Value of Visualization

Nam Wook Kim

Visualization
Fall 2019

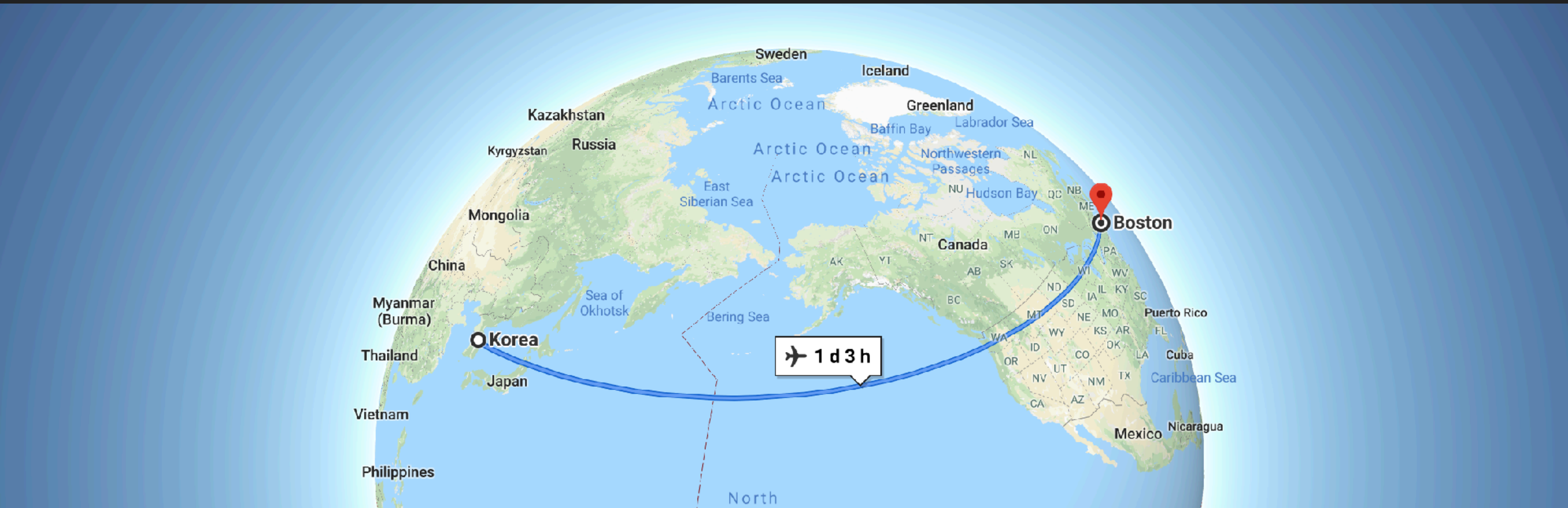
BOSTON COLLEGE

About Me

Nam Wook Kim

Assistant Professor, Computer Science Department

Information Visualization & Human-Computer Interaction



Goal

To understand the value of visualization

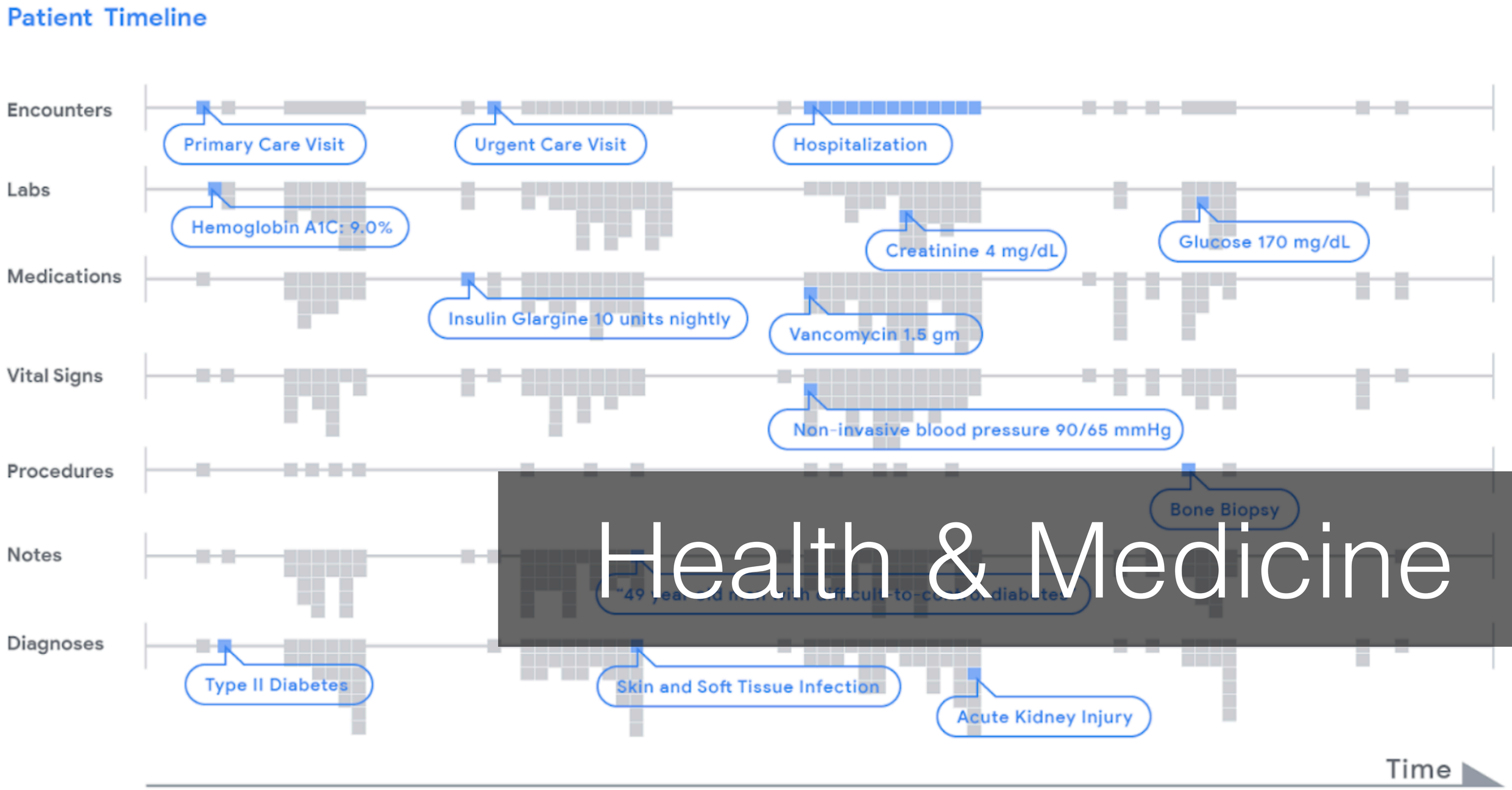
Outline

1. Rising Tide of Data
2. What Is Visualization?
3. Why Is It Important?
4. How Does It Help?
5. Three Functions of Visualization
6. Course Logistics

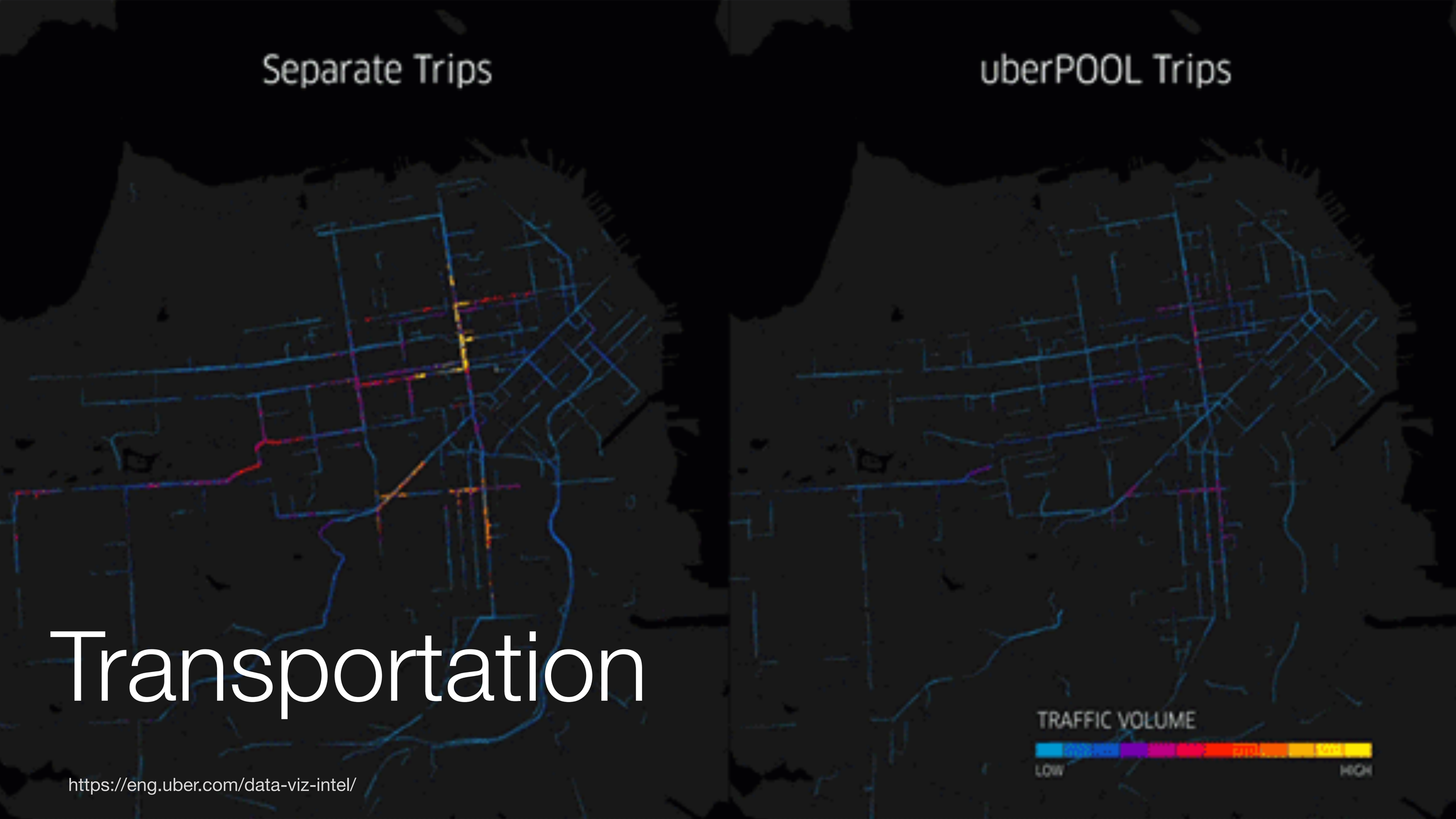
Rising Tide of Data

Big Data
Small Data
Data Everywhere

Health & Medicine



https://ai.googleblog.com/2018/05/deep-learning-for-electronic-health.html

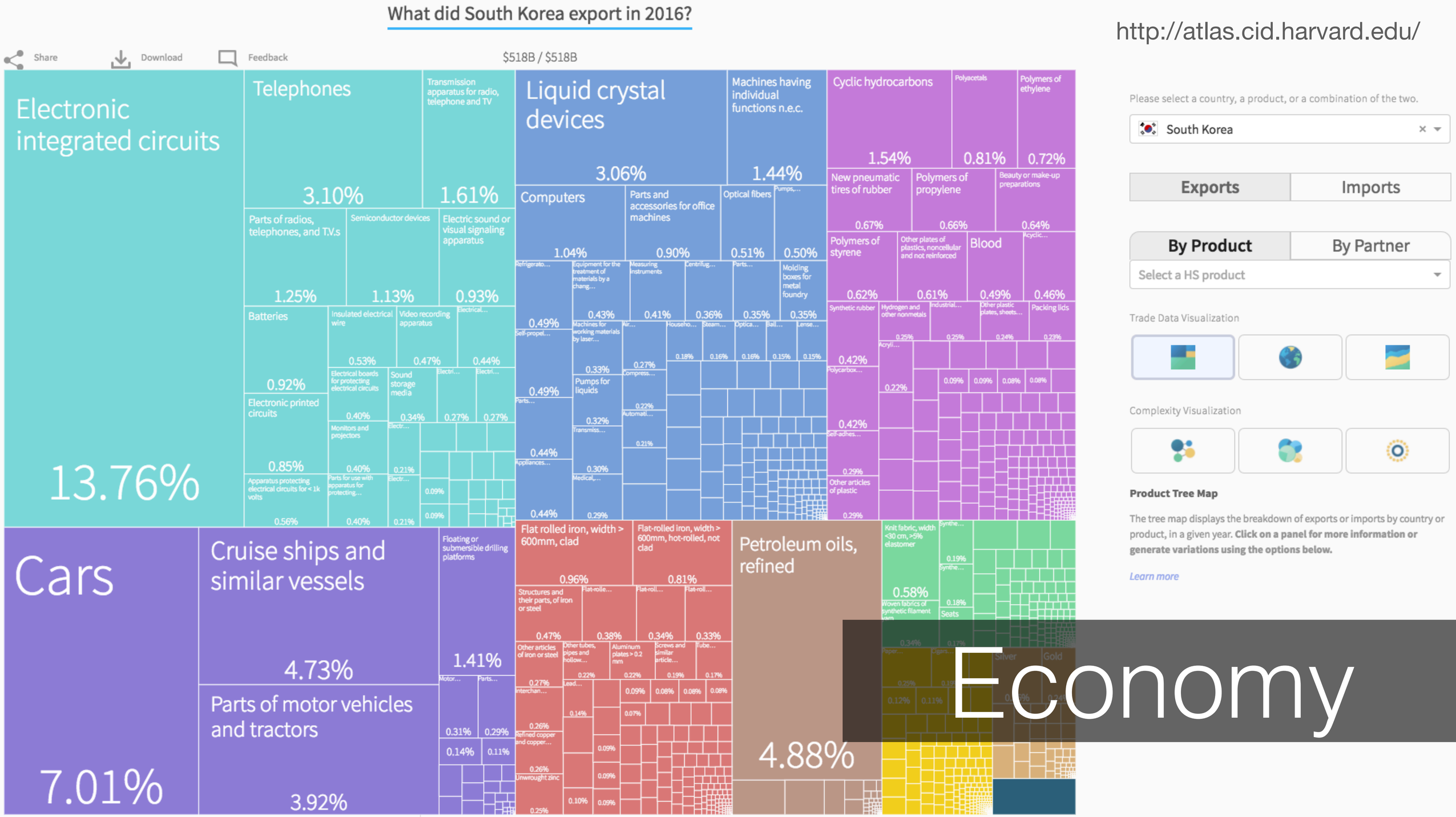


Transportation

https://eng.uber.com/data-viz-intel/

Economy

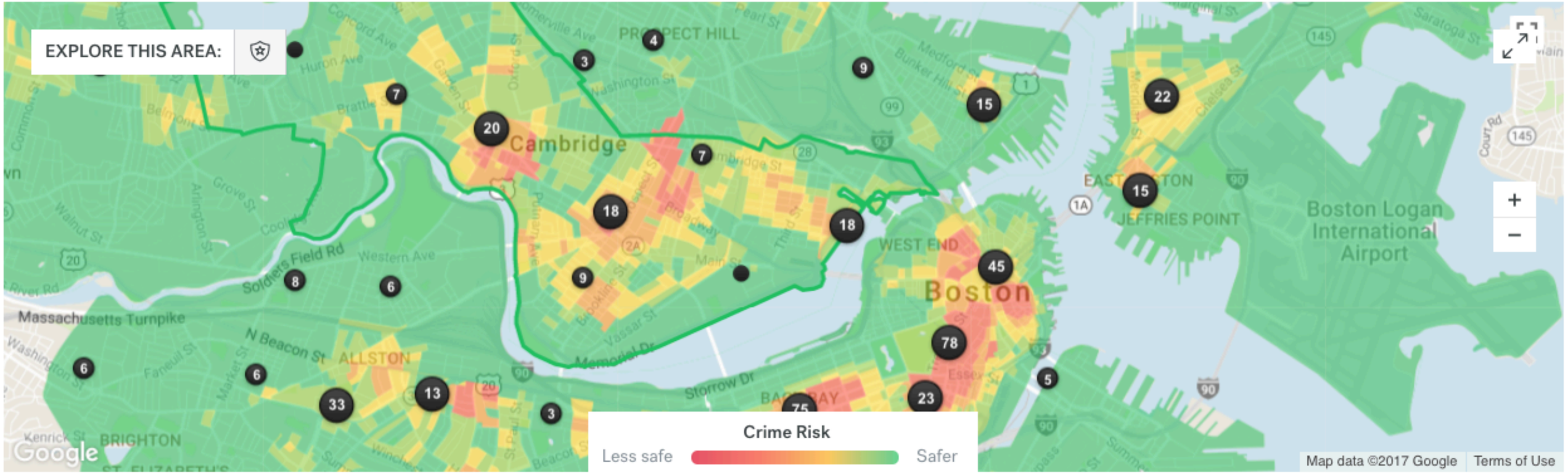
http://atlas.cid.harvard.edu/



Crime Data in Cambridge

https://www.trulia.com/real_estate/Cambridge-Massachusetts/crime/

Trulia uses crime reports to provide valuable information on the relative safety of homes in the U.S. Use the map below to learn more about crime activity in and around Cambridge.



Most frequent crimes in the past year

1,754	362	304	235	32
COUNTS OF THEFT	COUNTS OF BURGLARY	COUNTS OF ARREST	COUNTS OF ASSAULT	COUNTS OF VANDALISM

Public Safety

Human Activity

Timeline TODAY

2017 December 26

South Station Bus Terminal 7:34 AM - 7:56 AM

700 Atlantic Ave, Boston, MA 02111

4 hours 33 mins

Port Authority Bus Terminal (41st Street between 8th and 9th Avenues) 12:29 PM - 12:48 PM

41st Street, New York, NY 10018

Walking - 0.2 mi 2 mins

SEPHORA 12:50 PM - 1:16 PM

200 W 42nd St, New York, NY 10036

Walking - 0.2 mi 2 mins

Boston and New York
Dec 26 – 30, 2017

25 Monday December 2017

26 Tuesday December 2017

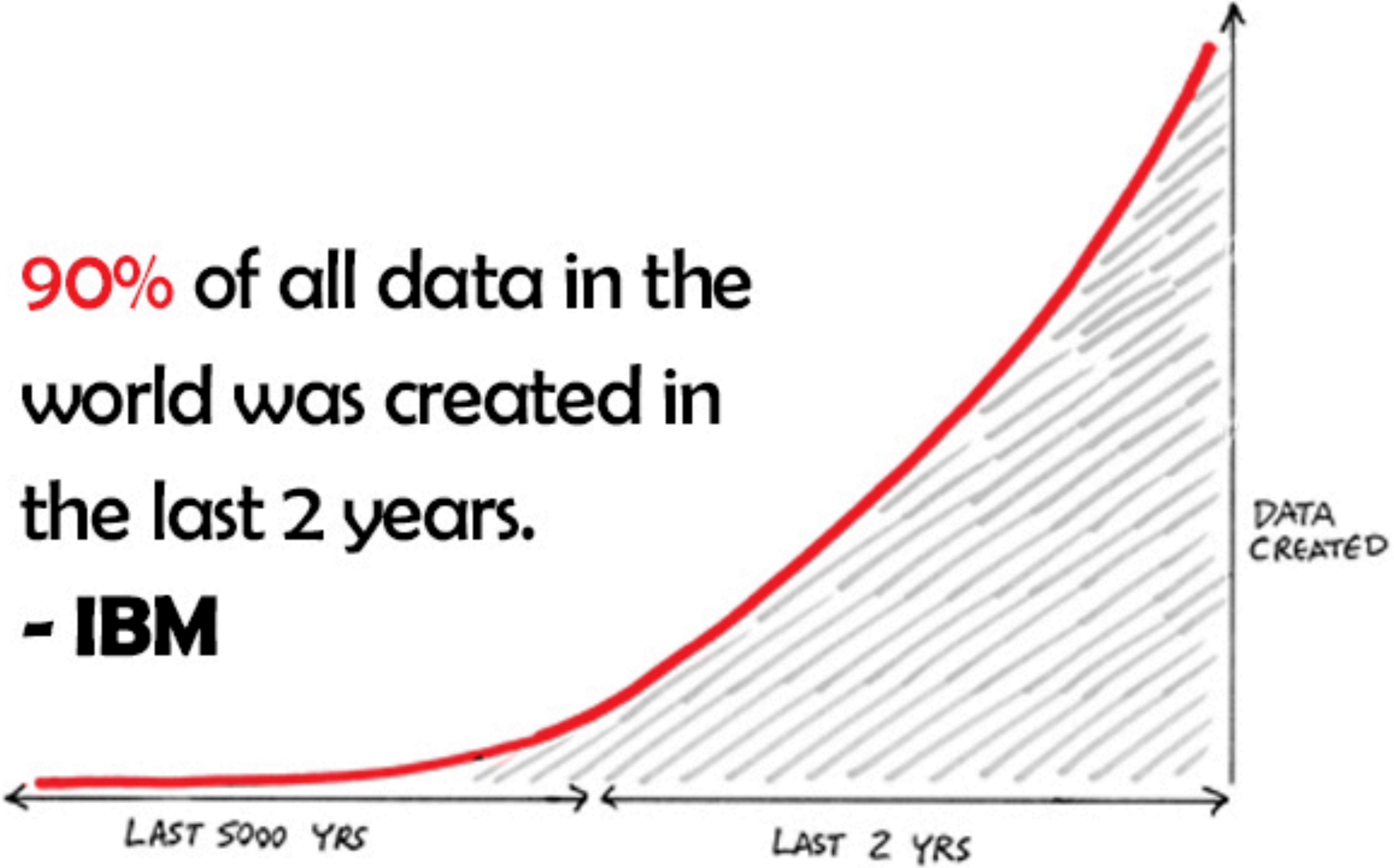
27 Wednesday December 2017

28 Thursday December 2017

29 Friday December 2017

90% of all data in the world was created in the last 2 years.

- IBM



The Industrial Revolution of Data

Joe Hellerstein
UC Berkley, 2008

Data Literacy

"The ability to take data — to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it — that's going to be a hugely important skill in the next decades …"

Hal Varian, Google's Chief Economist
The McKinsey Quarterly, January 2009

A Poverty of Attention

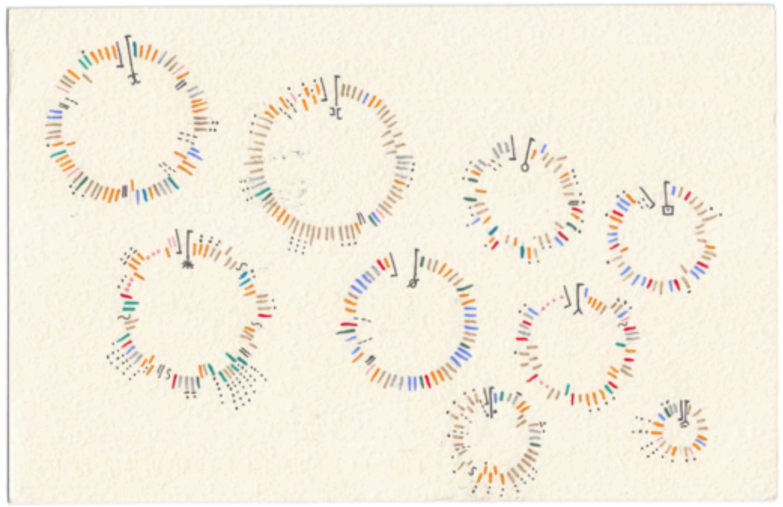
"…Information consumes the attention of its recipients. Hence … a need to allocate that attention efficiently among the overabundance of information sources that might consume it."

