CS448B :: 23 Oct 2012

Interaction

Jeffrey Heer Stanford University

Visualization Tools

How do people create visualizations?

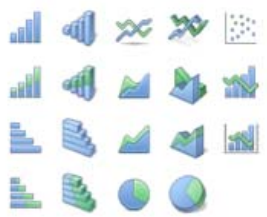


Chart Typology
Pick from a stock of templates
Easy-to-use but limited expressiveness
Prohibits novel designs, new data types

Component Architecture
Permits more combinatorial possibilities
Novel views require new operators,
which requires software engineering.

Today's first task is not to invent wholly new [*graphical*] techniques, though these are needed. Rather we need most vitally to recognize and reorganize **the essential of old techniques**, to make easy their **assembly in new ways**, and to modify their external appearances to **fit the new opportunities**.

J. W. Tukey, *The Future of Data Analysis*, 1962.

Efficiency

Chart Typologies
Excel, Many Eyes, Google Charts

Visual Analysis Languages
Tableau VizQL, ggplot2, HiVE

?

Component Model Architectures
Improvise, Prefuse, Flare

Graphics APIs
OpenGL, Java2D, GDI+, Processing

Expressiveness

Efficiency

Chart Typologies
Excel, Many Eyes, Google Charts

Visual Analysis Languages
Tableau VizQL, ggplot2, HiVE

Declarative Encoding Languages
Protovis, D3

Component Model Architectures
Improvise, Prefuse, Flare

Graphics APIs
OpenGL, Java2D, GDI+, Processing

Expressiveness

Designing Custom Tools and Graphics

Chart Typologies
Excel, Many Eyes, Google Charts

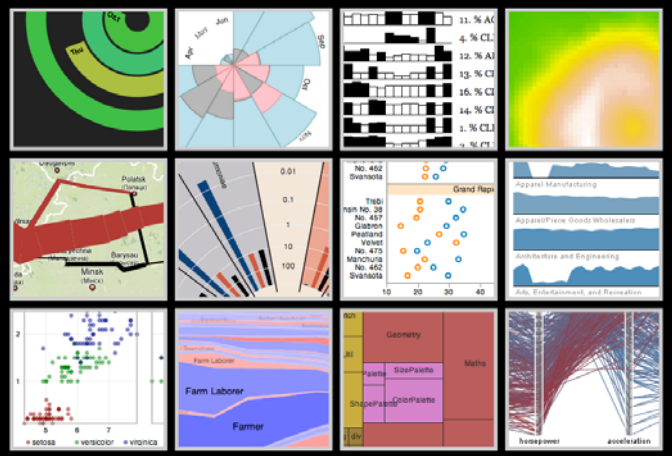
Visual Analysis Languages
Tableau VizQL, ggplot2, HiVE

Exploratory Data Analysis

Declarative Encoding Languages
Protovis, D3

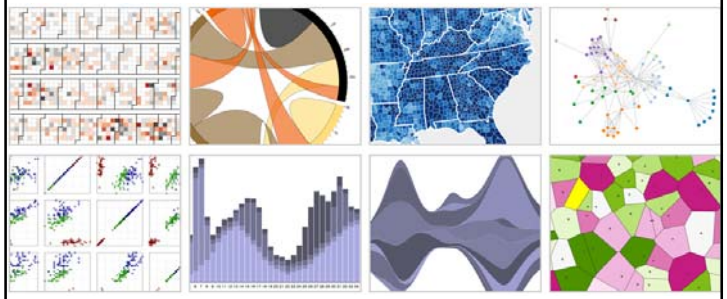
Component Model Architectures
Improvise, Prefuse, Flare

Graphics APIs
OpenGL, Java2D, GDI+, Processing



Protovis - A Graphical Toolkit for Visualization

d3.js Data-Driven Documents



http://mbostock.github.com/d3

Protovis

Specialized mark types

+ Streamlined design
- Limits expressiveness
- More overhead (slower)
- Harder to debug
- Self-contained model

Specify a scene (nouns)

+ Quick for static vis
- Animation, interaction are more cumbersome

D3

Bind data to DOM

- Exposes SVG/CSS/...
+ Exposes SVG/CSS/...
+ Less overhead (faster)
+ Debug in browser
+ Use with other tools

Transform a scene (verbs)

- More complex model
+ Dynamic data, animation, and interaction natural

Interaction

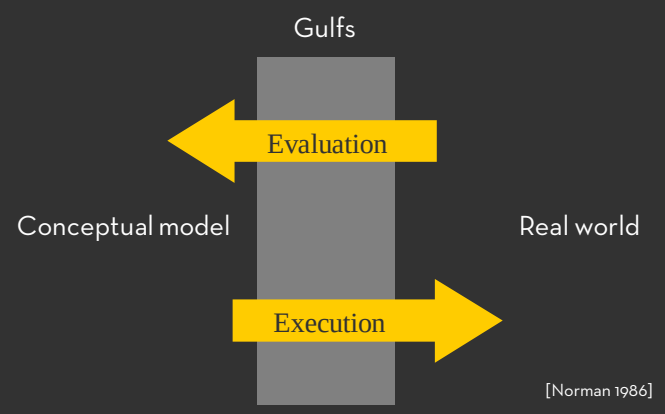
[There is an] apparent challenge that computational artifacts pose to the longstanding distinction between the physical and the social, in the special sense of those things that one designs, builds, and uses, on the one hand, and those things with which one communicates, on the other.

"Interaction"– in a sense previously reserved for describing a uniquely interpersonal activity – seems appropriately to characterize what goes on between people and certain machines as well.

Lucy Suchman, *Plans and Situated Actions*

Interaction between people and machines requires *mutual intelligibility* or *shared understanding*.

Gulfs of Execution & Evaluation



[Norman 1986]

Gulf of Execution

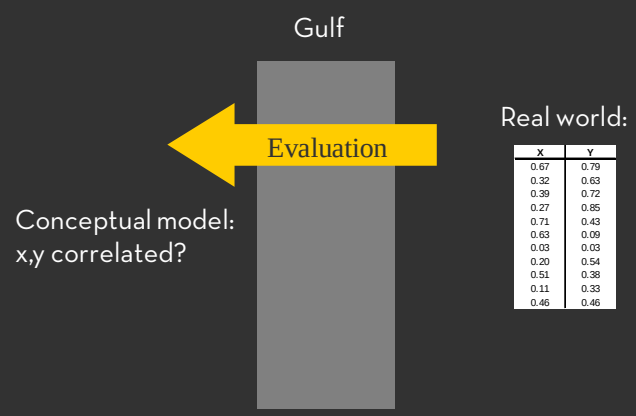
The difference between the user's intentions and the allowable actions.

Gulf of Evaluation

The amount of effort that the person must exert to interpret the state of the system and to determine how well the expectations and intentions have been met.

