Visualization

Johanna Beyer
jbeyer@seas.harvard.edu

cs171.org

Outline

- Who?
- How?
- What?
- Why?

Outline

- Who?
- How?
- What?
- Why?

Johanna Beyer

Lecturer, Research Associate

PhD in Computer Science at the Vienna University of Technology, Austria

Visual Computing Group, PI: Hanspeter Pfister

Visual Computing Group

Group Leader:
Professor Hanspeter Pfister

Our Research:
Information Visualization, Volume Visualization, Visual Analytics, Augmented Reality, Computer Graphics, Vision, Segmentation & Machine Learning

vcg.seas.harvard.edu

CS171 Staff

Ronell Sicat (Head TF) - Postdoctoral Fellow, Visual Computing Group

Michael Behrisch - Postdoctoral Fellow, Visual Computing Group
Zona Kostic - Postdoctoral Researcher, Web-based Visualization
Fritz Lekschas - PhD Student, Visual Computing Group
Charlene Hwang - College Student

Alain Ibrahim - Web Developer
Alexandra Abrahams - College Student
Katherine Harrison - College Student
Javier Cuan Martinez - College Student

About You

CS
171

Outline

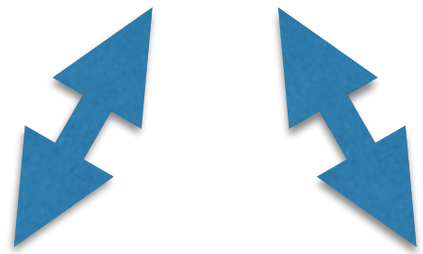
- Who?
- How?
- What?
- Why?

Course Structure

CS 171

CS 171 Goals

Theory apply fundamental principles & techniques

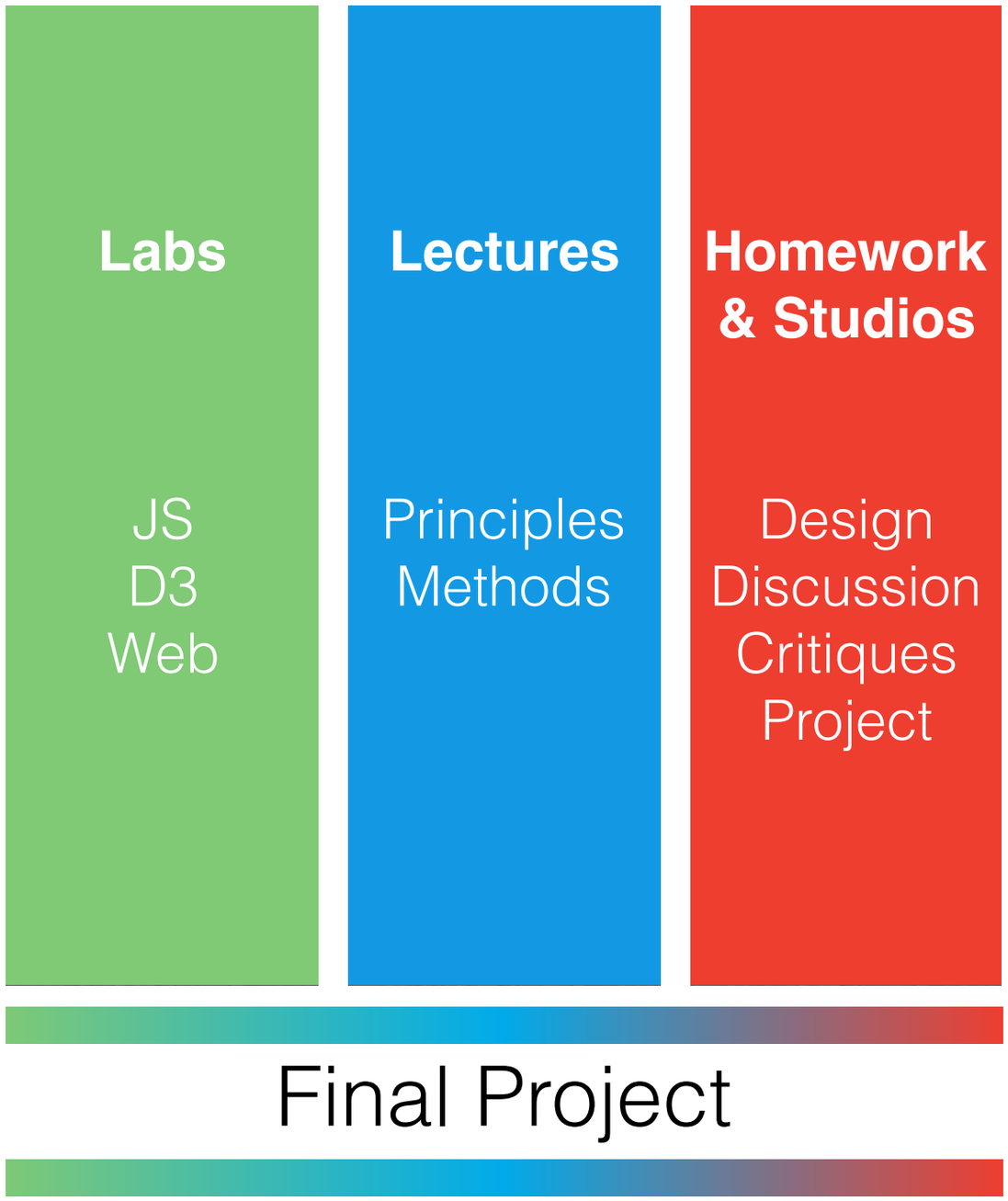


Design Skills  **Coding Skills**

create, evaluate and critique visualization designs

implement interactive web-based data visualizations

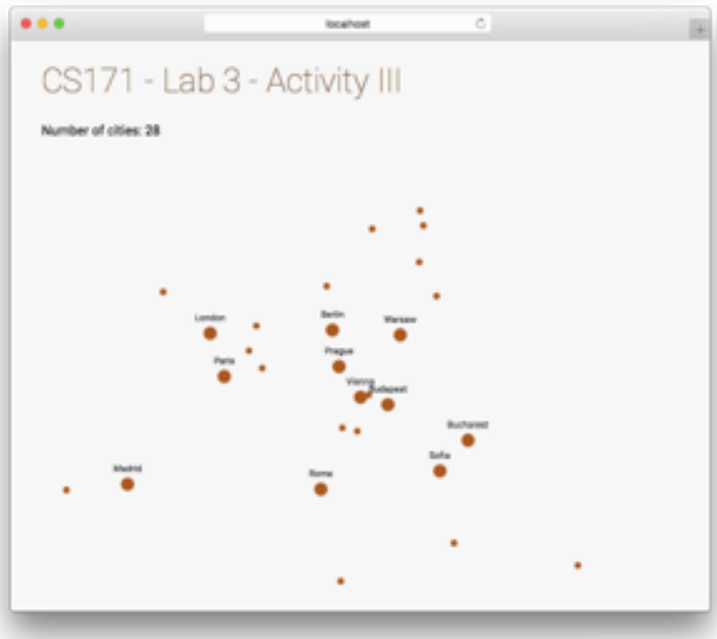
Develop a substantial visualization project!

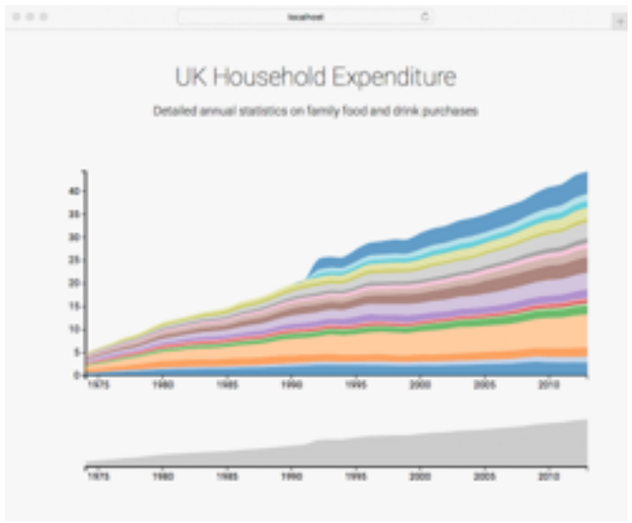

Labs
JS
D3
Web
Lectures
Principles
Methods
Homework
& Studios
Design
Discussion
Critiques
Project
Final Project

Labs

JS
D3
Web

Tue,
2:30-4:00





Lectures

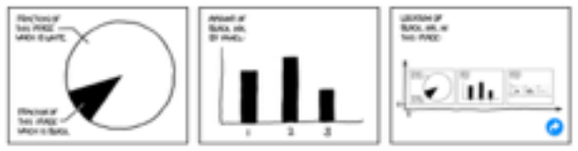
Principles
Methods

Thur,
2:30-4:00

CS 171

Design Principles

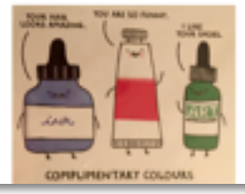
Hanspeter Pfister
pfister@seas.harvard.edu



CS 171

Perception

Hanspeter Pfister (pfister@g.harvard.edu)



