Lecture 01
Introduction to Information Visualisation

Uta Hinrichs

CS5044 – Information Visualisation

CS5044 – aims & objectives I

- **Principles & methods** of designing effective visualisations

- **Visualisation techniques** to visually represent certain data types

- **Interaction techniques** to allow for the exploration of data through visualization

- **Skills to critically assess** information visualisations

- **Application areas** of information visualisation

CS5044 – aims & objectives II

- Getting you excited about visualisation

- Getting you to think critically about
  - Data & data collection
  - How visualisation can be used to communicate data
  - How a visualisation is interpreted and what insights it can promote (or not)

- Training your visual thinking and creativity
  - Telling stories using visualisation
  - Intentional use of visualisation tools
  - Creating novel visualisations with and/or without the computer
  - Experiment: think beyond line & bar charts

CS5044 – learning outcomes

- Design effective visualisations of given data sets
  - Data types
  - Target audience
  - General purpose

- Communicate and justify your design choices and process

- Critically assess the design of a given data visualisation
  - Identify effective visualisations

- Make use of existing visualisation tools & libraries

→ Good for industry

→ Good for research

→ Good in general as a communication and problem solving skill

target audience

- Students interested in learning how to design and implement expressive, effective, and interactive visual representations of data
  - Based on scientific theory
  - Based on good design practice

- Programming in JavaScript will be required!

general philosophy of this module

- Theory + hands-on work

- Manual sketching + programming
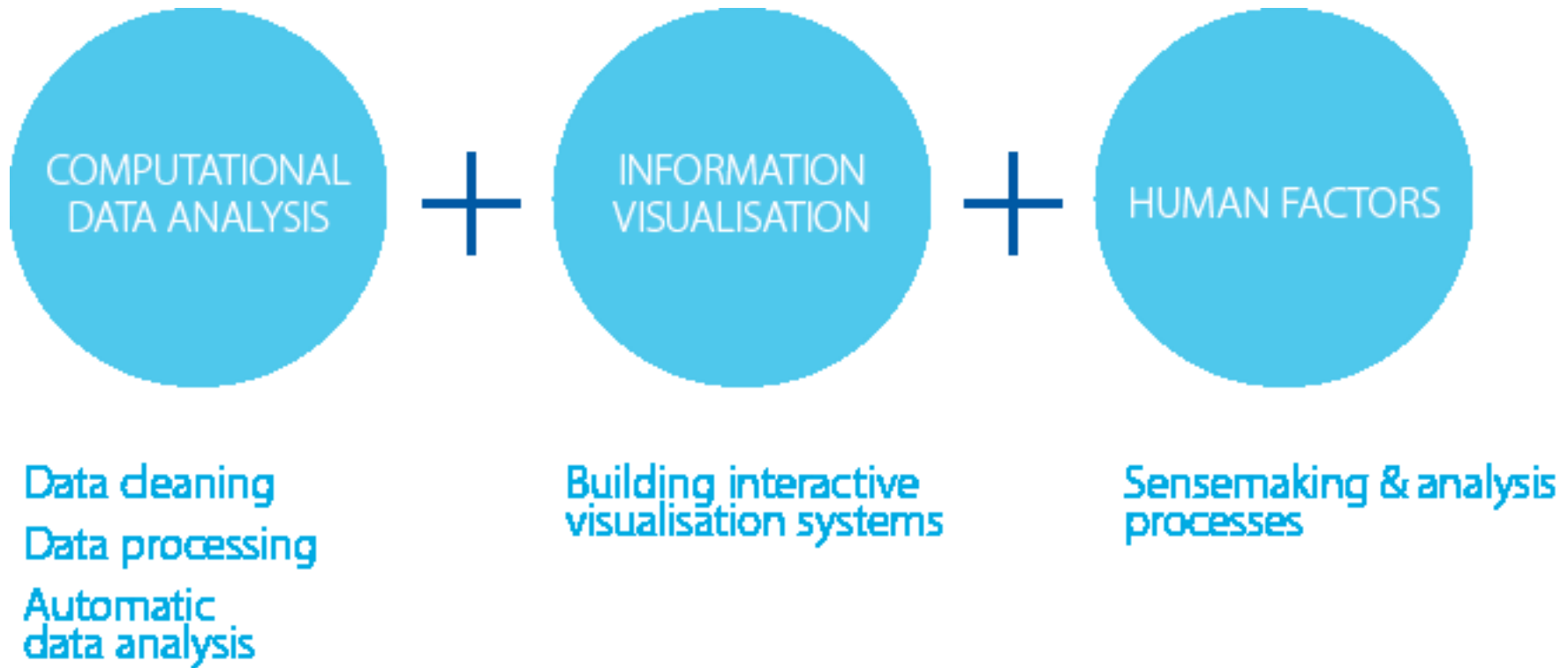
- Research-oriented

- Participatory

experimental approach

- Experiment with data and different data sources
- Experiment with different visualisation techniques
    - → Think beyond standard techniques
- Experiment with tools
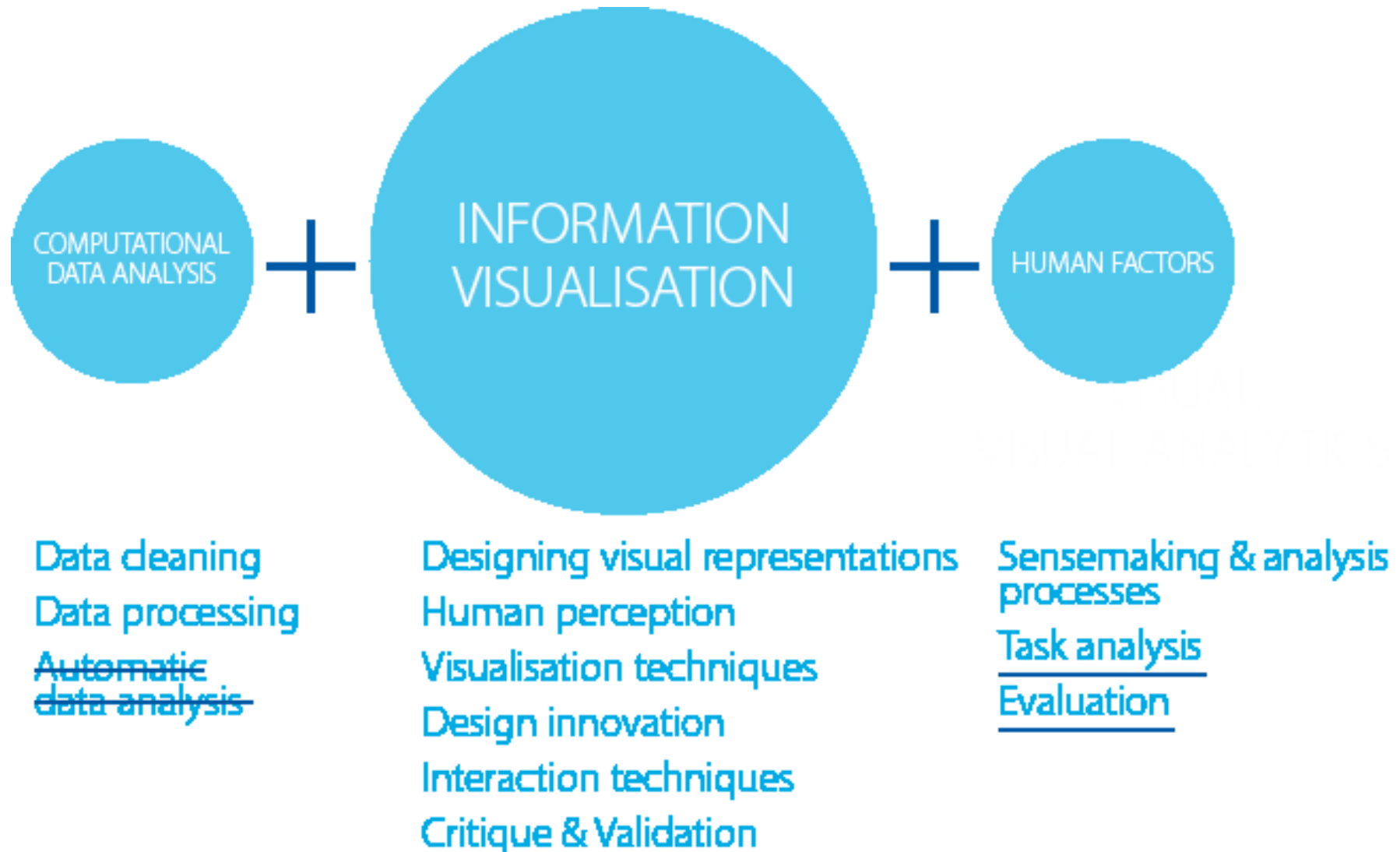   - Paper
    - Tableau Desktop
    - D3.js

what's the difference to Visual Analytics?

- CS5044 – Information visualisation
- CS4144 – Visual Analytics

what's the difference to Visual Analytics?

COMPUTATIONAL DATA ANALYSIS

+

INFORMATION VISUALISATION

+

HUMAN FACTORS

Data cleaning
Data processing
Automatic data analysis

Building interactive visualisation systems

Sensemaking & analysis processes

what's the difference to Visual Analytics?

COMPUTATIONAL DATA ANALYSIS + INFORMATION VISUALISATION + HUMAN FACTORS

Computational Data Analysis
- Data cleaning
- Data processing
- ~~Automatic data analysis~~

Information Visualisation
- Designing visual representations
- Human perception
- Visualisation techniques
- Design innovation
- Interaction techniques
- Critique & Validation

Human Factors
- Sensemaking & analysis processes
- Task analysis
- Evaluation

general information

CS5044

module structure

- Lectures
  - Theory & principles of information visualisation
  - Interactive quizzes
  - Group discussions

- Exercises & Tutorials
  - Design & implementation of visualisations
  - Data sketching
  - Tableau Desktop
  - D3.js

module structure

- 15 credits
  - Work expectation: 10h/week (on average, depending on your skills)
- Lecture + Exercises: 2.5h
- Time distribution
  - **Lecture:** Thursdays 10:00 am – 11:00 am; Jack Cole Building 1.33b
  - **Tutorial & Exercises:** Mondays 9:00 am – 10:30 am; John Honey Teaching Lab

Mondays

9:00

Tutorial & Exercises

10:30

11:00

Thursdays

10:00

Lecture

11:00

12:00
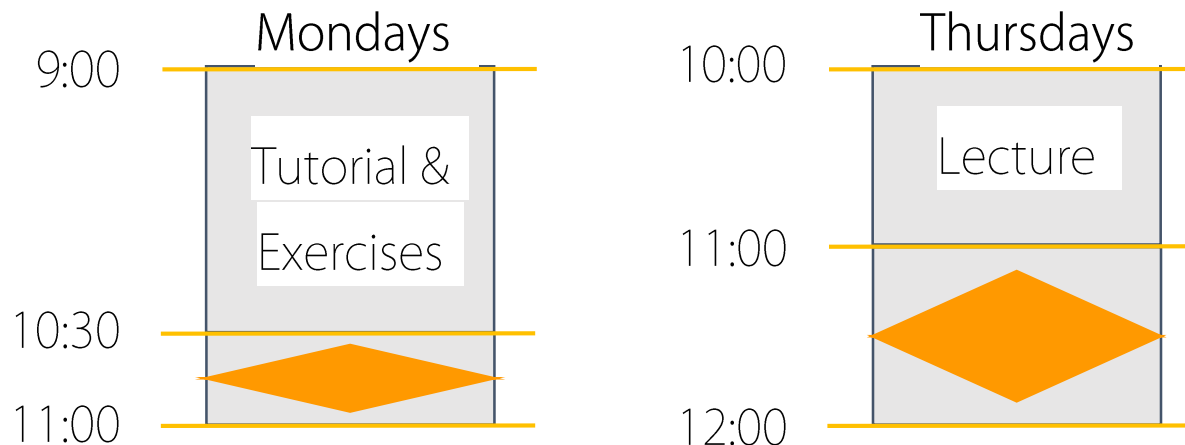
lecture etiquette

- Attending both lectures & exercise classes is more than strongly advised

- Questions are very welcome any time

- Suggestions are welcome: "more of…" "less of…"

- Active participation is expected!

reading material

- Main reading
  - Tamara Munzner: Visualisation Analysis and Design
  - → Available at the library as e-book

- If you can afford it, get your own copy
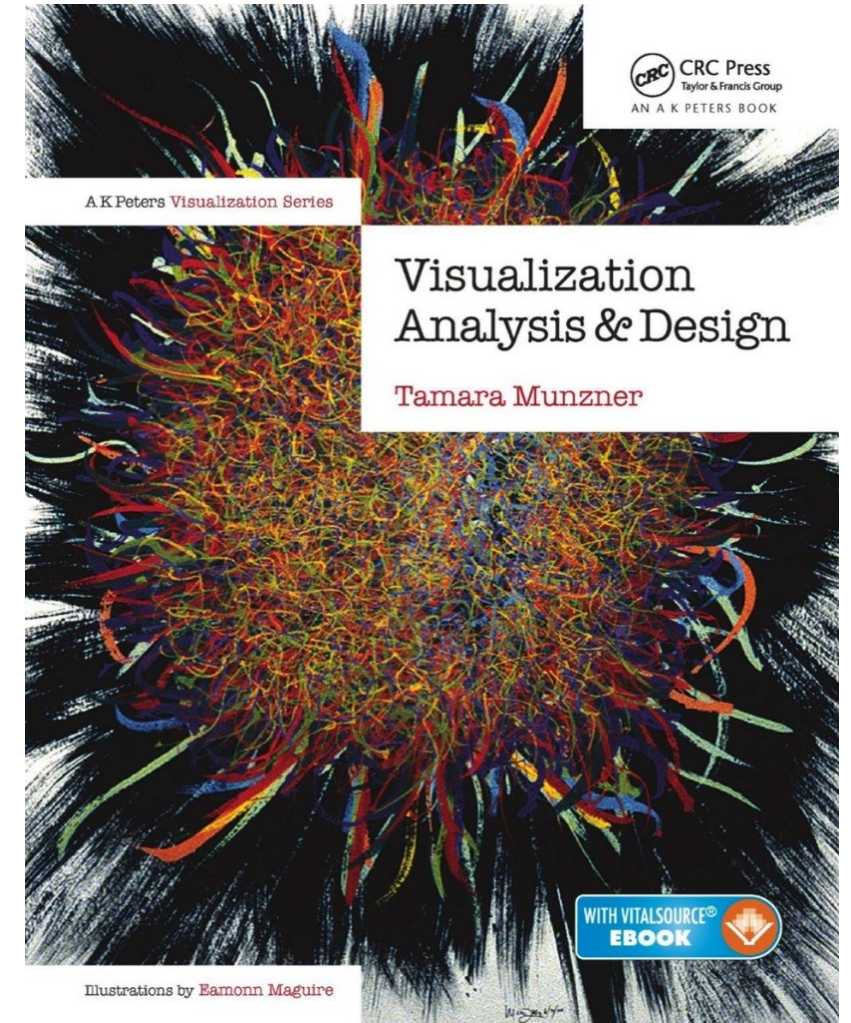
reading material

