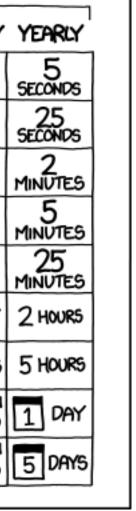
CS-5630 / CS-6630 Uisualization for Data Science Uiews



Alexander Lex <u>alex@sci.utah.edu</u>

HOW LONG CAN YOU WORK ON MAKING A ROUTINE TASK MORE EFFICIENT BEFORE YOU'RE SPENDING MORE TIME THAN YOU SAVE? (ACROSS FIVE YEARS)

			H0W	OFTEN YO	UDOTHE	TA5K
		50/DAY	5/DAY	DAILY		MONTHLY
	1 SECOND	1 DAY	2 HOURS	30 MINUTES	4 MINUTES	1 MINUTE
	5 SECONDS	5 DAYS	12 HOURS	2 HOURS	21 MINUTES	5 MINUTES
	30 SECONDS	4 WEEKS	3 DAYS	12 HOURS	2 HOURS	30 MINUTES
HOL MUCH		8 WEEKS	6 DAYS	1 DAY	4 HOURS	1 HOUR
TIME YOU		9 MONTHS	4 WEEKS	6 DAYS	21 HOURS	5 HOURS
Shave Off	ALL MINULES		6 MONTHS	5 WEEKS	5 DAYS	1 DAY
	1 HOUR		IO MONTHS	2 MONTHS	10 DAYS	2 DAYS
	6 HOURS				2 MONTHS	2 WEEKS
	1 DAY					8 WEEKS





Multiple Views

Eyes over Memory: Trade-off of display space and working memory

Juxtapose and Coordinate Multiple Side-by-Side Views €

- → Share Encoding: Same/Different
 - → Linked Highlighting



→ Share Data: All/Subset/None

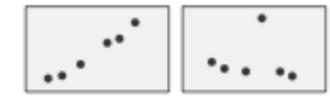


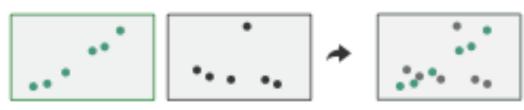
➔ Share Navigation

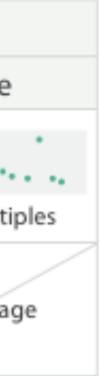


			Data		
		All	Subset	None	
Encoding	Same	Redundant	Overview/ Detail	Small Multi	
	Different	Multiform	Multiform, Overview/ Detail	No Linka	

→ Partition into Side-by-Side Views







Linked Views

Multiple Views that are simultaneously visible and linked together such that actions in one view affect the others.

Linked Views Options

highlighting: to link, or not navigation: to share, or not

encoding: same or multiform dataset: share all, subset, or none

- → Share Encoding: Same/Different
 - → Linked Highlighting

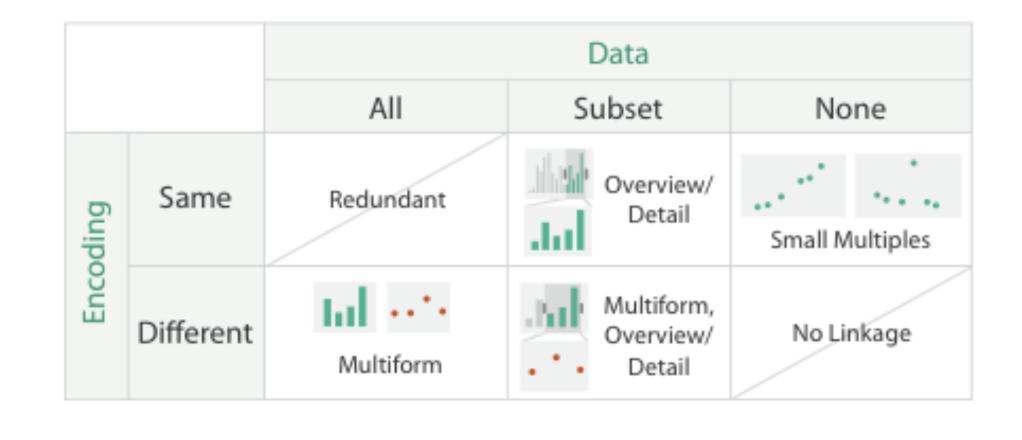


→ Share Data: All/Subset/None

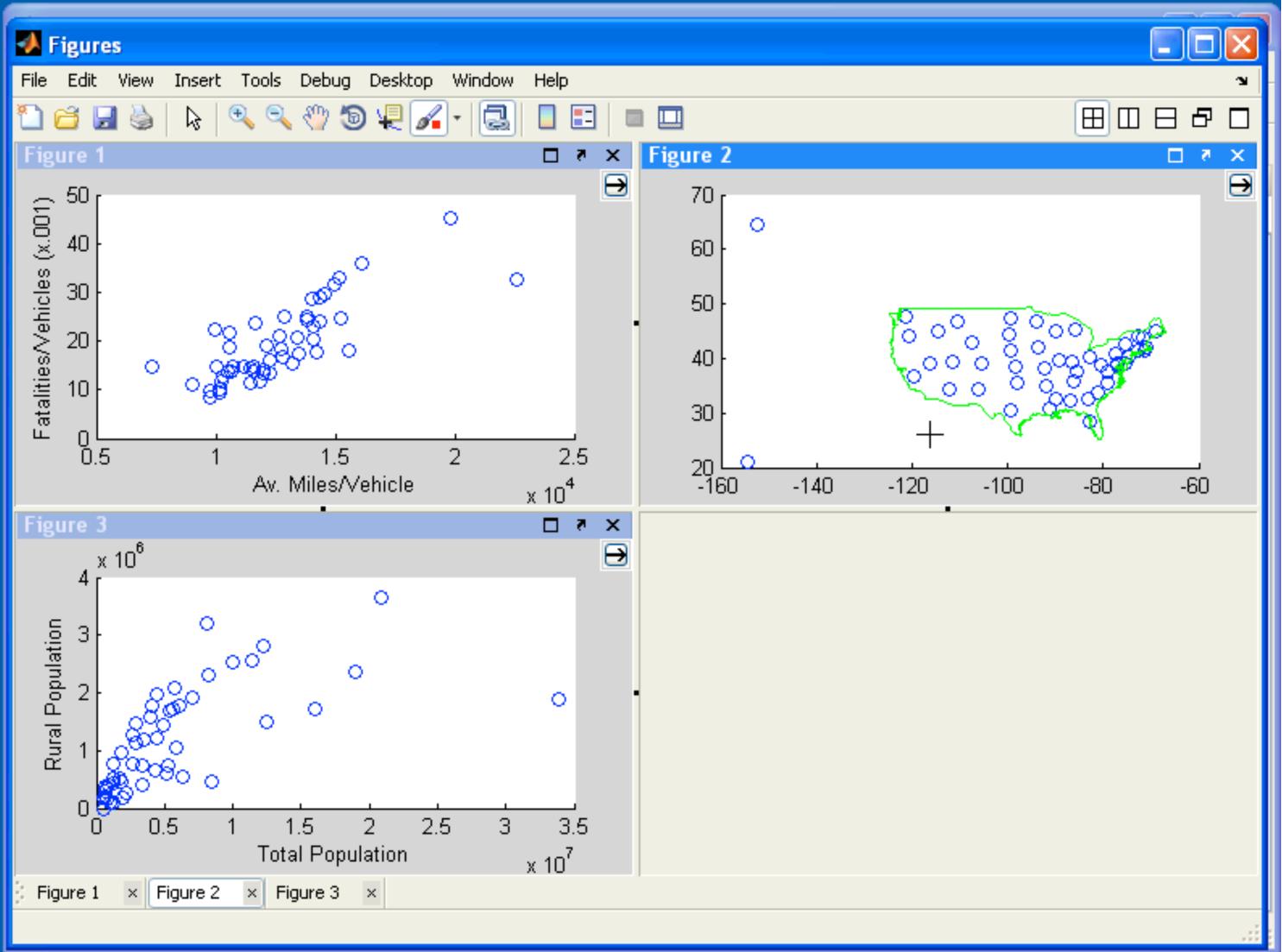


→ Share Navigation





Linked Highlighting



Linked Highlighting

i ne tastest-growing private companies in America.

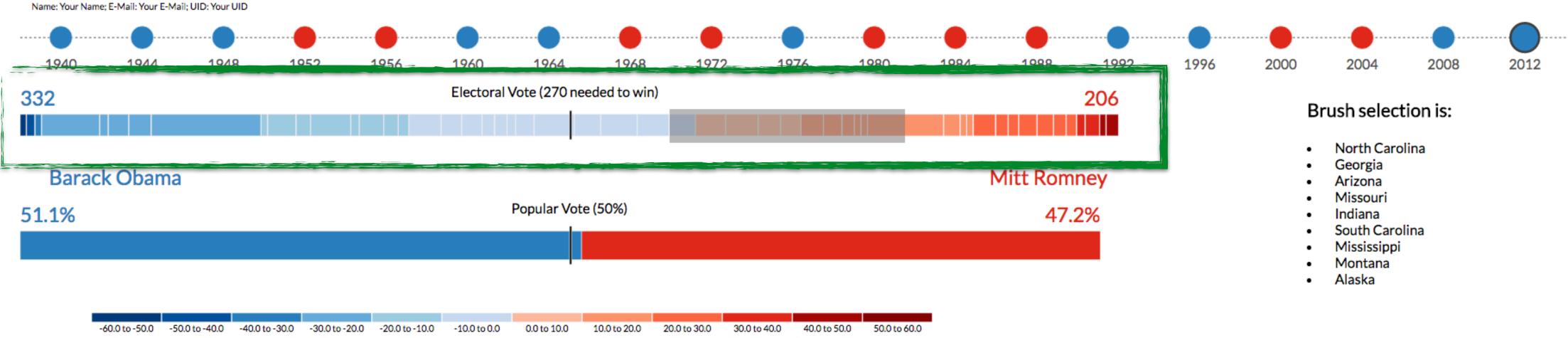
🗐 Industry	y	Growth	ᅌ 🔍 Q Search: c
Q Search	200 400 600 800	7 160k%	🖲 Fuhu 🗹
IT Services	301		🗐 Consumer Prod
Advertising & Marketing	181		El Segundo (Los
Business Products & Services	159		227 workers
	152	2 57k%	Quest Nutrition
	150		Food & Beverag
			-
Consumer Products & Services	84		El Segundo (Los 101 workers)
Government Services	81		191 workers
Manufacturing	64	3 55k%	Reliant Asset M
Retail	91		🗐 Business Produ
Human Resources	77		Arlington (Wash
Construction	87		📽 145 workers
Telecommunications	54	4 26k%	❶ Superfish [⊡]
Logistics & Transportation	46		())Software
Food & Beverage	64		Palo Alto (San J
Real Estate	58		In a lo Alto (Sali 5 In a lo Alto (Sali 5 In a lo Alto (Sali 5) In a lo Alto (Sali 5
Energy	46		-
Education	31	5 21k%	Acacia Commu
Security	29		💼 Telecommunica
Insurance Engineering			Maynard (Bosto)
25 Rows ~3 More	# 200 400 600 800	-	警 92 workers
State		6 20k%	Provider Power
		9	🗐 Energy
California	200 400 600 800	-	Q Auburn (Lewisto
	169		*50 workers
New York		7 18k%	Crescendo Bio
	116	/ TOK/0	
Virginia	113		(iii) Health
Illinois	98		South San Fran
Georgia	80		129 workers
Pennsylvania	73	<i>8</i> 16k%	Plexus Worldw
Massachusetts	75		🗐 Health
Ohio	59		Scottsdale (Pho
New Jersey			130 workers



