

# CS-5630 / CS-6630 Visualization for Data Science Views

Alexander Lex  
[alex@sci.utah.edu](mailto:alex@sci.utah.edu)



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

		HOW OFTEN YOU DO THE TASK					
		50/DAY	5/DAY	DAILY	WEEKLY	MONTHLY	YEARLY
HOW MUCH TIME YOU SHAVE OFF	1 SECOND	1 DAY	2 HOURS	30 MINUTES	4 MINUTES	1 MINUTE	5 SECONDS
	5 SECONDS	5 DAYS	12 HOURS	2 HOURS	21 MINUTES	5 MINUTES	25 SECONDS
	30 SECONDS	4 WEEKS	3 DAYS	12 HOURS	2 HOURS	30 MINUTES	2 MINUTES
	1 MINUTE	8 WEEKS	6 DAYS	1 DAY	4 HOURS	1 HOUR	5 MINUTES
	5 MINUTES	9 MONTHS	4 WEEKS	6 DAYS	21 HOURS	5 HOURS	25 MINUTES
	30 MINUTES		6 MONTHS	5 WEEKS	5 DAYS	1 DAY	2 HOURS
	1 HOUR		10 MONTHS	2 MONTHS	10 DAYS	2 DAYS	5 HOURS
	6 HOURS				2 MONTHS	2 WEEKS	1 DAY
	1 DAY					8 WEEKS	5 DAYS

# 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

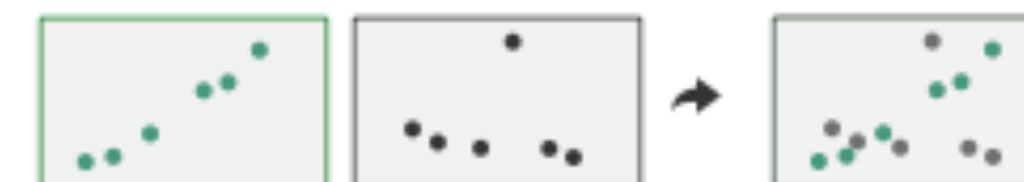


		Data		
		All	Subset	None
Encoding	Same	Redundant	 Overview/ Detail	 Small Multiples
	Different	 Multiform	 Multiform, Overview/ Detail	No Linkage

## ➔ Partition into Side-by-Side Views



## ➔ Superimpose Layers



# Linked Views

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

# Linked Views Options

encoding: same or multiform

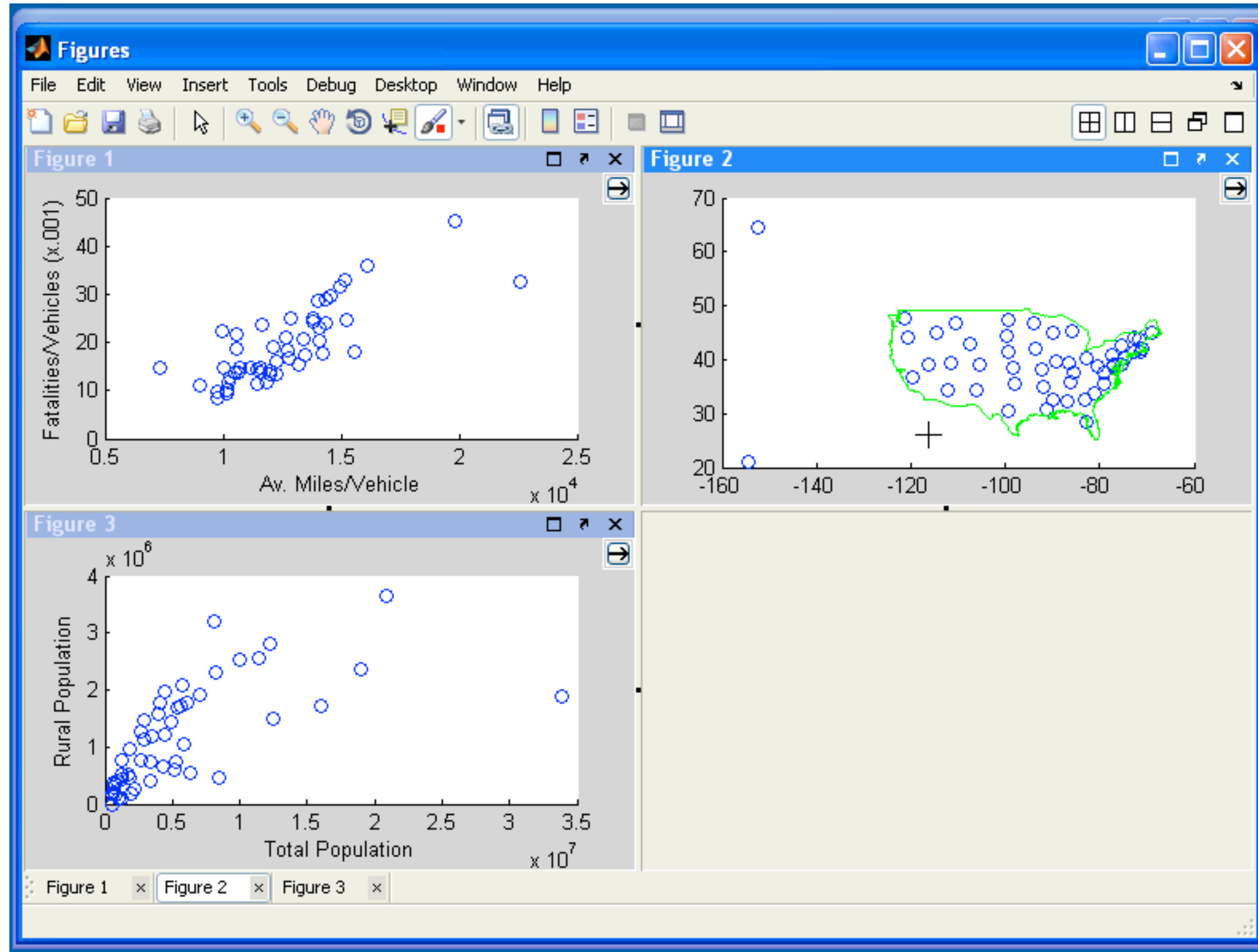
dataset: share all, subset, or none

highlighting: to link, or not

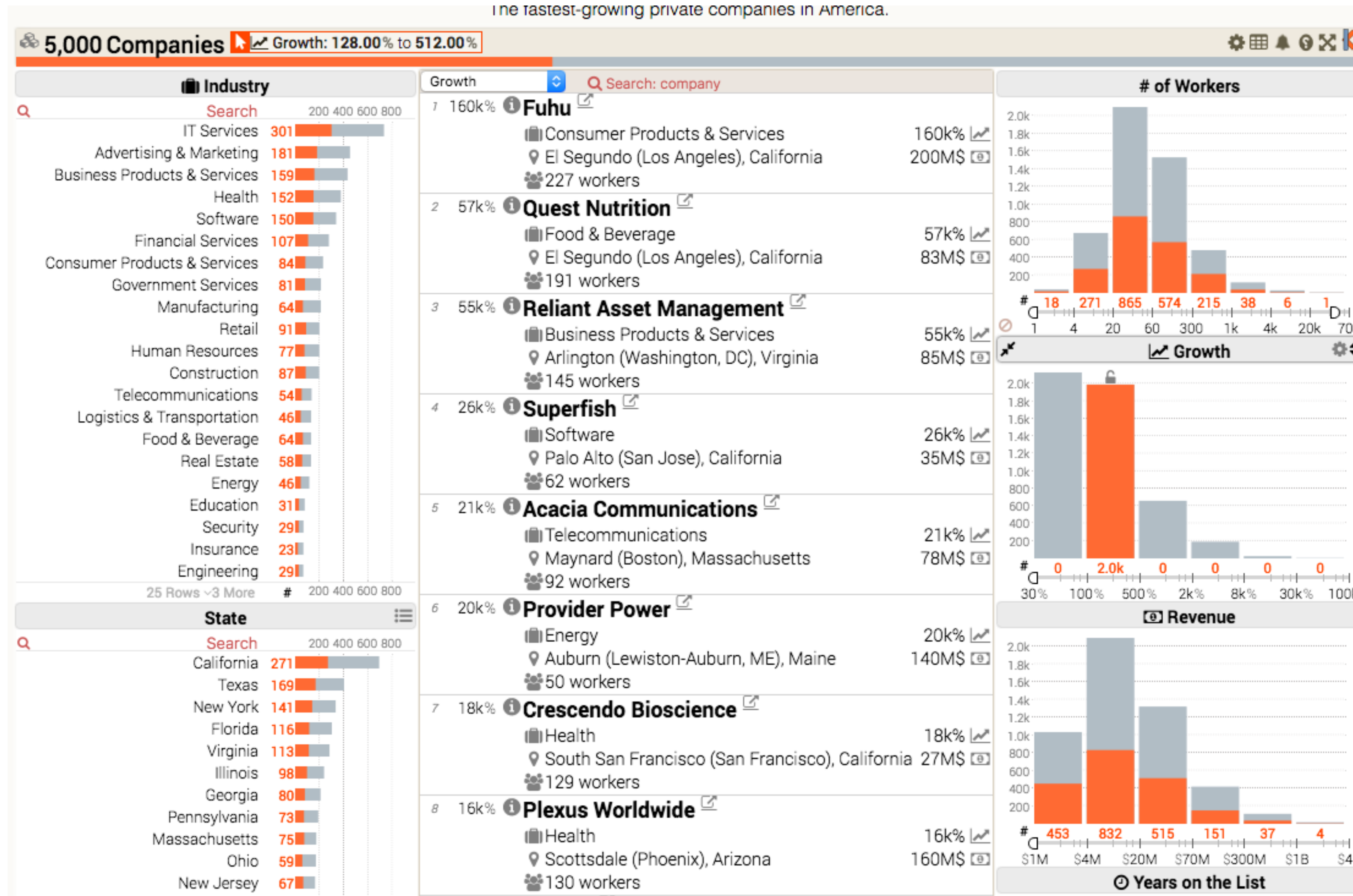
navigation: to share, or not



# Linked Highlighting



# Linked Highlighting



# Multiform

difference visual encodings are used between the views

- implies shared data

- either all data

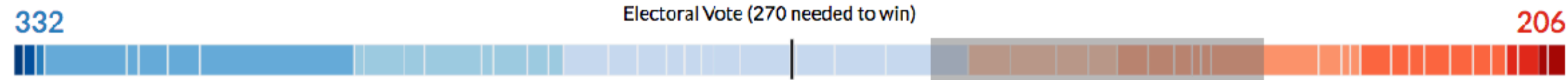
- or subset of data (overview + detail)

