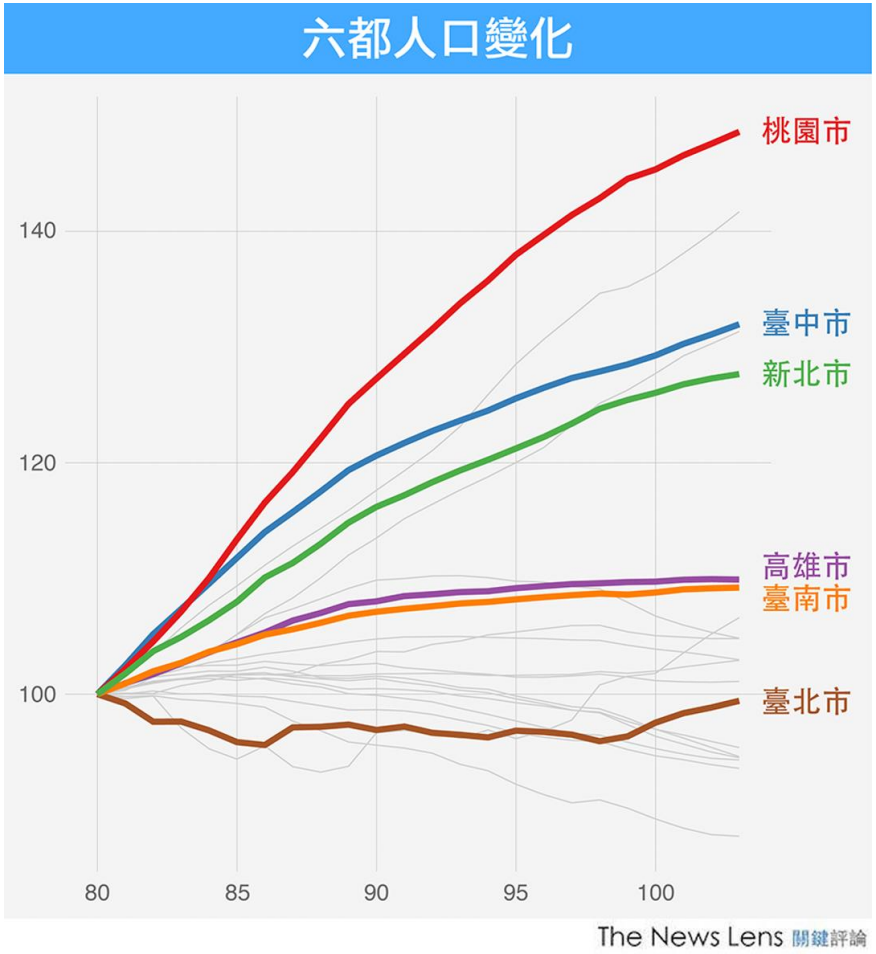
鏡周刊-數讀政治獻金

排名	姓名	政治獻金總收入	營利事業佔總收入比	當選	政治世家（註）
1	黃國昌	32,986,008	14.93%	☮	
2	洪慈庸	28,413,178	17.59%	☮	
3	蘇巧慧	20,398,298	32.89%	☮	✓
4	鍾易仲	16,367,803	15.21%		✓
5	張鎔麒	15,464,175	54.05%		✓
6	陳文彬	13,876,594	39.09%		
7	賴瑞隆	13,801,207	47.6%	☮	
8	蔣萬安	13,466,432	30.63%	☮	✓
9	林昶佐	12,580,453	19.51%	☮	
10	黃韻涵	10,241,024	30.72%		✓

