

Introduction to Information Visualization

SWEN422

Human Computer Interaction

Dr Craig Anslow

craig.anslow@ecs.vuw.ac.nz

SWEN 422 – Lecture Schedule

- Week 6 – Information Visualization
- Week 7 – Information Visualization
- Week 8 – Information Visualization
- Week 9 – Gestural Interfaces
- Week 10 - AR/VR
- Week 11 - AR/VR
- Week 12 - no lectures (work on project)

https://ecs.wgtn.ac.nz/Courses/SWEN422_2024T1/LectureSchedule

Visualization

“The ability to take data –

to be able to **understand** it,
to **visualize** it,
to **communicate** it –

that’s going to be a hugely important skill in the next decades,

... because now we really do have essentially free and ubiquitous data.

So the complimentary **scarce factor is the ability to understand**
the data and extract value from it.”

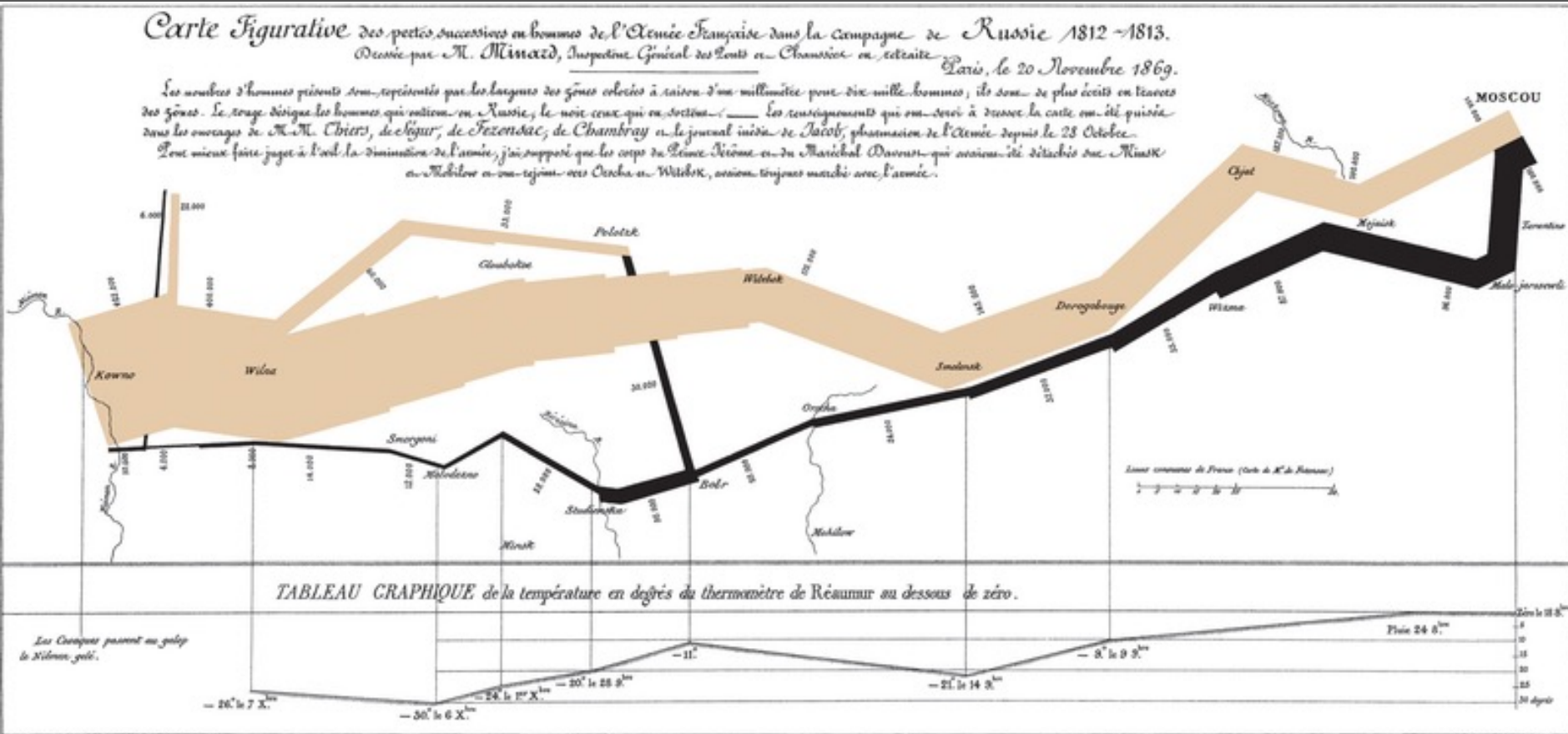
Hal Varian, Google’s Chief Economist
The McKinsey Quarterly, Jan 2009

Charles Minard – Napoleon 1812

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dessiné par M. Minard, Ingénieur Général des Ponts et Chaussées en retraite, Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en traces des zones. Le rouge désigne les hommes qui ont péri en Russie, le noir ceux qui en sont restés. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Elégar, de Fozendac, de Chambray et le journal inédit de Jacoby, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust, qui avaient été détachés sur Minsk et Mielow et qui n'ont pas été ramenés en France, avaient toujours marché avec l'armée.



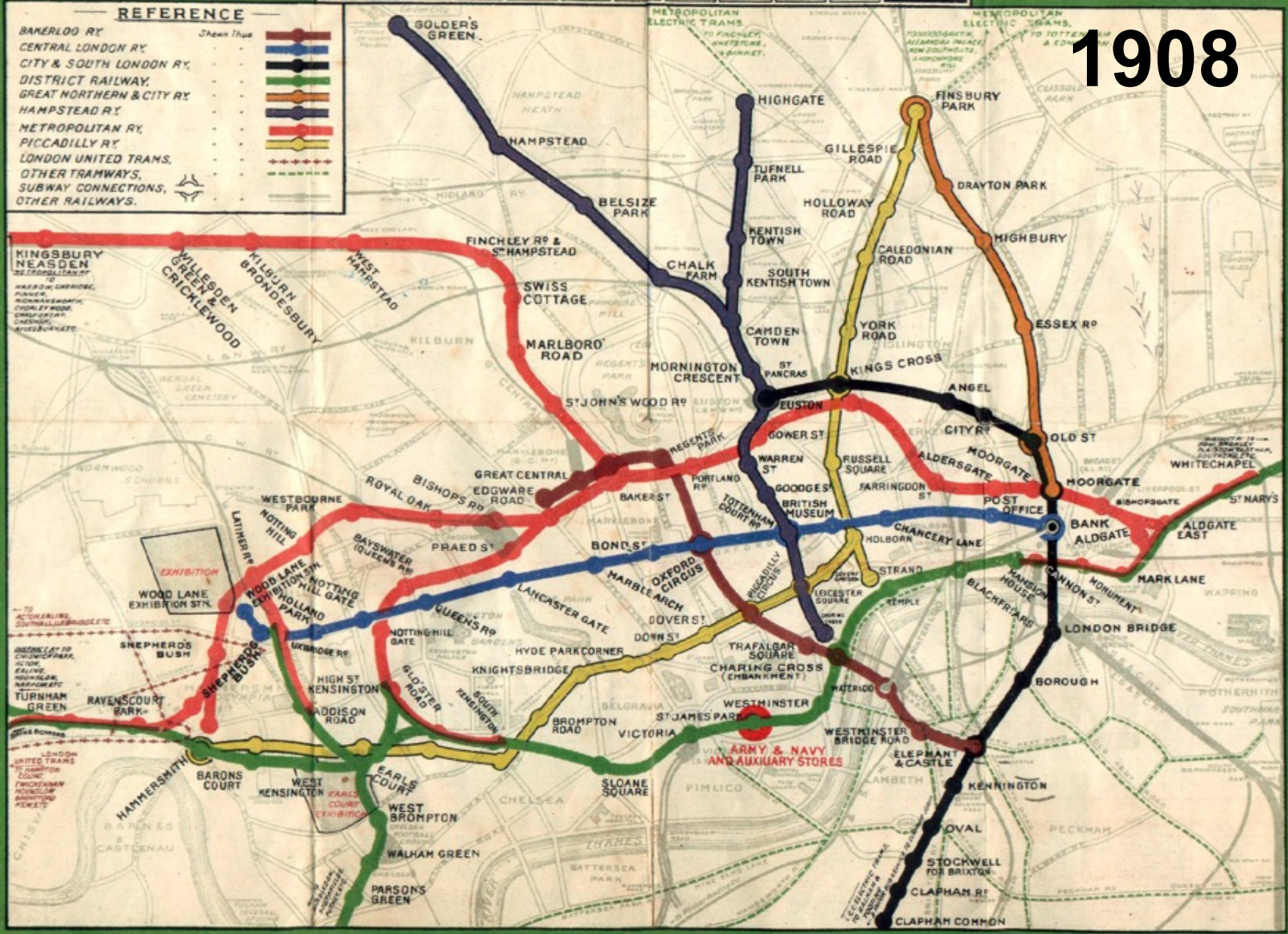
Charles Minard's map of Napoleon's disastrous Russian campaign of 1812. Graphic notable for representation in 2D of 6 types of data: # of troops; distance; temperature; the latitude and longitude; direction of travel; and location relative to dates.