**rational:**

single, monolithic view has strong limits on the number of attributes that can be shown simultaneously

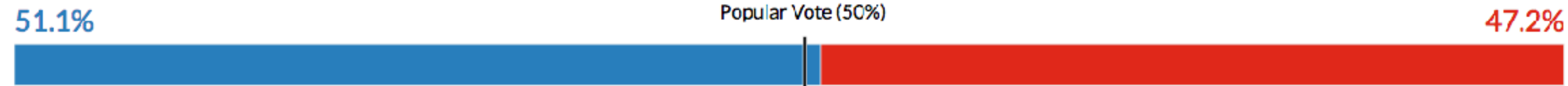
# US Presidential Elections from 1940 to 2012

Name: Your Name; E-Mail: Your E-Mail; UID: Your UID



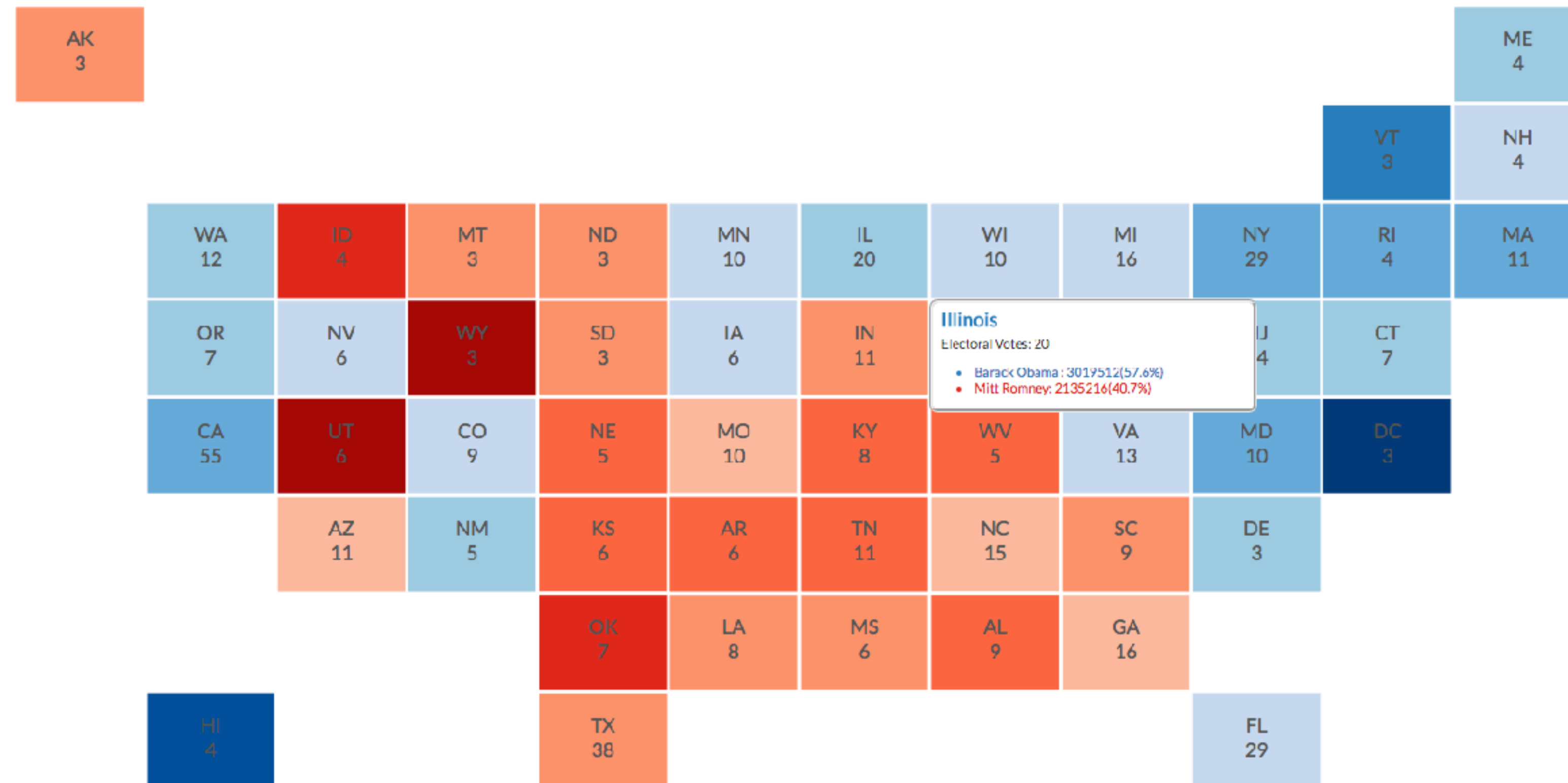
Barack Obama

Mitt Romney



Brush selection is:

- North Carolina
- Georgia
- Arizona
- Missouri
- Indiana
- South Carolina
- Mississippi
- Montana
- Alaska



**Multiform**  
Different Views  
here also same data

# SHARED-DATA

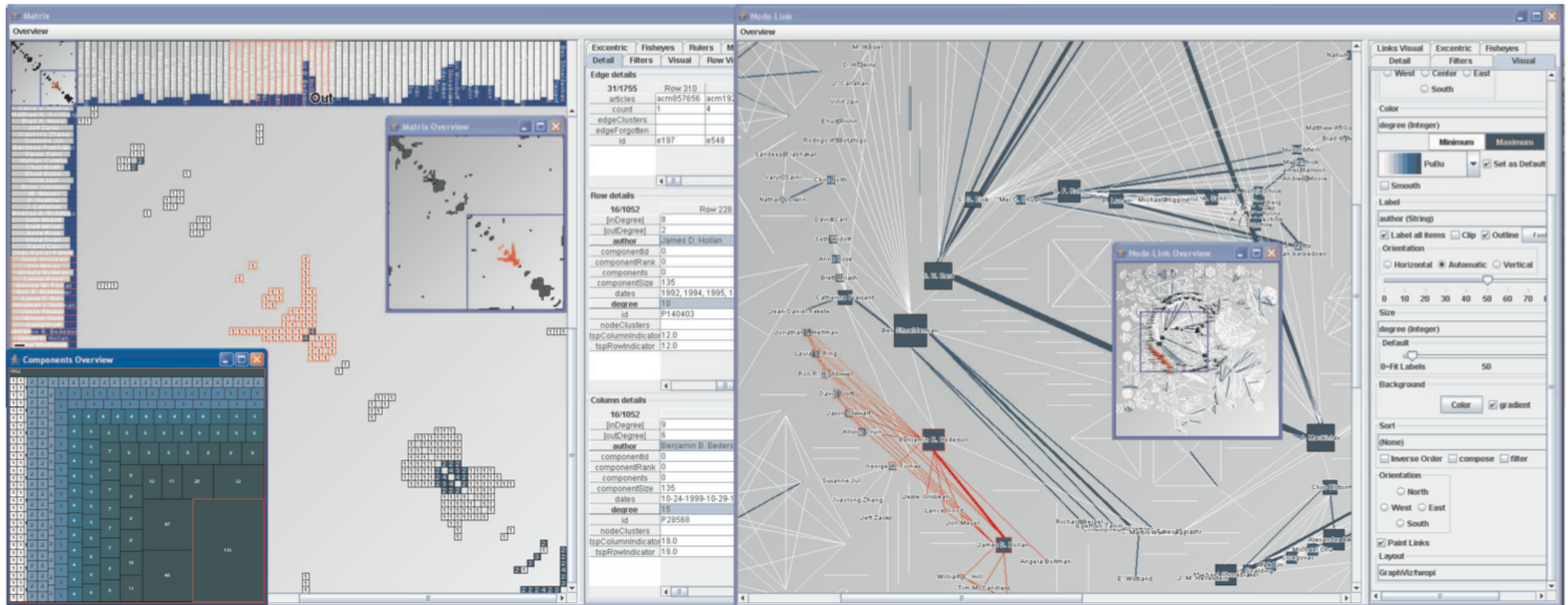
showing all data in each view, but with different encoding schemes

**rational**

different views support different tasks



# MatrixExplorer



Same Data - Different Idioms (Multiform)

