ΠΡΟΣΧΕΔΙΑΣΜΕΝΟΣ & ΕΥΕΛΙΚΤΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ

PLAN-DRIVEN &AGILE PROGRAMMING

Κωδικός Θ: , Κωδικός Ε:
Ώρες (Θ - ΑΠ - Ε): 2 - 1 - 3
Προ-απαιτούμενα: Προγραμματισμός, Αρχές Τεχνολογίας Λογισμικού Βάσεις Δεδομένων

Data Visualization

ΘΕΩΡΙΑ & ΕΡΓΑΣΤΗΡΙΟ 3 & 4

Σημερινό μάθημα

Data Visualization

- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Data Visualization

Acknowledgement

Ένα μεγάλο μέρος του υλικού που παρουσιάζεται στις παρακάτω διαφάνειες είναι από τις διαφάνειες του «Course: CS 171 Visualization» του πανεπιστημίου Harvard University.
Η ιστοσελίδα του μαθήματος είναι:
http://www.cs171.org/

Data Visualization

Ενδεικτικά Μαθήματα δύο κορυφαίων Πανεπιστημίων

Stanford University

Course: cs448b Data Visualization

https://graphics.stanford.edu/wikis/cs448b-10-fall

Course Syllabus

- The Value of Visualization
- Data and Image Models
- Visualization Design
- Exploratory Data Analysis
- Multidimensional Data Visualization
- Graphical Perception
- Interaction
- Interactive Visualization
- JavaScript / Protovis Tutorial
- Flash / Flare Tutorial
- Animation
- Color
- Mapping & Cartography
- Using Space Effectively
- Graph Layout and Network Analysis
- Text Visualization
- Identifying Design Principles

Harvard University

Course: CS 171 Visualization

http://www.cs171.org/

Course Syllabus **© 2009 Harvard University**

Act I: Foundations

1. **Introduction** (What is visualization? Course overview.)
2. **Design Principles** (Graphical Integrity, Lie Factor, Data-Ink Ratio, Chartjunk, Context, Small Multiples, Space & Time, Layering.)
3. **Process & Validation** (Visualization Process, Validation Models.) **Reading**: A Nested Model for Visualization Design and Validation. Tamara Munzner. IEEE TVCG 15(6) (Proc. InfoVis 2009) [PDF]
4. **Data & Image Models** (Data Types. Data Scraping. Image Models.) **Reading**: Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases (extended paper) Chris Stolte, Diane Tang and Pat Hanrahan, IEEE Transactions on Visualization and Computer Graphics, Vol. 8, No. 1, January 2002. [PDF]
5. **Perception** (Looking vs. Seeing. The Eye. The Retina. Visual Cortex. Things that Pop.) **Reading**: Ware Chapters 1 & 2
6. **Patterns** (Contours. Textures. Gestalt Principles.) **Reading**: Ware Chapter 3
7. **Color** (Color Processing. Color Spaces. Color Design.) **Reading**: Ware Chapter 4
8. **Cognition** (Image Gist. Scene Understanding. Long-term Memory.) **Reading**: Ware Chapter 6
9. **Interaction** (Overview & Detail. Zooming. Focus & Context. Dynamic Queries. Brushing & Linking. Animation. Off-The-Desktop Interaction.) **Reading**: A review of overview+detail, zooming, and focus+context interfaces. Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. ACM Computing Surveys 41(1), 2008. [PDF]

Act II: Methods

1. Statistical Graphs (Aspect Ratio. Multiscale Banking to 45 Degrees. Scales. Nominal Comparisons. Percentage and Proportion Data. Scatter Plots. Box & Whisker Plots.) Reading: Multi-Scale Banking to 45 Degrees Jeffrey Heer, Maneesh Agrawala. Proc. InfoVis 2006, published as IEEE Transactions on Visualization and Computer Graphics (TVCG), 12(5), Sep/Oct 2006, pages 701-708. [PDF]
2. Maps (Google Maps. Map Projections. Cartograms. Choropleth Maps. Proportional Symbol Maps. Flow Maps. Real-Time Maps. Thematic Maps.) Reading: Unfolding the Earth: Myriahedral Projections. Jarke J. van Wijk. The Cartographic Journal, Vol. 45, No. 1, pp.32-42, February 2008. [PDF]
3. Trees & Networks (Definitions. Visualizing Trees. Visualizing Networks.) Reading: Graph Visualisation in Information Visualisation: a Survey Ivan Herman, Guy Melancon, M. Scott Marshall. IEEE Transactions on Visualization and Computer Graphics, 6(1), pp. 24-44, 2000. [PDF]
4. High Dimensionality (Vis Techniques. Parallel Coordinates. Glyphs. Dimensionality Reduction. MDS. PCA.) Reading: Glimmer: Multilevel MDS on the GPU. Stephen Ingram, Tamara Munzner and Marc Olano. IEEE TVCG, 15(2):249-261, Mar/Apr 2009. [Project Page], A tutorial on Principal Components Analysis, Lindsay I Smith [PDF]

Act III: Applications

1. Visualization and the Arts (Ben Fry)
2. Communicating with Visualizations (Bang Wong, Broad Institute)
3. Visualization in Biology (Miriah Meyer, Harvard and Janet Iwasa, HMS)
4. Network Visualization (Fernanda Viegas and Martin Wattenberg, IBM)

Data Visualization: Ορισμός

Data visualization

Είναι η σπουδή της γραφικής αναπαράστασης δεδομένων,

"information which has been abstracted in some schematic form, including attributes or variables for the units of information".

Data Visualization: Ορισμός

According to Friedman (2008)

The **main goal of data visualization** is:

to communicate information clearly and effectively through graphical means.

Ο κύριος στόχος της οπτικοποίησης δεδομένων είναι να **μεταδώσει πληροφορία** με τρόπο:

Ξεκάθαρο &
αποτελεσματικό

χρησιμοποιώντας διάφορα **γραφικά τεχνουργήματα**.

Data Visualization: Ορισμός

Για να μεταφέρουμε την πληροφορία αποτελεσματικά στον τελικό αναγνώστη, μέσω ενός γράφου, θα πρέπει να δώσουμε έμφαση :

- στην **αισθητική απεικόνιση** της πληροφορίας &
- στην **λειτουργικότητα** του γράφου.

παρέχοντας στον αναγνώστη **ενόραση** στην ωφέλιμη για αυτόν κάθε φορά πληροφορία από ένα μεγάλο σύνολο πληροφοριών.

Data Visualization

Data visualization είναι στενά συνδεδεμένο με:

- Information graphics,
- Information visualization,
- Scientific visualization and
- Statistical graphics.

Στην νέα χιλιετία το data visualization συμμετέχει ενεργά στην έρευνα, την εκπαίδευση και την ανάπτυξη λογισμικού.

Σύμφωνα με τον «Post et al (2002)» σύνδεσε τα γνωστικά πεδία “scientific and information visualization”

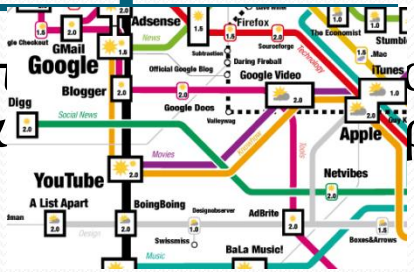
Προηγμένες τεχνικές οπτικοποίησης δεδομένων

Η οπτική αναπαράσταση δεδομένων και πληροφοριών είναι μια δυναμικά εξελισσόμενη περιοχή η οποία επεκτείνεται από απλές οπτικές αναπαραστάσεις σε αναπαραστάσεις που εφαρμόζονται σύνθετα μαθηματικά μοντέλα. Για να εκπληρωθούν αυτές οι γραφικές αναπαραστάσεις, έχουν αναπτυχθεί προηγμένες τεχνικές που επιτρέπουν αποτελεσματική πρόσβαση, γρήγορη επισκόπηση και διεισδυτική θεώρηση των δεδομένων με διαδραστικούς τρόπους. Ενδεικτικά παραδείγματα των τεχνικών αυτών επιτρέπουν:

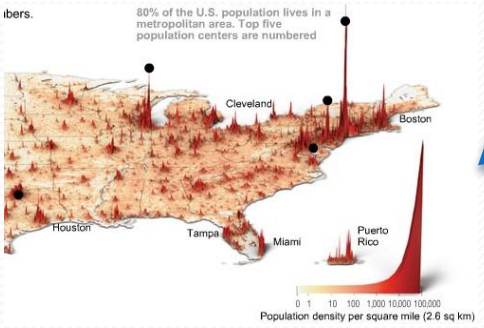
- γρήγορη επισκόπηση σε πρώτο χρόνο και εμβάθυνση στα δεδομένα σε δεύτερο χρόνο,
- εστίαση και φιλτράρισμα σε δεδομένα υψηλού ή πρωτεύοντος ενδιαφέροντος (zoom and filter) αφαιρώντας τα ασήμαντα στοιχεία και
- προοδευτική εξόρυξη πληροφορίας (details on demand) μέσω επιλογής αντικειμένων των οποίων η λεπτομέρεια ενδιαφέρει.

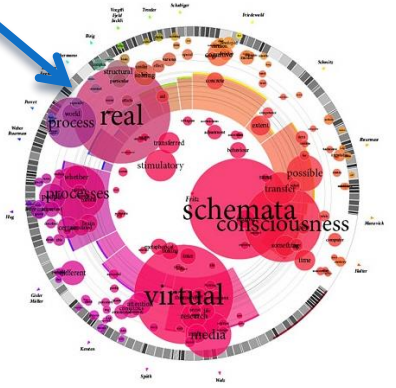
Προηγμένες τεχνικές οπτικοποίησης δεδομένων

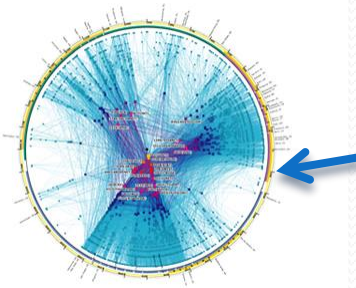
Κάπ ο ενδιαφέρουσες προσεγγίσεις σ
αναπ ροφοριών συμπεριλαμβάνουν



- **Mindmaps,**
- **Displaying News,**
- **Displaying Data,**
- **Displaying connections,**
- **Displaying websites,**
- **Articles**







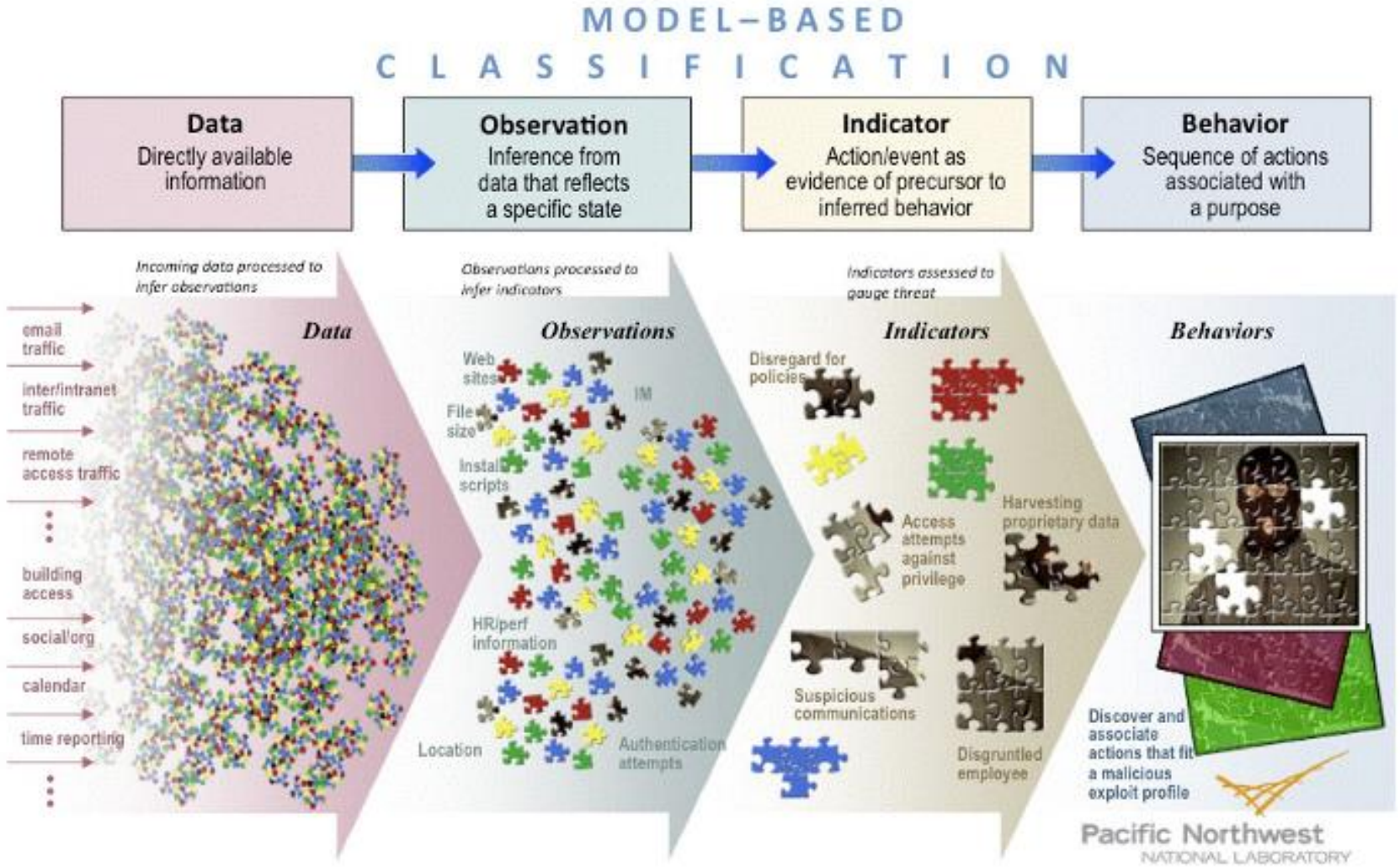
Data Visualization: Ένα Μοντέλο

Η γραφική αναπαράσταση των δεδομένων βασίζεται σε ένα μοντέλο τεσσάρων σταδίων

1. Το πρώτο στάδιο αφορά την συλλογή των πληροφοριών από τις πηγές (**Data**),
2. Το δεύτερο στάδιο διατυπώνει την διαδικασία για την ερμηνεία και εξαγωγή συμπερασμάτων αυτών των δεδομένων (**Observation**),
3. Το τρίτο στάδιο ασχολείται με την ανάδειξη των ενεργειών / γεγονότων με σκοπό να προκύψουν συμπεριφορές (**Indicator**) και
4. Το τελευταίο στάδιο ακολουθεί μια σειρά από ενέργειες συνδυασμένες από ένα λόγο /σκοπό (**Behavior**) και είναι το στάδιο που είναι φανερό στον χρήστη.

Data Visualization: Ένα Μοντέλο

Αρχιτεκτονική Οπτικοποίησης Δεδομένων



Data Visualization:
Τι Περιλαμβάνει

Data Visualization: Τι περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Data Visualization: Τι περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Design

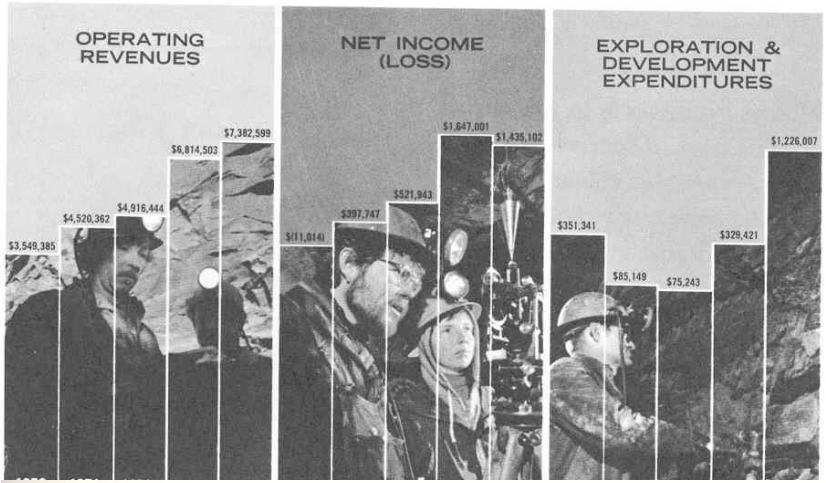
Data Visualization:

Design

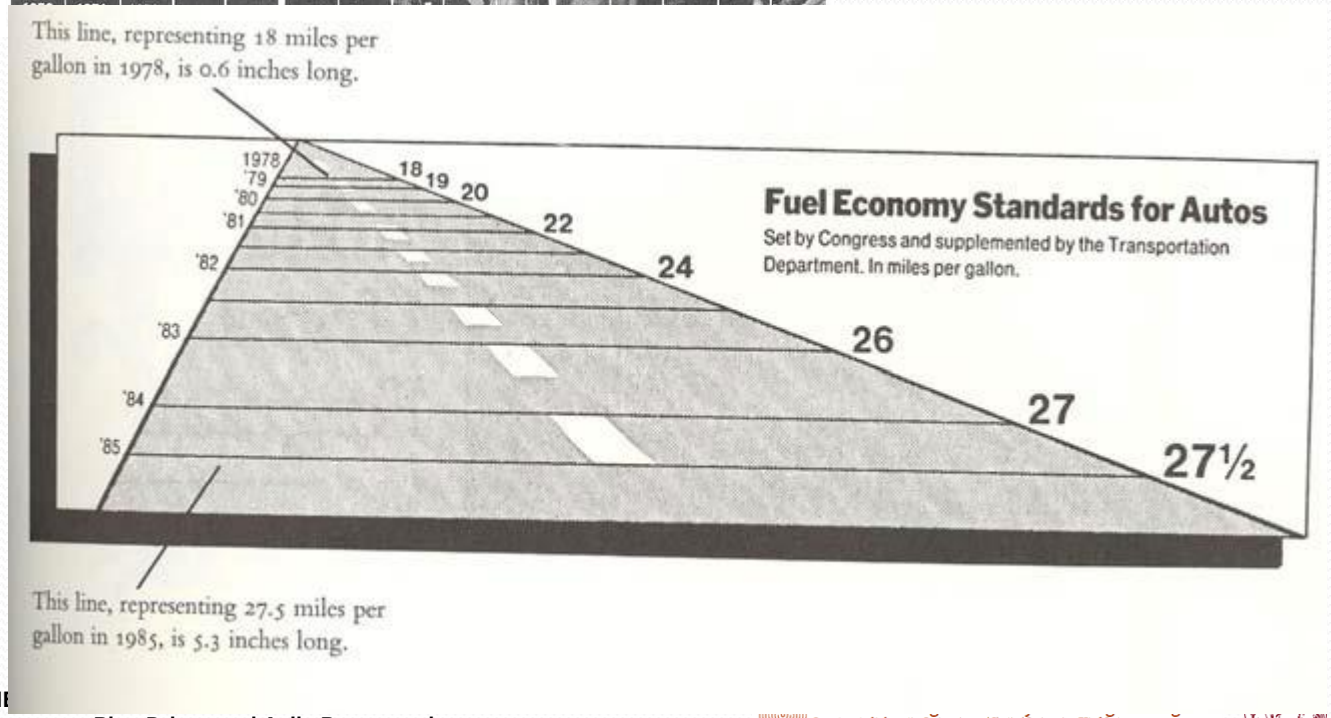
- Graphical Integrity
- Design Principles
- Design Elements

Data Visualization: Design - Graphical Integrity

Missing Scales



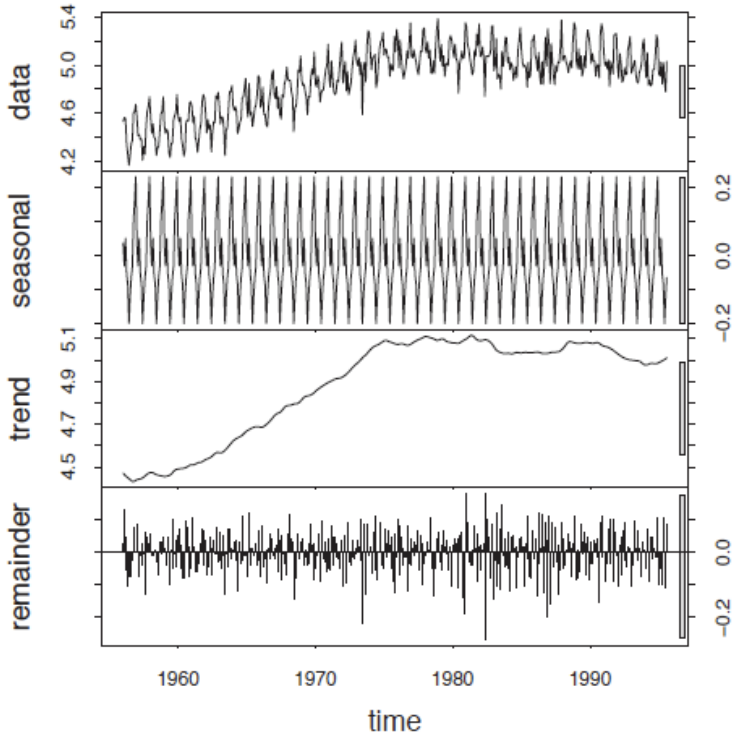
The Lie Factor



Data Visualization: Design - Graphical Integrity

Design Distortions

Decomposition



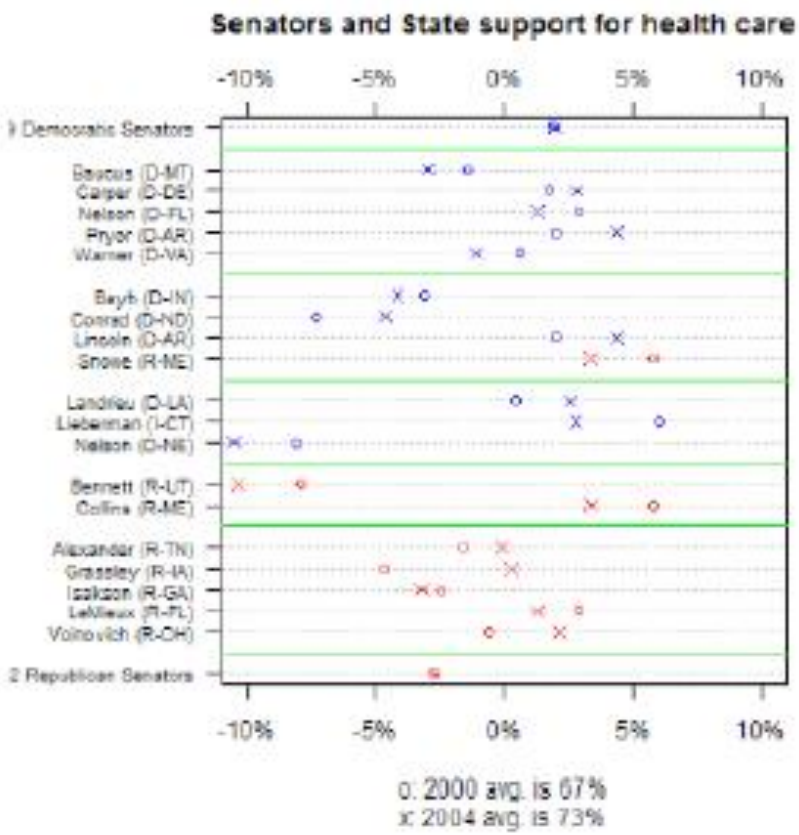
Tufte, VDQI

Data Visualization: Design - Graphical Principles

Tufte's Design Principles

- Above all else show the data
- Maximize data-ink ratio
- Erase non-data ink
- Erase redundant data ink
- Revise and edit

Data Visualization: Design - Graphical Principles



Andrew Gelman, Nov. 2009

Junk Charts

Data Visualization: Design - Graphical Principles

Data Visualization: Design - Graphical Principles

Subjective Dimensions

- Aesthetics: Attractive things are perceived as more useful than unattractive ones
- Style: Communicates brand, process, who the designer is
- Playfulness: Encourages experimentation and exploration
- Vividness: Can make a visualization more memorable

Pat Hanrahan, Nov 2007

Data Visualization: Design - Elements

CRAP

- **C**ontrast
- **R**epetition
- **A**lignment
- **P**roximity

Data Visualization: Design - Elements

CRAP- Contrast

Peter's cake metaphor ties in nicely with Galls Law

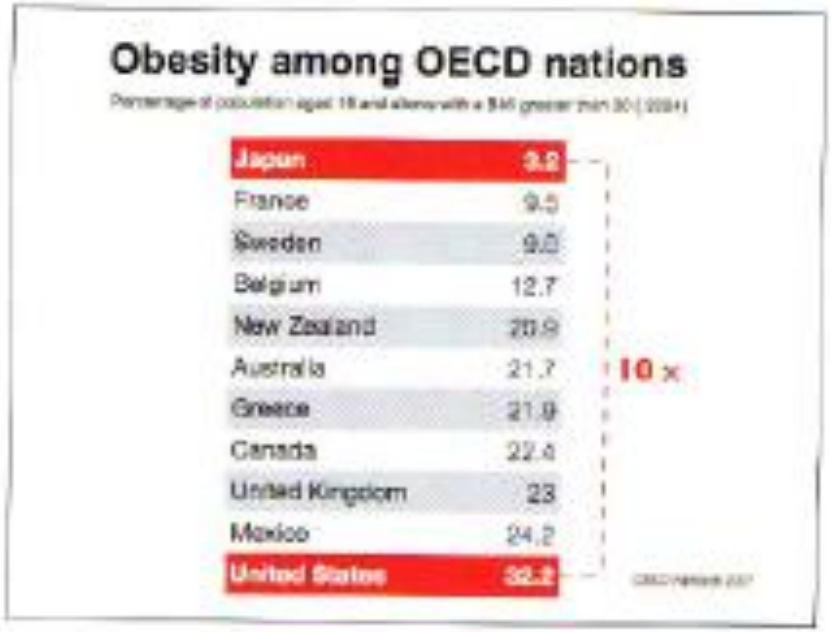
> A complex system that works is invariably found to have evolved from a simple system that worked. The inverse proposition also appears to be true: A complex system designed from scratch never works and cannot be made to work. You have to start over, beginning with a working simple system.

