Information Visualization

Making Data Understandable

Petra Isenberg
petra.isenberg@inria.fr

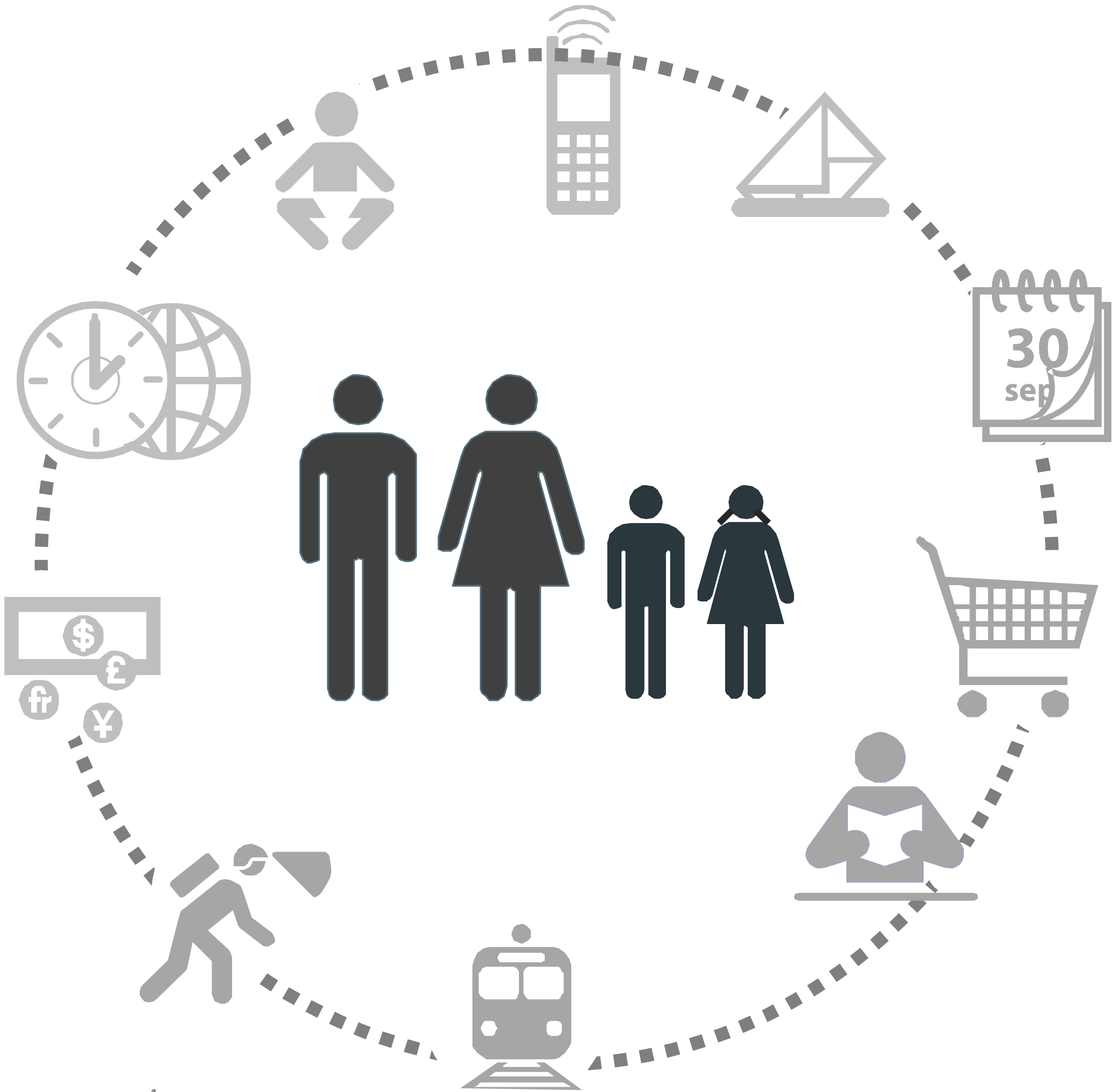
INRIA

After today you will...

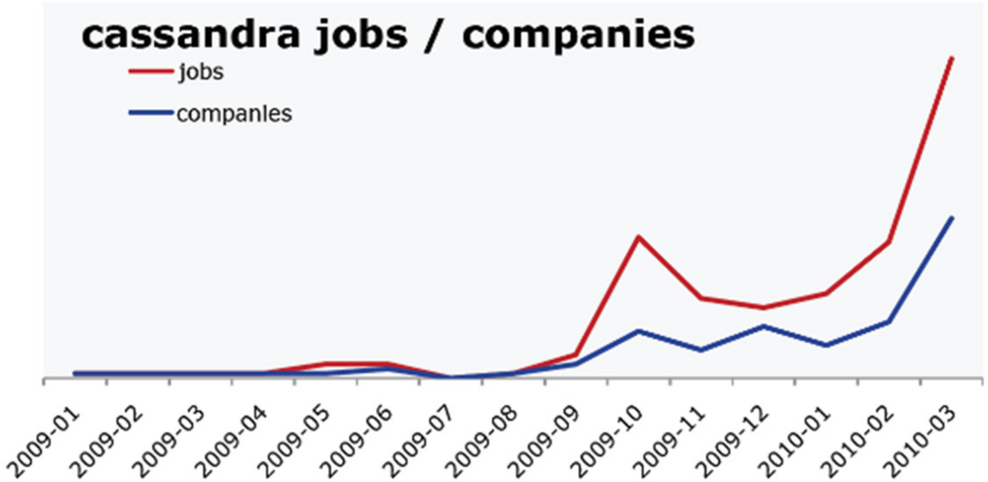
- have gained an overview of the research area
- learned principles of data representation and interaction
- learned about existing applications, tools, and techniques
- develop an understanding of and ability to critique information visualizations

It is estimated that 800 exabyte (**800x 10^19)**
of digital information will be generated this year

[source: The Diverse and Exploding Digital Universe, IDC, 2008]
[credit: Did You Know; Fisch, McLeod, Brenman]


30
sep

Hiring trends for data science



It's not easy to get a handle on jobs in data science. However, data from O'Reilly Research shows a steady year-over-year increase in Hadoop and Cassandra job listings, which are good proxies for the "data science" market as a whole. This graph shows the increase in Cassandra jobs, and the companies listing Cassandra positions, over time.

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

Hal Varian, chief economist at Google

Question

how can we effectively access data?

- understand its structure?
- make comparisons?
- make decisions?
- gain new knowledge?
- convince others?
- ...

Today…

Information Visualization

Example

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Raw Data from Anscombe’s Quartet

[Source: Anscombe's quartet, Wikipedia]

Statistical Analysis

For all four columns, the statistics are identical

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

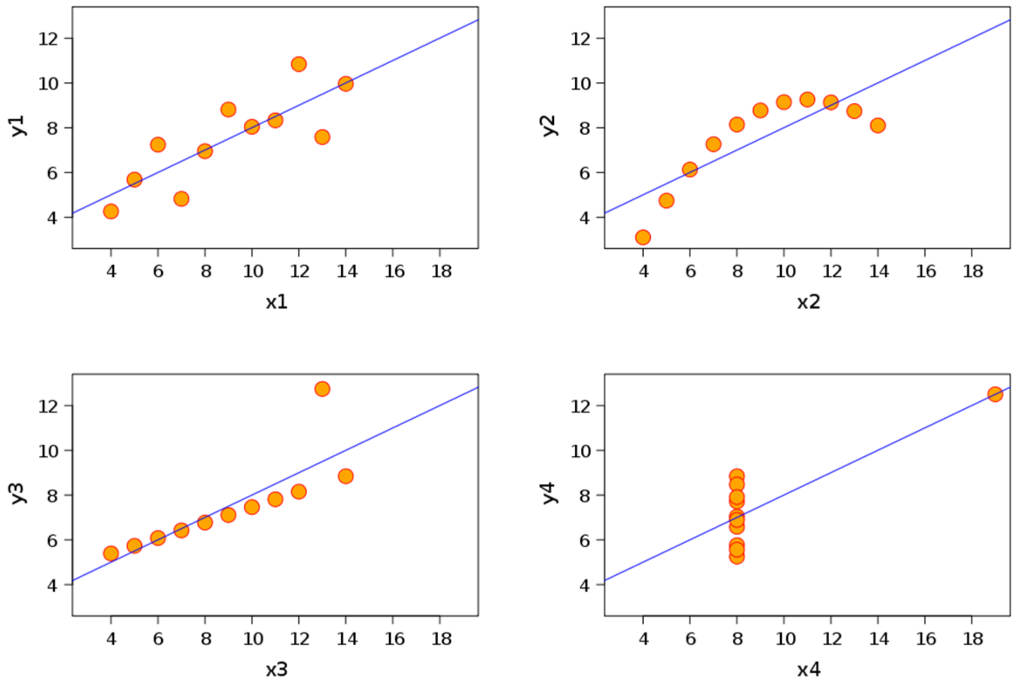
Mean of x	9.0
Variance of x	11.0
Mean of y	7.5
Variance of y	4.12
Correlation between x and y	0.816
Linear regression line	$y = 3 + 0.5x$

[Source: Anscombe's quartet, Wikipedia]

Visual Representation of the Data

Visual representation reveals a different story

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

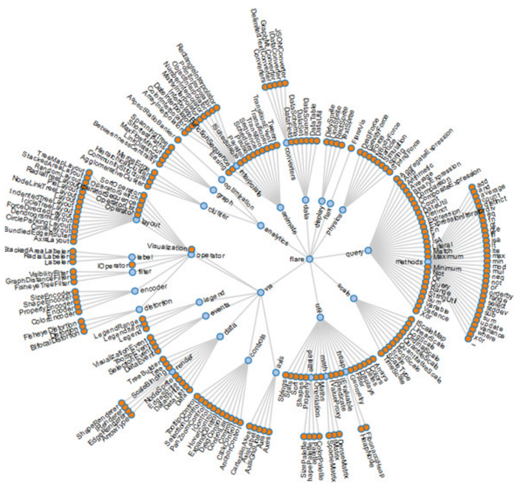


[Source: Anscombe's quartet, Wikipedia]

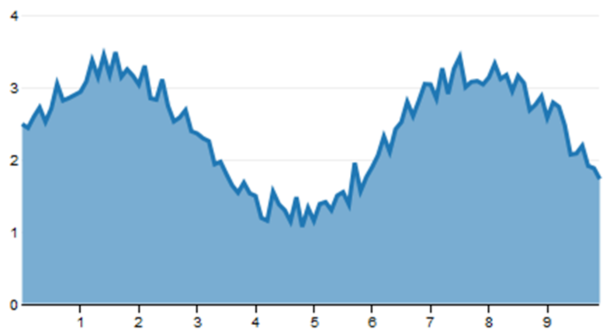
Why visual data representations?

- Vision is our most dominant sense
- We are very good at recognizing visual patterns
- We need to see and understand in order to explain, reason, and make decisions

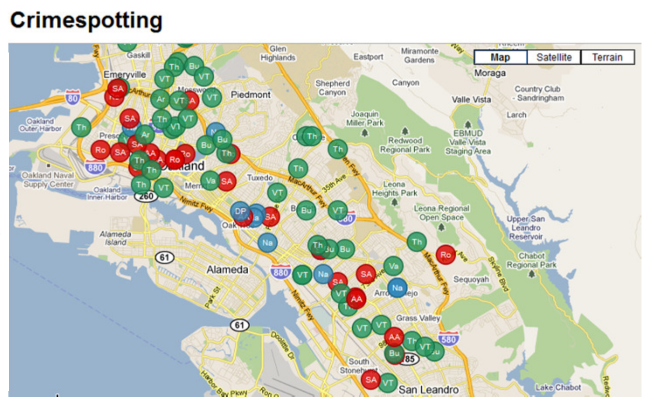
common examples:



graphs / hierarchies



charts



maps

all examples from: http://vis.stanford.edu/protovis/

Other benefits of visualization

- expand human working memory
  - offload cognitive resources to the visual system,
- reduce search
  - by representing a large amount of data in a small space,
- enhance the recognition of patterns
  - by making them visually explicit
- aid monitoring of a large number of potential events
- provides a manipulable medium & allows exploration of a space of parameter values.

百 聞 不 如 一 見

"One hundred rumors are not comparable to one look."

An Old Chinese Inscription

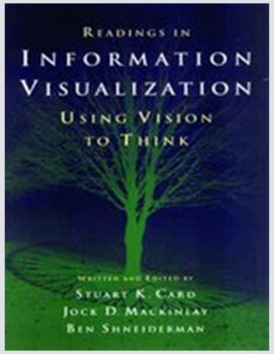
Via Brinton, Graphic Presentation, 1939

Information visualization

- Create visual representation
- Concentrates on abstract data
- Includes interaction

Official Definition:

The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.
[Card et al., 1999]



Functions of Visualizations

- Recording information
  - Tables, blueprints, satellite images
- Processing information
  - needs feedback and interaction
- Presenting information
  - share, collaborate, revise
  - for oneself, for one’s peers and to teach
- Seeing the unseen

Information Visualization Research Field

- ~25 years old
- came out of scientific visualization, computer graphics, HCI
- major conferences

IEEE VisWeek (3 main conferences: Vis, InfoVis, VAST)

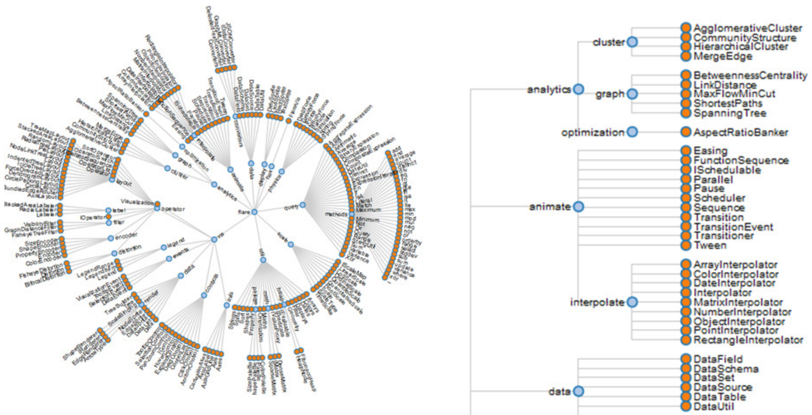
EG EuroVis (both SciVis and InfoVis)

ACM CHI (general HCI)

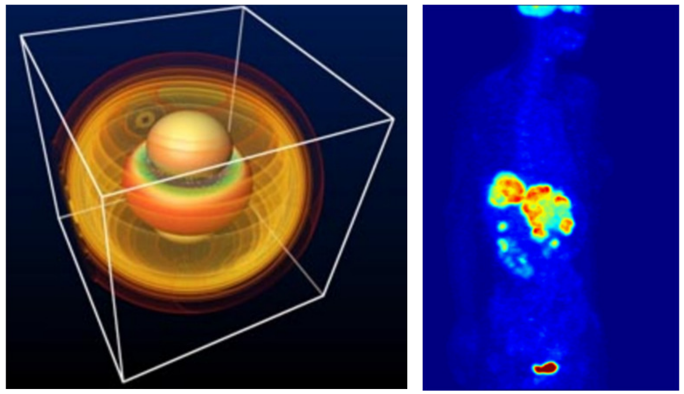
Information vs. Scientific Visualization

InfoVis
- focus on abstract data
- position of data in space can be chosen freely
- typical techniques:
  - (we will see later)

SciVis
- focus on scientific data
- position of data in space is typically fixed
- typical techniques
  - flow visualization
  - volume visualization (CT, MRT)



different spatial graph layouts of abstract data



spatial representation of gravity waves & a PET scan

there exists a gray region between both fields (e.g. maps)

Visual Analytics

- Related to both information visualization and scientific visualization
- Focus on whole data analysis process

More information: Illuminating the Path

Visualization of abstract data has been practiced for hundreds of years…

HISTORICAL EXAMPLES

The Broadway Street Pump

- In 1854 cholera broke out in London
  - 127 people near Broad Street died within 3 days
  - 616 people died within 30 days
- People thought it spread by “miasma in the atmosphere”
- Dr. John Snow was the first to link contaminated water to the outbreak of cholera
- How did he do it?
  - he talked to local residents
  - identified a water pump as a likely source
  - used maps to illustrate his theory
  - convinved authorities to disable the pump

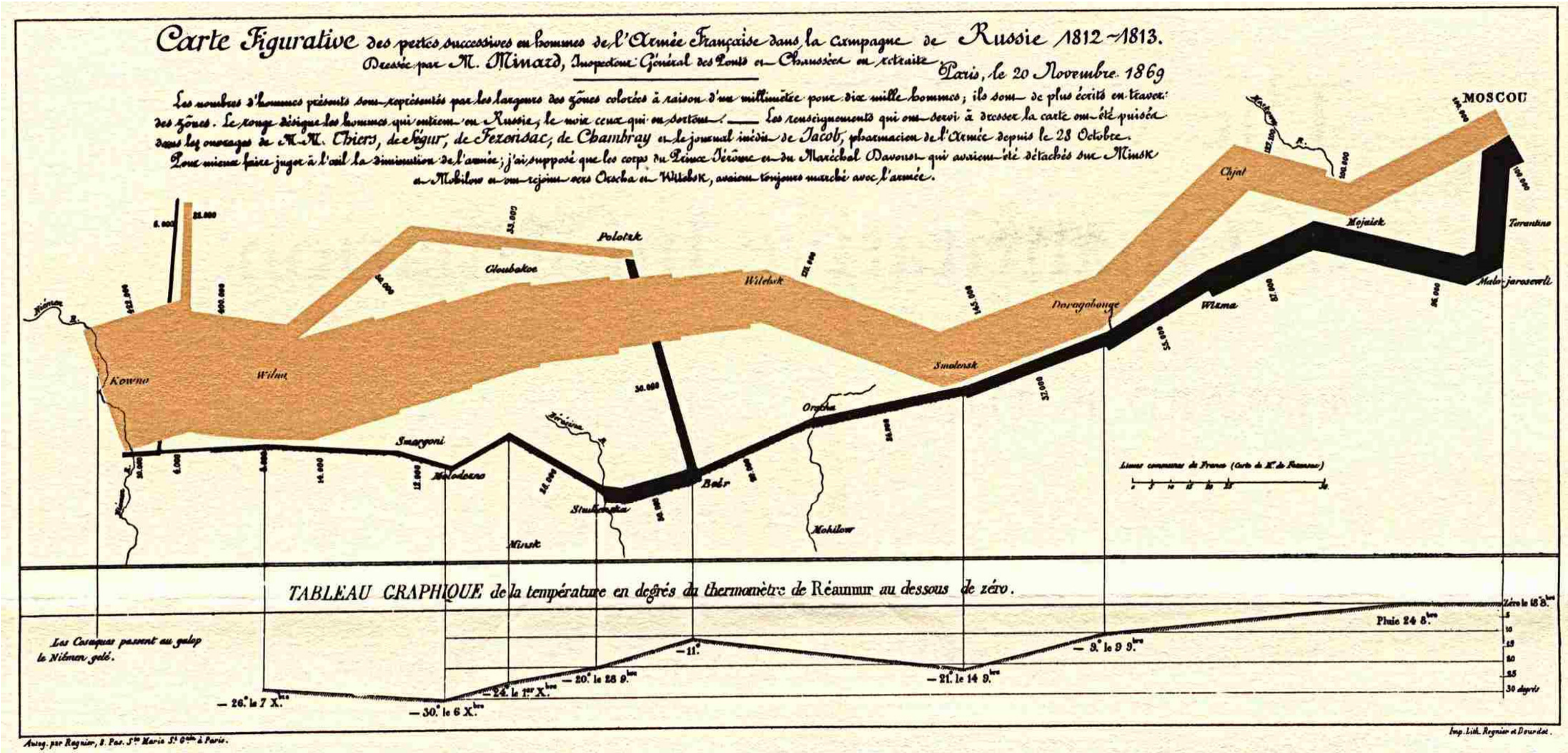
More info here: http://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak

John Snow, 1854

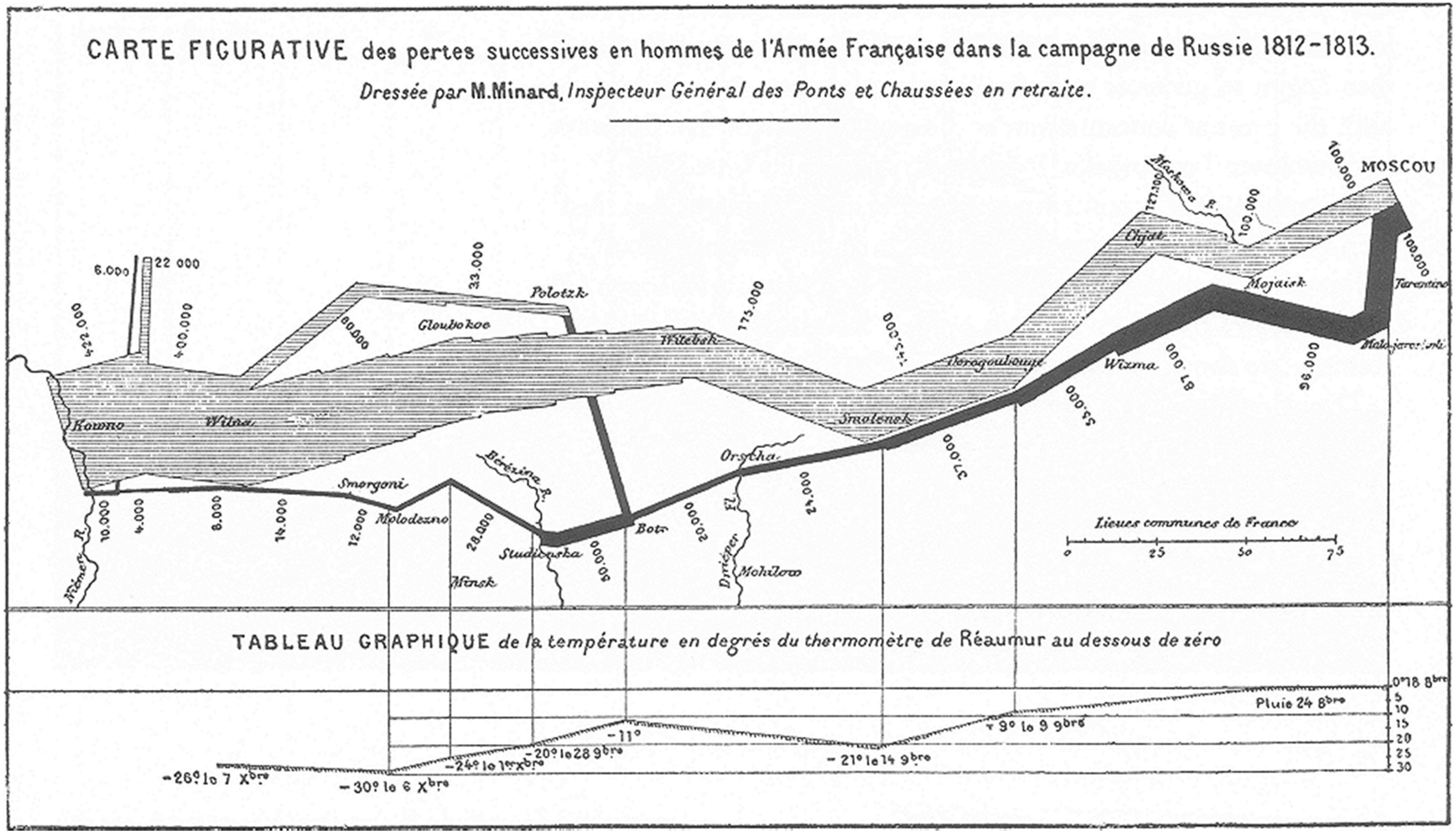
Napoleon's March on Moscow Charles Minard, 1869

Named the best statistical graphic ever drawn (by Edward Tufte)

- Includes: spatial layout linked with stats on: army size, temperature, time
- Tells a story in one overview

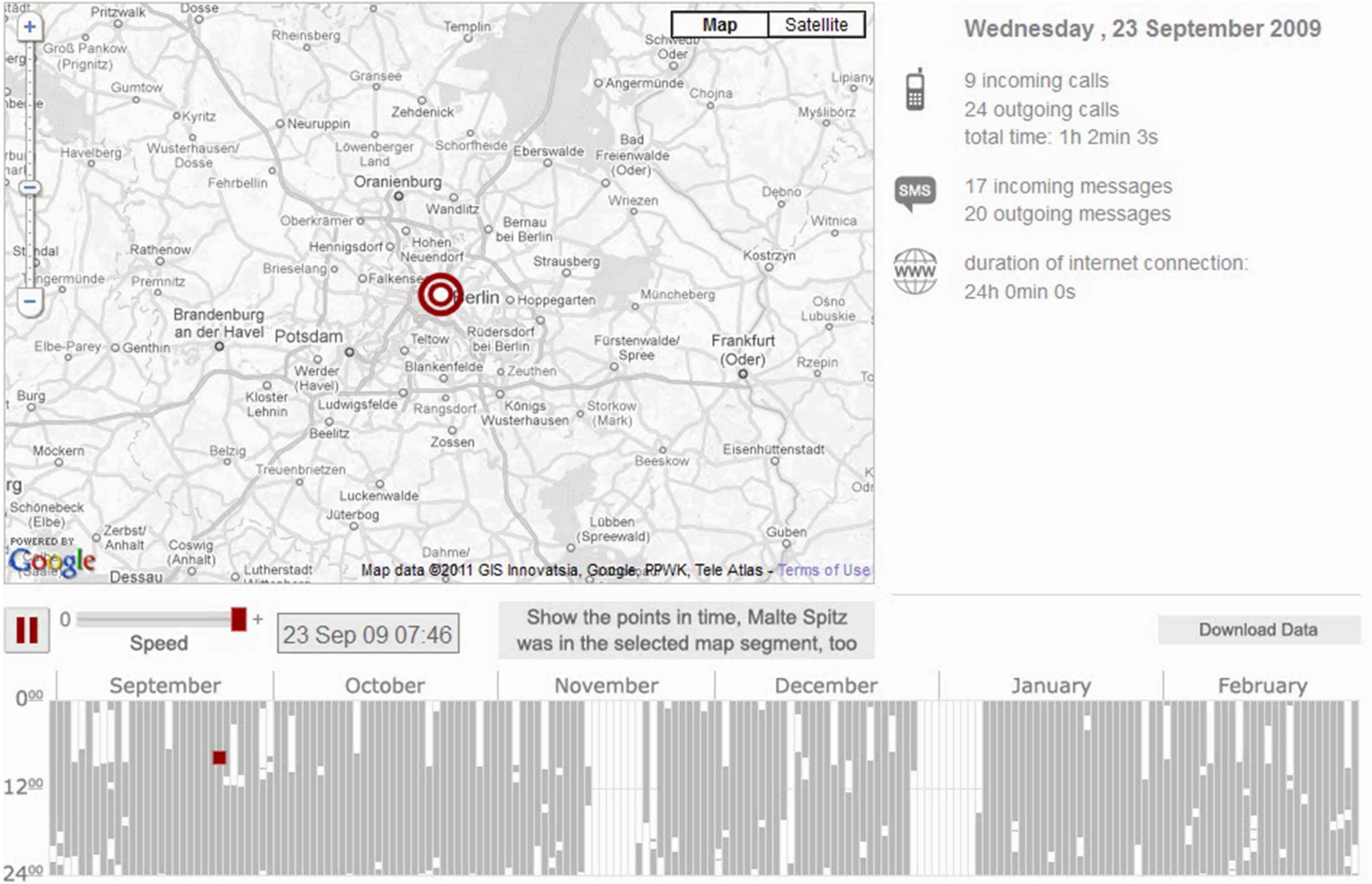


More info: The Visual Display of Quantitative Information (Tufte)


CARTE FIGURATIVE des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813.
Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.
MOSCOU
Mojaisk
Wizma
Smolensk
Orscha
Witebsk
Polotzk
Gloubokoe
Wilna
Kowno
Smorgoni
Molodezno
Studienska
Minsk
Bobr
Mohilow
Dorogobouge
Chjat
Malo-jarosewli
Tarantino
Niémen R.
Bérézina R.
Dniéper Fl.
422.000
400.000
22 000
6.000
60.000
33.000
175.000
145.000
127.000
100.000
100.000
100.000
96.000
87.000
55.000
37.000
24.000
20.000
50.000
28.000
12.000
14.000
8.000
4.000
10.000
Lieues communes de France
0
25
50
75
TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro
0° le 18 8bre
Pluie 24 8bre
+9° le 9 9bre
−21° le 14 9bre
−11°
−20° le 28 9bre
−24° le 1er Xbre
−30° le 6 Xbre
−26° le 7 Xbre
5
10
15
20
25
30

… AND VERY RECENTLY

Mobile Phones…



http://www.zeit.de/datenschutz/malte-spitz-data-retention openpaths.cc

UPDATE: Apple has released update 4.3.3 that will delete this data. Register now to preserve your data.

Register with OpenPaths. 1

Get our desktop app to upload your data. 2

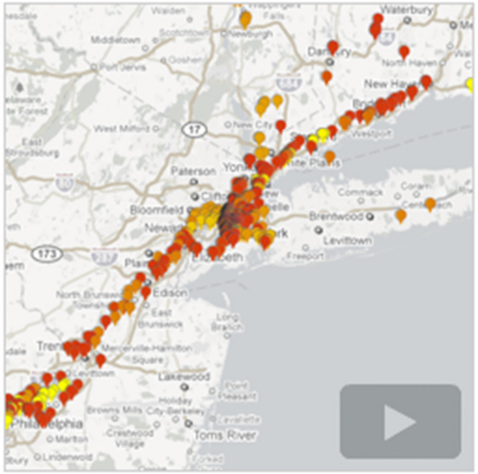
Start visualizing, using, and sharing your location data. 3

Your device.

On April 20, 2011, **researchers announced** that they had discovered a file on Apple iOS devices, like the iPhone and iPad, that show where you've been over the past year.

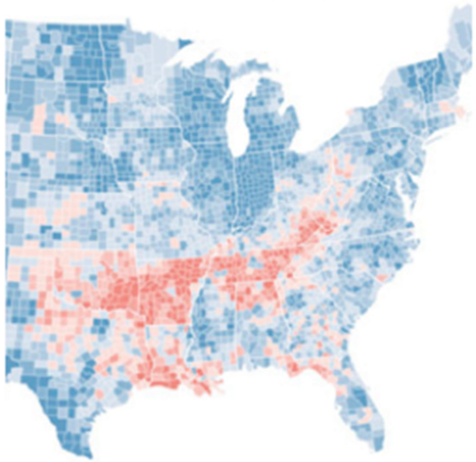
Apple has released an update that will delete this data. Register now to preserve your data.

Your data.



With OpenPaths, you can preserve your iPhone or iPad's location information as well as visualize where you've been. You can even download your data in CSV and JSON format so you can remix it and use it in your own projects.

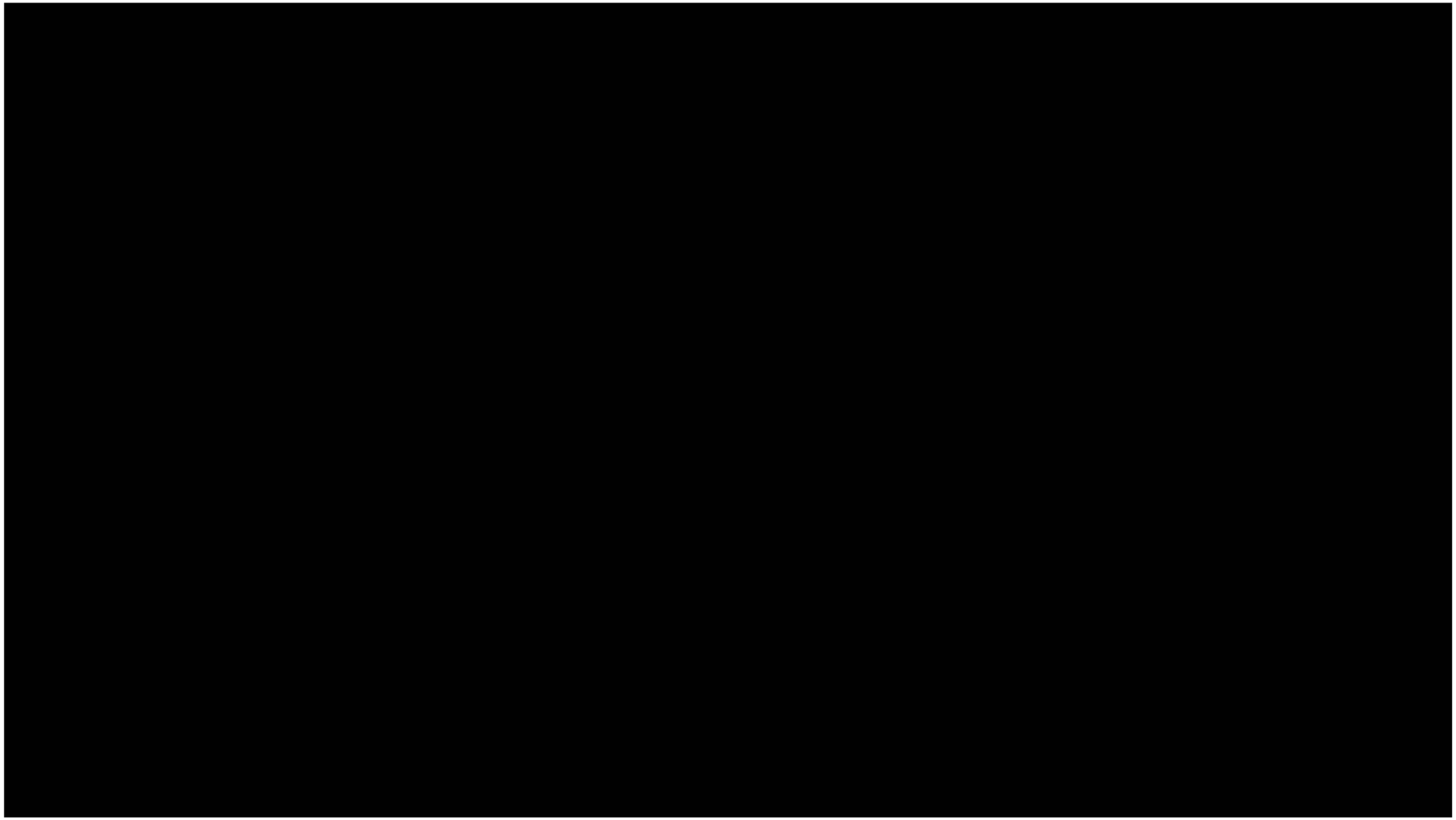
Your decision.



What's more, OpenPaths allows you to securely and anonymously donate your data to researchers who could use it to study mobility, transportation, land use, epidemiology, and overall make the world a better place.

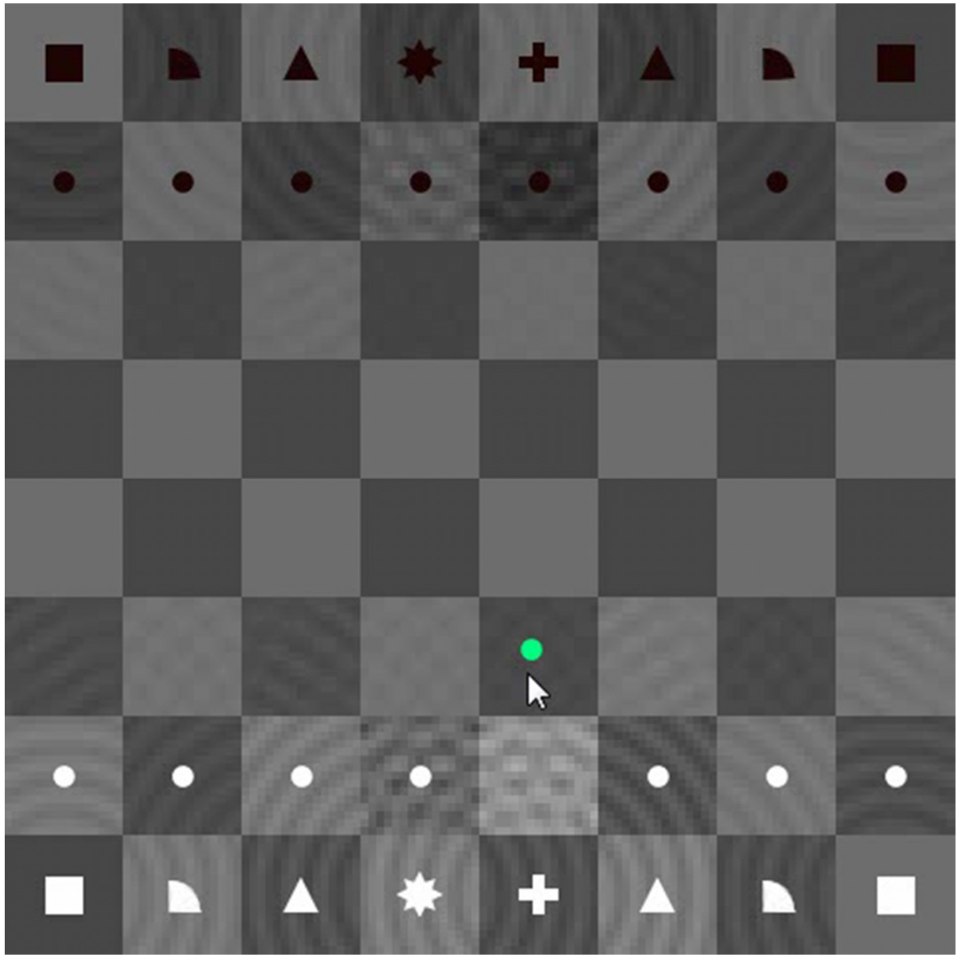
Learn more >>

TrashTrack



Winner of the NSF International Science & Engineering Visualization Challenge!
http://senseable.mit.edu/trashtrack/

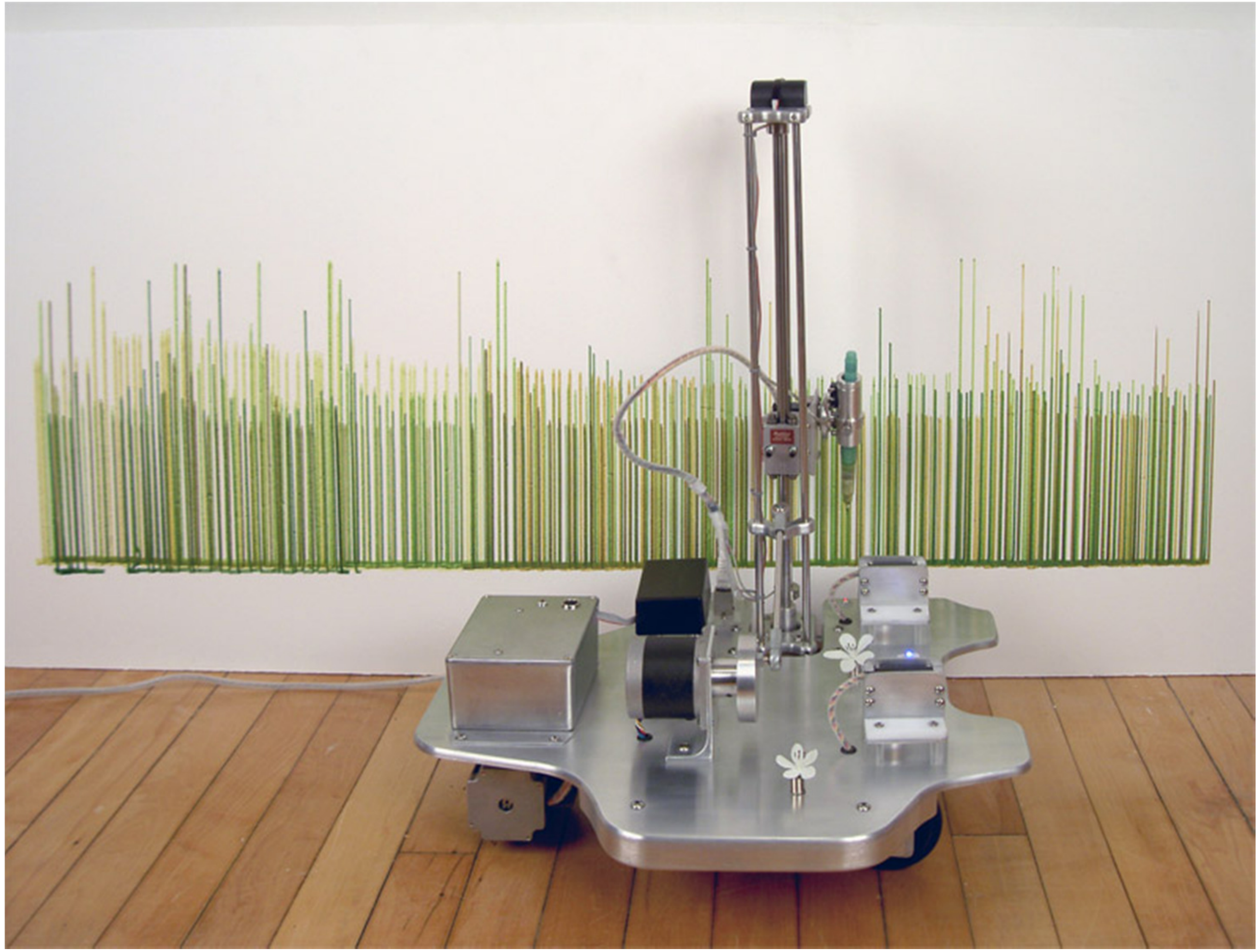
Artificial Intelligence



http://www.turbulence.org/spotlight/thinking/chess.html

HRI

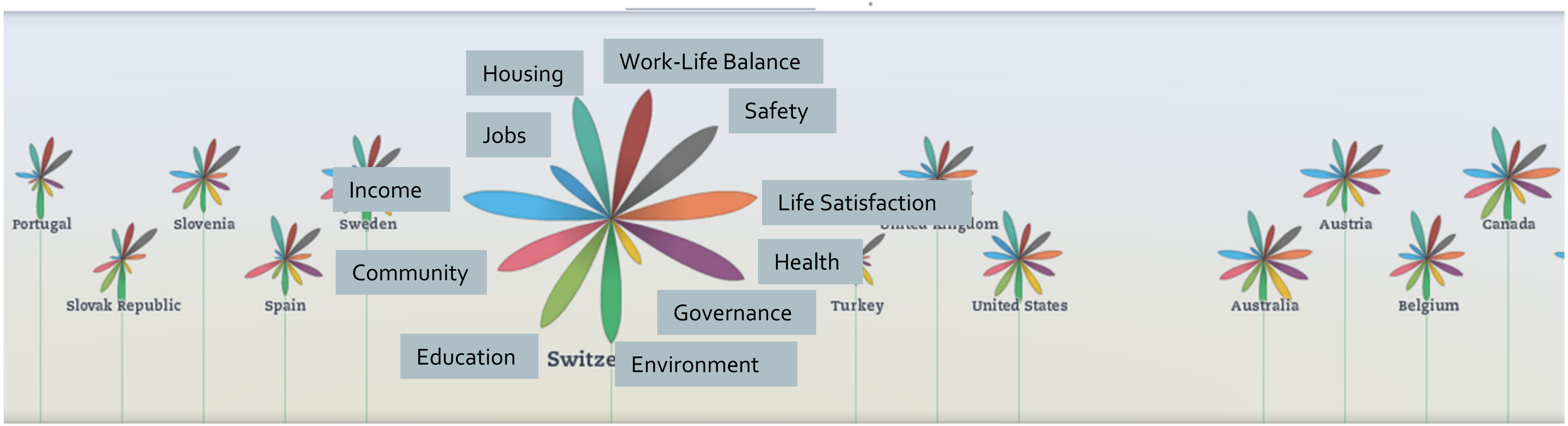
Translator II: Robot visualizing CO_2 emissions



http://www.youtube.com/watch?v=Q3b9LlAYqAw

Open Data

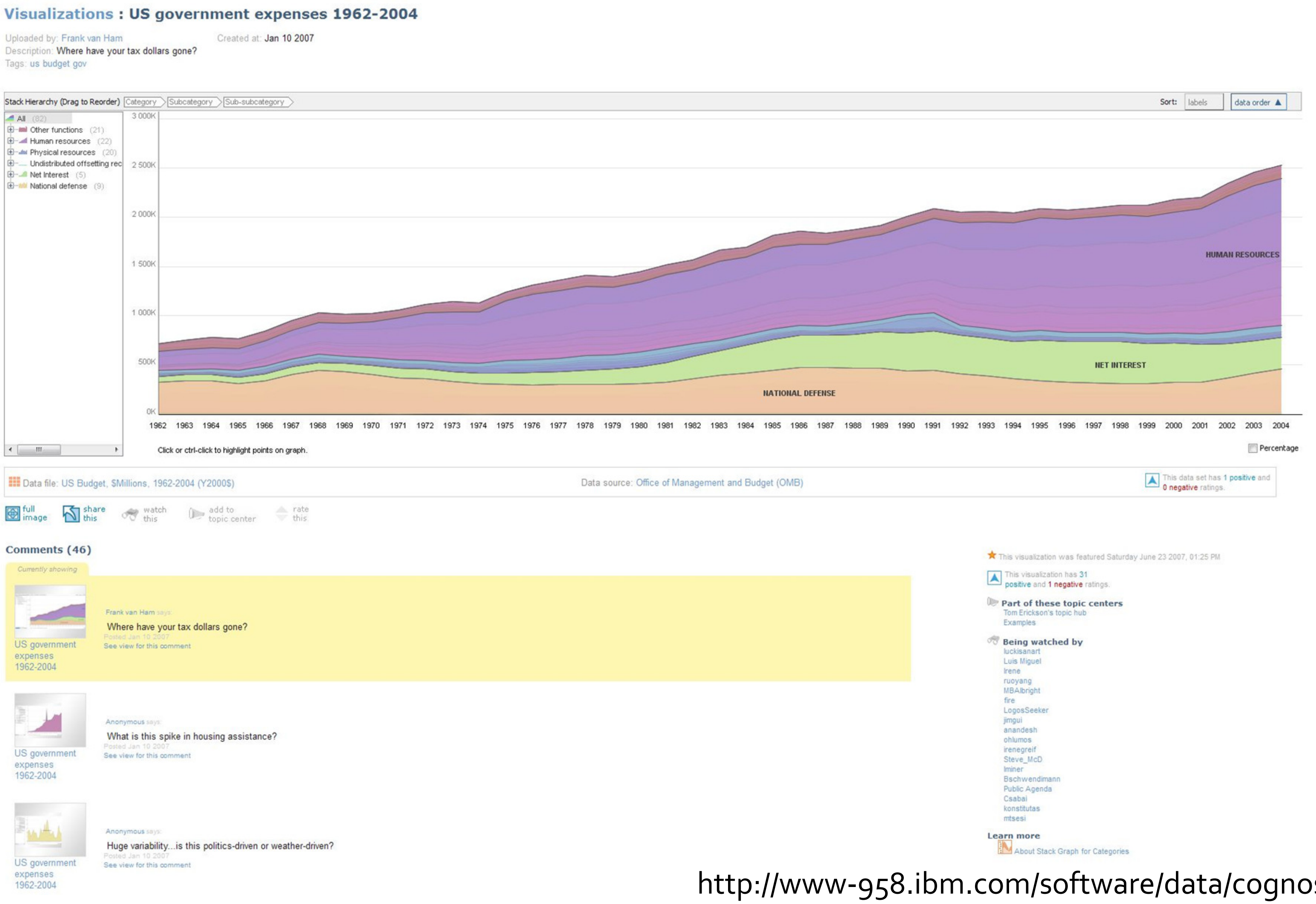
- Movement making government data freely available
- Encourage participation by everyone



OECD Better Life Index: http://www.oecdbetterlifeindex.org/

Many Eyes

- Upload data, create visualizations, discuss
- Distributed asynchronous collaboration

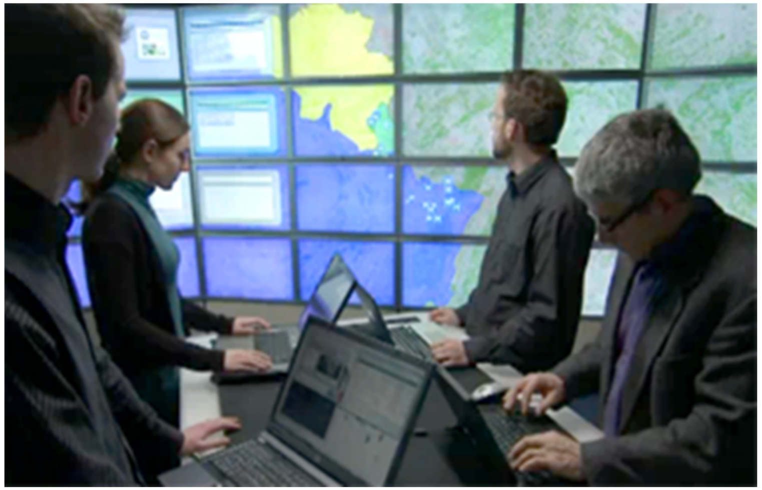


http://www-958.ibm.com/software/data/cognos/manyeyes/

Specific Visualization Environments

Molecular visualisation in the Reality Cube
University of Groningen, NL

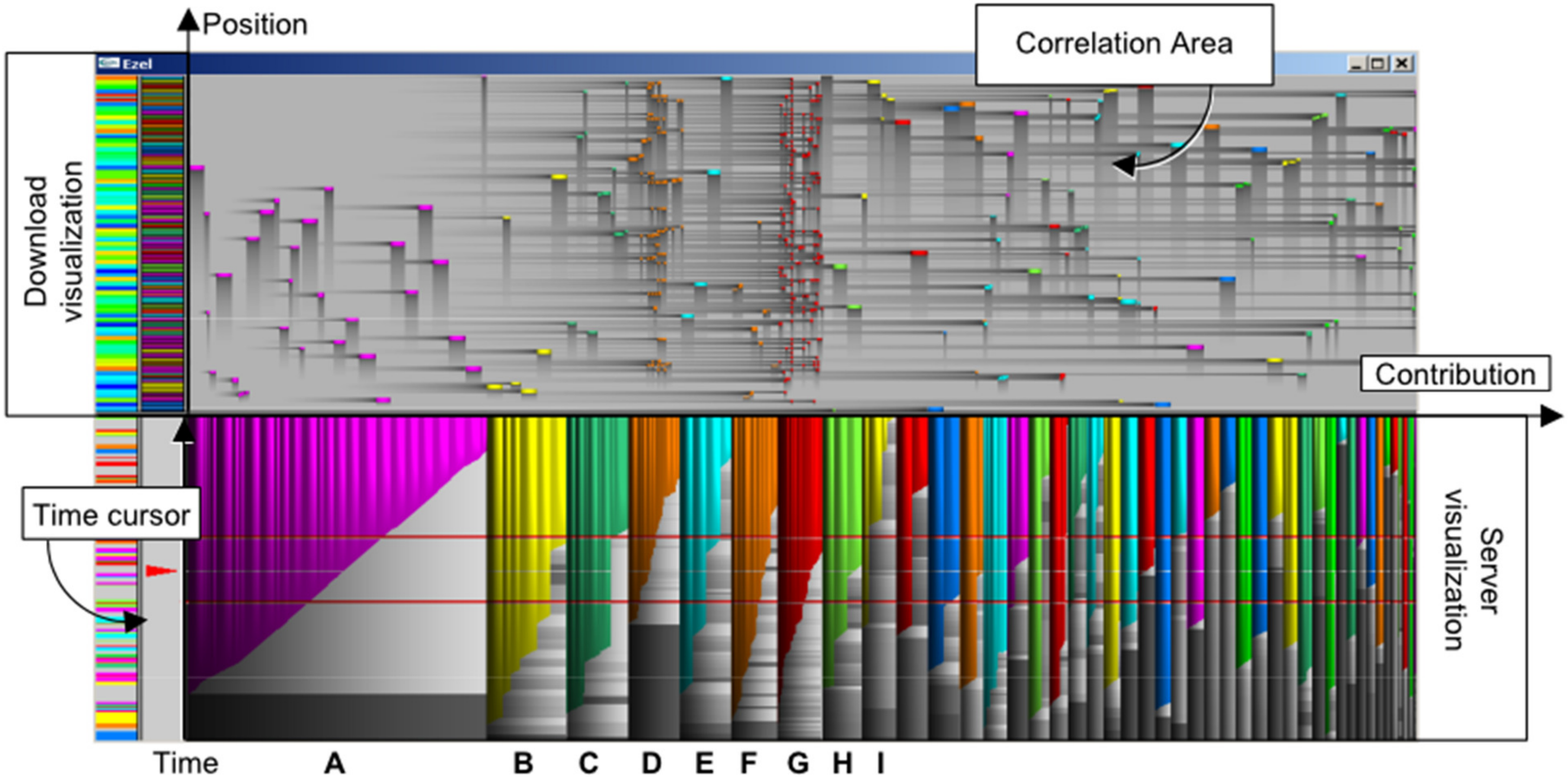
Tabletops for Visualization
University of Calgary



Wall, INRIA

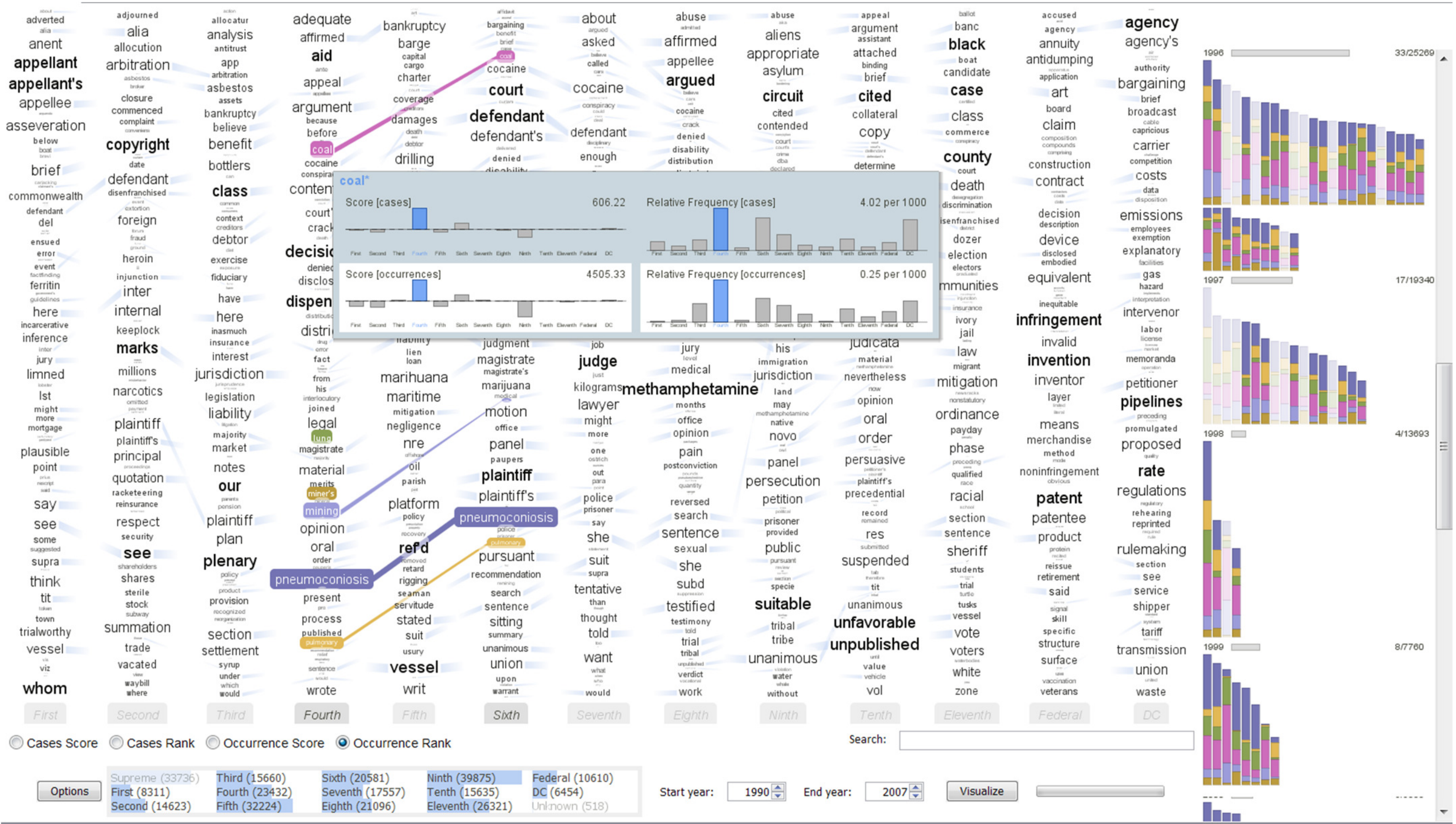
Software Visualization

EZEL: a Visual Tool for Performance Assessment of Peer-to-Peer File-Sharing Networks (Voinea et al., InfoVis, 2004)

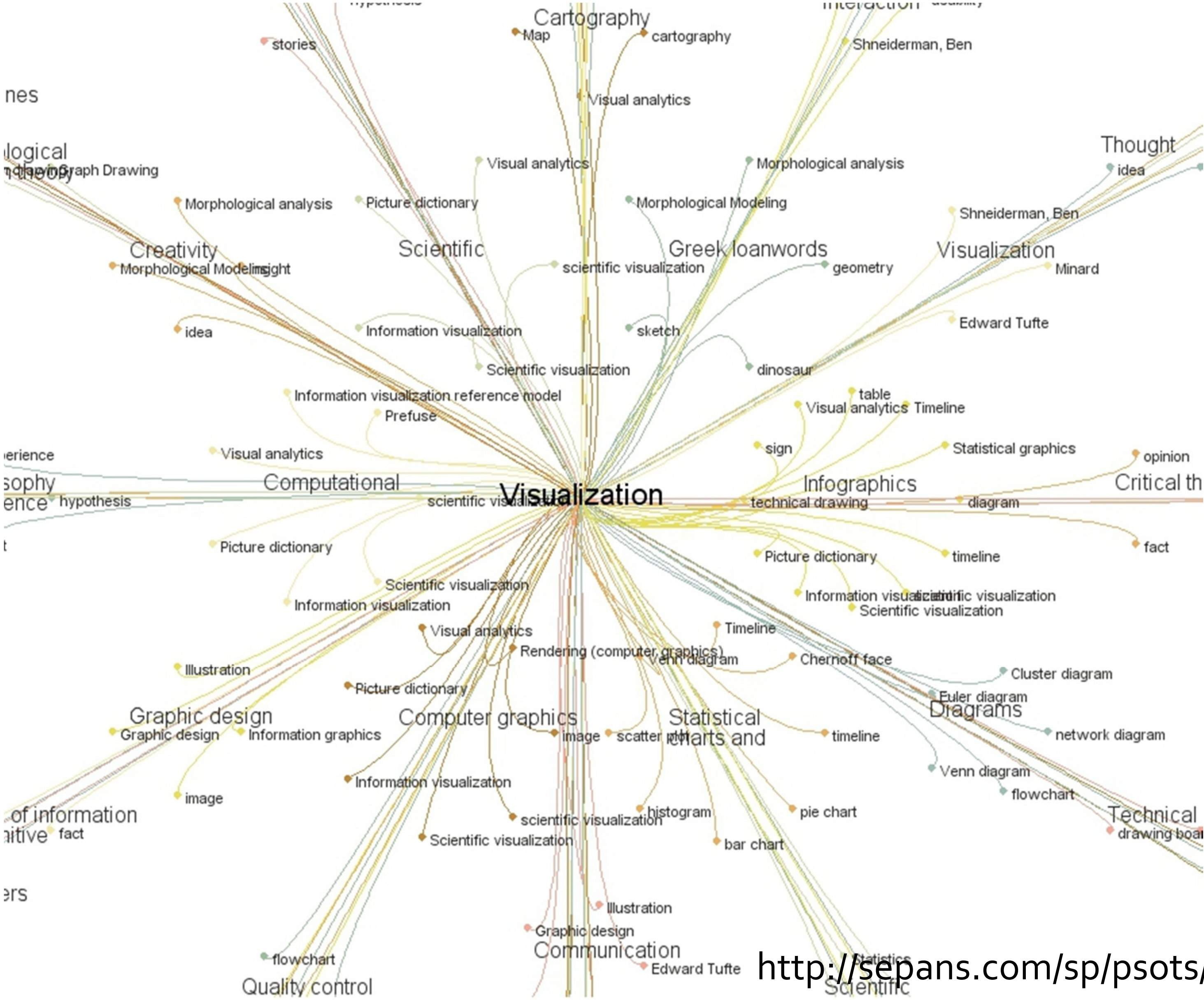


Text Visualization

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

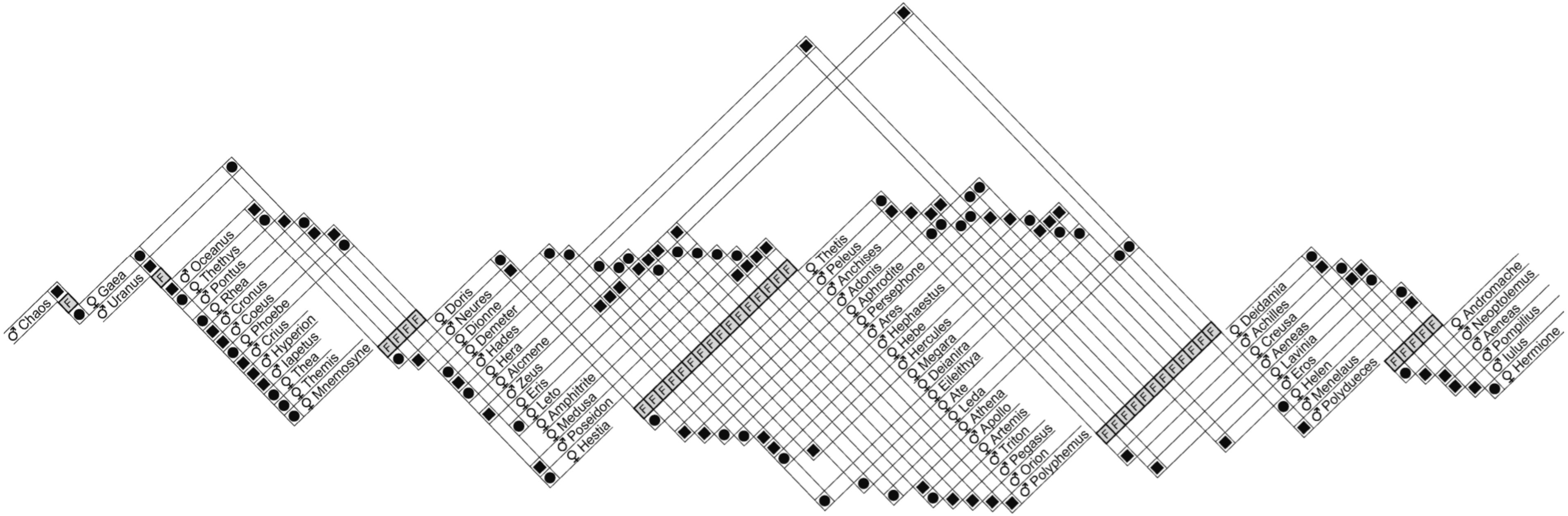


Graphs



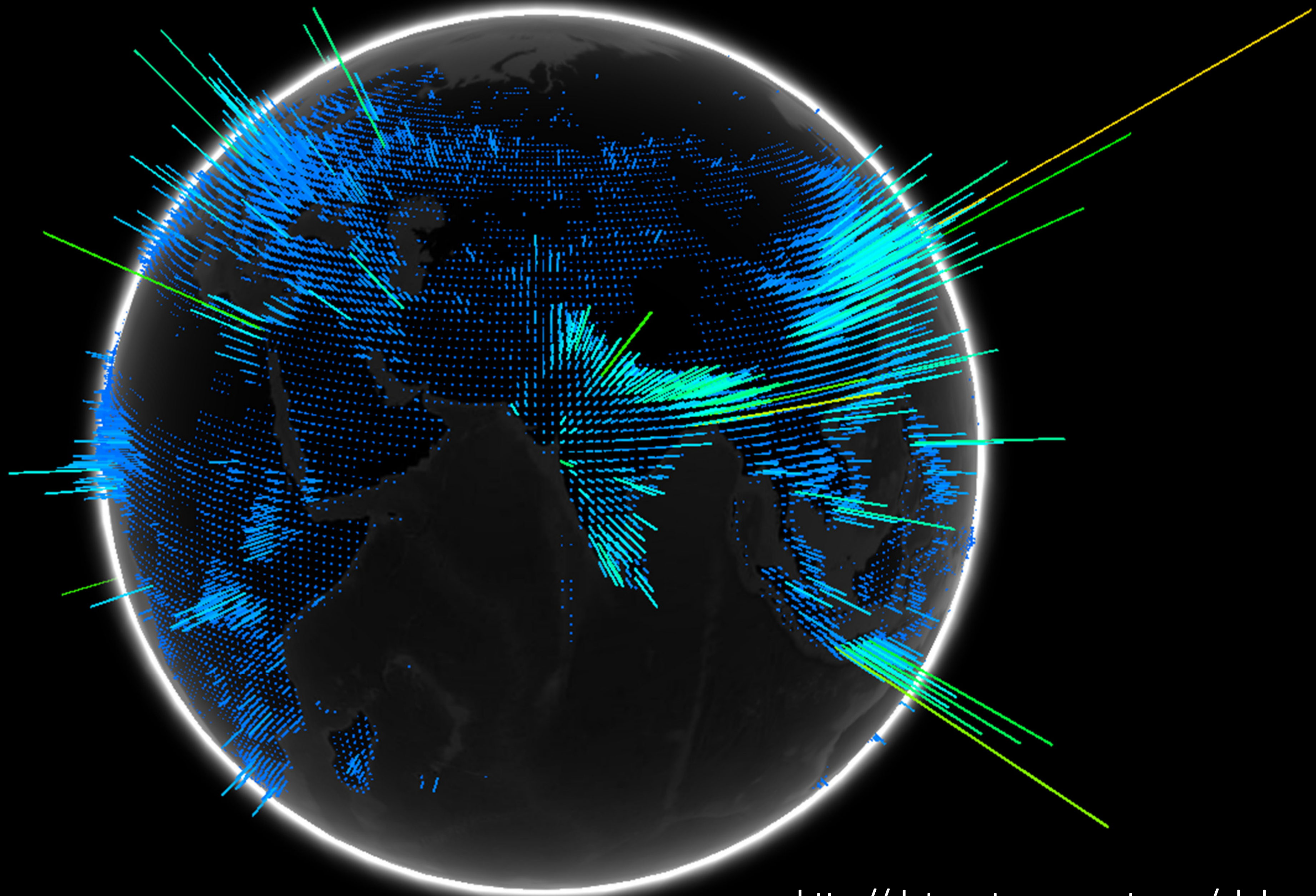
Here Wikipedia
http://sepans.com/sp/psots/wiki_category/

Family Trees



http://www.aviz.fr/geneaquilts/

Geographic Visualization

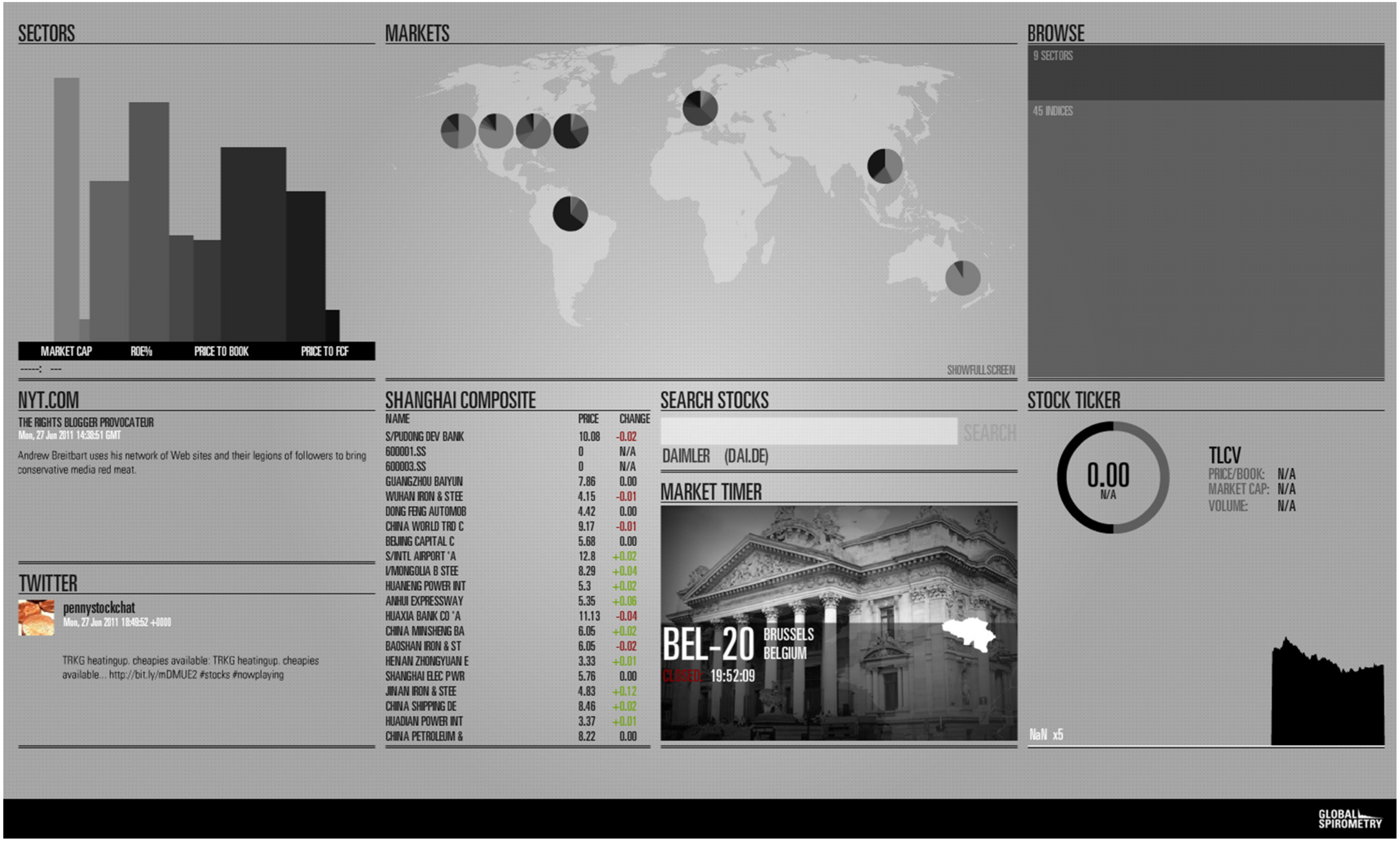


http://data-arts.appspot.com/globe

Weather

http://weatherspark.com/

Data Dashboards



http://globalspirometry.com

Resources for more examples

- Visualization conferences
- Blogs
  - http://infosthetics.com/
  - http://fellinlovewithdata.com/
  - http://eagereyes.org/
  - http://flowingdata.com/
  - http://www.informationisbeautiful.net/
- Books
  - Textbooks
    - Readings in Information Visualization: Using Vision to Think (a bit old now but good intro)
    - Information Visualization (Robert Spence – a light intro, I recommend as a start)
    - Information Visualization Perception for Design (Colin Ware, focused on perception and cognition)
    - Interactive Data Visualization: Foundations, Techniques, and Applications (Ward et al. – most recent)
  - Examples
    - Beautiful Data (McCandless)
    - Now You See it (Few)
    - Tufte Books: Visual Display of Quantitative Information (and others)
    - … (many more, ask me for details)

It is difficult to create

CREATE VISUALIZATIONS

good


Cem

AND ALL WERE IN THE WRONG!

Good Presentation Should Be Susceptible to Only One Interpretation

It was six men of Indostan
  To learning much inclined,
Who went to see the Elephant
  (Though all of them were blind.)
That each by observation
  Might satisfy his mind.

The First (side) "Is very like a wall!"

The Second (tusk) "Is very like a spear!"

The Third (trunk) "Is very like a snake!"

The Fourth (knee) "Is very like a tree!"

The Fifth (ear) "Is mighty like a fan!"

The Sixth (tail) "Is very like a rope!"

And so these men of Indostan
  Disputed loud and long,
Each in his own opinion
  Exceeding stiff and strong
Though each was partly in the right,
  And all were in the wrong!

From John Godfrey Saxe, "The Blind Men and the Elephant", *Clever Stories of Many Nations Rendered in Rhime*, 1865.

Via Brinton, Graphic Presentation, 1939

What is a representation?

- A representation is
  - a formal system or mapping by which the information can be specified (D. Marr)
  - a sign system in that it stands for something other than its self.
- for example: the number thirty-four

34	100010	XXXIV
decimal	binary	roman

Presentation

- different representations reveal different aspects of the information

    decimal: counting & information about powers of 10,

    binary: counting & information about powers of 2,

    roman: counting & adding and subtracting

- presentation

    how the representation is placed or organized on the screen

    34, **34**, <u>34</u>

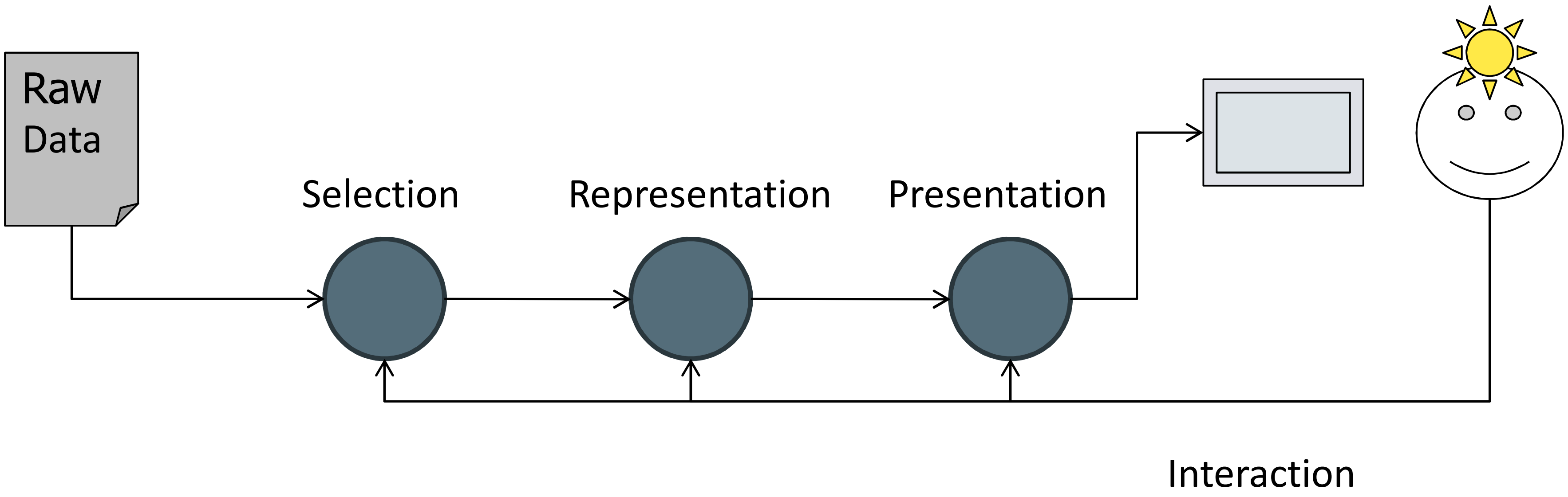
Principles of Graphical Excellence

- Well-designed presentation of interesting data – a matter of *substance*, *statistics*, *design*
- Complex ideas communicated with clarity, precision, efficiency
- Gives the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
- Involves almost always multiple variables
- Tell the truth about the data

Or a bit more simply…

- Solving a problem simply means representing it so as to make the solution transparent … (Simon, 1981)
- Good representations:
  - allow people to find relevant information
    - information may be present but hard to find
  - allow people to compute desired conclusions
    - computations may be difficult or “for free” depending on representations

How do we arrive at a visualization?

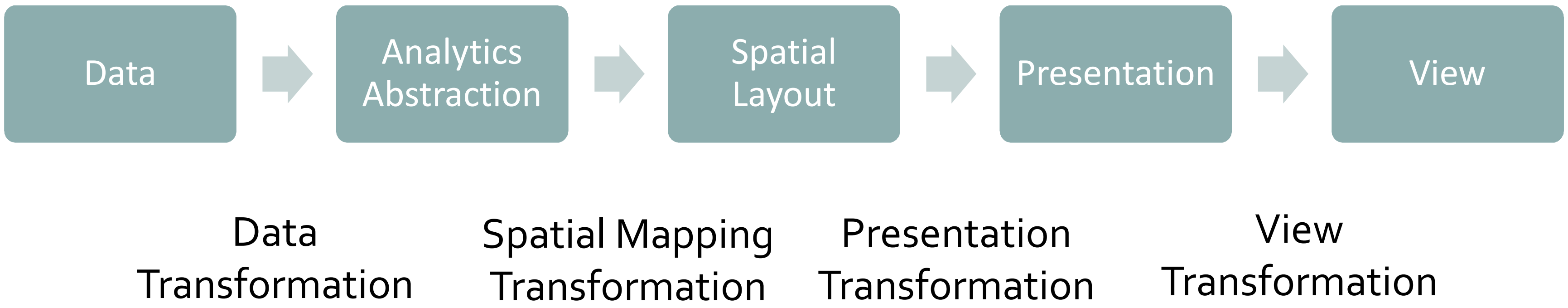


The Visualization Pipeline

From [Spence, 2000]

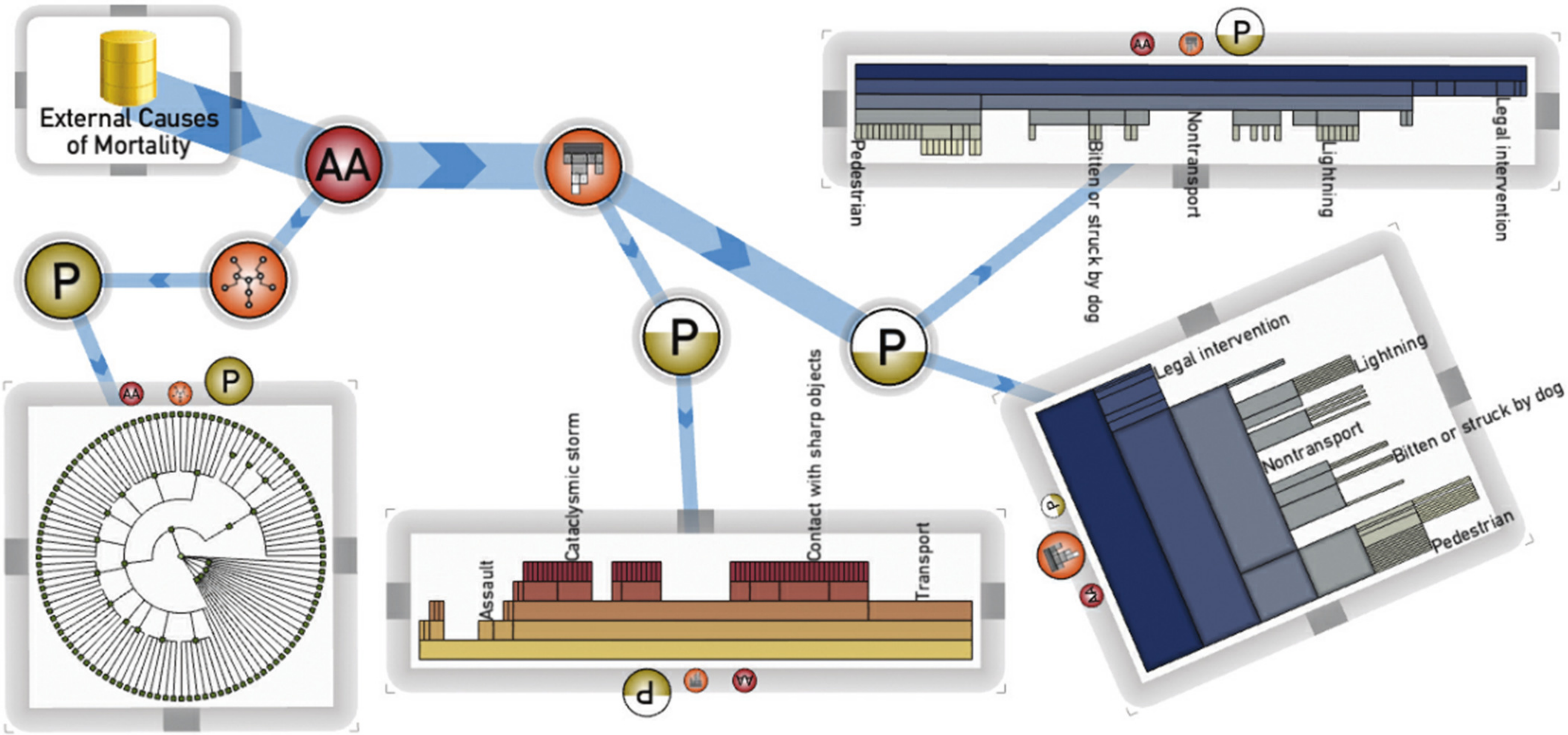
Visualization Reference Model

Also a visualization pipeline a bit expanded



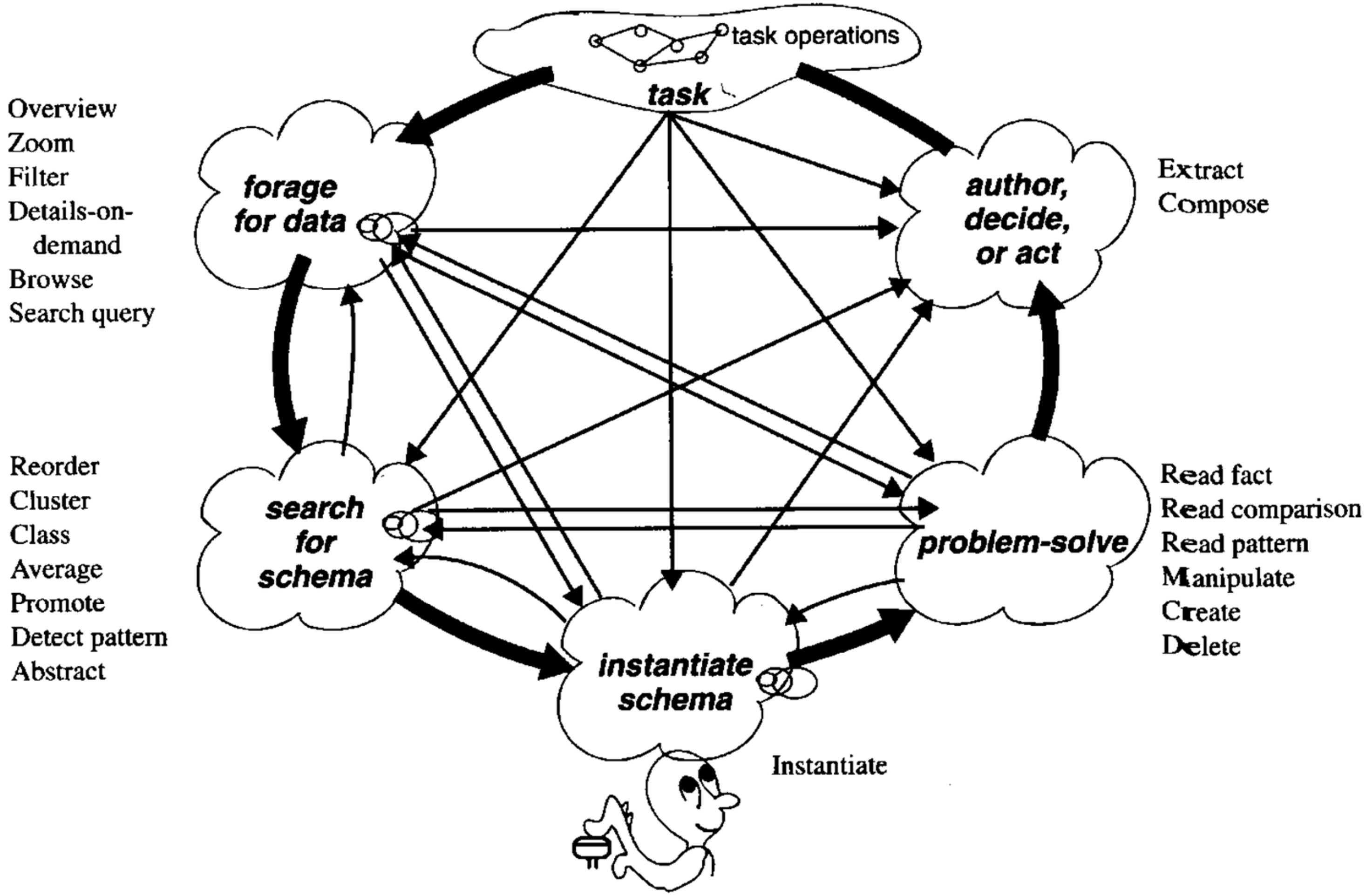
From [Card et al., Readings in Information Visualization]

Visualization pipeline in an image



[Tobiasz et al., 2009]

Knowledge Crystallization Cycle



Working with visualizations in NOT a linear process

[Card et al., 1999]

Ben Fry's "Seven Stages of Visualizing Data"

Visualizing complex data sets often requires insights from diverse fields of knowledge, such as statistics, data mining, graphic design, et cetera. Ben Fry suggests a seven stage design process, reconciling all stages into a single process.

