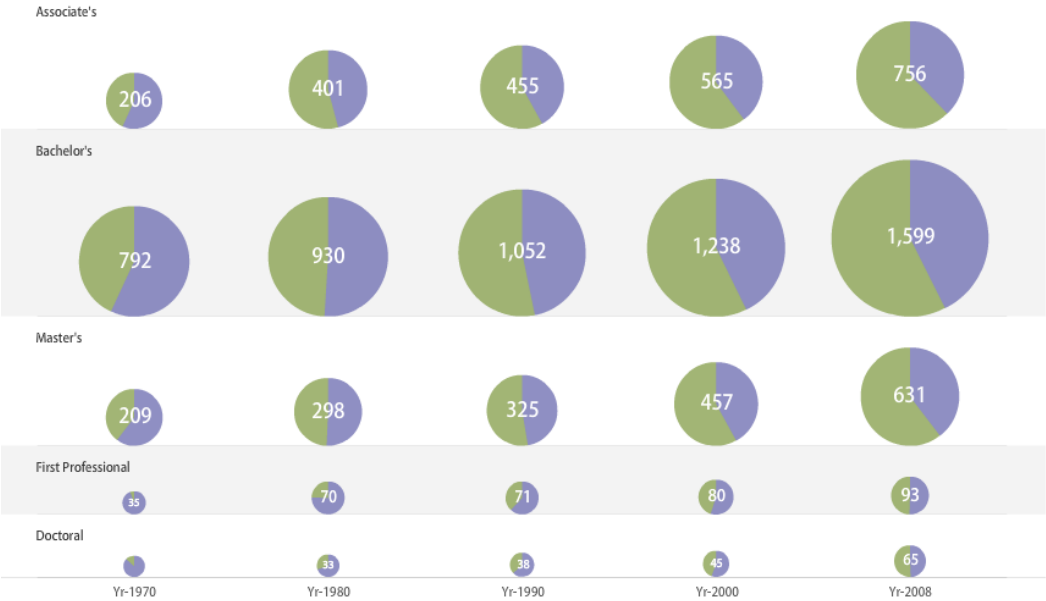
Accessible Analytics:
Complex Charts, Large Datasets, and Node Diagrams

Introduction

- Graphical visualizations communicate information simultaneously
- Compact
- Conceptualize
- Show relationships
  - Differences
  - Similarities
  - Connections
  - Patterns



- *What is the best mode to communicate the information to the blind?*

©2008 IBM Corporation

Accessible analytics - Outline

- Principles in approaches
- Specific approaches to several complex visualizations
  - In development, now
  - Current application environment: web
- Proposed approaches for other sticky problems
  - Need validation
  - Additional anticipated application environments:
    - Desktop solutions
    - Small devices (mobile)
- Next problems

©2008 IBM Corporation

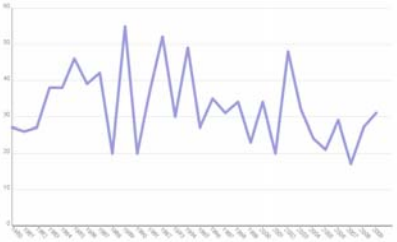
Principles of Complex Visualization Alternate Modes

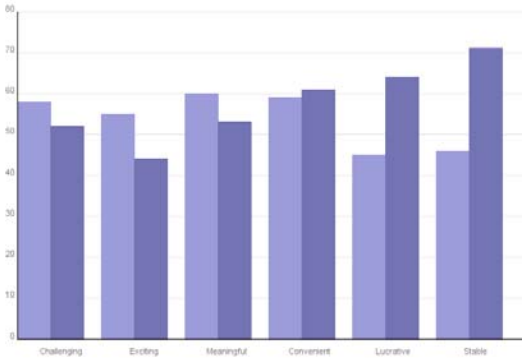
- Organization
- Navigable structure
- Filtering
- Sorting
- User interaction and query
- Mental mapping
- Automated analysis
- (Future) Automated conclusions

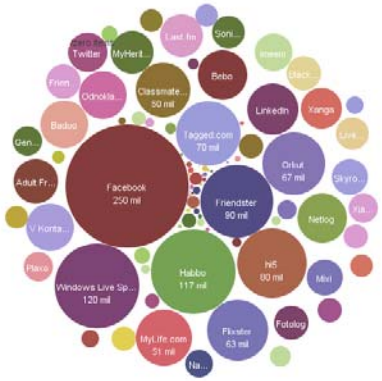
©2008 IBM Corporation

Types of complex visualizations

- Pie, Bar, and Scatter Charts, Line Graphs
- Tag and Word Clouds
- Bubble Charts
- Time Lines, Gantt Charts
- Node Diagrams, Org Charts
- Maps







©2008 IBM Corporation

Pie & Bar Charts, Scatter Plots, Line Graphs

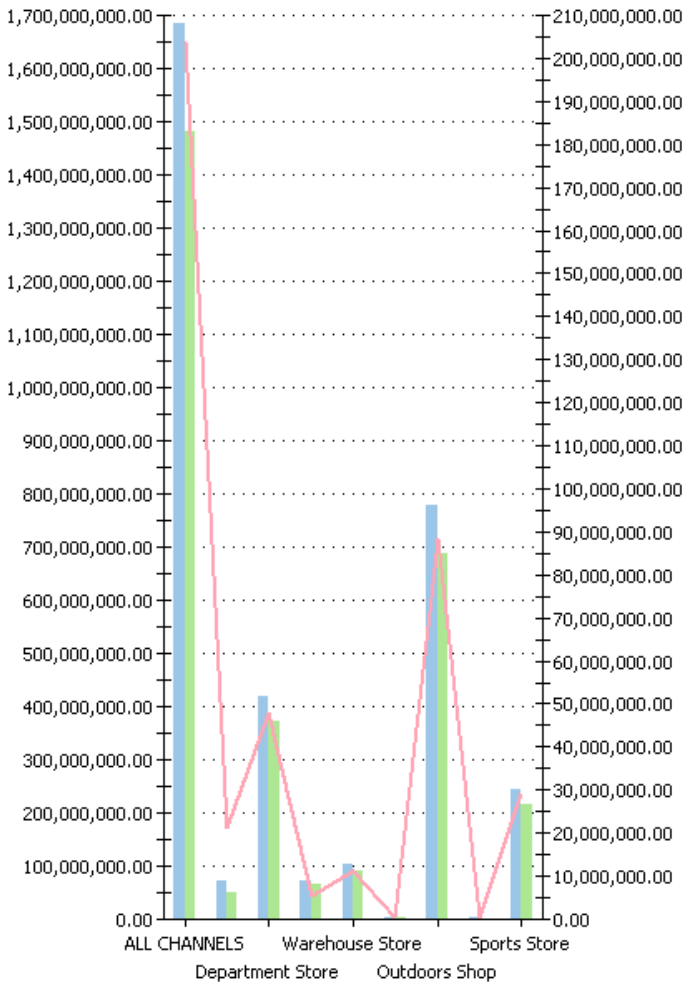
Per accessibility Guidelines, text should convey the information in the chart