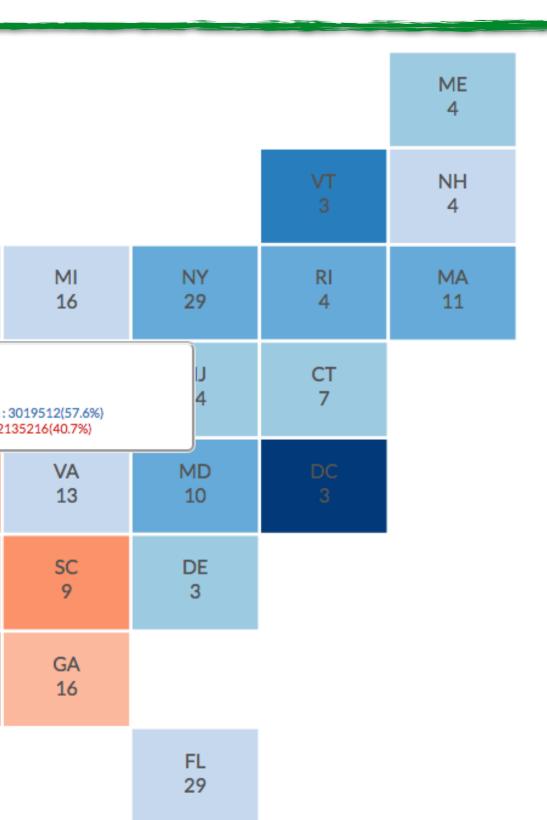
Multiform

difference visual encodings are used between the views implies shared data either all data or subset of data (overview + detail) rational: single, monolithic view has strong limits on the number of attributes that can be shown simultaneously different views support different tasks

US Presidential Elections from 1940 to 2012



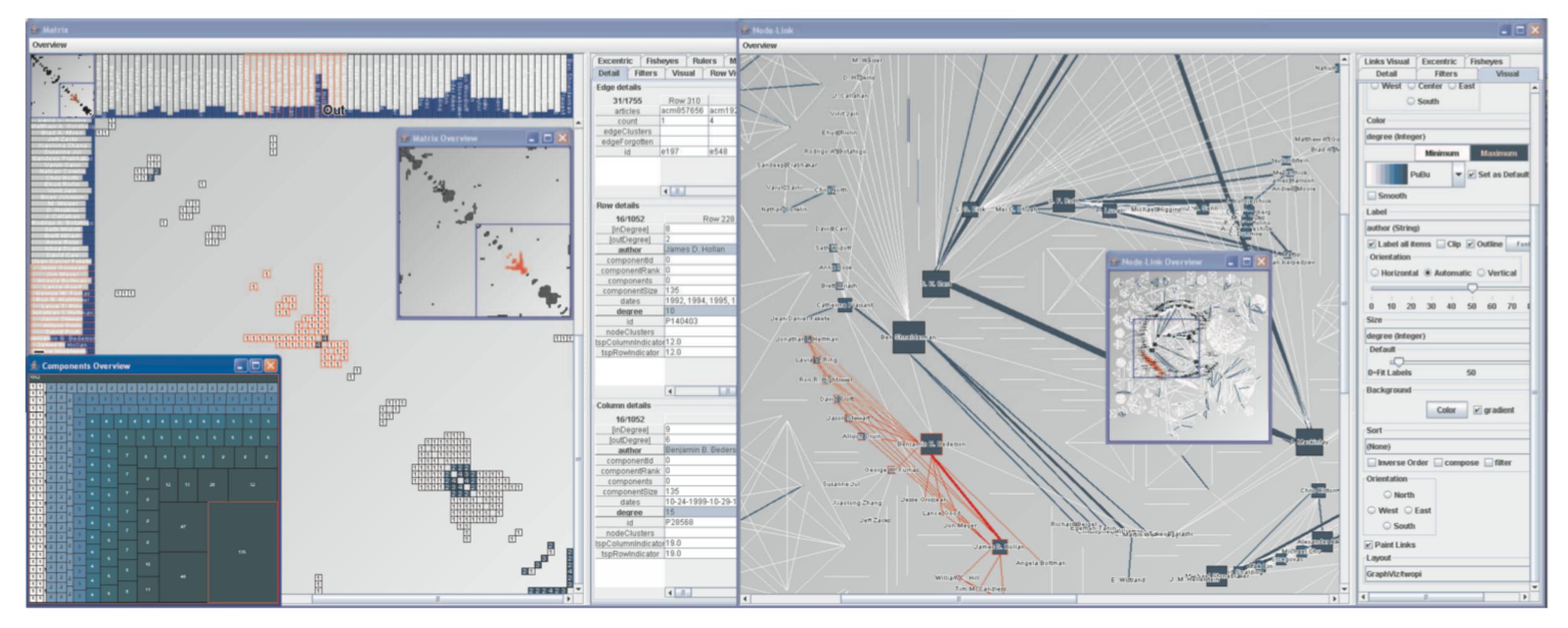
AK 3							
	WA 12	ID 4	MT 3	ND 3	MN 10	IL 20	WI 10
	OR 7	NV 6	WY 3	SD 3	IA 6	IN 11	Illinois Electoral Votes: 20 • Barack Obama : 3 • Mitt Romney: 21
	CA 55	UT 6	CO 9	NE 5	MO 10	KY 8	WV 5
		AZ 11	NM 5	KS 6	AR 6	TN 11	NC 15
				OK 7	LA 8	MS 6	AL 9
	HI 4			TX 38			



Multiform **Different Views** here also same data

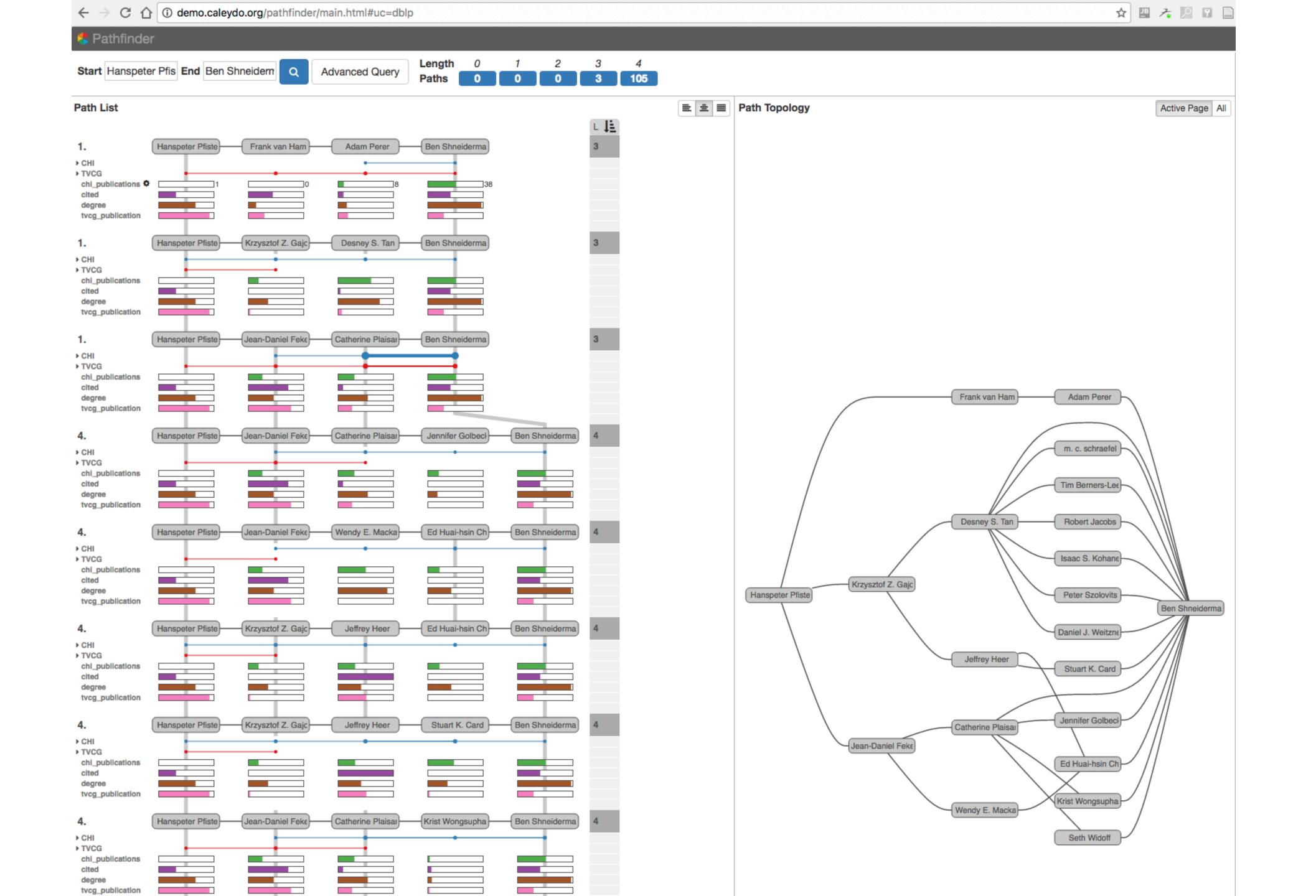


MatrixExplorer



Same Data - Different Idioms (Multiform)

