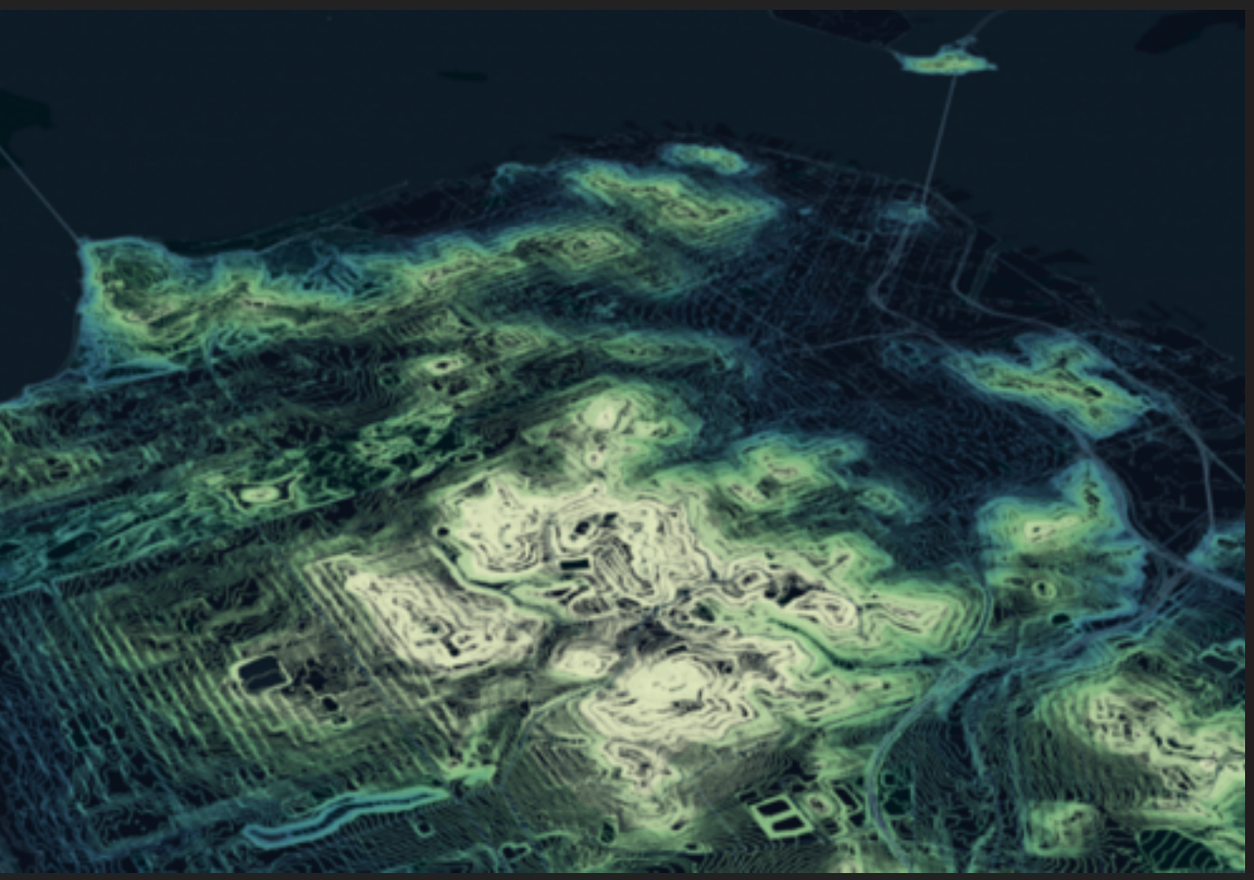
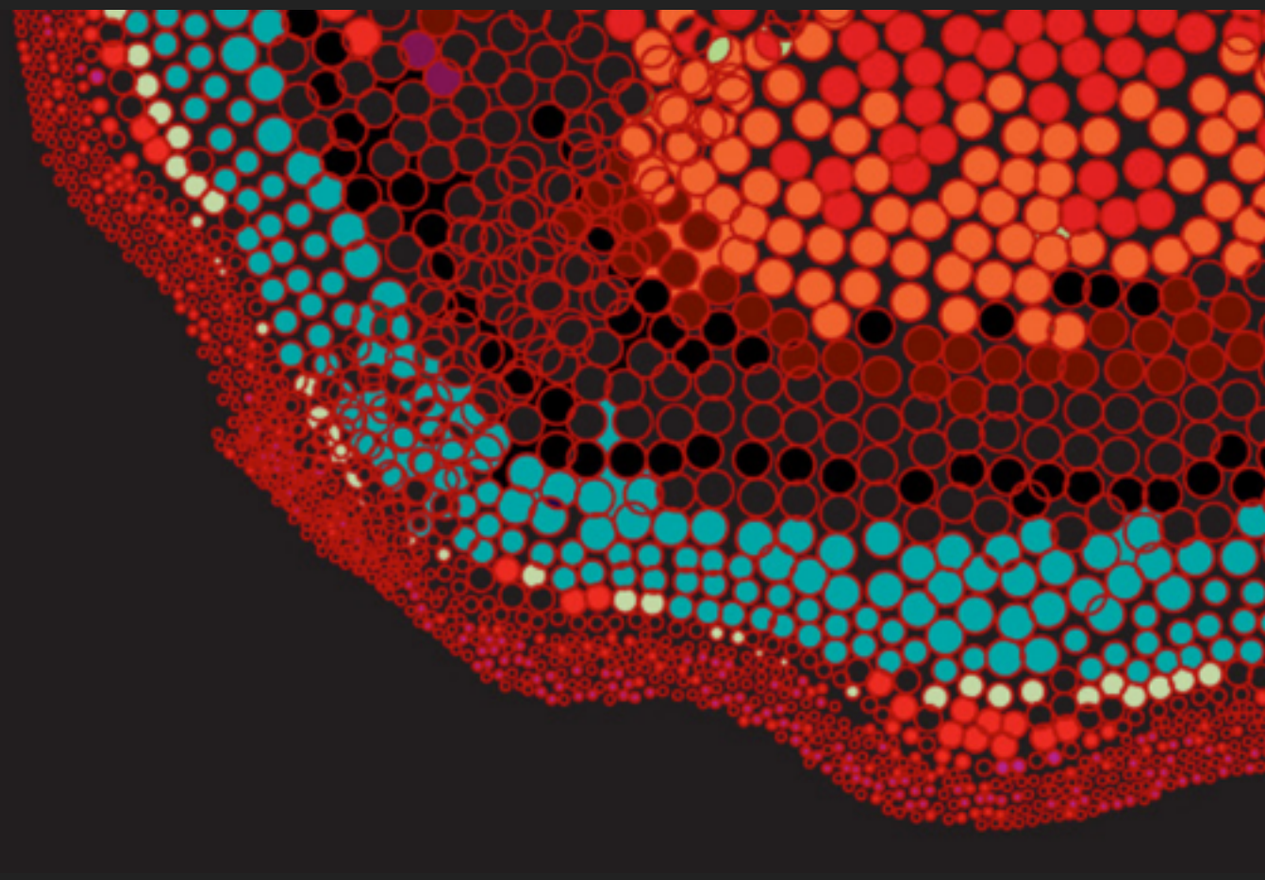




Lecturer: Min Lu, Qian Zheng

2020 Autumn

DATA VISUALIZATION

UISZU

Course Overview

Keywords

Fun!

Practice :)

Agenda

Final Project

Week 5 Week 8 Week 16, 17

- Week 1 - 7
- Week 9 - 15

Lectures + Tutorials

Assignments

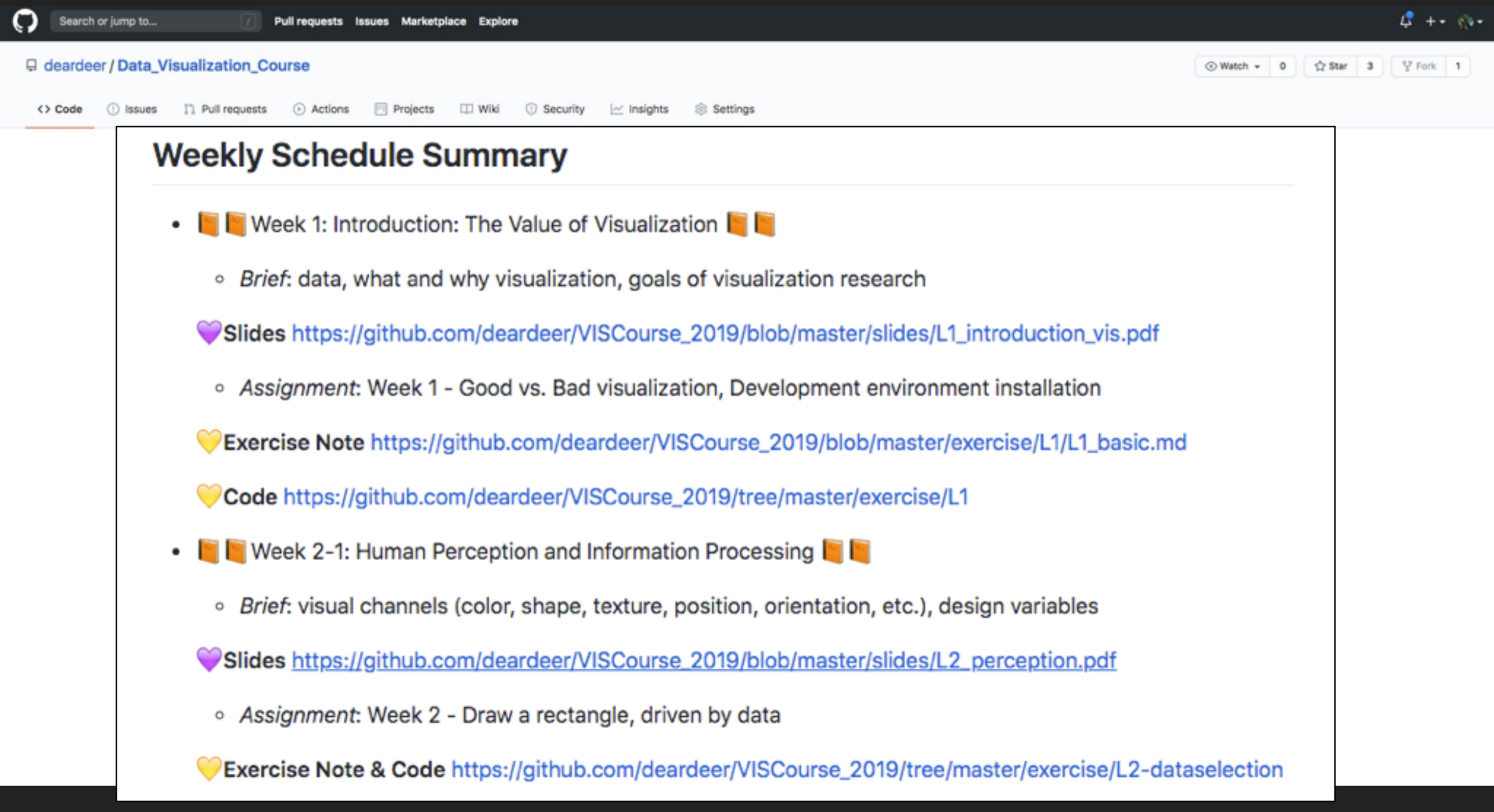
40% + 60%

Assignments (cont.)

- Class Participation (20%)
- Mini-Project (20%): Dear Data - Due National Holiday
- Final Project (60%): - Due Jan. 4
  - Design（20%）- Due Nov. 2
  - Implementation（40%）- Due Jan. 4

