

lecture 02: Visualization Fundamentals

September 13, 2017

Data

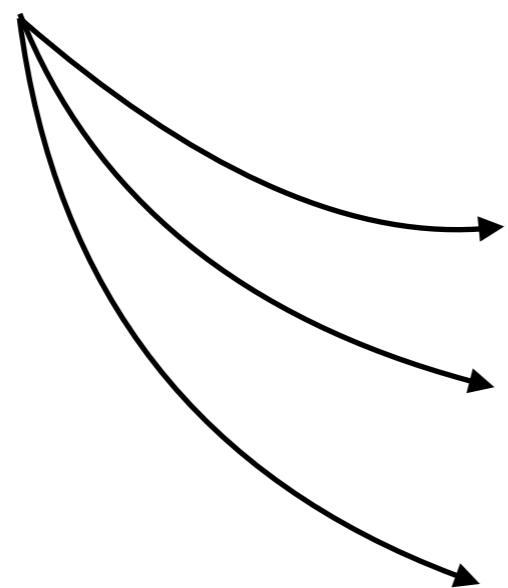
Last class we talked about “tidy data” but there are other ways data can be represented

Hadley Wickham (author of ggplot2, an R package for data viz)

has an idea of

Tidy Data

Rows are observations



Columns are variables

There are other ways to format and represent data

Another common method is a hierarchical, or list-based structure



We often make the
distinction between

Numerical

discrete

continuous

and

Categorical

“normal”

ordinal

Let's brainstorm some variables that could be recorded
about you that are numerical and categorical

A big question for this course is how to best map variables to visual attributes

Some (all?) of the visual attributes we have to play with

	Points	Lines	Areas	Best to show
Shape		possible, but too weird to show	cartogram	qualitative differences
Size			cartogram	quantitative differences
Color Hue				qualitative differences
Color Value				quantitative differences
Color Intensity				qualitative differences
Texture				qualitative & quantitative differences

Some history

(old white guy alert)

Laxcaux, France

cave paintings

15,000 B.C.



(via Jordan Crouser)

900s

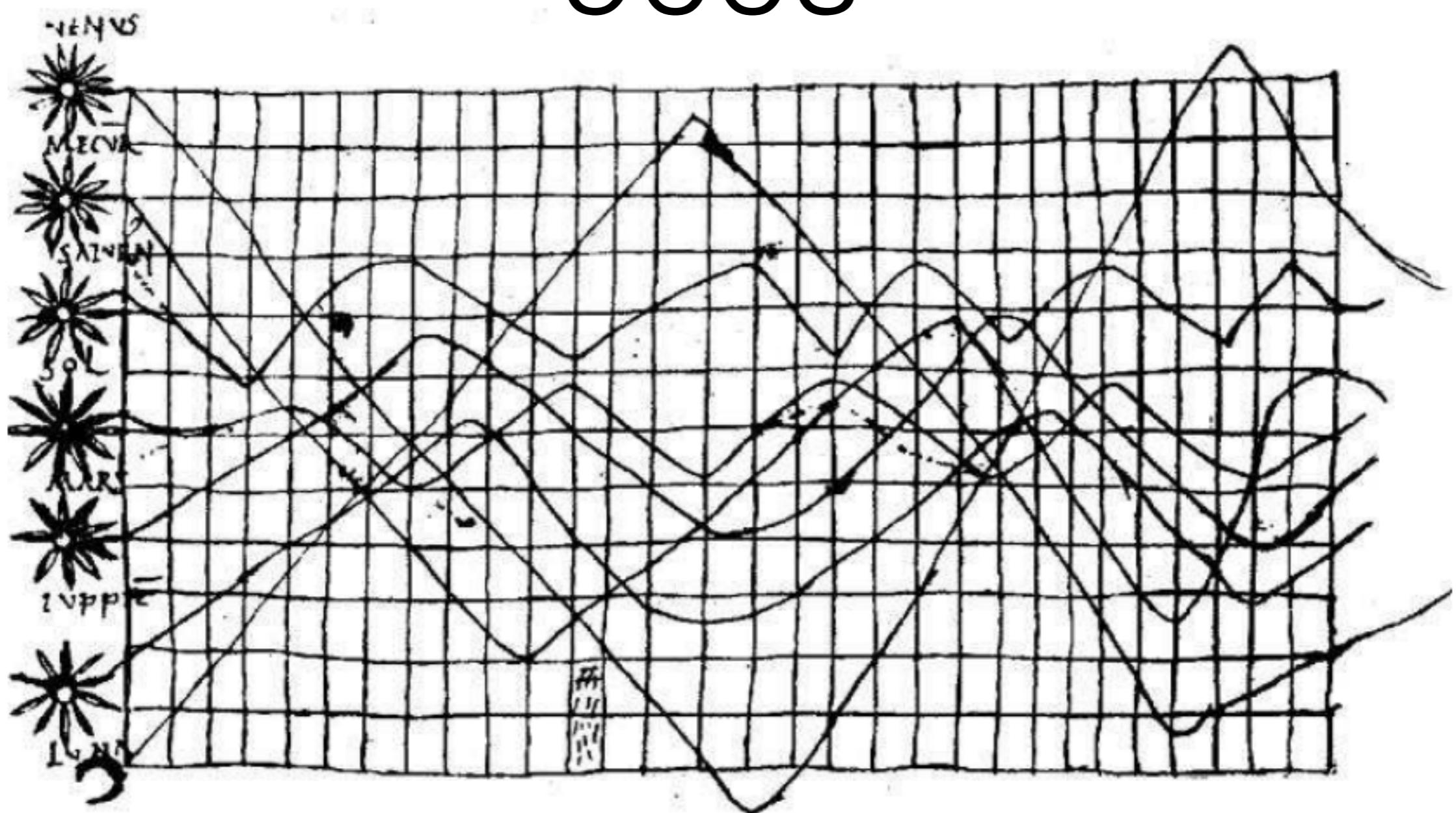
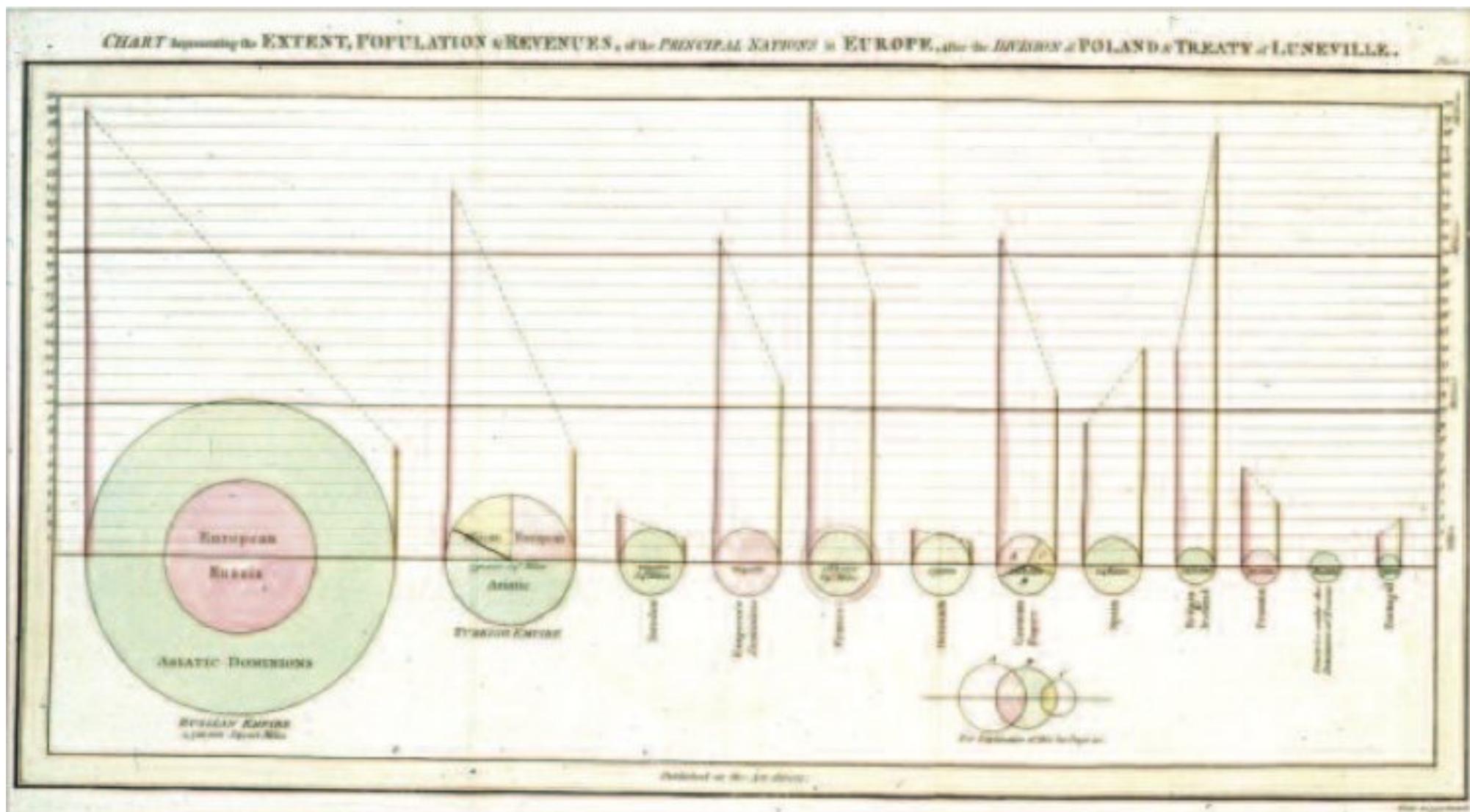


Figure 2: Planetary movements shown as cyclic inclinations over time, by an unknown astronomer, appearing in a 10th century appendix to commentaries by A. T. Macrobius on Cicero's *In Somnium Scipionis*. Source: [Funkhouser \(1936, p. 261\)](#).

