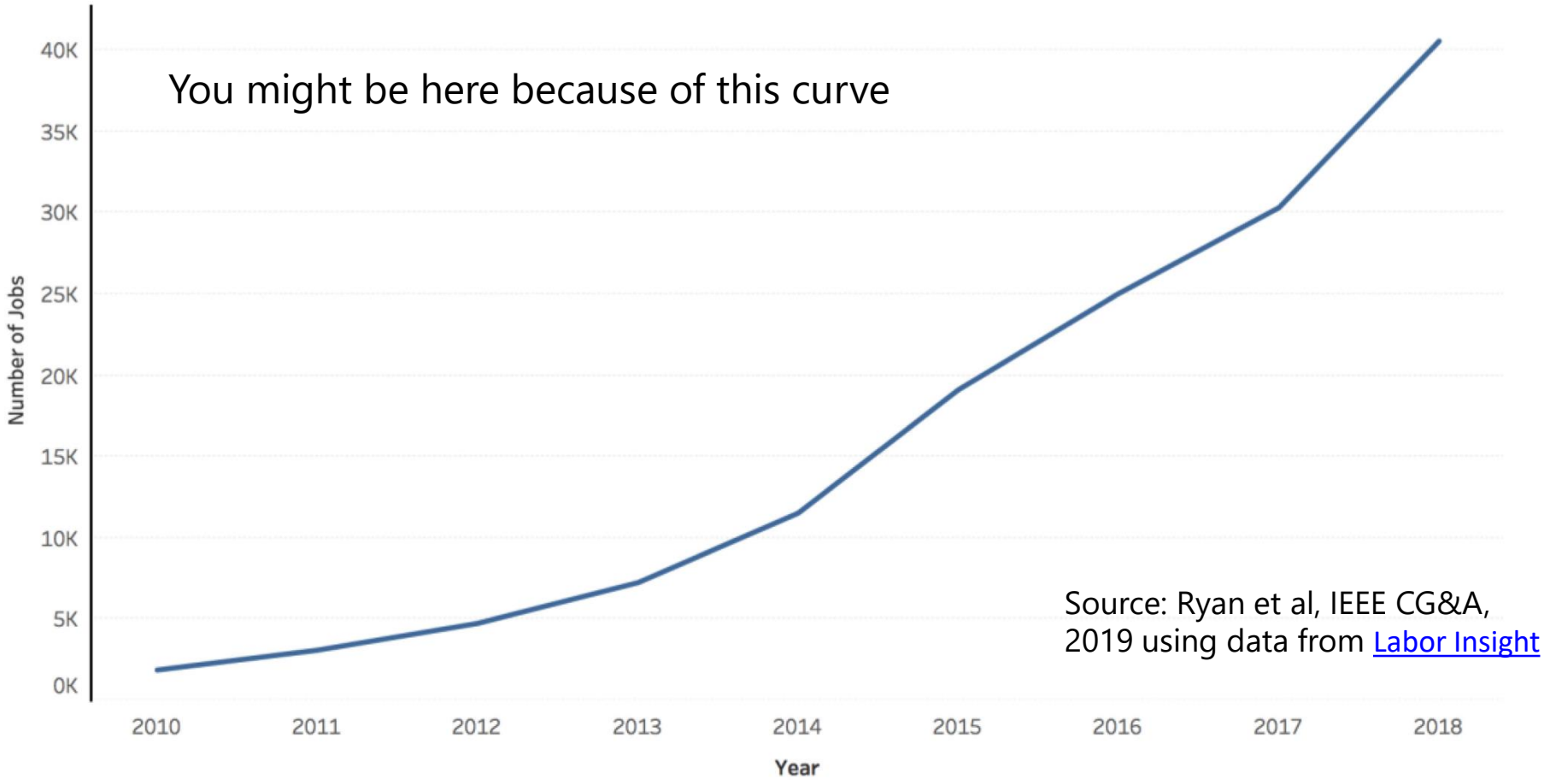
CSE 564
Visualization & Visual Analytics

Introduction

Klaus Mueller

Computer Science Department
Stony Brook University

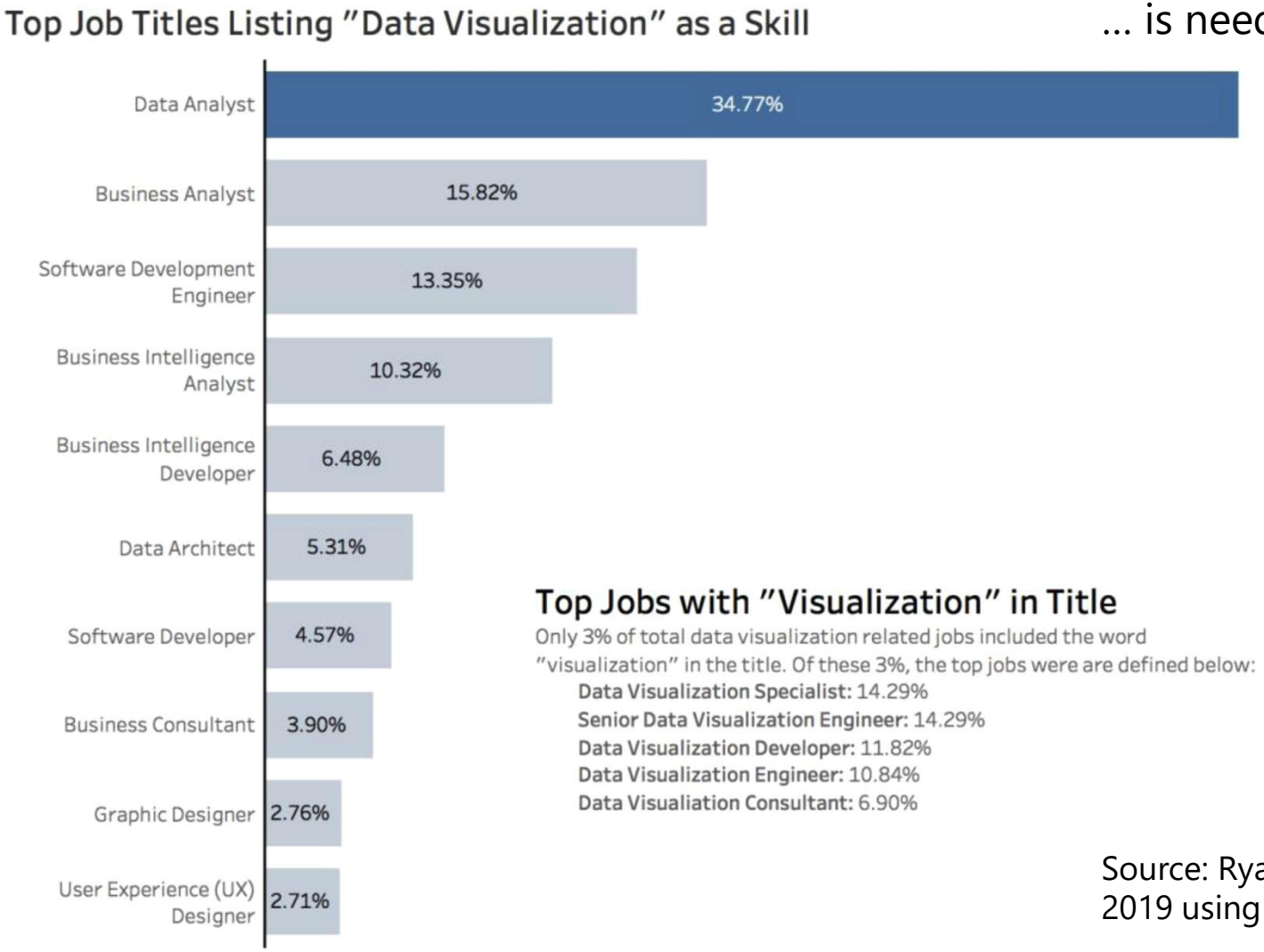
Why Are You Here?



The growth of jobs mentioning “data visualization” as a skill from 2010 through 2017 has steadily increased from only 1,888 jobs in 2010 to 30,327 jobs in 2017 (16×growth)

"Visualization" Skill…

… is needed everywhere



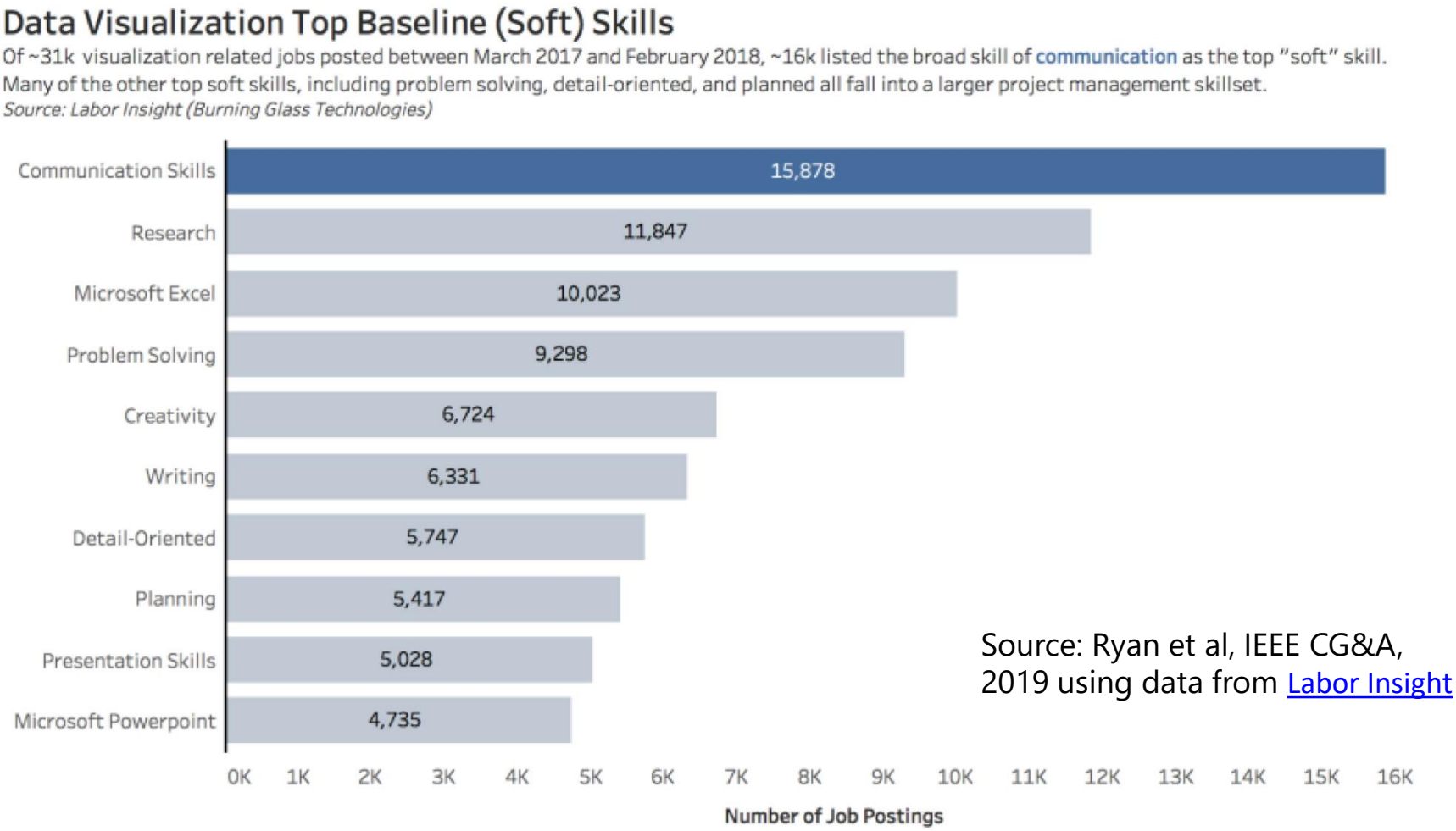
Top Jobs with "Visualization" in Title

Only 3% of total data visualization related jobs included the word "visualization" in the title. Of these 3%, the top jobs were are defined below:

- **Data Visualization Specialist:** 14.29%
- **Senior Data Visualization Engineer:** 14.29%
- **Data Visualization Developer:** 11.82%
- **Data Visualization Engineer:** 10.84%
- **Data Visualiation Consultant:** 6.90%

Source: Ryan et al, IEEE CG&A, 2019 using data from Labor Insight

What Other Skills?



Source: Ryan et al, IEEE CG&A, 2019 using data from [Labor Insight]

Baseline, or “soft” skills listed for these 30k “Data Visualization” jobs.

Skills, Reading Between The Lines

Communication, when mentioned in conjunction with data visualization really means:

- communication of information derived from data
- *visual* story telling with data
- half of the data analytics projects fail due to poor communication
  (according to L. Kart, N. Neudecker, F. Buytendijk, Gartner Report GG0255160, 2013)

Apart from the specialized skills, these general skills (or proficiencies) are also often listed:

- SQL
- Tableau (41%),
- Excel (34%), PowerPoint (16%)
- Python (30%), SAS (22%), R (16%), Plotly (?%)
- JavaScript & JavaScript-based data-driven documents D3.js (13%)

Source: Ryan et al, IEEE CG&A, 2019 using data from Labor Insight

Skills, Reading Between The Lines

Communication, when mentioned in conjunction with data visualization really means:

- communication of information derived from data
- *visual* story telling with data
- half of the data analytics projects fail due to poor communication
  (according to L. Kart, N. Neudecker, F. Buytendijk, Gartner Report GG0255160, 2013)

Apart from the specialized skills, these general skills (or proficiencies) are also often listed:

- SQL
- Tableau (41%),
- Excel (34%), PowerPoint (16%)
- Python (30%), SAS (22%), R (16%), Plotly (?%)
- JavaScript & JavaScript-based data-driven documents D3.js (13%)

Source: Ryan et al, IEEE CG&A, 2019 using data from Labor Insight

Skills, Reading Between The Lines

Communication, when mentioned in conjunction with data visualization really means:

- communication of information derived from data
- *visual* story telling with data
- half of the data analytics projects fail due to poor communication
  (according to L. Kart, N. Neudecker, F. Buytendijk, Gartner Report GG0255160, 2013)

Apart from the specialized skills, these general skills (or proficiencies) are also often listed:

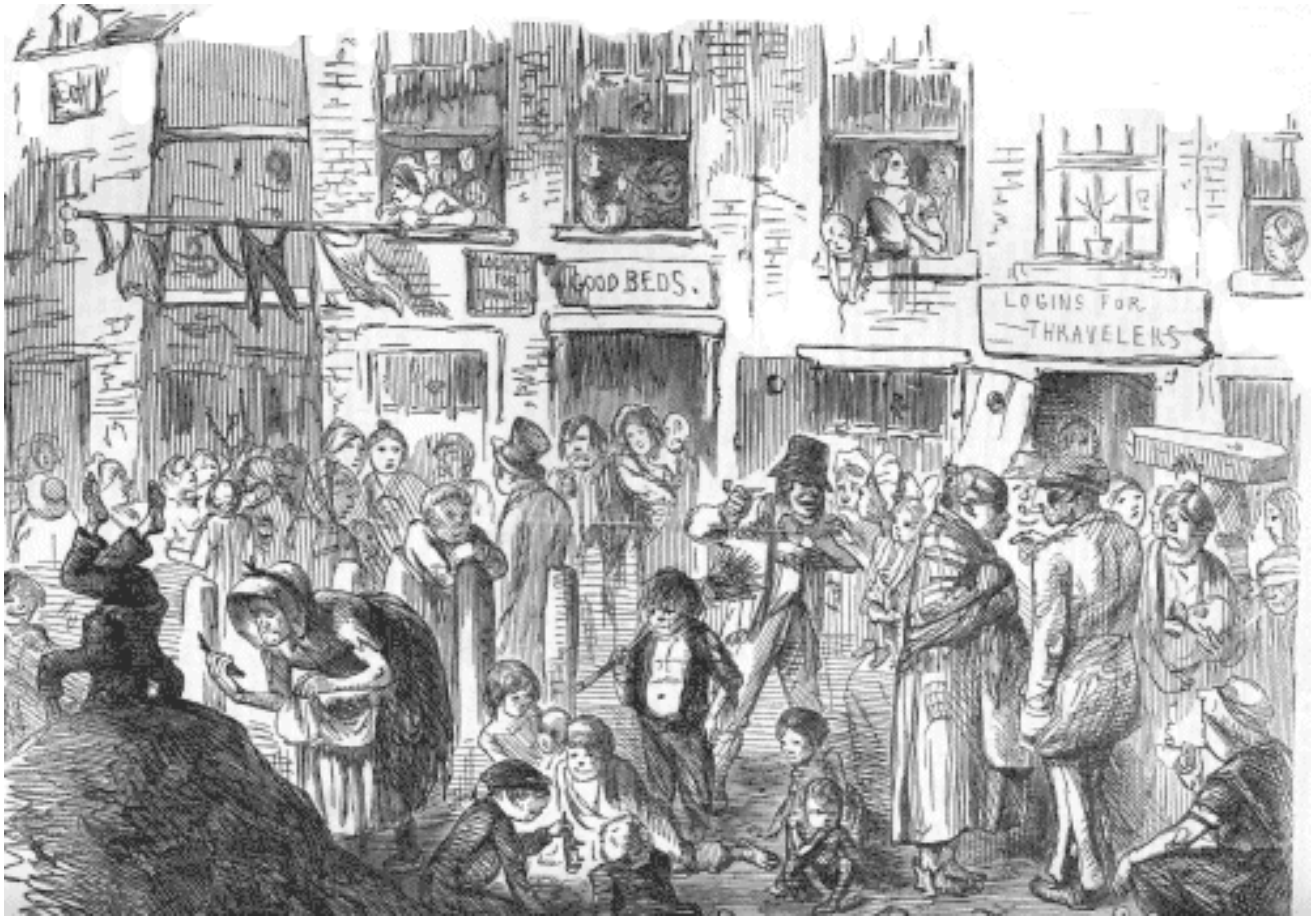
- SQL
- Tableau (41%),
- Excel (34%), PowerPoint (16%)
- Python (30%), SAS (22%), R (16%), Plotly (?%)
- JavaScript & JavaScript-based data-driven documents D3.js (13%)

Source: Ryan et al, IEEE CG&A, 2019 using data from Labor Insight

Visualization is not new

Rich History

Let’s go back some 170 years to 1854, London, England



Newsflash, 1854

The most terrible outbreak of cholera which ever occurred in this kingdom, is probably that which is taking place in Broad Street, Golden Square, and adjoining streets.

Within two hundred and fifty yards of the spot where Cambridge Street joins Broad Street, there are upwards of five hundred fatal attacks of cholera in ten days.