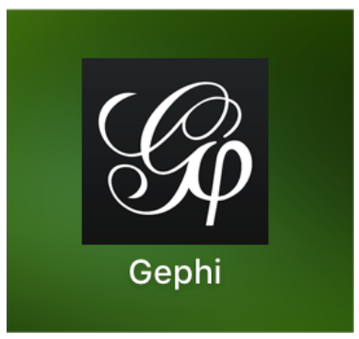
Course Homepage

https://github.com/deardeer/Data_Visualization_Course



Tools To Master

Off-line, Non-programing Tools



Programming Tools

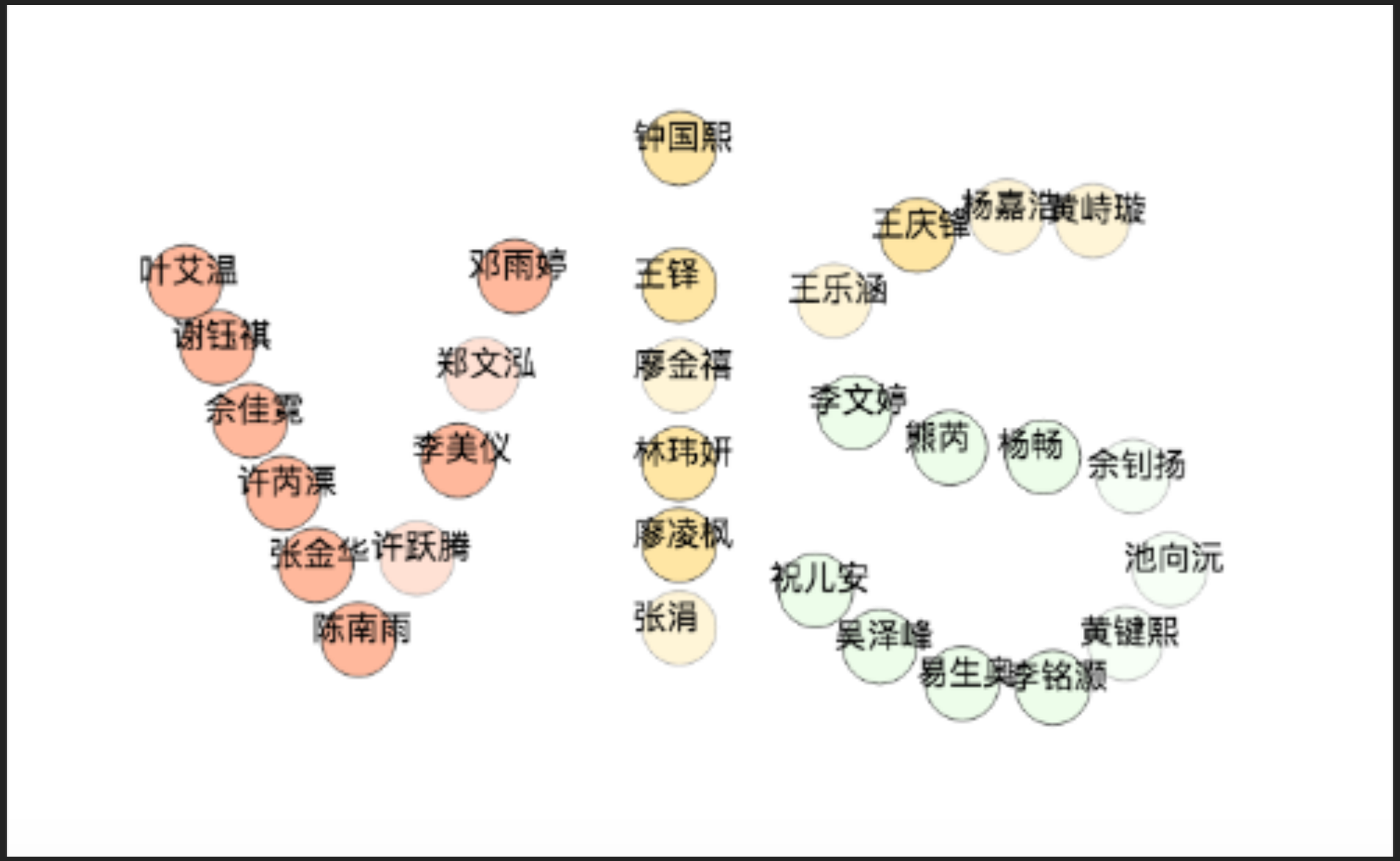
On-line, Non-program Tools

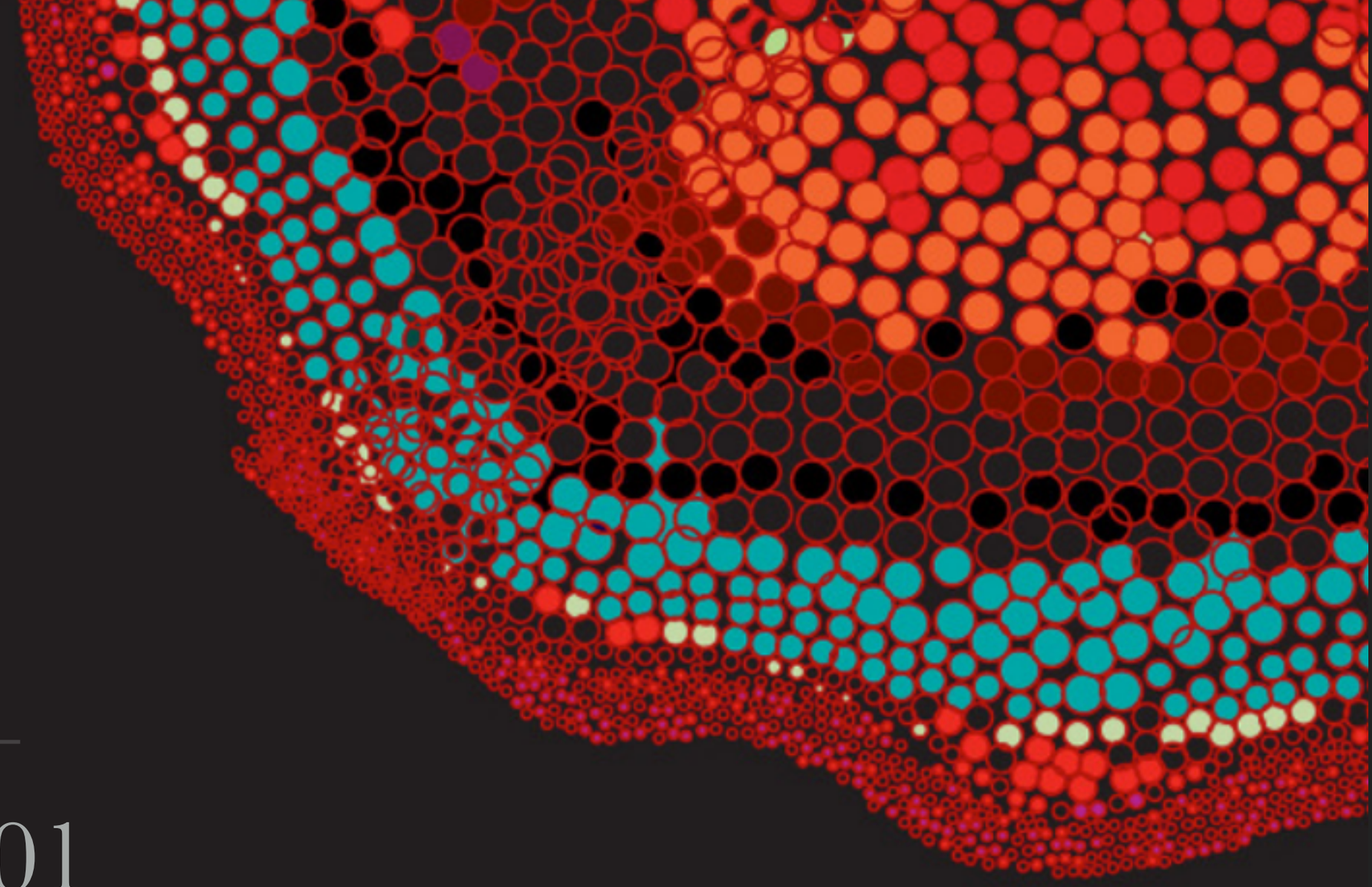
Who Am I

- 陆 旻 Min Lu, *Ph.D*, visualization, visual analytics
- Homepage: *https://deardeer.github.io/*
- Email: *lumin.vis@gmail.com*
- Office: *ScienceTec-building 1402*

Office Hour: Every Tuesday, 3:00 - 5:00 pm

Who Are You


钟国熙
王庆锷
杨嘉浩
黄峙璇
王乐涵
叶艾温
邓雨婷
王铎
谢钰祺
郑文泓
廖金禧
佘佳霓
李美仪
林玮妍
李文婷
熊芮
杨畅
佘钊扬
许芮溧
张金华
许跃腾
廖凌枫
池向沅
祝儿安
张涓
陈南雨
吴泽峰
黄键熙
易生奥
李铭灏



LECTURE 01

The Value Of Visualization

How Much Data (Bytes) Do We Create Everyday?

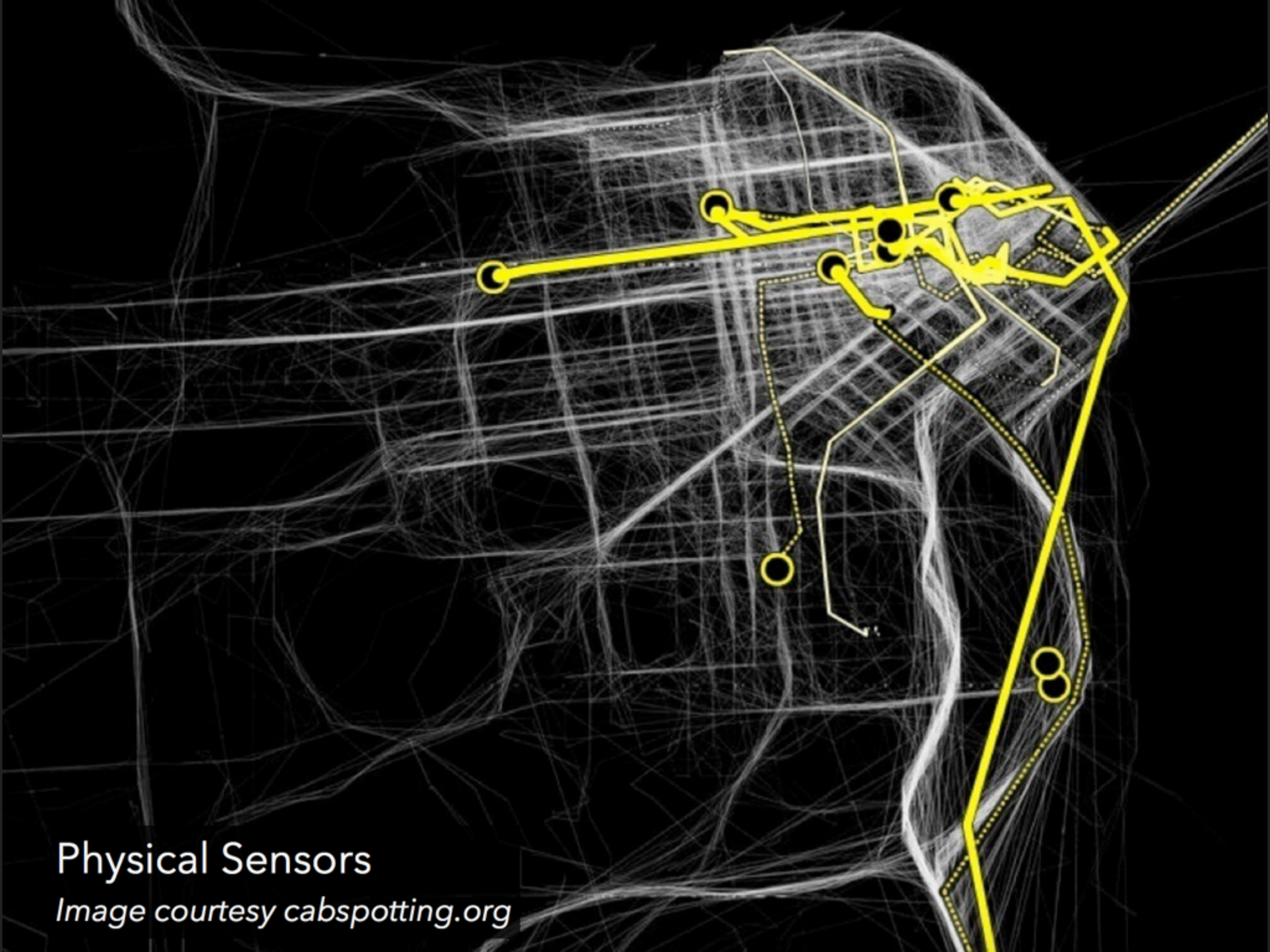
$10^{?}$

REFERENCE: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#1836349e60ba

$$1G = 10^9$$

$$10^{18} = 10^9 * 1G$$


INFORMATION TECHNOLOGY
ADVERTISING
WWW
SOCIAL MEDIA
SERVICES
PROJECTS
BIG DATA
PROMOTION
CONTENT
TARGET
CONSUMER
ORGANIZATION
PLANNING
EVENTS
PROGRAMMING
CODING
SOFTWARE
PLAN
E-MARKETING
COMMUNICATION
BRANDING
CONSUMER DEMAND
WEB MARKETING
SOCIAL
BIG DATA
BUSINESS
VISION
ENGINEERING
WEB DEV
BRAND
WEB SERVICES
BUZZ
BIG DATA
STRATEGY
WORLDWIDE
ORGANIZATION
BEHAVIOUR
PRO
PRICING
B2B TARGET
MEDIA
PLANNING
MOBILE
INFORMATION
DIGITAL
SOCIAL NETWORKS
SOLUTIONS
PROJECTS
SERVICE
CODING

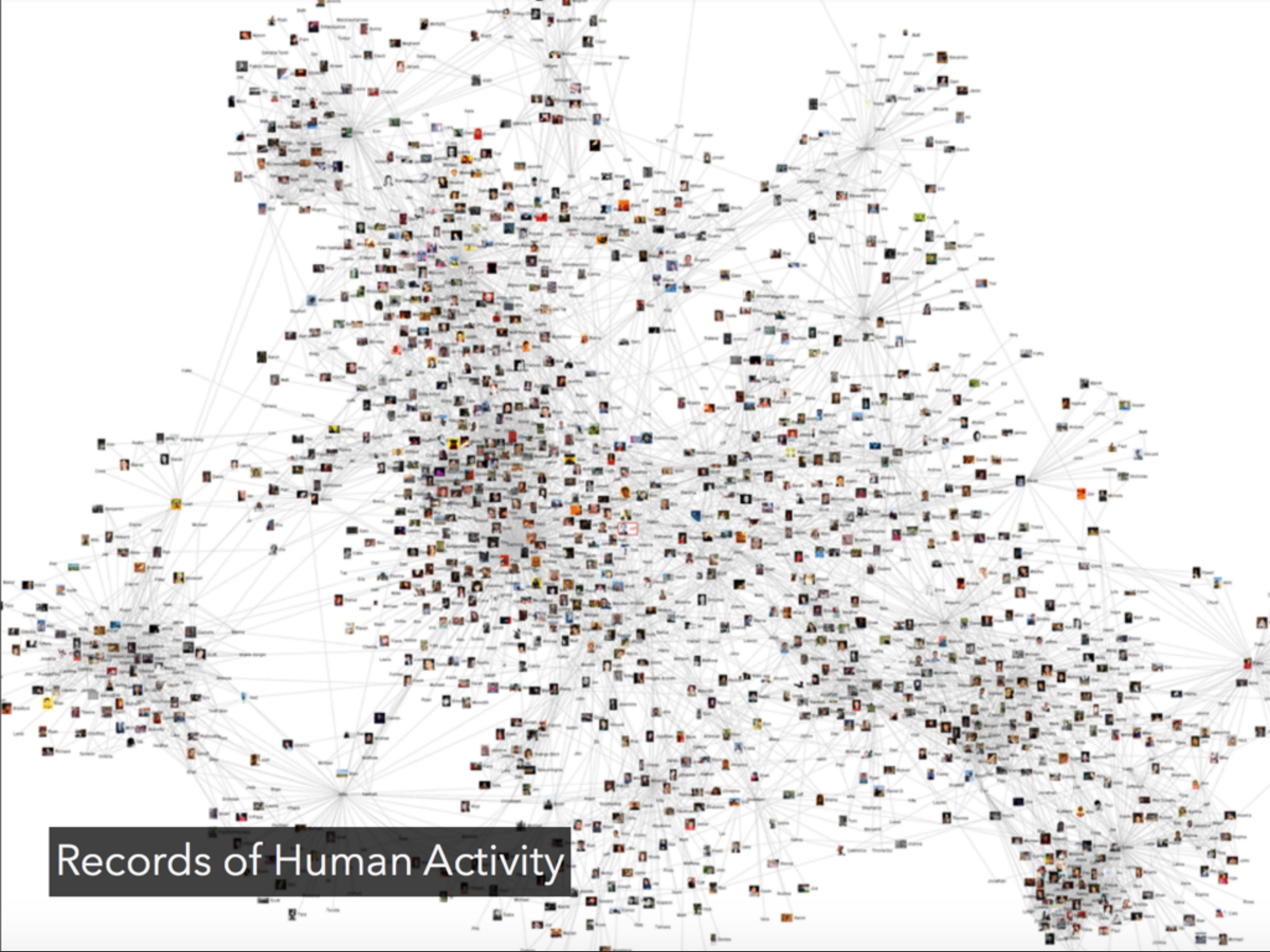


Physical Sensors

Image courtesy cabspotting.org

HR/ECG
1/min
137
Art
mmHg sys/dia
130/65
SpO2
%
RR/CO2
1/min
HR/ECG
1/min
97
Art
mmHg sys/dia
82/60
SpO2
%
99
RR/CO2
1/min
HR/ECG
1/min
79
Art
mmHg sys/dia
152/79
SpO2
%
95
RR/CO2
1/min
HR/ECG
1/min
64
Art
mmHg sys/dia
93/55
SpO2
%
99
RR/CO2
1/min
HR/ECG
1/min
99
mmHg sys/dia
SpO2
%
RR/Imp
1/min