William Playfair

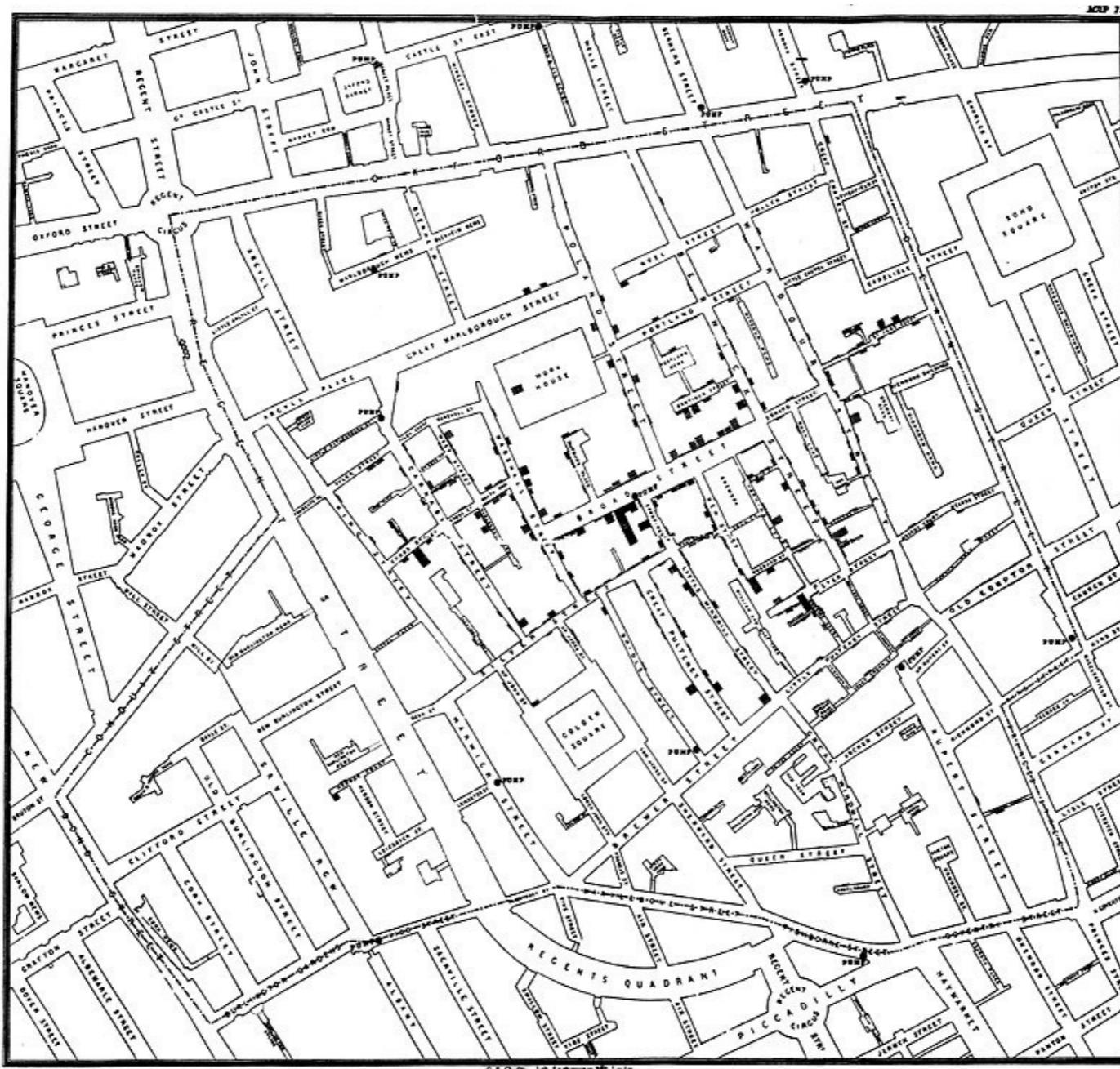
1759-1823

Credited with the invention of many common data visualizations: the pie chart, the bar chart, the line and area chart



John Snow

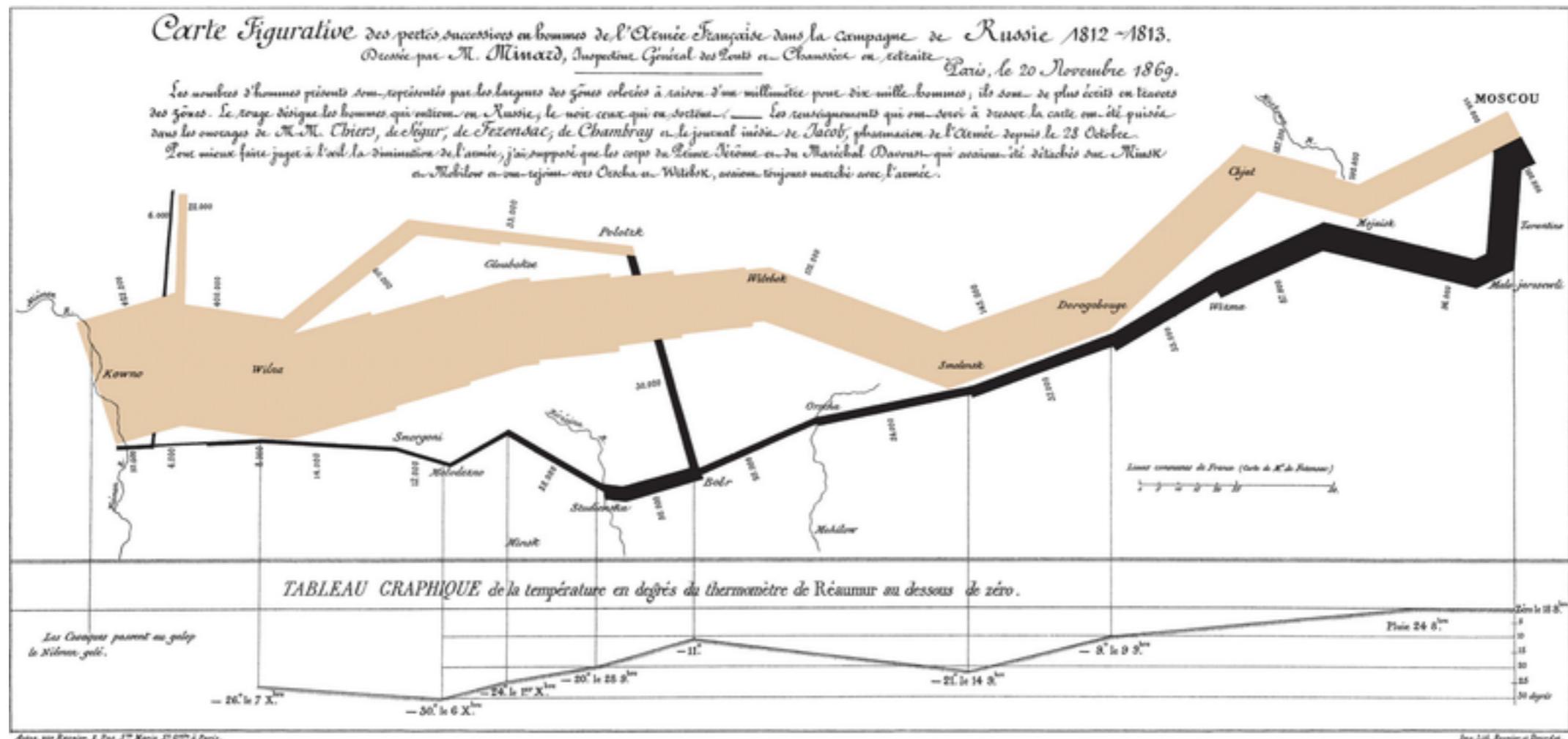
(no, not the one you're thinking about)
Used mapping to solve a cholera epidemic in London



(poor Charles Cheffins, who drew this graphic and gets no credit)