Data Visualization: Design - Elements

CRAP- Contrast

Before

After

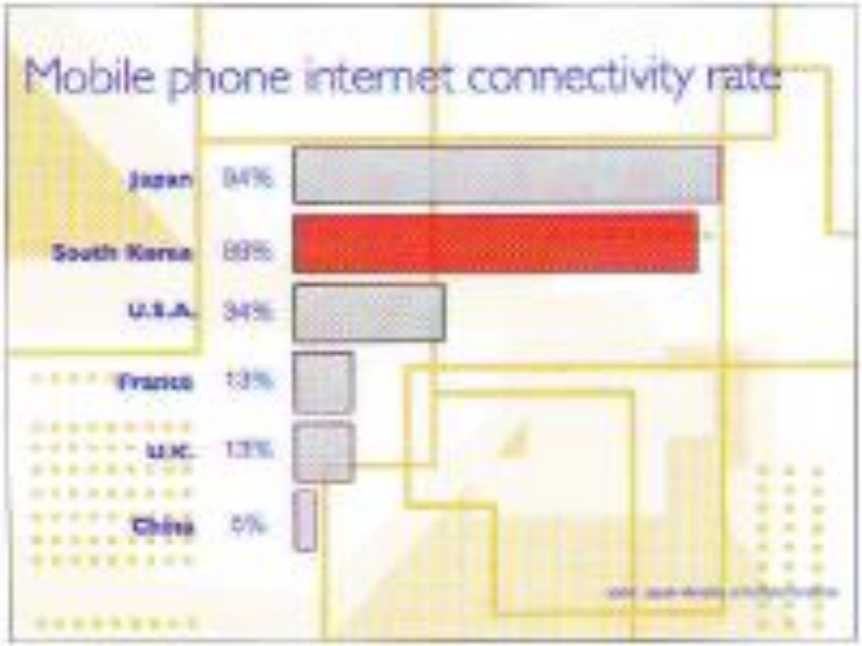


G. Reynolds, Presentation Zen

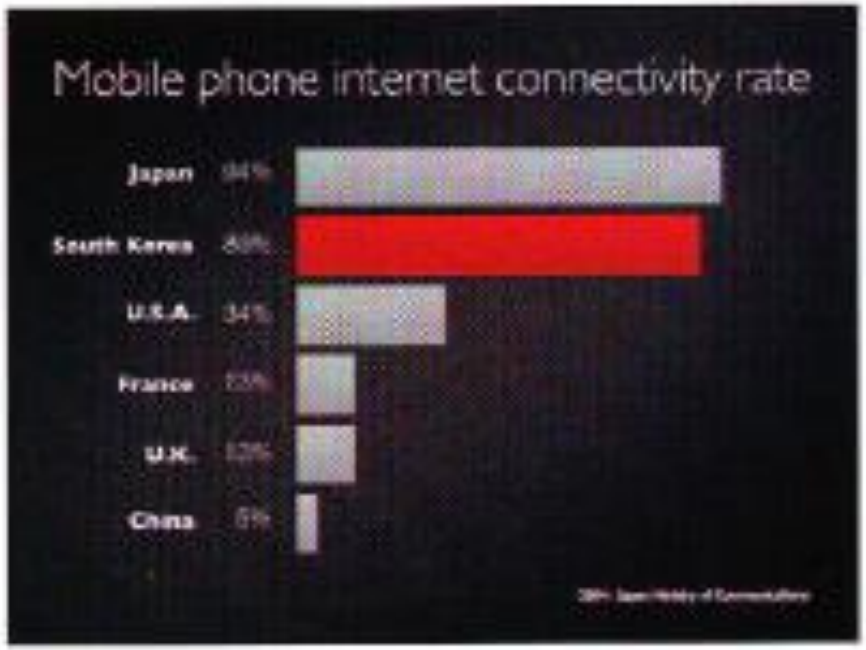
Data Visualization: Design- Elements

CRAP- Contrast

Before



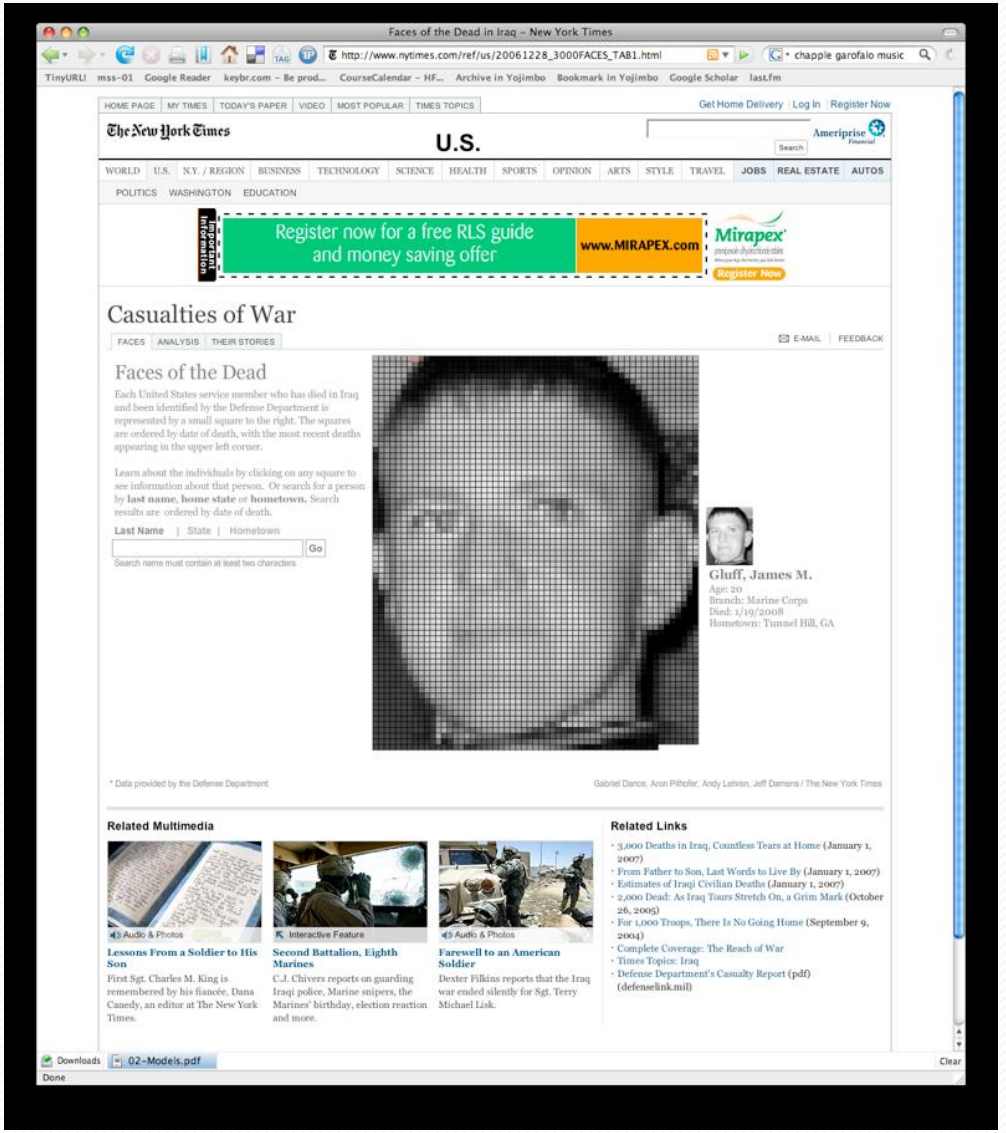
After



G. Reynolds, Presentation Zen

Data Visualization: Design- Elements

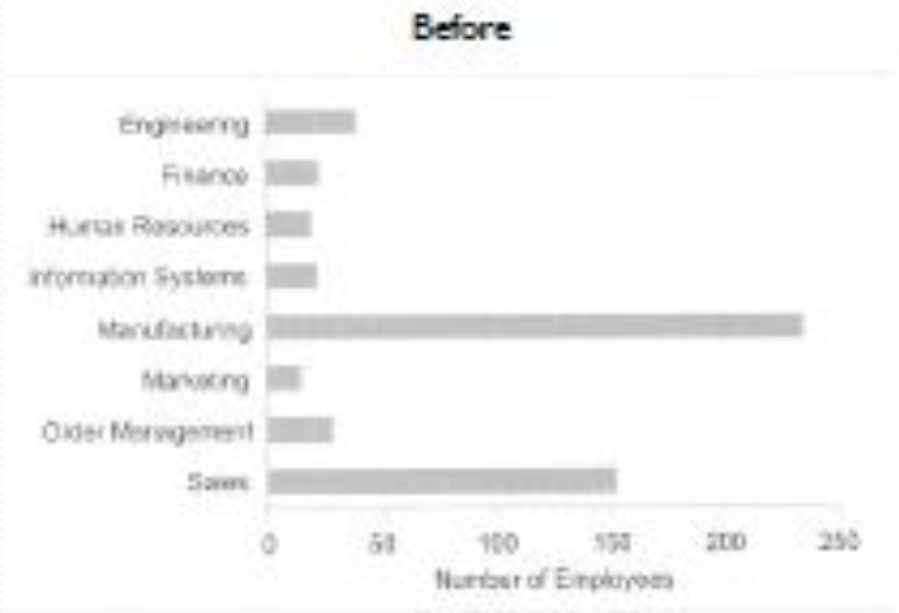
CRAP - Repetition

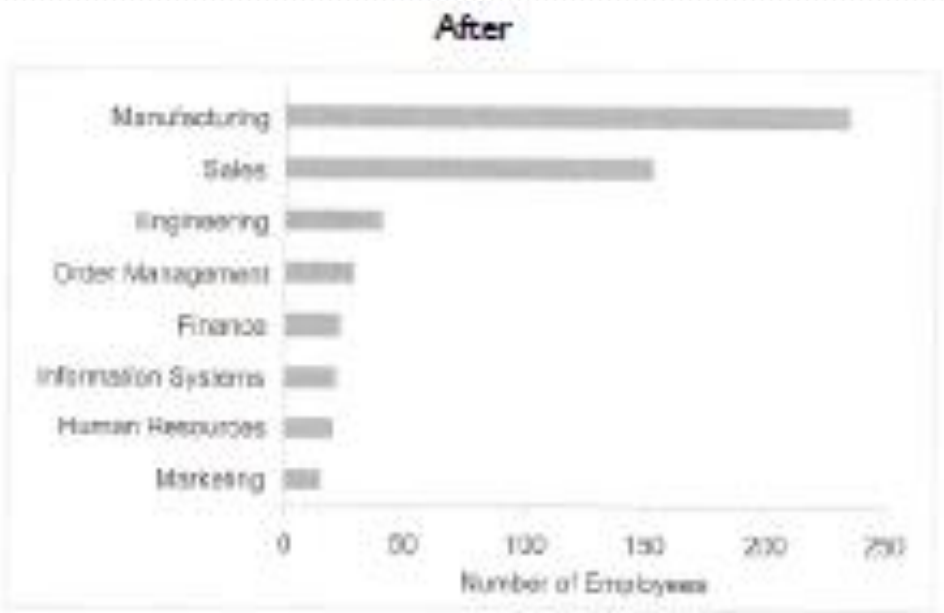


Faces of the Dead in Iraq, NY Times

Data Visualization: Design- Elements

CRAP - Alignment

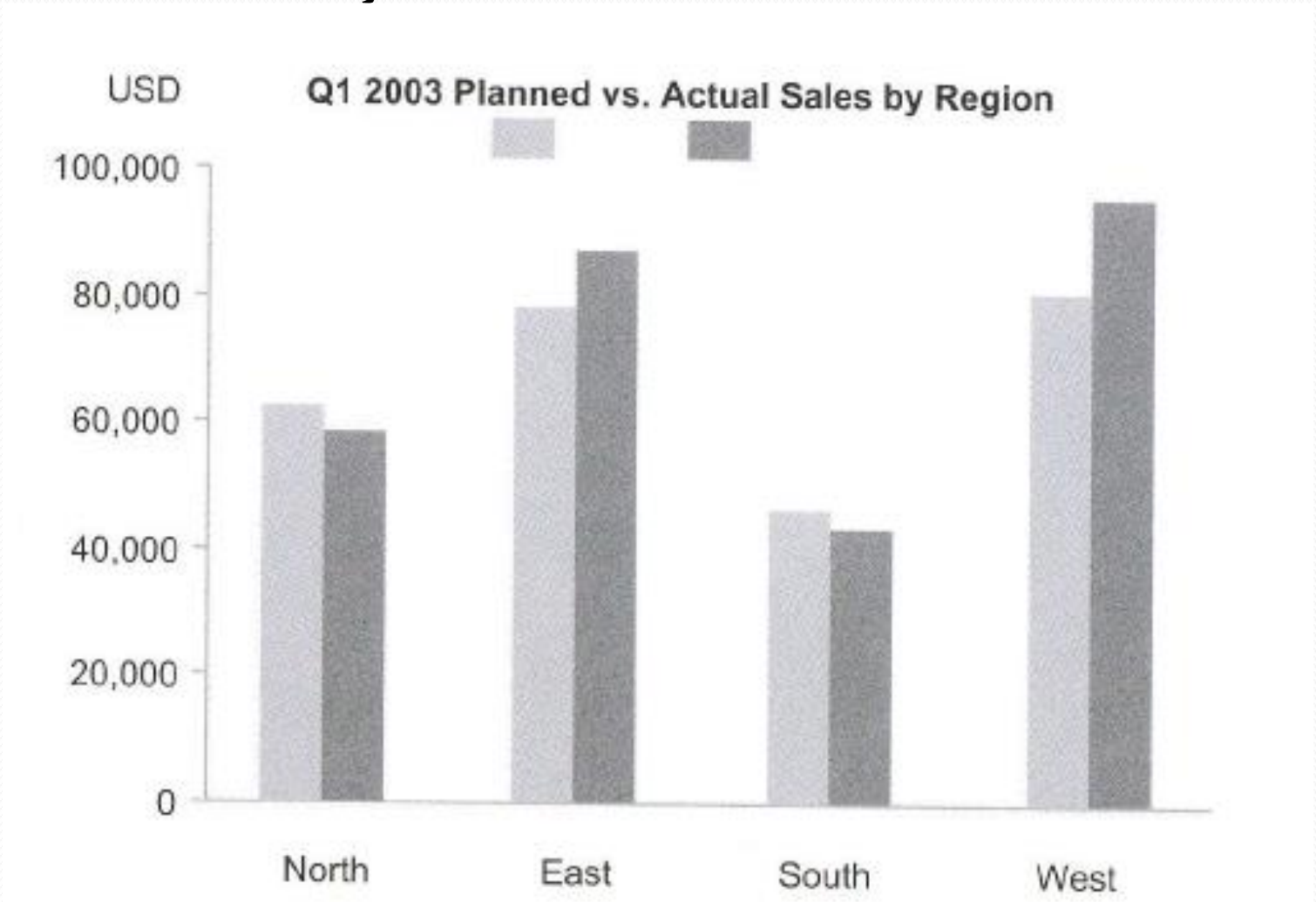




S. Few, Show Me The Numbers

Data Visualization: Design- Elements

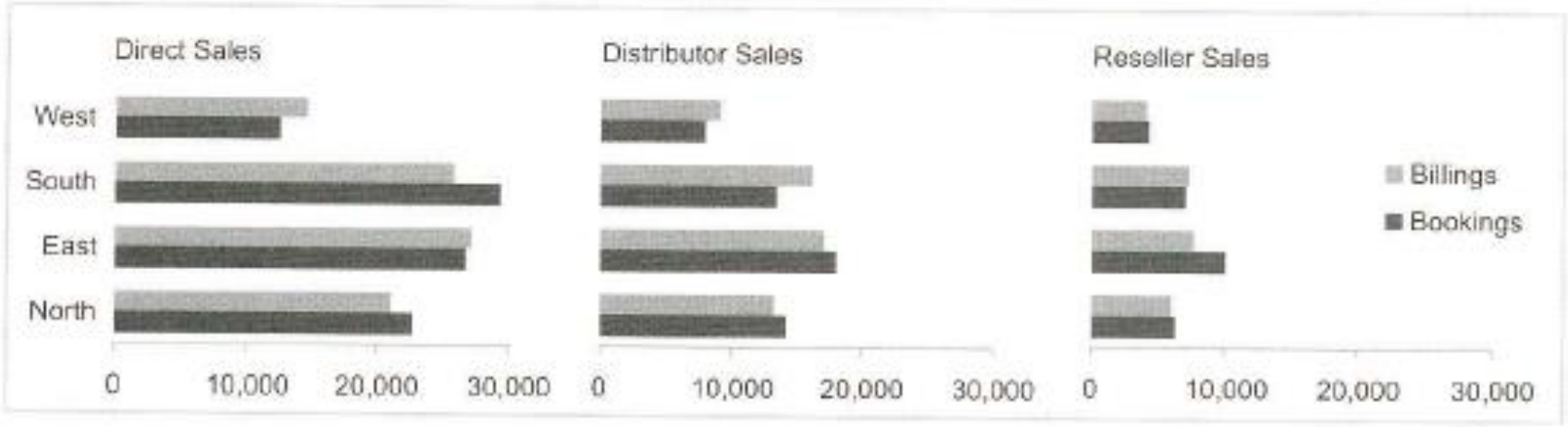
CRAP- Proximity



S. Few, Show Me The Numbers

Data Visualization: Design- Elements

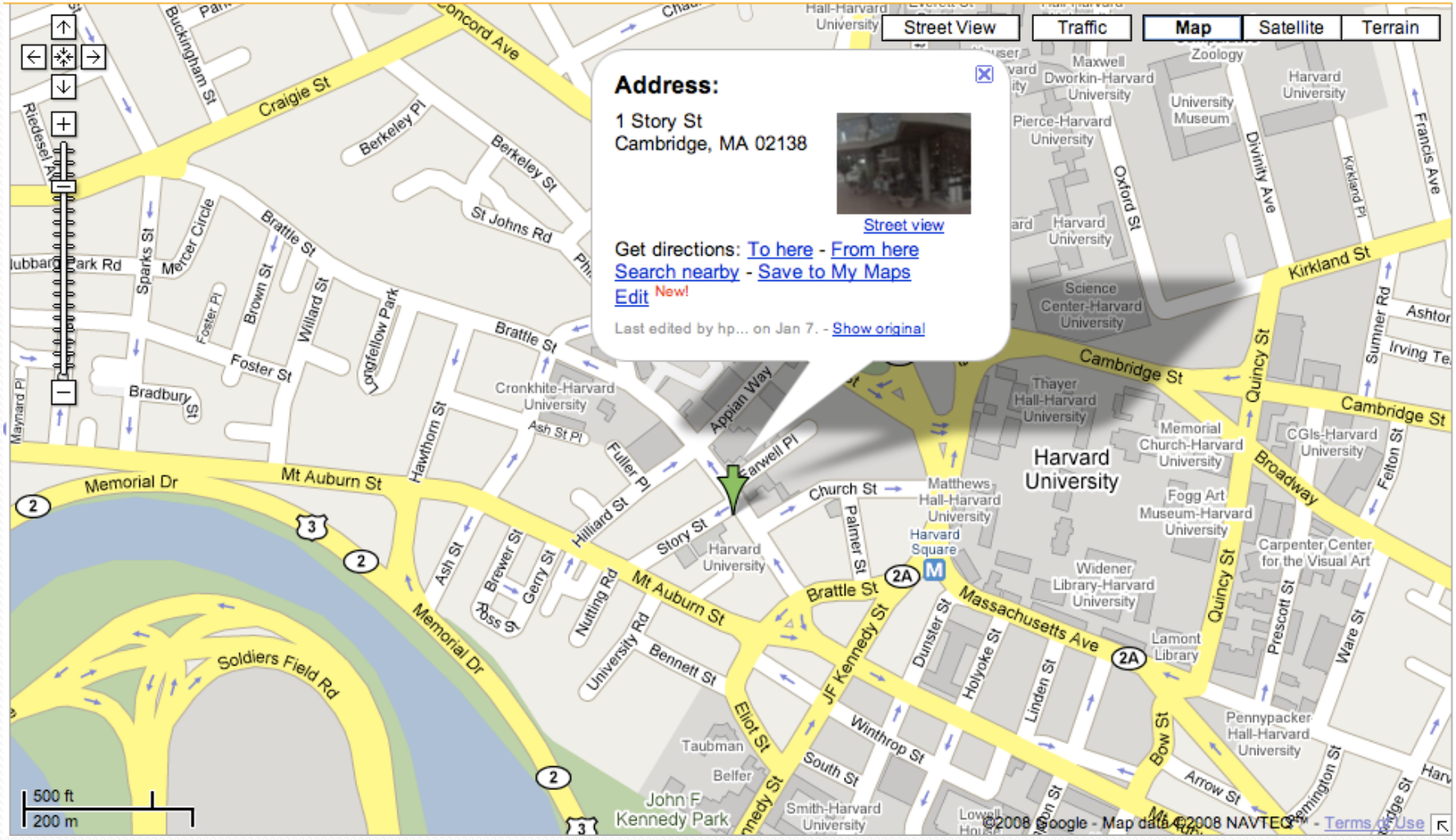
Small Multiples



S. Few, Show Me the Numbers, p. 220

Data Visualization: Design- Elements

Layering and Separation



Tufte, VDQI (Vol. 1) p. 174

Data Visualization: Design Περίληψη

Graphical Integrity

- Missing scales
- Lie factor
- Distortions

Design Principles

- Hierarchical Display
- Context
- Maximize Data-Ink Ratio
- Avoid Chartjunk
- Subjective Dimensions

Design Elements

- CRAP
- Small Multiples
- Layering & Separation
- Negative Space

Data Visualization: Τι περιλαμβάνει

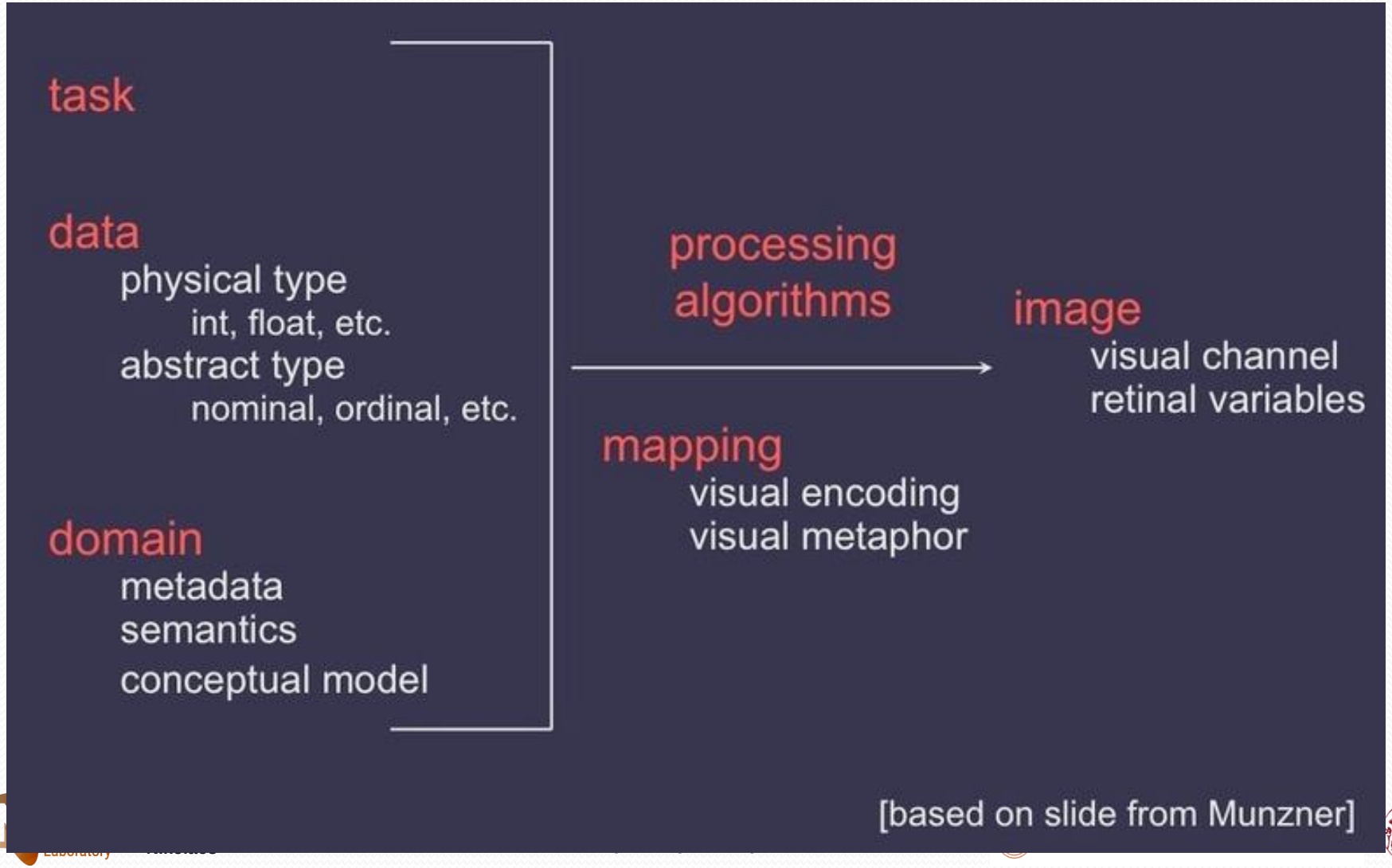
- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Process

Data Visualization: Process

[Maneesh Agrawala]
[Tamara Munzner]

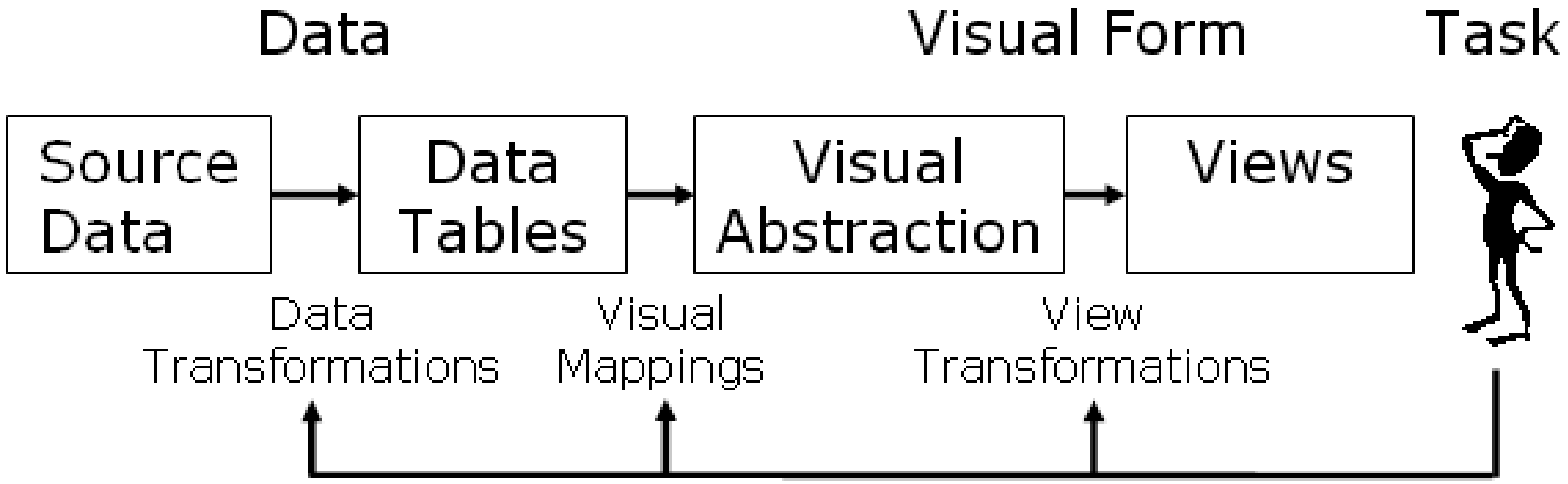
Pipeline Model



Data Visualization: Process

[J. Heer, Prefuse]

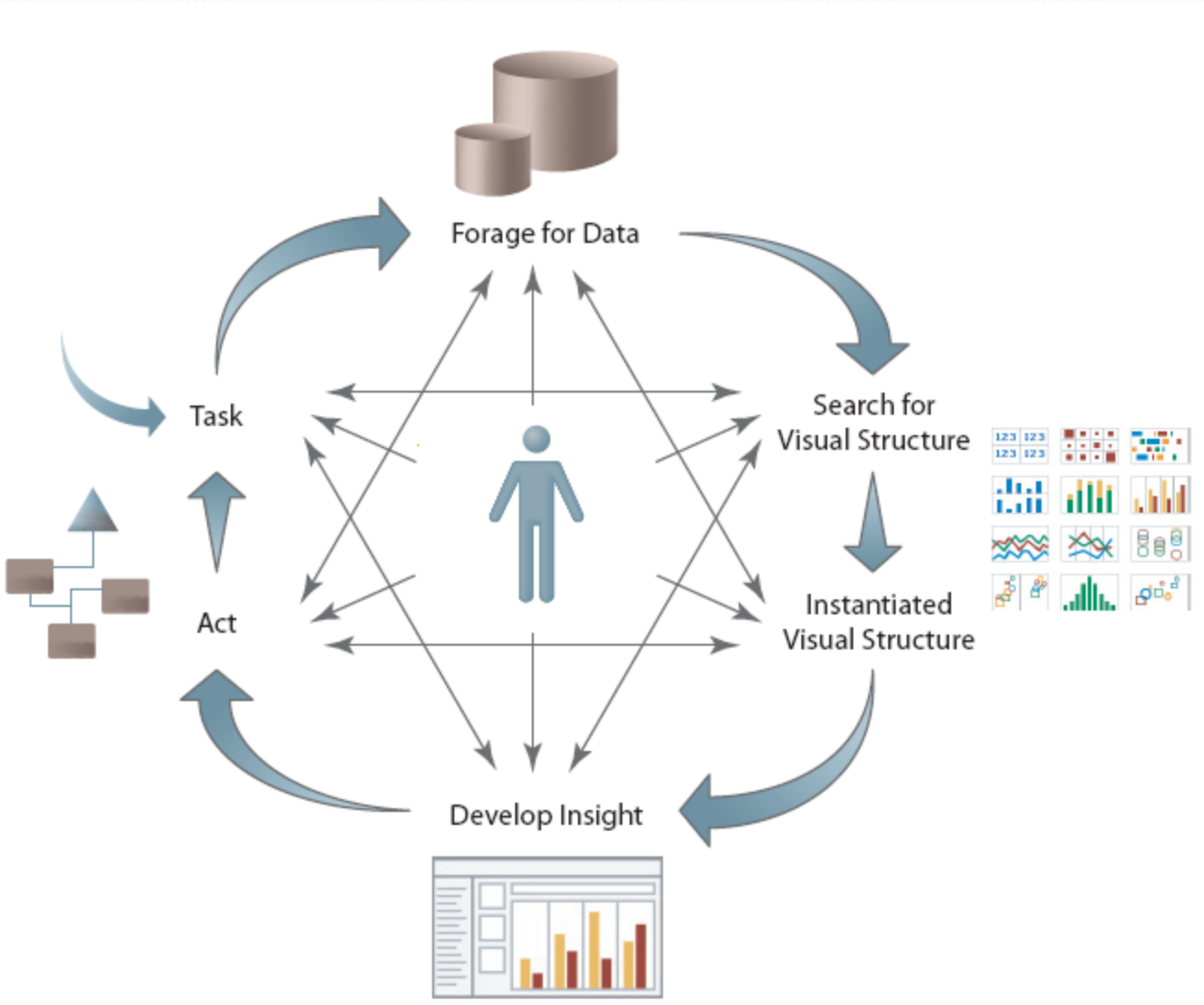
Pipeline Model



Data Visualization: Process

Cyclical Model

[C. Stolte, Tableau]



Data Visualization: Process

Cyclical Model

[J. van Wijk, The Value of Visualization]

