Lesson 1: Introduction to Visual Analytics

Dr. Kam Tin Seong
Assoc. Professor of Information Systems

School of Computing and Information Systems,
Singapore Management University

2019-04-28 (updated: 2022-04-12)
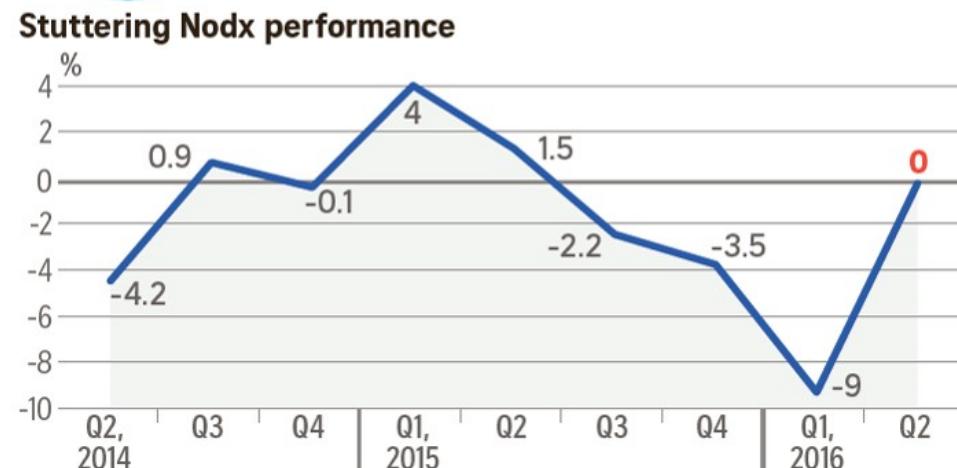
What will you learn from this lesson?

- Motivation of Visual Analytics
- What is Visual Analytics
- Learning from the history of visual analytics
- Data Visualisation tools

Motivation of Visual Analytics
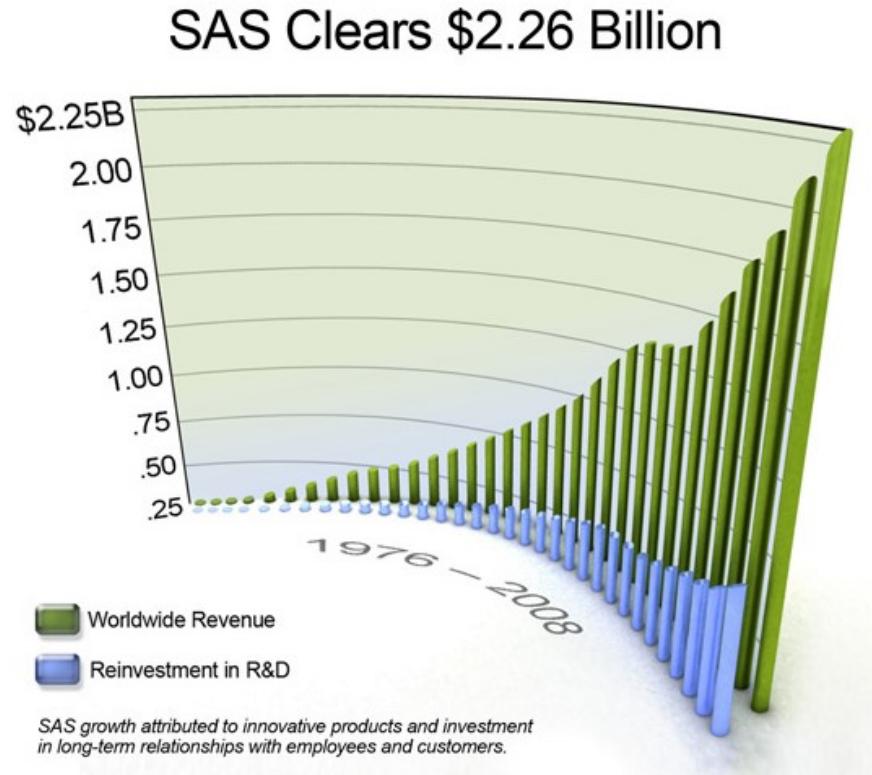
The World is full of charts that lie!

Non-oil domestic exports (Nodx) growth was flat in the second quarter. For the first half this year,

Nodx fell 4.5 per cent

compared with the same period last year.

Stuttering Nodx performance

Motivation of Visual Analytics

Junk charts are good for advertising but not for understanding.