UNDERGROUND

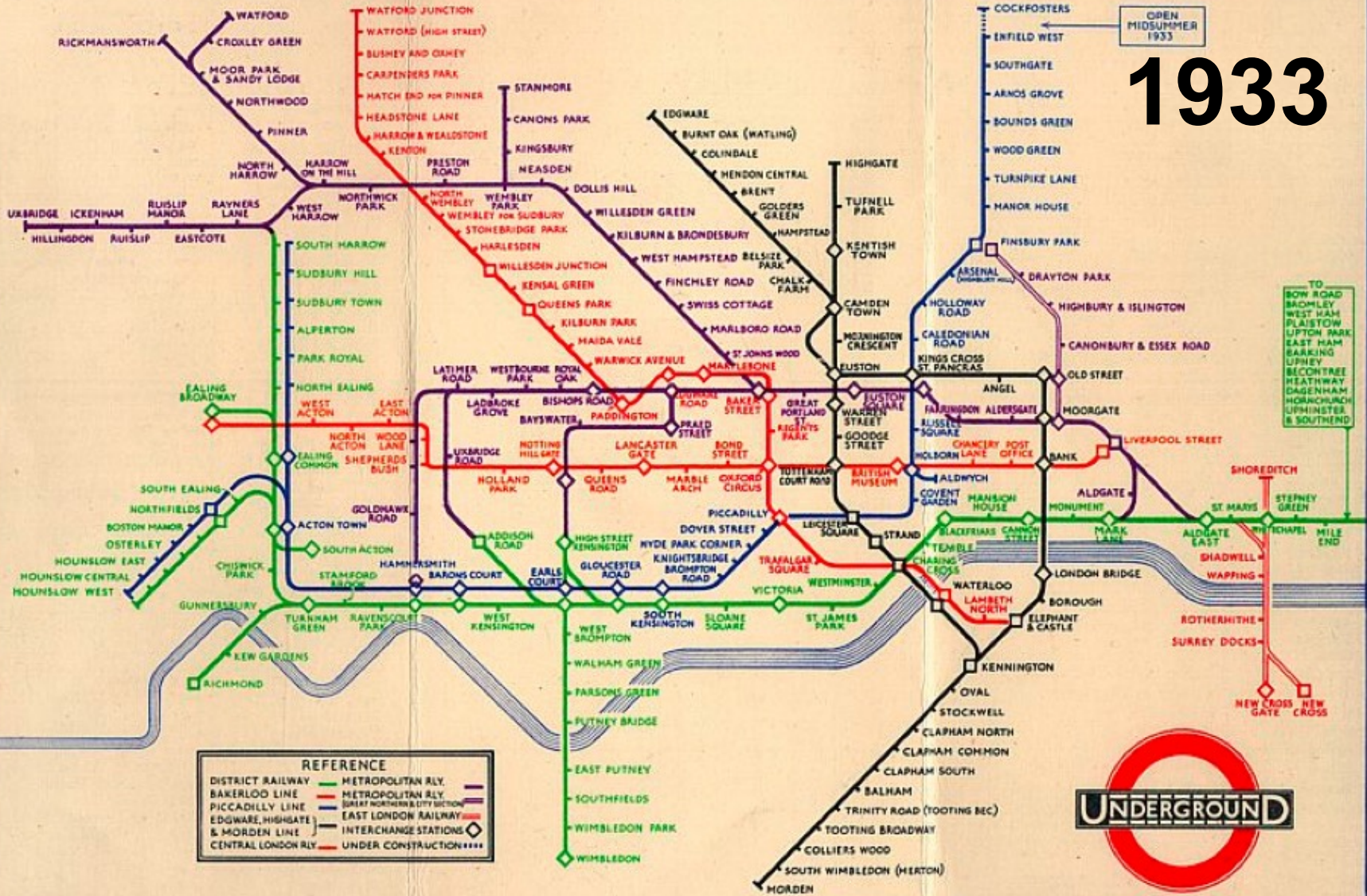
1908

REFERENCE

- BAKERLOO RY *Shown Thus*
- CENTRAL LONDON RY
- CITY & SOUTH LONDON RY
- DISTRICT RAILWAY
- GREAT NORTHERN & CITY RY
- HAMPSTEAD RY
- METROPOLITAN RY
- PICCADILLY RY
- LONDON UNITED TRAMS
- OTHER TRAMWAYS
- SUBWAY CONNECTIONS
- OTHER RAILWAYS



1933

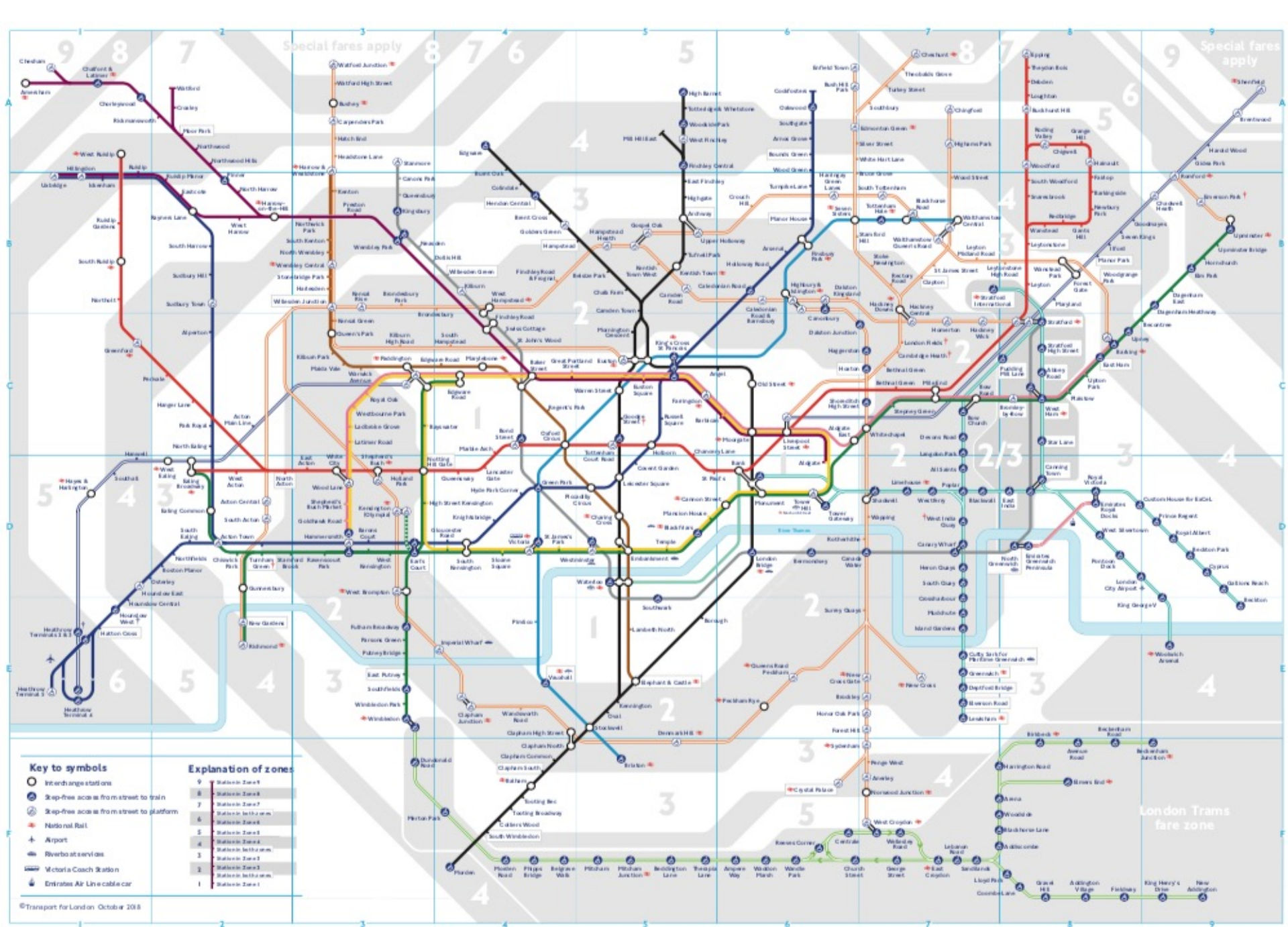


TO
BOW ROAD
BADMLEY
WEST HAM
PLAISTOW
UPTON PARK
EAST HAM
BARKING
LIPNEY
BECHTREA
HEATHWAY
DAGENHAM
HORSHAM
UPMINSTER
& SOUTHBEND