Henry 2006



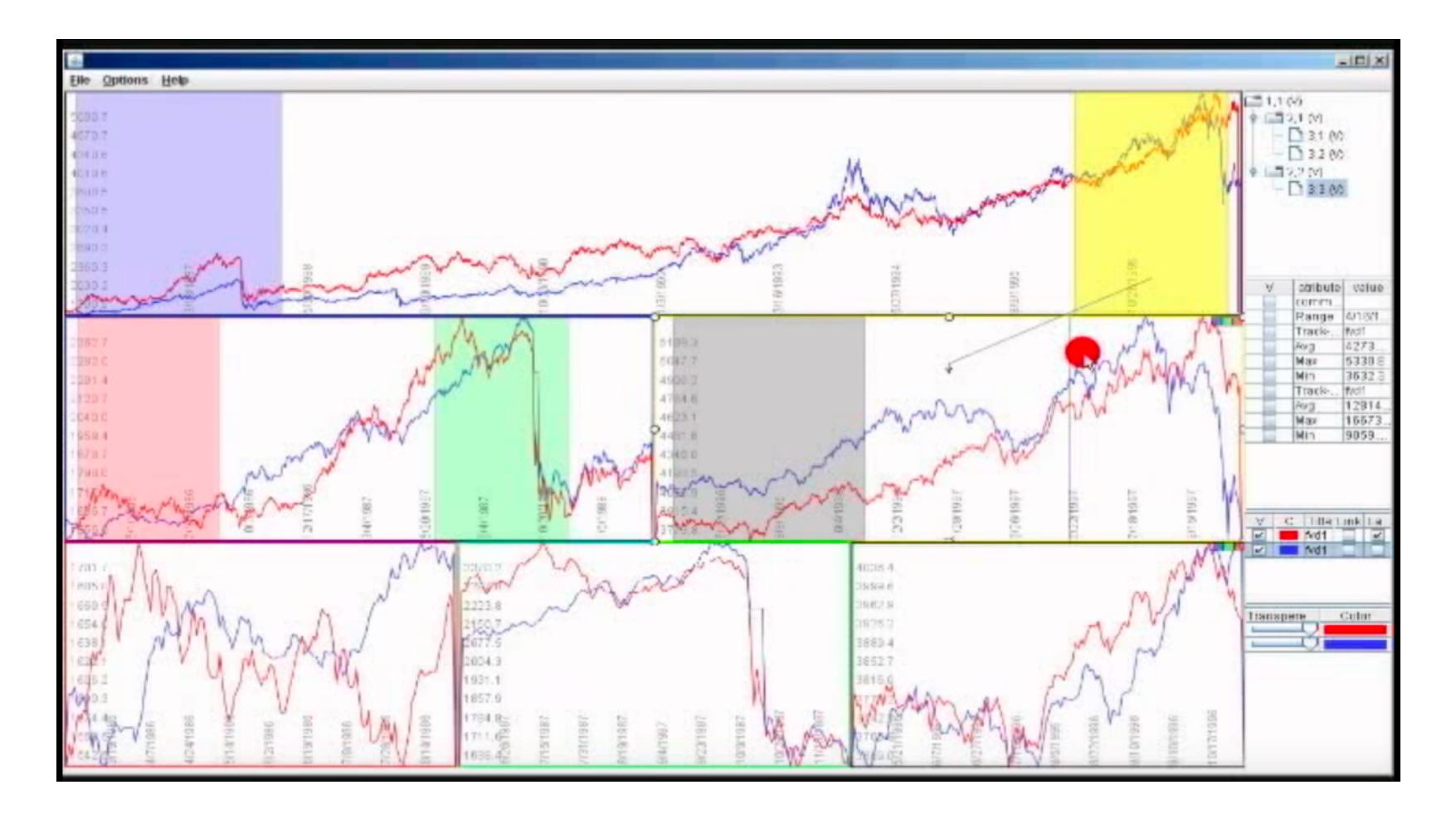
OVERVIEW + DETAIL

one view shows (often summarized) information about entire dataset, while additional view(s) shows more detailed information about a subset of the data

rational

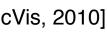
for large or complex data, a single view of the entire dataset cannot capture fine details

Stack Zooming

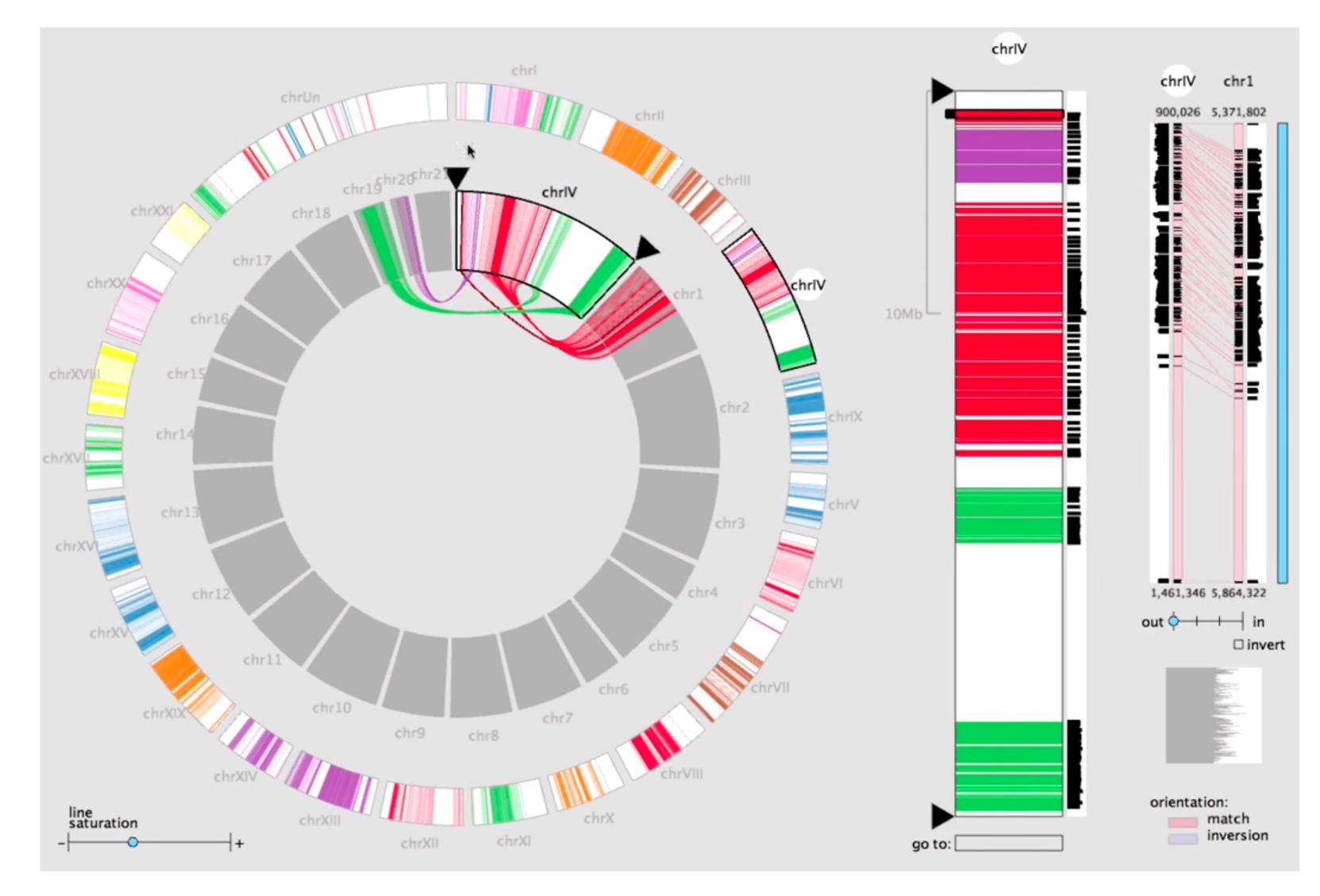


Same Data - Same Encoding, Different Resolution

[Javed & Emlqvist, PacificVis, 2010]



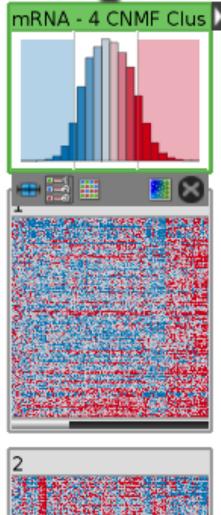
MizBee

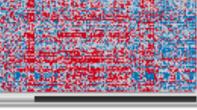


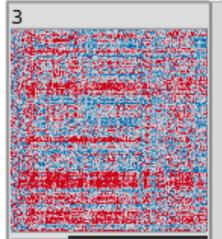
Multiform Overview & Detail

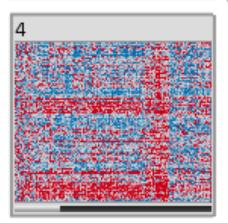


StratomeX

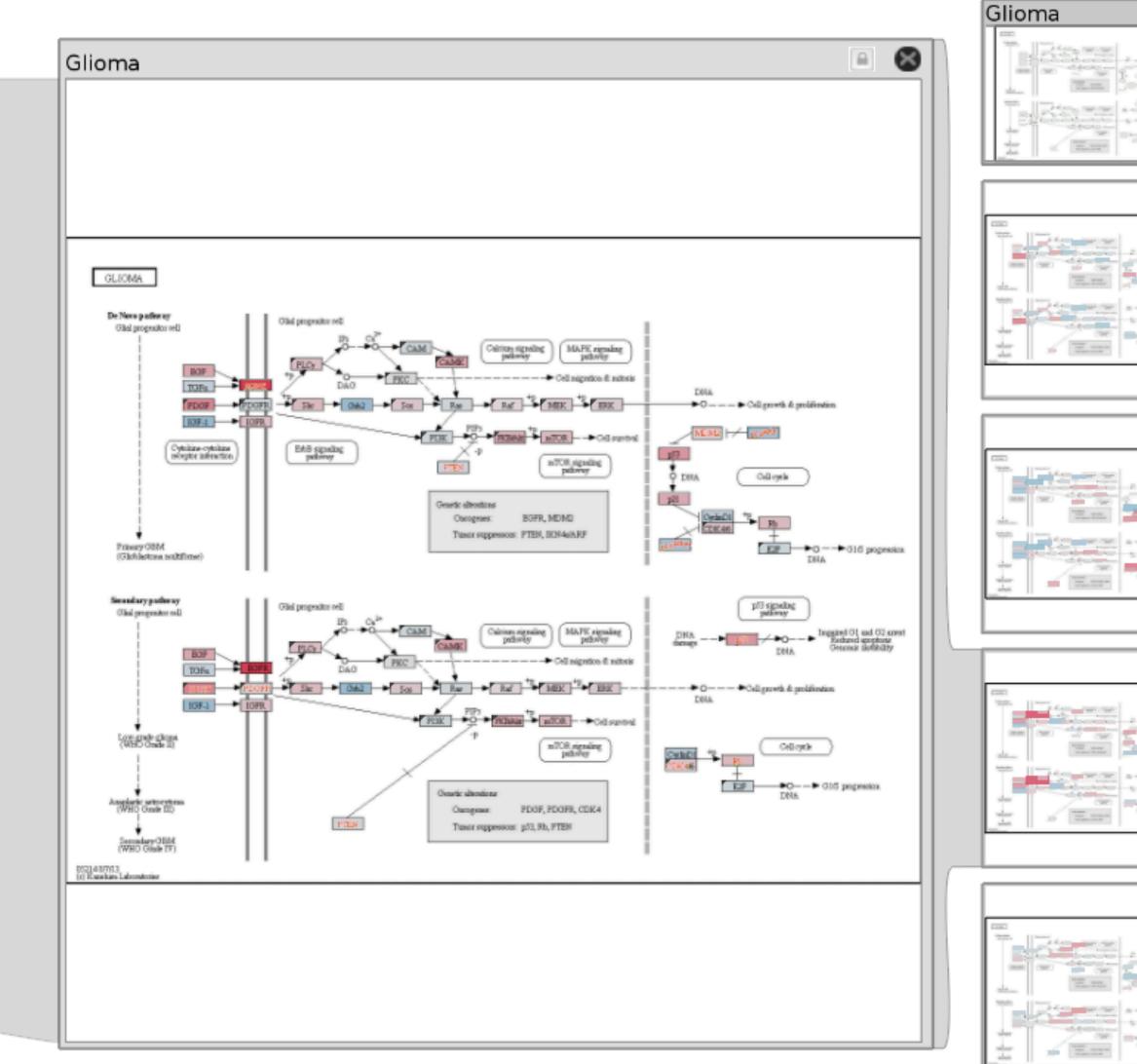








•



- 11	
-	
080	
S	
(M) 1.5.*500	
Carlo Carlos	
	L
in the second	
080	
- 10 m - 10 m	
1980	
(10)	
(S#)	