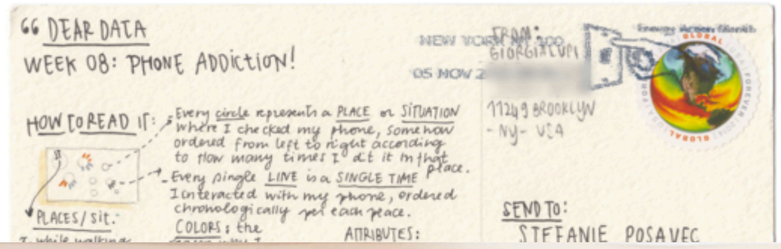
Herbert A. Simon
Economist & Psychologist

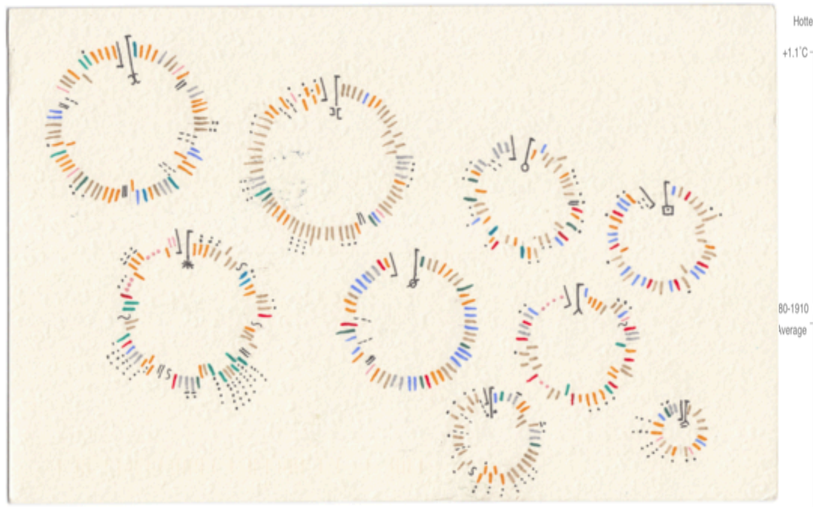
Visualization can help!

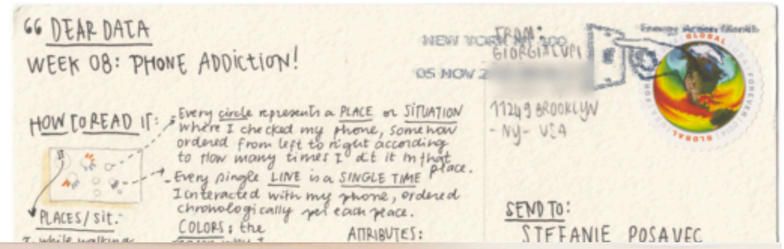
provides a powerful yet accessible way to make sense of large and complex data

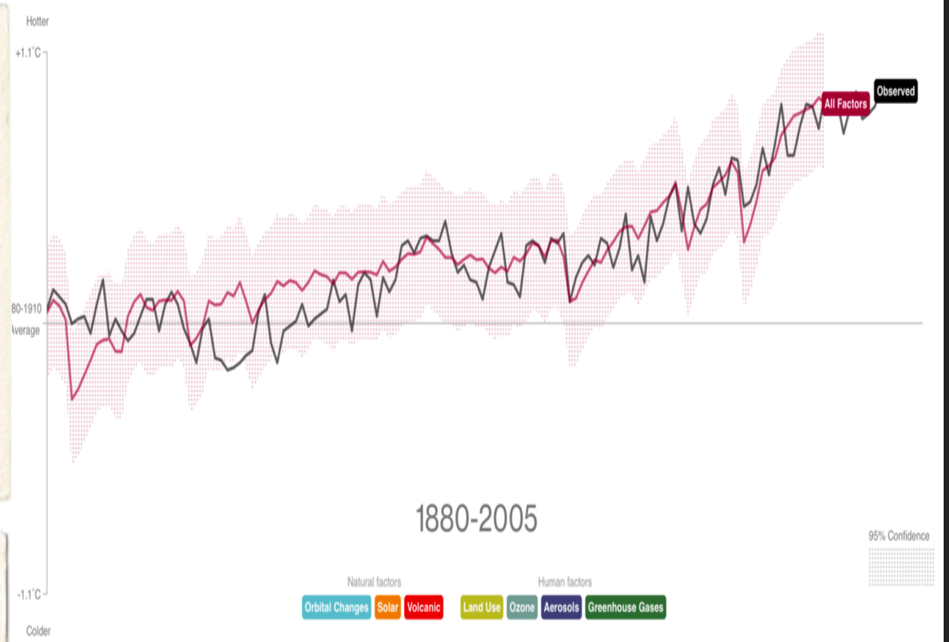
What Is Visualization?

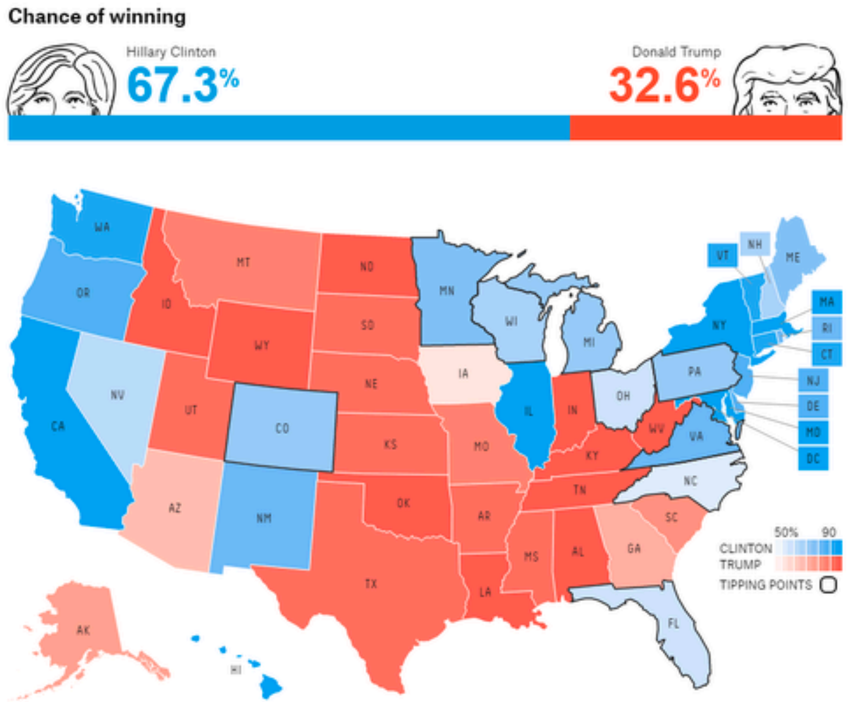


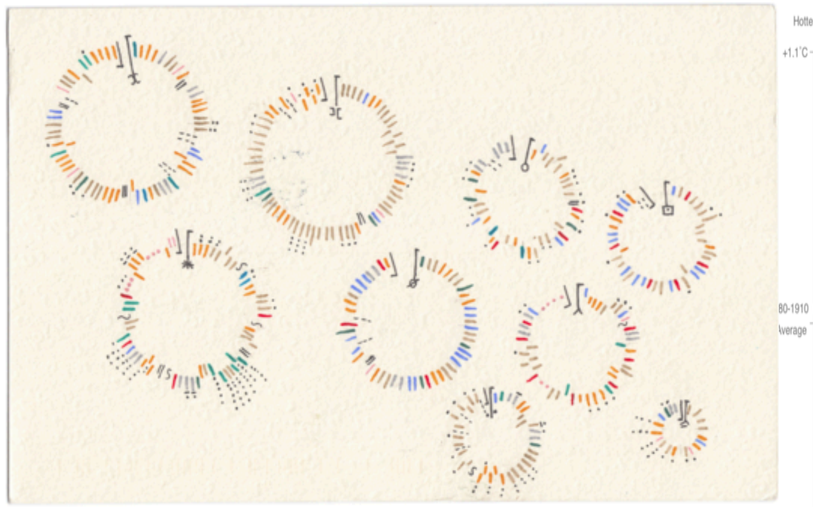

DEAR DATA
WEEK 08: PHONE ADDICTION!
FROM:
GIORGIA LUPI
05 NOW 2
11249 BROOKLYN
- NY- USA
HOW TO READ IT:
Every circle represents a PLACE or SITUATION where I checked my phone, somehow ordered from left to right according to how many times I'd it in that place.
Every single LINE is a SINGLE TIME I interacted with my phone, ordered chronologically per each place.
PLACES/SIT.
COLORS: the
ATTRIBUTES:
SEND TO:
STEFANIE POSAVEC

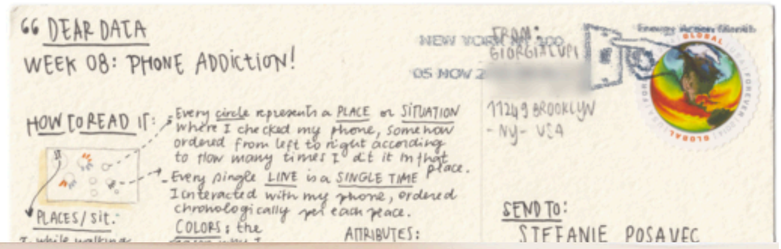


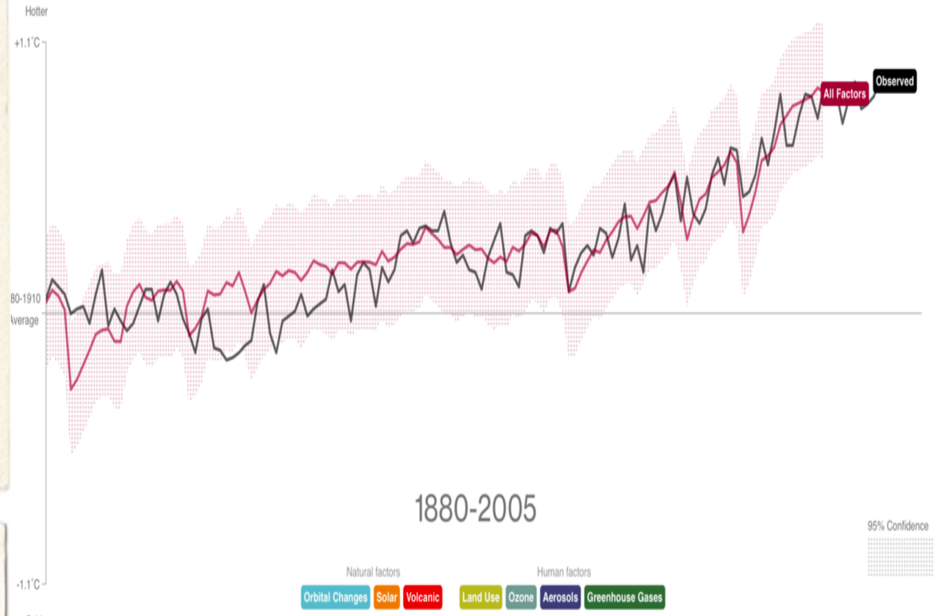

DEAR DATA
WEEK 08: PHONE ADDICTION!
HOW TO READ IT:
Every circle represents a PLACE or SITUATION where I checked my phone, somehow ordered from left to right according to how many times I'd it in that place.
Every single LINE is a SINGLE TIME I interacted with my phone, ordered chronologically per each peace.
PLACES / SIT.
COLORS: the
ATTRIBUTES:
SEND TO:
STEFANIE POSAVEC
11249 BROOKLYN
- NY- USA

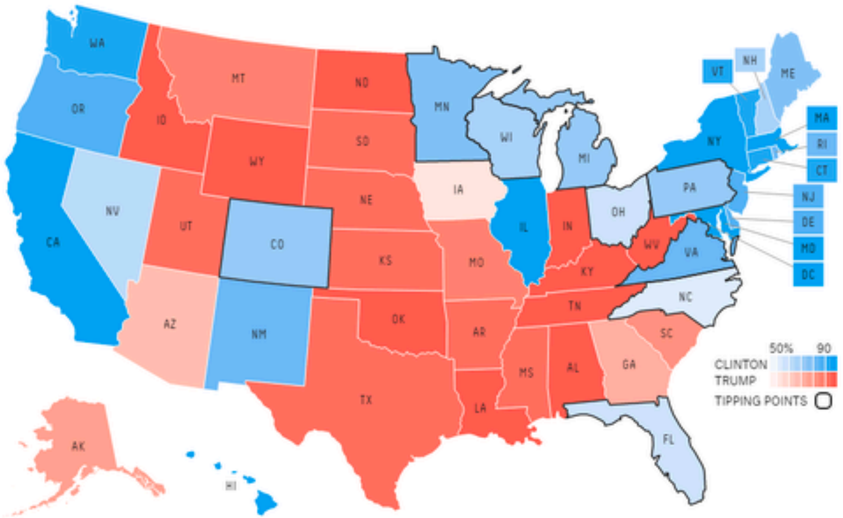

Hotter
+1.1°C
80-1910
Average
-1.1°C
Colder
All Factors
Observed
1880-2005
95% Confidence
Natural factors
Orbital Changes
Solar
Volcanic
Human factors
Land Use
Ozone
Aerosols
Greenhouse Gases


Chance of winning
Hillary Clinton
67.3%
Donald Trump
32.6%
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MI
IN
KY
TN
MS
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
VT
NH
ME
MA
RI
CT
NJ
DE
MD
DC
AK
HI
50%
90
CLINTON
TRUMP
TIPPING POINTS
Electoral votes
Popular vote



