

intro to viz

What is visualization?



visualization



All News Videos Books More

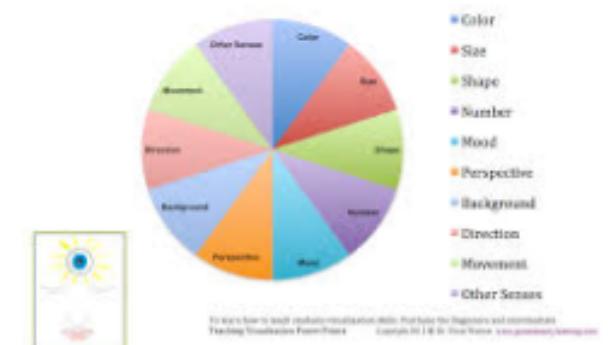
Settings Tools

View saved SafeSearch ▾

process math success relationship progress connection comparison time art power security color taste money network graph time series correlation

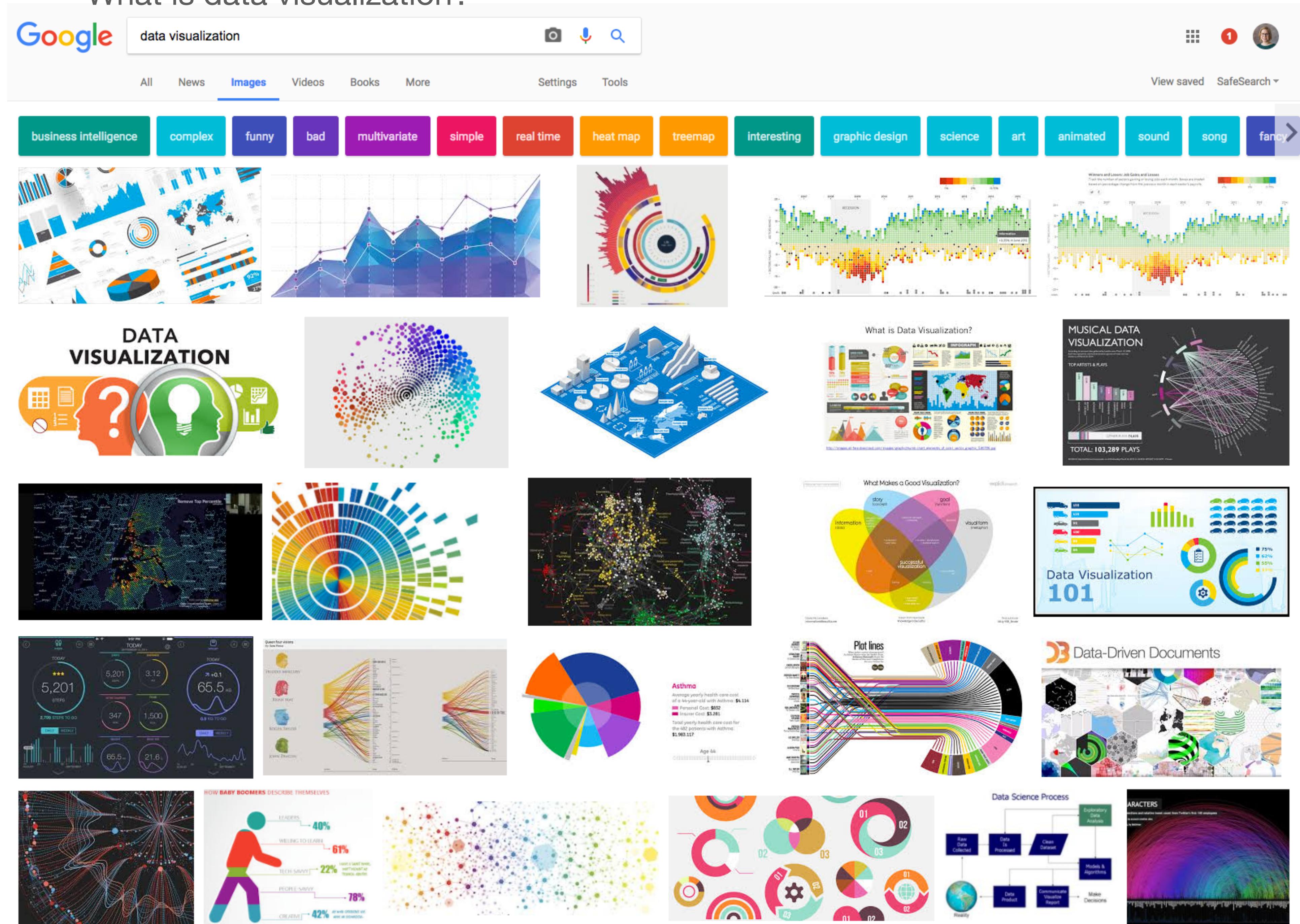


There are Ten Visualization Skills



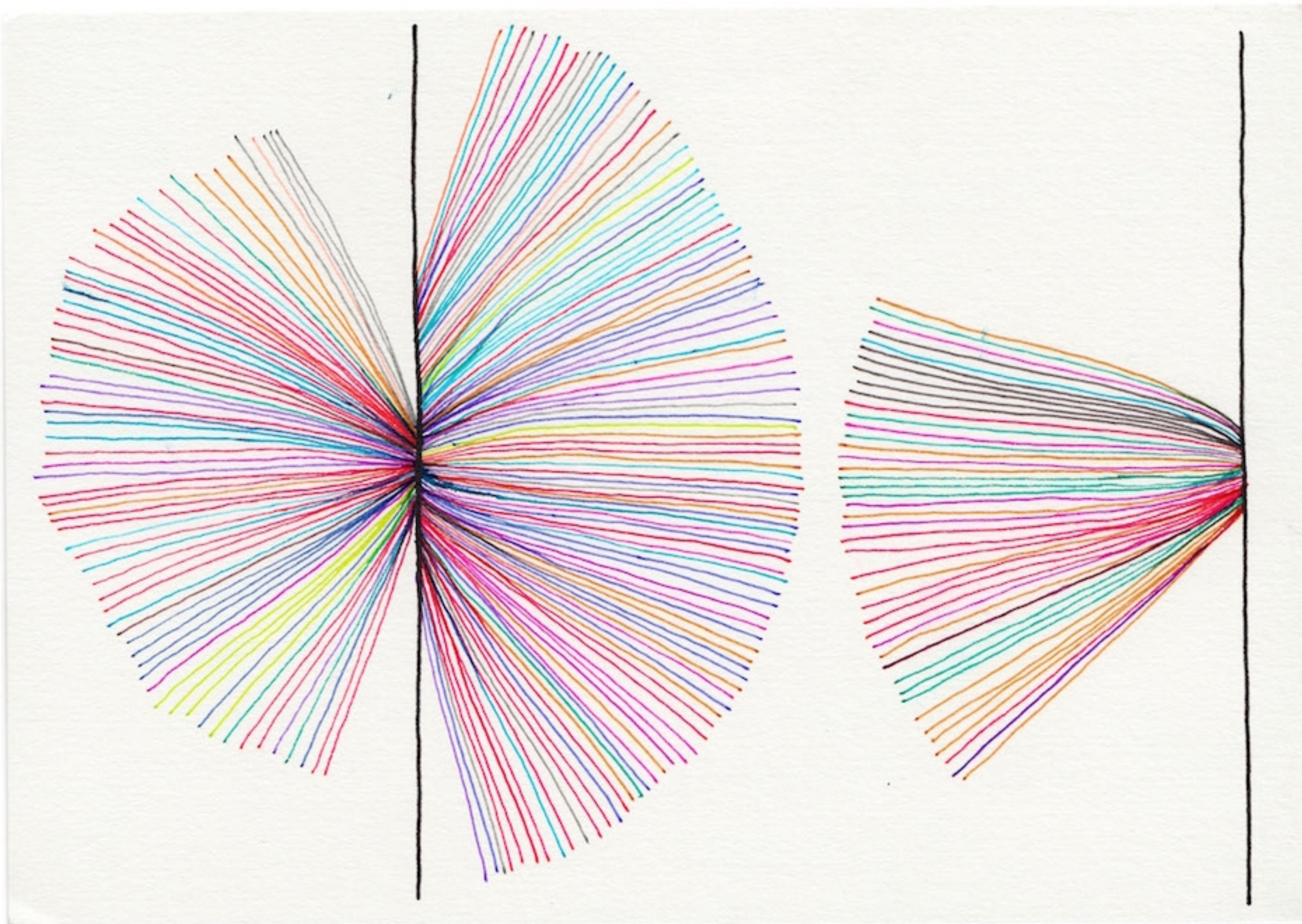
VISUALIZE Responsibly

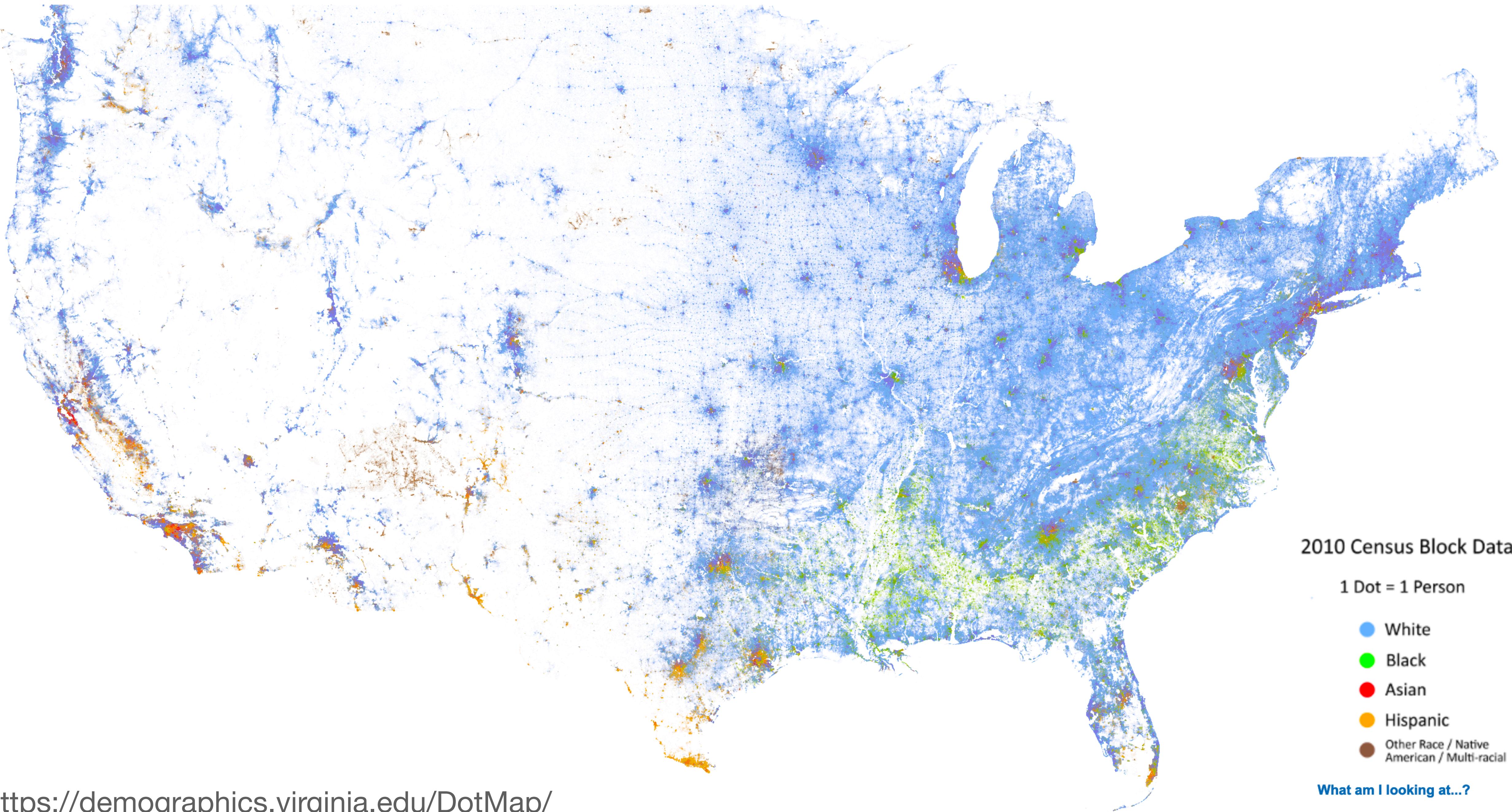
What is data visualization?



