VISUALIZING MULTI-ATTRIBUTE DATA

DATA TABLES

Petra Isenberg

informatics mathematics
Inria

RECAP

you have learned about

- visual channels and marks
- that their perceptual properties matter

RECAP

Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

Same (Color luminance, Color saturation)

Same (Curvature, Volume (3D size))

Effectiveness: Most ▲ … ▼ Least

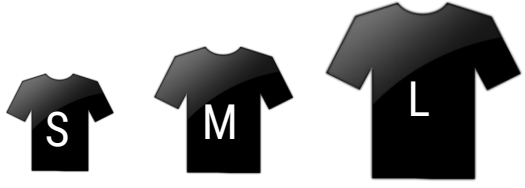
Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape

RECAP

DATA TYPES

ORDINAL (ranking)

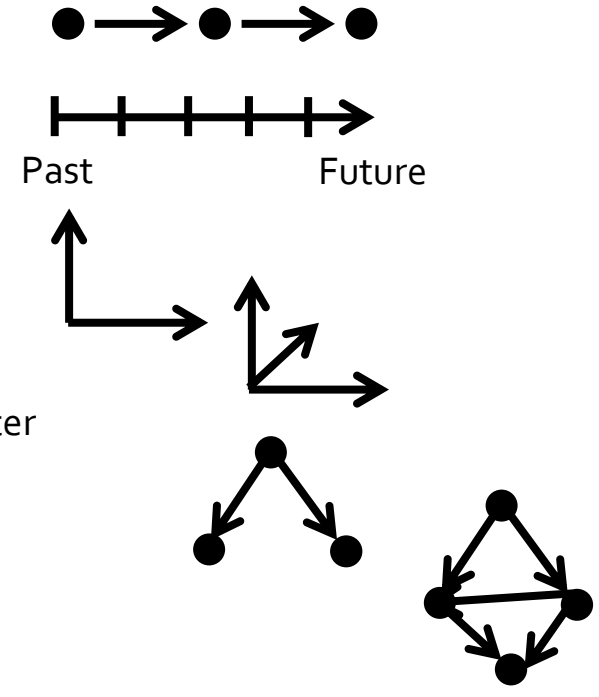


NOMINAL (categorical)

QUANTITATIVE (numerical)

OTHER VIEWS ON DATA TYPES

- 1D (linear)
- Temporal
- 2D (maps)
- 3D
- nD (relational)
- Trees (hierarchies)
- Networks (graphs)

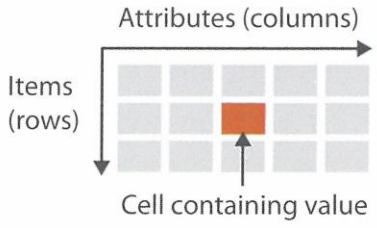


vis examples later

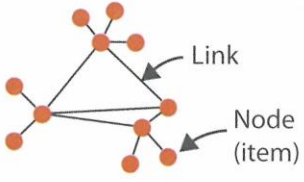
Shneiderman: The Eyes Have It

ANOTHER VIEW

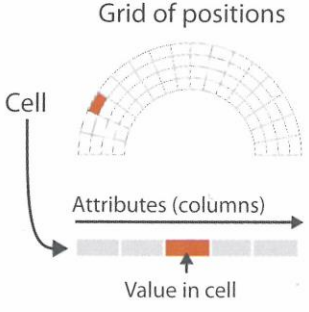
➔ Tables



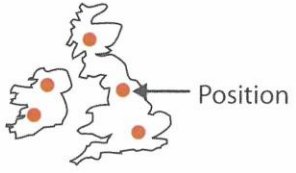
➔ Networks



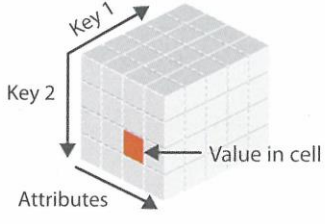
➔ Fields (Continuous)



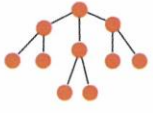
➔ Geometry (Spatial)

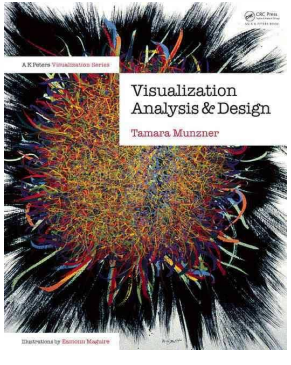


➔ *Multidimensional Table*



➔ *Trees*





TODAY

How to turn something like this … into a muti-dimensional data representation

Vélib' : Disponibilité temps réel

Informations | Tableau | Carte | Analyse | Export | API

	Code de la station	Nom de la station	Etat des stations	Etat du Totem	Nombres de bornes en station	Nombre de bornes disponibles	Nombre de vélo mécanique	Nombre vélo électrique	Achat possible en
1	11037	Faubourg Du Temple - Republique	Close	yes	39	36	2	1	yes
2	11104	Charonne - Robert et Sonia Delauney	Operative	yes	20	17	2	1	no
3	14111	Cassini - Denfert-Rochereau	Close	yes	25	25	0	0	yes
4	12109	Mairie du 12ème	Operative	yes	30	30	0	0	no
5	5110	Lacépède - Monge	Operative	yes	23	16	6	1	yes
6	17038	Grande Armée - Brunel	Operative	yes	62	40	20	2	yes
7	10152	Gare du Nord - Place de Valenciennes	Operative	yes	25	18	6	1	yes
8	13007	Le Brun - Gobelins	Operative	yes	48	47	1	0	yes
9	41301	Bois de Vincennes - Gare	Operative	yes	51	39	8	4	yes
10	31024	Romainville - Vaillant-Couturier	Operative	yes	38	35	2	1	no
11	15028	Grenelle - Dr Finlay	Operative	yes	71	63	7	2	yes
12	16118	Michel-Ange - Parent de Rosan	Operative	yes	26	25	0	1	no
13	20035	Pyrénées - Ménilmontant	Operative	yes	26	24	0	2	no
14	10027	Dunkerque - Alsace	Operative	yes	18	10	8	0	no
15	8048	Marceau - Chaillot	Operative	yes	21	8	9	4	no
16	14013	Liard - Amiral Mouchez	Operative	yes	1	1	1	0	yes
17	5024	Place Monge	Close	yes	21	21	0	0	no
18	7018	Ségur - d'Estrées	Operative	yes	19	8	10	1	no
19	10029	Dunkerque - Rocroy	Operative	yes	23	16	6	1	no
20	8009	Gare Saint-Lazare - Isly	Operative	yes	27	14	10	3	yes
21	8036	Lisbonne - Monceau	Operative	yes	33	7	25	1	yes
22	17040	Pereire - Ternes	Operative	yes	48	44	2	2	yes
23	31708	Noisy le Sec - Jean-Baptiste Clément	Operative	yes	22	21	0	1	no
24	10105	Mazagran - Bonne Nouvelle	Operative	yes	26	15	9	2	yes

SOME INSPIRING EXAMPLES

TOP SONGS

by album
spotify popularity

80
60
40
20

Greetings from AP 1973
E Street Shuffle 1973
Born to Run 1975
Darkness on Edge of Town 1978
The River 1980
Nebraska 1982
Born in the USA 1984
Tunnel of Love 1987
Human Touch 1992
Lucky Town 1992
Ghost of Tom Joad 1995
The Rising 2002
Devils & Dust 2005
Magic 2007
Working on Dream 2009
Wrecking Ball 2012
High Hopes 2014

song types
energy & valence

5
10
15

us album sales rank
petals = E Street member

TOP ALBUMS

the history of BRUCE SPRINGSTEEN

Over a 45 year career Bruce Springsteen has released over 200 songs and 17 studio albums. Born in Long Branch NJ and raised in working class Freehold NJ he is known for his poetic lyrics and energetic stage performances. He has sold 135 million albums worldwide. The following is a musical history of Bruce Springsteen using excerpts from his autobiography "Born to Run", interviews, and Wikipedia.

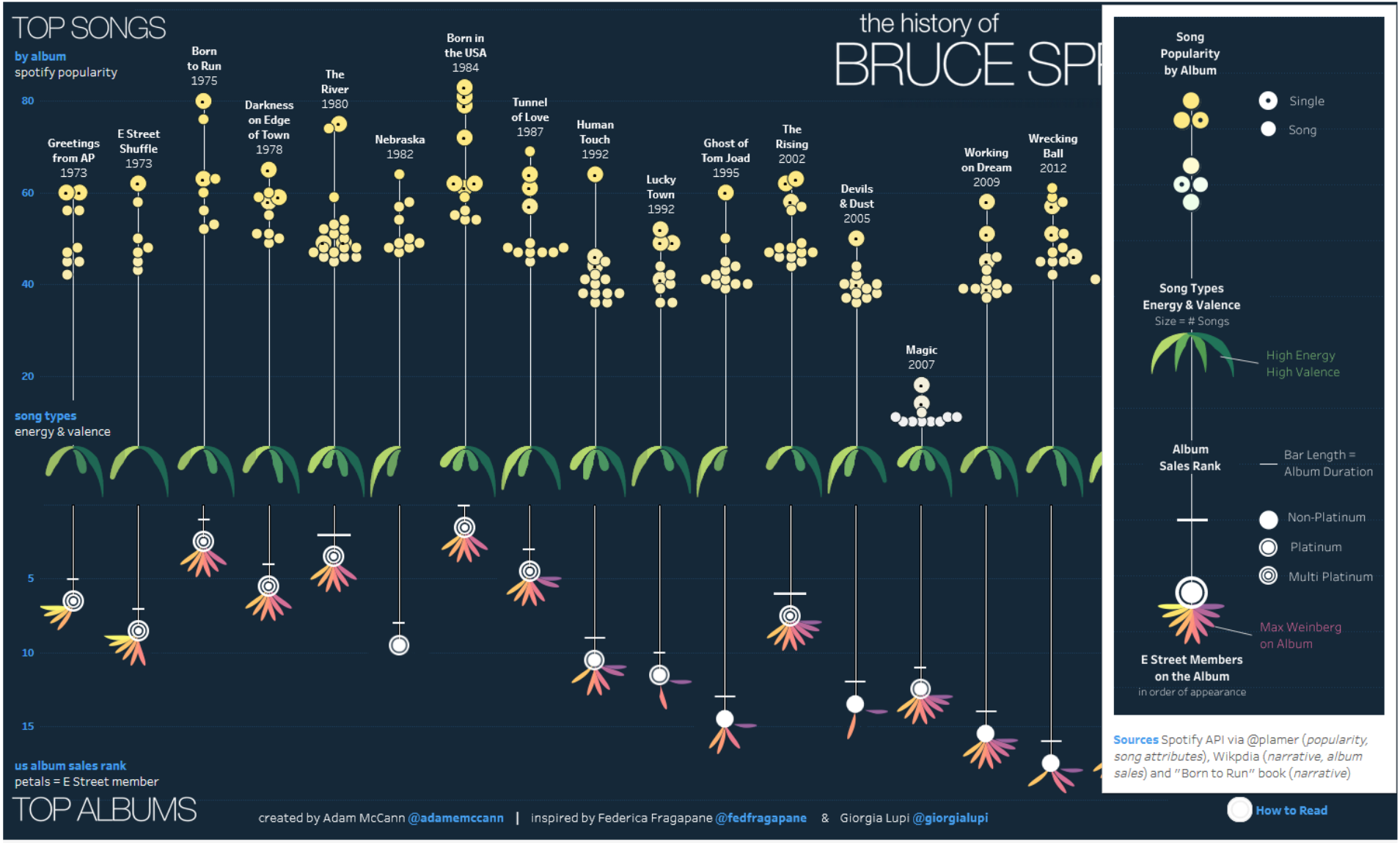
John Hammond of Columbia Records signed Bruce after he played Saint in the City for him in his office. Columbia wanted a solo artist like Bob Dylan and James Taylor but Bruce insisted on recording with a band to get the new sound he was pursuing: "an amalgam of good song-writing mixes with a soul-and-R&B influced, rock music" as Bruce said in his autobiography "Born to Run".

Greetings from Asbury Park, N.J. is Bruce's debut studio album released on January 5, 1973. After completing the album Clive Davis of Columbia Records was worried there were no hits. So Bruce wrote "Spirit in the Night" and "Blinded by the Light" the next day at the beach. They were the two biggest hits on the album. Early reviews of the album made favorable comparisons to Dylan.

The Wild, the Innocent & the E Street Shuffle is Bruce's second album released on November 11, 1973. Bruce recorded this album over 3 months at the 914 Sound Studios in Blauvelt NY. Bruce and Clarence Clemons slept in a tent behind the studio to save money. Like Greetings from Asbury Park the album was not a commercial success but received positive

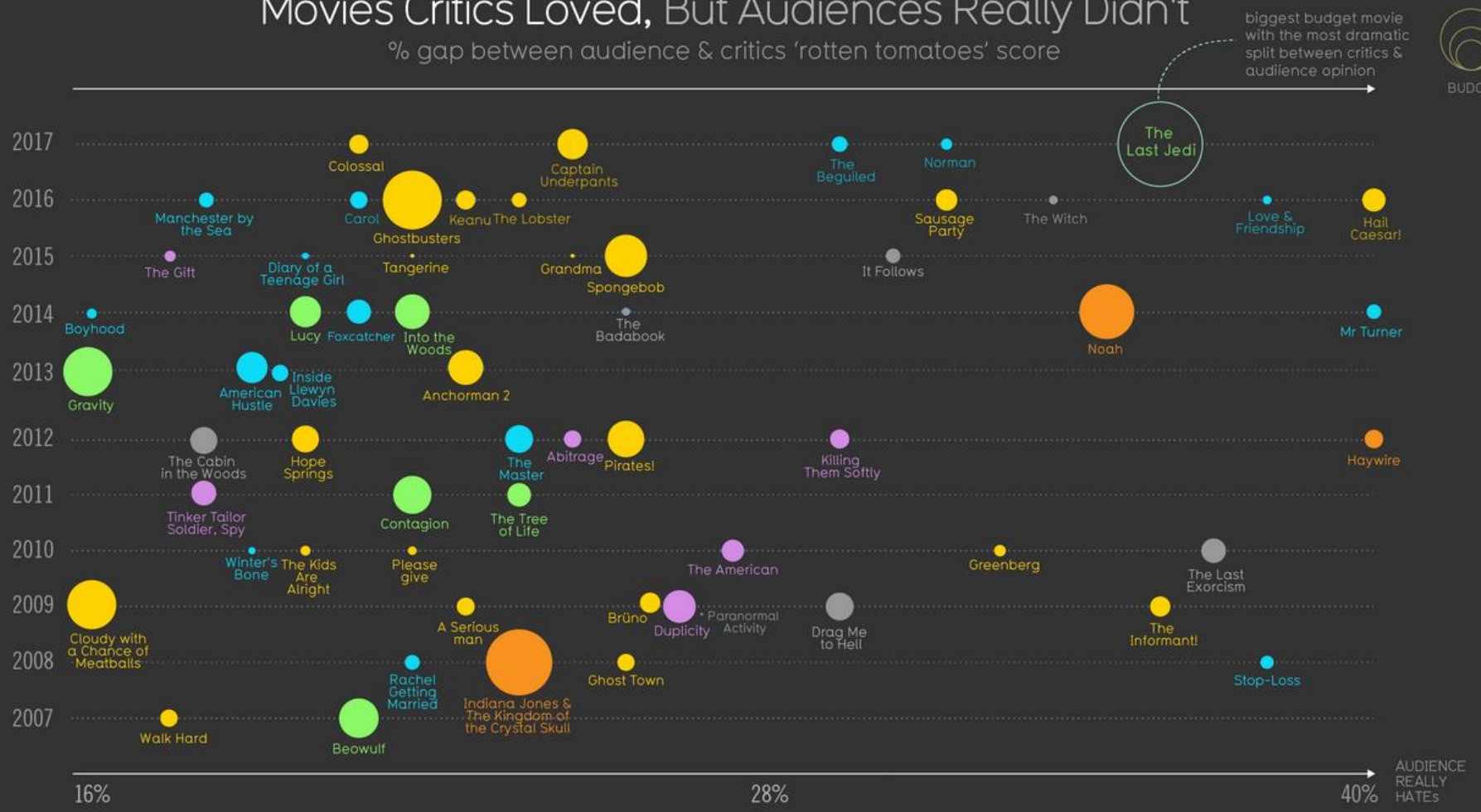
created by Adam McCann @adamemccann | inspired by Federica Fragapane @fedfragapane & Giorgia Lupi @giorgialupi

How to Read


TOP SONGS
by album
spotify popularity
the history of
BRUCE SPR
Greetings from AP 1973
E Street Shuffle 1973
Born to Run 1975
Darkness on Edge of Town 1978
The River 1980
Nebraska 1982
Born in the USA 1984
Tunnel of Love 1987
Human Touch 1992
Lucky Town 1992
Ghost of Tom Joad 1995
The Rising 2002
Devils & Dust 2005
Magic 2007
Working on Dream 2009
Wrecking Ball 2012
80
60
40
20
song types
energy & valence
5
10
15
us album sales rank
petals = E Street member
TOP ALBUMS
Song Popularity by Album
Single
Song
Song Types Energy & Valence
Size = # Songs
High Energy
High Valence
Album Sales Rank
Bar Length = Album Duration
Non-Platinum
Platinum
Multi Platinum
Max Weinberg on Album
E Street Members on the Album
in order of appearance
Sources Spotify API via @plamer (popularity, song attributes), Wikpdia (narrative, album sales) and "Born to Run" book (narrative)
created by Adam McCann @adamemccann | inspired by Federica Fragapane @fedfragapane & Giorgia Lupi @giorgialupi
How to Read

Movies Critics Loved, But Audiences Really Didn't

% gap between audience & critics 'rotten tomatoes' score



Comedy Sci-fi / Fantasy Action / Adventure Drama Thriller Horror

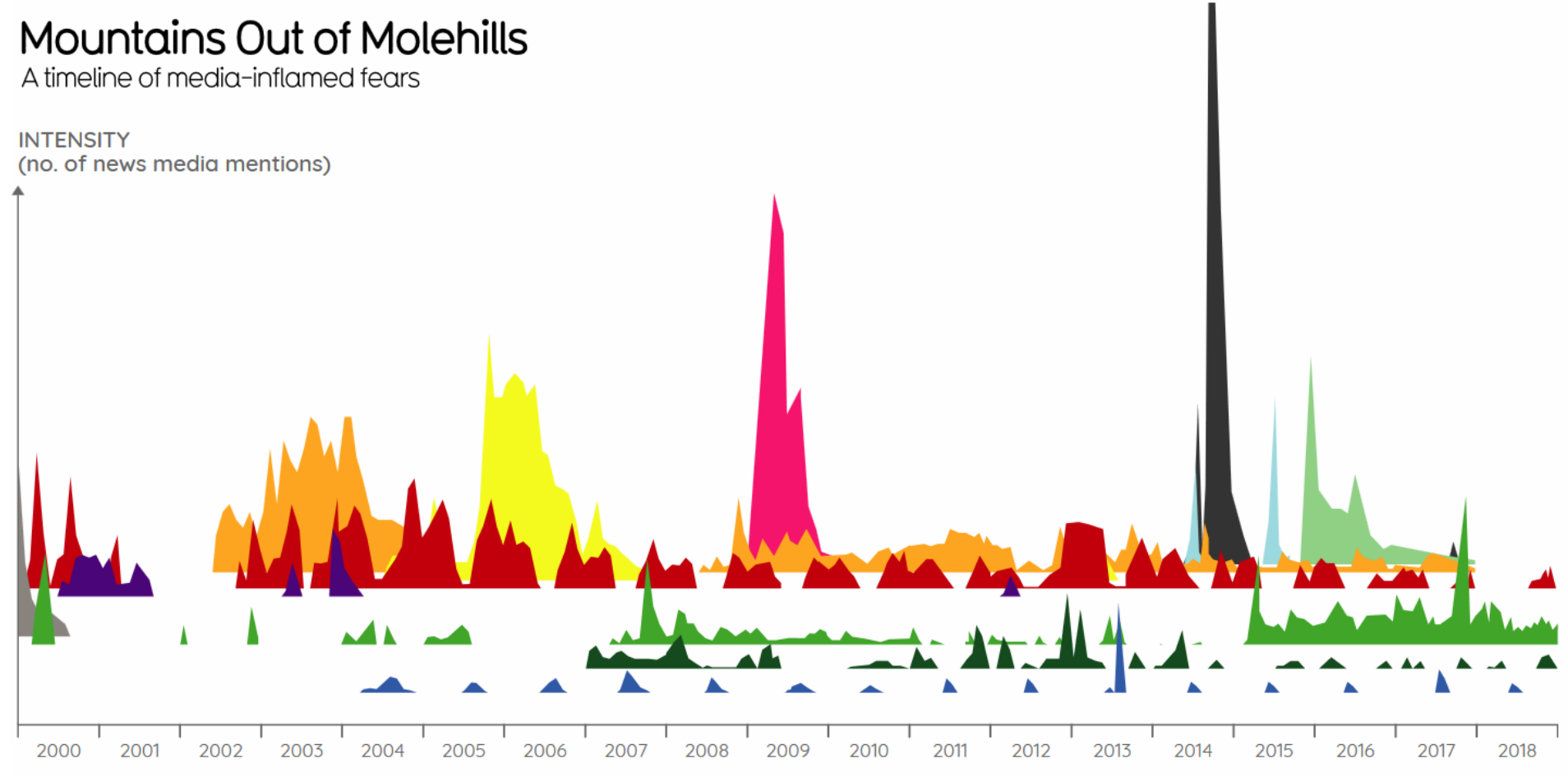
source: RottenTomatoes.com

data: bit.ly/IIB_MovieLove

informationisbeautiful.net

Mountains Out of Molehills

A timeline of media-inflamed fears

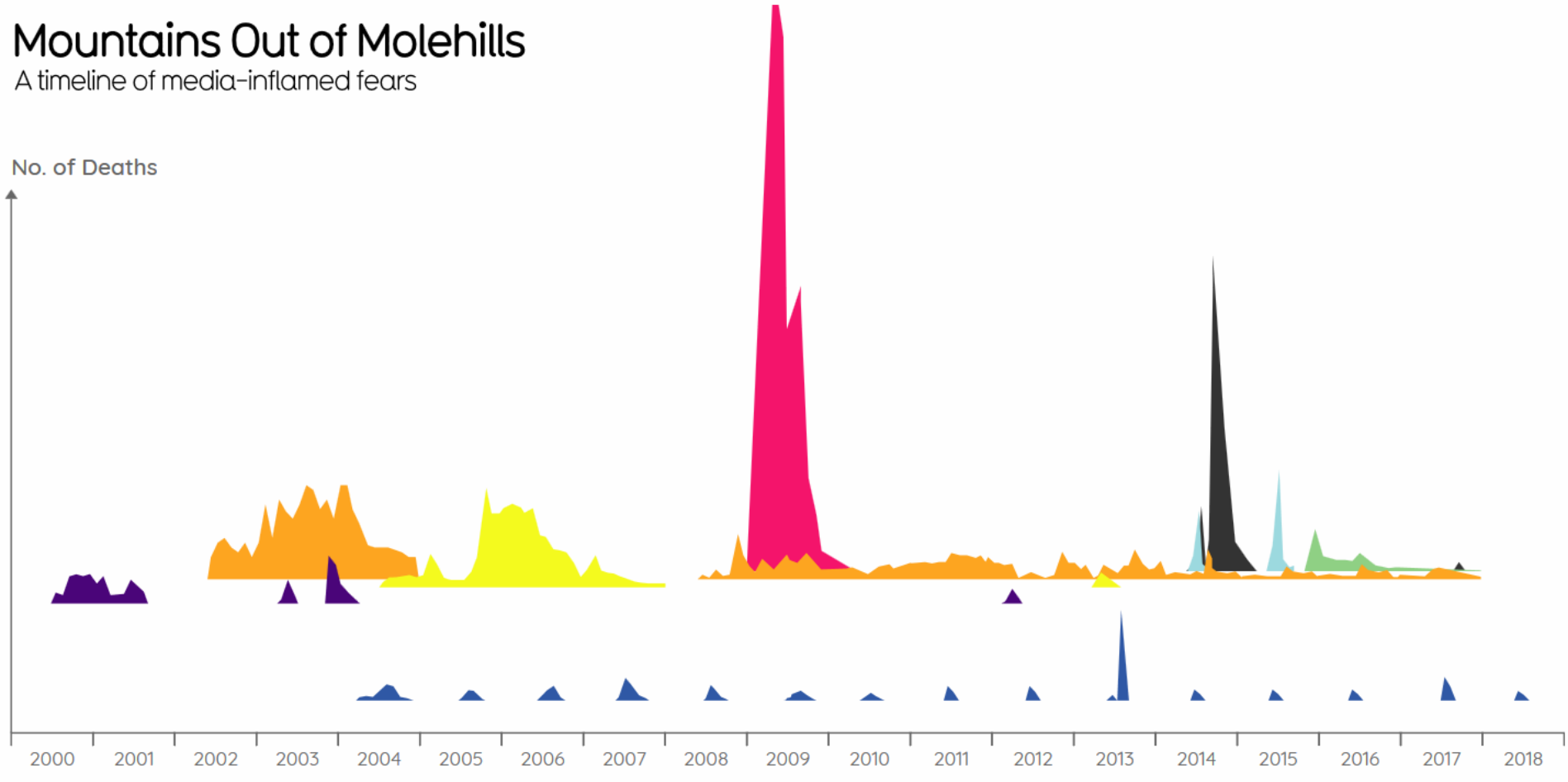


David McCandless, Fabio Bergamaschi // source: Google Trends

align to baseline scale to fit Ebola scale by deaths

Mountains Out of Molehills

A timeline of media-inflamed fears



ASTEROIDS | BIRD FLU | EBOLA | KILLER WASPS | MAD COW DISEASE | MERS | MILLENIUM BUG | SARS | SWINE FLU | VACCINES & AUTISM | VIOLENT VIDEO GAMES | ZIKA

David McCandless, Fabio Bergamaschi // source: Google Trends

align to baseline | scale to fit Ebola | scale by deaths

THE MANY MOONS OF JUPITER

IN 1610, Galileo Galilei gazed up through his telescope in the direction of Jupiter. In that moment he likely became the first person to see a moon other than our own, as light that had left the vicinity of the gas giant around half an hour earlier crashed into his pupils and revealed four dotted silhouettes. These Galilean moons, one of which is even larger than the planet Mercury, became the opening entries into a collection that is still increasing today. In fact in 2018, 407 years after the Italian polymath made his discovery, scientists confirmed the existence of 12 more moons locked in slow rotation with the largest planet in our solar system. These newly found satellites form part of a diverse family, many of which share little commonality other than their gravitational anchor. Their orbital shapes range from near perfect circles to highly eccentric and inclined. Their scales vary hugely, from the size of planets to just a kilometer across. Some may have been asteroids captured by Jupiter's powerful gravitational pull, while others were likely a by-product of the very formation of the planet itself.

This data visualization displays every currently known moon of Jupiter, each featuring the year of discovery, discoverer and a representation of scale. Additionally, on the right are some additional insights about the moons. Finally, while all information is correct as of 2018, scientists are finding new wonders in our solar system every day; so who knows how many new Jovian moons are out there right now, held in endless revolutions, just waiting for eyes to meet them for the first time?

79
Moons discovered as of 2018

18
The number of moons that are prograde

The number of moons that are retrograde **61**

The Voyager probes, launched in 1977, discovered three of Jupiter's large inner moons

99.997%
The approx. % of the total mass in orbit around Jupiter that comes from the four Galilean moons

One of the newly discovered moons has an odd prograde orbit which sees its path cross several other retrograde moons. This means a collision is very likely, although scientists predict it could take another billion years to actually happen

As of 2018, five of the moons are considered lost

Ganymede, Jupiter's largest moon, is actually 8% larger than the planet Mercury, making it the 9th largest object in our solar system

It's thought that several of the larger moons could feature subsurface oceans, leading to some exciting possibilities about the existence of life there!

54
The most moons have been found by a team led by Scott Sheppard

A German astronomer called Simon Marius independently discovered the four Galilean moons at the same time as Galileo. While he didn't receive the title of discoverer, he is responsible for their names, which are still used today

Reading the data visualization

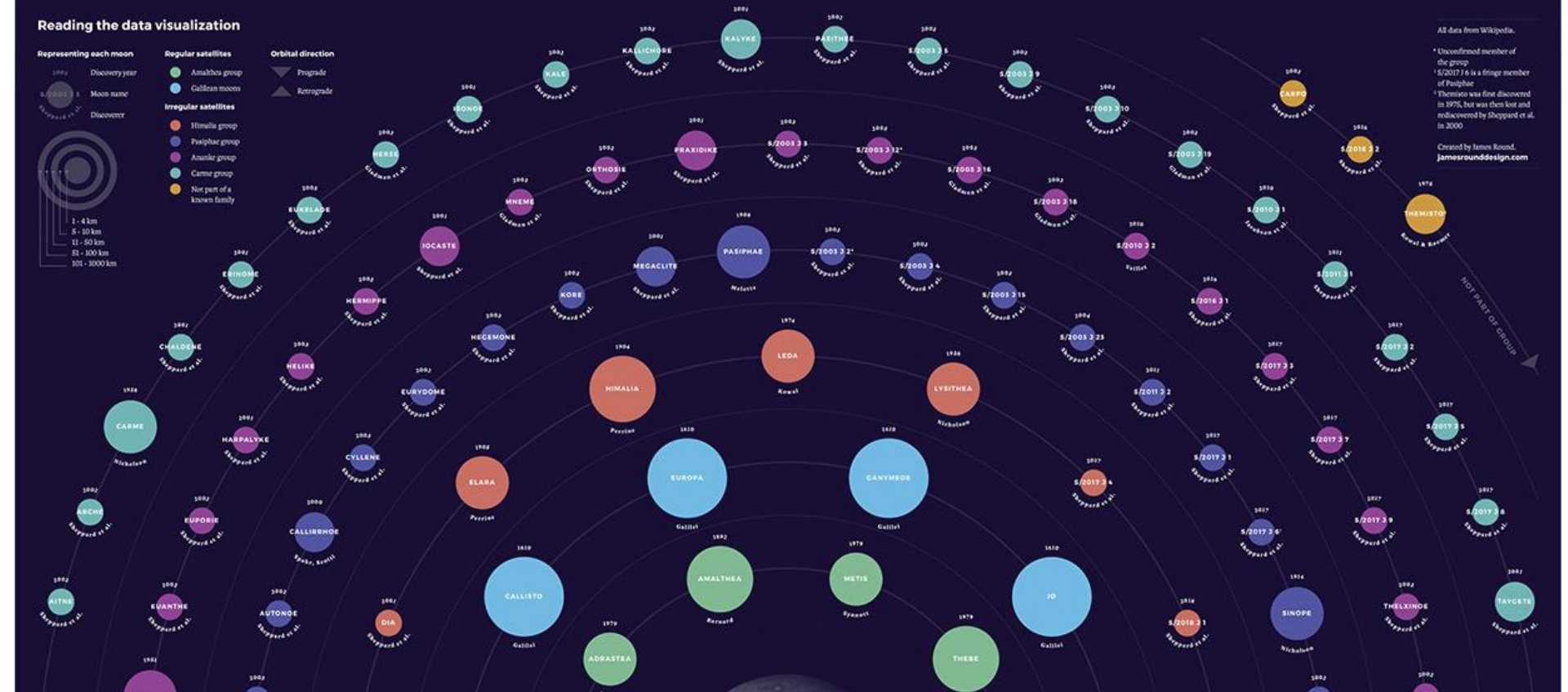
Representing each moon
- Discovery year
- Moon name
- Discoverer

Regular satellites
- Amalthea group
- Galilean moons

Irregular satellites
- Himalia group
- Pasiphae group
- Ananke group
- Carme group
- Not part of a known family

Orbital direction
- Prograde
- Retrograde

1 - 4 km
5 - 10 km
11 - 50 km
51 - 100 km
101 - 1000 km



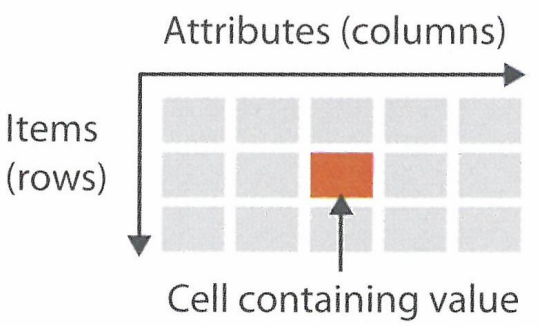
All data from Wikipedia.