Health & Medicine


Records of Human Activity

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, … because now we really do have **essentially free and ubiquitous data**. So the complimentary scarce factor is the ability to understand that data and extract value from it.

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

What is Visualization?

"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]

"The use of computer-generated, interactive, visual representations of data to amplify cognition."
[Card, Mackinlay, & Shneiderman 1999]

Set A		Set B		Set C		Set D	
X	Y	X	Y	X	Y	X	Y
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

$u_X = 9.0$ $\sigma_X = 3.317$

$u_Y = 7.5$ $\sigma_Y = 2.03$

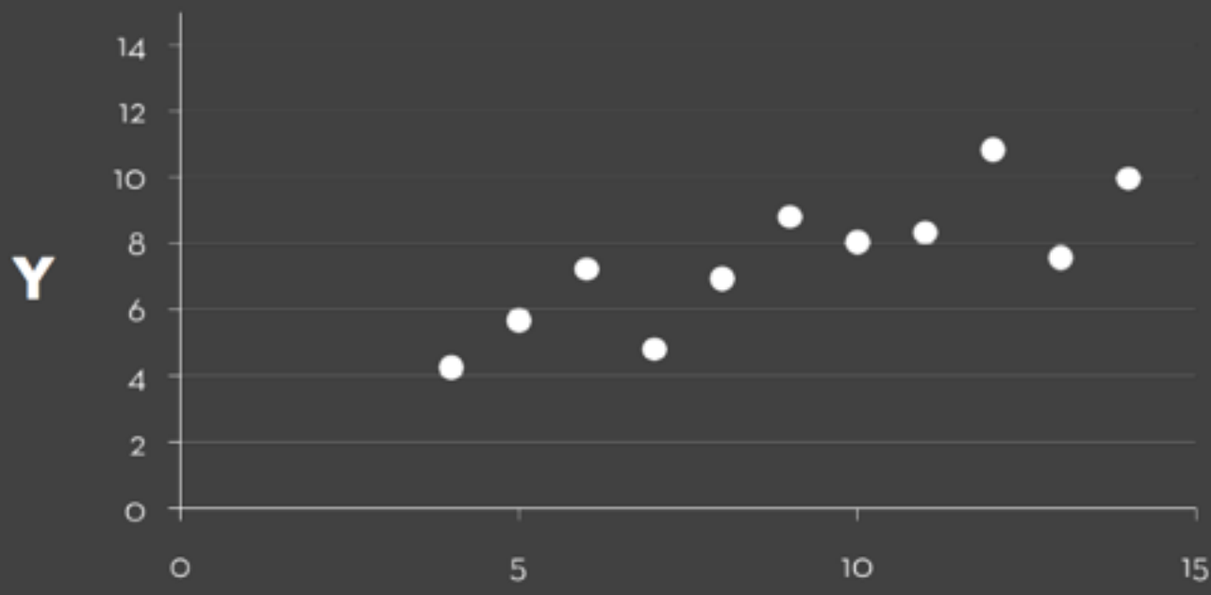
Linear Regression

$Y' = 3 + 0.5\,X$

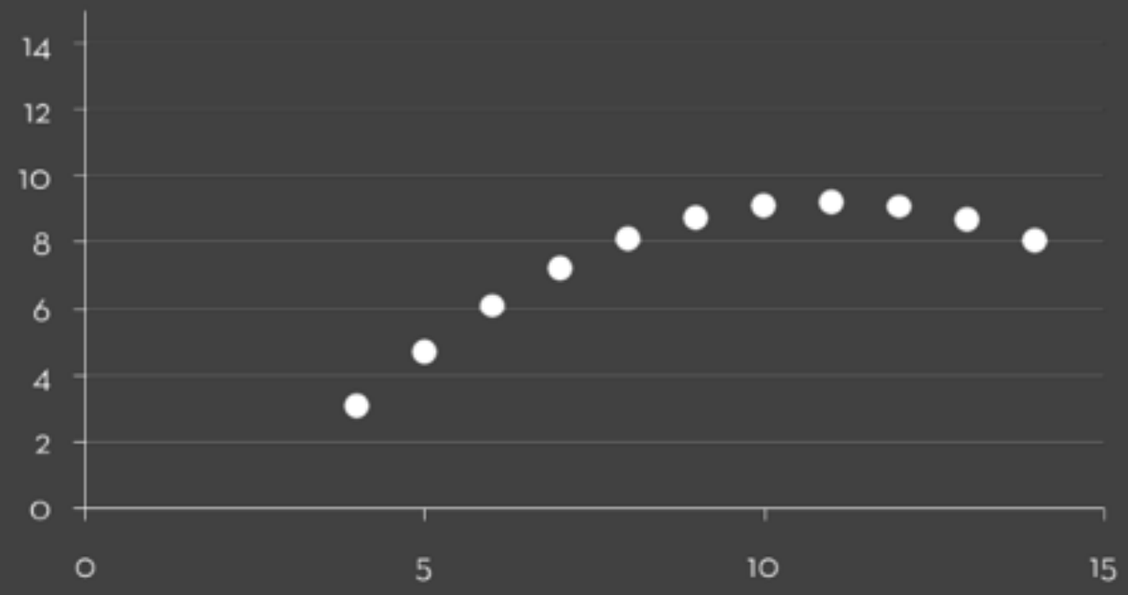
$R^2 = 0.67$

[Anscombe 1973]