We will also consider
information visualization
infographics
statistical graphics
and more

A taste of data visualization





U.S. GUN DEATHS IN

2013 2010

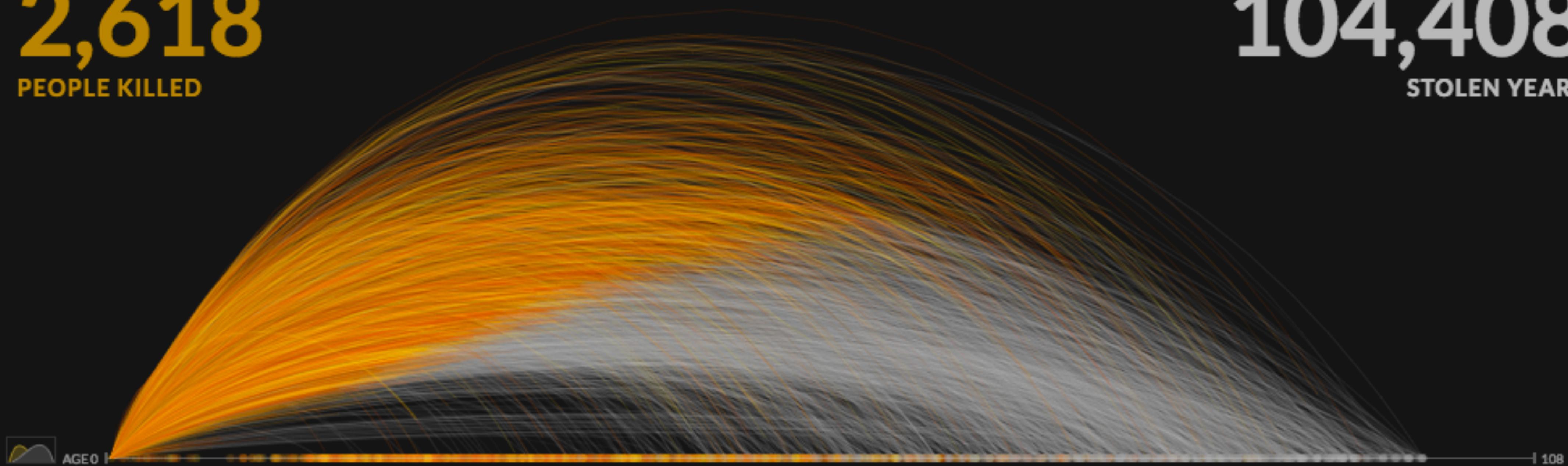
APRIL

2,618

PEOPLE KILLED

104,408

STOLEN YEARS



Recall: tidy data

country	year	cases	pop
Afghanistan	1990	745	1012731
Afghanistan	2000	666	2012520
Afghanistan	2010	112	3112522
Afghanistan	2011	112	3112523
Afghanistan	2012	112	3112522
Afghanistan	2013	112	3112523
Afghanistan	2014	112	3112522
Afghanistan	2015	112	3112523
Afghanistan	2016	112	3112522
Afghanistan	2017	112	3112523
Afghanistan	2018	112	3112522
Afghanistan	2019	112	3112523
Afghanistan	2020	43700	12072363
India	2020	43700	12072363

A data set is **tidy** iff:

1. Each **variable** is in its own **column**
2. Each **case** is in its own **row**
3. Each **value** is in its own **cell**

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

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

Some history

15,000 B.C.

Laxcaux, France

cave paintings



900s

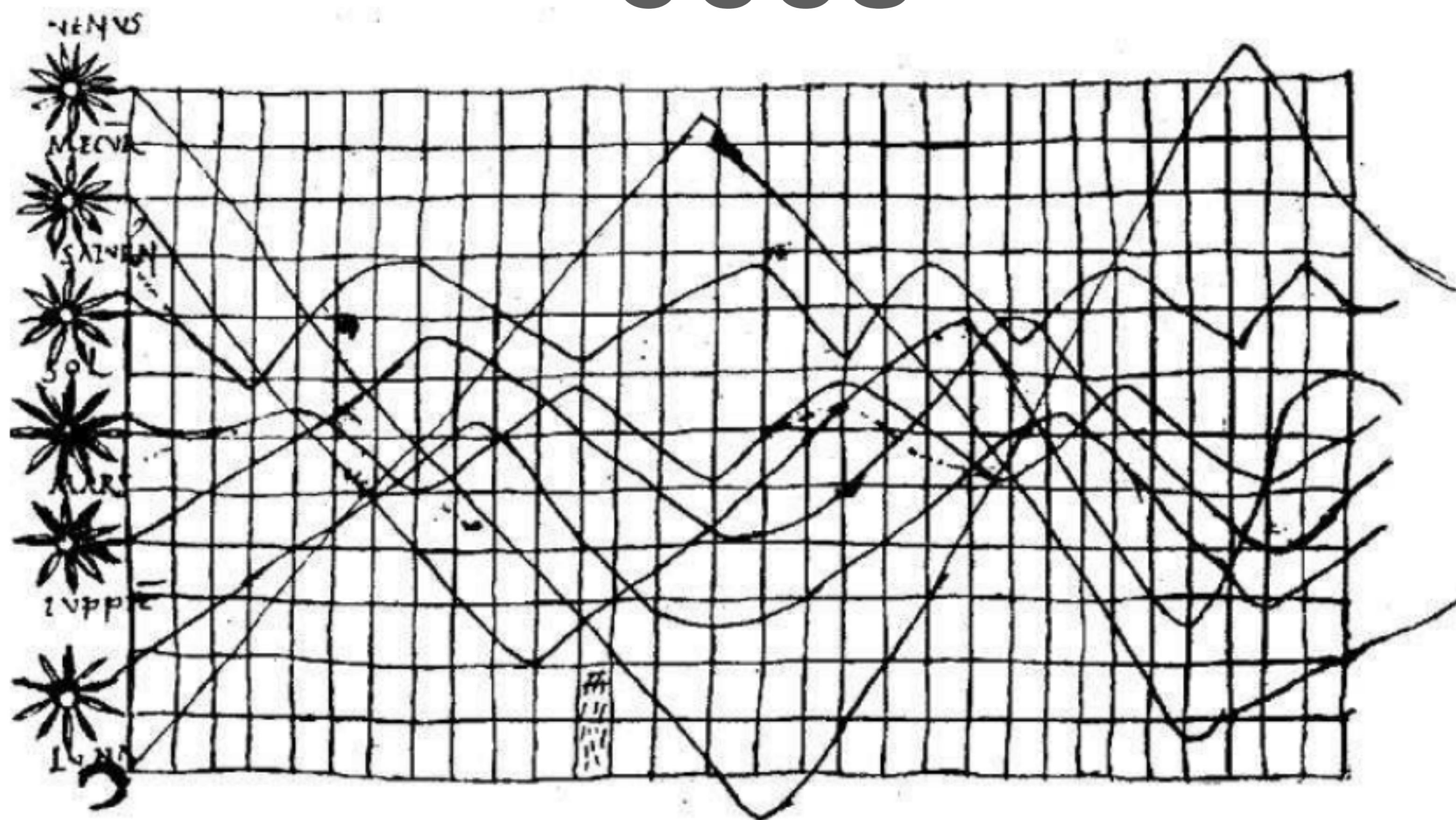
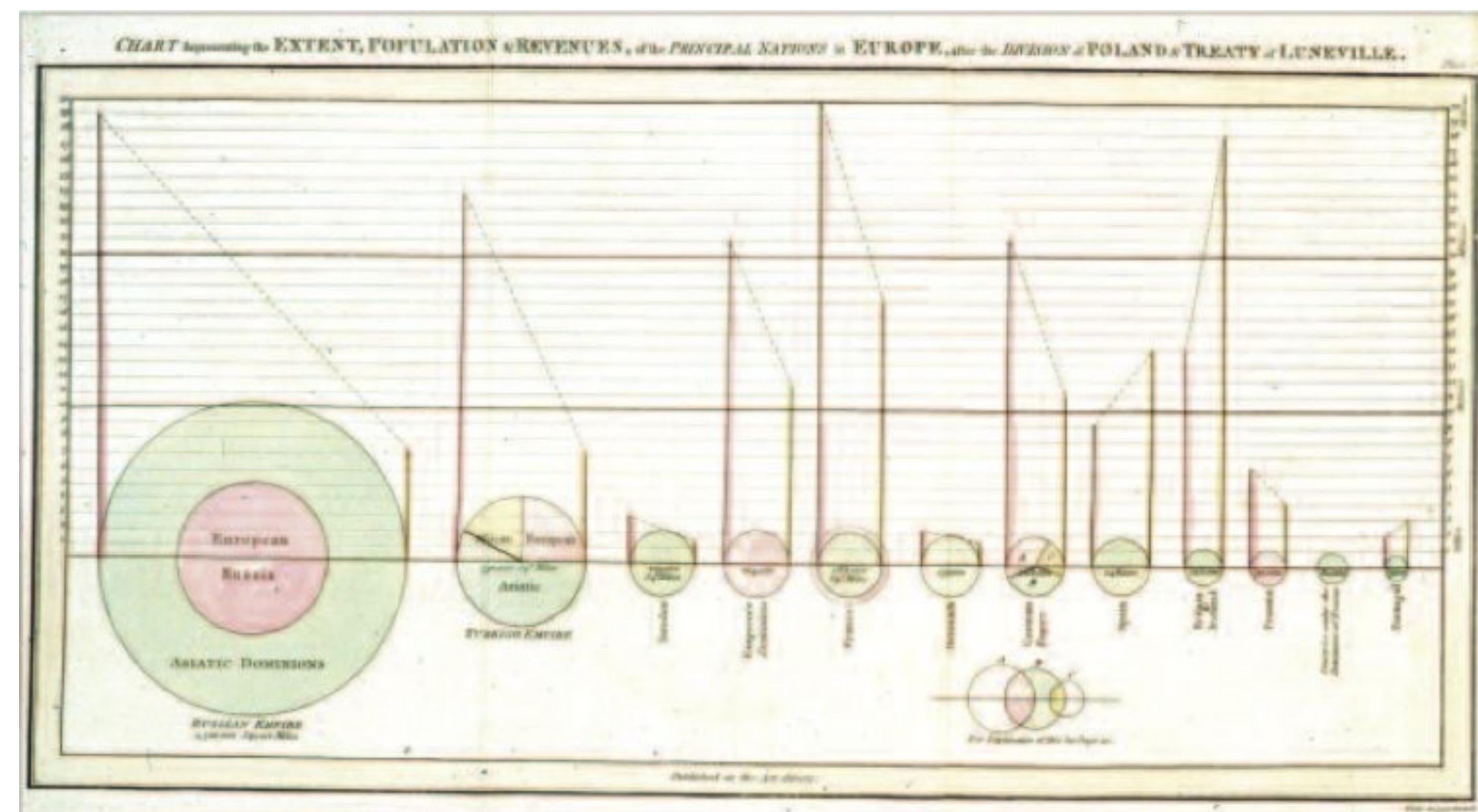


Figure 2: Planetary movements shown as cyclic inclinations over time, by an unknown astronomer, appearing in a 10th century appendix to commentaries by A. T. Macrobius on Cicero's *In Somnium Scipionis*. Source: [Funkhouser \(1936, p. 261\)](#).