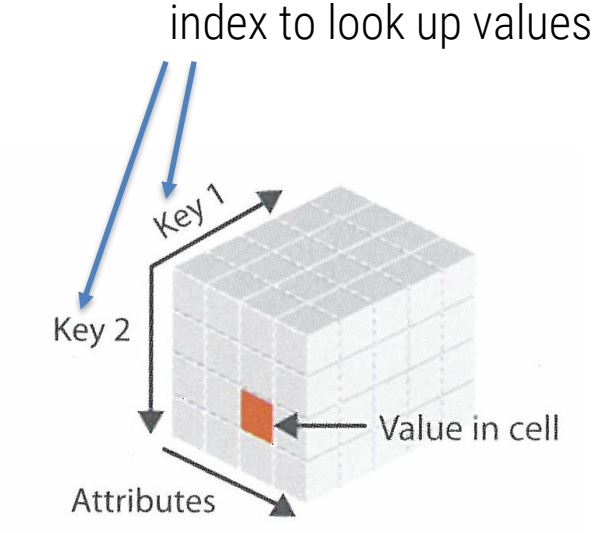
* Unconfirmed member of the group
[1] S/2017 J 6 is a fringe member of Pasiphae
[2] Themisto was first discovered in 1975, but was then lost and rediscovered by Sheppard et al. in 2000

Created by James Round.
jamesrounddesign.com

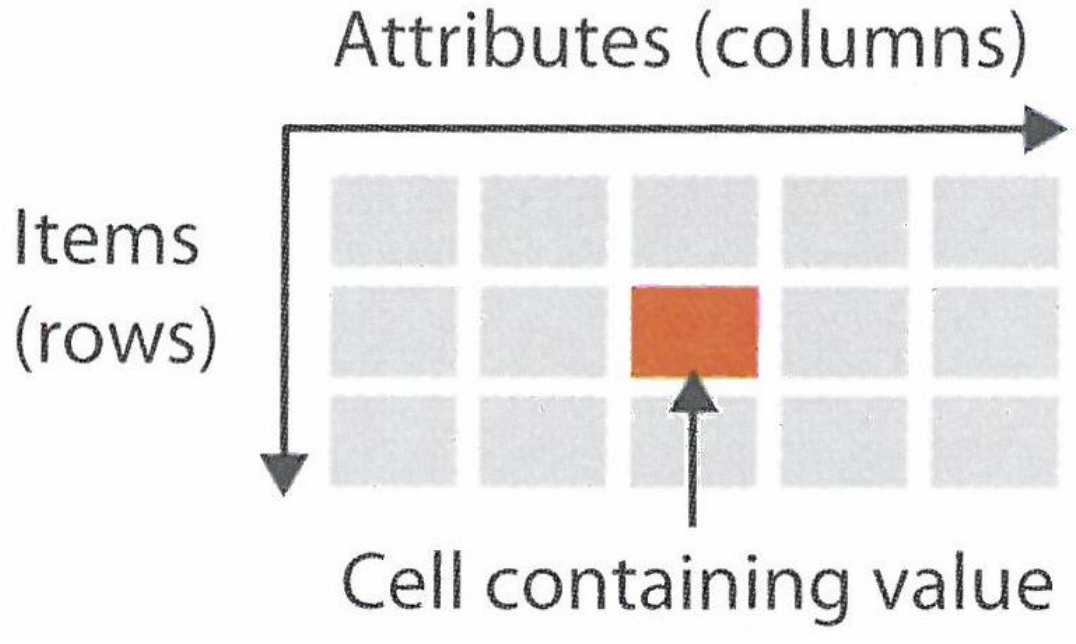
BACK TO THE BASICS

DATA TABLES - TERMINOLOGY

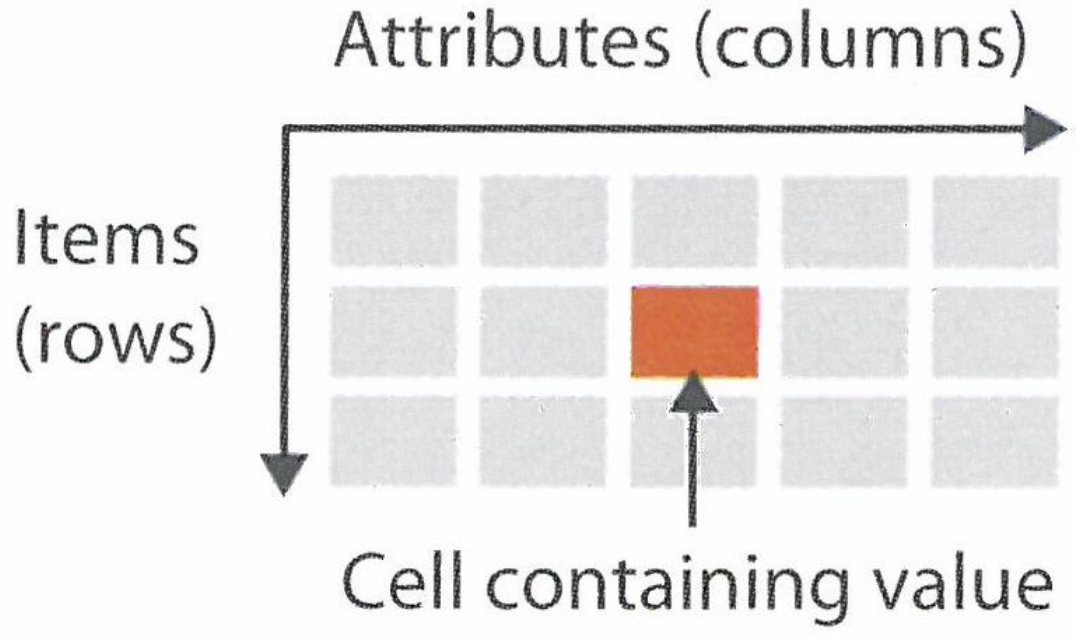




WHAT COULD BE THE KEY HERE?



WHAT DATA TYPE IS SUITABLE FOR A KEY?



KEYS VS. VALUES

key attributes are also sometimes called:

- independent attribute
- dimension

value attributes are also sometimes called:

- dependent attribute
- measure

LEVELS

= unique values for a categorical or ordered attribute

Abc Vispubdata-Grobid-min-c... Conference	# Vispubdata-... Year	Abc Vispubdata-Grobid-min-clean.... Paper.Title
InfoVis	2015	A comparative study ...
InfoVis	2015	A Linguistic Approach...
InfoVis	2015	A Psychophysical Inv...
InfoVis	2015	A Simple Approach fo...
InfoVis	2015	Acquired Codes of Me...
InfoVis	2015	AggreSet: Rich and Sc...
InfoVis	2015	AmbiguityVis: Visuali...
InfoVis	2015	Automatic Selection ...
InfoVis	2015	Beyond Memorability...
InfoVis	2015	Beyond Weber's Law:...
InfoVis	2015	Evaluation of Parallel...
InfoVis	2015	Guidelines for Effecti...
InfoVis	2015	High-Quality Ultra-Co...
InfoVis	2015	HOLA: Human-like Ort...
InfoVis	2015	How do People Make ...

CONFERENCE:
InfoVis, Vis, SciVis, VAST

YEAR:
1990 – 2015

PAPER.TITLE:
>2500 different

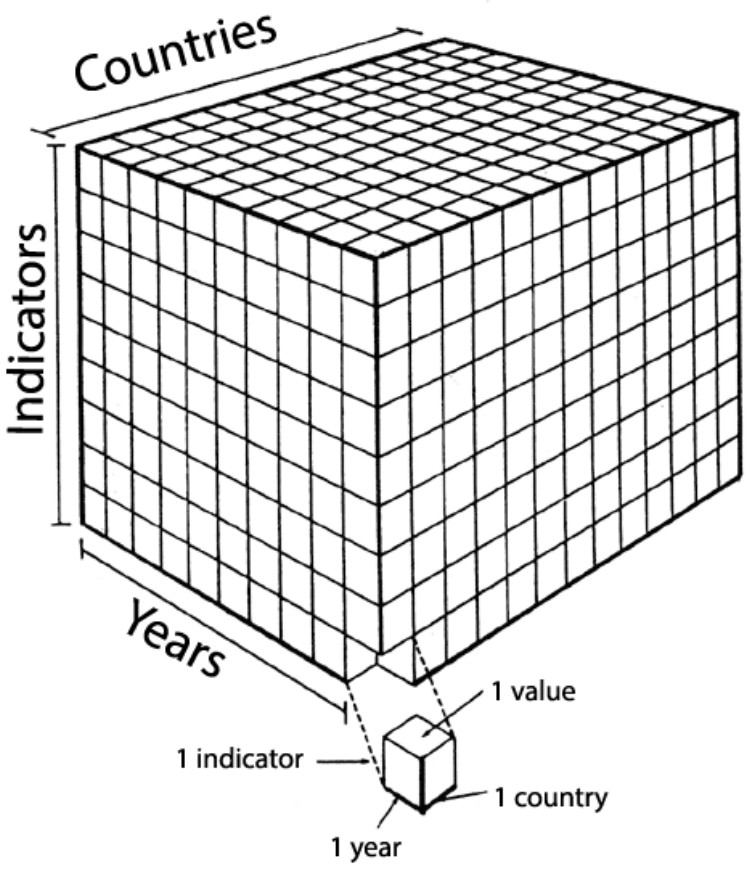
VISPUBDATA

ATTRIBUTES

ITEMS

	Year	Paper.Title	Paper.DOI	Link	First.page	Last.page	Paper.type..C.conf...	Abstract	Author.Names	First.Author.Affilia...	Deduped.author.n...	References	Author.Keywords	OCR.Authors
	2015	A comparative study ...	10.1109/TVCG.2015....	http://dx.doi.org/10....	619	628	J	RadViz and star coord...	Rubio-Sanchez, M.;Ra...	;;;	Rubio-Sanchez, M.;Ra...	10.1109/VAST.2010....	RadViz, Star coordina...	Rubio-S ´ Anchez,Ma...
	2015	A Linguistic Approach...	10.1109/TVCG.2015....	http://dx.doi.org/10....	698	707	J	When data categorie...	Setlur, V.;Stone, M.C.	;	Setlur, V.;Stone, M.C.	*null*	linguistics, natural la...	Setlur,Vidya;Stone,M...
	2015	A Psychophysical Inv...	10.1109/TVCG.2015....	http://dx.doi.org/10....	479	488	J	Physical visualization...	Jansen, Y.;Hornbaek, K.	Univ. of Copenhagen, ...	Jansen, Y.;Hornbaek, K.	10.1109/TVCG.2012....	Data physicalization, ...	Jansen,Yvonne;Hornb...
	2015	A Simple Approach fo...	10.1109/TVCG.2015....	http://dx.doi.org/10....	678	687	J	General methods for ...	Simonetto, P.;Archam...	;;	Simonetto, P.;Archam...	10.1109/TVCG.2011....	Euler diagrams, Boun...	Simonetto,Paolo;Arc...
	2015	Acquired Codes of Me...	10.1109/TVCG.2015....	http://dx.doi.org/10....	509	518	J	While information vis...	Byrne, L.;Angus, D.;W...	;;	Byrne, L.;Angus, D.;W...	10.1109/TVCG.2013....	Visual Design, Taxono...	Byrne,Lydia;Angus,D...
	2015	AggreSet: Rich and Sc...	10.1109/TVCG.2015....	http://dx.doi.org/10....	688	697	J	Datasets commonly i...	Yalcin, M.A.;Elmqvist,...	Univ. of Maryland, Co...	Yalcin, M.A.;Elmqvist,...	10.1109/TVCG.2011....	Multi-valued attribut...	Adil Yalçın,M;Beders...
	2015	AmbiguityVis: Visuali...	10.1109/TVCG.2015....	http://dx.doi.org/10....	359	368	J	Node-link diagrams p...	Yong Wang;Qiaomu S...	;;;;;;	Yong Wang;Qiaomu S...	10.1109/TVCG.2006....	Visual Ambiguity, Vis...	Wang,Yong;Shen,Qia...
InfoVis	2015	Automatic Selection ...	10.1109/TVCG.2015....	http://dx.doi.org/10....	669	677	J	Effective small multi...	Anand, A.;Talbot, J.	;	Anand, A.;Talbot, J.	10.1109/VAST.2010....	Small multiple displa...	Anand,Anushka;Talbo...
InfoVis	2015	Beyond Memorability:...	10.1109/TVCG.2015....	http://dx.doi.org/10....	519	528	J	In this paper we mov...	Borkin, M.A.;Bylinskii...	;;;;;;;;	Borkin, M.;Bylinskii, Z....	10.1109/TVCG.2012....	Information visualiza...	*null*
InfoVis	2015	Beyond Weber's Law:...	10.1109/TVCG.2015....	http://dx.doi.org/10....	469	478	J	Models of human per...	Kay, M.;Heer, J.	;	Kay, M.;Heer, J.	10.1109/TVCG.2014....	Weber's law, percept...	Kay,Matthew;Heer,Je...
InfoVis	2015	Evaluation of Parallel...	10.1109/TVCG.2015....	http://dx.doi.org/10....	579	588	J	The parallel coordina...	Johansson, J.;Forsell,...	Norrkoping Visualiza...	Johansson, J.;Forsell,...	10.1109/TVCG.2014....	Survey, evaluation, g...	Johansson,Jimmy;For...
InfoVis	2015	Guidelines for Effecti...	10.1109/TVCG.2015....	http://dx.doi.org/10....	489	498	J	Semi-automatic text ...	Strobelt, H.;Oelke, D.;...	;;;;	Strobelt, H.;Oelke, D.;...	10.1109/TVCG.2012....	Text highlighting tec...	Strobelt,Hendrik;Oel...
InfoVis	2015	High-Quality Ultra-Co...	10.1109/TVCG.2015....	http://dx.doi.org/10....	339	348	J	Prior research into ne...	Yoghourdjian, V.;Dwy...	;;;;;	Yoghourdjian, V.;Dwy...	10.1109/TVCG.2008....	Network visualizatio...	Yoghourdjian,Vahan;...
InfoVis	2015	HOLA: Human-like Ort...	10.1109/TVCG.2015....	http://dx.doi.org/10....	349	358	J	Over the last 50 year...	Kieffer, S.;Dwyer, T.;...	;;;	Kieffer, S.;Dwyer, T.;...	10.1109/TVCG.2006....	Graph layout, orthog...	Kieffer,Steve;Dwyer,...
InfoVis	2015	How do People Make ...	10.1109/TVCG.2015....	http://dx.doi.org/10....	499	508	J	In this paper, we wou...	Sukwon Lee;Sung-He...	Sch. of Ind. Eng., Purd...	Sukwon Lee;Sung-He...	10.1109/TVCG.2013....	Sensemaking model, i...	Lee,Sukwon;Kim,Sun...
InfoVis	2015	Improving Bayesian R...	10.1109/TVCG.2015....	http://dx.doi.org/10....	529	538	J	Decades of research ...	Ottley, A.;Peck, E.M.;...	;;;;;;;;	Ottley, A.;Peck, E.M.;...	10.1109/TVCG.2014....	Bayesian Reasoning, ...	Ottley,Alvitta;Peck,E...
InfoVis	2015	Matches, Mismatche...	10.1109/TVCG.2015....	http://dx.doi.org/10....	449	458	J	The energy performa...	Brehmer, M.;Ng, J.;Ta...	;;;	Brehmer, M.;Ng, J.;Ta...	10.1109/TVCG.2011....	Design study, design ...	Brehmer,Matthew;N...

THE DATA CUBE



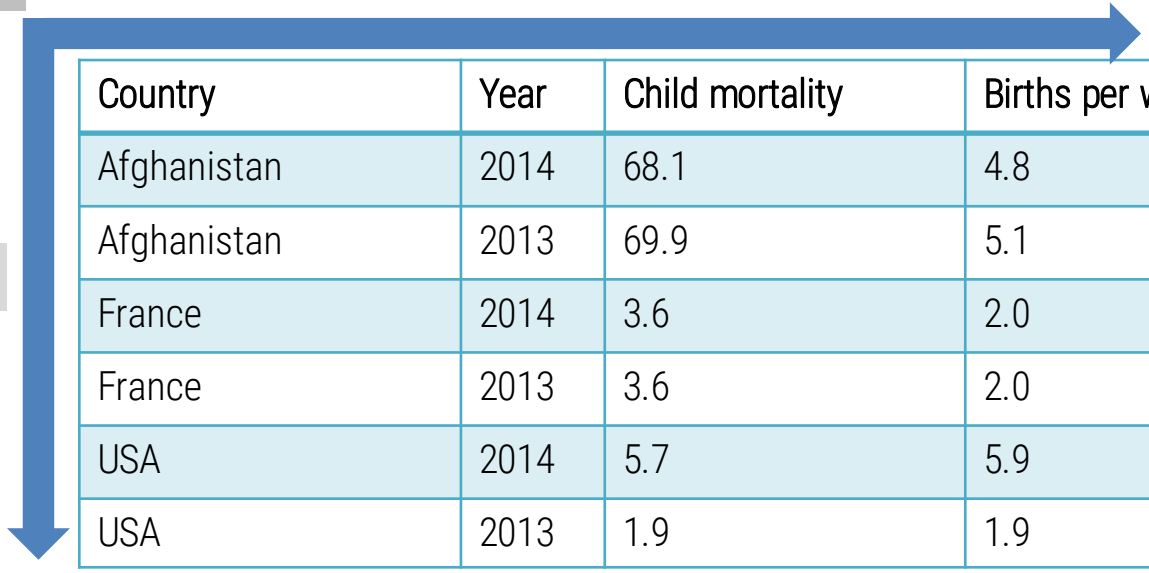
Country	Year	Child mortality	Births per woman
Afghanistan	2014	68.1	4.8
Afghanistan	2013	69.9	5.1
France	2014	3.6	2.0
France	2013	3.6	2.0
USA	2014	5.7	5.9
USA	2013	1.9	1.9

MULTI-ATTRIBUTE DATA – OUR VIEW TODAY

n x d matrix

n attributes

d items (data points)

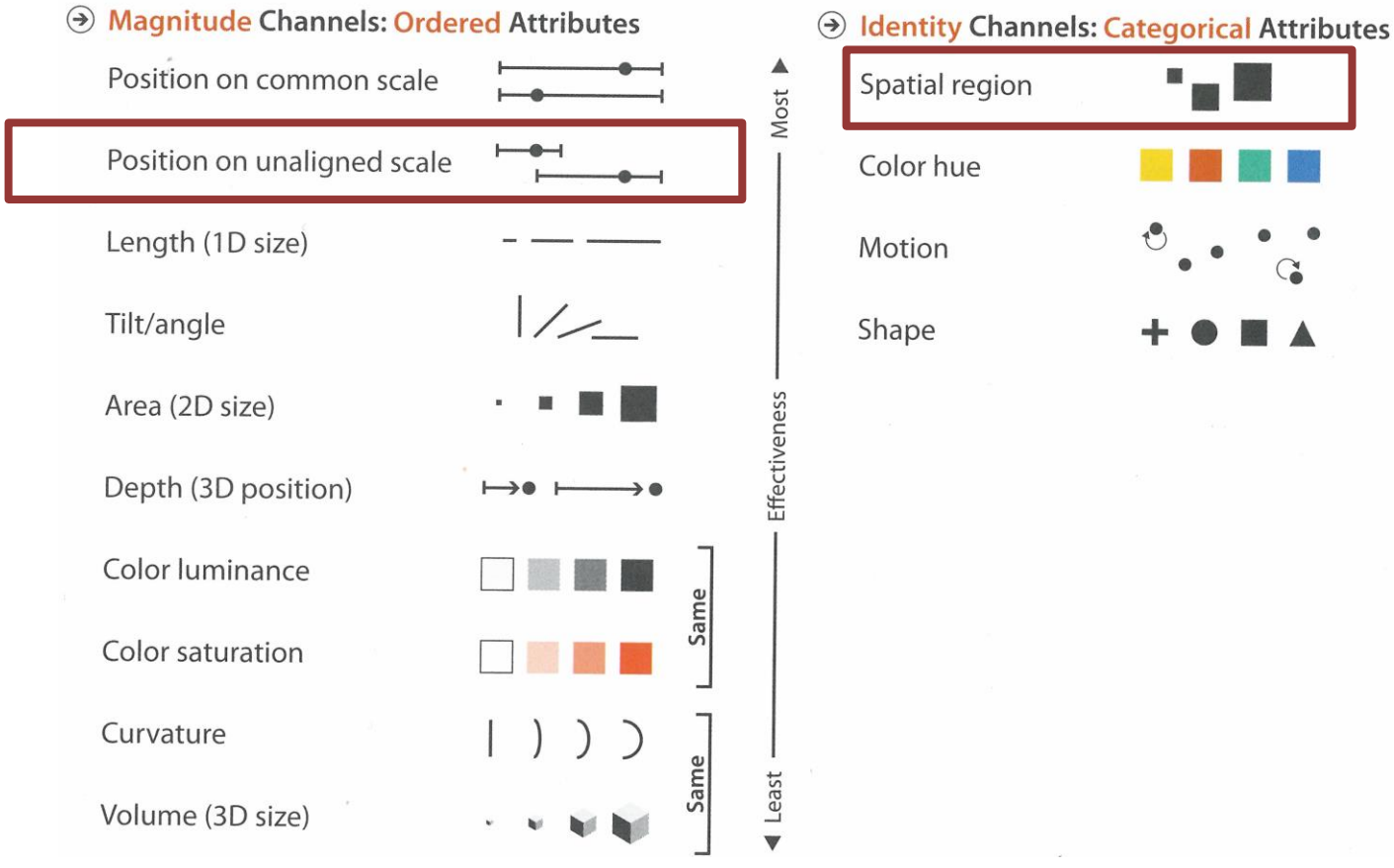


Country	Year	Child mortality	Births per woman
Afghanistan	2014	68.1	4.8
Afghanistan	2013	69.9	5.1
France	2014	3.6	2.0
France	2013	3.6	2.0
USA	2014	5.7	5.9
USA	2013	1.9	1.9

ARRANGING TABULAR DATA

In Space

WHY ARRANGING DATA



QUANTITATIVE VALUES

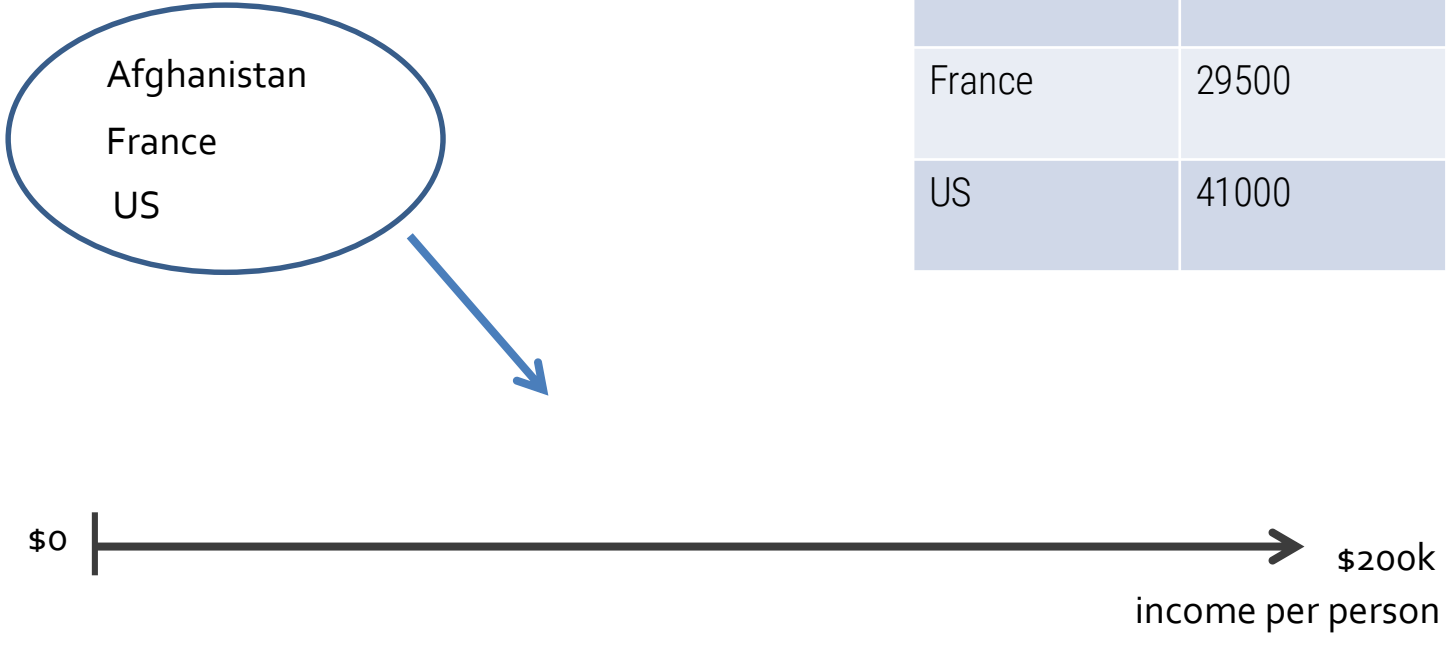
APPROACH

Let’s start with two attributes: country & income per person

Country	Income per person
Afghanistan	850
France	29500
US	41000

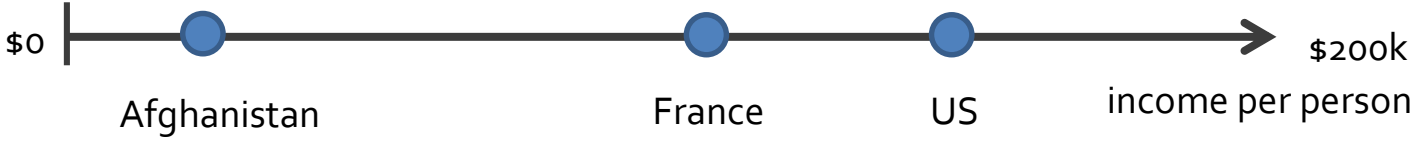
1. FIND A LAYOUT

Country	Income per person
Afghanistan	850
France	29500
US	41000



2. CHOOSE A VISUAL ENCODING & MARK

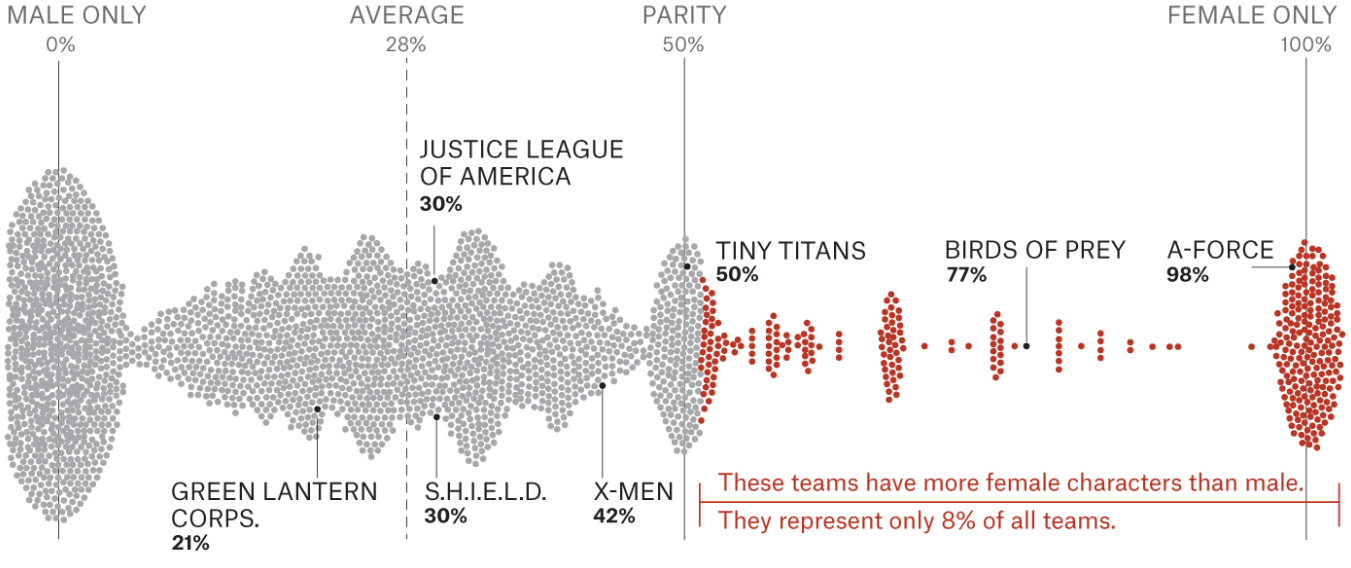
E.g. position + circle



ThePudding

Analyzing the Gender Representation of 34,476 Comic Book Characters

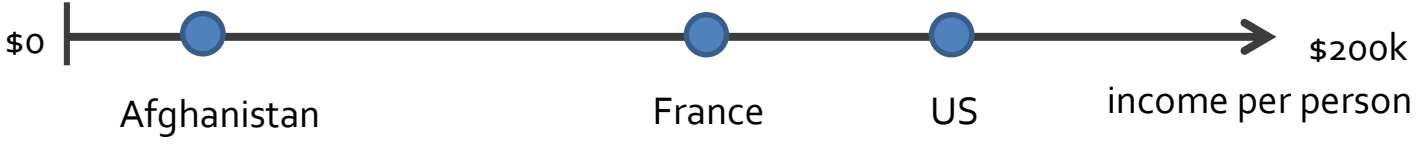
Female percentage of every team
Each dot represents one of 2,862 teams in DC and Marvel.



