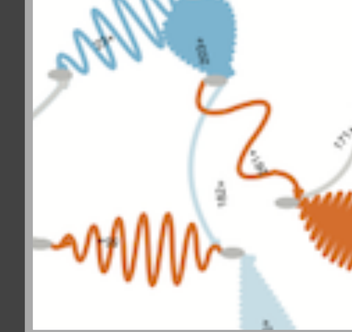
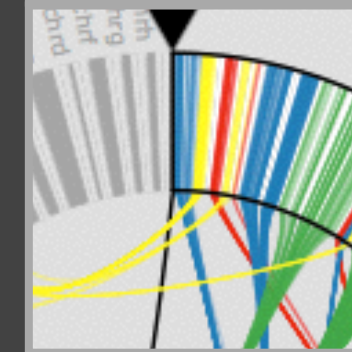
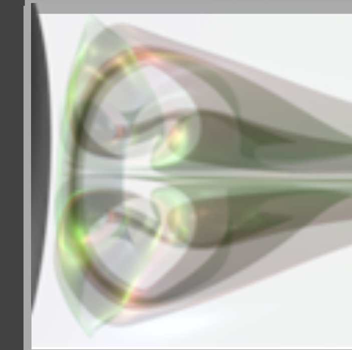
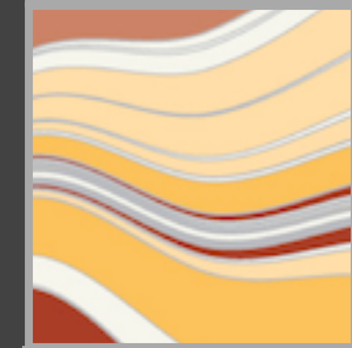
# VISUALIZATION

cs2420 | Spring 2015

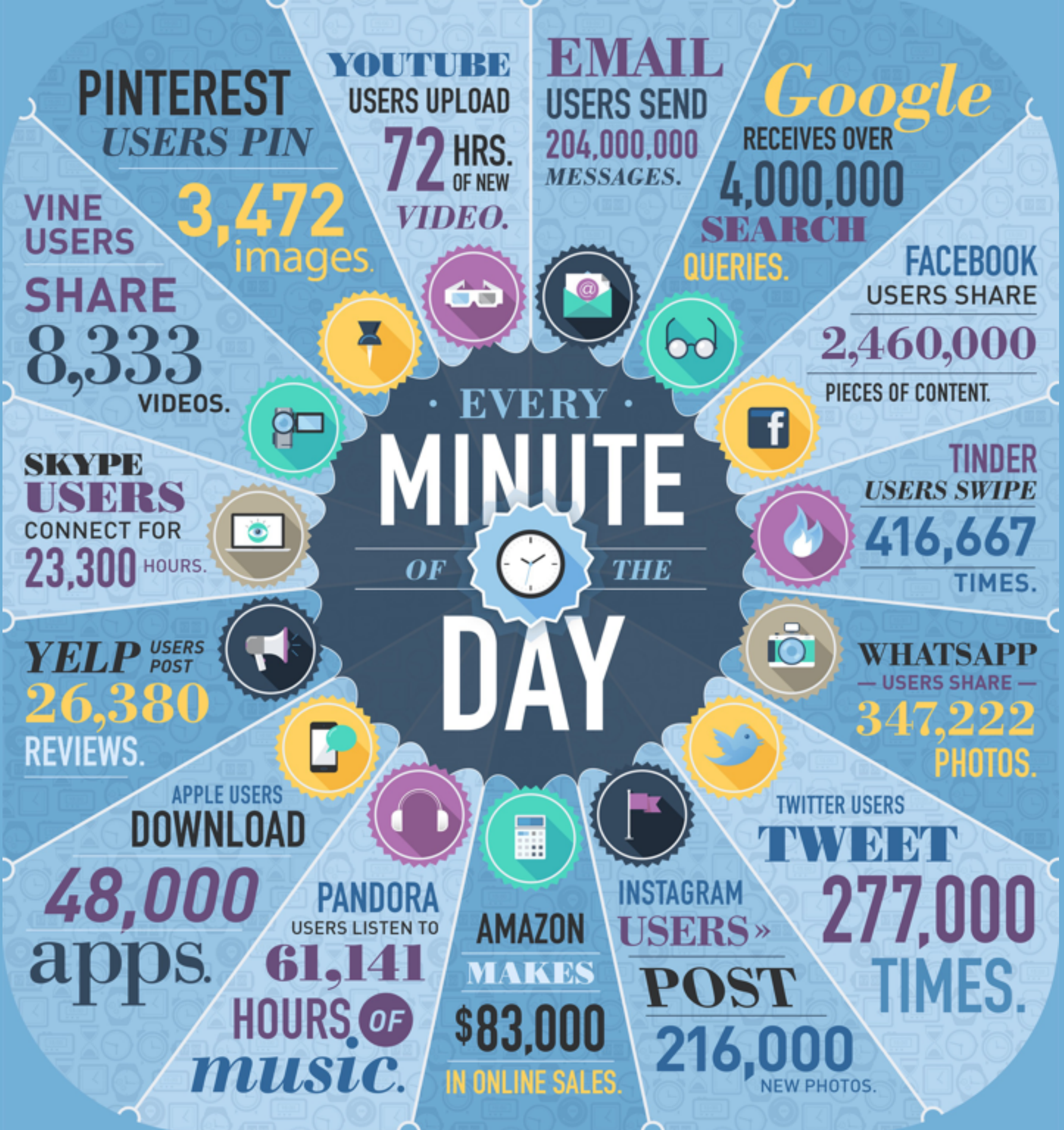

# administrivia…

-assignment 12 is due Tuesday

-upcoming lectures


# INDUSTRIAL REVOLUTION OF DATA

Joe Hellerstein, UC Berkley


EVERY MINUTE OF THE DAY
PINTEREST USERS PIN 3,472 images.
YOUTUBE USERS UPLOAD 72 HRS. OF NEW VIDEO.
EMAIL USERS SEND 204,000,000 MESSAGES.
Google RECEIVES OVER 4,000,000 SEARCH QUERIES.
FACEBOOK USERS SHARE 2,460,000 PIECES OF CONTENT.
TINDER USERS SWIPE 416,667 TIMES.
WHATSAPP — USERS SHARE — 347,222 PHOTOS.
TWITTER USERS TWEET 277,000 TIMES.
INSTAGRAM USERS » POST 216,000 NEW PHOTOS.
AMAZON MAKES $83,000 IN ONLINE SALES.
PANDORA USERS LISTEN TO 61,141 HOURS OF music.
APPLE USERS DOWNLOAD 48,000 apps.
YELP USERS POST 26,380 REVIEWS.
SKYPE USERS CONNECT FOR 23,300 HOURS.
VINE USERS SHARE 8,333 VIDEOS.

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

Hal Varian, Google's Chief Economist

1440

1150

1350

8:1

1440

25:1

1150

1350

12:1

8:1

1440

$56,000

25:1

$71,000

1150

1350

12:1

$84,000

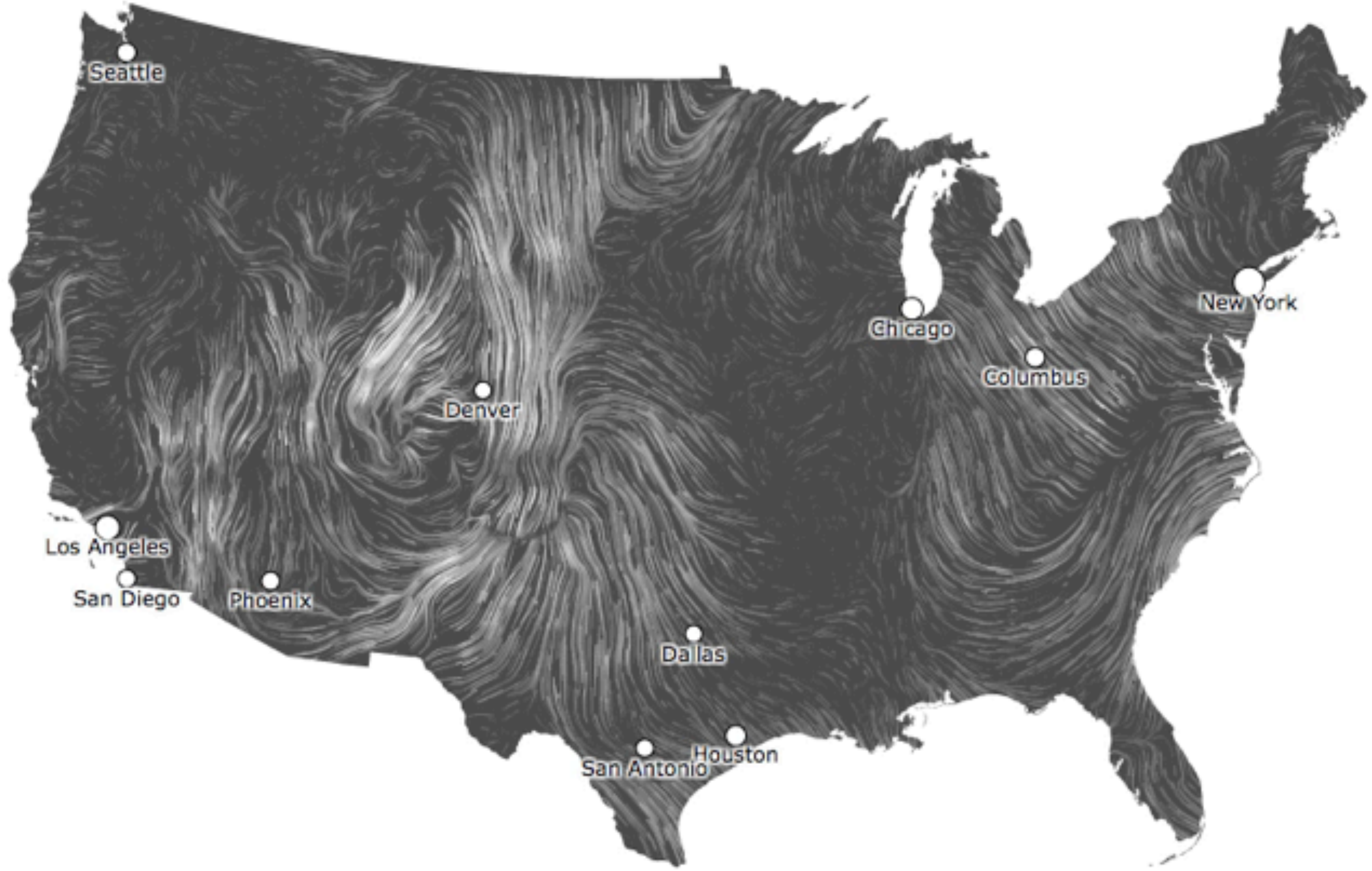
22,000
12,500
8:1
30%
11.5
1440
94%
4.5
13
8%
4,200
$56,000
25:1
1%
8
650
82%
$9.5M
$71,000
1150
33%
1210
45%
5%
3,500
45
44,000
28
1350
12:1
1,248
110
15
4
$300M
16
$5.5M
$84,000
25%
985
250
37
98%
1095

22,000
12,500
8:1
30%
1440
11.5
94%
4.5
13
8%
4,200
$56,000
25:1
8
650
1%
82%
$9.5M
$71,000
1150
33%
1210
45%
3,500
45
5%
44,000
28
1350
110
12:1
1,248
15
4
$300M
16
$5.5M
$84,000
25%
985
250
37
98%
1095

.
.
.
47
48
49
.
.
.

22,000
12,500
8:1
30%
1440
11.5
94%
4.5
13
8%
4,200
$56,000
25:1
8
650
1%
82%
$9.5M
$71,000
1150
33%
1210
45%
3,500
45
5%
44,000
28
1350
12:1
110
1,248
15
4
$300M
16
$5.5M
$84,000
25%
985
250
37
98%
1095

•
•
•
47
48
49
•
•
•





sources: AbyssExplorer, hint.fm/wind, Wordle

# visualization
uses perception to point out interesting things.

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPSNDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

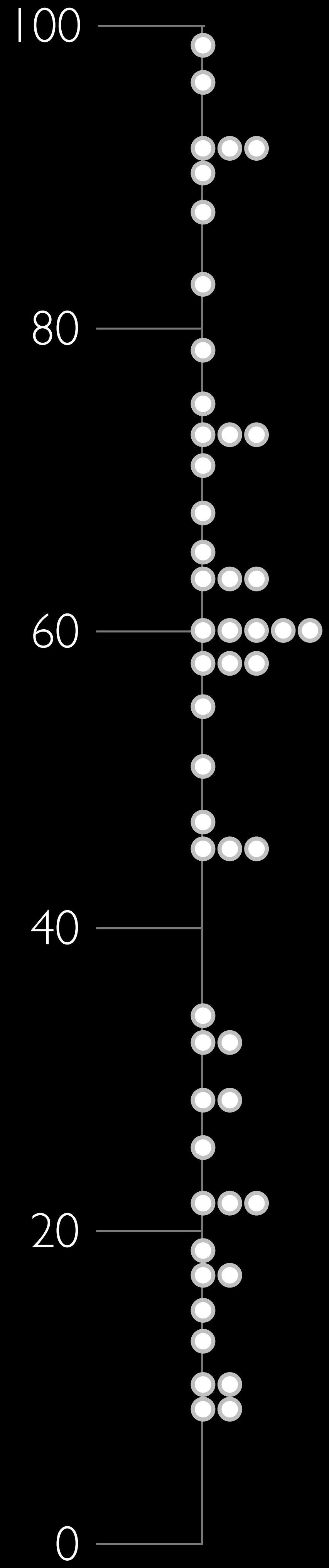
MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPSNDRCEQDIFPS
GHLMFHKMVPSTWYACEQTWRN

**visualization**
uses pictures to enhance working memory.

| | | |
|---|---|---|
| 15 | 19 | 60 |
| 33 | 11 | 75 |
| 57 | 34 | 79 |
| 18 | 51 | 92 |
| 73 | 22 | 13 |
| 71 | 60 | 22 |
| 17 | 10 | 68 |
| 73 | 18 | 55 |
| 65 | 46 | 29 |
| 60 | 73 | 22 |
| 46 | 92 | 97 |
| 10 | 58 | 46 |
| 57 | 17 | 83 |
| 26 | 99 | 33 |
| 88 | 92 | 60 |
| 91 | 29 | 57 |
| 96 | 12 | 47 |

**given these 50 numbers . . .**

. . . what number appears most often?




**given these 50 numbers . . .**

. . . what number appears most often?


The "Door" Study - YouTube
http://www.youtube.com/watch?v=FWSxSQsspiQ&list=UUoUA-CpKaFCCV2Uz__qNJZw&index=8&feature=plcp
Google
Apple Yahoo! Google Maps YouTube Wikipedia News (436) Popular Google Scholar
YouTube
Browse Movies Upload
Create Account Sign In
The "Door" Study
profsimons Subscribe 10 videos
Copyright (c) 1998 Simons & Levin. All rights reserved.
0:20 / 1:37
360p
Like Add to Share
260,166
Uploaded by profsimons on Mar 13, 2010
This video shows footage from a 1998 study by Daniel Simons and Daniel Levin in which a participant fails to notice when the person he is talking to is replaced by someone else. The study was among the first to demonstrate that the phenomenon of "change blindness" can occur outside the laboratory.
249 likes, 5 dislikes
As Seen On:
DefensiveCarry Concea...
Show more
Top Comments
This explains a lot about one of my ex-boyfriends. "If it looks like a girl, and it feels like a girl, and it smells like a girl, and it acts like a girl... it must be my girlfriend."
ataaah 1 year ago 141
Change Blindness 1
by yeblind
36,759 views
Featured Video
0:22
Gradual Change Test 1
by profsimons
34,443 views
1:20
Change Blindness
by trutapes
25,498 views
5:57
Test Your Awareness.....
by beepsquick
43,847 views
1:34
Perception of beauty
by andreic27
92,589 views
1:14
Amazing Fire & Gas Trick!
by brusspup
1,078,932 views
1:20
Try To Watch This Without Laughing Or
by 88ownsnascar
2,042,315 views
1:15
How much is:
75 + 26
1:29
Sociopath Test
by Daanando
213,997 views
Awareness Test
by J0EKthePANDA

**vi·su·al·i·za·tion**

***noun, plural*** -s

1. formation of mental visual images

2. the act or process of interpreting in visual terms or of putting into visible form

"Computer-based **visualization** systems provide visual representations of datasets intended to help people carry out tasks more effectively."

Prof. Tamara Munzner

# ANALYZE DATA

# THE CHALLENGER DISASTER

source: Wikipedia

## HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

| | | Cross Sectional View | | | Top View | | |
|---|---|---|---|---|---|---|---|
| | SRM No. | Erosion Depth (in.) | Perimeter Affected (deg) | Nominal Dia. (in.) | Length Of Max Erosion (in.) | 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 |

Oct 30, 1985

AFT

Jan 85

July

*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.

| MOTOR | O-RING |
|---|---|
| DM-4 | 47 |
| DM-2 | 52 |
| QM-3 | 48 |
| QM-4 | 51 |
| SRM-15 | 53 |
| SRM-22 | 75 |
| SRM-25 | 29<br>27 |

BLOW BY HISTORY

SRM-15 WORST BLOW-BY
- O 2 CASE JOINTS (80°), (110°) ARC
- O MUCH WORSE VISUALLY THAN SRM-22

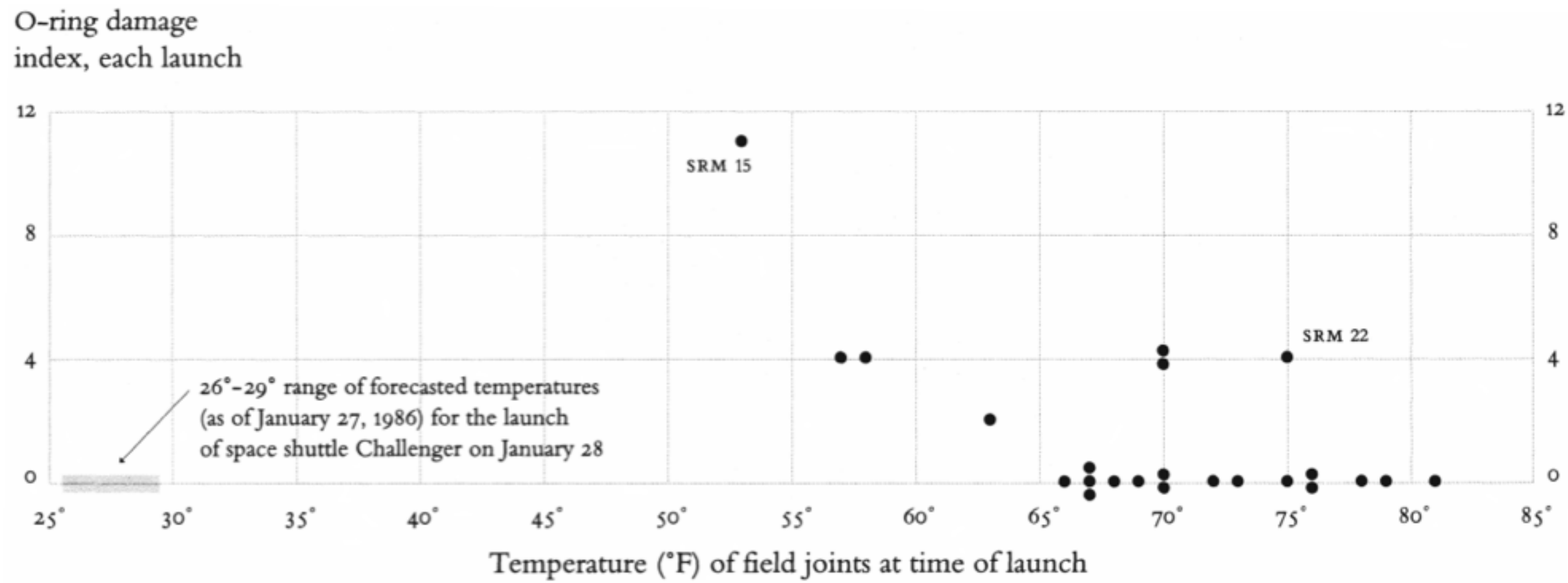
SRM 22 BLOW-BY
- O 2 CASE JOINTS (30-40°)