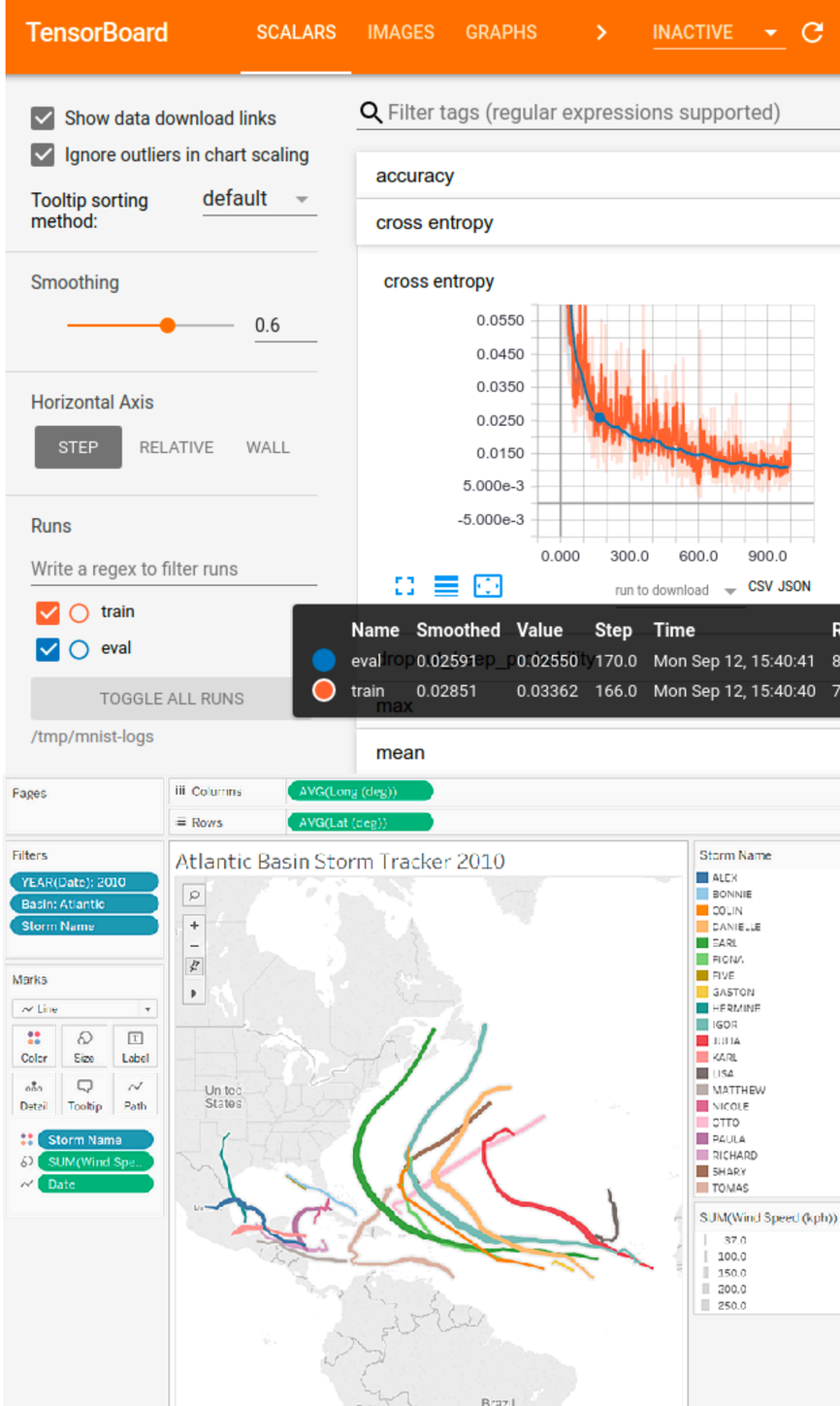




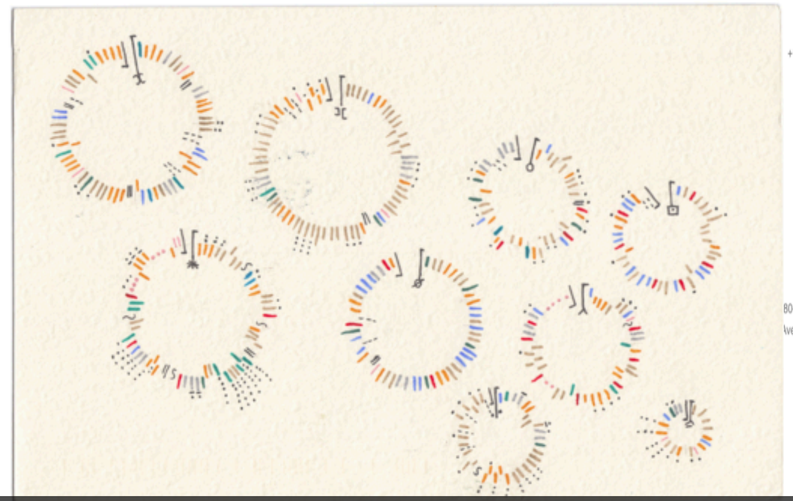


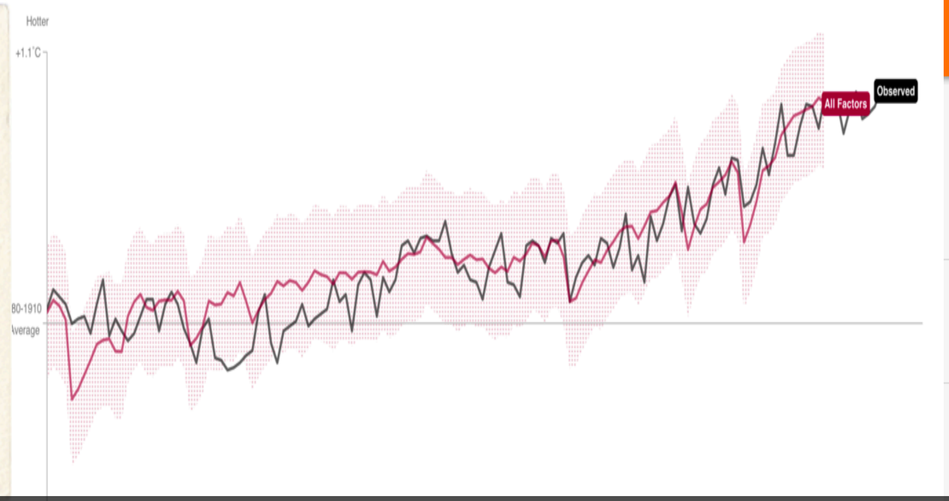
Electoral votes

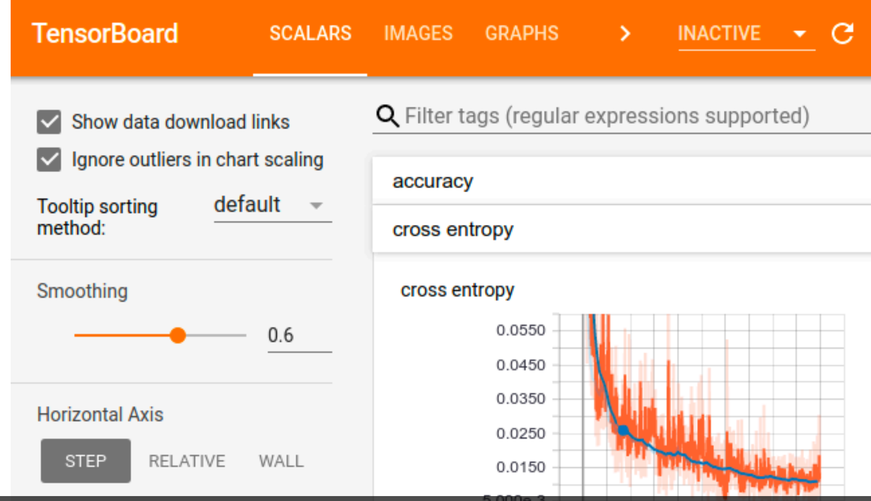
Popular vote

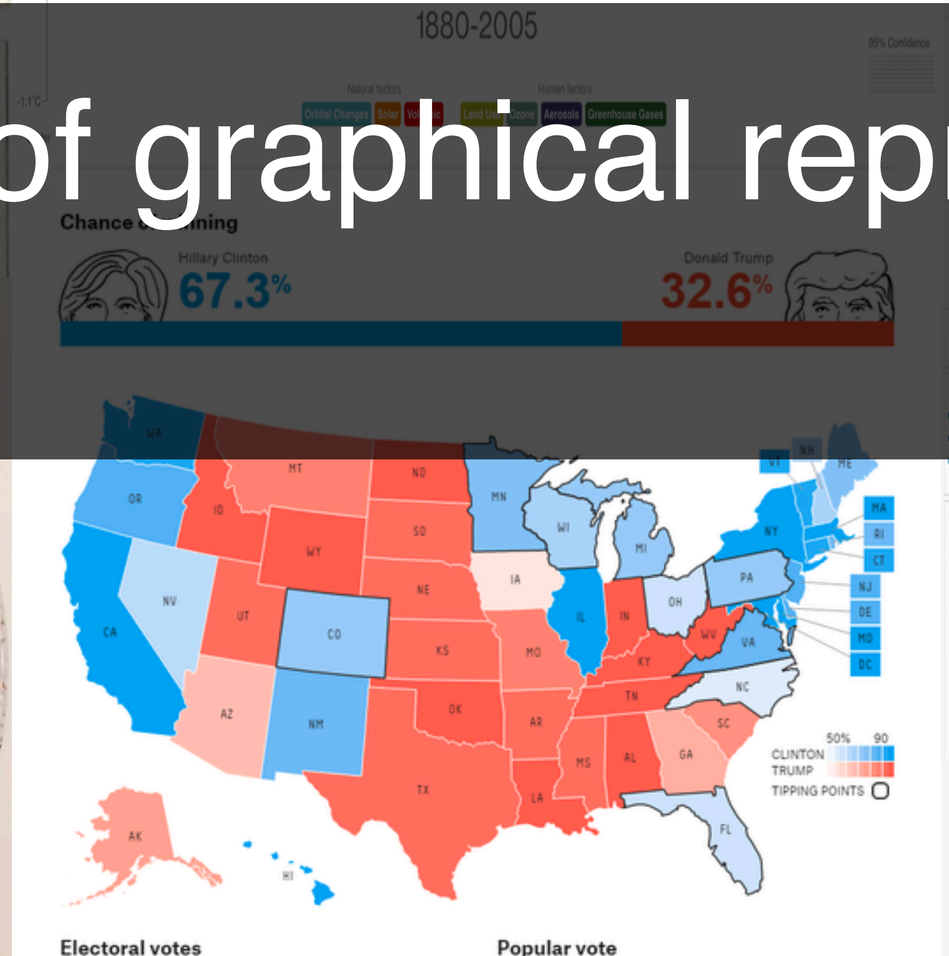


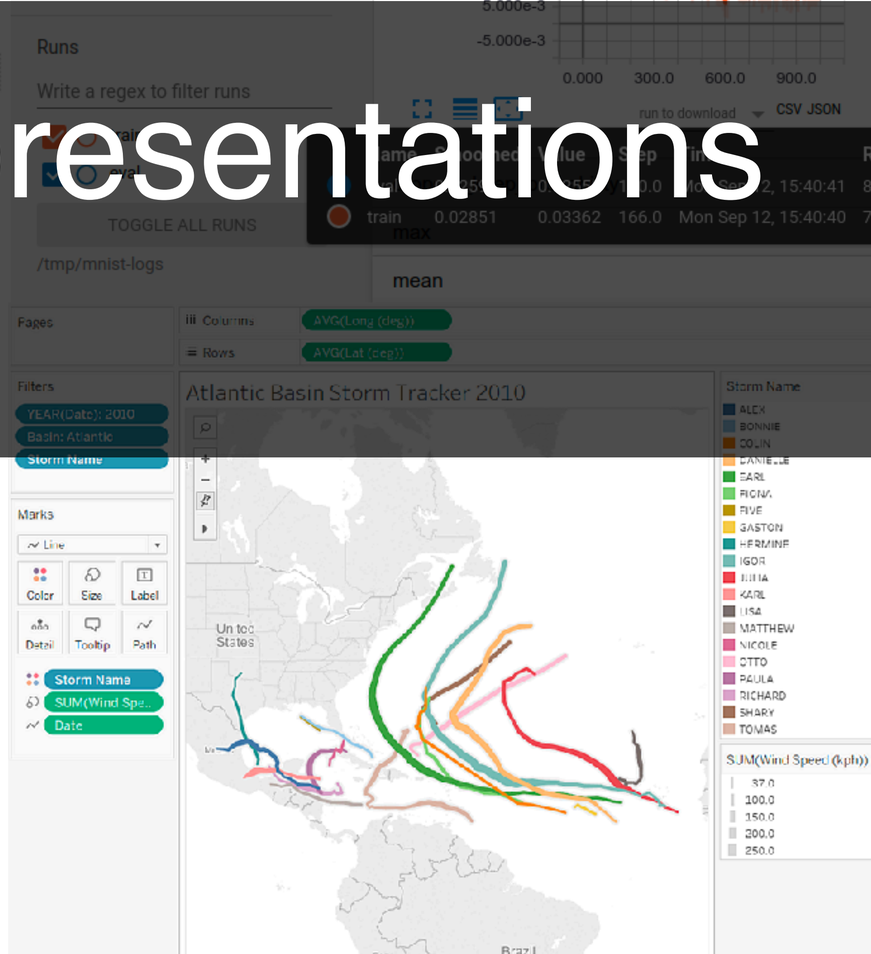
Any kinds of graphical representations



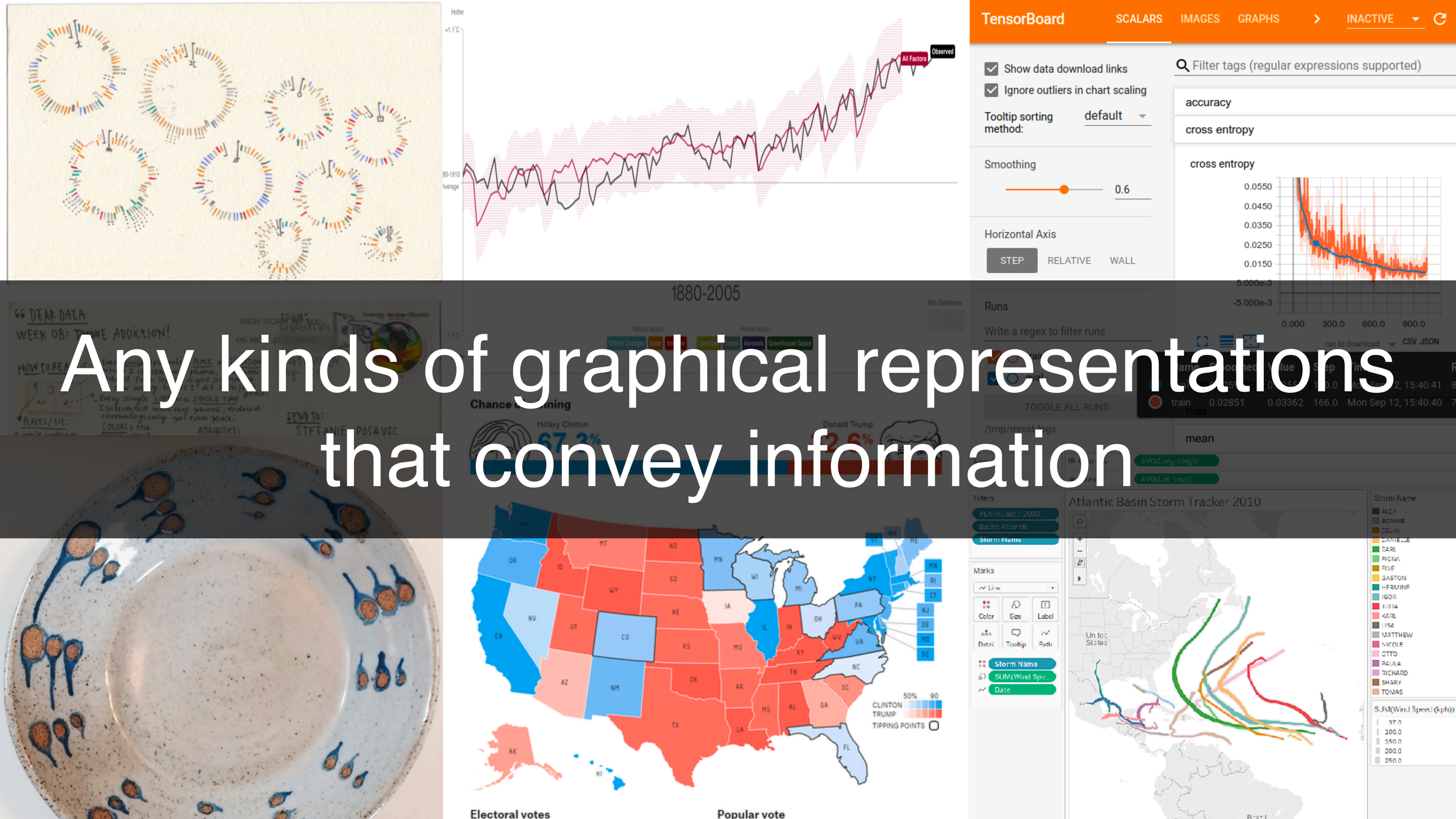








Any kinds of graphical representations that convey information



“Transformation of the **symbolic** into the **geometric**”
—McCormick et al. 1987

“... finding the **artificial memory** that best supports our natural means of perception.”
—Bertin 1967

“… visual representations of data to **amplify cognition**.”
—Card, Mackinlay, & Shneiderman 1999

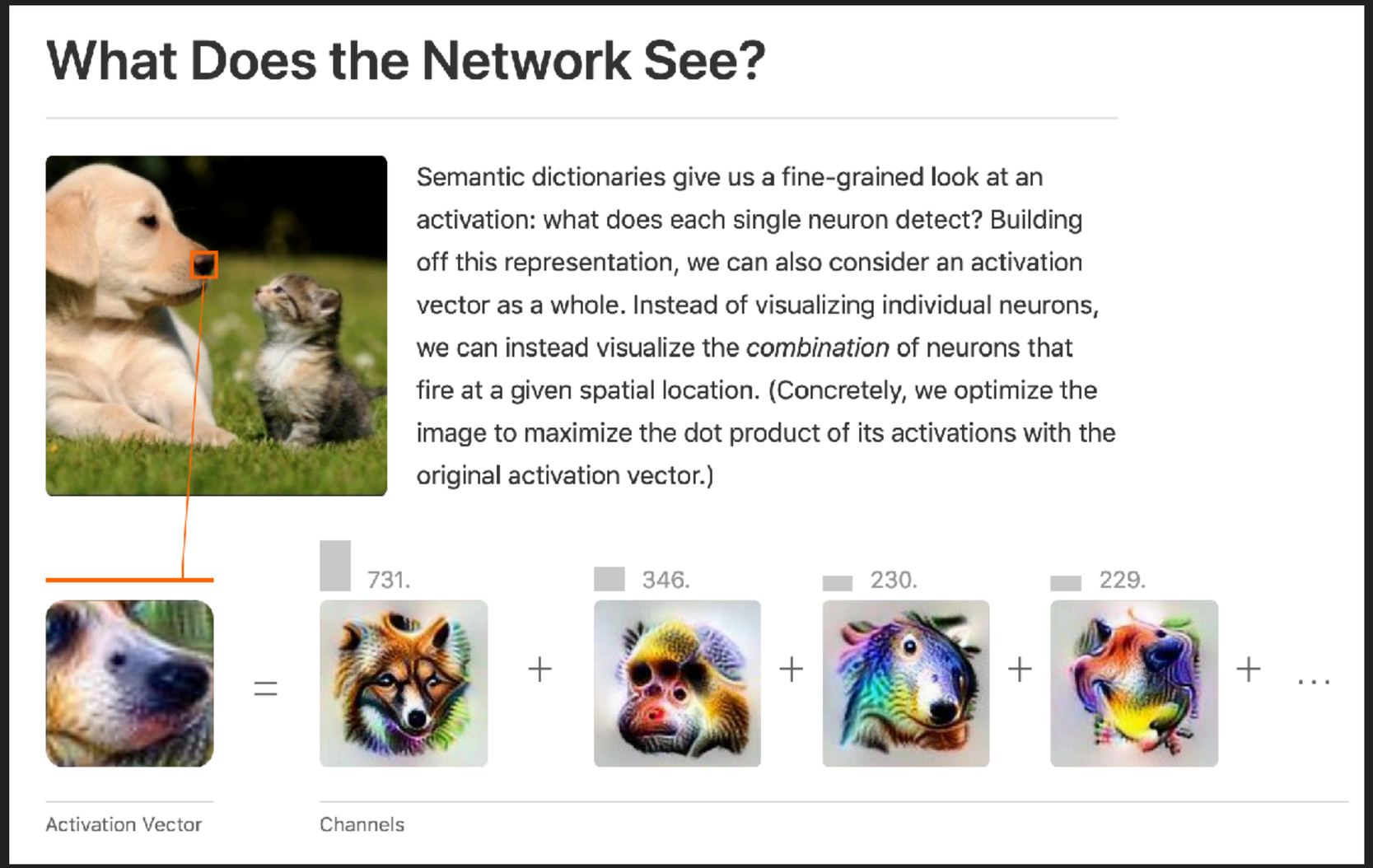
Often also known as …

- Information visualization
- Data visualization
- Information graphics
- Data graphics
- Information design

What about natural images?

Natural scene images

Information visualization



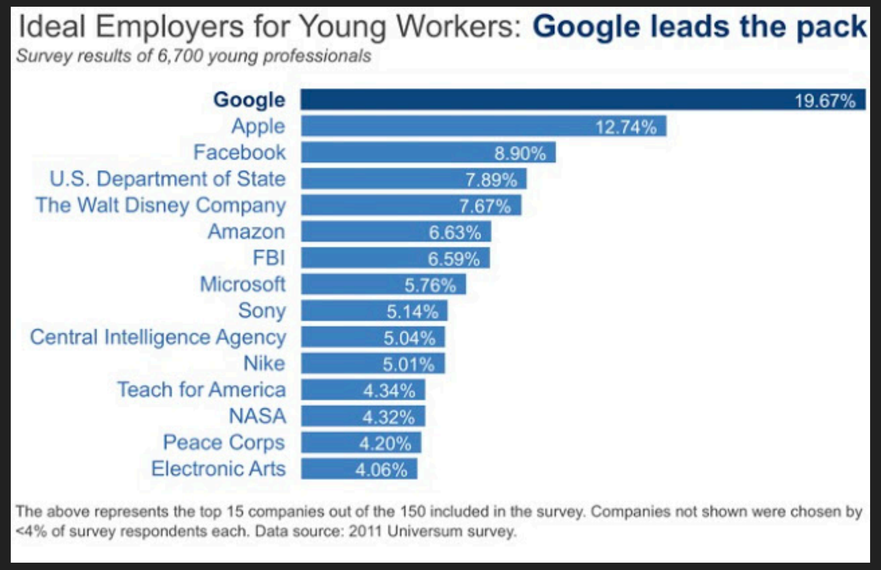
What about tables?

What's Your Dream Company?

Consulting firm Universum asked some 6,700 young professionals with one to eight years of work experience to choose their ideal employers from a list of firms. Here's the percentage of respondents that chose each company.

<< first < prev 1 2 3 next > last >>

Company	Rank	Percent
Google	1	19.67%
Apple	2	12.74%
Facebook	3	8.90%
U.S. Department of State	4	7.89%
The Walt Disney Company	5	7.67%
Amazon	6	6.63%
FBI	7	6.59%
Microsoft	8	5.76%
Sony	9	5.14%
Central Intelligence Agency	10	5.04%
Nike	11	5.01%
Teach for America	12	4.34%



http://www.storytellingwithdata.com/blog/2011/11/visual-battle-table-vs-graph

Why is it Important?

Anscombe's Quartet

A		B		C		D	
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	6.8

Summary Statistics

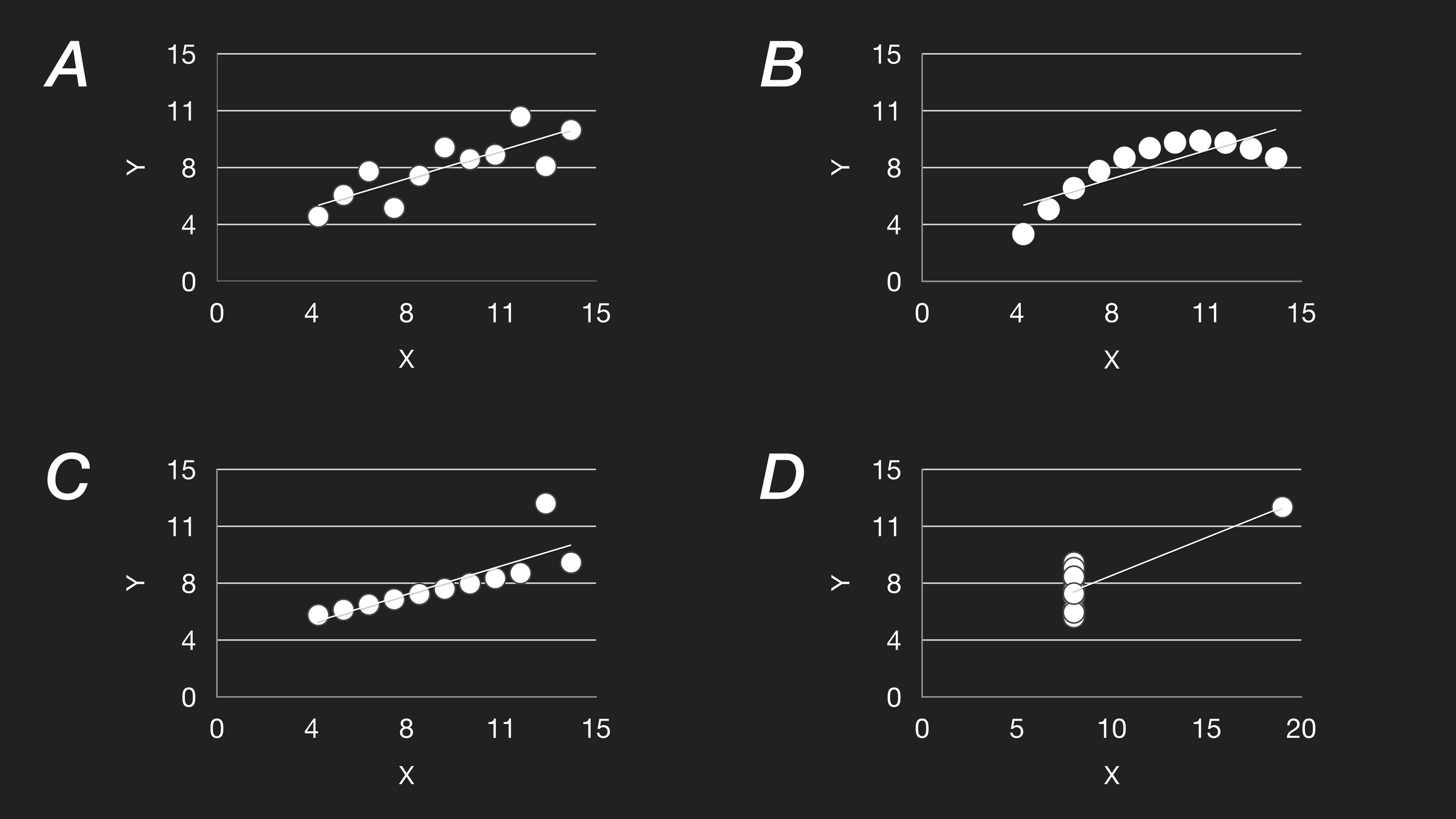
$u_X = 9.0$ $\sigma_X = 3.317$

$u_Y = 7.5$ $\sigma_Y = 2.03$

Linear Regression

$Y = 3 + 0.5\,X$

$R^2 = 0.67$


A
15
11
8
4
0
Y
0
4
8
11
15
X
B
15
11
8
4
0
Y
0
4
8
11
15
X
C
15
11
8
4
0
Y
0
4
8
11
15
X
D
15
11
8
4
0
Y
0
5
10
15
20
X

“

...make both calculations and graphs. Both sorts of output should be studied; each will contribute to understanding.

F. J. Anscombe, 1973

X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

All distinct datasets with same statistical properties

Matejka & Fitzmaurice 2017

The purpose of computing is insight, not numbers.

R.W. Hamming, 1962

The purpose of visualization is insight, not pictures.

Ben Shneiderman

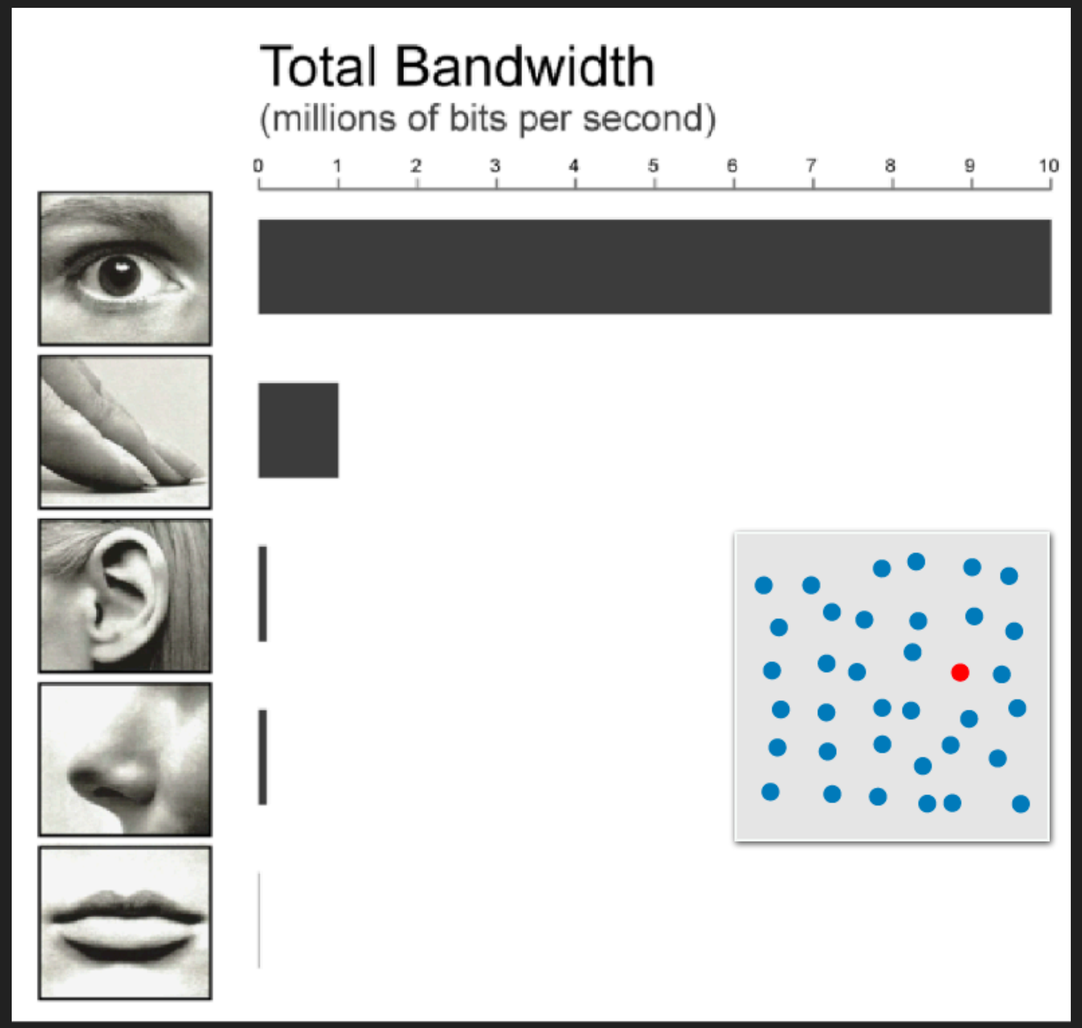
Why is it Important?

Not because of pretty pictures

Visualization
allows us to see the unseen

How Does It Help?

Leverage Visual Perception



- Vision is the most powerful channel for processing information.
- Parallel vs serial pattern processing (e.g., visual popout v.s. sound streaming)
- Strong pattern-recognition capabilities.