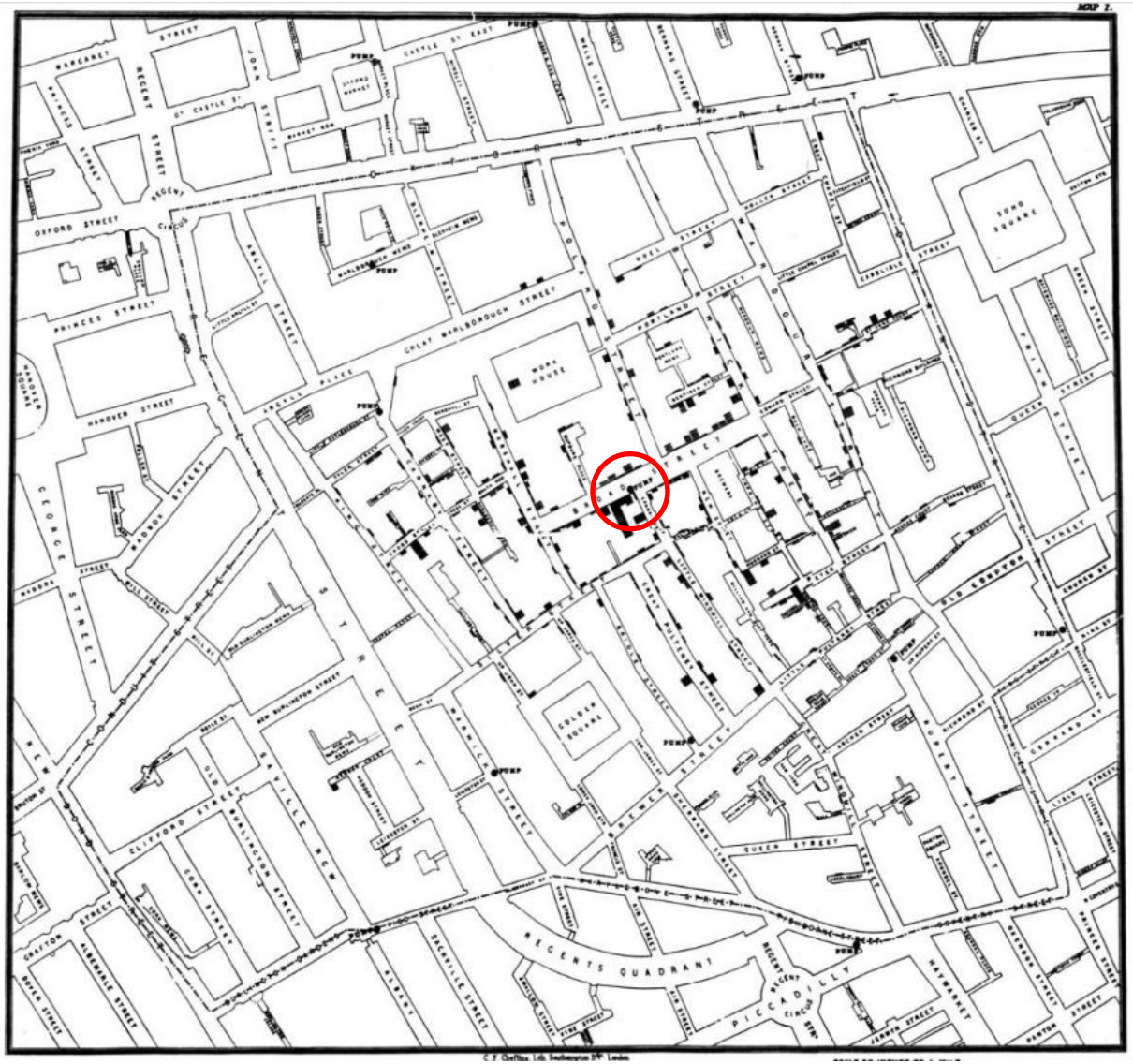
The mortality in this limited area probably equals any that was ever caused in this country, even by the plague; and it is much more sudden, as the greater number of cases terminated in a few hours.

What Can We Do?

What Is The Cause?

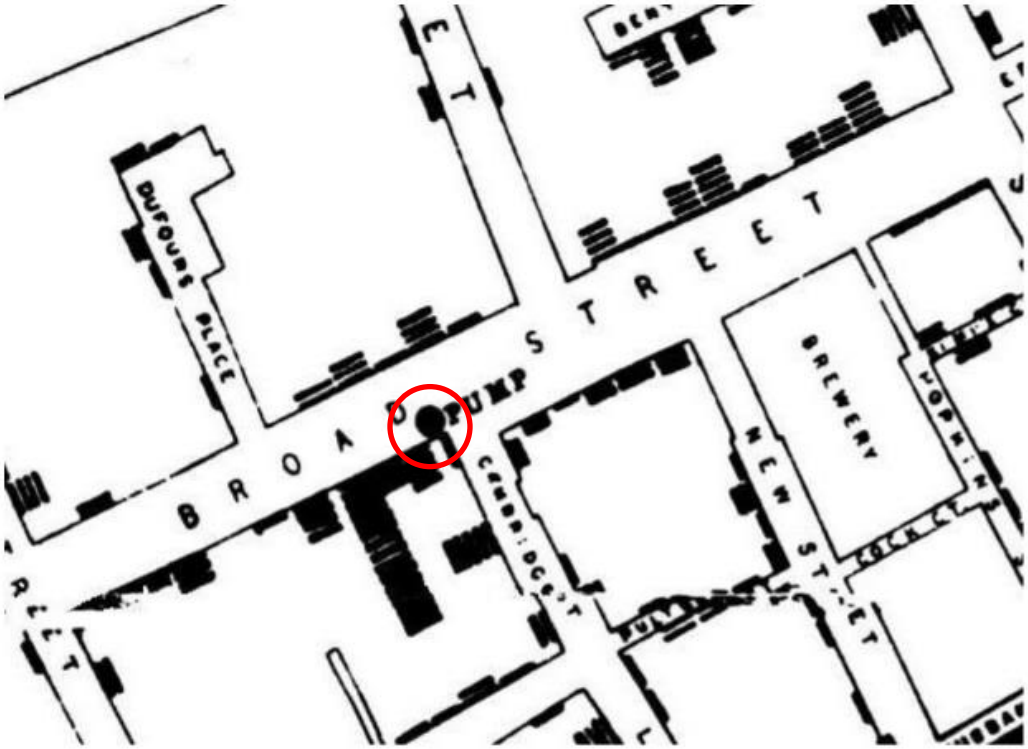
How Can We Eliminate It?

Time For "Imagination"



John Snow

Time For "Imagination"



John Snow

Proved The Hypothesis

Hypothesis: cholera spreads through water

- and not via some other fantastic causes
- one said it rose out of the burying grounds of plague victims from two centuries earlier
- the bacteria was discovered later, in 1886

A real-life experiment (often the case with observational data)

- established the mode of cholera transmission
- and consequently the method of prevention: keep drinking water, food, and hands clear of infected sewage

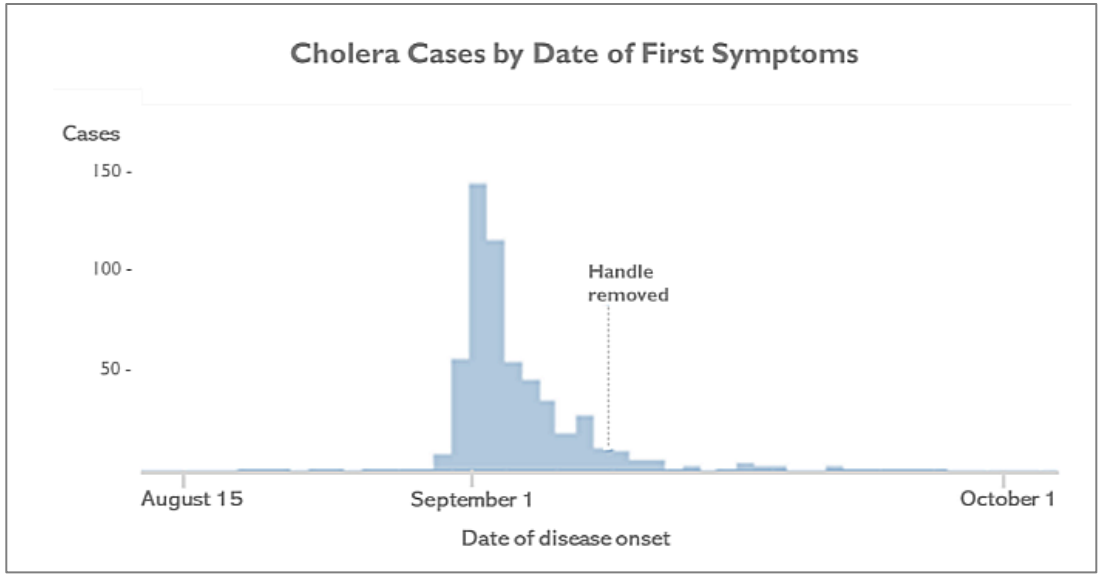
Visualization provided

- inspiration
- convincing arguments to justify actions (removing the pump handle)
- led to Dr. John Snow's historic immortality
- a bar near the old Broad Street pump bears his name (safe drinking)

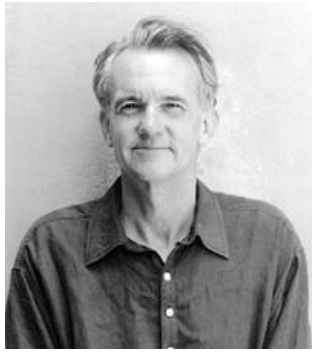
Graphed Over Time

Turns out that the handle was removed at the end of the outbreak

- graphing deaths over time revealed this
- also done by Dr. Snow but far less publicized
- but likely prevented a new outbreak



Much Later

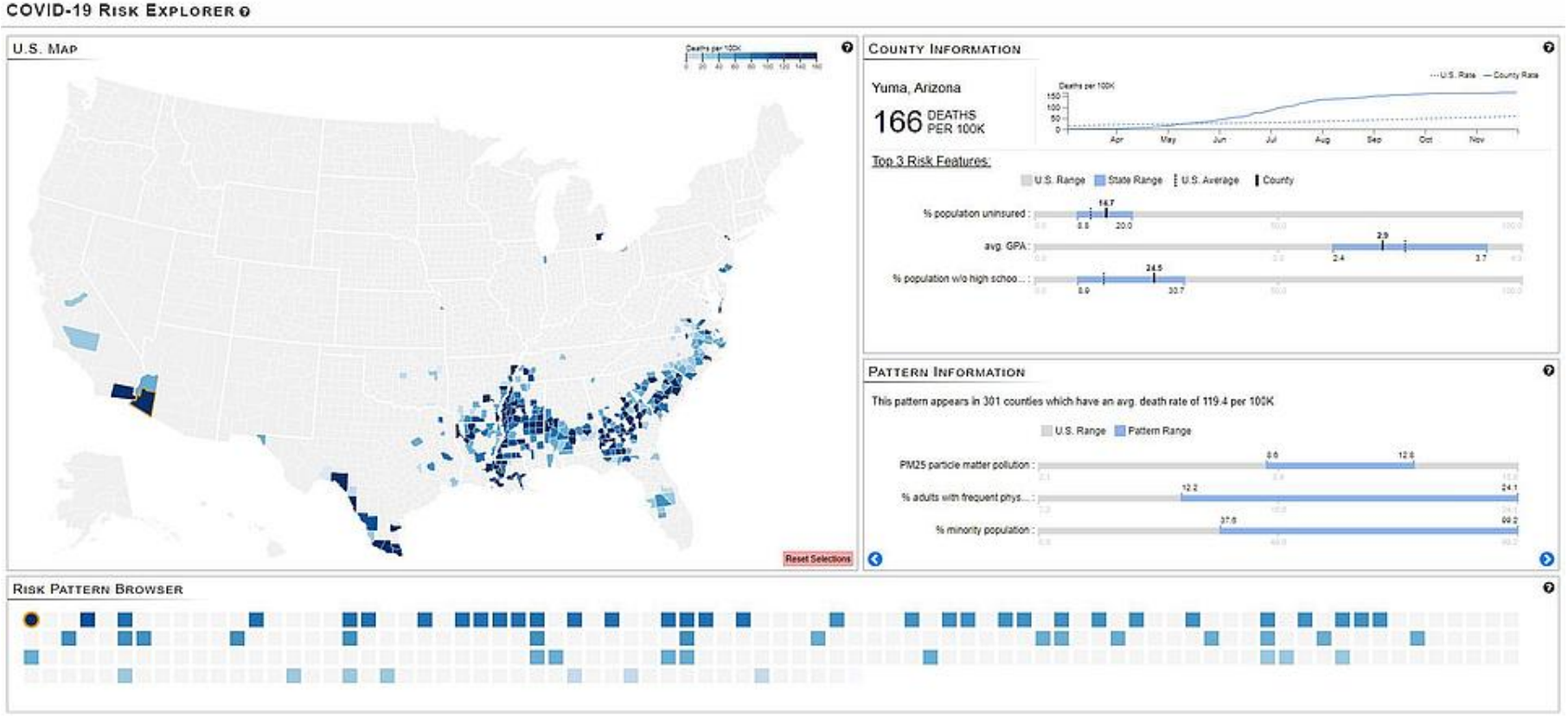


Edward Tufte redrew the map

- only kept the most critical street and building details
- switched out Dr. Snow's dashes for dots
- focused the visual emphasis on Cholera victims and well locations, and not the features of the ground
- better data-to-ink ratio

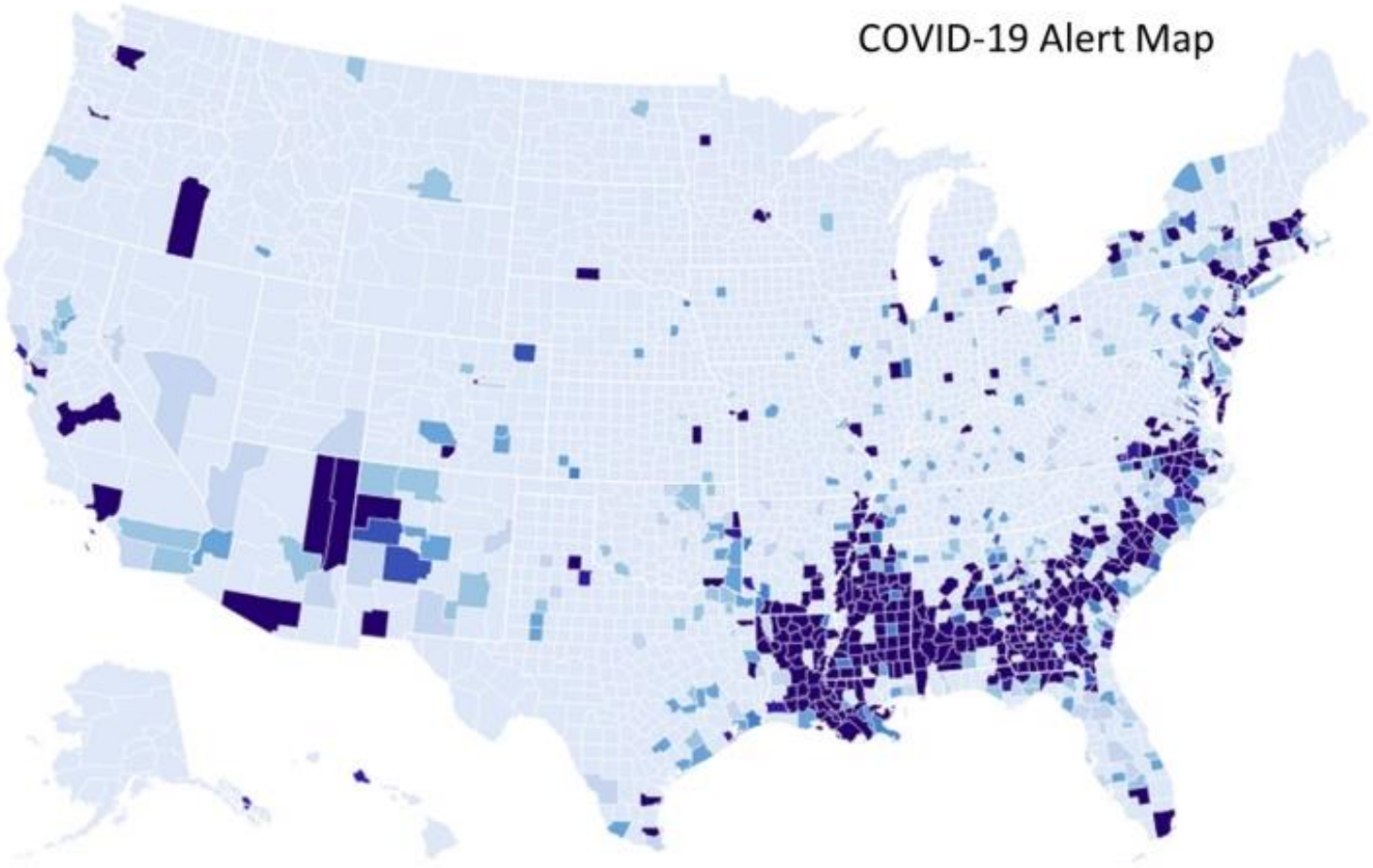
COVID-19 Risk Map

Use pattern analysis of US county socio-economic vulnerability risk factors to predict the initial spread of the virus



COVID-19 Risk Map

created
May 10, 2020

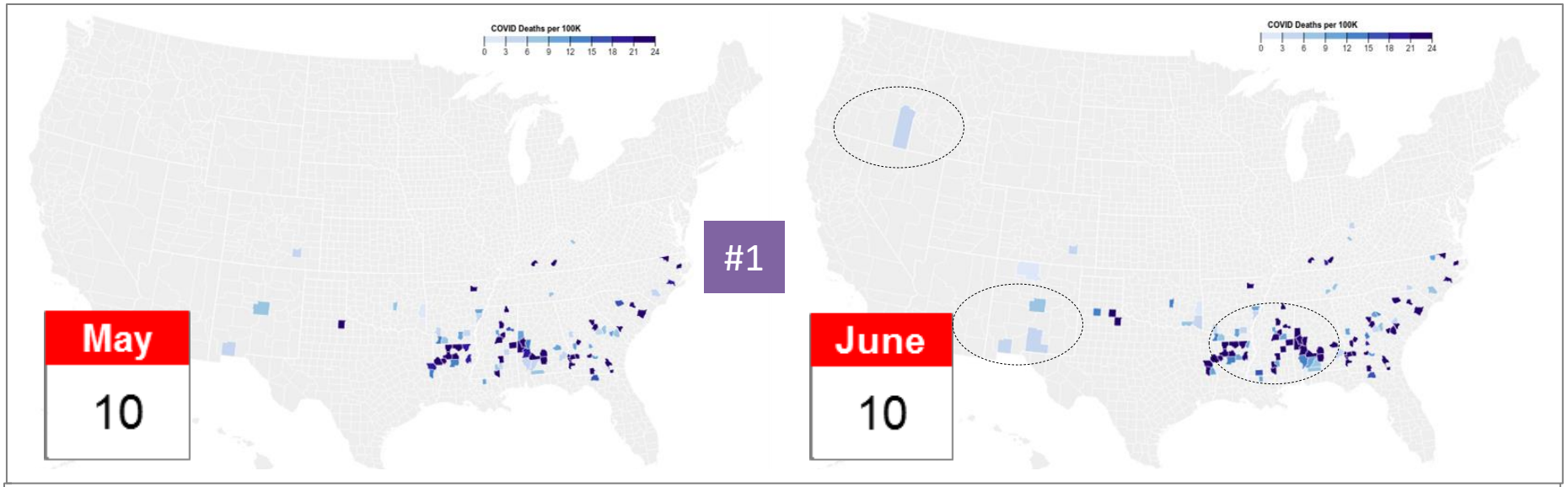


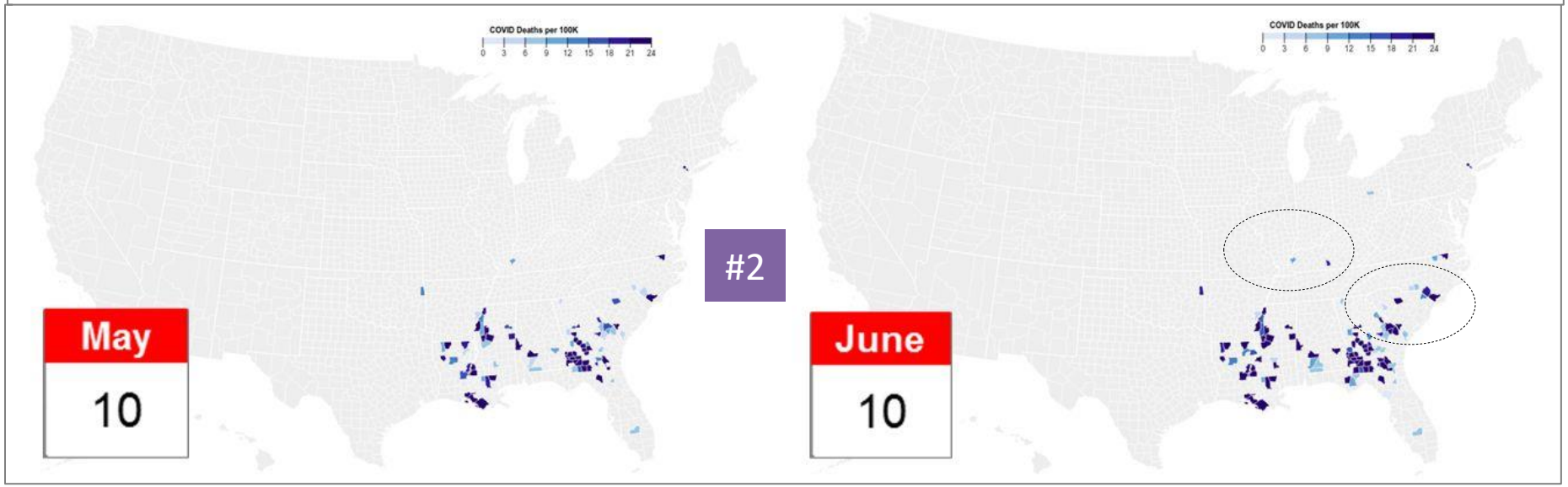
Color mapping:

- the number of times a U.S. county is part of a “high risk” set
- the higher level of risk a county has for high COVID-19 death rates the darker the color

Only counties with at least 1 death on May 10, 2020 are shown

Pattern-Based Predictions (2020)





Published in…

K. Mueller, E. Papenhausen, "Demographic Pattern Analysis to Predict COVID-19 Fatalities on the US County Level," *ACM Digital Government: Research and Practice,* 2 (1): 1-11, 2020.