SRM-13A, 15, 16A, 18, 23A 24A
- O NOZZLE BLOW-BY

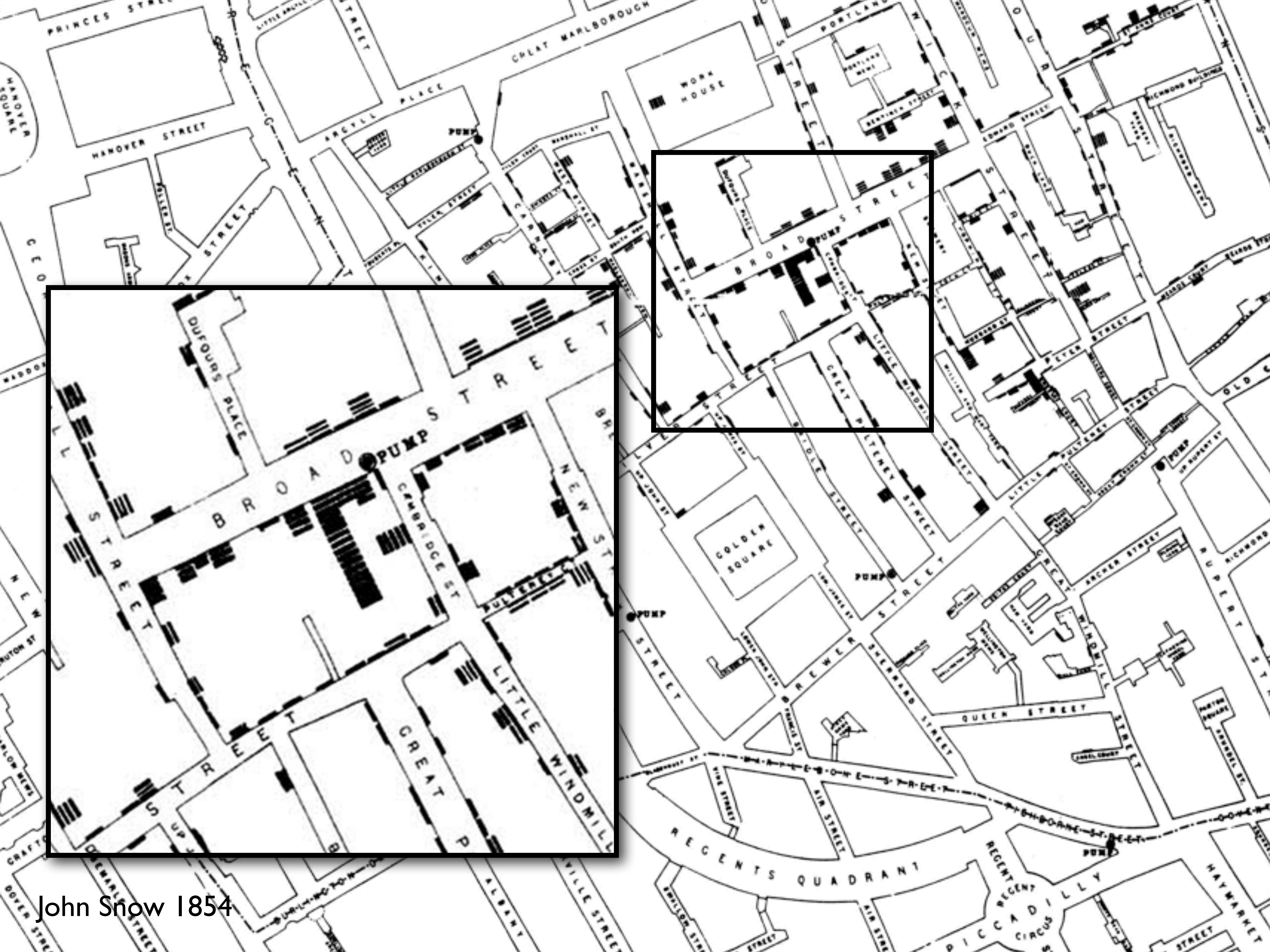
HISTORY OF O-RING TEMPERATURES
(DEGREES - F)

| MOTOR | MBT | AMB | O-RING | WIND |
|---|---|---|---|---|
| DM-4 | 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<br>27 | 10 MPH<br>25 MPH |

source: E. Tufte




source: E. Tufte

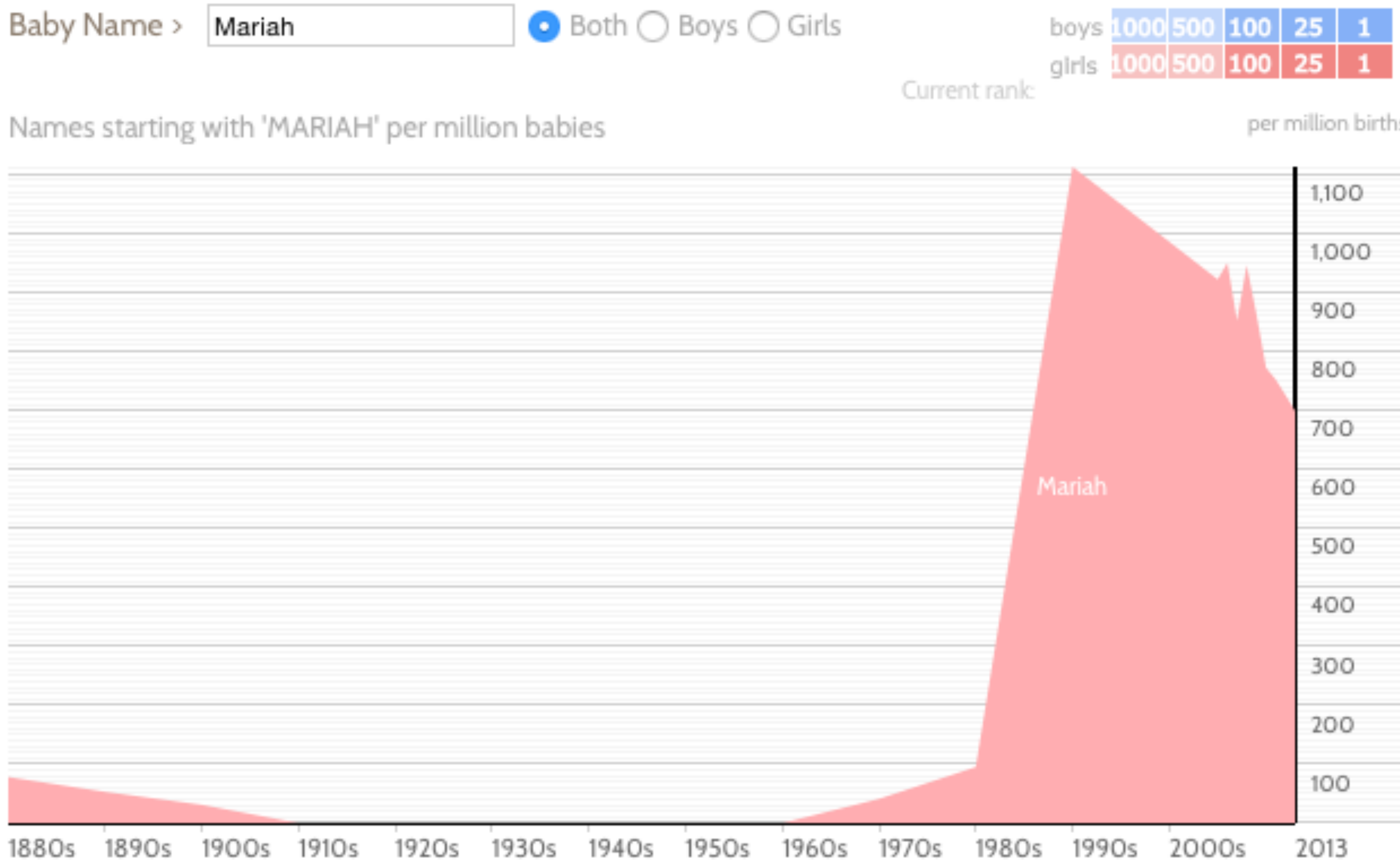


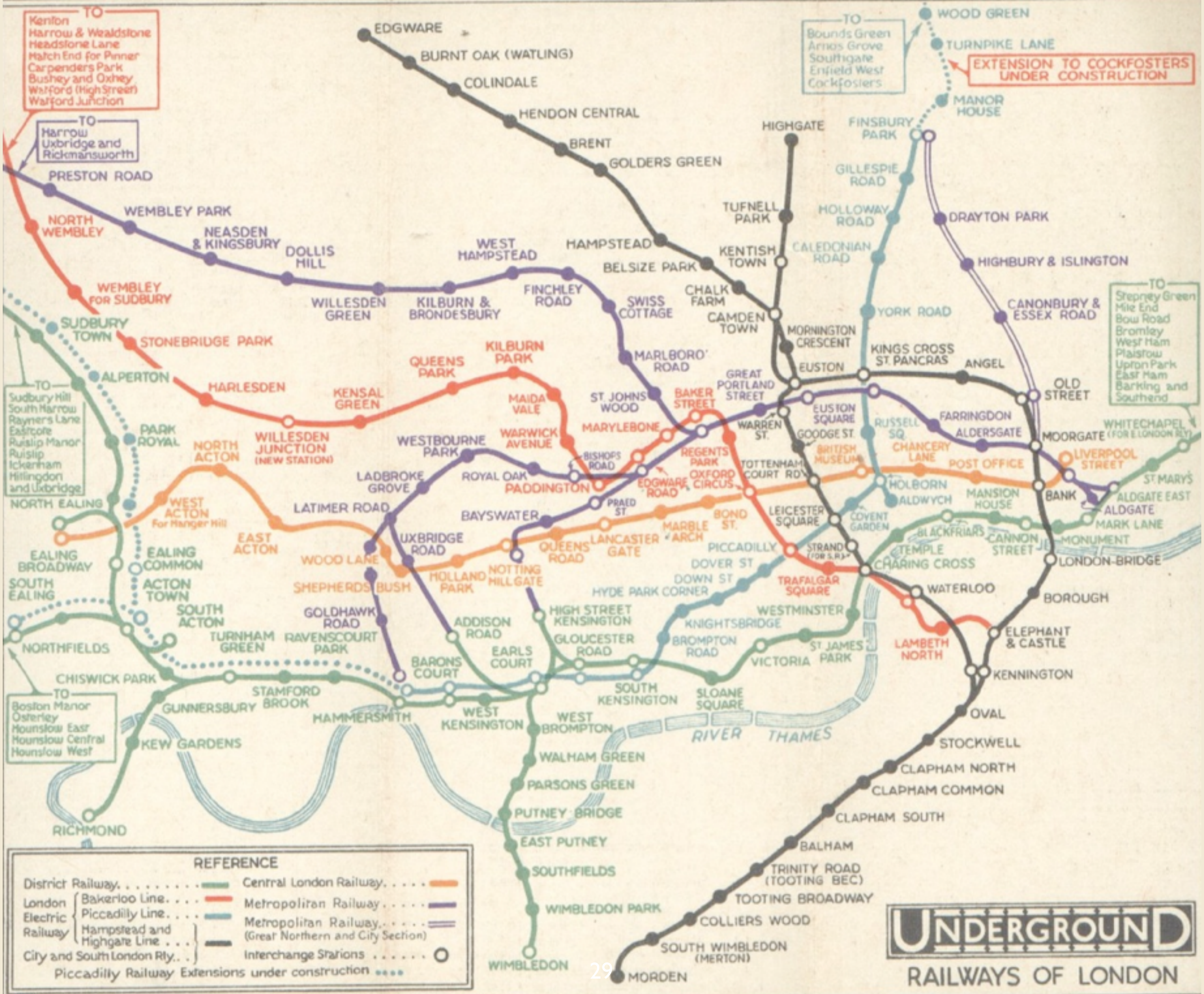

John Snow 1854

**NameVoyager**: Explore baby names and name trends letter by letter

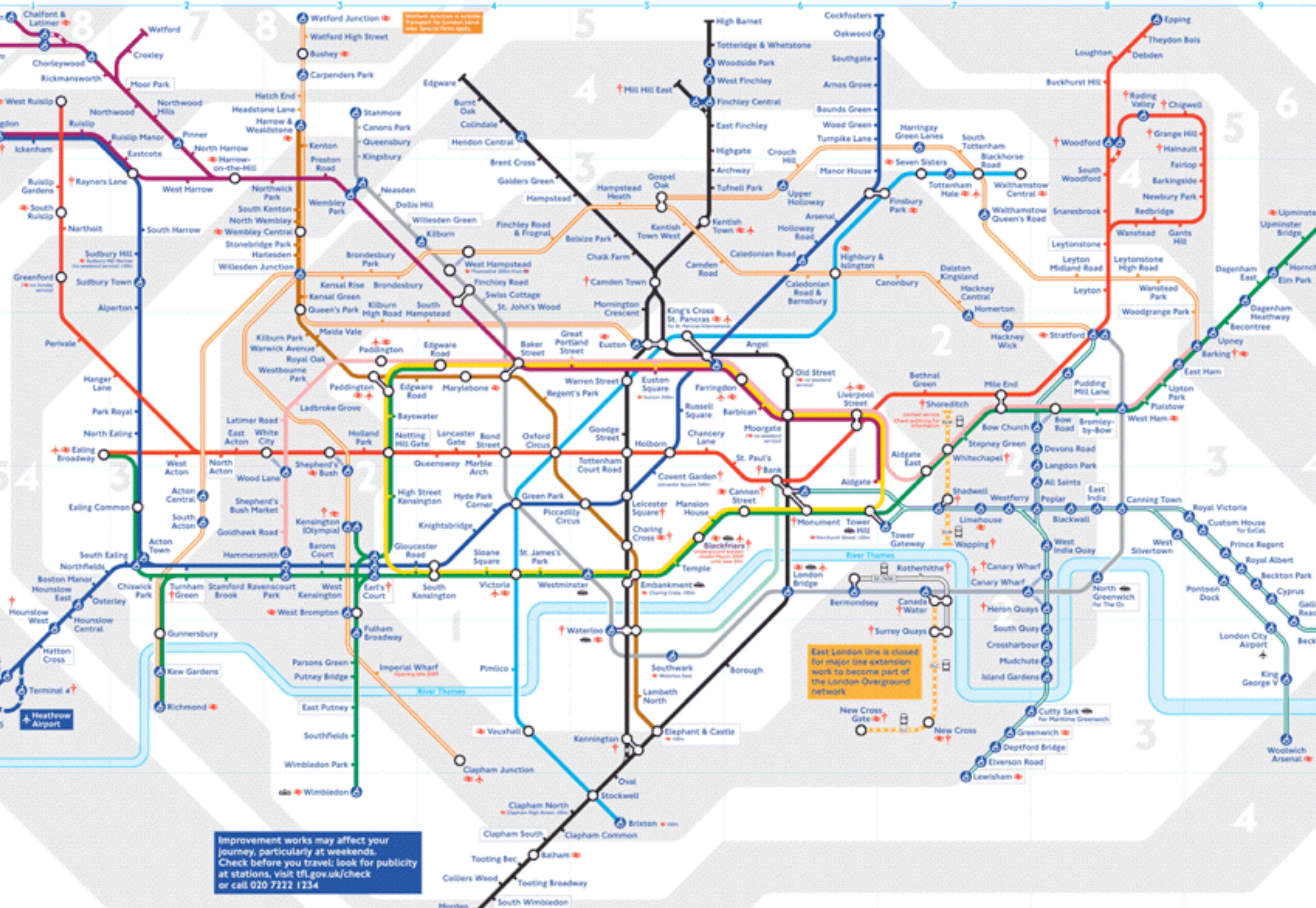
Looking for the perfect baby name? Sign up for free to receive access to our expert tools!

Baby Name › Mariah ◉ Both ○ Boys ○ Girls

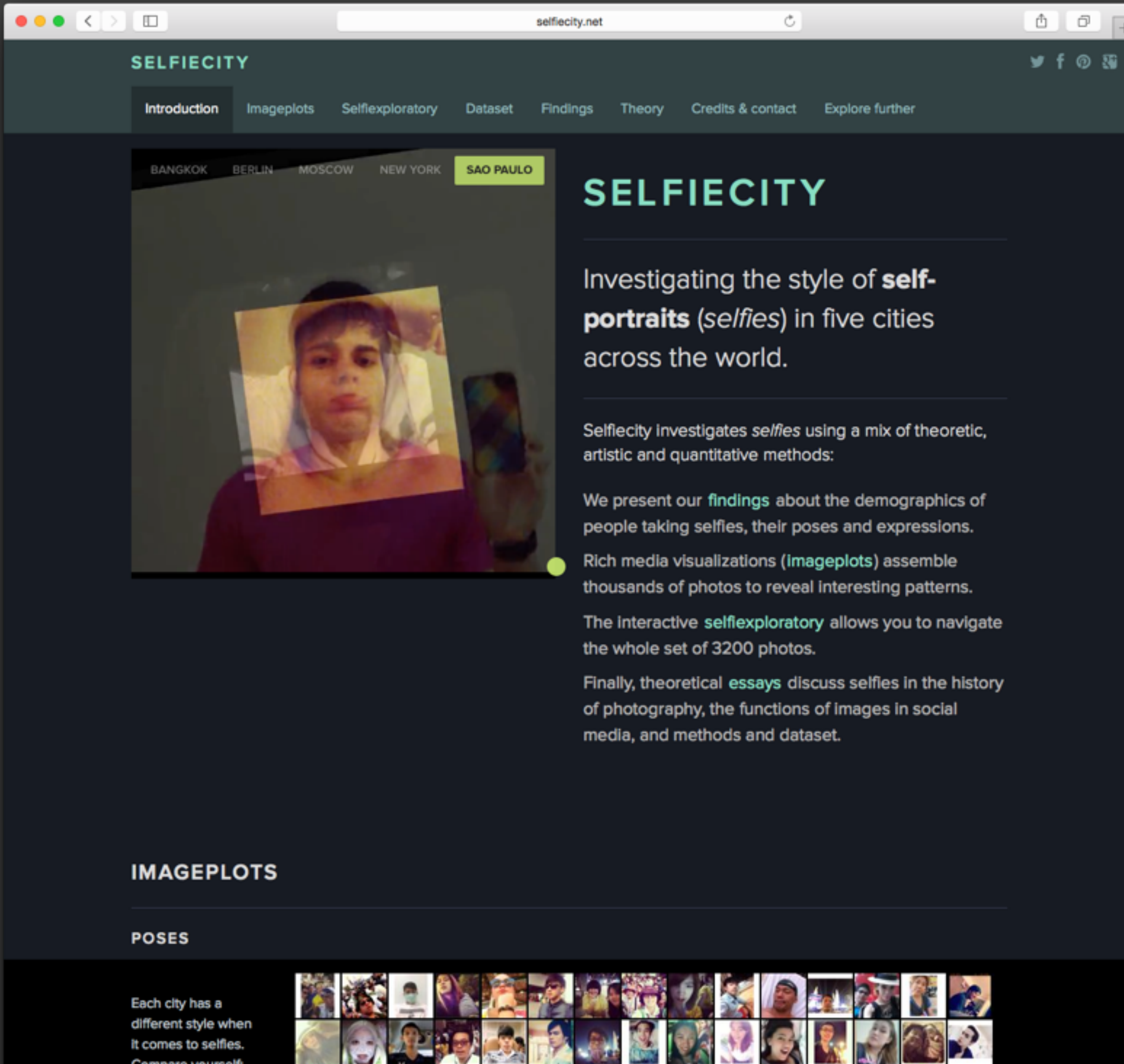




TO Kenton Harrow & Wealdstone Headstone Lane Hatch End for Pinner Carpenders Park Bushey and Oxhey Watford (High Street) Watford Junction
TO Harrow Uxbridge and Rickmansworth
EDGWARE
BURNT OAK (WATLING)
COLINDALE
HENDON CENTRAL
BRENT
GOLDERS GREEN
HAMPSTEAD
BELSIZE PARK
CHALK FARM
HIGHGATE
TUFNELL PARK
KENTISH TOWN
CAMDEN TOWN
MORNINGTON CRESCENT
EUSTON
TO Bounds Green Arnos Grove Southgate Enfield West Cockfosters
WOOD GREEN
TURNPIKE LANE
EXTENSION TO COCKFOSTERS UNDER CONSTRUCTION
MANOR HOUSE
FINSBURY PARK
GILLESPIE ROAD
HOLLOWAY ROAD
CALEDONIAN ROAD
YORK ROAD
DRAYTON PARK
HIGHBURY & ISLINGTON
CANONBURY & ESSEX ROAD
TO Stepney Green Mile End Bow Road Bromley West Ham Plaistow Upton Park East Ham Barking and Southend
KINGS CROSS ST. PANCRAS
ANGEL
OLD STREET
PRESTON ROAD
NORTH WEMBLEY
WEMBLEY PARK
NEASDEN & KINGSBURY
DOLLIS HILL
WILLESDEN GREEN
KILBURN & BRONDESBURY
WEST HAMPSTEAD
FINCHLEY ROAD
SWISS COTTAGE
MARLBORO' ROAD
ST. JOHNS WOOD
WEMBLEY FOR SUDBURY
SUDBURY TOWN
STONEBRIDGE PARK
ALPERTON
HARLESDEN
KENSAL GREEN
QUEENS PARK
KILBURN PARK
MAIDA VALE
WARWICK AVENUE
WILLESDEN JUNCTION (NEW STATION)
TO Sudbury Hill South Harrow Rayners Lane Eastcote Ruislip Manor Ruislip Ickenham Hillingdon and Uxbridge
PARK ROYAL
NORTH ACTON
WESTBOURNE PARK
BAKER STREET
MARYLEBONE
GREAT PORTLAND STREET
EUSTON SQUARE
WARREN ST.
GOODGE ST.
RUSSELL SQ.
FARRINGDON
ALDERSGATE
MOORGATE
WHITECHAPEL (FOR E. LONDON RLY.)
LIVERPOOL STREET
BRITISH MUSEUM
CHANCERY LANE
POST OFFICE
REGENTS PARK
OXFORD CIRCUS
TOTTENHAM COURT RD.
BISHOPS ROAD
ROYAL OAK
PADDINGTON
EDGWARE ROAD
PRAED ST.
LADBROKE GROVE
NORTH EALING
WEST ACTON For Hanger Hill
EAST ACTON
LATIMER ROAD
BAYSWATER
BOND ST.
MARBLE ARCH
LANCASTER GATE
QUEENS ROAD
HOLBORN
ALDWYCH
MANSION HOUSE
BANK
ST. MARY'S
ALDGATE EAST
ALDGATE
MARK LANE
MONUMENT
CANNON STREET
BLACKFRIARS
TEMPLE
LEICESTER SQUARE
COVENT GARDEN
STRAND (FOR S.R.)
CHARING CROSS
LONDON BRIDGE
EALING BROADWAY
EALING COMMON
WOOD LANE
UXBRIDGE ROAD
SHEPHERDS BUSH
HOLLAND PARK
NOTTING HILL GATE
PICCADILLY
DOVER ST.
DOWN ST.
HYDE PARK CORNER
TRAFALGAR SQUARE
WATERLOO
BOROUGH
SOUTH EALING
ACTON TOWN
SOUTH ACTON
GOLDHAWK ROAD
ADDISON ROAD
HIGH STREET KENSINGTON
KNIGHTSBRIDGE
WESTMINSTER
ELEPHANT & CASTLE
NORTHFIELDS
TURNHAM GREEN
RAVENSCOURT PARK
BARONS COURT
EARLS COURT
GLOUCESTER ROAD
BROMPTON ROAD
ST. JAMES PARK
VICTORIA
LAMBETH NORTH
KENNINGTON
CHISWICK PARK
TO Boston Manor Osterley Hounslow East Hounslow Central Hounslow West
GUNNERSBURY
STAMFORD BROOK
HAMMERSMITH
WEST KENSINGTON
SOUTH KENSINGTON
SLOANE SQUARE
OVAL
KEW GARDENS
WEST BROMPTON
RIVER THAMES
STOCKWELL
WALHAM GREEN
CLAPHAM NORTH
PARSONS GREEN
CLAPHAM COMMON
PUTNEY BRIDGE
CLAPHAM SOUTH
RICHMOND
EAST PUTNEY
BALHAM
TRINITY ROAD (TOOTING BEC)
SOUTHFIELDS
TOOTING BROADWAY
WIMBLEDON PARK
COLLIERS WOOD
SOUTH WIMBLEDON (MERTON)
WIMBLEDON
MORDEN
REFERENCE
District Railway
London Electric Railway: Bakerloo Line, Piccadilly Line, Hampstead and Highgate Line, City and South London Rly.
Central London Railway
Metropolitan Railway
Metropolitan Railway (Great Northern and City Section)
Interchange Stations
Piccadilly Railway Extensions under construction
UNDERGROUND
RAILWAYS OF LONDON