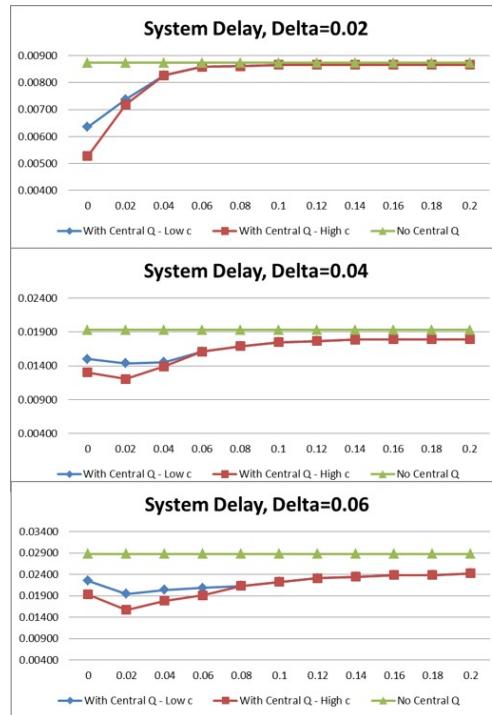# Motivation of Visual Analytics

Data rich, information poor

Source: http://omnitele-com.s3.frantic.com/2015/04/big-data-ocean-crop.jpeg

Motivation of Visual Analytics
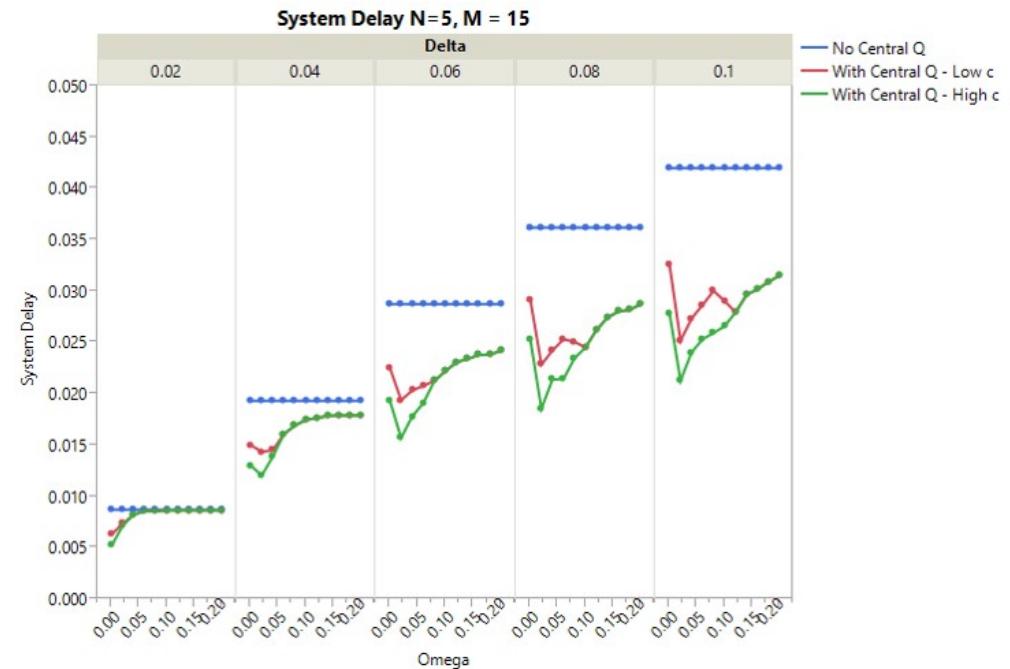
The World is buried in data,

Just SHOW me the number

Motivation of Visual Analytics
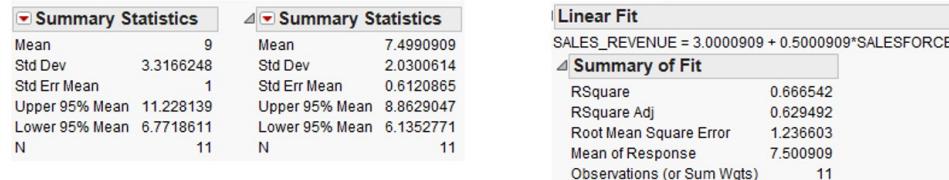
Spreadsheet: A great hammer but lousy screwdriver

Visual analytics come to the rescue

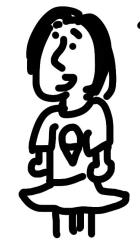# Motivation of Visual Analytics

When statistics failed!

SALESFORCE A	SALES_REVENUE A	SALESFORCE B	SALES_REVENUE B	SALESFORCE C	SALES_REVENUE C	SALESFORCE D	SALES_REVENUE D
10	8.04	10	9.14	10	7.46	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

Mean	9
Std Dev	3.3166248
Std Err Mean	1
Upper 95% Mean	11.228139
Lower 95% Mean	6.7718611
N	11

Summary Statistics

Mean	7.4990909
Std Dev	2.0300614
Std Err Mean	0.6120865
Upper 95% Mean	8.8629047
Lower 95% Mean	6.1352771
N	11

Linear Fit

SALES_REVENUE = 3.0000909 + 0.5000909*SALESFORCE

Summary of Fit

RSquare	0.666542
RSquare Adj	0.629492
Root Mean Square Error	1.236603
Mean of Response	7.500909
Observations (or Sum Wgts)	11

I just can't make sense of this data.

Have you tried looking at the pictures?

fresh spectrum

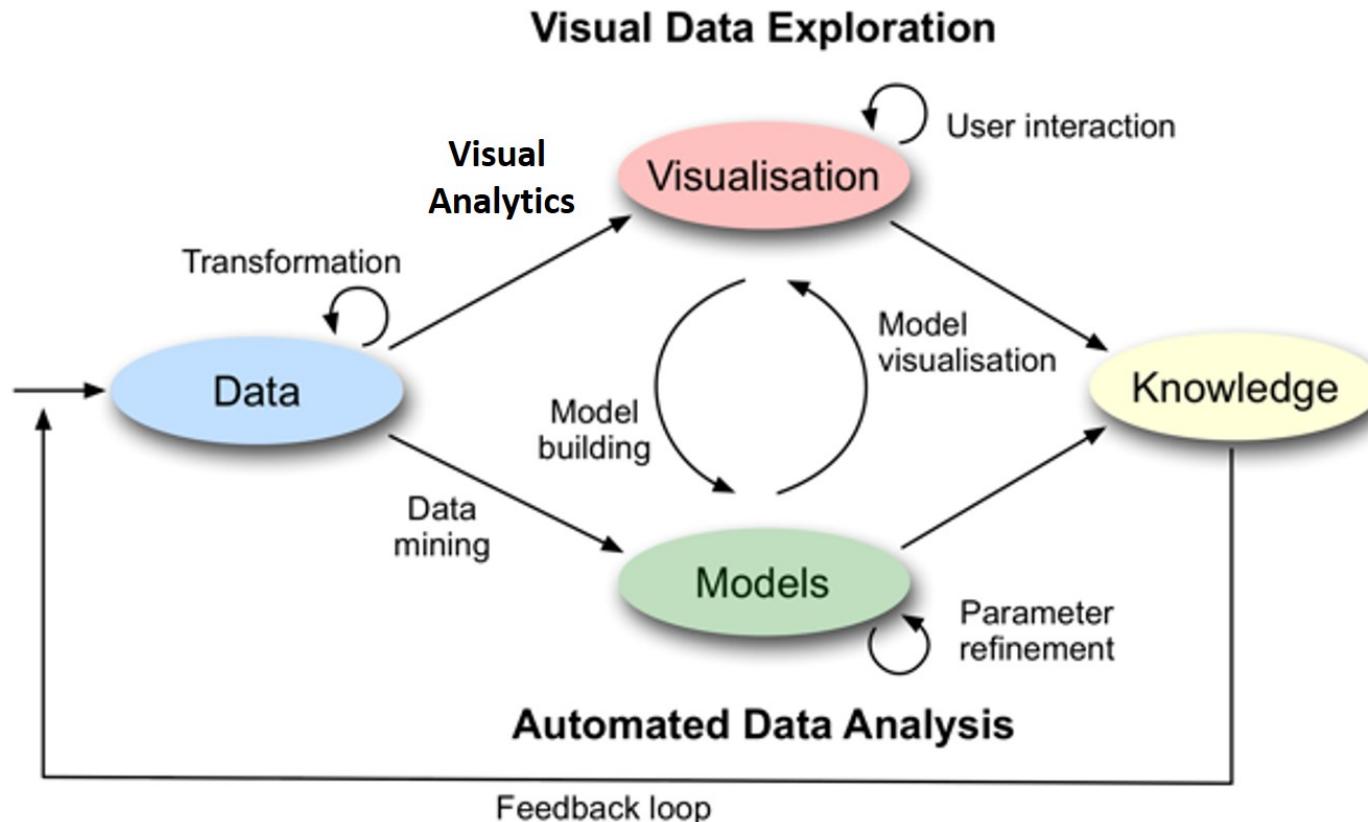Data analysis without data visualisation is no data analysis

