

INTERACTION

SPECTRUM

Static content

Infographics

Books

Dynamic content

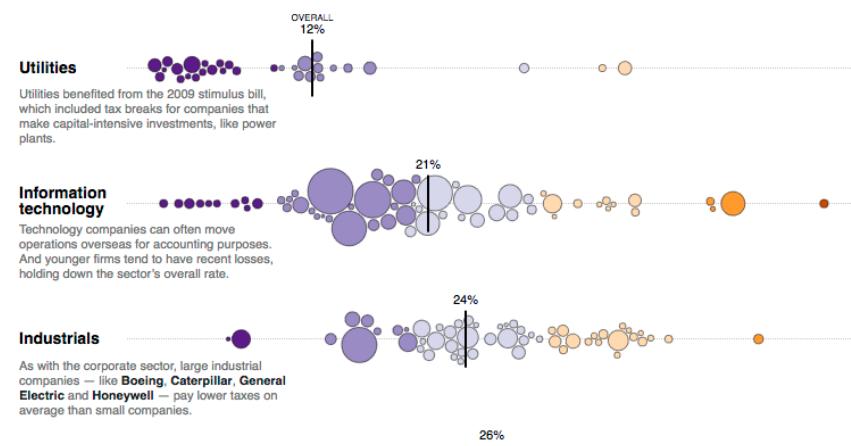
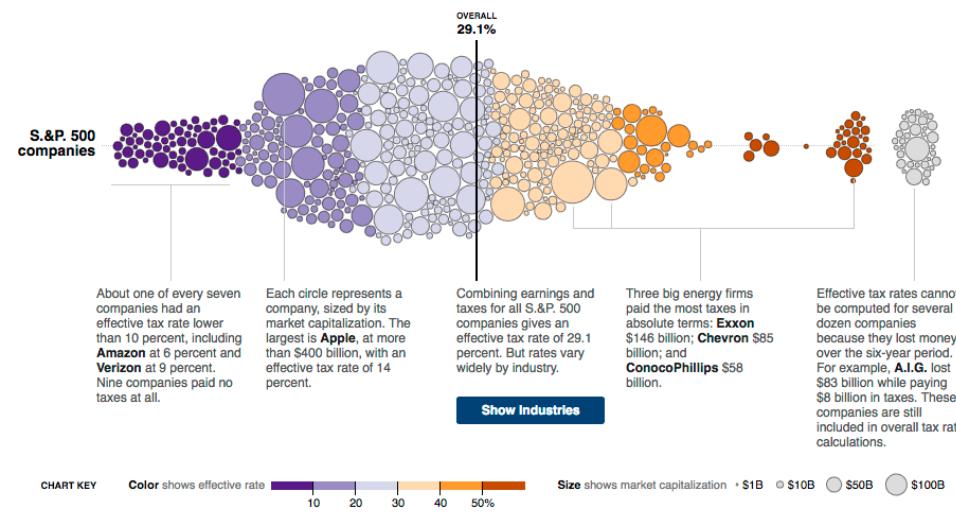
1. Animated content
“Auto-play”, user is not in control
2. Interactive content
Changes are a result of user actions

WHY INTERACT WITH VISUALIZATION?

Need to explore data that is big/complex

Too much data

Too many ways to show it

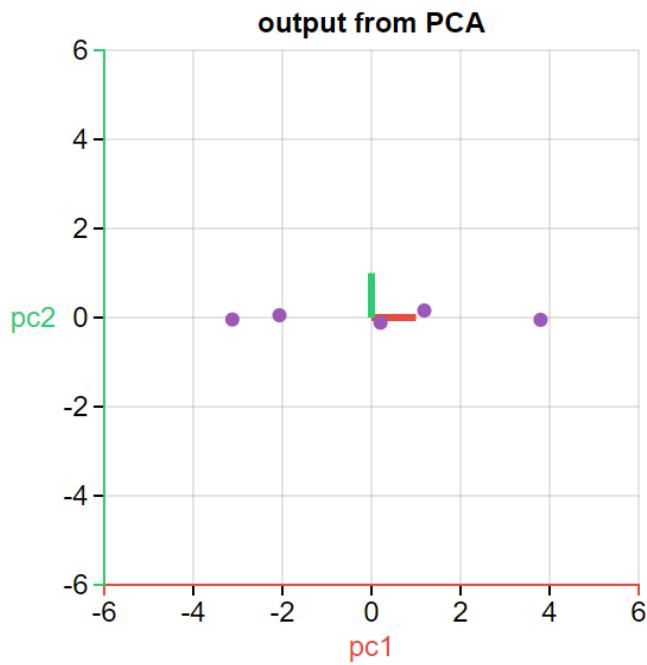
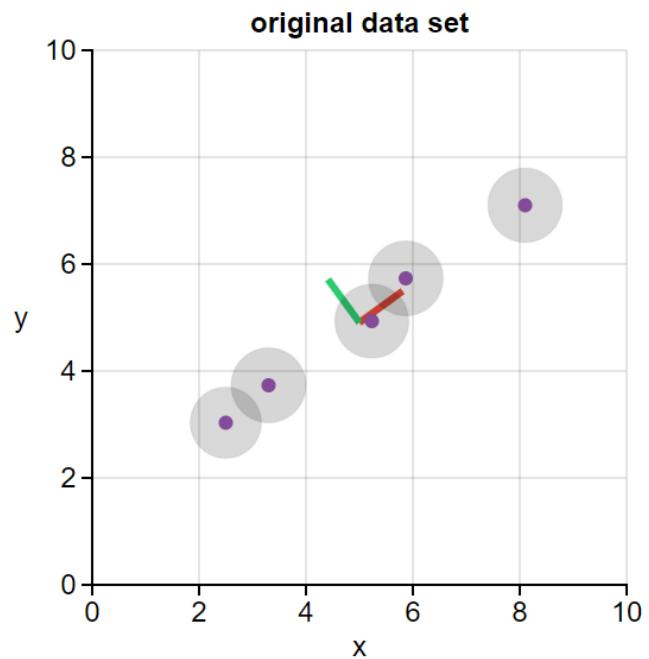


WHY INTERACTION WITH VIS?

Interaction amplifies cognition

Understand things better if we can touch them

When we can observe cause and effect



INTERACTION METHODS

What do you design
for?

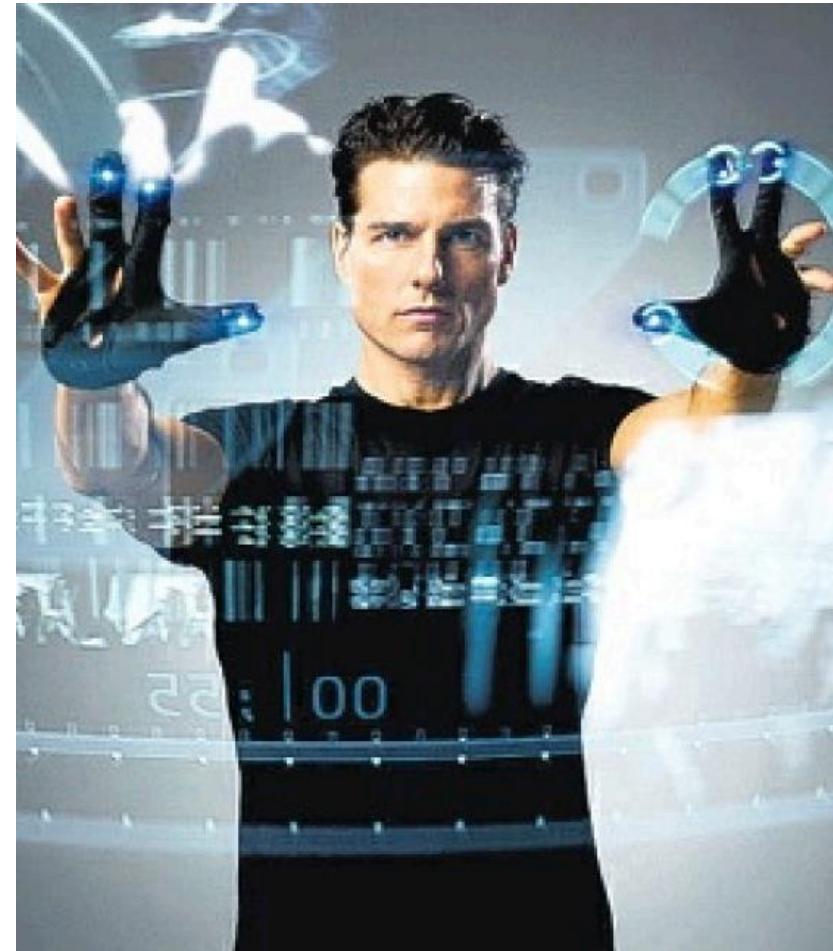
Mouse, keyboard?

Touch interaction / mobile?

Gestures?

Eye movement?

Speech?

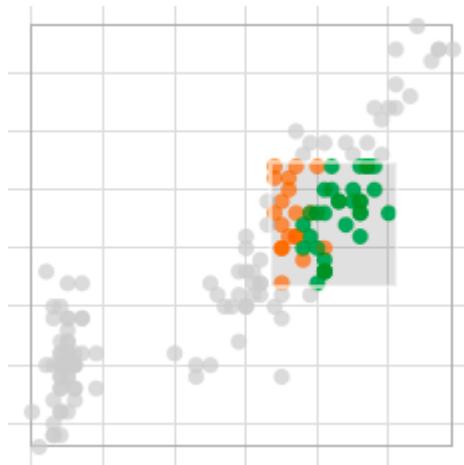


DIRECT MANIPULATION

Interact directly with objects

Continuous feedback/updates

Indirect interact
Using a query, a slides, etc.



TYPES OF INTERACTION

Single view

Change over time

Navigation

Semantic zooming

Filtering and Querying

Focus + Context

Multiple views

Selection (Details on demand)

Linking & brushing

Adapting representation

PURPOSES OF INTERACTION

DOI:10.1145/2133806.2133821

Article development led by ACM Queue
queue.acm.org

A taxonomy of tools that support the fluent
and flexible use of visualizations.

BY JEFFREY HEER AND BEN SHNEIDERMAN

Interactive Dynamics for Visual Analysis

TABLE 1: Taxonomy of interactive dynamics for visual analysis

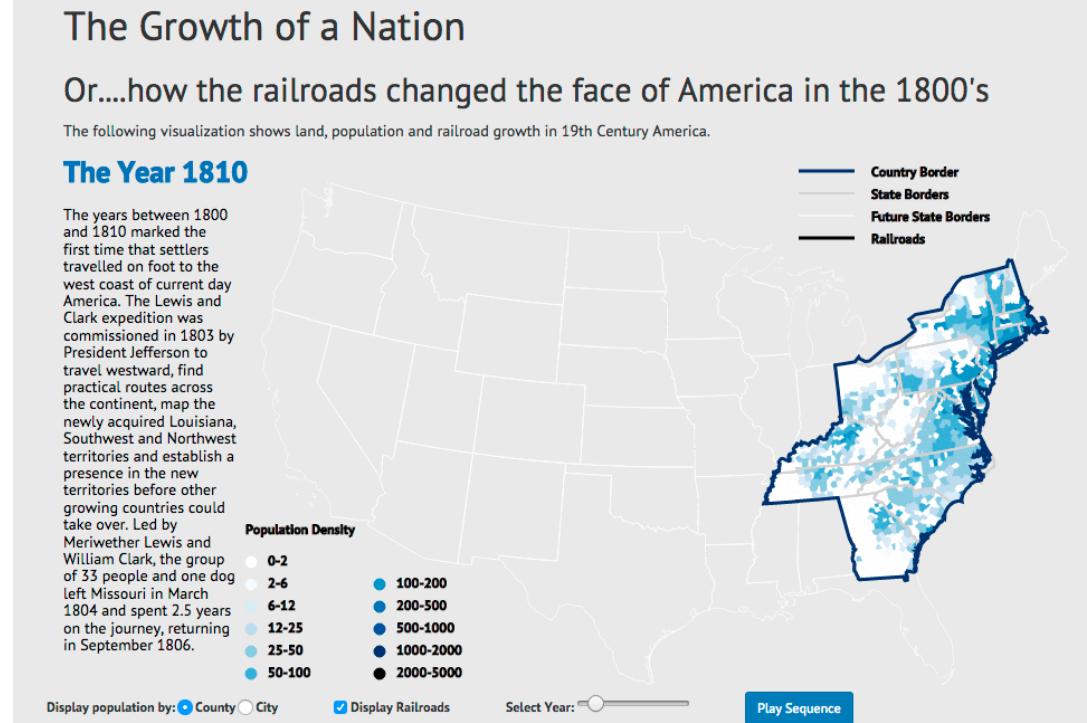
Data & View Specification	Visualize data by choosing visual encodings. Filter out data to focus on relevant items. Sort items to expose patterns. Derive values or models from source data.
View Manipulation	Select items to highlight, filter, or manipulate them. Navigate to examine high-level patterns and low-level detail. Coordinate views for linked, multi-dimensional exploration. Organize multiple windows and workspaces.
Process & Provenance	Record analysis histories for revisit, review and sharing. Annotate patterns to document findings. Share views and annotations to enable collaboration. Guide users through analysis tasks or stories.

CHANGE OVER TIME/TRANSITIONS

CHANGE OVER TIME

Use slides to see view with data at different times

Sometimes better to show difference explicitly

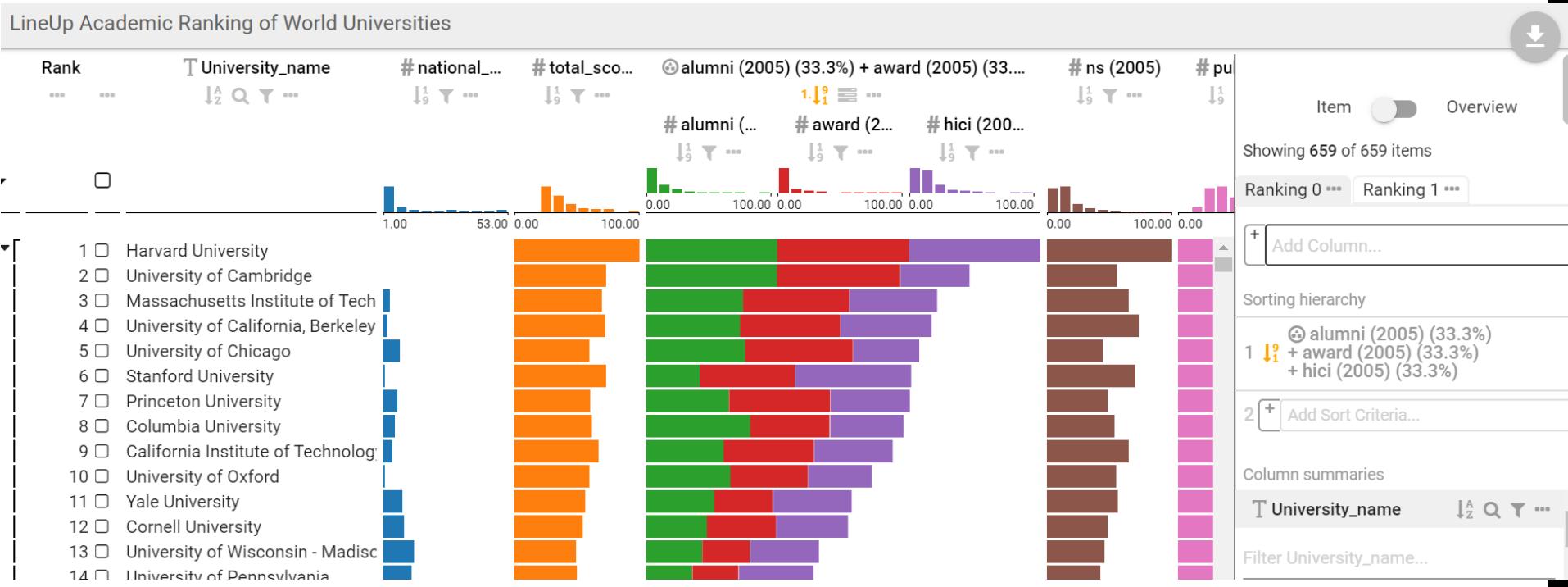


CHANGE OVER TIME

Doesn't have to be literal time:

change as you go

as part of an analysis process



<https://lineup.js.org/app/#shanghai>

<https://caleydo.org/tools/lineup/>

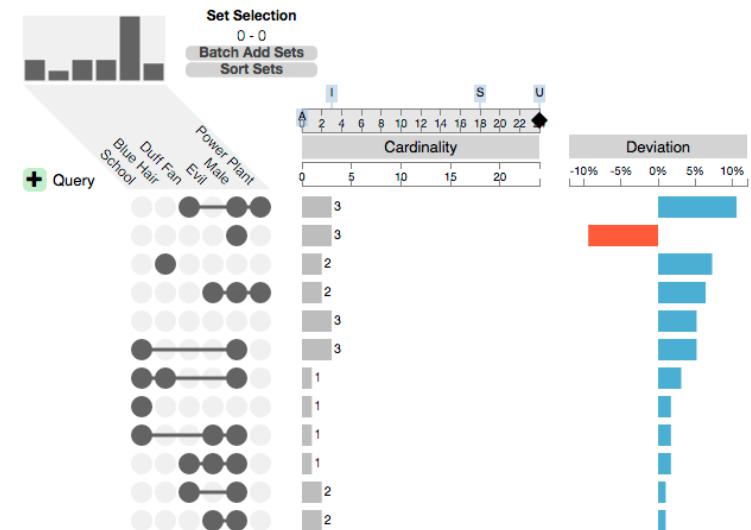
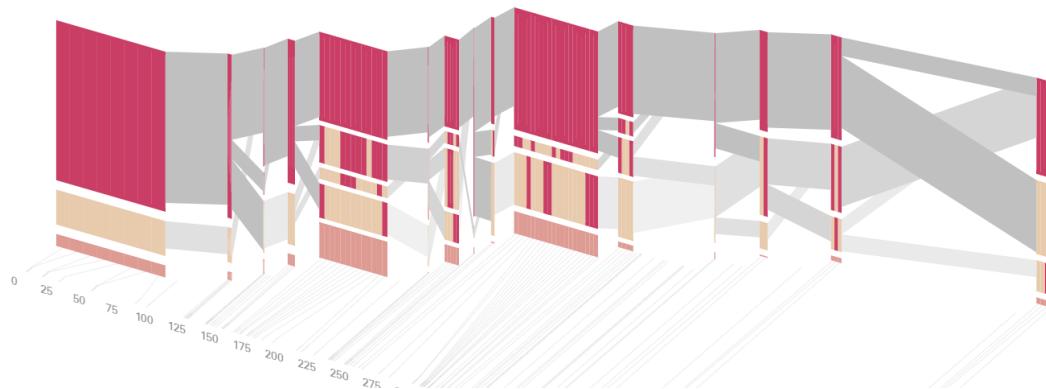
WHY TRANSITION?

Different representations support different tasks

Bar chart vs stacked bar chart

Change ordering

Transition make it possible for users to track what is going on

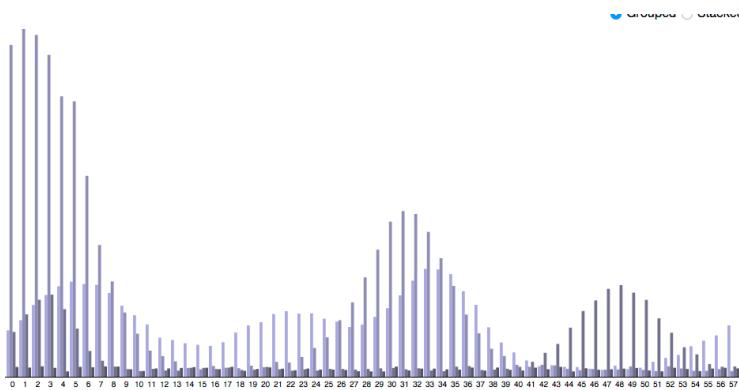


<https://benfry.com/isometricblocks/>

<http://vcg.github.io/upset/>

ANIMATED TRANSITIONS

Smooth interpolation between states or visualization techniques

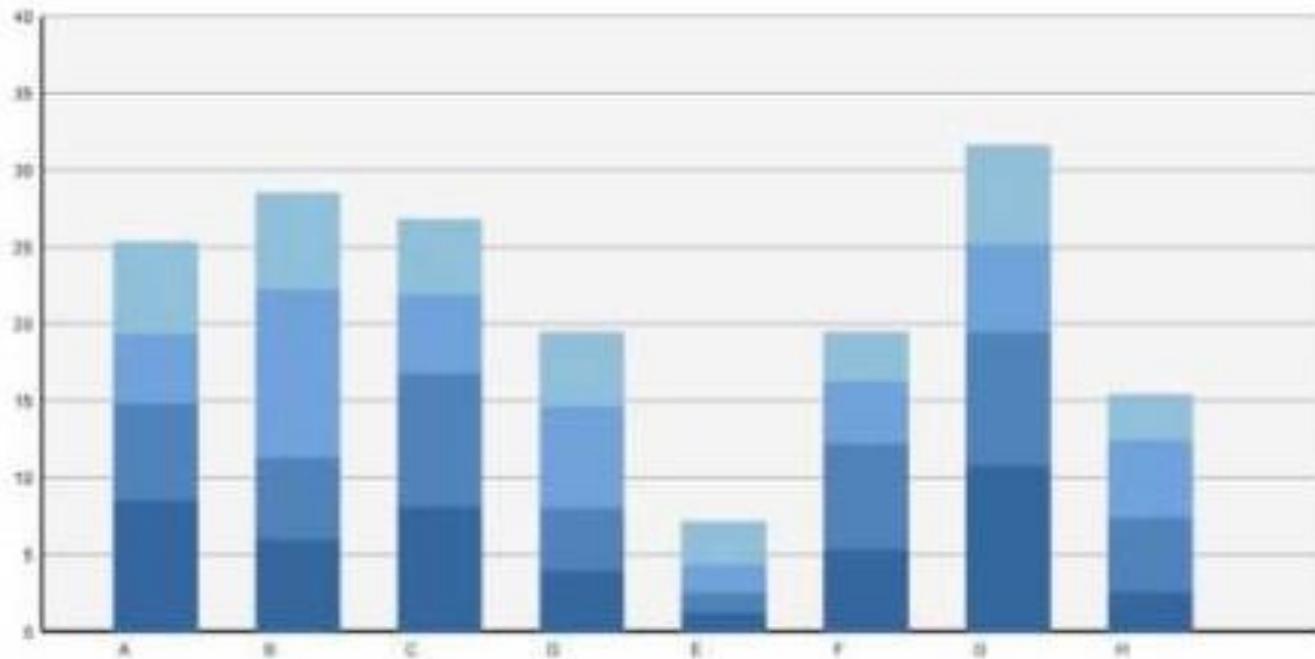


<http://bl.ocks.org/mbostock/3943967>



<https://observablehq.com/@d3/zoomable-sunburst>

WHY ANIMATED TRANSITION?



Animated Transitions in Statistical Data Graphics

Jeffrey Heer, George G. Robertson

Abstract—In this paper we investigate the effectiveness of animated transitions between common statistical data graphics such as bar charts, pie charts, and scatter plots. We extend theoretical models of data graphics to include such transitions, introducing a taxonomy of transition types. We then propose design principles for creating effective transitions and illustrate the application of these principles in DynaVis, a visualization system featuring animated data graphics. Two controlled experiments were conducted

<https://youtu.be/vLk7mlAtEXI>

ANIMATION CAVEATS (WARNING)

Changes can be hard to track

Eyes over memory

Show all states in multiple views

NAVIGATION

NAVIGATION

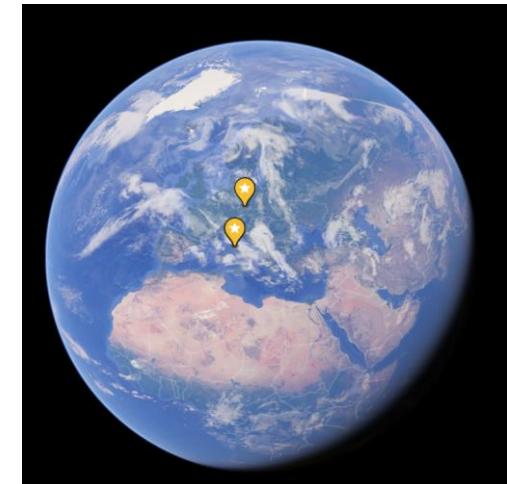
Pan

Move around

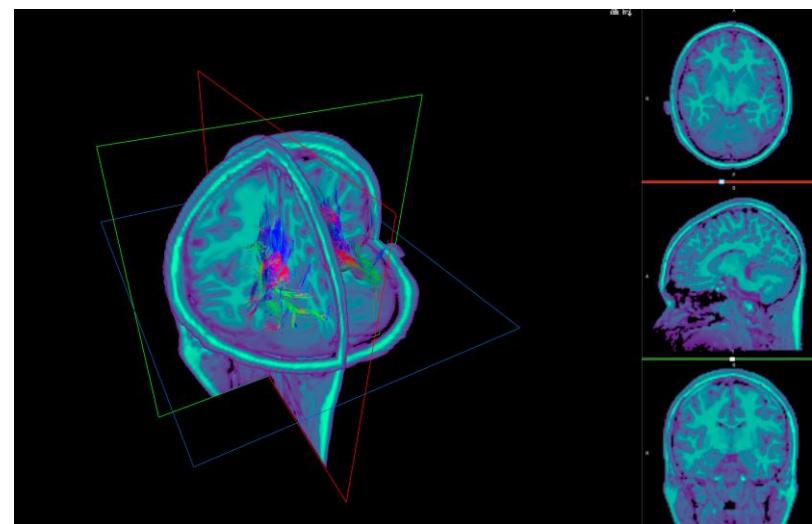
Zoom

Enlarge/make smaller

Rotate



<https://www.google.com/maps/@40.5388417,19.7863469,14278660m/data=!3m1!1e3>



<http://x.babymri.org/example1/?scene=http://x.babymri.org/example1/scene.json>

SCROLLTELLING

Telling an interactive story

Interaction by scrolling

Nice but

Continuous scrolling vs discrete states

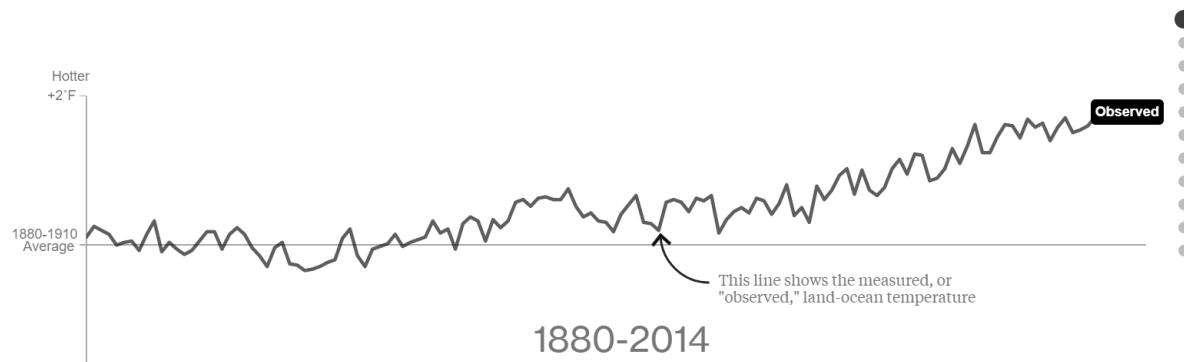
Direct access

Unexpected behavior

What's Really Warming the World?

By Eric Roston  and Blacki Migliozzi  | June 24, 2015

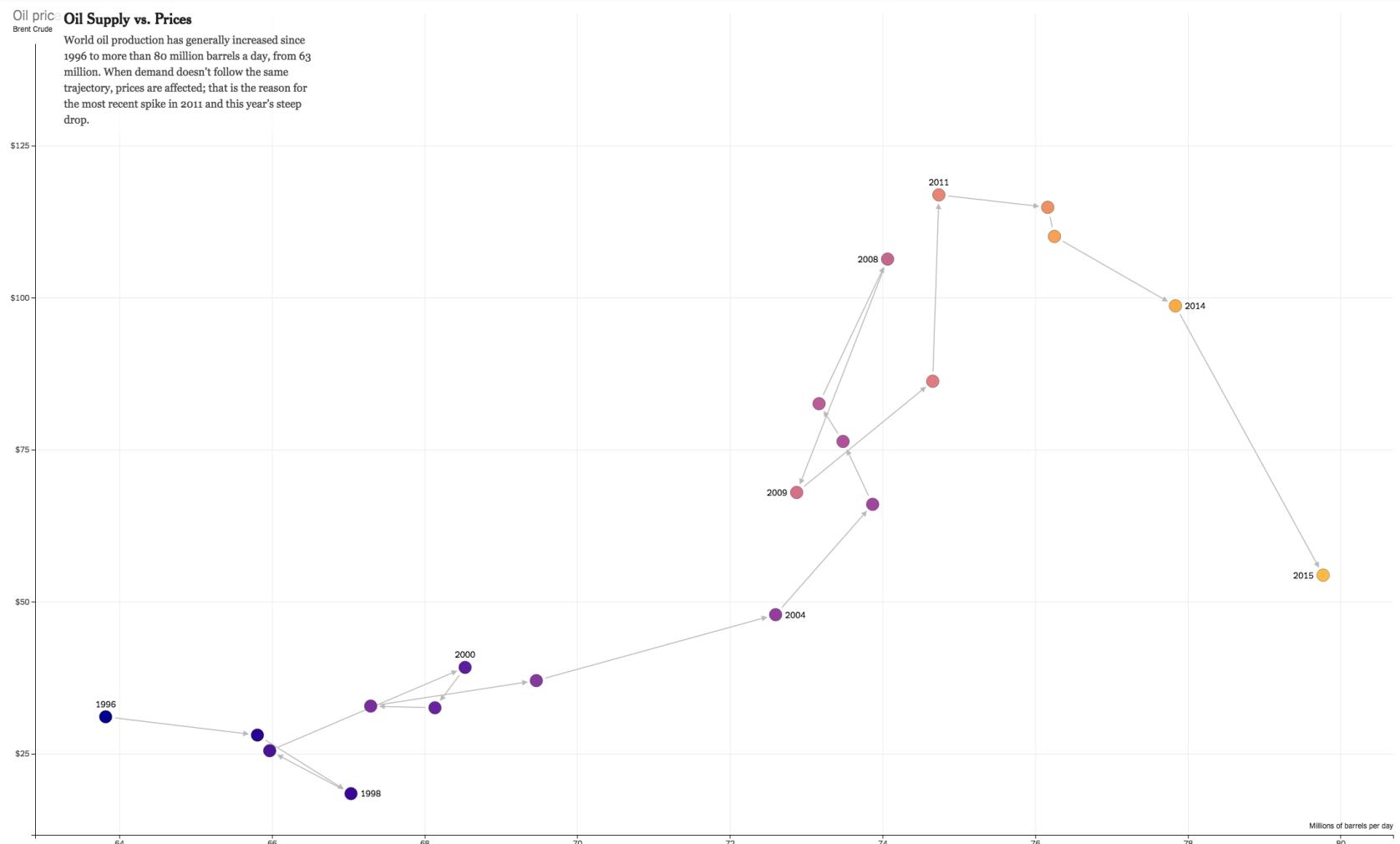
Skeptics of manmade climate change offer various natural causes to explain why the Earth has warmed 1.4 degrees Fahrenheit since 1880. But can these account for the planet's rising temperature? Scroll down to see how much different factors, both natural and industrial, contribute to global warming, based on findings from NASA's Goddard Institute for Space Studies.



<https://www.bloomberg.com/graphics/2015-whats-warming-the-world/>

<https://eagereyes.org/blog/2016/the-scrollytelling-scourge>

SCROLLTELLING: OIL PRICES

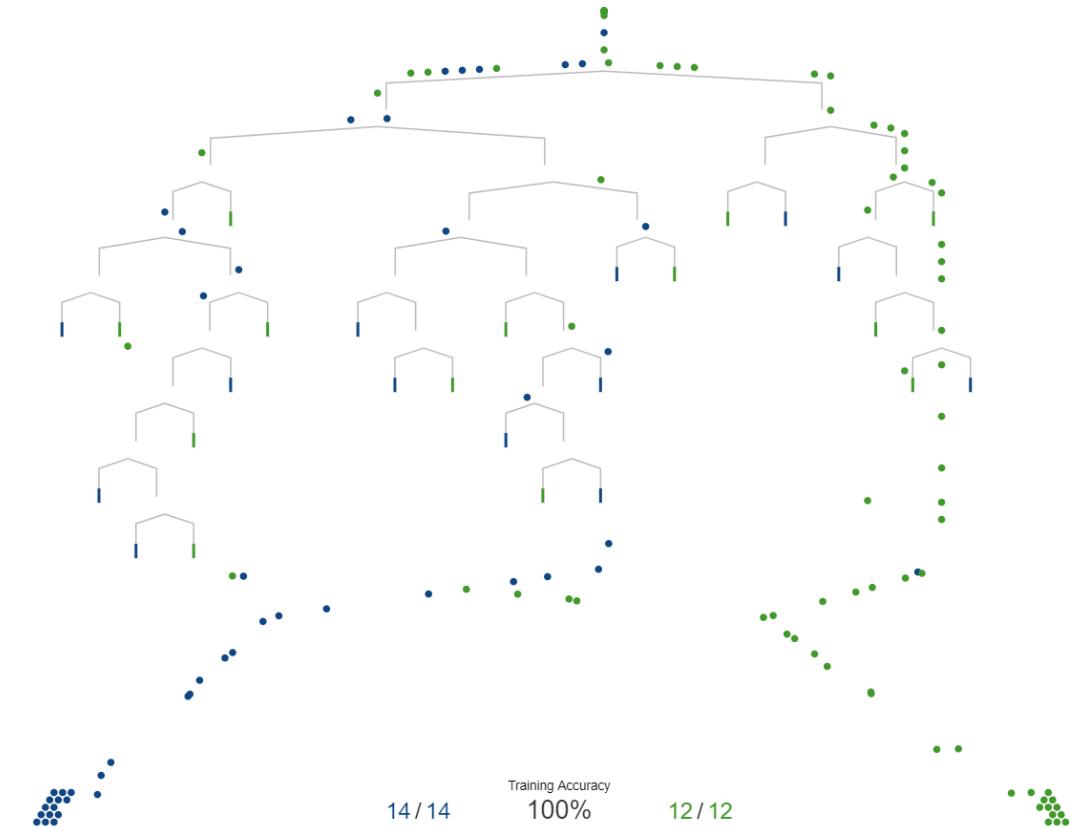


http://www.nytimes.com/interactive/2015/09/30/business/how-the-us-and-opec-drive-oil-prices.html?_r=0

SCROLLTELLING: INTRO TO ML

Making predictions

The newly-trained decision tree model determines whether a home is in San Francisco or New York by running each data point through the branches.



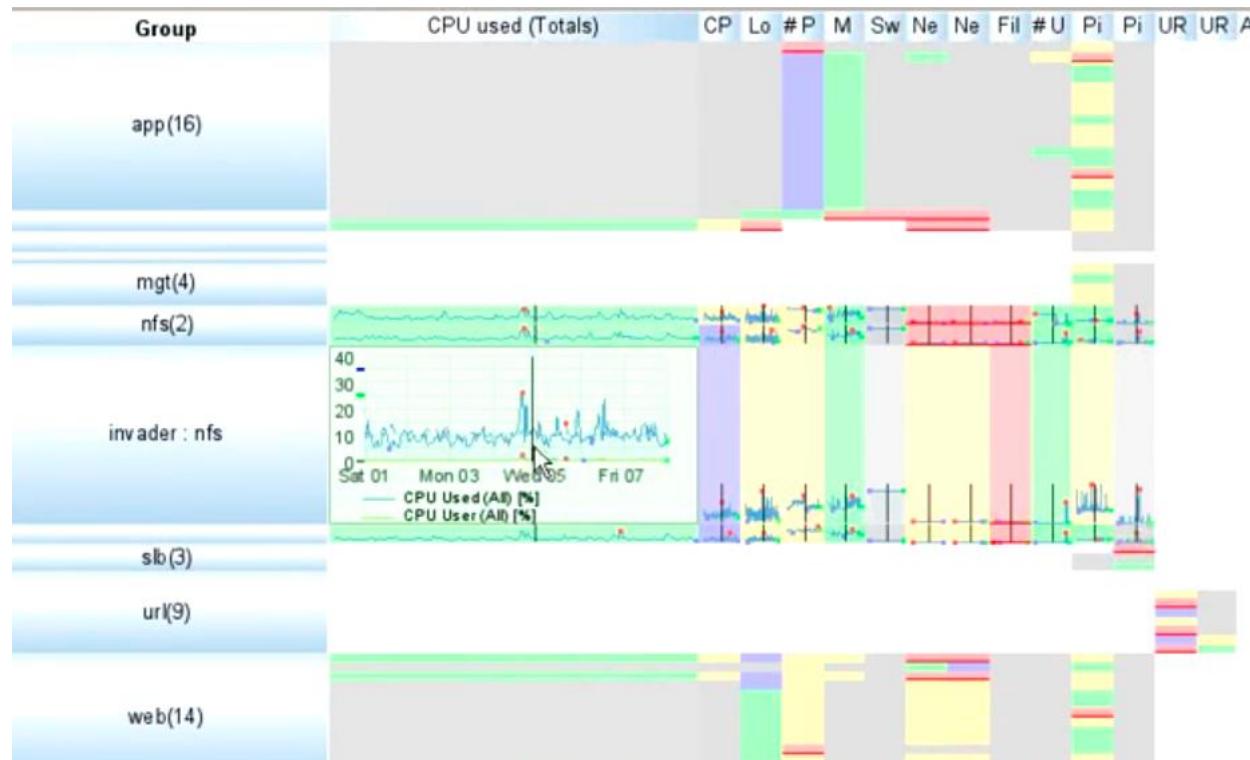
SEMANTIC ZOOMING

SEMANTIC ZOOMING

Update content on zooming

More detail as more space becomes available

Ideally readable at multiple resolutions



<https://youtu.be/lD0c3H0VSkw>

**FOCUS +
CONTEXT**

FOCUS + CONTEXT

Carefully pick what to show

Hint at what you are not showing

FOCUS + CONTEXT

Synthesis of **visual encoding** and **interaction**

User selects regions of interest (focus) through navigation or selection

Provide context through

aggregation

reduction

layering

④ Embed

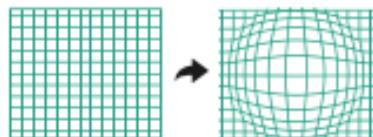
→ Elide Data



→ Superimpose Layer



→ Distort Geometry

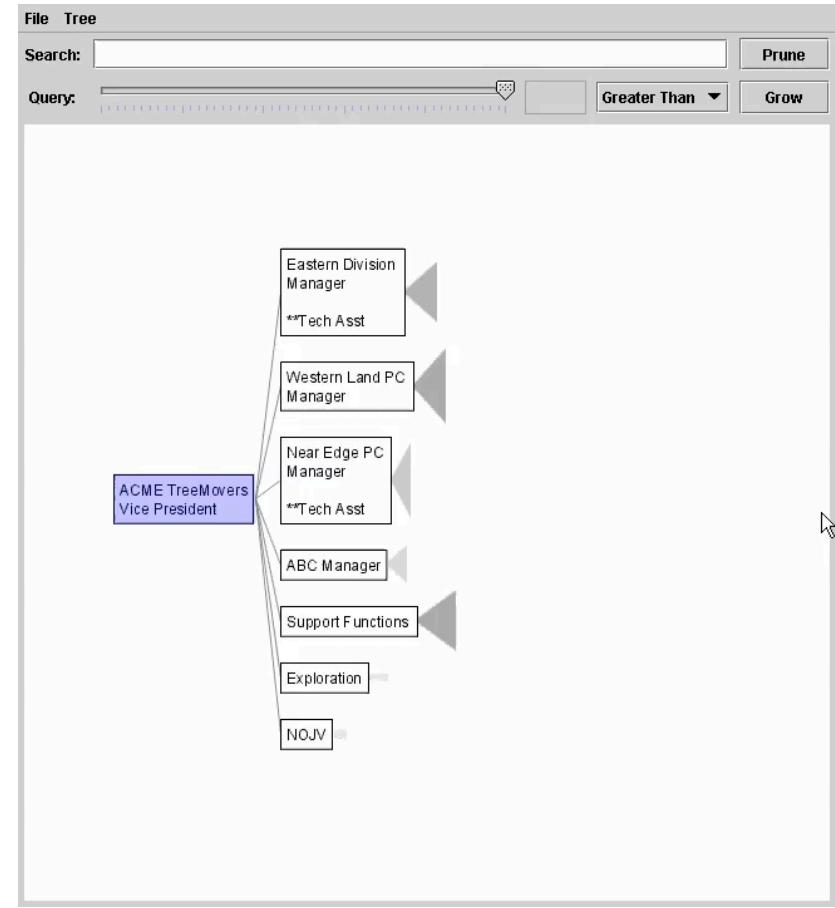


ELISION

Focus item shown in detail

Other items summarized for context

Example: SpaceTree



<https://www.cs.umd.edu/hcil/spacetree/>

<https://youtu.be/F7-vJBsFBWw>

DEGREE OF INTEREST (DOI)

Represent objects in the neighborhood in detail, and only major landmarks far away

Balance between local detail and global context

$$\text{DOI}(x) = I(x) - D(x,y)$$

$I(x)$: interest in object x

D – a distance function to the current focus y of x

There may have many foci

DOI TREE

Interactive tree with animated transitions that fit within a bounded region of space layout depends on the user's estimated DOI

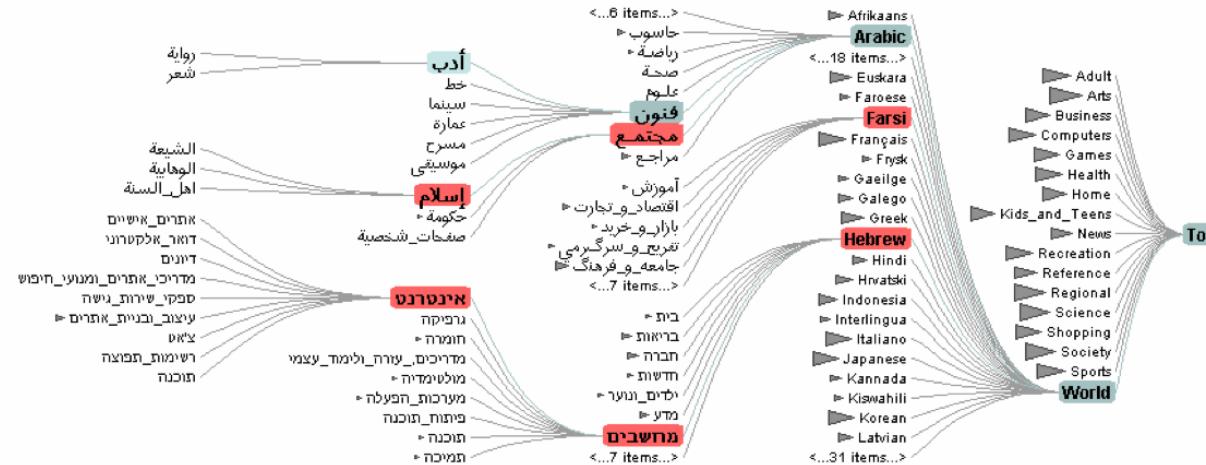
USE:

logical filtering based on DOI

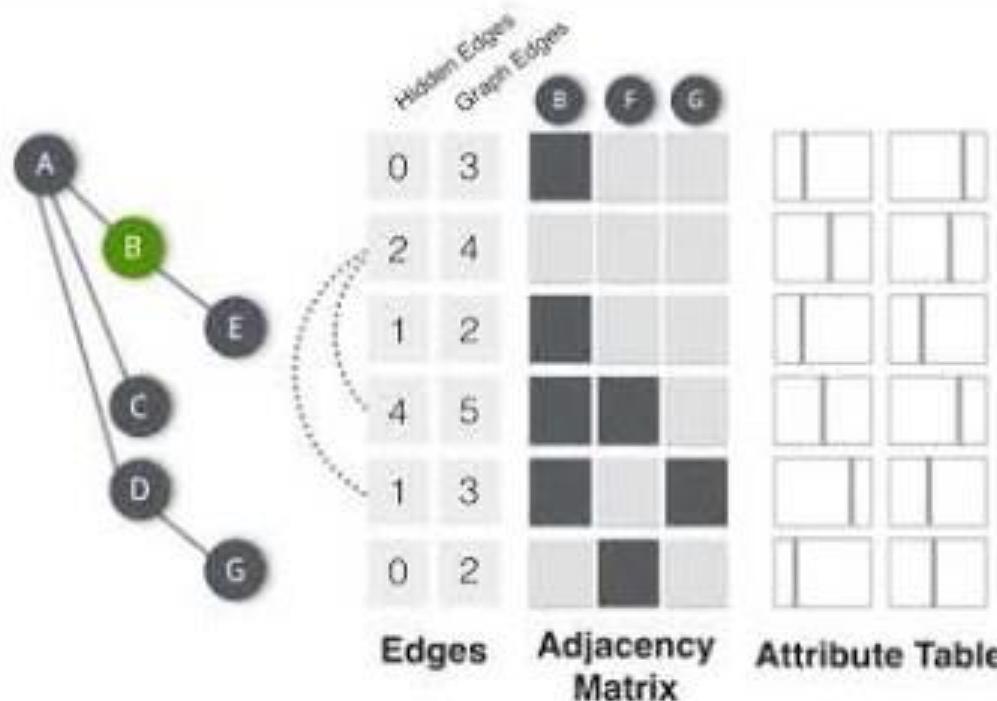
Geometric distortion of node size based on DOI

Semantic zooming on content based on node size

Aggregate representations of elided subtrees



DOI: JUNIPER EXAMPLE



<https://youtu.be/EAjNxFgsJ58>

SUPERIMPOSE

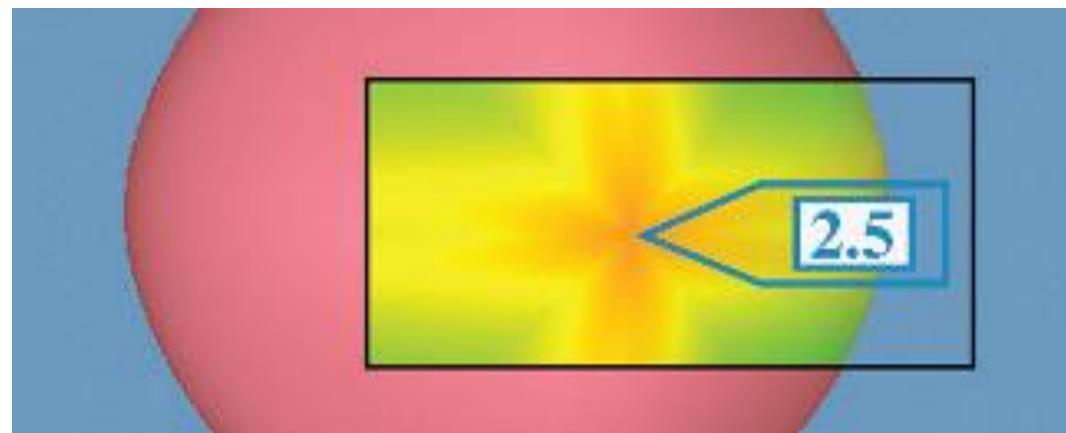
Focus layer limited to a local region of view

Instead of stretching across the entire view

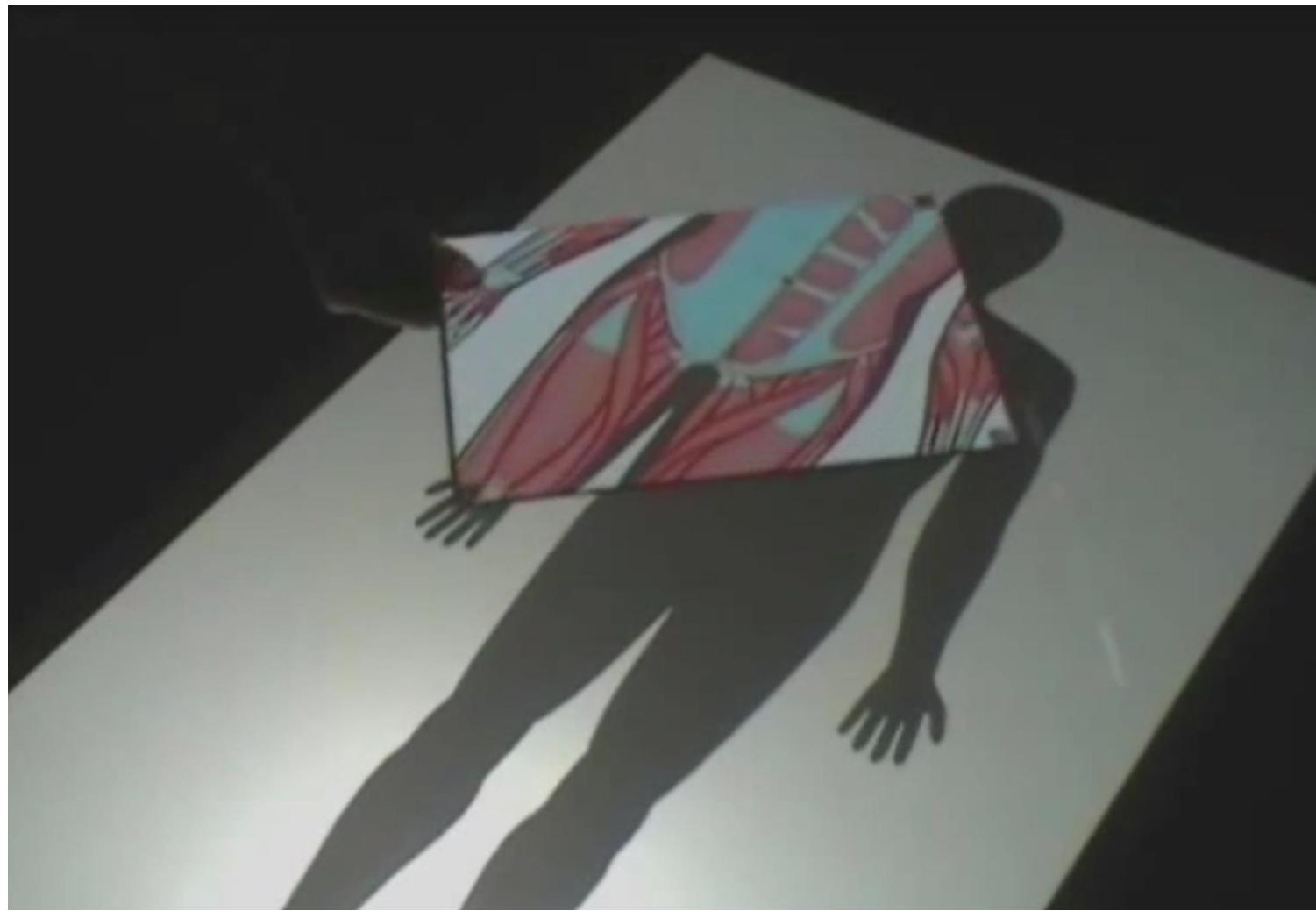
TOOLGLASS & MAGIC LENSES

Magic lense:

details/different data is shown
when moving a lens
over a scene

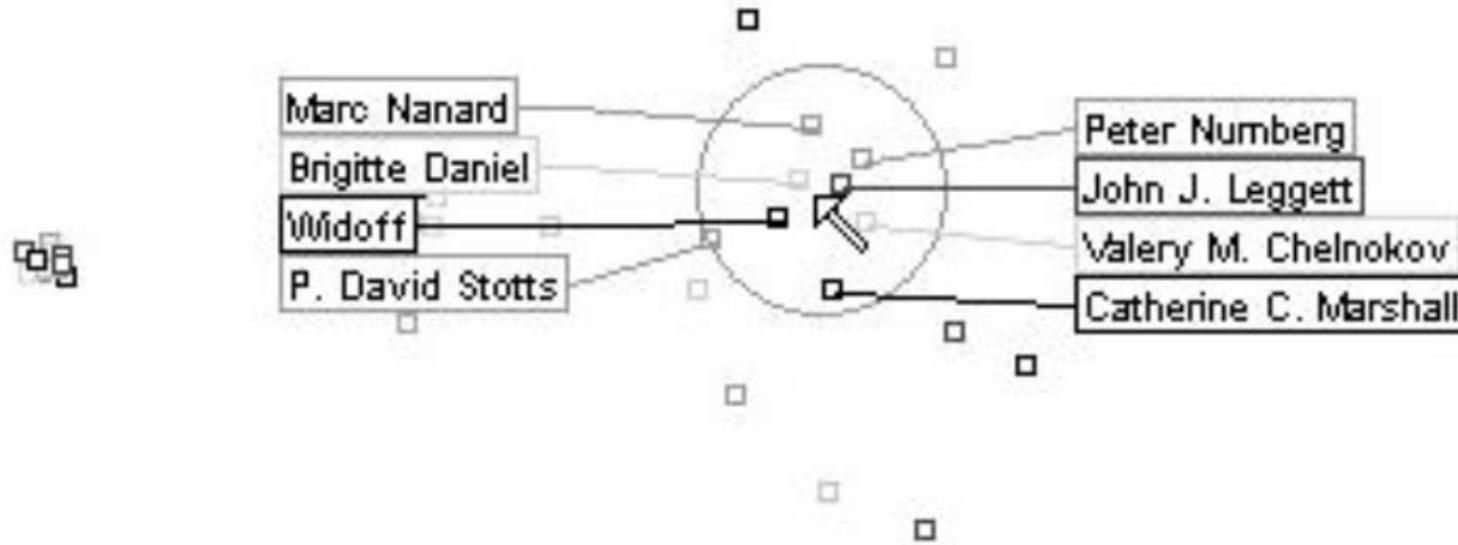


MAGIC LENSE



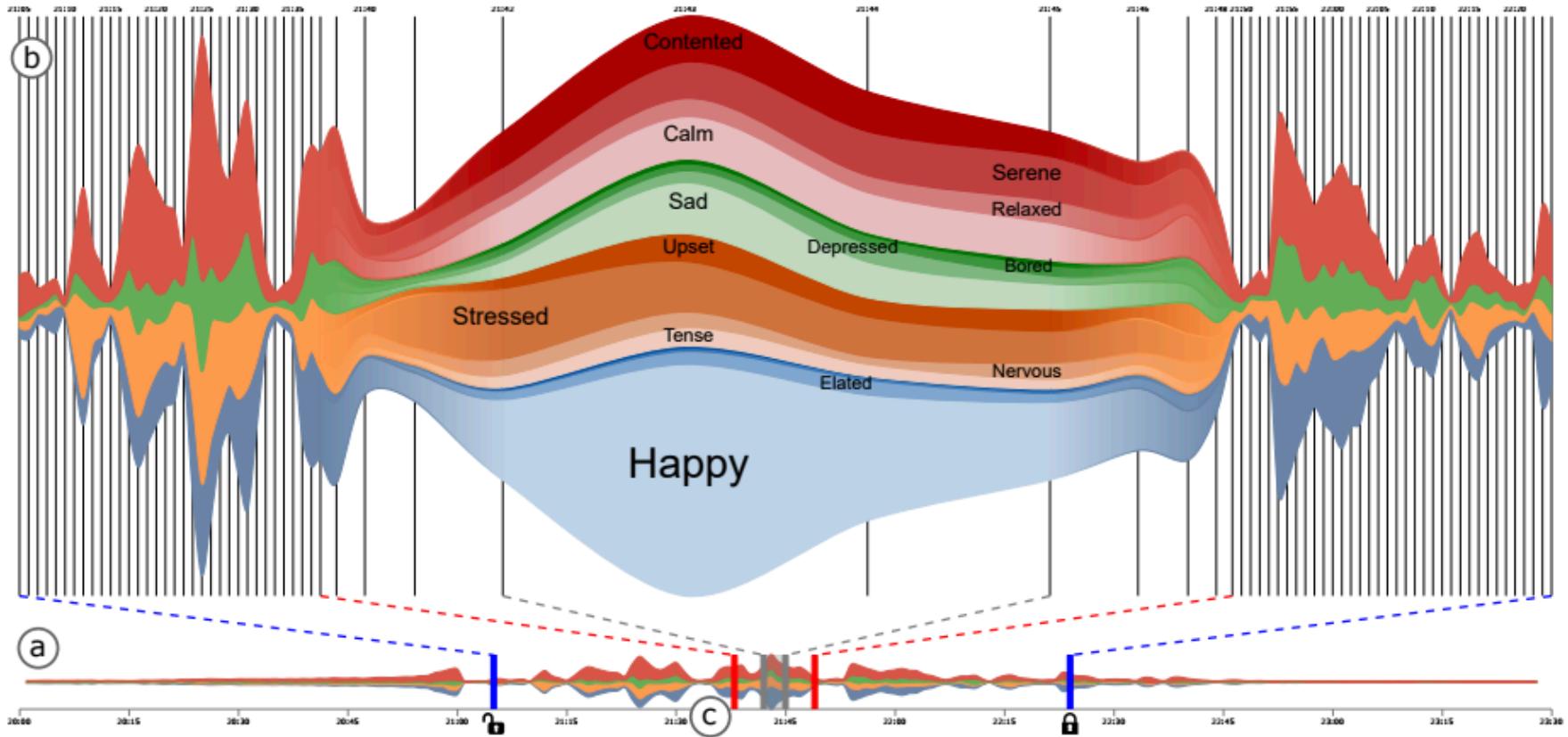
https://www.youtube.com/watch?v=h-mF4_OAhU0