Charles Joseph Minard

1871-1870



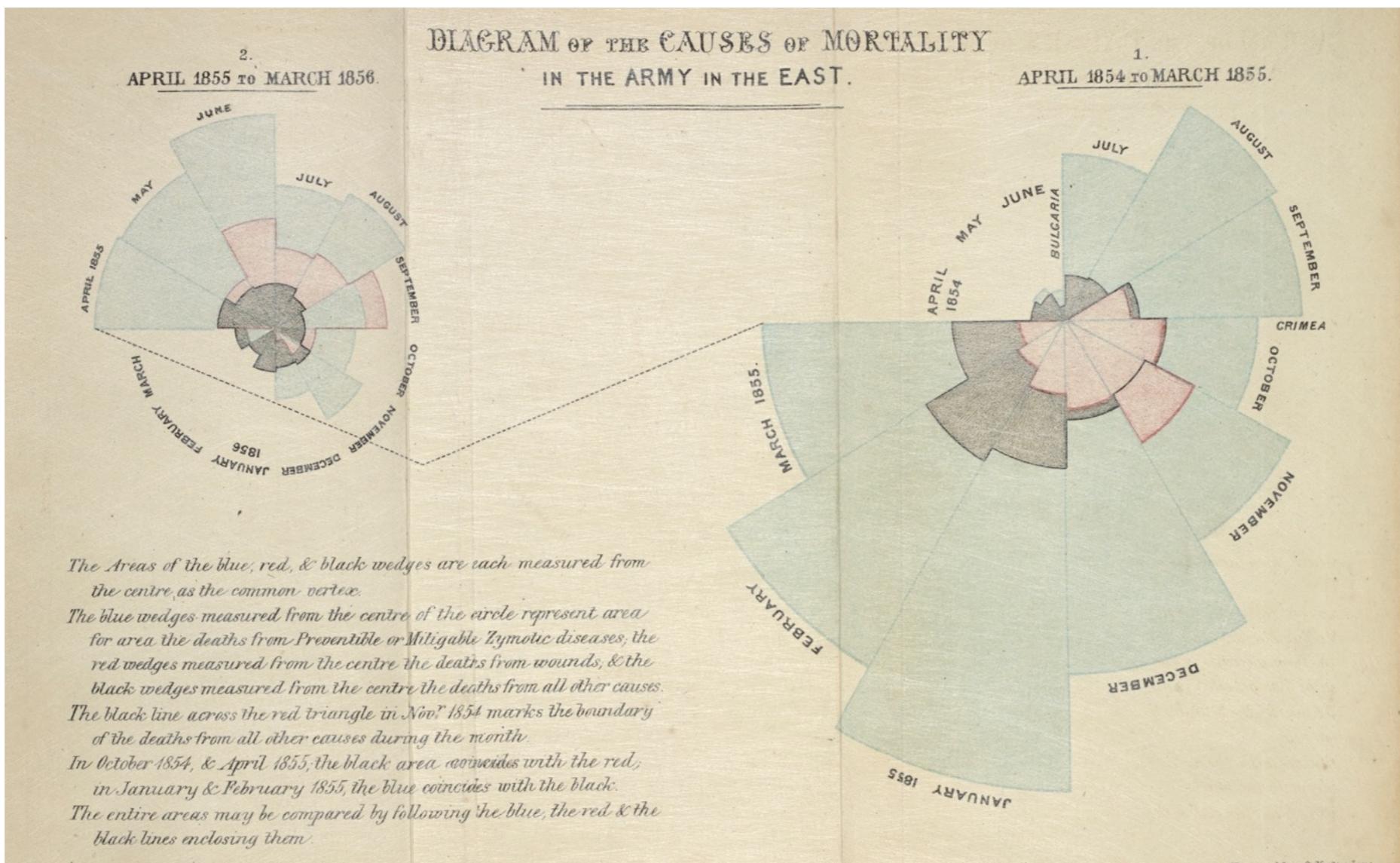
Napoleon's march on Russia

“The best statistical graphic ever drawn”?

Florence Nightingale

1820-1910

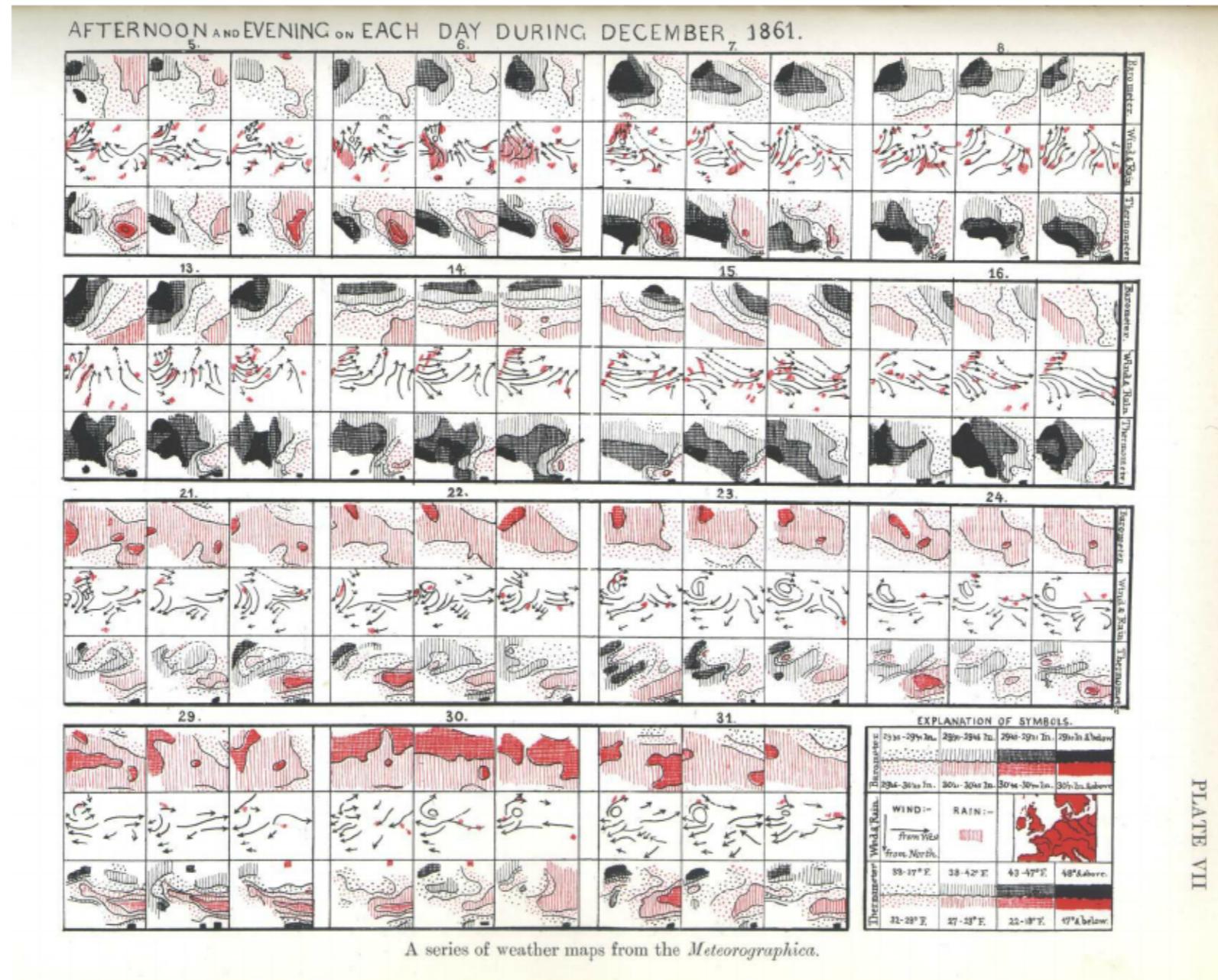
In addition to her work as a nurse, Nightingale was a statistician and invented the “coxcomb,” a variation on the pie chart



Francis Galton

1822-1911

Super-famous statistician 😊
and eugenicist 🤢



Statistical atlases

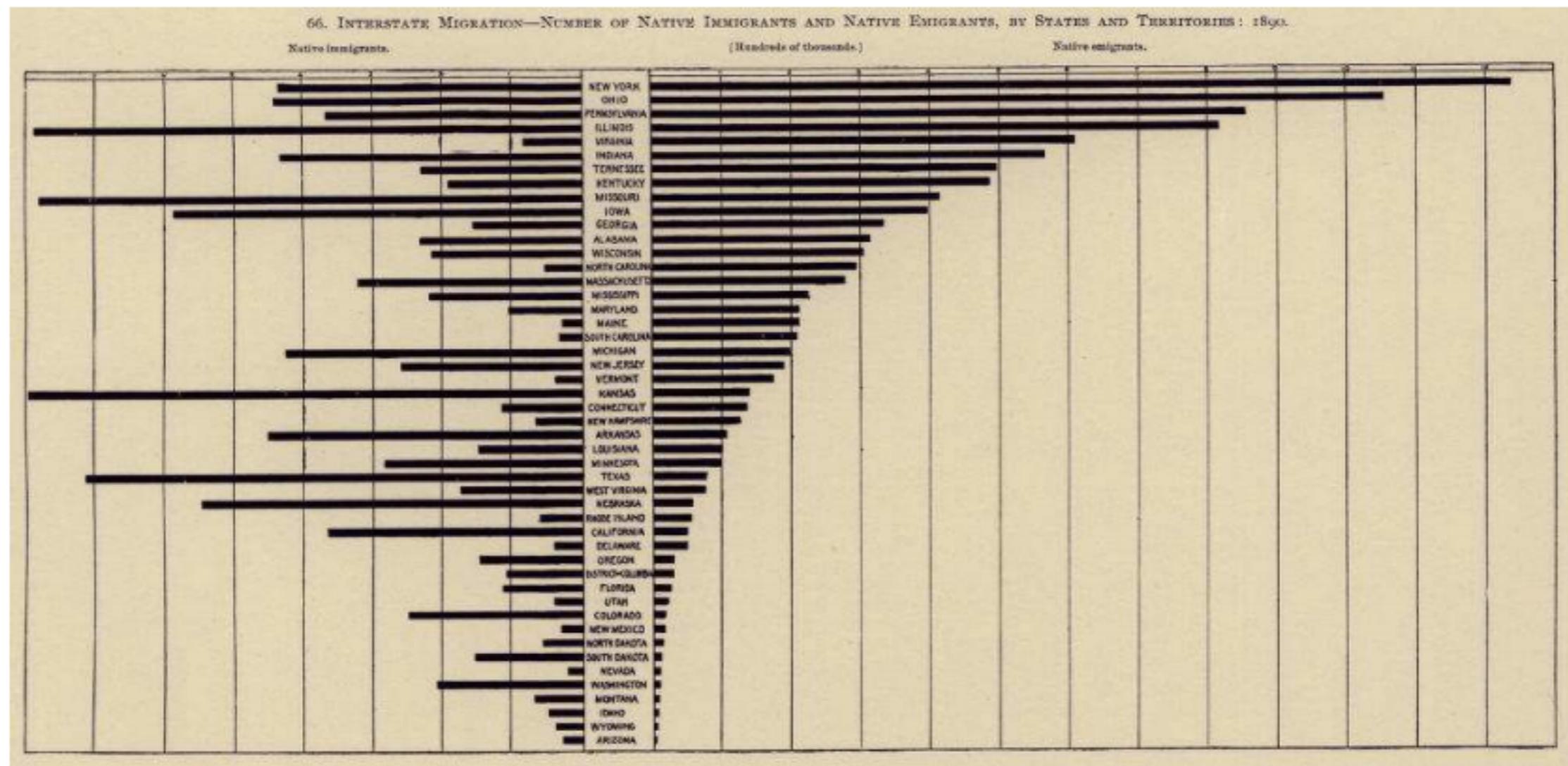


Figure 13: Interstate migration shown by back-to-back bar charts, sorted by emigration. *Source:* Statistical Atlas of the Eleventh Census, 1890, diagram 66, p. 23 (author's collection).