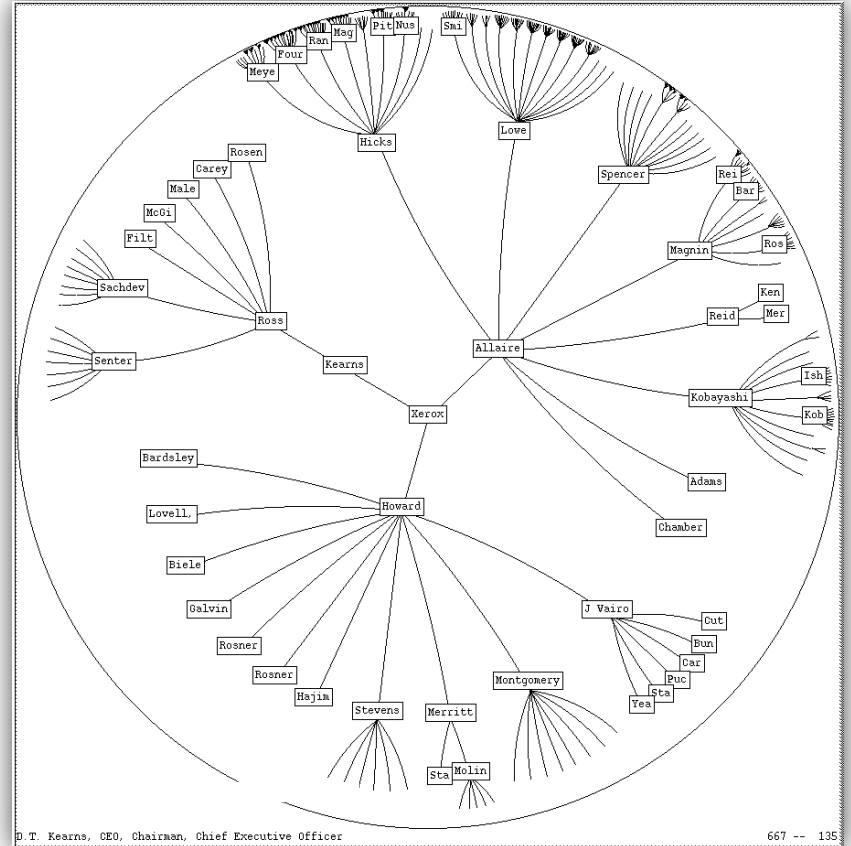


H.C. BECK

https://en.wikipedia.org/wiki/Harry_Beck



InfoGraphics vs Visualization



A **Visualization** can be applied to many datasets; an **Infographic** is created for a particular dataset.

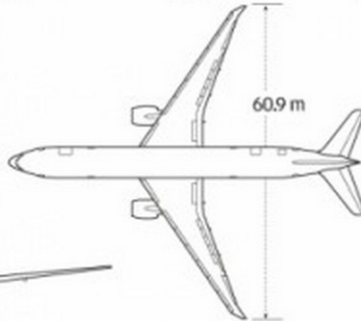
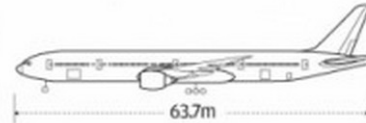
Malaysia Airlines flight MH370

Still a mystery

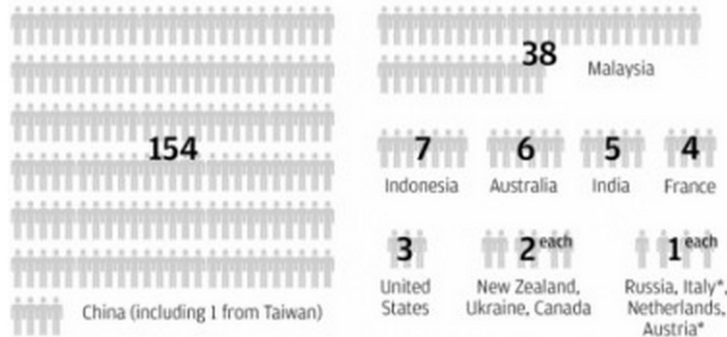
Flight MH370 with 239 people on board vanished yesterday over the South China Sea near the Vietnam coastline. There were more than 150 Chinese nationals on board.

Boeing 777-200ER (Flight MH 370)

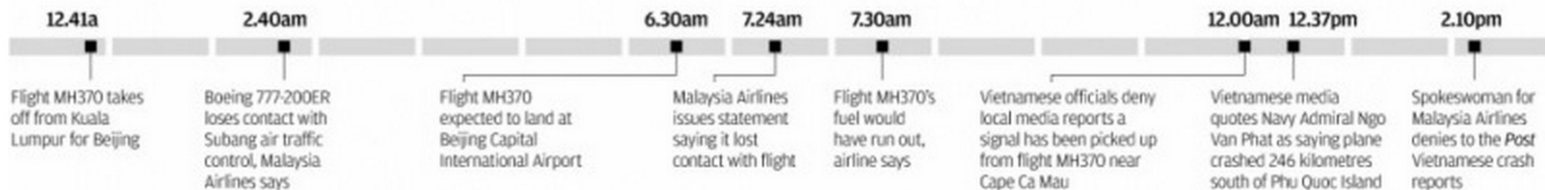
Passengers	Up to 292
Engines (Max. thrust)	Rolls-Royce Trent 895 42,365kg
Fuel	171,170 litres
Range	7,725 nautical miles (14,305km)
Cruise speed	0.84 Mach



On board flight MH370



*Italian and Austrian passports were reported stolen



What is Visualization?

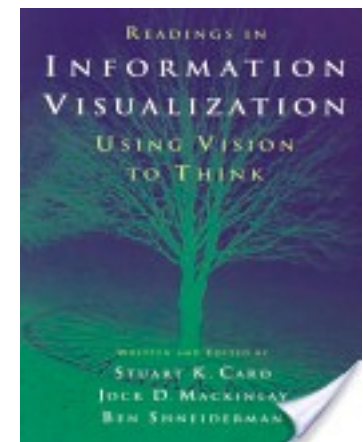
- Visual representation of data
- “Transformation of the symbolic into the geometric”
[McCormick et al, 1987]
- “... artificial memory that best supports our natural means of perception” [Bertin, 1967]
- “Use of computer-generated, interactive, visual representations of data to amplify cognition”
[Card, Mackinlay, & Shneidermann, 1999]
- “The use of visual representations to explore, make sense of, and communicate data.” [Few, 2014]