SMALL MULTIPLES

each view uses the same visual encoding, but shows a different subset of the data

rational

quickly compare different parts of a data set, relying on eyes instead of memory

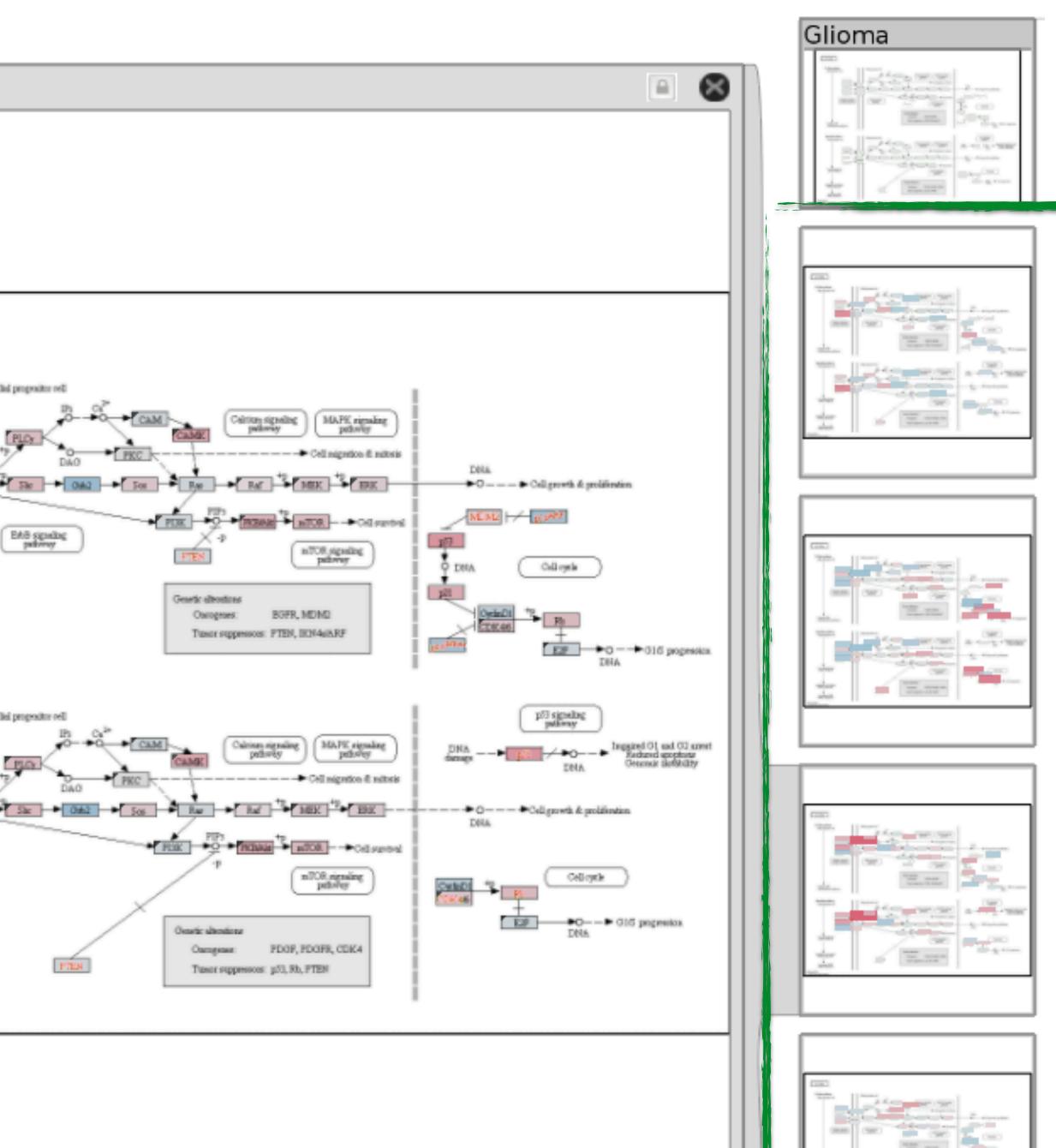


Small Multiples for Graph Attributes



StratomeX

\otimes Glioma GLIOMA. De Nova patieray Olid propositor reli-Gilal progration reli DOF:1 Cytokize-cytokize progitize inferention Princy OBM (Globlestone multifume) Secondary pathway Gilei propositor rell Obd progenitor cell Low-grade glicona (WHO Oracle II) Anglatic attorner (WHO Goale III) ٠ Secondary OBM (WHO Gitale IV) 0521410713 (c) Katakara Laboratorian



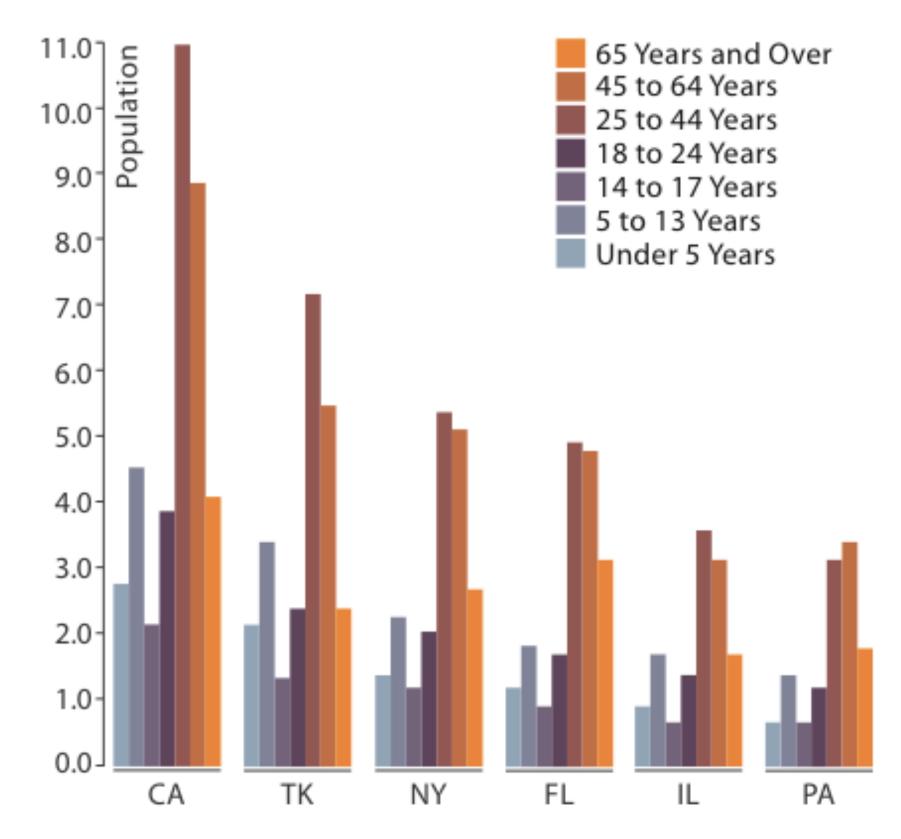
Partitioning

PARTITIONING

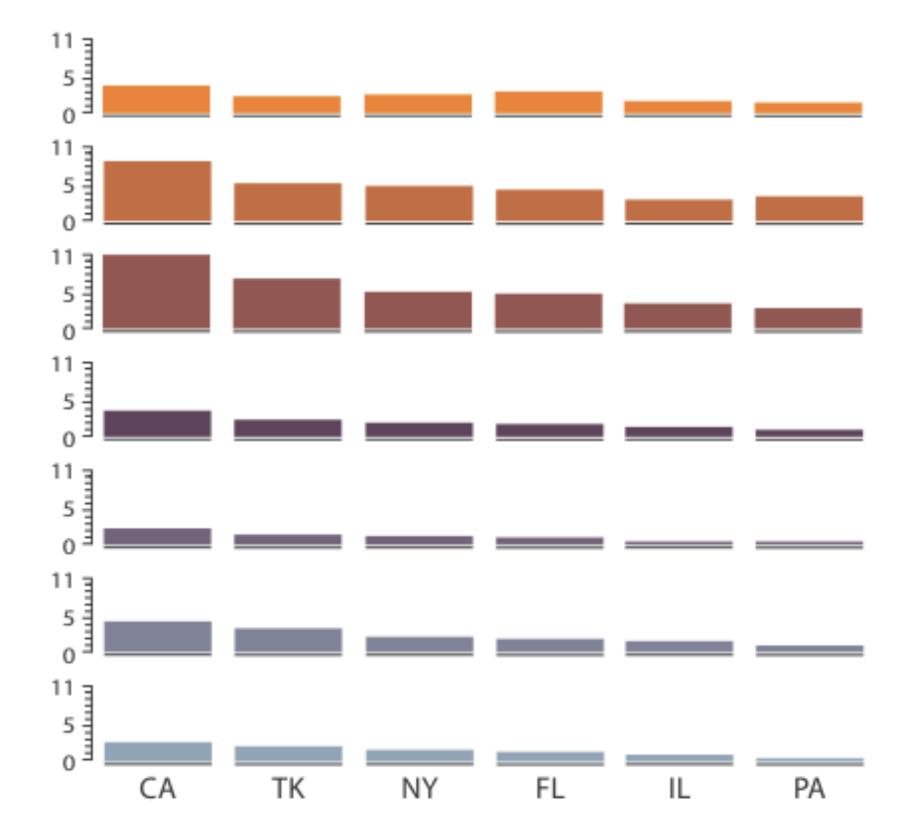
action on the dataset that separates the data into groups design choices

how to divide data up between views, given a hierarchy of attributes how many splits, and order of splits how many views (usually data driven) partition attribute(s) typically categorical