Motivation of Visual Analytics
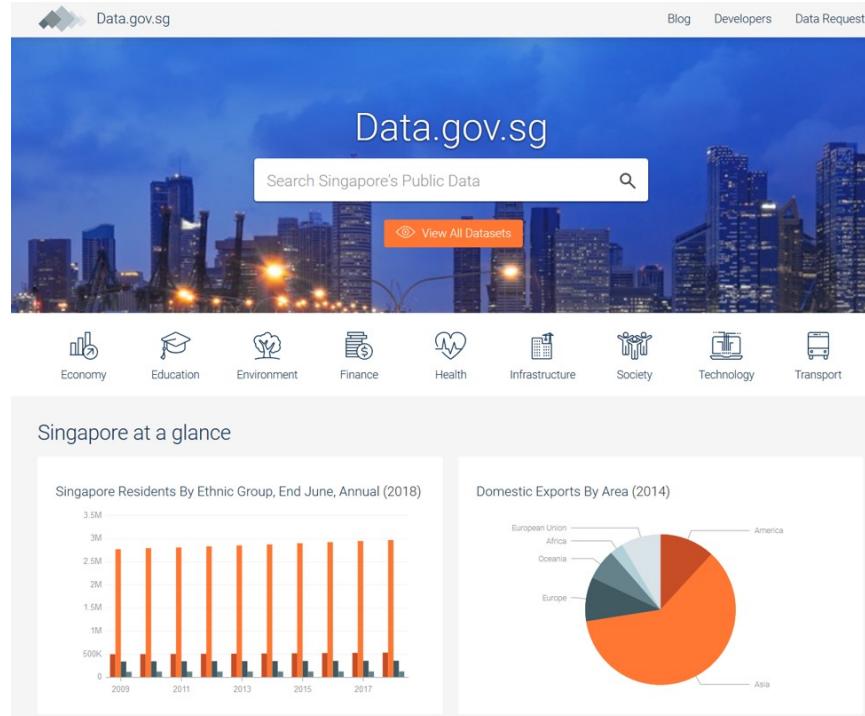
Analytical graphics reveal the truth

Motivation of Visual Analytics

Is there a place for Visual Analytics in knowledge discovery?

Visual Data Exploration
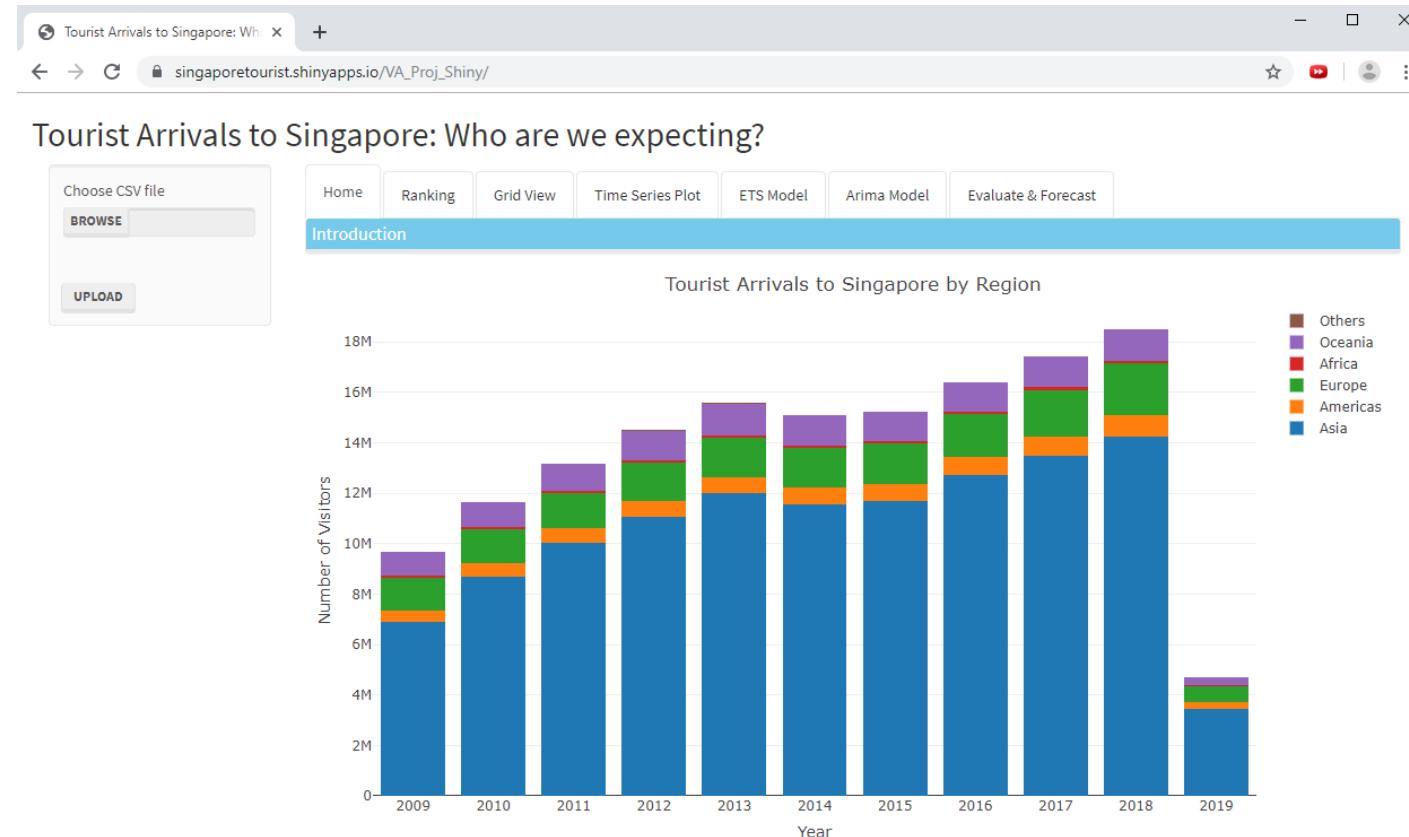
Visualisation — User interaction