Support External Cognition

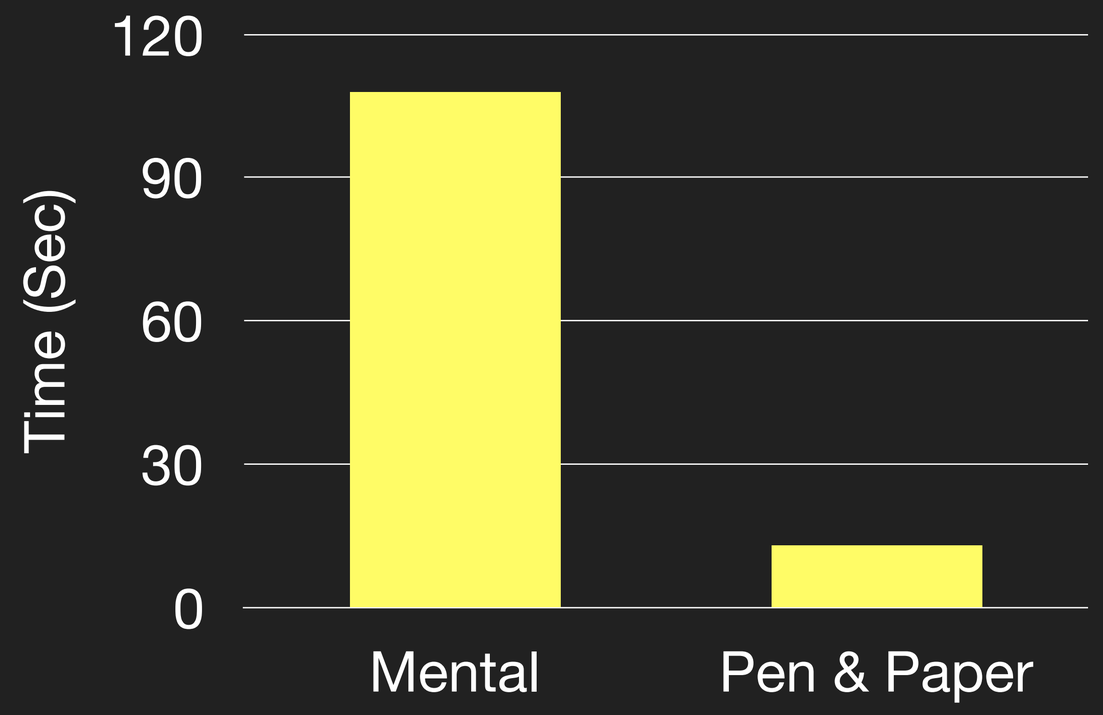
A visual memory aid to overcome our limited cognitive capabilities

$$\begin{array}{r} 34 \\ \times\ 72 \\ \hline \end{array}$$

Support External Cognition

A visual memory aid to overcome our limited cognitive capabilities

```
   34
 x 72
 ----
   68
 2380
 ----
 2448
```



Three
Functions of Visualization

Three Functions of Visualization

1. Record Information
2. Support Analytic Reasoning
3. Communicate Information to Others

Three Functions of Visualization

1. **Record Information**
2. Support Analytic Reasoning
3. Communicate Information to Others