48 名候選人（20%）掌握了 50% 的政治獻金



關鍵評論-中南部人口減少有多嚴重？



Dashboard

Tesla Model S

COCKPIT DISPLAY

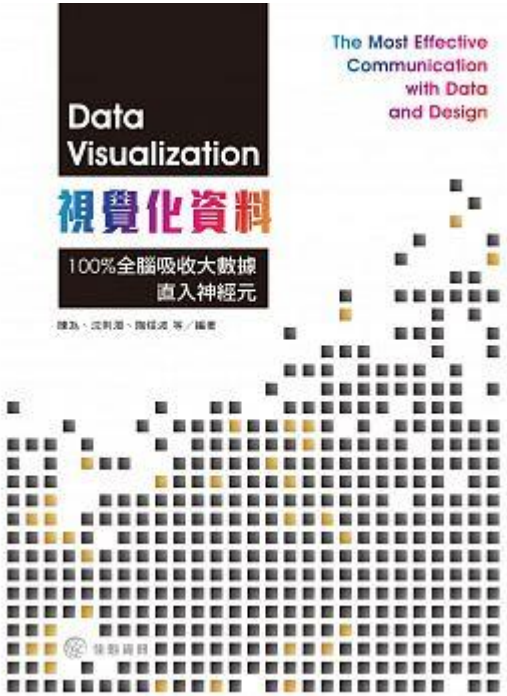
SPACEX- CREW DRAGON DISPLAYS

InfoVis Resource

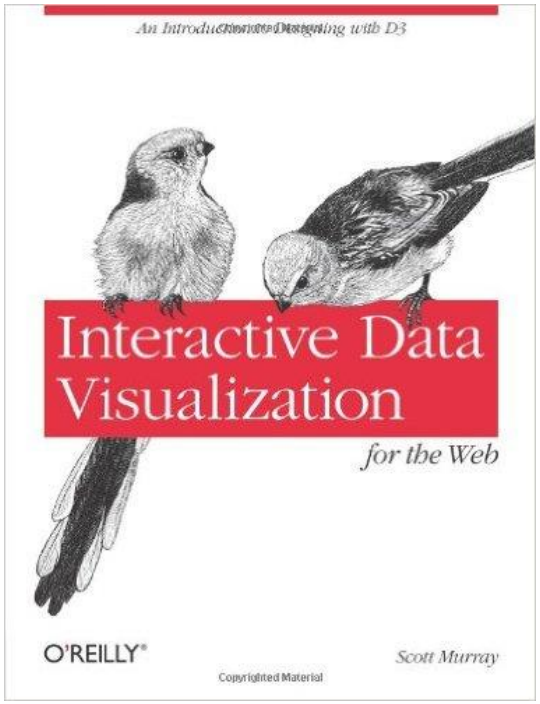
- NY Times - Flowingata
- Information is beautiful awards
- D3.js - gallery

RESOURCE

Reference book



視覺化資料 –100% 全腦吸收大數據，直入神經元。
陳為、沈則潛、陶煜波。
佳魁資訊。2014



Interactive Data Visualization for the Web: An Introduction to Designing with D3
Scott Murray
O'Reilly Media; 1 edition (April 5, 2013) (2ed Oct, 2017)

Reference conference

- IEEE VIS conference
  - VAST, InfoVis, SciVIs
- IEEE PacificVis
- ACM Siggraph

SciVis

- continuous data
- inherent spatial position

- from computer graphics
- algorithmic focus

InfoVis

- discrete data
- abstract

- from HCI
- usability focus

Reference

- Stanford Vis Group
  - http://vis.stanford.edu/papers/
- UCDavis center for visualization
  - http://vis.ucdavis.edu/
- 浙江大学可视分析小组
  - http://www.cad.zju.edu.cn/home/vagblog/

Chapter 1

INTRODUCTION

Definition

- Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

  Tamara Munzner

- Why?