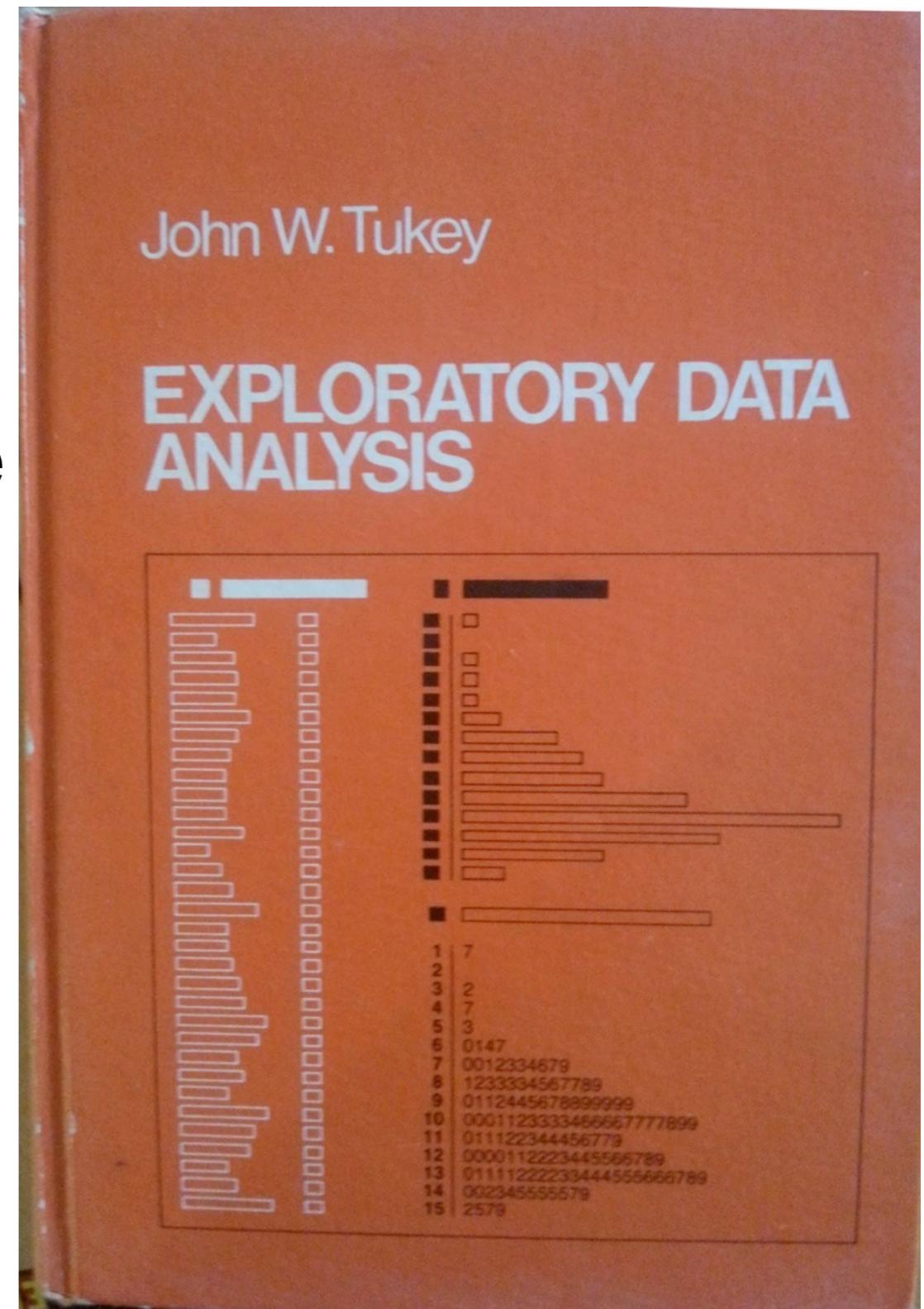
John Tukey

1915–2000



Statistician who rocked the boat

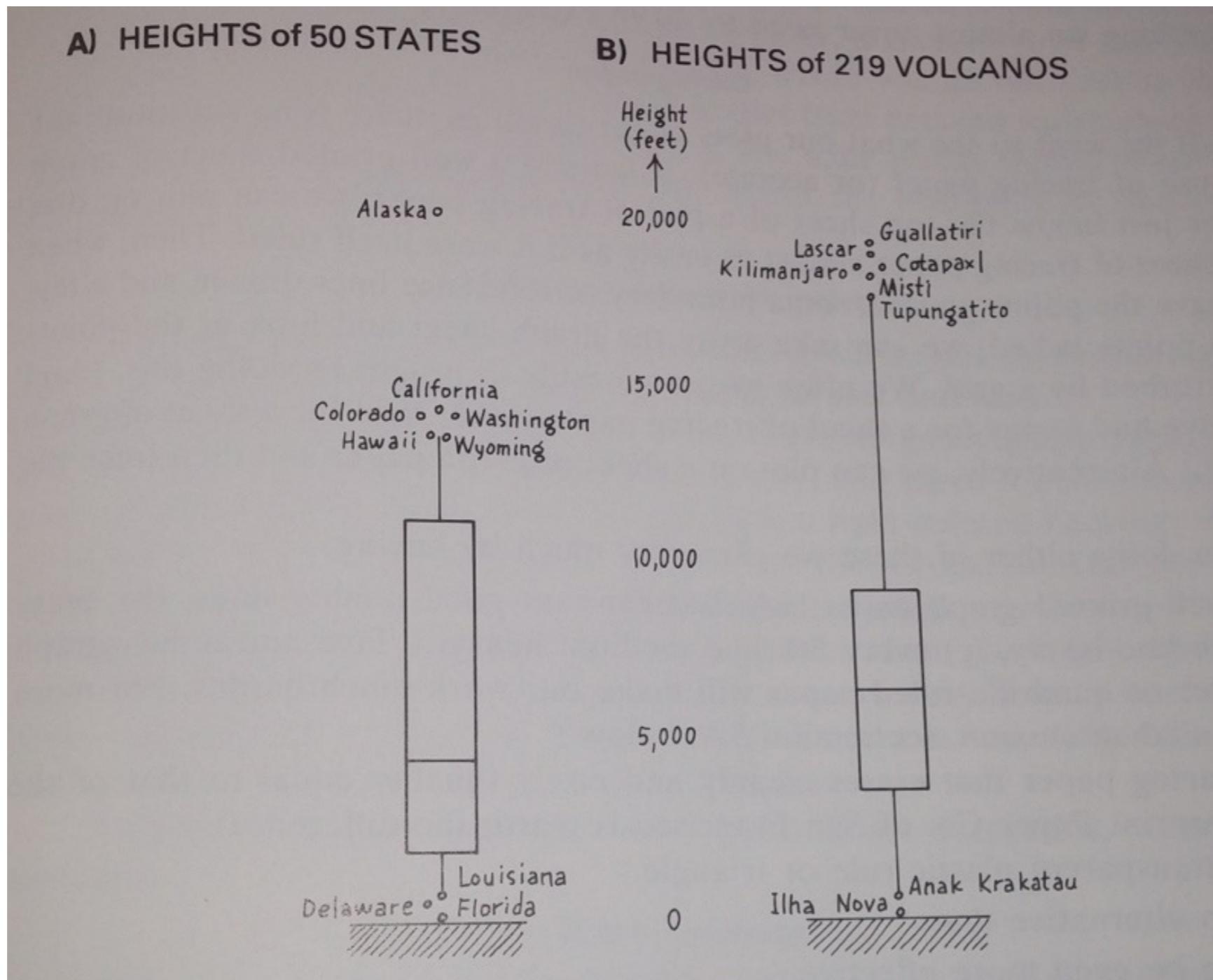
Proposed a method called
Exploratory Data Analysis (EDA),
which involves making many simple
graphs to understand data



