

How to data 101

How to scrape, clean, and visualise data.

Scrape. theory + python
Clean. python
Visualise. theory + rawGraphs+

Add some data sketches?

Data Visualization

Data Visualization

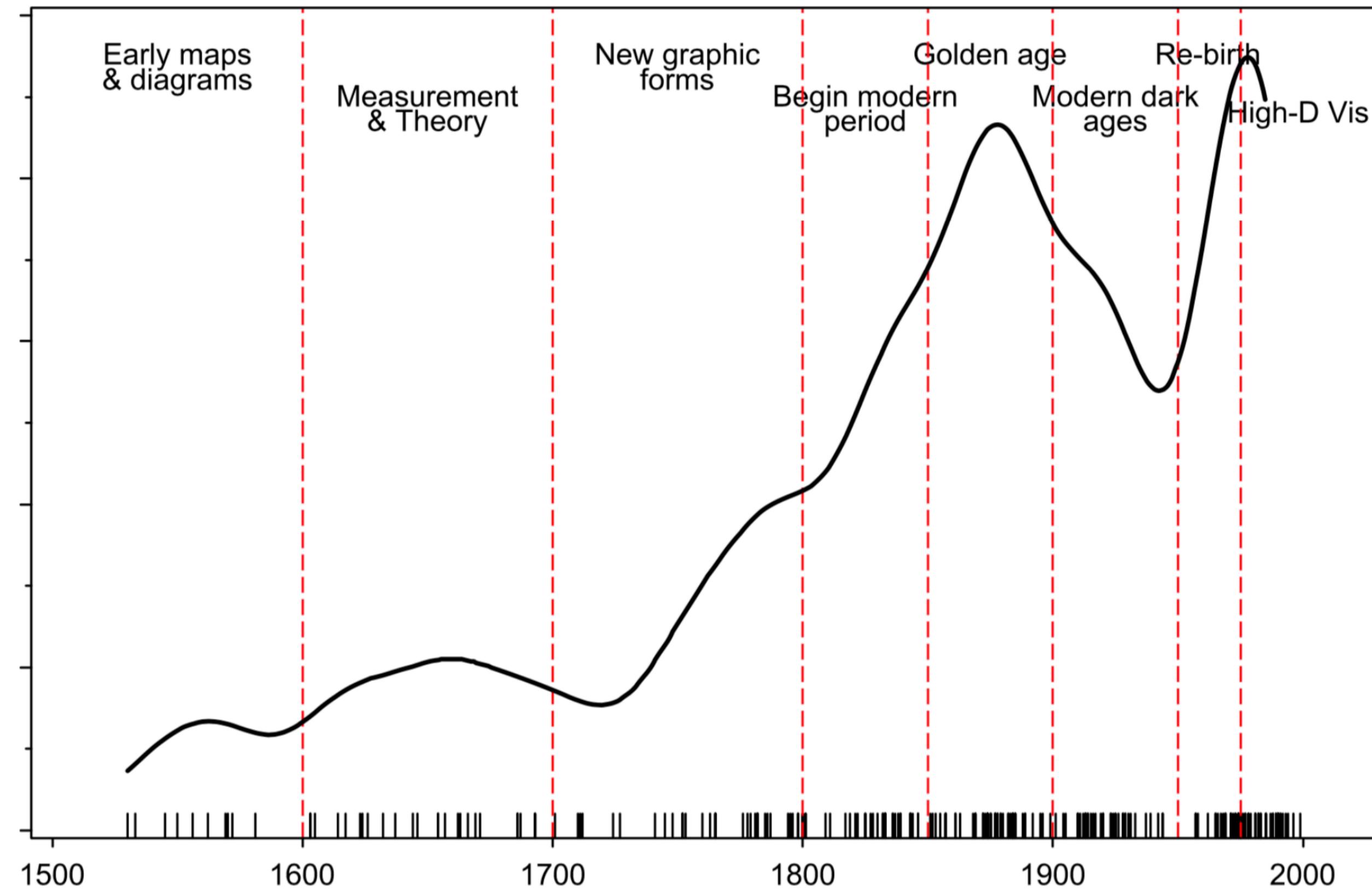
“The use of computer-supported, interactive visual representations of data to amplify cognition”¹

“Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.”²

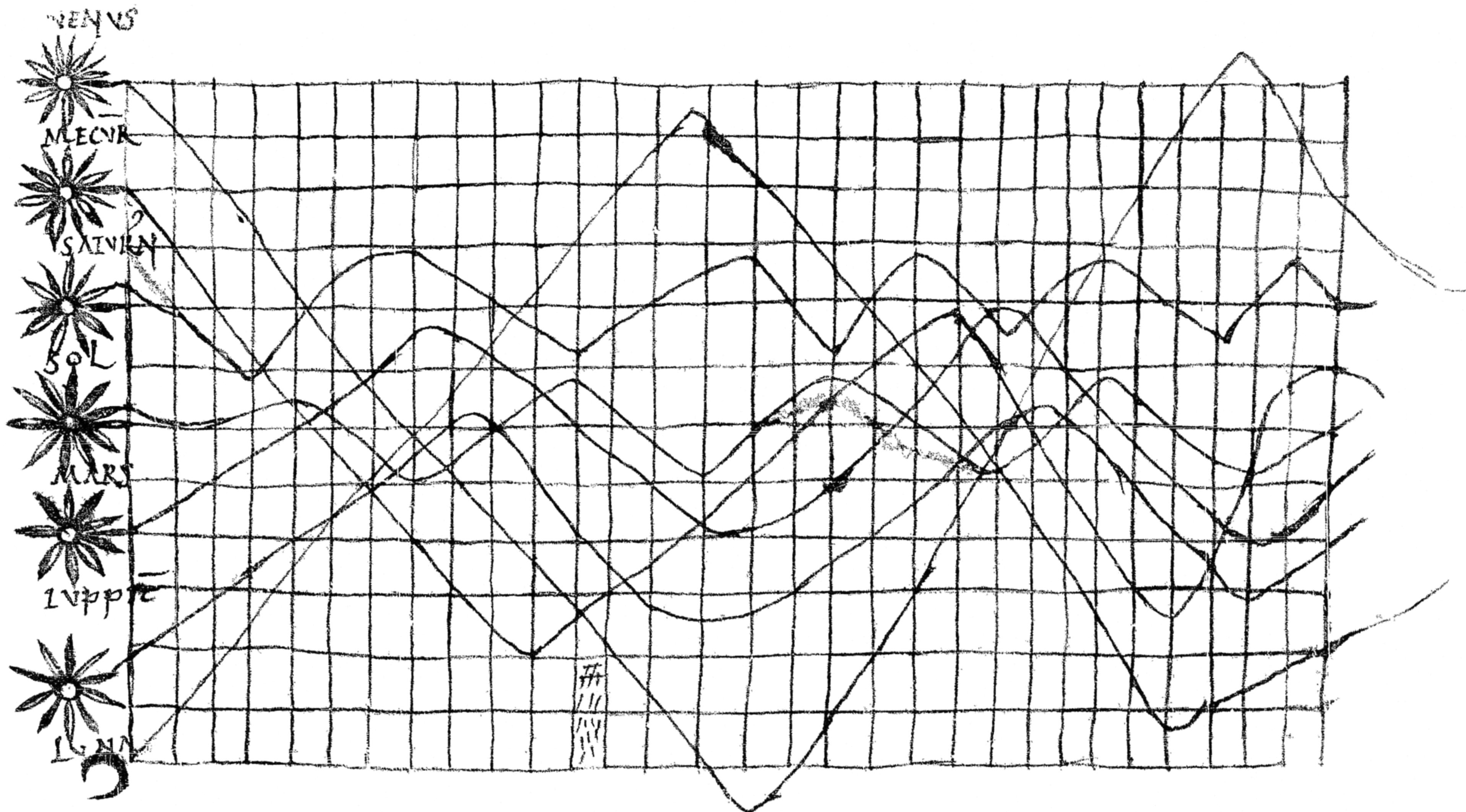
1 Card, S. K., Mackinlay, J. D., and Shneiderman, B. (1999). Using Vision to Think, chapter 1: Information Visualization, pages 1–34. Morgan Kaufmann.