MAGIC LENSE: LABELING



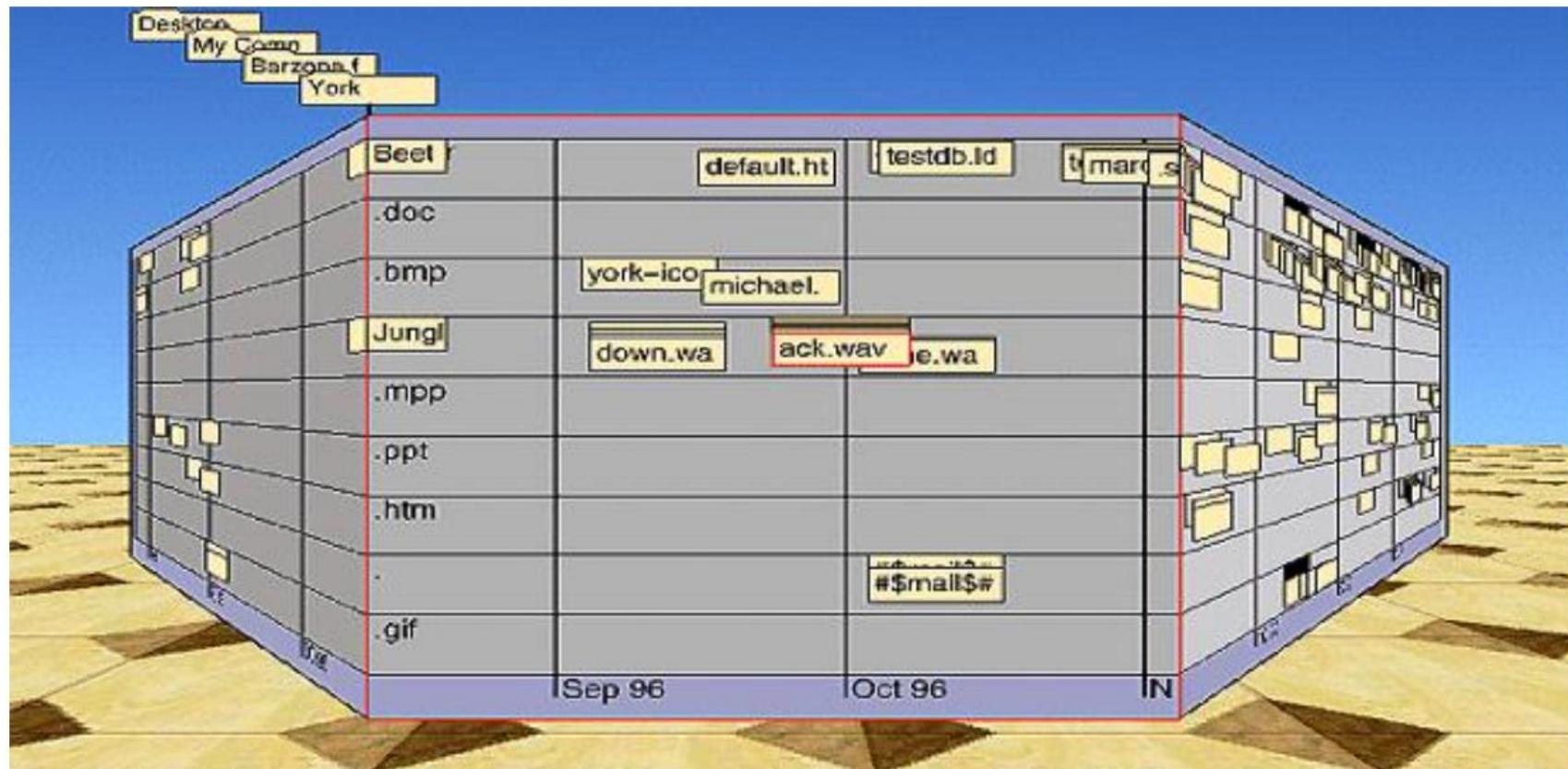
DISTORTION

Use geometric distortion of the contextual regions to make room for the details in the focus regions(s)



<http://advanse.lirmm.fr/multistream/>

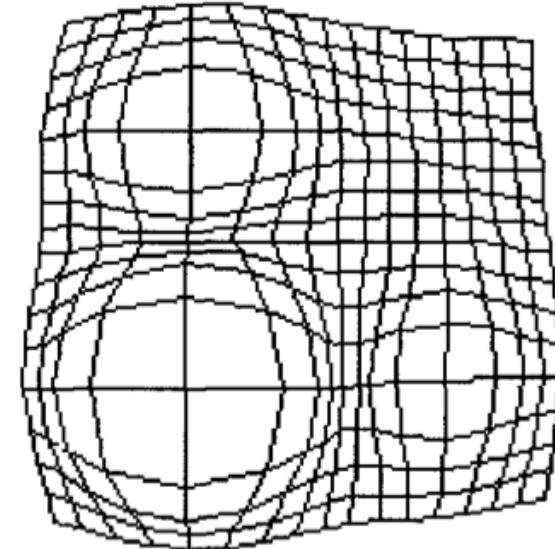
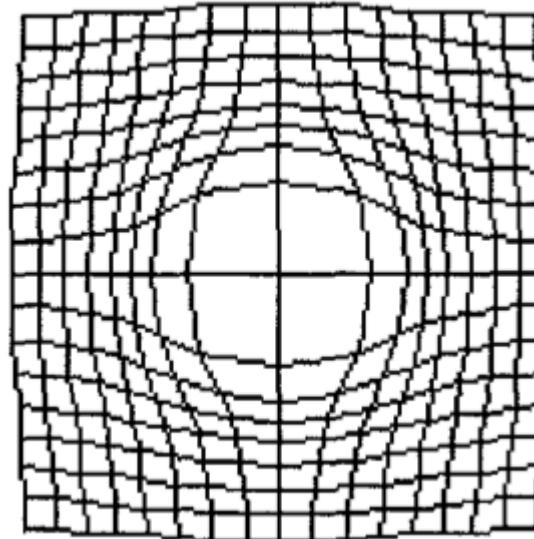
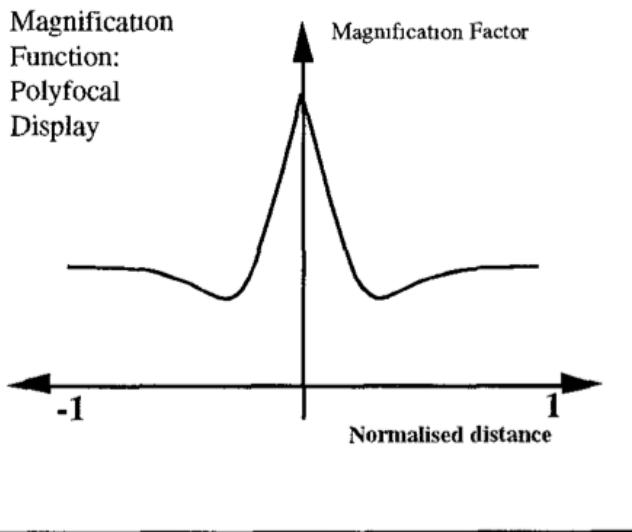
PERSPECTIVE WALL



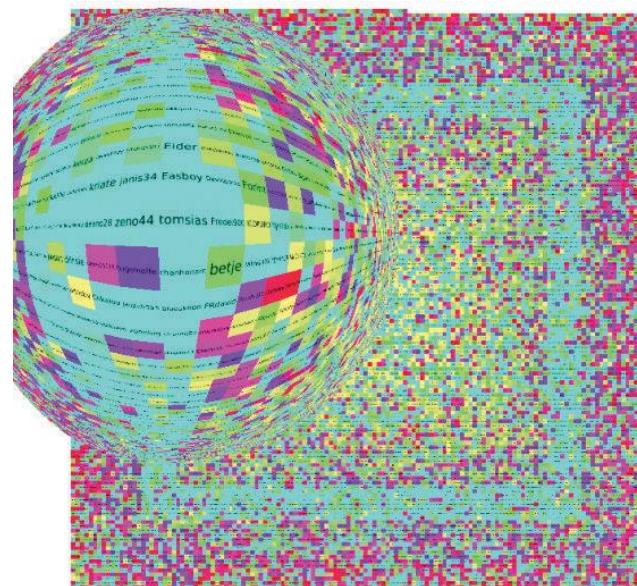
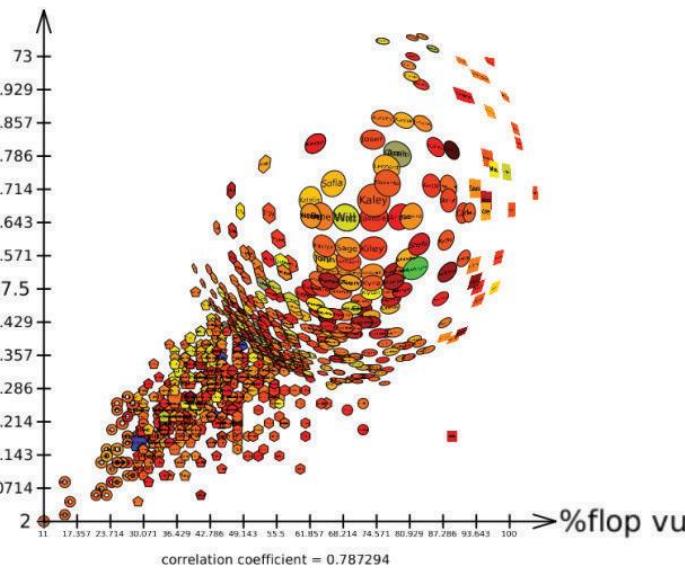
<https://www.youtube.com/watch?v=hYUZbrWtCZg>

FISHEYE

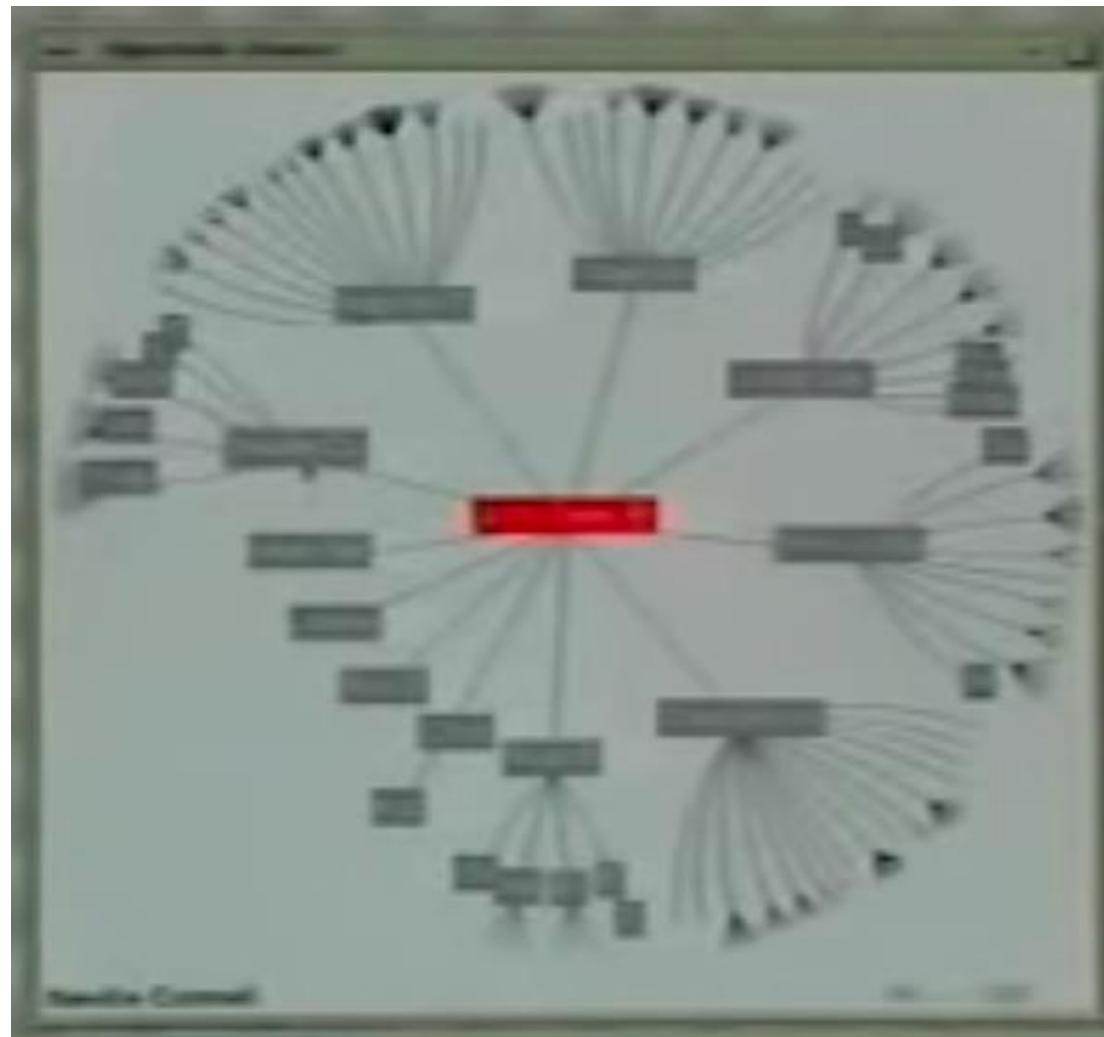
Magnification
Function:
Polyfocal
Display



%show vu

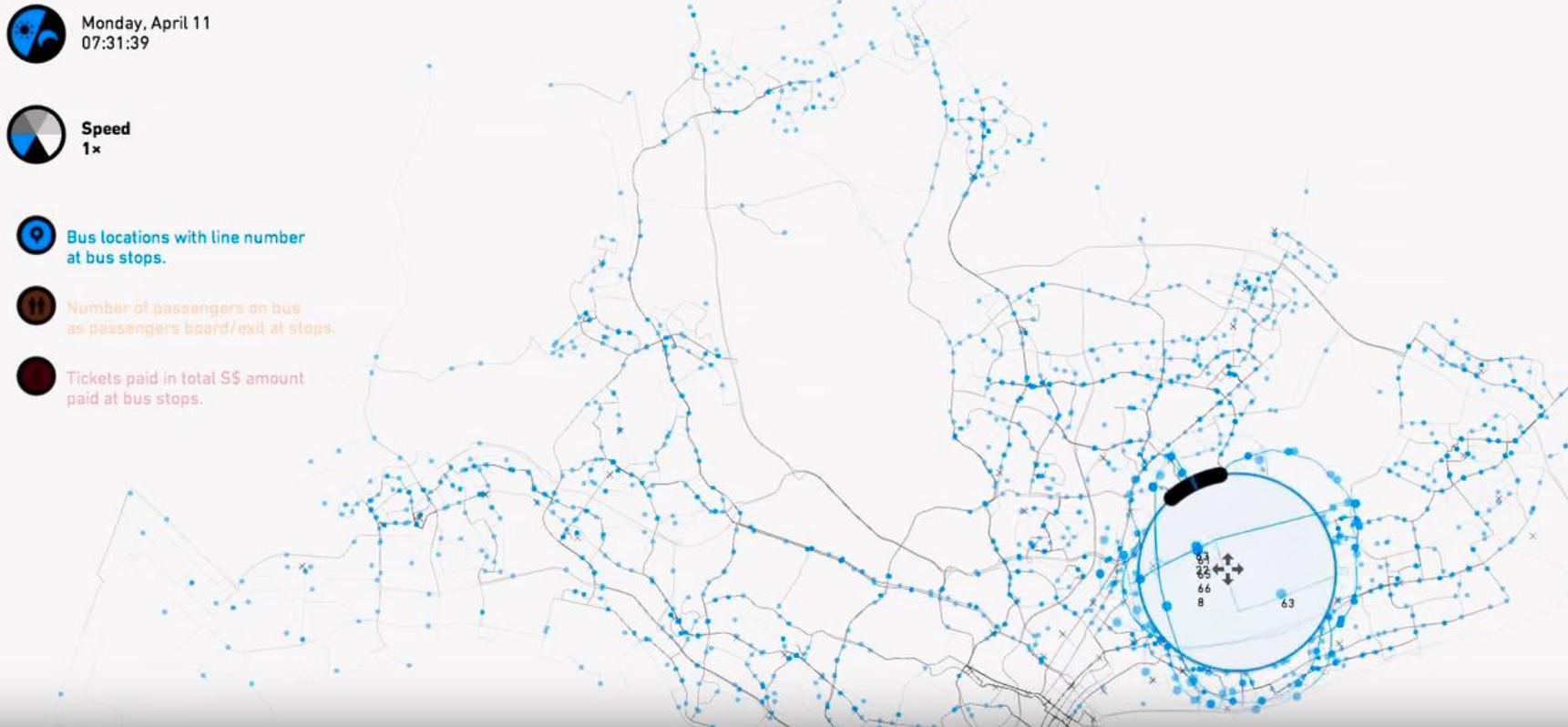


HYPERBOLIC GEOMETRY

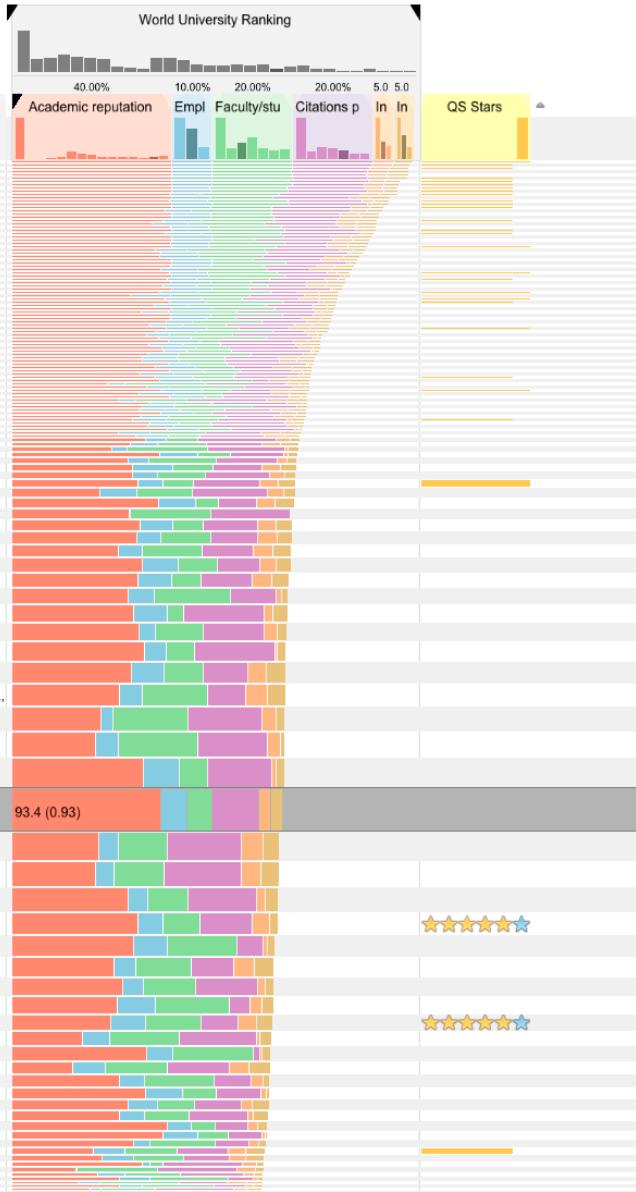


<https://www.youtube.com/watch?v=pwpze3RF55o>

EXPLORING PUBLIC TRANSIT -BUSES AT BUS STOPS



<http://pmcruz.com/information-visualization/data-lenses>



Fisheye Tree View



ctominski

 [Subscribe](#)

100 views

Add to Share More

<https://www.youtube.com/watch?v=mTf7bPvdeJo>

DISTORTION CONCERNS

Unsuitable for relative spatial judgements

Overhead of tracking distortion

Visual communication of distortion

gridlines, shading

Target acquisition problem

lens displacing items away from screen location

Mixed result compared to separate views and temporal navigation

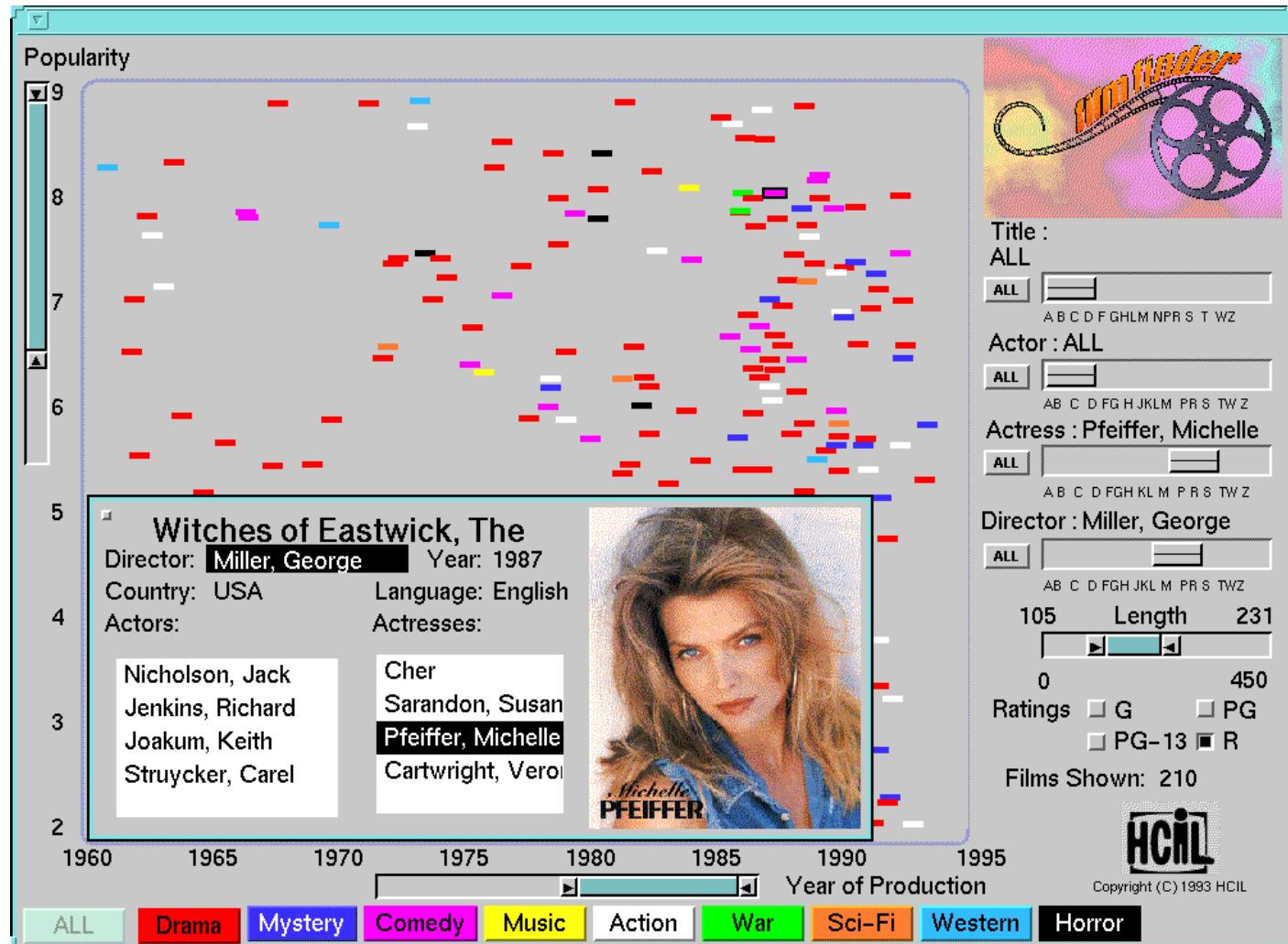
OVERVIEW + DETAIL

OVERVIEW AND DETAIL

One view shows overview

Other show detail





[FilmFinder, Ahlberg & Shneiderman, 1994]

FILTERING & DYNAMIC QUERYING

MANTRA

Visual information seeking matra (Shneiderman, 1996)

Overview first,

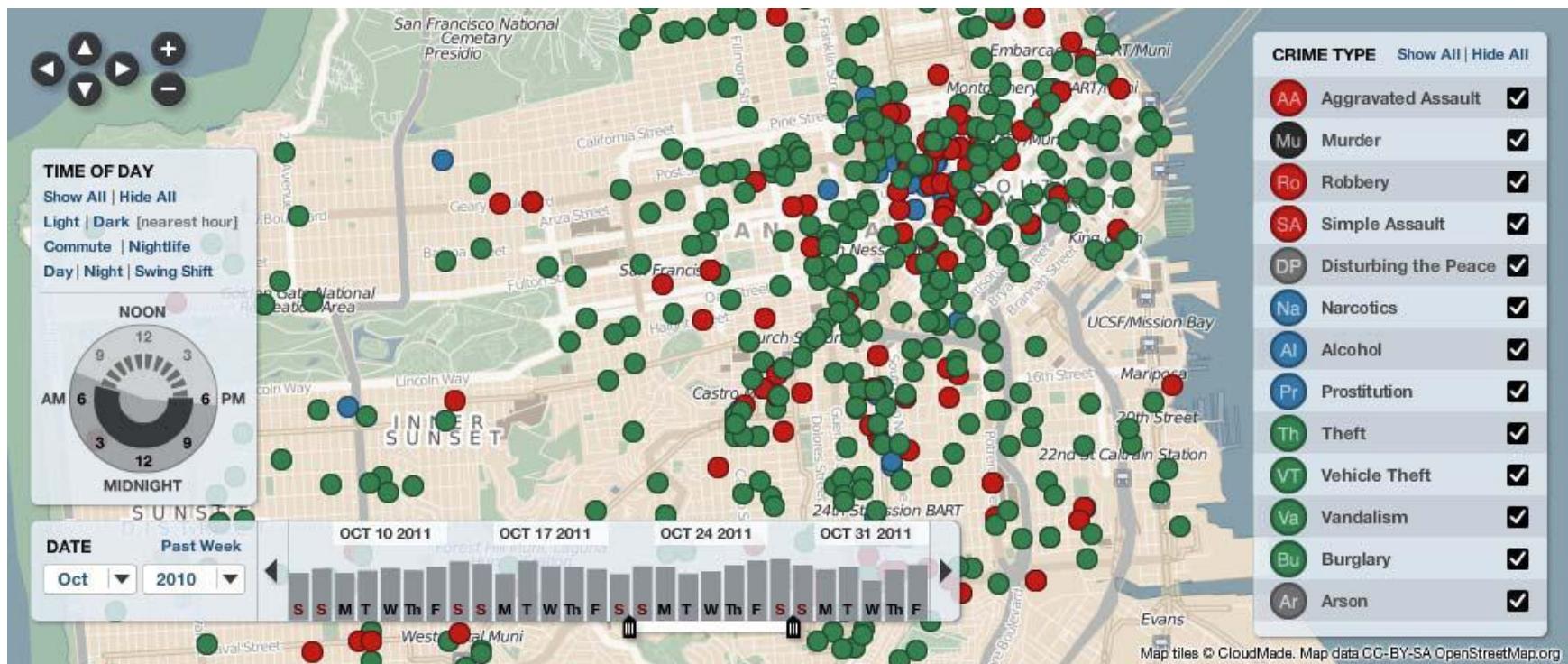
Zoom and filter,

Then details on demand

Related, history, extract

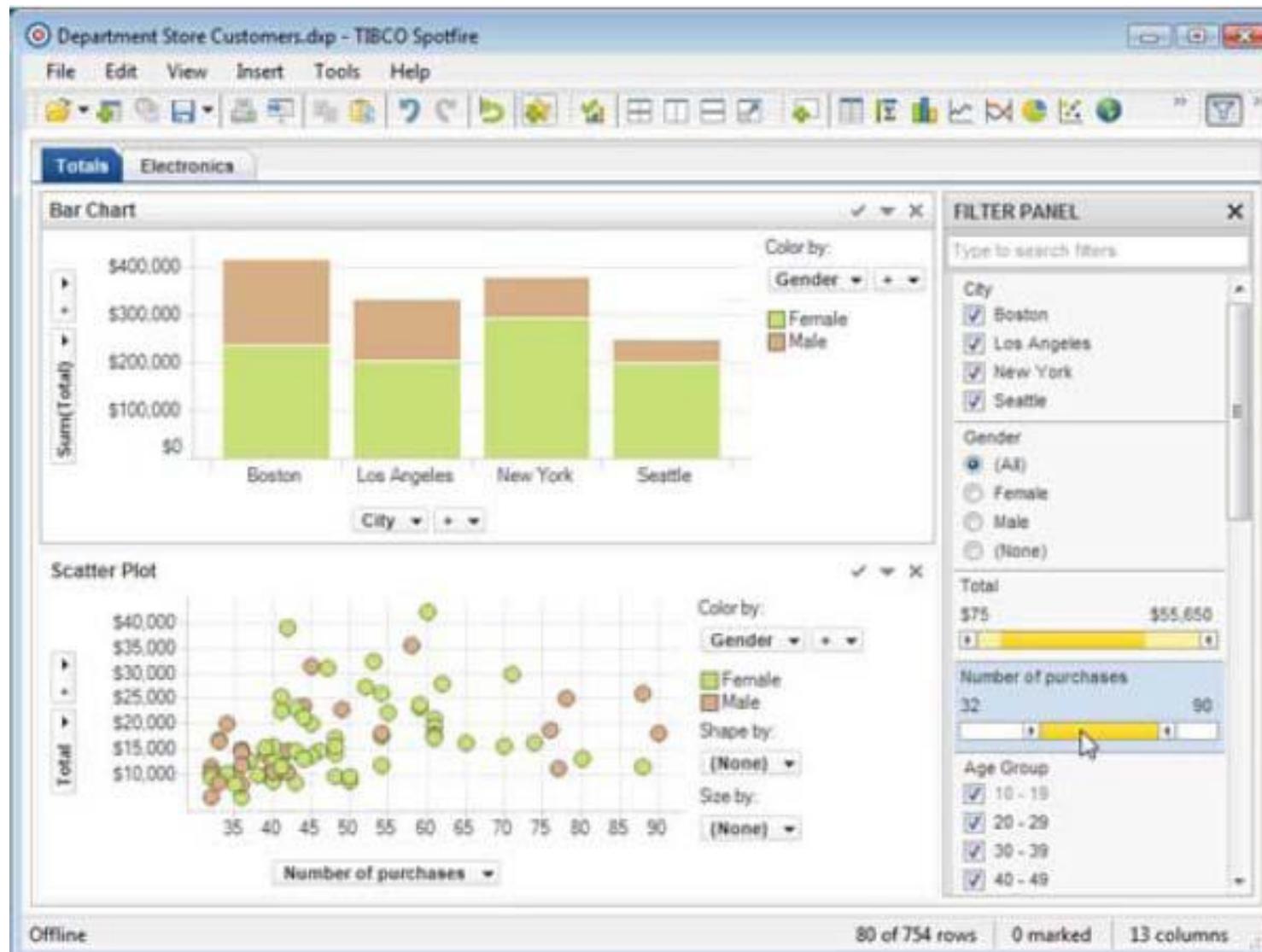
DYNAMIC QUERIES

Define criteria for inclusion/exclusion

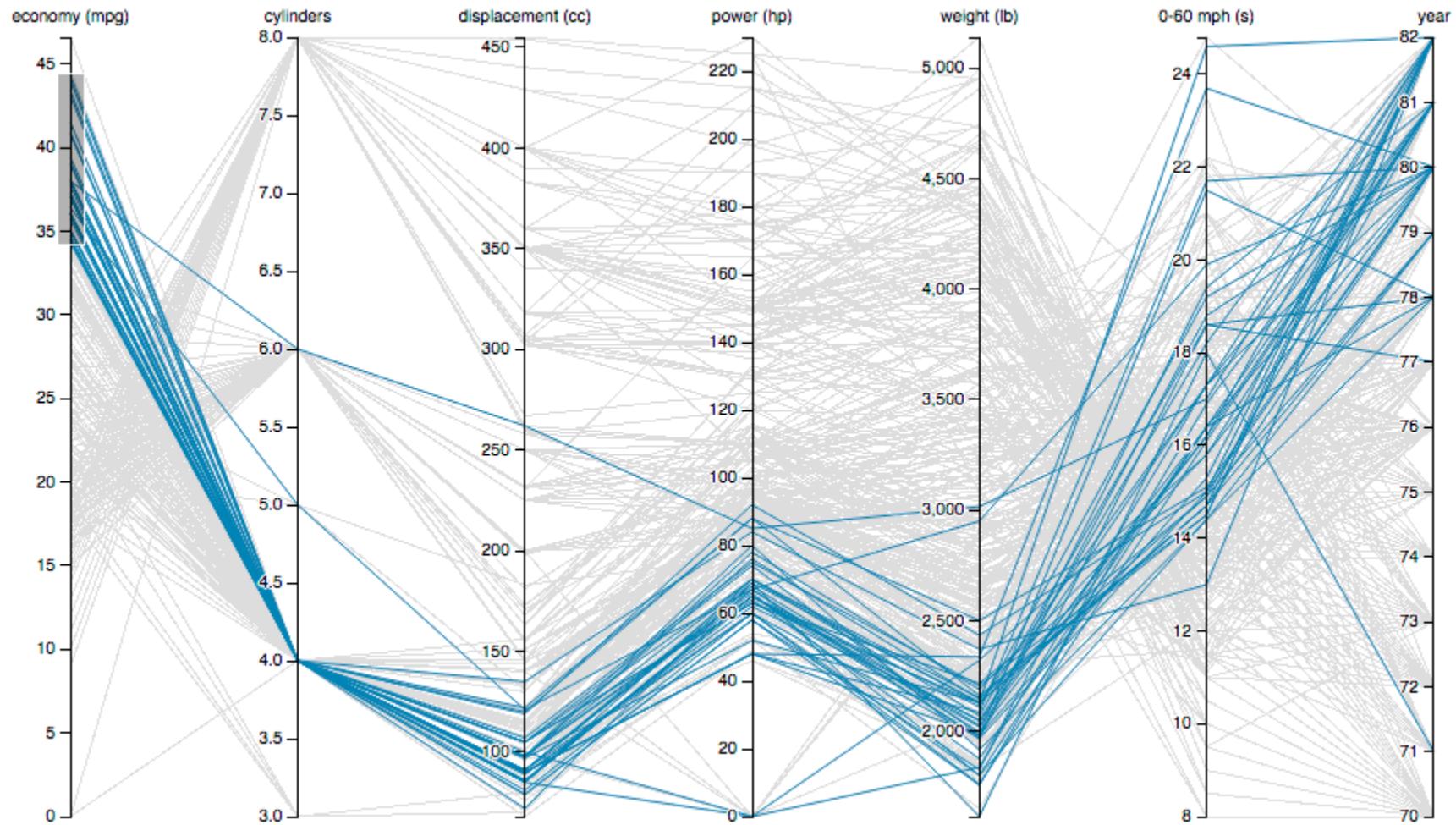


DYNAMIC QUERY WITH FILTER

Spotfire

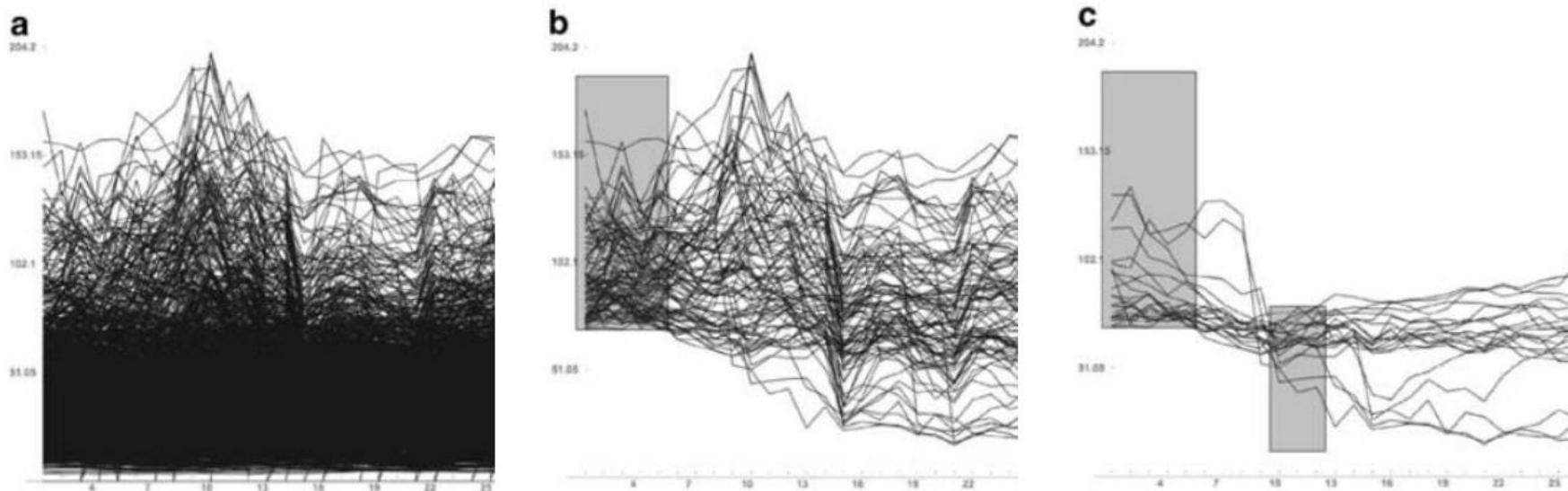


VISUAL QUERIES



<http://bl.ocks.org/jasondavies/1341281>

VISUAL QUERIES

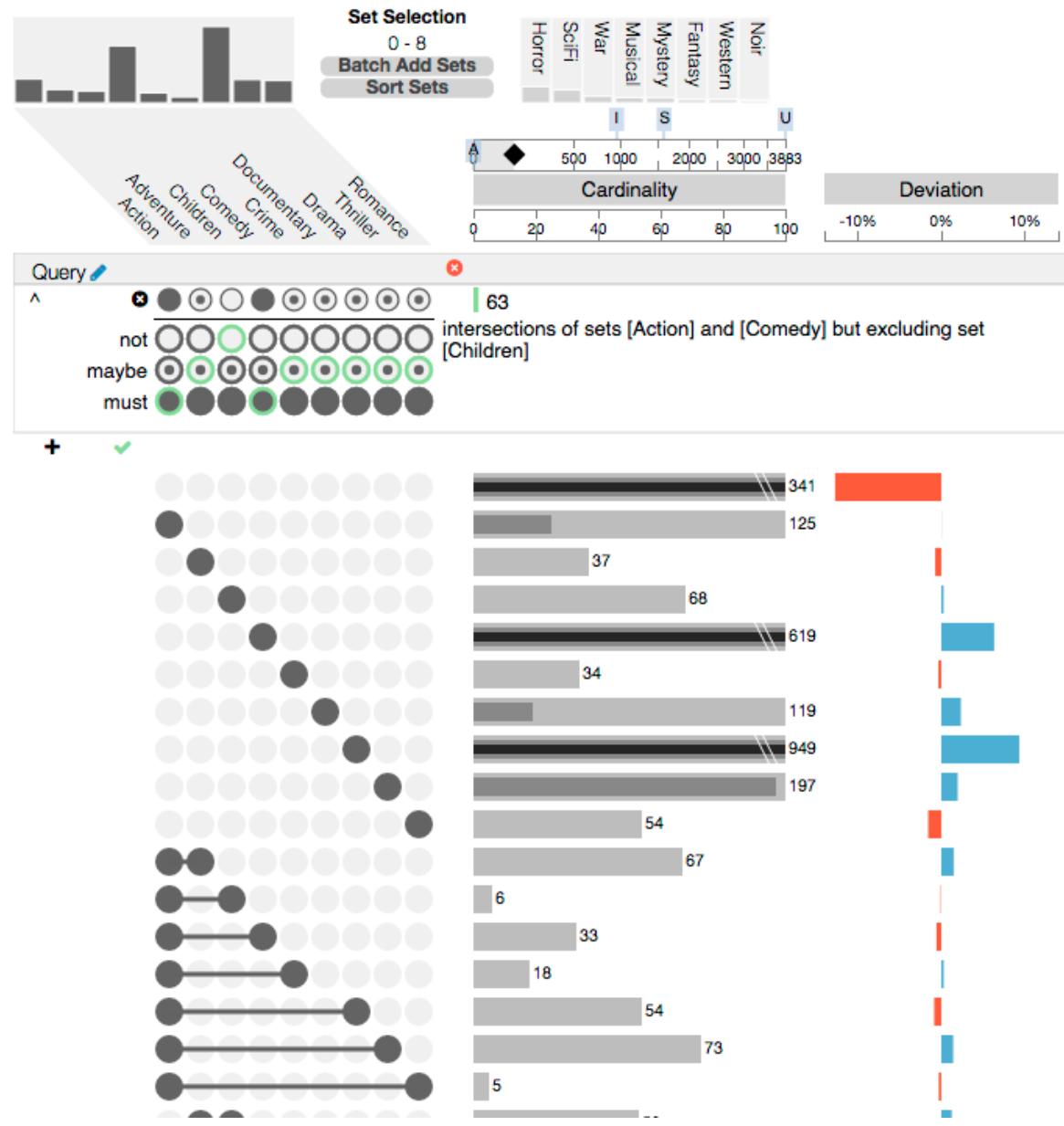


Time Searcher (Hocheiser, 2003)

6

51

QUERY INTERFACES



VIEWS

DESIGN CHOICES FOR MULTIPLE VIEWS

④ Juxtapose and Coordinate Multiple Side-by-Side Views

→ Share Encoding: Same/Different

→ *Linked Highlighting*



→ Share Data: All/Subset/None



→ Share Navigation



		Data		
		All	Subset	None
Encoding	Same	Redundant	Overview/Detail	Small Multiples
	Different	Multiform	Multiform, Overview/Detail	No Linkage

④ Partition into Side-by-Side Views



④ Superimpose Layers



WHY MULTIPLE VIEWS?

Eyes beat memory

- Showing two views side by side are easier to compare than changing views over time

No single visual encoding is optimal for all possible tasks

- Use different encoding for one data

Too many to shown in one view

LINKED VIEWS

Multiple views are simultaneously visible and linked together

Actions in one view affect the others

LINKED VIEWS OPTIONS

Four options

- Highlighting: to link, or not
- Navigation: to share, or not
- Encoding: same or multiform
- Dataset: Shared all, subset, or none

→ Share Encoding: Same/Different

→ *Linked Highlighting*



→ Share Data: All/Subset/None

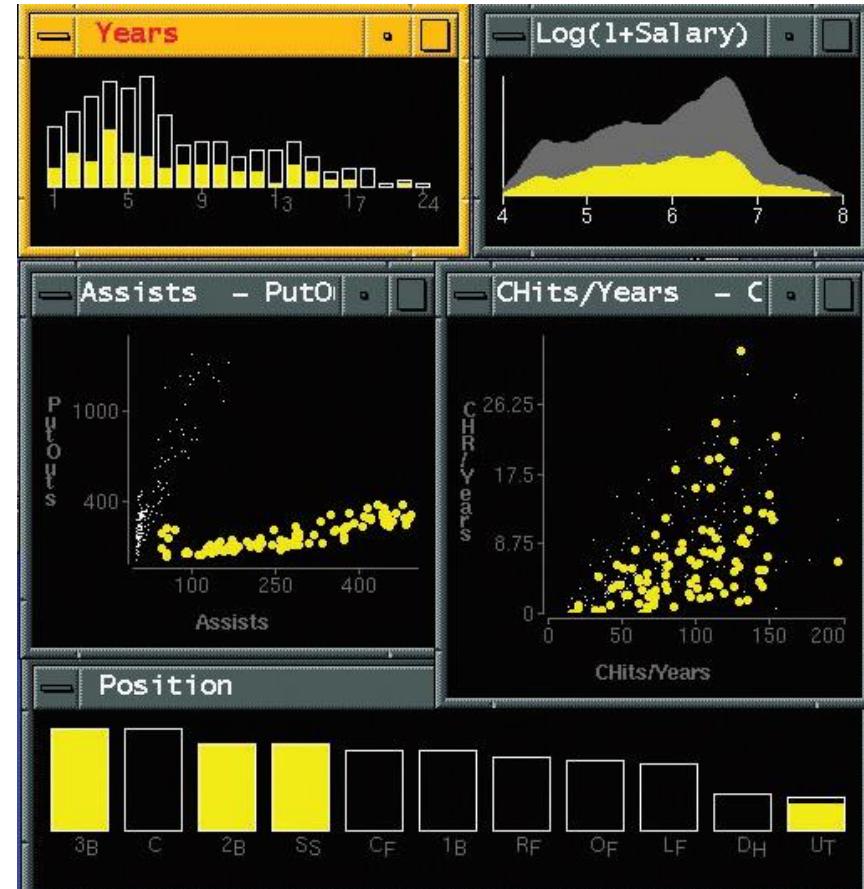
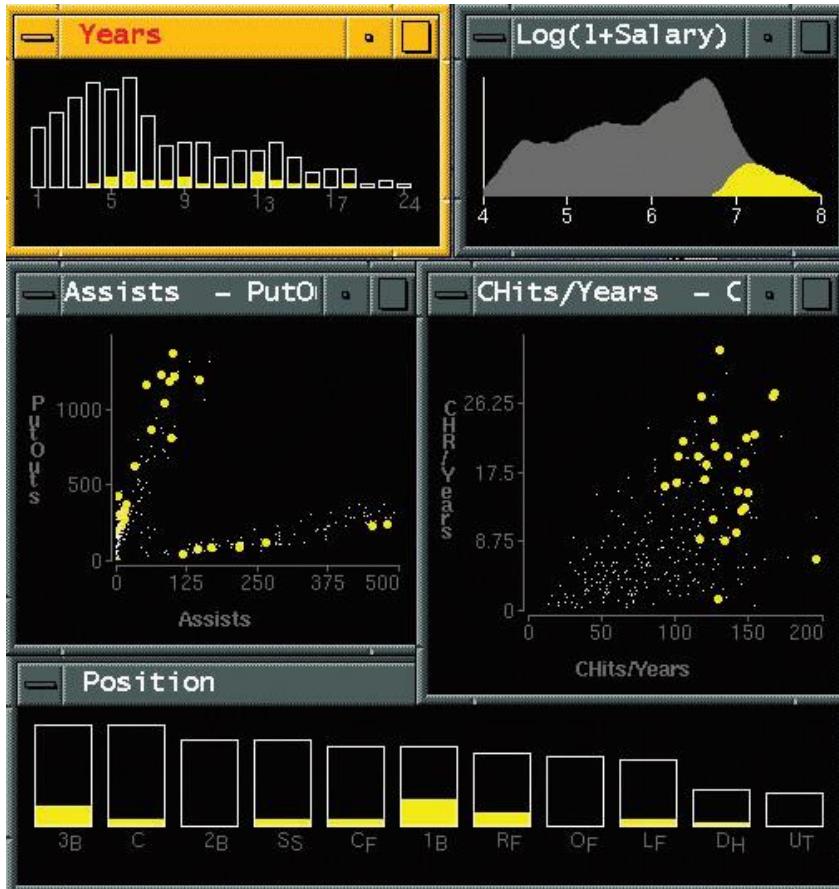


→ Share Navigation

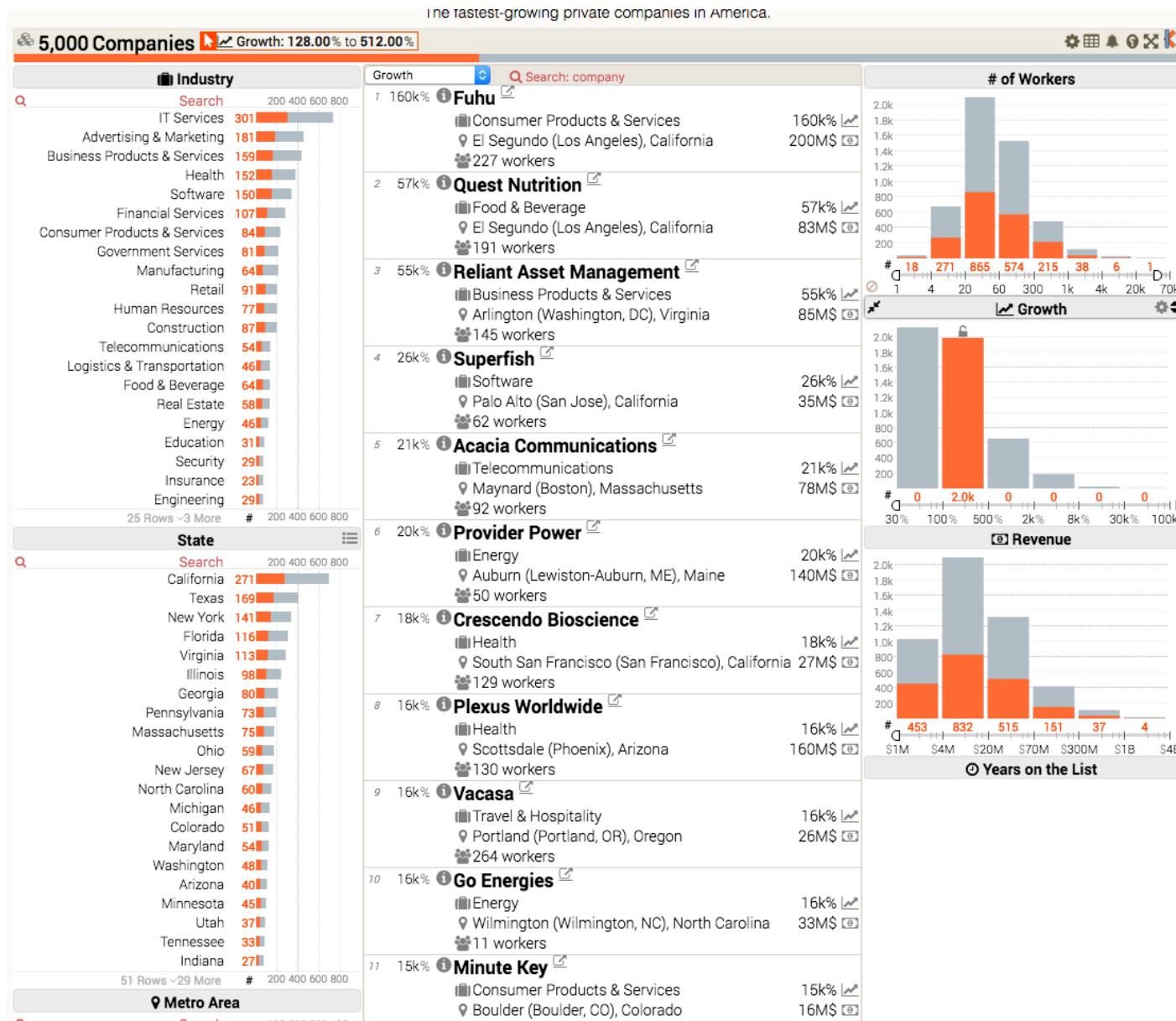


		Data		
		All	Subset	None
Encoding	Same	Redundant		
	Different	Multiform	Multiform, Overview/Detail	No Linkage

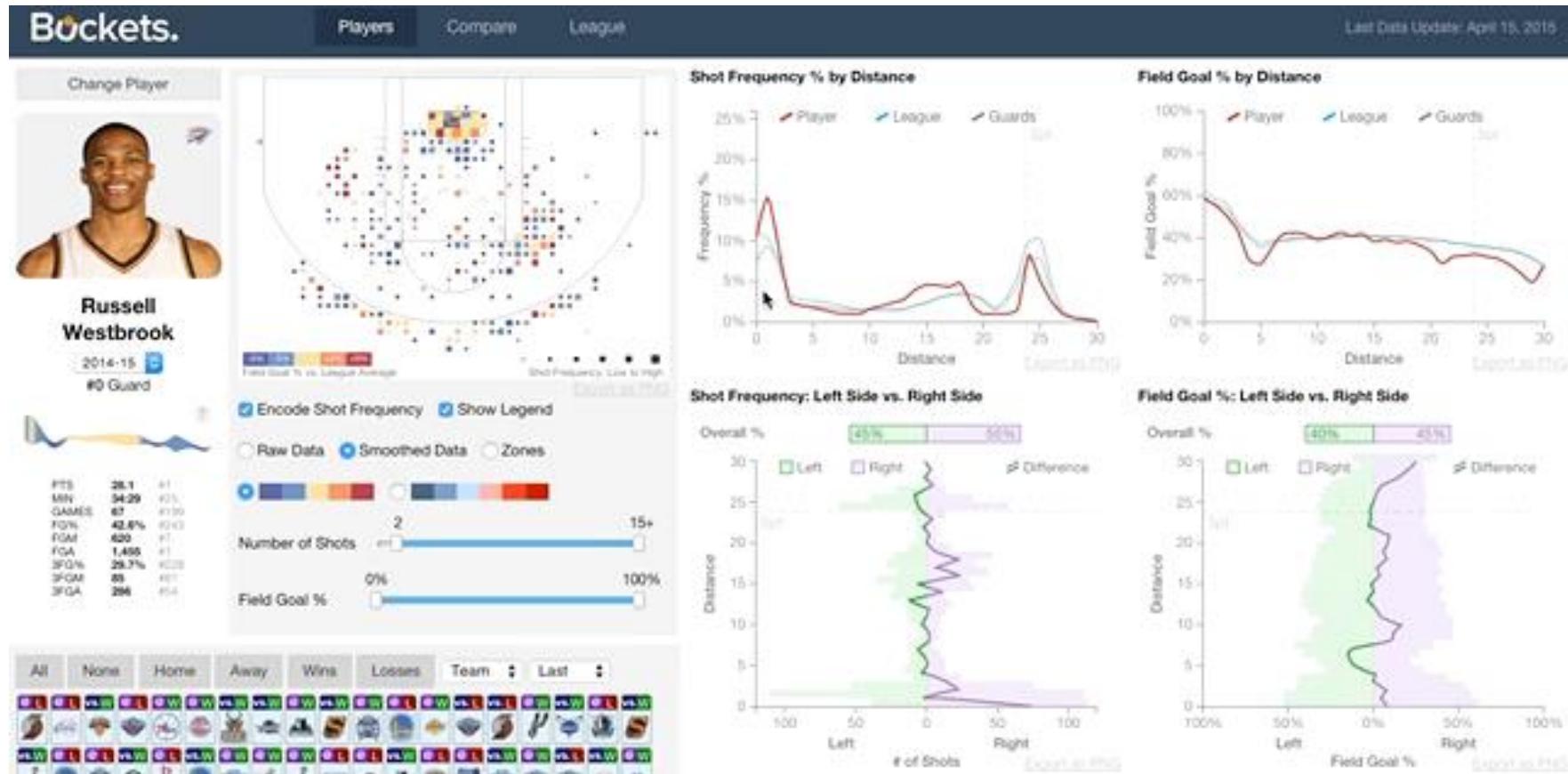
LINKED HIGHLIGHTING



LINKED HIGHLIGHTING



MULTIDIRECTIONAL LINKING



https://buckets.peterbeshai.com/app/#/playerView/201935_2015

MULTIFORM

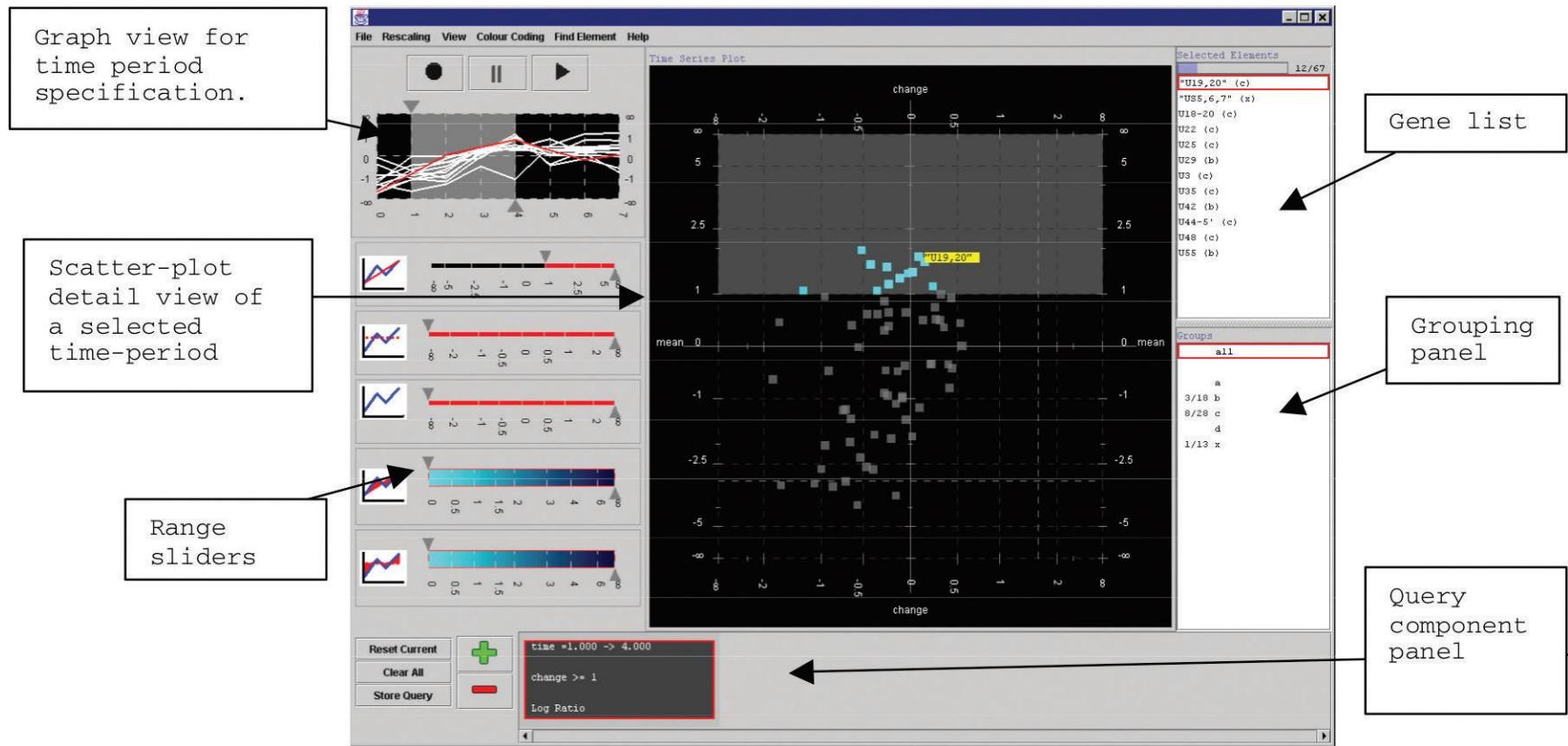
Different visual encodings are used between views

- shared data: either all data, or subset of data (overview + detail)