- WGBH National Center for Accessible Media
  - Science, Technology, Engineering, Math (STEM) Guidelines
  - http://ncam.wgbh.org/experience_learn/educational_media/stemdx/intro
- Text alternative provides salient information
- Table of actual data as an alternative
  - Useful when all users need access to the detailed data
  - May be the only option for real-time charts

But even accessible tables are limited in their effectiveness

©2008 IBM Corporation

Graphics vs Tables



Subsidiaries : **TOTAL COMPANY** | Products : **TOTAL PRODUCTS**

	GROSS SALES REVENUE	NET SALES REVENUE	Gross-Net
ALL CHANNELS	1,683,410,998.65	1,479,704,578.29	203,706,420.36
Golf Shop	72,778,687.72	51,386,851.12	21,391,836.60
Department Store	418,974,311.23	371,247,324.00	47,726,987.23
Direct Marketing	72,325,384.53	67,029,056.09	5,296,328.44
Warehouse Store	103,105,808.36	92,055,344.40	11,050,463.96
Equipment Rental Store	1,897,345.68	1,661,281.51	236,064.17
Outdoors Shop	777,231,362.26	689,020,113.37	88,211,248.89
Eyewear Store	1,801,848.51	1,330,259.78	471,588.73
Sports Store	245,221,358.38	215,899,456.04	29,321,902.34

- ALL CHANNELS
- Golf Shop
- Department Store
- Direct Marketing
- Warehouse Store
- Equipment Rental Store
- Outdoors Shop
- Eyewear Store
- Sports Store

©2008 IBM Corporation

Tables as Alternatives to Bar Charts

Revenue		Camping Equipment	Golf Equipment	Outdoor Protection	Products
2004	Telephone	80,467,596.88	44,244,120.93	8,141,169.76	**178,793,580.36**
	Sales visit	62,069,990.87	12,046,323.14	6,455,901.27	**101,072,721.10**
	Web	125,829,519.92	49,583,401.41	13,735,716.85	**473,771,464.65**
	Special	5,844,211.12	3,705,129.72	858,775.41	**13,905,918.75**
	Mail	10,015,621.15	8,396,407.78	1,402,530.23	**22,766,850.51**
	E-mail	39,124,634.73	29,322,537.92	4,135,915.35	**95,402,796.21**
	Fax	9,634,763.39	6,255,930.08	1,435,512.20	**28,639,472.14**
	Order method	**332,986,338.06**	**153,553,850.98**	**36,165,521.07**	**914,352,803.72**
2005	Telephone	47,562,256.31	27,340,352.57	3,203,287.70	**107,160,284.09**
	Sales visit	43,352,874.97	7,720,215.67	2,672,442.62	**79,721,524.37**
	Web	270,463,415.88	116,939,694.38	16,479,270.80	**881,315,747.68**
	Special	5,012,986.47	1,259,633.25	267,356.86	**10,769,180.34**
	Mail	8,846,759.96	2,355,735.00	594,355.43	**16,013,779.49**
	E-mail	21,291,005.31	8,851,232.61	1,406,531.47	**44,318,886.43**
	Fax	6,228,274.27	3,539,563.59	385,329.20	**19,896,187.76**
	Order method	**402,757,573.17**	**168,006,427.07**	**25,008,574.08**	**1,159,195,590.16**
2006	Telephone	17,715,451.40	6,411,233.64	507,485.63	**37,199,842.80**
	Sales visit	37,660,178.39	10,562,039.68	746,934.80	**73,918,652.38**
	Web	426,353,675.75	203,385,896.61	8,570,078.91	**1,339,714,172.77**
	Special	332,095.44		10,798.96	**1,006,100.01**
	Mail	2,482,266.26	1,941,144.70	101,506.05	**6,905,730.44**
	E-mail	10,612,304.02	5,401,733.78	289,343.01	**23,701,042.57**
	Fax	5,226,451.57	2,408,222.14	123,028.48	**13,445,559.93**
	Order method	**500,382,422.83**	**230,110,270.55**	**10,349,175.84**	**1,495,891,100.90**

- Many rows and many columns
- Complex – cells have multiple column and row headers
- Numbers are quite large
- Screen reader users quickly become lost in the data
- Screen magnifier users lose context
- Can’t gain insight that chart provides

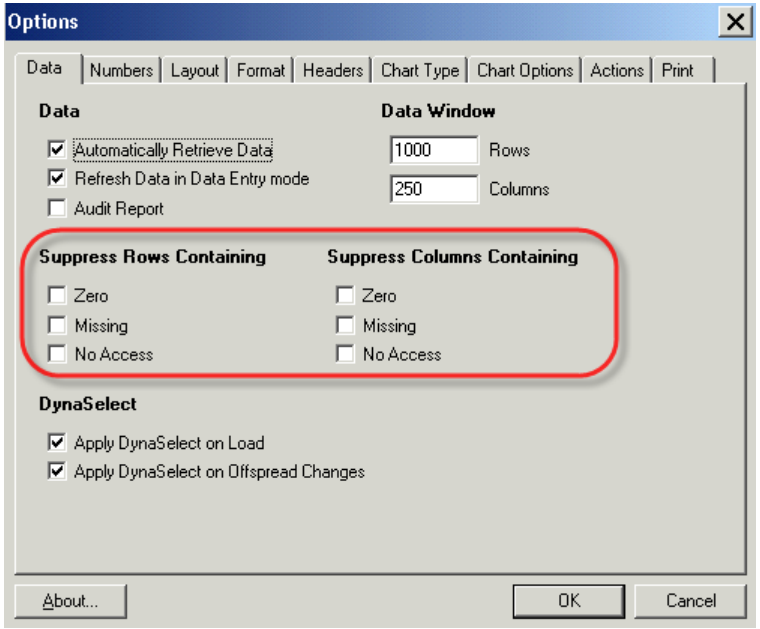
©2008 IBM Corporation

Principle: Provide filters to create a smaller view

- Ability to display a small set of rows or columns
- Ability to exclude rows or columns
- For sorted data:
  - ability to display every nth row

Filters are applied. See the Properties pane for more details.