- Trade-off

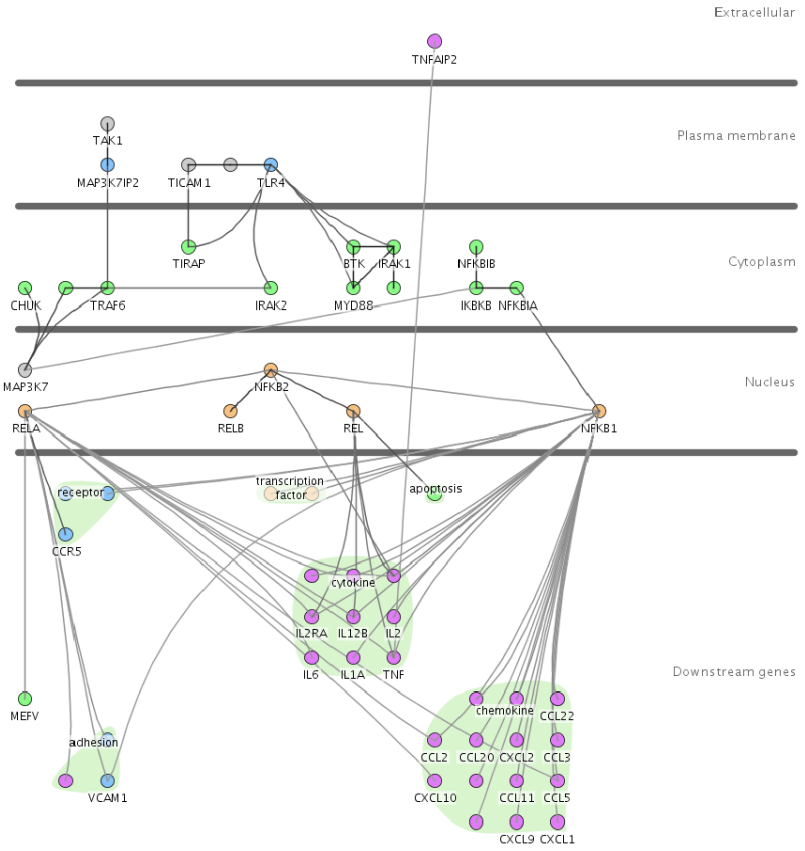
Why have a human in the loop?

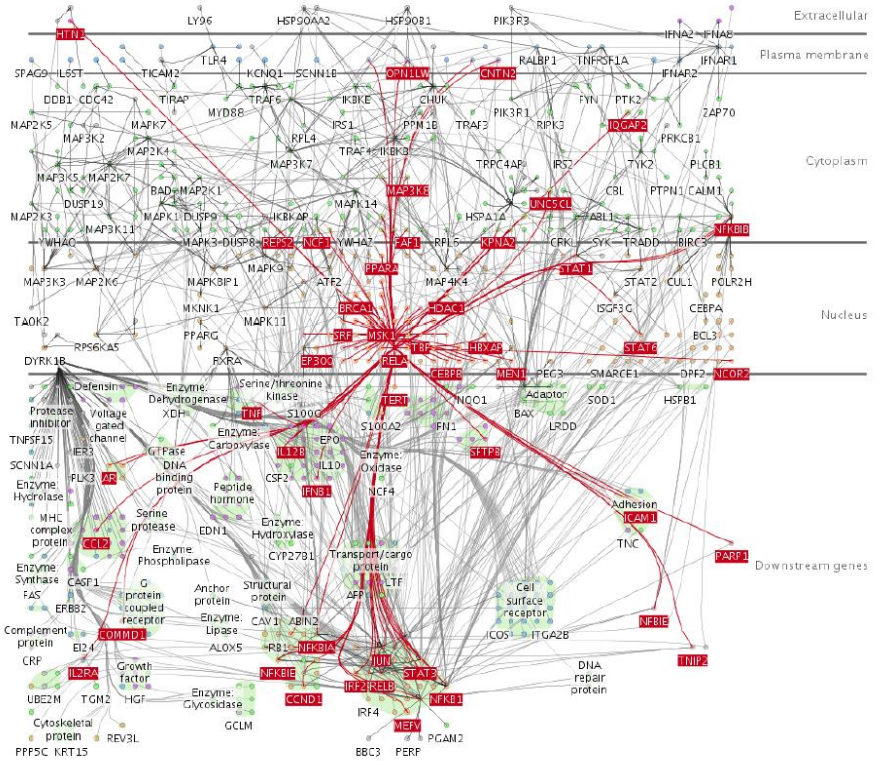
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- don't need vis when fully automatic solution exists and is trusted
- many analysis problems ill-specified
  - don't know exactly what questions to ask in advance
- possibilities
  - long-term use for end users (e.g. exploratory analysis of scientific data)
  - presentation of known results
  - stepping stone to better understanding of requirements before developing models
  - help developers of automatic solution refine/debug, determine parameters
  - help end users of automatic solutions verify, build trust