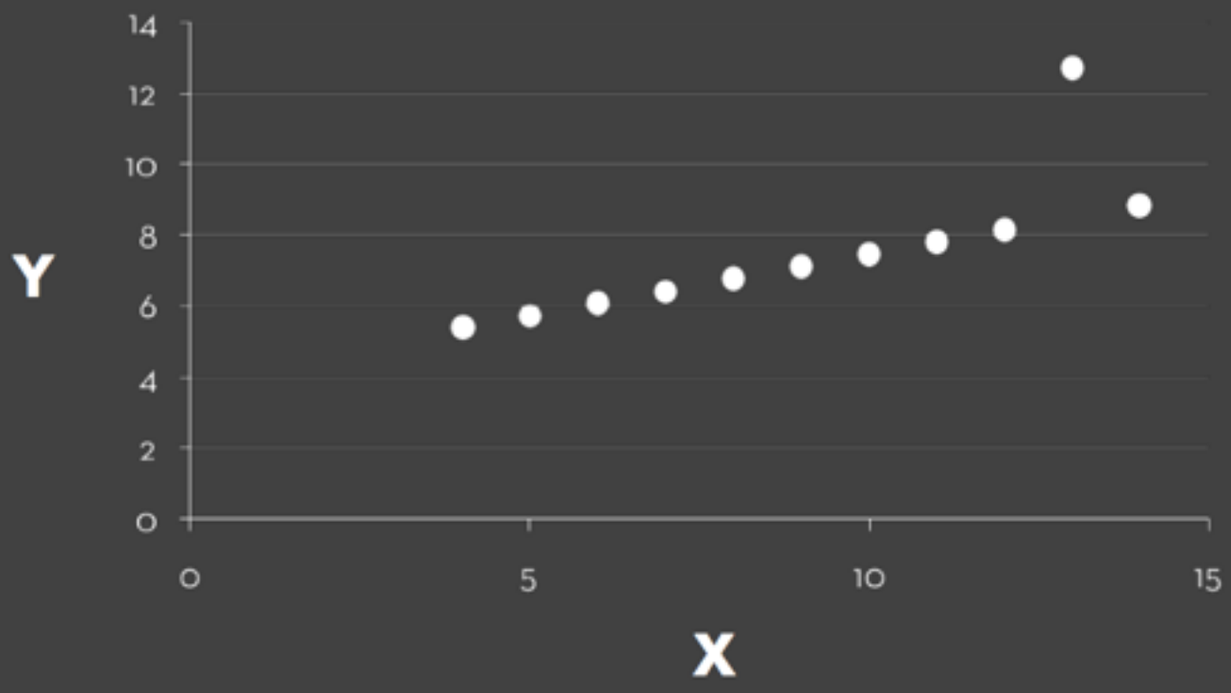
Set A



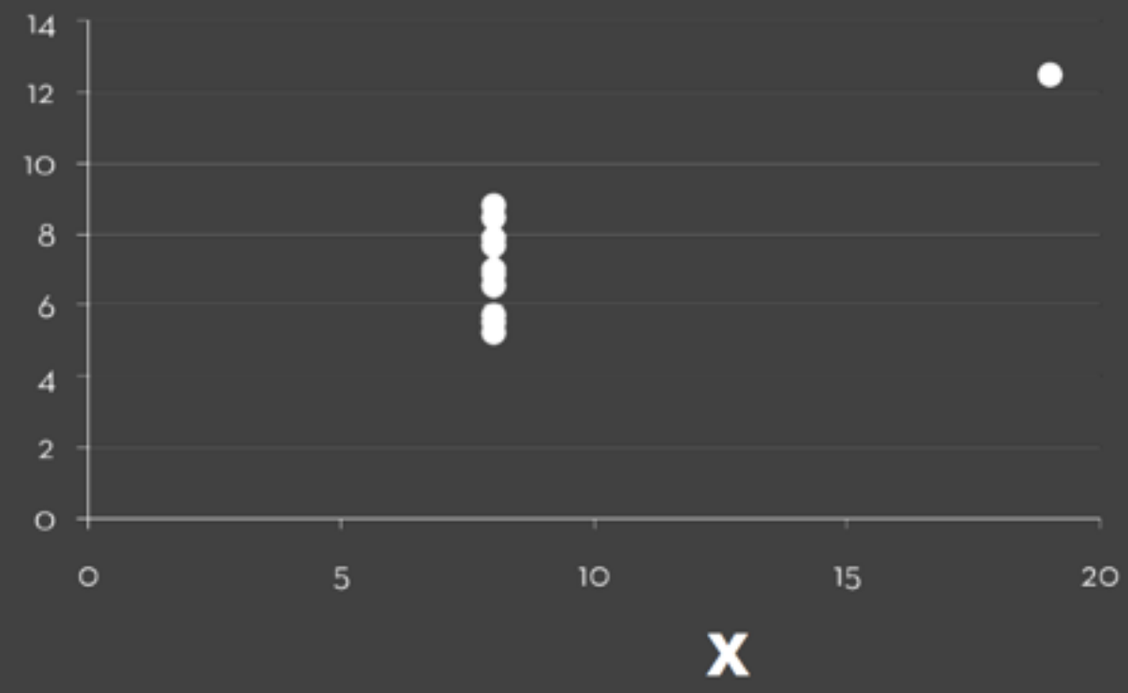
Set B



Set C



Set D

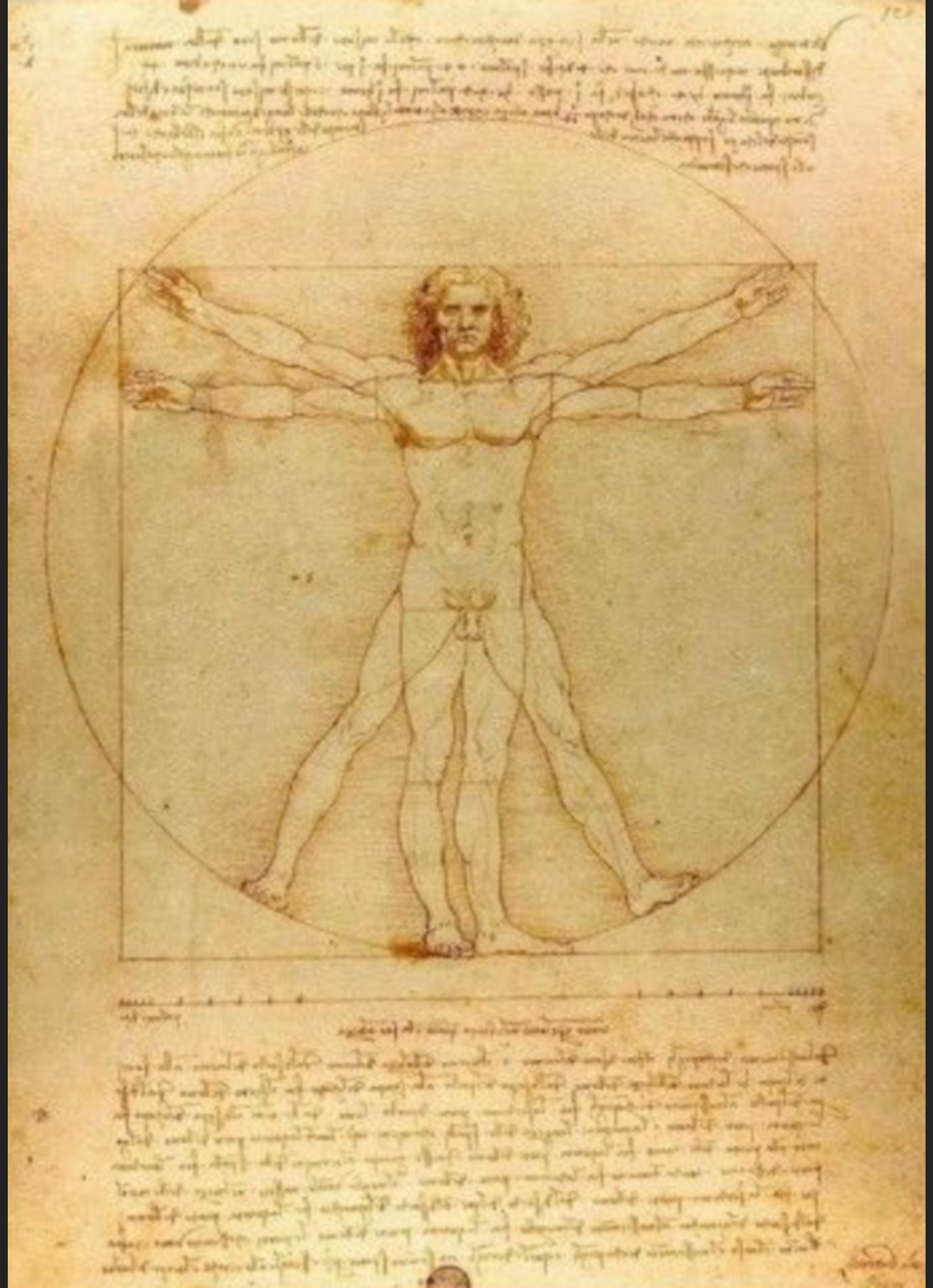


Why Create Visualizations?

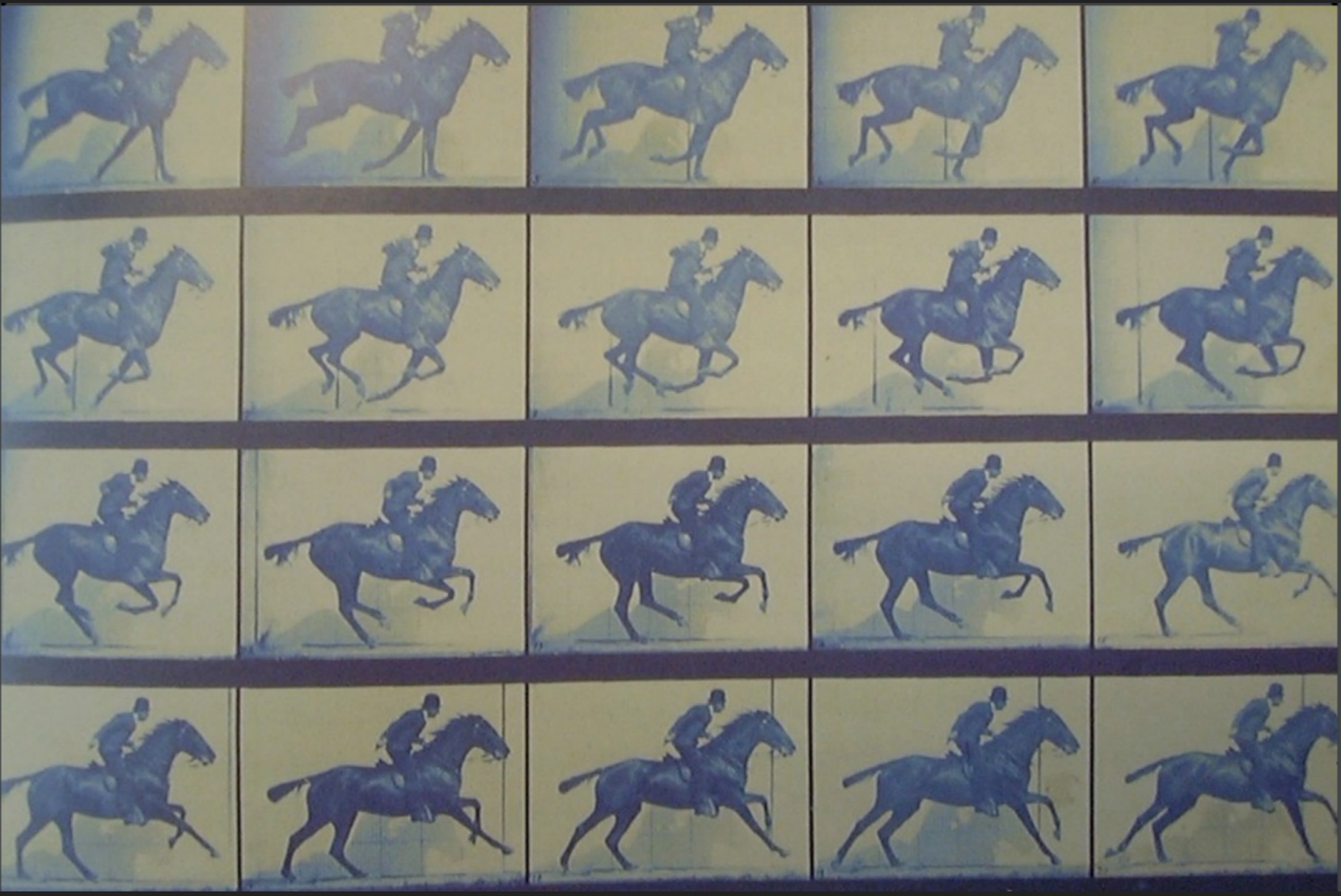
Answer questions (or discover them)
Make decisions
See data in context
Expand memory
Support graphical calculation
Find patterns
Present argument or tell a story
Inspire

Record Information

Spanish Los Caballos



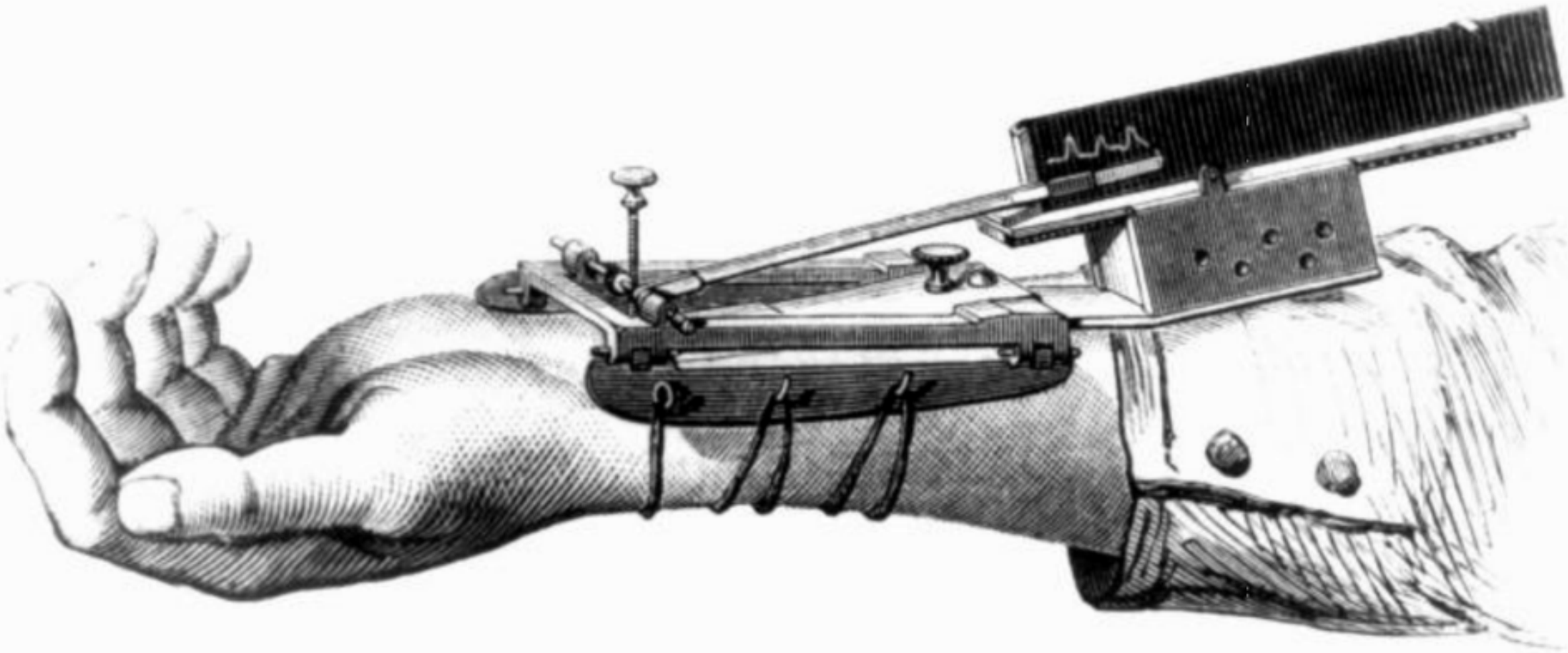
The Vitruvian Man, by
Leonardo da Vinci, 1487



Gallop, Bay Horse “Daisy” [Muybridge 1884-86]

May 2011

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	2	**3** NM	4	5	6	7
8	9	**10** FQ	11	12	13	14
15	16	**17** FM	18	19	20	21
22	23	**24** LQ	25	26	27	28
29	30	31				

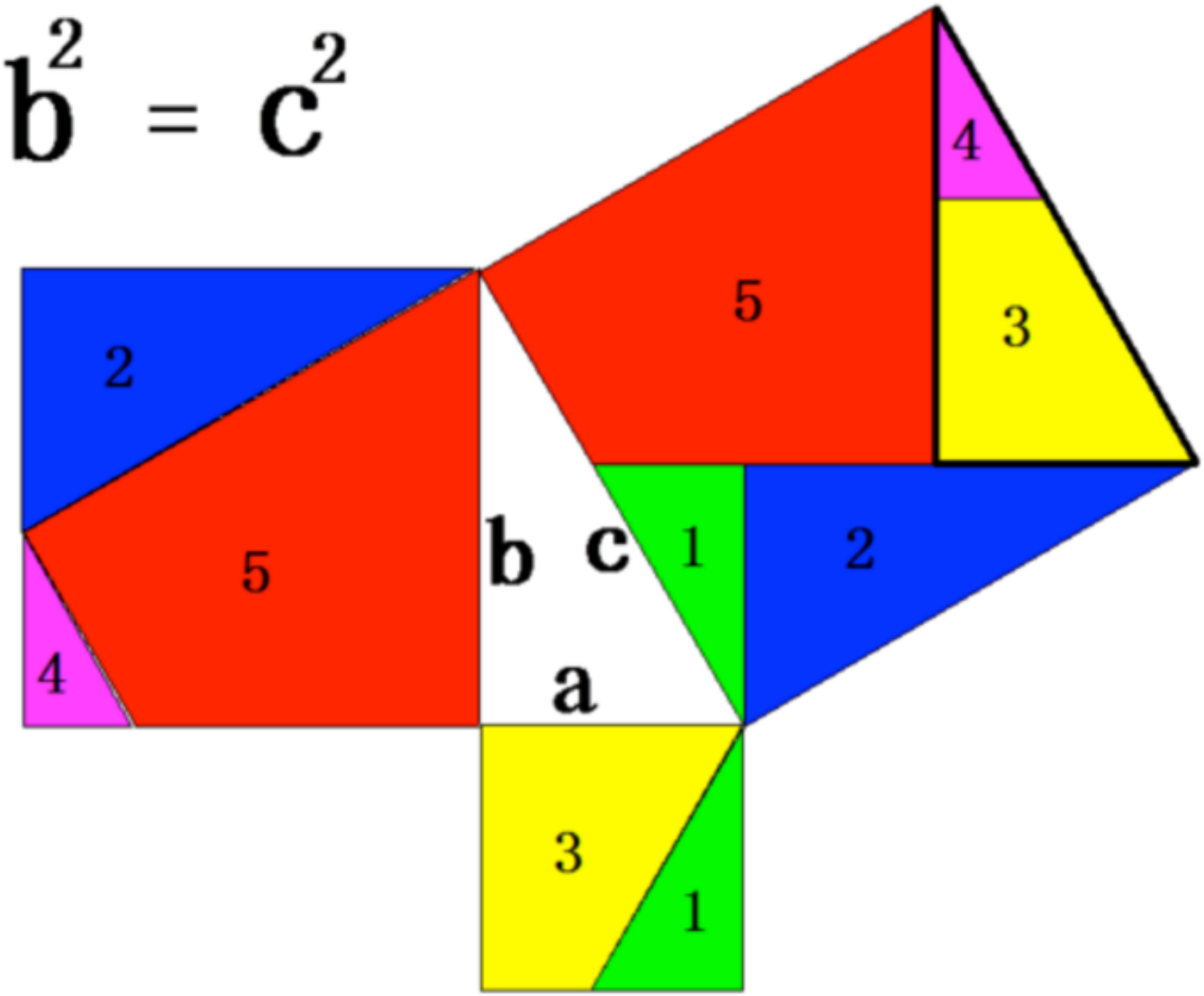


1.
Marey's sphygmograph in use, 1860. *La méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine.*

E.J. Marey's sphygmograph [from Braun 83]