CS 171

Cognition

Hendrik Strobelt

CS 171

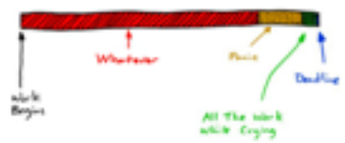
Interaction for Visualization

Hanspeter Pfister

CS 171

Visualization Process

Hendrik Strobelt



Team Work

CS 171

CS 171

Vis Exploration

Hendrik Strobelt / Hanspeter Pfister

CS 171

Evaluation & Innovation

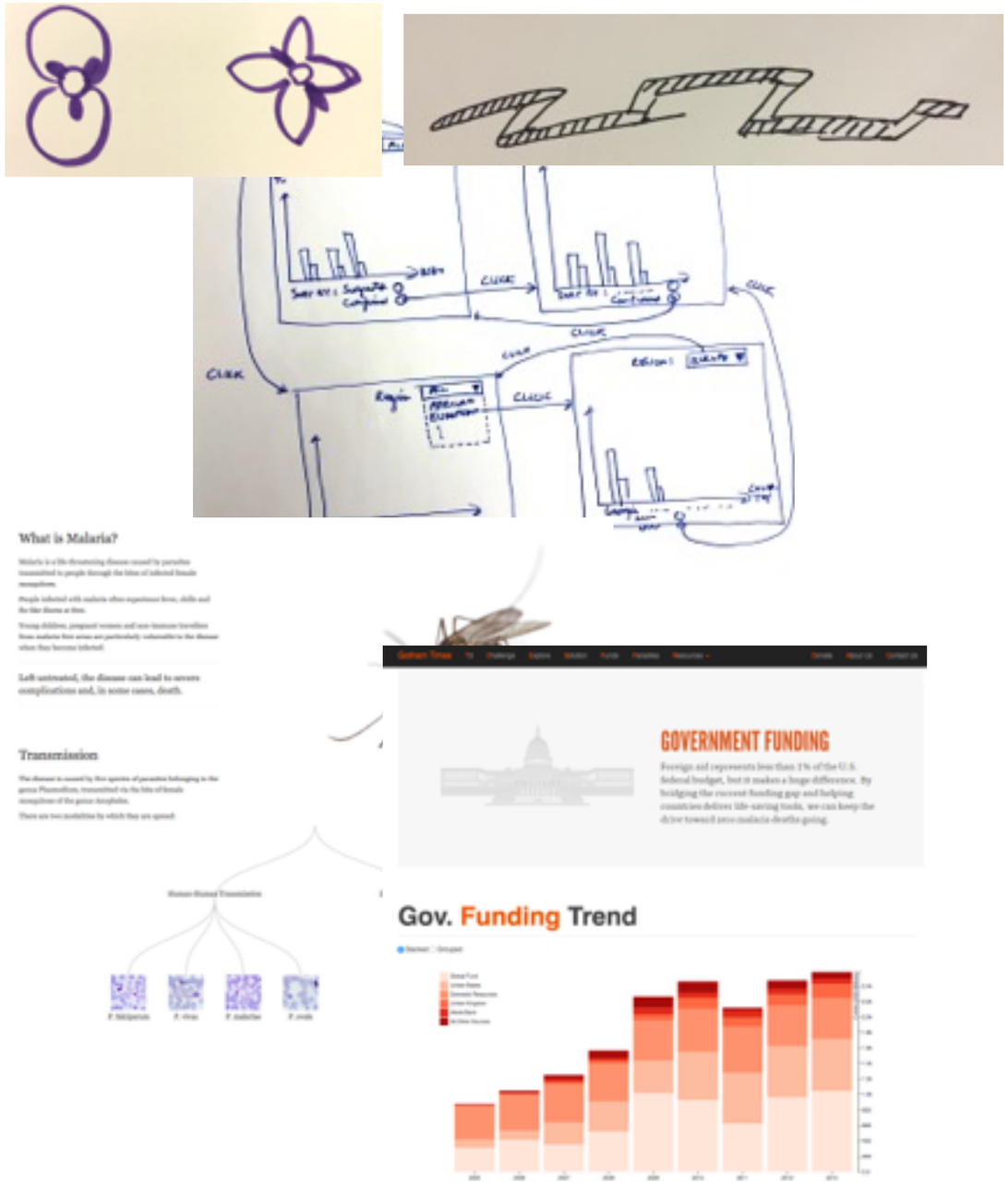
Hanspeter Pfister / Hendrik Strobelt

Homework & Studios

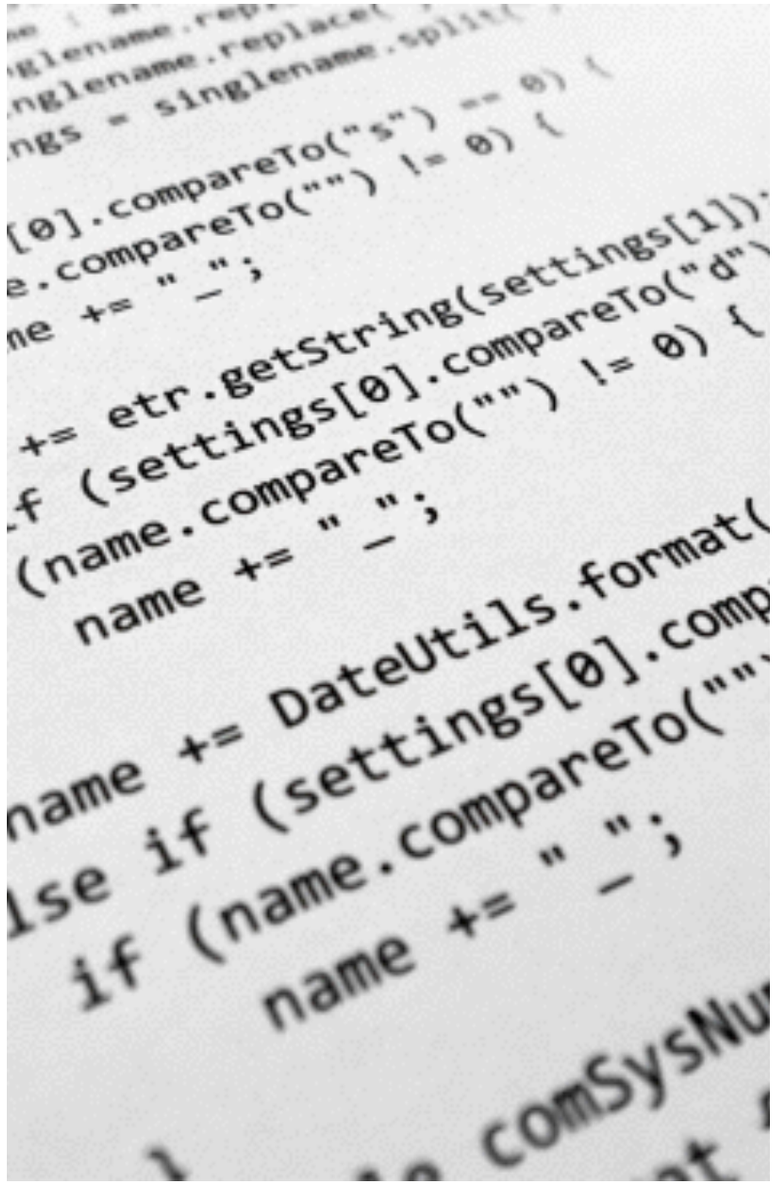
Design
Discussion
Critiques
Project

Studios Mandatory

Homework
Due Mo



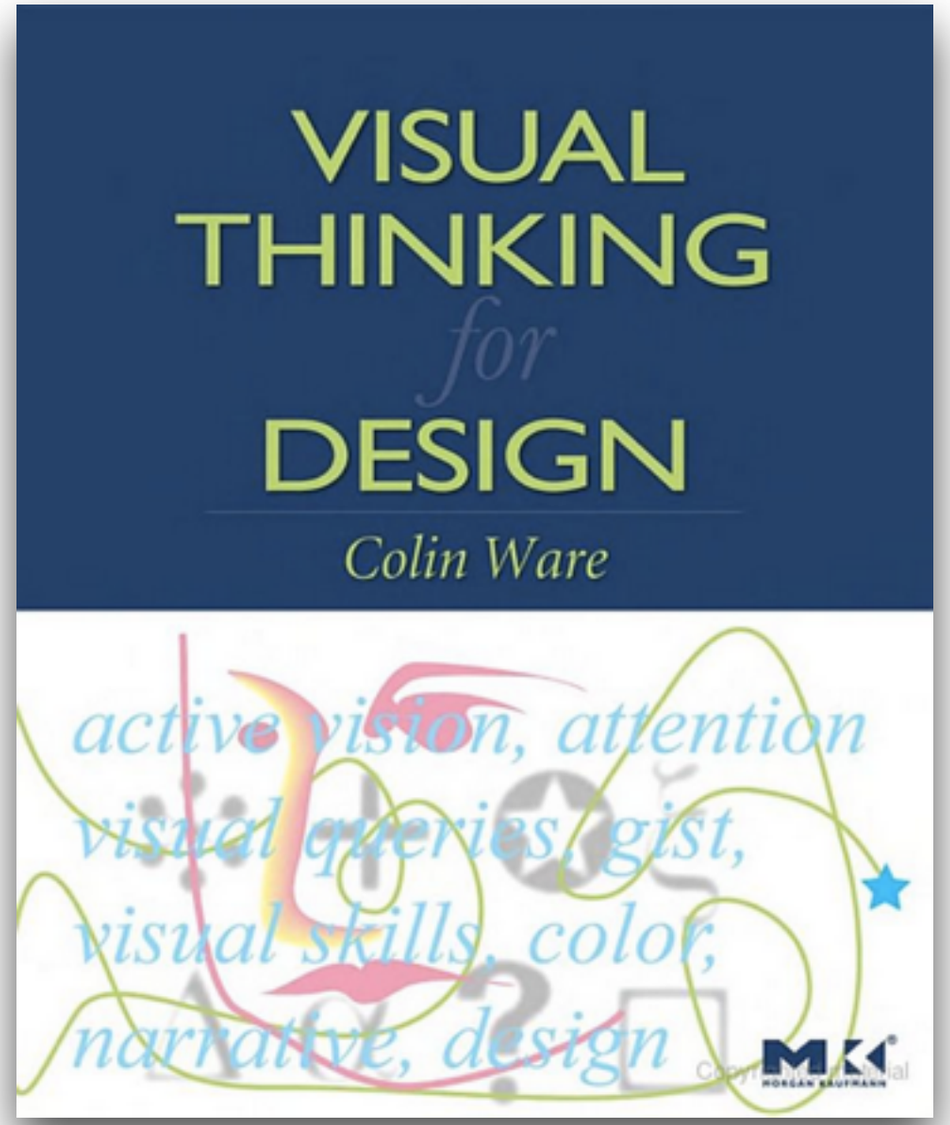
Design



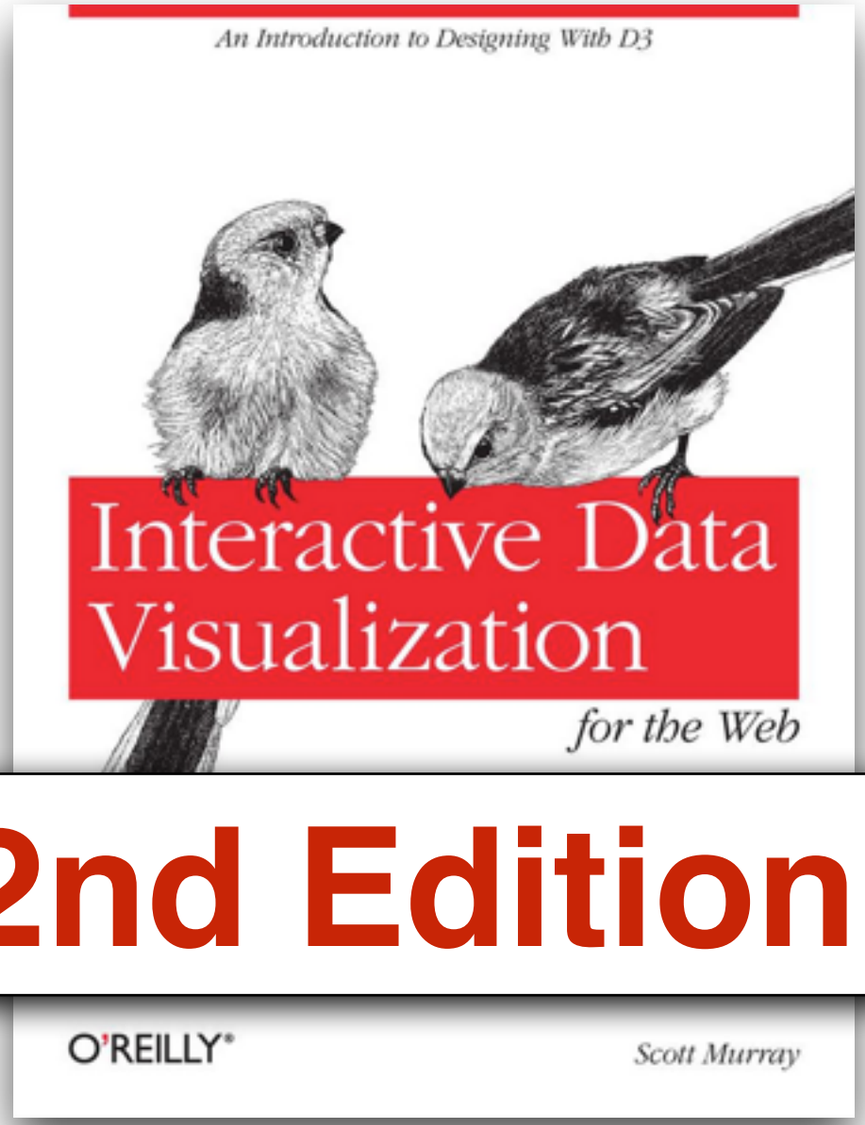
Coding

Readings & Quizzes

Complete the assigned readings before class
Short online quizzes, due prior to each class



Design



Coding

What about exams and grading ???

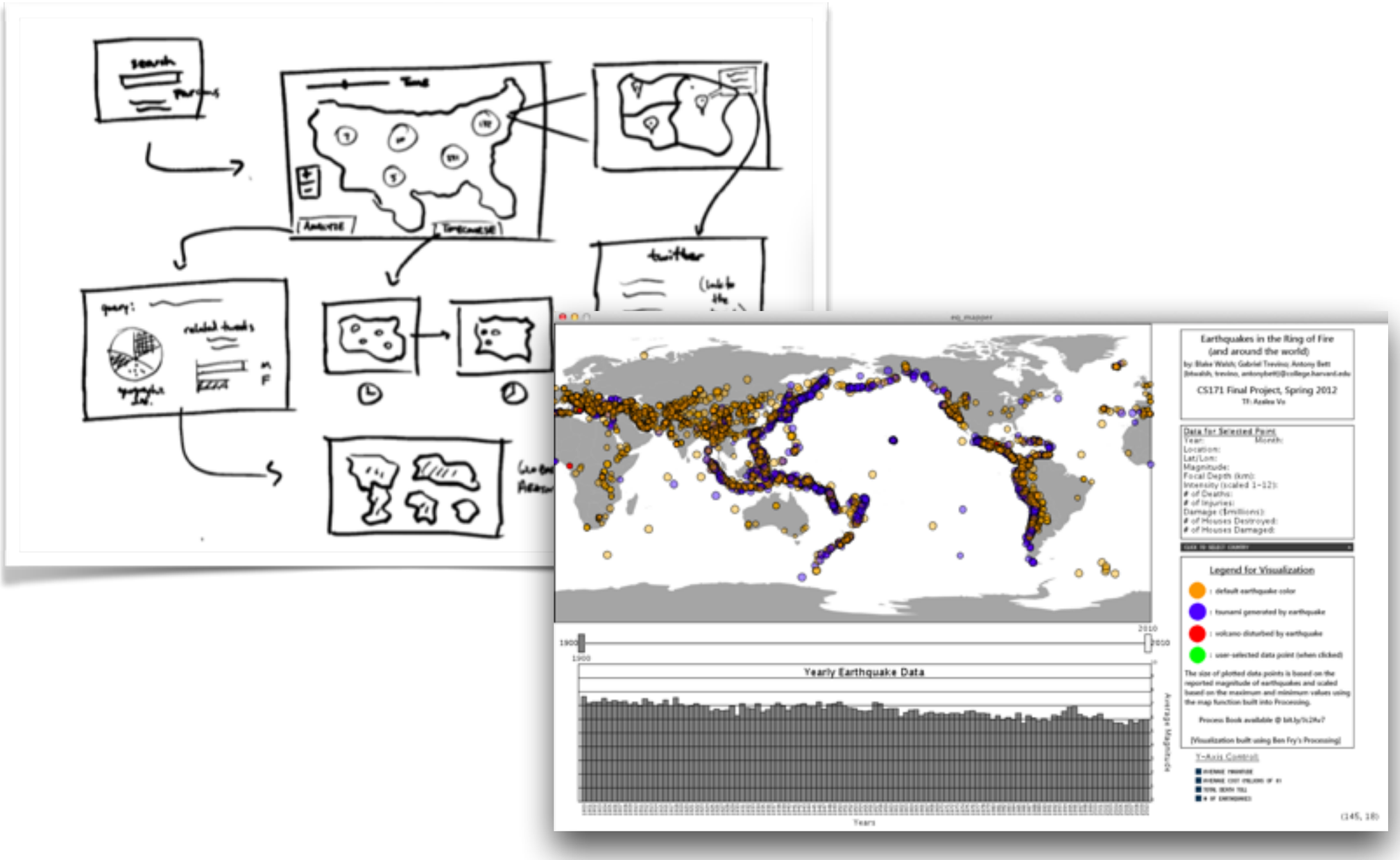
Grading

- Participation (attendance, pre-quizzes, class, studios, Piazza) (15%)
- Midterm (20%)
- Homework Assignments (30%)
- Project Assignments (35%)

- Graduate students: Extra credit assignments in homework are required for you, as well as more effort on the final project.

Midterm

Week Of 10/23 - 10/29

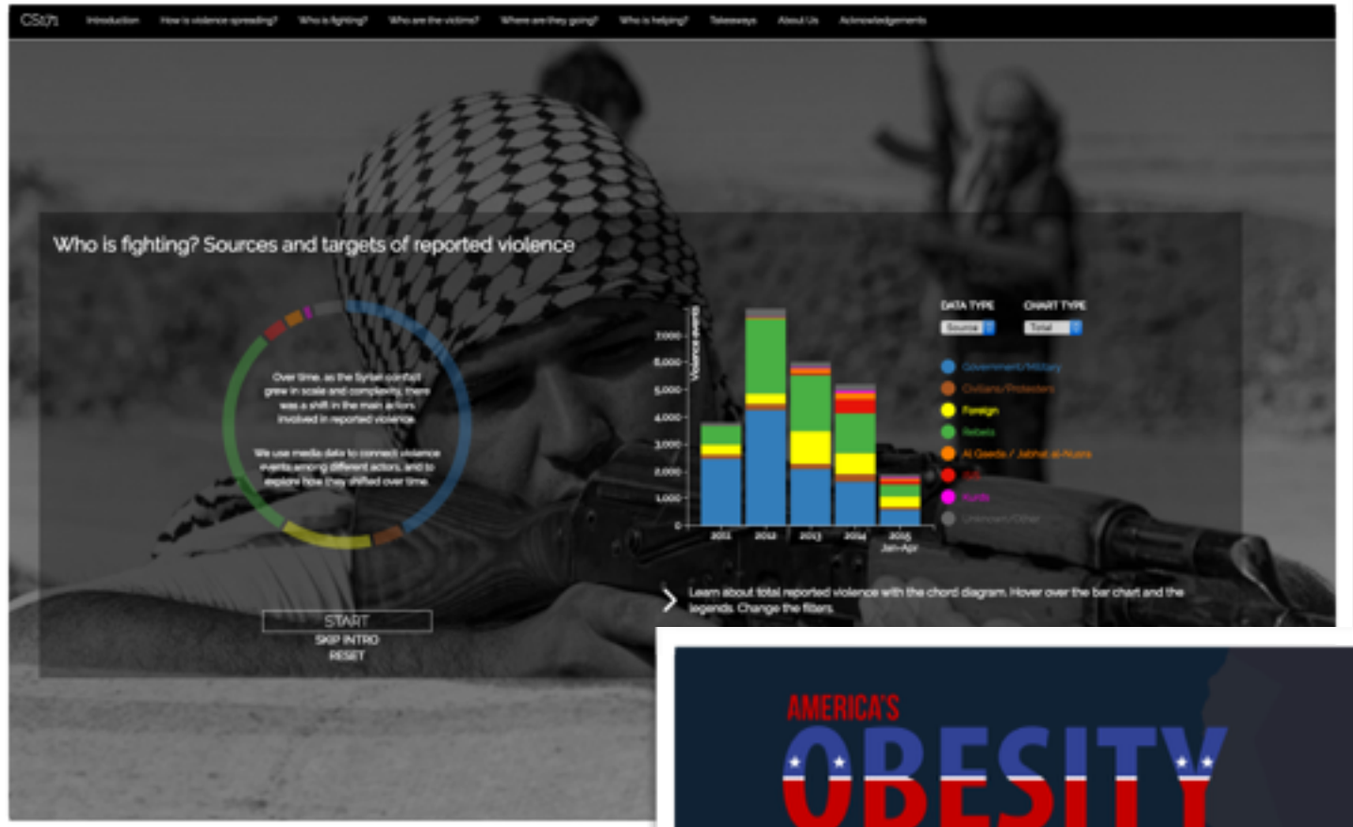


Part I: 90 minute design exercise (in class)
Part II: 5-10 hours D3 implementation (at home)

Group Project

Teams of 3 Students

www.syria-visualized.com



http://obesity-epidemic.github.io/

Prerequisites

Programming experience

- At least at the level of CS50
- C, C++, Java, Python, etc.

Willingness to learn new software & tools

- This can be **time consuming**
- You will need to build skills by yourself

Policies

- Late policy: Assignments will not be accepted late
- Attendance: Mandatory for studios, midterm, several required classes (see schedule)
- Cheating: Follow the Harvard Honor Code
- **Read the syllabus** at www.cs171.org!
- **Check Canvas** for all assignments (weekly modules)

Device Policy

- Bring laptop to lecture, lab, and studio
- Please only use it for in-class activities!

Enrollment

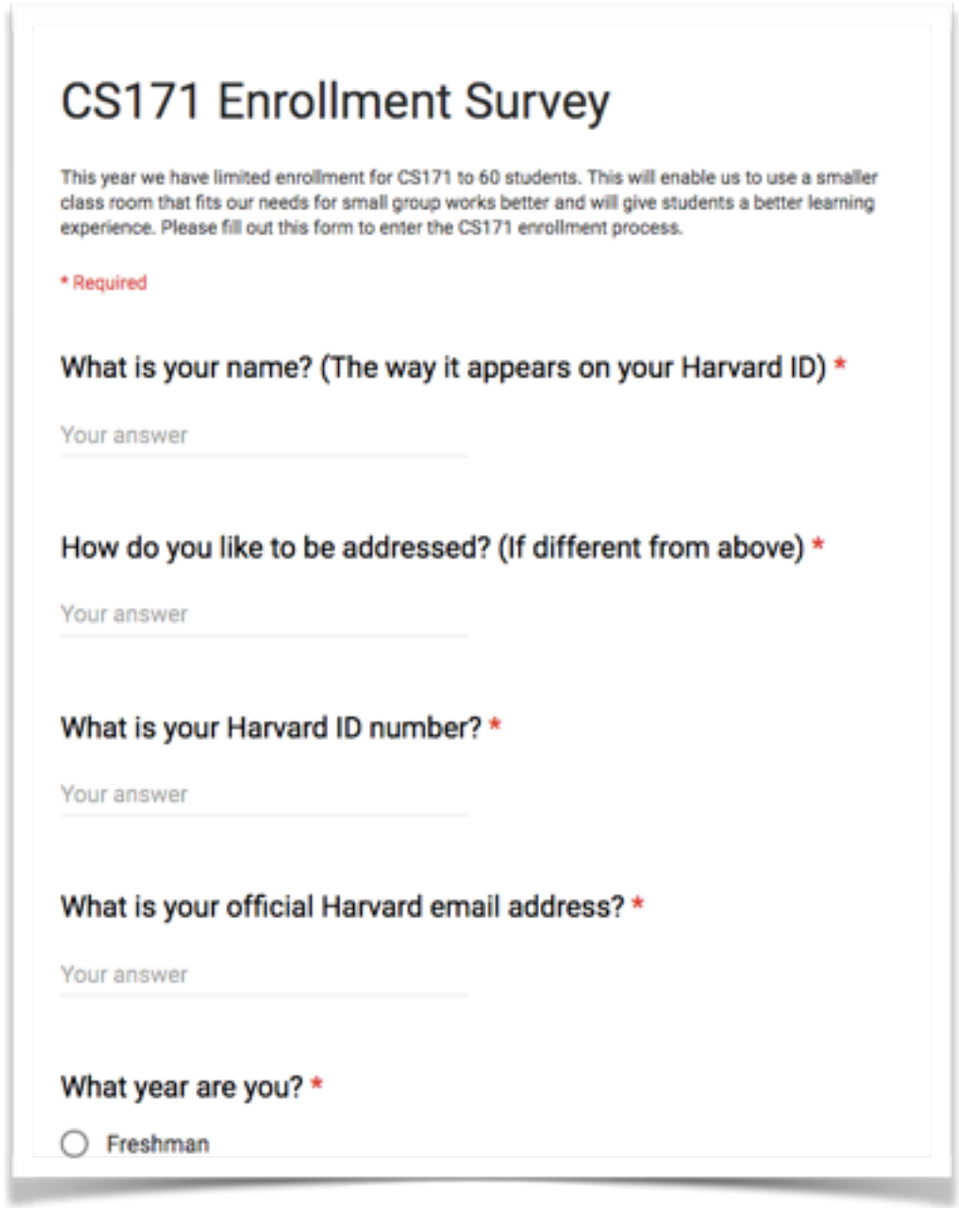
Capped at 70

Finish HW0 by Monday (incl. enrollment survey)

We e-mail you by Tuesday evening

You have to reply by Wednesday noon

You need Canvas access !!!



CS171 Enrollment Survey

This year we have limited enrollment for CS171 to 60 students. This will enable us to use a smaller class room that fits our needs for small group works better and will give students a better learning experience. Please fill out this form to enter the CS171 enrollment process.

* Required

What is your name? (The way it appears on your Harvard ID) *

Your answer

How do you like to be addressed? (If different from above) *

Your answer

What is your Harvard ID number? *

Your answer

What is your official Harvard email address? *

Your answer

What year are you? *

Freshman

Communicate

- Studios
- Piazza
- Office Hours (starting next week)

- Email
  staff@cs171.org
  jbeyer@seas.harvard.edu
  sicat@g.harvard.edu

CS 171 Visualization

HARVARD
John A. Paulson
School of Engineering
and Applied Sciences



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

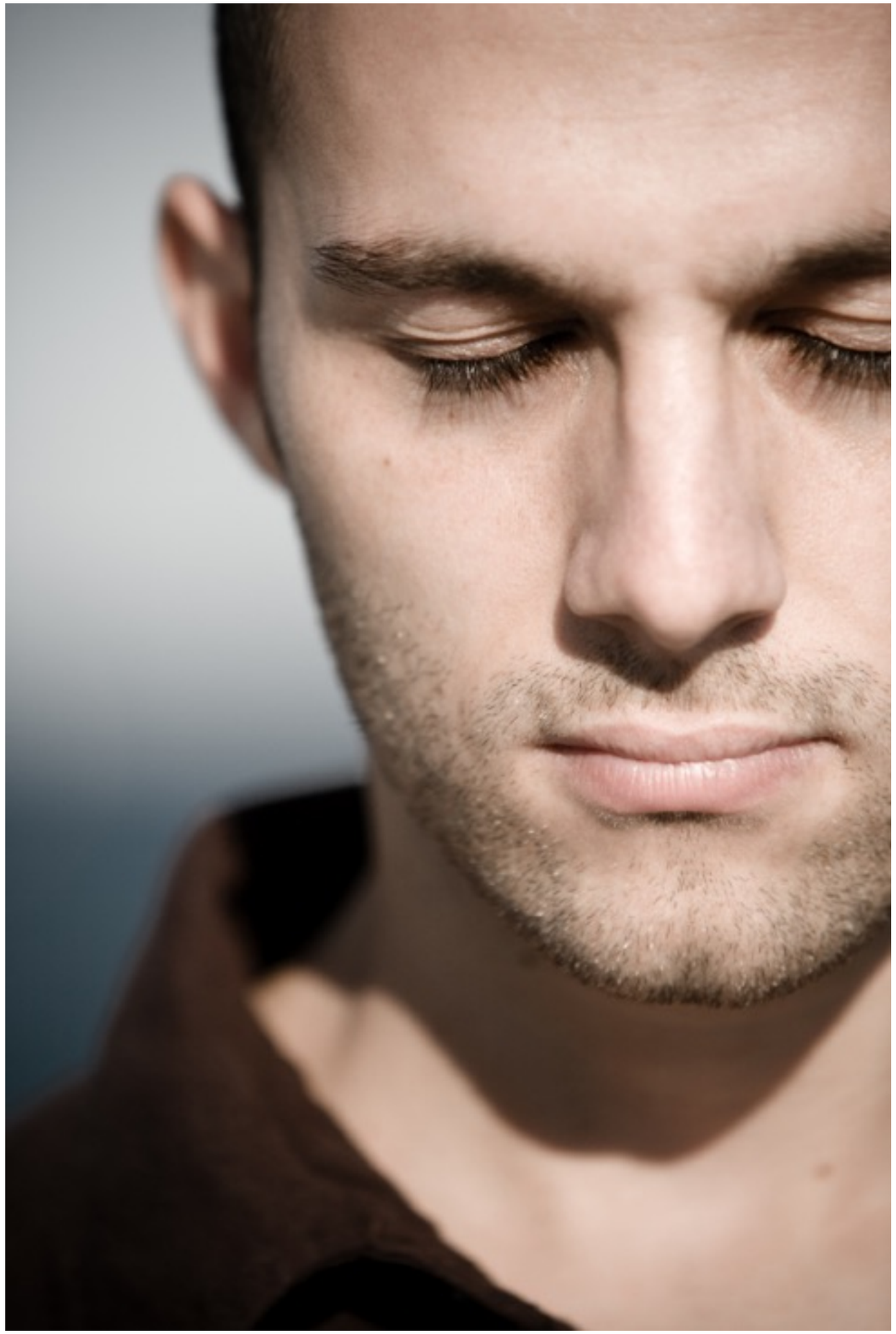
The amount and complexity of information produced in science, engineering, business, and everyday human activity is increasing at staggering rates. The goal of this course is to expose you to visual representation methods and techniques that increase the understanding of complex data. Good visualizations 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, methods for visualization of data from a variety of fields, and programming of interactive web-based visualizations using D3.

cs171.org

Outline

- Who?
- How
- What?
- Why?