D. Coelho, N. Gupta, E. Papenhausen, **K. Mueller**, "Patterns of Social Vulnerability – An Interactive Dashboard to Explore Risks to Public Health on the US County Level," *Workshop on Visual Analytics in Healthcare (VAHC, jointly held with AMIA)*, Washington, DC, November, 2022. (won Best Paper Award)

link

What Is Needed for Visualization?

Data (wide variety)

Algorithms

- data mining
- data analytics

Computer

- run those algorithms
- data storage

Humans

- with a purpose/need to understand their data
- endowed with cognitive faculties, creative thought, intuition
- domain expertise

Understanding of humans

- perception, cognition, HCI issues
- we can gain it through experimentation with humans

What Is Needed for Visualization?

Data (wide variety)

Algorithms

- data mining
- data analytics

Computer

- run those algorithms
- data storage

= Visual Analytics

Humans

- with a purpose/need to understand their data
- endowed with cognitive faculties, creative thought, intuition
- domain expertise

Understanding of humans

- perception, cognition, HCI issues
- we can gain it through experimentation with humans

Dr. John Snow: A Visual Analytics Pioneer

Dr. John Snow's London Cholera Map of 1854

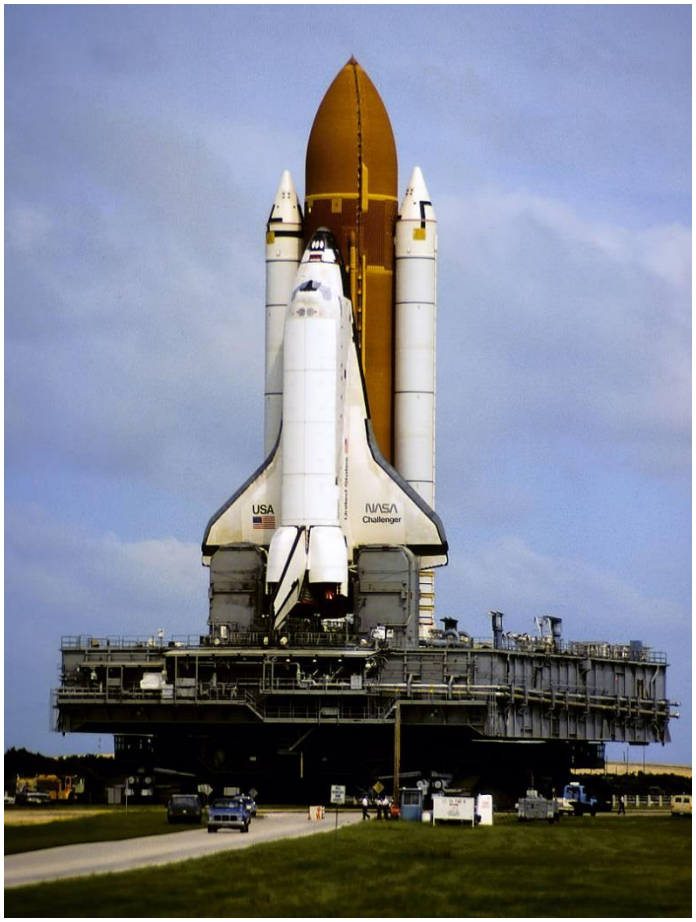
- data collection
- data assimilation
- statistical testing
- visualization
- computational analysis (brain)
- domain knowledge

Very early example of visual analytics

More Recent History

Let's go back some 40 years to 1986, JFK Space Center, FL

The crew of Space Shuttle mission STS-51-L 11/15/85. Back row, left to right: Ellison S. Onizuka, Sharon Christa McAuliffe, Greg Jarvis, Judy Resnik. Front row, left to right: Michael J. Smith, Dick Scobee, Ron McNair.



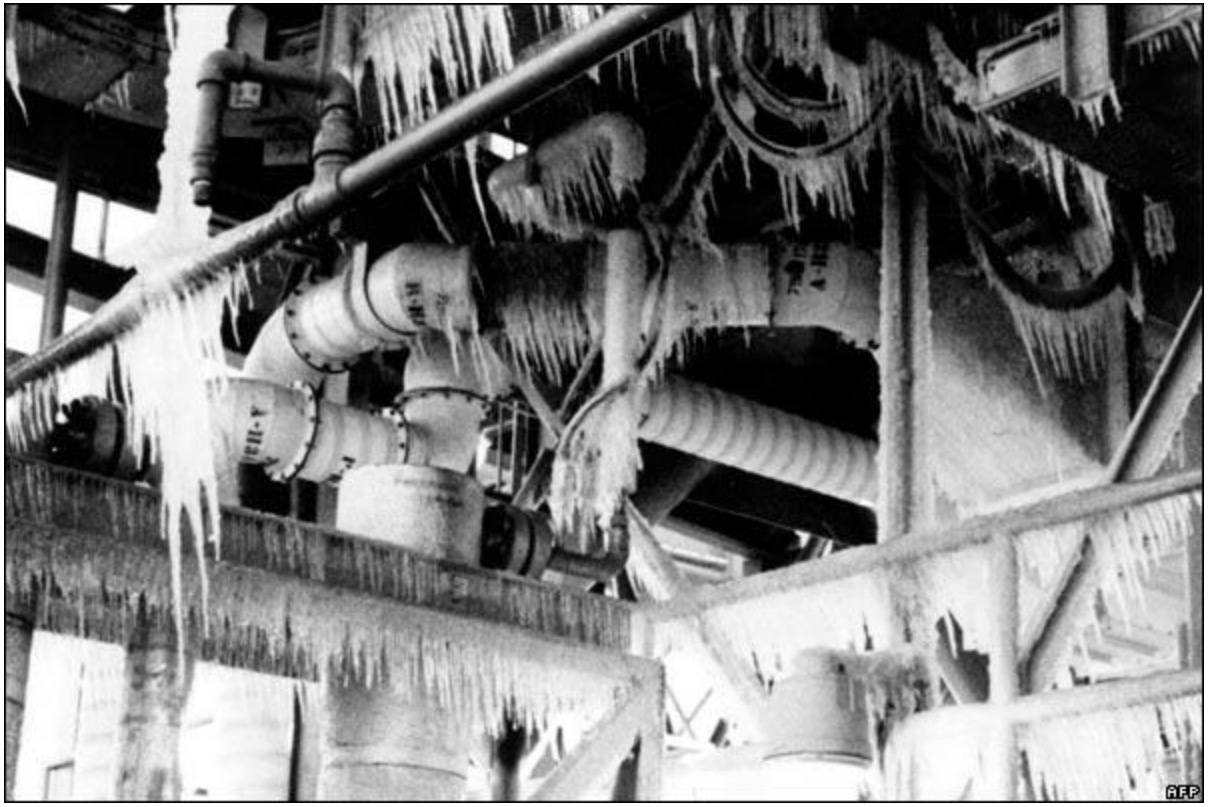
73 Seconds After Lift-off

What Happened?

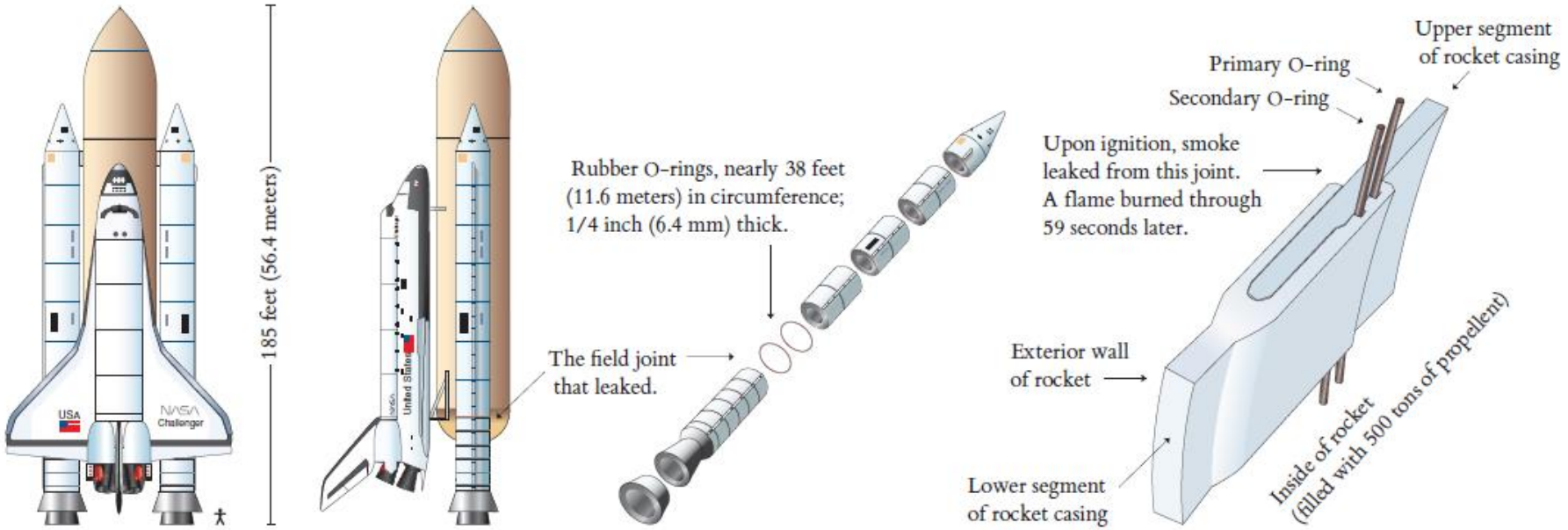
What Was The Cause?

The Day of The Launch

36 degrees F on Launch Pad 39



Space Shuttle 101



Fast Forward
58 Seconds After Ignition

What Happened?

What Was The Cause?

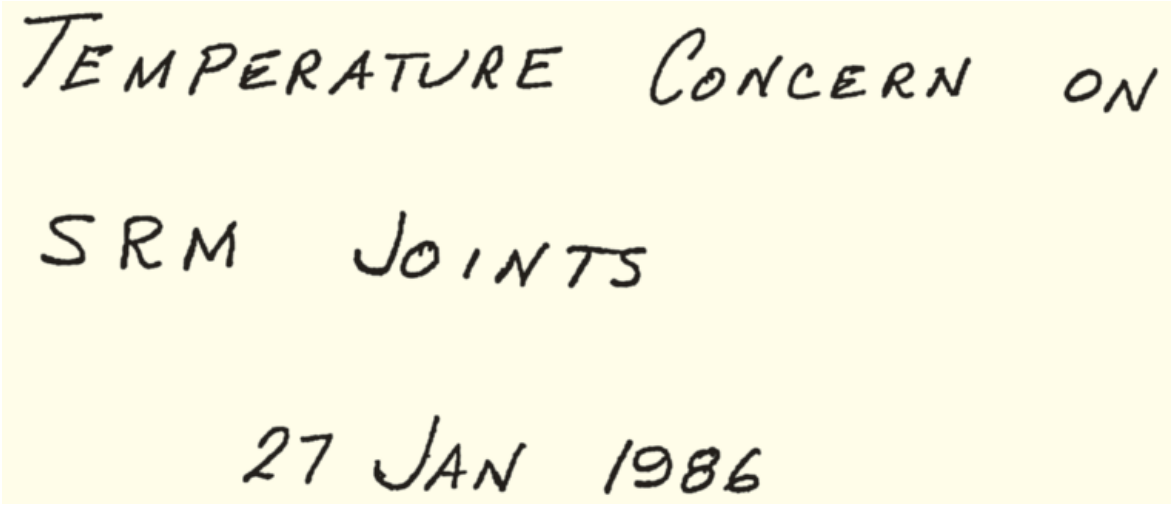
Could It Have Been Prevented?

Engineers At Thiokol Had A Hunch

Two days before launch they presented their concerns

- created 13 charts to make their case

Slide #1:



- SRM – Solid Rocket Motor

SLIDE #2

Teaches about past damages to O-ring

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.

SLIDES #2 AND 3

Teaches about O-ring damage mechanics and erosion

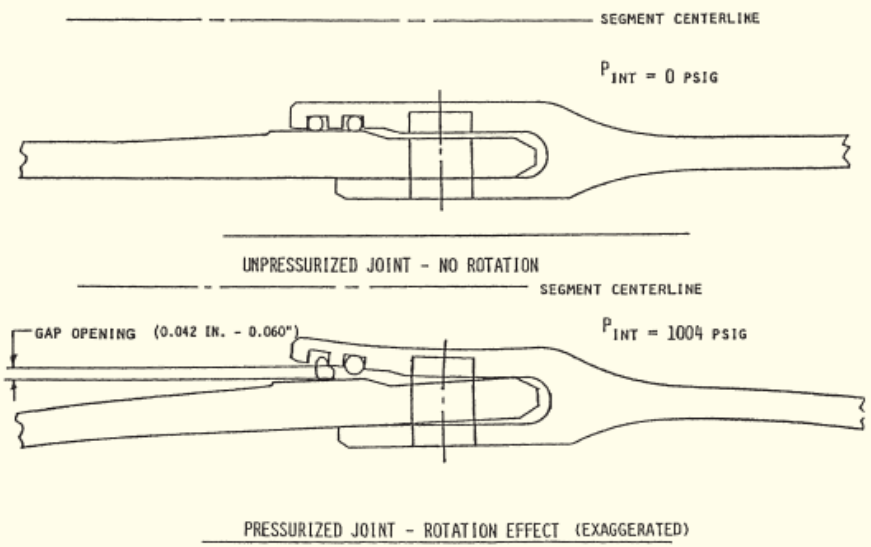
PRIMARY CONCERNS -

FIELD JOINT - HIGHEST CONCERN

- EROSION PENETRATION OF PRIMARY SEAL REQUIRES RELIABLE SECONDARY SEAL FOR PRESSURE INTEGRITY
  - IGNITION TRANSIENT - (0-600 MS)
    - (0-170 MS)HIGH PROBABILITY OF RELIABLE SECONDARY SEAL
    - (170-330 MS) REDUCED PROBABILITY OF RELIABLE SECONDARY SEAL
    - (330-600 MS) HIGH PROBABILITY OF NO SECONDARY SEAL CAPABILITY

- STEADY STATE - (600 MS - 2 MINUTES)
  - IF EROSION PENETRATES PRIMARY O-RING SEAL - HIGH PROBABILITY OF NO SECONDARY SEAL CAPABILITY
    - BENCH TESTING SHOWED O-RING NOT CAPABLE OF MAINTAINING CONTACT WITH METAL PARTS GAP OPENING RATE TO MEOP
    - BENCH TESTING SHOWED CAPABILITY TO MAINTAIN O-RING CONTACT DURING INITIAL PHASE (0-170 MS) OF TRANSIENT

PRIMARY CONCERNS - CONT



SLIDES #4 AND 5

Lists temperature and blow-by history for two SRMs

BLOW BY HISTORY

SRM-15 WORST BLOW-BY

- 2 CASE JOINTS (80°), (110°) ARC
- MUCH WORSE VISUALLY THAN SRM-22

SRM 22 BLOW-BY

- 2 CASE JOINTS (30-40°)

SRM-13A, 15, 16A, 18, 23A 24A

- NOZZLE BLOW-BY

HISTORY OF O-RING TEMPERATURES (DEGREES-F)

MOTOR	MBT	AMB	O-RING	WIND
DM-1	68	36	47	10 MPH
DM-2	76	45	52	10 MPH
QM-3	72.5	40	48	10 MPH
QM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29 27	10 MPH 25 MPH

Assume You're a NASA Manager

Given the information provided in the company slides

- would you vote for a launch?
- ignore you know about the consequences

Be keenly aware of the immense PR pressures

- President Reagan's upcoming State of the Union speech
- the first civilian in space
- NASA's funding problems

Launch:

- No: OK with a PR disaster & possible budget cuts down the road
- Yes: the rocket company is too cautious & concerns are unproven

Why The Recommendation Failed

Presentation only has exactly two shuttle flights

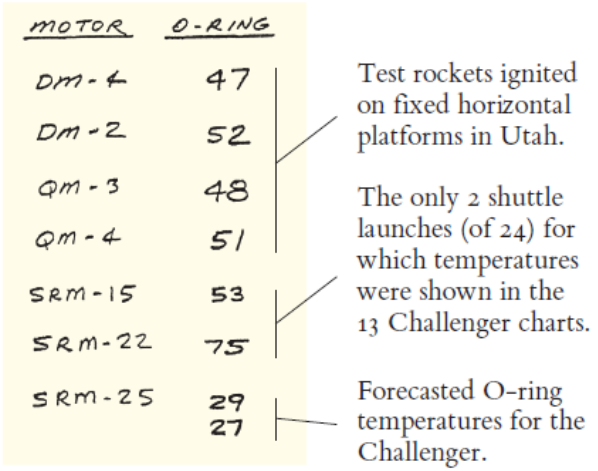
- one with two blow-by's and high temperature
- one with two blow-by's and low temperature
- ignores all other 22 shuttle flights (SRM)

Statistically weak

Recommendation

- "O-ring temp must be >53°F at launch"

- is only based on a sample size of 1
- context of other flights is missing
- no statistical leverage

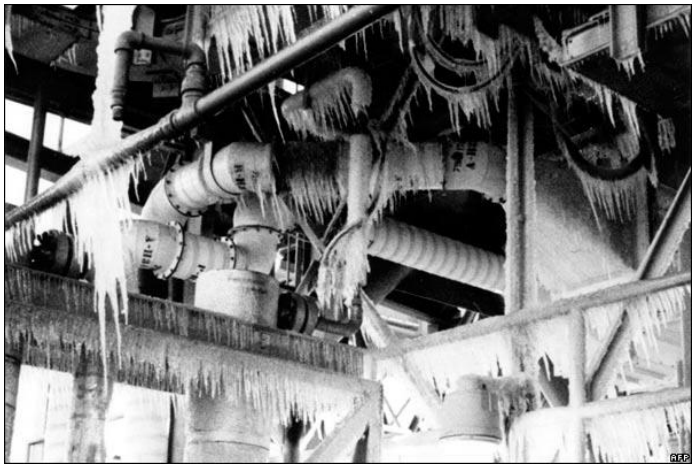


MOTOR	O-RING	
DM-4	47	Test rockets ignited on fixed horizontal platforms in Utah.
DM-2	52	
QM-3	48	The only 2 shuttle launches (of 24) for which temperatures were shown in the 13 Challenger charts.
QM-4	51	
SRM-15	53	
SRM-22	75	
SRM-25	29	Forecasted O-ring temperatures for the Challenger.
	27	

Deficiencies

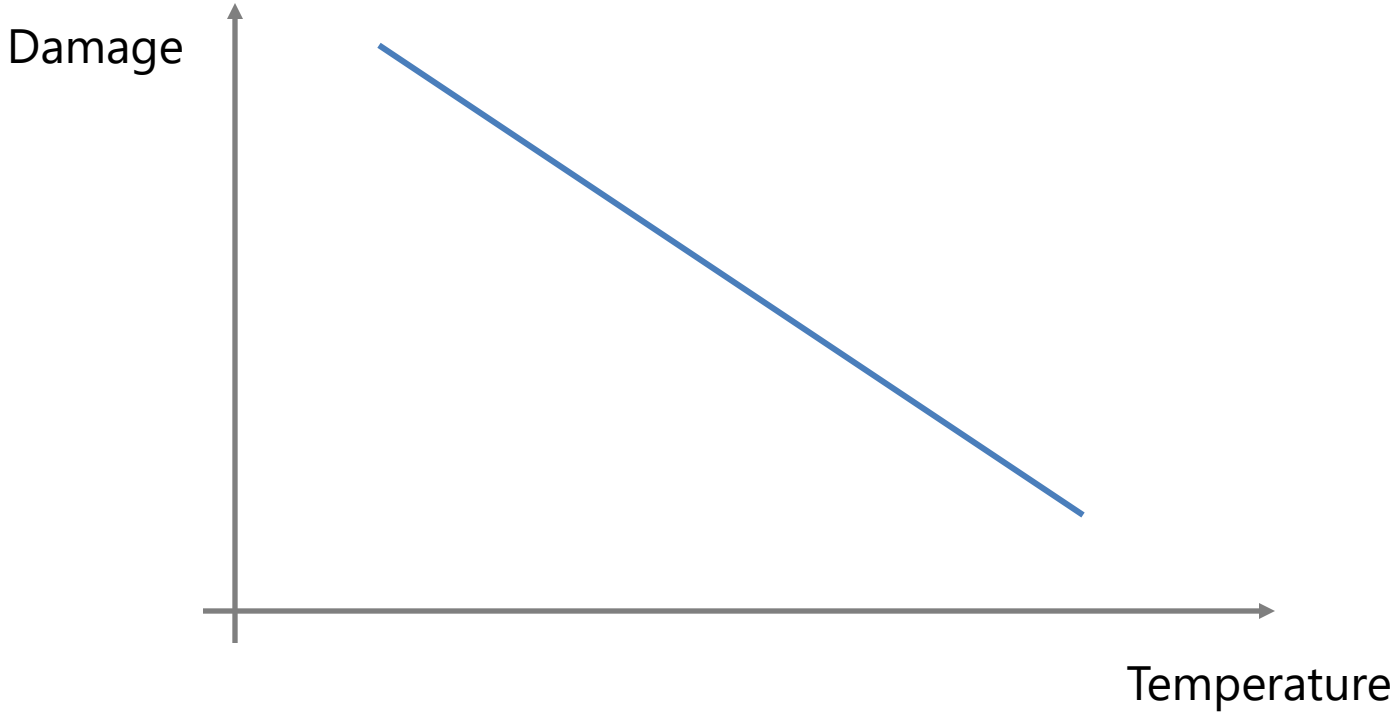
Lots of numbers and facts

But no causal evidence that could predict



What is needed?