Visual Analytics

Transformation

Data

Model visualisation

Model building

Data mining

Knowledge

Models — Parameter refinement

Automated Data Analysis

Feedback loop

Motivation of Visual Analytics

Open data need to be unlocked!

Visit the link at https://data.gov.sg

Motivation of Visual Analytics

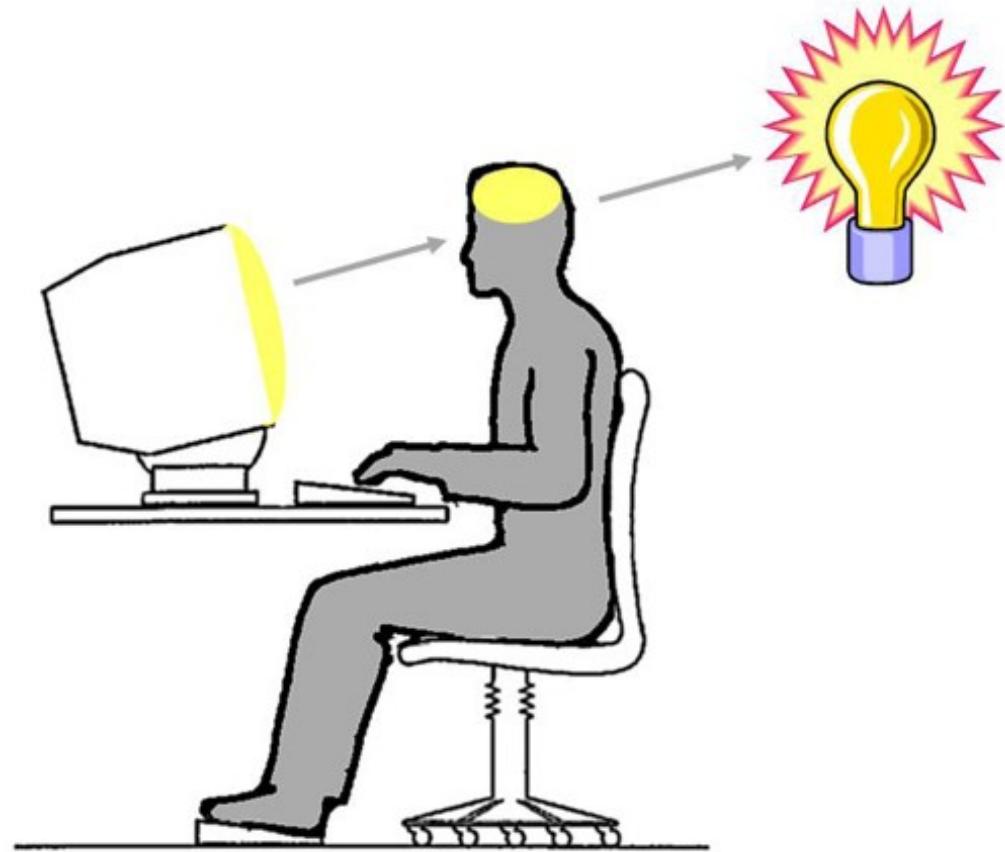## Democratising Data and Analytics with Visually-driven Web Apps

Visit the link at https://singaporetourist.shinyapps.io/VA_Proj_Shiny/

What is Visual Analytics?

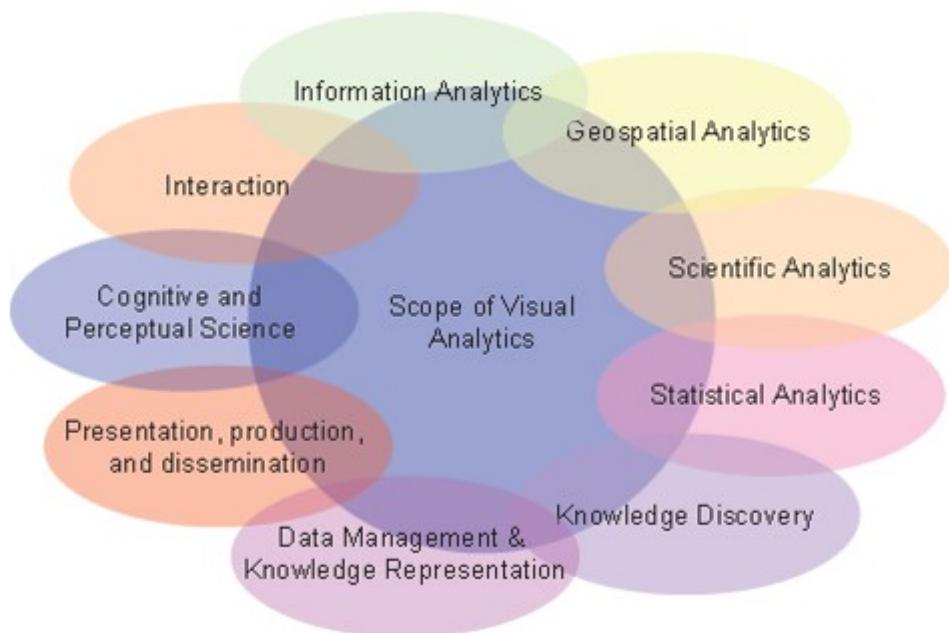Visual Analytics is the craft of **analytical reasoning** supported by:

- Data Science and Analytics
- Data Visualisation
- Human-Computing-Interaction (HCI)
- Interface Design
- Visual Storytelling

The marriage of data analytics and visualizations

What is Visual Analytics?