Revenue		Products
2007	Telephone	17,832,073.81
	Sales visit	55,481,936.15
	Web	1,017,434,523.30
	Special	1,670,121.15
	Mail	404,978.53
	E-mail	16,420,318.95
	Fax	8,092,322.18
	Order method	**1,117,336,274.07**

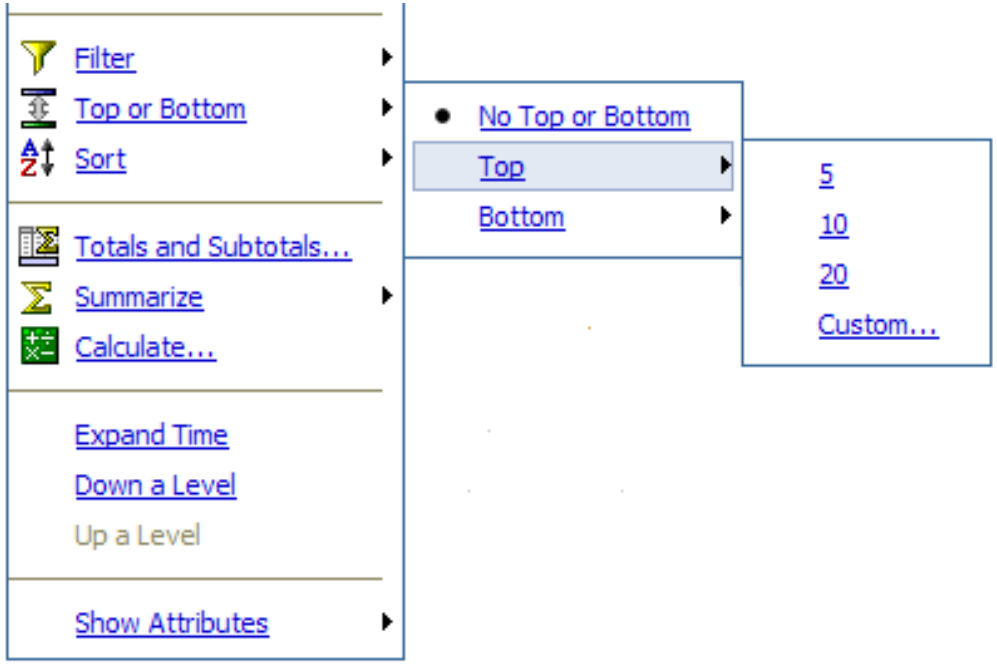


Source: IBM Cognos product

©2008 IBM Corporation

Principle: Provide interactive queries

- Find the row or column that contains the largest or smallest number.
- Find the column that has the largest rate of change from top to bottom.
- Find the set of rows that have numbers larger than 1,000,000.
- Find the 10 rows with the largest numbers



Source: IBM Cognos product

©2008 IBM Corporation

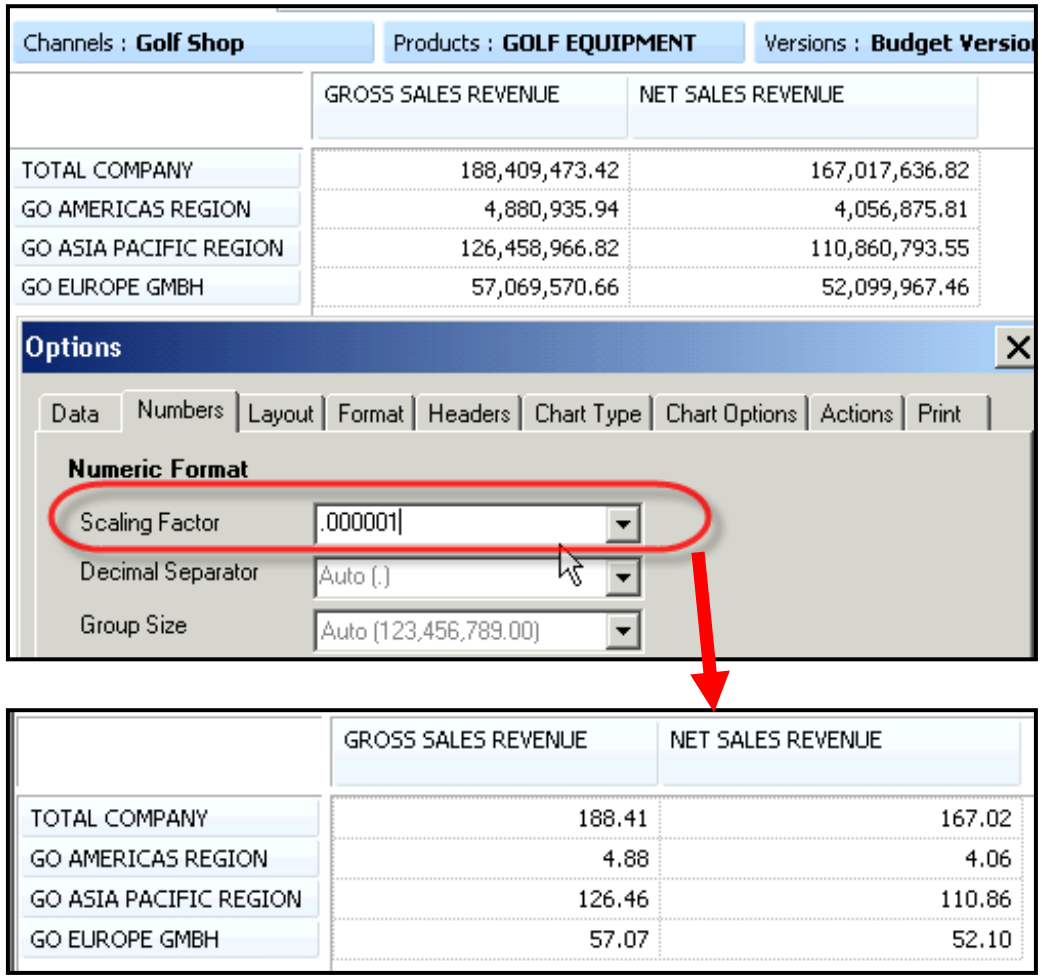
Principle: Reduce complexity with scaling

For large or similar numbers:

- Allow rounded numbers view
- User controlled
- Drill down in interesting data

Simulates sighted scanning of large sets of numbers.

Allows focus on the significant differences.



©2008 IBM Corporation

Tag Cloud Usability

Usable Tag Could features:

- Alternative views
- Enable filtering and sorting
- Enable search
- Limit number of tags displayed
- All tags legible
- Provide contextual information to screen readers
- Link to items of that tag
- Use CSS classes for font, size, etc. – no inline styles

Tagging

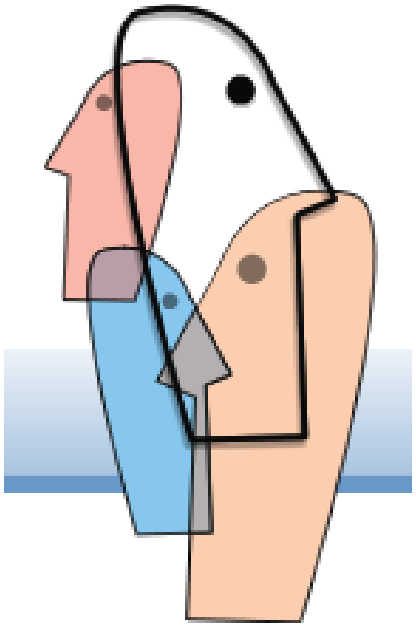
- Sensible, descriptive tags.
- Be clear and concise
- Use existing tags where possible

©2008 IBM Corporation

Interactive Text View - IBM Many-Eyes

Attempt to capture the sighted person's visual investigation

1. Summary
   - Data bounds and information
   - Rankings of selected subset of the view
   - (future) Analysis
2. Properties – manipulation and filtering of Visualizatic
3. Detailed Data - navigation of filtered data
4. Available for:
   1. Word charts
   2. Bubble charts
   3. Block histograms



Source: Stephen Levy, lead IBM Researcher

©2008 IBM Corporation

Word Clouds - *Wordles*

Text View

Summary

Number of words in dataset: 547
Number of displayed words: 150

[table of top 5 words]
[table of bottom 5 words]