1759-1823

William Playfair

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

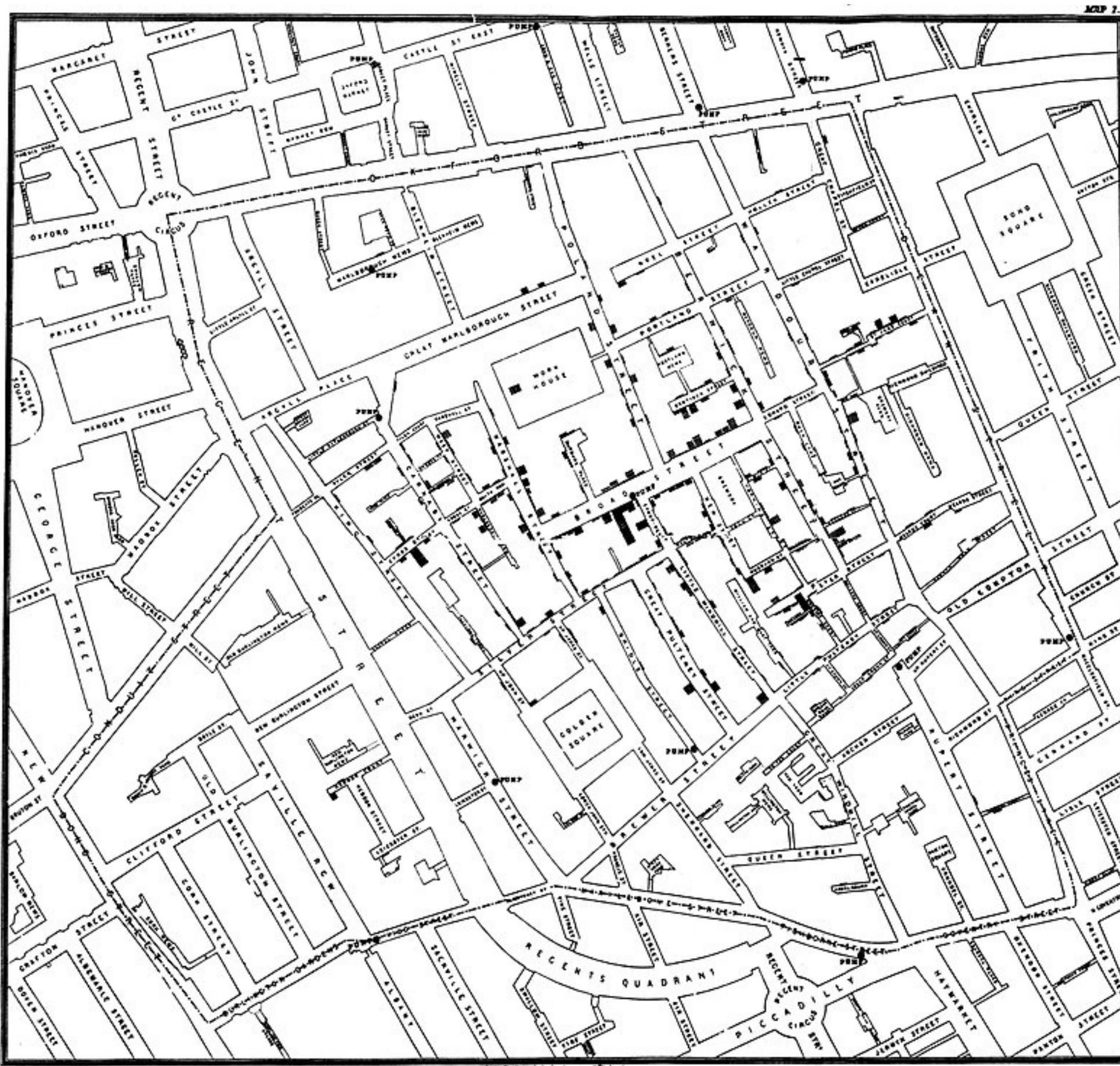


1813-1858

John Snow

(no, not the one you're thinking about)

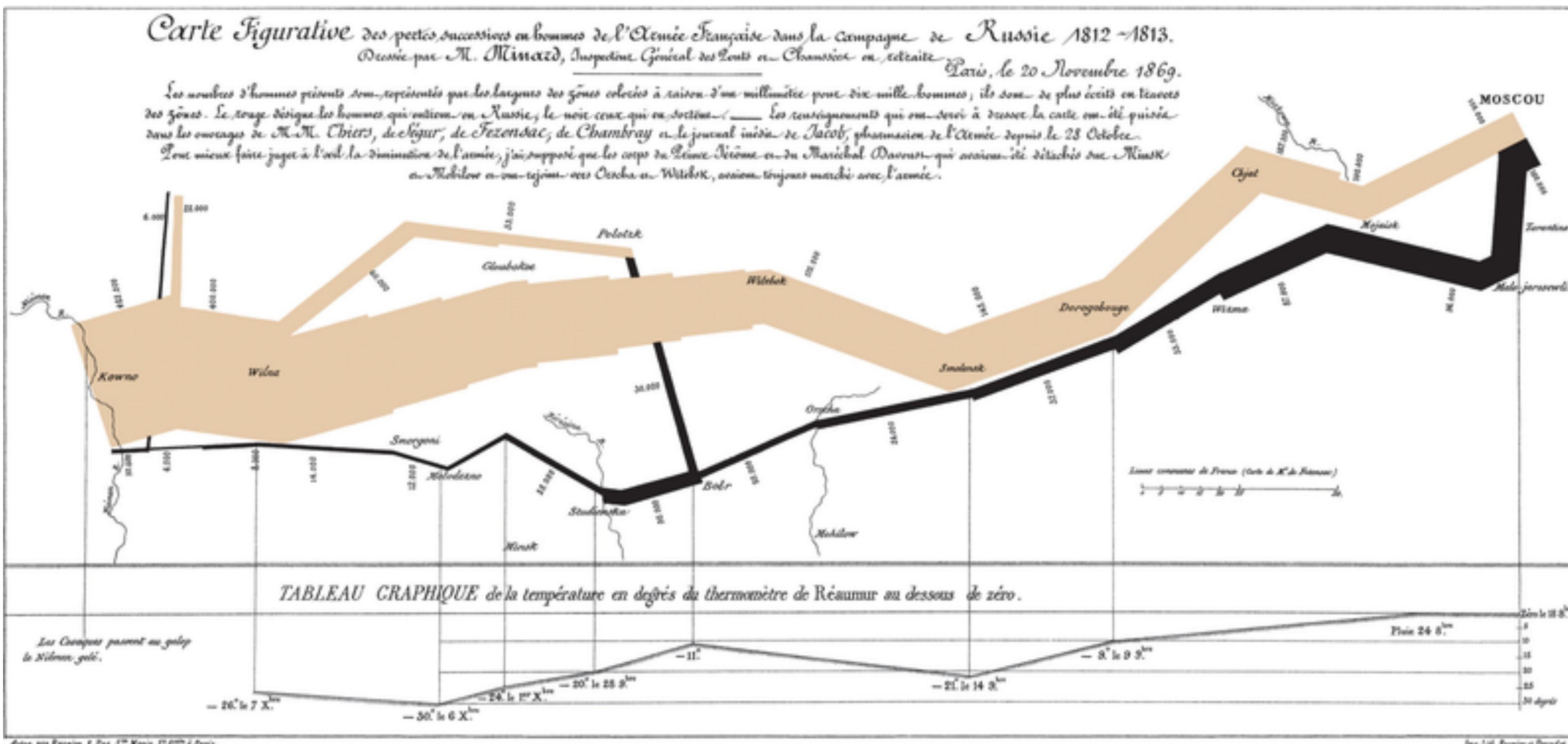
Used mapping to solve a cholera epidemic in London