Science of Visual Analytics

Aims of Visual Analytics

Detecting the expected, discovering the unexpected

What is Visual Analytics?

Visual Data Sense-making

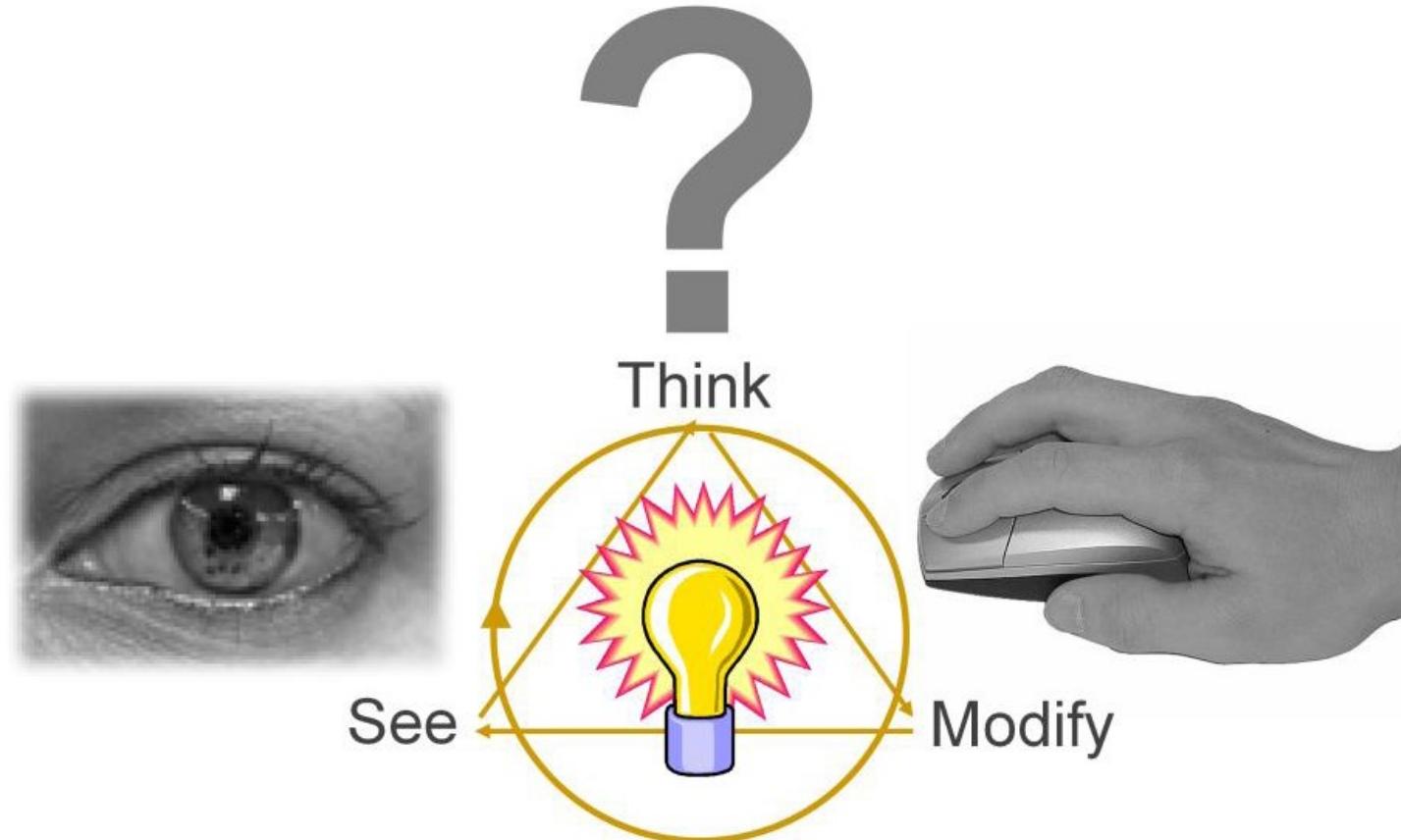Visual Analytics' greatest strength is its use for **sense-making**.