1. FIND A LAYOUT

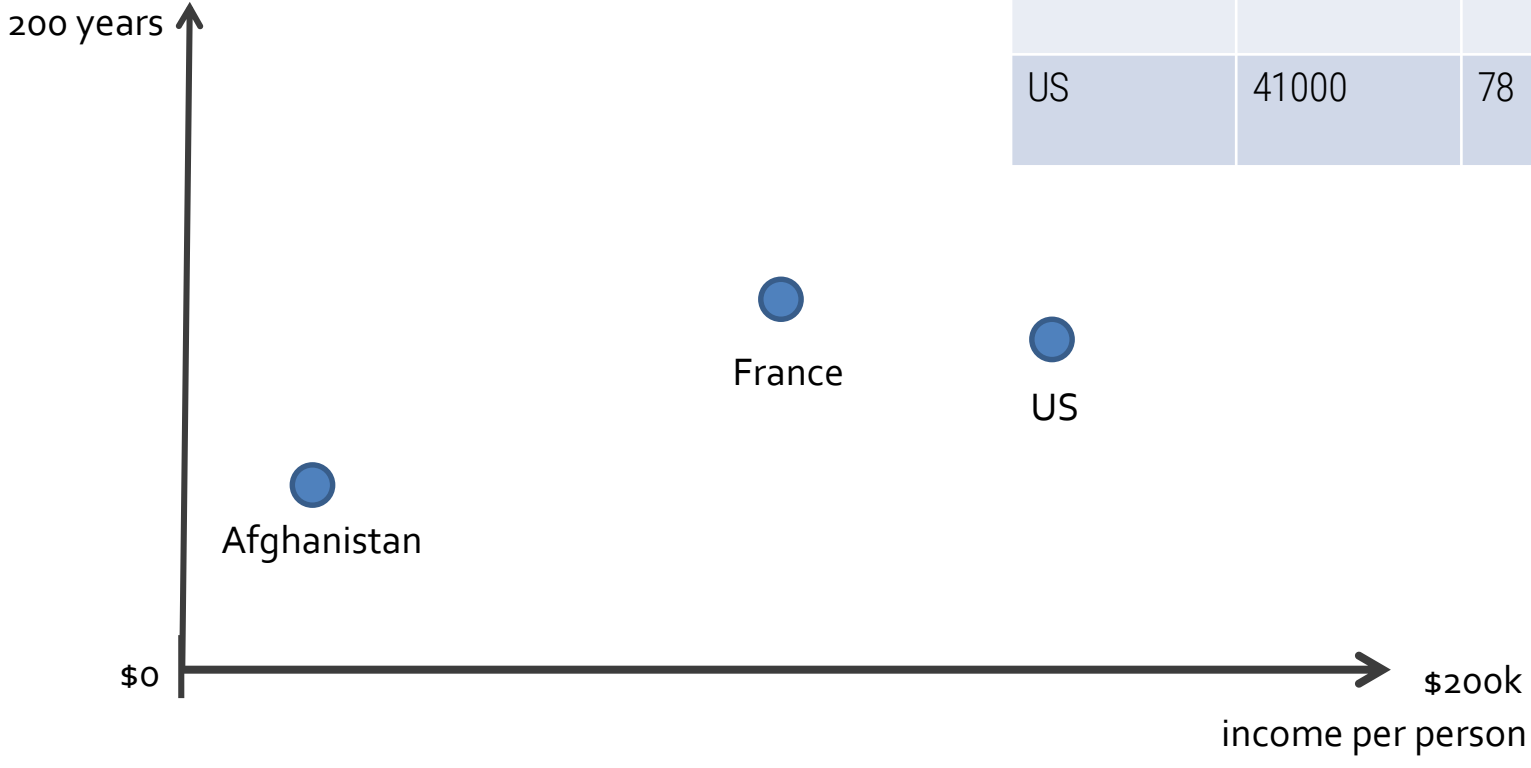
Country	Income per person	Life expectancy
Afghanistan	850	57
France	29500	81
US	41000	78

How do we extend this to 3 data attributes?



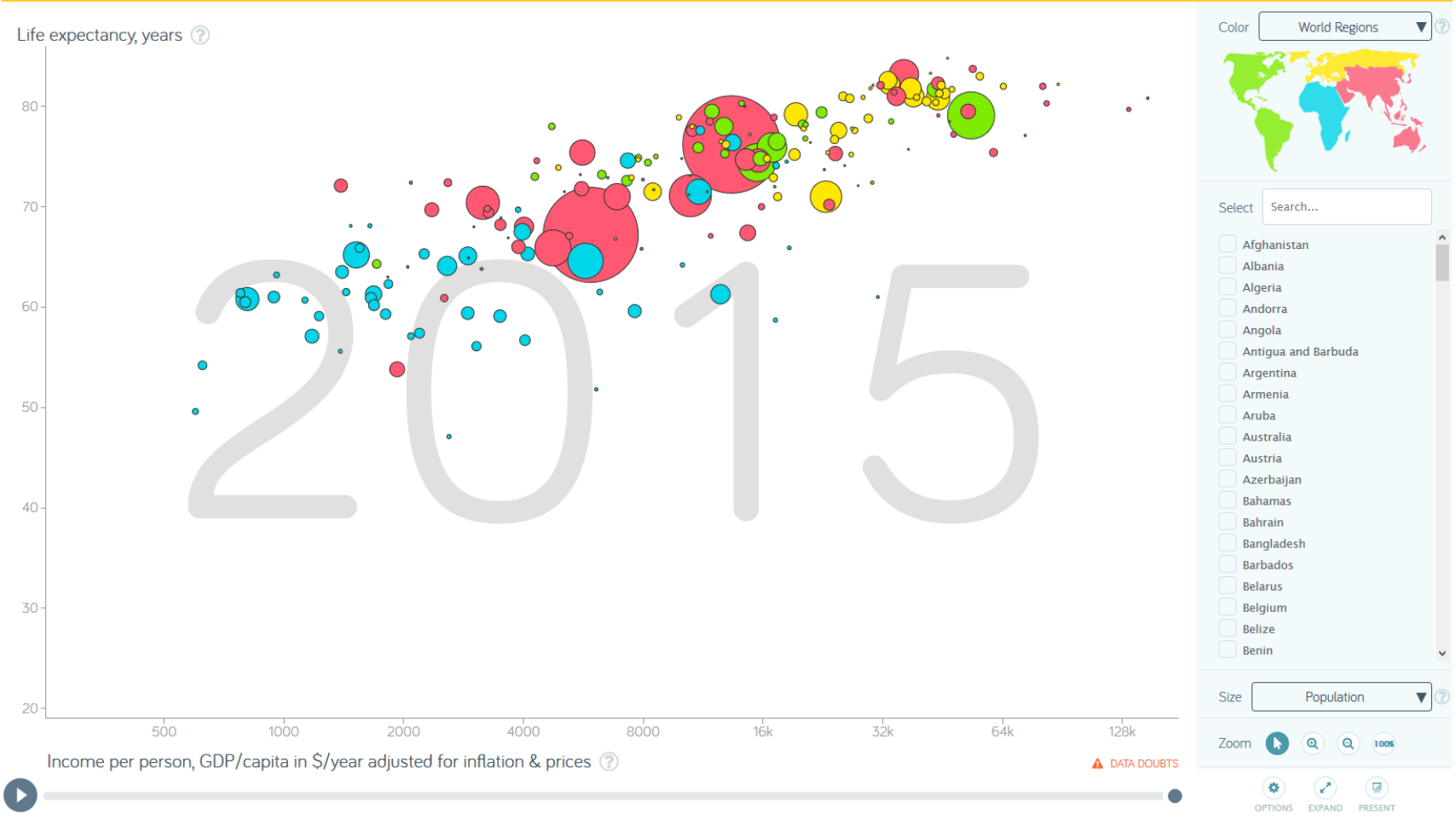
1. FIND A LAYOUT

Country	Income per person	Life expectancy
Afghanistan	850	57
France	29500	81
US	41000	78



SCATTERPLOTS

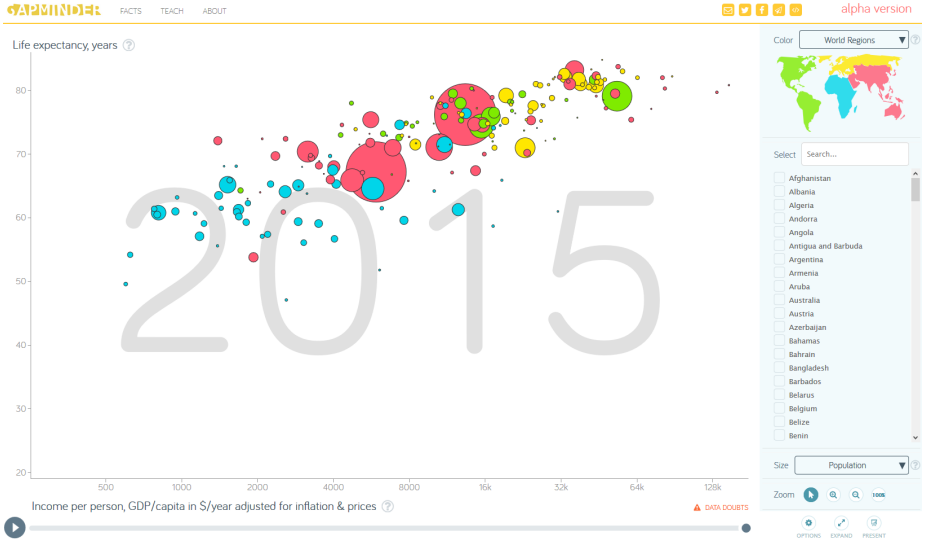
- two quantitative values
- horizontal and vertical spatial dimensions
- mark type = point



when marks are sized, the chart is often called a bubble chart or bubble plot

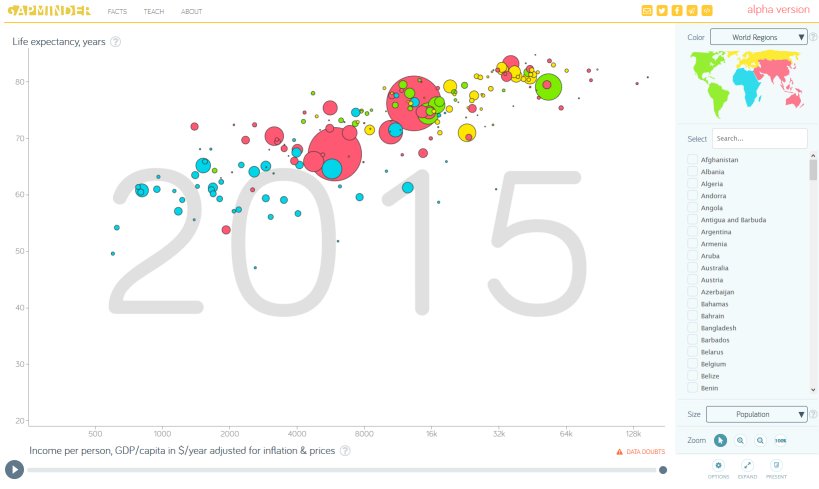
Why choose a scatterplot?

What can you show with them?

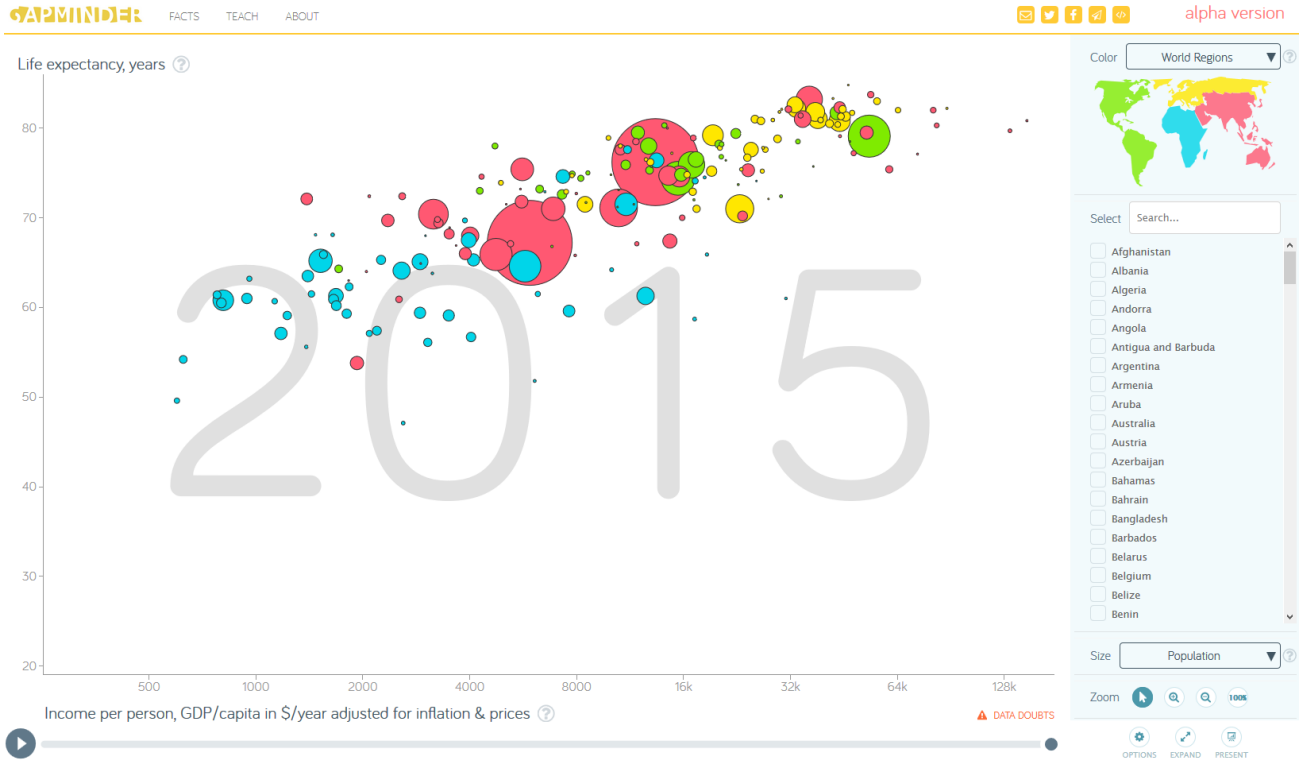


TASKS

- find trends
- find outliers
- show distribution
- show correlation
- locate clusters



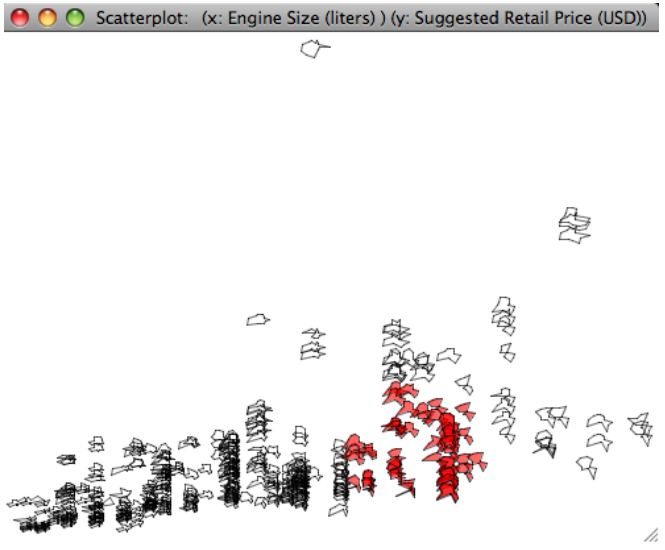
how many items are reasonable to put on a scatterplot?



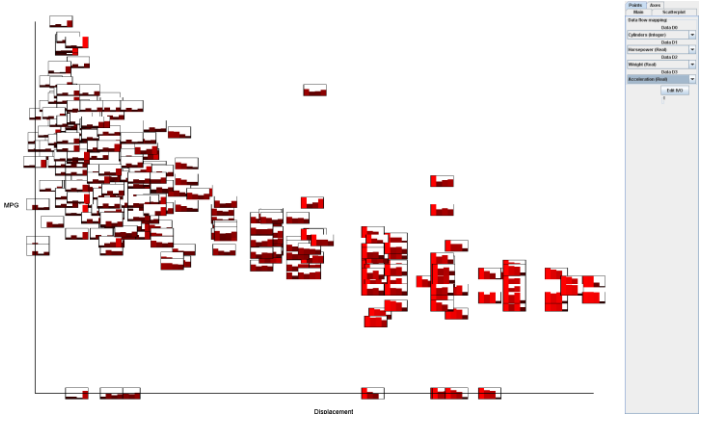
GLYPHS

marks can be replaced with glyphs

glyphs are themselves composed of multiple marks



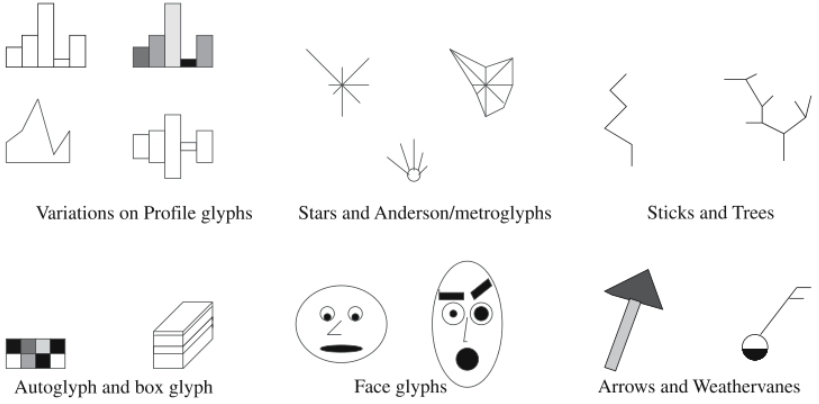
http://rosuda.org/software/Gauguin/gauguin.html



https://engineering.purdue.edu/~elm/projects/gpuvis.html

GLYPHS

- Small composite visual representations of multi-dimensional data points
- Characterized generally by lack of reference structures (grid lines, axes labels, …)



From Ward, 2002
A taxonomy of glyph placement strategies for multidimensional data visualization

EXAMPLE: CHERNOFF FACES

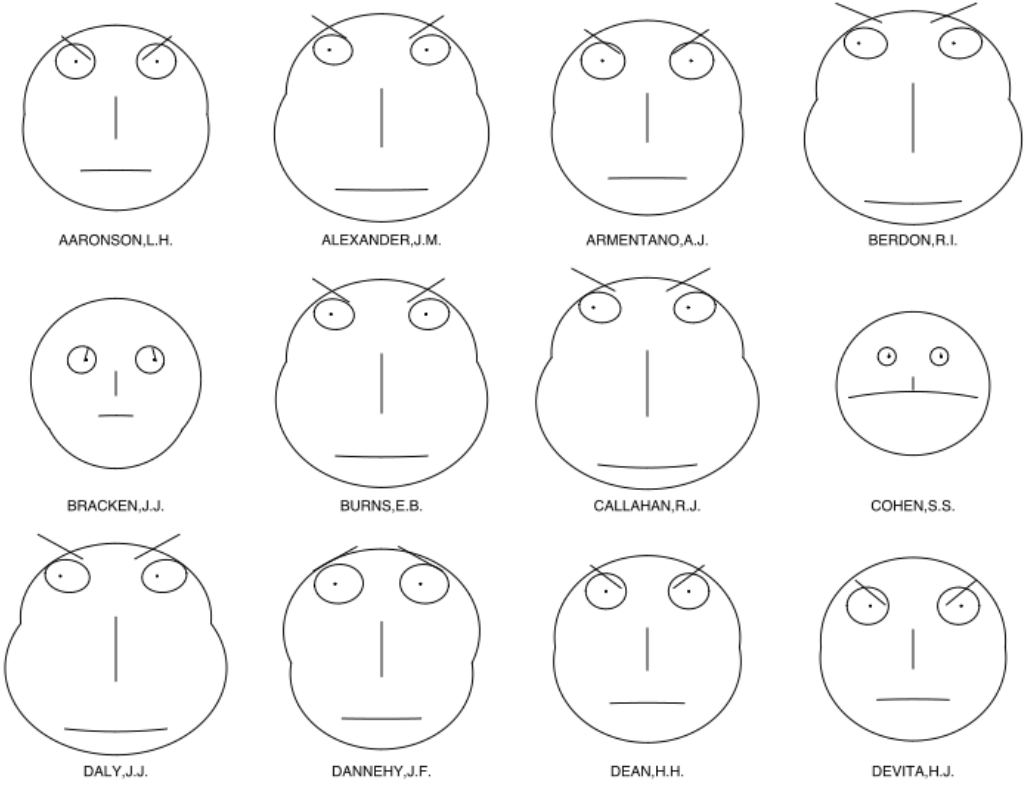
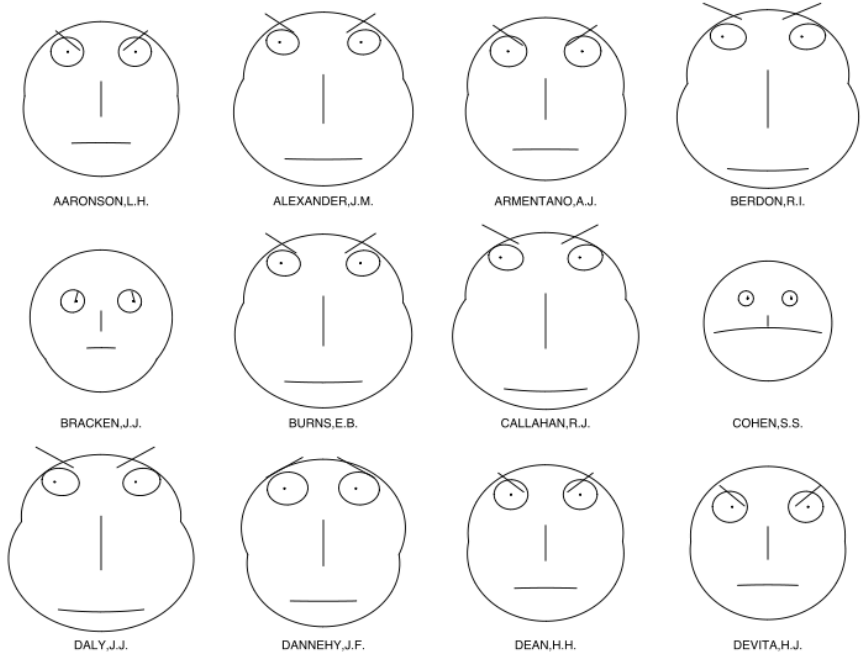


Image source: Wikipedia

Herman Chernoff, The Use of Faces to Represent Points in K-Dimensional Space Graphically, 1973.

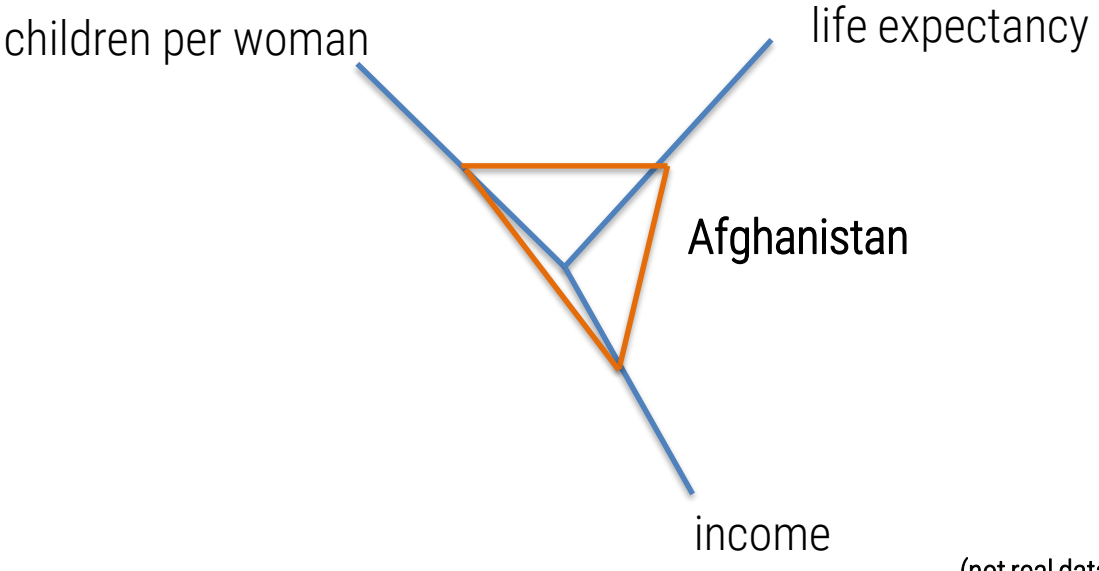
CHERNOFF FACES

- features of a human face encode data values (e.g. slant of eye brows, size of eyes, …)
- reasoning: humans are good at differentiating faces and reading face features
- problem: chernoff faces have generally been found not to be very effective



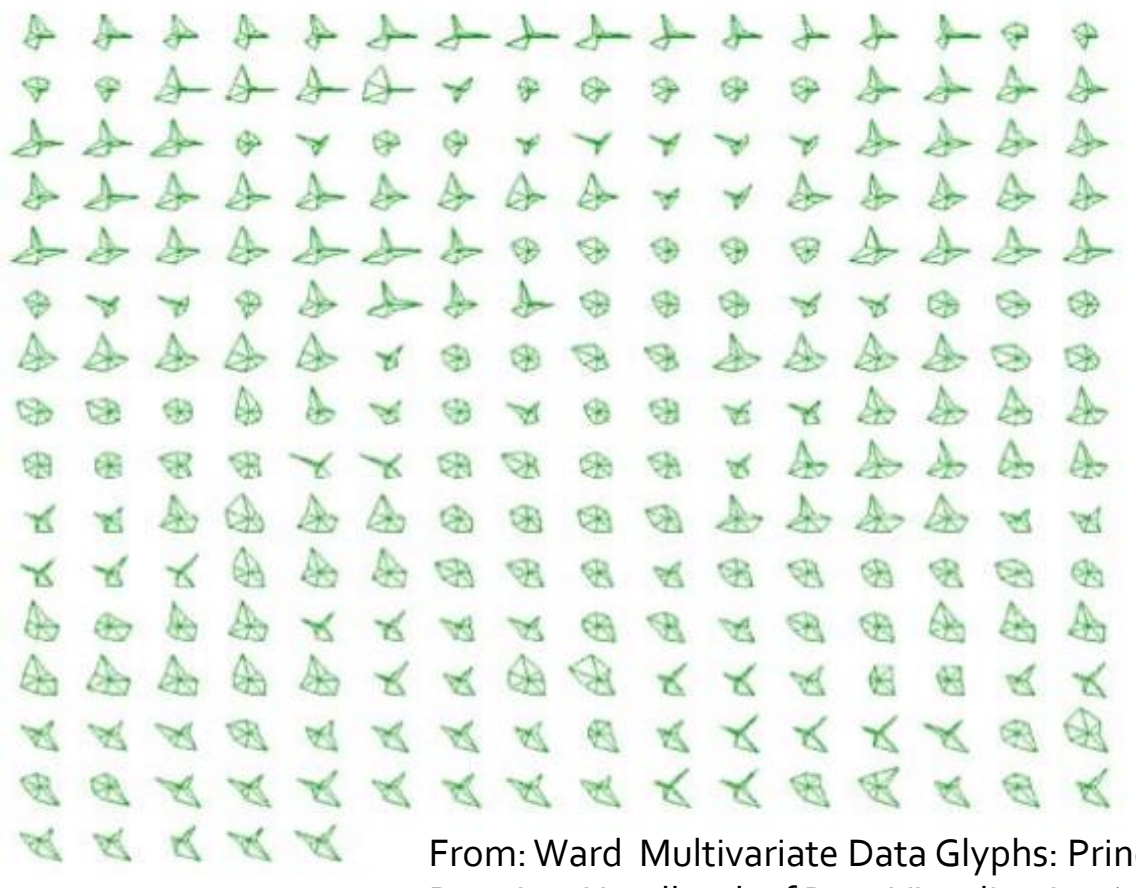
EXAMPLE: STAR GLYPHS

- Lay out dimension in radial fashion
- Draw each point as a ring



(not real data, here)

STAR GLYPHS



From: Ward Multivariate Data Glyphs: Principles and Practice. Handbook of Data Visualization (2008)

It's gettin hot out here

2015: WARMEST DECEMBER

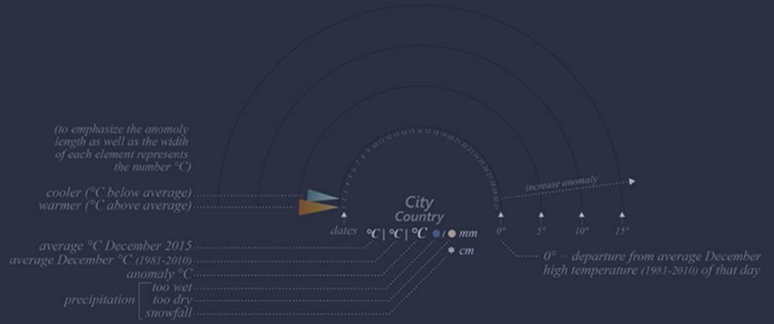
http://www.studioterp.nl/its-gettin-h

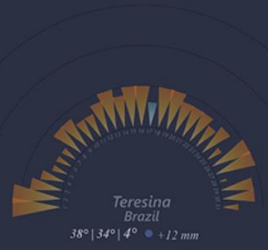
Across the globe, record warm temperatures were observed over every continent, including a large swath of eastern North America, southern Mexico through northern South America, western and central Europe, most of southern Africa, parts of central and southeastern Asia, and a large section of southeastern Australia.
The link between the tumultuous weather events experienced around the world in December are likely to be down to the natural phenomenon known as El Niño making the effects of man-made climate change worse. The 2015 El Niño is one of the strongest on record, leading to record temperatures, rainfall and weather extremes.