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

1781-1870

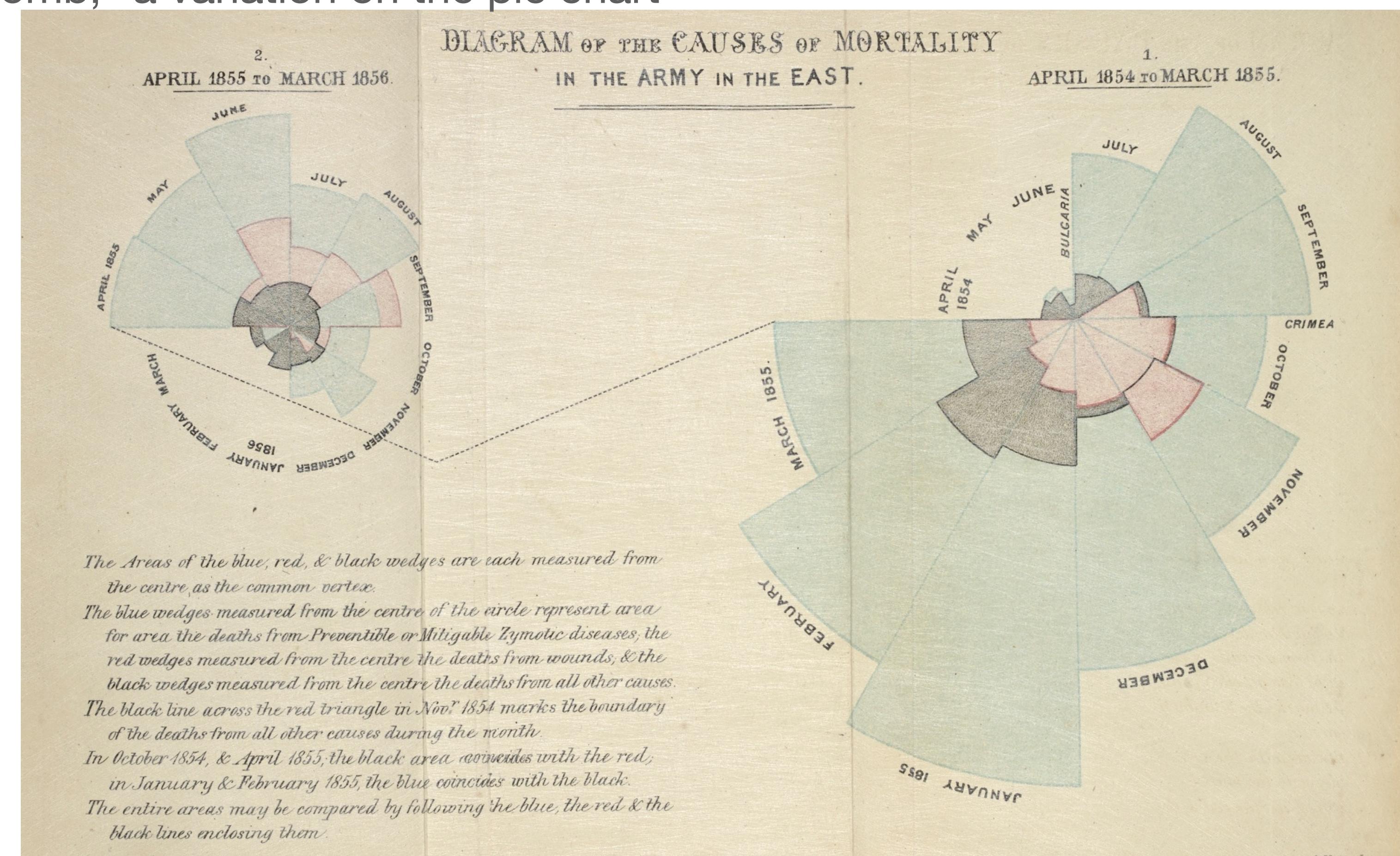
Charles Joseph Minard



1820-1910

Florence Nightingale

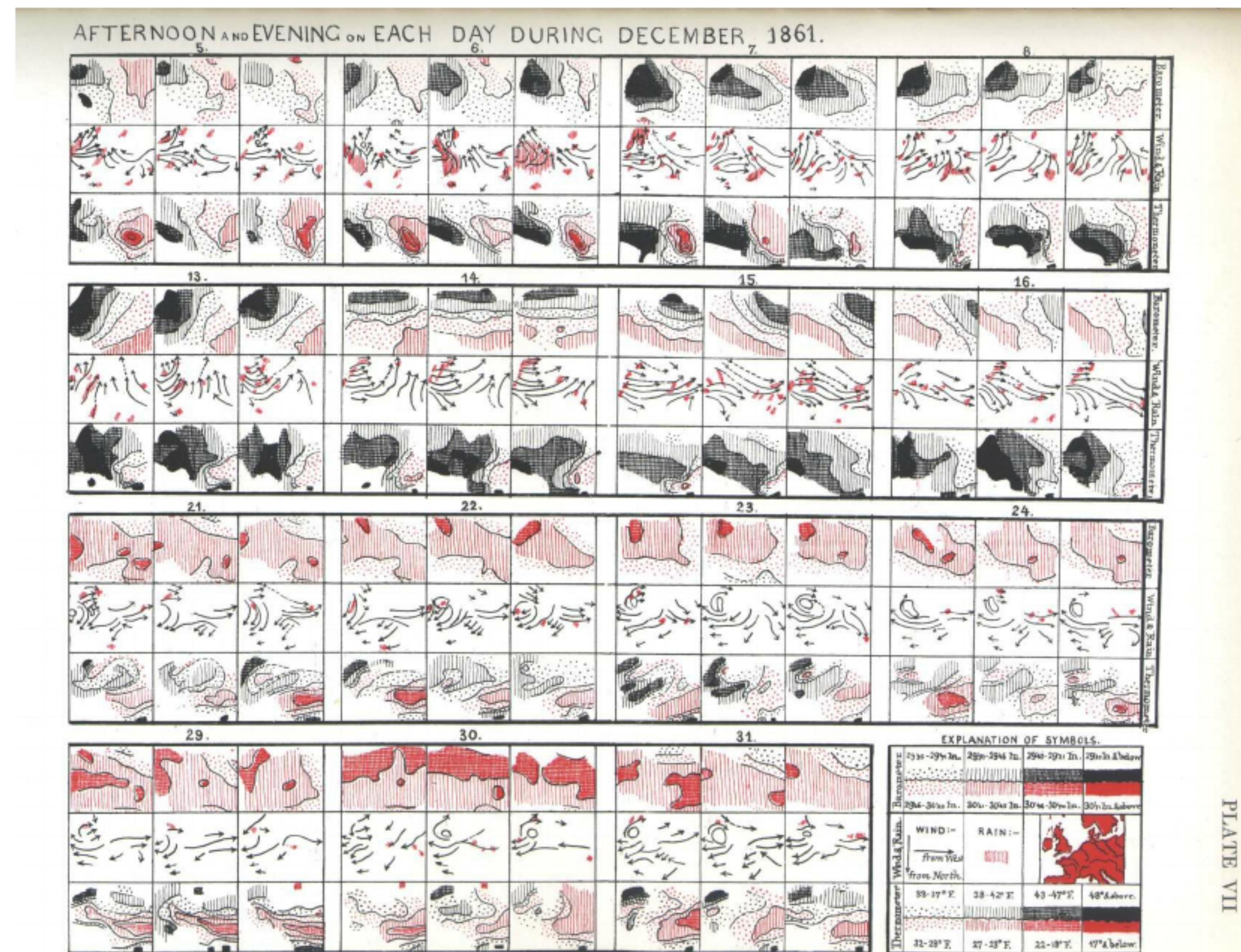
In addition to her work as a nurse, Nightingale was a statistician and invented the “coxcomb,” a variation on the pie chart



1822-1911

Francis Galton

Super-famous statistician 😊
and eugenicist 🤢



1800s

Statistical atlases

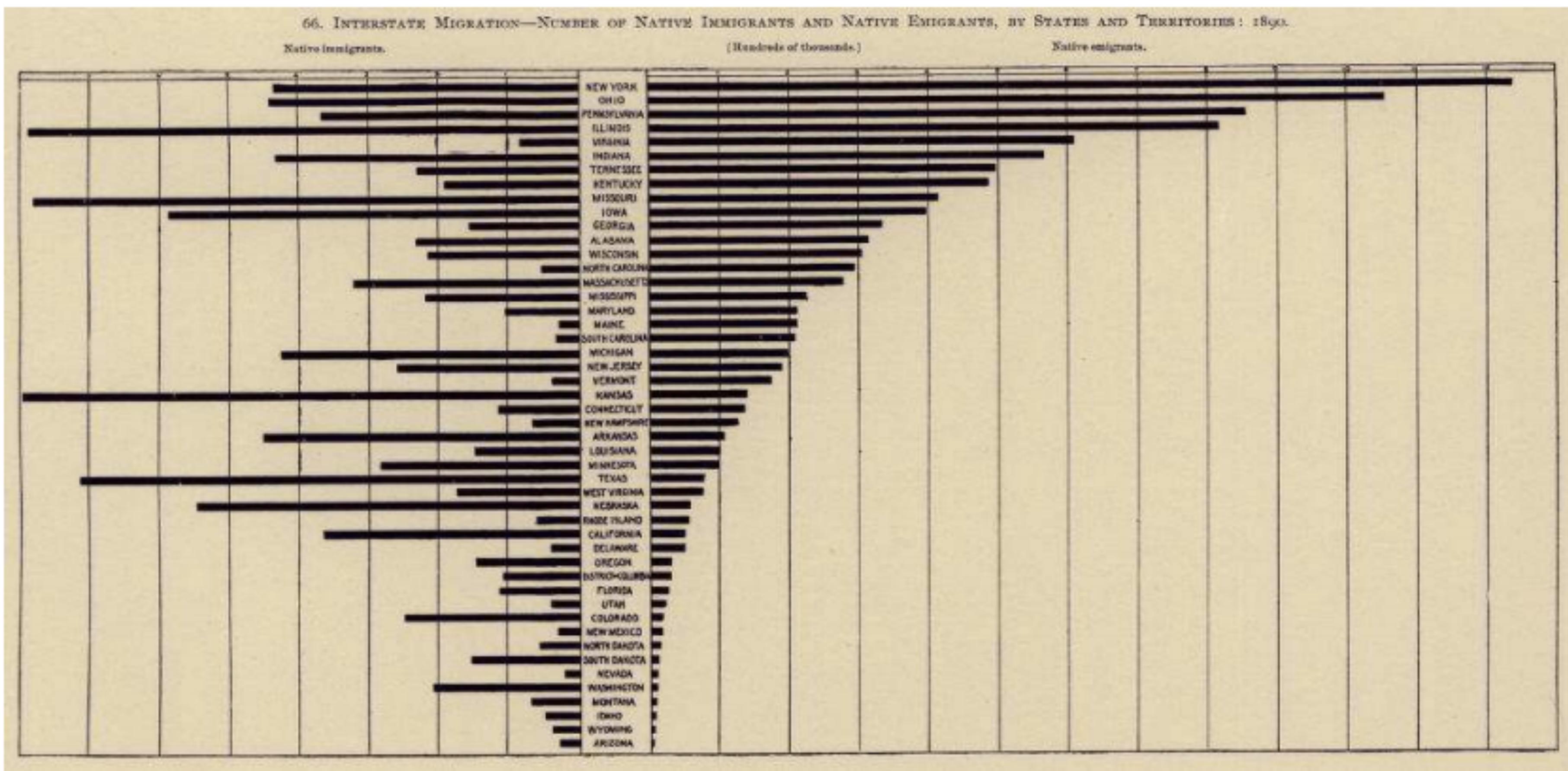


Figure 13: Interstate migration shown by back-to-back bar charts, sorted by emigration. *Source:* Statistical Atlas of the Eleventh Census, 1890, diagram 66, p. 23 (author's collection).

1868-1963

W. E. B. Du Bois

(yes, the same Du Bois you're thinking of)

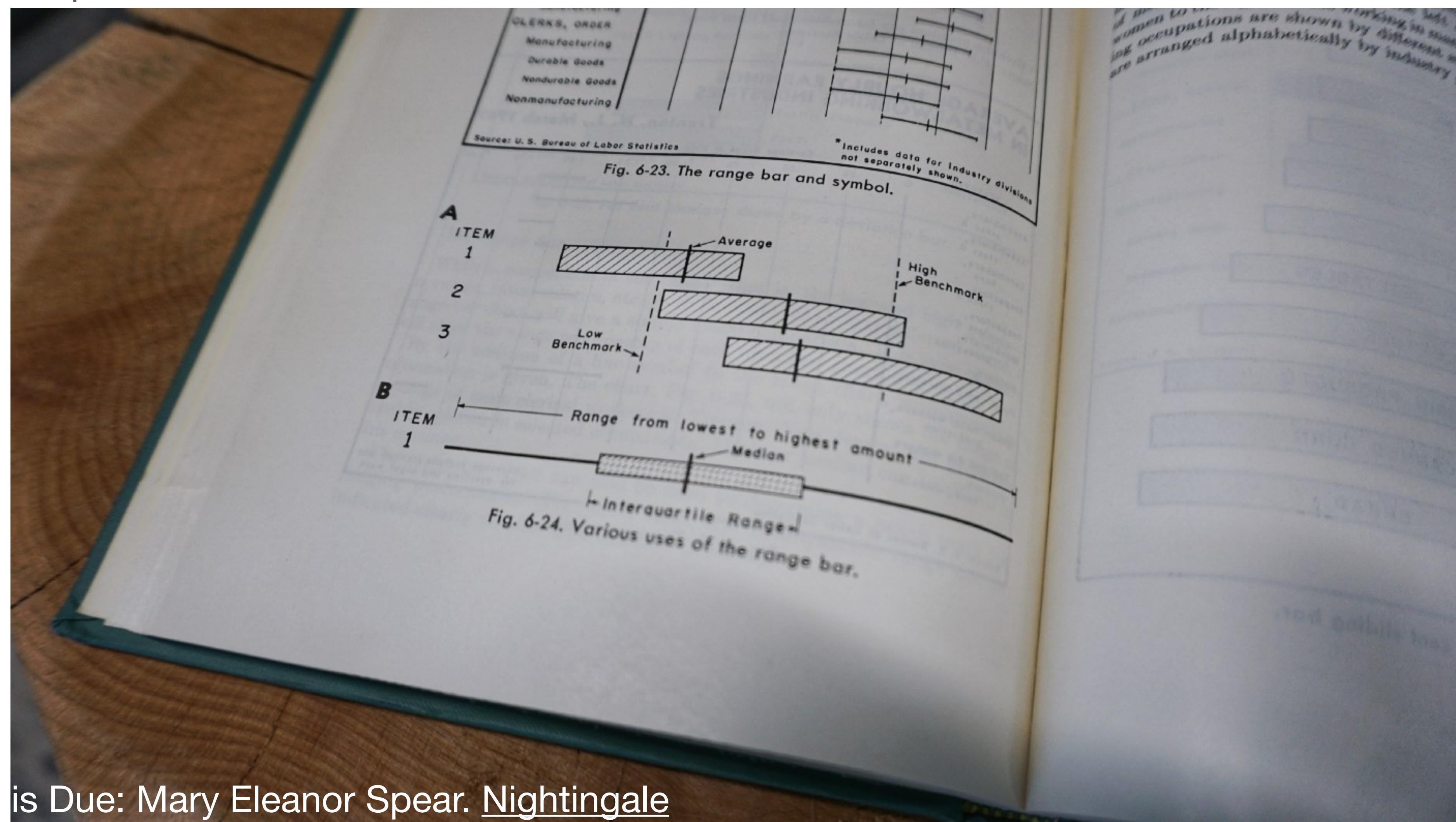
Data Portraits at the Paris Exhibition



1897-1986

Mary Eleanor Spear

Statistician who developed the box plot and bar chart!