Henry 2006



Start Hanspeter Pfister End Ben Shneiderman

Advanced Query

Length Paths 0 0 0 3 105

Path List

1.	Hanspeter Pfister	Frank van Ham	Adam Perer	Ben Shneiderman	3	
CHI						
TVCG						
chi_publications	1	0	8	38		
cited						
degree						
tvog_publication						
1.	Hanspeter Pfister	Krzysztof Z. Gajc	Desney S. Tan	Ben Shneiderman	3	
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						
1.	Hanspeter Pfister	Jean-Daniel Fekete	Catherine Plaisant	Ben Shneiderman	3	
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						
4.	Hanspeter Pfister	Jean-Daniel Fekete	Catherine Plaisant	Jennifer Golbeck	Ben Shneiderman	4
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						
4.	Hanspeter Pfister	Jean-Daniel Fekete	Wendy E. Macka	Ed Hui-hsin Chi	Ben Shneiderman	4
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						
4.	Hanspeter Pfister	Krzysztof Z. Gajc	Jeffrey Heer	Ed Hui-hsin Chi	Ben Shneiderman	4
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						
4.	Hanspeter Pfister	Krzysztof Z. Gajc	Jeffrey Heer	Stuart K. Card	Ben Shneiderman	4
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						
4.	Hanspeter Pfister	Jean-Daniel Fekete	Catherine Plaisant	Krist Wongsupha	Ben Shneiderman	4
CHI						
TVCG						
chi_publications						
cited						
degree						
tvog_publication						