1. Aquire

Obtain the data, whether from an Excel document, an XML feed, et cetera.

2. Parse

Data will not always be organized ideally for visualizing it. Give your data structure by ordering it into categories.

3. Filter

Be careful to prevent information overload, remove all but the data of interest

4. Mine

Apply methods from statistics or data mining as a way to find patterns or meaning in the data.

5. Represent

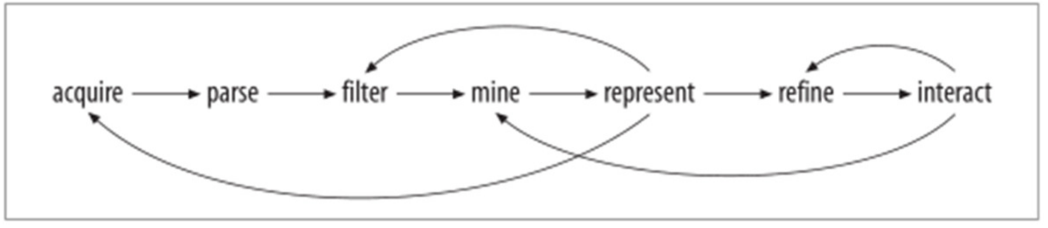
Choose a basic visual model to visual the data. (see Chapter 5)

6. Refine

Improve the basic representation to make it more clear and more visually engaging. (see Chapter 6)

7. Interact

Add methods for manipulating the data. Allows users to control what they see or even possibly how they see it.



Visualization Pitfalls

- **Data Design**
  - Select the right data dimensions
  - Pitfall: Display irrelevant data relationships
- **Visual Design**
  - Consider perceptual capabilities
  - Pitfall: Difficult to interpret, lead users to misinterpretation
- **Interface Design**
  - Mode of interaction with visual and data appropriate
  - Pitfall: Poor interaction negates benefits from data and visual design
- **Understanding Target Users**
  - Visualization design accounts for stakeholder needs and characteristics
  - Pitfall: Mistaken assumptions of cultural norms or user abilities leads to misinterpretation

Pitfalls

- Selecting the wrong data
- Selecting the wrong data structure
- Filtering out important data
- Failed understanding of the types of things that need to be shown
- Choosing the wrong representation
- Choosing the wrong presentation format
- Inappropriate interactions provided to explore the data

Exercise

- Think about the cognitive tasks that people would perform with the aid of this data.
- What would a person want to know about this data set?
- Create a visualization of this data that would assist those tasks

	mary	joe	paul	sue	betty	lyn	jack	fred
mary	0	1	2	1	5	3	5	4
joe	2	0	1	5	1	4	1	3
paul	2	2	0	2	5	2	2	2
sue	1	5	1	0	1	1	1	1
betty	5	1	5	1	0	2	2	1
lyn	1	4	1	2	1	0	2	2
jack	5	1	2	1	3	1	0	1
fred	2	3	1	2	1	2	1	0

0 self
1 knows
2 likes
3 dislikes
4 loves
5 does not know

Critique

- End first 1.5h

Recap

- So far you
  - learned what information visualization is
  - learned about the advantages of visualization
  - saw a number of examples (historical and new)
  - tried to create your own first visualization from a dataset

- Next
  - you will get to know your data
  - you will learn about the basic components of visualization
  - try another example

Data

- Data is the foundation of any visualization
- The visualization designer needs to understand
  - the data properties
  - know what meta-data is available
  - know what people want from the data

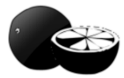
Nominal, Ordinal and Quantitative

- Nominal (labels)
  - Fruits: apples, oranges
- Ordered
  - Quality of meat: grade A, AA, AAA
  - Can be counted and ordered, but not measured
- Quantitative: Interval
  - no clear zero (or arbitrary)
  - e.g. dates, longitude, latitude
  - usually compare differences (intervals)
- Quantitative: Ratio
  - meaningful origin (zero)
  - physical measurements (temperature, mass, length)
  - counts and amounts

S.S. Stevens, On the theory of scales of measurements, 1946

Nominal, Ordinal and Quantitative

- Nominal (labels)
  - Operations: =, ≠

 ≠ 

- Ordered
  - Operations: =, ≠, <, >

 > 

- Quantitative: Interval
  - Operations: =, ≠, <, >, -, +
  - Can measure distances or spans

[1989 – 1999] + [2002 – 2012]

- Quantitative: Ratio
  - Operationrs: =, ≠, <, >, - , +, •, ÷
  - Can measure ratios or proportions

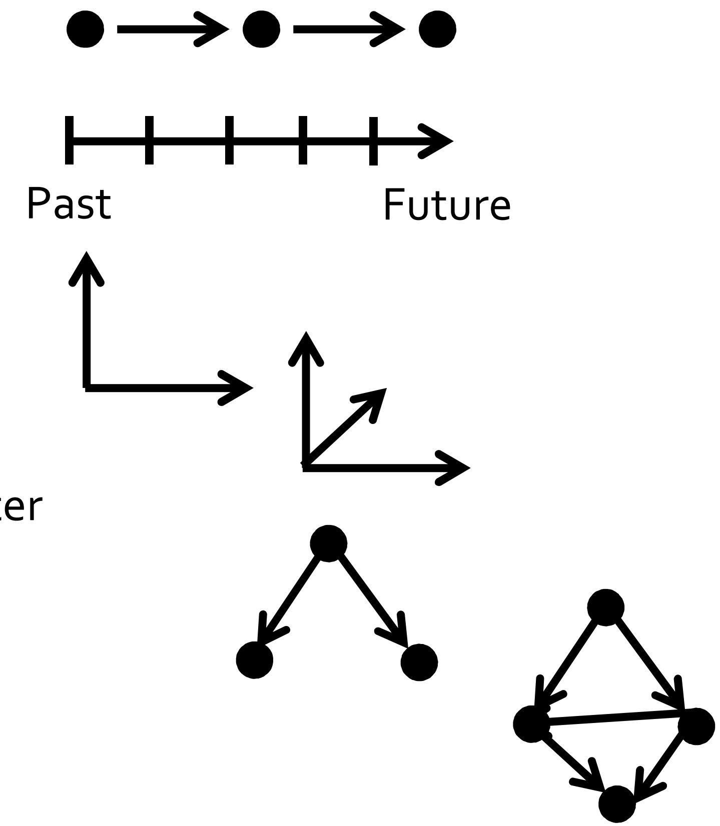
10kg / 5kg

S.S. Stevens, On the theory of scales of measurements, 1946

Data-Type Taxonomy

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

vis examples later



Shneiderman: The Eyes Have It

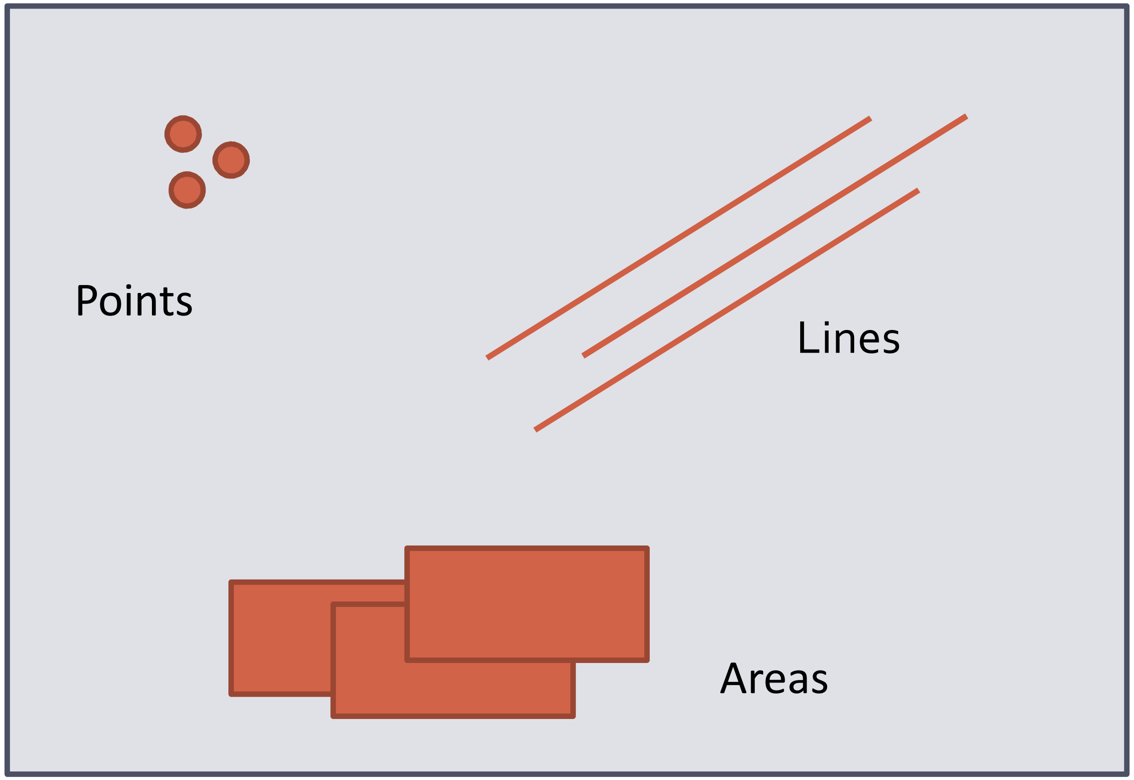
Why is this important?

- Nominal, ordinal, and quantitative data are best expressed in different ways visually
- Data types often have inherent tasks
  - temporal data (comparison of events)
  - trees (understand parent-child relationships)
  - ...

- But:
  - any data type (1D, 2D,...) can be expressed in a multitude of ways!

Visualization's Main Building Blocks

Marks which represent:

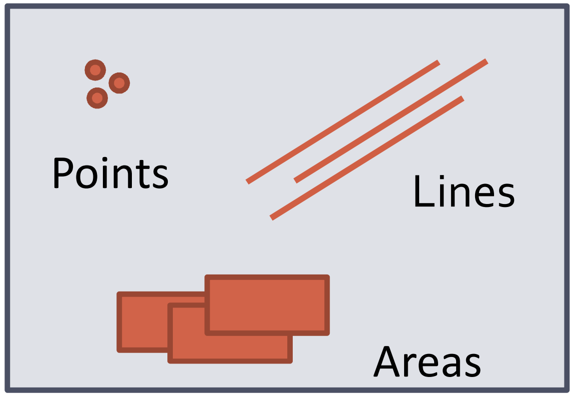


From Semiology of Graphics (Bertin)

The following slides on the topic adapted from Sheelagh Carpendale

Points

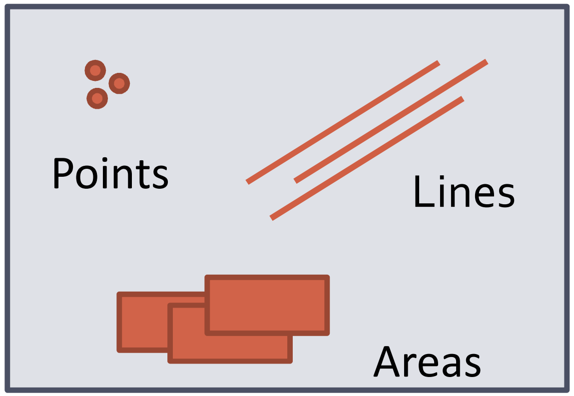
- "A point represents a location on the plane that has **no theoretical length or area**. This signification is independent of the size and character of the mark which renders it visible."
- a location
- marks that indicate points can vary in all visual variables



From Semiology of Graphics (Bertin)

Lines

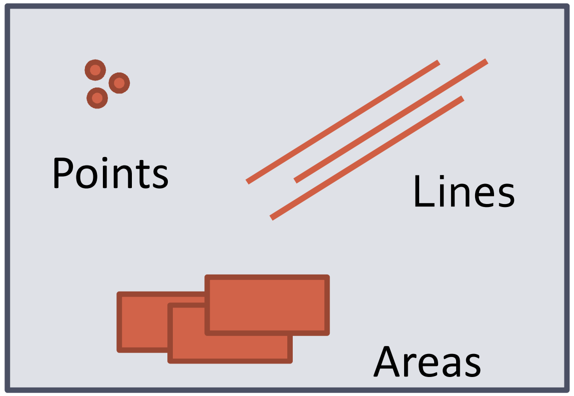
- “A line signifies a phenomenon on the plane which has **measurable length but no area**. This signification is independent of the width and characteristics of the mark which renders it visible.”
- a boundary, a route, a connection



From Semiology of Graphics (Bertin)

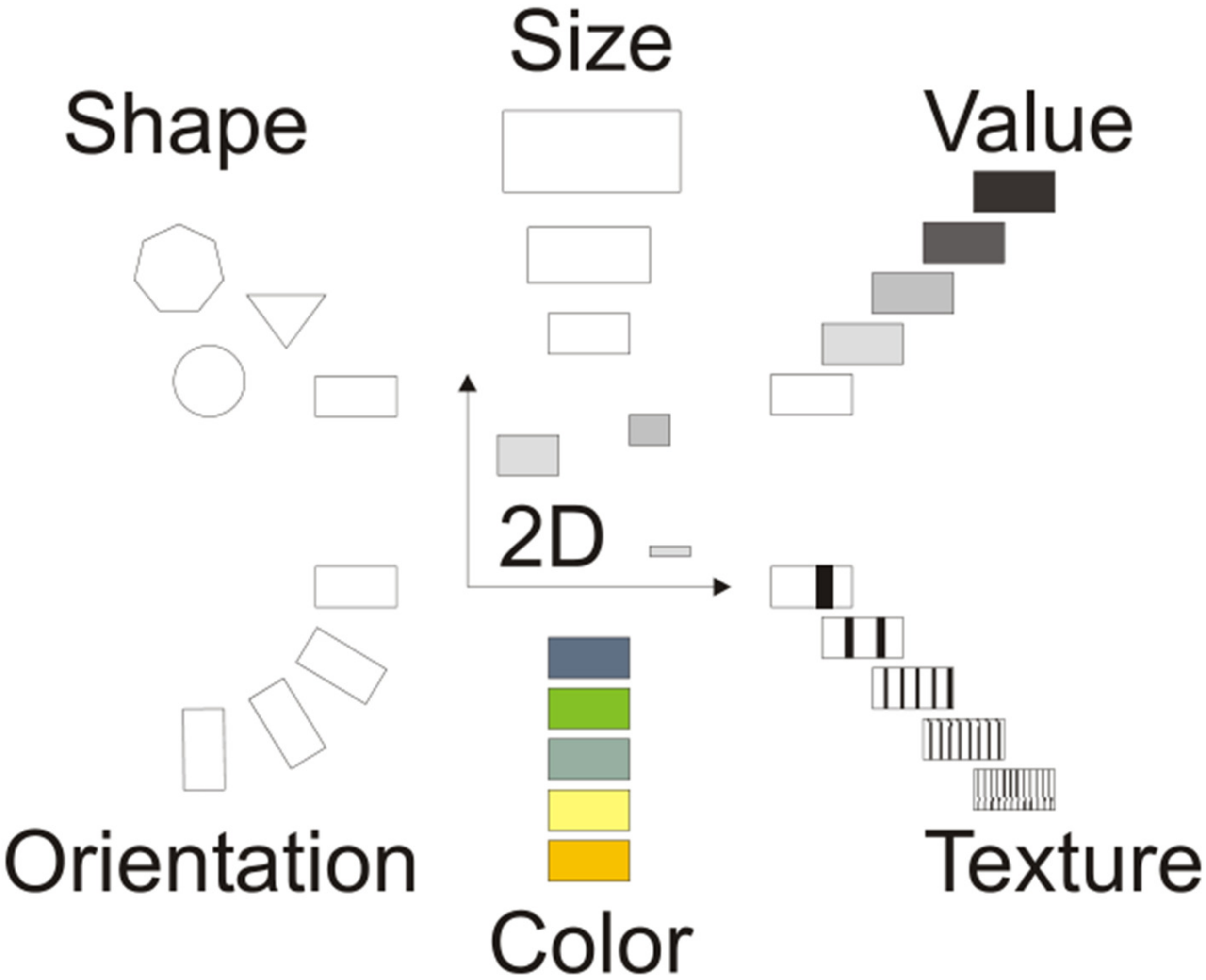
Areas

- "An area signifies something on the plane that **has measurable size**.
  This signification applies to the entire area covered by the visible mark."
- an area can change in position but not in size, shape or orientation without making the area itself have a different meaning



From Semiology of Graphics (Bertin)

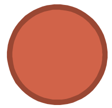
Visual Variables Applicable to Marks



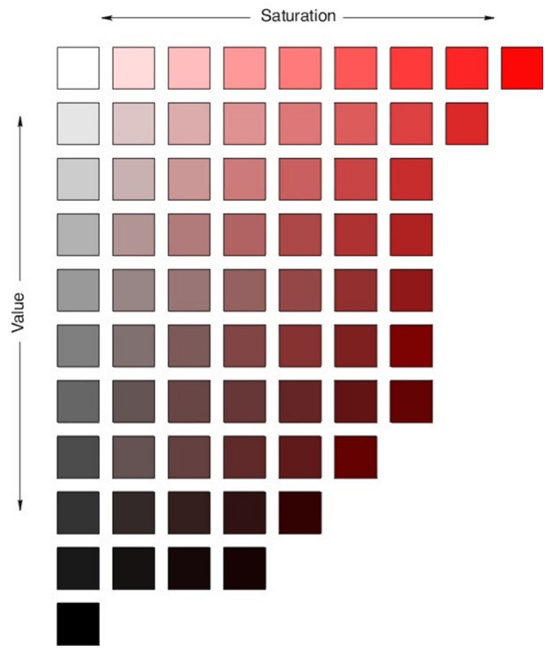
From Semiology of Graphics (Bertin)

Additional Variables for Computers

- **motion**
  - direction, acceleration, speed, frequency, onset, ‘personality’



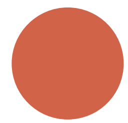
- **saturation**
  - colour as Bertin uses largely refers to hue, saturation != value



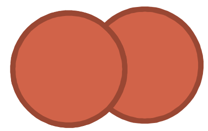
Extending those from Semiology of Graphics (Bertin)

Additional Variables for Computers

- **flicker**
  - frequency, rhythm, appearance

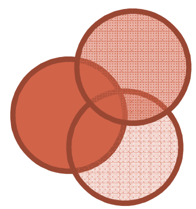


- **depth?** 'quasi' 3D
  - depth, occlusion, aerial perspective, binocular disparity



- **Illumination**

- **transparency**



From Semiology of Graphics (Bertin)

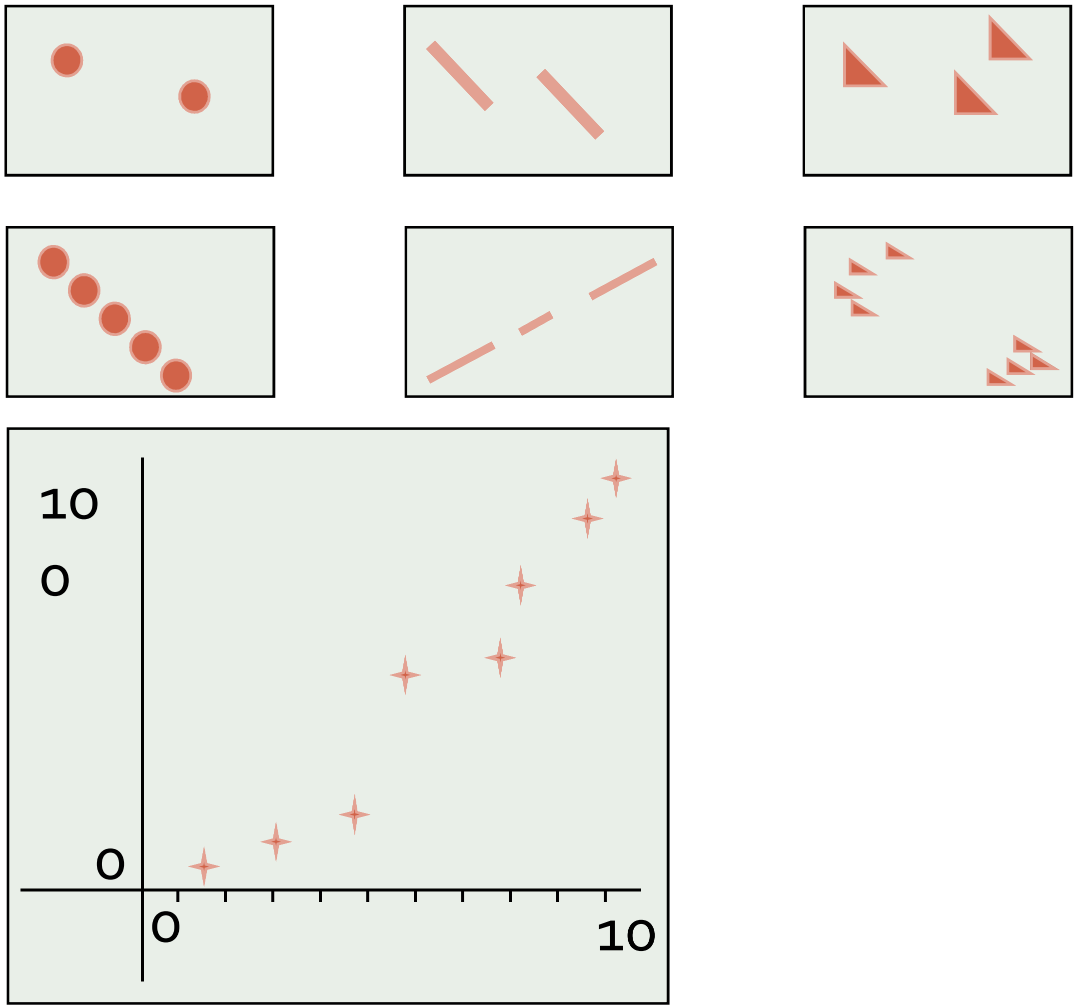
Characteristics of Visual Variables

- Selective:
  Is a change in this variable enough to allow us to select it from a group?
- Associative:
  Is a change in this variable enough to allow us to perceive them as a group?
- Quantitative:
  Is there a numerical reading obtainable from changes in this variable?
- Order:
  Are changes in this variable perceived as ordered?
- Length:
  Across how many changes in this variable are distinctions possible?

From Semiology of Graphics (Bertin)

Visual Variable: Position

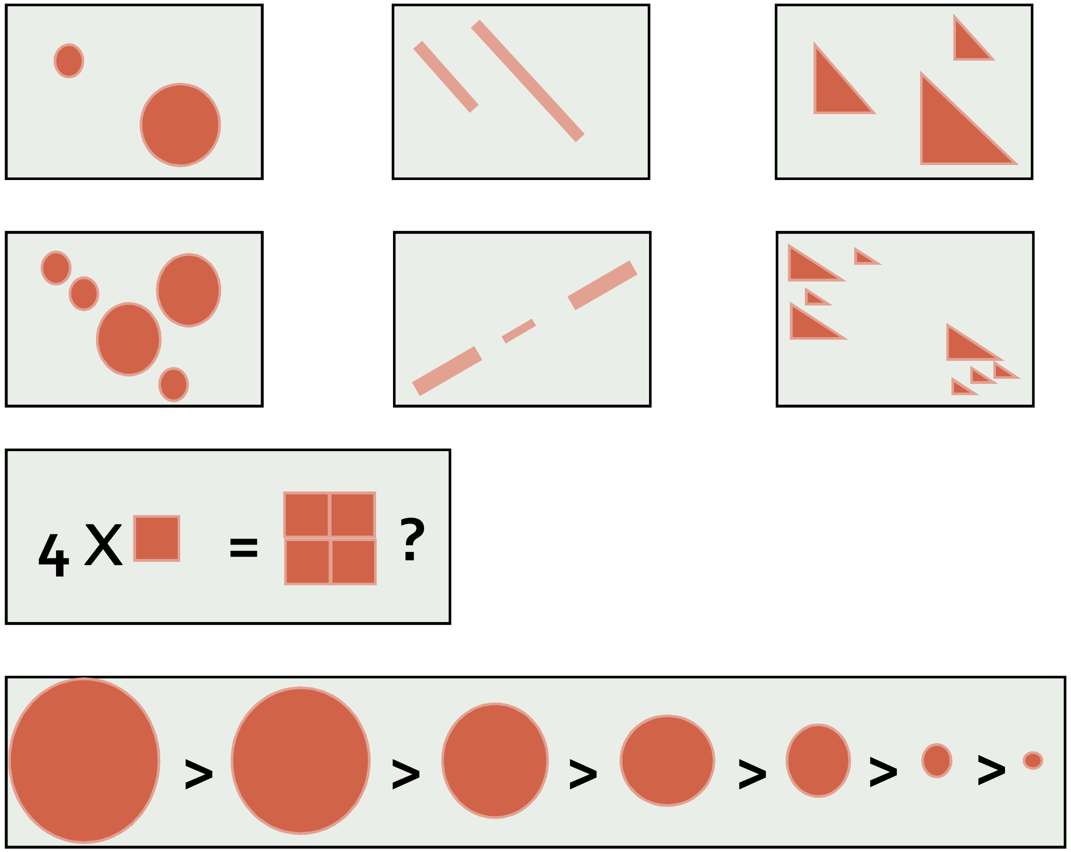
- selective
- associative
- quantitative
- order
- length



From Semiology of Graphics (Bertin)

Visual Variable: Size

- selective
- associative
- quantitative
- order
- Length
  - theoretically infinite but practically limited
  - association and selection ~ 5 and distinction ~ 20



Size

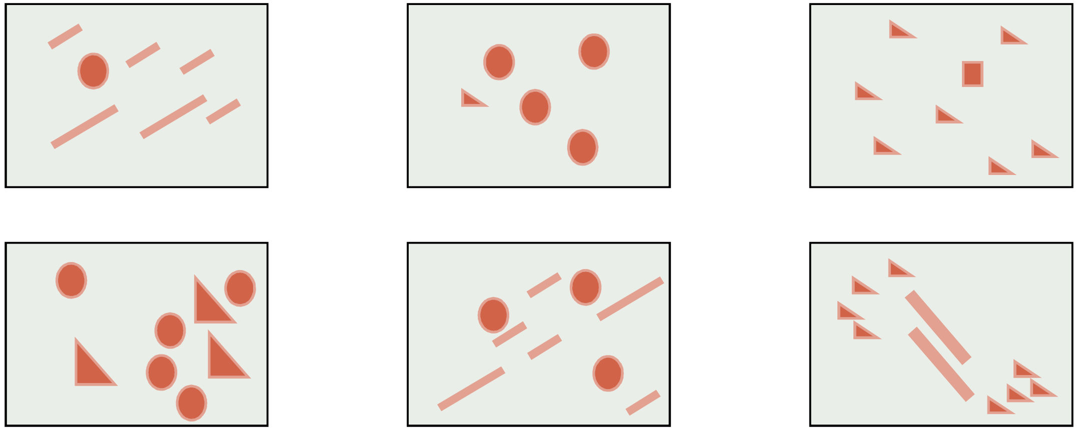
points

lines

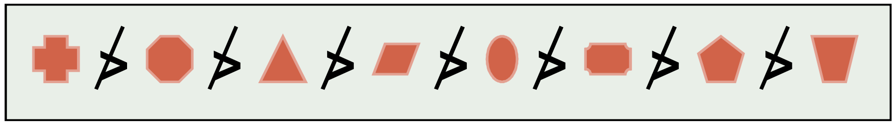
areas

Visual Variable: Shape

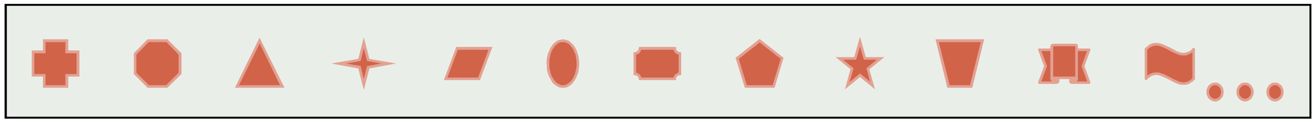
- selective
- associative



- quantitative

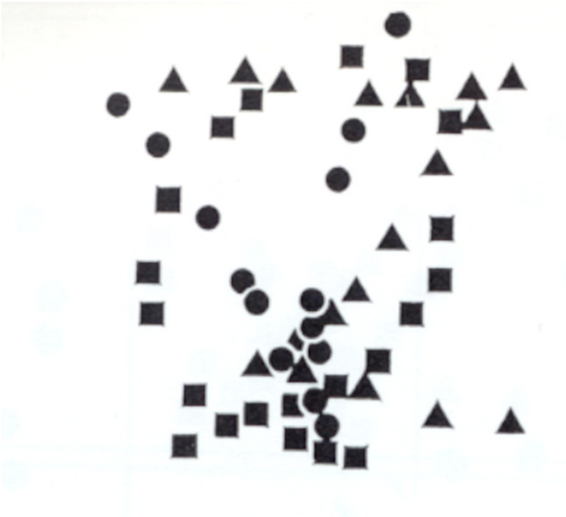


- order

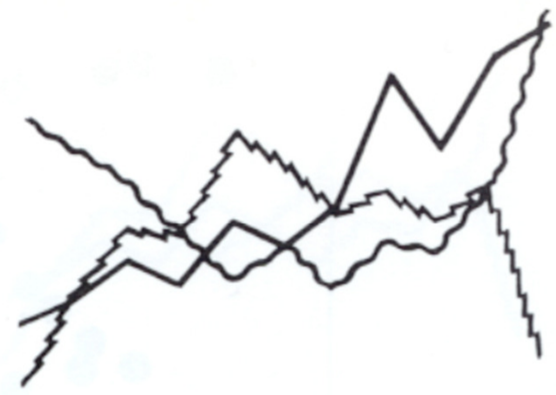


- length
  - infinite

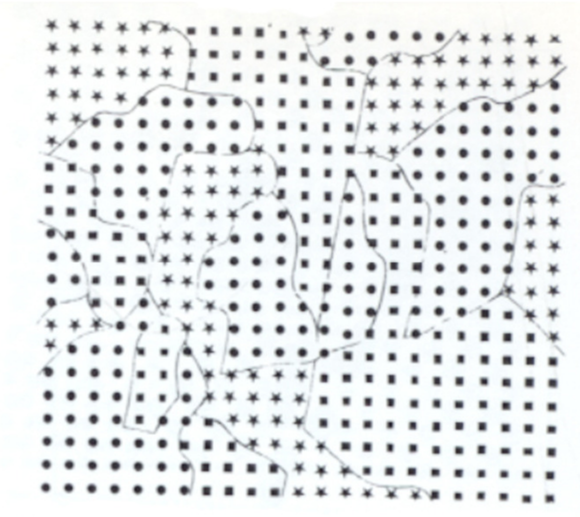
Shape



points



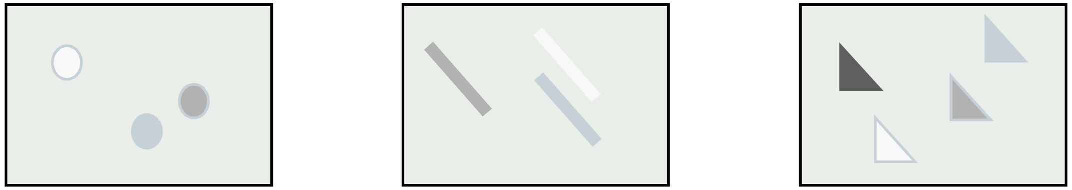
lines



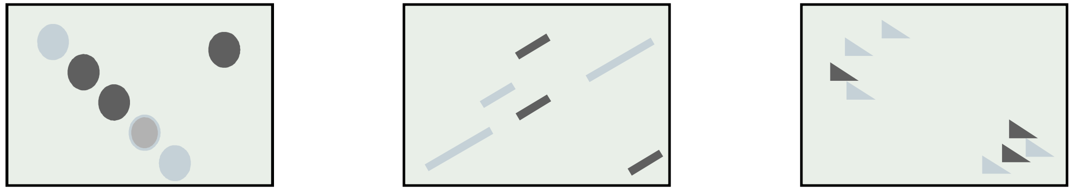
areas

Visual Variable: Value

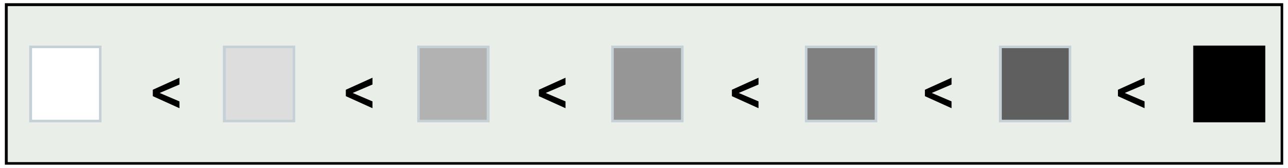
- selective



- associative

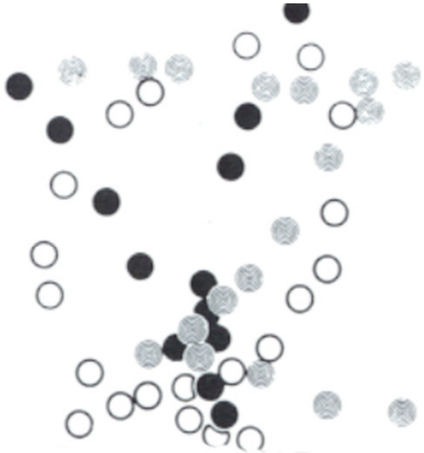


- quantitative
- order



- length
  - theoretically infinite but practically limited
  - association and selection ~ < 7 and distinction ~ 10

Value



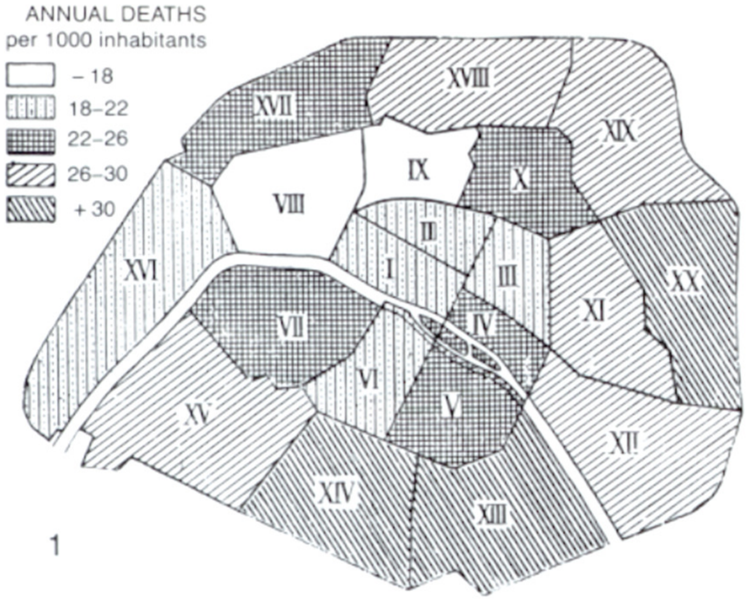
points

lines

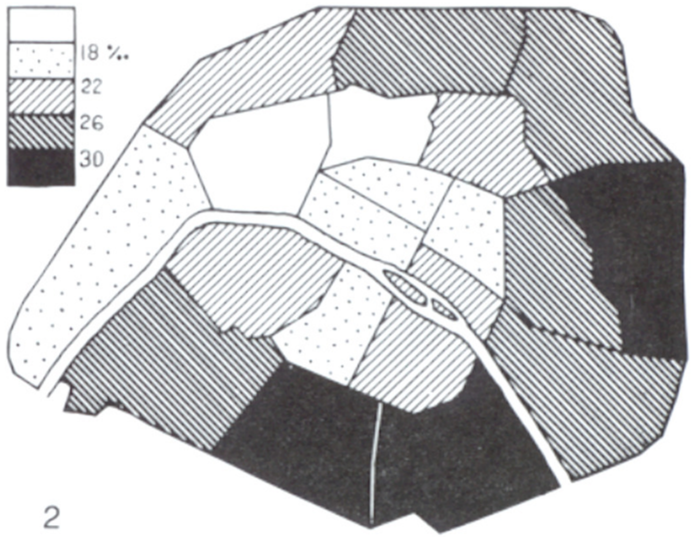
areas

Value

- Ordered, cannot be reordered



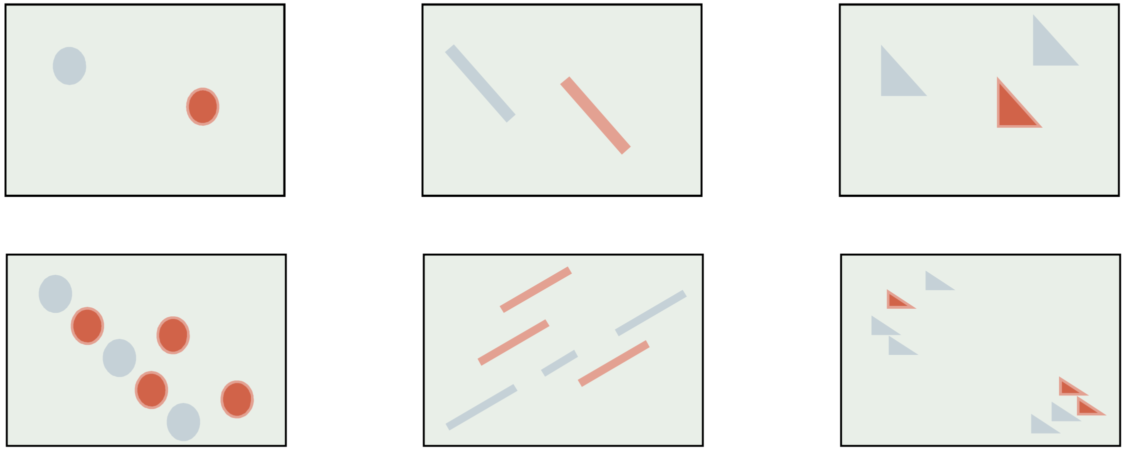
Values not ordered correctly according to scale
Information has to be read point by point



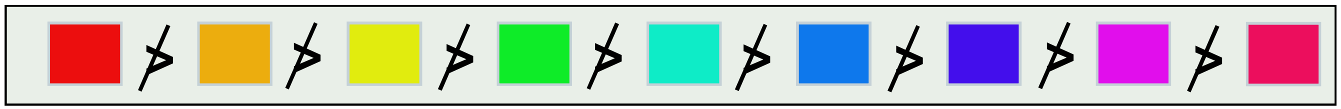
Values ordered correctly
Image much more useful

Visual Variable: Colour

- ✓ selective
- ✓ associative
- ≠ quantitative



- ≠ order

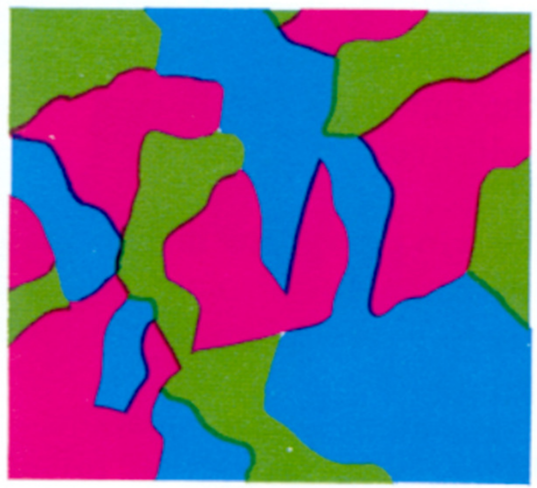


- ✓ length
  - theoretically infinite but practically limited
  - association and selection ~ < 7 and distinction ~ 10

Colour

points

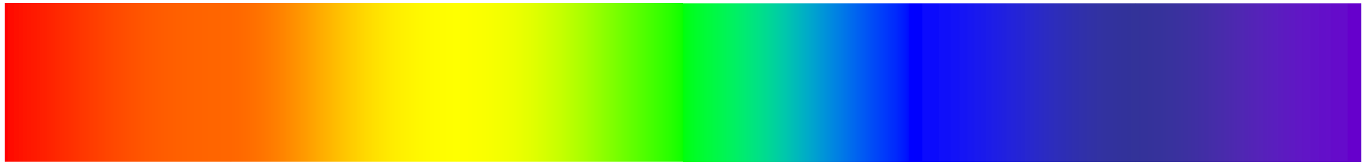
lines



areas

Colour Scales

- Common to use a rainbow scale… caution!

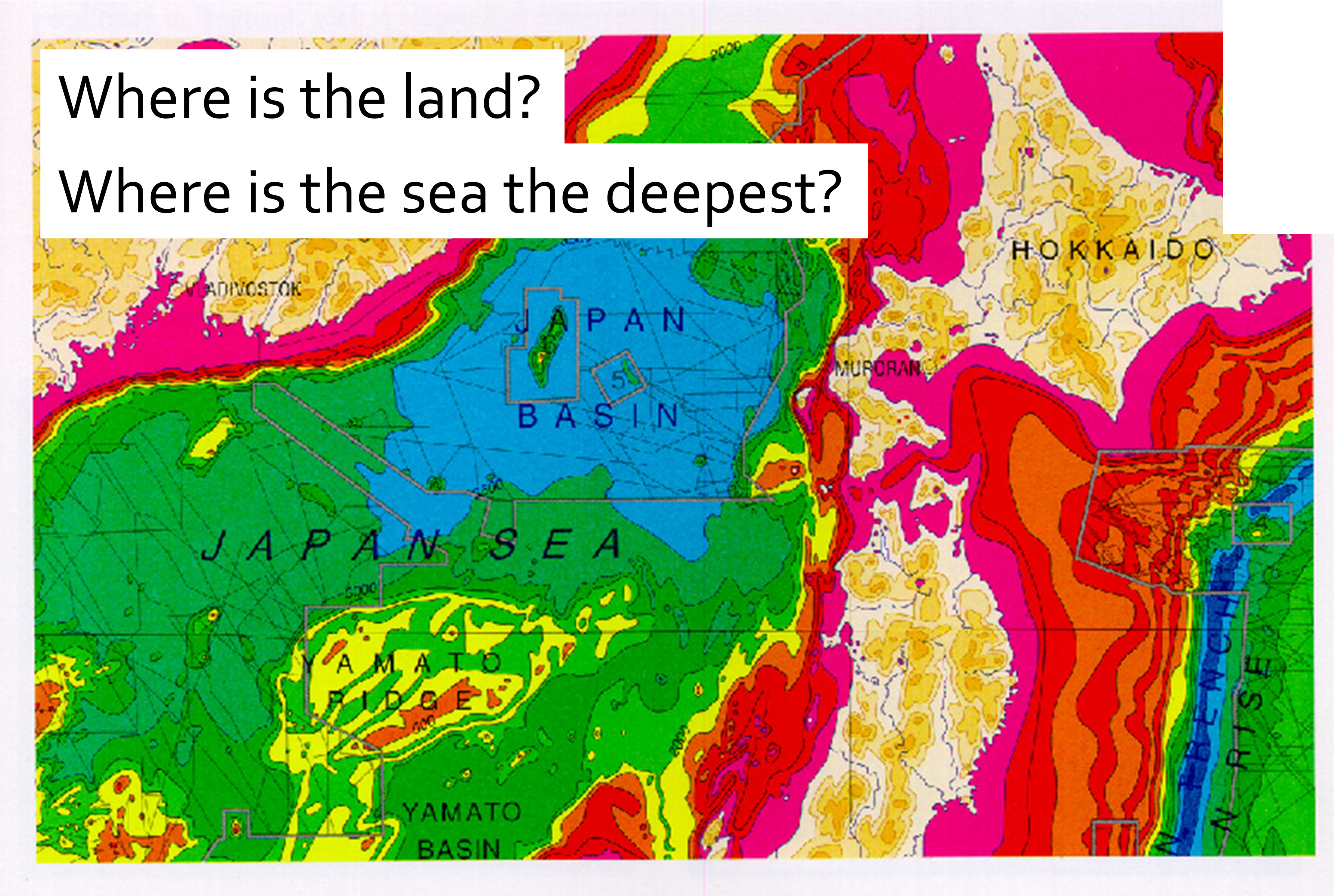


General Bathymetric Chart of the Ocean

Every color mark signals:
longitude, latitude, sear/land, depth/altitude

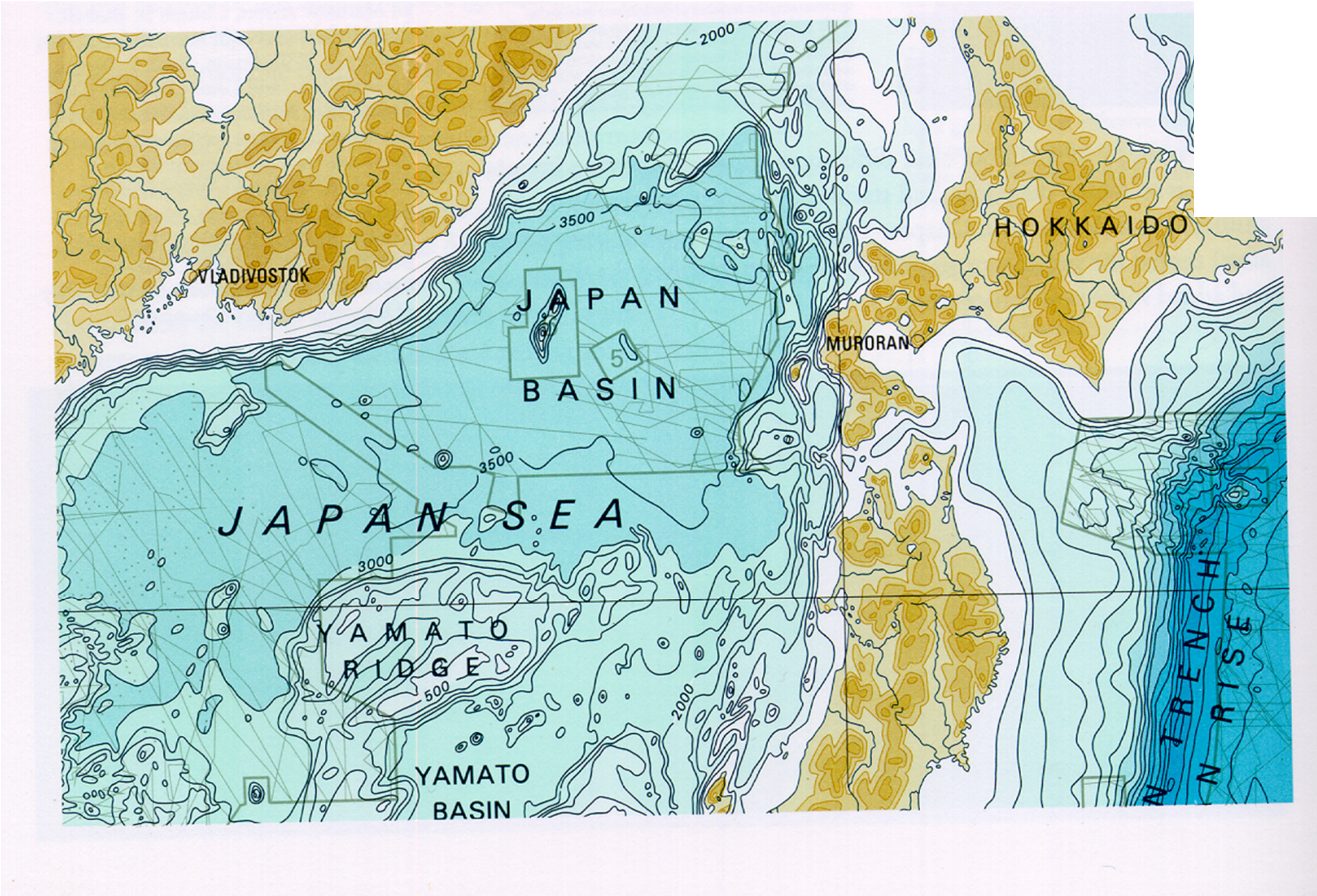
Where is the land?

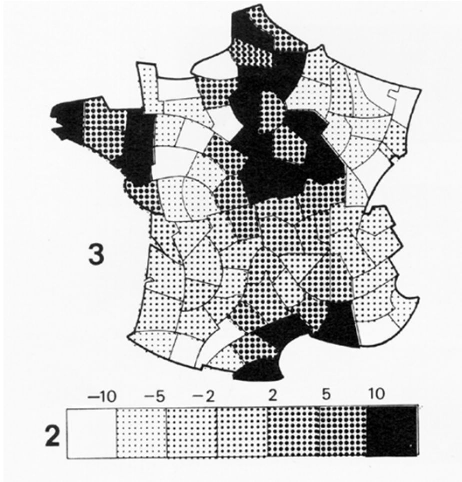
Where is the sea the deepest?



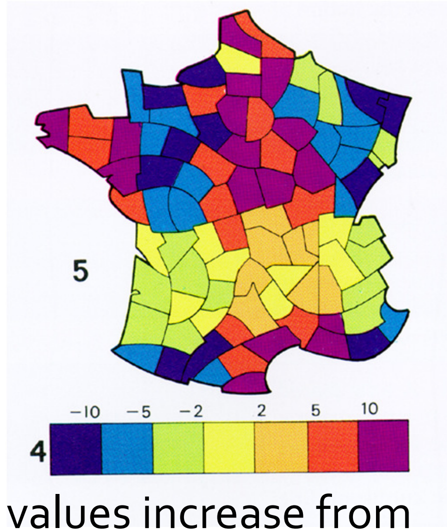
General Bathymetric Chart of the Ocean

Now describe what kind of color scale was possibly used here

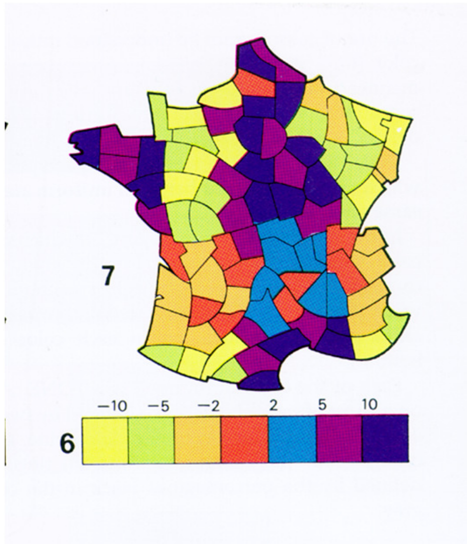




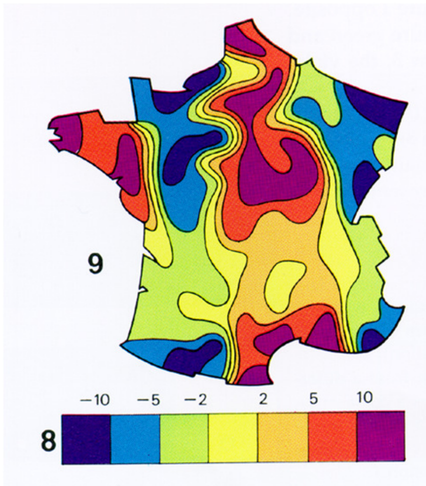
only represented by value
shows north-south band



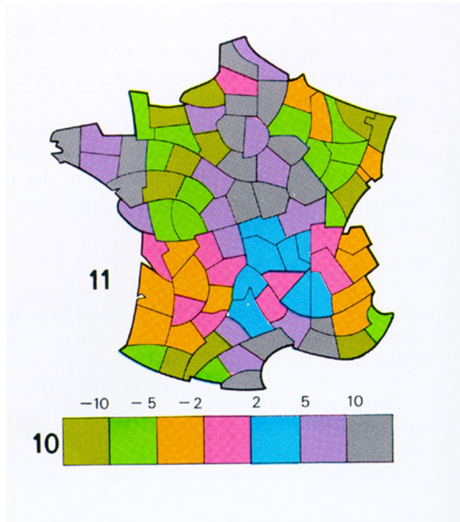
values increase from
yellow to left & right
creates east-west band



colors ordered by value
north-south band reappears



rainbow scale only works
when colors not next to each
other in the scale are not
brought together on the plane

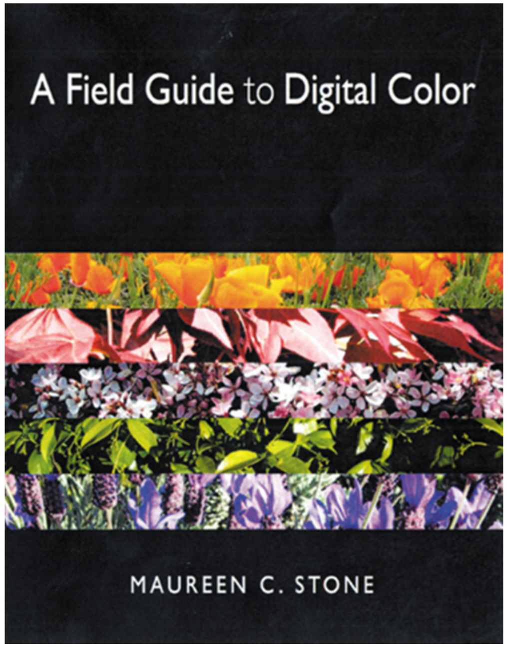


scale with no value changes
cannot represent orders
only good for selective and
associative depiction

Controls Legibility

Colour Resources

- Maureen Stone's Resources
  - http://www.stonesc.com
  - *A Field Guide to Digital Color (A. K. Peters)*

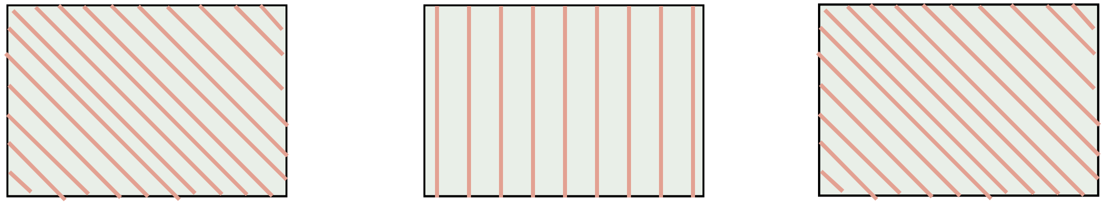


- Cindy Brewer's *ColorBrewer*
  - http://colorbrewer.org
- ColourLovers Community Palette Sharing
  - http://www.colourlovers.com

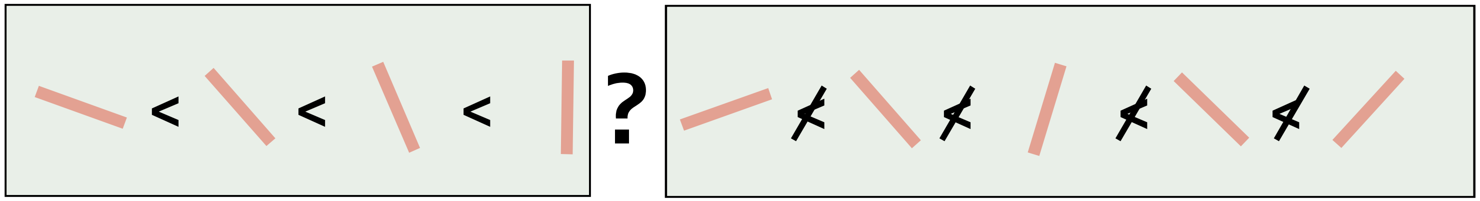
Visual Variable: Orientation

- ✓ selective

- ✓ associative

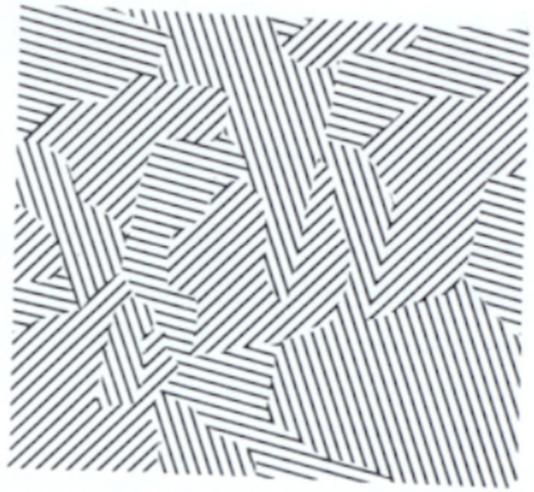


- ≠ quantitative
- ≠ order



- ✓ length
  - ~5 in 2D; ? in 3D

Orientation

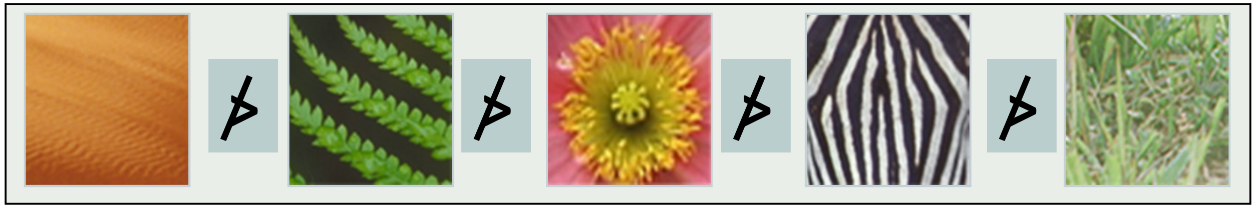
points lines areas

Visual Variable: Texture

- ✓ selective

- ✓ associative

- ≠ quantitative
- ≠ order

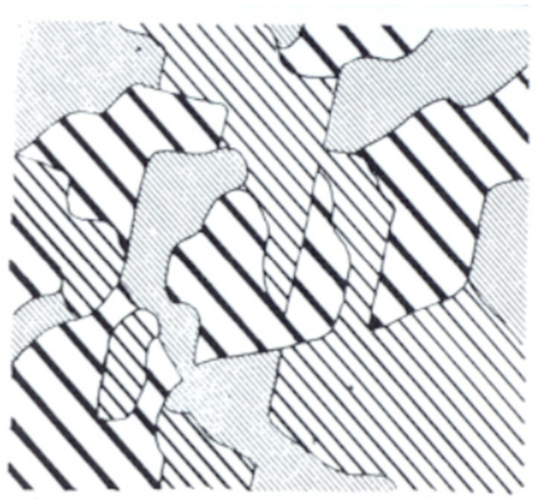


- ✓ length
  - theoretically infinite

Texture

points

lines



areas

Visual Variable: Motion

- ✓ selective
  - motion is one of our most powerful attention grabbers
- ✓ associative
  - moving in unison groups objects effectively
- $\neq$ quantitative
  - subjective perception
- $\neq$ order
- ? length
  - distinguishable types of motion?

Motion

Visual Variables

Visual Variable	Selective	Associative	Quantitative	Order	Length
Position	Yes	Yes	Yes	Yes	Dependant on resolution
Size	Yes	Yes	Approximate	Yes	Association: 5; Distinction: 20
Shape	With Effort	With Effort	No	No	Infinite
Value	Yes	Yes	No	Yes	Association: 7; Distinction: 10
Hue	Yes	Yes	No	No	Association: 7; Distinction: 10
Orientation	Yes	Yes	No	No	4
Grain	Yes	Yes	No	No	5
Texture	Yes	Yes	No	No	Infinite
Motion	Yes	Yes	No	Yes	Unknown

Short Break

- Summary:
  - Now you know the main building blocks are **marks**
  - Marks are modified by **visual variables**
  - Visual variables have **specific characteristics**
  - These characteristics influence how the data will be perceived
- Next you will
  - See that the vision system is **quicker and better** at detecting certain visual variables
  - Learn that these influences how visualizations will be read

Preattentive Processing

- Some things can be extracted without need for focused attention
- Generally within 200-250 ms
- Seems to be done in parrel by low-level vision system
- Preattentive processing is influenced by visual encoding

How many 3's do you see?

1281768756138976546984506985604982826762
9809858458224509856458945098450980943585
9091030209905959595772564675050678904567
8845789809821677654876364908560912949686

From: Ware, Information Visualization using Vision to Think

How about now?

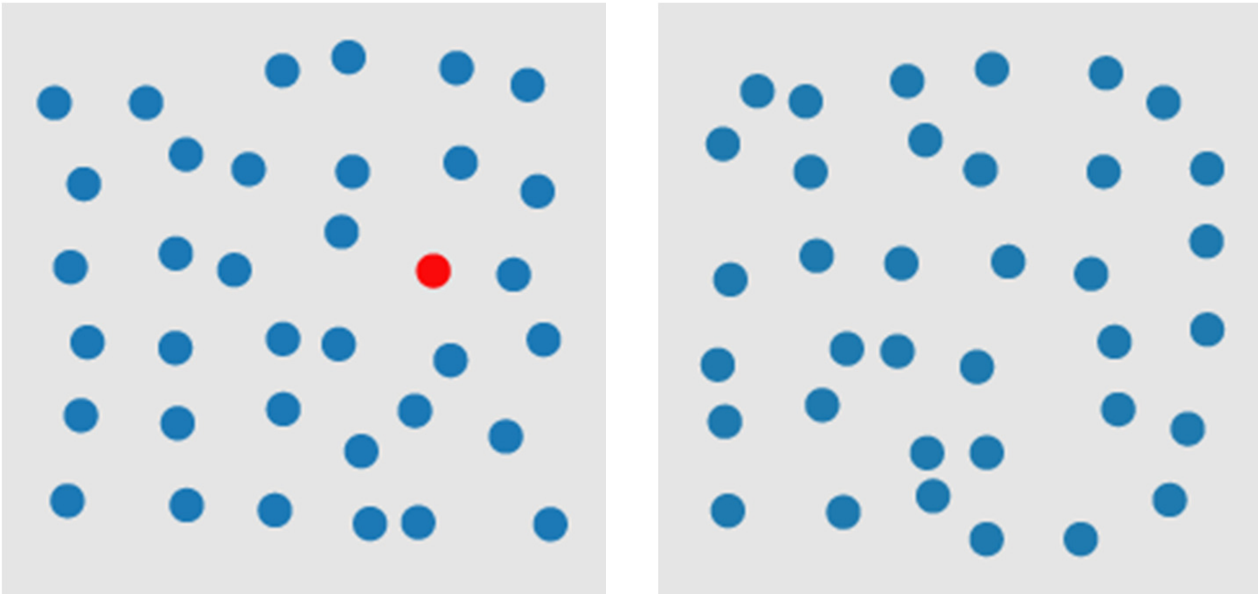
1281768756138976546984506985604982826762
9809858458224509856458945098450980943585
9091030209905959595772564675050678904567
8845789809821677654876364908560912949686

From: Ware, Information Visualization using Vision to Think

Visual cariables influence preattentive processing

DETERMINE IF A RED CIRCLE IS PRESENT

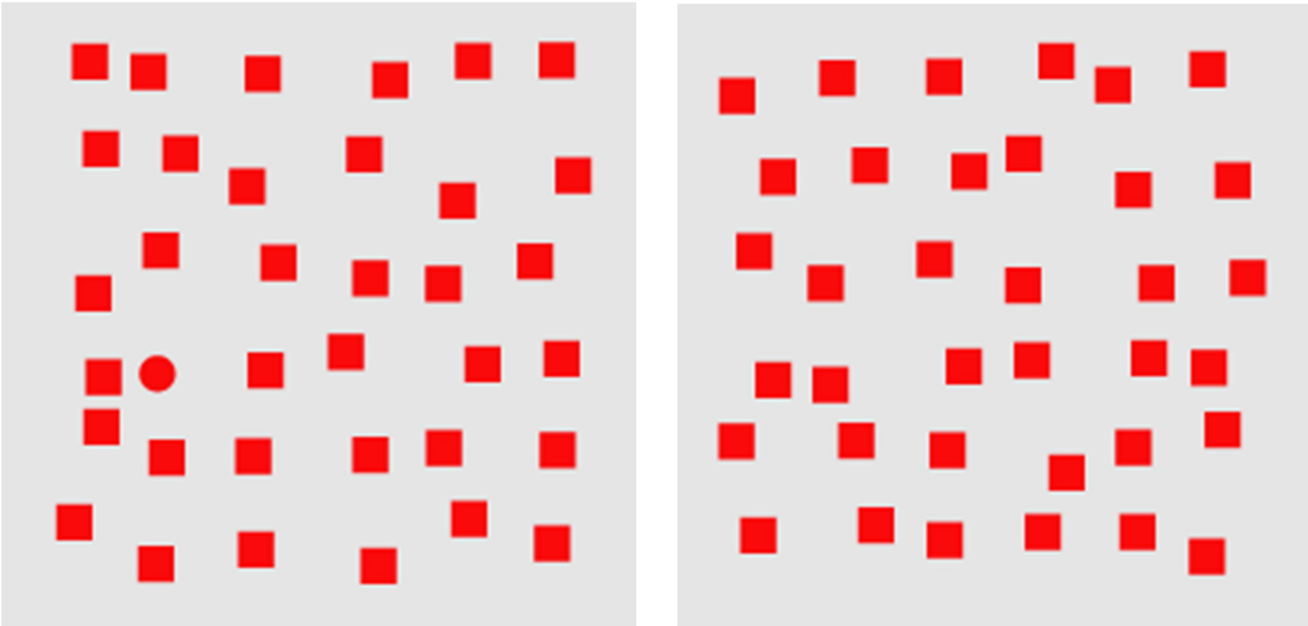
Hue



Yes, can be done preattentively

From: Healey, Perception in Visualization

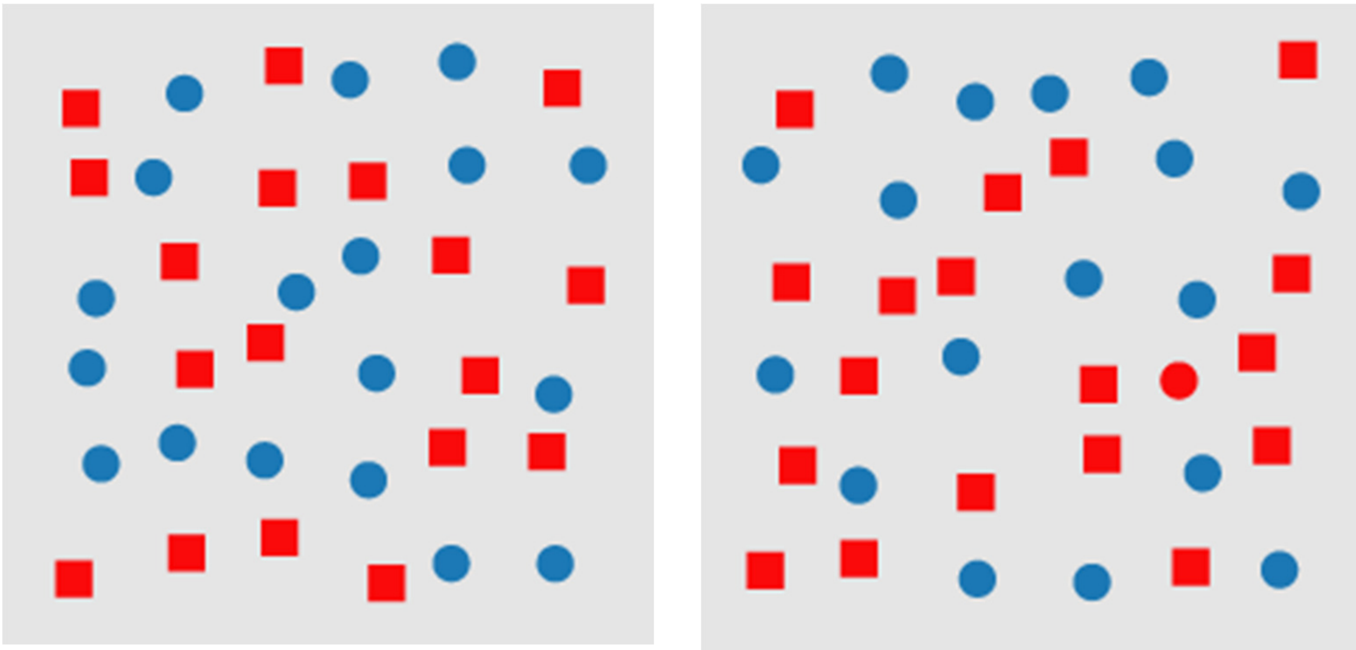
Shape



Yes, can be done preattentively

From: Healey, Perception in Visualization

Hue and Shape



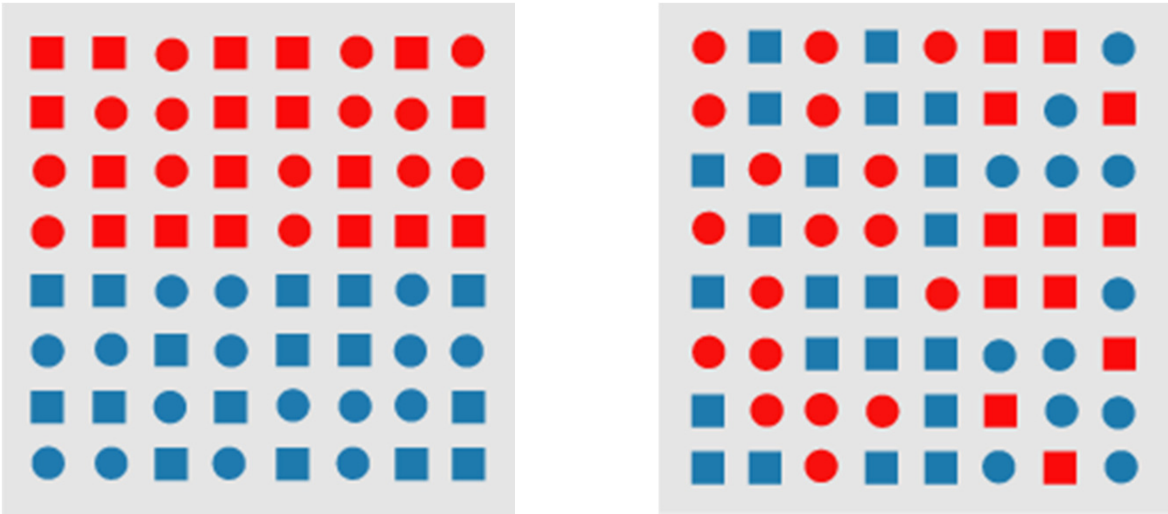
Cannot be done preattentively due to conjunction of shape and hue
→ need to search

From: Healey, Perception in Visualization

Visual cariables influence preattentive processing

DETERMINE IF THERE IS A BOUNDARY IN THE DISPLAY

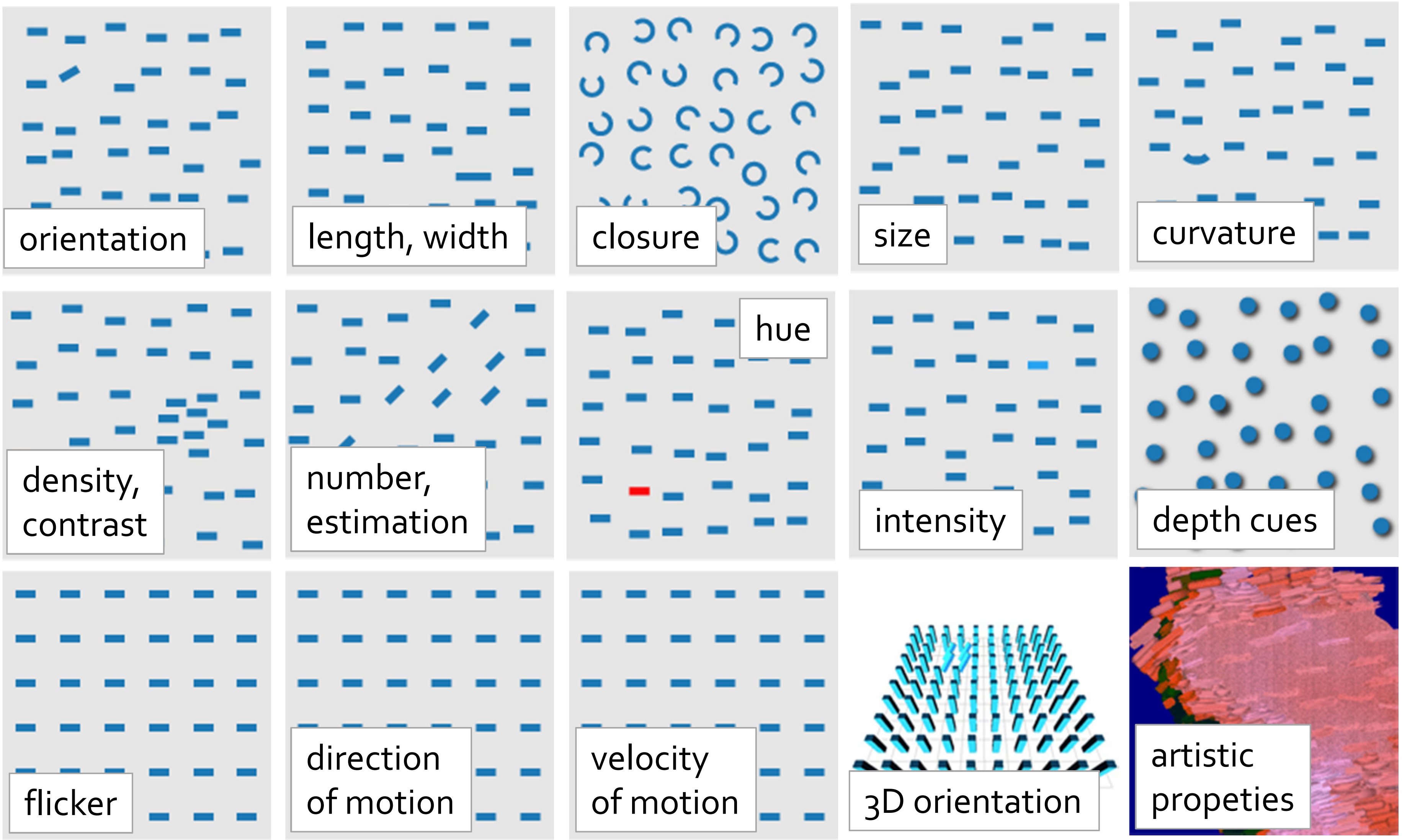
Fill and Shape



- Left can be done preattentively (groups are distinguished by one unique feature)
- Right cannot since groups are distinguished by two features

From: Healey, Perception in Visualization

Preattentive visual features *(some)*



From: Healey, Perception in Visualization

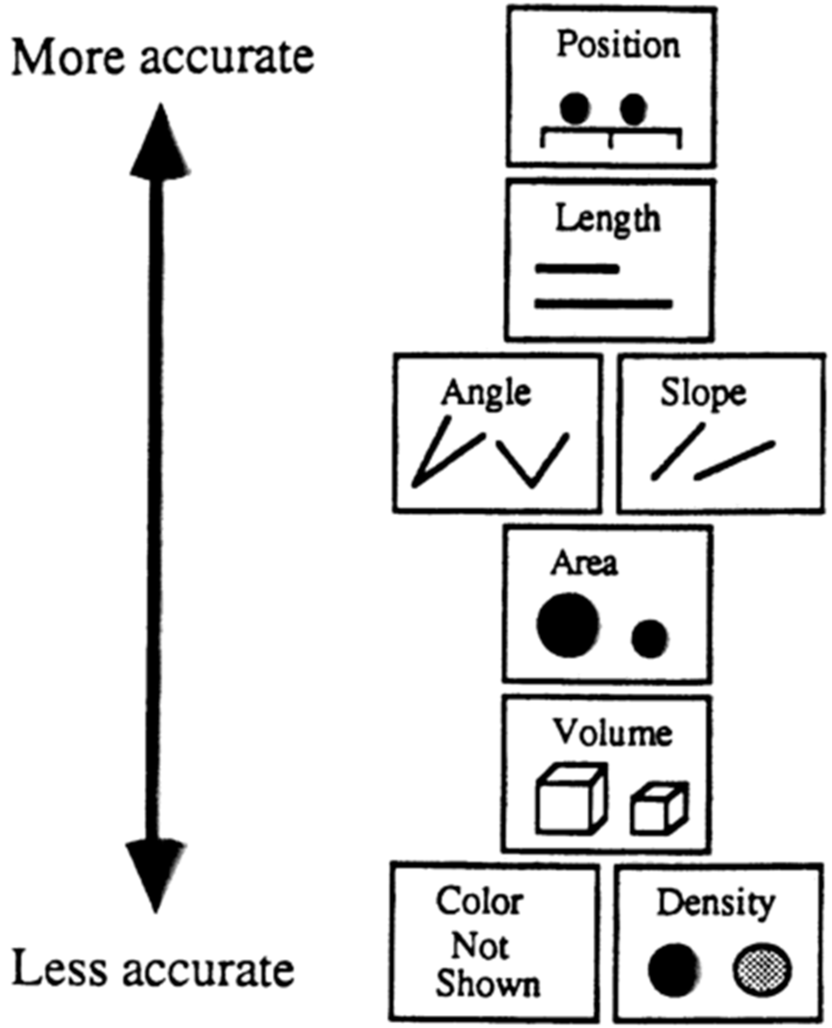
Elementary Graphical Perception Tasks

William S. Cleveland 1980s

- Performed controlled experiments to find out how effectively people could judge changes in visual features
- Focus on quantitative information
- Variables used: angle, area, color hue, color saturation, density (value), length, position, slope, volume

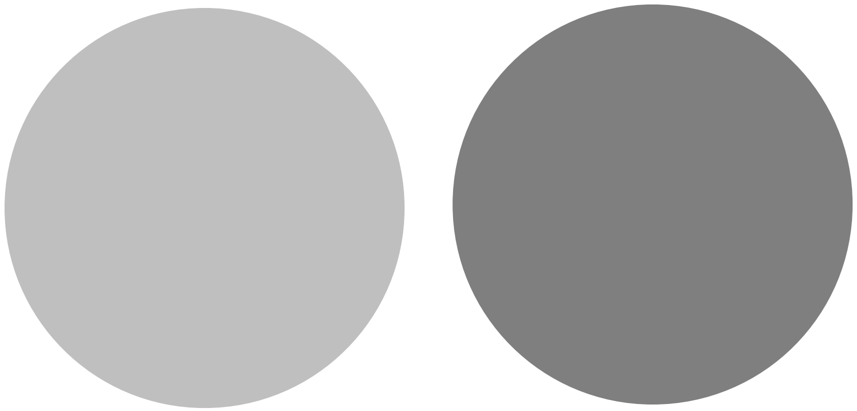
Elementary Graphical Perception Tasks

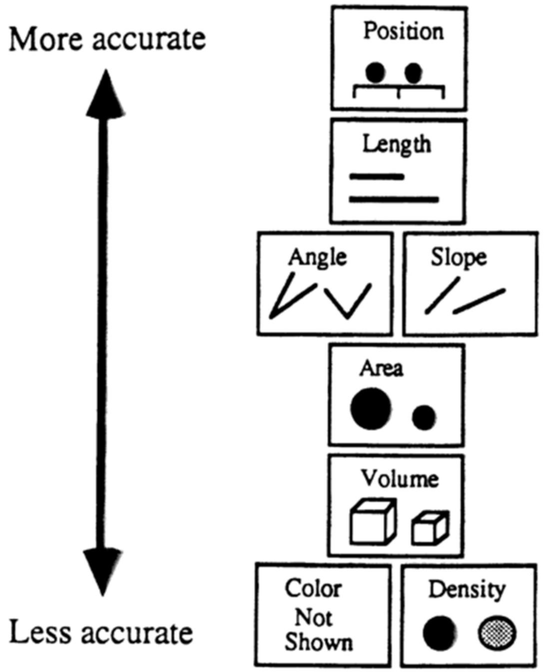
William S. Cleveland 1980s



Value

- What percentage in value is the right from the left (=100%) ?