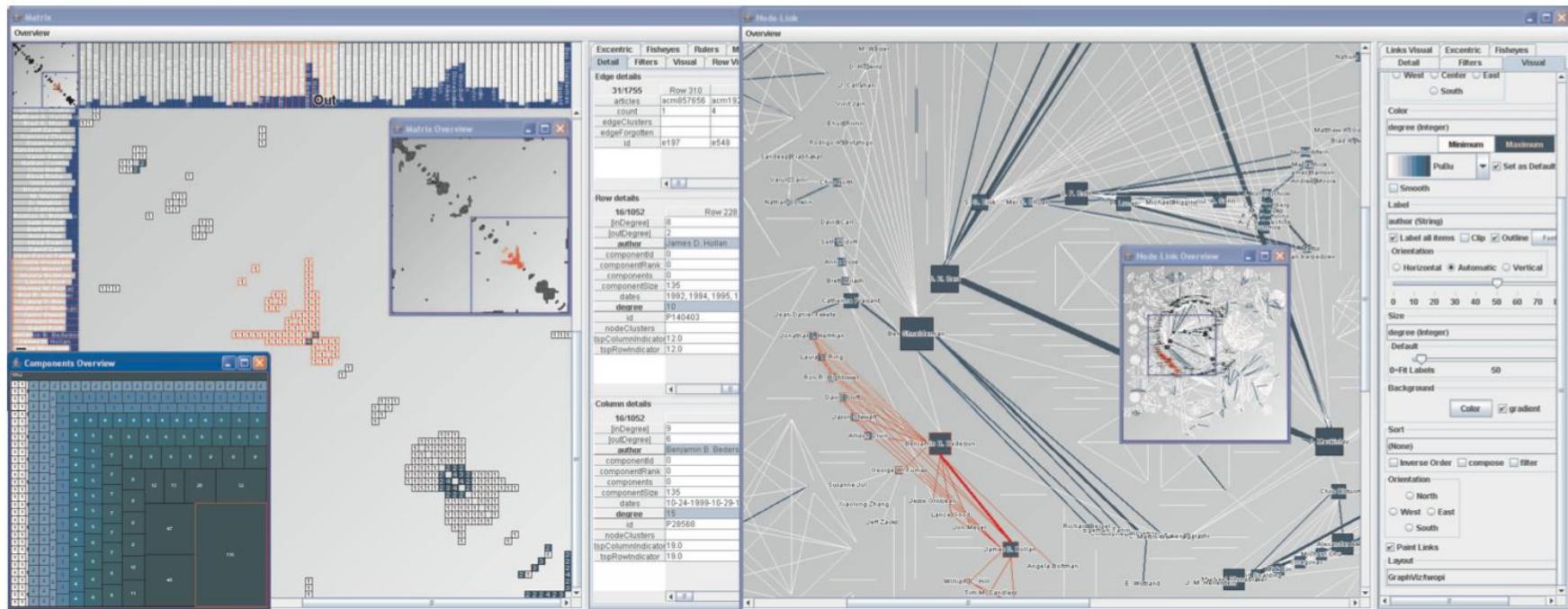
Rational:

- Single view has strong limits on the numbers of attributes that can be shown simultaneously
- Different views support different tasks

MULTIFORM OVERVIEW-DETAIL



MATRIX EXPLORER



MULTIFORM

demo.caleydo.org/pathfinder/main.html#uc=dblp

Pathfinder

Start Hanspeter Pfleis | End Ben Shneiderma

Advanced Query

Length Paths 0 1 2 3 4

Path List

1. Hanspeter Pfleis → Frank van Ham → Adam Perer → Ben Shneiderma (3 paths)

1. Hanspeter Pfleis → Krzysztof Z. Gajc → Desney S. Tan → Ben Shneiderma (3 paths)

1. Hanspeter Pfleis → Jean-Daniel Fekc → Catherine Plaisai → Ben Shneiderma (3 paths)

4. Hanspeter Pfleis → Jean-Daniel Fekc → Catherine Plaisai → Jennifer Golbed → Ben Shneiderma (4 paths)

4. Hanspeter Pfleis → Jean-Daniel Fekc → Wendy E. Macka → Ed Huai-hsin Ch → Ben Shneiderma (4 paths)

4. Hanspeter Pfleis → Krzysztof Z. Gajc → Jeffrey Heer → Ed Huai-hsin Ch → Ben Shneiderma (4 paths)

4. Hanspeter Pfleis → Krzysztof Z. Gajc → Jeffrey Heer → Stuart K. Card → Ben Shneiderma (4 paths)

4. Hanspeter Pfleis → Jean-Daniel Fekc → Catherine Plaisai → Krist Wongsupha → Ben Shneiderma (4 paths)

Path Topology

OVERVIEW + DETAIL

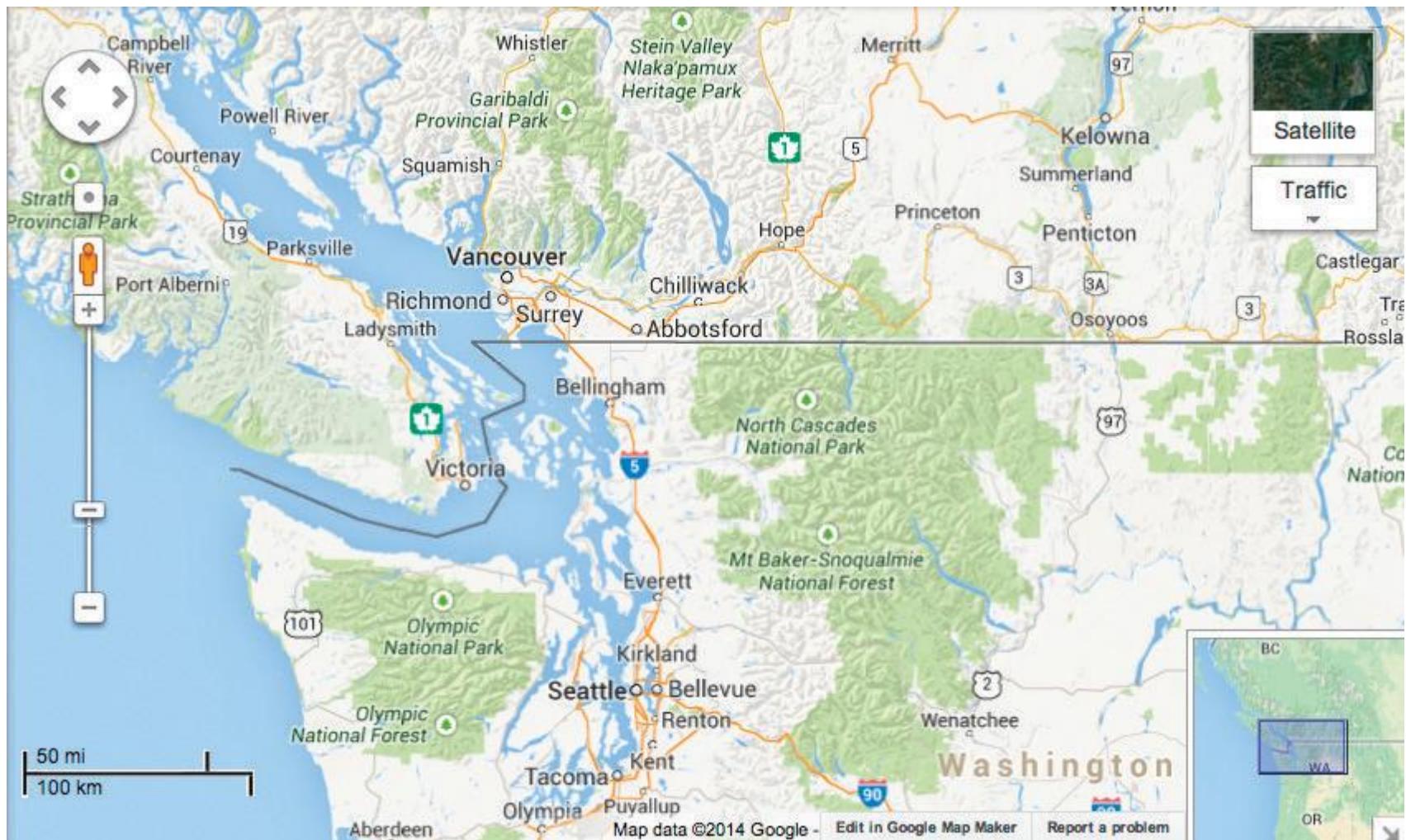
One view shows (often summarized) information about entire dataset

Additional view(s) show more detailed information about a subset of the data

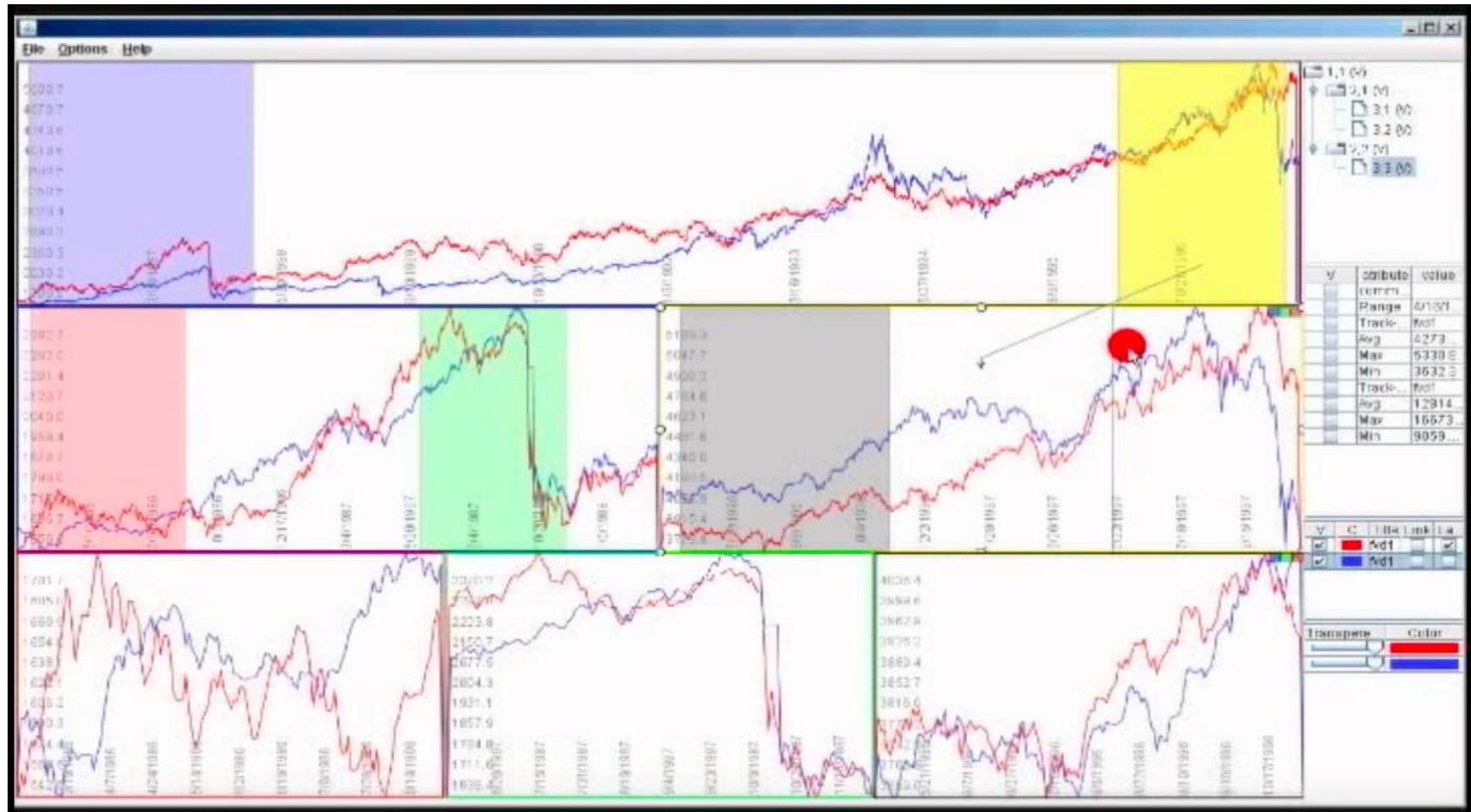
Rational:

For large or complex data, a single view of the entire dataset cannot capture fine details

MULTIFORM: OVERVIEW-DETAIL

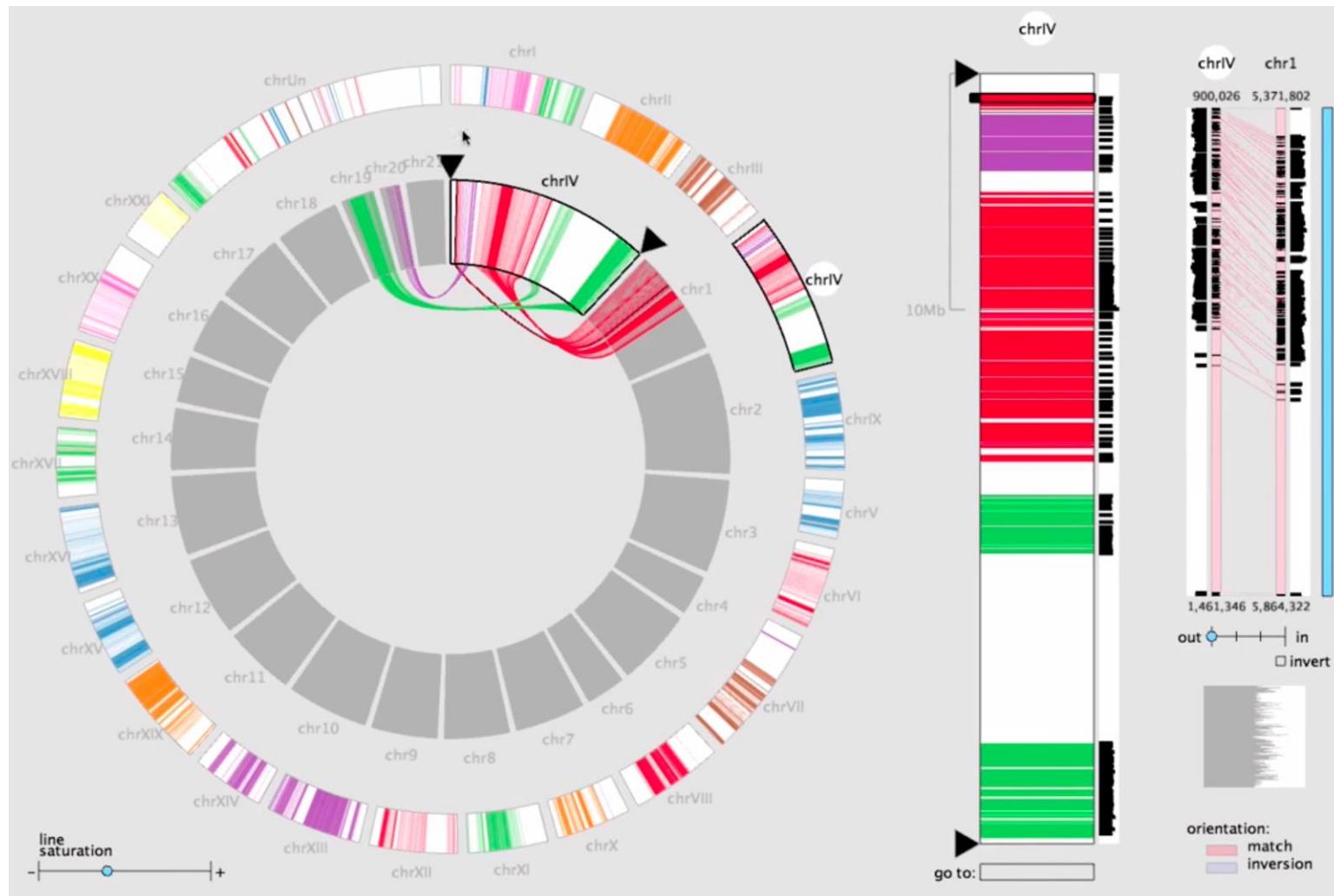


STACK ZOOMING

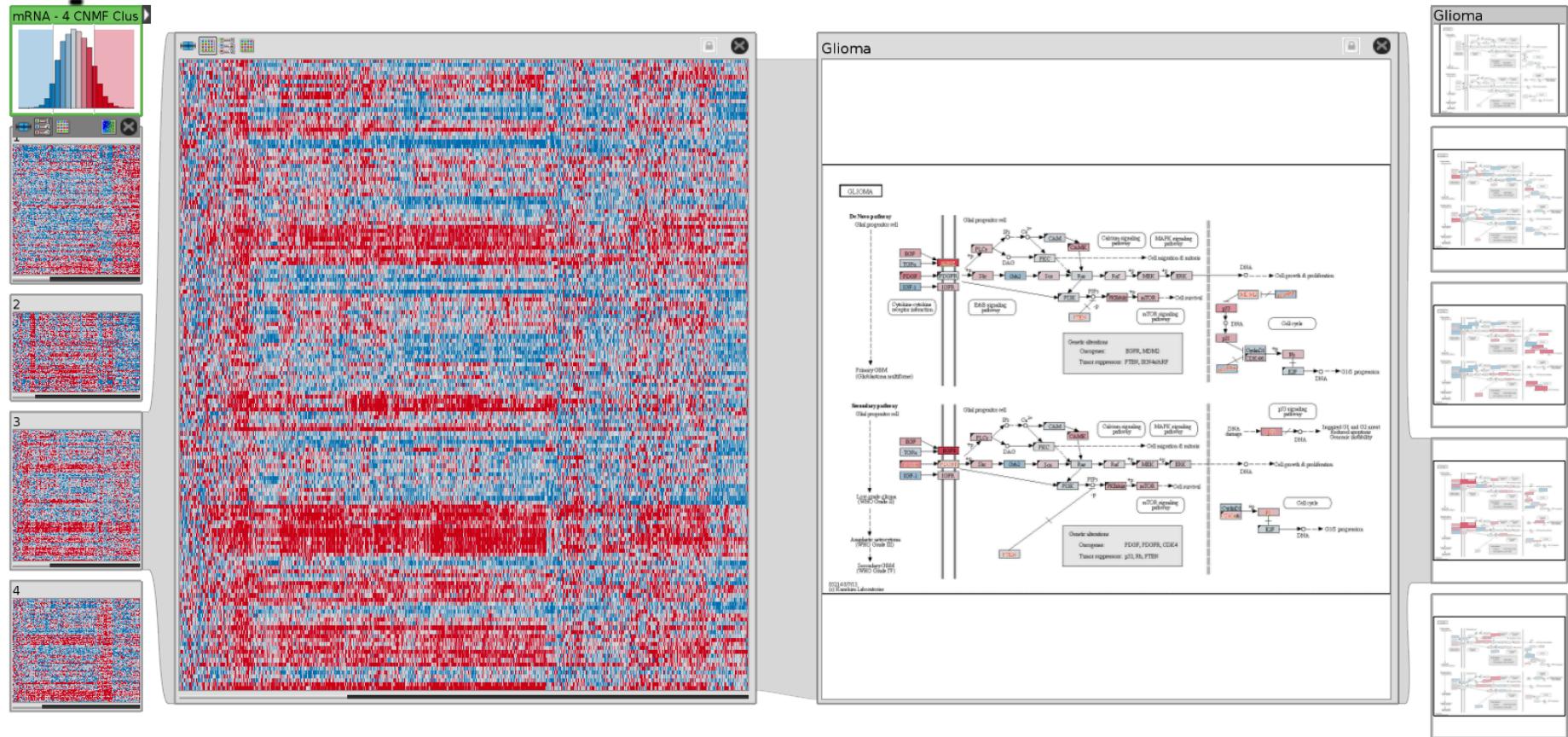


Same Data - Same Encoding, Different Resolution
[Javed & Emlqvist, PacificVis, 2010]

MIZBEE



STRATOMEX



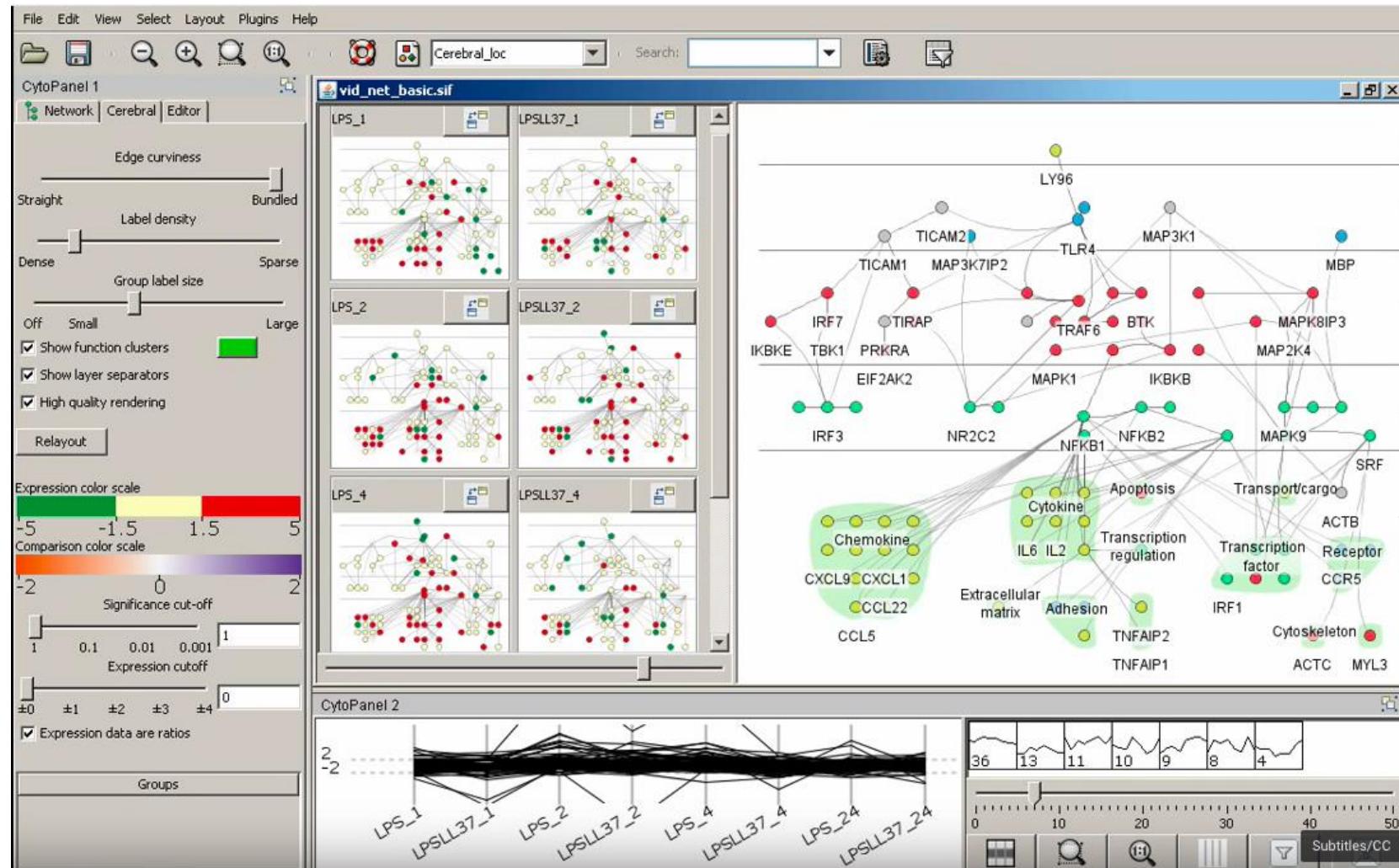
SMALL MULTIPLES

Each view uses the same visual encoding, but shows a different subset of the data

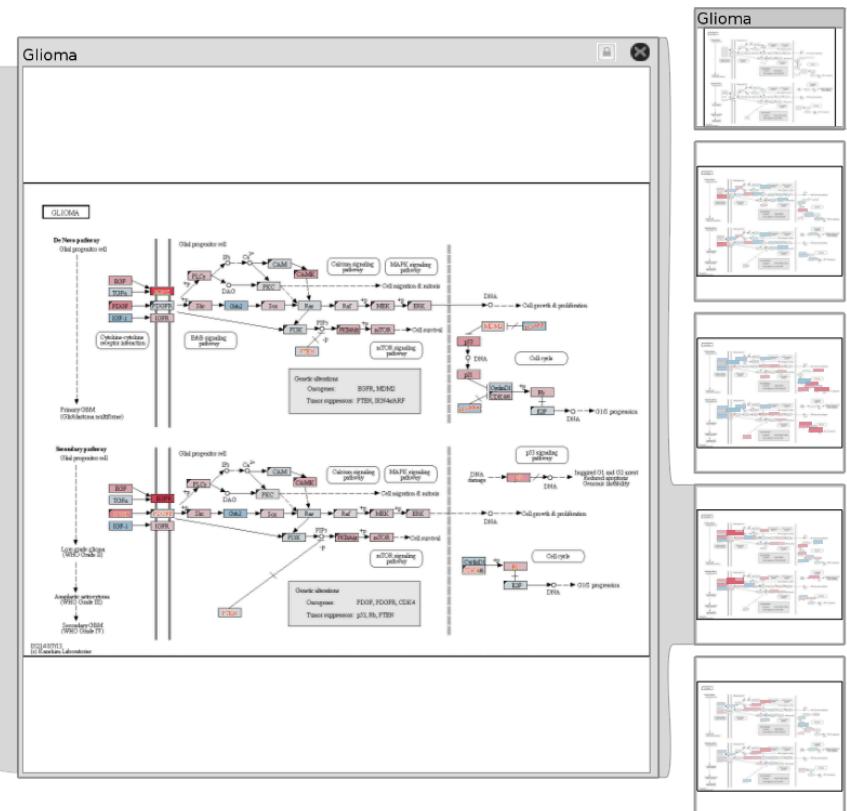
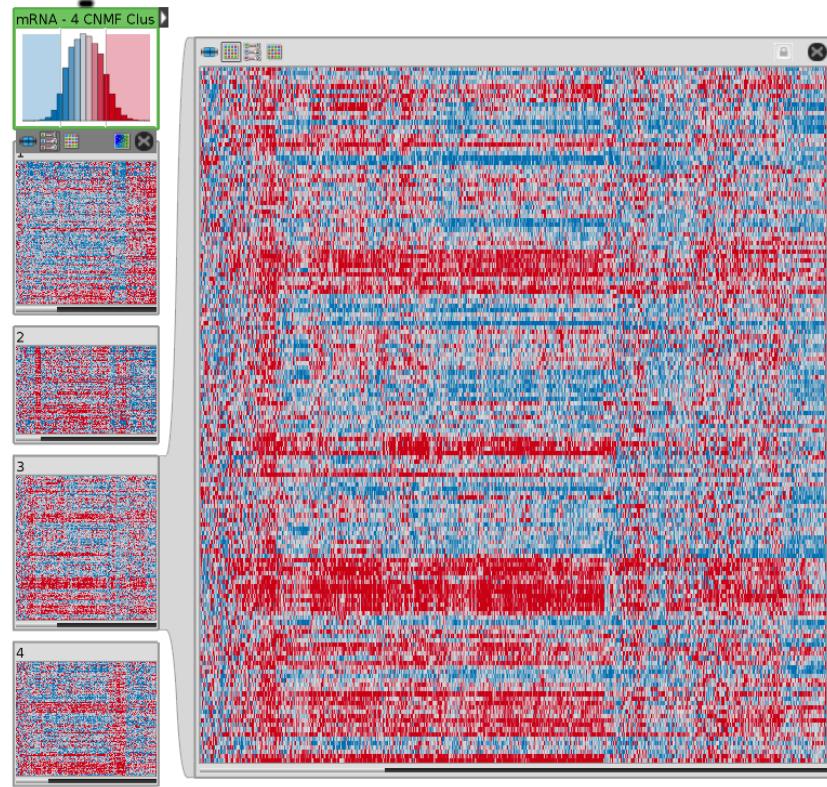
Rational:

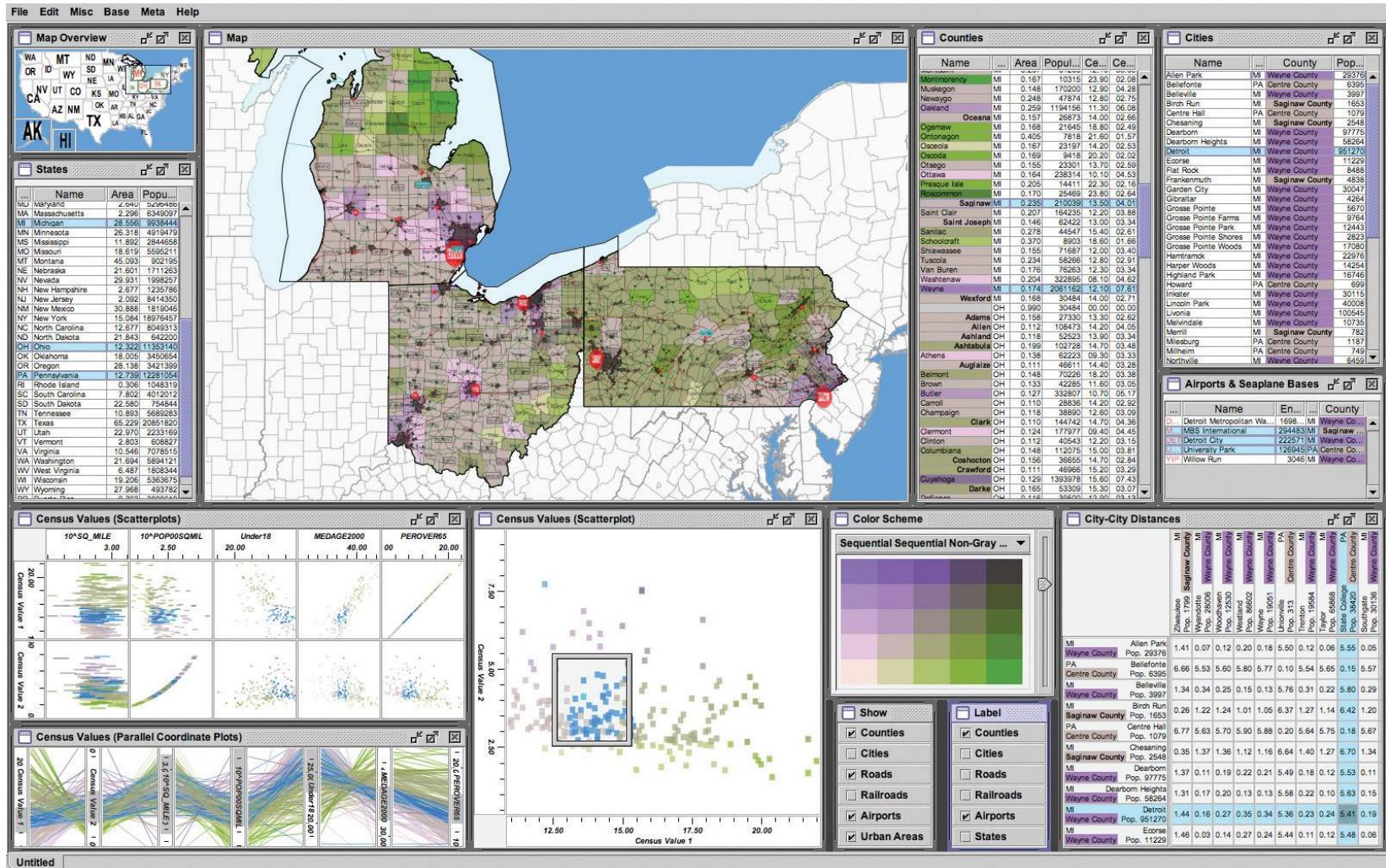
Quickly compare different parts of a dataset,
relying on eyes instead of memory

SMALL MULTIPLES FOR GRAPH ATTRIBUTES

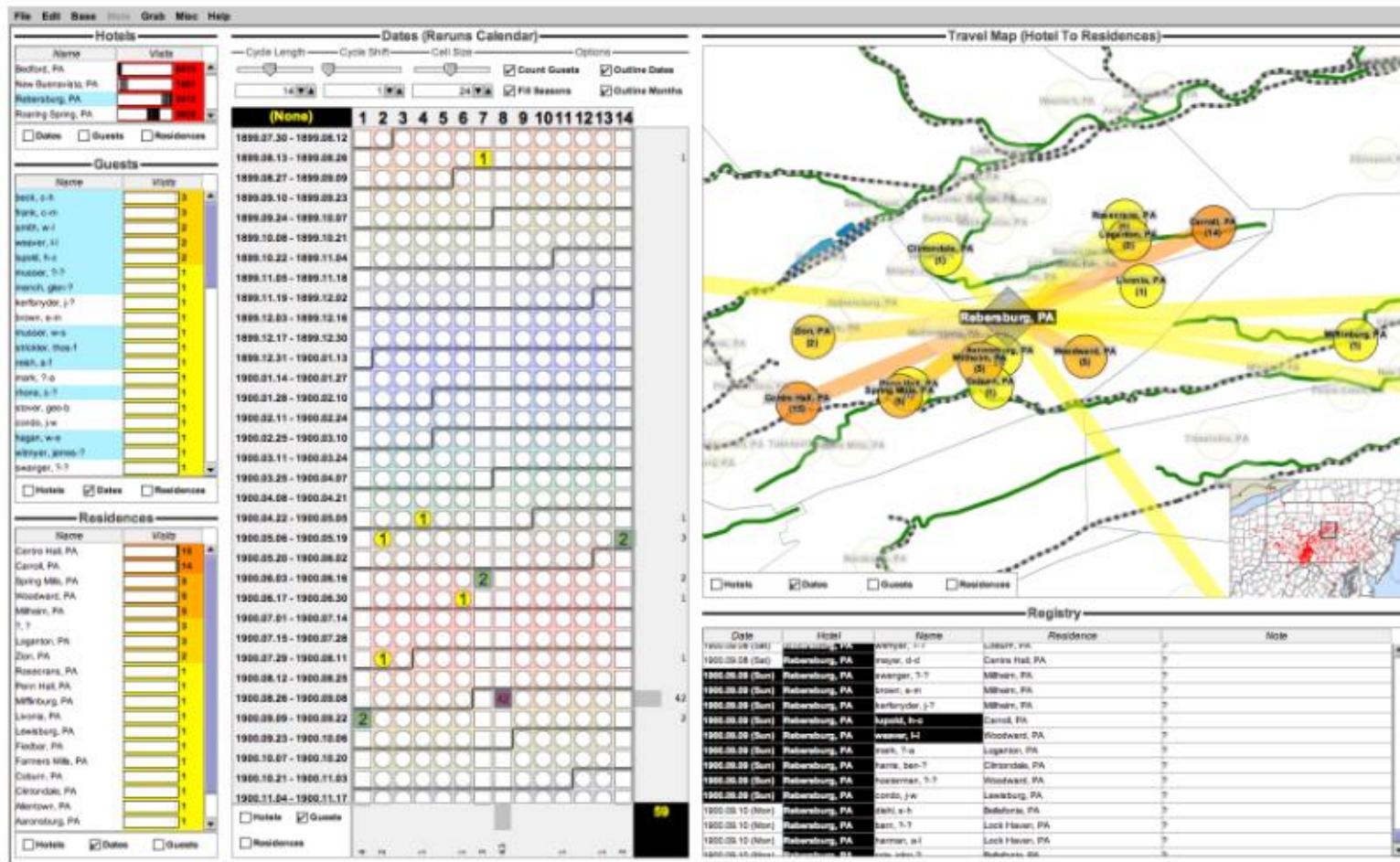


STRATOMEX





VISUAL ANALYSIS OF HISTORICAL HOTEL VISITATION PATTERNS



Click image for full-size version.

<https://www.youtube.com/watch?v=Tzsv6wkZoiQ>

PARTITIONING

PARTITIONING

Action on the dataset that separates the data into groups

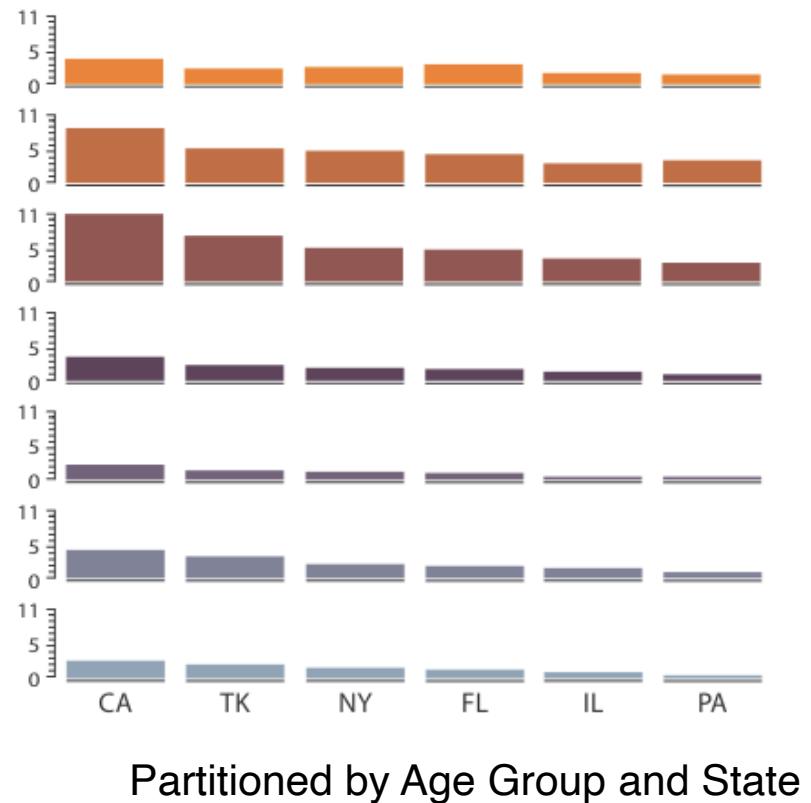
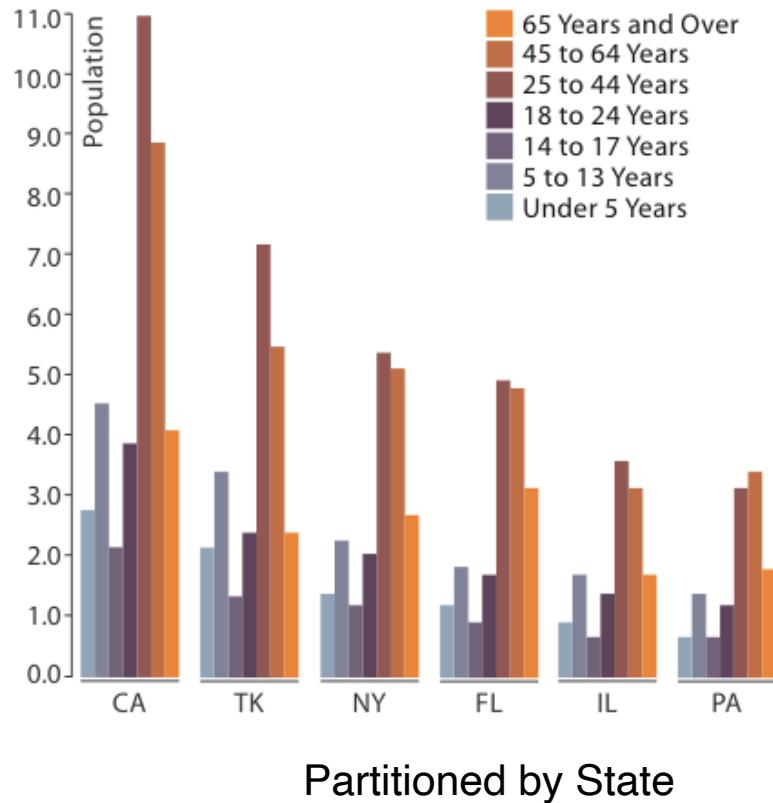
Design choices

- How to divide data up between views, given a hierarchy of attributes
- How many splits, and order of splits
- How many views (usually data driven)

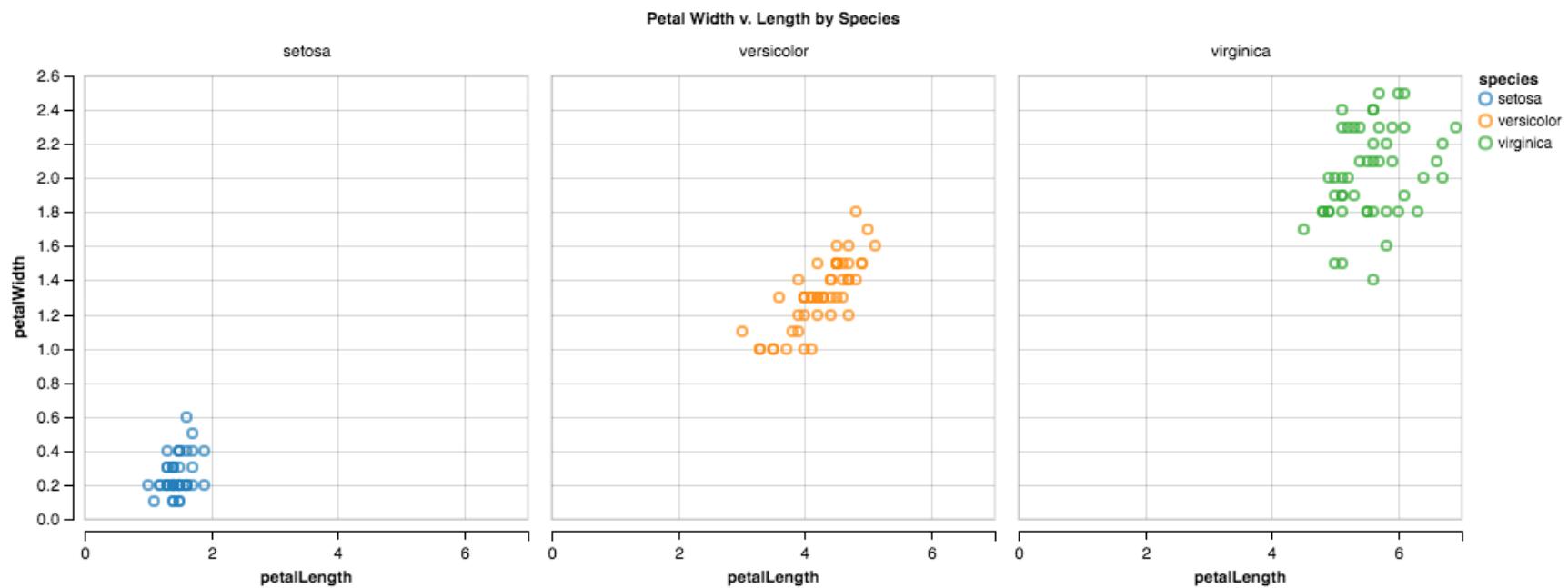
Partition attribute(s)

Typically categorical

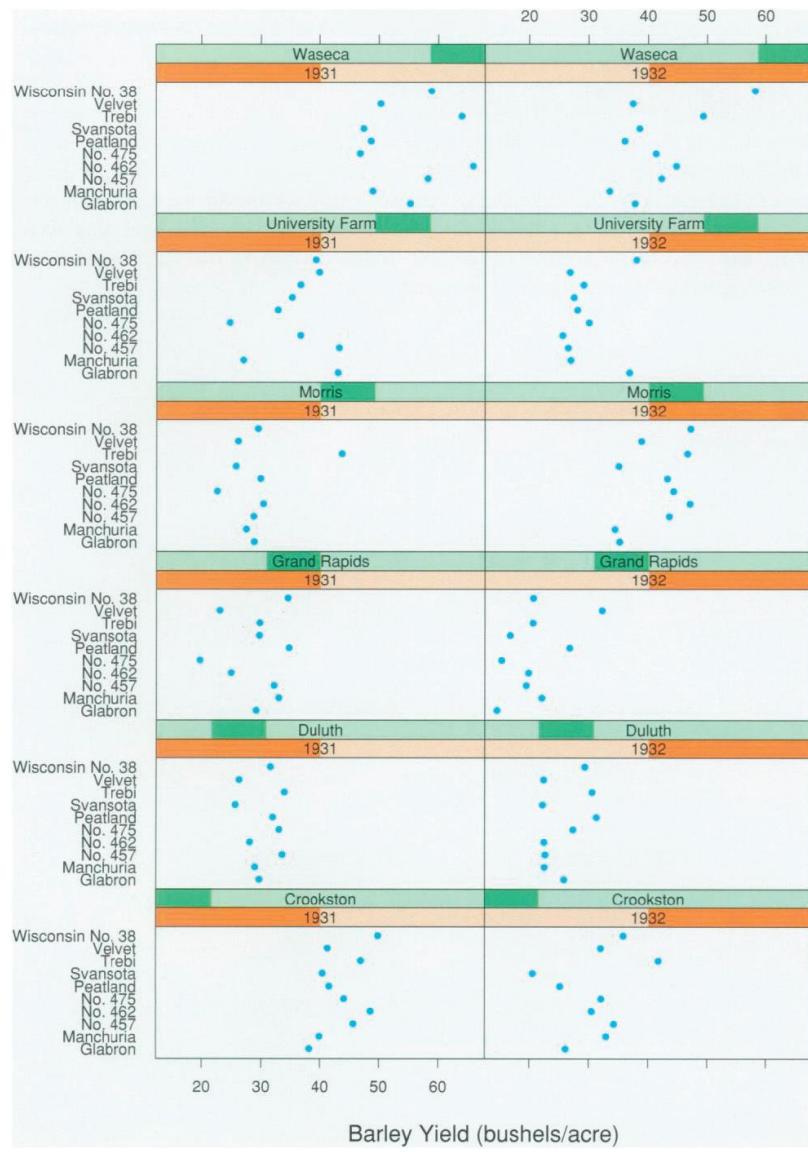
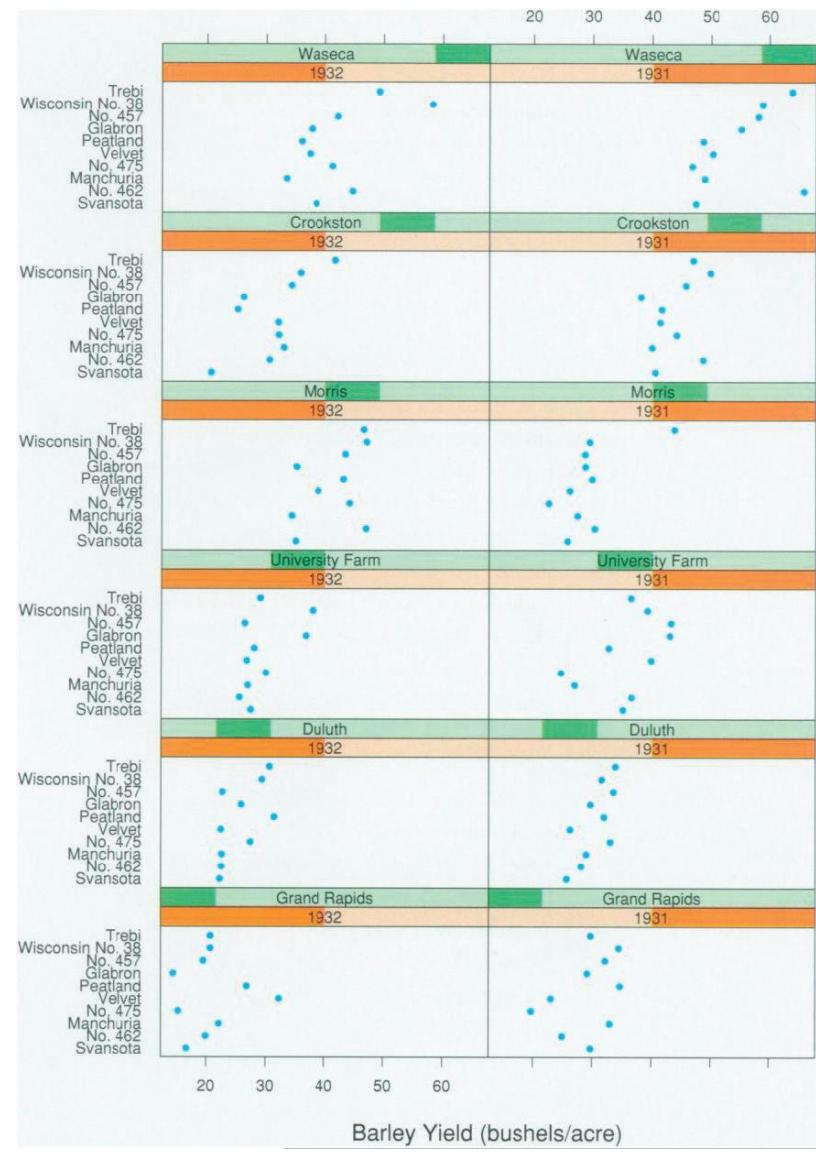
AGE DISTRIBUTION BY STATE



PARTITION BY CATEGORY



TRELLIS



TRELLIS PLOTS

panel variables

attributes encoded in individual views

partitioning variables

partitioning attributes assigned to columns
and rows

main-effects ordering

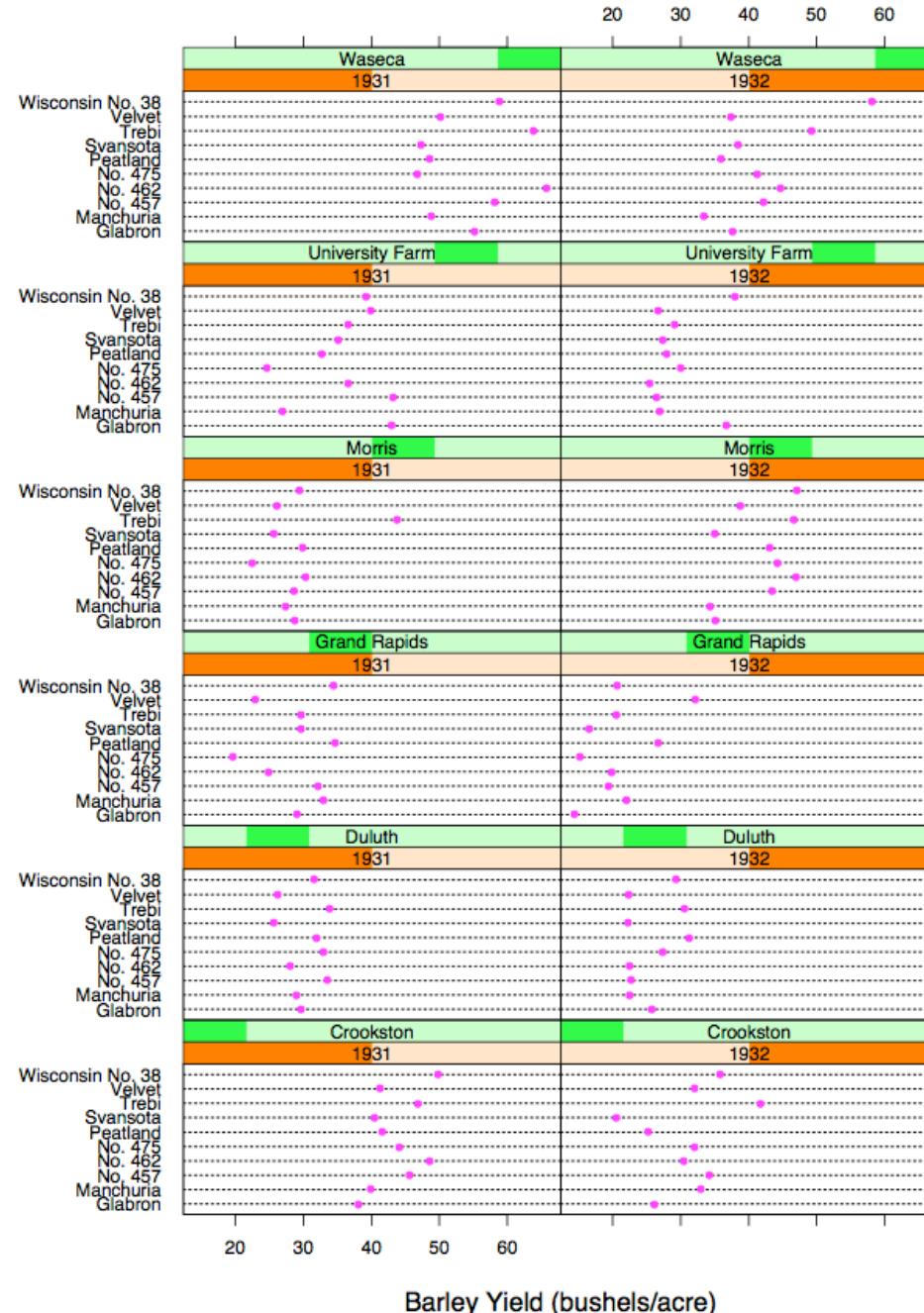
order partitioning variable based on derived
data

support perception of trends and structure in
data

Data

Barley Yields in two years across multiple farms for multiple barley strains

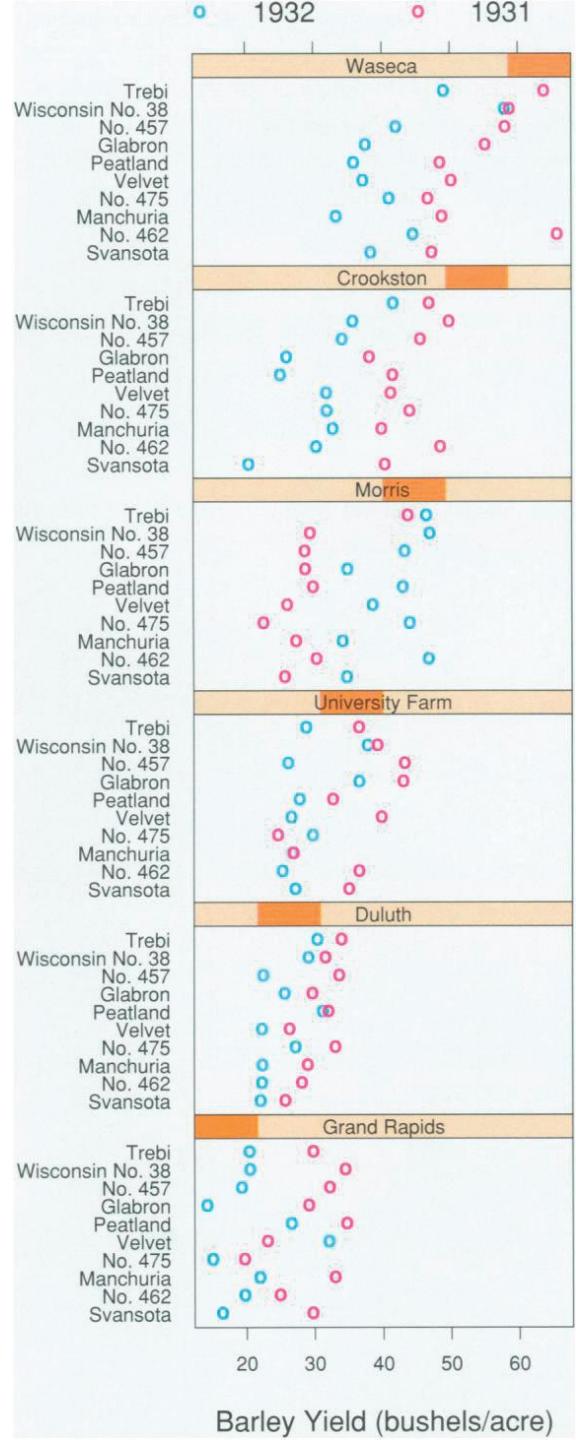
partitioning variables
Columns partitioned by year
Rows partitioned by farm



TRELLIS

Years in a single plot with year encoded by color

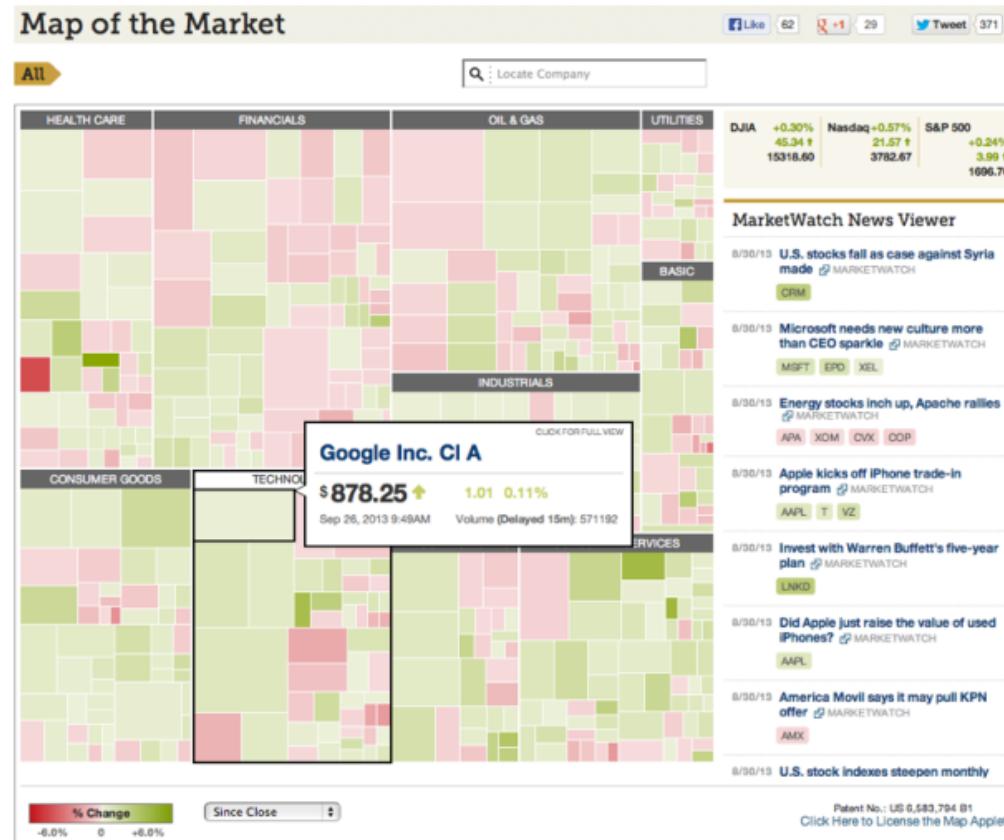
Showing strong evidence of an anomaly in the data
(the 3rd part)



RECURSIVE SUBDIVISION

partitioning:

Flexibly transform data attributes into a hierarchy use **treemaps** as spacefilling rectangular layouts