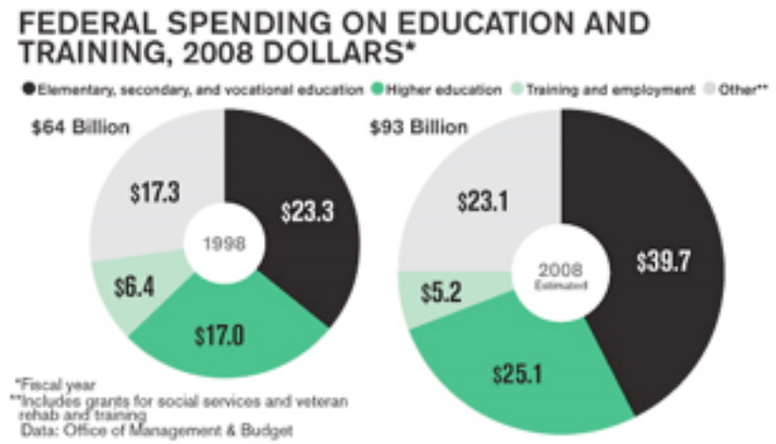


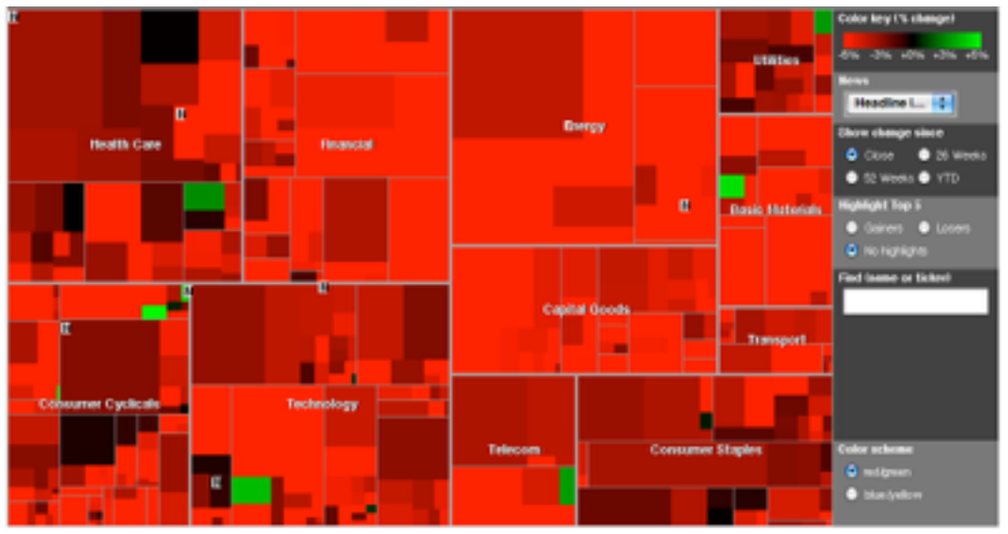
vi·su·al·i·za·tion

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

The American Heritage Dictionary

…to convey information through graphical representations of data





Visualization Goals

Data Exploration - find the unknown

Data Analysis - check hypotheses

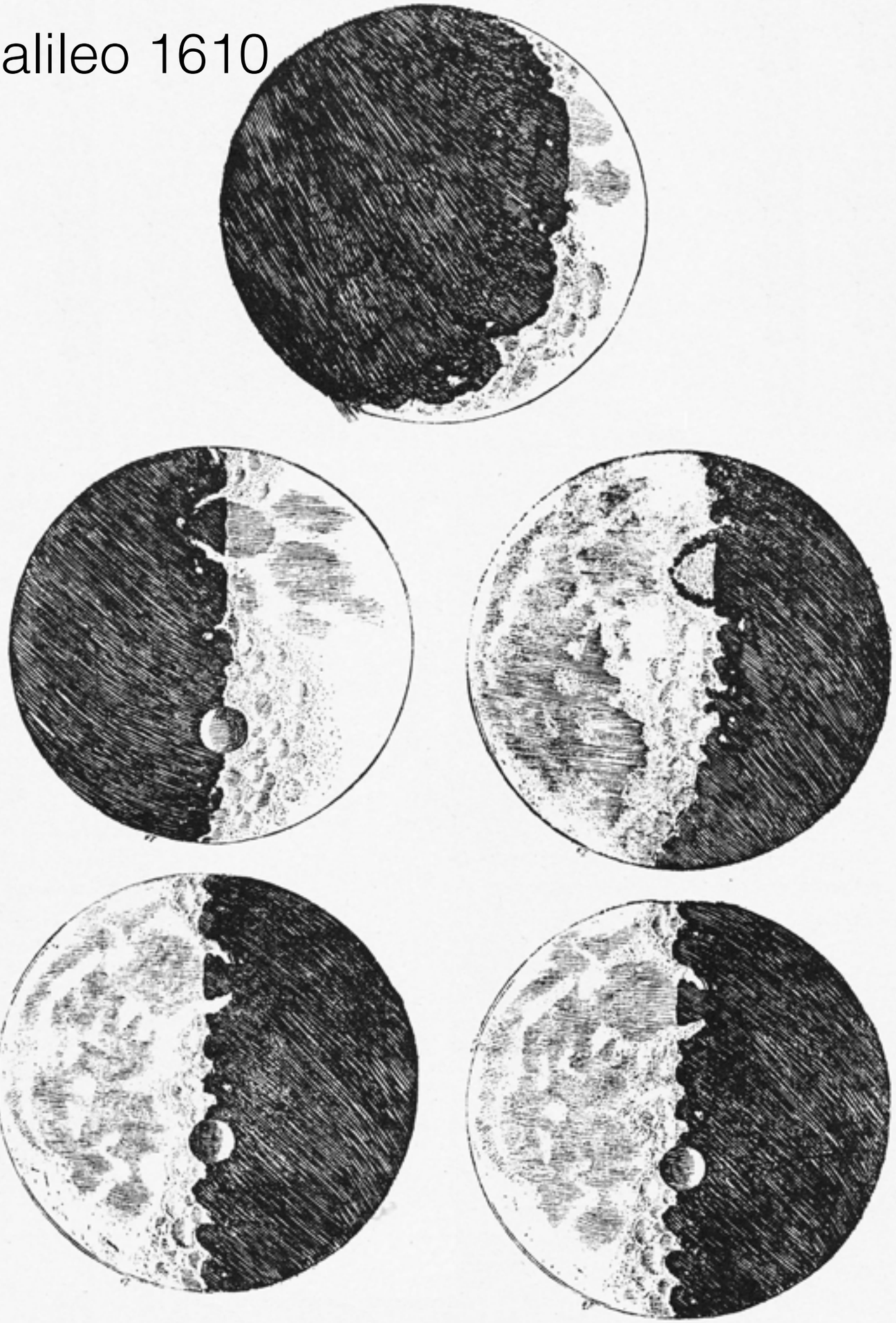
Presentation - communicate and disseminate

Explore & Analyze

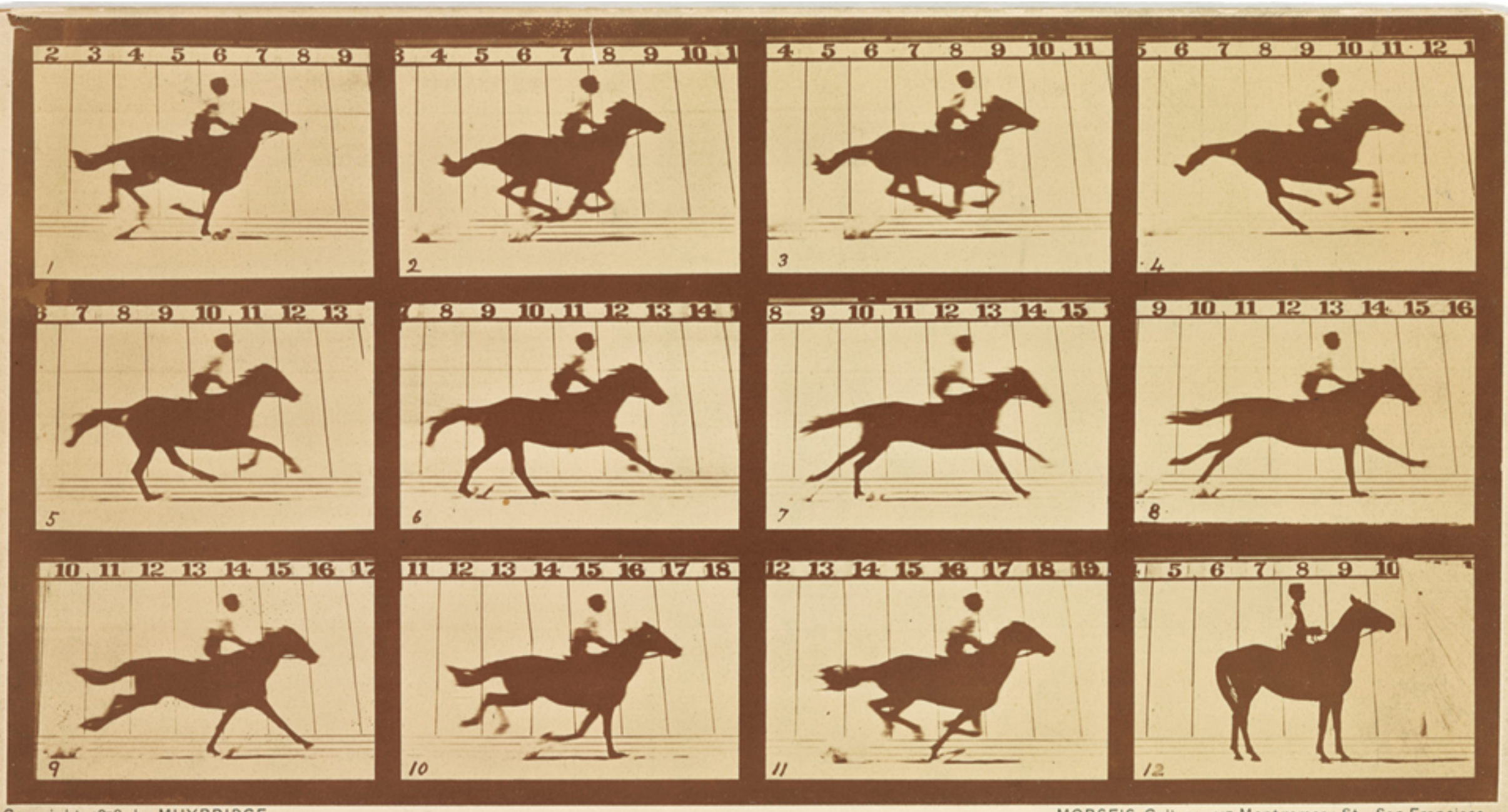
CS
171

Leonardo da Vinci 1485

Galileo 1610



E. J. Muybridge 1878



MORSE'S Gallery, 417 Montgomery St., San Francisco

The Horse in Motion.

Illustrated by

MUYBRIDGE.

Patent for apparatus applied for.

Automatic Electro-Photograph.

"SALLIE GARDNER," owned by LELAND STANFORD; ridden by G. DOMM, running at a 1.40 gait over the Palo Alto track, 19th June, 1878.

The negatives of these photographs were made at intervals of twenty-seven inches of distance, and about the twenty-fifth part of a second of time; they illustrate consecutive positions assumed during a single stride of the mare. The vertical lines were twenty-seven inches apart; the horizontal lines represent elevations of four inches each. The negatives were each exposed during the two-thousandth part of a second, and are absolutely "untouched."

Edmond Halley1686

IIIh Variable Winds

Madera IIh

BARBARIA

IIh MEDI

THE NORTH SEA

Bahama Isles

Canary Isles

Cuba

Spagniola

or

C Blanco

Jamaica

Caribe Isles

ATLANTICK

C Verd Isles

Gambia

AFRI

C Verd

OCEAN

Calms and Tornado's

GUINEA

ÆTHIO

T OF

GUIANA

PERU

FICK

AMERICA

BRASILIA

THE ÆTHIOPICK

OCEAN

Ascension

St Helena

Congo

Angola

Merid: Londini

TH SEA

Monon

IIIIh

IIh

Variable Winds

Cape of good Hope

William Playfair 1820

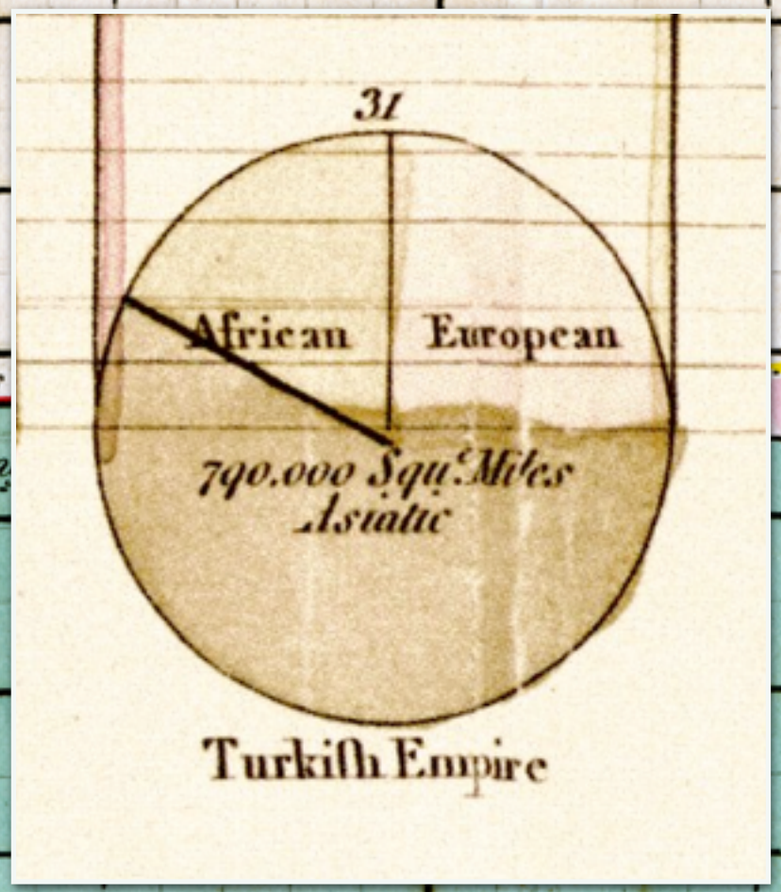
17th. Century

No. 3.

18th. Century

CHART

Shewing the Value of the Quarter of Wheat
In Shillings & in Days Wages
OF A GOOD MECHANIC
from
1565 to 1821.



Shillings

in Shillings

in Wages

in Shillings

in Wages

1625 1650 1675 1700 1725 1750 1775 1800

NameVoyager: Explore name trends letter by letter FAQ/Help | More on Visualization

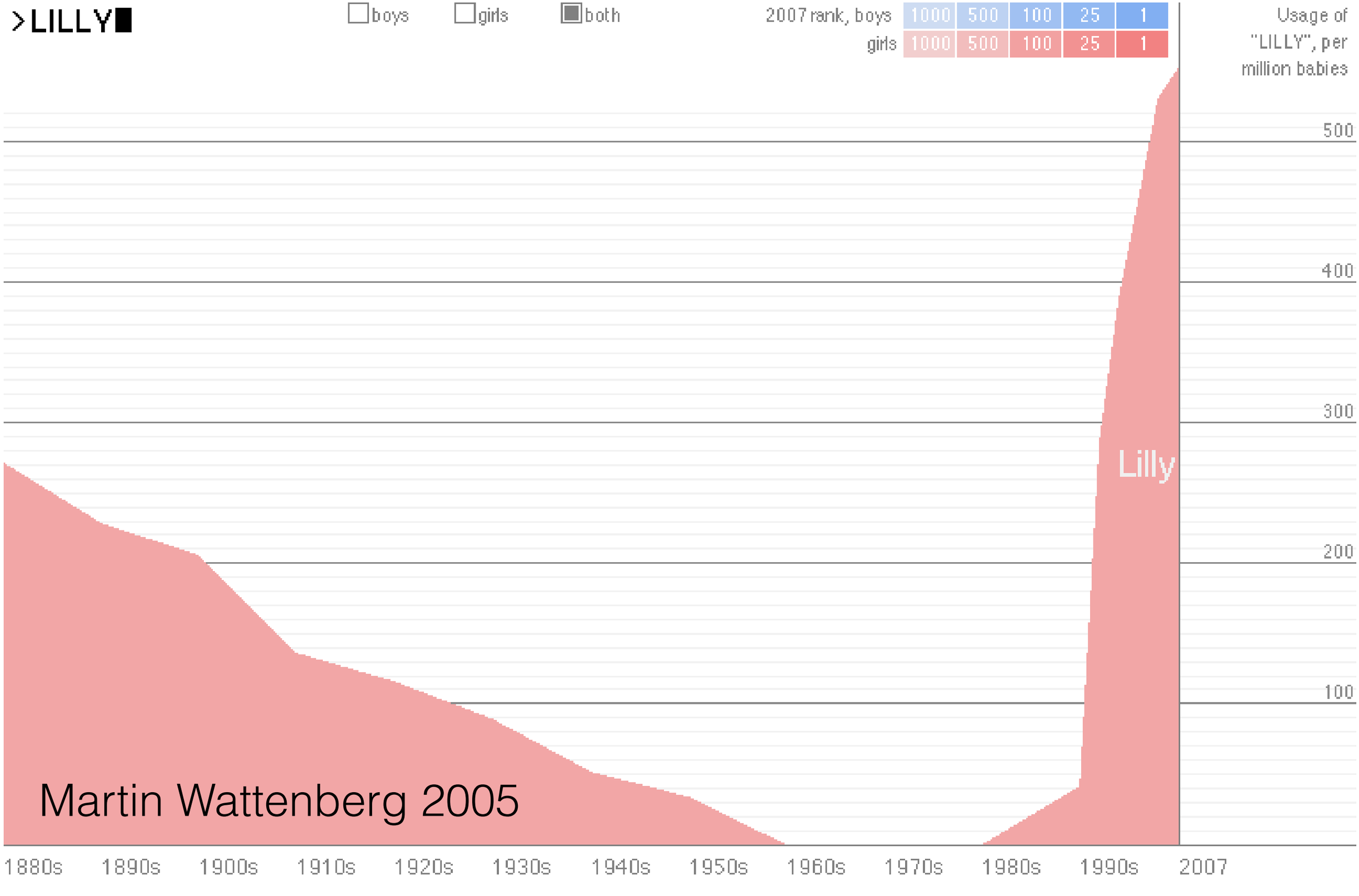
New! Try the **NameMapper** to see where your favorite names are being used, and **Namipedia** for full info on every name!