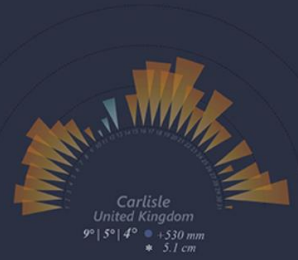
During December, the globally-averaged land surface temperature was almost 2°C above the 20th century average. This was the highest for December in the 1880–2015 record, surpassing the previous record of 2006 by 0.5°C. The December temperature departure from average was also the highest departure among all months in the historical record and the first time a monthly departure has reached little over +1°C from the 20th century average.

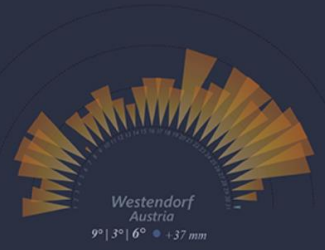
HOW TO READ IT

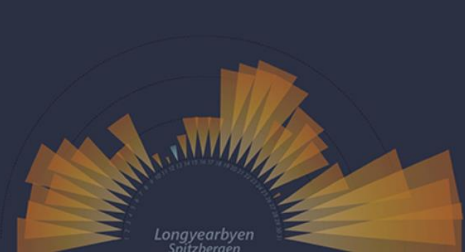
This visualization shows 8 places around the globe chosen for their location in areas where anomalies occured. Shown are the number of °C departing from the average temperature of each December day.











SHOW CATEGORICAL DATA

Using Regions: Separate, Order, and Align

CATEGORICAL VALUES

- spatial position is an ordered magnitude visual channel
- categorical attributes are unordered identities (no magnitude)

→cannot be encoded with spatial position

- BUT: can be differentiated with a spatial region

REGIONS

- contiguous bounded areas
- distinct from one another
- need to be separated, ordered, and aligned

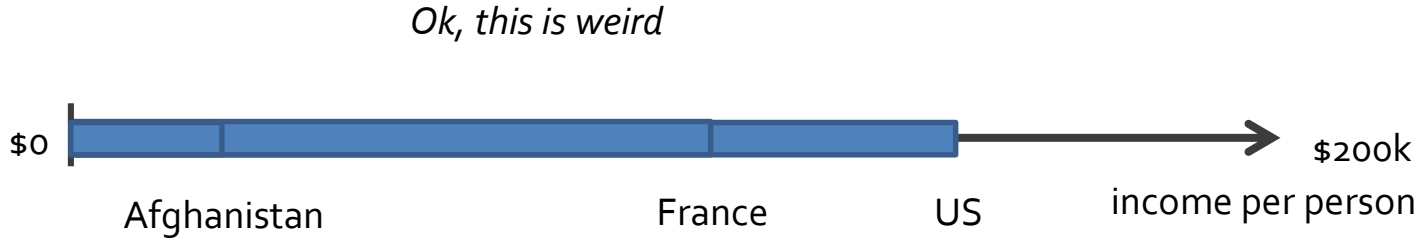
LIST ALIGNMENT

ONE KEY

LIST ALIGNMENT

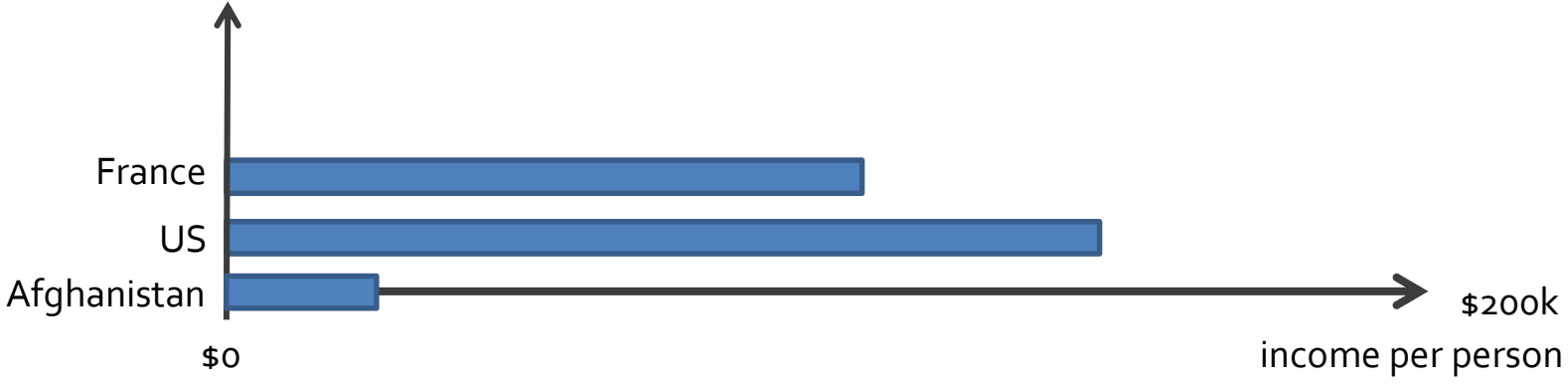
separate into regions by key

E.g. length + rectangle



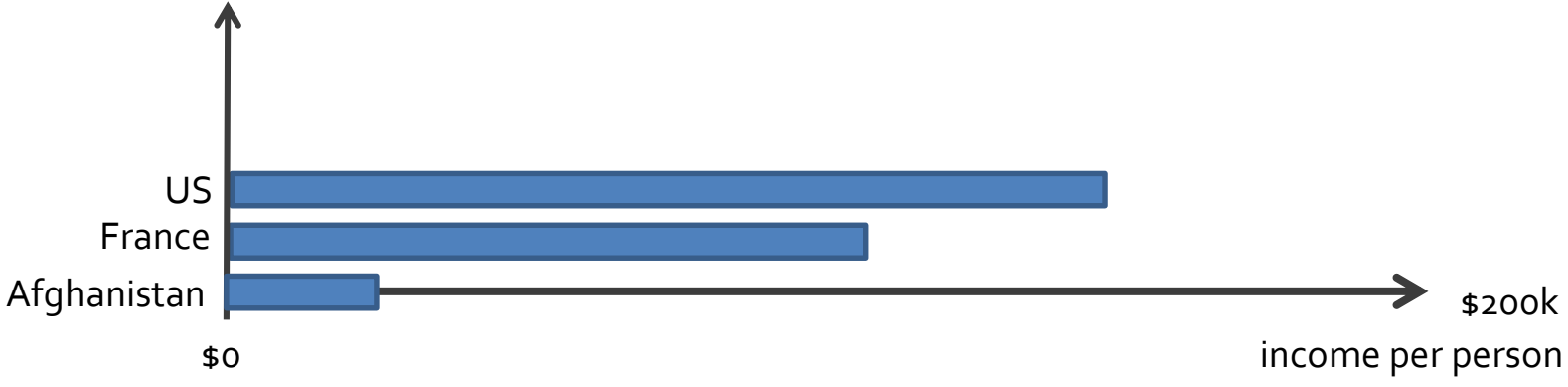
ALIGN

align regions of key categorical values along one axis in a common frame



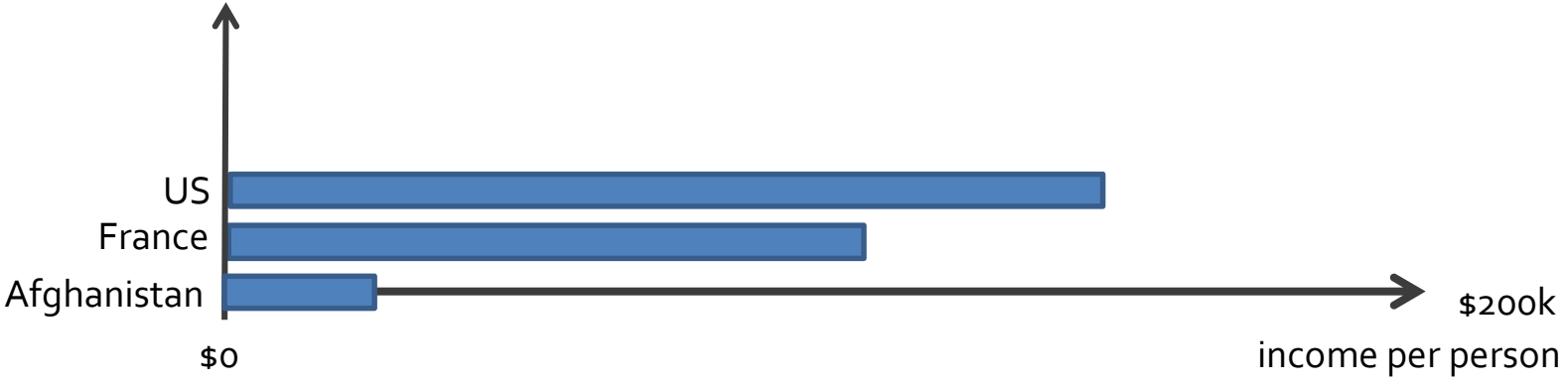
ORDER

- using a derived attribute such as alphabet
- and/or using dependent data values



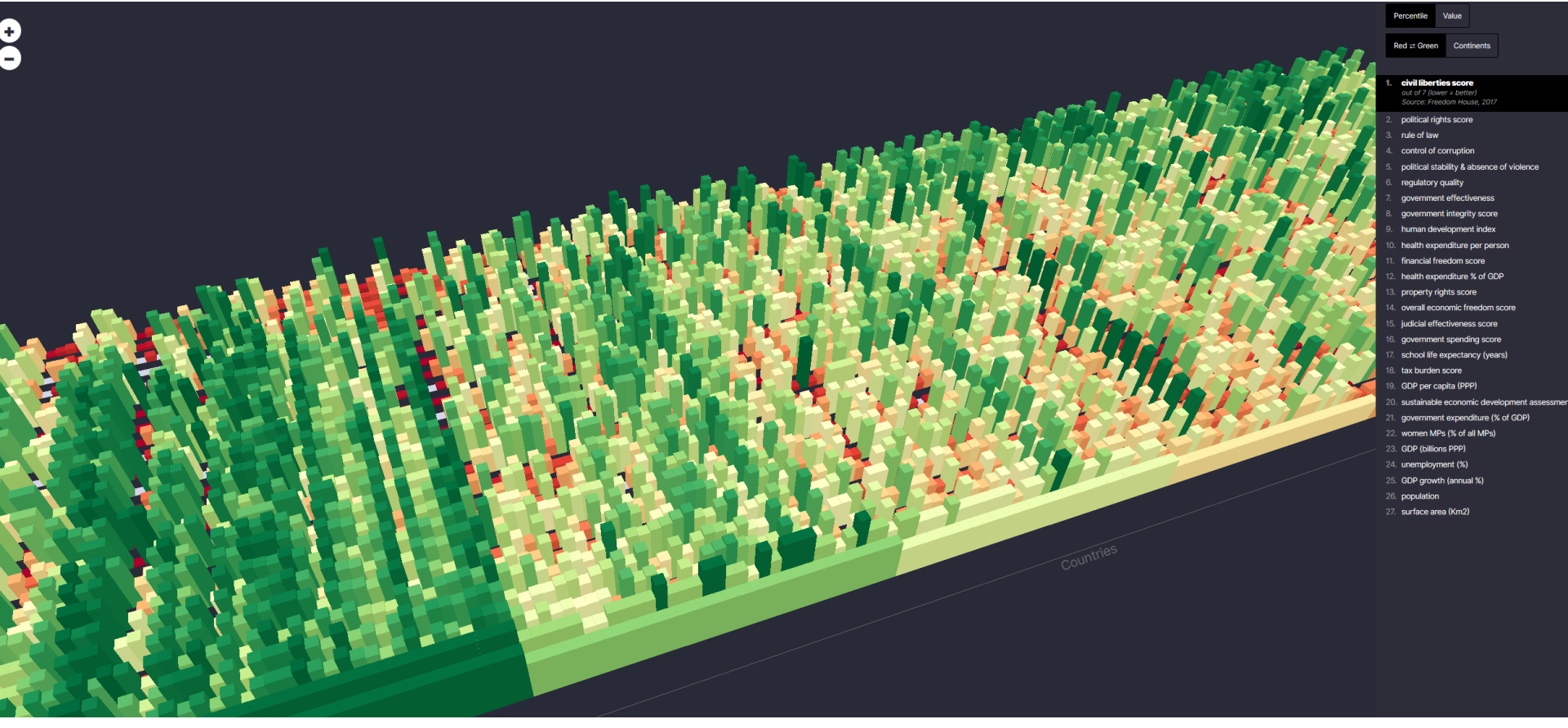
BAR CHARTS

DATA	one quantitative value attribute, one categorical key attribute
ENCODE	line marks, express value attribute with aligned vertical position (length), separate key attribute with horizontal position
TASK	lookup and compare values
SCALE	key attribute: dozens to hundreds of levels

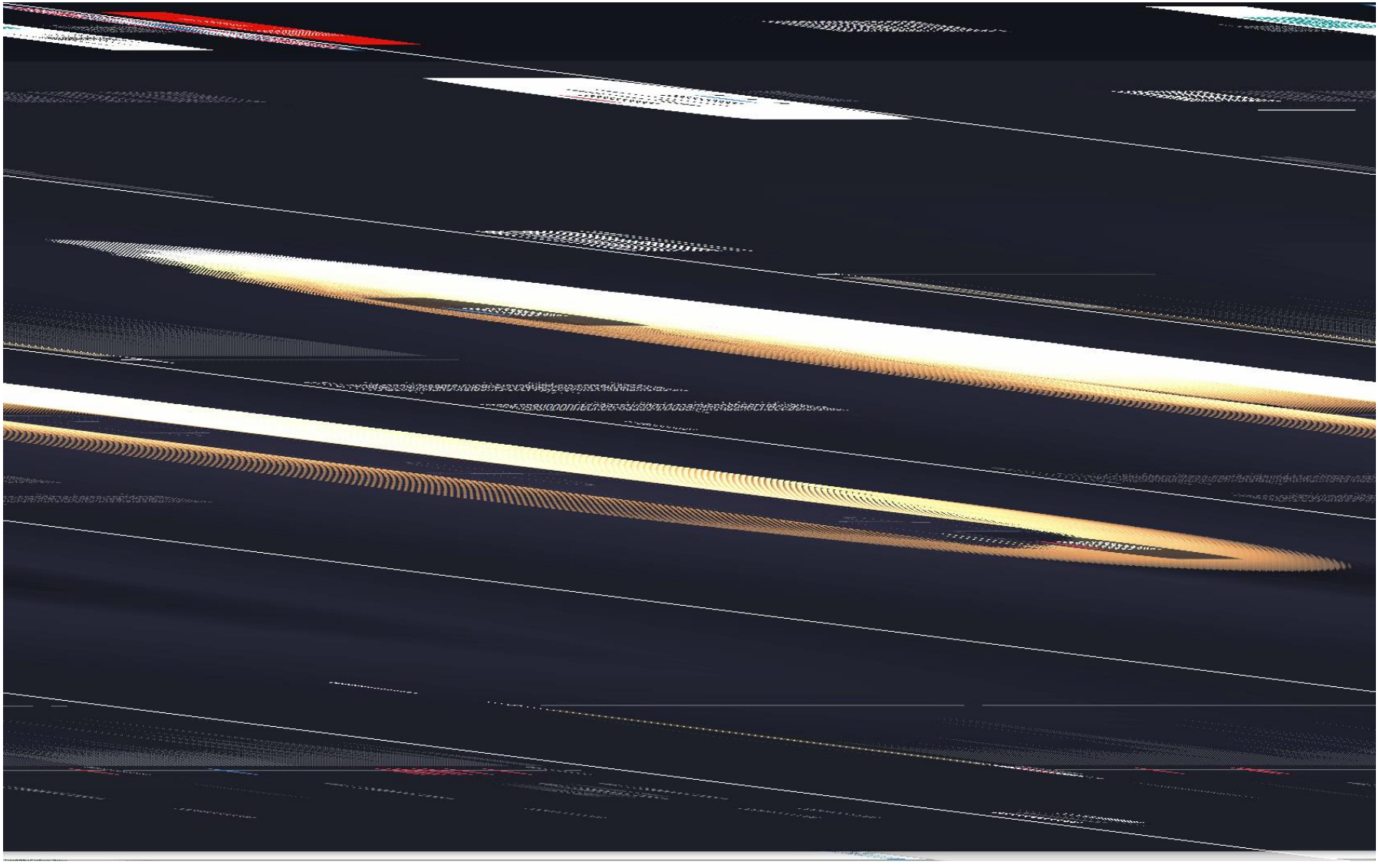


https://pudding.cool/2017/03/incarceration/index.html

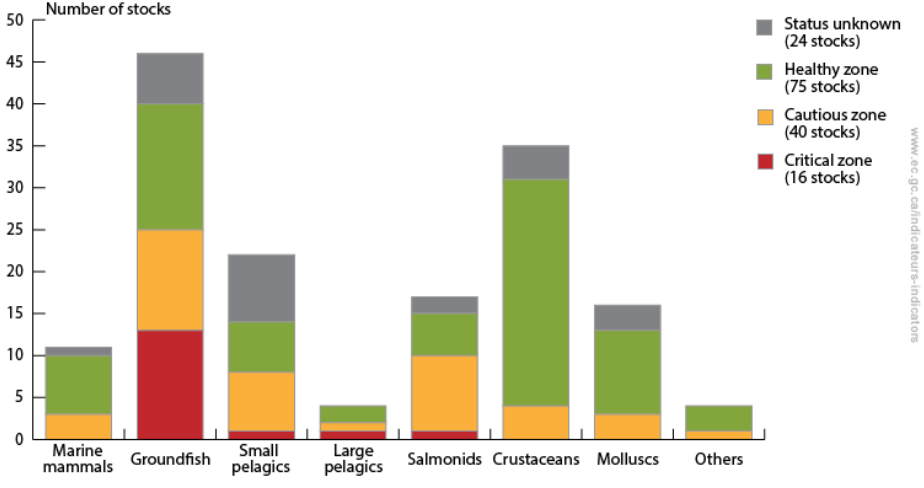
Be careful with 3D bar charts. Only use them when you know what you are doing.



https://wattenberger.com/wdvp



ALTERNATIVE

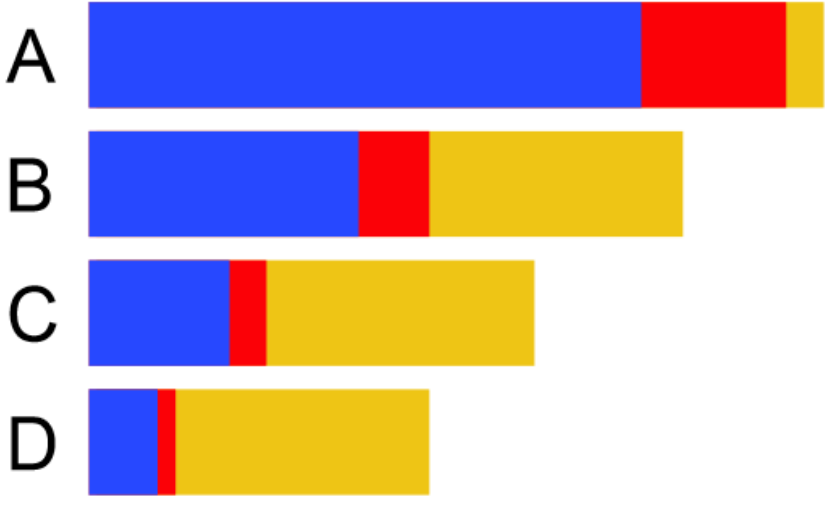


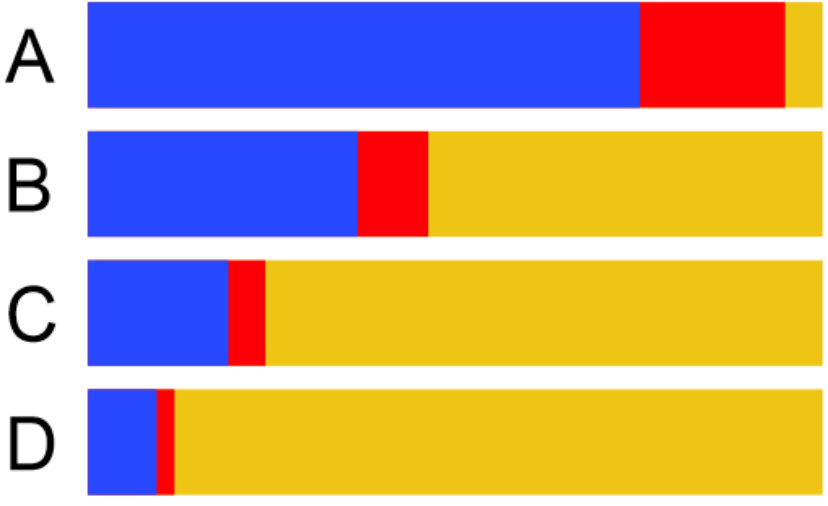
https://www.ec.gc.ca/indicateurs-indicators/default.asp?lang=en&n=1BCD421B-1

Stacked bar chart

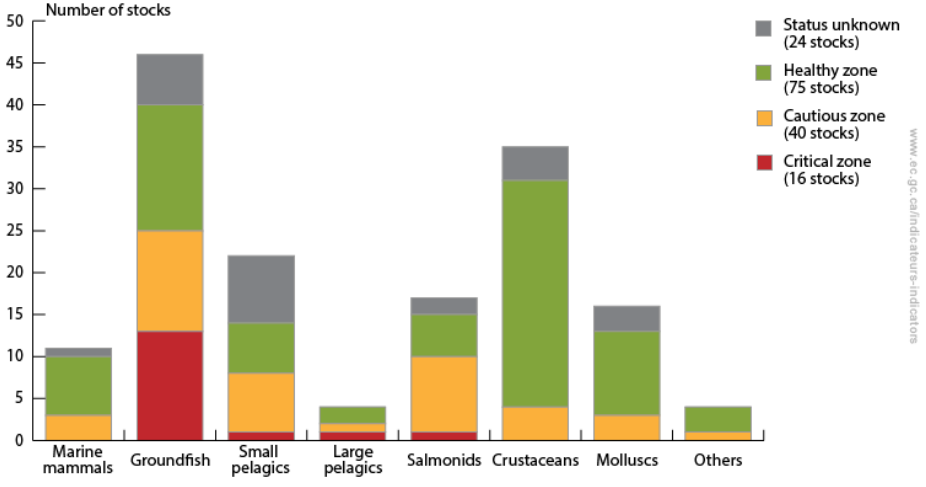
- each bar is a composite glyph
- each bar part encodes a value
- composite glyphs arranged as a list according to primary key
- color used to distinguish secondary key
- typically used for absolute values (use a normalized stacked bar for proportions)

STACKED BARS VS. NORMALIZED STACKED BARS





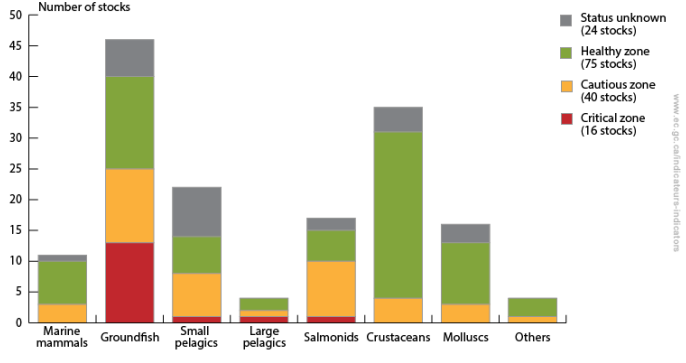
STACKED BARS



- ADVANTAGE
  - can compare totals and lowest level well
- DISADVANTAGE
  - upper levels of secondary key require comparison against non-aligned scale

STACKED BARS

DATA	MD table; one quantitative value attribute, two categorical key attributes
ENCODE	bar glyph: length-encoded subcomponents for each level of secondary key attribute separate bars by category of primary key
TASK	part-to-whole relationship, lookup values, find trends
SCALE	key attribute (main axis): dozens to hundreds of levels key attribute (stacked glyph axis): several to one dozen

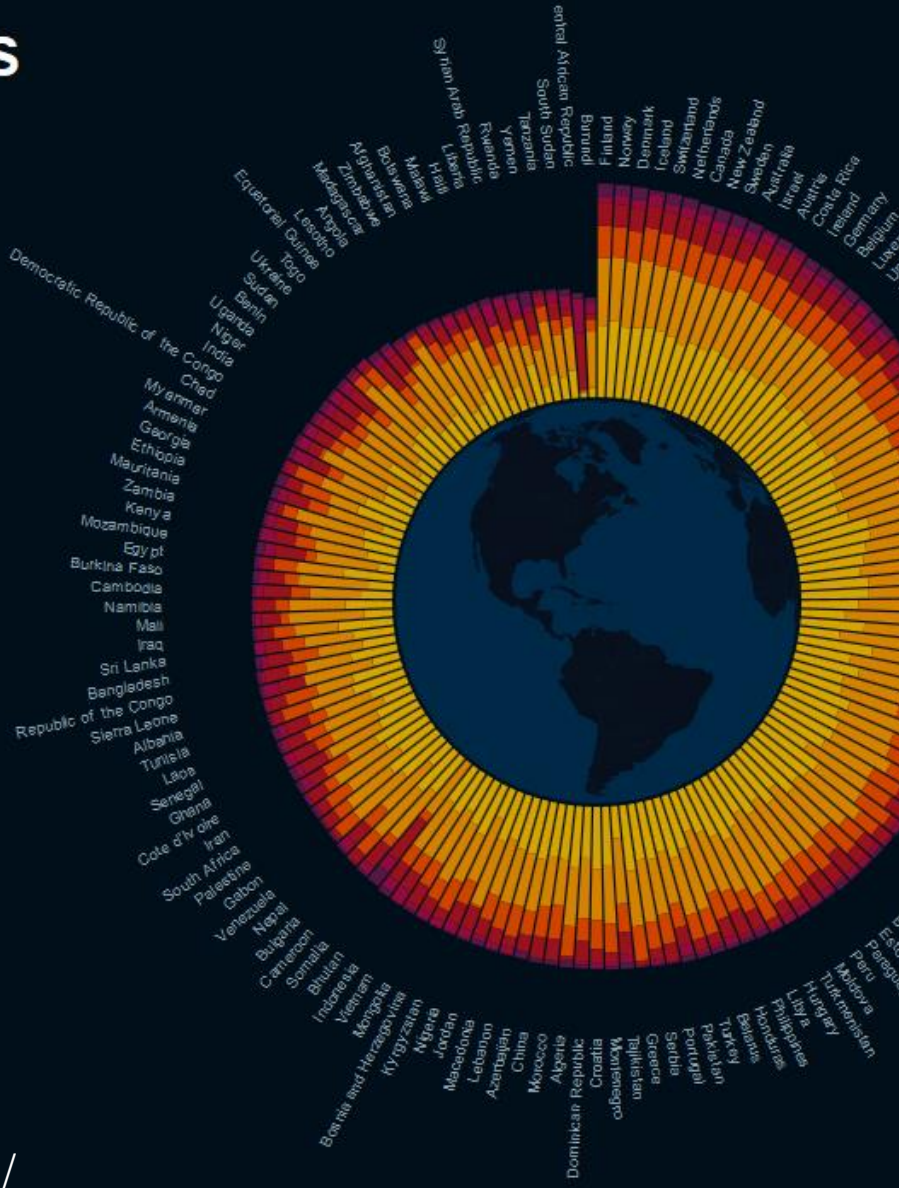


A WORLD OF HAPPINESS

According to the UN World Happiness Report, these factors combined contribute to national happiness.

Explore the globe and see how your country measures up.