Path Topology

Active Page All





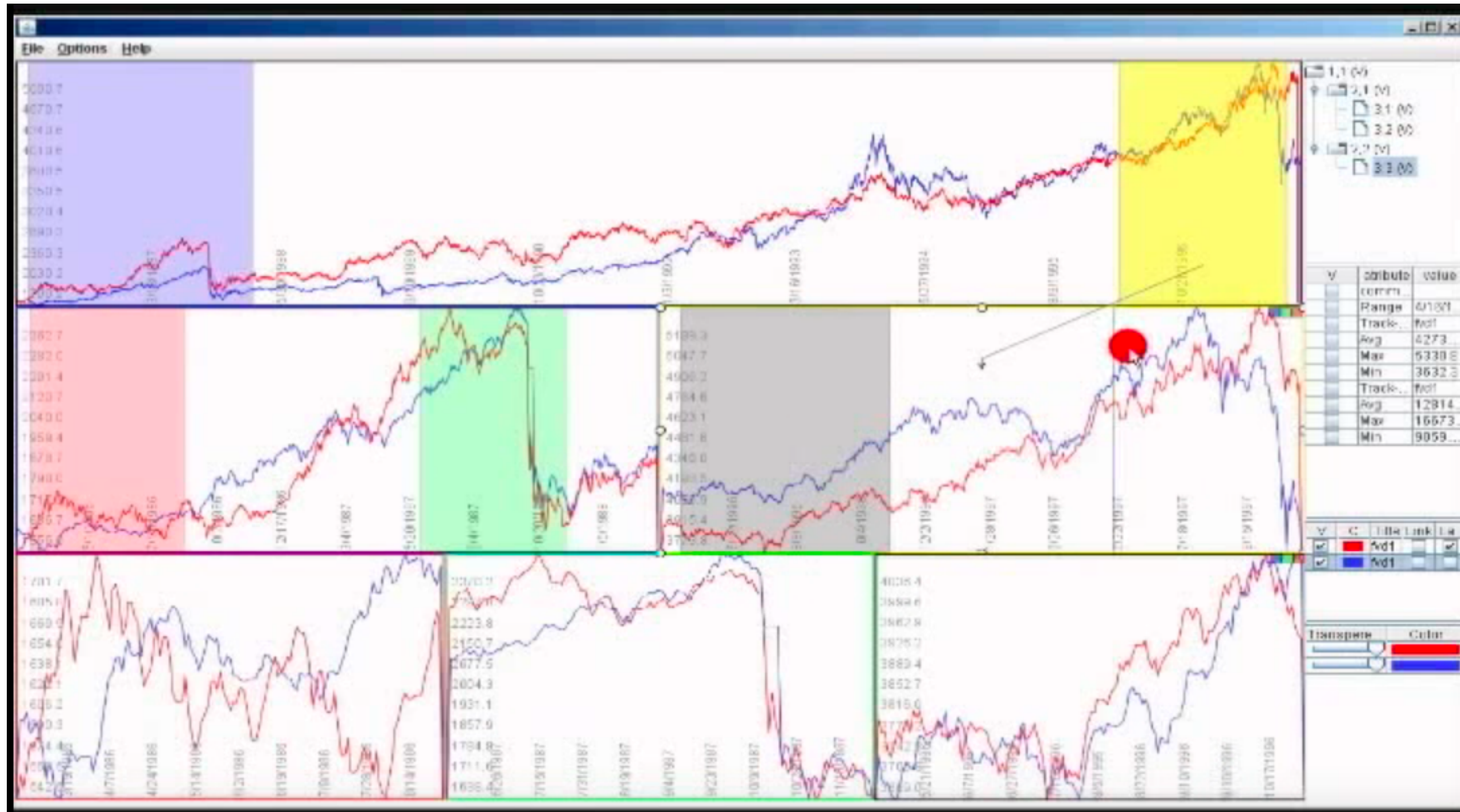
# OVERVIEW + DETAIL

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

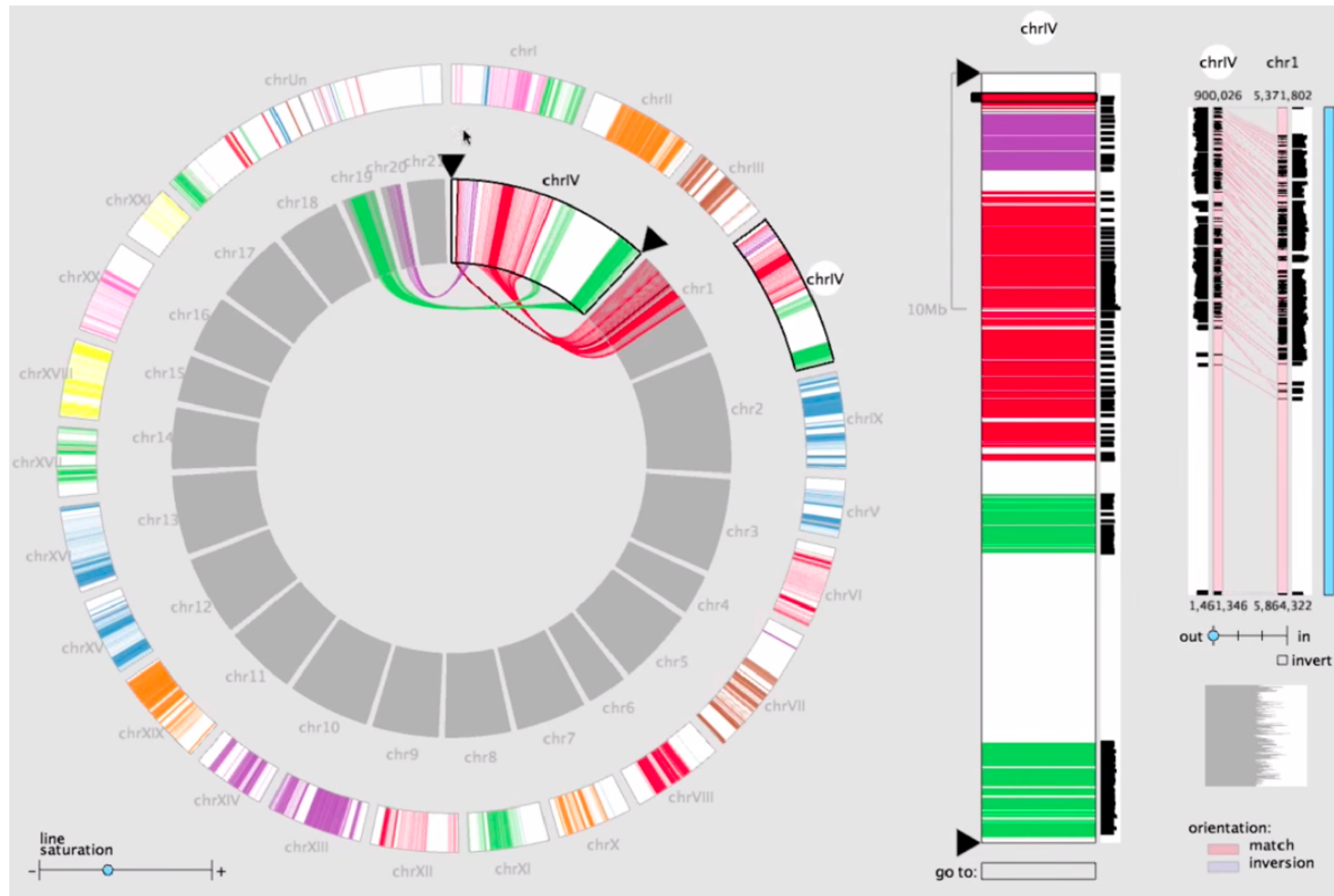
# Stack Zooming



Same Data - Same Encoding, Different Resolution



# MizBee

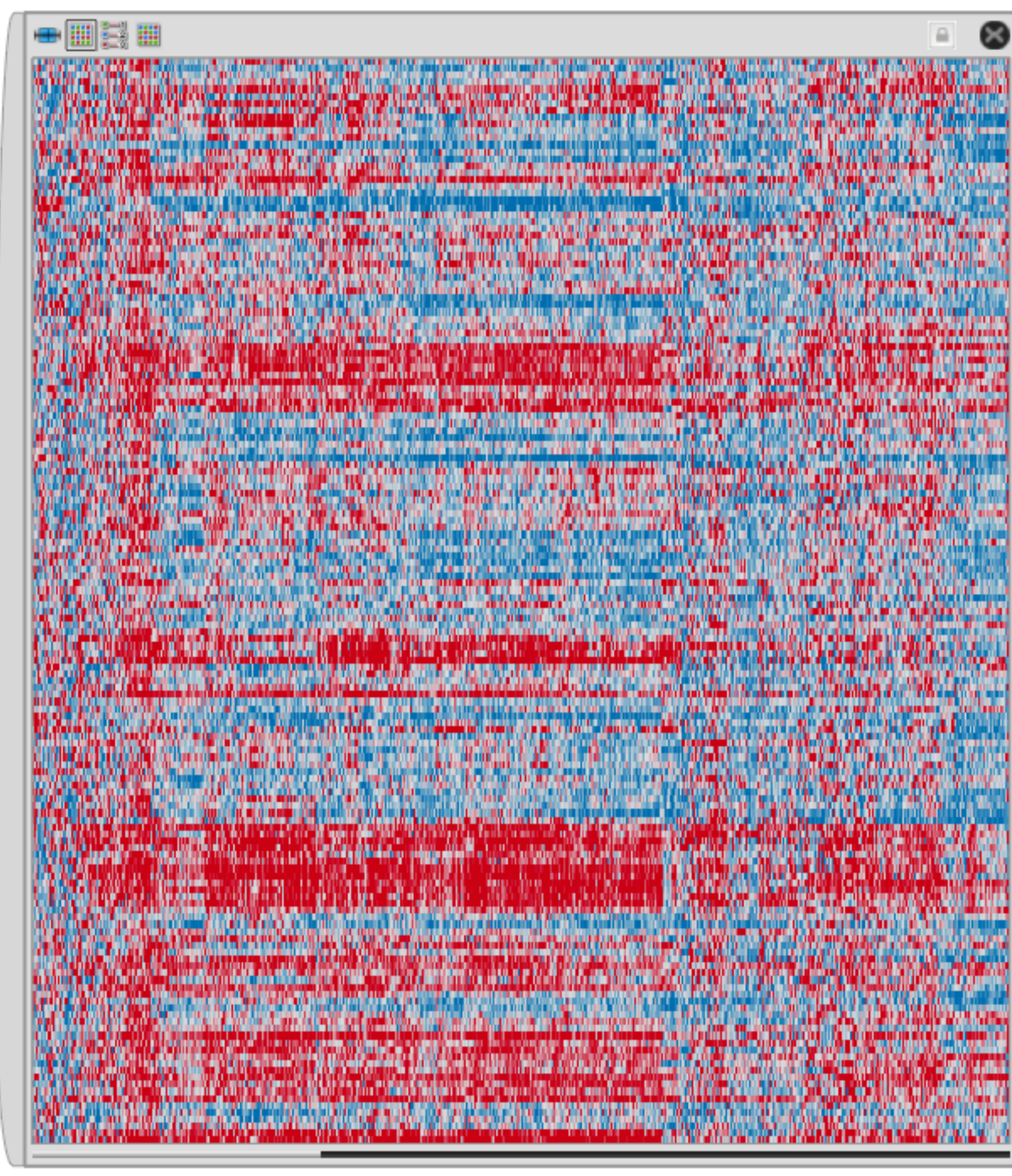
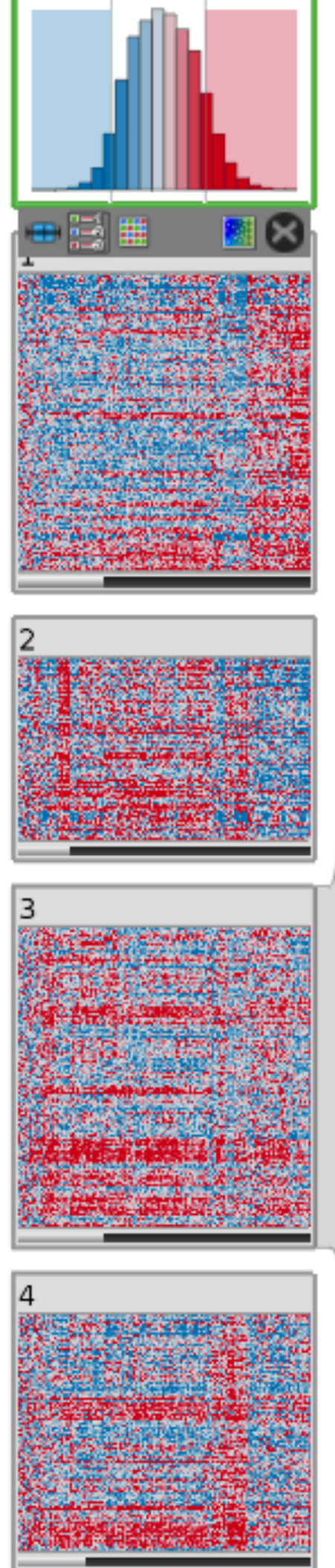