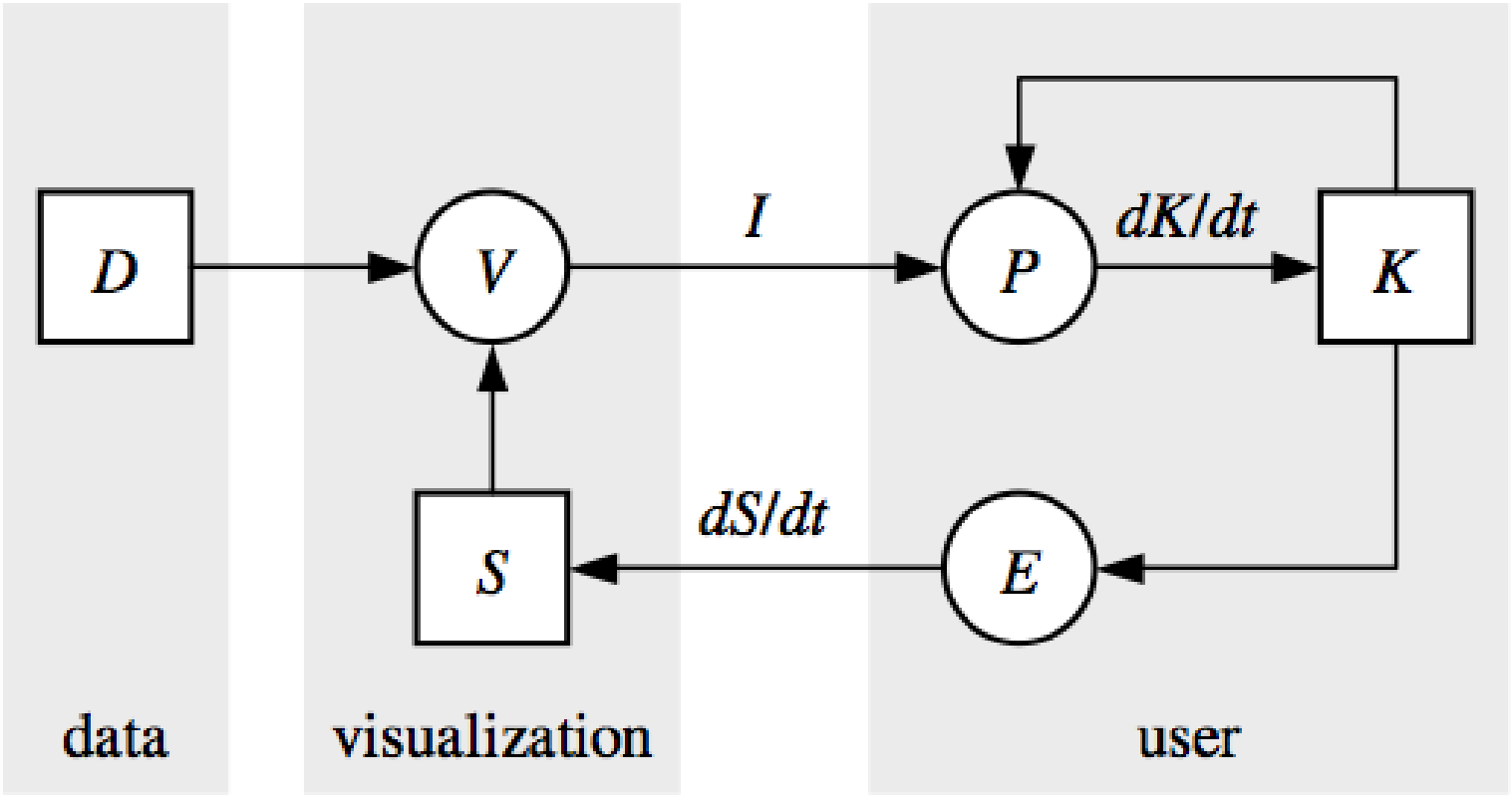
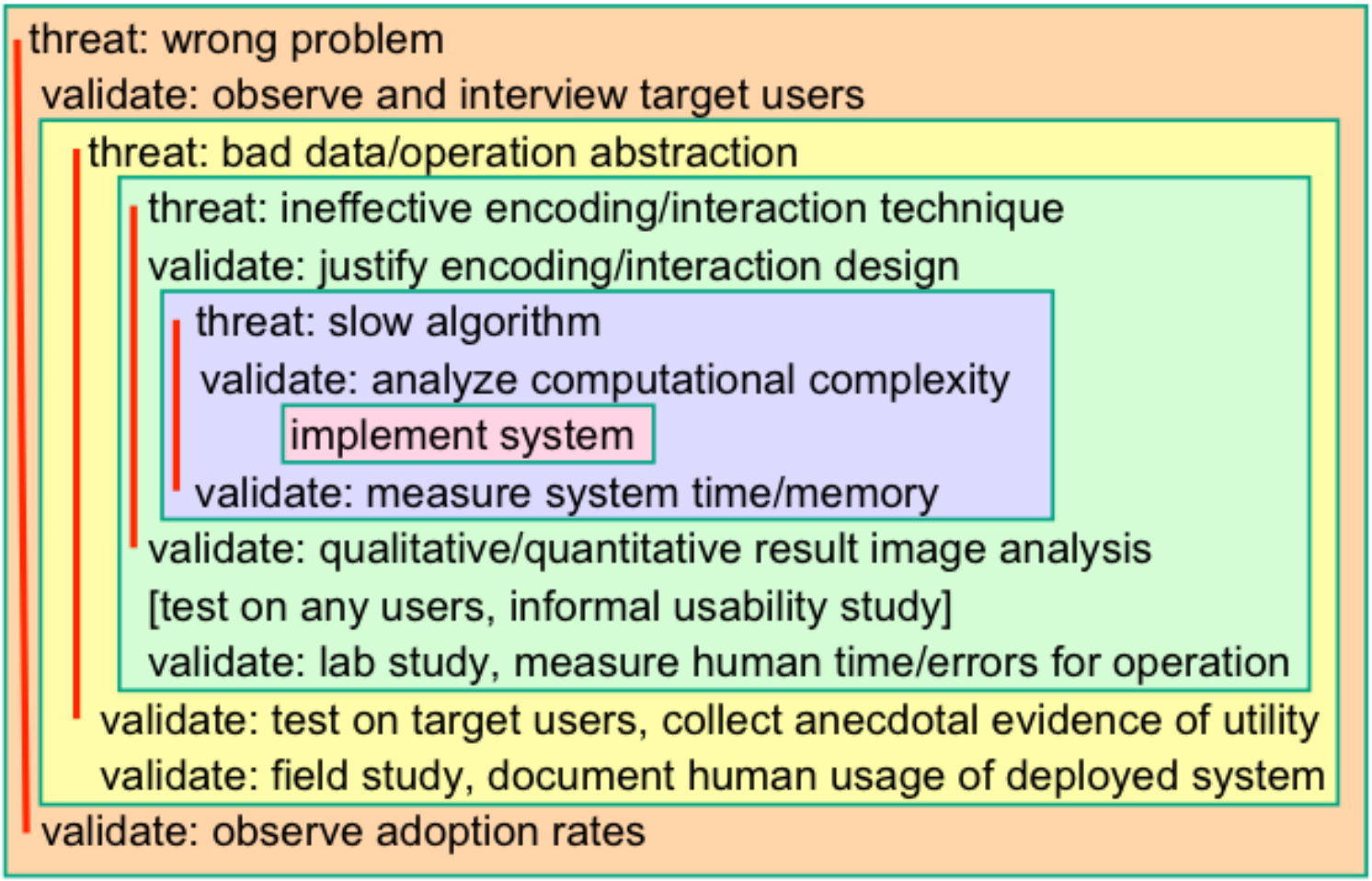


Figure 1: A simple model of visualization

Data Visualization: Process

Nested Model

[T. Munzner, A Nested Model for Visualization Design and Validation]



Data Visualization: Process

Nested Model

- What is the problem?

  Understand the questions & domain concepts

- What are the data and tasks?

  Abstract questions into data types and operations

- How do I show it?

  Design visual encodings and interactions

- How do I implement it?

  Implement the system with efficient algorithms

Data Visualization: Process

What is a threat at the abstraction level?

1. You're showing them the wrong thing
2. Your code is too slow
3. They don't do that
4. Nobody is using your tool
5. The way you show it does not work

Data Visualization: Process

What can go wrong?

- Wrong problem - they don't do that
- Wrong abstraction - you're showing them the wrong thing
- Wrong encoding / interaction - the way you show it does not work
- Wrong algorithm - your code is too slow

Data Visualization: Process

What validation is useful to test your abstraction?

1. Wall-clock timings
2. Observe usage of deployed system
3. Quantitative image quality analysis
4. Wizard Of Oz testing
5. Observe adoption rates

Data Visualization: Process

Validation Approaches

- Observing and talking to your users
  - Ethnographic field studies, interviews
- Have users try your tool
  - Anecdotal evidence, lab study, long-term field study
- Justify your design based on known principles
  - Discussion in paper, heuristic evaluation
- Measure quantitatively
  - Human time/errors for operations, run time, memory

Data Visualization: Process - Task Abstraction

Table 1. Questions for the analysis of conserved syntenic data, with the scale and relationship addressed by each. The scales are: *g*, genome; *c*, chromosome; *b*, block; and *f*, feature. The relationships are: *p*, proximity/location; *z*, size; *o*, orientation; and *s*, similarity.

	question	*scale*				*relationship*			
		g	*c*	*b*	*f*	*p*	*z*	*o*	*s*
1	Which chromosomes share conserved blocks?	X				X			
2	For one chromosome, how many other chromosomes does it share blocks with?	X	X			X			
3	What is the density of coverage and where are the gaps on: chromosomes? blocks?	X	X	X		X			
4	Where are the blocks: on chromosomes? around a specific location on a chromosome?	X	X			X			
5	What are the sizes and locations of other genomic features near a block?		X			X	X		
6	How large are the blocks?		X				X		
7	Do neighboring blocks go to the same: chromosomes? relative location on a chromosome?	X	X			X			
8	Are the orientations matched or inverted for: block pairs? feature pairs?		X	X				X	
9	Do the orientations match for pairs of: neighboring blocks? features within a block?		X	X				X	
10	Are similarity scores alike: with respect to neighboring blocks? within a block?		X	X					X
11	Are the paired features within a block contiguous?			X		X			
12	How large is a feature relative to other genes within a block?			X			X		
13	What are the sizes, locations, and names of features within a block?			X		X	X		
14	What are the differences between individual nucleotides of feature pairs?				X				X

[Meyer et al., MizBee: A Multiscale Synteny Browser, 2009]

Data Visualization: Process

4 An Analytic Task Taxonomy

The ten tasks from the affinity diagramming analysis are:

- Retrieve Value
- Filter
- Compute Derived Value
- Find Extremum
- Sort
- Determine Range
- Characterize Distribution
- Find Anomalies
- Cluster
- Correlate

[Amar, Eagan, & Stasko, 2005]

Data Visualization: Process

Examples:
- Order the cars by weight.
- Rank the cereals by calories.

[Amar, Eagan, & Stasko, 2005]

1) **Filter**: Find data that satisfies conditions
2) **Find Extremum**: Find data with extreme values
3) **Sort**: Rank data according to some metric
4) **Determine Range**: Find span of data values
5) **Find Anomalies**: Find data with unexpected / extreme values

Data Visualization:

Design Exercise

Η ιδέα της άσκησης είναι από τις σημειώσεις του *Hanspeter Pfister*

CS 171: Visualization, Process & Validation, Hanspeter Pfister

Data Visualization:

Συνταγή Βασιλόπιτας

Συστατικά
250 gr. βούτυρο μαλακό ή γάλακτος
8 αυγά
900 gr. αλεύρι
100 gr. κονιάκ
500 gr. ζάχαρη
6 - 7 πορτοκάλια μόνο το χυμό
1 κουτ. γλ. σόδα
1 κουτ. γλ. ξύσμα πορτοκάλι
2 κουτ. γλ. μπέικιν πάουντερ
ζάχαρη άχνη

http://www.sintagespareas.gr/sintages/basilopita.html#ixzz1ldr2cMRb

Συστατικά
600 γρ. ζάχαρη
600 γρ. αλεύρι
300 γρ. βούτυρο
300 γρ. γάλα κρύο
6 αβγά
3/4 κ.σ. κρεμόριο
3/4 κ.γ. σόδα μαγειρική
1 φλυτζανάκι κονιάκ
1 βανίλια
ξύσμα πορτοκαλιού

http://www.sintagespareas.gr/sintages/basilopita-2.html#ixzz1ldrNO7ic

Συστατικά
1 1/2 κούπα μαργαρίνη σε θερμοκρασία δωματίου
2 1/2 κούπες ζάχαρη
6 χωρισμένα αβγά
4 1/2 κούπες φαρίναπ
2 κουταλάκια baking powder
1/4 κουταλάκι αλάτι
το ξύσμα & το χυμό ενός πορτοκαλιού
1 κούπα ψημένα & χοντροκομμένα φουντούκια
125 γρ. ψηφίδες κουβερτούρας ή σοκολάτας γάλακτος
άχνη ζάχαρη

http://www.sintagespareas.gr/sintages/basilopita-3.html#ixzz1ldrVZP2N

Συστατικά
1 φλιτζάνι τσαγιού φρέσκο βούτυρο ή ένα πακέτο βιτάμ
2 φλιτζάνια τσαγιού ζάχαρη
6 αυγά
4 κουταλιές σούπας κονιάκ
4 φλιτζάνια τσαγιού κόκκινη φαρίνα ΓΙΩΤΗΣ
1 φλιτζάνι τσαγιού γάλα ή χυμό πορτοκαλιού
Ξύσμα πορτοκαλιού και ζάχαρη άχνη

http://www.sintagespareas.gr/sintages/basilopita-4.html#ixzz1ldrdBXsM

Data Visualization:

Συνταγή Βασιλόπιτας

Ερωτήσεις

1. Δεδομένων κάποιων συγκεκριμένων συστατικών, ποια συνταγή μπορώ να εκτελέσω;
2. Ποια συνταγή είναι καταλληλότερη για δίαιτα;
3. Ποια συνταγή είναι οικονομικότερη;
4. Ποιες διαφορές θα έχουν οι βασιλόπιτες; Γεύση; Μορφή;
5. Σε ποιο βαθμό διαφέρουν οι ποσότητες; Τι απόκλιση υπάρχει;

Data Visualization:

Συνταγή Βασιλόπιτας

Οπτικοποίηση Συνταγής

- Διαλέξτε μια ερώτηση
- Ποιοι τύποι δεδομένων συμμετέχουν;
- Ποιες είναι οι δραστηριότητες (tasks);
- Ποια είναι τα ποιο αποτελεσματικά «visual encodings»;
- Φτιάξτε σκίτσο του σχεδίου σας
- Πώς θα αξιολογούσατε το σχέδιο σας;

Data Visualization:

Development Exercise

- Υλοποιείστε ένα visualization στο Netbeans για την σχεδίαση σας.

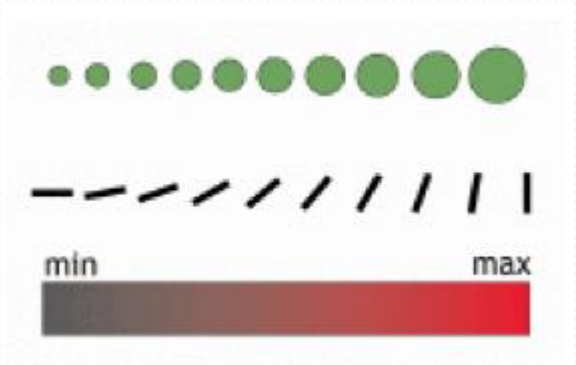
περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color

Models

- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Data Visualization: Data Model – Data Types



- Quantitative (Q)

  10 inches, 23 inches, etc.

- Ordinal (ordered) (O)

  Small, medium, large

- Nominal (categorical) (N)

  Apples, Oranges, Bananas, …

Data Visualization: Data Model – Data Types

Quantitative

- Q - Interval (location of zero arbitrary)

  Dates: Jan 19; Location: (Lat, Long)

  Only differences (i.e., intervals) can be compared

- Q - Ratio (zero fixed)

  Measurements: Length, Mass, Temp, …

  Origin is meaningful, can measure ratios & proportions

Data Visualization: Data Model – Data Types

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
32	7/16/07	2-High	Medium Box	0.6	7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
166	9/12/07	2-High	Small Box	0.55	9/14/07
193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

Record

Data Visualization: Data Model – Data Types

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
32	7/16/07	2-High	Medium Box	0.6	7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
166	9/12/07	2-High	Small Box	0.55	9/14/07
193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

Field

Data Visualization: Data Model – Data Types

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
32	7/16/07	2-High	Medium Box	0.6	7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69	6/4/05	4-Not Specified		0.44	6/6/05
69	6/4/05	4-Not Specified		0.6	6/6/05
70	12/18/06	5-Low		0.59	12/23/06
70	12/18/06	5-Low		0.82	12/23/06
96	4/17/05	2-High		0.55	4/19/05
97	1/29/06	3-Medium		0.38	1/30/06
129	11/19/08	5-Low		0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08			0.82	5/3/08
135	10/21/07			0.64	10/23/07
166	9/12/07			0.55	9/14/07
193	8/8/06			0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

1 = Quantitative

2 = Nominal

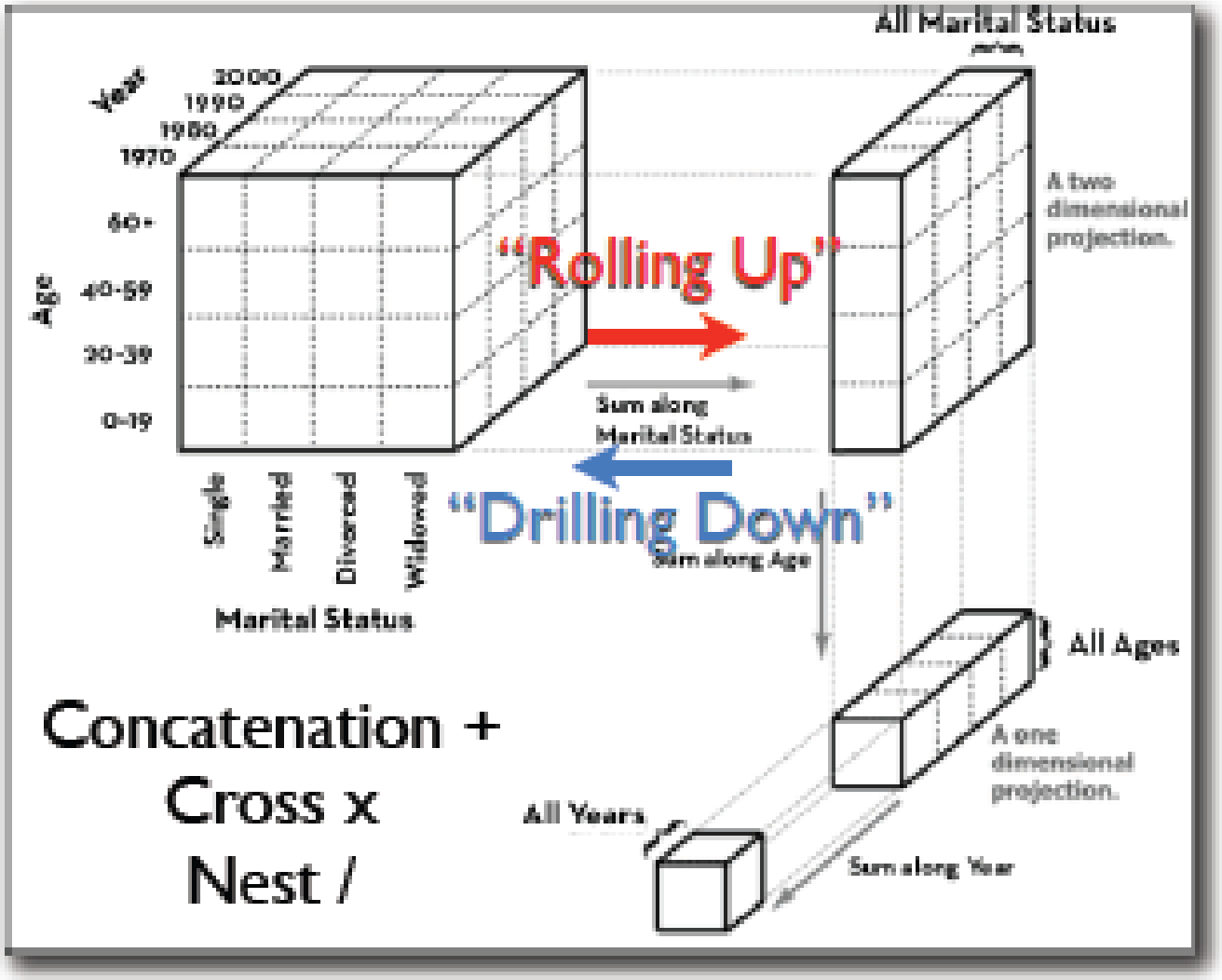
3 = Ordinal

Quantitative = Measures

Nominal /Ordinal = Dimensions

Data Visualization: Data Model – Data Types

Relational Data Cubes



Data Visualization: Data Model

Data vs. Conceptual Models

- Data Model: Low-level description of the data

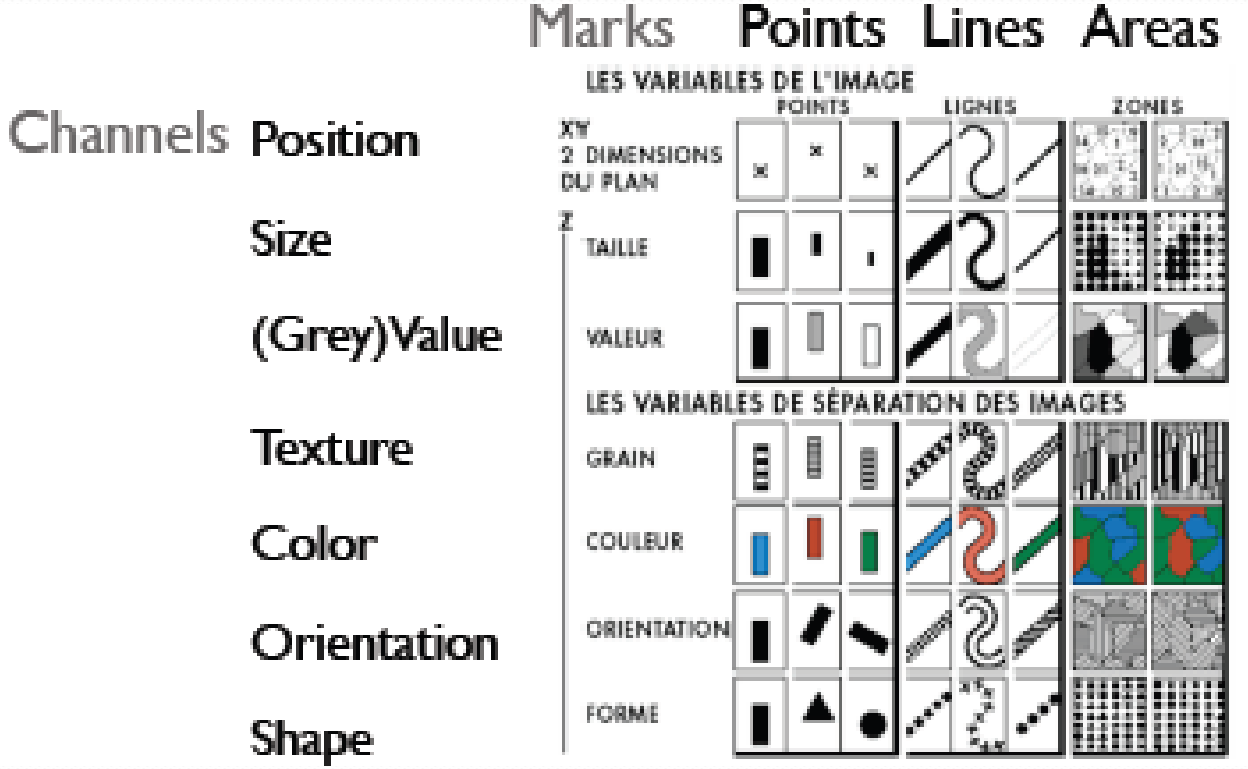
  Set with operations, e.g., floats with +, -, /, *

- Conceptual Model: Mental construction

  Includes sen ning

Data	Conceptual
1D floats	temperature
3D vector of floats	space

Data Visualization: Image Model



Semiology of Graphics [J. Bertin, 67]

Data Visualization: Image Model

(Grey)Value is perceived as ordered (O)

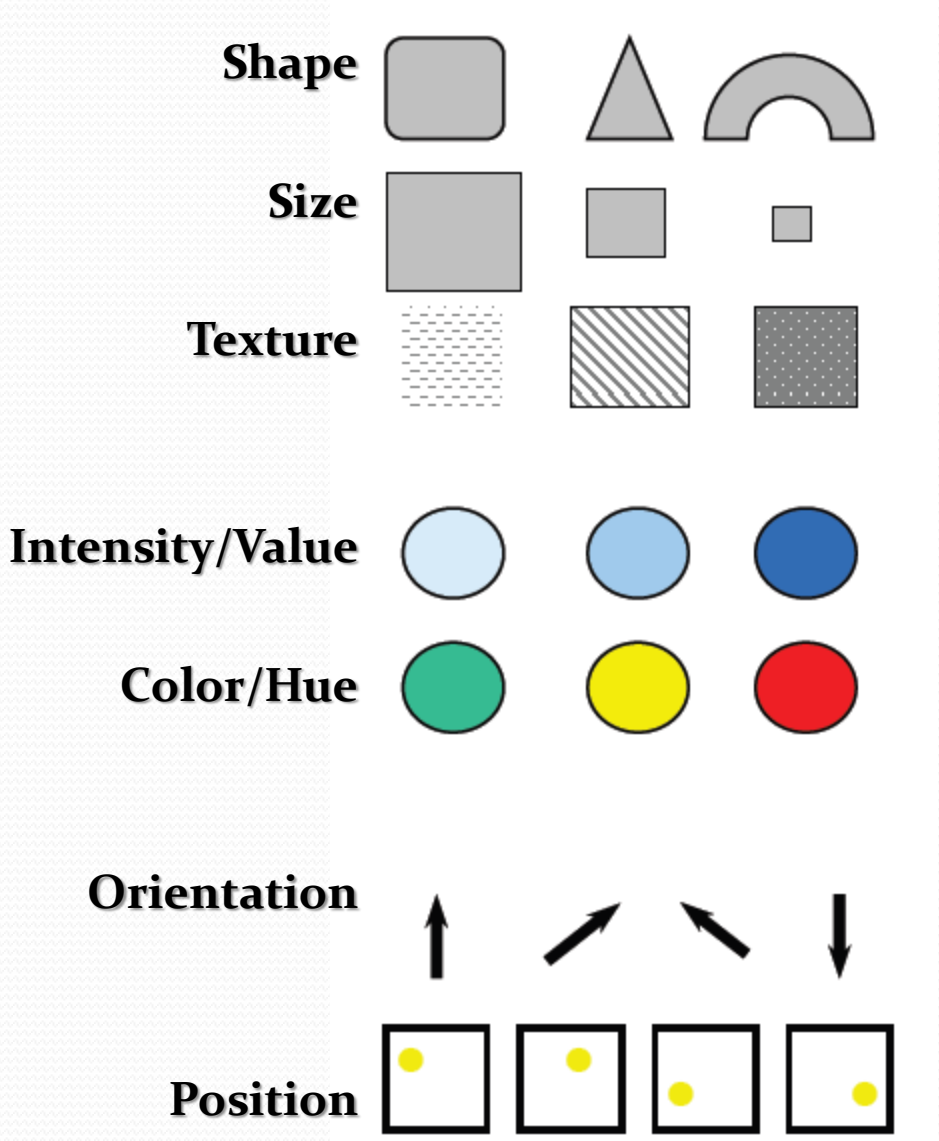
Can encode quantitative values (Q) [not as well]

Hue is normally perceived as unordered (N)

Data Visualization: Data Model

Bertin's Retinal Variables

Visualizable Parameters = Retinal Variables

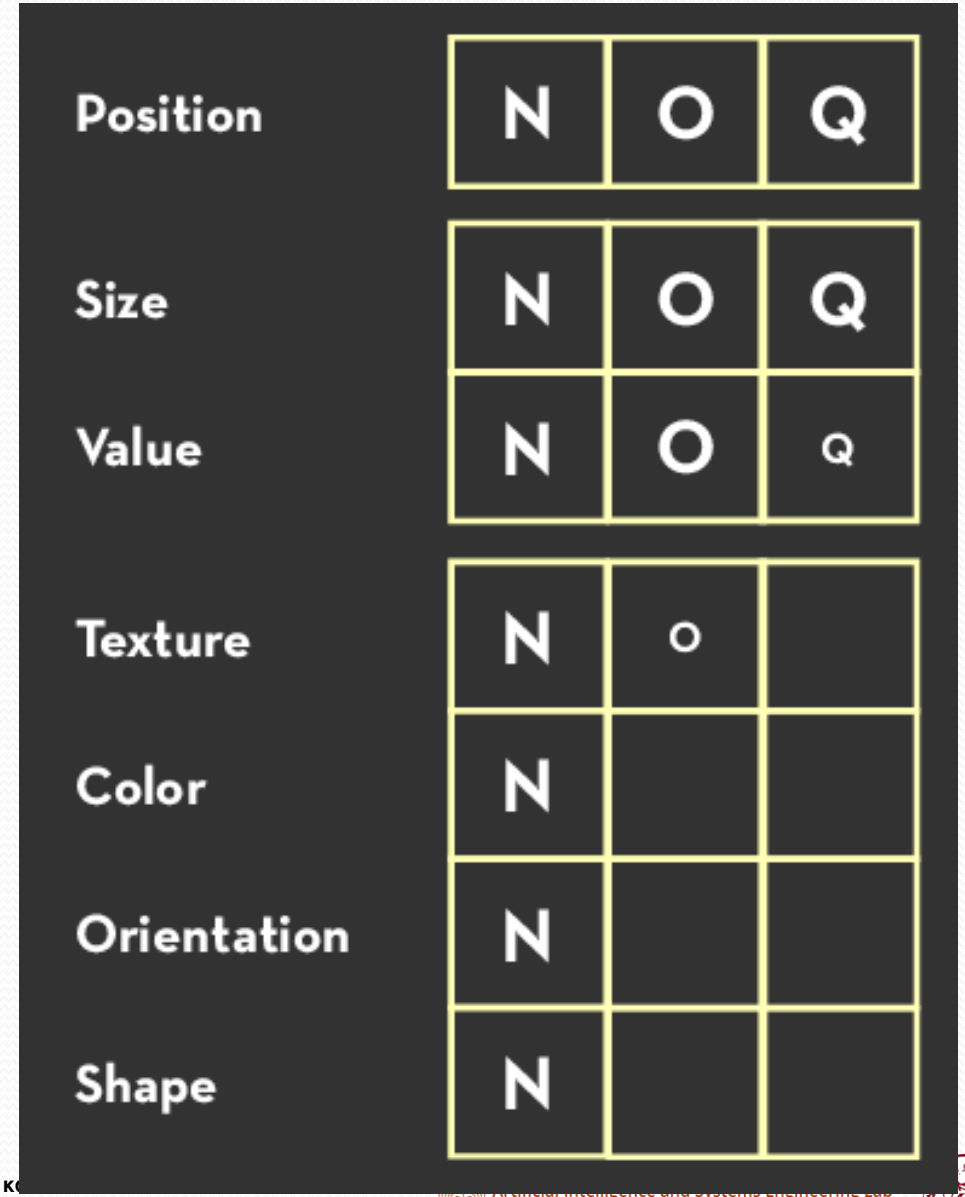


Bertin, J. (1983). Semiology of graphics. Madison, Wis.: University of Wisconsin Press.