- GDP per capita
- Social support
- Healthy life expectancy
- Freedom to make life choices
- Generosity
- Perceptions of corruption



http://www.benscott.co.uk/wdvp/

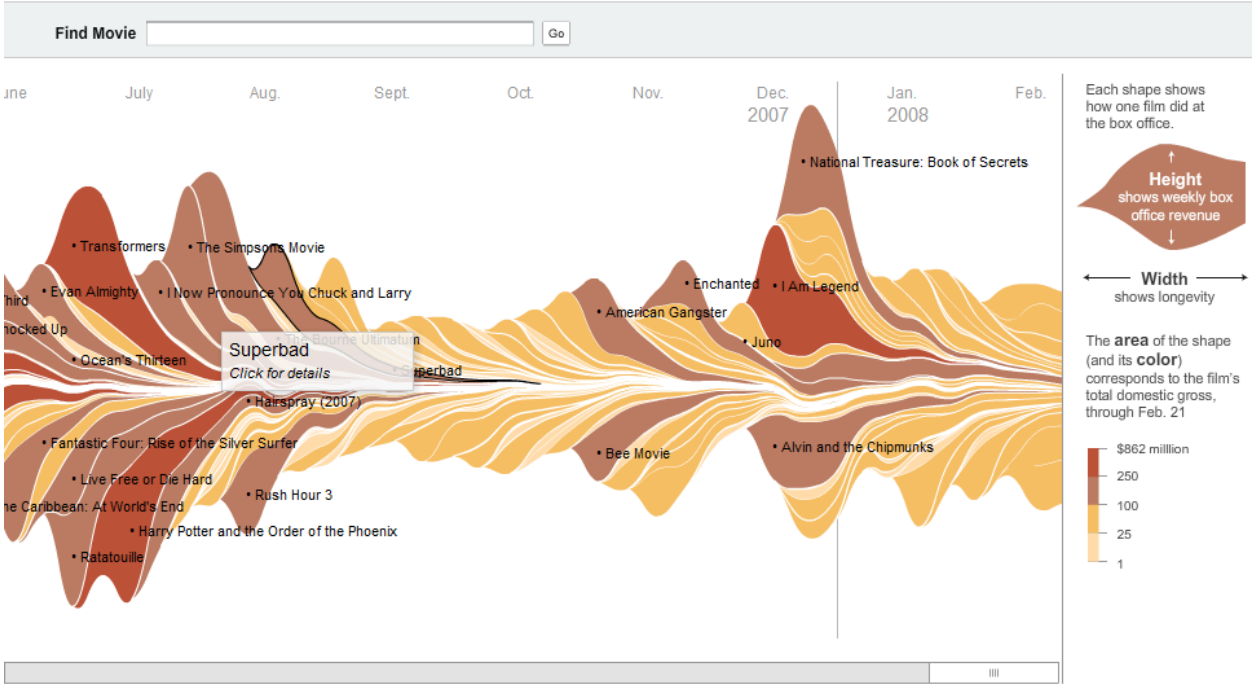
STREAMGRAPH

February 23, 2008

SIGN IN TO E-MAIL OR SAVE THIS | FEEDBACK

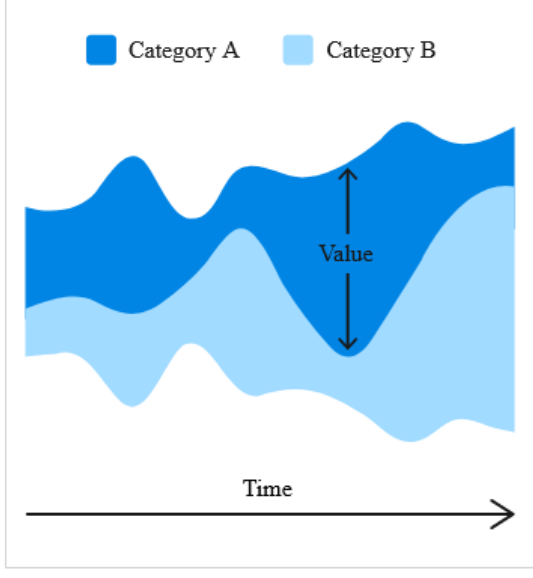
The Ebb and Flow of Movies: Box Office Receipts 1986 — 2008

Summer blockbusters and holiday hits make up the bulk of box office revenue each year, while contenders for the Oscars tend to attract smaller audiences that build over time. Here's a look at how movies have fared at the box office, after adjusting for inflation.



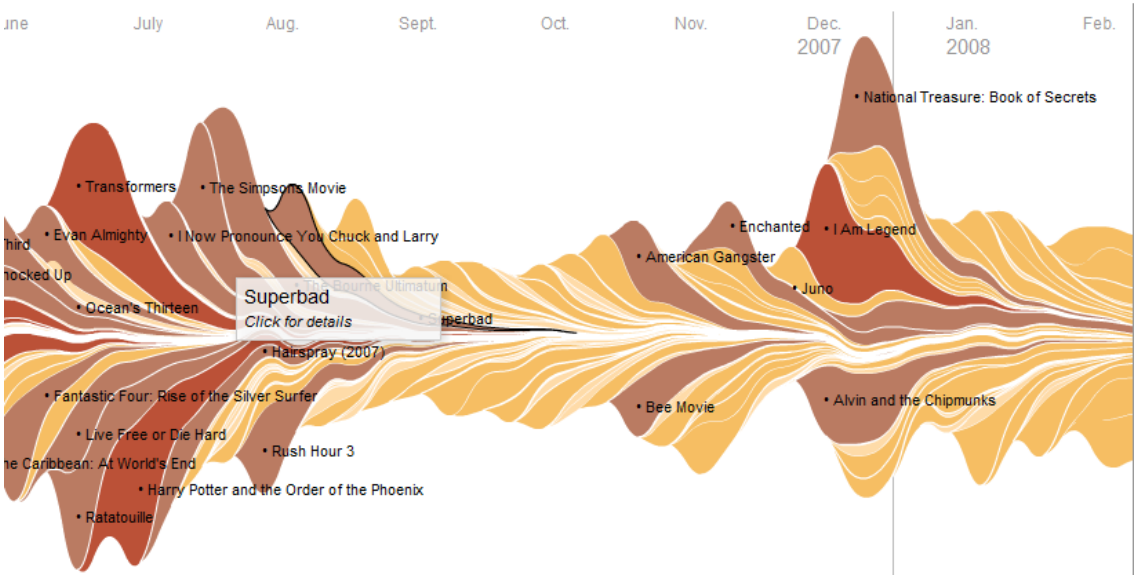
Sources: Baseline StudioSystems; Box Office Mojo

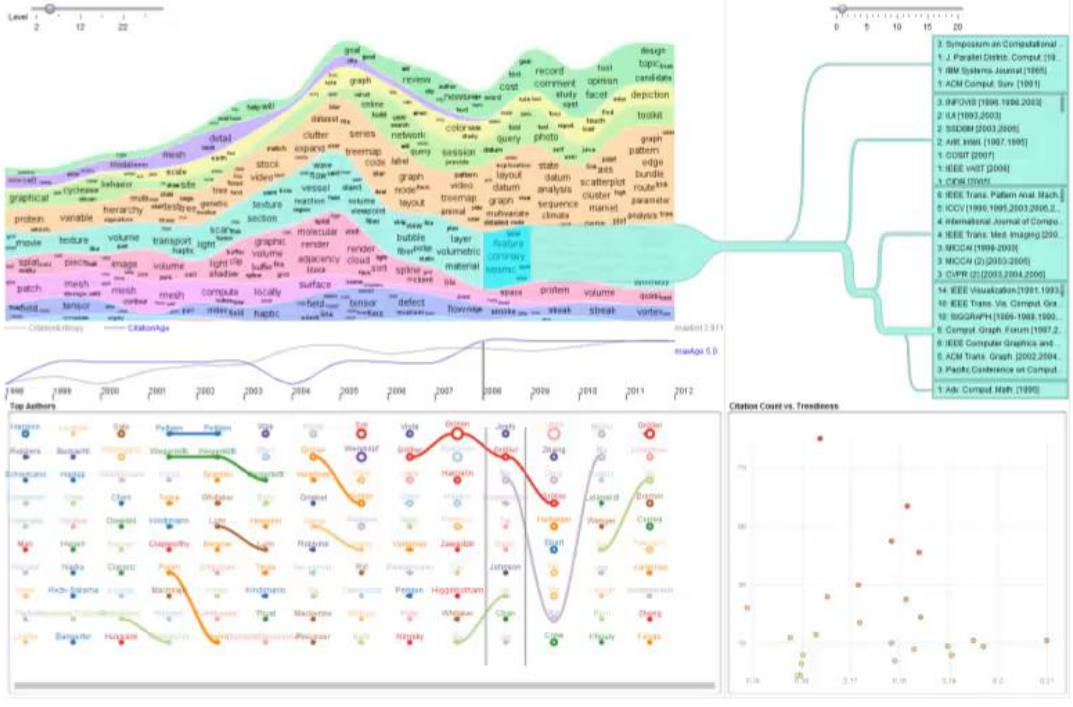
Mathew Bloch, Lee Byron, Shan Carter and Amanda Cox



https://datavizcatalogue.com/methods/stream_graph.html

DATA	MD table; one quantitative value attribute (e.g. counts), one ordered key attribute (e.g. time), one categorical key attribute (e.g. film)
DERIVE	order of layers is derived from a quantitative attribute
ENCODE	use derived geometry to show layers across time, layer height encodes count
SCALE	key attributes (time, main axis): hundreds of time points key attributes (short axis): dozens to hundreds





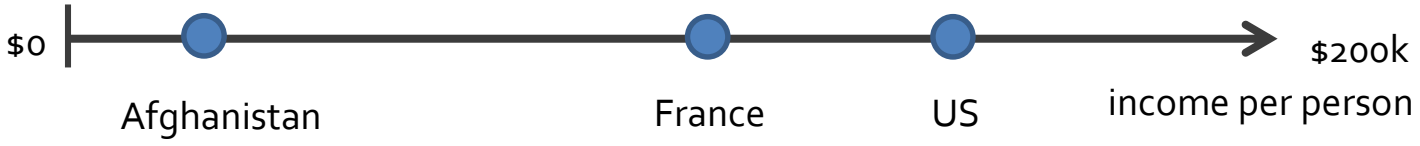
CiteRivers

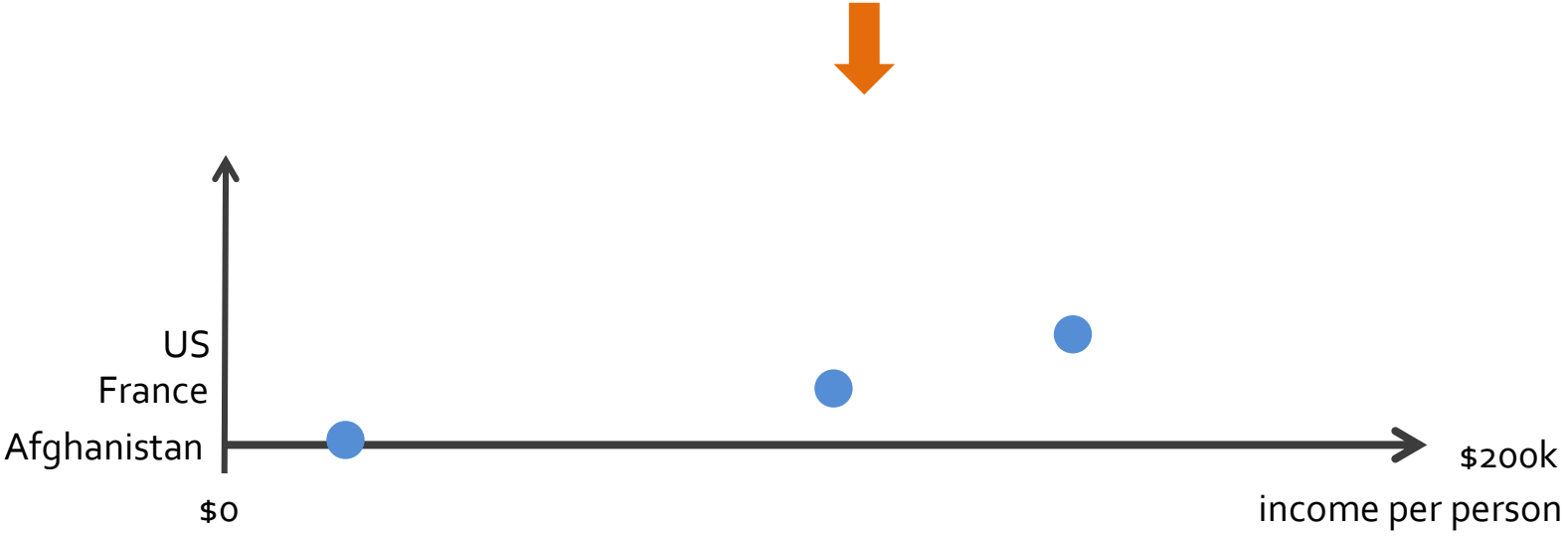
Florian Heimerl, Qi Han,
Steffen Koch, Thomas Ertl
University of Stuttgart
florian.heimerl@vis.uni-stuttgart.de

IEEE VAST 2015

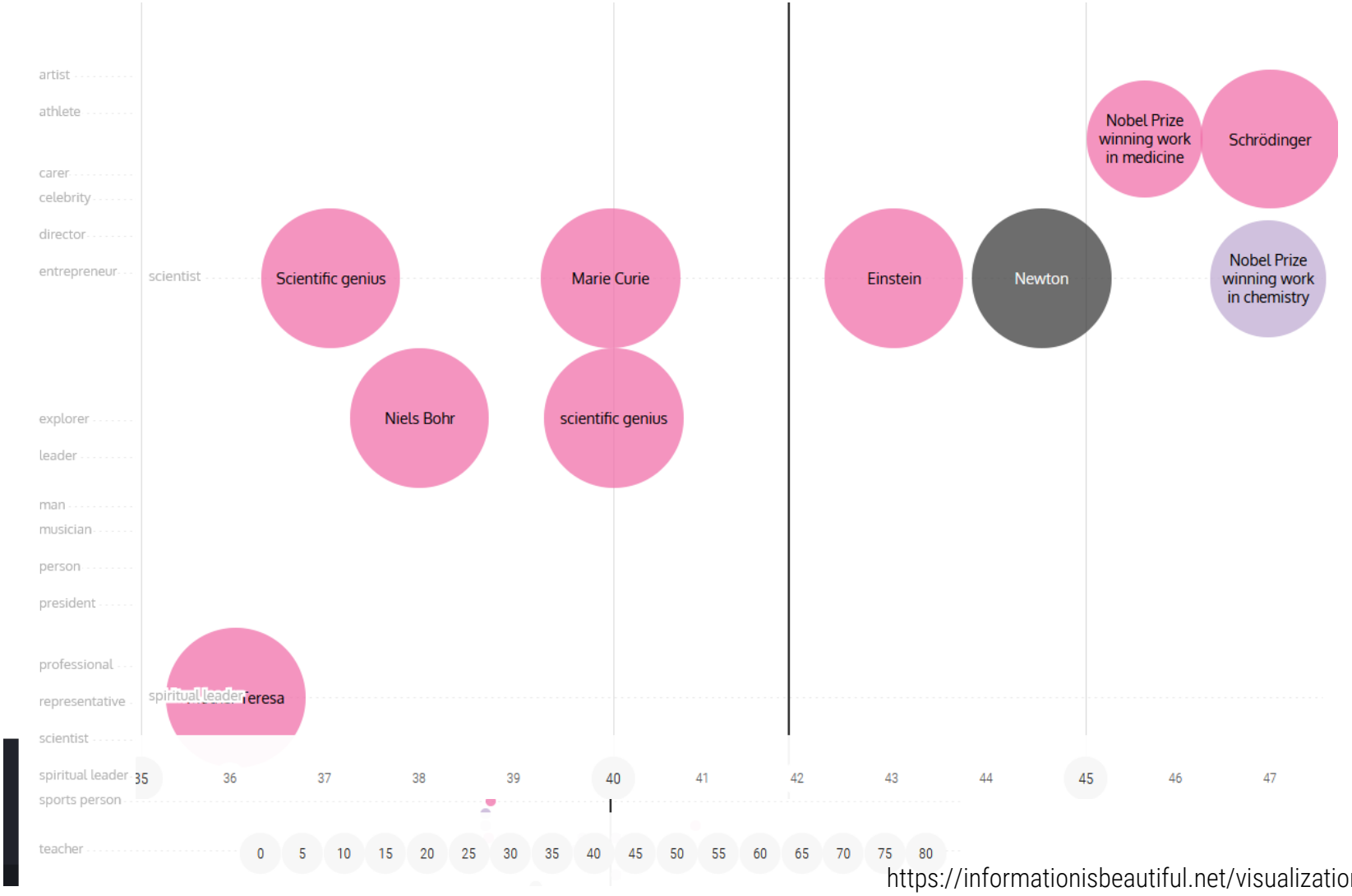
Institute for Visualization and Interactive Systems

DOT CHART/PLOT





Who Old Are You?



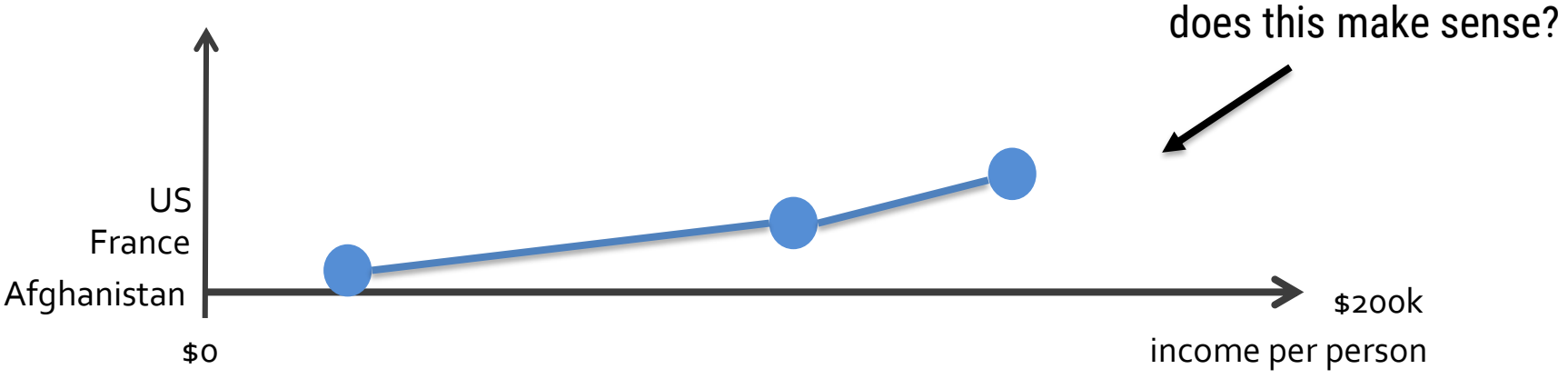
https://informationisbeautiful.net/visualizations/who-old-are-you/

LINE CHART

augment with line connection marks

emphasize the ordering and show trends

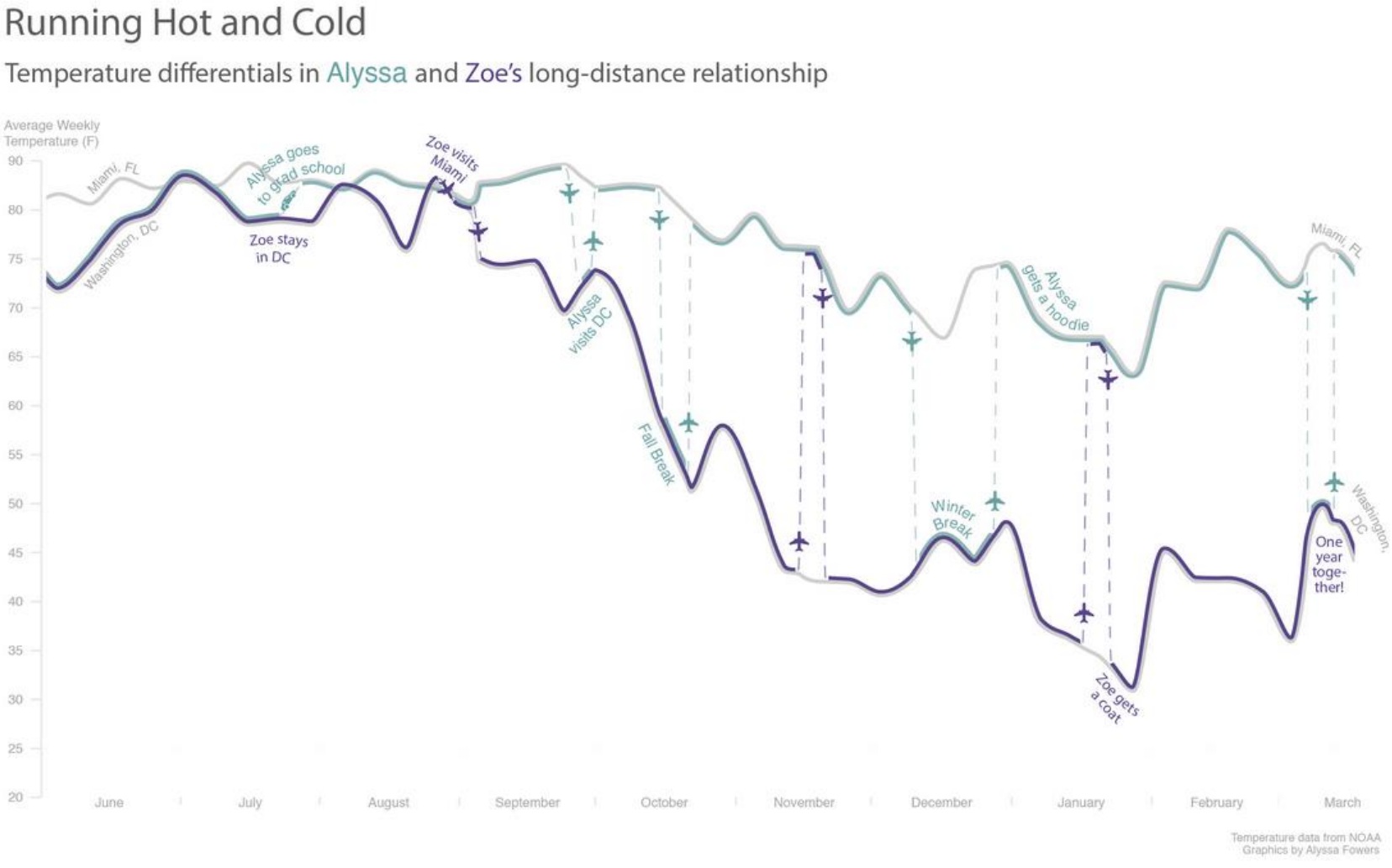
should not be used with categorical keys



Housing market declines since 1980

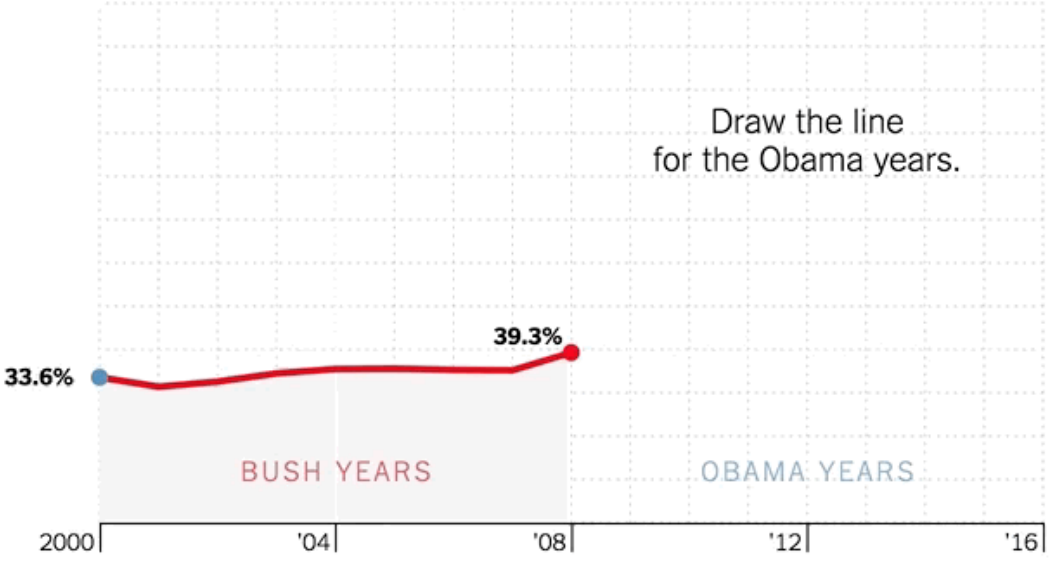
Each line represents a downturn.
(Longer line = longer downturn; steeper line = steeper downturn)

https://www.informationisbeautifulawards.com/showcase/3917-house-of-cards



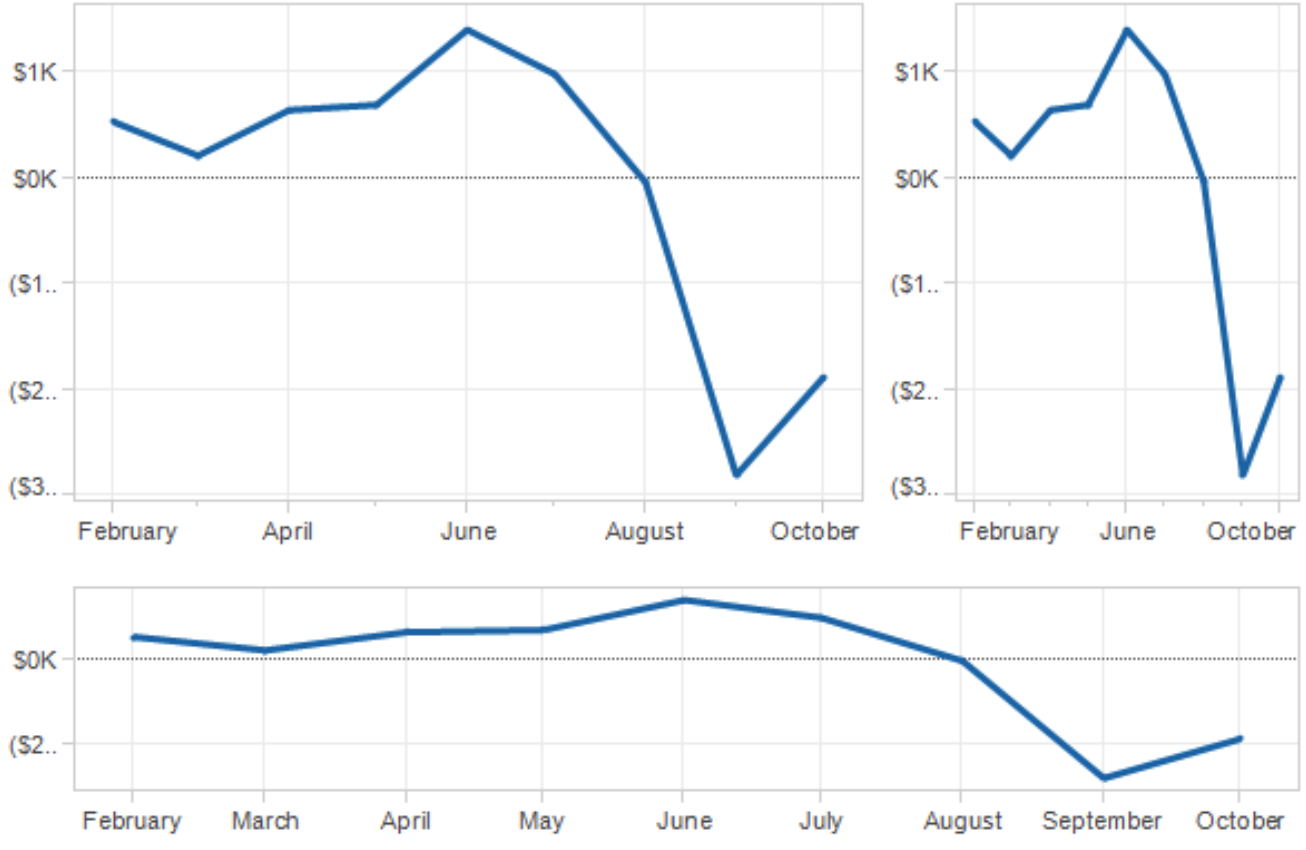
https://www.alyssafowers.com/portfolio/#/running-hot-and-cold/

Under Mr. Obama, the **national debt** as a percentage of the gross domestic product ...