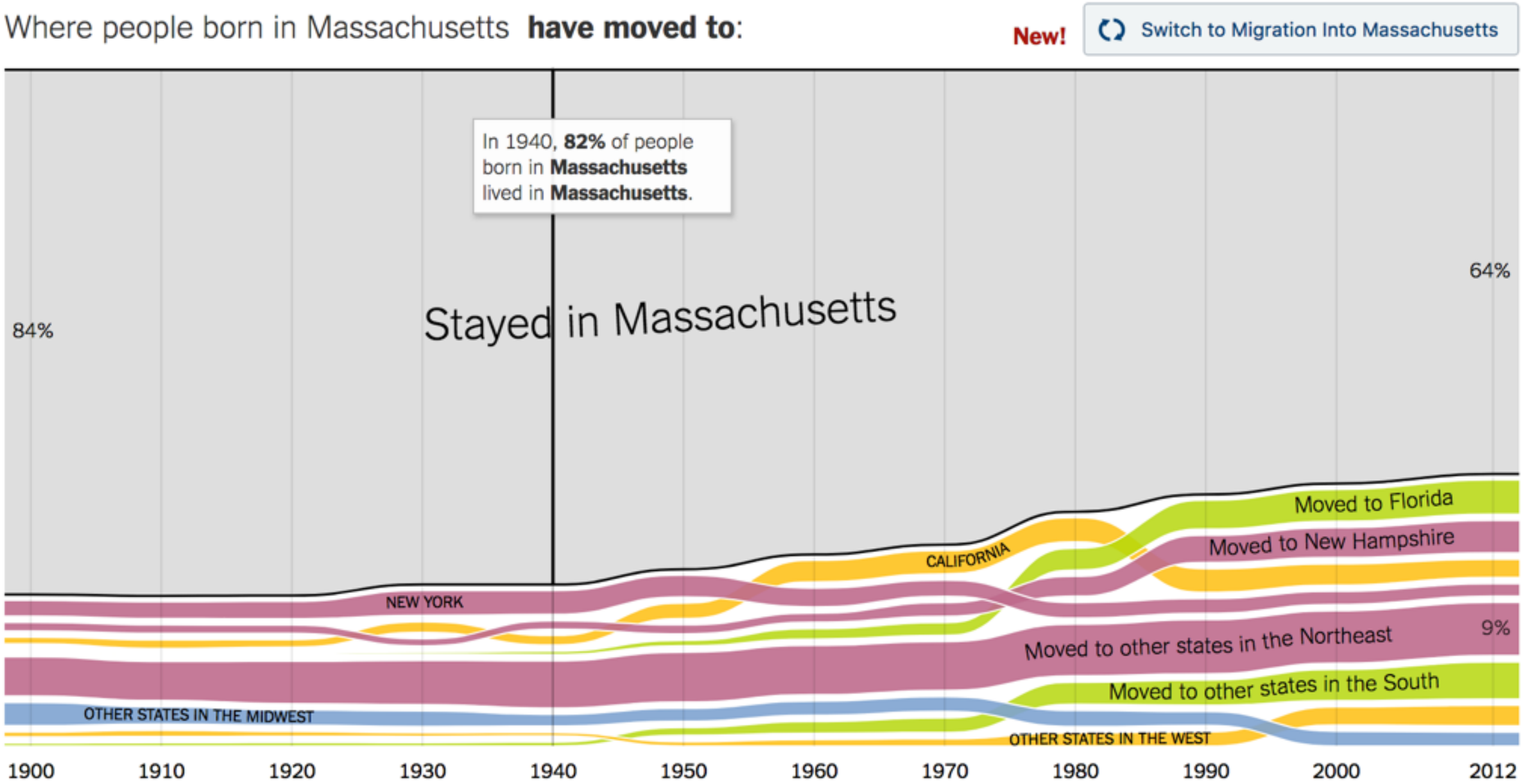
>LILLY

☐ boys ☐ girls ■ both

2007 rank, boys 1000 500 100 25 1

girls 1000 500 100 25 1





New York Times

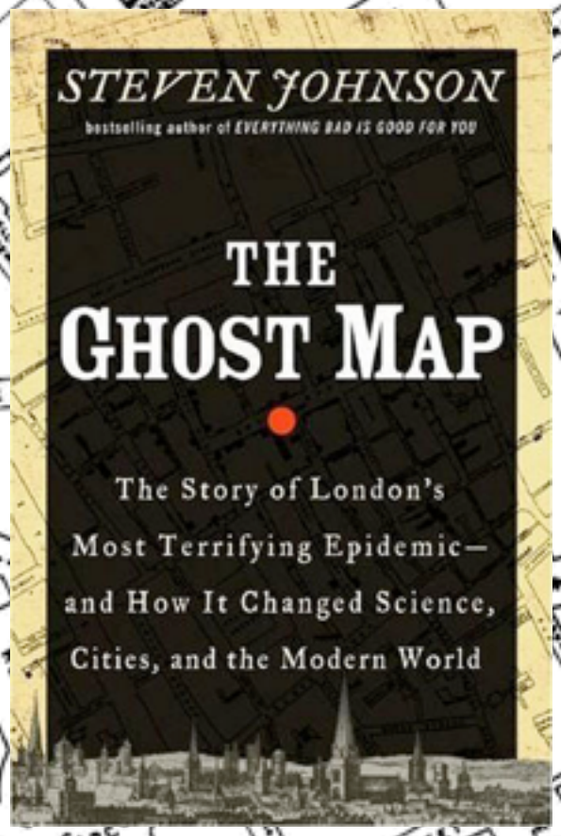
Communicate

CS
171

Hans Rosling 2006

Konya town map, Turkey, c. 6200 BC

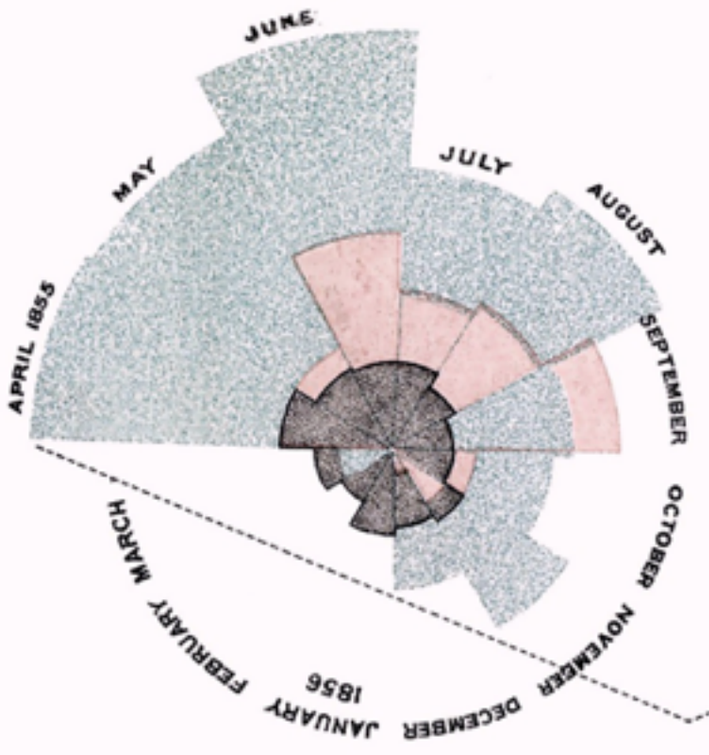
John Snow 1854




The Broad Street Well
GREAT MARLBOROUGH STREET
ARGYLL PLACE
ARGYLL STREET
REGENT STREET
CIRCUS
GLENHAM MEWS
GLENHAM STREET
PORTLAND STREET
NOEL STREET
WORK HOUSE
BROAD STREET
GREAT PULTENEY STREET
LITTLE PULTENEY STREET
PETER STREET
OLD COMPTON STREET
SOHO SQUARE
CARLISLE STREET
RICHMOND BUILDINGS
NEW BURLINGTON STREET
SAVILE ROW
GOLDEN SQUARE
HANOVER STREET
MADDOX STREET
ARCHER STREET
RICHMOND ST

DIAGRAM of the CAUSES of MORTALITY IN THE ARMY IN THE EAST.

2.
APRIL 1855 to MARCH 1856.

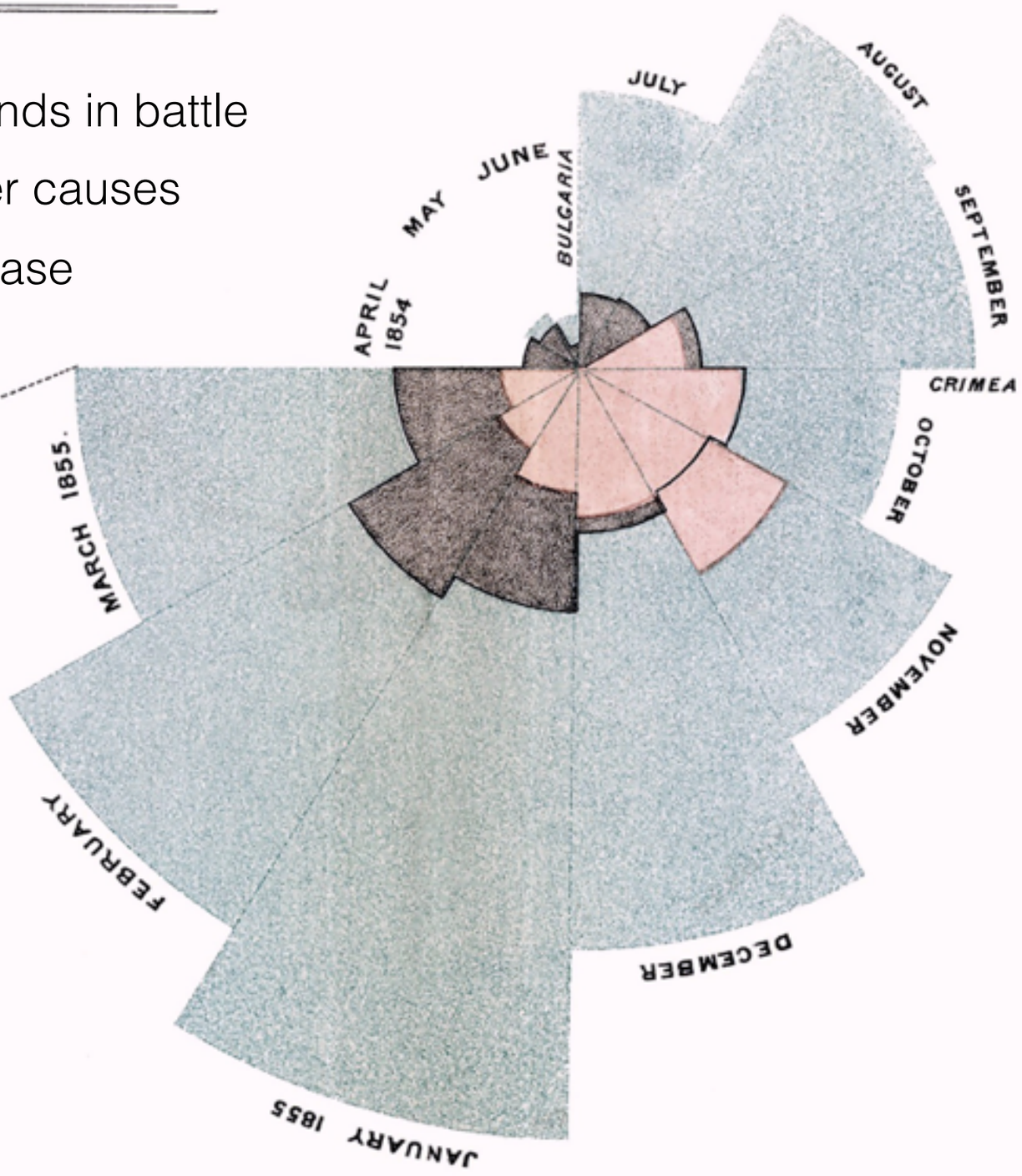


Death from wounds in battle

Death from other causes

Death from disease

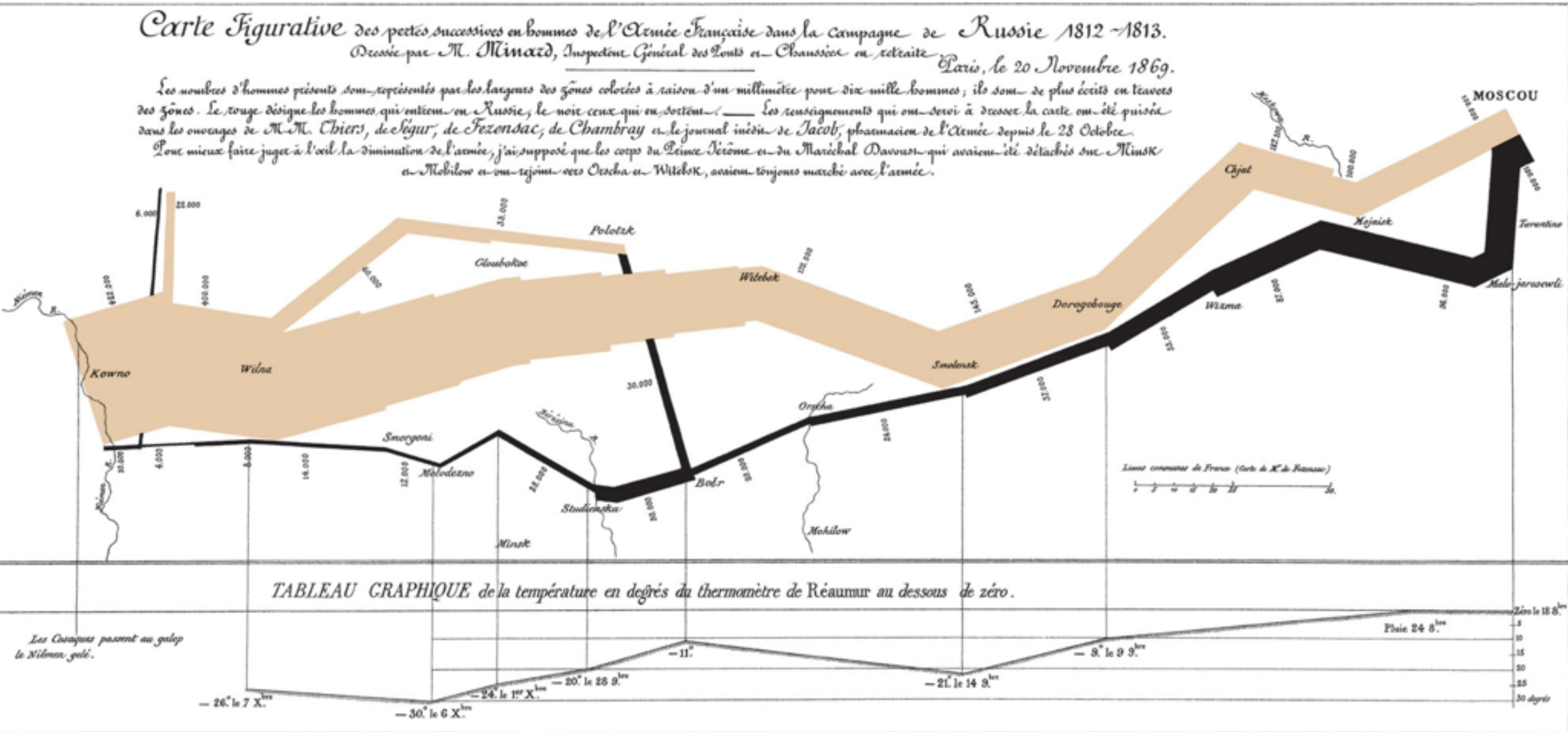
1.
APRIL 1854 to MARCH 1855.



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.
The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases; the red wedges measured from the centre the deaths from wounds; & the black wedges measured from the centre the deaths from all other causes.
The black line across the red triangle in Nov.r 1854 marks the boundary of the deaths from all other causes during the month.
In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.
The entire areas may be compared by following the blue, the red & the black lines enclosing them.

Florence Nightingale 1856

Joseph Minard 1861



Autog. par Regnier, 8. Pas. St.e Marie St. Gen.ve à Paris.
Imp. Lith. Regnier et Dourdet.

Evaluating Visualizations

CS 171

UNDERGROUND

RAILWAYS OF LONDON

1927

TO
Kenton
Harrow and Wealdstone
Headstone Lane
Hatch End for Pinner
Carpenders Park
Bushey and Oxhey
Watford (High Street)
Watford Junction

TO
Harrow
Uxbridge and
Rickmansworth

EDGWARE
BURNT OAK
COLINDALE
HENDON CENTRAL
BRENT
GOLDERS GREEN
PRESTON ROAD
WEMBLEY PARK
NORTH WEMBLEY
NEASDEN & KINGSBURY
DOLLIS HILL
WEST HAMPSTEAD
HAMPSTEAD
BELSIZE PARK
WEMBLEY for SUDBURY
WILLESDEN GREEN
KILBURN BRONDESBURY
FINCHLEY ROAD
SWISS COTTAGE
CHALK FARM
CAMDEN TOWN
SUDBURY TOWN
STONEBRIDGE PARK
QUEENS PARK
KILBURN PARK
MARLBORO' ROAD
MORNINGTON CRESCENT
EUSTON (FOR L.M.&S.R.)
GREAT PORTLAND STREET
KINGS CROSS ST. PANCRAS (FOR L.N.E.R. & L.M.&S.R.)
ALPERTON
HARLESDEN
KENSAL GREEN
ST. JOHNS WOOD
BAKER STREET
MAIDA VALE
EUSTON SQUARE
GOODGE ST.
RUSSELL SQUARE
FARRINGDON
ALDERSGATE
OLD STREET
WHITECHAPEL (FOR E. LONDON R'WAY)
MOORGATE
LIVERPOOL STREET
ST. MARYS
ALDGATE EAST
MARK LANE
MONUMENT
BANK
ALDGATE
MANSION HOUSE
POST OFFICE
CHANCERY LANE
HOLBORN
ALDWYCH
COVENT GARDEN
LEICESTER SQUARE
TOTTENHAM COURT RD.
BRITISH MUSEUM
WARREN ST.
REGENTS PARK
OXFORD CIRCUS
BOND ST.
PICCADILLY CIRCUS
STRAND (FOR S.R.)
TEMPLE
CHARING CROSS (FOR S.R.)
BLACKFRIARS
CANNON STREET (FOR S.R.)
LONDON BRIDGE (FOR S.R.)
WATERLOO (FOR S.R.)
BOROUGH
ELEPHANT & CASTLE
KENNINGTON (NEW STREET)
LAMBETH (NORTH)
OVAL
STOCKWELL
CLAPHAM NORTH
CLAPHAM COMMON
CLAPHAM SOUTH
BALHAM
TRINITY ROAD (TOOTING BEC)
TOOTING BROADWAY
COLLIERS WOOD
SOUTH WIMBLEDON
MORDEN
HIGHGATE
TUFNELL PARK
KENTISH TOWN
CALEDONIAN ROAD
GILLESPIE ROAD
HOLLOWAY ROAD
FINSBURY PARK
DRAYTON PARK
HIGHBURY & ISLINGTON
CANONBURY & ESSEX ROAD
YORK ROAD
ANGEL

TO
Stepney Green
Mile End
Bow Road
Bromley
West Ham
Plaistow
Upton Park
East Ham
Barking and
Southend

TO
Sudbury Hill
South Harrow
Rayners Lane
Eastcote
Ruislip Manor
Ruislip
Ickenham
Hillingdon
and Uxbridge

NORTH EALING
PARK ROYAL
NORTH ACTON
WILLESDEN JUNCTION (NEW STATION)
WESTBOURNE PARK
WARWICK AVENUE
MARYLEBONE (FOR L.&N.E.R.)
BISHOPS ROAD
LADBROKE GROVE
ROYAL OAK
PADDINGTON (FOR G.W.R.)
EDGWARE ROAD
LATIMER ROAD
BAYSWATER
PRAED ST
WEST ACTON
EAST ACTON
UXBRIDGE ROAD
WOOD LANE
SHEPHERDS BUSH
HOLLAND PARK
NOTTING HILL GATE
QUEENS ROAD
LANCASTER GATE
MARBLE ARCH
DOVER ST
DOWN ST
HYDE PARK CORNER
TRAFALGAR SQUARE
WESTMINSTER
KNIGHTSBRIDGE
EALING BROADWAY
EALING COMMON
SOUTH EALING
ACTON TOWN
SOUTH ACTON
GOLDHAWK ROAD
ADDISON ROAD
HIGH STREET KENSINGTON
BROMPTON ROAD
ST. JAMES' PARK
VICTORIA (FOR S.R.)
NORTHFIELDS
TURNHAM GREEN
RAVENSCOURT PARK
BARONS COURT
EARLS COURT
GLOUCESTER ROAD
CHISWICK PARK
STAMFORD BROOK
GUNNERSBURY
HAMMERSMITH
WEST KENSINGTON
SOUTH KENSINGTON
SLOANE SQUARE
WEST BROMPTON