Treemap

HIVE EXAMPLE: LONDON PROPERTY

partitioning attributes
house type
neighborhood
sale time

encoding attributes
average price (color)
number of sales (size)
results

between neighborhoods,
different housing
distributions
within neighborhoods,
similar prices



HIVE EXAMPLE

partitioning attributes
neighborhood
house type
sale time (year)

sale time (month)

encoding attributes
neighborhood location
(approximate)

average price (color)
n/a (size)

results

expensive neighborhoods
near center of city



Configuring Hierarchical Layouts to Address Research Questions



Aidan Slingsby, Jason Dykes and Jo Wood
giCentre, Department of Information Science, City University London
http://www.gicentre.org/hierarchical_layouts/



<https://vimeo.com/9870257>

LAYERING

Combining multiple views on top of one another

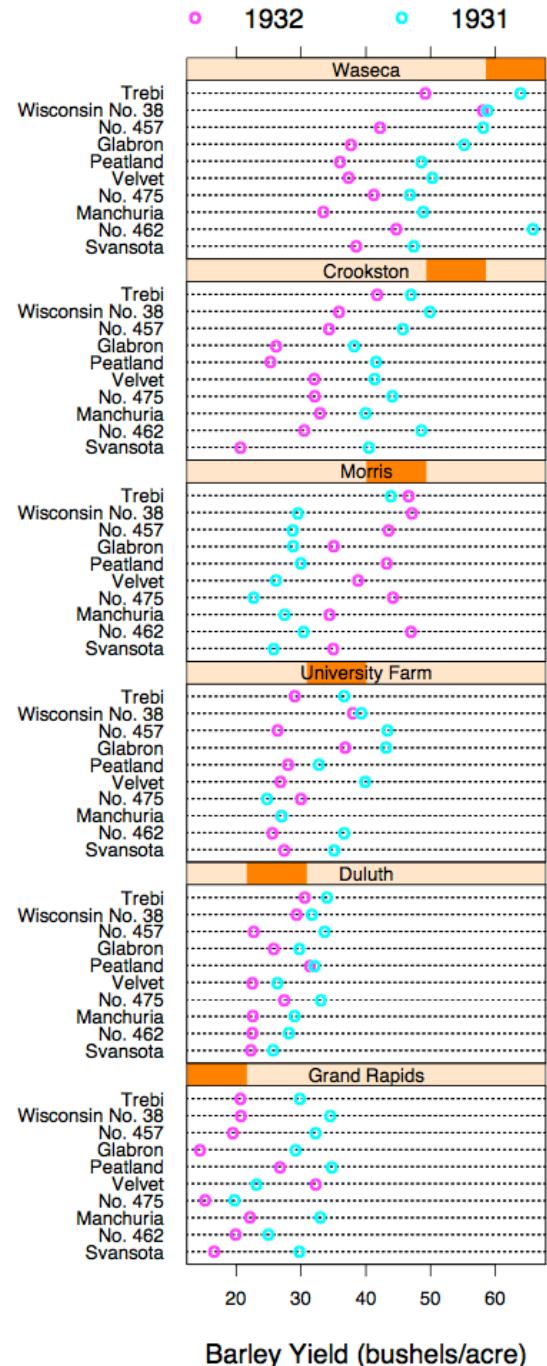
To form a composite view

Rational

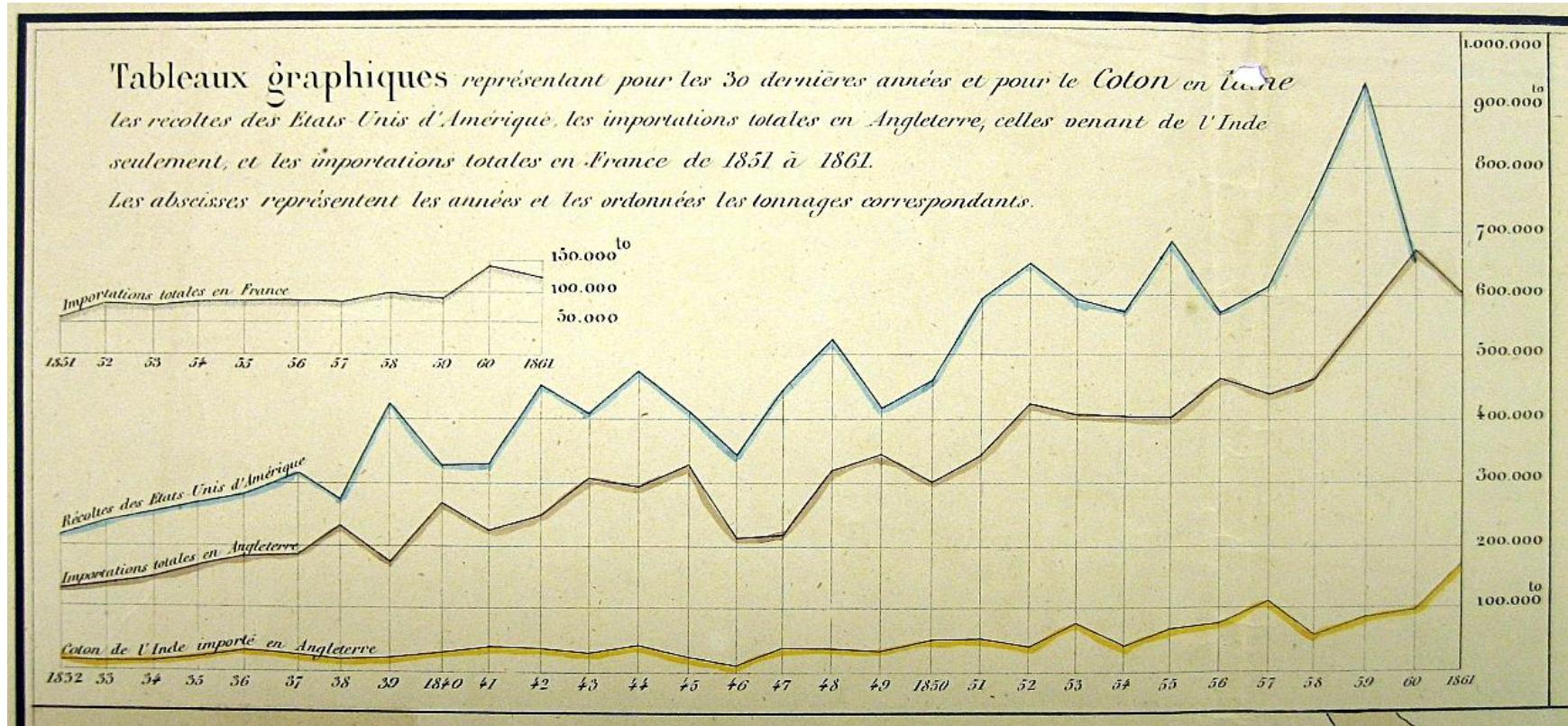
Supports a larger, more detailed view than using multiple views

Trade-off

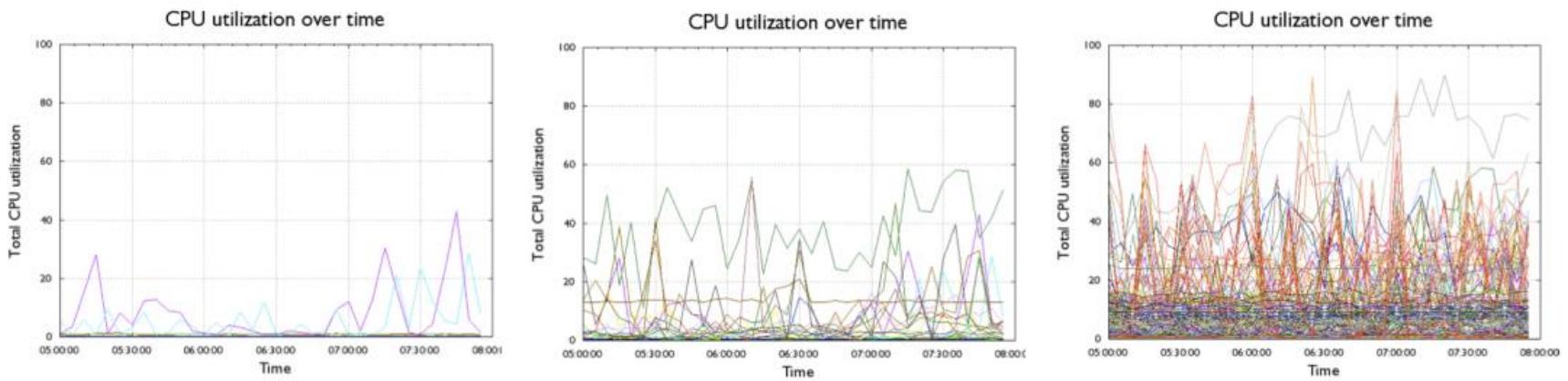
Layering imposes constraints on visual encoding choice as well as number of layers that can be shown



JOSEPH MINARD



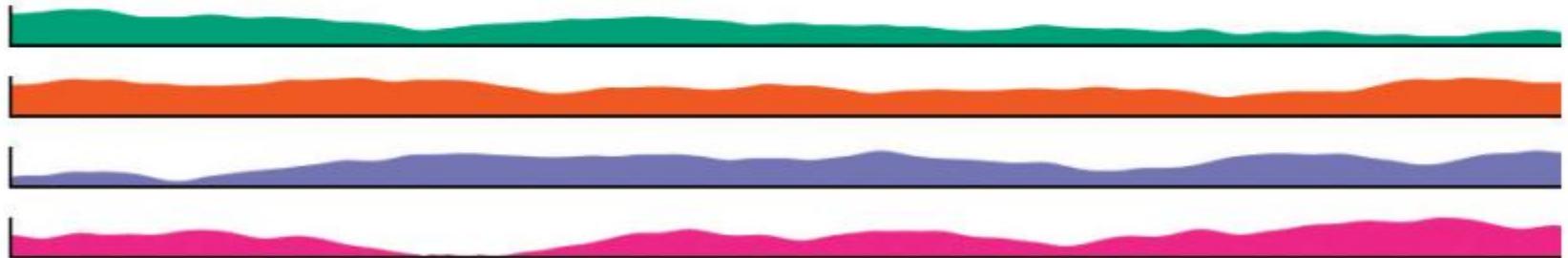
OVERLAYS



SUPERIMPOSED VS JUXTAPOSED



(a)



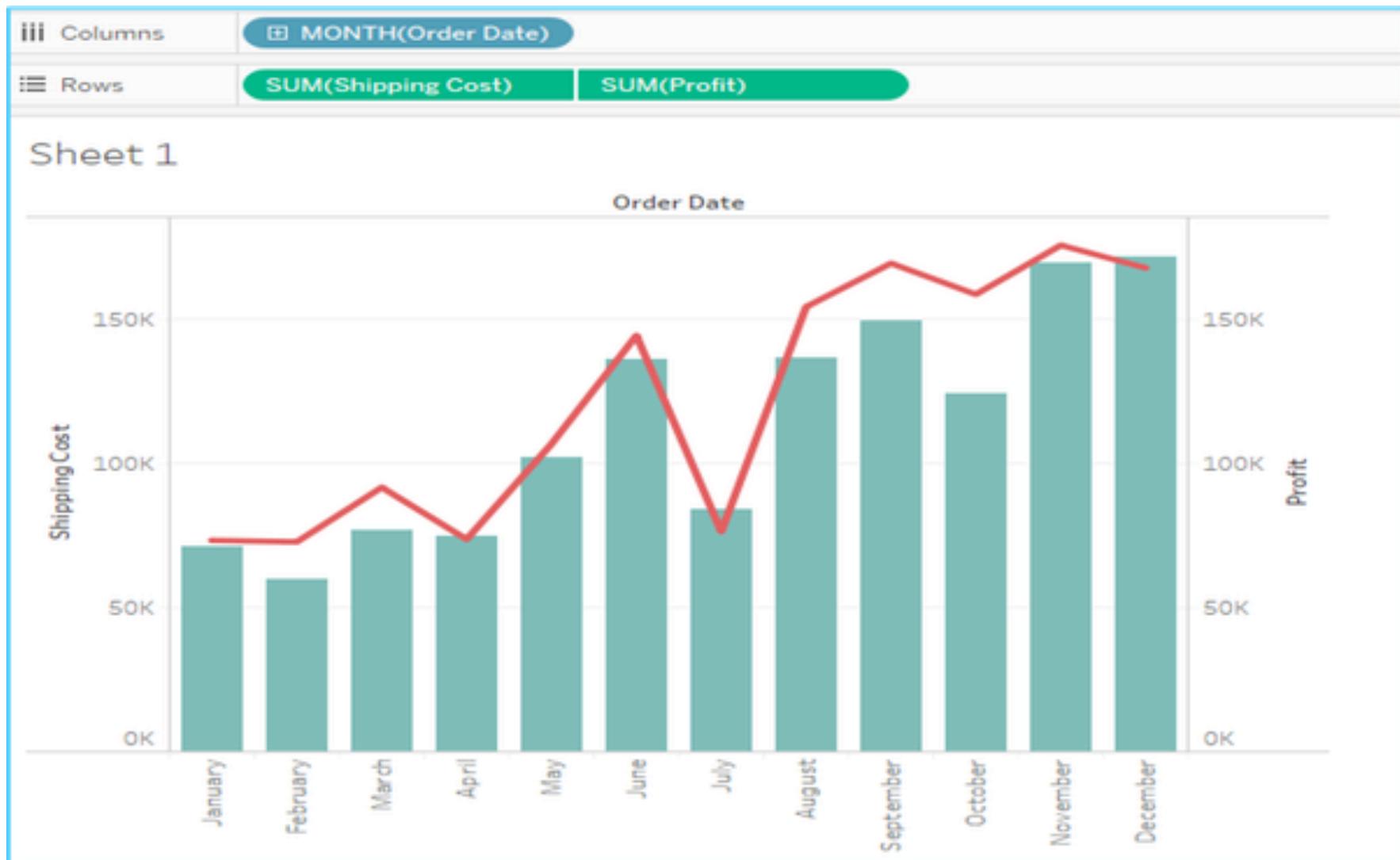
(b)

Empirical study comparing superimposed line charts to juxtaposed filled-area line charts. (a) Superimposed line charts performed best for tasks carried out within a local visual span. (b) Juxtaposed filled area charts were best for global tasks, especially as the number of time series increased.

LAYERING

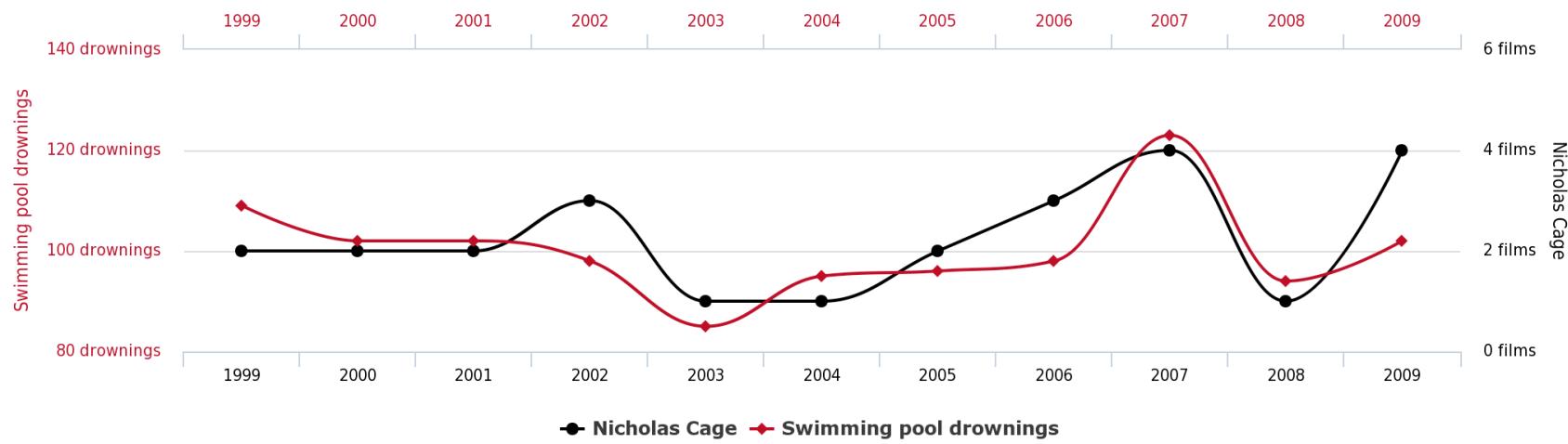


DUAL AXIS



DUAL AXIS (DON'T)

Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in

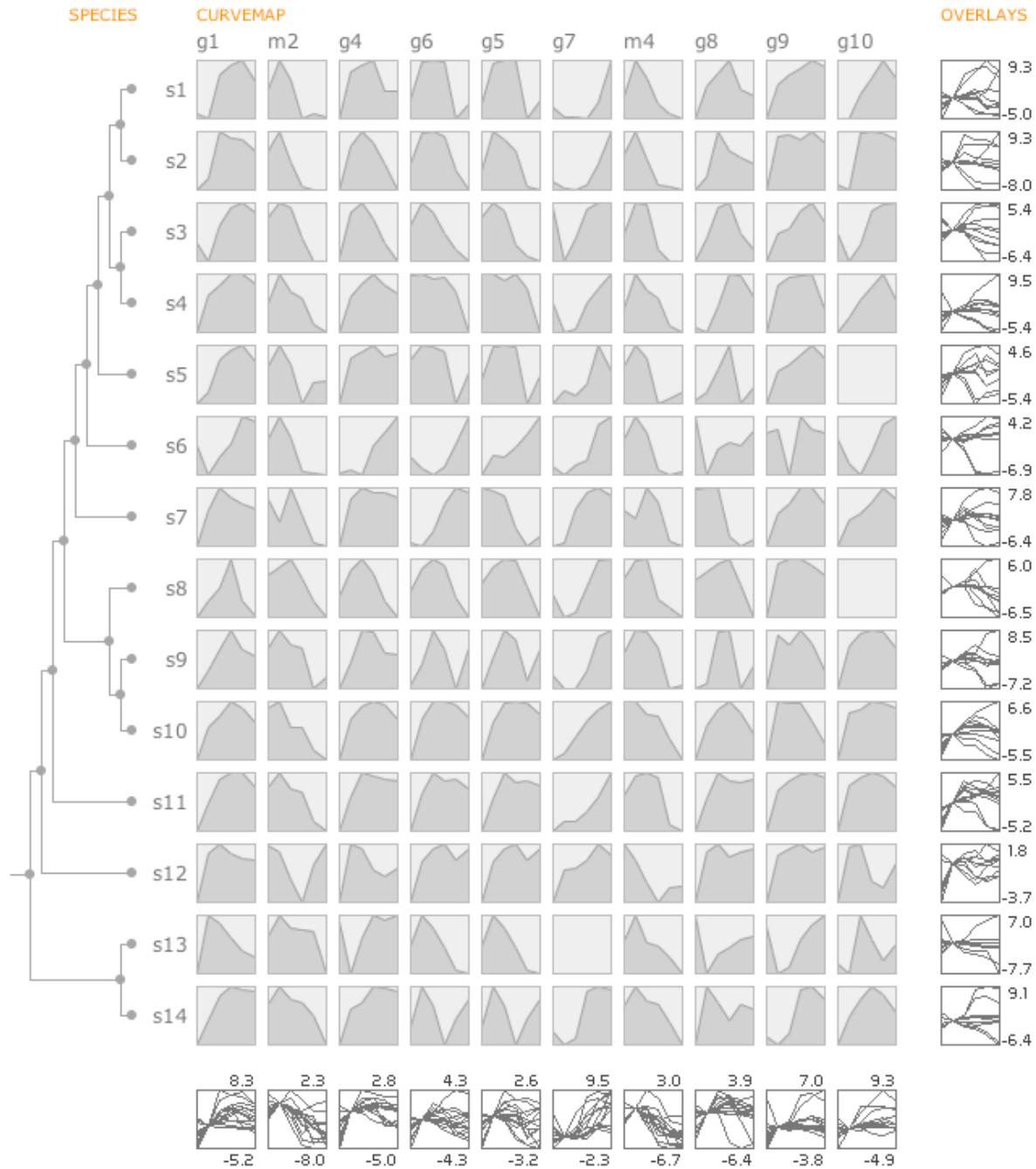


tylervigen.com

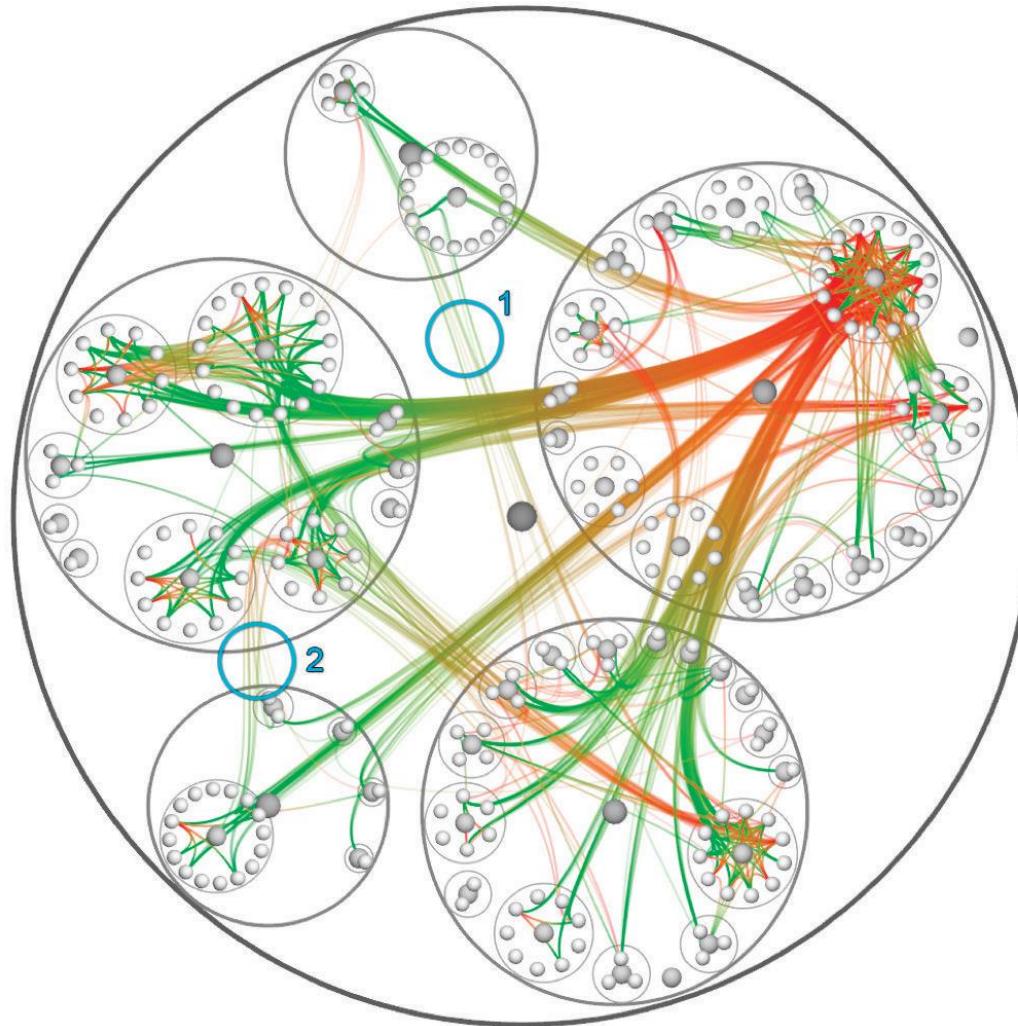
COMBINED

Partitioned +
layered graph

Synchronized
through
highlighting



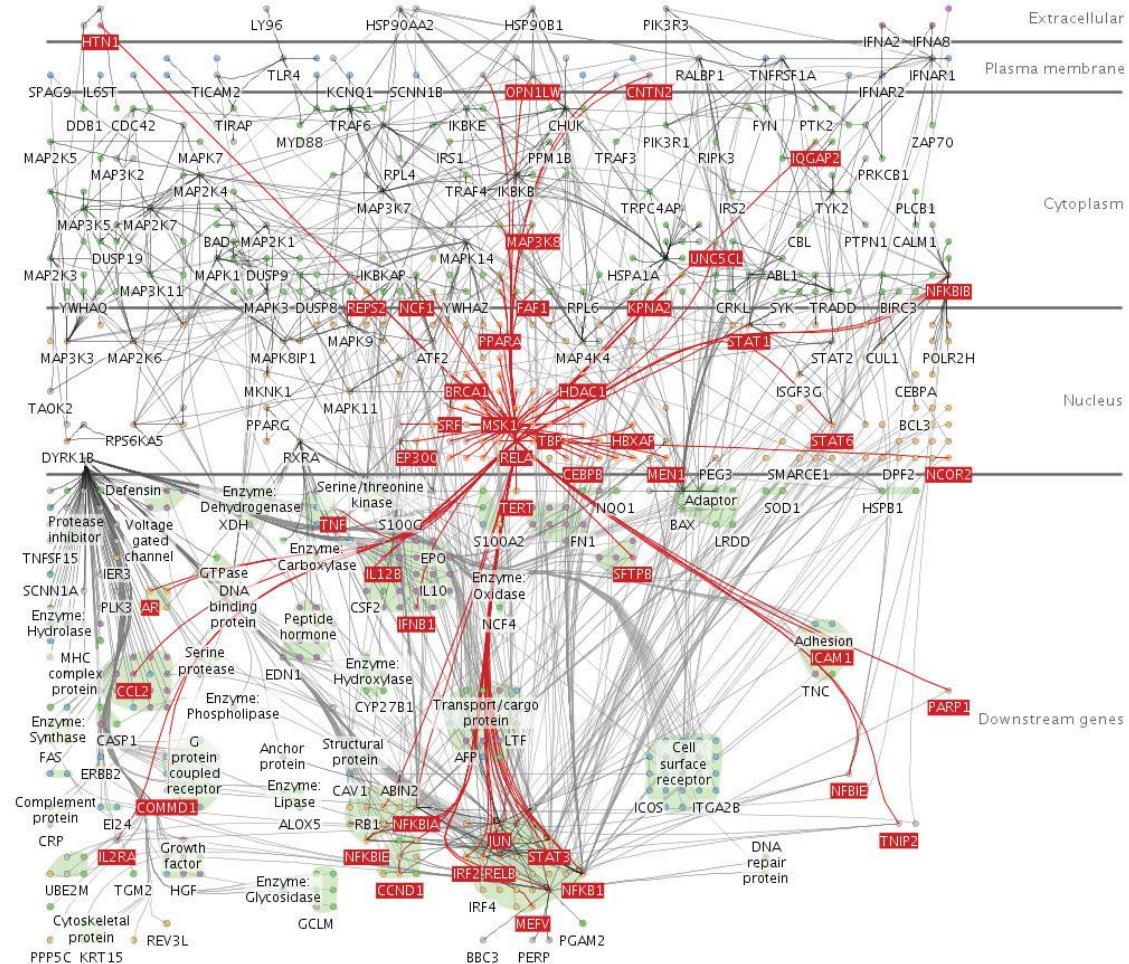
LAYERS



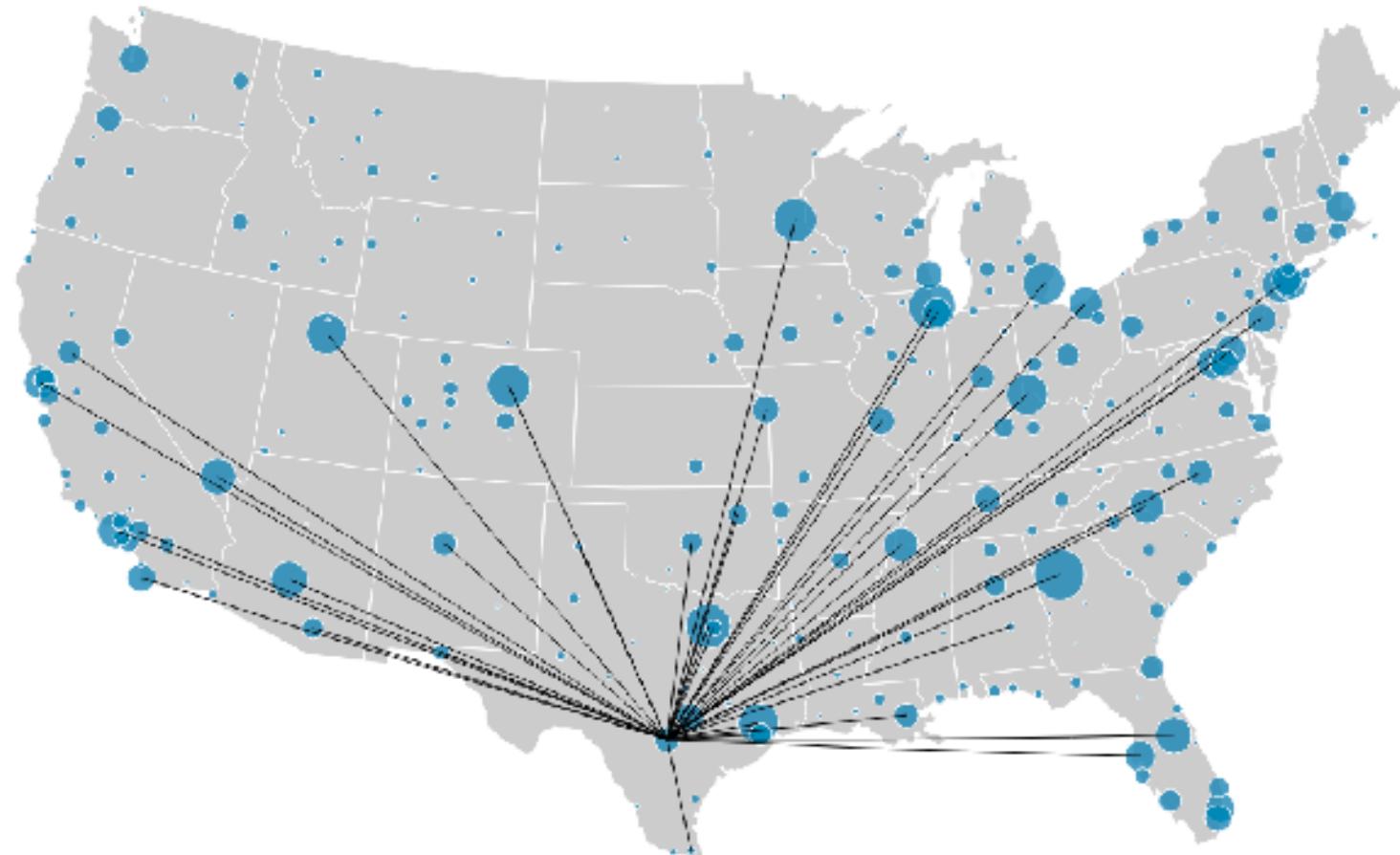
DYNAMIC LAYERS

in response to user selection.

The number of possible layers can be huge



DYNAMIC VISUAL LAYERING



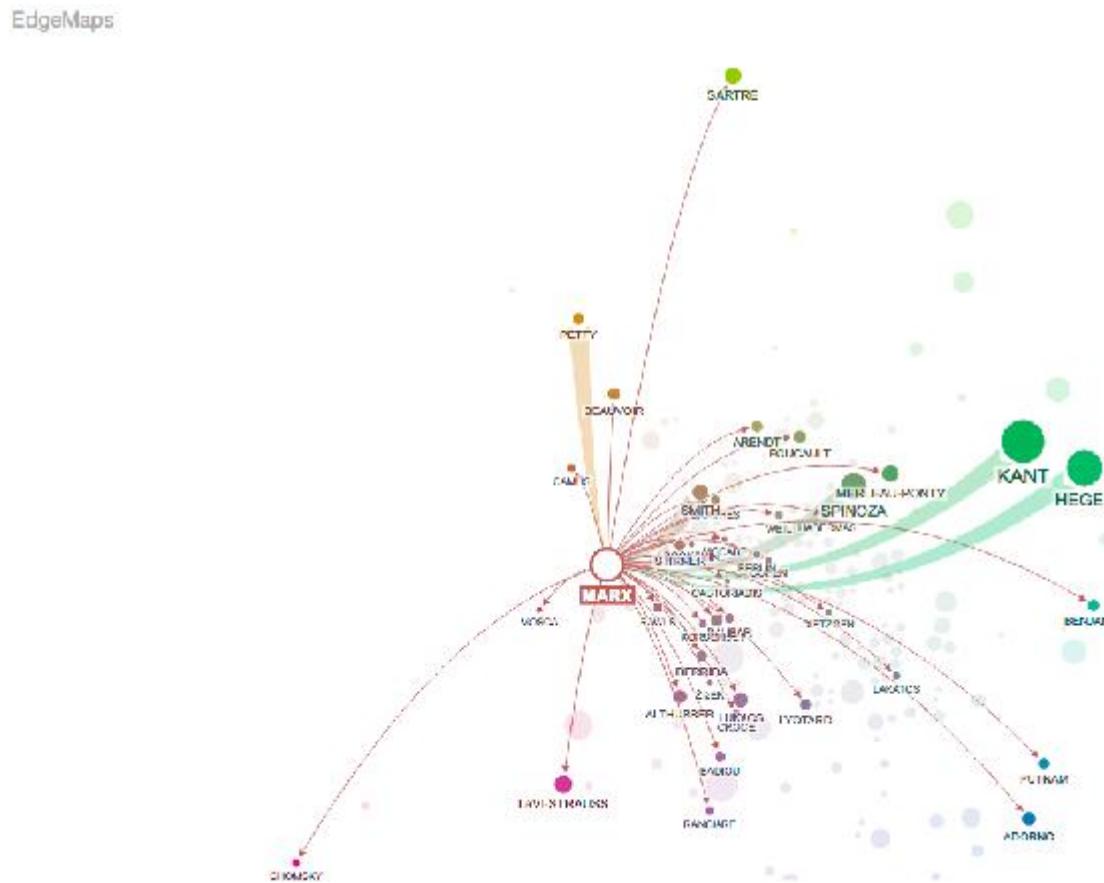
show Voronoi

San Antonio International, 2008
great arcs and symbol map

<http://mbostock.github.io/d3/talk/20111116/airports.html>

DYNAMIC VISUAL LAYERING

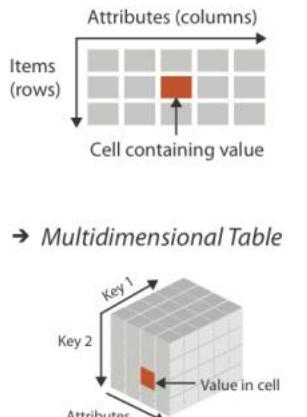
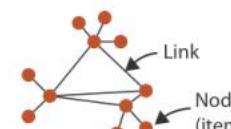
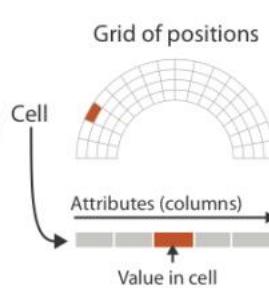
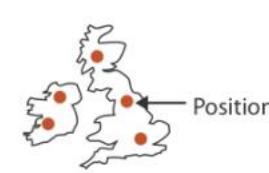
One-hop neighbor highlighting



<https://mariandoerk.de/edgemaps/demo/>

TABLES

DATASET TYPES

Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items Attributes	Items (nodes) Links Attributes	Grids Positions Attributes	Items Positions	Items
 <p>Attributes (columns) Items (rows) Cell containing value</p> <p>→ Multidimensional Table</p> <p>Key 1 Key 2 Attributes Value in cell</p>	 <p>→ Trees</p> <p>Link Node (item)</p>	 <p>Grid of positions Cell Attributes (columns) Value in cell</p>	 <p>Position</p>	

EXERCISE: SKETCH 2 WAYS TO VIS EACH TABLE

	Age	Best 100 m	Furthest Jump	Sex
Amy	16	13.2	5.2	F
Basil	18	12.4	4.2	F
Clara	14	14.1	2.5	F
Desmond	22	10.01	6.3	M
Charles	19	11.3	5.3	M

	BPM T1	BPM T2	BPM T3
Amy	90	130	150
Basil	70	110	109
Clara	60	140	141
Desmond	84	100	108
Charles	81	110	130

SCALE OF TABLES

Need different approaches for “normal” and “high-dimensional” tables

How many dimensions?

- ~ 50 – tractable with “just” visualization
- ~1000 – need analytical methods

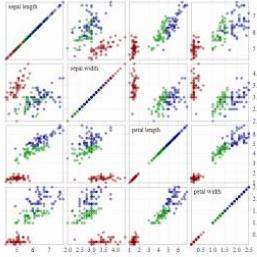
How many records?

- ~ 1000 – “just” vis is fine
- >> 10 000 – need analytical methods

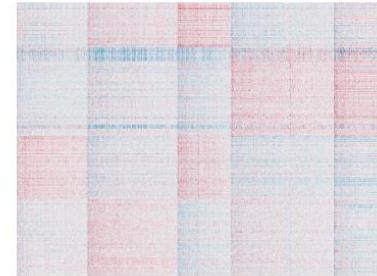
Homogeneity

- Same data type?
- Same scales?

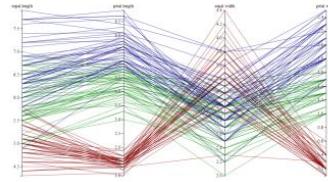
ANALYTIC COMPONENT



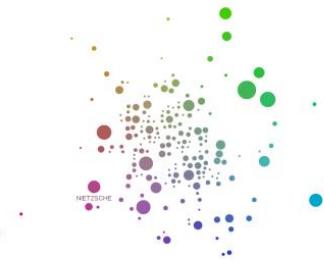
Scatterplot Matrices
[Bostock]



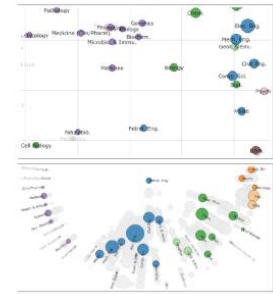
Pixel-based visualizations /
heat maps



Parallel Coordinates
[Bostock]



Multidimensional Scaling
[Doerk 2011]



[Chuang 2012]

no / little analytics

strong analytics
component

TECHNIQUES AND TASKS

Magnitude

Distribution

Ranking

Part to whole

Deviation

Correlation

Change over time



MAGNITUDE

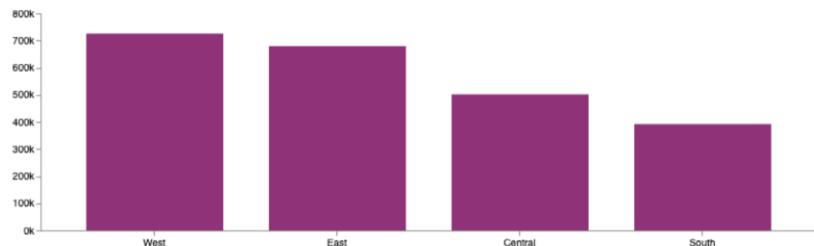
Show size comparison

- Relative: to see bigger/larger
- Absolute: to see differences

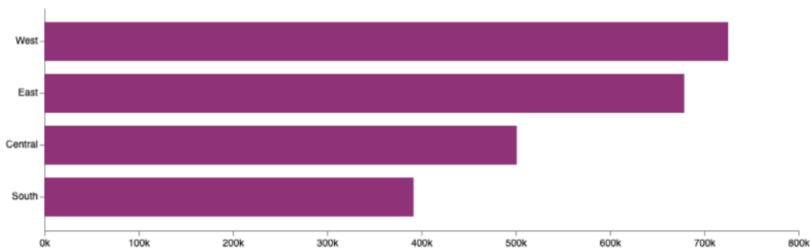
Usually for “counted” numbers

Rather than a calculated rate or per cent

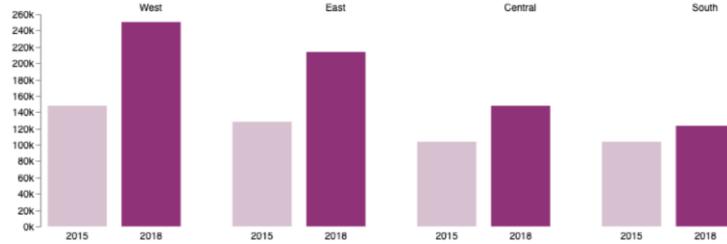
BAR CHARTS



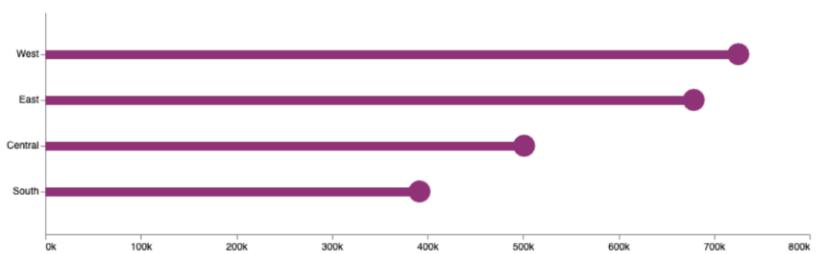
Vertical Bar Chart / Column Chart



Horizontal Bar Chart

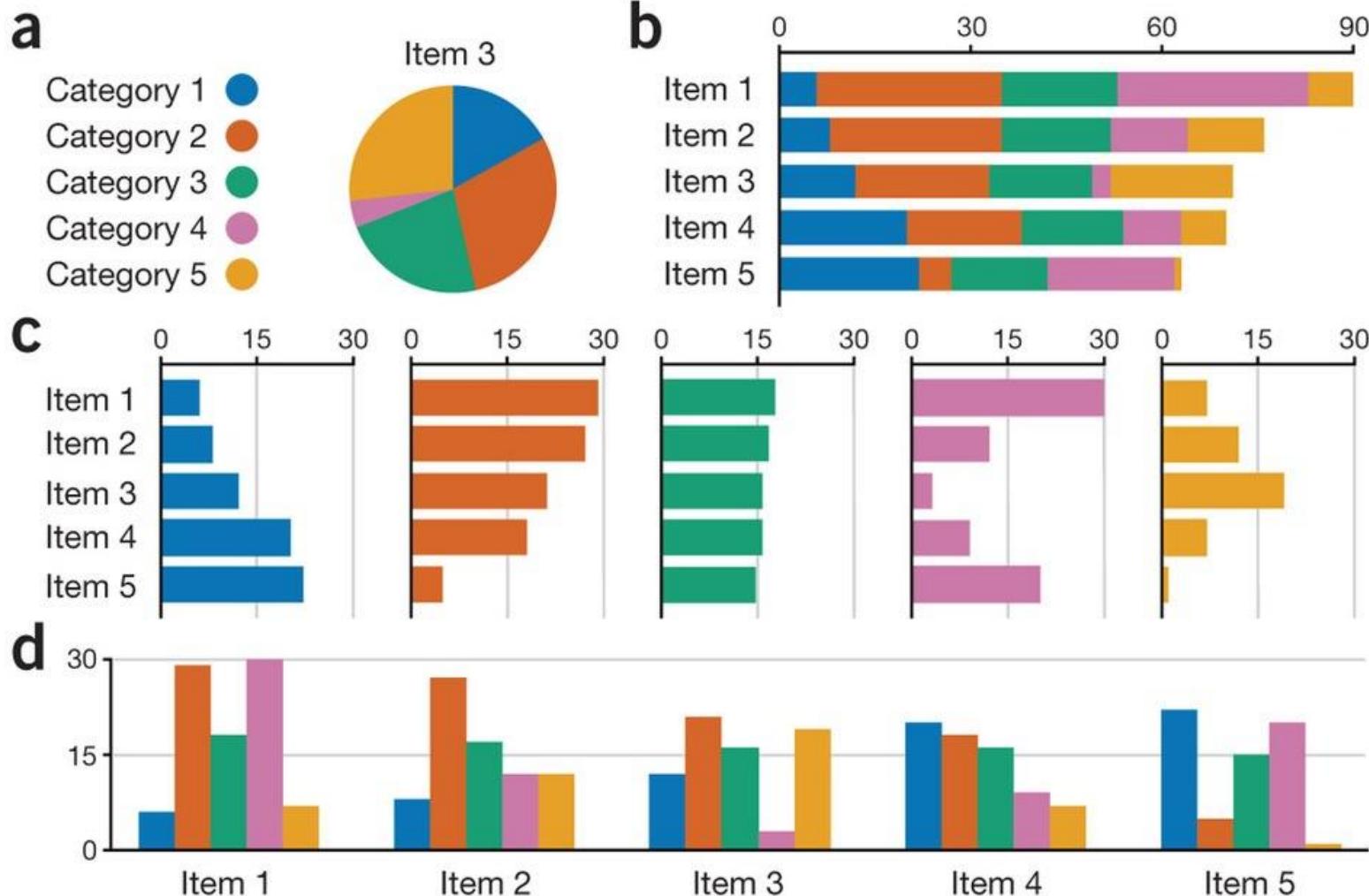


Grouped Bar Chart

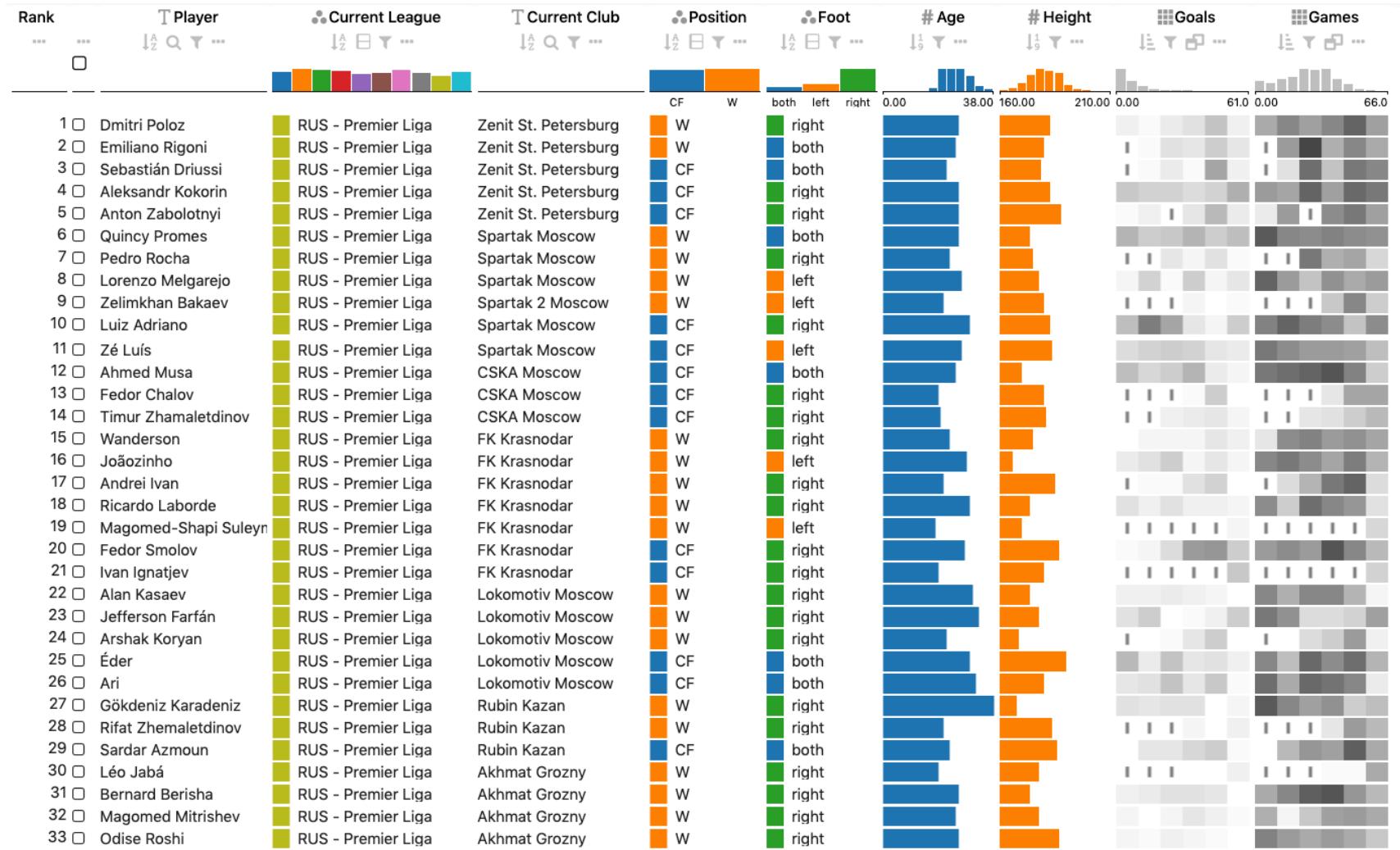


Lollipop Chart

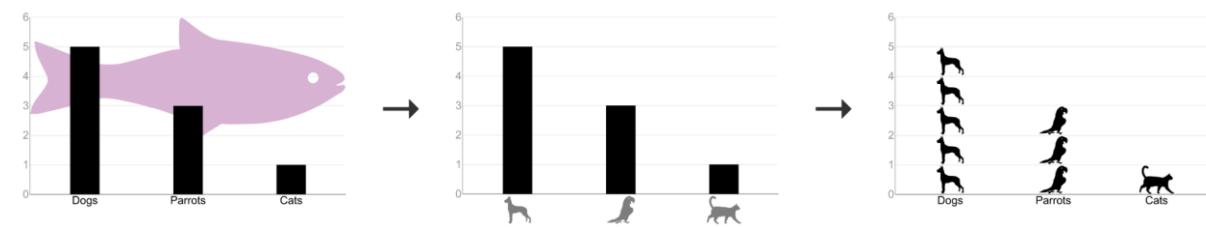
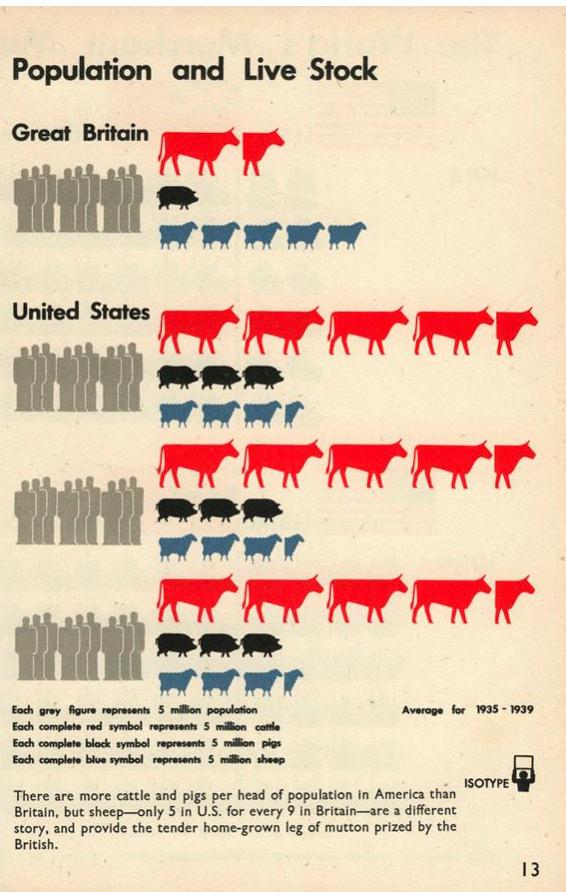
BAR CHARTS TYPES



BAR CHARTS IN COMPLEX VIS



ISO TYPE VISUALIZATION



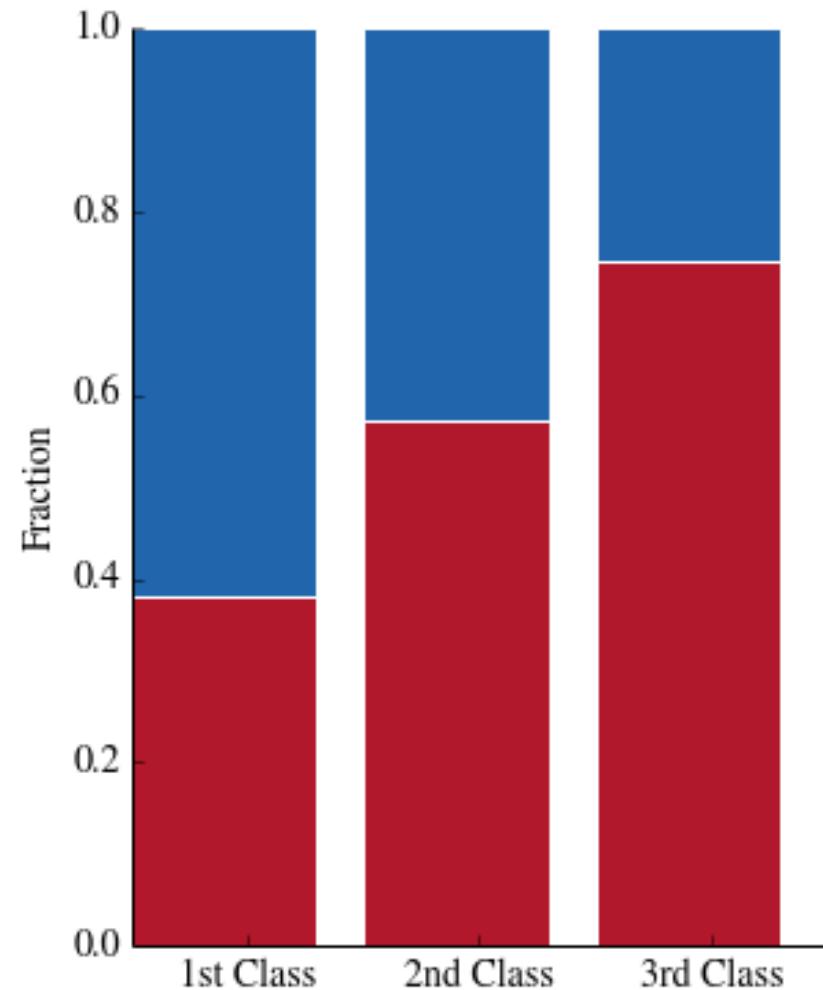
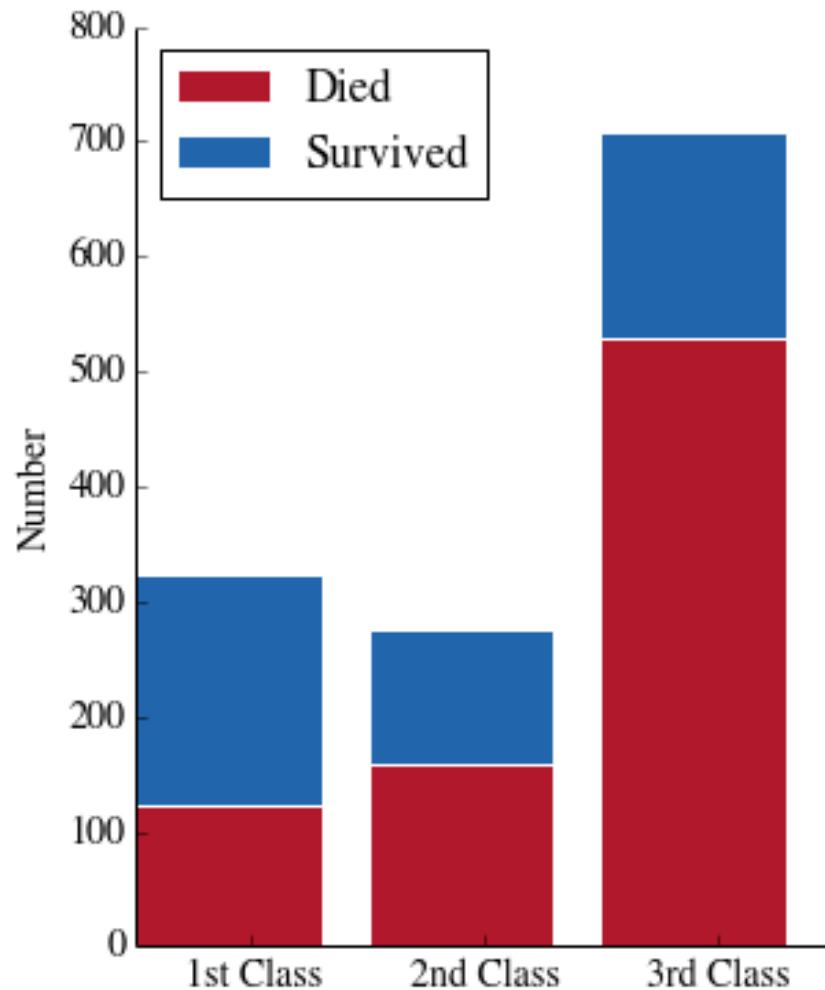
PART OF WHOLE

Show how a single entity can be broken down into its component elements

If the interest is solely in the size of the components

Consider magnitude-type chart

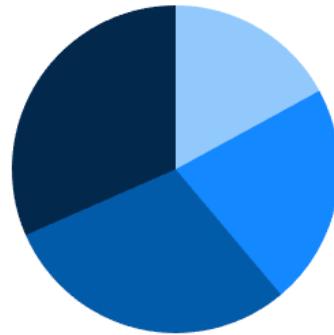
STACKED BAR CHART



PIE AND DONUT CHARTS

Pie

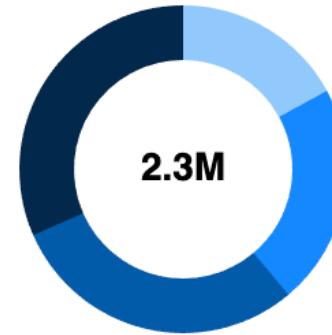
A common way of showing part-to-whole data - but be aware that it's difficult to accurately compare the size of the segments.



