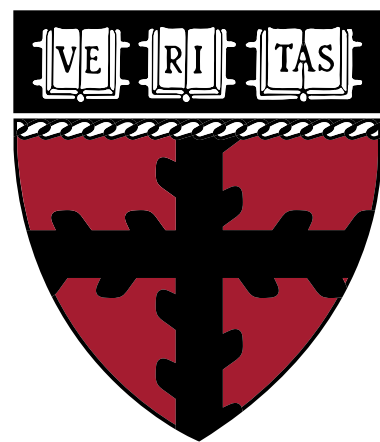
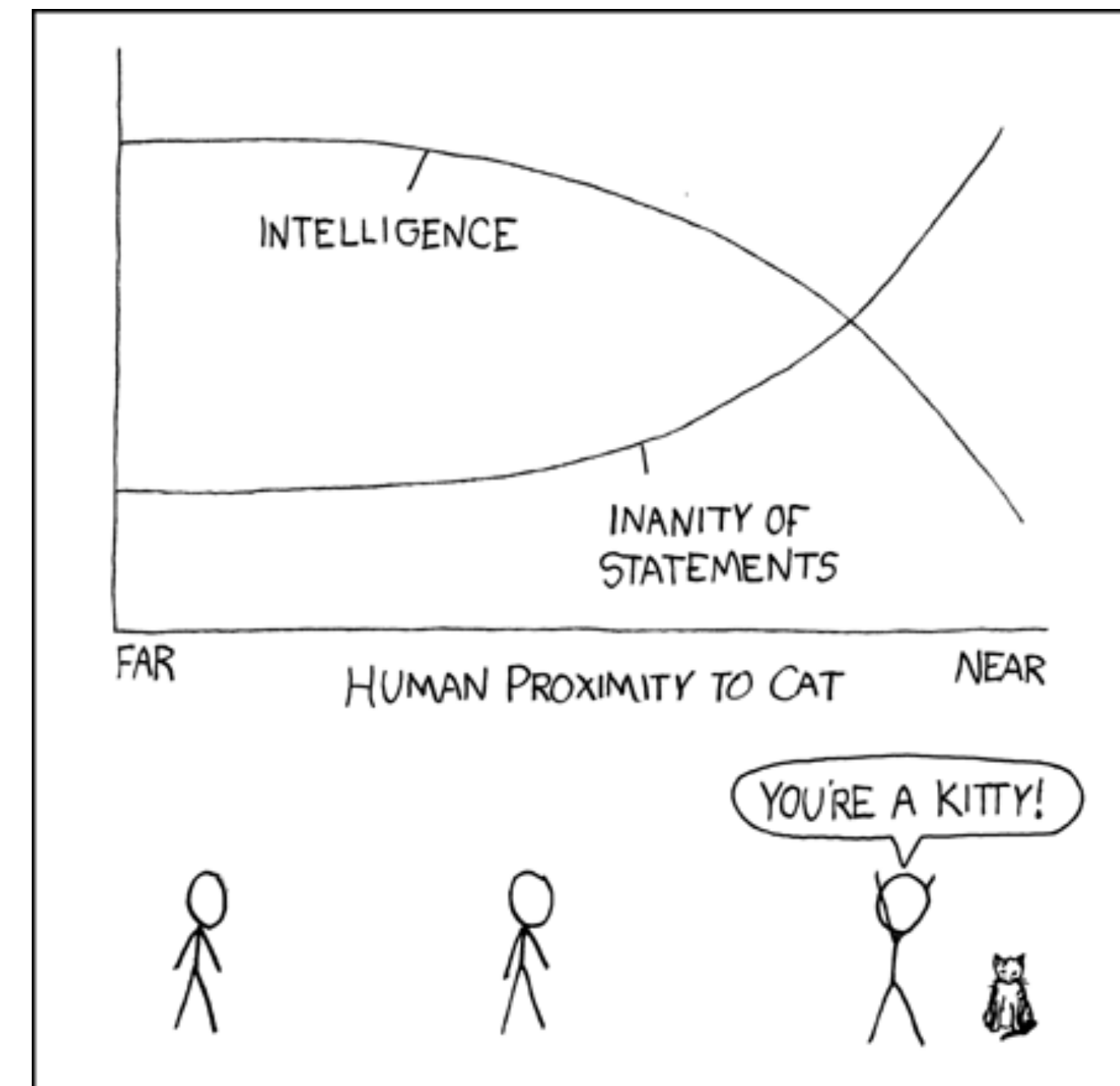


CS171 Visualization

Alexander Lex
alex@seas.harvard.edu



HARVARD
School of Engineering
and Applied Sciences



[xkcd]



vi · su · al · i · za · tion

1. Formation of mental visual images
2. The act or process of interpreting in visual terms or of putting into visible form

Visualization Definition

Visualization is the process that **transforms** (abstract) **data** into **interactive graphical representations** for the purpose of **exploration, confirmation, or presentation.**

Why Visualize?

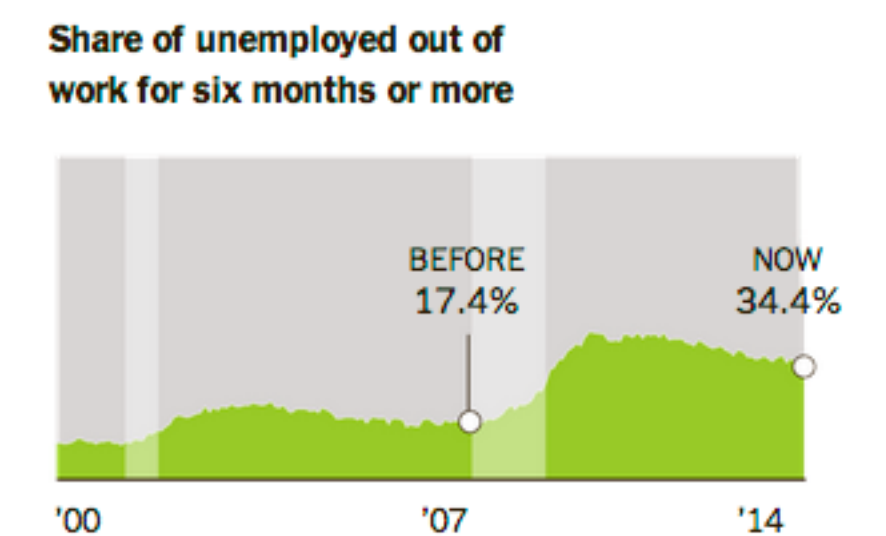
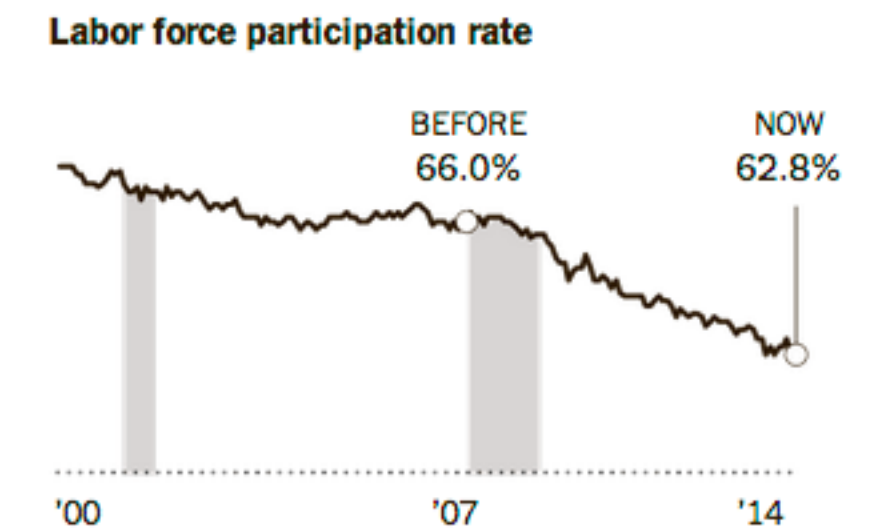
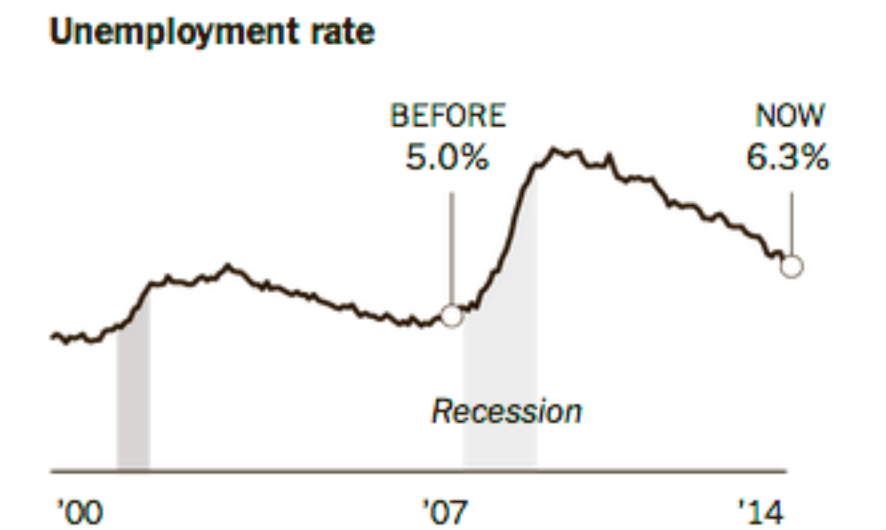
To inform humans: **Communication**

How did the unemployment and labor force develop over the last years?

When questions are not well defined: **Exploration**

Which combination of genes causes cancer?

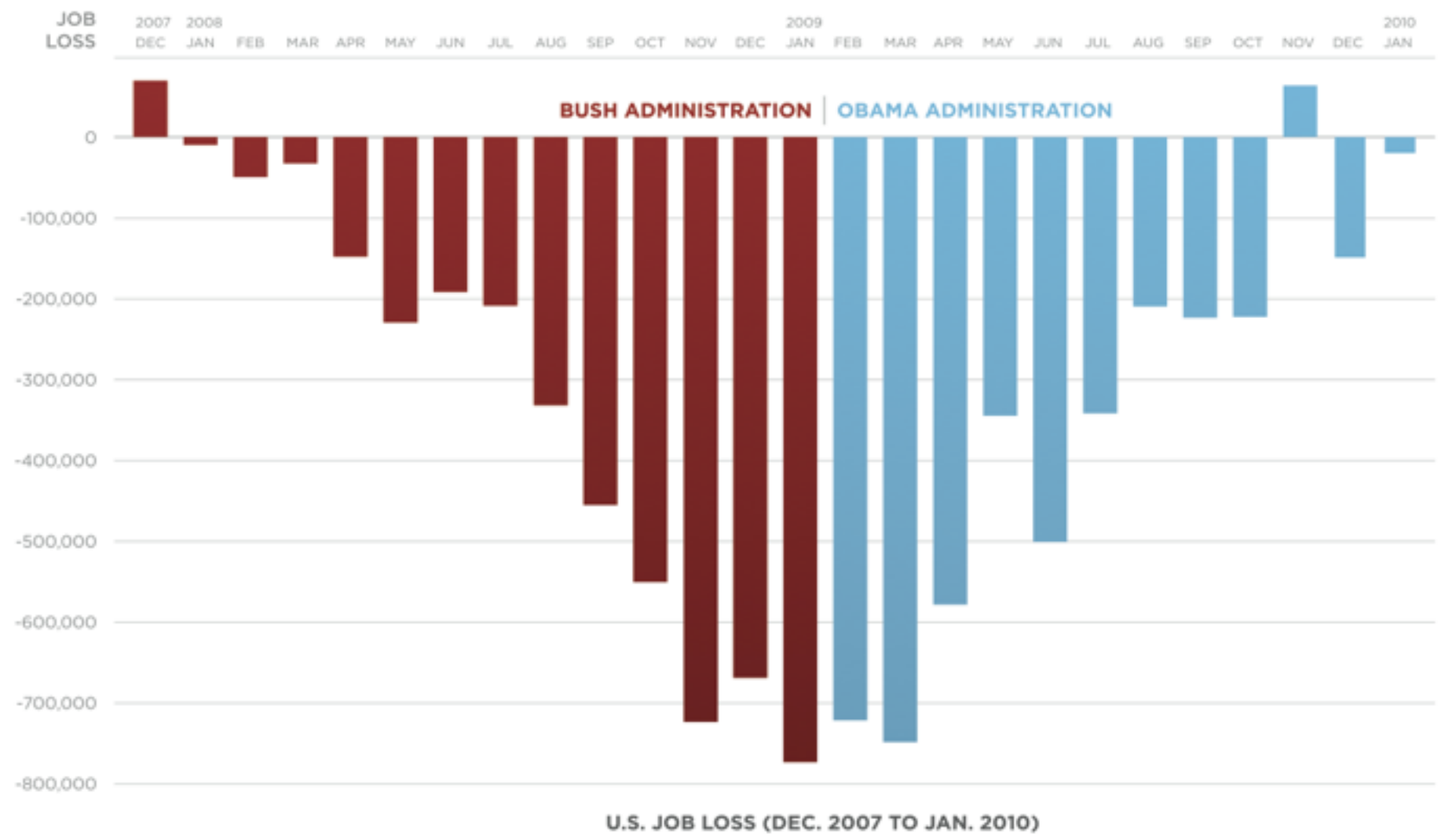
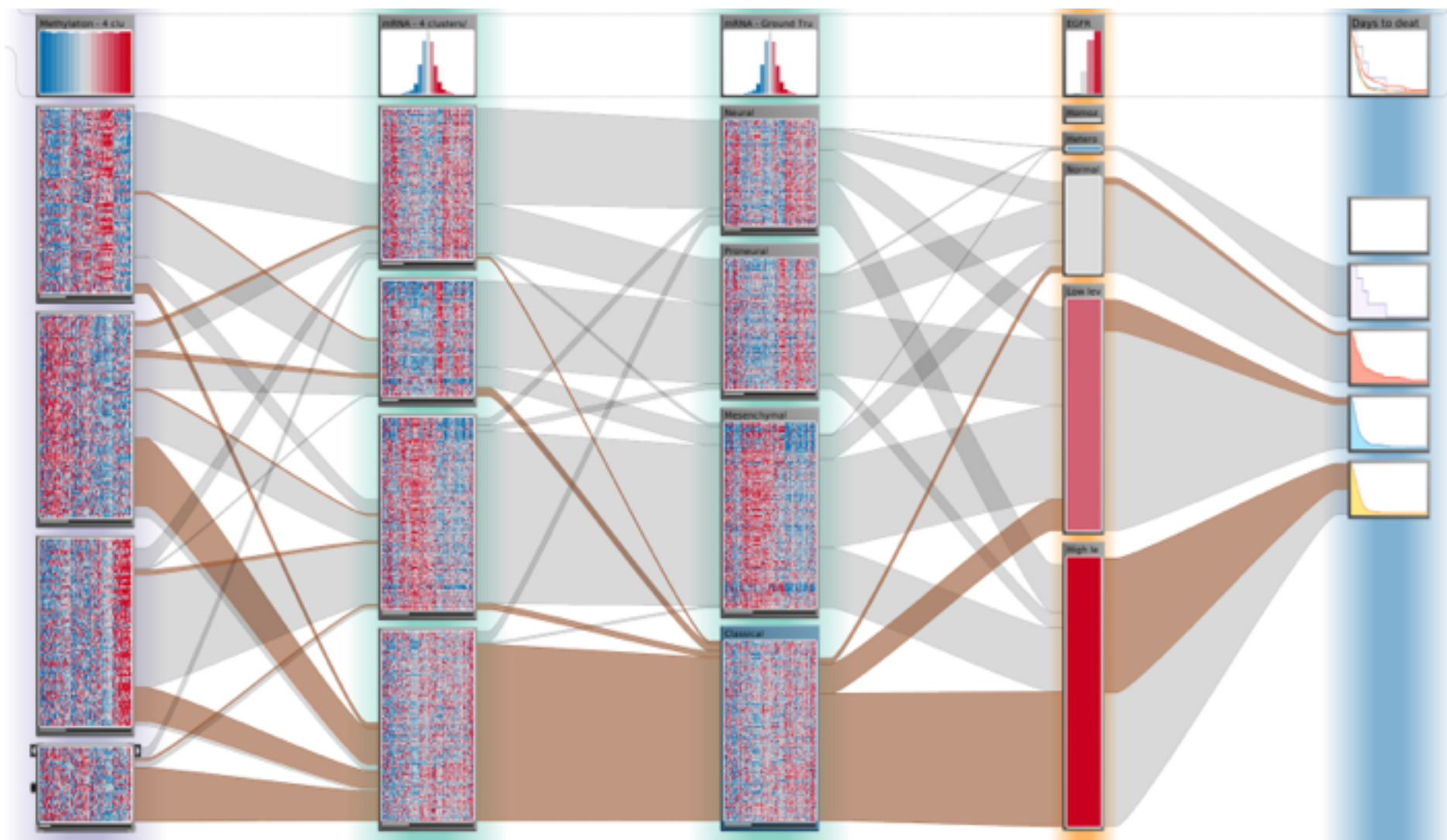
Which drug can help patient X?