- Additional reading (optional)
  - Isabel Meirelles: Design for Information
  - → Should be available at the library soon
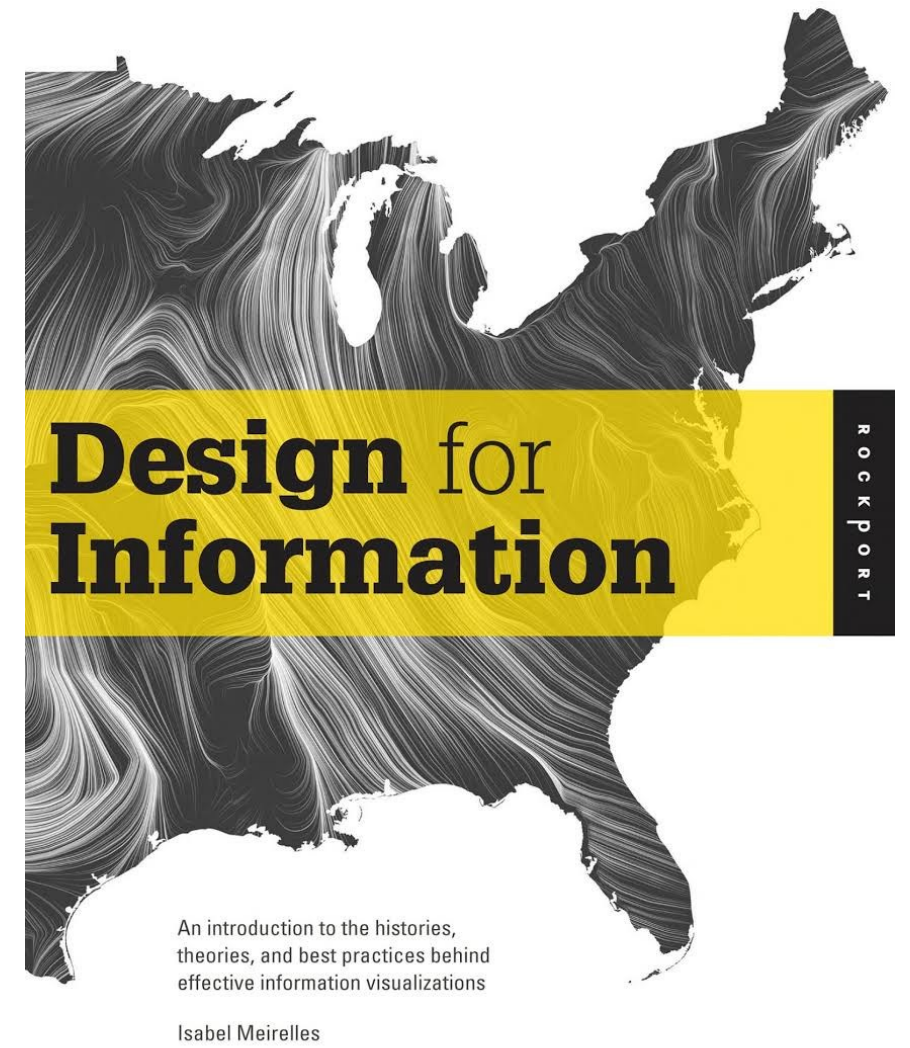
reading material

- More readings available on ResourceList
  - http://resourcelists.st-andrews.ac.uk
  - → Look for CS5044

- There will be weekly readings made available before each lecture
  - Book chapters
  - Research papers
  - Videos

- Do your readings before class!

resources

- Resources regarding coding, including tutorials
  - Tableau Desktop & D3.js
  - http://resourcelists.st-andrews.ac.uk
    - → CS5044

- StudRes
  - All lecture material
  - All materials for in-class exercises

assessment
60% coursework + 40% exam

Module	100%	
Exam	40%	
P1 VIS CRITIQUE	10%	Due Week 3
P2 VIS IN TABLEAU	20%	Due Week 6
P3 VIS IN D3.JS	30%	Due Week 11; Group work
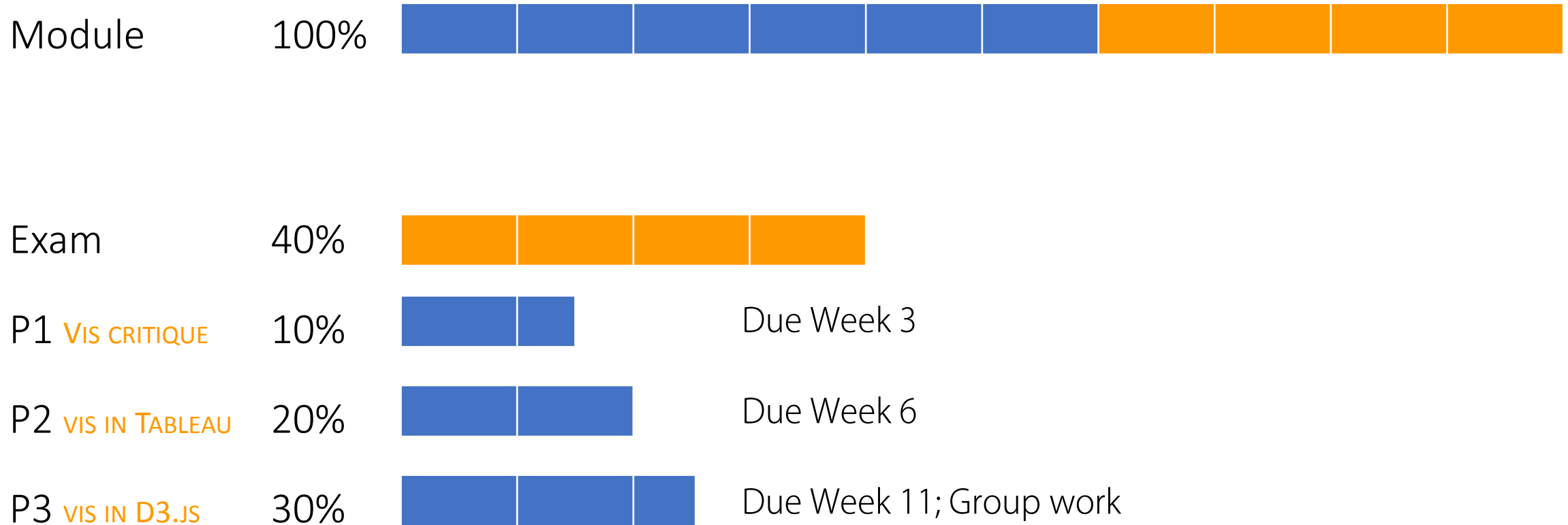
ethics and academic misconduct

https://www.st-andrews.ac.uk/students/rules/academicpractice/

- What you hand in should reflect on your own work

- It is NOT ok to submit the same work to two different modules without consent and knowledge from the module coordinators

- Source of code and inspiration need to be clearly reported, for example:
  - Code for the radial tree layout is based on Michael Bostock's online example
  - Inspiration for the visual representation of countries is based on Mia Newman's cartogram on international connections on Facebook

- How much you "borrow" influences how your work is evaluated
  - → Go for unique content + good referencing

- If in doubt, ask!! Don't risk a misconduct hearing.

key policy points

https://info.cs.st-andrews.ac.uk/student-handbook/key-points.html

- You are assumed to be familiar with the student handbook

- Read the Good Academic Practice policy.

- Check that coursework submitted to MMS has been received successfully, and that it is the right piece of work.

- Coursework submitted after the deadline is subject to automatic penalty.

- Any special circumstances must be documented immediately through the self-certification system.

- You must be available in St Andrews for the entire exam period.

contact

- Uta Hinrichs | JC 1.09 | uh3@st-andrews.ac.uk

- Best practice: get in touch on Mondays or Thursdays after the lecture/exercise class

- When to contact me
  – Questions regarding the module's content procedure or practical aspects that you cannot resolve

- When NOT to contact me
  – To ask questions about the lecture or tutorials, if you have not tried to work through the material yourself (I can only try to help with specific questions)
  – To request extensions (contact dopgt-cs@st-andrews.ac.uk)

- I may make questions and my answers available to all students if it makes sense

quiz

hands-on exercise

given a data set that consists of
two numbers: 75 and 37

try to find as many possible
ways to visualize this data set

discuss your visualisations with the
person sitting next to you
– How to read the visualisations?
– How do they differ from each other

among the two of you, pick the best
visualisation!

[3 minutes]

45 ways to communicate two quantities

1. writing, number notation

75, 37

a

b

c

The familiar indo-arabic numerals notation in *a*, and the babylonian numbers notation in *b*.

2. squares

3. repeated icon
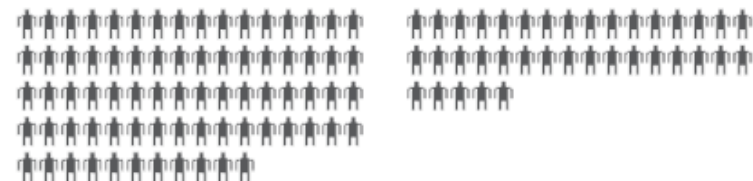
75 and 35

4. hundreds, tens, units, decimals... represented by squares

5. bars

a

b

c

6. line graph

75
37
0
a

75
37
b

75
37
0
c
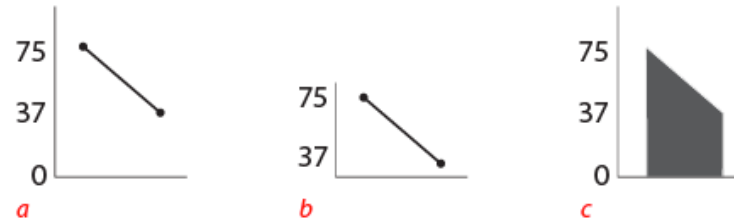
7. percentages bars

a

b

75 and 35

8. spliced bar

75 | 37

a

b

9. proportion

75

37
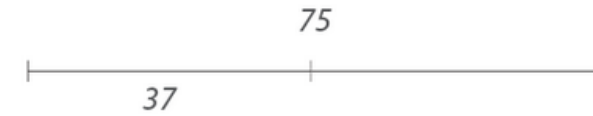
For the case one value is a part of the other.

10. interval

37 | 75

a

b
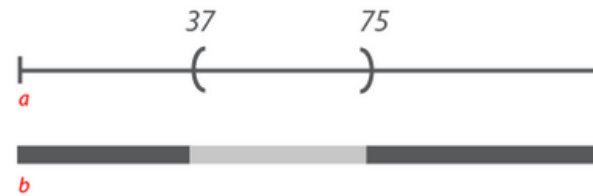
11. squares merged
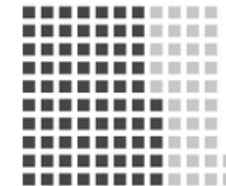
12. percentages in squares
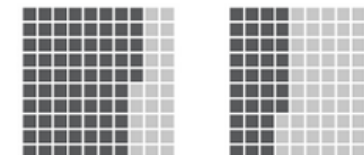
75 and 35

13. pie charts

a   b   c

14. donut chart

It differs from the pie chart because it can be nested into other donut charts in order to create a more complex visualization.

15. circle areas

16. semi-circle areas

Can be placed on a bi-dimensional space associating the two values to a single point.

https://visual.ly/blog/45-ways-to-communicate-two-quantities/

30

75 and 35

20. shape divided

21. square surfaces

22. shape surfaces
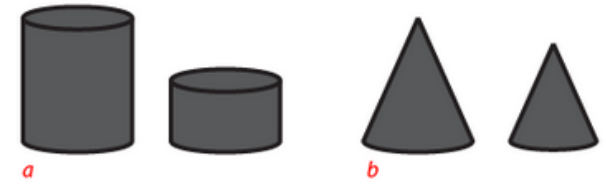

23. different shape surfaces
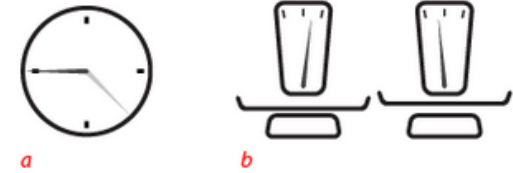

24. icon surfaces

75 and 35

26. volumes

In *a* the volume is proportional to the object's height, not the case for *b*, which is very difficult to read.

27. special metaphors

...for special units: a gauge for speeds, balances for weights
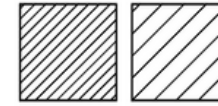
28. gray tones

29. color scale

75 and 35

38. dashed

similar to *28* gray tones and *36* density.

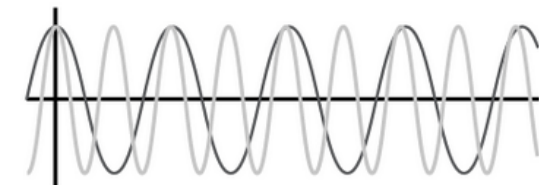39. nodes and connections in a network

As we don't have more than the quantitative information, this visualization is only a way to have a first glance of the external structure of the network (the relations are chosen randomly).

40. parameters of a mathematical function

$y = cos(75x)cos(37x)$

41. harmonic frequencies

reflections on the exercise

- There are many, many ways to visualize very small datasets
  - If we add more numbers, the amount of options even increases
  - → the design space for visualisation is huge, even just in 2D

- There is no "best" way to visualize a set of data
  - The semantics of the data matter (what do the numbers stand for?)
  - Size of the data set matters
  - The context in which the data is to be interpreted
  - The audience that will read and interpret the data

- In this module, you will learn about
  - The design space of visualisation: what are the options?
  - How to navigate this space and make sensible choices
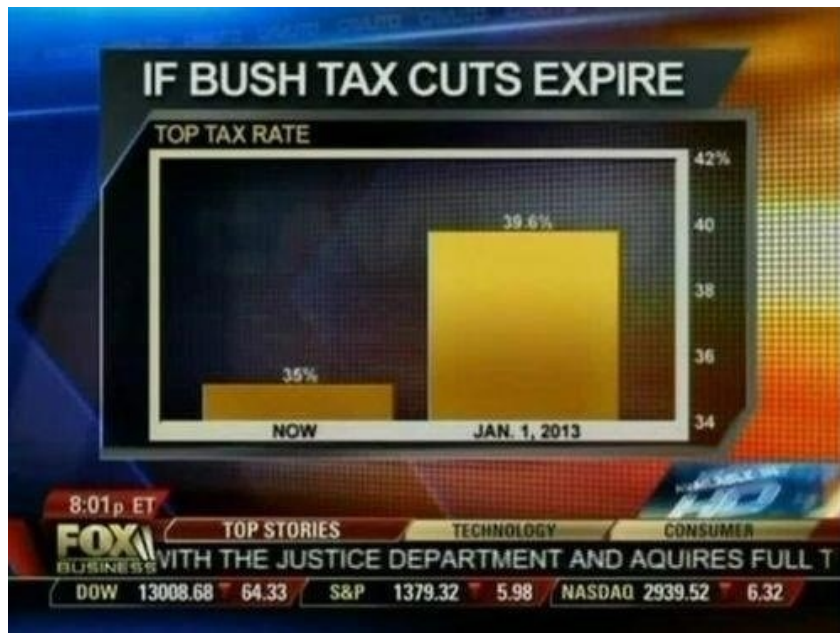
how to avoid mistakes

Image source:
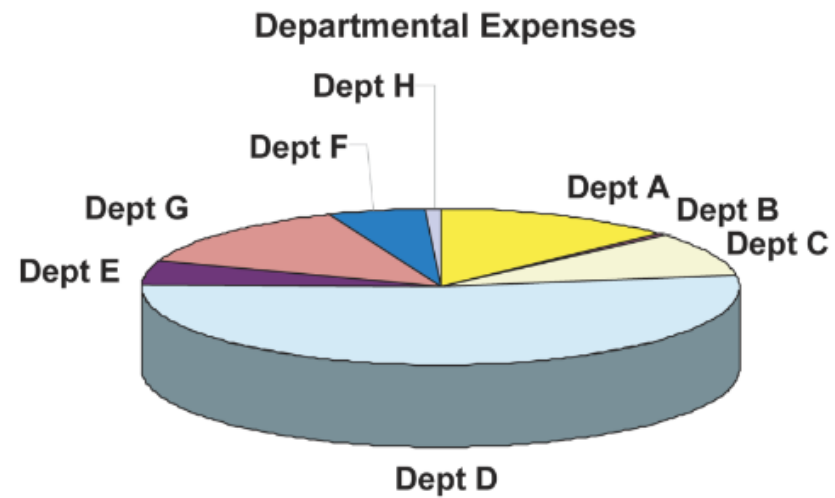https://blog.heapanalytics.com/wp-content/uploads/2014/04/misleading1_fox.jpg

Image source:
Show me the numbers, Steven Few

what is information visualisation?

what is information visualisation?

