VISUALIZING MULTI-ATTRIBUTE DATA

DATA TABLES

Petra Isenberg

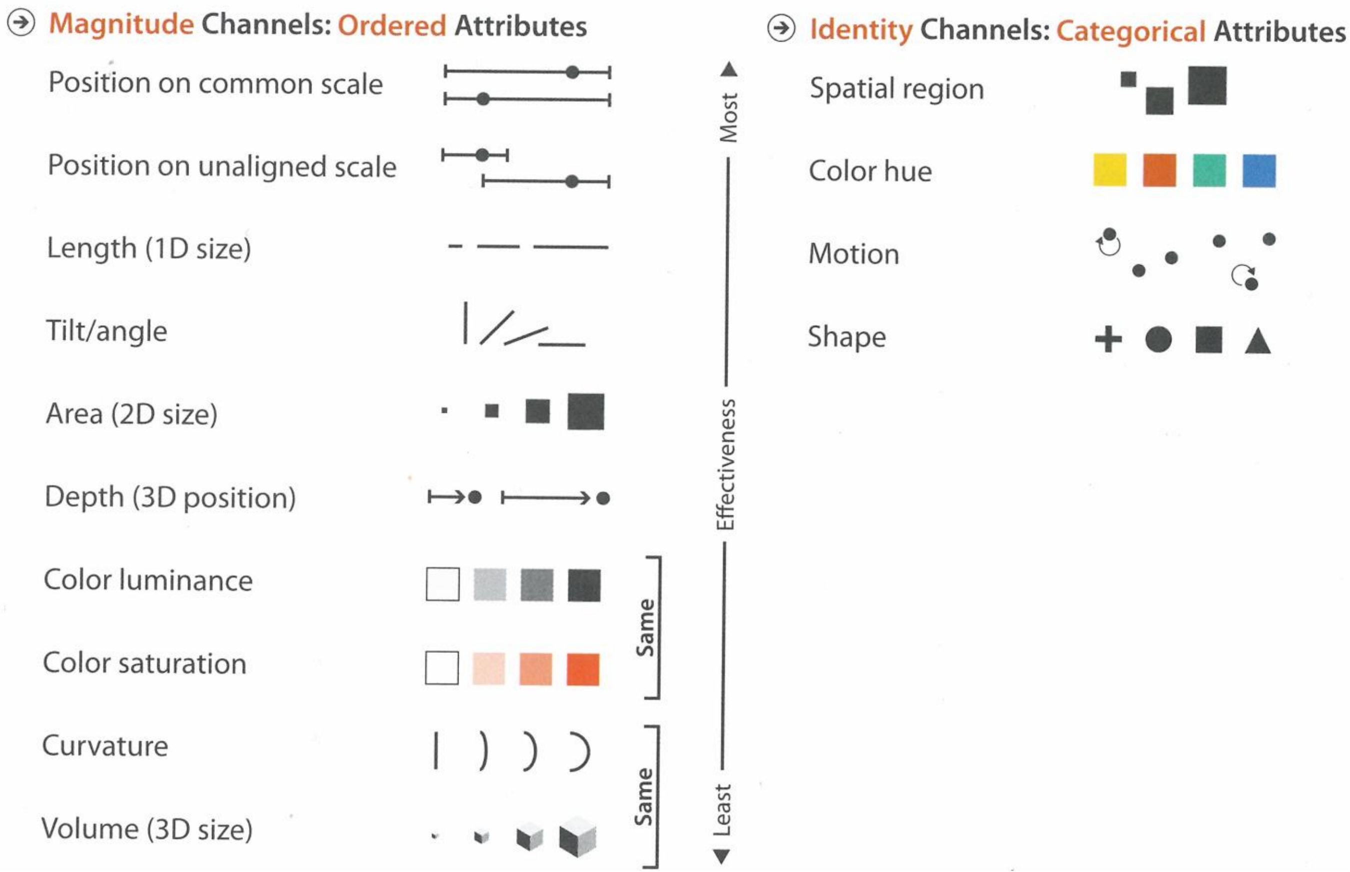
informatics mathematics

RECAP

you have learned about

- visual variables and marks
- that their perceptual properties matter

RECAP

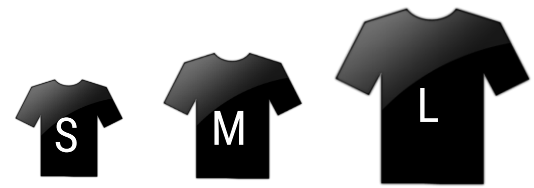


Magnitude Channels: Ordered Attributes
Position on common scale
Position on unaligned scale
Length (1D size)
Tilt/angle
Area (2D size)
Depth (3D position)
Color luminance
Color saturation
Curvature
Volume (3D size)
Same
Same
Most
Effectiveness
Least
Identity Channels: Categorical Attributes
Spatial region
Color hue
Motion
Shape

RECAP

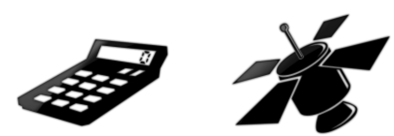
DATA TYPES

ORDINAL (ranking)



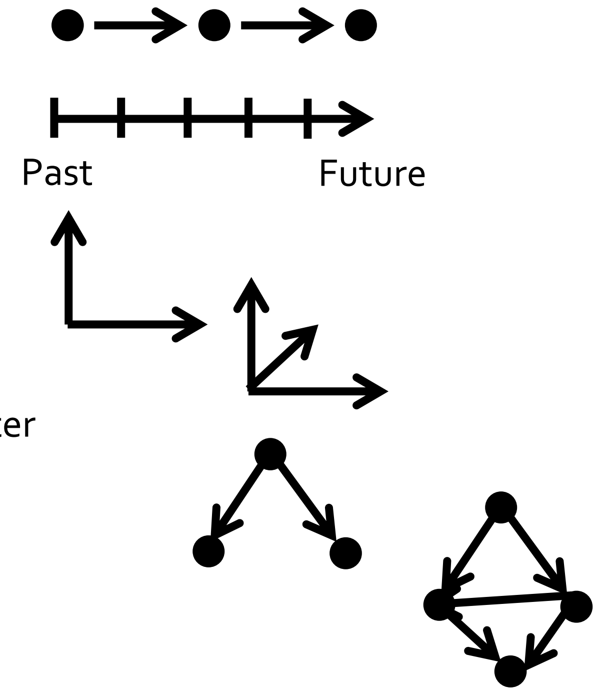
NOMINAL (categorical)

QUANTITATIVE (numerical)



RECAP

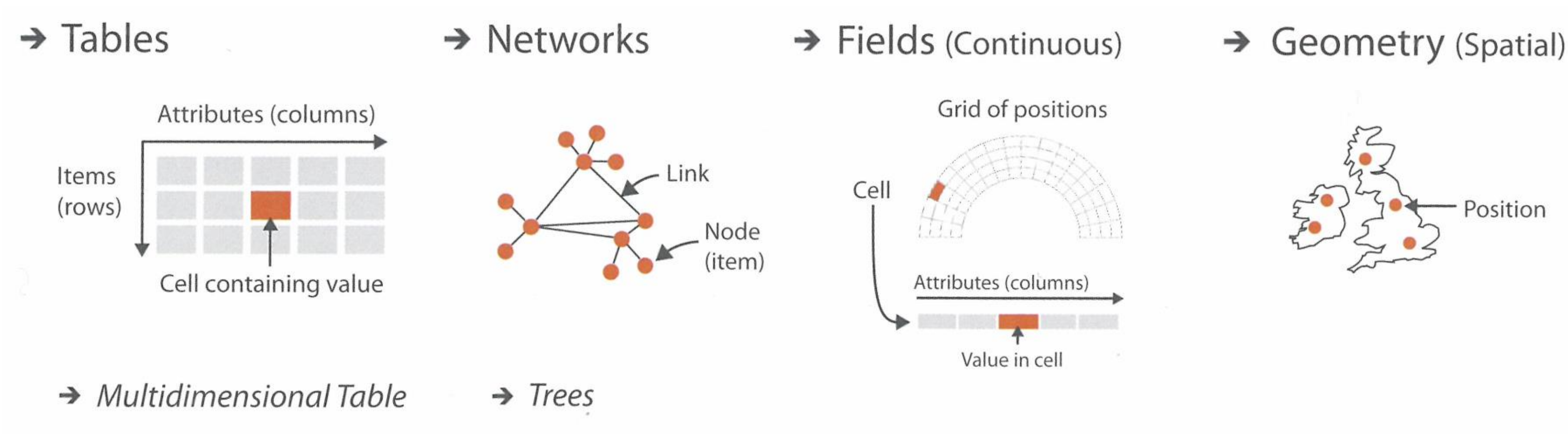
- 1D (linear)
- Temporal
- 2D (maps)
- 3D
- nD (relational)
- Trees (hierarchies)
- Networks (graphs)

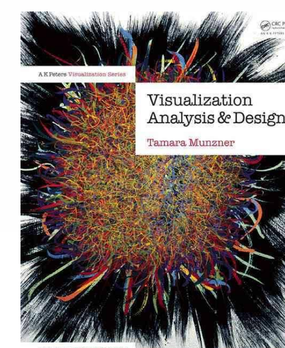


vis examples later

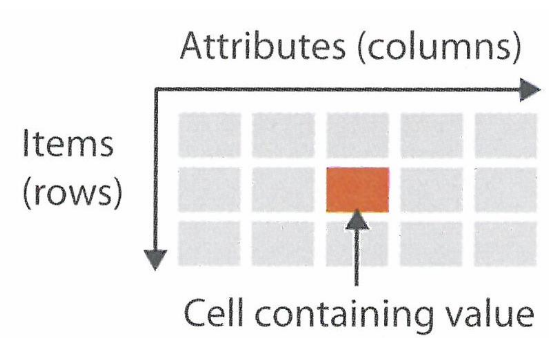
Shneiderman: The Eyes Have It

ANOTHER VIEW

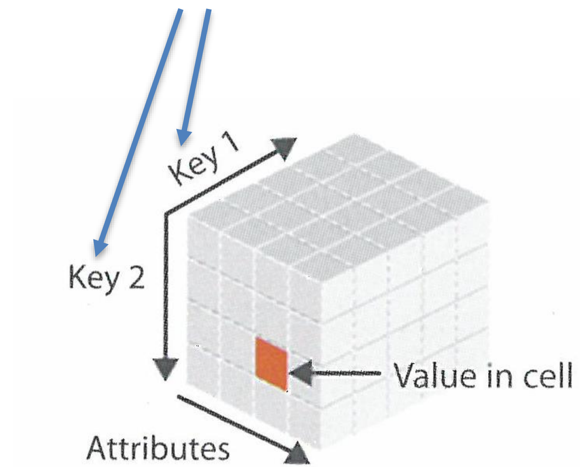




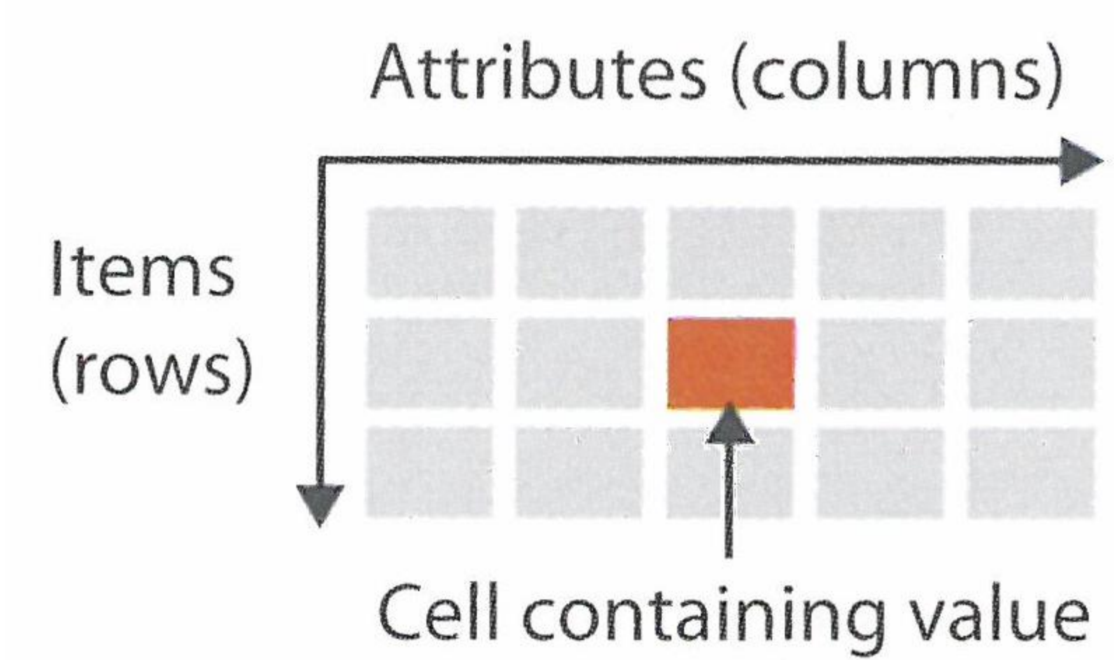
DATA TABLES - TERMINOLOGY



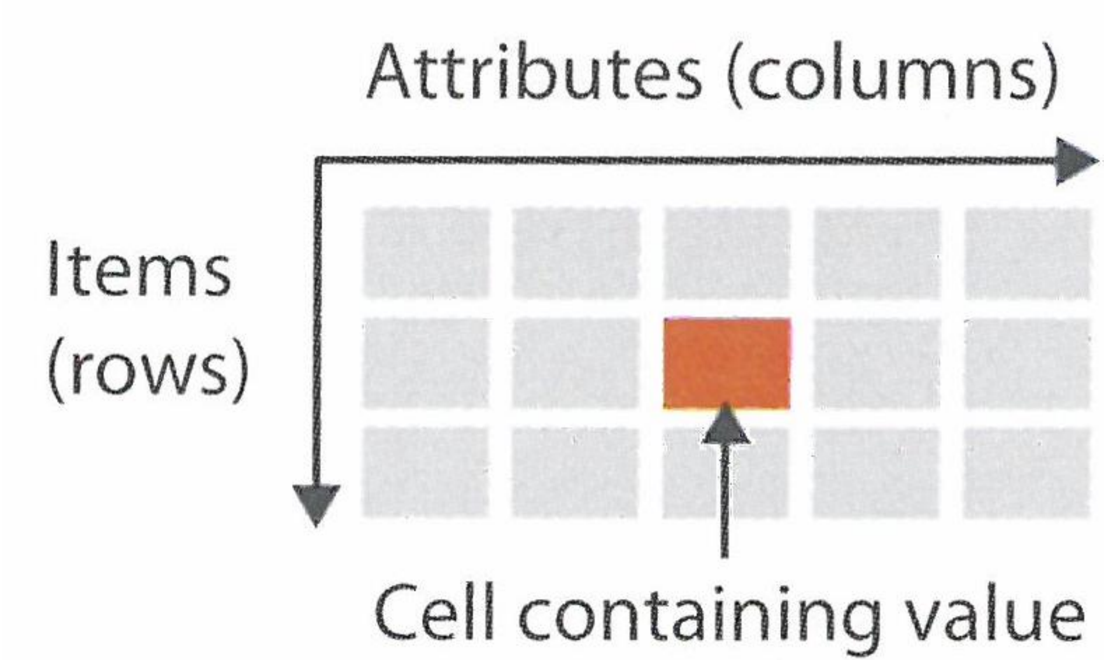
index to look up values



WHAT COULD BE THE KEY HERE?



WHAT DATA TYPE IS SUITABLE FOR A KEY?



KEYS VS. VALUES

key attributes are also sometimes called:

- independent attribute
- dimension

value attributes are also sometimes called:

- dependent attribute
- measure

LEVELS

= unique values for a categorical or ordered attribute

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

CONFERENCE:
InfoVis, Vis, SciVis, VAST

YEAR:
1990 – 2015

PAPER.TITLE:
>2500 different

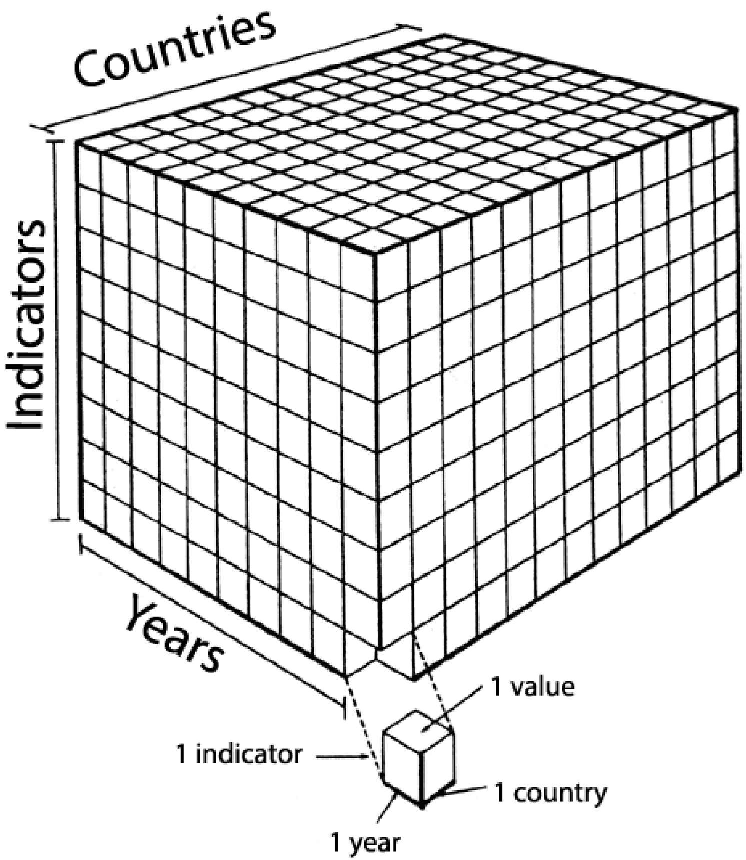
VISPUBDATA

ATTRIBUTES

ITEMS

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

THE DATA CUBE



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

MULTI-ATTRIBUTE DATA – OUR VIEW TODAY

n x d matrix

n attributes

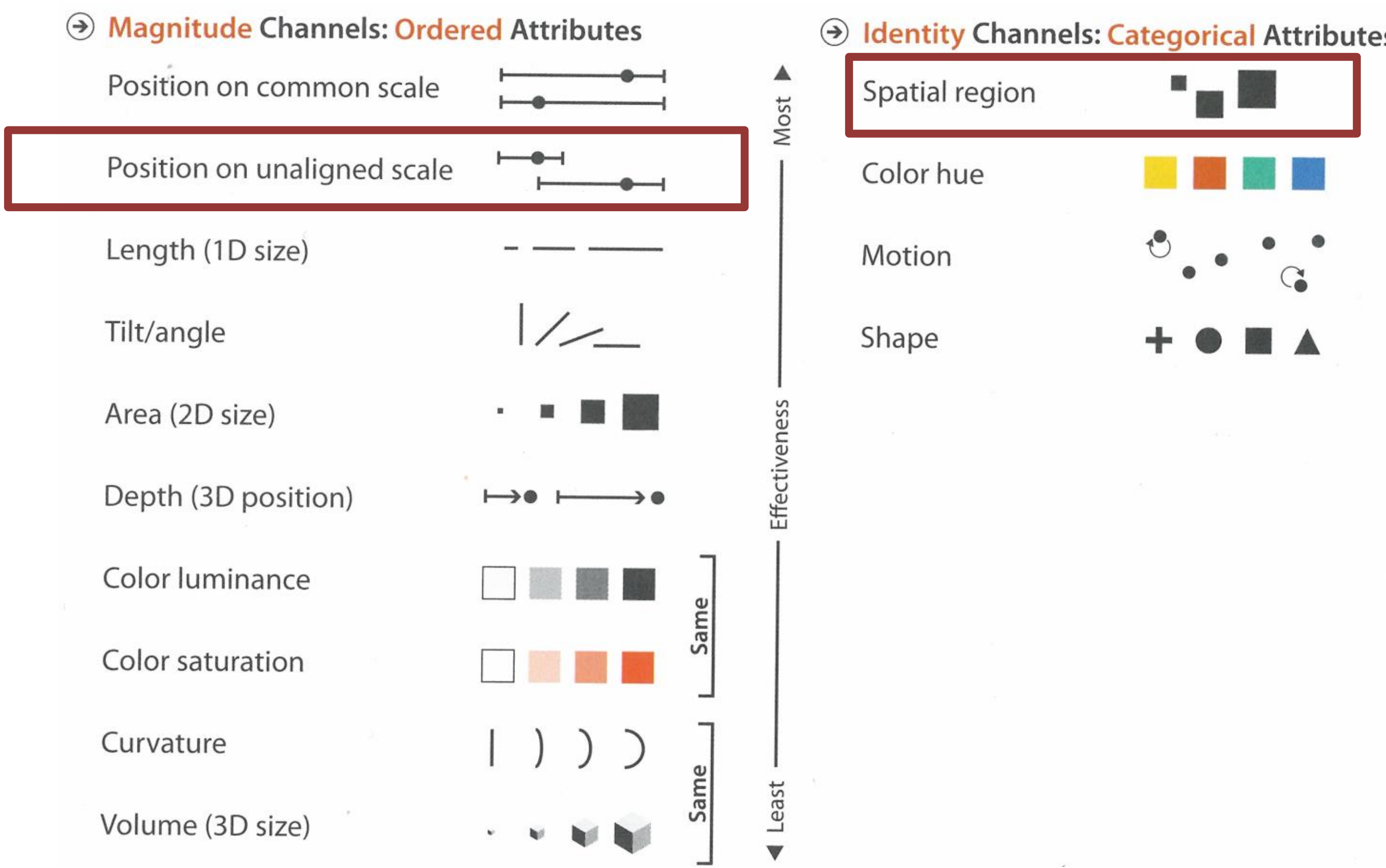
d items (data points)

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

ARRANGING TABULAR DATA

In Space

WHY ARRANGING DATA



QUANTITATIVE VALUES

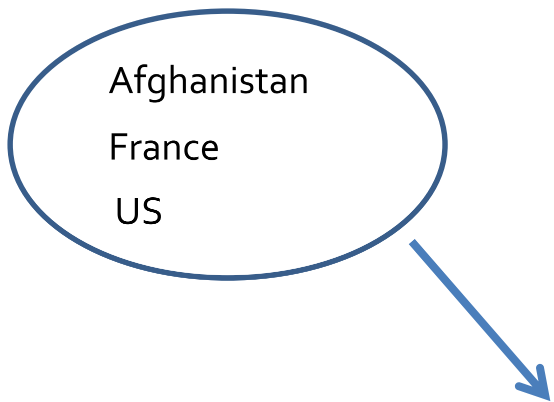
APPROACH

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

Country	Income per person
Afghanistan	850
France	29500
US	41000

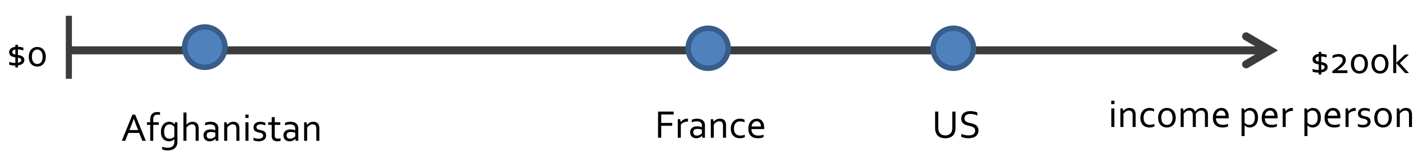
1. FIND A LAYOUT

Country	Income per person
Afghanistan	850
France	29500
US	41000



2. CHOOSE A VISUAL ENCODING & MARK

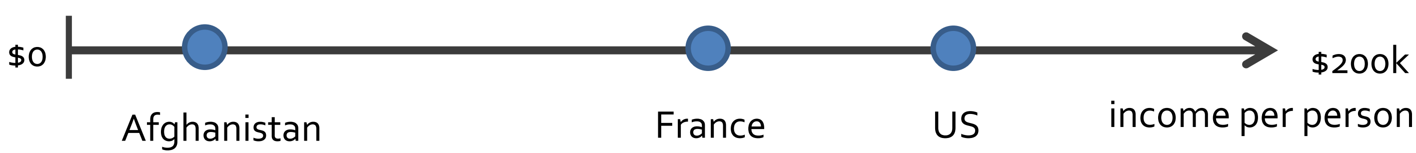
E.g. position + circle



1. FIND A LAYOUT

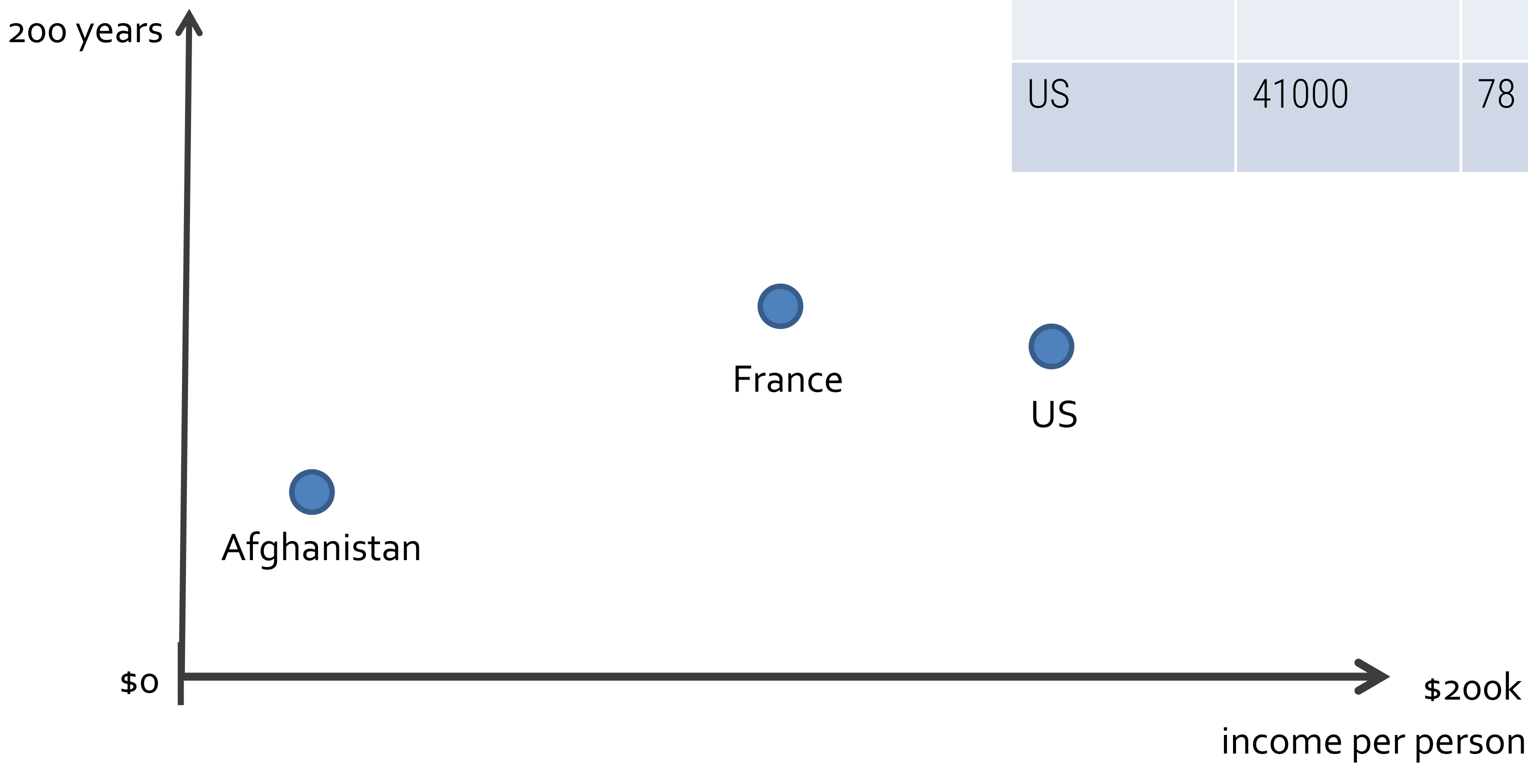
Country	Income per person	Life expectancy
Afghanistan	850	57
France	29500	81
US	41000	78

How do we extend this to 3 data attributes?