~ 6200 BC

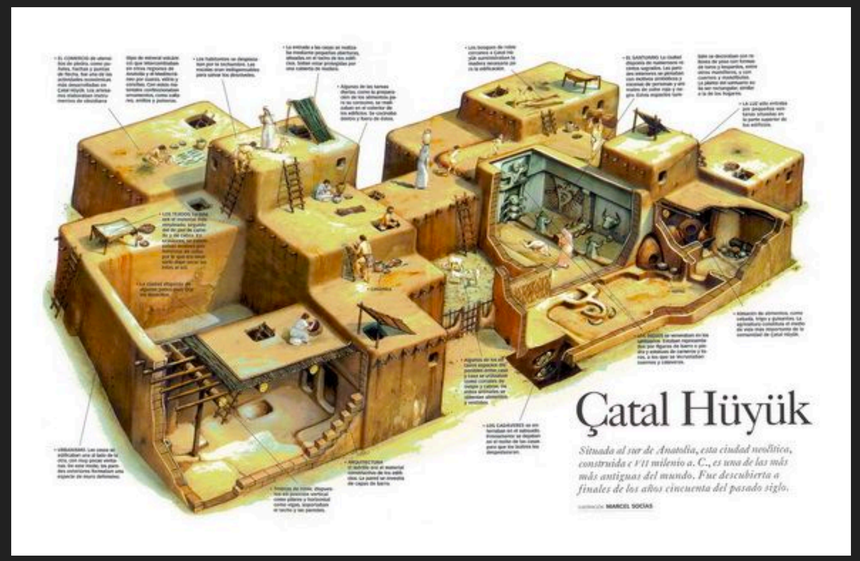
Konya Town Map

BC AD

~ 6200 BC

Konya Town Map

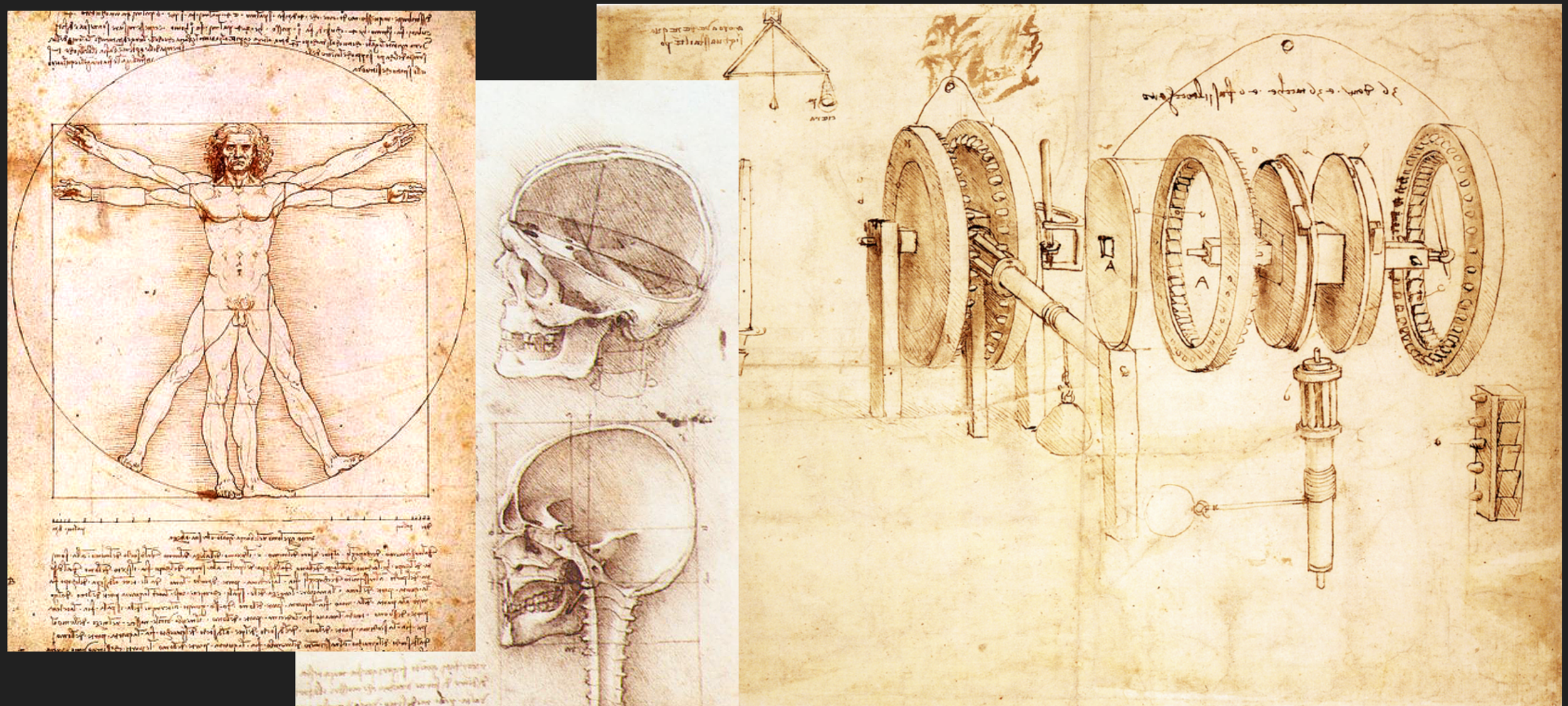
BC | AD



Leonardo da Vinci' Sketches - 1485

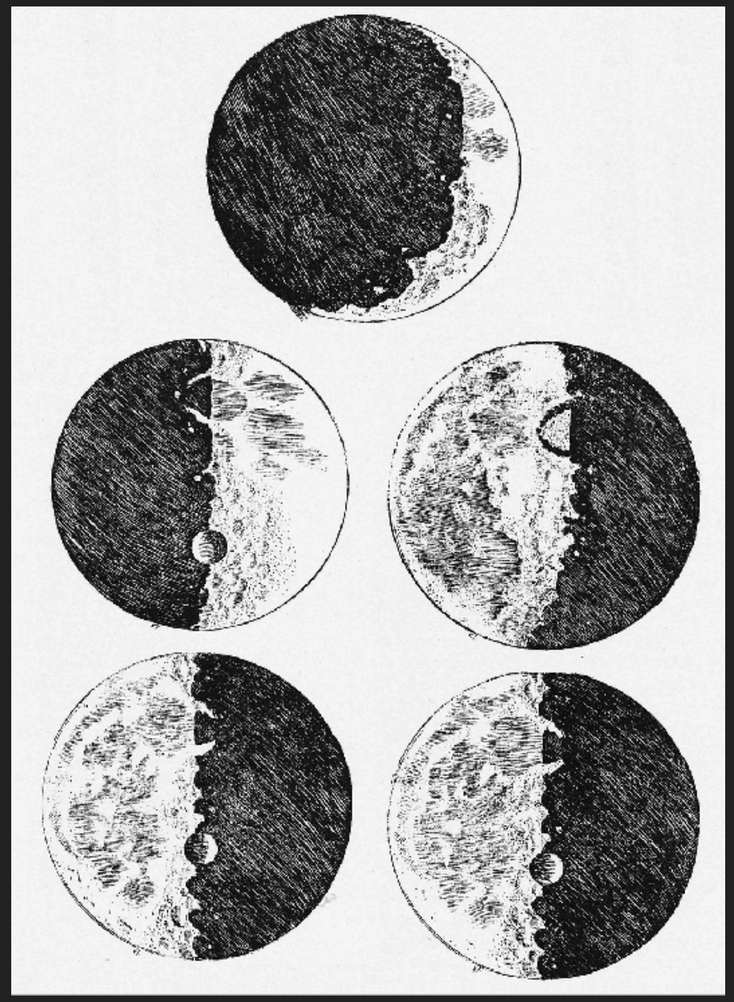
AD

500 1000 1500 2000

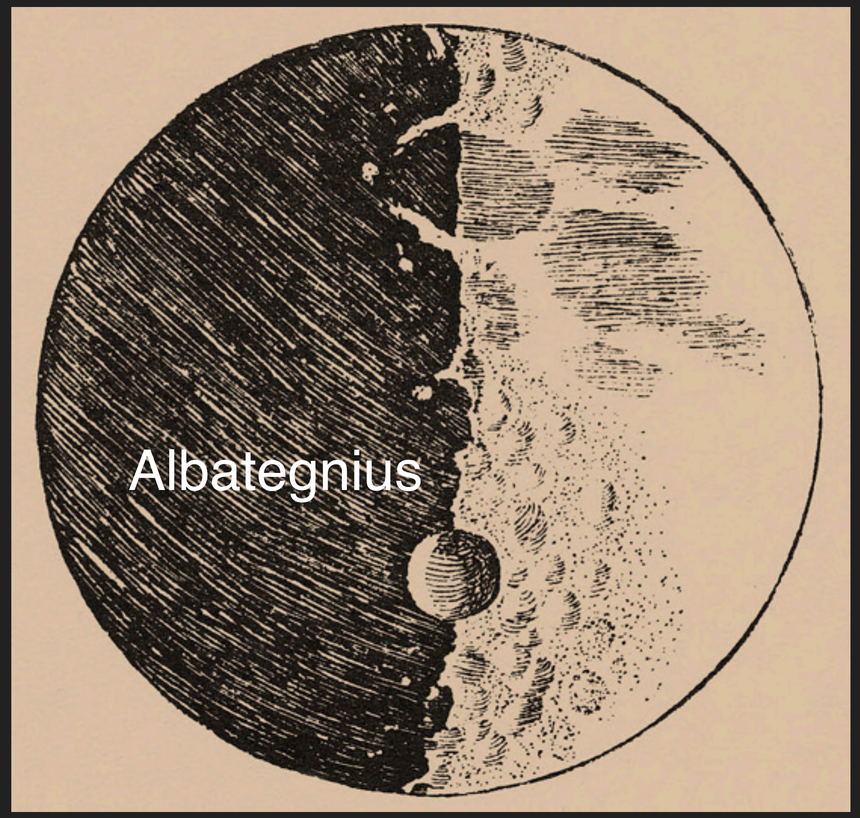


Drawings of the Moon, Galileo Galilei - 1609

1000 1500 2000



Phases of the Moon

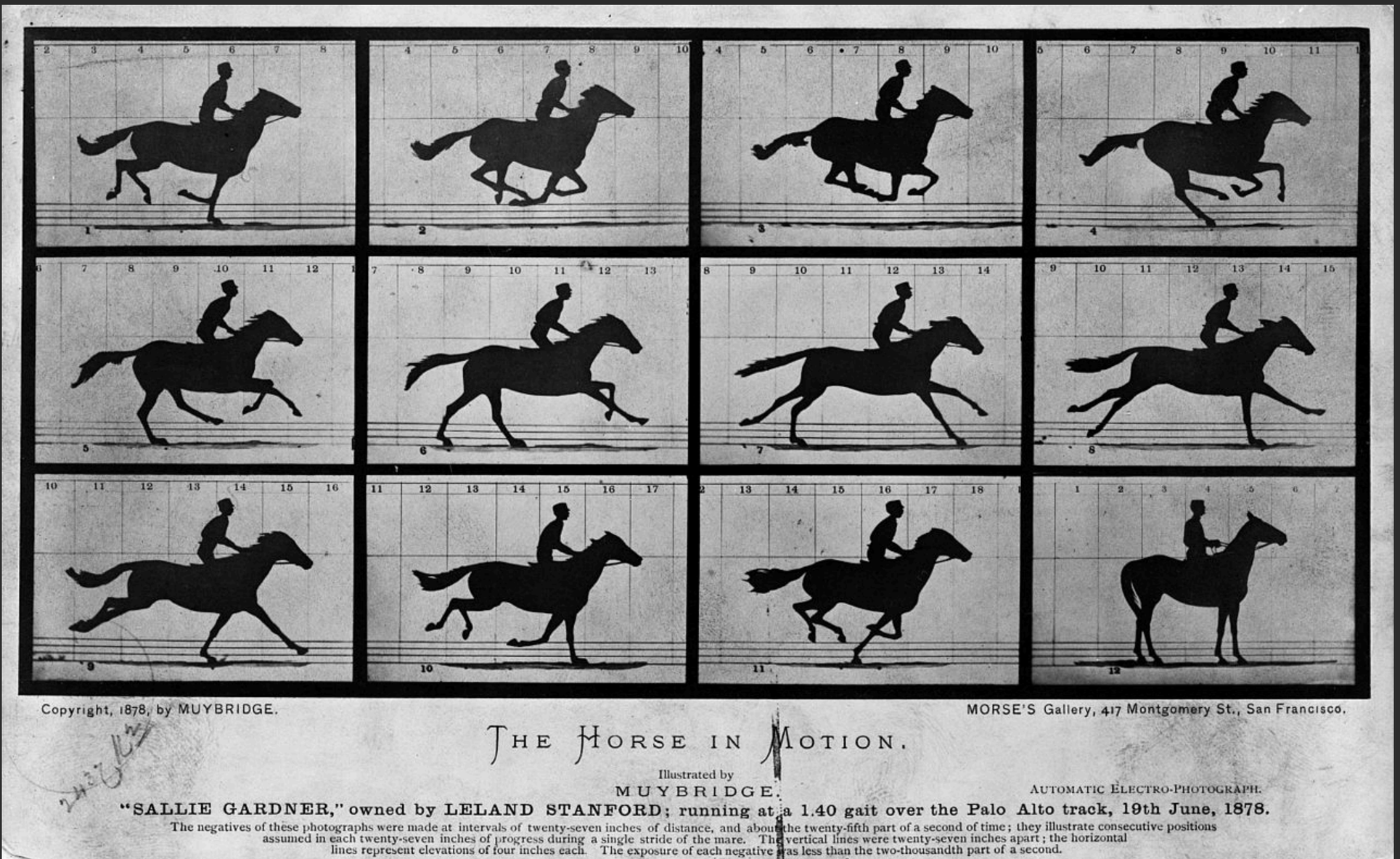


Sketch

Satellite image

The House In Motion, Muybridge - 1878

1000 1500 2000

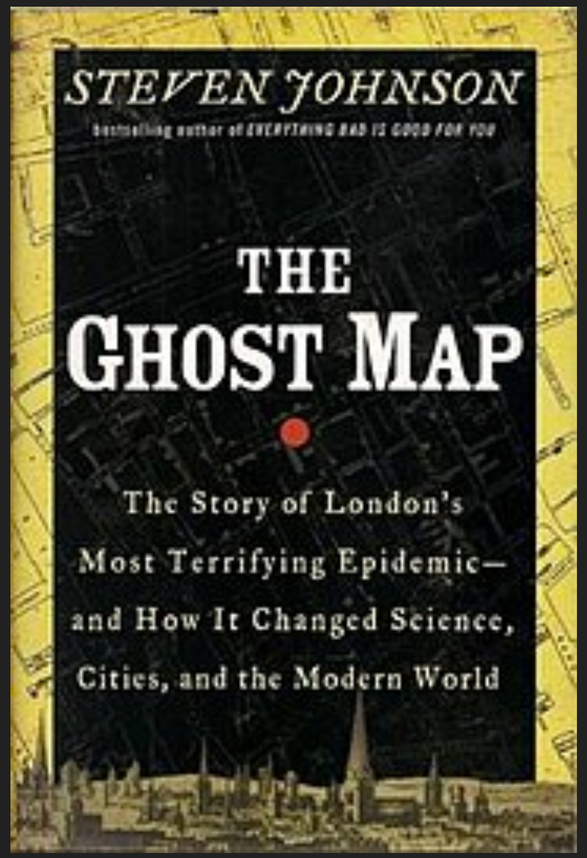


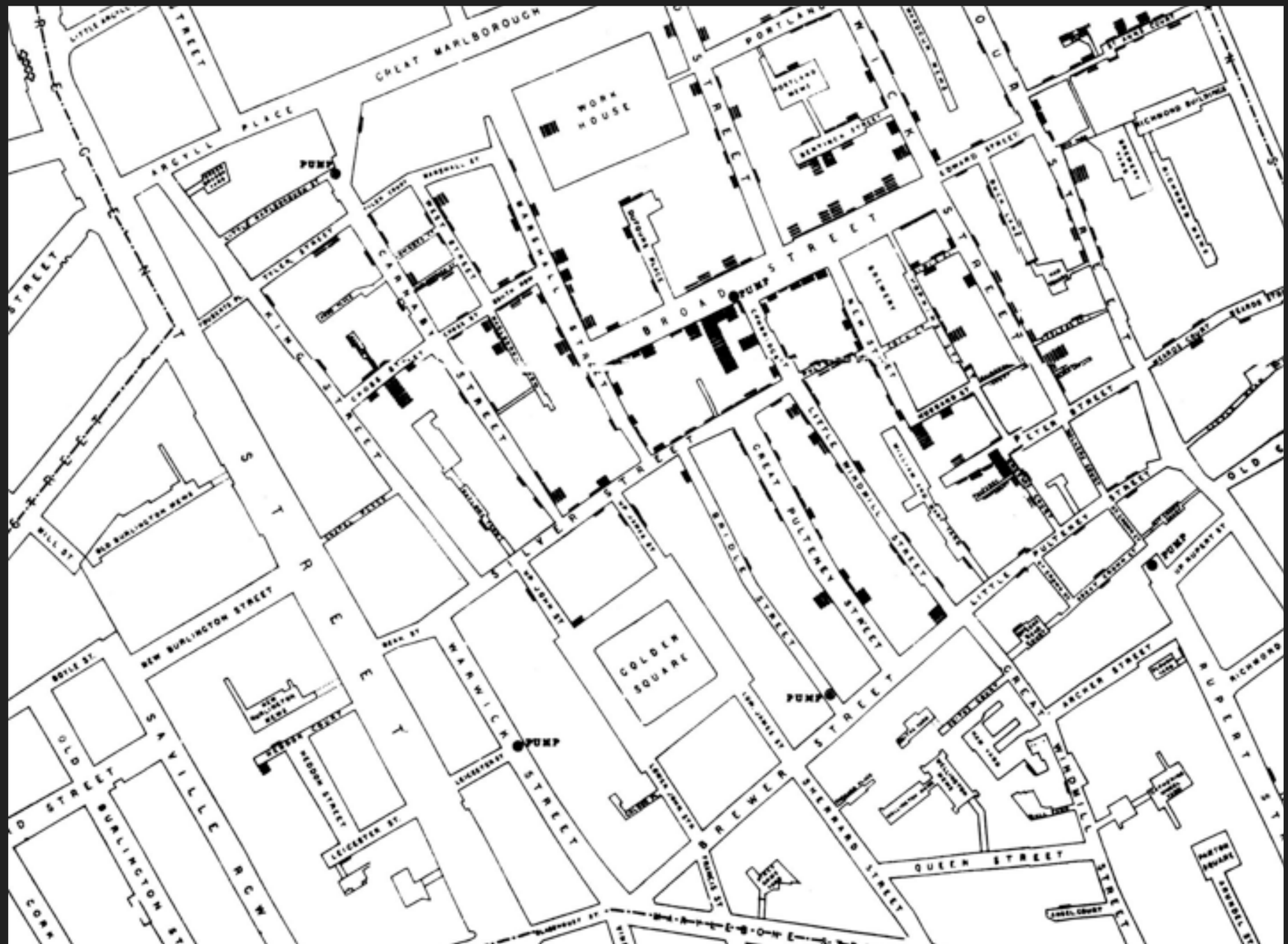
Three Functions of Visualization

1. Record Information
2. **Support Analytic Reasoning**
3. Communicate Information to Others

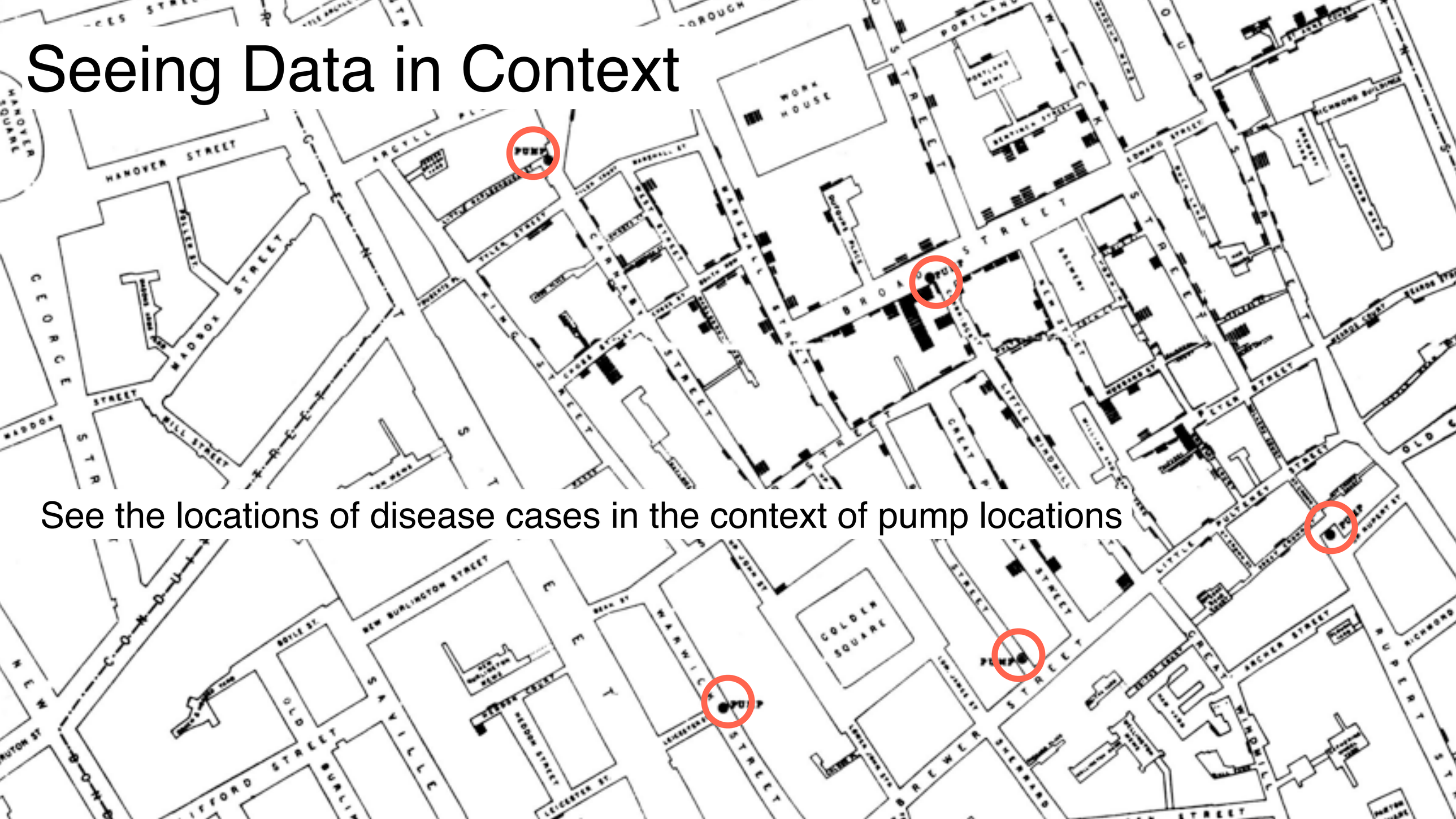
John Snow' Cholera Map - 1854

1000 1500 2000

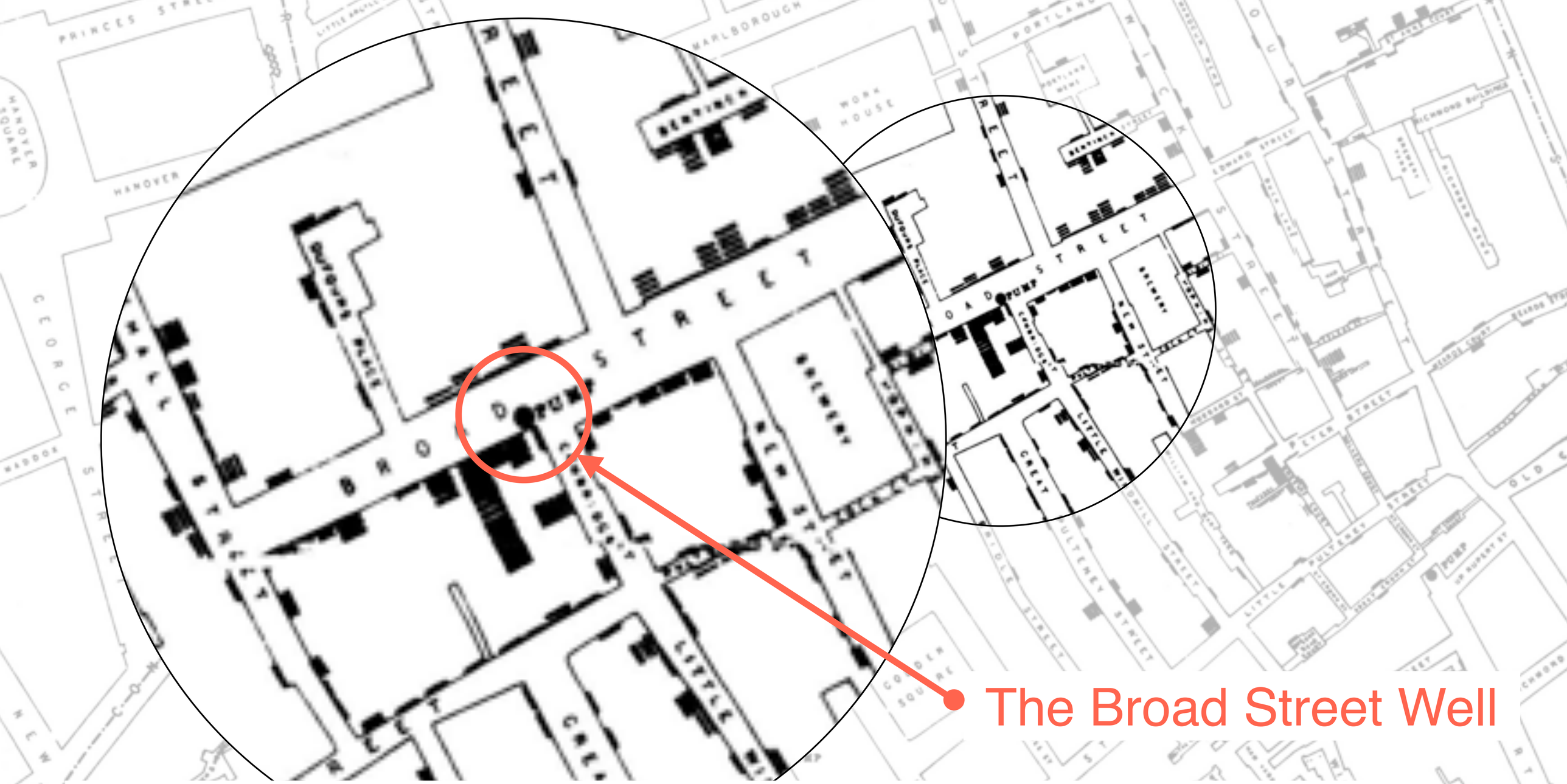




Seeing Data in Context



See the locations of disease cases in the context of pump locations



Allows us to reason about the relationship between the disease and the cause.

Feynman Diagram - 1948

1000 1500 2000

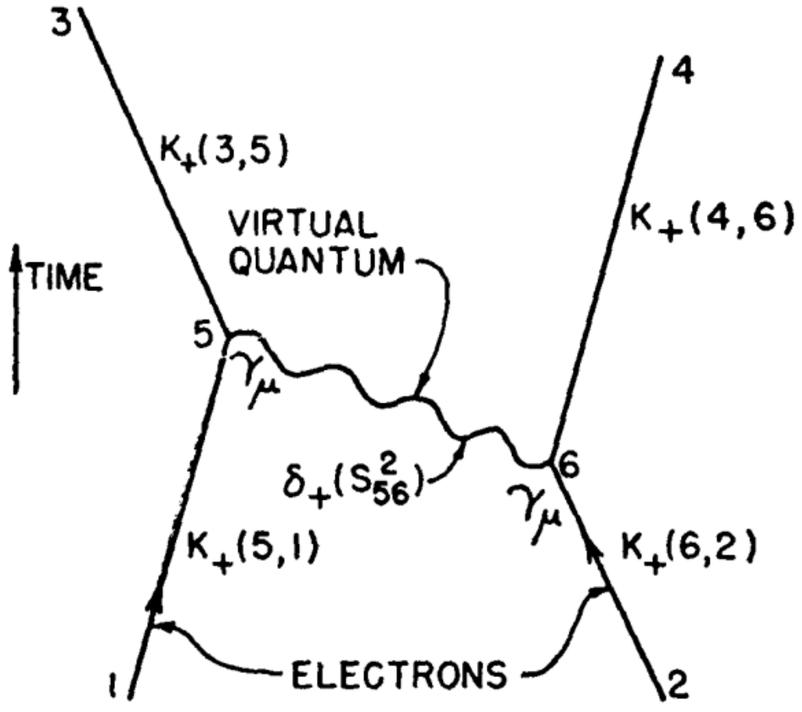
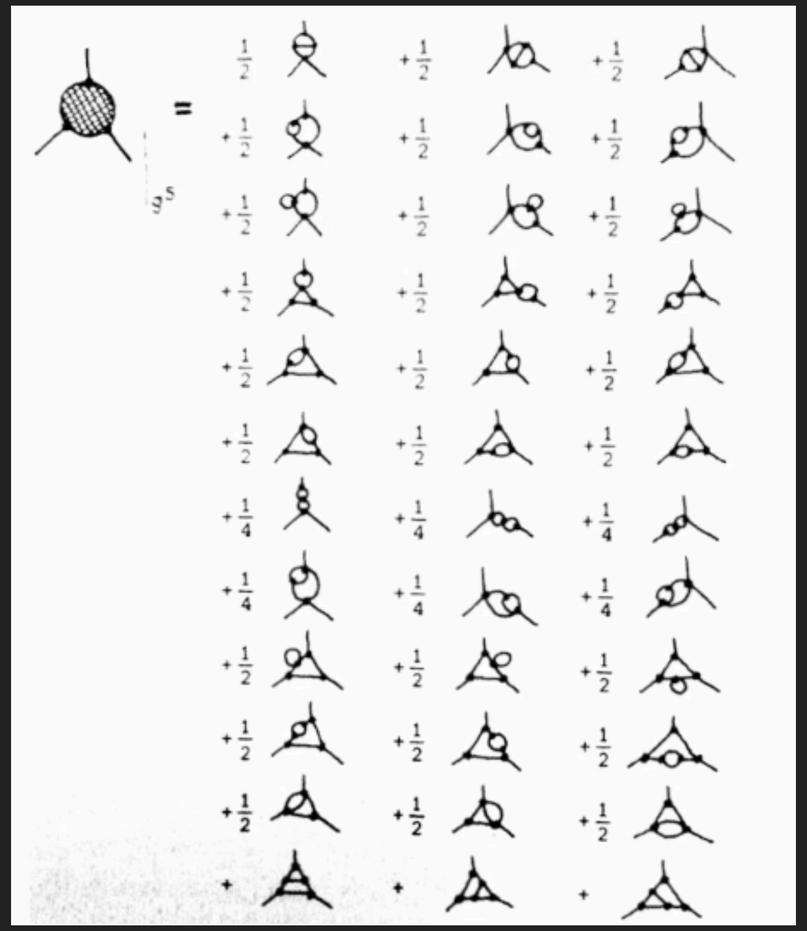
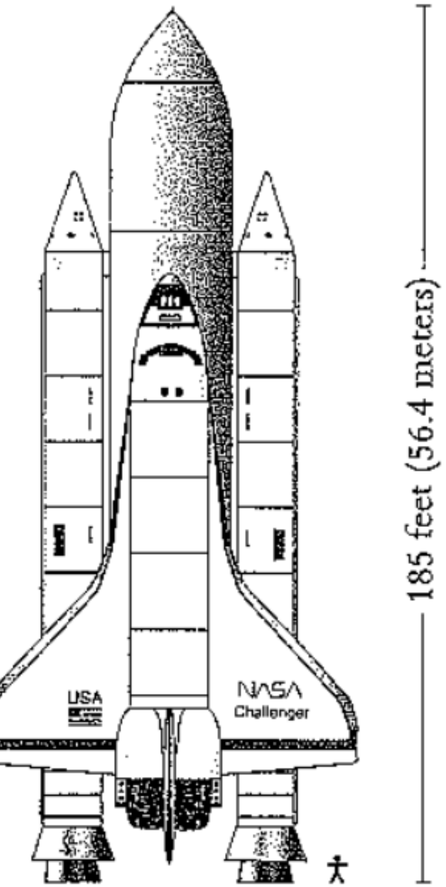


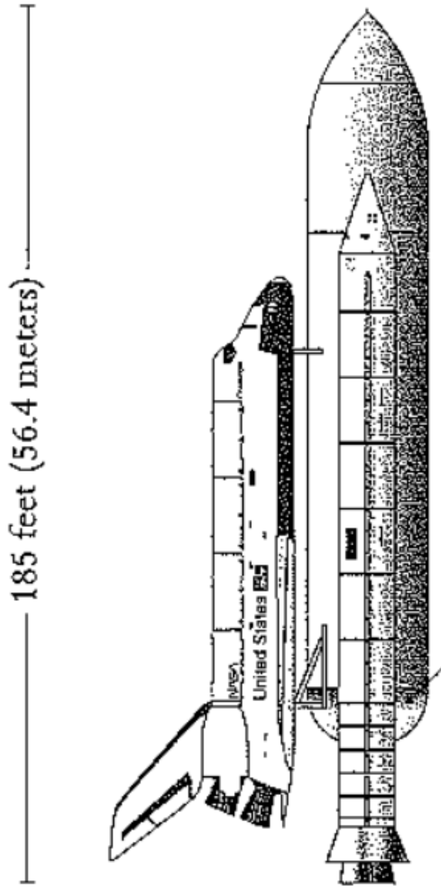
Figure 3. Electron-electron scattering is described by one of the earliest published Feynman diagrams (fea-

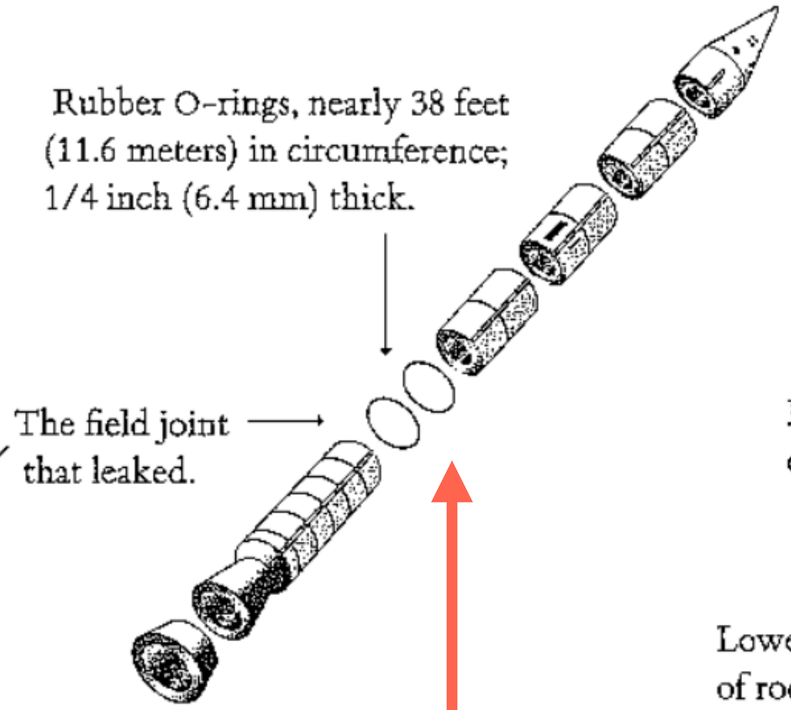


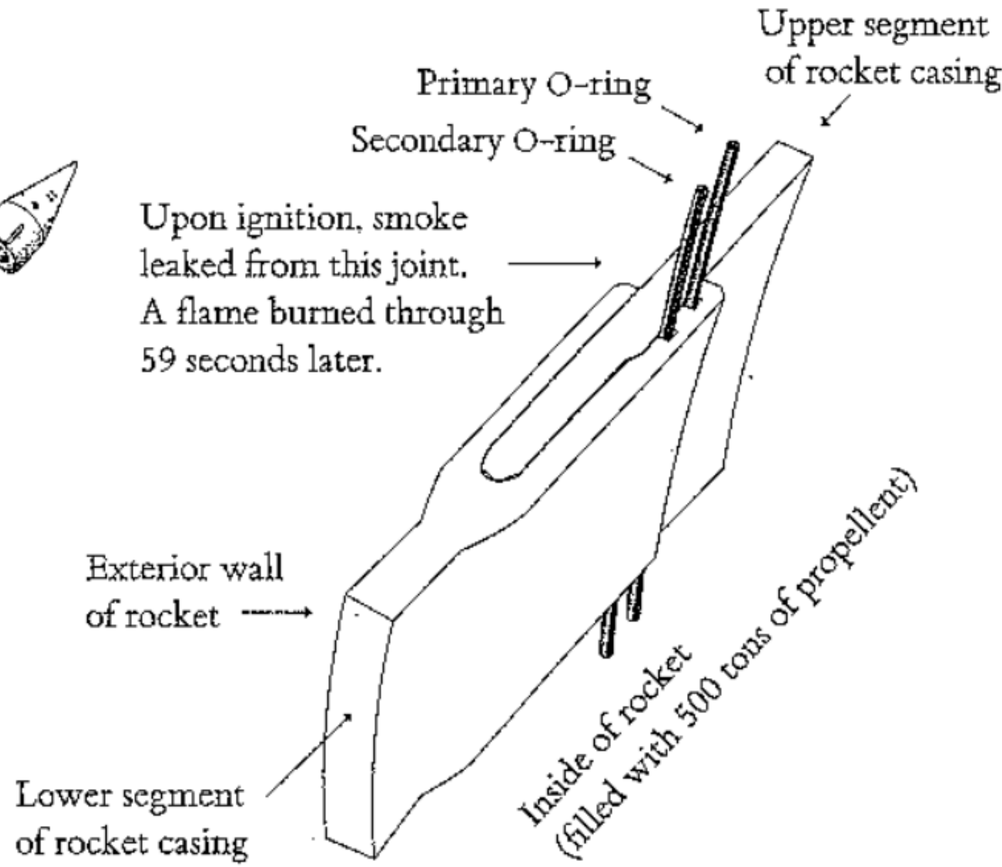
"since the middle of the 20th century, theoretical physicists have increasingly turned to this tool to help them undertake critical calculations" — David Kaiser

Space Shuttle Challenger Disaster 1986









Rubber O-rings
had problems with cold temperatures.

One of original reports sent to NASA officials before launch

HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

	SRM No.	Cross Sectional View: Erosion Depth (in.)	Cross Sectional View: Perimeter Affected (deg)	Cross Sectional View: Nominal Dia. (in.)	Top View: Length Of Max Erosion (in.)	Top View: Total Heat Affected Length (in.)	Clocking Location (deg)
61A LH Center Field**	22A	None	None	0.280	None	None	36°--66°
61A LH ~~CENTER~~ FIELD**	22A	NONE	NONE	0.280	NONE	NONE	338°-18°
51C LH Forward Field**	15A	0.010	154.0	0.280	4.25	5.25	163
51C RH Center Field (prim)***	15B	0.038	130.0	0.280	12.50	58.75	354
51C RH Center Field (sec)***	15B	None	45.0	0.280	None	29.50	354
41D RH Forward Field	13B	0.028	110.0	0.280	3.00	None	275
41C LH Aft Field*	11A	None	None	0.280	None	None	--
41B LH Forward Field	10A	0.040	217.0	0.280	3.00	14.50	351
STS-2 RH Aft Field	2B	0.053	116.0	0.280	--	--	90

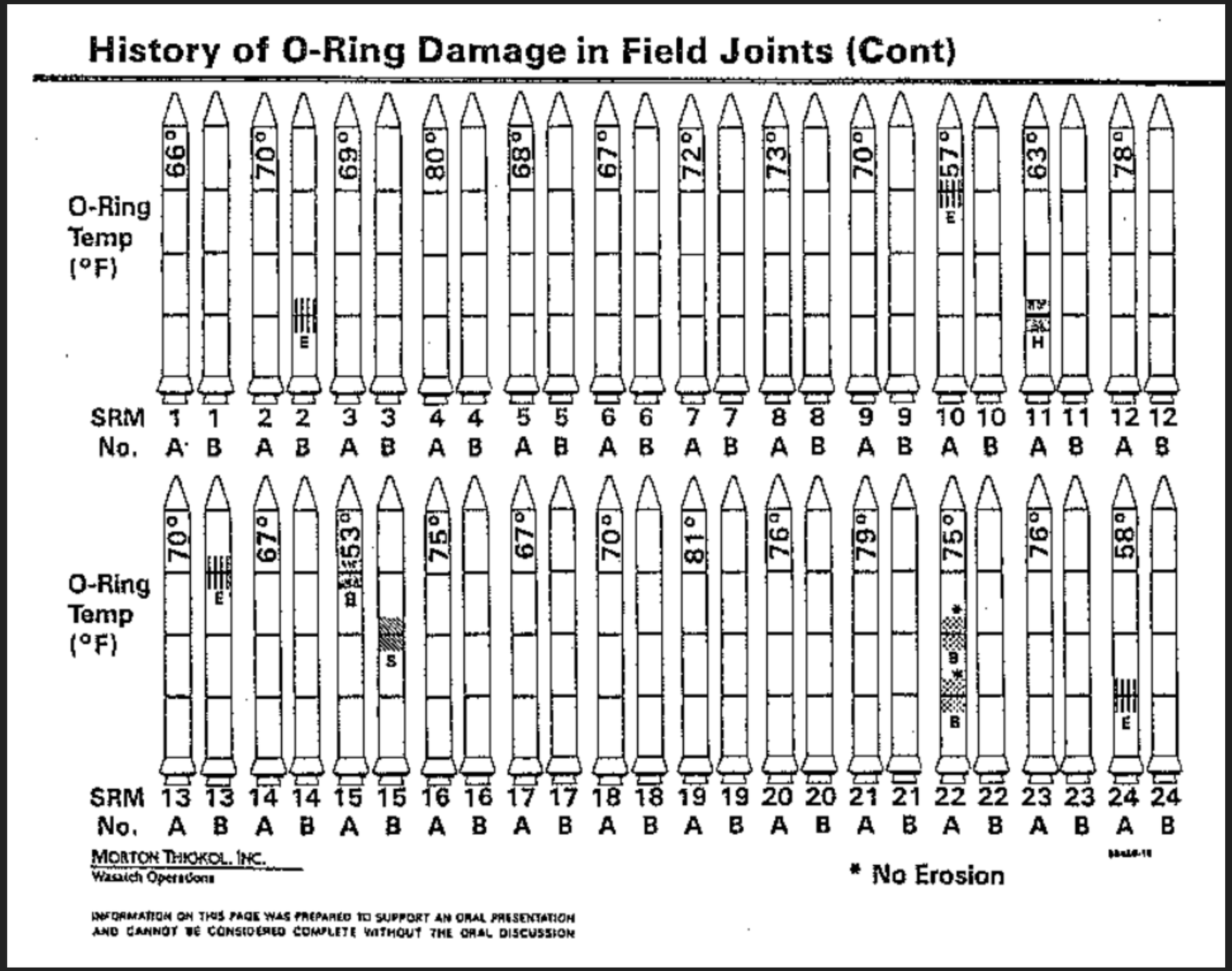
*Hot gas path detected in putty. Indication of heat on O-ring, but no damage.
**Soot behind primary O-ring.
***Soot behind primary O-ring, heat affected secondary O-ring.

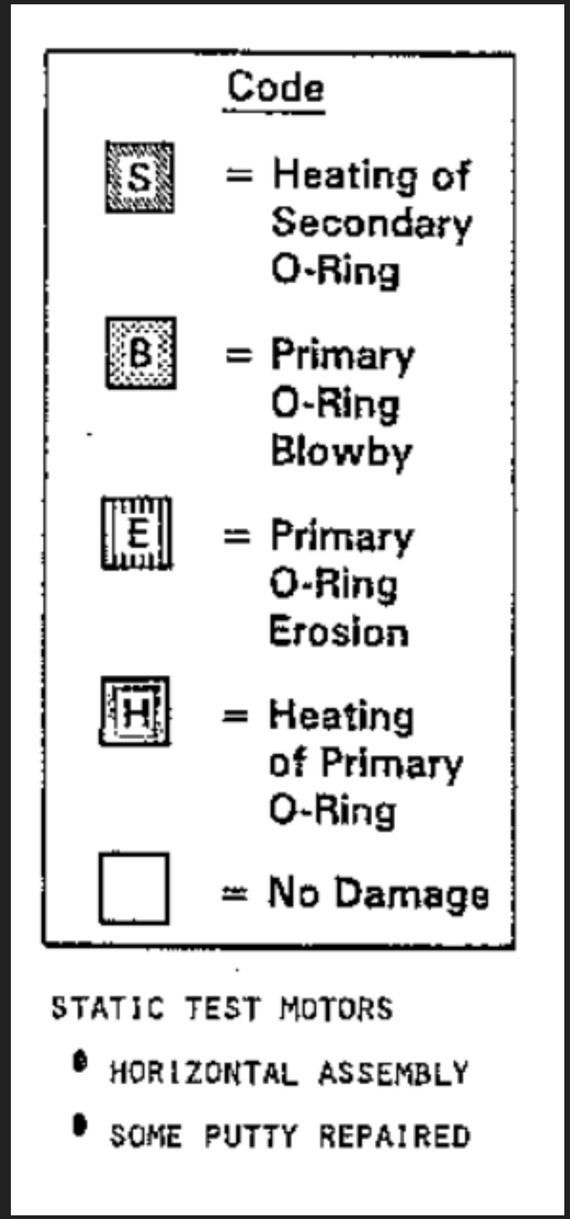
Clocking location of leak check port - 0 deg.

OTHER SRM-15 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY AND NO SOOT NEAR OR BEYOND THE PRIMARY O-RING.