Numbers

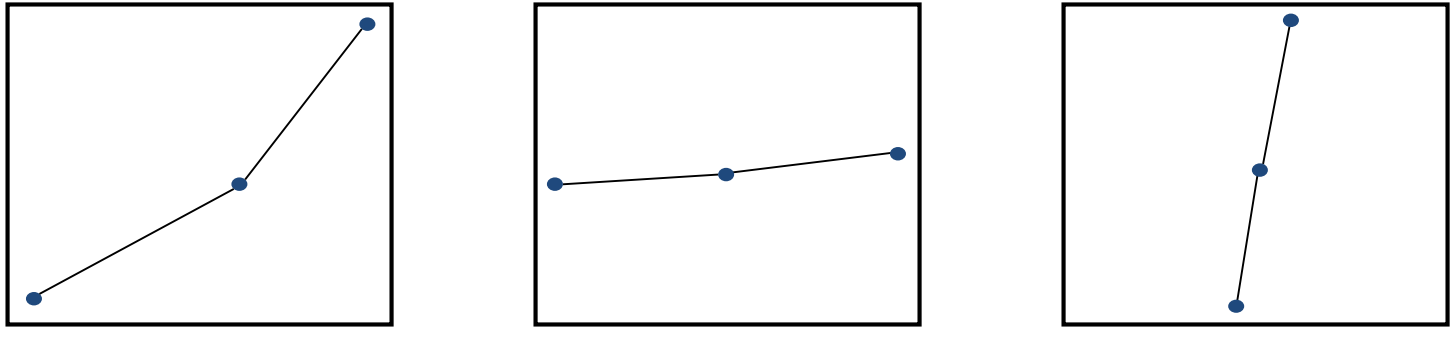
ASPECT RATIO SELECTION



BANKING TO 45°

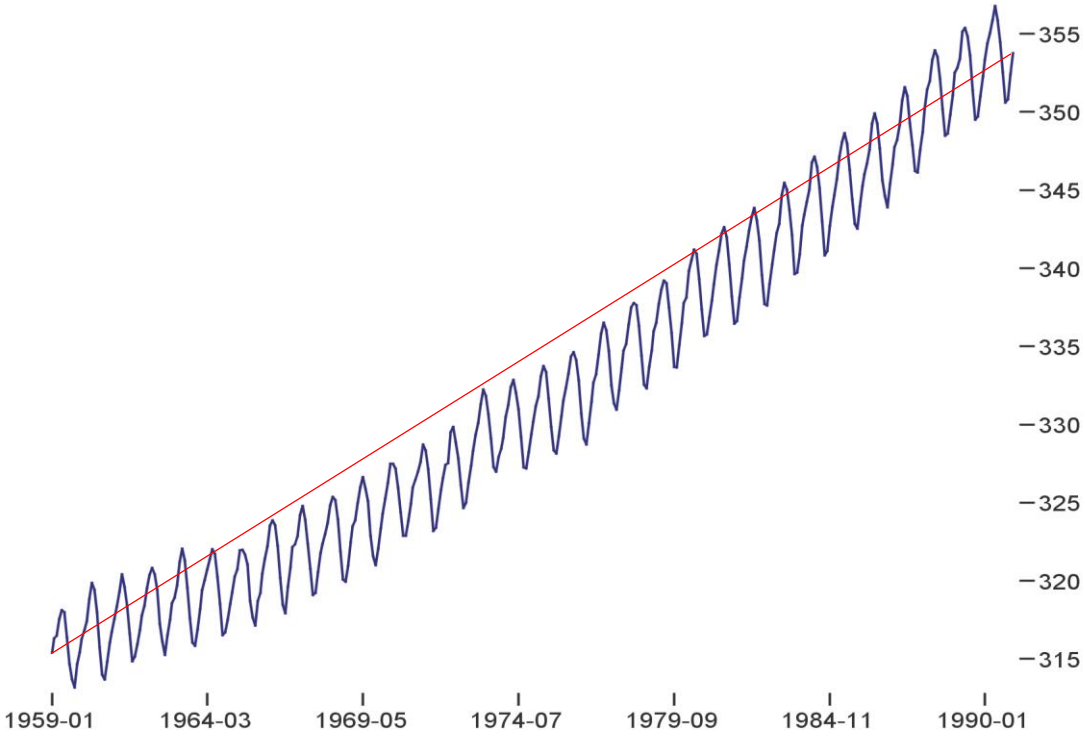
[Cleveland]

TO FACILITATE PERCEPTION OF TRENDS, MAXIMIZE THE DISCRIMINABILITY OF LINE SEGMENT ORIENTATIONS

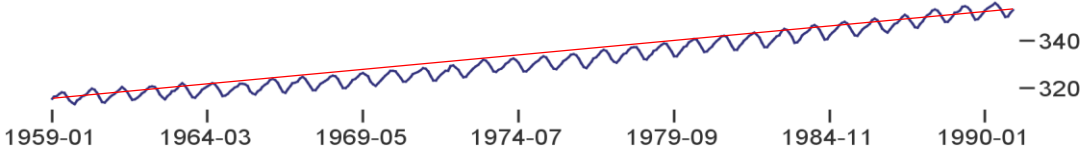


TWO SEGMENTS ARE MAXIMALLY DISCRIMINABLE WHEN THEIR AVG ABSOLUTE ANGLE IS 45°

OPTIMIZE THE *ASPECT RATIO* TO BANK TO 45°

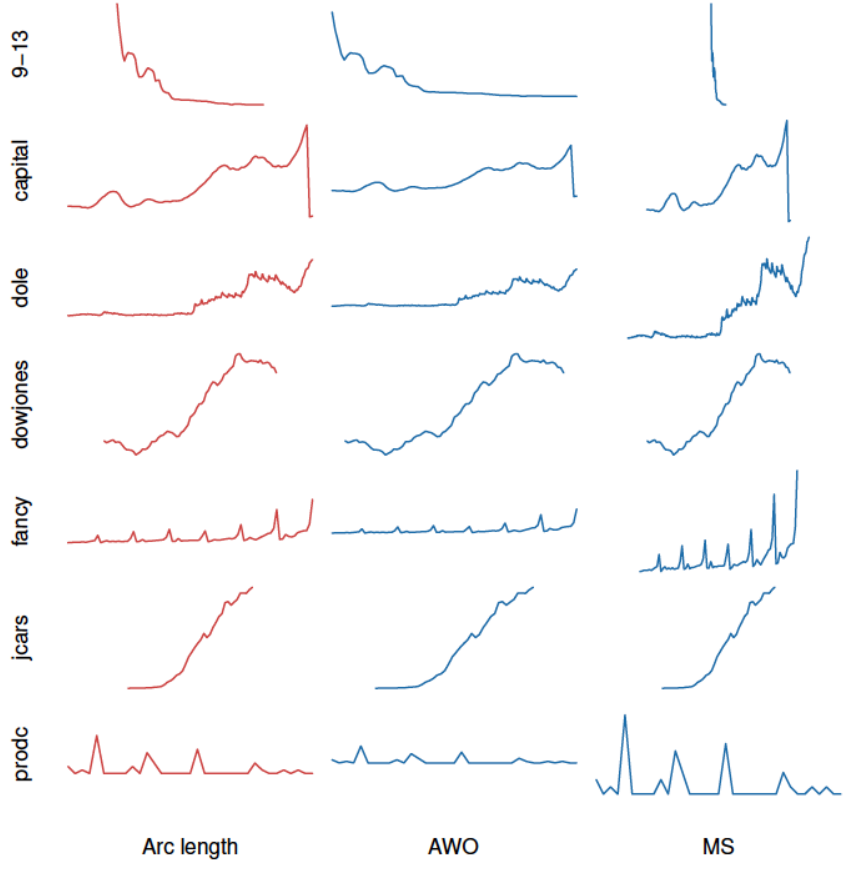


ASPECT RATIO = 1.17



ASPECT RATIO = 7.87

ALTERNATIVE METHODS



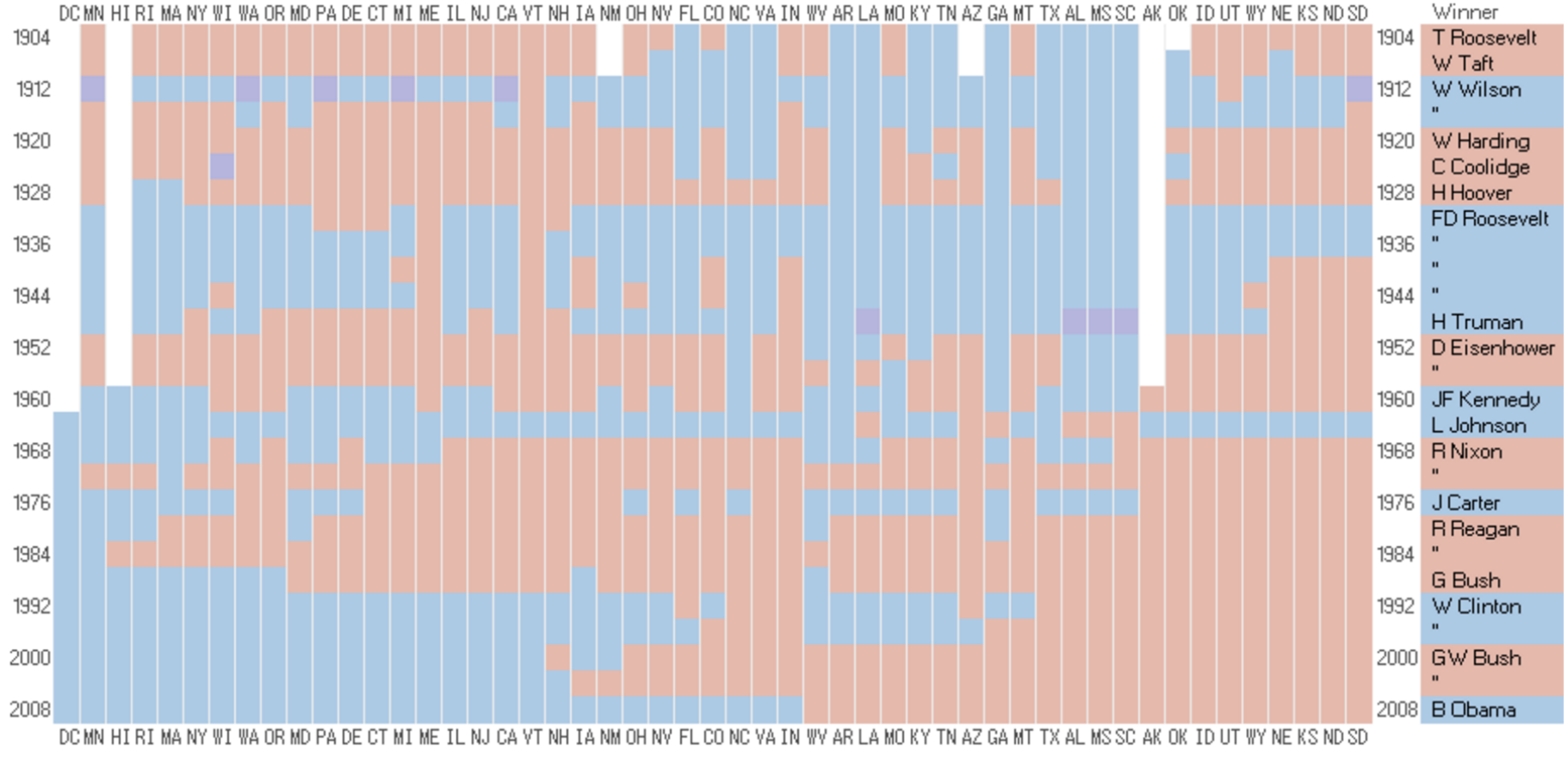
Practical advice:

CHOOSE AN ASPECT RATIO THAT EMPHASIZES THE IMPORTANT DETAILS FOR YOUR TASK

[TALBOT ET AL, 2011]

MATRIX ALIGNMENT

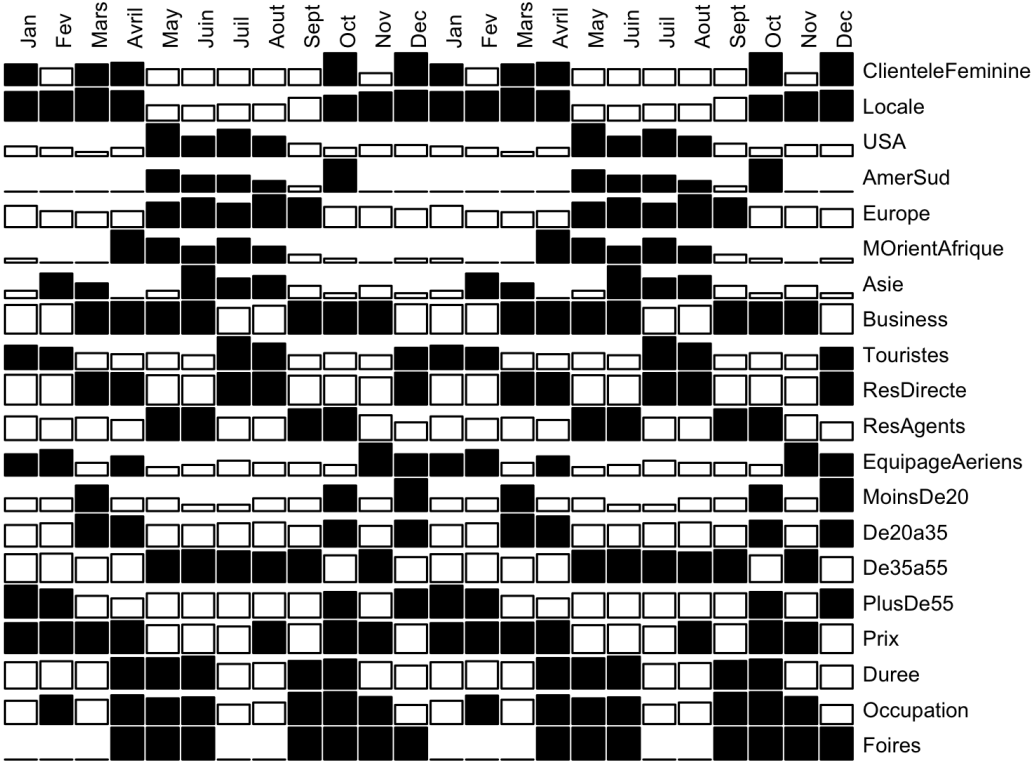
Two keys



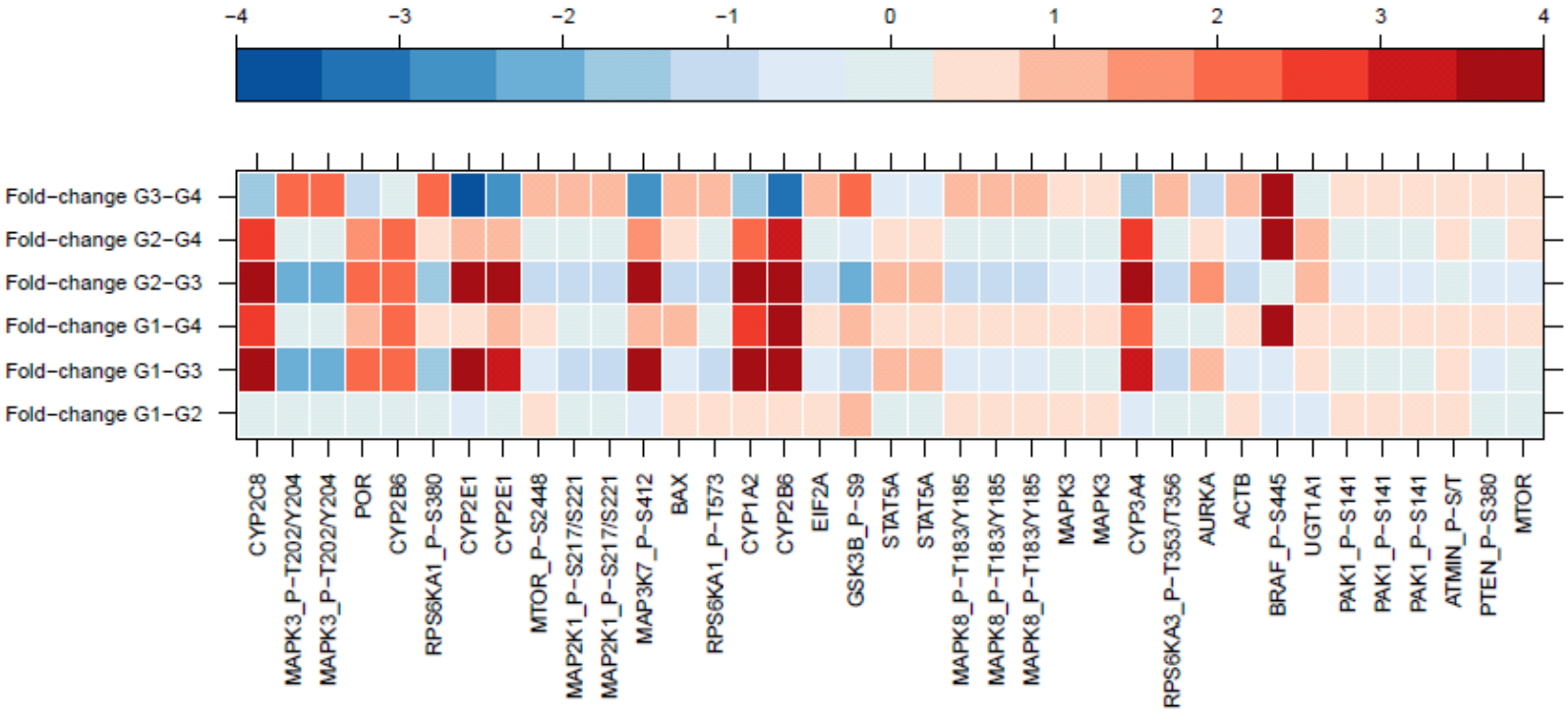
https://ldld.samizdat.cc/2016/tag/catalog/

HEATMAP

Hotel 2



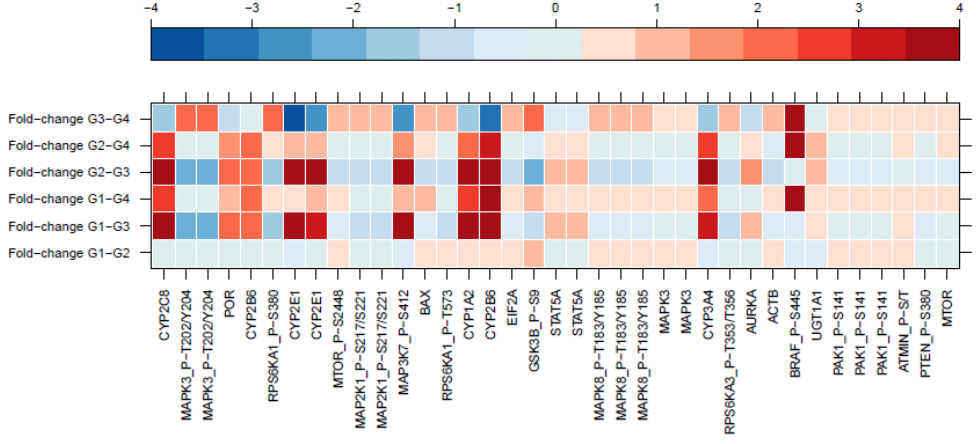
HEATMAP



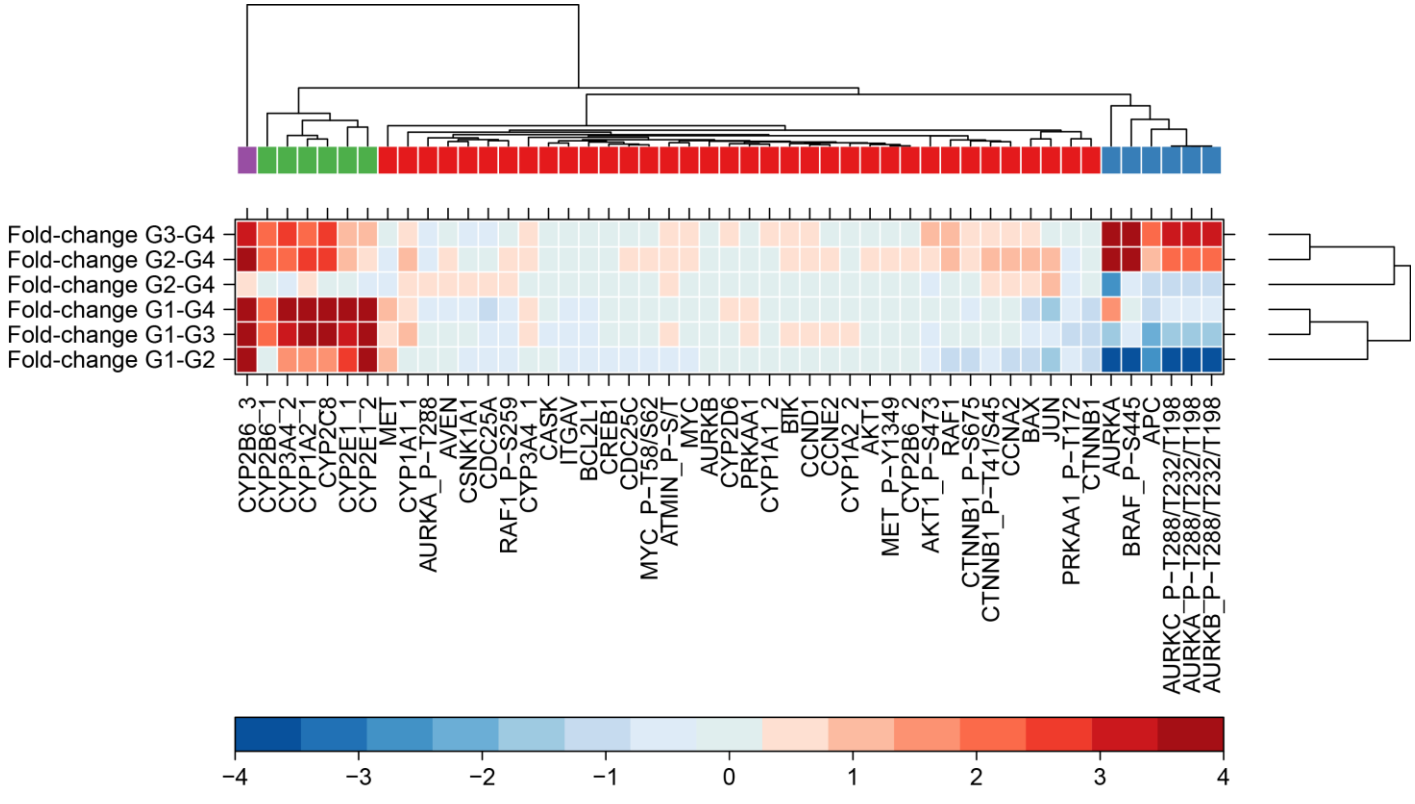
http://www.ra.cs.uni-tuebingen.de/software/RPPApipe/doc/documentation.htm

HEATMAP

DATA	Table; two categorical key attributes, one quantitative value attribute
ENCODE	2D matrix alignment of area marks, e.g. with diverging color map
TASK	find clusters, outliers; summarize
SCALE	items: ~1 million (on 1000x1000px), categorical attribute levels: hundreds, quantitative attribute levels: 3-11



CLUSTERED HEATMAP



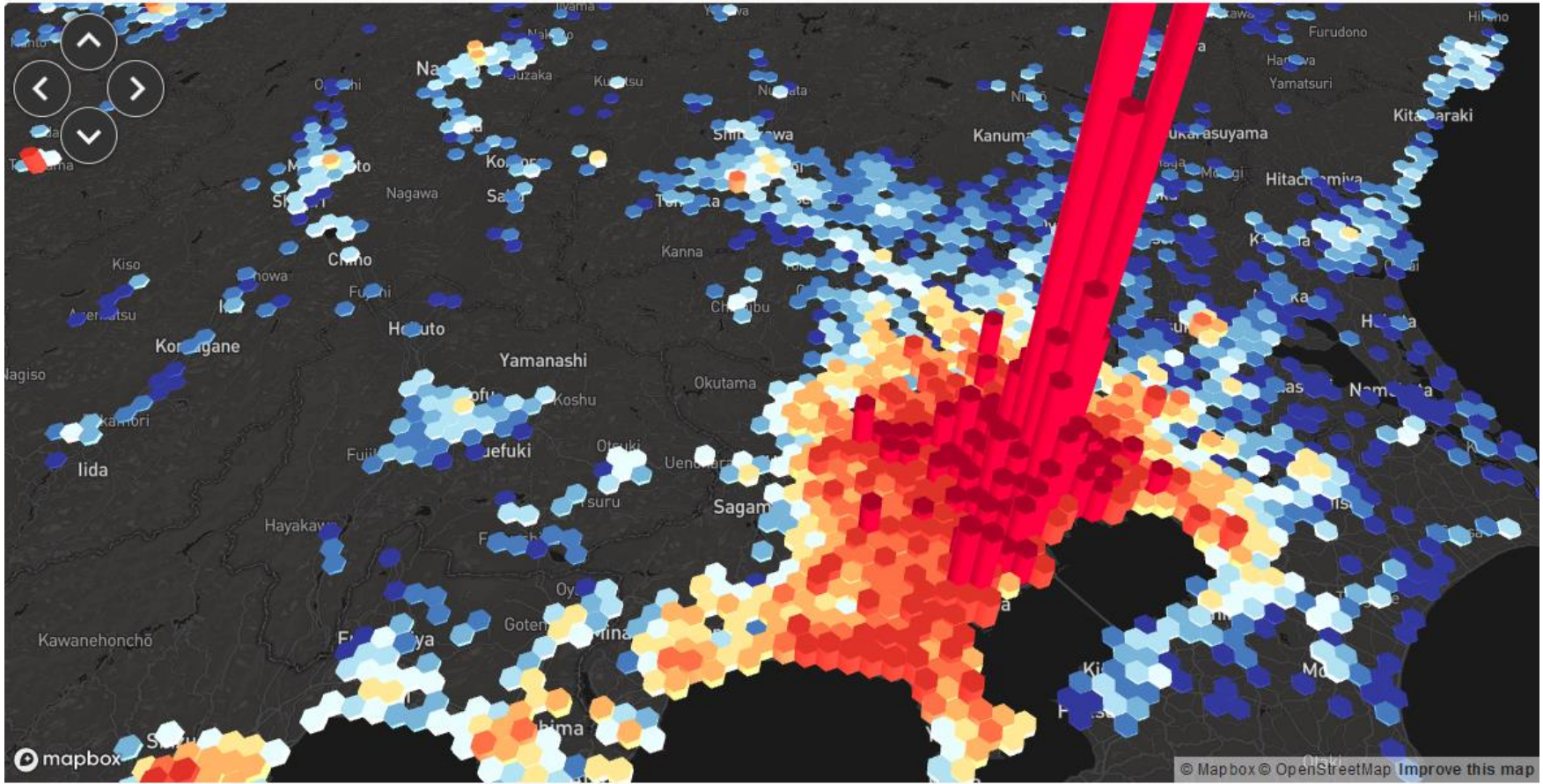
http://www.ra.cs.uni-tuebingen.de/software/RPPApipe/doc/documentation.htm