When you have eliminated the impossible, whatever remains, however improbable, must be the truth.
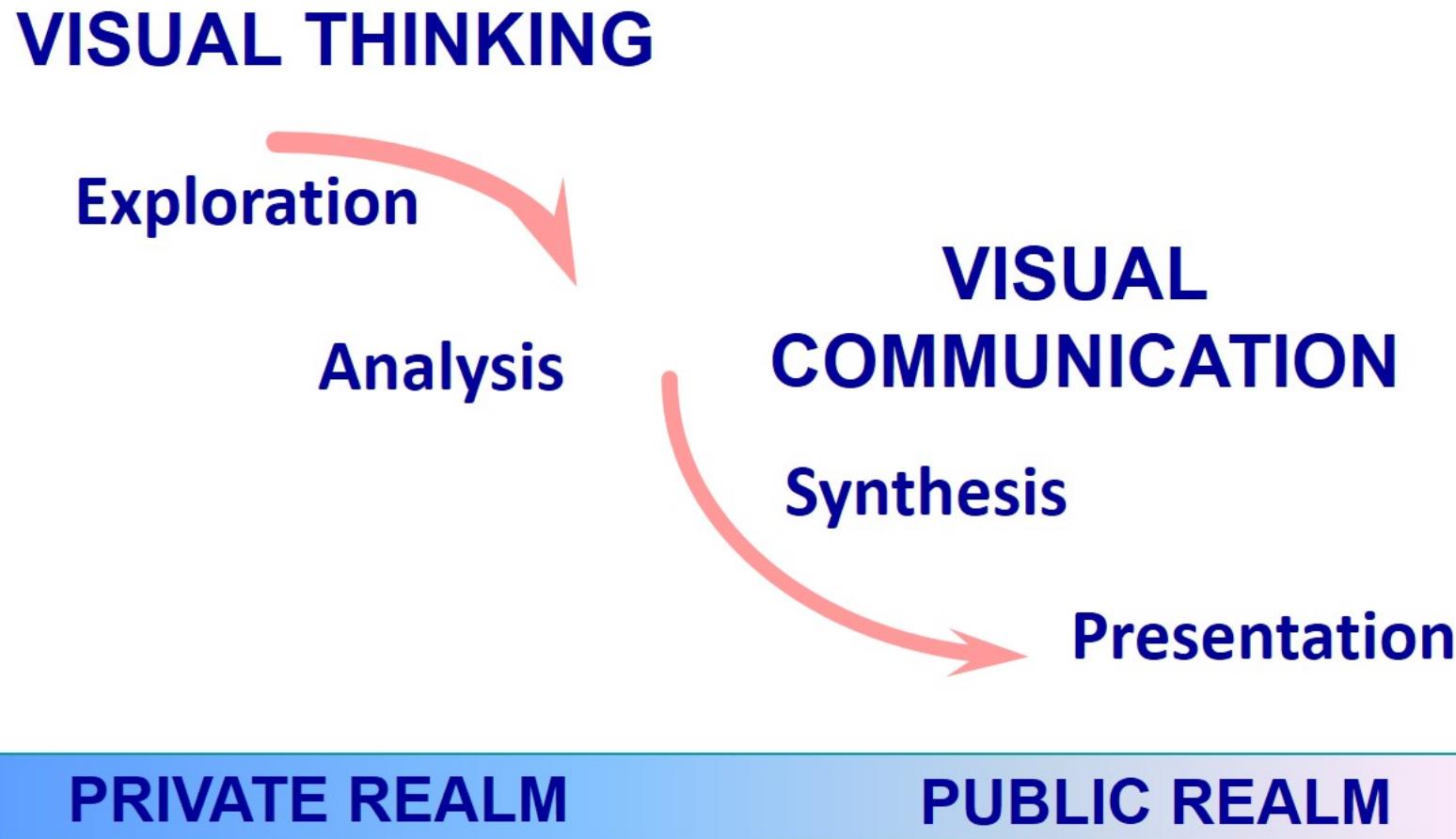- Sherlock Holmes

What is Visual Analytics?

Visual Analytics Framework
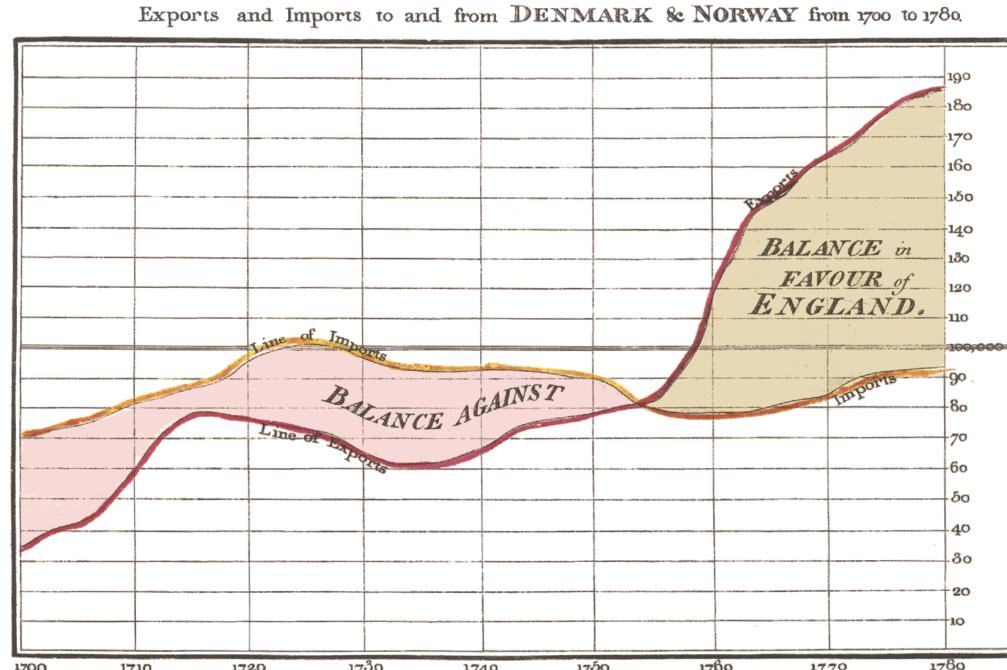
What is Visual Analytics?

Visual Analytics require a shift of mindset

VISUAL THINKING

Exploration

Analysis

VISUAL COMMUNICATION

Synthesis

Presentation

PRIVATE REALM   **PUBLIC REALM**

What is Visual Analytics?

Learning from the history of Visual Analytics
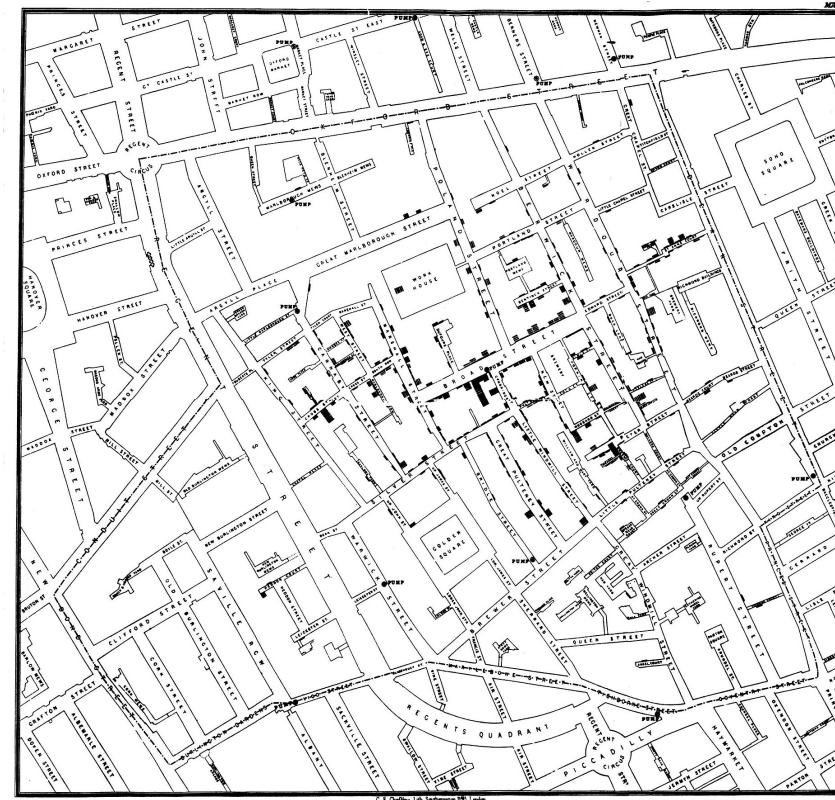
Commercial and Political Atlas by William Playfair (22 September 1759 – 11 February 1823)

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.