Data Visualization: Data Model

Bertin's Retinal Variables

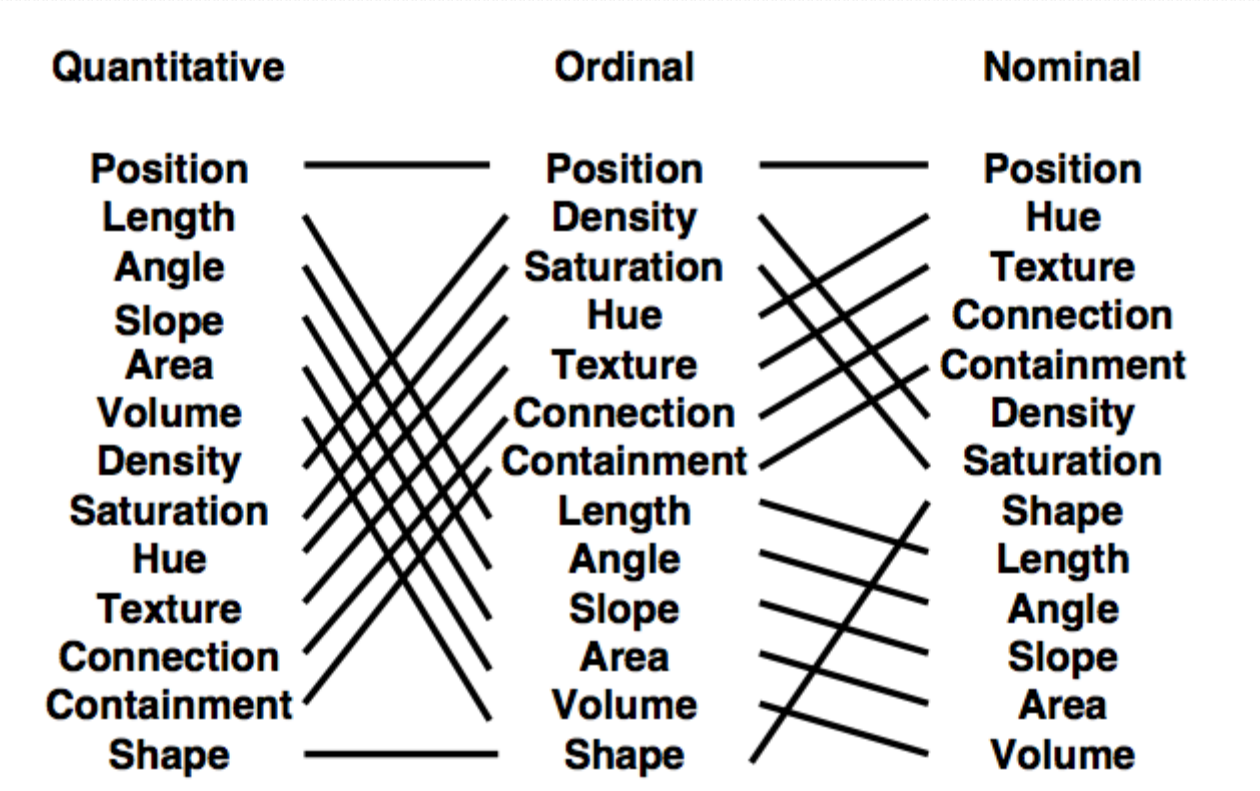
Visualizable Parameters = Retinal Variables



Bertin, J. (1983). Semiology of graphics. Madison, Wis.: University of Wisconsin Press.

Data Visualization: Data Model

Mackinlay's Retinal Variables



[**Mackinlay,** Automating the Design of Graphical Presentations of Relational Information, ACM TOG 5:2, 1986]

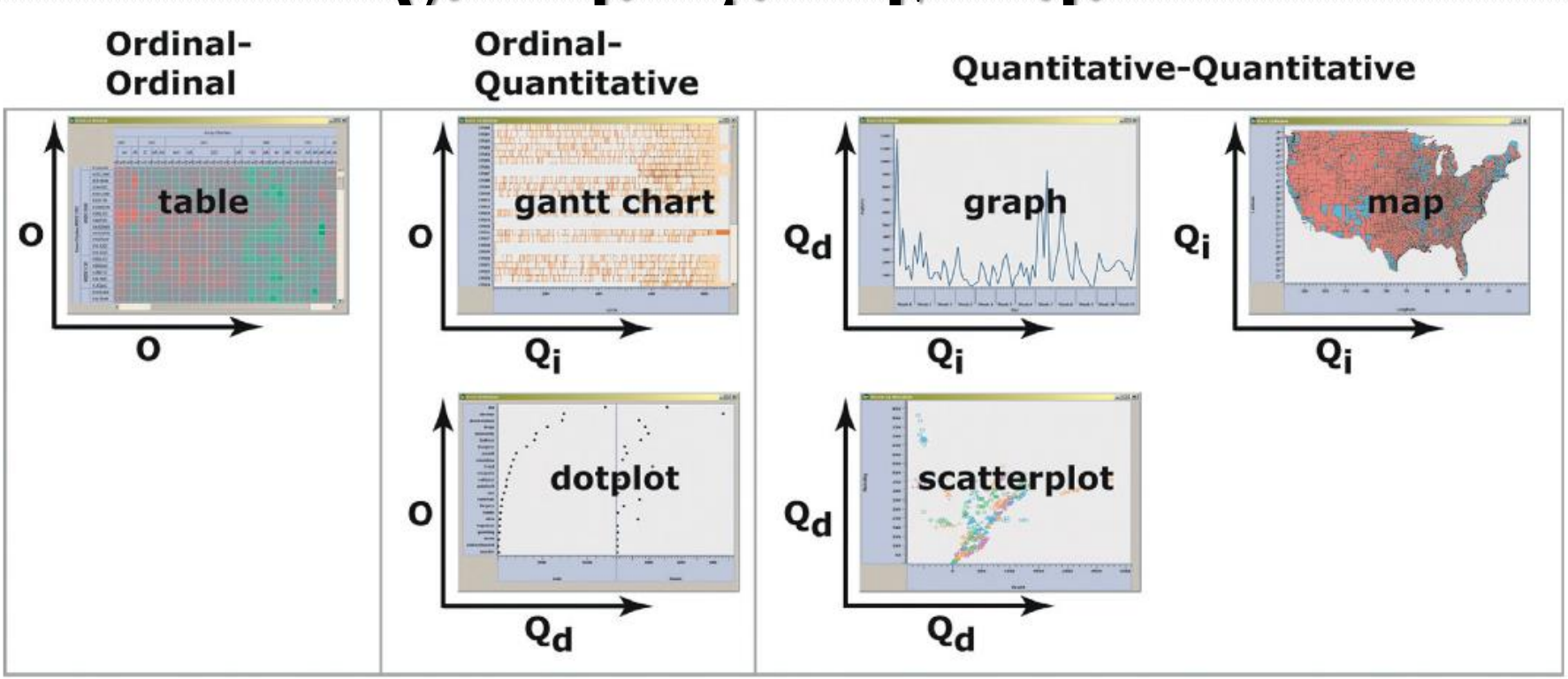
Data Visualization: Data Model

Tableau's Retinal Variables

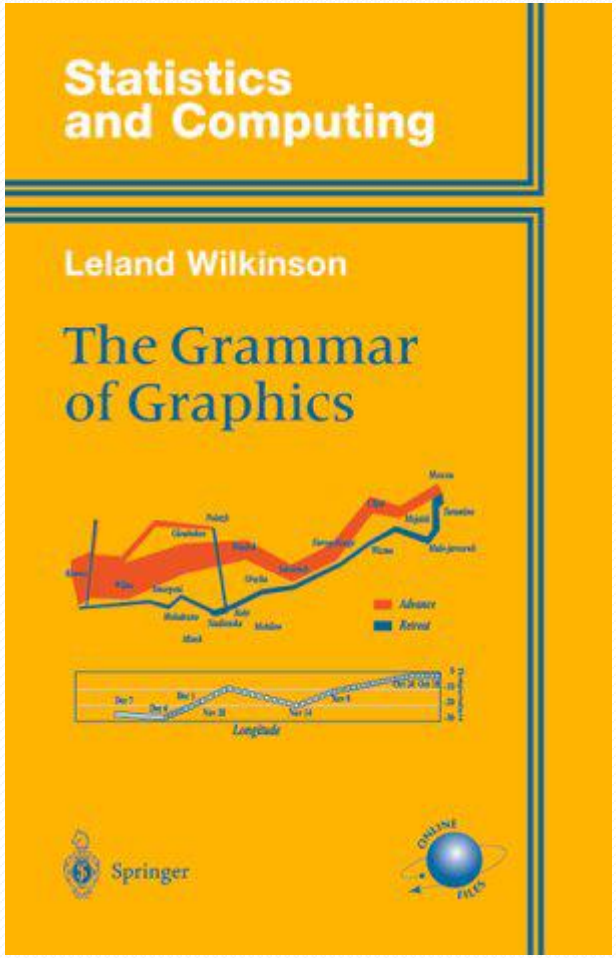
property	marks	ordinal/nominal mapping	quantitative mapping
shape	glyph	○ □ + △ S U	
size	rectangle, circle, glyph, text		
orientation	rectangle, line, text		
color	rectangle, circle, line, glyph, y-bar, x-bar, text, gantt bar		

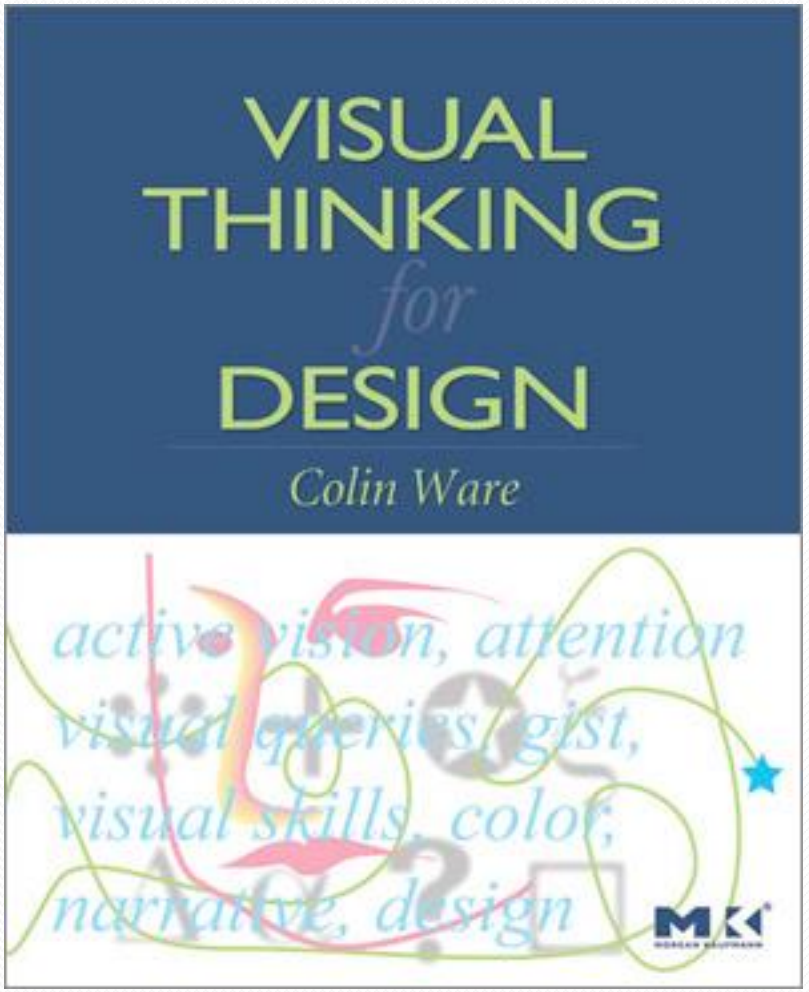
["**Polaris:** A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases" Chris Stolte, Diane Tang, and Pat Hanrahan, 2002]

Data Visualization: Data Model



Data Visualization: Data Model Books





περιλαμβάνει

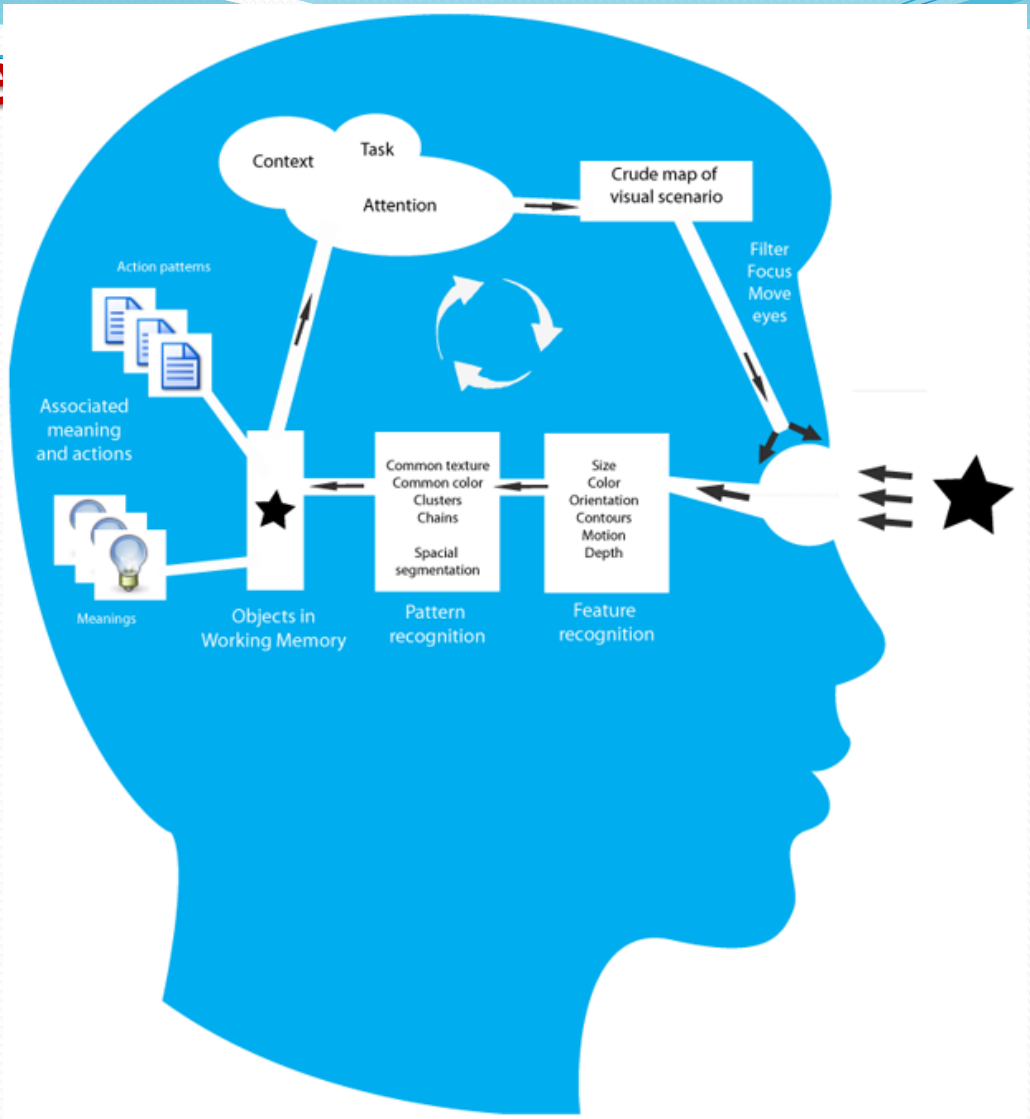
- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Perception

Data Visualization: Perce

"Visualization is really about external cognition, that is, how resources outside the mind can be used to boost the cognitive capabilities of the mind."

Stuart Card



A primitive model of visual perception

Data Visualization: Perception

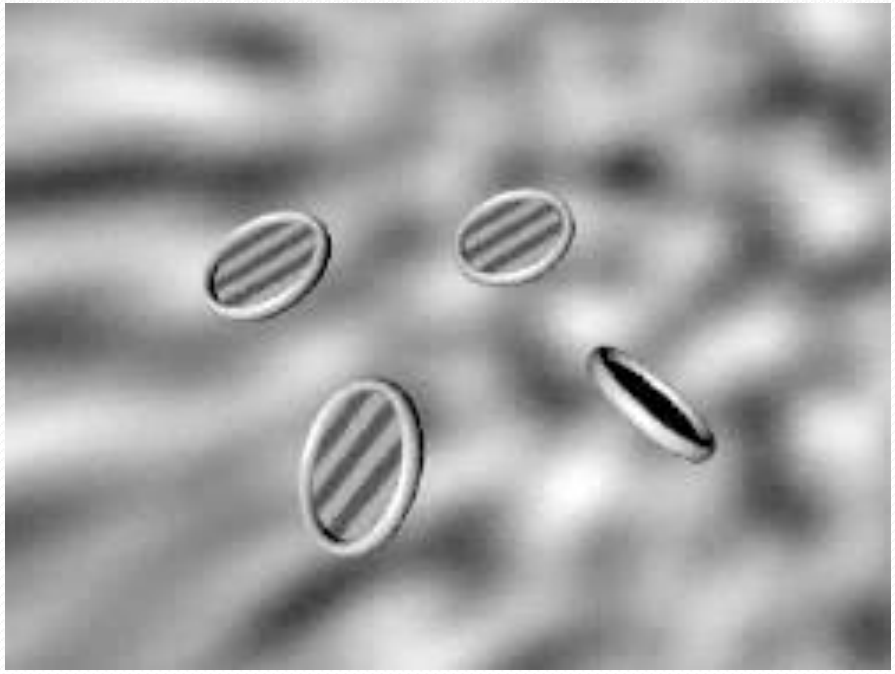
Fact #1

Memory plays an important role in human cognition, but working memory is extremely limited

Visualization must serve as an external aid to augment working memory

Data Visualization: Perception

- Which disk rotates differently?



Fact #2

To see an object change, it is necessary to attend to it

Make changes visible in visualizations to reduce the cognitive load

Data Visualization: Perception

- Simultaneous Contrast

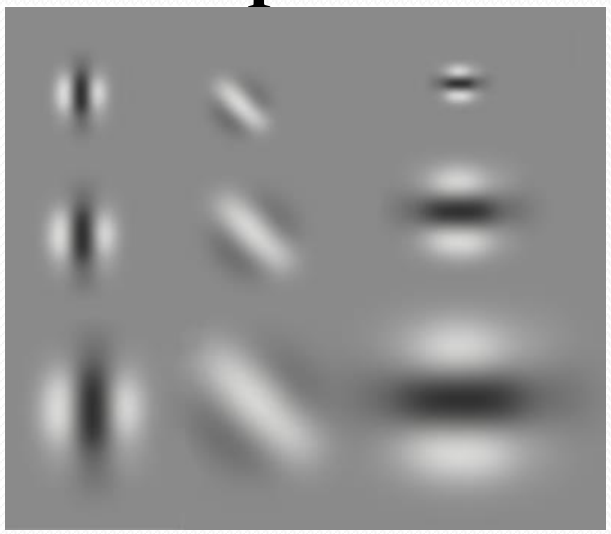
Fact #3

Our visual system sees differences, not absolute

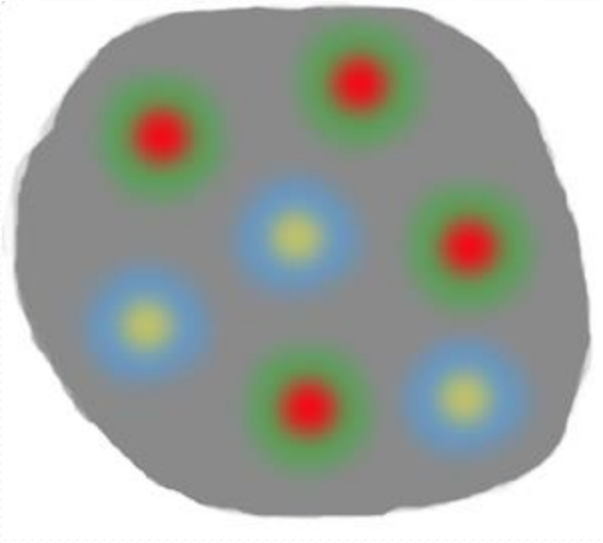
Use high contrast between objects that should be distinguishable

Data Visualization: Perception

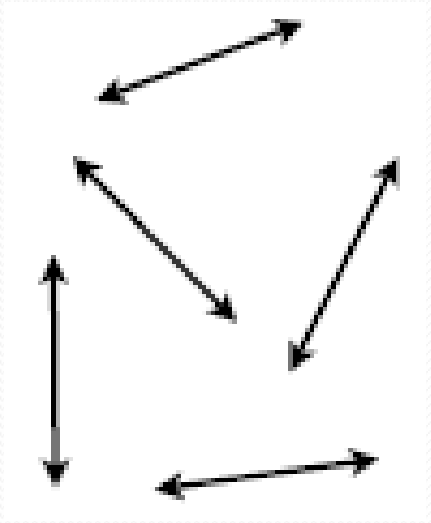
Receptive Fields



Shape



Color Differences



Motion

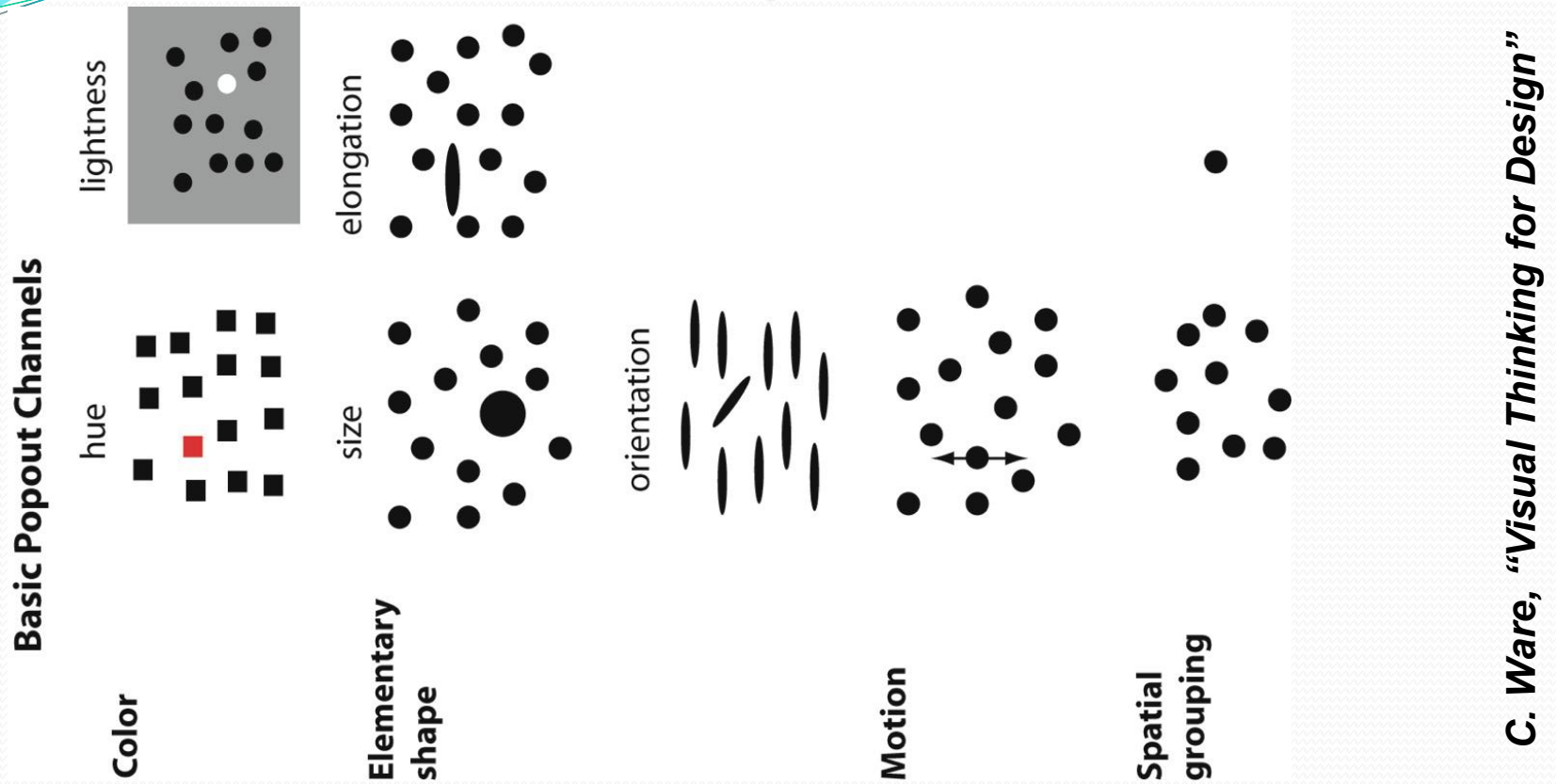
C. Ware, "Visual Thinking for Design"

Fact #4

The visual system has different channels for shape, color, and motion

Use different visual channels to separate pieces of information

Data Visualization: Perception



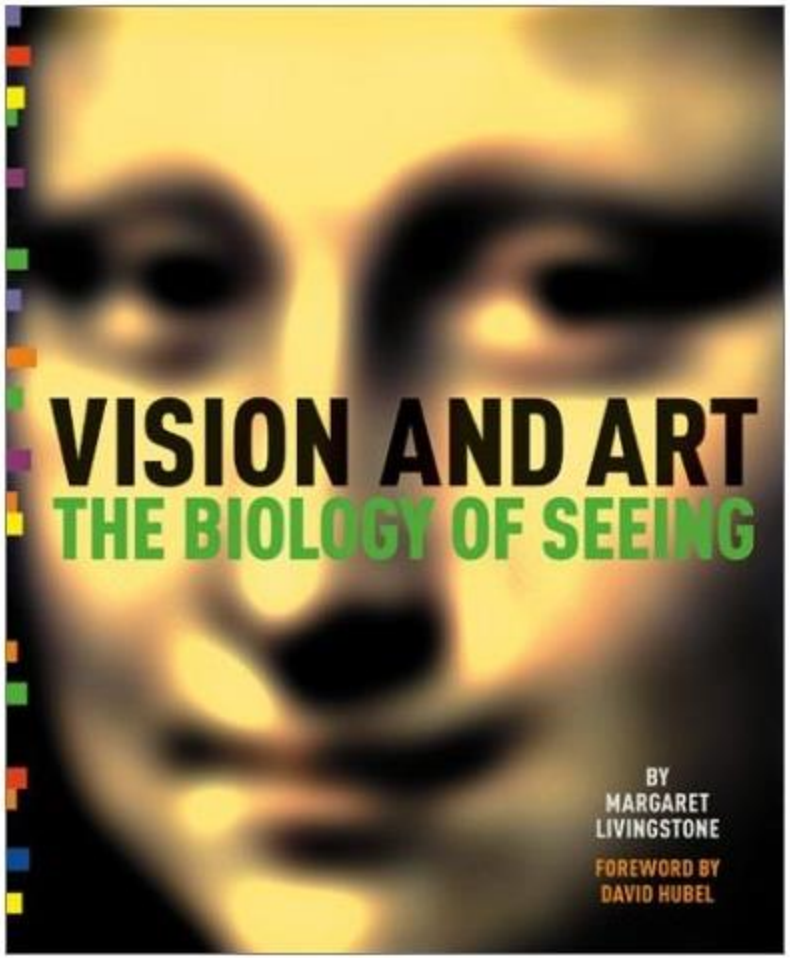
Fact #5

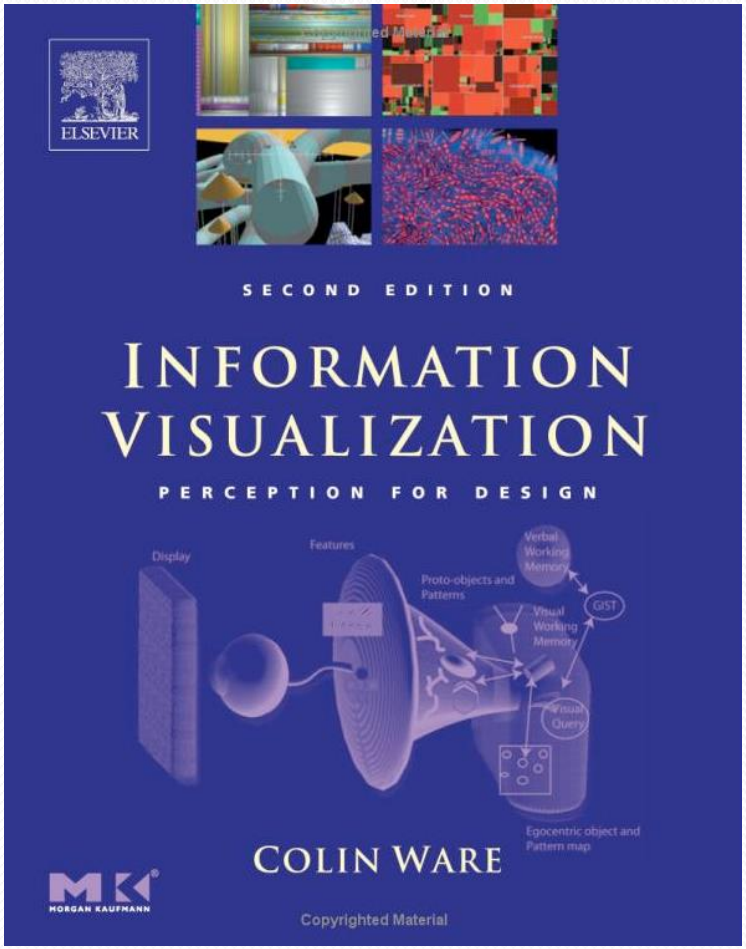
Visual perception is selective, as awareness of everything would overwhelm us

We must encode data in ways that allow what's interesting and important to stand out

Data Visualization: Perception

Books – Further reading





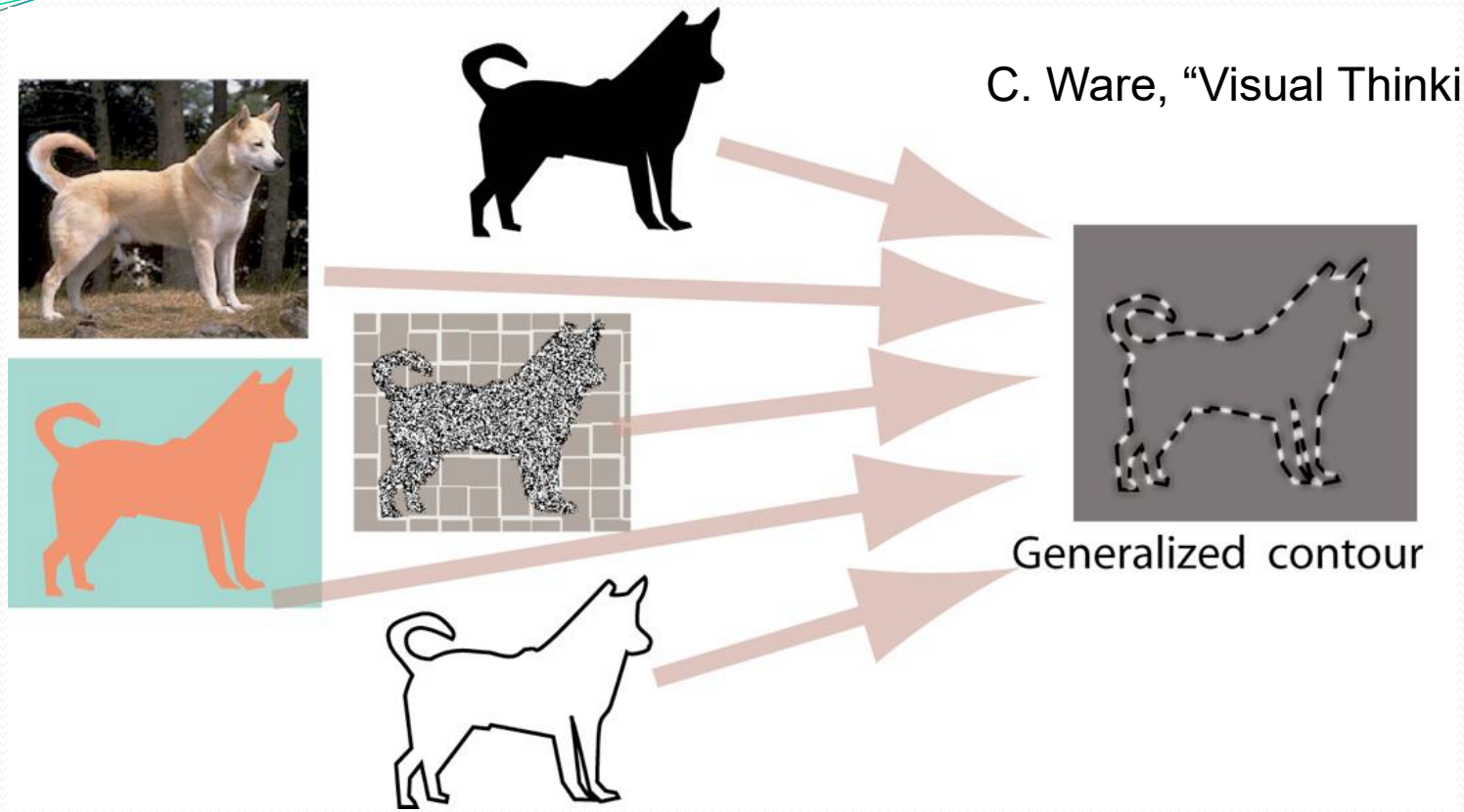
περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color

Patterns

- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Data Visualization: Patterns

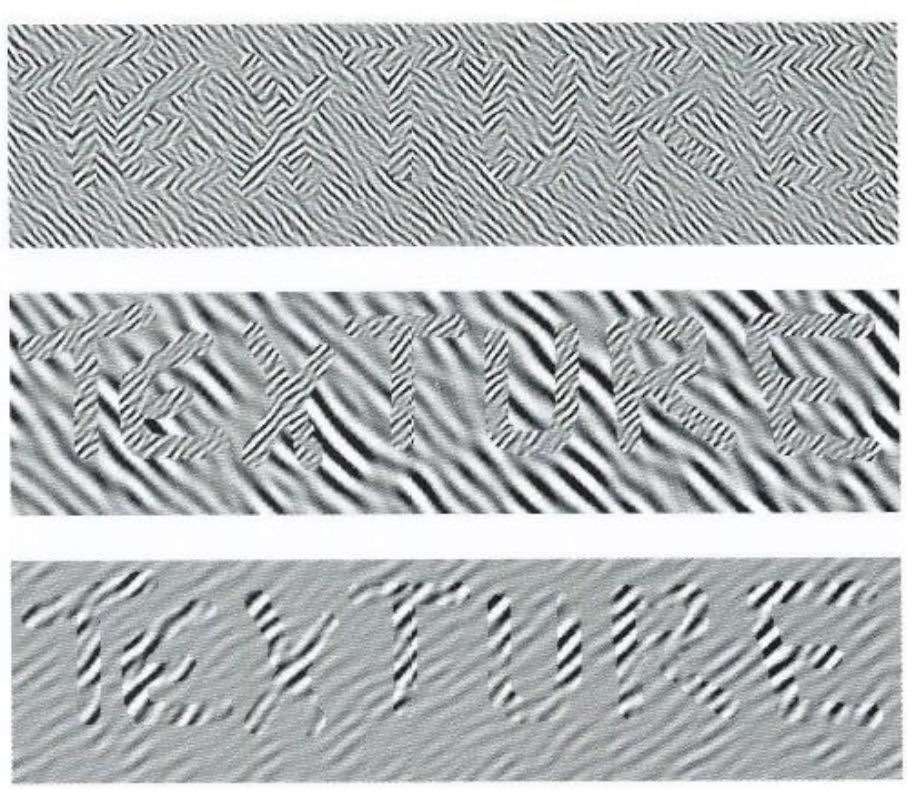


Fact #6

The brain constructs surface color based largely on edge contrast information

To make areas of interest stand out adjust the background to make the edges more distinct.