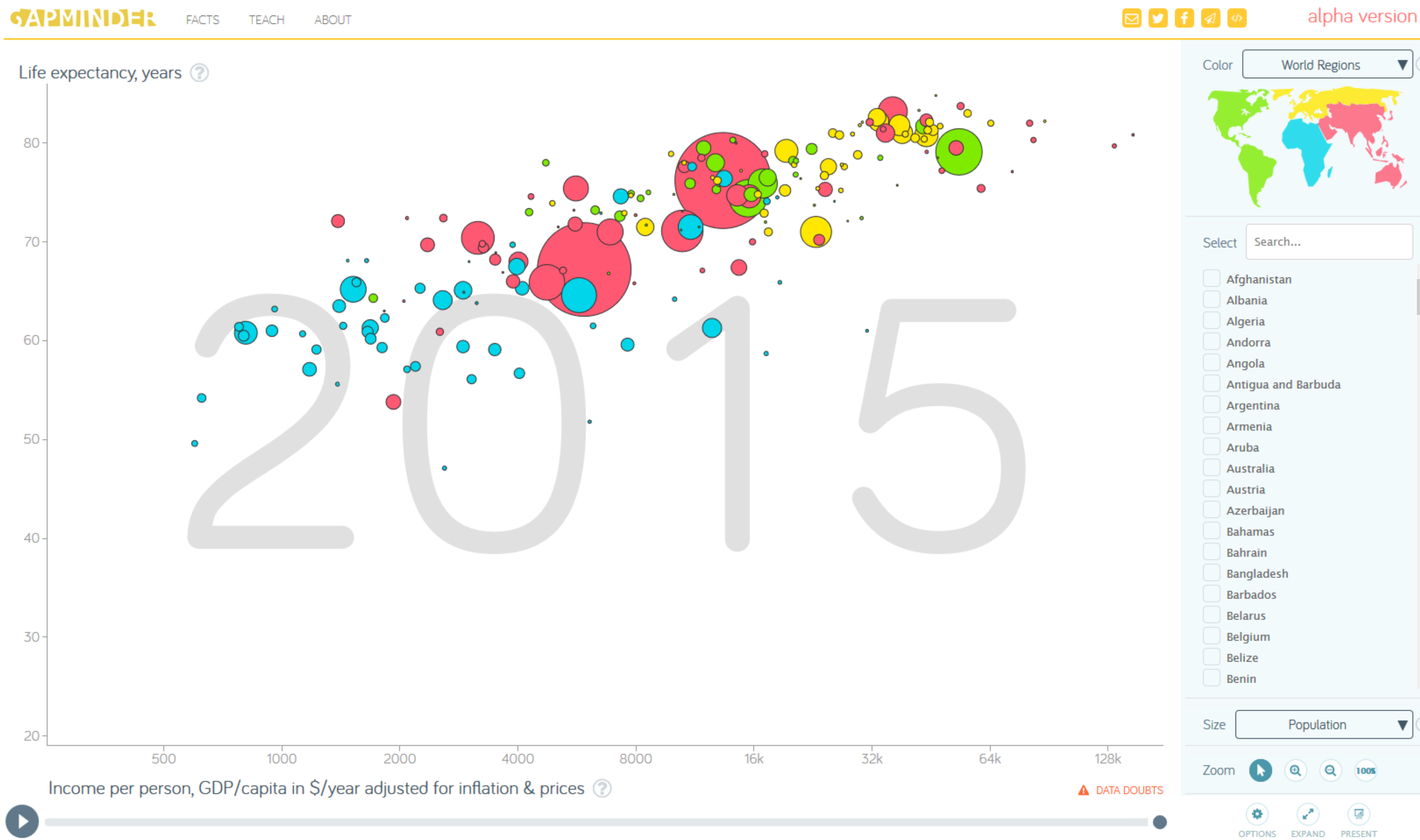
1. FIND A LAYOUT

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



SCATTERPLOTS

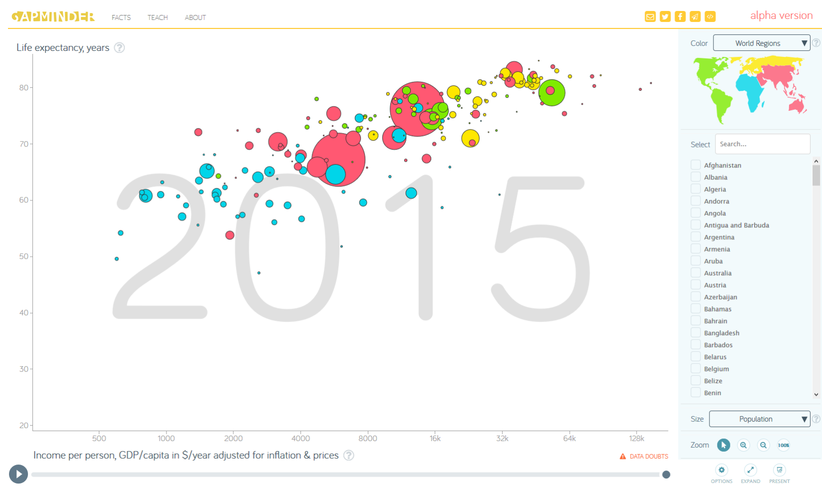
- two quantitative values
- horizontal and vertical spatial dimensions
- mark type = point



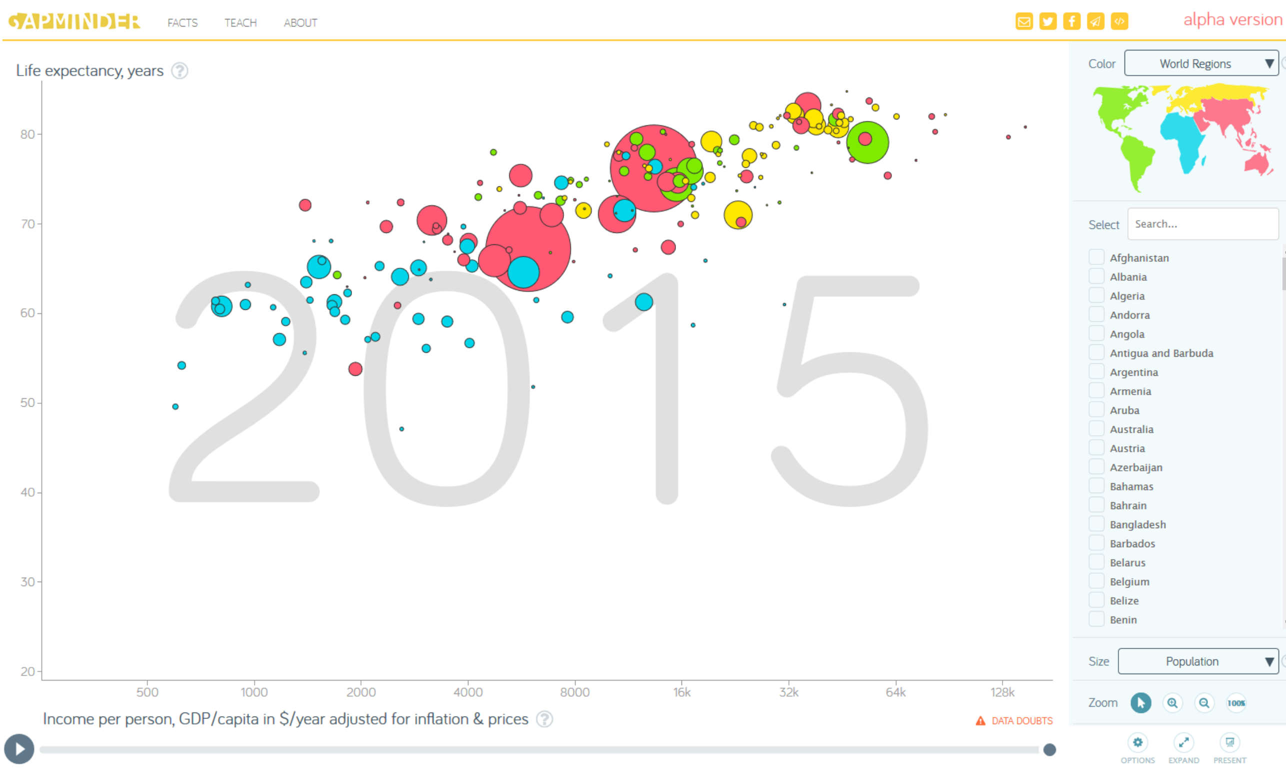
when marks are sized, the chart is often called a bubble chart or bubble plot

TASKS

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

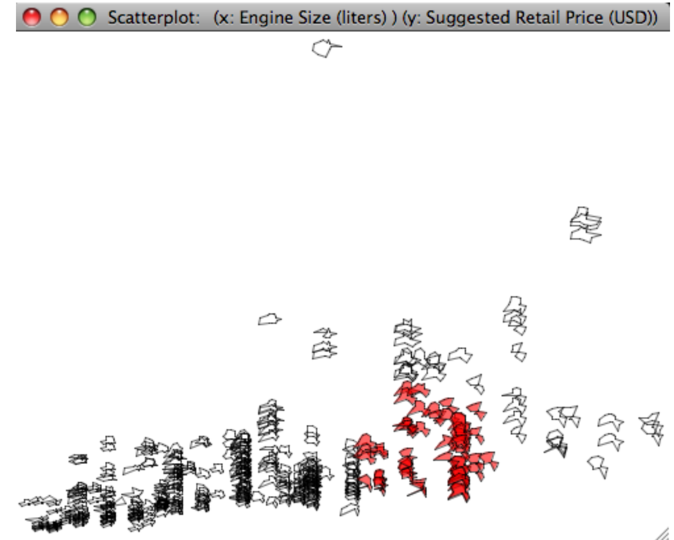


how many items are reasonable to put on a scatterplot?



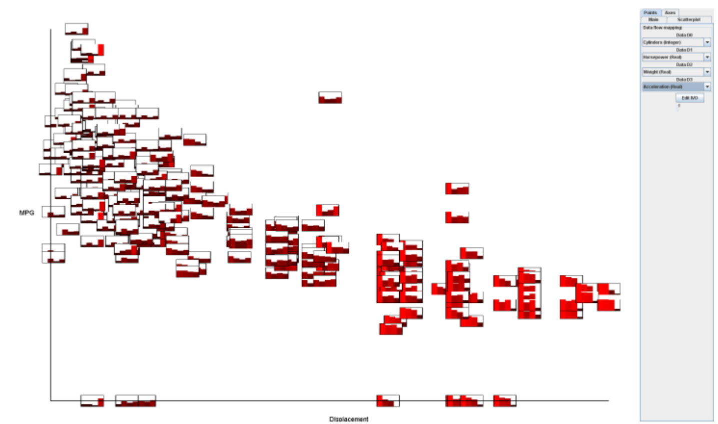
GLYPHS

marks can be replaced with glyphs



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

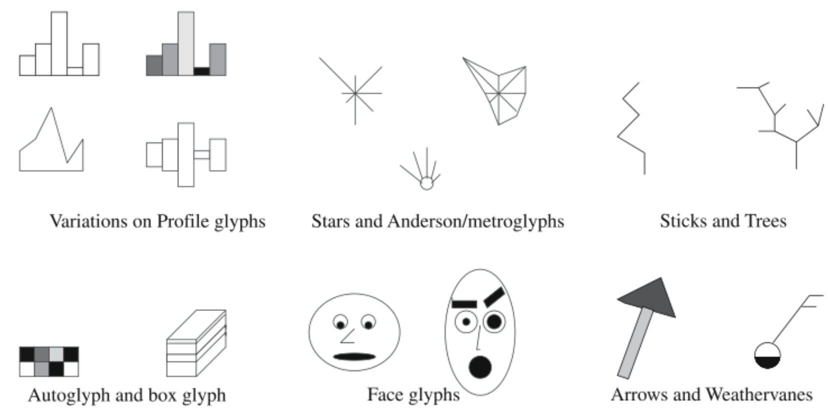
glyphs are themselves composed of multiple marks



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

GLYPHS

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



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

EXAMPLE: CHERNOFF FACES

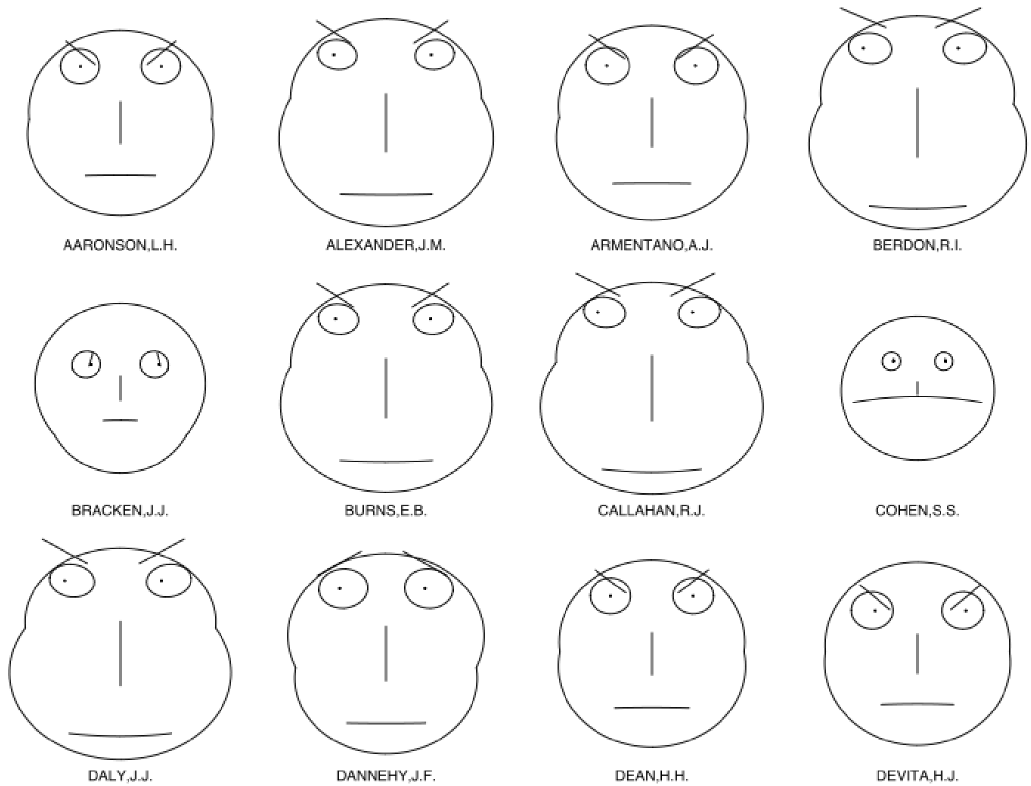
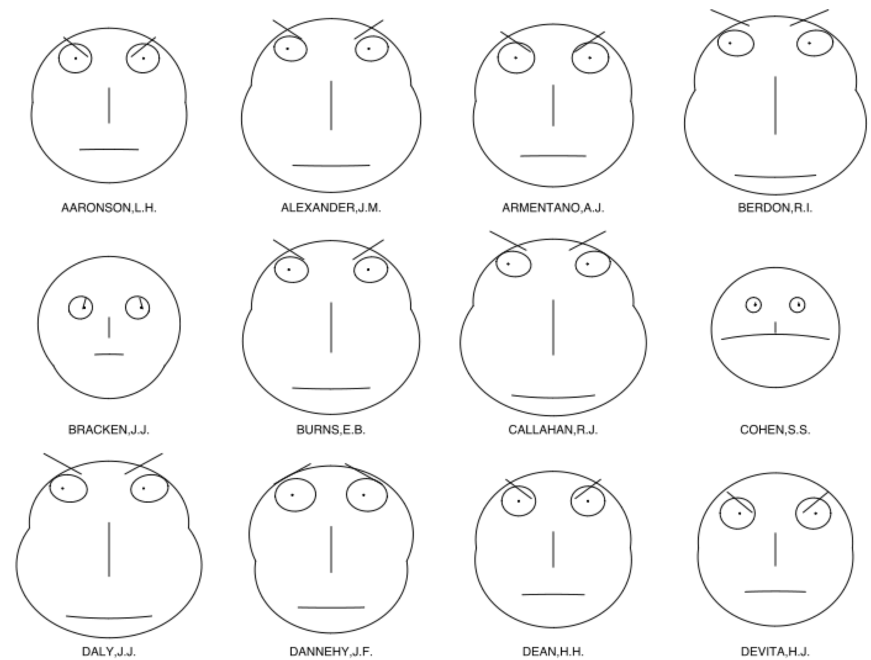