Why have a Computer in the loop?

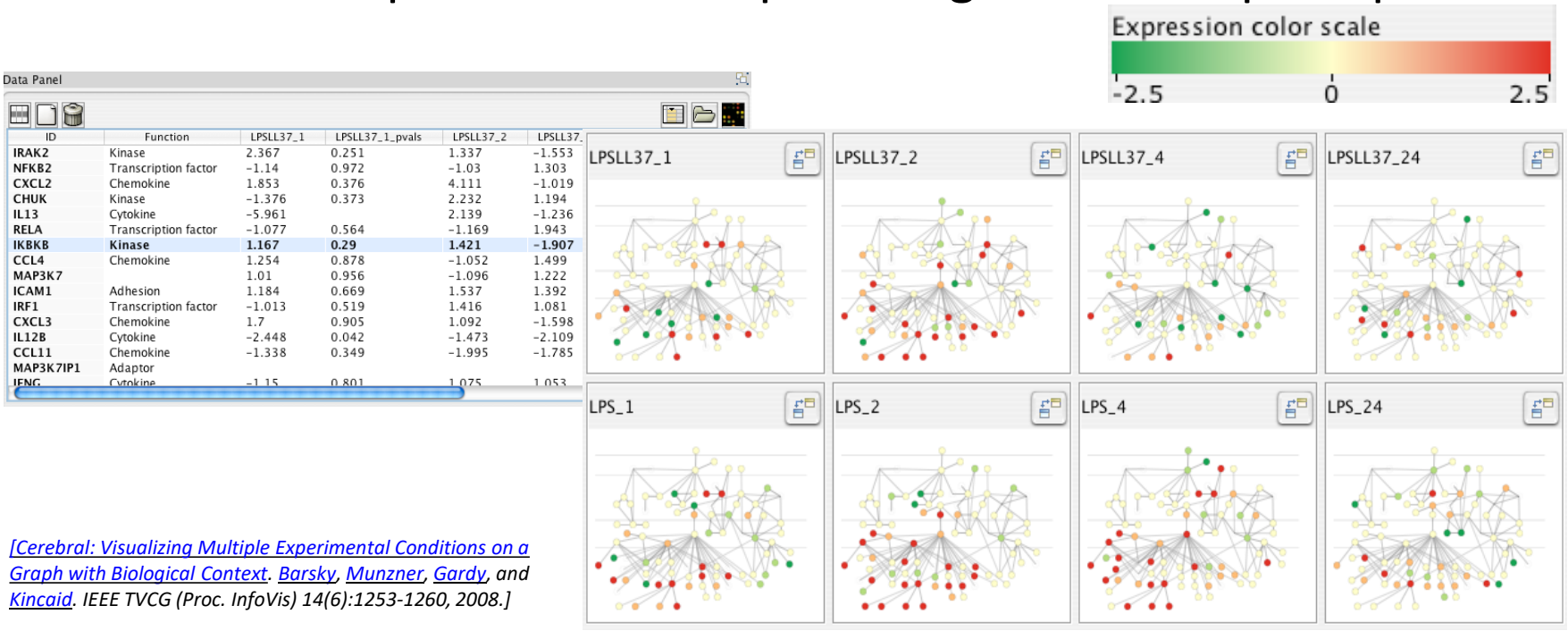




Why use an external representation?