2 Munzner, T., (2014). Visualization Analysis and Design, A K Peters/CRC Press.

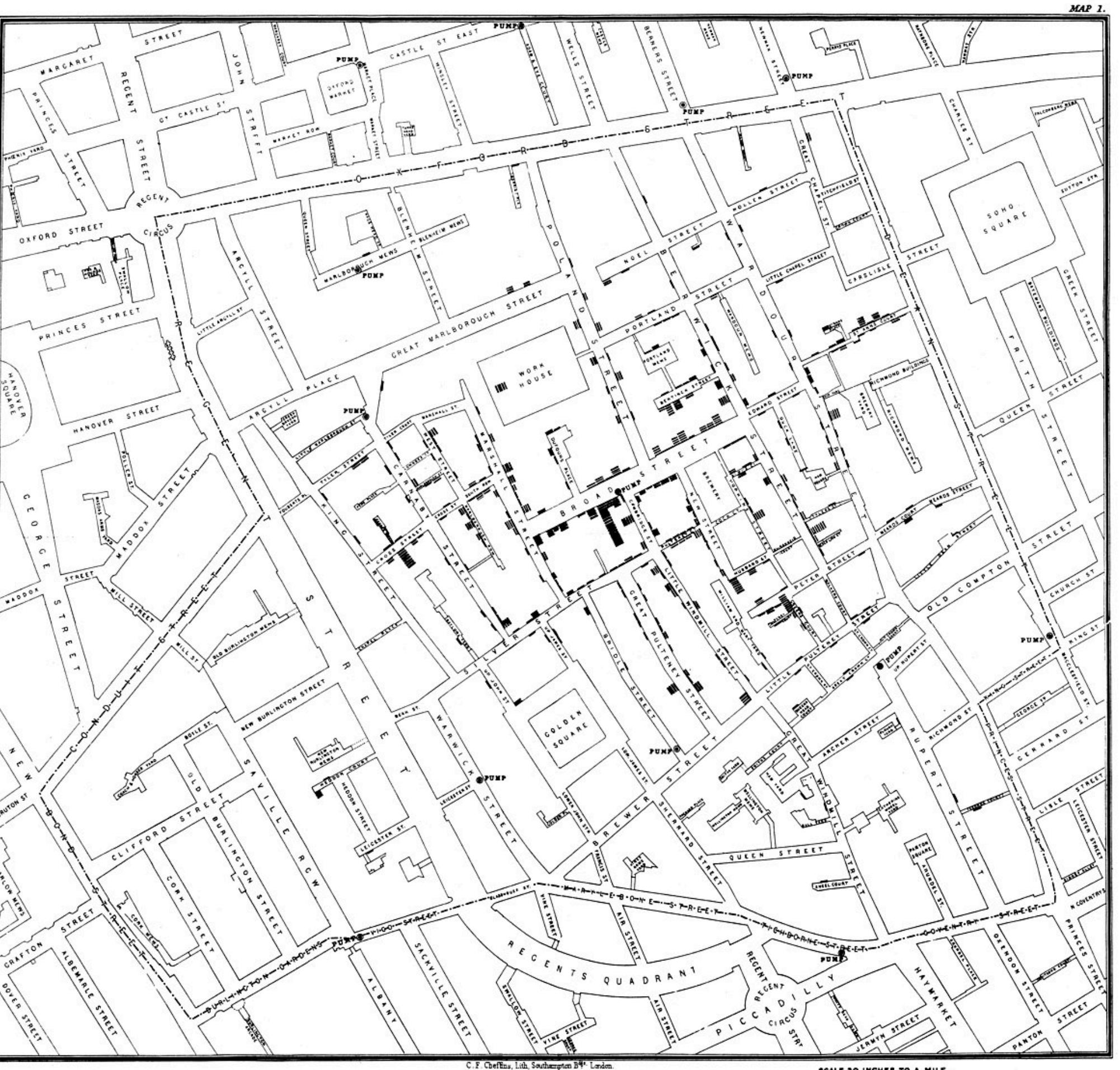
A short history



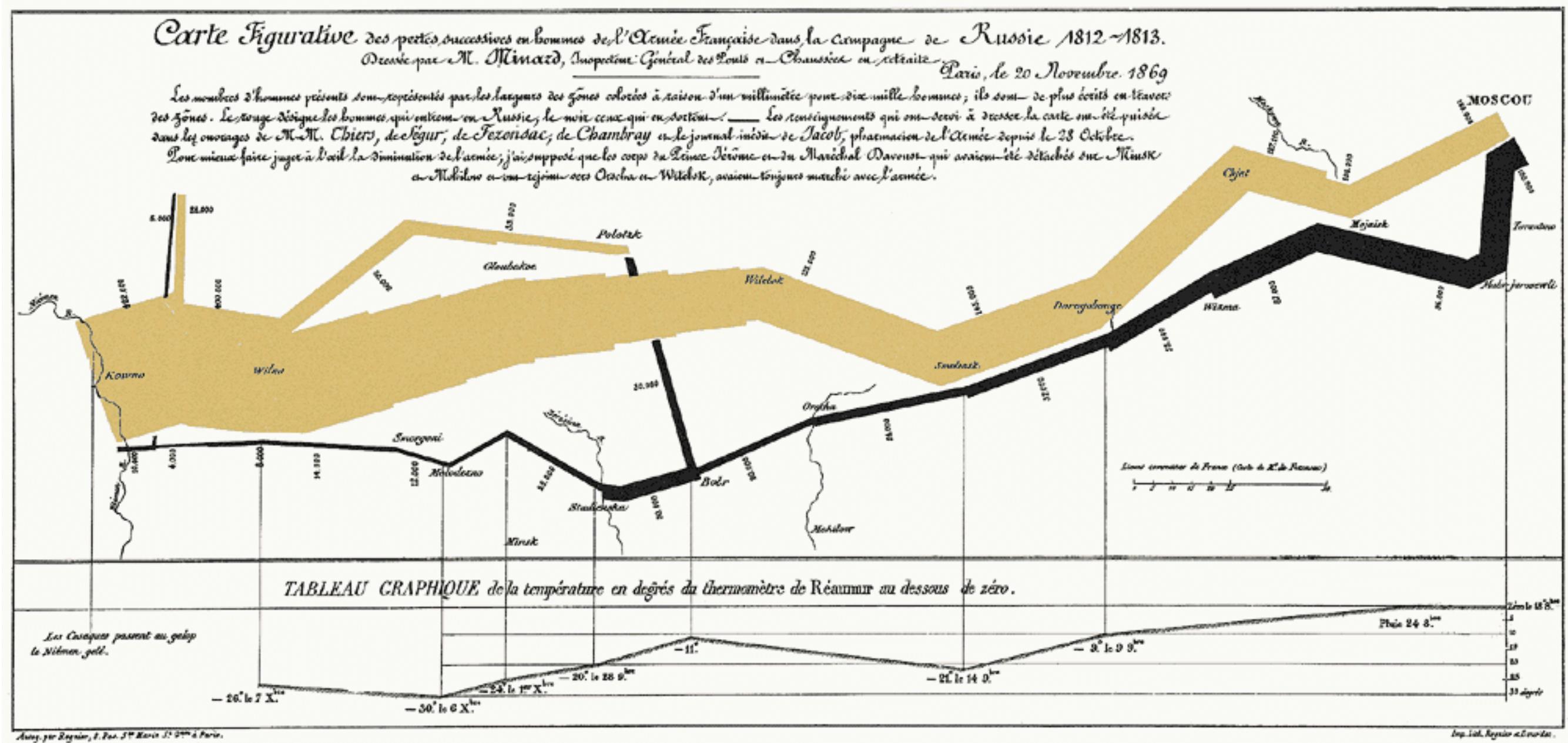
A short history



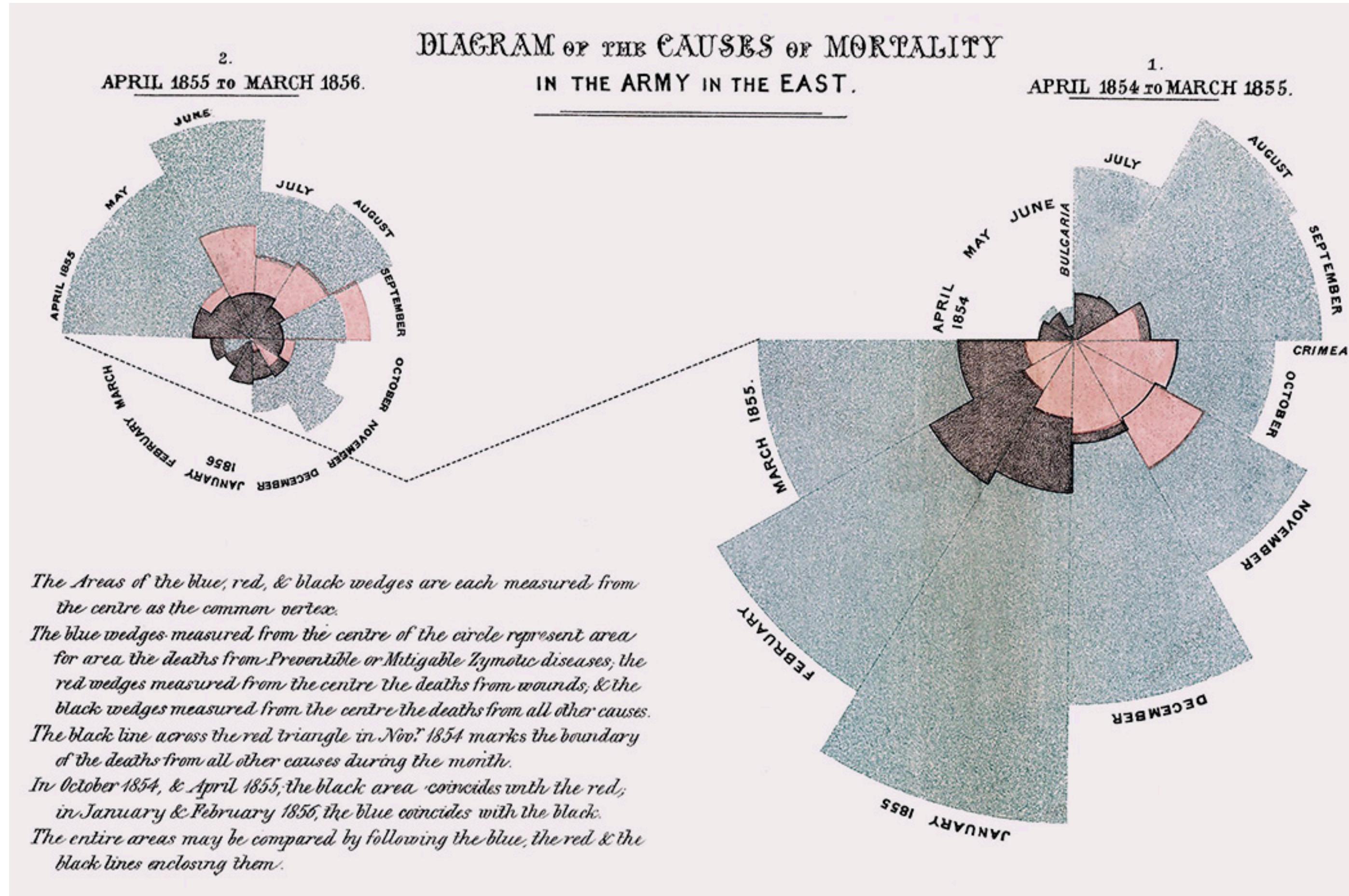
A short history



A short history



A short history



A short history

Something here to point towards a more global look at datavis?

What is it for?

Communication.

Insights.

What else?

Entertainment.