John Tukey



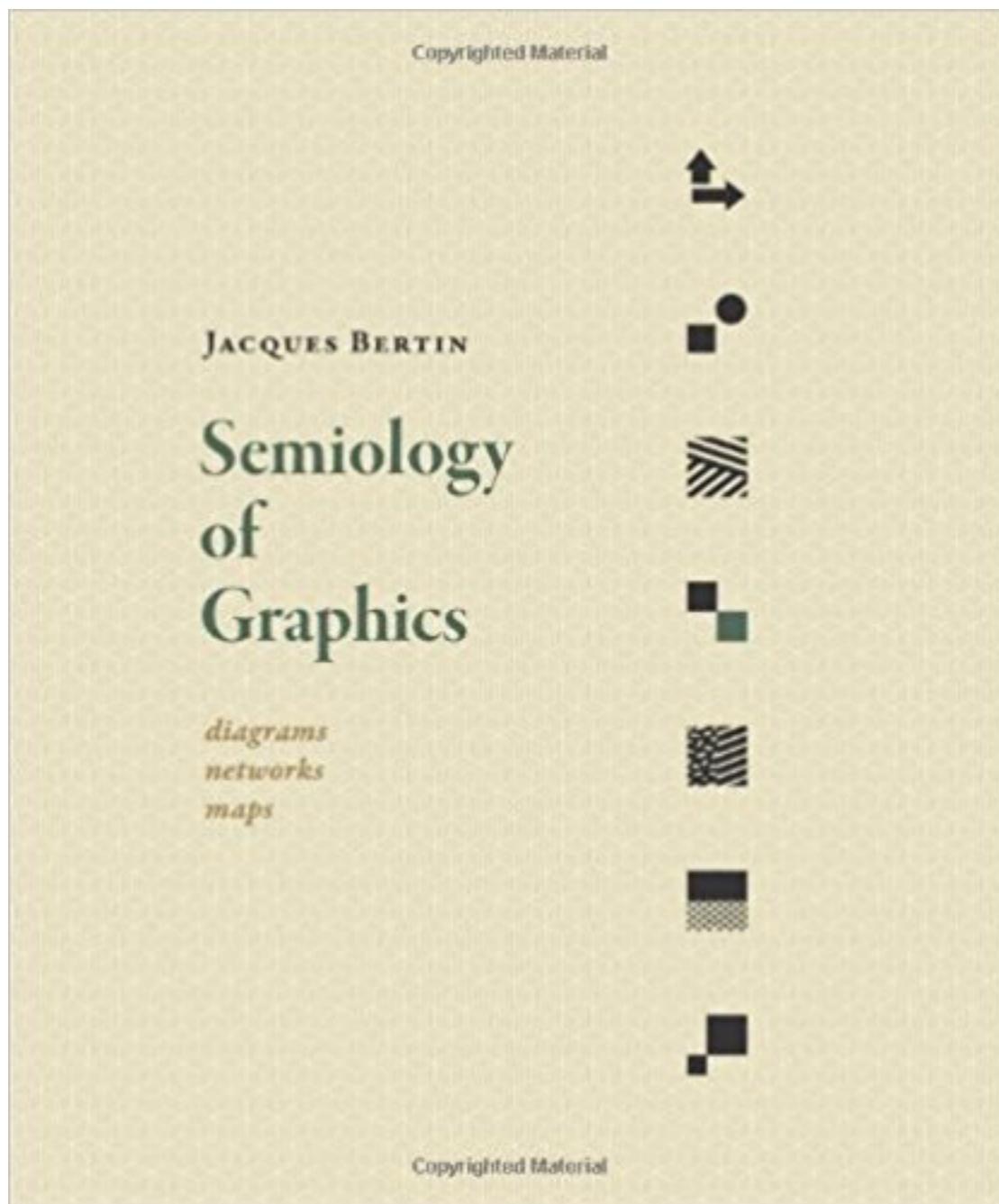
1915–2000



Jacques Bertin

1918-2010

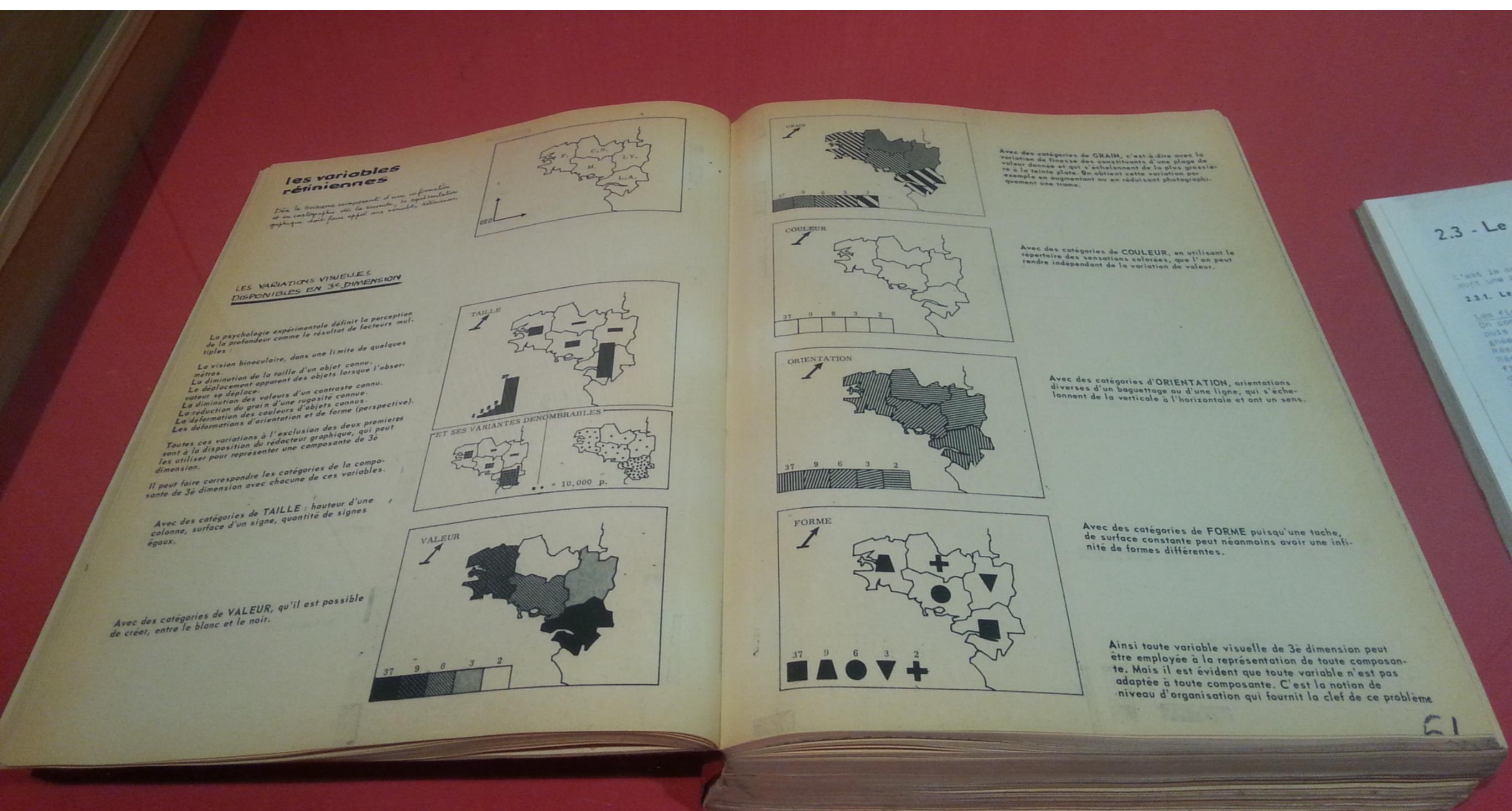
“Cartographer and theorist”



	Points	Lines	Areas	Best to show
Shape		<i>possible, but too weird to show</i>	<i>cartogram</i>	<i>qualitative differences</i>
Size			<i>cartogram</i>	<i>quantitative differences</i>
Color Hue				<i>qualitative differences</i>
Color Value				<i>quantitative differences</i>
Color Intensity				<i>qualitative differences</i>
Texture				<i>qualitative & quantitative differences</i>

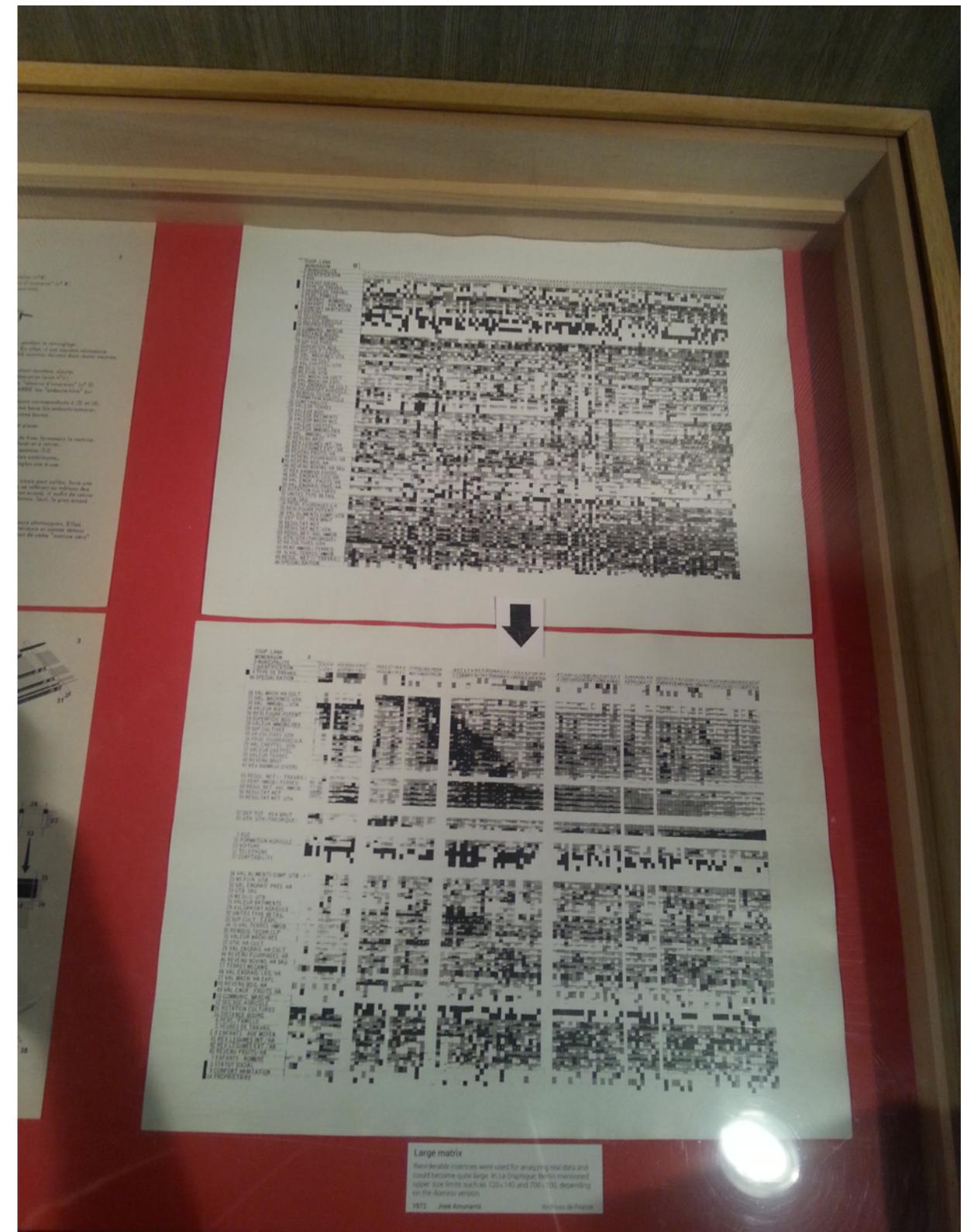
Jacques Bertin

1918-2010



Jacques Bertin

Bertin matrices



William Cleveland

1943-

Professor of statistics at
Purdue
Did famous research about
effectiveness of visualizations