Data and information visualisation is concerned with showing quantitative and qualitative information, so that a viewer can see patterns, trends or anomalies, constancy or variation, in ways that other forms—text and tables—do not allow.

Michael Friendly, 2008

M. Friendly. The Golden Age of Statistical Graphics. Statistical Science, 2008, Vol.23, No. 4, 502-535.

what is information visualisation?

Data and information visualisation is concerned with showing quantitative and qualitative information, so that a viewer can see patterns, trends or anomalies, constancy or variation, in ways that other forms—text and tables—do not allow.

Michael Friendly, 2008

M. Friendly. The Golden Age of Statistical Graphics. Statistical Science, 2009, Vol.23, No. 4, 502-535.

what is information visualisation?

Data and information visualisation is concerned with showing quantitative and qualitative information, so that a viewer can see patterns, trends or anomalies, constancy or variation, in ways that other forms—text and tables—do not allow.

Michael Friendly, 2008

M. Friendly. The Golden Age of Statistical Graphics. Statistical Science, 2009, Vol.23, No. 4, 502-535.

what is information visualisation?

Data and information visualisation is concerned with showing quantitative and qualitative information, so that a viewer can see patterns, trends or anomalies, constancy or variation, in ways that other forms—text and tables—do not allow.

Michael Friendly, 2008

M. Friendly. The Golden Age of Statistical Graphics. Statistical Science, 2009, Vol.23, No. 4, 502-535.

showing trends, patterns, and anomalies

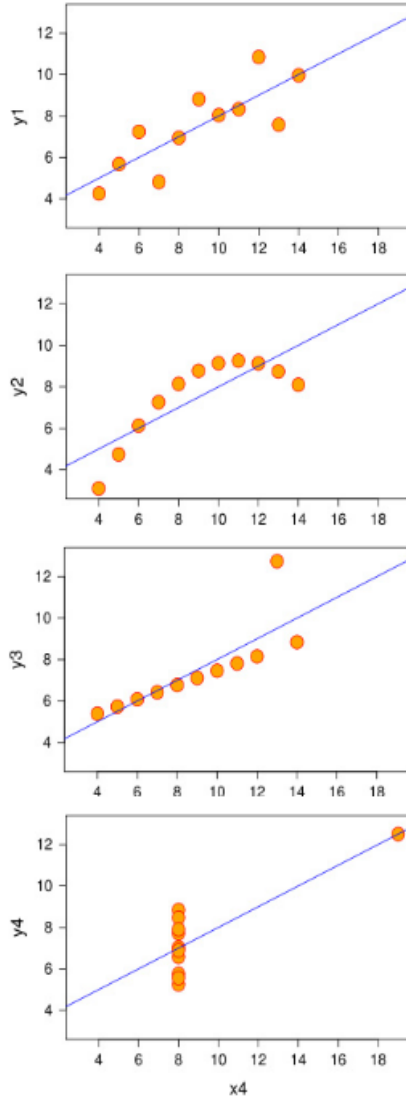| I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	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

Anscobe's Quartet

showing trends, patterns, and anomalies

- Mean
  - x = 9.0
  - y = 7.5
- Variance
  - x = 10
  - y = 3.75
- Correlation
  - 0.816

I		II		III		IV	
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

Anscobe's Quartet

showing trends, patterns, and anomalies

| I		II		III		IV	
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

Anscobe's Quartet

historic examples
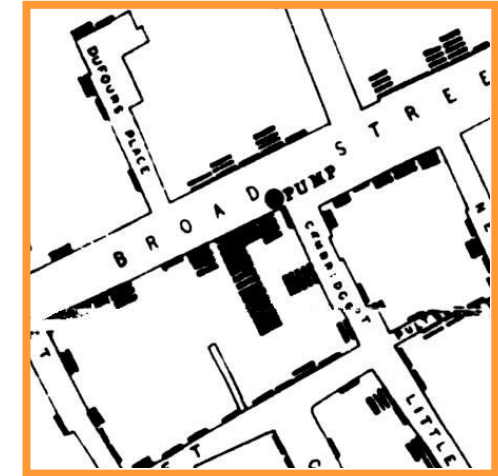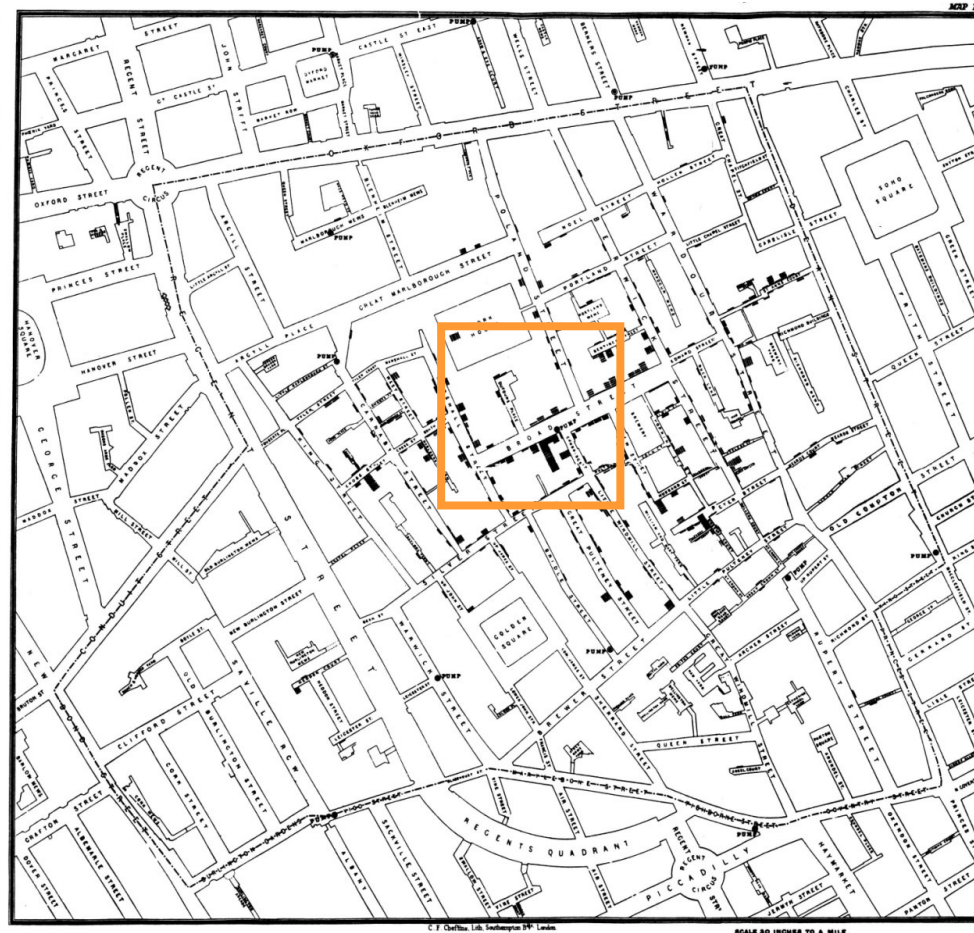
John Snow's map of cholera cases in the London epedemic of 1854