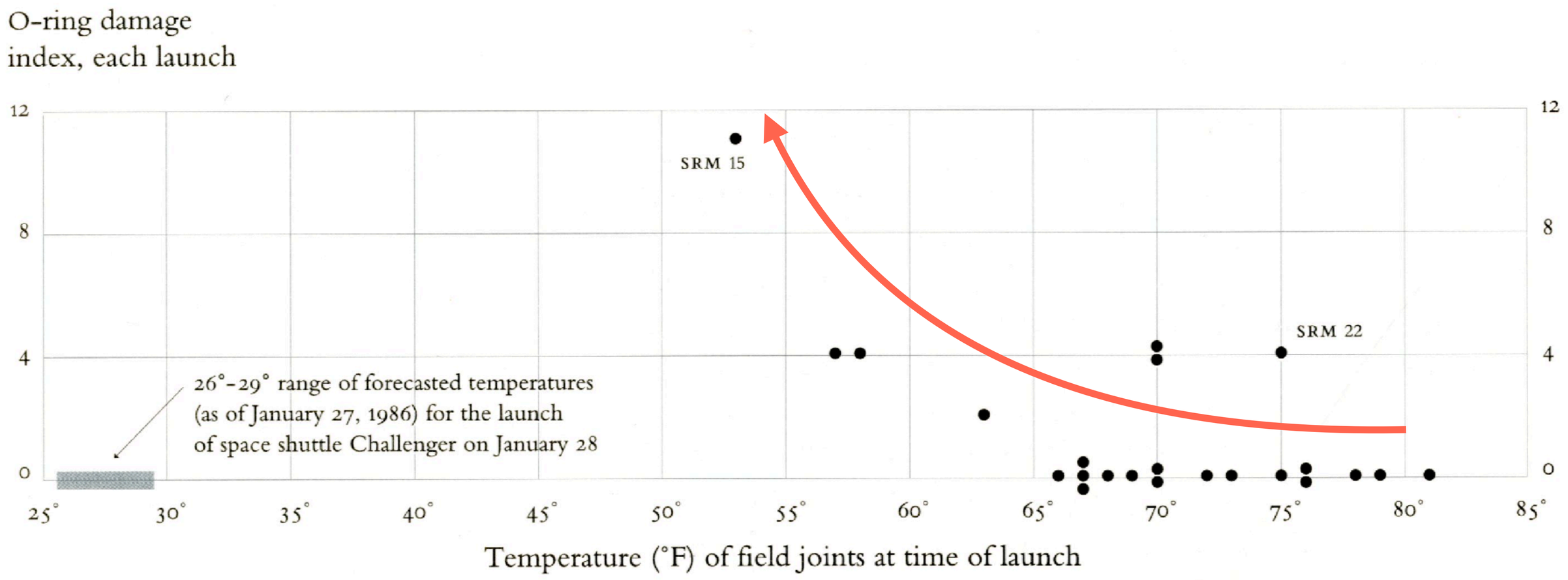
SRM-22 FORWARD FIELD JOINT HAD PUTTY PATH TO PRIMARY O-RING, BUT NO O-RING EROSION AND NO SOOT BLOWBY. OTHER SRM-22 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY.

Even after launch …





Use a right visualization to make a right decision

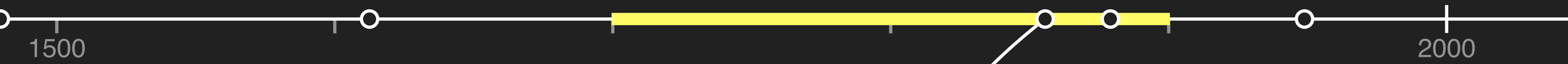


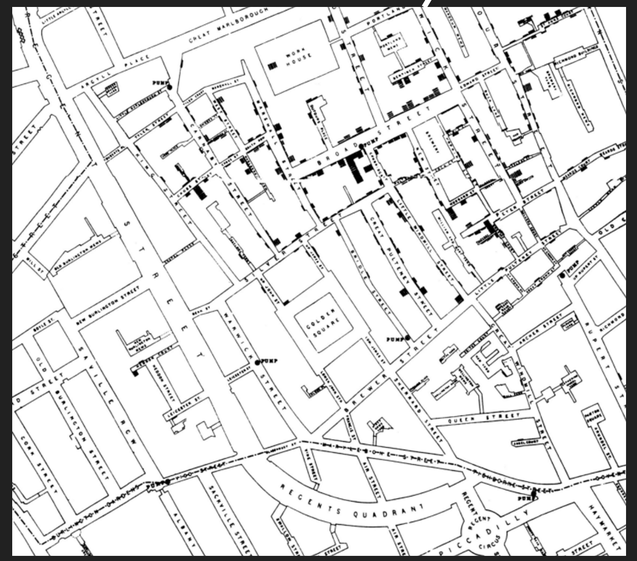
Edward Tufte 1997

Three Functions of Visualization

1. Record Information
2. Support Analytic Reasoning
3. **Communicate Information to Others**

Golden Age of Visualization - 1700 - 1900

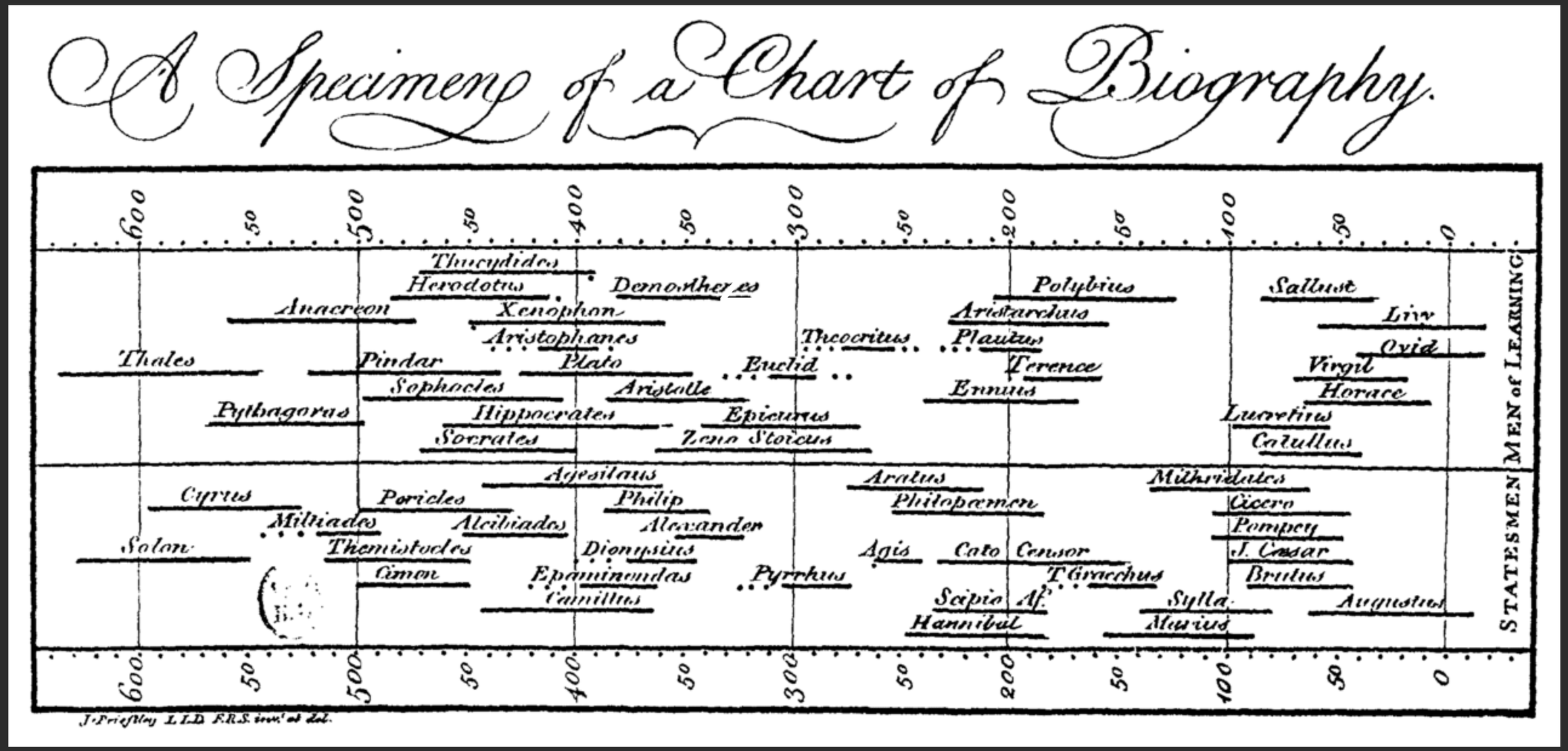




John Snow' Cholera Map

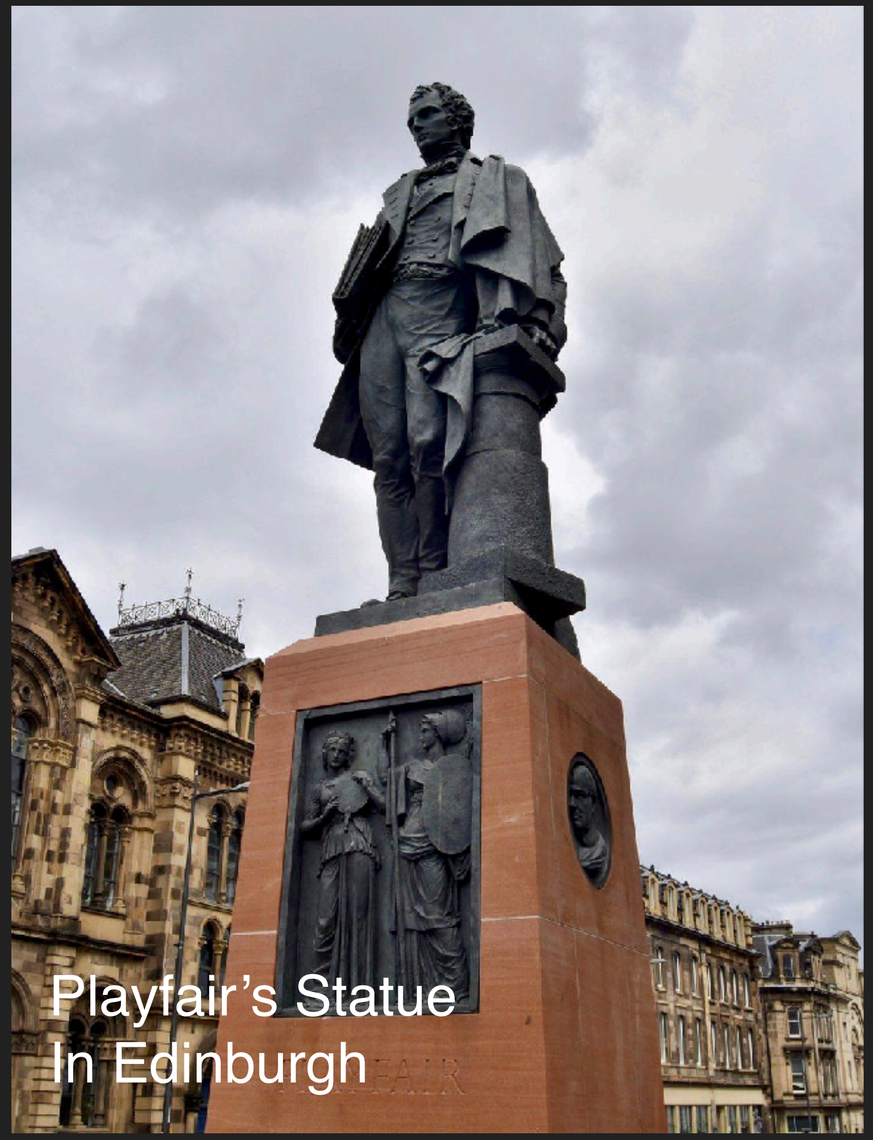
Joseph Priestley's Chart of Biography - 1765

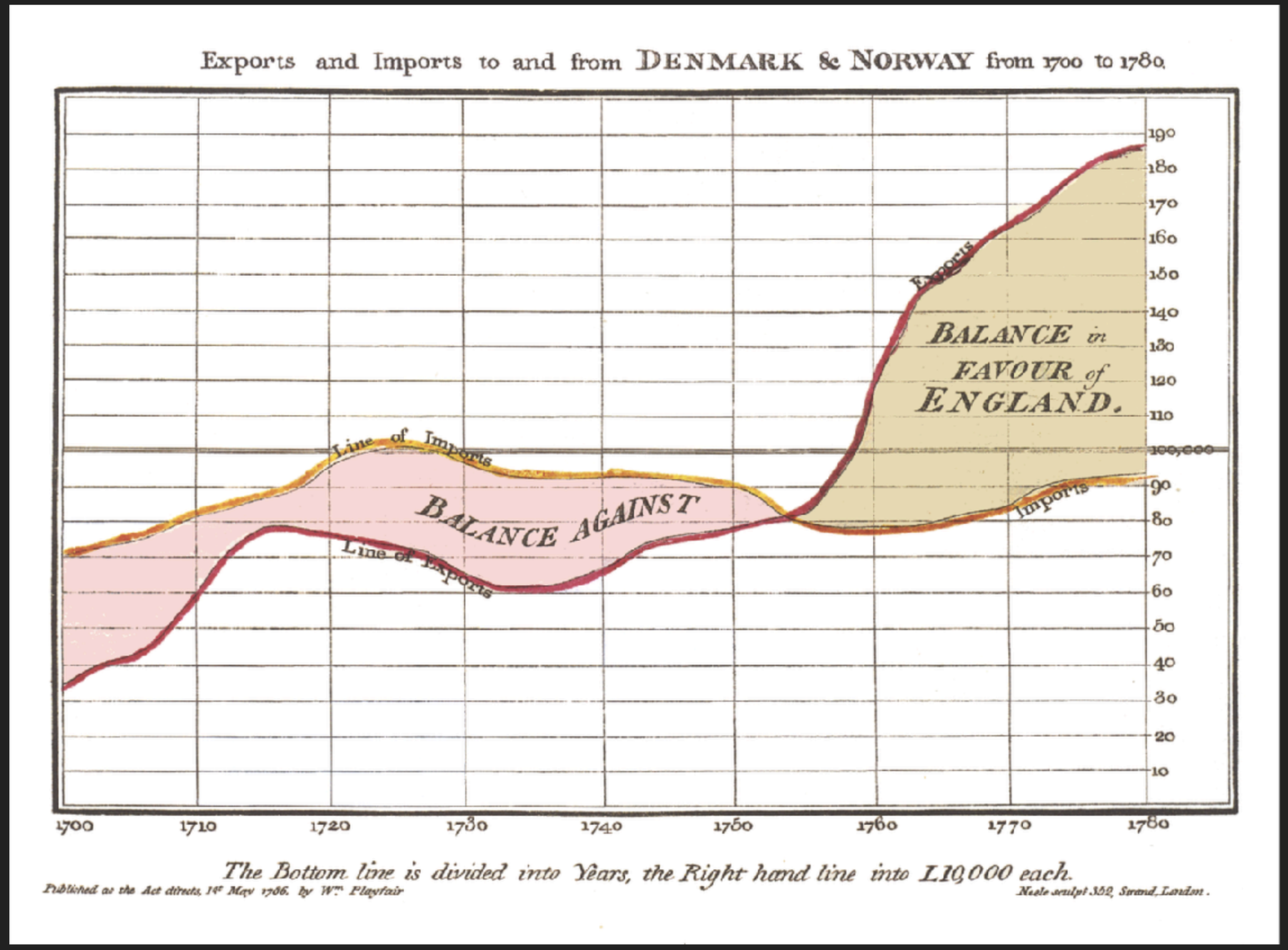
1500 2000



The Commercial and Political Atlas, William Playfair - 1786

1500 2000

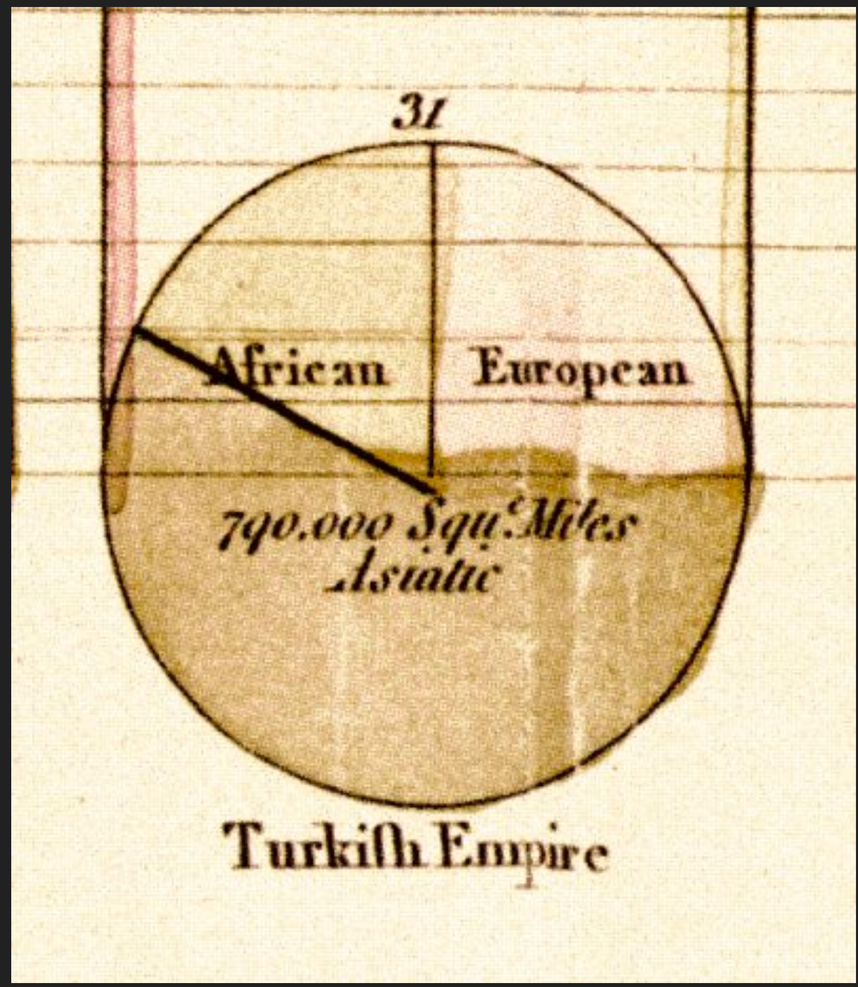




Statistical Breviary, William Playfair - 1801

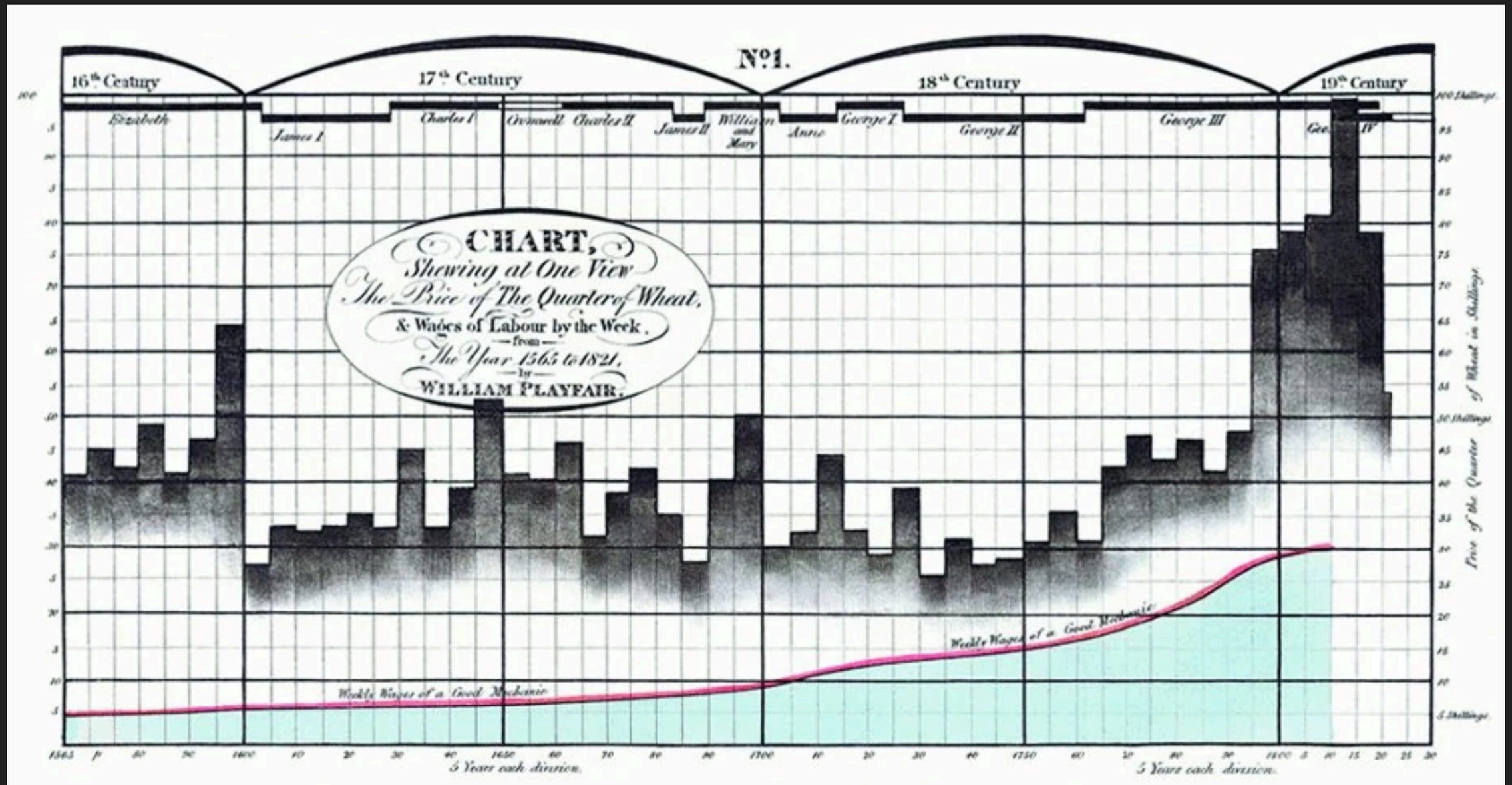
1500

2000



Agricultural Distresses - William Playfair - 1821

1500 2000

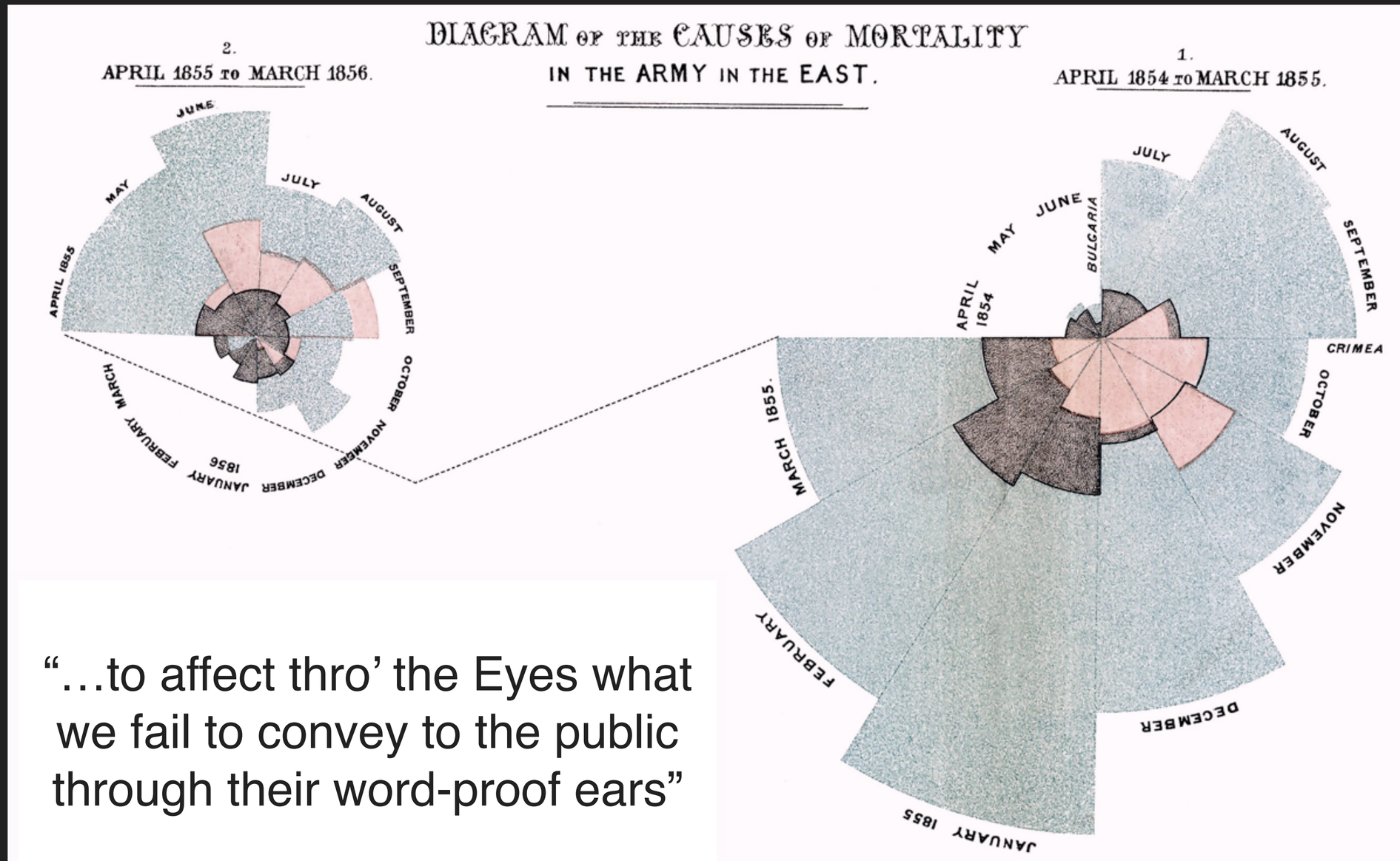


Coxcombs of the British casualties in the Crimean War, Florence Nightingale - 18

1500 2000

Crimean War Deaths

- Death from battle wounds
- Death from other causes
- Death from disease



"…to affect thro' the Eyes what we fail to convey to the public through their word-proof ears"