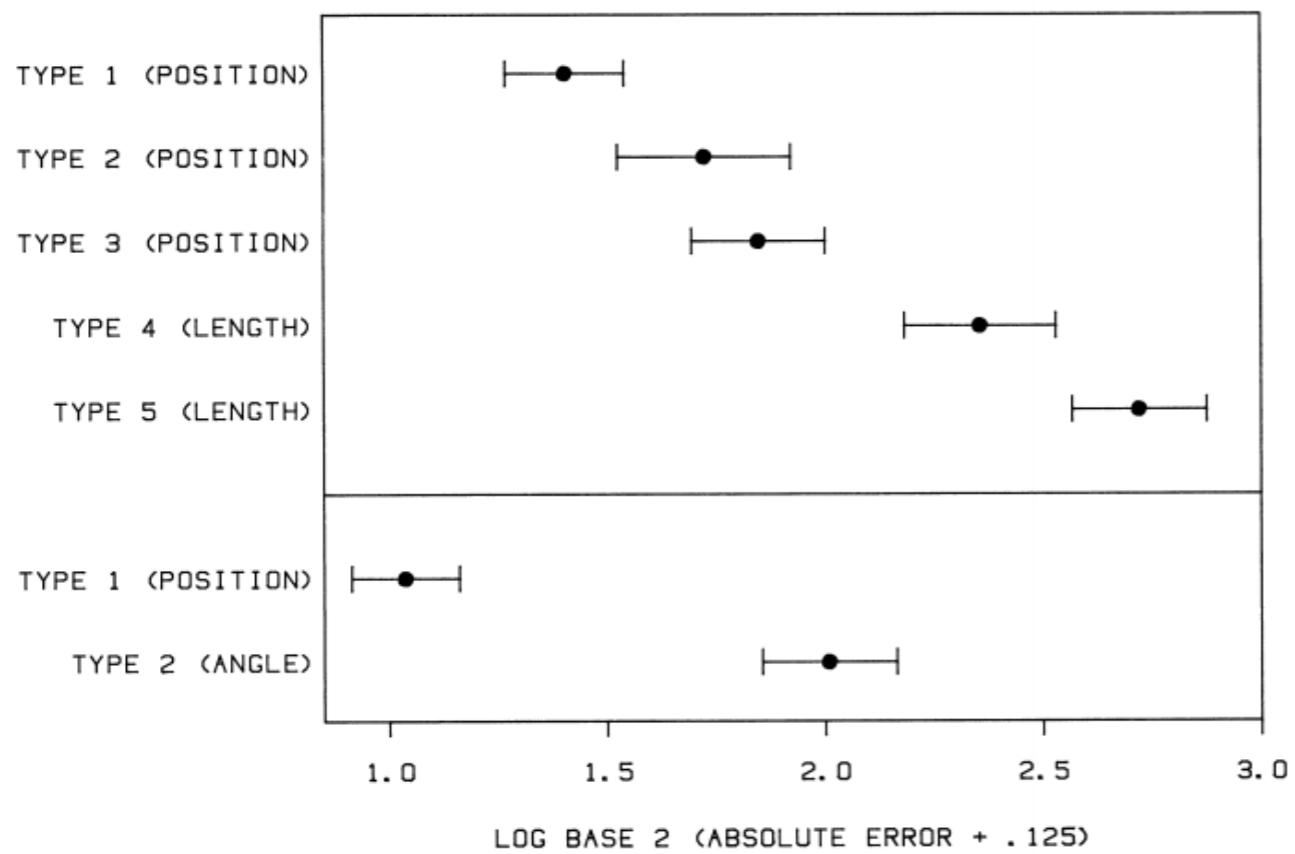
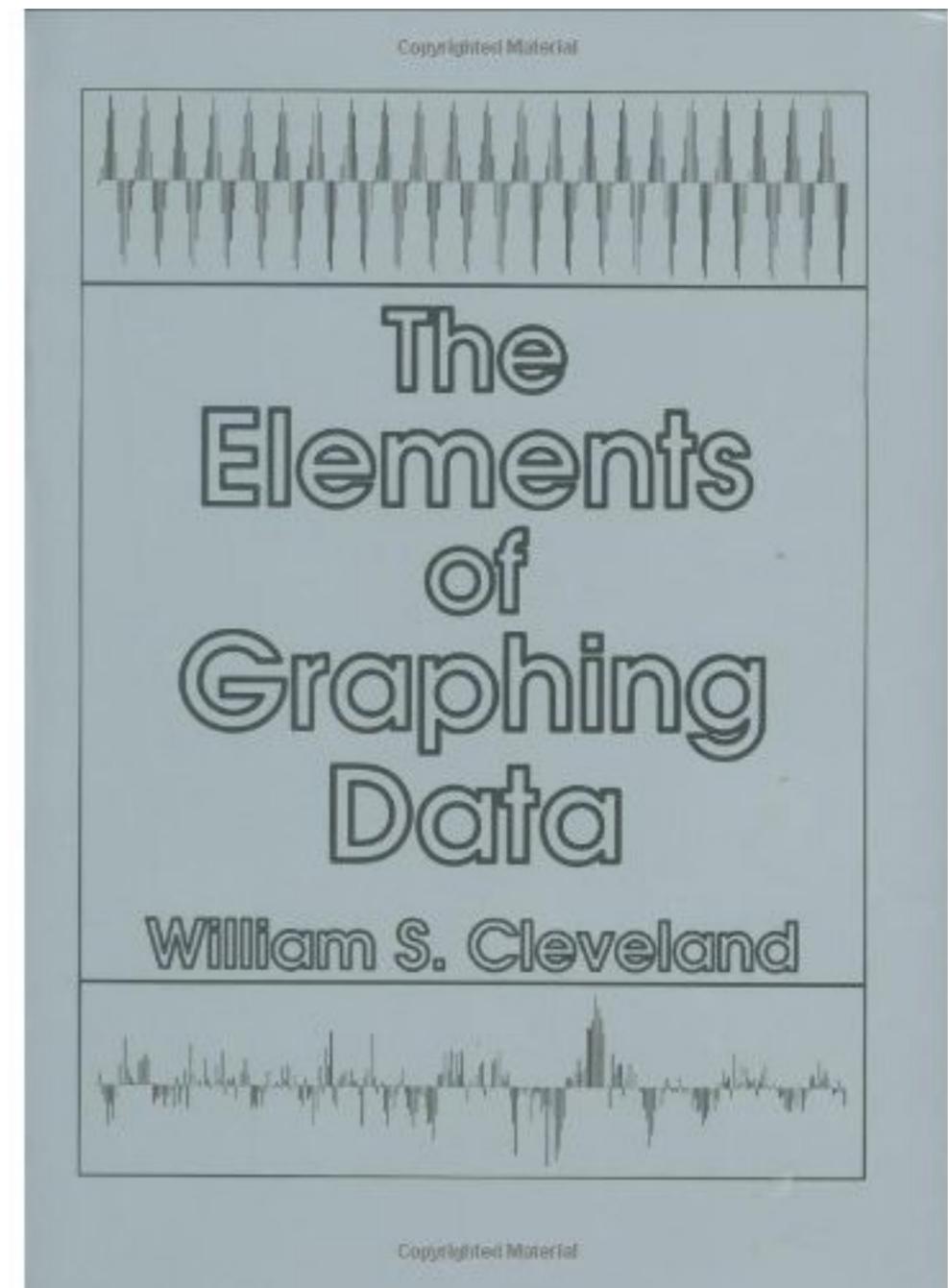
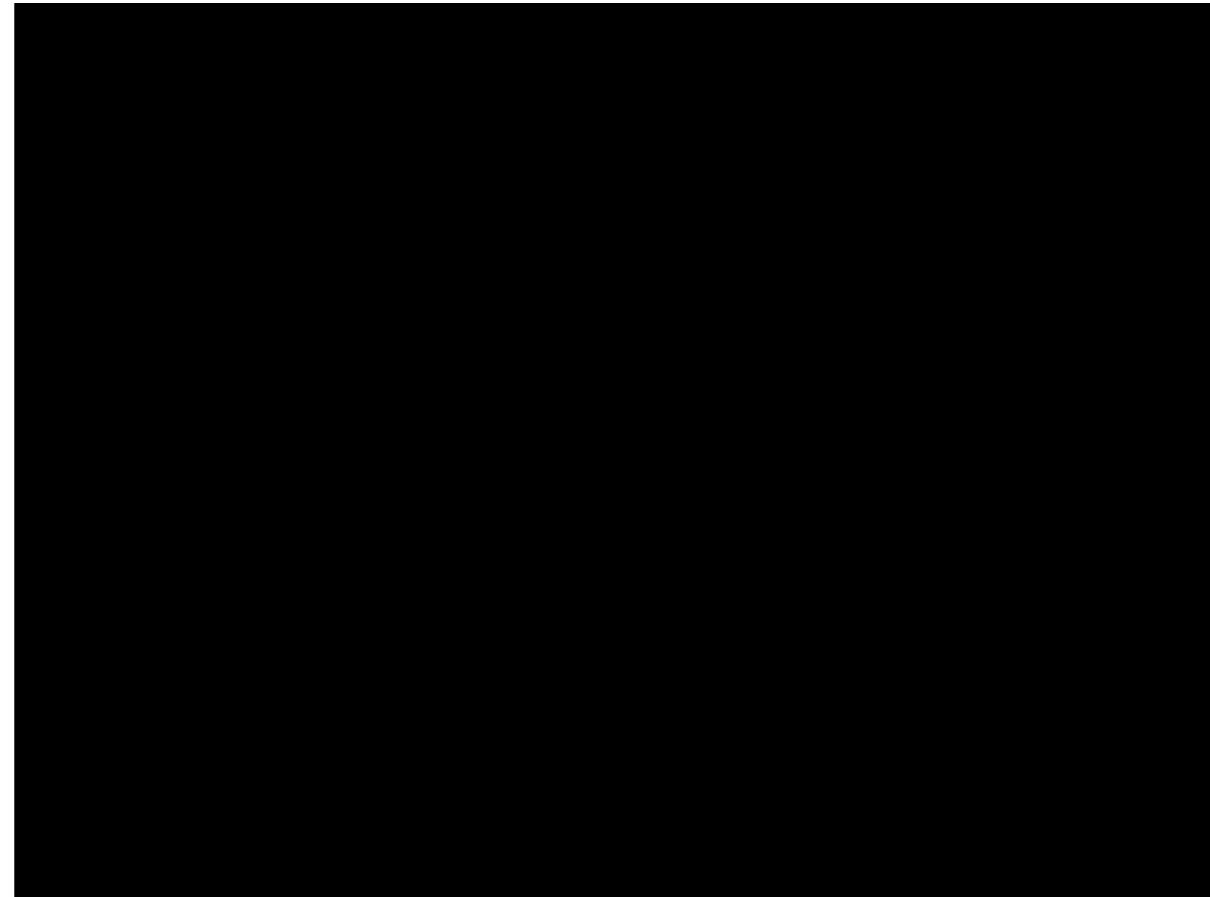


Figure 16. Log absolute error means and 95% confidence intervals for judgment types in position-length experiment (top) and position-angle experiment (bottom).