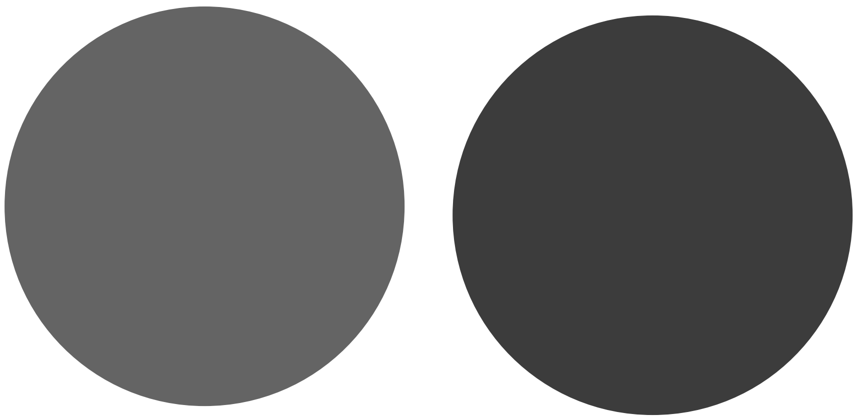




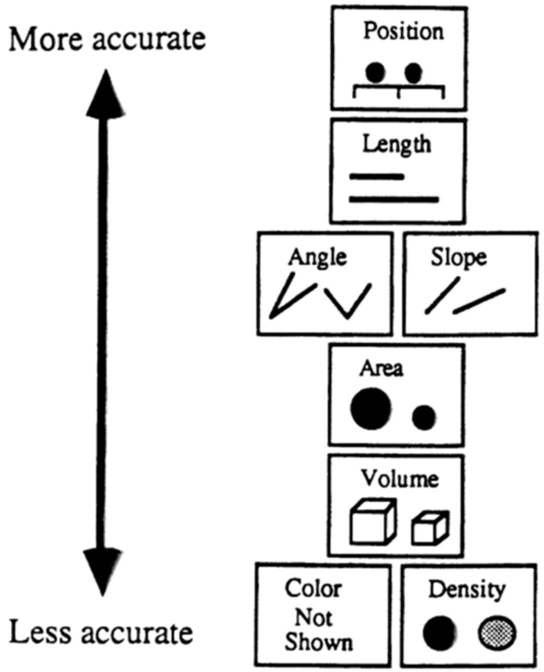
66%

Value

- What percentage in
  value is the right from the left
  (=100%)?

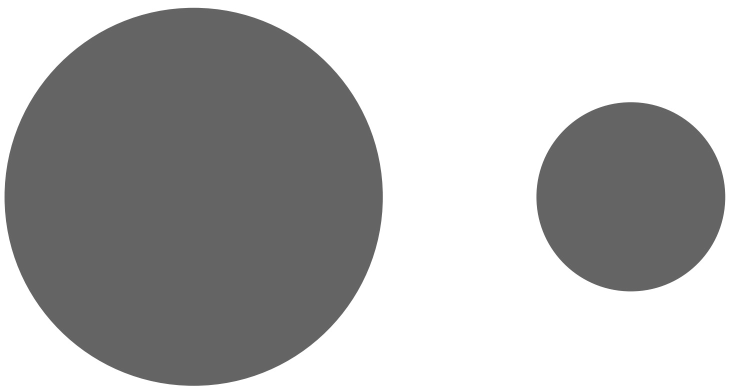


60%

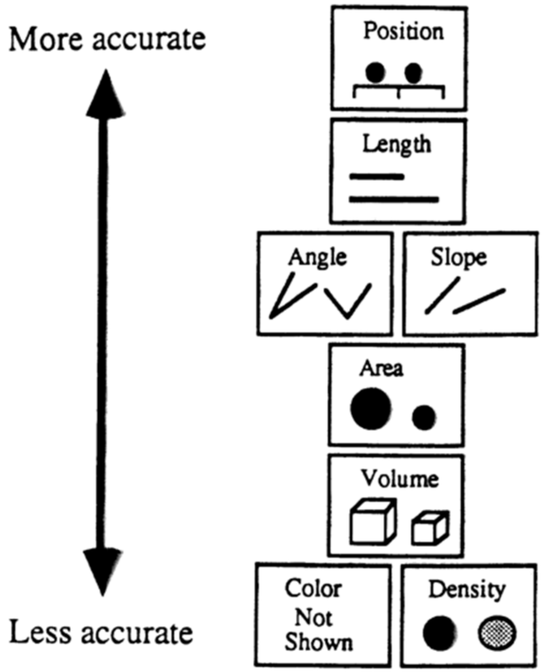


Area

- What percentage in
  size is the right from the left
  (=100%)?

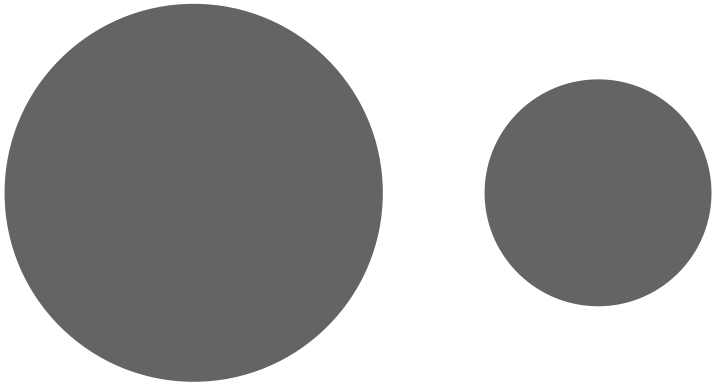


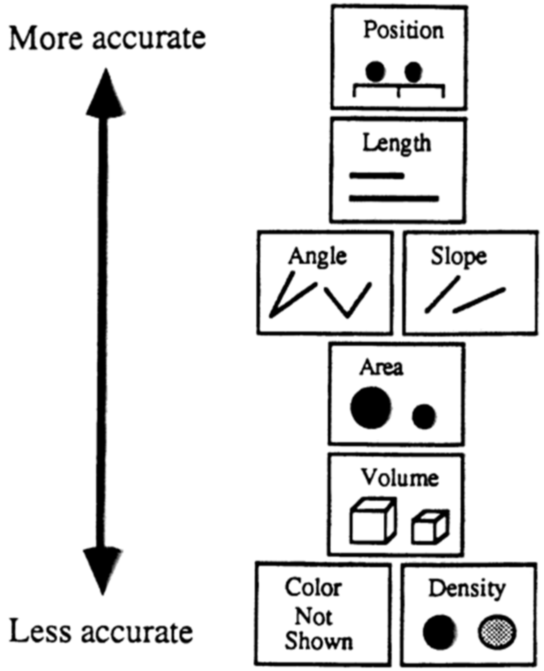
25%



Area

- What percentage in size is the right from the left (=100%)?

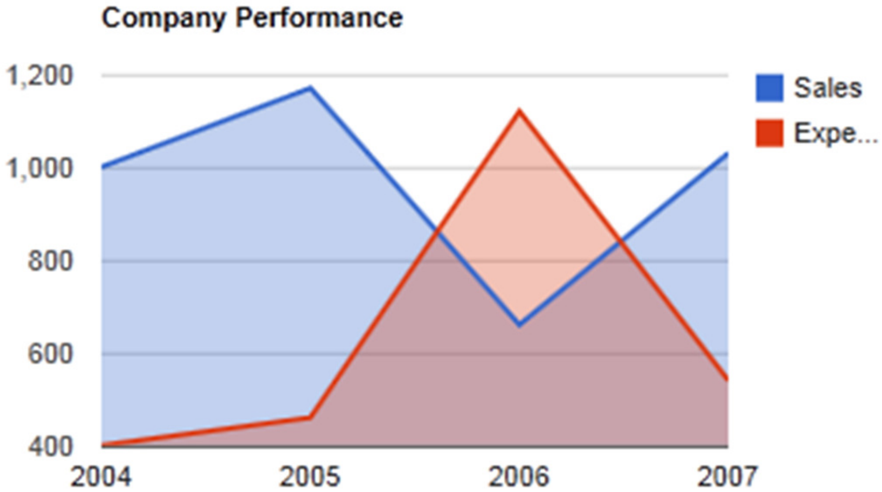


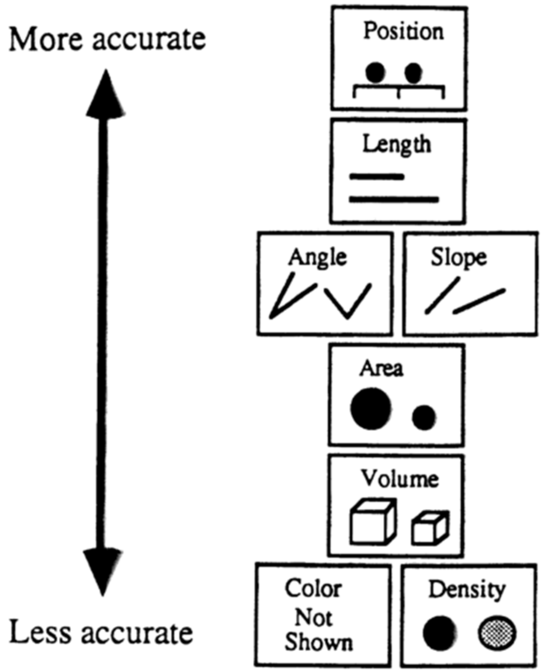


36%

Area

- What percentage in size is the red from the blue (=100%)?



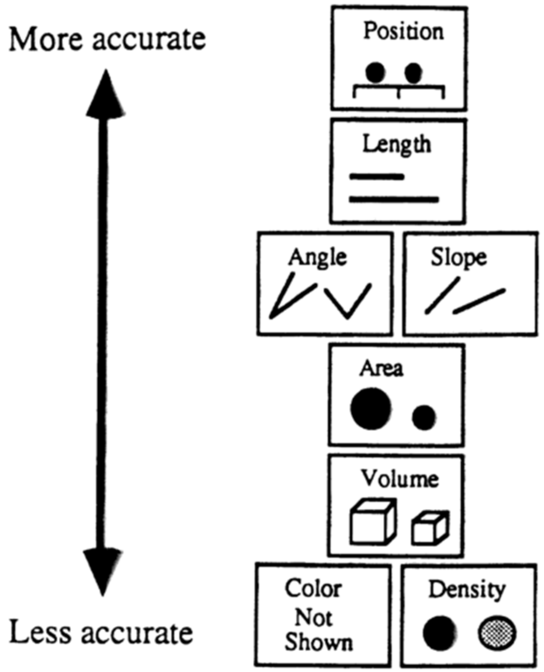


no idea – this is very difficult

Length

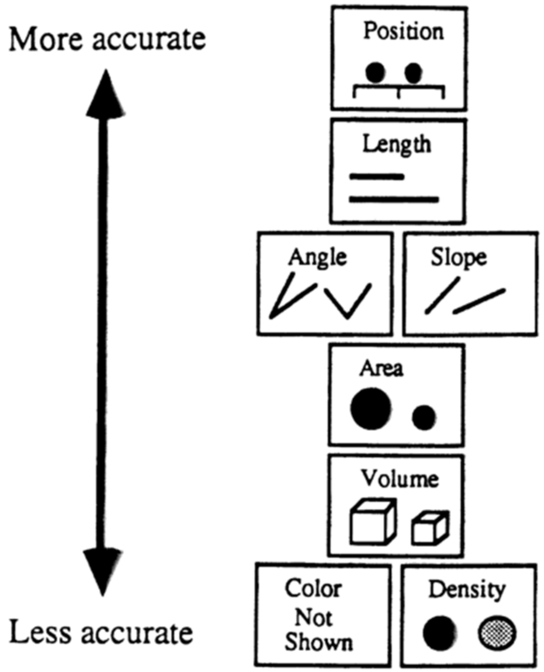
- What percentage in length is the right from the left (=100%)?

75%



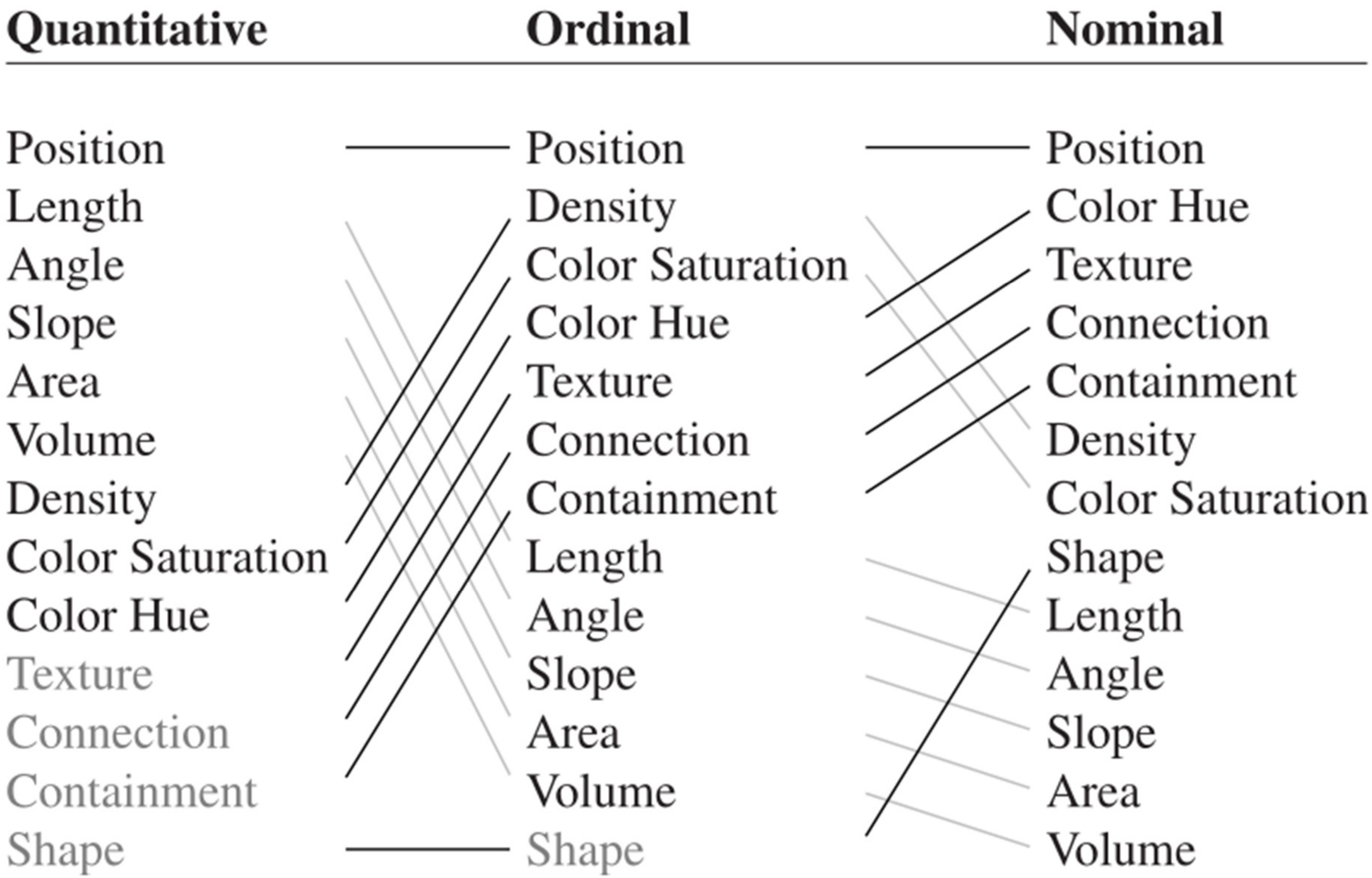
Length / Position

- What percentage in length is the right from the left (=100%)?



25%

Effectiveness of Data Encodings



Mackinlay 1986

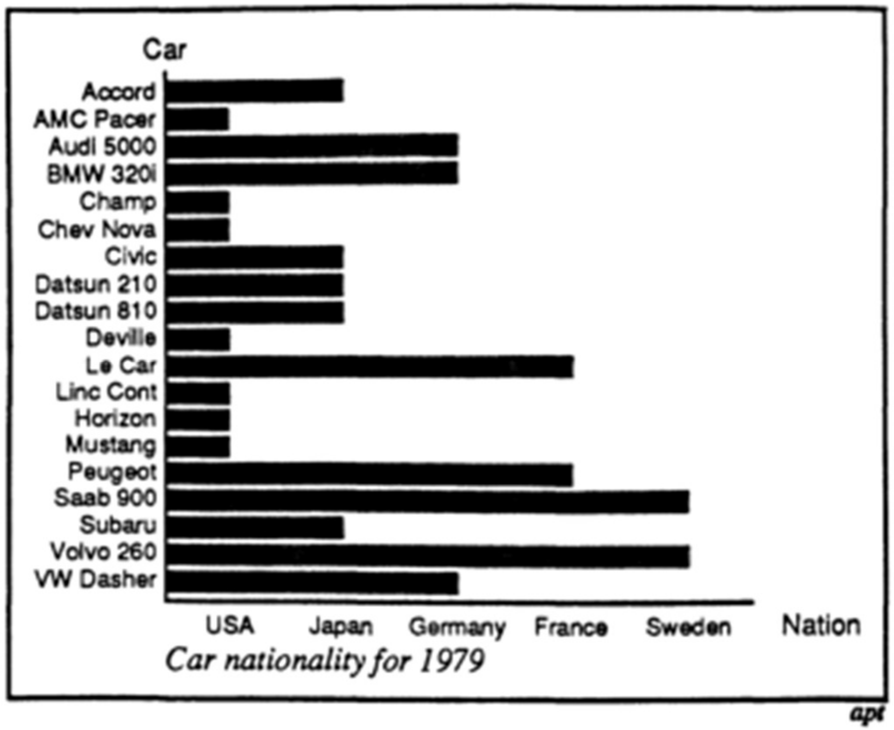
Let's evaluate…

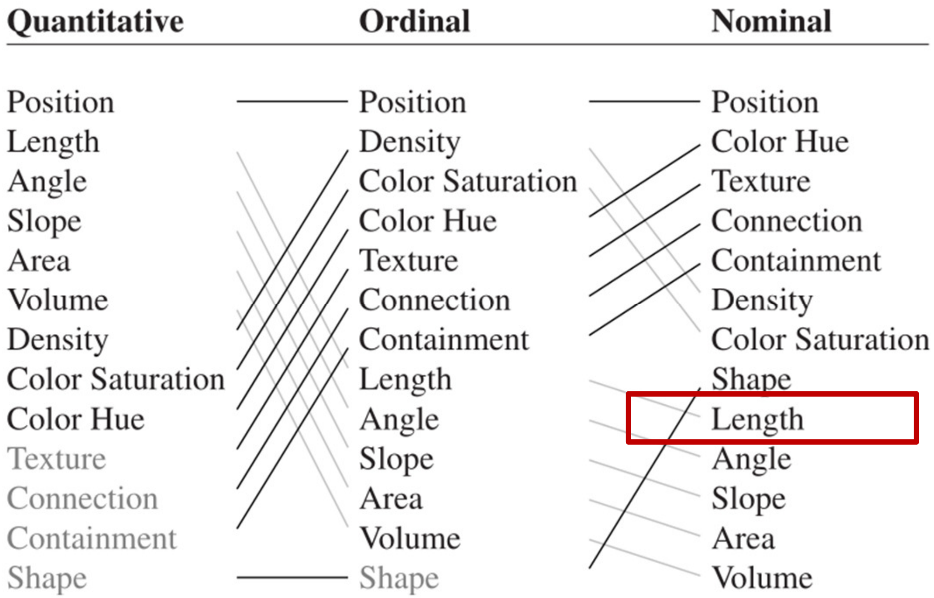
Car / Nation	USA	Japan	Germany	France	Sweden
Accord		x			
AMC Pacer	x				
Audi 5000			x		
BMW 320i			x		
Champ	x				
Chev Nova	x				
Saab 9000				x	

What kind of data are we looking at?

Nations: Nominal
Cars: Nominal
(Nation,Car): Nominal

Let's evaluate…



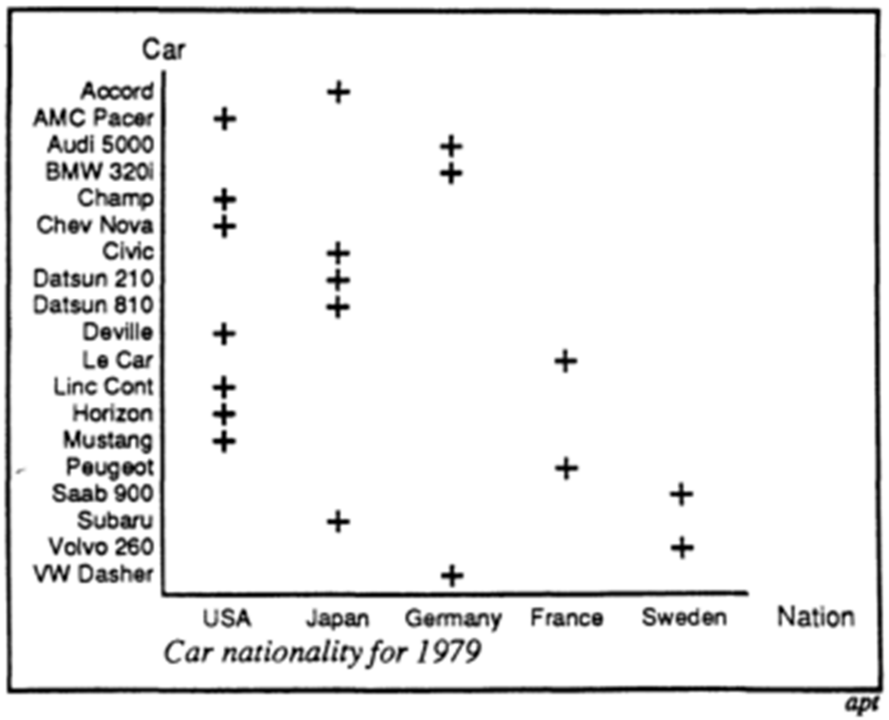


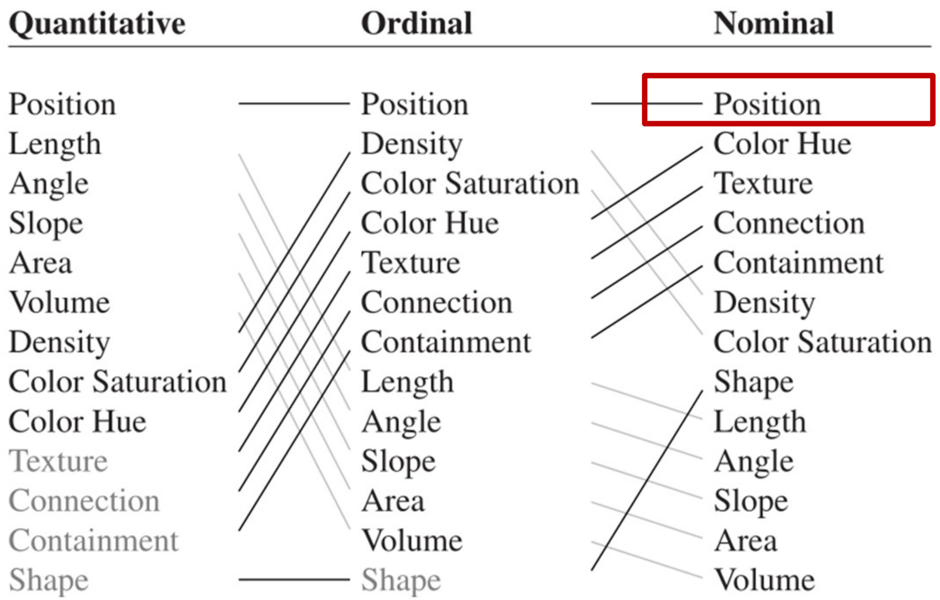
Problem:

Length of bar suggests an order or quantity

(e.g. Swedish cars are better)

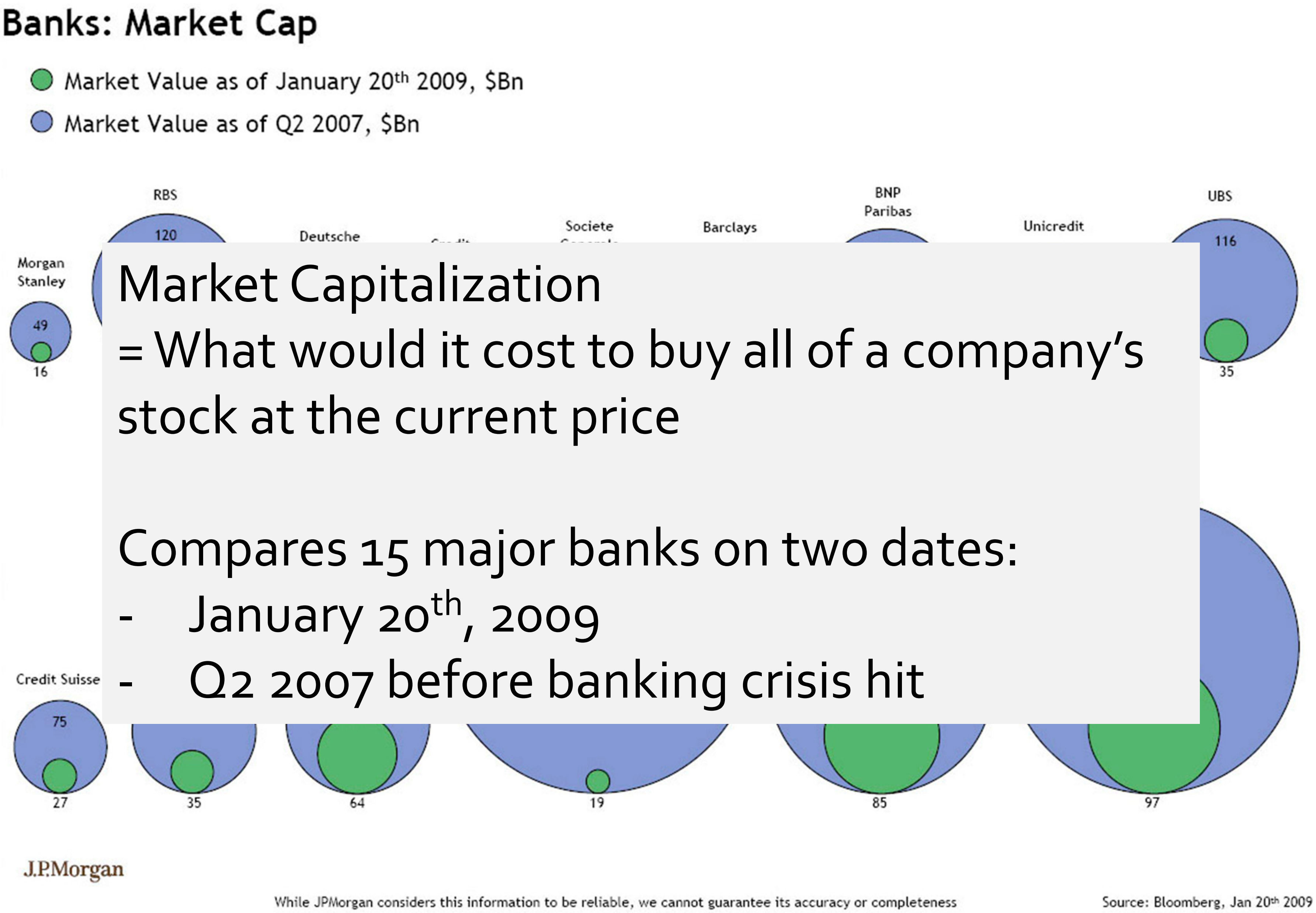
Let's evaluate…





Better!

Let's evaluate…



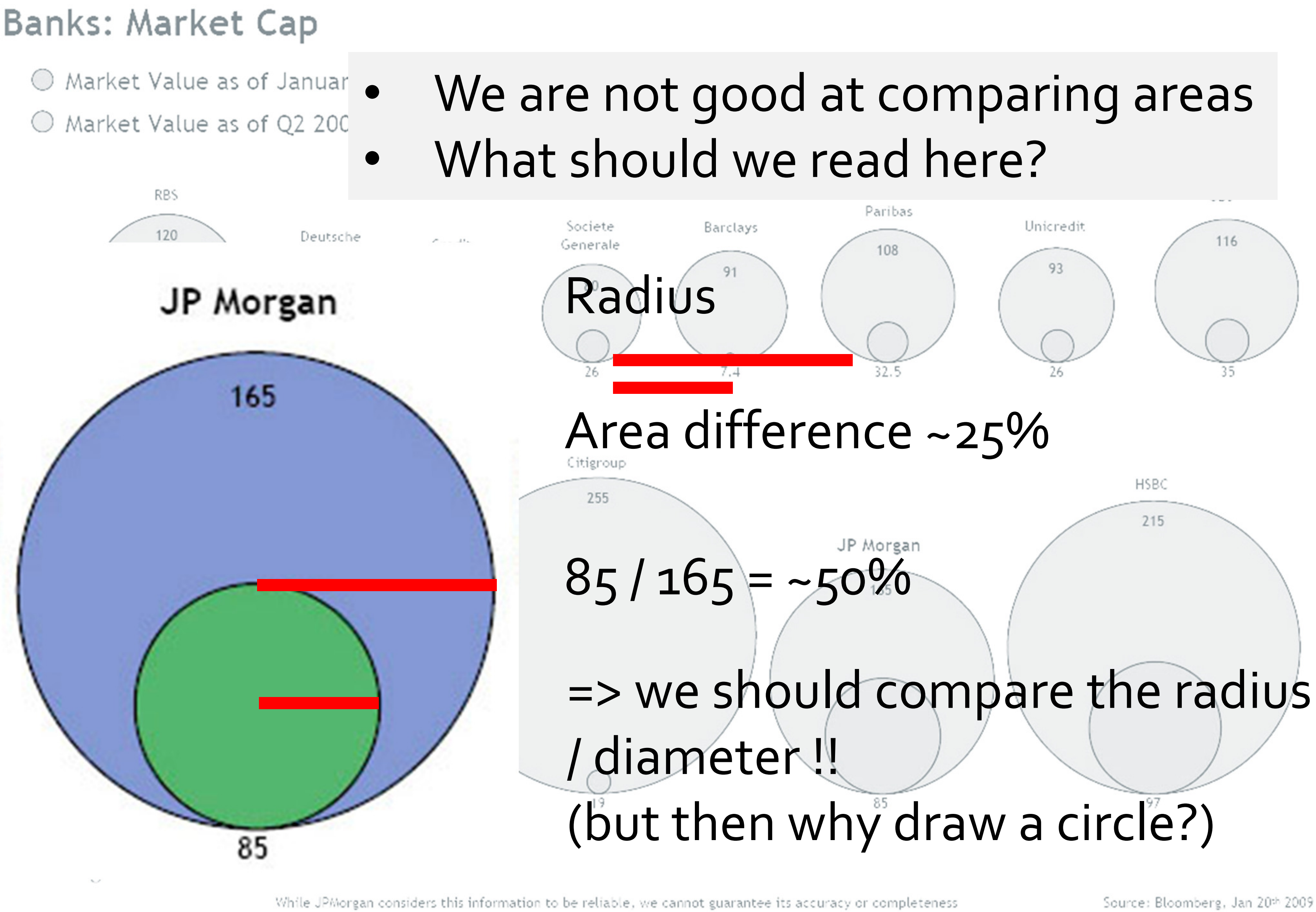
Market Capitalization
= What would it cost to buy all of a company's stock at the current price

Compares 15 major banks on two dates:
- January 20th, 2009
- Q2 2007 before banking crisis hit

Problems here?

- We are not good at comparing areas

Problems here?



- We are not good at comparing areas
- What should we read here?

Radius

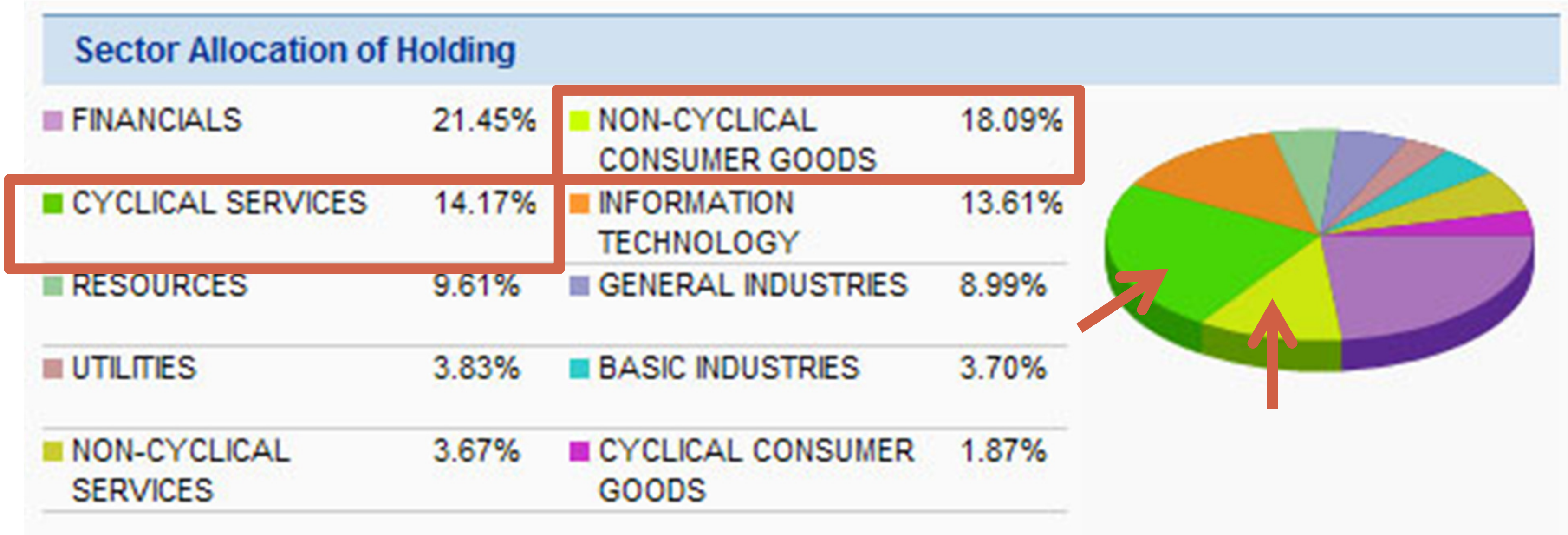
Area difference ~25%

85 / 165 = ~50%

=> we should compare the radius / diameter !!

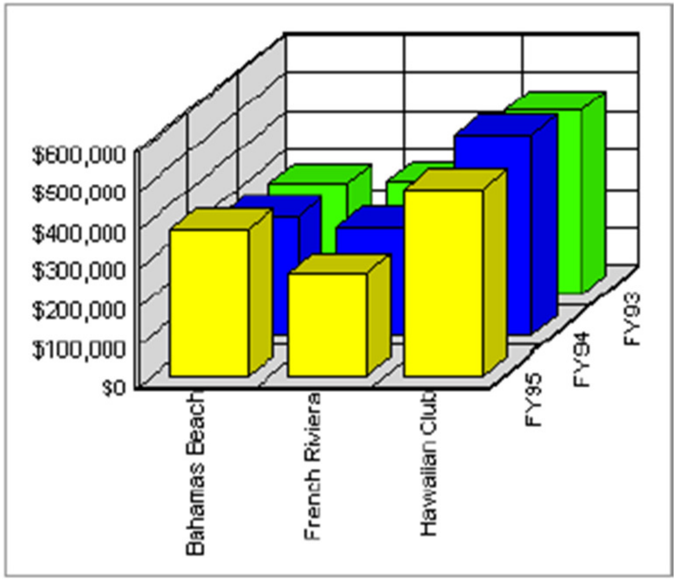
(but then why draw a circle?)

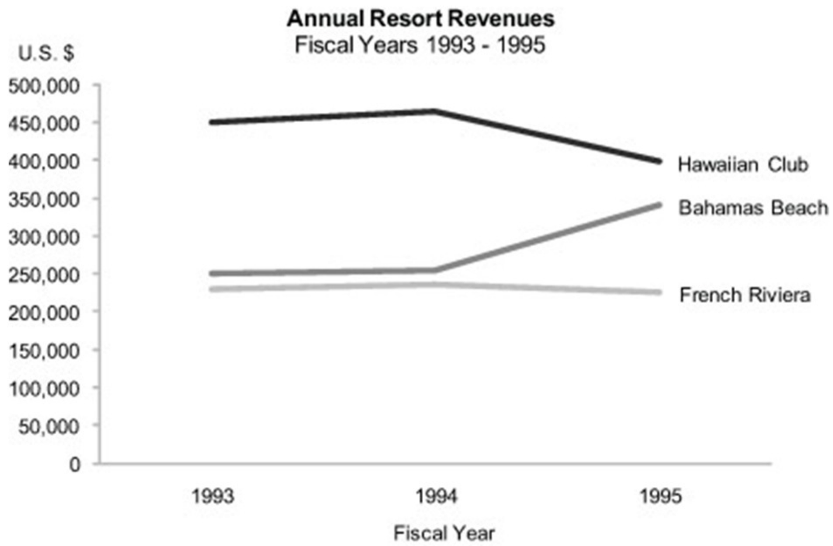
Problem here?



- Pie slices are difficult to compare in area
- There is likely a bug or error in the data
- Perspective distortion adds to the problem
- Colors are difficult to distinguish

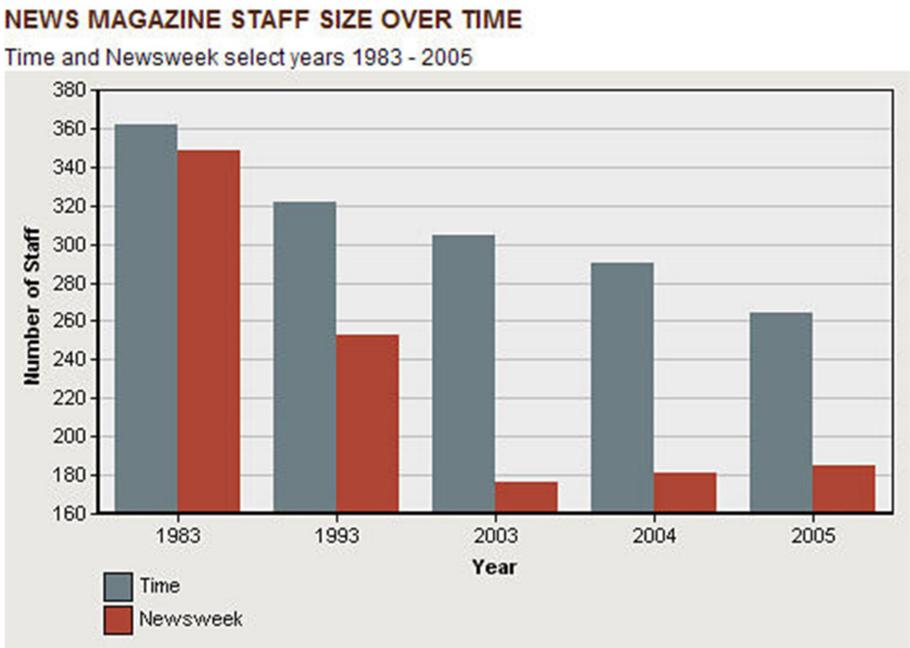
Similarly…3D bar charts are not recommended

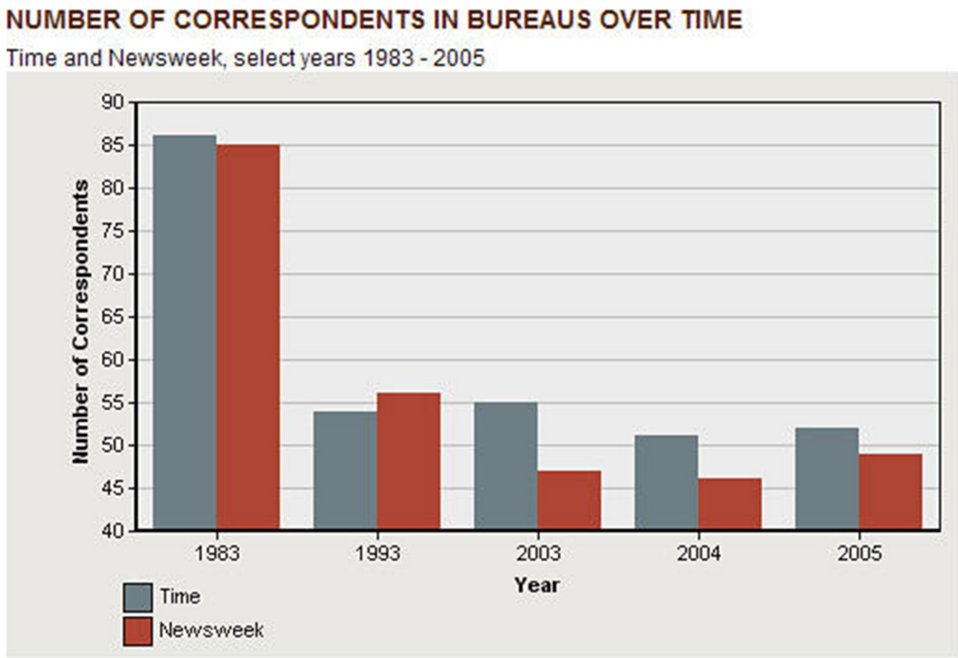


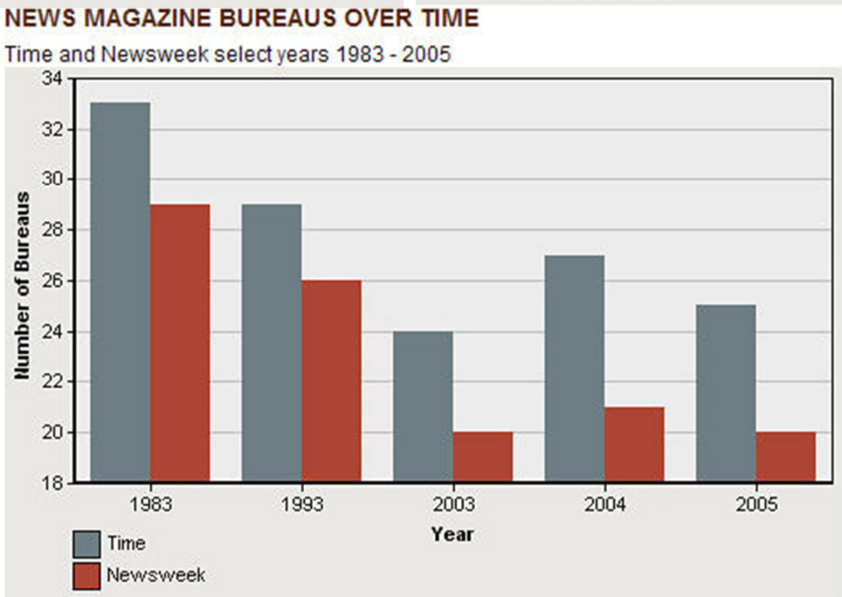


This is much easier to see and compare!

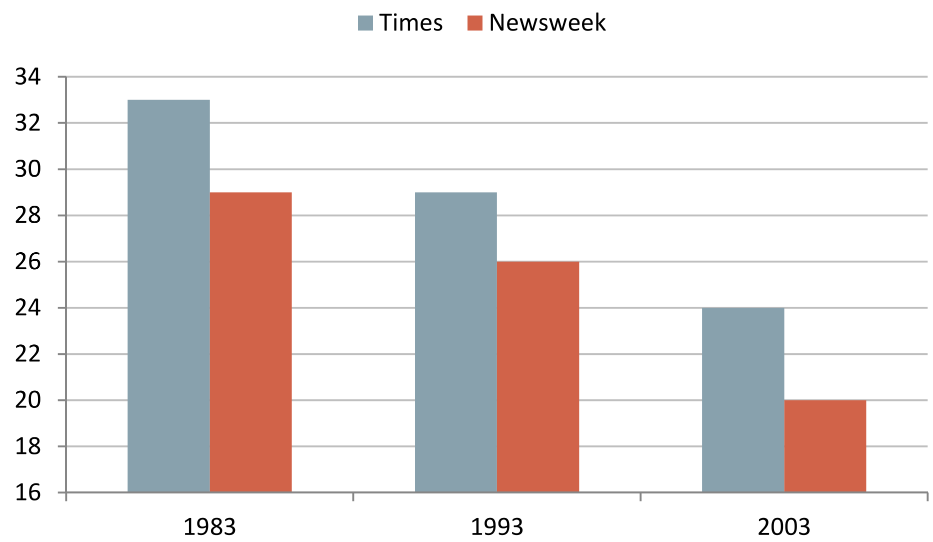
Problem here?

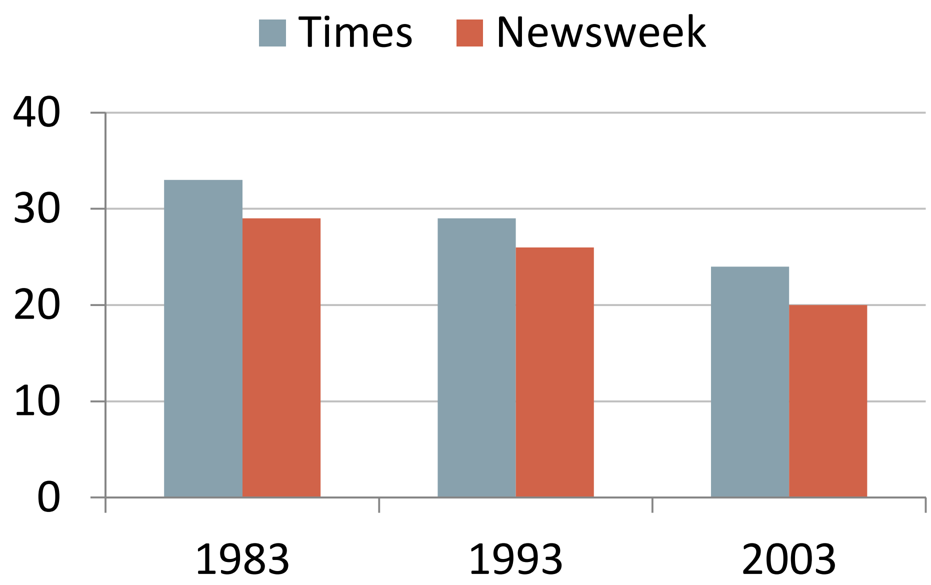






Length Comparison

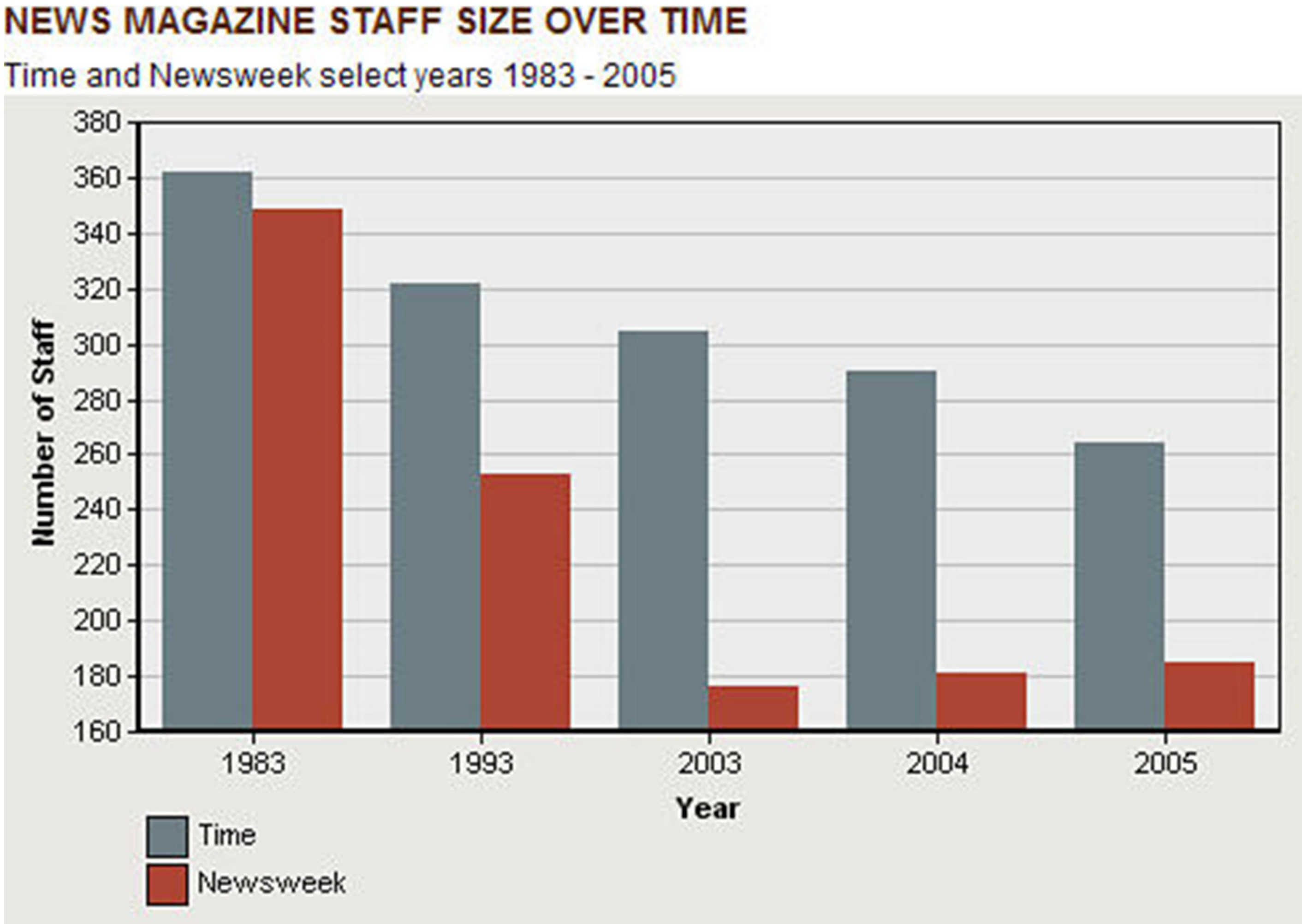




At first glance:
2003: Newsweek is 50% of Times

If we add a proper 0:
2003: Newsweek is ~80% of Times

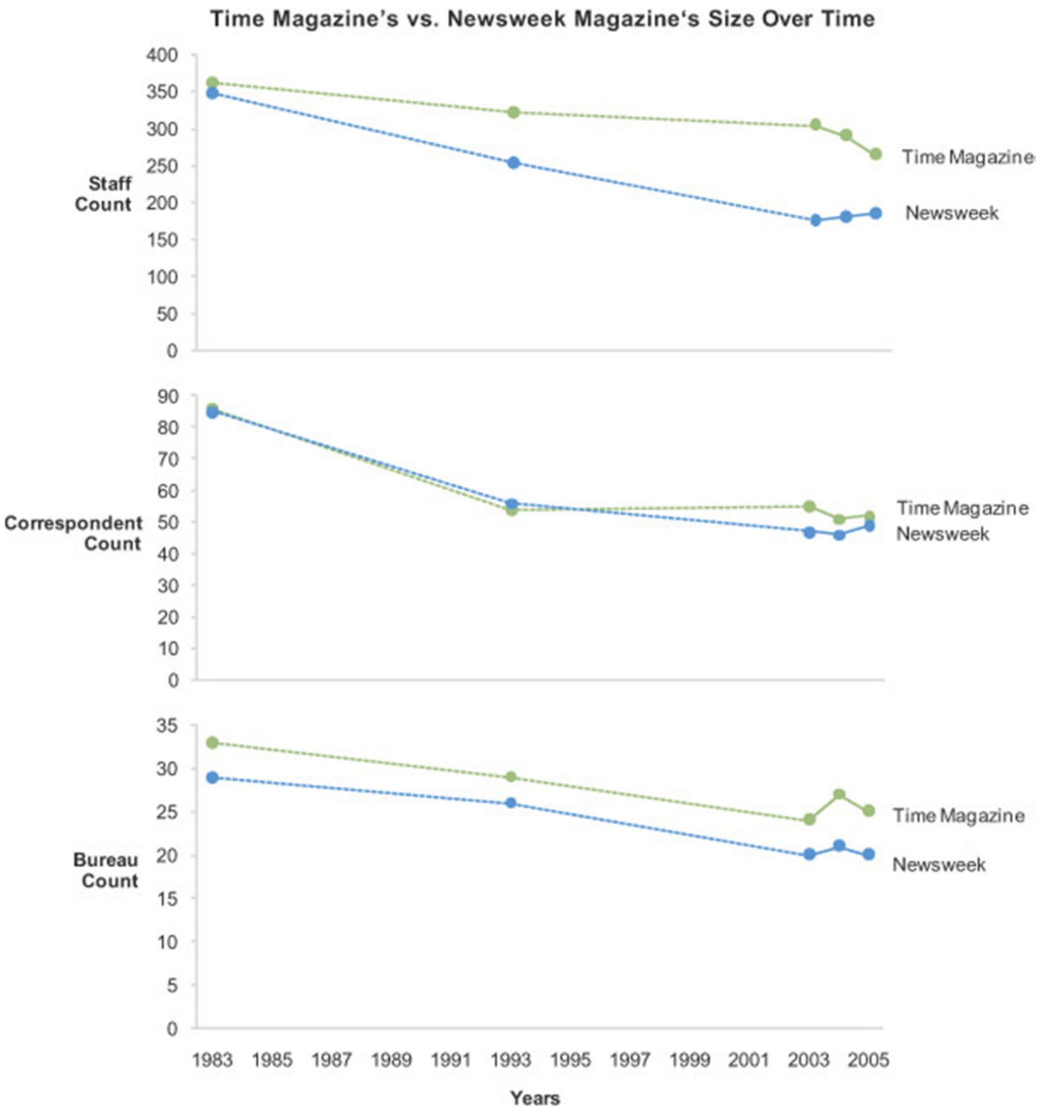
Moreover…



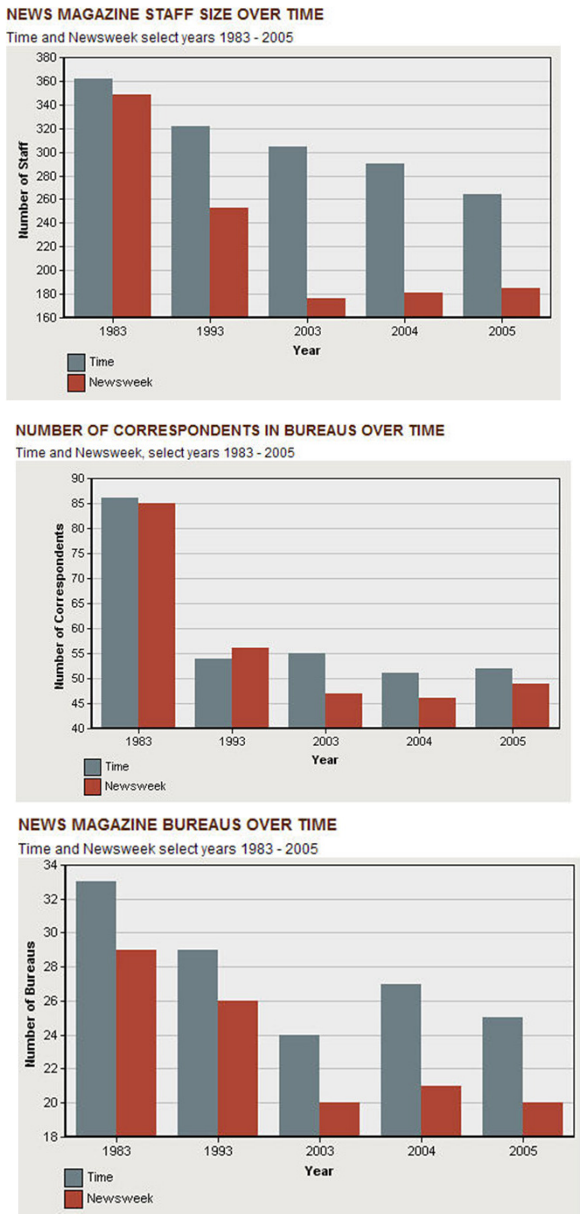
10 years

1 year

Redesign (by Stephen Few)



Note: A dashed line connecting two points indicates that there are years between the points for which values were not available. If the values were available, the shape of the lines might vary significantly.

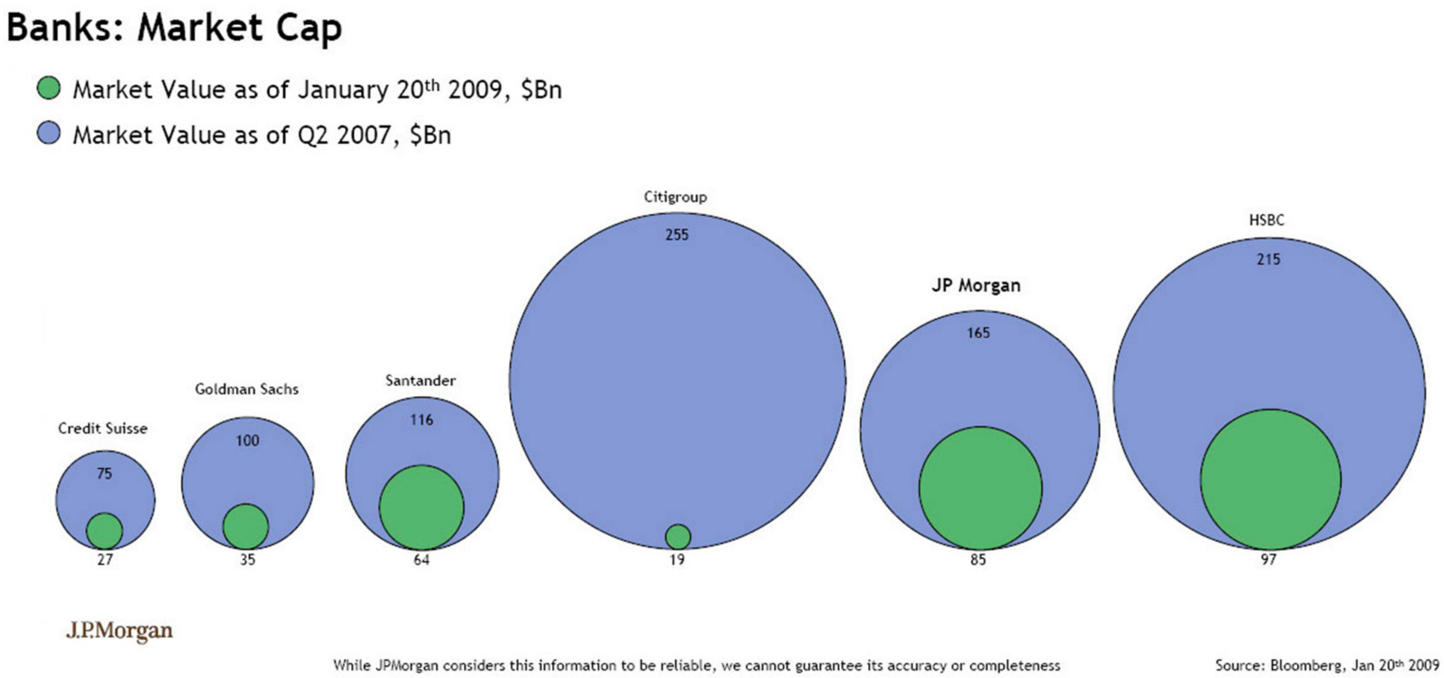


Many more examples and critiques can be found:

- In Stephen Few's books
- on his website http://www.perceptualedge.com (where the previous examples are from)
- In Edward Tufte's books
- http://junkcharts.typepad.com/junk_charts/

Exercise

- ~20mins
- Re-design this graph
- Keep in mind what people would want to learn from this graph
- Provide brief reason for encoding



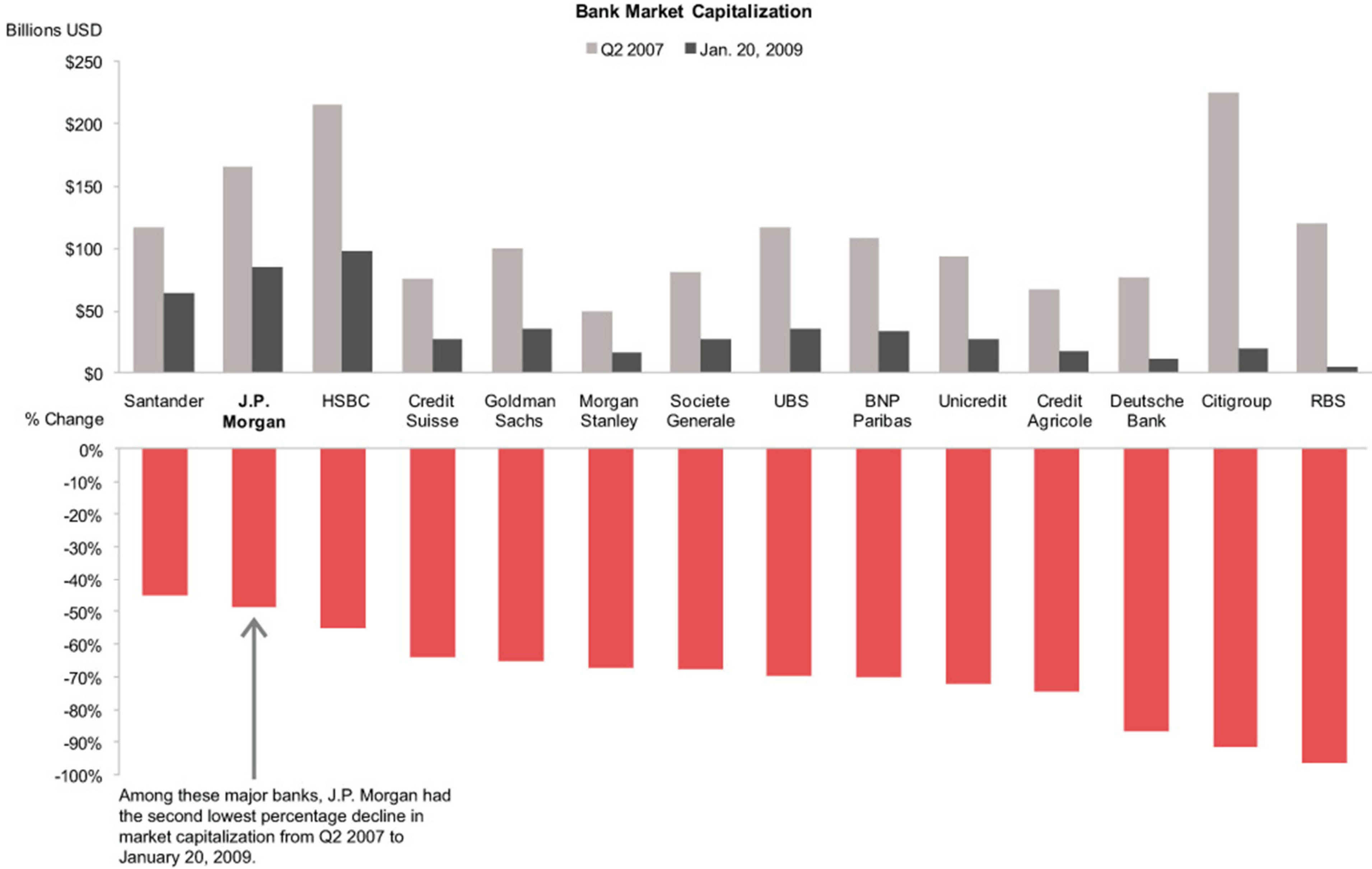
Banks: Market Cap

J.P.Morgan

While JPMorgan considers this information to be reliable, we cannot guarantee its accuracy or completeness

Source: Bloomberg, Jan 20th 2009

Stephen Few's Redesign



Short Break

- Summary
  - You saw that the vision system is **quicker and better** at detecting certain visual variables
  - Learned about the **effectiveness** of variables for encoding certain types of data
  - Learned to **critique basic data encodings**
- Next
  - Well known visualization techniques

INFORMATION VISUALIZATION TECHNIQUES

1D DATA

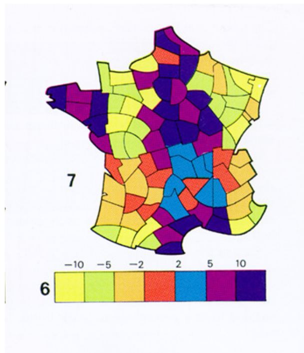
We have already talked about…

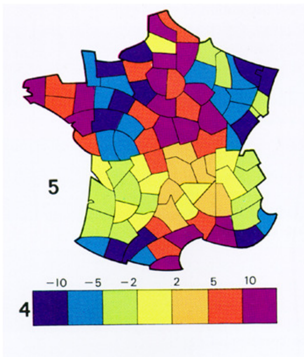
- How to encode 1D values with
  - length
  - position
  - value
  - size
  - texture
  - color
  - …

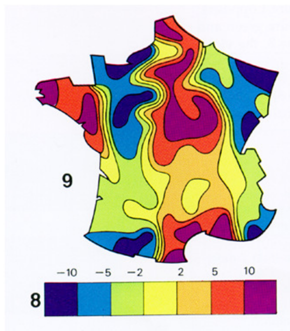
Let's look a bit more in depth at…

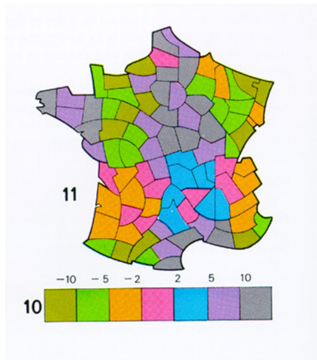
- Color Scales

- So far you know:
  - Rainbow scales can be very harmful
  - Value changes in color are important

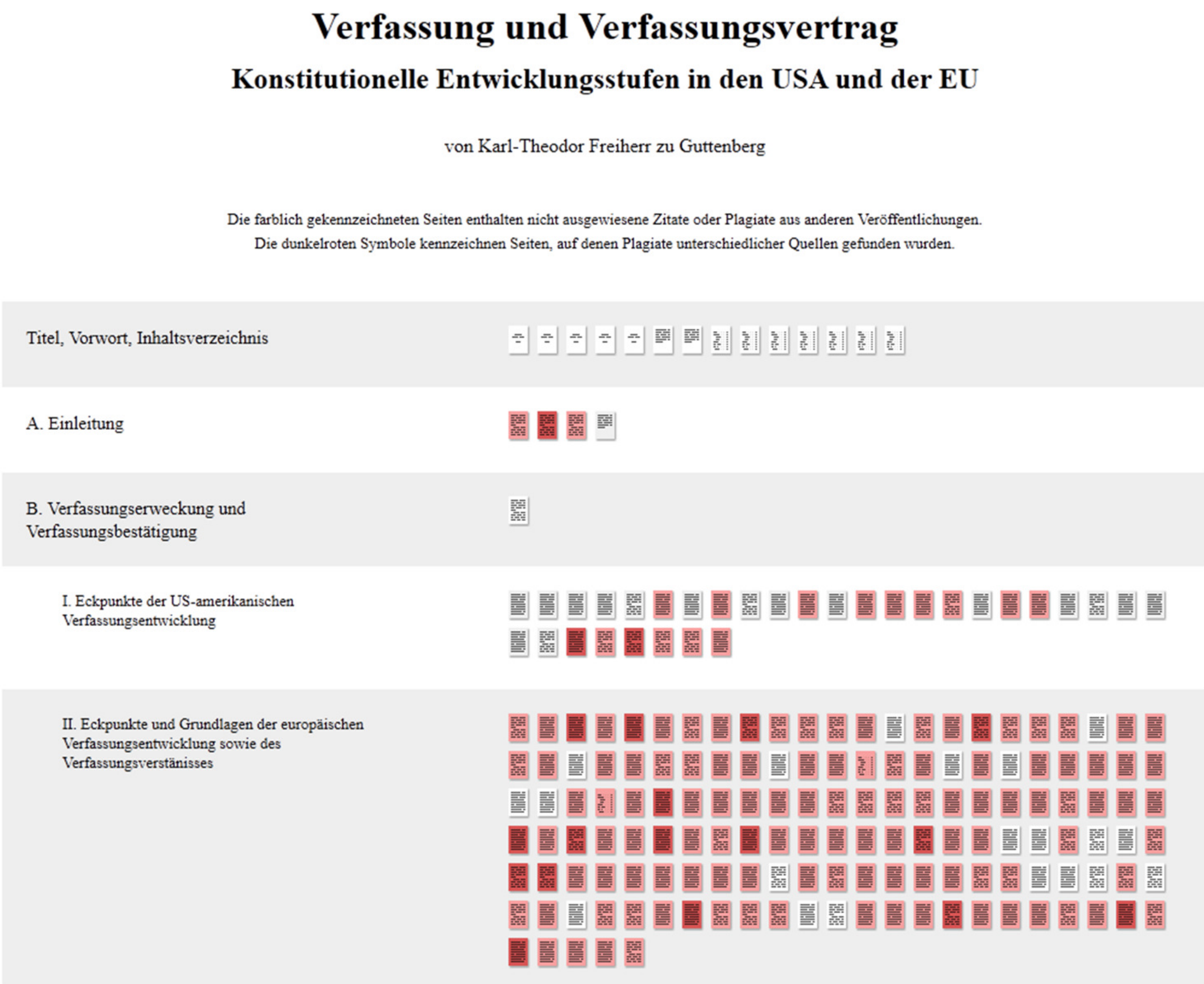








Color for Quantitative, Nominal, Ordered Data



Here:
Visualization of pages in the Doctoral thesis of Germany's former Defense Minister

page with non-cited text from one source

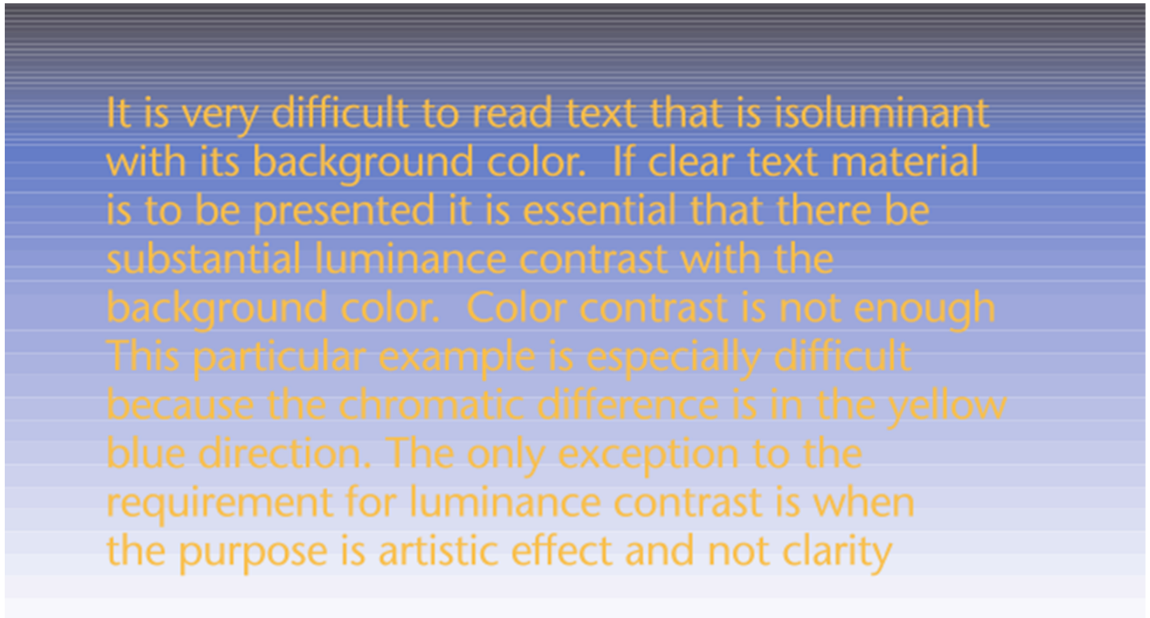
non-cited text from several sources

Visualization helped to show how much of the thesis was plagiarized

http://labs.vis4.net/guttenplag/

More general rules about color

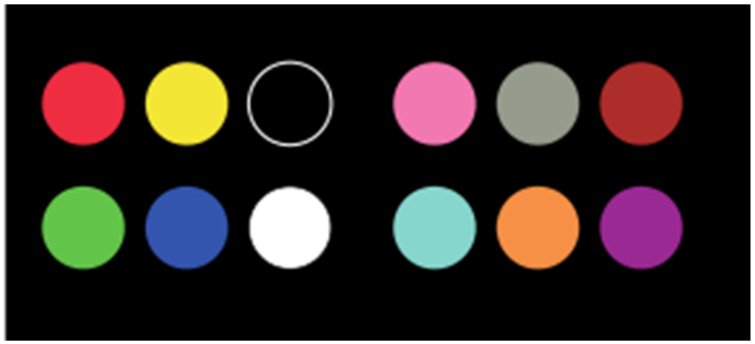
- Always have high luminance contrast between foreground and background



From Ware, Information Visualization

More general rules about color

- Always have high luminance contrast between foreground and background
- Use only few distinct colors

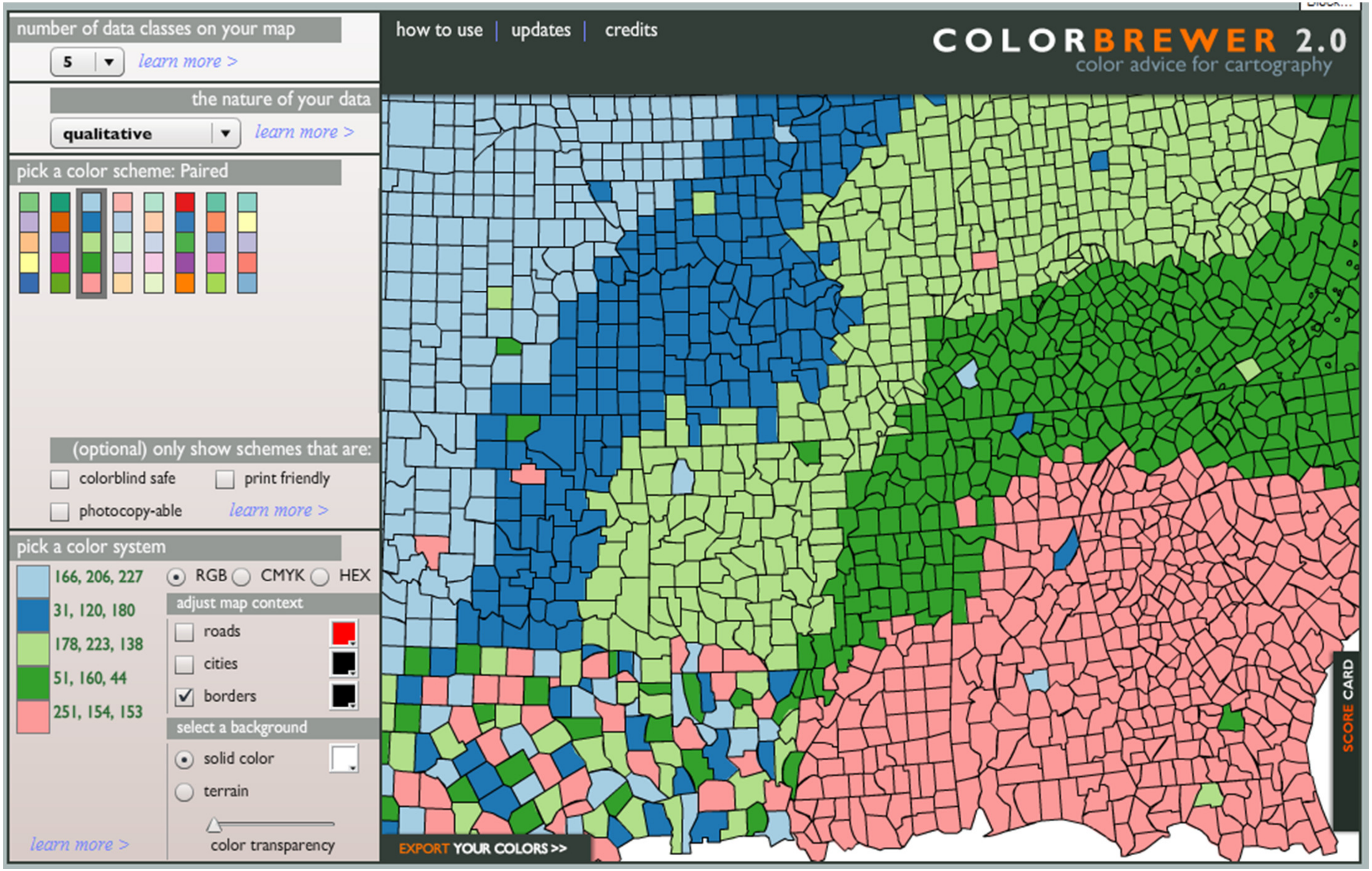


- > 12 colors will likely not work
- ~ 5 colors recommended

From Ware, Information Visualization

ColorBrewer http://colorbrewer2.org/

- Highly recommended resource
- Designed originally for maps but will also work well for other types of visualizations


number of data classes on your map
5
learn more >
the nature of your data
qualitative
learn more >
pick a color scheme: Paired
(optional) only show schemes that are:
colorblind safe
print friendly
photocopy-able
learn more >
pick a color system
166, 206, 227
31, 120, 180
178, 223, 138
51, 160, 44
251, 154, 153
RGB
CMYK
HEX
adjust map context
roads
cities
borders
select a background
solid color
terrain
color transparency
learn more >
how to use | updates | credits
COLORBREWER 2.0
color advice for cartography
EXPORT YOUR COLORS >>
SCORE CARD

More general rules about color

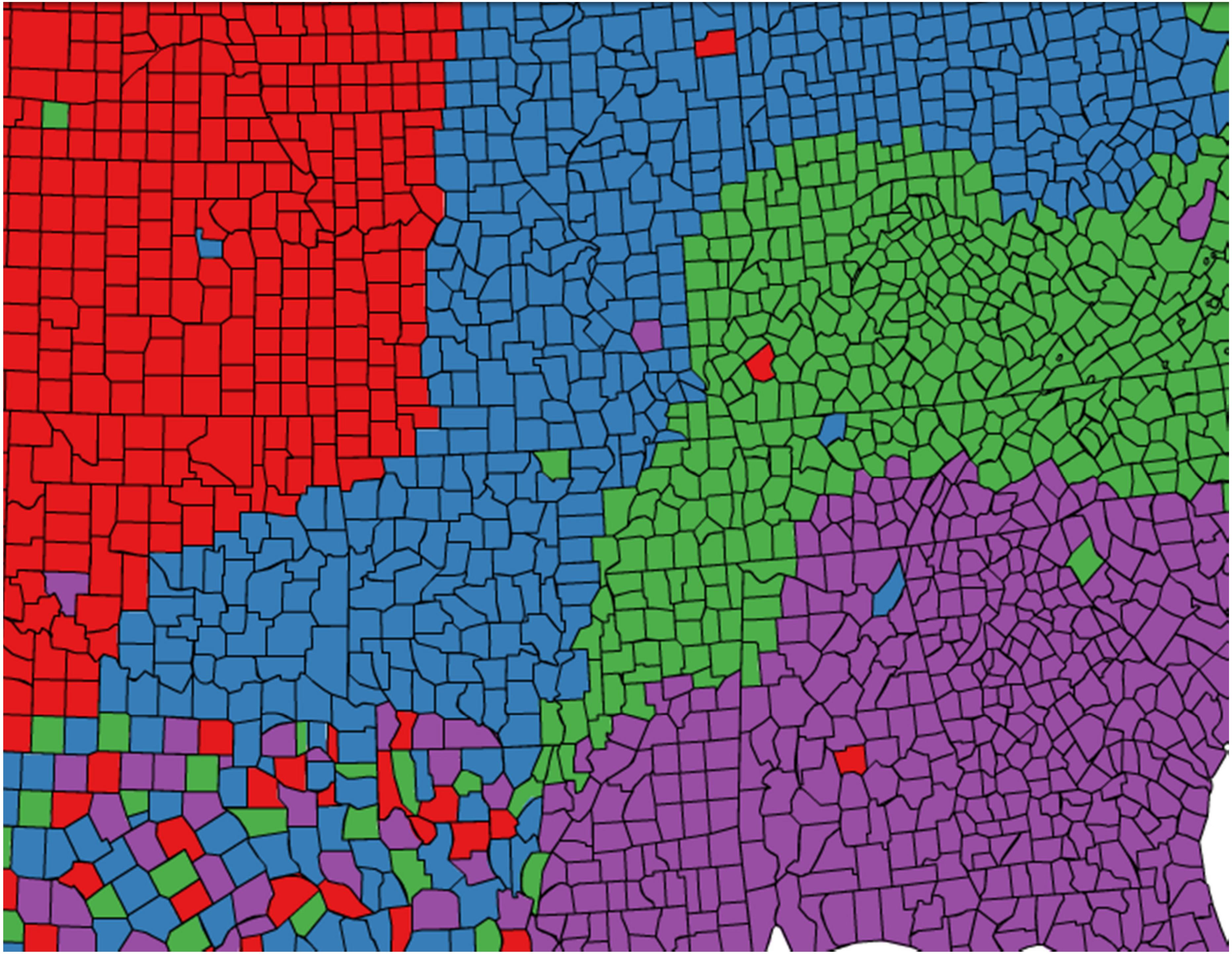
- Always have high luminance contrast between foreground and background
- Use only few distinct colors
- Red, green, yellow, blue are hard-wired into the brain. Use them first.