Supporting cognition

Increases resources: working memory, parallel processing

Reduce search: high data density, including “overview”

Pattern recognition: abstraction and aggregation

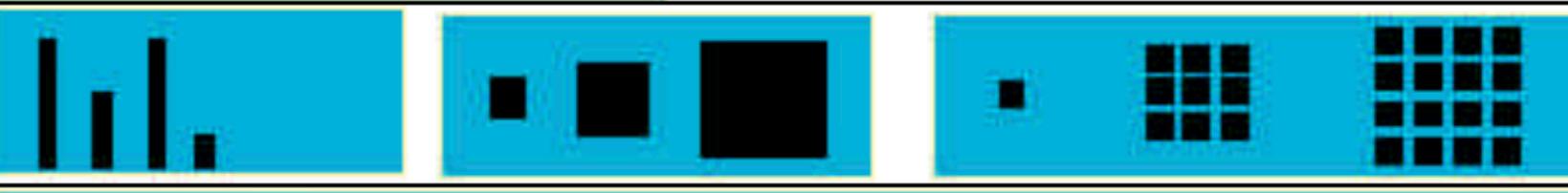
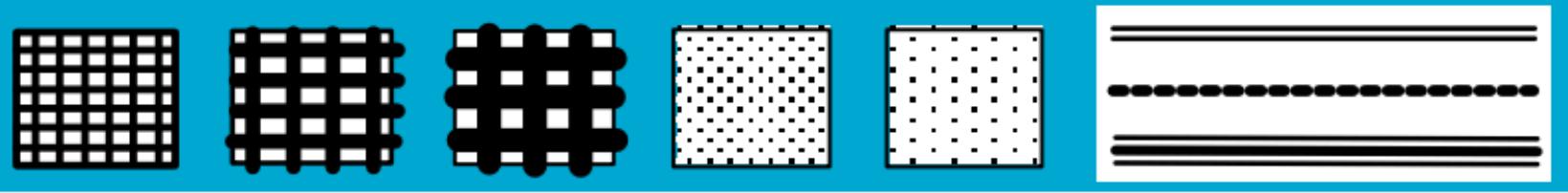
Perceptual inferences: making possible relationships obvious

Perceptual monitoring: revealing trends and changes

Interactive manipulation: adjusting parameters, filters and views

Visual Variables

Position	X-Y location
Size	Location
Value	Length or area brightness
Texture	Visual pattern
Colour	Hue & Saturation
Orientation	Alignment
Shape	Visual form
Animation	Change visual vars

Bertin's Original Visual Variables	
Position changes in the x, y location	
Size change in length, area or repetition	
Shape infinite number of shapes	
Value changes from light to dark	
Colour changes in hue at a given value	
Orientation changes in alignment	
Texture variation in 'grain'	

Bertin, J. (1983). Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press.

Carpendale, S. (2003). Considering visual variables as a basis for information visualisation.
Technical Report 2001-693-16, Department of Computer science, University of Calgary.

Visualization types



Prepare data  Process data  Present data

10 minutes break :)

Hands on tutorials.

<https://github.com/uclab-potsdam/filmuni-workshop/>

15 minutes break :)

Next week