http://en.wikipedia.org/wiki/John_Snow_%28physician%29
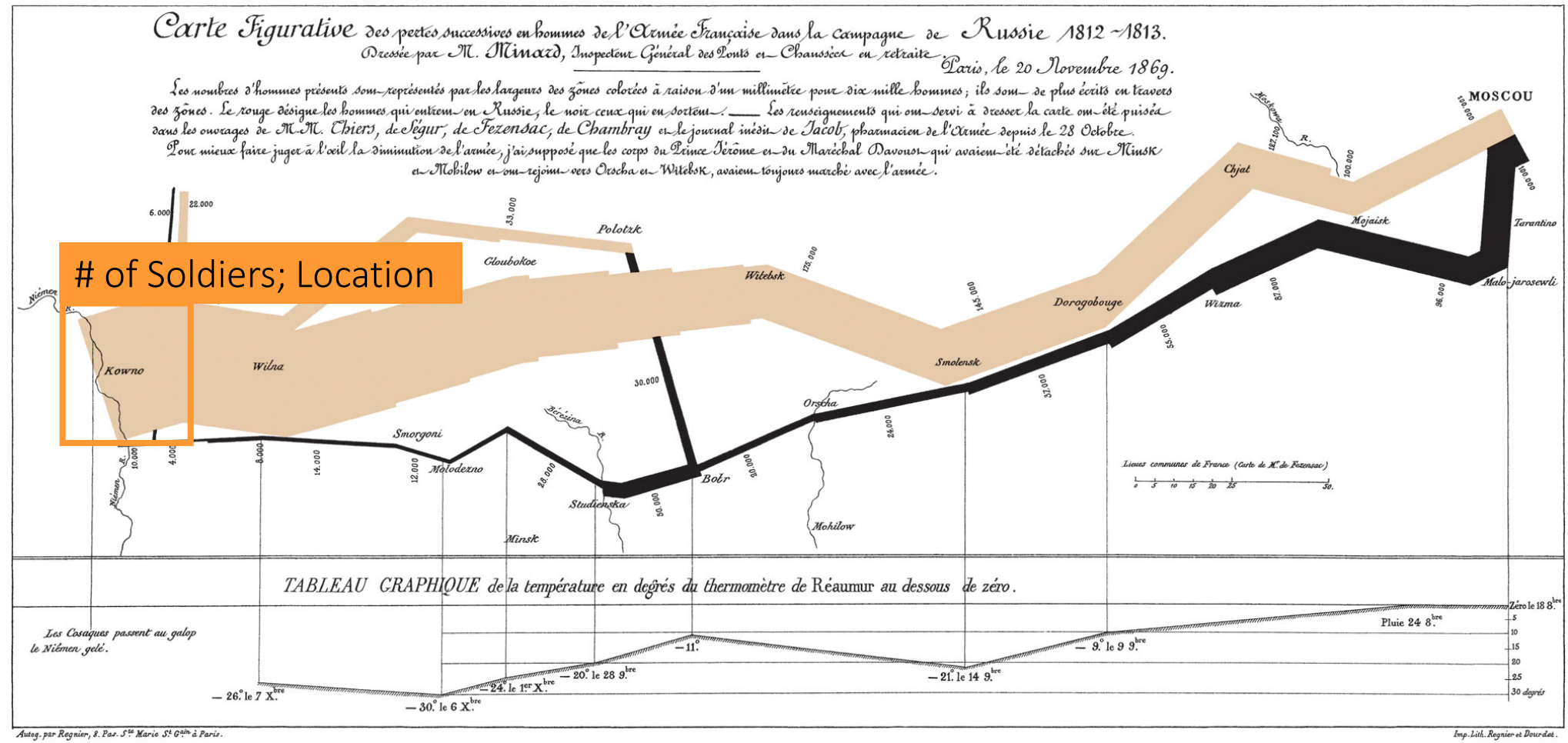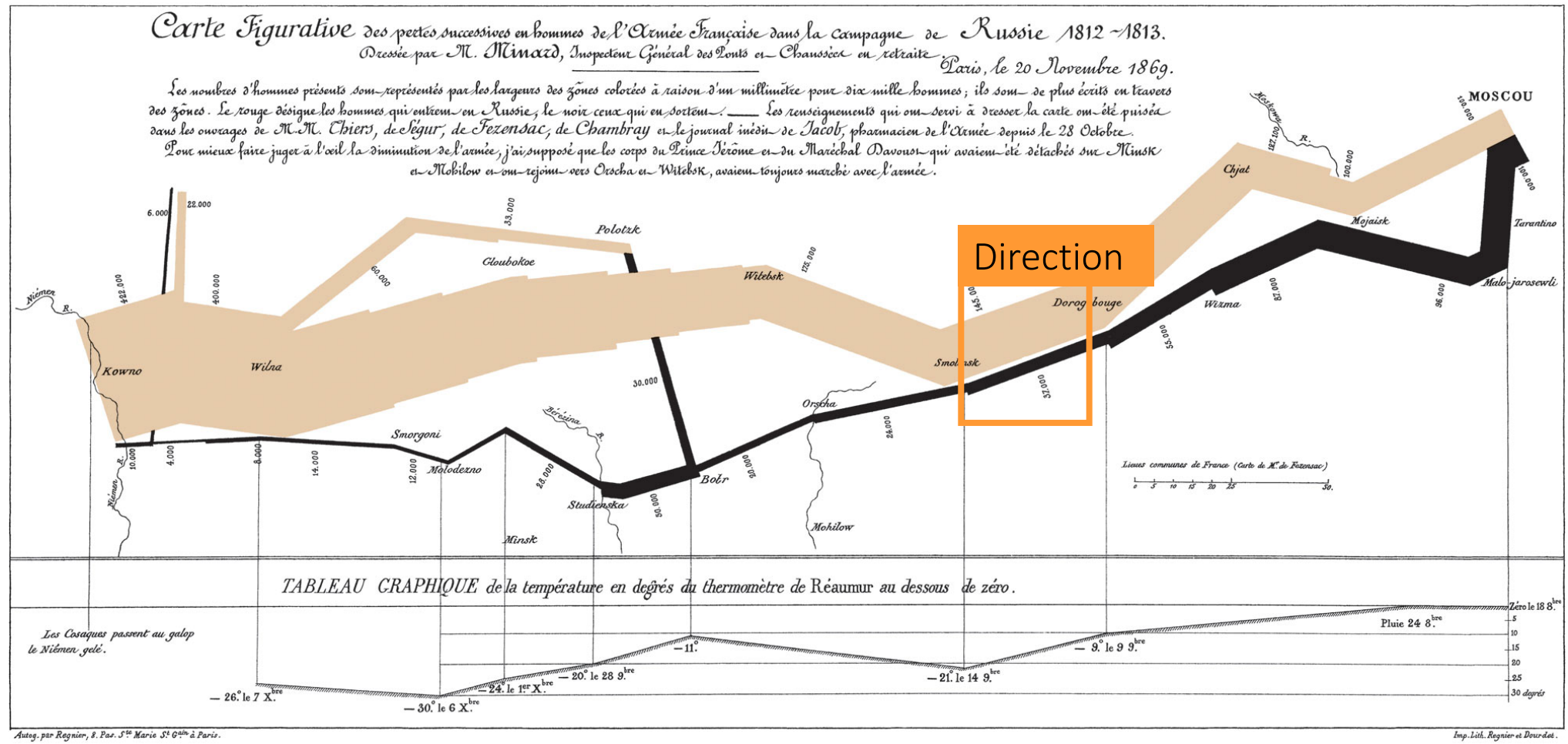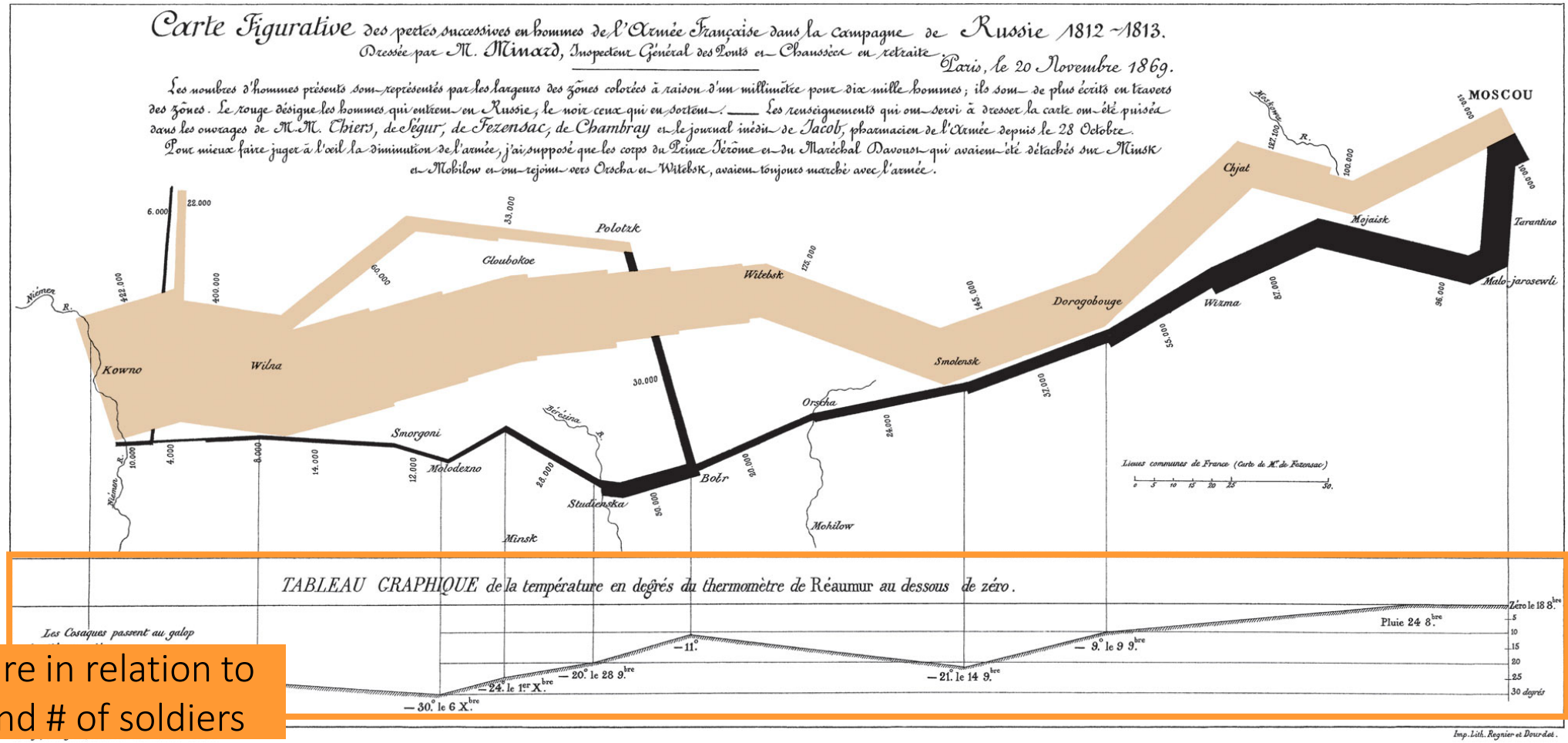
historic examples
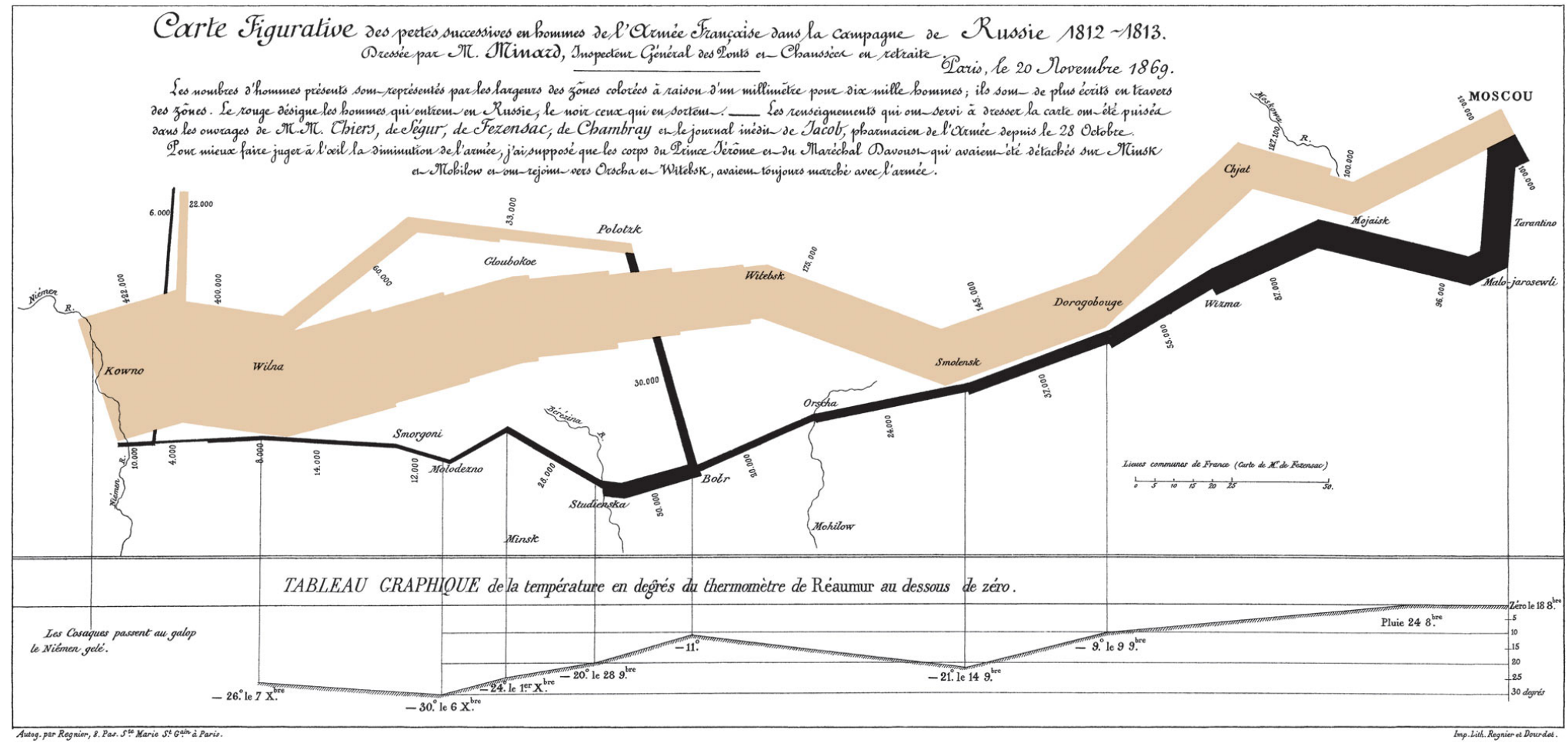
Charles Minard's map of Napoleon's disastrous Russian campaign of 1812

http://en.wikipedia.org/wiki/Charles_Joseph_Minard

historic examples

# of Soldiers; Location

Charles Minard's map of Napoleon's disastrous Russian campaign of 1812

http://en.wikipedia.org/wiki/Charles_Joseph_Minard

46

historic examples

Charles Minard's map of Napoleon's disastrous Russian campaign of 1812

http://en.wikipedia.org/wiki/Charles_Joseph_Minard

historic examples

Temperature in relation to location and # of soldiers

Charles Minard's map of Napoleon's disastrous Russian campaign of 1812

http://en.wikipedia.org/wiki/Charles_Joseph_Minard

48

what is information visualisation?

Data visualisation is the graphical display of abstract information for two purposes: sense-making (also called data analysis) and communication.
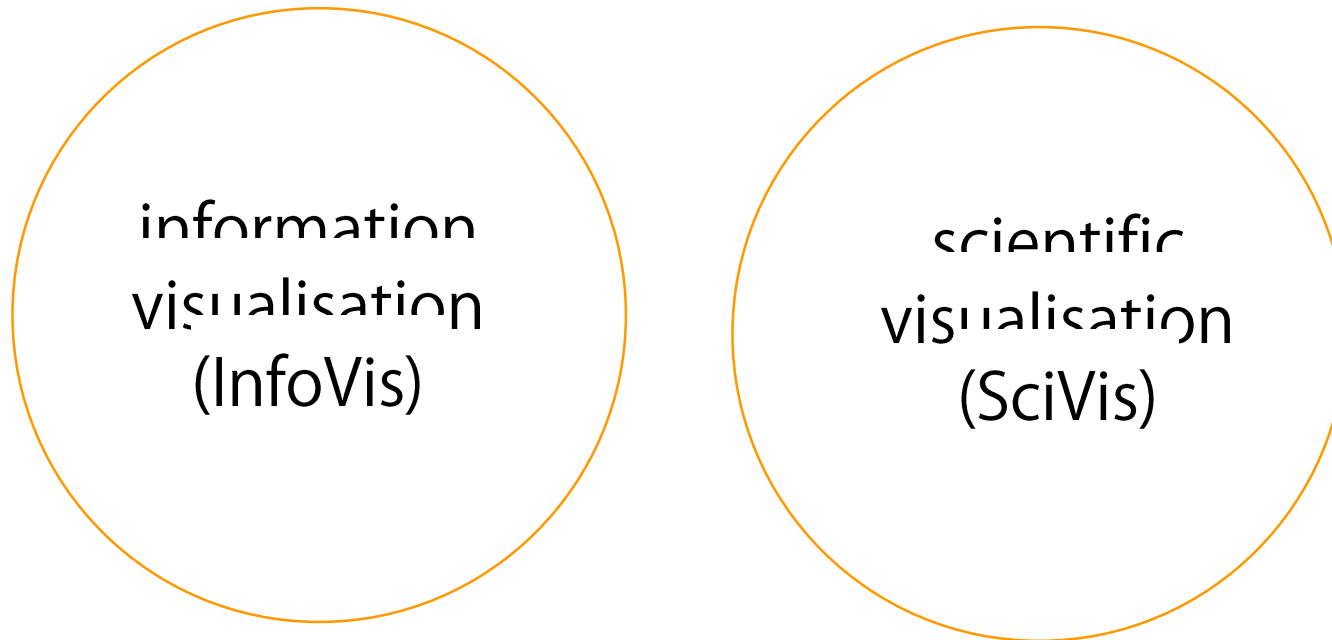
Stephen Few

Data visualisation for Human Perception. The Encyclopedia of Human-Computer Interaction, 2nd Edition

https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualisation-for-human-perception

what is information visualisation?

Data visualisation is the graphical display of abstract information for two purposes: sense-making (also called data analysis) and communication.

Stephen Few

what is information visualisation?

- In visualization, we distinguish between

information visualisation (InfoVis)

scientific visualisation (SciVis)

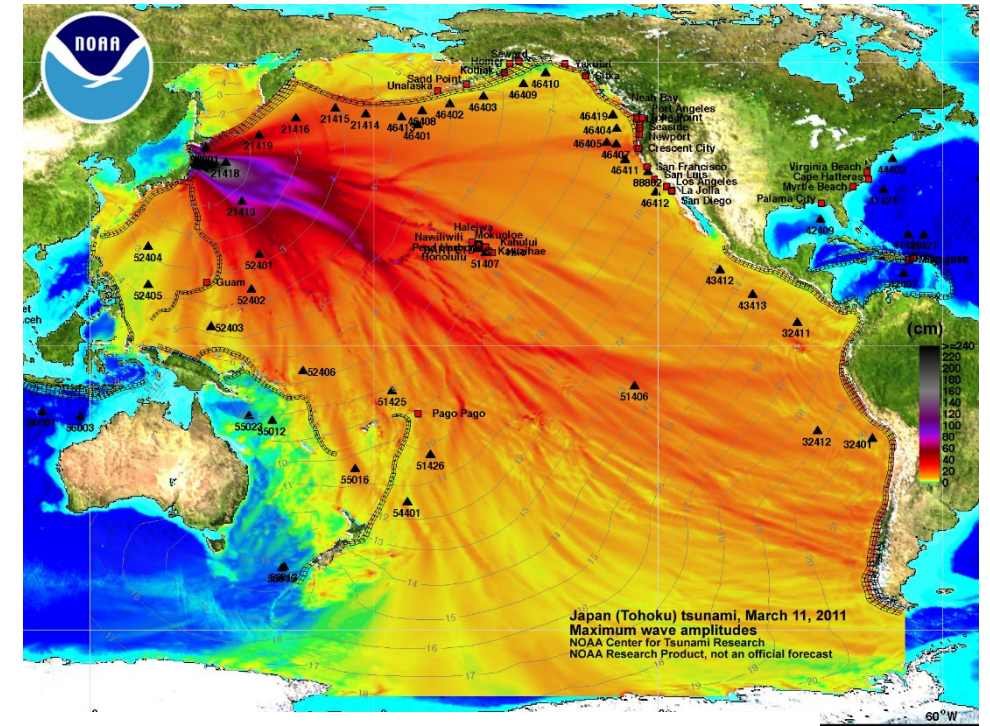Card et al. Readings in information visualisation: using vision to think; p. 7