[Norman 1986]

Gulf of Evaluation



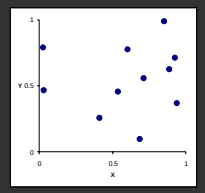
X	Y
0.67	0.79
0.32	0.63
0.39	0.72
0.27	0.85
0.71	0.43
0.63	0.09
0.03	0.03
0.20	0.54
0.51	0.38
0.11	0.33
0.46	0.46

Gulf of Evaluation

Gulf

Evaluation

Real world:



Conceptual model:
x,y correlated?

Gulf of Evaluation

Gulf

Evaluation

Real world:

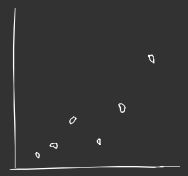
$\rho = -.29$

Conceptual model:
x,y correlated?

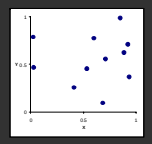
Gulf of Execution

Gulf

Conceptual model:
Draw a scatterplot



Execution



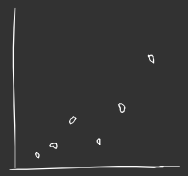
Real world

Move 90 30
Rotate 35
Pen down
…

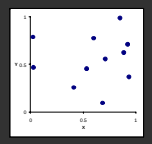
Gulf of Execution

Gulf

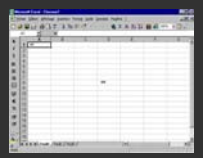
Conceptual model:
Draw a scatterplot



Execution



Real world



Interaction Techniques

Are there "essential" interactive operations for supporting exploratory data visualization?

Interaction Techniques

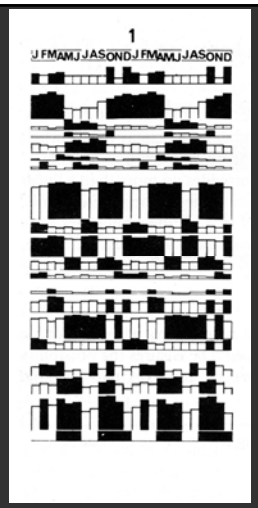
Are there "essential" interactive operations for supporting exploratory data visualization?

- View Specification (map data to visual vars)
- Navigation (pan, zoom, scale, rotate)
- Selection / Highlighting
- Filtering
- Sorting
- Extract Data

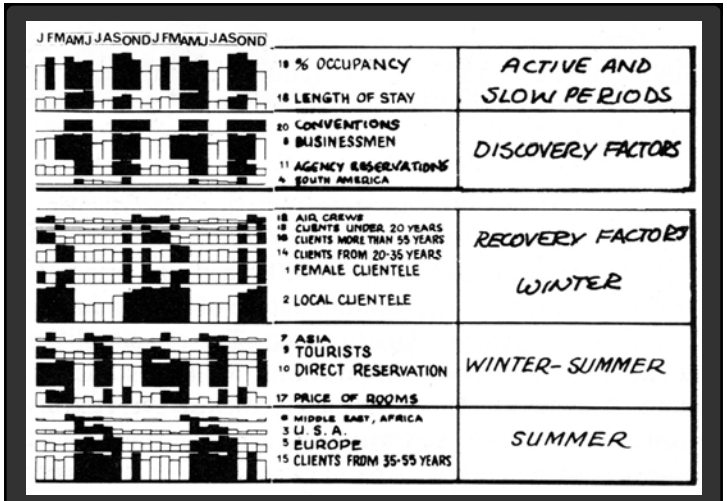
Interactive Visualization

J	F	M	A	M	J	J	A	S	O	N	D		
26	21	26	28	20	20	20	20	20	40	15	40	1	% CLIENTELE FEMALE
69	70	77	71	37	36	39	39	55	60	68	72	2	% —"— LOCAL
7	6	3	6	23	14	19	14	9	6	8	8	3	% —"— U.S.A.
0	0	0	0	8	6	6	4	2	12	0	0	4	% —"— SOUTH AMERICA
20	15	14	15	23	27	22	30	27	19	19	17	5	% —"— EUROPE
1	0	0	8	6	4	6	4	2	1	0	1	6	% —"— M.EAST, AFRICA
3	10	6	0	3	13	8	9	5	2	5	2	7	% —"— ASIA
78	80	85	86	85	87	70	76	87	85	87	80	8	% BUSINESSMEN
22	20	15	14	15	13	30	24	13	15	13	20	9	% TOURISTS
70	70	75	74	69	68	74	75	68	68	64	75	10	% DIRECT RESERVATIONS
20	18	19	17	27	27	19	19	26	27	21	15	11	% AGENCY —"—
10	12	6	9	4	5	7	6	6	5	15	10	12	% AIR CREWS
2	2	4	2	2	1	1	2	2	4	2	5	13	% CLIENTS UNDER 20 YEARS
25	27	37	35	25	25	27	28	24	30	24	30	14	% —"— 20-35 —"—
48	49	42	48	54	55	53	51	55	46	55	43	15	% —"— 35-55 —"—
25	22	17	15	19	19	19	19	19	20	19	22	16	% —"— MORE THAN 55 —"—
163	167	166	174	152	155	145	170	157	174	165	156	17	PRICE OF ROOMS
1.65	1.71	1.65	1.91	1.90	2.	1.54	1.60	1.73	1.82	1.66	1.44	18	LENGTH OF STAY
67	82	70	83	74	77	56	62	90	92	78	55	19	% OCCUPANCY
			X	X	X			X	X	X	X	20	CONVENTIONS

[Graphics and Graphic Information Processing, Bertin 81]

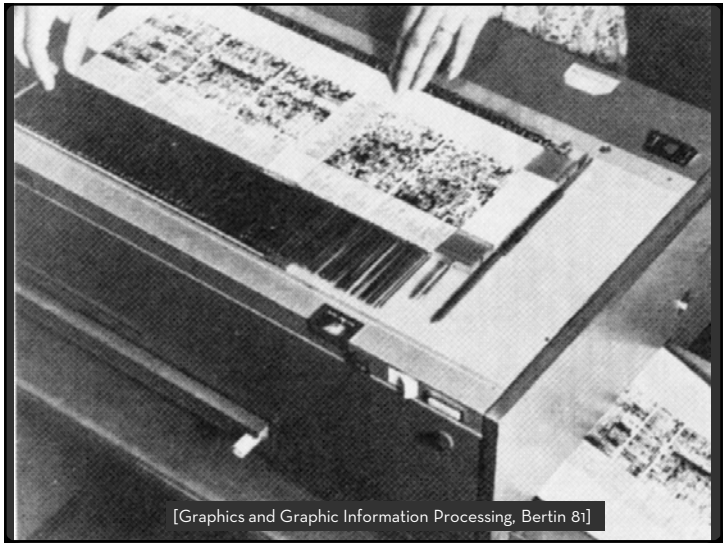


[Graphics and Graphic Information Processing, Bertin 81]