What is Visual Analytics?

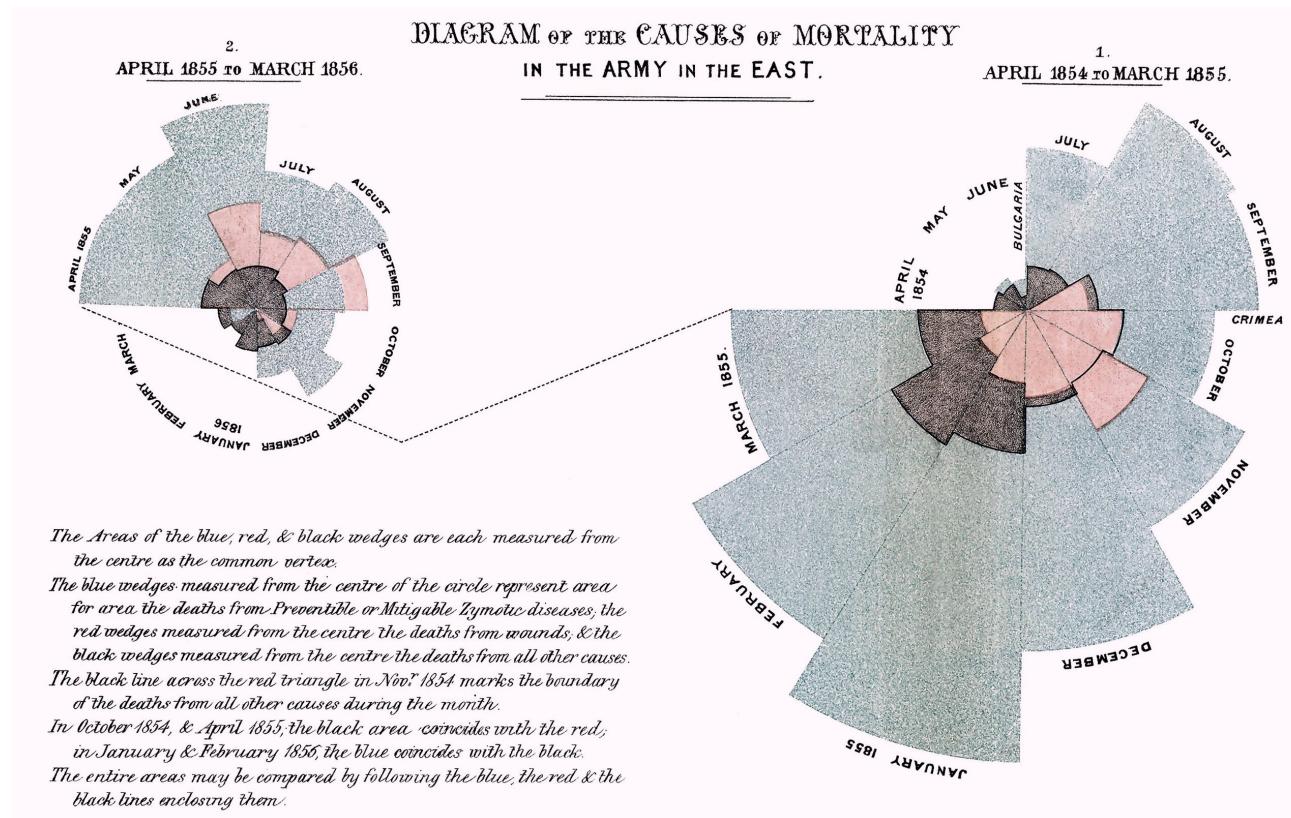## Learning from the history of Visual Analytics

Cholera and the Pump on Broad Street(John Snow, 1854).

What is Visual Analytics?