Florence Nightingale is a Design Hero

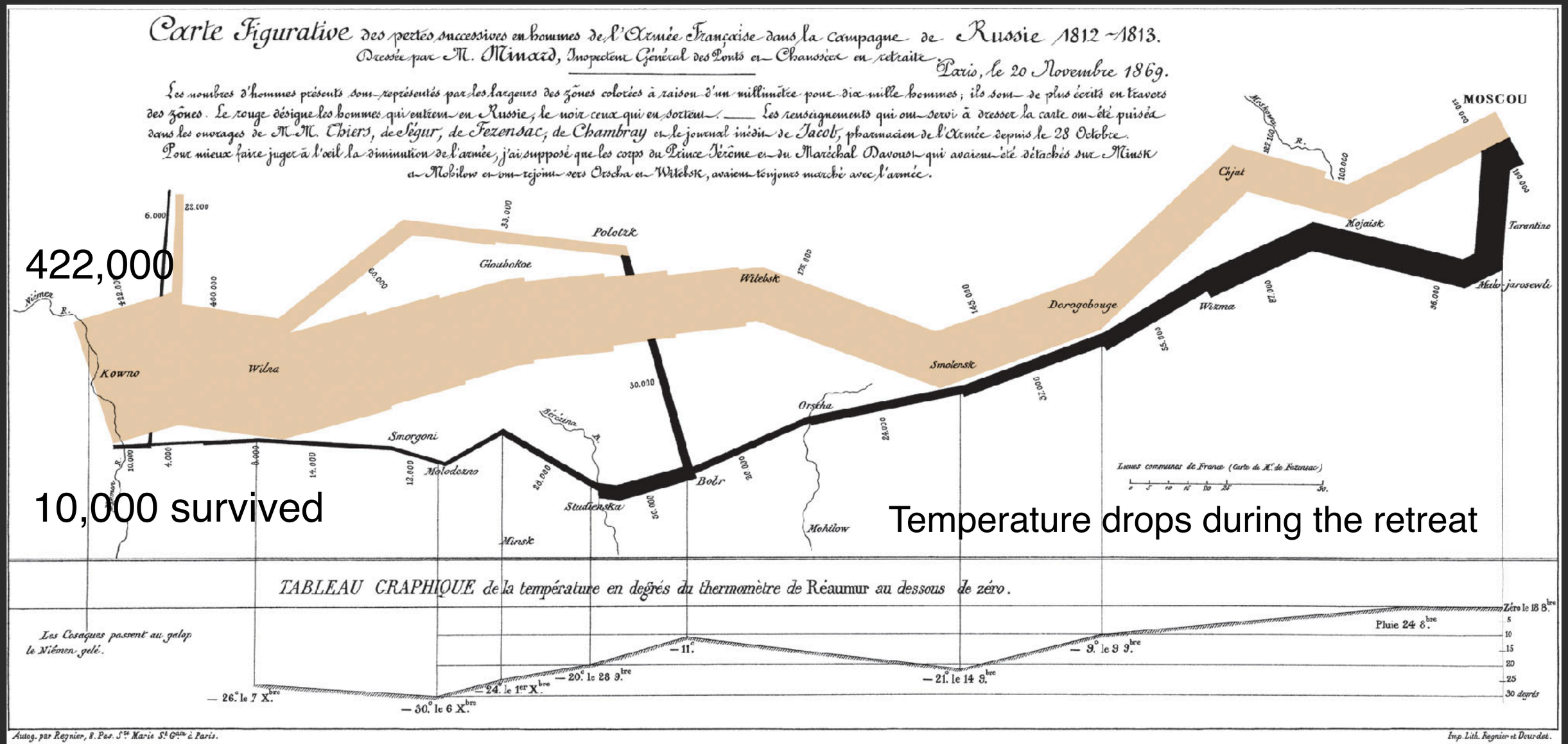
May her light forever shine bright.

RJ Andrews Follow
Jul 15 · 13 min read ★

Founder of modern nursing. Feminist champion. Celebrity entrepreneur. Passionate statistician. Political operator. Data visualization pioneer.

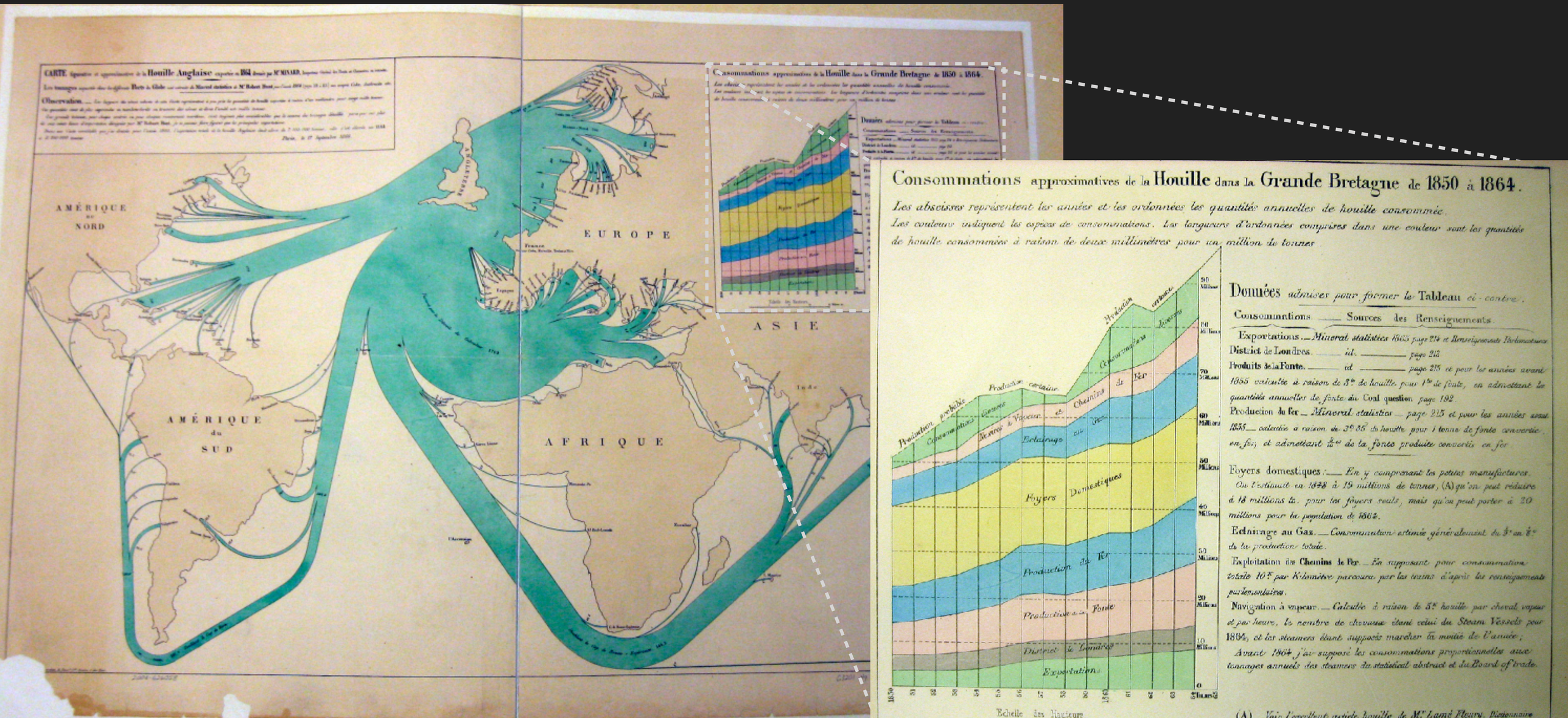
Napoleon's March to Moscow, Charles Joseph Minard - 1869

1500 2000



British Coal Exports, Charles Joseph Minard - 1864

1500 2000



About

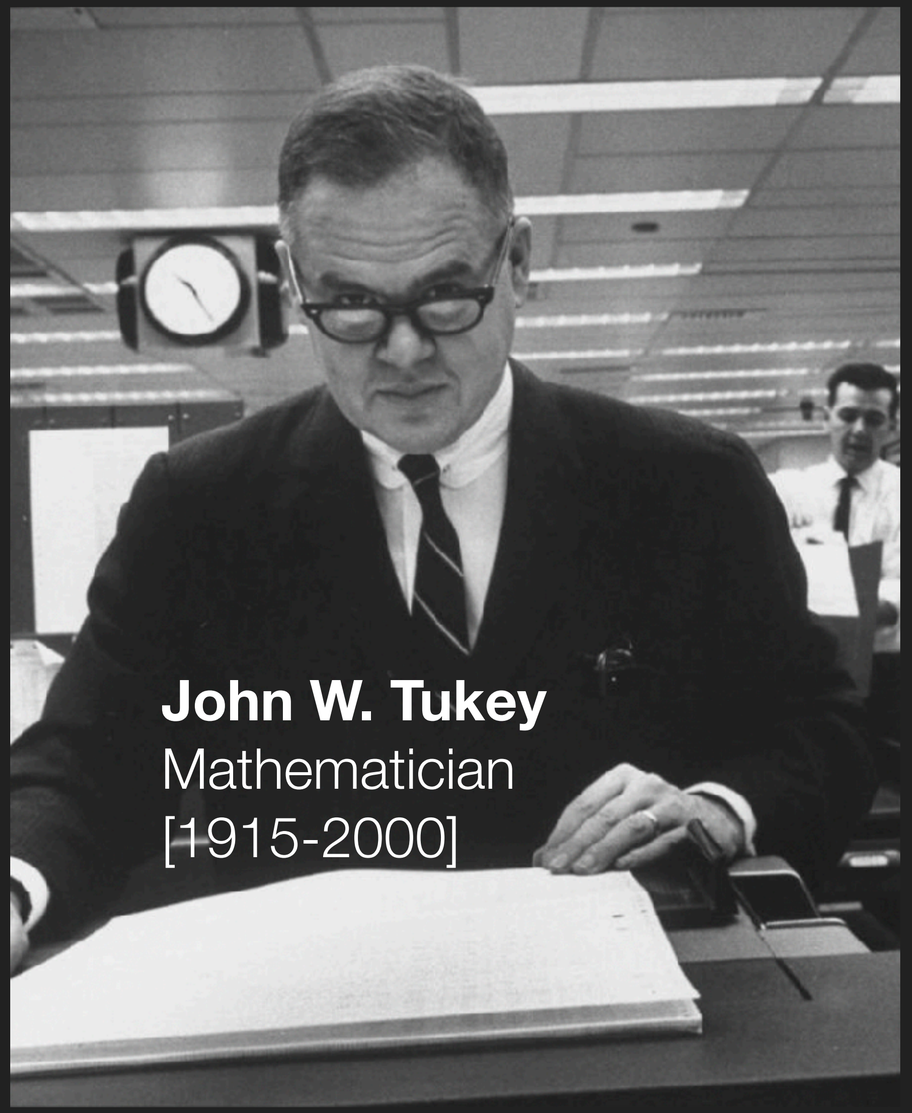
© 2017 RJ Andrews.
All rights reserved.

The Modern Dark Age of Visualization : 1900 ~ 1950

- Rise of quantification and formal statistical models
- Pictures considered not robust and reliable
- Few graphical innovations but widely adopted and used
- Dormancy awaiting new ideas and technology

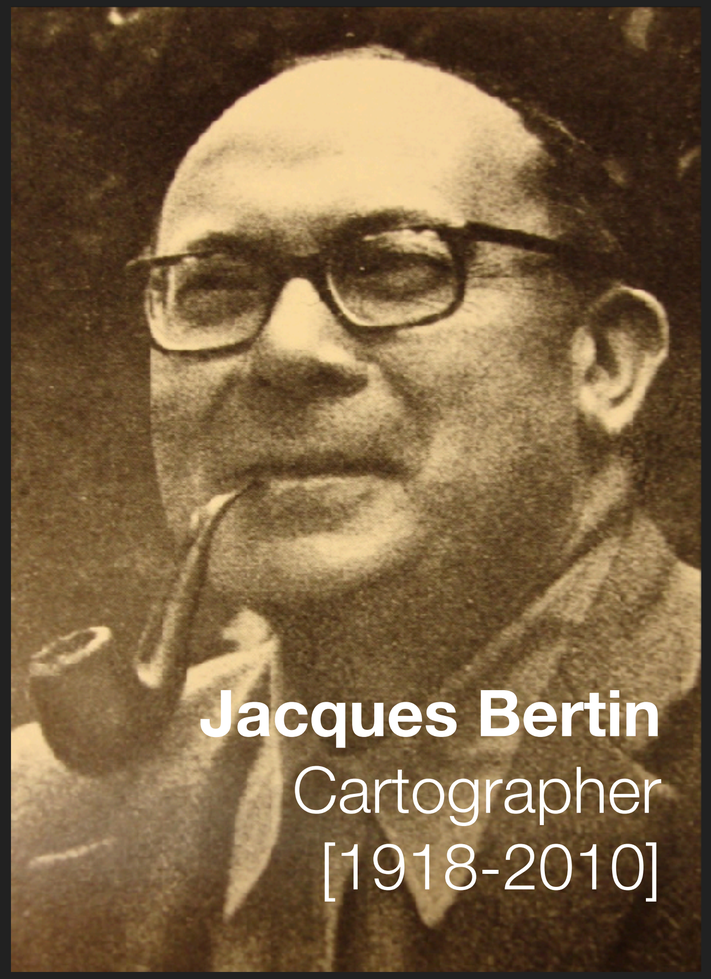
Exploratory Data Analysis (1977)

Use of graphical methods to rapidly explore data



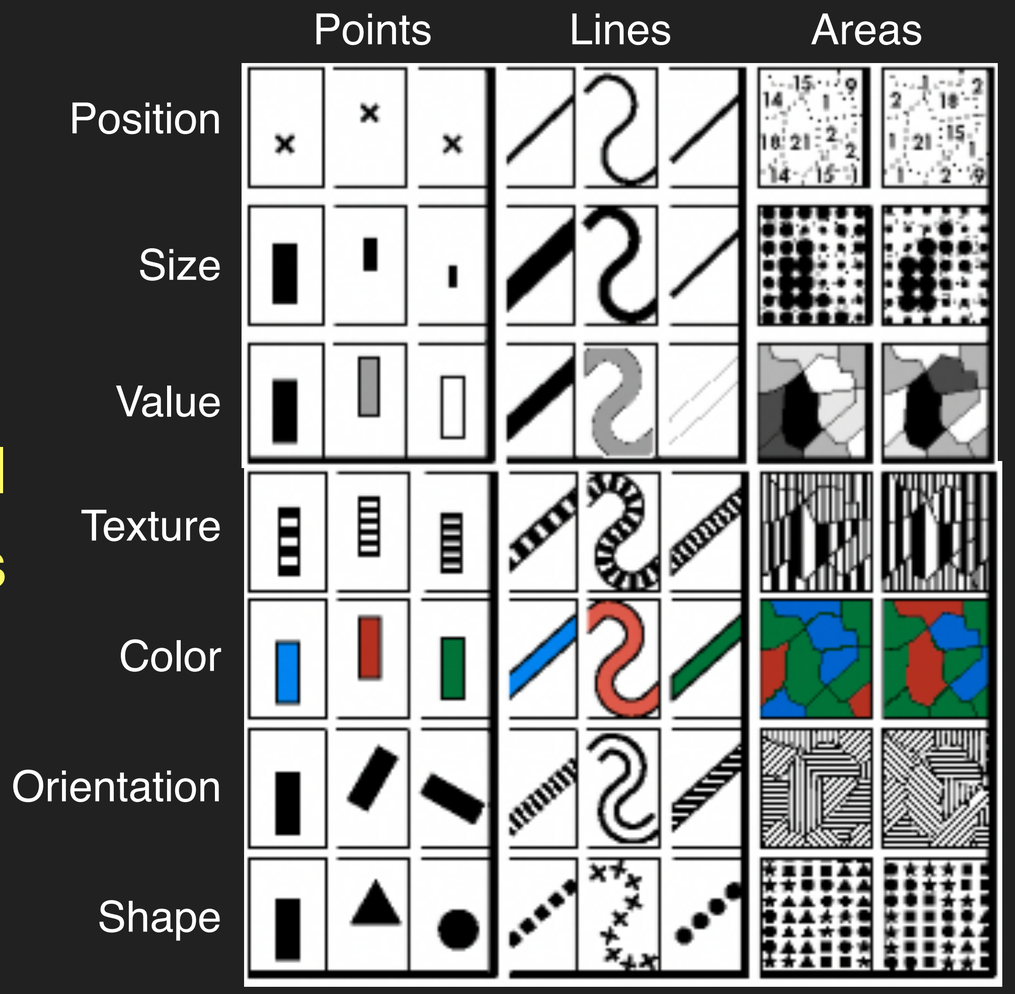
Semiology of Graphics (1967)