Properties

Edit Properties

Undo Redo

Language Properties

Word Case Leave Words as Spelled Apply

Language Properties Remove common English Words Apply

Word Layout Properties

Prefer Alphabetical Order Apply

Maximum words to layout: 150 Apply

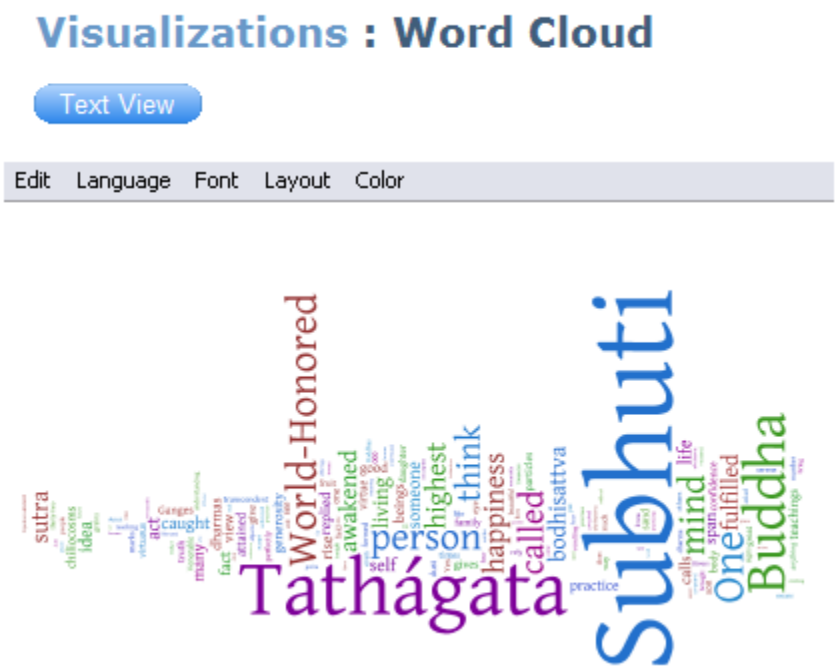
Additional words to remove (separate with commas): Apply

Details

Detailed Data Sort Weight Descending Apply

First Previous Displaying row 1 thru 20 of 150 Next Last

Word	Weight
8217	1000
8220	448



©2008 IBM Corporation

Wordles Example

Text View of Word Cloud Generator

Summary

Number of words in dataset: 819
Number of displayed words: 150

Top 5 Words

Order or Occurrence	Word	Weight
1	Subhuti	1000
2	Tathágata	669
3	Buddha	484
4	One	376
5	World-Honored	353

Bottom 5 Words

Order or Occurrence	Word	Weight
146	different	30
147	endurance	30
148	greater	30
149	heard	30
150	immeasurable	30

Properties

Edit Properties
Undo Redo

Language Properties
Word Case
Leave Words as Spell Apply

Language Properties
Remove common Eng Apply

Word Layout Properties
Prefer Alphabetical Order Apply

Maximum words to layout:
150 Apply

Additional words to remove (separate with commas):
Apply

Detailed Data

Detailed Data Sort Weight Descending Apply

First Previous Displaying row 1 thru 20 of 150 Next Last

Word	Weight
Subhuti	1000
Tathágata	669
Buddha	484
One	376
World-Honored	353
person	330
mind	315
called	292
think	292
happiness	246
highest	238
awakened	207
sutra	207
bodhisattva	200
fulfilled	200
living	184
life	161
act	153
idea	153
self	153

©2008 IBM Corporation

Bubble Chart

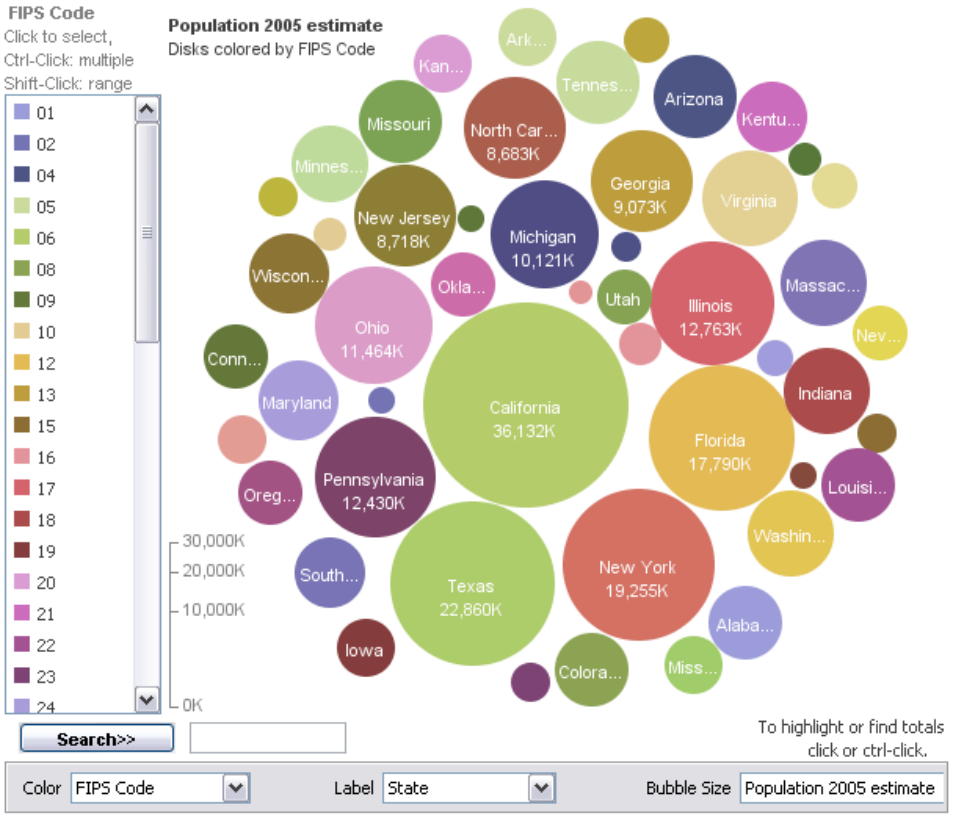
Detailed Data

Detailed Data Sort

Population 2005 estim | Descending | Apply

First | Previous | Displaying row 1 thru 20 of 50 | Next | Last

Selected	State	Population 2005 estimate	Aggregation	Distribution
☐	California	36,132K	1 item.	36,132K in 06
☐	Texas	22,860K	1 item.	22,860K in 48
☐	New York	19,255K	1 item.	19,255K in 36
☐	Florida	17,790K	1 item.	17,790K in 12
☐	Illinois	12,763K	1 item.	12,763K in 17
☐	Pennsylvania	12,430K	1 item.	12,430K in 42
☐	Ohio	11,464K	1 item.	11,464K in 39
☐	Michigan	10,121K	1 item.	10,121K in 26
☐	Georgia	9,073K	1 item.	9,073K in 13
☐	New Jersey	8,718K	1 item.	8,718K in 34
☐	North Carolina	8,683K	1 item.	8,683K in 37
☐	Virginia	7,567K	1 item.	7,567K in 51
☐	Massachusetts	6,399K	1 item.	6,399K in 25
☐	Washington	6,288K	1 item.	6,288K in 53
☐	Indiana	6,272K	1 item.	6,272K in 18
☐	Tennessee	5,963K	1 item.	5,963K in 47
☐	Arizona	5,939K	1 item.	5,939K in 04
☐	Missouri	5,800K	1 item.	5,800K in 29