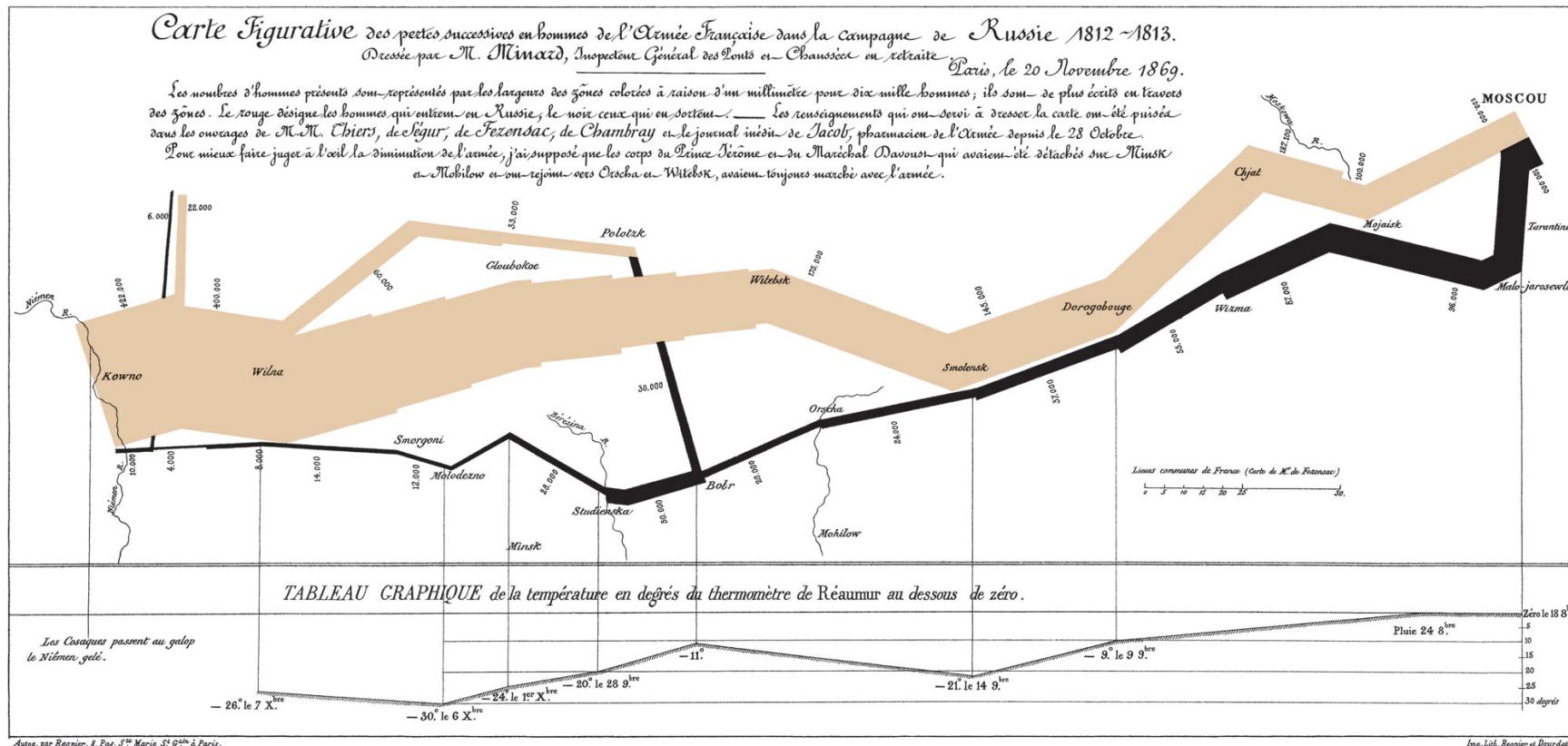## Learning from the history of Visual Analytics

Florence Nightingale: The compassionate statistician (12 May 1820 – 13 August 1910)

What is Visual Analytics?

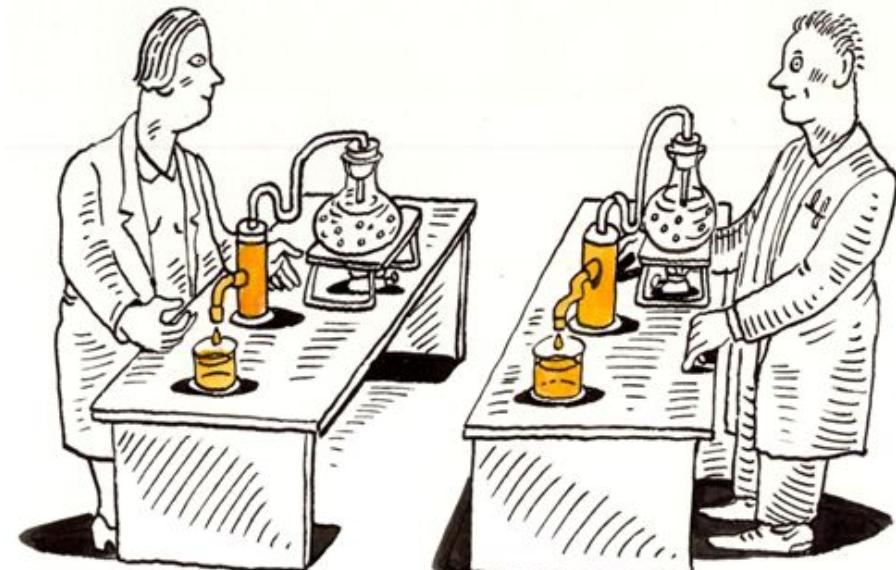## Learning from the history of Visual Analytics

Plot of Napoleon's ill-fated Russian campaign of 1812-1813 (Charles Joseph Minard,1861)

What is Visual Analytics?

Visual Analytics must be Reprducible

- For every result, keep track of how it was produced.
- Avoid manual data manipulation steps.
- Archive the excat versions of all external programmes used.
- Version control all custom scripts.
- Record all intermediate results, when possible in standard formats.
- For analysis that include randomness, note underlying random seeds.
- Always store raw data behind plots.
- Generate hierarchical analysis output, allowing layers of increasing detail to be inspected.

- Connect textual statements to underlying results.
- Provide public access to scripts, runs, and results.

Data Visualisation Toolkits

- Commercial-off-the-shelf
  - Visual analytics tookit such as Tableau, JMP, PowerBI, Qlik Sense/View & TIBCO SpotFire.
  - Specialised data visualisation applications such as TreeMap from Macrofocus (https://www.macrofocus.com/)
- Open source
  - Specialised data visualisation applications such as Gephi (https://gephi.org/)
  - Data visualisation API library such as ggplot2, plotly, and D3.js (https://d3js.org/)

COS Data Visualisualisation Toolkit

Magic Quadrant for BI and Analytics Platforms

COS Data Visualisation Toolkit

Data Visualisation Centric

BI Centric

Tableau: A data visualisation tool for everyone

- Theory to practice
  - The Grammar of Graphics
- Science to product
  - VisQL
  - Show Me
  - Polaris

Programmable Data Visualisation Library

- Javascript
  - D3.js (https://d3js.org/)
- Python
  - Maplotlib (http://matplotlib.org/)
  - Bokeh (http://bokeh.pydata.org/en/latest/)
  - Plotly (https://plot.ly/python/)
- R, Base Graphics
  - lattice (https://cran.r-project.org/web/packages/lattice/index.html)
  - ggplot (https://ggplot2.tidyverse.org/)
  - plotly R (https://plot.ly/r/)
  - R Shiny (https://shiny.rstudio.com/)

Data Visualisation Toolkit for Data Journalism

- Flourish

- Datawrapper

Highly recommended blogs

- Perceptual Edge (https://www.perceptualedge.com/)
- Junk Charts (http://junkcharts.typepad.com/junk_charts/)
- The Functional Art (http://www.thefunctionalart.com/)