[New York Times]

Purpose of Visualization

[Obama Administration]



Open Exploration

Confirmation

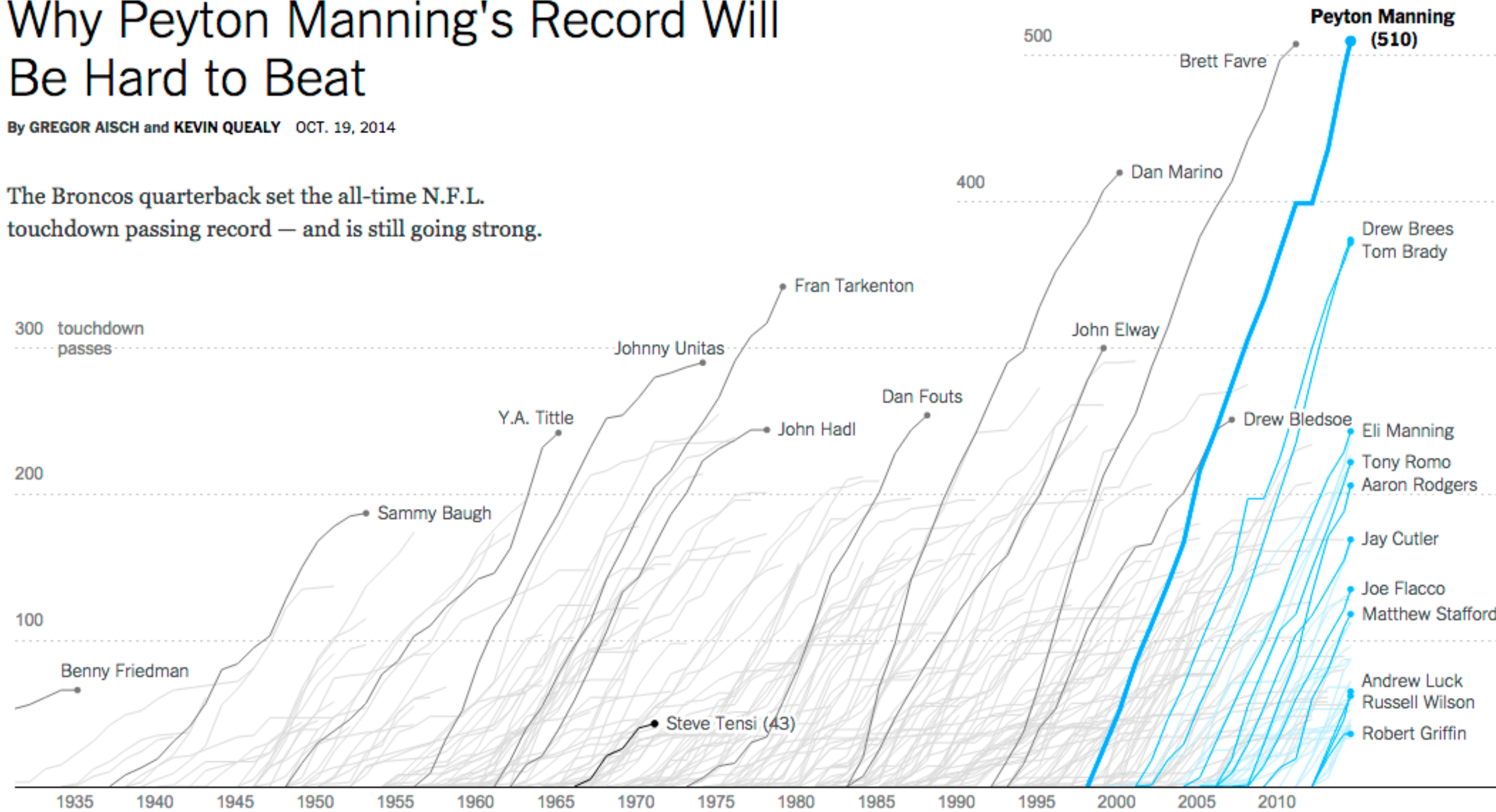
Communication

Example Communication

Why Peyton Manning's Record Will Be Hard to Beat

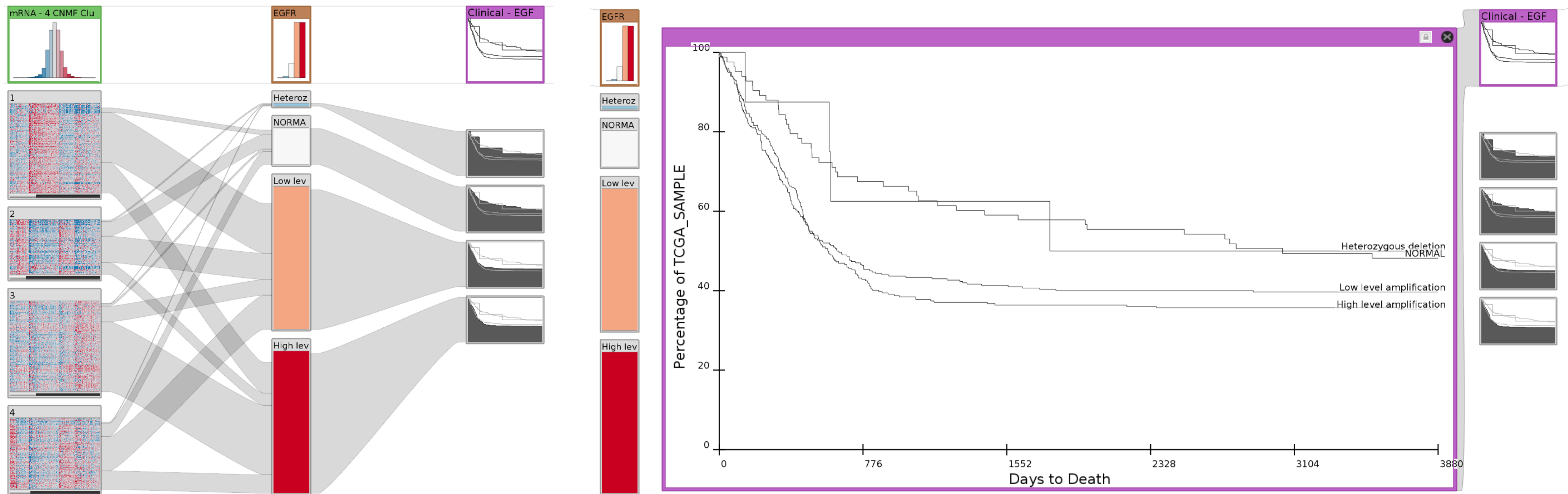
By GREGOR AISCH and KEVIN QUEALY OCT. 19, 2014

The Broncos quarterback set the all-time N.F.L. touchdown passing record — and is still going strong.



[New York Times]

Example Exploration: Cancer Subtypes



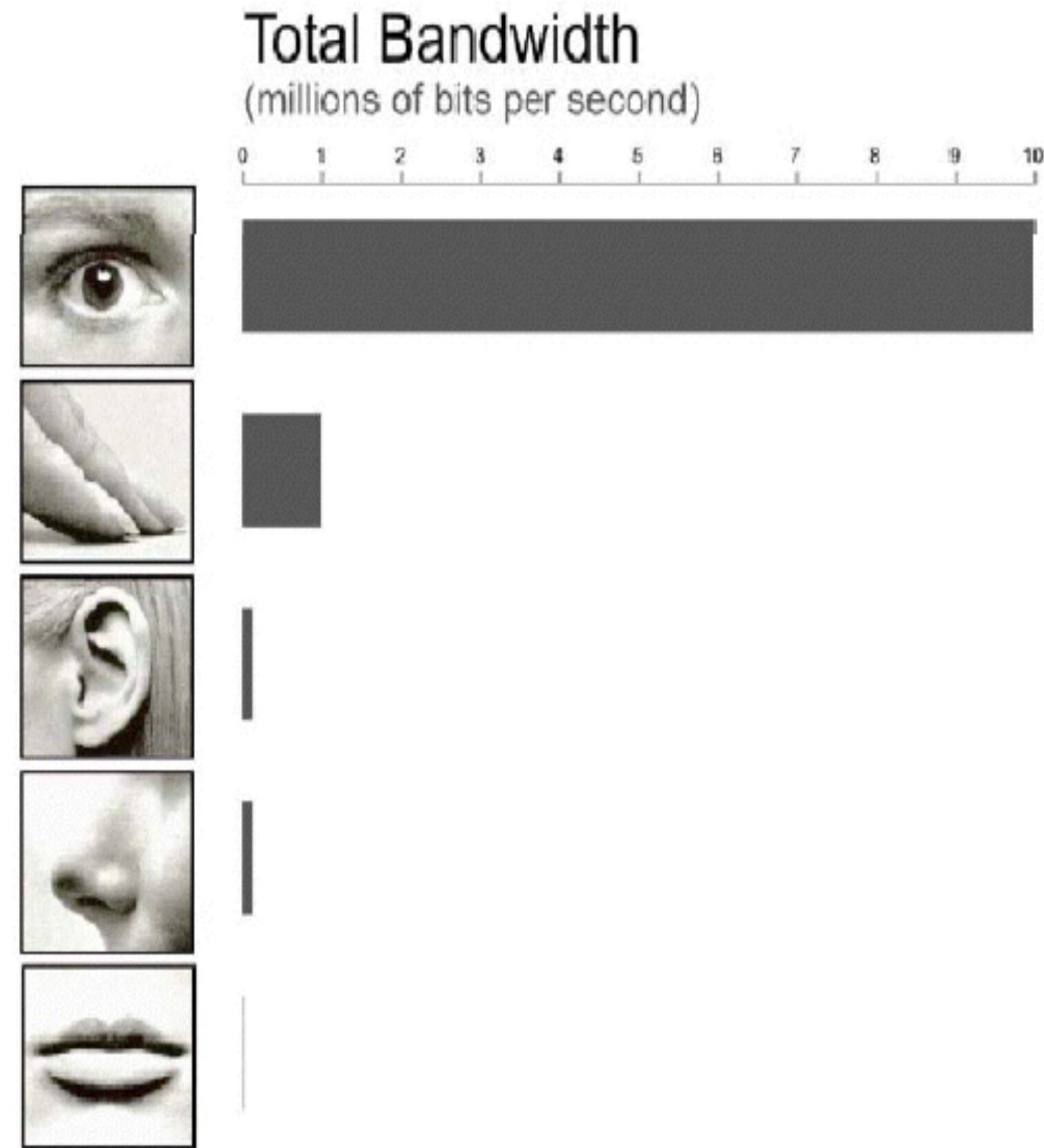
Why Graphics?

Figures are **richer**; provide more information with less clutter and in less space.

Figures provide the *gestalt* effect: they give an overview; **make structure more visible**.

Figures are **more accessible**, easier to understand, **faster to grasp**, more comprehensible, **more memorable**, more fun, and less formal.

list adapted from: [Stasko et al. 1998]



When not to visualize?

When to automate?

Well defined question on well-defined dataset

Which gene is most frequently mutated in this set of patients?

What is the current unemployment rate?

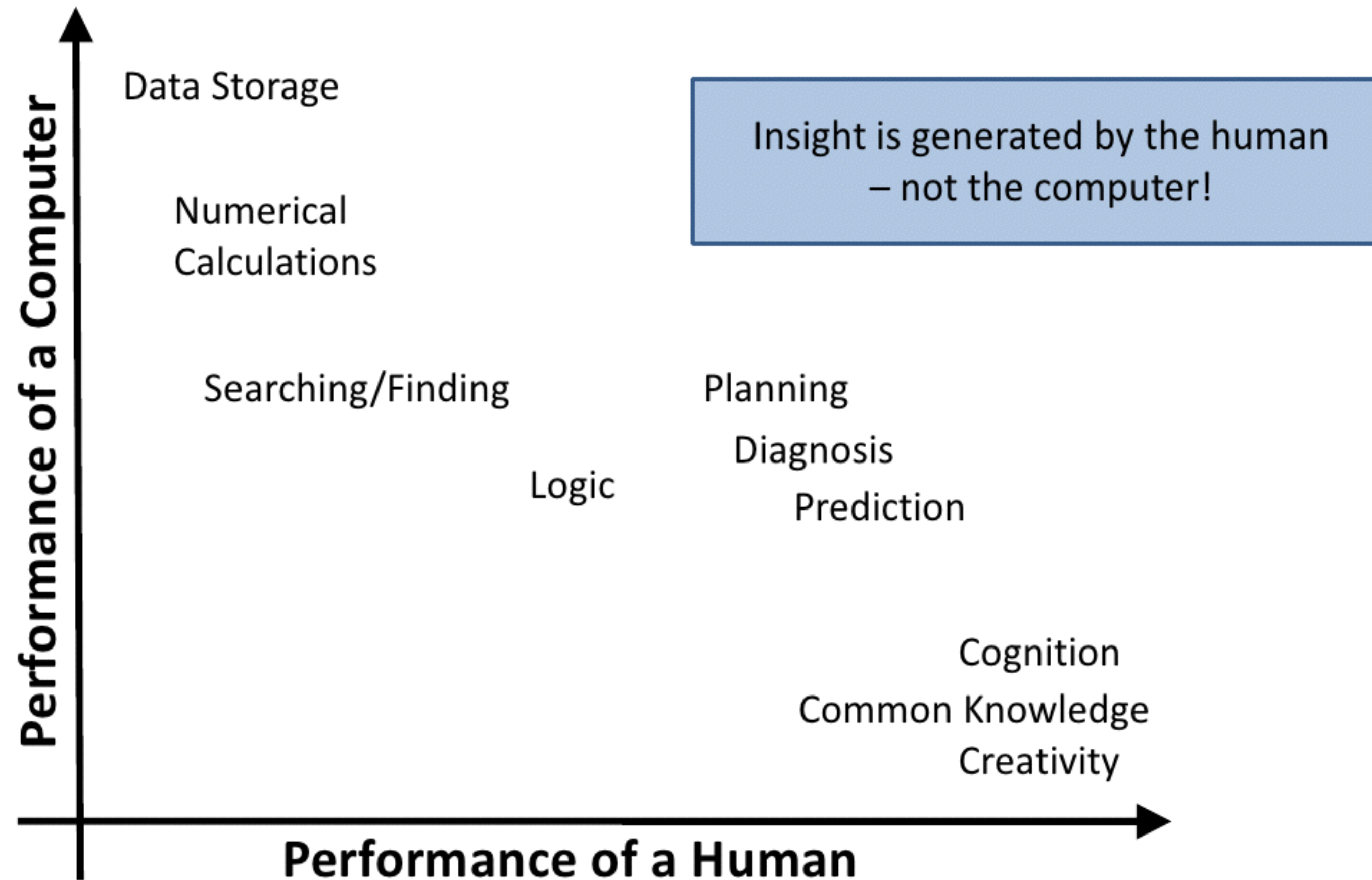
Decisions needed in minimal time

High frequency stock market trading: which stock to buy/sell?

Manufacturing: is bottle broken?



The Ability Matrix



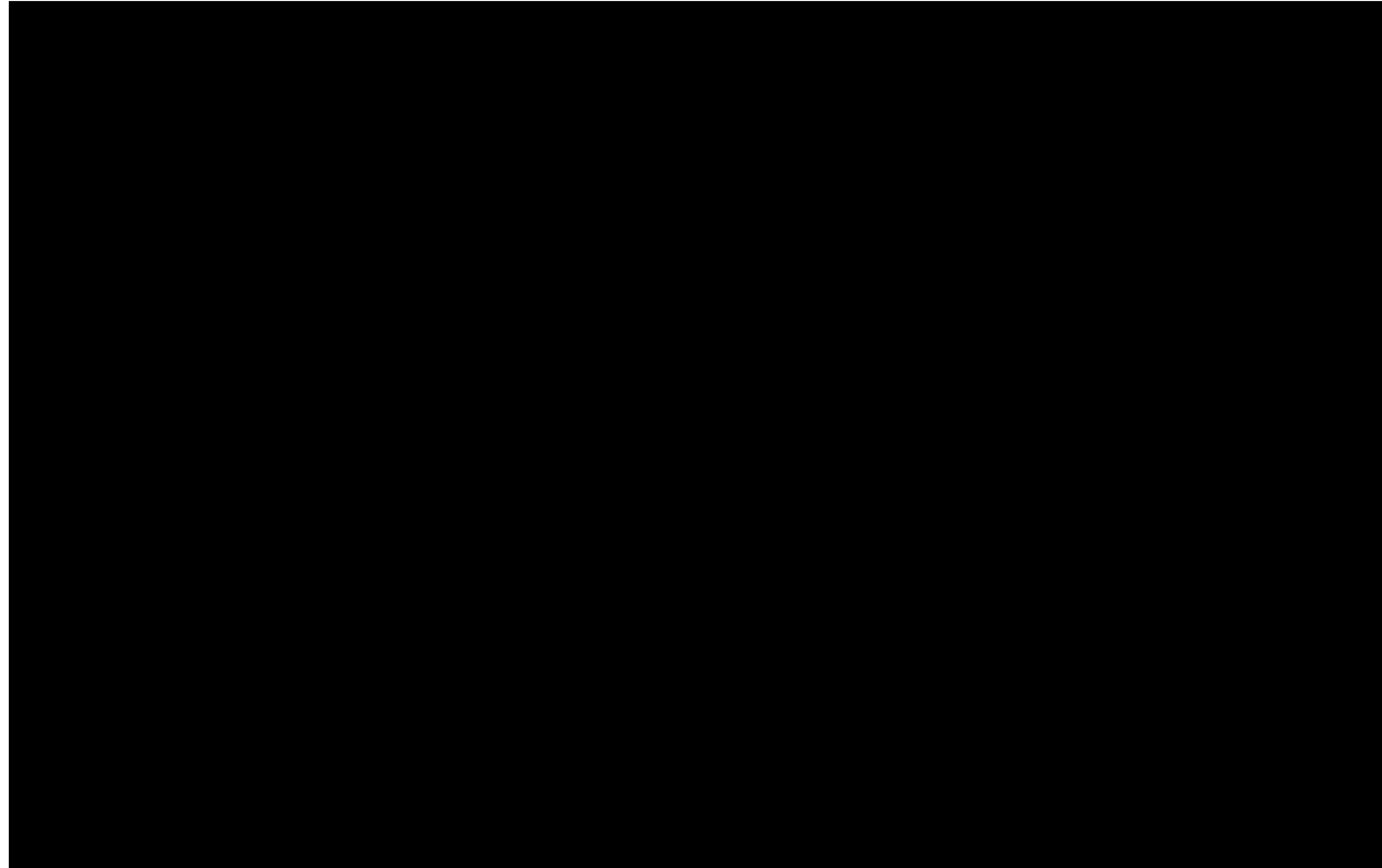
Why User Computers?

Scale

Drawing by hand infeasible

Interaction allows to “drill down” into data

Integration with algorithms



Why User Computers?

Efficiency

Re-use charts for different datasets

Quality

Precise data driven rendering

Storytelling

Use time

Tell Stories

[New York Times]

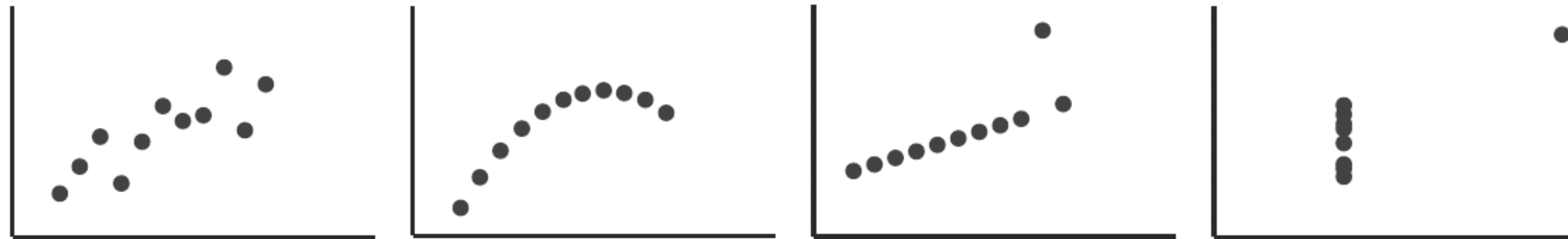


Why not just use Statistics?