TO
Boston Manor
Osterley
Hounslow East
Hounslow Central
(for Heston)
Hounslow West

KEW GARDENS
WALHAM GREEN
PARSONS GREEN
PUTNEY BRIDGE
EAST PUTNEY
SOUTHFIELDS
WIMBLEDON PARK
WIMBLEDON
RICHMOND

THAMES RIVER

REFERENCE

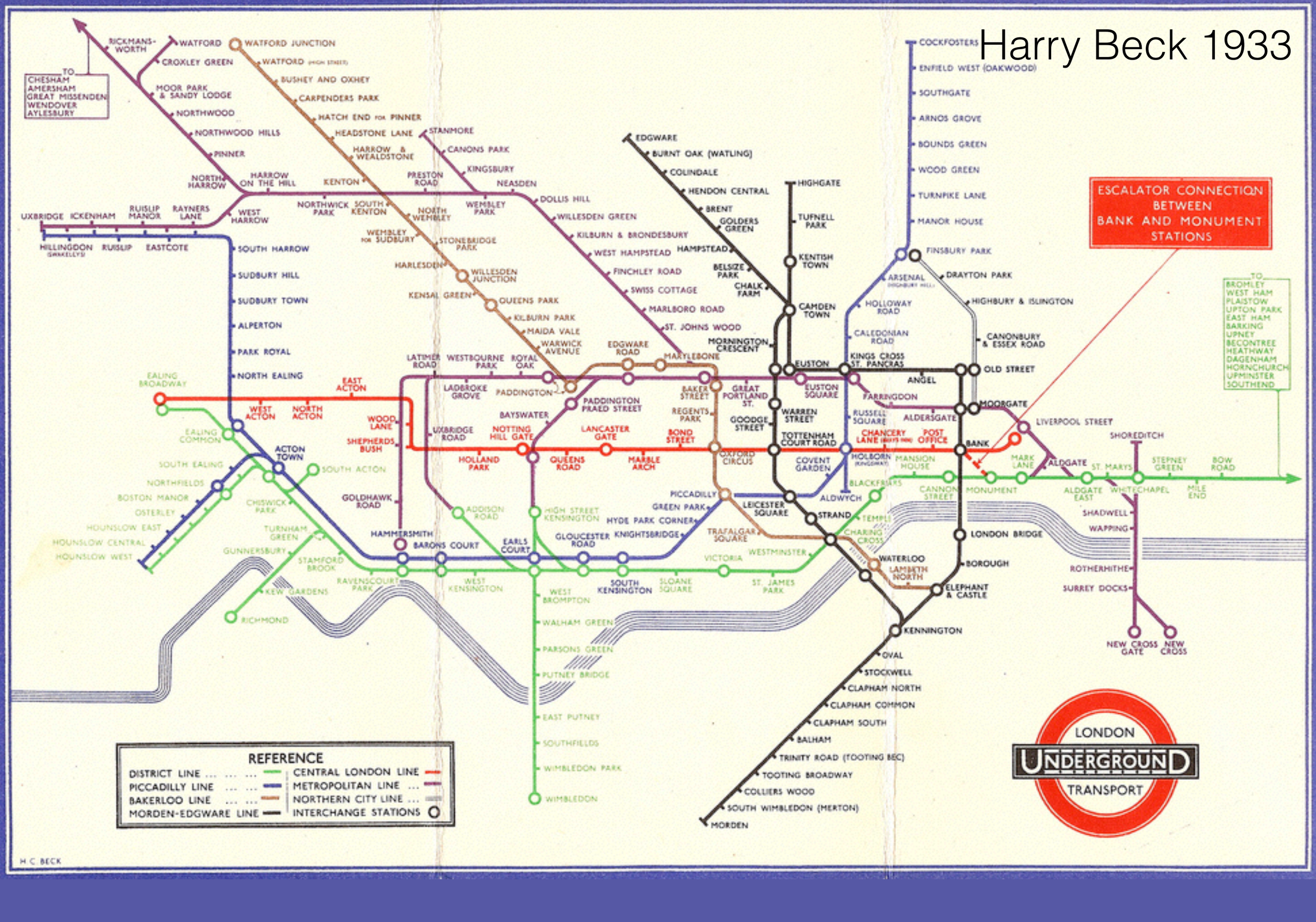
- District Railway
- London Electric Railway: Bakerloo Line
- London Electric Railway: Piccadilly Line
- London Electric Railway: Hampstead and Highgate Line
- City and South London Rly.
- Central London Railway
- Metropolitan Railway
- Metropolitan Railway (Great Northern and City Section)
- Interchange Stations

FRS

WATERLOW & SONS LIMITED LONDON, DUNSTABLE & WATFORD

(JANUARY 1927)

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

BRONX

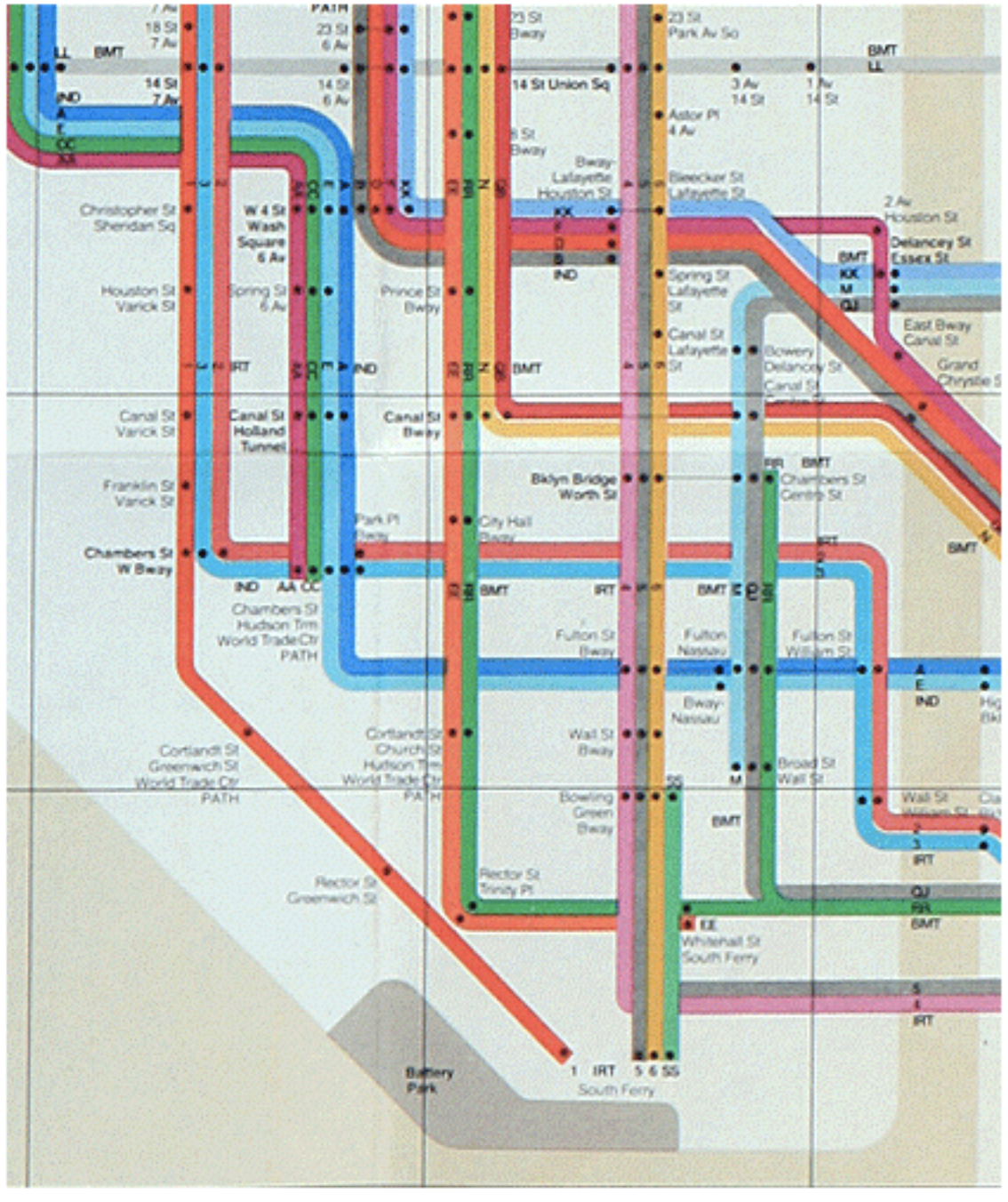
BROOKLYN

MANHATTAN

QUEENS

The New Map

1972: Vignelli's Classic



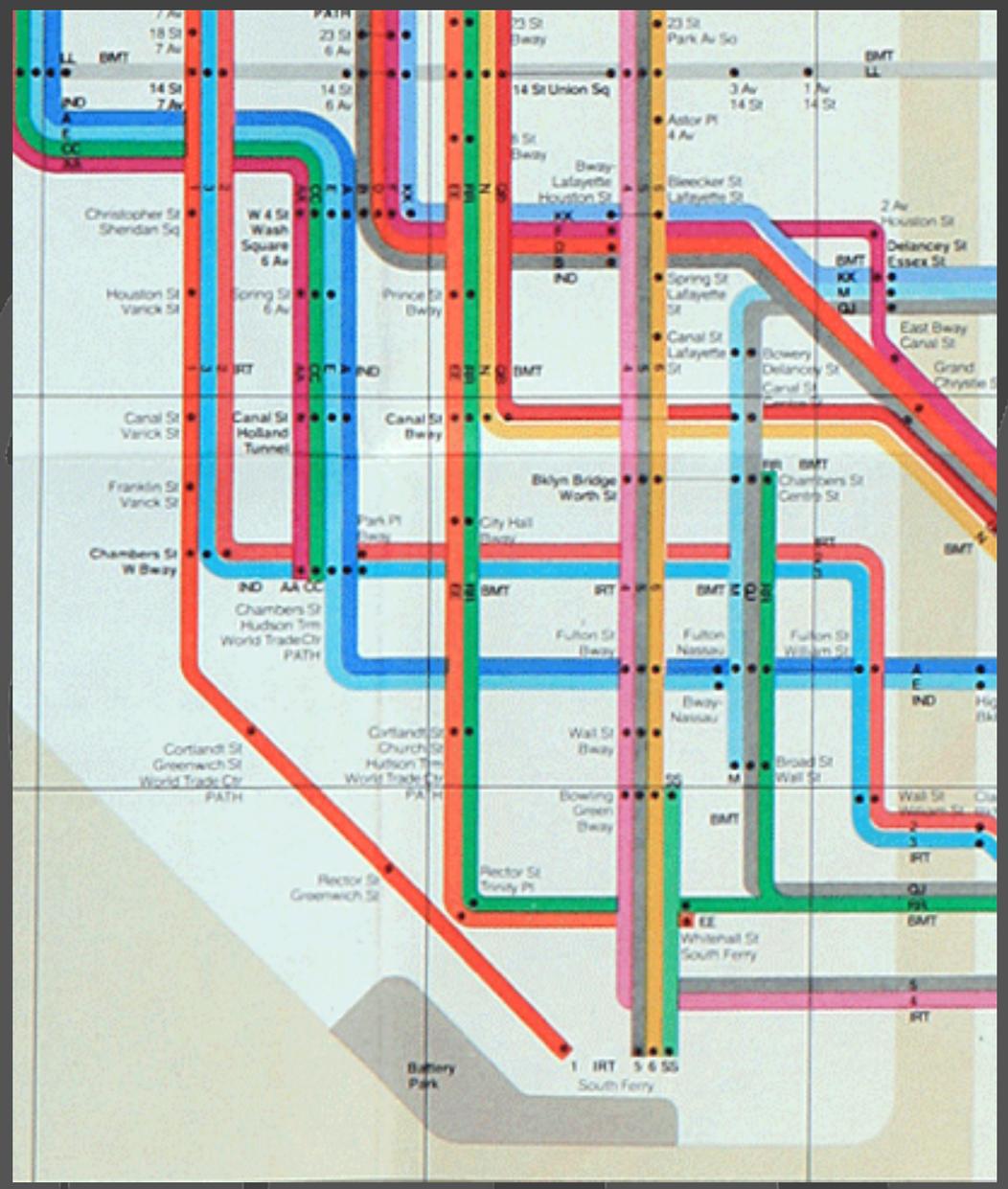
New York Times 2010

Activity

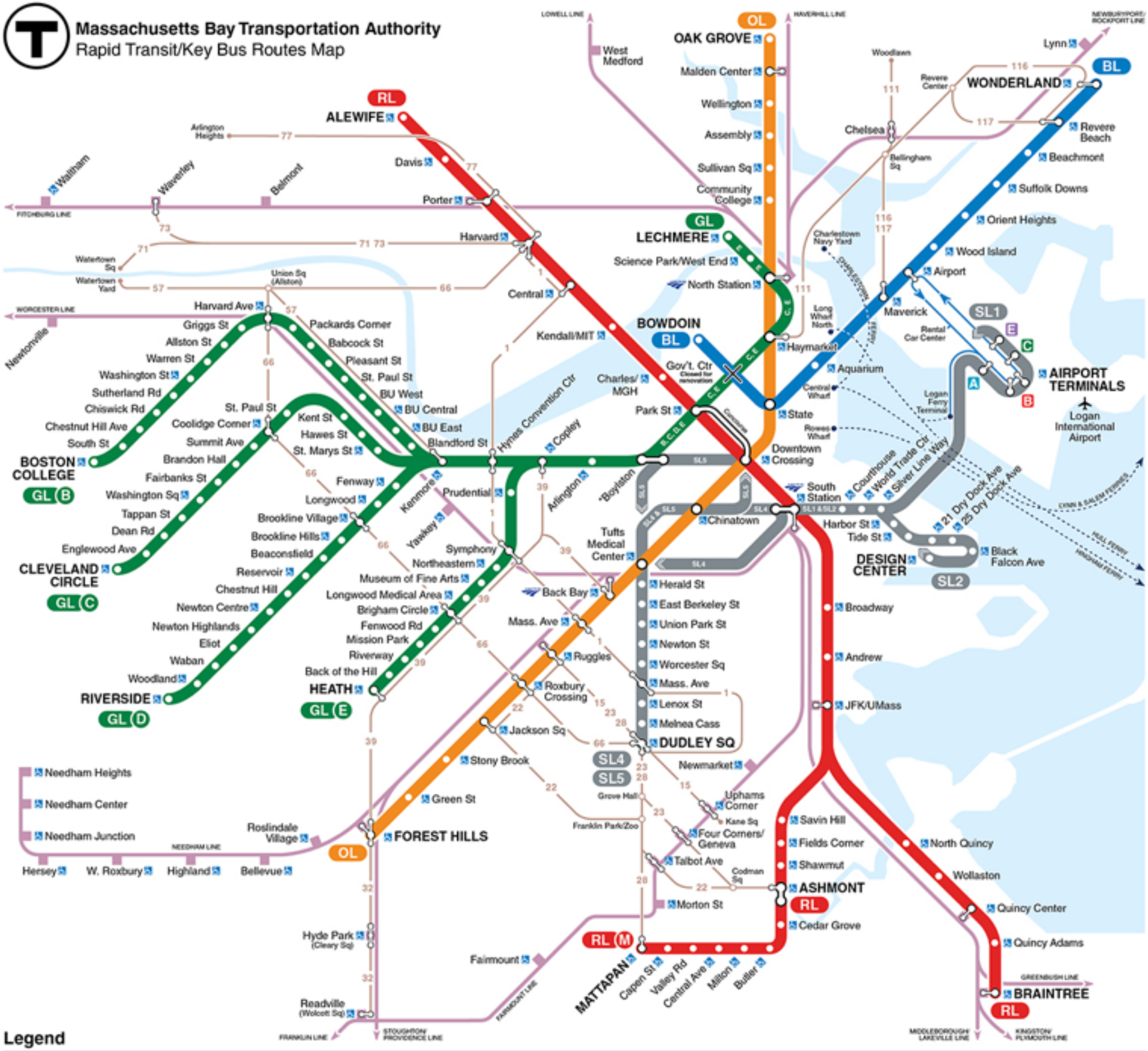
Which of these two subway maps is better? Why? [2 min]

A

B



Massachusetts Bay Transportation Authority
Rapid Transit/Key Bus Routes Map



Legend

- RL RED LINE
- M MATTAPAN LINE
- OL ORANGE LINE
- BL BLUE LINE
- SL SILVER LINE and branches (SL1 SL2 SL3 SL4 SL5)
- GL GREEN LINE and branches
  - B Terminates at Park St
  - C Terminates at N. Station
  - D Terminates at Park St (N. Station during off-peak)
  - E Terminates at Lechmere
- COMMUTER RAIL
- 000 KEY BUS ROUTE Frequent Service
- FERRYS
- Accessible station — All MBTA and Massport bus and ferry services are accessible
- Rapid Transit transfer station
- Commuter Rail transfer station
- Free Logan Airport shuttle bus
- Amtrak service — Back Bay, North & South Stations
- *Boylston: Accessible for Silver Line only

Customer Communications & Travel Info
617-222-3200, 1-800-392-6100,
TTY 617-222-5146, www.mbta.com