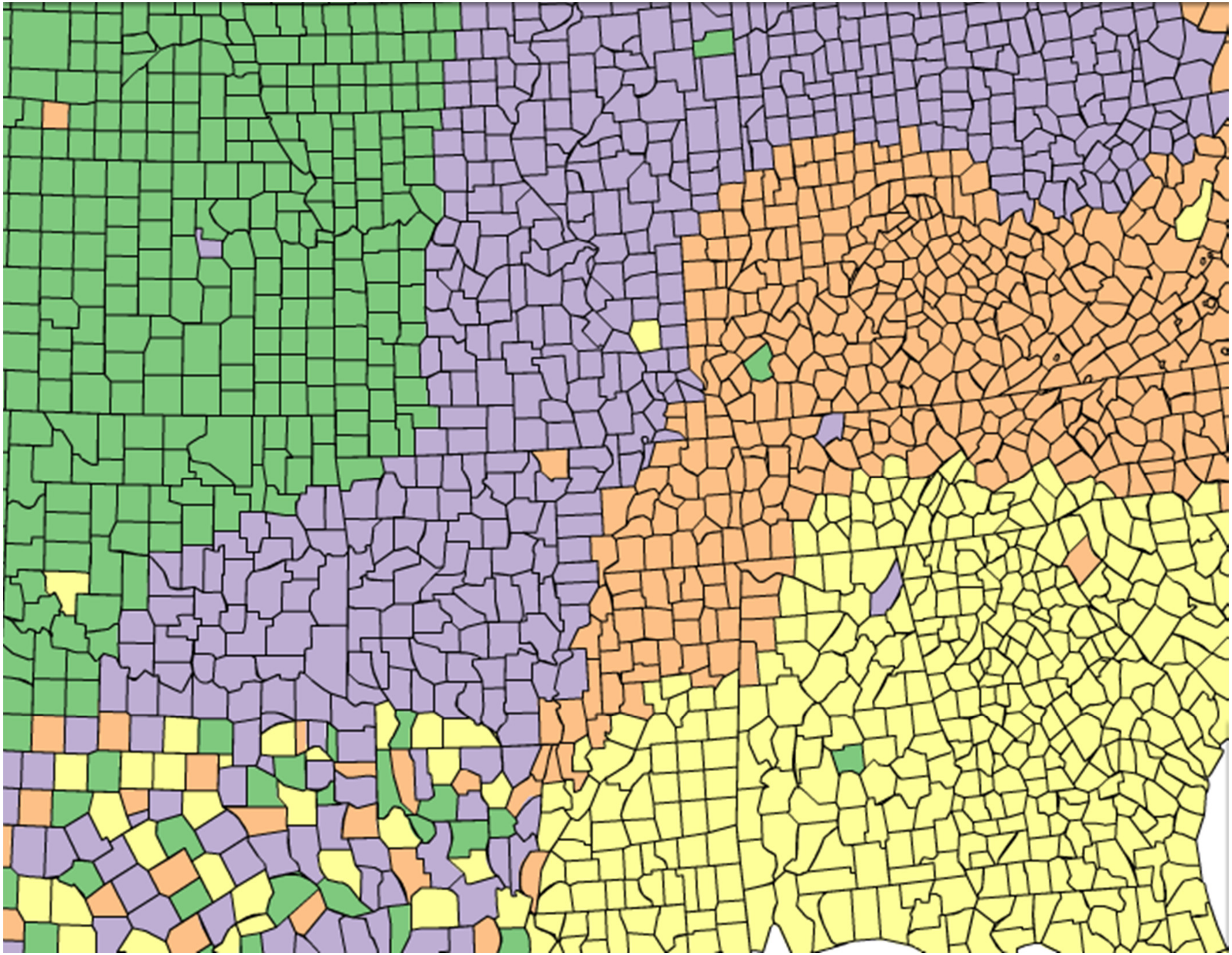
From Ware, Information Visualization

More general rules about color

- Always have high luminance contrast between foreground and background
- Use only few distinct colors
- Red, green, yellow, blue are hard-wired into the brain. Use them first.
- For large areas use muted colors

From Ware, Information Visualization



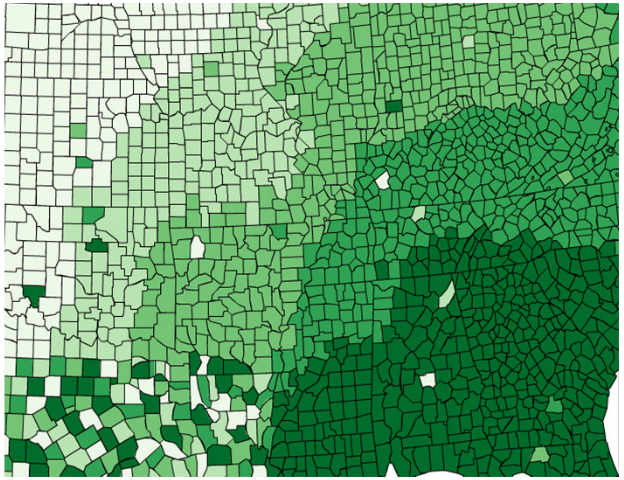


More general rules about color

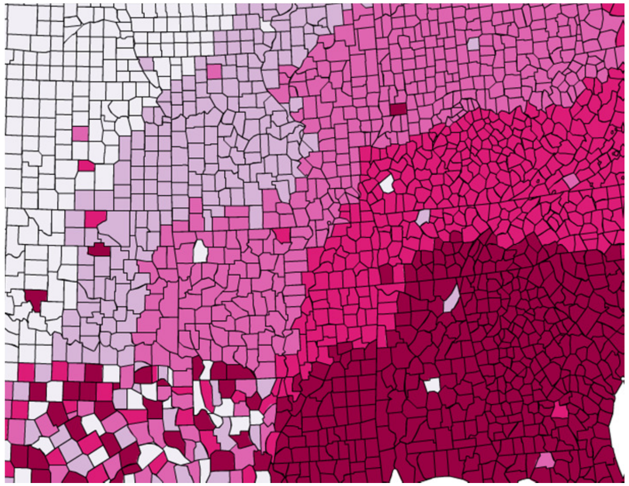
- Always have high luminance contrast between foreground and background
- Use only few distinct colors
- Red, green, yellow, blue are hard-wired into the brain. Use them first.
- For large areas use muted colors
- When an ordering is required, recommended: red-green, yellow-blue, low saturation-high saturation, dark-light

From Ware, Information Visualization

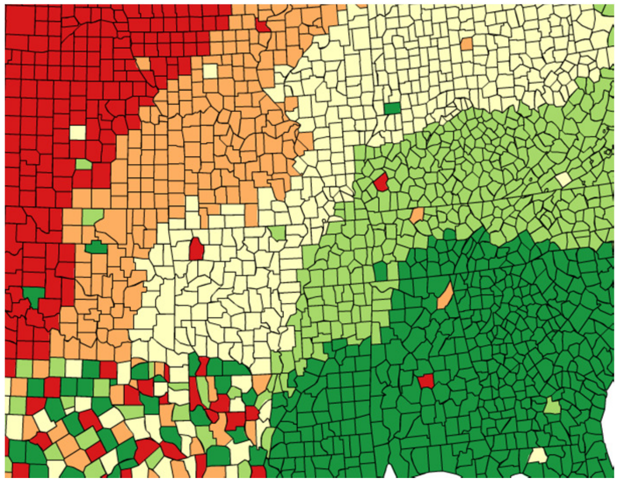
Ordered Color Scales



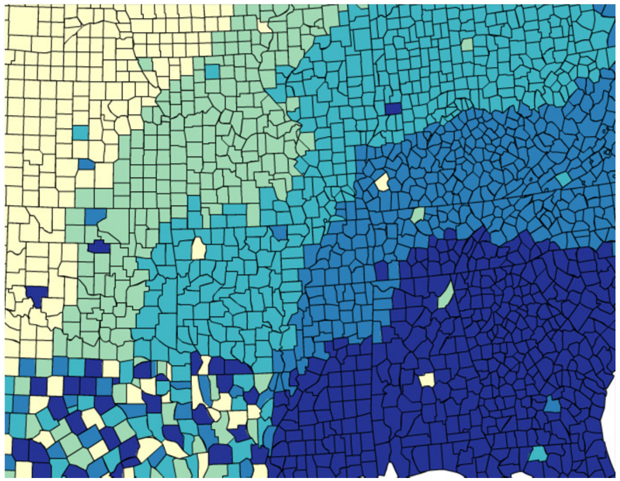
low - high saturation



dark-light



red-green



yellow-blue

If you use red-green 7% of the viewers may not see anything

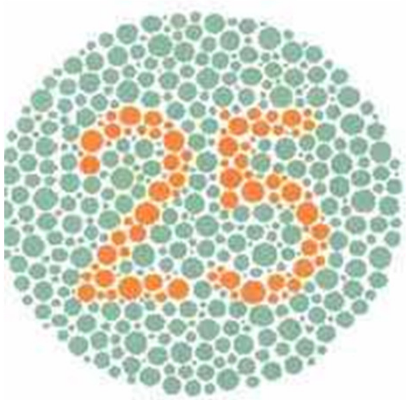
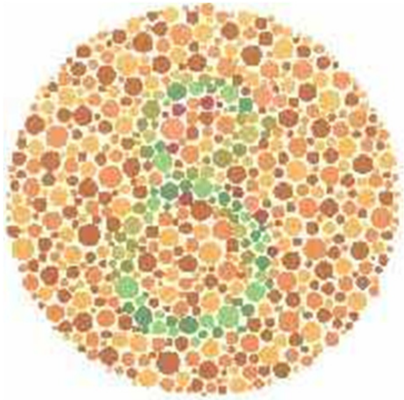
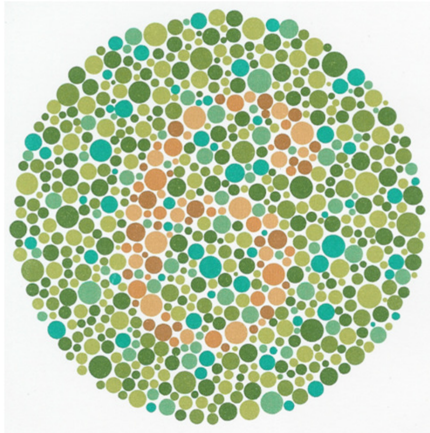
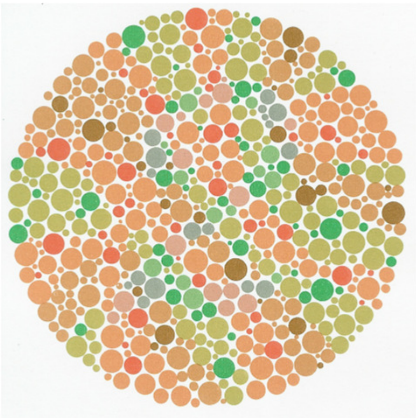
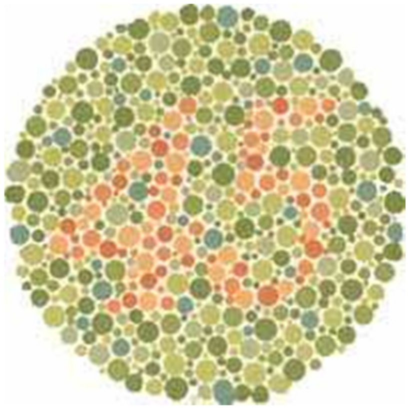
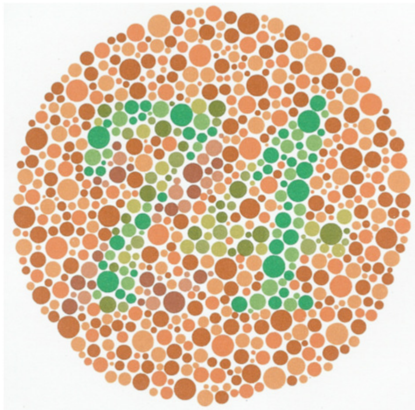
ONE WARNING ABOUT RED-GREEN

The following slides on the topic are adapted from Tobias Isenberg's

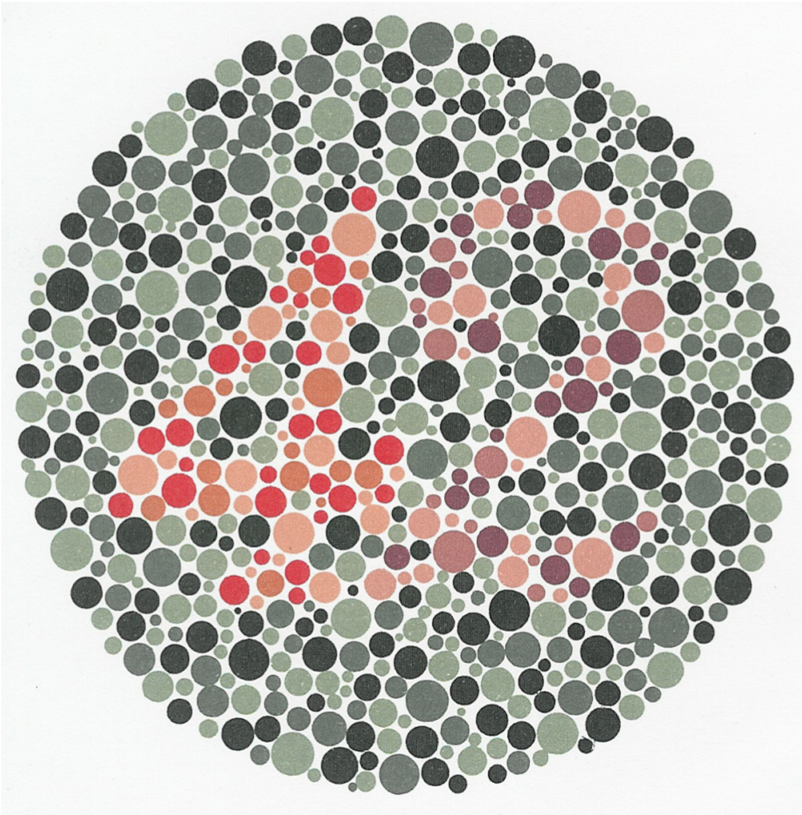
Color Deficiency

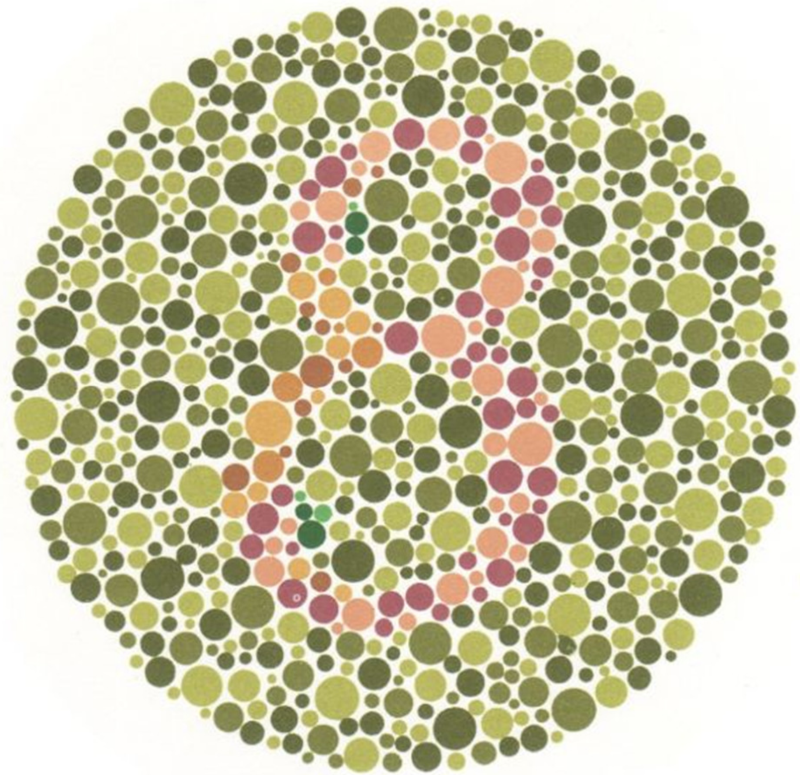
- ca. 7% of male population color-deficient
- mostly red-green color deficiency (Deuteranopy or Protanopy); other forms as well (e.g., Tritanope, very rare)
- avoid red-green color contrasts for visualization purposes!
- simulate: GIMP, http://www.vischeck.com/
- side note: there are (very, very few) people with more than three cone types

Color Deficiency Test

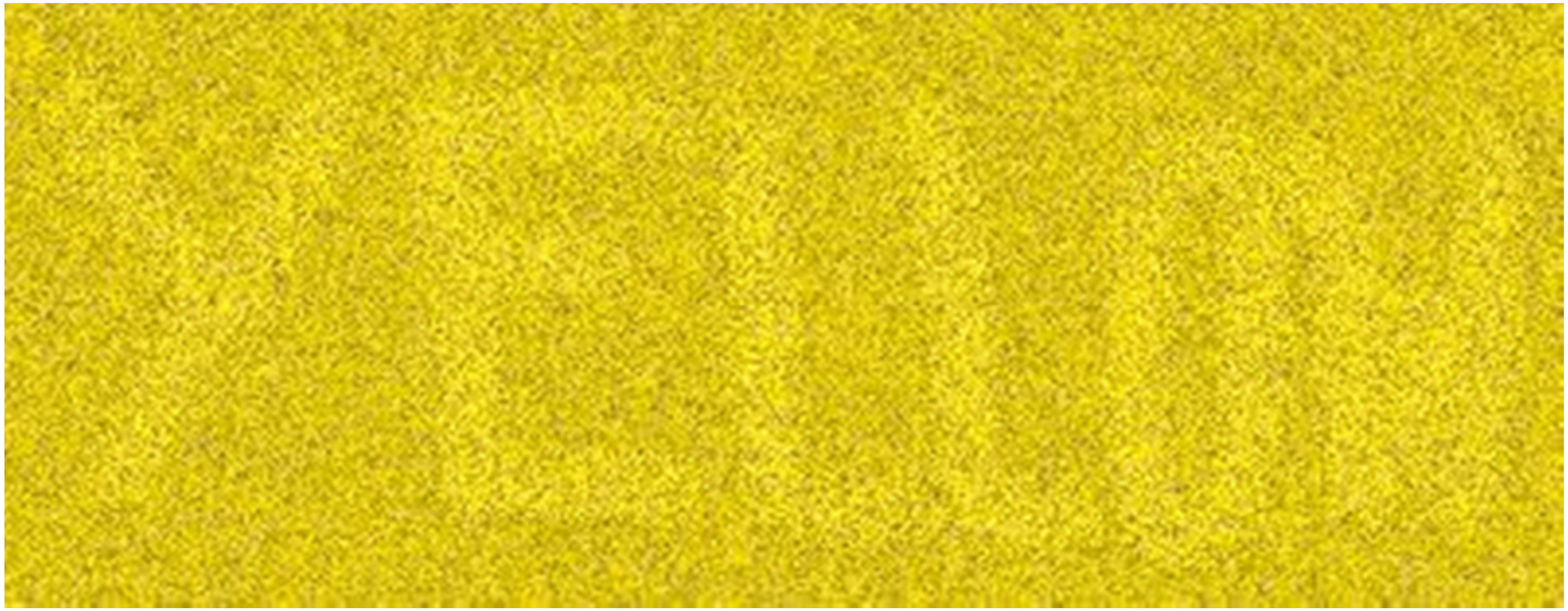







Color Deficiency Test

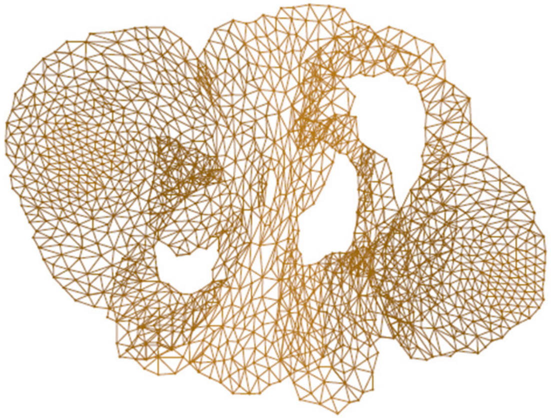
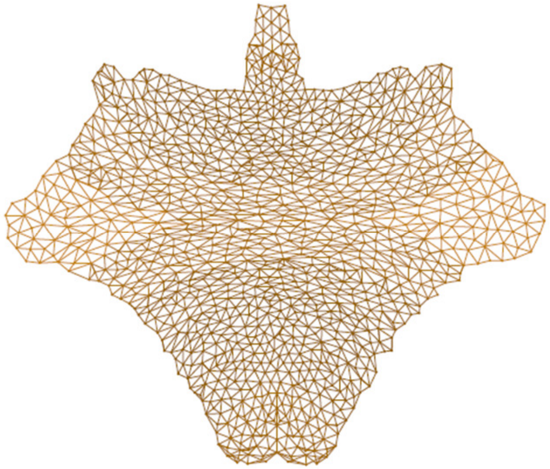
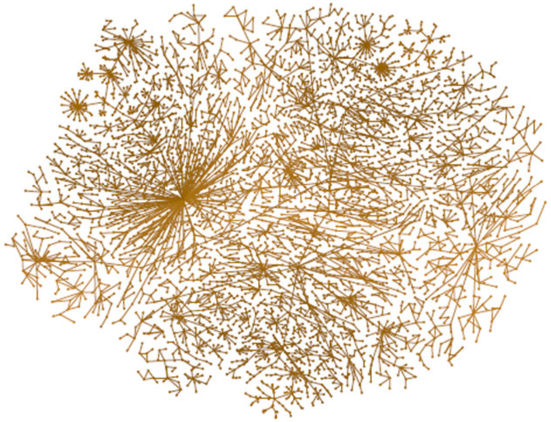
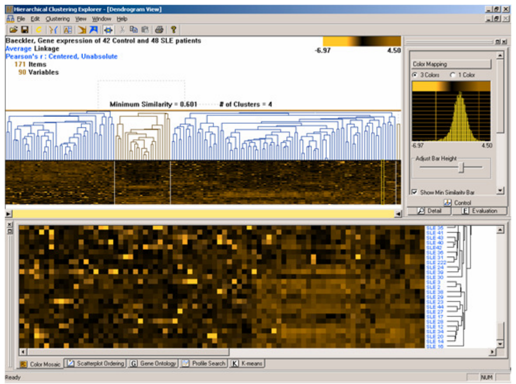
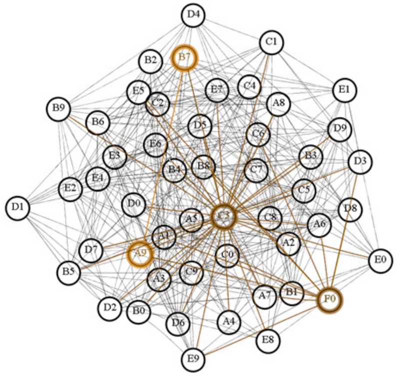
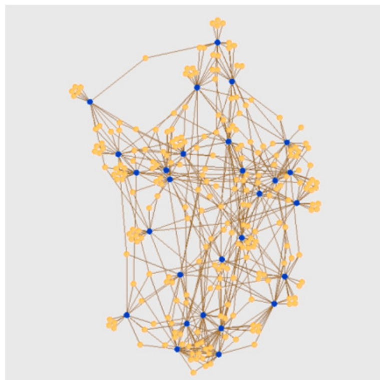




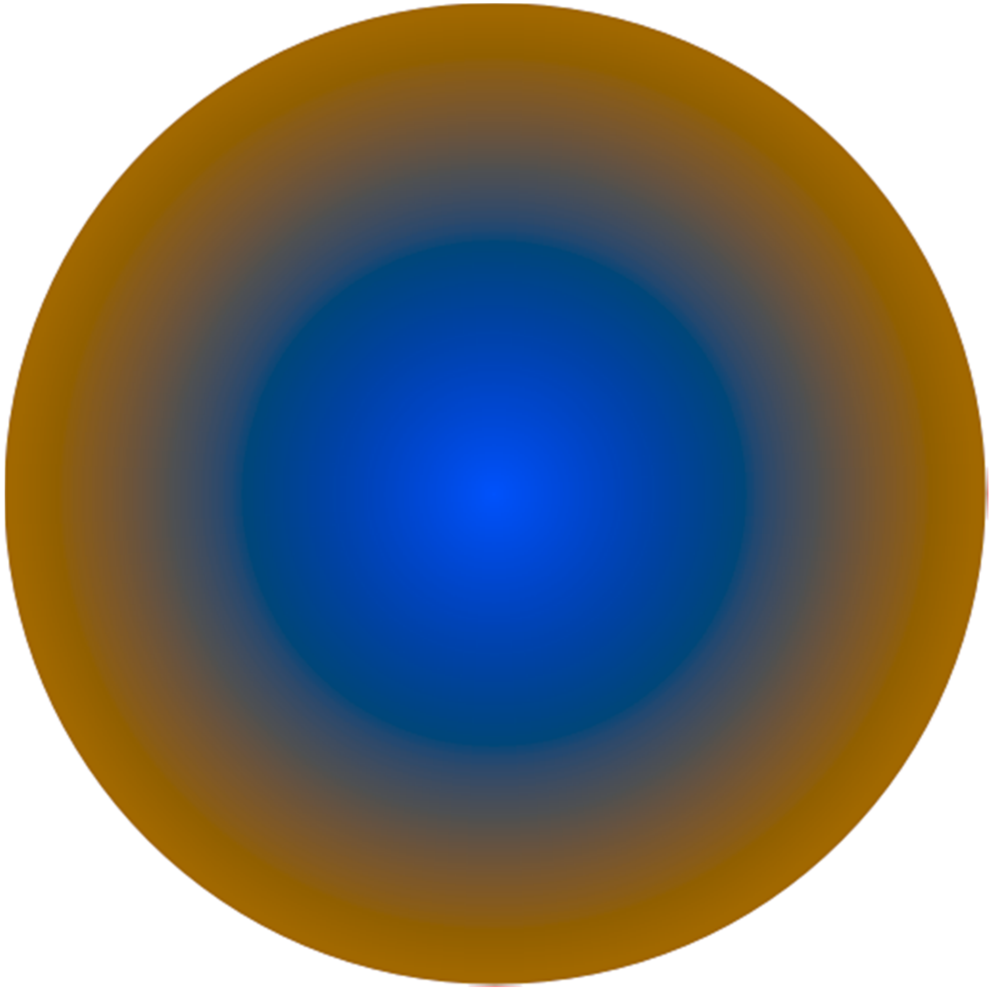
Color Deficiency



Examples from VIS/InfoVis 2004


Better: Red-Blue Contrast



More general rules about color

- Always have high luminance contrast between foreground and background
- Use only few distinct colors
- Red, green, yellow, blue are hard-wired into the brain. Use them first.
- For large areas use muted colors
- When an ordering is required, recommended: red-green, yellow-blue, low saturation-high saturation, dark-light
- **… there exist many many more specialized rules (when you get more advanced…)**

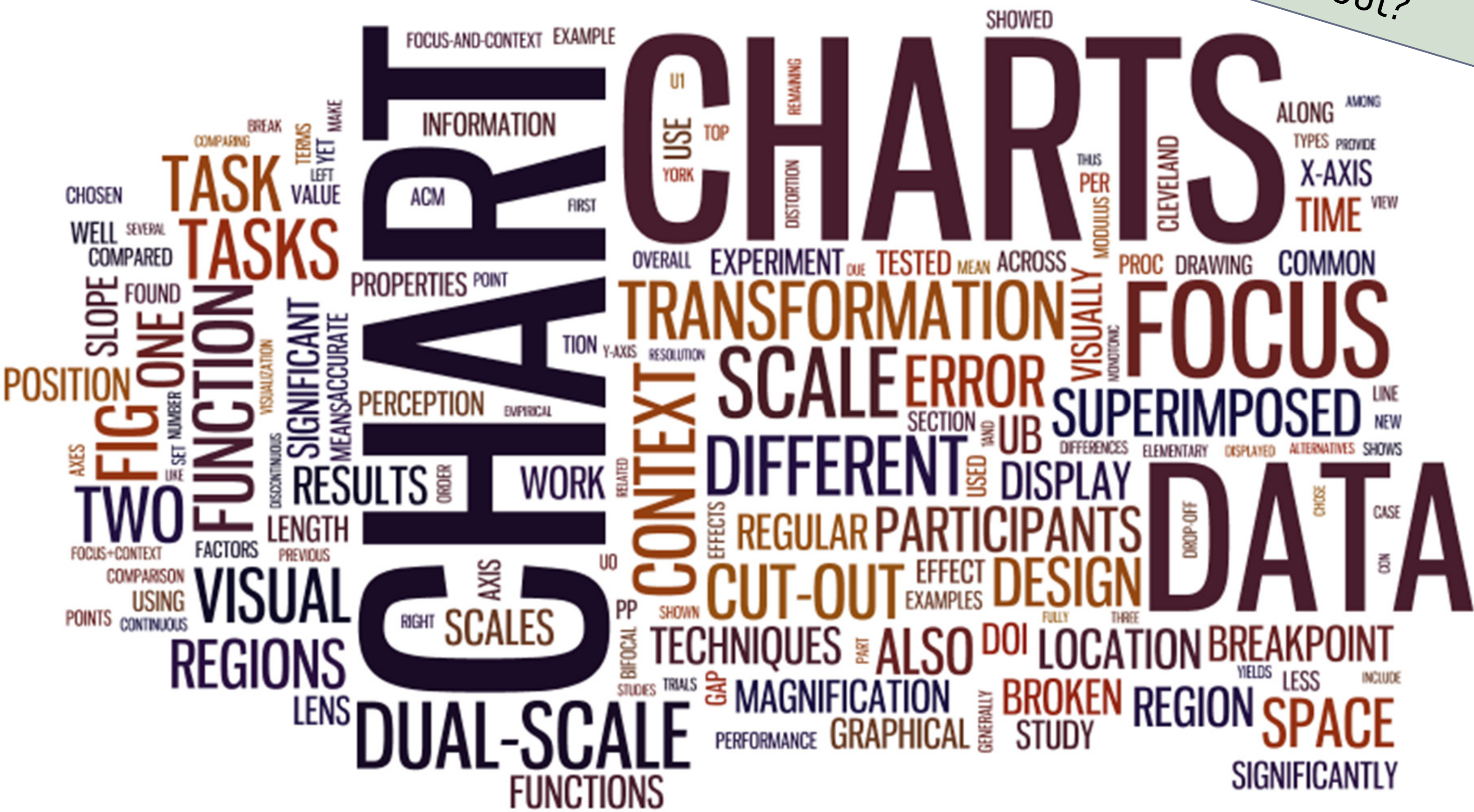
From Ware, Information Visualization

Visualization techniques for text and a few others for

LINEAR DATA

Tag Clouds

My recent InfoVis paper
What is it about?

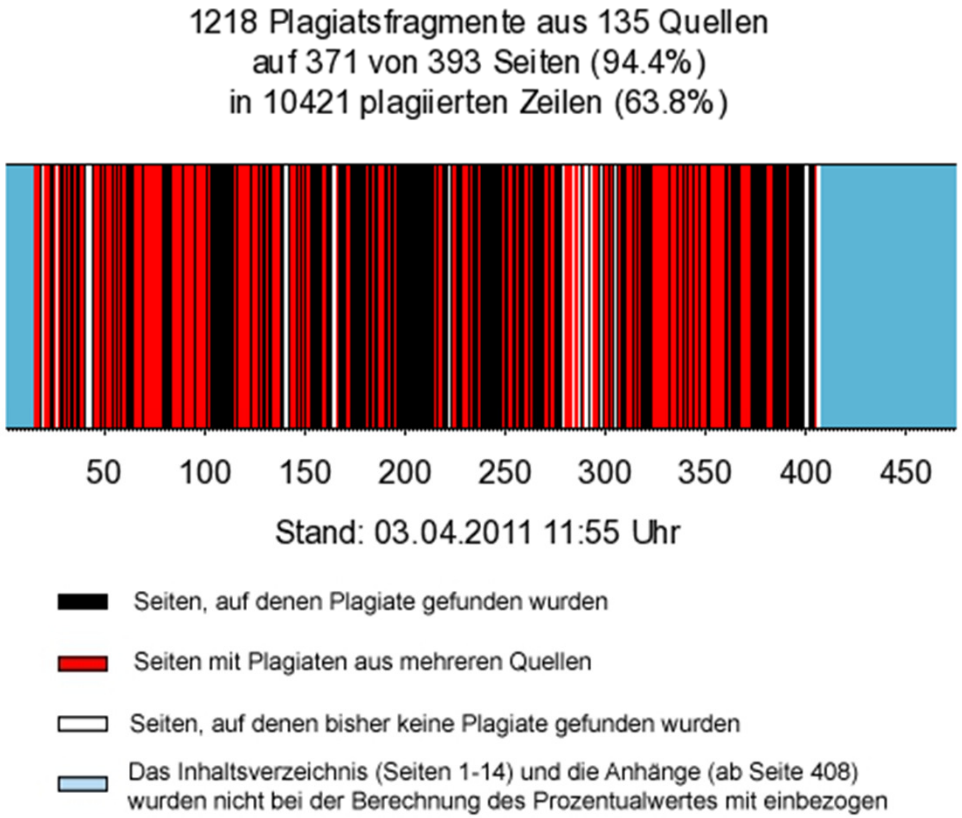


Visualizes word count
Color & position have no meaning
Size encodes count

http://www.wordle.net/

Other Vis from the Guttenberg Plagiarism Wiki

- Very simple color+order visualization



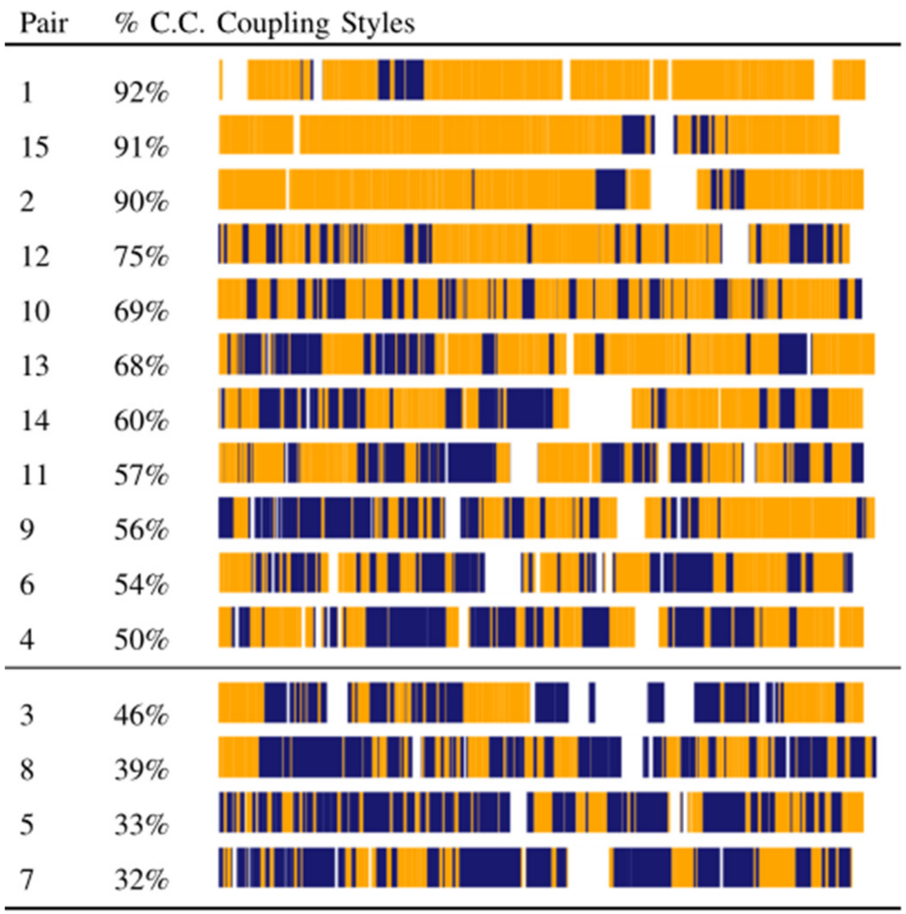
Note: I do not think the colors are well chosen

pages with plagiarism
pages with plagiarism from >1 source
pages without plagiarism
pages not counted in % plagiarized
e.g. TOC & Appendix

http://de.guttenplag.wikia.com/wiki/GuttenPlag_Wiki/English

More examples

- Useful for comparison

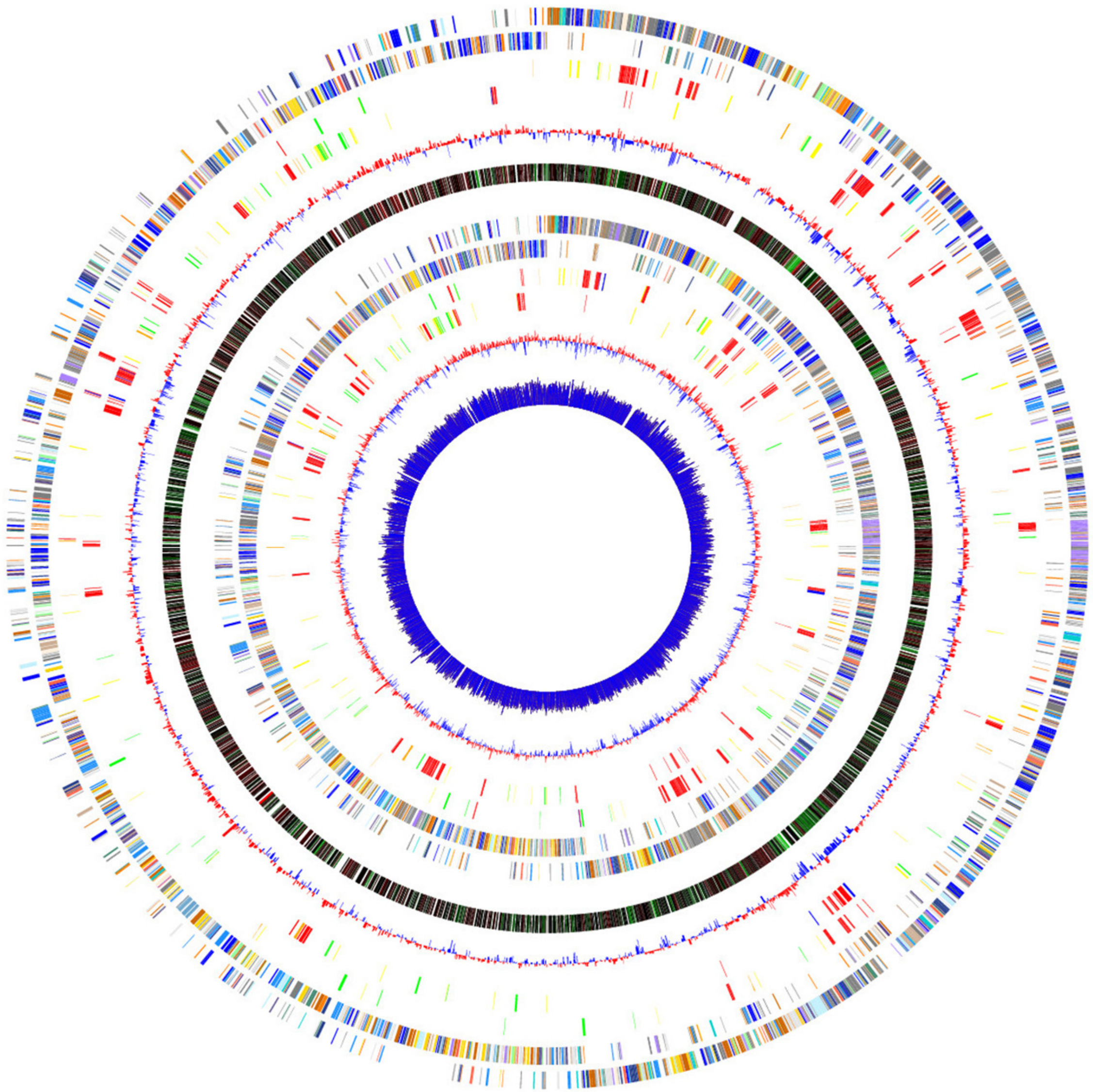


E.g. to compare the behavior of groups over time

<- according to two main behavior types (one in blue and one in yellow)

Isenberg et al., TVCG 2011, to appear

GenomeViz



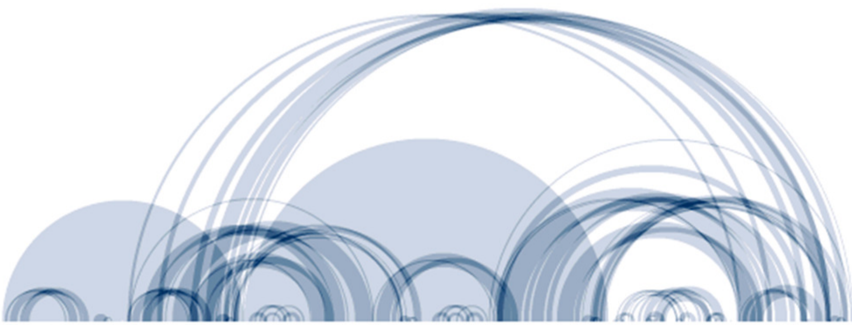
- Comparison of several 1D viz in one view
- Genomic data makes sense in a circle

Ghai et al. 2004

ArcDiagrams

Figure 1. Connecting an essential matching pair

Figure 4. Repetition of different substrings


Chopin, Mazurka in F# Minor
The image illustrates the complex, nested structure of the piece.


Madonna, Like A Prayer
Pop music has its own style of repetition.

http://turbulence.org/Works/song

ArcDiagrams

- Very old idea applied to current data

Salinas, De musica, 119

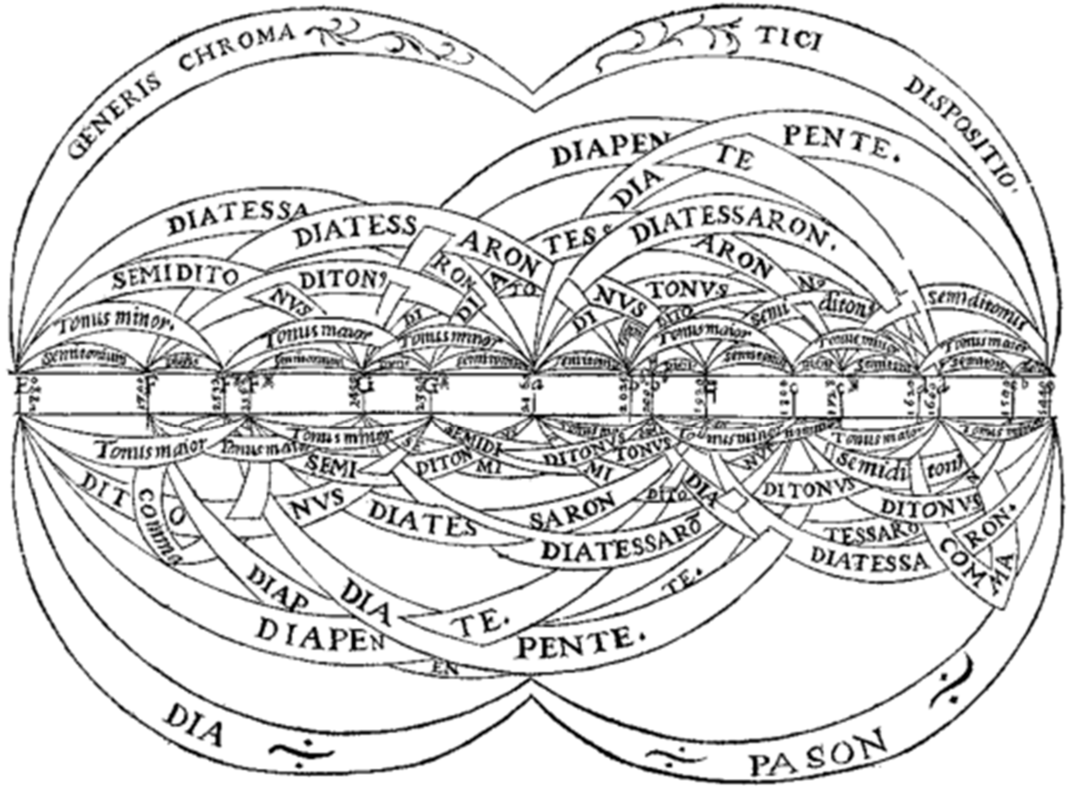


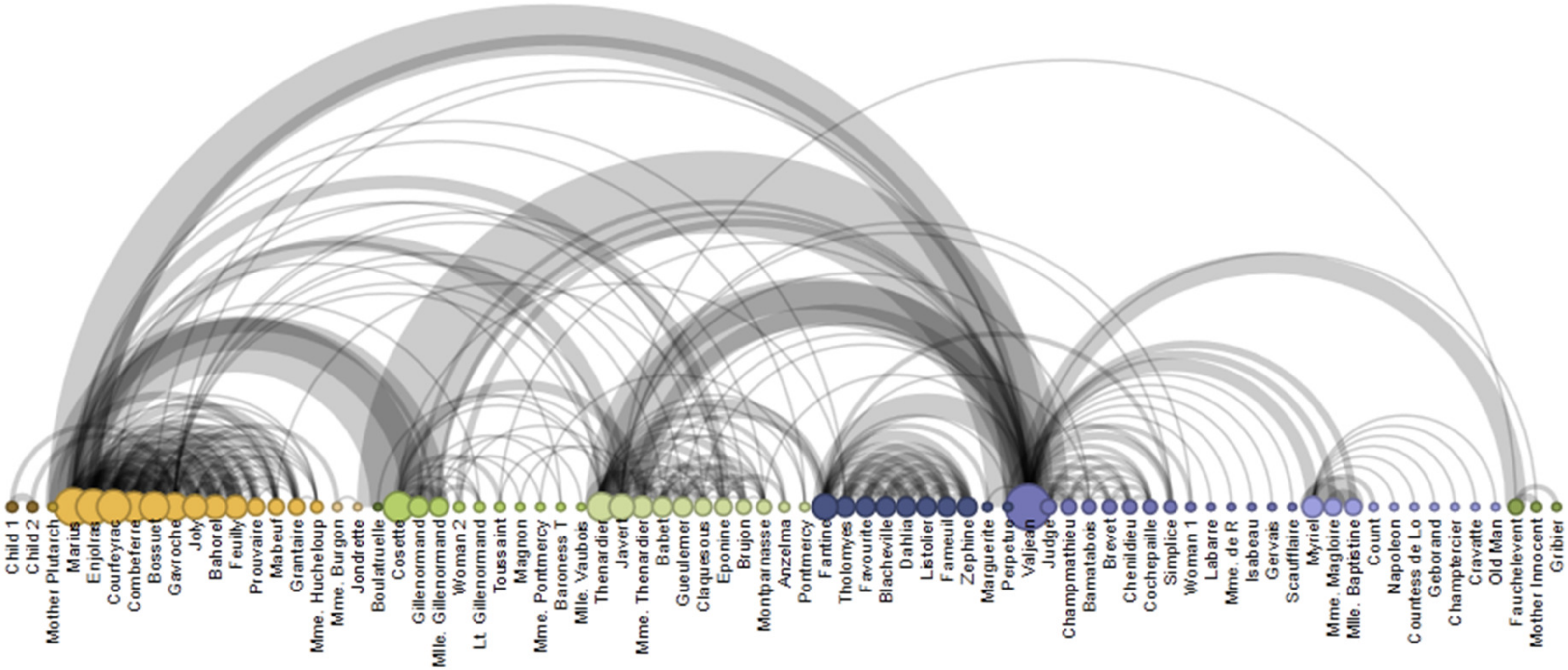
Figure 2.17: Historical diagram of a meantone tuning taken from (DE SALINAS, 1577).

ArcDiagrams

- Can be easily extended

character co-occurrence in the chapters of Victor Hugo's classic novel, *Les Misérables*

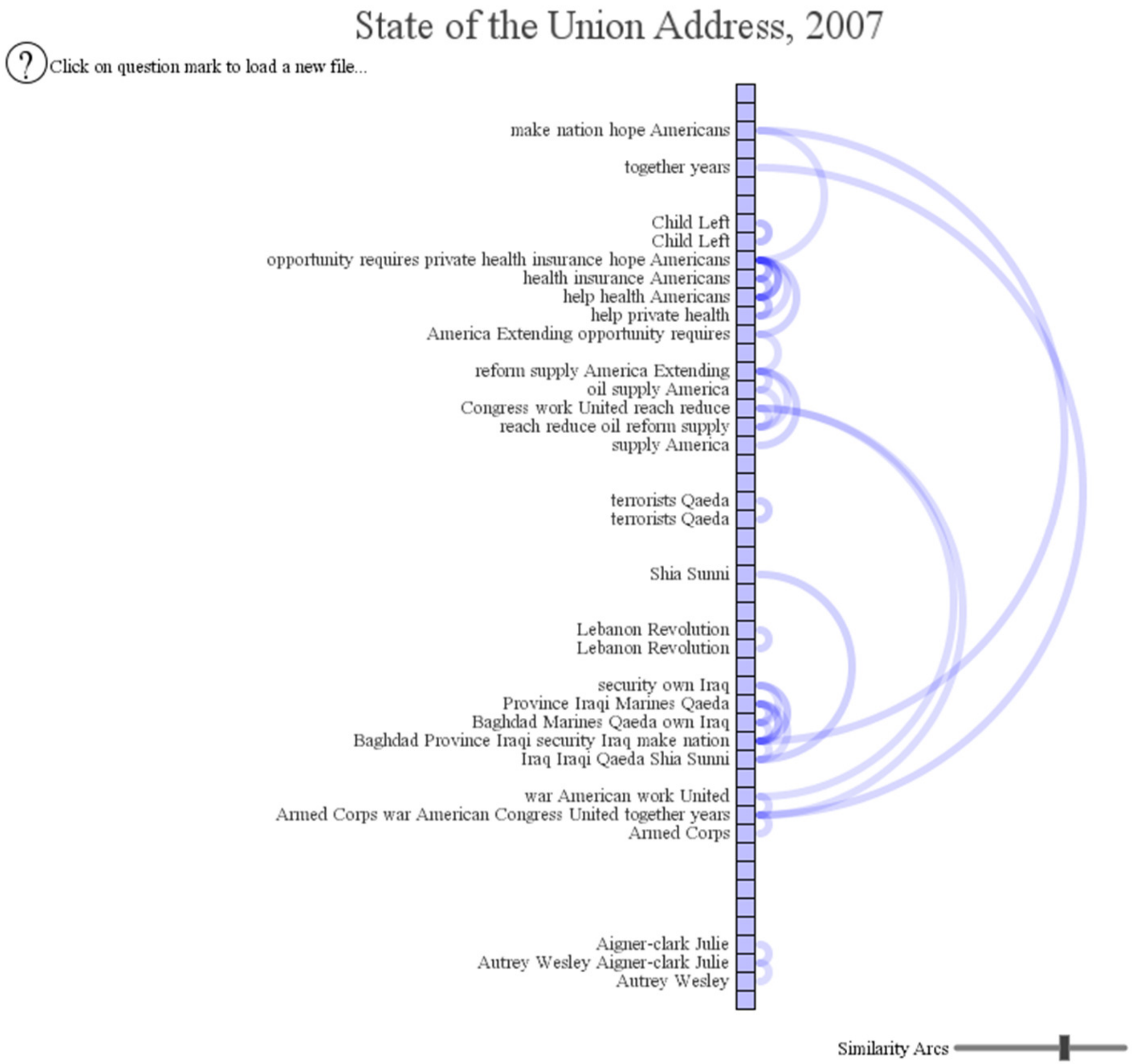
Arc Diagrams



Note that this is actually a graph data structure.
(More on graphs later)

http://mbostock.github.com/protovis/ex/arc.html

DocumentArcDiagrams



Similarity within a text document

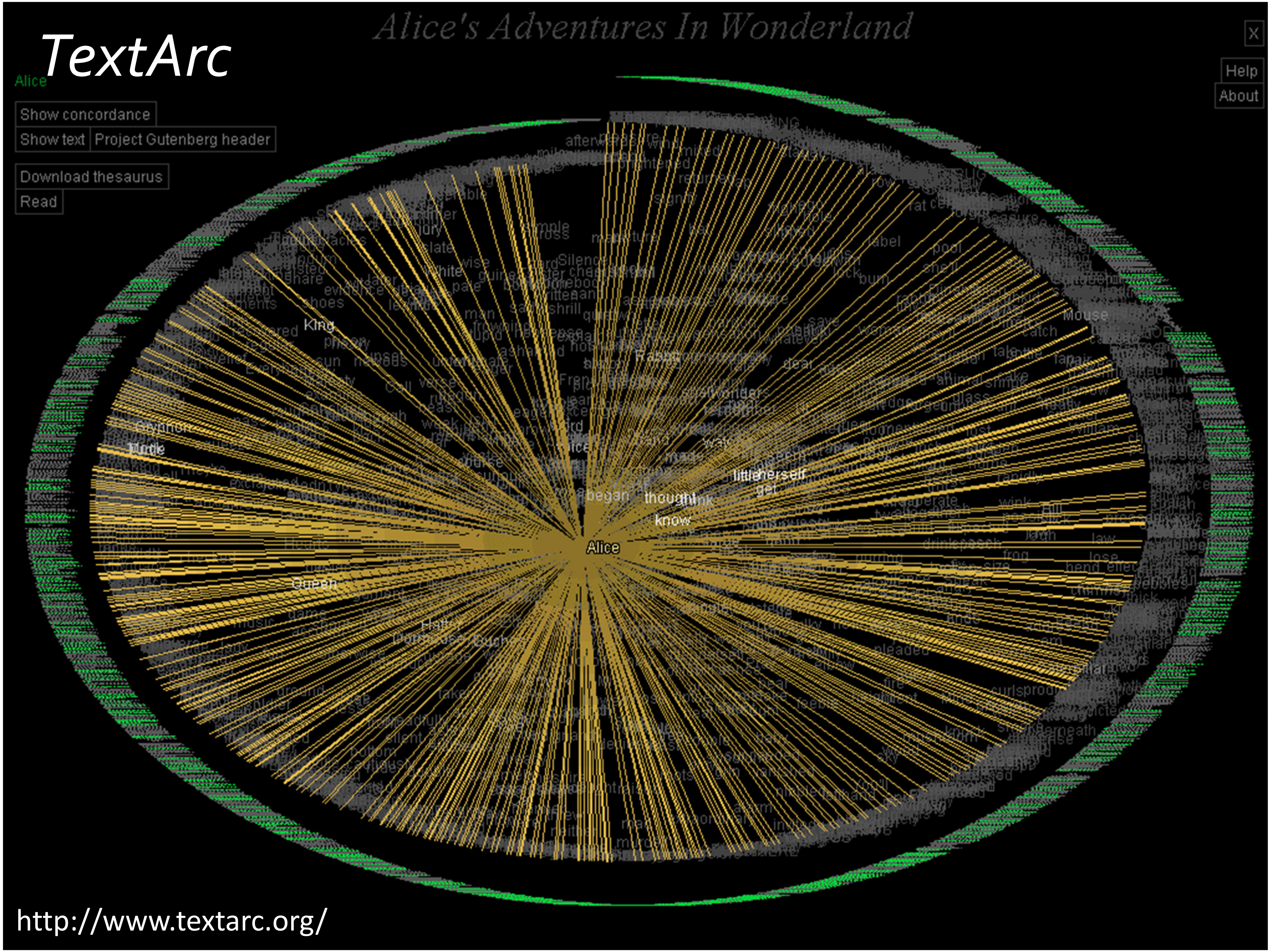
ThreadArcs

Kerr 2004

TextArc

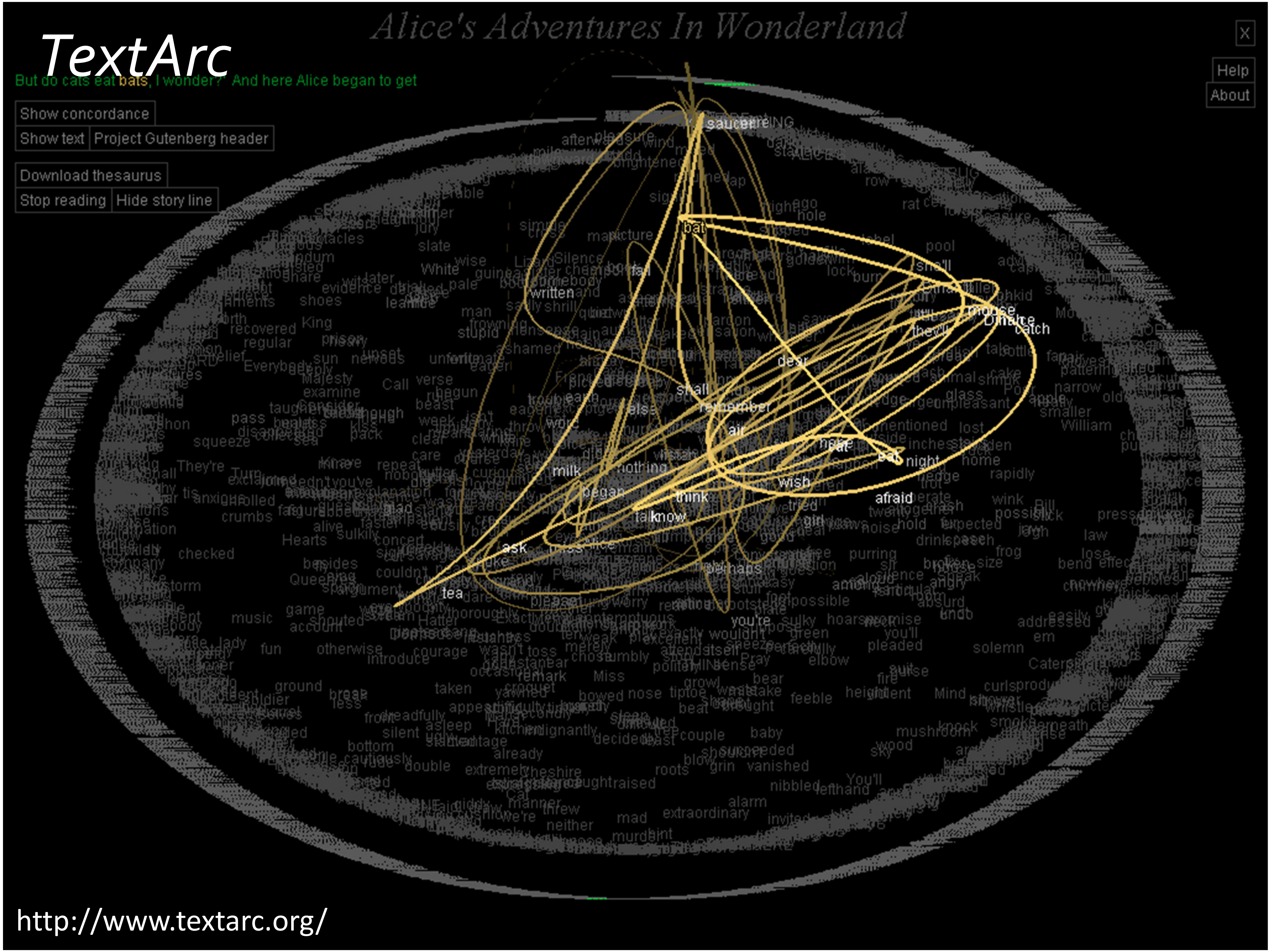
http://www.textarc.org/

TextArc



http://www.textarc.org/

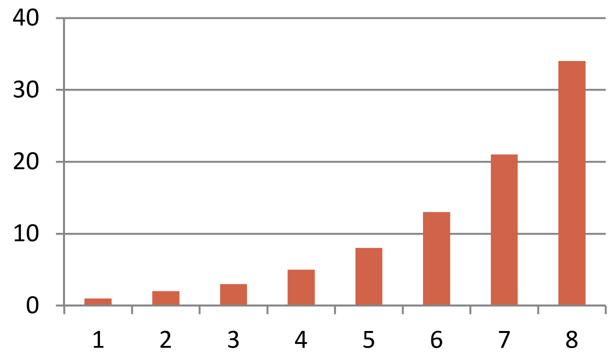
TextArc

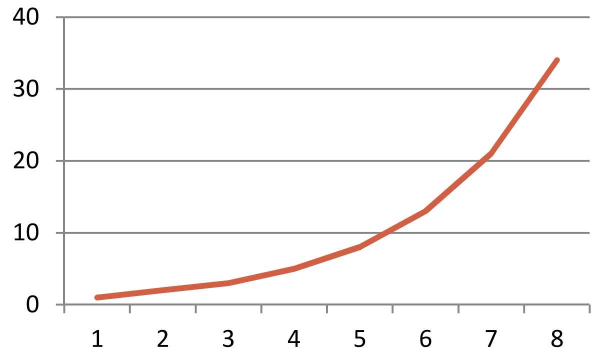


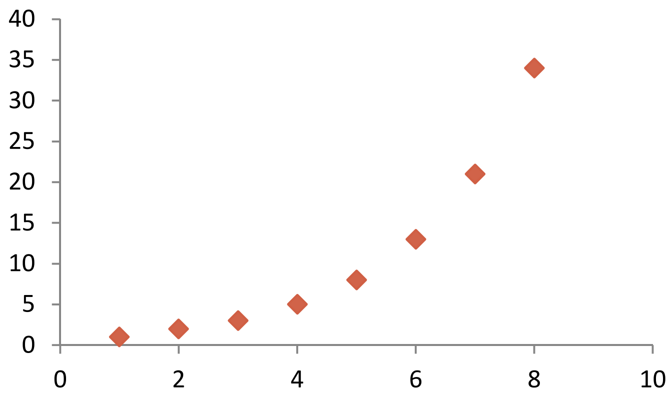
http://www.textarc.org/

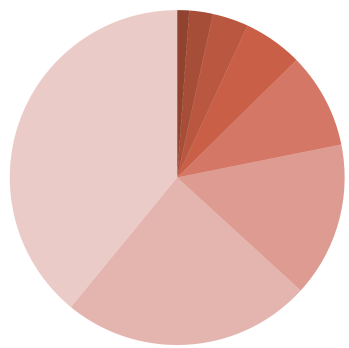
Also many diagrams can show 1D data

- Bar Charts
- Line Charts
- Pie Charts
- Dot Charts
- …

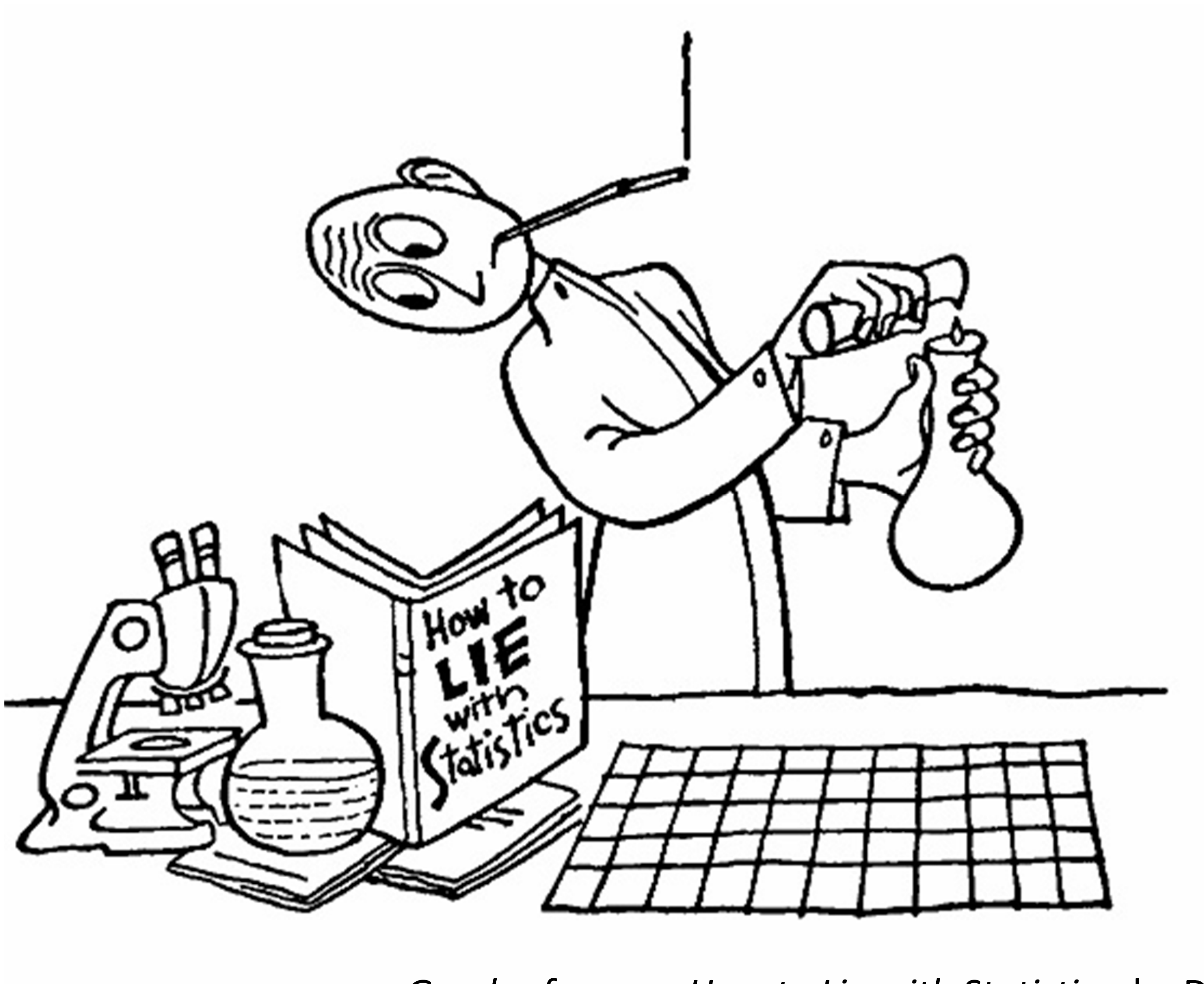








A few more words on charts



Good reference: How to Lie with Statistics, by Darrell Huff

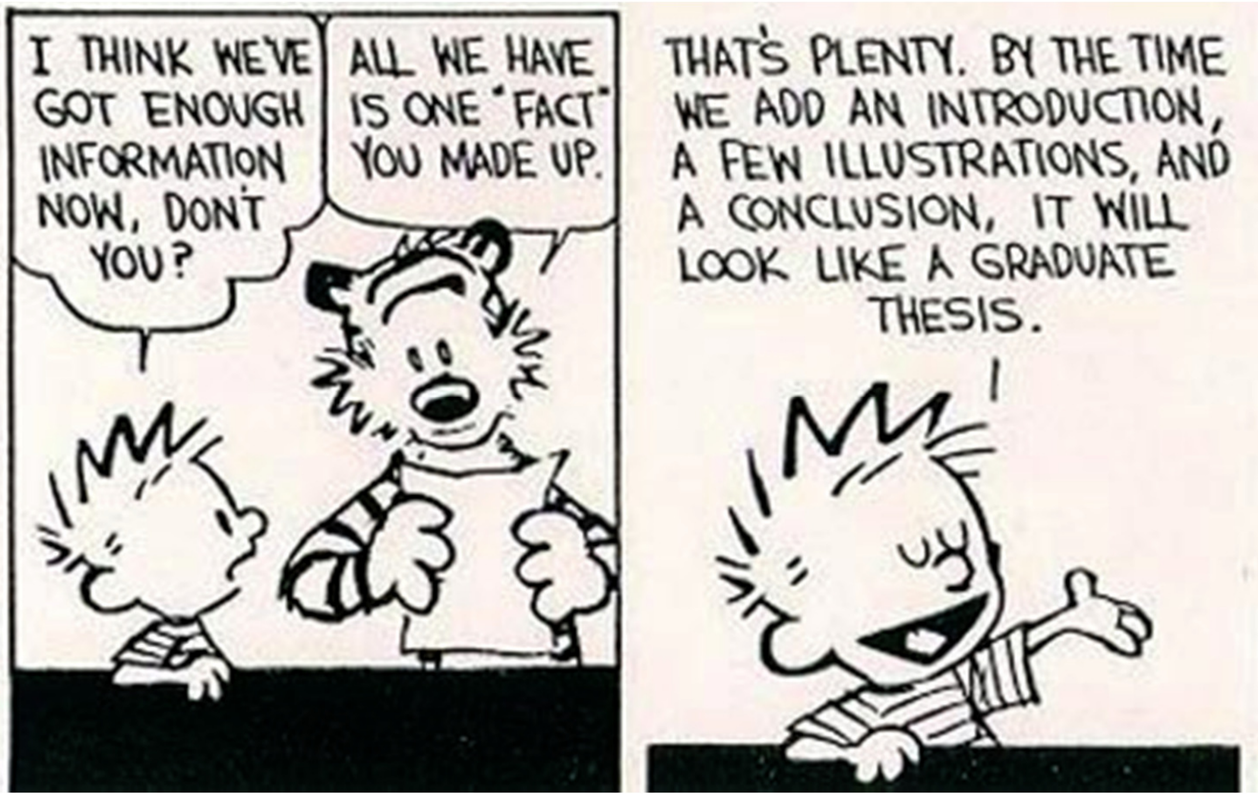
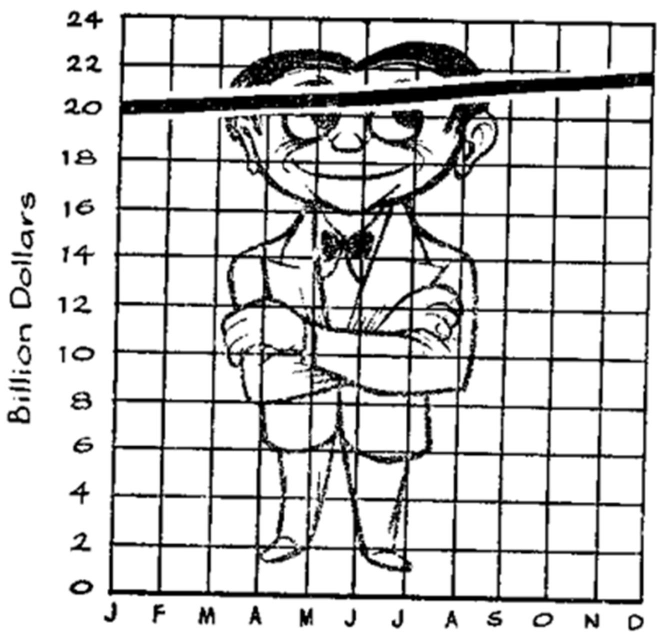

I THINK WE'VE GOT ENOUGH INFORMATION NOW, DON'T YOU?
ALL WE HAVE IS ONE "FACT" YOU MADE UP.
THAT'S PLENTY. BY THE TIME WE ADD AN INTRODUCTION, A FEW ILLUSTRATIONS, AND A CONCLUSION, IT WILL LOOK LIKE A GRADUATE THESIS.

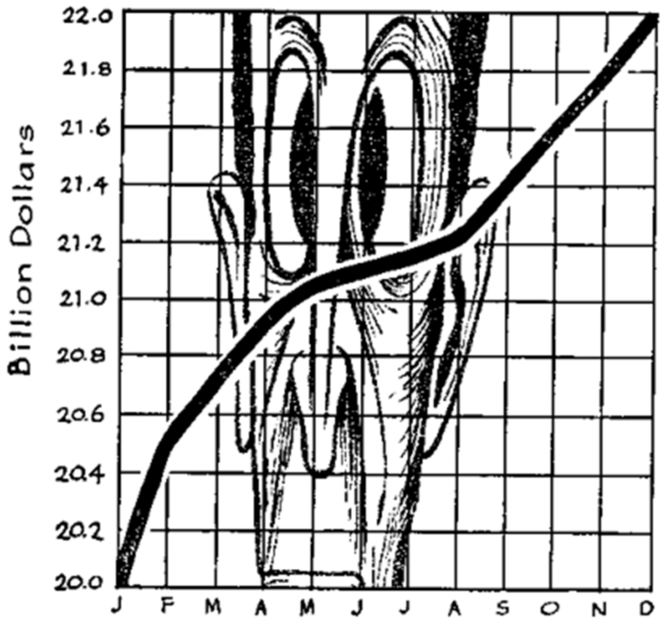
Chart Rules

- Provide a proper baseline



A 10% increase. Good!

Already looks more impressive



Wow!

Chart Rules

- Provide a proper baseline & label your axes

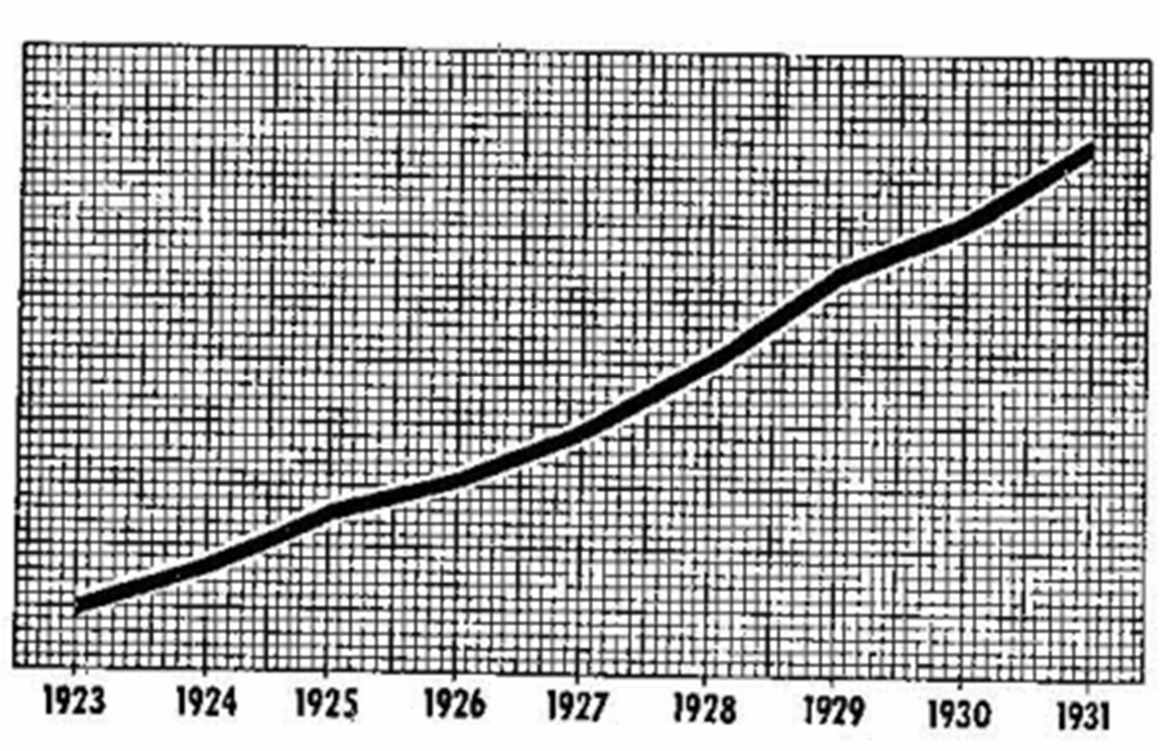
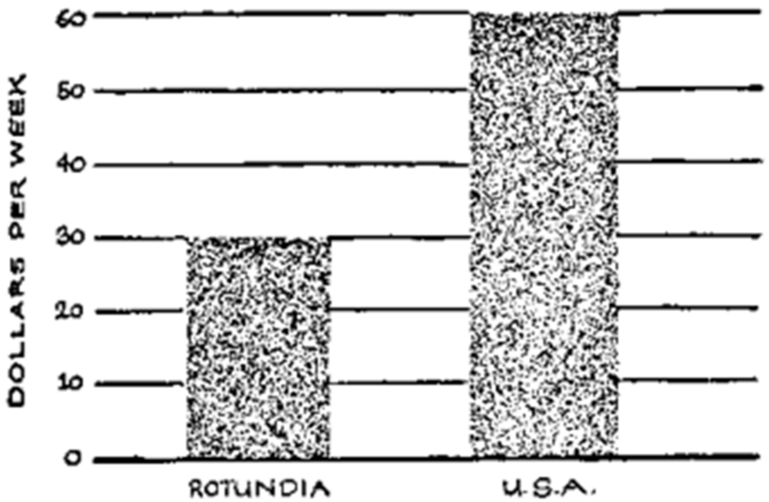


Chart Rules

- Provide a proper baseline & label your axes
- Avoid eye-candy



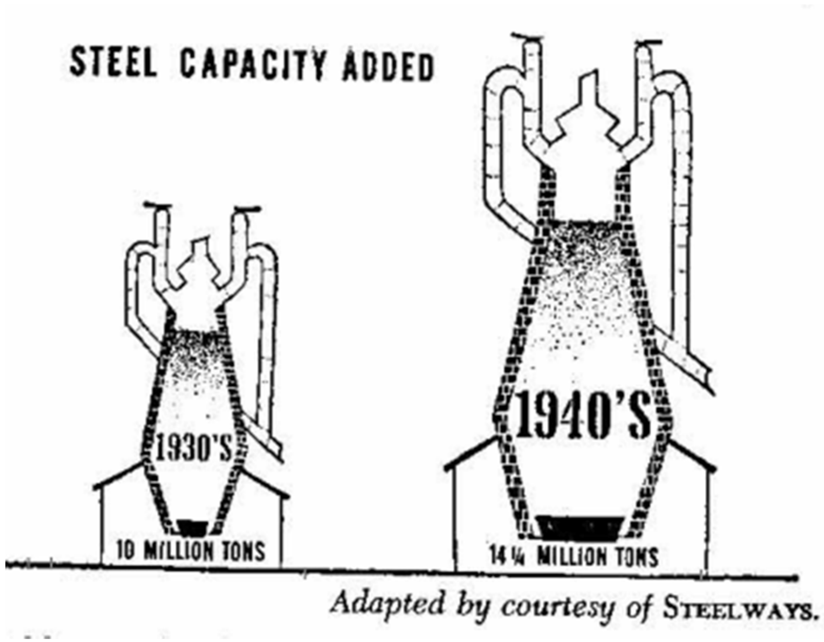
True data

same data with eye-candy & no numbers
but tells the same general story

impressive but a lie!

Chart Rules

- Provide a proper baseline & label your axes
- Avoid eye-candy
- Don't make people compare areas when not necessary



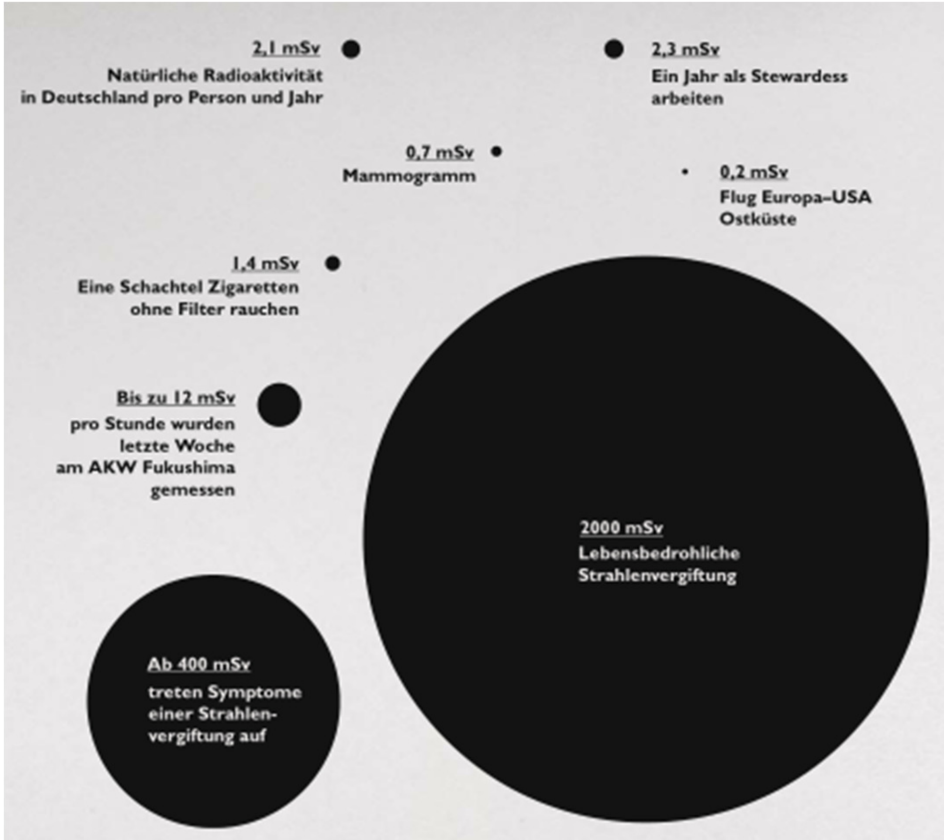
http://images.zeit.de/wissen/2011-04/s41-infografik-schiffsverkehr.pdf

Do the boxes represent the little white numbers??