Computer-based visualization systems provide **visual representations** of datasets designed to help people carry out tasks more effectively.

- external representation: replace cognition with perception



[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. IEEE TVCG (Proc. InfoVis) 14(6):1253-1260, 2008.]

Why depend on vision?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- human visual system is high-bandwidth channel to brain
  - overview possible due to background processing
    - subjective experience of seeing everything simultaneously
    - significant processing occurs in parallel and pre-attentively
- sound: lower bandwidth and different semantics
  - overview not supported
    - subjective experience of sequential stream
- touch/haptics: impoverished record/replay capacity
  - only very low-bandwidth communication thus far
- taste, smell: no viable record/replay devices

Why represent all the data?

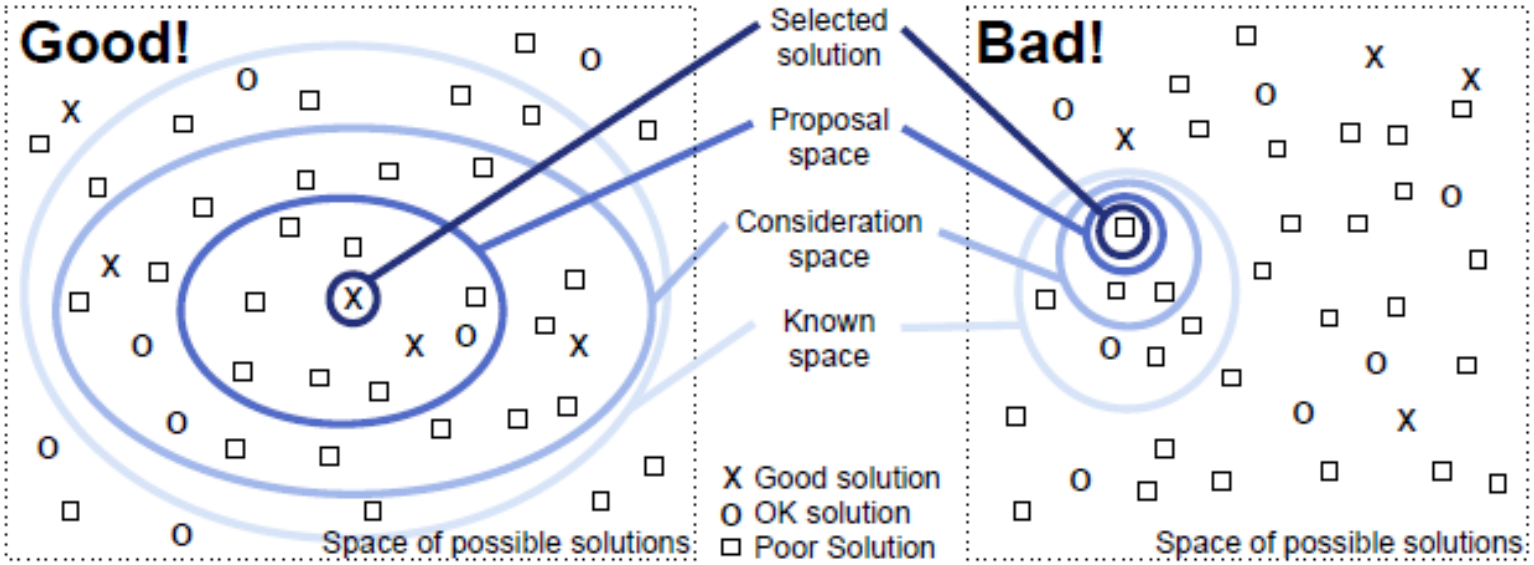
	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Why focus on tasks and effectiveness?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- tasks serve as constraint on design (as does data)
  - idioms do not serve all tasks equally!
  - challenge: recast tasks from domain-specific vocabulary to abstract forms
- most possibilities ineffective
  - validation is necessary, but tricky
  - increases chance of finding good solutions if you understand full space of possibilities
- what counts as effective?
  - novel: enable entirely new kinds of analysis
  - faster: speed up existing workflows

Why are most designs ineffective?