Support Reasoning

$$a^2 + b^2 = c^2$$

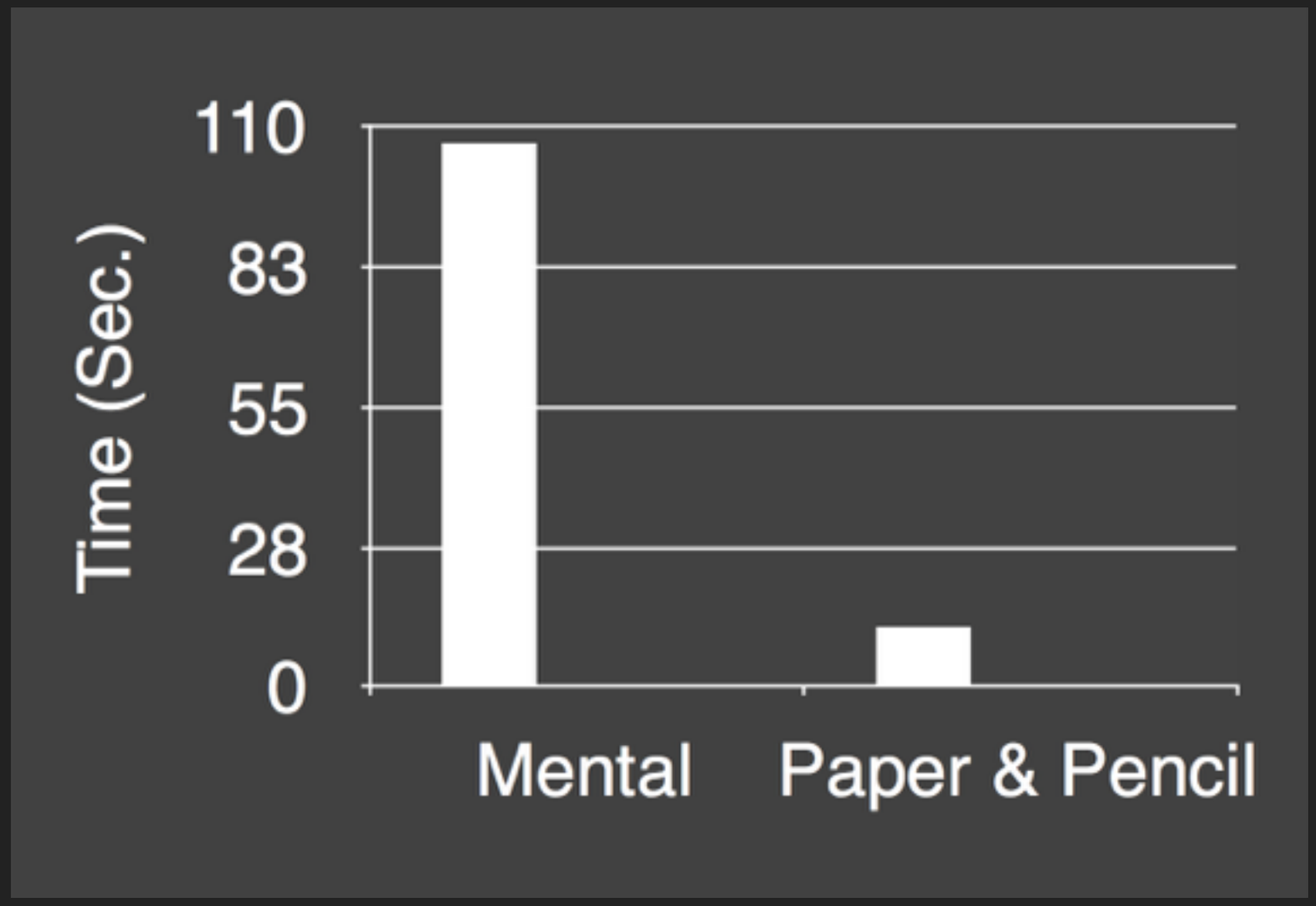


In 1854 John Snow plotted the position of each cholera case on a map. [from Tufte 83]

Expand Memory

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

$$
\begin{array}{r}
34 \\
\times\ 72 \\
\hline
68 \\
2380 \\
\hline
2448
\end{array}
$$



Find Patterns

New York City's Weather in 2003

JANUARY FEBRUARY MARCH APRIL MAY JUNE JULY AUGUST SEPTEMBER OCTOBER NOVEMBER DECEMBER

Temperature

Bars represent range between the daily high and low. Average temperature for the year was 53.4°, making 2003 the coldest year since 1978.

110° 100° 90° 80° 70° 60° 50° 40° 30° 20° 10° 0° -10° -20°

RECORD HIGH: 79°

TIED RECORD LOW: 11°

RECORD HIGH
ACTUAL HIGH
NORMAL RANGE
ACTUAL LOW
RECORD LOW

Precipitation Cumulative monthly precipitation in inches compared with normal monthly precipitation. Total precipitation in 2003 was 58.51 inches, 8.82 inches more than normal, which makes the year the sixth wettest on record.

10 8 6 4 2 0

JANUARY: NORMAL 4.13; ACTUAL 2.30

FEBRUARY: 4th snowiest Feb.; RECORD SNOWFALL: 16.3 in.; 3.15; 4.55

MARCH: 4.37; 4.57

APRIL: 11th snowiest April; 4.28; 3.20

MAY: 4.69; RECORD RAINFALL: 1.26 in.; 3.40

JUNE: RECORD RAINFALL: 2.75 in.; 10.26; 3.84; Wettest June on record

JULY: 4.62; 3.76

AUGUST: 5.85; 4.22

SEPTEMBER: 4.23; 6.03

OCTOBER: 3.85; RECORD RAINFALL: 1.88 in.; 4.90

NOVEMBER: 4.36; 4.27

DECEMBER: 3.95; 6th snowiest Dec.; RECORD SNOWFALLS; 5.42; RECORD RAINFALL: 1.42 in.

Microsoft Excel - animal.xls

File Edit View Insert Format Tools Data Window Help

A1 fx ID

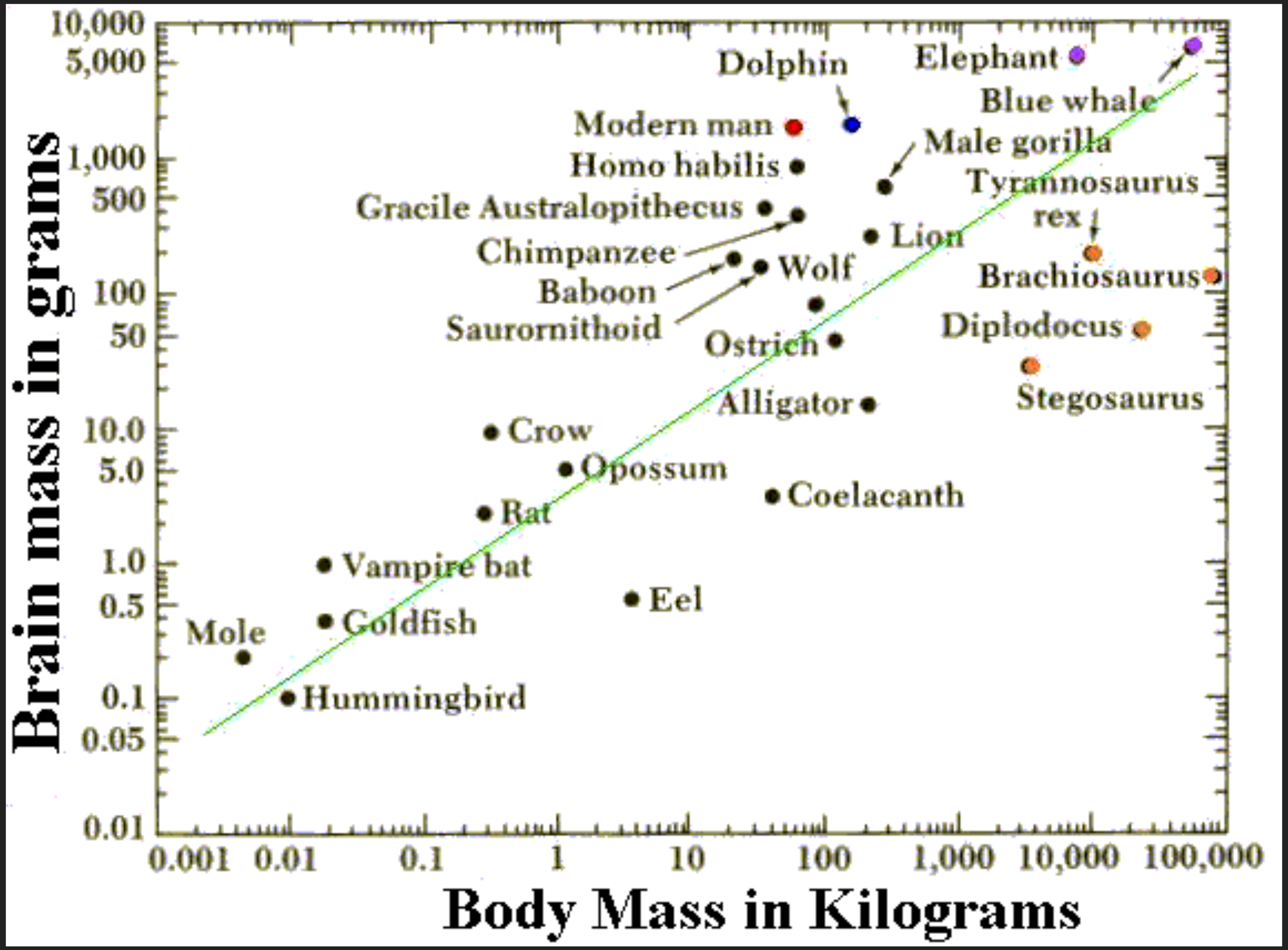
	A	B	C	D	E
1	**ID**	**Name**	**Body Weight**	**Brain Weight**	
2	1	Lesser Short-tailed Shrew	5	0.14	
3	2	Little Brown Bat	10	0.25	
4	3	Mouse	23	0.3	
5	4	Big Brown Bat	23	0.4	
6	5	Musk Shrew	48	0.33	
7	6	Star Nosed Mole	60	1	
8	7	Eastern American Mole	75	1.2	
9	8	Ground Squirrel	101	4	
10	9	Tree Shrew	104	2.5	
11	10	Golden Hamster	120	1	
12	11	Mole Rate	122	3	
13	12	Galago	200	5	
14	13	Rat	280	1.9	
15	14	Chinchilla	425	6.4	
16	15	Desert Hedgehog	550	2.4	
17	16	Rock Hyrax (a)	750	12.3	
18	17	European Hedgehog	785	3.5	
19	18	Tenrec	900	2.6	
20	19	Arctic Ground Squirrel	920	5.7	
21	20	African Giant Pouched Rat	1000	6.6	
22	21	Guinea Pig	1040	5.5	
23	22	Mountain Beaver	1350	8.1	
24	23	Slow Loris	1400	12.5	
25	24	Genet	1410	17.5	
26	25	Phalanger	1620	11.4	

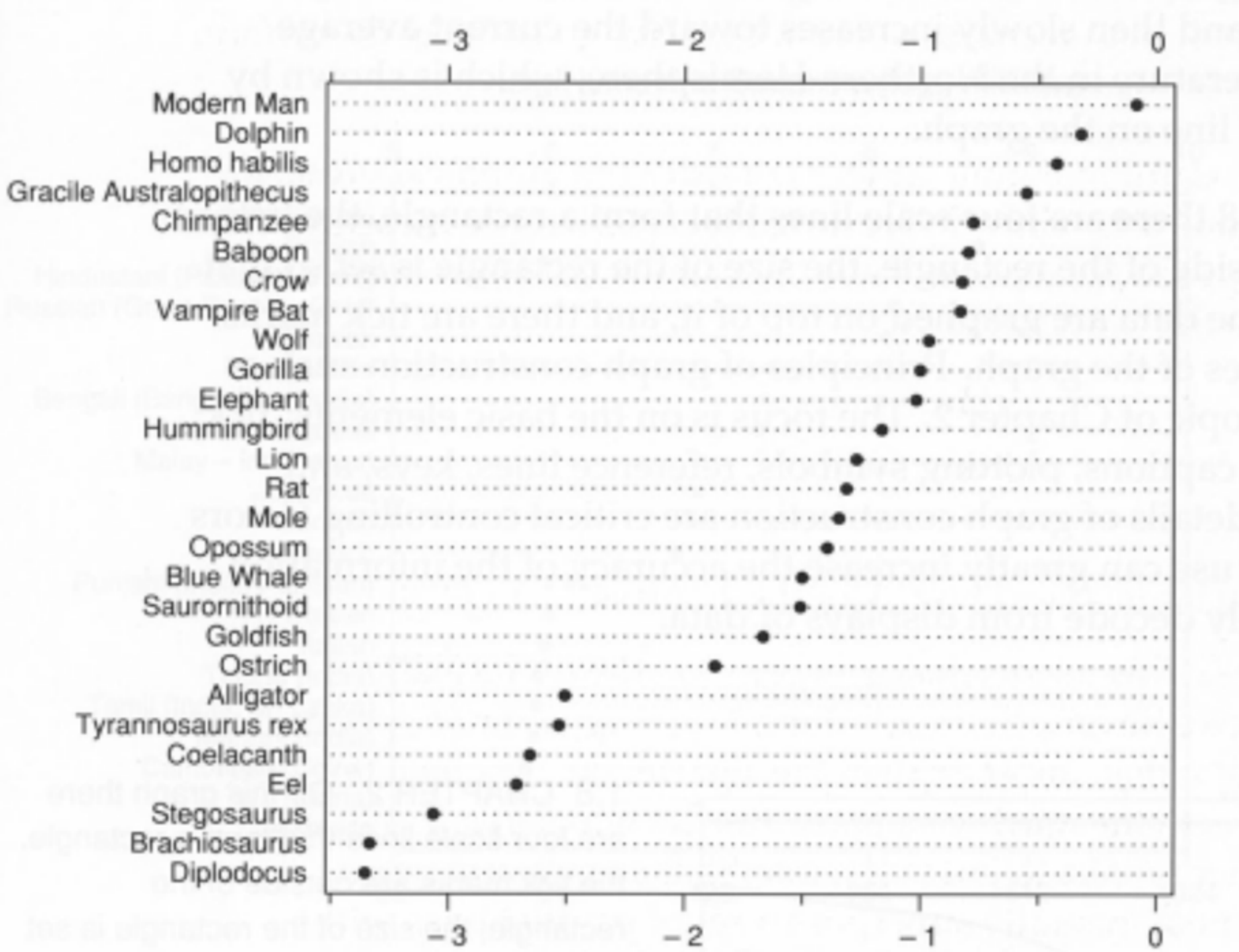
animal

Ready

Which animal is smartest?