Information Visualization

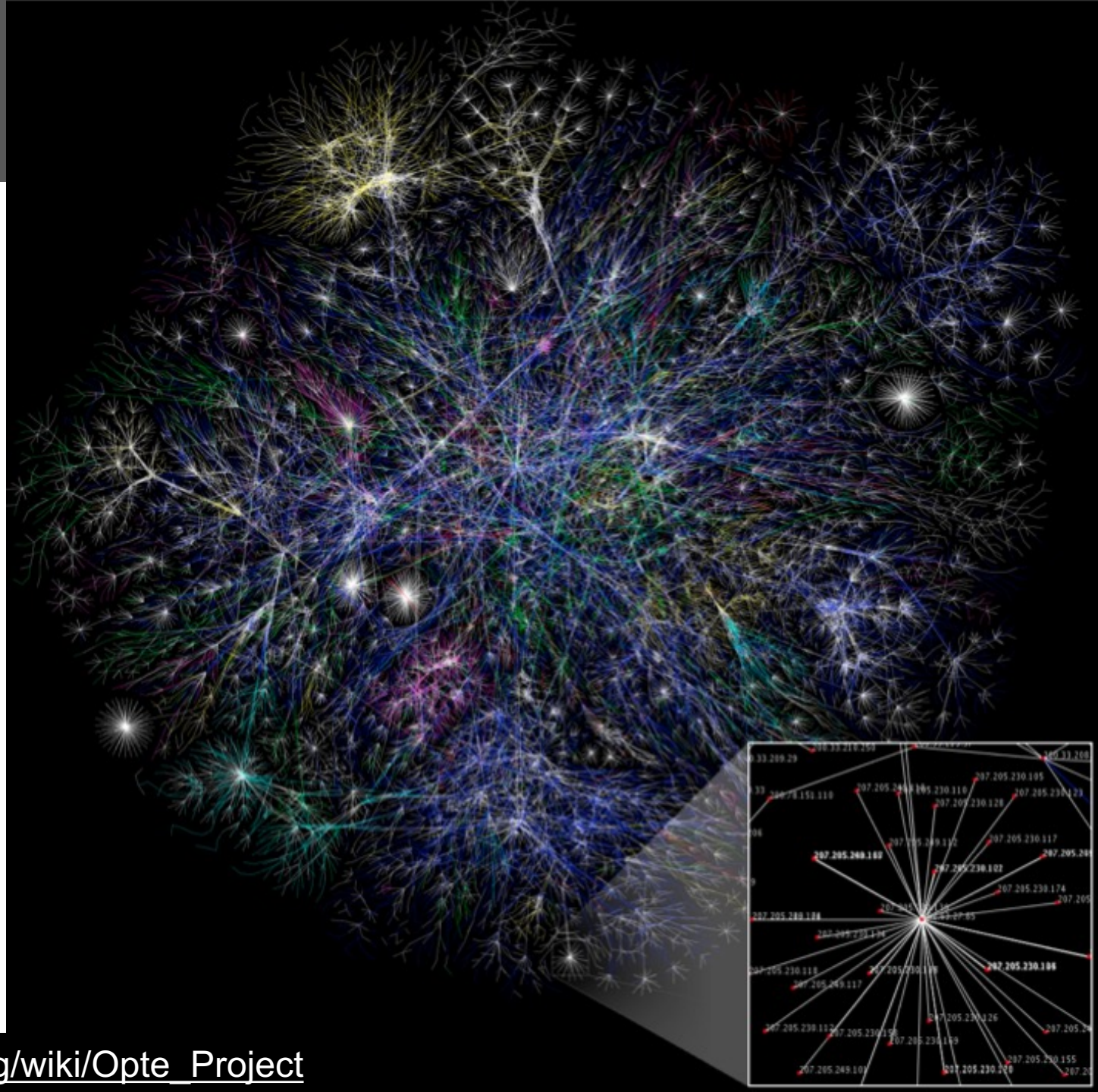
Information Visualization is the use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image, it is mapped to screen space. The image can be changed by users as they proceed working with it. This interaction is important as it allows for constant redefinition of goals when new insight into the data has been gained.

Card, Mackinlay, Shneiderman.

Readings in Information Visualization: Using Vision to Think
Morgan Kaufmann, 1999

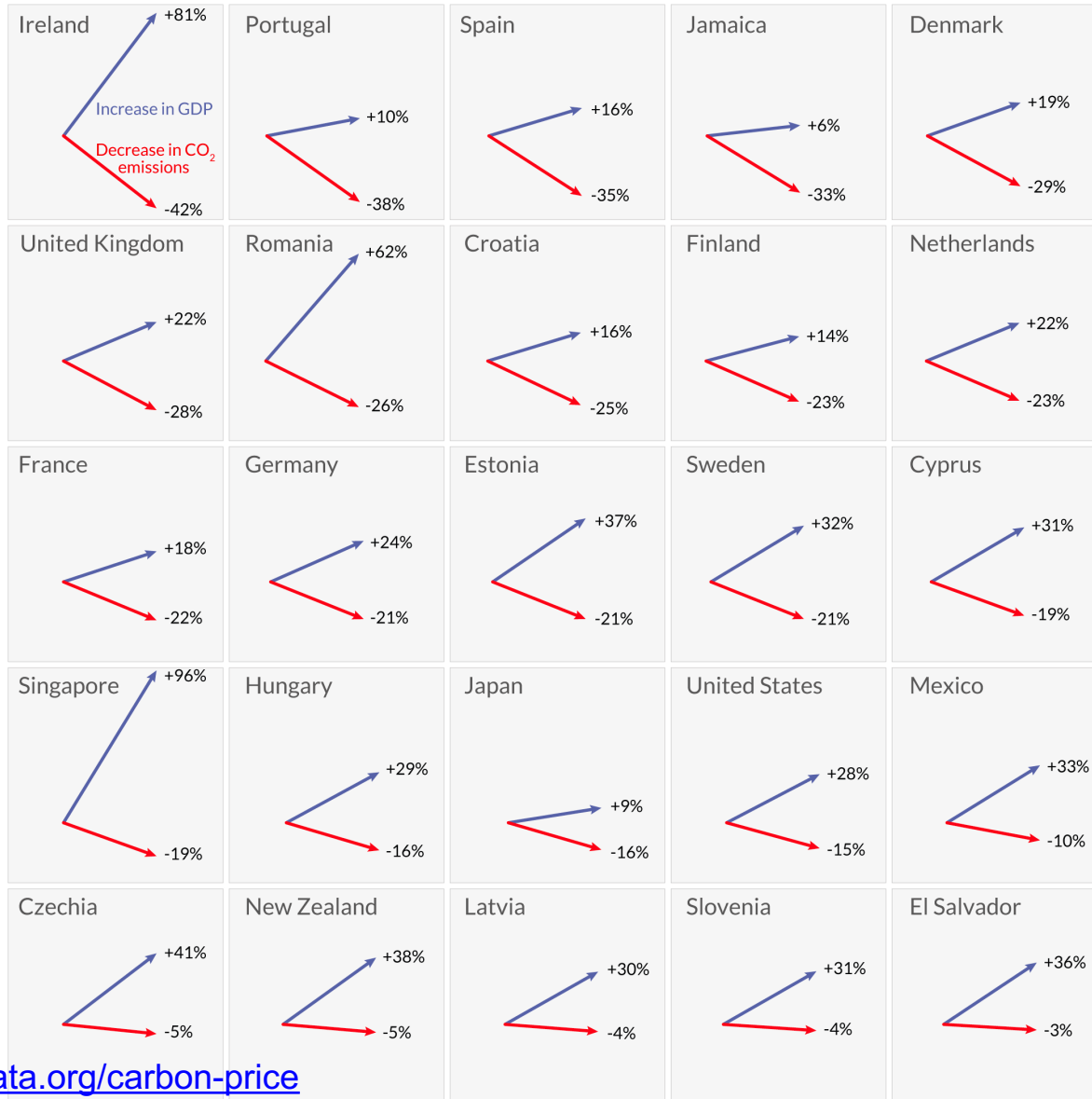


OPTE



Decoupling: Countries that achieved economic growth while reducing CO₂ emissions, 2005–19

Emissions are adjusted for trade. This means that CO₂ emissions caused in the production of imported goods are added to its domestic emissions – and for goods that are exported the emissions are subtracted.



<https://ourworldindata.org/carbon-price>

Interactive Examples

- **Word Clouds:**

- <https://worditout.com/>
- <http://www.wordle.net/> Not Wordle Game <https://wordlegame.org/>

- **Data:**

- Arden <https://www.newsroom.co.nz/2018/09/27/256105/full-text-pms-speech-to-the-united-nations>
- Luxon <https://www.national.org.nz/speechstateofthenation>

- **Baby Name Wizard:** <http://www.babynamewizard.com/>

- **Sorting Algorithms:**

- <http://www.sorting-algorithms.com/>
- <https://softvis.wordpress.com/sorting-algorithms>

Gource

Gource

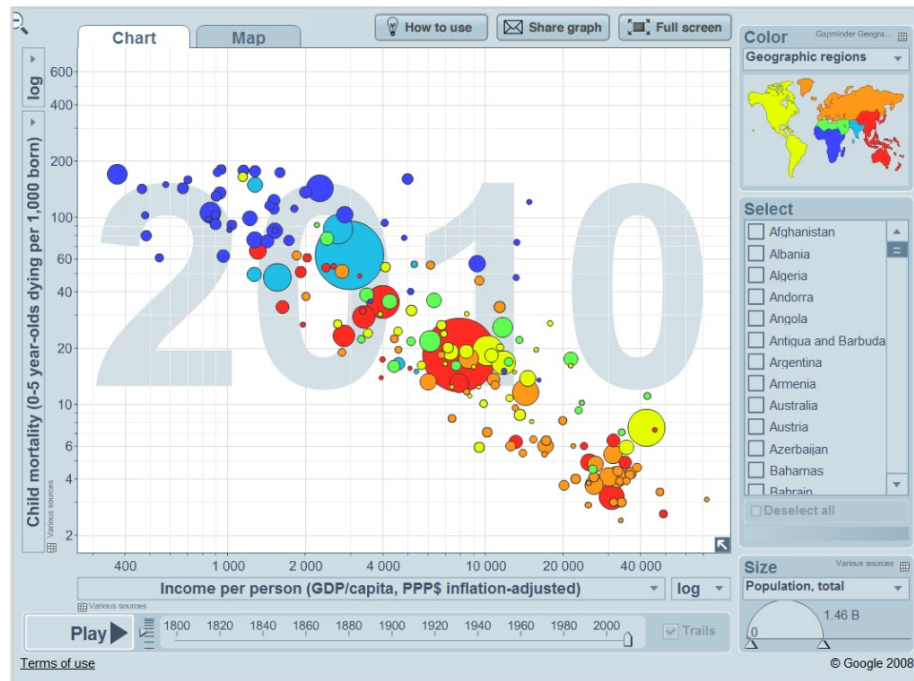
Software Version Control Visualization
gource.googlecode.com/