Use of graphical methods to rapidly explore data



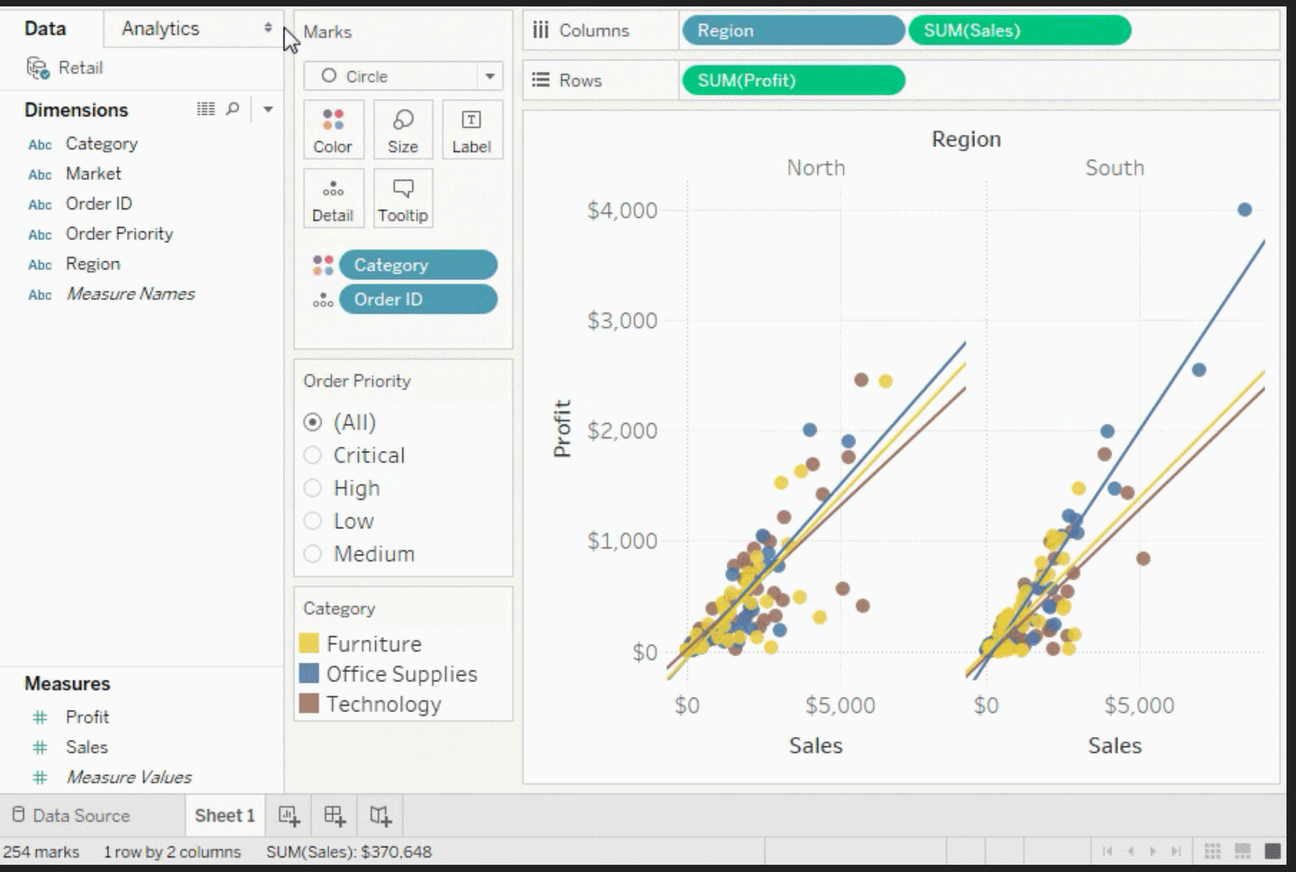
Visual Marks

Perceptual Channels

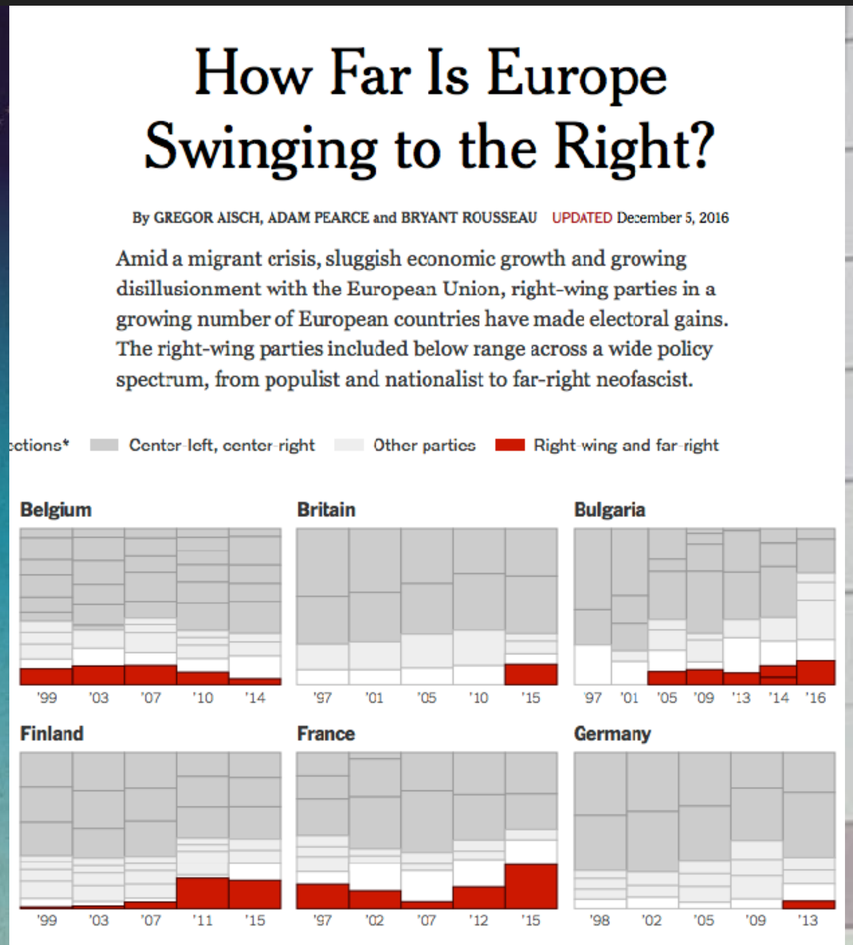


Going Mainstream : 1975 - today

Interactive Data Science - Support Analytic Reasoning



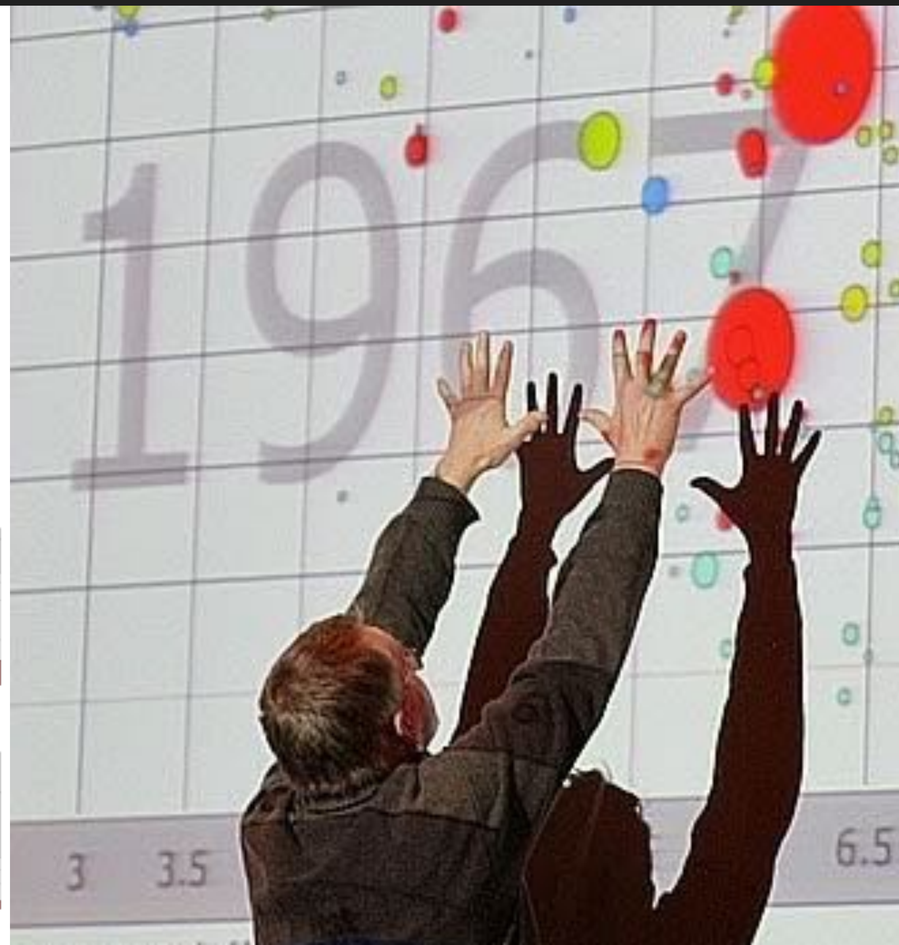
Data Journalism - Communicate Information to Others



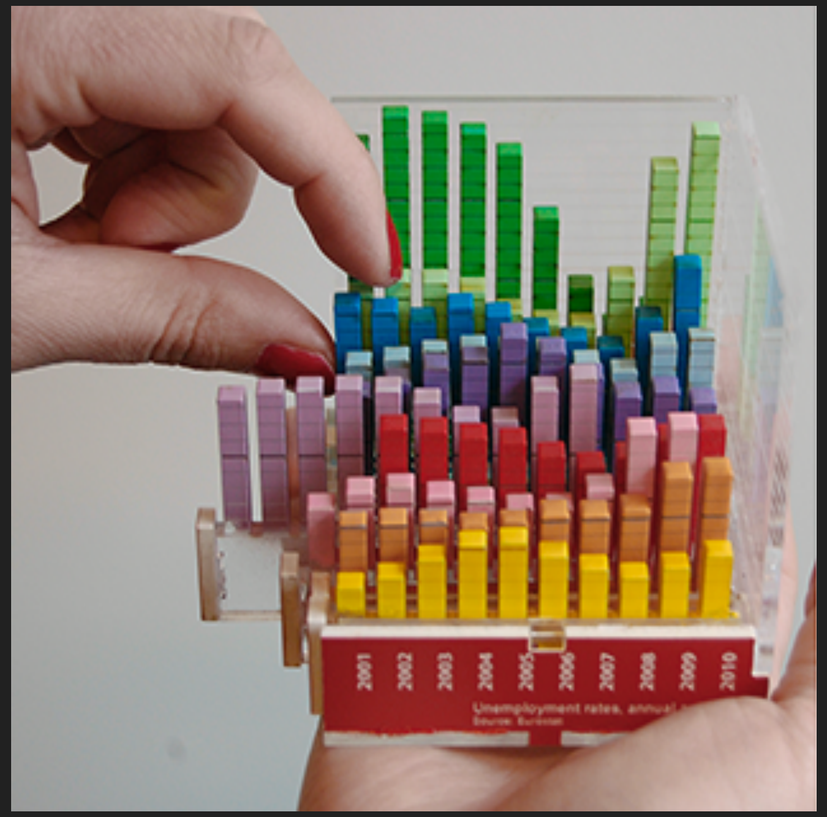
How Far Is Europe Swinging to the Right?

By GREGOR AISCH, ADAM PEARCE and BRYANT ROUSSEAU UPDATED December 5, 2016

Amid a migrant crisis, sluggish economic growth and growing disillusionment with the European Union, right-wing parties in a growing number of European countries have made electoral gains. The right-wing parties included below range across a wide policy spectrum, from populist and nationalist to far-right neofascist.



Personal Visualization - Record Information

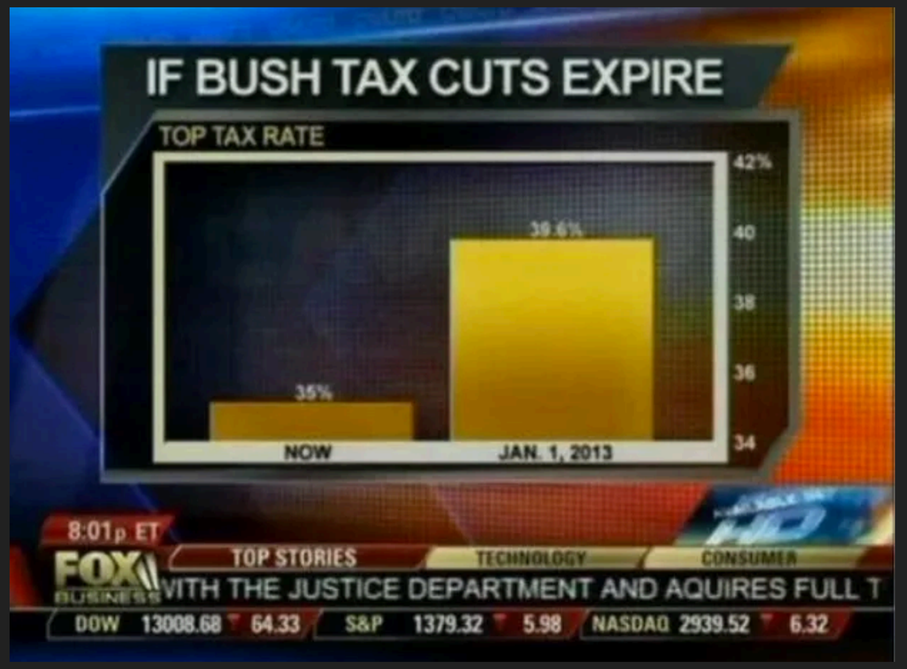


Summary

- What? Visual representations that convey information
- Why? Amplify our cognition to see the unseen
- How? By leveraging human visual perception system
- Three functions: Record, analyze, and communicate data

Next

Design Principles



Is this good, bad or weird?

Course Logistics

Learning Objectives

- Design effective visualizations
- Articulate rationales for the designs
- Build useful and usable interactive visualizations

Structure

- Lectures (Tue)
  - Fundamentals of visualization design
  - In-class activities
- Lab (Thu)
  - Hands-on training for building visualizations
  - HTML/CSS/Javascript/SVG and D3.js/Vega-lite
- Minute papers after every lecture and lab for feedback

Technology

- Canvas — Submissions and Grades
- Piazza — Q&A
- Github Classrooms — Code Hosting & Management
- Website — Course Schedule, Readings, & Slides

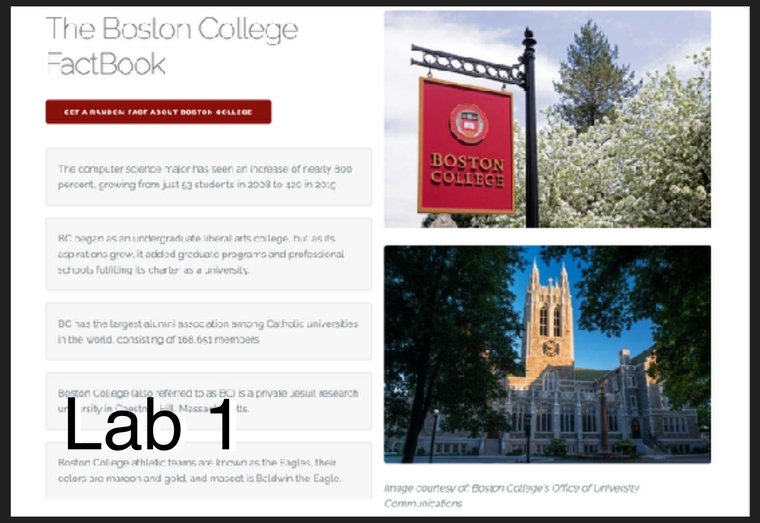
Lecture Topics

1. Value of Visualization
2. Design Principles
3. Visual Encoding Design
4. Exploratory Data Analysis
5. Storytelling with Data
6. Task Analysis and Evaluation
7. Interaction
8. Perception & Cognition
9. Networks and Maps
10. Multi-Dimensional Data
11. Uncertainty and Interpretability
12. Research in Visualization

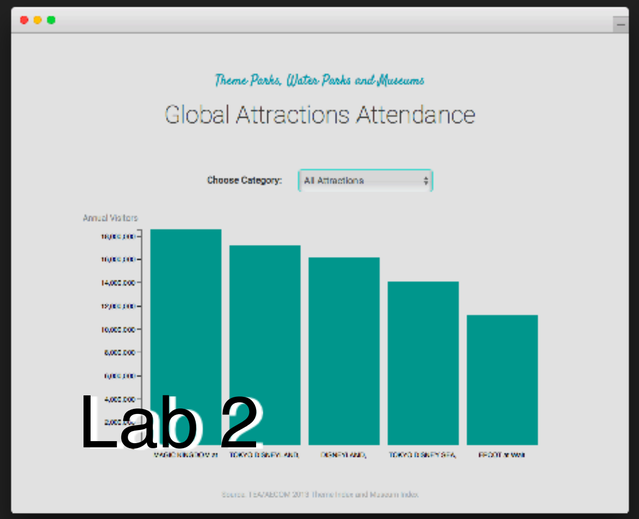
Lab Topics

1. Web Development Basics
2. Introduction to Javascript
3. Introduction to SVG & D3
4. D3 Scales and Axes
5. D3 Update Pattern
6. D3 Layouts
7. D3 Brushing and Linking
8. D3 Custom Visualizations
9. D3 Multiple View Compositions
10. Introduction to Vega-Lite
11. Advanced Vega-lite

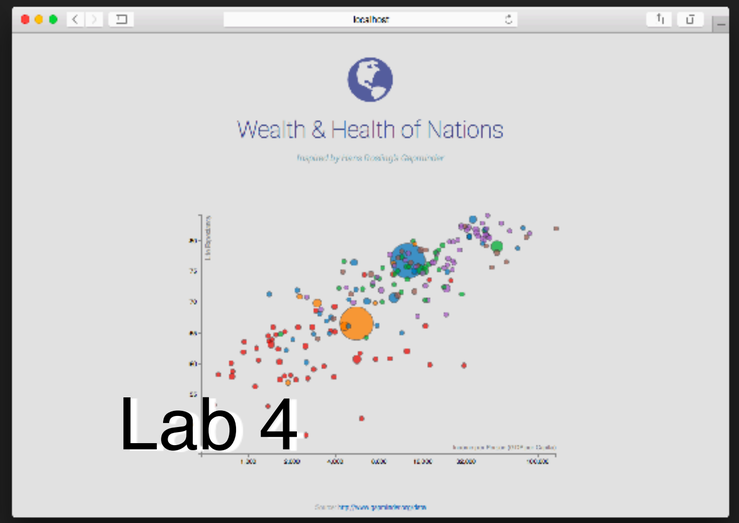
Lab Previews



Lab 1

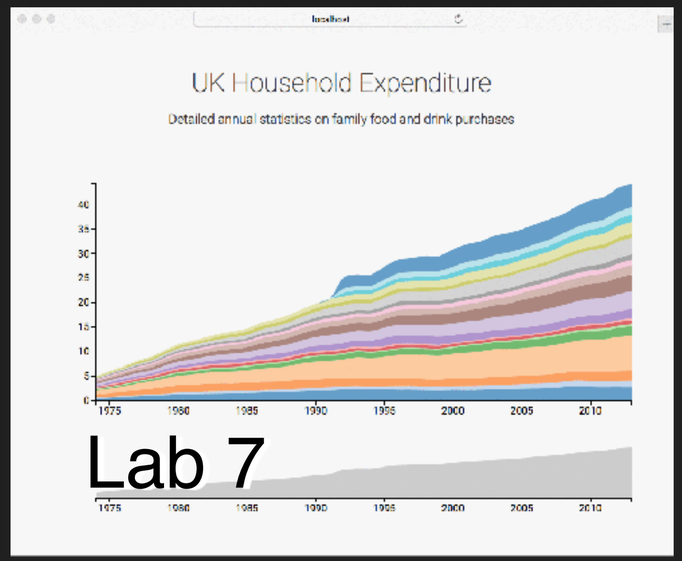


Lab 2



Lab 4

Lab 6



Lab 7

Lab 8

Prerequisites

- No hard prerequisites
- Any programming experience preferred
- Having experience in web development is a plus.

Office Hours

- Instructor - Nam Wook Kim
  - St. Mary's Hall S256, Th 1:30-3:30 pm
  - Reserve a 15 min slot on Calendly
- TA - Pedro De Almeida Rosa Guimaraes
  - Computer Science Lab - Fulton 160
  - Mo 7:00-9:00 pm

Readings

- Each lecture and lab require pre-class readings. Students are required make commentaries before each class on Canvas.

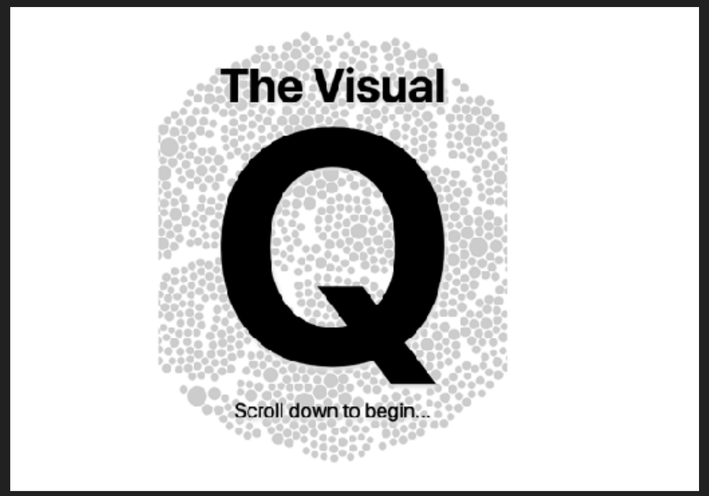
Homework

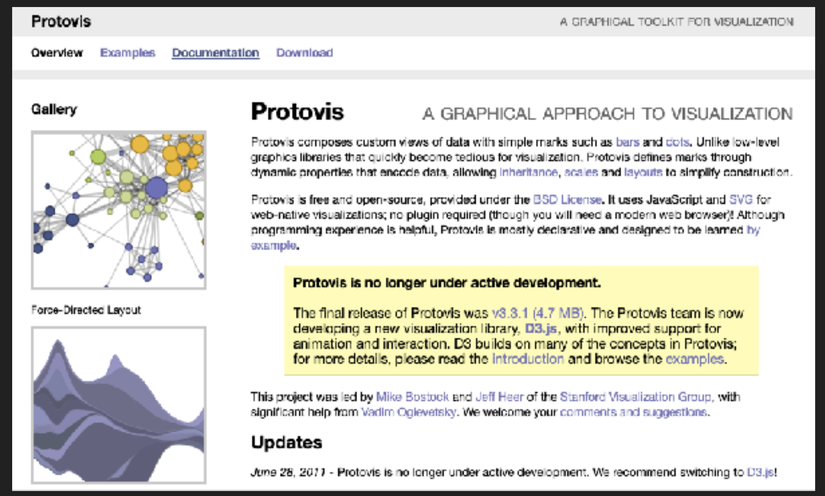
- HW1: Visualization critique & redesign (out next week)
- HW2: Exploratory and explanatory analysis with Tableau
- HW3: Interactive visualization using D3.js (Group project)

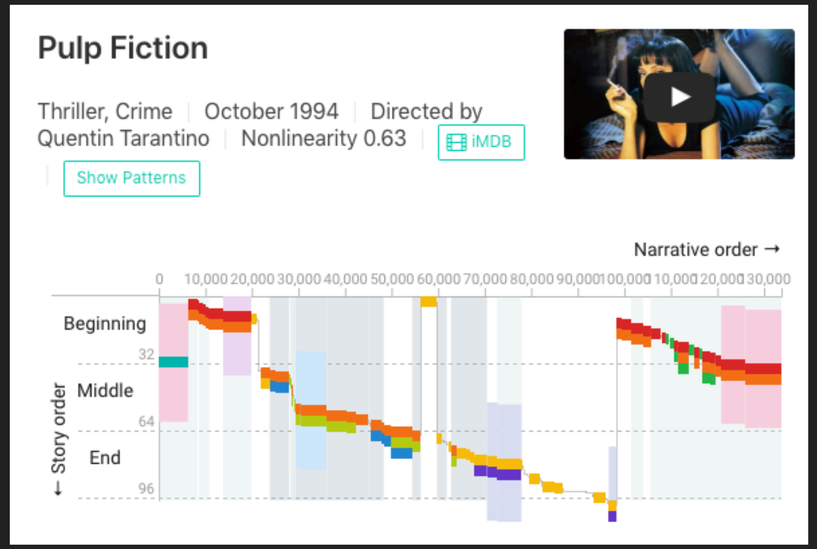
Final Project

- Build creative visualization products of your own interest.
- Design feedback sessions
- Final presentations
- Final report in the form of a conference abstract paper.

Inspirations







Your Project
in the Next Year

Gradings

- Readings — 5% : Before every lecture and lab
- Labs — 10% : Due by every next Tuesday
- Homeworks — 45% : 1: 8%, 2: 12%, 3: 25%
- Final Project — 40%

In-class activities and minute papers are not graded, but they will be referenced when your grades are on the boundaries.

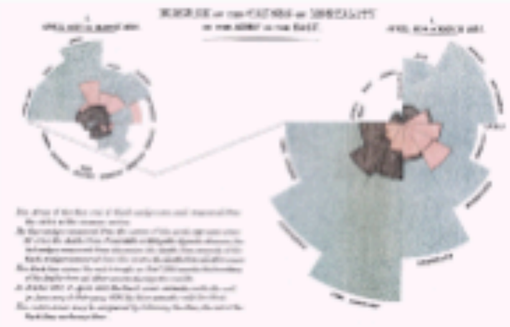
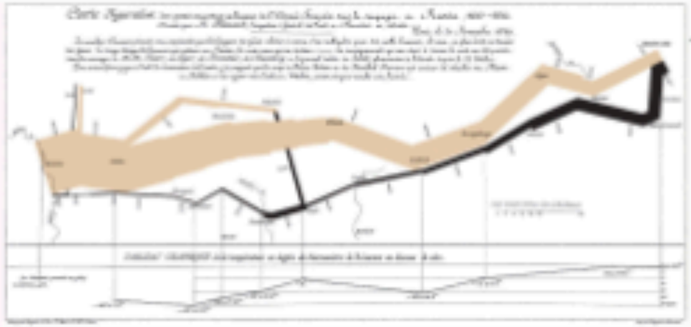
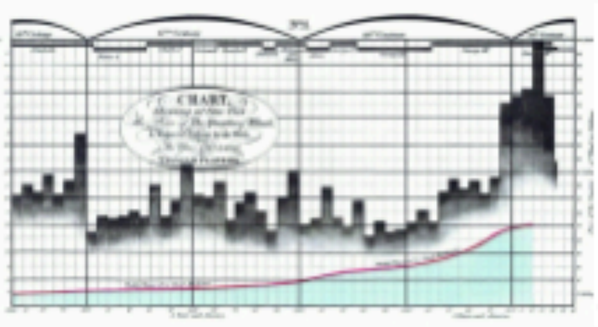
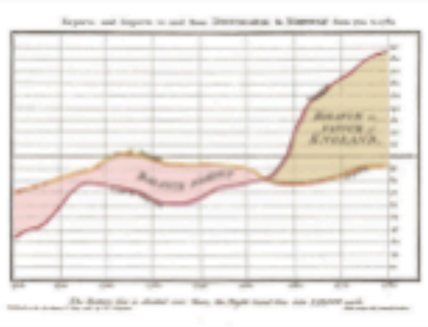
Course Policy

- Laptop policy
  - Bring your laptop for class activities and labs.
  - Only use it when asked.
- Mandatory attendance
- Late submission generally not accepted
- Regrade policy
  - Requests within 7 days after the initial grade

Code of Conduct

- Demonstrate professionalism (esp. for team projects)
- Promote inclusive and collaborative learning environment
- Reject behavior that strays into any form of harassments

CSCI339001 Visualization

Data can capture a snapshot of the world and allow us to understand ourselves and our communities better. With ever-increasing amounts of data, the ability to understand and communicate data is becoming essential for everyone. Visualization leverages our visual perception to provide a powerful yet accessible way to make sense of large and complex data. It has been widely adopted across disciplines, from science and engineering to business and journalism, to combat the overabundance of information in our society.

In this course, students will learn to acquire foundational knowledge about how to design effective visualizations for analysis and presentation based on theories and principles from graphic design, perceptual psychology, and cognitive science. Students will also learn practical skills about how to rapidly explore and communicate data using Tableau and build interactive visualization products (e.g., articles, tools, and systems) using web-based frameworks including D3.js and Vega-Lite.

https://bcviscourse.github.io

Time: TuTh 12:00-1:15 pm

Location: Merkert Center 130

Lastly…

- Please fill out the 1-minute paper.
  Available on Canvas, open until midnight tonight
- For waitlisted students, please let me know your name, major (CS?), year (senior?)