1915–2000

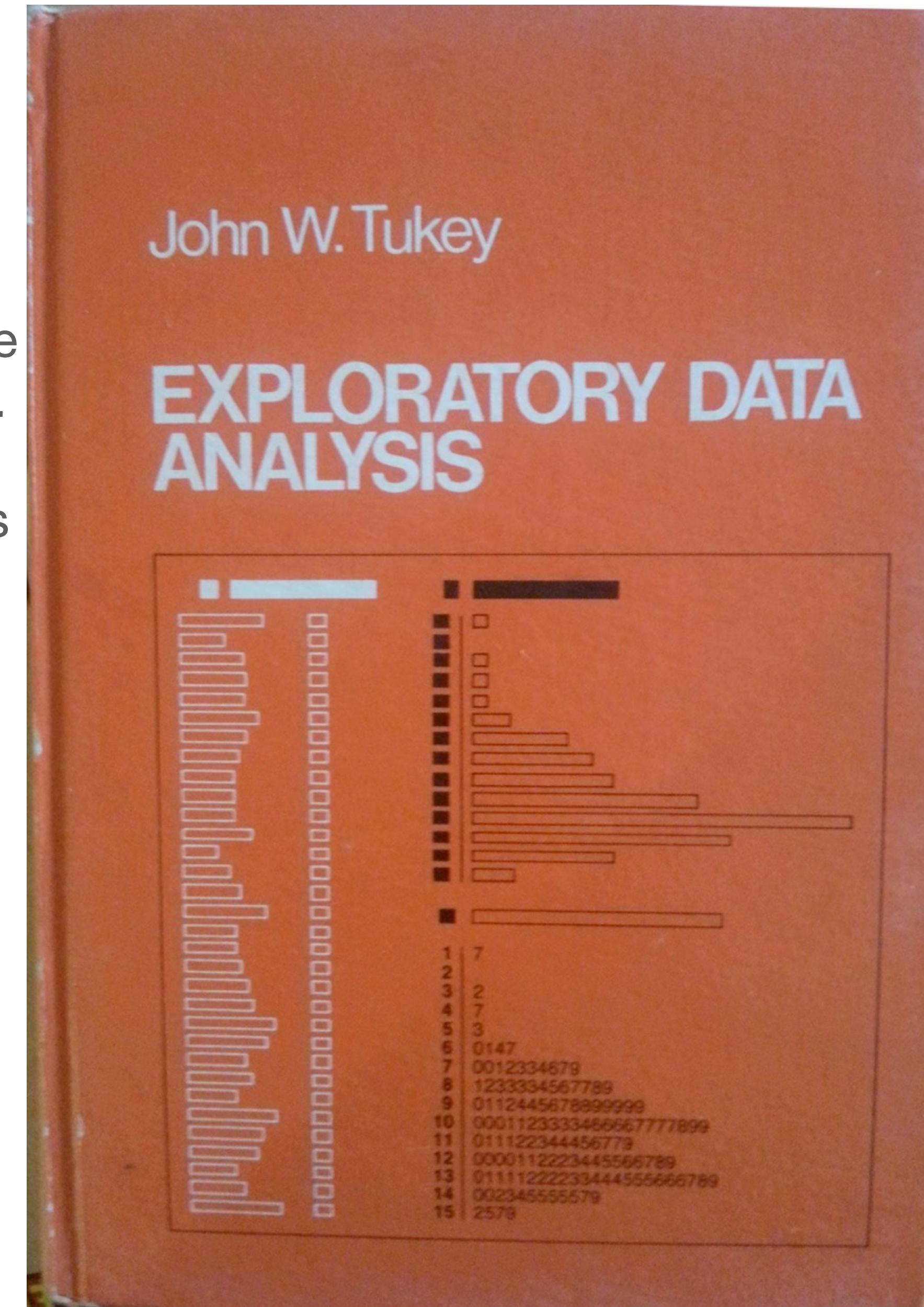
John Tukey

Statistician who rocked the boat

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

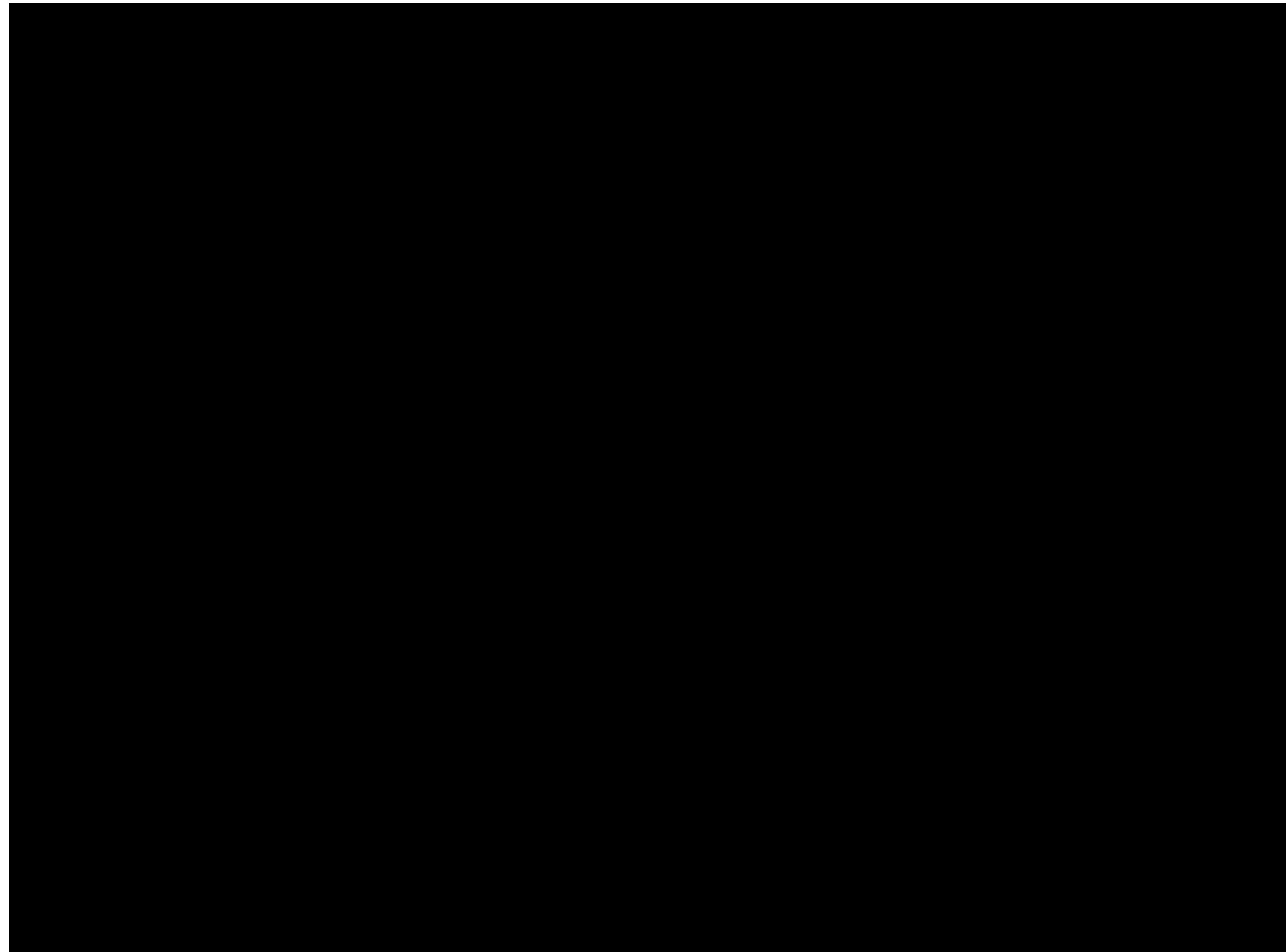
“The greatest value of a picture is when it forces us to notice what we never expected to see.”