- Carl Segan


Brain mass in grams
10,000
5,000
1,000
500
100
50
10.0
5.0
1.0
0.5
0.1
0.05
0.01
Dolphin
Elephant
Blue whale
Modern man
Male gorilla
Homo habilis
Tyrannosaurus rex
Gracile Australopithecus
Lion
Chimpanzee
Wolf
Brachiosaurus
Baboon
Saurornithoid
Diplodocus
Ostrich
Stegosaurus
Alligator
Crow
Opossum
Coelacanth
Rat
Vampire bat
Eel
Goldfish
Mole
Hummingbird
0.001
0.01
0.1
1
10
100
1,000
10,000
100,000
Body Mass in Kilograms



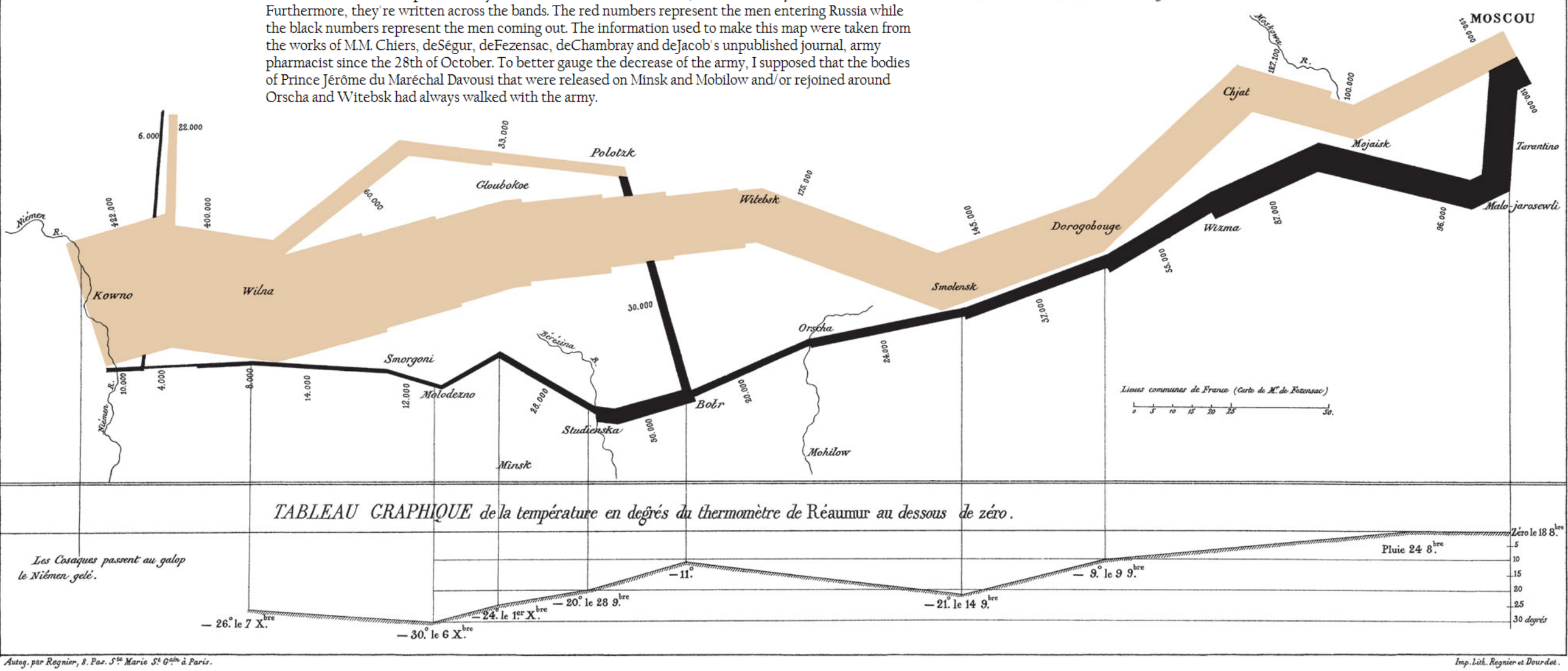
The Elements of Graphing Data
[Cleveland]

Convey Information To Others

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite. Paris, le 20 Novembre 1869.

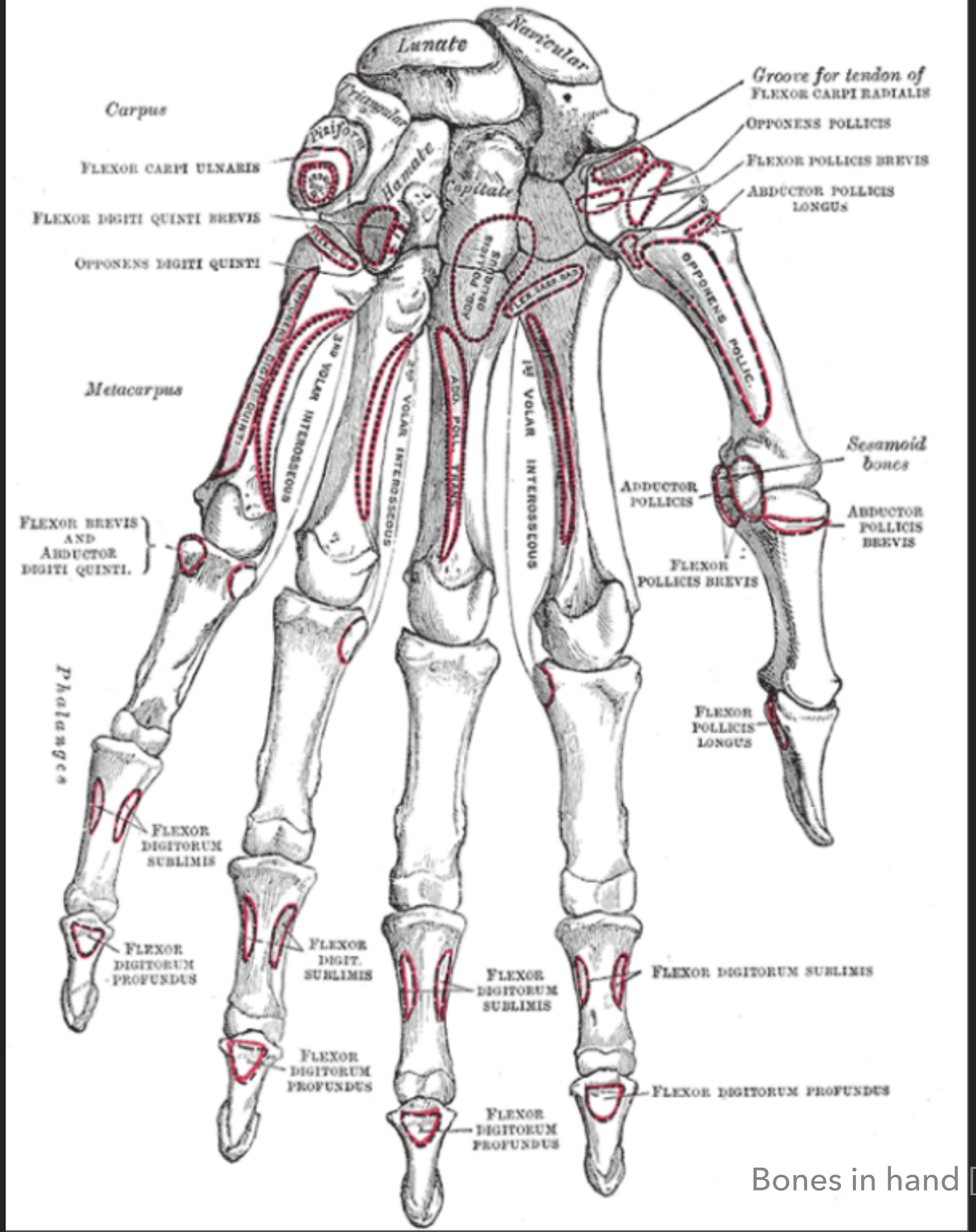
The number of men is represented by the width of the coloured bands, where 1 mm equals 6000 men. Furthermore, they're written across the bands. The red numbers represent the men entering Russia while the black numbers represent the men coming out. The information used to make this map were taken from the works of M.M. Chiers, deSégur, deFezensac, deChambray and deJacob's unpublished journal, army pharmacist since the 28th of October. To better gauge the decrease of the army, I supposed that the bodies of Prince Jérôme du Maréchal Davousi that were released on Minsk and Mobilow and/or rejoined around Orscha and Witebsk had always walked with the army.



Autog. par Regnier, 8. Pas. Ste Marie St Gain à Paris.

Imp. Lith. Regnier et Dourdet.

Napoleon's army in the Russian campaign of 1812 by Charles Joseph Minard



Bones in hand [from 1918 edition]

Double helix model [Watson, Crick, Franklin]

The Value of Visualization

Record information

- Blueprints, photographs, seismographs, …

Analyze data to support reasoning

- Develop and assess hypotheses
- Find patterns / Discover errors in data
- Expand memory

Communicate information to others

- Share and persuade
- Collaborate and revise

Goals of Visualization Research

1 **Understand** how visualizations convey information

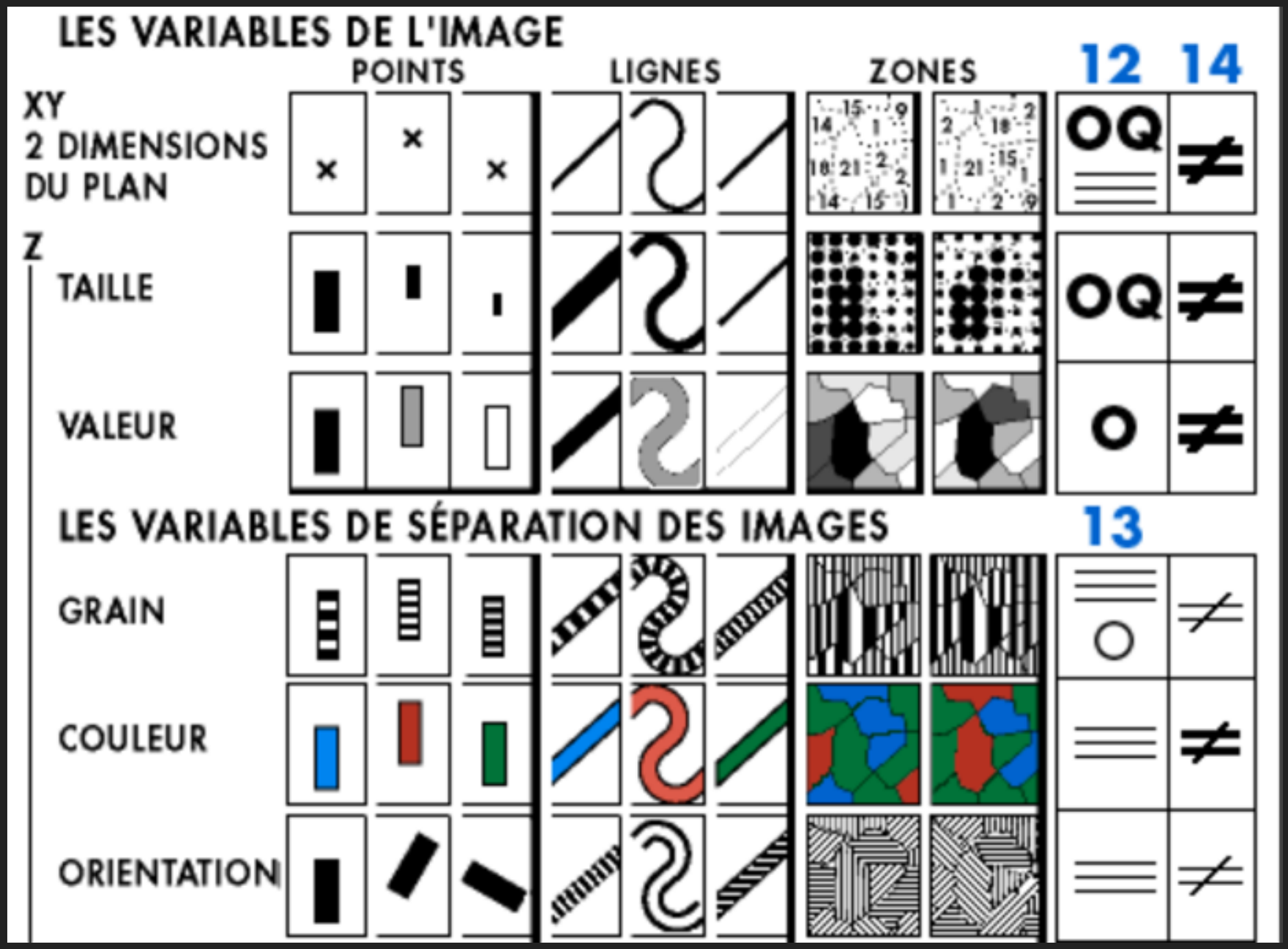
What do people perceive / comprehend?
How do visualizations inform mental models?