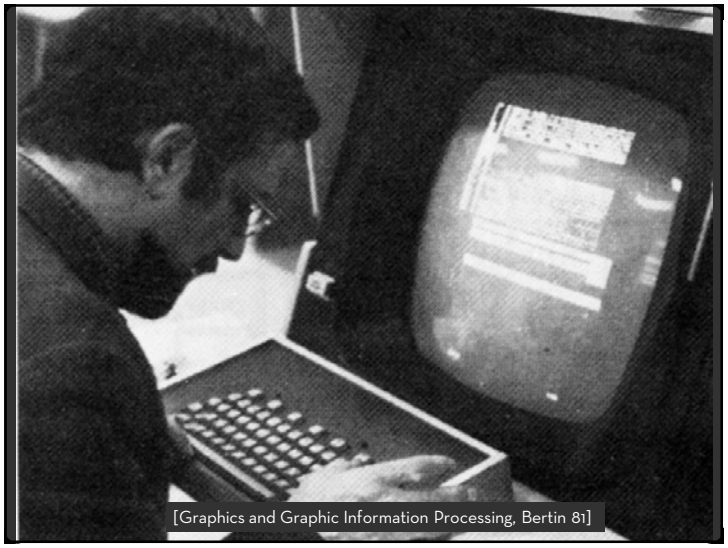


[Graphics and Graphic Information Processing, Bertin 81]

[Graphics and Graphic Information Processing, Bertin 81]

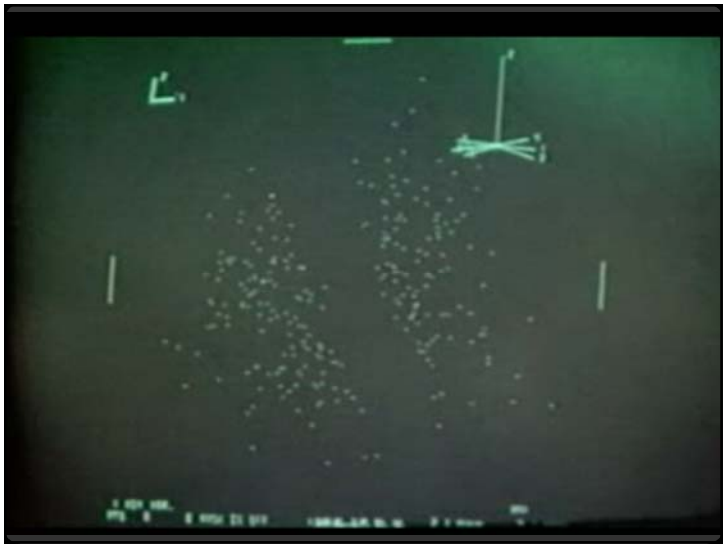


[Graphics and Graphic Information Processing, Bertin 81]



[Graphics and Graphic Information Processing, Bertin 81]

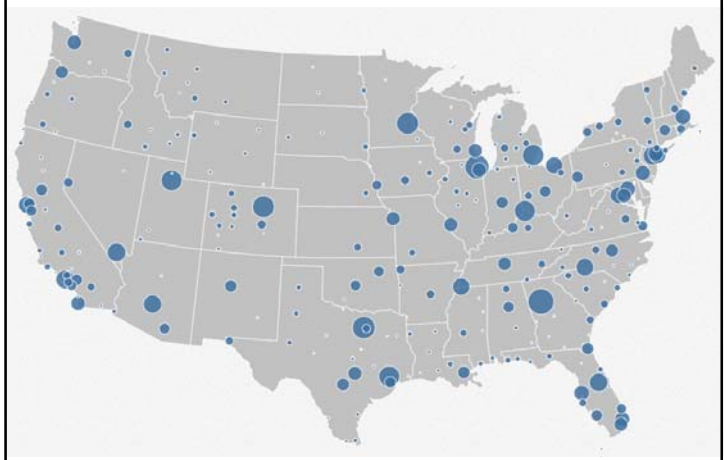
PRIM-9, Tukey, Fisherkeller, Friedman 1972

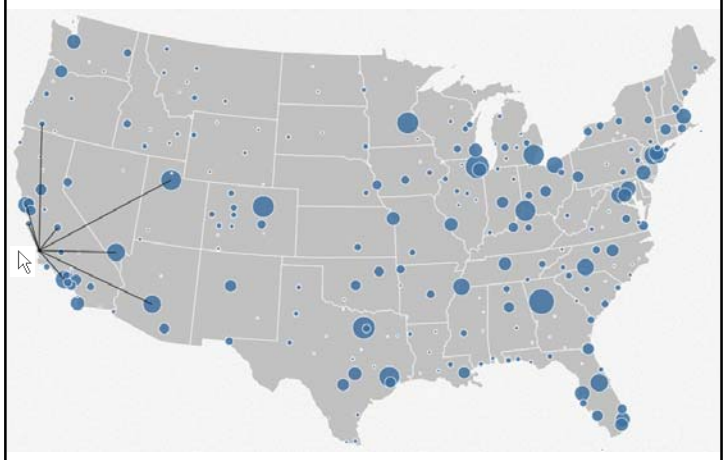


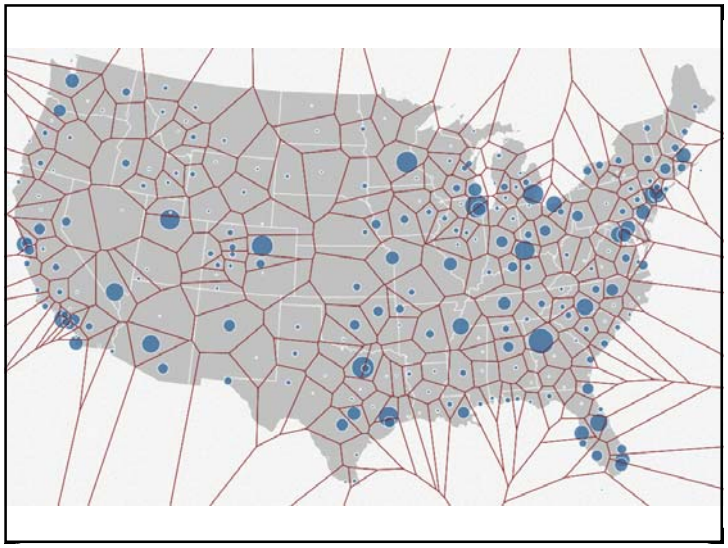
Pointing

Basic Pointing Methods

Point Selection
Mouse Hover / Click
Touch / Tap
Select Nearby Element (e.g., Bubble Cursor)







Basic Pointing Methods

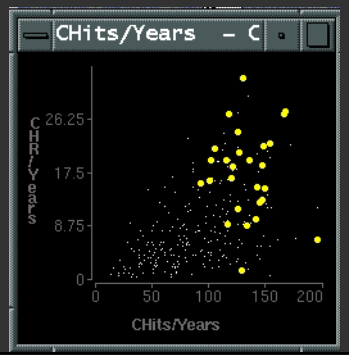
Point Selection
Mouse Hover / Click
Touch / Tap
Select Nearby Element (e.g., Bubble Cursor)