A few more recent chart sins

All from www.zeit.de/grafik

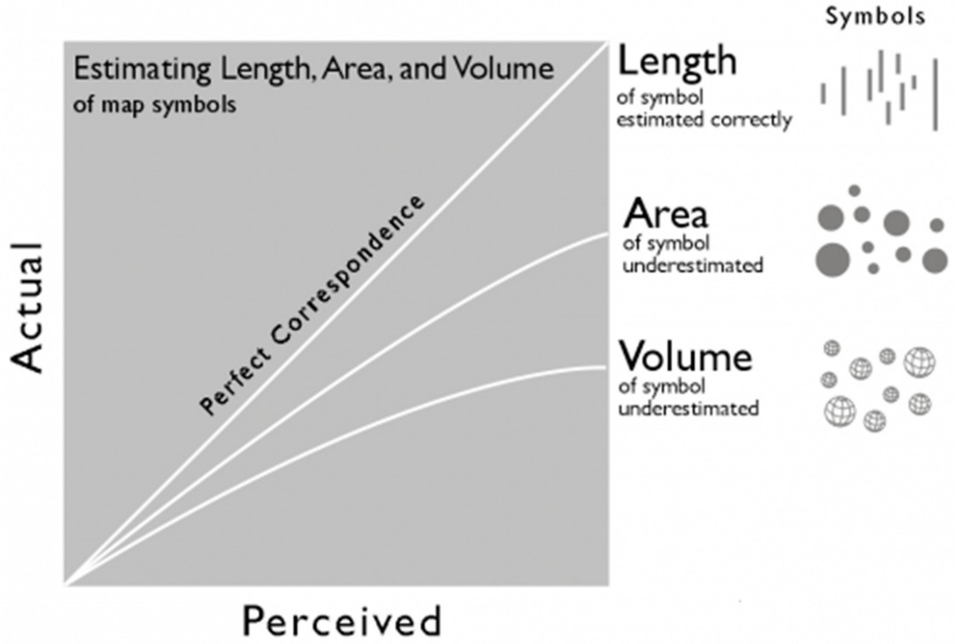
The numbers are the only useful part here.
The areas are not comparable



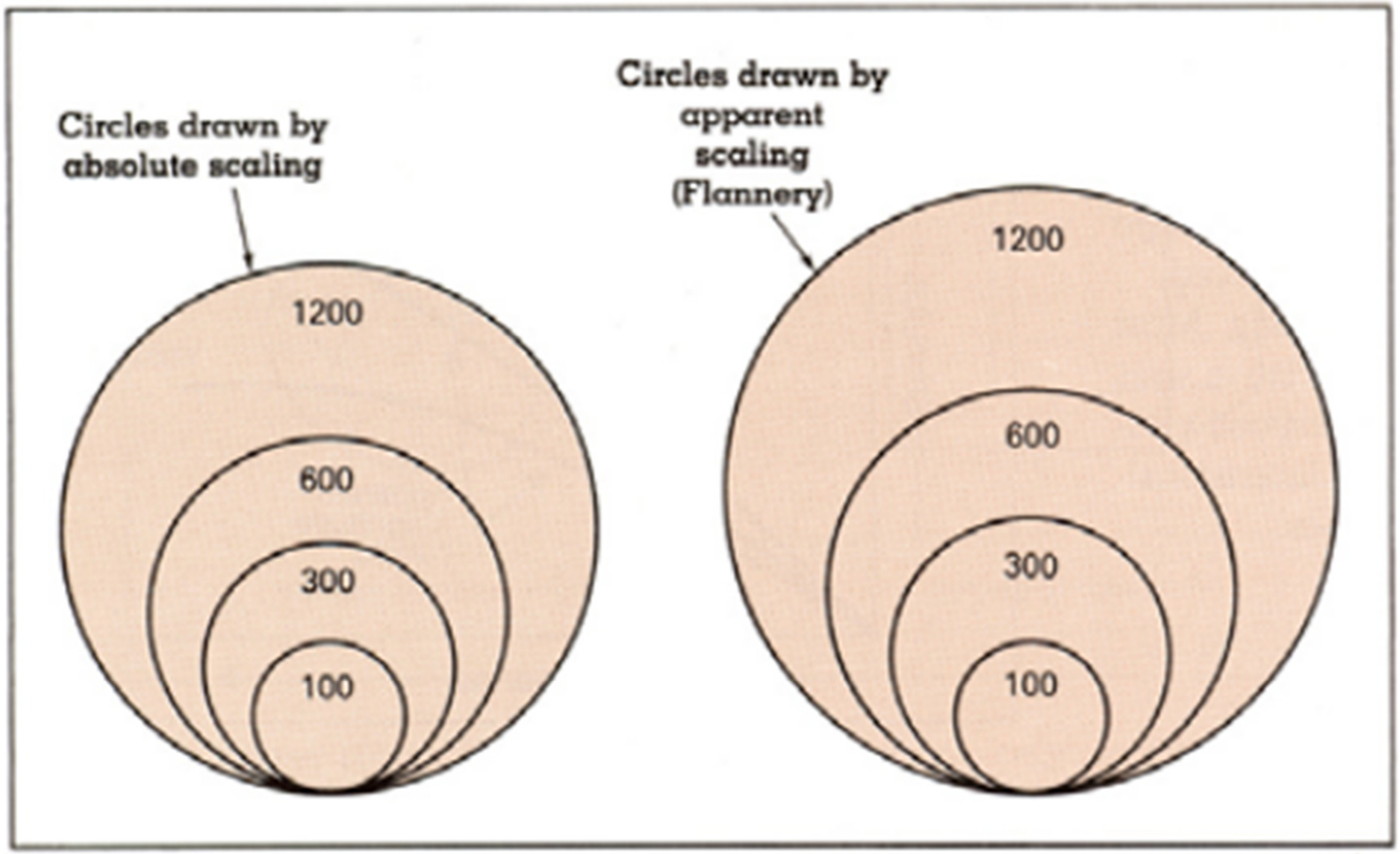
We already talked about problems comparing circles
→ There are actually even more problems…

Interlude

- People tend to **correctly estimate lengths**
- They tend to **underestimate areas and volumes.**
  - When asked to pick a circle that is two times the size of another most people would pick a circle ~1.8 times the size. This tendency gets worse with larger areas, and is worse in general for estimations of volumes.

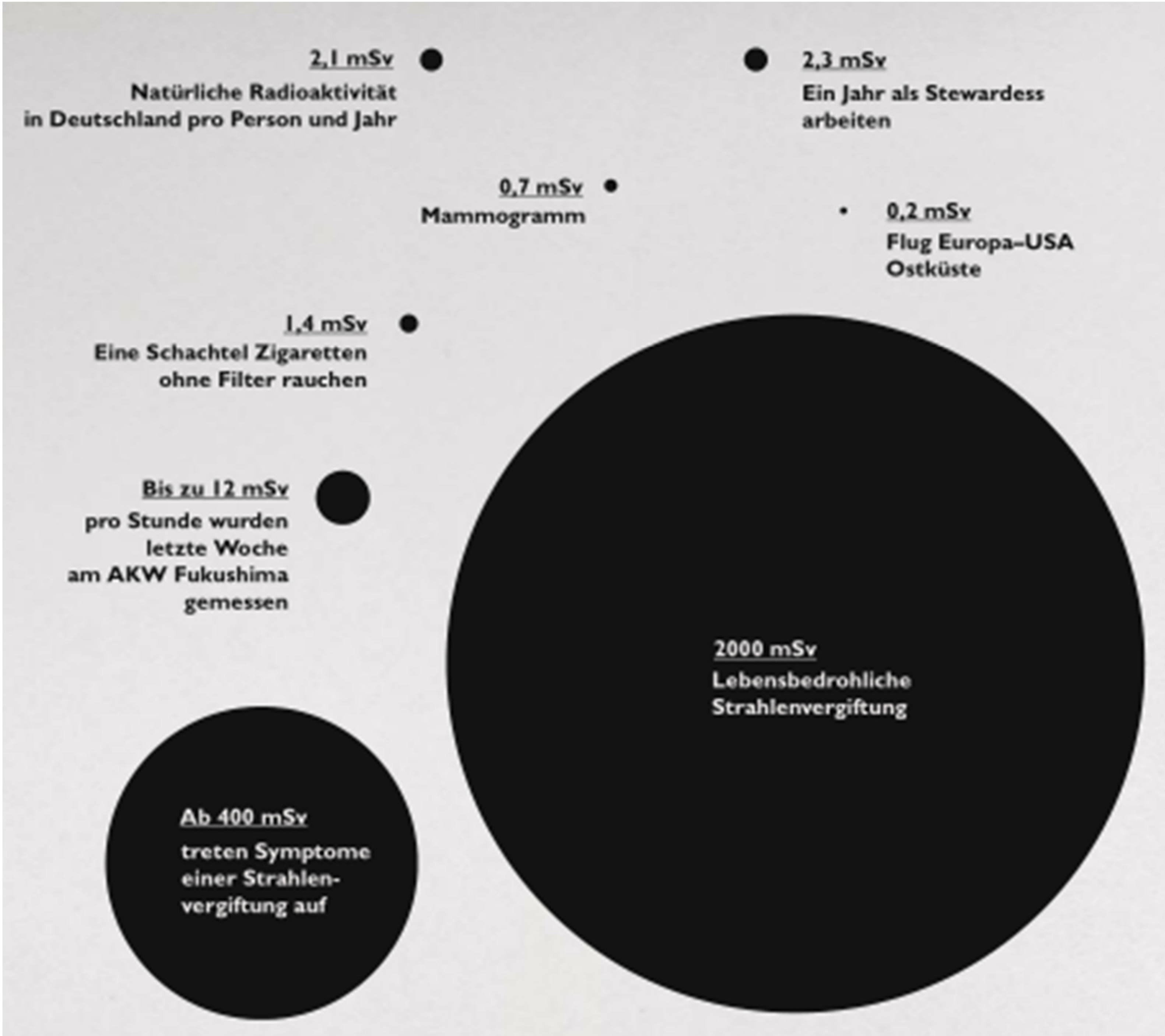


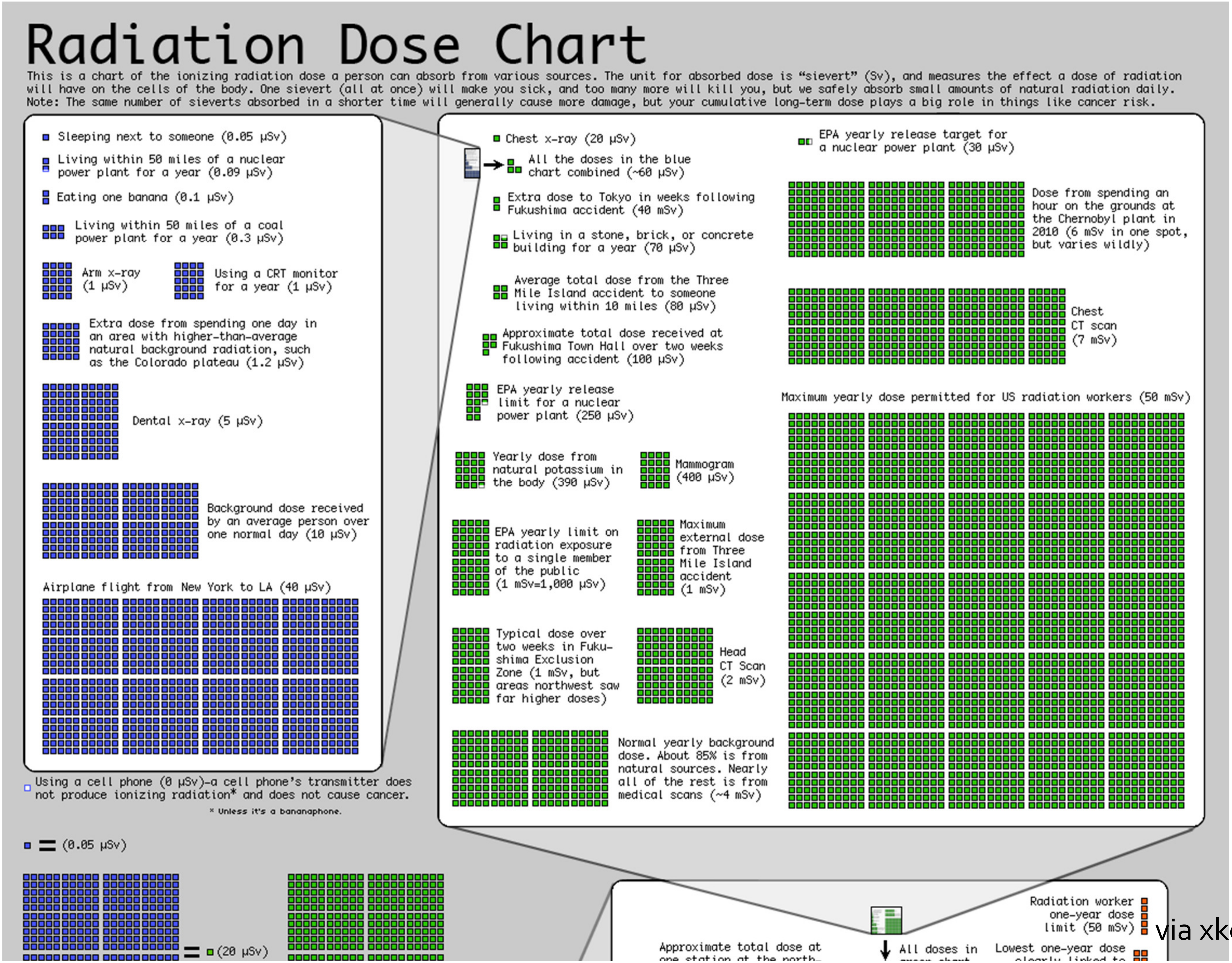
http://makingmaps.net/2007/08/28/perceptual-scaling-of-map-symbols/



[Cartography: Thematic Map Design, Figure 8.6, p. 170, Dent, 96]

$S = 0.98A^{0.87}$ [from Flannery 71]


2,1 mSv
Natürliche Radioaktivität in Deutschland pro Person und Jahr
2,3 mSv
Ein Jahr als Stewardess arbeiten
0,7 mSv
Mammogramm
0,2 mSv
Flug Europa–USA Ostküste
1,4 mSv
Eine Schachtel Zigaretten ohne Filter rauchen
Bis zu 12 mSv
pro Stunde wurden letzte Woche am AKW Fukushima gemessen
2000 mSv
Lebensbedrohliche Strahlenvergiftung
Ab 400 mSv
treten Symptome einer Strahlen-vergiftung auf


Radiation Dose Chart
This is a chart of the ionizing radiation dose a person can absorb from various sources. The unit for absorbed dose is "sievert" (Sv), and measures the effect a dose of radiation will have on the cells of the body. One sievert (all at once) will make you sick, and too many more will kill you, but we safely absorb small amounts of natural radiation daily. Note: The same number of sieverts absorbed in a shorter time will generally cause more damage, but your cumulative long-term dose plays a big role in things like cancer risk.
Sleeping next to someone (0.05 μSv)
Living within 50 miles of a nuclear power plant for a year (0.09 μSv)
Eating one banana (0.1 μSv)
Living within 50 miles of a coal power plant for a year (0.3 μSv)
Arm x-ray (1 μSv)
Using a CRT monitor for a year (1 μSv)
Extra dose from spending one day in an area with higher-than-average natural background radiation, such as the Colorado plateau (1.2 μSv)
Dental x-ray (5 μSv)
Background dose received by an average person over one normal day (10 μSv)
Airplane flight from New York to LA (40 μSv)
Using a cell phone (0 μSv)–a cell phone's transmitter does not produce ionizing radiation* and does not cause cancer.
* Unless it's a bananaphone.
Chest x-ray (20 μSv)
All the doses in the blue chart combined (~60 μSv)
Extra dose to Tokyo in weeks following Fukushima accident (40 mSv)
Living in a stone, brick, or concrete building for a year (70 μSv)
Average total dose from the Three Mile Island accident to someone living within 10 miles (80 μSv)
Approximate total dose received at Fukushima Town Hall over two weeks following accident (100 μSv)
EPA yearly release limit for a nuclear power plant (250 μSv)
Yearly dose from natural potassium in the body (390 μSv)
Mammogram (400 μSv)
EPA yearly limit on radiation exposure to a single member of the public (1 mSv=1,000 μSv)
Maximum external dose from Three Mile Island accident (1 mSv)
Typical dose over two weeks in Fukushima Exclusion Zone (1 mSv, but areas northwest saw far higher doses)
Head CT Scan (2 mSv)
Normal yearly background dose. About 85% is from natural sources. Nearly all of the rest is from medical scans (~4 mSv)
EPA yearly release target for a nuclear power plant (30 μSv)
Dose from spending an hour on the grounds at the Chernobyl plant in 2010 (6 mSv in one spot, but varies wildly)
Chest CT scan (7 mSv)
Maximum yearly dose permitted for US radiation workers (50 mSv)
= (0.05 μSv)
= (20 μSv)
Radiation worker one-year dose limit (50 mSv)
Approximate total dose at one station at the north-
All doses in
Lowest one-year dose

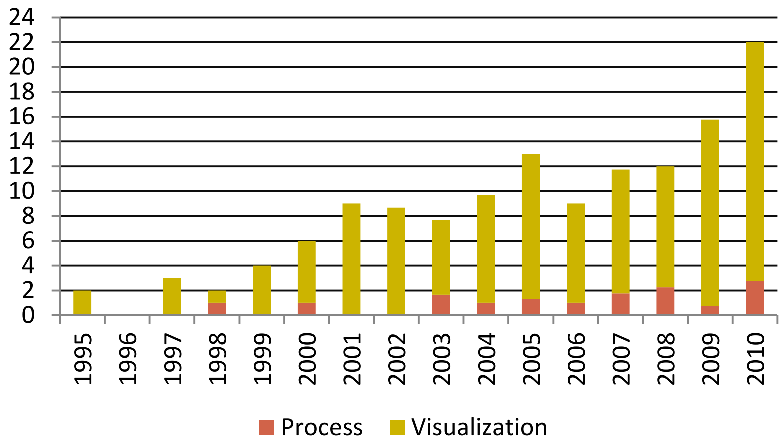
via xkcd

Chart Rules

- Provide a proper baseline & label your axes
- Avoid eye-candy
- Don’t make people compare areas when not necessary
- Provide legends

Chart Rules

- Provide a proper baseline & label your axes
- Avoid eye-candy
- Don't make people compare areas when not necessary
- Provide legends
- Grids help but make them subtle (no black lines!)



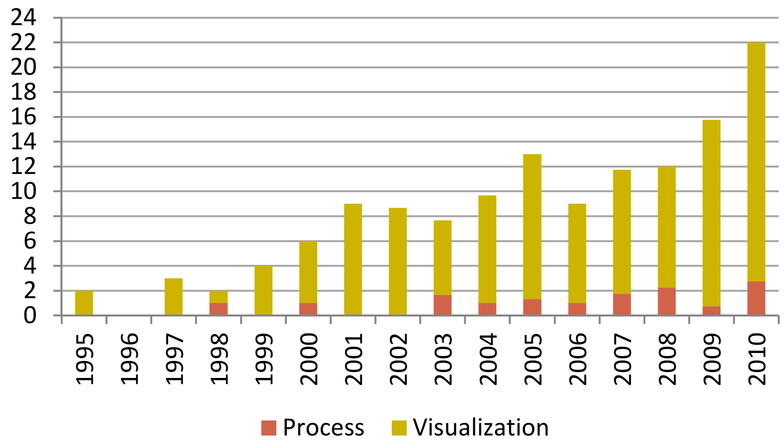
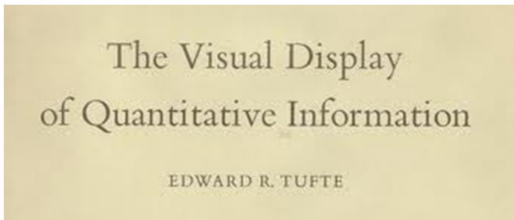
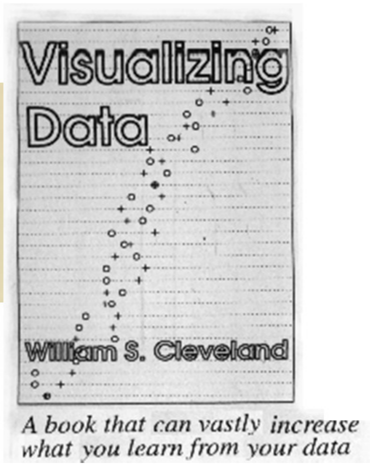
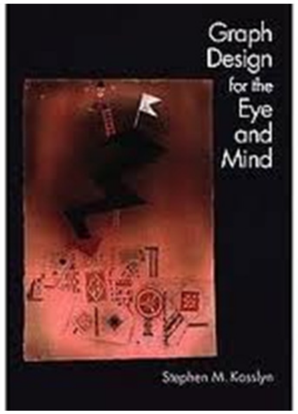


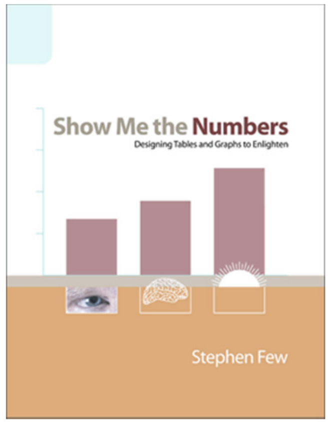
Chart Rules

- Provide a proper baseline & label your axes
- Avoid eye-candy
- Don't make people compare areas when not necessary
- Provide legends
- Grids help but make them subtle (no black lines!)
- Many more…









SPATIAL VISUALIZATION TECHNIQUES

Heatmaps

Hotmap (Fisher, 2007)

http://hotmap.msresearch.us/

Figure 1. Hotmap, showing all trails. A pixel is red if at least one user has looked at it. Note the frequent horizontal and vertical lines from users holding down scroll buttons.

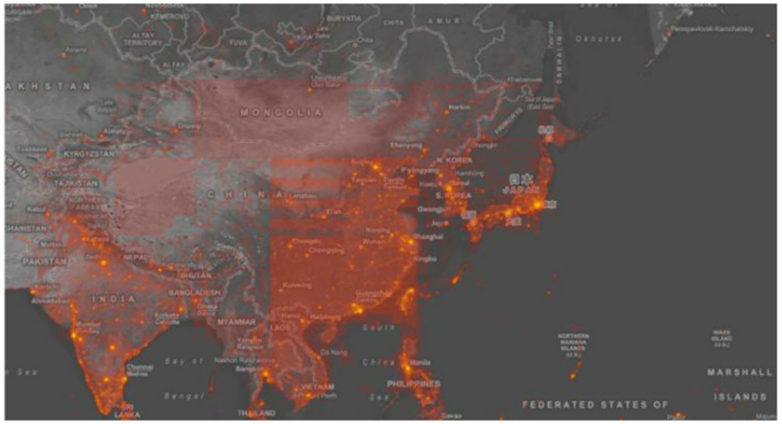
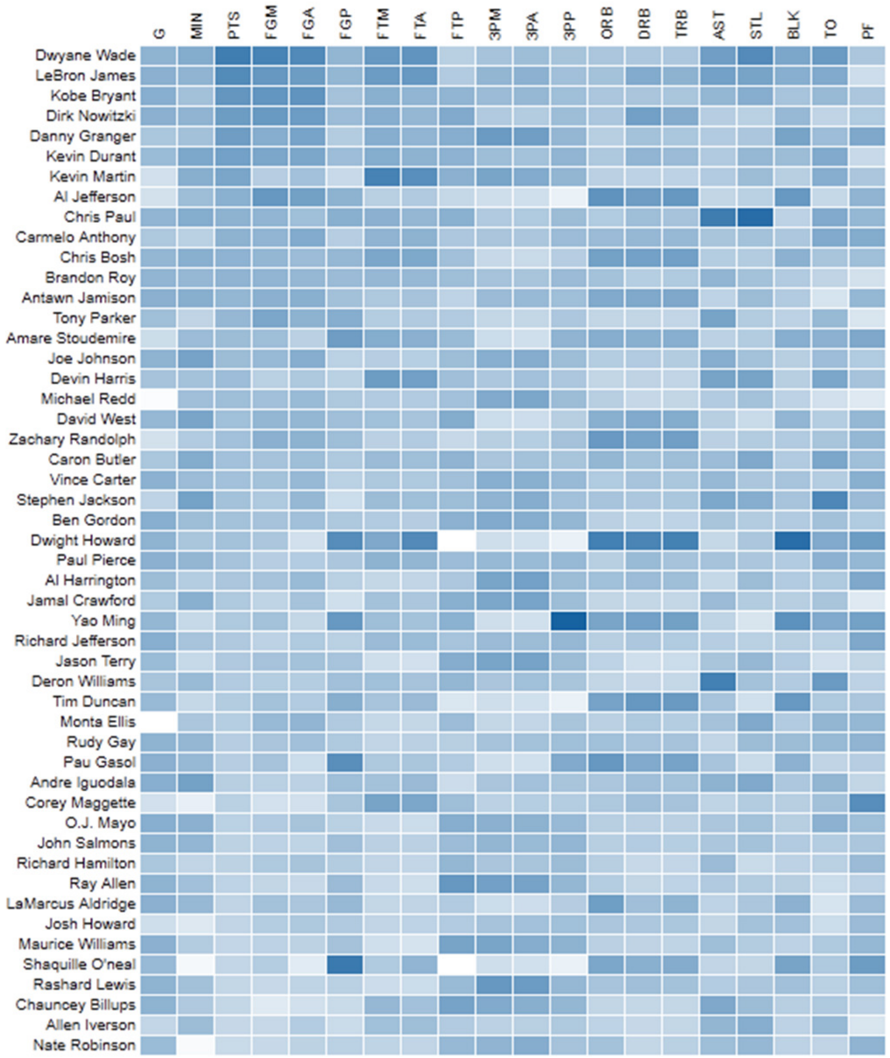


Figure 3. Hotmap in China. Note the faint highlight over greater China, and the brighter colors over the coastal area, suggesting repeated scraping of the data.

Heatmaps

Also work for non-spatial data

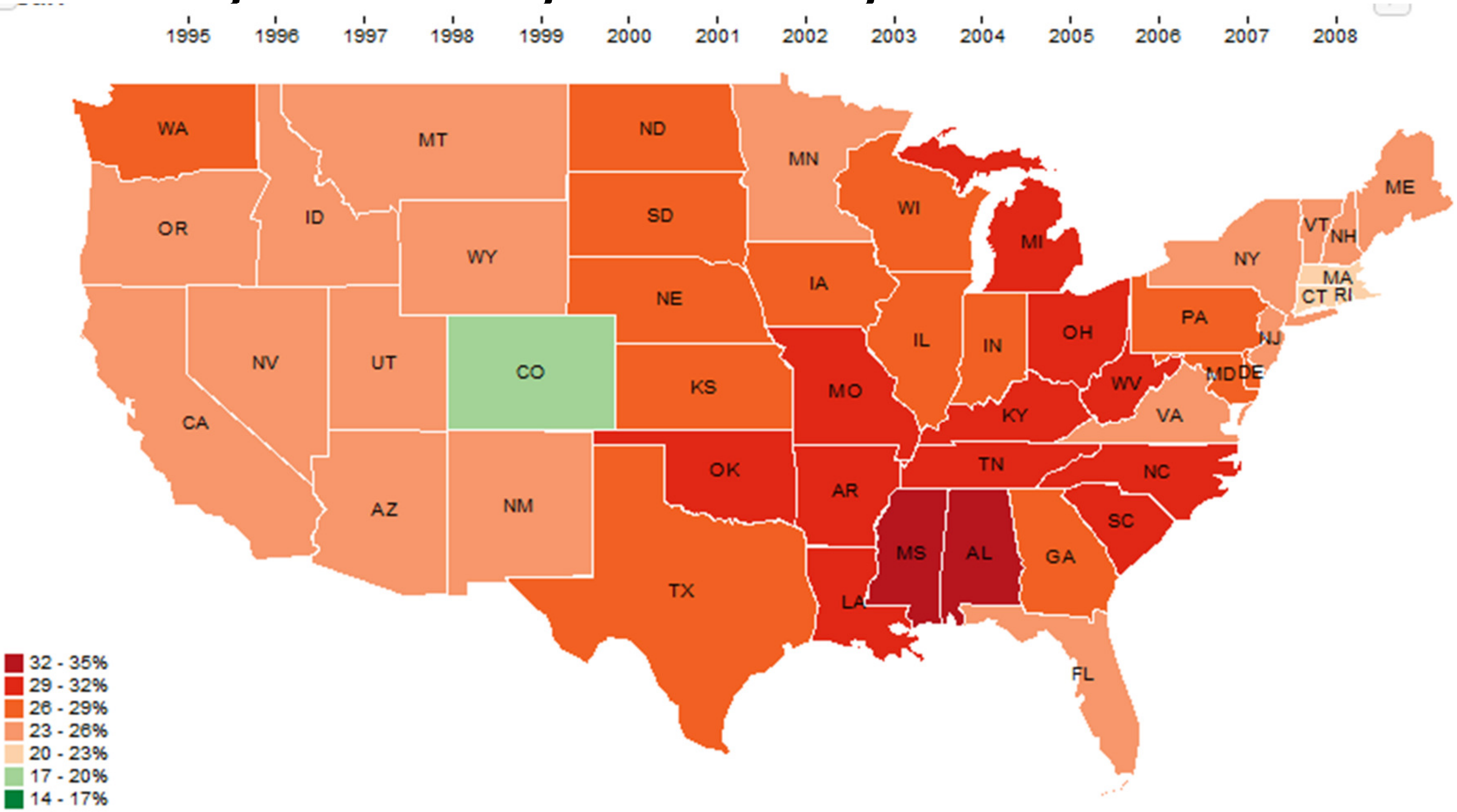


Source: FlowingData

top scorer in the NBA

http://mbostock.github.com/protovis/ex/nba.html

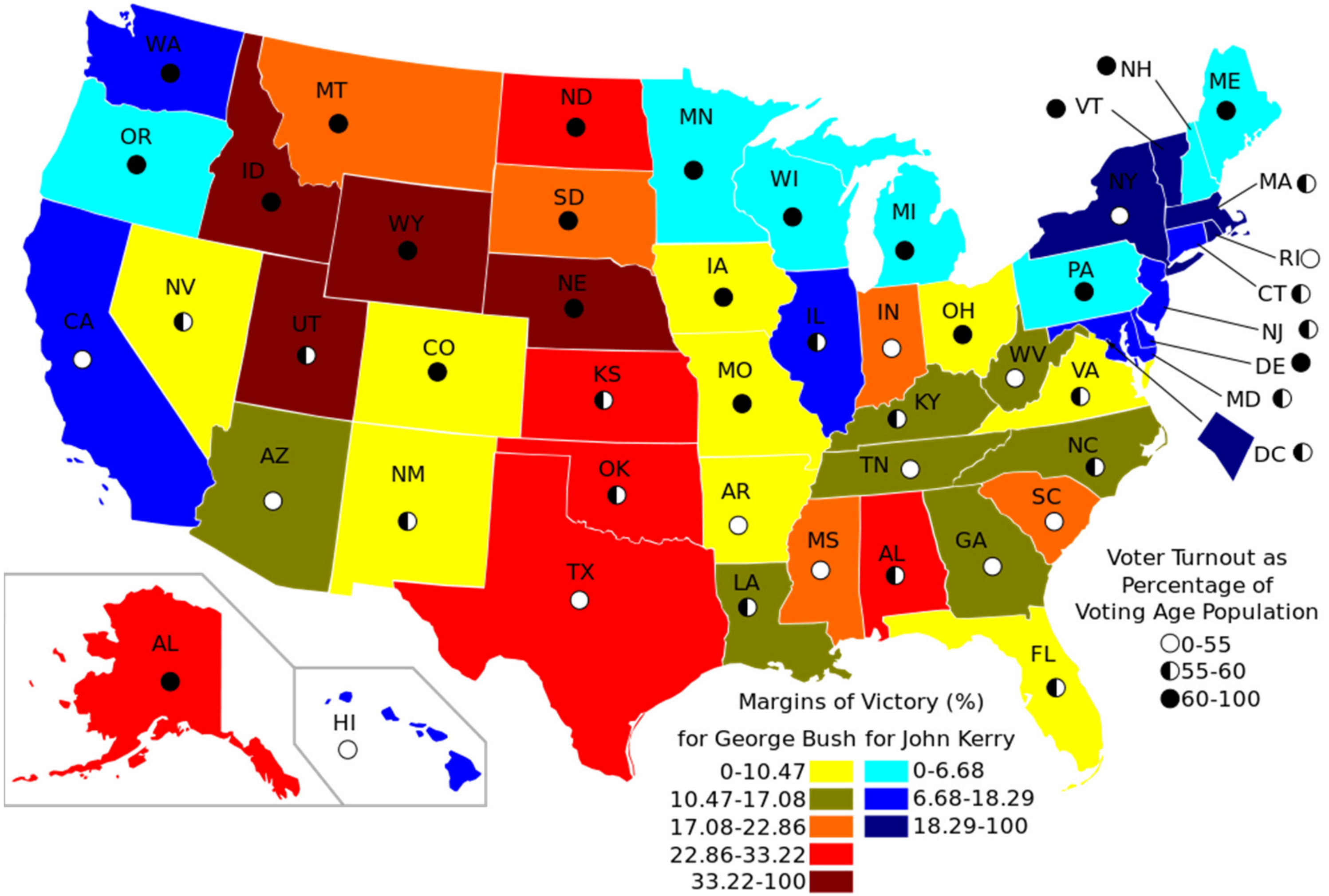
HeatMap – Cloropleth Map



Problems: If area is small, values are hard to read and color perception may be difficult

http://mbostock.github.com/protovis/ex/choropleth.html

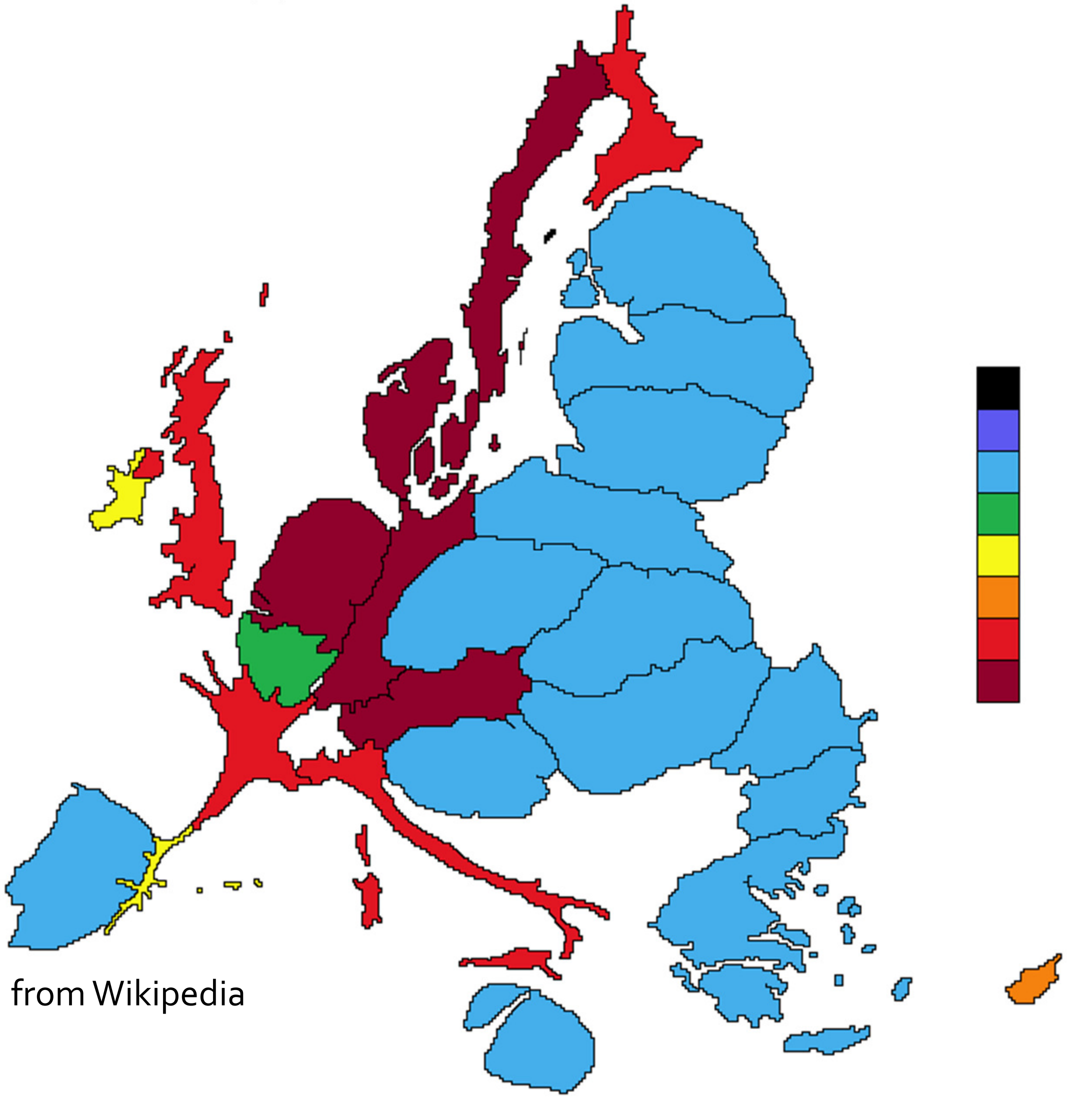
Remember to use good color – always!



P.S. This is not good use of color!

http://en.wikipedia.org/wiki/File:2004US_election_map.svg

Cartogram



from Wikipedia

Cartogram showing Open Europe estimate of total European Union *per capita* net budget expenditure in euros for the whole period 2007-2013 per capita, based on Eurostat 2007 population estimates (Luxembourg not shown).

Net contributors

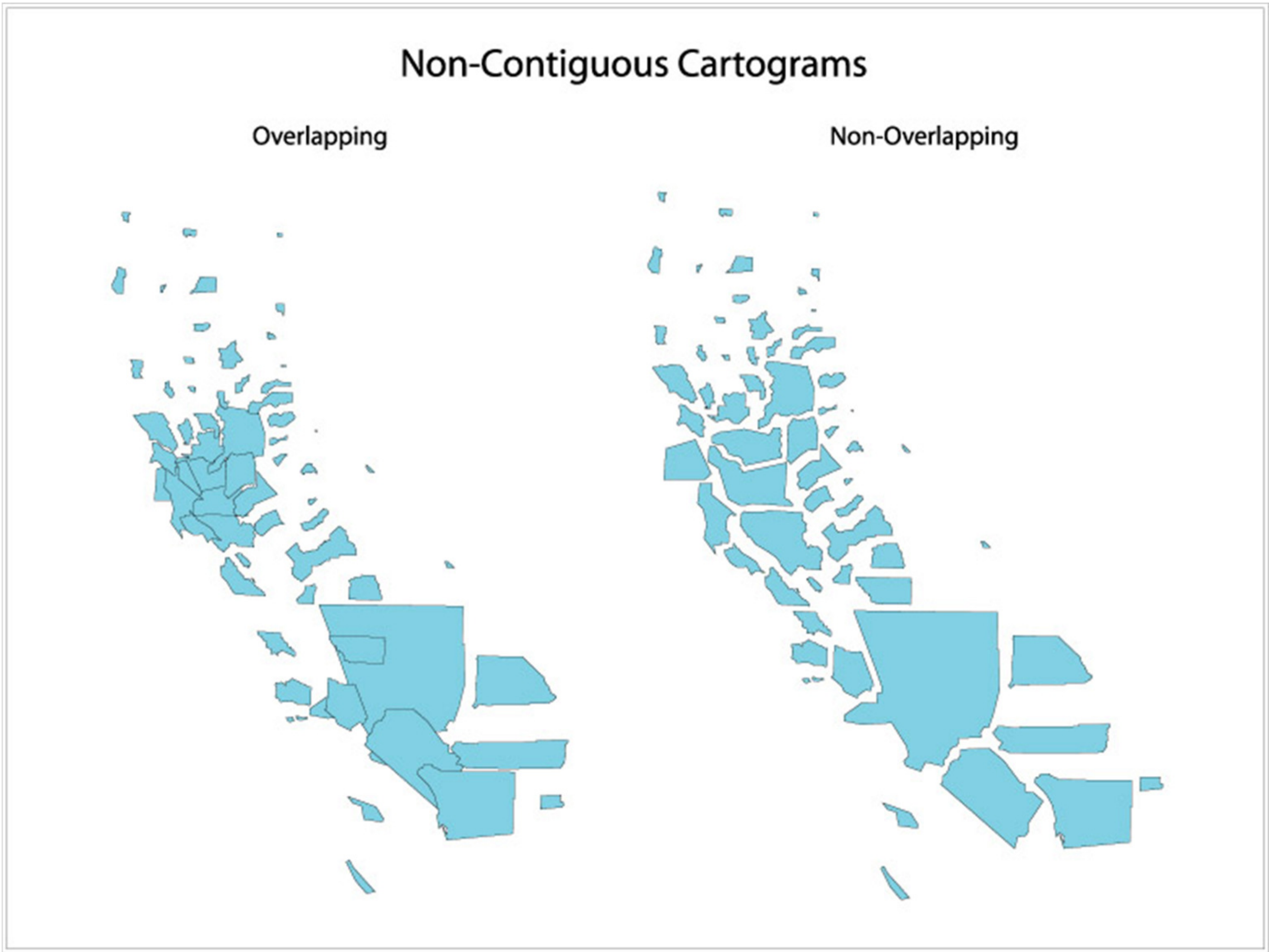
- -5000 to -1000 euro per capita
- -1000 to -500 euro per capita
- -500 to 0 euro per capita

Net recipients

- 0 to 500 euro per capita
- 500 to 1000 euro per capita
- 1000 to 5000 euro per capita
- 5000 to 10000 euro per capita
- 10000 euro plus per capita
- n/a

Continuous cartogram (tries to optimize size constraints as much as possible)

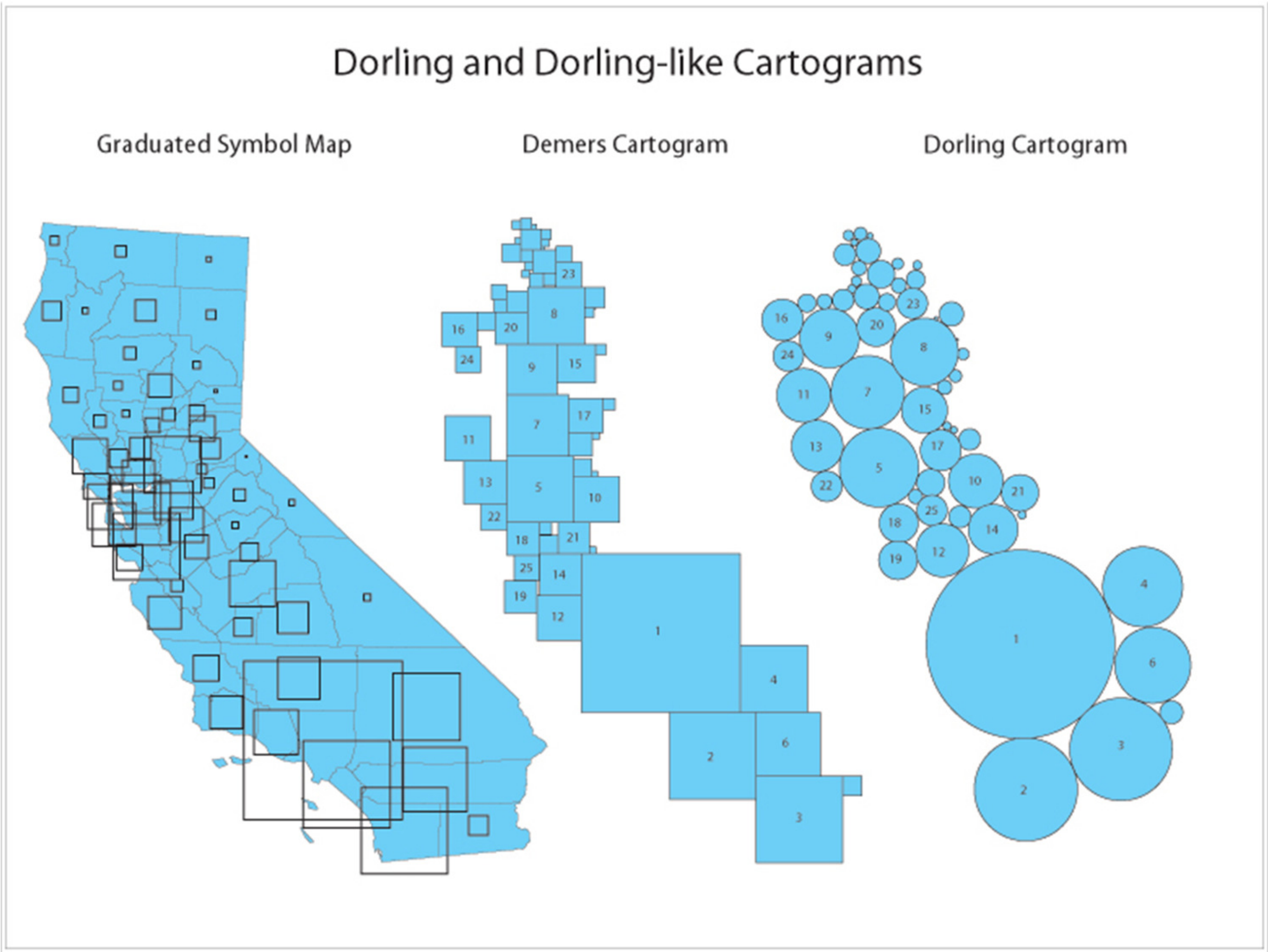
Cartogram



Size can be set easily but map shape and connectivity is lost

http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html

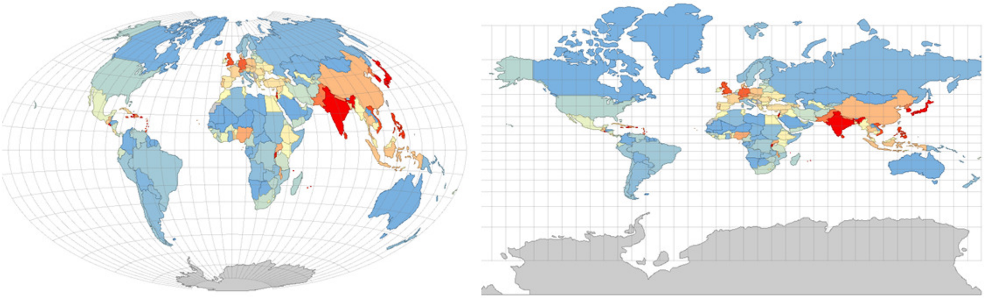
Cartogram



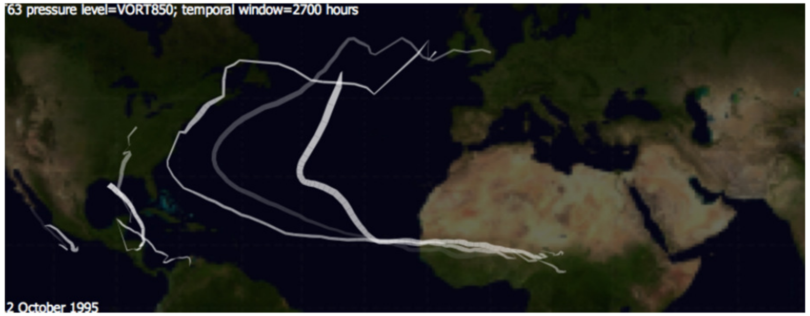
Size comparison easier but map topology largely lost

http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html

GeoVis is its own field with many specific rep problems



Projections



Slingsby et al, 2010

Movement in Space (here storms)

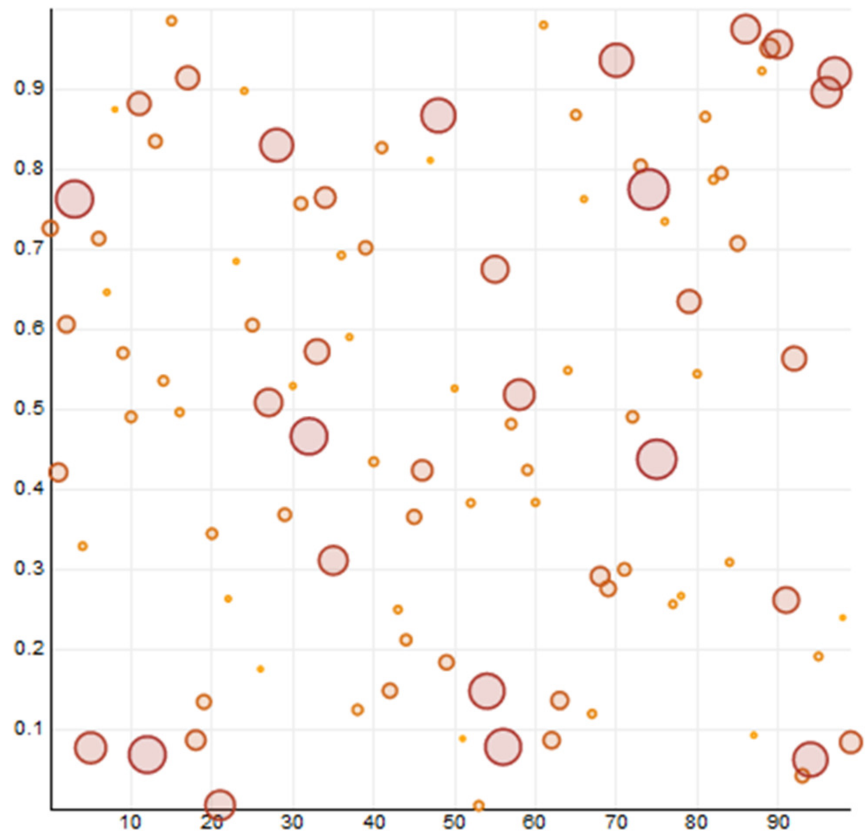
Consult the specialized conferences and literature for more information (too detailed for now)

Visualization Techniques for

MULTIVARIATE DATA

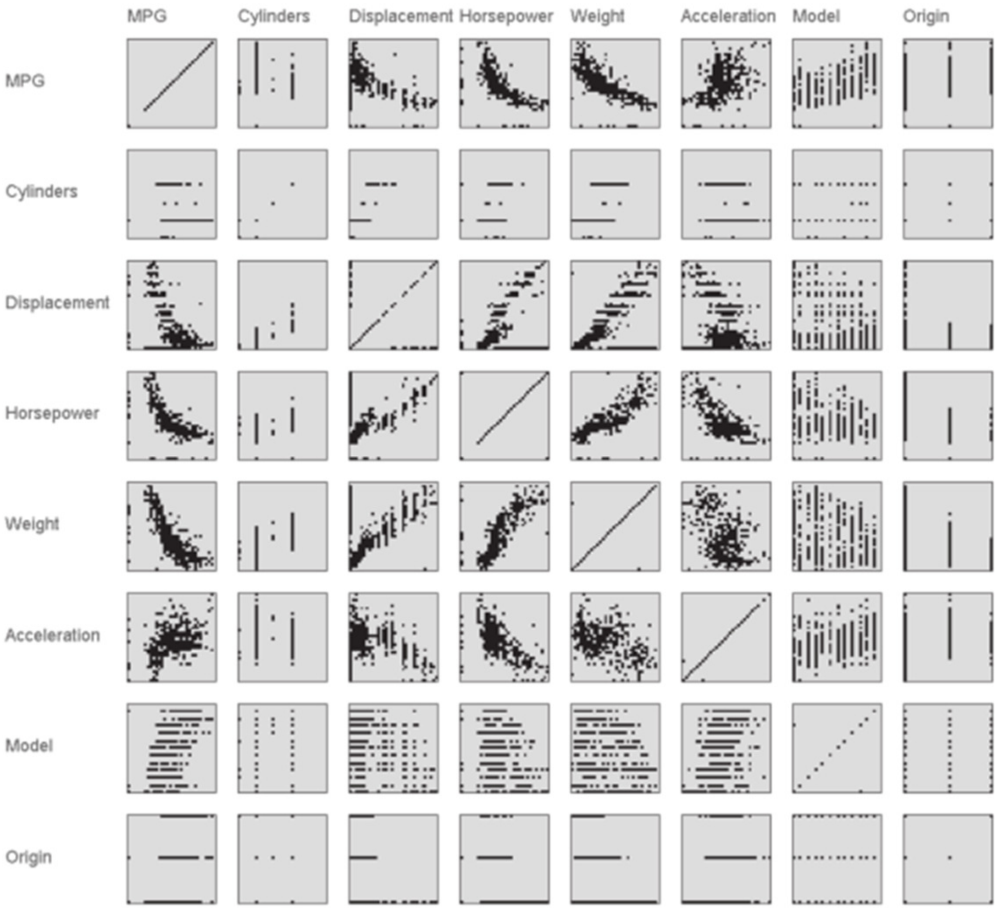
Scatterplot Matrices

- Regular scatterplot
  - 2D position
    (2 data dimension)
  - Size can encode additional dimension
  - Color of circles can encode another dimension



- Hans Rosling TED Talk Video

Scatterplot Matrices



7-dimensional car dataset

correlation of 2 dimensions per square

Interlude – interaction with ScatterDice

Rolling the Dice

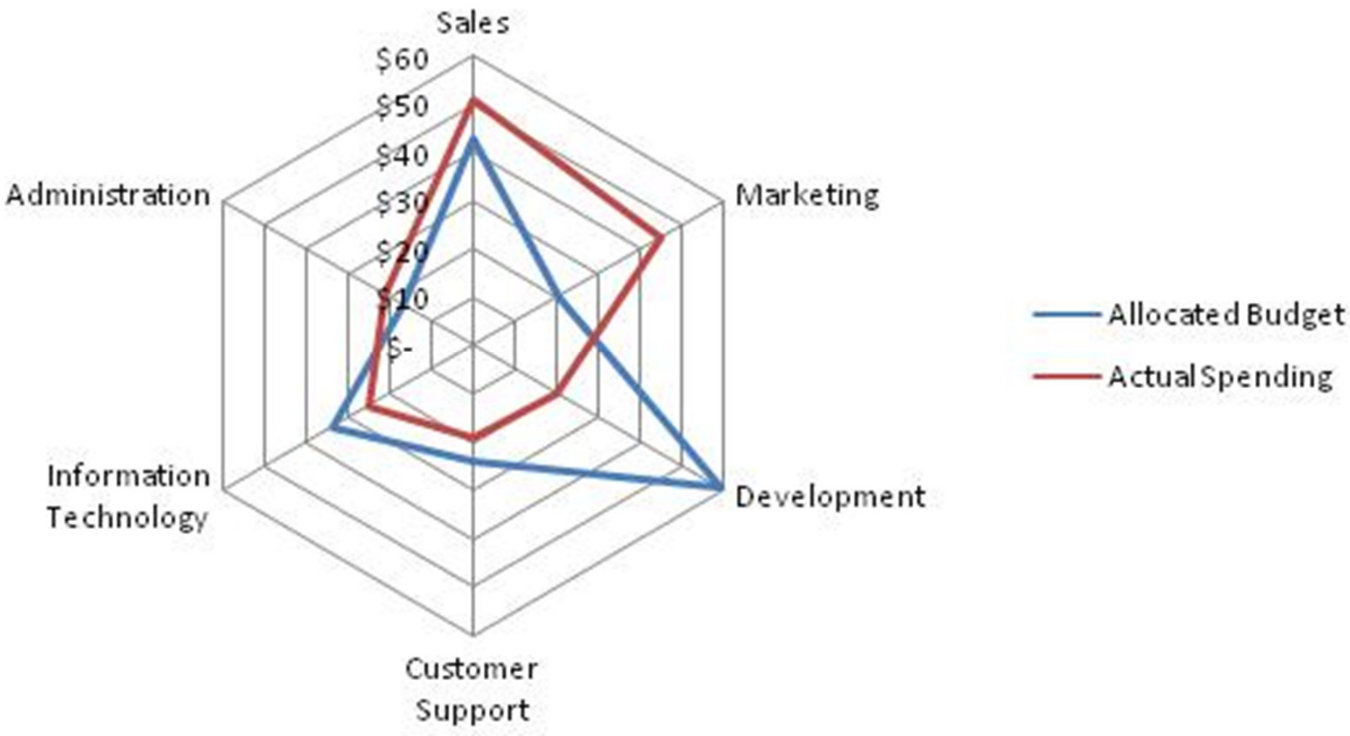
Multidimensional Visual Exploration
using Scatterplot Matrix Navigation

Niklas Elmqvist
Pierre Dragicevic
Jean-Daniel Fekete

INRIA

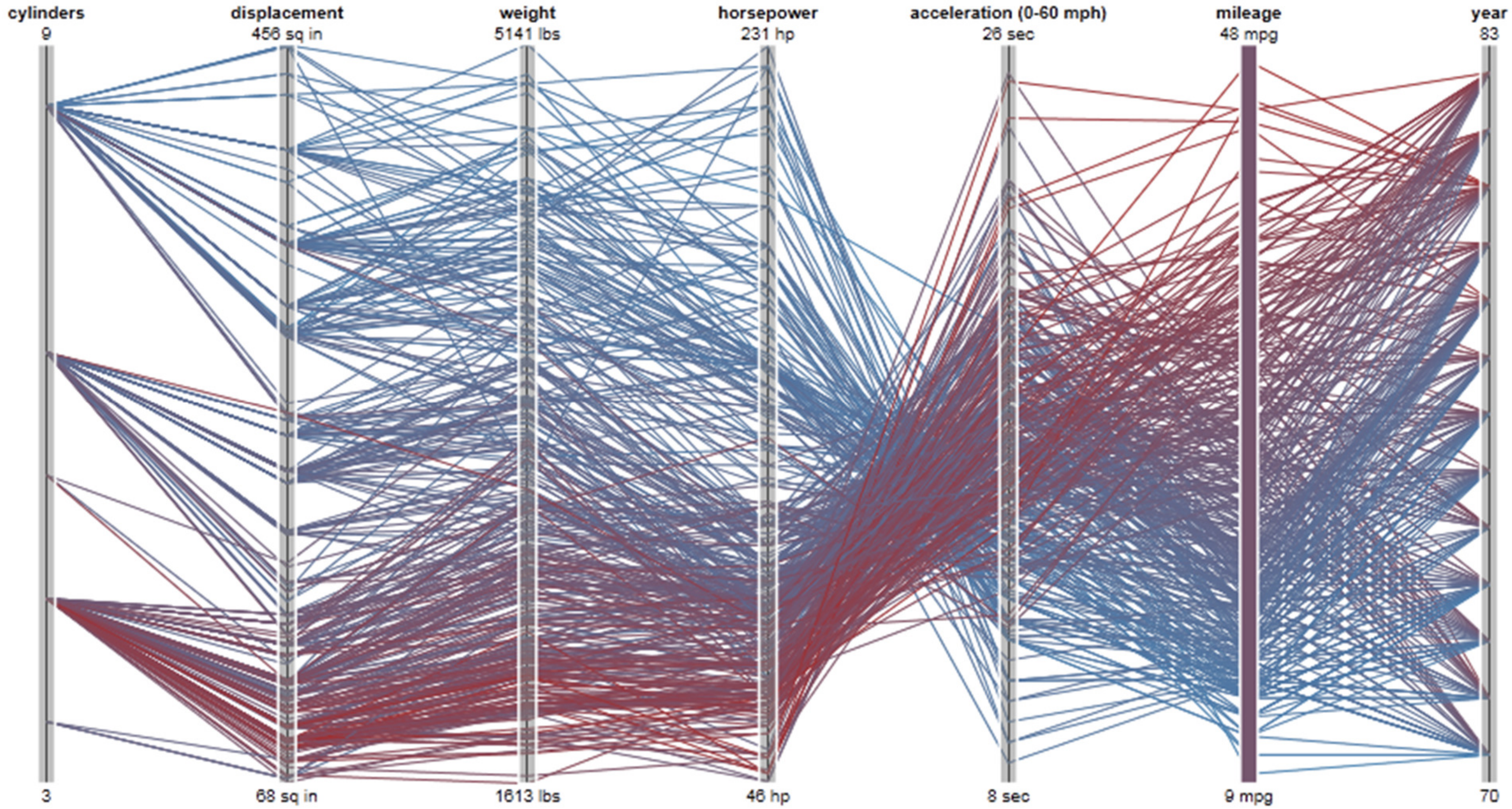
Star Plot / Spider Chart / Radar Chart

- In this case sharing similar scales
- But also works for different scales



Parallel Coordinates

- Extremely well researched method with many possibilities to fine-tune



http://mbostock.github.com/protovis/ex/cars.html

Parallel Coordinates

Google scholar "parallel coordinates" Search Advanced Scholar Search

Scholar Articles and patents ▾ anytime ▾ include citations ▾ Create email alert Results **1** - **10** of about **4,880**. (**0.08** sec)

Parallel coordinates: a tool for visualizing multi-dimensional geometry [PDF] from tuwien.ac.at
A Inselberg… - Proceedings of the 1st conference on …, 1990 - portal.acm.org
Here a methodology based on a multi-dimensional system of **Parallel Coordinates** whose development began in 1978 is described. Preliminary results on some representations and construction algorithms for N-Dimensional Lines and Hyperplanes ap- peared in 1981 [28]. The results were …
Cited by 700 - Related articles - All 10 versions

The plane with **parallel coordinates**
A Inselberg - The Visual Computer, 1985 - citeulike.org
Abstract By means ofParallel Coordinates planar “graphs” of multivariate relations are obtained. Certain properties of the relationship correspond tothe geometrical properties of its graph. On the plane a point ←→ line duality with several interesting properties is induced. …
Cited by 614 - Related articles - Cached - All 3 versions

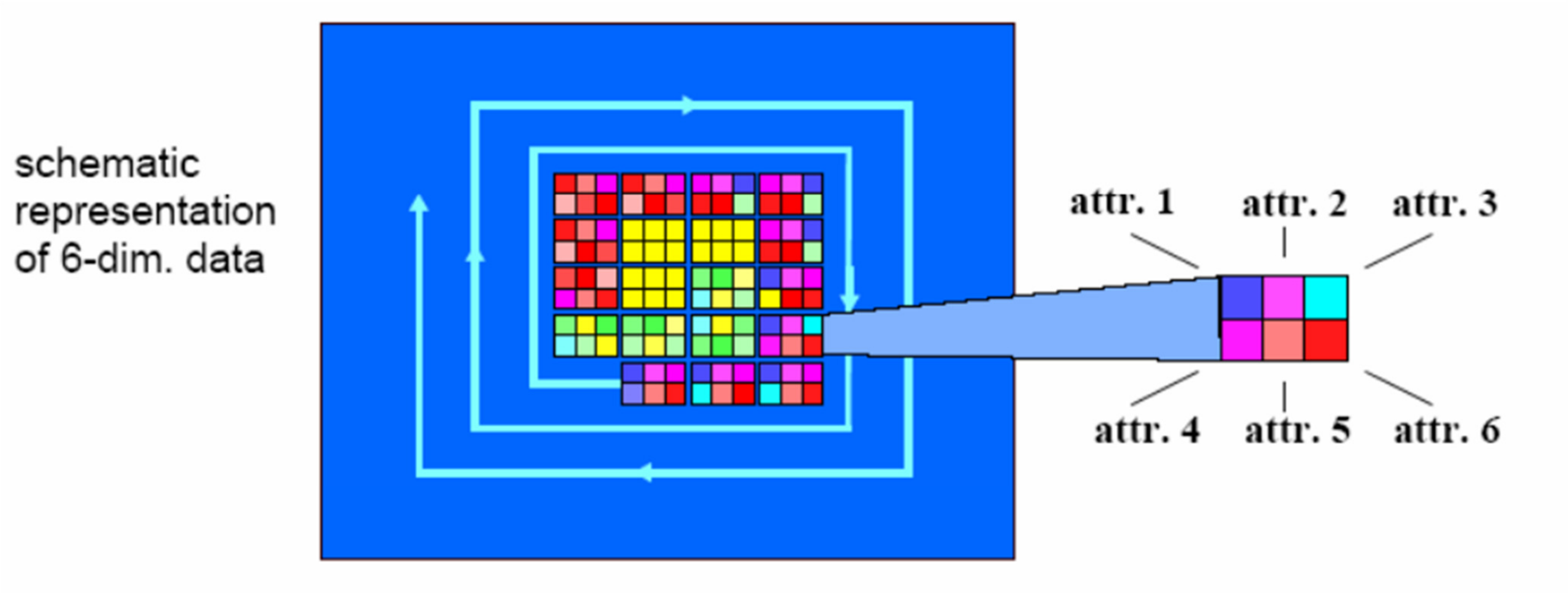
Tools to try them

- http://davis.wpi.edu/~xmdv/visualizations.html
- http://www2.research.att.com/areas/stat/xgobi/
- http://www.wallinfire.net/picviz/
- R: http://stat.ethz.ch/R-manual/R-patched/library/MASS/html/parcoord.html
- Visualization Toolkits…

- http://www.kdnuggets.com/software/parallax.html (from the „inventor" – but not free)

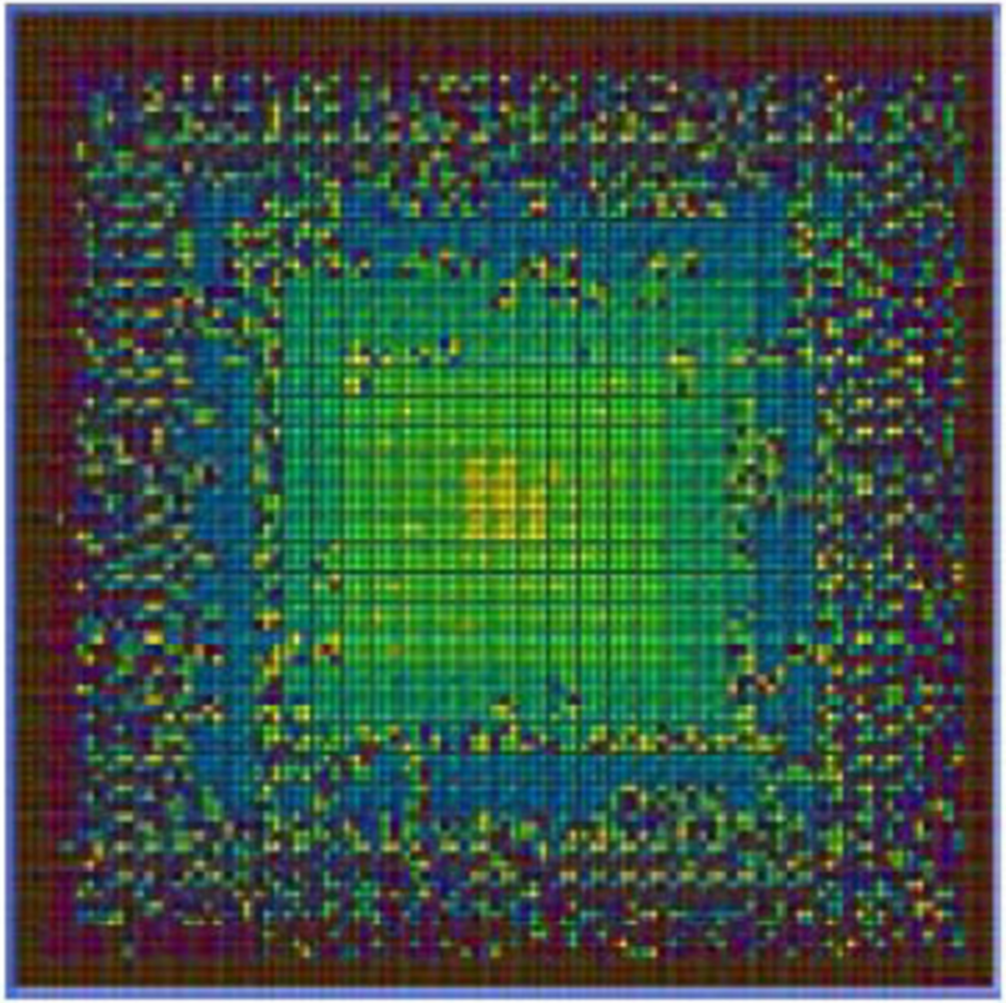
Disclaimer: I haven't tested any of them

Dense Pixel Displays



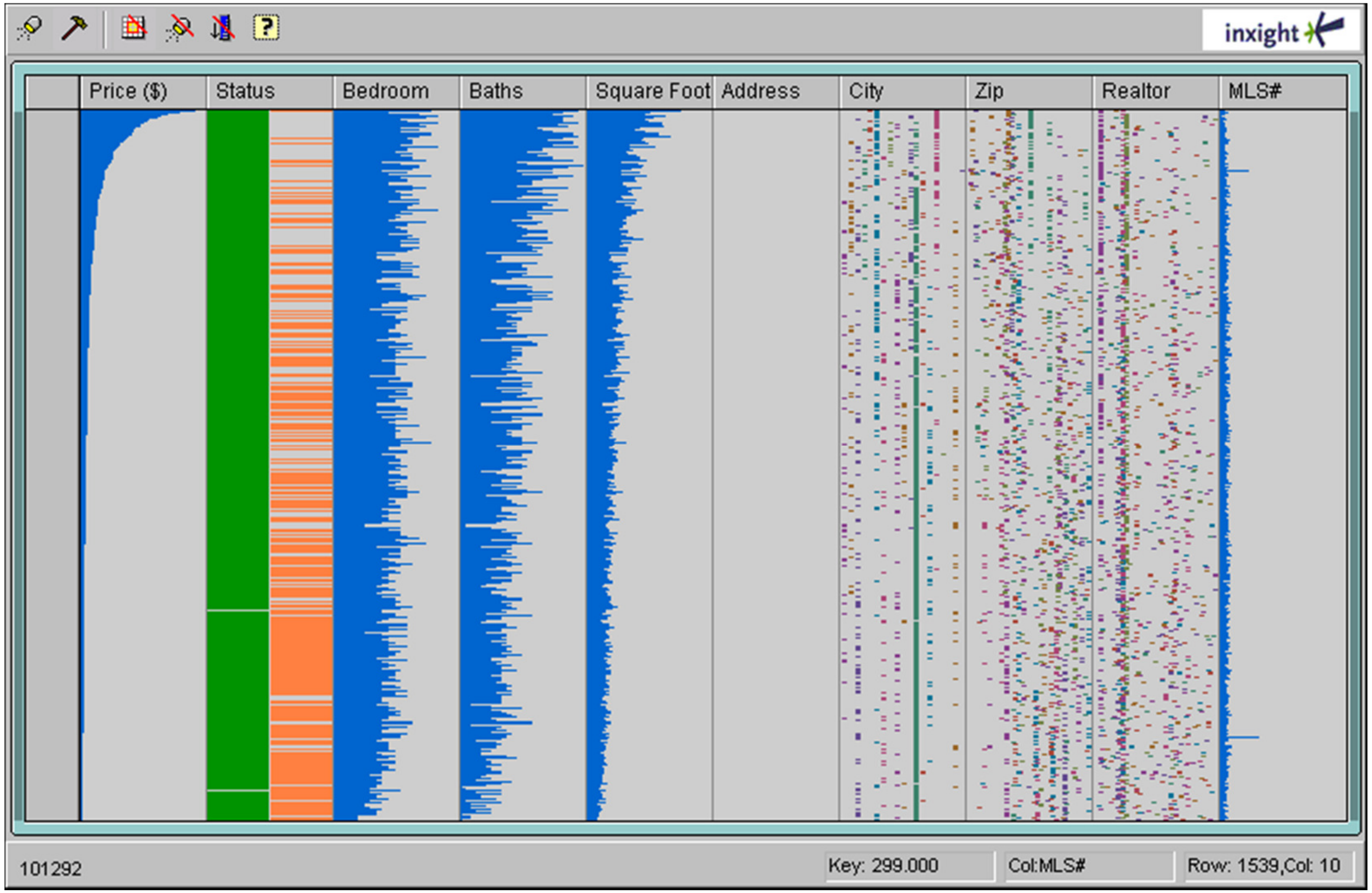
Levkowitz, Vis '91

Dense Pixel Displays



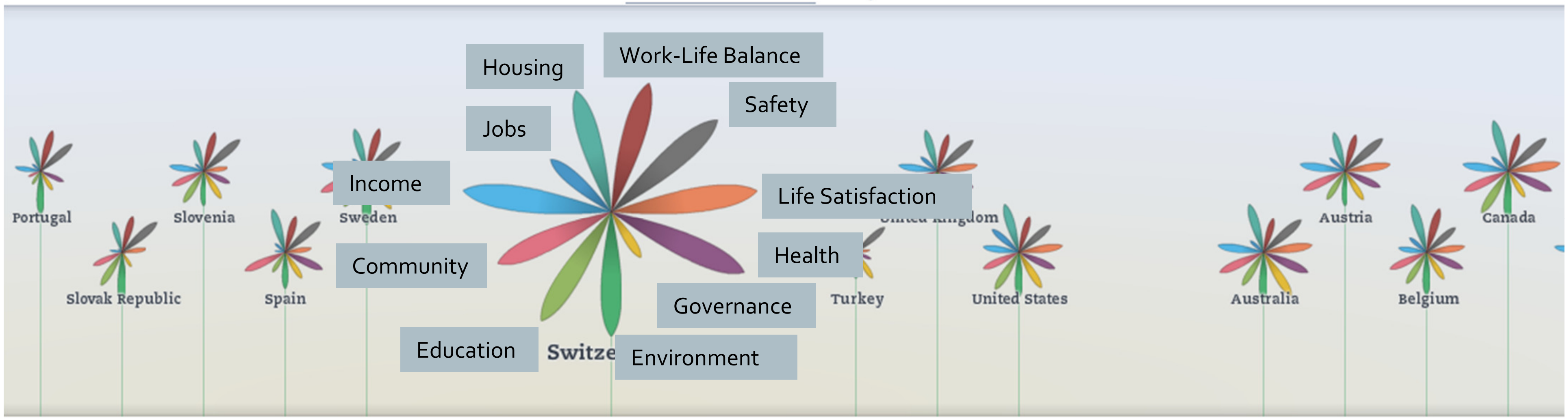
Tabular Data Displays

- TableLens



Can do some of this in Excel now

Glyphs



Housing
Work-Life Balance
Safety
Jobs
Income
Life Satisfaction
Community
Health
Governance
Education
Environment
Portugal
Slovenia
Sweden
Slovak Republic
Spain
Switze
Turkey
United States
Australia
Austria
Belgium
Canada

People Garden

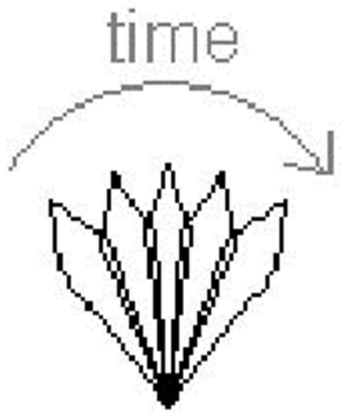
- Visualization technique for portraying on-line interaction environments (Virtual Communities)
- Provides both individual and societal views
- Utilizes garden and flower metaphors

Xiong and Donath
UIST '99

The following slides on PeopleGarden are adapted from John Stasko's

Data Portrait: Petals

Fundamental view of an individual



His/Her postings are represented as petals of the flower, arranged by time in a clockwise

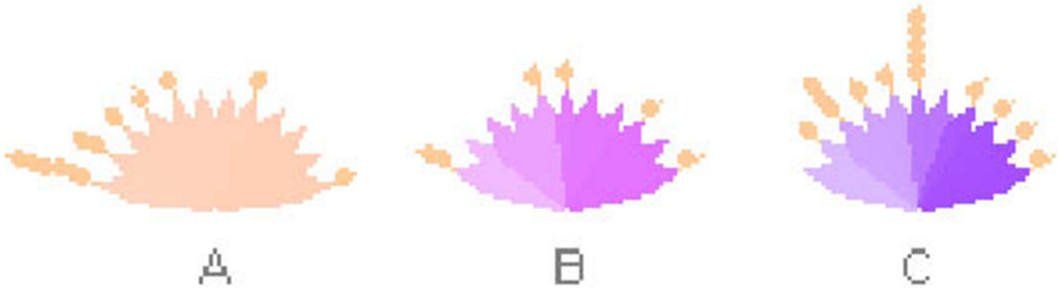
Data Portrait: Postings

Time of Posting

- New posts are added to the right
- Slide everything back so it stays symmetric
- Each petal fades over time showing time since posting
- A marked difference in saturation of adjacent petals denotes a gap in posting

Data Portrait: Responses

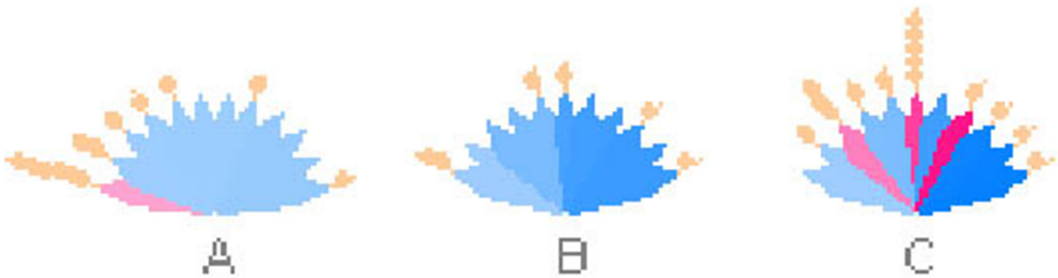
Response to posting



Small circle drawn on top of a posting to represent each follow-up response

Data Portrait: Color

Initial post vs. reply



Color can represent original/reply

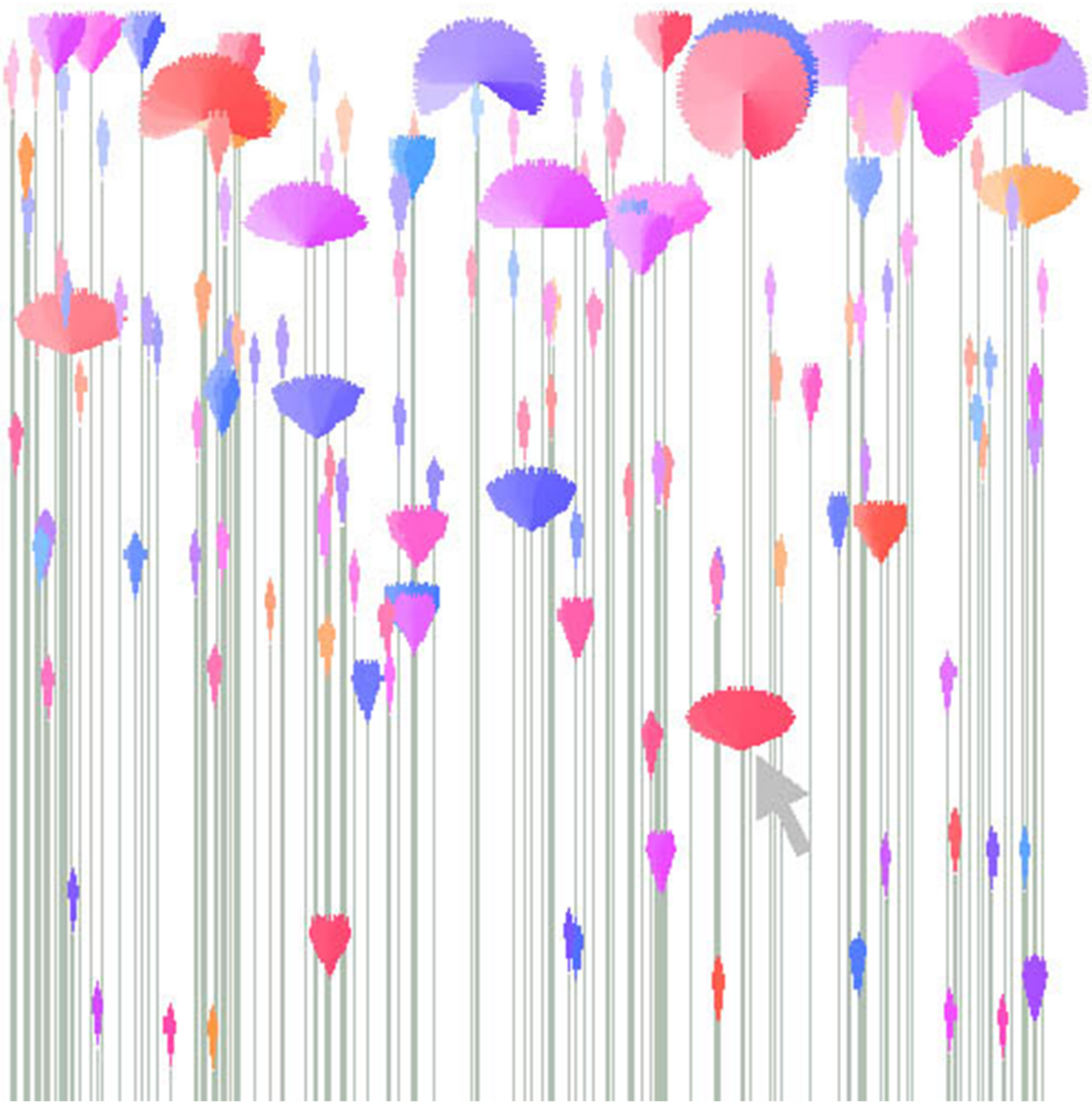
Here magenta is original post, blue is reply

Garden

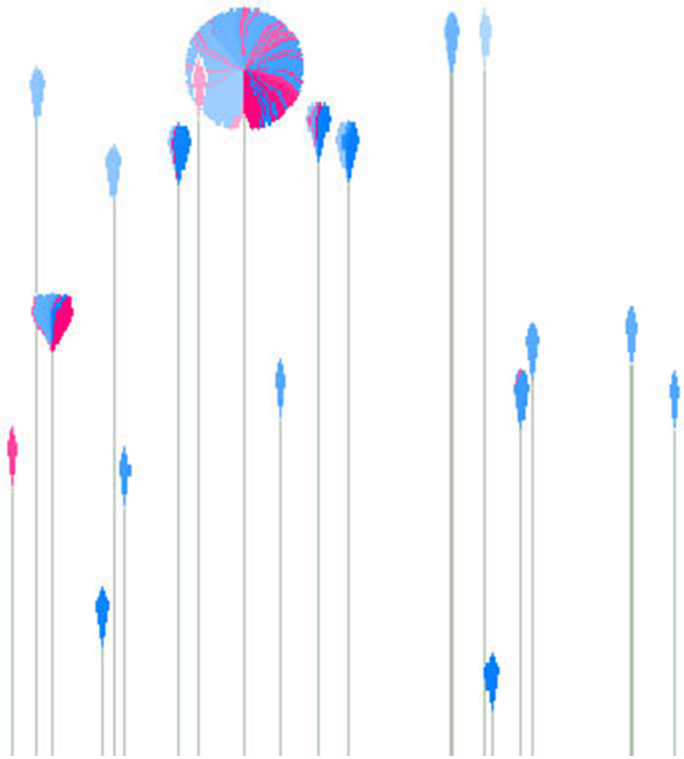
Combine many portraits to make a garden

Message board with 1200 postings over 2 months

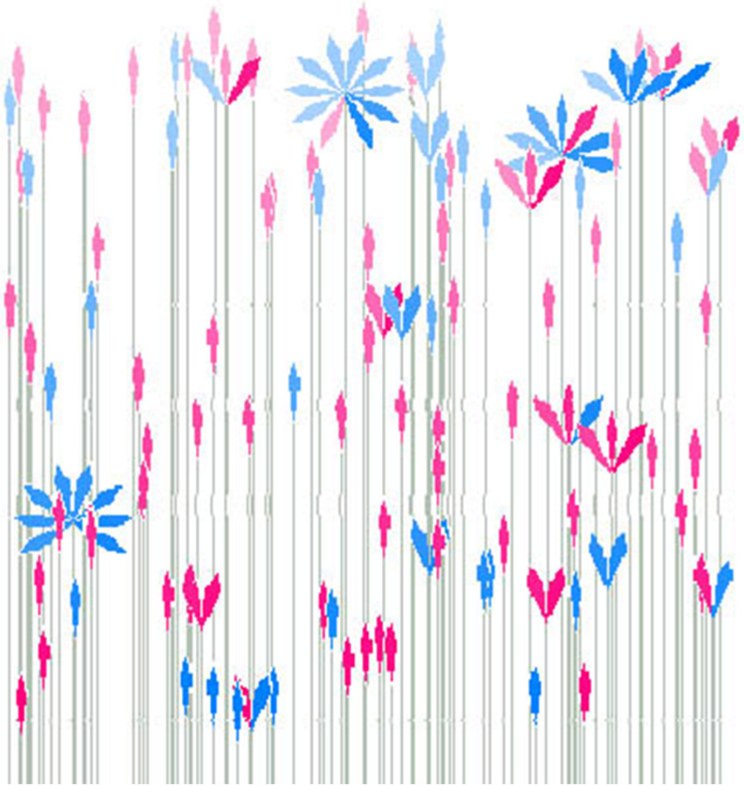
Each flower is a different user
Height indicates length of time at the board



Interpreting Displays



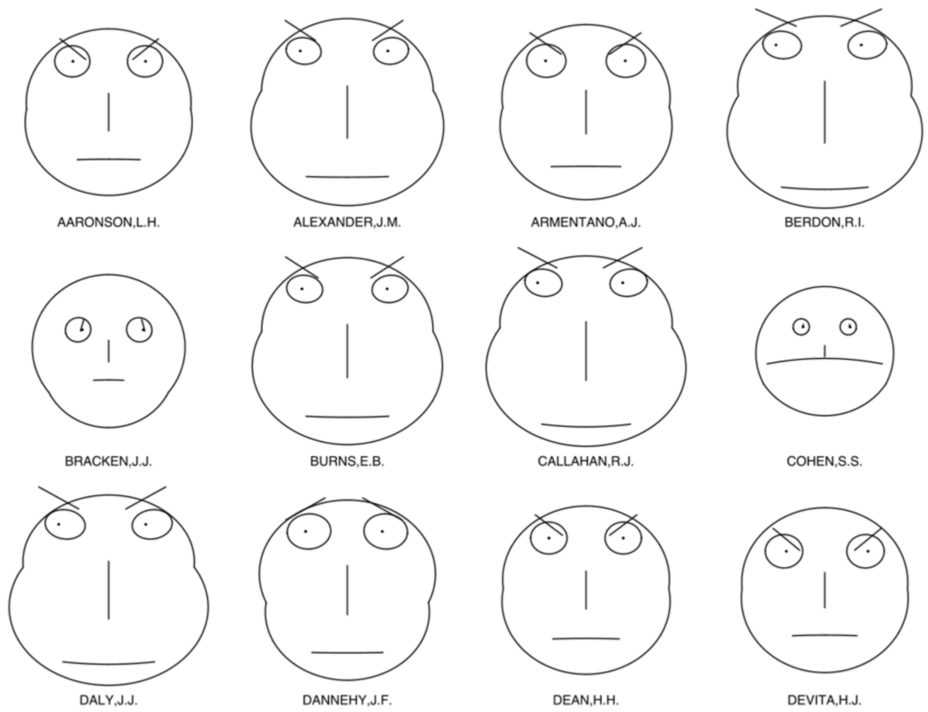
Group with one dominating person



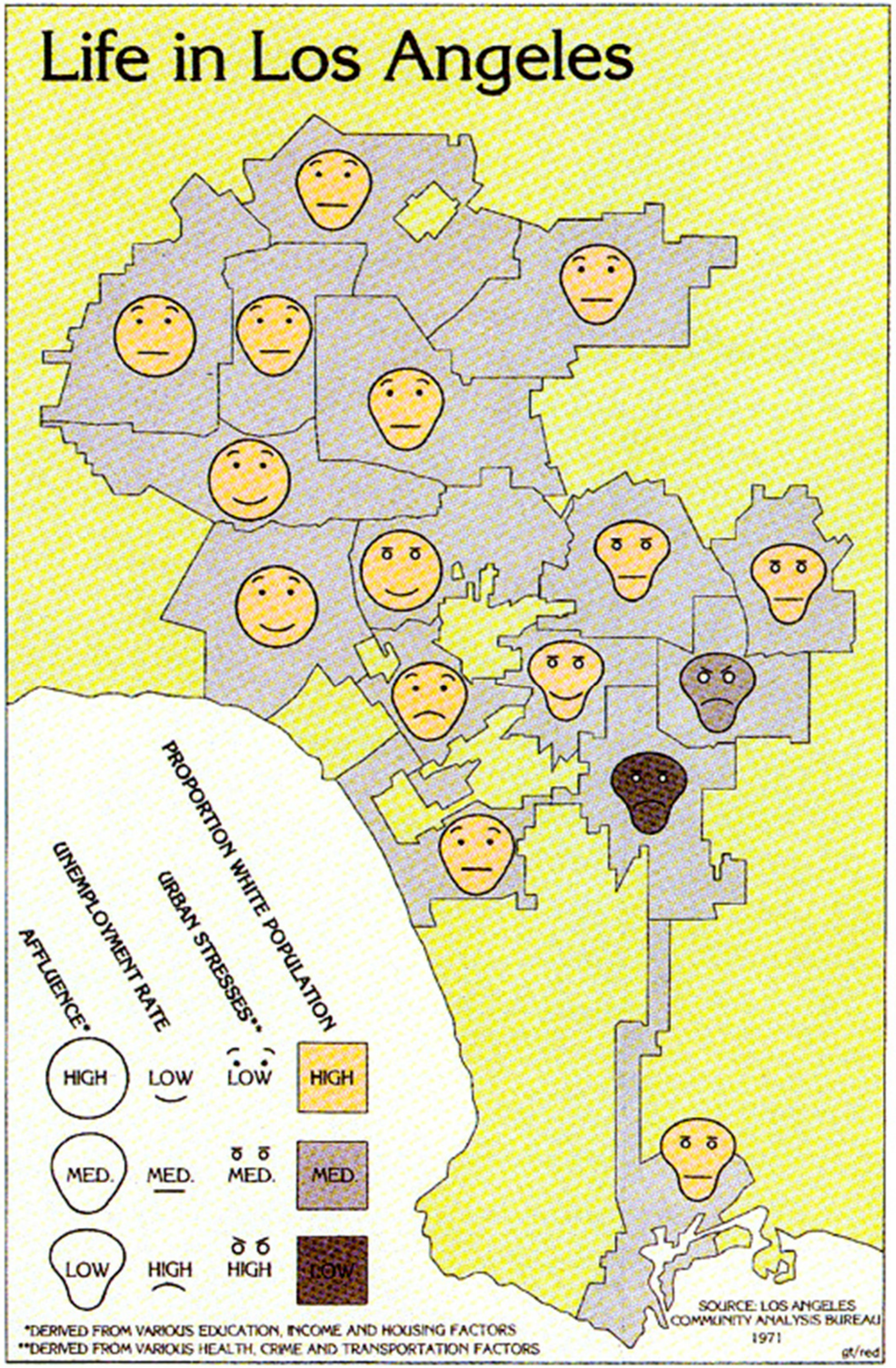
More democratic group

Chernoff Faces

- eyes, ears, mouth and nose represent values of the variables by their shape, size, placement and orientation
- Based on the idea that we are very very good at recognizing facial expression



Chernoff 1973



Life in Los Angeles
PROPORTION WHITE POPULATION
URBAN STRESSES**
UNEMPLOYMENT RATE
AFFLUENCE*
HIGH
LOW
LOW
HIGH
MED.
MED.
MED.
MED.
LOW
HIGH
HIGH
LOW
*DERIVED FROM VARIOUS EDUCATION, INCOME AND HOUSING FACTORS
**DERIVED FROM VARIOUS HEALTH, CRIME AND TRANSPORTATION FACTORS
SOURCE: LOS ANGELES COMMUNITY ANALYSIS BUREAU 1971
gt/red

Many more exist but…

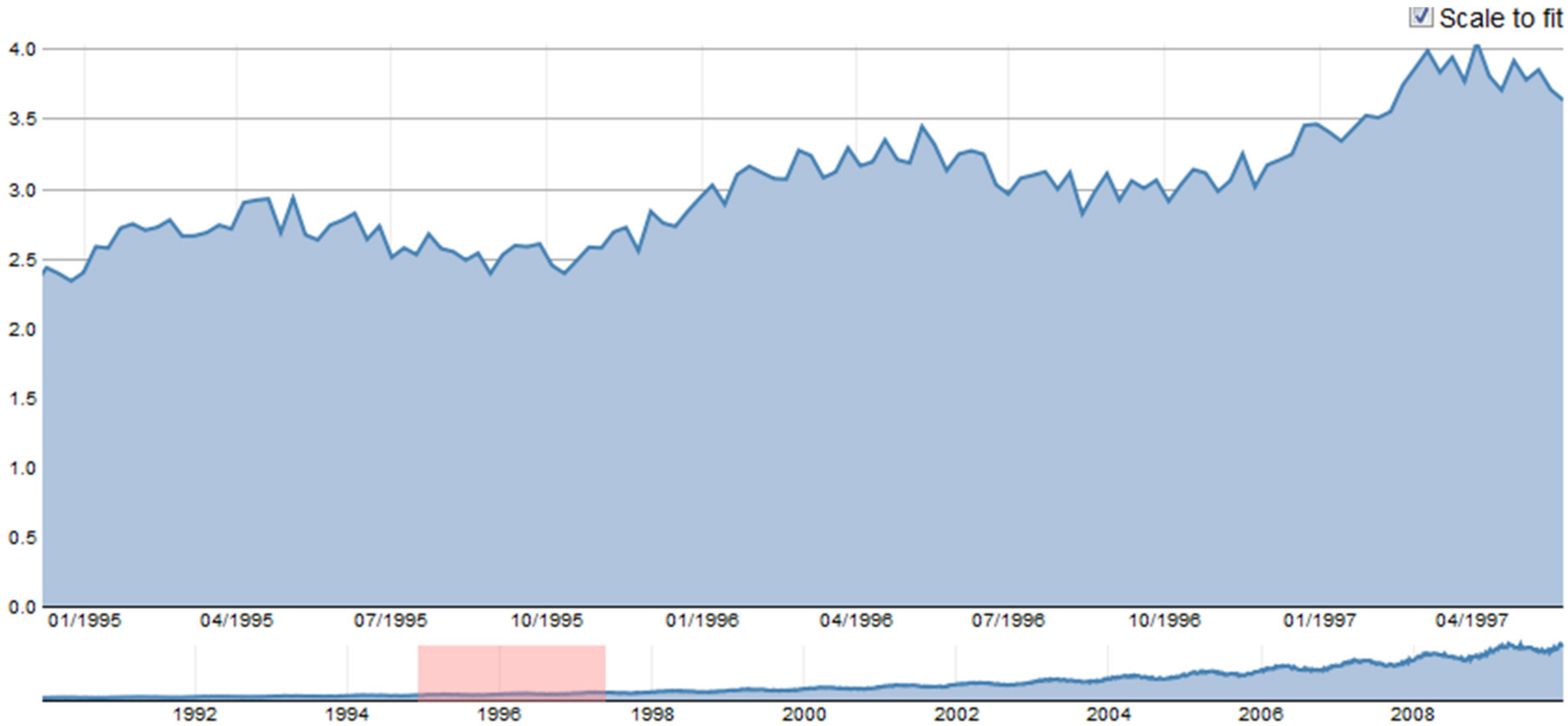
WE PROBABLY NEED A BREAK?