Got credit for the boxplot, but didn't create it



1915–2000

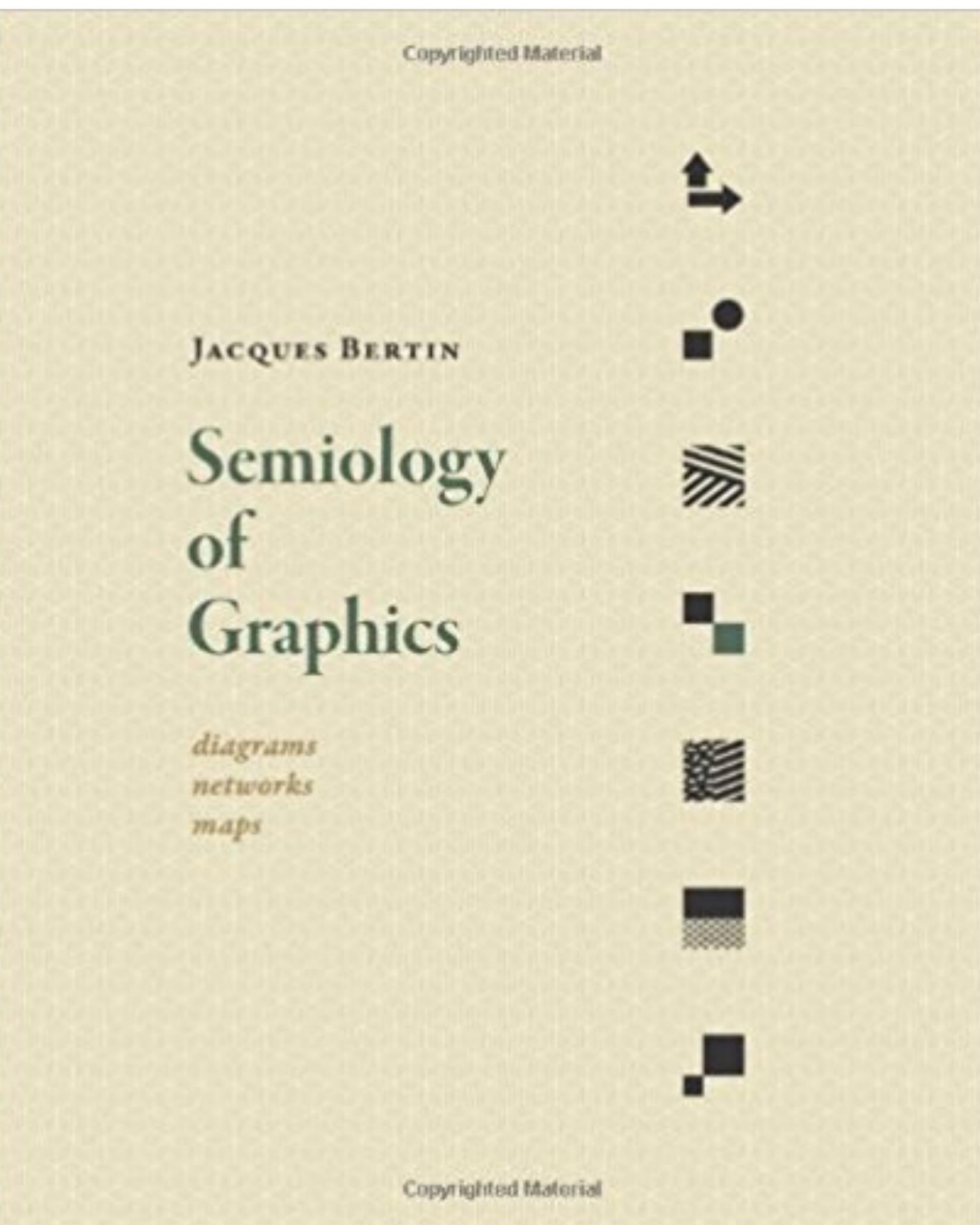
John Tukey



1918-2010

Jacques Bertin

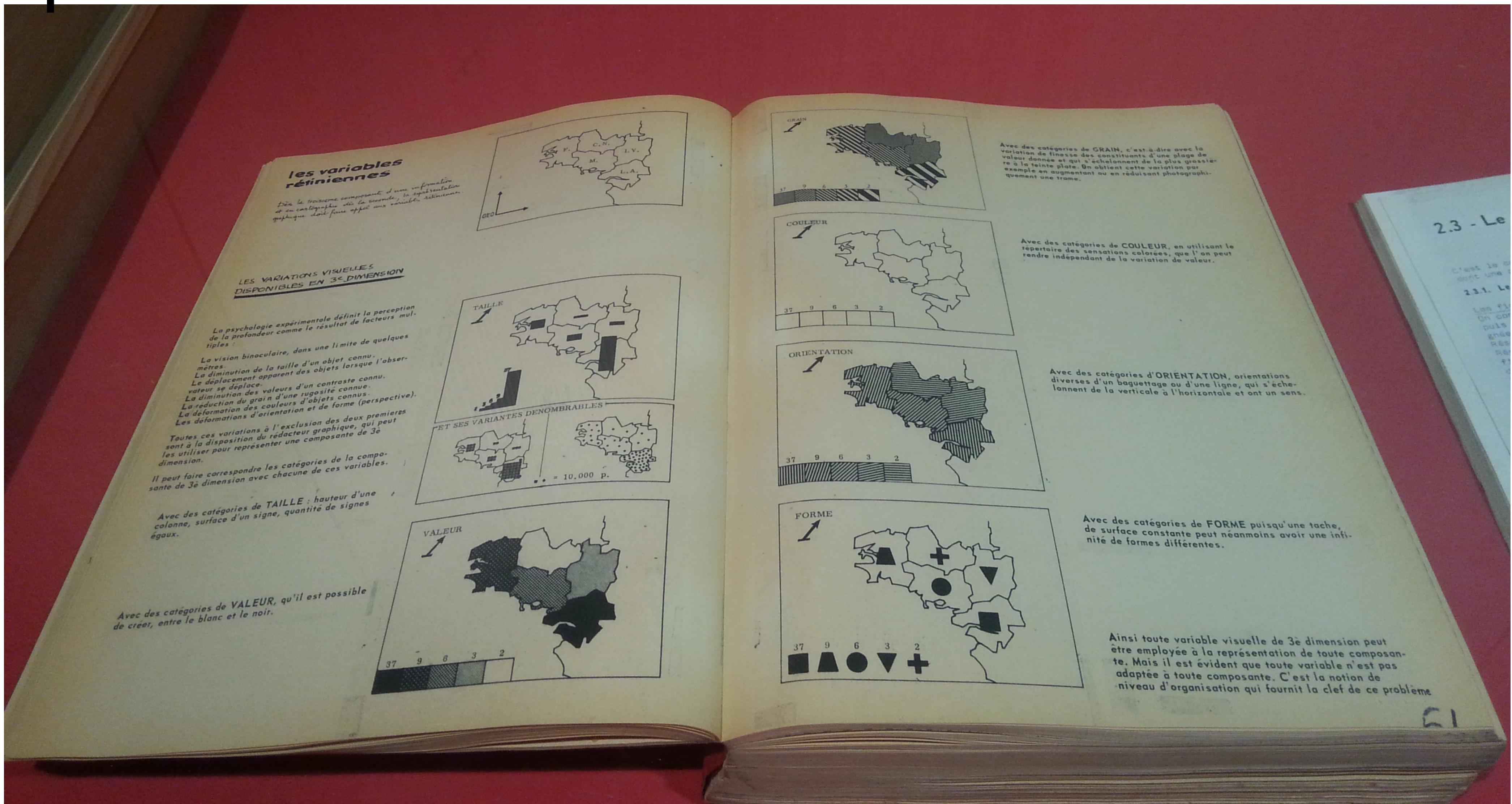
“Cartographer and theorist”