Data Visualization: Patterns



Orientation

Orientation & Size

Contrast

C. Ware, "Information Visualization"

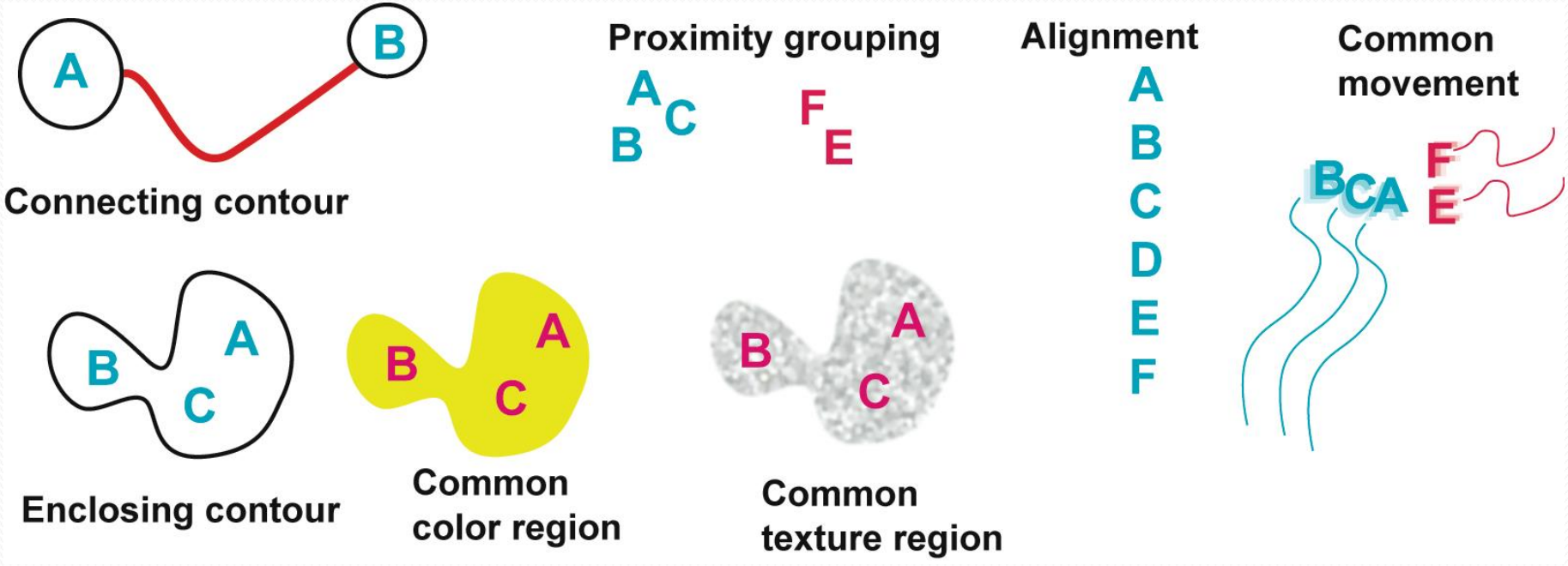
Fact #7

Texture information depends on the size, orientation, and frequency of the features.

To make textured areas of interest distinct adjust the orientation and frequency appropriately.

Data Visualization: Patterns

Gestalt Principles



Fact #8

Gestalt principles are powerful and influence the way we see patterns.

Use the Gestalt principles to arrange your visual patterns to maximize perception.

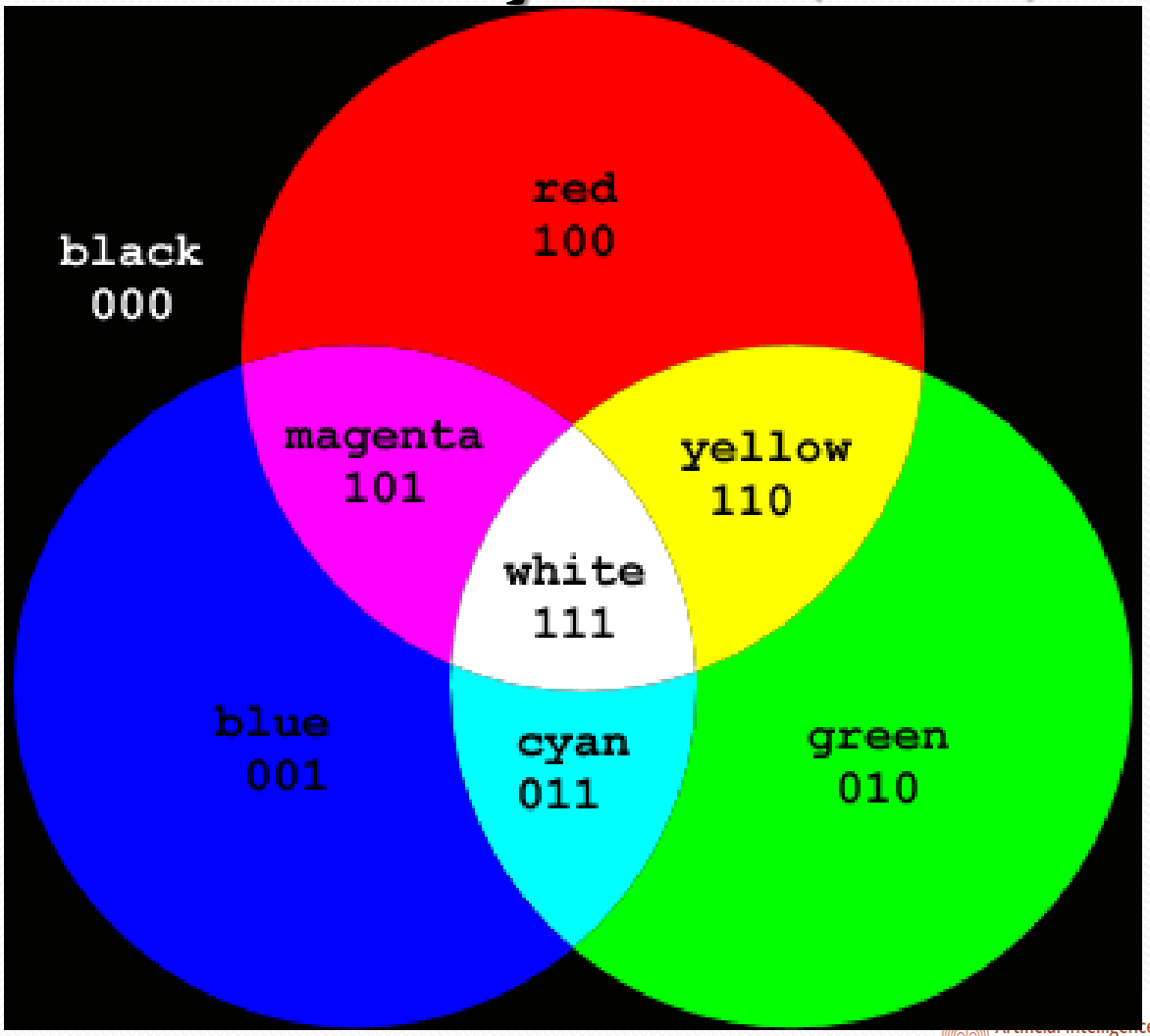
περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Color

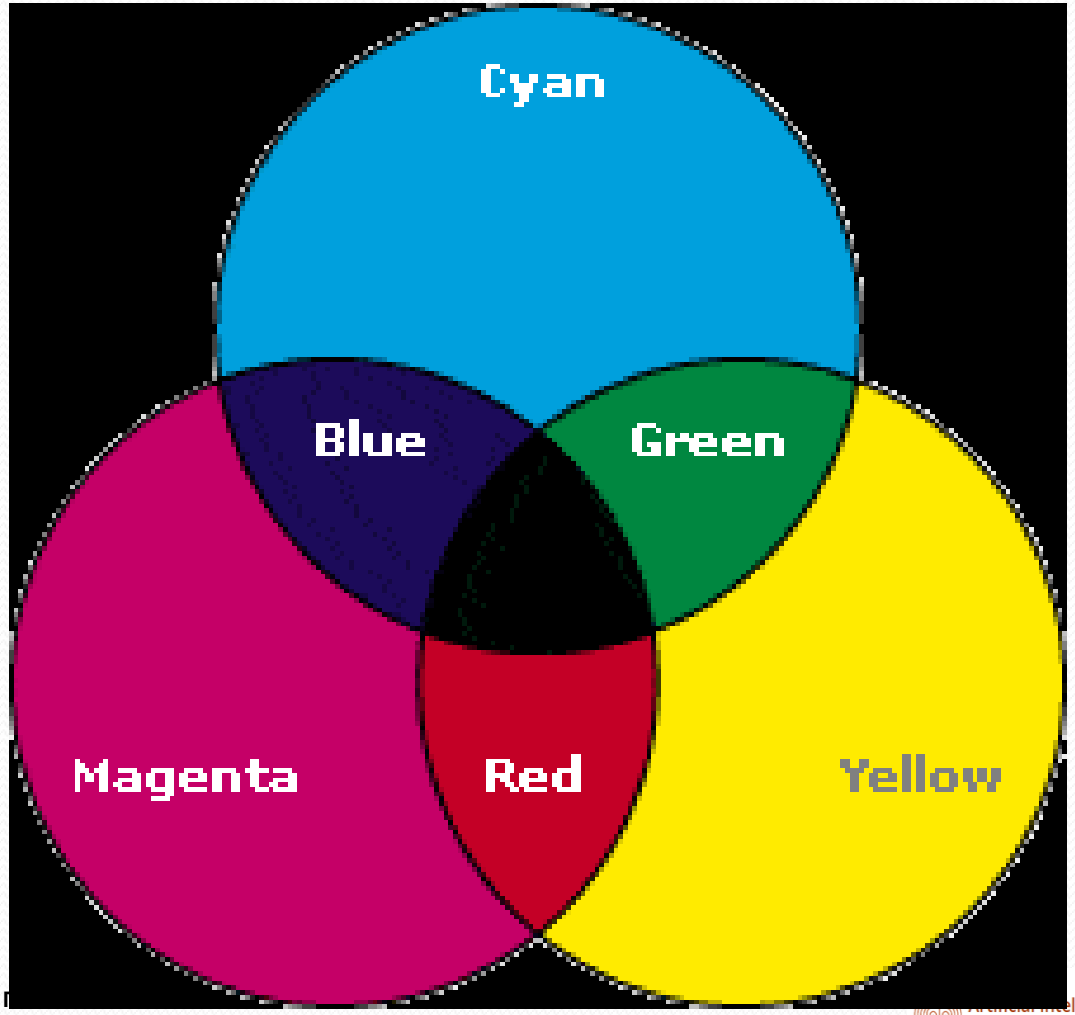
Data Visualization: Color

Additive System (RGB)



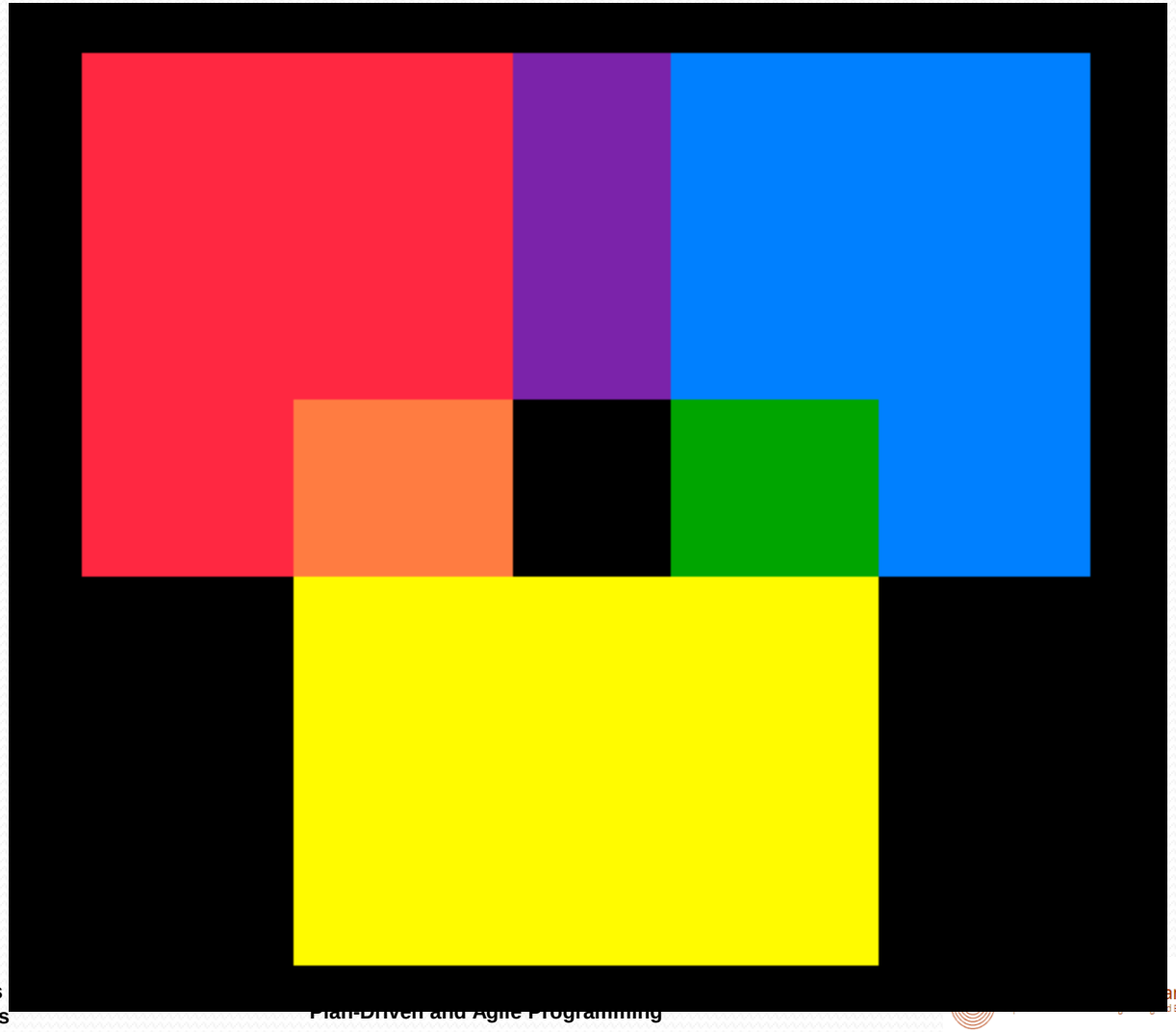
Data Visualization: Color

Subtractive System (CMYK)



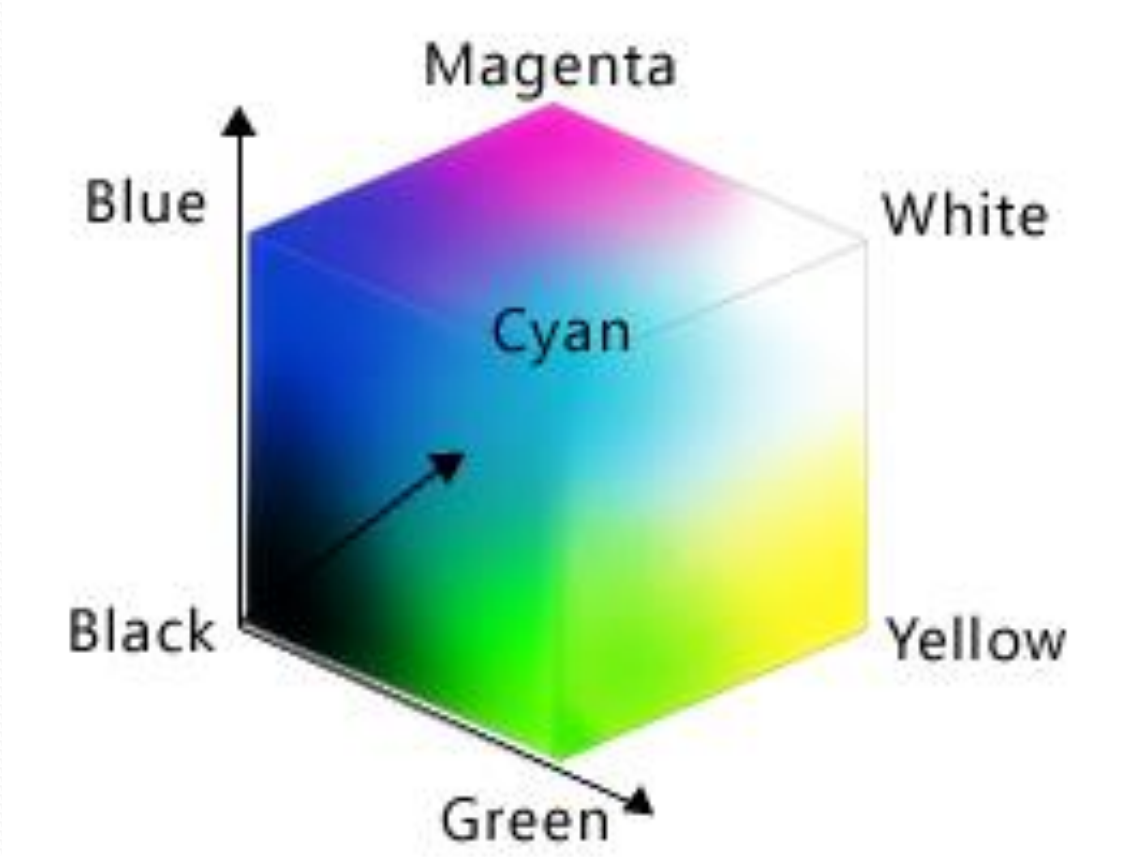
Data Visualization: Color

Subtractive System (RYB)



Data Visualization: Color -Models

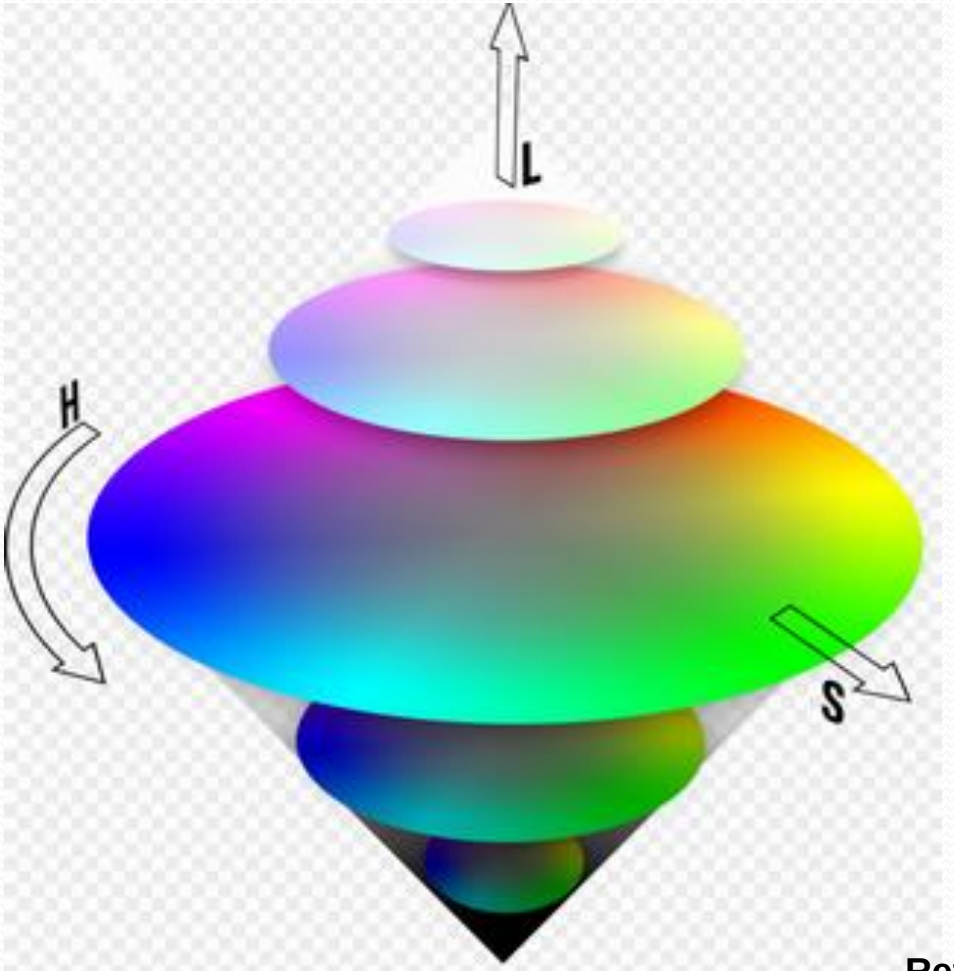
RGB Color Space



Reference: http://msdn.microsoft.com/en-us/library/aa511283.aspx

Data Visualization: Color -Models

HSL Color Space



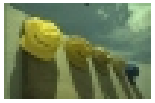
Reference: wikipedia.org

Data Visualization: Color - Blindness

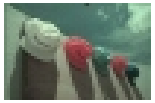
www.vischeck.com

Try Vischeck on a Webpage

Select the type of color vision to simulate:


(•) Deuteranope (a form of red/green color deficit)


( ) Protanope (another form of red/green color deficit)


( ) Tritanope (a blue/yellow deficit- very rare)

Enter the URL of any webpage- eg. **www.google.com**.

URL: [ ] Run Vischeck!

Data Visualization: Color - Design

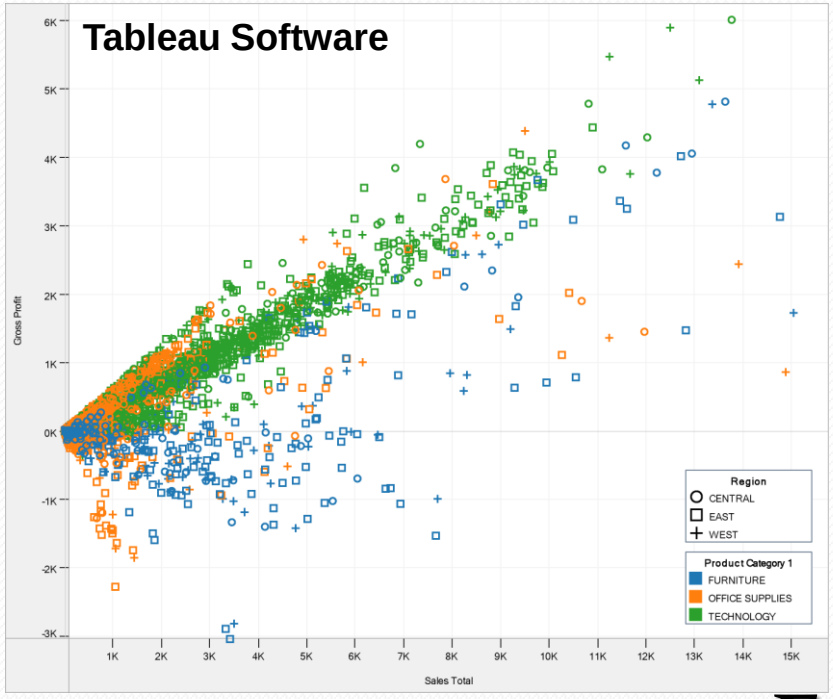
"… avoiding catastrophe becomes the first principle in bringing color to information: *Above all, do no harm*."