What Is Needed?

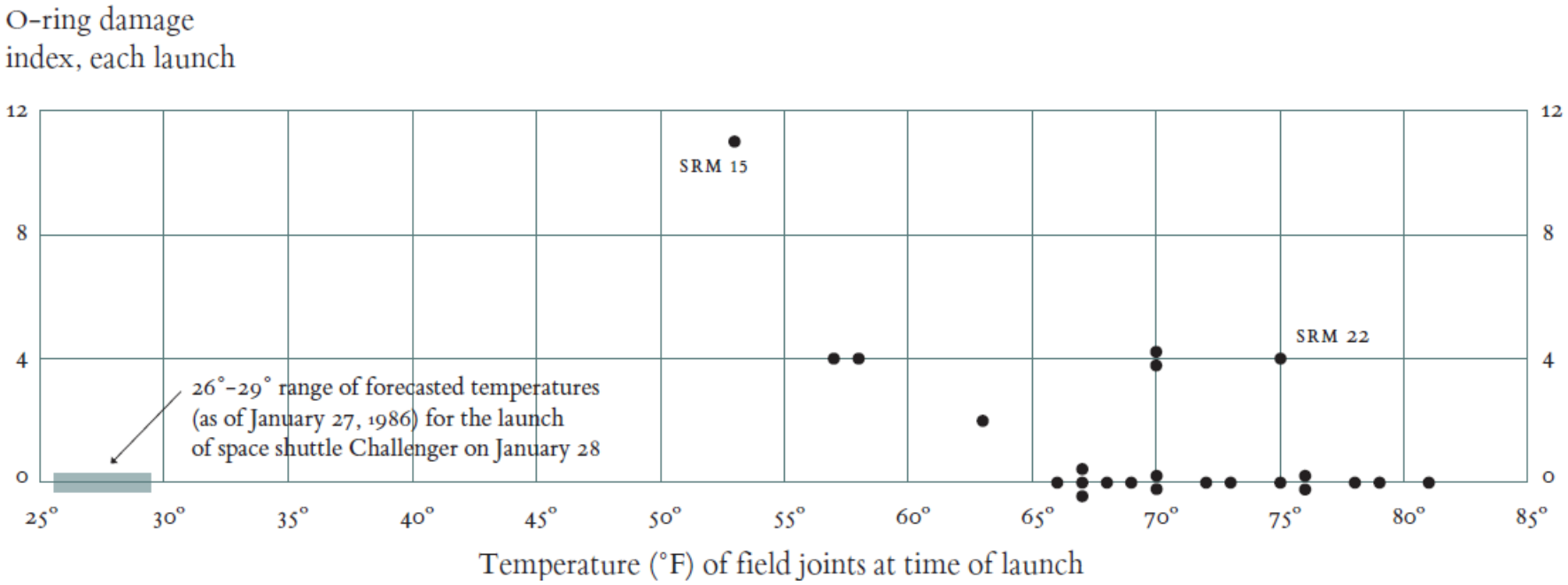


Need a measure for damage

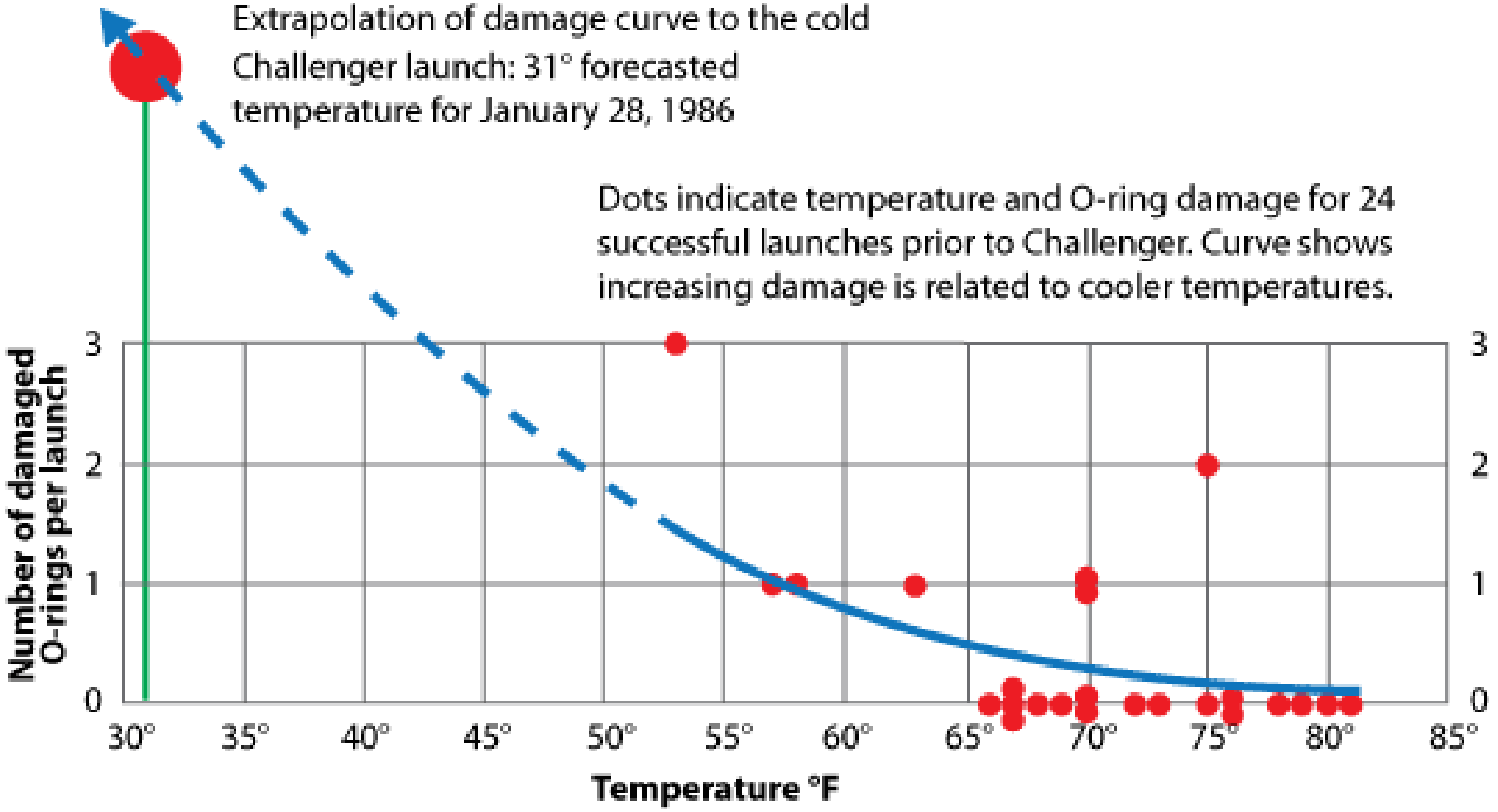
Damage Index

Flight	Date	Temperature °F	Erosion incidents	Blow-by incidents	Damage index	Comments
51-C	01.24.85	53°	3	2	11	Most erosion any flight; blow-by; back-up rings heated.
41-B	02.03.84	57°	1		4	Deep, extensive erosion.
61-C	01.12.86	58°	1		4	O-ring erosion on launch two weeks before Challenger.
41-C	04.06.84	63°	1		2	O-rings showed signs of heating, but no damage.
1	04.12.81	66°			0	Coolest (66°) launch without O-ring problems.
6	04.04.83	67°			0	
51-A	11.08.84	67°			0	
51-D	04.12.85	67°			0	
5	11.11.82	68°			0	
3	03.22.82	69°			0	
2	11.12.81	70°	1		4	Extent of erosion not fully known.
9	11.28.83	70°			0	
41-D	08.30.84	70°	1		4	
51-G	06.17.85	70°			0	
7	06.18.83	72°			0	
8	08.30.83	73°			0	
51-B	04.29.85	75°			0	
61-A	10.30.85	75°		2	4	No erosion. Soot found behind two primary O-rings.
51-I	08.27.85	76°			0	
61-B	11.26.85	76°			0	
41-G	10.05.84	78°			0	
51-J	10.03.85	79°			0	
	06.27.82	80°			?	O-ring condition unknown; rocket casing lost at sea.
51-F	07.29.85	81°			0	

Visualize It – Just The Facts

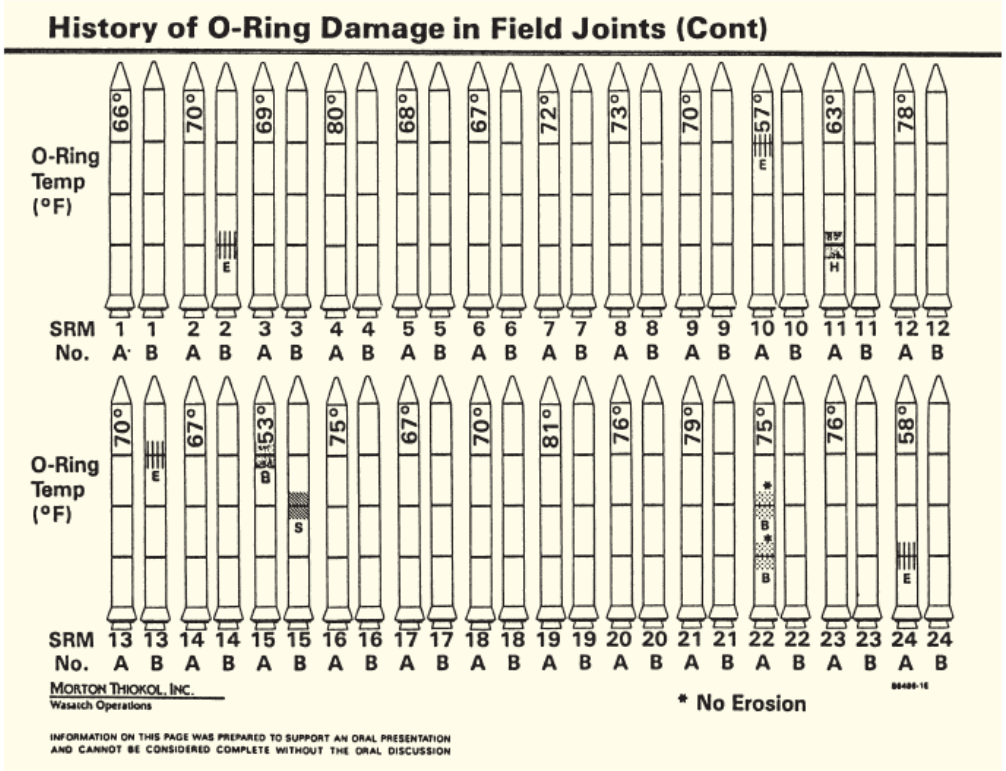


Visualize It – Tell The Story



Shown at Congressional Hearings

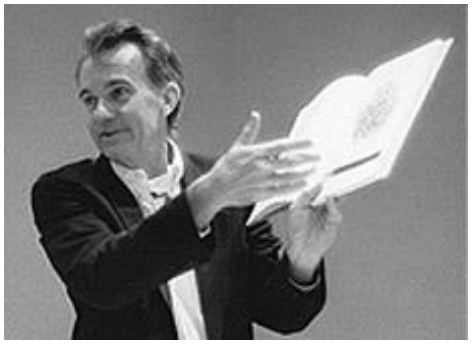
Used these charts



All information is there

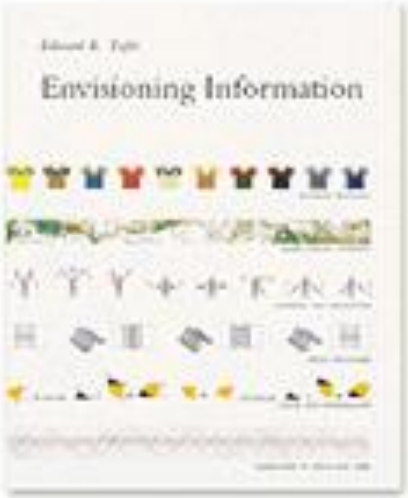
- but very hard to identify and assimilate
- why?

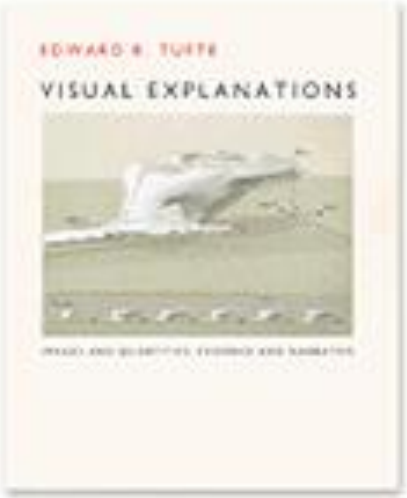
Source: Edward Tufte



Four seminal books

- standard literature for every visualization enthusiast
- written 1983, 1990, 1997, 2006





- taught information design at Princeton University
- now a professor at Yale University

Course Topics

Spatial Data

shock wave

virtual frog

spiral flow

nerve cell

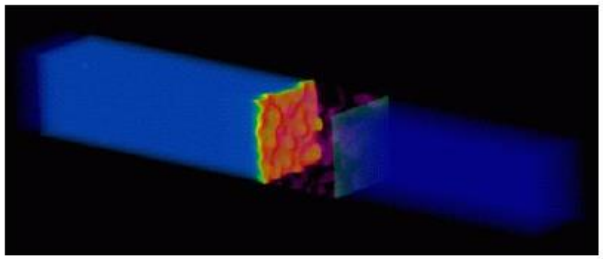
wind flow

transparent MRI head

semi-transparent
tomato

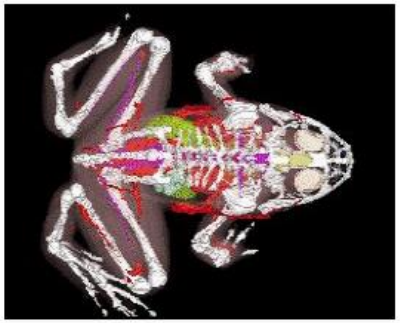
MRI head

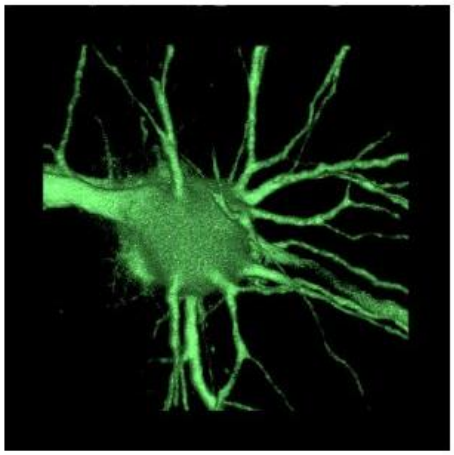
Spatial Data



shock wave

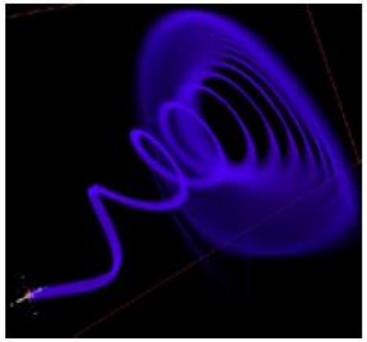
virtual frog

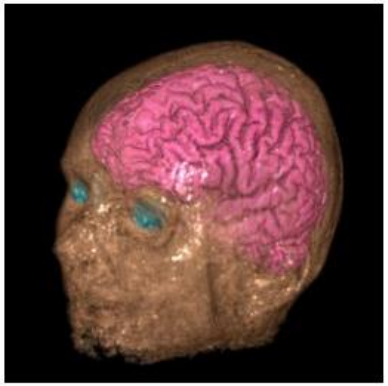




nerve cell

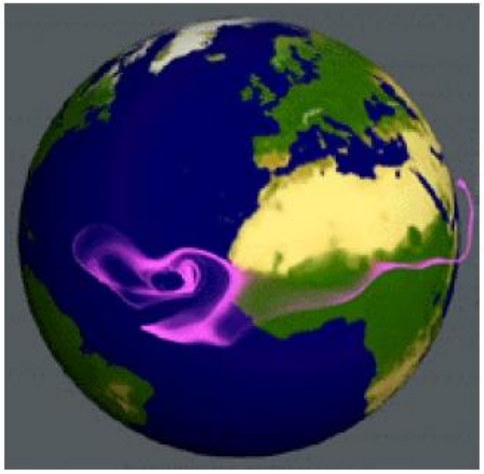
spiral flow





transparent MRI head

wind flow



MRI head

semi-transparent tomato

Spatial Data

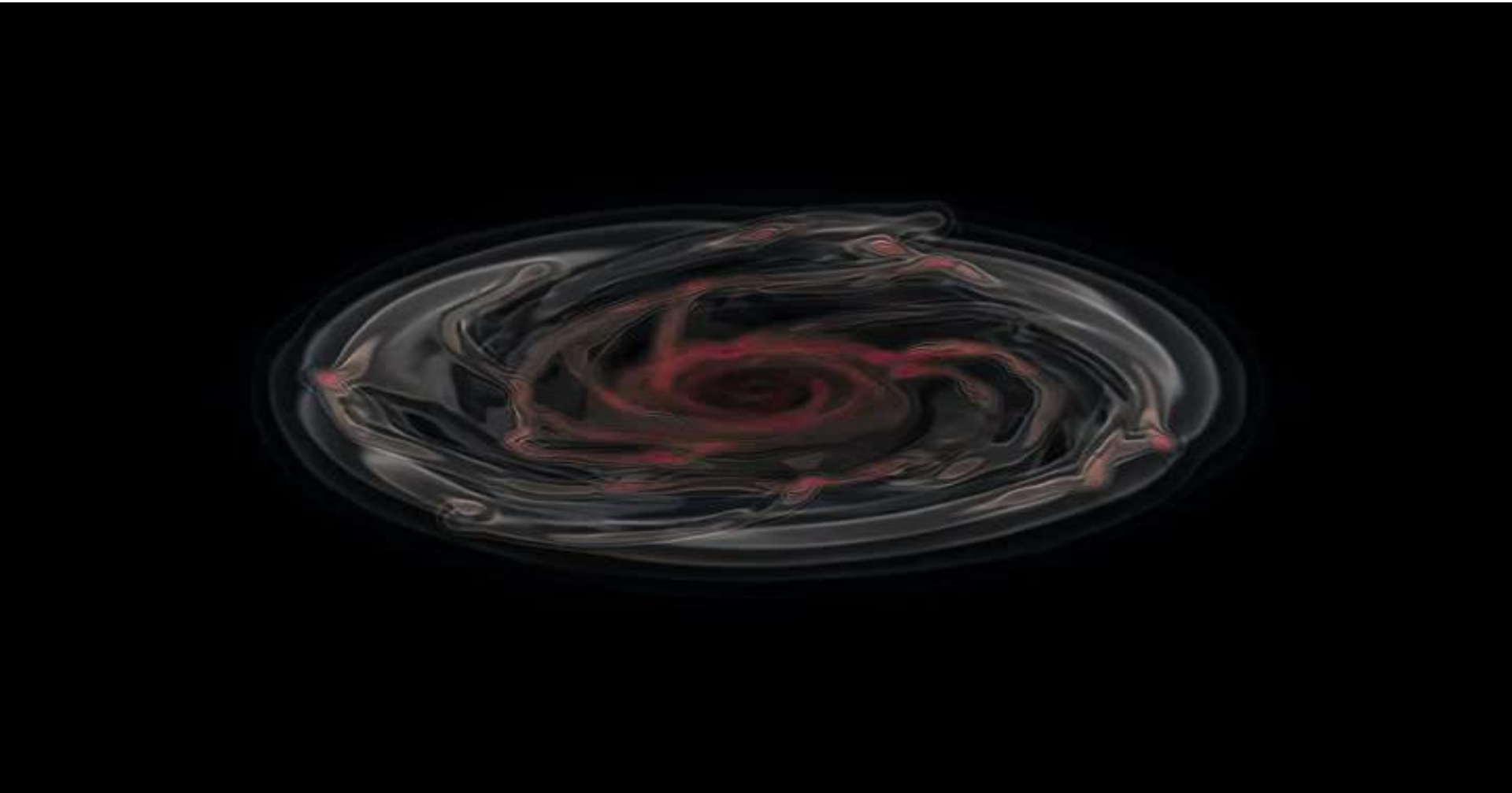
Example: Datasets obtained by 3D volumetric scans (CT, MRI)

- what are some questions you might have?