## Multiform Overview & Detail

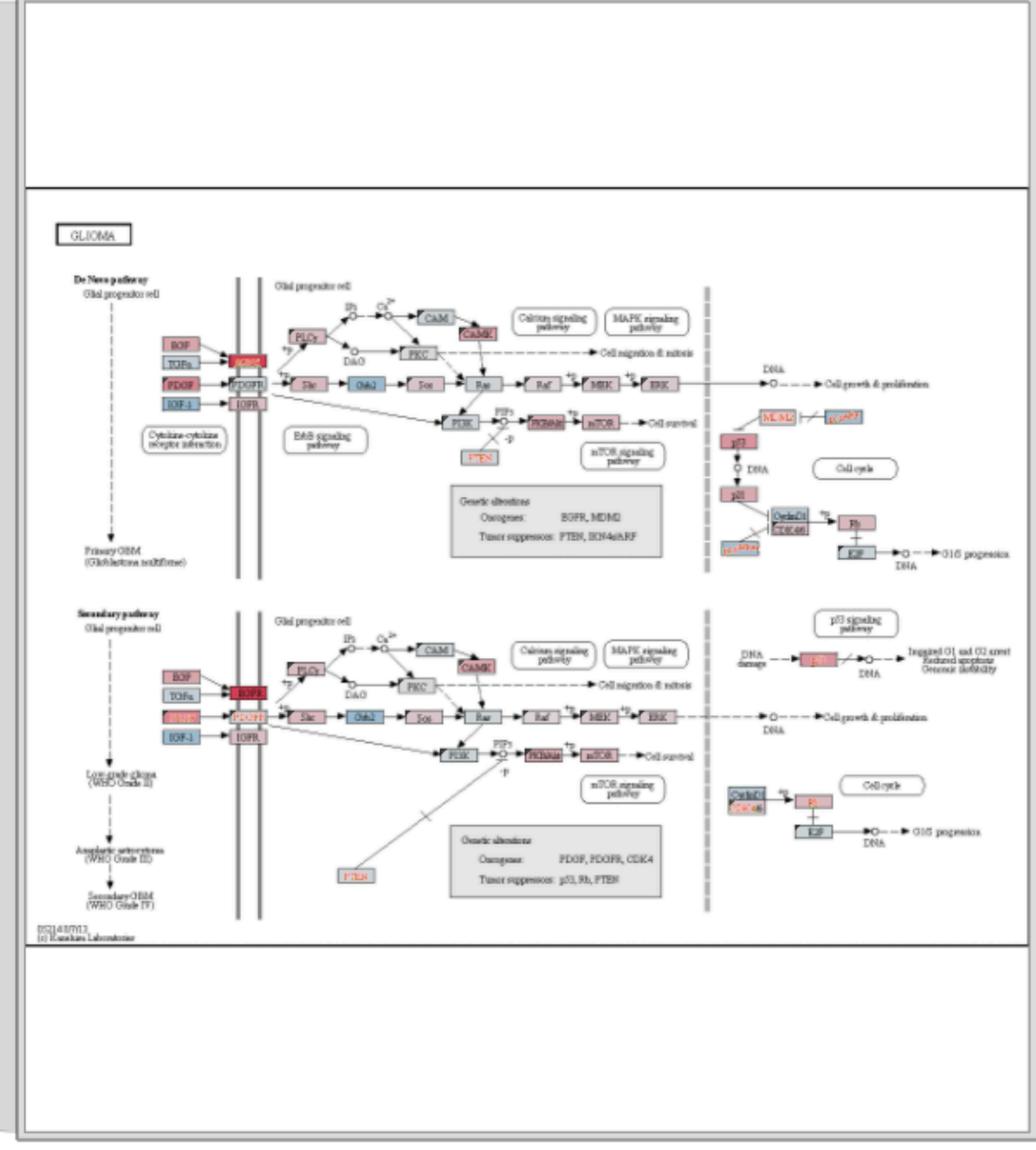


# StratomeX

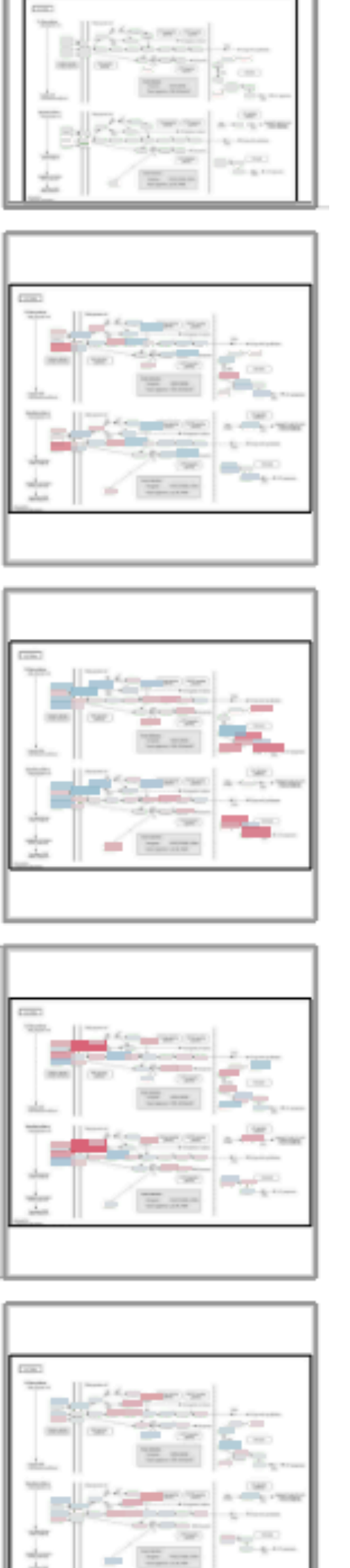
mRNA - 4 CNMF Clus



Glioma



Glioma





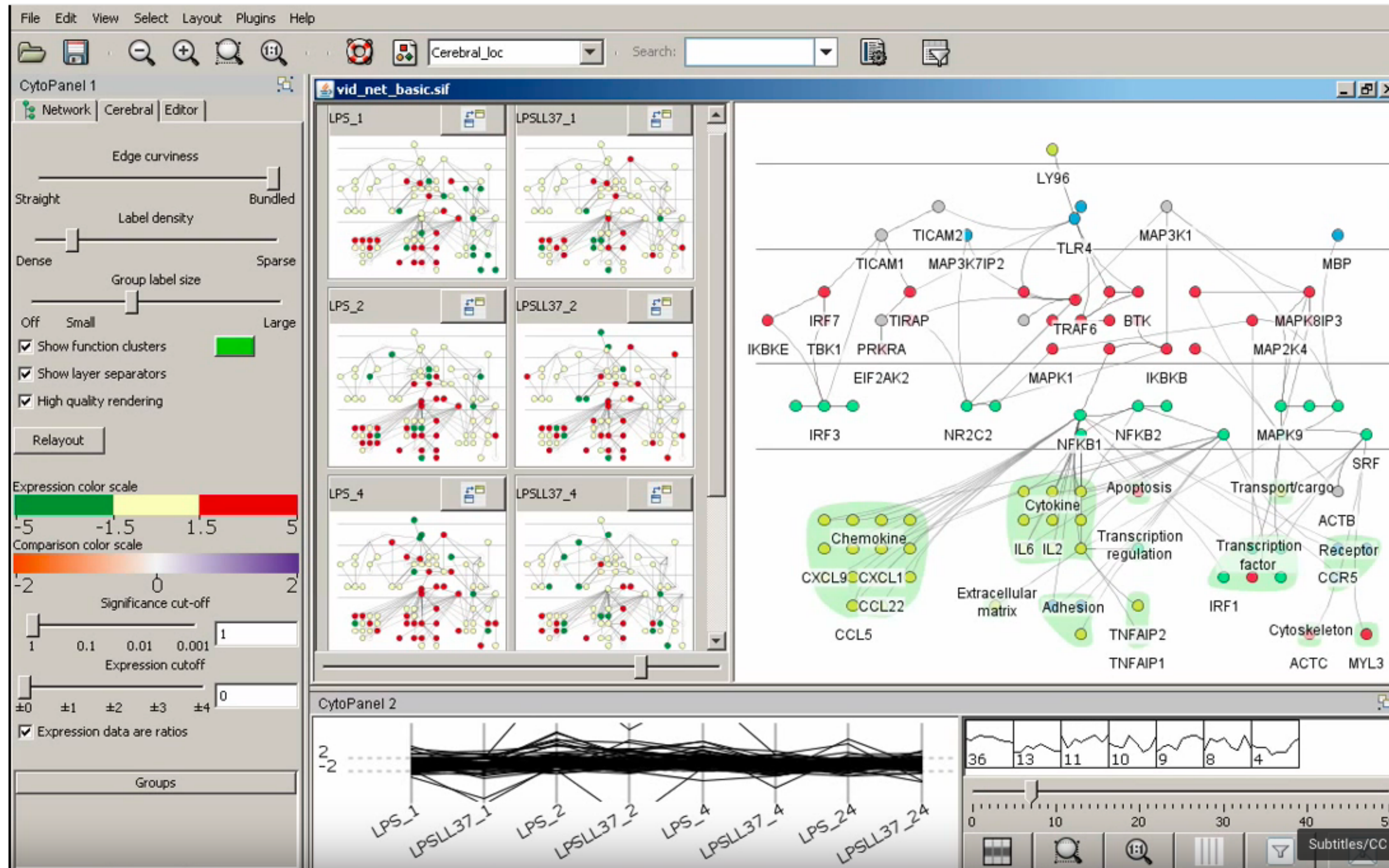
# SMALL MULTIPLES

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

## **rational**

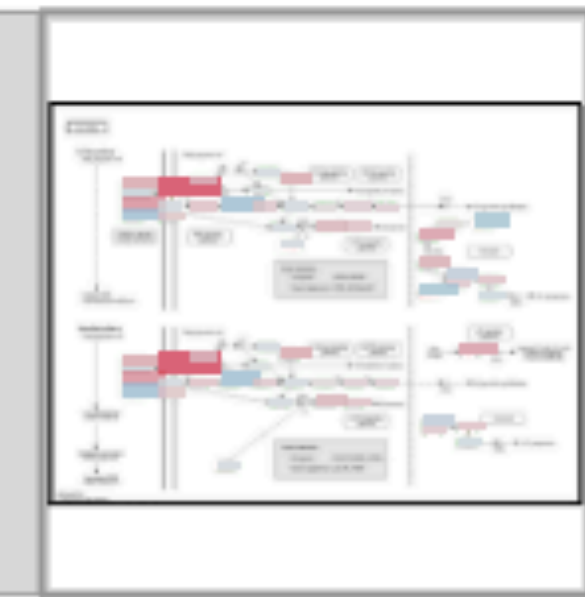
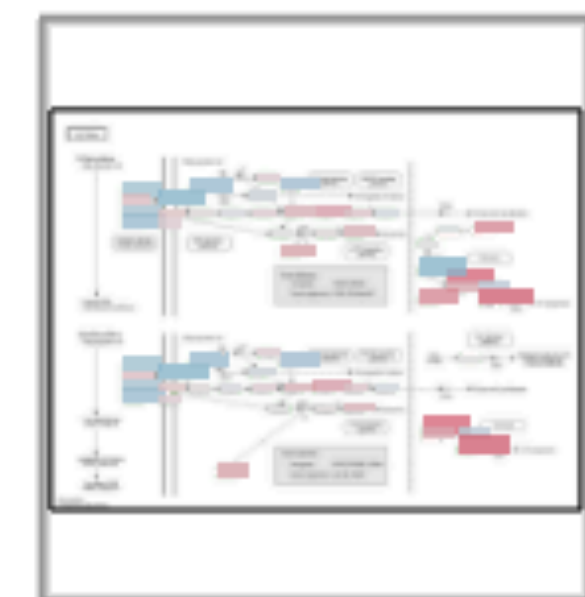
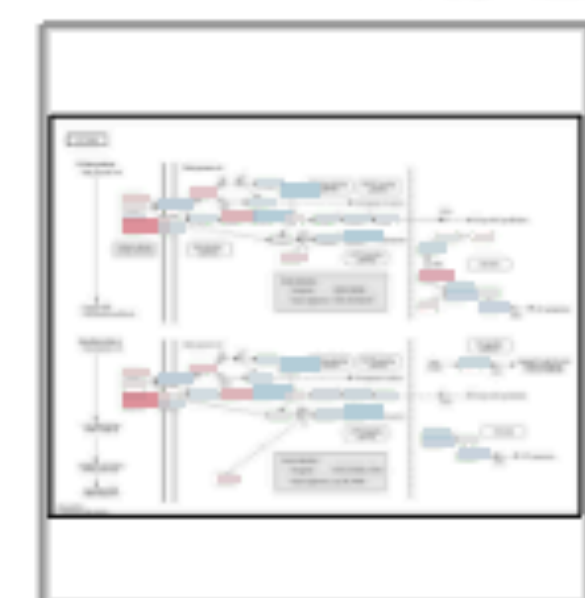
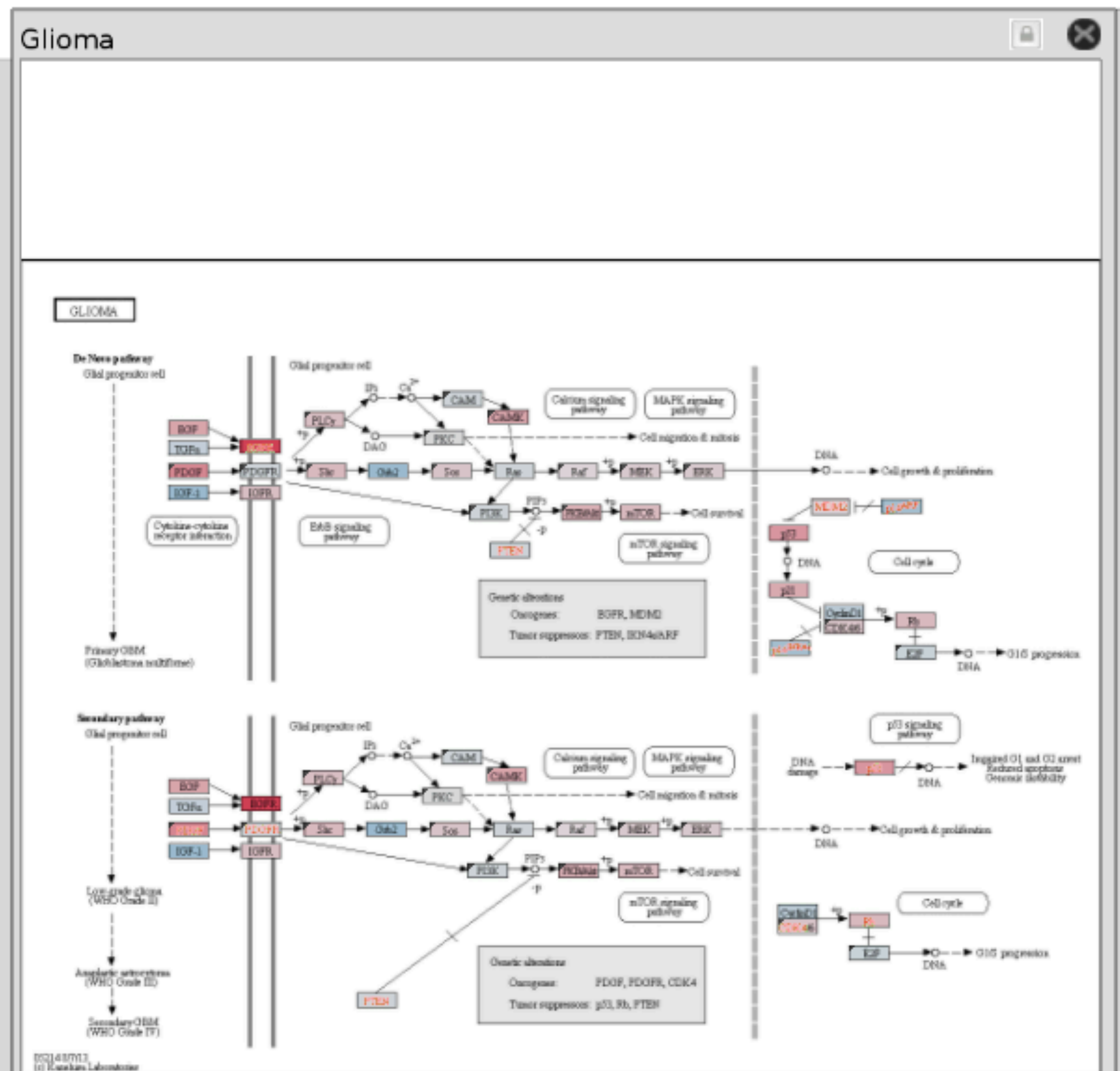
quickly compare different parts of a data set, relying on eyes instead of memory