<http://gource.io/>

GapMinder – Hans Rosling

Debunking Myths About the “Third World”

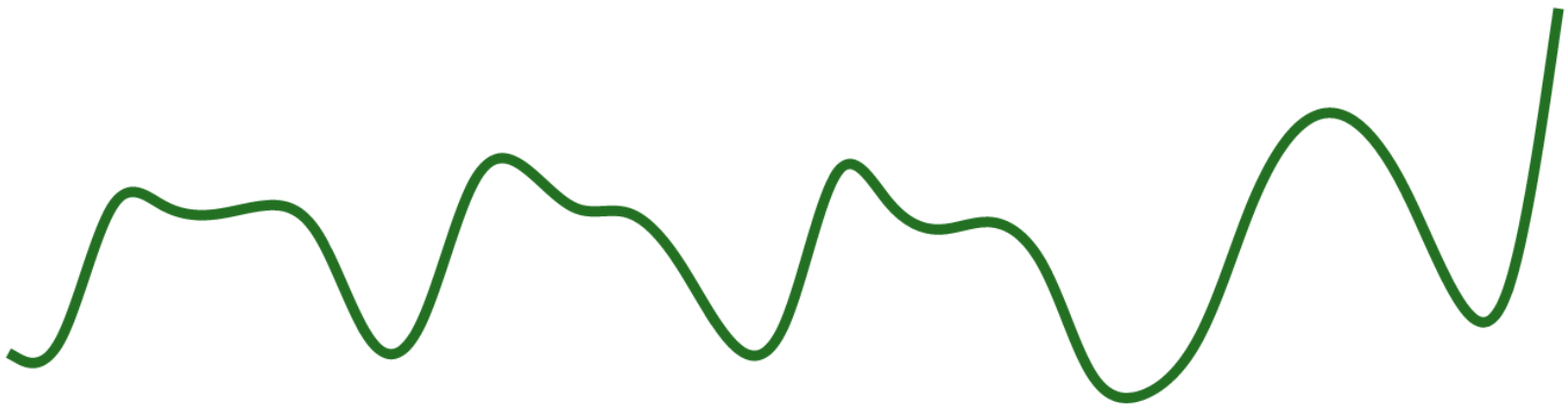


http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen
<http://www.gapminder.org/>

What is Visualization?

1. Based on non-visual data
2. Produce an image
3. Result must be readable and recognizable

Robert Kosara



Why do we Visualize?

- **Visual bandwidth** is enormous

Why do we Visualize?

- **Visual bandwidth** is enormous
- **Human perceptual skills** are remarkable
 - Trend, cluster, gap, outlier, pattern ...
 - Color, size, shape, proximity...

Perception

How many 3s?

1 8 4 7 9 5 3 2 1 2 4 6 7 8 9 5 6 4 3
4 8 0 6 4 8 0 3 2 8 8 7 9 6 2 3 1 0 6
9 9 6 3 4 4 2 6 8 1 5 6 8 7 9 0 3 2 1
1 5 6 8 7 9 6 5 1 2 3 5 9 9 7 8 9 6 5
4 3 2 1 3 2 1 5 4 9 8 3 4 2 5 8 4 8 9
2 2 1 5 6 7 8 6 5 6 3 1 4 5 1 3 4 5 1

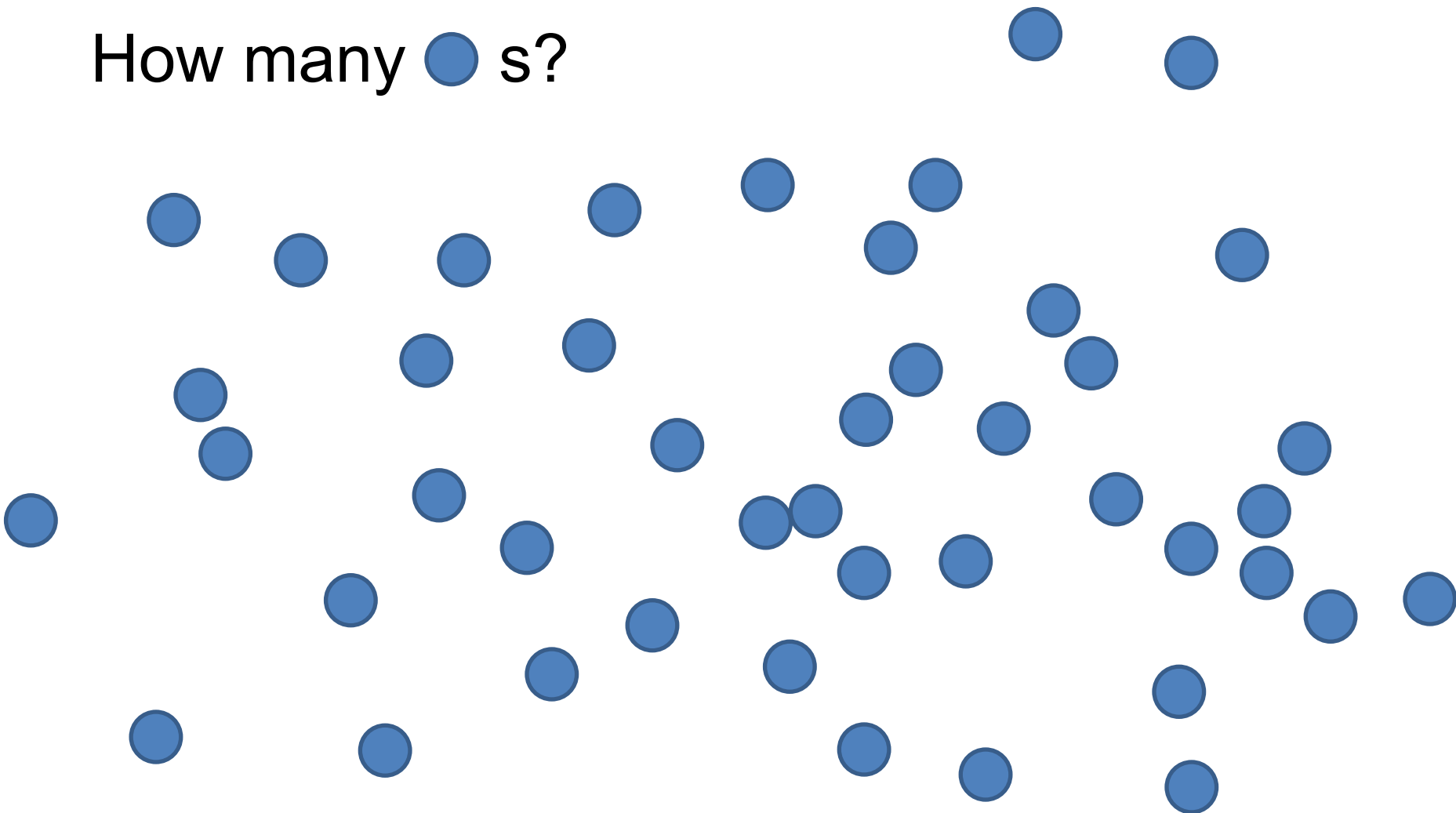
Perception

How many 3s?

1 8 4 7 9 5 **3** 2 1 2 4 6 7 8 9 5 6 4 **3**
4 8 0 6 4 8 0 **3** 2 8 8 7 9 6 2 **3** 1 0 6
9 9 6 **3** 4 4 2 6 8 1 5 6 8 7 9 0 **3** 2 1
1 5 6 8 7 9 6 5 1 2 **3** 5 9 9 7 8 9 6 5
4 **3** 2 1 **3** 2 1 5 4 9 8 **3** 4 2 5 8 4 8 9
2 2 1 5 6 7 8 6 5 6 **3** 1 4 5 1 **3** 4 5 1

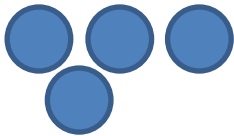
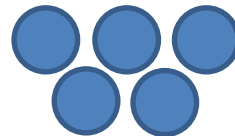
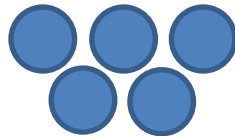
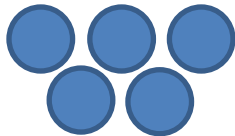
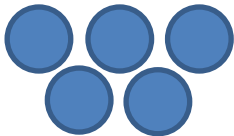
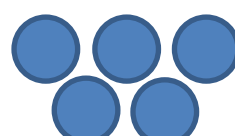
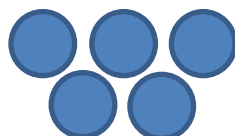
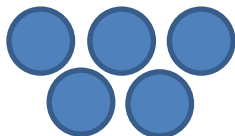
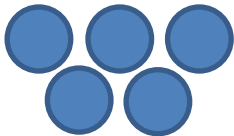
Perception

How many ● s?