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.0	10	9.1	10	7.4	8	6.5
8	6.9	8	8.1	8	6.7	8	5.7
13	7.5	13	8.7	13	12.	8	7.7
9	8.8	9	8.7	9	7.1	8	8.8
11	8.3	11	9.2	11	7.8	8	8.4
14	9.9	14	8.1	14	8.8	8	7.0
6	7.2	6	6.1	6	6.0	8	5.2
4	4.2	4	3.1	4	5.3	19	12.
12	10.	12	9.1	12	8.1	8	5.5
7	4.8	7	7.2	7	6.4	8	7.9
5	5.						6.8

Mean x: 9 y: 7.50
Variance x: 11 y: 4.122
Correlation x - y: 0.816
Linear regression: $y = 3.00 + 0.500x$

Anscombe's Quartett



Mean x: 9 y: 7.50

Variance x: 11 y: 4.122

Correlation x - y: 0.816

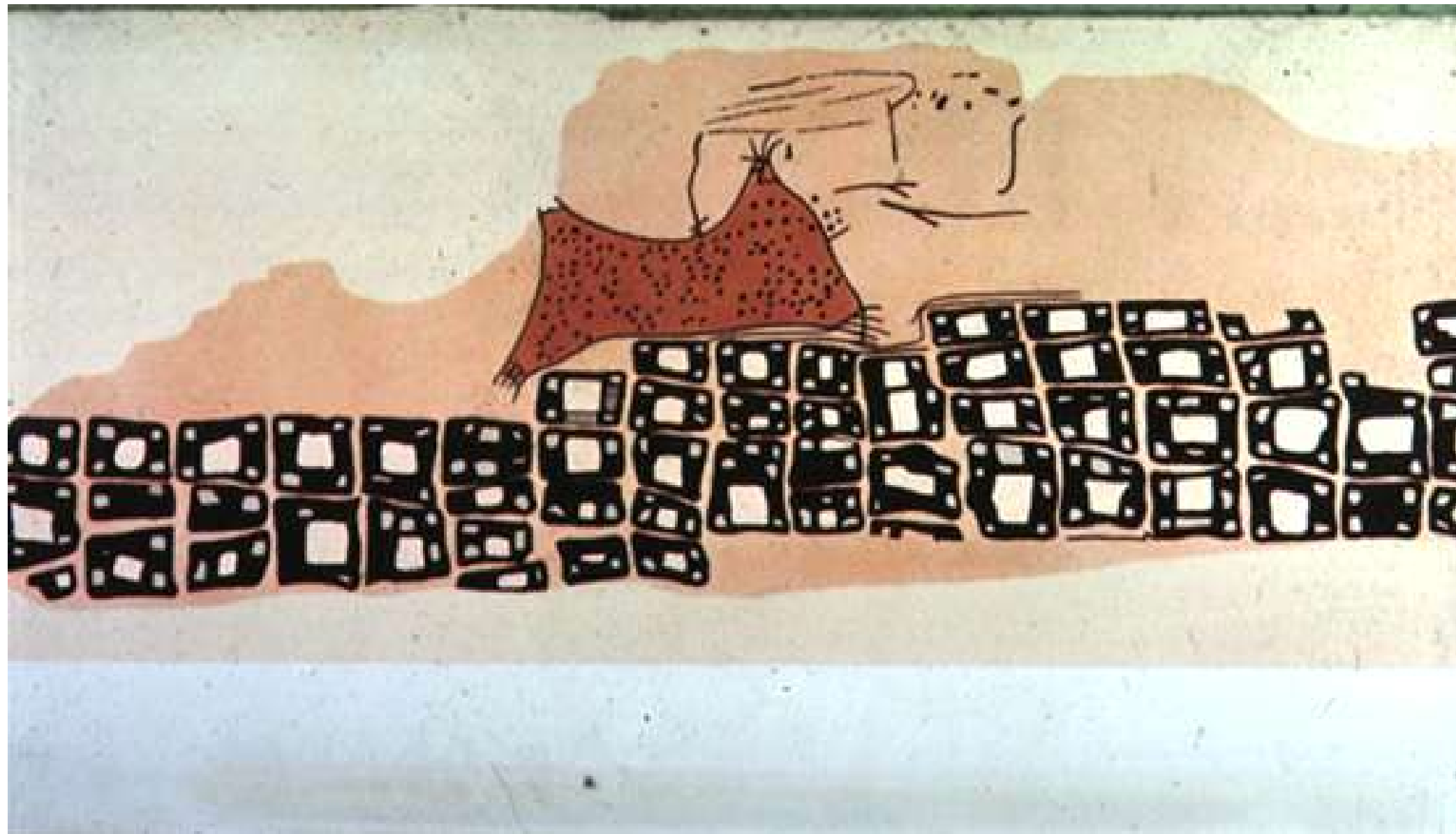
Linear regression: $y = 3.00 + 0.500x$

Good Data Visualization

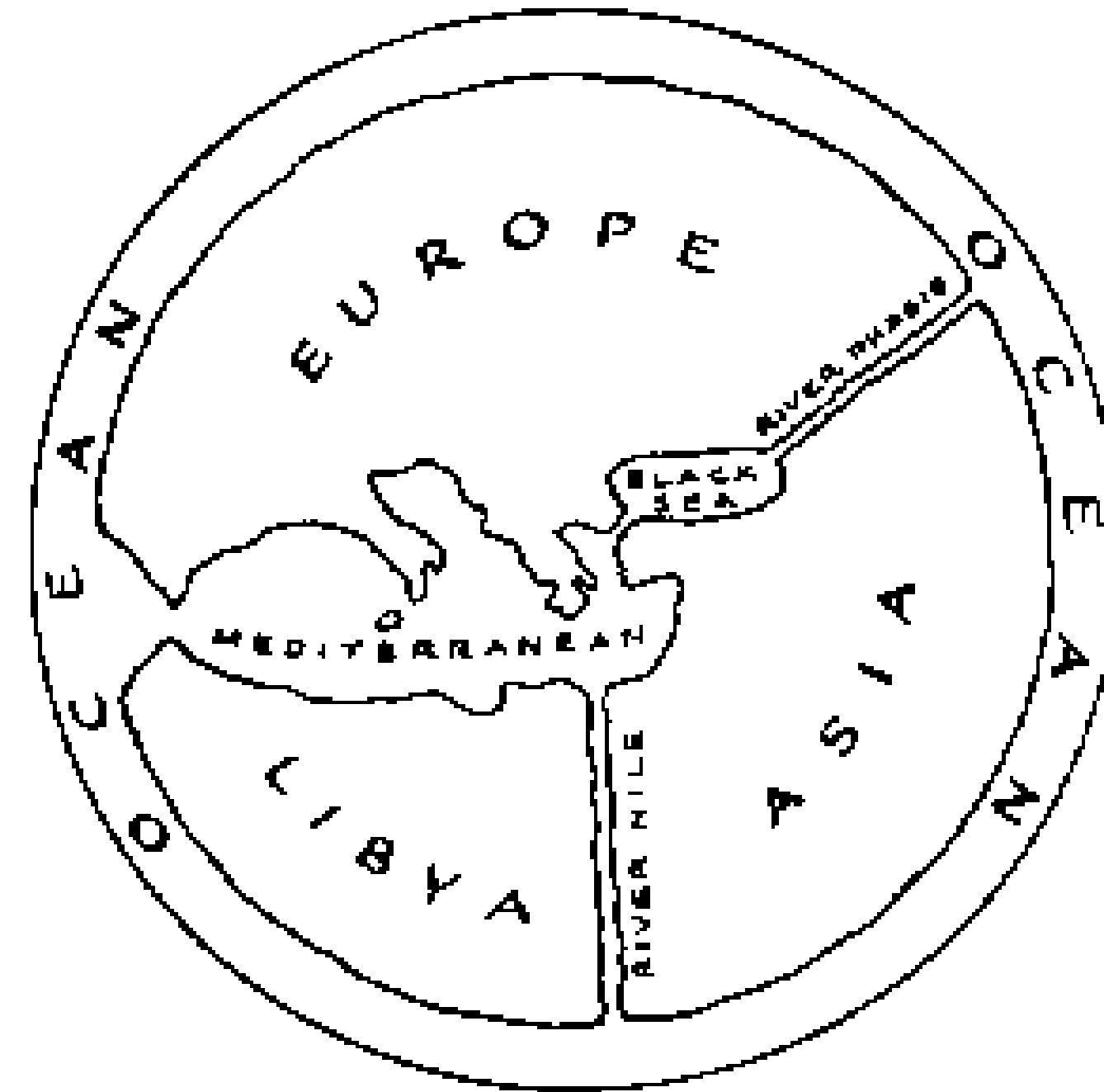
- ... makes data **accessible**
- ... combines strengths of **humans and computers**
- ... enables **insight**
- ... **communicates**

How did we get here?

Record



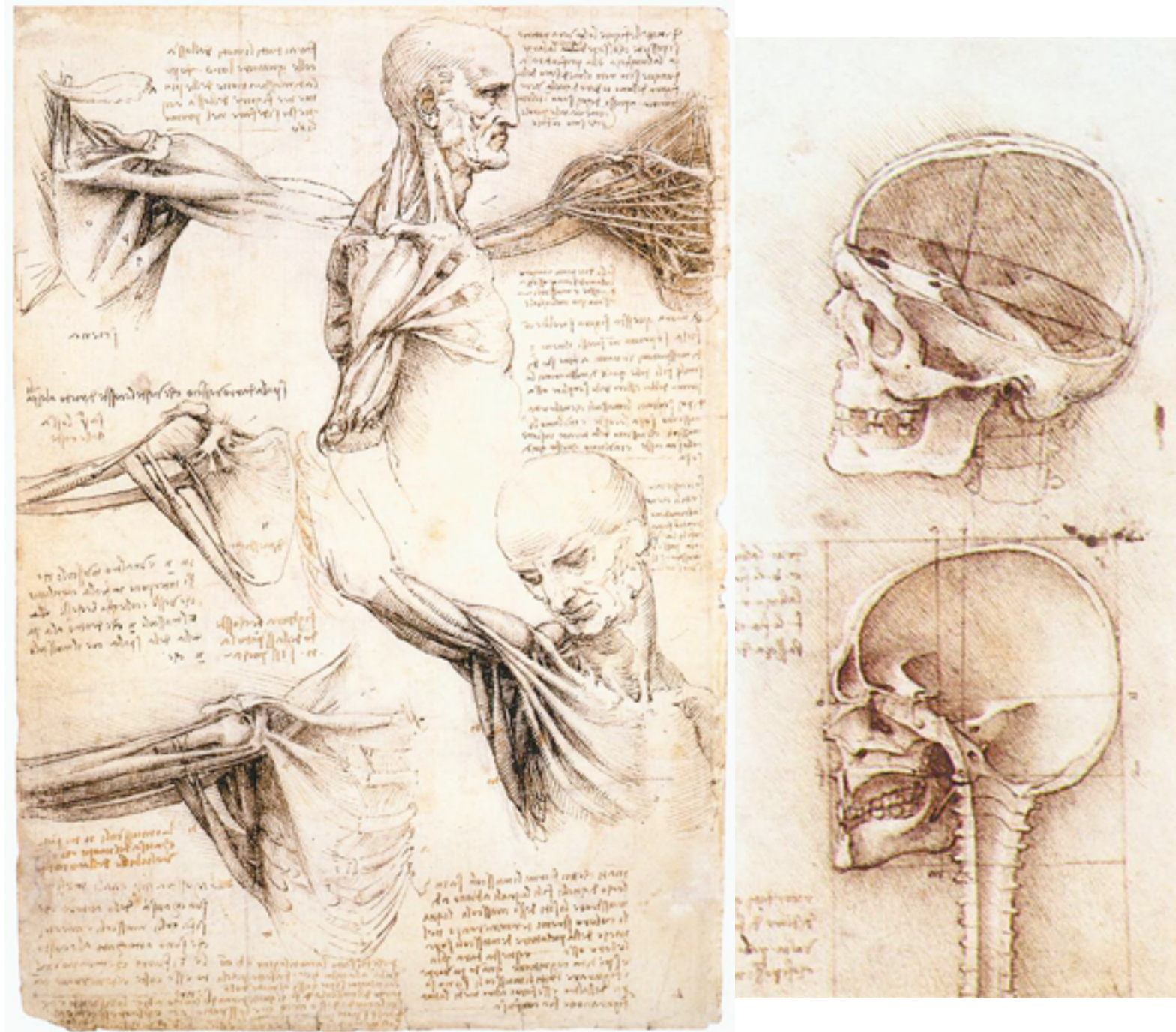
Konya town map, Turkey, c. 6200 BC



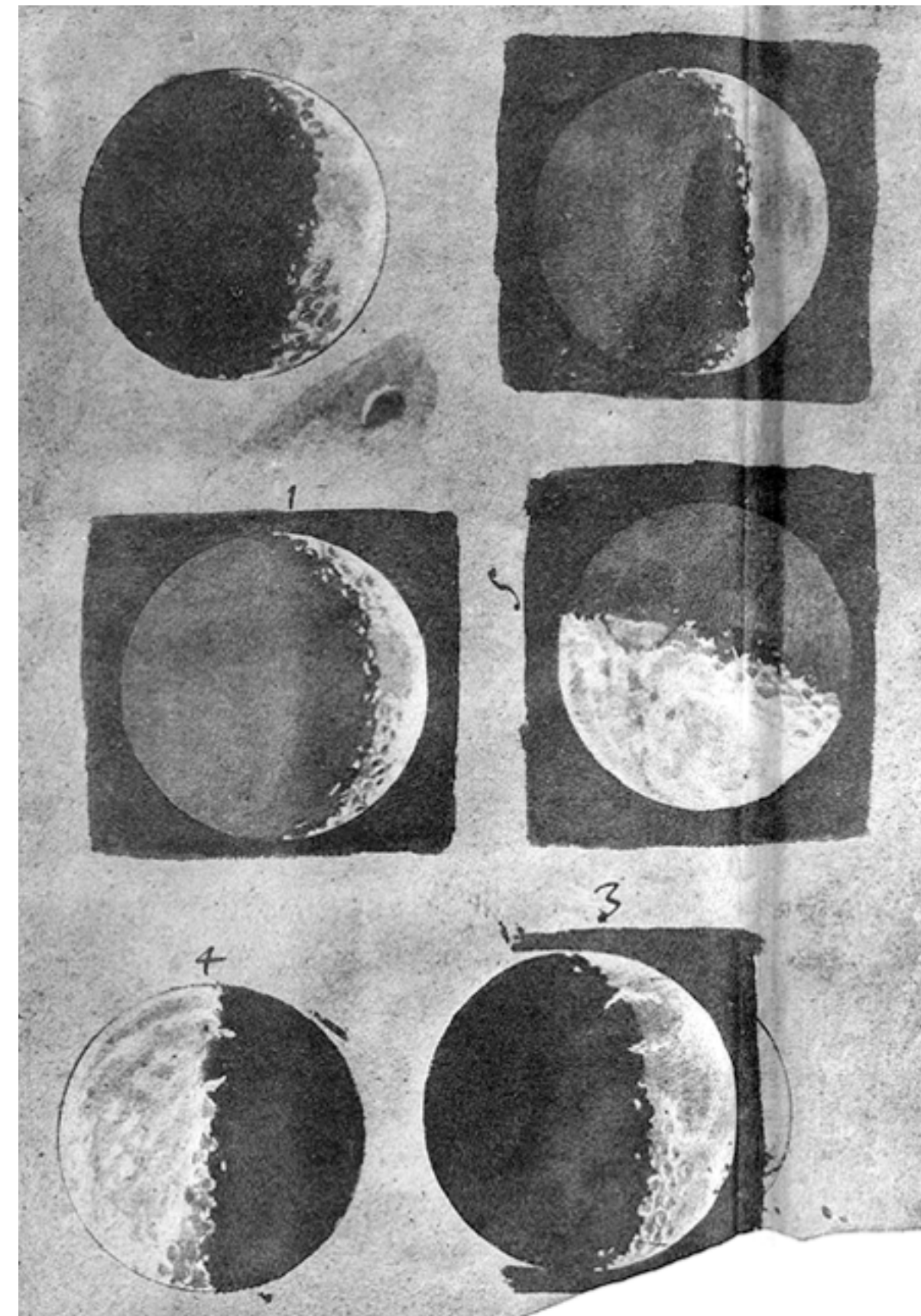
Anaximander's Map of the World

Anaximander of Miletus, c. 550 BC

Record



Leonardo Da Vinci, ca. 1500



Galileo Galilei, 1616

Donald Norman

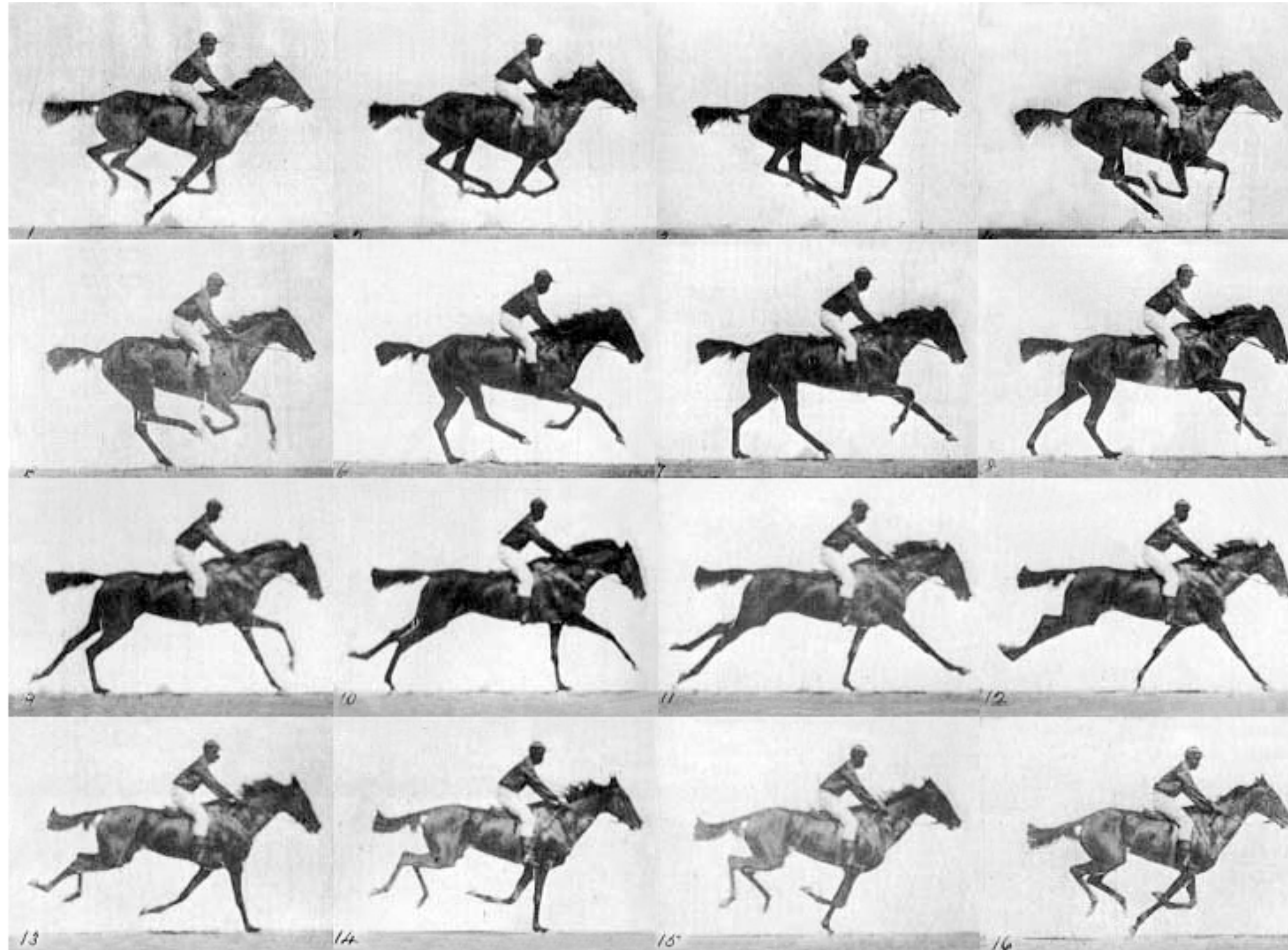


William Curtis (1746-1799)