3D Map of Land Values in Japan: 1989 - 2019

26,000 points of official land prices of the whole country for the past 30 years

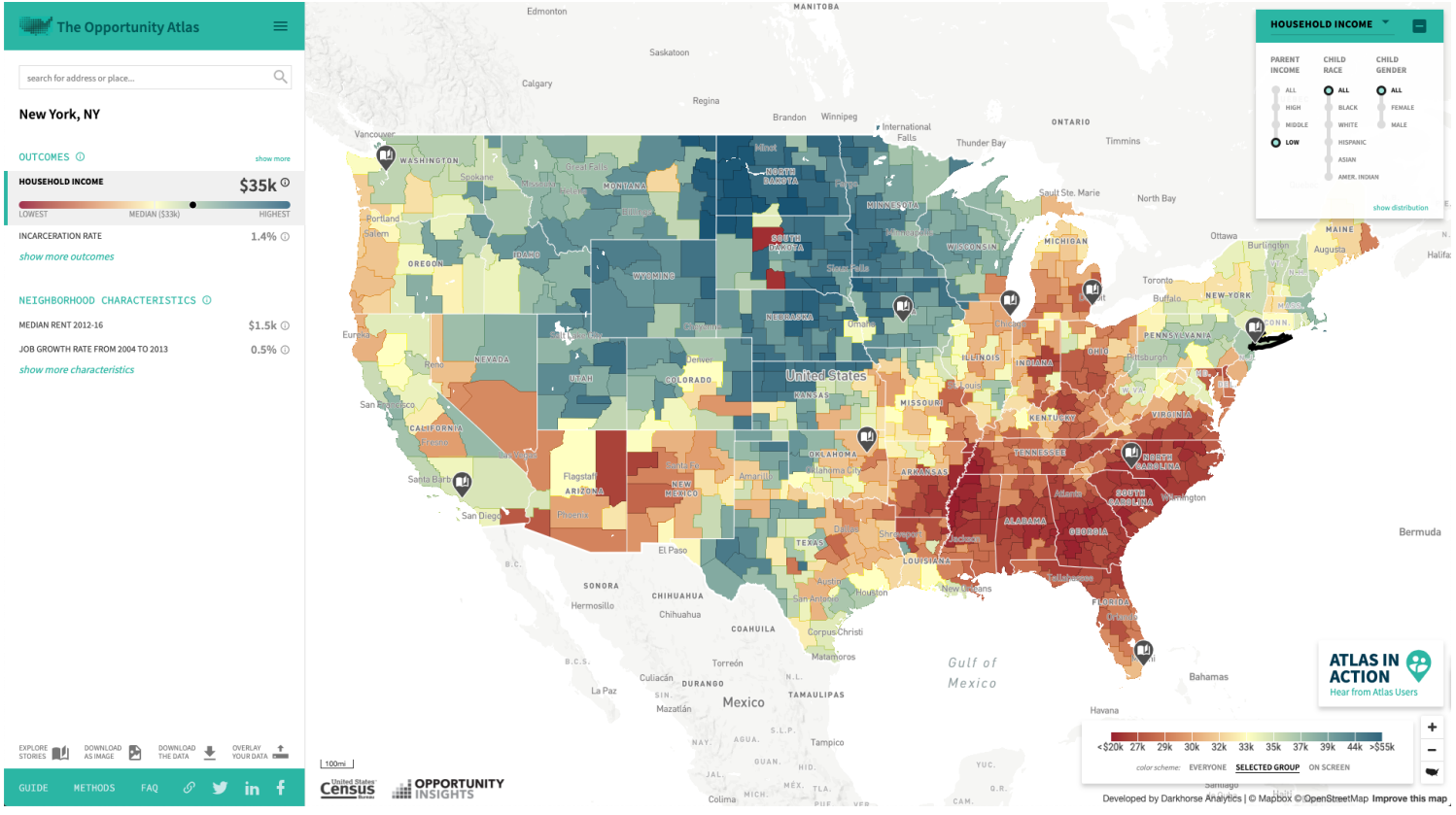
3rd April 2019 JAPANESE

Latest value | All points



https://toyokeizai.net/sp/visual/tko/landprices3/en.html

CLOROPLETH MAP



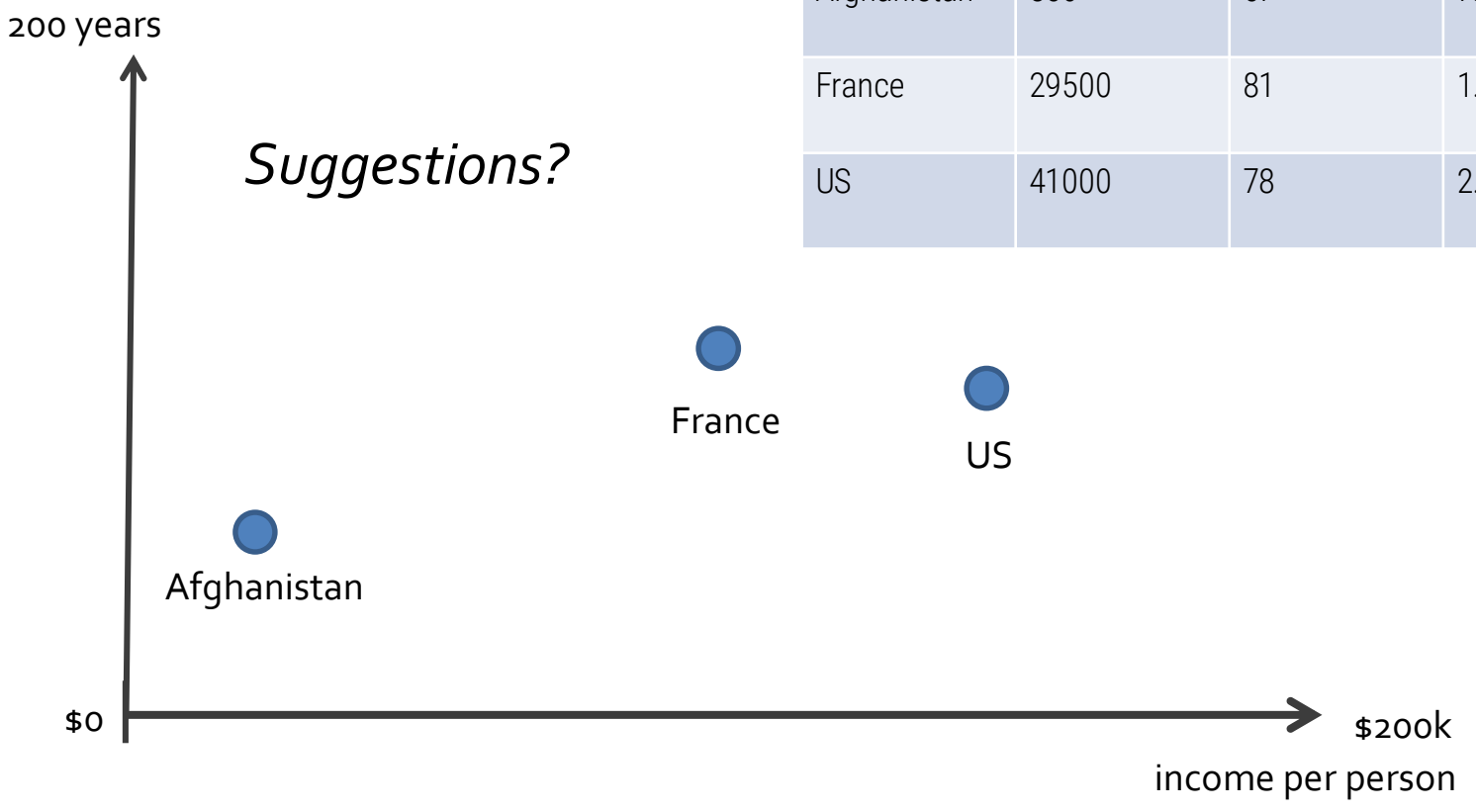
Uses heatmap idea

BACK TO OUR ORIGINAL EXAMPLE

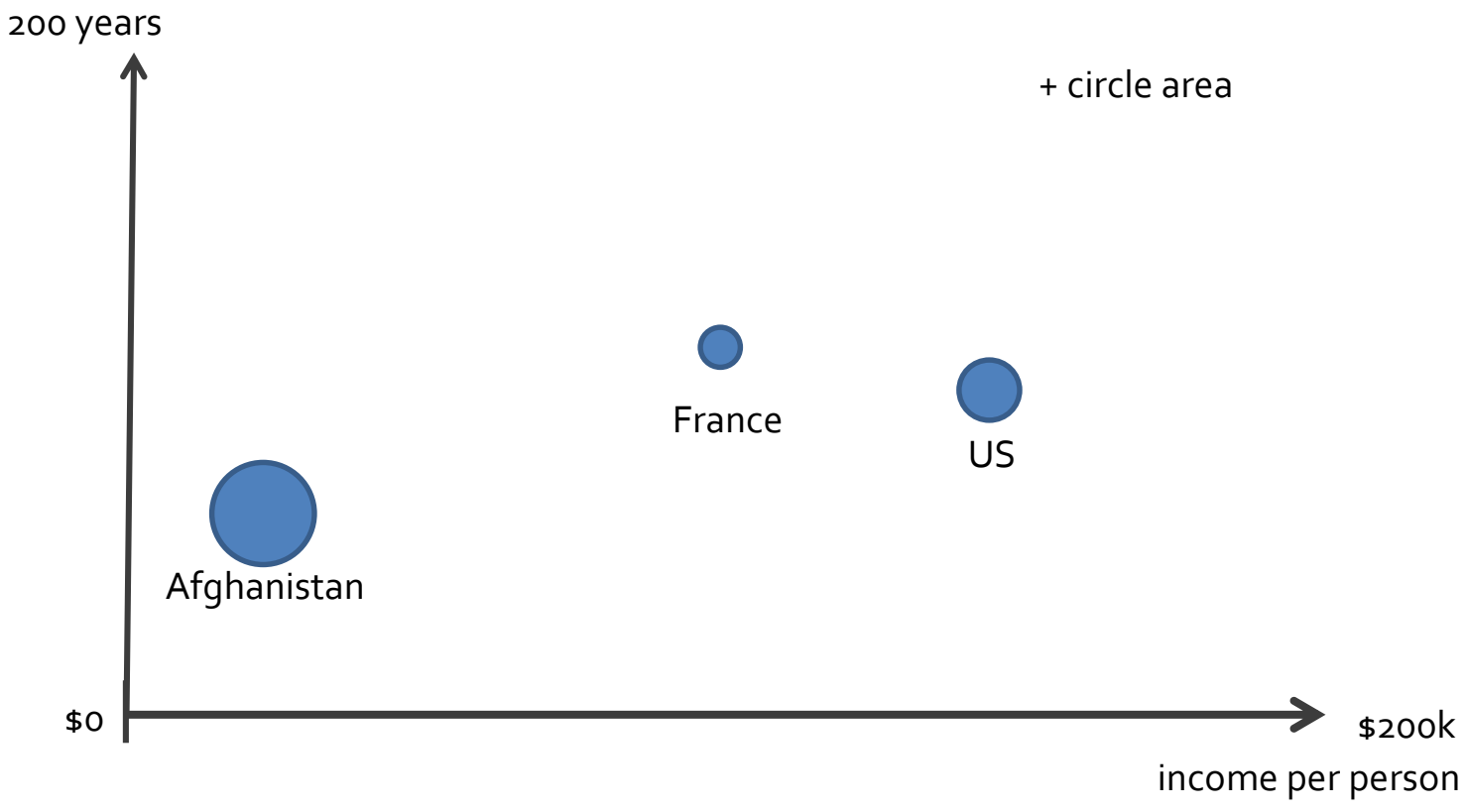
Country	Income per person	Life expectancy	Children per woman
Afghanistan	850	57	7.1
France	29500	81	1.9
US	41000	78	2.1

now with 4 attributes

Country	Income per person	Life expectancy	Children per woman
Afghanistan	850	57	7.1
France	29500	81	1.9
US	41000	78	2.1

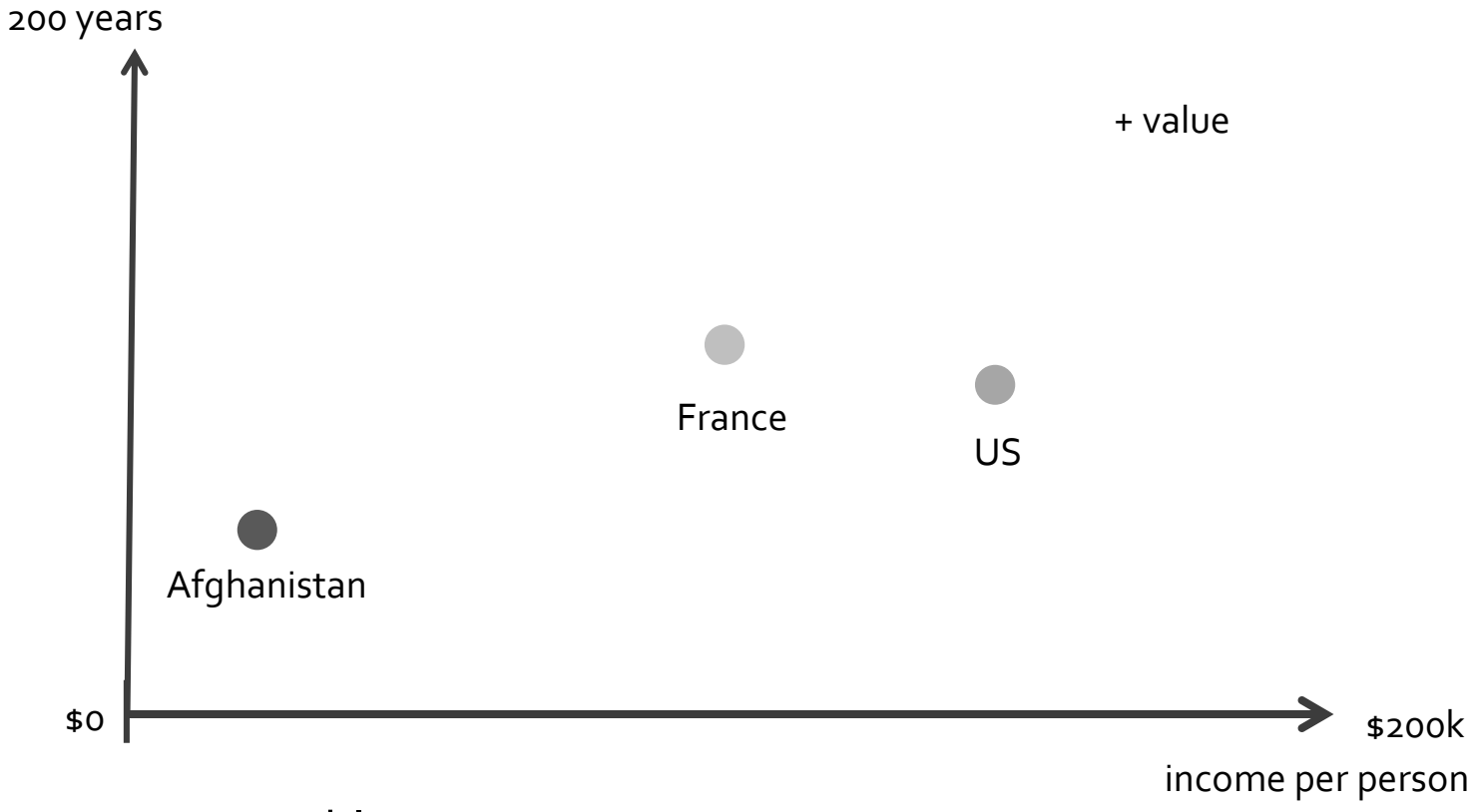


ADD ANOTHER VISUAL ENCODING



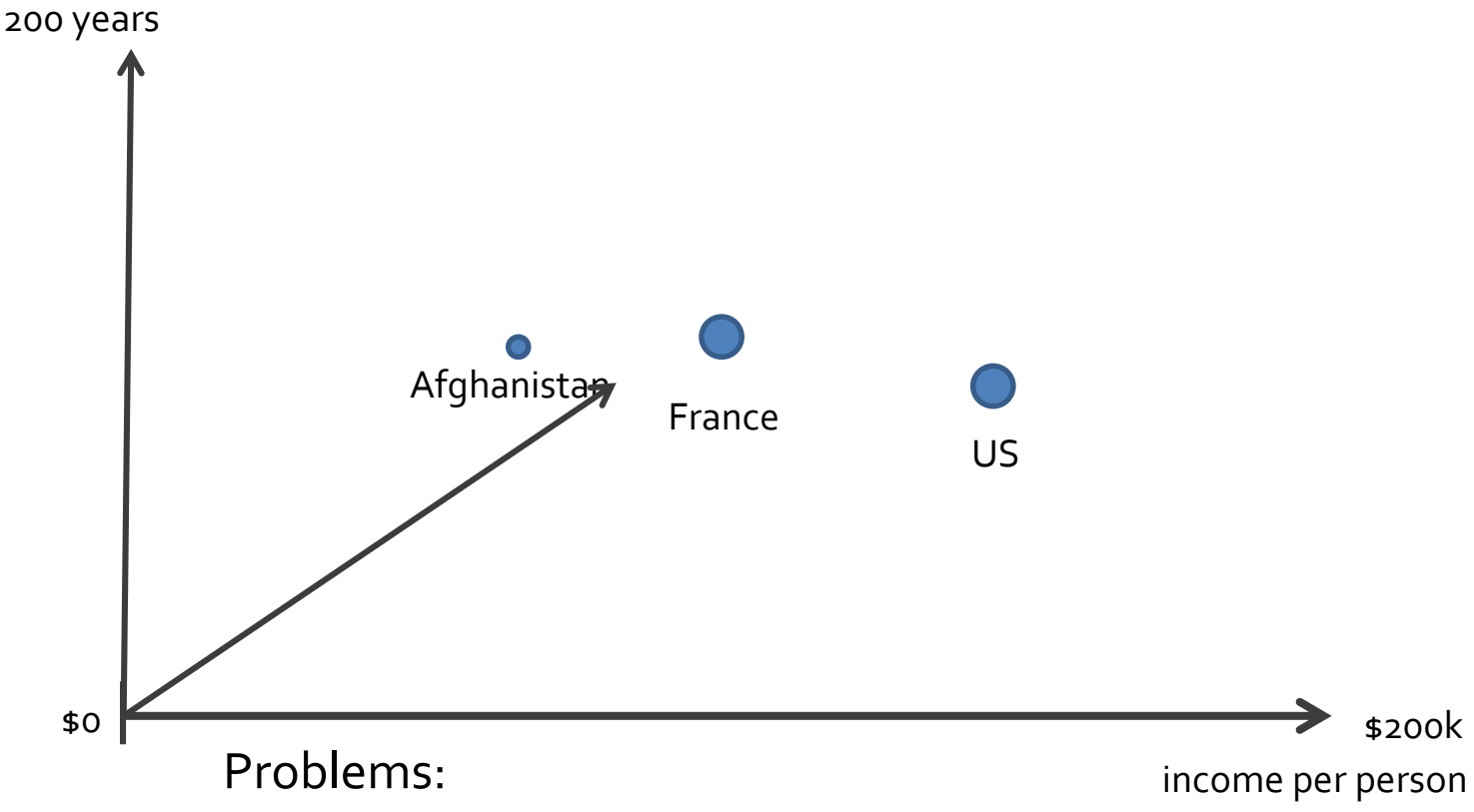
Problem:
Does not scale well to more attributes

ADD ANOTHER VISUAL ENCODING



Problem:
Does not scale well to more attributes

ADD AN AXIS

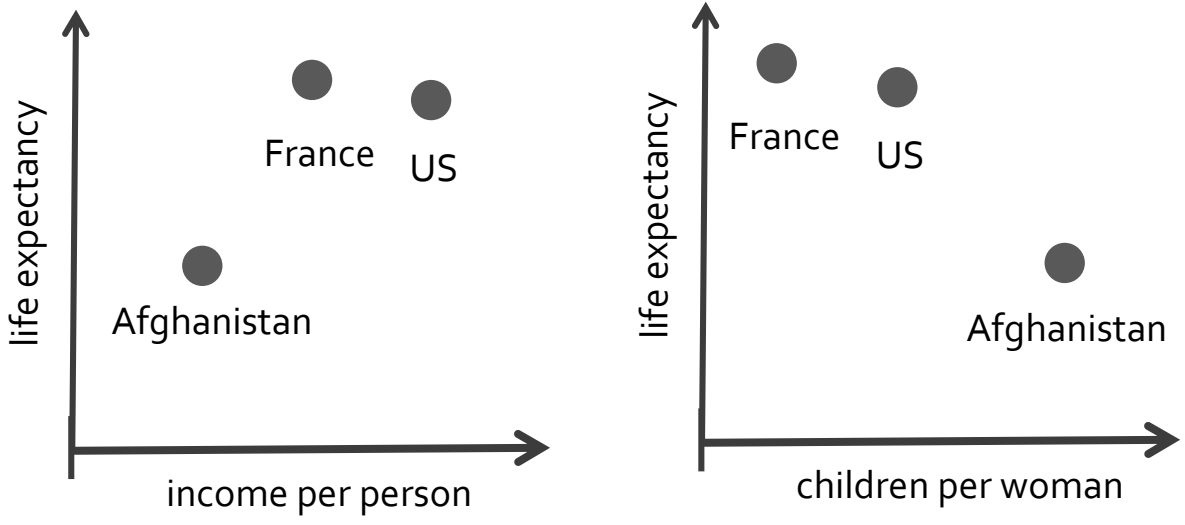


Problems:

Occlusion, perspective distortion, does not scale

→ Not usually recommended

ADD AN AXIS


life expectancy
France
US
Afghanistan
income per person
life expectancy
France
US
Afghanistan
children per woman

SCATTERPLOT MATRIX

This idea scales relatively well

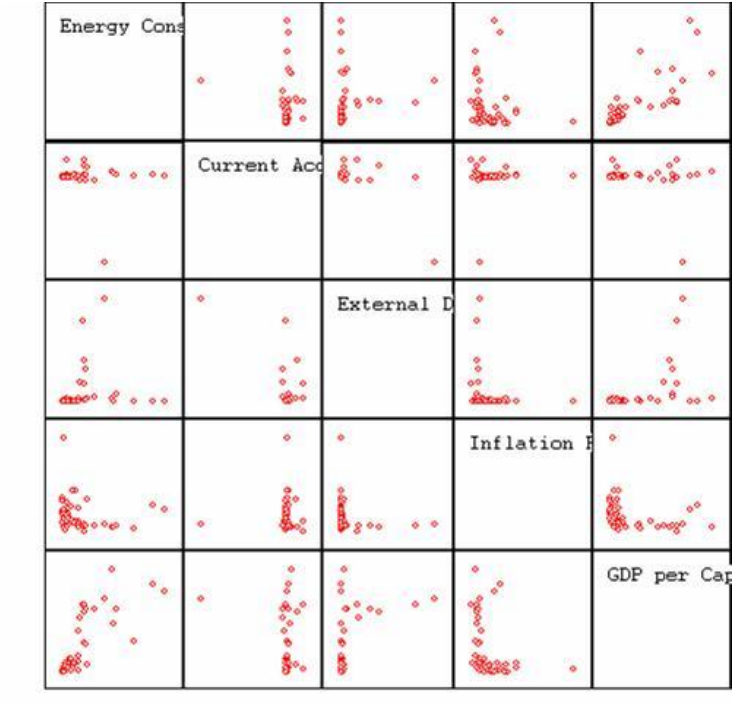


Image Source: Wikipedia

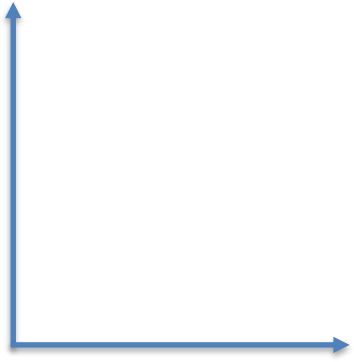
GraphDice: A System for Exploring Multivariate Social Networks

A. Bezerianos
F. Chevalier
P. Dragicevic
N. Elmqvist
J-D. Fekete

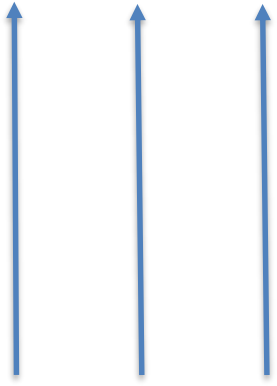
INRIA
École Centrale Paris
Purdue University

SPATIAL AXIS ORIENTATION

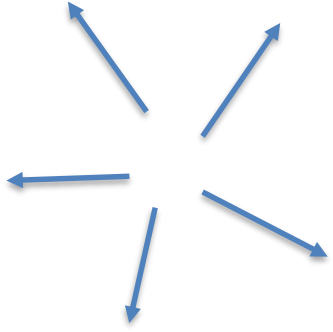
An additional design choice



rectilinear



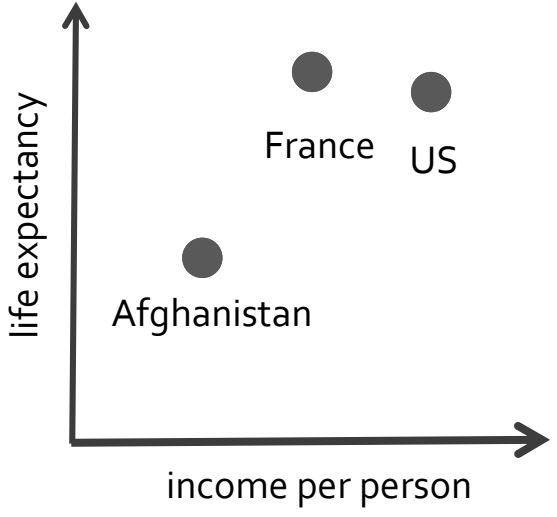
parallel



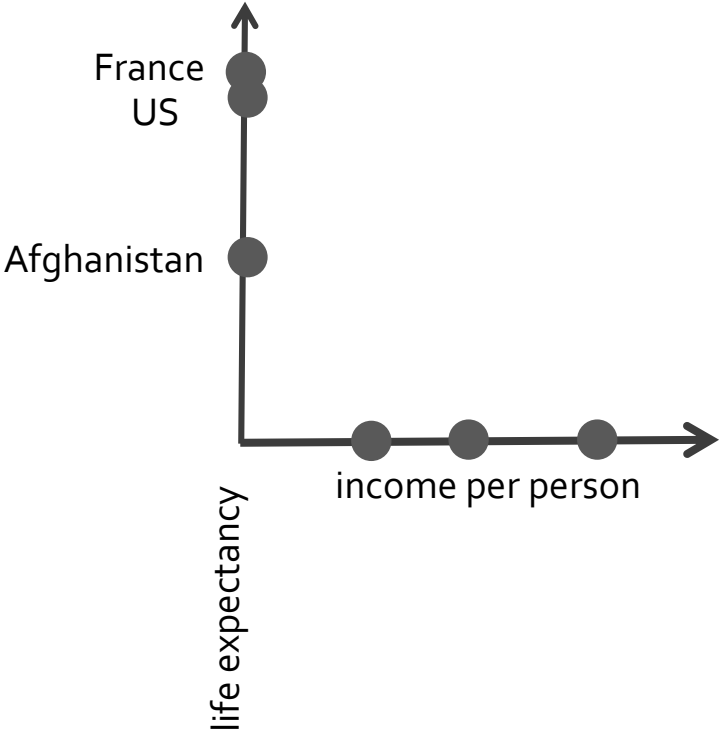
radial

parallel coordinates

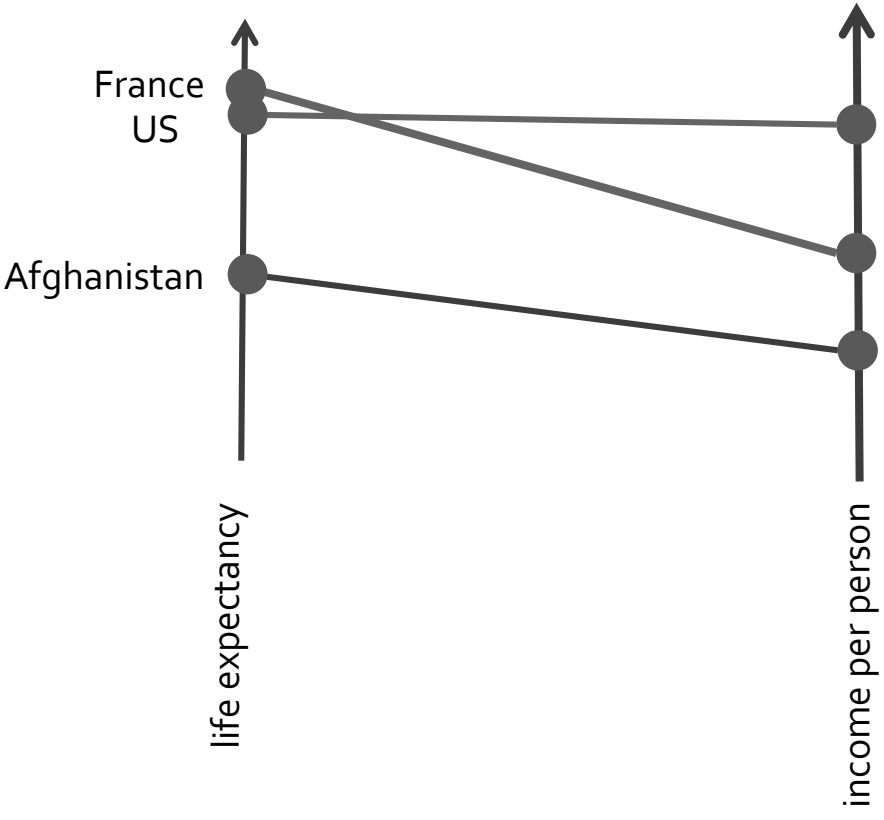
Back to our original example



Parallel Coordinates



parallel coordinates



- show correlations between neighboring axes

MULTIDIMENSIONAL DETECTIVE

Alfred Inselberg*, Multidimensional Graphs Ltd†
&
Computer Science Department
Tel Aviv University, Israel
aiisreal@math.tau.ac.il

Abstract

The display of multivariate datasets in parallel coordinates, transforms the search for *relations* among the variables into a 2-D pattern recognition problem. This is the basis for the application to *Visual Data Mining*. The Knowledge Discovery process together with some general guidelines are illustrated on a dataset from the production of a VLSI chip. The special strength of parallel coordinates is in modeling **relations**. As an example, a simplified Economic Model is constructed with data from various economic sectors of a real country. The visual model shows the interelationship and dependencies between the sectors, circumstances where there is competition for the same resource, and feasible economic policies. Interactively, the model can be used to do trade-off analyses, discover sensitivities, do approximate optimization, monitor (as in a Process) and Decision Support.

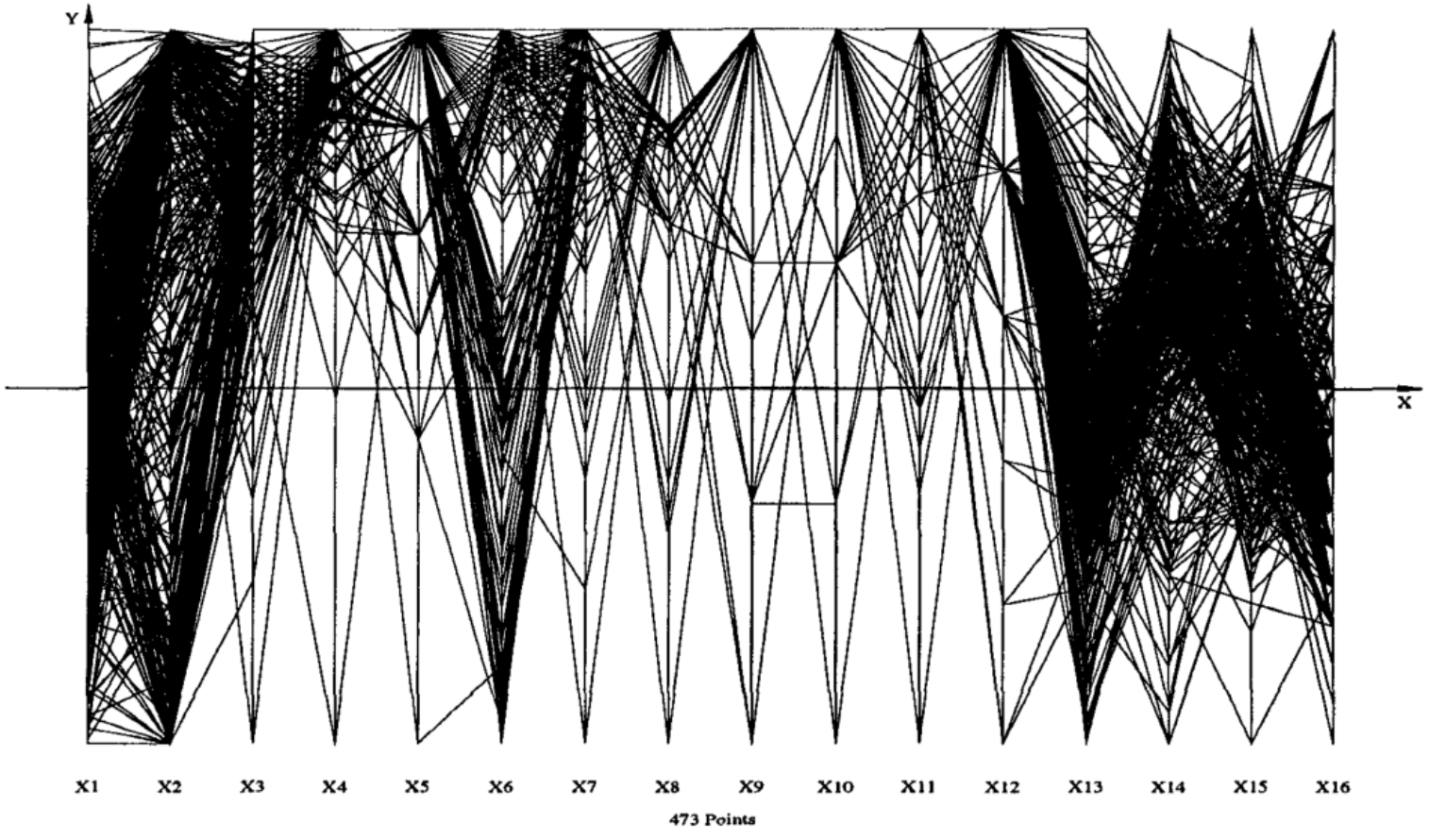
Introduction

In Geometry parallelism, which does not require a notion of angle, rather than orthogonality is the more fundamental concept. This, together with the fact that orthogonality "uses-up" the plane very fast, was the inspiration in 1959 for "Parallel" Coordinates. The systematic development began in 1977 [4]. The goals of the program were and still are (see [6] and [5] for short reviews) the visualization of multivariate/multidimensional problems without loss of information and having the properties:

1. Low representational complexity. Since the number of axes, N equals the number of dimensions (variables) the complexity is $O(N)$,
2. Works for any N,
3. Every variable is treated uniformly (unlike "Chernoff Faces" and various types of "glyphs"),
4. The displayed object can be recognized under projective transformations (i.e. rotation, translation, scaling, perspective),
5. The display easily/intuitively conveys information on the properties of the N-dimensional object it represents,
6. The methodology is based on rigorous mathematical and algorithmic results.