MBTA Transit Police: **911**
TTY 617-222-1200

Elevator/escalator/lift updates: 800-392-6100

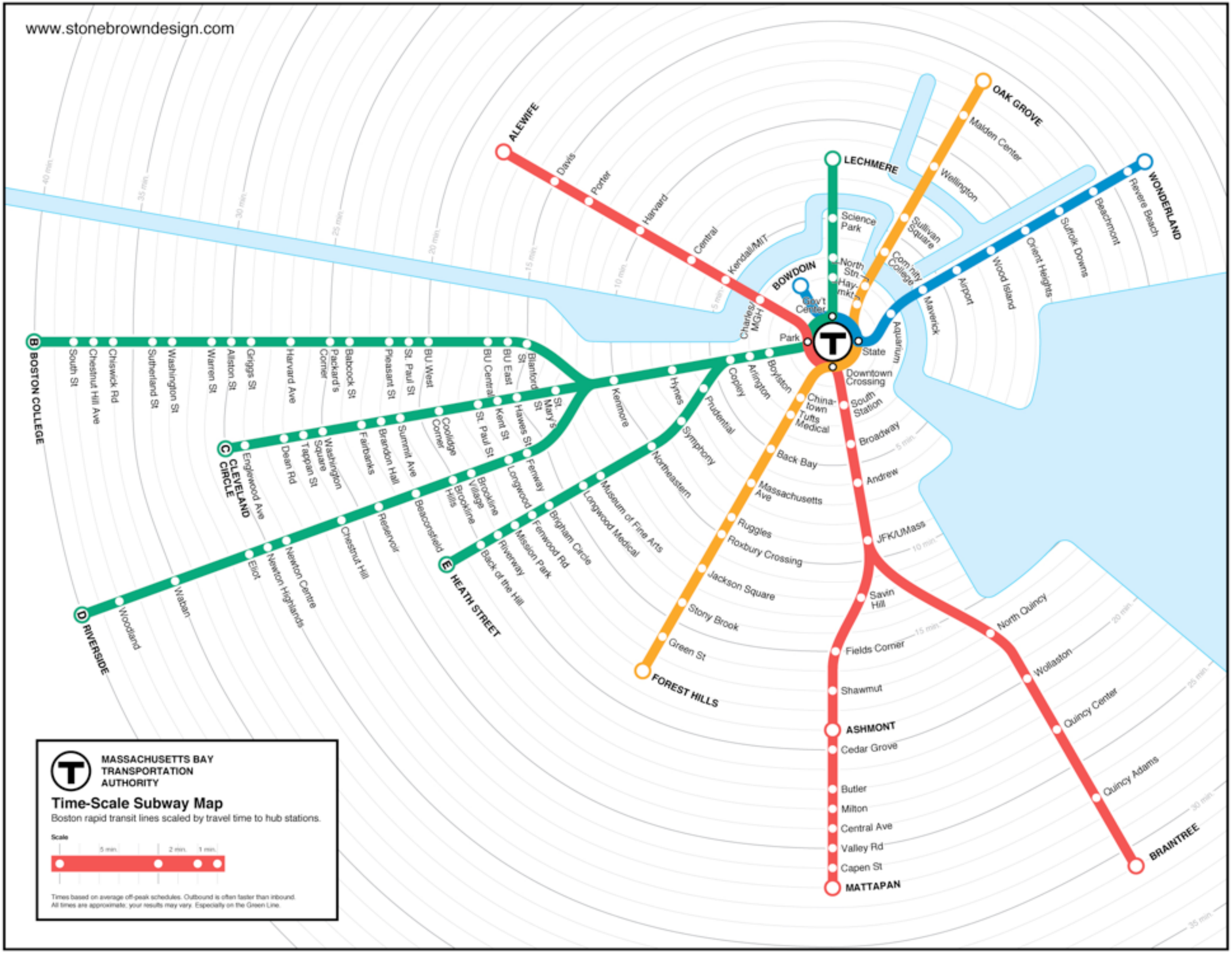
Not to scale

Geographic vs Topological Metro Map

FATHOM, 2013

Terrence Fradet 2013

www.stonebrowndesign.com



MASSACHUSETTS BAY TRANSPORTATION AUTHORITY

Time-Scale Subway Map

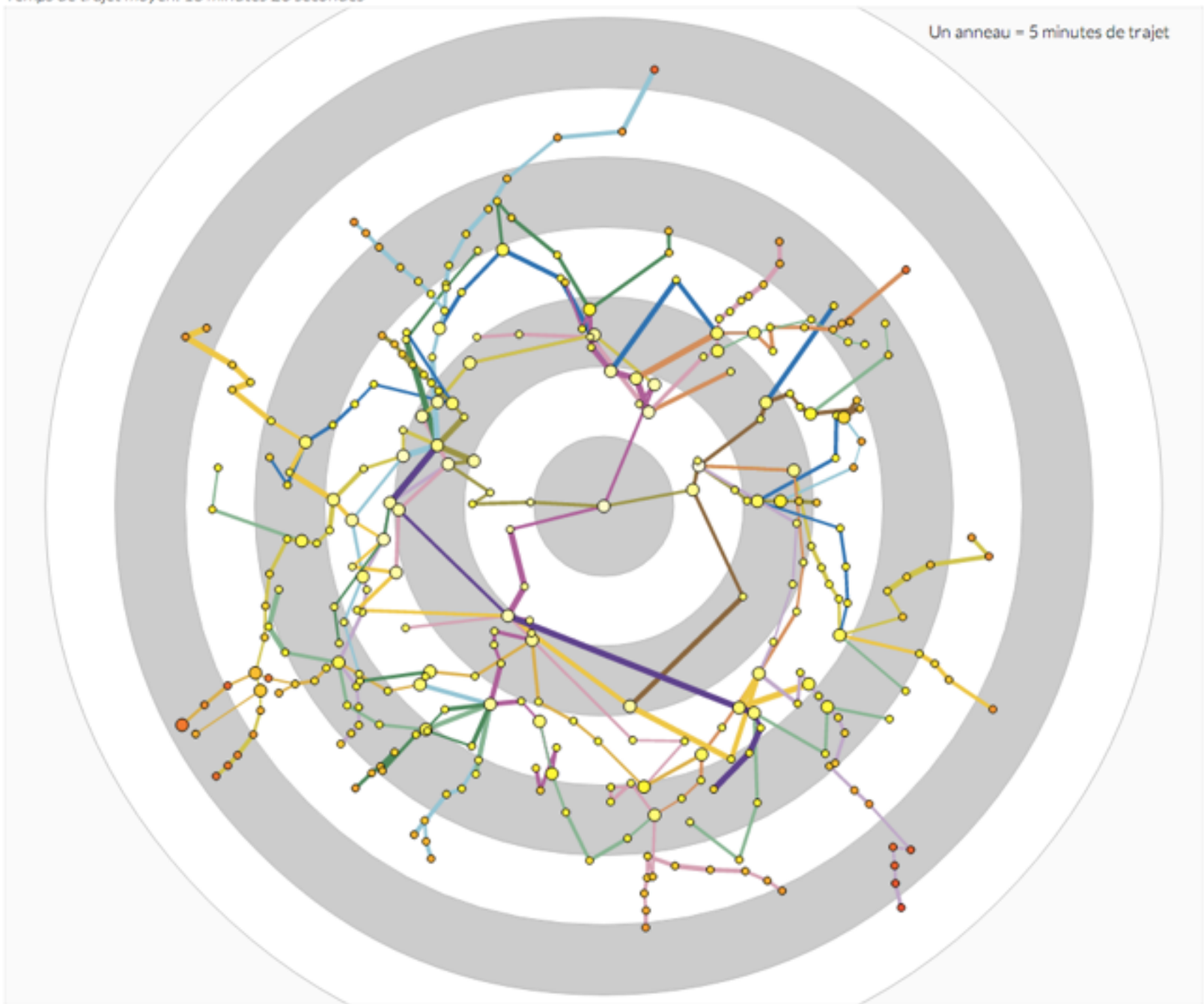
Boston rapid transit lines scaled by travel time to hub stations.

Scale

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 ?

Jerome Cukier, D3
Writeup About the Map

Outline

- Who?
- How?
- What?
- Why?

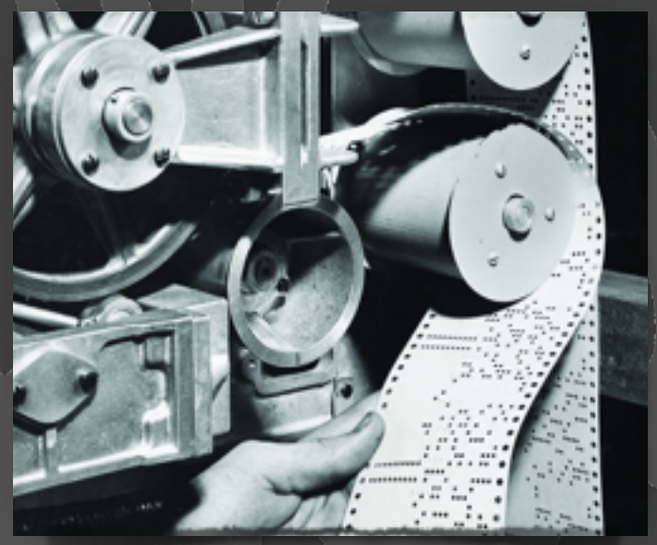
INDUSTRIAL REVOLUTION OF DATA

Joe Hellerstein, UC Berkley, 2008

Activity

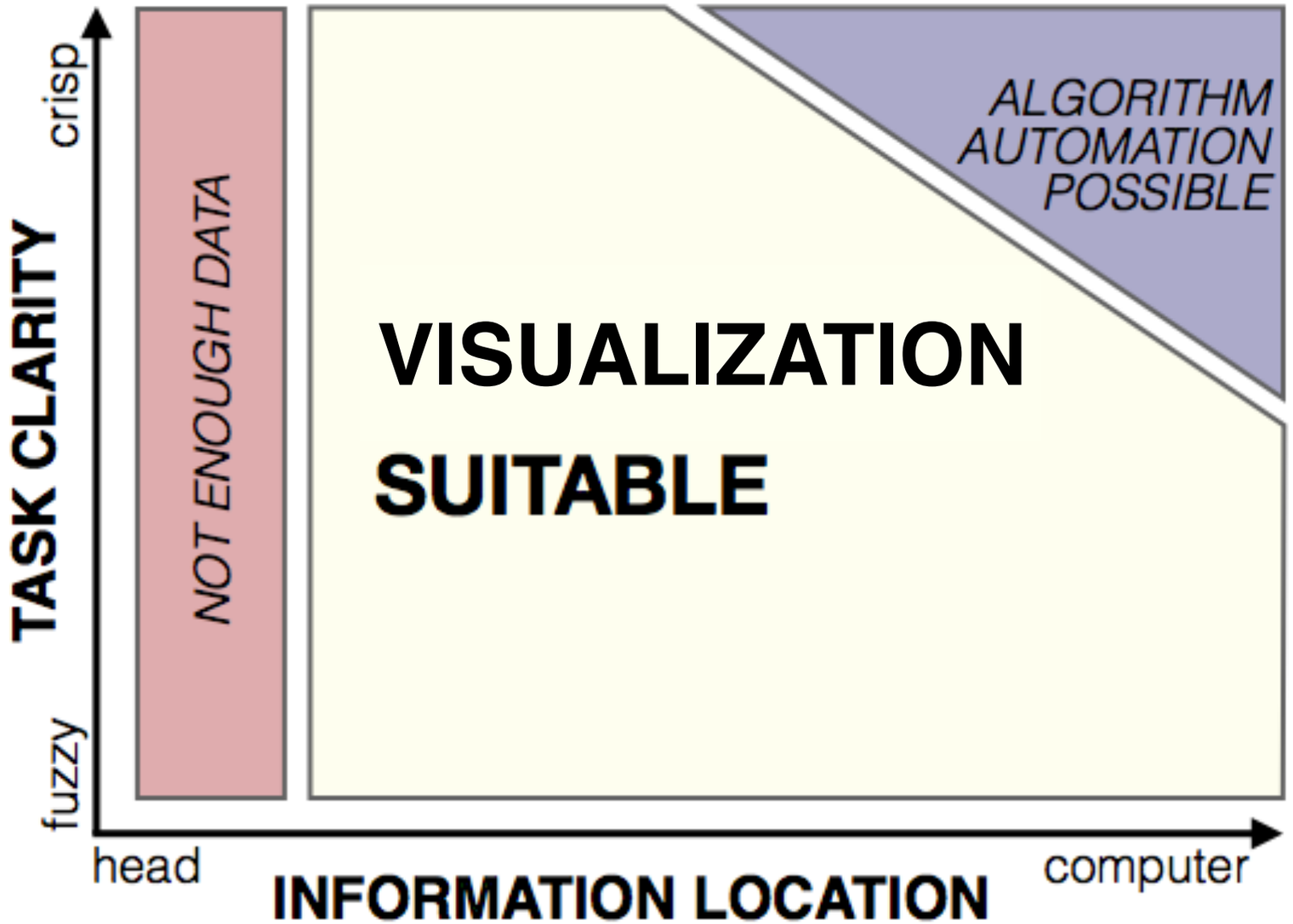
Imagine a system to analyze large amounts of data. Collect and discuss arguments answering the questions:

Why would you have a human in the loop?



Why would you have a computer in the loop?

Is visualization the appropriate method to solve the problem?



Design Study Methodology: Reflections from the Trenches and the Stacks, Michael Sedlmair, Miriah Meyer, and Tamara Munzner. IEEE Trans. Visualization and Computer Graphics, 2012.

Ascombe's Quartet

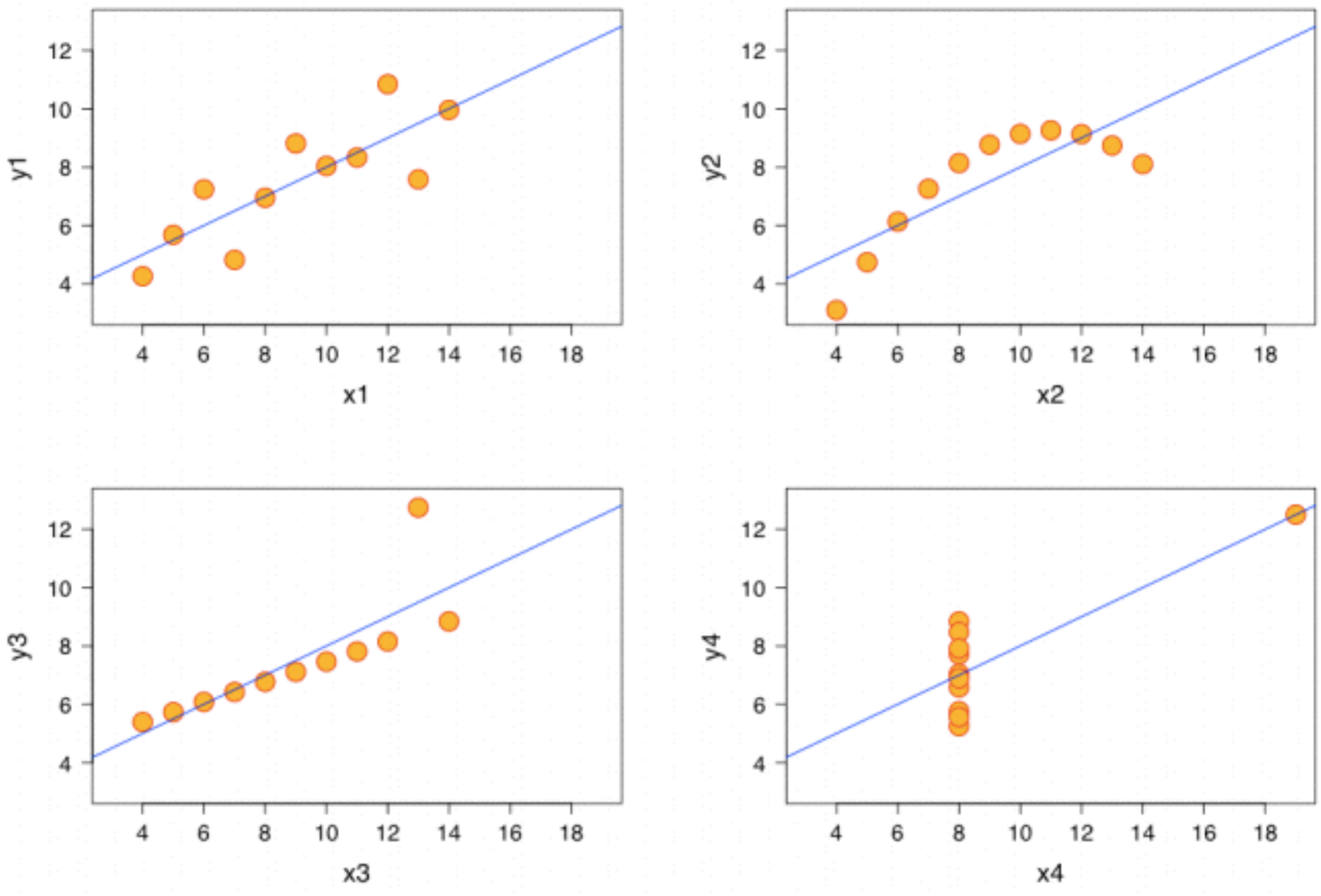
Anscombe's Quartet: Raw Data

I		II		III		IV	
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	6.89
mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
var.	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
corr.	0.816		0.816		0.816		0.816	

Anscombe '73

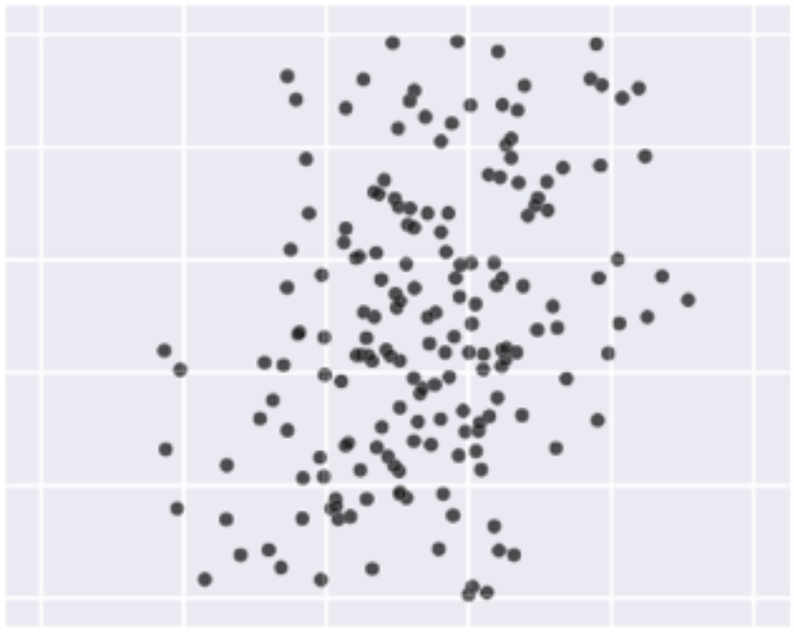
Ascombe’s Quartet

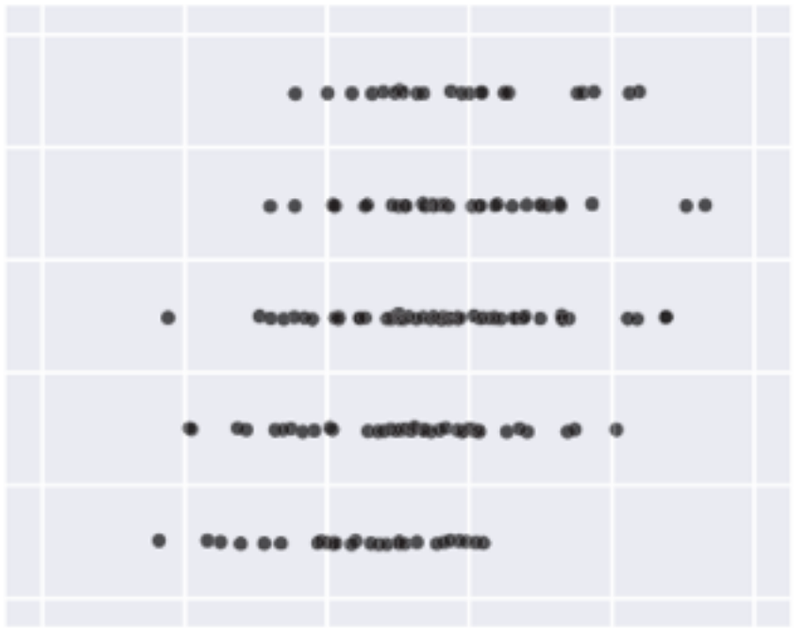
Same mean, variance, correlation, and linear regression line