	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>

1918-2010

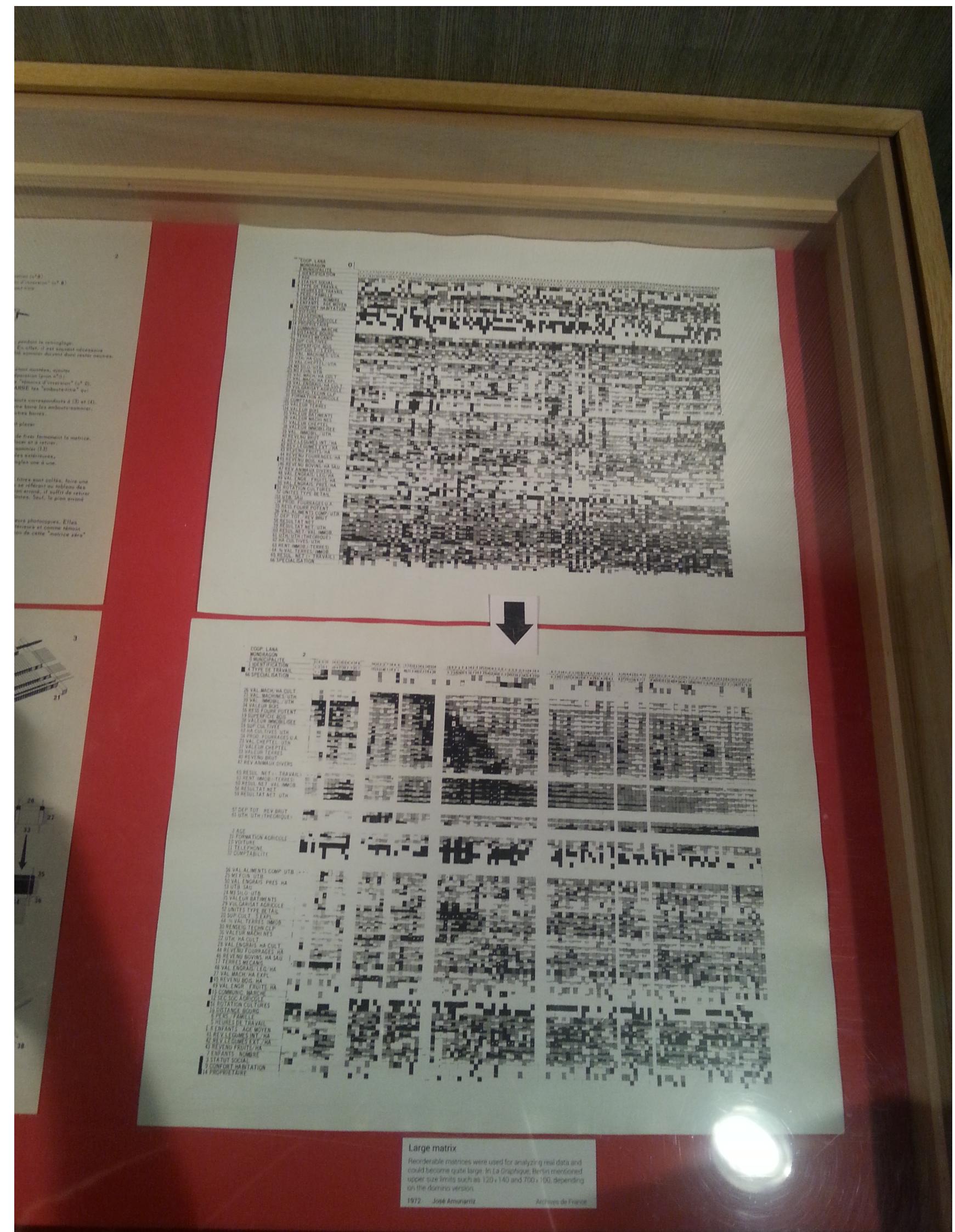
Jacques Bertin



1918-2010

Jacques Bertin

Bertin matrices



1943-

William Cleveland

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

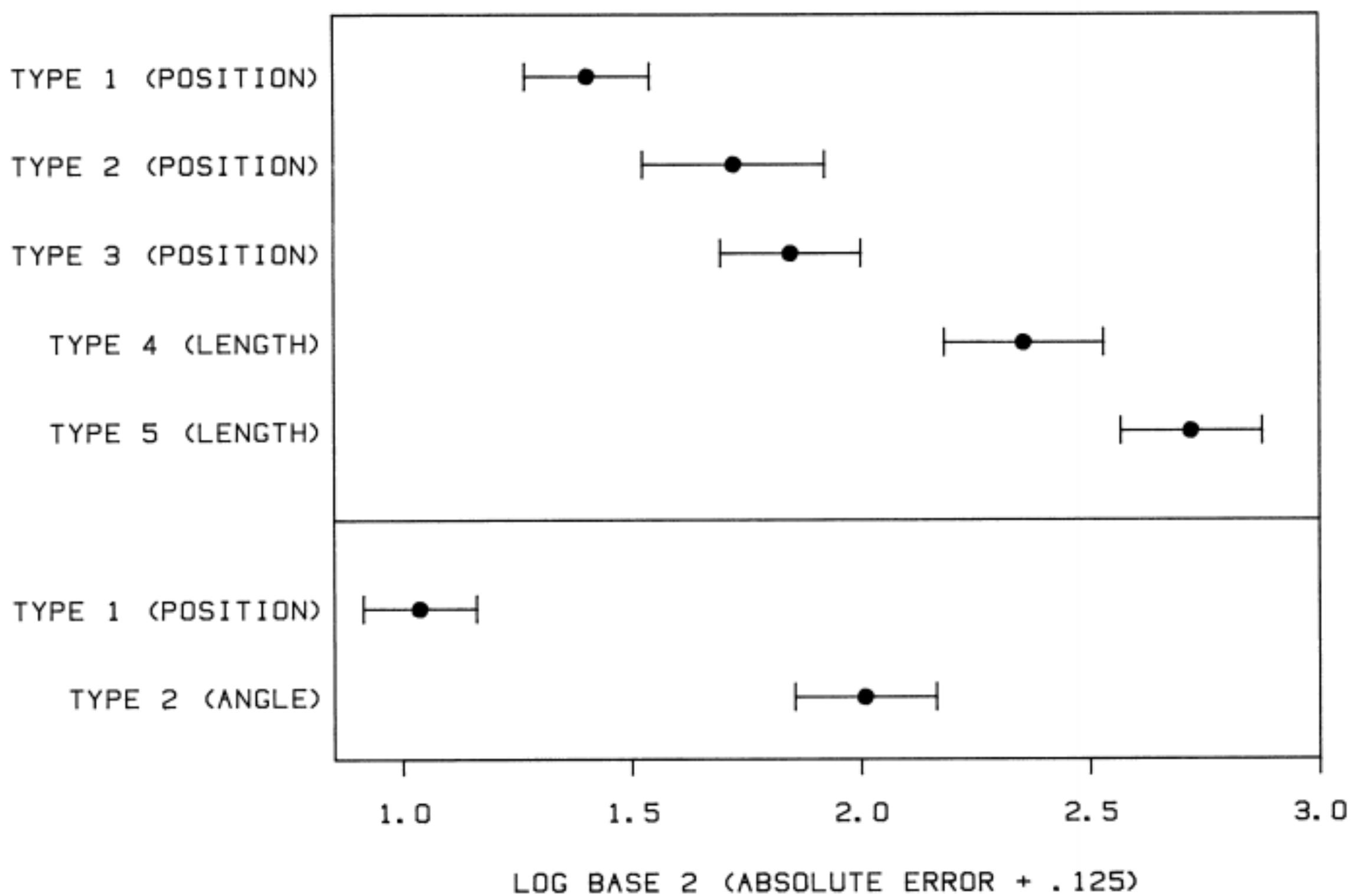
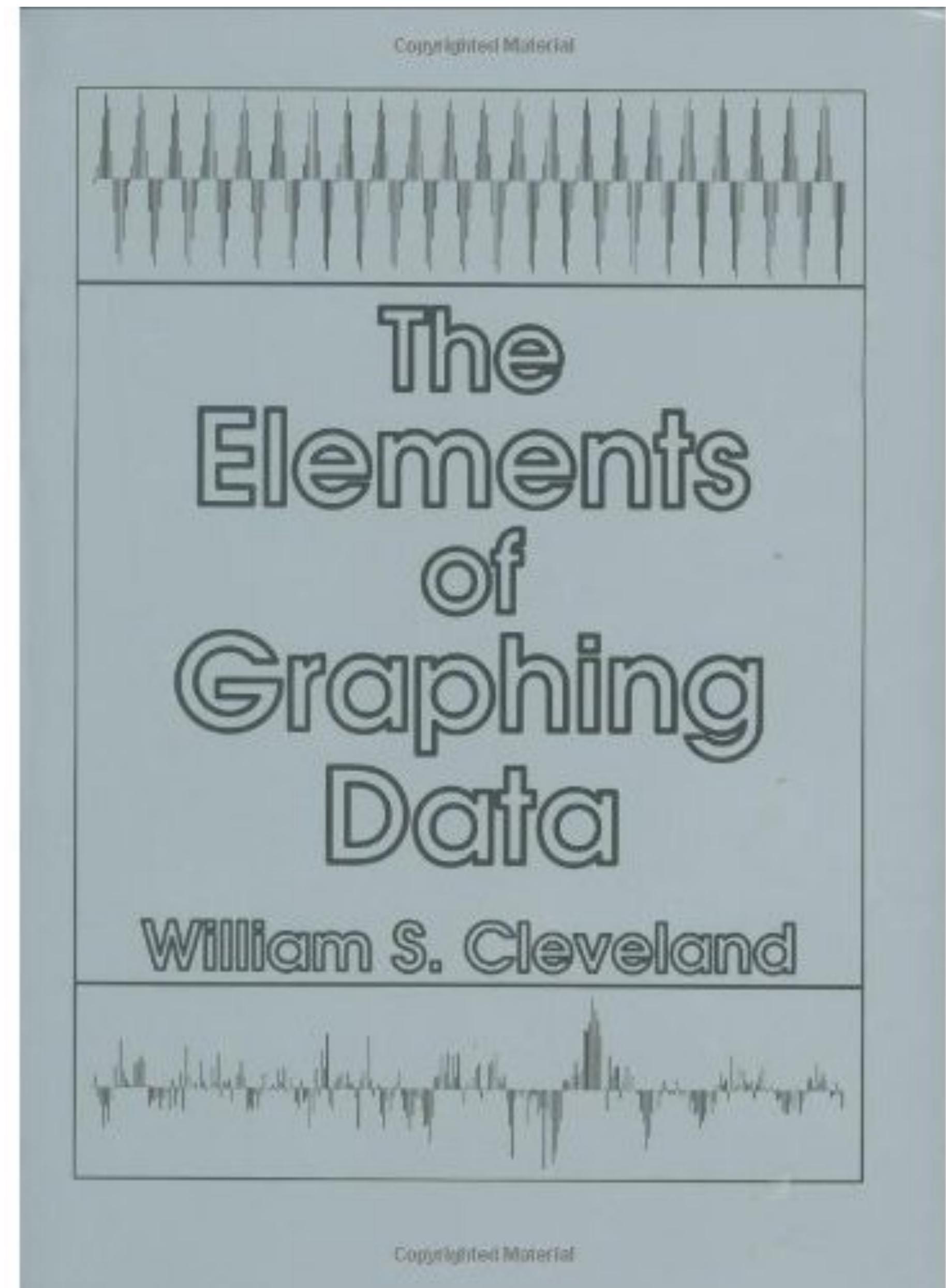
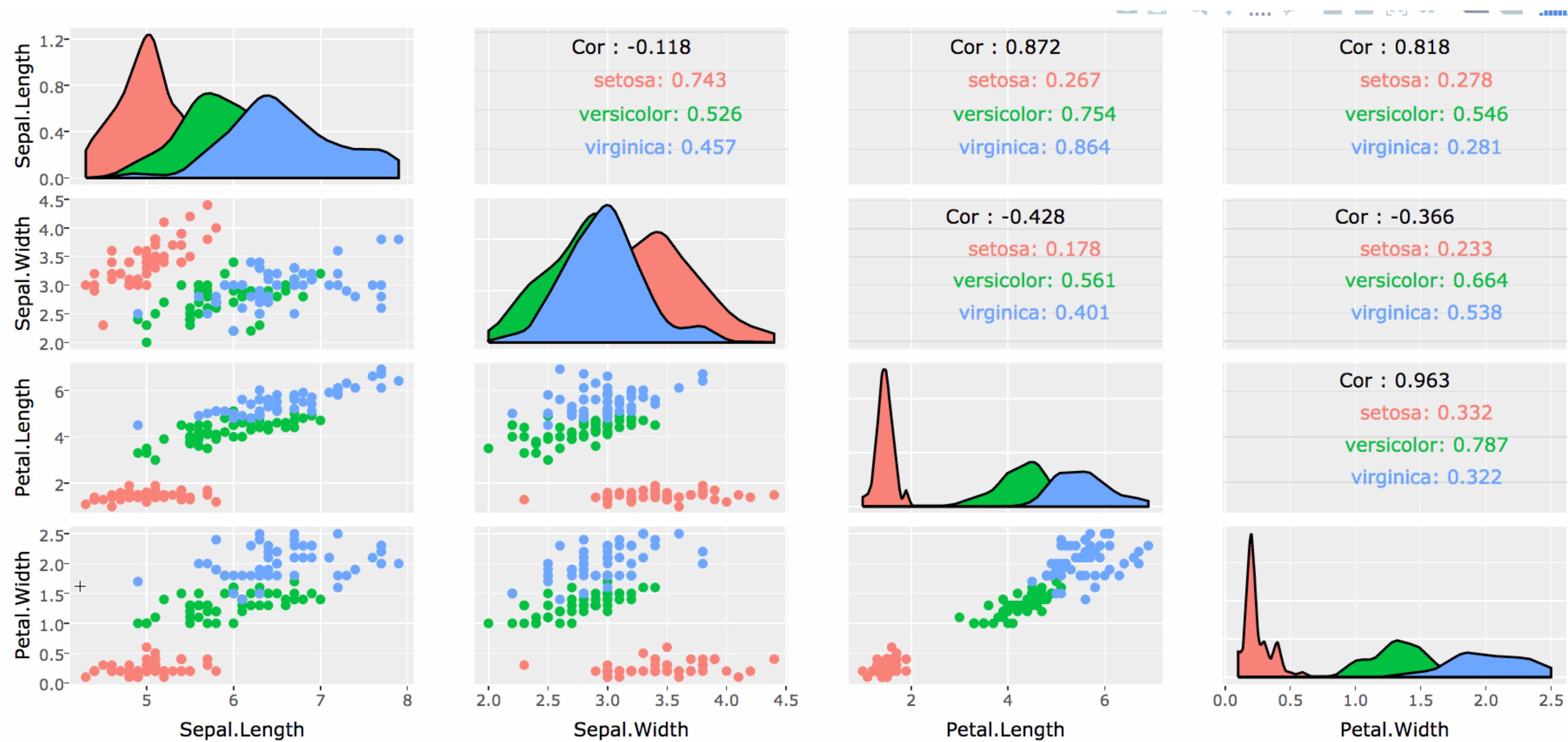


Figure 16. Log absolute error means and 95% confidence intervals for judgment types in position-length experiment (top) and position-angle experiment (bottom).