The History of Visual Communication

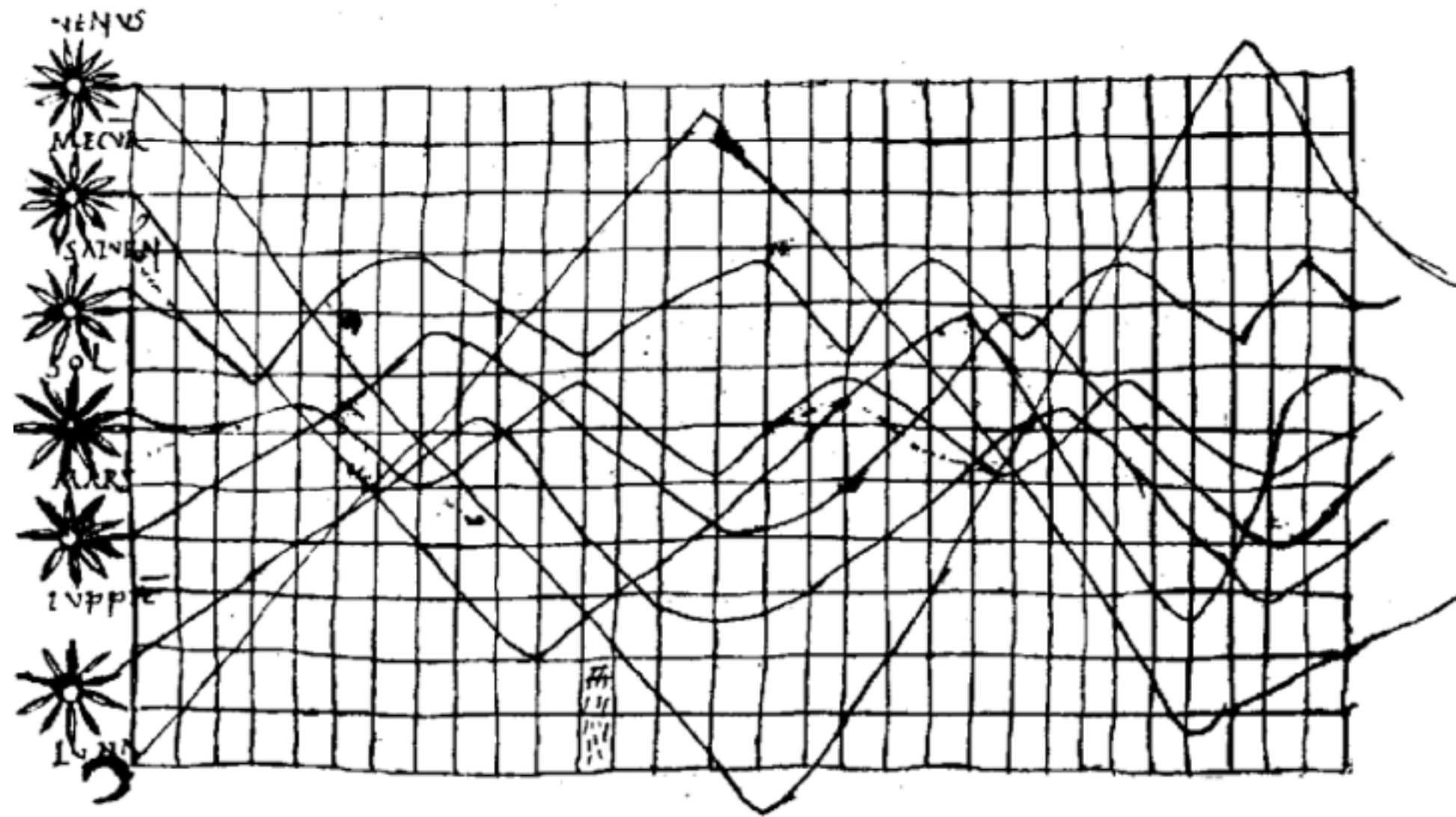
The Galileo Project, Rice University

Record

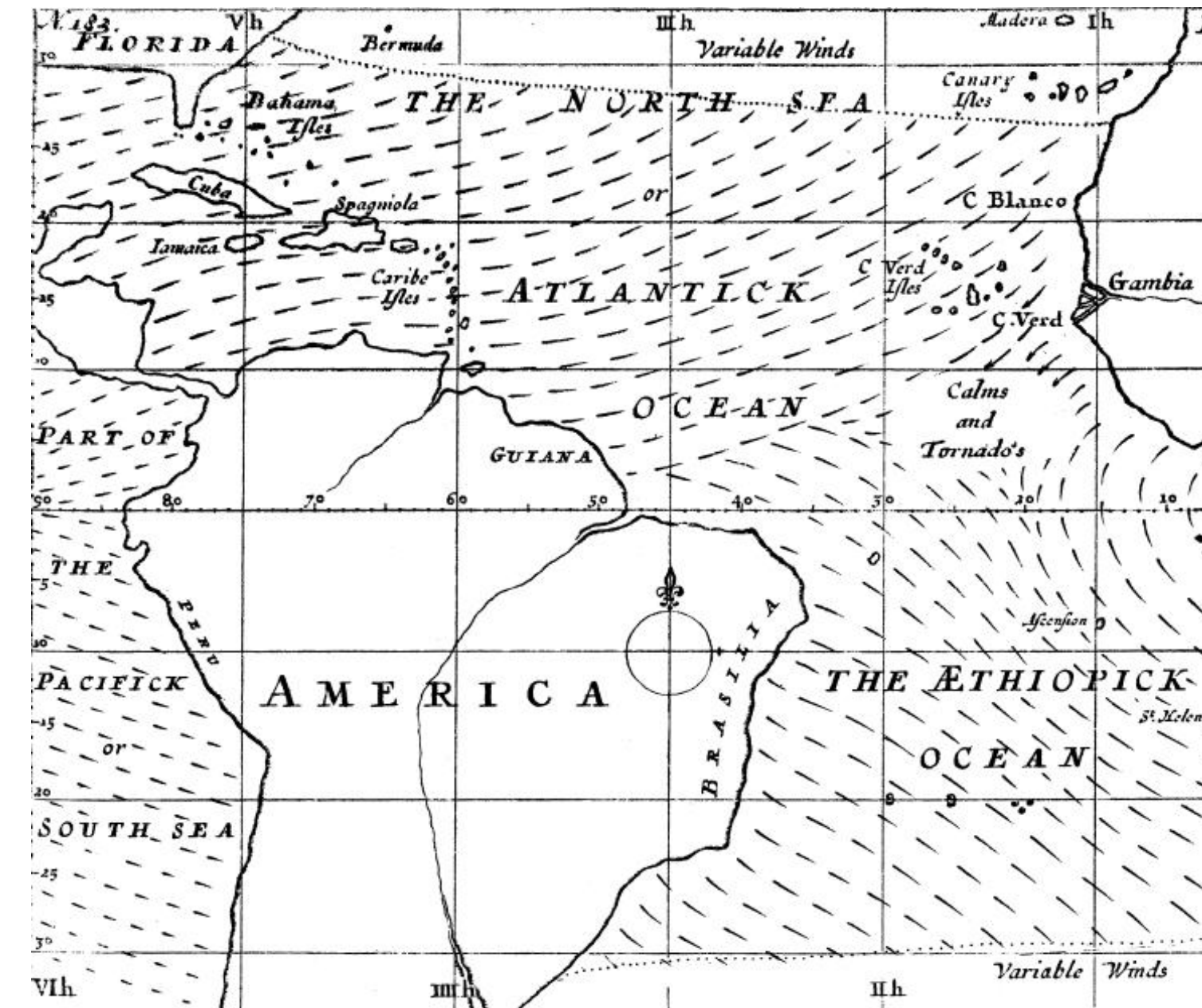


E. J. Muybridge, 1878

Analyze



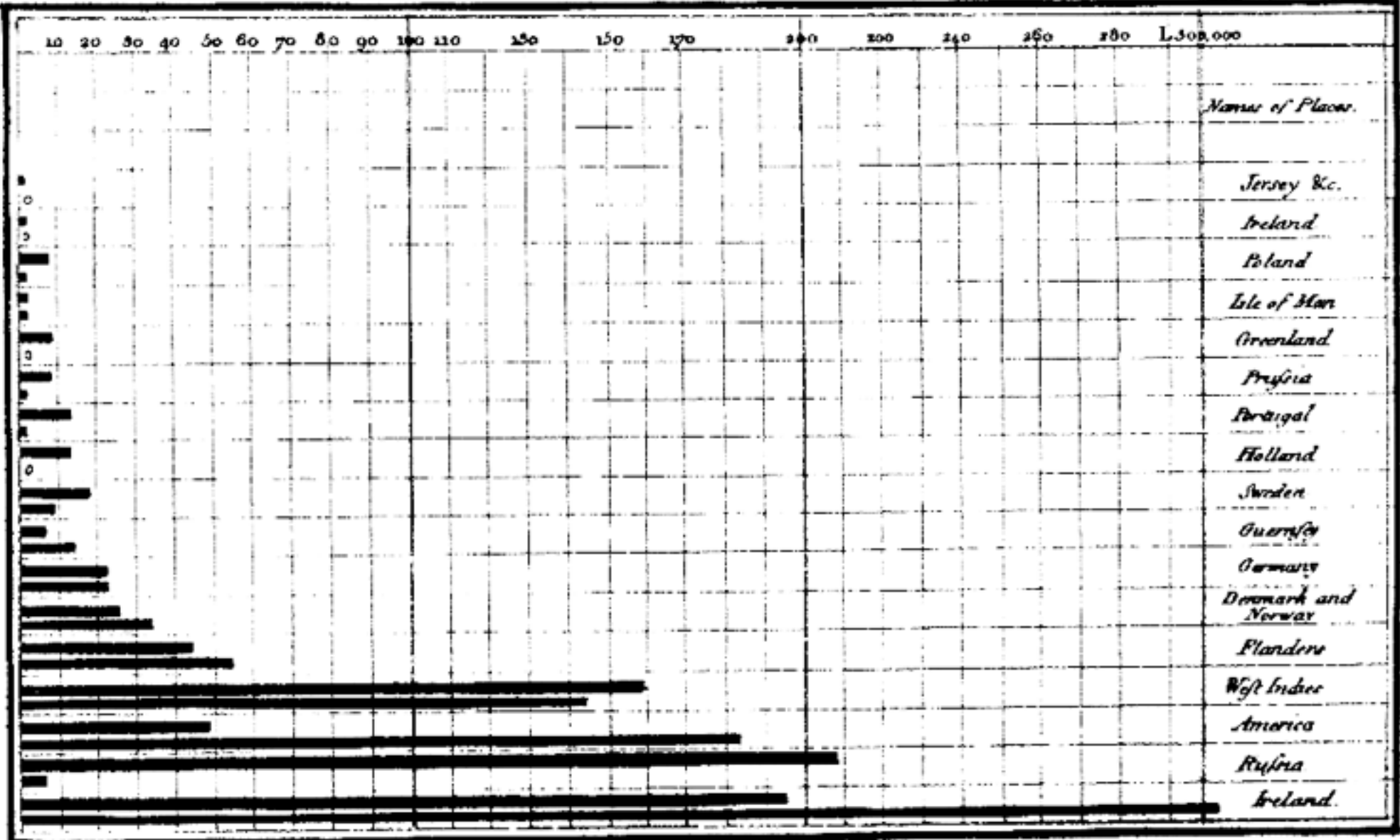
Planetary Movement Diagram, c. 950



Halley's Wind Map, 1686

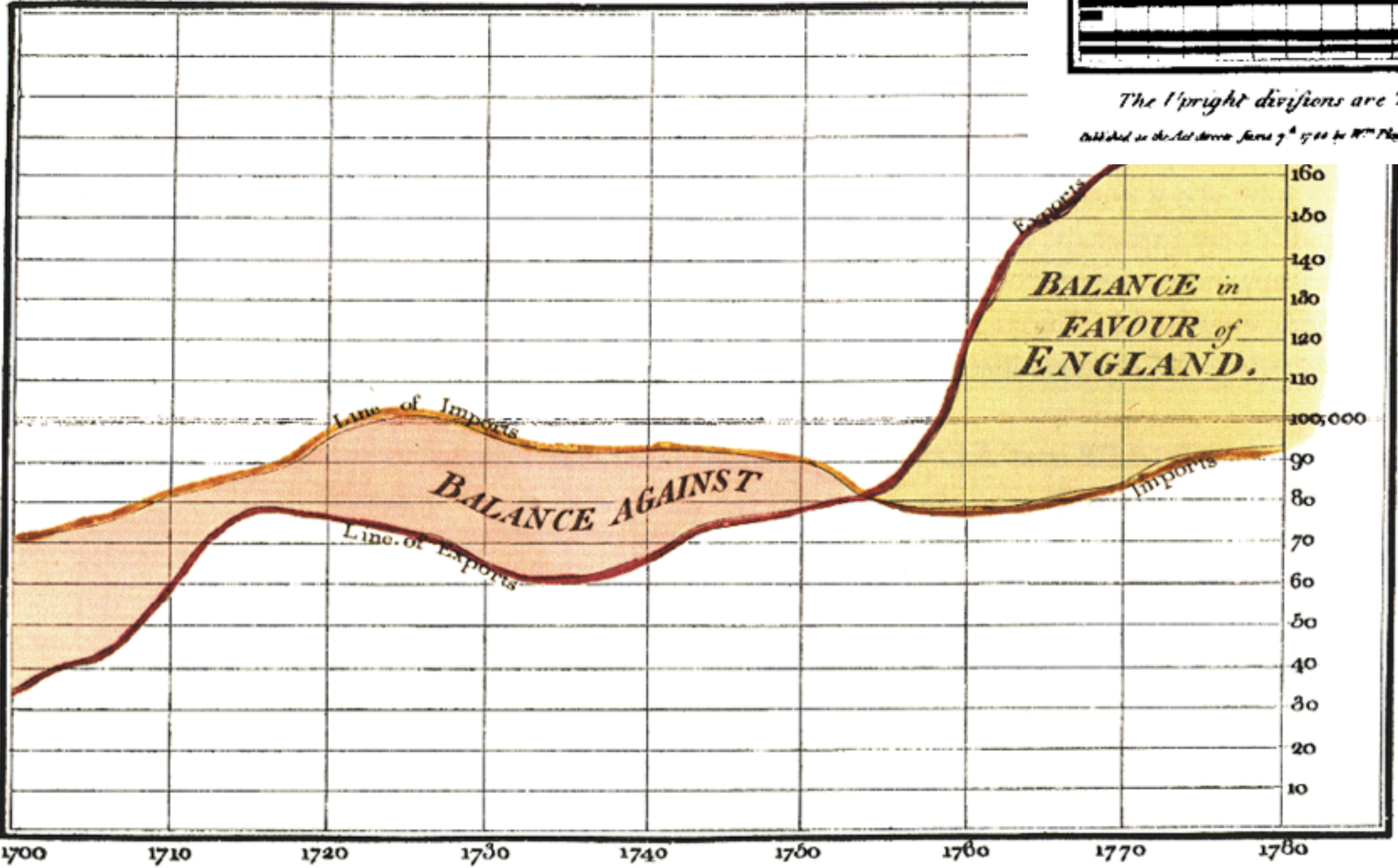
Analyze

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

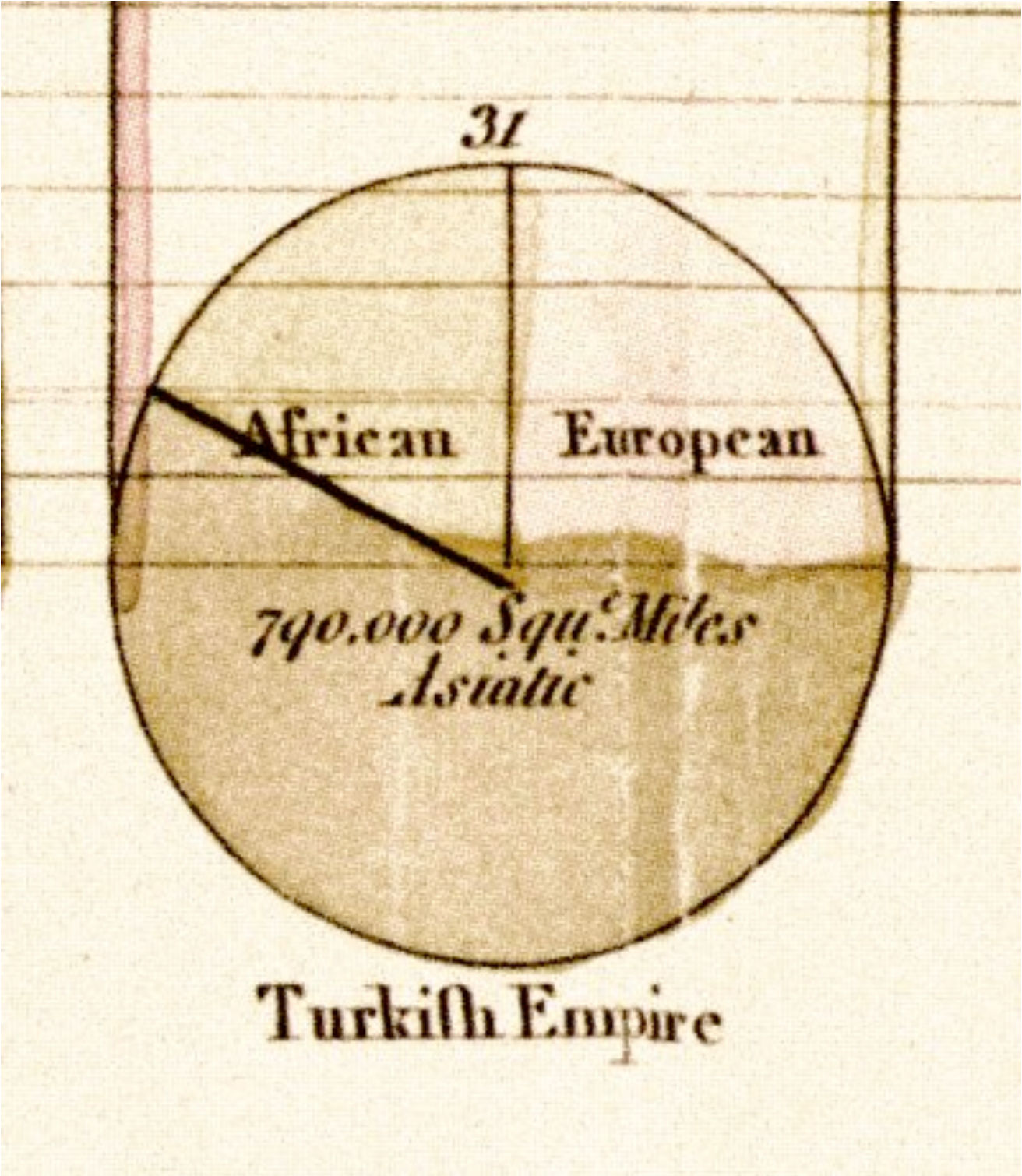


The upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed lines Imports.
 Published in the Advertiser June 7th 1781 by W. Playfair
 Made up by J. G. Bond, London.