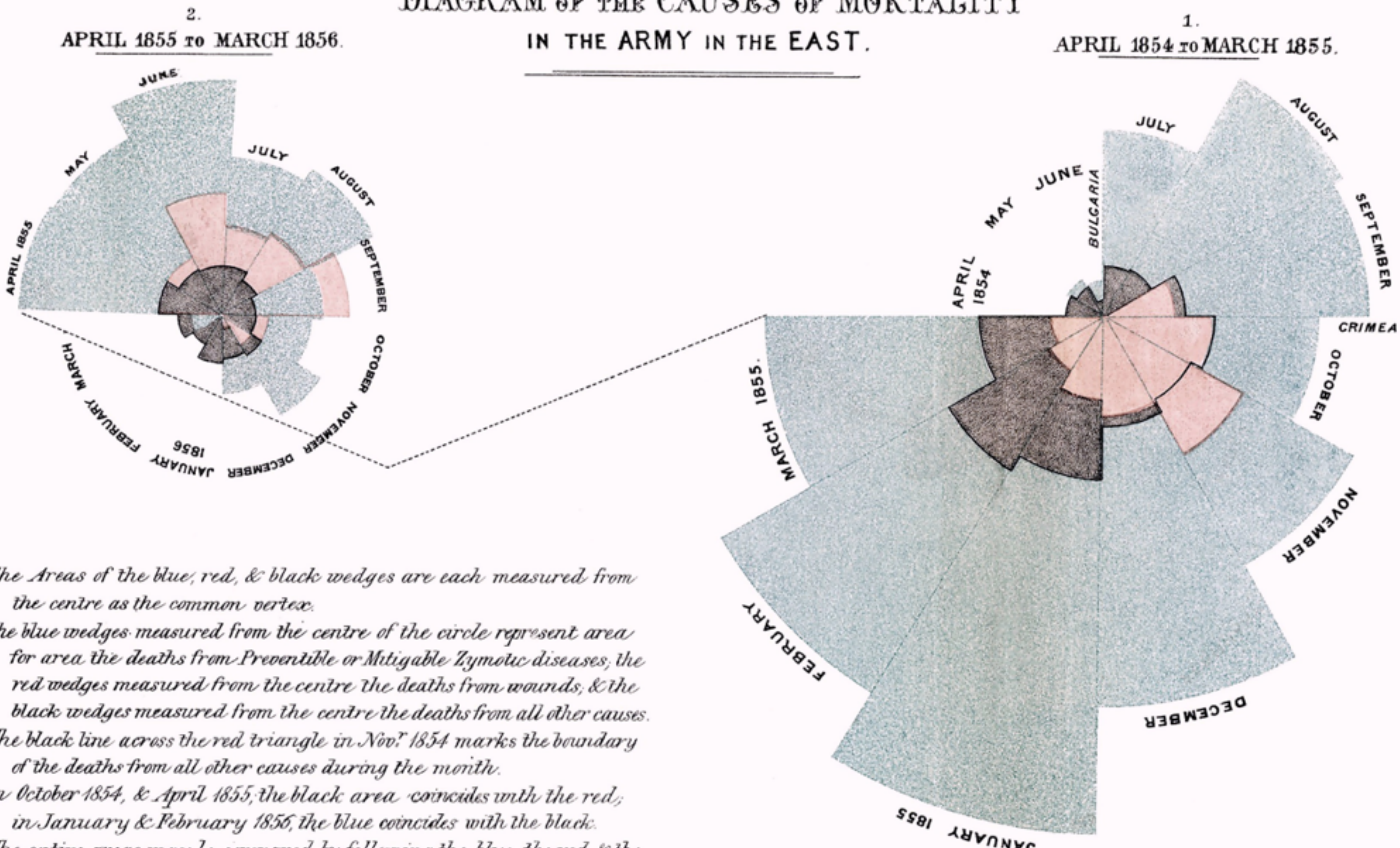


Chalfont & Latimer
Chorleywood
Rickmansworth
Watford
Croxley
Moor Park
West Ruislip
Northwood
Northwood Hills
Ruislip
Ickenham
Ruislip Manor
Eastcote
Pinner
North Harrow
Harrow-on-the-Hill
Ruislip Gardens
Rayners Lane
West Harrow
South Ruislip
Northolt
South Harrow
Sudbury Hill
Greenford
Sudbury Town
Alperton
Perivale
Hanger Lane
Park Royal
North Ealing
Ealing Broadway
Ealing Common
West Acton
North Acton
East Acton
White City
Acton Central
South Acton
Acton Town
South Ealing
Northfields
Boston Manor
Osterley
Hounslow East
Hounslow West
Hounslow Central
Hatton Cross
Terminal 4
Heathrow Airport
Chiswick Park
Turnham Green
Stamford Brook
Ravenscourt Park
Gunnersbury
Kew Gardens
Richmond
Watford Junction
Watford Junction is outside Transport for London fare zones. Special fares apply.
Watford High Street
Bushey
Carpenders Park
Hatch End
Headstone Lane
Harrow & Wealdstone
Kenton
Preston Road
Northwick Park
South Kenton
North Wembley
Wembley Central
Stonebridge Park
Harlesden
Willesden Junction
Wembley Park
Stanmore
Canons Park
Queensbury
Kingsbury
Neasden
Dollis Hill
Willesden Green
Kilburn
Brondesbury Park
Brondesbury
Kensal Rise
Kensal Green
Queen's Park
Kilburn Park
Warwick Avenue
Maida Vale
Royal Oak
Westbourne Park
Ladbroke Grove
Latimer Road
Shepherd's Bush
Wood Lane
Shepherd's Bush Market
Goldhawk Road
Hammersmith
Kensington (Olympia)
Barons Court
West Kensington
Earl's Court
West Brompton
Fulham Broadway
Parsons Green
Putney Bridge
East Putney
Southfields
Wimbledon Park
Wimbledon
Imperial Wharf
Clapham Junction
Edgware
Burnt Oak
Colindale
Hendon Central
Brent Cross
Golders Green
Hampstead
Finchley Road & Frognal
West Hampstead
Finchley Road
Swiss Cottage
St. John's Wood
Kilburn High Road
South Hampstead
Belsize Park
Chalk Farm
Camden Town
Mornington Crescent
Hampstead Heath
Gospel Oak
Kentish Town West
Kentish Town
Mill Hill East
High Barnet
Totteridge & Whetstone
Woodside Park
West Finchley
Finchley Central
East Finchley
Highgate
Archway
Tufnell Park
Crouch Hill
Upper Holloway
Camden Road
Caledonian Road & Barnsbury
Caledonian Road
Holloway Road
Arsenal
Highbury & Islington
Canonbury
Cockfosters
Oakwood
Southgate
Arnos Grove
Bounds Green
Wood Green
Turnpike Lane
Manor House
Harringay Green Lanes
Seven Sisters
South Tottenham
Blackhorse Road
Tottenham Hale
Walthamstow Central
Walthamstow Queen's Road
Finsbury Park
Dalston Kingsland
Hackney Central
Homerton
Hackney Wick
Paddington
Edgware Road
Baker Street
Great Portland Street
Euston
King's Cross St. Pancras
Angel
Old Street
Marylebone
Warren Street
Regent's Park
Euston Square
Farringdon
Barbican
Russell Square
Goodge Street
Chancery Lane
Moorgate
Liverpool Street
Bethnal Green
Shoreditch
Holland Park
Notting Hill Gate
Lancaster Gate
Bond Street
Oxford Circus
Tottenham Court Road
Holborn
St. Paul's
Bank
Queensway
Marble Arch
Covent Garden
Bayswater
High Street Kensington
Hyde Park Corner
Green Park
Piccadilly Circus
Leicester Square
Mansion House
Cannon Street
Monument
Tower Hill
Aldgate
Aldgate East
Tower Gateway
Knightsbridge
Gloucester Road
Sloane Square
South Kensington
St. James's Park
Victoria
Westminster
Charing Cross
Embankment
Temple
Blackfriars
Waterloo
Southwark
Borough
London Bridge
Lambeth North
Elephant & Castle
Pimlico
Vauxhall
Kennington
Oval
Stockwell
Brixton
Clapham North
Clapham Common
Clapham South
Balham
Tooting Bec
Tooting Broadway
Colliers Wood
South Wimbledon
Morden
Mile End
Bow Road
Bow Church
Devons Road
Langdon Park
All Saints
Pudding Mill Lane
Stepney Green
Whitechapel
Shadwell
Wapping
Limehouse
Westferry
Poplar
Blackwall
East India
West India Quay
Canary Wharf
Heron Quays
South Quay
Crossharbour
Mudchute
Island Gardens
Cutty Sark for Maritime Greenwich
Greenwich
Deptford Bridge
Elverson Road
Lewisham
Bermondsey
Rotherhithe
Canada Water
Surrey Quays
New Cross Gate
New Cross
East London line is closed for major line extension work to become part of the London Overground network
North Greenwich for The O2
Canning Town
West Silvertown
Pontoon Dock
London City Airport
King George V
Woolwich Arsenal
Royal Victoria
Custom House for ExCeL
Prince Regent
Royal Albert
Beckton Park
Cyprus
Stratford
West Ham
Plaistow
Upton Park
East Ham
Bromley-by-Bow
Barking
Upney
Becontree
Dagenham Heathway
Dagenham East
Elm Park
Hornchurch
Upminster Bridge
Upminster
Leytonstone
Leyton
Leyton Midland Road
Leytonstone High Road
Wanstead Park
Woodgrange Park
Wanstead
Redbridge
Gants Hill
Newbury Park
Barkingside
Fairlop
Hainault
Grange Hill
Chigwell
Roding Valley
Woodford
South Woodford
Snaresbrook
Buckhurst Hill
Loughton
Debden
Theydon Bois
Epping
River Thames
Improvement works may affect your journey, particularly at weekends. Check before you travel; look for publicity at stations, visit tfl.gov.uk/check or call 020 7222 1234


selfiecity.net
SELFIECITY
Introduction
Imageplots
Selfiexploratory
Dataset
Findings
Theory
Credits & contact
Explore further
BANGKOK
BERLIN
MOSCOW
NEW YORK
SAO PAULO
SELFIECITY
Investigating the style of **self-portraits** (*selfies*) in five cities across the world.
Selfiecity investigates *selfies* using a mix of theoretic, artistic and quantitative methods:
We present our findings about the demographics of people taking selfies, their poses and expressions.
Rich media visualizations (imageplots) assemble thousands of photos to reveal interesting patterns.
The interactive selfiexploratory allows you to navigate the whole set of 3200 photos.
Finally, theoretical essays discuss selfies in the history of photography, the functions of images in social media, and methods and dataset.
IMAGEPLOTS
POSES
Each city has a
different style when
it comes to selfies.

# COMMUNICATE IDEAS

# DIAGRAM OF THE CAUSES OF MORTALITY IN THE ARMY IN THE EAST.




*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.*

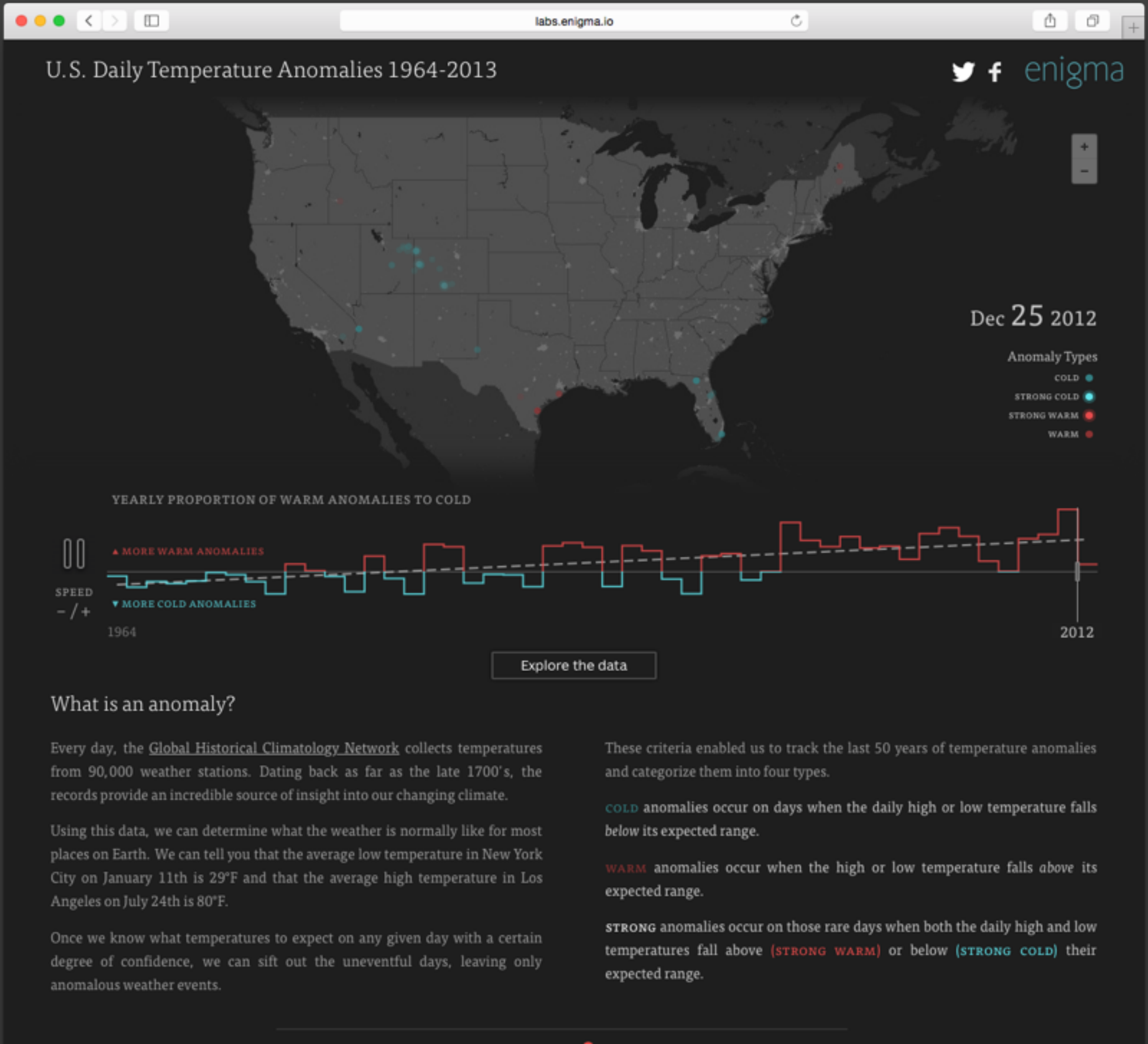
*The black line across the red triangle in Nov.r 1854 marks the boundary of the deaths from all other causes during the month.*

*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.*

F. Nightingale 1856


nationalgeographic.com
What the World Eats
Daily Diet
Meat Consumption
16543
World
Daily diets vary considerably around the world. Select different countries or regions to see how consumption patterns have changed in the last 50 years. Click on grams to understand the quantities of food consumed per person in each place. Alternatively, view the breakdown by calories to measure how that balance of food translates into fuel and energy. Percentages in the pie chart reflect the proportion of each item in the overall diet.
Calories per Person
Grams per Person
Dairy & Eggs
Meat
Other
Produce
Sugar & Fat
Grain
World
Daily Calories
2,870
2011
World
China
United States
India
Brazil
Hong Kong
Argentina
Russia*
Cuba
North Korea
South Korea
Kuwait
Libya


labs.enigma.io
U.S. Daily Temperature Anomalies 1964-2013
enigma
Dec 25 2012
Anomaly Types
COLD
STRONG COLD
STRONG WARM
WARM
YEARLY PROPORTION OF WARM ANOMALIES TO COLD
▲ MORE WARM ANOMALIES
▼ MORE COLD ANOMALIES
SPEED
– / +
1964
2012
Explore the data
What is an anomaly?
Every day, the Global Historical Climatology Network collects temperatures from 90,000 weather stations. Dating back as far as the late 1700's, the records provide an incredible source of insight into our changing climate.
Using this data, we can determine what the weather is normally like for most places on Earth. We can tell you that the average low temperature in New York City on January 11th is 29°F and that the average high temperature in Los Angeles on July 24th is 80°F.
Once we know what temperatures to expect on any given day with a certain degree of confidence, we can sift out the uneventful days, leaving only anomalous weather events.
These criteria enabled us to track the last 50 years of temperature anomalies and categorize them into four types.
COLD anomalies occur on days when the daily high or low temperature falls below its expected range.
WARM anomalies occur when the high or low temperature falls above its expected range.
STRONG anomalies occur on those rare days when both the daily high and low temperatures fall above (STRONG WARM) or below (STRONG COLD) their expected range.

bost.ocks.org

June 26, 2014 / Mike Bostock

# Visualizing Algorithms

> *The power of the unaided mind is highly overrated... The real powers come from devising external aids that enhance cognitive abilities.* —Donald Norman

Algorithms are a fascinating use case for visualization. To visualize an algorithm, we don't merely fit data to a chart; there is no primary dataset. Instead there are logical rules that describe behavior. This may be why algorithm visualizations are so unusual, as designers experiment with novel forms to better communicate. This is reason enough to study them.

But algorithms are also a reminder that visualization is more than a tool for finding patterns in data. Visualization leverages the human visual system to augment human intellect: we can use it to better understand these important abstract processes, and perhaps other things, too.

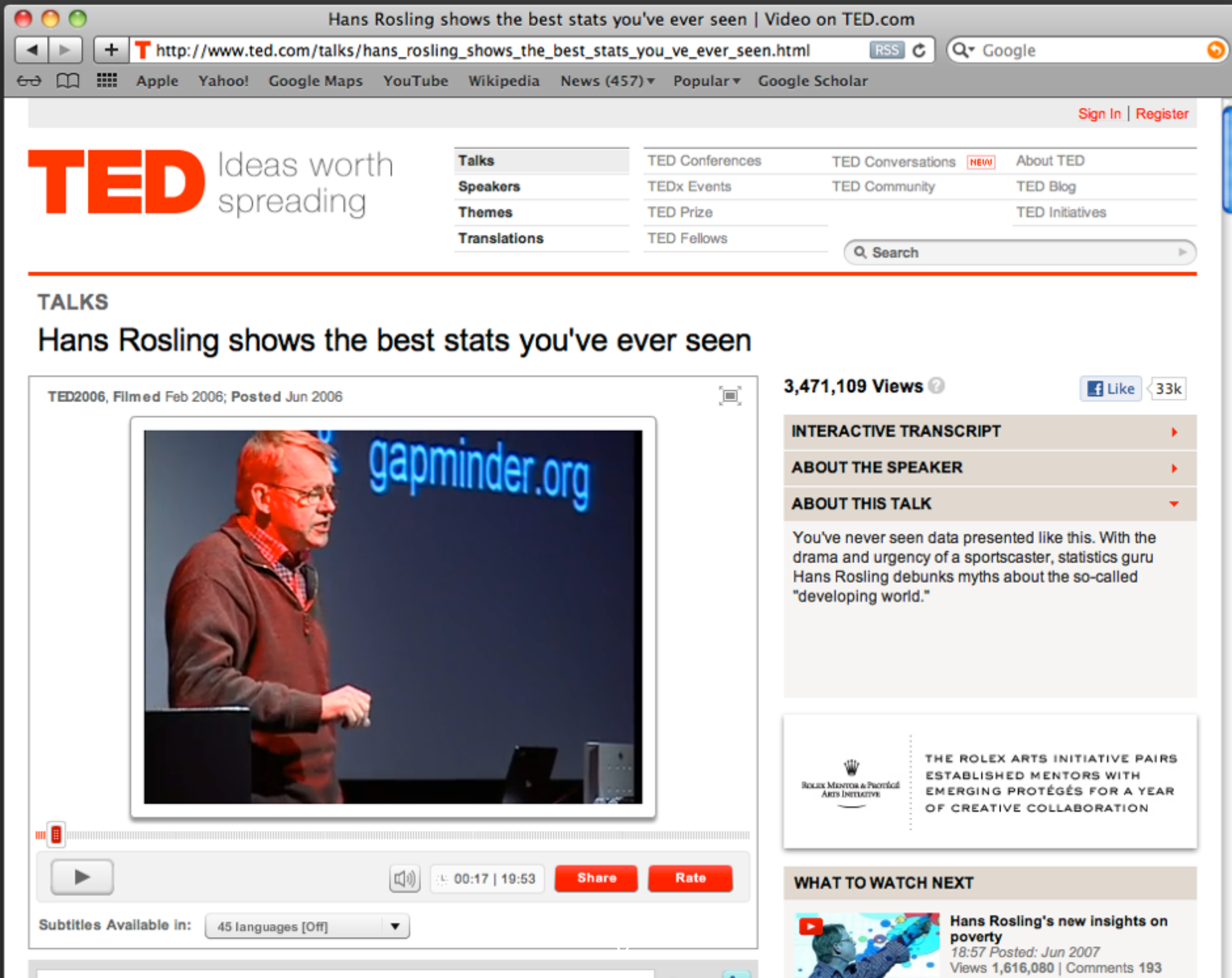
This is an adaption of my talk at Eyeo 2014. A video of the talk is available on Vimeo. (Thanks, Eyeo folks!)