Image source: Wikipedia

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

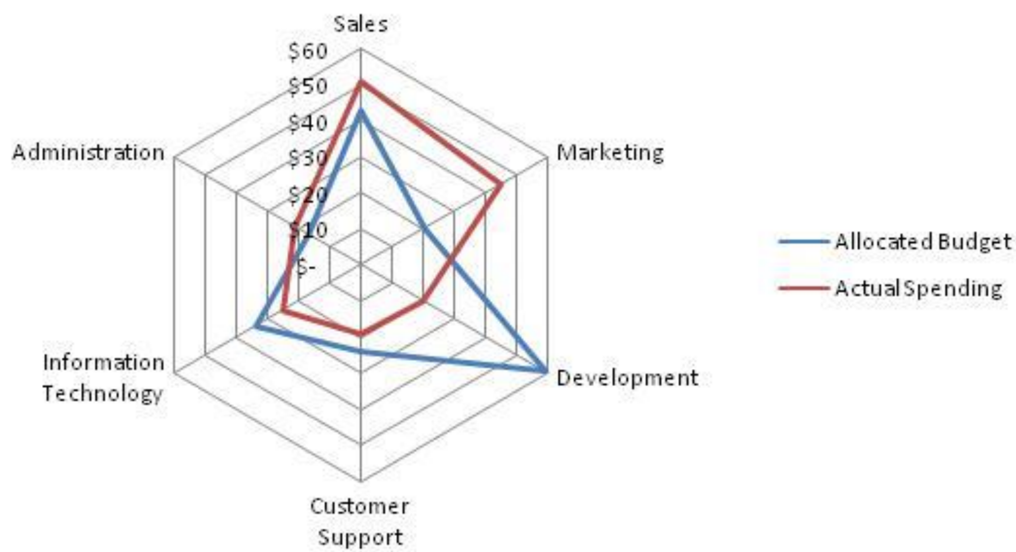
CHERNOFF FACES

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

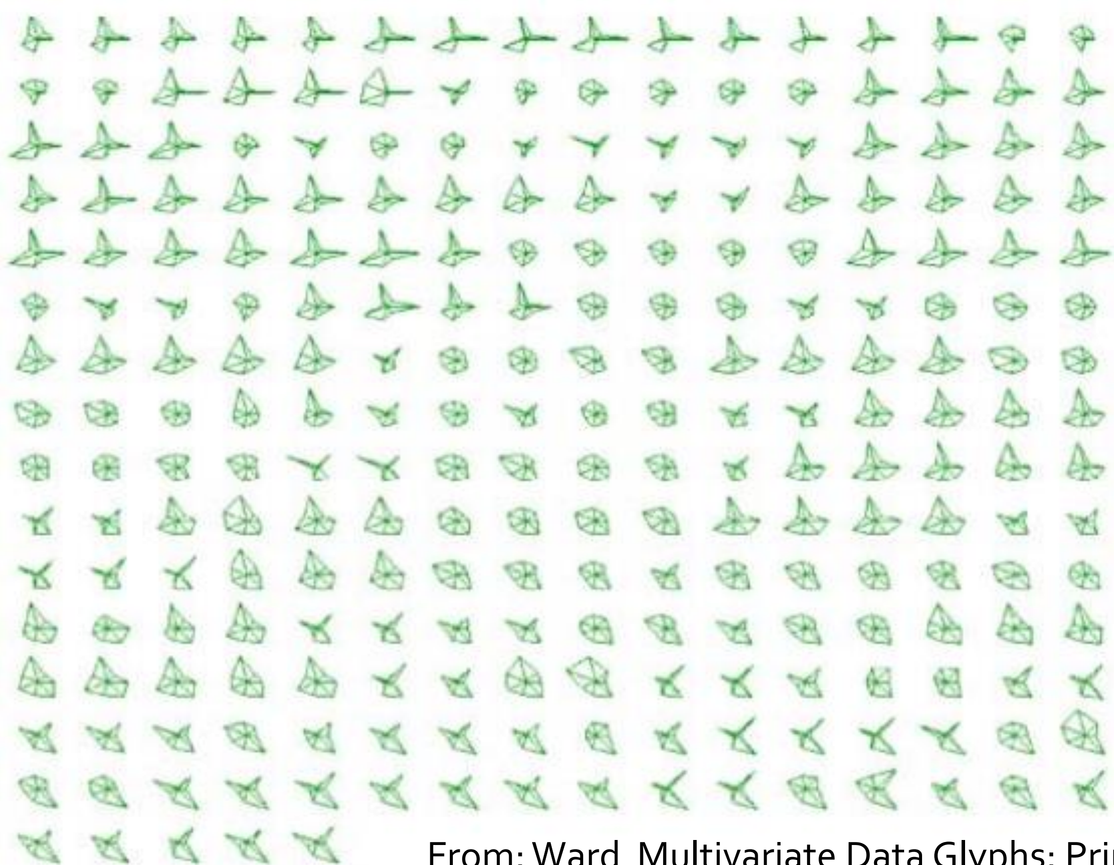


EXAMPLE: STAR GLYPHS

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



STAR GLYPHS



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

SHOW CATEGORICAL REGIONS

Separate, Order, and Align

CATEGORICAL VALUES

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

→cannot be encoded with spatial position

- BUT: can be expressed with a spatial region

REGIONS

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

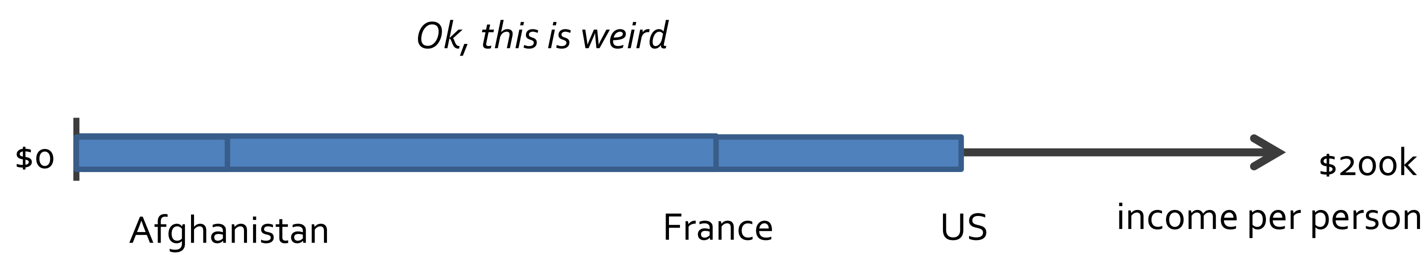
LIST ALIGNMENT

ONE KEY

LIST ALIGNMENT

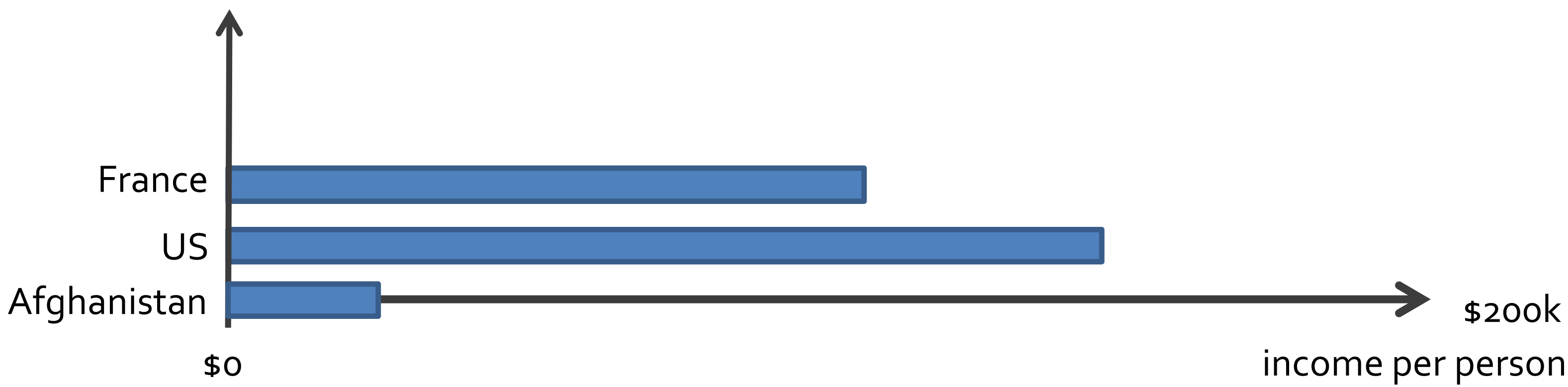
separate into regions by key

E.g. length + rectangle



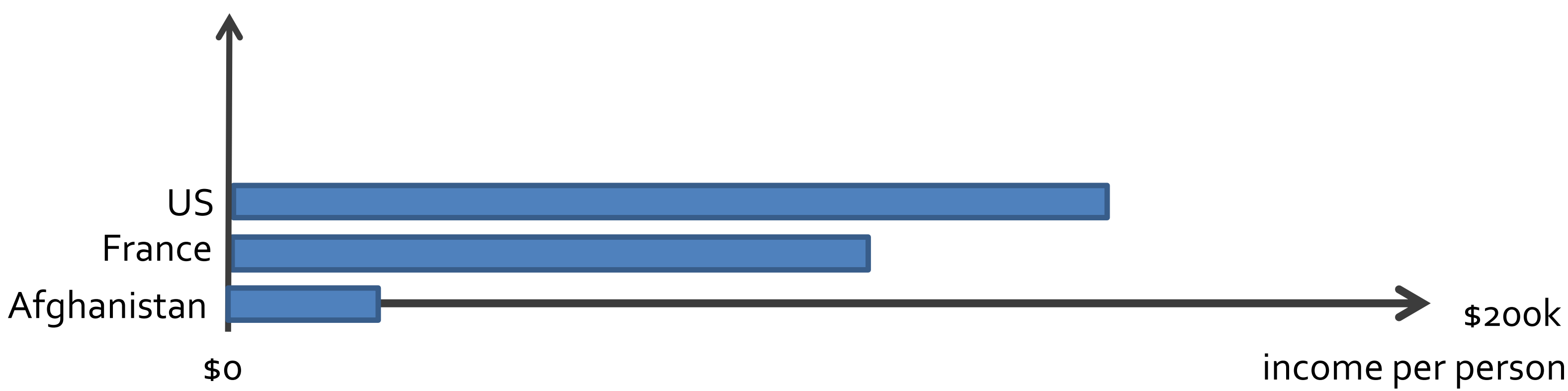
ALIGN

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



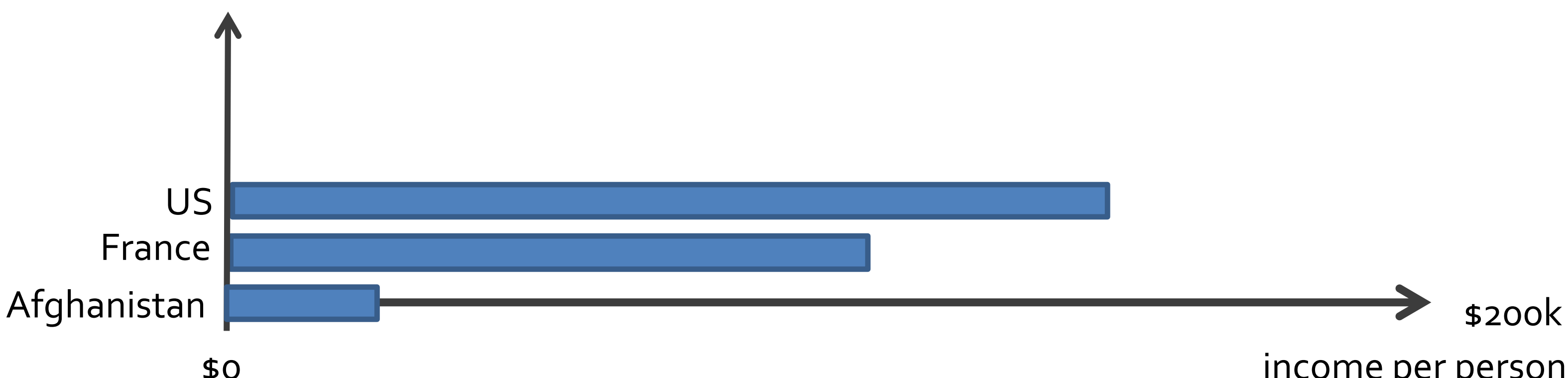
ORDER

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

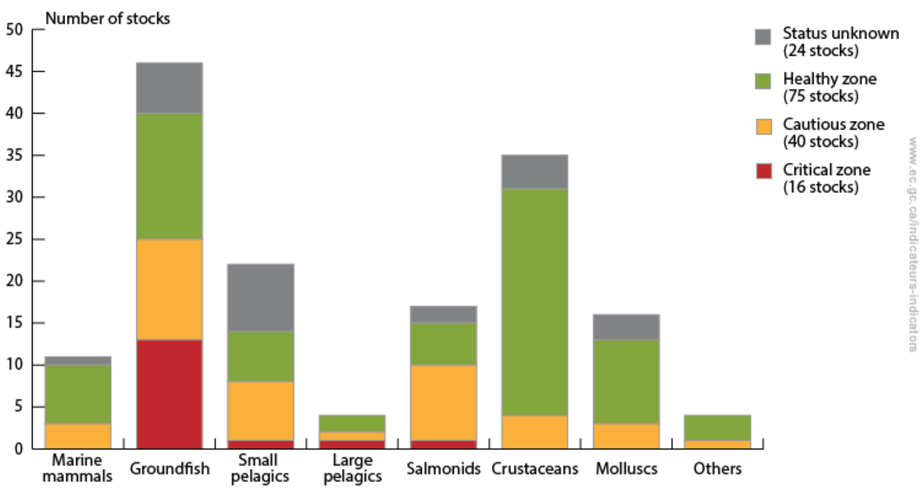


BAR CHARTS

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



ALTERNATIVE

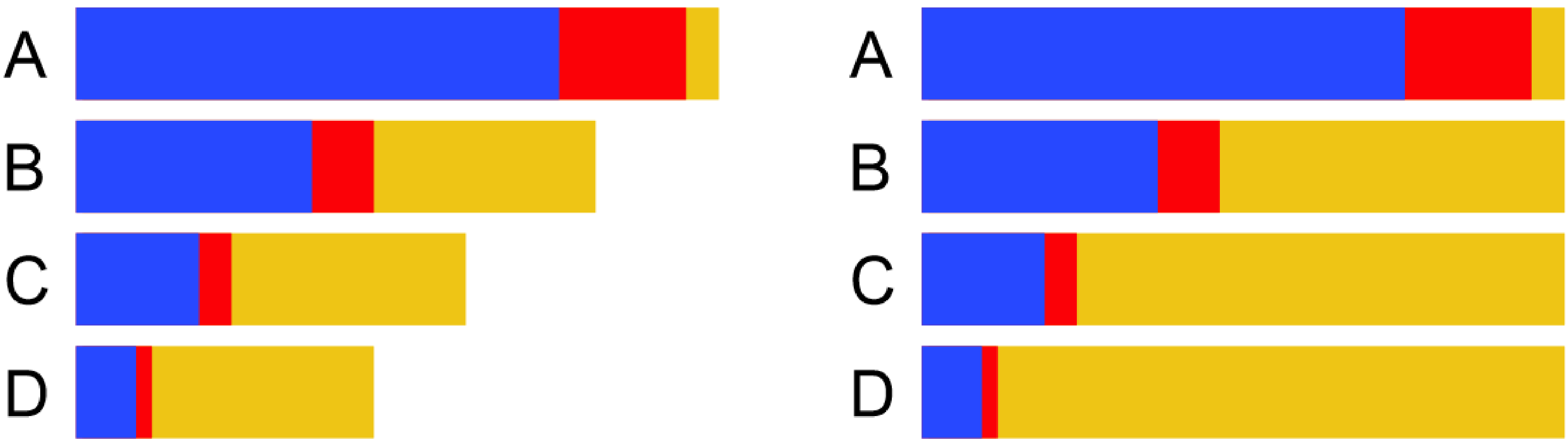


https://www.ec.gc.ca/indicateurs-indicators/default.asp?lang=en&n=1BCD421B-1

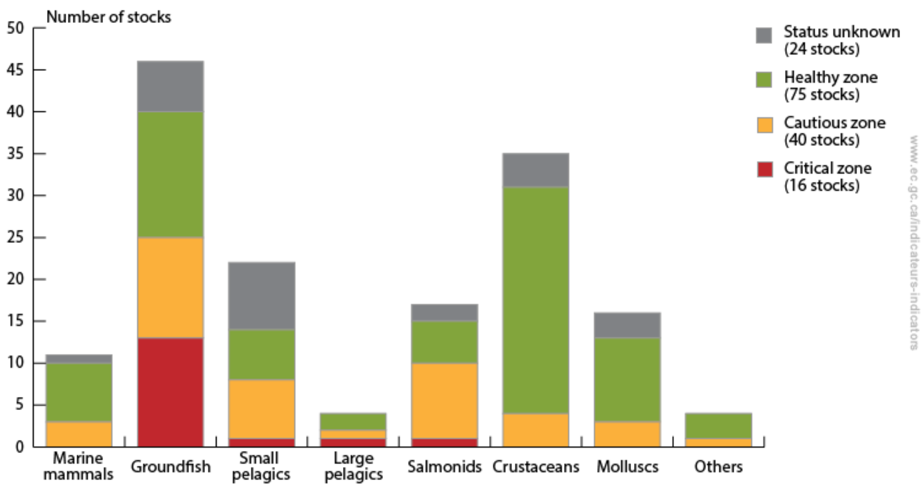
Stacked bar chart

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

STACKED BARS VS. NORMALIZED STACKED BARS



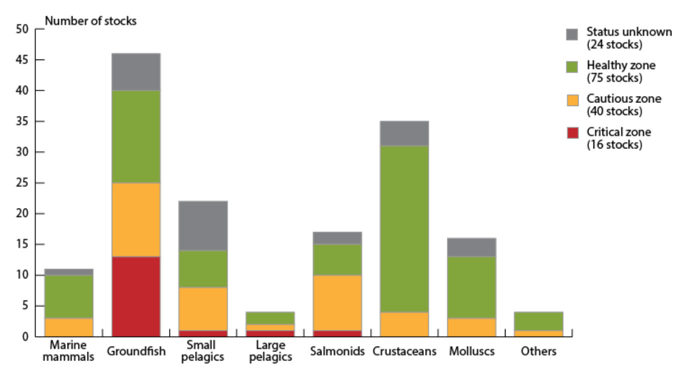
STACKED BARS



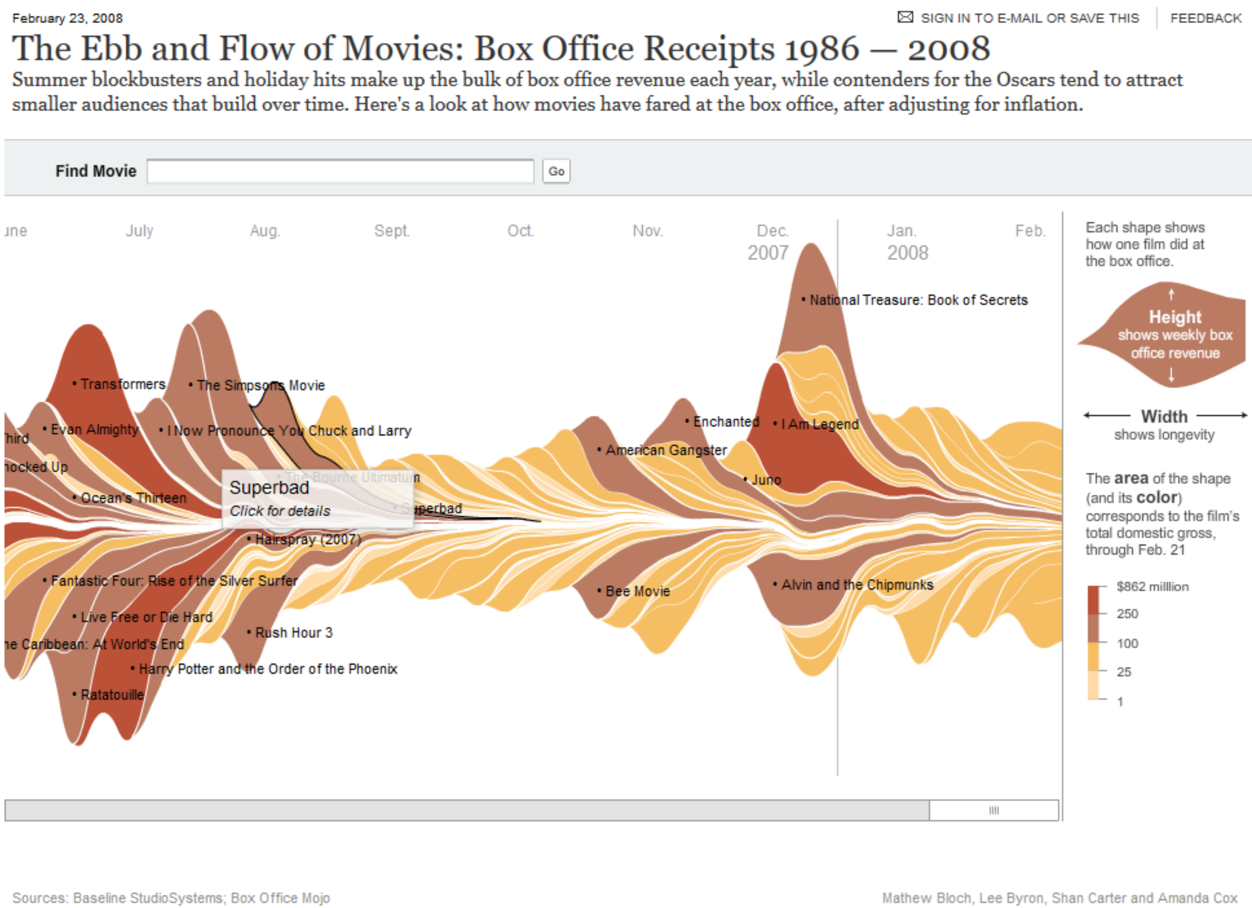
- ADVANTAGE
  - can compare totals and lowest level well
- DISADVANTAGE
  - upper levels of secondary key require comparison against non-aligned scale

STACKED BARS

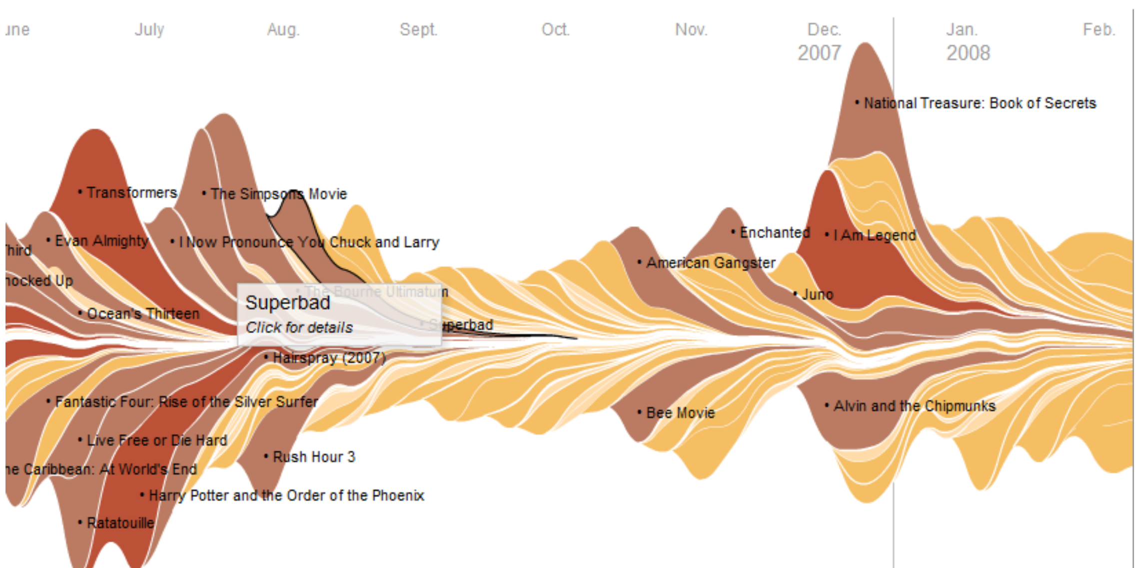
DATA	MD table; one quantitative value attribute, two categorical key attributes
ENCODE	bar glyph: length-encoded subcomponents for each level of secondary key attribute separate bars by category of primary key
TASK	part-to-whole relationship, lookup values, find trends
SCALE	key attribute (main axis):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