Exports and Imports to and from DENMARK & NORW



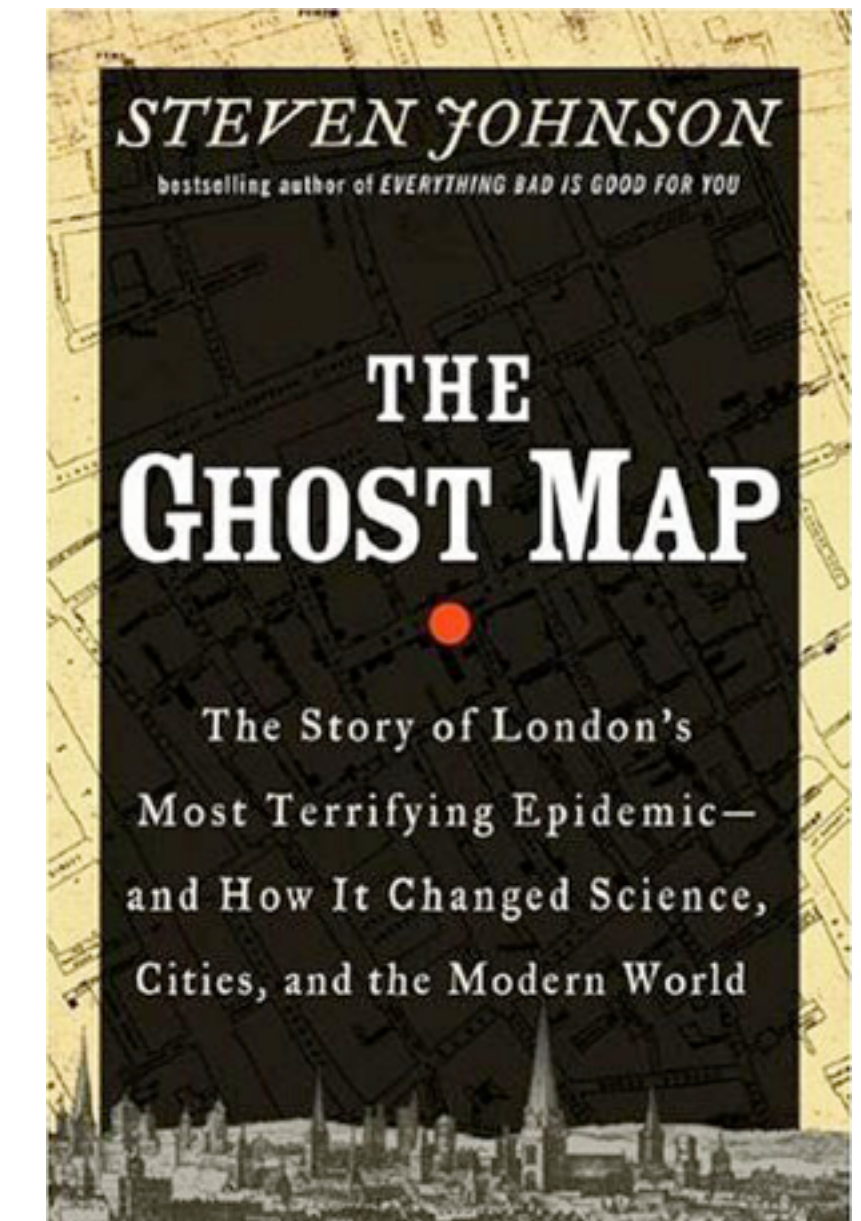
W. Playfair, 1786



Find Patterns

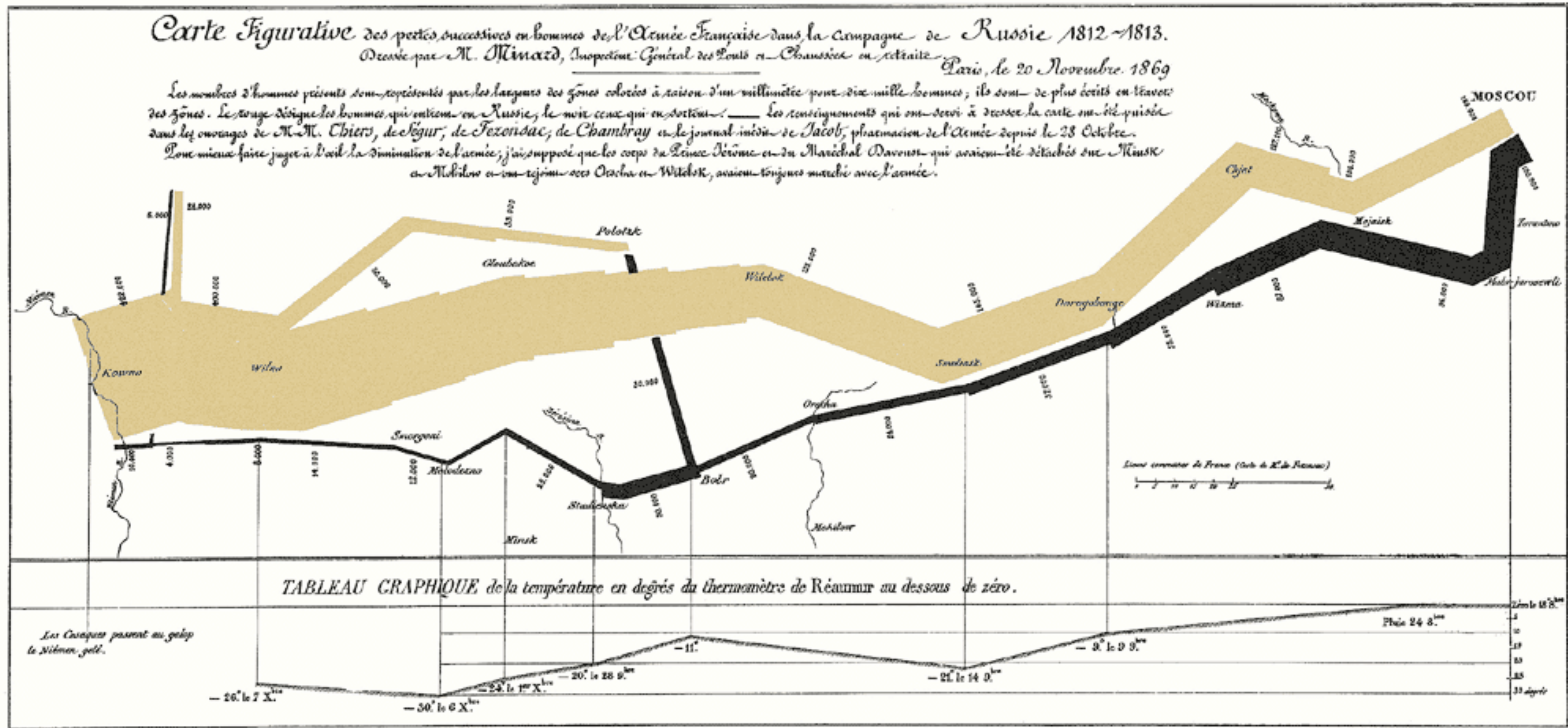


John Snow, 1854

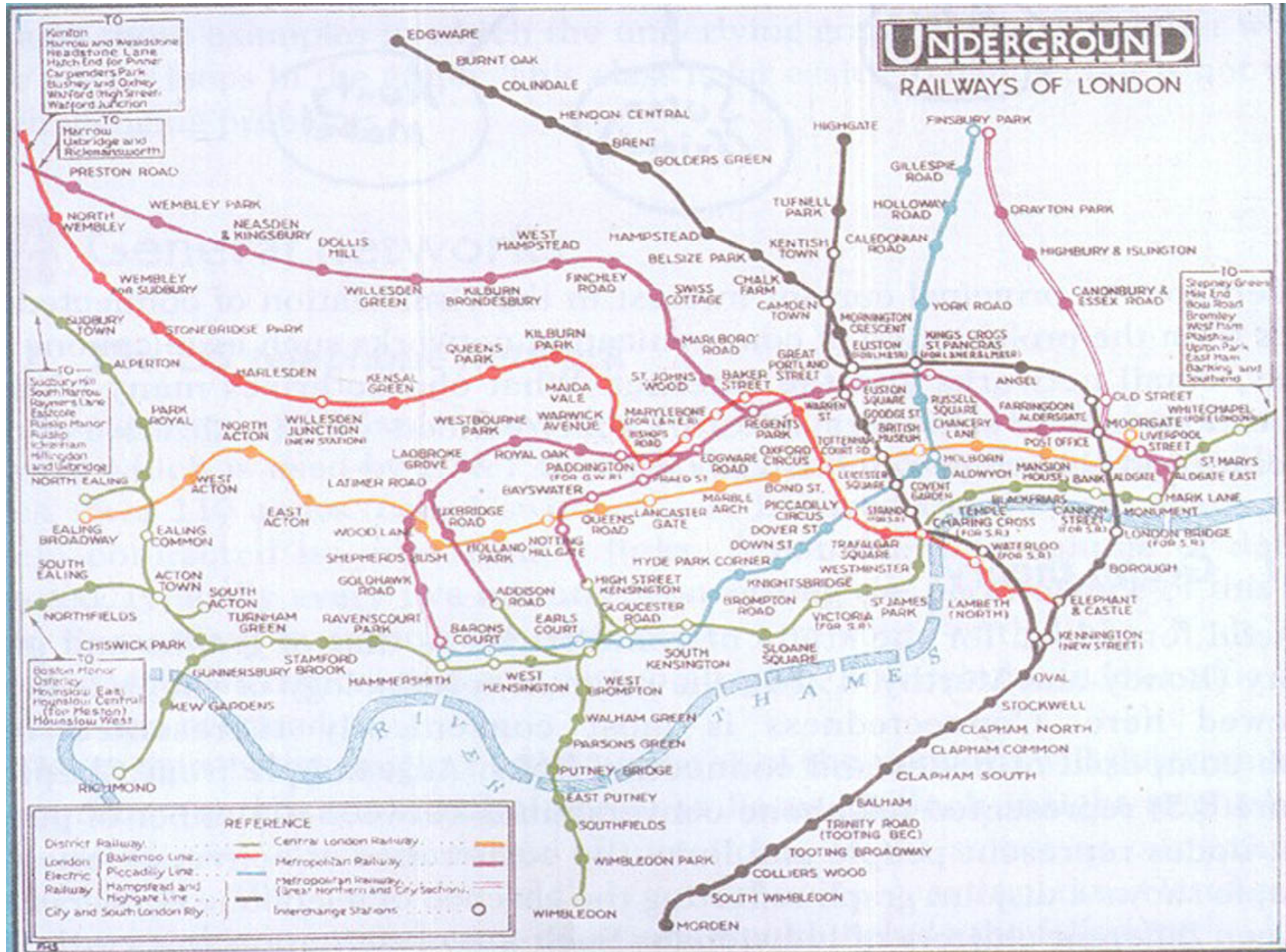


E. Tufte, Visual Explanations, 1997

Communicate



Communicate



London Subway Map, 1927


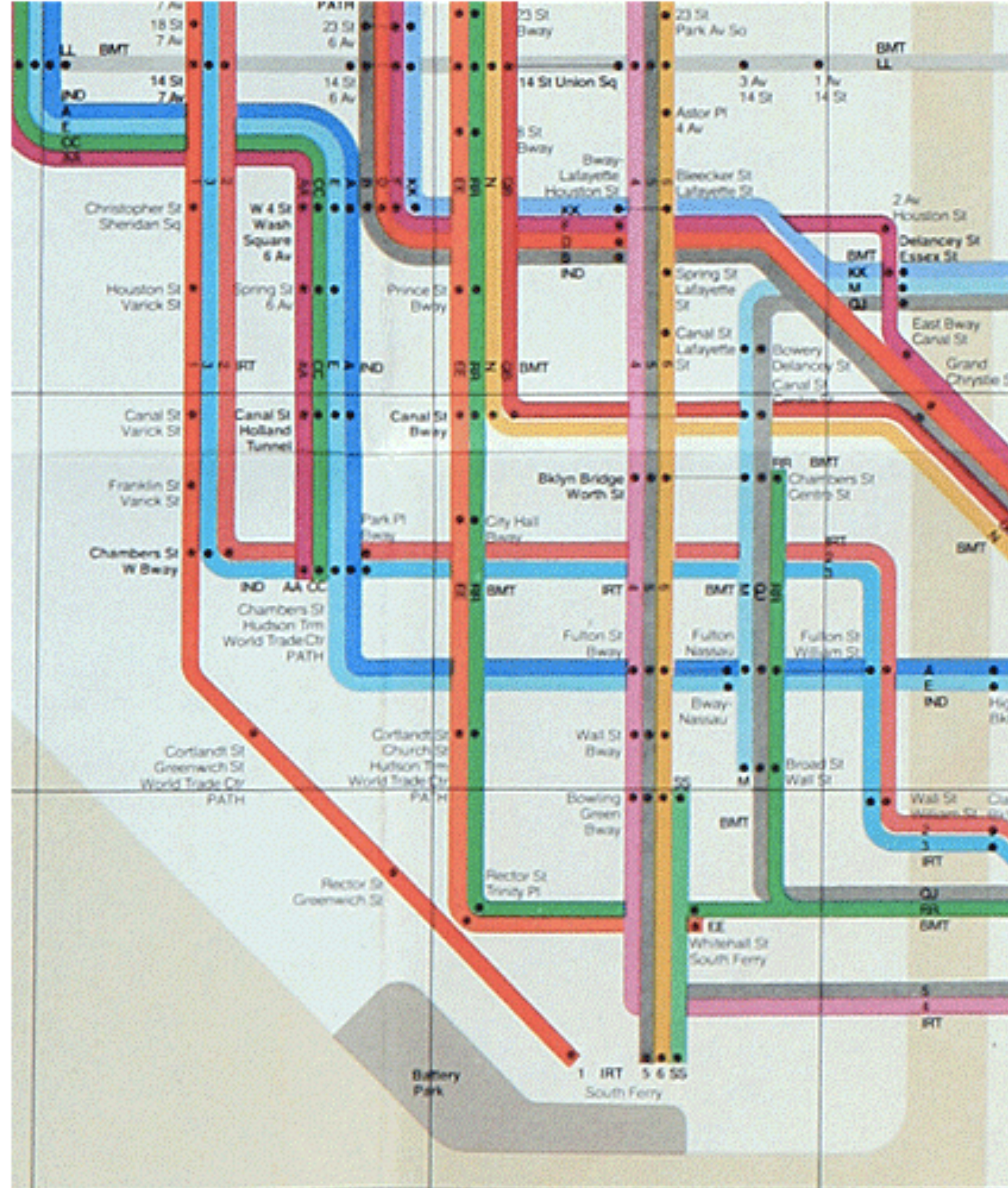
Communicate

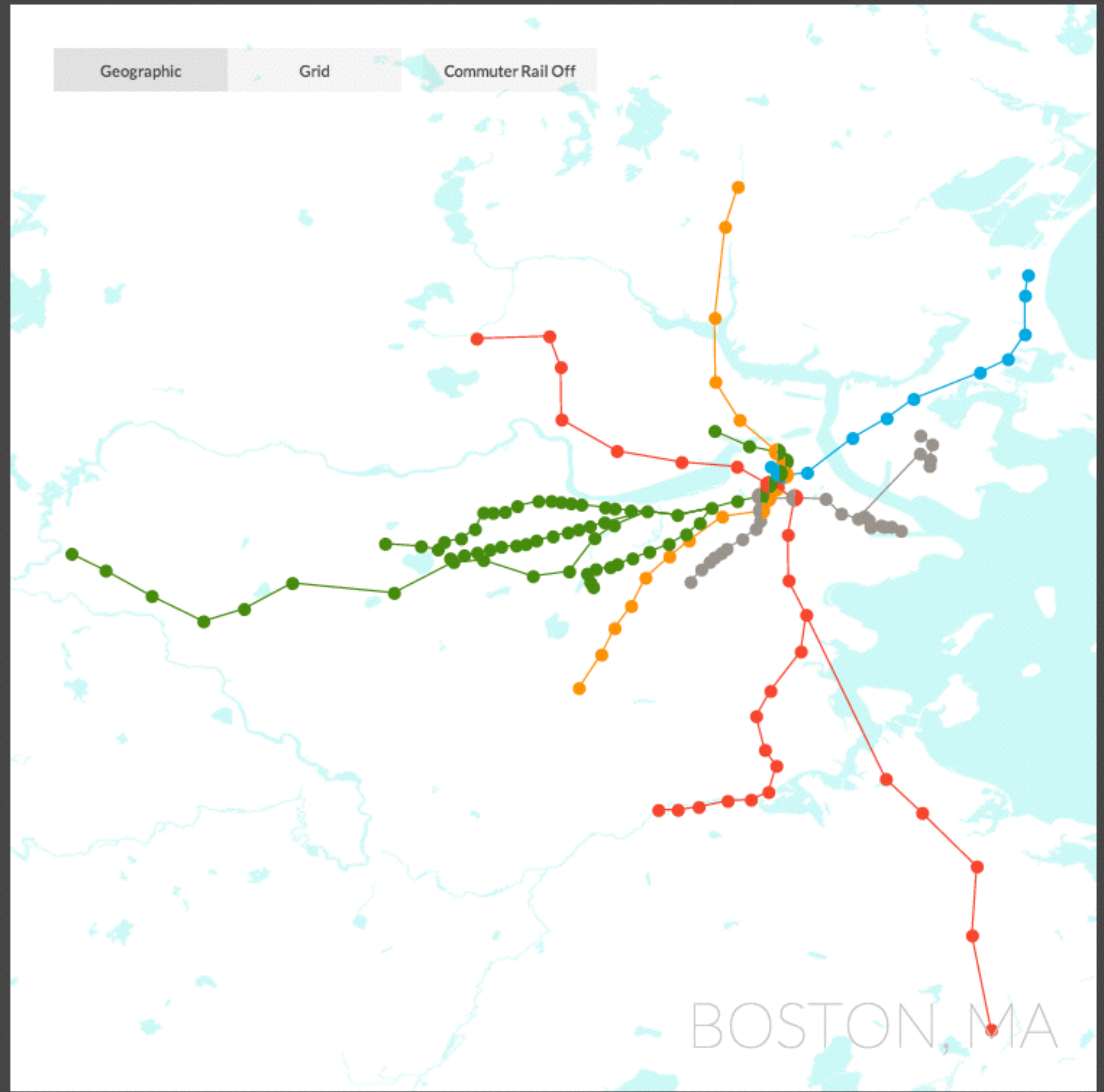


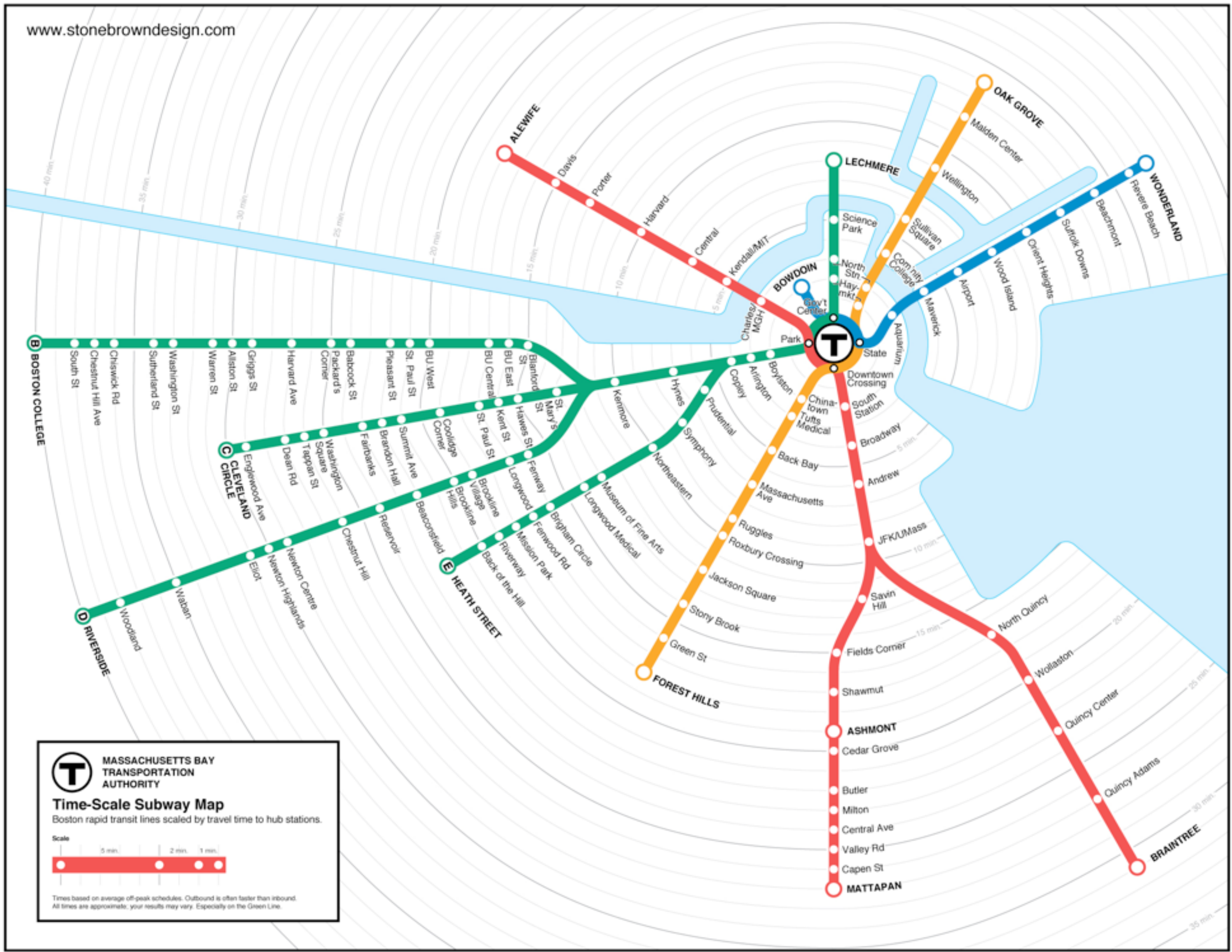
Harry Beck, 1933

An Overhaul of an Underground Icon

Next month, the Metropolitan Transportation Authority will unveil a resized, recolored and simplified edition of the well-known map, its first overhaul in more than a decade. [Related Article »](#)

FULL MAP	The New Map	1972: Vignelli's Classic
BRONX		
BROOKLYN		
MANHATTAN		
QUEENS		





T MASSACHUSETTS BAY
TRANSPORTATION
AUTHORITY

Time-Scale Subway Map
Boston rapid transit lines scaled by travel time to hub stations.