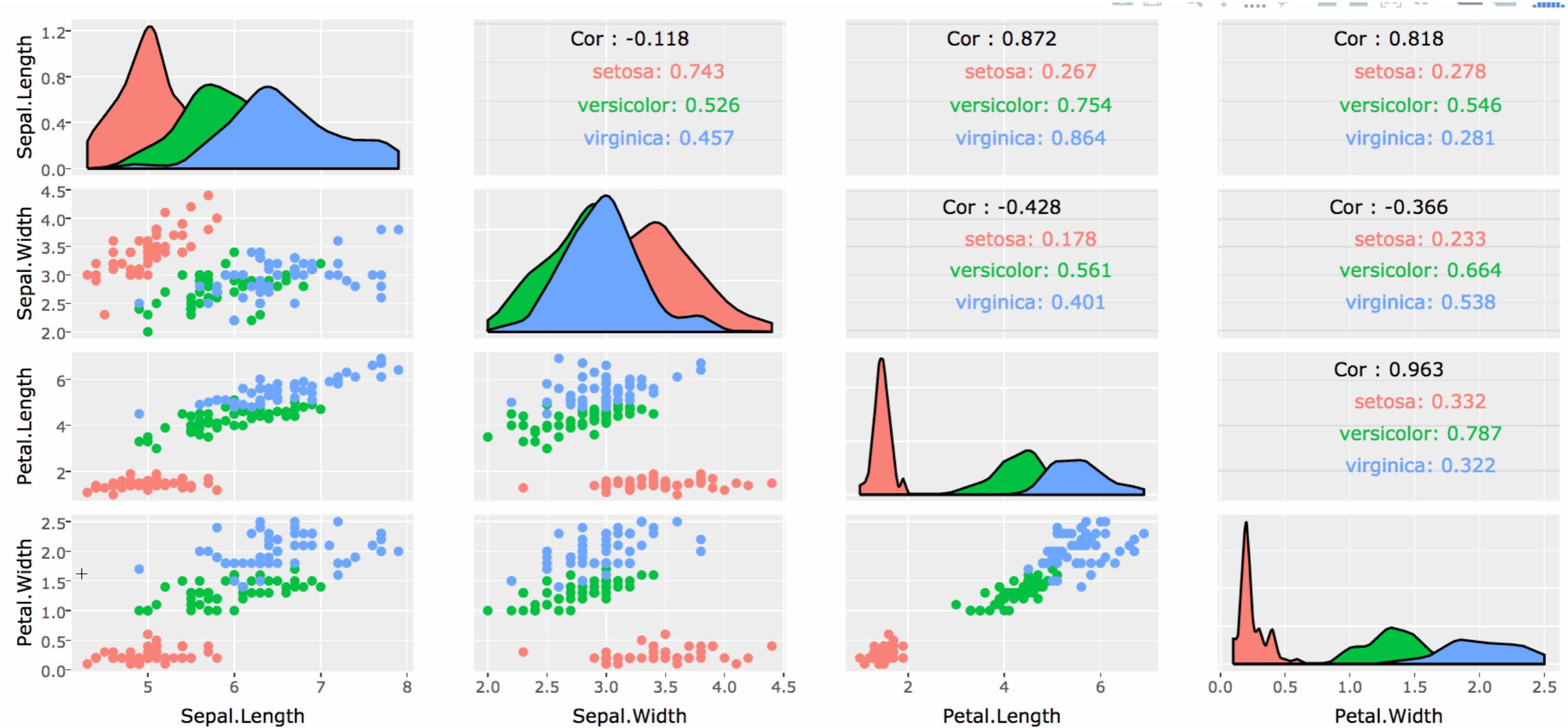


John Tukey

1915–2000



Interactivity, brushing and linking



Luke Tierney

xlisp-stat

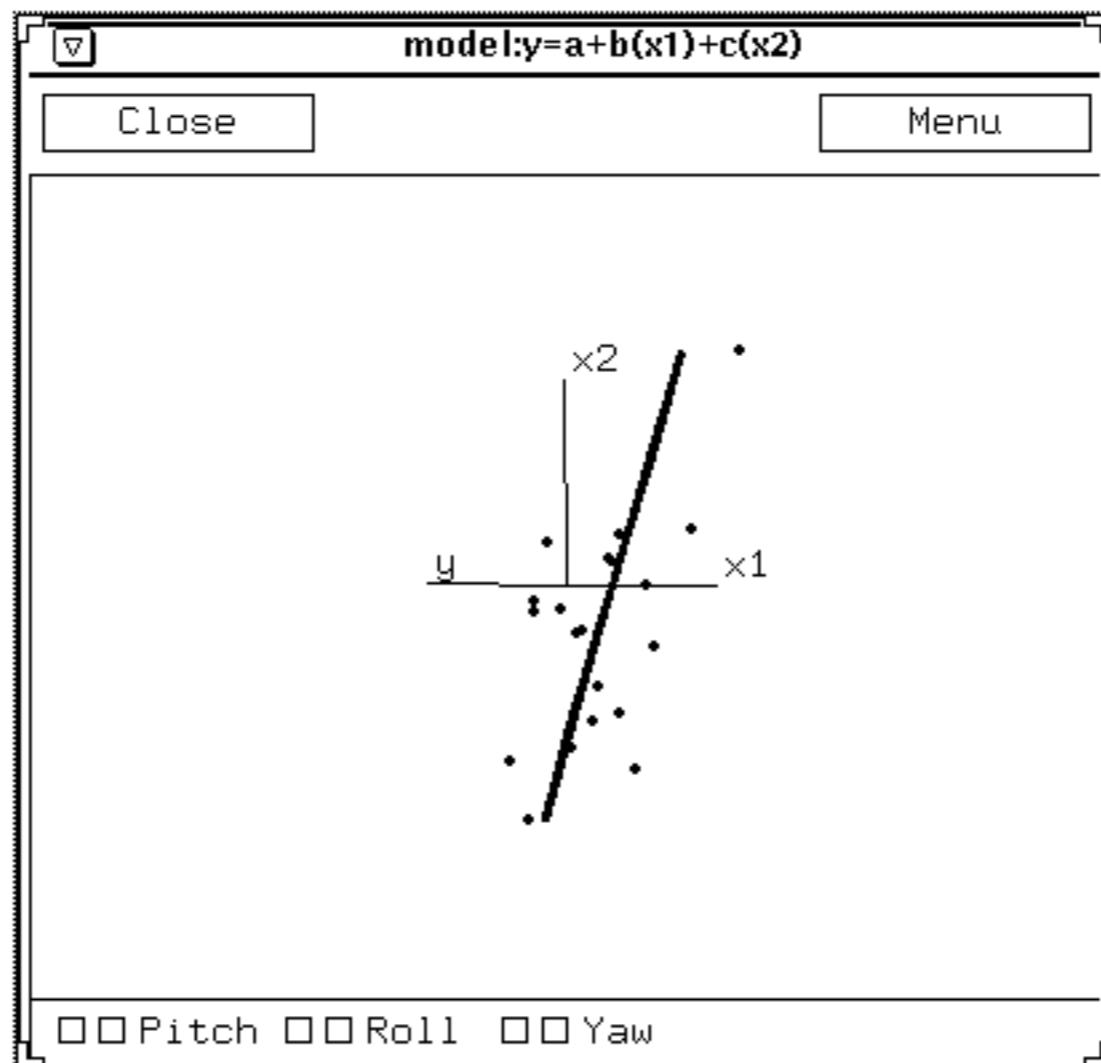
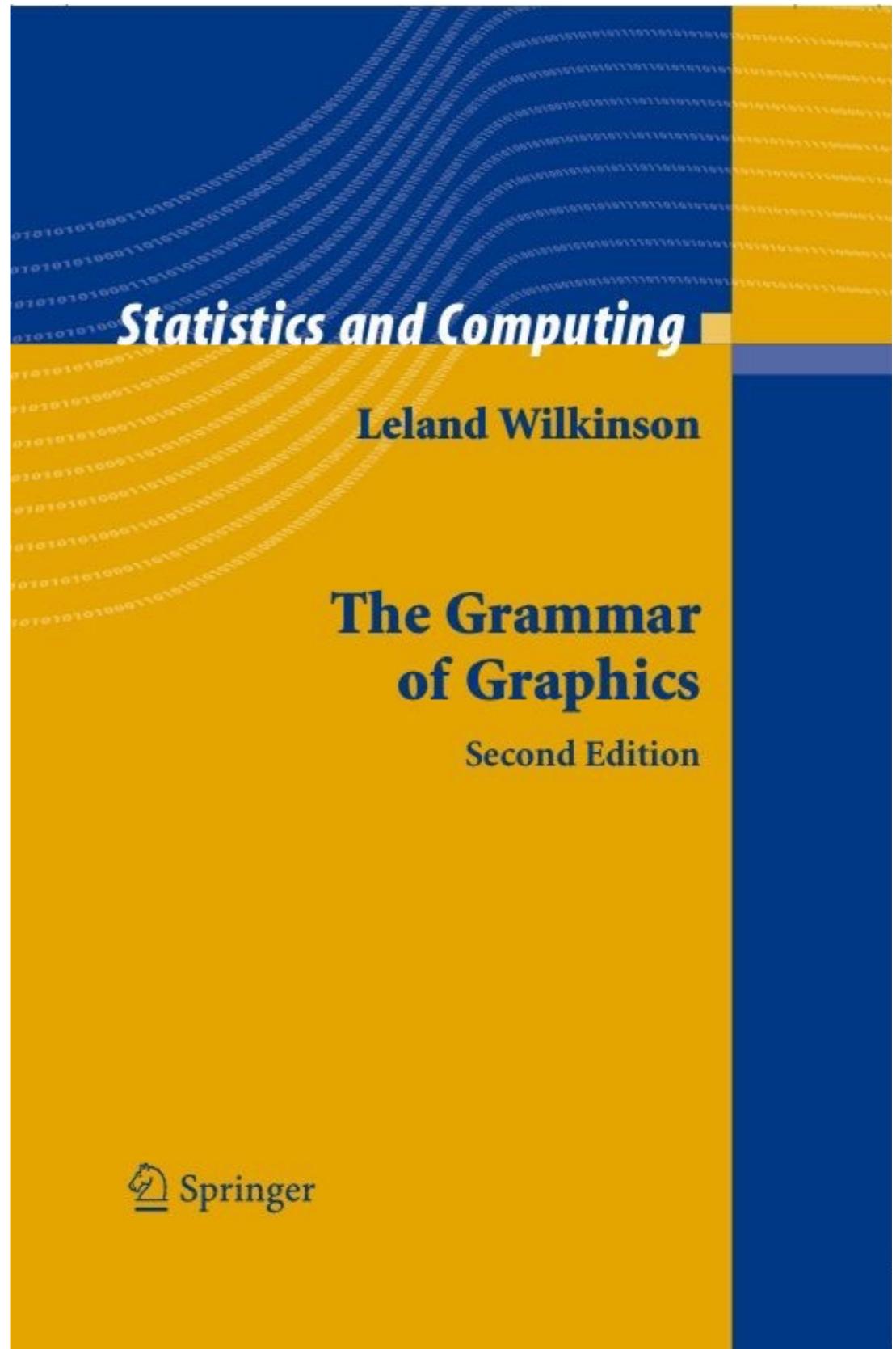