Perception

How many ● s?



Set A		Set B		Set C		Set D	
X	Y	X	Y	X	Y	X	Y
10	8.08	10	9.14	10	7.47	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.11	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

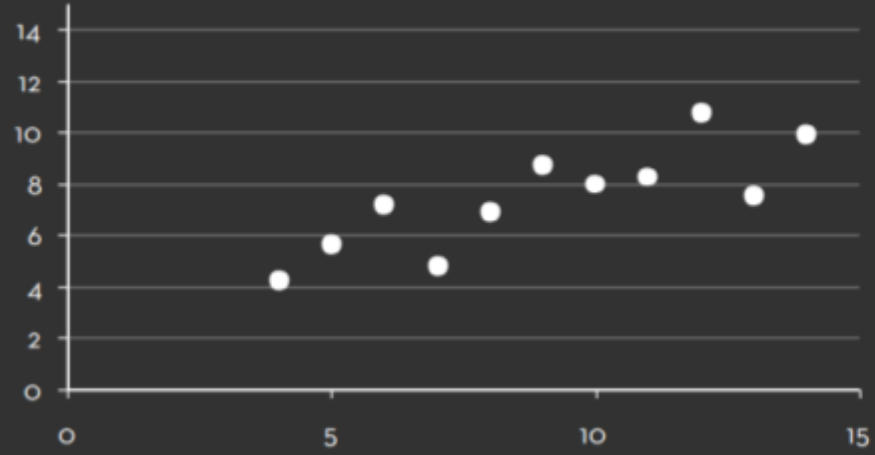
Summary Statistics

$u_X = 9.0$ $\sigma_X = 3.317$

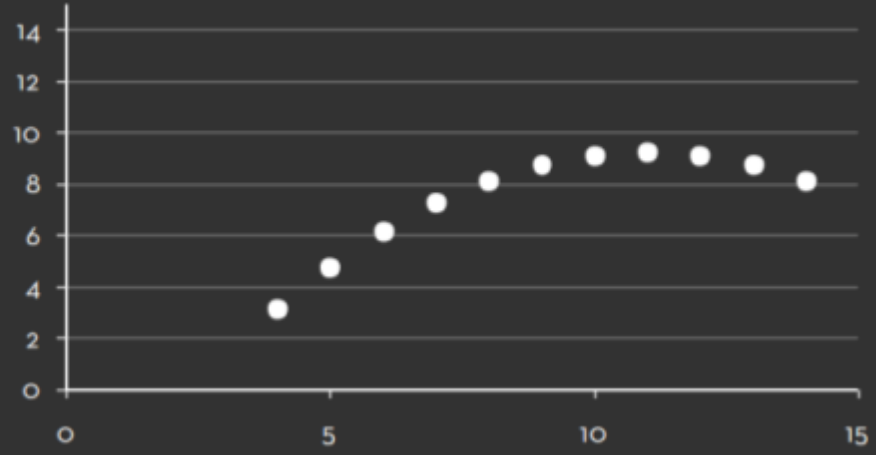
$u_Y = 7.5$ $\sigma_Y = 2.03$

Anscombe's Quartet (Anscombe, Francis J., 1973)

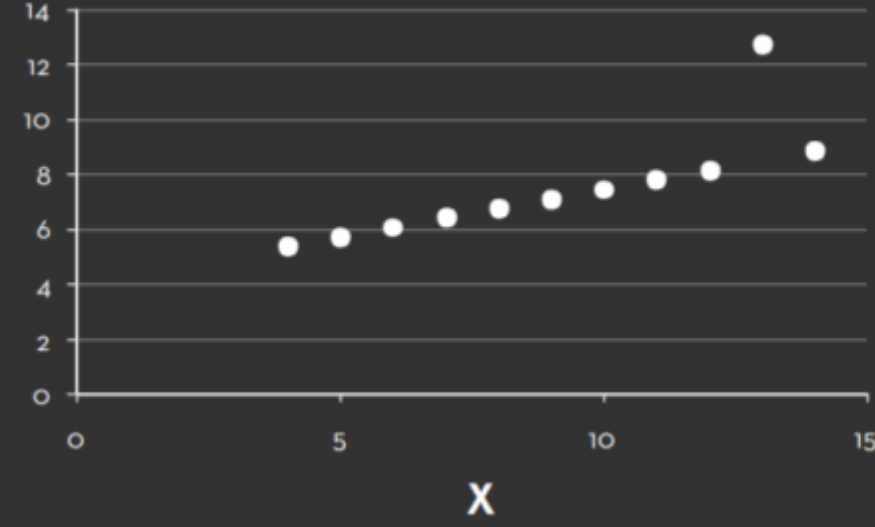
Set A



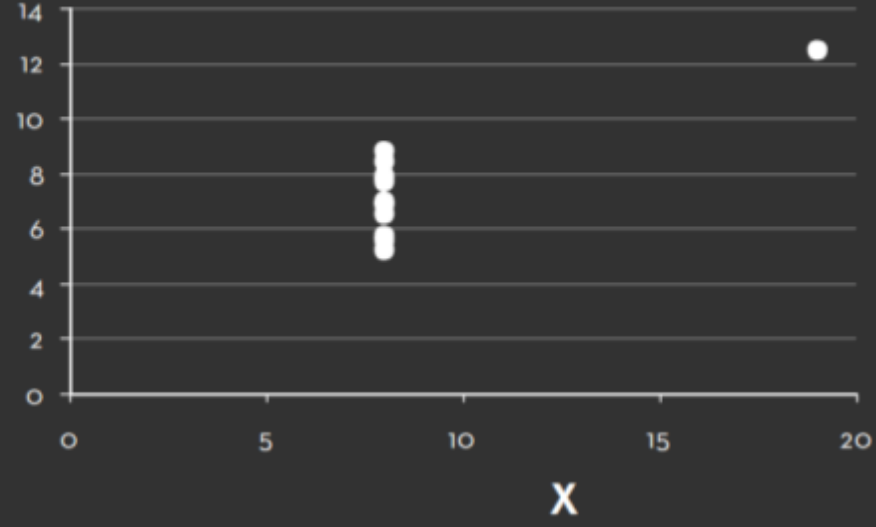
Set B



Set C



Set D



Why do we Visualize?

- **Visual bandwidth** is enormous
- **Human perceptual skills** are remarkable
 - Trend, cluster, gap, outlier, pattern...
 - Color, size, shape, proximity...
- **External representation**
 - **Reduces load** on working memory
 - **Offload** cognition

External Representation

Paper

Mental Buffer

57

X 48

External Representation

Paper

Mental Buffer

57
X 48

$$7 * 8 = 56$$

External Representation

Paper	Mental Buffer
$\begin{array}{r} 5 \\ 57 \\ \times 48 \\ \hline 6 \end{array}$	$7 * 8 = 56$

External Representation

Paper	Mental Buffer
$\begin{array}{r} 5 \\ 57 \\ X 48 \\ \hline 6 \end{array}$	

External Representation

Paper

Mental Buffer

$$\begin{array}{r} 5 \\ 57 \\ \times 48 \\ \hline 456 \end{array}$$

$$5 * 8 = 40 + 5 = 45$$

External Representation

Paper

Mental Buffer

5

57

X 48

456

External Representation

Paper	Mental Buffer
$\begin{array}{r} \cancel{25} \\ 57 \\ \times 48 \\ \hline 456 \\ 8 \end{array}$	$4 * 7 = 28$

External Representation

Paper

Mental Buffer

$$\begin{array}{r} \cancel{25} \\ 57 \\ \times 48 \\ \hline 456 \\ 8 \end{array}$$

External Representation

Paper	Mental Buffer
$\begin{array}{r} \cancel{25} \\ 57 \\ \times 48 \\ \hline 456 \\ 228 \end{array}$	$4 * 5 = 20 + 2 = 22$

External Representation

Paper

Mental Buffer

~~25~~

57

X 48

456

228

External Representation

Paper	Mental Buffer
$\begin{array}{r} \cancel{25} \\ 57 \\ \times 48 \\ \hline \overset{1}{4}56 \\ 228 \\ \hline 2736 \end{array}$	$\begin{array}{l} 6 + 0 = 6 \\ 5 + 8 = 13 \\ 2 + 4 = 6 + 1 = 7 \\ 2 + 0 = 2 \end{array}$