Region Selection
Rubber-band or Lasso
Area Cursors ("Brushes")

Brushing and Linking

Highlighting / Brushing

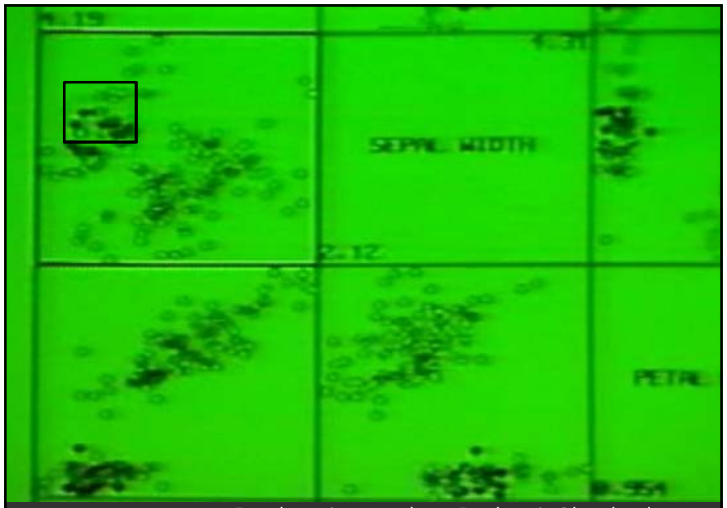
Direct attention to a data subset within a graph [Wills 95]



Brushing and Linking

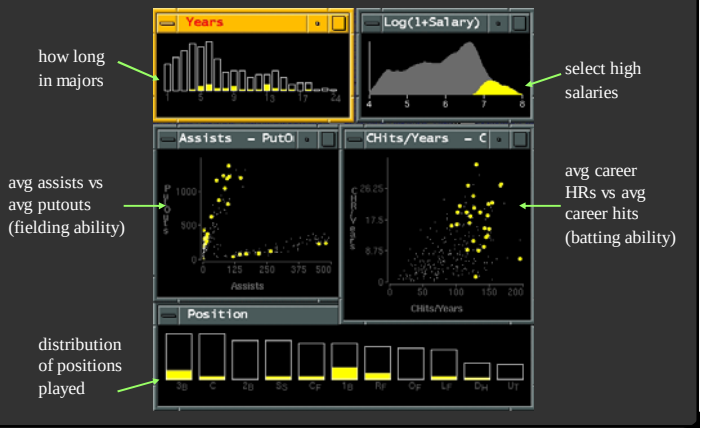
Select ("brush") a subset of data
See selected data in other views

The components must be *linked*
by *tuple* (matching data points), or
by *query* (matching range or values)

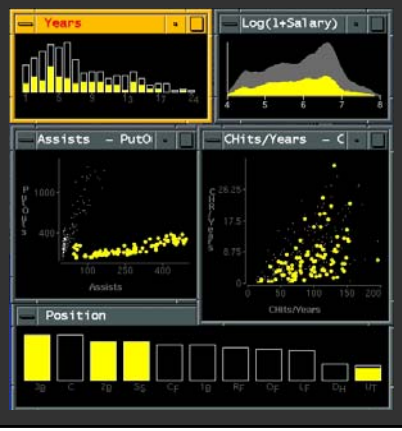


Brushing Scatterplots, Becker & Cleveland 1982

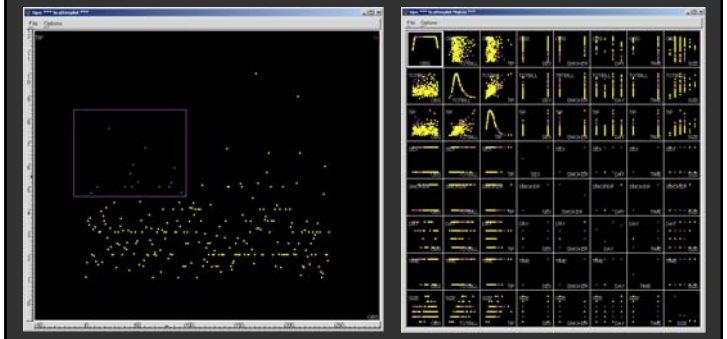
Baseball Statistics [from Wills 95]



Linking Assists to Positions



GGobi: Brushing

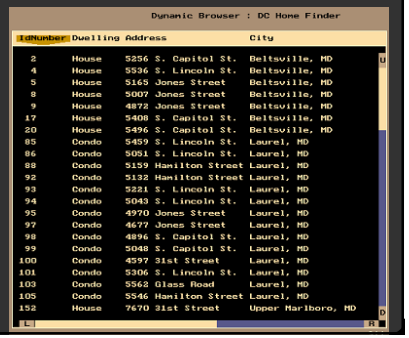


http://www.ggobi.org/

Dynamic Queries

Query and Results

SELECT house FROM palo_alto_homes
WHERE price < 1,000,000 AND bedrooms > 2
ORDER BY price



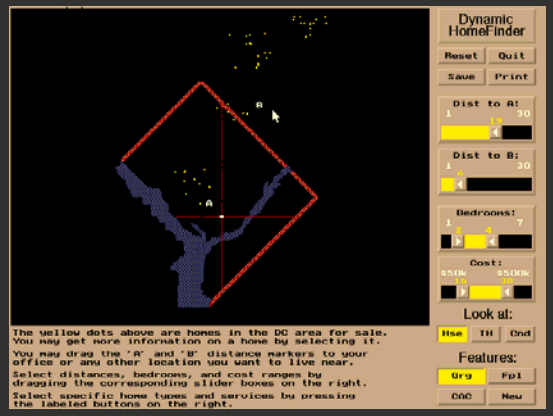
Dynamic Browser : DC Home Finder

IdNumber	Dwelling	Address	City
2	House	5256 S. Capitol St.	Beltsville, MD
4	House	5536 S. Lincoln St.	Beltsville, MD
5	House	5165 Jones Street	Beltsville, MD
8	House	5007 Jones Street	Beltsville, MD
9	House	4872 Jones Street	Beltsville, MD
17	House	5408 S. Capitol St.	Beltsville, MD
20	House	5496 S. Capitol St.	Beltsville, MD
85	Condo	5459 S. Lincoln St.	Laurel, MD
86	Condo	5051 S. Lincoln St.	Laurel, MD
88	Condo	5159 Hamilton Street	Laurel, MD
92	Condo	5132 Hamilton Street	Laurel, MD
93	Condo	5221 S. Lincoln St.	Laurel, MD
94	Condo	5043 S. Lincoln St.	Laurel, MD
95	Condo	4970 Jones Street	Laurel, MD
97	Condo	4677 Jones Street	Laurel, MD
98	Condo	4896 S. Capitol St.	Laurel, MD
99	Condo	5048 S. Capitol St.	Laurel, MD
100	Condo	4597 31st Street	Laurel, MD
101	Condo	5306 S. Lincoln St.	Laurel, MD
103	Condo	5563 Glass Road	Laurel, MD
105	Condo	5546 Hamilton Street	Laurel, MD
152	House	7670 31st Street	Upper Marlboro, MD