©2008 IBM Corporation

Bubble Chart Interaction

Detailed Data

Detailed Data Sort

Population 2005 estim | Descending | Apply

First | Previous | Displaying row 1 thru 20 of 50 | Next | Last

Selected	State	Population 2005 estimate	Aggregation	Distribution
☑	California	36,132K	1 item.	36,132K in 06
☑	Texas	22,860K	1 item.	22,860K in 48
☑	New York	19,255K	1 item.	19,255K in 36
☑	Florida	17,790K	1 item.	17,790K in 12
☑	Illinois	12,763K	1 item.	12,763K in 17
☐	Pennsylvania	12,430K	1 item.	12,430K in 42
☐	Ohio	11,464K	1 item.	11,464K in 39

Apply Bubble Selections

Summary

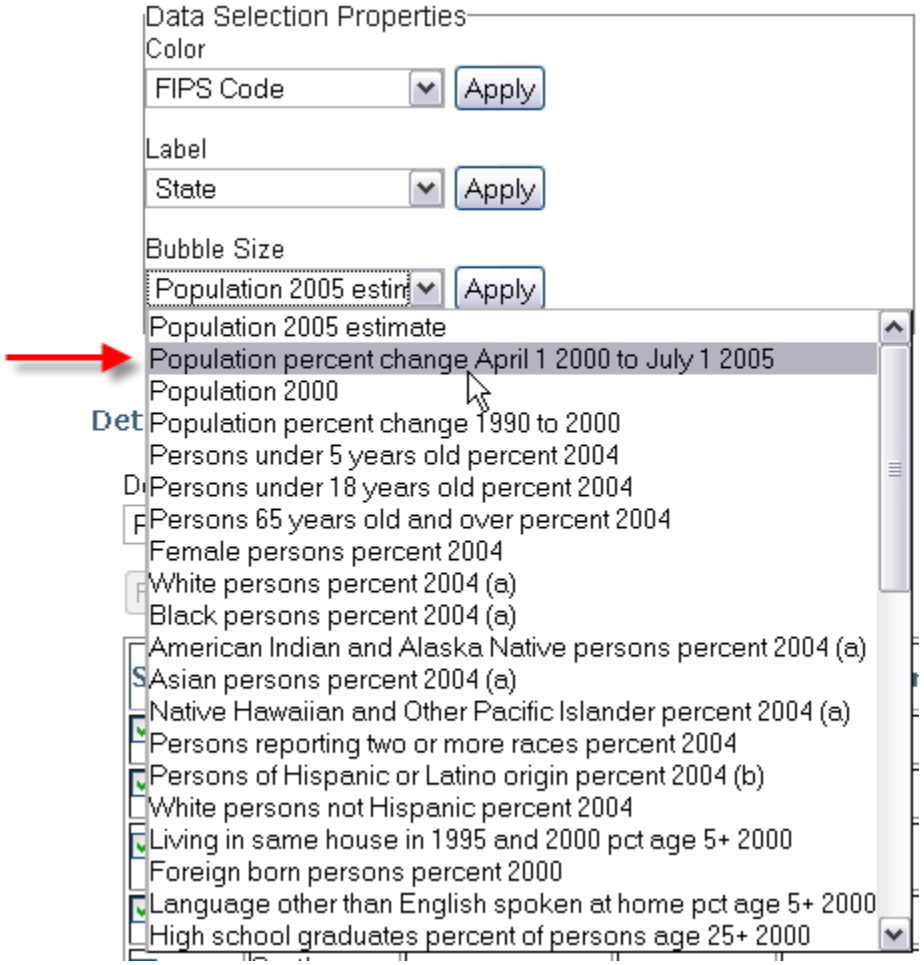
Number of Bubbles: 50
Bubble size based on: Population 2005 estimate
Disks colored by FIPS Code
Total Population 2005 estimate selected: 108,800K = 36%
Number of selected bubbles: 5

Selected Bubbles

State	Population 2005 estimate	Aggregation	Distribution
California	36,132K	1 item.	36,132K in 06
Florida	17,790K	1 item.	17,790K in 12
Texas	22,860K	1 item.	22,860K in 48
Illinois	12,763K	1 item.	12,763K in 17
New York	19,255K	1 item.	19,255K in 36

©2008 IBM Corporation

Bubble Chart Interaction (2)



Summary

Number of Bubbles: 50
Bubble size based on: Population percent change April 1 2000 to July 1 2005
Disks colored by FIPS Code
Total Population percent change April 1 2000 to July 1 2005 selected: 0.33
Number of selected bubbles: 5

Selected Bubbles

State	Population percent change April 1 2000 to July 1 2005	Aggregation	Distribution
California	0.07	1 item.	0.07 in 06
Florida	0.11	1 item.	0.11 in 12
Texas	0.10	1 item.	0.10 in 48
Illinois	0.03	1 item.	0.03 in 17
New York	0.02	1 item.	0.02 in 36

©2008 IBM Corporation

Histogram

Text View of Block Histogram

Summary

Item measure based on: Persons 65 years old and over percent 2004
Item measure range: 0.06 - 0.18
Item count based on: State
Item count range: 0 - 16
Number of selected items: 8

Selected items

State	Persons 65 years old and over percent 2004
Wyoming	0.12
Arizona	0.13
Colorado	0.10
Montana	0.14
Nevada	0.11
Idaho	0.11
Utah	0.09
New Mexico	0.12

Properties

Search Properties
Name of item
Search

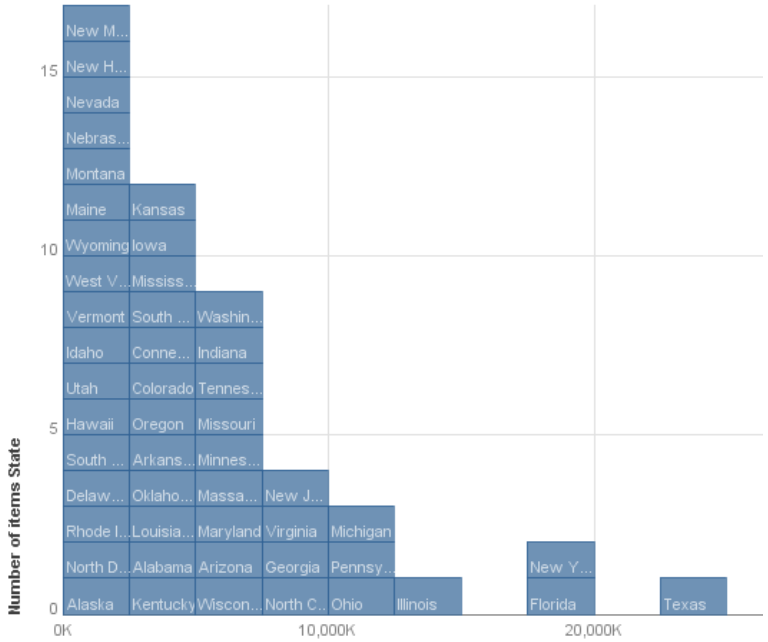
Data Selection Properties
Values
Persons 65 years old Apply