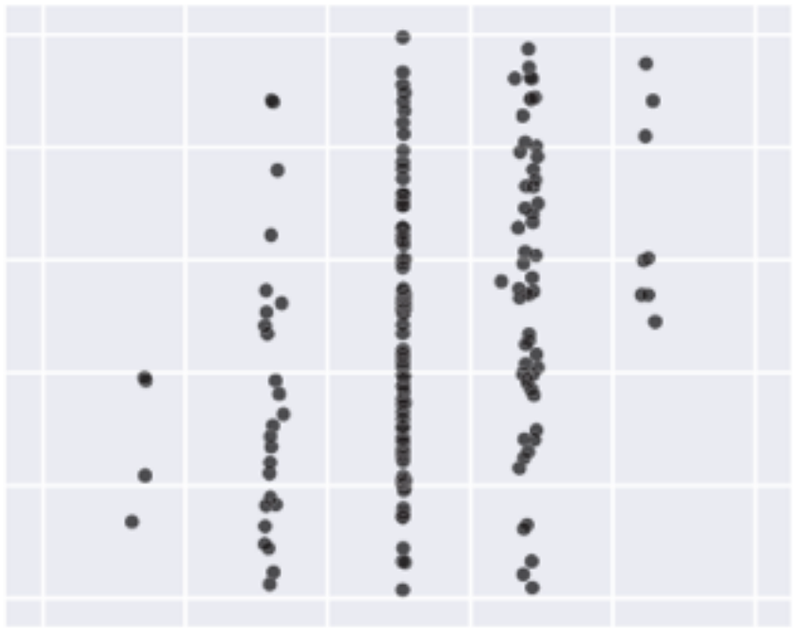


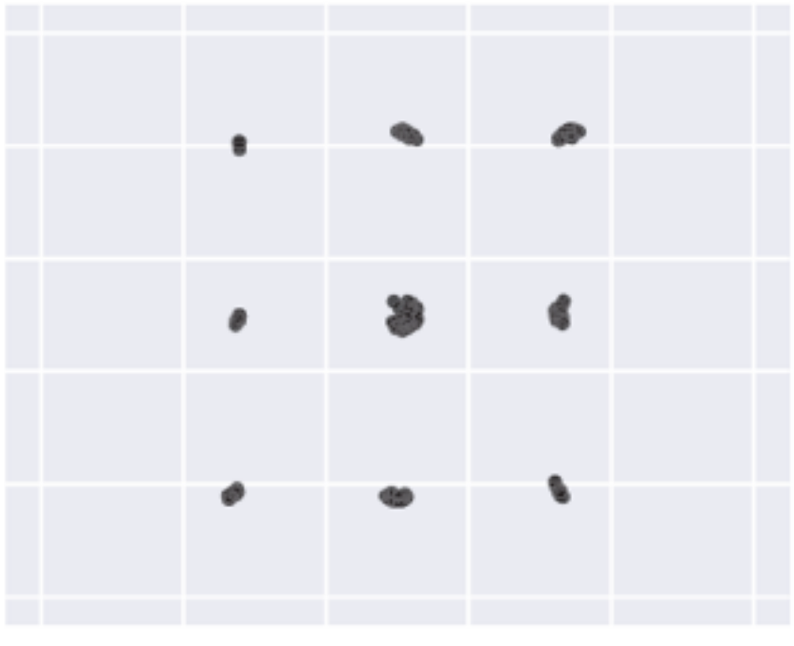
Anscombe ’73

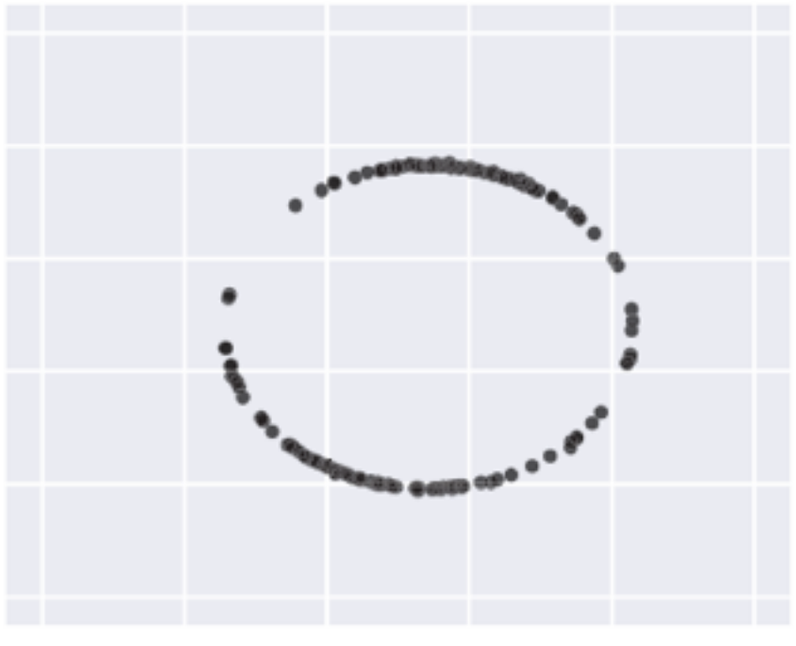
Same Stats, Different Graphs

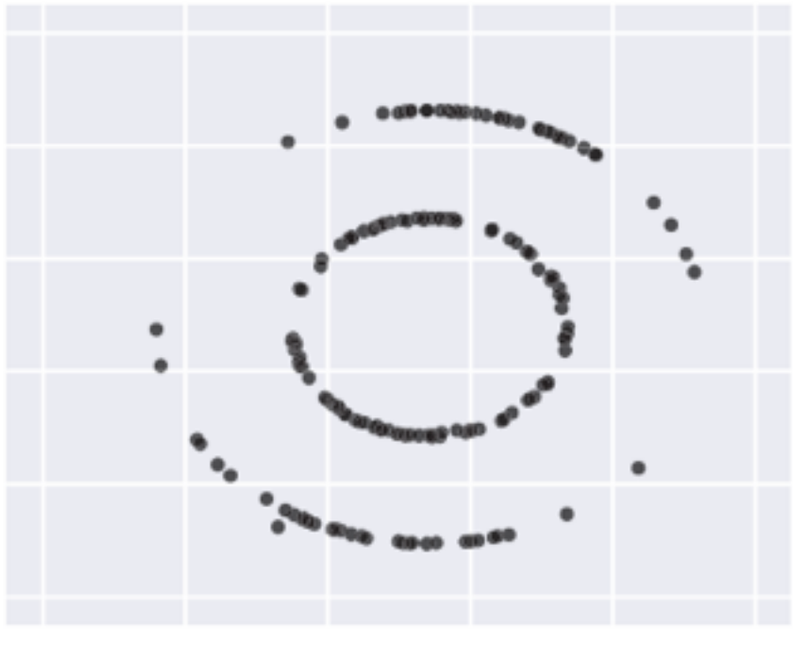












Daniel Simons 1998

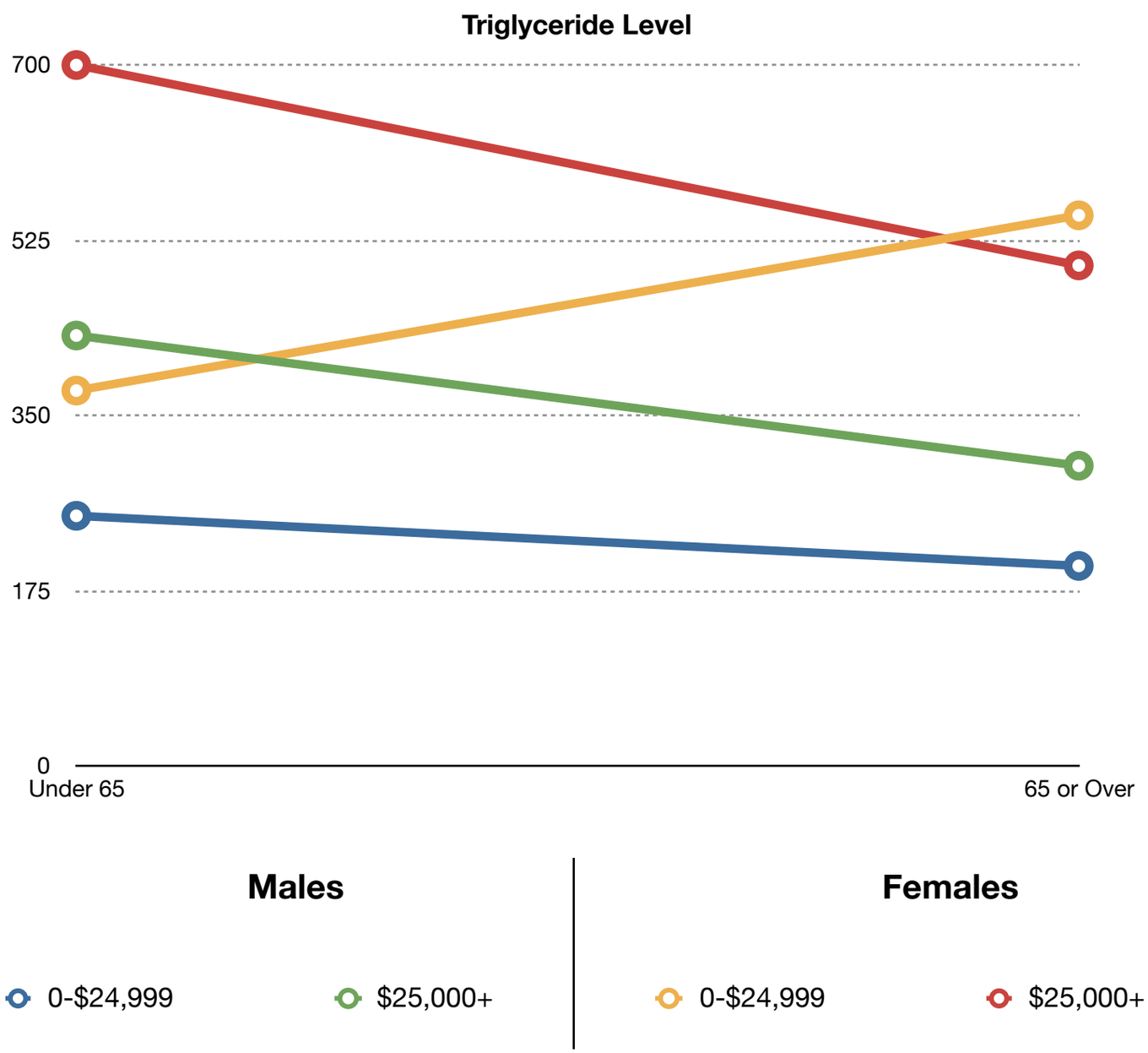
The power of graphical representations of data

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

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

Stephen Kosslyn “Clear and to the Point”

The power of graphical representations of data



Stephen Kosslyn “Clear and to the Point”

The power of perception

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPSNDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

The power of perception

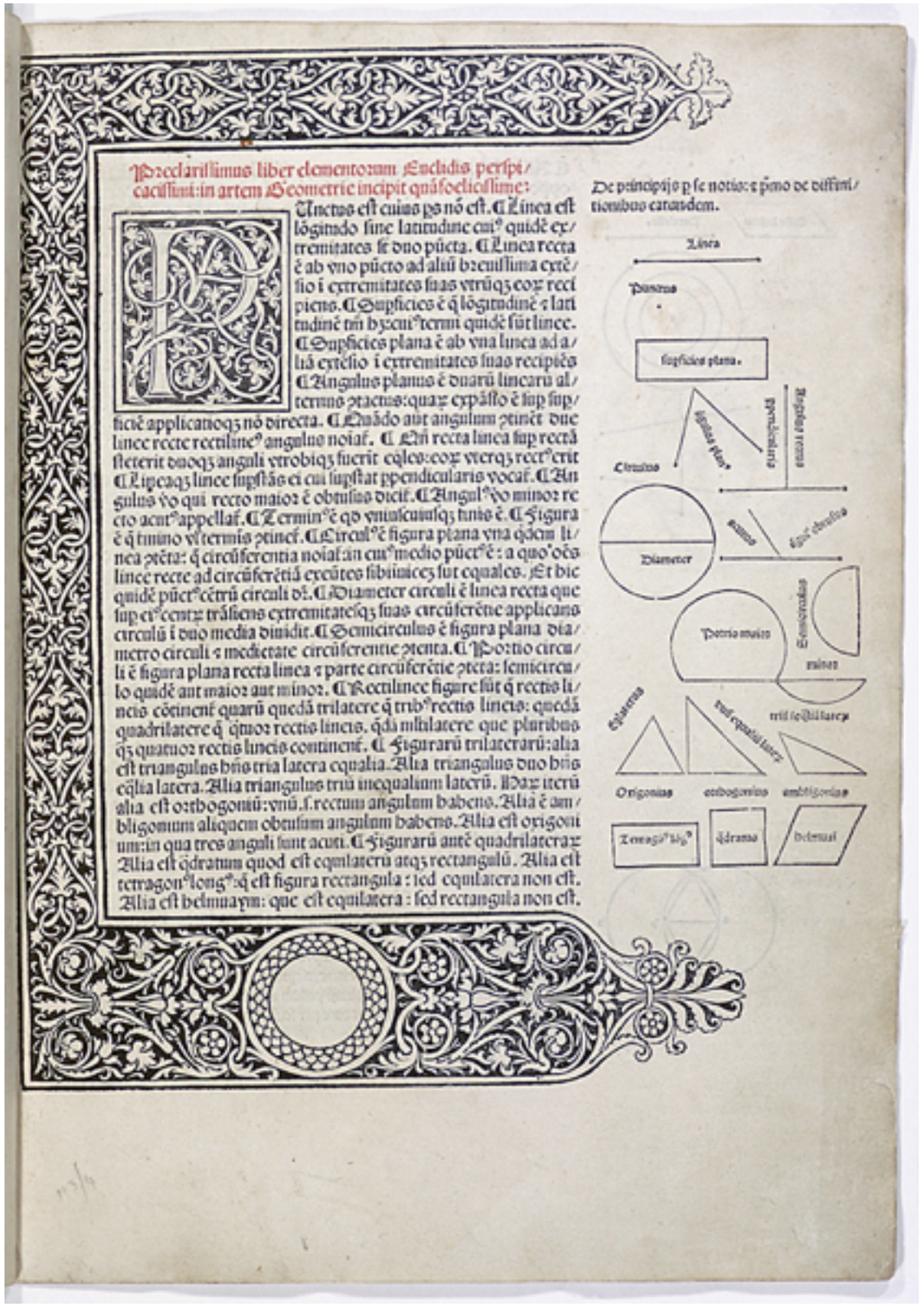
MTHI**V**LWYADCEQGHKILKMTWYN

ARDCAIREQGHL**V**KMFPSTWYARN

GFPS**V**CEILQGKMFPSNDRCEQDIFPS

GHLMFHKM**V**PSTWYACEQTWRN

Extending our cognitive powers



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

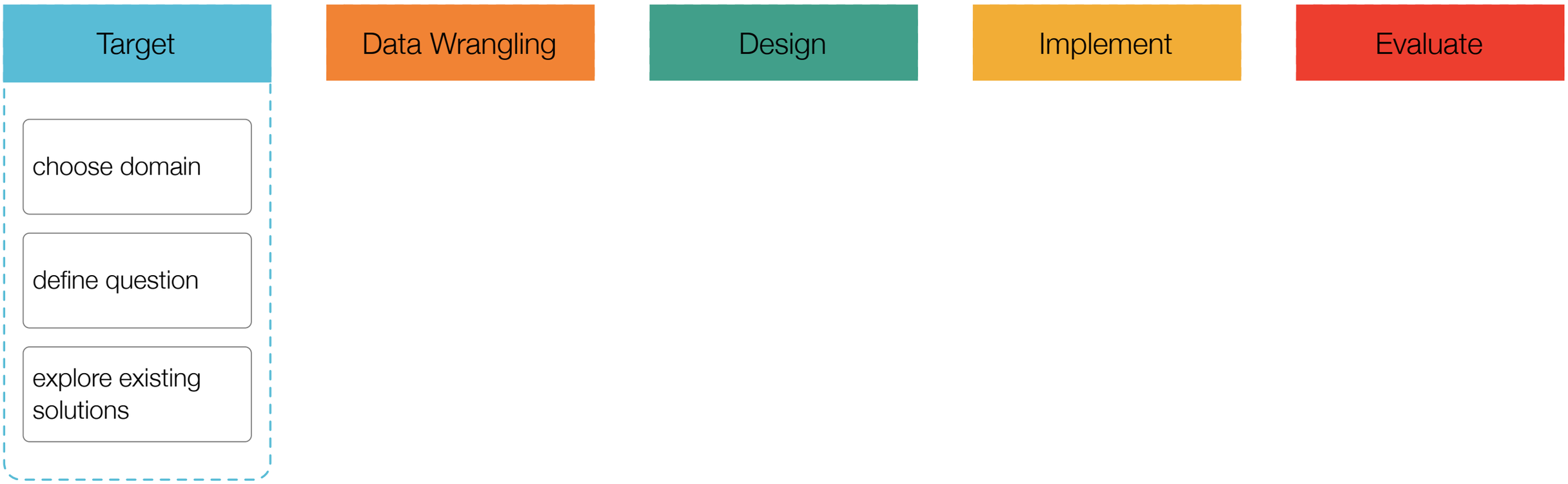
–Stuart Card

Overview of the Visualization Process

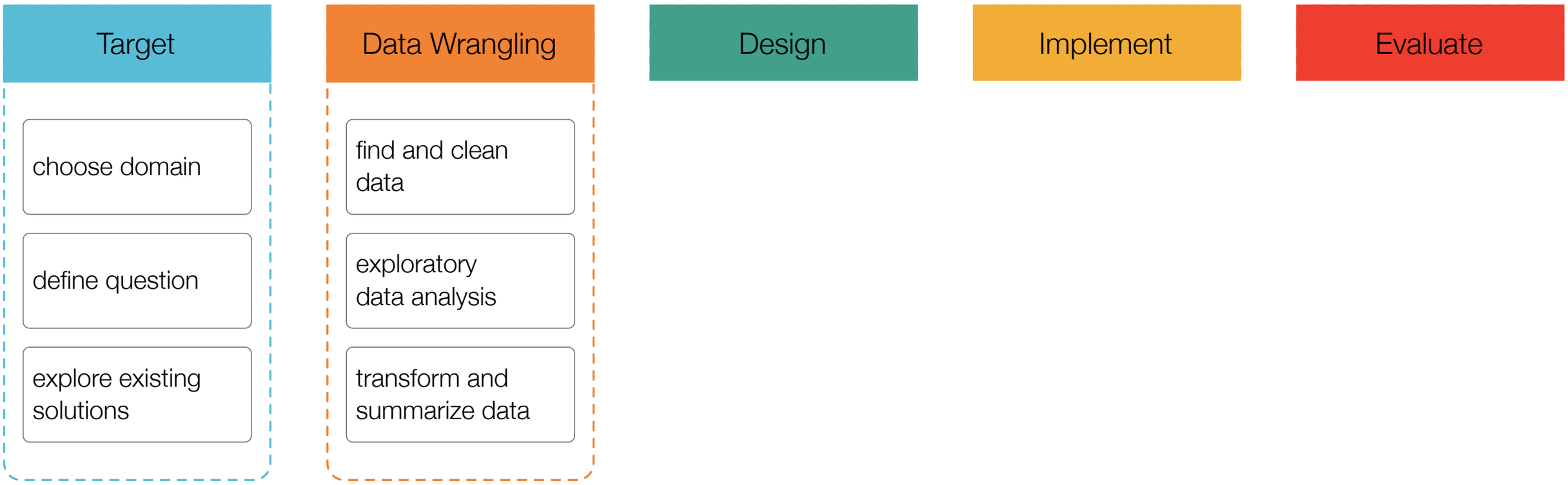
5-Step Model

Target | Data Wrangling | Design | Implement | Evaluate

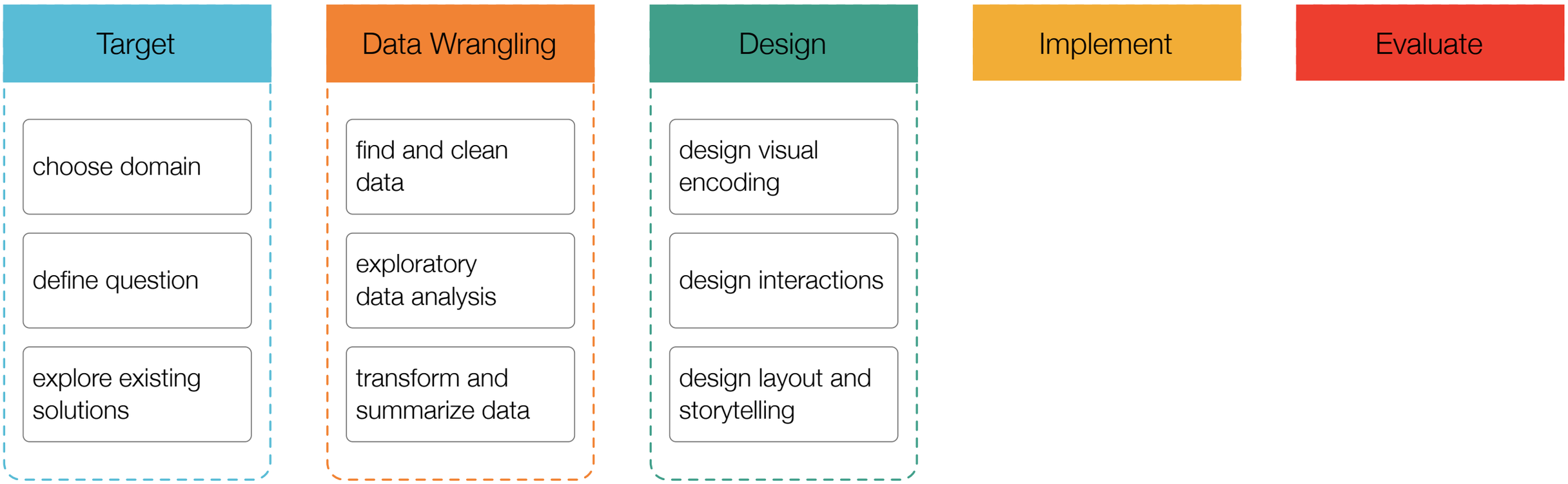
Our Model


Target
Data Wrangling
Design
Implement
Evaluate
choose domain
define question
explore existing solutions