Issues

1. For programmers
2. Rigid syntax
3. Only shows exact matches
4. Too few or too many hits
5. No hint on how to reformulate the query
6. Slow question-answer loop
7. Results returned as table

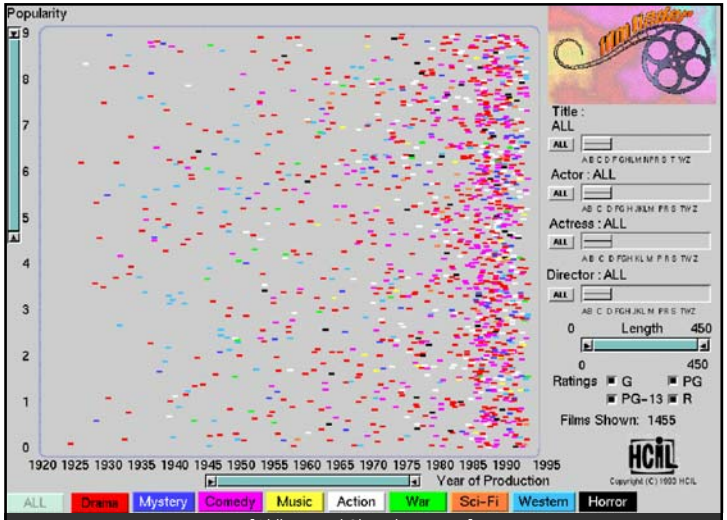
HomeFinder



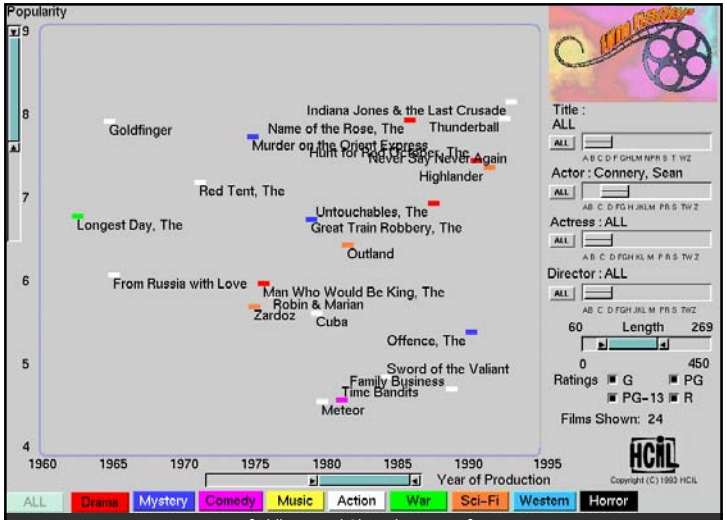
[Williamson and Shneiderman 92]

Direct Manipulation

1. Visual representation of objects and actions
2. Rapid, incremental and reversible actions
3. Selection by pointing (not typing)
4. Immediate and continuous display of results

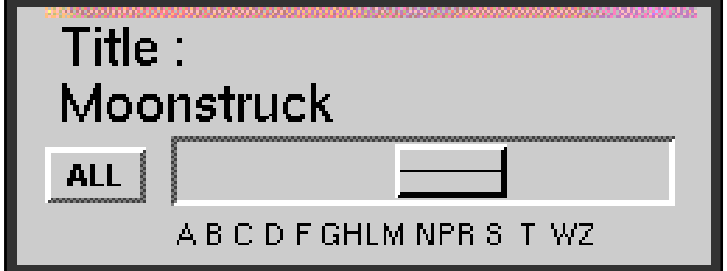


[Ahlberg and Shneiderman 94]

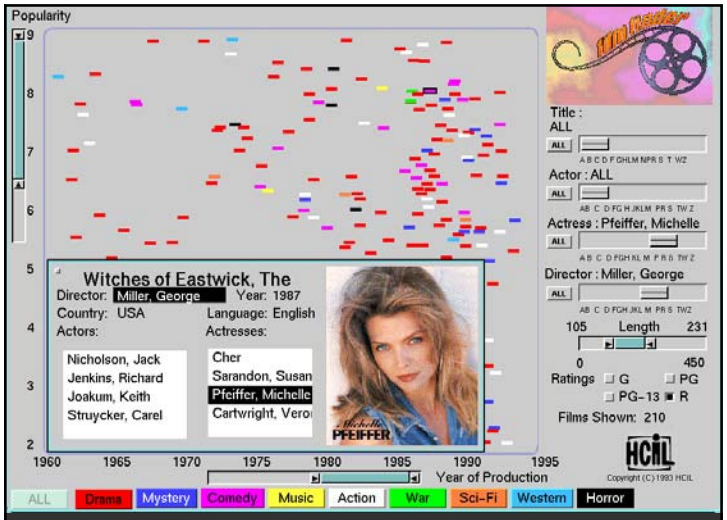


[Ahlberg and Shneiderman 94]

Alphaslider



[Ahlberg and Shneiderman 94]



[Ahlberg and Shneiderman 94]

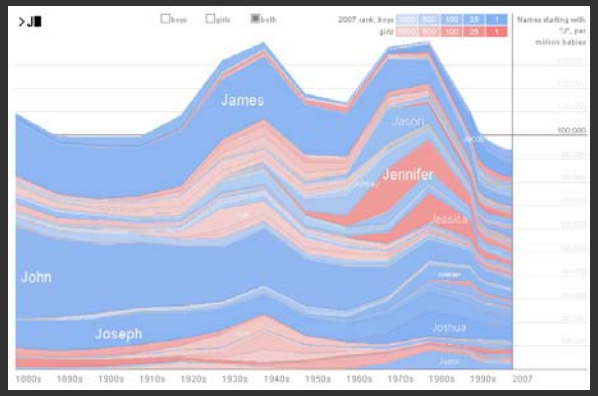
Attribute Explorer [Spence and Tweedie 98]

- Video Clip

Zipdecode [Fry 04]