Spatial Data

Example: Datasets obtained by 3D Simulations

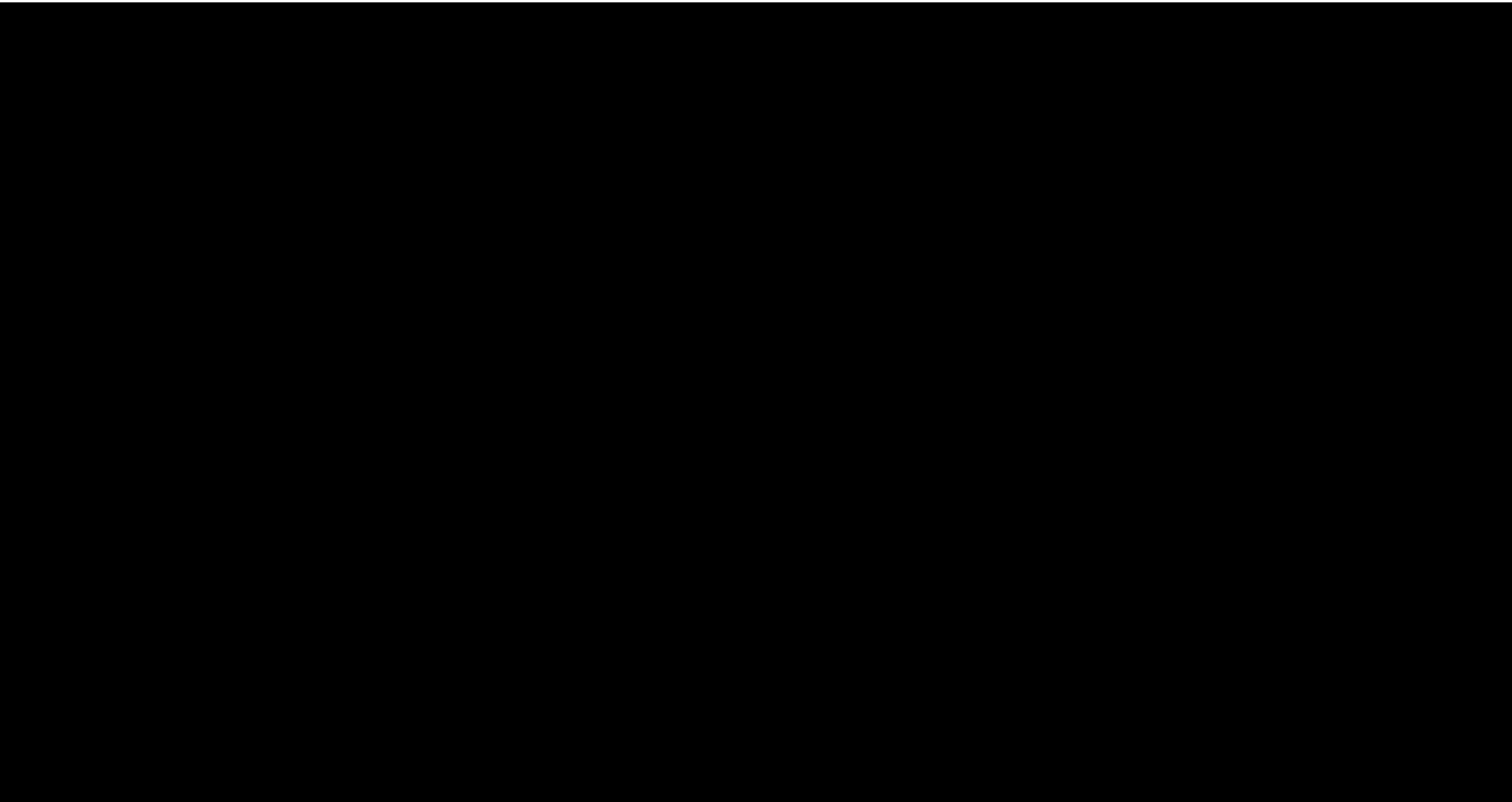
- what are some questions you might have?



Spatial Data

Example: Data obtained by observation-supported simulations

- what are some questions you might have?



Non-Spatial Data

The salient features of a car:

- miles per gallon (MPG)
- top speed
- acceleration
- number of cylinders
- horsepower
- weight
- year
- country origin
- brand
- number of seats
- number of doors
- reliability (# of breakdowns)
- and so on...

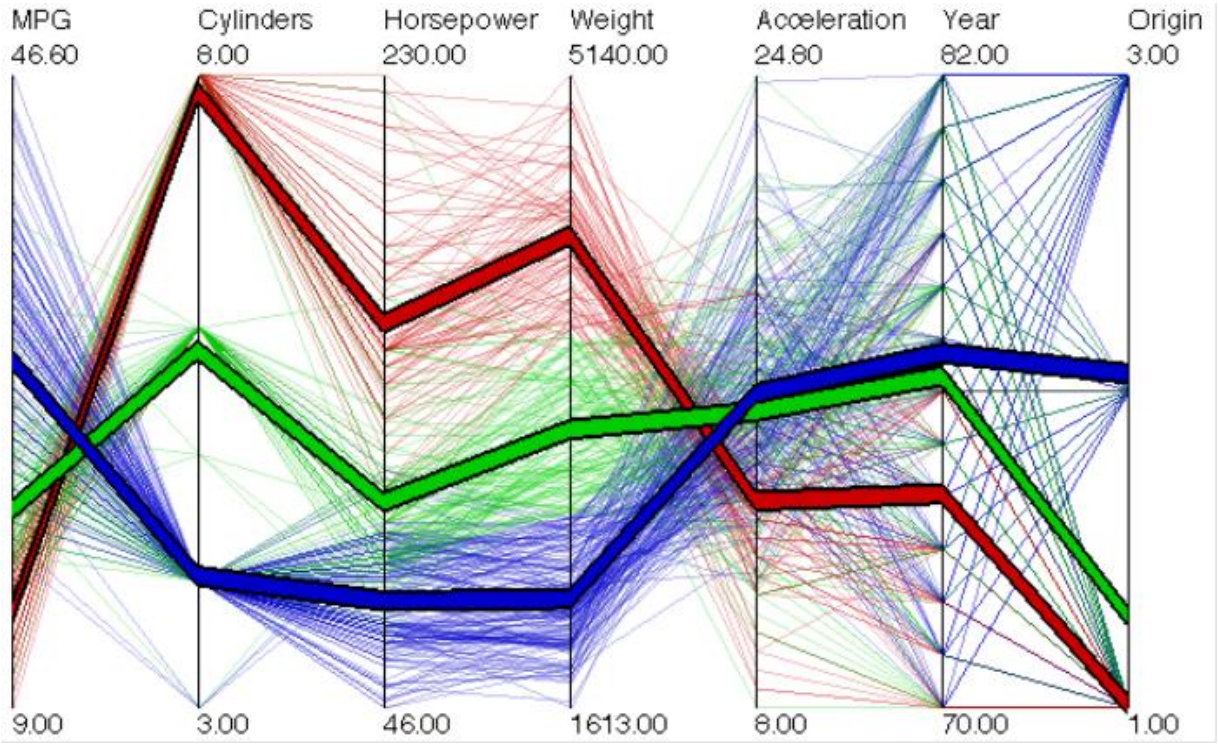
Can You Visualize Them Like This?

Total rows: 428 Total columns: 15 Rows 1-100

	Make	Model	Type	Origin	DriveT...	MSRP	Invoice	Engin...	Cylind...	Horse...	MPG_...	MPG_...	Weight	Wheel...	Leng...
1	Acura	MDX	SUV	Asia	All	$36,945	$33,337	3.5	6	265	17	23	4451	106	189
2	Acura	RSX Type	Sedan	Asia	Front	$23,820	$21,761	2	4	200	24	31	2778	101	172
3	Acura	TSX 4dr	Sedan	Asia	Front	$26,990	$24,647	2.4	4	200	22	29	3230	105	183
4	Acura	TL 4dr	Sedan	Asia	Front	$33,195	$30,299	3.2	6	270	20	28	3575	108	186
5	Acura	3.5 RL 4dr	Sedan	Asia	Front	$43,755	$39,014	3.5	6	225	18	24	3880	115	197
6	Acura	3.5 RL w/I	Sedan	Asia	Front	$46,100	$41,100	3.5	6	225	18	24	3893	115	197
7	Acura	NSX coup	Sports	Asia	Rear	$89,765	$79,978	3.2	6	290	17	24	3153	100	174
8	Audi	A4 1.8T 4c	Sedan	Europe	Front	$25,940	$23,508	1.8	4	170	22	31	3252	104	179
9	Audi	A41.8T co	Sedan	Europe	Front	$35,940	$32,506	1.8	4	170	23	30	3638	105	180
10	Audi	A4 3.0 4dr	Sedan	Europe	Front	$31,840	$28,846	3	6	220	20	28	3462	104	179
11	Audi	A4 3.0 Qu	Sedan	Europe	All	$33,430	$30,366	3	6	220	17	26	3583	104	179
12	Audi	A4 3.0 Qu	Sedan	Europe	All	$34,480	$31,388	3	6	220	18	25	3627	104	179
13	Audi	A6 3.0 4dr	Sedan	Europe	Front	$36,640	$33,129	3	6	220	20	27	3561	109	192
14	Audi	A6 3.0 Qu	Sedan	Europe	All	$39,640	$35,992	3	6	220	18	25	3880	109	192
15	Audi	A4 3.0 cor	Sedan	Europe	Front	$42,490	$38,325	3	6	220	20	27	3814	105	180
16	Audi	A4 3.0 Qu	Sedan	Europe	All	$44,240	$40,075	3	6	220	18	25	4013	105	180
17	Audi	A6 2.7 Tur	Sedan	Europe	All	$42,840	$38,840	2.7	6	250	18	25	3836	109	192
18	Audi	A6 4.2 Qu	Sedan	Europe	All	$49,690	$44,936	4.2	8	300	17	24	4024	109	193

How are MPG, weight, HP, and reliability related? Are there tradeoffs?
Which car is best for me?

High-Dimensional Data Visualization



Big Data

12+ TBs
of tweet data
every day
? TBs of
data every day
25+ TBs of
log data every day
30 billion RFID
tags today
(1.3B in 2005)
4.6
billion
camera
phones
world wide
100s of
millions
of GPS
enabled
devices
sold
annually
2+
billion
people on
the Web
by end
2011
76 million smart
meters in 2009…
200M by 2014
Google
Reader
talk
more »
Google
Google | Analytics
YouTube
130 m
http://www.

The Scientific Method

In the Age of Data Science

Formulate Question

Generate Hypothesis

Form Experiment (find data sources)

Form Testable Prediction

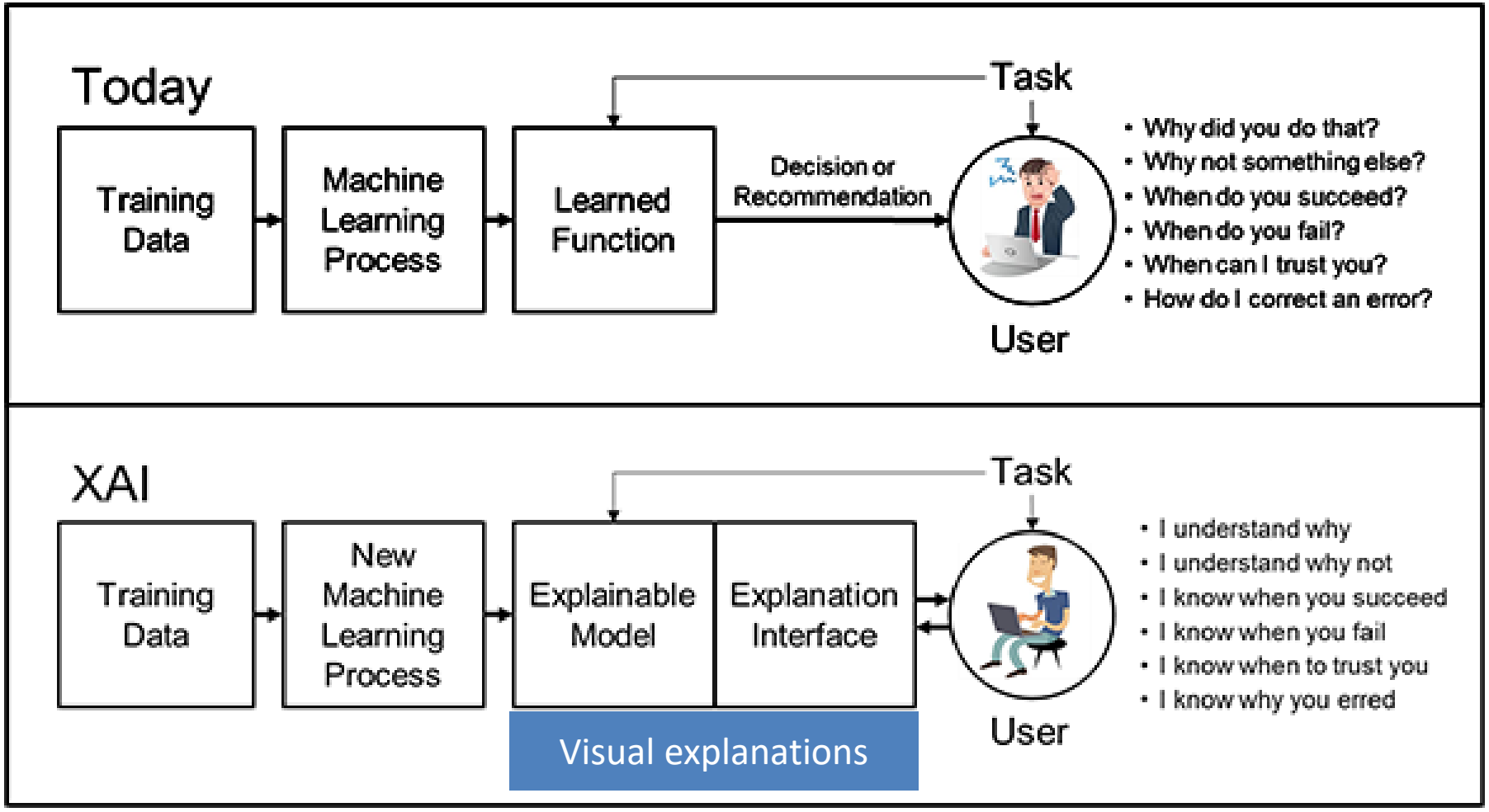
Collect Data (scrape, mine)

Analyze Data

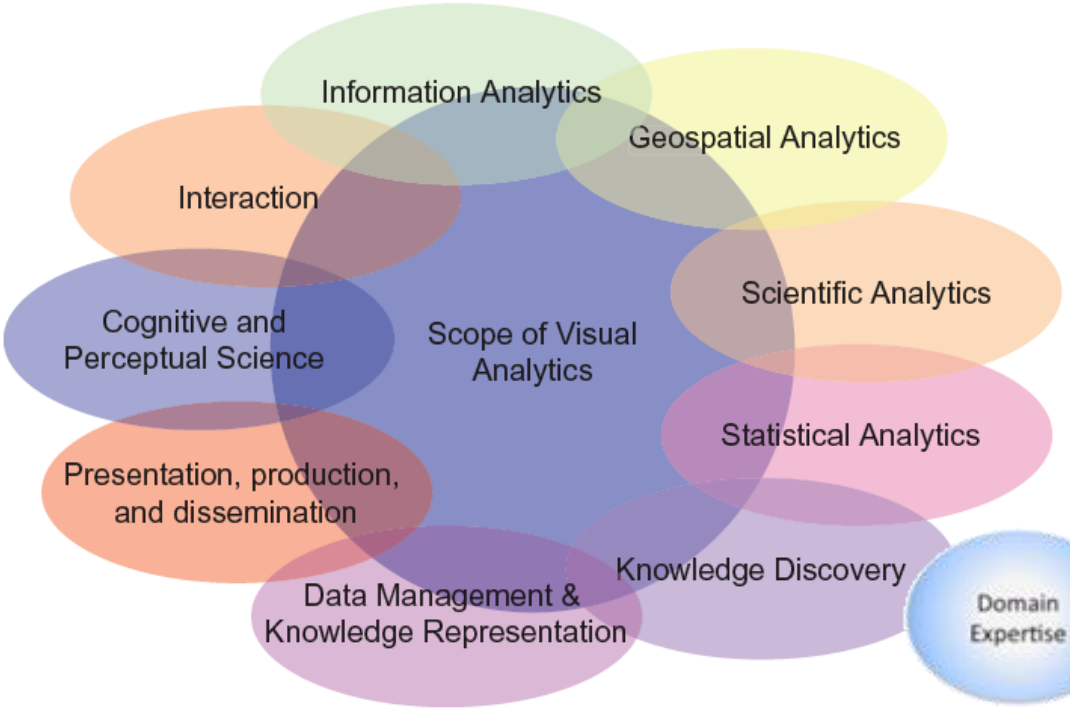
Test Prediction (visualize)