Partitioning - Age Distribution by State



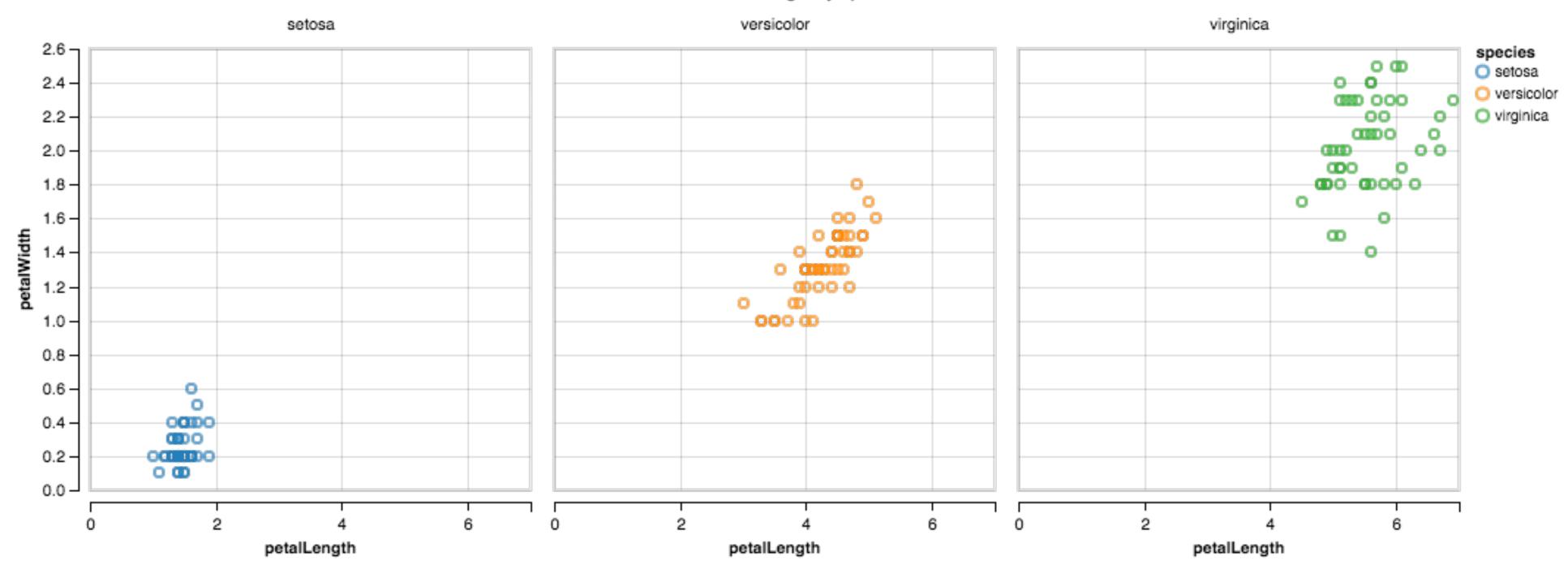
Partitioned by State



Partitioned by Age Group and State

Partition by Category

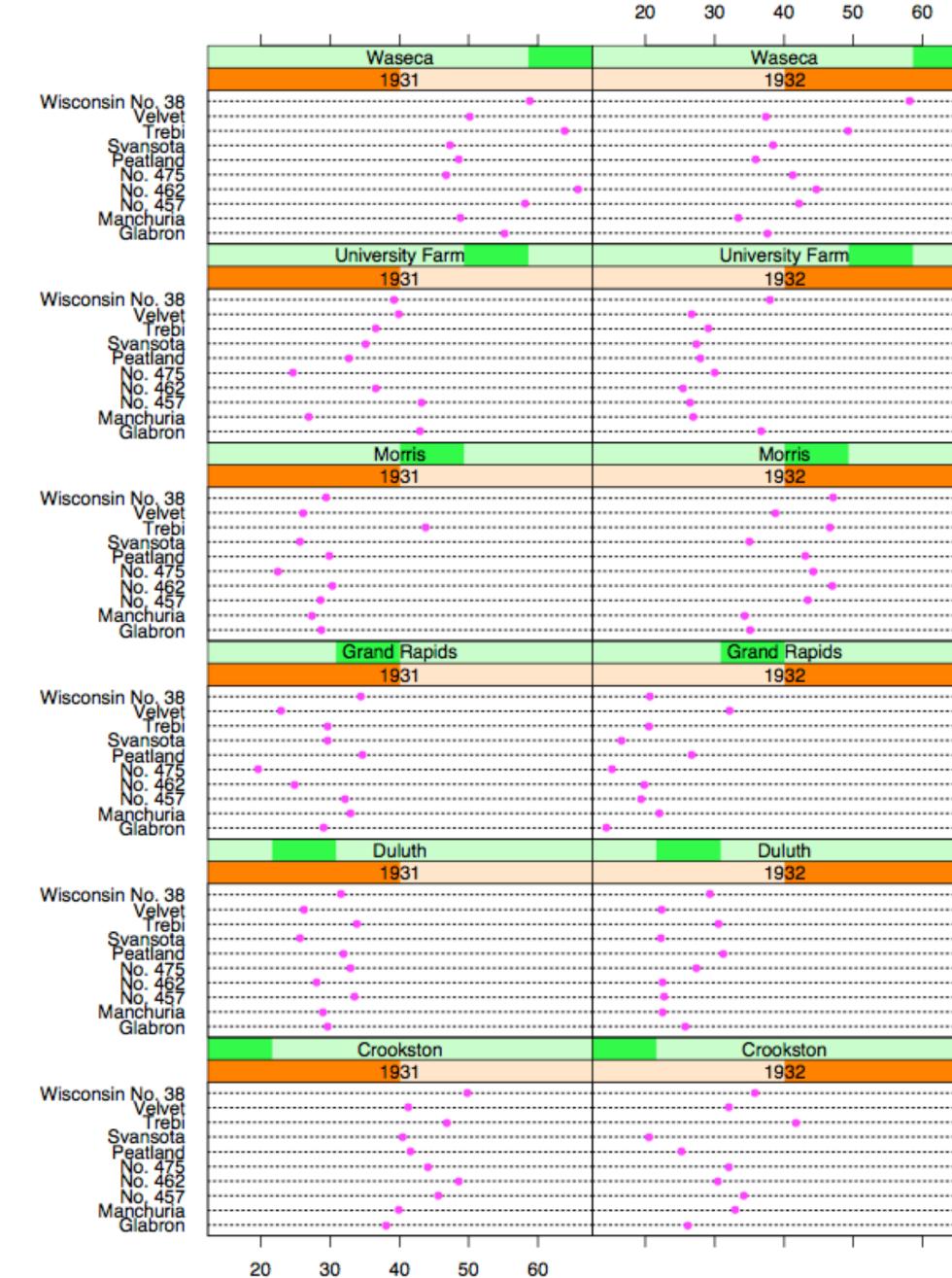
Petal Width v. Length by Species



Trellis Plots

panel variables

- attributes encoded in individual views
- partitioning variables
 - partitioning attributes assigned to columns and rows
- main-effects ordering
 - order partitioning variable based on derived data
 - support perception of trends and structure in data



Barley Yield (bushels/acre)

	_	_	_	
				ļ
		-		ł
•	•	-	•	1
		=		ł
-		_		ļ
				1
-	•	-	•	1
•		-		ł
				J
				1
-	•	-	•	1
-	•	-	•	ł
_	_	_		J
"	-	-		1
				1
				1
				1
=	•	-	•	1
		-		ł
_	_	_	_	l
				1
-	•	-	•	1
		=		ł
-	•	-		1
•	•	-	•	ł
	e.		e.	ļ
	Ĩ	Ĩ	ĵ	J
-	•	-	•	ţ
-	-	-	•	ļ
				ļ
				J
				1
				1
-	-	-	1	1
1	1	1	1	1
•	-	-	•	ţ
				ļ
-	•	-	•	1
-	•	-	•	ł
_		_		ļ
				1
=	*	=	-	1
-		-		ł
_	_	_	_	J
			1	1
-	•	=		1
	-		-	ł
				1
			_	
-		-		
		-		
	-			
	-			
	-			
	-			
	-			
-	-	-		
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-	-	-	
-	-			
-	-		-	
	-			
	-			
	-			
	-			
	-			
	-			
	-			