Visualization Techniques for

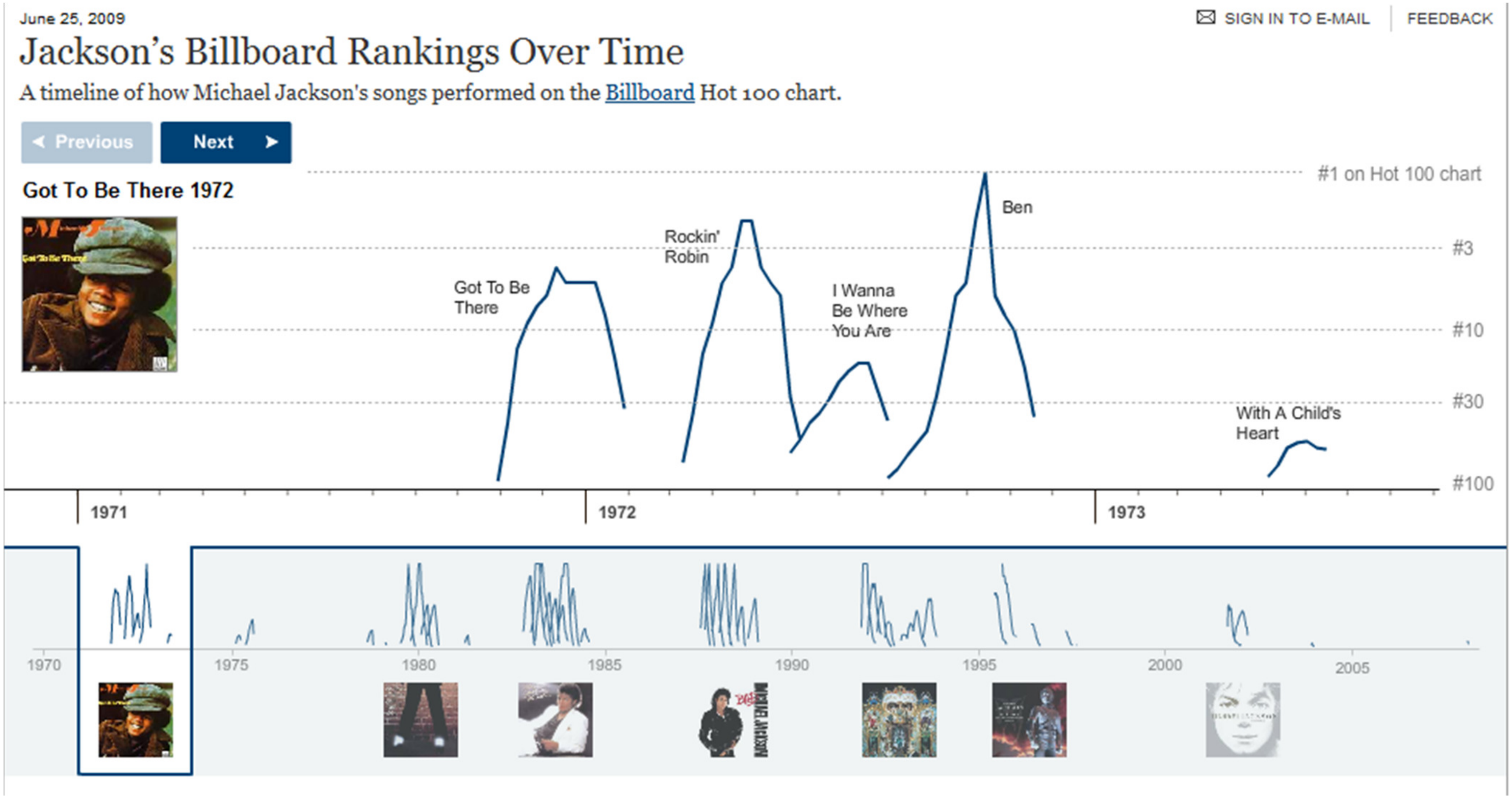
TEMPORAL DATA

Charts

- Focus+Context Chart



ProtoVis Example
http://vis.stanford.edu/protovis/ex/zoom.html



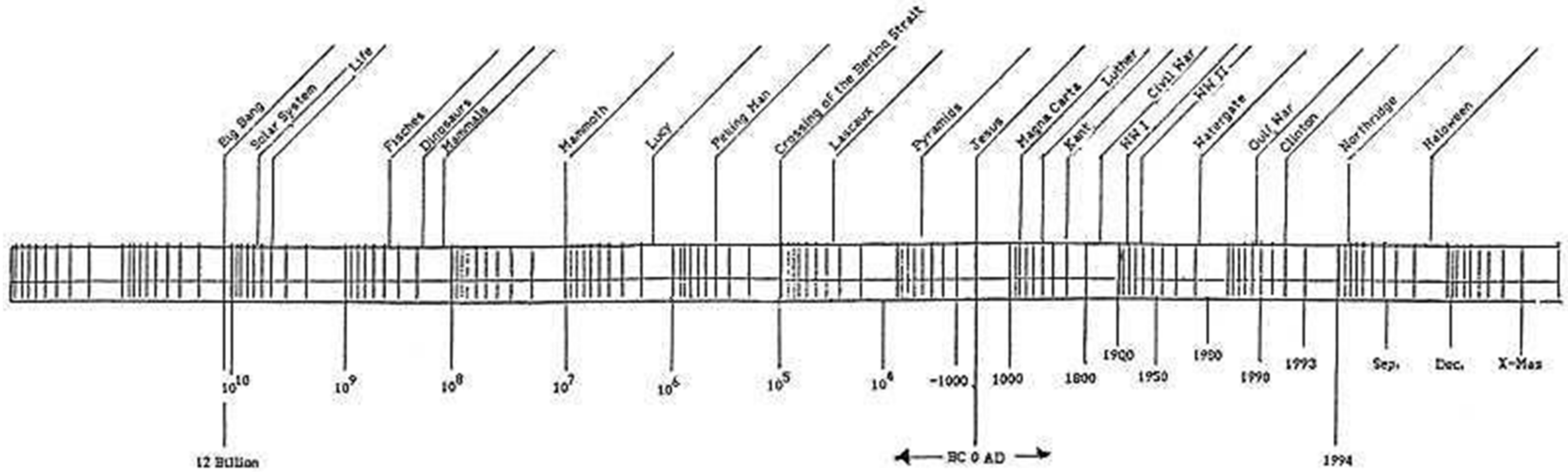
NYTimes

http://www.nytimes.com/interactive/2009/06/25/arts/0625-jackson-graphic.html

Log Scales

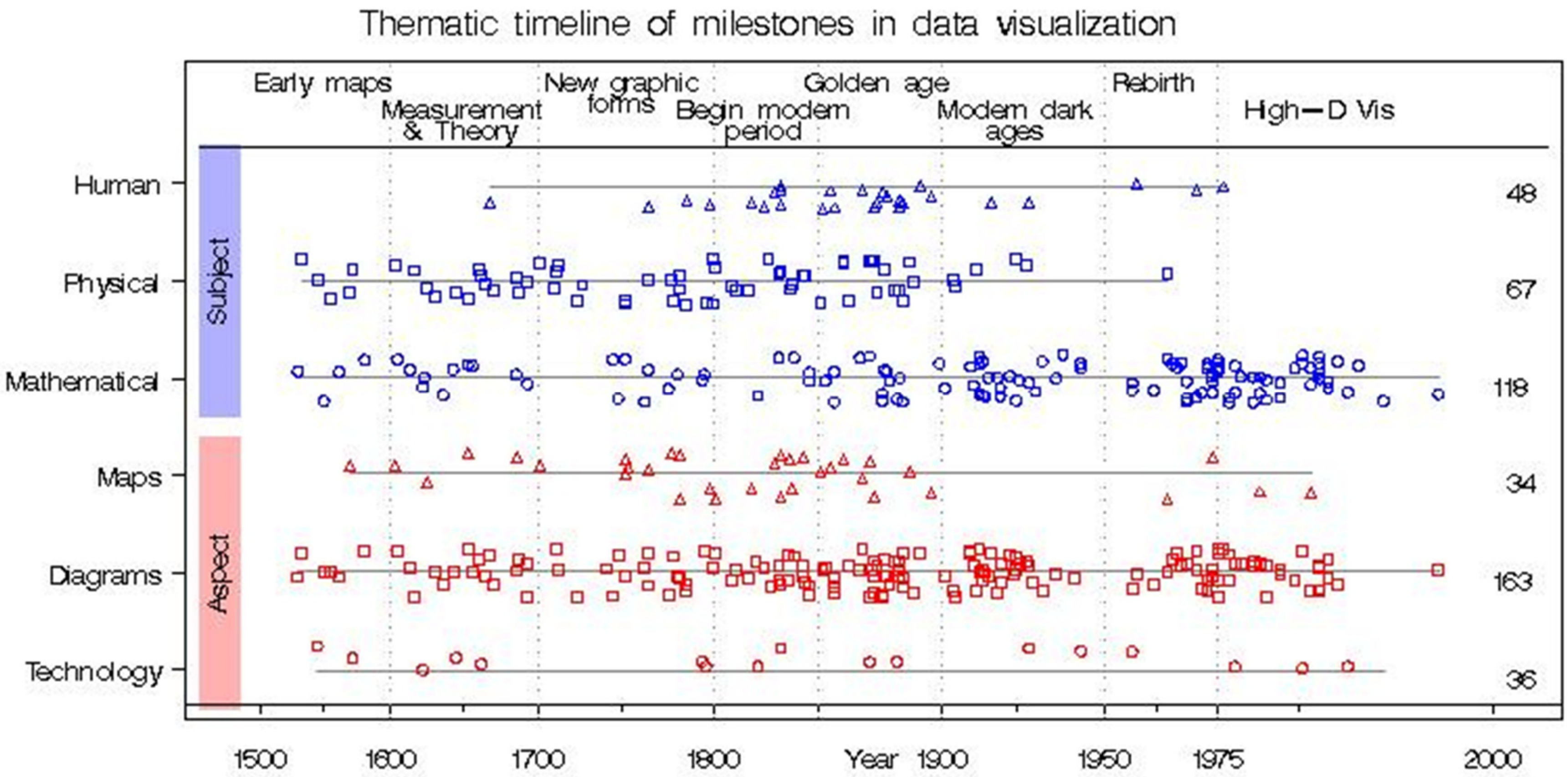
- If timescale too large to fit in one view

Logarithmic



http://www.math.yorku.ca/SCS/Gallery/timelines.html#hyperhistory

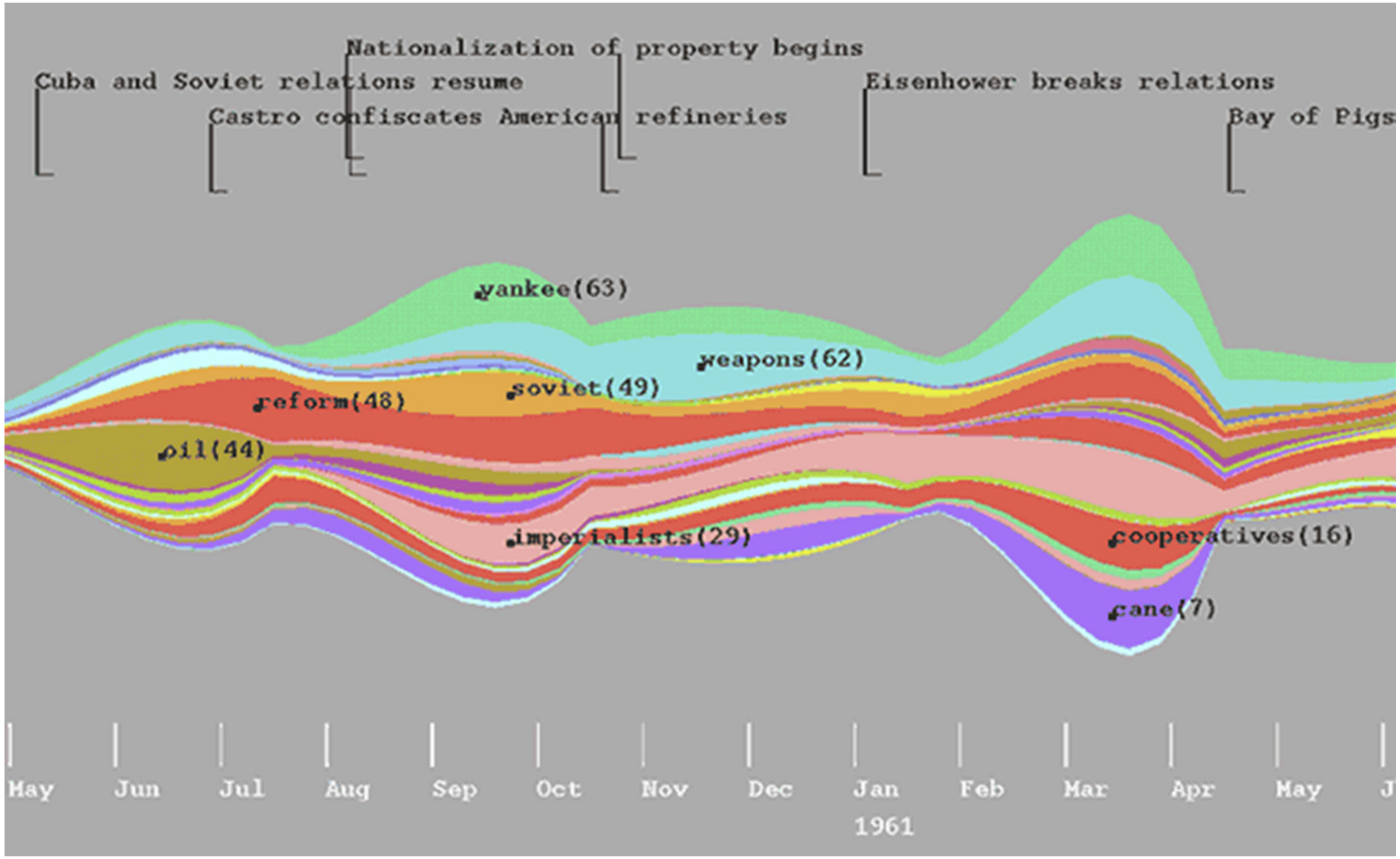
Reverse Square-Root Scale



http://www.math.yorku.ca/SCS/Gallery/images/timelines/milecatline.jpg

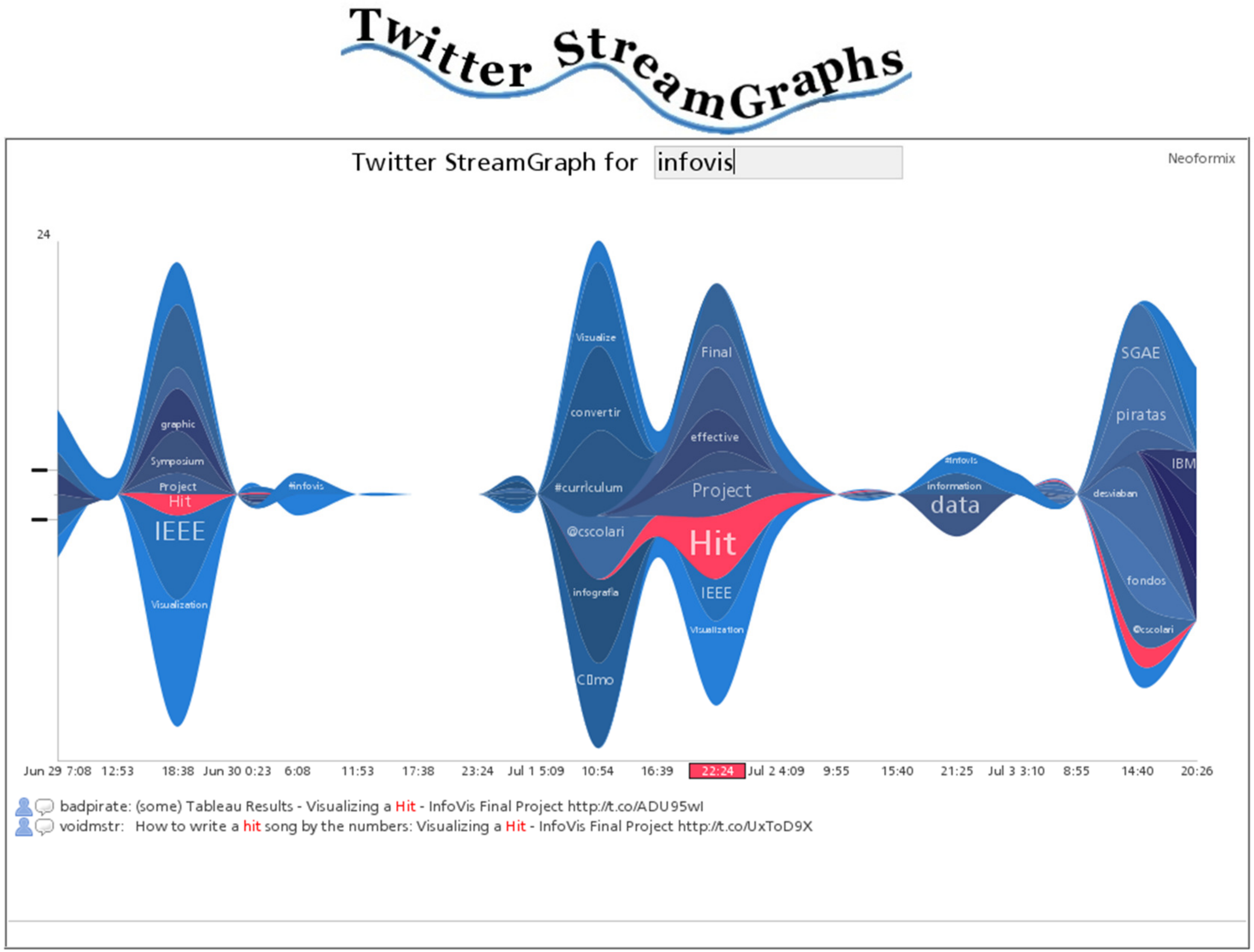
ThemeRiver

- Topic layered on top of each other around central line



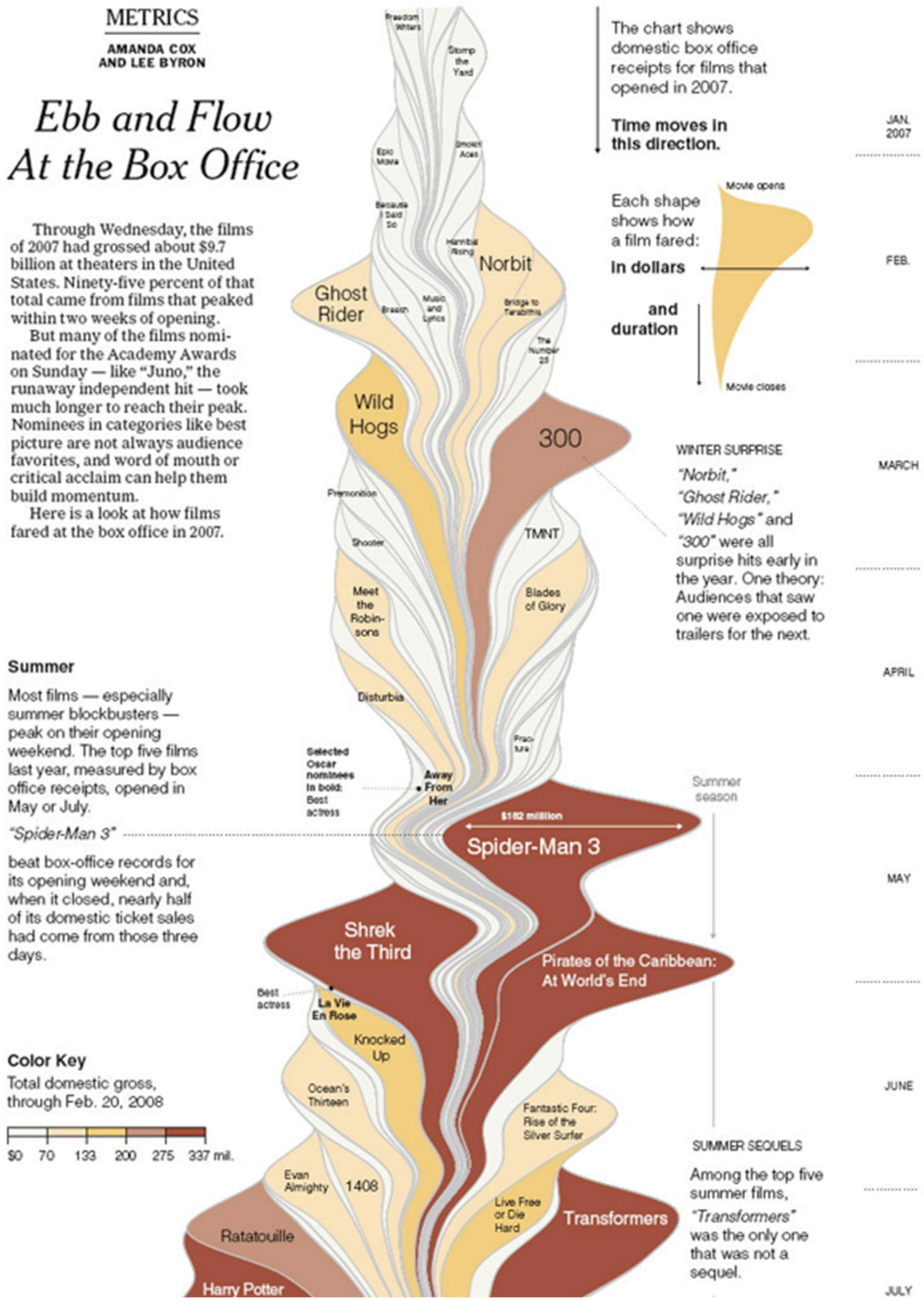
Havre et al., 2002

StreamGraph

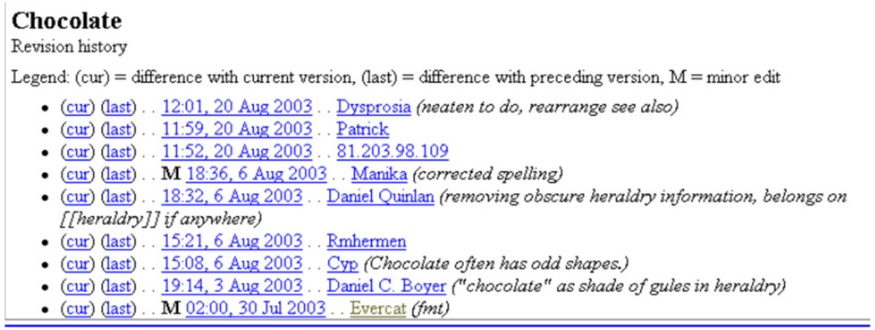


http://www.neoformix.com/Projects/TwitterStreamGraphs/view.php

StreamGraph



HistoryFlow

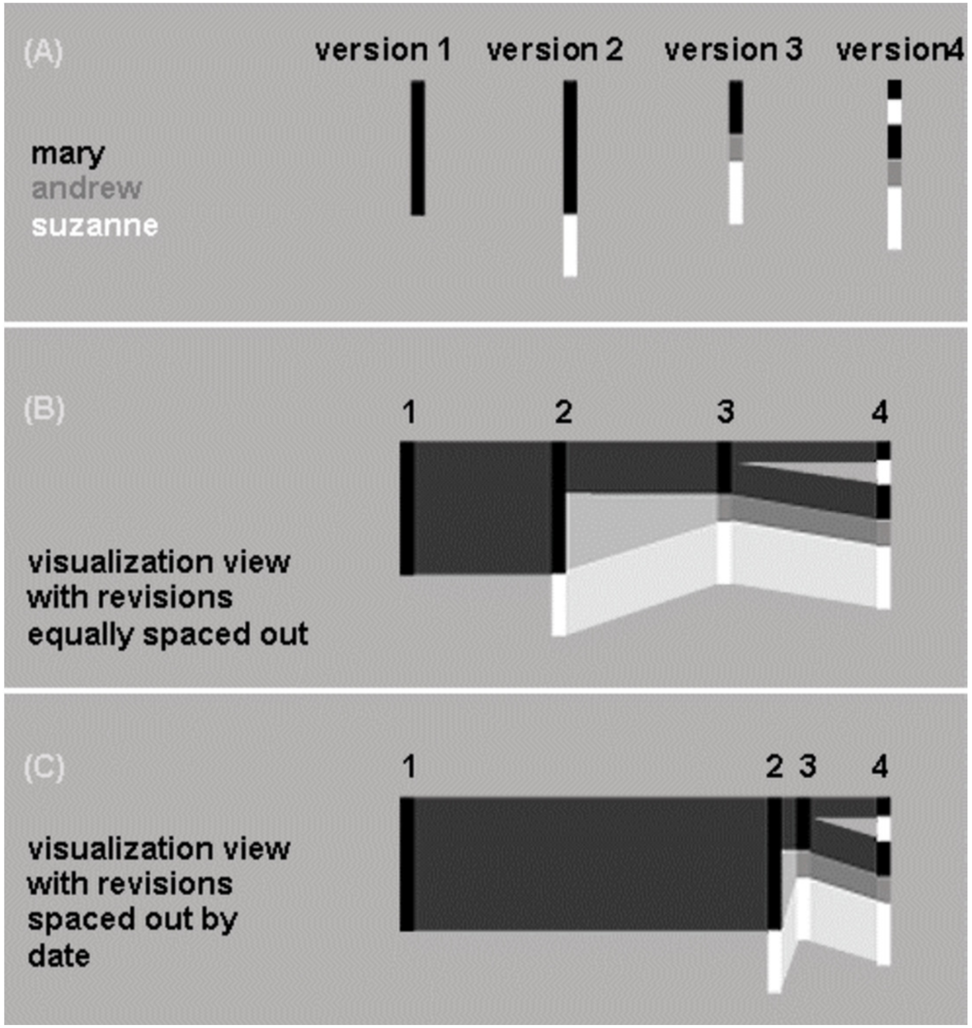

Chocolate

Revision history

Legend: (cur) = difference with current version, (last) = difference with preceding version, M = minor edit

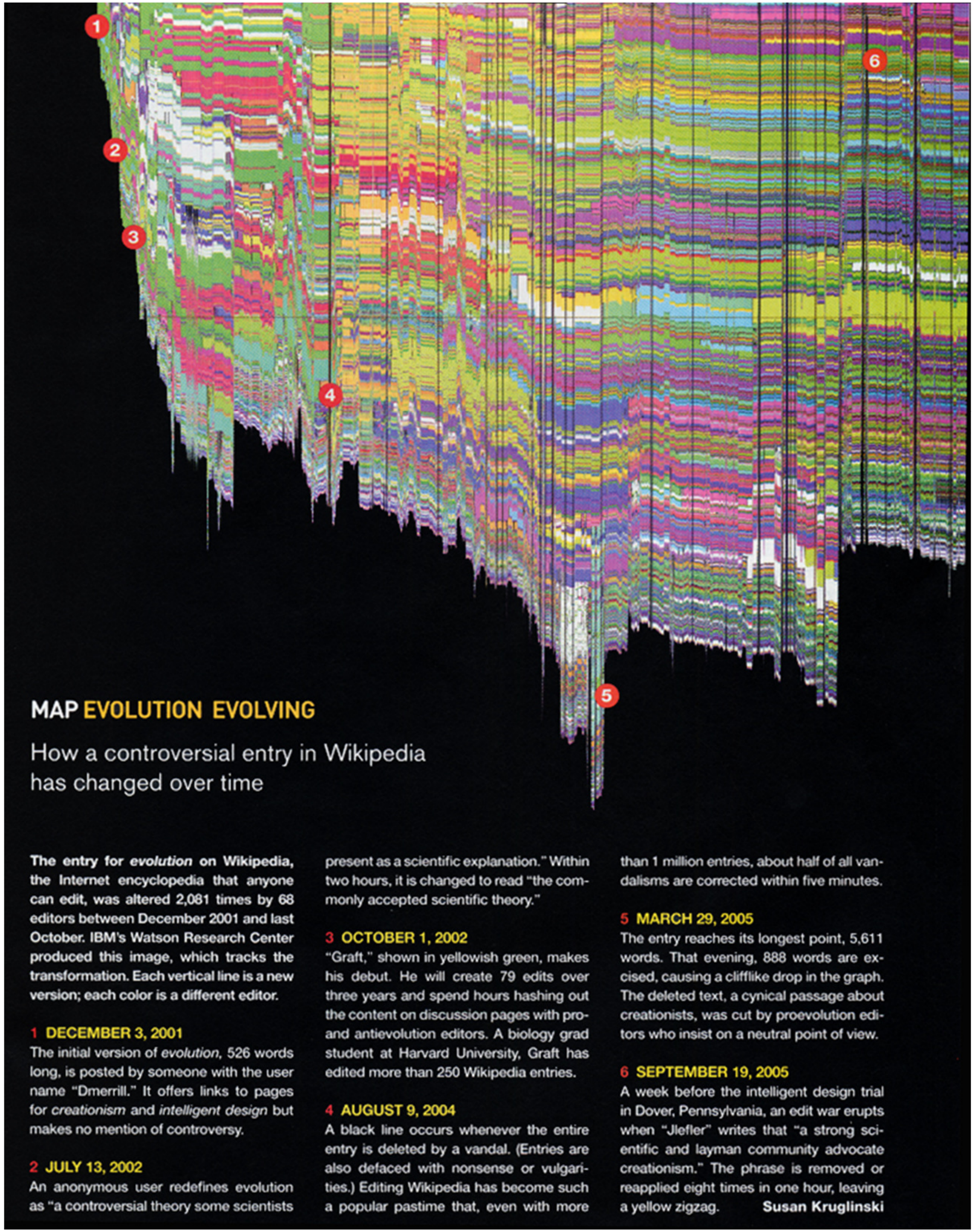
- (cur) (last) . . 12:01, 20 Aug 2003 . . Dysprosia *(neaten to do, rearrange see also)*
- (cur) (last) . . 11:59, 20 Aug 2003 . . Patrick
- (cur) (last) . . 11:52, 20 Aug 2003 . . 81.203.98.109
- (cur) (last) . . **M** 18:36, 6 Aug 2003 . . Manika *(corrected spelling)*
- (cur) (last) . . 18:32, 6 Aug 2003 . . Daniel Quinlan *(removing obscure heraldry information, belongs on [[heraldry]] if anywhere)*
- (cur) (last) . . 15:21, 6 Aug 2003 . . Rmhermen
- (cur) (last) . . 15:08, 6 Aug 2003 . . Cyp *(Chocolate often has odd shapes.)*
- (cur) (last) . . 19:14, 3 Aug 2003 . . Daniel C. Boyer *("chocolate" as shade of gules in heraldry)*
- (cur) (last) . . **M** 02:00, 30 Jul 2003 . . Evercat *(fmt)*

Fig 1: *Detail of revision history of Wikipedia's Chocolate page.*



Viegas et al., 2004

HistoryFlow



MAP **EVOLUTION EVOLVING**

How a controversial entry in Wikipedia has changed over time

The entry for *evolution* on Wikipedia, the Internet encyclopedia that anyone can edit, was altered 2,081 times by 68 editors between December 2001 and last October. IBM's Watson Research Center produced this image, which tracks the transformation. Each vertical line is a new version; each color is a different editor.

1 DECEMBER 3, 2001
The initial version of *evolution*, 526 words long, is posted by someone with the user name "Dmerrill." It offers links to pages for *creationism* and *intelligent design* but makes no mention of controversy.

2 JULY 13, 2002
An anonymous user redefines evolution as "a controversial theory some scientists present as a scientific explanation." Within two hours, it is changed to read "the commonly accepted scientific theory."

3 OCTOBER 1, 2002
"Graft," shown in yellowish green, makes his debut. He will create 79 edits over three years and spend hours hashing out the content on discussion pages with pro- and antievolution editors. A biology grad student at Harvard University, Graft has edited more than 250 Wikipedia entries.

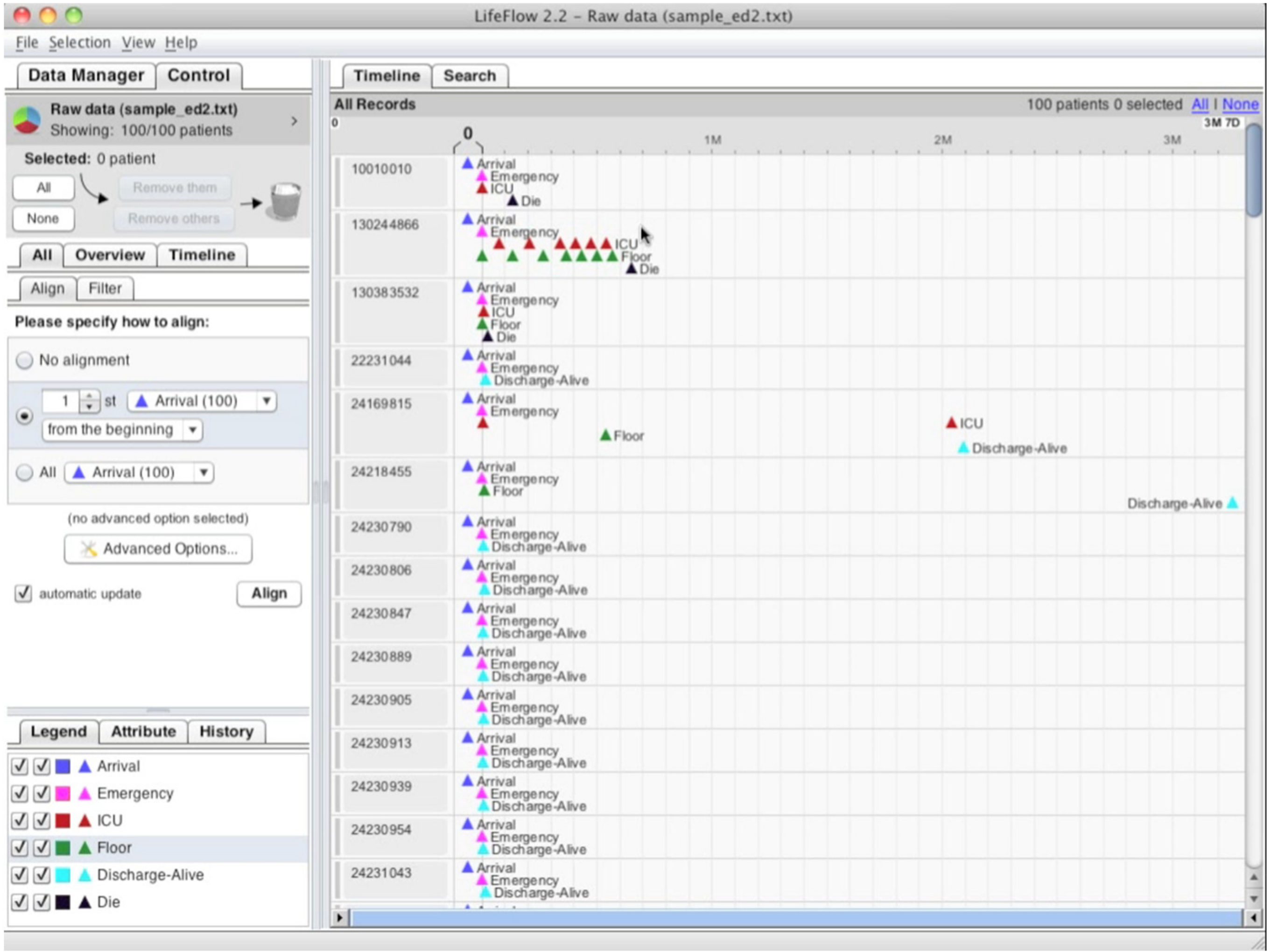
4 AUGUST 9, 2004
A black line occurs whenever the entire entry is deleted by a vandal. (Entries are also defaced with nonsense or vulgarities.) Editing Wikipedia has become such a popular pastime that, even with more than 1 million entries, about half of all vandalisms are corrected within five minutes.

5 MARCH 29, 2005
The entry reaches its longest point, 5,611 words. That evening, 888 words are excised, causing a clifflike drop in the graph. The deleted text, a cynical passage about creationists, was cut by proevolution editors who insist on a neutral point of view.

6 SEPTEMBER 19, 2005
A week before the intelligent design trial in Dover, Pennsylvania, an edit war erupts when "Jlefler" writes that "a strong scientific and layman community advocate creationism." The phrase is removed or reapplied eight times in one hour, leaving a yellow zigzag. **Susan Kruglinski**

Viegas et al., 2004

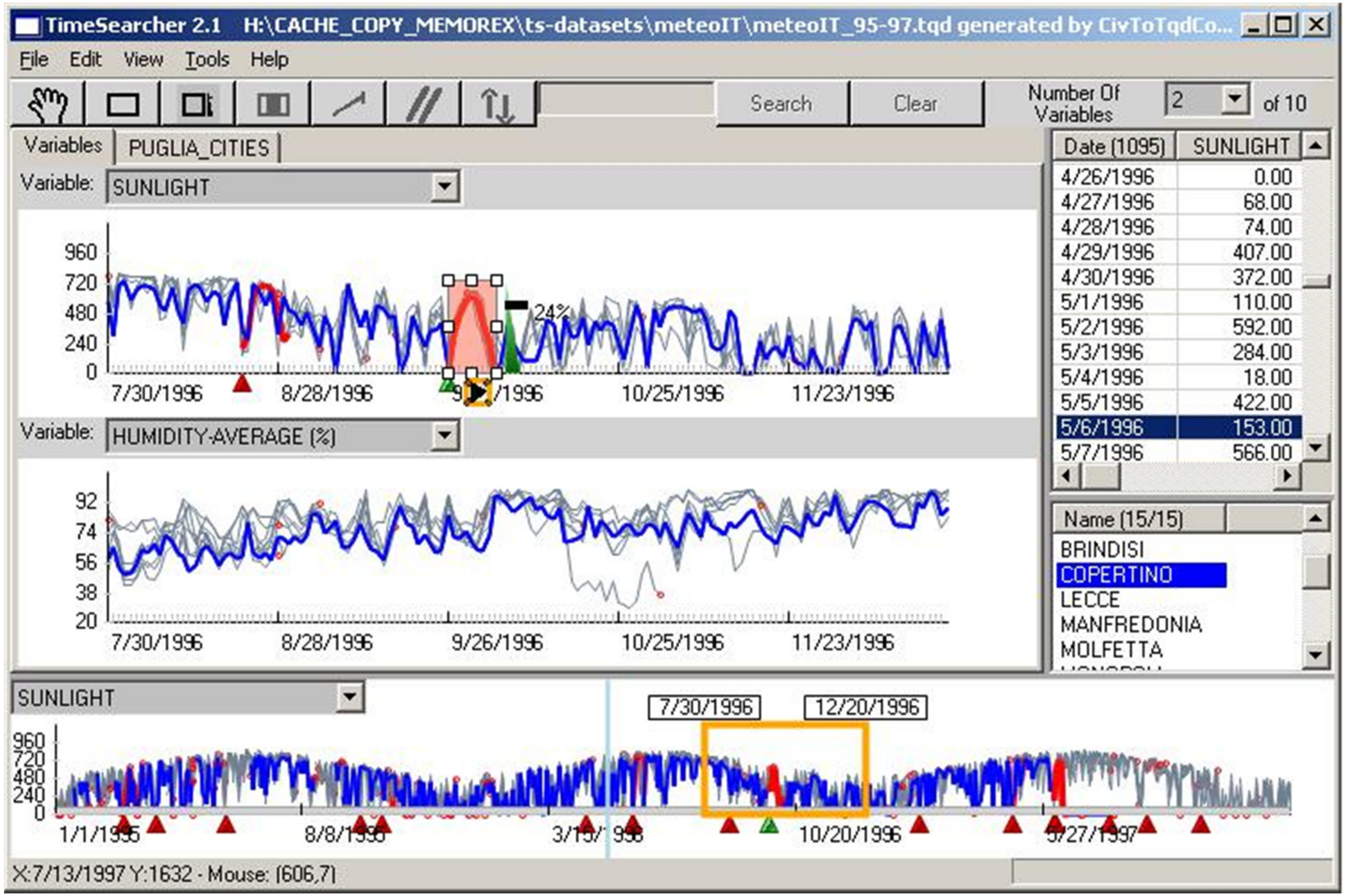
LifeFlow



http://www.cs.umd.edu/hcil/lifeflow/

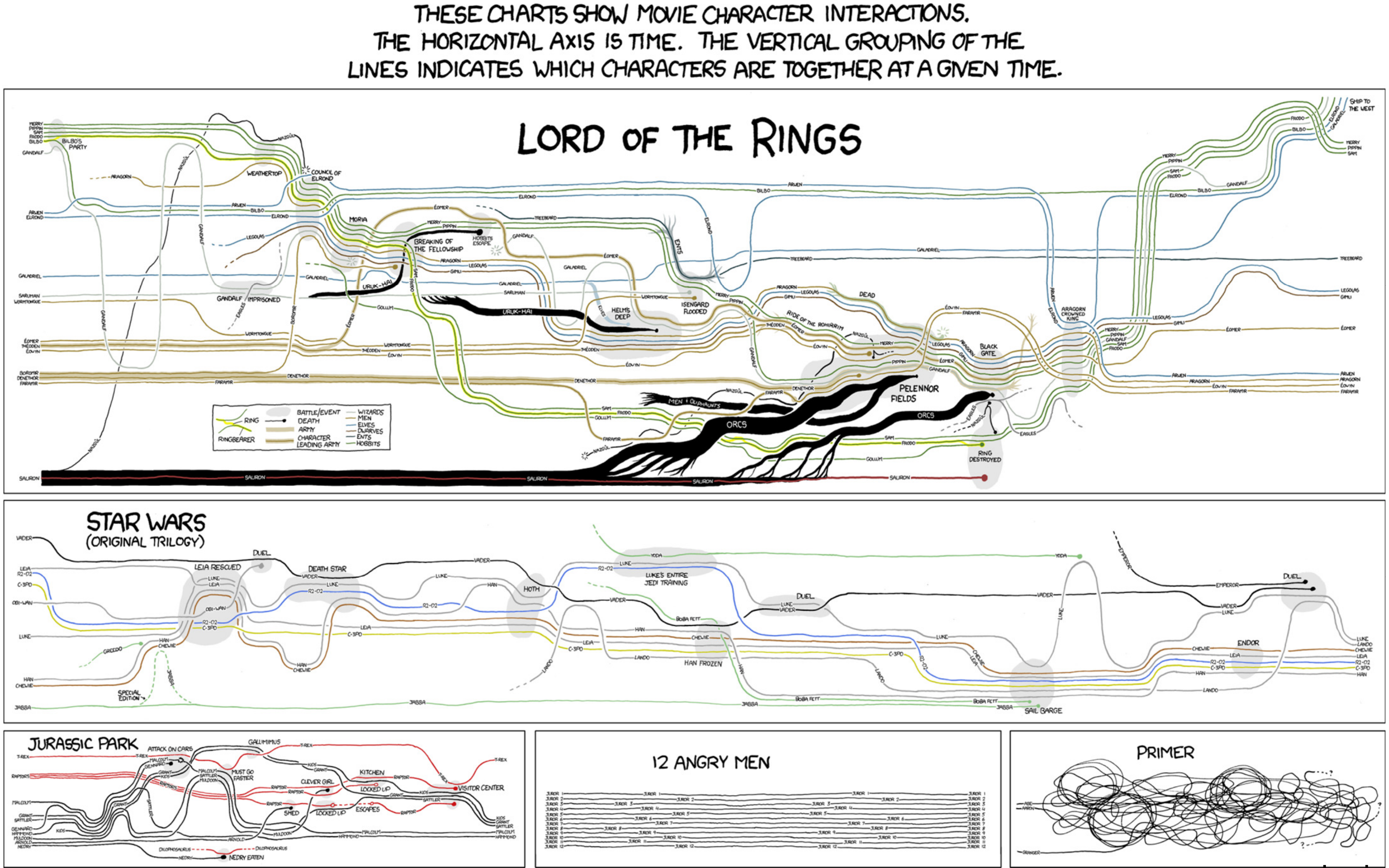
Wang et al., 2011

TimeSearcher



Download running version here: http://www.cs.umd.edu/hcil/timesearcher/

Narrative Chart



xkcd

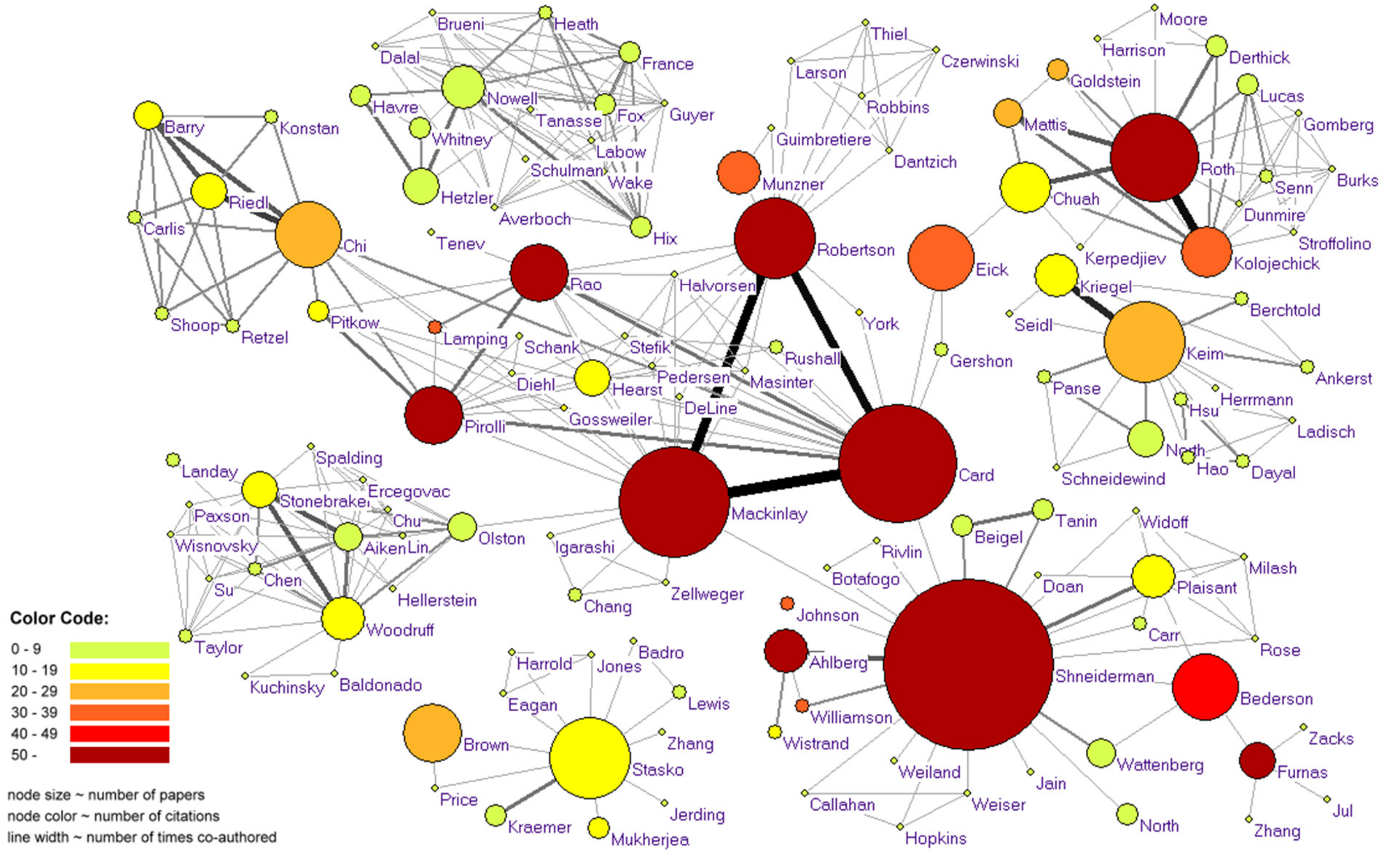
Visualization Techniques for

GRAPHS

Graphs

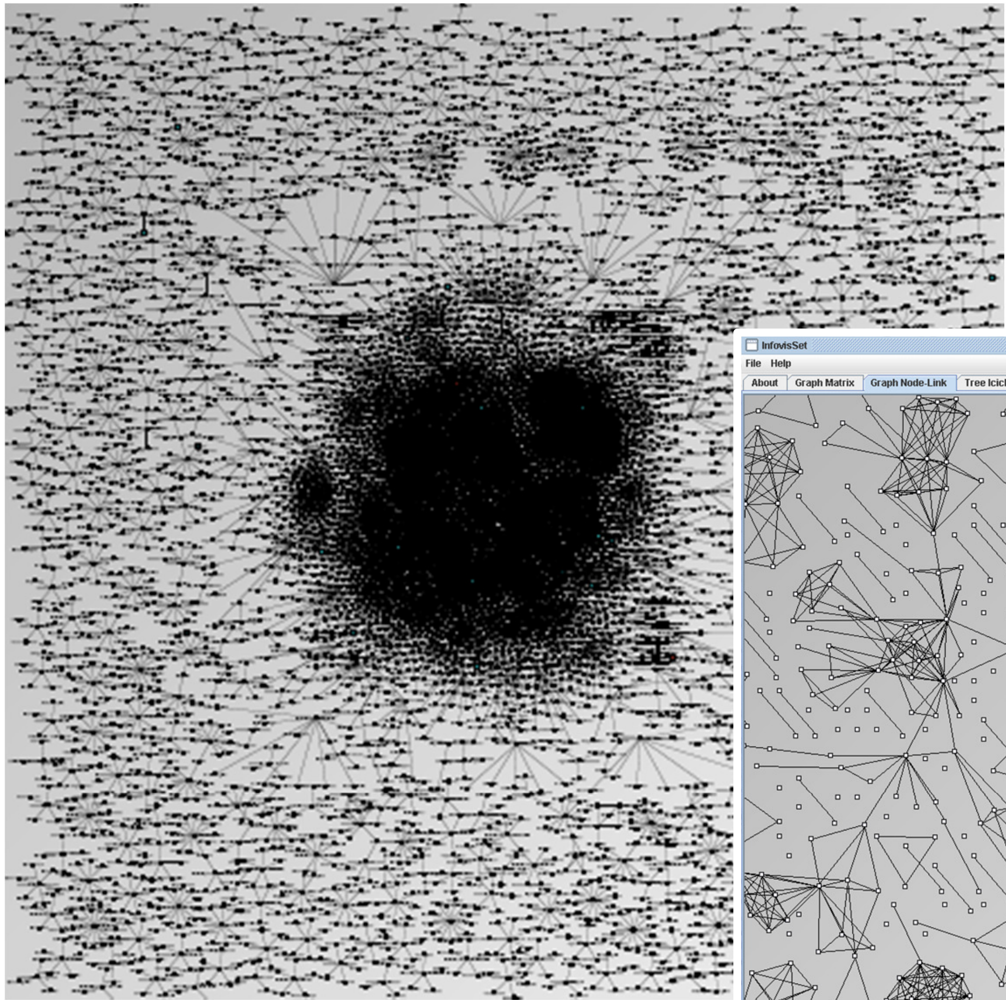
- Represent relationships among data
- Consist of nodes and edges
- Graph Drawing is a big research area in infovis
  - has its own conferences

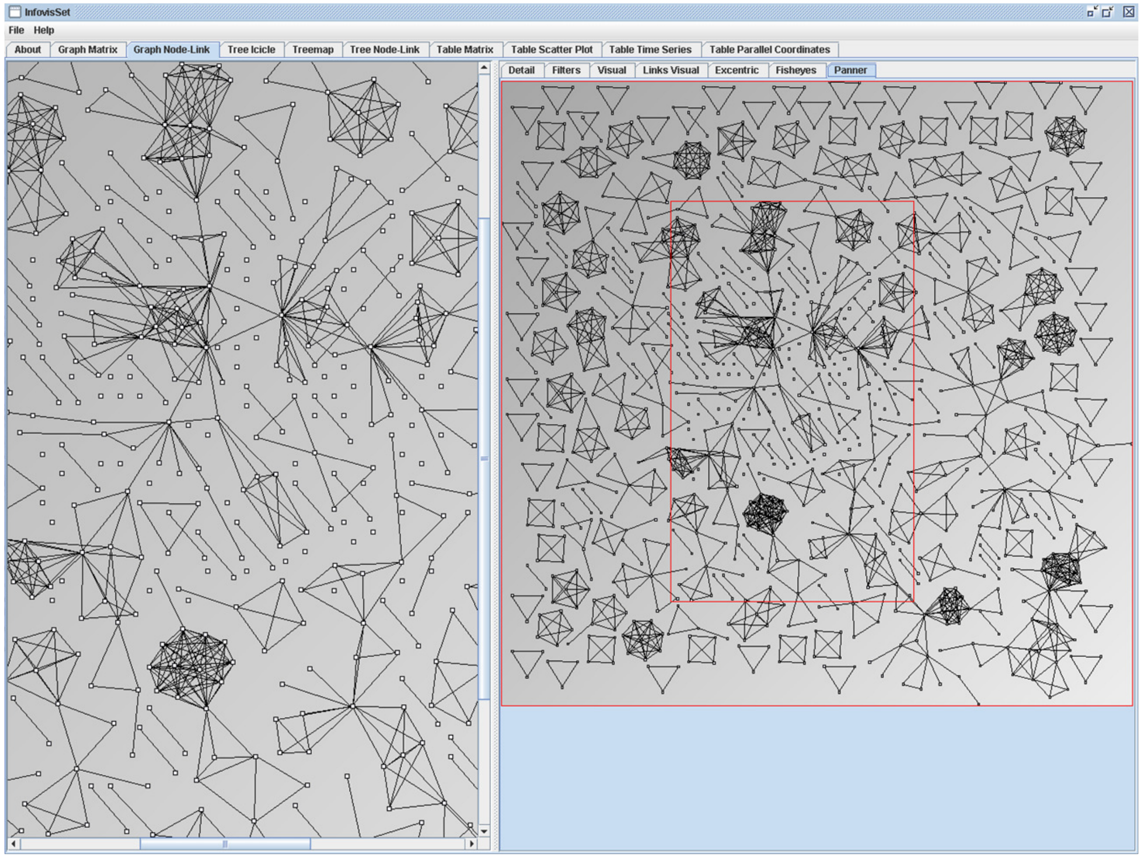
InfoVis Co-authoring (K. Börner et al.)

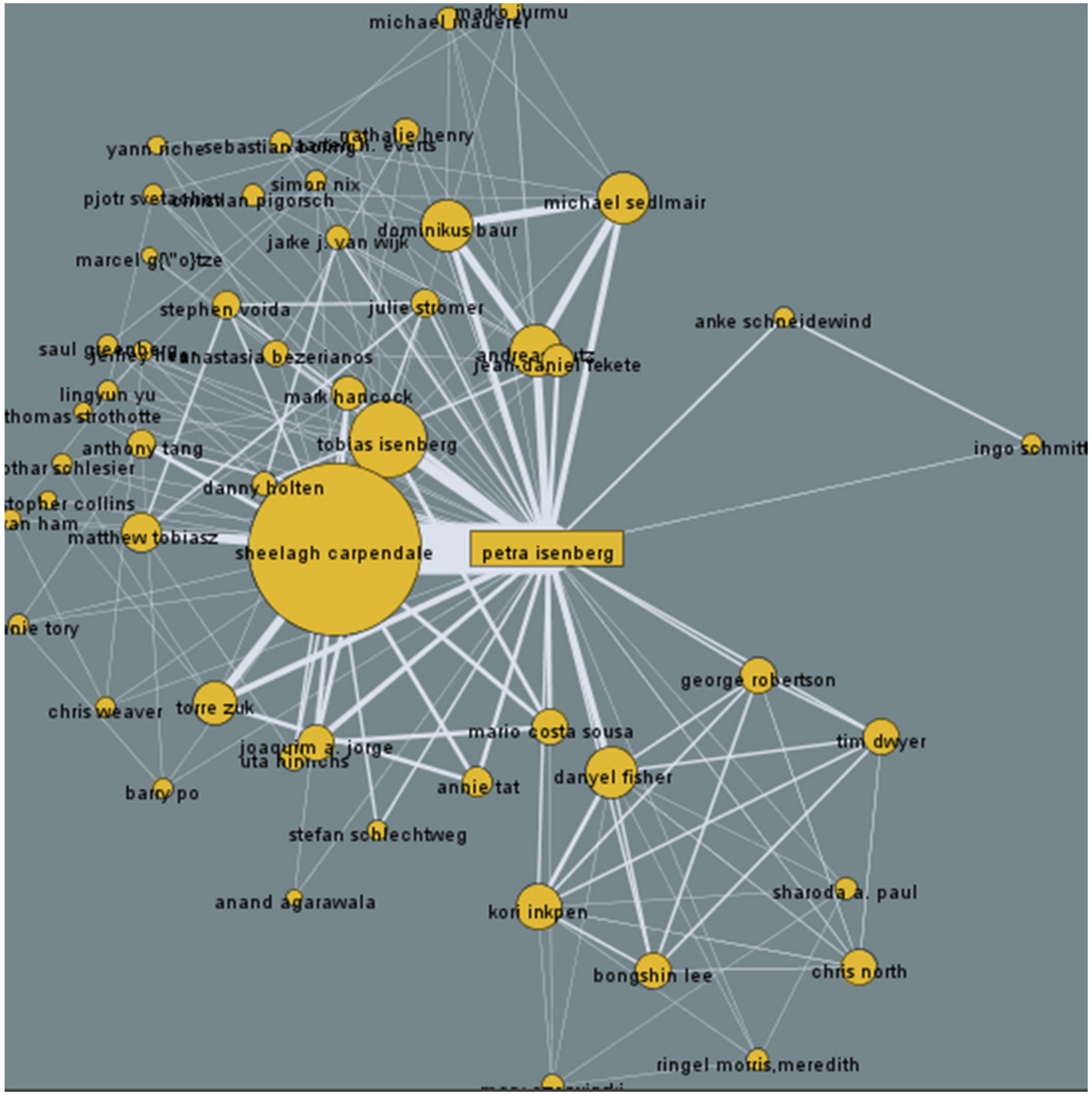


The following slides on graphs vs. matrices adapted from Jean-Daniel Fekete

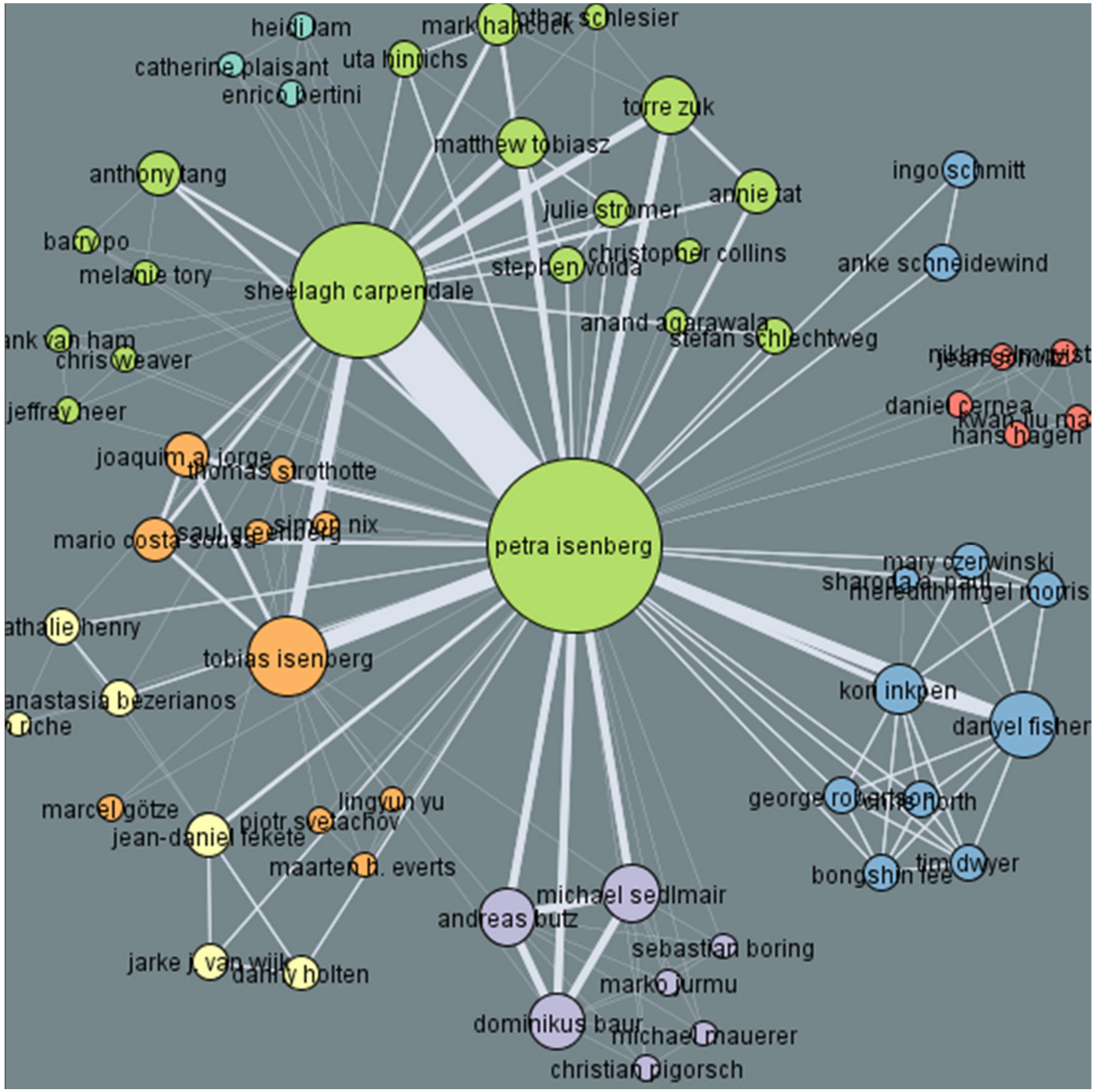
Generally, after loading…







“dummy” force-directed layout of my co-authorship network



Same data using LinLog Layout – a specialized layout for social networks

Graph Drawing Criteria

Crossings:	Minimization of edge crossings. Ideally a planar graph.
Area:	Minimization of the area of the drawing.
Total Edge Length:	Minimization of the sum of the lengths of the edges.
Maximum Edge Length:	Minimization of the maximum length of an edge.
Uniform Edge Length:	Minimization of the variations in edge length.
Total Bends:	Minimization of the total number of bends along an edge.
Maximum Bends:	Minimization of the maximum number of bends on an edge.
Uniform Bends:	Minimization of the number of bends on an edge.
Aspect Ratio:	Minimization of the aspect ratio of the drawing.
Symmetry:	Display symmetries of the graph in the drawing.
Angular Resolution:	Maximization of the smallest angle between two edges incident at a node.