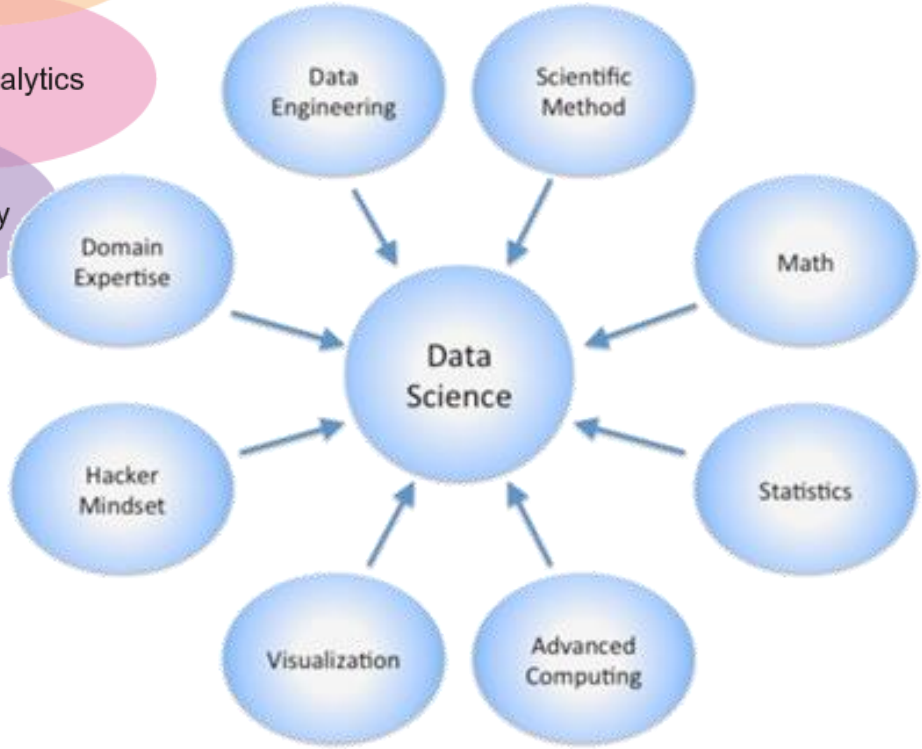
Publish Results

Explainable AI



Visual Analytics vs. Data Science





MODERN DATA SCIENTIST

21th century, requires a mixture of
, computer science, communicatio
who a data scientist is, is equally h
the modern data scientist really i

MATH & STATISTICS

- Machine learning
- Statistical modeling
- Experiment design
- Bayesian inference
- Supervised learning: decision trees, random forests, logistic regression
-
-

DOMAIN KNOWLEDGE & SOFT SKILLS

- Passionate about the business
- Curious about data
- Influence without authority
- Hacker mindset
- Problem solver
- Strategic, proactive, creative, innovative and collaborative

PROGRAMMING & DATABASE

- Computer science fundamentals
- Scripting language e.g. Python
- Statistical computing packages, e.g., R
- Databases: SQL and NoSQL
- Relational algebra
-

COMMUNICATION & VISUALIZATION

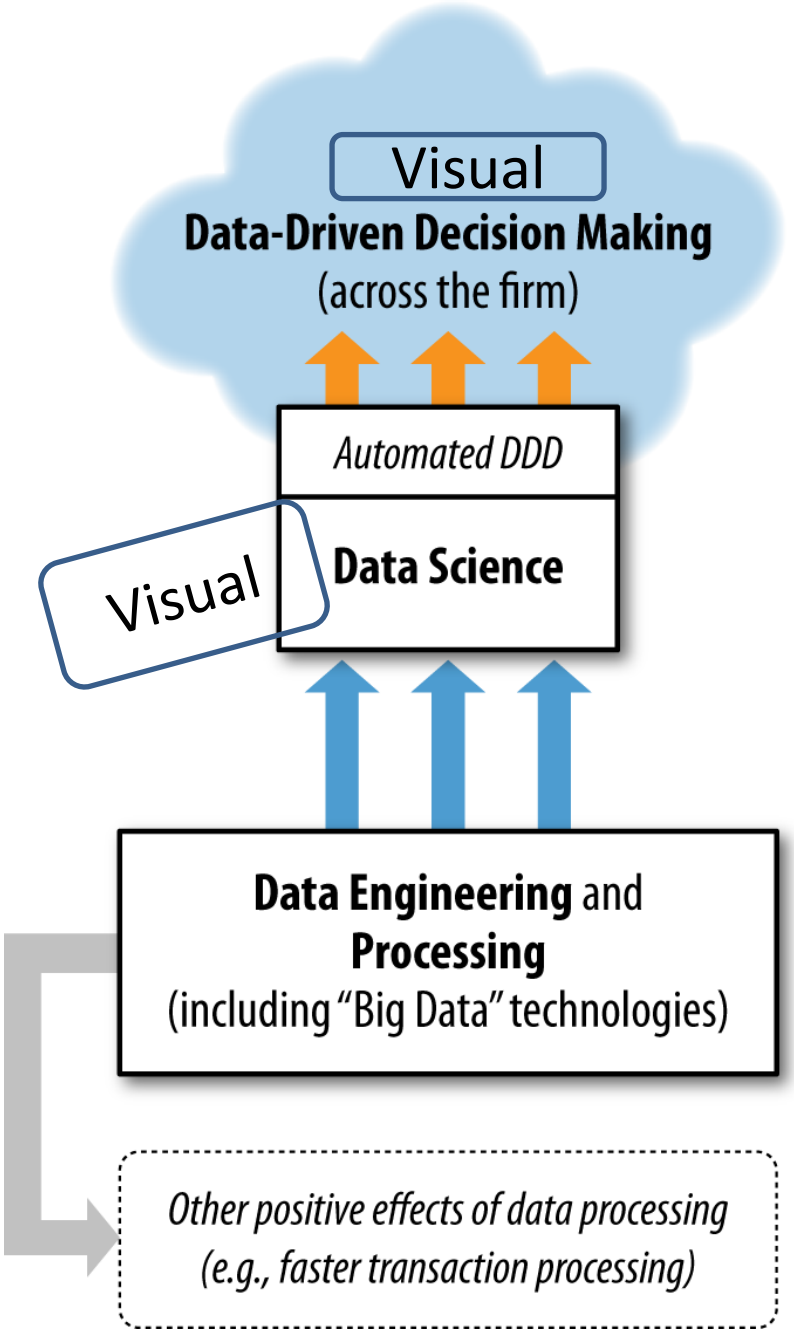
- Able to engage with senior management
- Story telling skills
- Translate data-driven insights into decisions and actions
- Visual art design
- R packages like ggplot or lattice
- Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Marketing
DISTILL
(c) Krzysztof Z

VDDD =

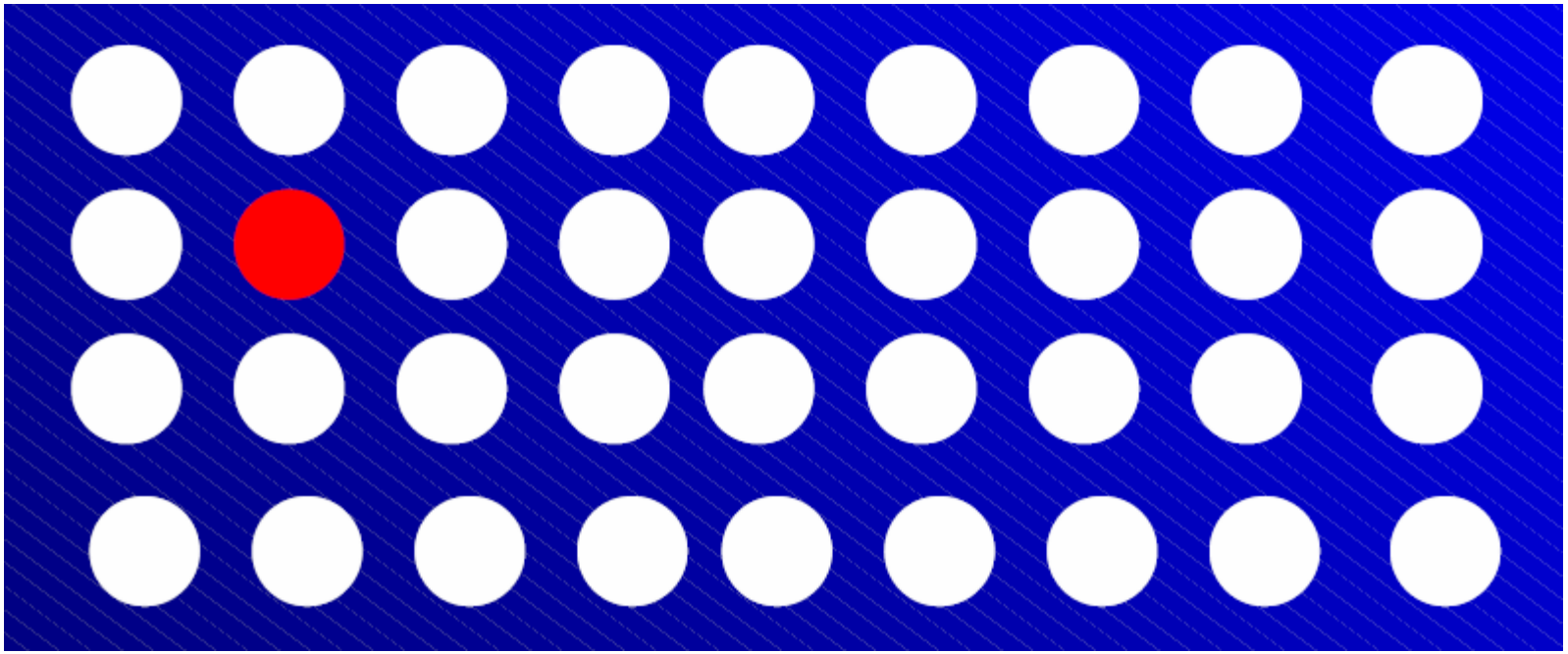
Make decisions based on data

- not purely on intuition and long business experience
- use a combination of these