# Small Multiples for Graph Attributes





# StratomeX





# 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

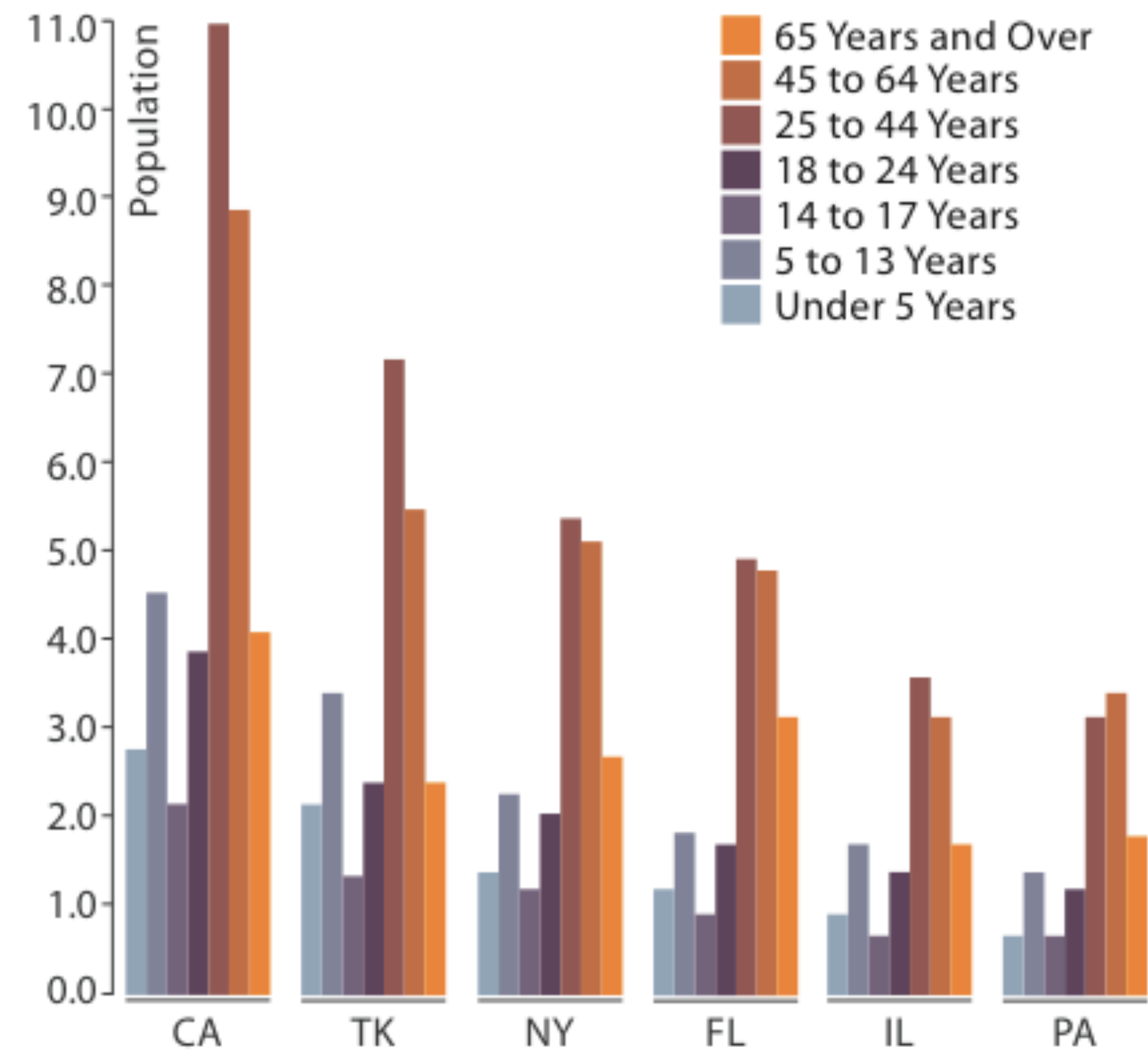
- how many splits, and order of splits

- how many views (usually data driven)