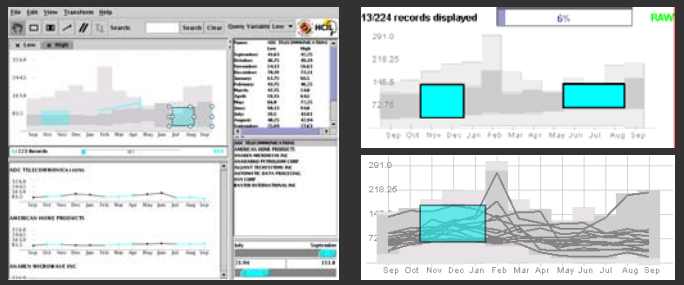
http://benfry.com/zipdecode/

NameVoyager



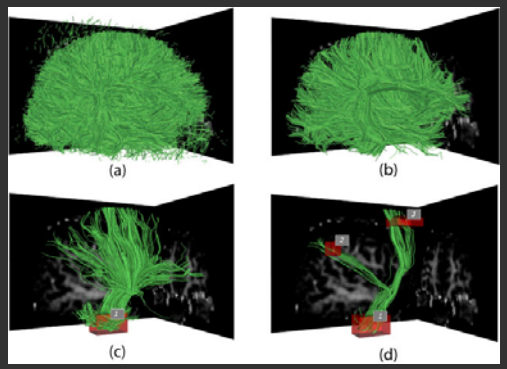
http://www.babynamewizard.com/voyager

TimeSearcher [Hochheiser & Shneiderman 02]

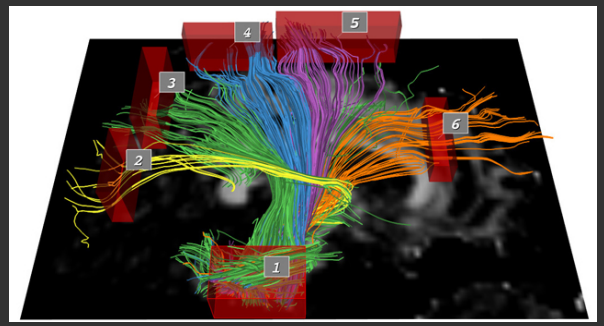


Based on Wattenberg's [2001] idea for sketch-based queries of time-series data.

3D dynamic queries [Akers et al. 04]



3D dynamic queries [Akers et al. 04]



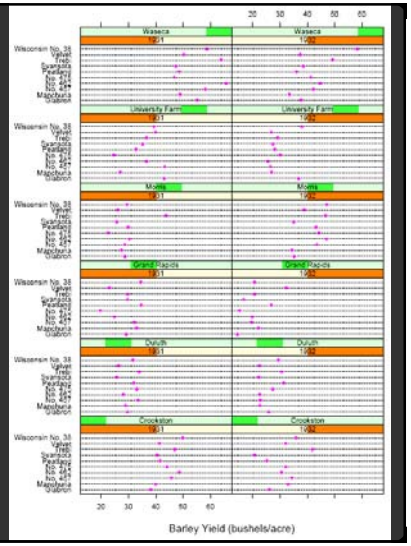
Pros and Cons

Pros
Controls useful for both novices and experts
Quick way to explore data

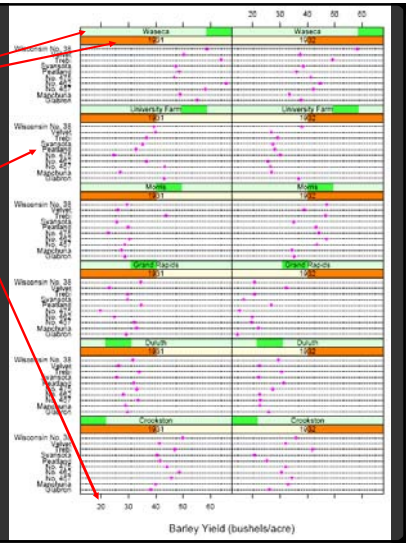
Cons
Simple queries
Lots of controls
Amount of data shown limited by screen space

Who would use these kinds of tools?

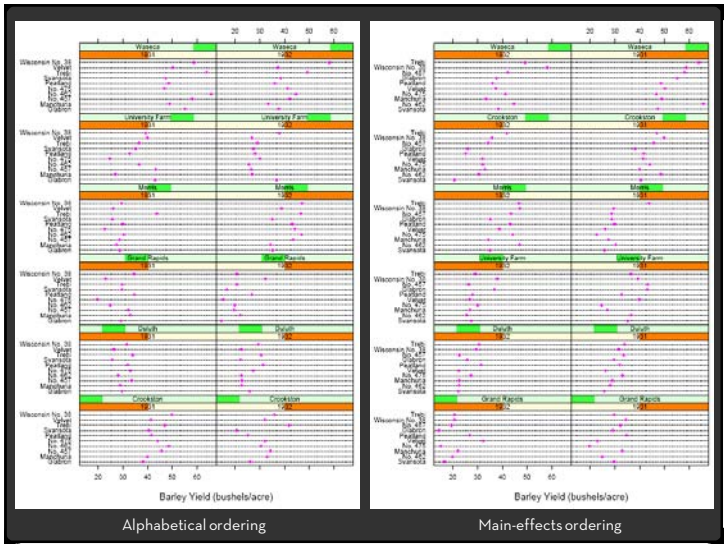
Sorting



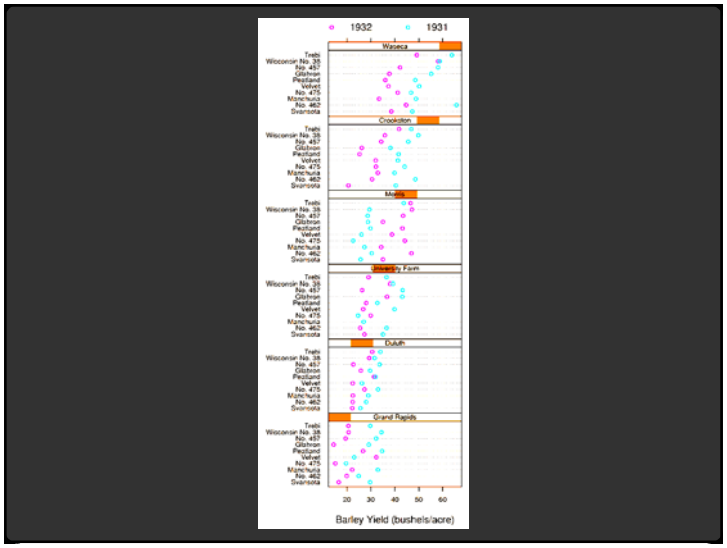
Trellis Display
[Becker, Cleveland, and Shyu 96]