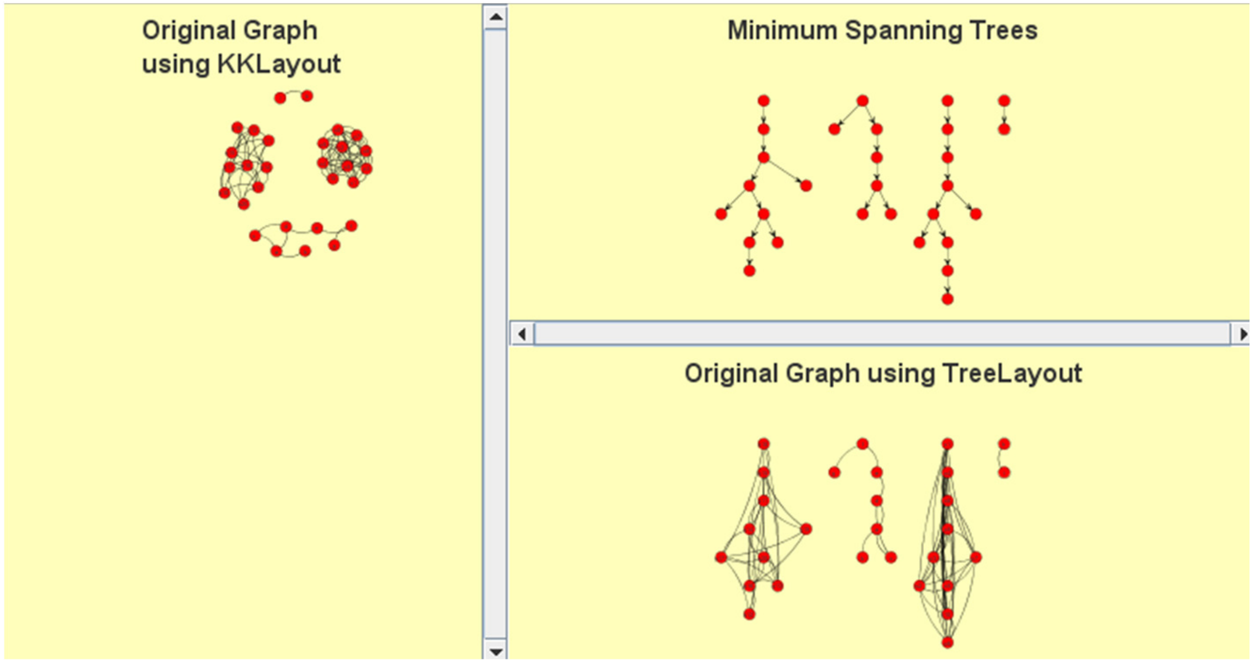
Battista, 1999

Approaches to Graph Layout

- Direct layout calculation using graph structure
  - Tree layout on spanning tree
  - Hierarchical layout
  - Adjacency matrix layout
- Optimization-based layout
  - Constraint satisfaction
  - Force-directed layout
- Attribute-driven layout
  - Layout using data attributes, not linkage

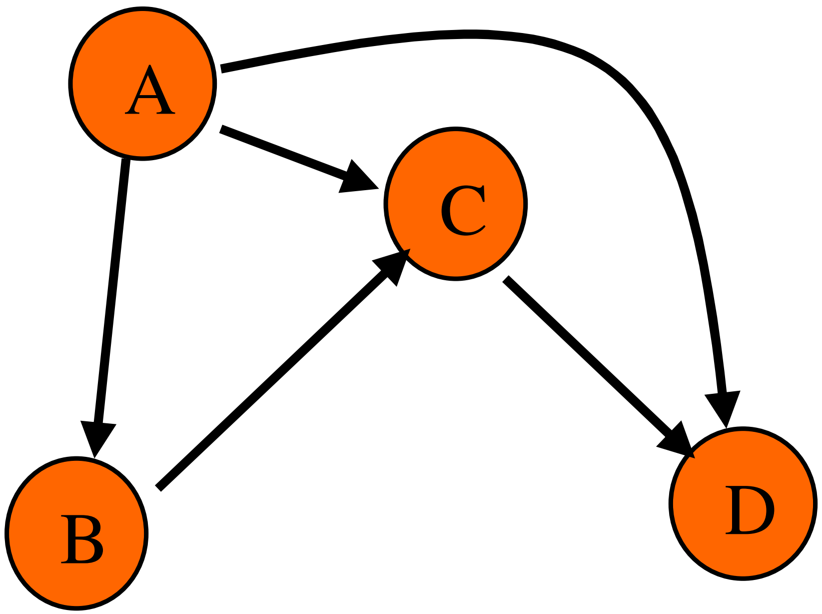
Spanning-Tree Layout

- Many graphs are tree-like or have useful spanning trees
  - Websites, Social Networks



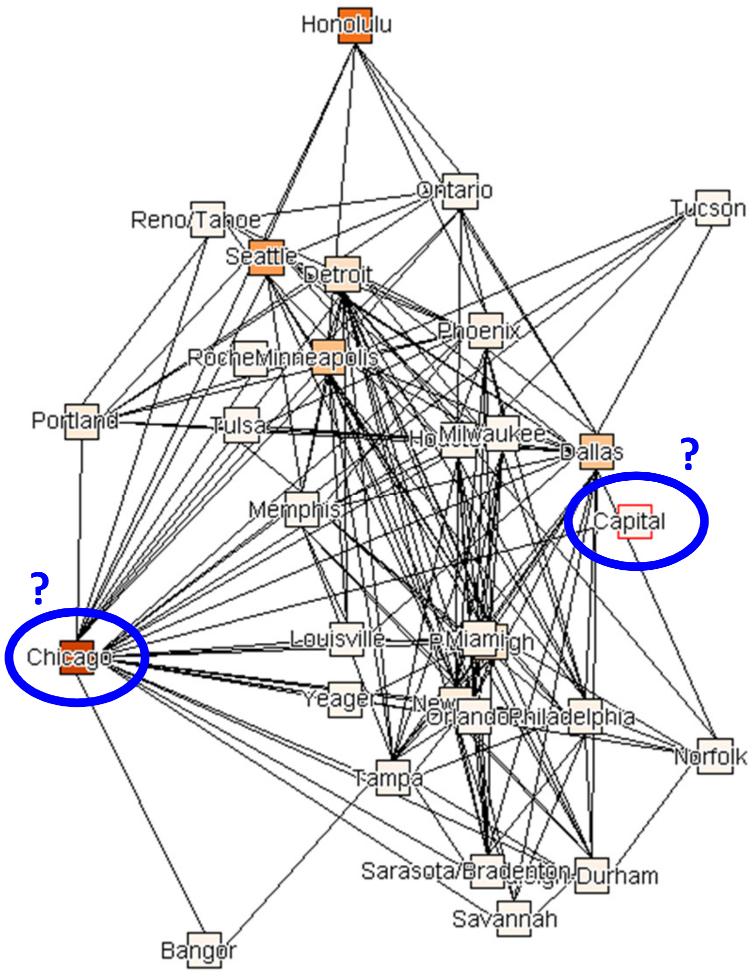
http://jung.sourceforge.net/applet/mstdemo.html

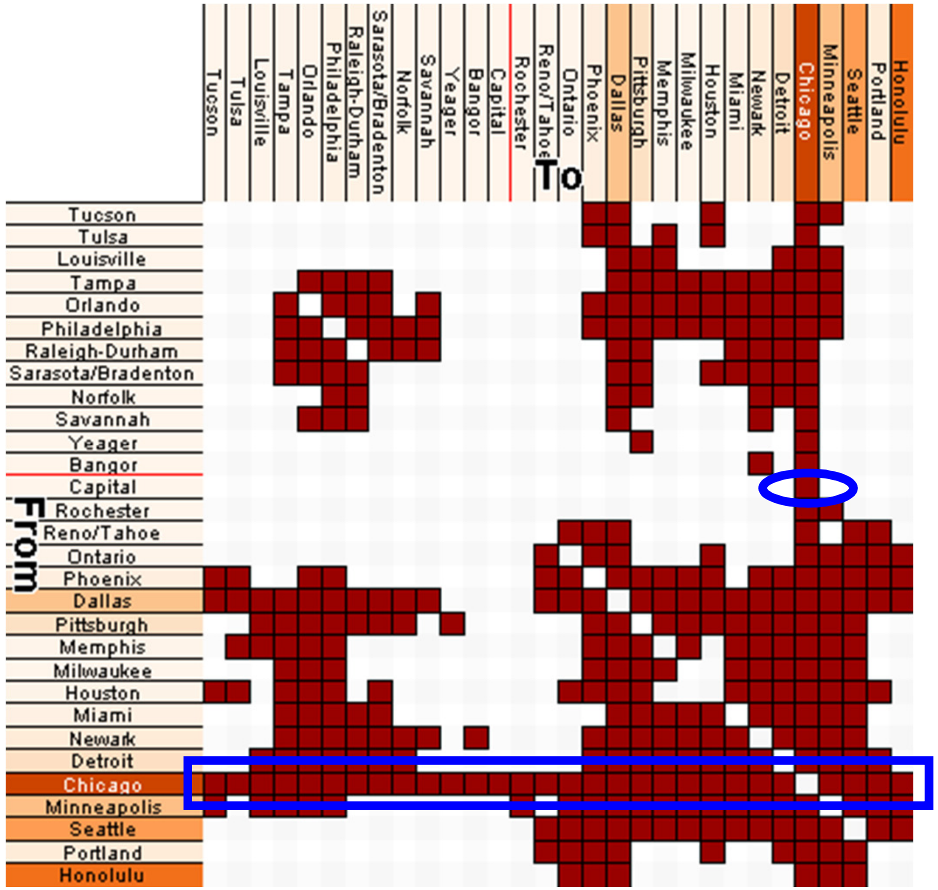
Alternative: Matrix Representation



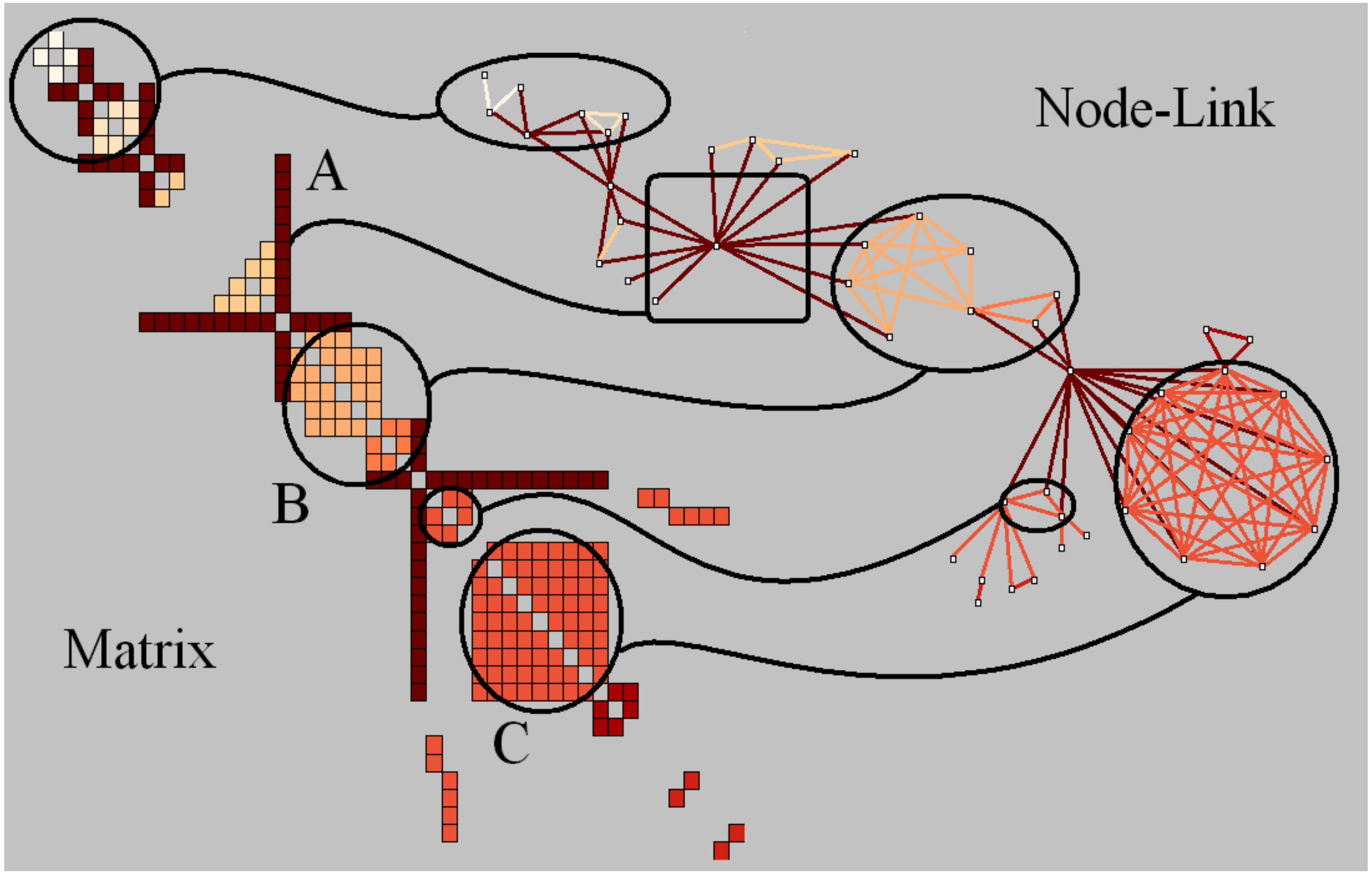
	A	B	C	D
A		X	X	X
B			X	
C				X
D				

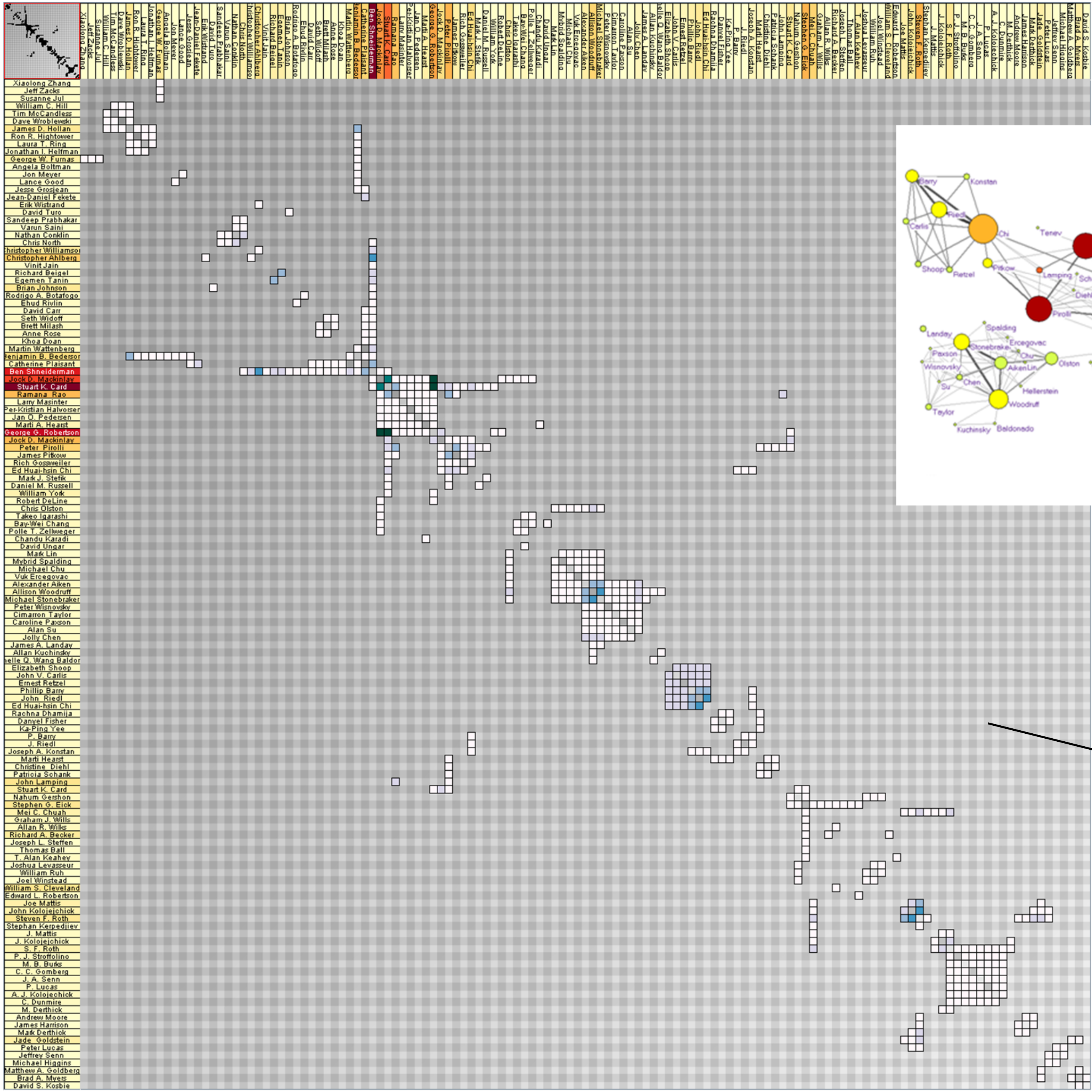
Matrix Visualization

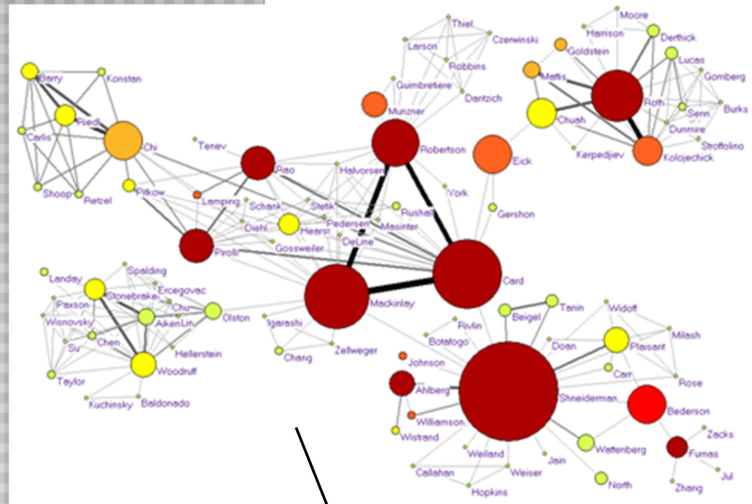




Visual Patterns with Ordered Matrices







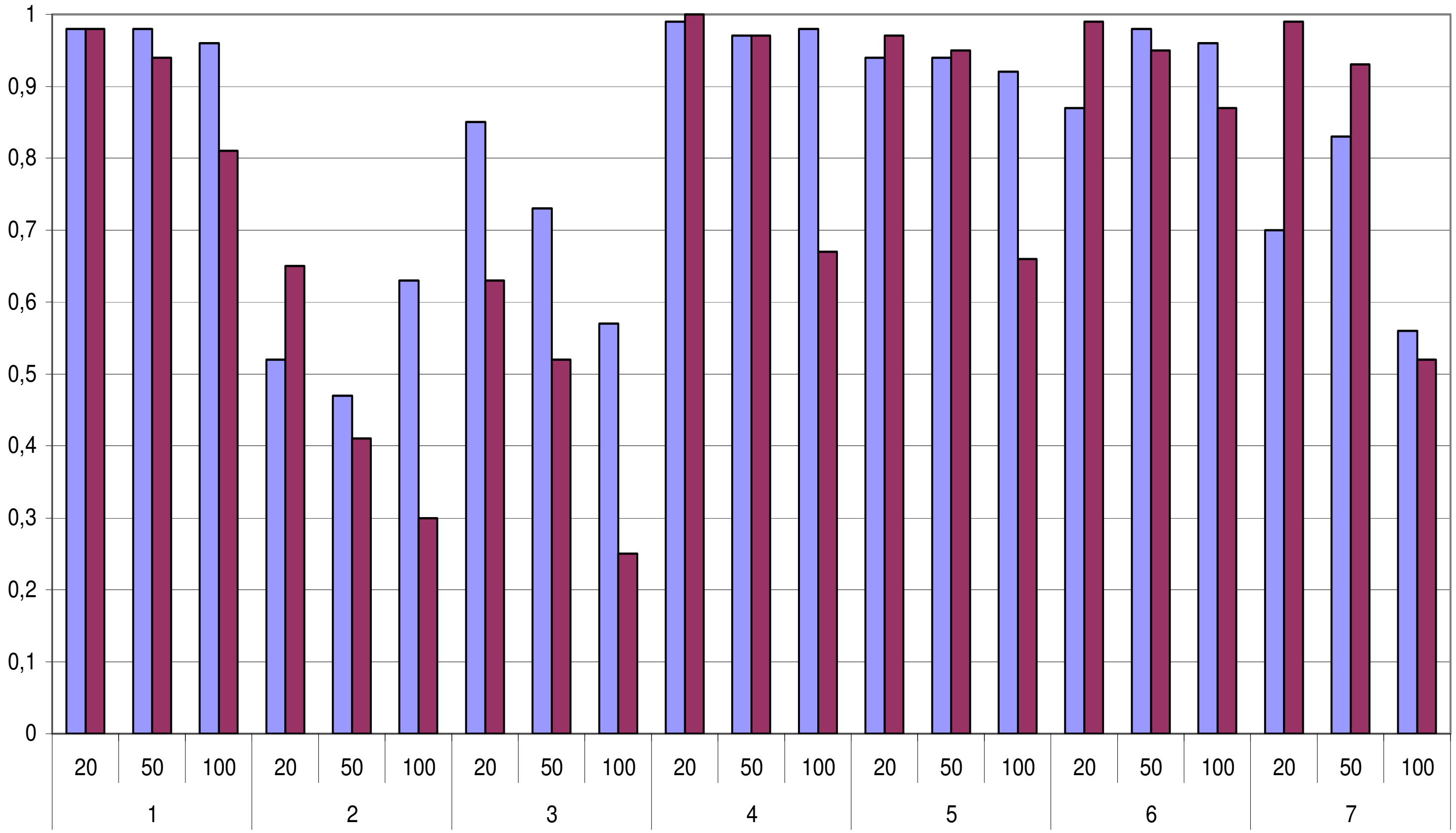
Communicate

Explore

User Study

- Tasks related to the overview
  - Number of vertices
  - Number of arcs
- Tasks related to graph elements
  - Finding an element (vertex, link)
  - Finding the most connected vertex (central actor, pivot, hub)
  - Finding a common neighbor
  - Finding a path
- Random graphs (3 sizes & 3 densities)
- 2 representations: Node-Link + Matrix

User Study



Completion time for the 7 tasks, 3 densities and 2 representations
(Node-Link in blue, Matrix in red)

User Study

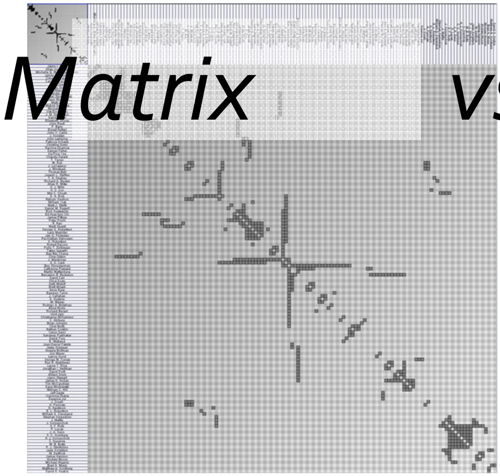
Results:

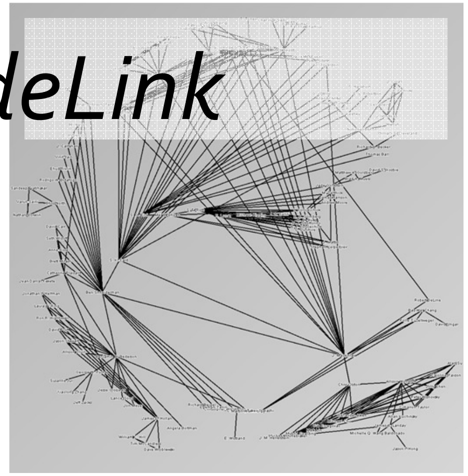
- Node-link diagrams are preferable for small sparse graphs (20 vertices)
- Matrices are more readable *wrt* dense graphs and medium/large graphs (> 20 vertices) *wrt* the selected tasks, except path finding

Reference:

Mohammad Ghoniem, Jean-Daniel Fekete and Philippe Castagliola *Readability of Graphs Using Node-Link and Matrix-Based Representations: Controlled Experiment and Statistical Analysis*, Information Visualization Journal, 4(2), Palgrave Macmillan, 2005, pp. 114-135.

Matrix vs. *NodeLink*

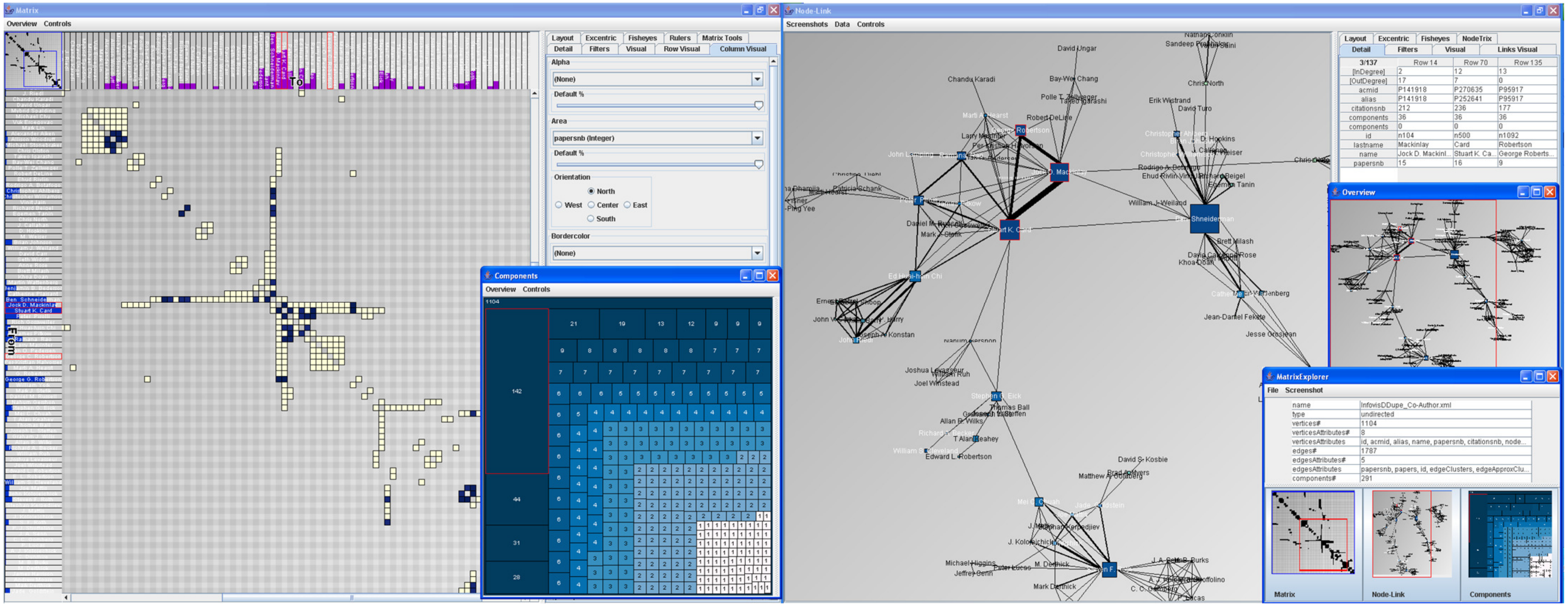




	Matrix	NodeLink
+	• Usable without reordering • No node overlapping No edge crossing ➔ Readable for dense graphs • Fast navigation • Fast manipulation ➔ Usable interactively • More readable for some tasks	• Familiar • Compact • More readable for path following • More effective for small graphs • More effective for sparse graphs
–	• Less familiar • Use more space • Weak for path following tasks	• Useless without layout • Node overlapping Edge crossing ➔ Not readable for dense graphs • Manipulation requires layout computation

MatrixExplorer *[Henry&Fekete06]*

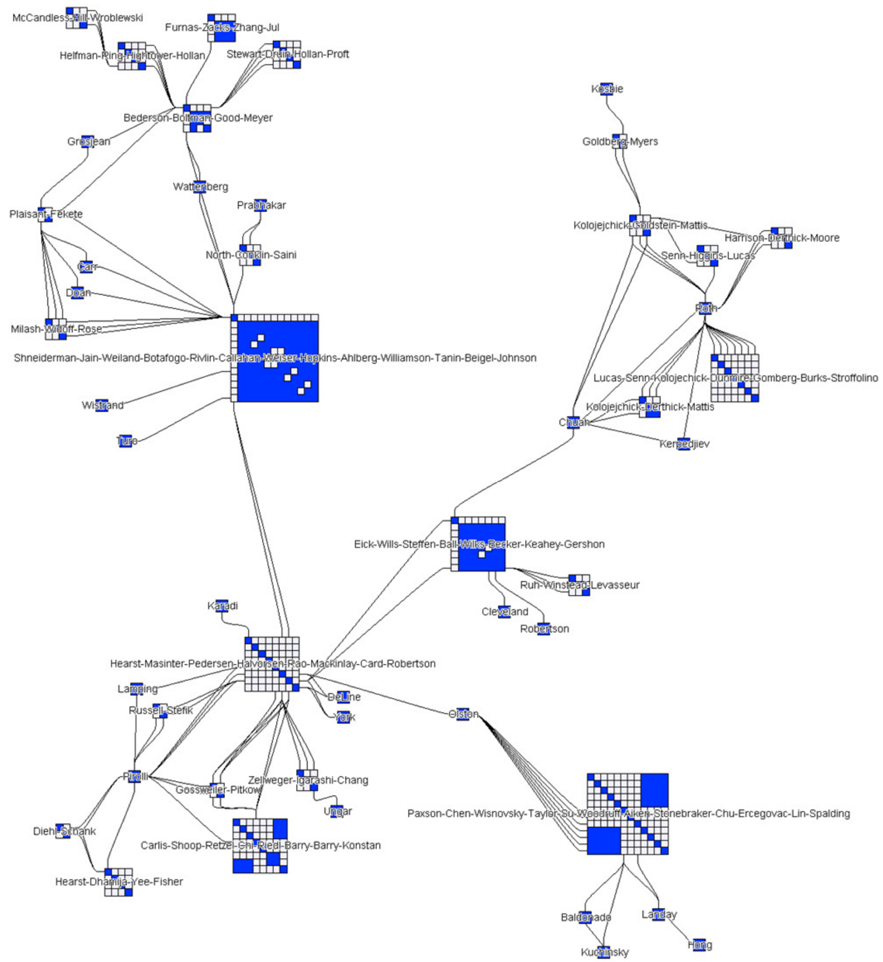
Combined representation



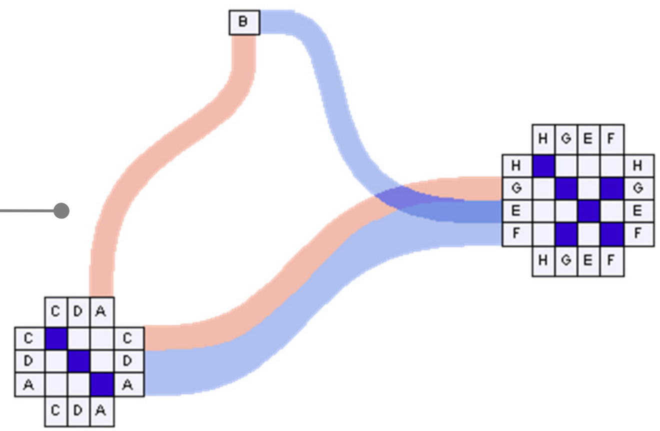
- Matrices to explore
- Node-Link diagrams to present findings

NodeTrix [Henry et al.07]

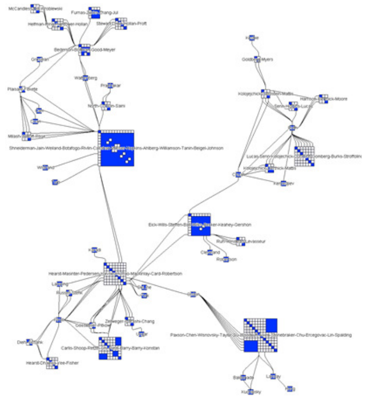
Hybrid representation

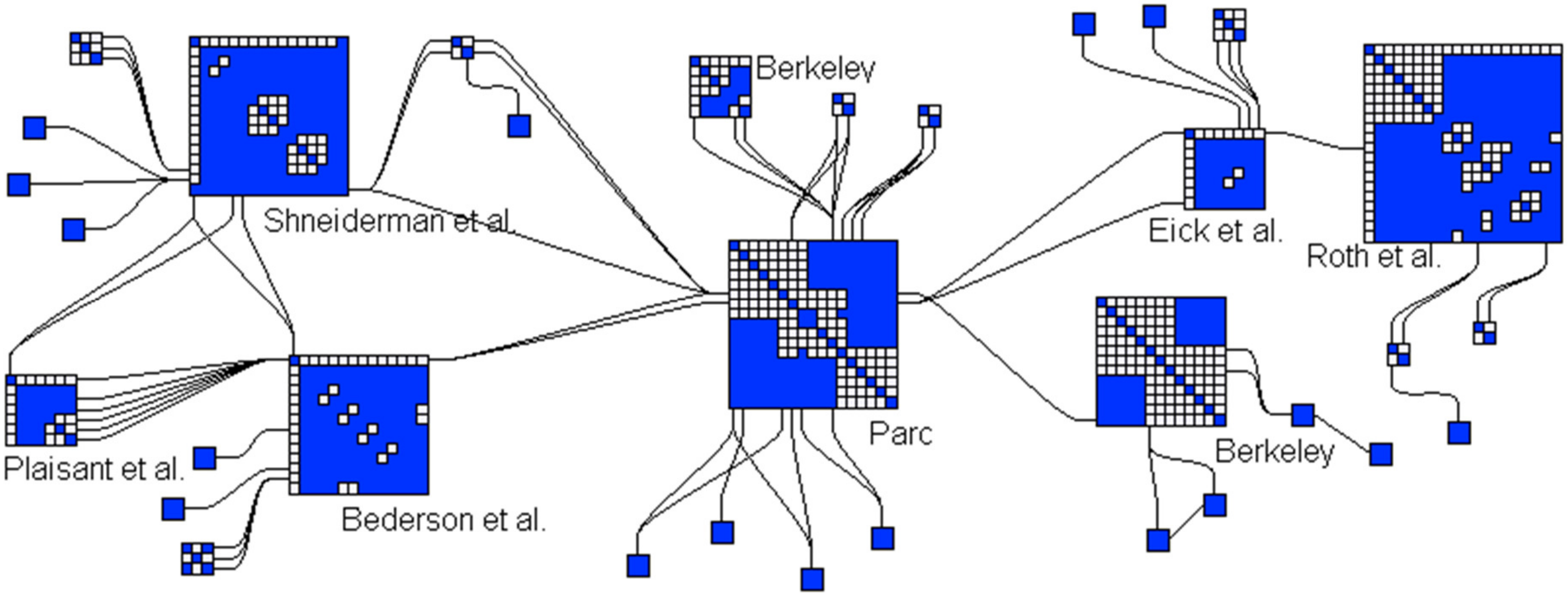


- Designed for small-world networks
  - Globally sparse
  - Locally dense
- Visualizing dense sub-graphs as matrices
- Interact to create, edit and remove the matrices

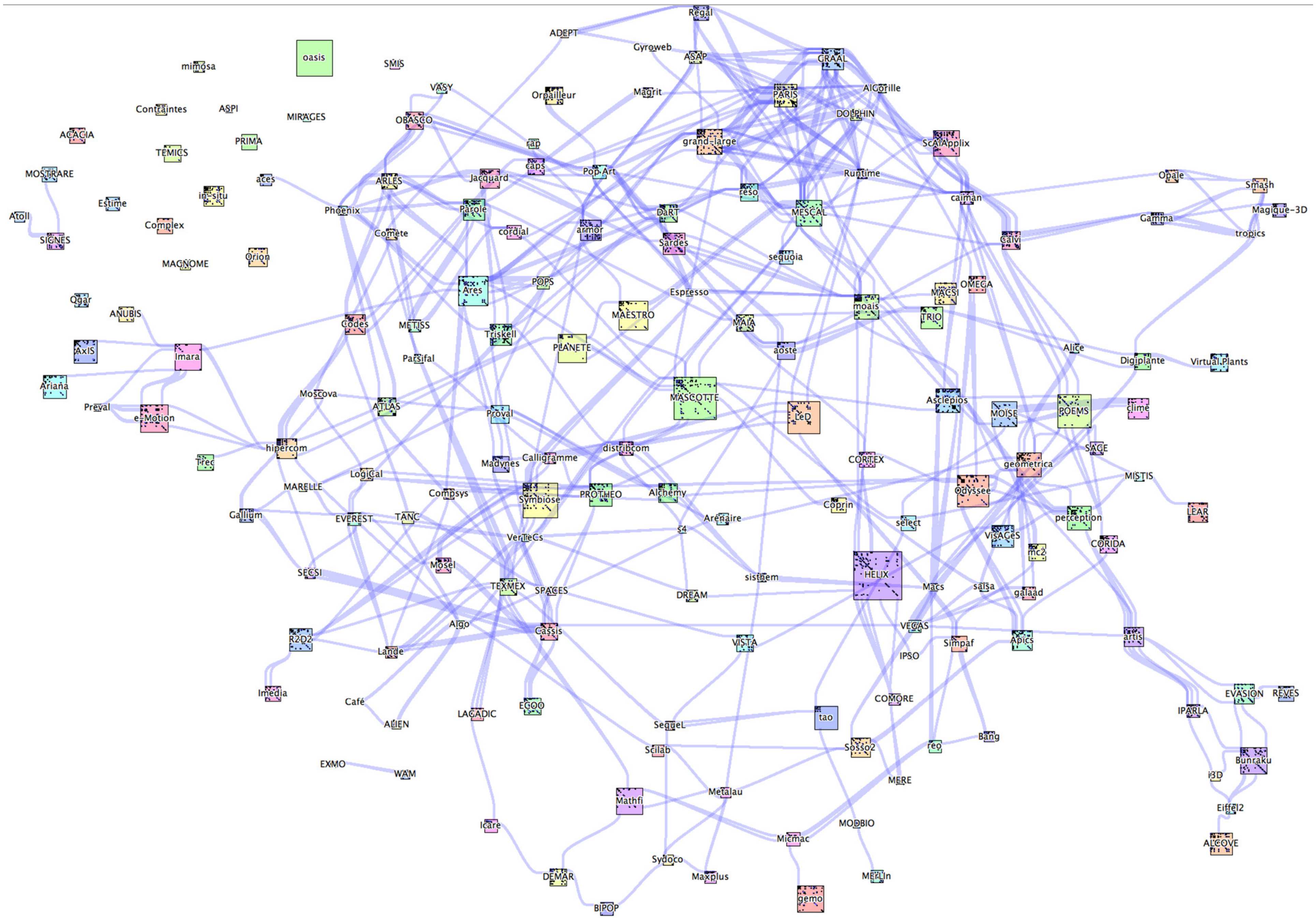


Visual Patterns

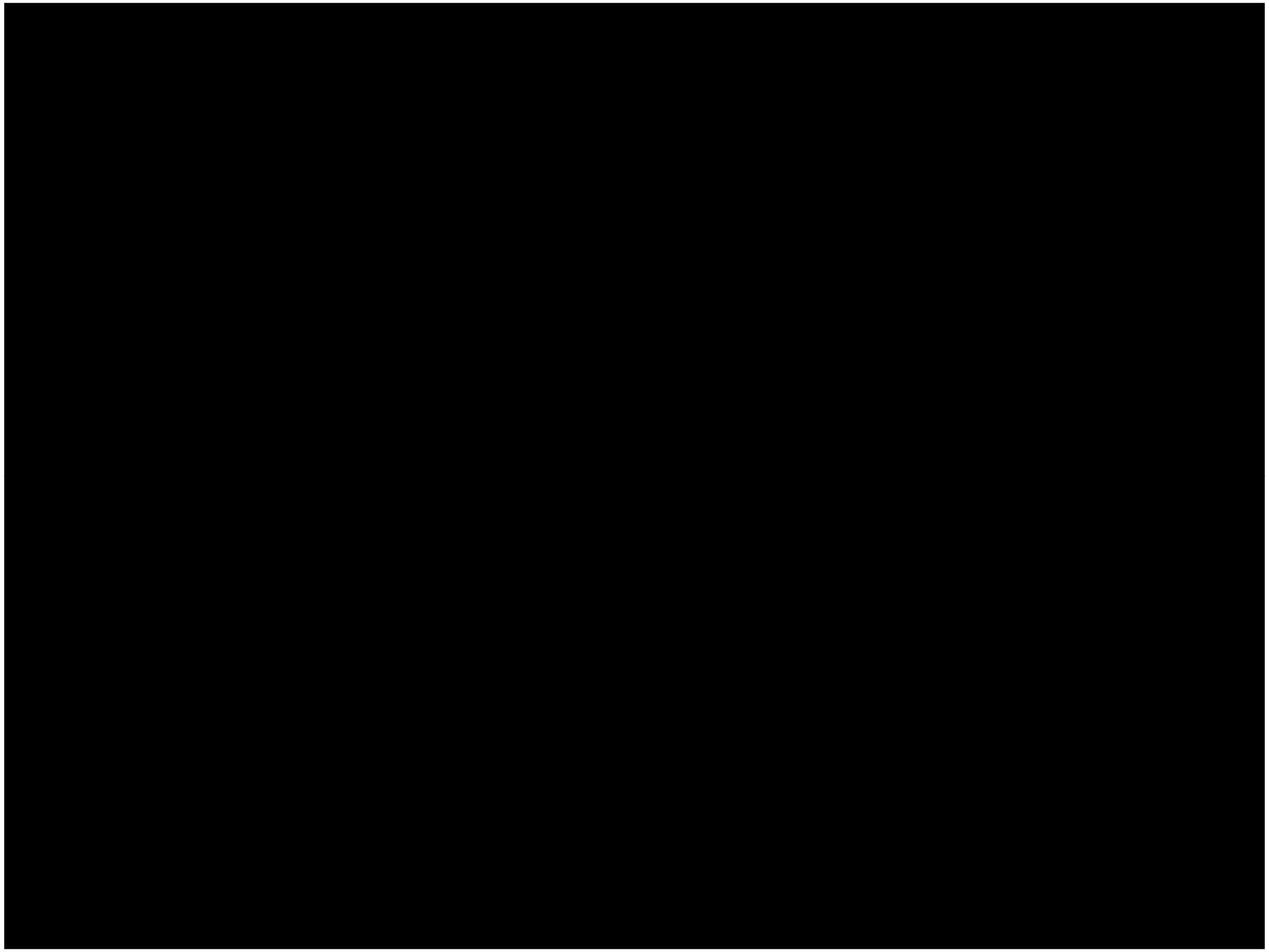




Infovis Coauthorship (133 actors)



Regal
ADEPT
Gyroweb
ASAP
GRAAL
oasis
mimosa
SMIS
VASY
Orpailleur
Magrit
PARIS
AlGorille
Contraintes
ASPI
MIRAGES
OBASCO
DOLPHIN
ACACIA
PRIMA
rap
grand-large
ScAlApplix
TEMICS
caps
Pop-Art
Runtime
Opale
Smash
MOSTRARE
aces
ARLES
Jacquard
reso
caiman
Magique-3D
in-situ
Phoenix
Parole
DaRT
MESCAL
Gamma
Atoll
Estime
Complex
Comète
cordial
armor
Sardes
Calvi
tropics
SIGNES
sequoia
MAGNOME
Orion
Espresso
OMEGA
Qgar
POPS
MACSI
Ares
moais
ANUBIS
MAESTRO
Codes
METISS
MAIA
TRIO
Triskell
PLANETE
aoste
Alice
AxIS
Imara
Digiplante
Virtual Plants
Parsifal
Ariana
MASCOTTE
Moscova
Asclepios
ATLAS
Preval
Proval
LeD
MOISE
POEMS
clime
e-Motion
hipercom
distribcom
Madynes
Calligramme
SAGE
CORTEX
Trec
geometrica
Logical
MARELLE
MISTIS
Odyssee
Compsys
Alchemy
PROTHEO
Symbiose
Coprin
Gallium
LEAR
EVEREST
TANC
Arenaire
perception
select
s4
VerTeCs
VISAGES
CORIDA
mc2
Mosel
HELIX
SECSI
sisthem
TEXMEX
SPACES
Macs
salsa
galaad
DREAM
Algo
VEGAS
Cassis
R2D2
VISTA
artis
Simpaf
Apics
Lande
IPSO
Imedia
EVASION
REVES
COMORE
Café
IPARLA
LAGADIC
tao
EGOO
ALIEN
SequeL
Bang
Scilab
Sosso2
reo
Bunraku
EXMO
WAM
i3D
Metalau
MERE
Mathfi
Eiffel2
Icare
MODBIO
Micmac
ALCOVE
Sydoco
DEMAR
Maxplus
MErLIn
gemo
BIPOP



Video on NodeTrix
http://www.lri.fr/~nhenry/nodetrix/NodeTrix_h264.mov

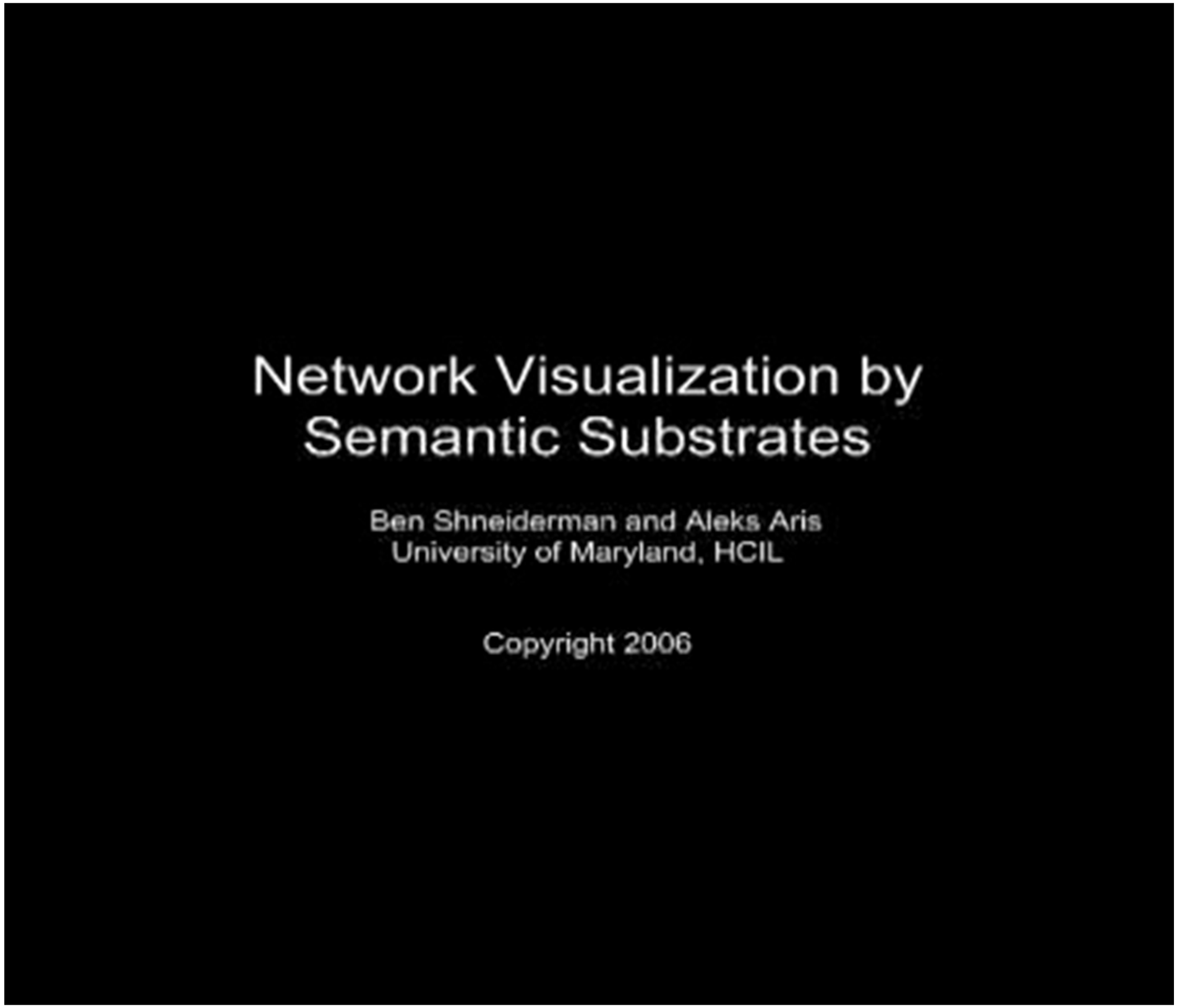
Force-Directed Layout

- Edges = springs
- Nodes = charged particles

- Repeatedly calculate forces, update node positions

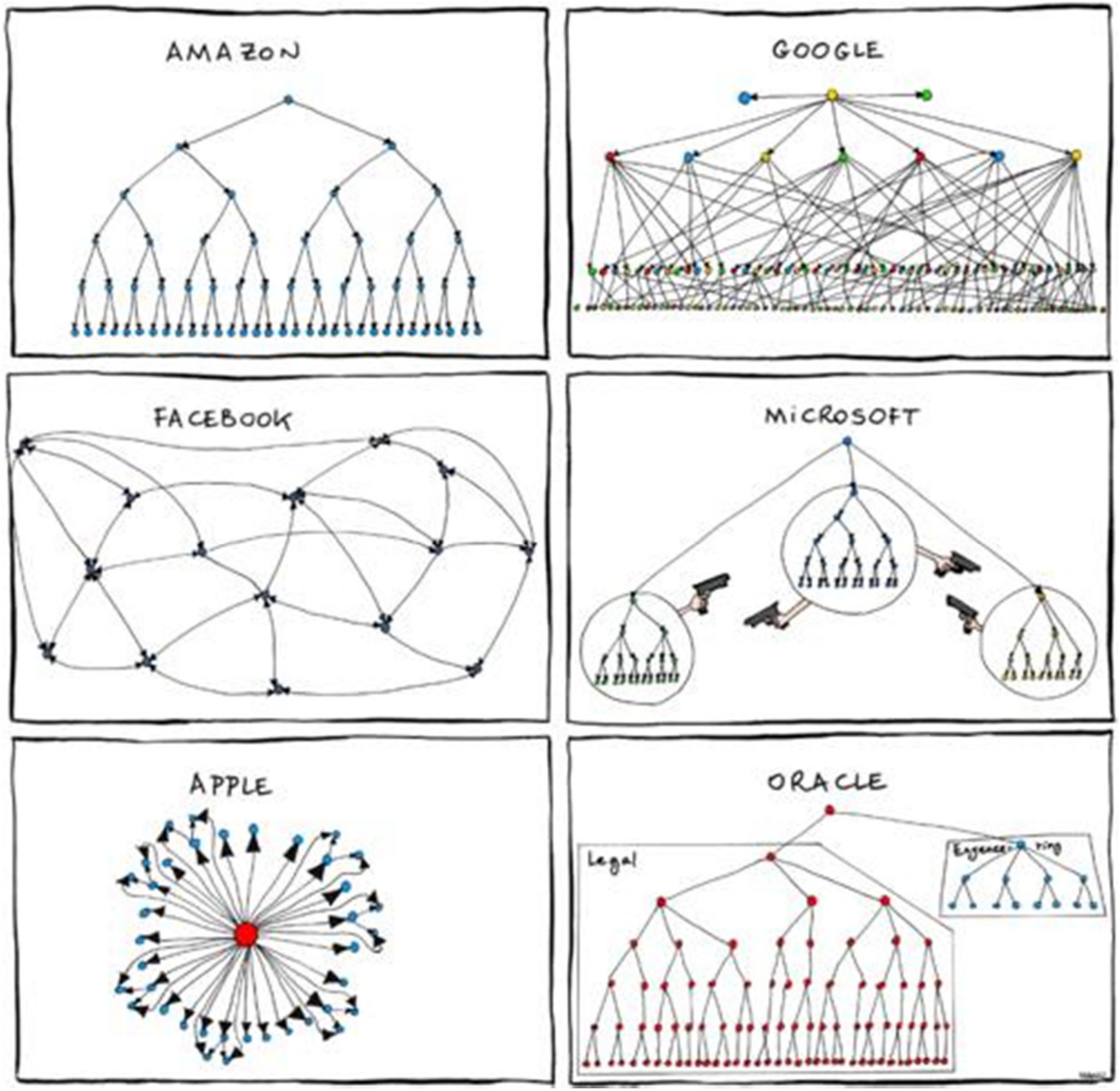
http://mbostock.github.com/d3/ex/force.html

Attribute-Driven Layout



http://www.cs.umd.edu/hcil/nvss/

And finally...

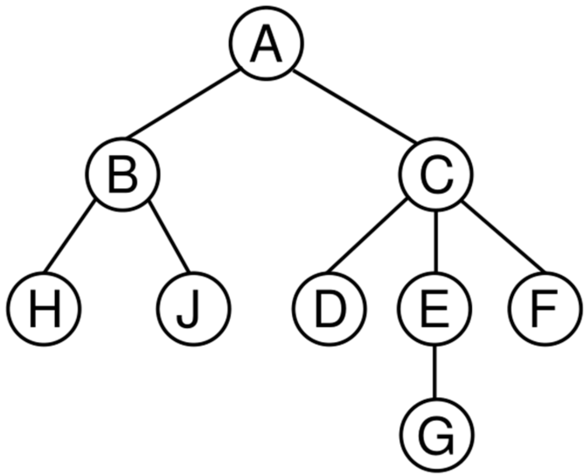


Visualization Techniques for

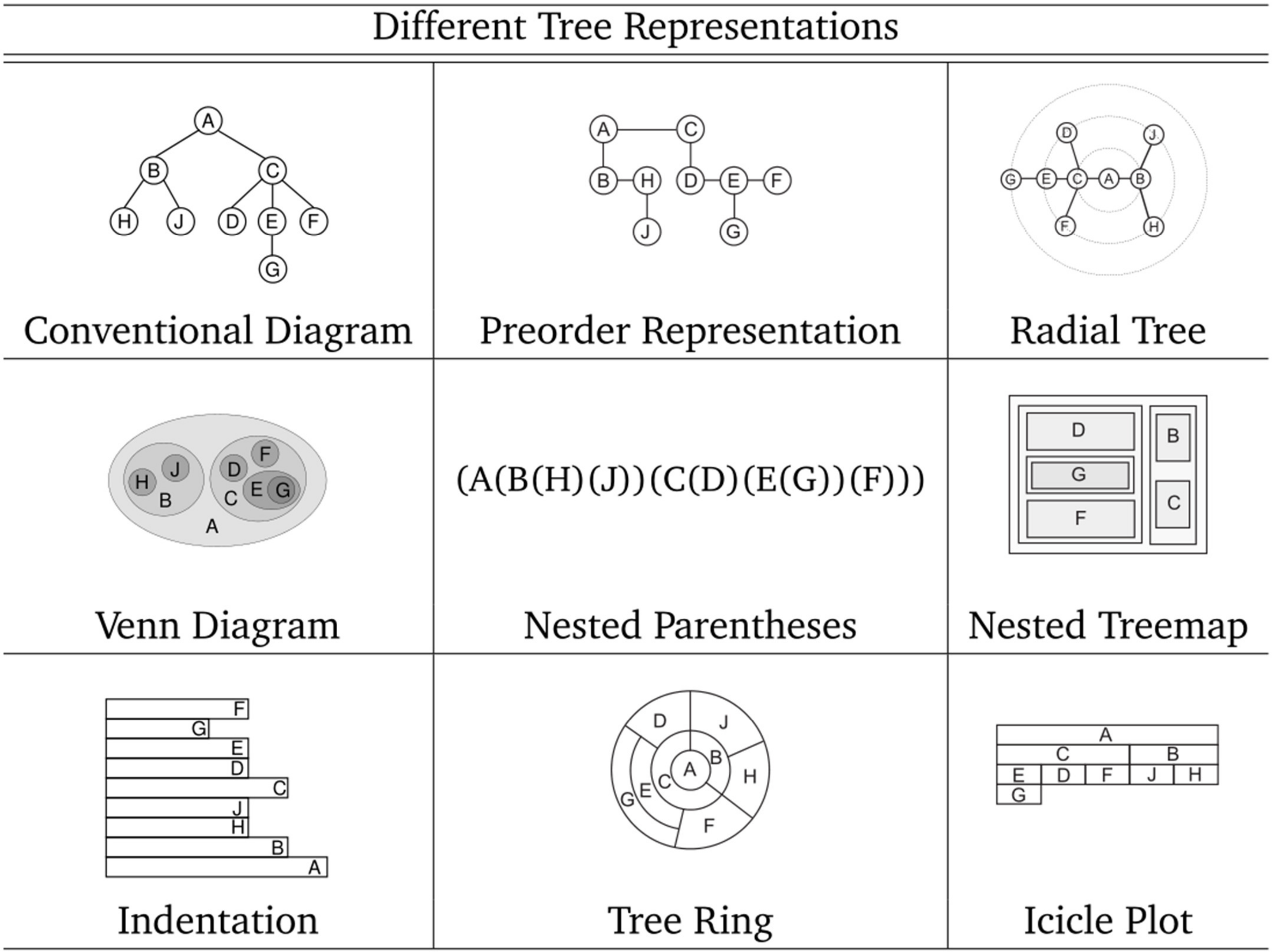
HIERARCHICAL DATA

Node-Link Layout

- The traditional approach
- Pick a layout algorithm
- Draw a rooted tree
- Well known layout algorithm: Reingold-Tilford



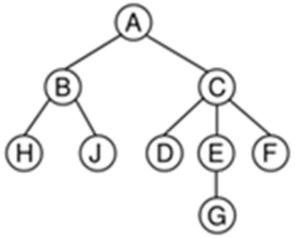
- Many more exist which try to balance a number of aesthetic criteria


Different Tree Representations
Conventional Diagram
Preorder Representation
Radial Tree
Venn Diagram
(A(B(H)(J))(C(D)(E(G))(F)))
Nested Parentheses
Nested Treemap
Indentation
Tree Ring
Icicle Plot

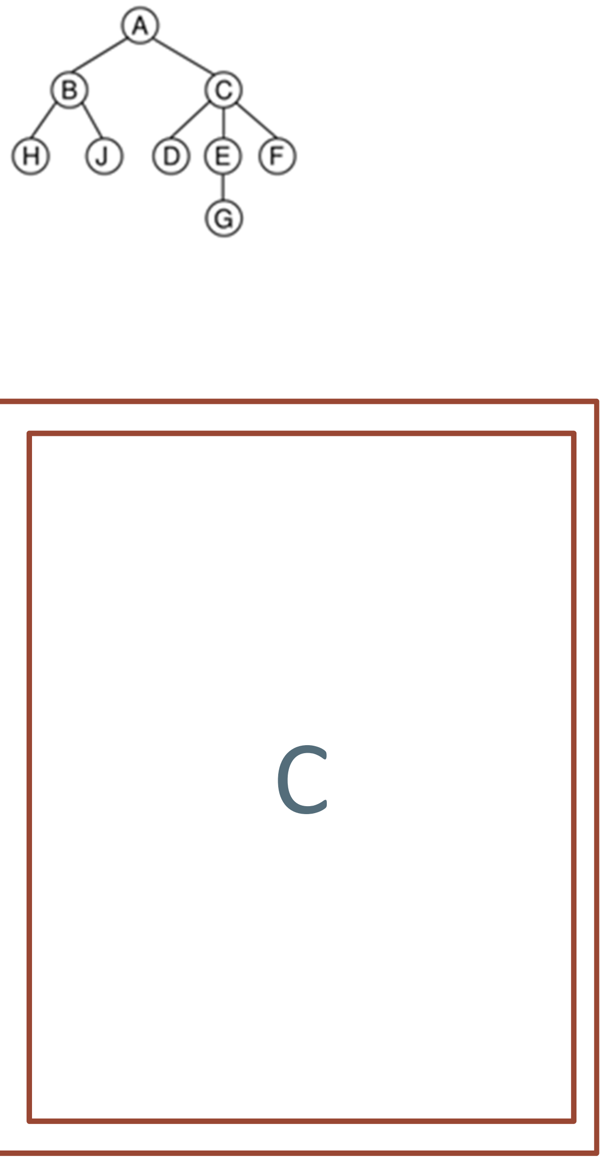
TreeMap

- Space-filling
- Free tool here:
  http://www.cs.umd.edu/hcil/treemap/

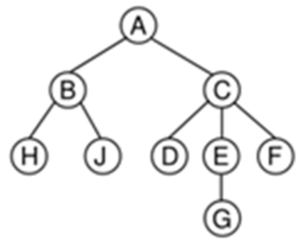
TreeMap

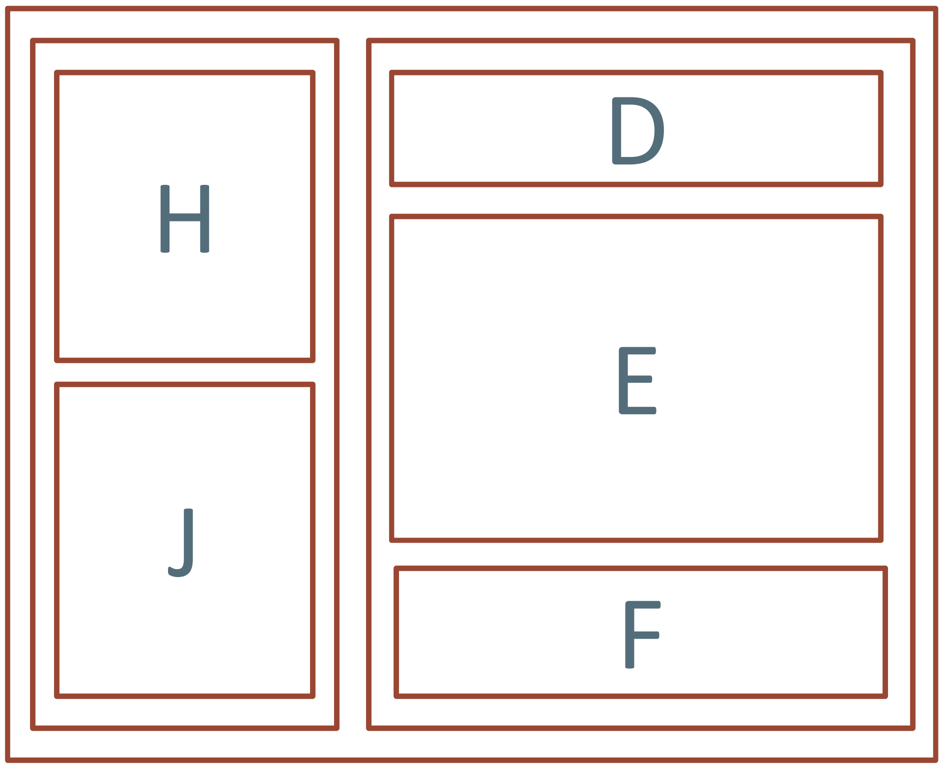


TreeMap

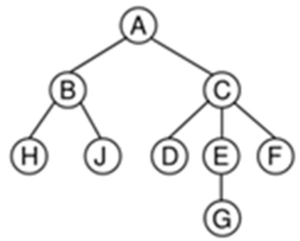


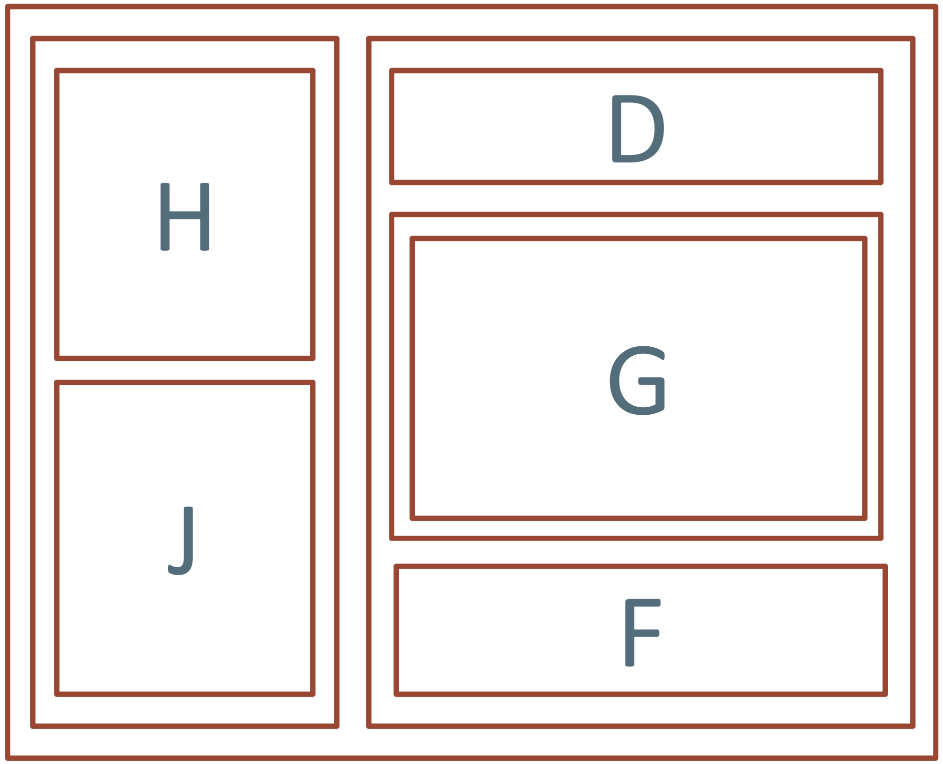
TreeMap



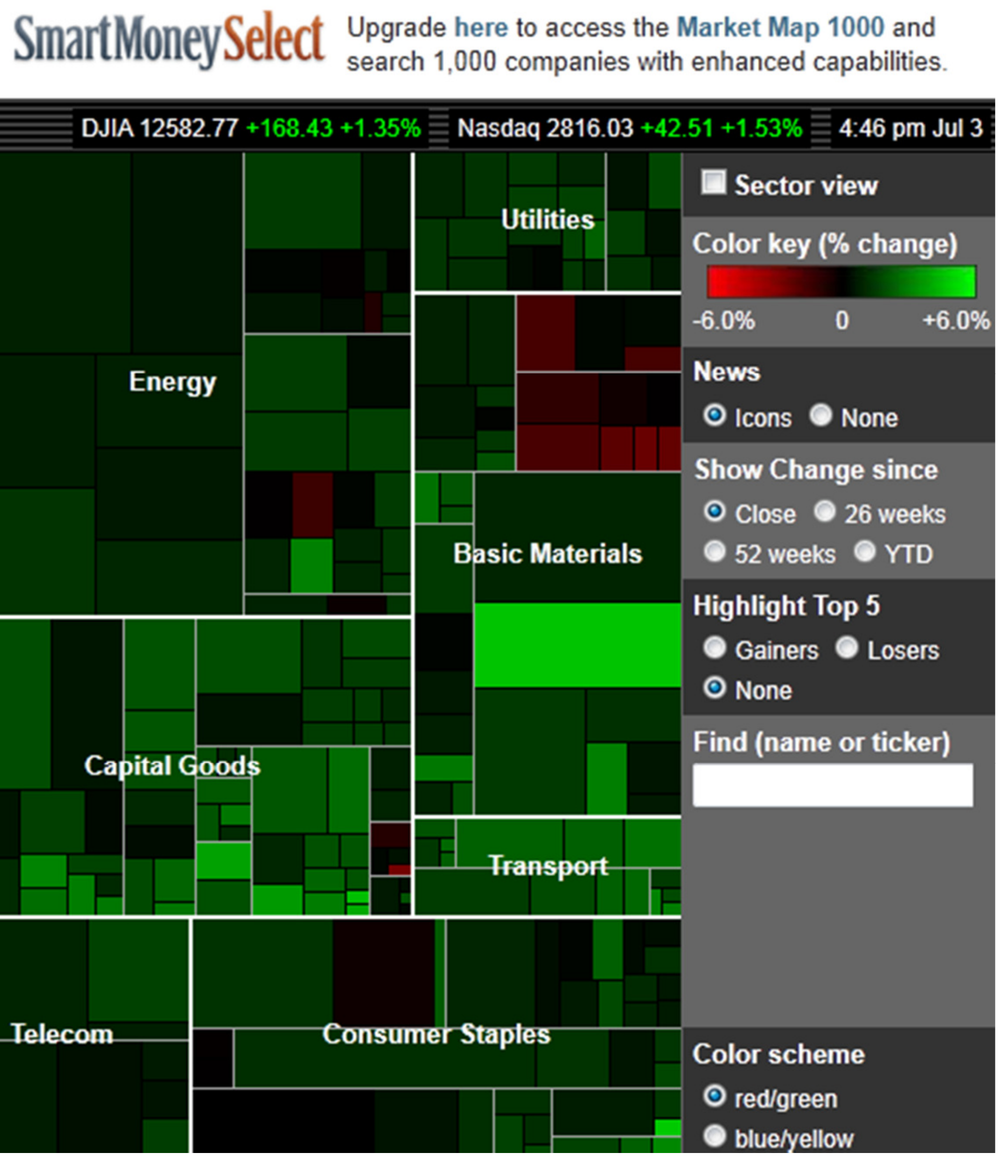


TreeMap

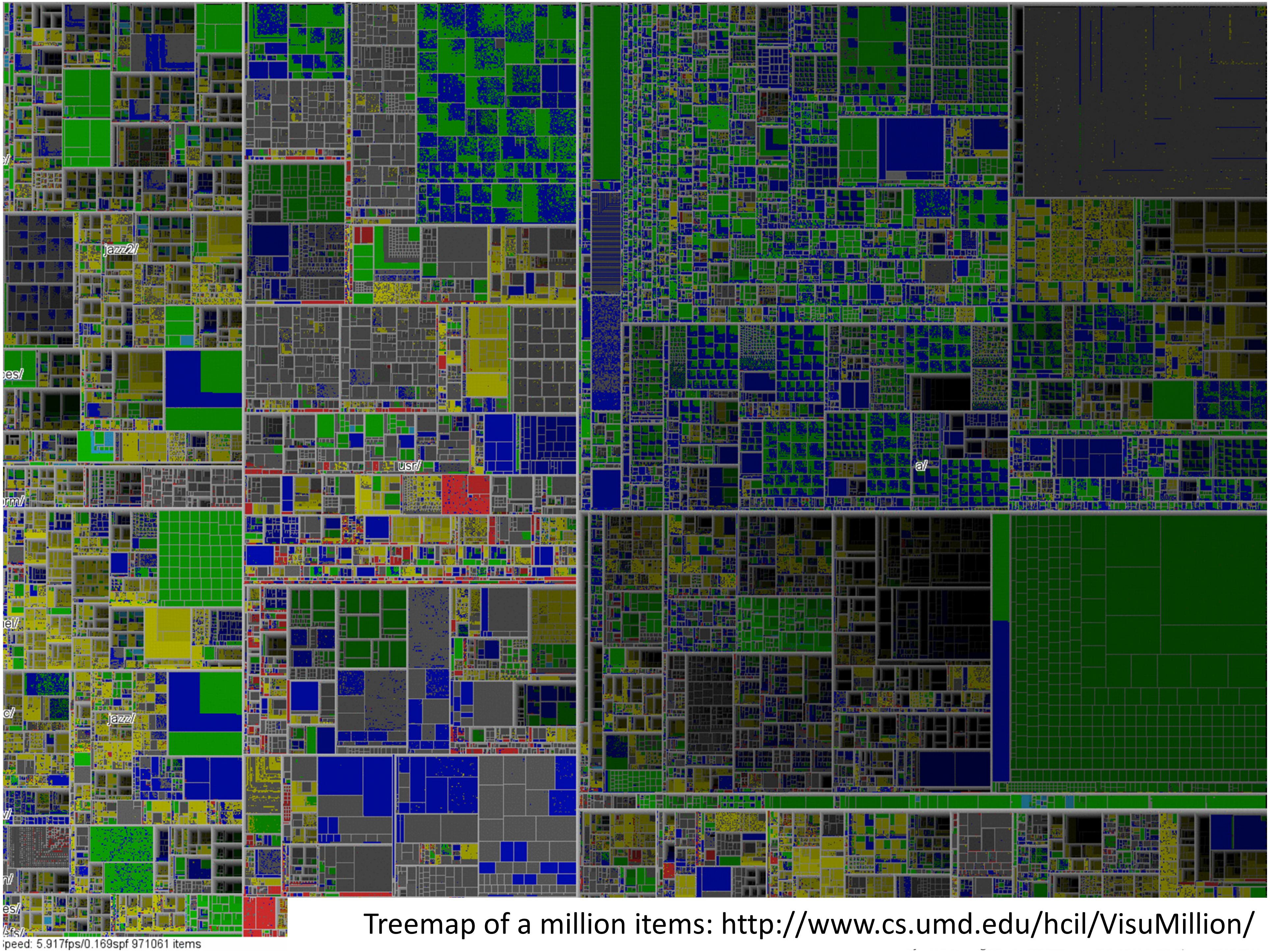




Map of the Market

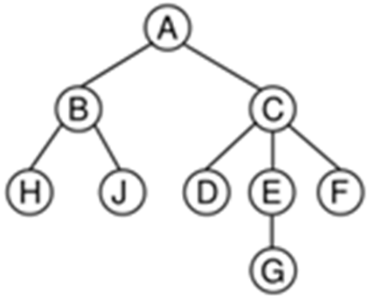


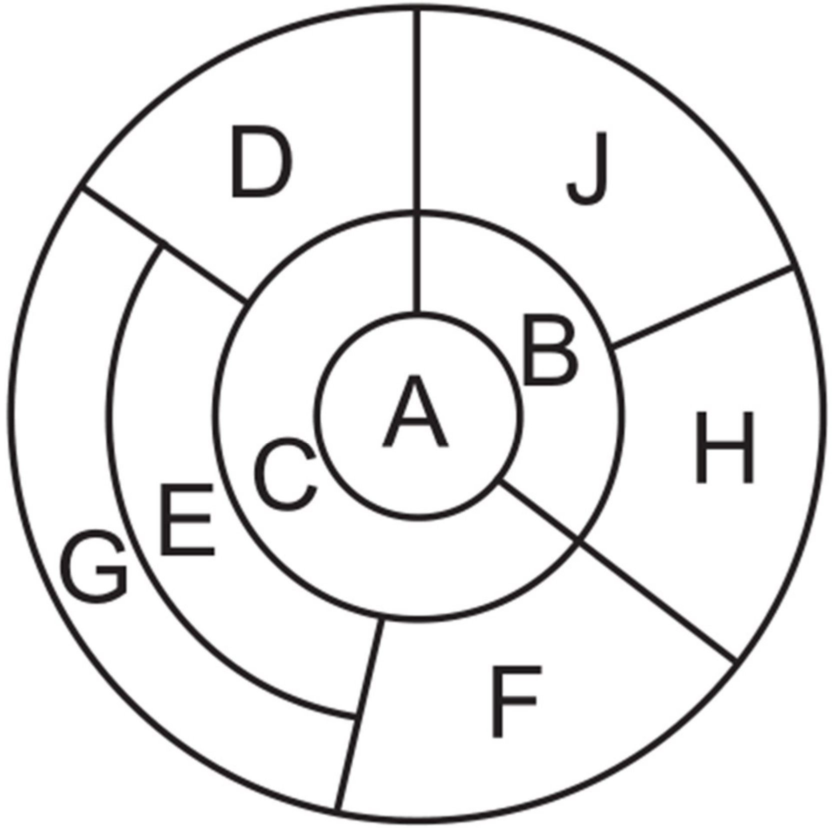
http://www.smartmoney.com/map-of-the-market/



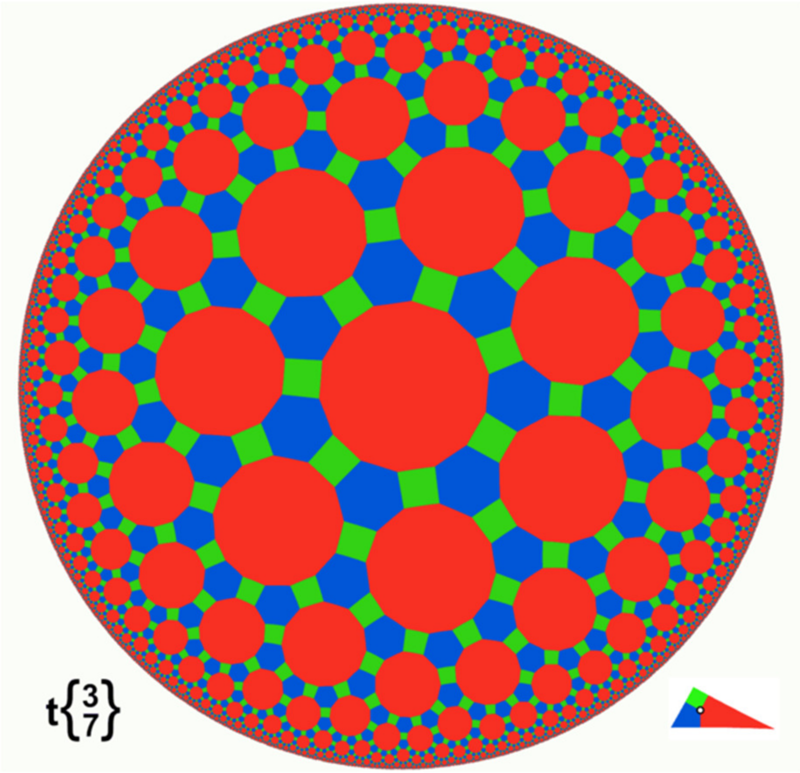
Treemap of a million items: http://www.cs.umd.edu/hcil/VisuMillion/

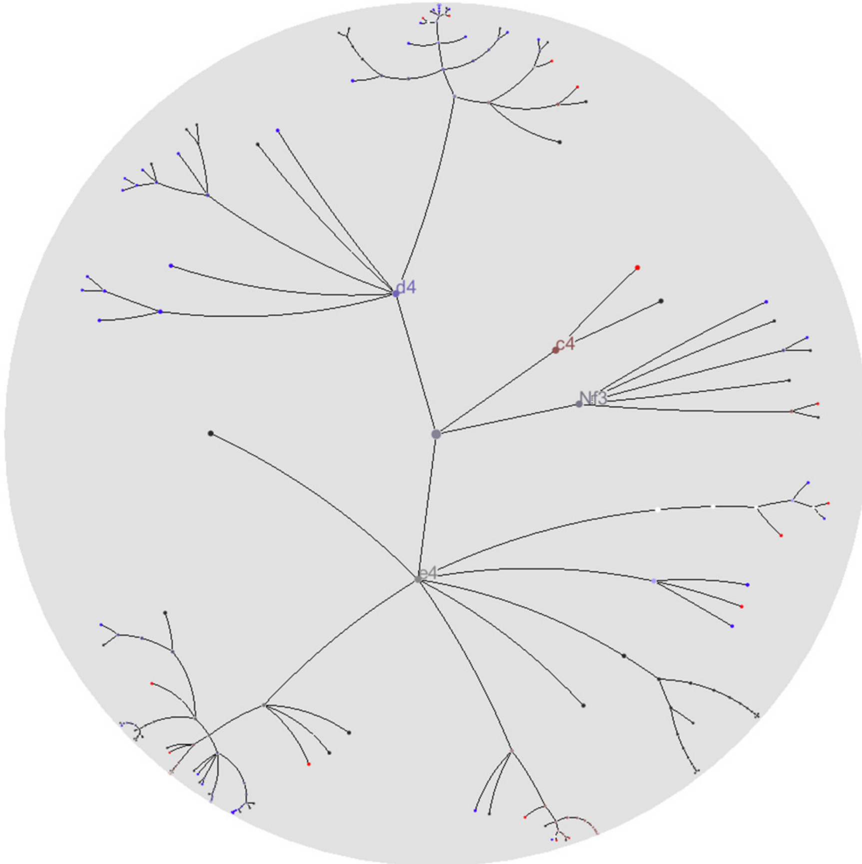
TreeRing





Hyperbolic Tree





Poincaré Disk -> Lay out a tree on this space
Advantage: Tree will definitely fit in view! (clever!) But nodes quite squished near edge

Geneaquilts

GeneaQuilts

A System for Exploring
Large Genealogies

A.Bezerianos P.Dragicevic J.-D.Fekete J.Bae B.Watson

http://www.aviz.fr/geneaquilts/

Last but not Least!