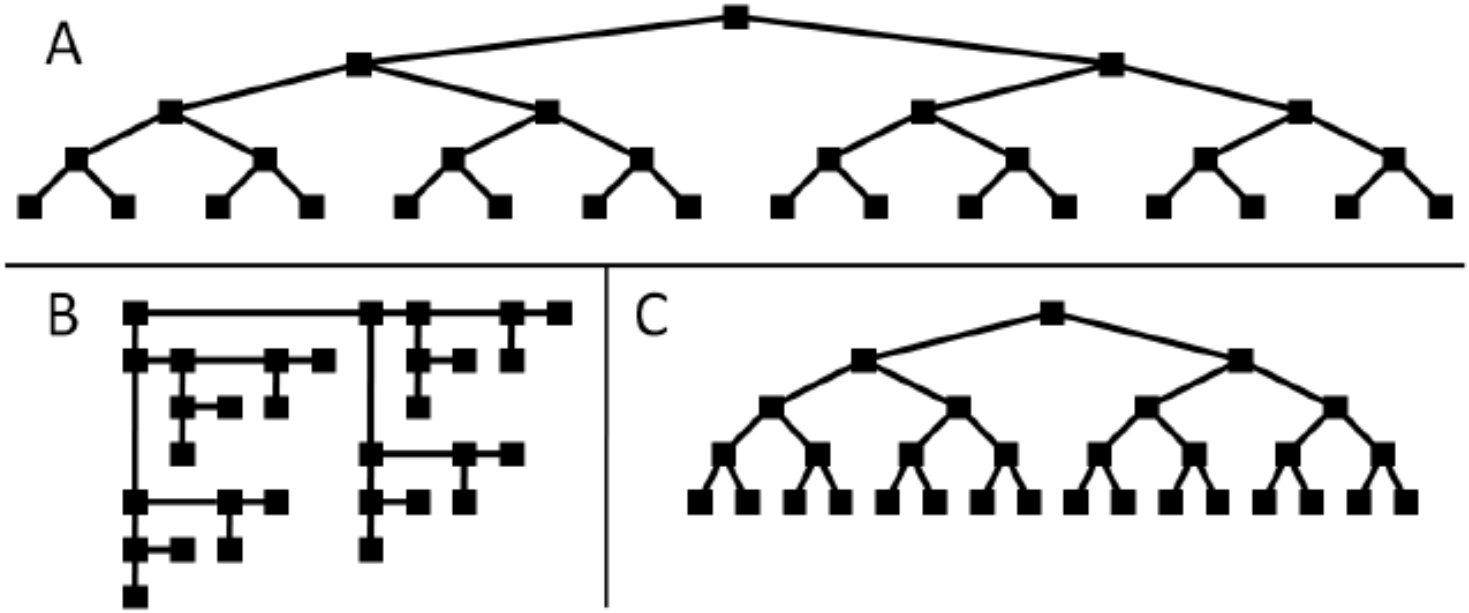


What resource limitations are we faced with?

Vis designers must take into account three very different kinds of resource limitations: those of computers, of humans, and of displays.

- Computational limits
  - processing time
  - system memory
- Human limits
  - human attention and memory
- Display limits
  - pixels are precious resource, the most constrained resource
  - information density: ratio of space used to encode info vs unused whitespace
    - tradeoff between clutter and wasting space, find sweet spot between dense and sparse

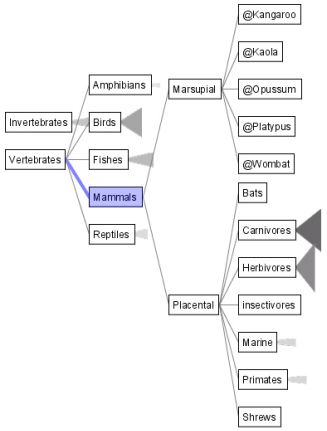
Display limits



Why analyze?