Spatial information/data

- Physically-based data with an inherent spatial mapping

- Spatial mapping is typically defined by nature

- Examples
  - Medical data
  - Geospatial data

→ Scientific visualisation

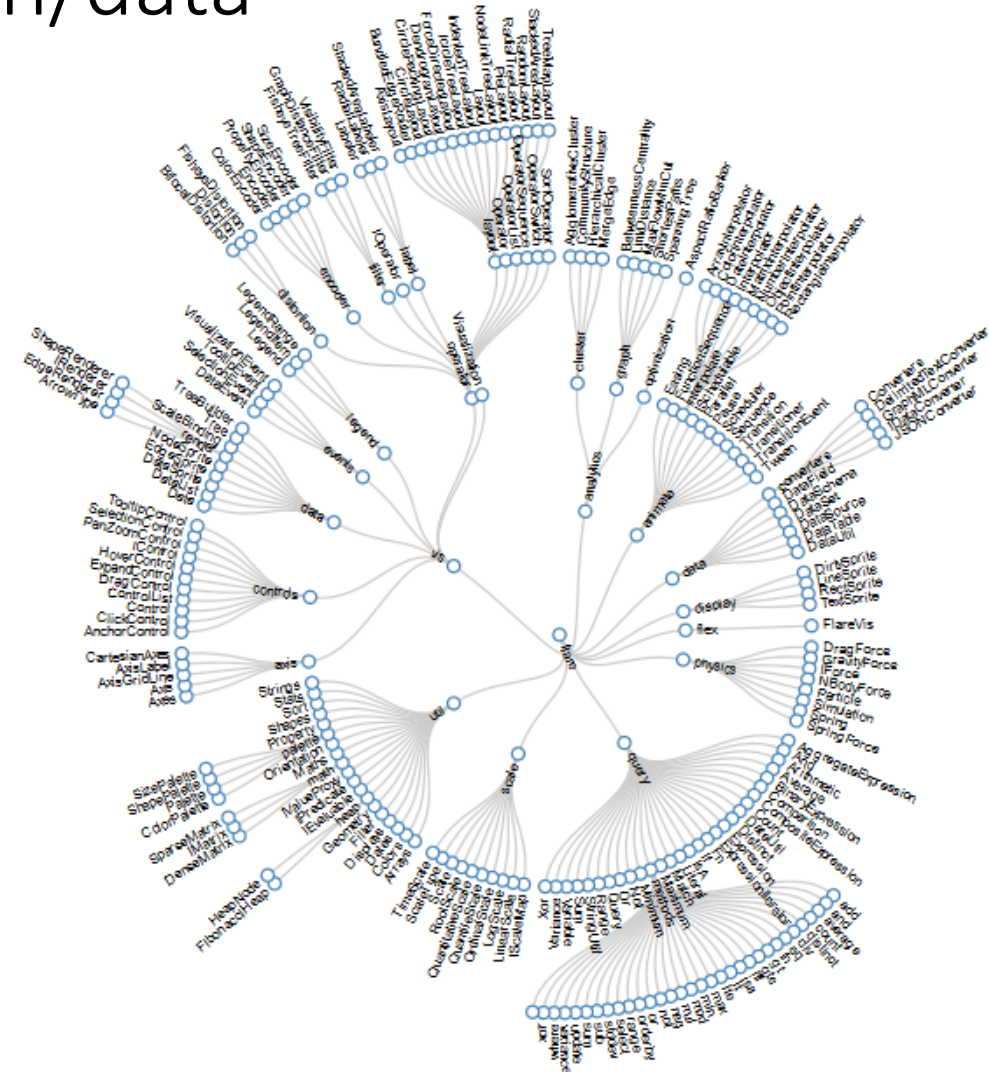http://en.wikipedia.org/wiki/Neuroimaging
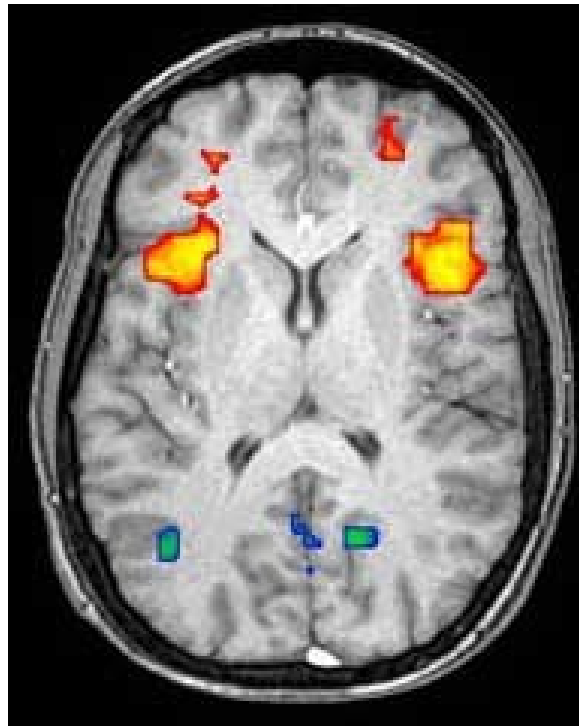
http://www.noaa.gov/features/03_protecting/japantsunami_oneyearlater.html

abstract information/data

- Abstract data
  - Does not have an inherent spatial mapping
  - We have to decide how to map data to visuals
  - There are many possible solutions

- Examples
  - Financial data
  - Biographical data
  - Text and documents
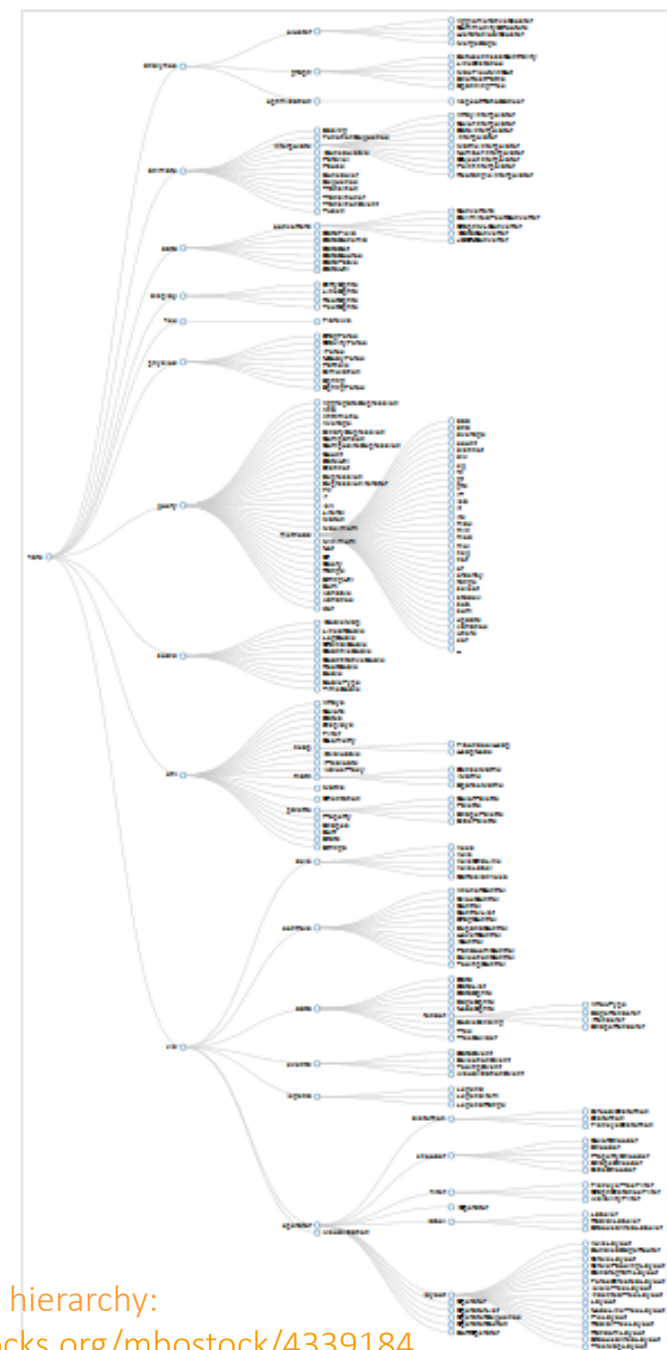 - Software structure
  - …

abstract vs. spatial information/data

- Abstract data
  - Does not have an inherent spatial mapping
  - We have to decide how to map data to visuals
  - There are many possible solutions

Flare class hierarchy:
http://bl.ocks.org/mbostock/4063550

54

abstract vs. spatial information/data

- Abstract data
  - Does not have an inherent spatial mapping
  - We have to decide how to map data to visuals
  - There are many possible solutions

Flare class hierarchy:
http://bl.ocks.org/mbostock/4339184
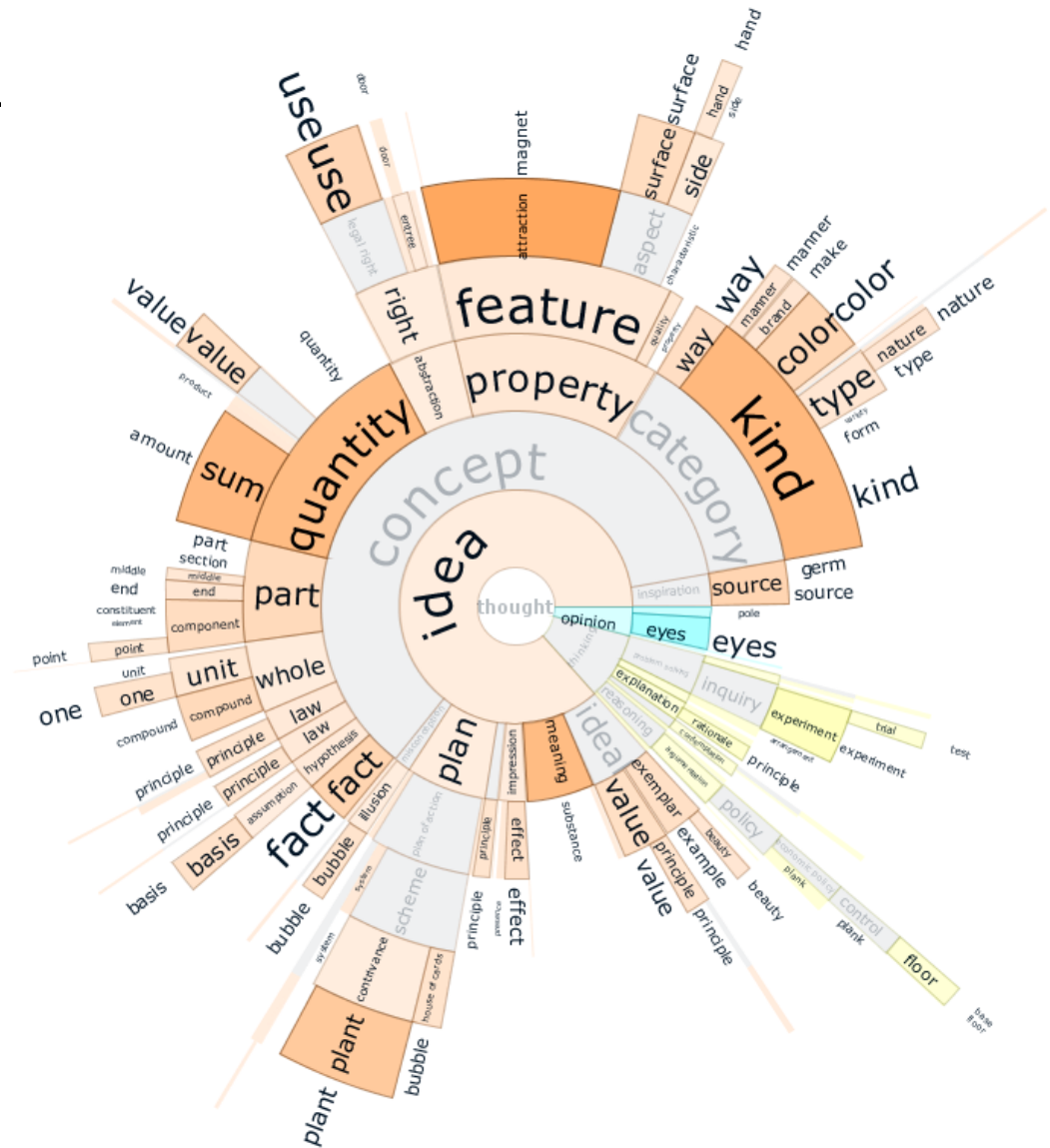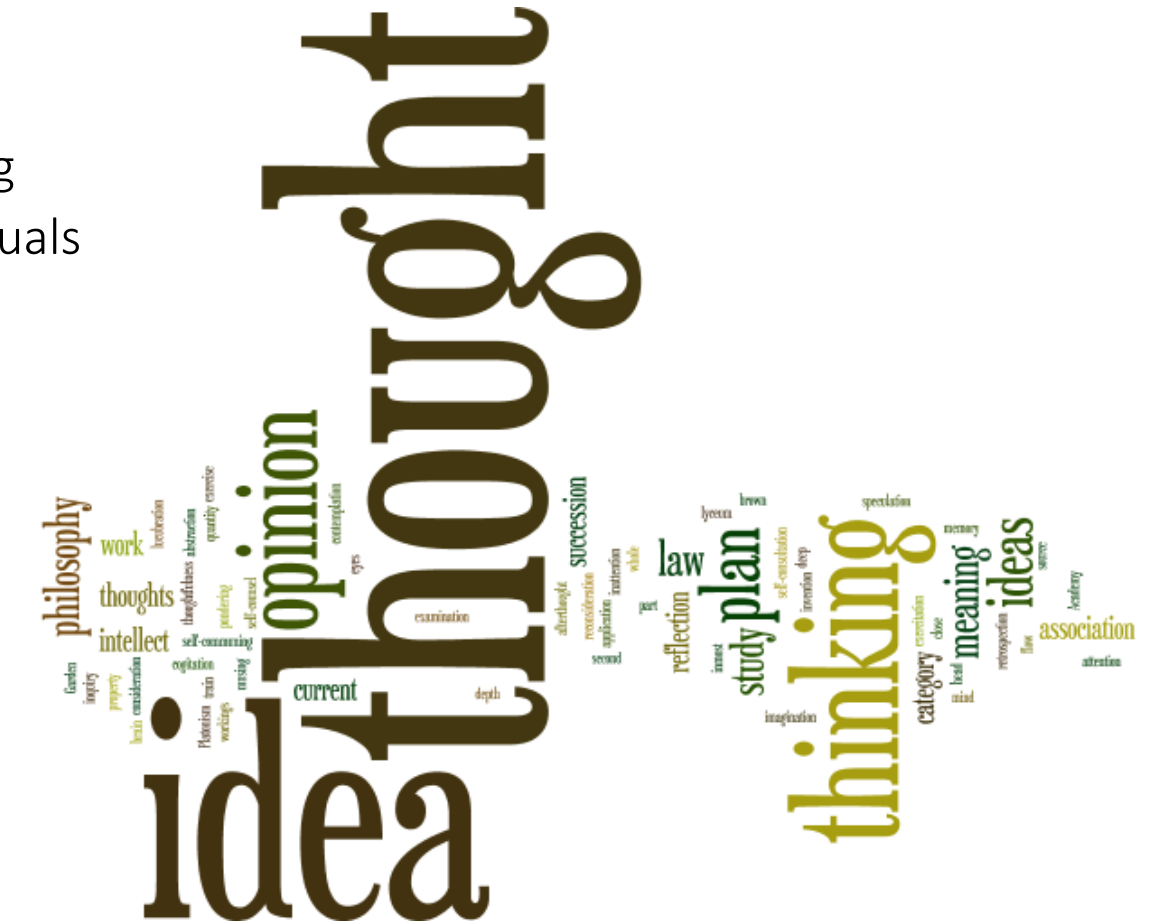
abstract vs. spatial informatior