INTERACTION

Interaction Techniques

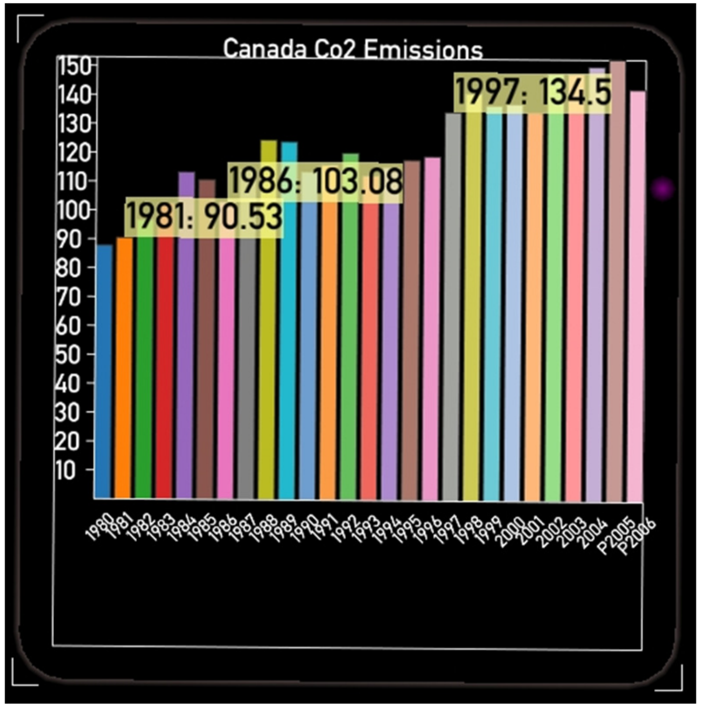
Based on user intent

- Select – mark something as interesting
- Explore – show me something else
- Reconfigure – show me a different arrangement
- Encode – show me a different representation
- Abstract/Elaborate – more or less detail
- Filter – show me something conditionally
- Connect – show me related items

After [Yi et al., InfoVis 2007]

Selection

- Mark something as interesting
- Often combined with other techniques



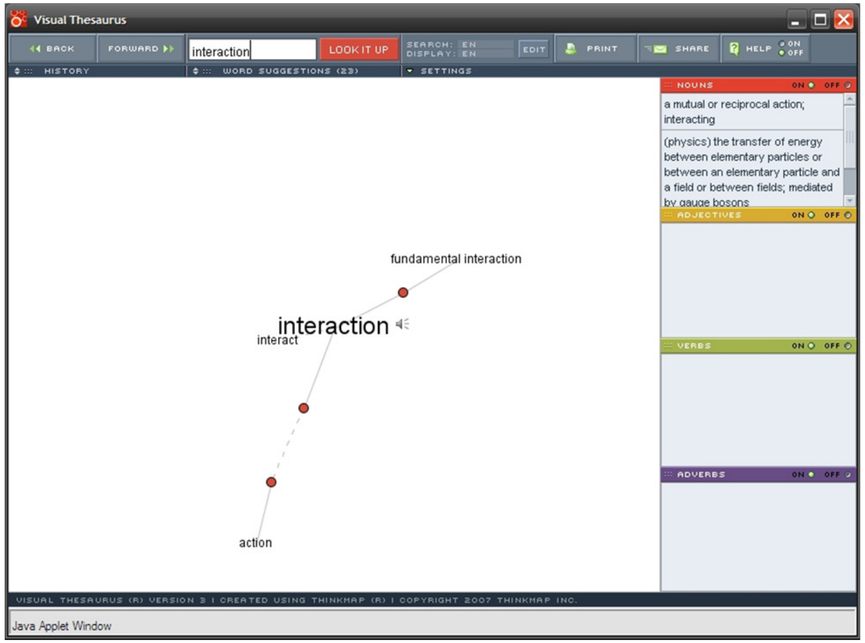
Selection

http://www.cs.ubc.ca/~tmm/papers/tj/

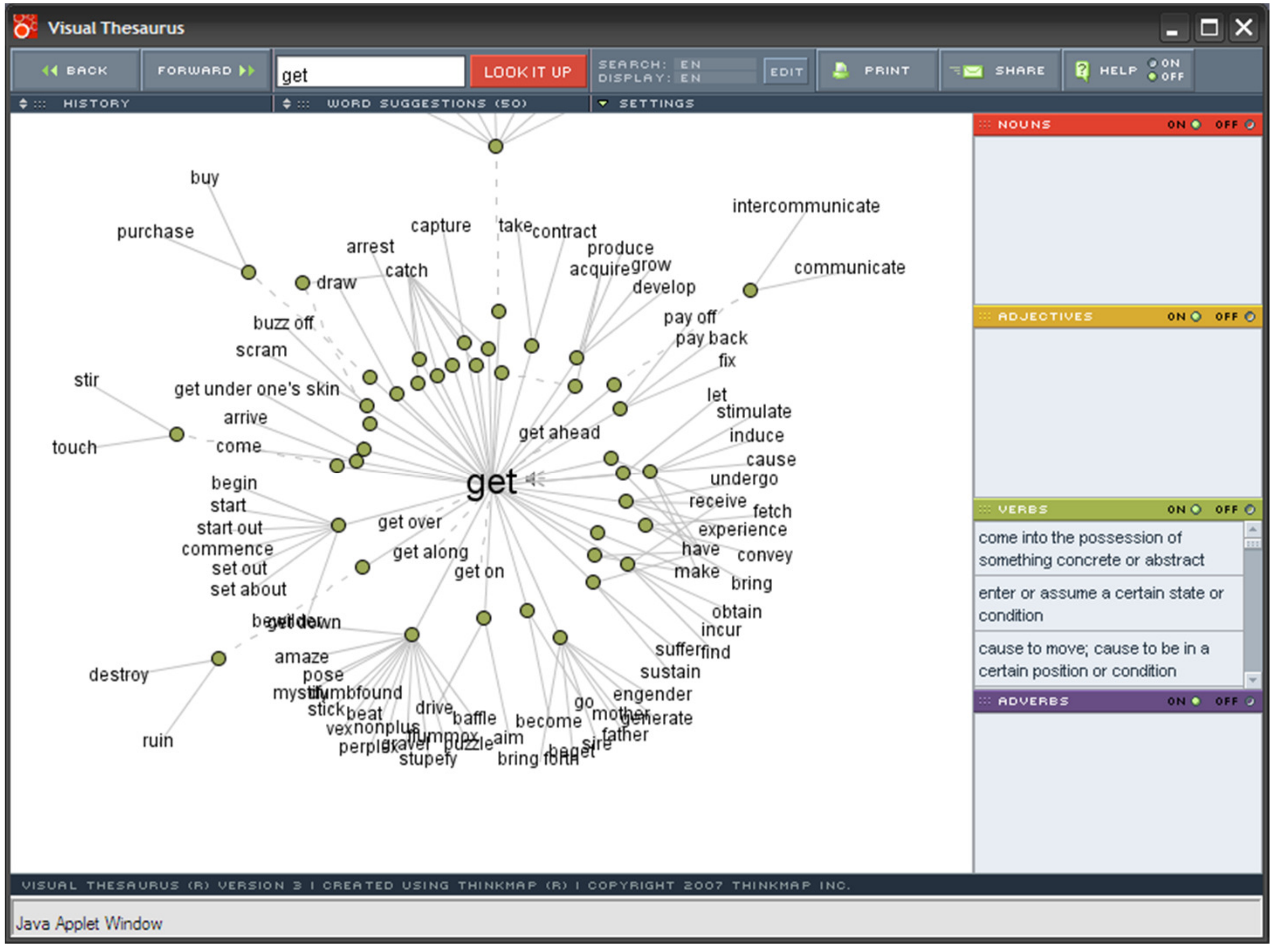
[Munzner et al, 2003]

Explore

- Show me something else
- Examine subset of data cases (view-based)
  - E.g. Panning (move viewpoint across representation)
  - E.g. Direct Walk (move viewing focus through clicks)



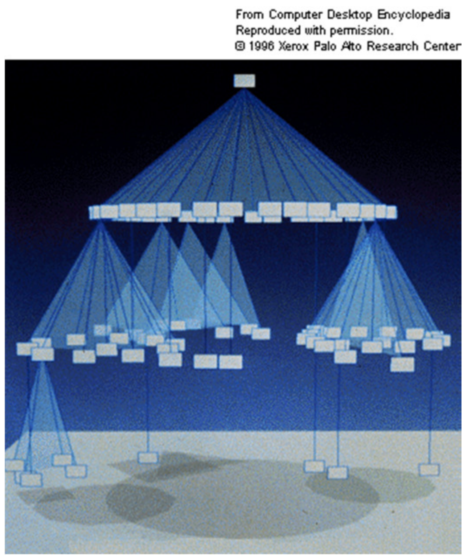
Explore

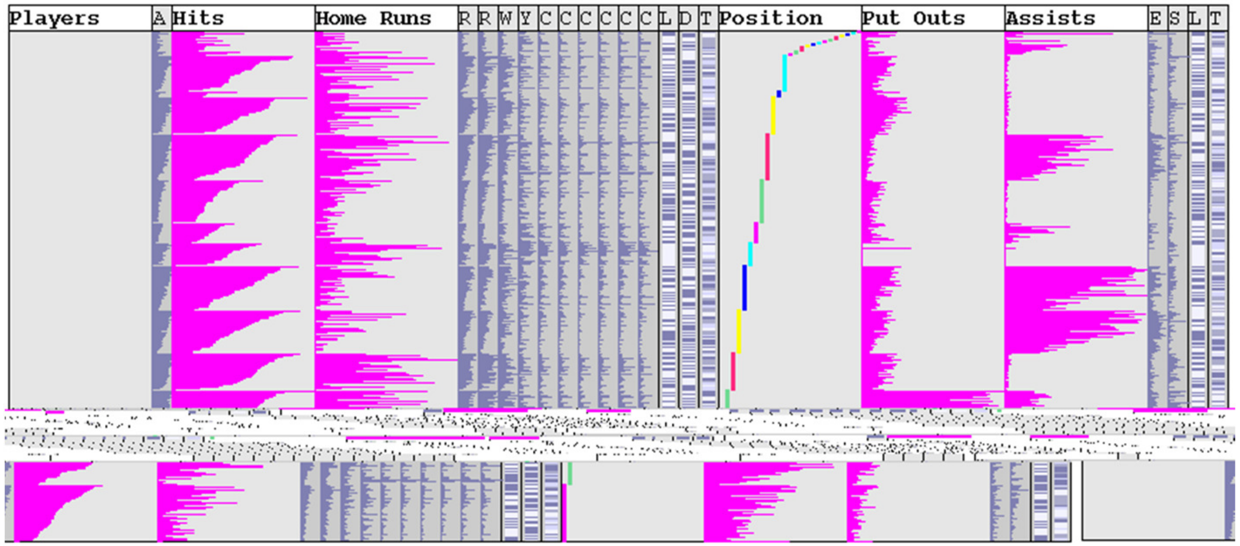


VisualThesaurus.avi

Reconfigure

- Show a different arrangement
  - Move data items to
    - Enable better comparison
    - Avoid occlusion
    - Correspond to some mental model of the data

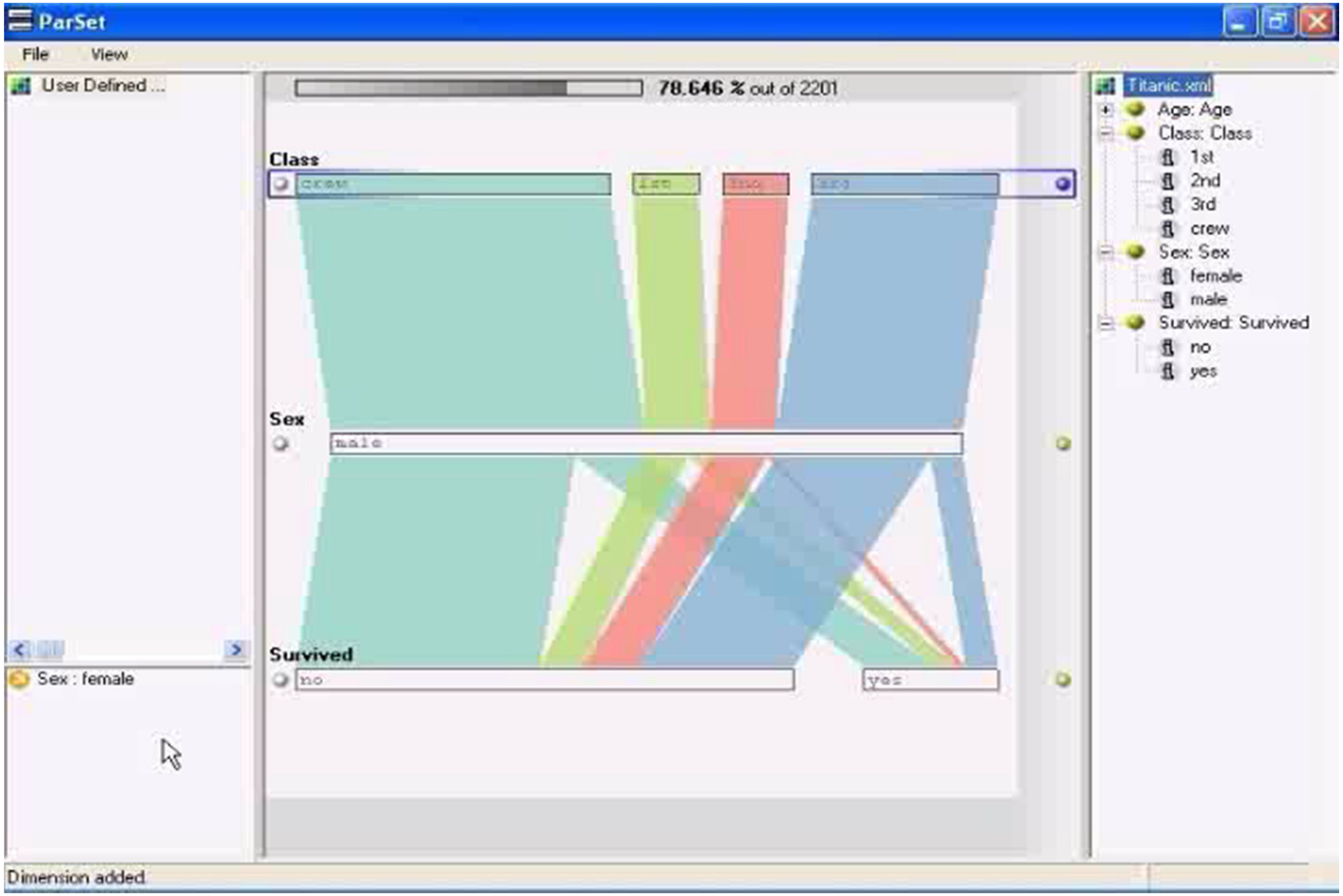




Cone Trees

Table Lens

Reconfigure



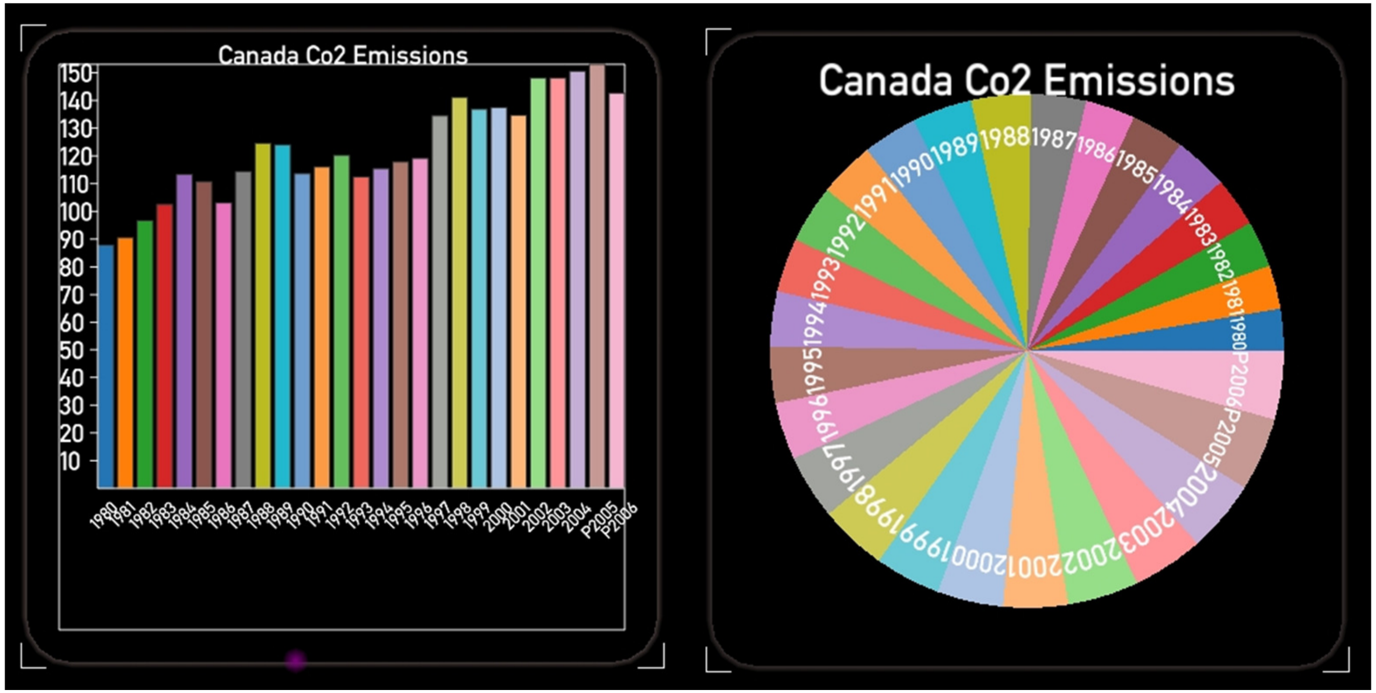
[ParallelSets - Bendix et al., 2005]

Reconfigure

TreeJuxtaposer Video

Encode

- Show a different:
  - Representation Type
  - Visual appearance: Colour, Size, Shape,…



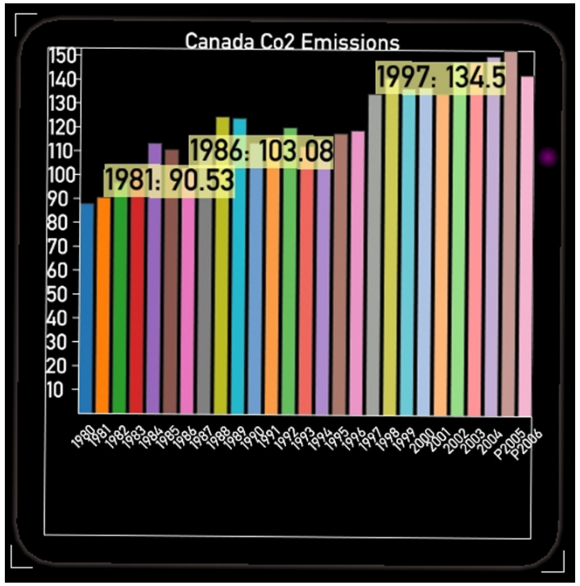
Encode

- Animation is important!

http://vis.berkeley.edu/papers/animated_transitions/ [DynaVis - Heer & Robertson, 2007]

Abstract/Elaborate

- Show me more or less detail
  - Adjust level of abstraction
  - Detail-on-demand
  - Zooming (as long as representation isn't fundamentally altered)

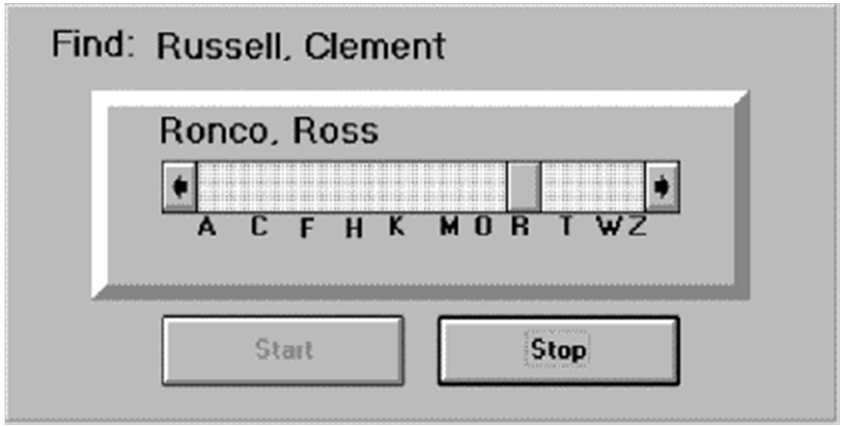


Warning: Not every technique belongs to just one category.

Abstract/Elaborate

Filter

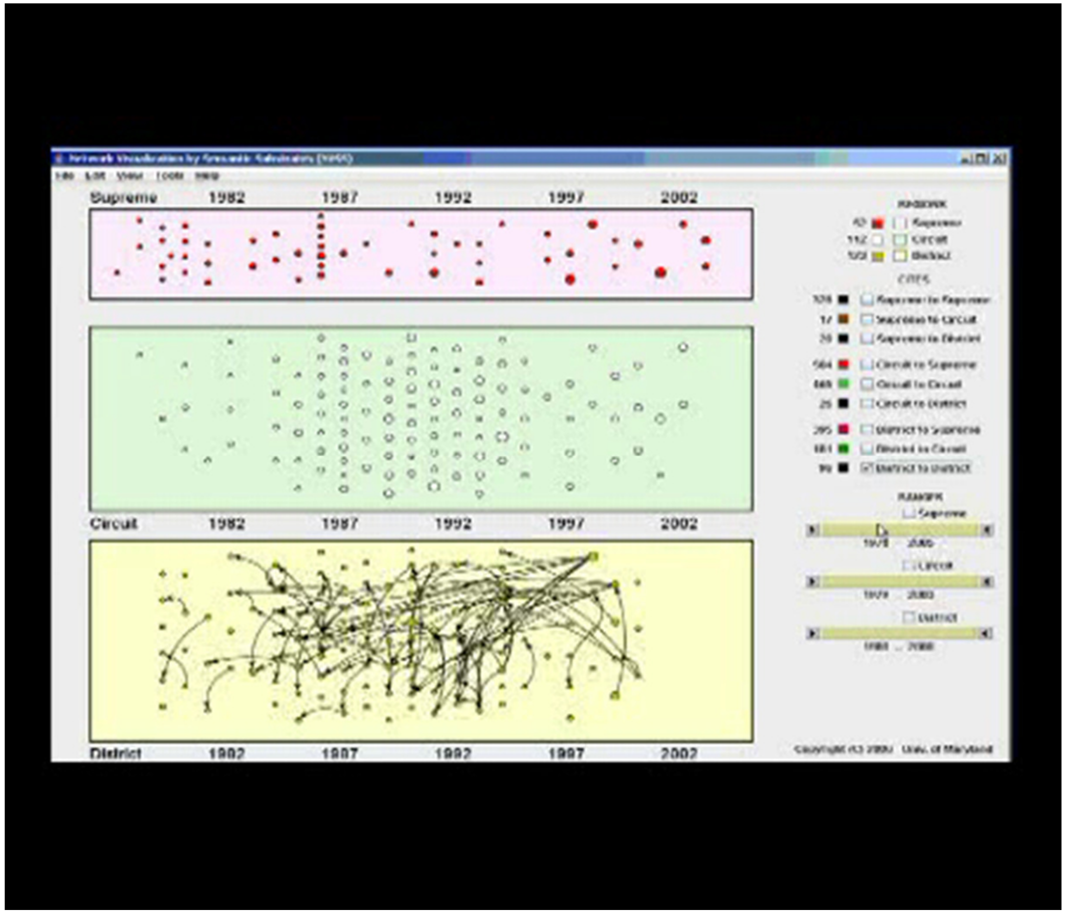
- Show subset of data based on condition
  - E.g. by selecting a data range



[AlphaSlider; Ahlberg & Shneiderman, 1993]

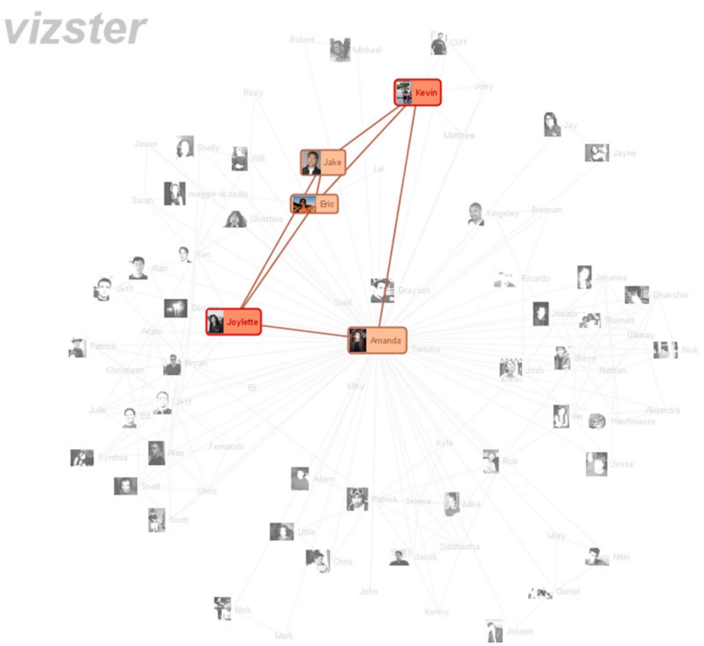
Filter

Remember the Range Sliders in this video



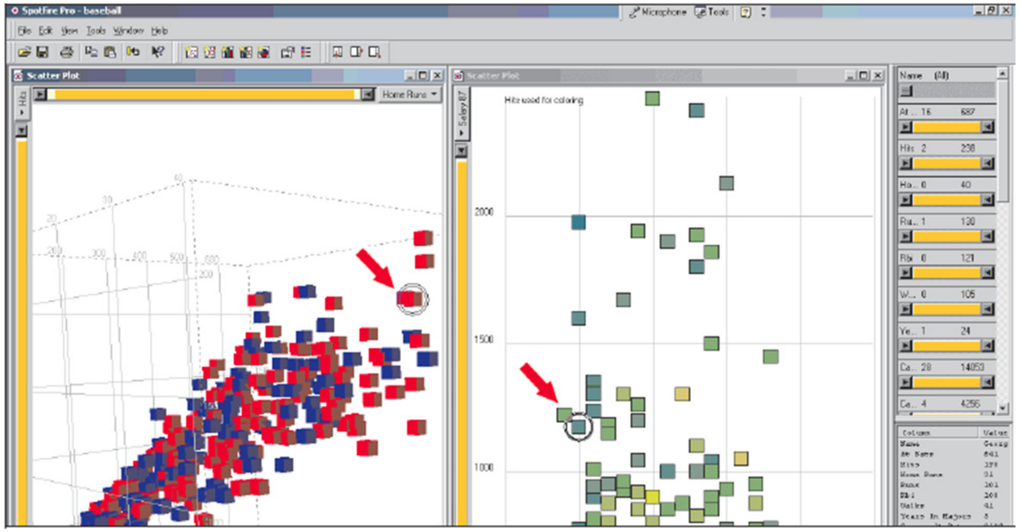
Connect

- Show related items
  - E.g. brushing



Single view

[Heer&boyd, InfoVis '05]

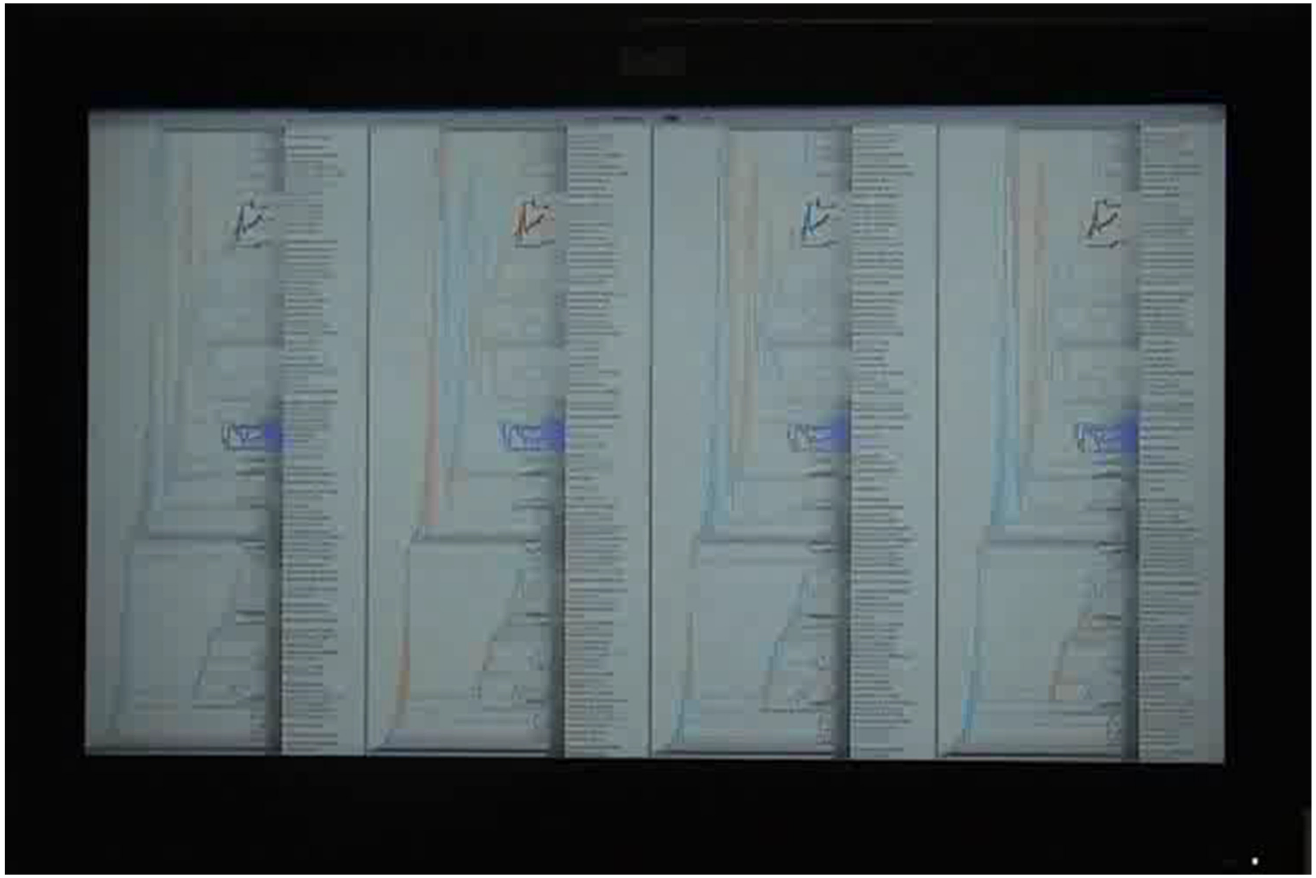


Multiple view

[Spotfire]

Connect

Which video previously showed a connect interaction?



Exercise & Discussion

Magic Lenses

Exercise & Discussion

Select

Explore

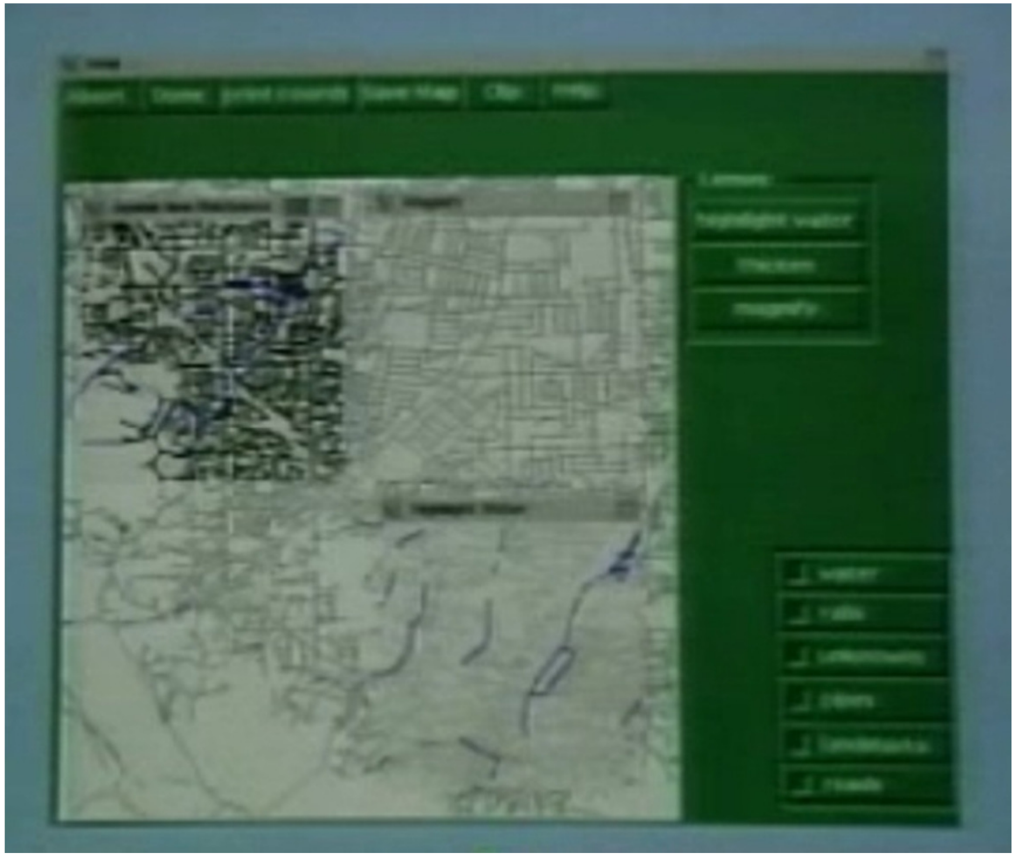
Reconfigure

Encode

Abstract/Elaborate

Filter

Connect

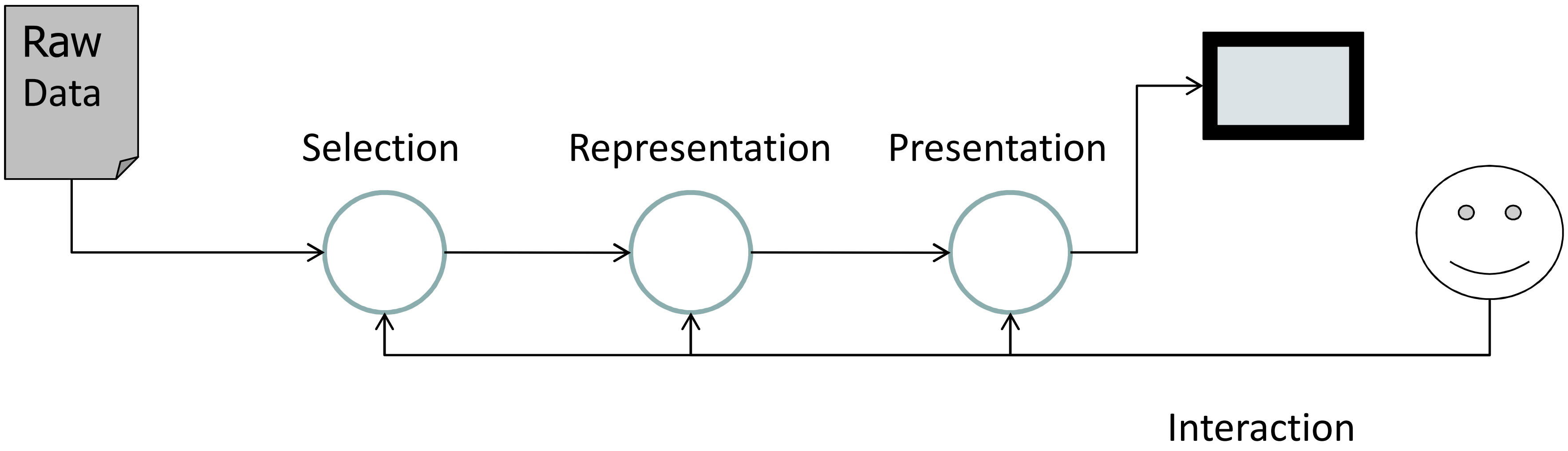


Toolglass & Magic Lenses

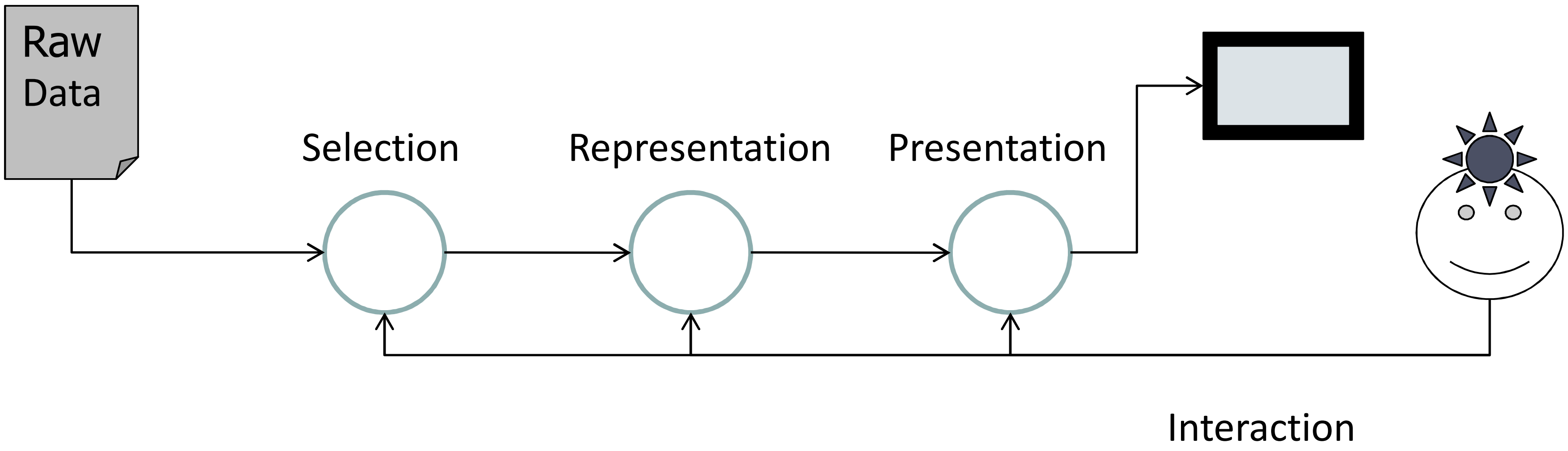
Part II

When & Where to interact

Spence's Infovis Model



Spence's Infovis Model

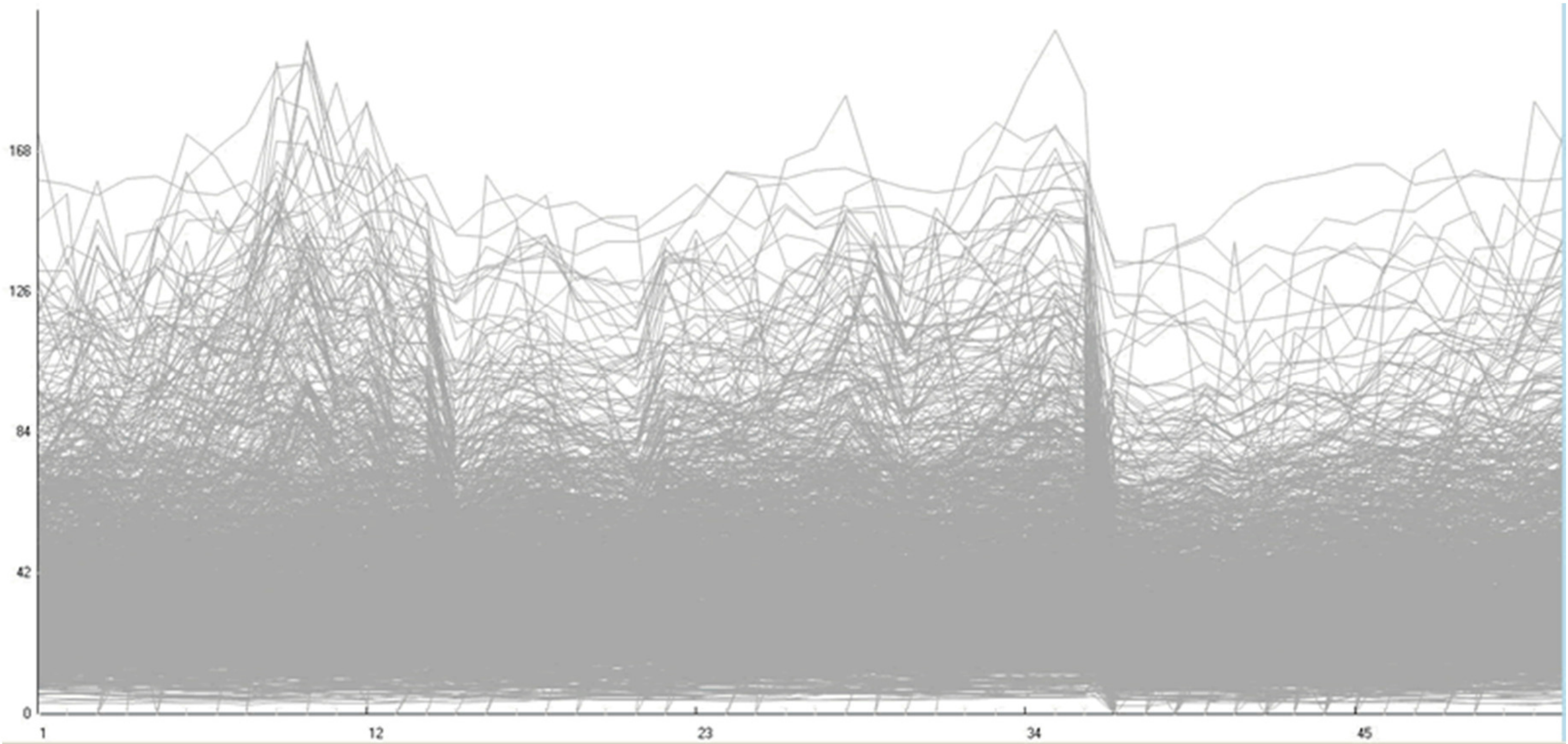


Focus on the "data" that is interacted with

Visual Information Seeking Mantra

[Shneiderman, 1996]

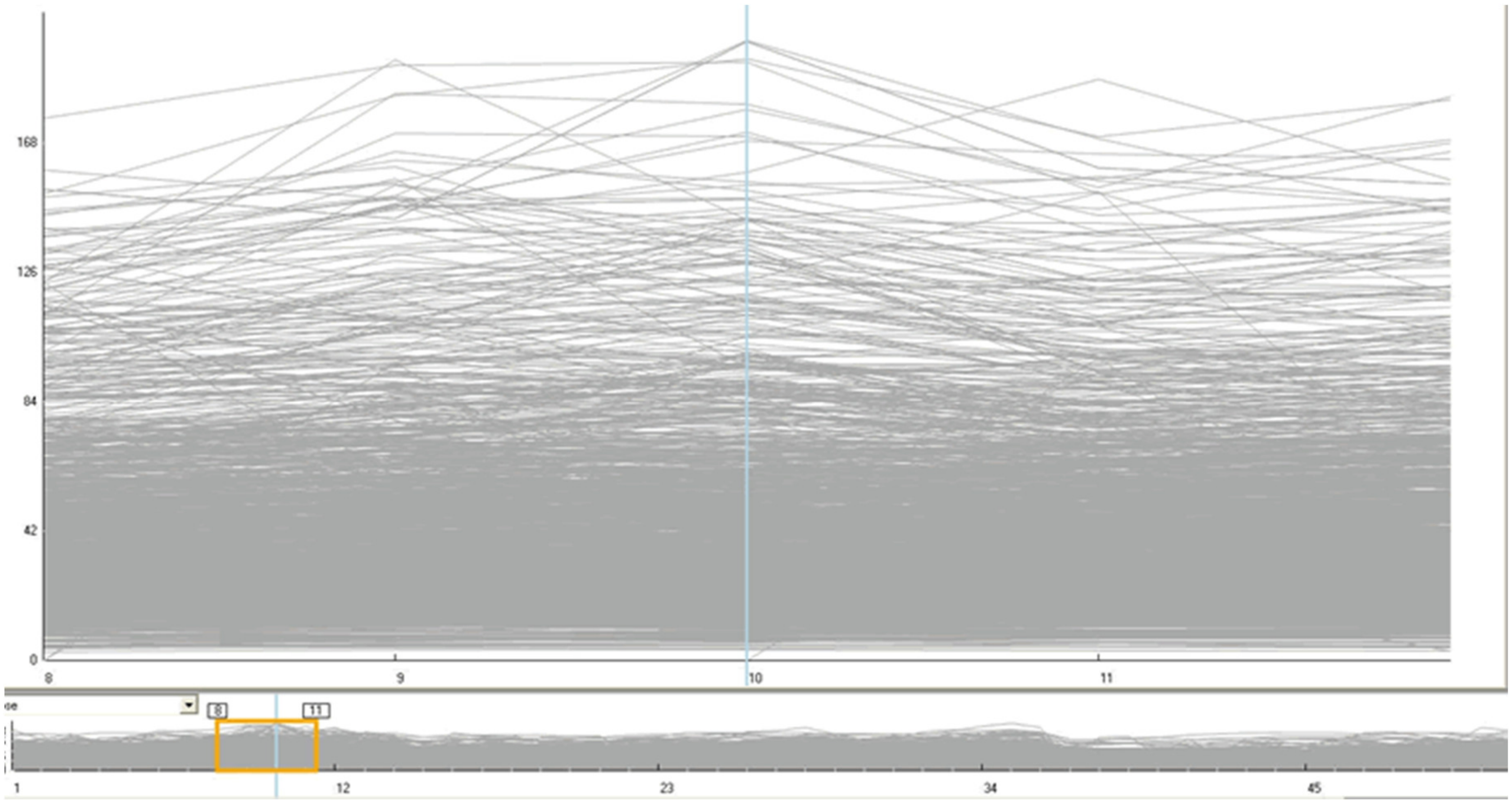
Describes the order of interaction operations



Visual Information Seeking Mantra

[Shneiderman, 1996]

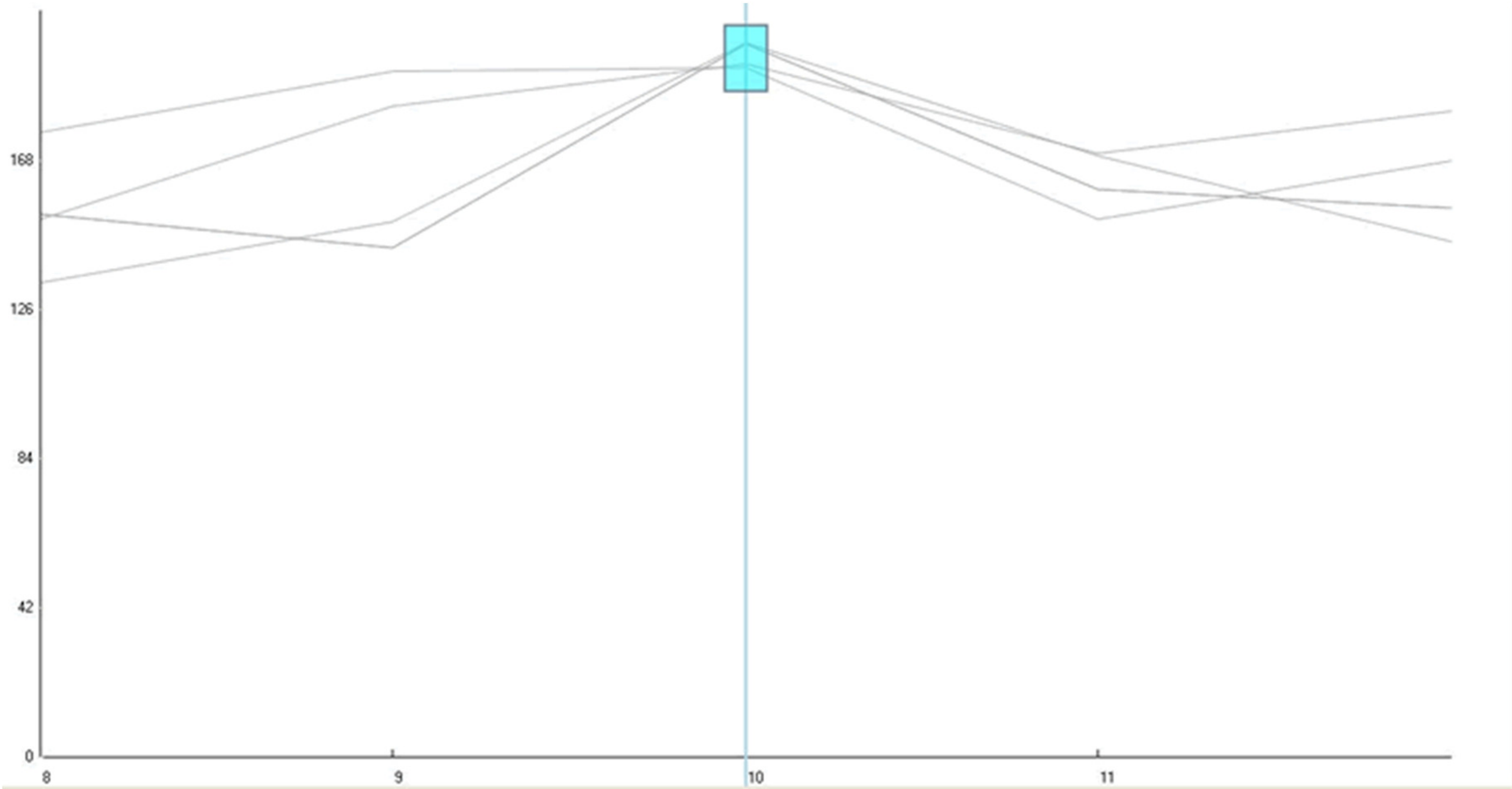
Describes the order of interaction operations



Visual Information Seeking Mantra

[Shneiderman, 1996]

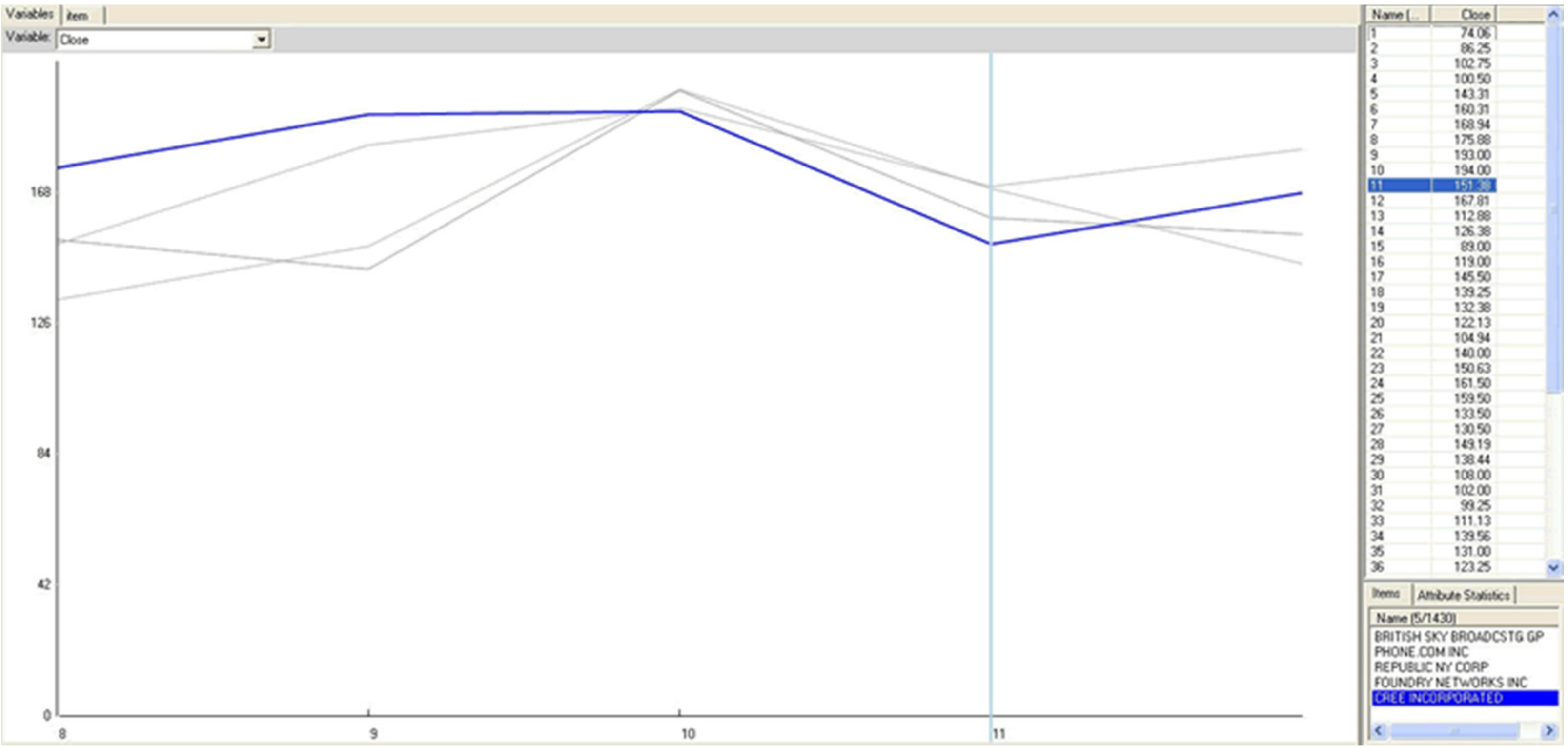
Describes the order of interaction operations



Visual Information Seeking Mantra

[Shneiderman, 1996]

Describes the order of interaction operations



Visual Information Seeking Mantra

Describes the order of interaction operations

- Overview first
- Zoom & filter
- Details on demand

→ useful for many (but not all) infovis applications

Knowledge Precepts

Problem: How to adequately support decision making?

→ More than simple queries needed!

Learning a domain

World-view Gap → What is shown != what needs to be shown

Identify the nature of trends

Predict the future

Rationale Gap → Deliver confidence to make a decision

Complex decision-making

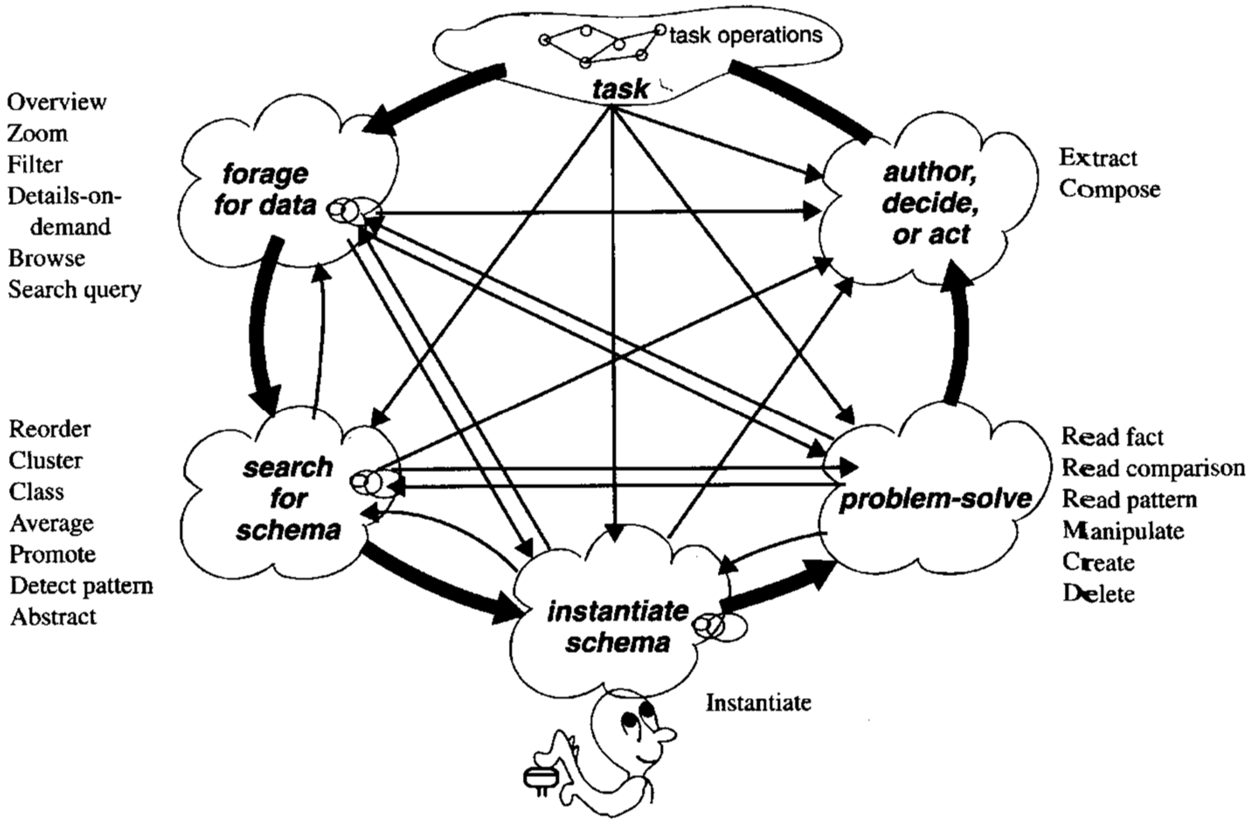
[Amar & Stasko, 2005]

Knowledge Precepts

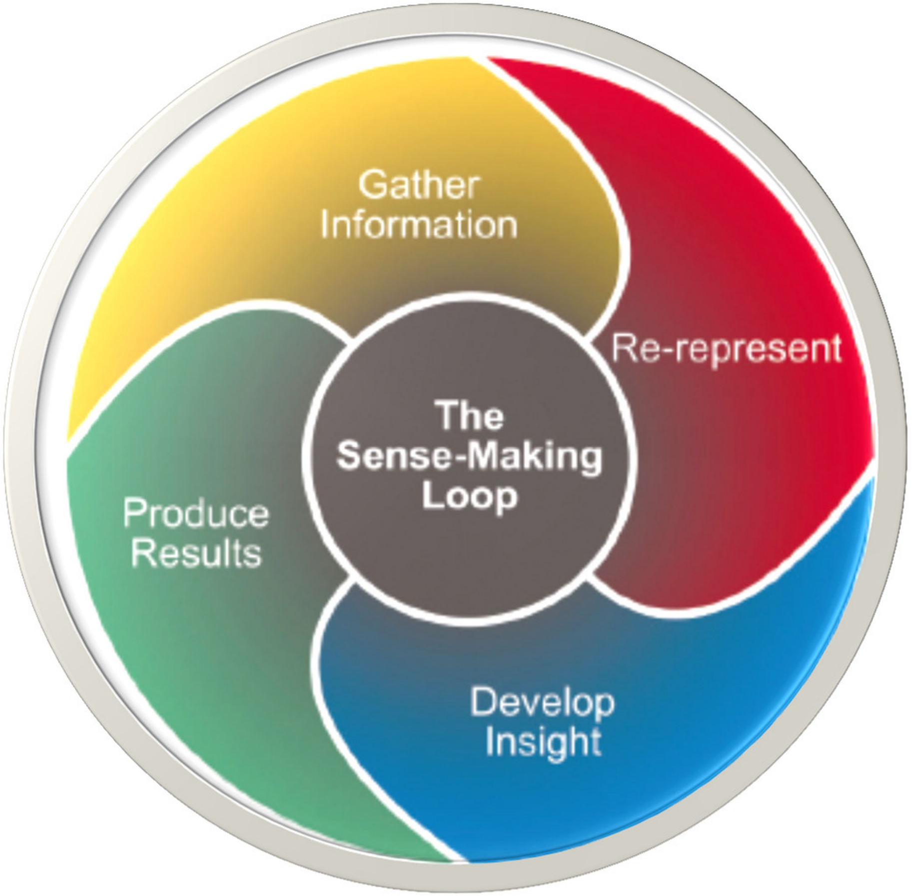
- World-view based precepts
  - Provide facilities to create, acquire, transfer knowledge about important domain parameters or
  - Provide support for discovery of multivariate explanations
  - Facilitate hypothesis testing
- Rationale-based precepts
  - Expose uncertainty
  - Concretize relationships
  - Expose cause & effect

Knowledge Crystallization Cycle

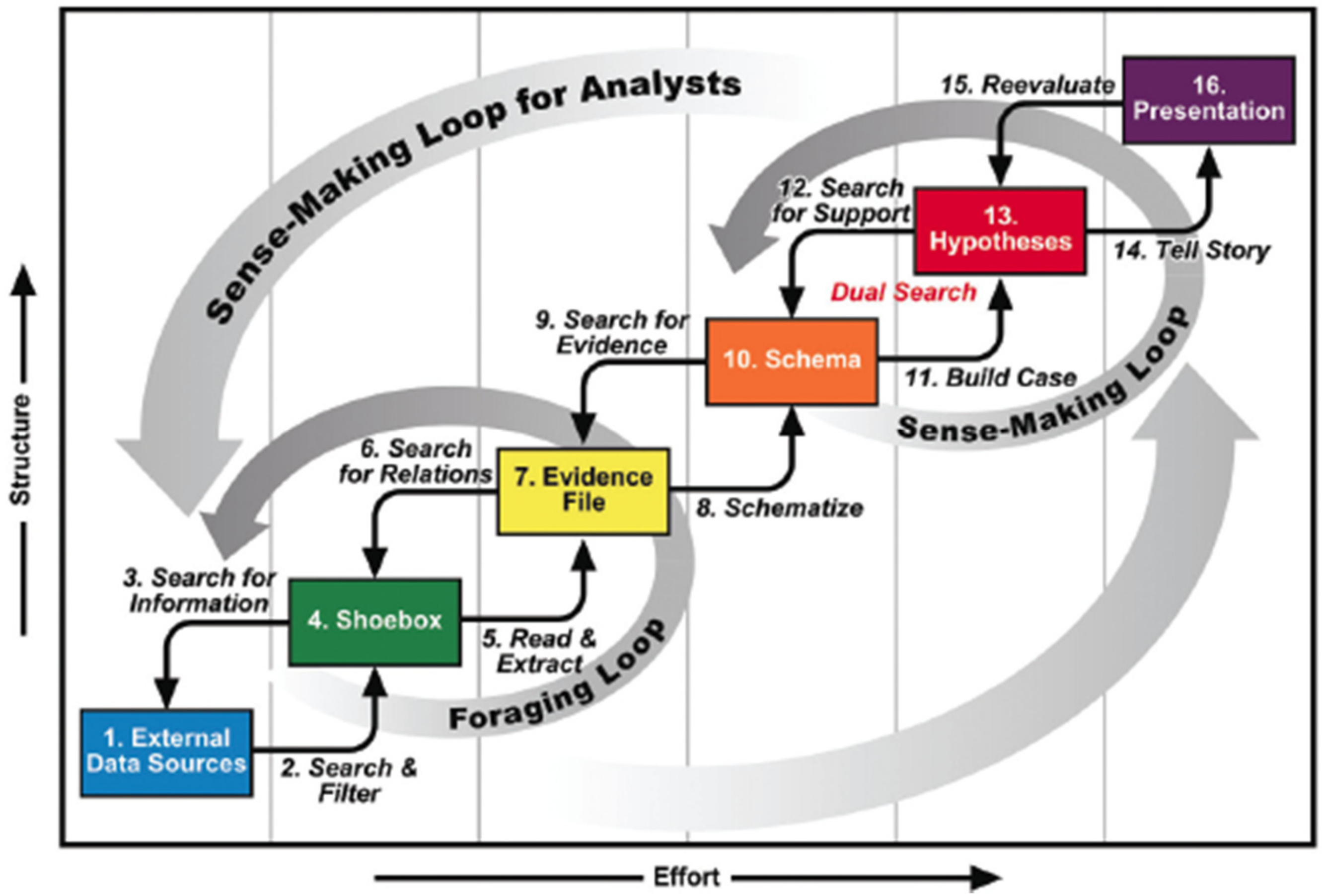
Focuses on process of knowledge extraction



The Analytic Reasoning Process



Sense-Making Loop



For some types of intelligence analysts

We’re out of time? Yes, this is the end. No, then…

HOW CAN I START…??

INFOVIS TOOLKITS

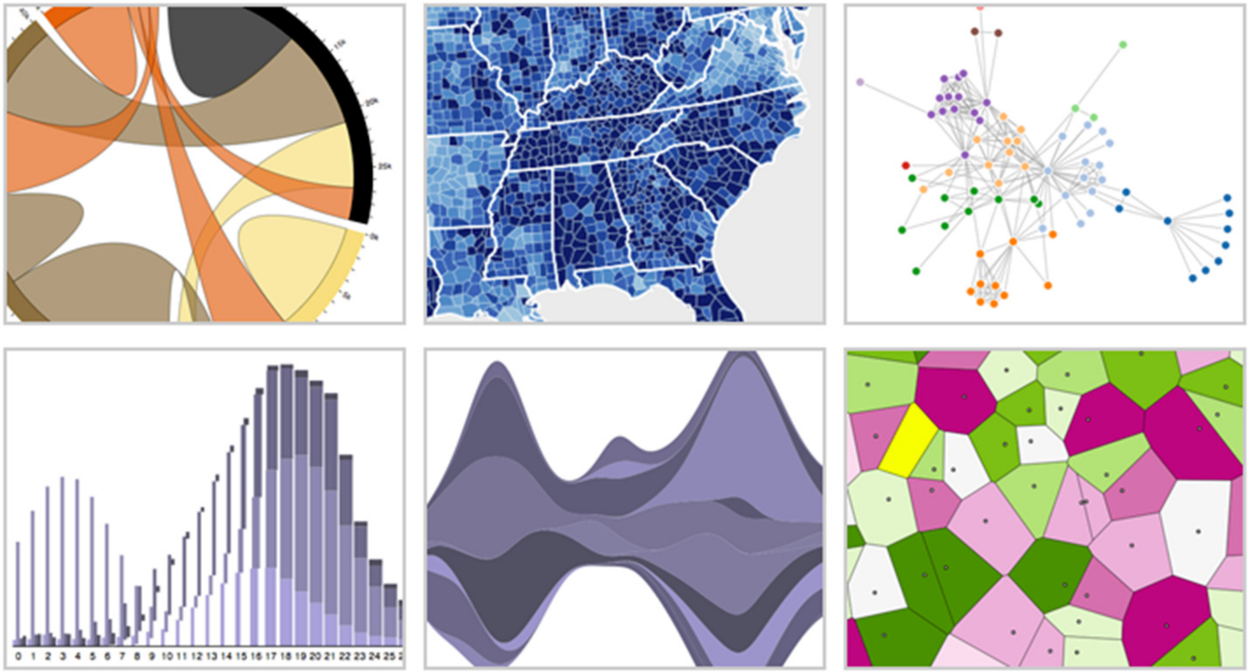
Web-based, open-source

http://mbostock.github.com/d3/

- d3 (javascript-based)

Data-Driven Documents

D3.js is a small, free JavaScript library for manipulating documents based on data.



D3 allows you to bind arbitrary data to a Document Object Model (DOM), and then apply data-driven transformations to the document. As a trivial example, you can use D3 to generate a basic HTML table from an array of numbers. Or, use the same data to create an interactive SVG bar chart with smooth transitions and interaction.

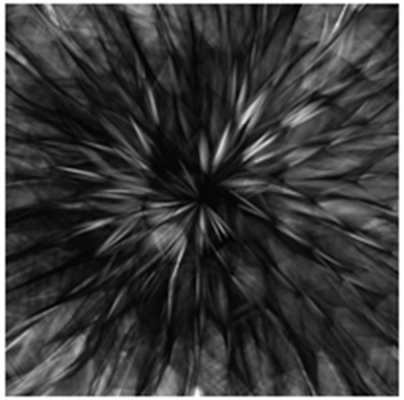
Open-source

Processing

Search

Cover \ Exhibition \ Reference \ Learning \ Download \ Shop \ About

»Feed »Forum »Wiki »Code



Announcing Processing 1.5+

A lot has changed in the nine months since 1.2.1 was released, the biggest being the introduction of "modes" for the editor, allowing you to switch between Android development and the classic desktop/web mode. More modes are coming in future releases. Please read revisions.txt to see all of the changes.

» Download Processing

» Explore the Exhibition

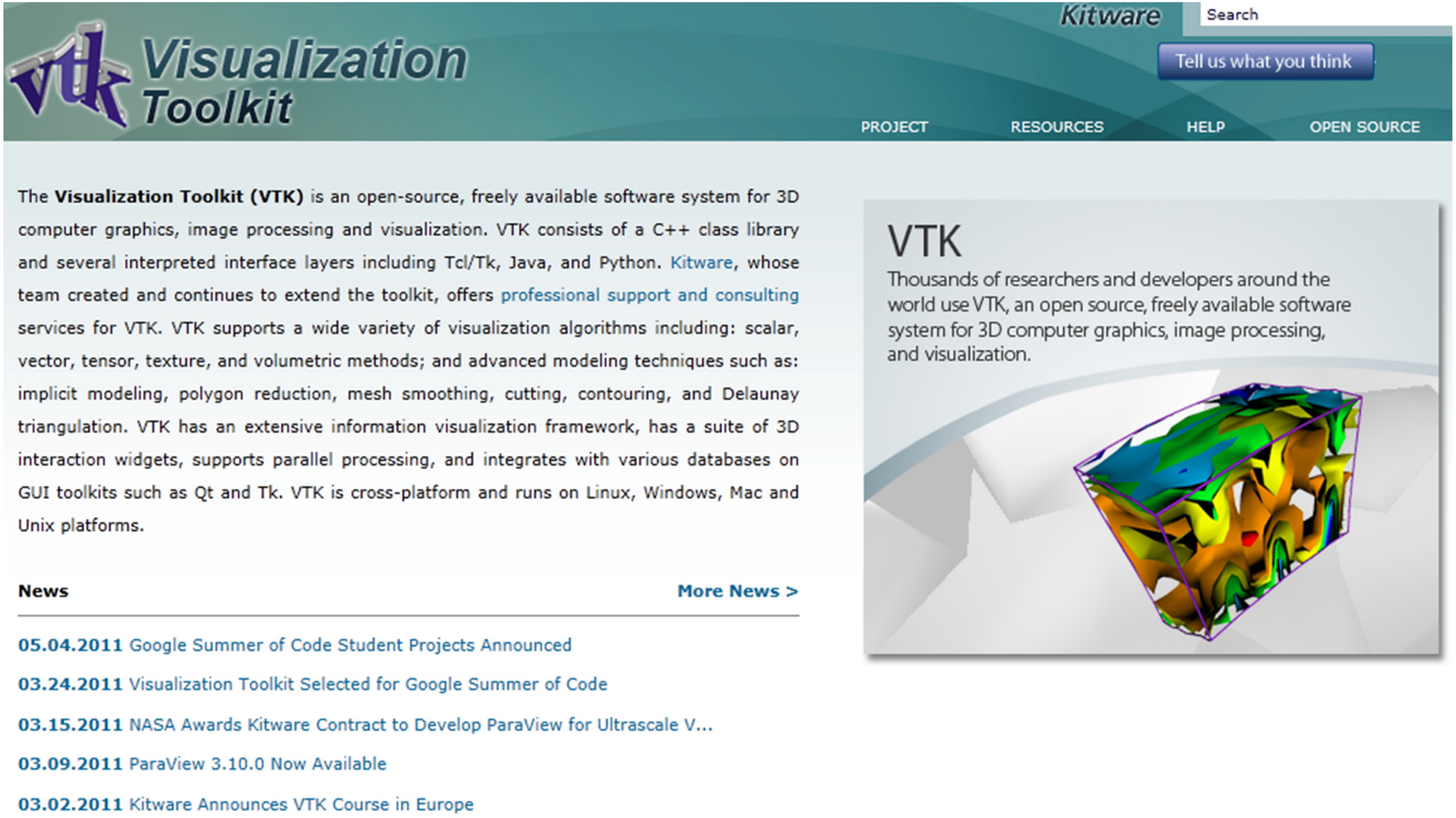
» Play with Examples

» Browse Tutorials

Processing is an open source programming language and environment for people who want to create images, animations, and interactions. Initially developed to serve as a software sketchbook and to teach fundamentals of computer programming within a visual context, Processing also has evolved into a tool for generating finished professional work. Today, there are tens of thousands of students, artists, designers, researchers, and hobbyists who use Processing for learning, prototyping, and production.

- Free to download and open source
- Interactive programs using 2D, 3D or PDF output
- OpenGL integration for accelerated 3D
- For GNU/Linux, Mac OS X, and Windows
- Projects run online or as double-clickable applications
- Over 100 libraries extend the software into sound, video, computer vision, and more...
- Well documented, with many books available

Mostly Sci-Vis


Kitware

Search

Tell us what you think

Visualization Toolkit

PROJECT RESOURCES HELP OPEN SOURCE

The **Visualization Toolkit (VTK)** is an open-source, freely available software system for 3D computer graphics, image processing and visualization. VTK consists of a C++ class library and several interpreted interface layers including Tcl/Tk, Java, and Python. Kitware, whose team created and continues to extend the toolkit, offers professional support and consulting services for VTK. VTK supports a wide variety of visualization algorithms including: scalar, vector, tensor, texture, and volumetric methods; and advanced modeling techniques such as: implicit modeling, polygon reduction, mesh smoothing, cutting, contouring, and Delaunay triangulation. VTK has an extensive information visualization framework, has a suite of 3D interaction widgets, supports parallel processing, and integrates with various databases on GUI toolkits such as Qt and Tk. VTK is cross-platform and runs on Linux, Windows, Mac and Unix platforms.

News **More News >**

05.04.2011 Google Summer of Code Student Projects Announced

03.24.2011 Visualization Toolkit Selected for Google Summer of Code

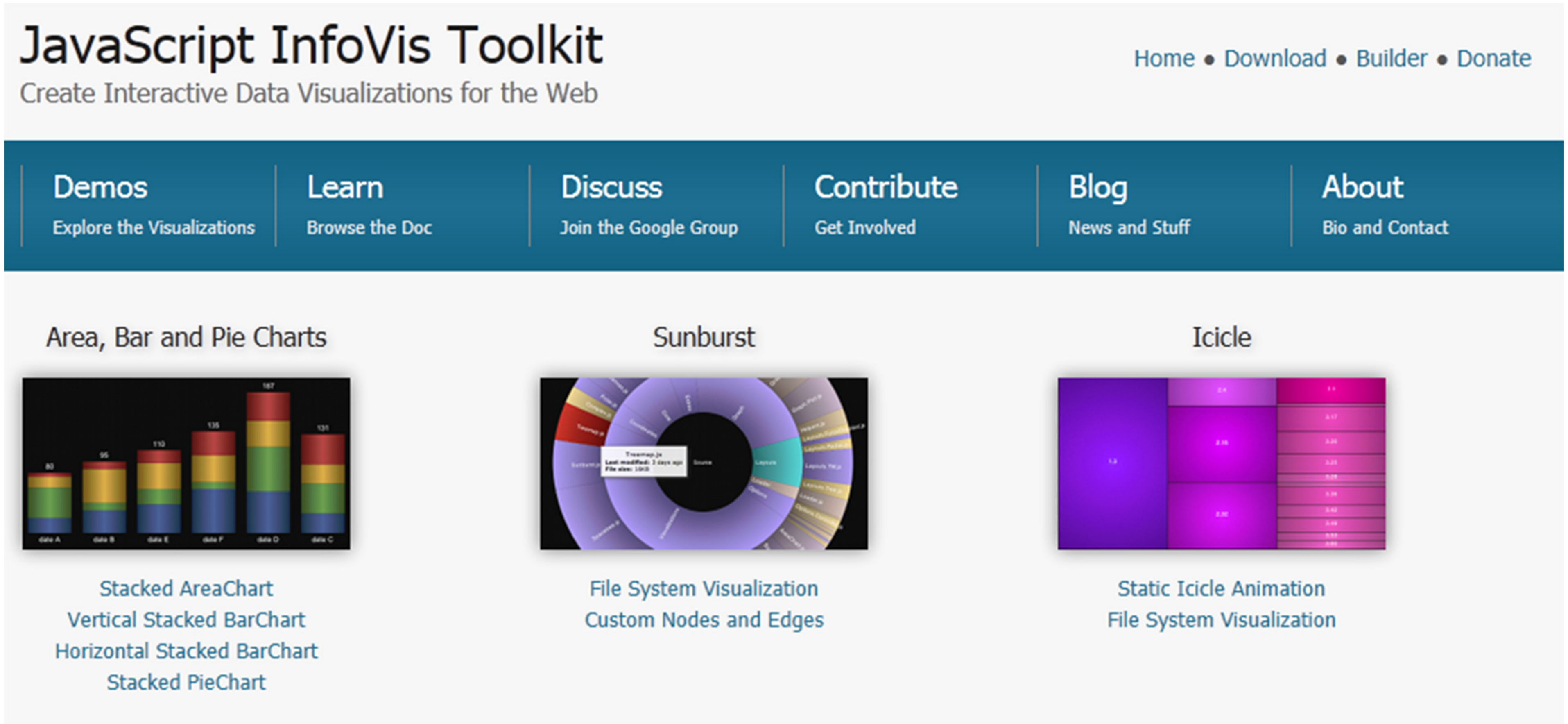
03.15.2011 NASA Awards Kitware Contract to Develop ParaView for Ultrascale V...

03.09.2011 ParaView 3.10.0 Now Available

03.02.2011 Kitware Announces VTK Course in Europe

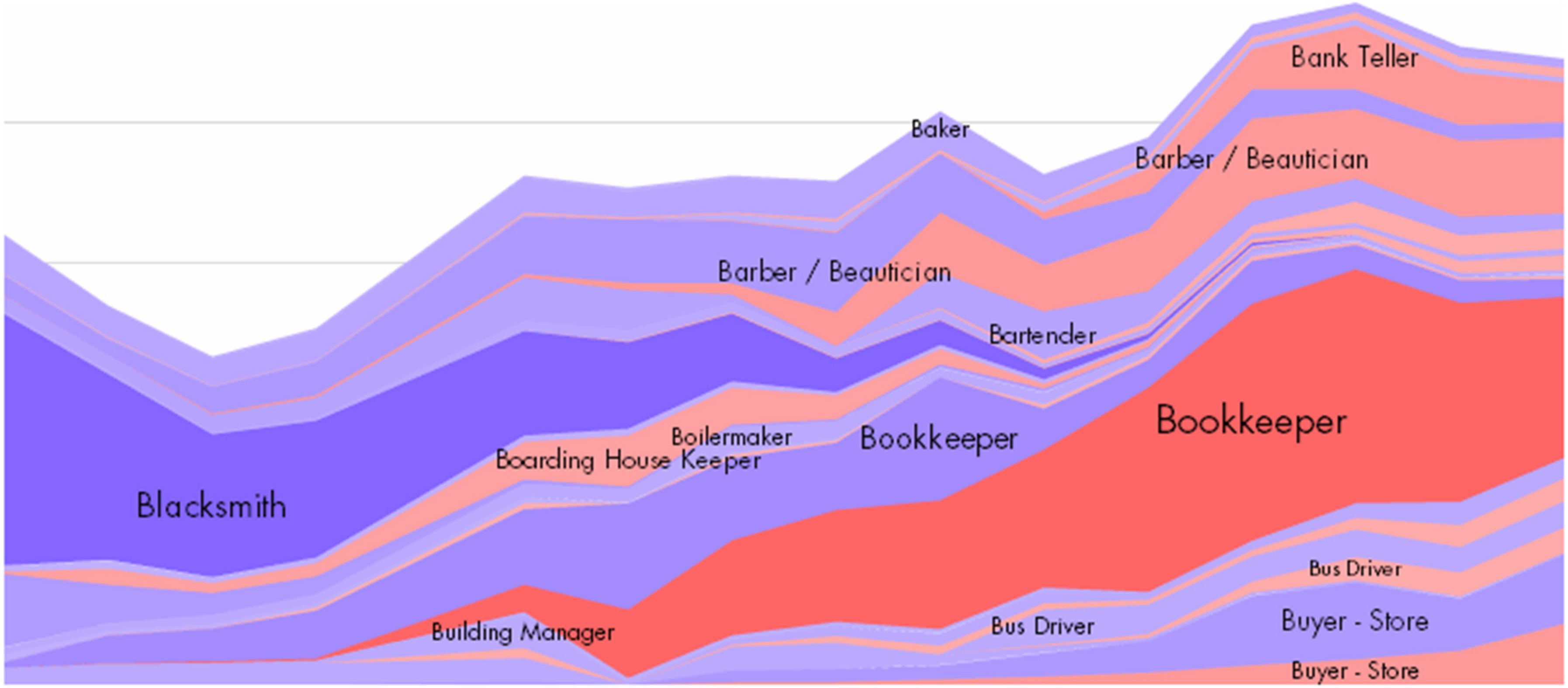
VTK

Thousands of researchers and developers around the world use VTK, an open source, freely available software system for 3D computer graphics, image processing, and visualization.



http://thejit.org/

flare DATA VISUALIZATION FOR THE WEB



Flare makes it easy to create interactive data visualizations.

Flare is an ActionScript library for creating visualizations that run in the Adobe Flash Player. From basic charts and graphs to complex interactive graphics, the toolkit supports data management, visual encoding, animation, and interaction techniques. Even better, flare features a modular design that lets developers create customized visualization techniques without having to reinvent the wheel.


Flare Demos

Java-based

prefuse

INFORMATION VISUALIZATION TOOLKIT

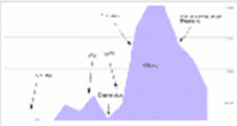
Home | Download | Gallery | Documentation | F

Sear

Download

prefuse beta
release 2007.10.21
source zip (.zip) 4.1mb

Gallery



sense.us

DocuBurst

the prefuse visualization toolkit

Prefuse is a set of software tools for creating rich interactive data visualizations. The original **prefuse** toolkit provides a visualization framework for the Java programming language. The **prefuse flare** toolkit provides visualization and animation tools for ActionScript and the Adobe Flash Player.

Prefuse supports a rich set of features for data modeling, visualization, and interaction. It provides optimized data structures for tables, graphs, and trees, a hos of layout and visual encoding techniques, and support for animation, dynamic queries, integrated search, and database connectivity. Prefuse is written in Java, using the Java 2D graphics library, and is easily integrated into Java Swing applications or web applets. Prefuse is licensed under the terms of a **BSD license**, a can be freely used for both commercial and non-commercial purposes.

The **visualization gallery** and **demonstration video** provide numerous examples of the types of applications that can be built with the prefuse toolkit.

To learn more about prefuse, take a look at the **user's manual** or the **frequently asked questions**. For users of the alpha version of the toolkit, there is also **porting guide** for migrating to the beta version.

Need help? Visit the **Help Forum on SourceForge.net** (You'll need a SourceForge login to post). Please be sure to include detailed information (e.g., stack traces, source code, etc) if you need debugging help.

If you are interested in tools for ActionScript and Flash, see the **prefuse flare** project instead.

Piccolo

Piccolo2D
Structured 2D Graphics Framework

Piccolo2D Home

- Play
- Learn
- Download
- Community
- Applications
- Press
- Sponsors

Found a Typo?

A Structured 2D Graphics Framework

Welcome to Piccolo2D! A revolutionary way to create robust, full-featured ***graphical applications*** in Java and C#, with striking visual effects such as ***zooming, animation*** and ***multiple representations***.

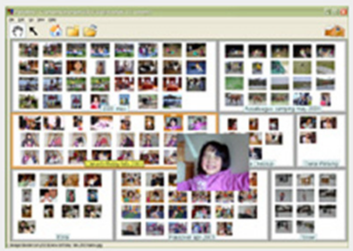
Piccolo2D is a toolkit that supports the development of 2D structured graphics programs, in general, and Zoomable User Interfaces (ZUIs), in particular. A ZUI is a new kind of interface that presents a huge canvas of information on a traditional computer display by letting the user smoothly zoom in, to get more detailed information, and zoom out for an overview. We use a "scene graph" model that is common to 3D environments. Basically, this means that Piccolo2D maintains a hierarchal structure of objects and cameras, allowing the application developer to orient, group and manipulate objects in meaningful ways.

Why use Piccolo2D? It will allow you to build structured graphical applications without worrying so much about the low level details. The infrastructure provides efficient repainting of the screen, bounds management, event handling and dispatch, picking (determining which visual object the mouse is over), animation, layout, and more. Normally, you would have to write all of this code from scratch. Additionally, if you want to build an application with zooming, that's built right into the framework too.

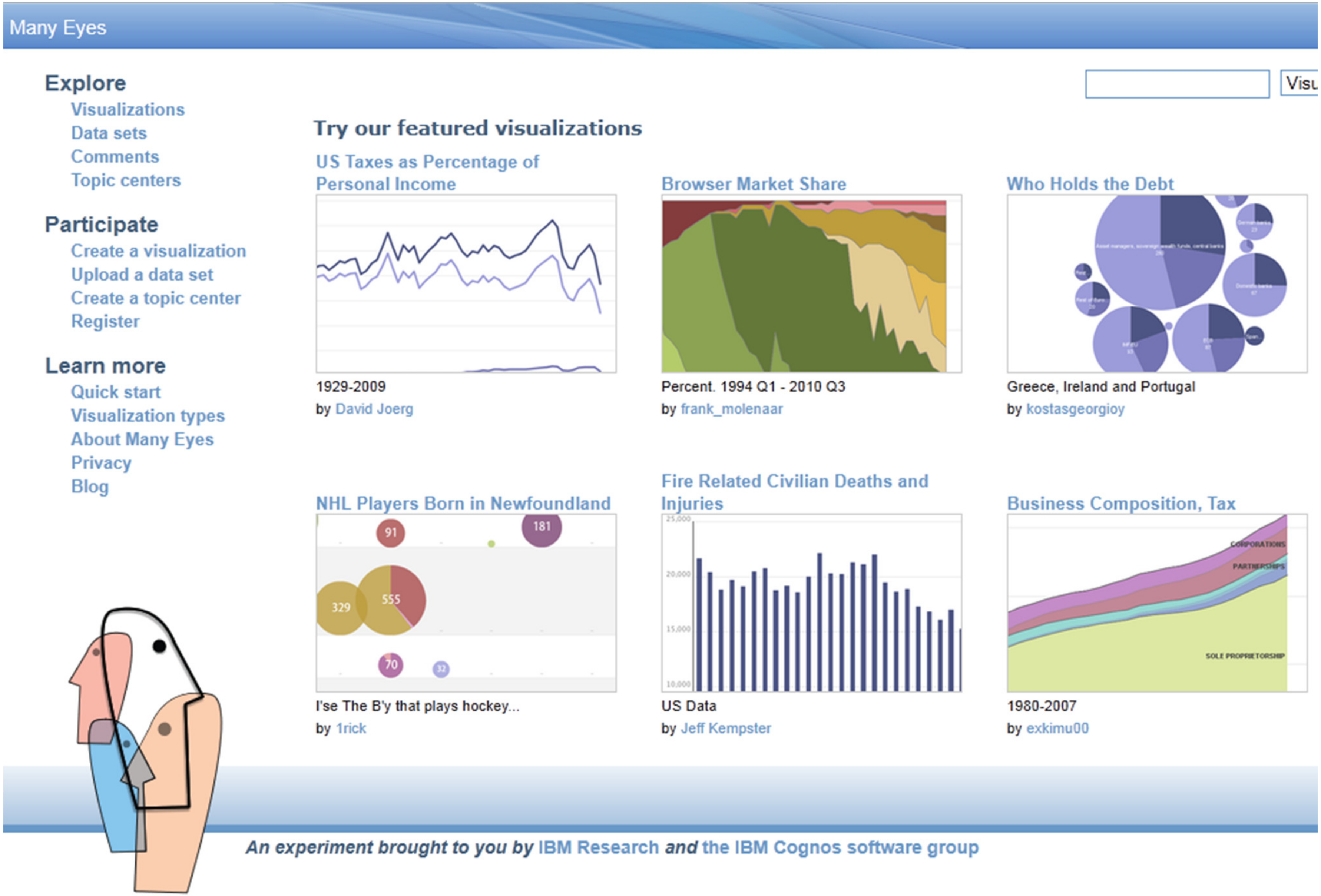
What exactly is it? Piccolo2D is a layer built on top of a lower level graphics API. There are currently three versions of the toolkit: Piccolo2D.Java, Piccolo2D.NET and PocketPiccolo2D.NET (for the .NET Compact Framework). The java version is built on Java 2 and relies on the Java2D API to do its graphics rendering. The .NET version is built on the .NET Framework and relies on the GDI+ API to do its graphics rendering. This makes it easy for Java and C# programmers, even those targeting PDAs, to build their own animated graphical applications. And best of all, Piccolo2D is ***free*** and ***open source!***

References

The primary paper describing and analyzing the Piccolo2D architecture: Bederson, B. B., Grosjean, J., & Meyer, J. (2004). Toolkit Design for Interactive Structured Graphics, IEEE Transactions on Software Engineering, 30 (8), pp. 535-546.



ManyEyes (but not really a "toolkit")



R

- Stats software
- Lots of data vis/plots packages

Google Charts API

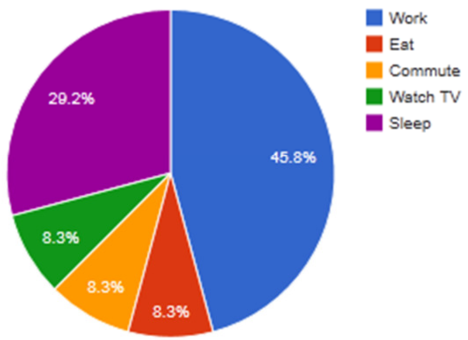
Google Chart Tools

Display live data on your site

Google chart tools are powerful, simple to use and free. Try out our rich gallery of interactive charts and data tools.

Get Started

Pie Chart - view source



more

Rich Gallery

Choose from a variety of charts. From simple scatter plots to hierarchical tree-maps, find the best fit to your data.

Customizable

Make the charts your own. Configure an extensive set of options to perfectly match the look and feel of your website.

HTML5 / SVG

Cross-browser compatibility (adopting VML for older IE versions) and cross platform portability to iOS and new Android releases. No plugins are needed.

Free

Completely free for all uses: commercial, governmental, personal or educational.

Interactive

Simple event handling enables a cohesive interactive experience with the surrounding web page.

Dynamic

Present real-time data using a variety of data connection tools.

Graph Vis Library

JUNG

Java Universal Network/Graph Framework

Overview
Download
Documentation
Examples
Wiki
Projects Using JUNG
FAQ
Support
Team
Presentations
Bug Tracker
Sourceforge
Acknowledgements

Links

SOURCEFORGE.NET

Overview

JUNG — the Java Universal Network/Graph Framework--is a software library that provides a common and extendible langua
graph or network. It is written in Java, which allows JUNG-based applications to make use of the extensive built-in capabili

The JUNG architecture is designed to support a variety of representations of entities and their relations, such as directed a
It provides a mechanism for annotating graphs, entities, and relations with metadata. This facilitates the creation of analyt
the metadata attached to each entity and relation.

The current distribution of JUNG includes implementations of a number of algorithms from graph theory, data mining, and
random graph generation, statistical analysis, and calculation of network distances, flows, and importance measures (centr

JUNG also provides a visualization framework that makes it easy to construct tools for the interactive exploration of networ
create their own custom layouts. In addition, filtering mechanisms are provided which allow users to focus their attention,

As an open-source library, JUNG provides a common framework for graph/network analysis and visualization. We hope that
anothers' development efforts, and thus avoid continually re-inventing the wheel.

— The JUNG Framework Development Team

Many More

- …

My sage advice

- Learn about your data
- Be creative & have fun building your own vis
- Build on what exists but also modify to adjust to your own data
- Do not neglect interaction
- Please choose good colors
  Permutations of 255,0,0 hardly ever look good!
- Be critical about visualizations you see!

Today We Covered

- Basics of Visualization
  – What is InfoVis
  – How do you generate an infovis
  – What are the basic building blocks
- Overview of Visualization Techniques
  – For a variety of datasets
  – Interaction
  – Toolkits

Acknowledgements

This lecture was inspired by slides sets from:

- Sheelagh Carpendale
- Christopher Collins
- Jean-Daniel Fekete
- Jeffrey Heer
- Tobias Isenberg
- John Stasko