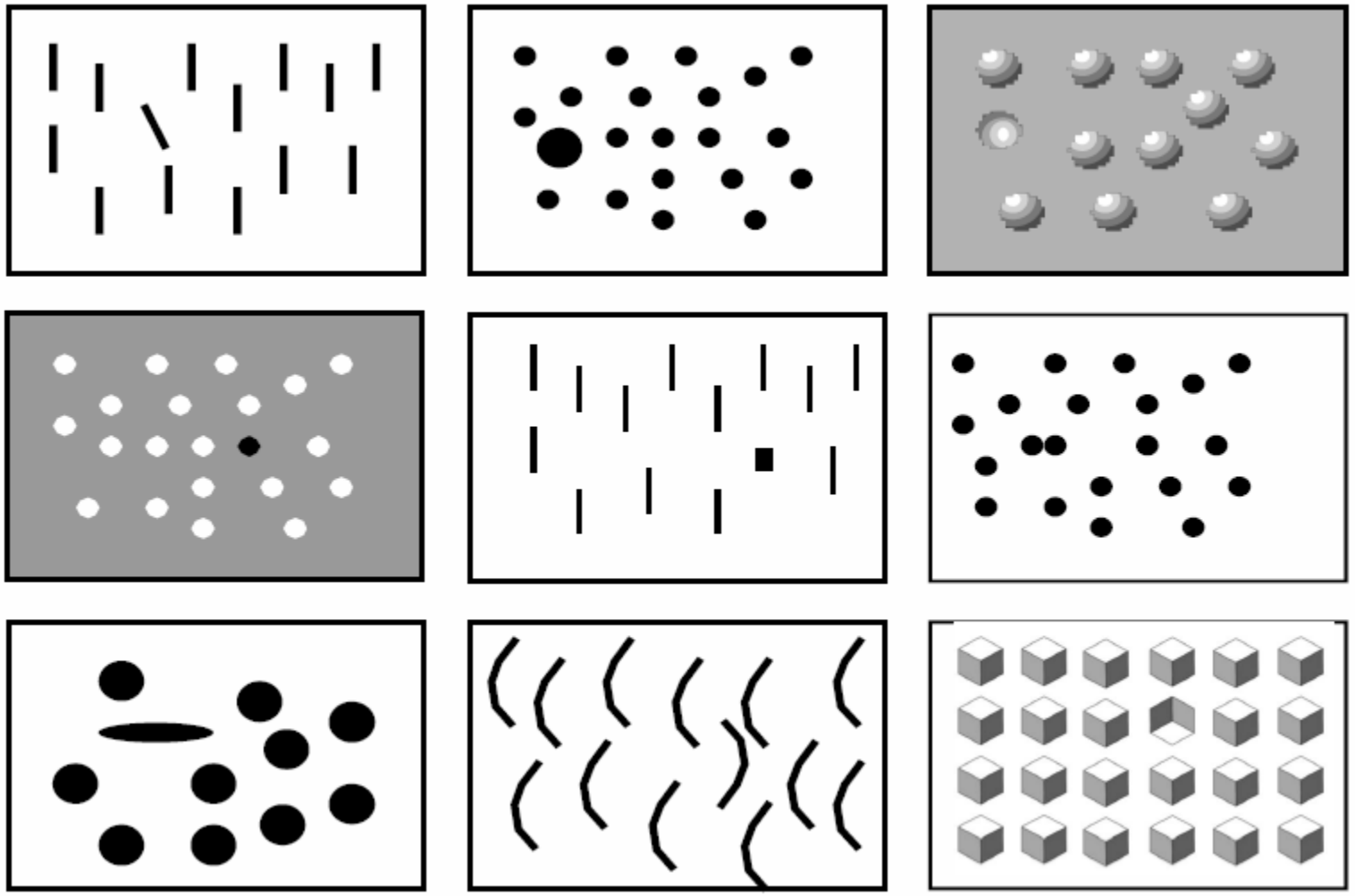
Visualization Can Be Beautiful

Visualization Is Fast

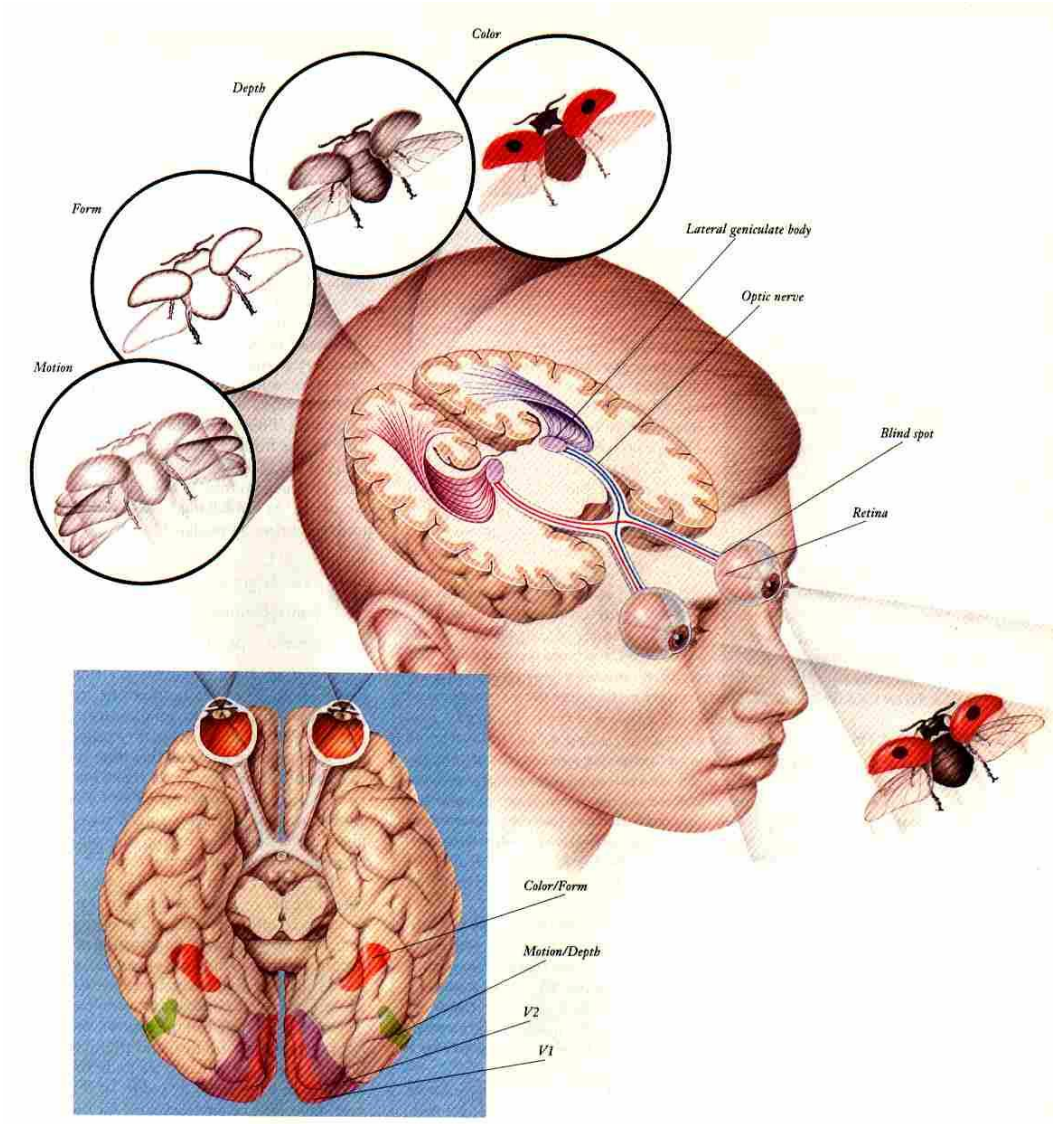


< 200 ms to recognize the red dot

Visualization Is Fast

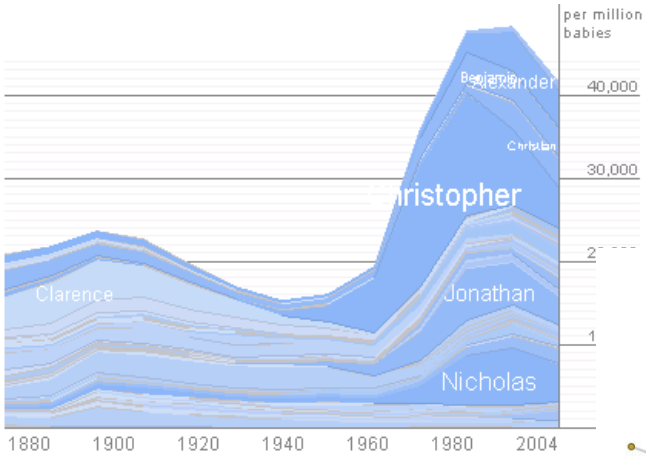


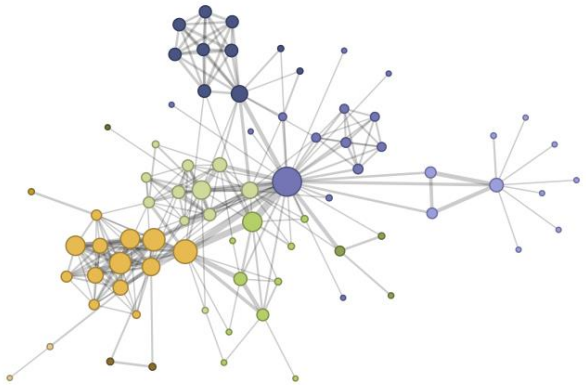
Vision Is Massively Parallel

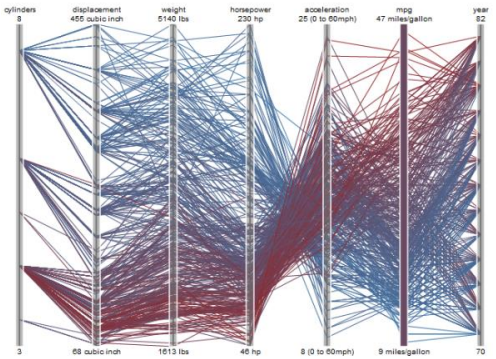


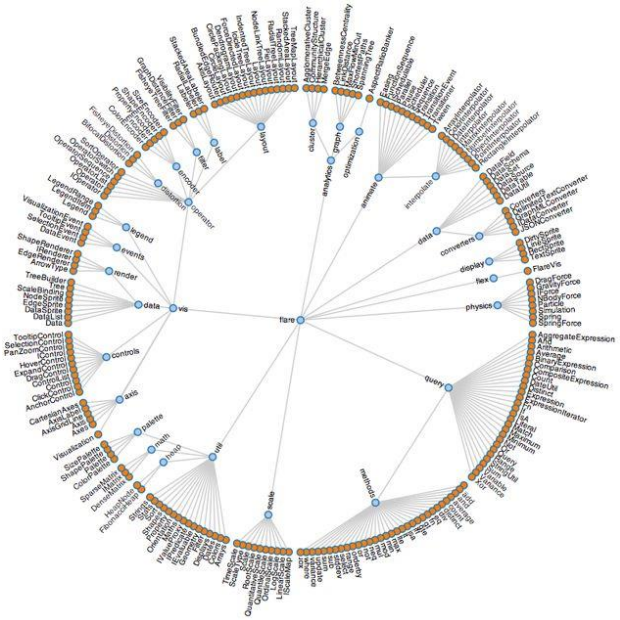
more than 50% of the brain

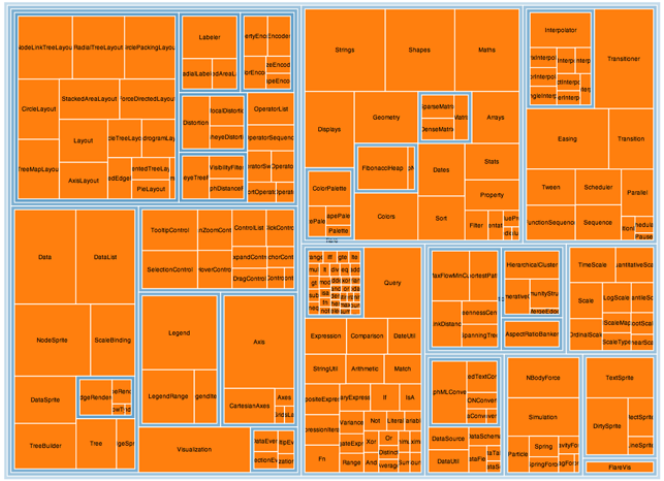
Visualization Can Be Beautiful

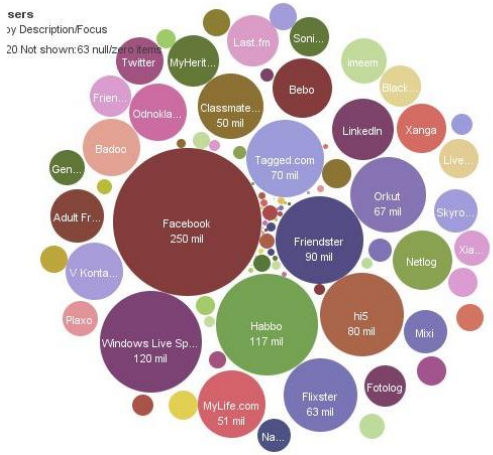










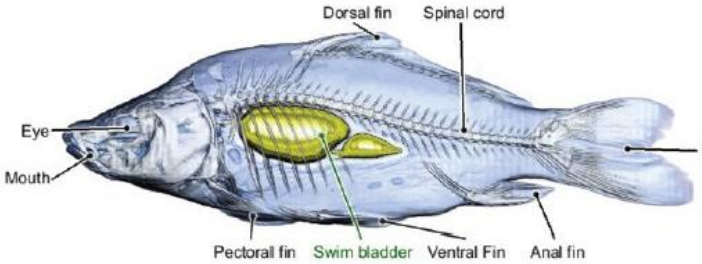


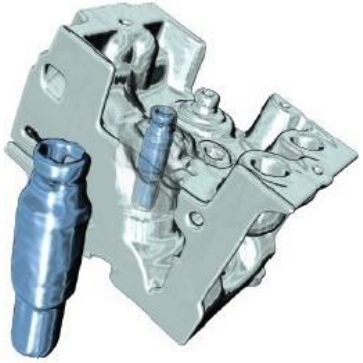
Visualization Can Be Interactive

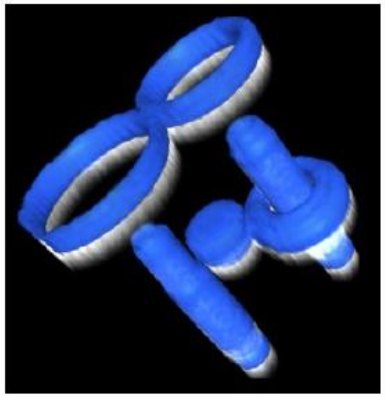
[D3 Demo]

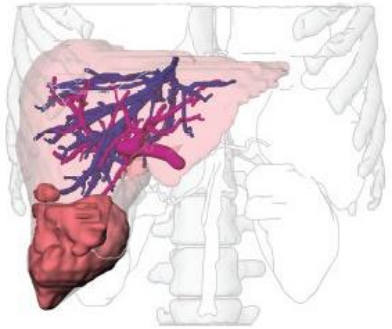
Visualization Has a Long History

Visualization Can be Inspired by Art

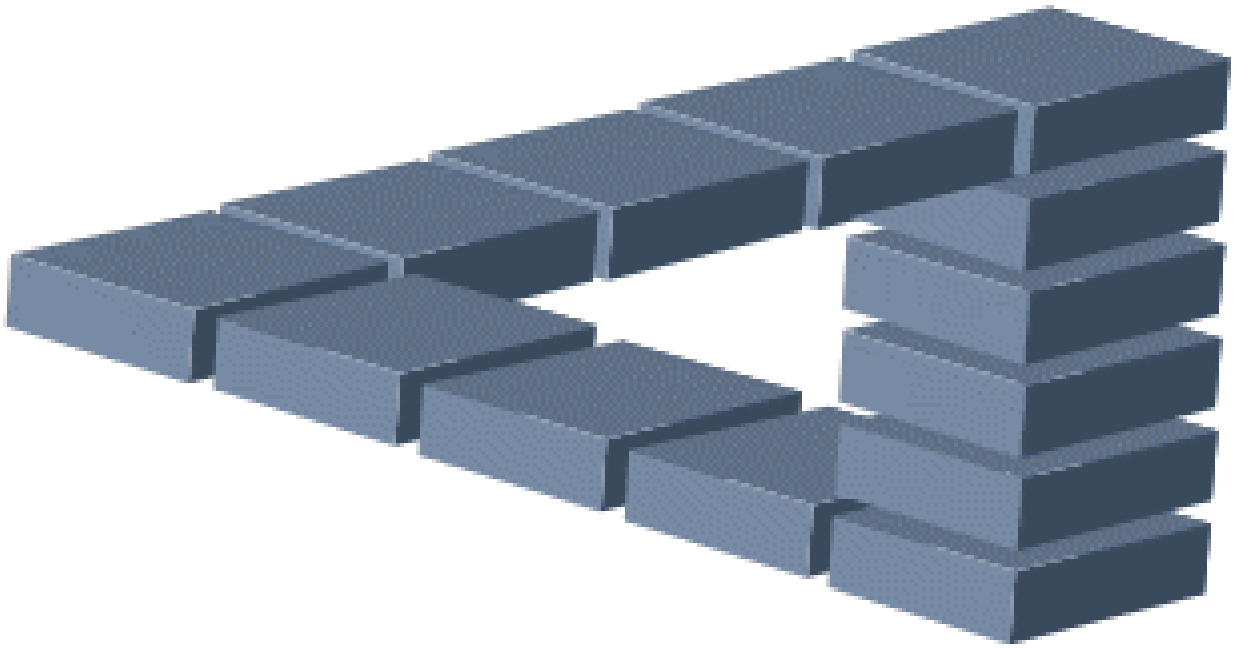






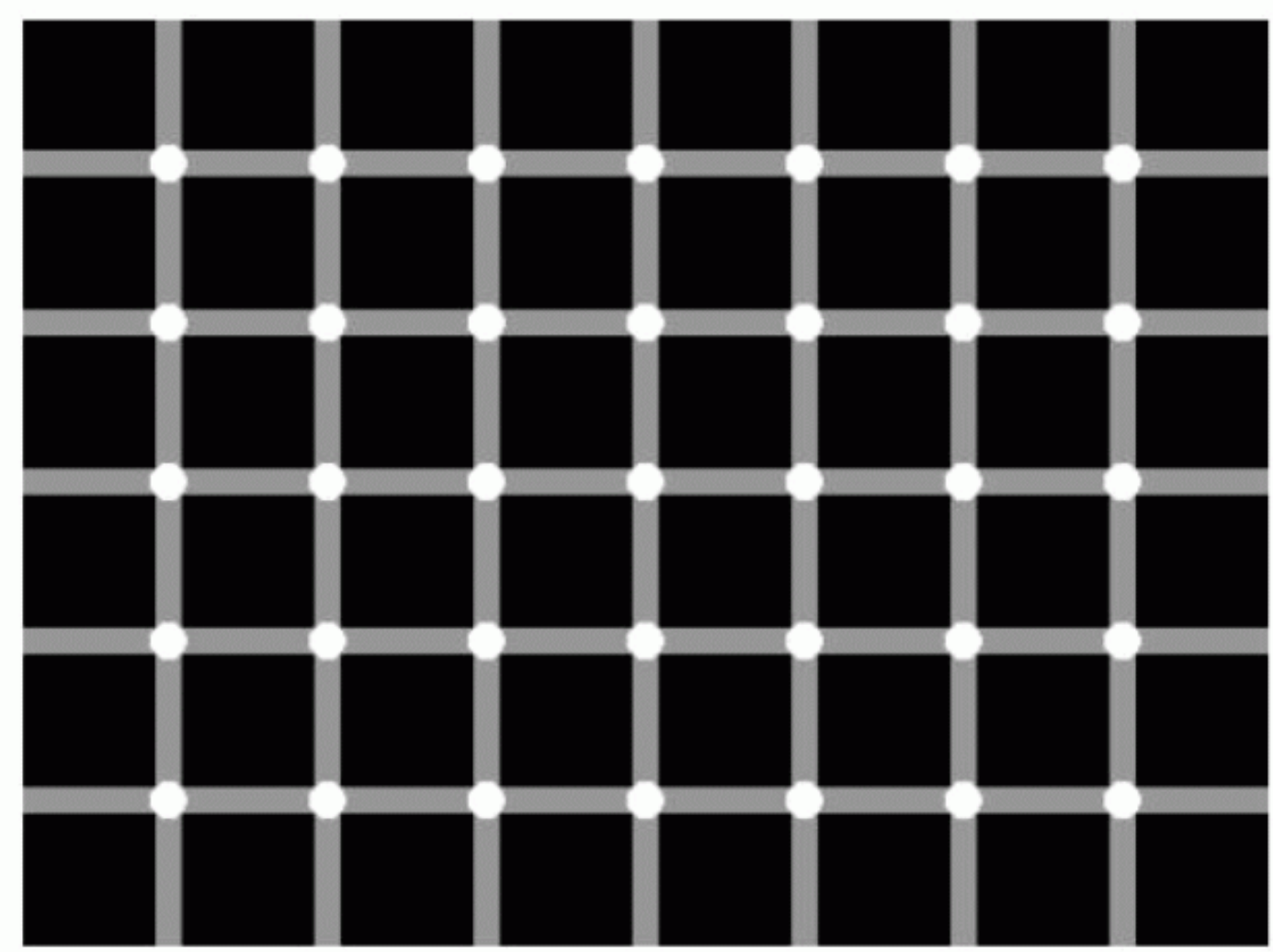


Visualization Can be Deceptive



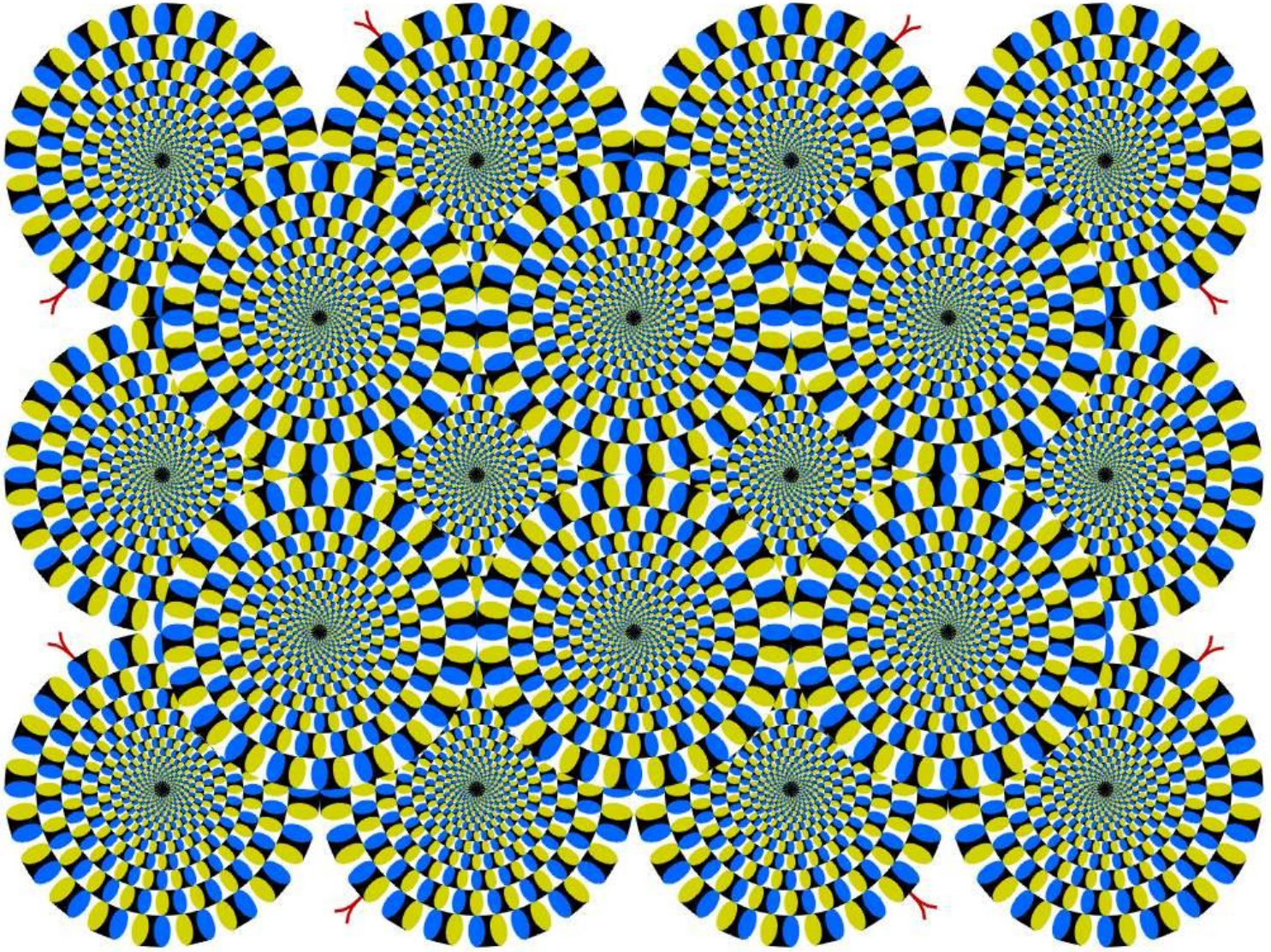
Visualization Can be Deceptive

Visualization Can be Deceptive

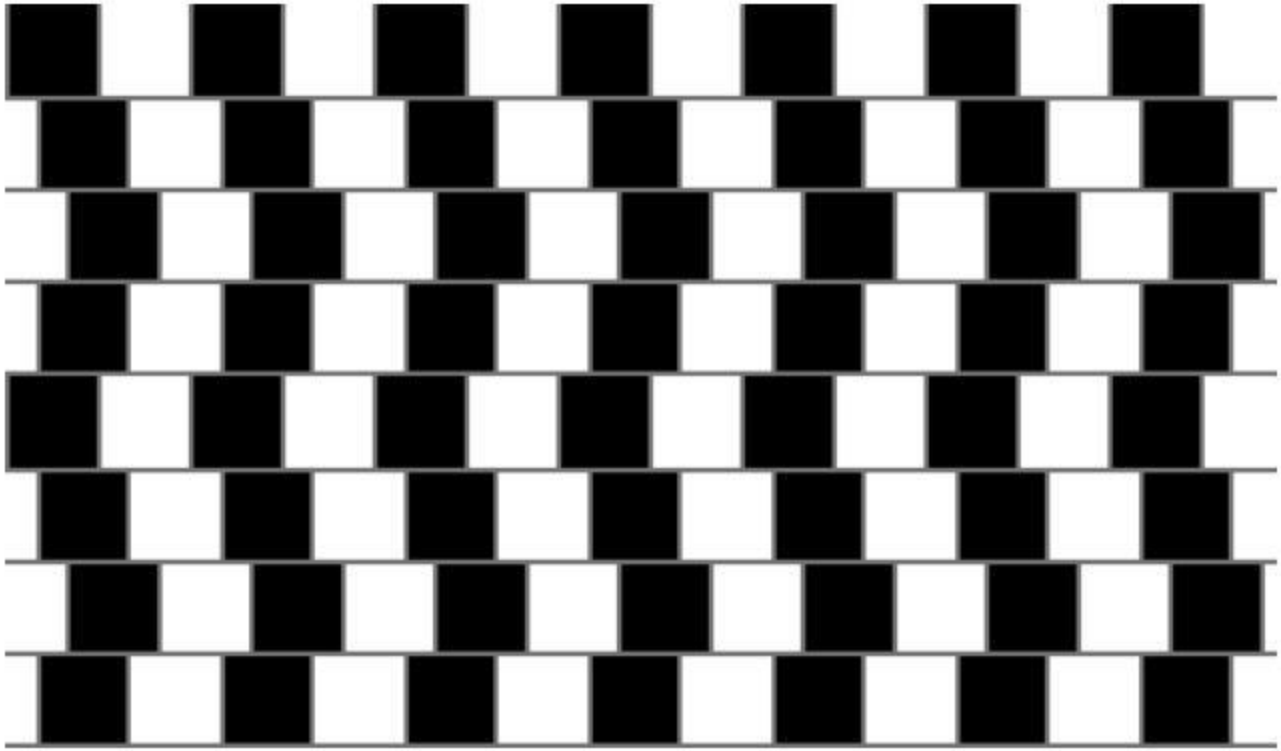


Count the number of black dots

Visualization Can be Deceptive

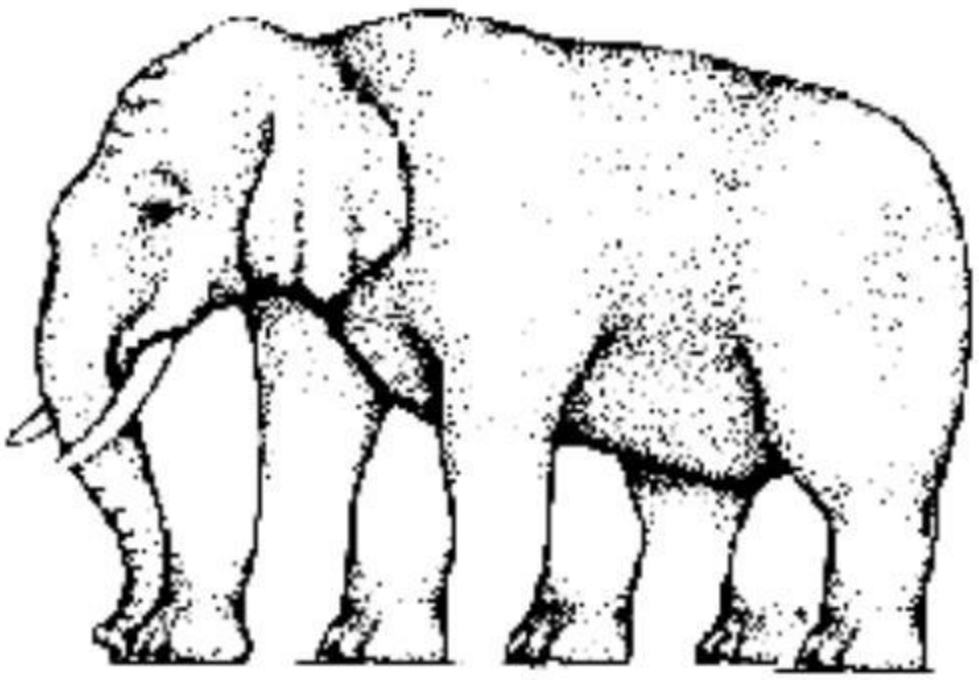


Visualization Can be Deceptive



Are the horizontal lines parallel or do they slope?