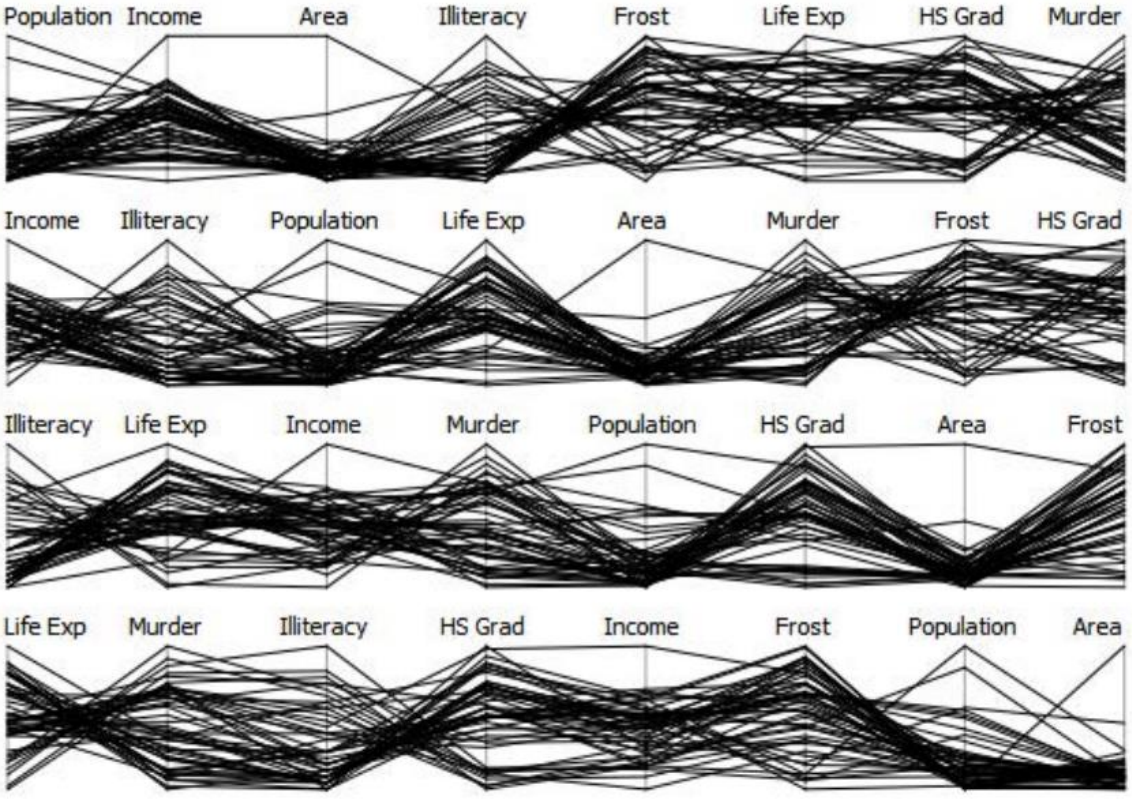
Parallel coordinates (abbr.||-coords) transform multivariate relations into 2-D patterns, a property that is well suited for Visual Data Mining.

*Senior Fellow San Diego SuperComputing Center
†36A Yehuda Halevy Street, Raanana 43556, Israel



Original Example from Inselberg 1997

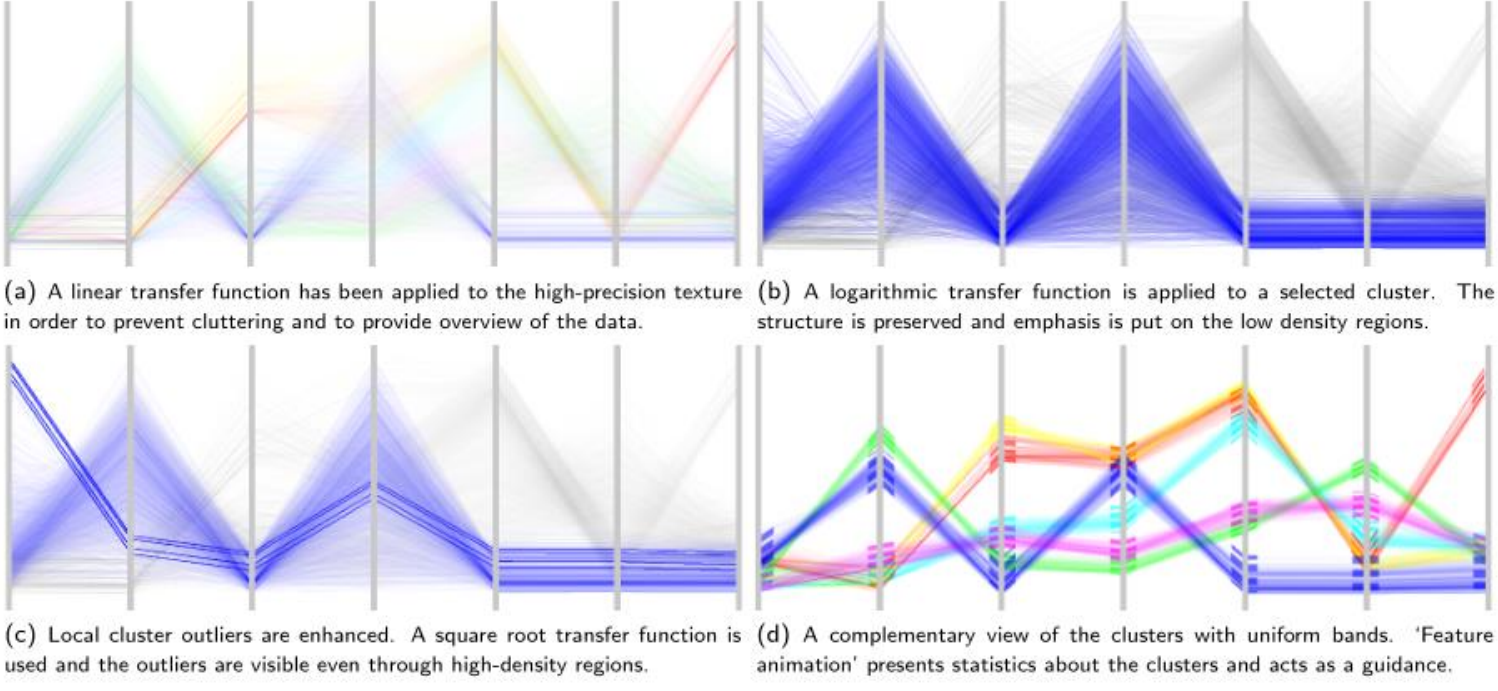
THE ORDER OF AXES MATTERS



Eurographics 2013, STAR Report
J. Heinrich, D. Weiskopf

REDUCE CLUTTER - HIGHLIGHT CLUSTERS

Lots of work on this. For example:



(a) A linear transfer function has been applied to the high-precision texture in order to prevent cluttering and to provide overview of the data.

(b) A logarithmic transfer function is applied to a selected cluster. The structure is preserved and emphasis is put on the low density regions.

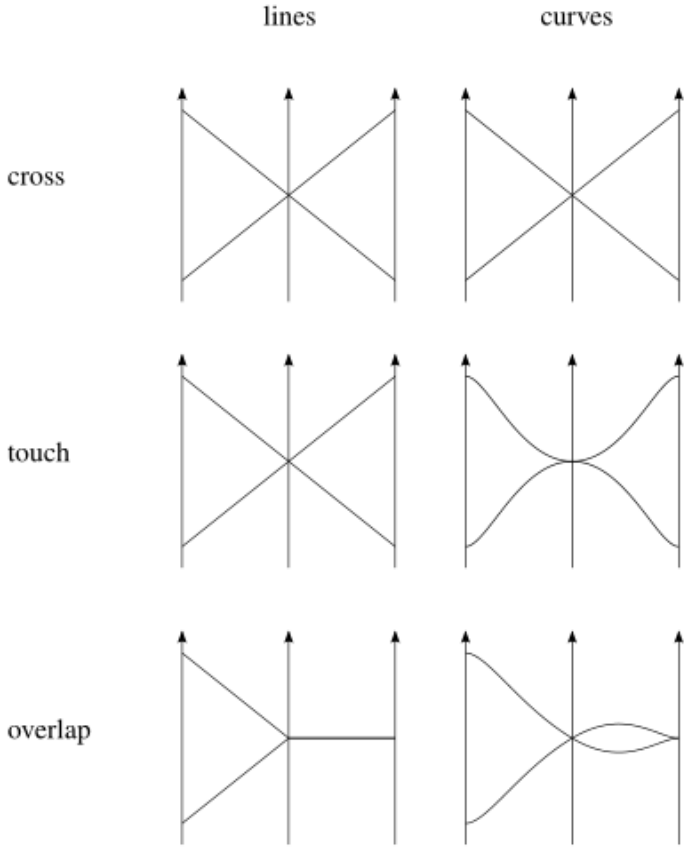
(c) Local cluster outliers are enhanced. A square root transfer function is used and the outliers are visible even through high-density regions.

(d) A complementary view of the clusters with uniform bands. 'Feature animation' presents statistics about the clusters and acts as a guidance.

Revealing Structure within Clustered Parallel Coordinates Displays, InfoVis 2005

HOW TO DRAW THE LINES

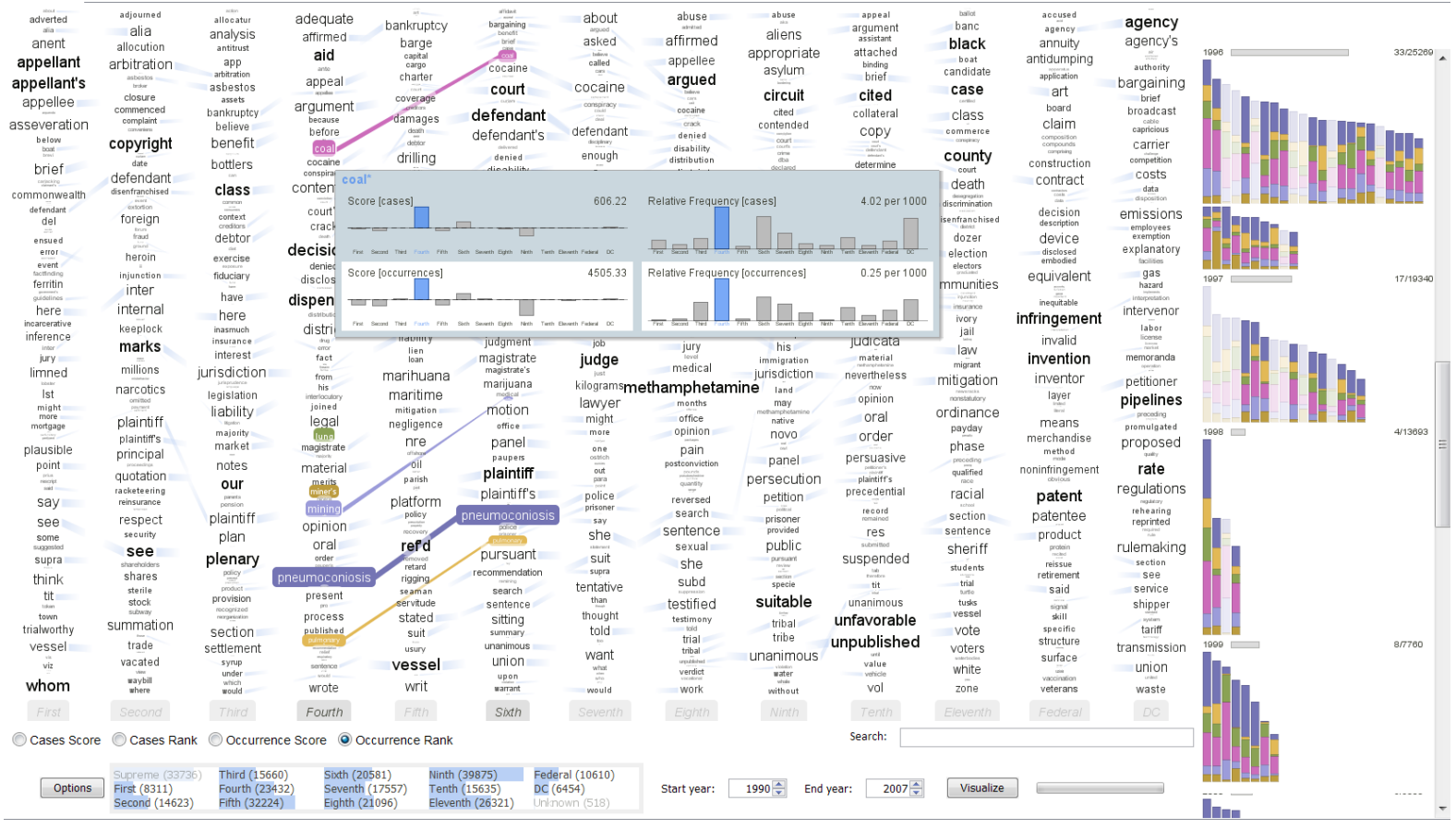
Goal: avoid ambiguity



Eurographics 2013, STAR Report
J. Heinrich, D. Weiskopf

COMBINE WITH OTHER VISUALIZATION TECHNIQUES

Parallel Tag Clouds to Explore Faceted Text Corpora (Collins et al., VAST 2009)



THERE IS MUCH MORE ON THIS…

Start here if you want more information

State of the Art of Parallel Coordinates

J. Heinrich and D. Weiskopf

Visualization Research Center, University of Stuttgart

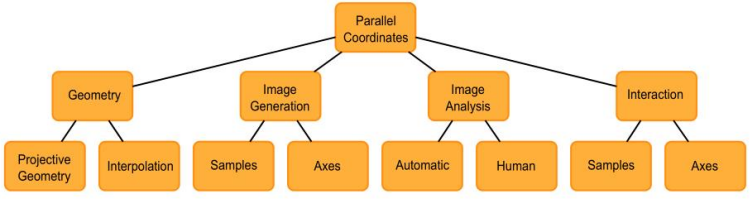


Figure 1: Taxonomy of topics for parallel coordinates in the scientific literature. The first-level nodes each represent a section in this paper, where the scope and definition of each topic will be explained.

Abstract

This work presents a survey of the current state of the art of visualization techniques for parallel coordinates. It covers geometric models for constructing parallel coordinates and reviews methods for creating and understanding visual representations of parallel coordinates. The classification of these methods is based on a taxonomy that was established from the literature and is aimed at guiding researchers to find existing techniques and identifying white spots that require further research. The techniques covered in this survey are further related to an established taxonomy of knowledge-discovery tasks to support users of parallel coordinates in choosing a technique for their problem at hand. Finally, we discuss the challenges in constructing and understanding parallel-coordinates plots and provide some examples from different application domains.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

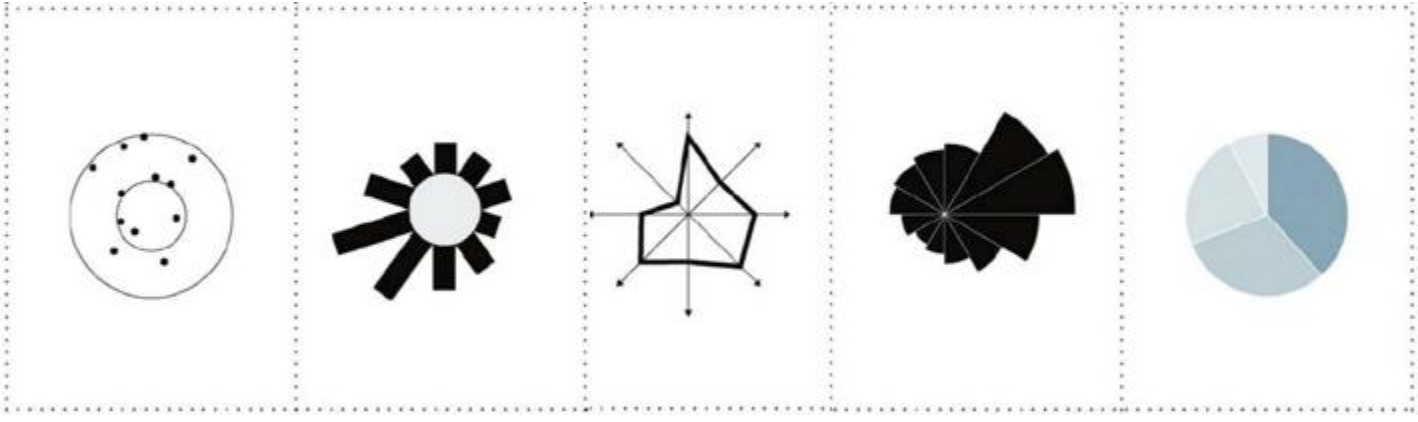
Scattering Points in Parallel Coordinates

Xiaoru Yuan[1], Peihong Guo[1], He Xiao[1], Hong Zhou[2], Huamin Qu[2]

1. Key Laboratary of Machine Perception (MOE), School of EECS, Peking University
2. Department of Computer Science and Engineering at Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

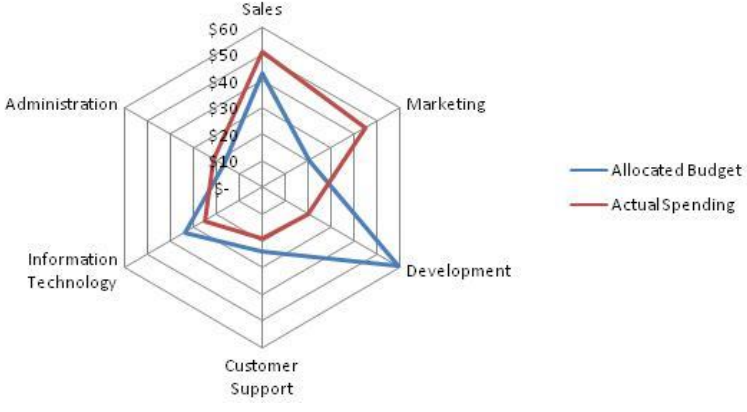
RADIAL AXES

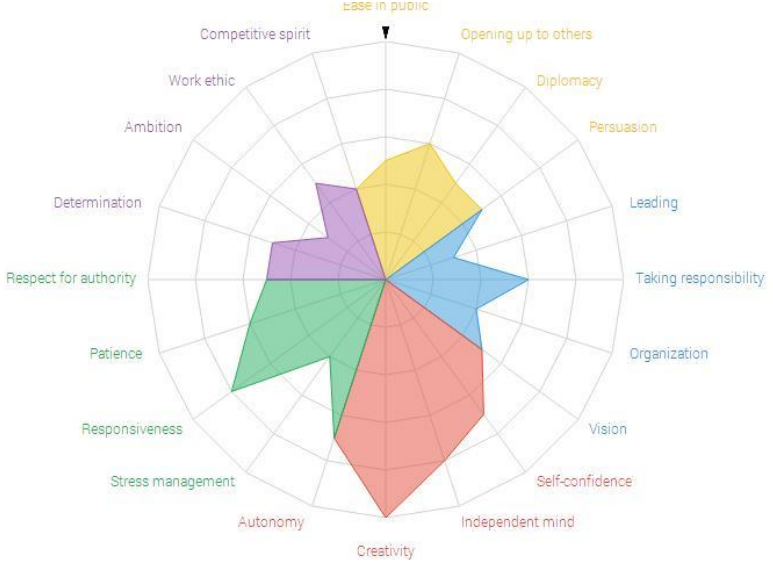
Polar



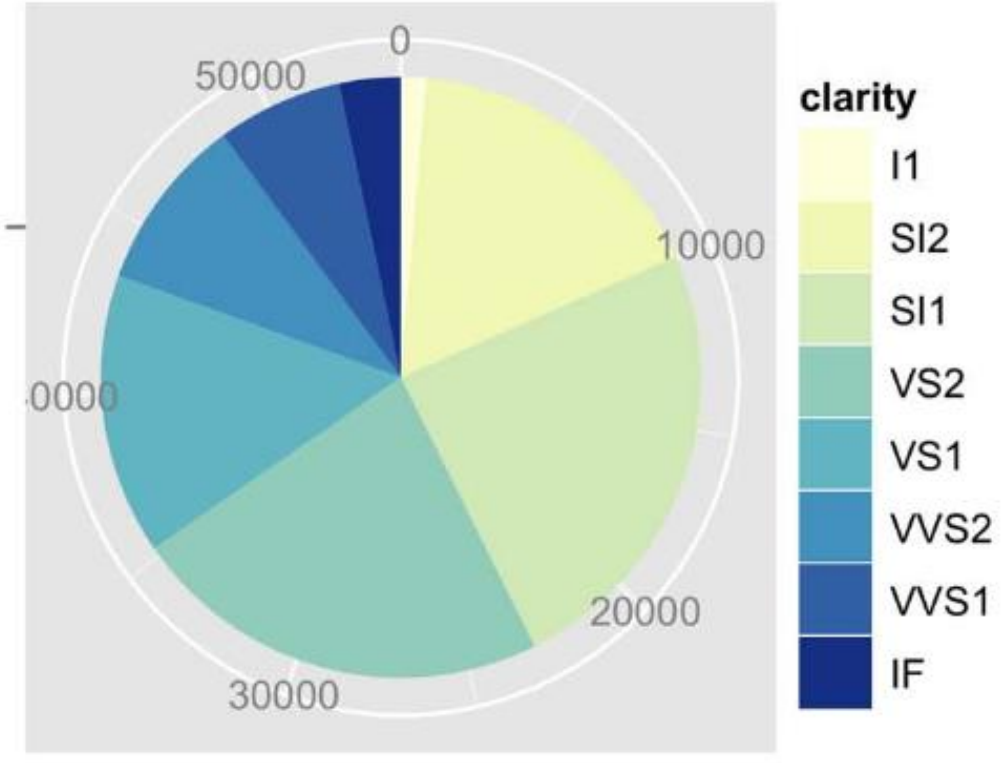
EXAMPLE: STAR PLOT

- = radial line chart

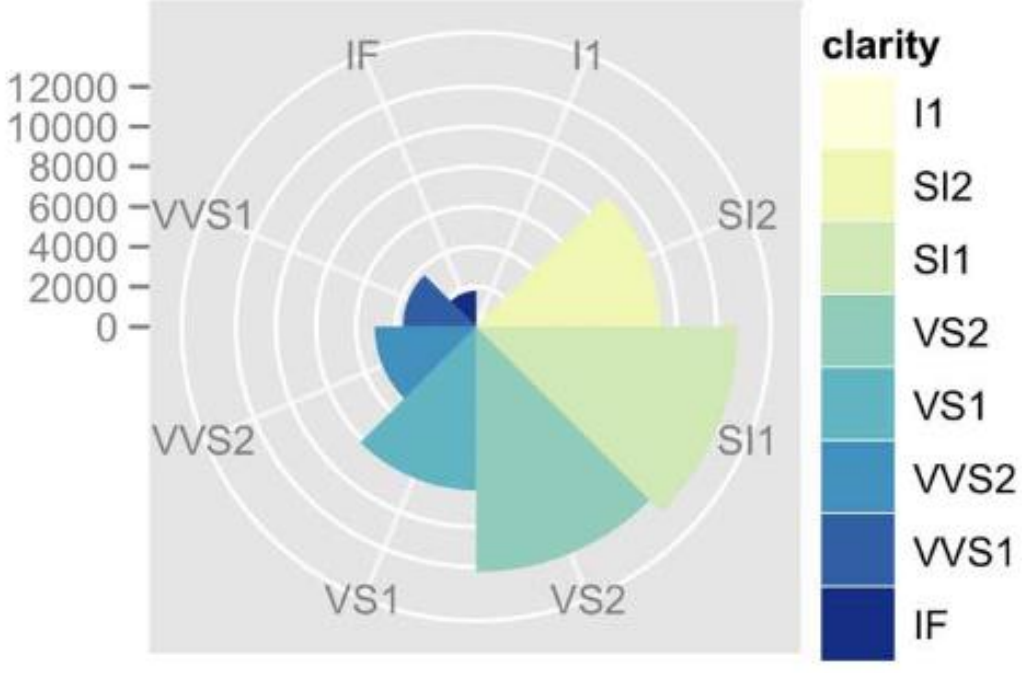




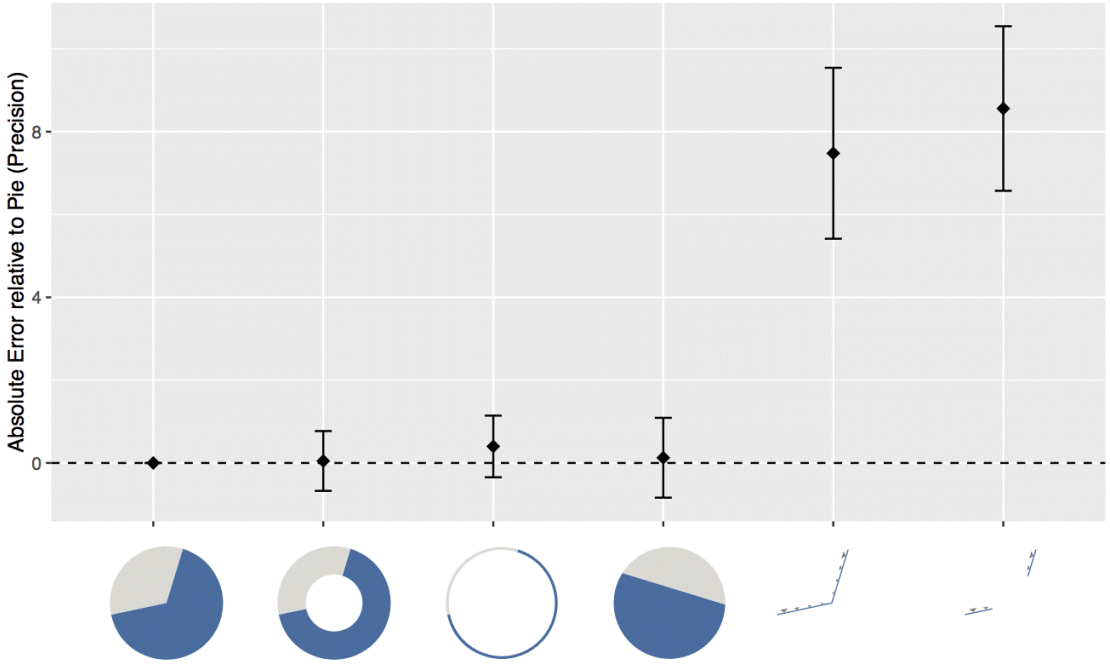
PIE CHARTS


0
10000
20000
30000
50000
clarity
I1
SI2
SI1
VS2
VS1
VVS2
VVS1
IF

POLAR AREA CHARTS

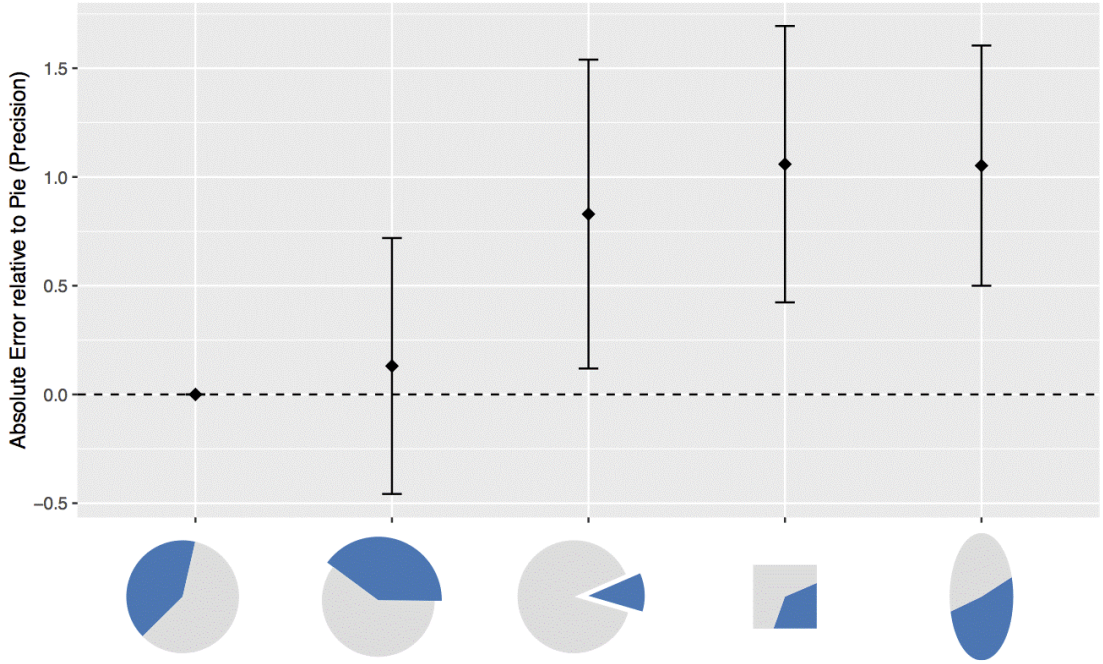


HOW DO PEOPLE READ PIE CHARTS?



https://eagereyes.org/blog/2016/an-illustrated-tour-of-the-pie-chart-study-results

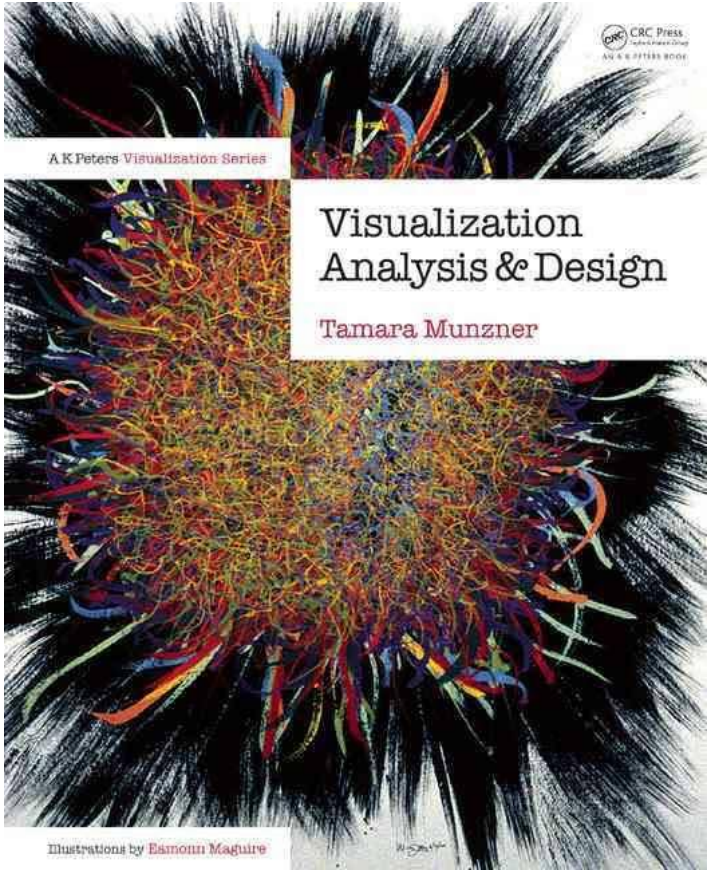
HOW DO PEOPLE READ PIE CHARTS?



https://eagereyes.org/blog/2016/an-illustrated-tour-of-the-pie-chart-study-results

WHAT IT ONE DIMENSION IS TIME?

READINGS


CRC Press
A K Peters Visualization Series
Visualization
Analysis & Design
Tamara Munzner
Illustrations by Eamonn Maguire

ACKNOWLEDGEMENTS

Slides in were inspired and adapted from slides by

- Wesley Willett (University of Calgary)
- Pierre Dragicevic (Inria)
- Uta Hinrichs (University of St. Andrews)