- Abstract data
  - Does not have an inherent spatial mapping
  - We have to decide how to map data to visuals
  - There are many possible solutions

DocuBurst, Christopher Collins
http://vialab.science.uoit.ca/portfolio/docuburst-visualizing-document-content-using-language-structure
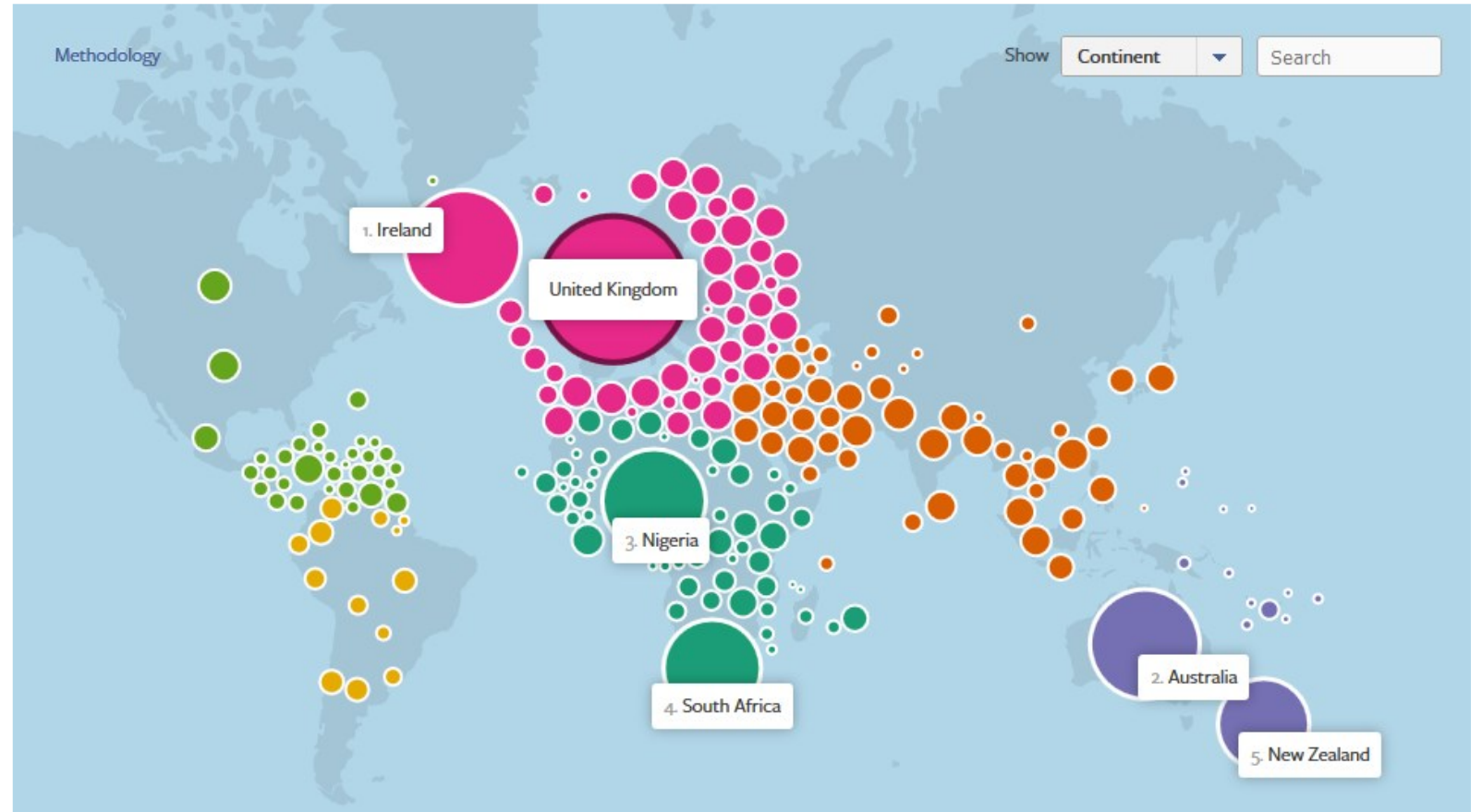
abstract vs. spatial information/data

- Abstract data
  - Does not have an inherent spatial mapping
  - We have to decide how to map data to visuals
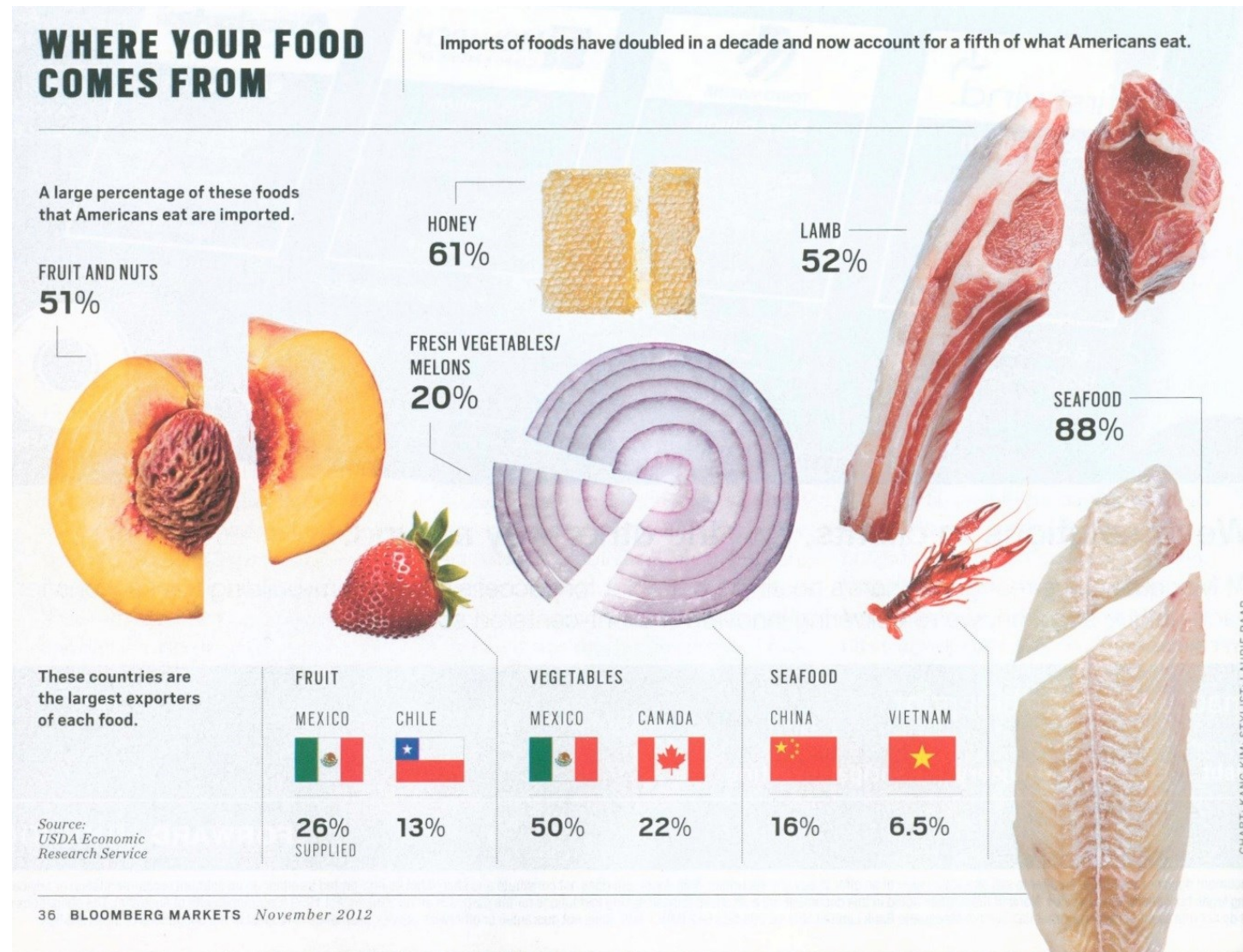 - There are many possible solutions

www.wordle.net/

hybrid cases exist!
abstract vs. spatial information

https://dabrownstein.com/category/facebook-friendship-map/

Is this a visualization?

WHERE YOUR FOOD COMES FROM

Imports of foods have doubled in a decade and now account for a fifth of what Americans eat.

A large percentage of these foods that Americans eat are imported.

FRUIT AND NUTS
51%

HONEY
61%

LAMB
52%

FRESH VEGETABLES/
MELONS
20%

SEAFOOD
88%

These countries are the largest exporters of each food.

	FRUIT		VEGETABLES		SEAFOOD	
	MEXICO	CHILE	MEXICO	CANADA	CHINA	VIETNAM
	26% SUPPLIED	13%	50%	22%	16%	6.5%

Source:
USDA Economic
Research Service

36 BLOOMBERG MARKETS November 2012

CHART: KANG KIM. STYLIST: LAURIE RAAB

Visualisation ≠ illustration

- Illustration
  - Using visuals to help communicate information, concepts or ideas
  - May be approximations, but not necessarily accurate representations of the data
  - Storytelling comes first

- Visualisation
  - Direct mapping between visual elements and data
  - If the data changes, the visuals also change
  - Accuracy comes first
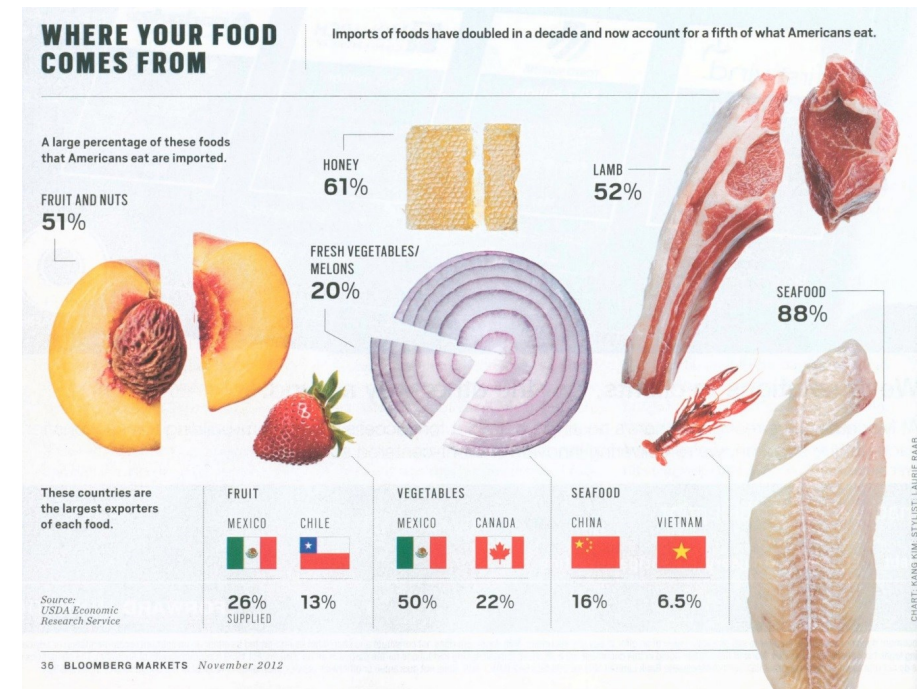
what is information visualisation?

Data visualisation is the graphical display of abstract information for two purposes: sense-making (also called data analysis) and communication.

Stephen Few

Data visualisation for Human Perception. The Encyclopedia of Human-Computer Interaction, 2nd Edition
https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualisation-for-human-perception

purpose of information visualisation

- Communicate (final results)
  - Summarize findings
  - Presenting insights
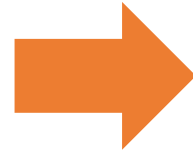
  → Reports & presentations

- Sense-making & active exploration (process)
  - Targeted analysis: questions are well-defined
  - Exploratory: questions are still evolving

  → Interactivity is important!

  → Research, business, personal life

what is information visualisation?

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

Card et al., 1999

Card et al. Readings in Information visualisation: Using Vision to Think; Chapter 1

what is information visualisation?

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

Card et al., 1999

Card et al. Readings in Information visualisation: Using Vision to Think; Chapter 1

what is information visualisation?