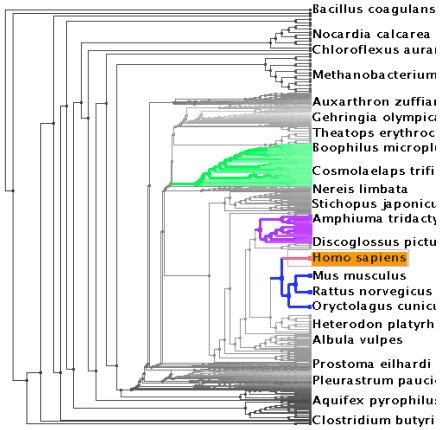
- imposes structure on huge design space
  - –scaffold to help you think systematically about choices
  - –analyzing existing as stepping stone to designing new
  - –most possibilities ineffective for particular task/data combination

SpaceTree

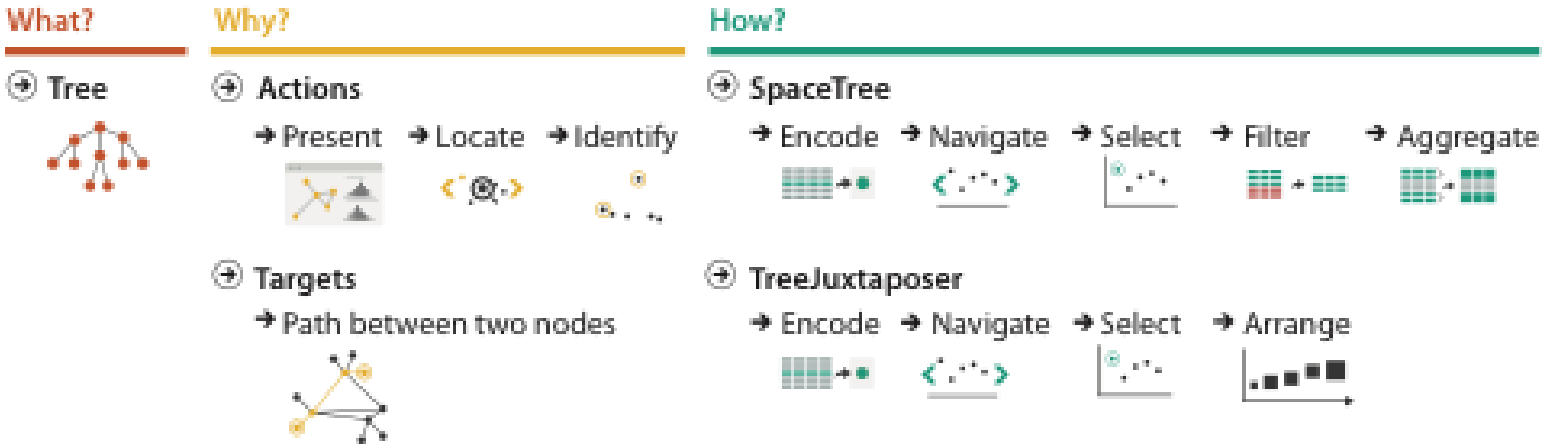


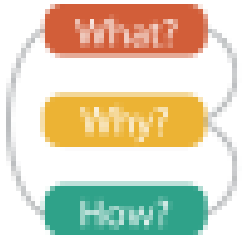
[SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation. Grosjean, Plaisant, and Bederson. Proc. InfoVis 2002, p 57–64.]

TreeJuxtaposer



[TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context With Guaranteed Visibility. ACM Trans. on Graphics (Proc. SIGGRAPH) 22:453– 462, 2003.]





Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
  - *Chap 1: What's Vis, and Why Do It?*
- A Tour through the Visualization Zoo. Communications of the ACM, 53(6), pp. 59-67, Jun 2010