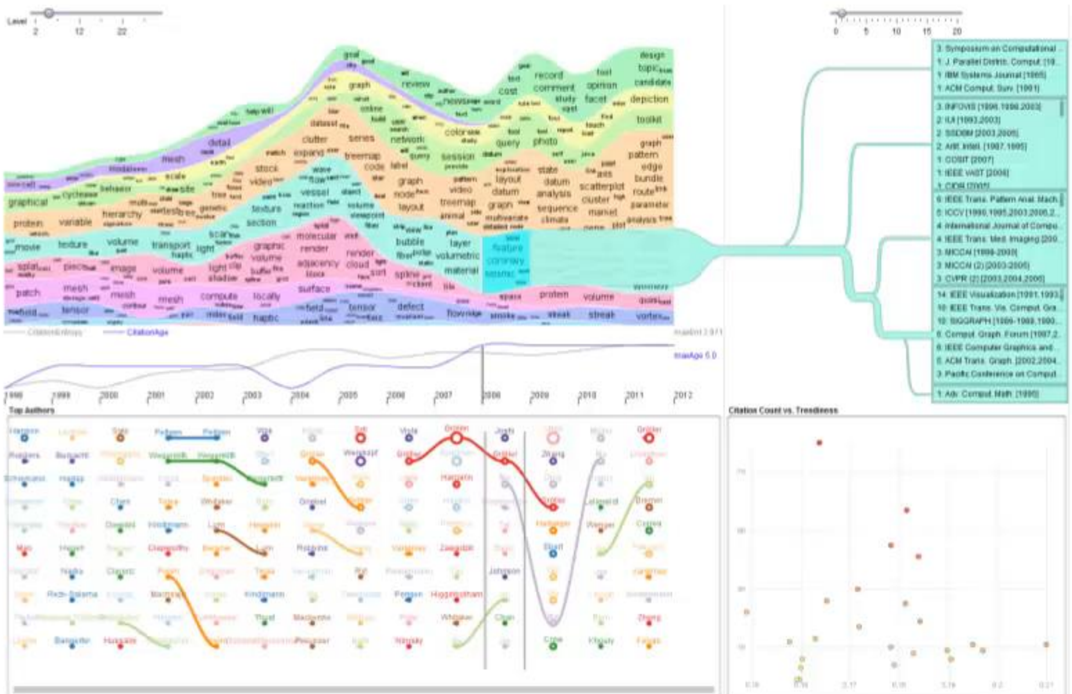


STREAMGRAPH



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





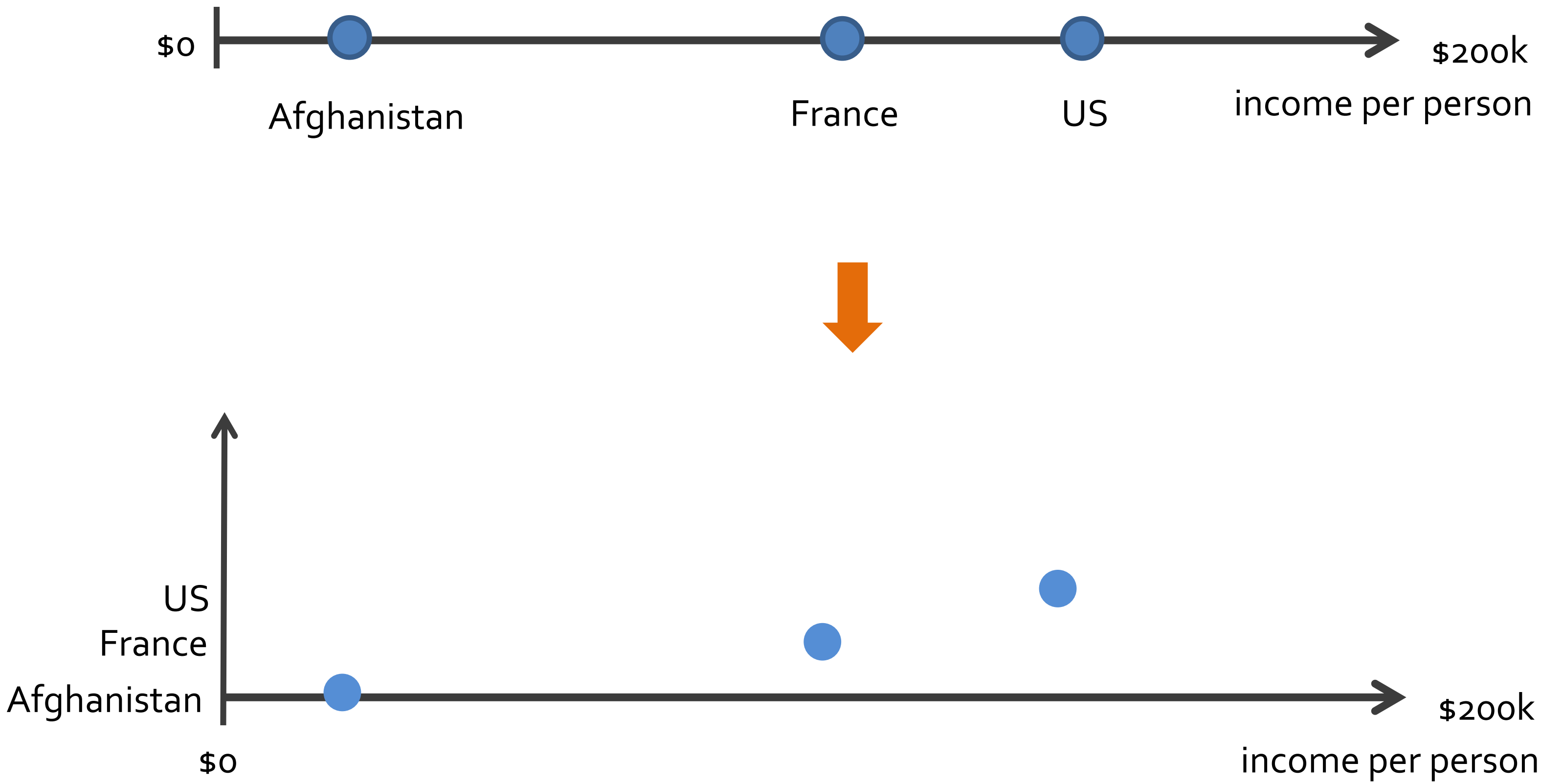
CiteRivers

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

IEEE VAST 2015

Institute for Visualization and Interactive Systems

DOT CHART/PLOT

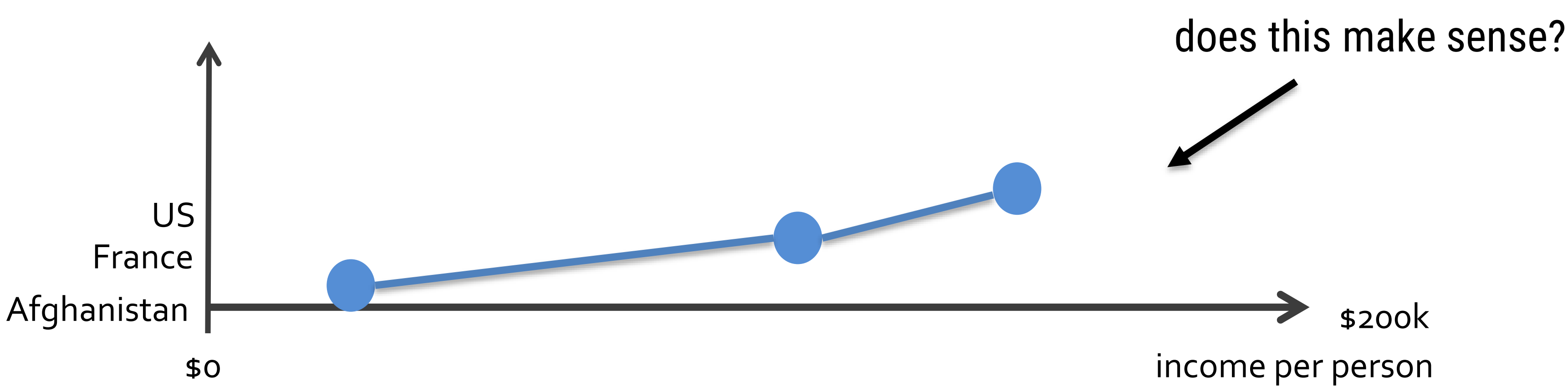


LINE CHART

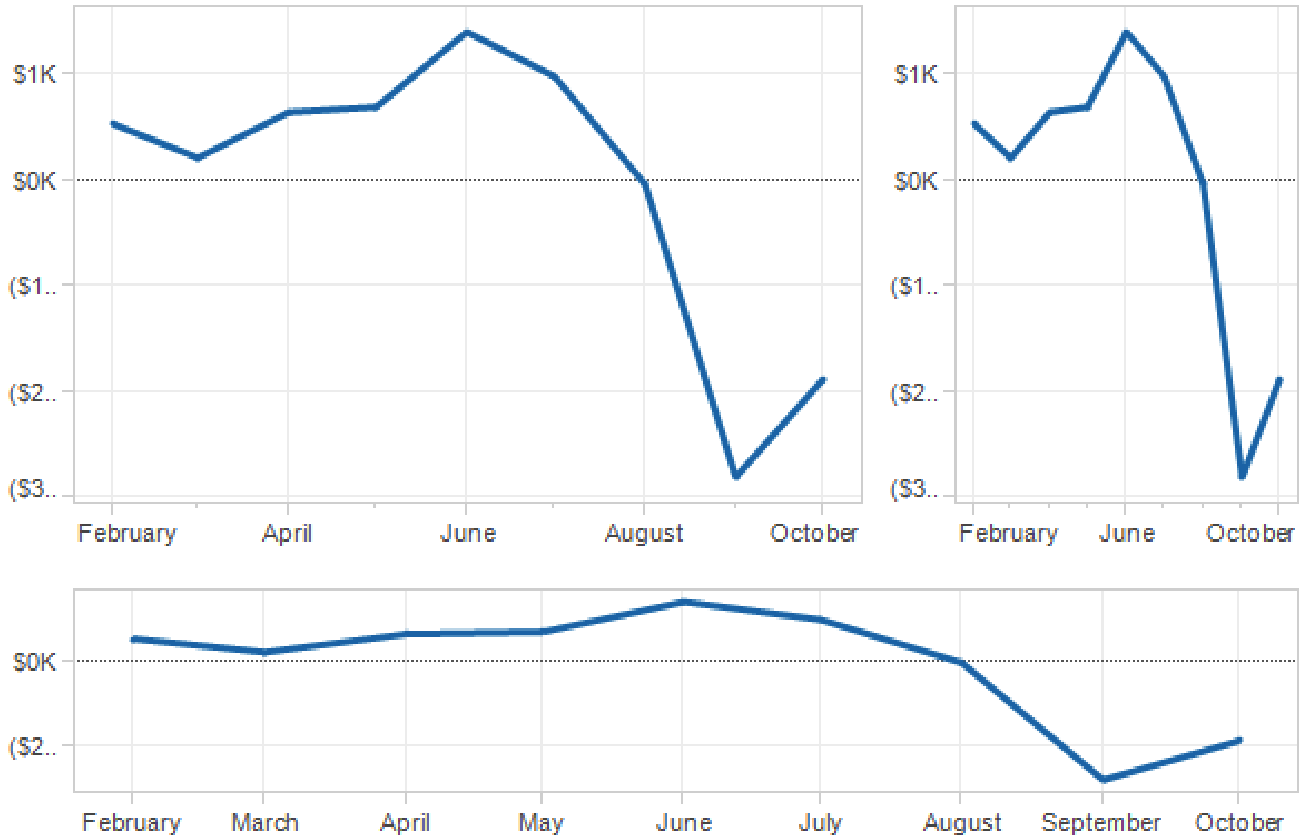
augment with line connection marks

emphasize the ordering and show trends

should not be used with categorical keys



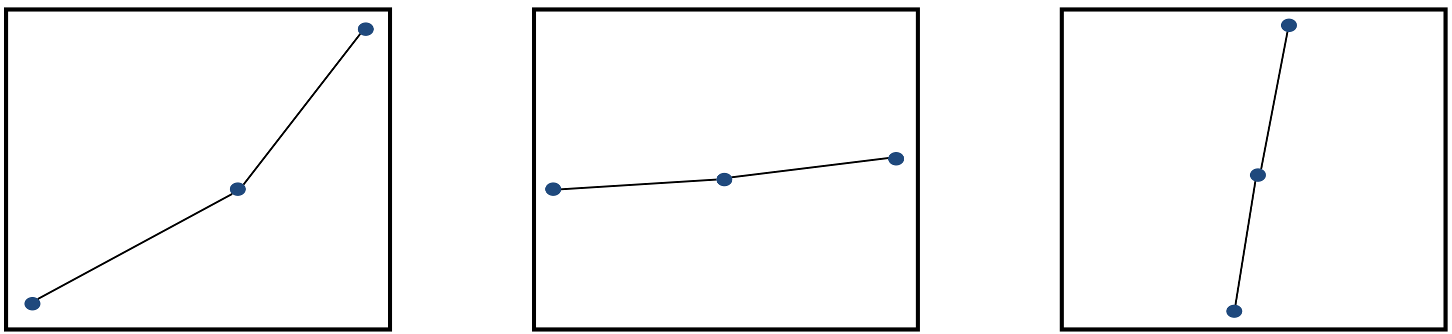
ASPECT RATIO SELECTION



BANKING TO 45°

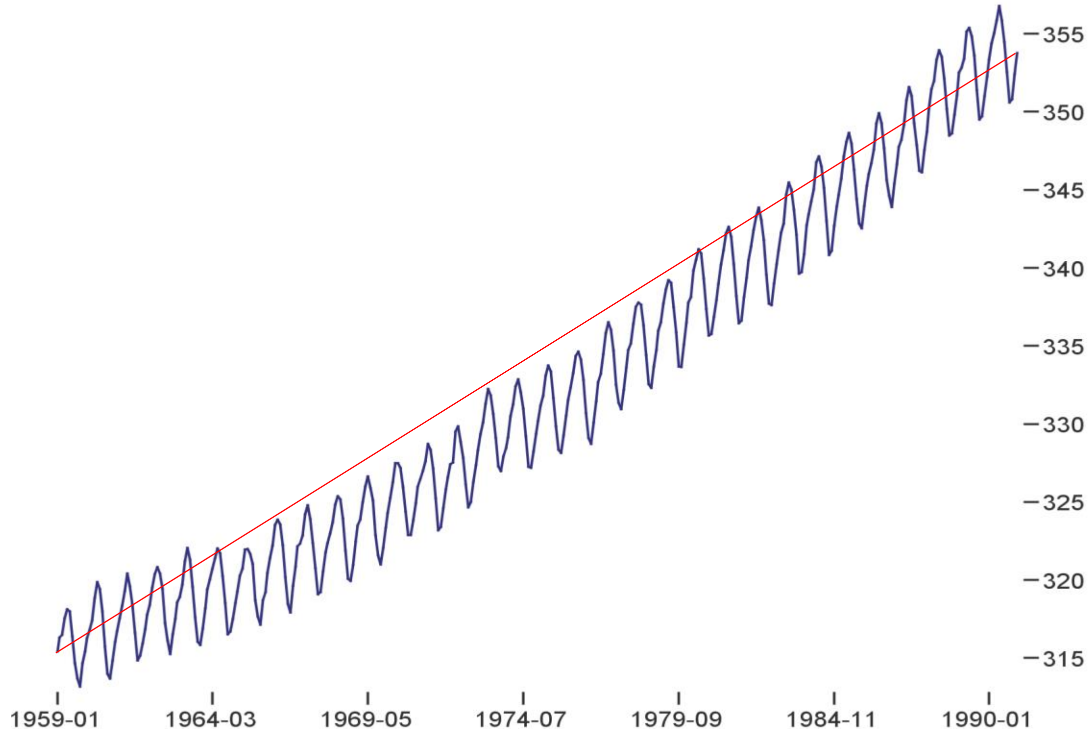
[Cleveland]

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

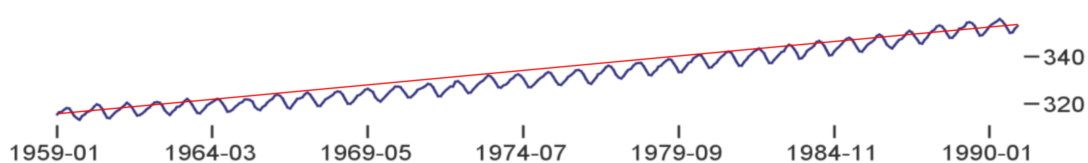


TWO SEGMENTS ARE MAXIMALLY DISCRIMINABLE WHEN THEIR AVG ABSOLUTE ANGLE IS 45°

OPTIMIZE THE *ASPECT RATIO* TO BANK TO 45°

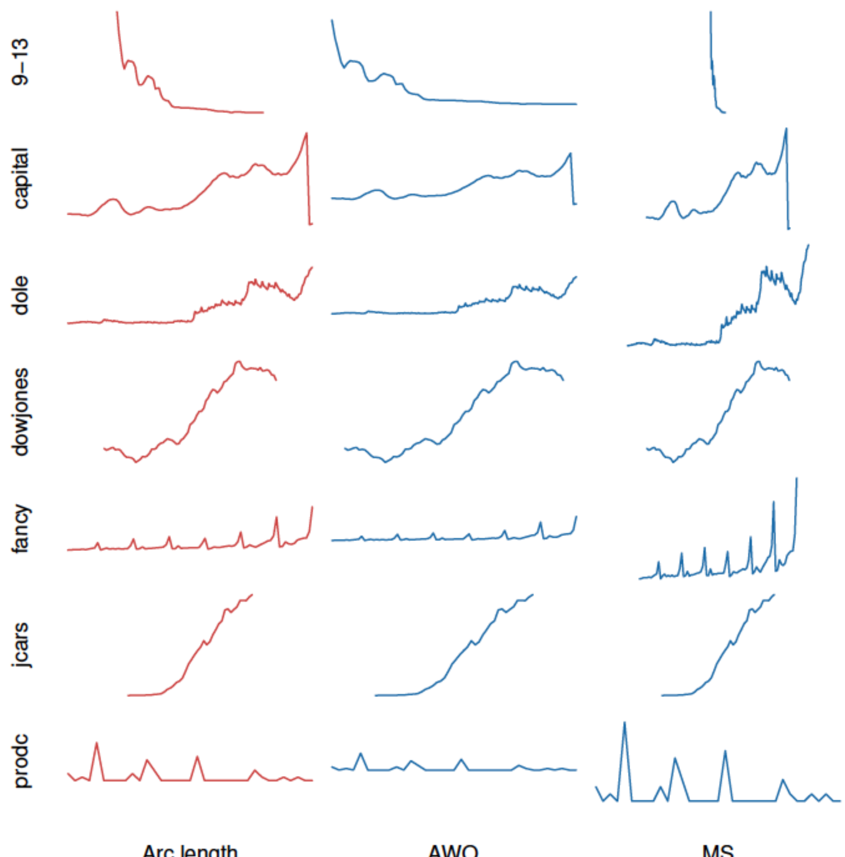


ASPECT RATIO = 1.17



ASPECT RATIO = 7.87

ALTERNATIVE METHODS



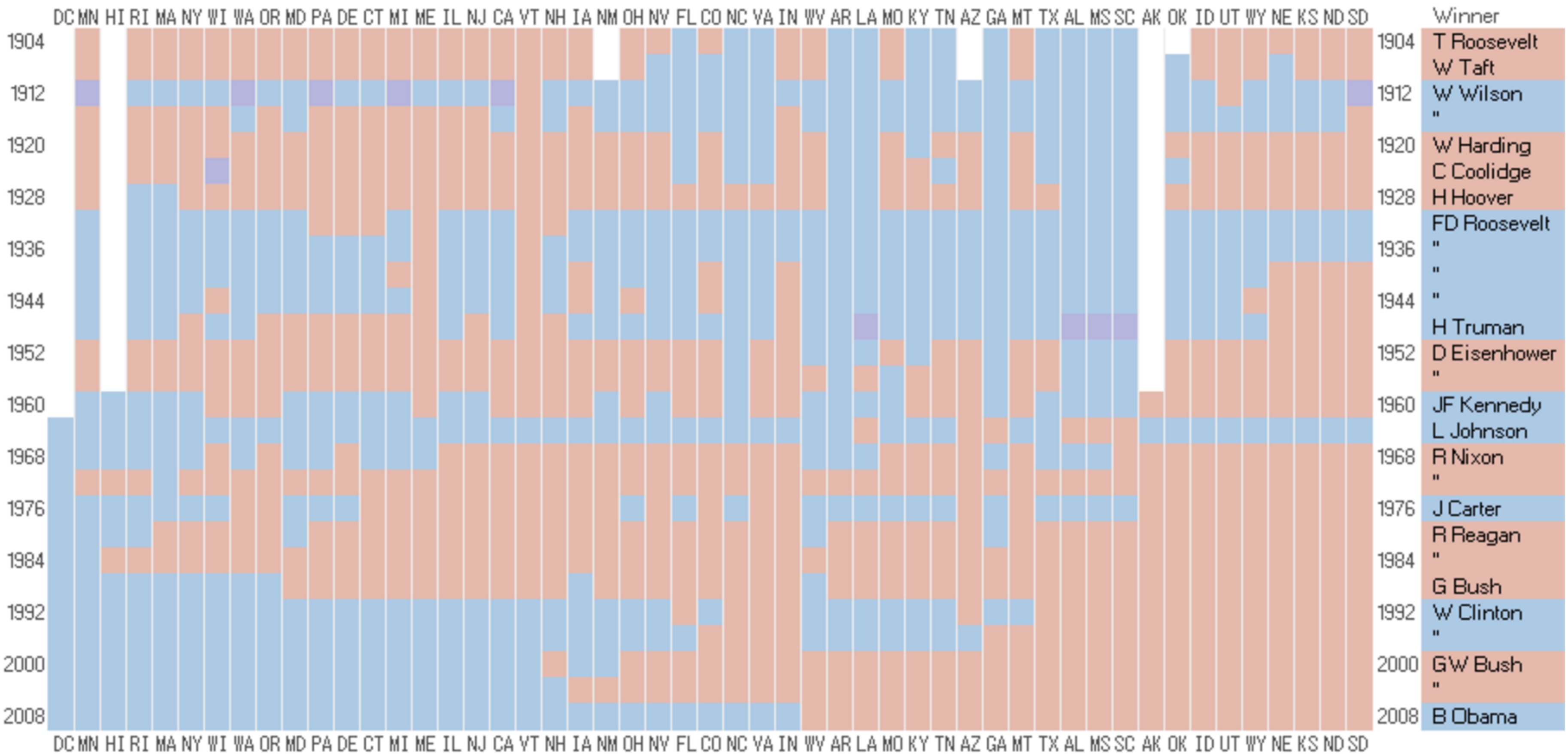
Practical advice:

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

[TALBOT ET AL, 2011]

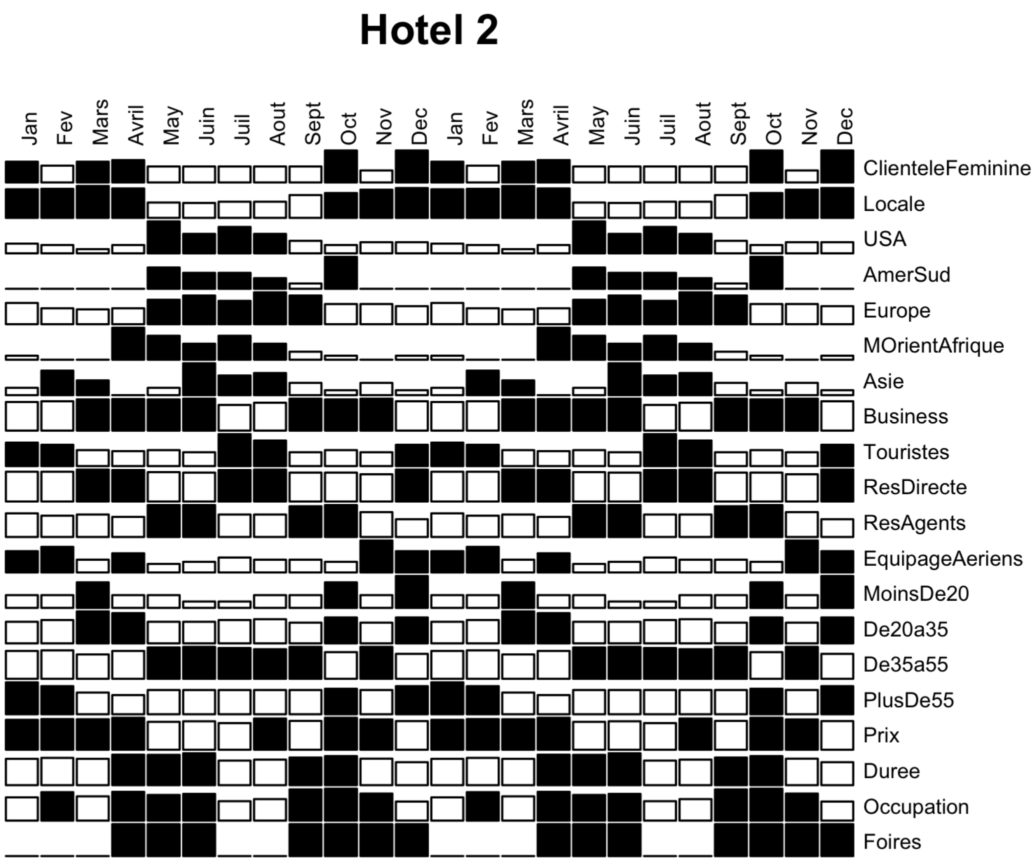
MATRIX ALIGNMENT

Two keys



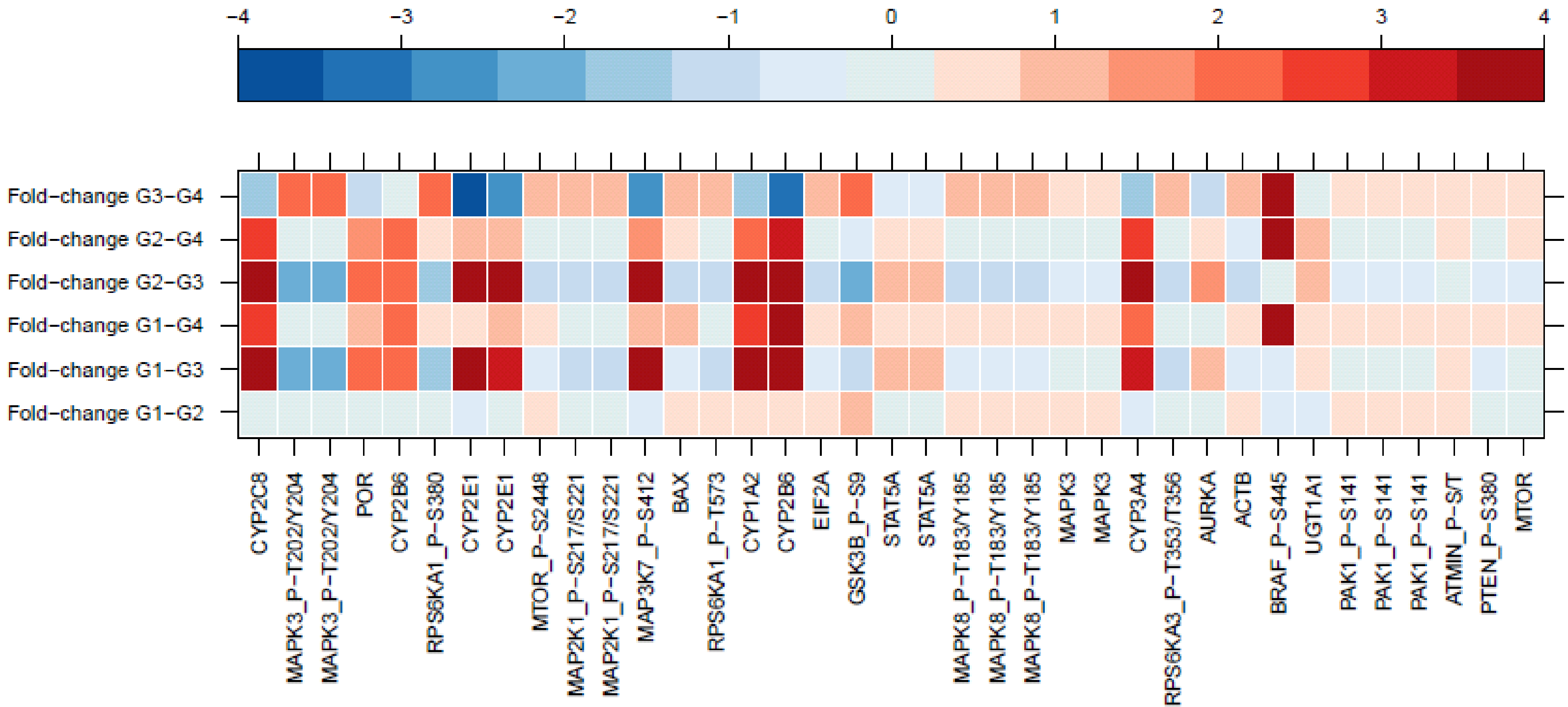
https://ldld.samizdat.cc/2016/tag/catalog/

HEATMAP



Hotel 2
Jan
Fev
Mars
Avril
May
Juin
Juil
Aout
Sept
Oct
Nov
Dec
Jan
Fev
Mars
Avril
May
Juin
Juil
Aout
Sept
Oct
Nov
Dec
ClienteleFeminine
Locale
USA
AmerSud
Europe
MOrientAfrique
Asie
Business
Touristes
ResDirecte
ResAgents
EquipageAeriens
MoinsDe20
De20a35
De35a55
PlusDe55
Prix
Duree
Occupation
Foires