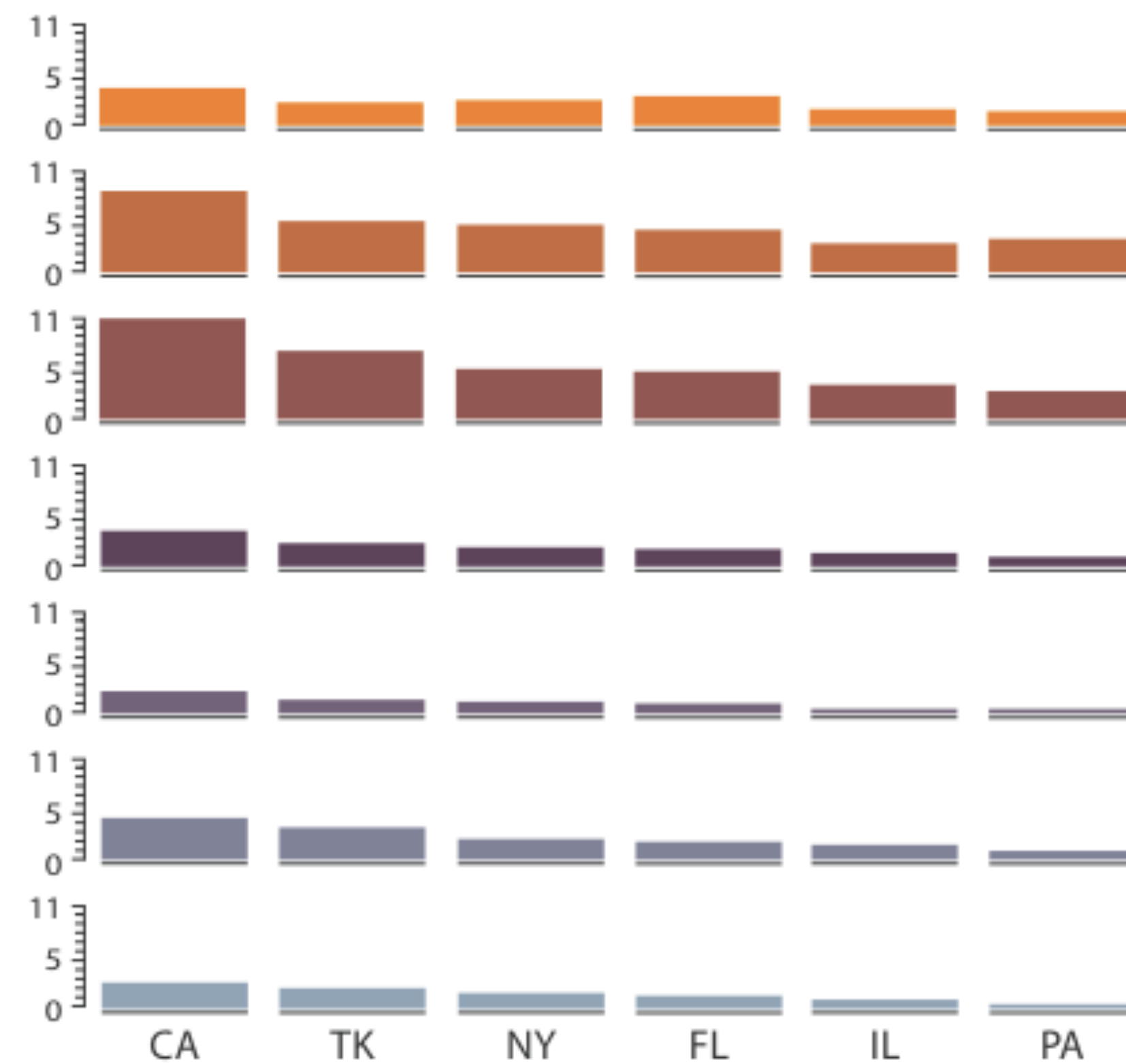
## **partition attribute(s)**

- typically categorical

# Partitioning

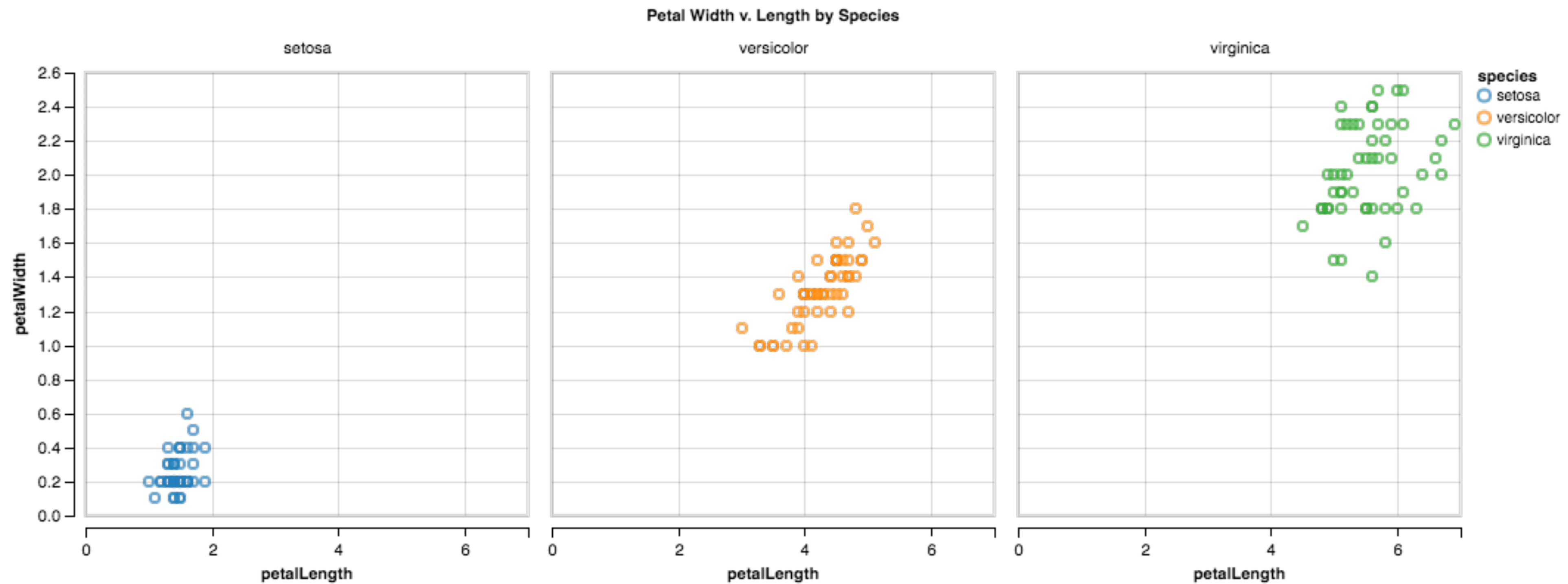


Partitioned by State



Partitioned by Age Group and State

# Partition by Category



# Trellis Plots

panel variables

attributes encoded in individual views

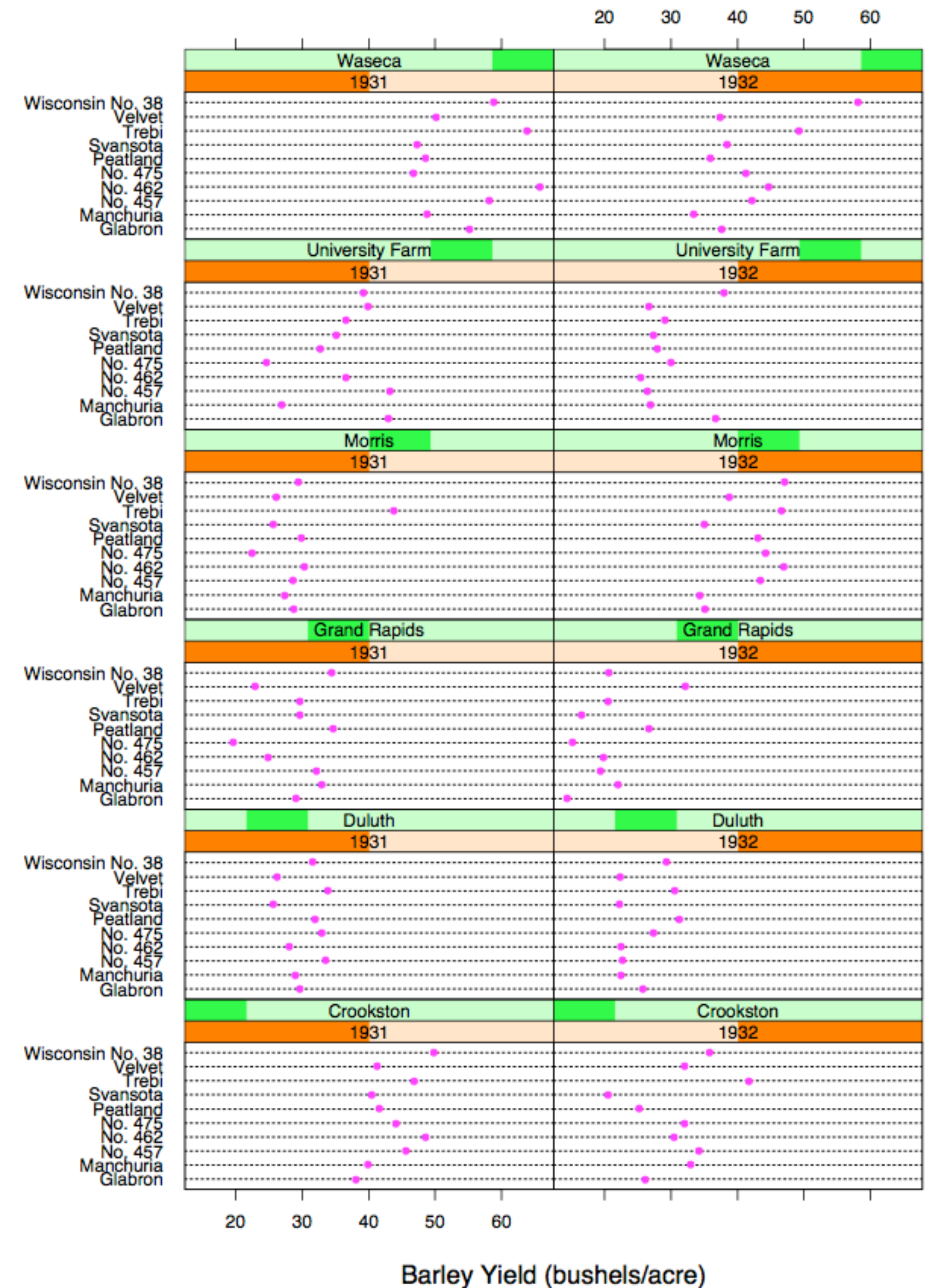
partitioning variables

partitioning attributes assigned to columns,  
rows, and pages

main-effects ordering

order partitioning variable levels/states  
based on derived data

support perception of trends and structure in  
data





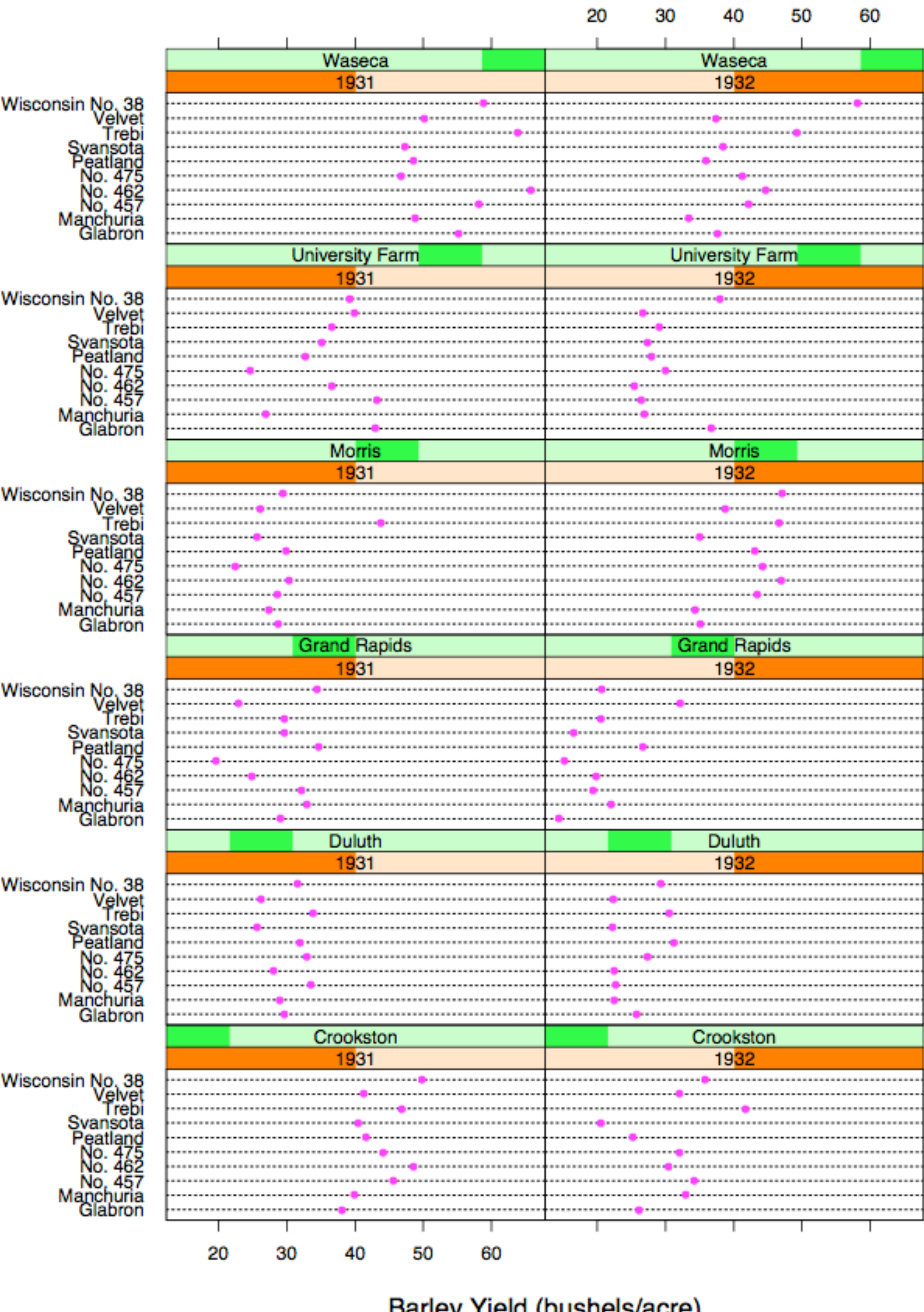
Data

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

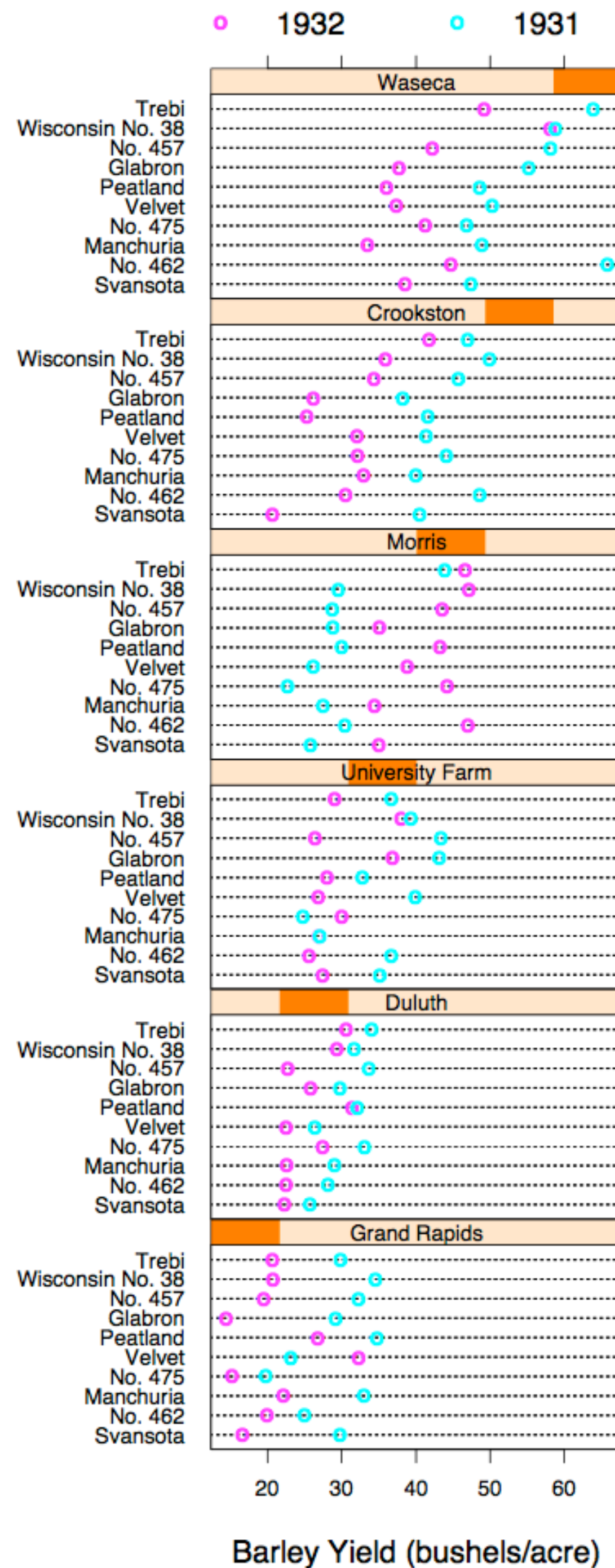
partitioning variables

Columns partitioned by year

Rows partitioned by farm









## Exploring Match Statistics for 2014 World Cup CS-S630/6630 Homework

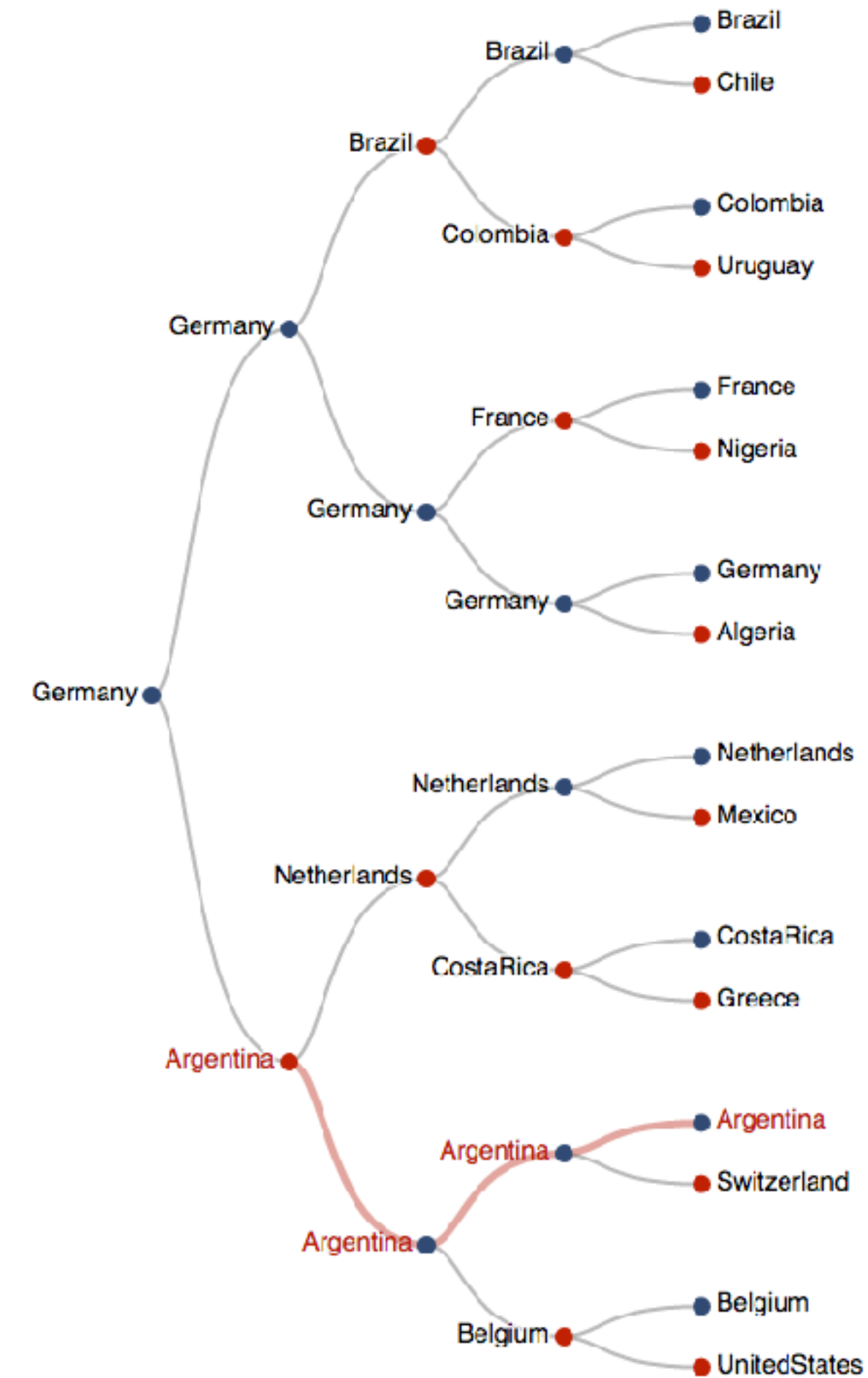
Name: YOURNAME; E-Mail: YOUREMAIL; UID: u0123456