## Sampling

Before I can explain the first algorithm, I first need to explain the problem it addresses.

Van Gogh's *The Starry Night*

Light — electromagnetic radiation — the light emanating from this screen, traveling through the air, focused by your lens and projected onto the retina — is a continuous signal. To be perceived, we must reduce light to discrete impulses by measuring its intensity and frequency distribution at different points in space.

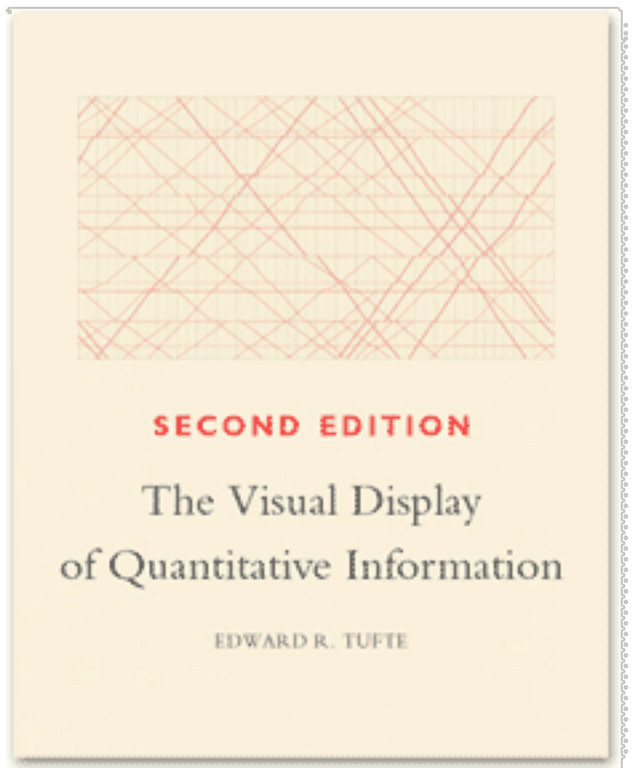

Hans Rosling shows the best stats you've ever seen | Video on TED.com
http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html
RSS
Google
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News (457)
Popular
Google Scholar
Sign In | Register
TED
Ideas worth spreading
Talks
Speakers
Themes
Translations
TED Conferences
TEDx Events
TED Prize
TED Fellows
TED Conversations NEW
TED Community
About TED
TED Blog
TED Initiatives
Search
TALKS
Hans Rosling shows the best stats you've ever seen
TED2006, Filmed Feb 2006; Posted Jun 2006
gapminder.org
00:17 | 19:53
Share
Rate
Subtitles Available in:
45 languages [Off]
3,471,109 Views
Like
33k
INTERACTIVE TRANSCRIPT
ABOUT THE SPEAKER
ABOUT THIS TALK
You've never seen data presented like this. With the drama and urgency of a sportscaster, statistics guru Hans Rosling debunks myths about the so-called "developing world."
THE ROLEX ARTS INITIATIVE PAIRS ESTABLISHED MENTORS WITH EMERGING PROTÉGÉS FOR A YEAR OF CREATIVE COLLABORATION
WHAT TO WATCH NEXT
Hans Rosling's new insights on poverty
18:57 Posted: Jun 2007
Views 1,616,080 | Comments 193

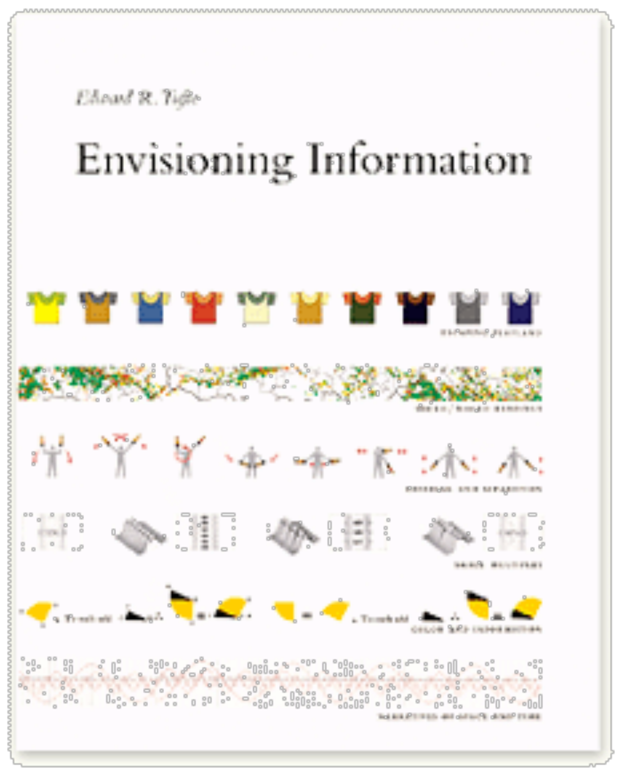
# DESIGN PRINCIPLES

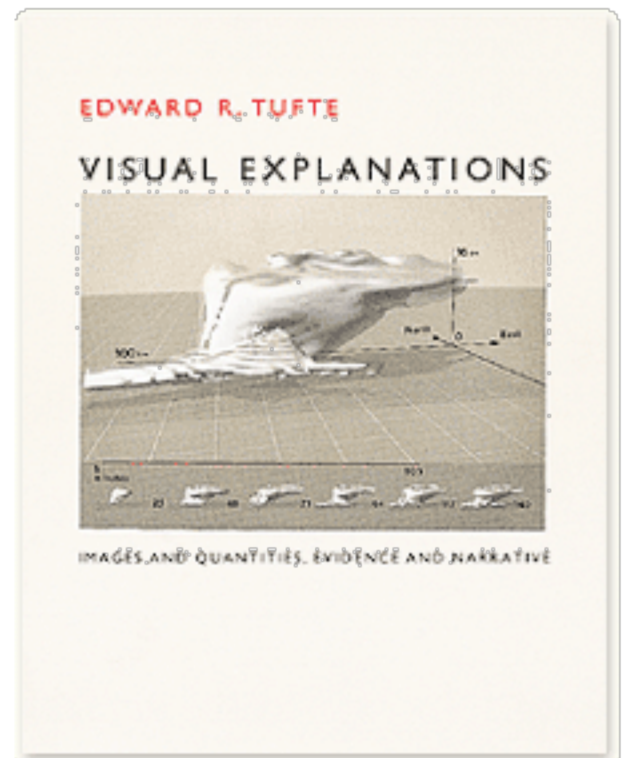
# design excellence

"Well-designed presentations of interesting data are a matter of substance, of statistics, and of design."

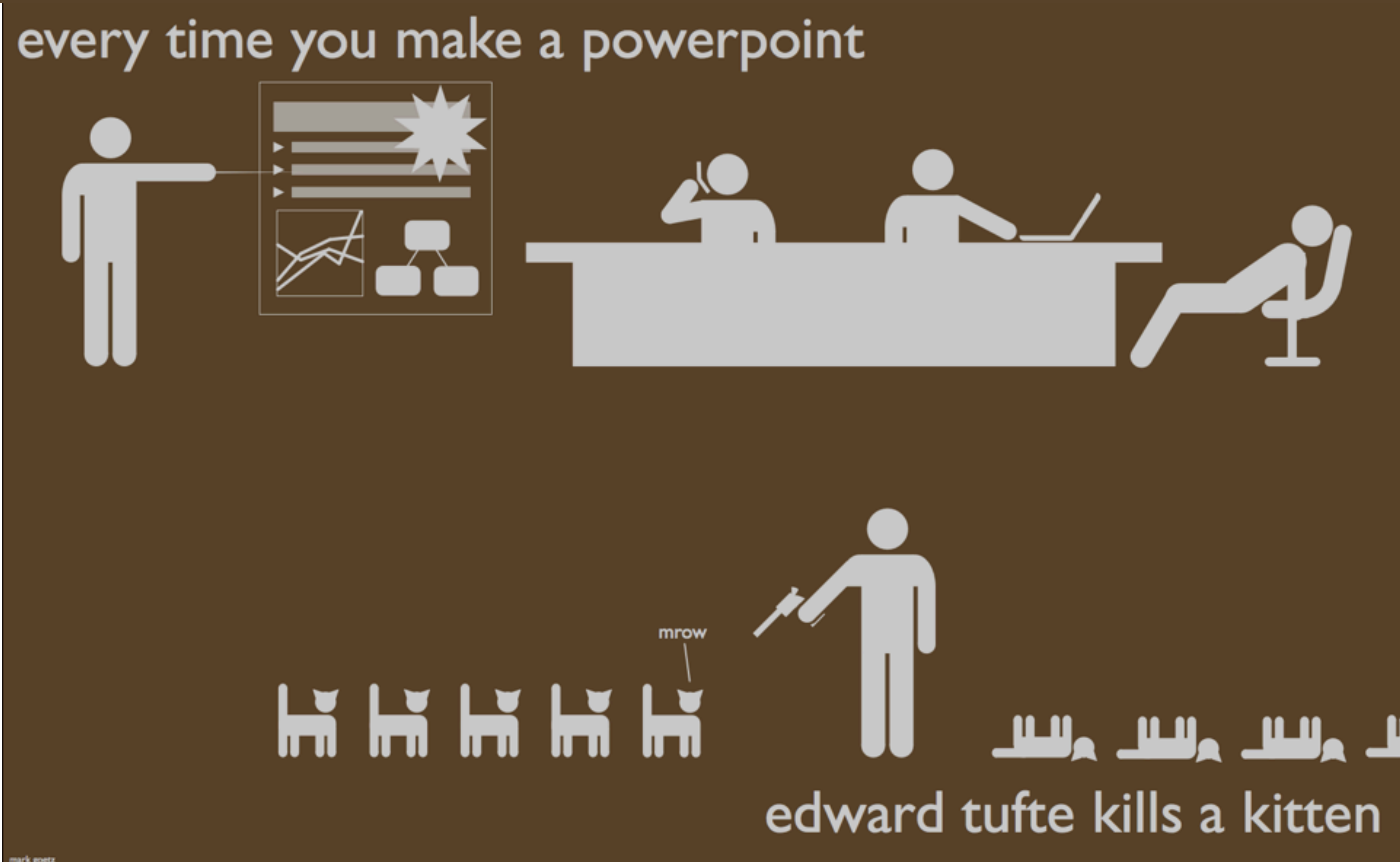
Edward Tufte


SECOND EDITION
The Visual Display
of Quantitative Information
EDWARD R. TUFTE


Edward R. Tufte
Envisioning Information

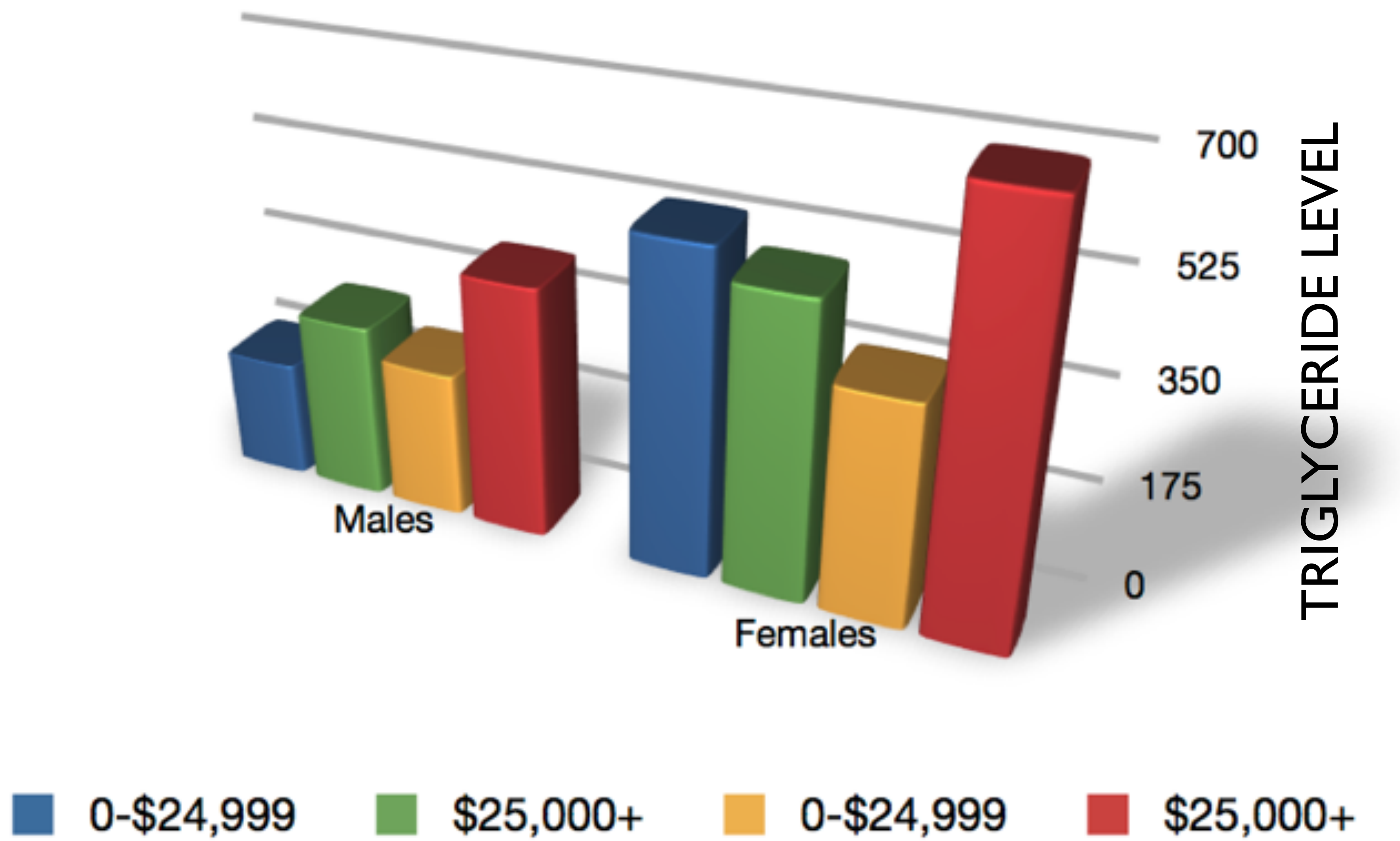

EDWARD R. TUFTE
VISUAL EXPLANATIONS
IMAGES AND QUANTITIES, EVIDENCE AND NARRATIVE


EDWARD TUFTE
BEAUTIFUL EVIDENCE


every time you make a powerpoint
mrow
edward tufte kills a kitten
mark goetz

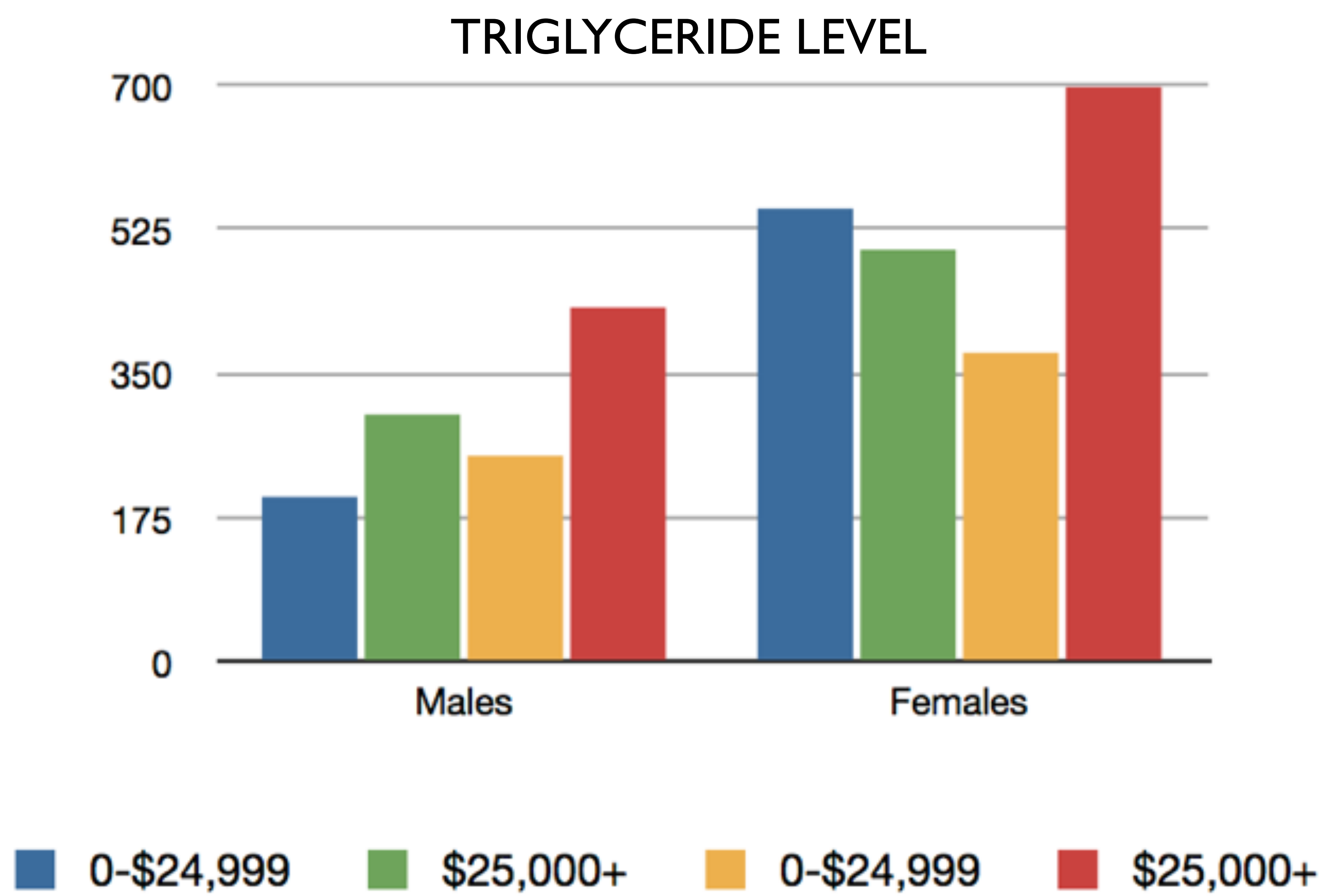
maximize the

**Data-ink Ratio** = $\dfrac{\text{data-ink}}{\text{total ink used in graphic}}$



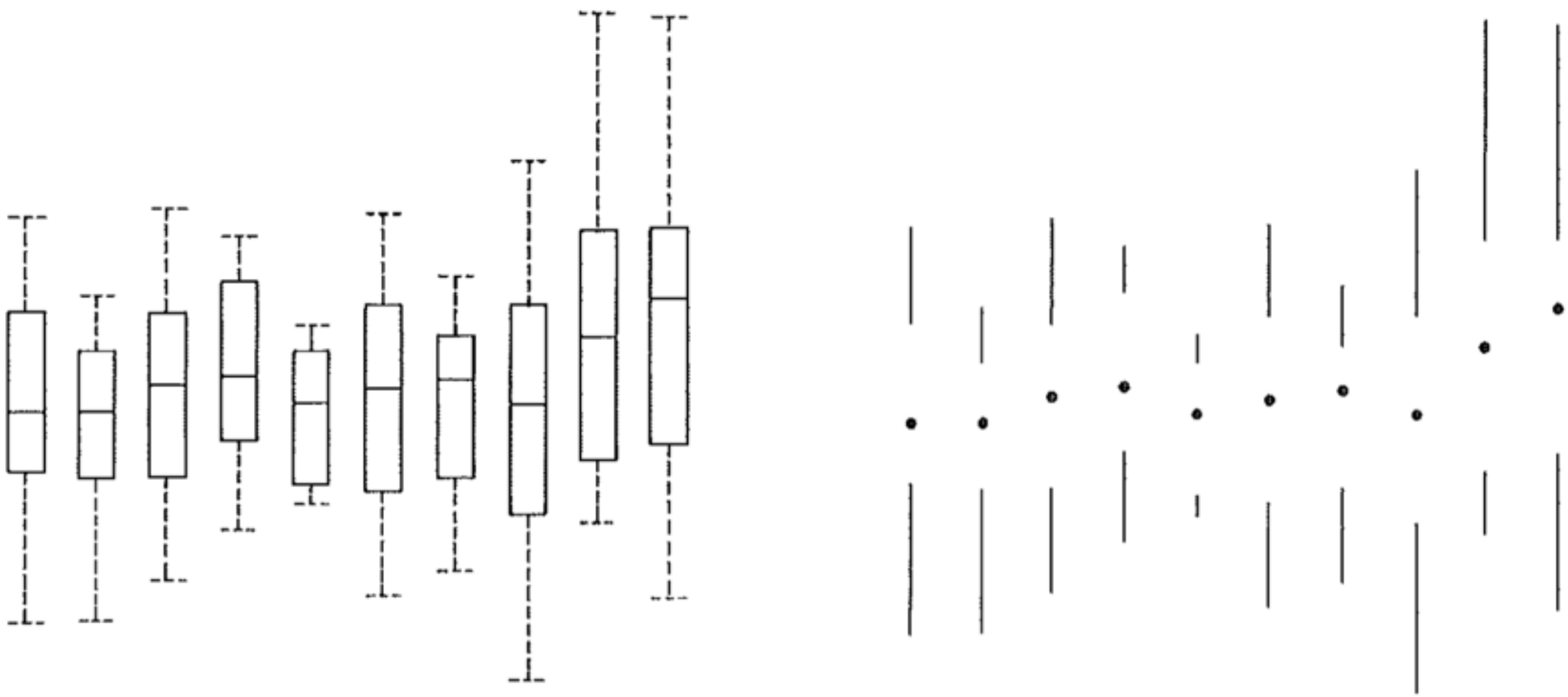
maximize the

$$\textbf{Data-ink Ratio} = \frac{\text{data-ink}}{\text{total ink used in graphic}}$$



maximize the

$$\textbf{Data-ink Ratio} = \frac{\text{data-ink}}{\text{total ink used in graphic}}$$