E. R. Tufte

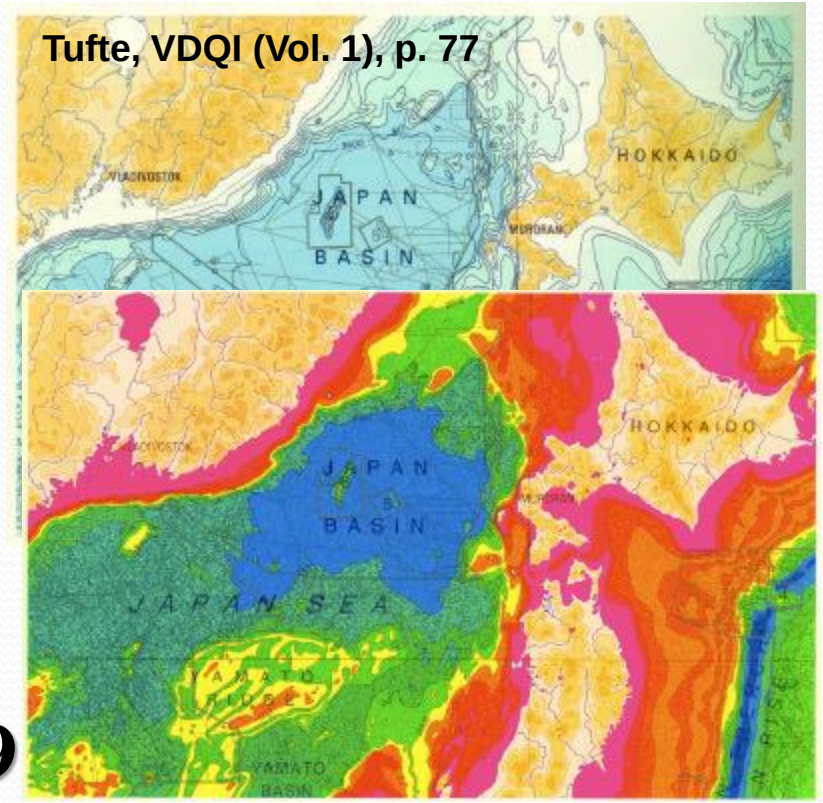
Tufte, "Envisioning Information"

Data Visualization: Color - Nominal

Small Areas



Large Areas

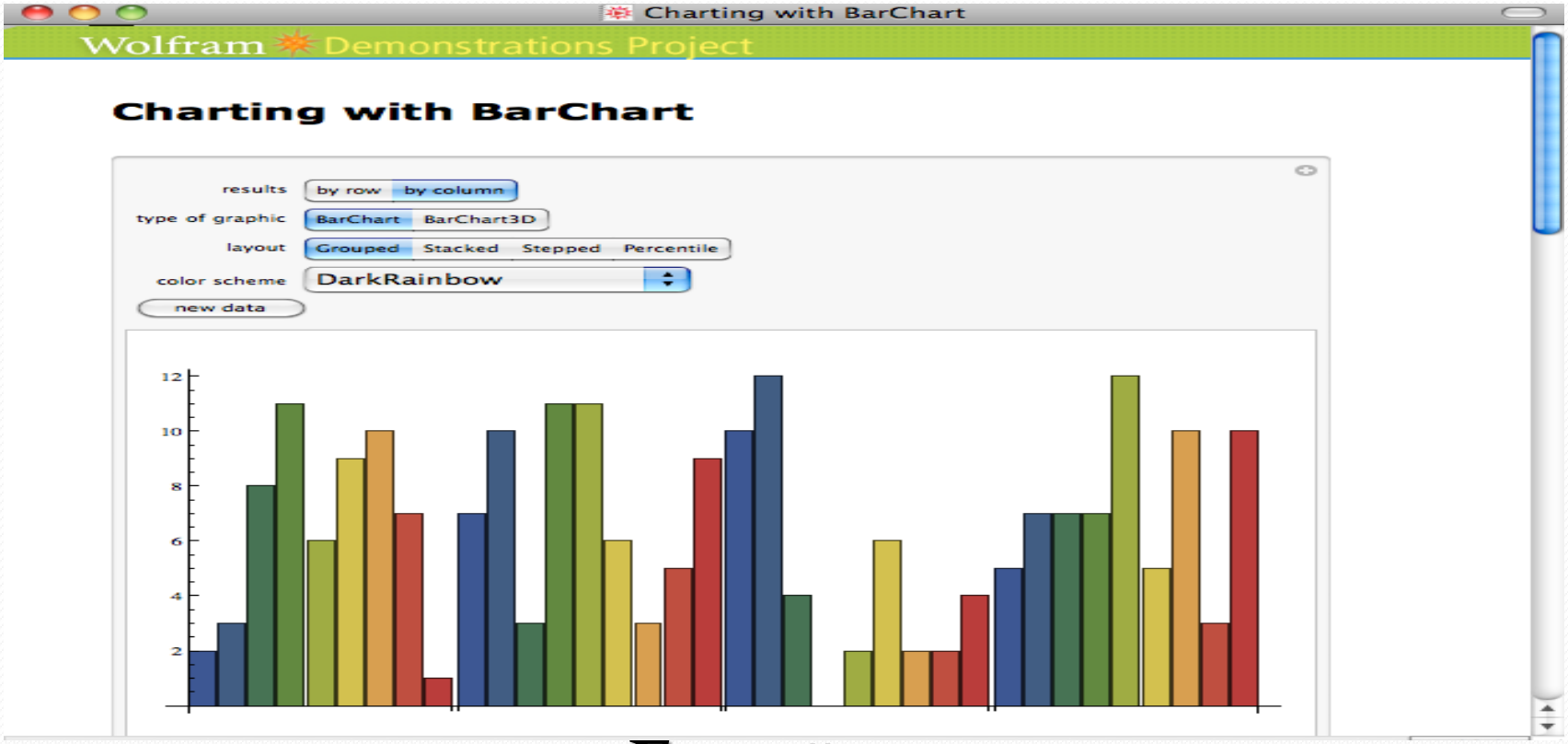


Fact #9

Color in small regions is difficult to perceive, and bright colors in large areas appear bigger.

Use bright, saturated colors for small regions, and use low saturation pastel colors for large regions and backgrounds.

Data Visualization: Color - Nominal



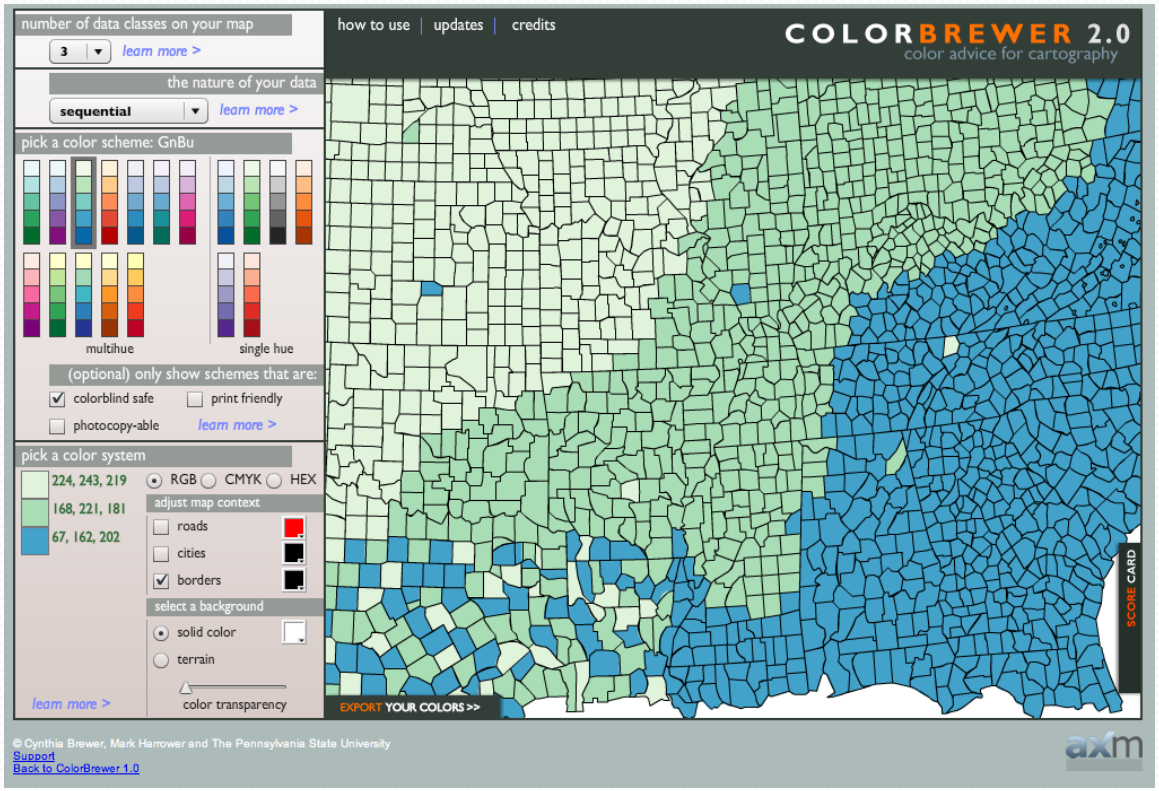
Fact #10

Only a small number of colors can be used effectively as nominal labels.

Keep the number of colors for nominal data to less than eight.

Use quiet medium grey backgrounds.

Data Visualization: Color -Ordinal



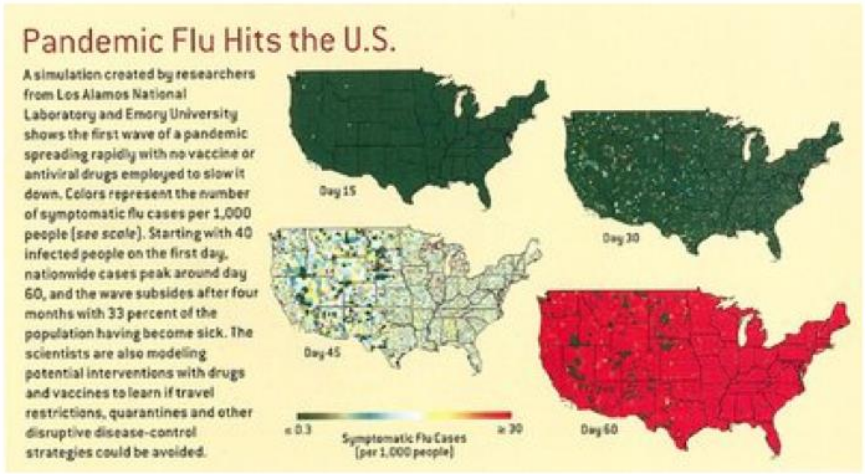
Fact #11

Lightness and saturation are effective for ordinal data because they have an implicit perceptual

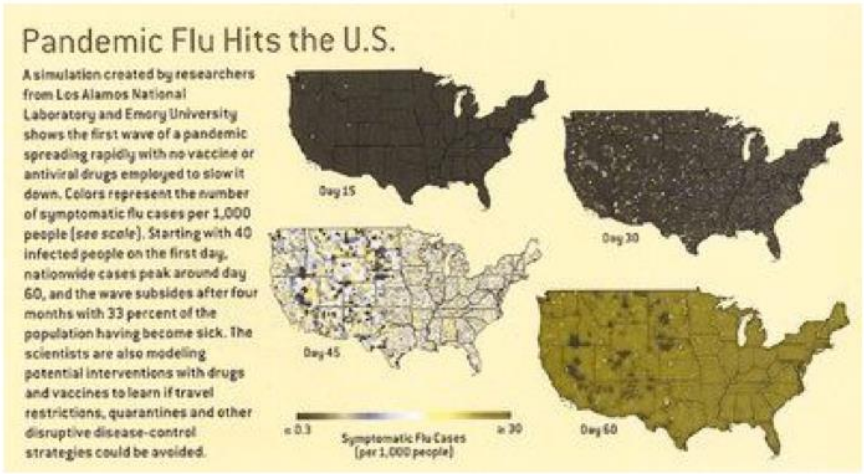
Show ordinal data with a discrete set of color values that change in lightness or saturation.

Data Visualization: Color - Quantitative

Original Image



Deuteranope Simulation



Fact #12

Quantitative data can be shown with a discrete or continuous colormap

Use colormaps with a limited hue palette and redundantly vary lightness and saturation.

Use discrete colormaps for accuracy.

Data Visualization: Τι περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

Cognition

Data Visualization: Cognition

principles of visual cognition

Perception:

about the nature of the signals coming in; what you *see*

Cognition:

about how you *understand* and *interpret* what you see

there's more to seeing than just what's there

REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Cognition

REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

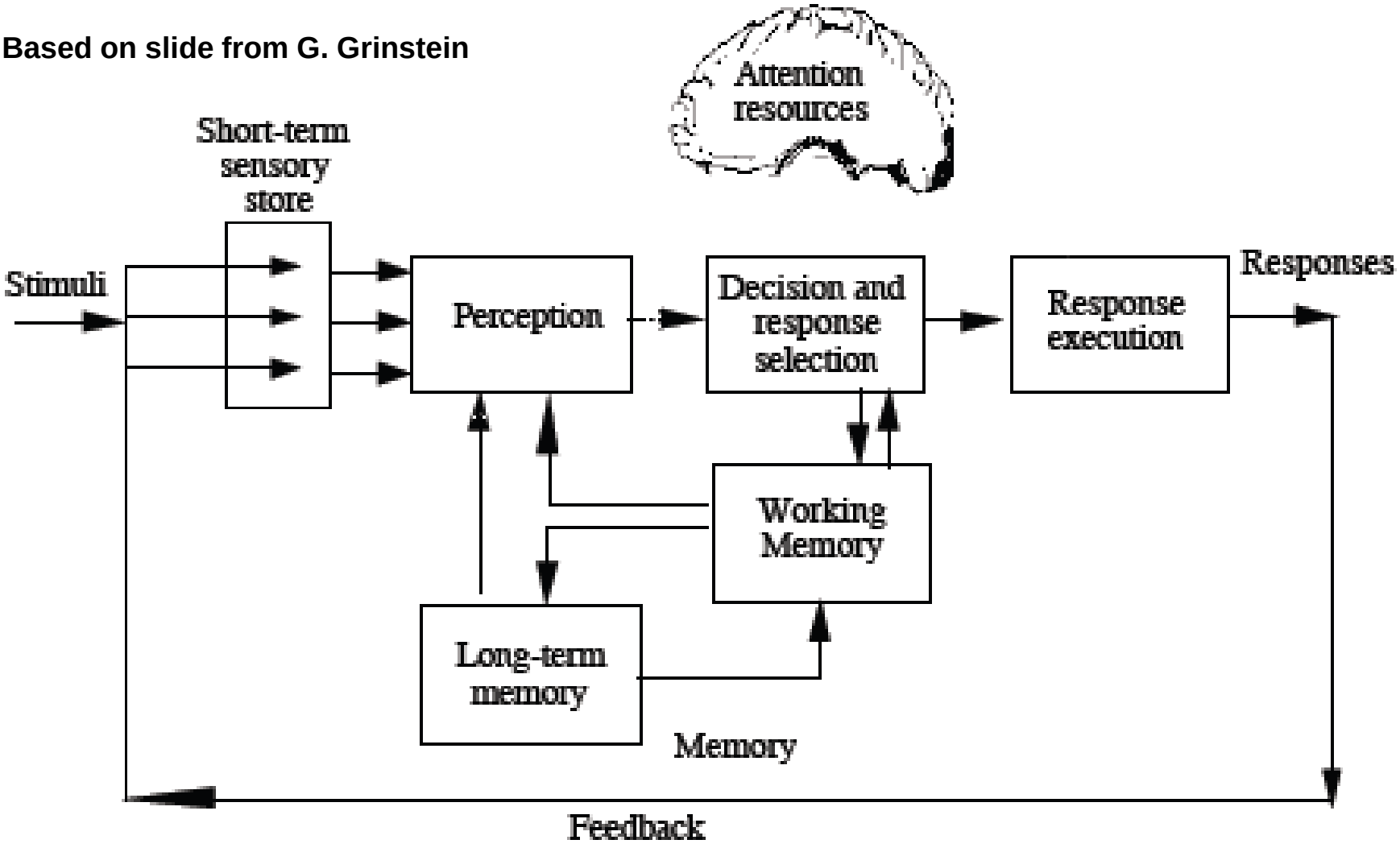
Data Visualization: Cognition

"What you see when you see a thing depends on what the thing is. What you see
the thing as depends on what you know
about what you are seeing."
Zenon Pylyshyn

REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Cognition

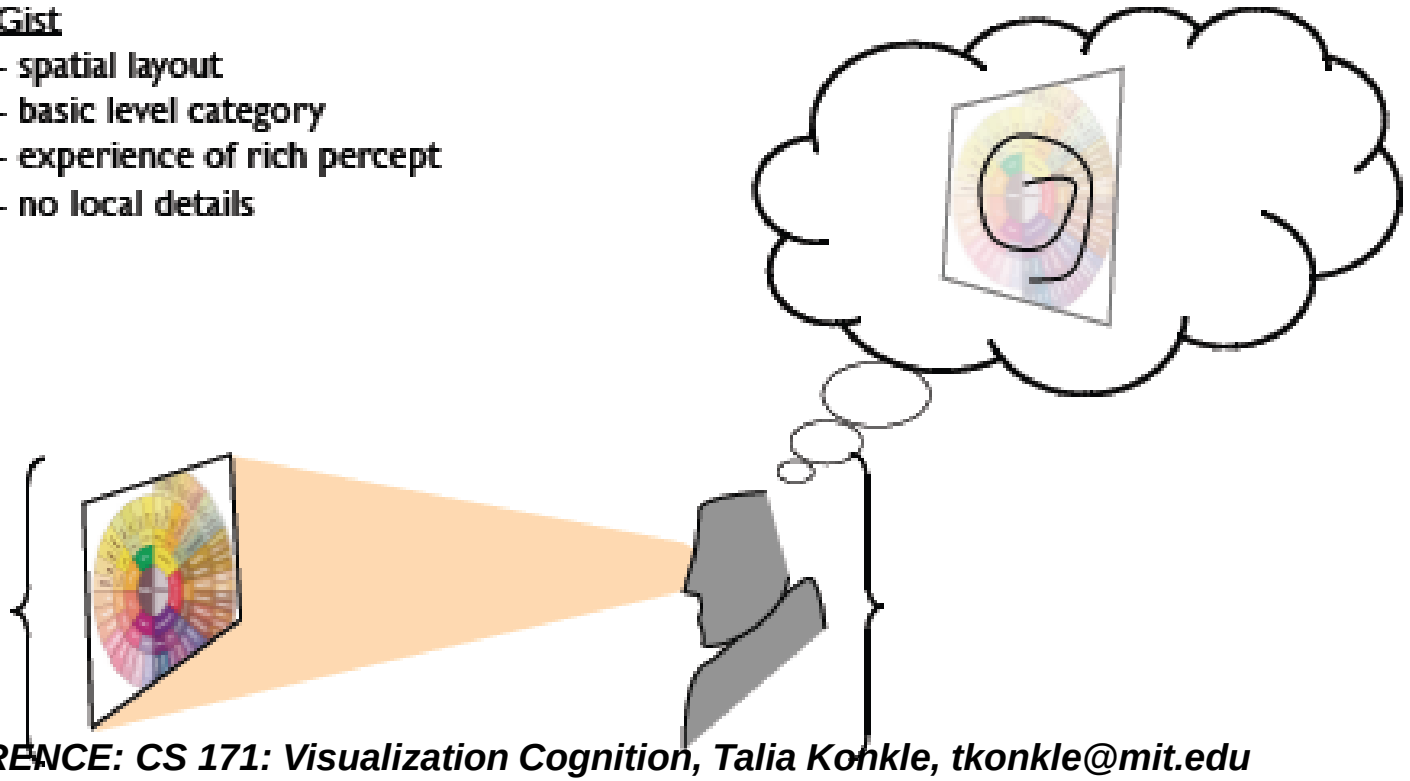
Based on slide from G. Grinstein



REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Cognition

Gist
- spatial layout
- basic level category
- experience of rich percept
- no local details

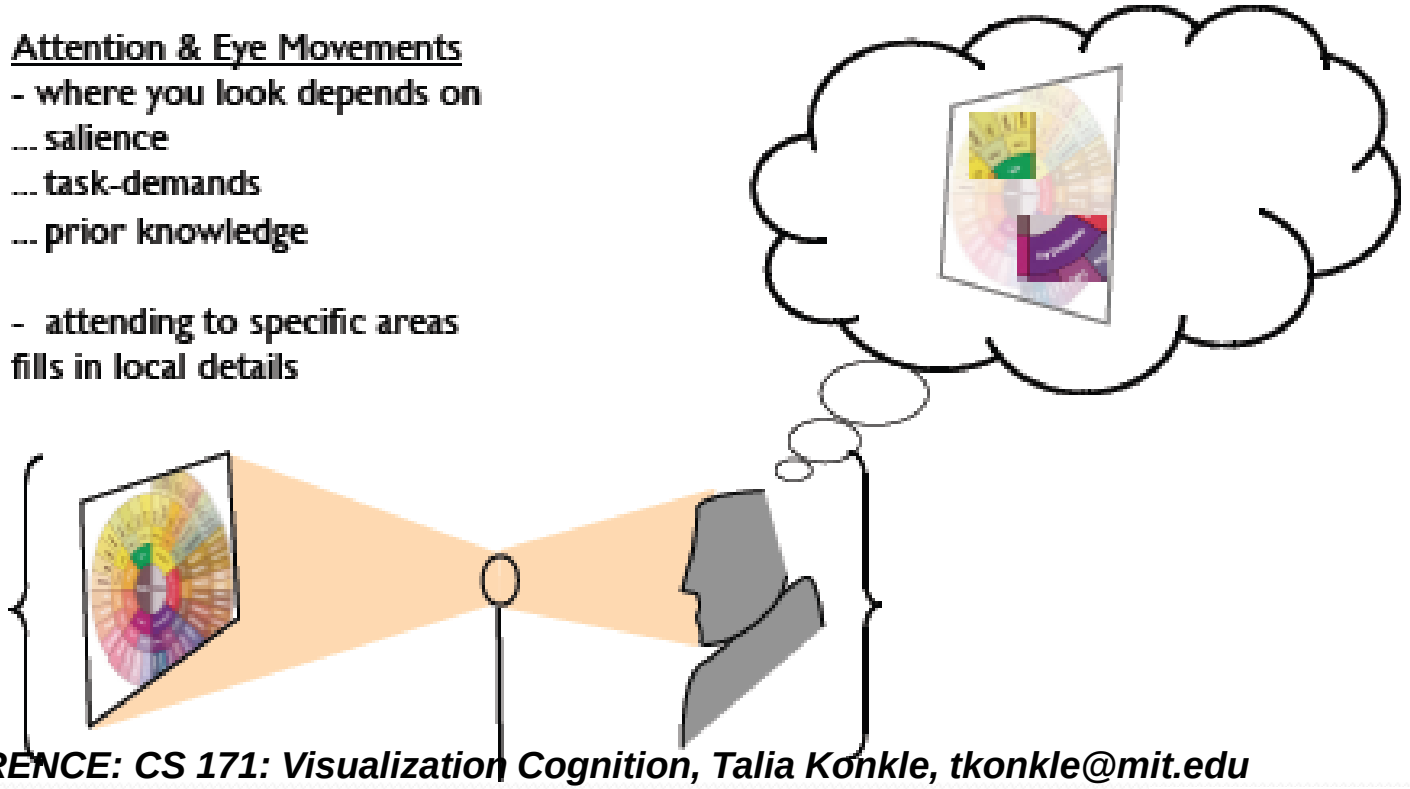


REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Cognition

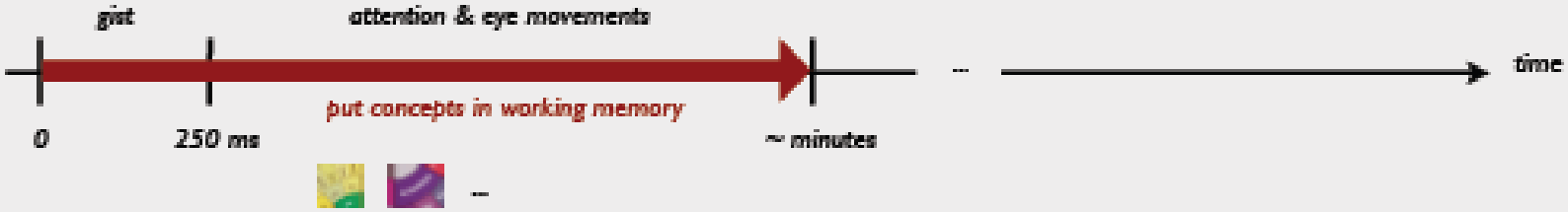
Attention & Eye Movements
- where you look depends on
... salience
... task-demands
... prior knowledge

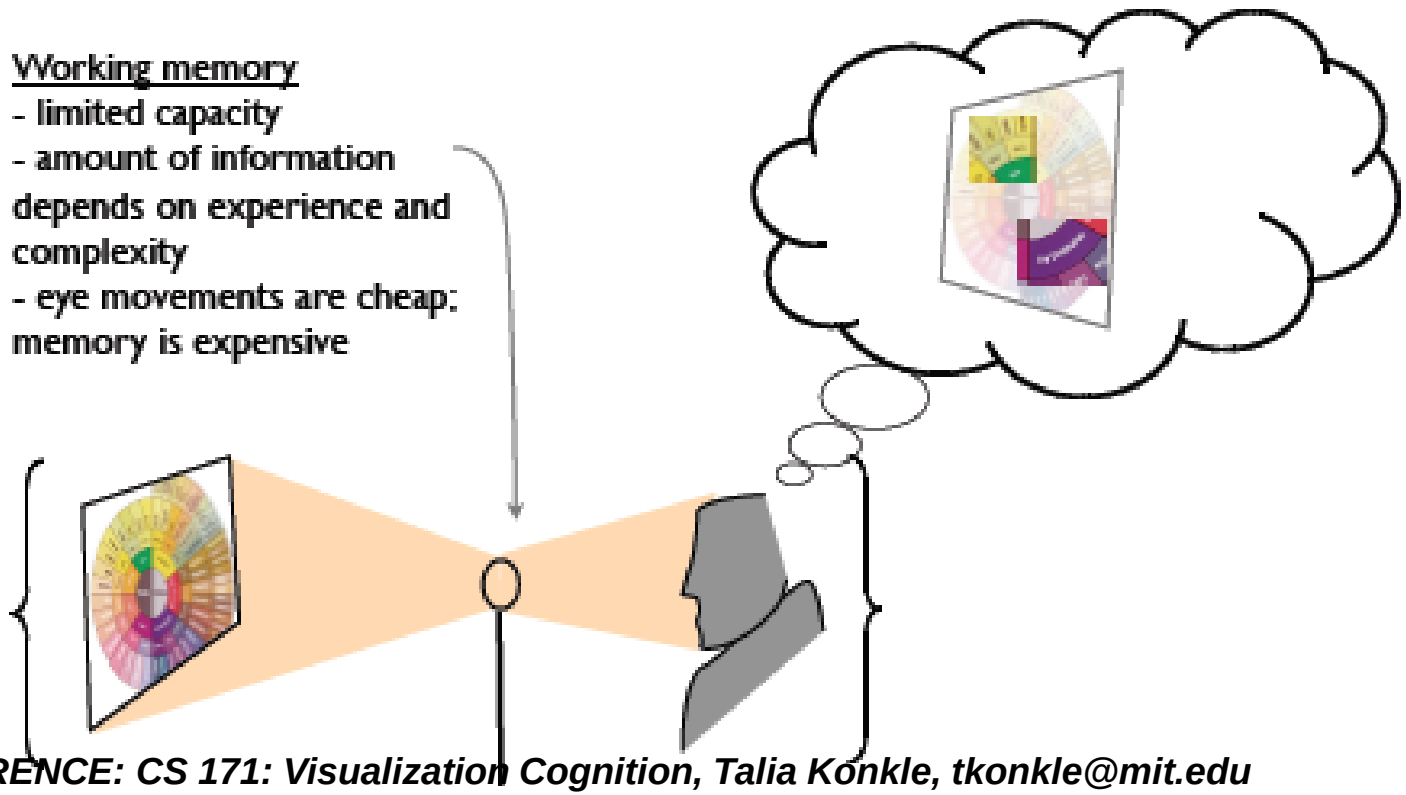
- attending to specific areas
fills in local details



REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

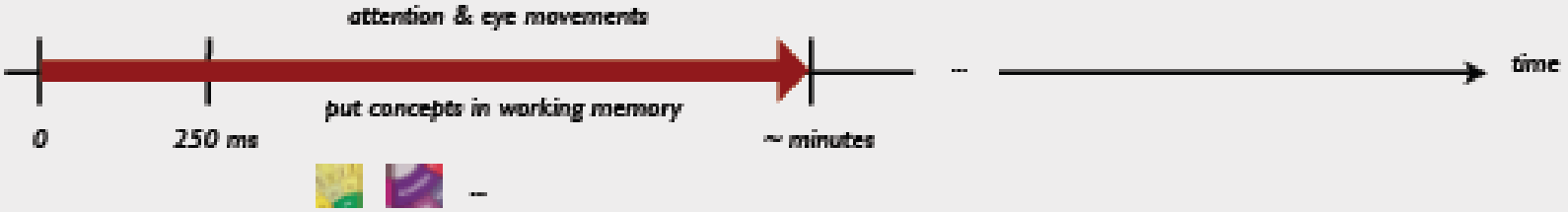
Data Visualization: Cognition





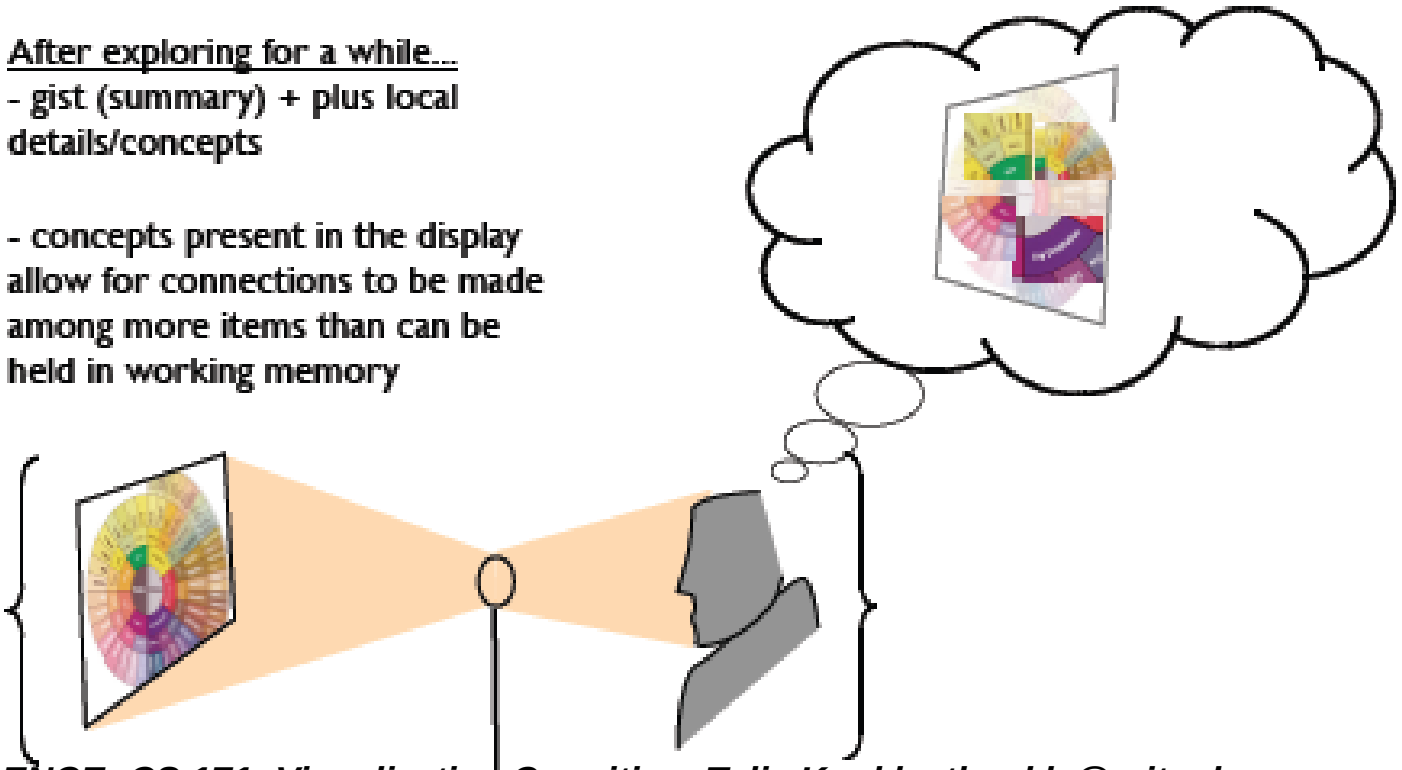
REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Cognition