Edit

Donut

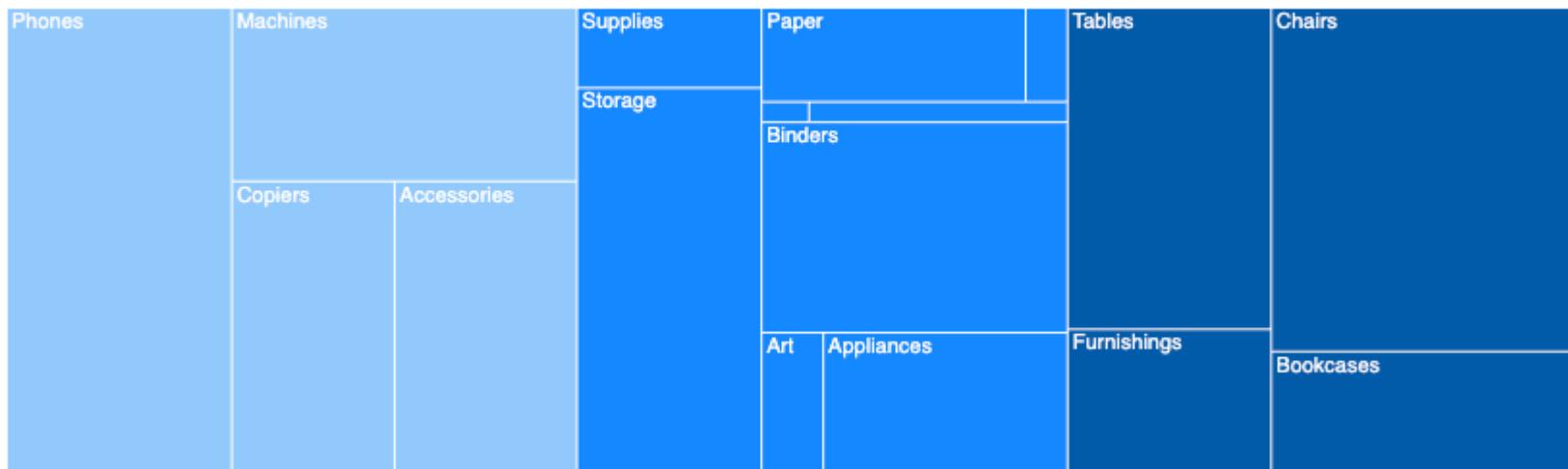
Similar to a pie chart - but the centre can be a good way of making space to include more information about the data (eg. total)



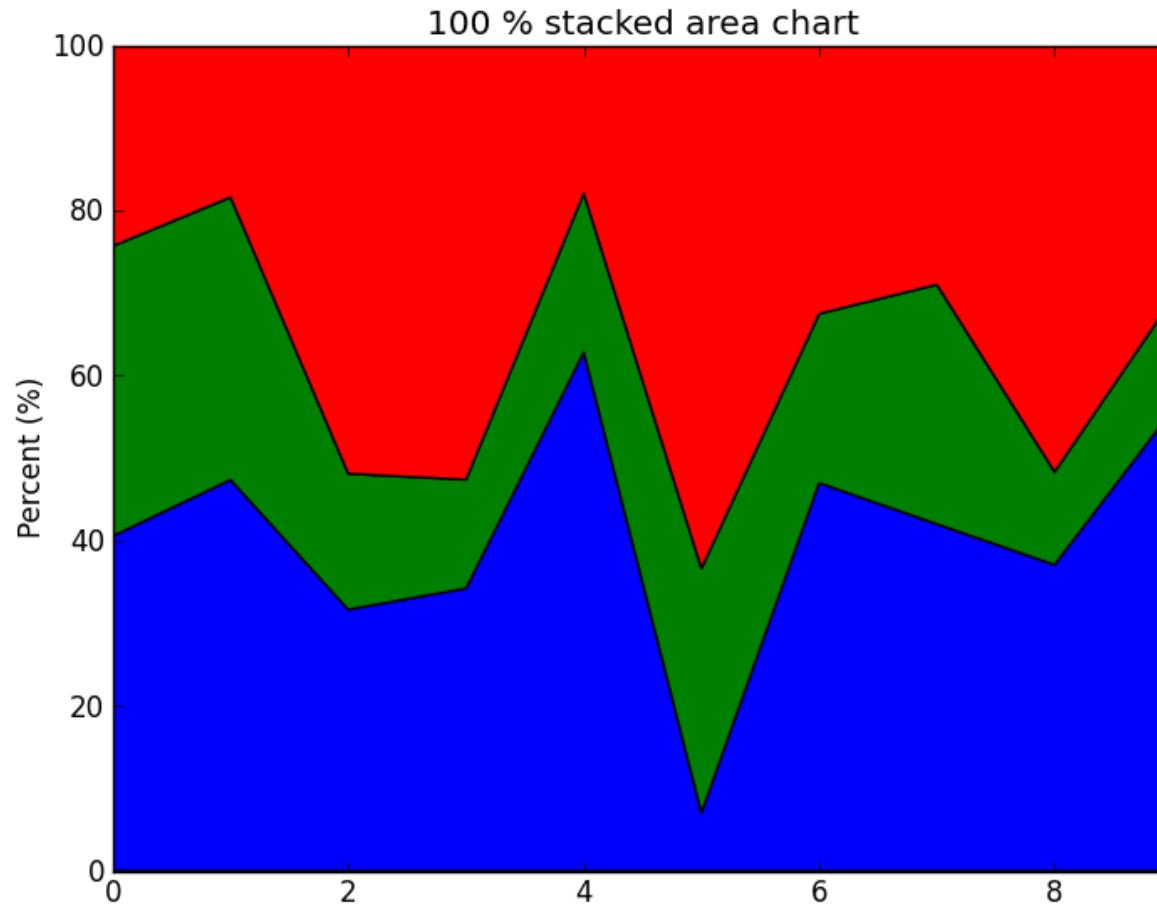
TREEMAP

Treemap

Use for hierarchical part-to-whole relationships; can be difficult to read when there are many small segments



STACKED AREA



DISTRIBUTION

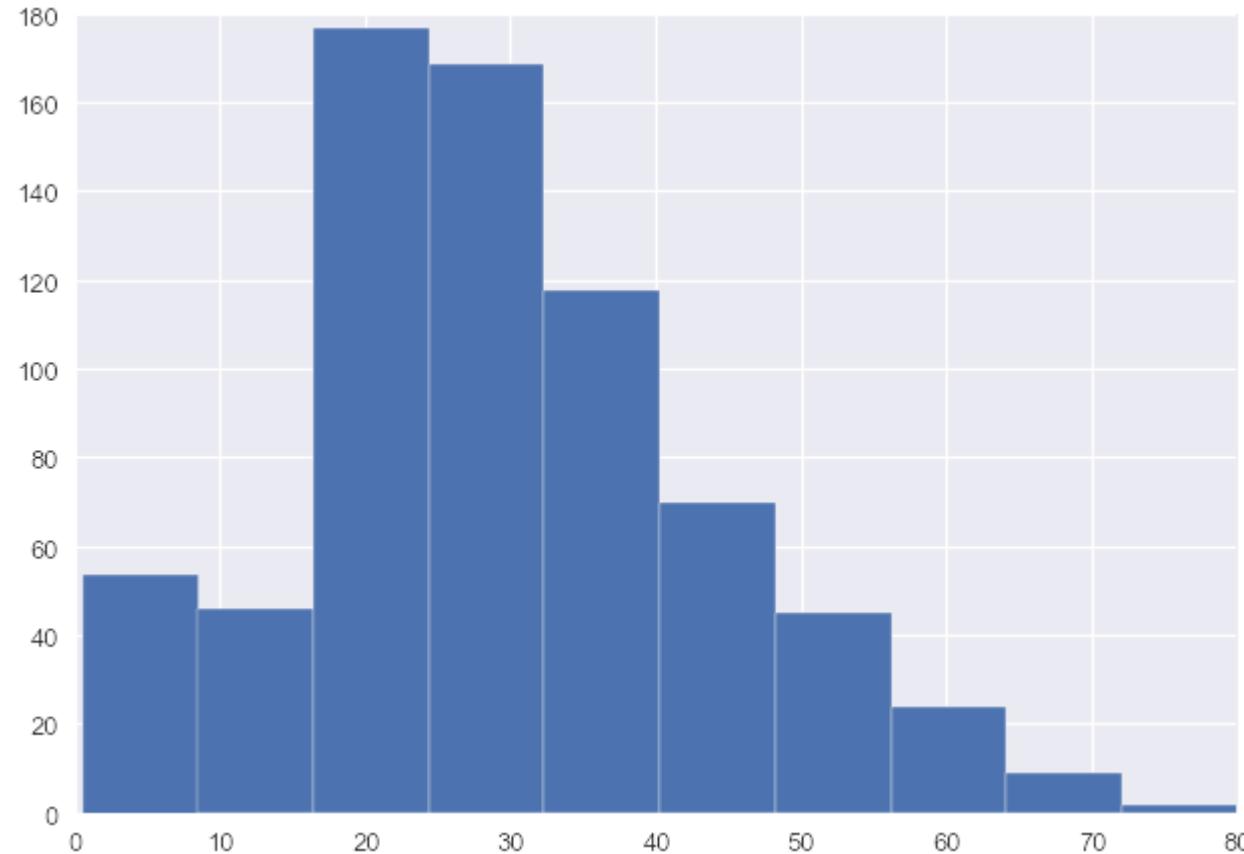
Aggregating large data vectors

Instead of showing all data points, show a data's distribution

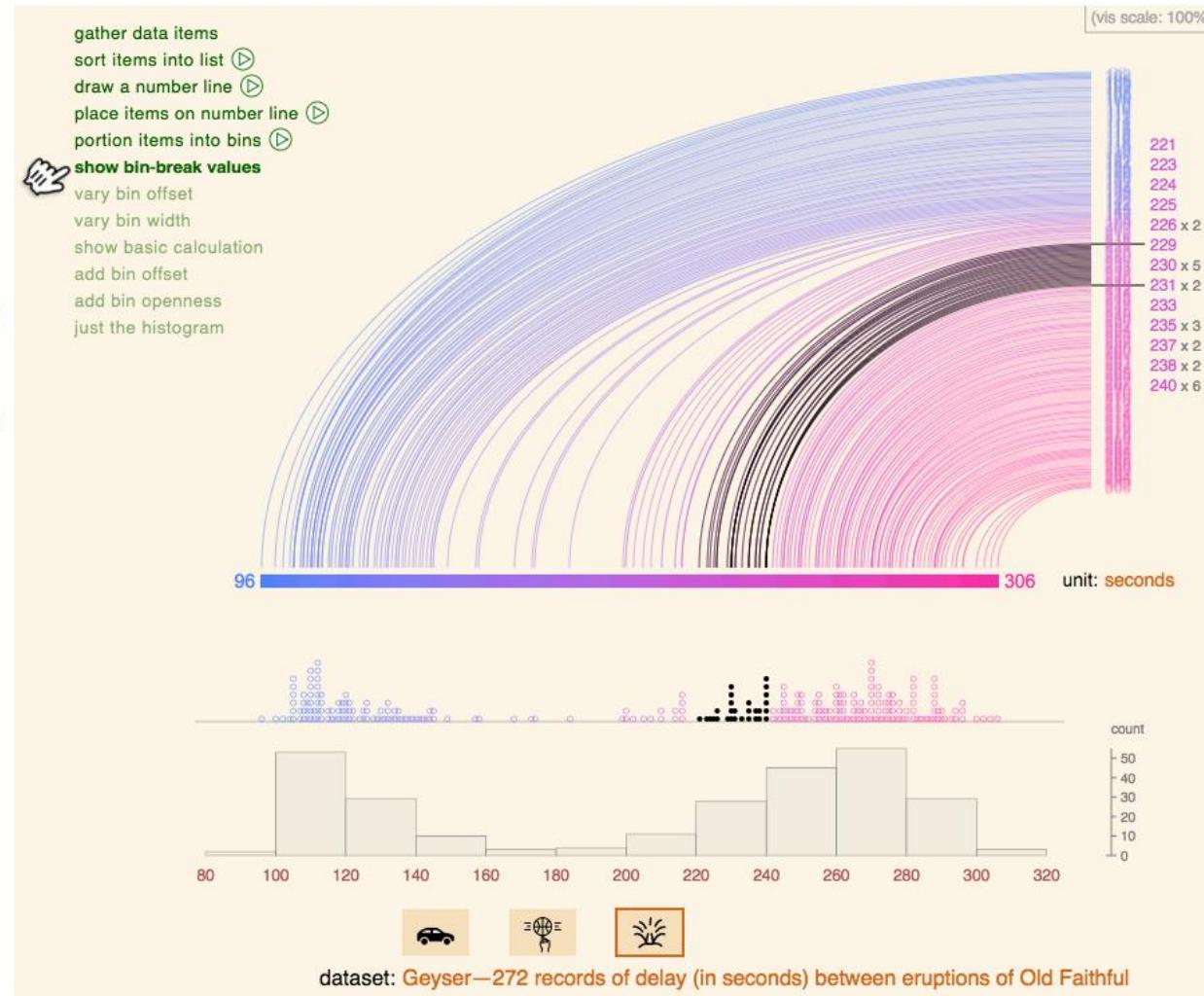
Pro: Compact representation

Con: Works only if data is “well behaved” for the type of distribution visualization

WHAT IS A HISTOGRAM



WHY HISTOGRAM IS HARD



<http://tinlizzie.org/histograms/>

HISTOGRAM

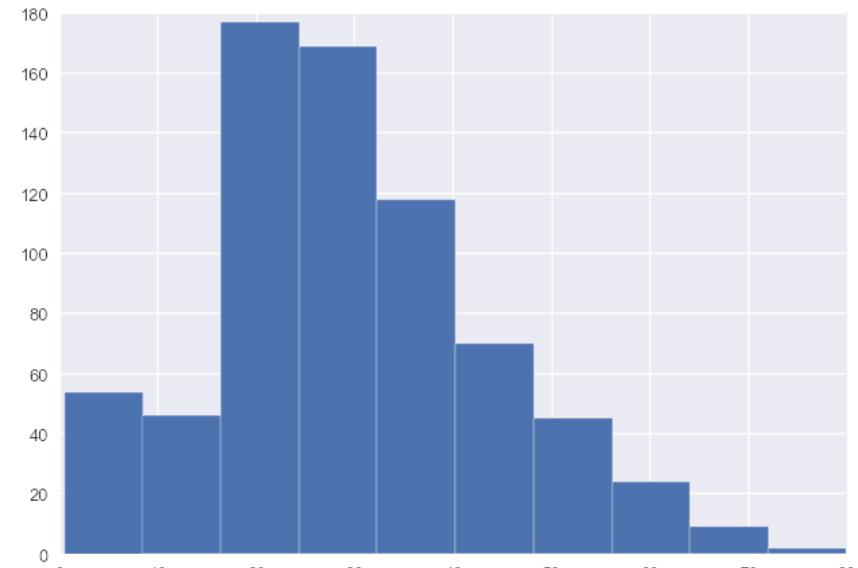
Good #bins hard to predict

Make interactive

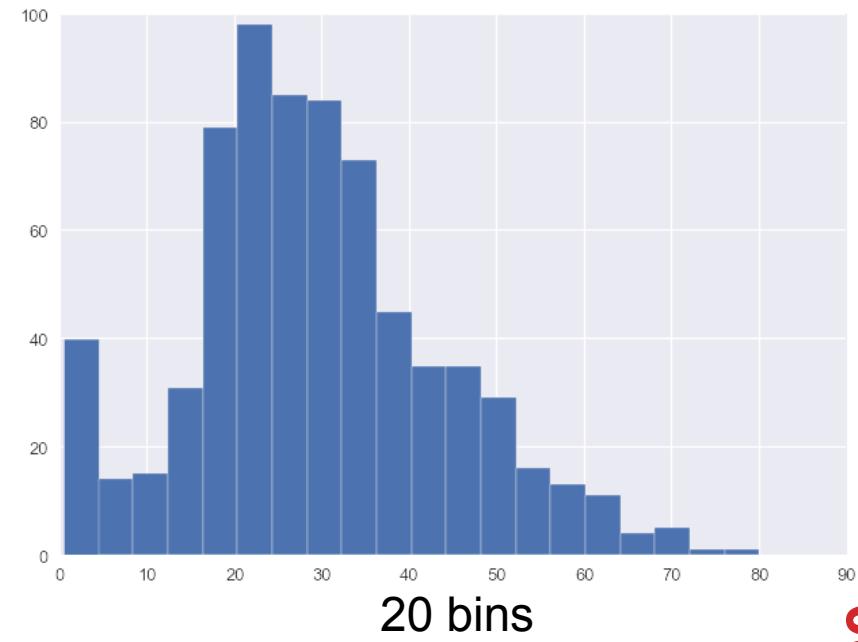
Rules of thumb:

$$\# \text{bins} = \sqrt{n}$$

$$\# \text{bins} = \log_2(2) + 1$$

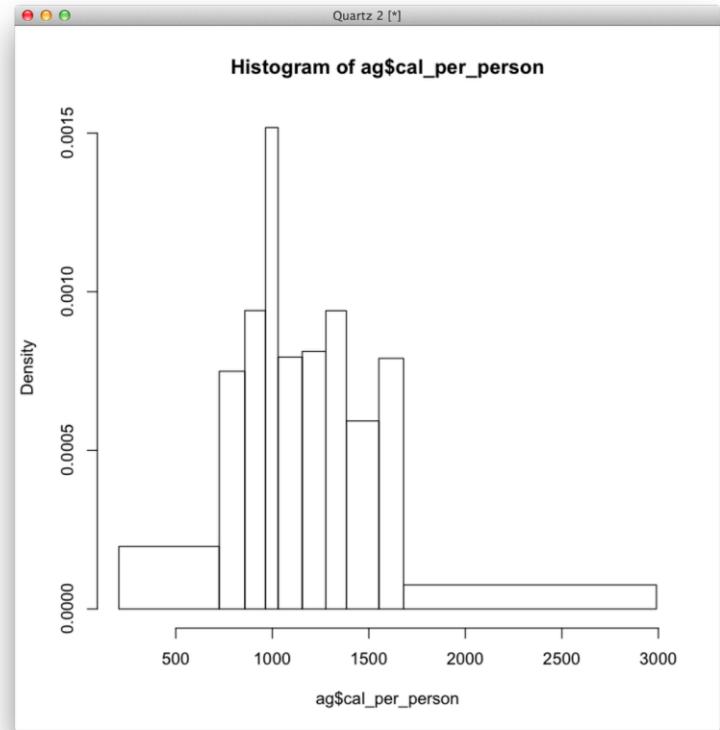
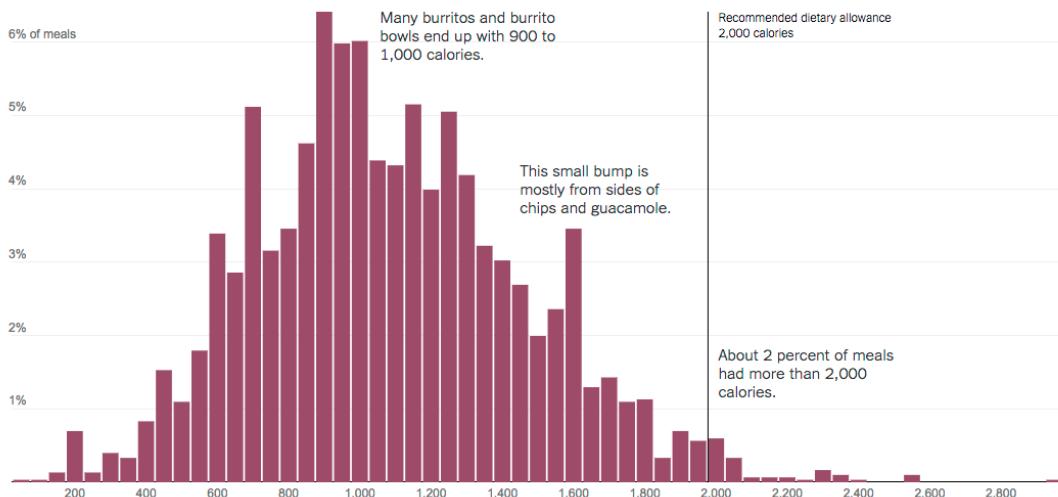


10 bins



20 bins

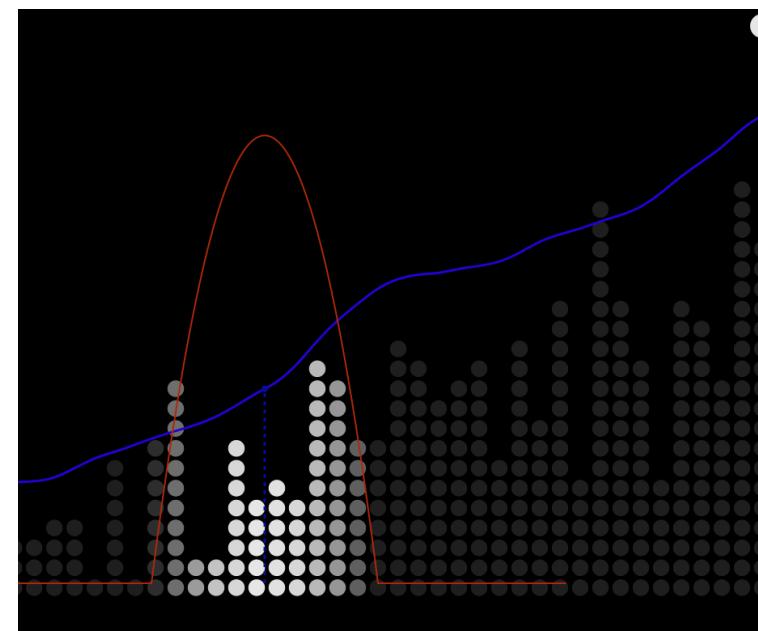
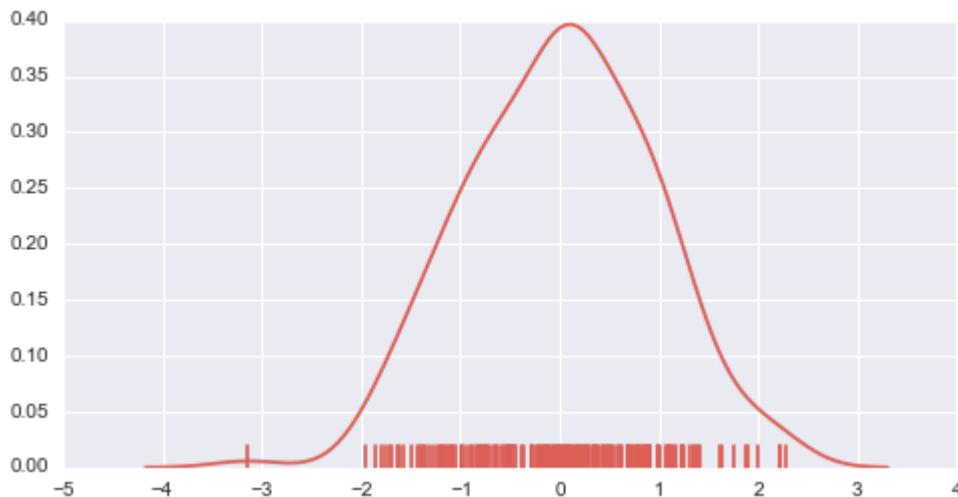
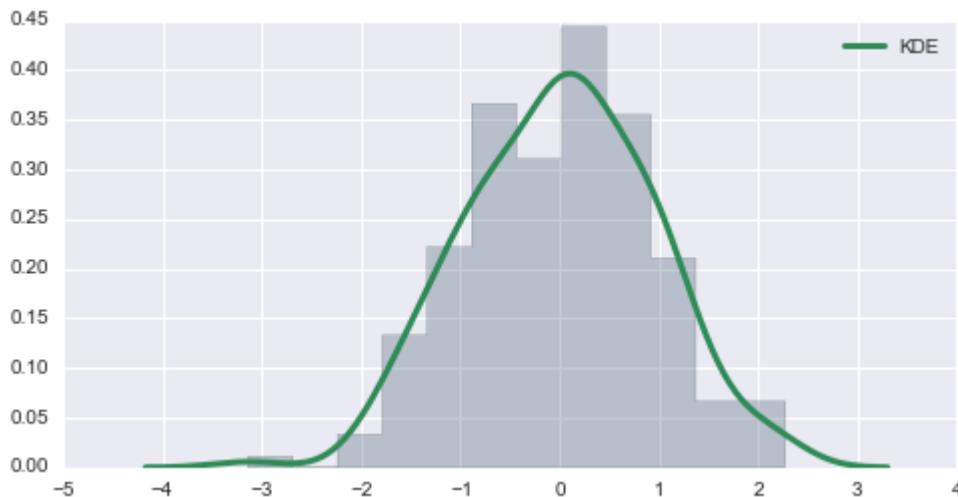
UNEQUAL BIN WIDTH



Can be useful if data is much sparser in some areas than others
Show density as area, not height

https://www.nytimes.com/interactive/2015/02/17/upshot/what-do-people-actually-order-at-chipotle.html?_r=1

DENSITY PLOTS (KERNEL DENSITY ESTIMATION)



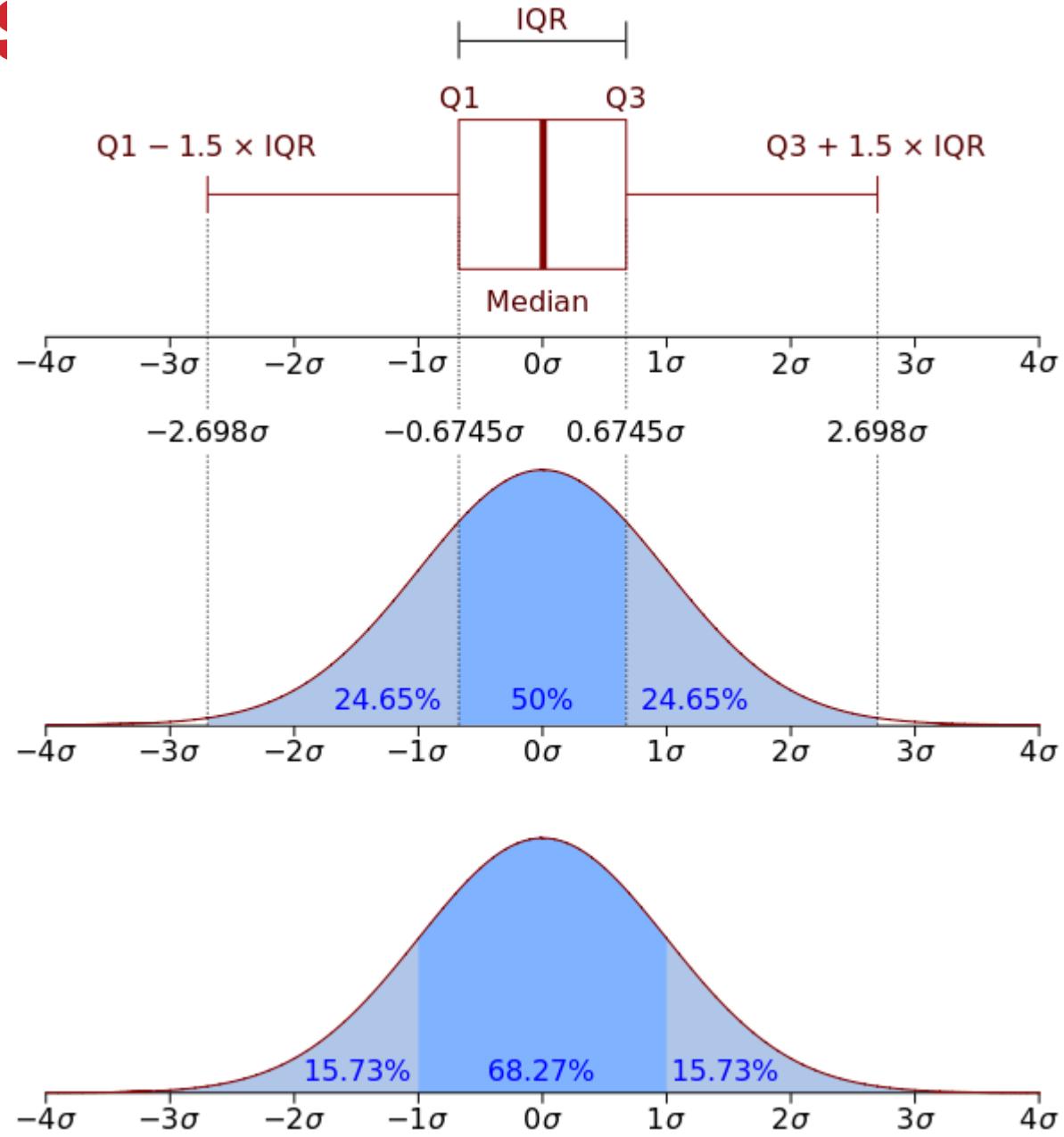
<https://mathisonian.github.io/kde/>

BOX PLOTS

aka Box-and-Whisker Plot

Show outliers as points!

Bad for non-normal distributed data
Especially bad for bi- or multimodal distributions



ONE BOXPLOT, FOUR DISTRIBUTIONS

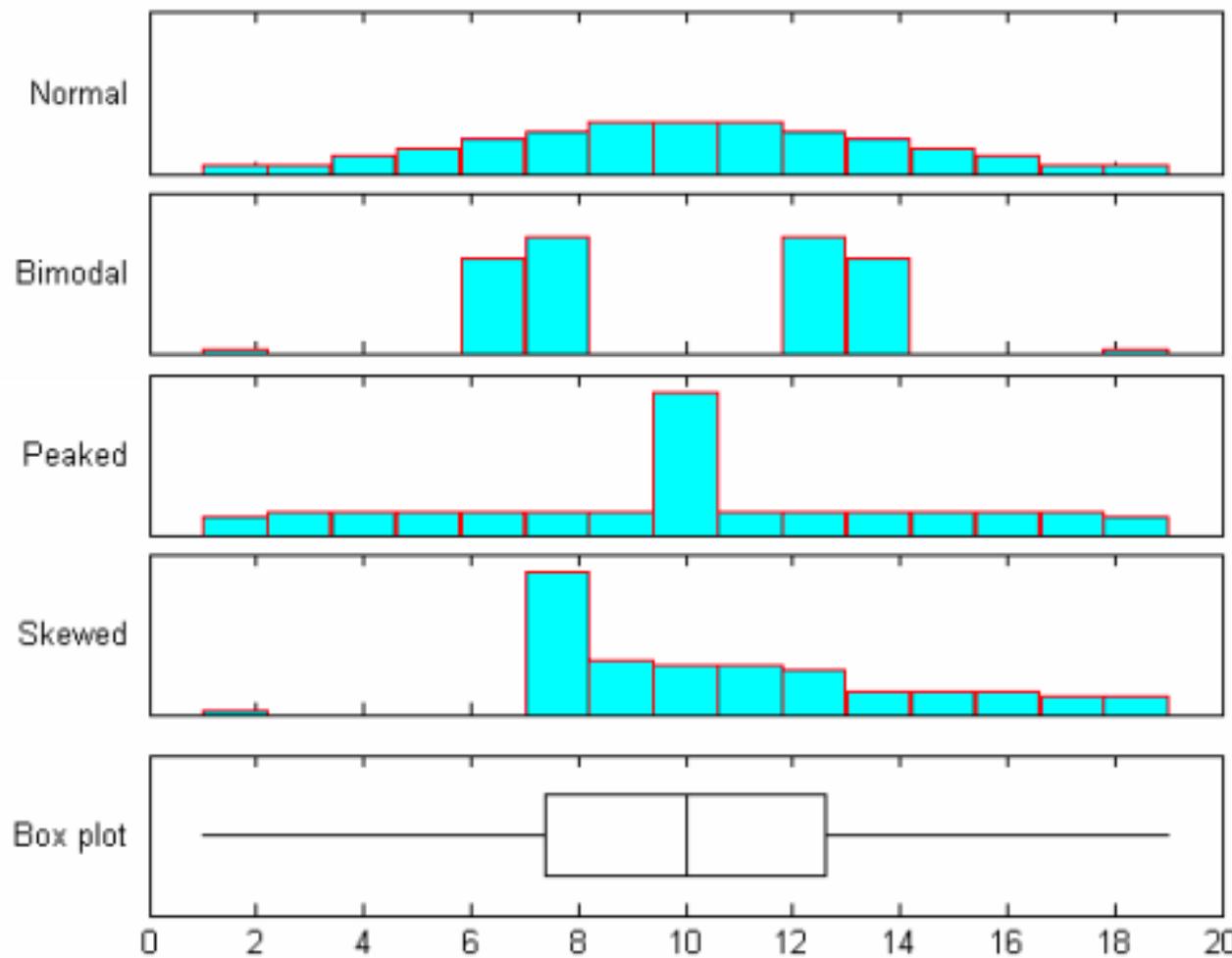
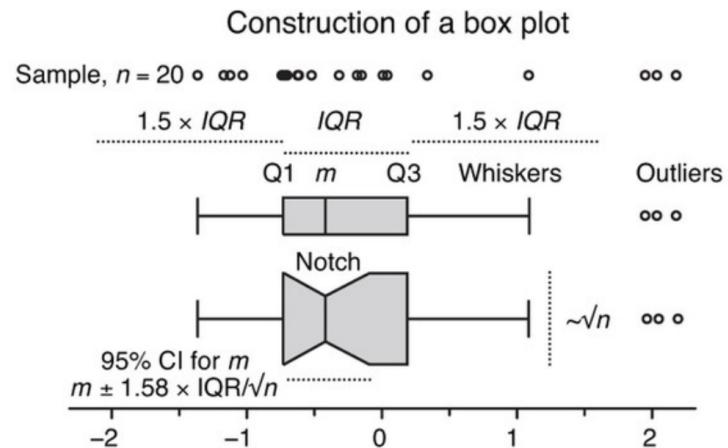


Figure 1: Histograms and box plot: four samples each of size 100

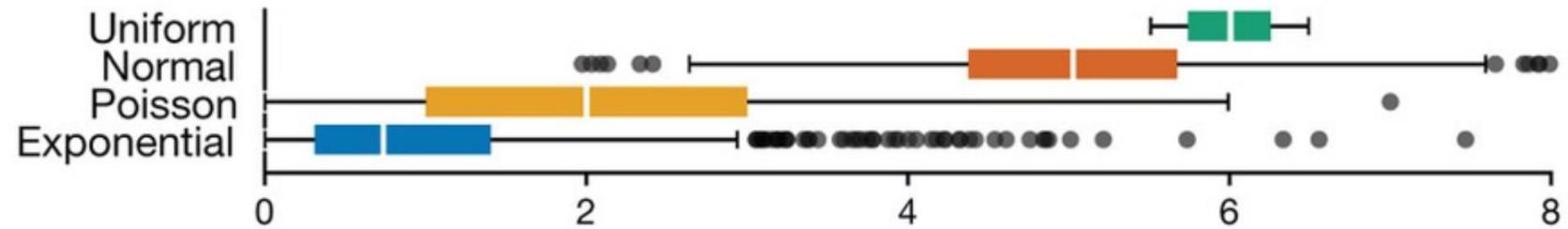
NOTCHED BOX PLOTS

Notch shows
 $m \pm 1.5 \times IQR/\sqrt{n}$
-> 95% Confidence Intervall

A guide to statistical significance.



BOX PLOTS

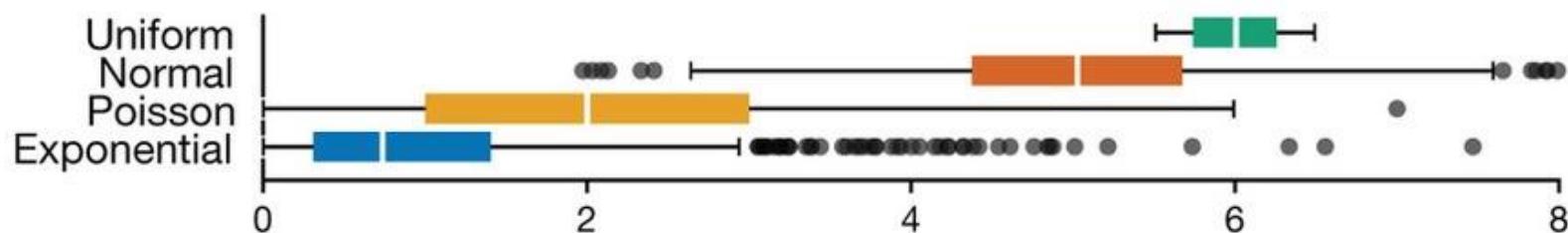


COMPARISON

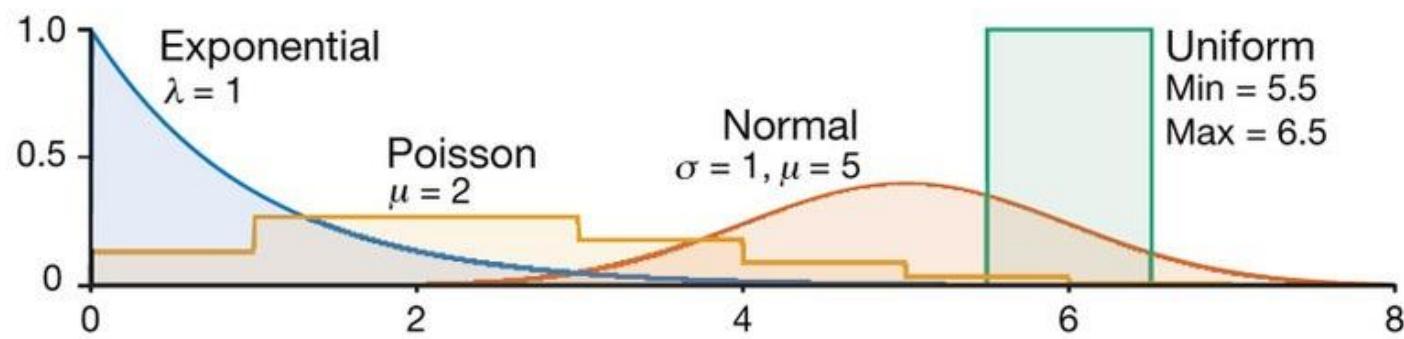
a



b

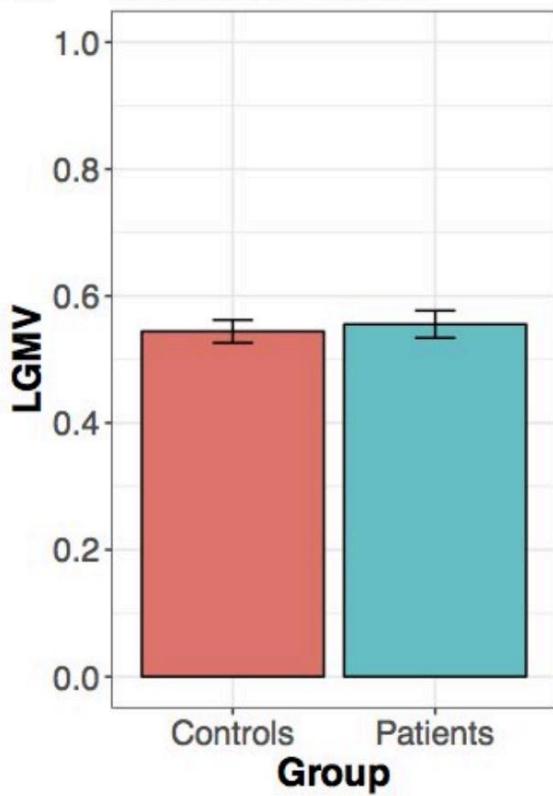


c

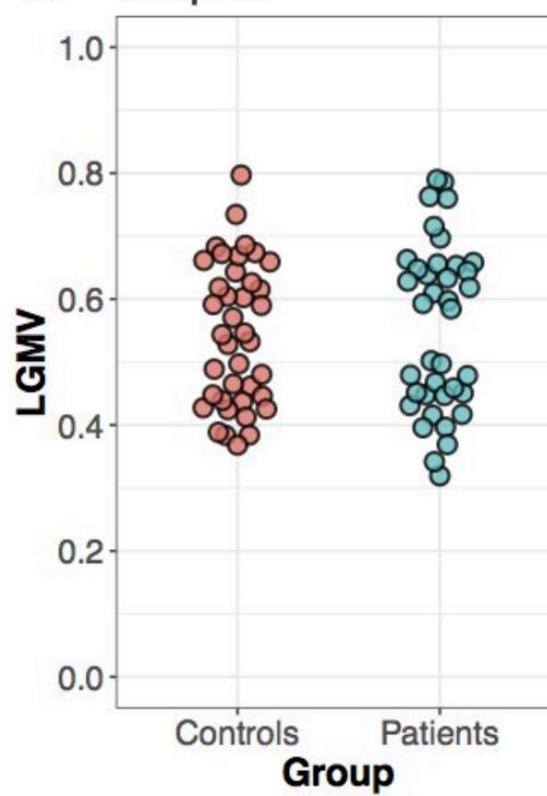


BAR CHARTS VS DOTPLOTS

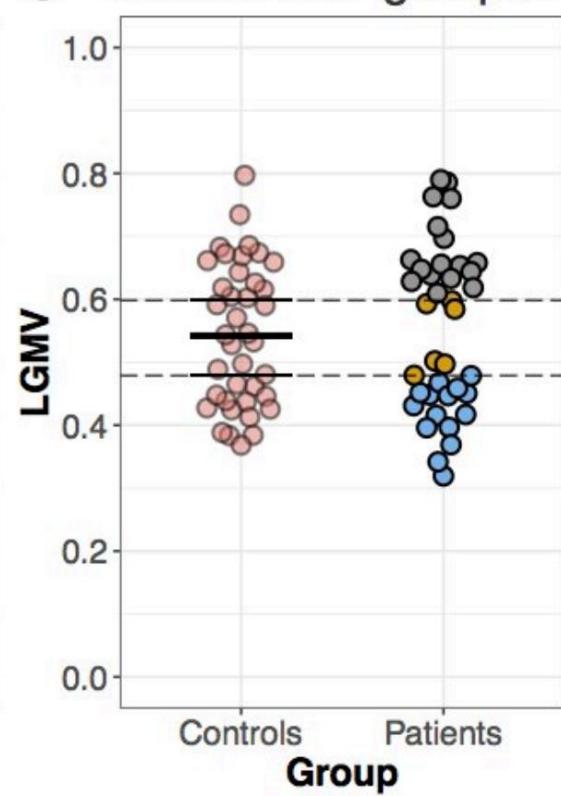
A Mean +/- SEM



B Dotplot

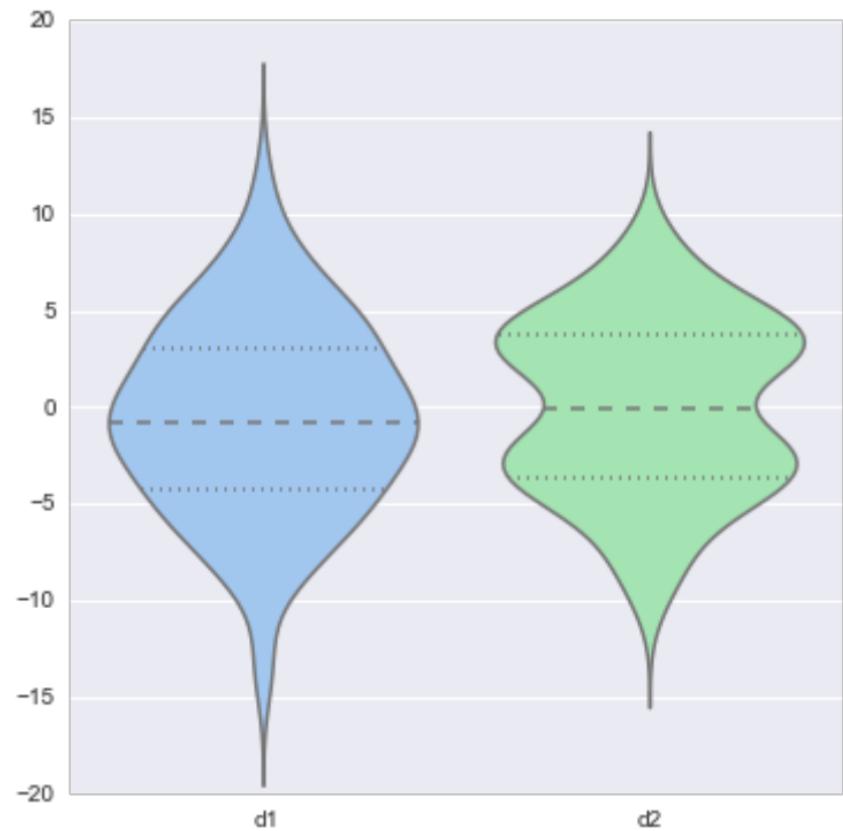
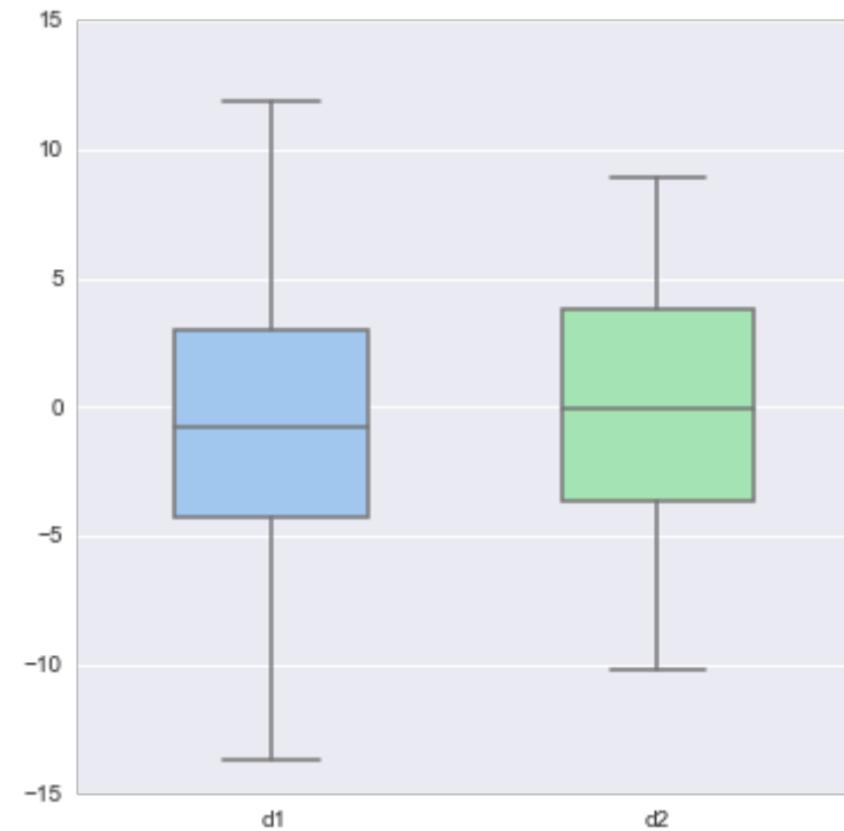


C Patient sub-groups?



VIOLIN PLOT

Violin plot = box plot + probability density function



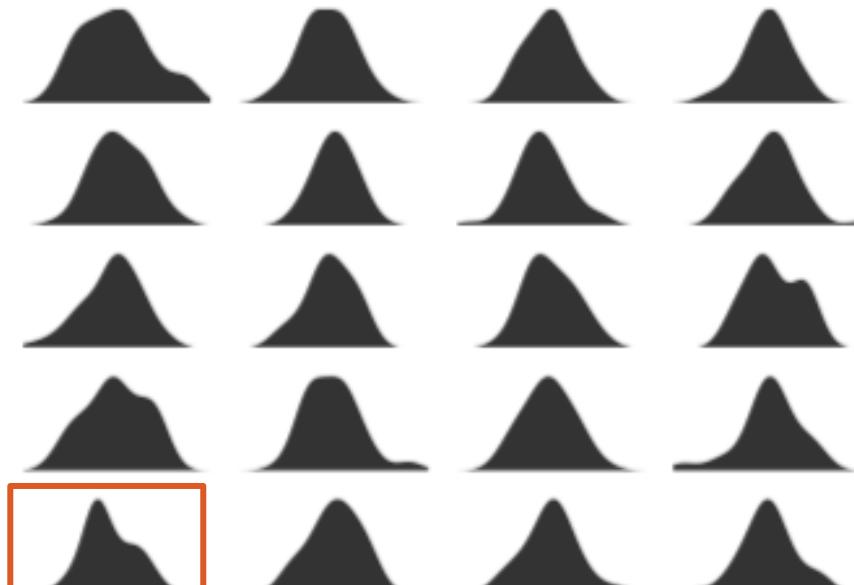
DIFFERENT DISTRIBUTIONS



<https://blog.bioturing.com/2018/05/16/5-reasons-you-should-use-a-violin-graph/>

DATA FLAWS

19 charts are random samples from a gaussian.
1 chart has 20% of samples with identical value



(a)

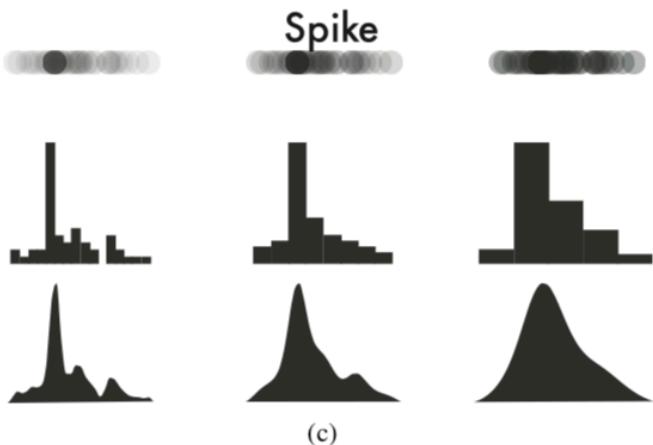
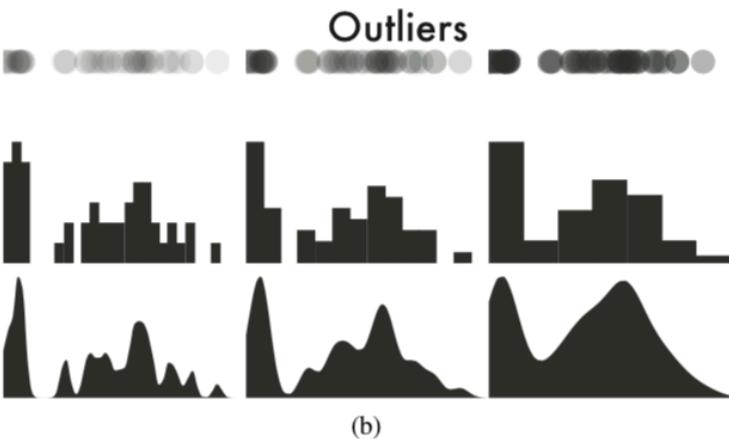
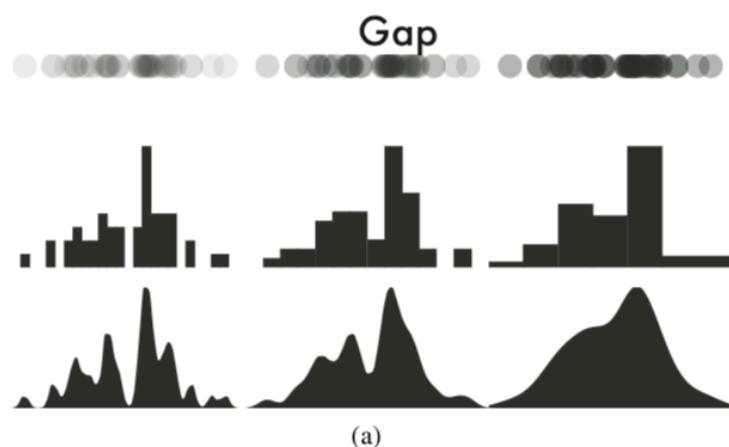


(b)

DETECTING DATA FLAWS

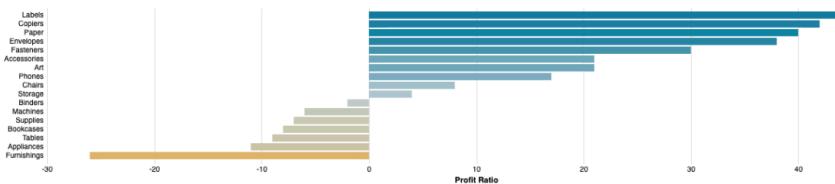
Tricky with aggregate visualization

Bin size / kernel type /
bandwidth / visualization
choice all affect different
situations

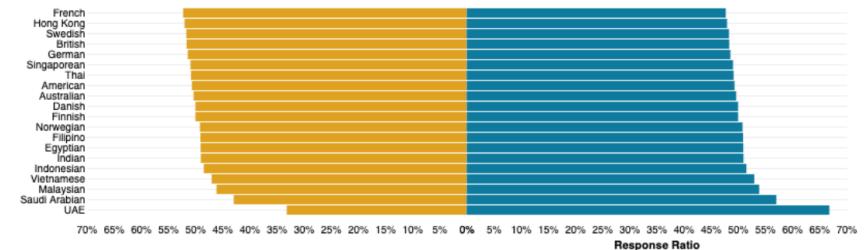


DEVIATION

COMPARISON TO REFERENCE POINT



Diverging Bar Chart



Juxtaposing Two Variables (male/female)

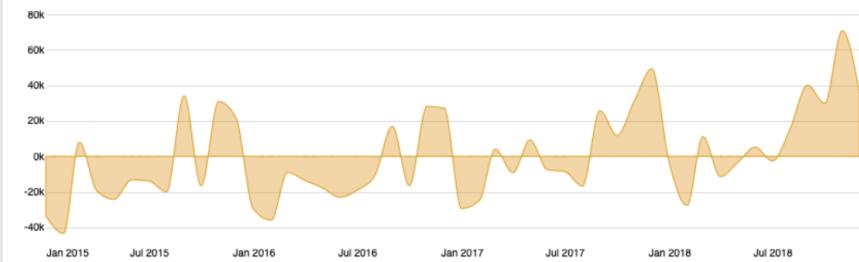
Surplus/deficit filled line

The shaded area of these charts allows a balance to be shown; either against a baseline or between two series



Surplus/deficit filled area

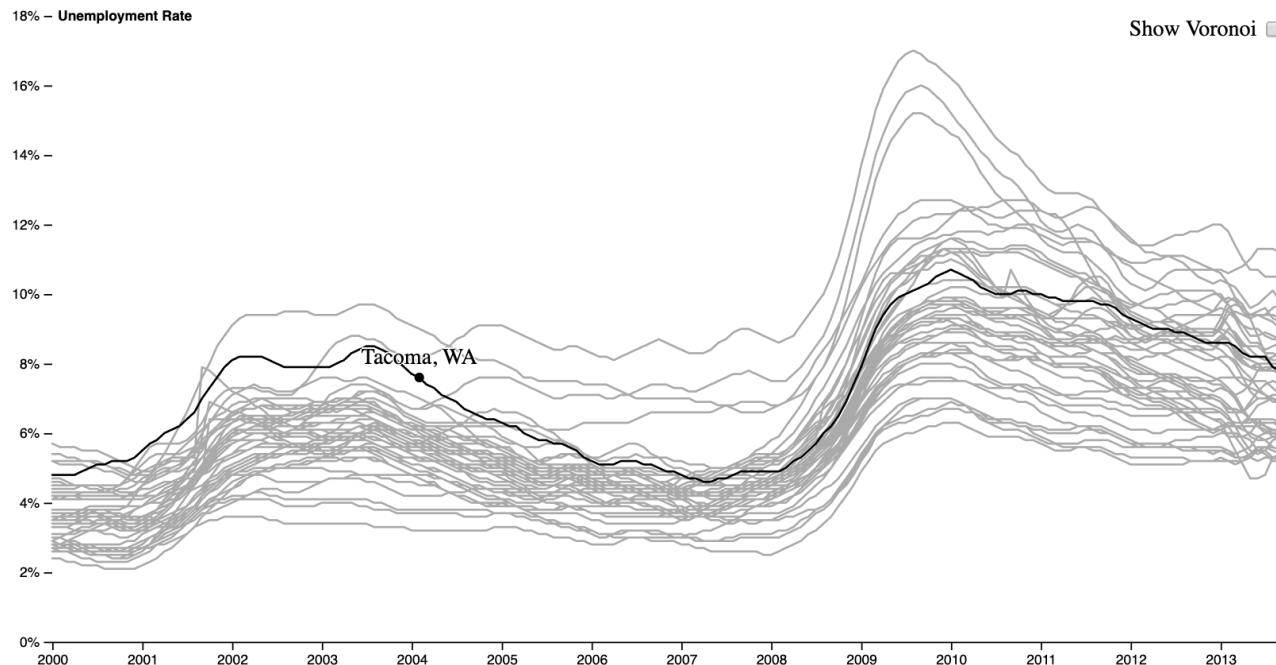
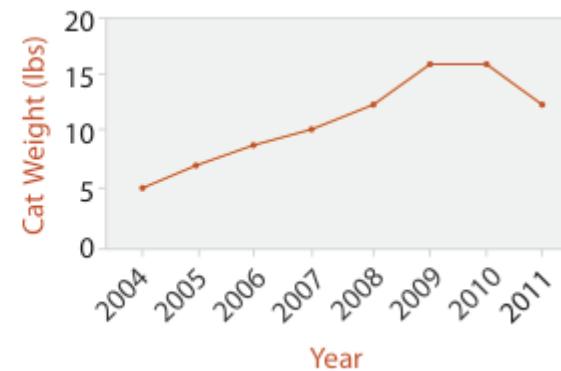
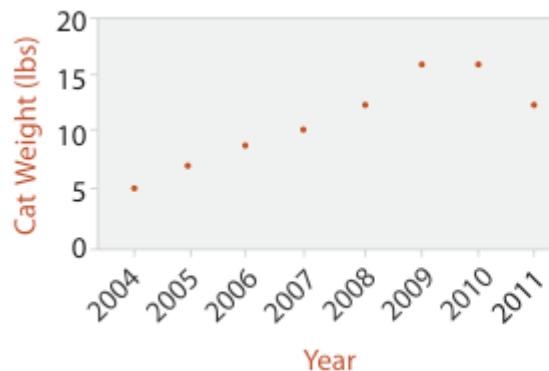
Same as before.