External Representation

Paper

Mental Buffer

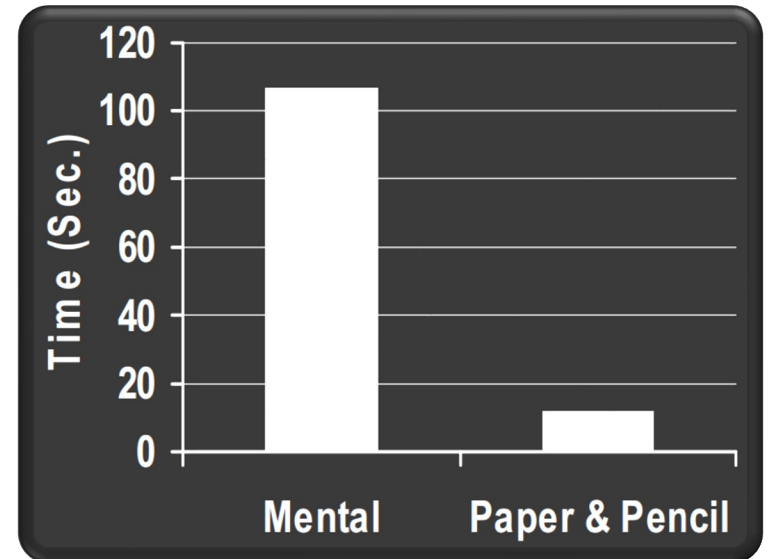
$$\begin{array}{r} \cancel{25} \\ 57 \\ \times 48 \\ \hline 456 \\ 228 \\ \hline 2736 \end{array}$$

$$6 + 0 = 6$$

$$5 + 8 = 13$$

$$2 + 4 = 6 + 1 = 7$$

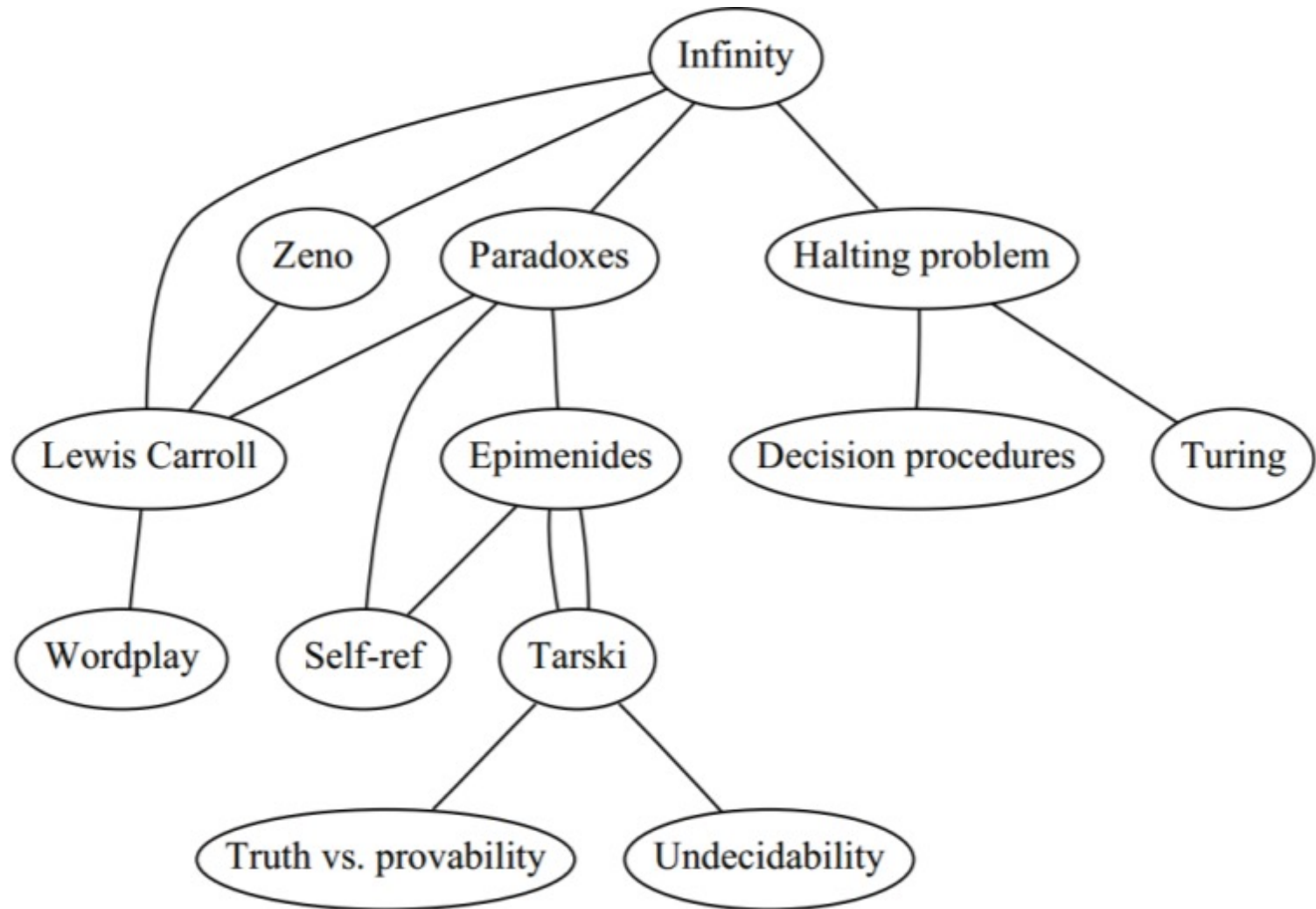
$$2 + 0 = 2$$



External Representation: Example

- Paradoxes – Lewis Carroll
- Paradoxes – Epimenides
- Paradoxes – Self-ref
- Paradoxes – Infinity
- Epimenides – Self-ref
- Epimenides – Tarski
- Halting Problem – Decision Procedure
- Halting Problem - Turing
- Infinity – Halting problem
- Infinity – Recursion
- Infinity – Zeno
- Infinity – Lewis Carroll
- Zeno – Lewis Carroll
- Lewis Carroll – Wordplay
- Tarski – Epimenides
- Tarski – Truth vs Provability
- Tarski - Undecidability

External Representation: Example



Why do we Visualize?

- **Visual bandwidth** is enormous
- **Human perceptual skills** are remarkable
 - Trend, cluster, gap, outlier, pattern...
 - Color, size, shape, proximity...
- **External representation**
 - **Reduces load** on working memory
 - **Offload** cognition
- **Fewer coded symbols** to decode

Decoding

Compare:

2947

vs

6621

vs

95.12

Decoding

Compare:

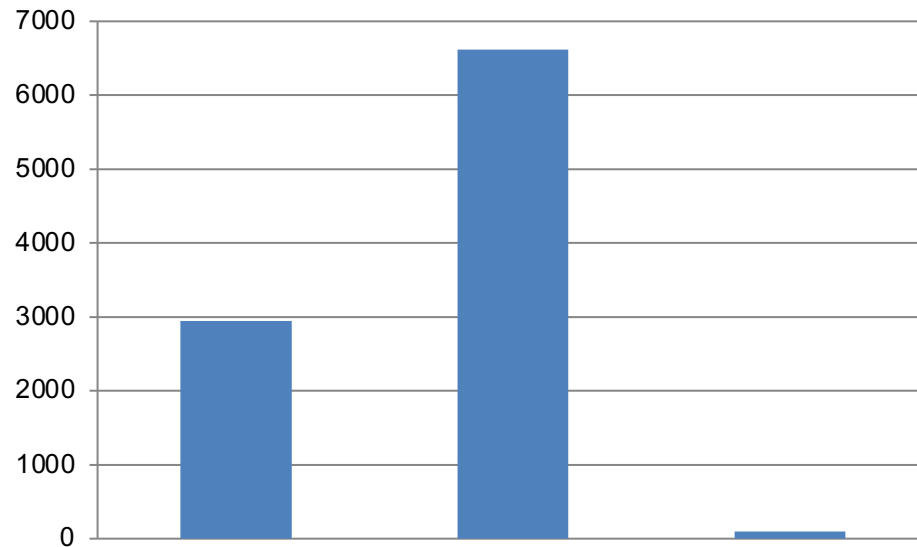
2947

vs

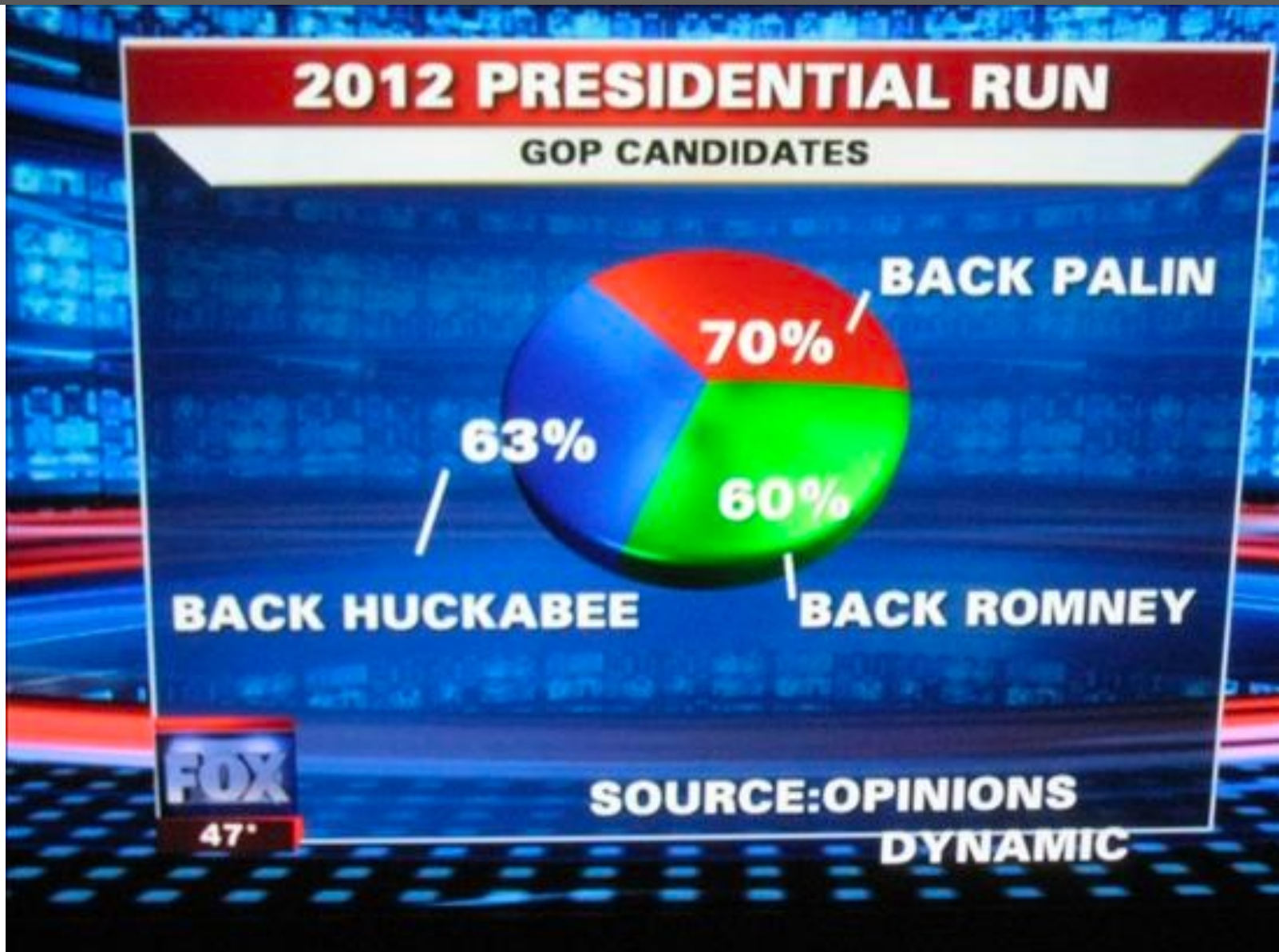
6621

vs

95.12



Why DON'T we Visualize?



Why DON'T we Visualize?

Myths:

- We visualize data because some people are visual learners.
- We visualize data for people who have difficulty understanding numbers.
- We visualize data to grab people's attention with eye-catching displays.
- Graphics provide the best means of telling stories contained in data.

Why do we Visualize?

- See the big picture
 - Don't have read & interpret each datum on its own
- Easily & rapidly compare values
 - Line graph
- See patterns among values
- Compare patterns

Visualization Resources - Books

- ***Semiology of Graphics*** by Jacques Bertin
 - *The book on visual variables*
- ***Visual Display of Quantitative Information, Beautiful Evidence, Visual Explanations, or Envisioning Information*** by Edward Tufte
 - *Beautiful examples of historic visualizations*
- ***Information Visualization: An Introduction*** by Robert Spence
 - *Overarching text book on information visualization*
- ***Visual Thinking for Design*** by Colin Ware
 - Ties perception theory and design processes to visualization practices.
- ***Visualization Analysis and Design*** by Tamara Munzner
 - Another current text book on visualization
- ***Visualizing Data*** by Ben Fry
 - *Text book on how to use Processing for visualization*
- ***Beautiful Visualization*** by Steele & Iliinsky
 - Combines techniques from artists, designers, scientists, and others.

Visualization Resources - Tools and Toolkits

- **List of Tools:**
 - <https://www.codewall.co.uk/best-javascript-chart-libraries/>
 - <http://datavisualization.ch/tools/>
 - <http://bigdata-madesimple.com/review-of-20-best-big-data-visualization-tools/>
 - <https://medium.com/javarevisited/top-javascript-chart-libraries-to-consider-in-2021-2a97c32fee3a>
 - <https://blog.logrocket.com/top-javascript-data-visualization-libraries-2021/>
- **Wordle** - <http://www.wordle.net/>
- **Tableau** - <http://www.tableausoftware.com/>
- **Power BI** - <https://powerbi.microsoft.com>
- **Google Charts Library** - <https://developers.google.com/chart/>
- **HighCharts** - <http://www.highcharts.com/>
- **D3** - <http://d3js.org/>
- **Observable** - <https://observablehq.com/>
- **Vega and Vega-Lite** - <https://vega.github.io/>
- **Prefuse** - <http://prefuse.org/>
- **Processing** - <http://www.processing.org/>
- **R** - <http://www.r-project.org/>
- **Shiny** - <https://shiny.rstudio.com>

Visualization Resources - Web Sites and Blogs

- **New York Times**
 - Dedicated team producing exceptional work.
 - <https://flowingdata.com/tag/new-york-times/>
- **Eagereyes**
 - Vis Researcher with criticism as well as overview from assorted research conferences
 - <https://eagereyes.org/>
- **Gapminder**
 - Hans Rosling's stat software & data.
 - <https://www.gapminder.org/>
- **Visual Business Intelligence**
 - Analytics blog, frequently discussing big data and dashboard design.
 - <https://www.perceptualedge.com/blog/>
- **Tableau Viz of the Day**
 - Daily example of a visualization created with Tableau software.
 - https://public.tableau.com/s/gallery?qt-overview_gallery=1

Visualization Resources

Websites and Blogs cont'd

- David McCandless - <http://www.informationisbeautiful.net/>
- Nathan Yau - <http://flowingdata.com>
- Manuel Lima - <http://www.visualcomplexity.com/vc/>
- Alark Joshi - <http://visualizeit.wordpress.com>
- Craig Anslow - <http://softvis.wordpress.com>
- Information Visualization Conference - <http://ieevis.org>
- Information Visualization Journal - <http://ivi.sagepub.com>

Readings

- [A Tour through the Visualization Zoo](#). Jeffrey Heer, Michael Bostock, Vadim Ogievetsky *Communications of the ACM*, 53(6), pp. 59-67, 2010
- [The eyes have it: a task by data type taxonomy for information visualizations](#). Ben Shneiderman. International Conference on Visual Languages 1996.
- [Considering Visual Variables as a Basis for Information Visualisation](#). Sheelagh Carpendale. Research report 2001-693-16, Department of Computer science, University of Calgary, Calgary, AB, Canada, 2003.

Assignment 2

- Essay
- Review one paper
- Topics:
 - Information Visualization
 - Gestural Interfaces
 - AR/VR
- Length: 3 pages
- Due: **2359 Friday 17 May**

https://ecs.wgtn.ac.nz/Courses/SWEN422_2024T1/Assignment2

Summary - How Do We Visualize?

- **Know the Data**
 - Number of attributes
 - Date types: ordinal vs ordered (ordinal or quantitative)
 - Trustworthiness: bad fields, inaccuracies, missing values
- **Know your purpose (& audience)**
 - What do you/they want to see?
 - What might you/they want to focus on?
- **Decide how to use visual variables & encode data**
 - Requires awareness of:
 - Human perceptual system
 - Display capacity
 - Characteristics of data (size, type)
 - Task

Dr. Craig Anslow

School of Engineering and Computer Science

Victoria University of Wellington

craig.anslow@ecs.vuw.ac.nz