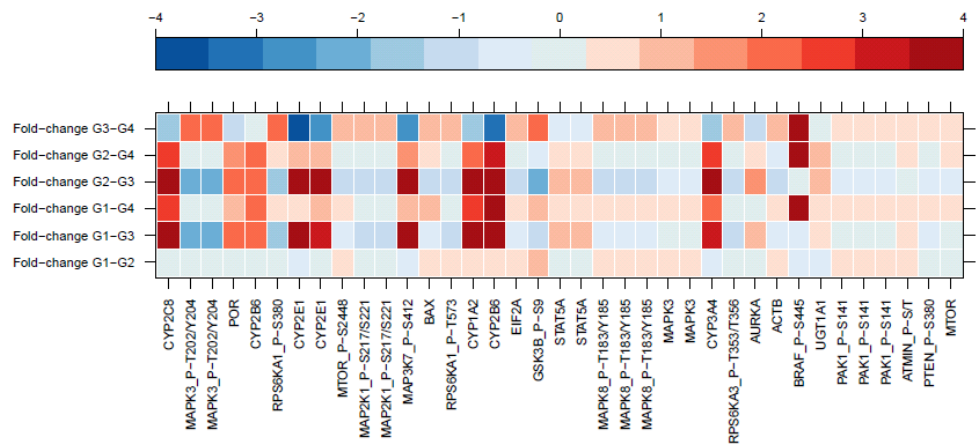
HEATMAP



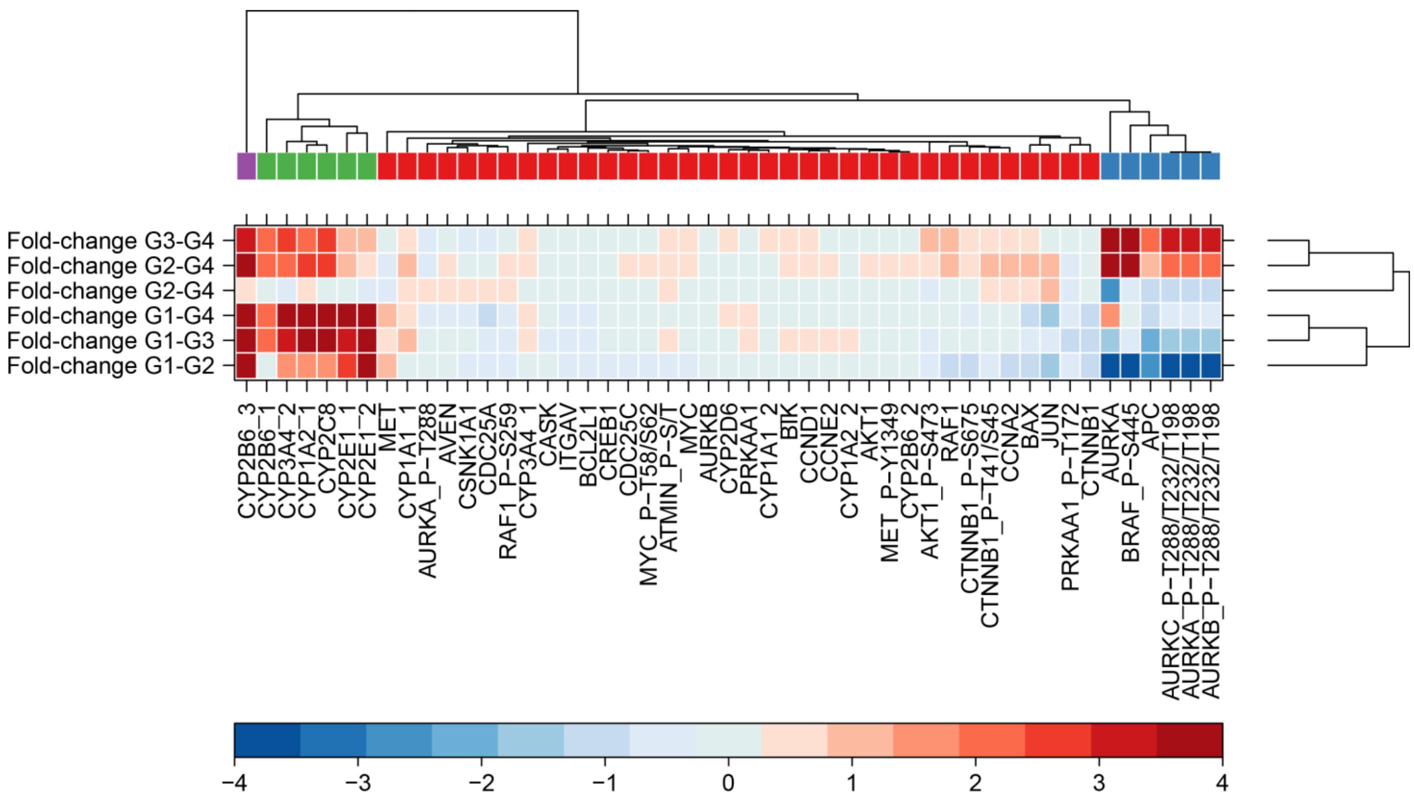
http://www.ra.cs.uni-tuebingen.de/software/RPPApipe/doc/documentation.htm

HEATMAP

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



CLUSTERED HEATMAP

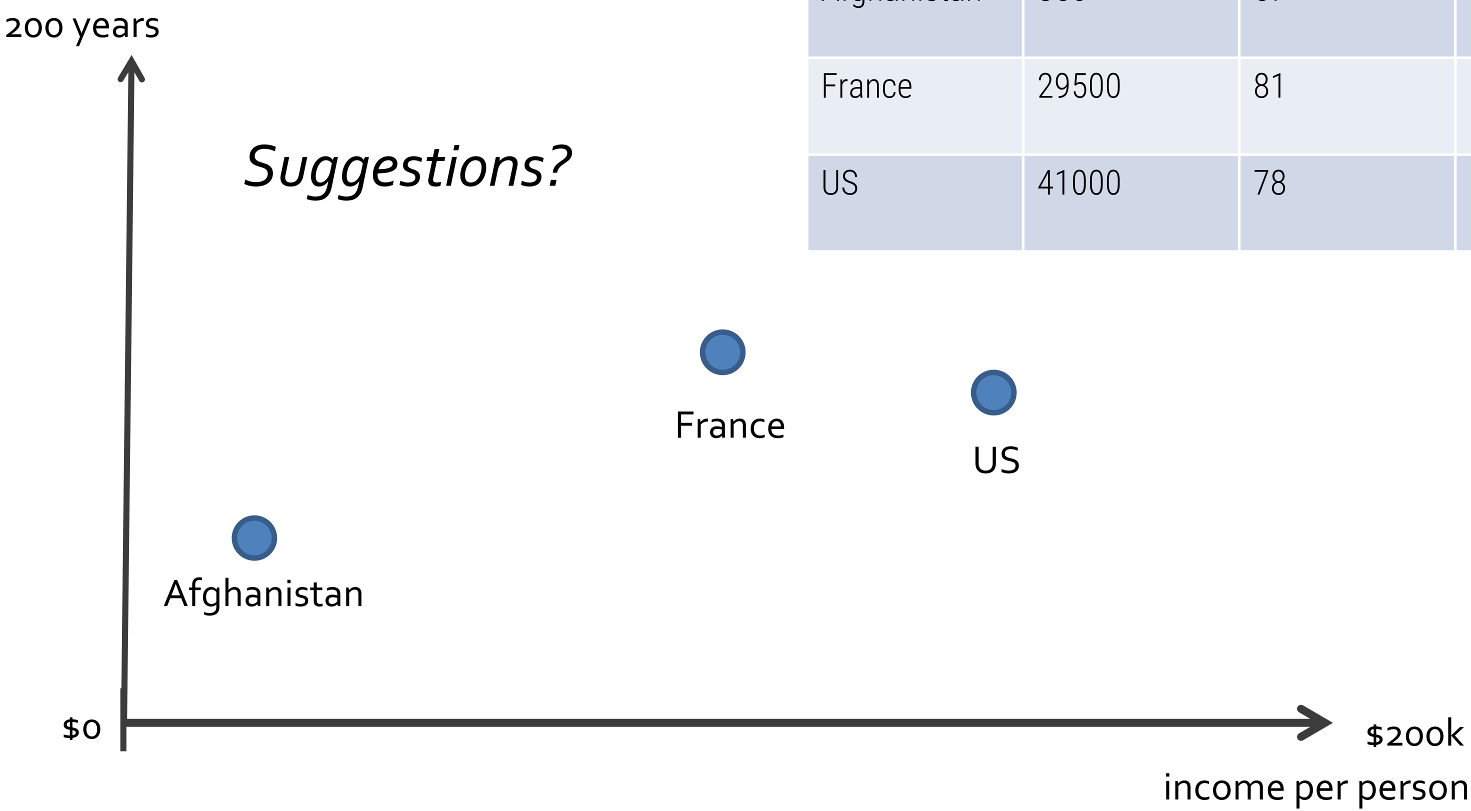


http://www.ra.cs.uni-tuebingen.de/software/RPPApipe/doc/documentation.htm

BACK TO OUR ORIGINAL EXAMPLE

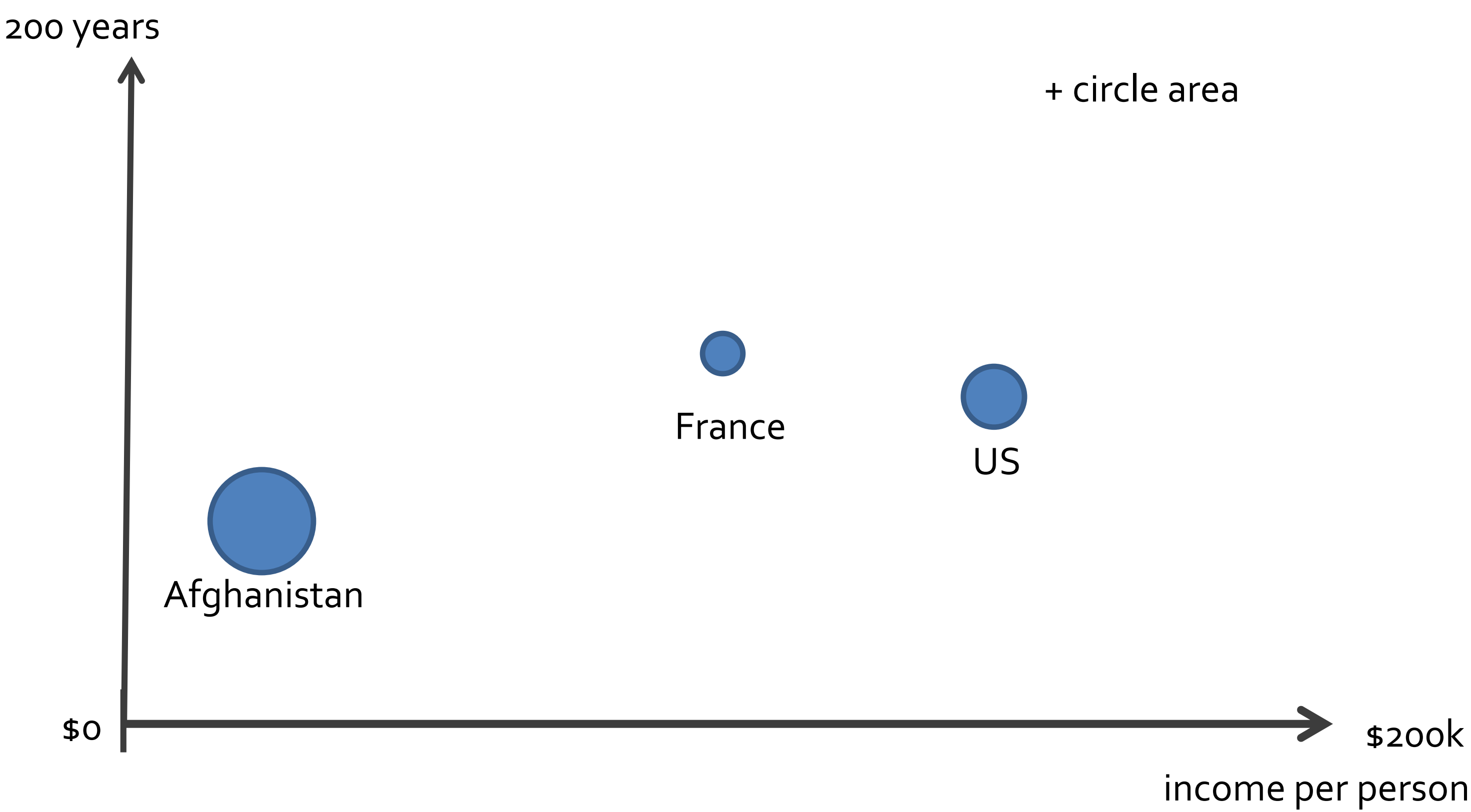
Country	Income per person	Life expectancy	Children per woman
Afghanistan	850	57	7.1
France	29500	81	1.9
US	41000	78	2.1

now with 4 attributes



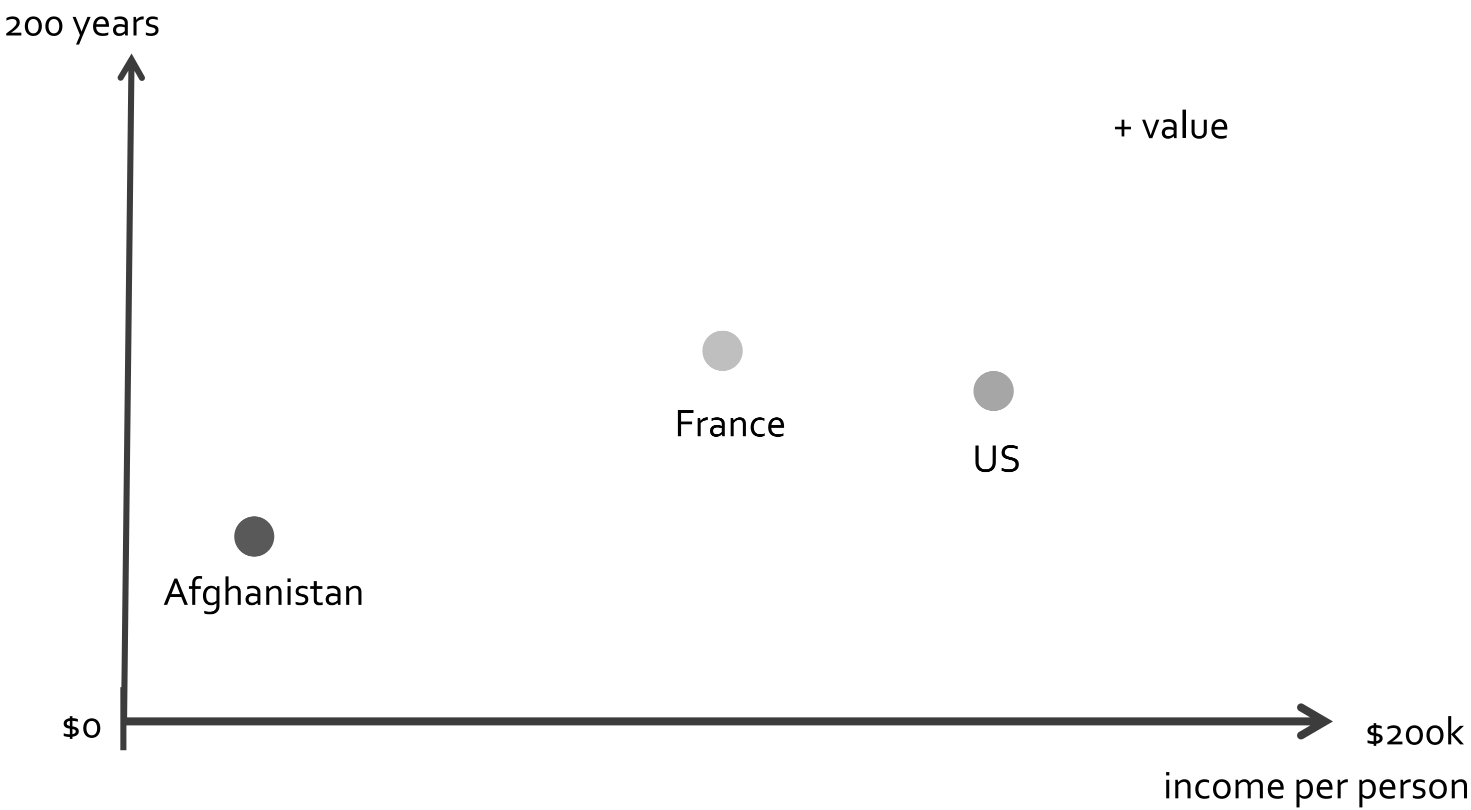
Country	Income per person	Life expectancy	Children per woman
Afghanistan	850	57	7.1
France	29500	81	1.9
US	41000	78	2.1

ADD ANOTHER VISUAL ENCODING



Problem:
Does not scale well to more attributes

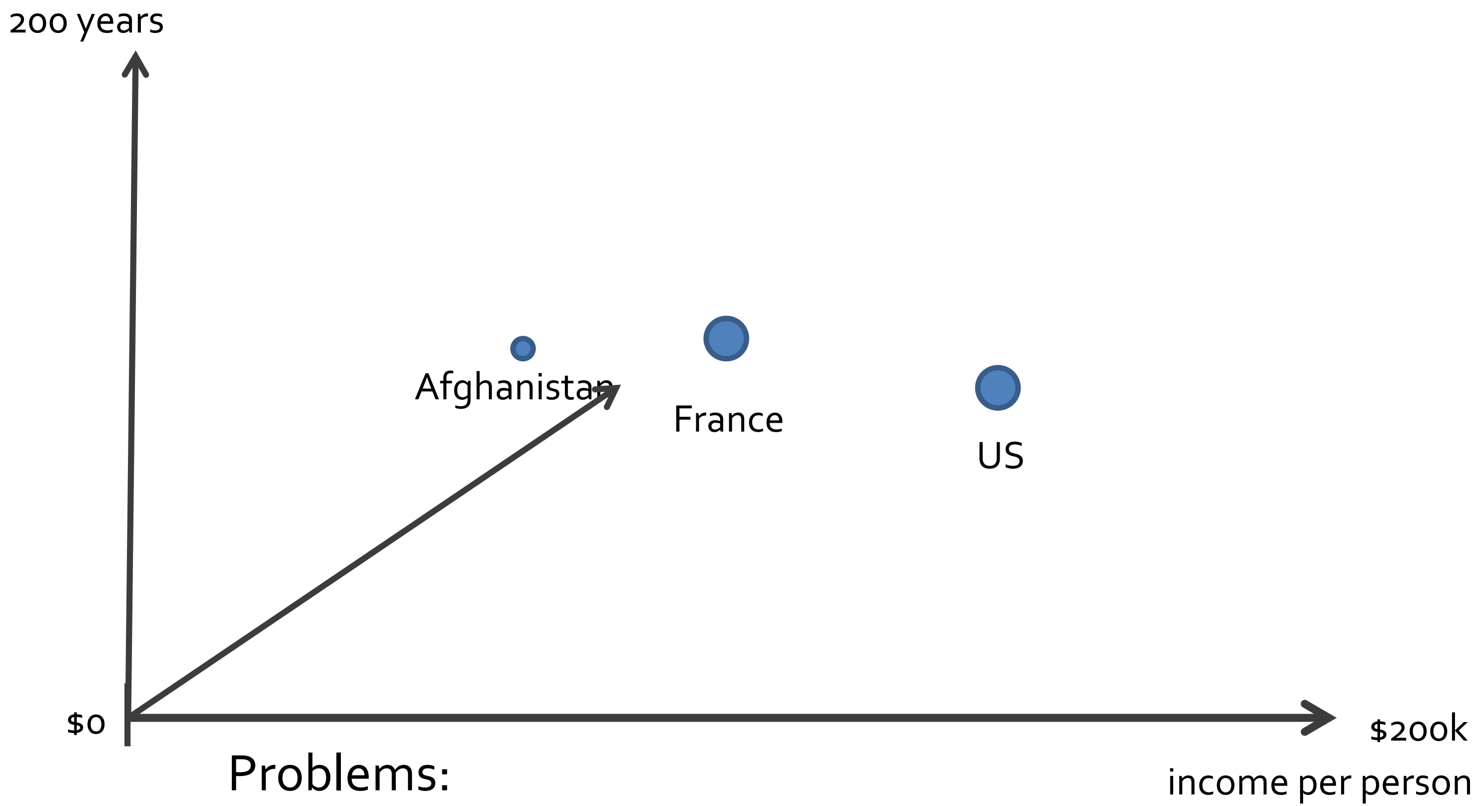
ADD ANOTHER VISUAL ENCODING



Problem:

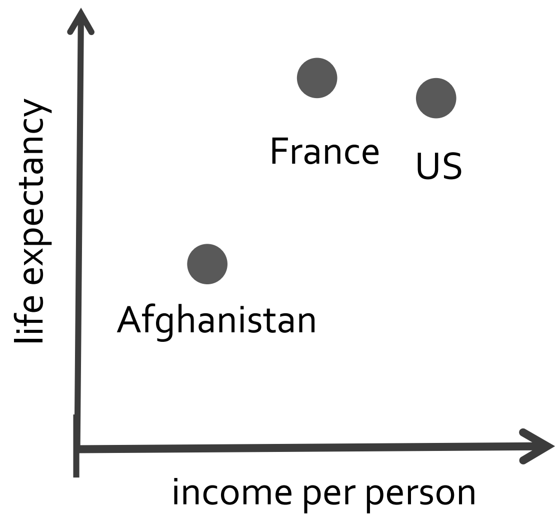
Does not scale well to more attributes

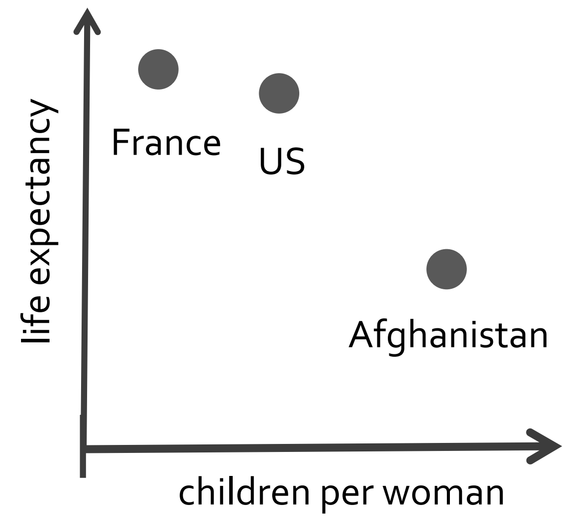
ADD AN AXIS



Problems:
Occlusion, perspective distortion, does not scale
→ Not usually recommended

ADD AN AXIS


life expectancy
France
US
Afghanistan
income per person


life expectancy
France
US
Afghanistan
children per woman

SCATTERPLOT MATRIX

This idea scales relatively well

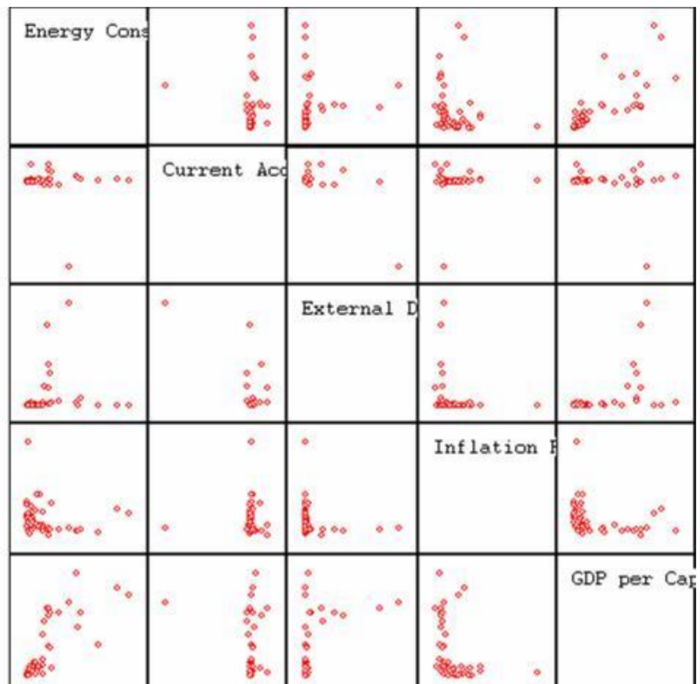
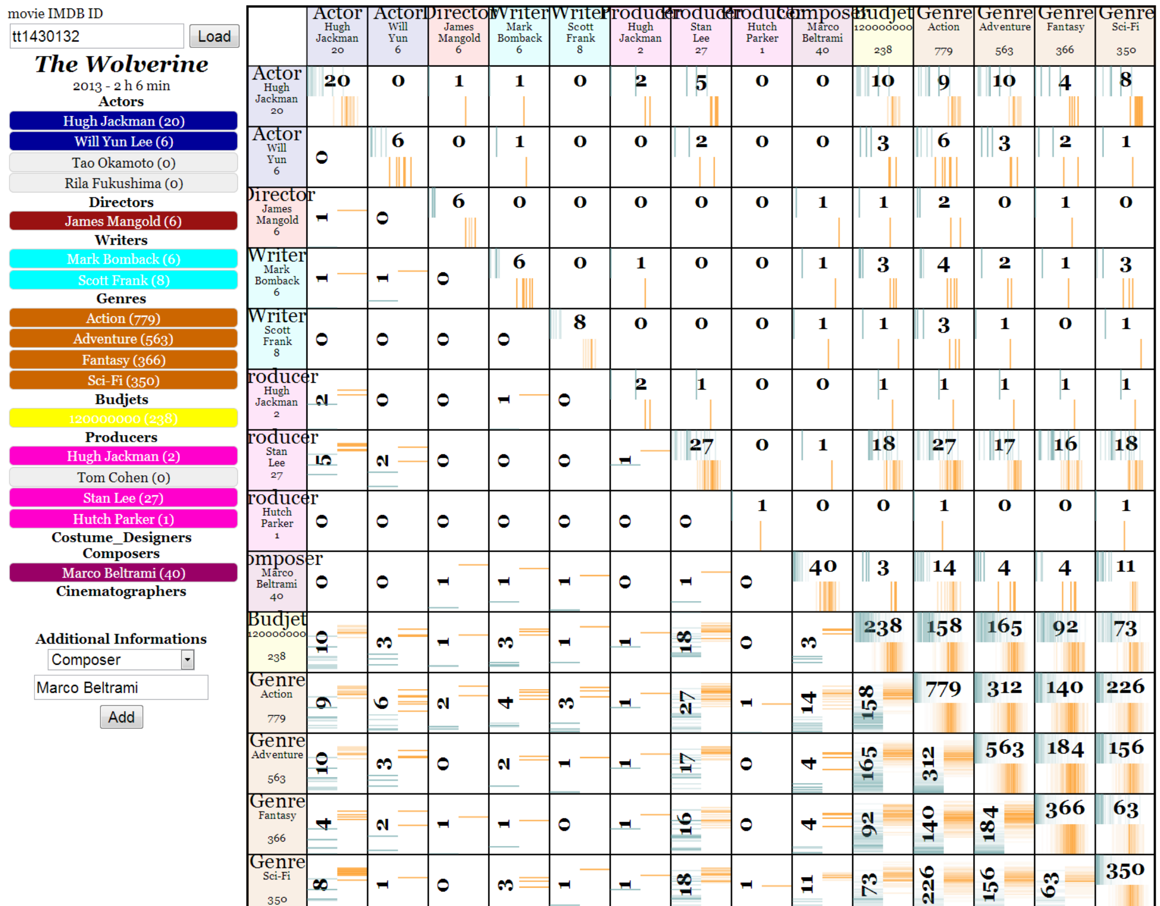