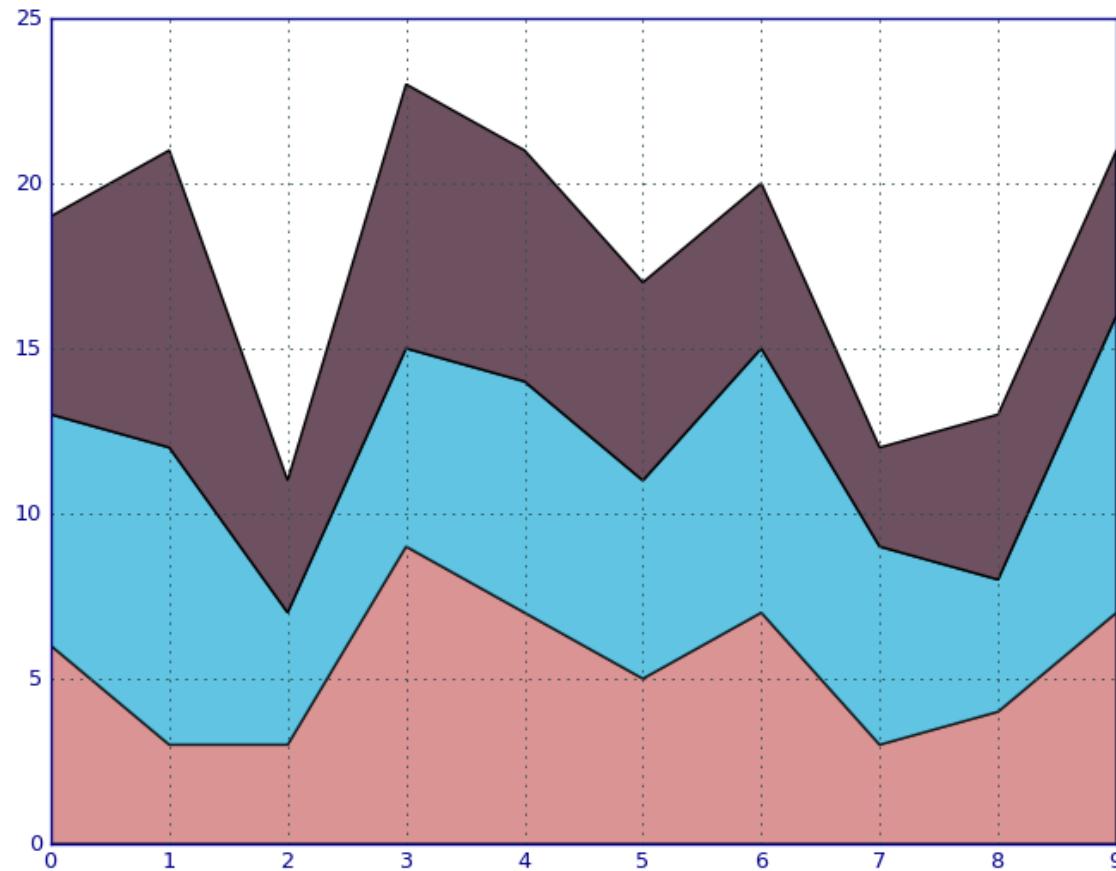
CHANGE OVERTIME

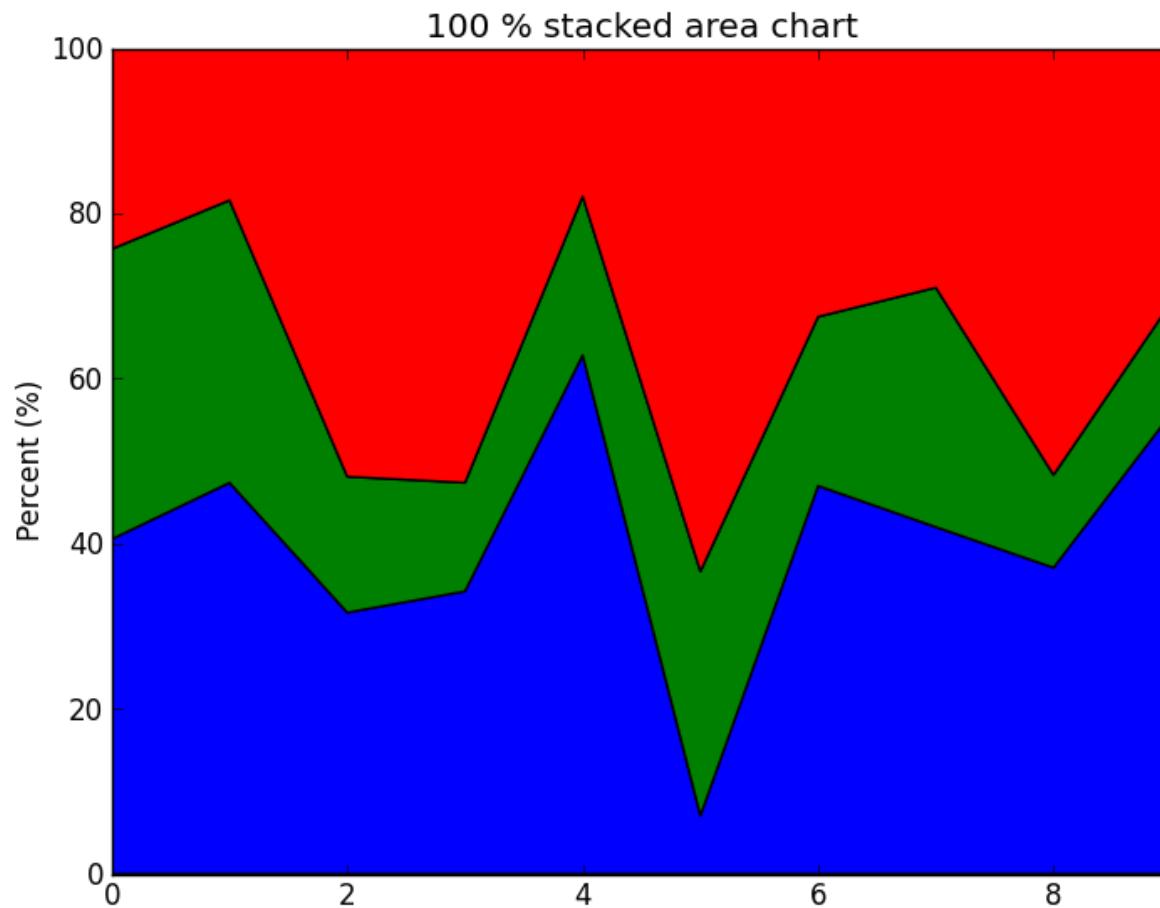
LINE CHART

Simple
Familiar
Accurate
Fairly Scalable

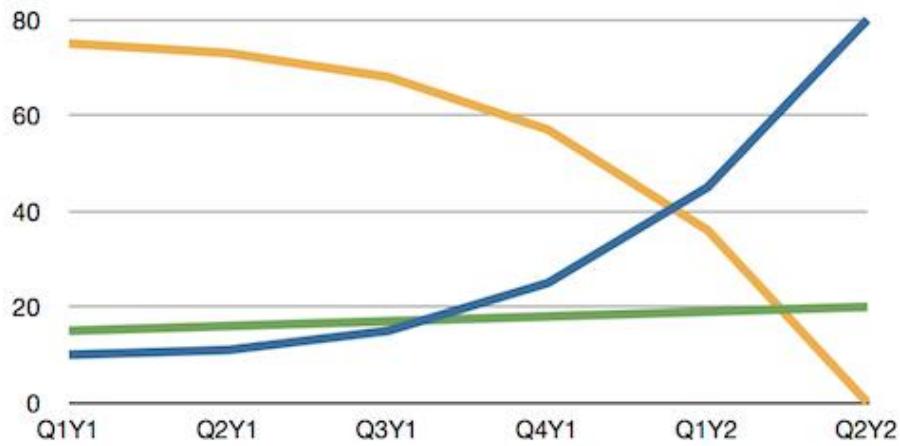
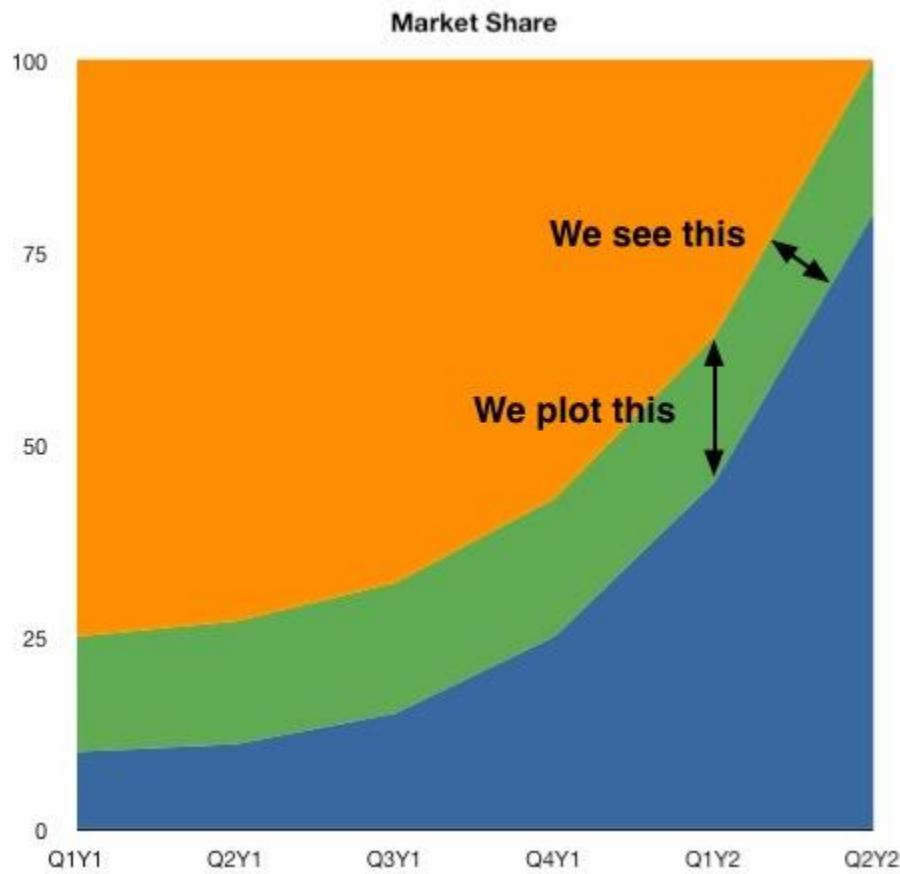


STACKED AREA

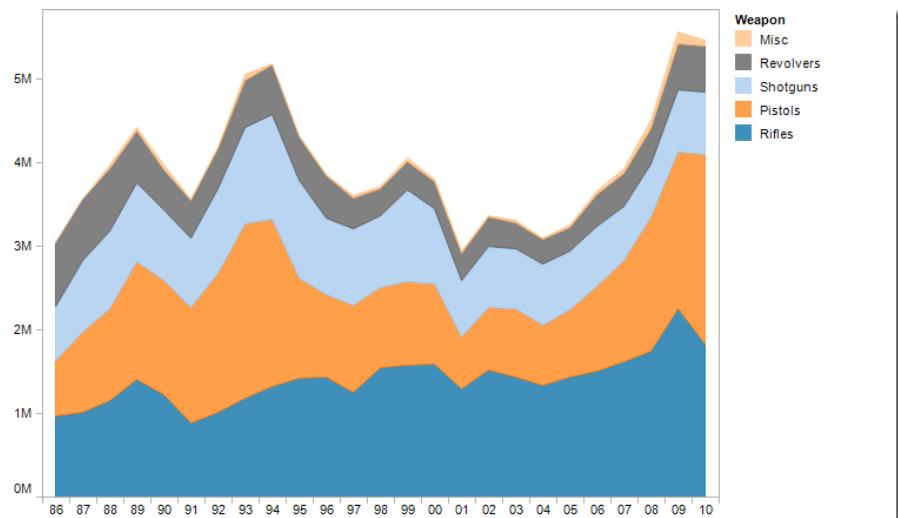
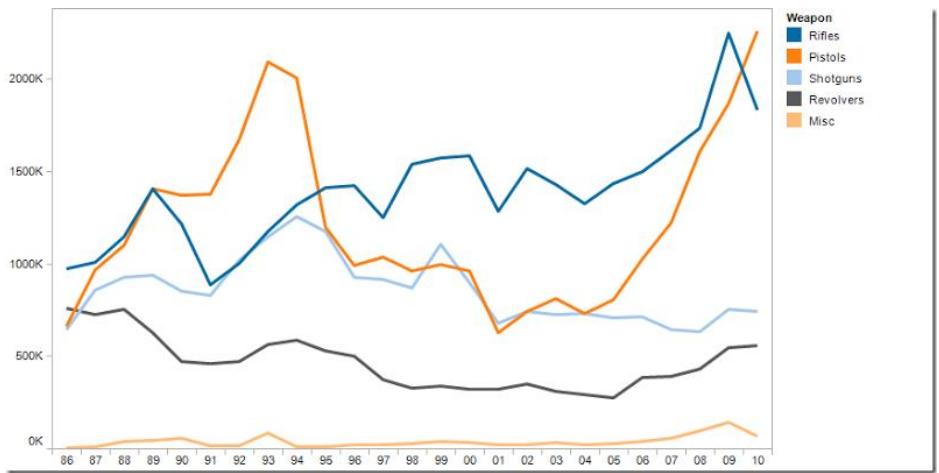
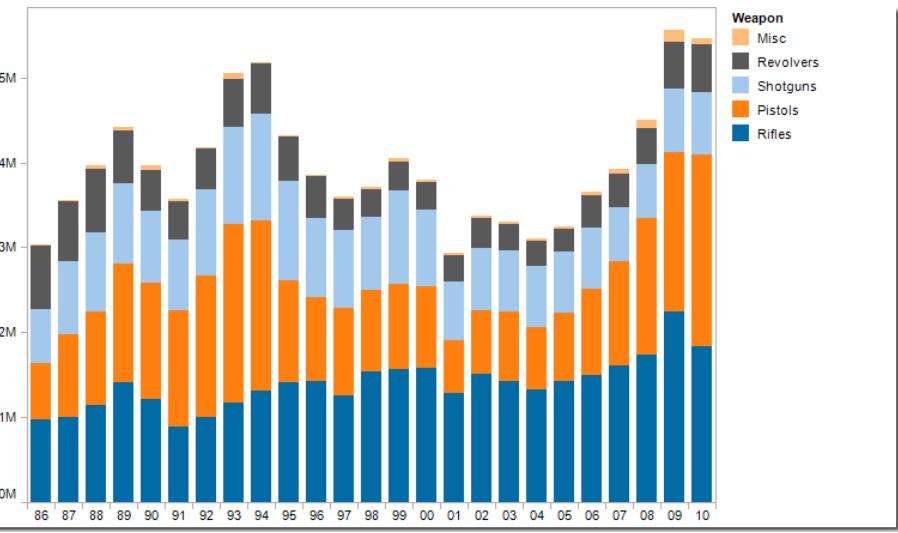




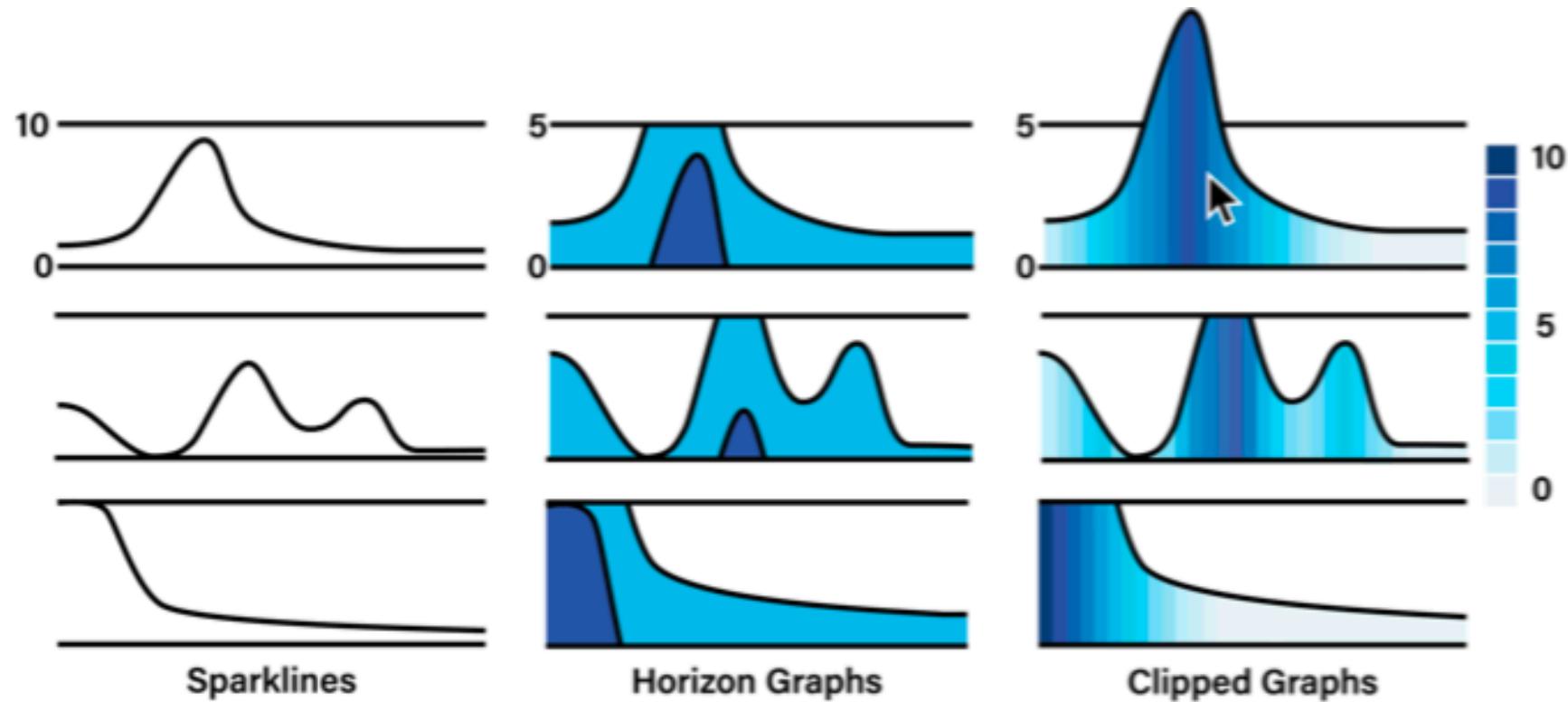
CAN YOU SPOT THE TREND?



CAN YOU SPOT THE TREND?

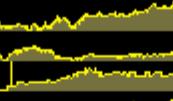
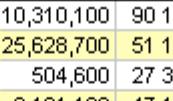
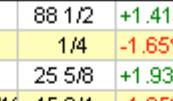
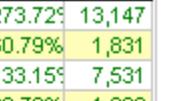
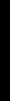


MULTIPLE LINE CHARTS



SPARKLINES

Small line charts. can be embedded in text or part of a table

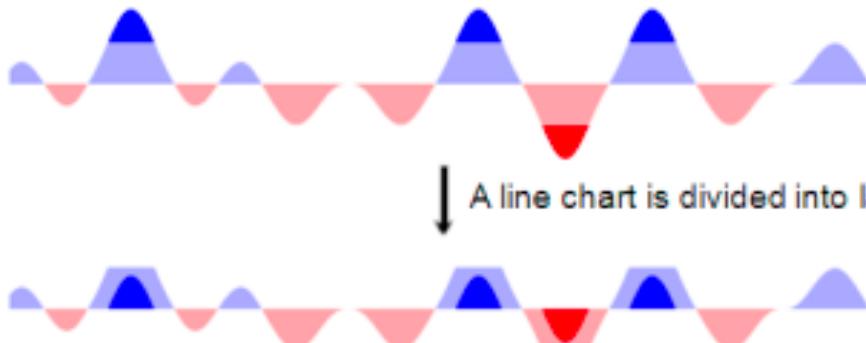
Symbol	Bid	Ask	Last	Change	T	Chart	Volume	High	Low	Value Change	Value	Gain	
DELL	89 3/4	89 13/16	89 3/4	+ 1 1/4	↑		10,310,100	90 1/8	88 1/2	+1.41%	250	17,950	+273.72% 13,147
CPQ	48 7/16	48 9/16	48 7/16	- 13/16	↓		25,628,700	51 1/4	1/4	-1.65%	-81	4,844	+60.79% 1,831
SDTI	26 1/4	26 3/8	26 3/8	+ 1/2	↓		504,600	27 3/8	25 5/8	+1.93%	250	13,188	+133.15% 7,531
COMS	46 1/2	46 9/16	46 9/16	- 25/32	↓		3,191,100	47 15/16	45 3/4	-1.65%	-102	6,053	+29.79% 1,389
LU	111 5/8	111 11/16	111 9/16	+ 1 9/16	↑		5,104,600	112 5/8	110	+1.42%	78	5,578	+22.76% 1,034
YHOO	368 1/16	368 1/2	368 1/2	+ 17 1/4	↓		3,787,800	381 3/16	280	+4.91%	431	9,213	-0.41% -38
AOL	162 13/16	163	163	+ 8	↑		10,008,500	164	158 1/2	+5.16%	280	5,705	+73.06% 2,408
CMGI	97 3/8	97 1/2	97 1/2	+ 5 7/8	↓		1,323,800	98 1/2	93	+6.41%	705	11,700	+186.76% 7,620
SPLN	33 13/16	33 15/16	33 13/16	+ 7/16	↓		300,200	34 3/4	33 5/8	+1.31%	88	6,763	+94.60% 3,288
BEAS	13 1/2	13 5/8	13 5/8	- 7/16	↓		389,200	14 1/4	13 1/8	-3.11%	-44	1,363	-9.17% -138
GNET	102	103 3/16	101 5/16	+ 6 1/8	↑		307,600	108	97	+6.43%	613	10,131	+130.26% 5,731
RNMK	67	67 1/4	67	+ 2 3/4	↓		1,233,900	69	64 15/16	+4.28%	275	6,700	+79.87% 2,975
MSFT	173 1/8	173 1/4	173 5/16	+ 1 3/4	↓		13,264,500	174 7/16	170	+1.02%	175	17,331	+54.74% 6,131
INTC	133 3/4	133 13/16	133 13/16	- 3 1/8	↓		8,094,300	137 1/2	133 3/8	-2.28%	-625	26,763	+65.20% 10,563
TOTAL					↑		205,302	80,993	+1.63%	2,293	143,280	+79.41% 63,377	

Mauricio Pochettino has lead Spurs on their best
run 8TH  2ND in 24
years of the Premier League

Alibaba stock is at 5 yr
high **93.89**  **152.11** as of July 2017

The FTSE100 Brexit
bounce **5562**  **7501** continues one year on
from the vote last summer

HORIZONTAL GRAPH



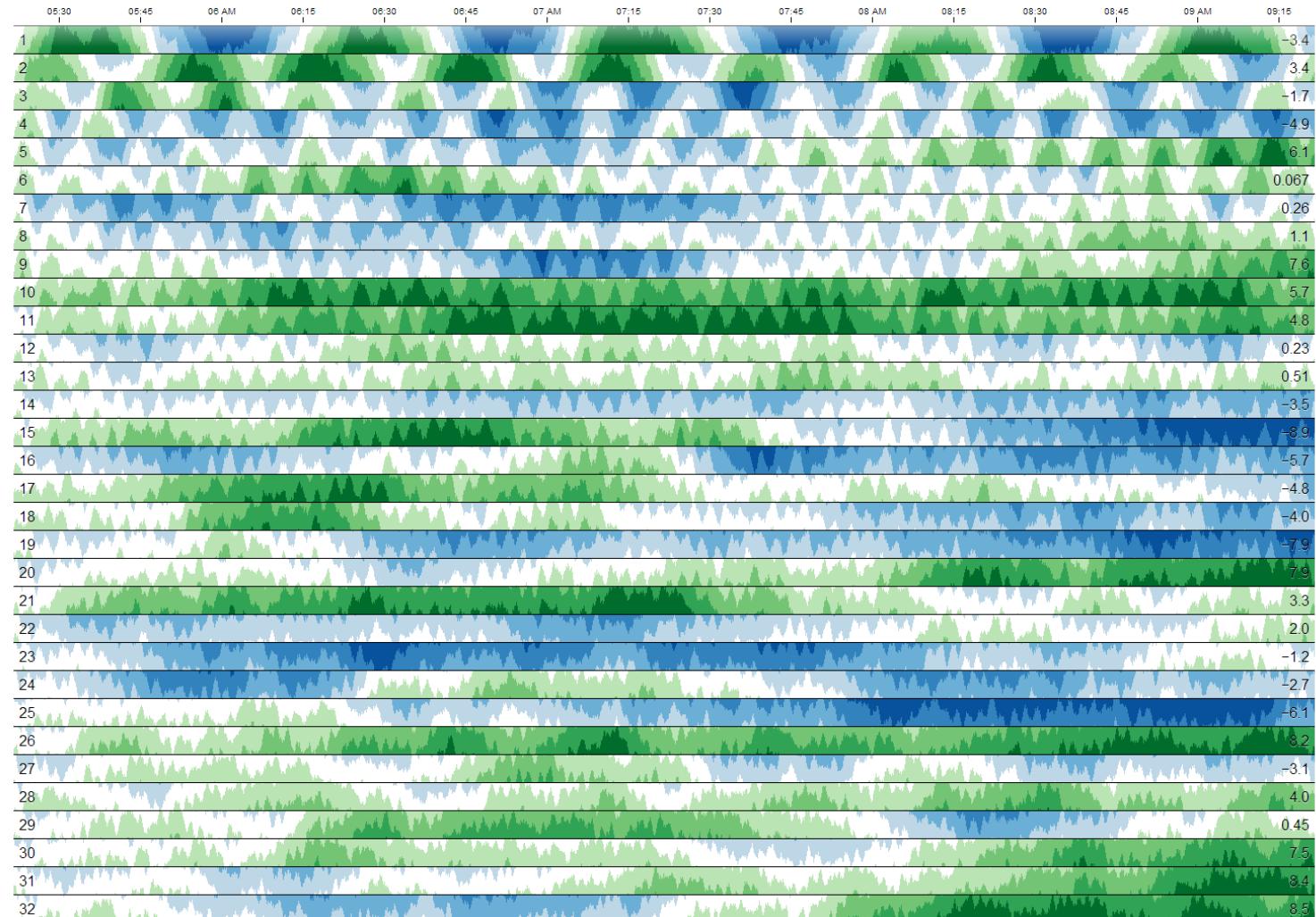
A line chart is divided into layered bands.

Negative values are mirrored.

Negative values are offset.

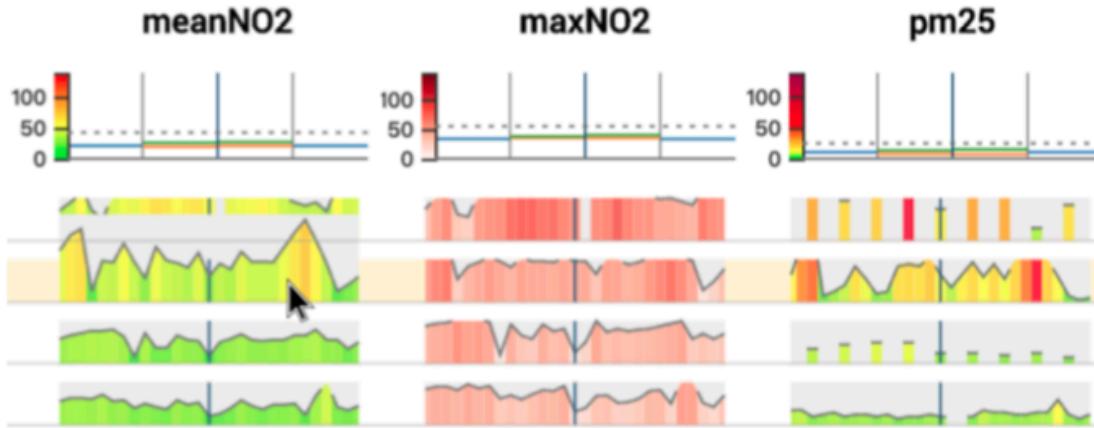


HORIZONTAL GRAPH

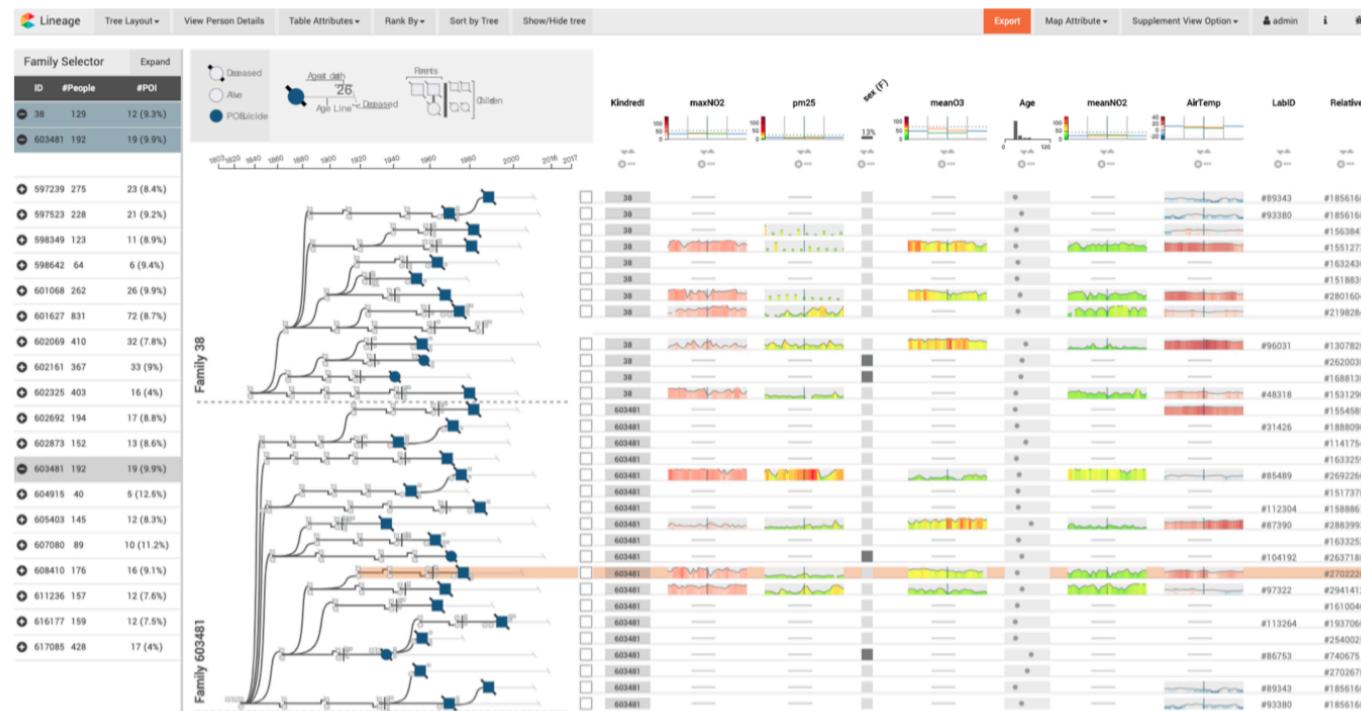


<http://square.github.io/cubism/>

CLIPPED GRAPHS



<https://youtu.be/IxoaFe6yBMg>



CONNECTED SCATTERPLOT

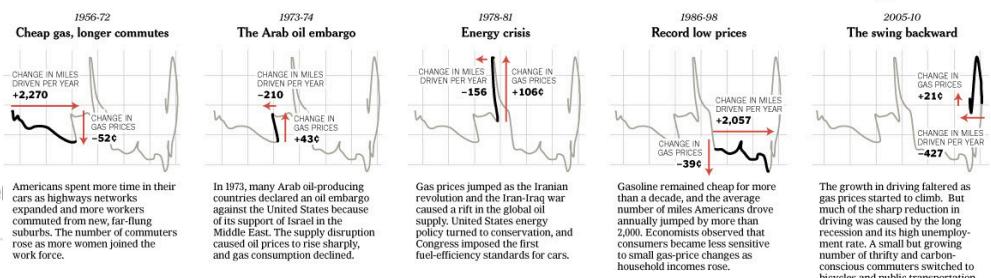
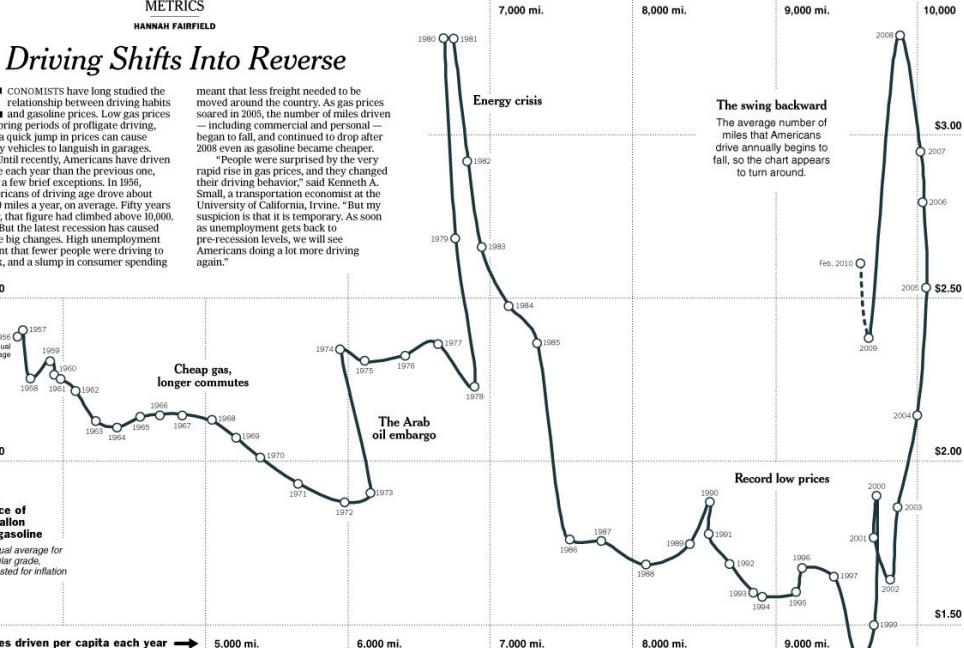
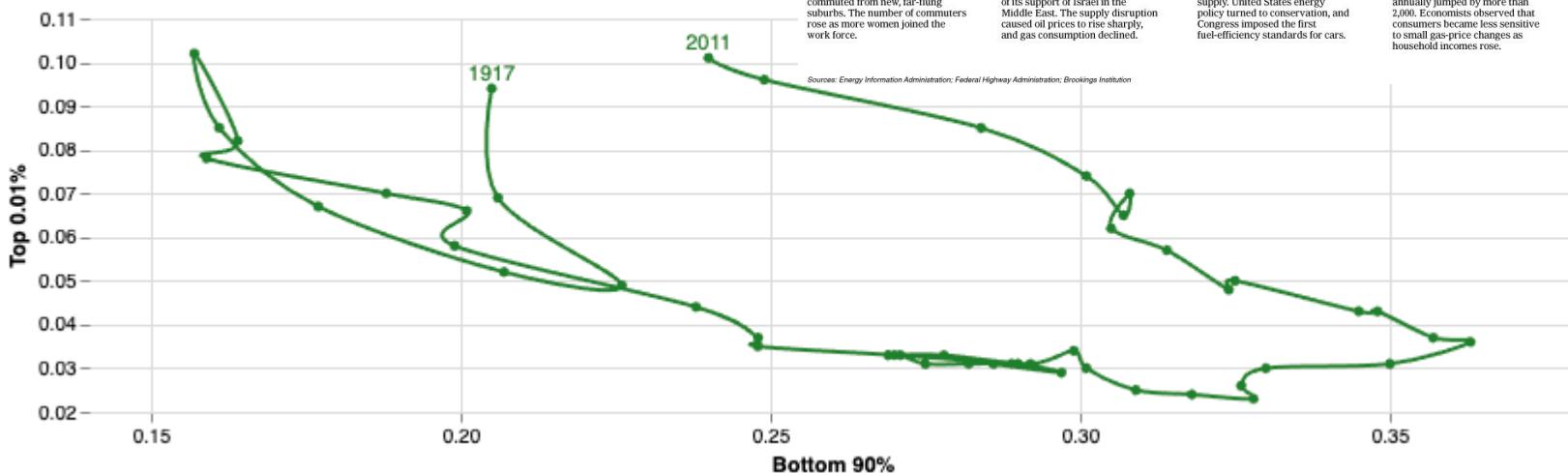
Two Variables + Time
Only one per Chart!
Labels important

http://steveharoz.com/research/connected_scatterplot/

<http://www.thefunctionalart.com/2012/09/in-praise-of-connected-scatter-plots.html>

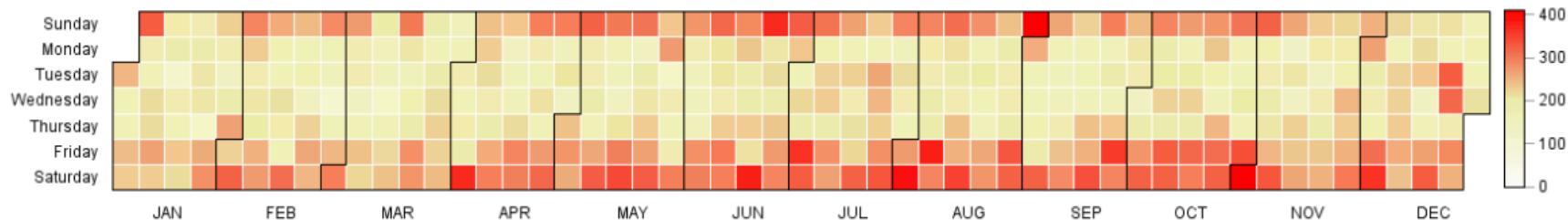
Connected scatterplot

A good way of showing changing data for two variables whenever the



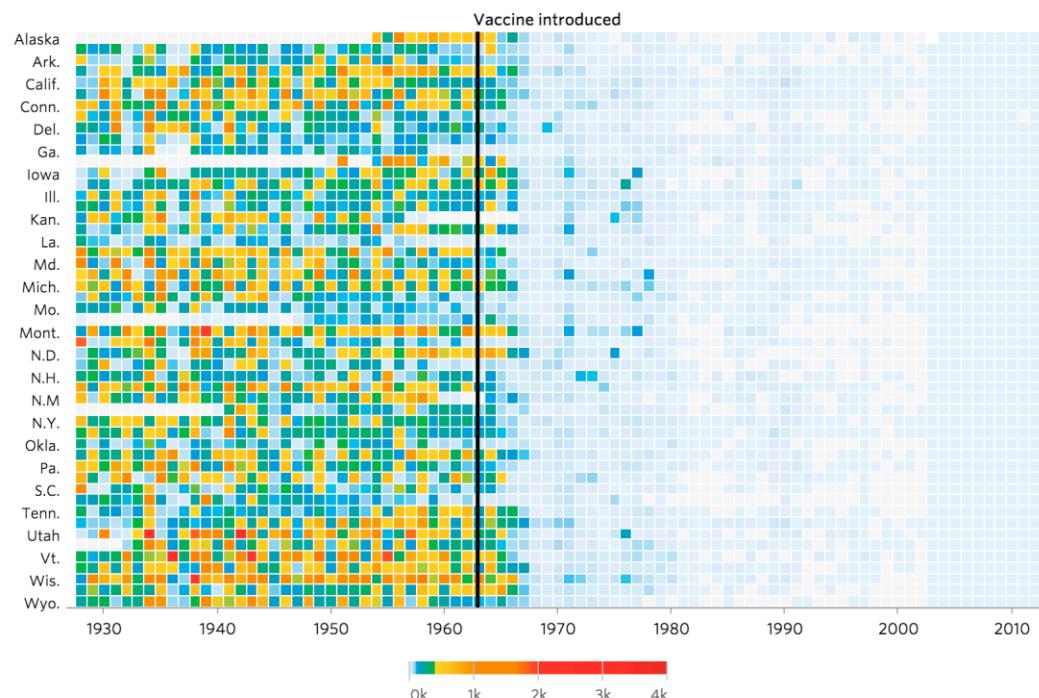
HEATMAP

Traffic Fatalities in the US during 2008



The heat maps below show number of cases per 100,000 people.

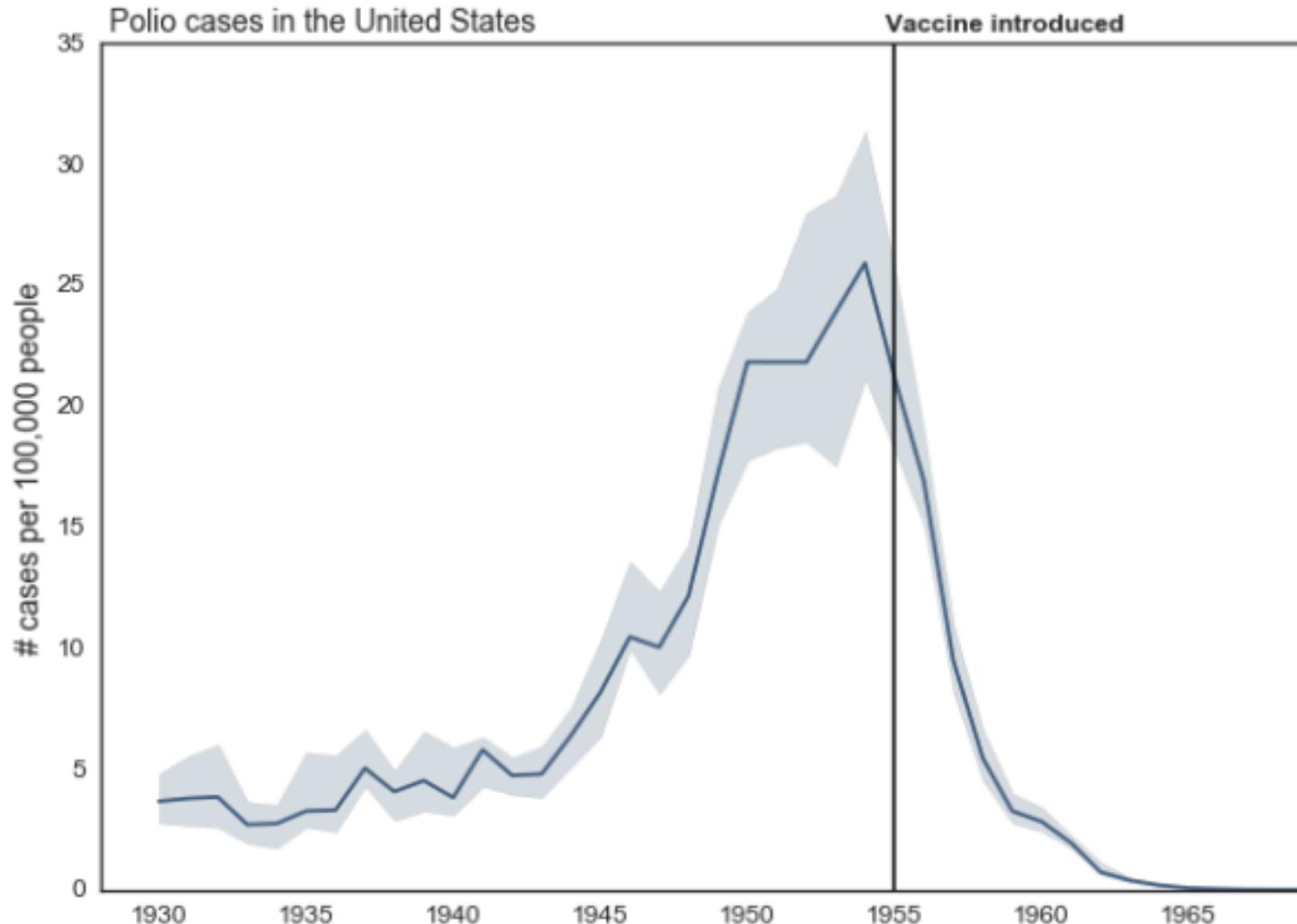
Measles



<https://www.informationisbeautifulawards.com/howcase/660-vaccines-and-infectious-diseases>

Note: CDC data from 2003-2012 comes from its Summary of Notifiable Diseases, which publishes yearly rather than weekly and counts confirmed cases as opposed to provisional ones.

SOMETIMES YOU CAN SHOW TOO MUCH DATA



Data source: Project TYCHO (tycho.pitt.edu) | Author: Randy Olson (randalolson.com / [@randal_olson](https://twitter.com/randal_olson))

<http://www.randalolson.com/2016/03/04/revisiting-the-vaccine-visualizations/>

RANKING

RANKING EXERCISE

	1	2	3	4	5	6	7
Bavaria	8	6	2	4	2	1	3
Dortmund	1	1	5	2	3	8	8
Leipzig	2	2	1	1	1	2	4
Leverkusen	5	5	4	8	7	6	7
Moenchengladbach	10	7	8	7	6	5	1
Wolfsburg	6	4	3	5	8	7	2

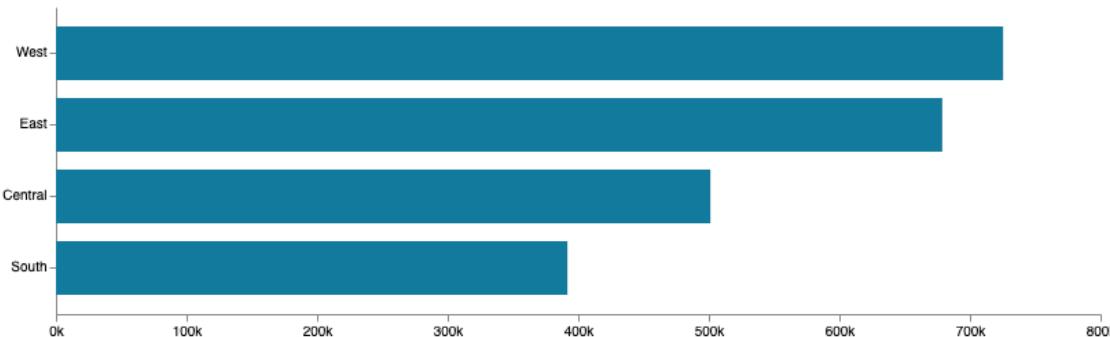
Design a visualization showing the ranking of these football clubs over time.

RANKING

Magnitude Visualization + Sorting

Ordered bar

Standard bar charts display the ranks of values much more easily when sorted into order

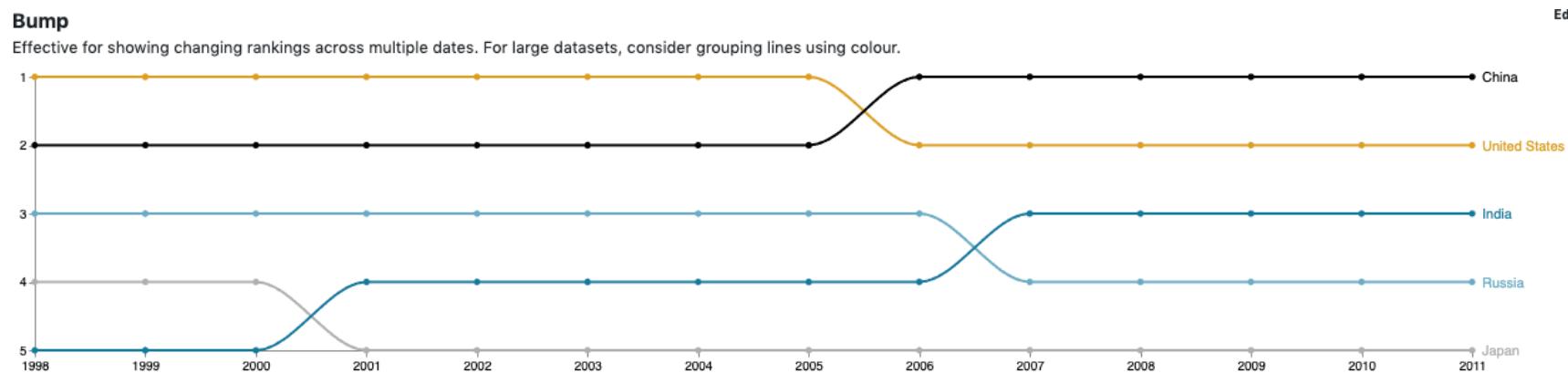


Edi

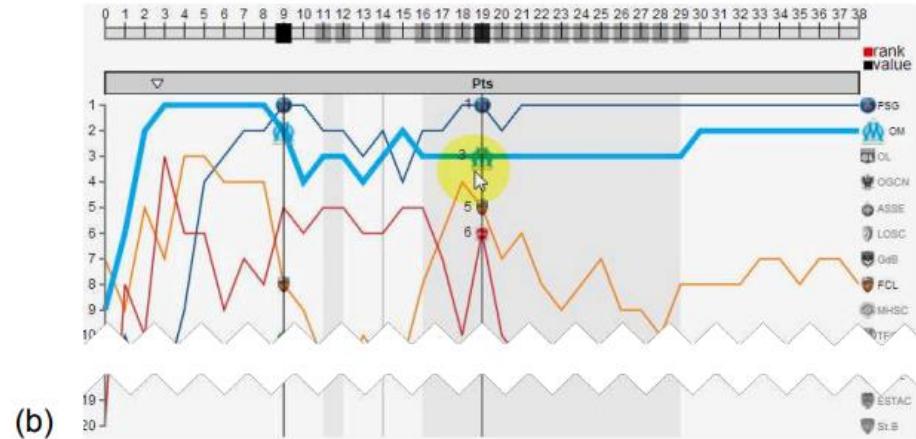
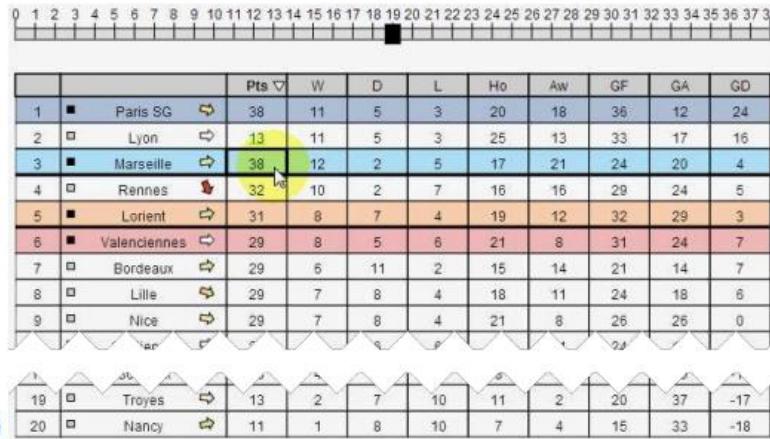
Bump Charts for Rankings over Time

Bump

Effective for showing changing rankings across multiple dates. For large datasets, consider grouping lines using colour.