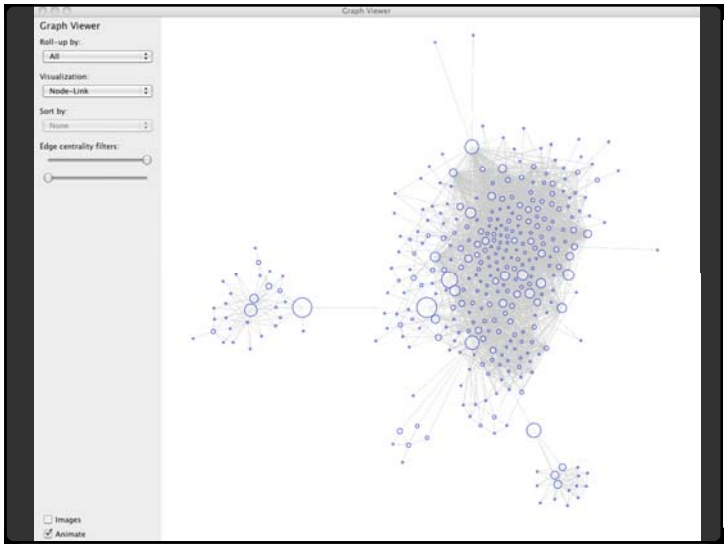


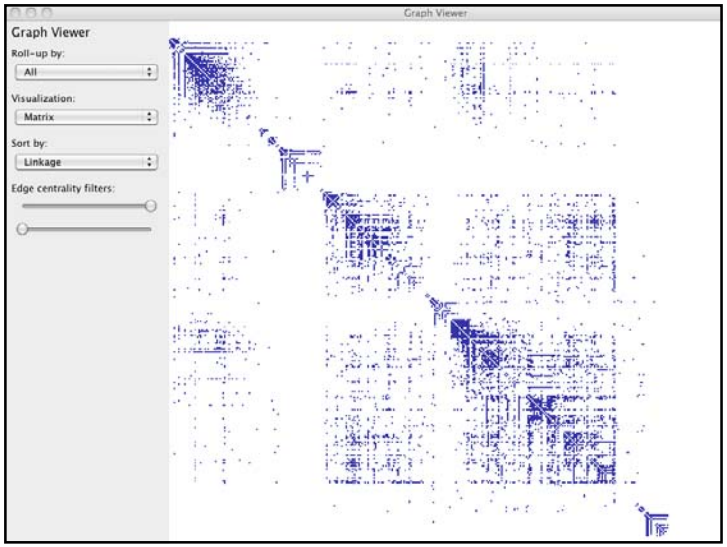
Trellis Display
[Becker, Cleveland, and Shyu 96]

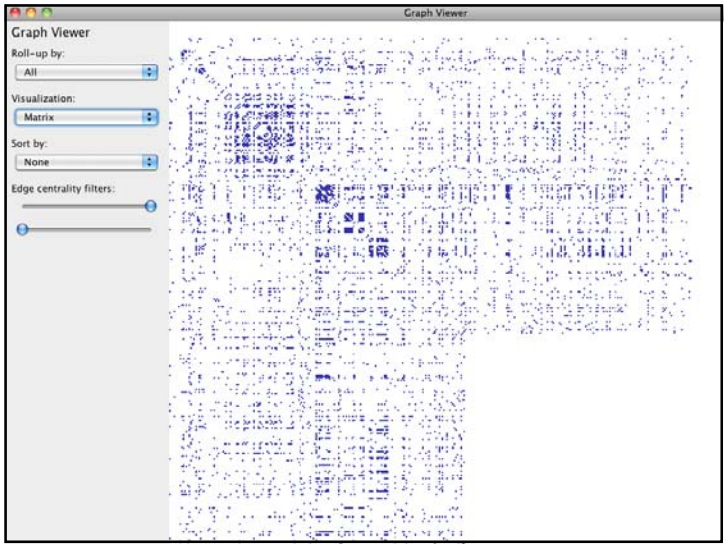


Alphabetical ordering Main-effects ordering









Administrivia

Assignment 2: Visual Data Analysis

Use visualization software (Tableau) to form & answer questions

First steps:

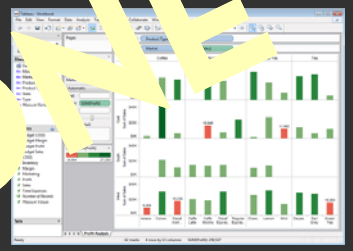
- Step 1: Pick a domain
- Step 2: Pose questions
- Step 3: Find Data
- Iterate

Create visualization

- Interact with data
- Refine your questions

Make working notebook

- Keep record of your analysis
- Prepare a final graphic and caption



DONE

Due by *1:00pm* on **Tuesday, October 16**

Assign 3: Interactive Visualization

Create an interactive visualization application. Choose a data domain and select an appropriate visualization technique.

1. Choose a data set and storyboard your interface
2. Implement the interface using tools of your choice
3. Submit your application and produce a final write-up

You should work in groups of 2.

Due by *end of day* on **Tuesday, October 30**

Assign 3: Project Partners

For A3, you should work in **groups of 2**.

If you do not have a partner, you should

1) Add your name + interests to the **ProjectPartners** wiki page
2) Stay after class to meet potential partners

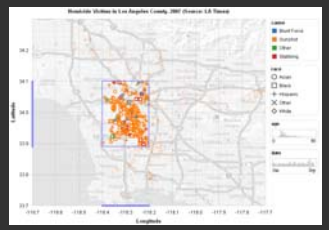
Assign 3 Tips

1) **Start now**. It will take longer than you think.
2) **Keep it simple**. (Eschew the kitchen sink.) Choose the minimal set of interactions that enables users to explore and generate interesting insights. Keep the design clean.
3) **Promote engagement**. How do your chosen interactions reveal interesting observations?

Generalized Selection

Visual Queries

Model selections as **declarative queries**.



↕

```
(-118.371≤lon AND lon≤-118.164)AND(33.915≤lat AND lat≤34.089)
```

Visual Queries

Model selections as **declarative queries** over the domain of visualized data.

Applicable to **dynamic, time-varying data**

Retarget selection **across visual encodings**

Perform operations on **query structure**

"Select items like this one."

Generalized Selection

Point to an example and define an abstraction based on one or more properties [Clark, Brennan]

"Blue like this"

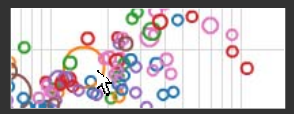
"The same shape as that"

Abstraction may occur over multiple levels

Generalized Selection

Provide *generalization mechanisms* that enable users to *expand a selection query* along *chosen dimensions* of interest