Image Source: Wikipedia

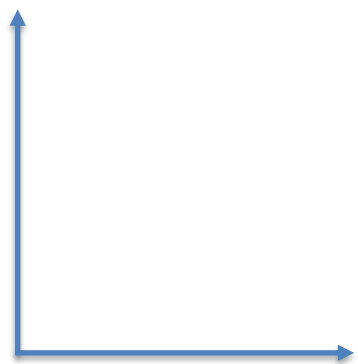
SCATTERPLOT MATRIX



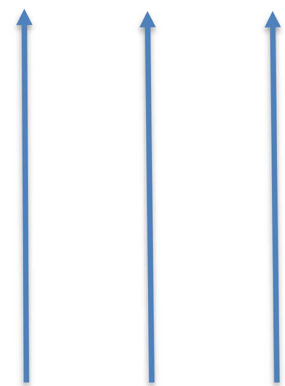
By Charles Perin

SPATIAL AXIS ORIENTATION

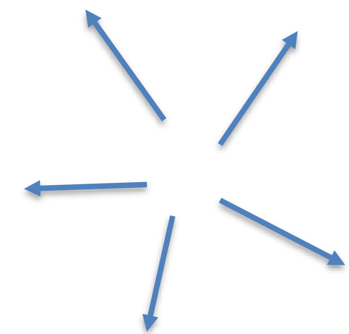
An additional design choice



rectilinear



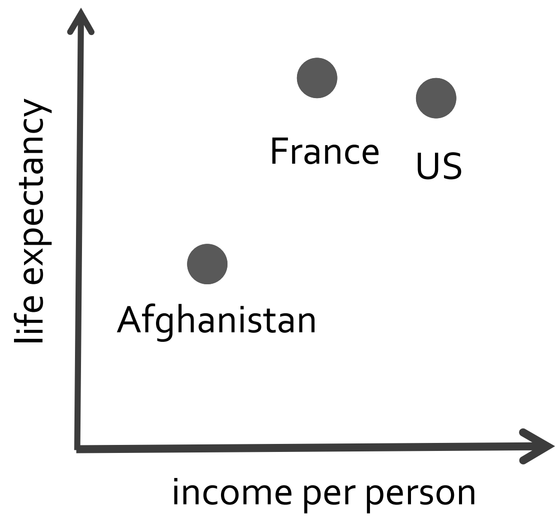
parallel



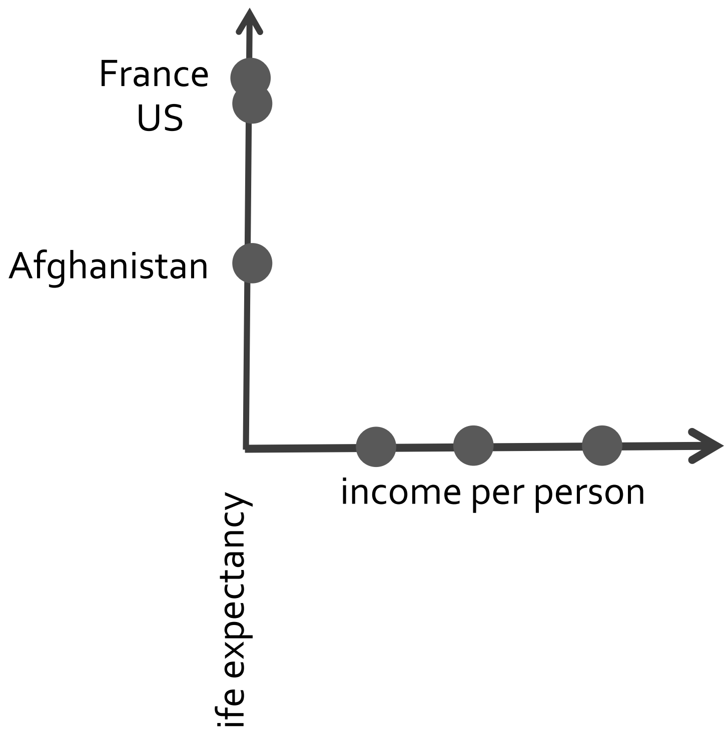
radial

parallel coordinates

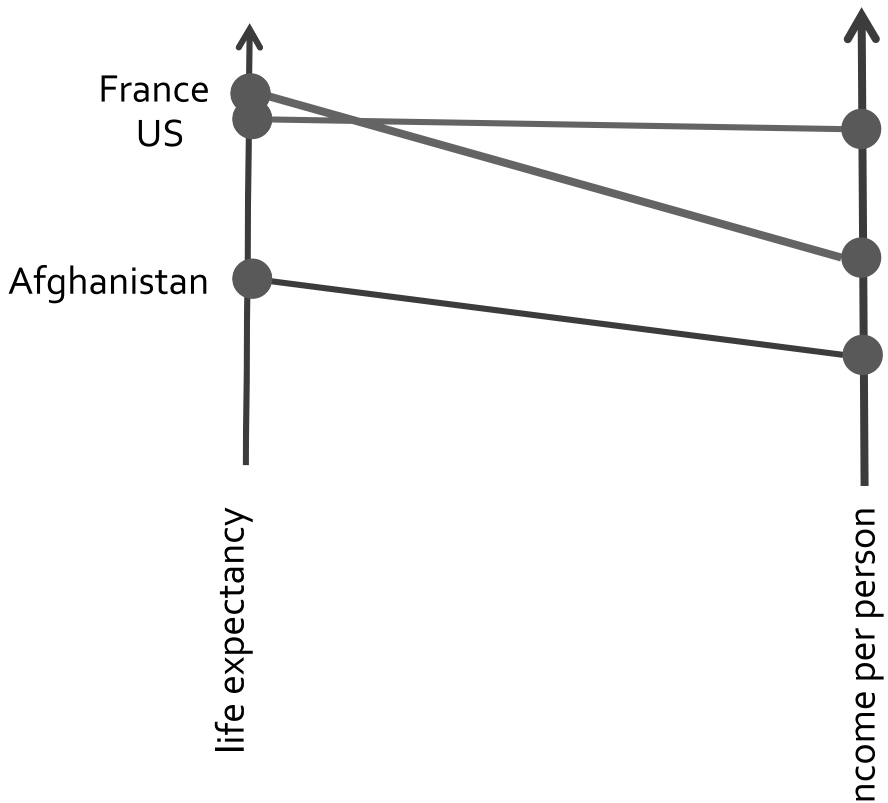
Back to our original example



Parallel Coordinates



parallel coordinates



- show correlations between neighboring axes

MULTIDIMENSIONAL DETECTIVE

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

Abstract

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

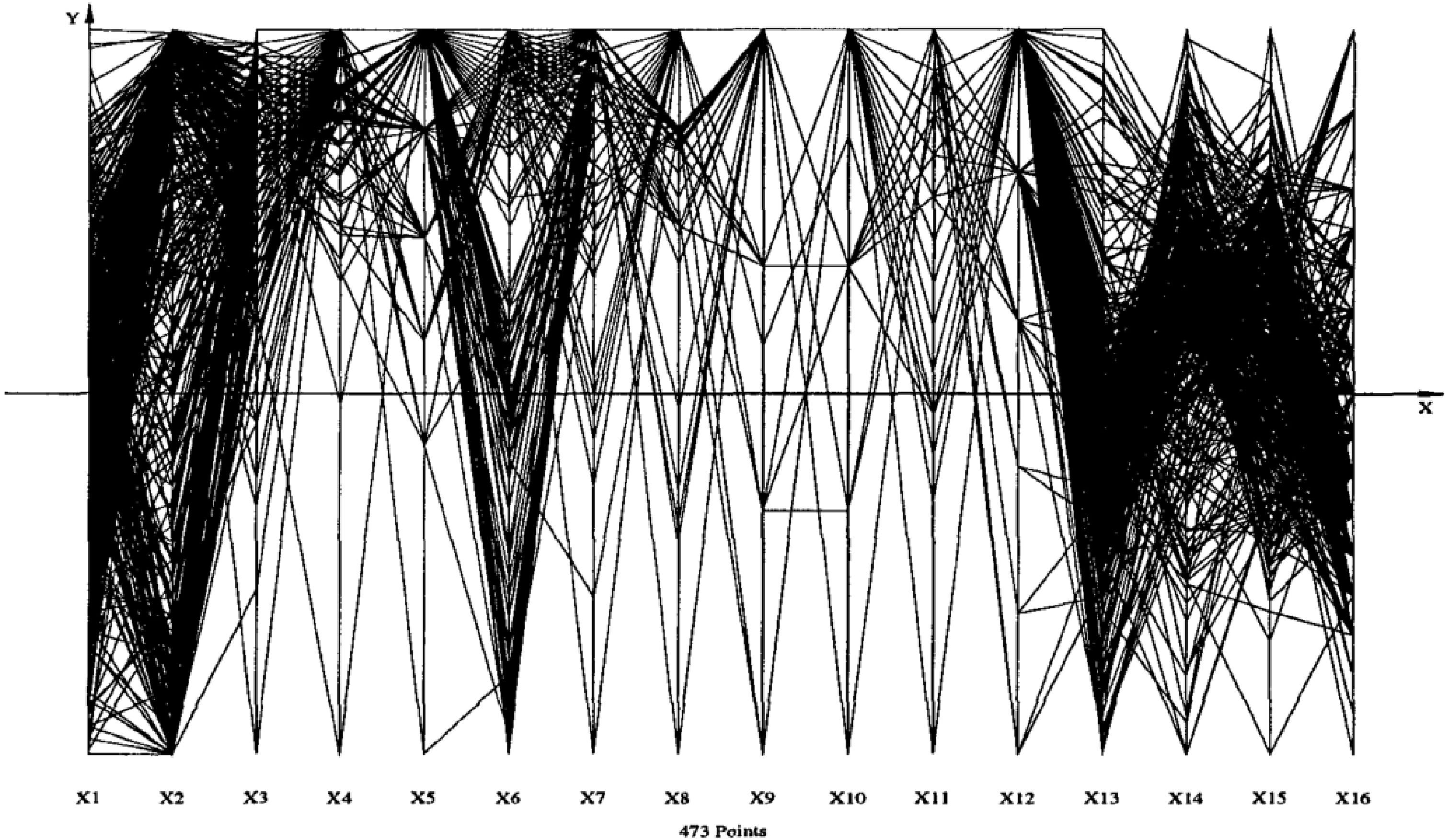
Introduction

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

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

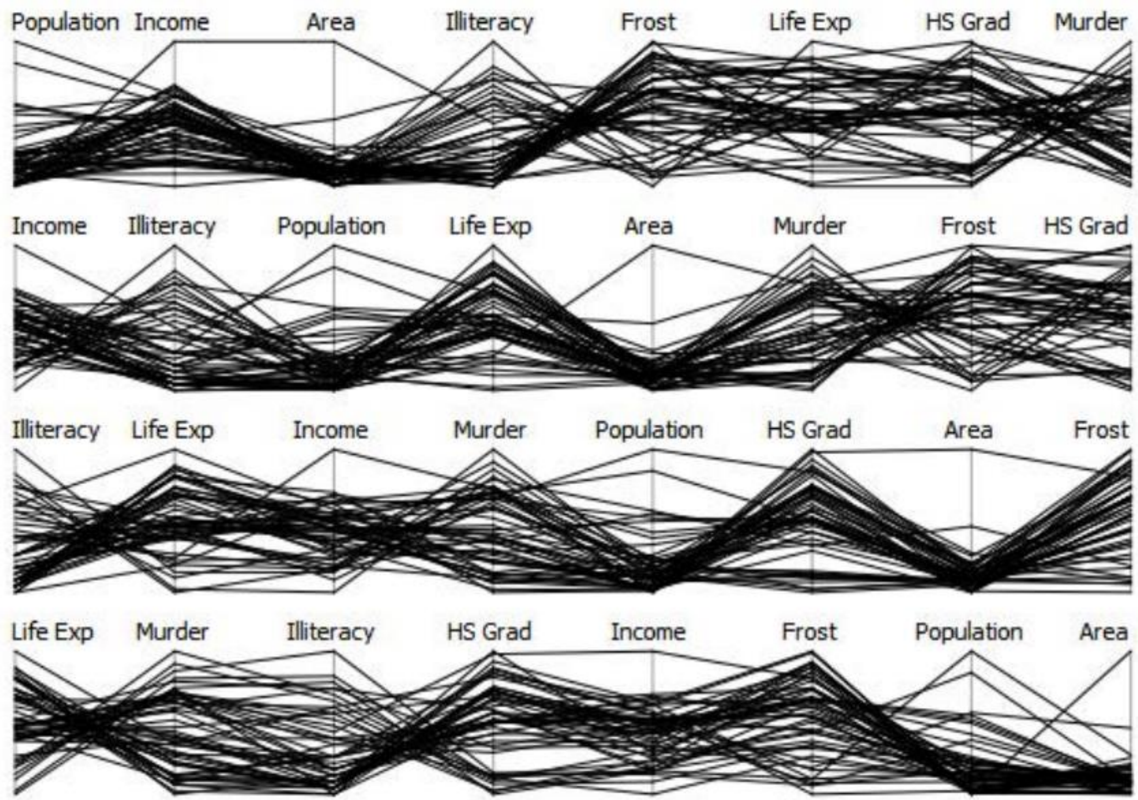
Parallel coordinates (abbr.||-coords) transform multivariate relations into 2-D patterns, a property that is well suited for Visual Data Mining.

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



Original Example from Inselberg 1997

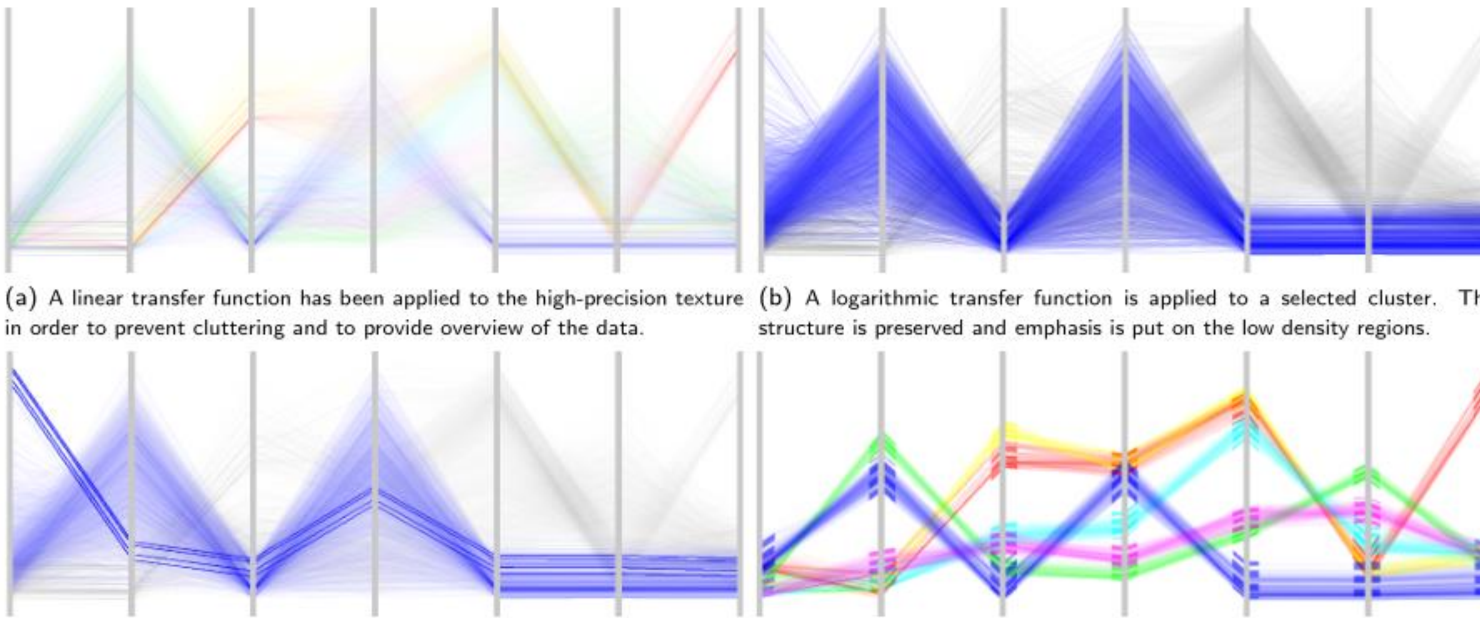
THE ORDER OF AXES MATTERS



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

REDUCE CLUTTER - HIGHLIGHT CLUSTERS

Lots of work on this. For example:



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

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

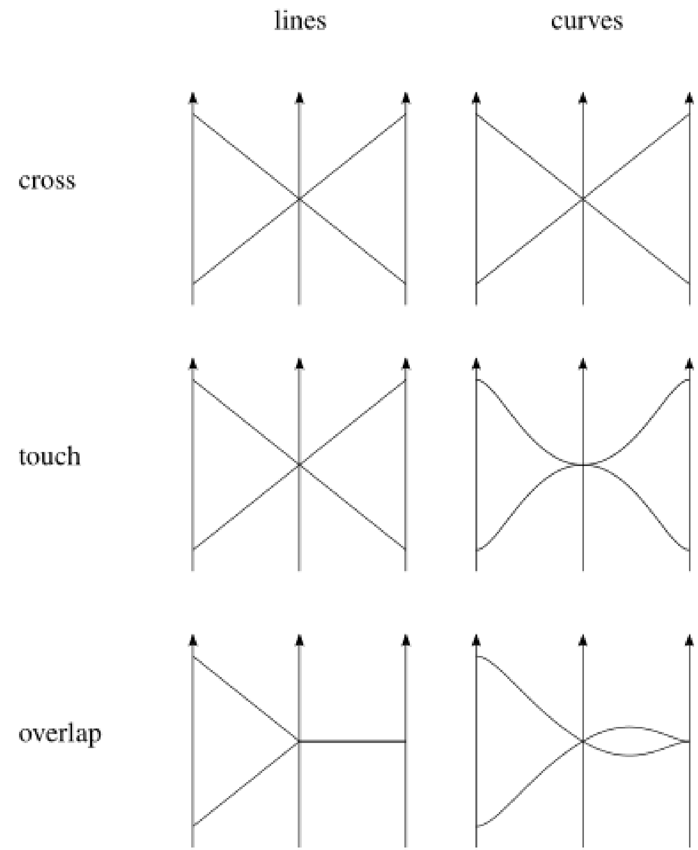
(c) Local cluster outliers are enhanced. A square root transfer function is used and the outliers are visible even through high-density regions.

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

Revealing Structure within Clustered Parallel Coordinates Displays, InfoVis 2005

HOW TO DRAW THE LINES

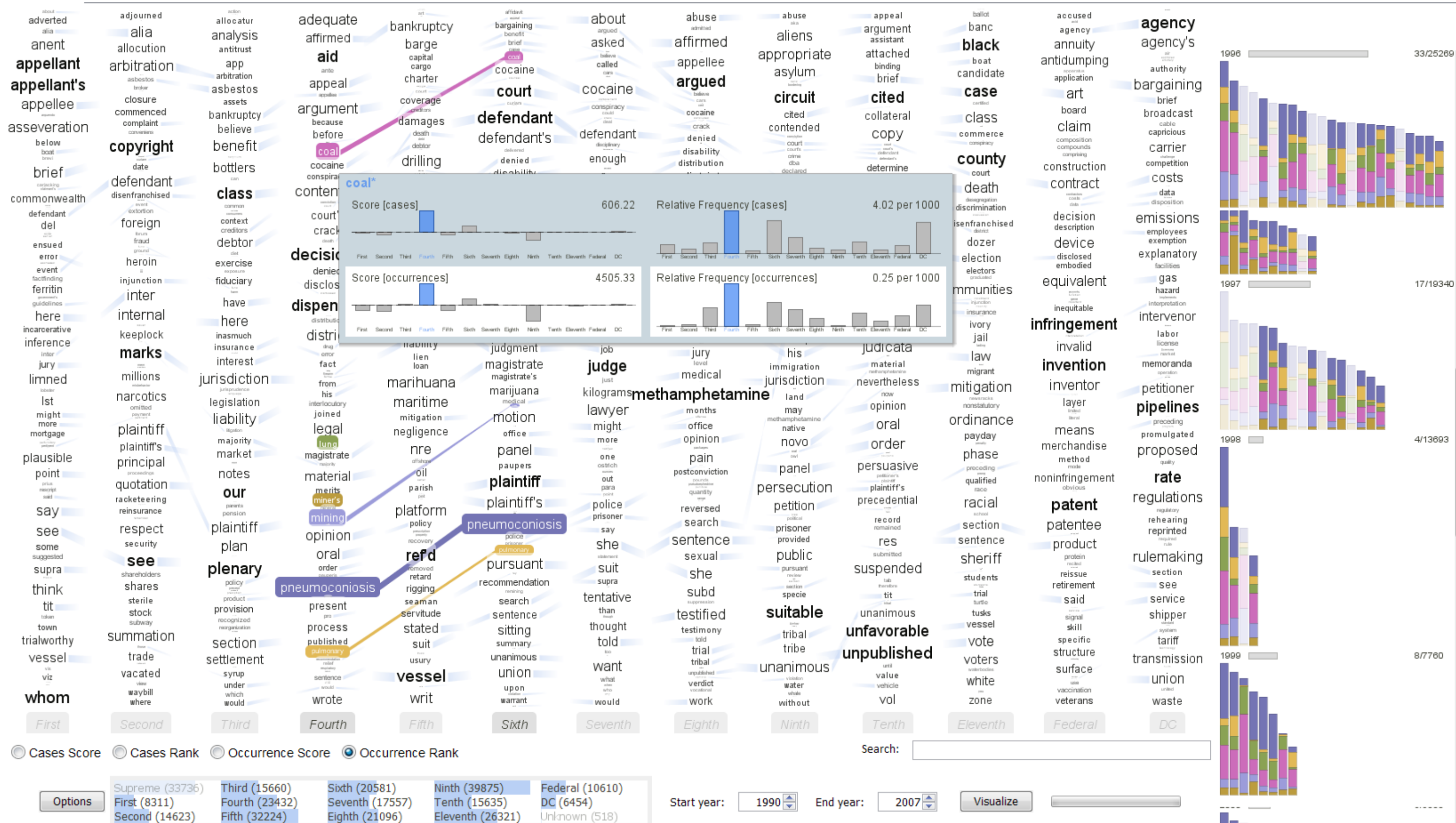
Goal: avoid ambiguity



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

COMBINE WITH OTHER VISUALIZATION TECHNIQUES

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



THERE IS MUCH MORE ON THIS…

Start here if you want more information

State of the Art of Parallel Coordinates

J. Heinrich and D. Weiskopf

Visualization Research Center, University of Stuttgart

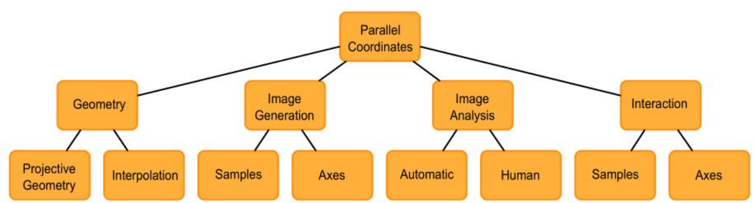


Figure 1: Taxonomy of topics for parallel coordinates in the scientific literature. The first-level nodes each represent a section in this paper, where the scope and definition of each topic will be explained.