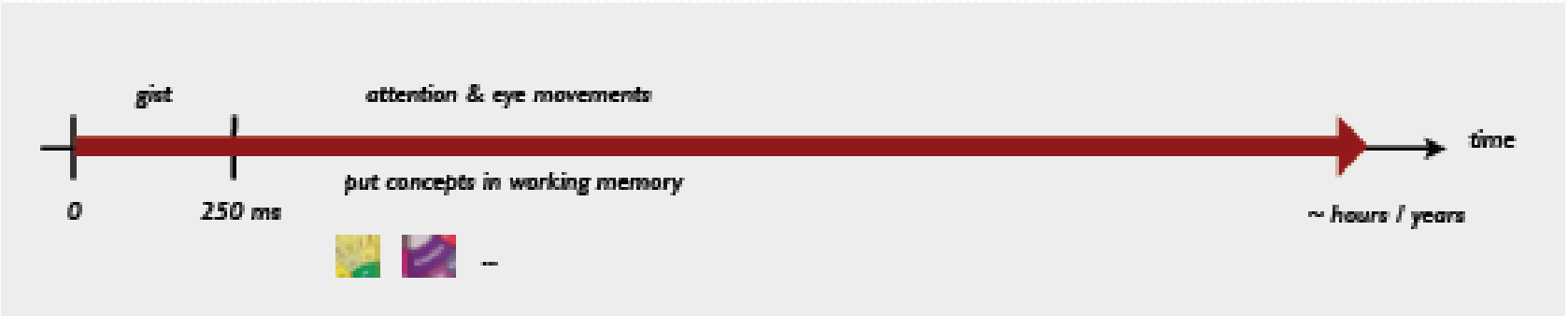
After exploring for a while...
- gist (summary) + plus local details/concepts

- concepts present in the display allow for connections to be made among more items than can be held in working memory



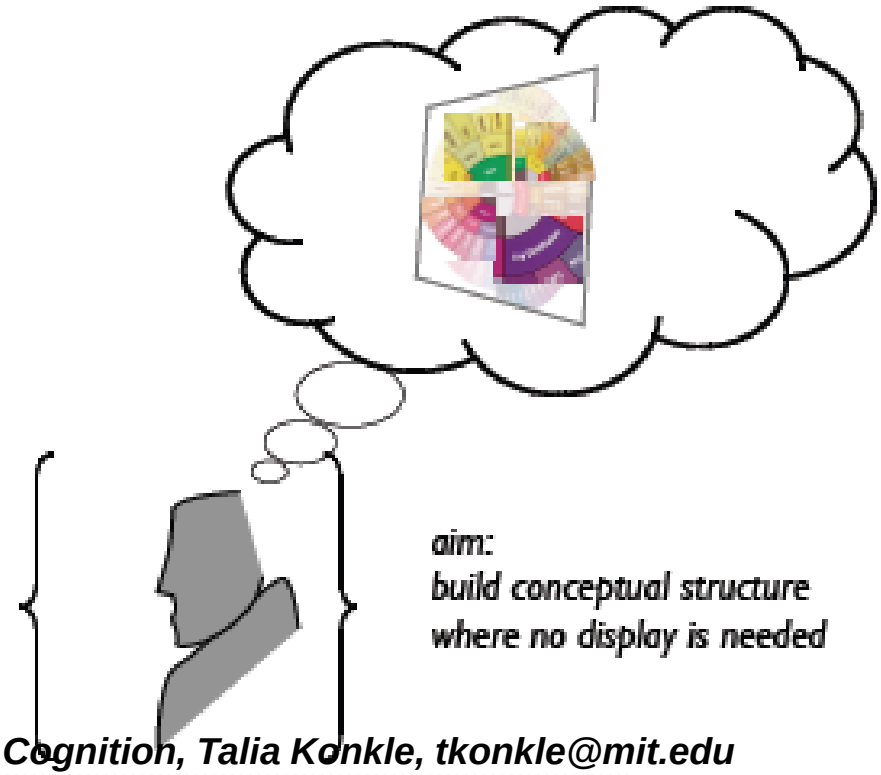
REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Cognition



Long-term memory
- massive capacity
- can be surprisingly detailed!
- aggregates visual experience

- provides 'chunks' for working memory and 'guidance' for attention



aim:
build conceptual structure
where no display is needed

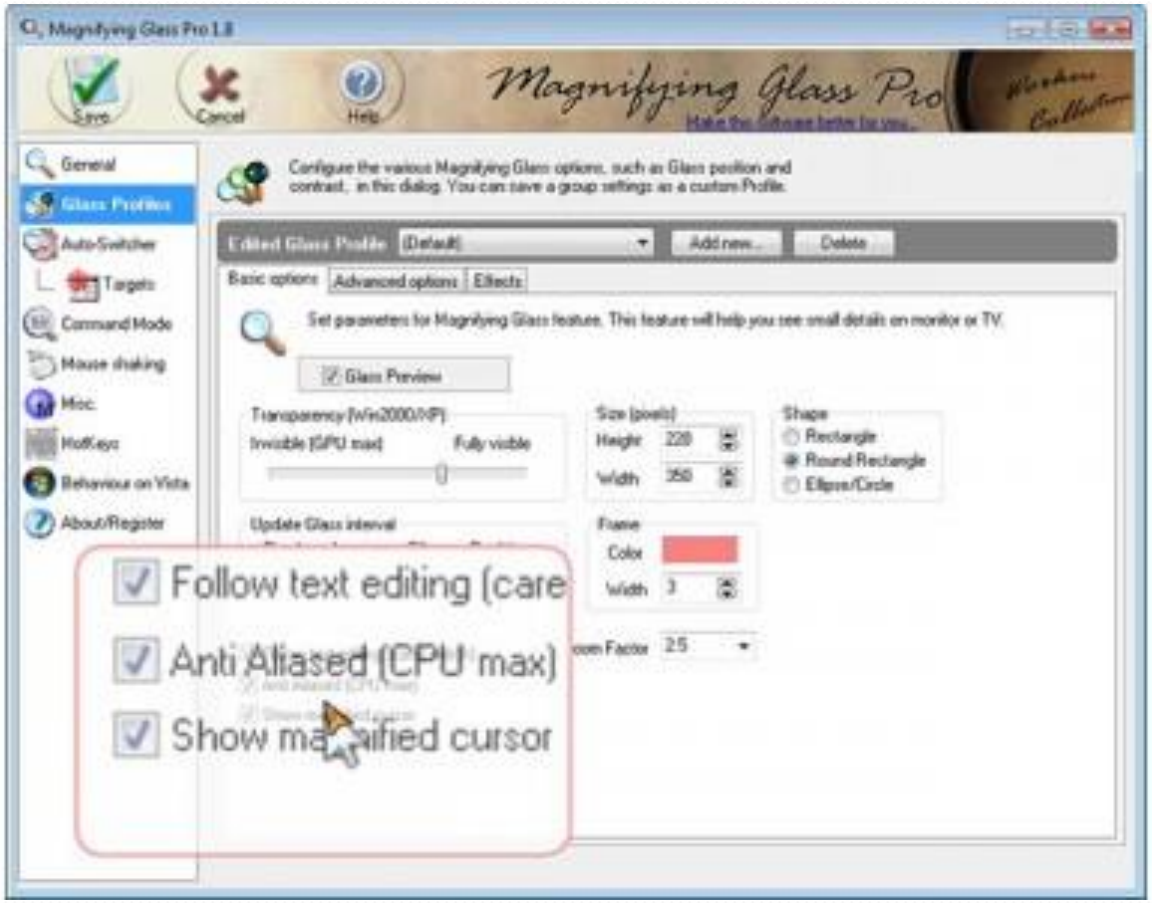
REFERENCE: CS 171: Visualization Cognition, Talia Konkle, tkonkle@mit.edu

Data Visualization: Τι περιλαμβάνει

- Design
- Process
- Models
- Perception
- Patterns
- Color
- Cognition
- Interaction
- Statistical Graphs
- Maps
- Trees and Networks
- Multi-Dimentional

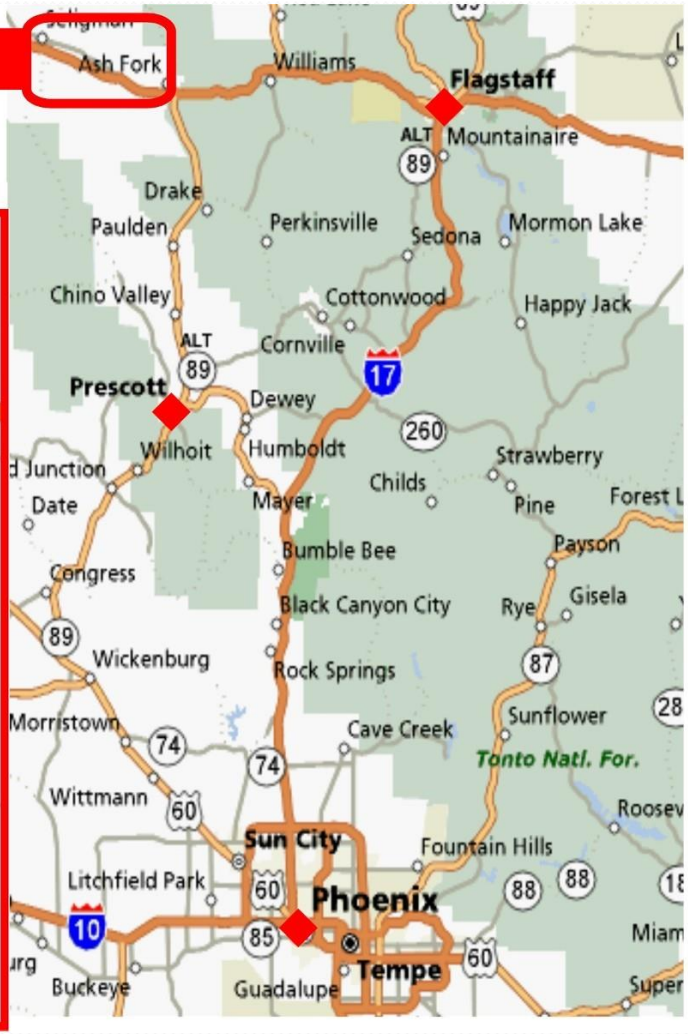
Interaction

Data Visualization: Interaction

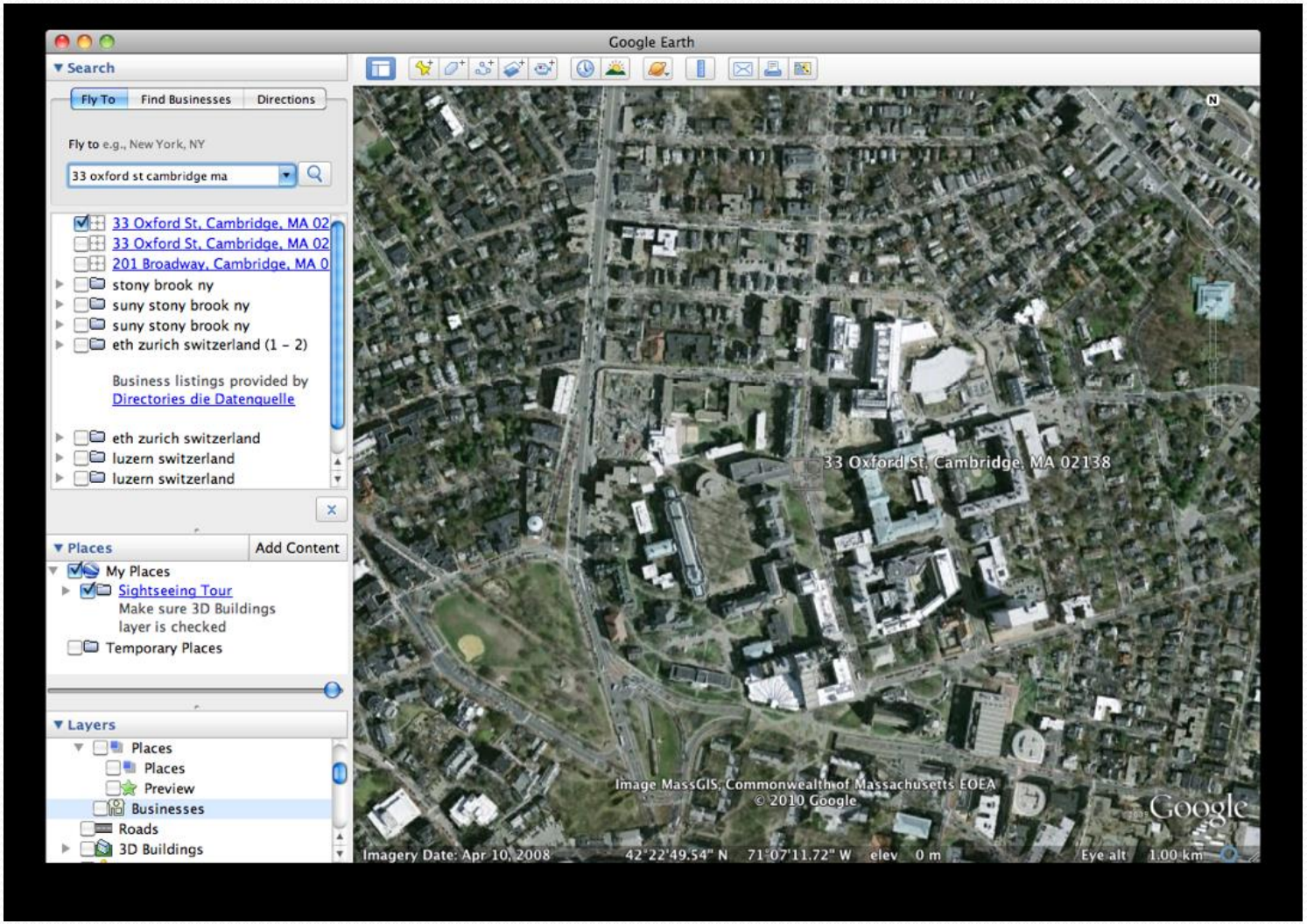


Data Visualization: Interaction

ReUnion Camp & Juniperwood Ranch Winery
7775 W. Stockman's Rd.
602-971-8586

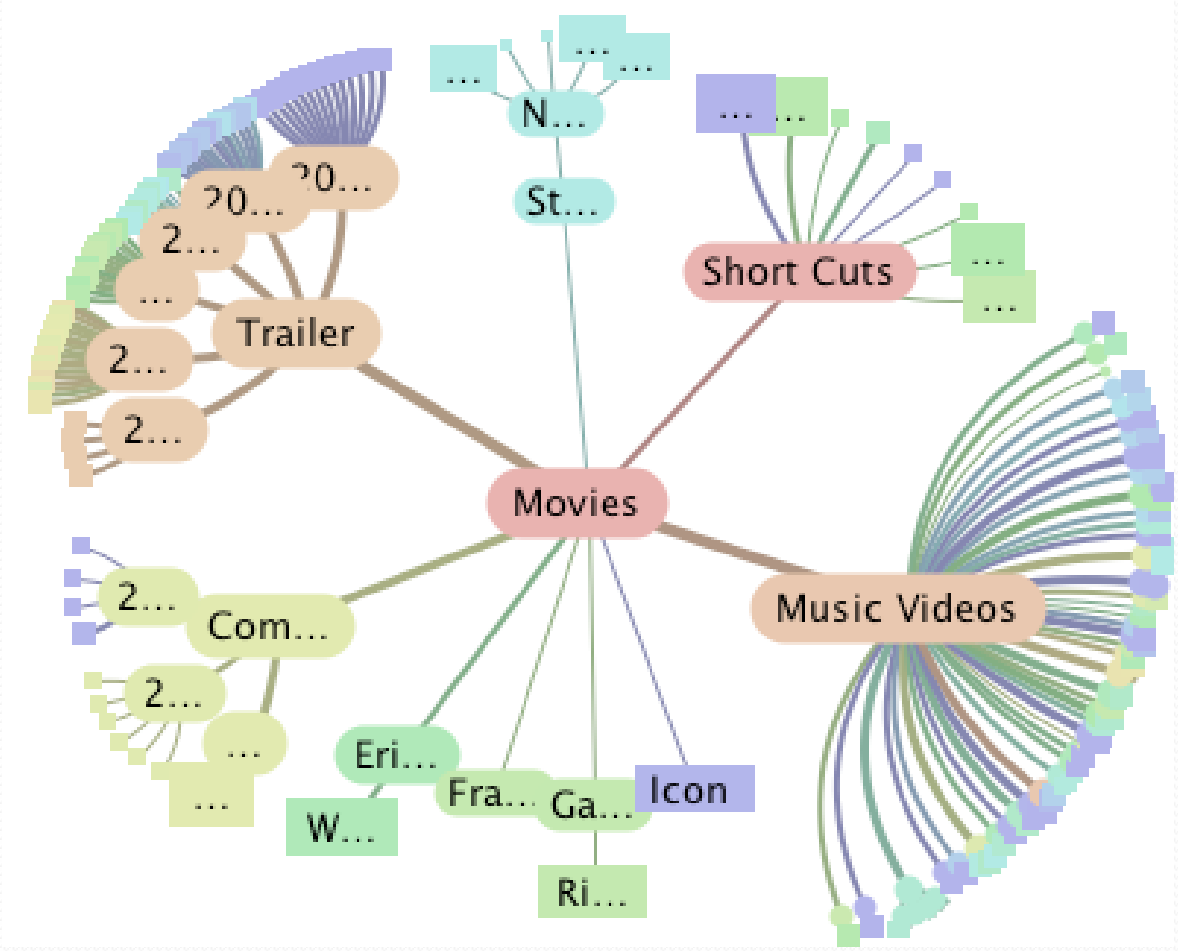


Data Visualization: Interaction



Data Visualization: Interaction

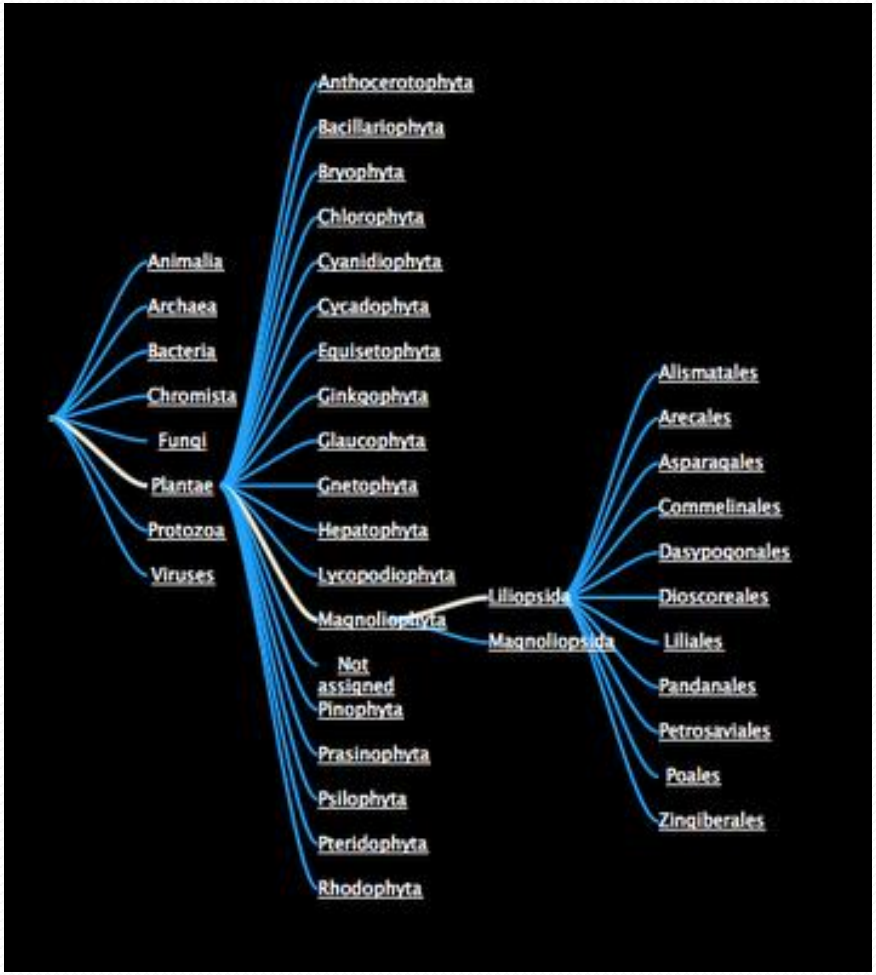
Focus & Context - Hyperbolic Tree 1995



http://www.randelshofer.ch/blog/2007/09/visualizing-large-tree-structures/

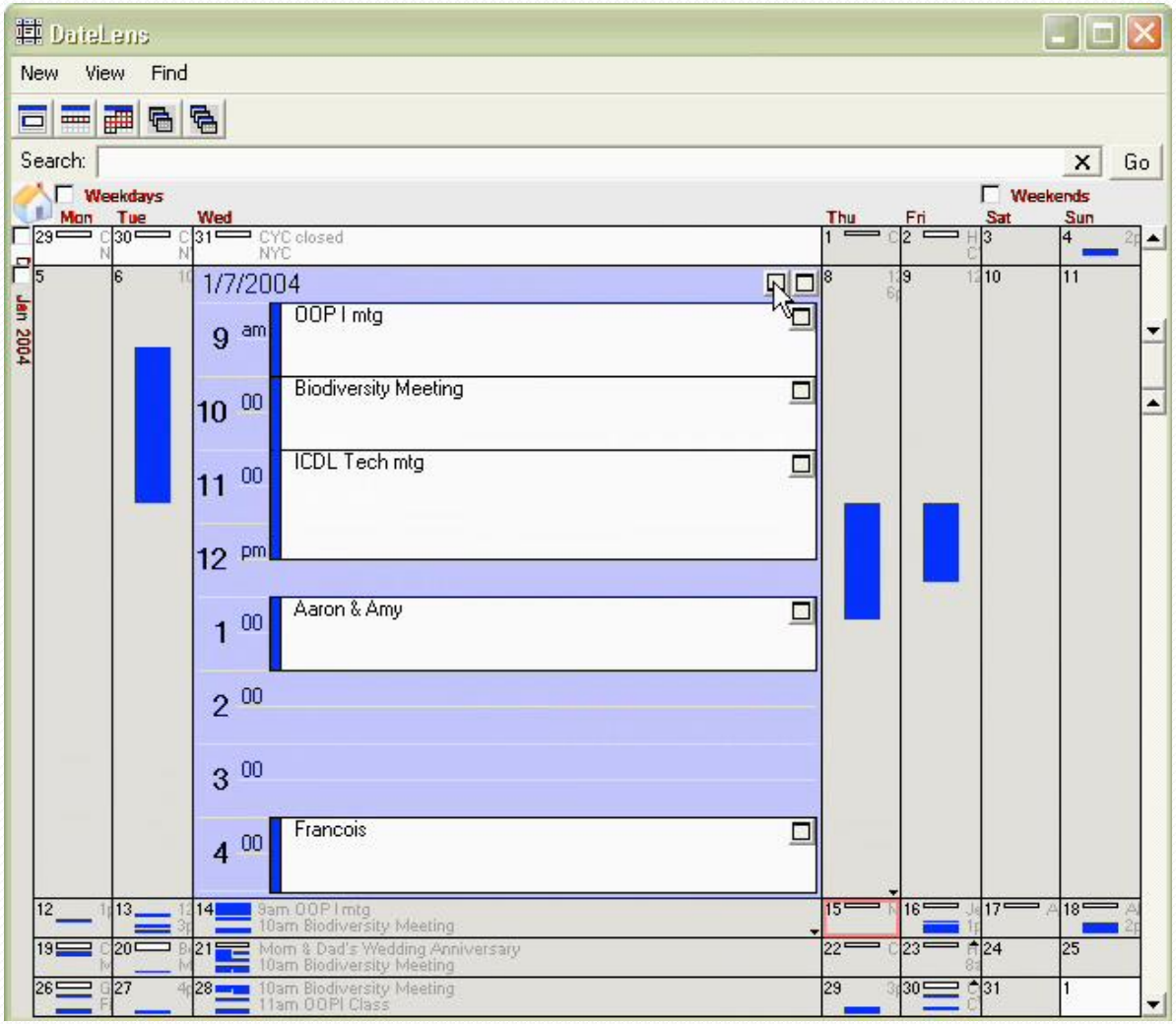
Data Visualization: Interaction

Focus & Context – Space Tree, 2002



Data Visualization: Interaction

Focus & Context – DataLens, 2003



Bederson et al., HCIL

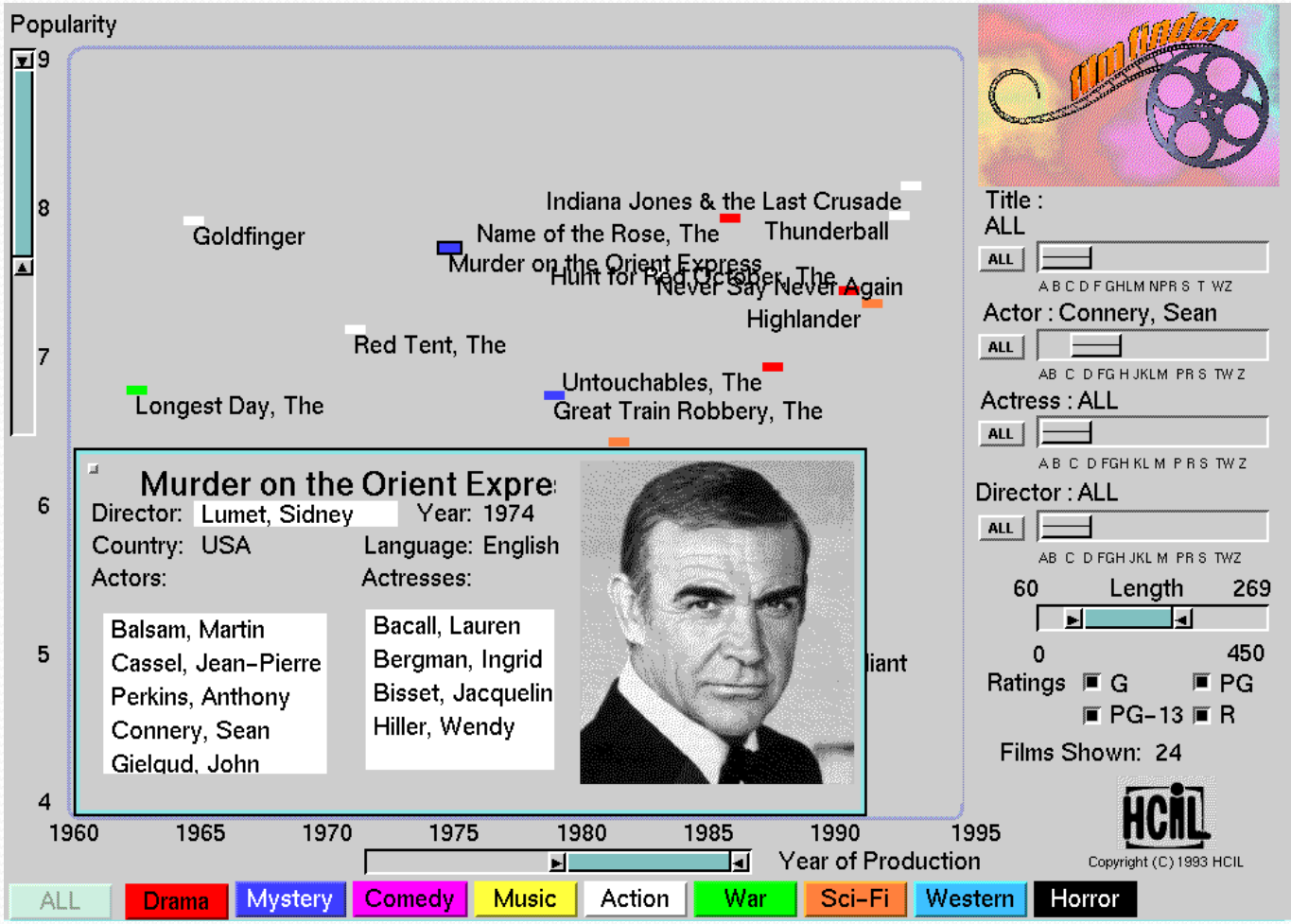
Data Visualization: Interaction

The Shneiderman Mantra:

"Overview first, zoom and filter, and details on demand."