Labels
State Apply

Visualizations : Histogram

Text View

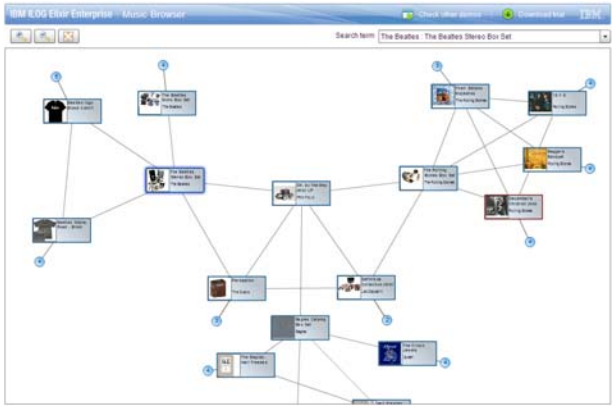
Uploaded by: Anonymous
Created at: Mar 1 2011
Description:

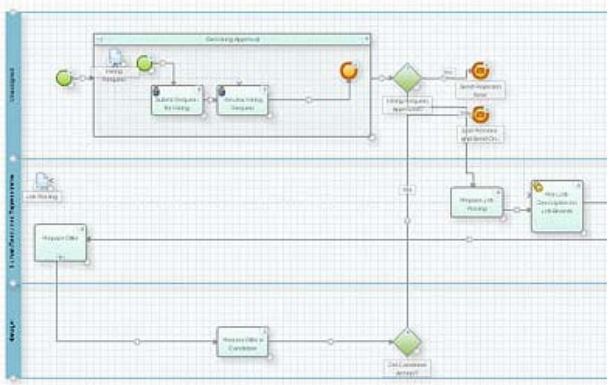


©2008 IBM Corporation

Node Diagrams

- Issues
  - Keyboard Navigation
    - No clear standards
    - User needs to know what the choices are without losing focus of current object.
    - Transition (the arrow) may have values / costs / limitations associated
    - Semantics for screen readers
  - Swim lanes
    - May communicate an attribute
    - May be inferred by organization and layout





©2008 IBM Corporation

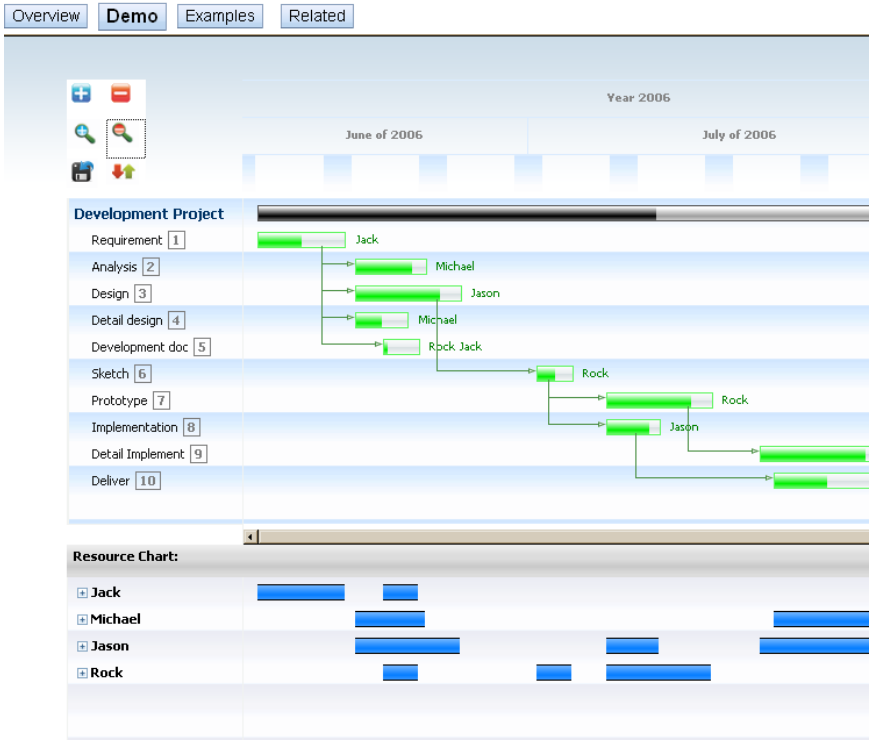
Node Diagrams

- Approaches
  - Stepping stone model for navigation
    - Focus remains on current object
    - Information provided about choices, for example, in a list
    - User selects the choice and moves the focus
  - Tree diagram has potential
    - What to do with nodes that return to the tree?
    - Need hyperlinks to “jump” to other points in tree
    - Need easy identification of the place “jumped to”

©2008 IBM Corporation

Timelines and Gantt Charts

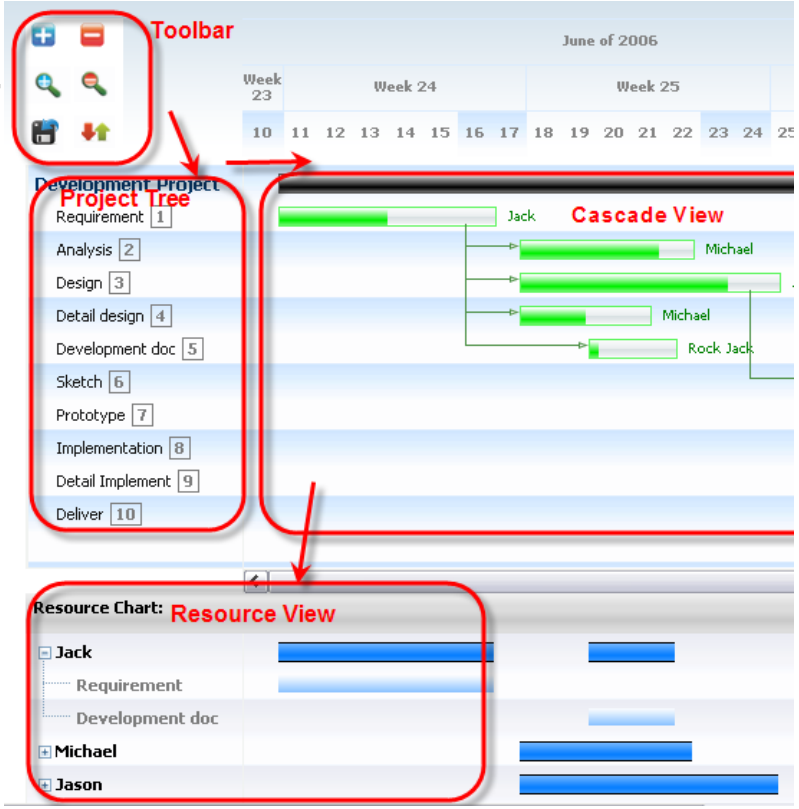
- Problems
  - Complex dependencies
    - Multiple inputs / outputs
    - Indication for critical path
  - Potential independent components not in critical path
  - Keyboard Navigation
    - No standards
    - User needs to know what the choices are without losing focus of current object.