Abstract

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

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

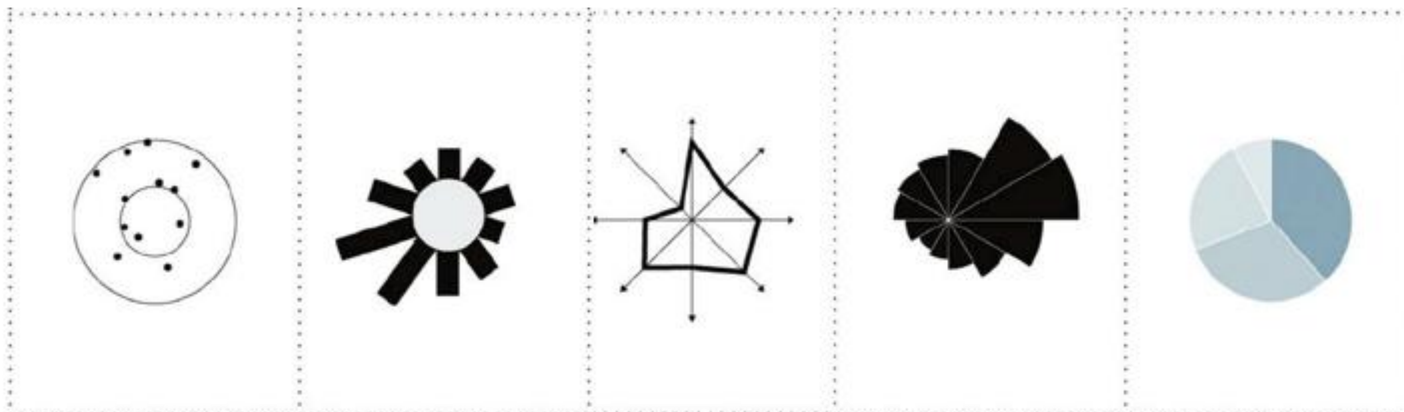
Scattering Points in Parallel Coordinates

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

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

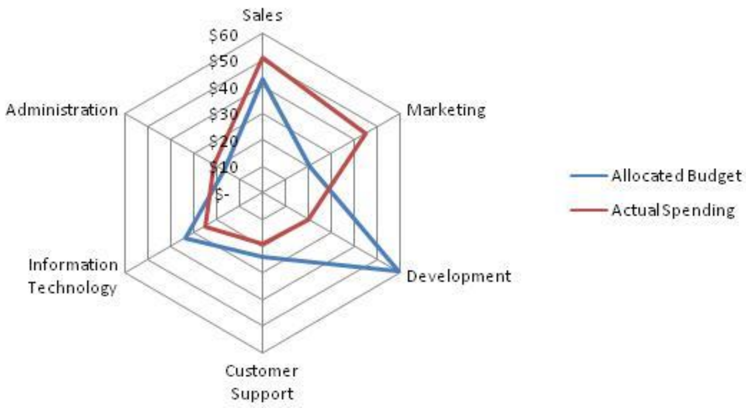
RADIAL AXES

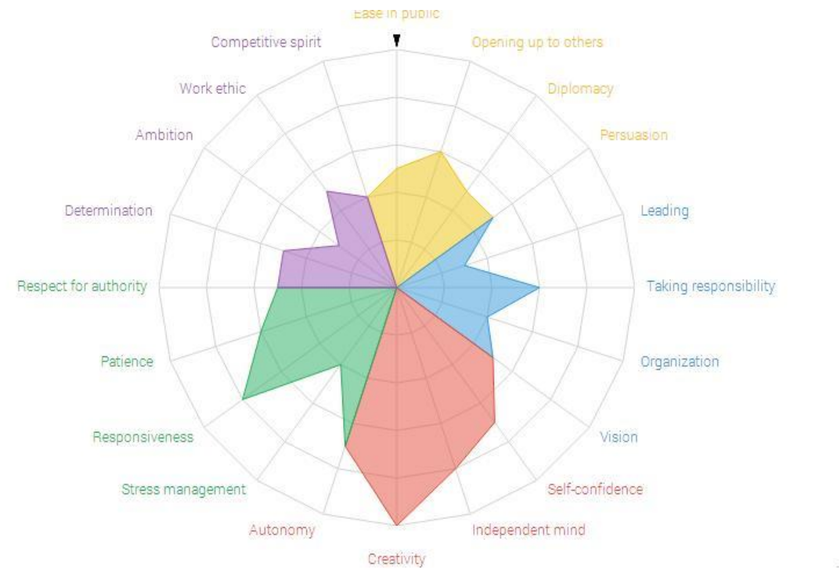
Polar



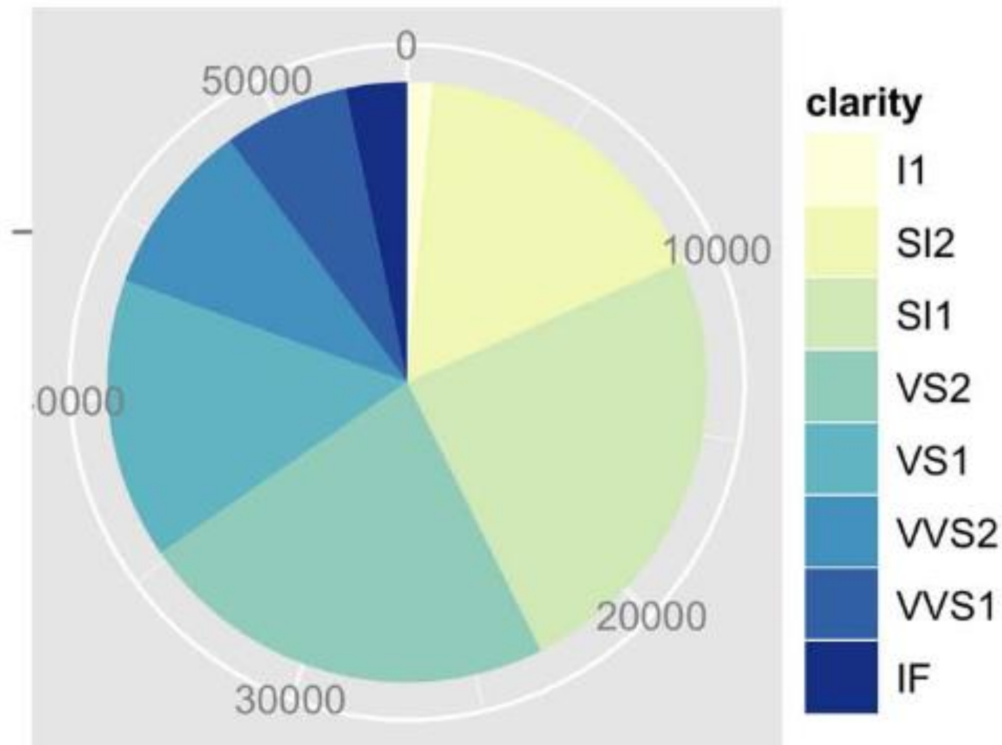
EXAMPLE: STAR PLOT

- = radial line chart

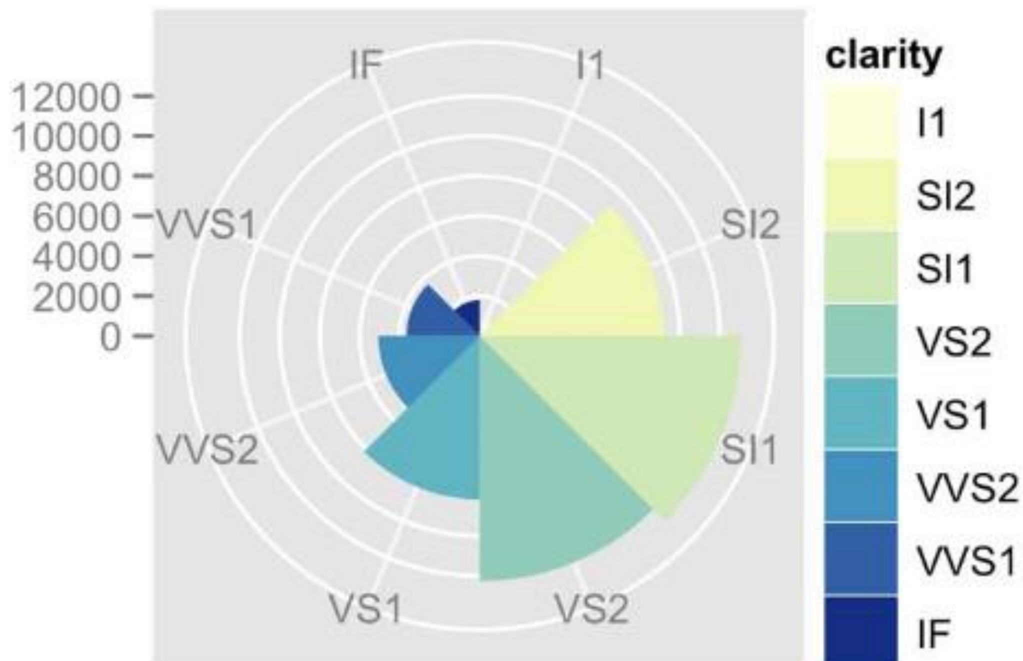




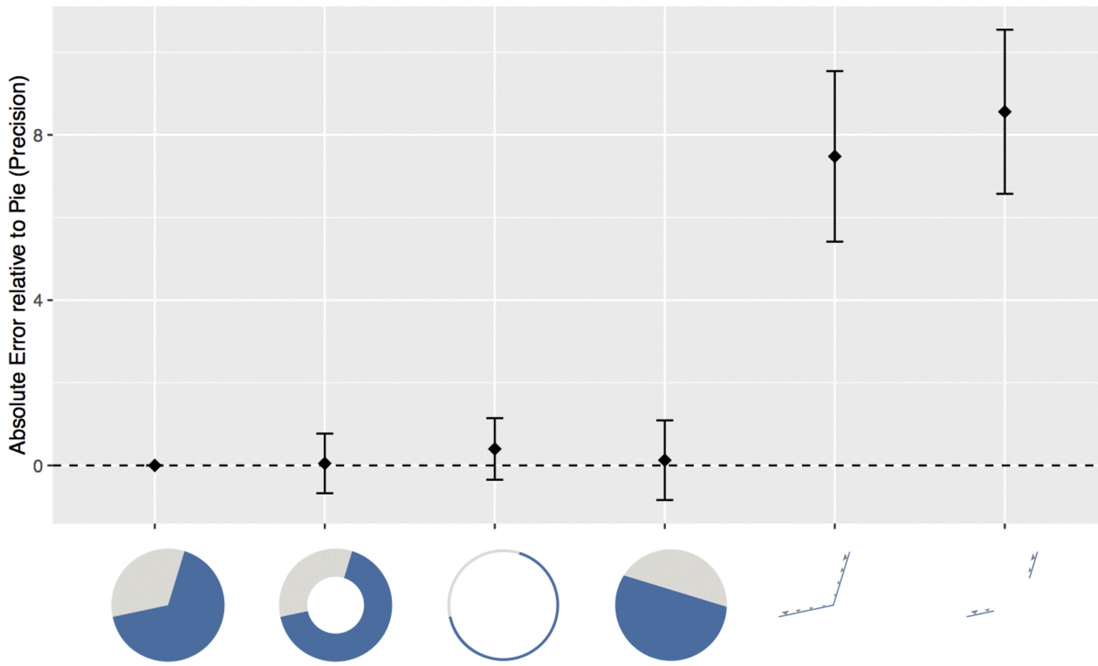
PIE CHARTS


0
10000
20000
30000
0000
50000
clarity
I1
SI2
SI1
VS2
VS1
VVS2
VVS1
IF

POLAR AREA CHARTS

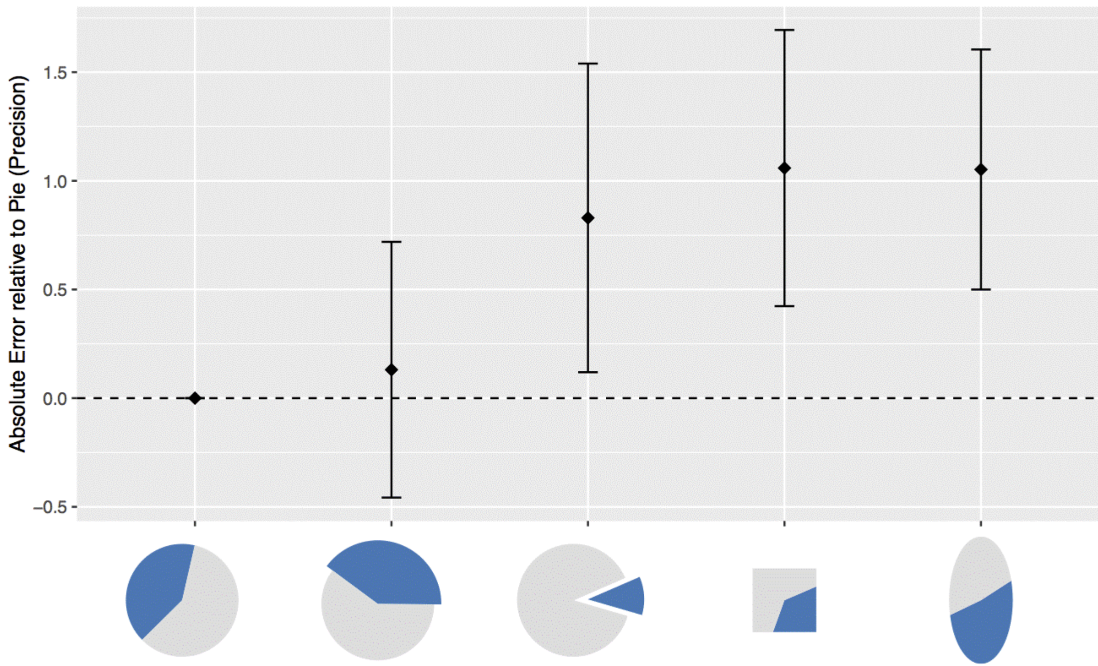


HOW DO PEOPLE READ PIE CHARTS?



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

HOW DO PEOPLE READ PIE CHARTS?



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

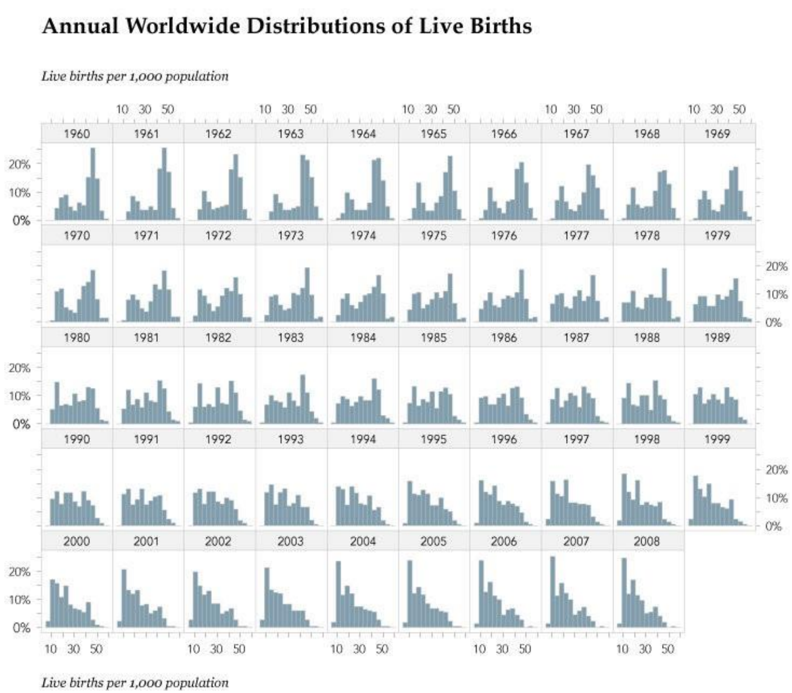
SPATIAL LAYOUT DENSITY

DATA DENSITY

MAXIMIZE THE RATIO OF:

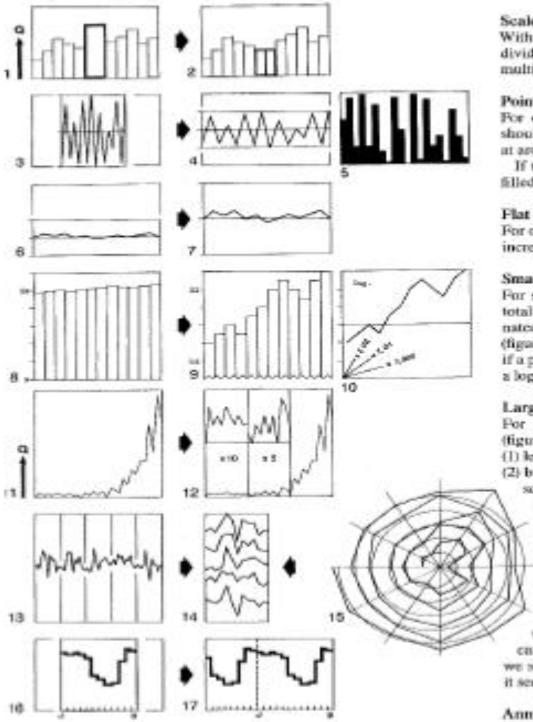
$$\frac{(\textbf{NUMBER OF ENTRIES}\text{ IN DATA})}{(\textbf{AREA}\text{ OF THE GRAPHIC})}$$

DATA DENSITY – SHRINK THE GRAPHICS



"SMALL MULTIPLES"

DATA DENSITY – SHRINK THE GRAPHICS



GRAPHIC PROBLEMS POSED BY TIME SERIES

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

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

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

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

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

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

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

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

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

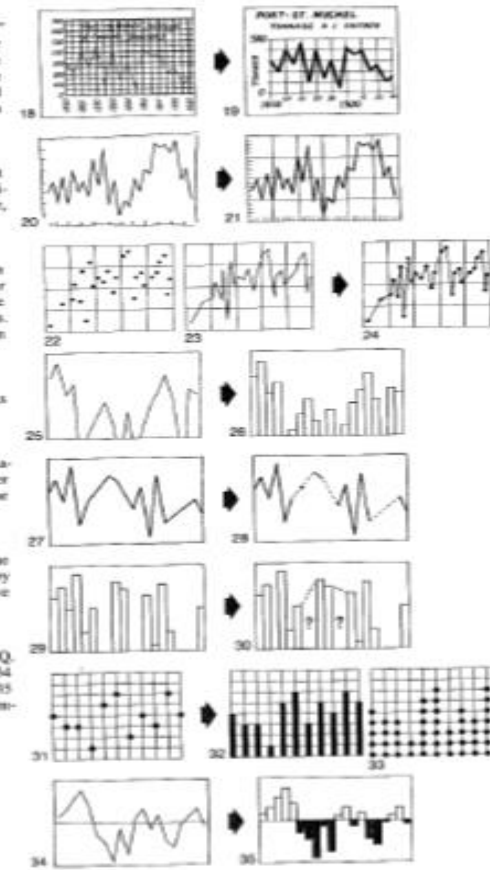
Precision reading
A precision reading (utilization on the elementary level, as in figure 24) is difficult in figure 22, which results in a poor reading of the order of the points, and in figure 23, where there is ambiguity concerning the position of the points. On the other hand, figure 22 does favor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

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

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

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

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



Bertin 1967

DATA DENSITY – SHRINK THE GRAPHICS

Placed in the relevant context, a single number gains meaning. Thus the most recent measurement of glucose should be compared with earlier measurements for the patient. This data-line shows the path of the last 80 readings of glucose:

 glucose 6.6

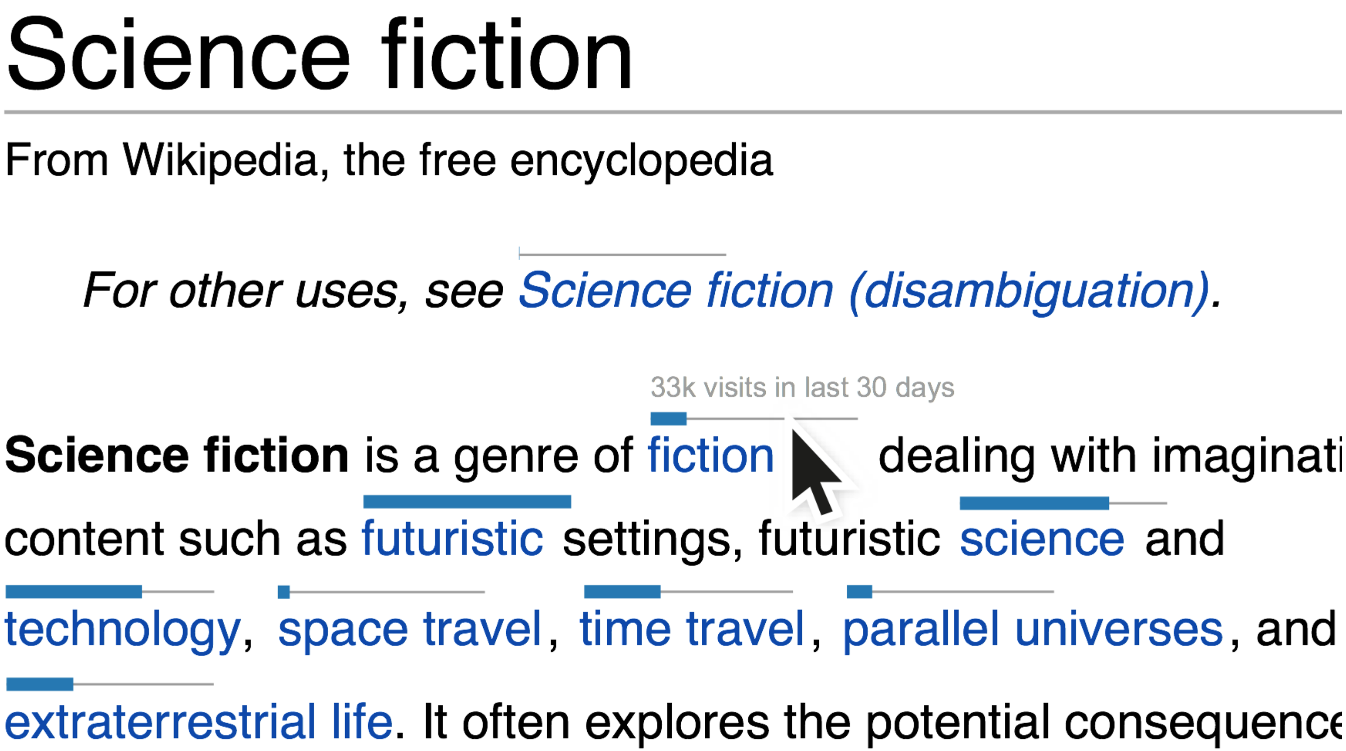
Lacking a scale of measurement, this free-floating line is dequantified. At least we do know the value of the line's right-most data point, which corresponds to the most recent value of glucose, the number recorded at far right. Both representations of the most recent reading are tied together with a color accent:

 glucose 6.6

Some useful context is provided by showing the *normal range* of glucose, here as a gray band. Compared to normal limits, readings above the band horizon are elevated, those below reduced:

 glucose 6.6 or glucose  6.6

SPARKLINES



Science fiction

From Wikipedia, the free encyclopedia

For other uses, see Science fiction (disambiguation).

33k visits in last 30 days

Science fiction is a genre of fiction dealing with imaginati
content such as futuristic settings, futuristic science and
technology, space travel, time travel, parallel universes, and
extraterrestrial life. It often explores the potential consequence

SPARKLINES

EASTERN EUROPE

Soviet cult and pragmatism in Transnistria

Experts worry that the next "Crimea"  could be the breakaway region of Transnistria  Many locals there don't share that fear, and if the last referendum holds, a large majority would welcome a Russian  annexation.

SPARKLINES

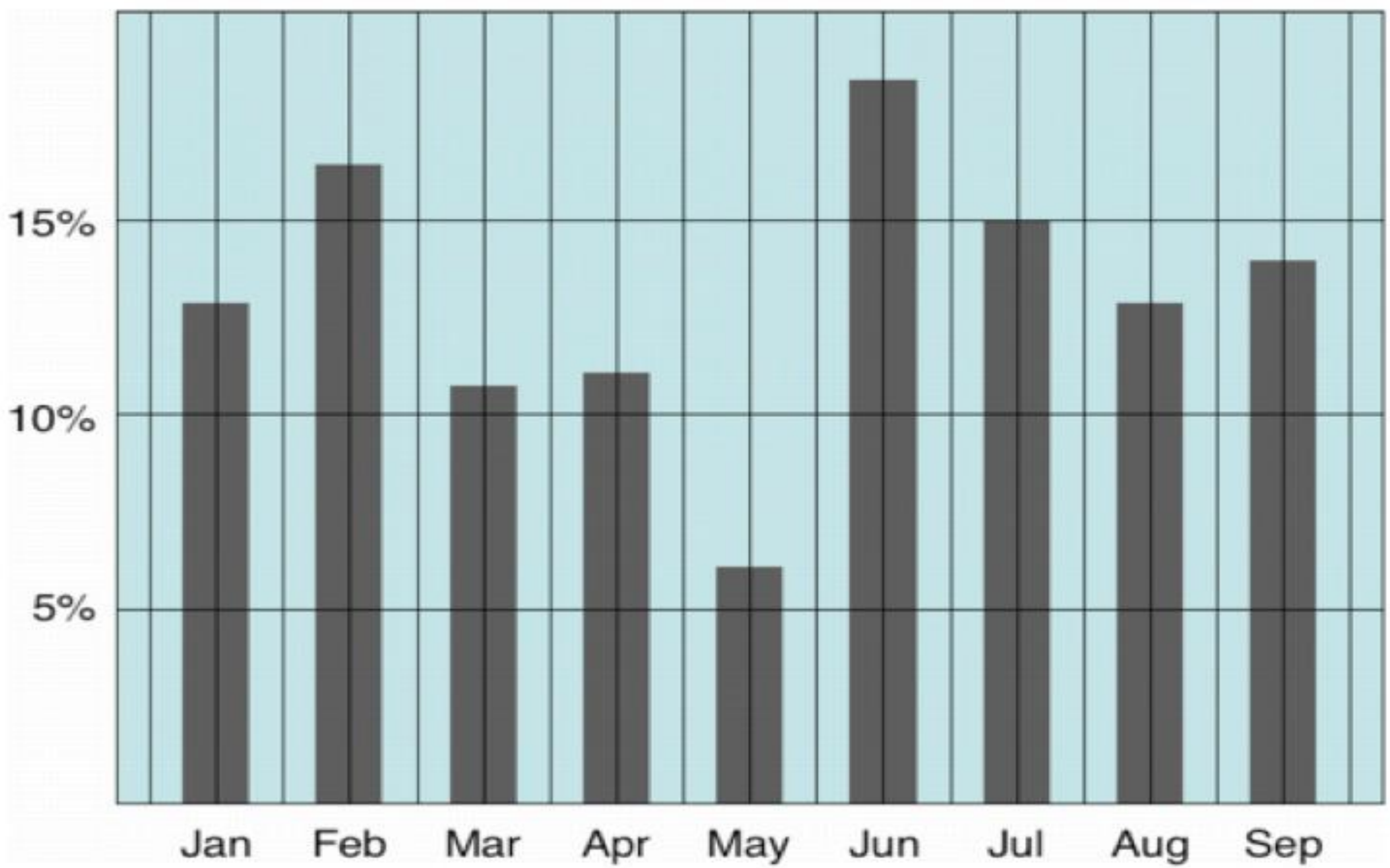
Gonzalo Higuaín slides a cross in from the right  and Ronaldo, at the front post, shoots off target.