2 **Develop principles and techniques** for creating effective visualizations and supporting analysis

Leverage perception & augment cognition
Improve ties between visualization & mental model

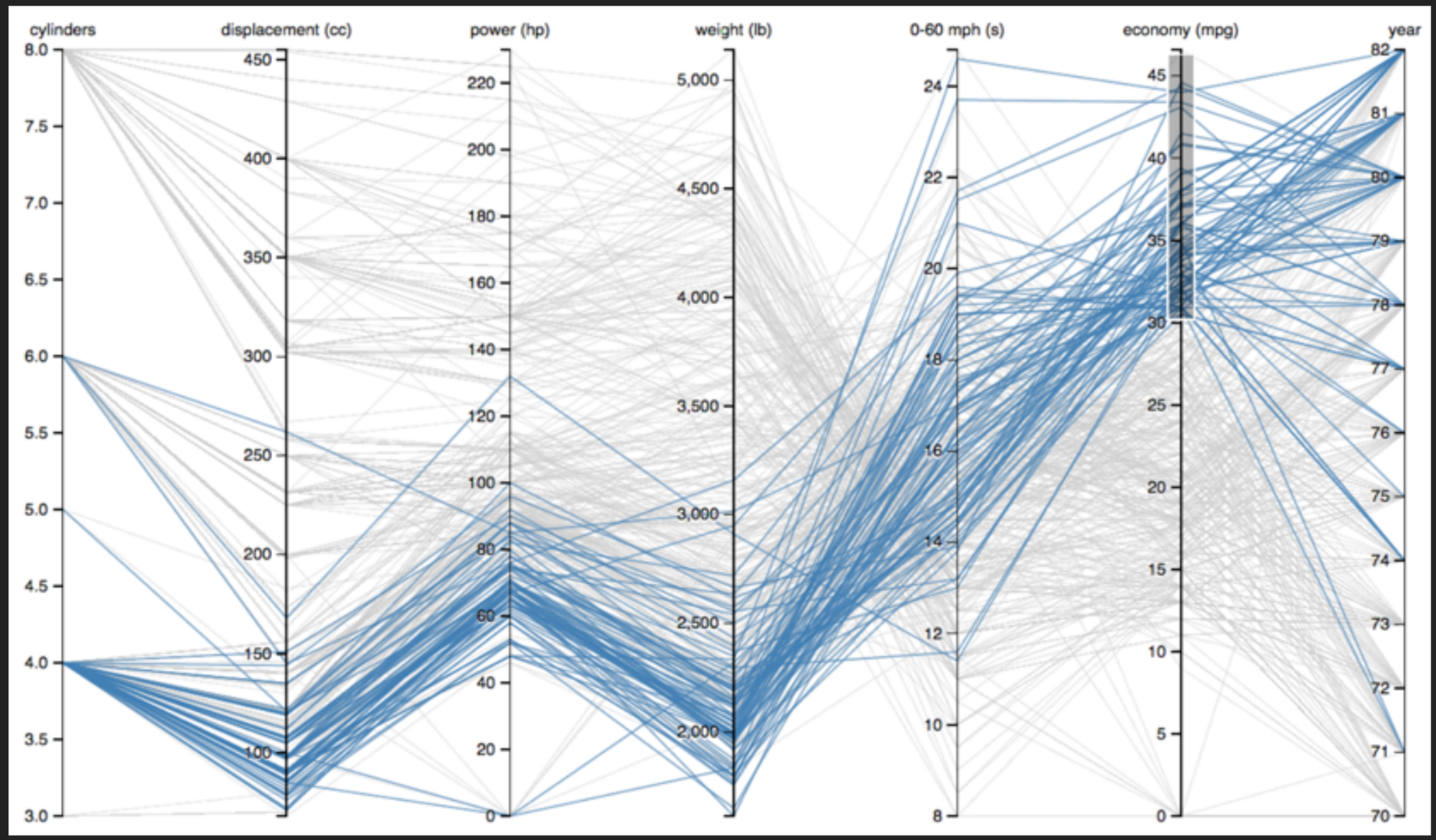
Topics In This Course

Visual Encoding: Marks & Channels



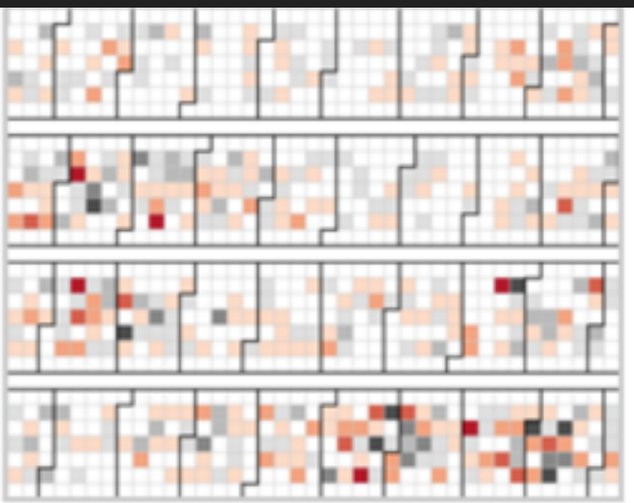
Sémiologie Graphique [Bertin 67]

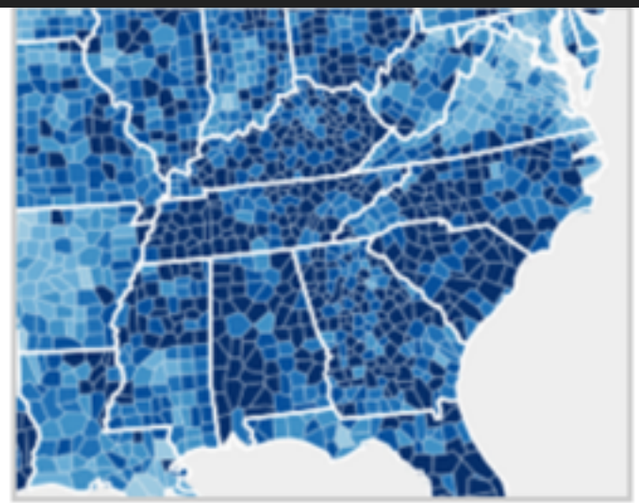
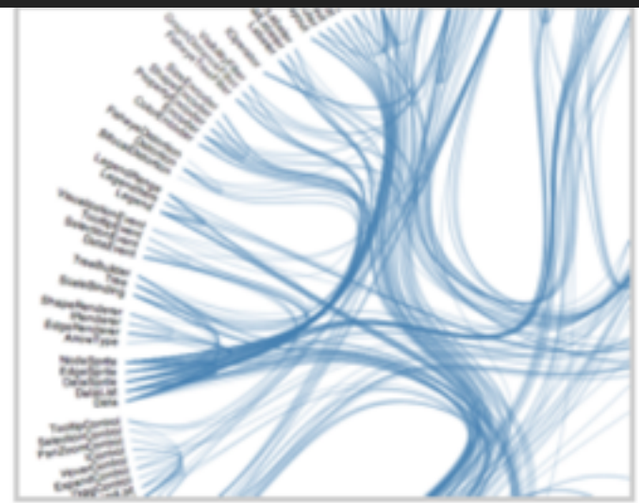
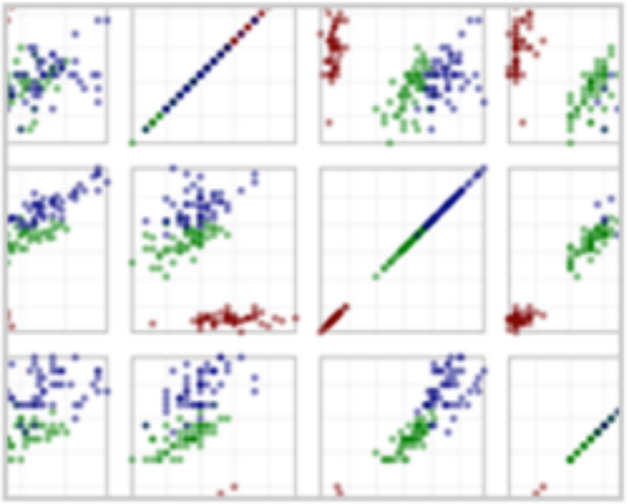
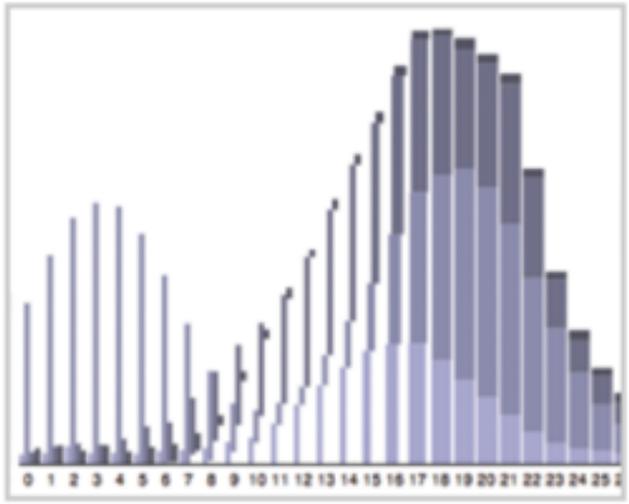
Exploratory Data Analysis: Data & Task Abstractions

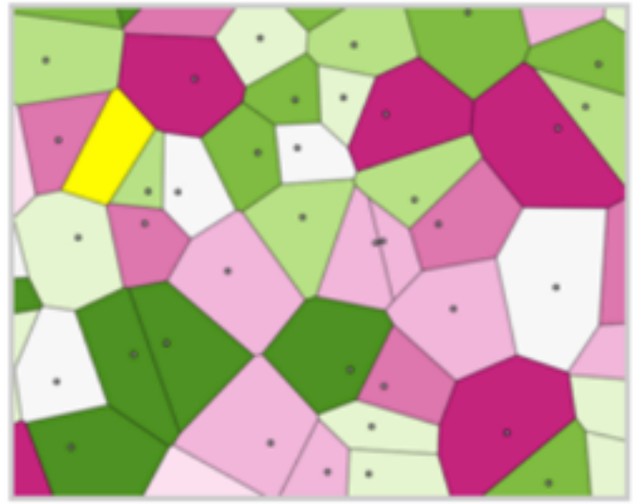


Sémiologie Graphique [Bertin 67]

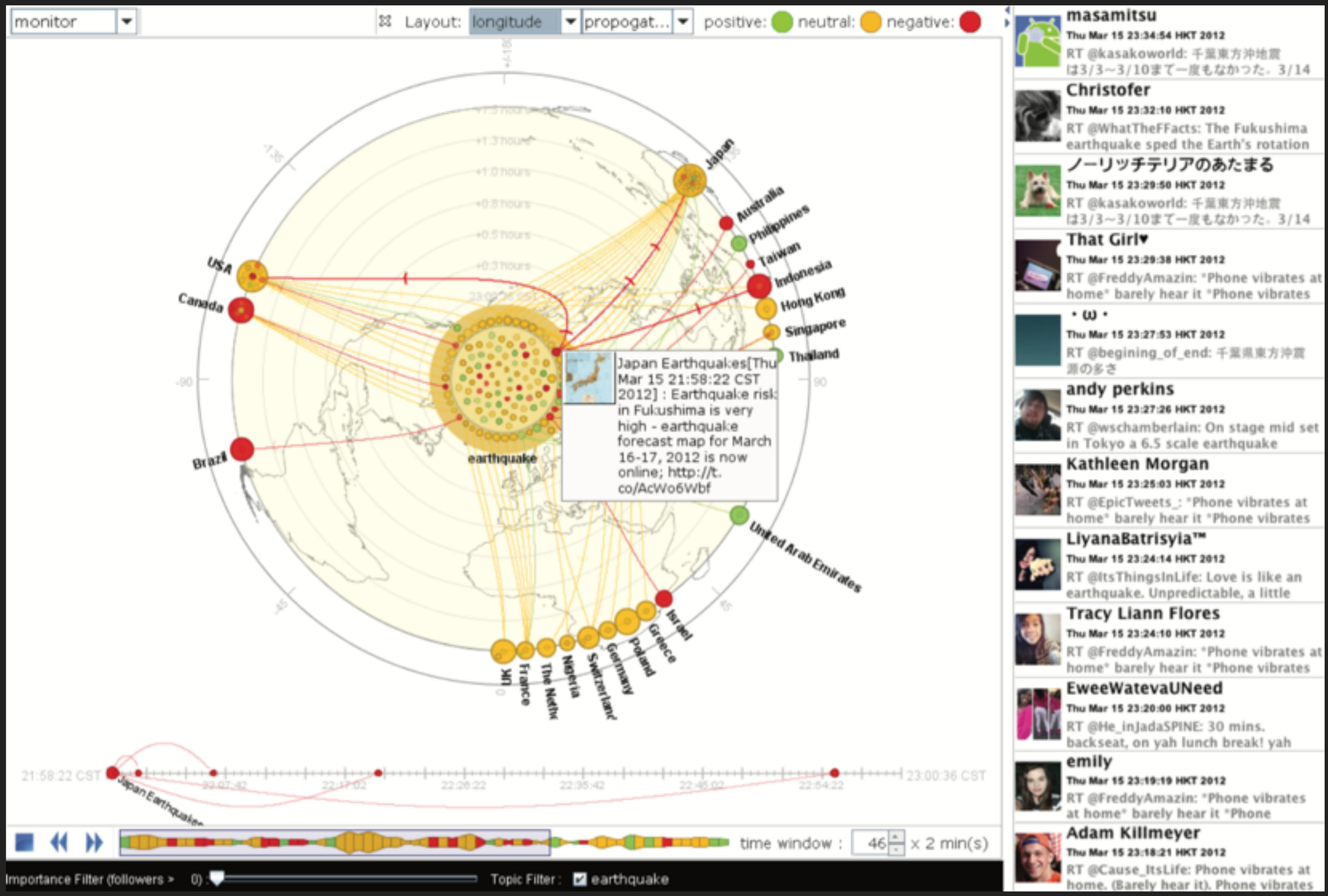
Data Representation: Classic Visualizations