Visualization Can be Deceptive

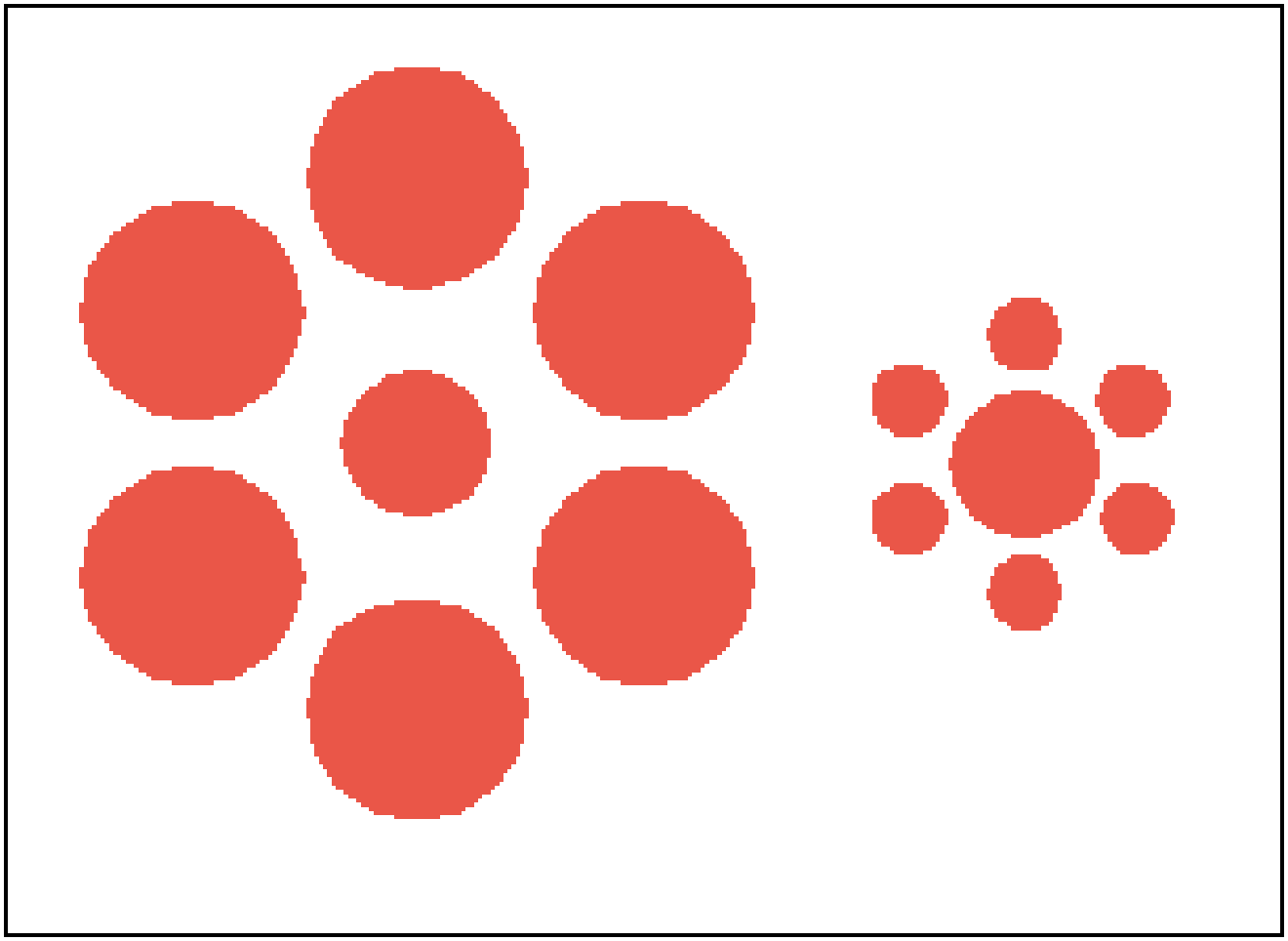


How many legs does this elephant have?

Visualization Can be Deceptive

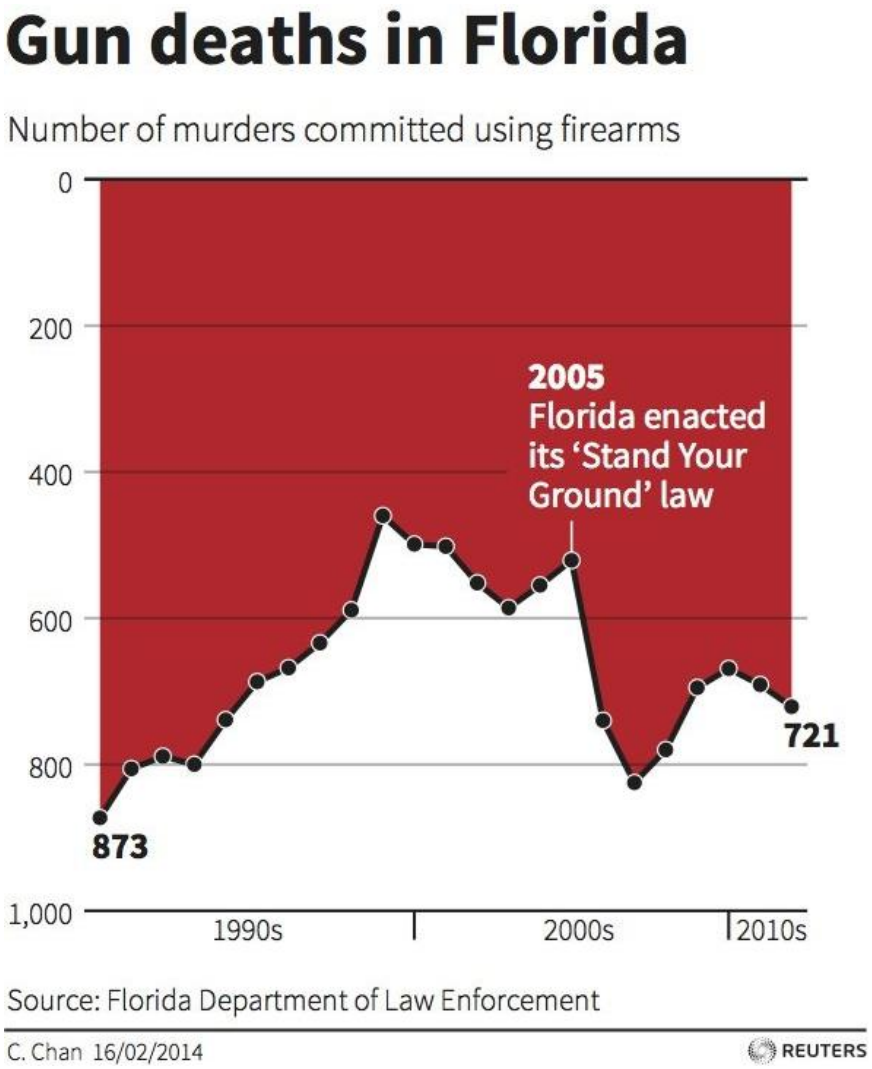
Julian Beever

Visualization Can be Deceptive



Which circle in the middle is bigger?

Visualization Can be Deceptive



The Power of The Visual System

The human visual system is not perfect, but it's extremely powerful

Vision is an integral part of life

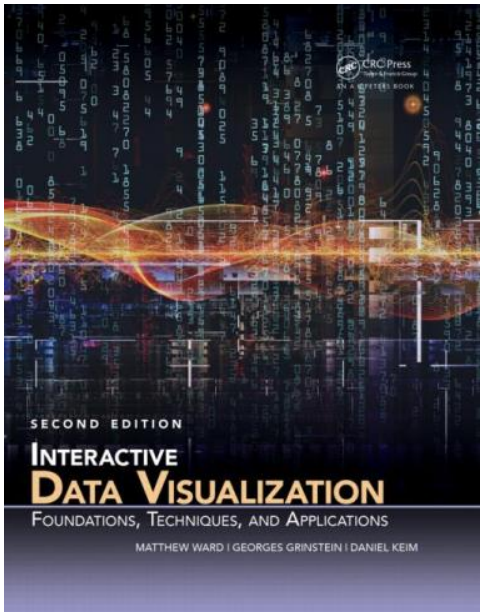
Vision is the gateway to higher-level regions of the brain

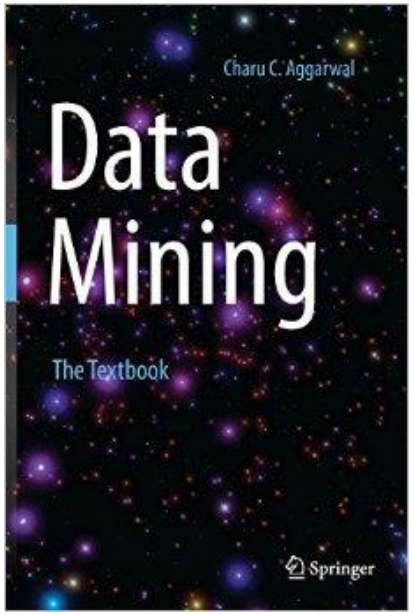
Exploit this fast and powerful processor for

- complex data analyses, creative tasks, communicating ideas

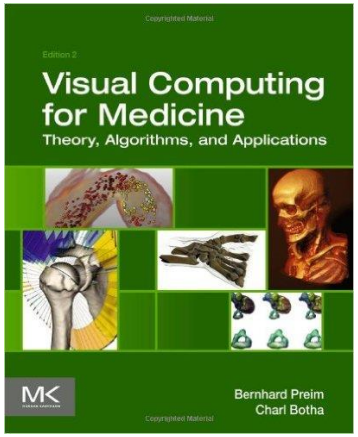
→ The science of visualization and visual analytics

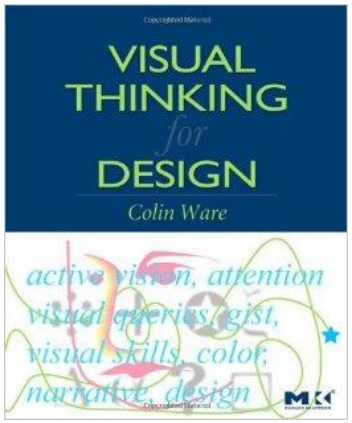
Text Books

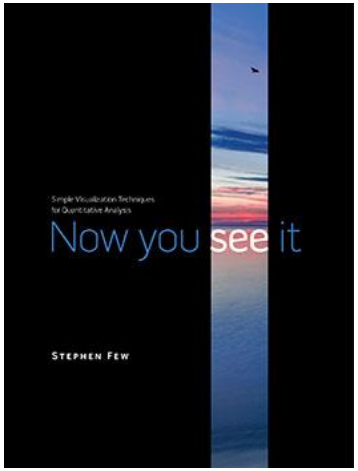


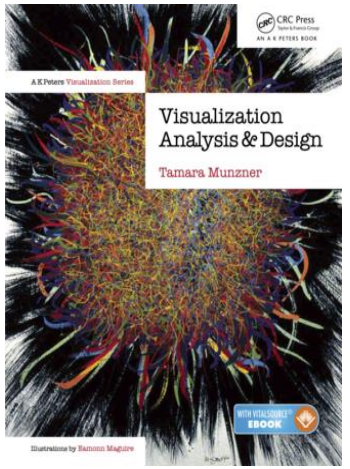


Required









Optional

Tentative Schedule

Lecture	Topic	Projects
1	Intro, schedule, and logistics	
2	Applications of visual analytics, basic tasks, data types	
3	Introduction to D3, basic vis techniques for non-spatial data	
4	Data assimilation and preparation	Project #1 out
5	Data assimilation and preparation	
6	Bias in visualization	
7	Data reduction and dimension reduction	
8	Visual perception	Project #2(a) out
9	Visual cognition	
10	Visual design and aesthetics	
11	Cluster analysis: numerical data	
12	Cluster analysis: categorical data	Project #2(b) out
13	High-dimensional data visualization	
14	Dimensionality reduction and embedding methods	
15	Principles of interaction	
16	Midterm #1	
17	Visual analytics	Final project proposal call out
18	The visual sense making process	
19	Maps	
20	Visualization of hierarchies	Final project proposal due
21	Visualization of time-varying and time-series data	
22	Foundations of scientific and medical visualization	
23	Volume rendering	Project 3 out
24	Scientific and medical visualization	Final Project preliminary report due
25	Visual analytics system design and evaluation	
26	Memorable visualization and embellishments	
27	Infographics design	
28	Midterm #2	

Course Website

http://www.cs.stonybrook.edu/~mueller/teaching/cse564/

Everything you need is there:

- syllabus
- course notes (slides) posted shortly after the lecture
- lab assignments
- course policy

There will also be (soon to be announced)

- a server for lab assignments
- piazza for online support

Grading

Projects (3): 10% each

Midterm (2) : 20% each

Capstone Project: 30%

- proposal: 5%
- prelim report: 5%
- final report and presentation: 20%

Extra credits

- will be given for projects but can only be applied in project grade

Participation

- not graded
- but I hope you will attend regularly and participate actively

For late submission policy see website

All communications will use Piazza

Final Project

Choose among two options:

A research project with some visual analytics theme

- a new technique to solve a human-in-the loop analytics task
- might even lead to a research paper for publication

A visual analytics (VA) dashboard that enables analytical tasks

- has synergy with one or more datasets you will identify
- needs to support brushing and linking and fit on the screen

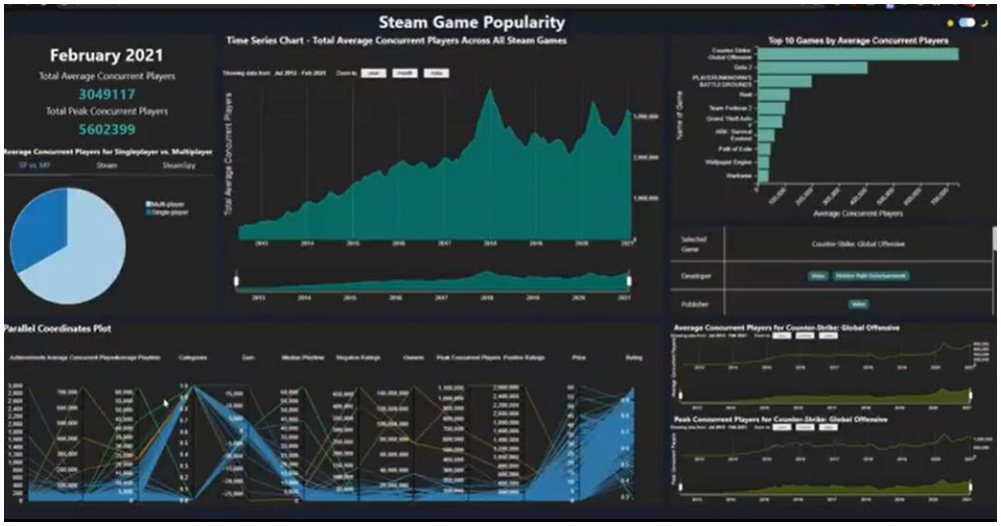
Both options will require a proposal

What's A VA Dashboard?

See a really good example on youtube

Programmed with:

- python
- html
- JavaScript
- D3 API



Your path to this:

- a dashboard is a collection of data visualizations linked together
- you will program most (but probably not all) of the individual dashboard components in lab 2 and 3
- the dashboard puts them all on one page and connect them in a meaningful way so they together can support users in interactive data analysis explorations

Not Ready? Here is What You Can Do

You will need to know html and js

- better get ready now if you do not know it

Fortunately there is a great and easy resource

- W3schools html
- W3schools JavaScript

HTML part, focus on:

- HTML Tutorial (specifically the sections *Home* to *Layout*)
- HTML Graphics
- will take you 2 days max

JavaScript part, focus on:

- JS Tutorial
- JS Objects, JS Functions, JS Async
- JS HTML DOM (Document Object Model)
- JS JSON (JavaScript Object Notation)
- will take you 2 weeks (one hour each day, ~15-20 hours total)

from here

Also Helpful: Code Editor

Several free code development environments are available

- Visual Studio Code (recommended)
- Atom
- Sublime Text

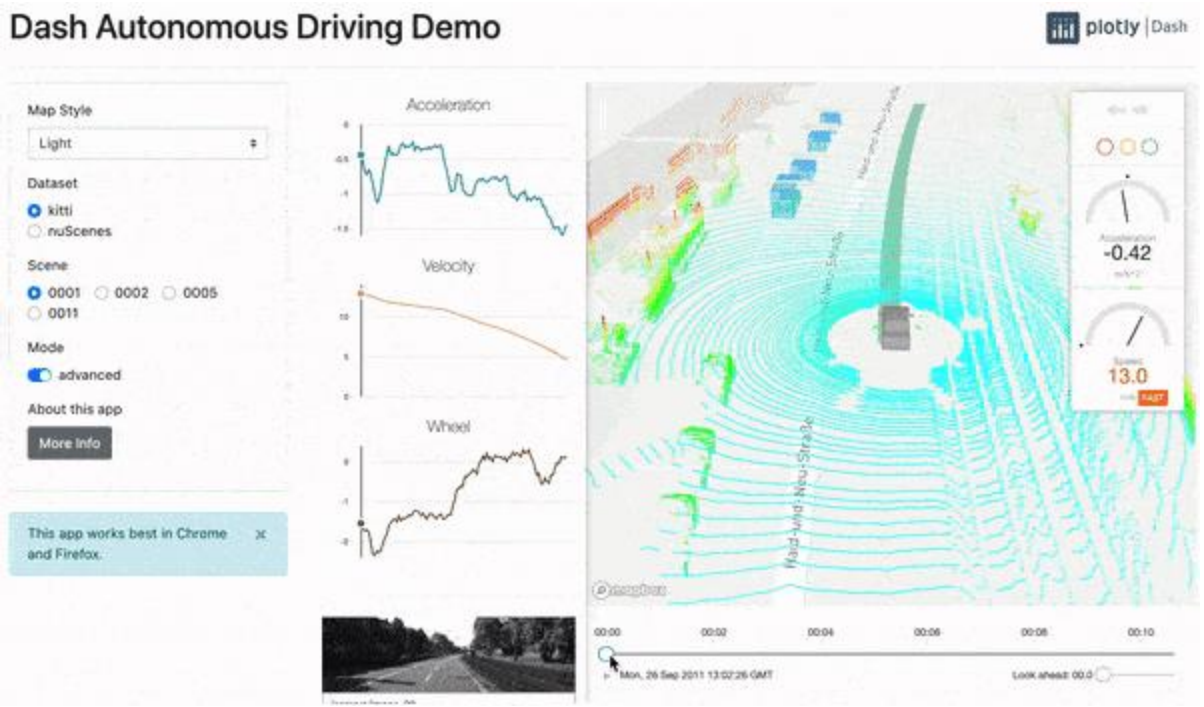
Browsers to run and develop your code

- Chrome
- Firefox
- IE and Edge are not overly suitable
- Chrome and Firefox also have panels where you can see and edit your code
- comes in handy when you want to change values of variables

Not A CS Student? This Is For You

If you are not a CS student can instead used [plotly]

- Plotly Dash lets you build/deploy analytic web apps via Python
- no JavaScript required.
- downside is that it does not support brushing and linking



Visualization Libraries etc.

D3, Vega, Vega-Lite

D3 – Data Driven Documents (we will use for this course)

- creates interactive webpages from data
- lots of creations are here

Vega (see here)

- higher-level visualization specification language on top of D3
- D3 is still more "expressive" and allows for more creative freedom

Vega-Lite (see here)

- a high-level grammar of interactive graphics
- built on top of Vega
- more concise & convenient form to author common visualizations
- supports data analytics (both data and visual transformations)
- better support for interactions

Tableau

Tableau is a leading commercial visual analytics platform

- founded in 2003 by a group of Stanford University researchers (Chris Stolte, Pat Hanrahan, and Christian Chabot)
- recently acquired by Salesforce
- goal was to make data more accessible through visualization
- key tech was VizQL – visualizes data by translating drag-and-drop actions into data queries through an intuitive interface

Example Tableau Dashboards

Account tracking

Quarterly results

Top accounts

Opportunity overview

Opportunity tracking

Marketing leads

D3 vs. Tableau

D3	Tableau
Open Source	Proprietary / Paid
Web Standards Focused	VizQL Language
Real-Time	Automated Updates but Not Real-Time
Expansive Viz Options	Limited Viz Choices*
Lots of Coding	Data to Viz in Seconds
Complex	Easy to Use
Limited Native Data Connections	Native Data Connections
Manual Calculations	Automated Calculations
Limited Data Manipulations	Strong Data Manipulations

source

Essentially, Tableau is great for expediently-developed in-house use
D3 is better for external use, real-time interactive web, and embedding into a product