DATA-INK RATIO

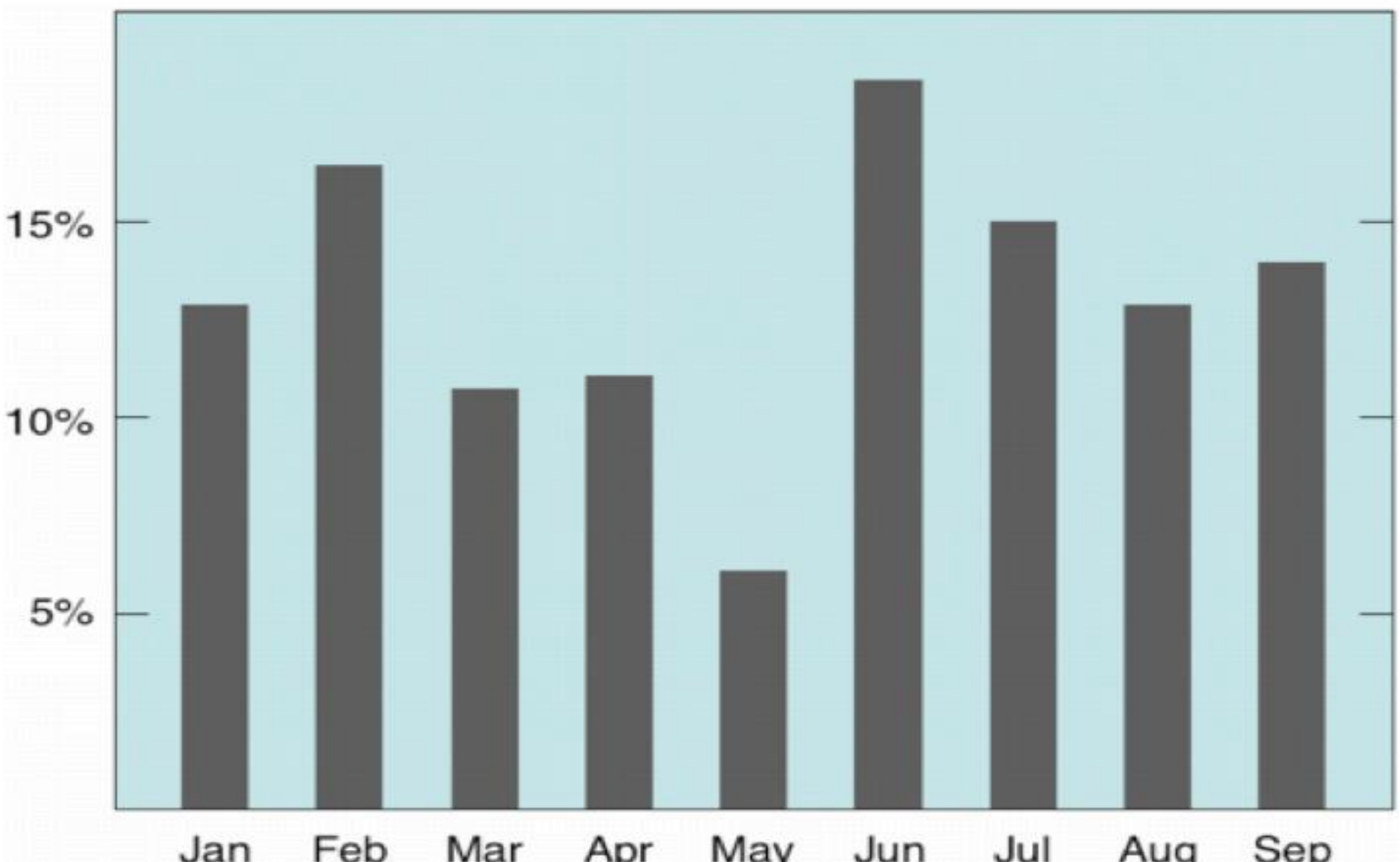
MAXIMIZE THE RATIO OF

$$\frac{(\textbf{INK} \text{ USED TO SHOW DATA})}{(\textbf{TOTAL INK} \text{ USED})}$$

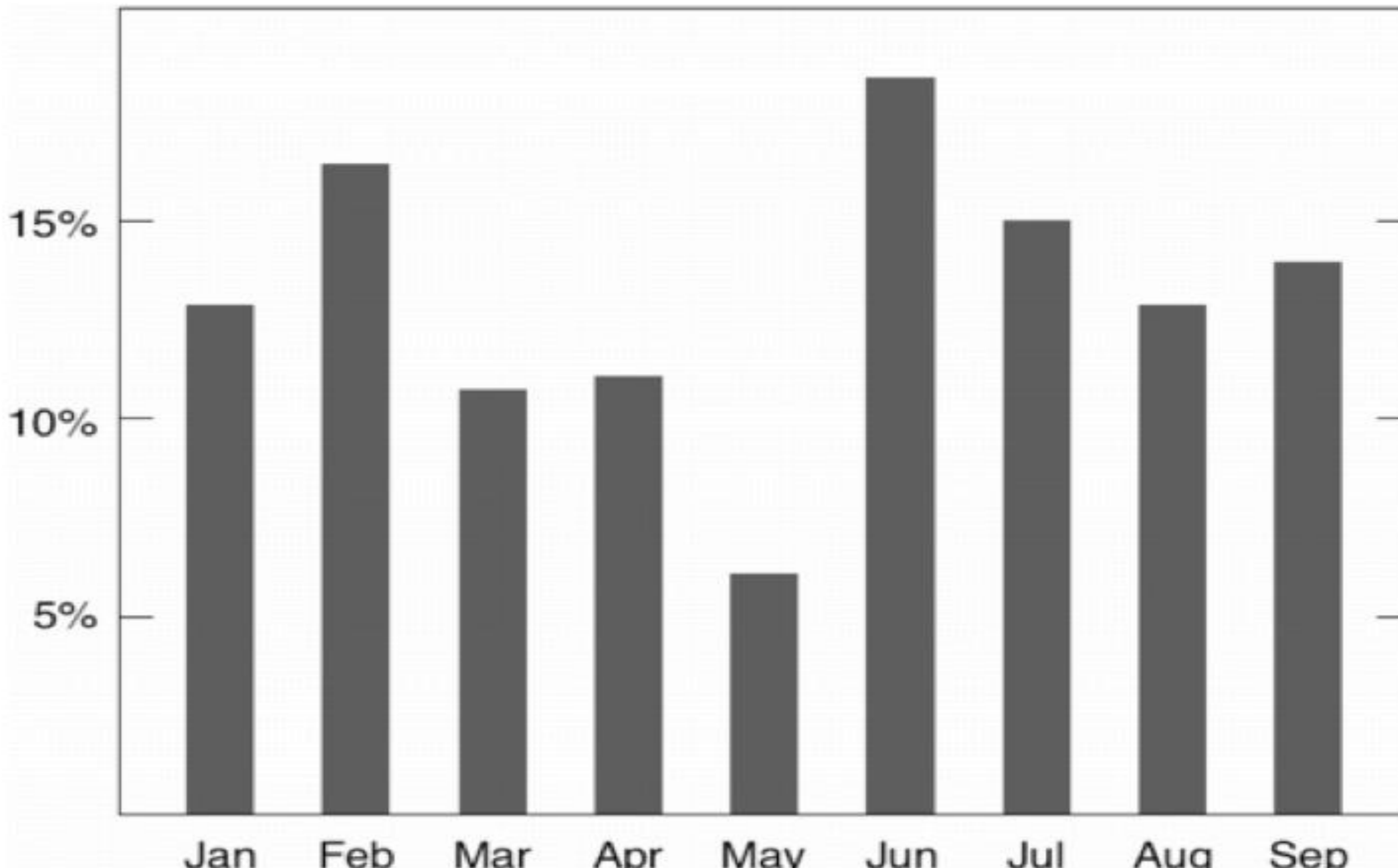
DATA-INK RATIO



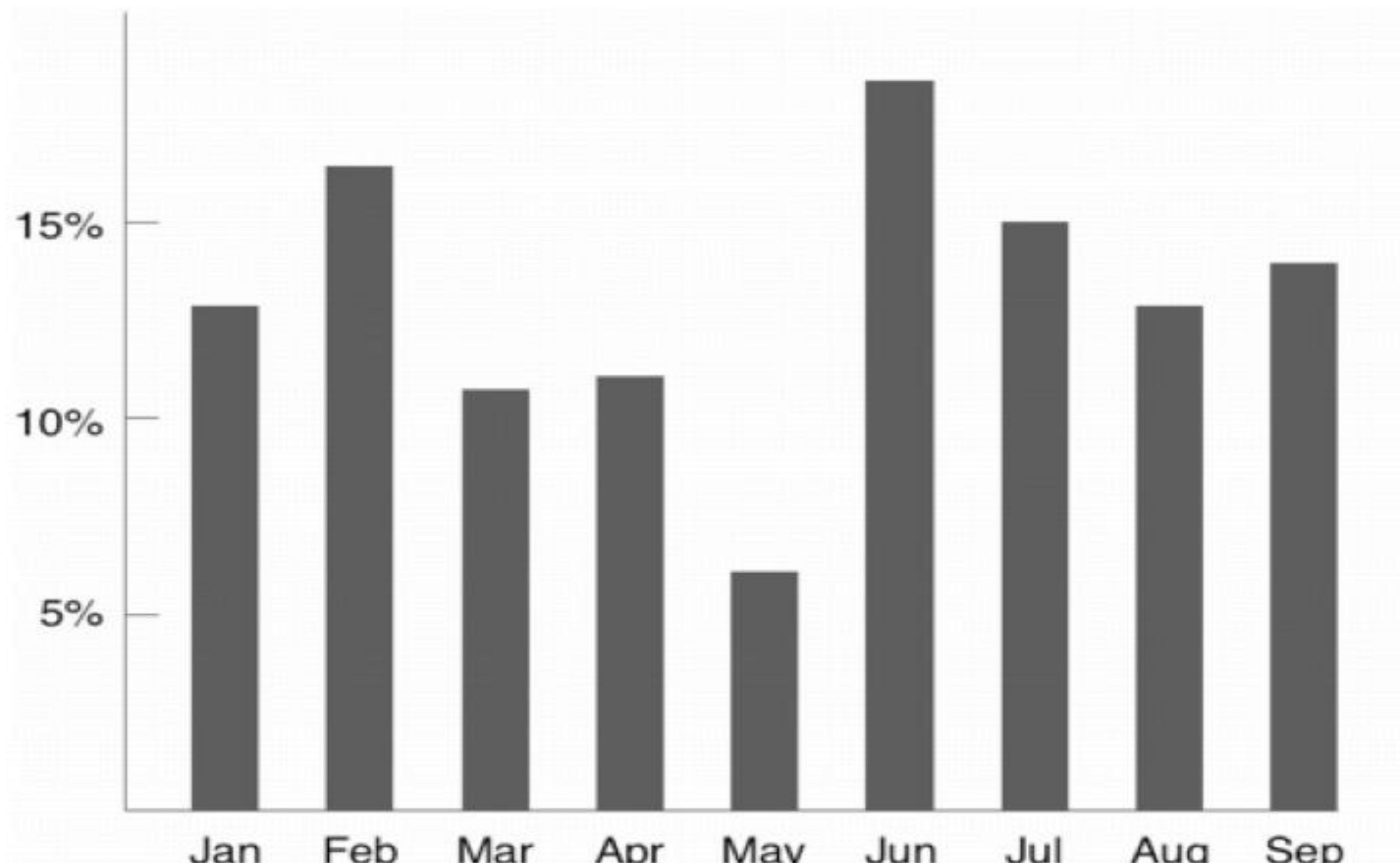
DATA-INK RATIO



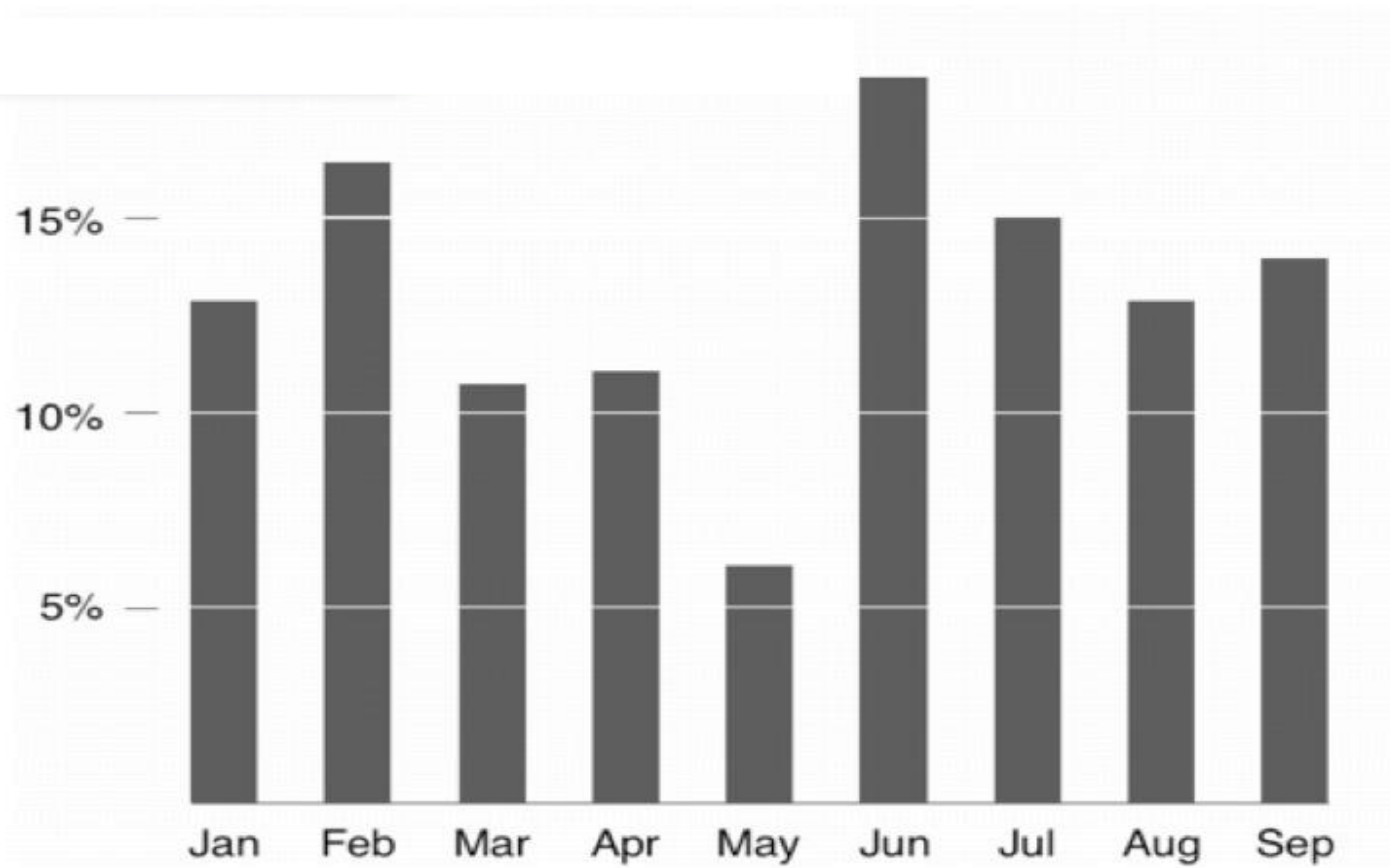
DATA-INK RATIO



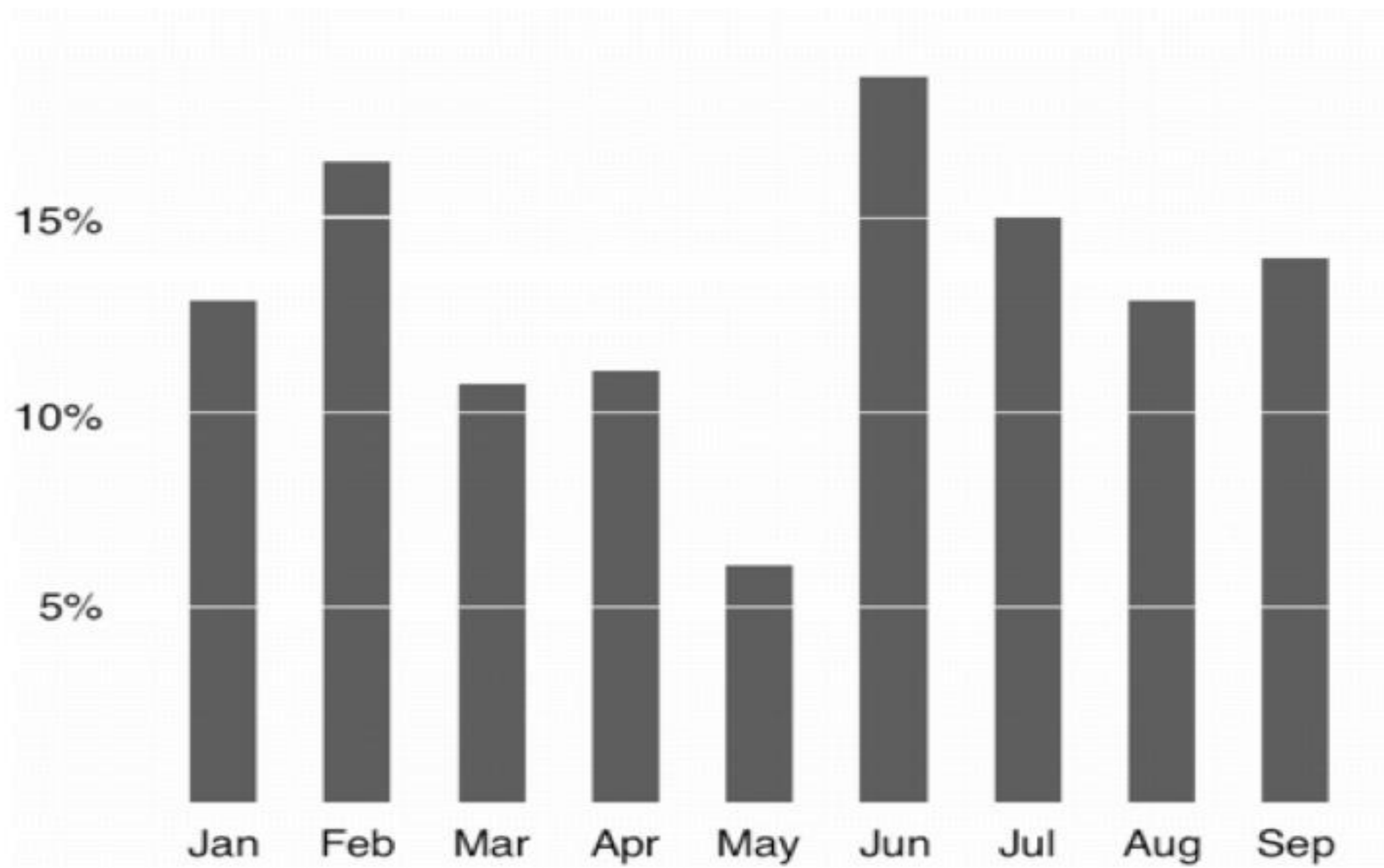
DATA-INK RATIO



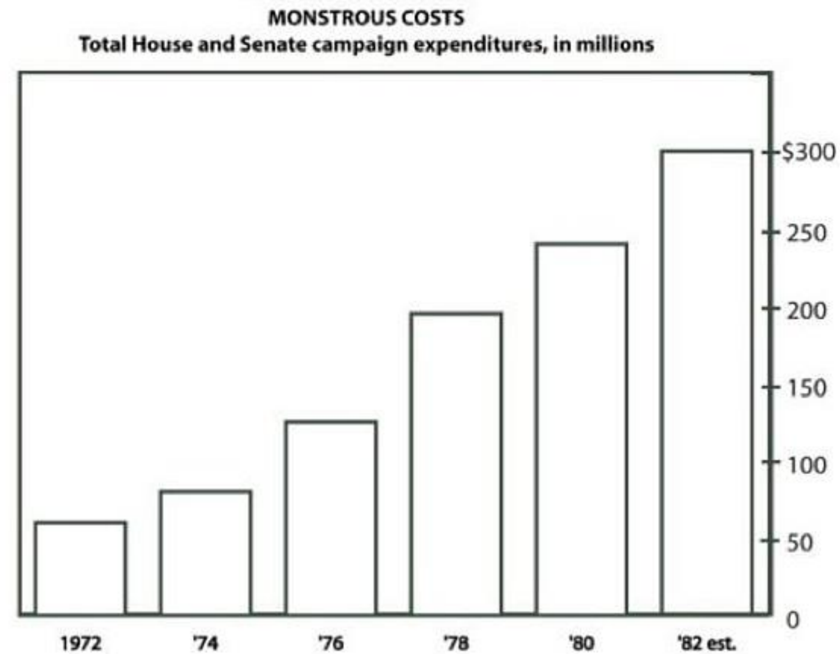
DATA-INK RATIO



DATA-INK RATIO



MINIMIZE CHART JUNK



Wayne Lytle

The Dangers of *GLITZINESS* and other Visualization Faux Pas

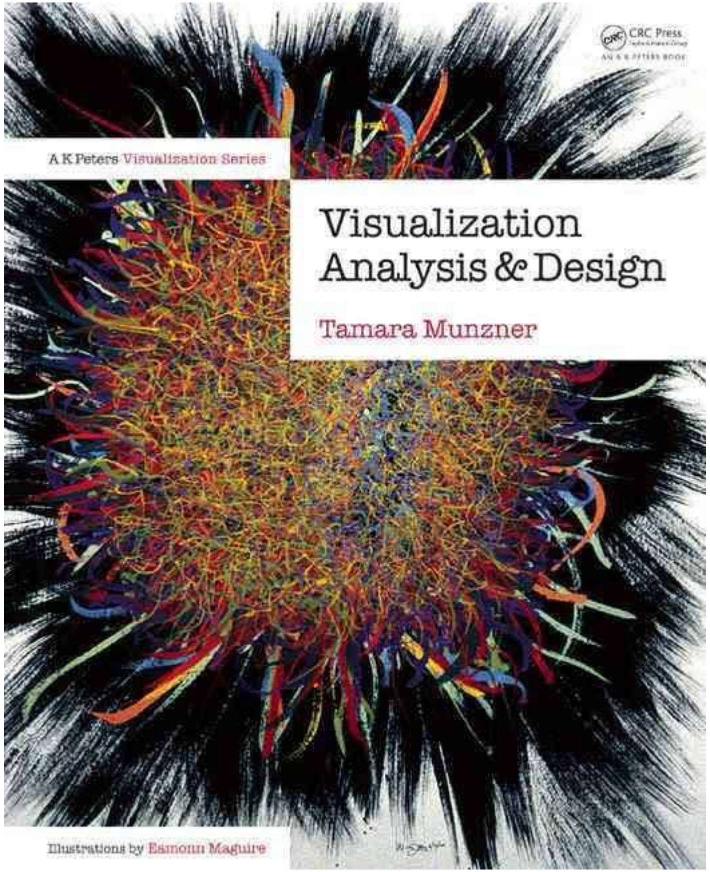
or... "What's Wrong with this Visualization?"

TUFTE'S INTEGRITY PRINCIPLES

- MAXIMIZE THE DATA-INK RATIO
- AVOID CHART JUNK *(SOMETIMES)*
- LAYER INFORMATION
- MAXIMIZE THE DATA DENSITY
  - *SHRINK THE GRAPHICS*
  - *MAXIMIZE THE AMOUNT OF DATA SHOWN (SOMETIMES)*

EDWARD TUFTE

READINGS



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

ACKNOWLEDGEMENTS

Slides in were inspired and adapted from slides by

- Nicolai Marquardt (University College London)
- Uta Hinrichs (University of St. Andrews)
- Saul Greenberg (University of Calgary)