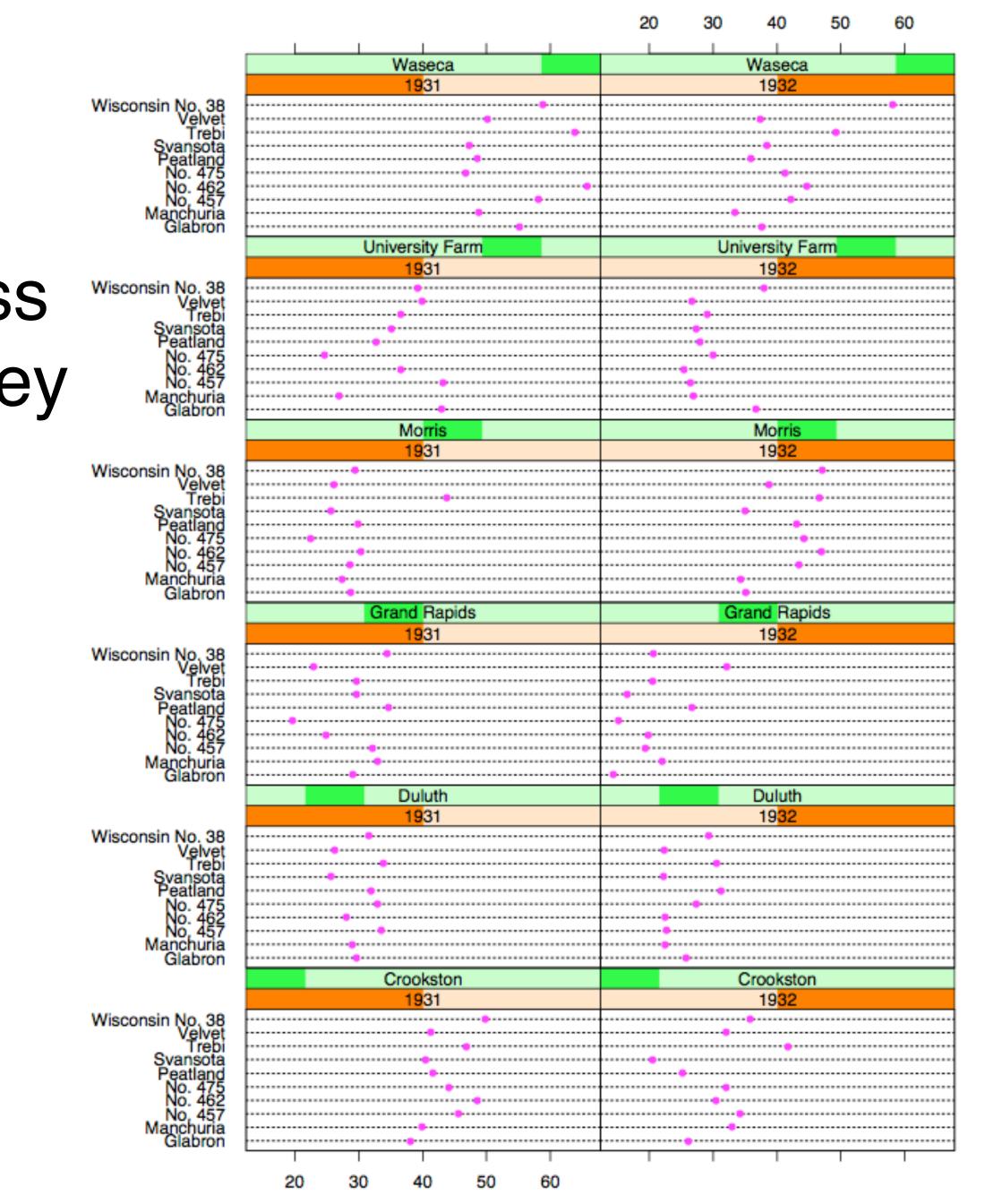
Data

Barley Yields in two years across multiple farms for multiples barley strains

partitioning variables

Columns partitioned by year

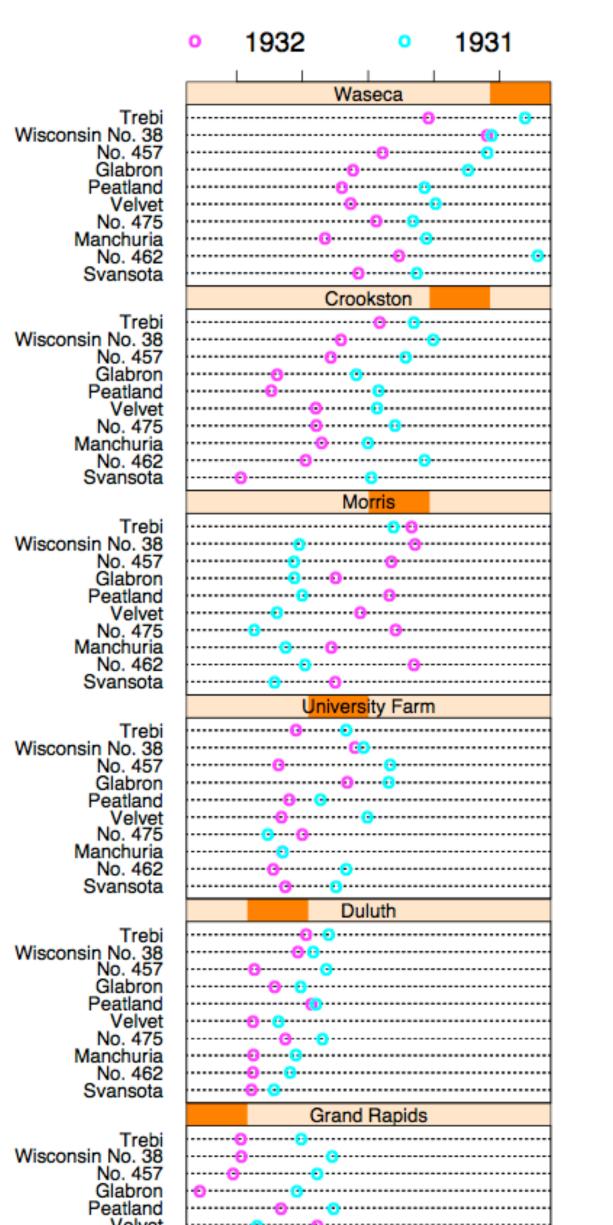
Rows partitioned by farm



Barley Yield (bushels/acre)

Becker 1996



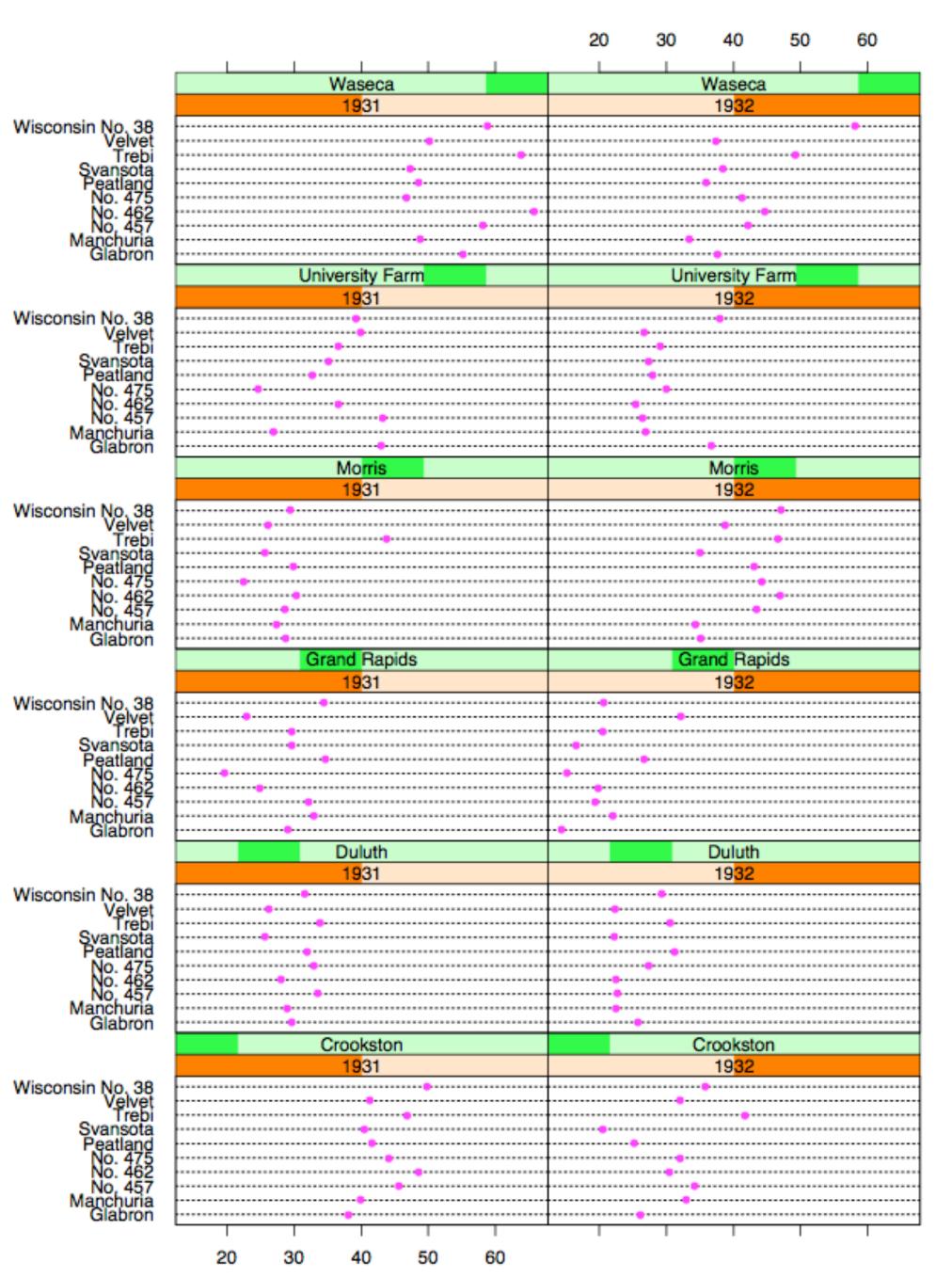


20 30 40 50

Peatland Velvet No. 475 Manchuria No. 462 Svansota

Barley Yield (bushels/acre)

60



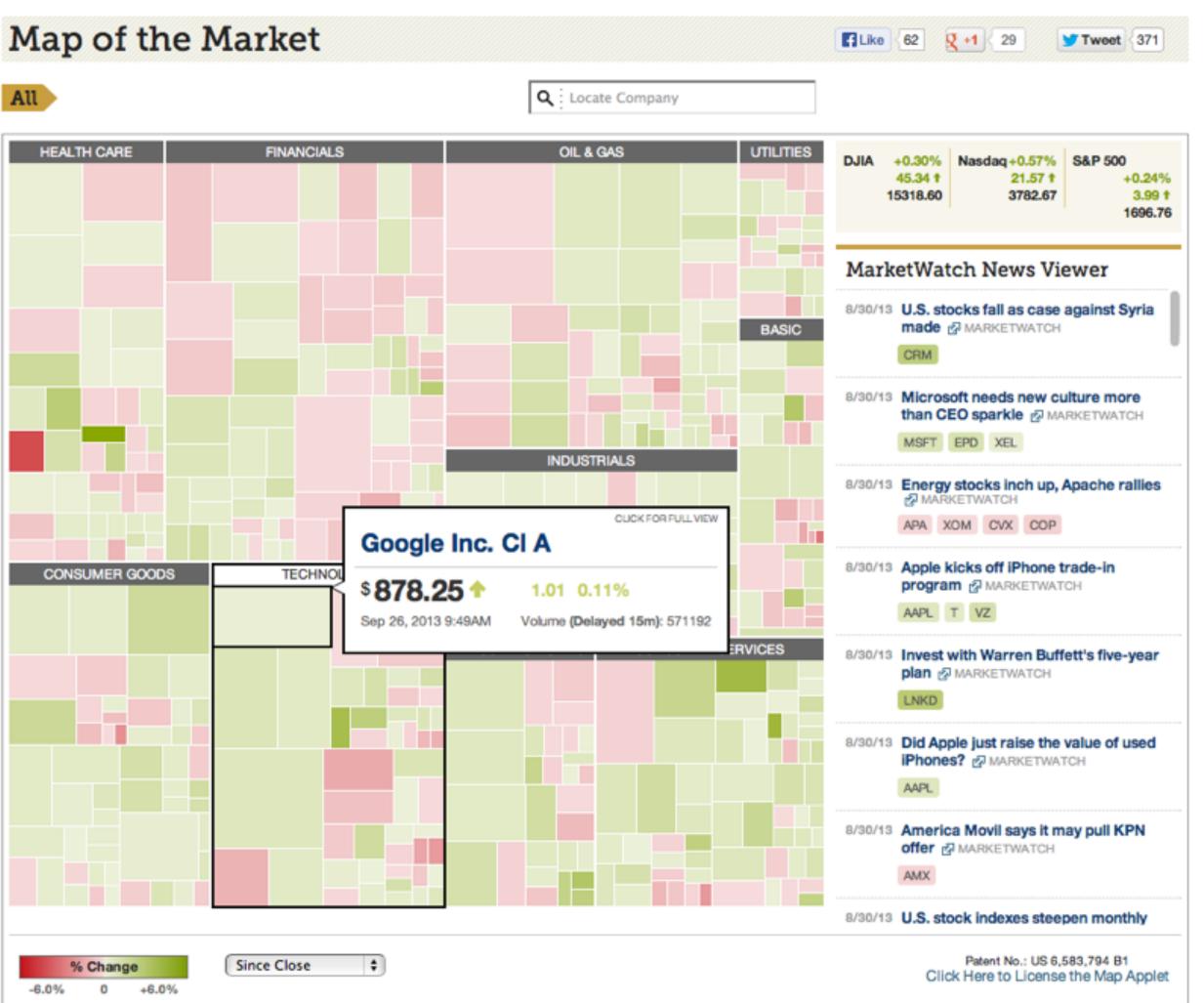
Barley Yield (bushels/acre)

Becker 1996

Recursive Subdivision

partitioning: flexibly transform data attributes into a hierarchy

use treemaps as spacefilling rectangular layouts



Treemap

HiVE example: London property

partitioning attributes

house type neighborhood sale time

encoding attributes

average price (color) number of sales (size)

results

between neighborhoods, different housing distributions within neighborhoods, similar prices