Tufte 2001

Eurographics / IEEE Symposium on Visualization 2011 (EuroVis 2011)
H. Hauser, H. Pfister, and J. J. van Wijk
(Guest Editors)

*Volume 30 (2011), Number 3*

# A User Study of Visualization Effectiveness Using EEG and Cognitive Load

E. W. Anderson[1], K. C. Potter[1], L. E. Matzen[2], J. F. Shepherd[2], G. A. Preston[3], and C. T. Silva[1]

[1]SCI Institute, University of Utah, USA
[2]Sandia National Laboratories, USA
[3]Utah State Hospital, USA

**Abstract**

*Effectively evaluating visualization techniques is a difficult task often assessed through feedback from user studies and expert evaluations. This work presents an alternative approach to visualization evaluation in which brain activity is passively recorded using electroencephalography (EEG). These measurements are used to compare different visualization [illegible] sources. In this paper, EEG signals and resp[illegible] of data distributions. This information is processed to provide insight into the cognitive load imposed on the viewer. This paper describes the design of the user study performed, the extraction of cognitive load measures from EEG data, and how those measures are used to quantitatively evaluate the effectiveness of visualizations.*

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: General—Human Factors, Evaluation, Electroencephalography

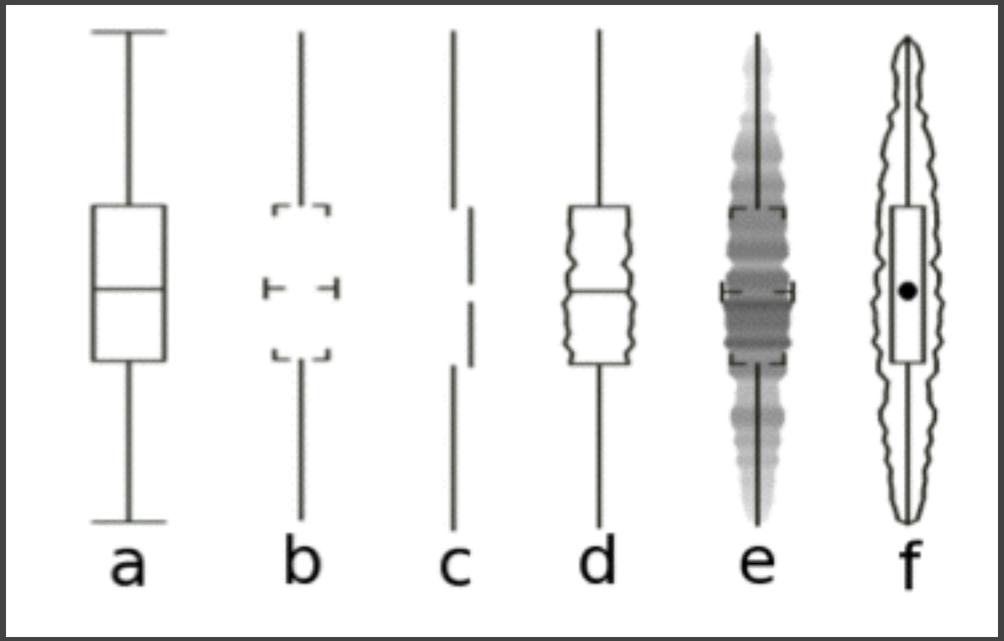
# COUNTER-POINT

## 1. Introduction

Efficient visualizations facilitate the understanding of data

this paper strives to evaluate visualization techniques objectively by using passive, non-invasive monitoring devices to measure the burden placed on a user's cognitive resources.

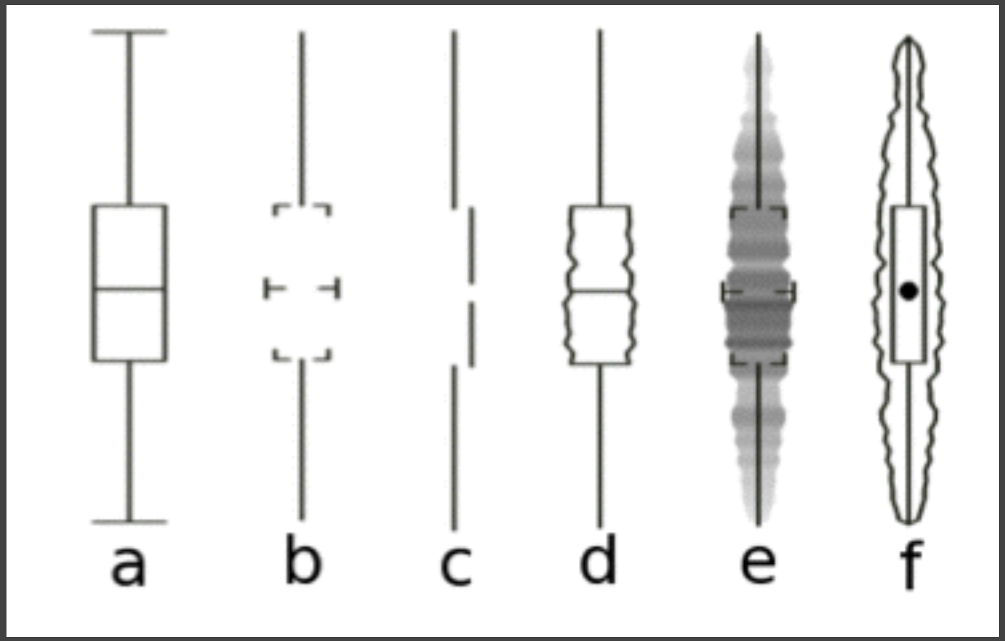
# EXPERIMENT



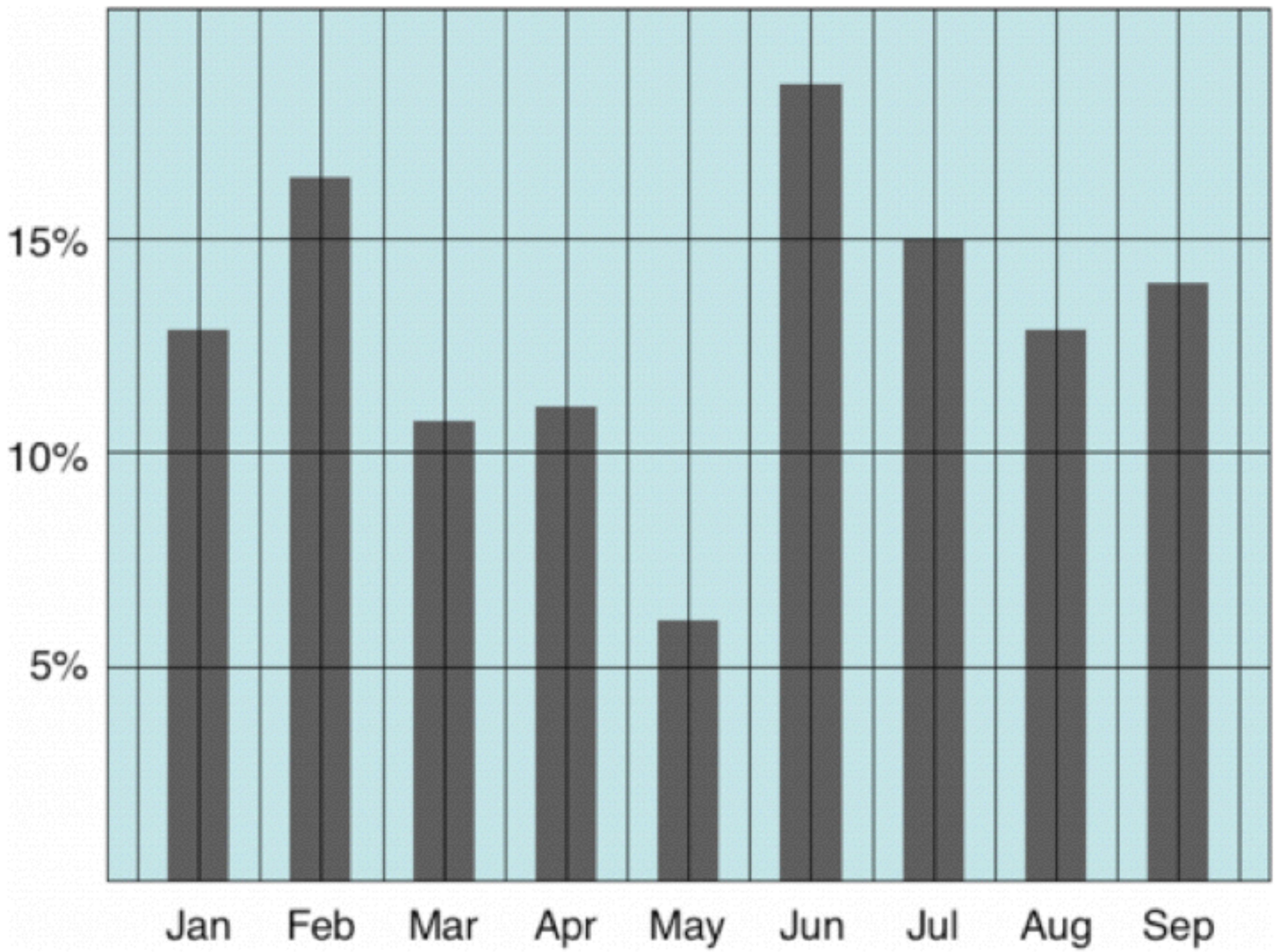

- asked participants to choose box plot with largest range from a set
- varied representation
- measured cognitive load from EEG brain waves

# EXPERIMENTAL RESULTS

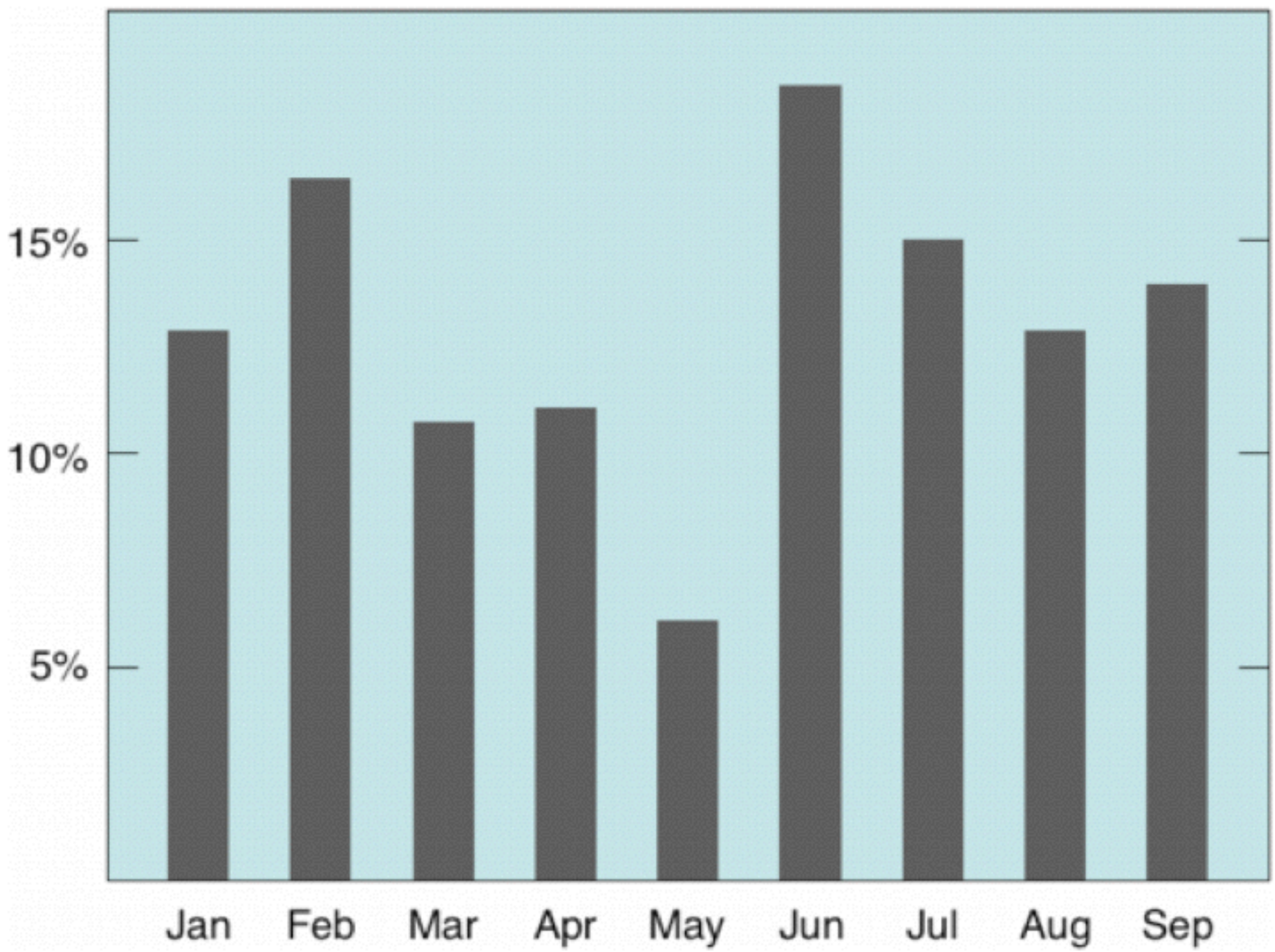
- paper focused on cognitive load as an evaluation method
- studies showed that the simplest box plot is hardest to interpret