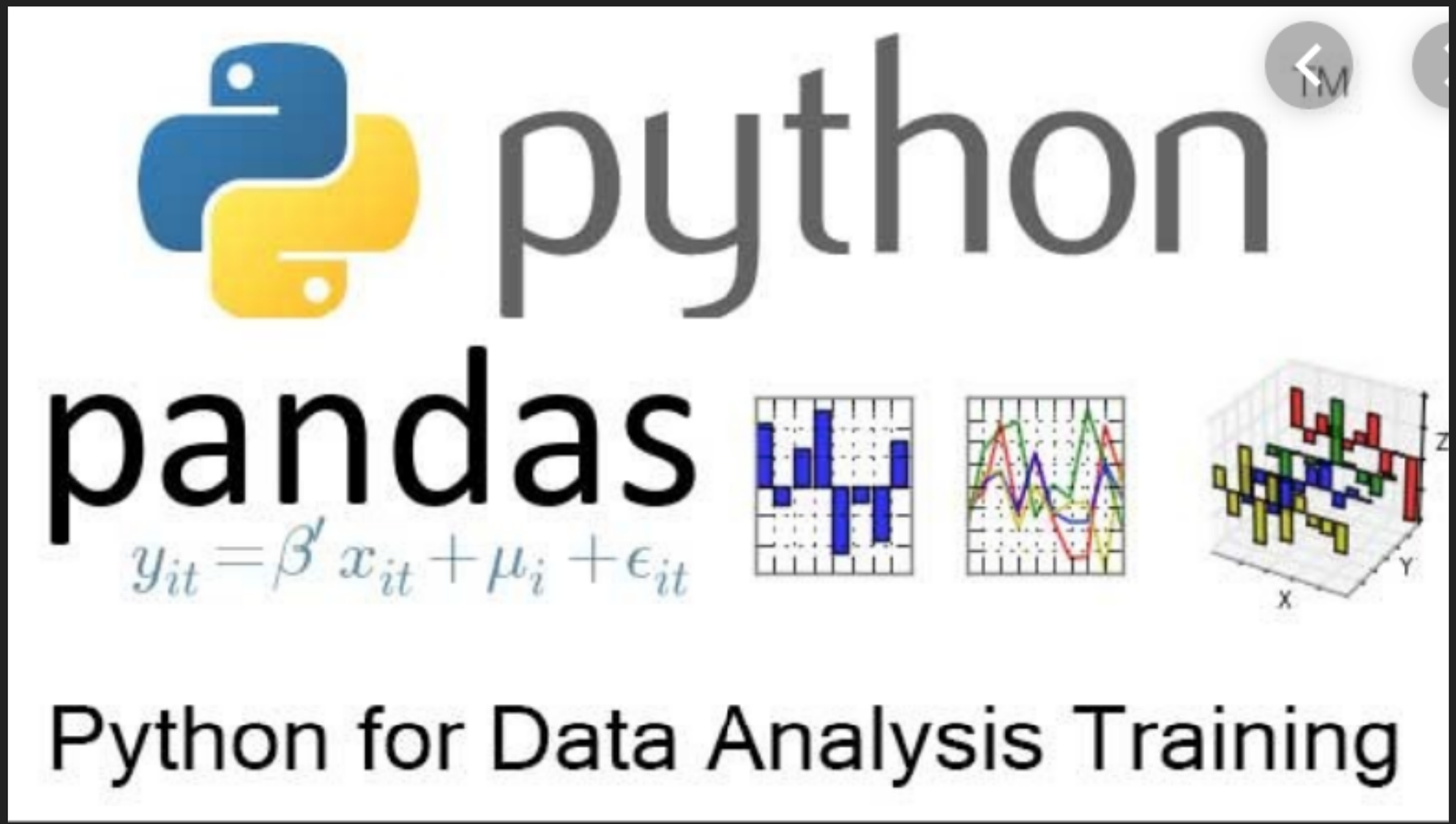



Facet Into Multiple Views



Whisper by Cao et al, 2013

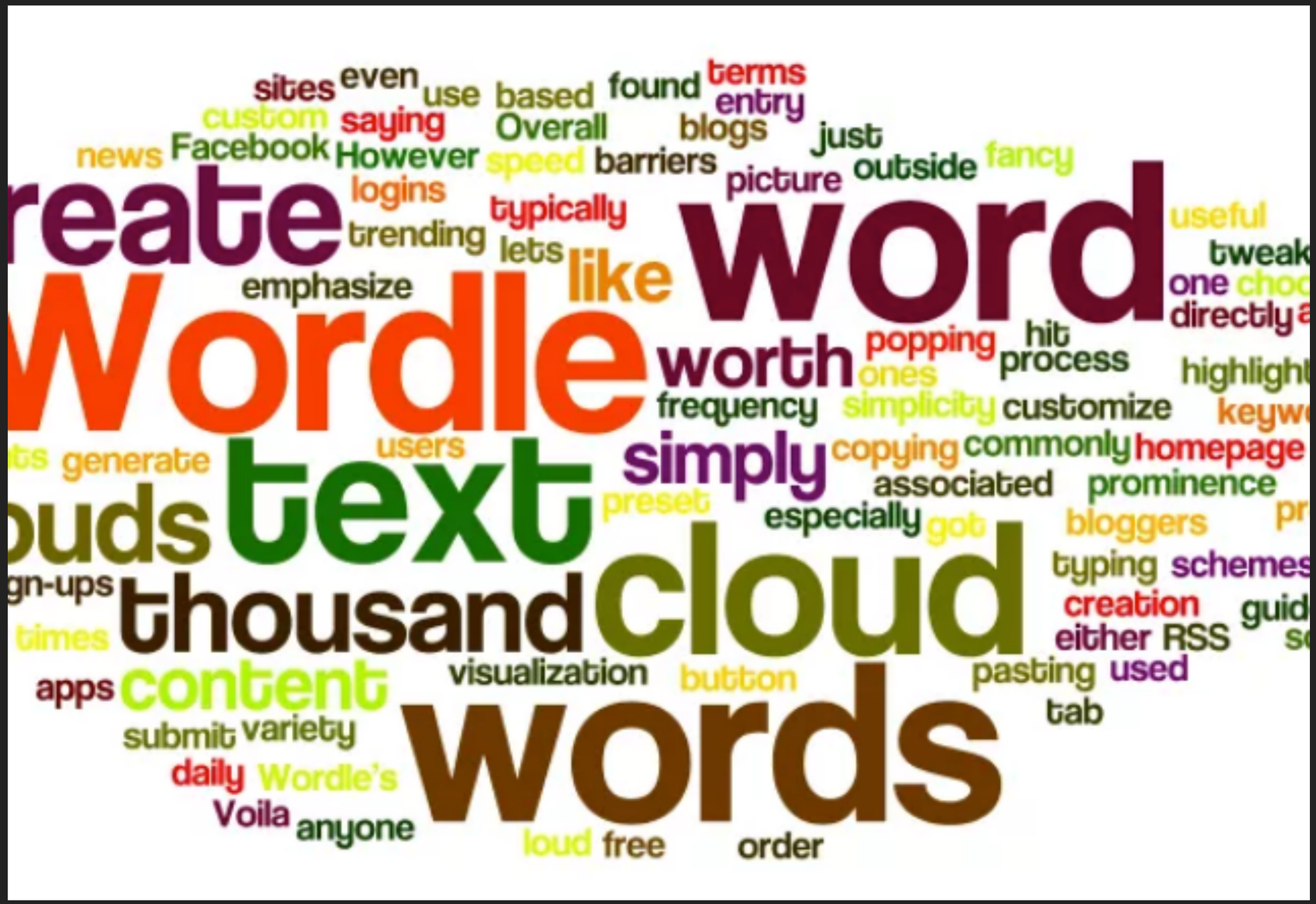
Data Analysis



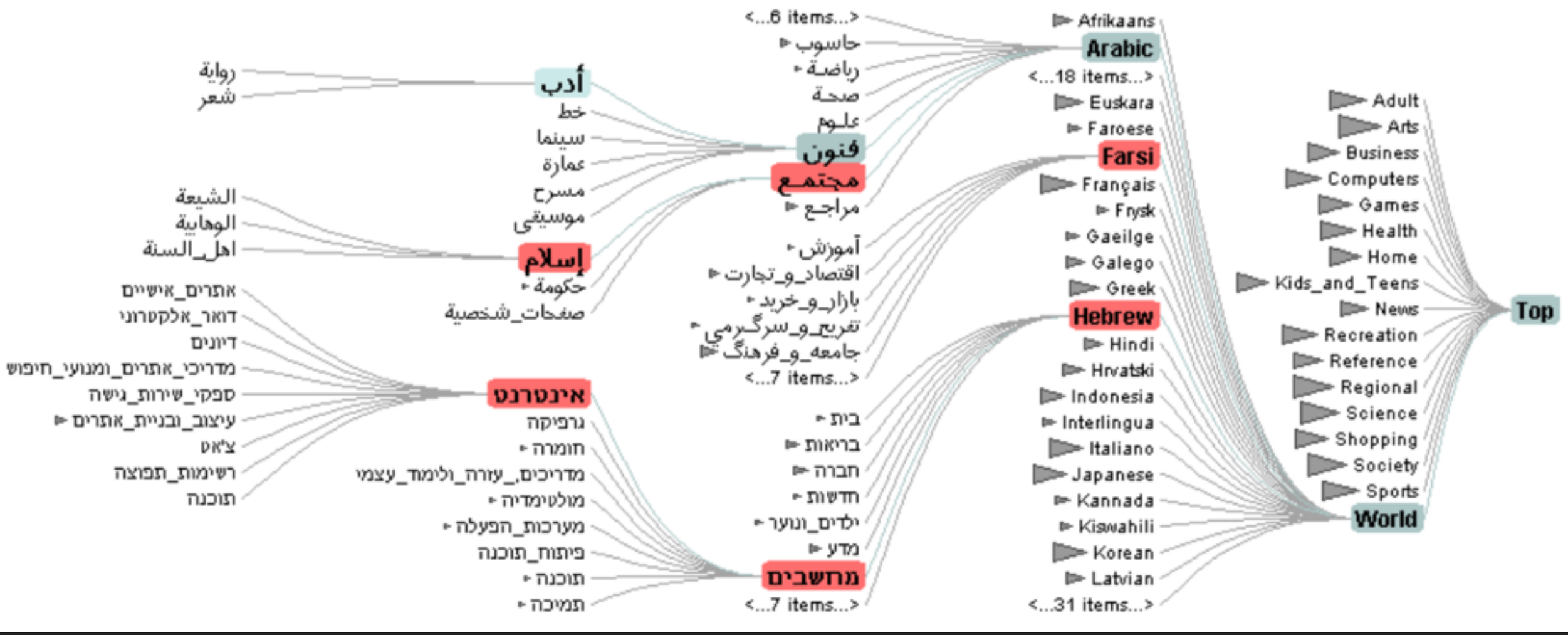
Cartographic Visualization

Temporal Visualization

Text Visualization

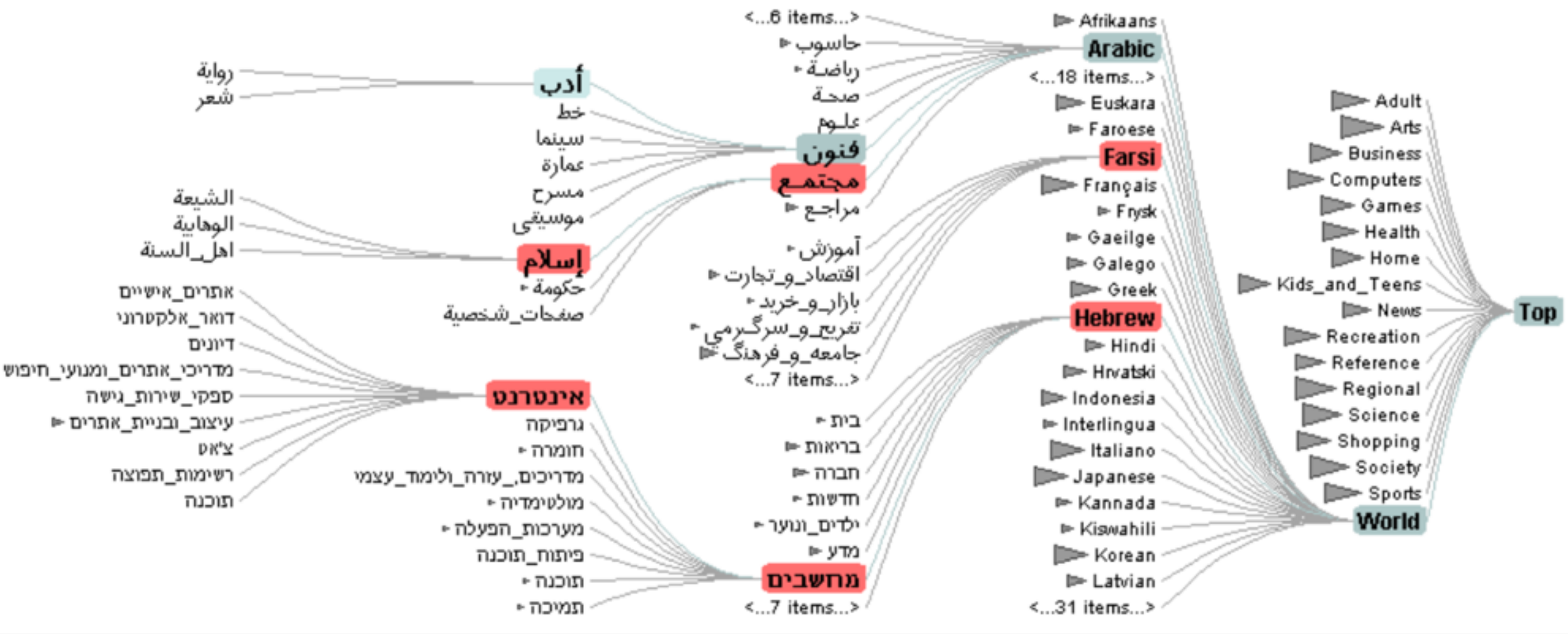


Network & Tree Visualization



Degree-Of-Interest Trees [Heer & Card 04]

High Dimensional Data Visualization



Degree-Of-Interest Trees [Heer & Card 04]

Toolkit Framework