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

Card et al., 1999

Card et al. Readings in Information visualisation: Using Vision to Think; Chapter 1

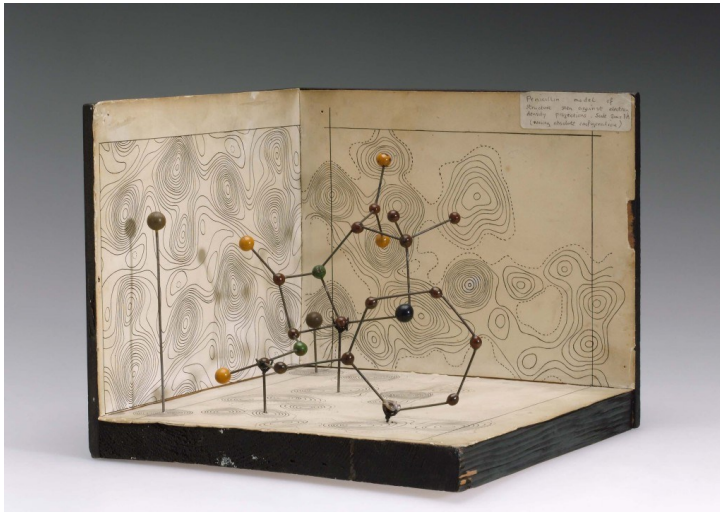# visual representation to amplify cognition

- visualisation leverages the human visual system to facilitate
  - Understanding
  - Discovery
  - Decision making
  - Communication
  - Memory
- Our brain is capable of quick and efficient background visual information processing
- Visual information processing happens at the preconscious level
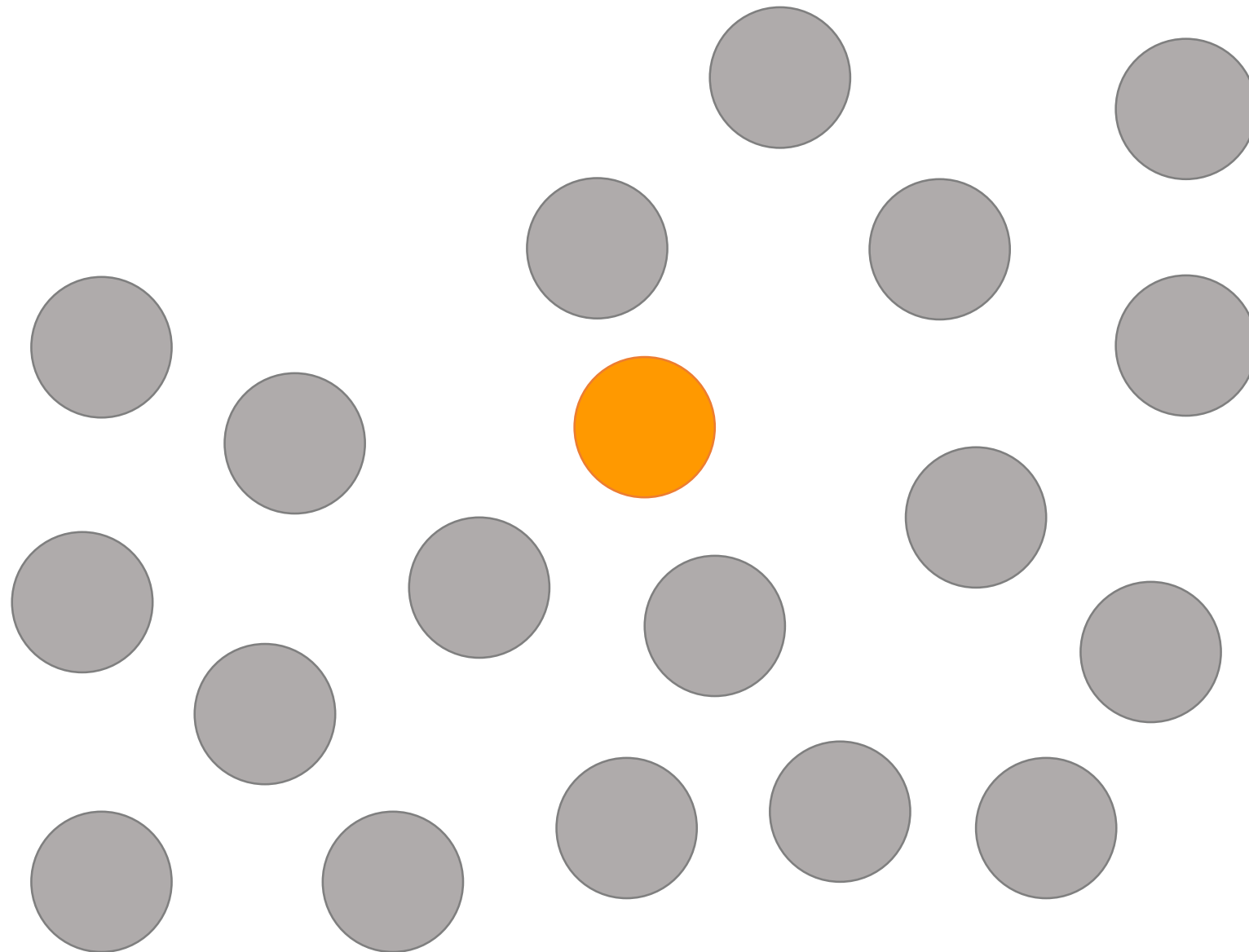
what about other senses?

- Sound is processed sequentially
  - → Sonification: representing data through sound

- Haptic/kinesthetic senses have a limited dynamic range
  - →Physicalization: physical artifacts whose geometry and material properties encode data

  Jansen et al., Opportunities and Challenges for Data Physicalization. Proc. of CHI'15
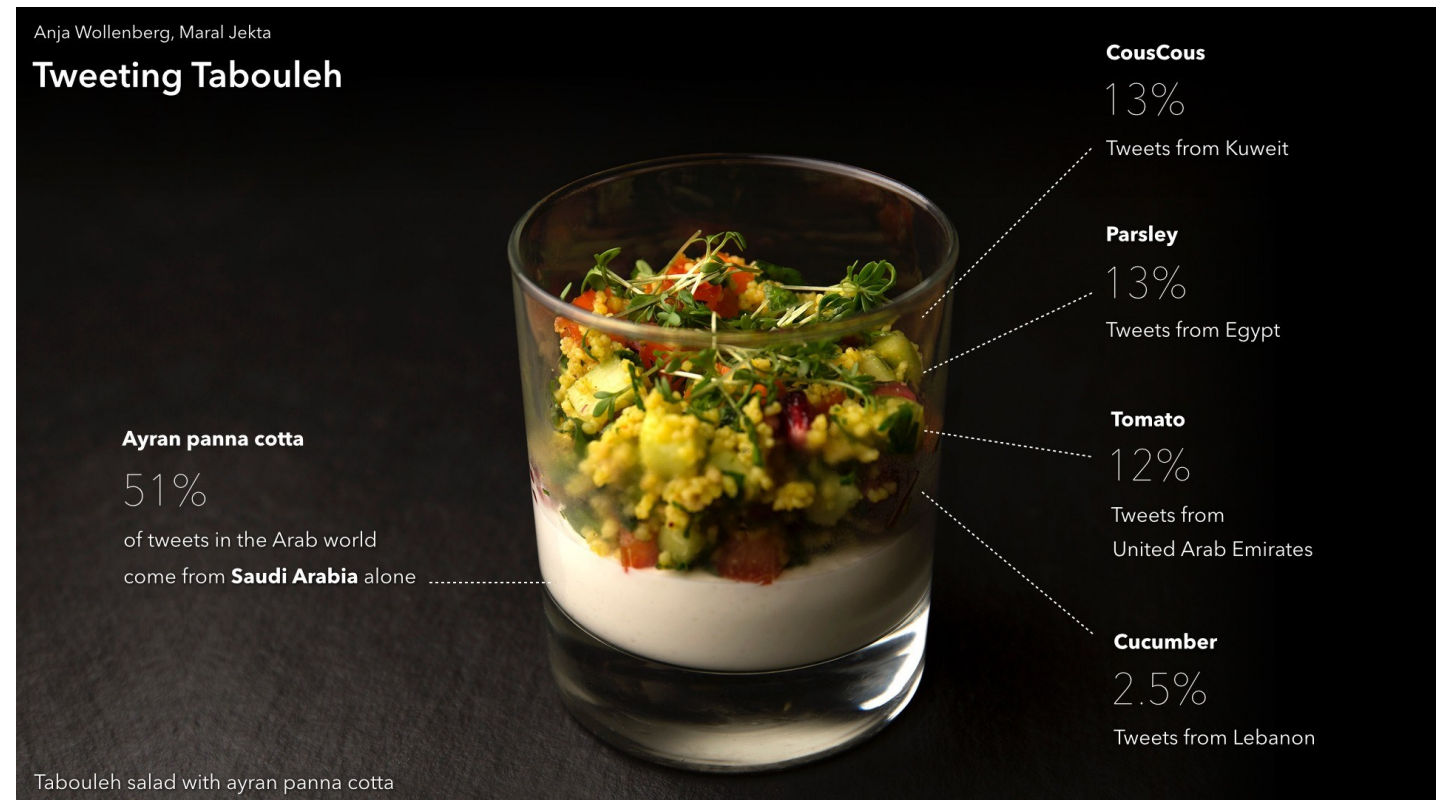 https://hal.inria.fr/hal-01120152/document

http://dataphys.org/list/electron-density-map-and-molecular-model-of-penicillin/

http://dataphys.org/list/walkable-age-pyramid/

data-cuisine-MiCT-food-slides-1024.003-e1423080202626.jpg

what about other senses?

- Smell and taste are limited through technology
  - Jaschko & Stefaner; Data cusine: food as a medium for information
  - http://data-cuisine.net/

Anja Wollenberg, Maral Jekta

Tweeting Tabouleh

CousCous
13%
Tweets from Kuweit

Parsley
13%
Tweets from Egypt

Tomato
12%
Tweets from United Arab Emirates

Cucumber
2.5%
Tweets from Lebanon

Ayran panna cotta
51%
of tweets in the Arab world come from **Saudi Arabia** alone

Tabouleh salad with ayran panna cotta

visual representation to amplify cognition

- There is a human in the loop!

- Visualisation is useful
  - When the questions are less targeted
  - When there are many possible questions
  - When the goal is to augment human capabilities rather than replace human capabilities

→ Otherwise, pure computational data analysis (without visualisation) might be sufficient

- Considering the audience and end user of your visualisation is highly important!

visual representation to amplify cognition

- Visualisations can be considered as external aids

The power of the unaided human mind is highly overrated. Without external aids, memory, thought and reasoning are all constraint.

But human intelligence is highly flexible and adaptive, superb at inventing procedures and objects that overcome its own limits.

The real powers come from devicing external aids that enhance cognitive abilities.

Don Norman, 1993

what is information visualisation?

The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.
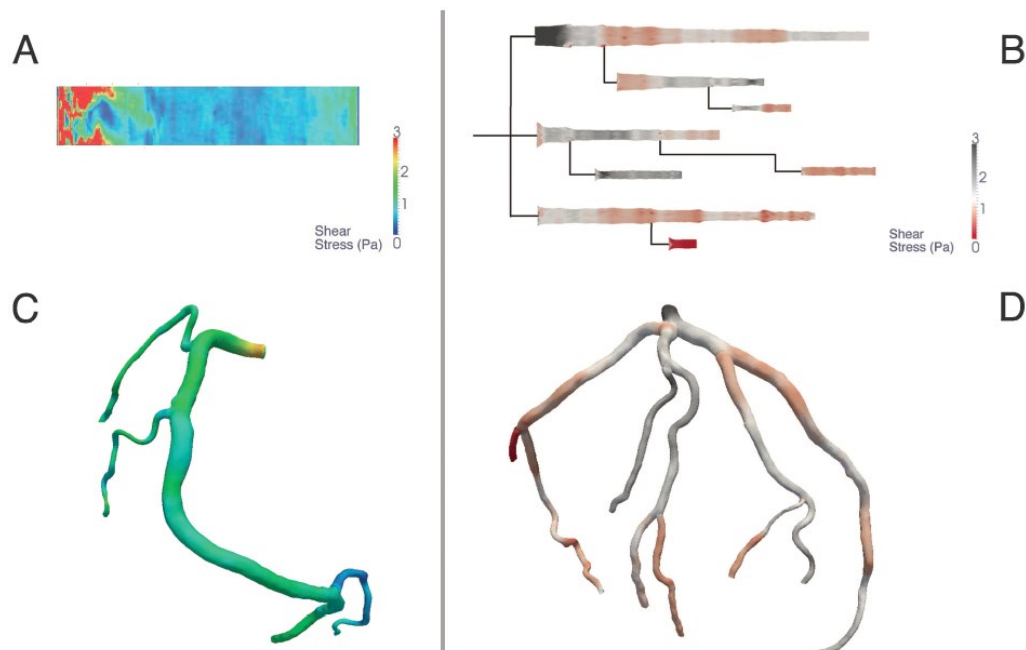
Card et al., 1999

Card et al. Readings in Information visualisation: Using Vision to Think; Chapter 1

why interactivity?

- Data can be large and/or complex!

- We cannot show everything at once!

- A single view can only show one aspect of the data – usually there is more to consider...

- Interactivity
  - to overcome limitations by people
  - to overcome technological limitations (e.g., display size and resolution)

- Common interaction paradigm in InfoVis
  - Overview, zoom-and-filter, details on demand [Shneiderman, 1996]
  - More about this later...

what is information visualisation?

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

→ working definition for this module   Card et al., 1999

- But we will also explore outside of the boundaries of this definition
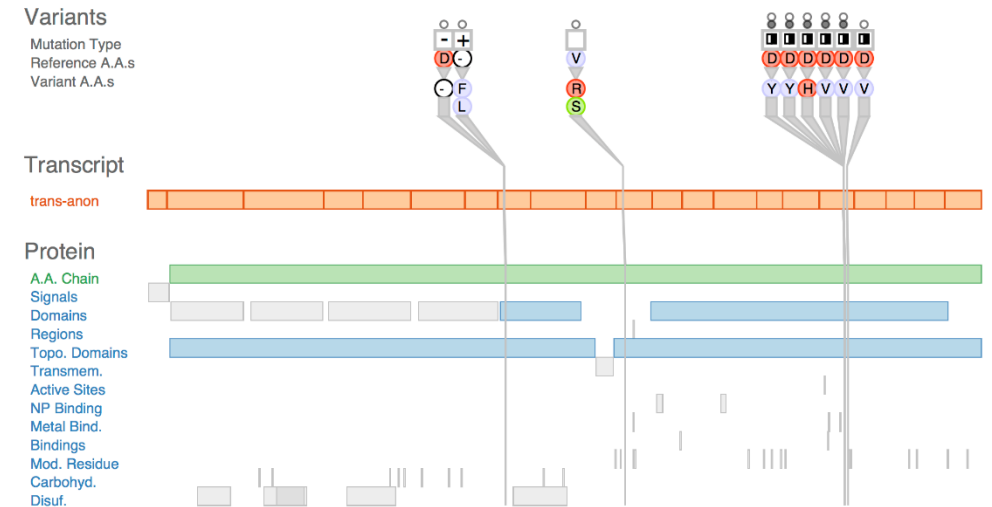 - Sketching data