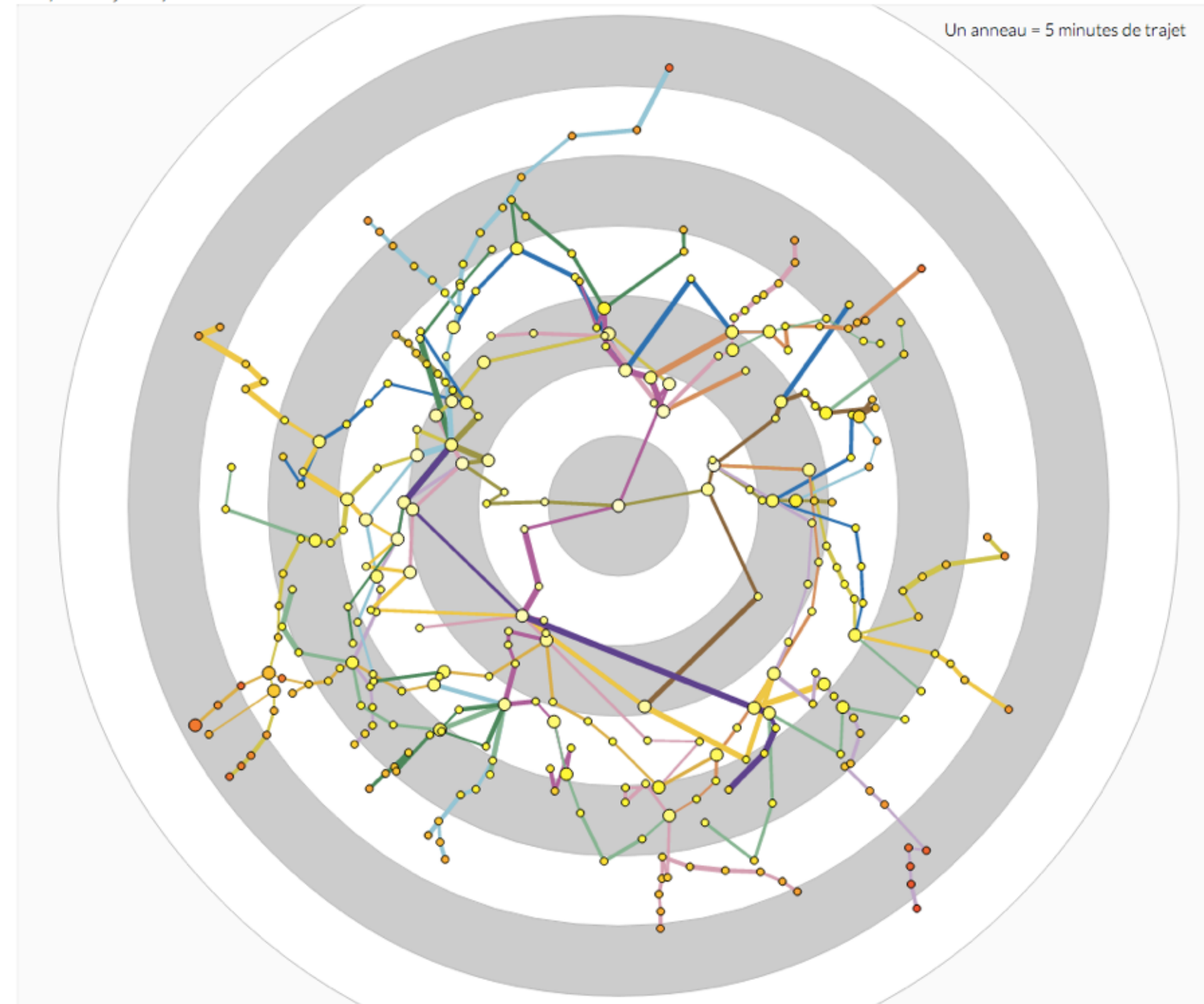
Scale

5 min. 2 min. 1 min.

Times based on average off-peak schedules. Outbound is often faster than inbound.
All times are approximate; your results may vary. Especially on the Green Line.

Réaumur — Sébastopol

Temps de trajet moyen: 18 minutes 26 secondes



Utiliser les positions exactes des stations

Améliorez le plan!

Vous connaissez bien la station Réaumur — Sébastopol ? Cliquez dans les zones grises pour améliorer les estimations.

Quelle distance entre la sortie et les quais ?



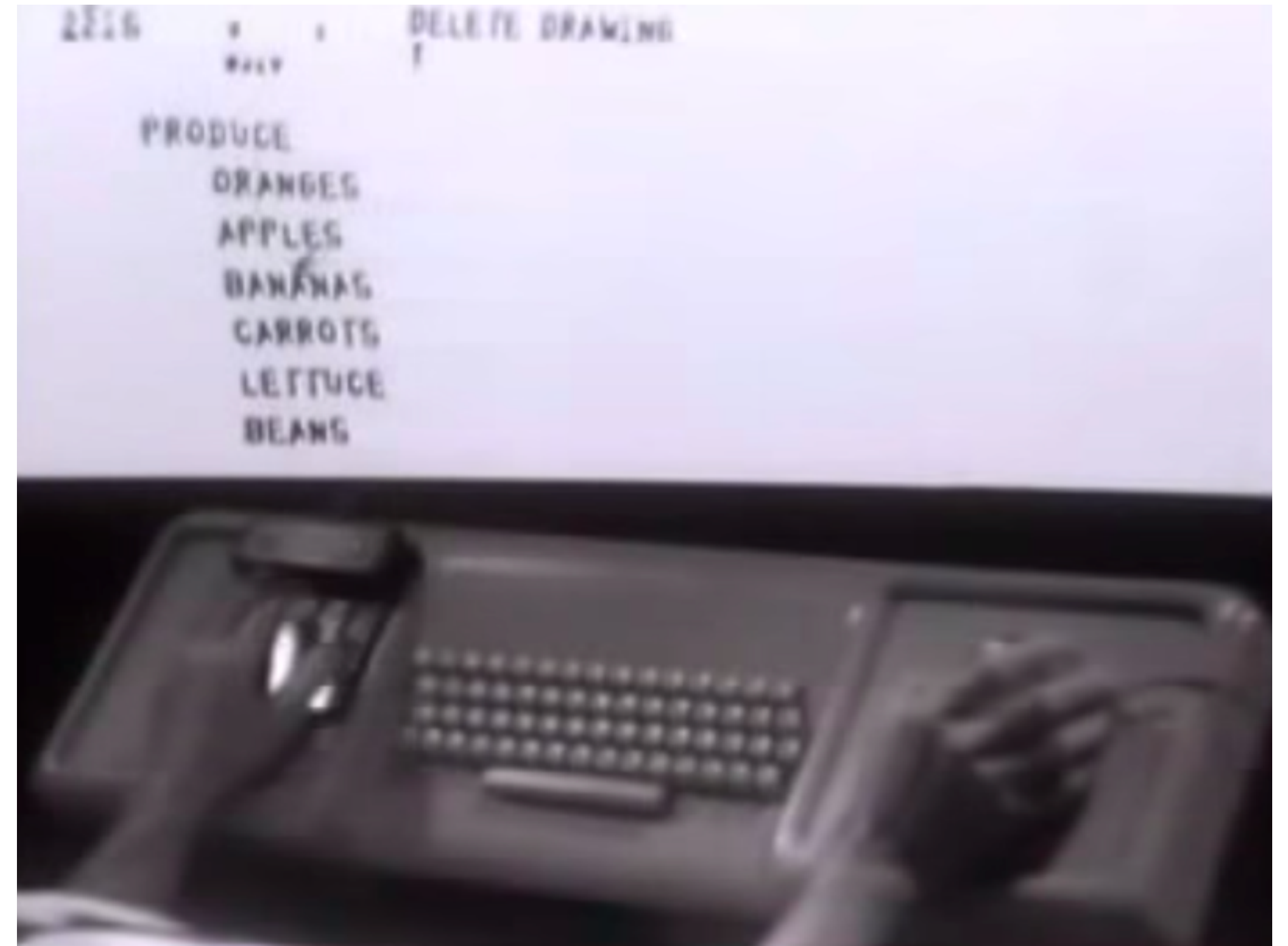
Quelle est la longueur des correspondances ?



Interact

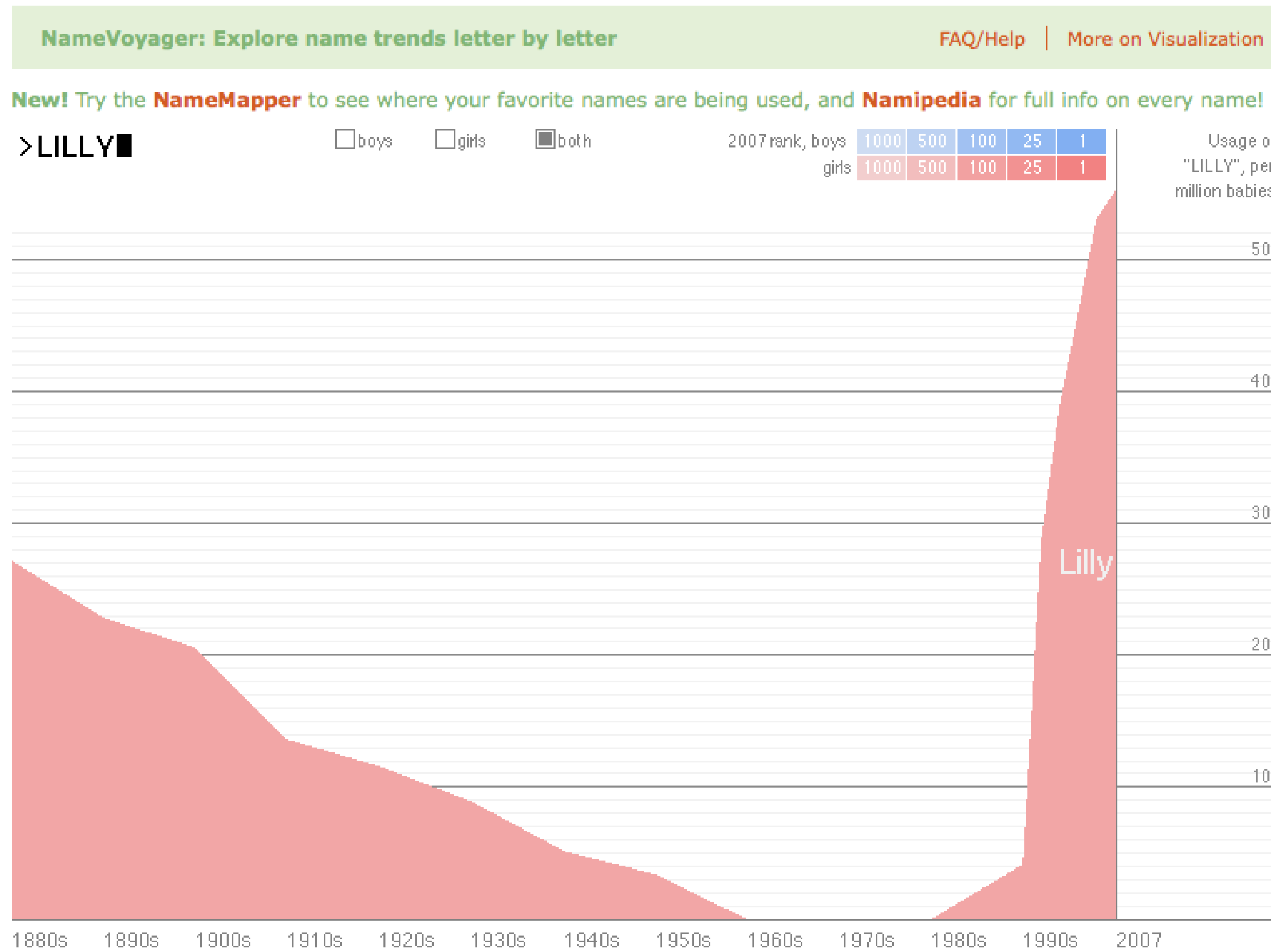


Ivan Sutherland, Sketchpad, 1963



Doug Engelbart, 1968

Analyze



M. Wattenberg, 2005

Communicate



Hans Rosling, TED 2006

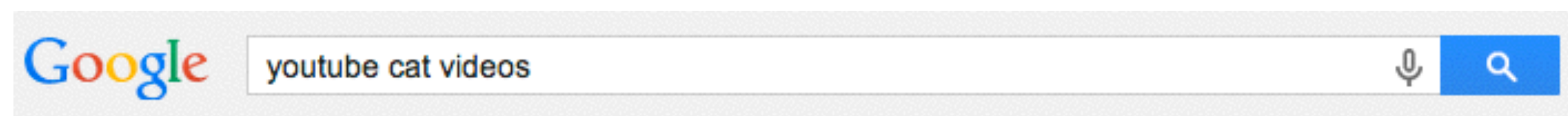
Big Data

2010: 1,200 exabytes, largely unstructured

Google stores ~10 exabytes (2013)

Hard disk industry ships ~8 exabytes/year

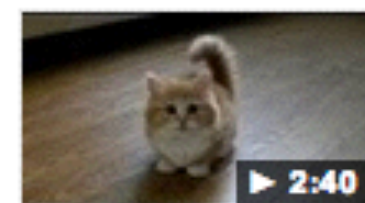
15 Exabytes in Punch Cards:
4.5 km over New England



Web Videos Shopping Images News More Search tools

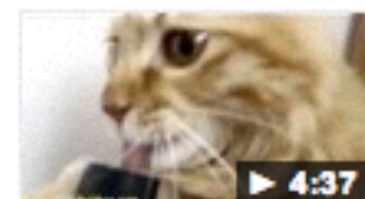
About 593,000,000 results (0.44 seconds)

[TOP 10 BEST CAT VIDEOS OF ALL TIME! - YouTube](#)



www.youtube.com/watch?v=... YouTube
Sep 6, 2012 - Uploaded by WatchTheDaily
We've scoured the internet and found the cutest and funniest cat videos of all time. Any we missed? Let us ...

[The World's Most Funny Cat Videos 2013 - YouTube](#)



www.youtube.com/watch?v=Ak... YouTube
Aug 15, 2013 - Uploaded by papiaanifails
FREE Cat Tree Construction Manual : <http://tinyurl.com/ovp252j>
Funniest Dog Videos <https://www.youtube.com> ...

In one second on the Internet there are...



“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.**”

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

Limits of Cognition



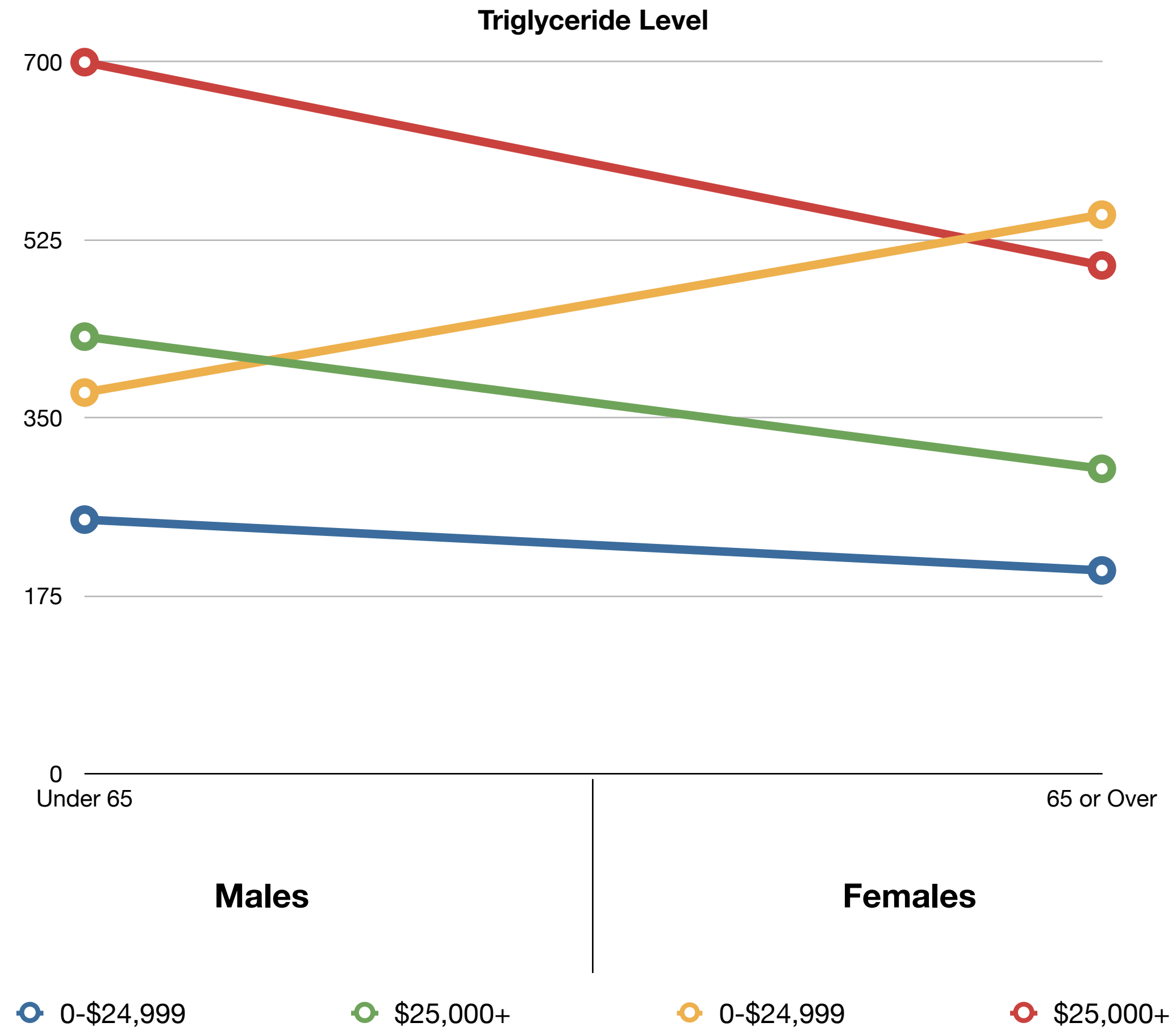
Daniel J. Simons and Daniel T. Levin, Failure to detect changes to people during a real world interaction, 1998

Limits of Cognition

Which gender or income level group shows different effects of age on cholesterol levels?

Income Group	Males		Females	
	Under 65	65 or Over	Under 65	65 or Over
0-\$24,999	250	200	375	550
\$25,000+	430	300	700	500

Visual Queries



“It is things that make us smart”

Donald A. Norman



15 a

E In terra pax hominibus bonae voluntatis. **Quidiam** re...

Domine deus rex caelestis deus pater omnipotens. **Virgine** re... **Sp** 7 al...

Domine deus agnus dei

maie uirginis marie. **Dea** mundi. **De** uili peccata mundi

30 The first Book of

describe the circle AEH ; and let DA be produced to the point G in the circumference thereof. Then $AG = CB$.

For $DGF = DE$, and $DAG = DC$. Wherefore $AG = CE = BC = AG$. Which was to be done.

The joining of the point A within or without the line BC varies the cases; but the construction, and the demonstration, are every where alike.

Q.E.D.

The line AG might be taken with a pair of compasses; but so doing answers to no postulate, as Ptolemy well intimates.

PROP. III.

Two right lines, A and B , being given, from the greater B cut away the right line BE equal to the lesser A .

At the point B draw the right line $BD = A$. The circle described from the center B in the distance of BD shall cut off $BE = BD = A = BE$. Which was to be done.