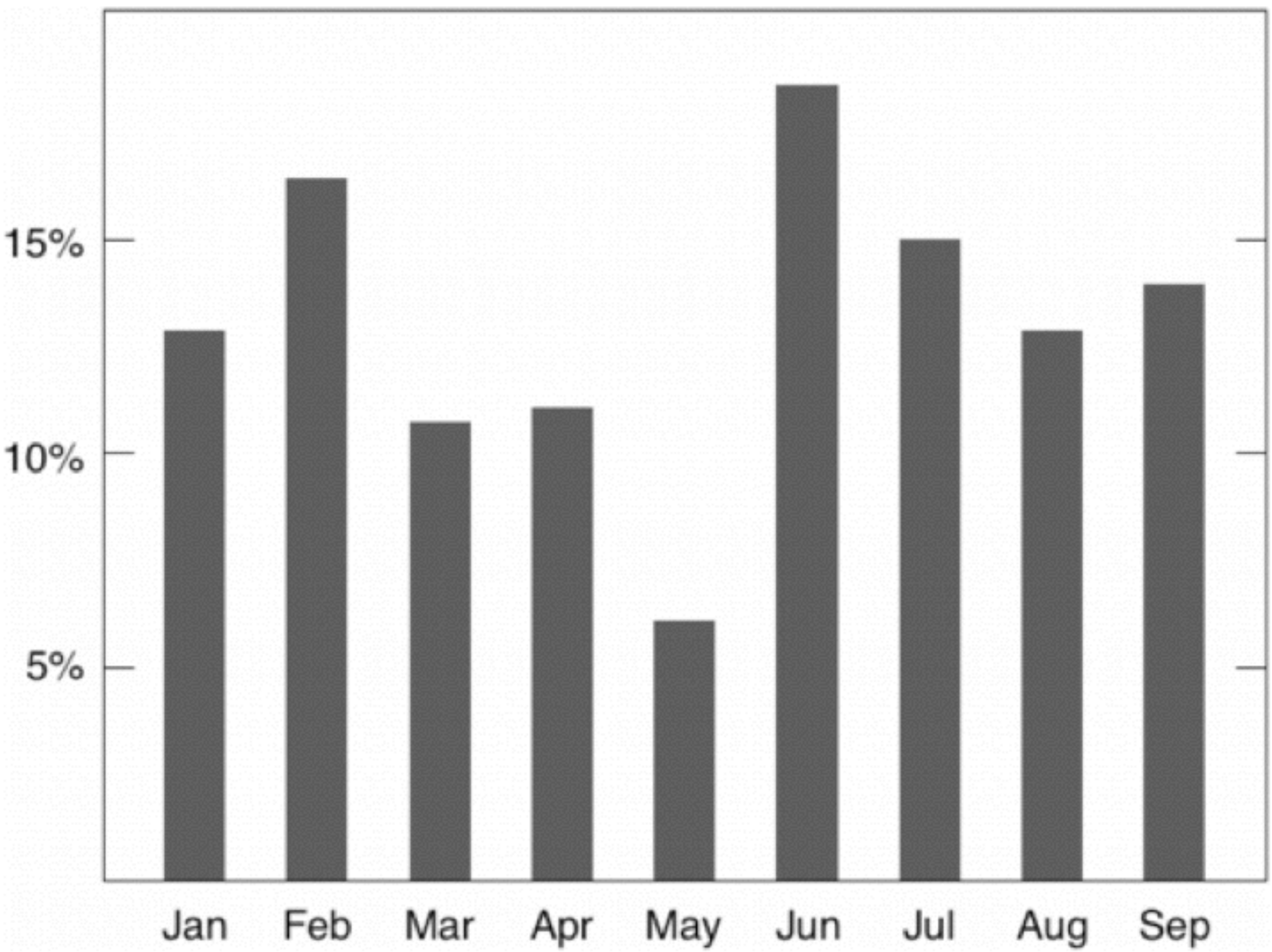
# AVOID CHART JUNK



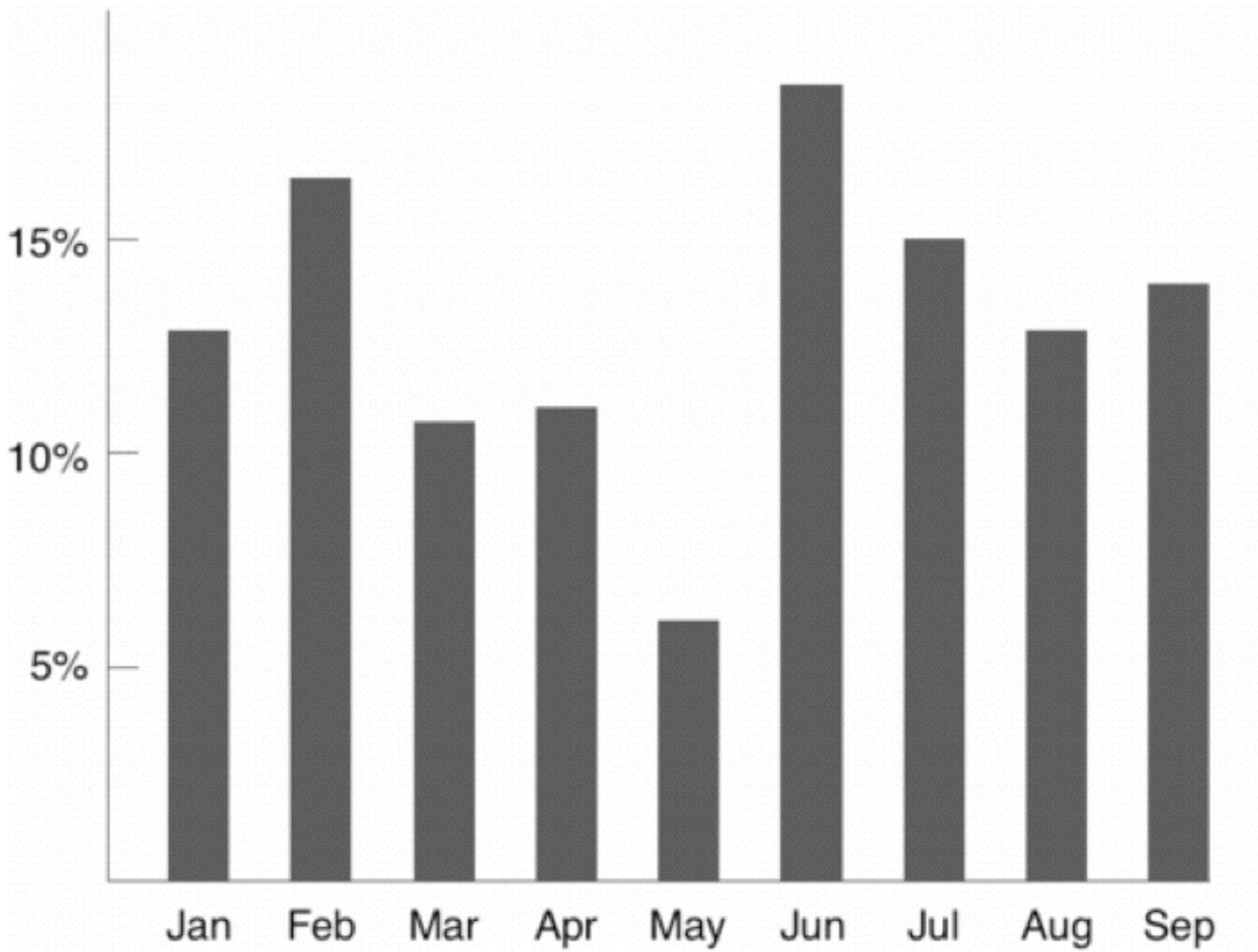
# AVOID CHART JUNK



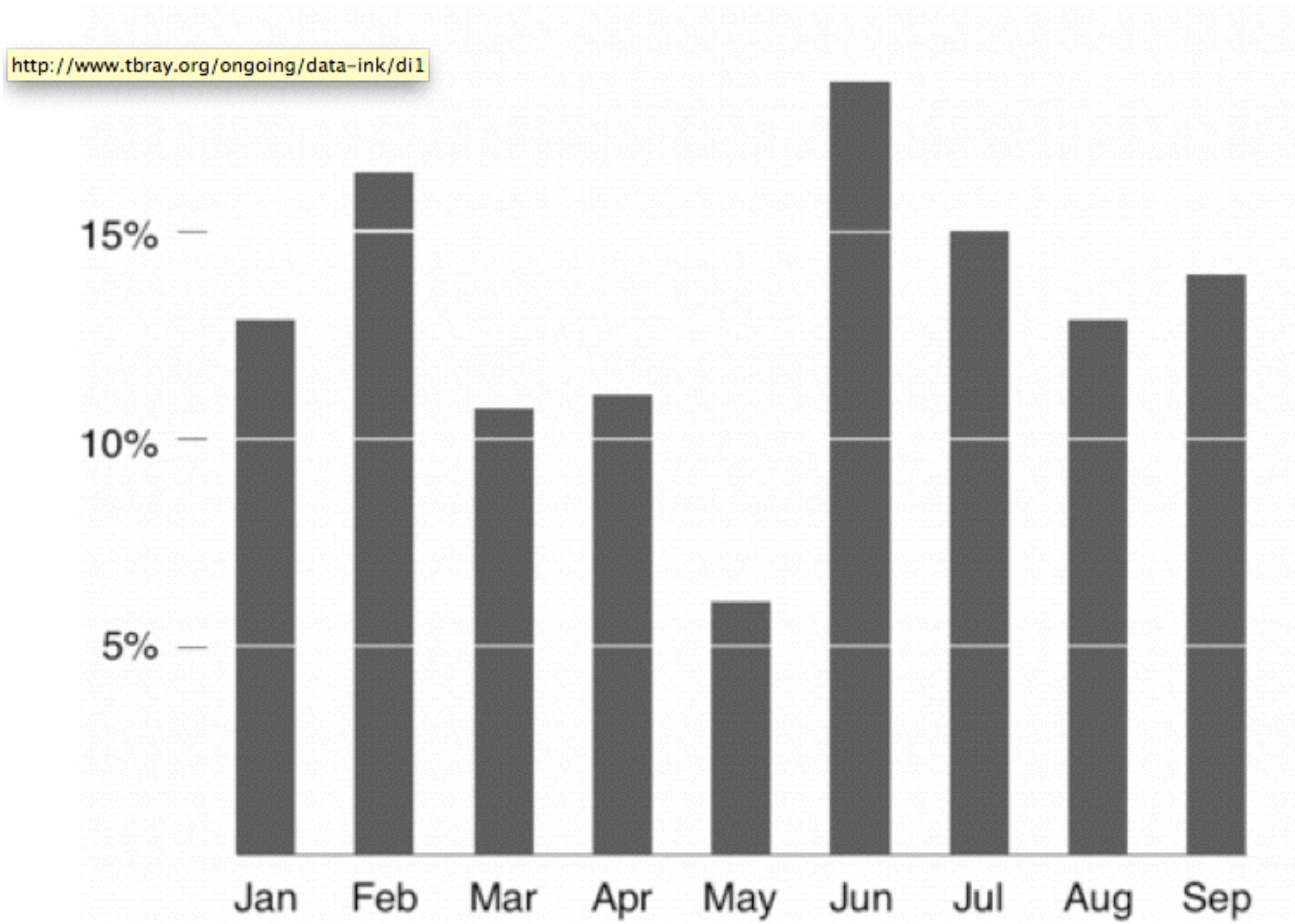
# AVOID CHART JUNK



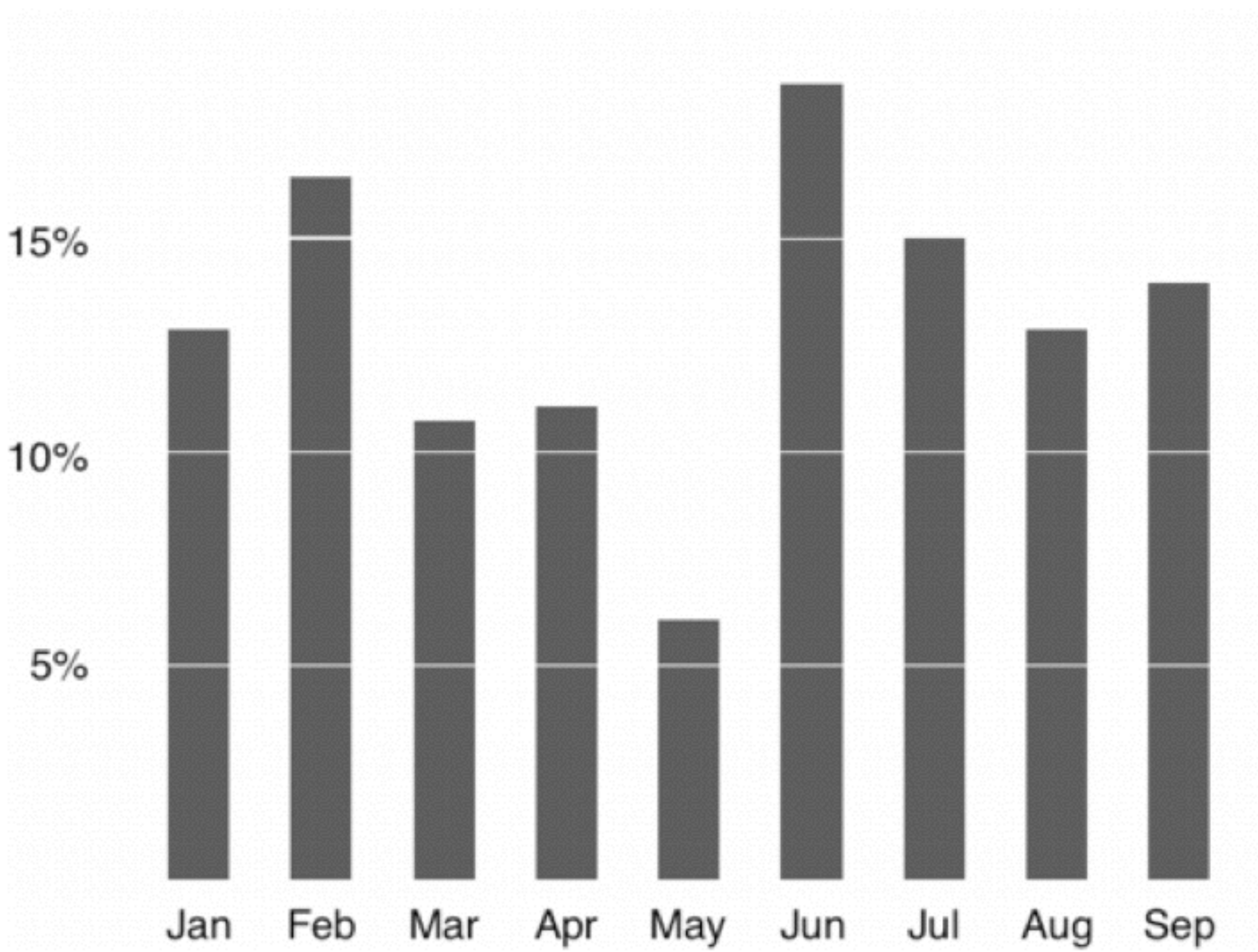
# AVOID CHART JUNK



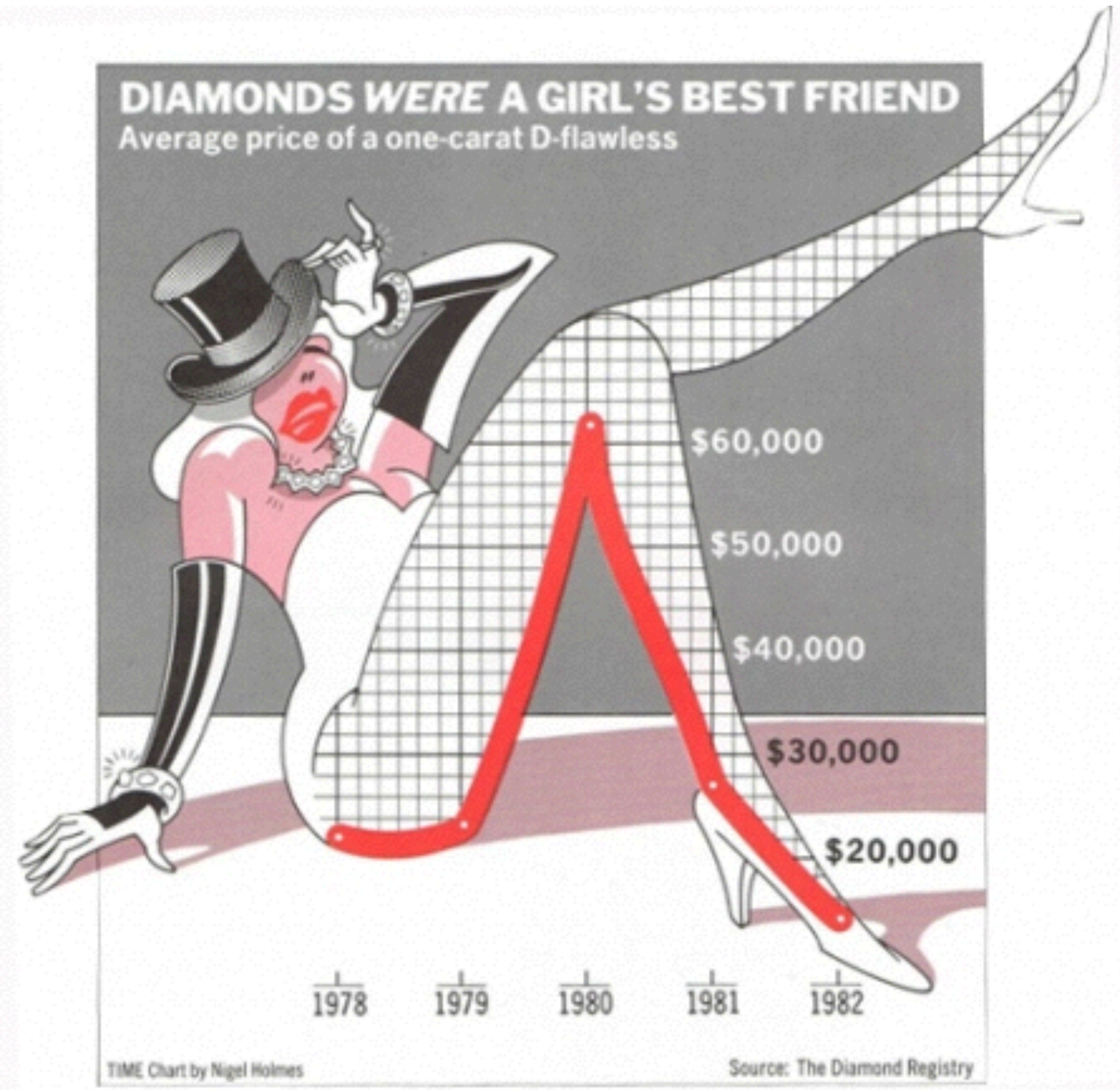
# AVOID CHART JUNK

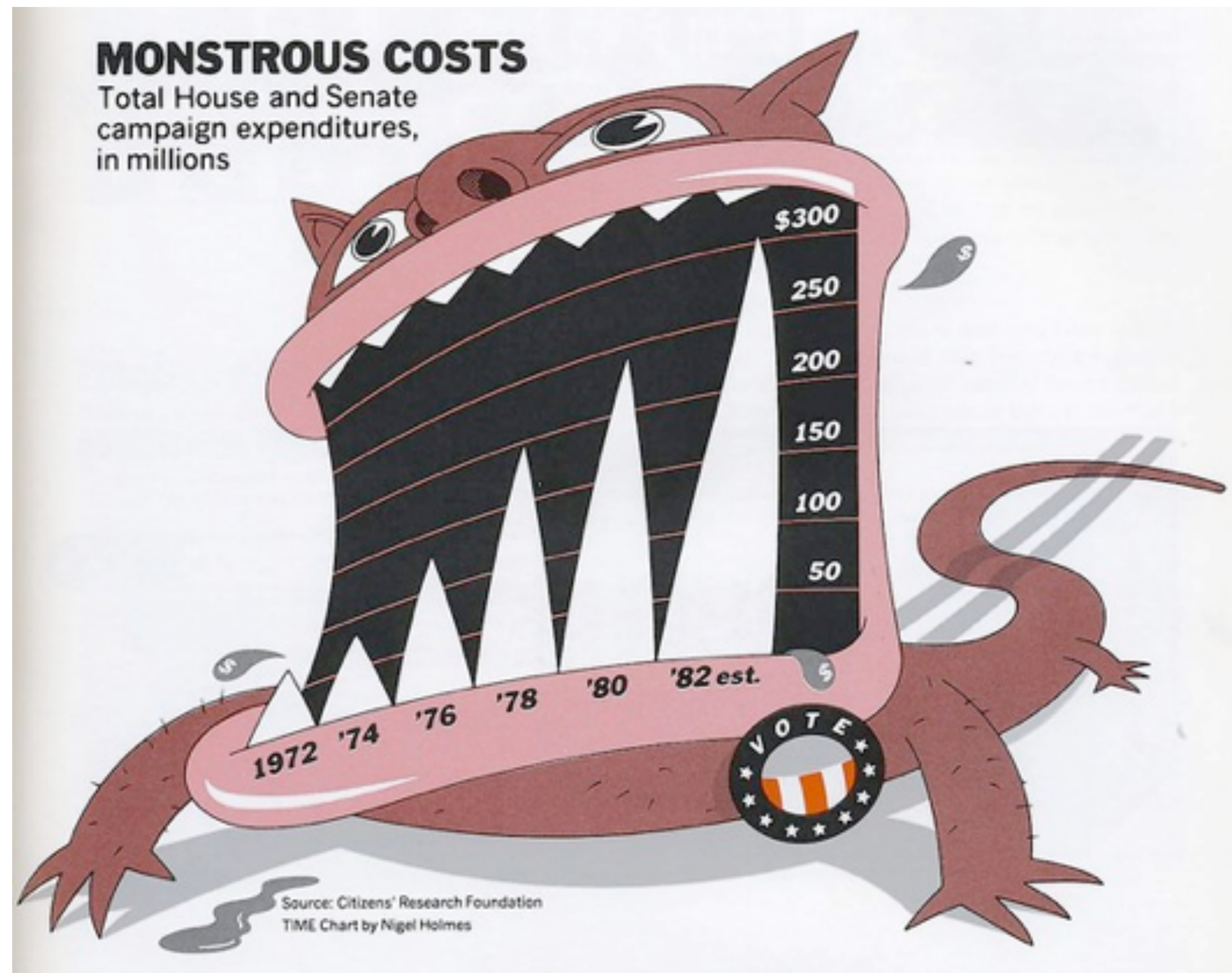


# AVOID CHART JUNK



# redesign exercise . . .




Nigel Holmes, TIME Magazine

# COUNTER-POINT

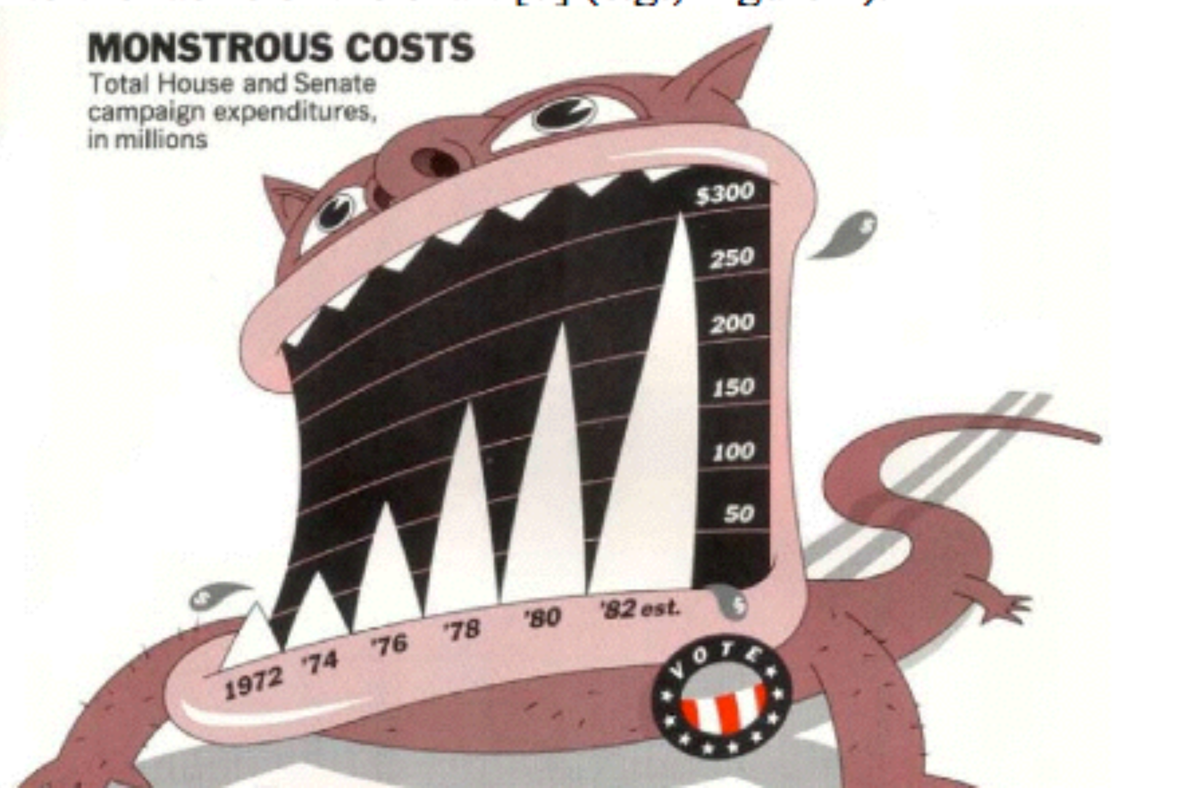
# Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

**Scott Bateman, Regan L. Mandryk, Carl Gutwin,**
**Aaron Genest, David McDine, Christopher Brooks**
Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca,
aaron.genest@usask.ca, dam085@mail.usask.ca, cab938@mail.usask.ca

**ABSTRACT**
Guidelines for designing information charts often state that the presentation should reduce 'chart junk' – visual embellishments that are not essential to understanding the data. In contrast, some popular chart designers wrap the presented data in detailed and elaborate imagery, raising the questions of whether this imagery is really as detrimental to understanding as has been proposed, and whether the visual embellishment may have other benefits. To investigate these issues, we conducted an experiment that compared embellished charts with plain ones, and measured both interpretation accuracy and long-term recall. We found that people's accuracy in describing the embellished charts was no worse than for plain charts, and that their recall after a two-to-three-week gap was significantly better. Although we are cautious about recommending that all charts be produced in this style, our results question some of the premises of the minimalist approach to chart design.

**Author Keywords**
Charts, information visualization, imagery, memorability.

Despite these minimalist guidelines, many designers include a wide variety of visual embellishments in their charts, from small decorations to large images and visual backgrounds. One well-known proponent of visual embellishment in charts is the graphic artist Nigel Holmes, whose work regularly incorporates strong visual imagery into the fabric of the chart [7] (e.g., Figure 1).