Figure 7a

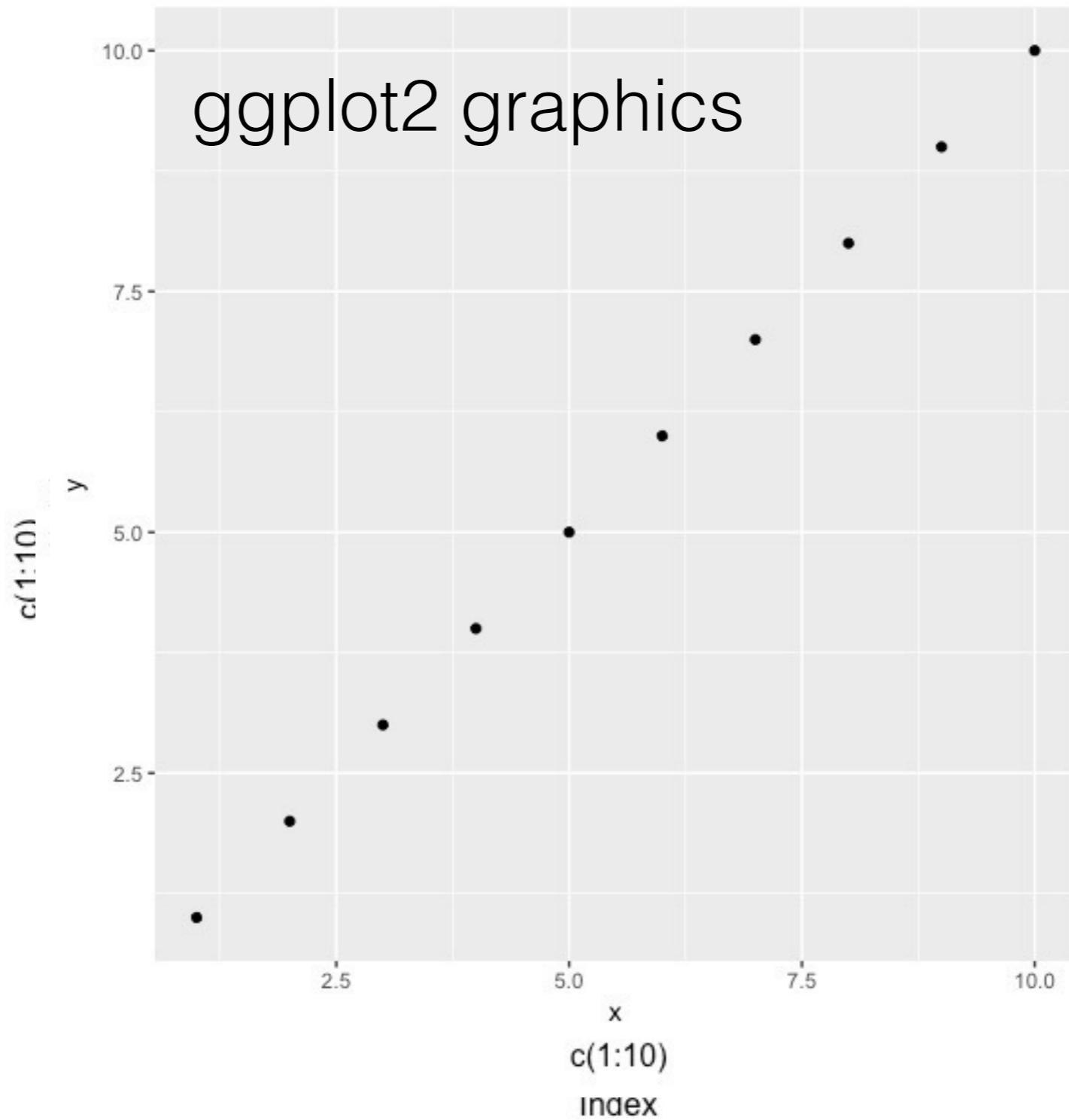
Leland Wilkinson

Statistician and software designer

Worked on SYSTAT, SPSS, Tableau, now H2O.ai



R



Hadley Wickham

Famous R programmer

Dissertation: “Practical tools
for exploring data and models”

Advisors: Di Cook and
Heike Hoffman



Mike Bostock

d3.js

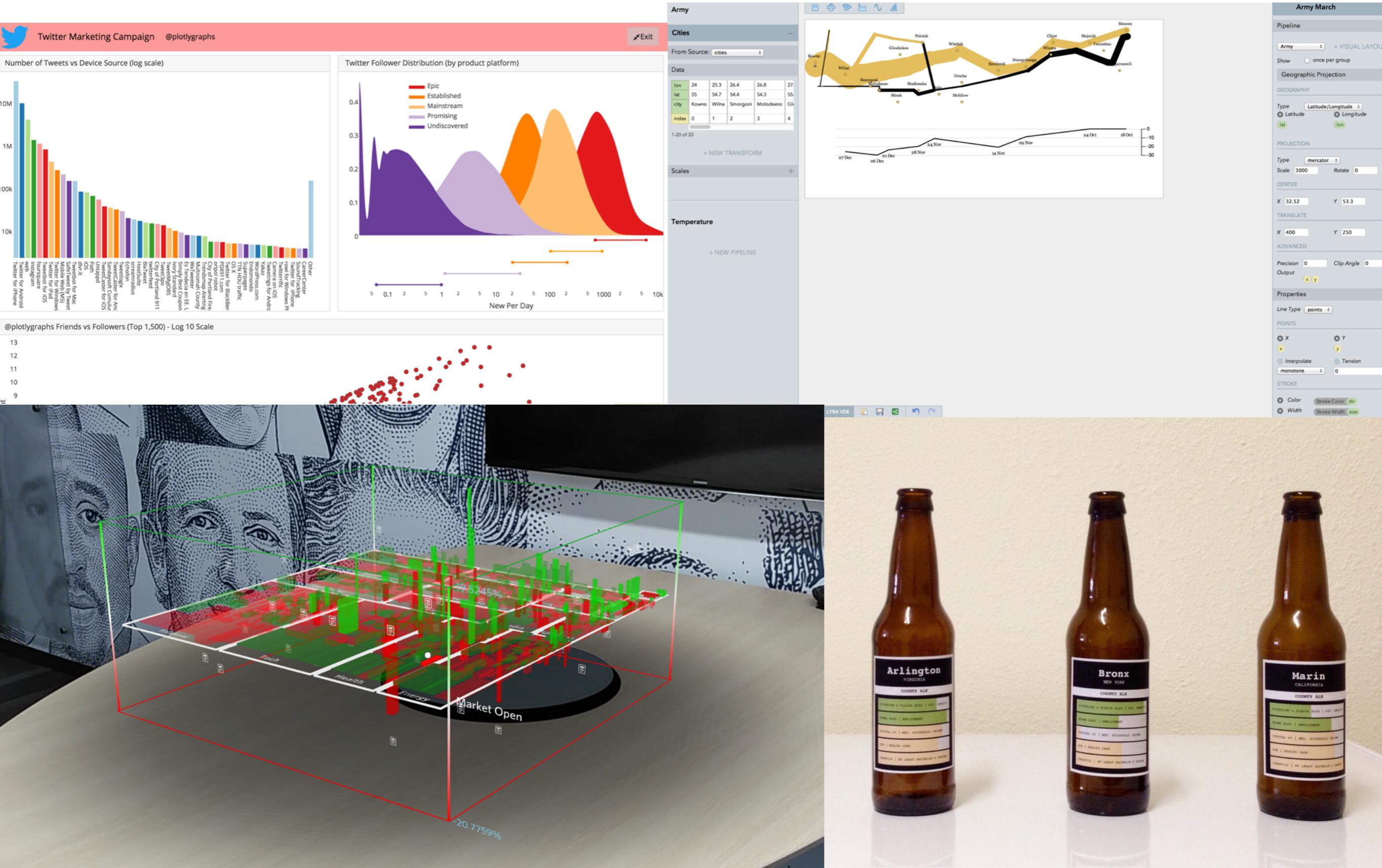
[Overview](#) [Examples](#) [Documentation](#) [Source](#)



Data-Driven Documents



Things are accelerating!



Before next class