Card et al. Readings in Information visualisation: Using Vision to Think; Chapter 1
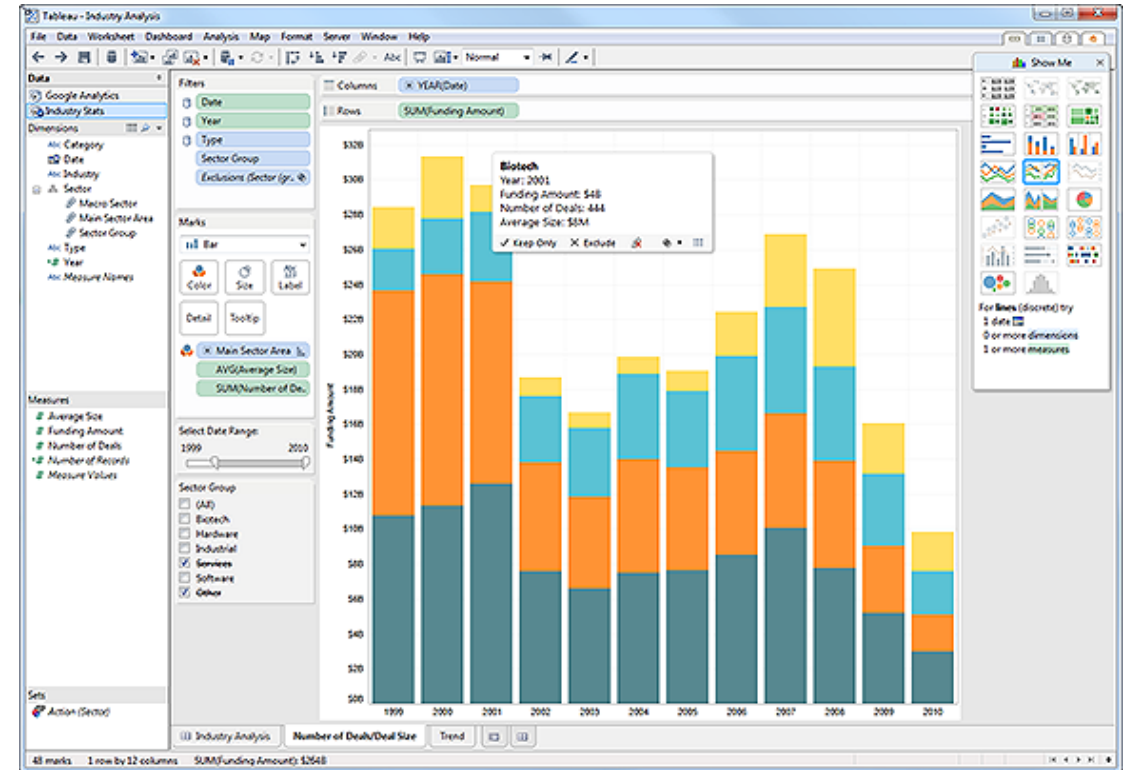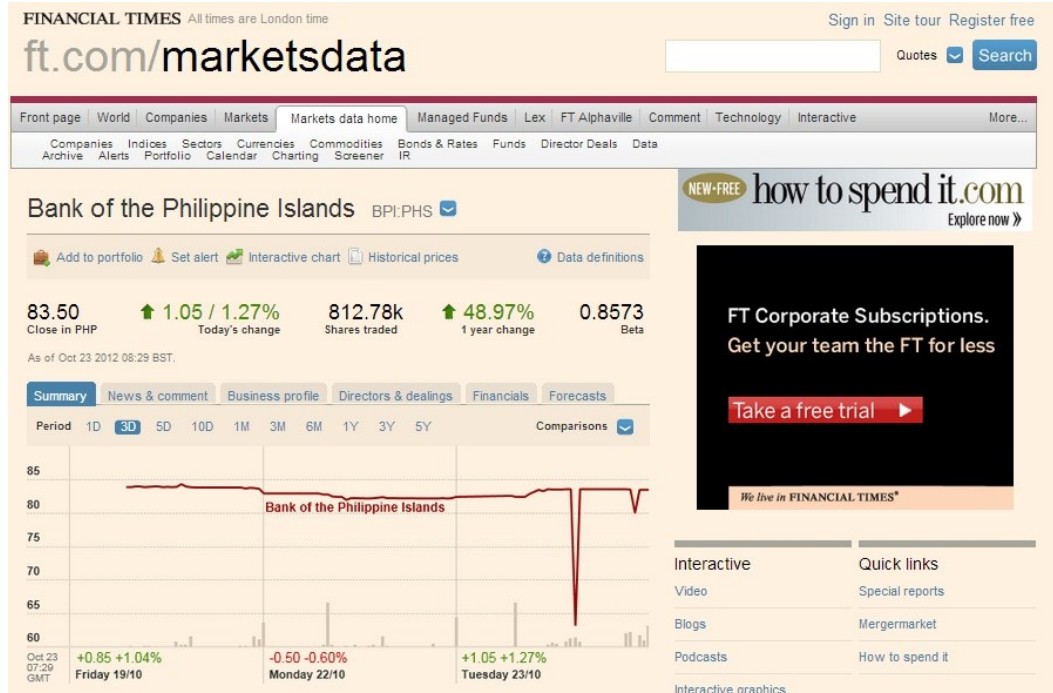
application areas

science & medicine

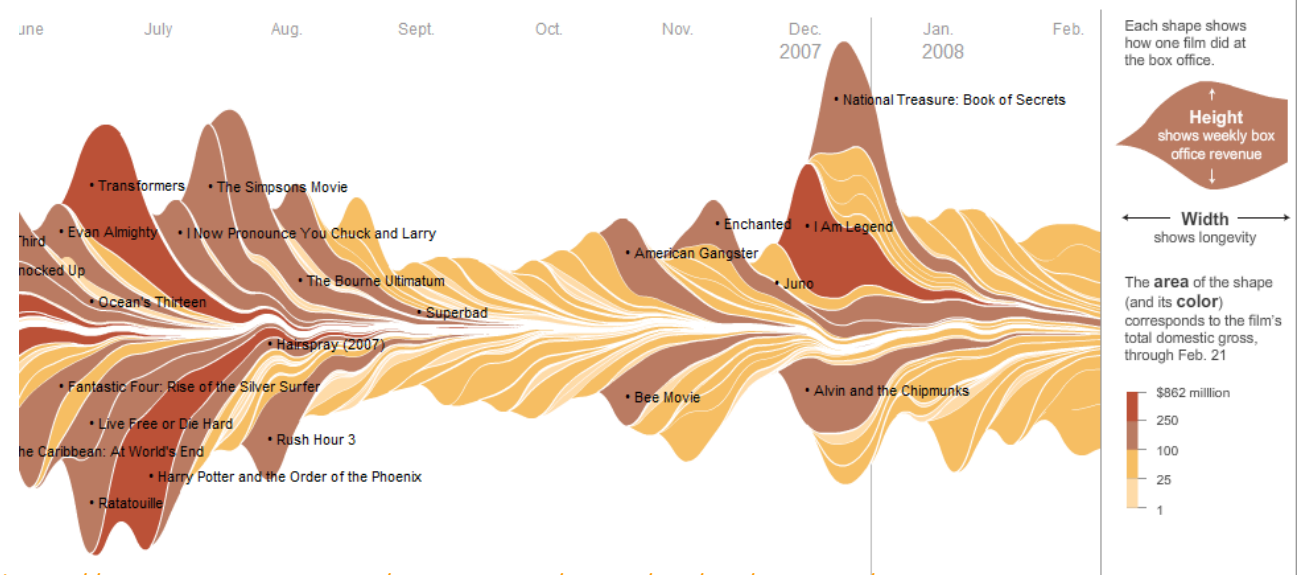http://www.cs.ubc.ca/labs/imager/tr/2013/VariantView/

http://iis.seas.harvard.edu/papers/2011/borkin11-infoviz.pdf

finance & industry

journalism

Each shape shows how one film did at the box office.

Height shows weekly box office revenue

Width shows longevity

The **area** of the shape (and its **color**) corresponds to the film's total domestic gross, through Feb. 21

$862 million
250
100
25
1

78

arts & humanities
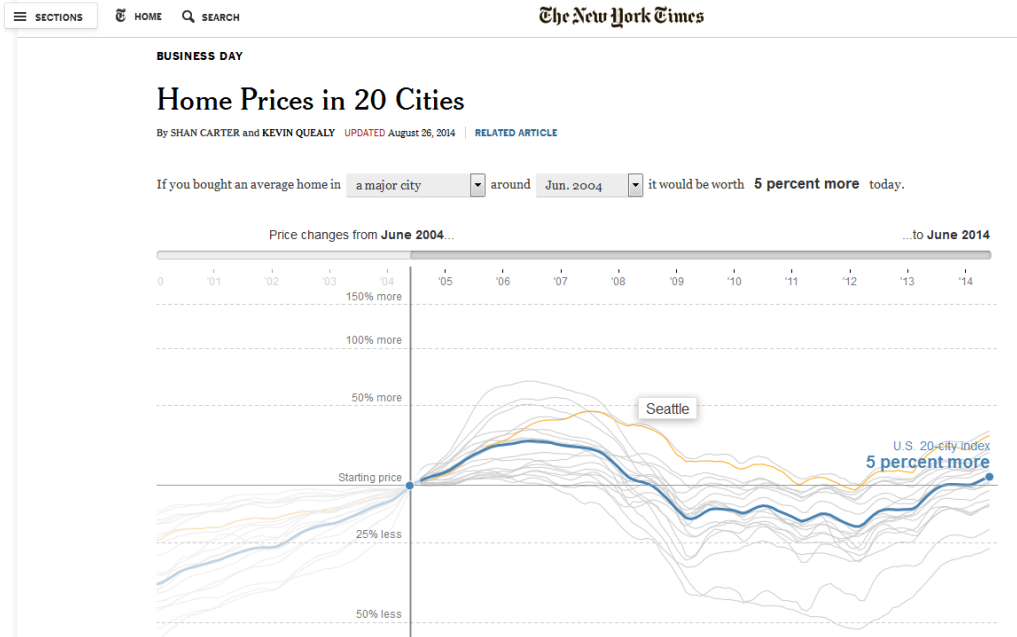
Artefacts of the Presence Era, 2006

Speculative W@nderverse, 2015

79

personal analytics

http://v.isits.in/

http://runkeeper.tumblr.com/post/47204746891/enhanced-visualisations-activity-entry-more-with
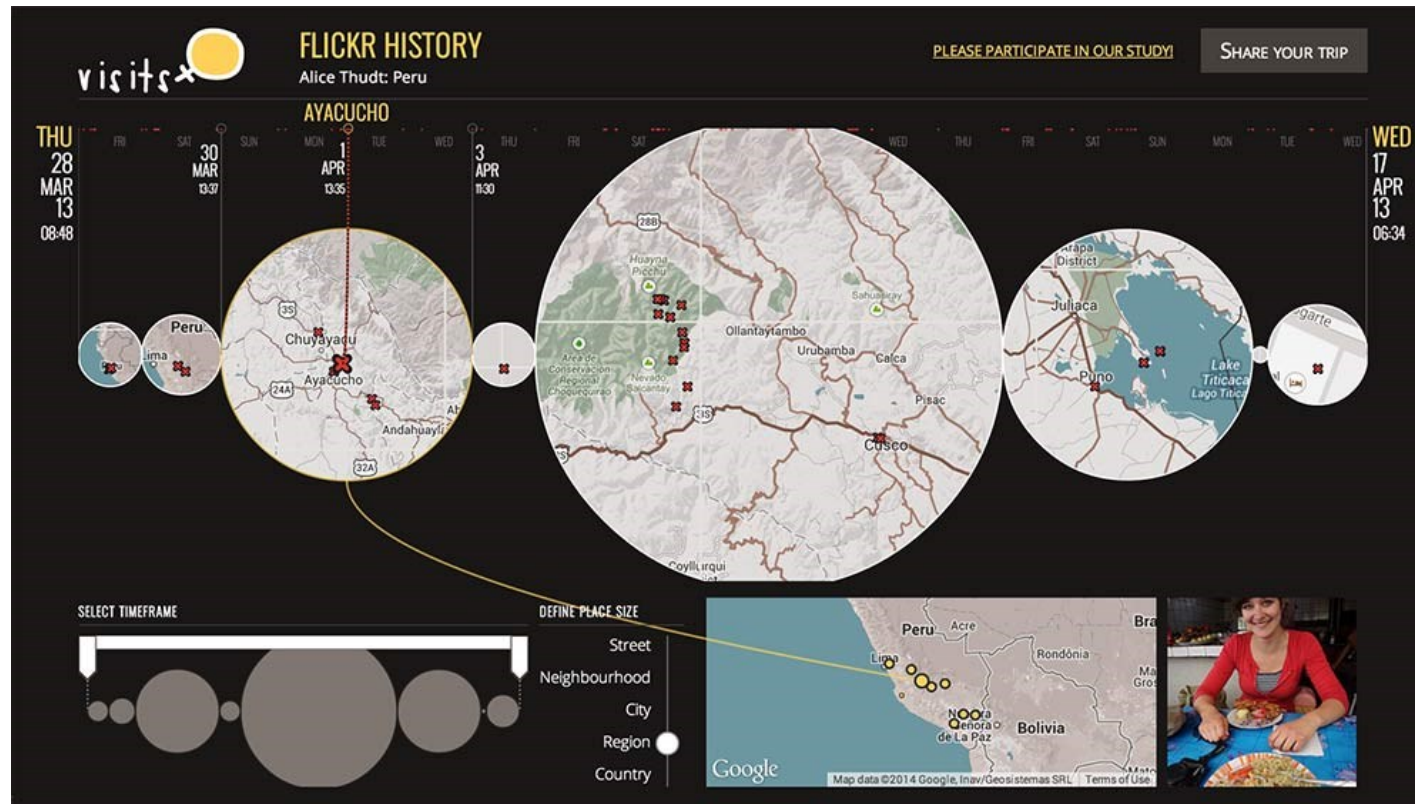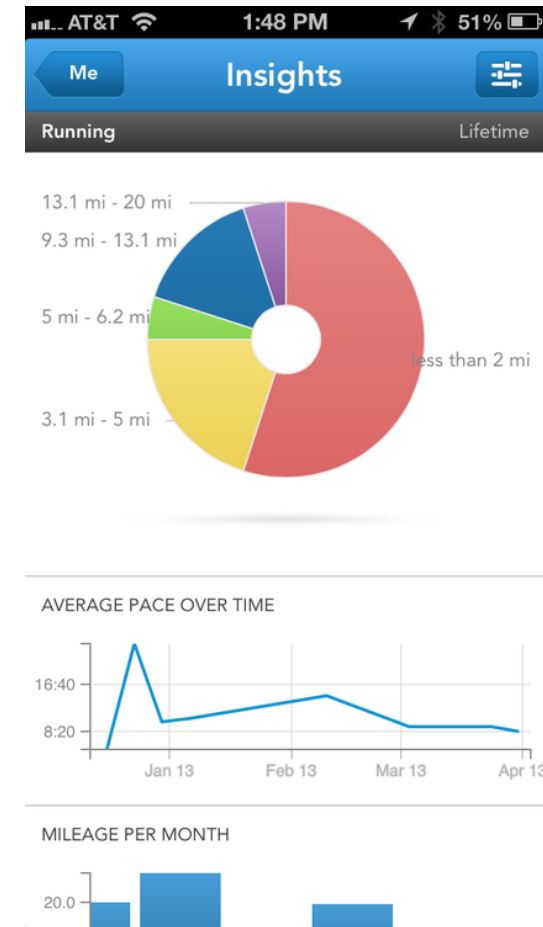
InfoVis – an interdisciplinary area

- Computer science
  - Computer graphics
  - Human-computer interaction

- Statistics

- Psychology
  - Human perception & cognition

- Visual & interaction design

- Domain experts

module overview

module overview

- Introduction to Information visualisation
- The process of information visualisation
    - What?
    - How?
    - Why?
- Guidelines to designing visualisations
- Visual perception
- Existing techniques to visualising different types of data
- Multidimensional data
- Interaction techniques for visualisation
- Validation & evaluation
- Critical interpretation / misleading with visualisation
- Narration & storytelling

Thursday – week 1

- The process of visualisation
  - What?
  - How?

- Readings for Thursday
  - Tamara Munzner: Visualisation Analysis and Design;
    - Chapter 2: Data abstraction
    - Chapter 5: Marks and channels (optional)
    - Chapter 1: What's Vis and Why Do it? (optional)