# EXPERIMENTAL QUESTIONS

1) whether visual embellishments do in fact cause comprehension problems

2) whether the embellishments may provide additional information that is valuable for the reader

# EXPERIMENTAL RESULTS

1) **No significant difference** between plain and image charts for interactive **interpretation accuracy**

2) **No significant difference** in **recall accuracy** after a five-minute gap

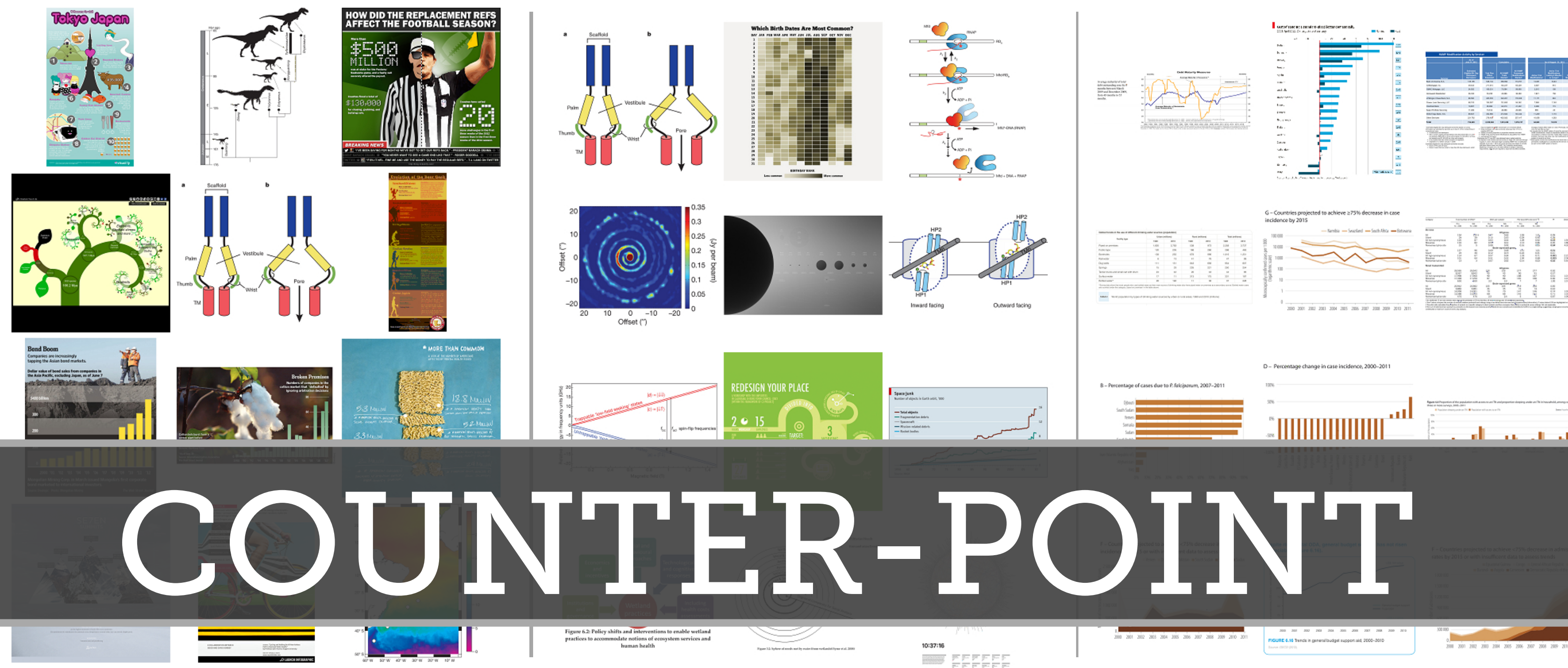
3) **Significantly better recall** for Holmes charts of both the chart topic and the details (categories and trend) after long-term gap (2-3 weeks).

4) Participants **saw value messages** in the Holmes charts **significantly more often** than in the plain charts.

5) Participants found the Holmes charts **more attractive**, **most enjoyed** them, and found that they were **easiest and fastest to remember**.

# What Makes a Visualization Memorable?

Michelle A. Borkin, *Student Member, IEEE*, Azalea A. Vo, Zoya Bylinskii, Phillip Isola, *Student Member, IEEE*, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister, *Senior Member, IEEE*



COUNTER-POINT

Fig. 1. **Left:** The top twelve overall most memorable visualizations from our experiment (most to least memorable from top left to bottom right). **Middle:** The top twelve most memorable visualizations from our experiment when visualizations containing human recognizable cartoons or images are removed (most to least memorable from top left to bottom right). **Right:** The twelve least memorable visualizations from our experiment (most to least memorable from top left to bottom right).

**Abstract**—An ongoing debate in the Visualization community concerns the role that visualization types play in data understanding. In human cognition, understanding and memorability are intertwined. As a first step towards being able to ask questions about impact and effectiveness, here we ask: "What makes a visualization memorable?" We ran the largest scale visualization study to date using 2,070 single-panel visualizations, categorized with visualization type (e.g., bar chart, line graph, etc.), collected from news media sites, government reports, scientific journals, and infographic sources. Each visualization was annotated with additional attributes, including ratings for data-ink ratios and visual densities. Using Amazon's Mechanical Turk, we collected memorability scores for hundreds of these visualizations, and discovered that observers are consistent in which visualizations they find memorable and forgettable. We find intuitive results (e.g., attributes like color and the inclusion of a human recognizable object enhance memorability) and less intuitive results (e.g., common graphs are less memorable than unique visualization types). Altogether our findings suggest that quantifying memorability is a general metric of the utility of information, an essential step towards determining how to design effective visualizations.

**Index Terms**—Visualization taxonomy, information visualization, memorability

# TAKE-AWAY

1) **intuitive findings:** color and human recognizable objects enhance memorability

2) **unintuitive findings:** common graphs are less memorable than unique visualization types

take away . . .

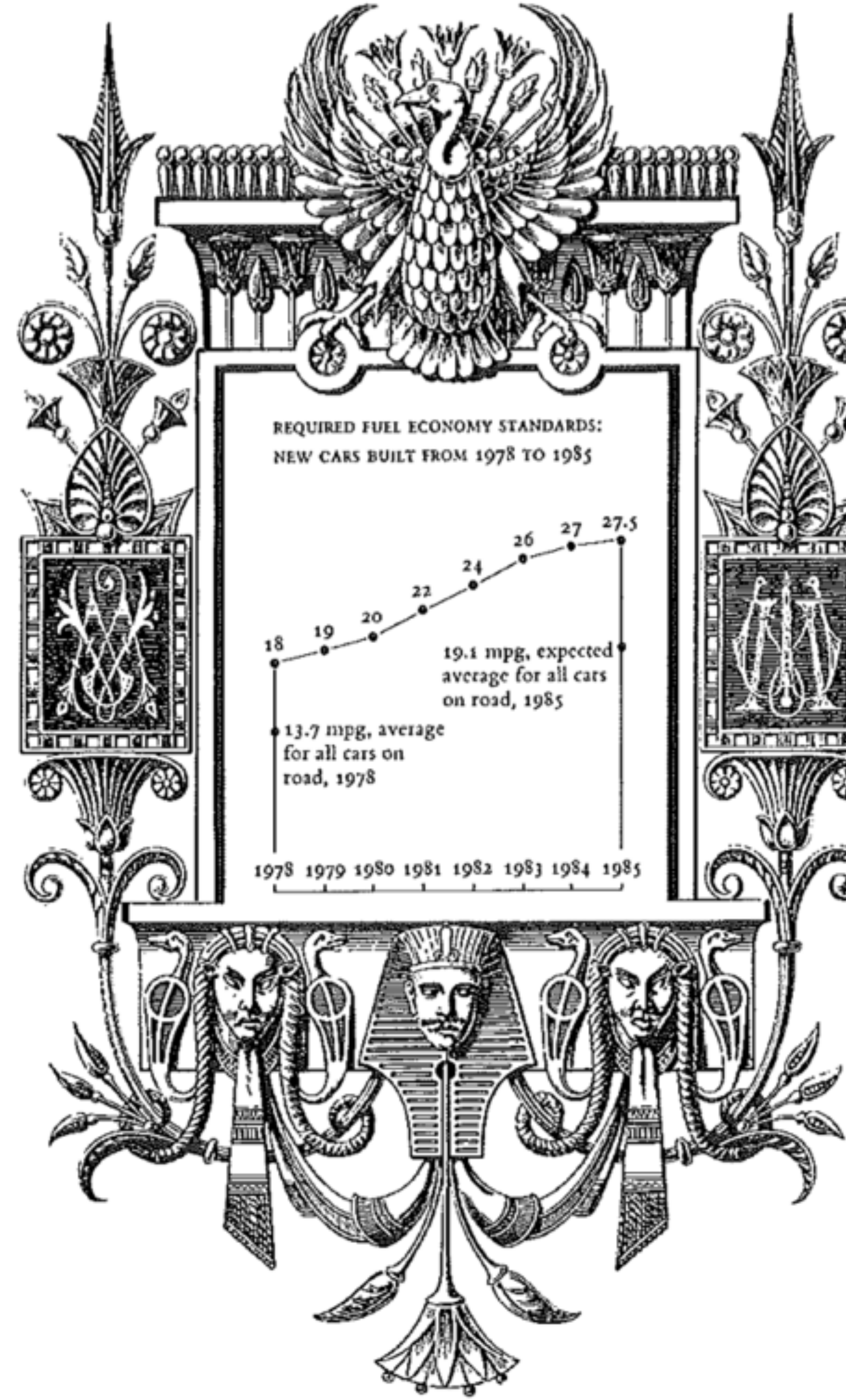
# CHART JUNK? IT DEPENDS

- persuasion
- memorability
- engagement

PROS

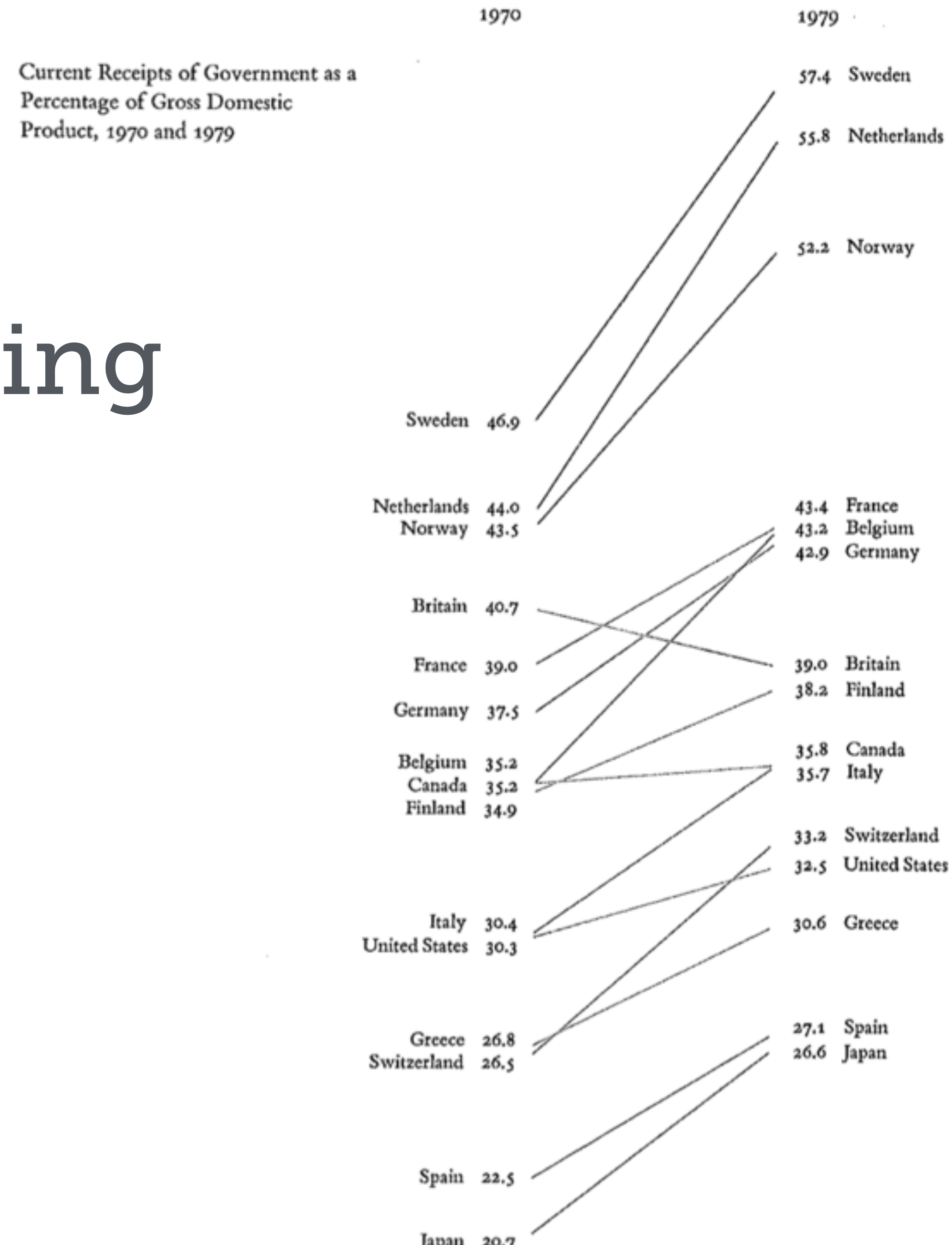
- unbiased analysis
- trustworthiness
- interpretability
- space efficiency

CONS

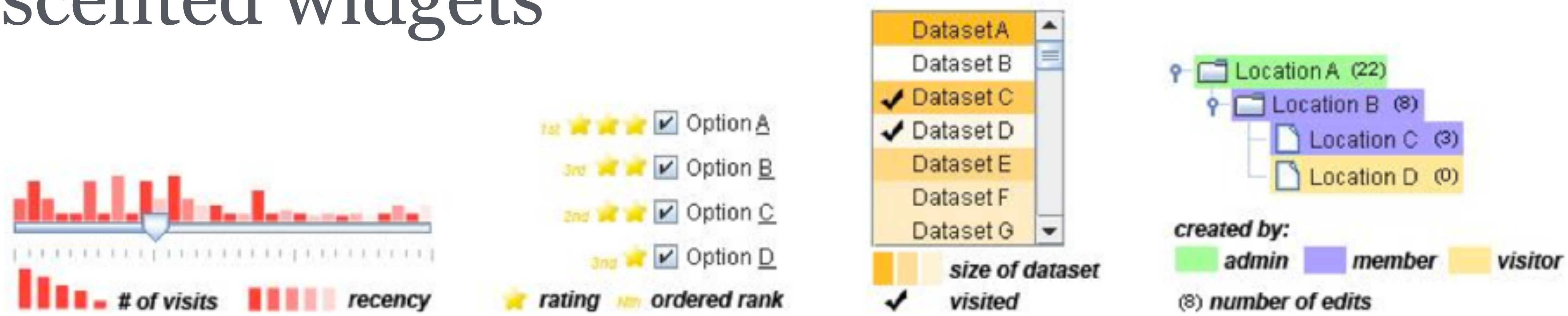

REQUIRED FUEL ECONOMY STANDARDS:
NEW CARS BUILT FROM 1978 TO 1985
18
19
20
22
24
26
27
27.5
13.7 mpg, average for all cars on road, 1978
19.1 mpg, expected average for all cars on road, 1985
1978
1979
1980
1981
1982
1983
1984
1985

Tufte 2001

# multifunctioning elements

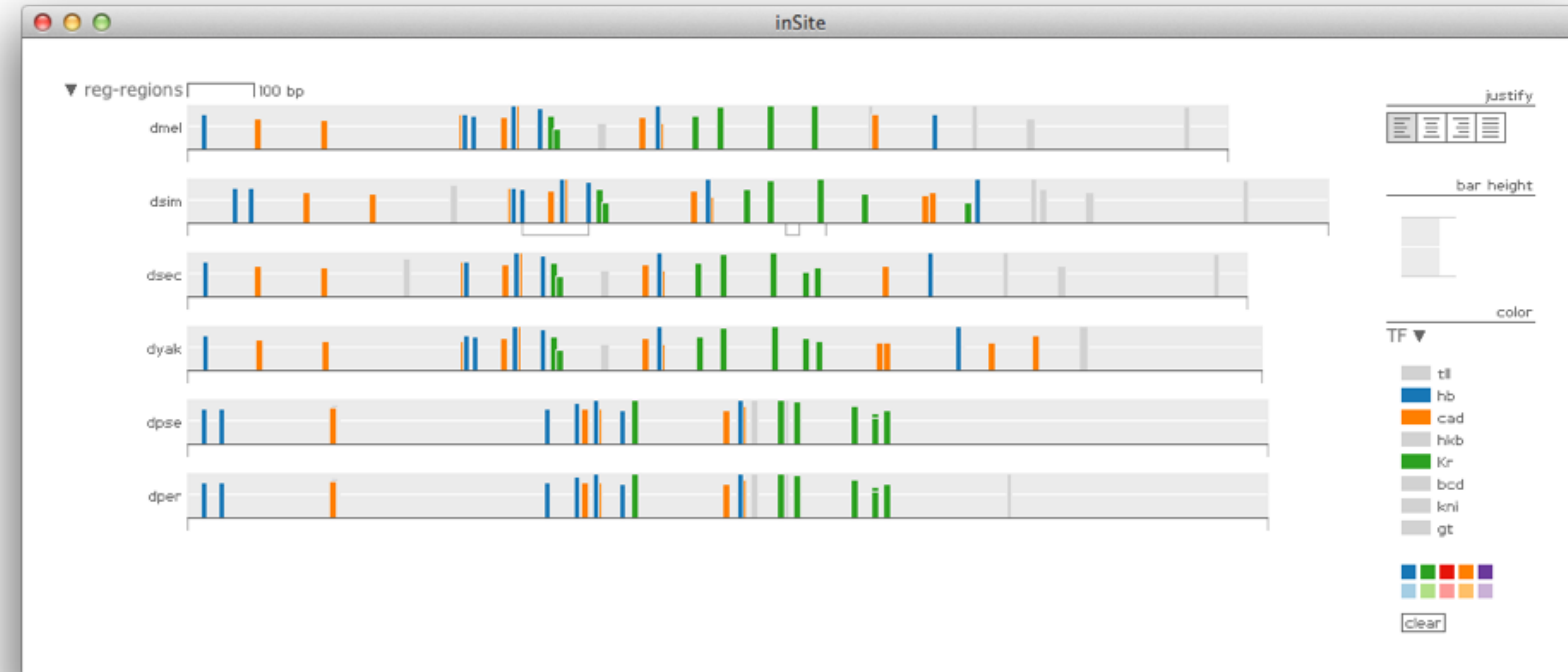


# scented widgets




Willett et al 2007

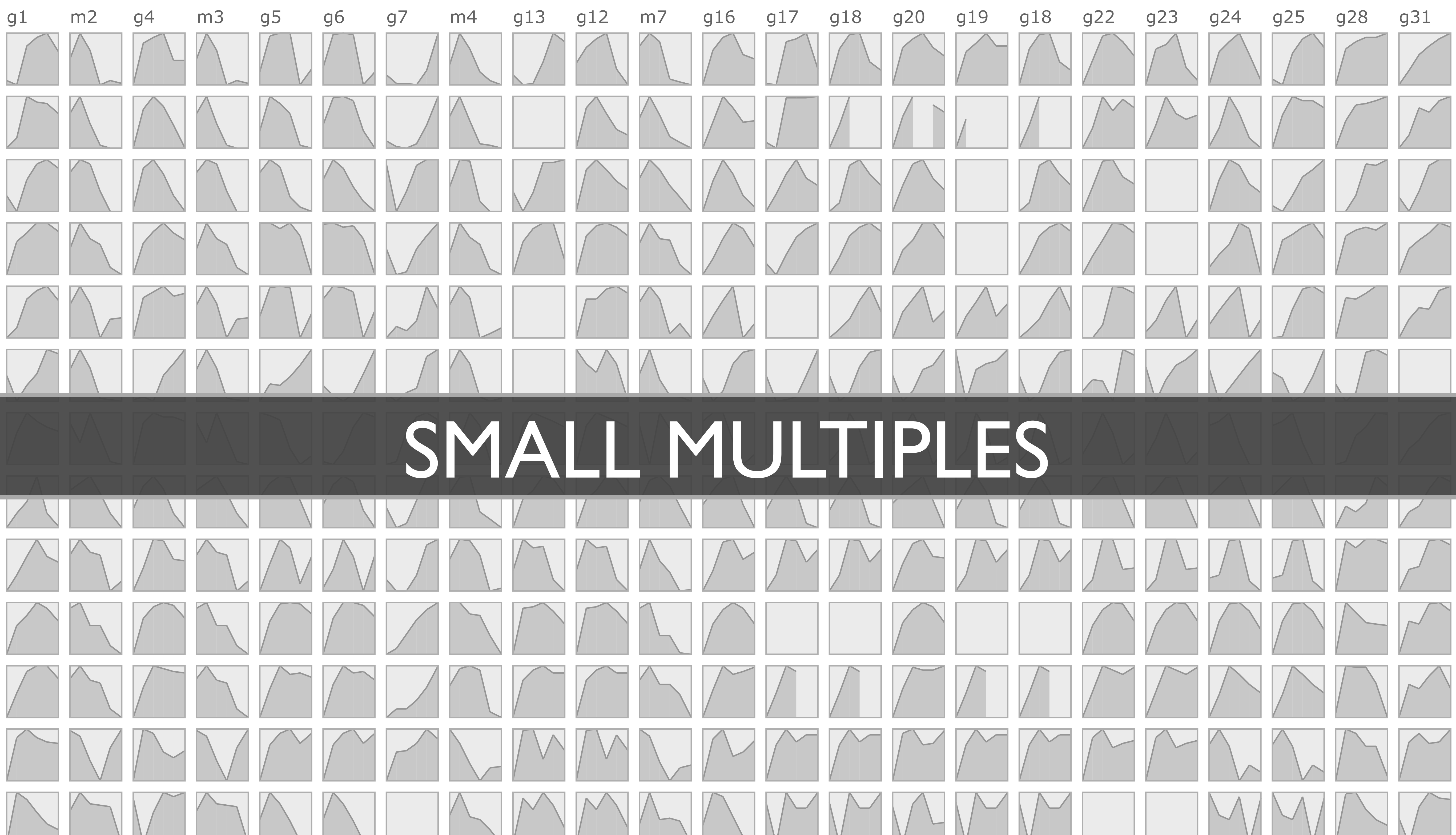
# interactive legend



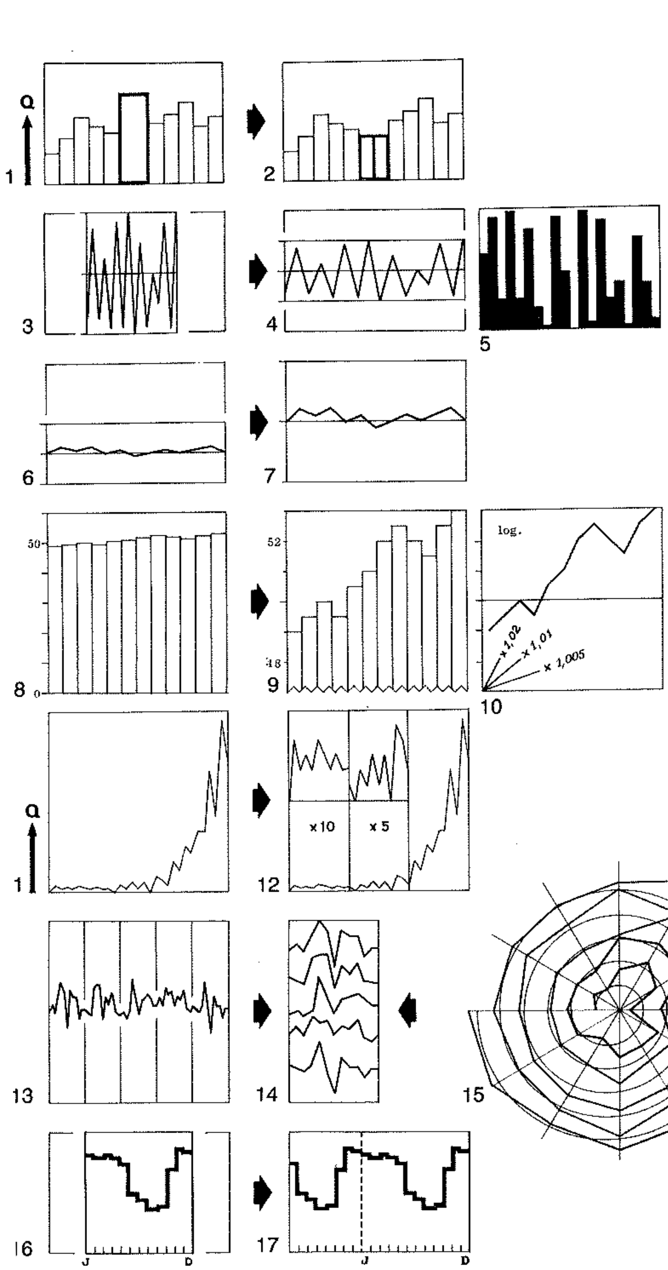
maximize the

$$\textbf{Data Density} = \frac{\text{number of entries in data array}}{\text{area of data graphic}}$$

# SHRINK THE GRAPHICS



# SHRINK THE GRAPHICS



GRAPHIC PROBLEMS POSED BY TIME SERIES

**Scale in years**
With a scale in years, a two-year total (figure 1) should be divided by 2 (figure 2). A total for six months should be multiplied by 2.