Slingsby 2009

partitioning attributes

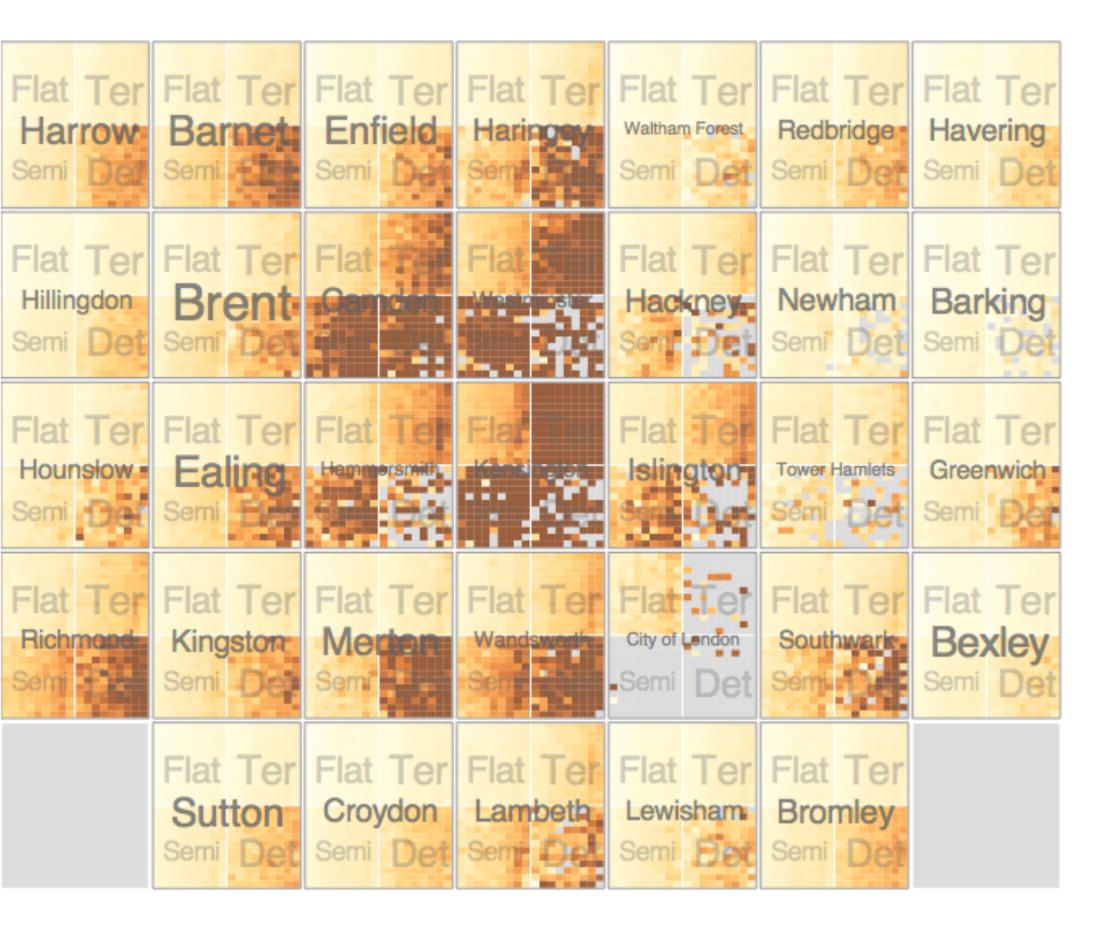
neighborhood house type sale time (year) sale time (month)

encoding attributes

neighborhood location (approximate) average price (color) *n/a* (size)

results expensive neighborhoods near center of city

HiVE example: London property



Slingsby 2009

Configuring Hierarchical Layouts to Address **Research Questions**



Aidan Slingsby, Jason Dykes and Jo Wood giCentre, Department of Information Science, City University London http://www.gicentre.org/hierarchical_layouts/

ONDON



https://vimeo.com/9870257



LAYERING

combining multiple views on top of one anoth to form a composite view rational

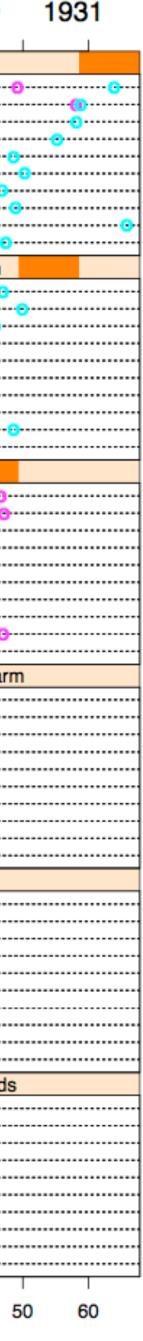
supports a larger, more detailed view than using multiple views

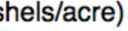
trade-off

layering imposes constraints on visual encoding choice as well as number of layers that can be show

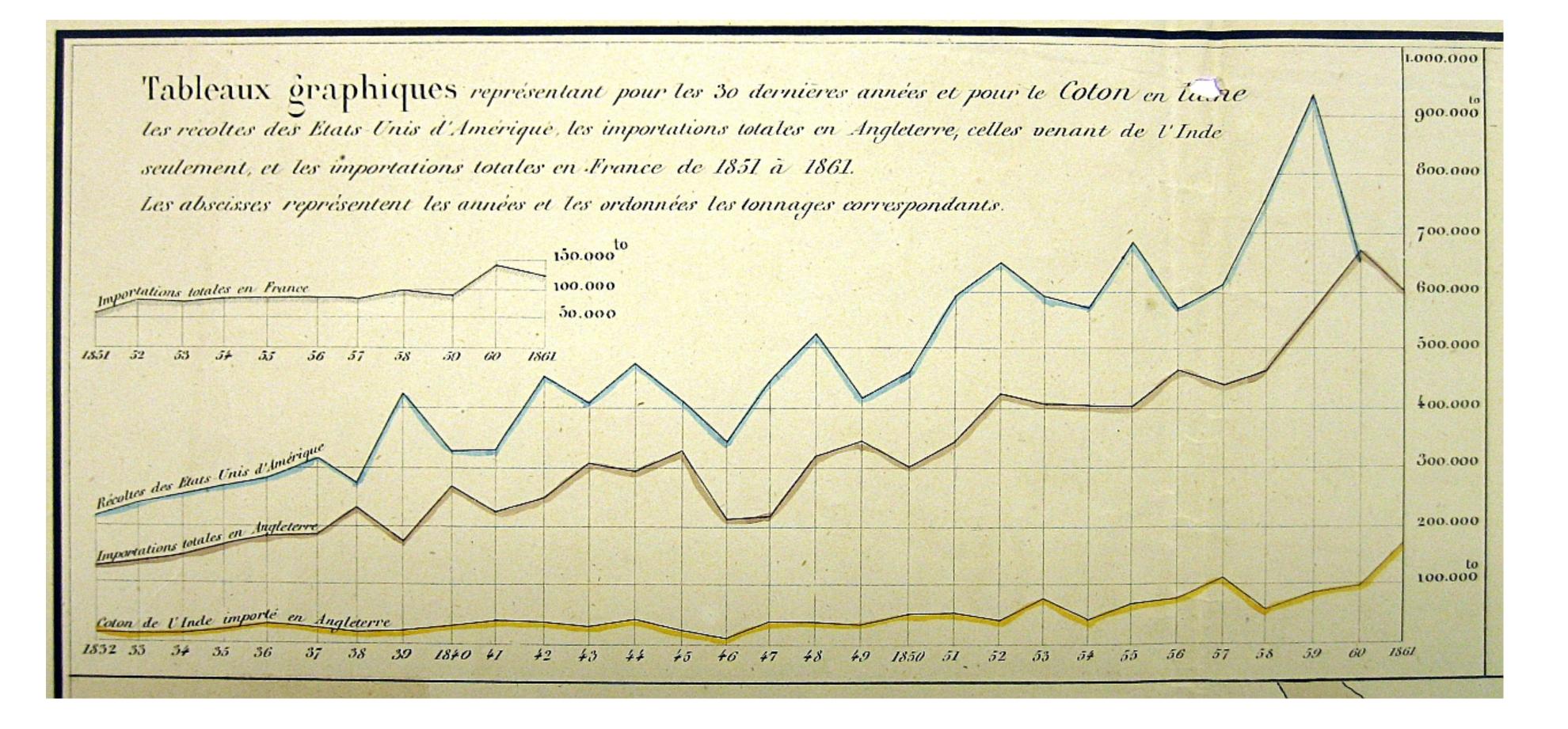
		0	193	2	•	2
			1	- N	/aseca	
	Trebi					···· 0 ···
	Wisconsin No. 38 No. 457					
	Glabron				0	
	Peatland Velvet					<mark>0</mark>
	No. 475					-
	Manchuria No. 462					•
	Svansota				•••••••	••••
	The lat			Cr	ooksto	n
	Trebi Wisconsin No. 38					•
	No. 457					9
	Glabron Peatland				•••••	
	Velvet					
	No. 475 Manchuria					
	No. 462 Svansota			•••		••••
her	Svansola		<u> </u>		Morris	
	Trebi					
	Wisconsin No. 38			0		•••••
	No. 457 Glabron			. 0		
	Peatland Velvet	l	·····	<u>G</u>		
	No. 475		-			
	Manchuria No. 462		••••••	· ·		
	Svansota			0		
				Unive	e <mark>rsi</mark> ty F	
	Trebi Wisconsin No. 38			0		
	No. 457		····· <mark>·</mark> ··		<mark>0</mark> -	
	Glabron Peatland		8		-	
	Velvet No. 475		•••••		•••	
	Manchuria		0			
	No. 462 Svansota		·····		••••••	
				[Duluth	
	Trebi			· ·		
	Wisconsin No. 38 No. 457		0			
	Glabron Peatland		·····	-		
	Velvet		0 0			
	No. 475 Manchuria		O	_		
	No. 462			-		
	Svansota		···· ·· ···	0		de.
	Trebi			Gran	nd Rap	
	Wisconsin No. 38		-			
	No. 457 Glabron		-			
	Peatland			·····•		
	Velvet No. 475	0	<mark>0</mark>		•••••	
	Manchuria		0	_		
NN	No. 462 Svansota					
					1	I
		2	20	30	40	50

Barley Yield (bushels/acre)