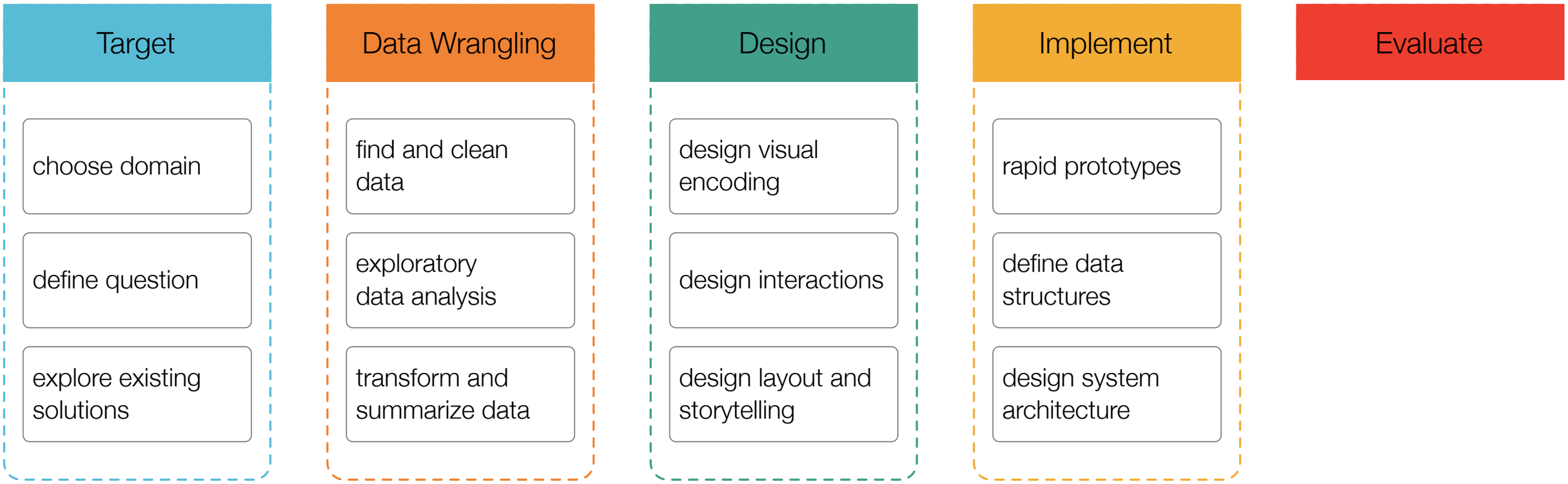
Our Model


Target
Data Wrangling
Design
Implement
Evaluate
choose domain
define question
explore existing solutions
find and clean data
exploratory data analysis
transform and summarize data

Our Model


Target
choose domain
define question
explore existing solutions
Data Wrangling
find and clean data
exploratory data analysis
transform and summarize data
Design
design visual encoding
design interactions
design layout and storytelling
Implement
Evaluate

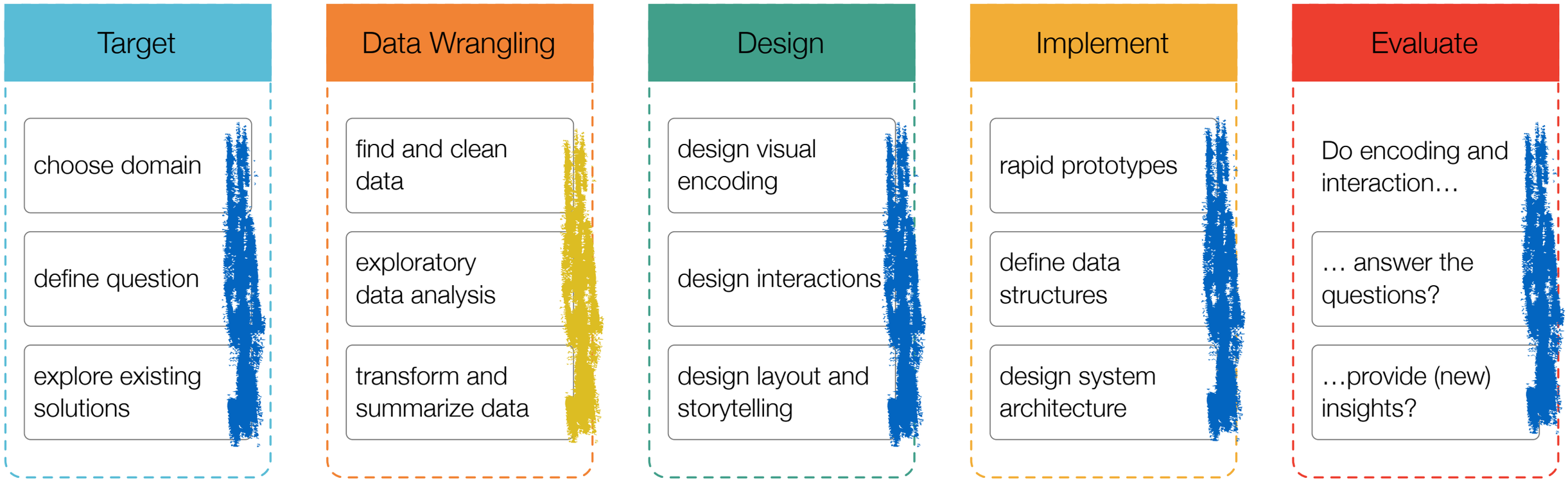
Our Model


Target
choose domain
define question
explore existing solutions
Data Wrangling
find and clean data
exploratory data analysis
transform and summarize data
Design
design visual encoding
design interactions
design layout and storytelling
Implement
rapid prototypes
define data structures
design system architecture
Evaluate

Our Model

Target	Data Wrangling	Design	Implement	Evaluate
choose domain	find and clean data	design visual encoding	rapid prototypes	Do encoding and interaction…
define question	exploratory data analysis	design interactions	define data structures	… answer the questions?
explore existing solutions	transform and summarize data	design layout and storytelling	design system architecture	…provide (new) insights?

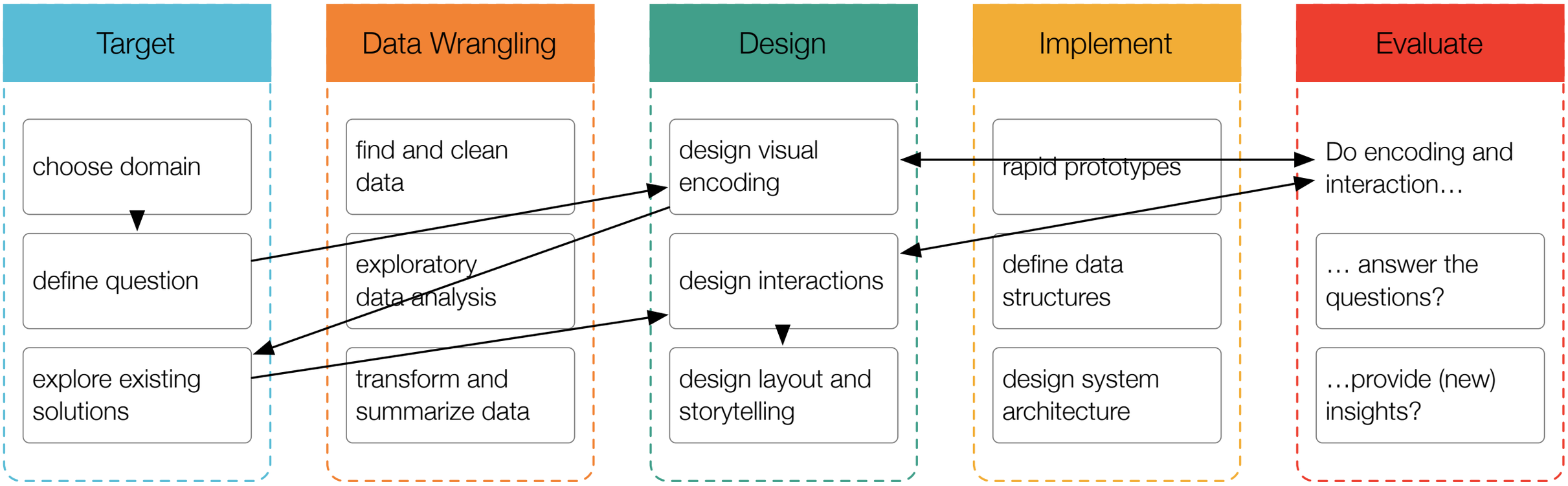
This Course



CS109 (www.cs109.org)

CS171 (www.cs171.org)

In Reality…

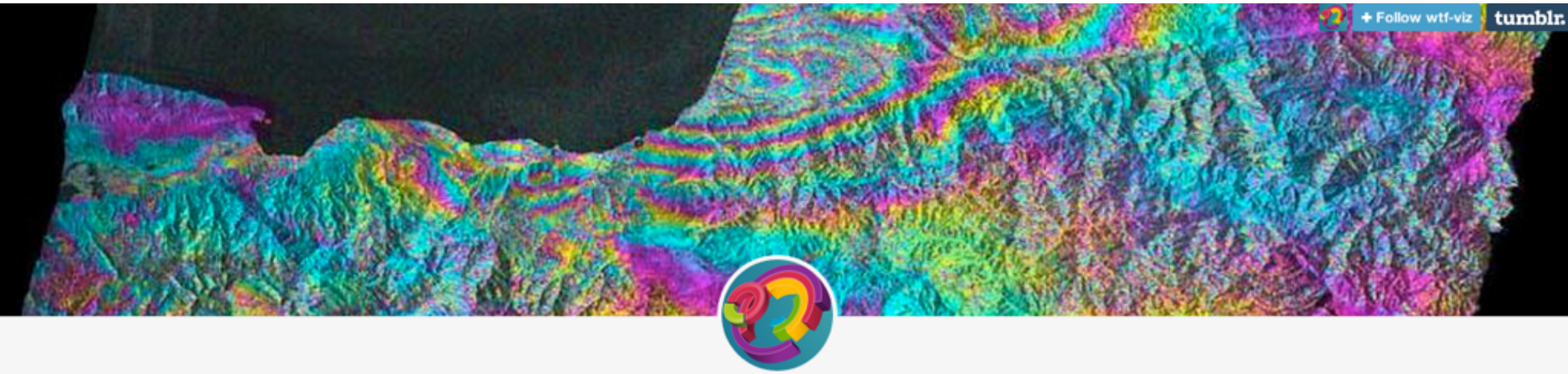

Target
Data Wrangling
Design
Implement
Evaluate
choose domain
define question
explore existing solutions
find and clean data
exploratory data analysis
transform and summarize data
design visual encoding
design interactions
design layout and storytelling
rapid prototypes
define data structures
design system architecture
Do encoding and interaction…
… answer the questions?
…provide (new) insights?

Design

“Design is a word that's come to mean so much that it's also a word that has come to mean nothing.”

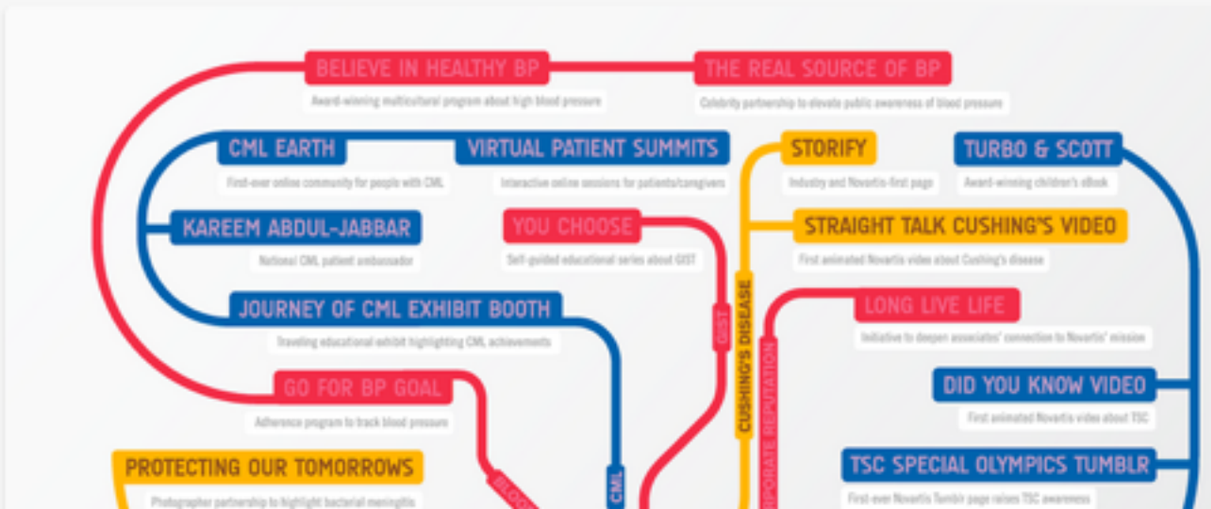
–Johathan Ive

http://wtfviz.net/



WTF Visualizations

Visualizations that make no sense.
For a discussion of what is wrong with a particular visualization, tweet at us @WTFViz.
Check out our friends Thumbs Up Viz and accidental aRt, or submit.

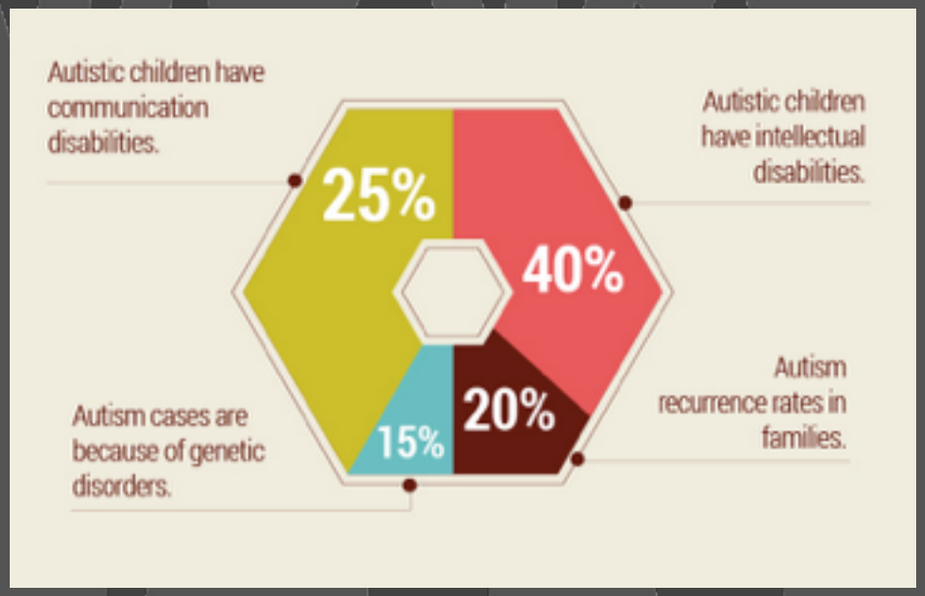


Activity

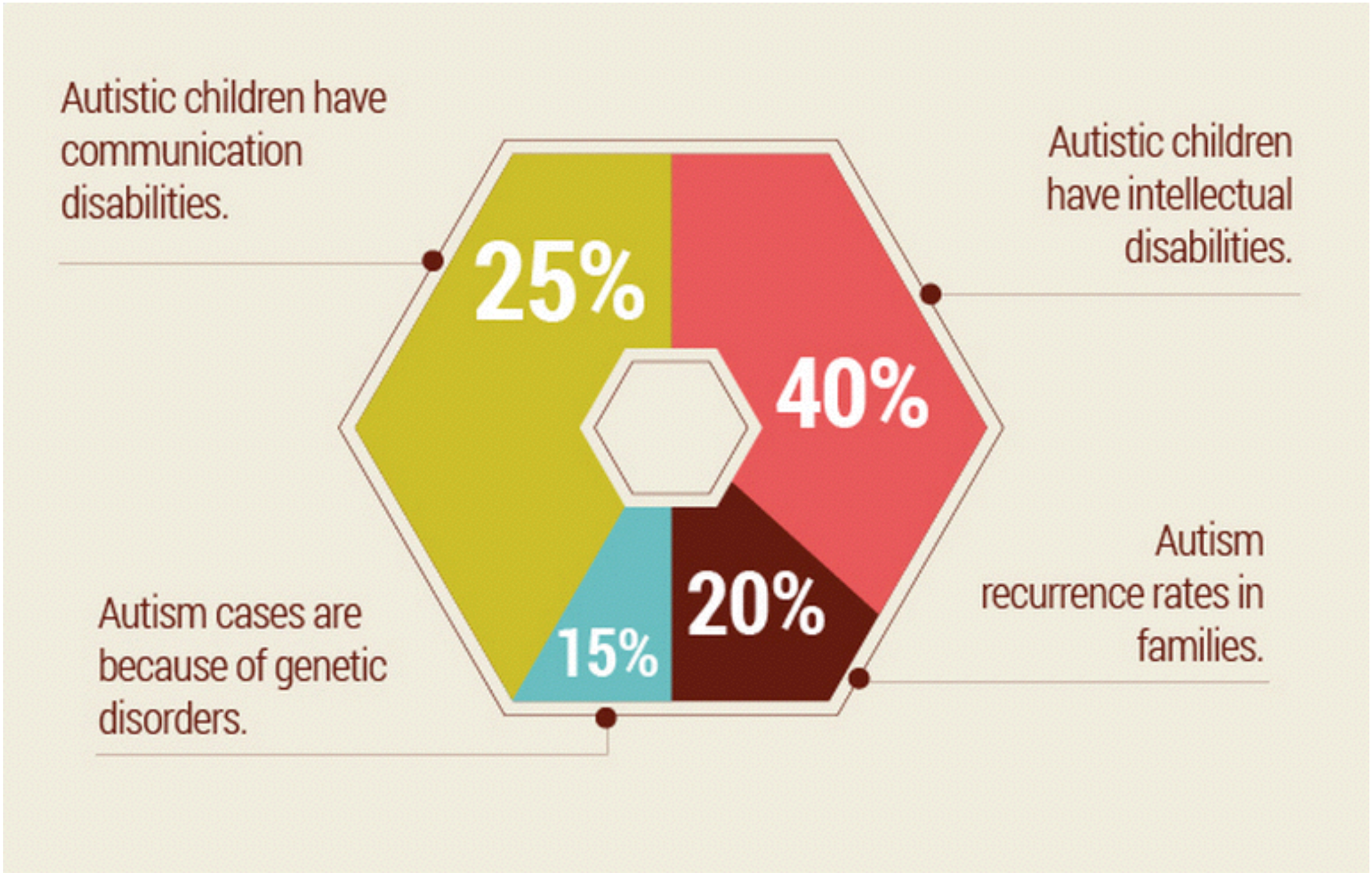
CS 171

Critique the following visualization answering the questions:

- What questions does this visualization answer?
- Why do you like / dislike about this visualization?



http://wtfviz.net/



http://wtfviz.net/

Please fill out the 1-minute paper!

Available on Canvas, open until midnight tonight!

Homework

- HW 0 due **Monday** including enrollment survey

Lab (next Tuesday)

- HTML / CSS / DOM
- **Reading and Pre-Quiz**: Murray, Chapter 3 (p. 17-36)

Lecture (next Thursday)

- Design & Design Principles
- **Reading and Pre-Quiz**: Ware, Chapter 8