**Pointed curves**
For overly pointed curves (figure 3), the scale of the Q should be reduced; optimum angular perceptibility occurs at around 70 degrees (figure 4).
If the curve is not reducible (large and small variations), filled columns can be used (figure 5).

**Flat curves**
For overly flat curves (figure 6), the scale of the Q should be increased (figure 7).

**Small variations**
For small variations in relation to the total (figure 8), the total loses its importance, and the zero point can be eliminated, provided the reader is made aware of this elimination (figure 9). The graphic can be interpreted as an acceleration if a precise study of the variations is necessary; here, we use a logarithmic scale (figure 10). (See also page 240.)

**Large range**
For a very large range between the extreme numbers (figure 11), we must either:
(1) leave out the smallest variations;
(2) be concerned only with relative differences (logarithmic scale), without knowing the absolute quantities;
(3) select different parts (periods) within the ordered component and treat them on different scales above the common scale (figure 12).

**Obvious periodicity**
If there is obvious periodicity (figure 13), and the study involves a comparison of the phases of each cycle, it is preferable to break up the cycles in order to superimpose them (figure 14). A polar construction can be used, preferably in a spiral shape (figure 15), but we should not begin with too small a circle. As striking as it seems, it is less efficient than an orthogonal construction.

**Annual curves**
For annual curves of rainfall or temperature, if a cycle has two phases (figure 17), why depict only one (figure 16)?

**A contrast**
Unlike what we see in figure 18, the pertinent or "new" information must be separated from the background or "reference" information. The background involves: (a) the invariant, highlighted by a heading (Port St. Michel); (b) the highly visible identification of each component (tonnage and dates). The new information (the curve) must stand out from the background (figure 19).

**Reference points**
It is impossible to utilize a graphic such as figure 20, except in a general manner. There is confusion concerning the position of the points, and no potential comparison is possible, as it is in figure 21.

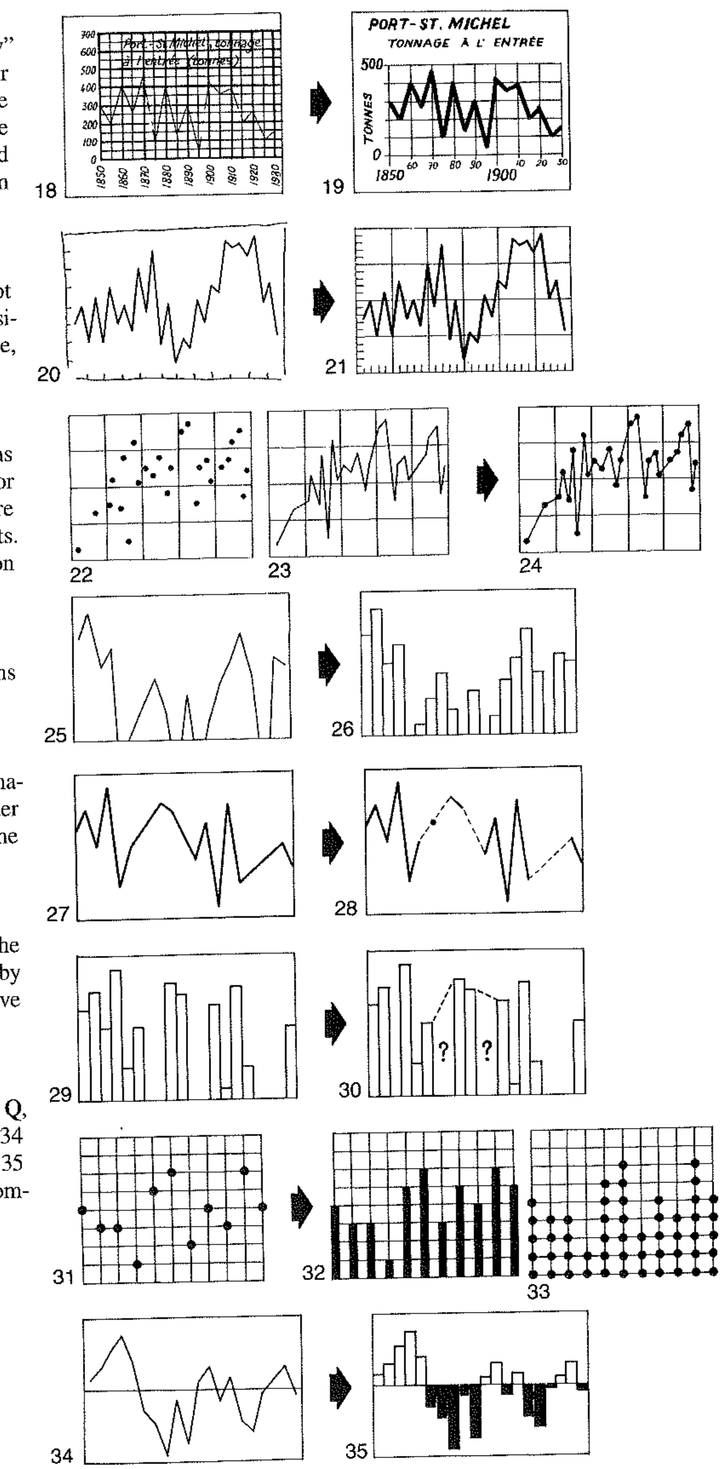
**Precision reading**
A precision reading (utilization on the elementary level, as in figure 24) is difficult in figure 22, which results in a poor reading of the order of the points, and in figure 23, where there is ambiguity concerning the position of the points. On the other hand, figure 22 does favor overall vision (correlation).

**Null boxes**
Curves accommodate null boxes poorly (figure 25). Columns (figure 26) are preferable.

**Unknown boxes**
The drawing must indicate the unknowns of the information in an unambiguous way (figures 28 and 30).The reader might interpret figure 27 as a change in the structure of the curve and figure 29 as involving null values.

**Very small quantities**
Except in seeking a correlation (quite improbable here) the number of ships entering into a port is represented better by figure 33 than by figures 31 or 32. The reader can perceive the numerical values at first glance.

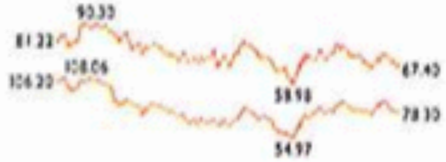
**Positive-negative variation**
This is in fact a problem involving three components O, Q, ≠ (+ –), and it must be visually treated as such. Figure 34 can be improved by utilizing a retinal variable (in figure 35 a value difference: black–white) to differentiate the ≠ component and thus highlight positive-negative variation.




Bertin 1967

# SHRINK THE GRAPHICS

*Dequantification* In exchange for an enormous increase in graphical resolving power, the wordlike size of sparklines precludes the overt labels and scaling of conventional statistical displays. Most of our examples have, however, depicted *contextual methods* for quantifying sparklines: the gray bar for normal limits and the red encoding to link data points in sparklines to exact numbers glucose 6.6 ; global scale bars and labels for sparkline clusters; and, probably best of all, surrounding a sparkline with an implicit data-scaling box formed by nearby numbers that label key data points (such as beginning/end, high/low) 1.1025 1.1907 1.0783 1.2858. And now and then sparklines might be scaled by very small type:



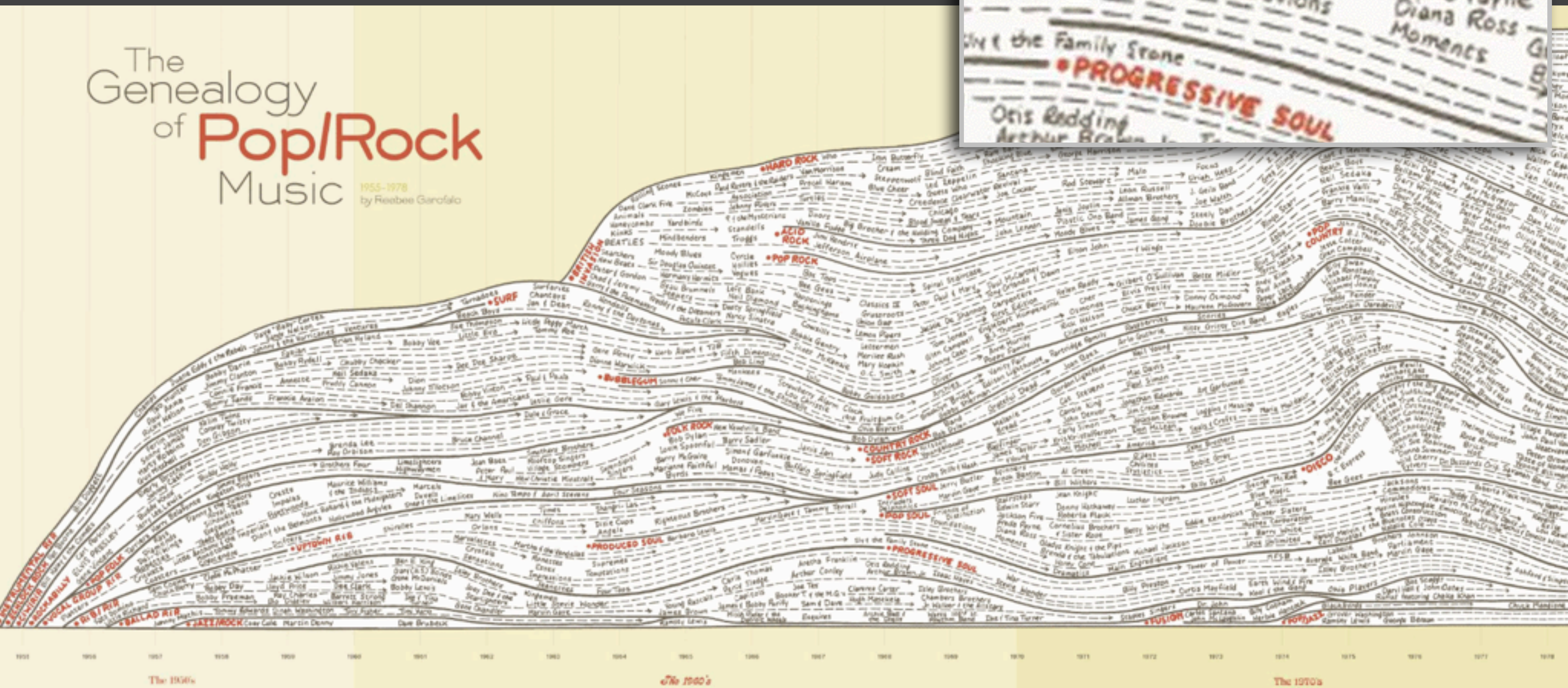
*Production methods* Data lines produced by conventional statistical graphics programs must be gathered together, rescaled, and resized into sparklines. Sometimes this can be quickly done by cutting and pasting data lines, then resizing the printed output to sparkline resolutions.

SPARKLINES

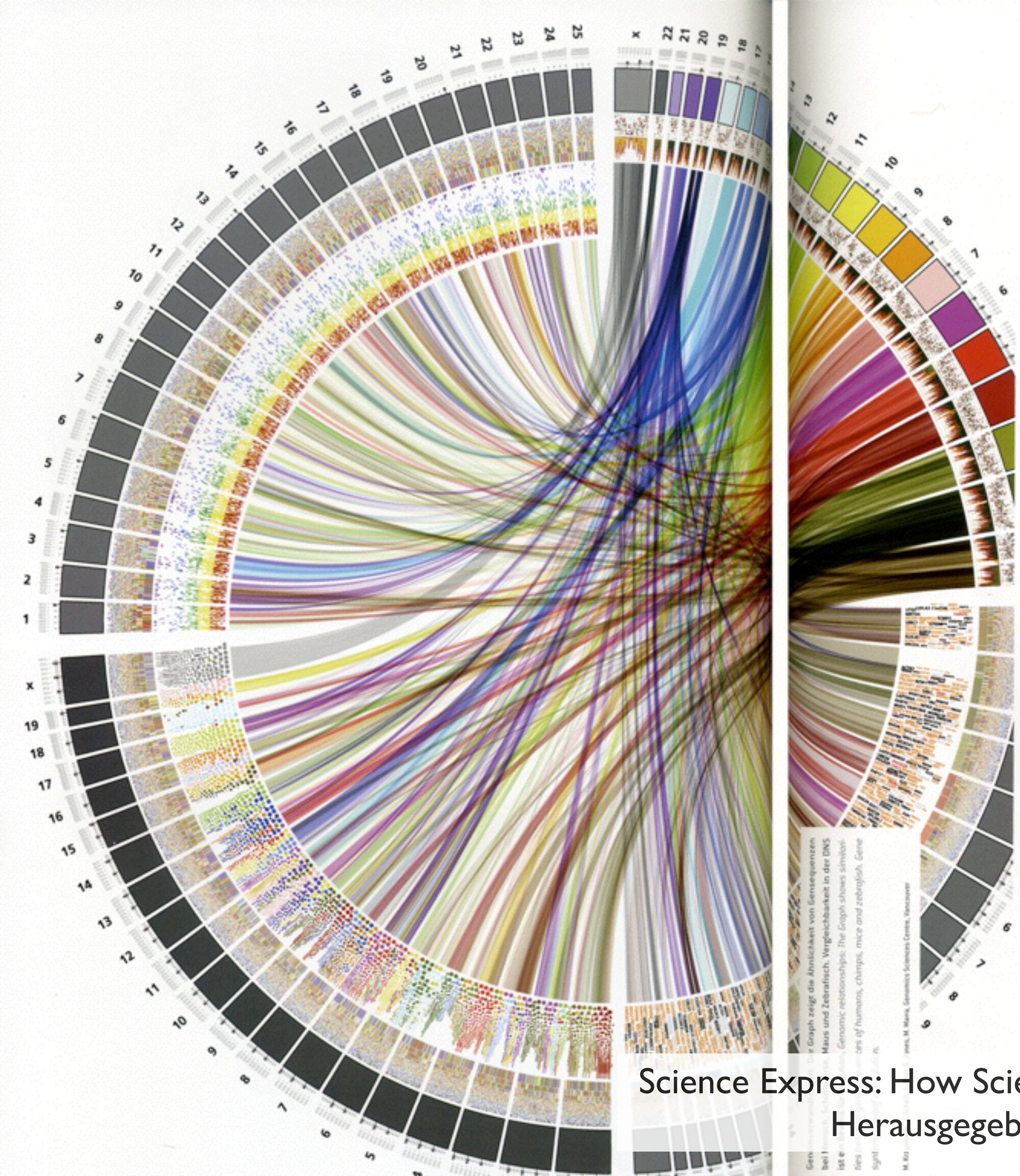
(3) a *statistical analysis* program to generate hundreds of chartjunk-free sparklines for export into design and layout operations. Once the basic templates for sparklines are worked out, then ongoing production and

Tufte 2006

# maximize the amount of data shown


Steve Chappel and Reebe Garofalo
in Rock 'N' Roll is Here to Pay: The History and Politics of the Music Industry, 1977



On the road to a digital society — **Computer technology** is an ubiquitous element of our world, and fast networks are spanning the globe. This is changing the way we live and work and communicate. A new digital world is emerging, an environment in which creativity and innovation can flourish in many new ways. As a result, science and research have a greater influence on our life in the 21st century than ever before. This is attributable to massive investments in research and development, but also to intensive cooperation and tough competition. The convergence of nano-, bio-, information- and neurotechnologies facilitates completely new applications. Taking its place beside the more traditional factors of land, capital and employment, knowledge is fast becoming the decisive factor for prosperity – and also for the resolution of global problems. In this, the appropriate balance between digital freedom and digital security must be maintained. — **Science 2020: Systematically surveying the world** *Millions of scientists are getting to the bottom of the secrets of our world, across the whole spectrum of space, time, energy and complexity. Fundamentally new knowledge is emerging from research into inter-disciplinary topics or extreme states of matter. Science long ago escaped the constraints of working only in the realm of our natural living conditions and our perceptions. Considerable investment is flowing into efforts to decode the smallest building blocks of our world and to understand how their interplay produces brand new qualities. The drivers of innovation in research today are data capture via digital sensors; storage, analysis and visualisation via computer and software; and the global exchange of information and knowledge.* — **The cost of new knowledge is rising** There is now no part of our life that is not the subject of research. At the same time, it is becoming ever more difficult to generate new knowledge. These days, new research methods and technologies enable us to study even the ›farthest frontiers‹ of the world: extremely fast or slow processes, the tiniest building blocks or the largest structures, extreme cold or extreme heat. — **Networked knowledge takes on global challenges** *Thanks to worldwide information and communication networks, the challenges our civilisation faces in the long term are known to us sooner and more clearly than ever before. We can start developing solutions together at an earlier stage. Research on many topics is global – taking place in close cooperation or in international competition for the fastest and best solutions. National boundaries are becoming irrelevant. Millions of scientists work across countries, continents and time zones in thousands of labs. Their global networking enhances the diversity and efficiency of science and technology. And this, in turn, reinforces globalisation and networking. In a world changing at such a pace, each country must redefine its place.* — **The end of distance** Mankind faces enormous challenges both locally and globally – the challenge of using resources sustainably and of organising a global economy. Across the globe, complex processes are being recorded in detail, collated in databases and analysed in computer networks. New visualisation techniques make it possible to analyse larger and larger data records and to draw conclusions from the results. — **Global networking as the driving force of science** In the early days, the Internet linked up scientists, large-scale equipment and information; now it networks computational power and enormous amounts of data through grid and cloud computing. A global Semantic Web is emerging, bringing together data, expertise and knowledge that had previously been distributed among virtual libraries and observatories. The information is being intelligently developed, new forms of cooperation are arising, and research is becoming more productive [illegible] development. Digital technologies enable people all over the world to use new [illegible] and improve equal opportunities, it is those in the least developed countries who must have access to the Internet. This is where new knowledge is freely available.



Science Express: How Science and Technology change our life.
Herausgegeben von der Max-Planck-Gesellschaft

# Unseen and Unaware: Implications of Recent Research on Failures of Visual Awareness for Human–Computer Interface Design

**D. Alexander Varakin** and **Daniel T. Levin**
*Vanderbilt University*

**Roger Fidler**
*Kent State University*

COUNTER-POINT

**ABSTRACT**

Because computers often rely on visual displays as a way to convey information to a user, recent research suggesting that people have detailed awareness of only a small subset of the visual environment has important implications for human–computer interface design. Equally important to basic limits of awareness is the fact that people often over-predict what they will see and become aware of. Together, basic failures of awareness and people's failure to intuitively understand

# ILLUSIONS OF VISUAL BANDWIDTH

people over-predict what they will see and become aware of

# next time…

DOI:10.1145/1743546.1743567

Article development led by acmqueue
queue.acm.org

**A survey of powerful visualization techniques, from the obvious to the obscure.**

BY JEFFREY HEER, MICHAEL BOSTOCK, AND VADIM OGIEVETSKY

# A Tour Through the Visualization Zoo

THANKS TO ADVANCES in sensing, networking, and data management, our society is producing digital information at an astonishing rate. According to one estimate, in 2010 alone we will generate 1,200

- homework
    - assignment 12 due Tuesday