PROP. IV.

If two triangles BAC , EDF , have two sides of the one BA , AC equal to two sides of the other ED , DF ; and the included angles B , D equal, then the third sides BC , EF shall be equal, and the other angles shall be equal, each to each.

EUCLIDE'S ELEMENTS.

11

and $AC = DF$ and have the angle B equal to the angle D contained under the equal sides BA , ED ; then shall the angle C be equal to the angle F ; and the side BC shall be equal to the side EF ; and the remaining angles E , F , each to each, and the remaining sides AC , DF shall be equal, each to each.

If the point D be applied to the point A , and the right line DE fall upon the right line AB , the point E shall fall upon B , because $DE = AB$; also the right line DF shall fall upon AC , because the angle $A = D$; moreover the point F shall fall upon the point C , because $AC = DF$. Therefore the right lines EF , BC shall agree, because they have the same terms, & so consequently are equal. Wherefore the triangles BAC , EDF , and the angles E , F , as also the angles C , F , do agree; and are equal, which was to be demonstrated.

PROP. V.

The angles BAC , ACB at the base of an isosceles triangle ABC , are equal one to the other; and if the equal sides AB , AC be produced, the angles CBD , BCE , under the base, shall be equal one to the other.

Take $AE = AD$, and join CE , BD , and BE .

Because, in the triangles ACD , AEB are $AC = AE$ and $AD = AB$, and the angle A common to them both; therefore the angle $ACE = ABE$, and the angle $AEB = ADC$; and the base $CE = BD$; also $EC = DB$. Therefore in the triangles BCE , BCD shall be the angle $ECB = DCB$, which was to be done. Also therefore the angle $EBC = DCE$; but the angle $BCE = ACD$; therefore the angle $BAC = ACD$, which was to be done.

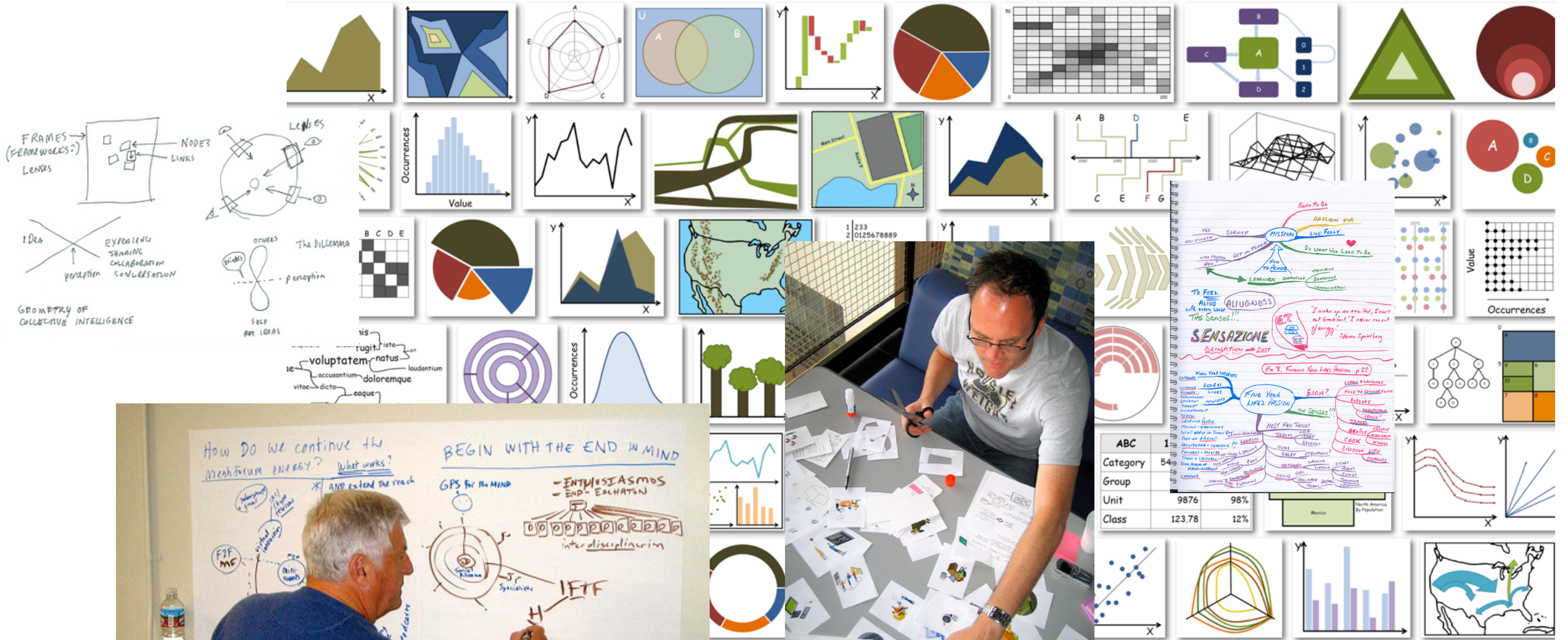
233

Incipit festiuitas sanctorum per dñi ceterum. **I**n isto sancti Saturnini martyris. **Oratio.**

Eus qui non deat Saturnini martyris cui concedis natalicio presertim non tribue meritis adiuuari. **P**ro domini nostri. **L. i.**

Rome nate faveri Saturnini martyris et senio; et sicut uiaconi sub maximiano. Et quo primo iter alios seruas dei caritati sunt ad fodendam harenas ad faciendam thermas vicentianam. Quorum unum sibi presertim: cum interrogasset idem maximianus quomodo uocaretur, respondit: Ego peccator; sicut seruus seruorum domini nostri iesu christi. Et ut eius post alia diceret maximianus; aut sacrificia tuo herculi aut carnis tuae cremato; respondit: Ego quidem semper hoc optaui; ueritate si meriti seruo coenam desideras acceptam.

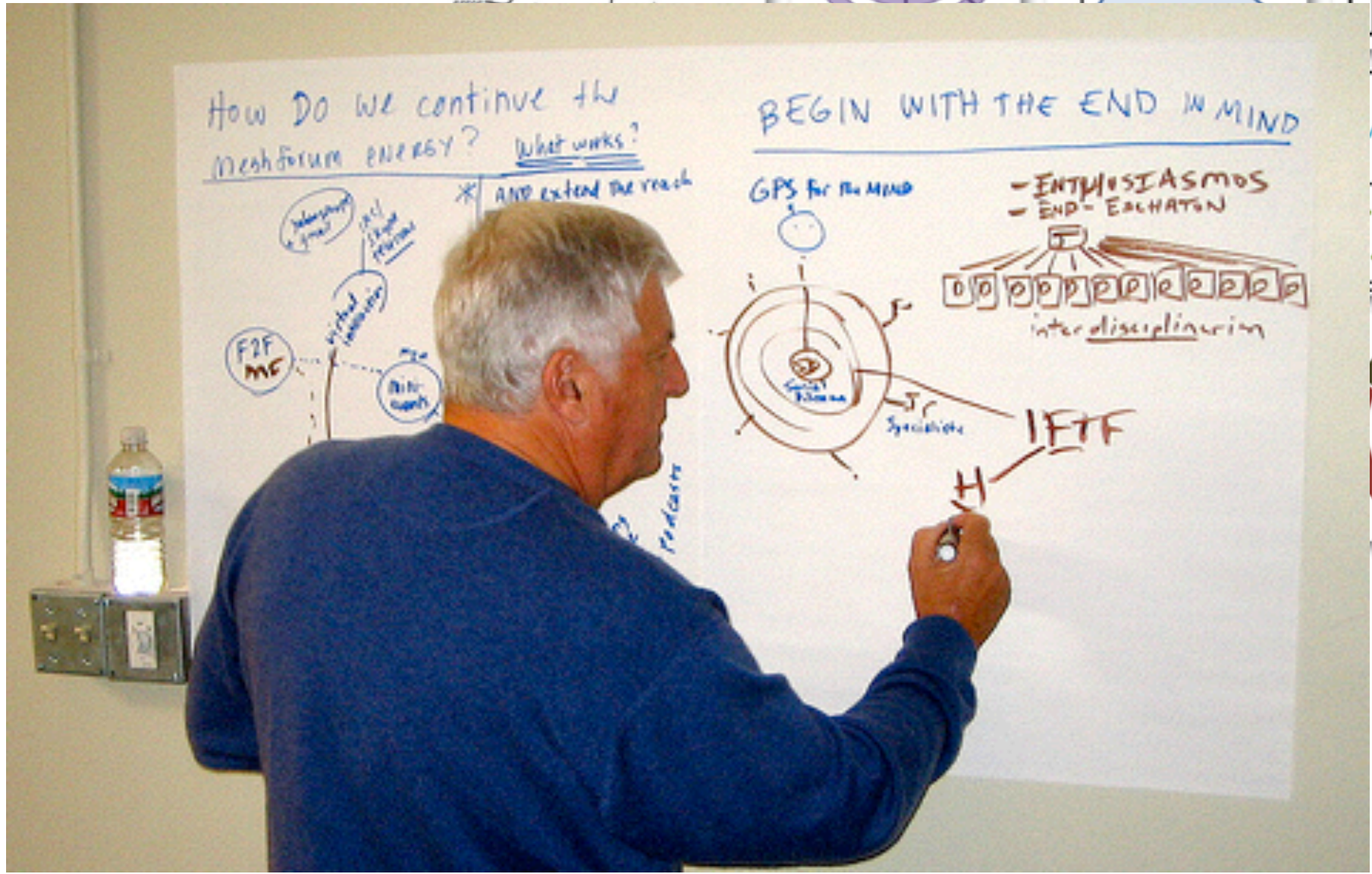
Maximus igitur laudato. **L. ceterum secunda.** **P**rofecto missas est in custodiam decem et



MISSION
 PASSIONE PER
 LUG FOLLY
 DO WHAT YOU LOVE TO DO
 TO FEEL ALIGNED WITH YOUR LIFE THE SENSES!!!
 SENSAZIONE
 SENSATION WITH ZEST
 "I wish you so excited, I can't eat breakfast. I never want of energy!" - Stearn Spielberg

FINO A CHE PUNTO?
 MOST FOR THINGS
 COOK
 DISCOVER LIFE

ABC
 Category 54
 Group
 Unit 9876 98%
 Class 123,78 12%



Visual Thinking Collection, Dave Grey

Idea Maps, by Jamie Nast

Visualization

“Visualization is really about external cognition, that is, how resources outside the mind can be used to boost the cognitive capabilities of the mind.”



Stuart Card

Who is CS 171?

Alexander Lex

Lecturer, Postdoctoral Fellow

PhD in Computer Science, Graz University of Technology

Visual Computing Group, PI: Prof. Hanspeter Pfister



Visual Computing Group

Prof. Dr. Hanspeter Pfister

Dr. Ray Jones

Dr. Johanna Beyer

Dr. Hendrik Stroblet

Dr. James Tompkin

Dr. Verena Kaynig

Dr. Seymour K.-B.

Dr. Dequin Sun

Dr. Michelle Borkin

Dr. Adi Suissa Peleg

Gaurav Bharaj

Daniel Haehn

Nam Wook Kim



<http://vcg.seas.harvard.edu/>

HARVARD UNIVERSITY GVI | HARVARD.EDU



Visual Computing Group

Professor Hanspeter Pfister



[News](#) [Projects](#) [Publications](#) [Presentations](#) [People](#) [Code and Data](#) [Classes](#) [Contact](#)



VISUAL COMPUTING

Our research in visual computing lies at the intersection of visualization, computer graphics, and computer vision. It spans a wide range of topics, including bio-medical visualization, image and video analysis, 3D fabrication, and data

OUR RESEARCH

Our goal is to combine interactive computer systems with the perceptual and cognitive power of human observers to solve practical problems in science and engineering. We are providing visual analysis tools and methods to

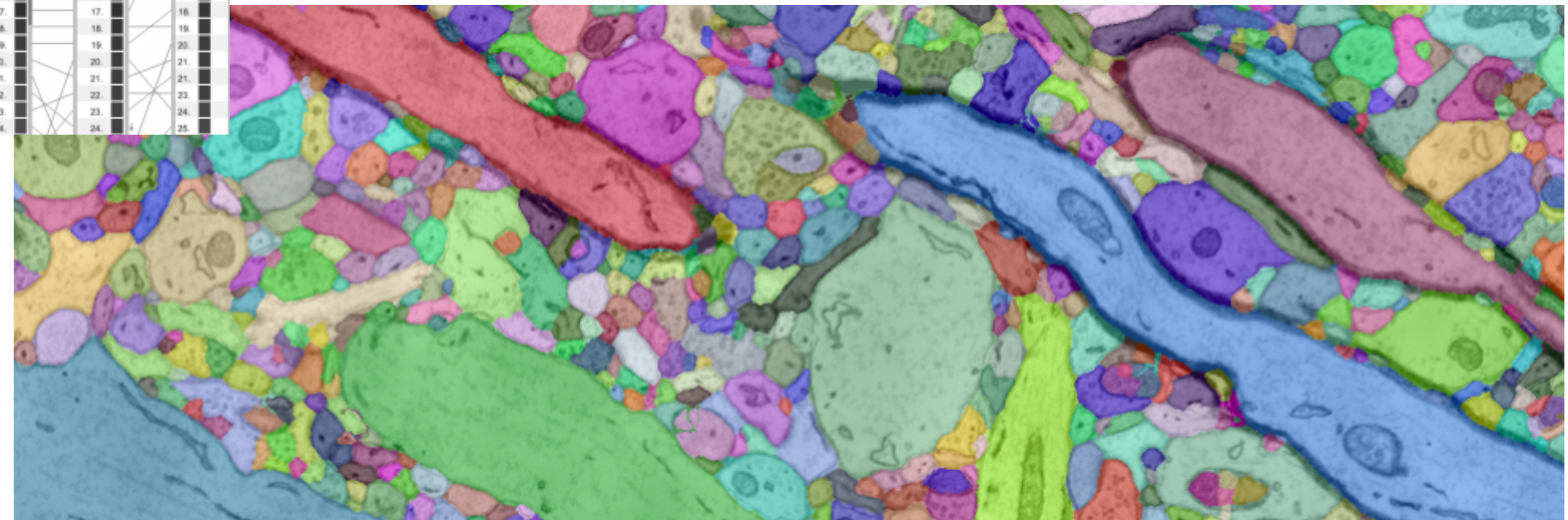
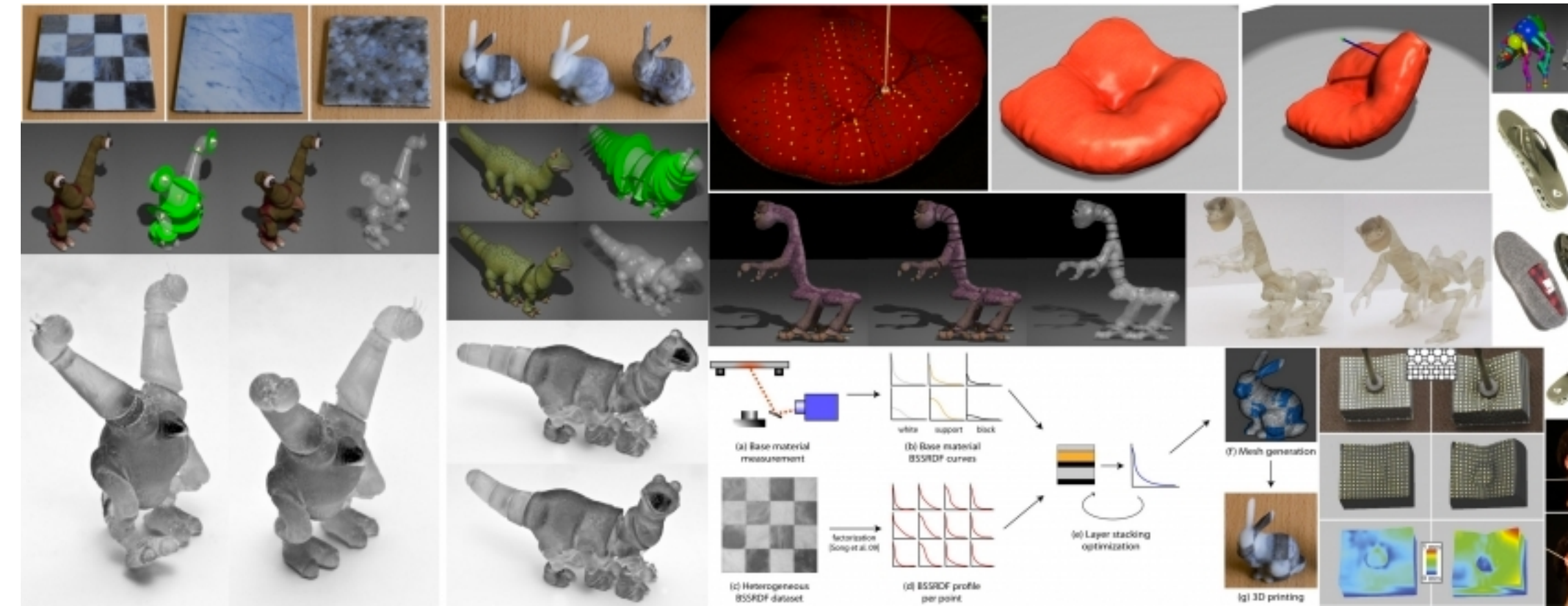
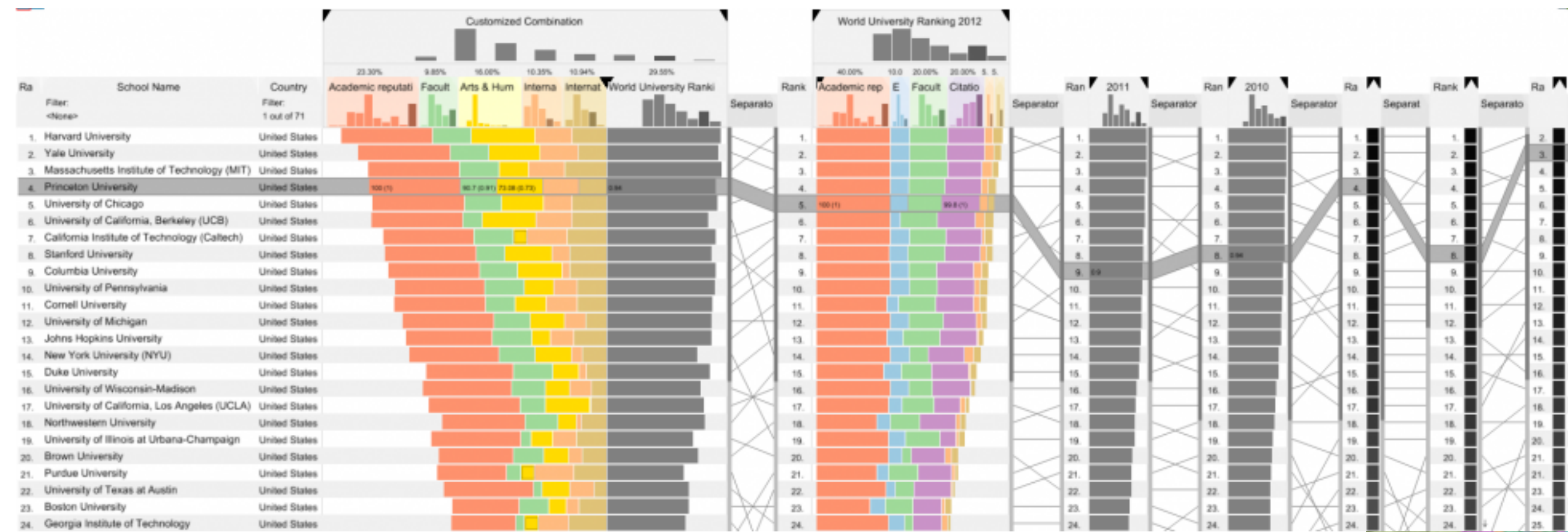
RECENT PUBLICATIONS

[Facial Performance Enhancement Using Dynamic Shape Space Analysis](#)

[Segmenting Planar Superpixel Adjacency Graphs w.r.t. Non-planar Superpixel Affinity Graphs](#)

Our Research

Visualization, Graphics, Vision



CS171 Staff

Dr. Johanna Beyer (Head TF) - Postdoctoral Fellow, Visual Computing Group

David Chouinard - Research Associate, Disney Research

Dr. Hendrik Strobelt - Postdoctoral Fellow, Visual Computing Group

Dr. Romain Vuillemot - Data Visualization Fellow, Center for International Development

Luciano Arango - A.B. candidate in Computer Science

Samuel Gratzl - PhD Student, Johannes Kepler University

Mohammad K. Hadhrawi - Graduate Student Research Assistant, MIT Media Lab

Daniel Haehn - PhD Student, Visual Computing Group

Alain Ibrahim - Senior Web Developer

Benjy Levin - Computer Science Concentrator

Andrew Mauboussin - Computer Science Concentrator

Kevin Sun - A.B. candidate in Applied Math

Dr. James Tompkin - Postdoctoral Fellow, Visual Computing Group

Mimi Lai

About You

Structure & Goals

CS 171 Goals

Evaluate and **critique** visualization designs

Implement interactive data visualizations

Apply fundamental principles & techniques

Design visual data analysis solutions

Develop a substantial visualization project

No Device Policy

No Computers, Tablets, Phones in lecture hall

except when used for exercises

Switch off, mute, flight mode

Why?