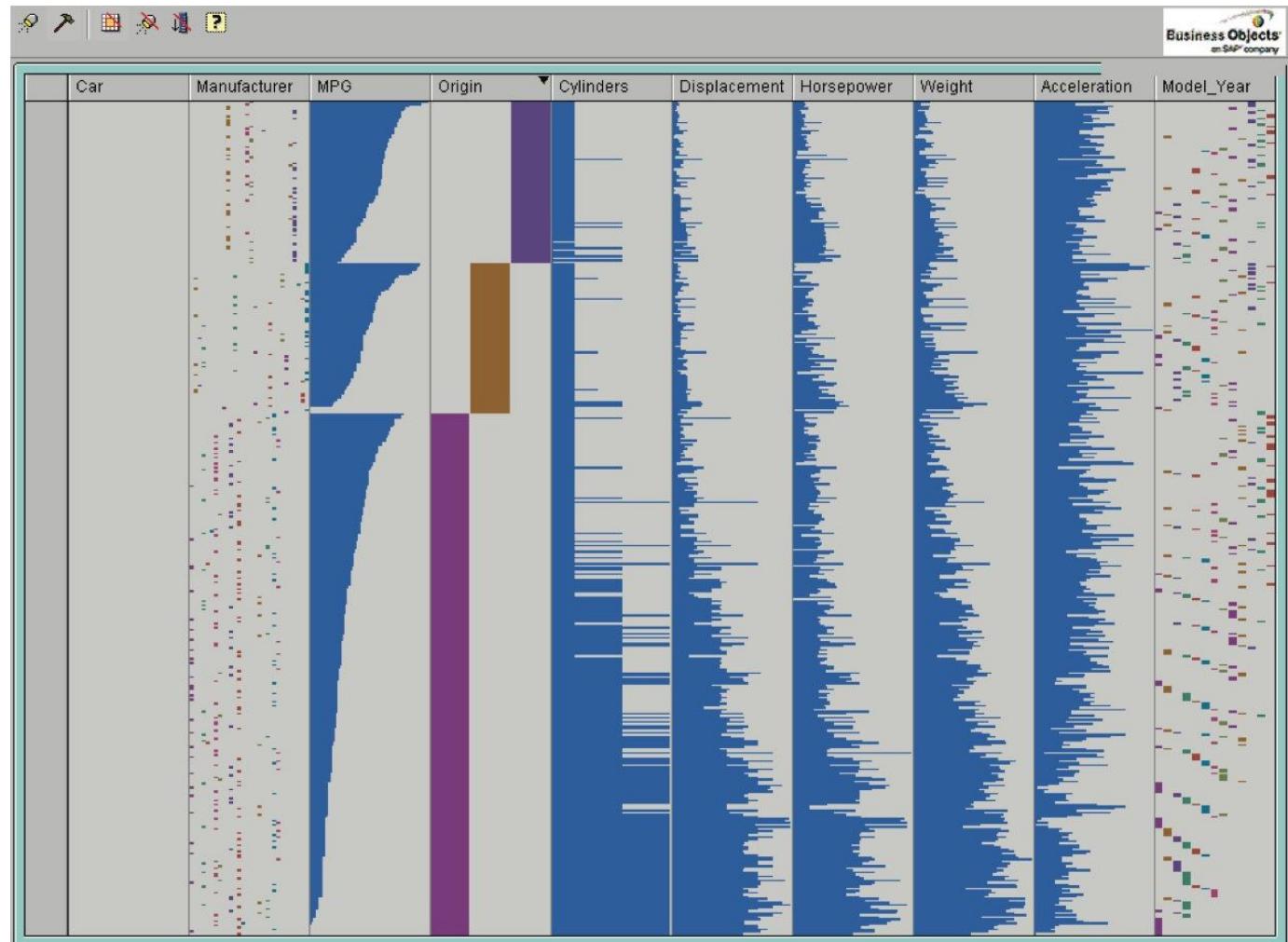
TEMPORAL RANKINGS



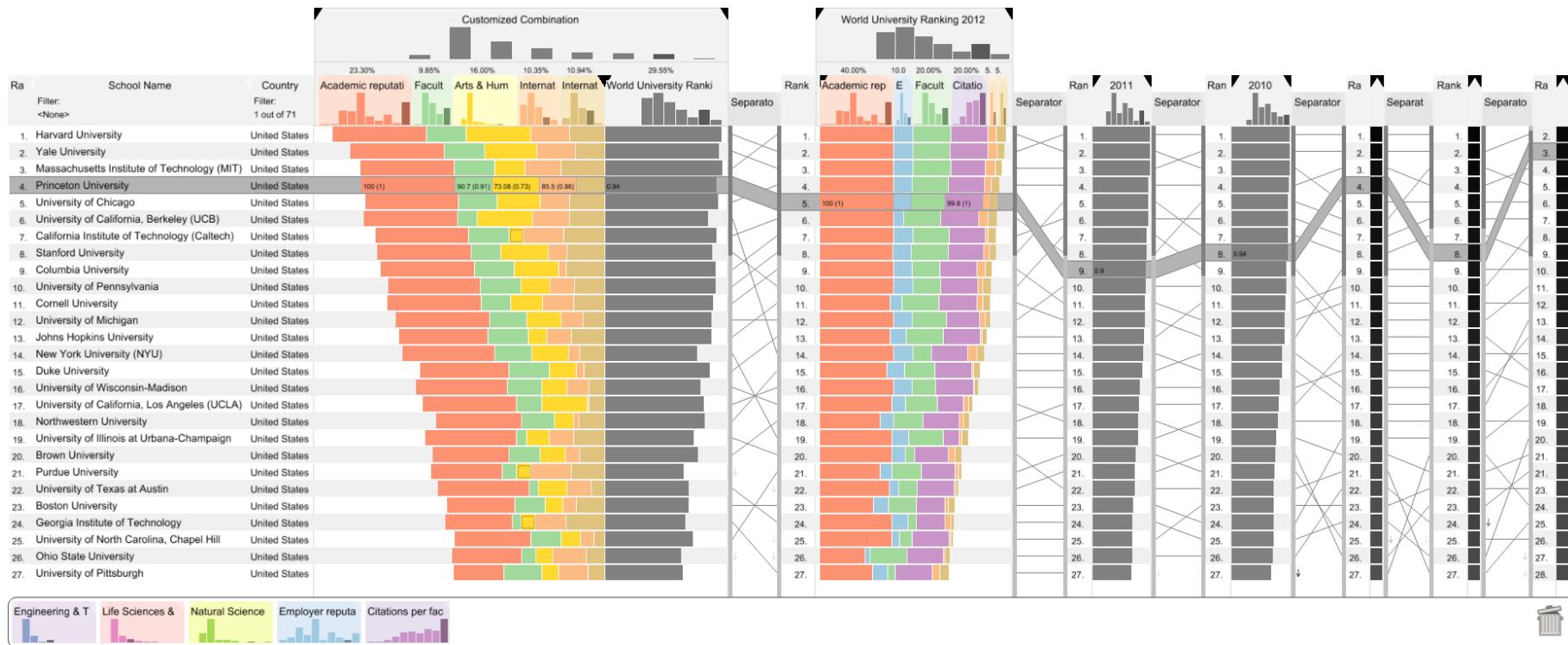
<https://youtu.be/a0duFWu3Zp0>

TABLE LENS

Interactive table based representation

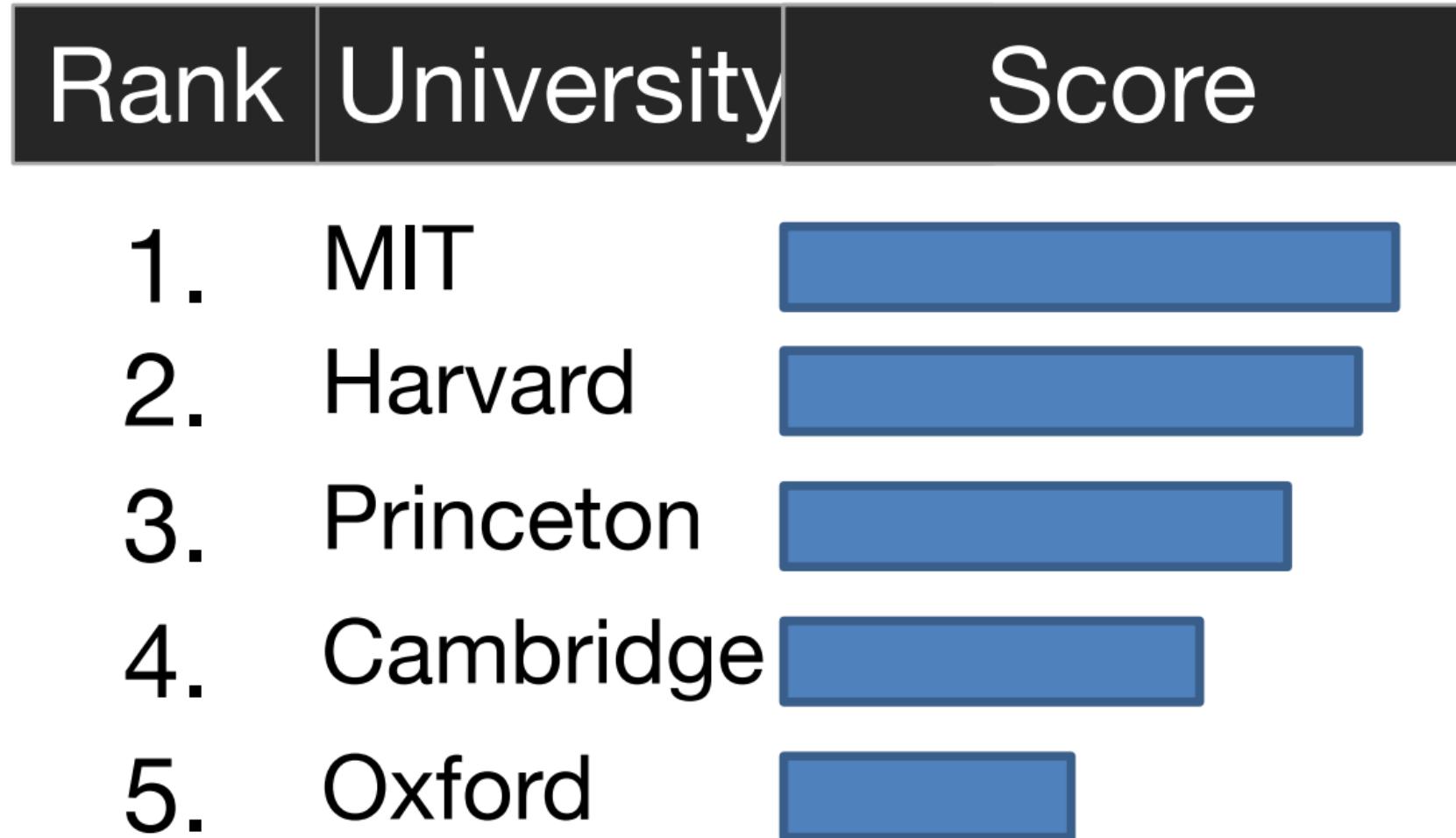


LINEUP



<http://lineup.caleydo.org>

RANKING



MULTIPLE ATTRIBUTES

COMBINER FUNCTION

(Weighted) sum

Score = $w_aA + w_bB + w_cC$ → Serial

Maximum

Score = $\max(A, B, C)$ → Parallel

Product
Nesting

→ Complex

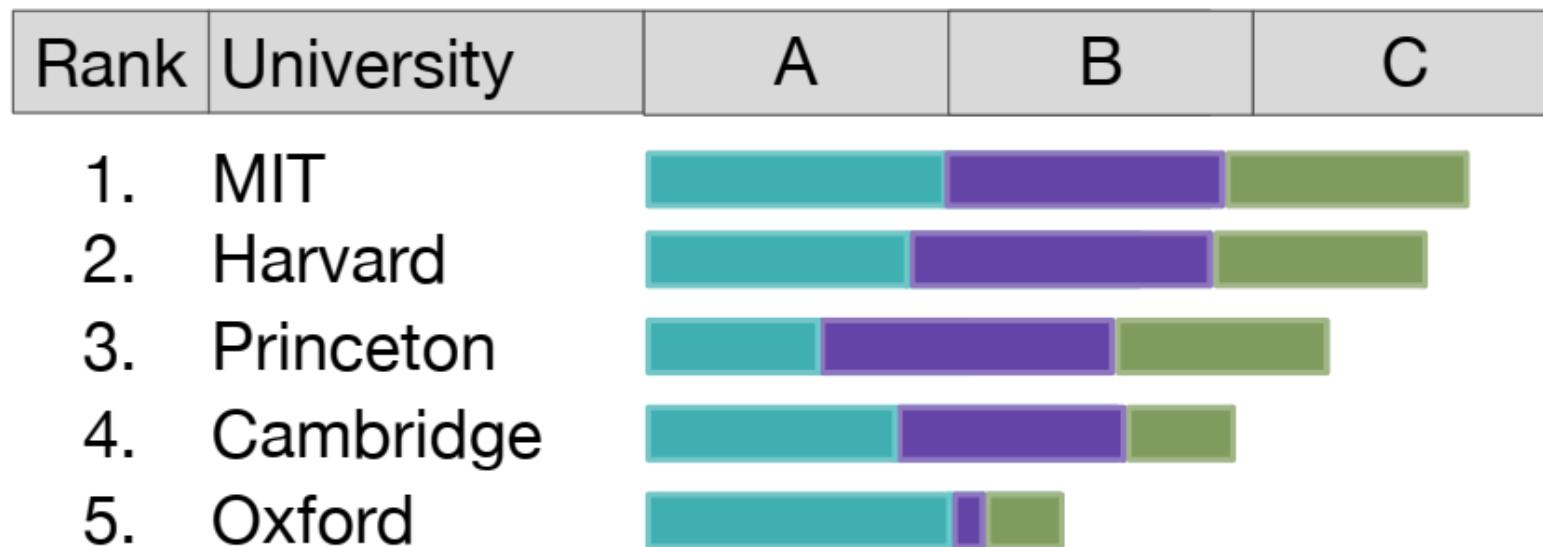
Serial Combiner(as Stacked Bar)



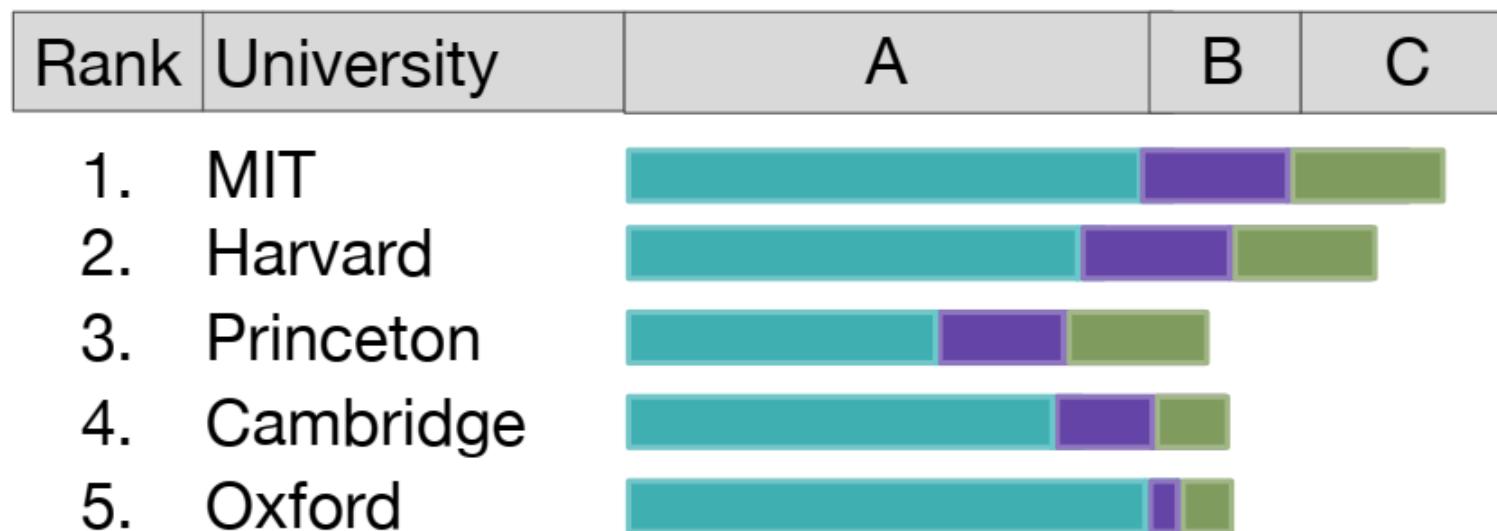
Rank	University	A	B	C
1.	MIT			
2.	Harvard			
3.	Princeton			
4.	Cambridge			
5.	Oxford			

Serial Combiner(as Stacked Bar)

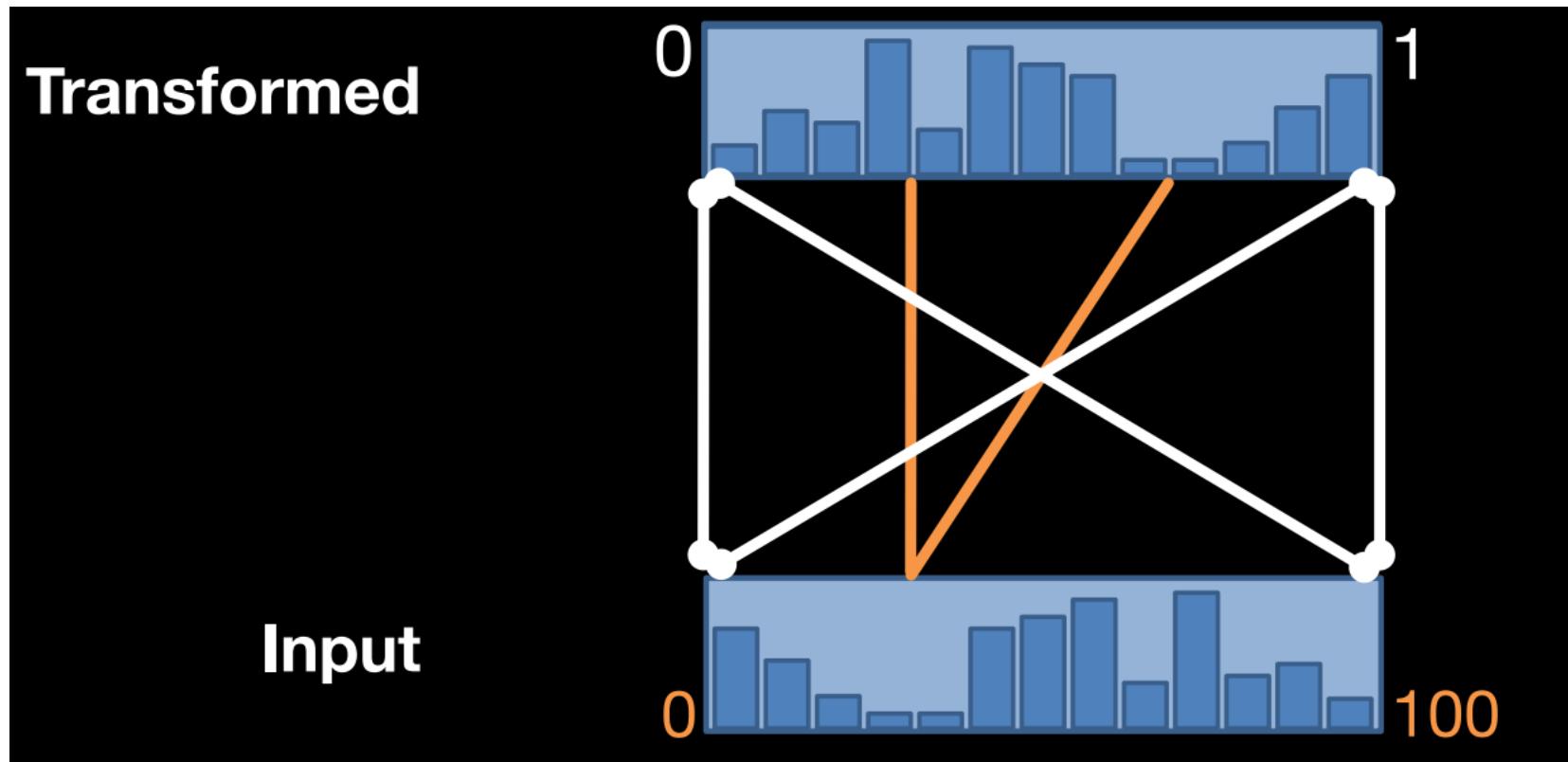
$$w_a A + w_b B + w_c C$$



Serial Combiner(as Stacked Bar)

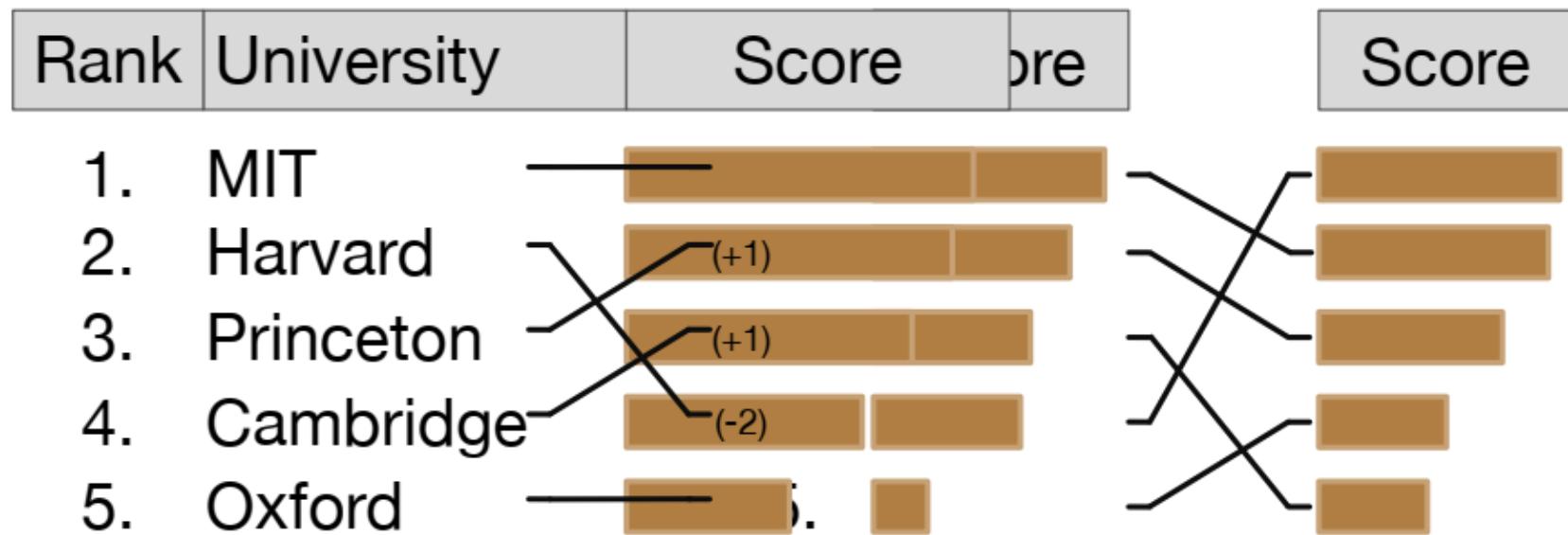


MAPPING ATTRIBUTES TO SCORES



BUMP CHARTS

Bump Charts

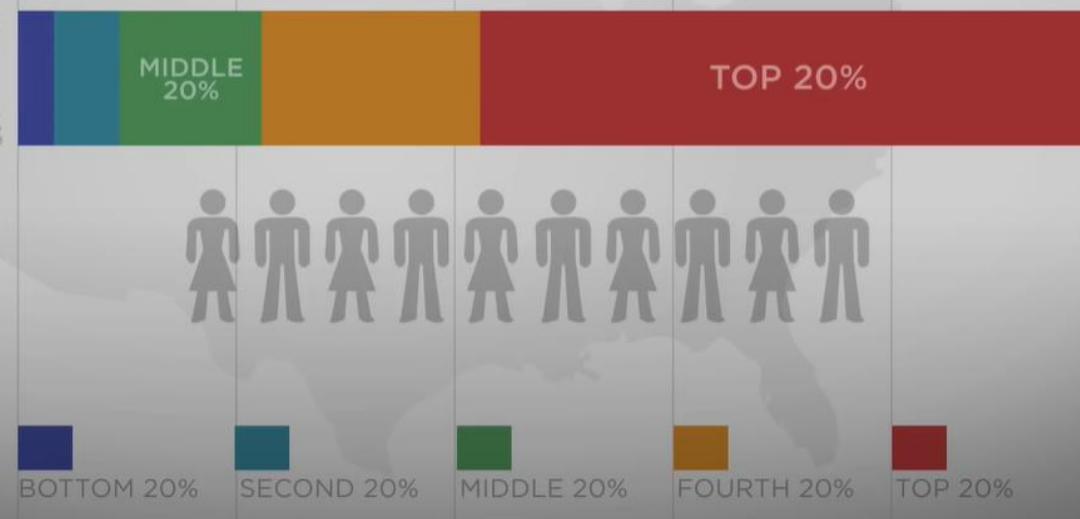


STORYTELLING

WEALTH INEQUALITY IN AMERICA

What Americans
THINK
The Distribution Is

Distribution
92% Choose as
IDEAL



<https://www.youtube.com/watch?v=QPKKQnijnsM>

STORYTELLING

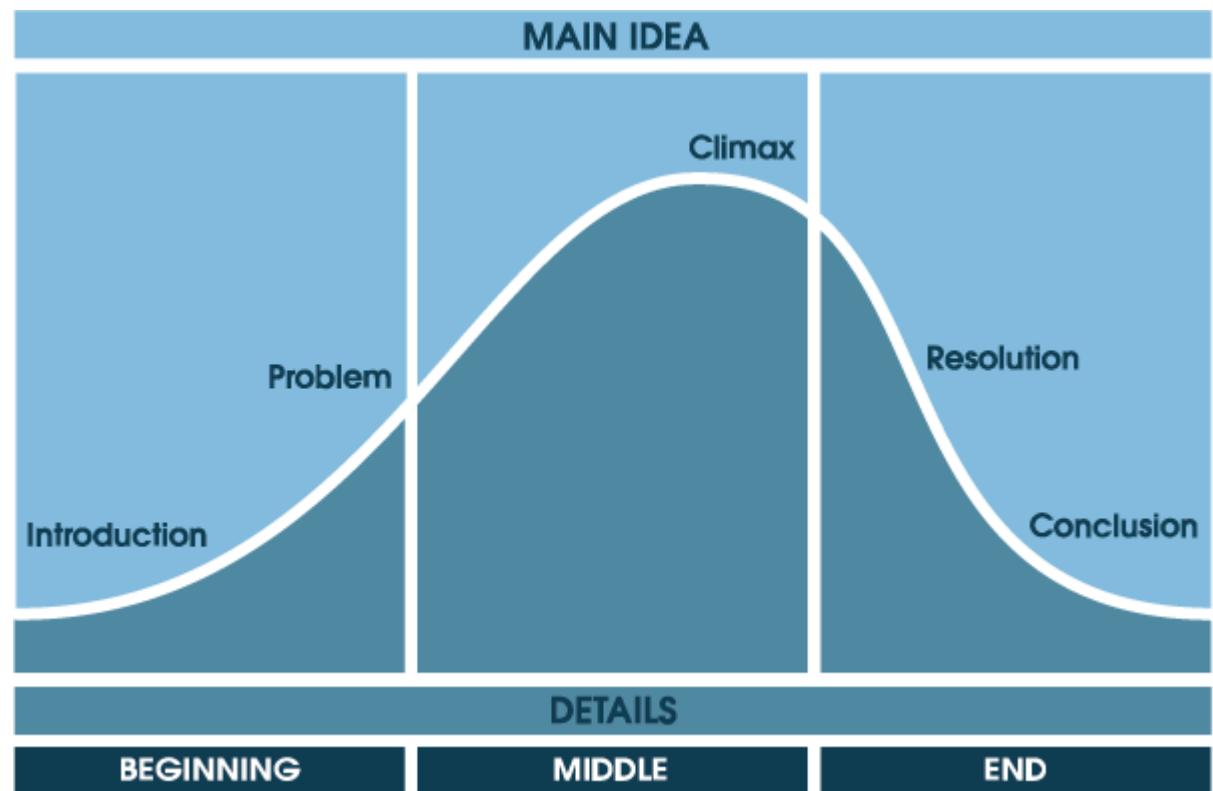
“The world’s second-oldest profession”

Good stories do more than provide facts and data

- They situate and give context
- They engage
- They educate

Who / What / Where

Why / How

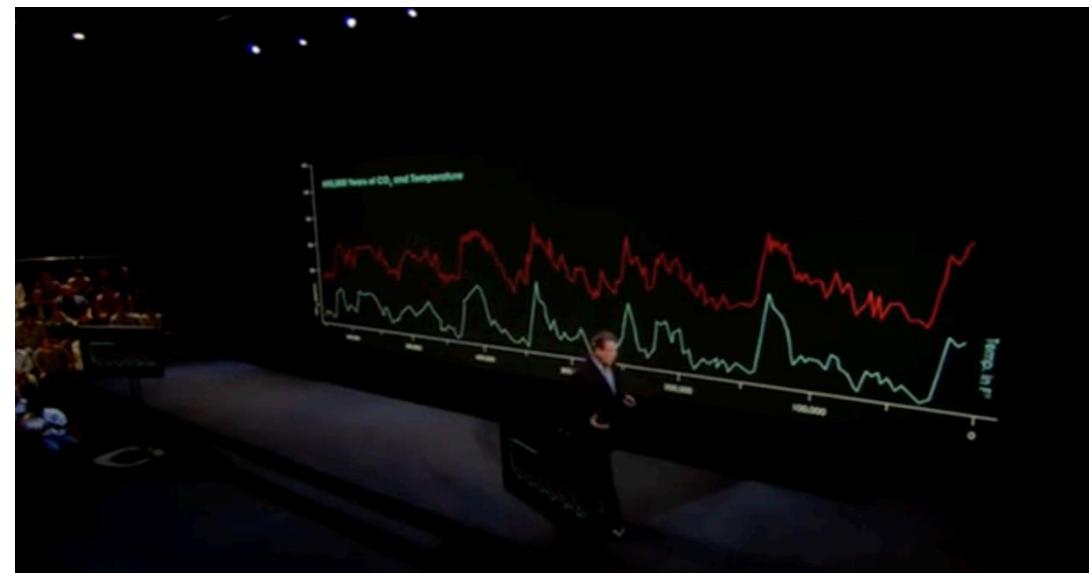


STORYTELLING WITH DATA/VIS

Underscore your arguments
with Data/Facts

Leverage the power of
Visualization

- Show trends
- Show correlations
- Show outliers
- Convey magnitudes



<https://youtu.be/-JluKjaY3r4>

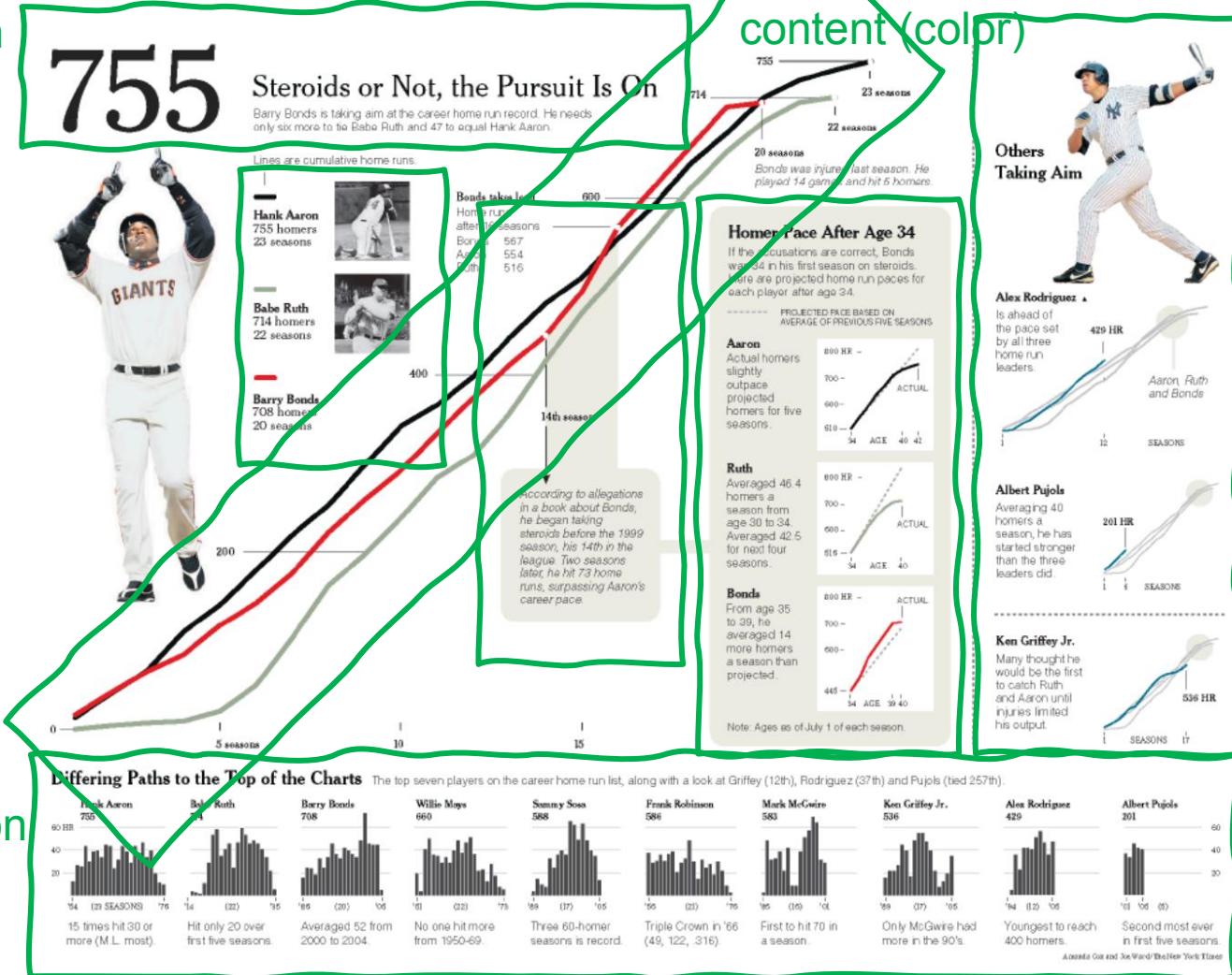
COMPONENTS OF A DATA STORY

Introduction

Context

Main story

Annotation
of key
points



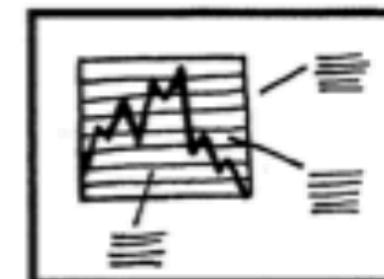
Another player
Context, matching on
content (color)

GENRE

Seven
Genres



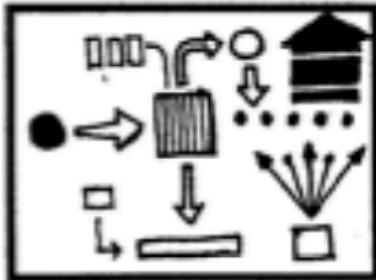
Magazine Style



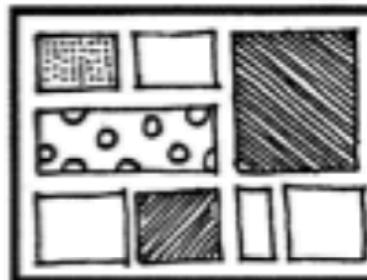
Annotated Chart



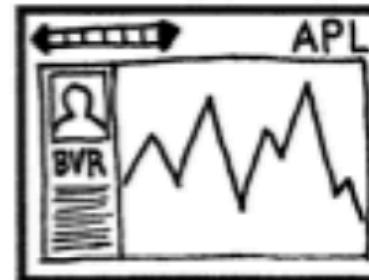
Partitioned Poster



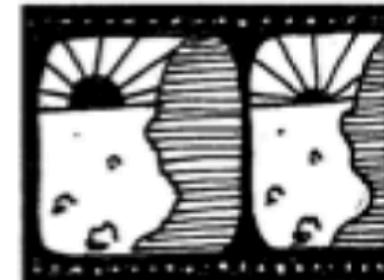
Flow Chart



Comic Strip



Slide Show



Film/Video/Animation

AUTHOR OR READER DRIVEN

Author driven

Reader driven



Linear ordering

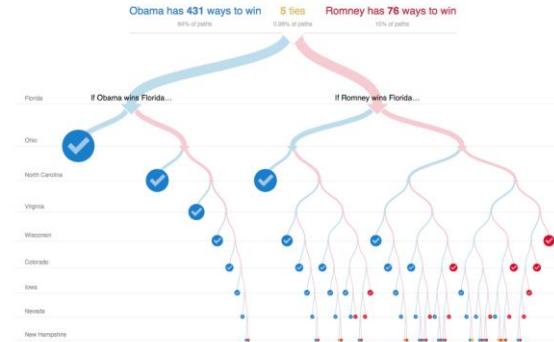
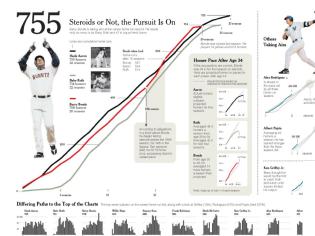
No ordering

Heavy messaging

No messaging

No interactivity

Free interactivity



AUTHOR OR READER DRIVEN

Martini Glass structure

start with author driven, open up for exploration



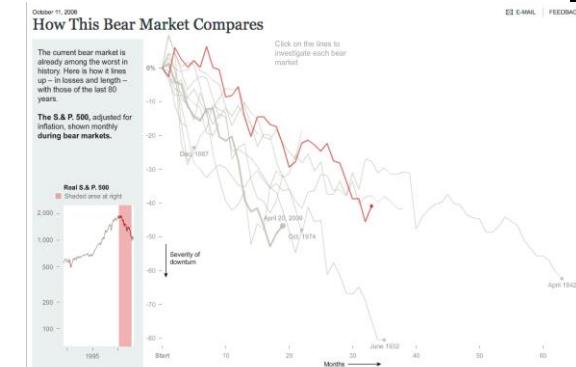
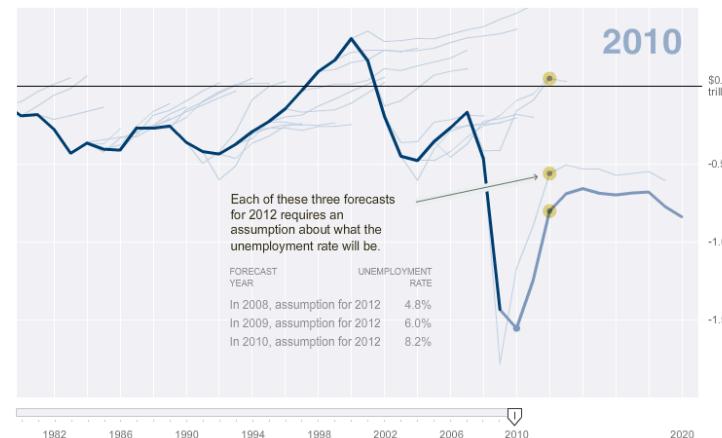
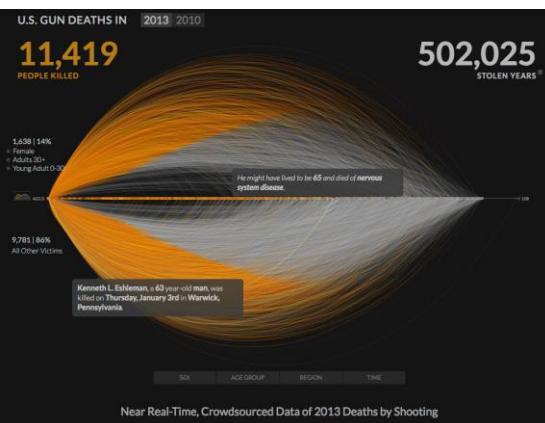
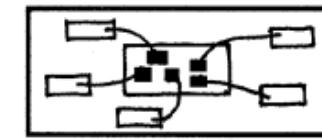
Interactive slideshow

Split into multiple scenes, allow interaction mid-way



Drill-down story

Let reader decide which path to follow, all paths are annotated



STRATEGY FOR STORYTELLING

LAYOUT PRINCIPLE

Descriptive titles

Descriptive subtitles

Annotations

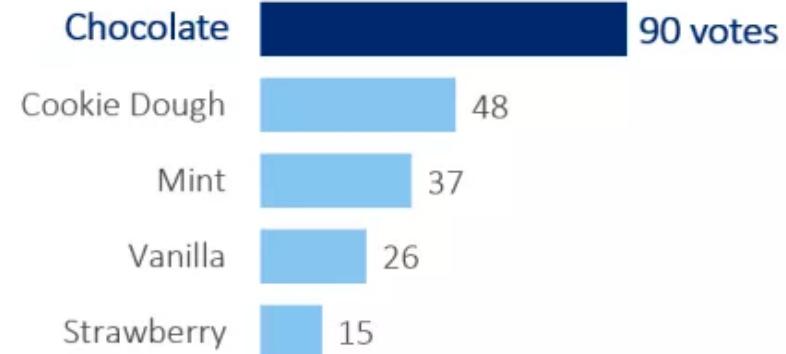
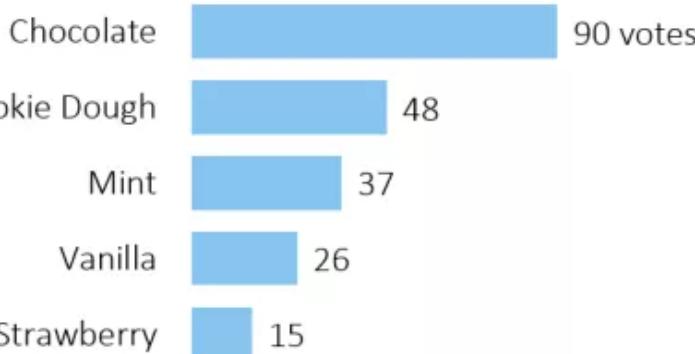
Saturation

DESCRIPTIVE TITLE AND SATURATION

Ice cream flavor preferences based on 2014 survey of elementary school students (n=216)

or

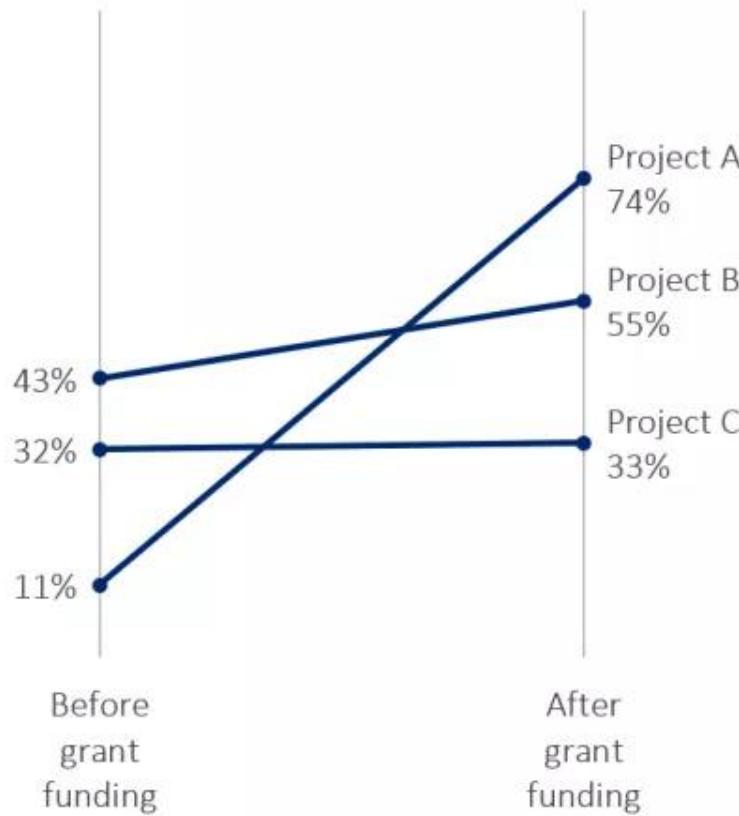
Chocolate was most popular flavor among elementary students surveyed



Source: 2014 survey of elementary school students (n=216)

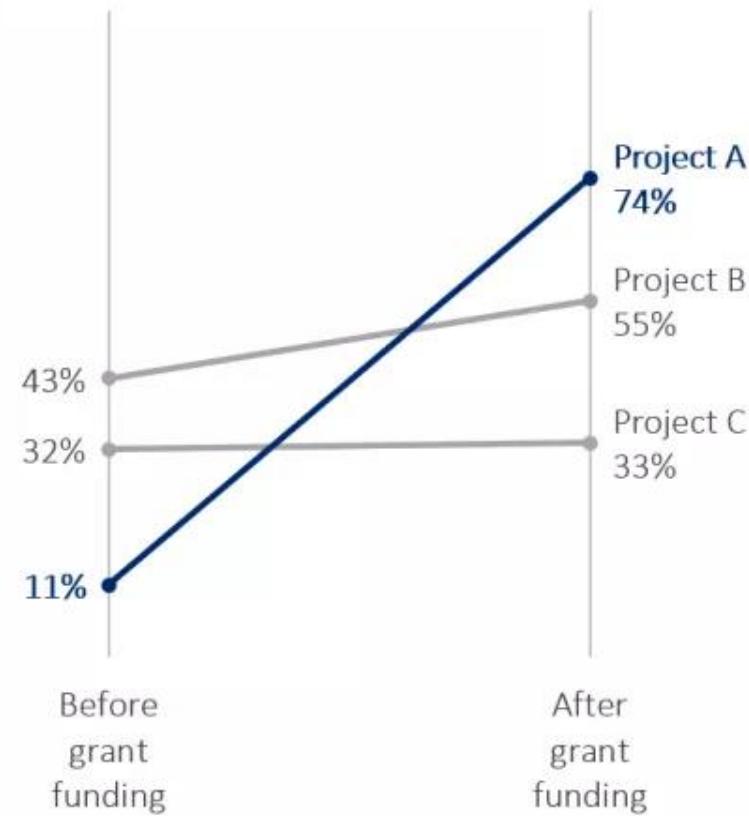
DESCRIPTIVE TITLE AND SATURATION EMPHASIZE HOW PROJECT A IS PERFORMING PARTICULARLY WELL

Project results before and after implementation of grant



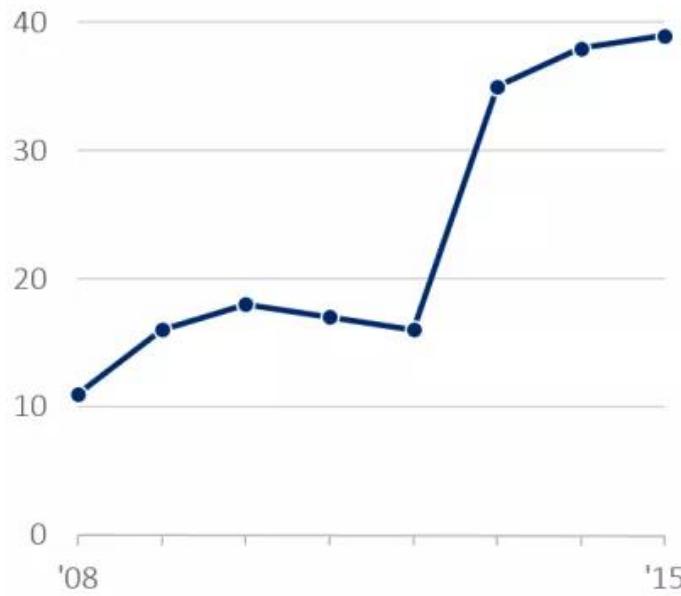
or

Project A had the greatest gains after the four-year grant funding



DESCRIPTIVE TITLE, DESCRIPTIVE SUBTITLE, AND ANNOTATION FOR INCREASED UNDERSTANDING

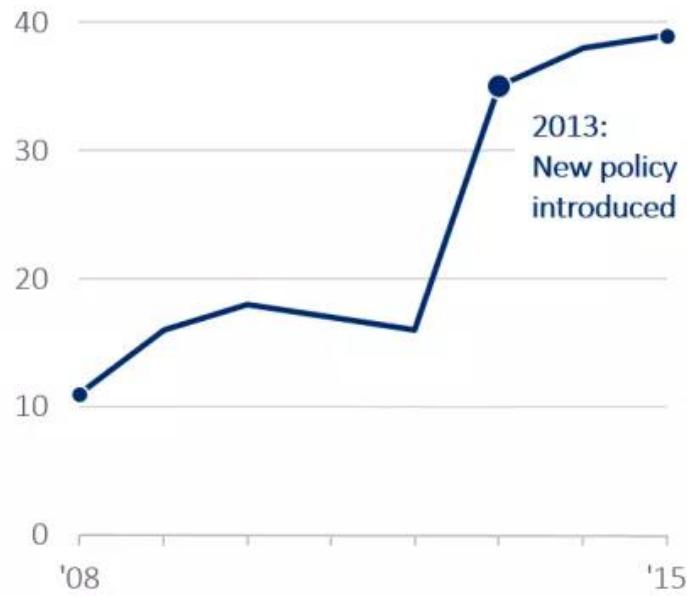
Number of studies funded each year



or

We're funding more studies each year

Beginning in 2013, we set aside new funding to measure the effectiveness of our initiatives – and we evaluated 39 of our programs in 2015 alone.



INTERACTIVITY

Navigation

Steppers, Scrolling,
Play/Fwd/Rwd, etc.



Explore the map to find out how life is across OECD regions and discover regions with similar well-being.

Each region is measured in eleven topics important for well-being. The values of the indicators are expressed as a score between 0 and 10. A high score indicates better performance relative to the other regions.

[Help](#)

Regions with similar well-being in other countries



Canada
Alberta



United Kingdom
Yorkshire and The Humber



Finland
Eastern and Northern Finland



Ireland
Southern and Eastern

Well-being in detail

Access to services

Utah reaches 8.5 /10 points in Access to services.



This puts the region in position 5 /51 regions in United States.



Civic Engagement

Utah reaches 2.4 /10 points in Civic Engagement.



This puts the region in position 44 /51 regions in United States.



Education

Utah reaches 9.5 /10 points in Education.



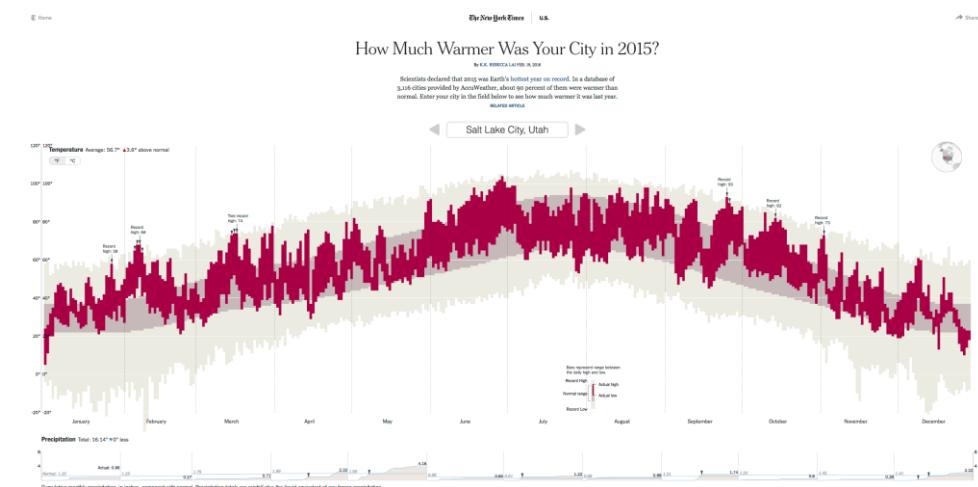
This puts the region in position 23 /51 regions in United States.



Details on Demand
Highlights when desired

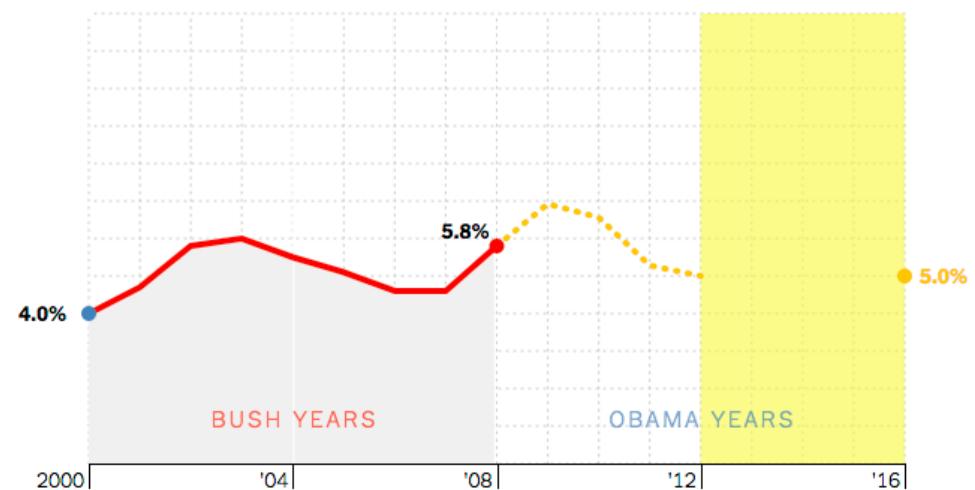
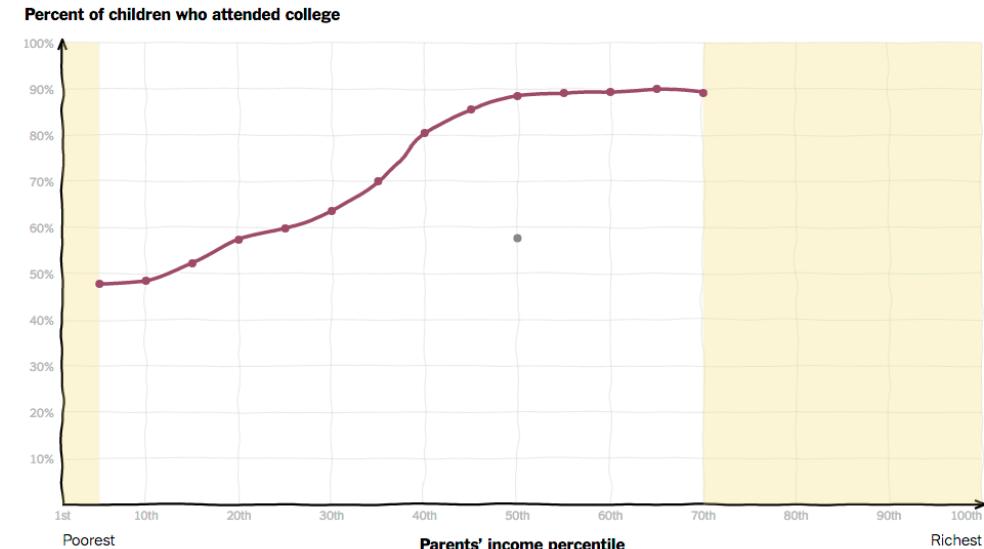
Making it relevant to the reader

What do you think?
Who are you?
Where do you live?



ASK FOR OPINIONS/ PRIOR KNOWLEDGE

Ask reader to draw a trend
Juxtapose with
reality
what other people think



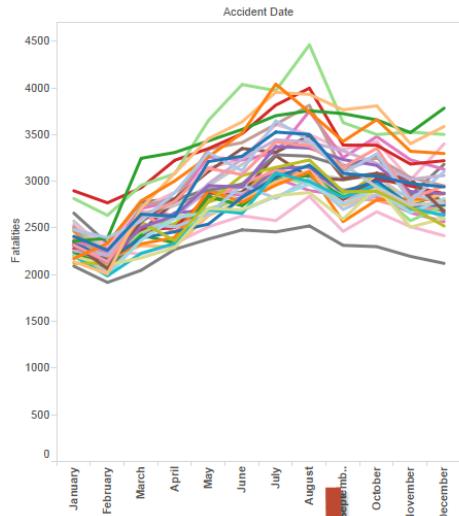
Show me how I did.

DESIGN CONSIDERATION

Fewer people are dying on US roads, but seasonal trends

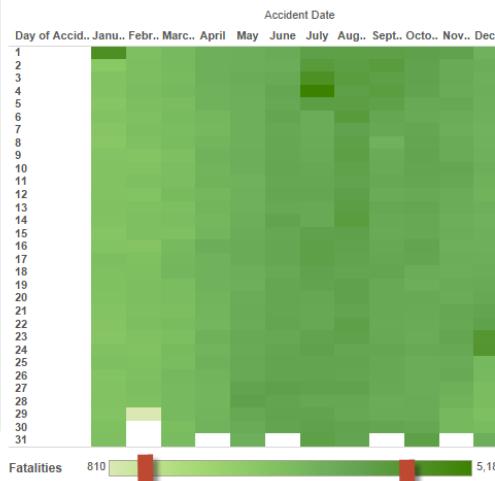
There have been 1,236,670 road fatalities in the US between 1979 and 2011. This dashboard explores the seasonality in the data.

1. Every year sees the same peaks and troughs in fatalities



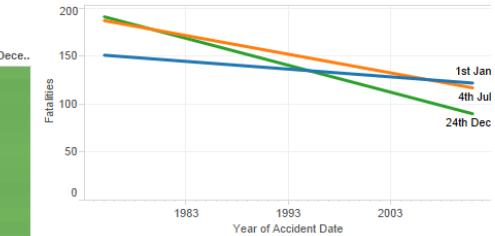
2. Three holidays (New Year, July 4th and Christmas) are particularly lethal

Highlight table of fatalities by day and month for years: All

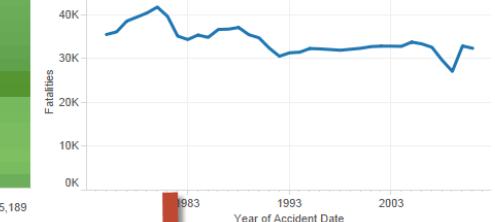


3. Fortunately, those holidays are getting safer...

Slope chart of deaths on the 3 most dangerous days



...which is part of an overall reduction in fatalities



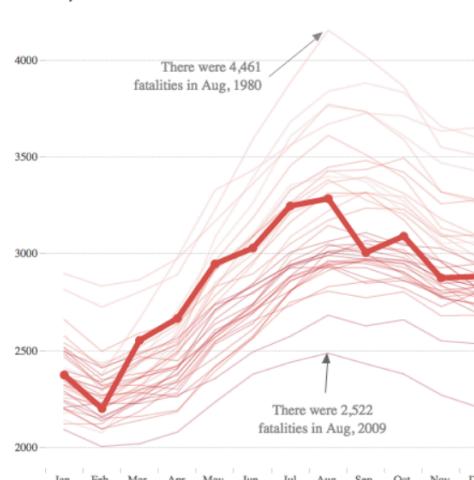
Data from the National Highway Traffic Safety Administration
<http://www-nrd.nhtsa.dot.gov/Cats/Index.aspx>

Fewer people are dying on US roads, but seasonal trends persist.

There have been 1,263,670 road fatalities in the US between 1979 and 2011. This dashboard explores the seasonality in the data.

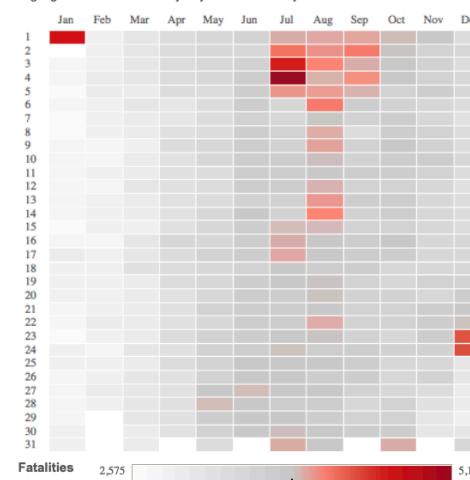
1. Every year sees the same peaks (August) and troughs (February) in fatalities

Fatalities by month. — Single year — Average (all years)



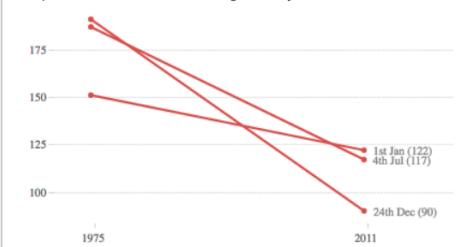
2. Three holidays (New Year, July 4th, Christmas) are particularly lethal

Highlight table of fatalities by day and month for years: All

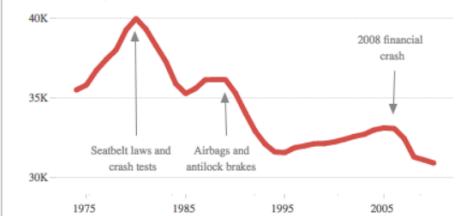


3. Fortunately, those holidays are getting safer...

Slope chart of deaths on 3 most dangerous days



... which is part of an overall reduction in fatalities



Fewer Colors
Averages for
Context
Better Scales
Richer
Annotations

ENGAGEMENT

KNOW YOUR AUDIENCE

People you don't know are difficult to influence

What do they know?

What motivates them?

What experiences do you share?

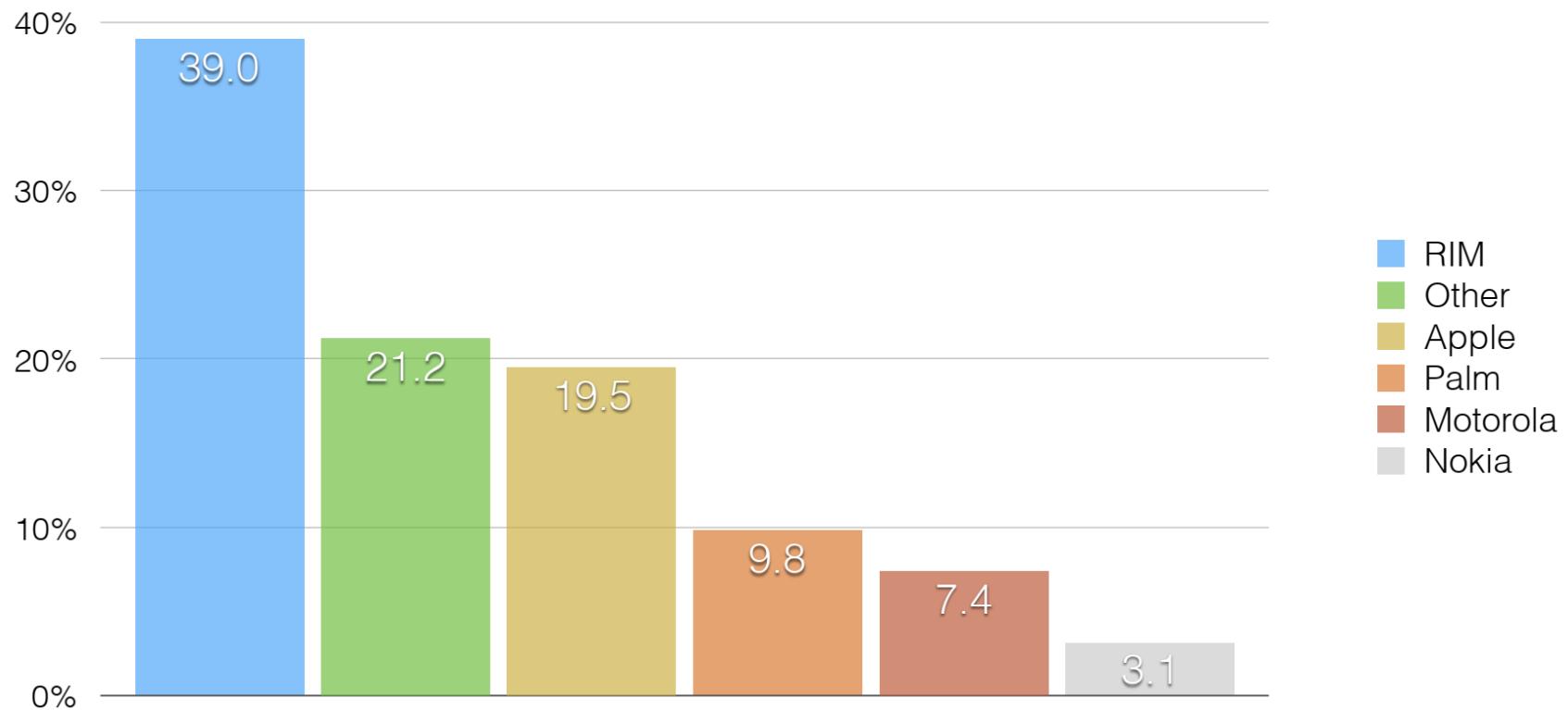
What are common goals?

What insights can you give them?