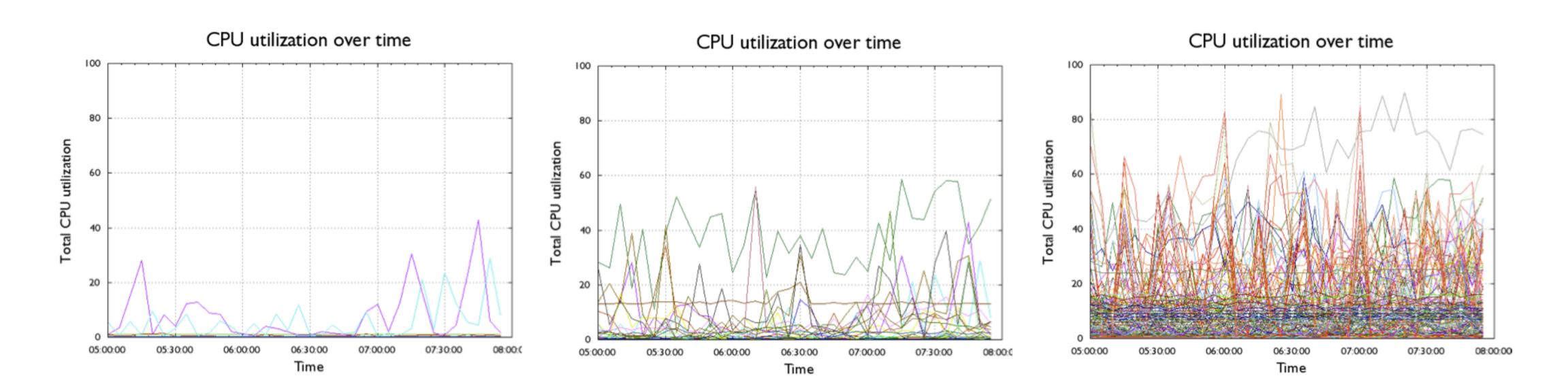


JOSEPH MINARD

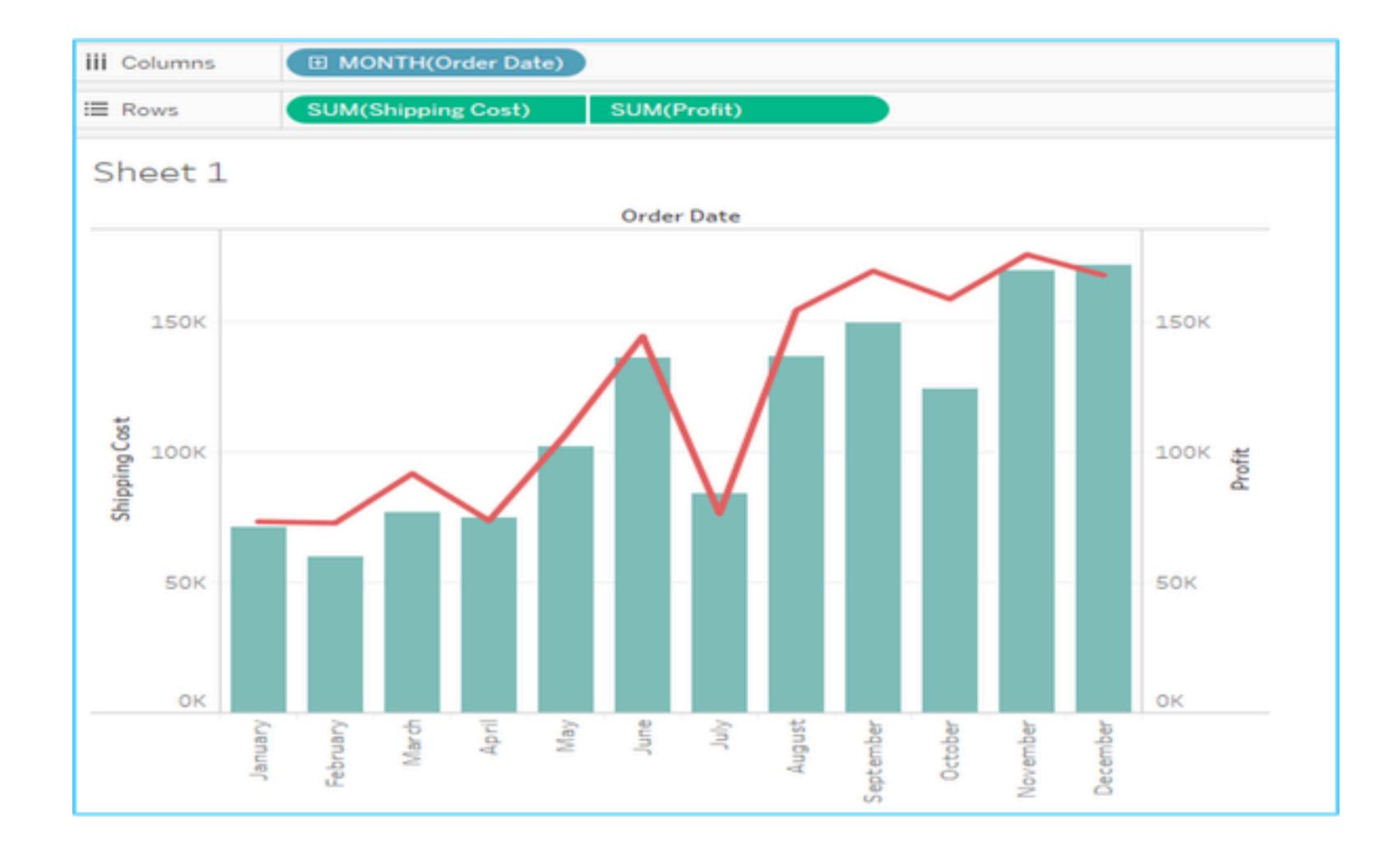




overlays

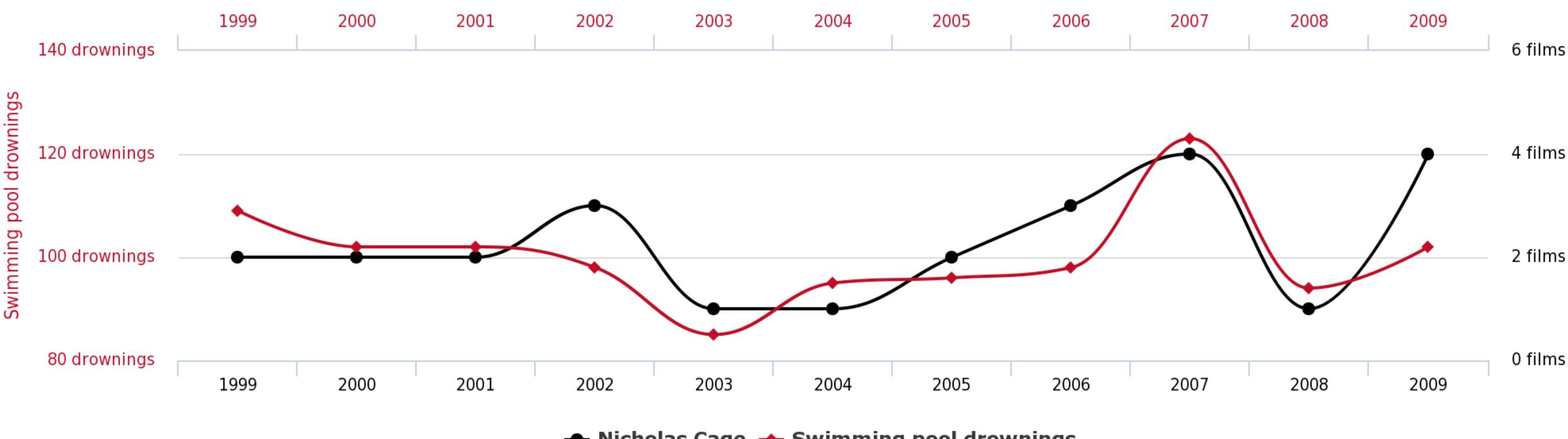


Dual Axis



Dual Axis (don't)

Number of people who drowned by falling into a pool correlates with **Films Nicolas Cage appeared in**





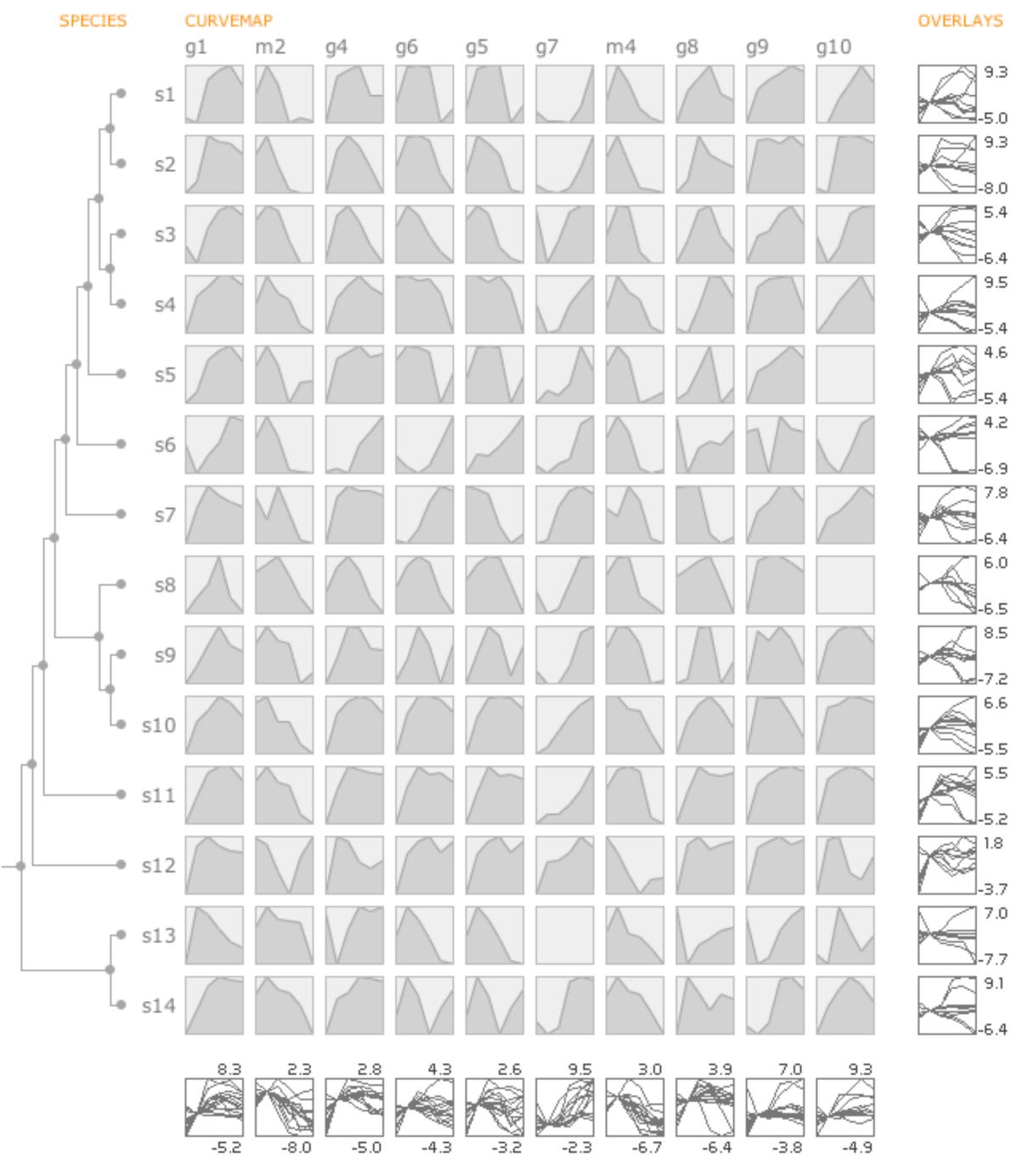
- Nicholas Cage - Swimming pool drownings





Combined

Partitioned + layered graph Synchronized through highlighting



)

=

MCV to the Max