©2008 IBM Corporation

Timelines and Gantt Charts (cont)

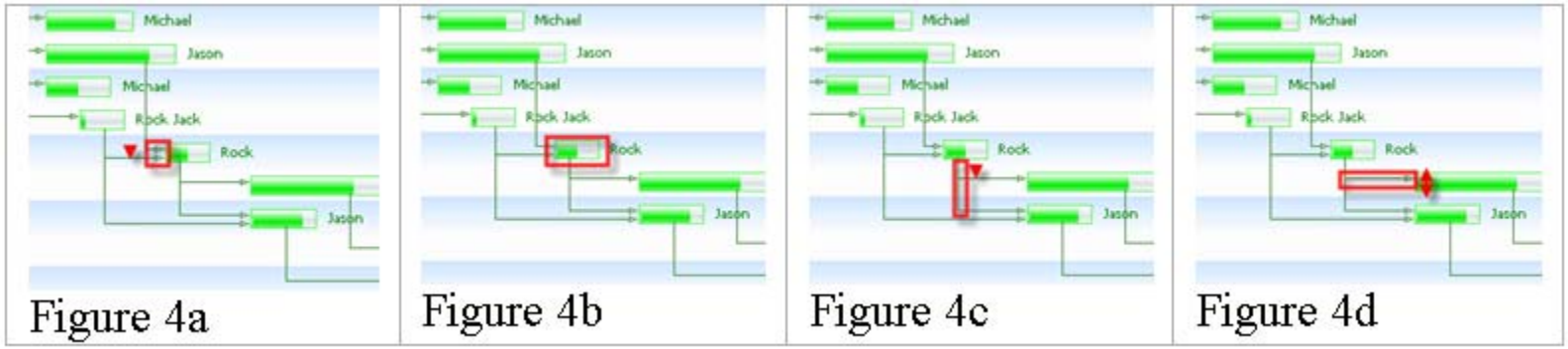
- Approach
  - Three widgets working together
    - Project tree
    - Cascade view
    - Resource view
  - Stepping stone navigation in cascade
  - Focus synchronized between project tree and cascade view, but not resource view
  - Icons and earcons for critical path



©2008 IBM Corporation

Timelines and Gantt Charts (continued)

- Each task has three focus areas:
  - Predecessors / task / successors
  - Each is a ARIA role="button"
  - Right / left arrow to navigate
- Successor (and predecessor) button opens list of links
  - Arrows to listen to successors, Enter to select
  - Circular design allows user to traverse entire cascade with right arrow
  - Maps to familiar ARIA navigation models



©2008 IBM Corporation

Timelines and Gantt Charts – Previous approach

- Tabular data format
- Hyperlinks for relationships
- Indication of critical path

Resource	Prereqs	Crit. Path	Event	Plan Begin	Actual	Duration	Overlap	Post Activities	Status
Kitchen		$	K01 Remove kitchen flooring	Mon 5/08 8am	Mon 5/08 8am	8 hrs		K02	Completed on time
Kitchen	K01	$	K02 Paint kitchen	Tues 5/09 8am	Tues 5/09 8am	8 hrs		K03	Completed on time
Kitchen	K02	$	K03 Kitchen dry	Wed 5/10 8am	Wed 5/10 8am	4 hrs		K04	Completed on time
Kitchen	K03	$	K04 Kitchen subfloor	Wed 5/10 1pm	Wed 5/10 2pm	8 hrs	C	K05	Behind! Must work overtime to catch up.
Kitchen	K04	$	K05 Kitchen tile	Thurs 5/11 1pm		8 hrs	D, E	K06	
Kitchen	K05	$	K06 Kitchen no walk wait	Fri 5/12 1pm		24 hrs	E, A	K07	
Bedrooms			Paint bedrooms	Fri 5/12 1pm		8 hrs	A		
Kitchen	K06	$	K07 Kitchen grout	Mon 5/15 8am		4 hrs	F	K08	
Kitchen	K07	$	K08 Kitchen grout cure wait	Mon 5/15 1pm		24 hrs	B	K09	
Kitchen	K08	$	K09 Kitchen grout clean up	Tues 5/16 1pm		4 hrs	G	K11	
Kitchen			K10 Kitchen lighting	Mon 5/15 1pm		4 hrs	B		
Kitchen	K09	$	K11 Kitchen grout sealant	Wed 5/17 8am		2 hrs			
Cabinets			C01 Cabinets cut wood	Wed 5/10 1pm	Wed 5/10 1pm	8 hrs	C	C02	Started on time.
Cabinets	C01		C02 Cabinets drill holes	Thu 5/11 1pm		2 hrs	D	C03	
Cabinets	C02		C03 Cabinets assemble	Fri 5/12 8am		8 hrs	E	C04	
Cabinets	C03		C04 Cabinets glue dry			24 hrs		C05	
Cabinets	C04		C05 Cabinets sand	Mon 5/15 8am		2 hrs	F	C06	

©2008 IBM Corporation

Maps

- Turn by turn works for directional.
- Next hard problems:
  - What works for a singular position?
  - Travel from a known point?
  - Multiple items of interest?
  - Relationships / distances matrix?

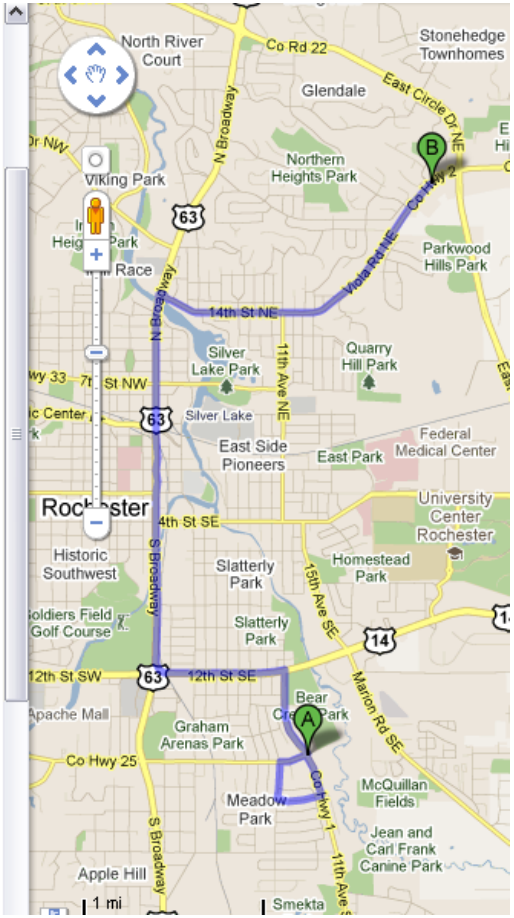
Driving directions to 2100 Viola Heights Dr NE, Rochester, MN 55906