It's better to take note by hand

Notifications are designed to grab your attention

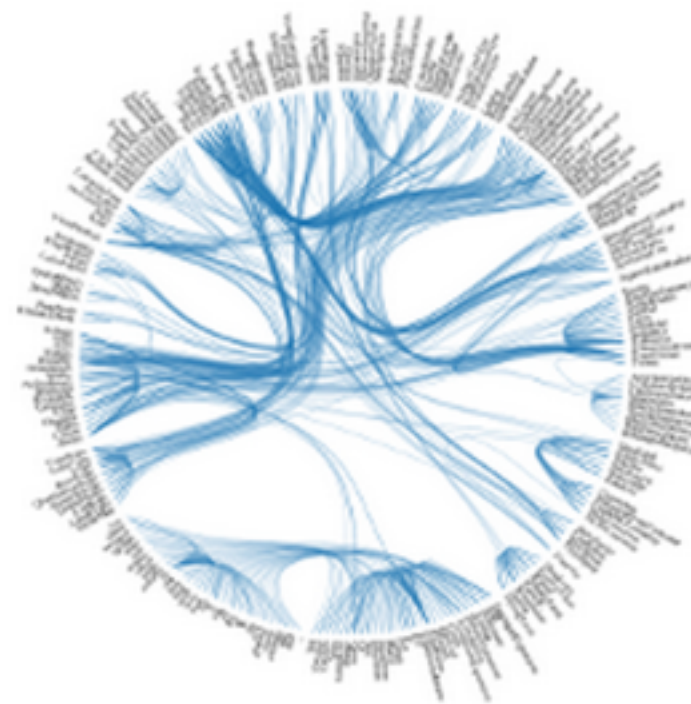
Information - <http://cs171.org>

CS 171 - Visualization



HARVARD
School of Engineering
and Applied Sciences

[Home](#) [Syllabus](#) [Schedule](#) [Homework](#) [Project](#) [Resources](#) [Fame](#)



Hierarchical edge bundling] | Wind map | How states have shifted

The amount and complexity of information produced in science, engineering, business, and everyday human activity is increasing at staggering rates. The goal of this course is to expose you to visual representation methods and techniques that increase the understanding of complex data. Good visualizations not only present a visual interpretation of data, but do so by improving comprehension, communication, and decision making.

In this course you will learn how the human visual system processes and perceives images, good design practices for visualization, tools for visualization of data from a variety of fields, and programming of interactive

Communicate

Office Hours
starting next week

Piazza

<https://piazza.com/harvard/cs171>

E-Mail

staff@cs171.org

alex@seas.harvard.edu

The screenshot shows a Piazza discussion for CS 171. The main thread is titled "Homework 0 online" and contains a note from the instructor: "Homework 0 online. See the hw0 subfolder in this repository: https://github.com/CS171/2015-cs171-homework". Below the note, there are several replies from students. One student, Pierre-Marie Meyitang, asks for help regarding a GitHub account issue. Another student, Alexander Lex, responds that it wasn't an issue last year. Other students, Matt Keane and Jack Golding, also provide information about their GitHub account issues. The discussion is moderated by the instructor, who has resolved the thread.

Course Components

Lecture
Reading
Discussion

Theory

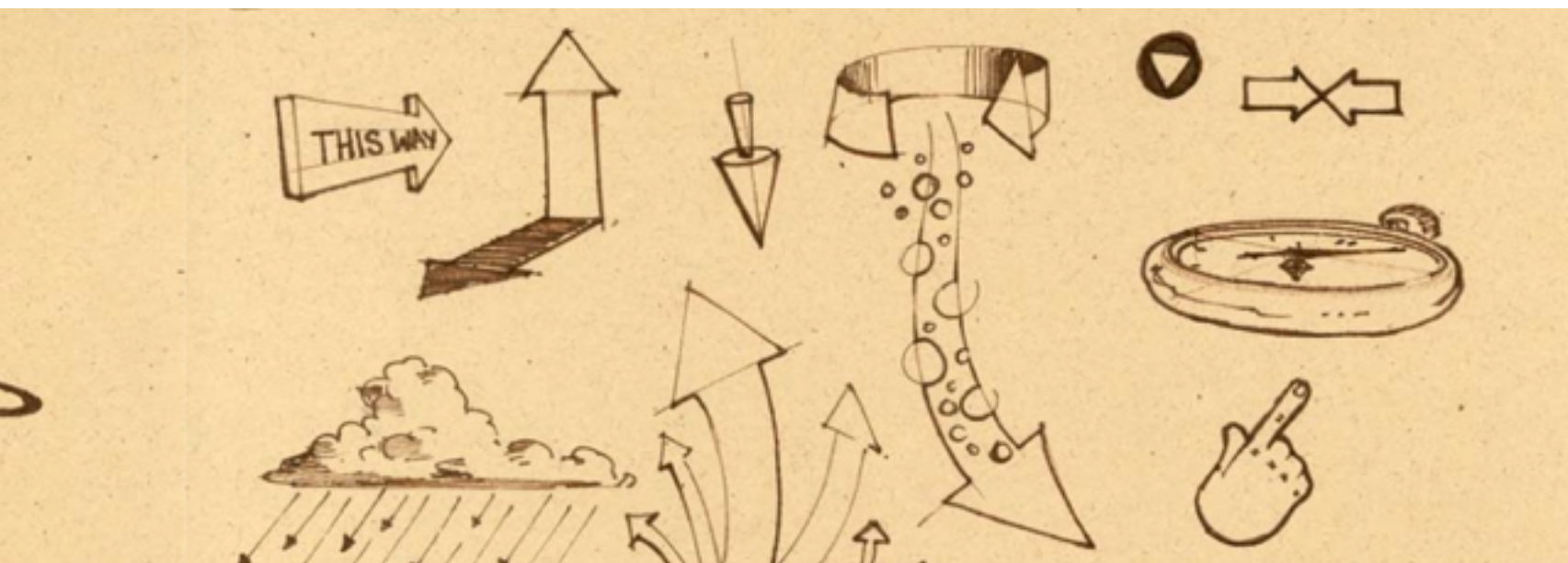
Sections
D3 reading
Self-study
Office hours

Design Lecture
Design Studios

Design Skills



Coding Skills



```
<!DOCTYPE html>  
<meta charset="utf-8">  
<style>  
  
text {  
  font: 10px sans-serif;  
}  
  
</style>  
<body>  
<script src="http://d3js.org/d3.v3.min.js"></script>  
</script>
```

Sections

Short coding tutorials in small groups

Based on a published script

Strongly related to homework assignments

One prototype section recorded

Schedule

CS 171 - Visualization



HARVARD
School of Engineering
and Applied Sciences

[Home](#) [Syllabus](#) [Schedule](#) [Homework](#) [Project](#) [Resources](#) [Fame](#)

Schedule

Schedule is Subject to Change

Week 1

Jan 26 - Feb 1

Reading

D3, Chapters 1-4; VAD, Chapter 1

Tuesday

Jan 27

Cancelled - Snow Day

Thursday

Jan 29

Lecture 1: Introduction. What is visualization? Why is it important? Who are we? Course overview.

Introduction to Homework 1

Week 2

Feb 2 - Feb 8

Reading

D3, Chapters 5-8; VAD, Chapters 1-2

Section

Section 1: GIT, HTML, SVG, CSS, the DOM

Tuesday

Feb 3

Lecture 2: SKILLS: Introduction to D3.

Guest Speaker: [Vadim Ogievetsky](#), co-creator of D3

Thursday

Feb 5

Lecture 3: Data Abstraction, Data Types.

Friday

Feb 6

Sections

<https://www.section.fas.harvard.edu/sectioning/>

Group 1: Mo 02:00-03:30, NW B150

Group 2: Mo 04:00-05:30, NW B150

Group 3: Tu 04:30-06:00, MD 123

Group 4: Tu 05:30-07:00, MD 223

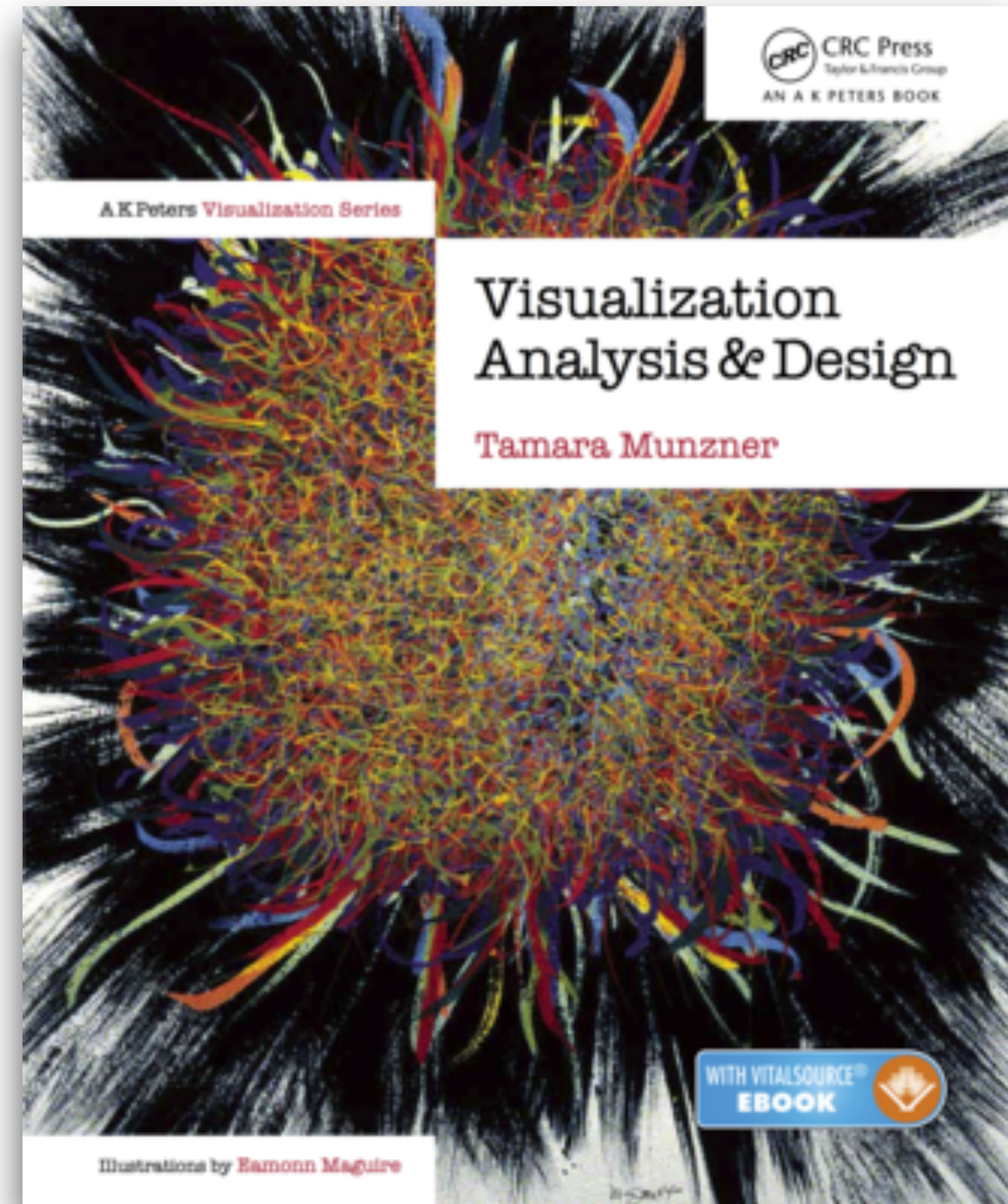
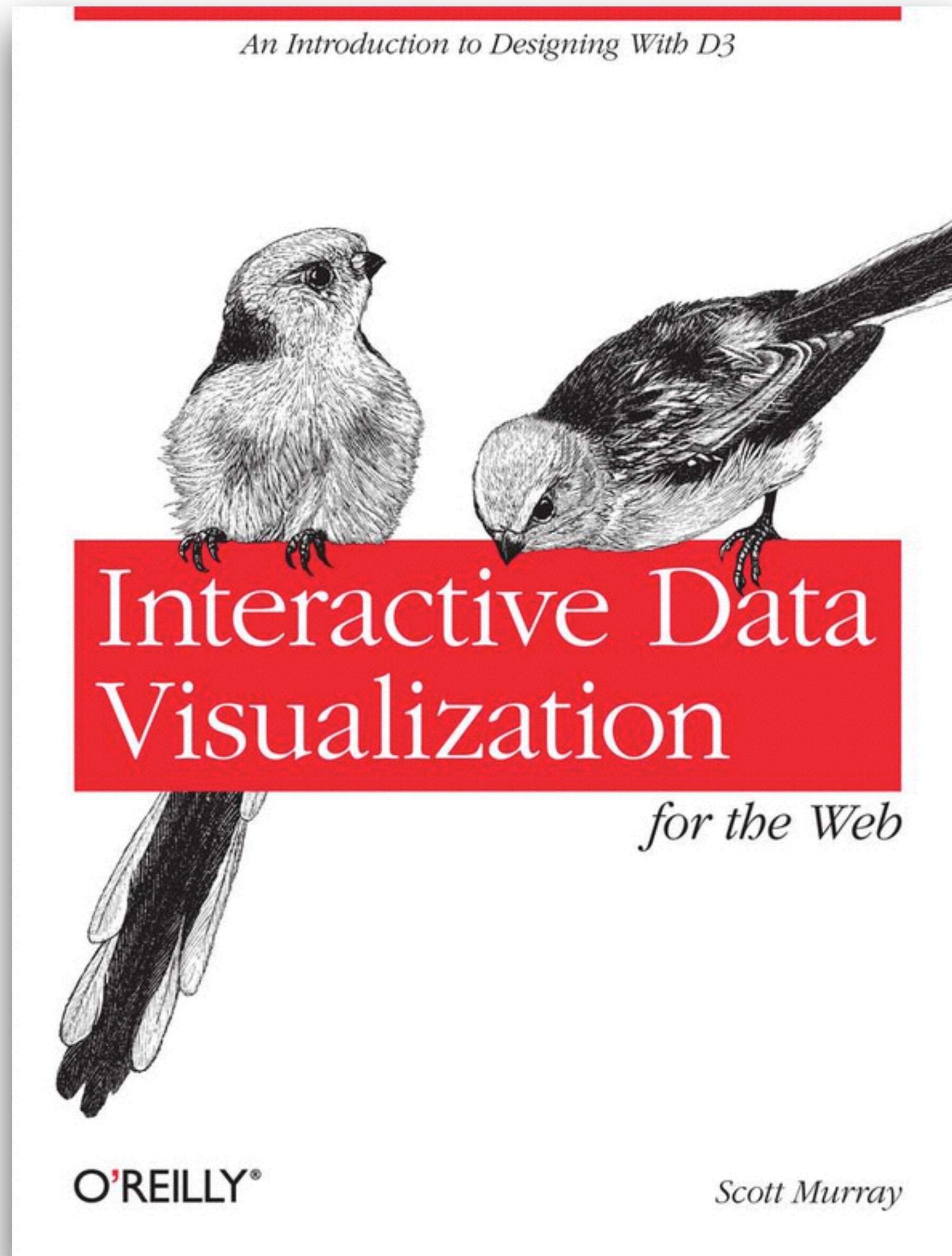
Group 5: We 10:00-11:30, MD 223

Group 6: We 03:30-05:00, NW B150

Online Students:

- recorded section
- material available
- dedicated time to discuss section with TFs

Required Books



Programming

HTML



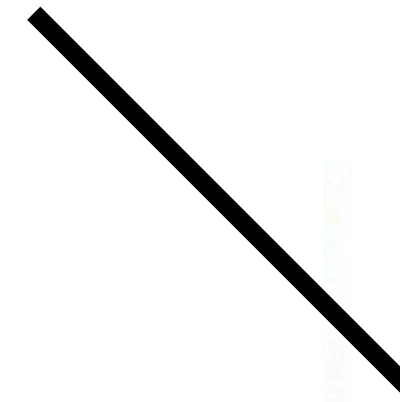
JS



Data-Driven Documents



Is this course for me ???



Prerequisites

Programming experience

C, C++, Java, Python, etc.

Willingness to learn new software & tools

This can be time consuming

You will need to build skills by yourself!

Engineering vs Computer Science

How are you graded?

4+1 Homework Assignments: 50%

Varying value, 2%-14%, depending on length/difficult

Start early! Will take long if you don't know JS/D3 yet

Due on Fridays, four late days

Final Project: 50%

Teams, two milestones

Attendance

Lectures and Sections: attendance appreciated but not required

Design Studios & Guest Lectures: **attendance mandatory**

This Week

HW0, including course survey

Readings

D3 Book, Chapters 1-4

VDA Book, Chapter 1

Table of Contents	
Preface.....	ix
1. Introduction.....	1
Why Data Visualization?	1
Why Write Code?	2
Why Interactive?	2
Why on the Web?	3
What This Book Is	3
Who You Are	4
What This Book Is Not	5
Using Sample Code	5
Thank You	6
2. Introducing D3.....	7
What It Does	7
What It Doesn't Do	8
Origins and Context	9
Alternatives	10
Easy Charts	10
Graph Visualizations	12
Geomapping	12
Almost from Scratch	13
Three-Dimensional	13
Tools Built with D3	14
3. Technology Fundamentals.....	15
The Web	15
HTML	17
Content Plus Structure	18

Next Week

Tuesday: Introduction to D3
Guest lecture by Vadim Ogievetsky

Sections starting Monday:

github, HTML / CSS, DOM

Office hours start!

CS 171 Homework 0

Due: Friday, February 6th, 11:59 pm.

Welcome to CS171. In this class, we will be using a variety of tools that will require some initial configuration. To ensure everything goes smoothly moving forward, we will setup the majority of those tools in this homework. This homework will not be graded **except** Problem 2, which will be graded. In Problem 2 you set up git and github for this course, which is essential before starting with HW 1.

Problem 1 - Class Survey, Signups, and Introduction

Sign up to github

You'll be using git and [GitHub](#) to manage homeworks and projects.

Sign up for a github account (if you don't already have one) and request a free account upgrade [on this page](#). You'll need to verify ownership of an **.edu e-mail address** (Harvard, MIT or any other) if you didn't sign up with your .edu account. You can use this suggested text to request the account

Sign up for GitHub now!