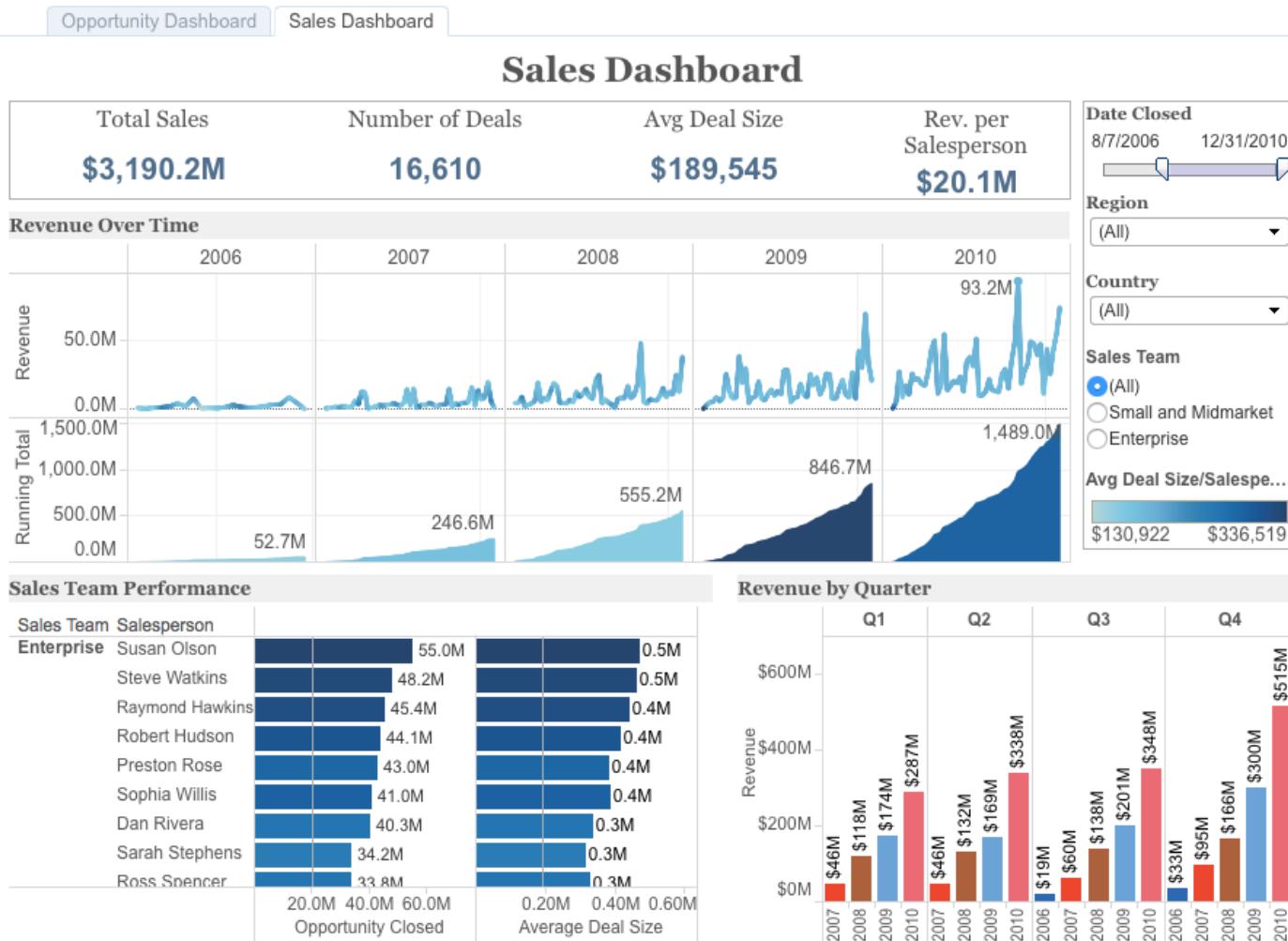
TARGET AUDIENCE



TARGET AUDIENCE?



TARGET AUDIENCE?



WHERE THERE'S SMOKE—THERE'S CANCER

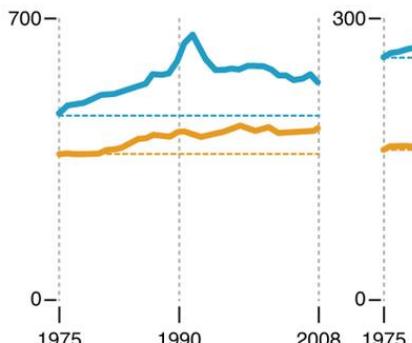
Cancer rates are up, but mortality is down. New diagnostics and treatments are responsible for part of this trend. But the greatest single contributing factor is the decline in smoking—rates are at their lowest level in 50 years.

Men Women

1 Increased incidence

An aging population contributes to rising incidence of cancer.

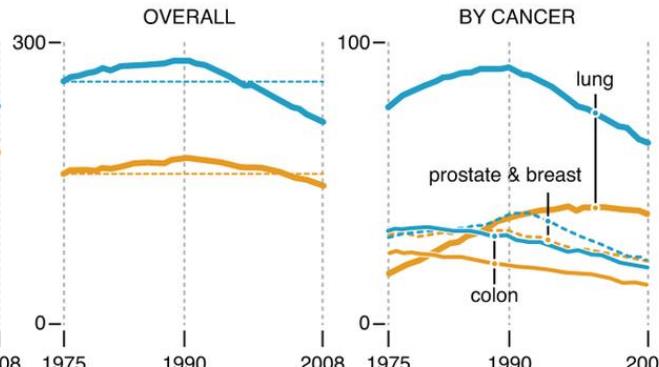
Cancer incidence rates (per 100,000)



2 Fewer deaths

Cancer deaths have been dropping since 1991, especially in males.

Cancer death rates (per 100,000)



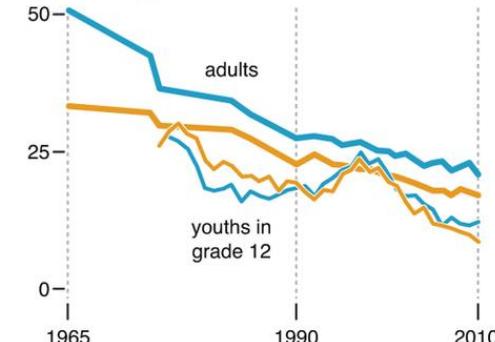
3 Decline of lung cancer

Drop in lung cancer deaths in males is the primary reason why death rates are down.

4 Decline in smoking

Since the 1964 first Surgeon General's report, smoking rates have been dropping. By 2010, the rate among males was down to 20%, from 50% at its peak. Among youths, rates have been on an even steeper decline since 1997.

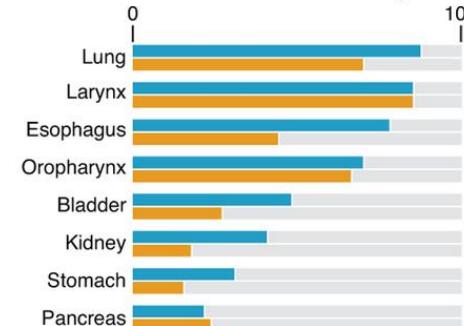
Smoking prevalence (%)



5 Impact of smoking on cancer deaths

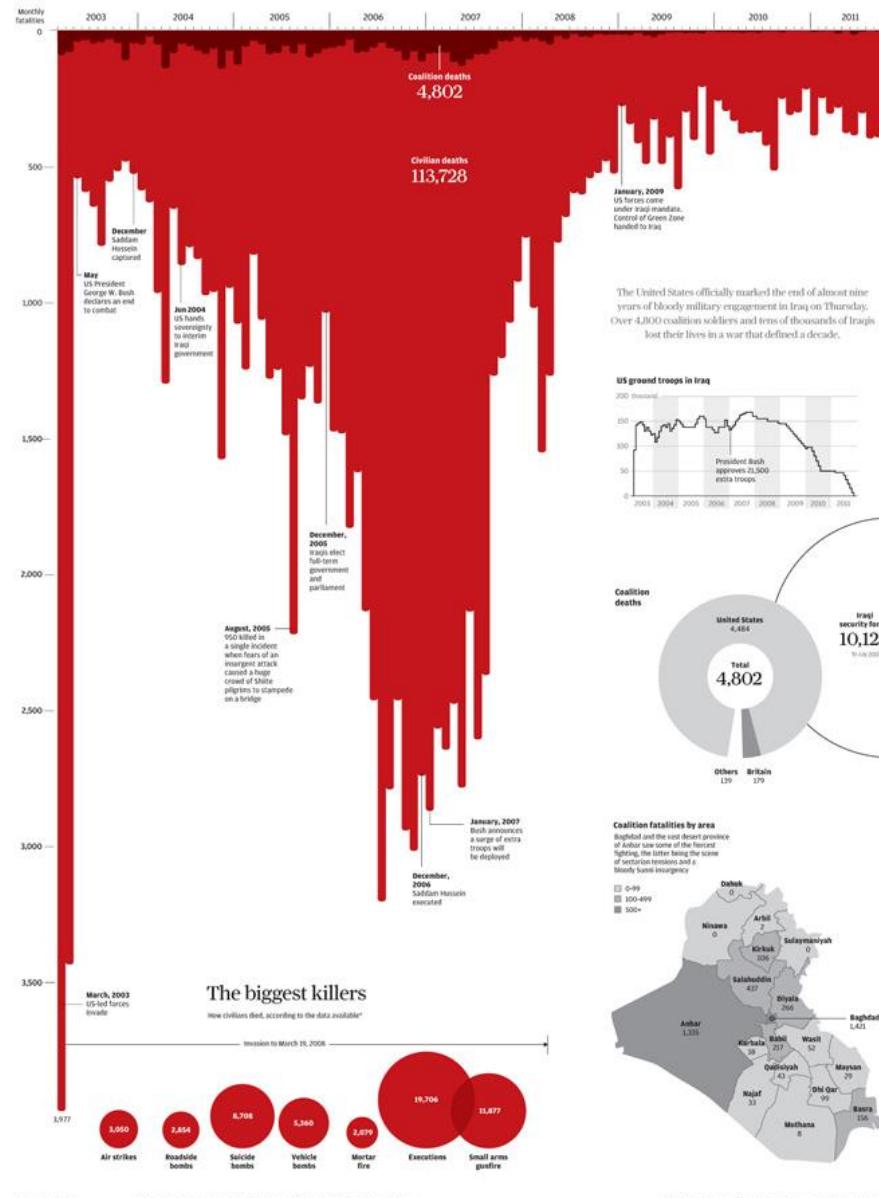
Smoking is a major risk factor for many types of cancer and significant contributor to cancer-related deaths. It remains the single largest preventable cause of disease and premature death in the US.

Percentage of cancer deaths attributable to smoking

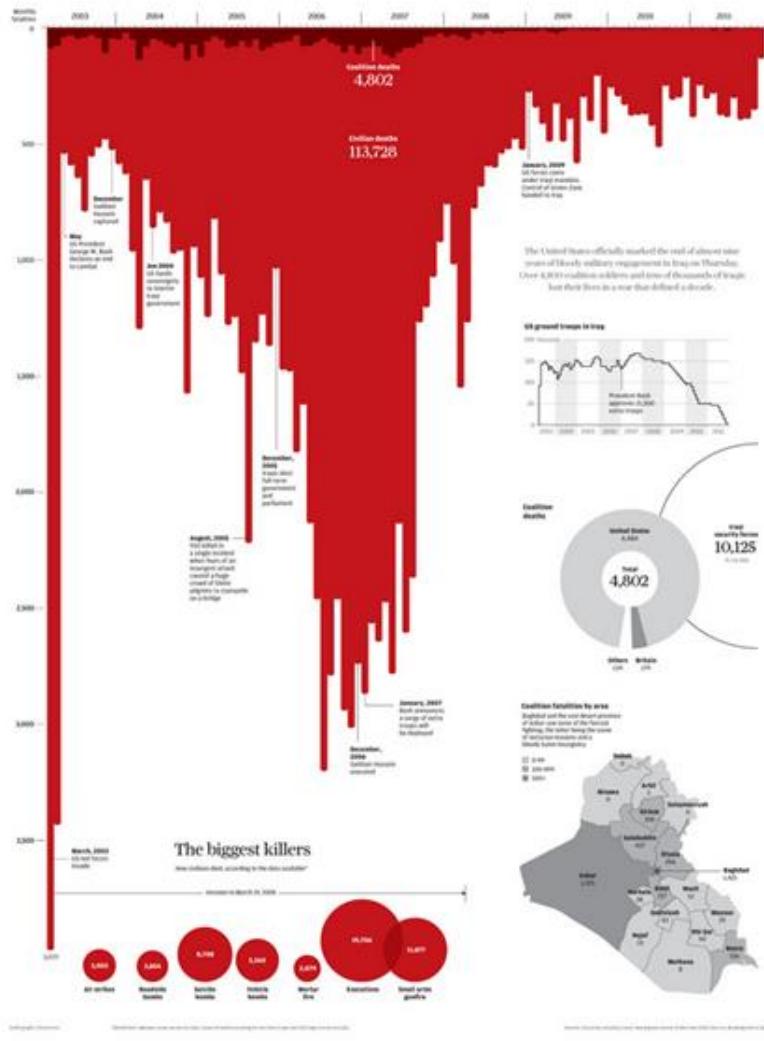


source: American Cancer Society Cancer Statistics 2012; Monitoring the Future (University of Michigan).

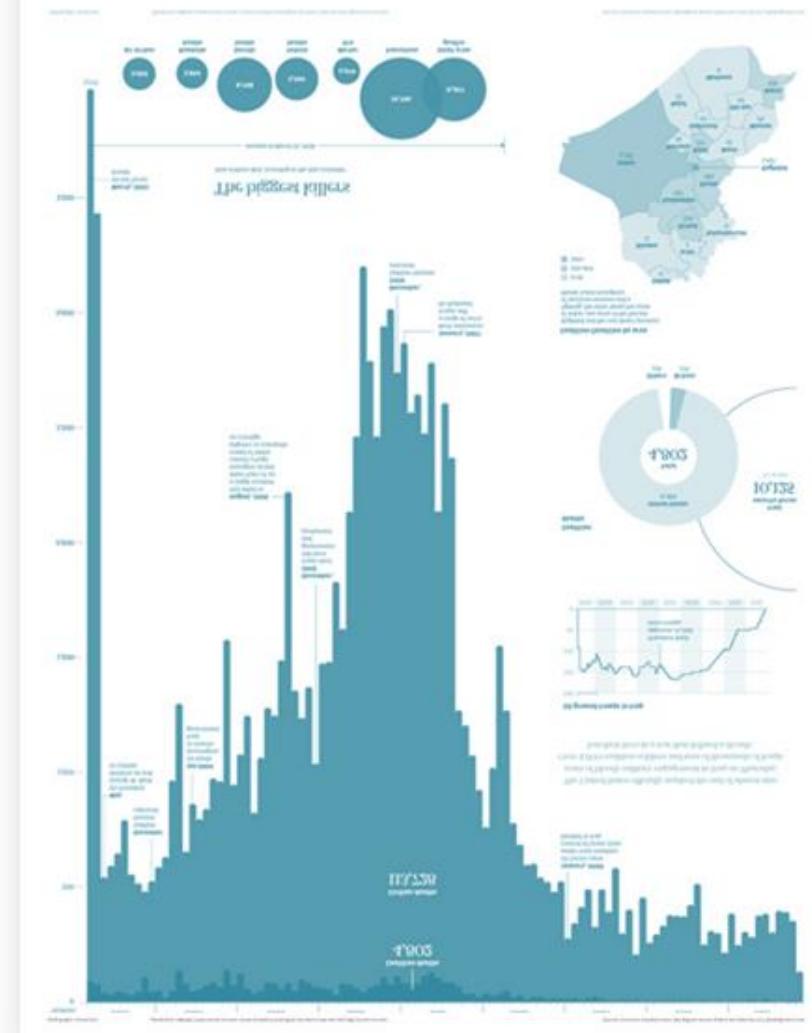
Iraq's bloody toll



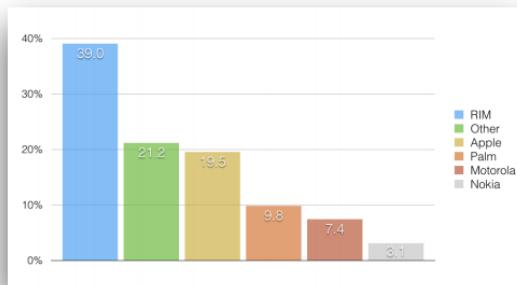
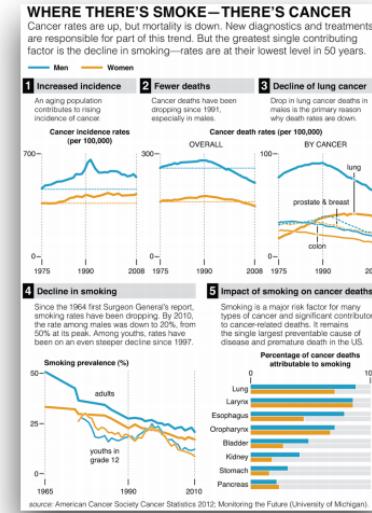
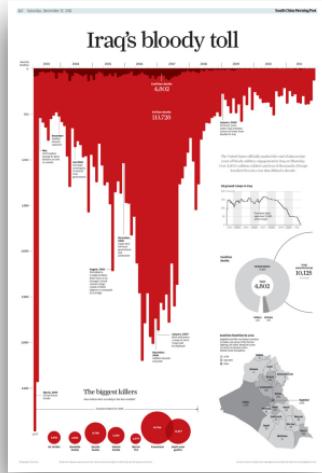
Iraq's bloody toll



Iraq: Deaths on the Decline



Opinionated

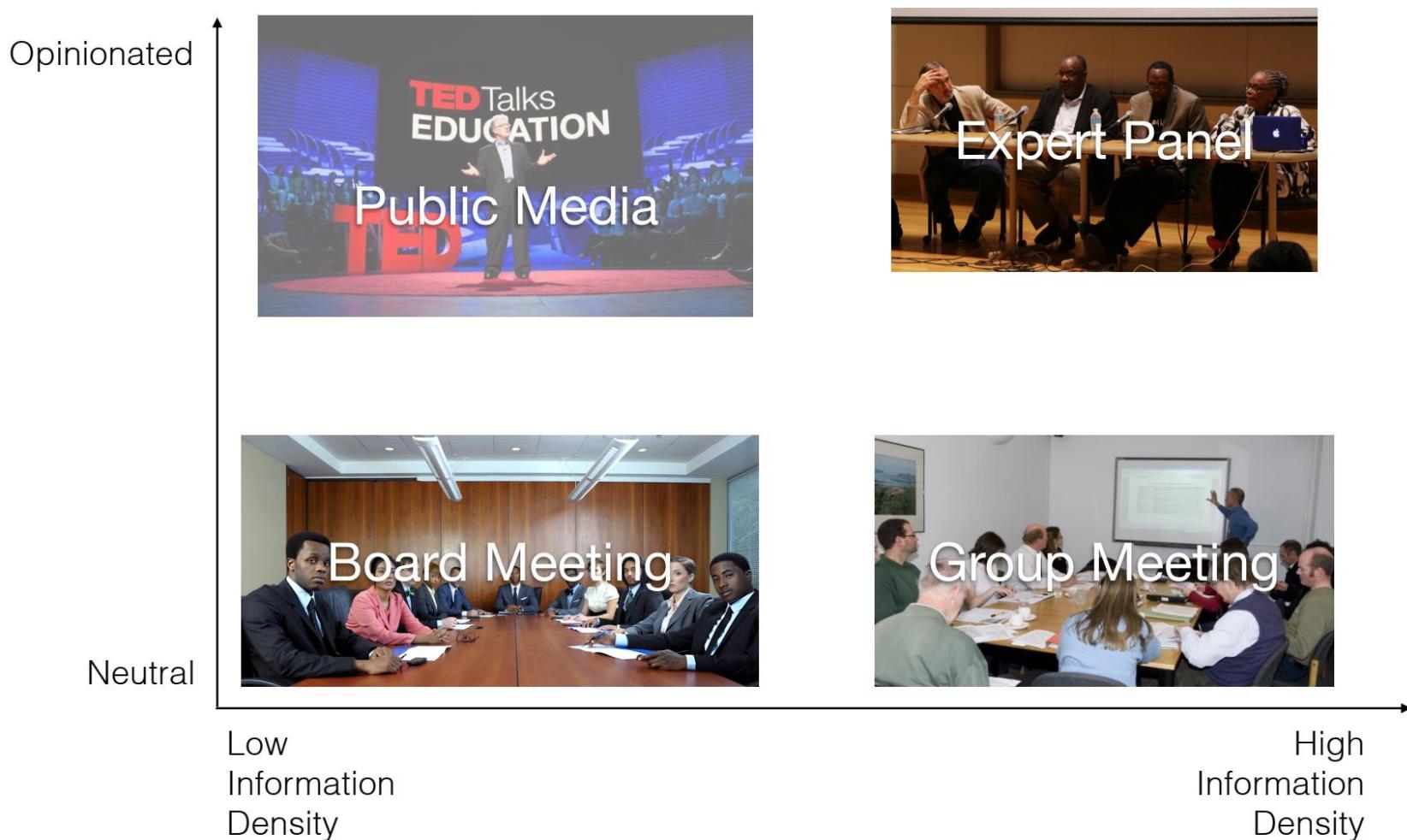


Neutral

Low
Information
Density

High
Information
Density

TARGET AUDIENCE



EXERCISES

What's the Genre?

(Magazine, Annotated Chart, Partitioned Poster,
Flow Chart, Comic Strip, Slide Show,
Video/Animation)

Is it Author or Reader Driven?

Can you make out a particular structure?
(Martini Glass, Interactive Slideshow,
Drill-Down Story)? Why?

Does it have a good Introduction?
What medium is used for the introduction?

Is the story opinionated or neutral?

Does it use titles, legends, and highlights well?

Does it use interactivity well?

Is the visualization well designed?

1.5 Million Missing Black Men

By JUSTIN WOLFERS, DAVID LEONHARDT and KEVIN QUEALY APRIL 20, 2015

For every 100 black women not in jail, there are only 83 black men. The remaining men – 1.5 million of them – are, in a sense, **missing**.



Among cities with sizable black populations, the largest single gap is in **Ferguson, Mo.**



North Charleston, S.C., has a gap larger than 75 percent of cities.



This gap – driven mostly by incarceration and early deaths – **barely exists among whites**.



Figures are for non-incarcerated adults who are 25 to 54.

What's the Genre?

(Magazine, Annotated Chart, Partitioned Poster,
Flow Chart, Comic Strip, Slide Show,
Video/Animation)

Is it Author or Reader Driven?

Can you make out a particular structure?
(Martini Glass, Interactive Slideshow,
Drill-Down Story)? Why?

Does it have a good Introduction?
What medium is used for the introduction?

Is the story opinionated or neutral?

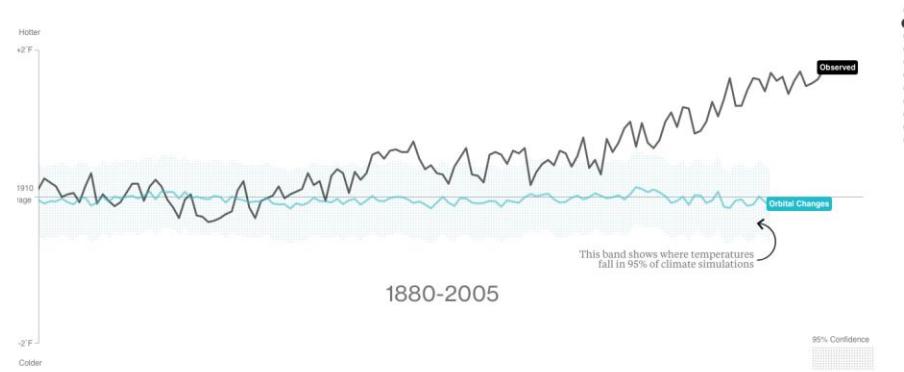
Does it use titles, legends, and highlights well?

Does it use interactivity well?

Is the visualization well designed?

Is It the Earth's Orbit?

The Earth wobbles on its axis, and its tilt and orbit change over many thousands of years, pushing the climate into and out of ice ages. Yet the influence of orbital changes on the planet's temperature over 125 years has been negligible.



What's the Genre?

(Magazine, Annotated Chart, Partitioned Poster, Flow Chart, Comic Strip, Slide Show, Video/Animation)

Is it Author or Reader Driven?

Can you make out a particular structure?
(Martini Glass, Interactive Slideshow, Drill-Down Story)? Why?

Does it have a good Introduction?
What medium is used for the introduction?

Is the story opinionated or neutral?

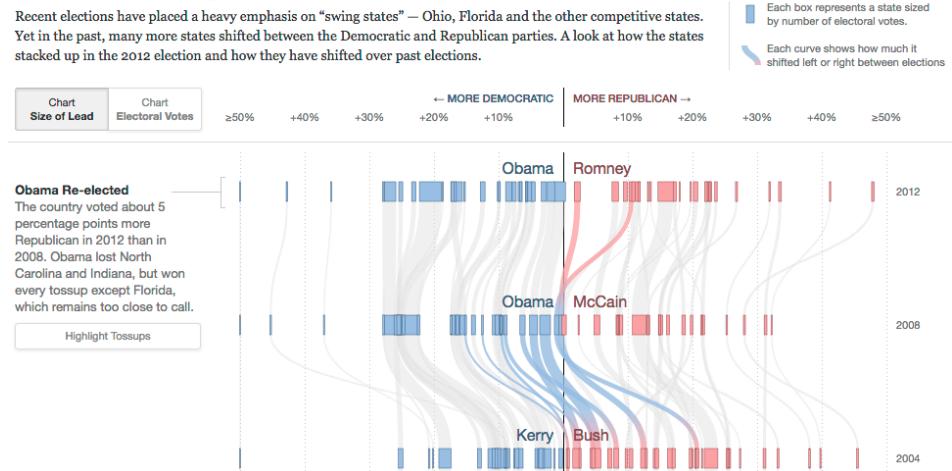
Does it use titles, legends, and highlights well?

Does it use interactivity well?

Is the visualization well designed?

Over the Decades, How States Have Shifted

Recent elections have placed a heavy emphasis on "swing states" — Ohio, Florida and the other competitive states. Yet in the past, many more states shifted between the Democratic and Republican parties. A look at how the states stacked up in the 2012 election and how they have shifted over past elections.



EVALUATION

PROBLEM-DRIVEN VS TECHNIQUE-DRIVEN

Problem-driven

top-down approach

identify a problem encountered by users

design a solution to help users work more effectively

sometimes called a design study

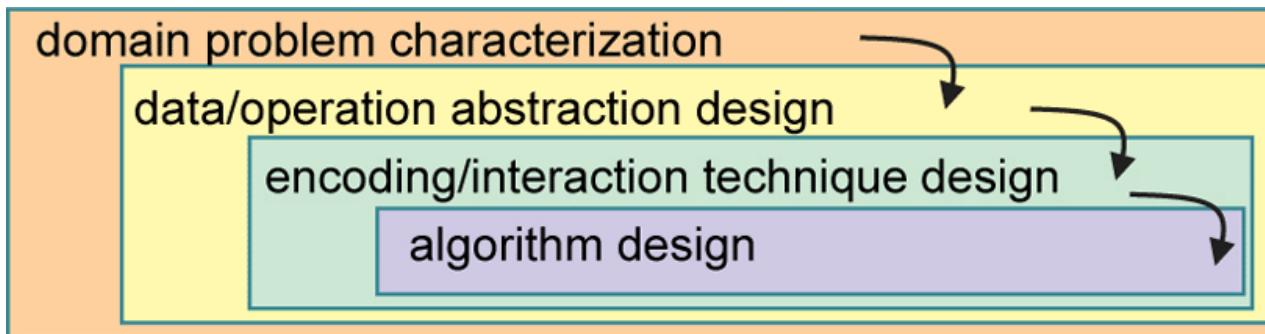
Technique-driven

bottom-up approach

invent new visualization techniques or algorithms

classify or compare against other idioms and algorithms

THE NESTED MODEL



Design

threat: wrong problem

validate: observe and interview target users

threat: bad data/operation abstraction

threat: ineffective encoding/interaction technique

validate: justify encoding/interaction design

threat: slow algorithm

validate: analyze computational complexity

implement system

validate: measure system time/memory

validate: qualitative/quantitative result image analysis

[test on any users, informal usability study]

validate: lab study, measure human time/errors for operation

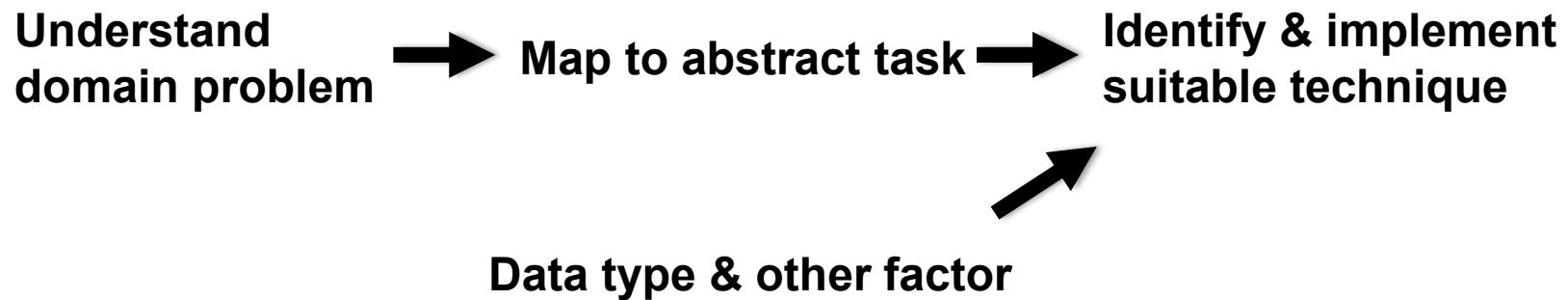
validate: test on target users, collect anecdotal evidence of utility

validate: field study, document human usage of deployed system

validate: observe adoption rates

Threats & evaluation

DESIGN PROCESS



DOMAIN CHARACTERIZATION

Details of an application domain

Group of users, target domain, their questions, their data

varies wildly by domain

must be specific enough to continue with

Cannot just ask people what they do

introspection is hard

DOMAIN PROBLEM CHARACTERIZATION

Infinite numbers of domain tasks

Can be broken down into simpler abstract tasks

We know how to address the abstract tasks

**Identify task – data combination:
solutions probably exist**

EXAMPLE: FIND GOOD MOVIES

I want to identify good movies in genres I like

Domain: general population, movie enthusiasts

DATA & TASK ABSTRACTION

The what-why, map into generalized terms

Identify tasks that users wish to perform or already do

Find data types and good model of the data

Sometimes must transform the data for a better solution

this can be varied and guided by the specific task

EXAMPLE: FIND GOOD MOVIES

What is a good movie for me?

Highly rated by critics?

Highly rated by audiences?

Successful at the box office?

Similar to movies I liked?

Specific genres?

Data sources: IMDB, Rotten Tomatoes, ...

ENCODINGS & INTERACTIONS

The design of visualization techniques

Visual encodings

Interactions

Ways to create and manipulate the visual representation of data

Decisions on these may be separate or intertwined

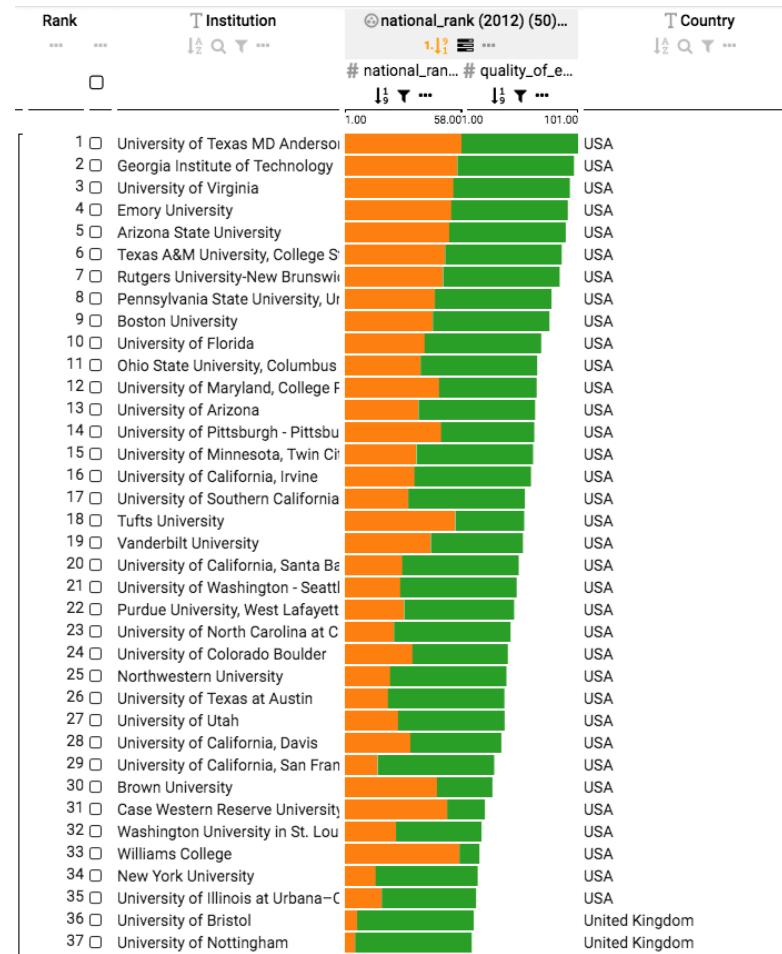
Visualization design principles drive decisions

EXAMPLE: FIND GOOD MOVIES

Combination of audience ratings and critics ratings, filtered by genre

Idiom:
stacked bar chart for ratings

filter interface for genre



TASKS

Analyze

High-level choices

Consume vs produce

Search

Find a known/unknown item

Query

Find out about characteristics of item

By itself or relative to others

HIGH-LEVEL ACTIONS: ANALYZE

Consume

Discover vs present

Classic split: explore vs explain

Enjoy: casual, social

Analyze

→ Consume

→ Discover



→ Present



→ Enjoy



Produce

Annotate, record

Derive: crucial design choice

→ Produce

→ Annotate



→ Record



→ Derive



MID-LEVEL ACTIONS: SEARCH, QUERY

Search: what does user know?

Target, location



Search

	Target known	Target unknown
Location known	 <i>Lookup</i>	 <i>Browse</i>
Location unknown	 <i>Locate</i>	 <i>Explore</i>

How much of the data matters?

One, some, all



Query

→ Identify



→ Compare



→ Summarize



LOW LEVEL: TARGETS

→ All Data

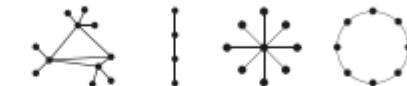
- Trends
 - Outliers
 - Features
-

→ Attributes

- One
 - Distribution
 - Extremes
 - Many
 - Dependency
 - Correlation
 - Similarity
-

→ Network Data

- Topology



- Paths



→ Spatial Data

- Shape



DESIGNING VISUALIZATIONS

WHAT IS DESIGN?

creating something new to solve a problem

can be used to make buildings, chairs, user interfaces, etc.

design is used in many fields

many possible users or tasks

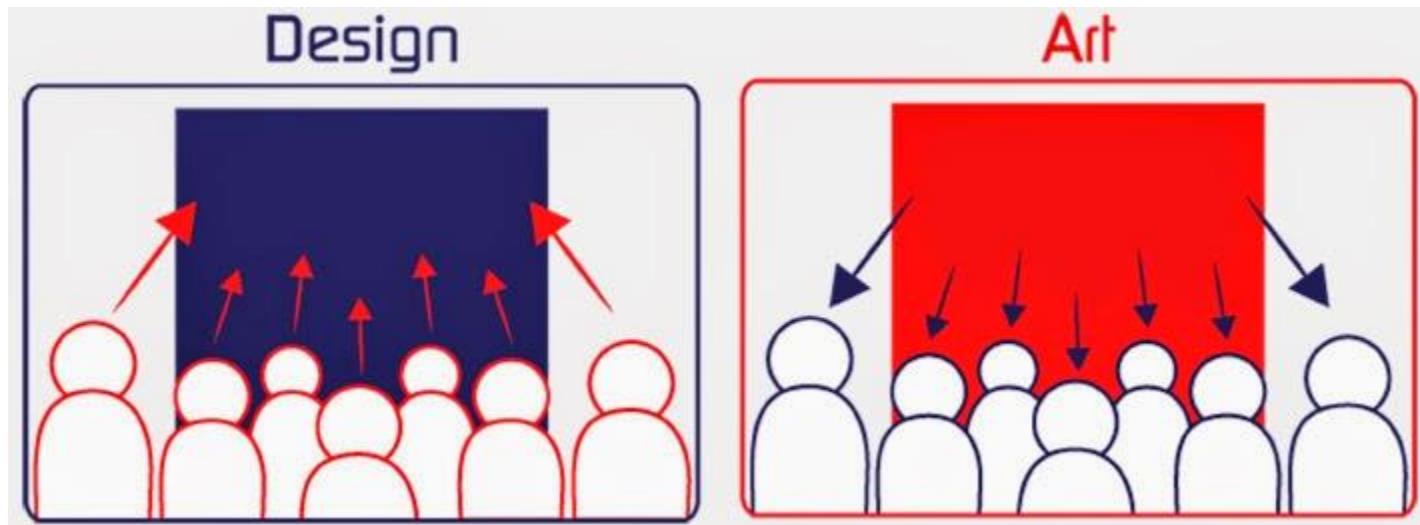
WHAT IS DESIGN NOT?

just making things pretty

art – appreciation of beauty or emotions invoked

something without a clear purpose

building without justification or evidence



FORM & FUNCTION

Commonly: “Form follows function”

Function can constrain possible forms

Form depends on tasks that must be achieved

**“The better defined the goals of an artifact,
the narrower the variety of forms it can adopt”**

- Alberto Cairo

WHEN DO WE DESIGN?

Wicked problems

No clear problem definition

Solutions are either good enough or not good enough

Multiple solutions exist, not true/false

No clear point to stop with a solution

Examples of non-wicked (“tame”) problems

Mathematics, chess, puzzles

WHY DOES DESIGN MATTER FOR VIS?

Many ineffective visualization combinations

Users with unique problems & data

Variations of tasks

Large design space

PITFALL

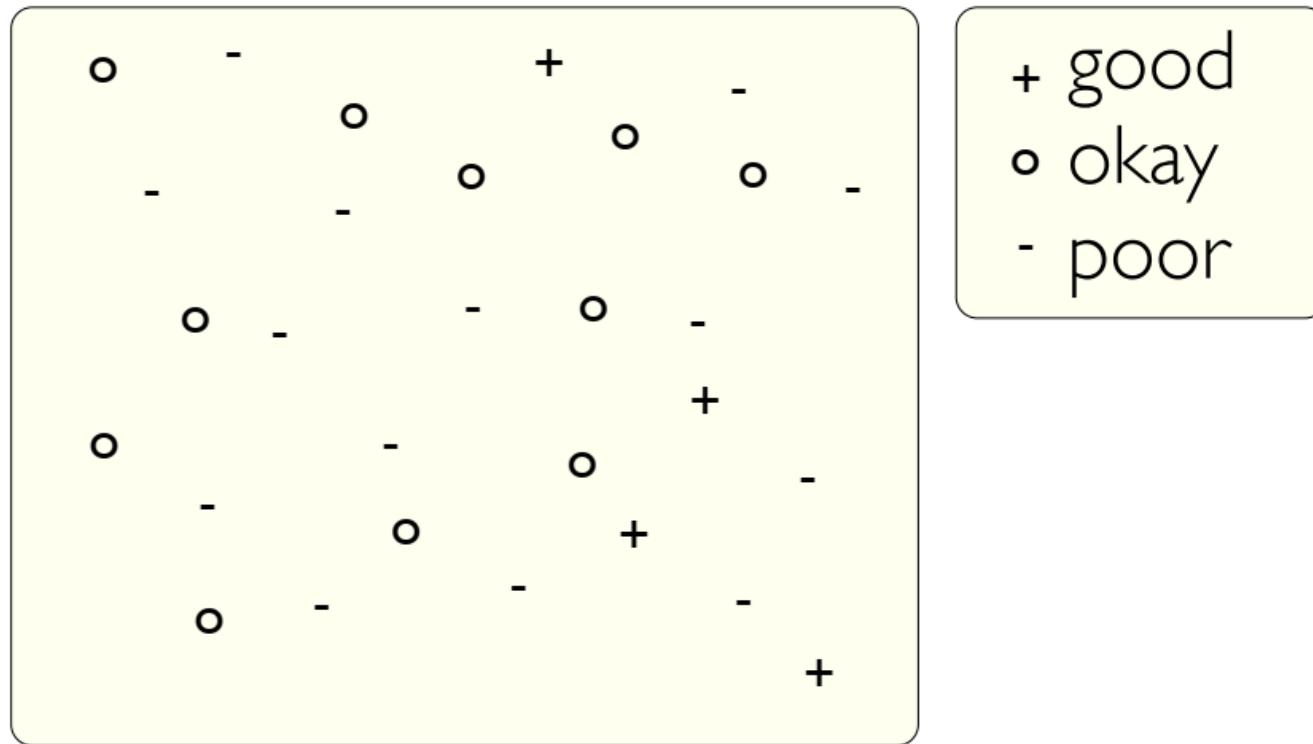
PREMATURE DESIGN COMMITMENT

Of course they need the cool
technique I built last year!



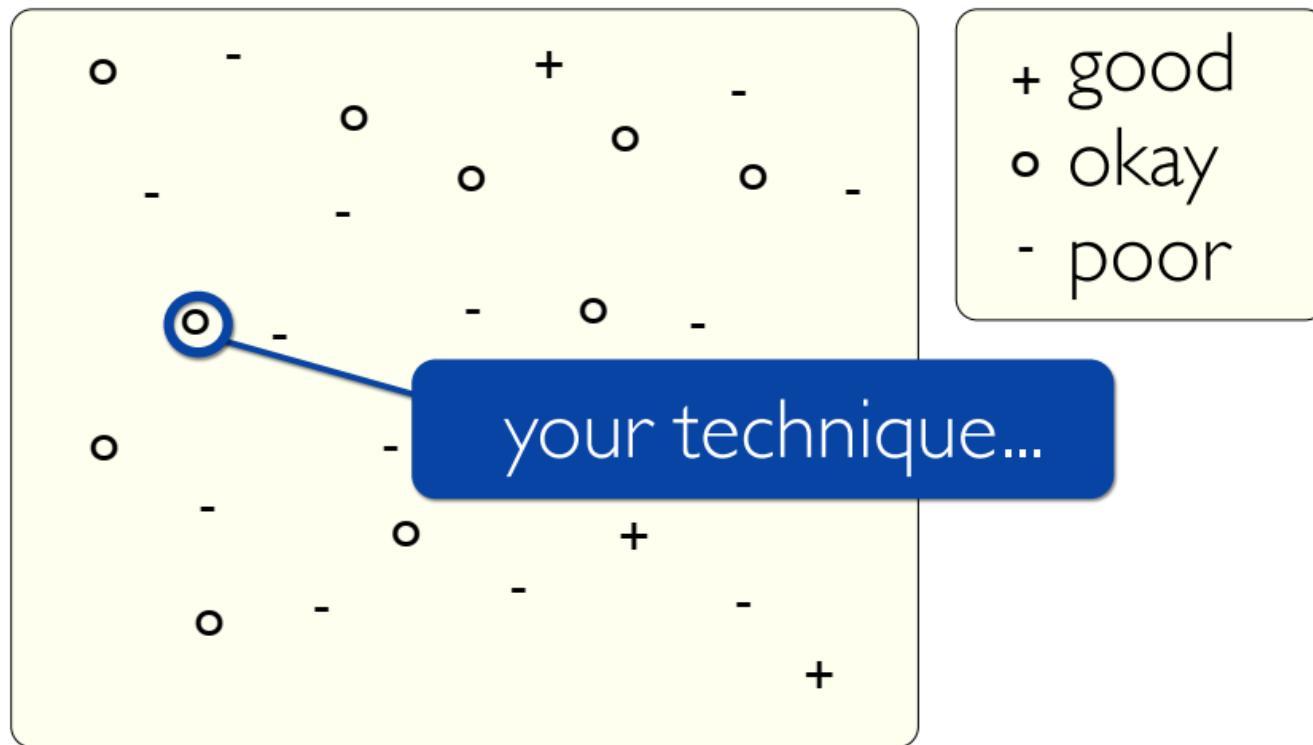
by T. Munzner

METAPHOR **Design Space**



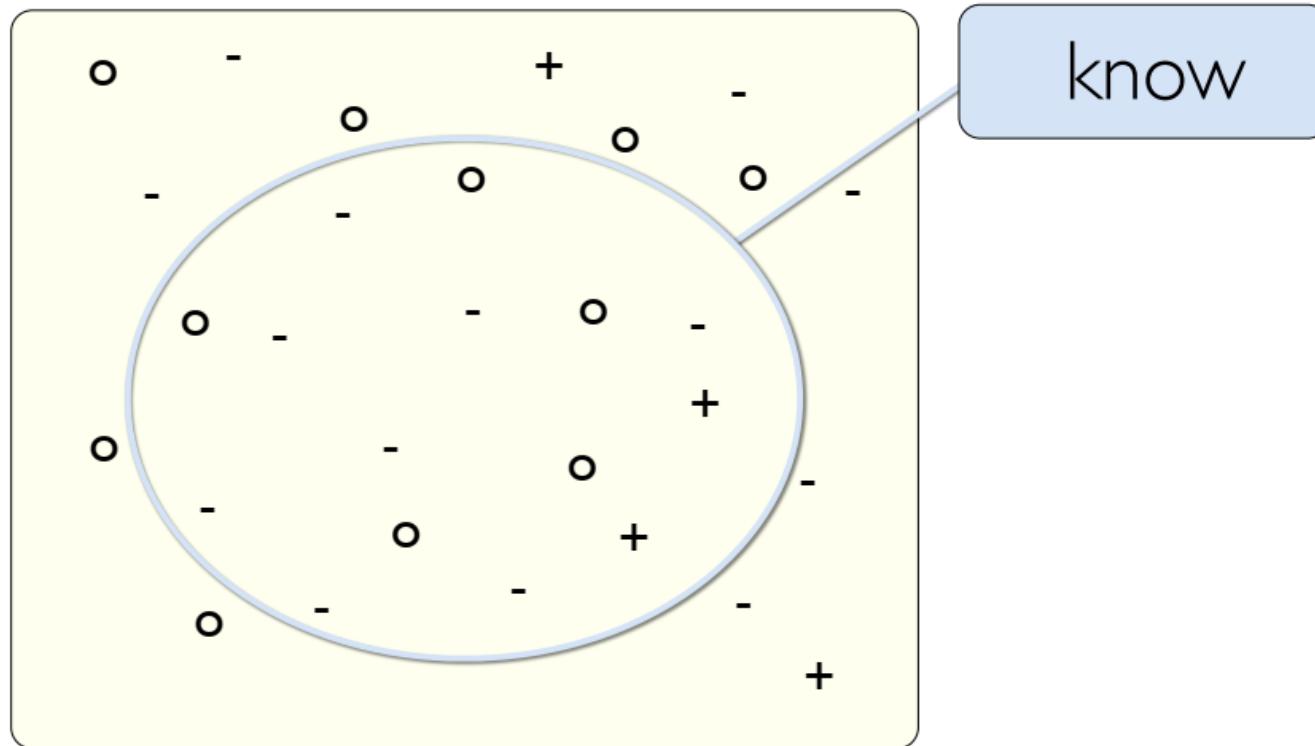
by T. Munzner

METAPHOR **Design Space**



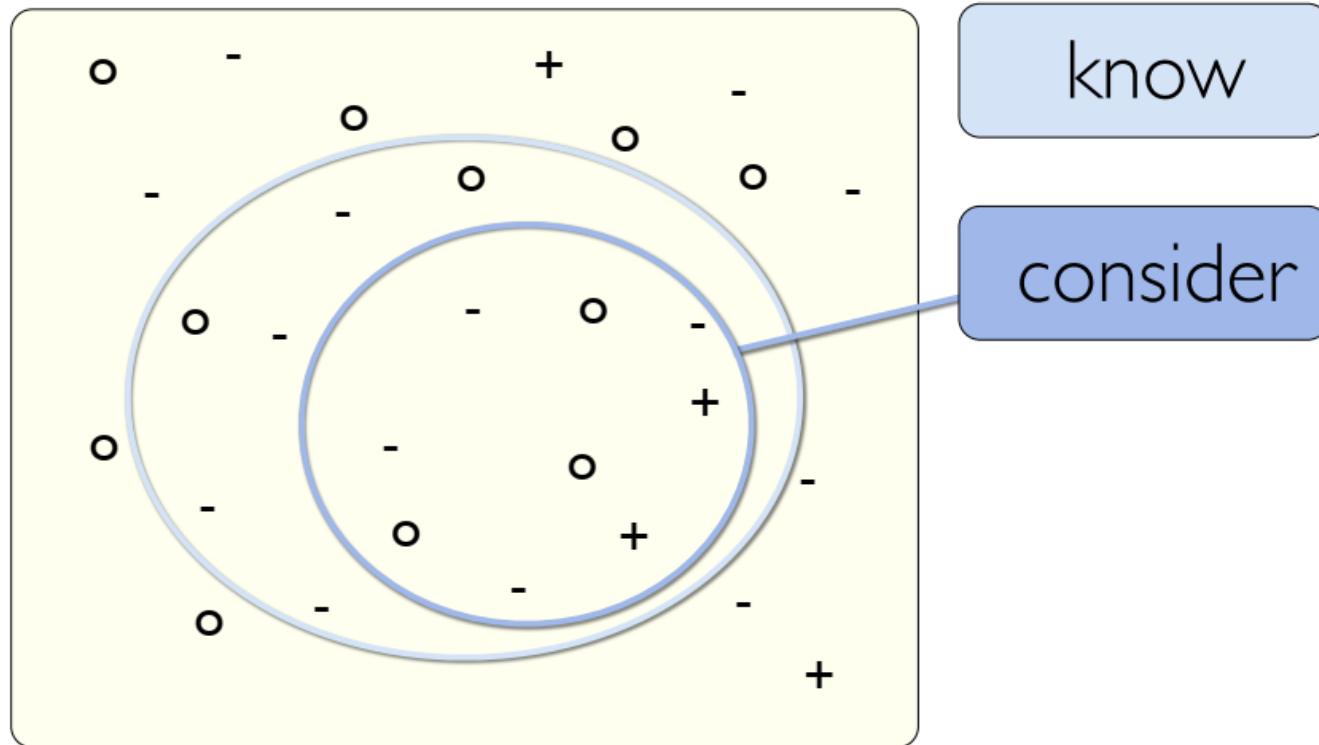
by T. Munzner

METAPHOR **Design Space**



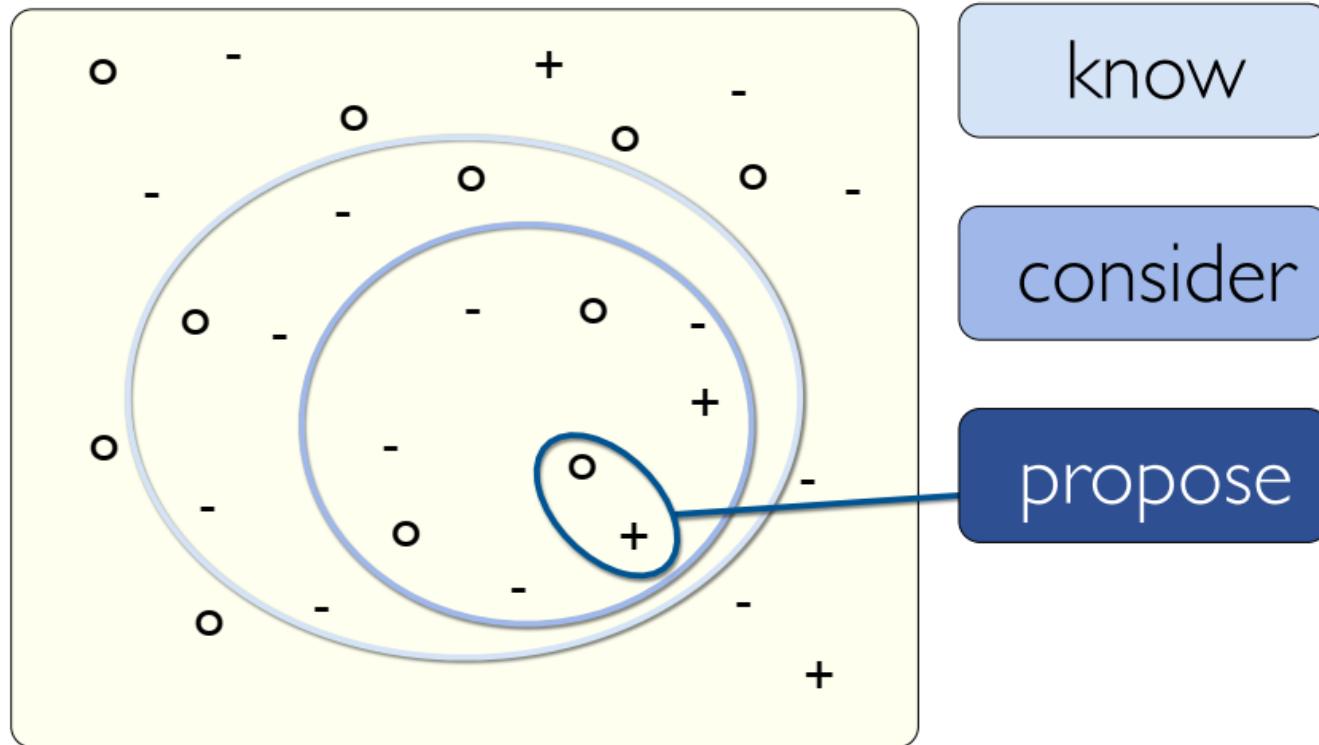
by T. Munzner

METAPHOR **Design Space**



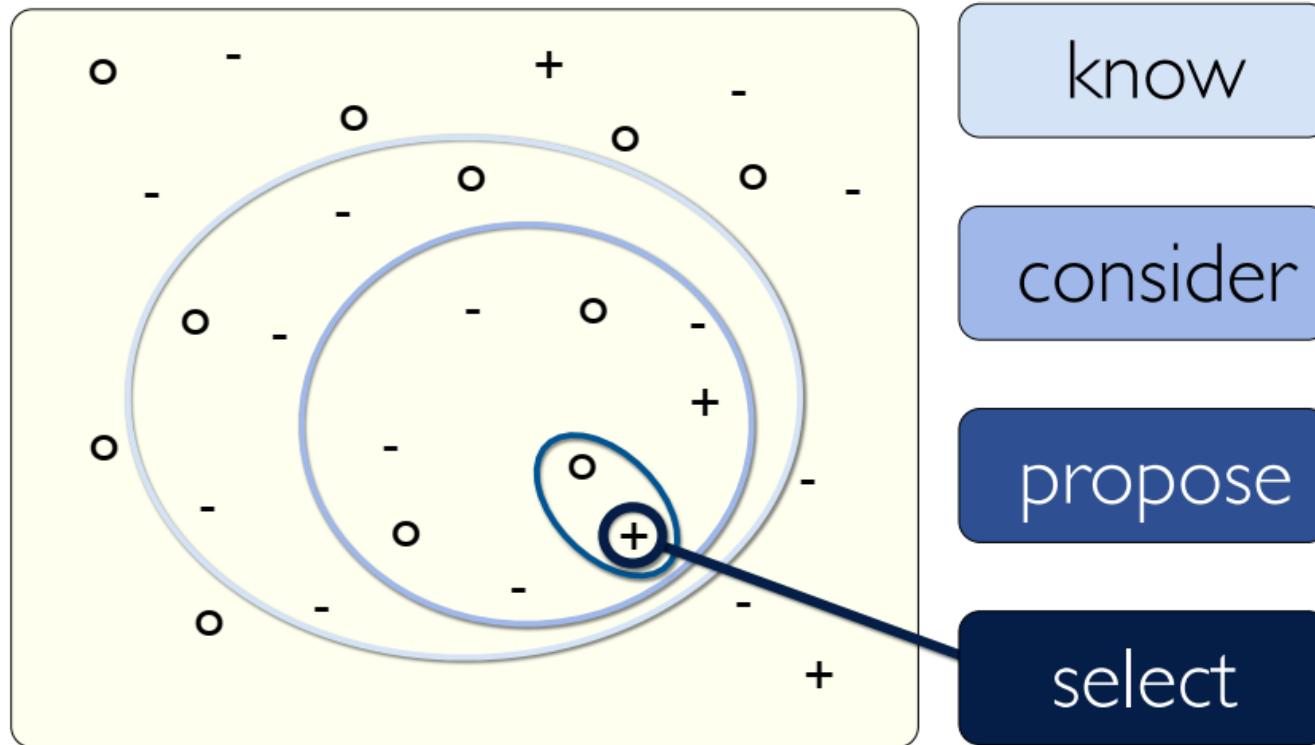
by T. Munzner

METAPHOR **Design Space**



by T. Munzner

METAPHOR **Design Space**



METAPHOR Design Space



EVALUATION METHODS

Controlled experiment

Laboratory, crowd-sourced

Interviews/questionnaires

Unstructured, structured, semi-structured

Field observation, lab observation

Video/audio analysis

Coding/classification of user behavior (speech, gestures)

Log analysis

Algorithmic performance measurement

EVALUATION METHODS

Heuristic evaluation

Judge compliance with recognized metrics/usability methods (heuristics)

Usability testing, e.g., thinking aloud tests

Wizard of Oz

Human simulates response of system

Test functionality before it's implemented

Eye tracker evaluation

Expert evaluation

Insight-based evaluation

Case studies

QUANTITATIVE VS QUALITATIVE EVALUATION

Quantitative methods

Objective metrics, measurements

Use numbers/statistics for interpreting data

Qualitative methods

Subjective metrics

Description of situations, events, people, interactions, and observed behaviors, the use of direct quotations from people about their experiences, attitudes, beliefs, and thoughts

Focused on understanding how people make meaning of and experience their environment or world

INTERNAL VS EXTERNAL VALIDITY

Internal validity – can you trust your experiment

High when tested under controlled lab conditions

Observed effects are due to the test conditions
(and not random variables)

External validity – is your experiment representative of real world usage

High when interface is tested in the field,
e.g. handheld device tested in museum

Results are valid in real world

The trade-off

The more akin to real-world situations, the more experiment is susceptible to uncontrolled sources of variation

SCOPE OF EVALUATION

threat: wrong problem

validate: observe and interview target users

threat: bad data/operation abstraction

threat: ineffective encoding/interaction technique

validate: justify encoding/interaction design

threat: slow algorithm

validate: analyze computational complexity

implement system

validate: measure system time/memory

validate: qualitative/quantitative result image analysis

[test on any users, informal usability study]

validate: lab study, measure human time/errors for operation

validate: test on target users, collect anecdotal evidence of utility

validate: field study, document human usage of deployed system

validate: observe adoption rates

SCOPE OF EVALUATION

Pre-design

To understand potential users' work env. And workflow

Design

To scope a visual encoding and interaction design space based on human perception and cognition

Prototype

to see if a visualization has achieved its design goals, to see how a prototype compares with the current state-of-the-art systems or techniques

Deployment

to see how a visualization influences workflow and work processes, to assess the visualization's effectiveness and use in the field

Re-design

to improve a current design by identifying usability problems