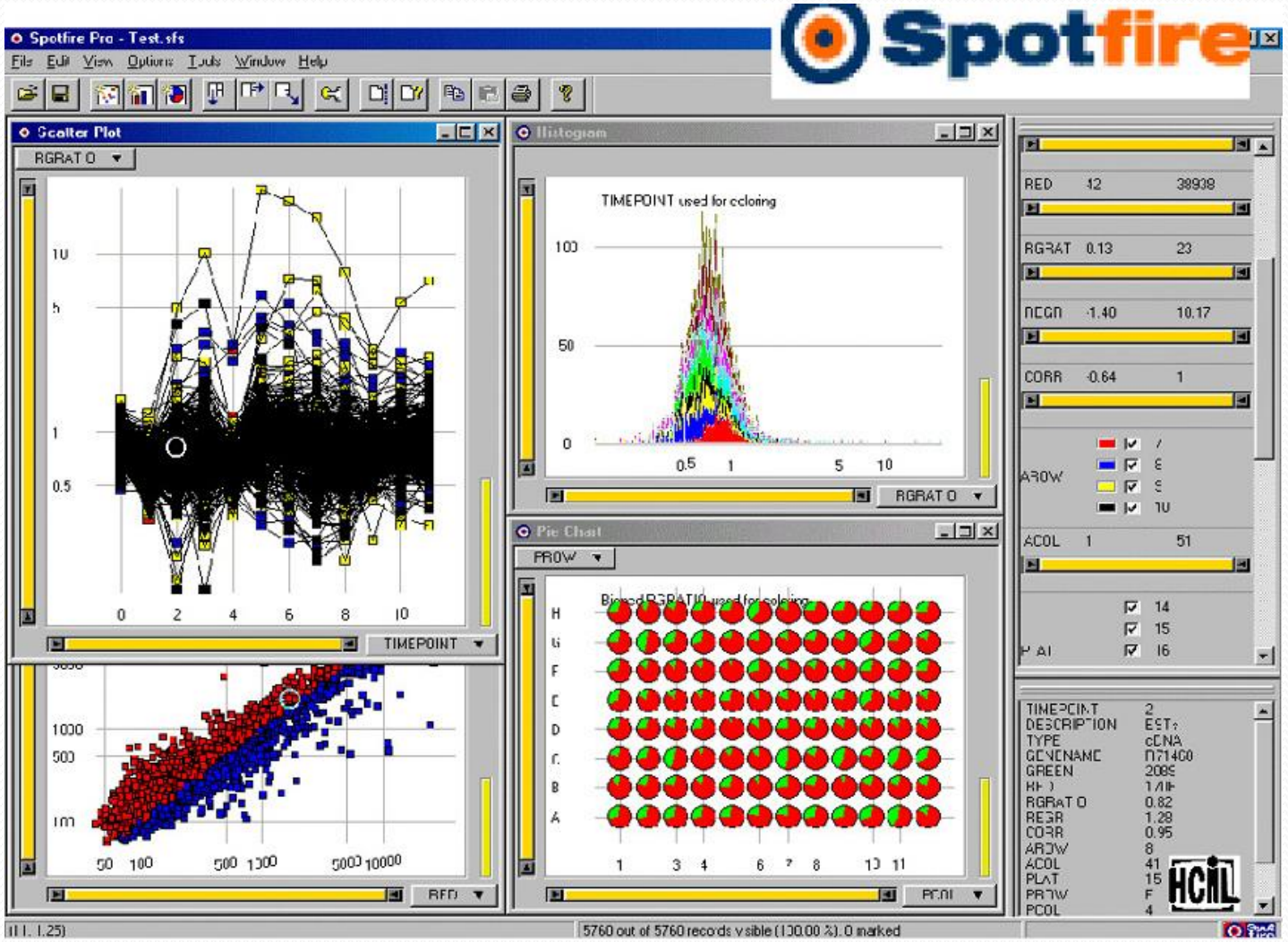
Ben Shneiderman

Data Visualization: Interaction



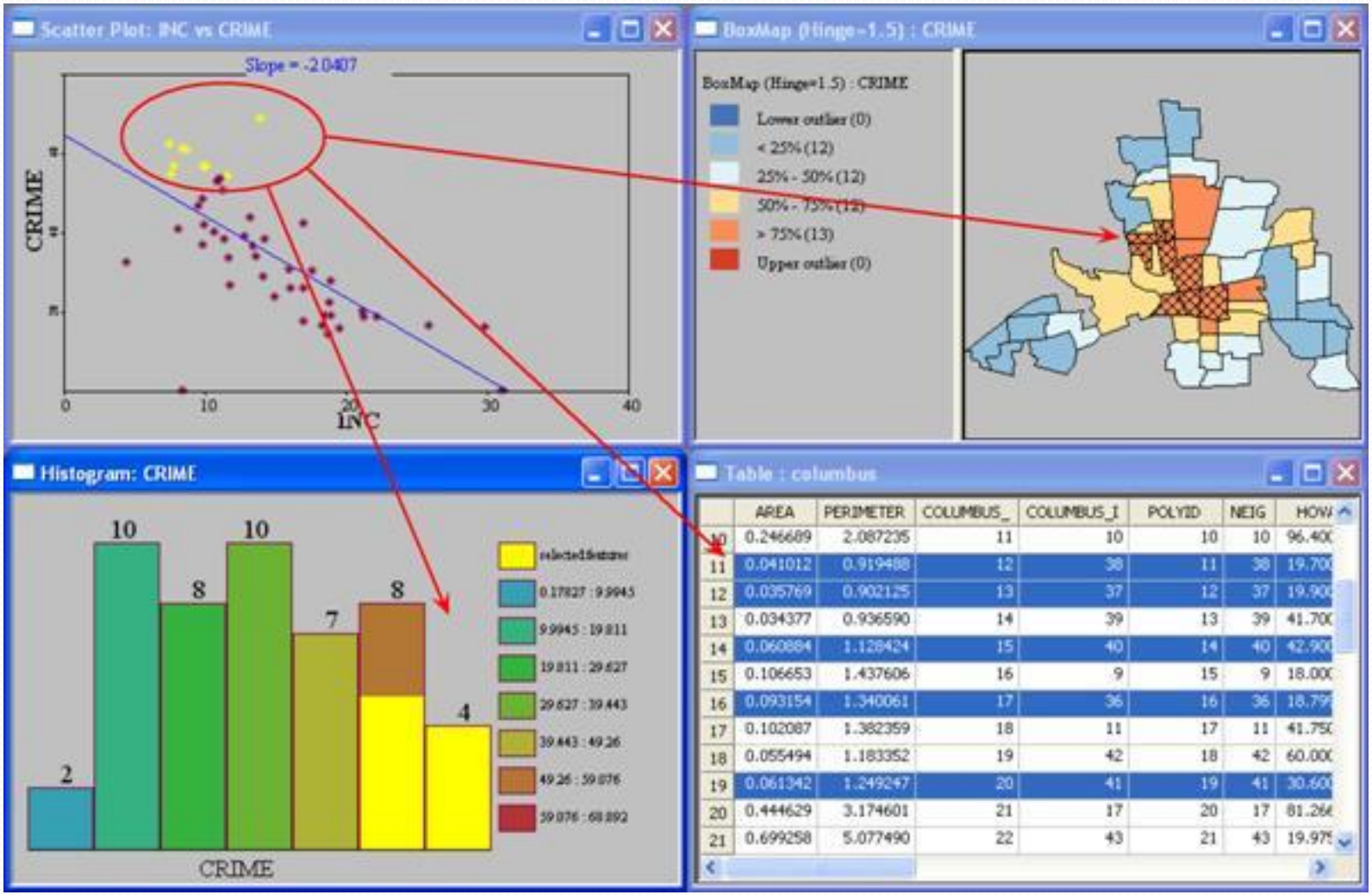
http://www.useit.com/papers/information_overload/

Data Visualization: Interaction



http://www.isr.umd.edu/research/research_briefs/older_accomplishments/039_Starfield/welcome.html

Data Visualization: Interaction



http://www.spatialanalysisonline.com/output/html/EDAESDAandESTDA.html

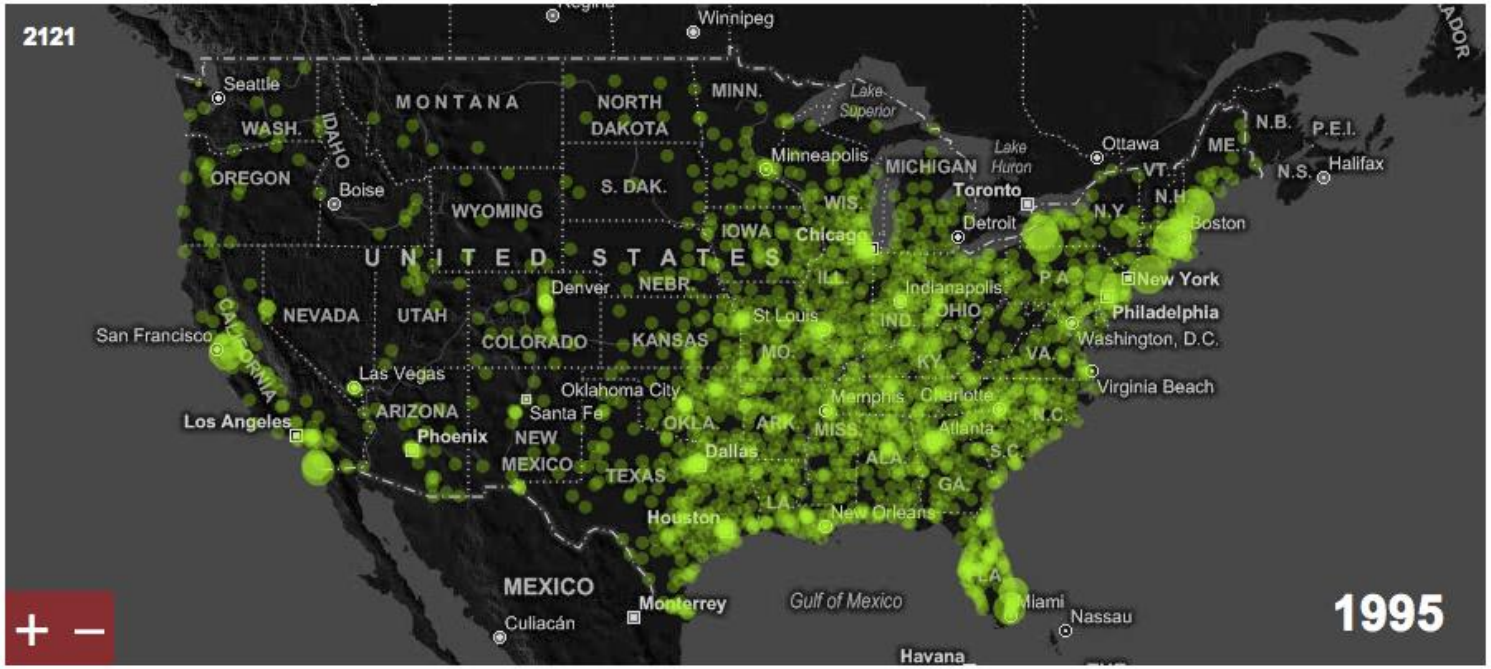
Data Visualization: Interaction

FlowingData
Strength in Numbers

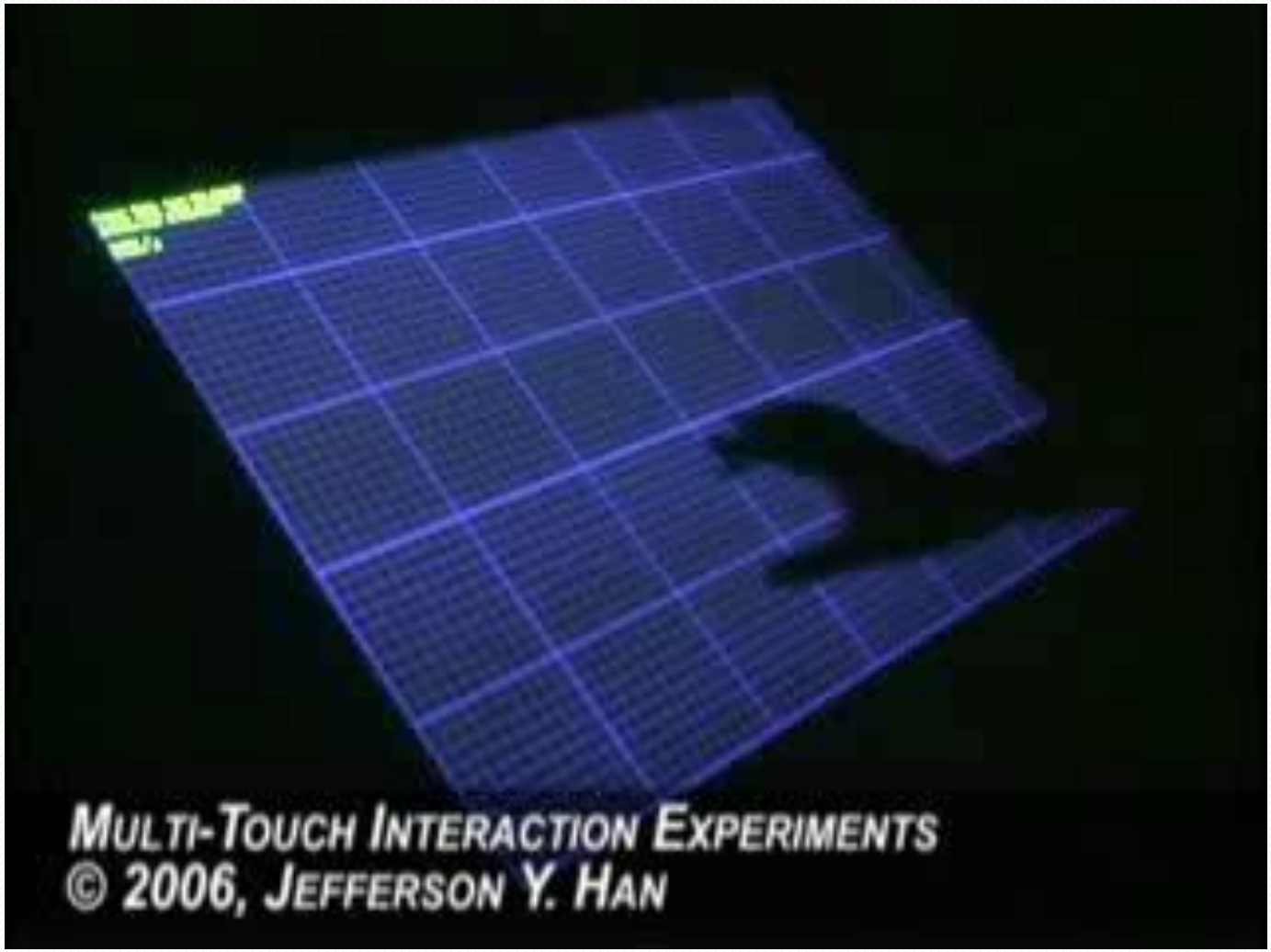
Type terms and press enter to search

Watching the Growth of Walmart Across America

Over the weekend, I mapped the spread of Walmart using Modest Maps. It starts slow and then spreads like wildfire in the southeast and makes its way towards the west coast. Subscribe to FlowingData / Read more...



Data Visualization: Interaction



Off-The-Desktop Interaction

Microsoft Surface

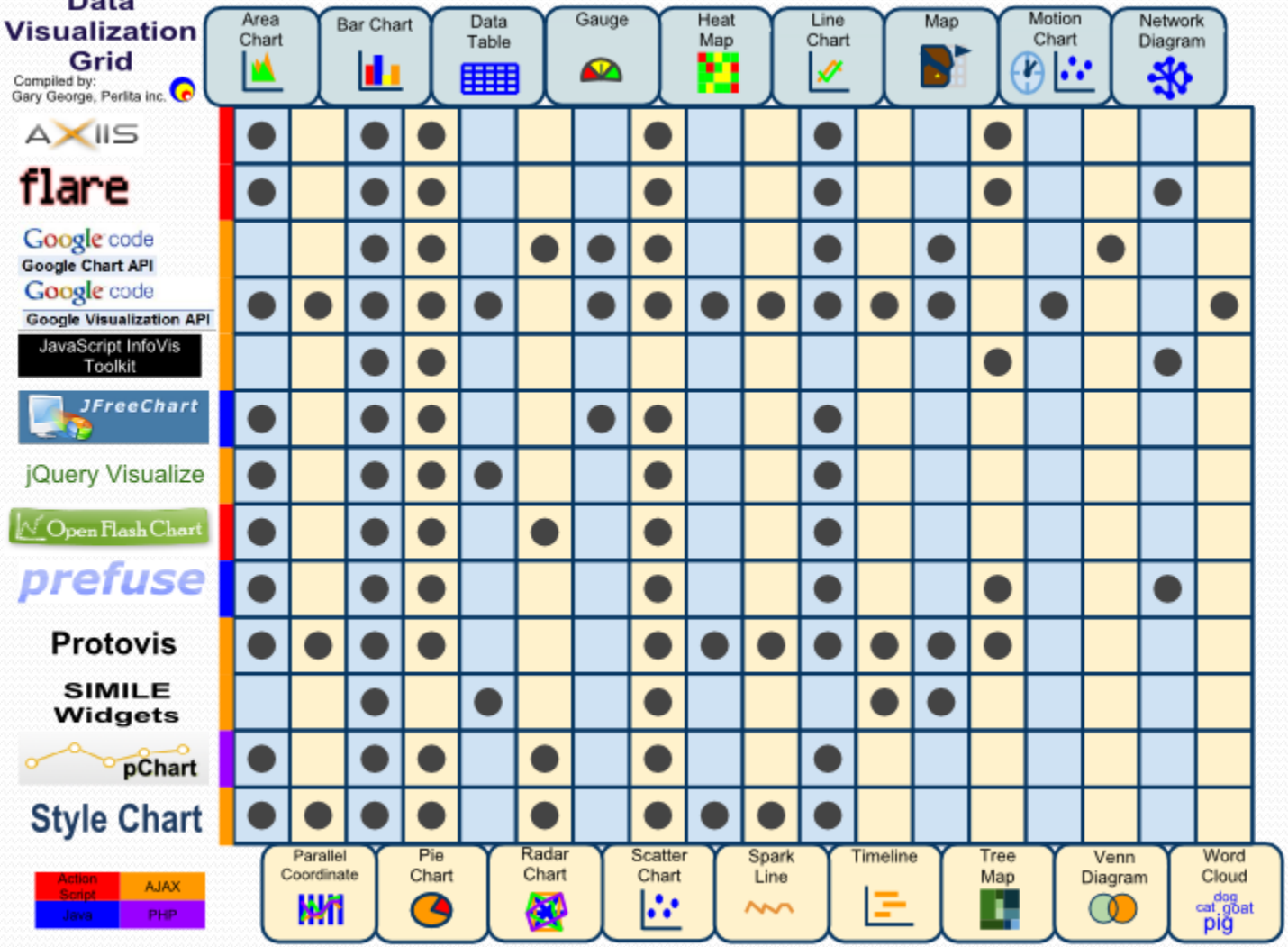
Data Visualization: Interaction

Off-The-Desktop Interaction

Vs.

Data Visualization: Toolkits

Data Visualization: Toolkits



http://www.perlitalabs.com/Infographics/Comparing_Data_Visualization_Tools.html

Data Visualization Toolkits:

Prefuse

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 host 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, and 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 a 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.

http://prefuse.org/

Data Visualization Toolkits:

Prefuse

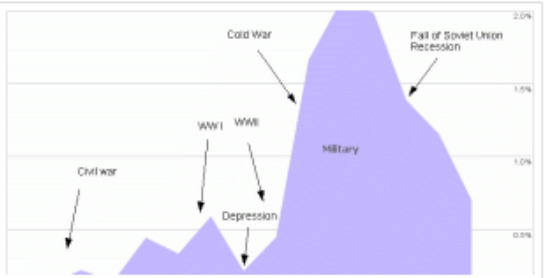
prefuse
INFORMATION VISUALIZATION TOOLKIT

Home | Dow

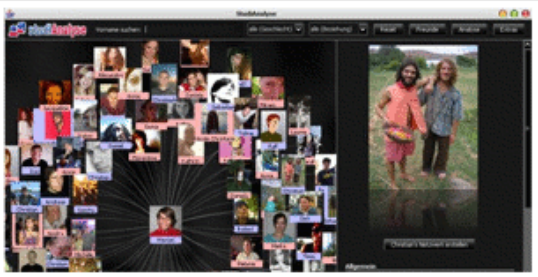
visualization gallery

DocuBurst by Christopher Collins

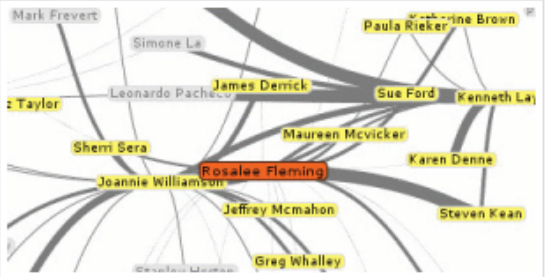
Uncertainty Lattices by Christopher Collins



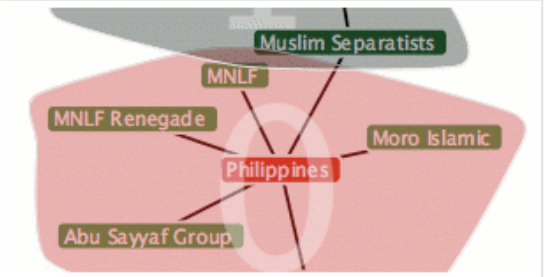
sense.us by Heer, Viégas, and Wattenberg



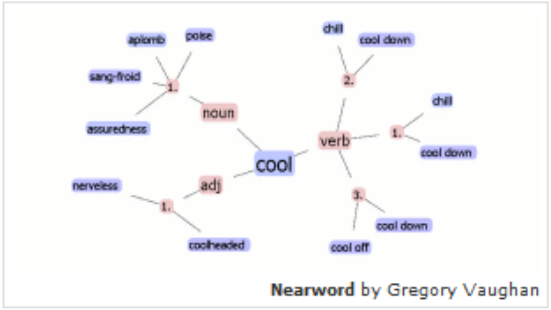
StudiAnalyse by Christoph Gerstle and Florian Moritz



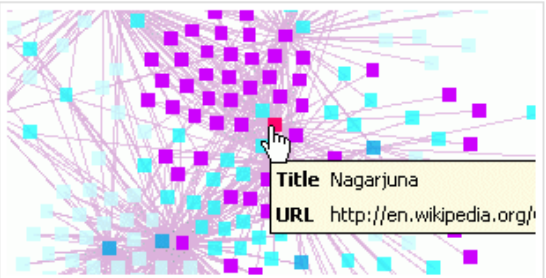
Enron Explorer by Trampoline Systems



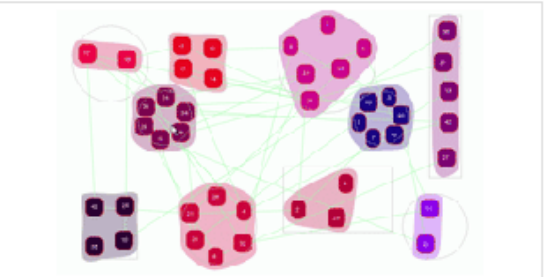
Social Action by Adam Perer



Nearword by Gregory Vaughan



34all by Martin Dudek



Zone Manager by Martin Dudek

Data Visualization Toolkits:

JFreeChart

JFreeChart is a free 100% Java chart library that makes it easy for developers to display professional quality charts in their applications. JFreeChart's extensive feature set includes:
a consistent and well-documented API, supporting a wide range of chart types;
a flexible design that is easy to extend, and targets both server-side and client-side applications;
support for many output types, including Swing components, image files (including PNG and JPEG), and vector graphics file formats (including PDF, EPS and SVG);
JFreeChart is "open source" or, more specifically, free software. It is distributed under the terms of the GNU Lesser General Public Licence (LGPL), which permits use in proprietary applications.
For a closer look at JFreeChart, please try our JFreeChart Demo (web start) or browse the Samples page.

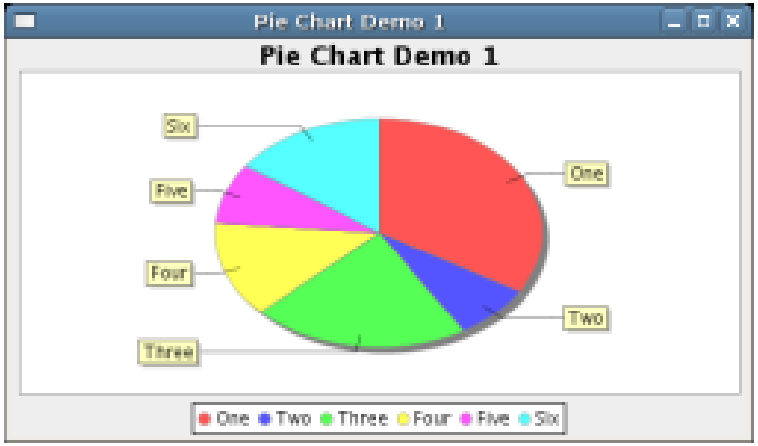
http://www.jfree.org/jfreechart/

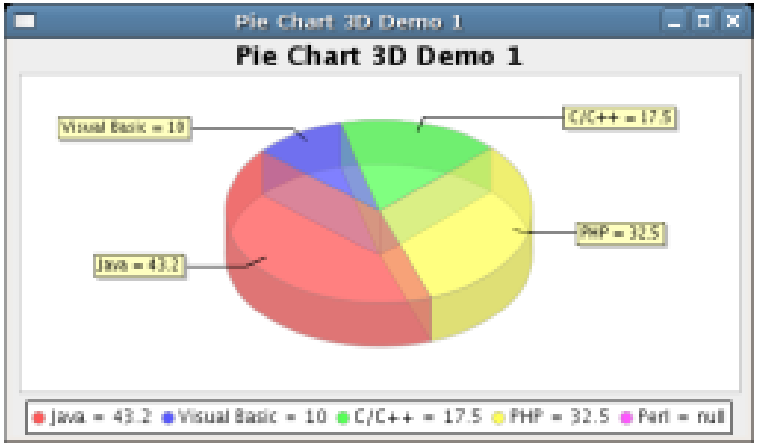
Data Visualization Toolkits:

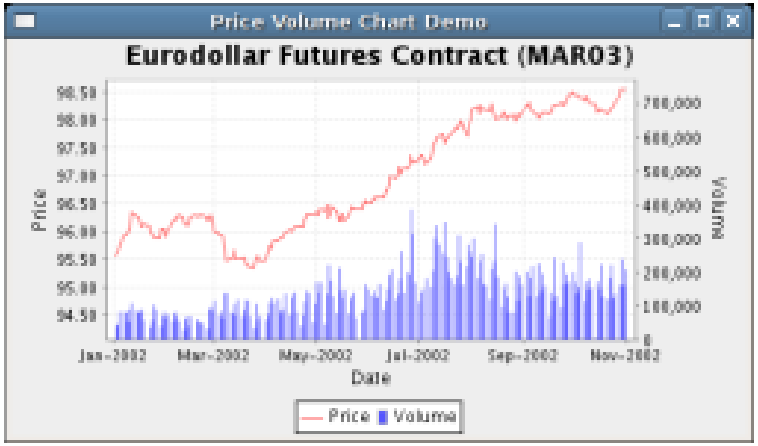
JFreeChart

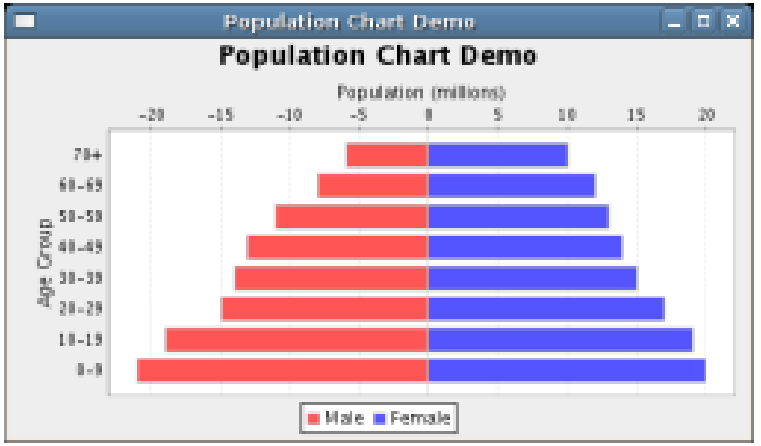
JFreeChart Samples

This page contains examples of the charts that can be produced using JFreeChart. If you'd prefer to see a live demo, please try our **JFreeChart Demo (web start)**.









Ασκήσεις(προαιρετικές)

- XXXXXXXXXXXXX
- XXXXXXXXXXX

Σύνοψη Μαθήματος

- YYYYY
  - YYYYY
- YYYYYYYYYYY
- YYYYYYYYY

Βιβλιογραφία - Where to Learn More

Several universities have developed dedicated graduate programs:

The University of Maryland, Stanford, the University of North Carolina, the University of California, Berkeley, and Georgia Tech are a few of the finest,

Only one academic journal: *Information Visualization Journal*, published quarterly by Palgrave Macmillan,

A few smaller publications: such as the *Visual Business Intelligence Newsletter,*

Conferences dedicated to the field are also few: IEEE's *VisWeek*, which includes the InfoVis and VAST (Visual Analytics Science and Technology) sub-conferences that are dedicated entirely to data visualization, r

Other conferences: CHI (Computer-Human Interaction) and SIGGRAPH.

CHI - Human Factors in Computing Systems

1982 1983 1985 1986 1987 1988 1989
1990 1991 1992 1993 1994 1995 1996
1997 1998 1999 2000 2001 2002 2003
2004 2005 2006 2007 2008 2009 2010

Next conference is coming up 07 May 2011 in Vancouver, BC, Canada

SIGGRAPH - International Conference on Computer Graphics and Interactive Techniques

1974 1975 1976 1977 1978 1979 1980
1981 1982 1983 1984 1985 1986 1987
1988 1989 1990 1991 1992 1993 1994
1995 1996 1997 1998 1999 2000 2001
2002

Next conference is coming up 09 Aug 2011 in Vancouver, Canada

InfoVis - IEEE Symposium on Information Visualization

1995 1997 1998 1999 2000 2001 2002
2003 2004 2005

The Information comes from:

Few, Stephen (2010). *Encyclopedia entry on Data Visualization for Human Perception*. Retrieved 21 March 2011 from Interaction-Design.org: http://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html

Βιβλιογραφία - Where to Learn More

Several good books have been written about data visualization. The following, in **chronological order**, are especially useful for surveying the field and as a source of basic instruction:

Tufte, Edward R. (1983): The Visual Display of Quantitative Information. Cheshire, CT, Graphics Press	All four of Tufte's books are exceptional, but his first is the best. It makes an inspiring case for graphical excellence.
Cleveland, William S. (1994): The Elements of Graphing Data. Hobart Press	Data visualization practices focused on the needs of statisticians.
Harris, Robert L. (2000): Information Graphics: A Comprehensive Illustrated Reference. Oxford University Press, USA	An encyclopedic reference for information graphics.
Card, Stuart K., Mackinlay, Jock D. and Shneiderman, Ben (eds.) (1999): Readings in Information Visualization: Using Vision to Think. Academic Press	An overview of the best academic research in the field as of the publication date.
Few, Stephen (2004): Show Me the Numbers: Designing Tables and Graphs to Enlighten. Analytics Press	An accessible, practical, and comprehensive guide to the design of tables and graphs for communication.
Ware, Colin (2008): Visual Thinking: for Design. Morgan Kaufmann	An eloquent introduction to visual perception and cognition as it relates to data visualization.
Few, Stephen (2009): Now You See It: Simple Visualization Techniques for Quantitative Analysis. Analytics Press	

The Information comes from:

Few, Stephen (2010). *Encyclopedia entry on Data Visualization for Human Perception*. Retrieved 21 March 2011 from Interaction-Design.org: http://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html

Βιβλιογραφία - Where to Learn More

Many blogs and **online discussion forums** feature data visualization - some thoughtfully, based on expertise, and some with the shallowness that is often found on the Web. Here are a few of the best:

Tufte.com (Edward Tufte)
Perceptual Edge (Stephen Few)
Eager Eyes (Robert Kosara)
Visual Complexity (Manuel Lima)
Flowing Data (Nathan Yau)
Pictures of Numbers (Mike Dickison)
Instant Cognition (Clint Ivy)

The Information comes from:
Few, Stephen (2010). *Encyclopedia entry on Data Visualization for Human Perception*. Retrieved 21 March 2011 from Interaction-Design.org: http://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html

Επόμενο μάθημα

- XXXXXXXXX