Expand selections via **query relaxation**



Interactor

↓

Query Builder

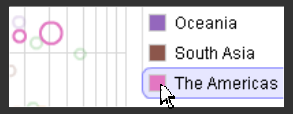
Query Builder

Click: Select Items

`(id = 'China')`

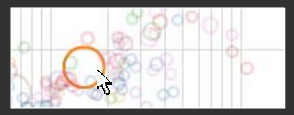
Drag: Select Range

`(2000 < gni AND gni < 10000) AND (.1 < internet AND internet < .2)`



Legend: Select Attributes

`(region = 'The Americas')`



Interactor

↓

Query Builder

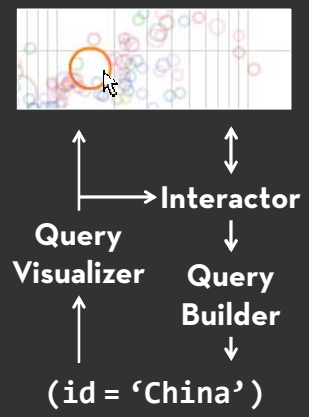
↓

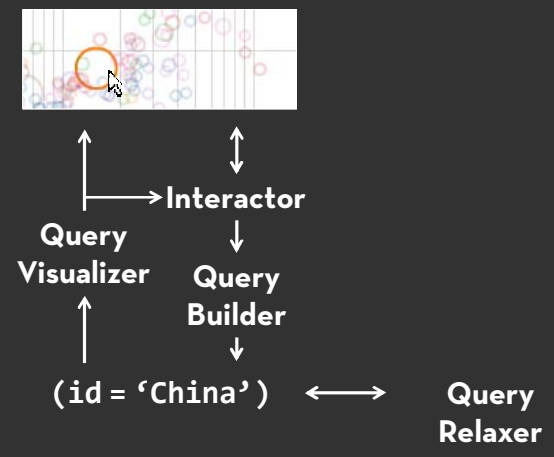
`(id = 'China')`

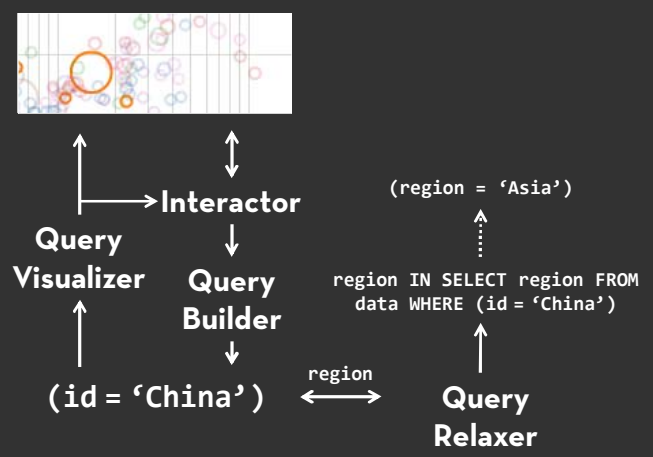
↑

Query Visualizer

↑



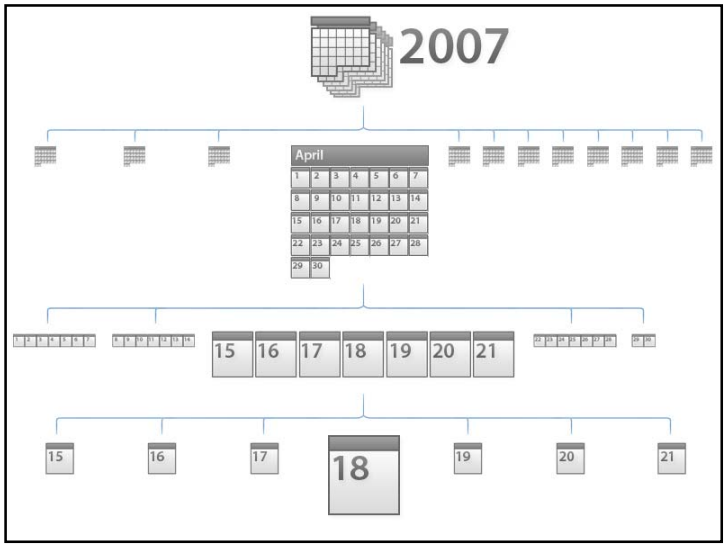




Query Relaxation

Generalize an input query to create an expanded selection, according to:

1. A *semantic structure* describing the data
2. A *traversal policy* for that structure

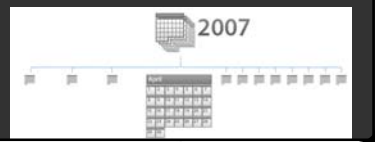


Relaxation using Hierarchies

Relax using **abstraction hierarchies** of the data
Traverse in direction of increasing generality

Examples

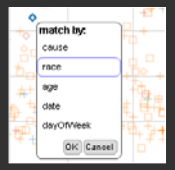
A Priori: Calendar, Categories, Geography
Data-Driven: Nearest-Neighbor, Clustering



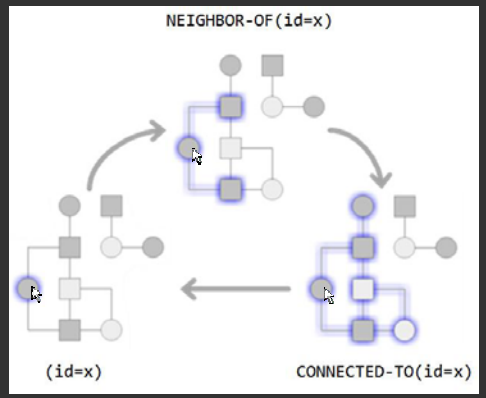
Relaxation using Attributes

If no explicit semantic structure is available, treat data itself as a "flat" hierarchy

Select all items with matching values along the attributes chosen for relaxation



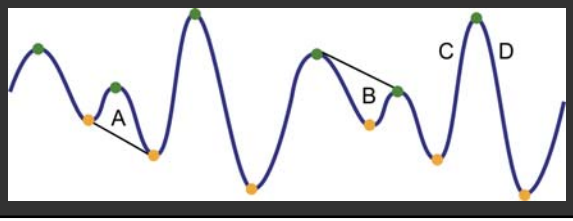
Relaxation of Networks



Lesson

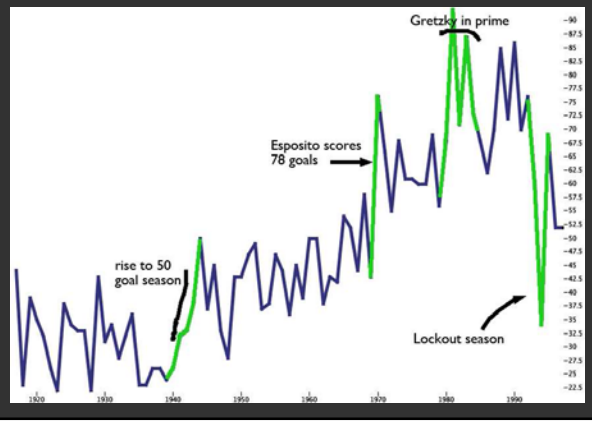
Consider how the structure and/or semantics of the data might be leveraged to aid analysis.

One extension: look beyond data features to incorporate perceptual features of the display.



Peaks, valleys, & slopes

Perceptual Annotation [Kong & Agrawala 09]



NHL Goals per Year

Summary

Most visualizations are interactive
Even passive media elicit interactions

Good visualizations are task dependent
Pick the right interaction technique
Consider the semantics of the data domain

Fundamental interaction techniques
Selection / Annotation, Sorting, Navigation, Brushing & Linking, Dynamic Queries