Interactivity, brushing and linking



Luke Tierney

xlisp-stat

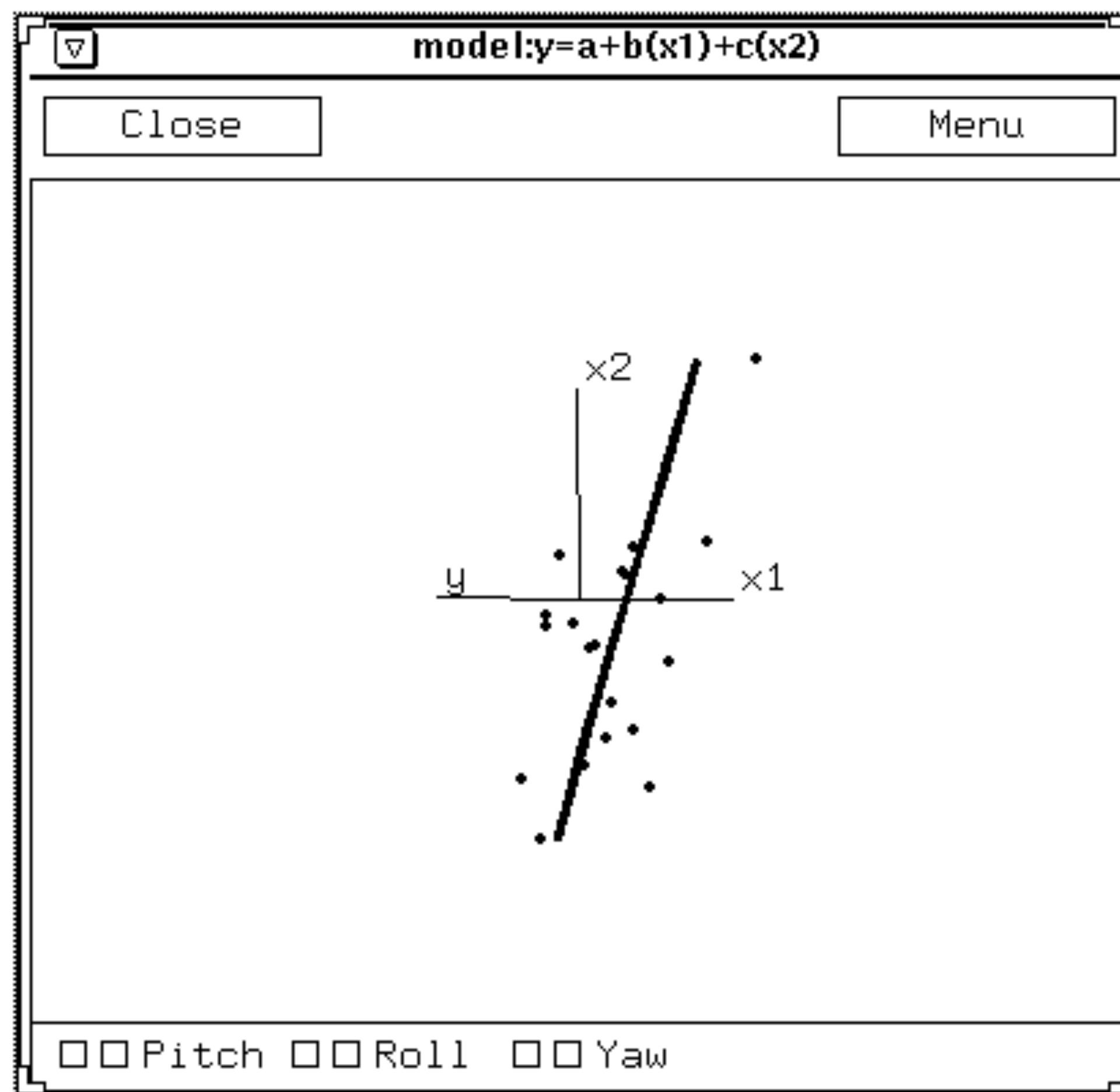


Figure 7a

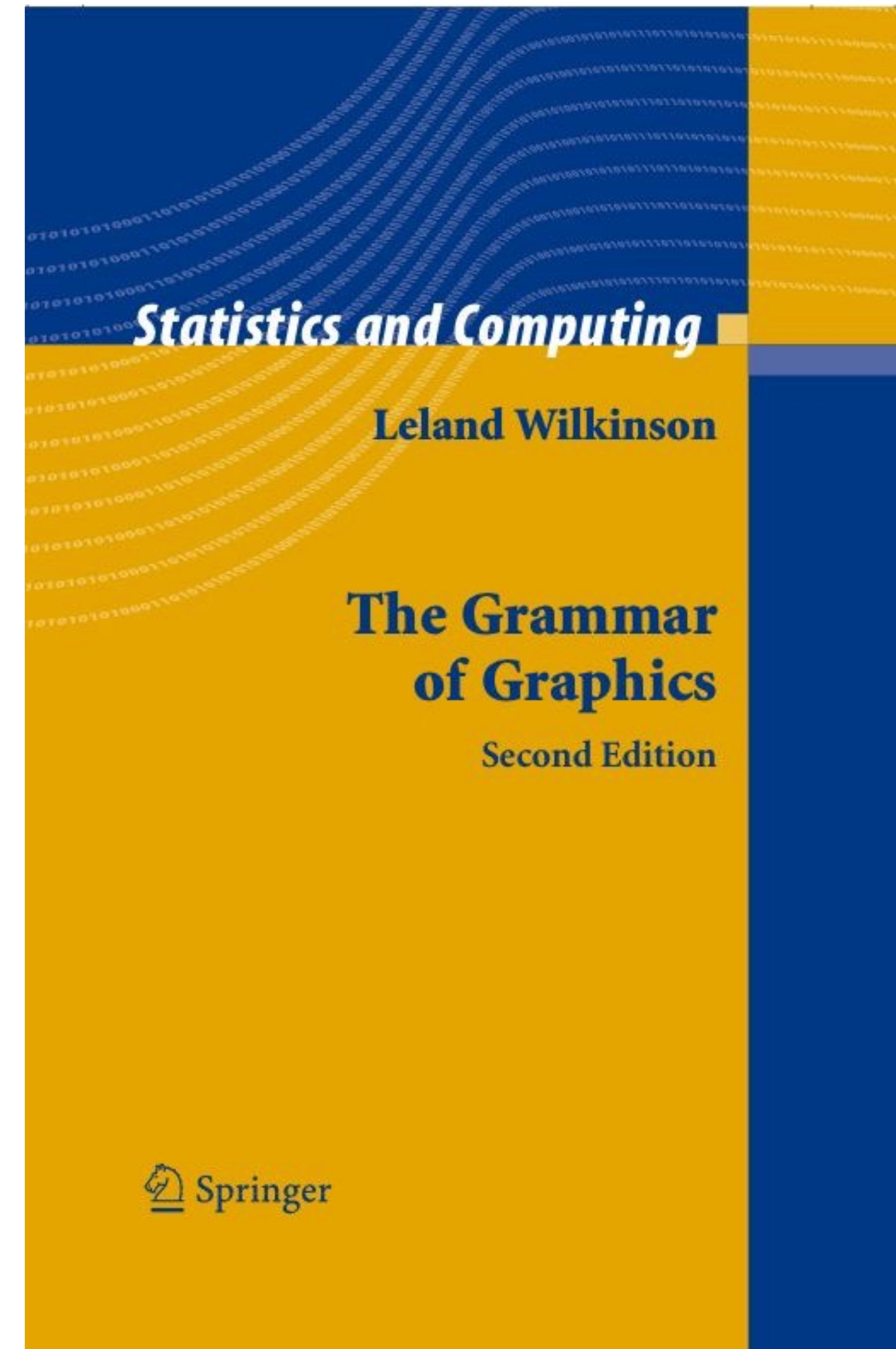
1945-

Leland Wilkinson

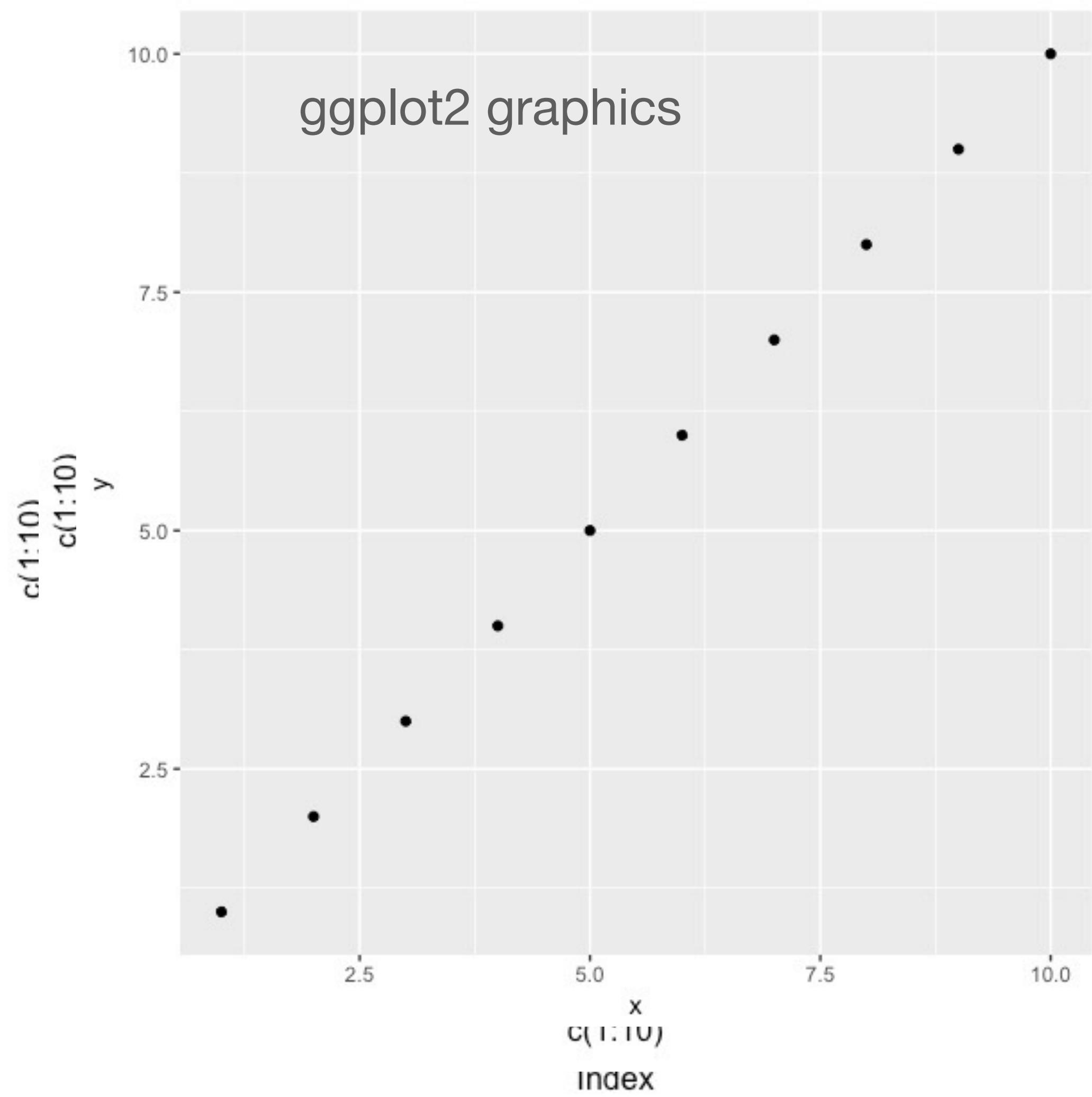
Statistician and software designer

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

*Coordinates
Statistics
Facets
Geometries
Aesthetics
Data*



R



1979-

Hadley Wickham

Famous R programmer

Implemented the grammar of graphics in R, ggplot2

Works at RStudio



Mike Bostock

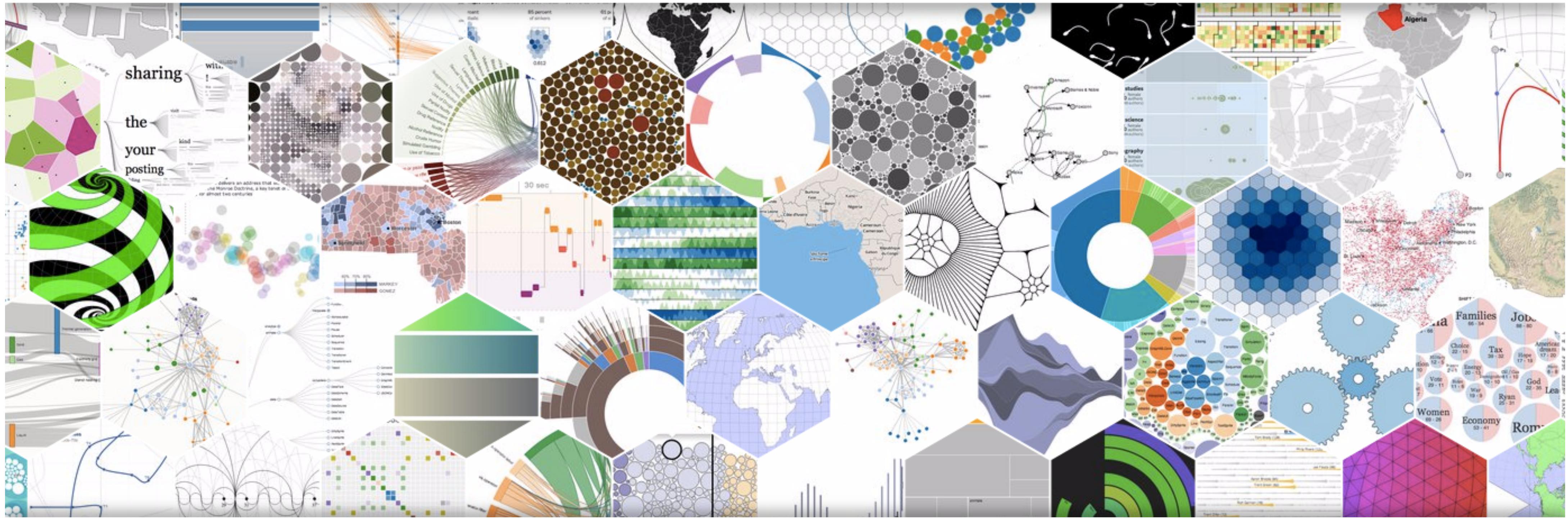
Created d3.js, a javascript implementation
of the grammar of graphics

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



Data-Driven Documents

A small orange rectangular button with white text that reads "Fork me on GitHub".



Things are accelerating!