Suggested routes

1. S Broadway **6.0 mi**	**12 mins**
2. US-14 E and Co Rd 22 5.6 mi	13 mins
3. S Broadway and Co Rd 22/East Circle Dr NE 7.5 mi	13 mins

1601 11th Ave SE
Rochester, MN 55904

1. Head **southeast** on **11th Ave SE** toward **16 1/2 St SE**	0.2 mi
2. Turn right at **16 1/2 St SE**	0.3 mi
3. Turn right at **9th Ave SE**	0.2 mi
4. Take the 2nd right onto **16th St SE**	0.2 mi
5. Take the 1st left onto **11th Ave SE**	0.5 mi
6. Turn left at **12th St SE**	0.7 mi
7. Turn right at **S Broadway**	2.1 mi
8. Turn right at **14th St NE**	0.9 mi



©2008 IBM Corporation

Next hard problems

- Automated importance
- Dense data
- Aggregations
- Automated analysis
- Automated descriptive text

©2008 IBM Corporation

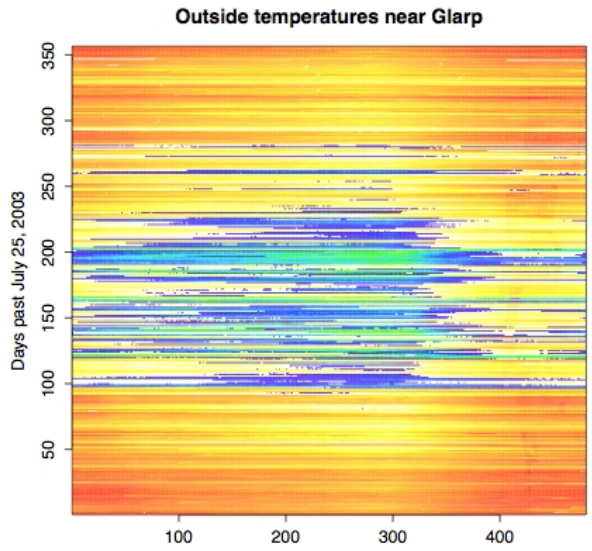
Future - Automated Importance Identification

- Identify significant data
  - High and low points
  - Deviations from trends
- Identify significant relationships
  - Largest differences
  - Similar data

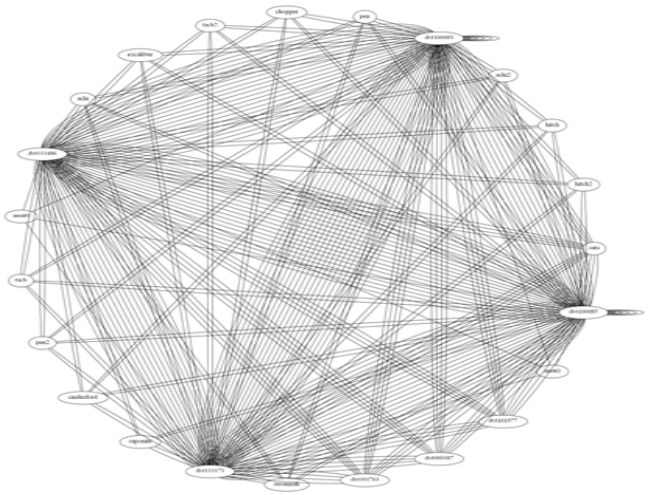
©2008 IBM Corporation

Dense data and Aggregations

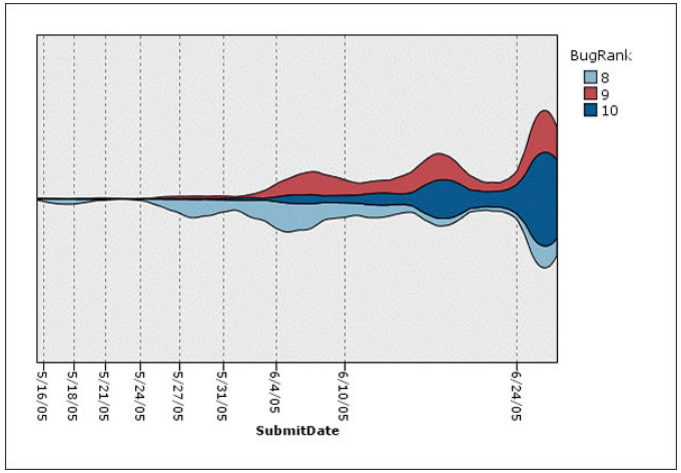
- Find visual patterns
- Repetition and cycles
- Filter data
- Map data to other modes



Daily Temperature



Dense data



Bug discovery

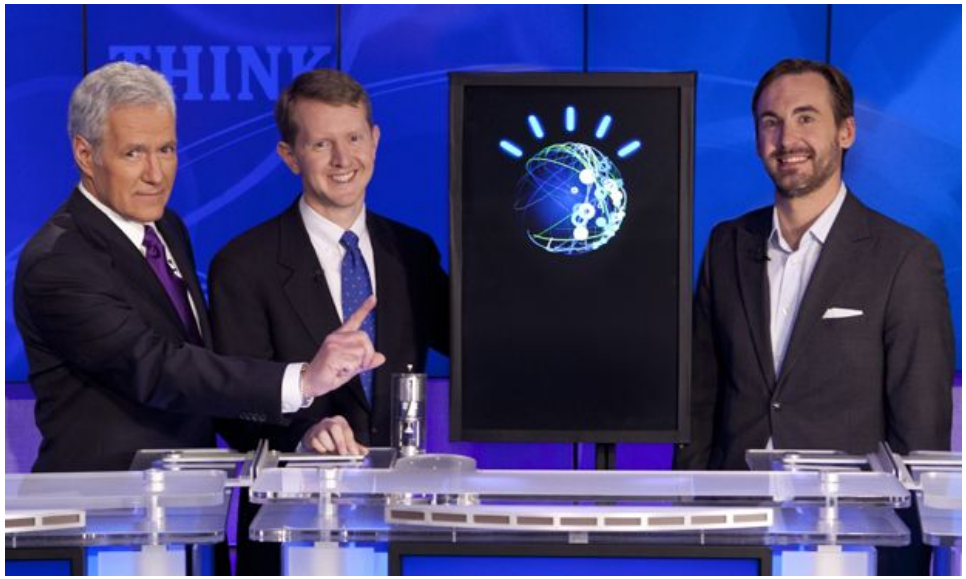
©2008 IBM Corporation

Automated analysis (analytics) + Descriptive text

Find correlations, relationships, in the data
Find the meaning *for* you.

Utilize advanced pattern matching to find similarities, differences, trends in massive data.

Express results in language.



©2008 IBM Corporation

Acknowledgements

- James Taylor
- Robert Kinsman
- Mary Jo Mueller
- Christophe Jolif
- Stephen Levy
- Susann Keohane

©2008 IBM Corporation

References

- WGBH STEM Guidelines
  - http://ncam.wgbh.org/experience_learn/educational_media/stemdx/intro

©2008 IBM Corporation