## Score Table

Team	Goals	Round/Result	Wins	Losses	Total Games
Brazil		Third Place	4	2	7
Germany		Winner	6	0	7
Argentina		Runner-Up	6	1	7
xGermany		Runner-Up			
xNetherlands		Semi Finals			
xBelgium		Quarter Finals			
xSwitzerland		Round of Sixteen			
xBosnia		Group			
xIran		Group			
xNigeria		Group			
Netherlands		Fourth Place	6	1	7
France		Quarter Finals	3	1	5
Costa Rica		Quarter Finals	3	1	5
Belgium		Quarter Finals	4	1	5
Chile		Round of Sixteen	2	2	4
Uruguay		Round of Sixteen	2	2	4
Nigeria		Round of Sixteen	1	2	4
Algeria		Round of Sixteen	1	2	4
Colombia		Quarter Finals	4	1	5
Mexico		Round of Sixteen	2	1	4
Greece		Round of Sixteen	1	2	4
Switzerland		Round of Sixteen	2	2	4
United States		Round of Sixteen	1	2	4
Spain		Group		2	3
England		Group		2	3
Cote d'Ivoire		Group	1	2	3
Iran		Group		2	3
Ghana		Group		2	3
Russia		Group		1	3
Australia		Group		3	3
Cameroon		Group		3	3
Japan		Group		2	3
Italy		Group	1	2	3
Honduras		Group		3	3
Korea		Group		2	3
Croatia		Group	1	2	3
Bosnia		Group	1	2	3
Ecuador		Group	1	1	3
Portugal		Group		1	3

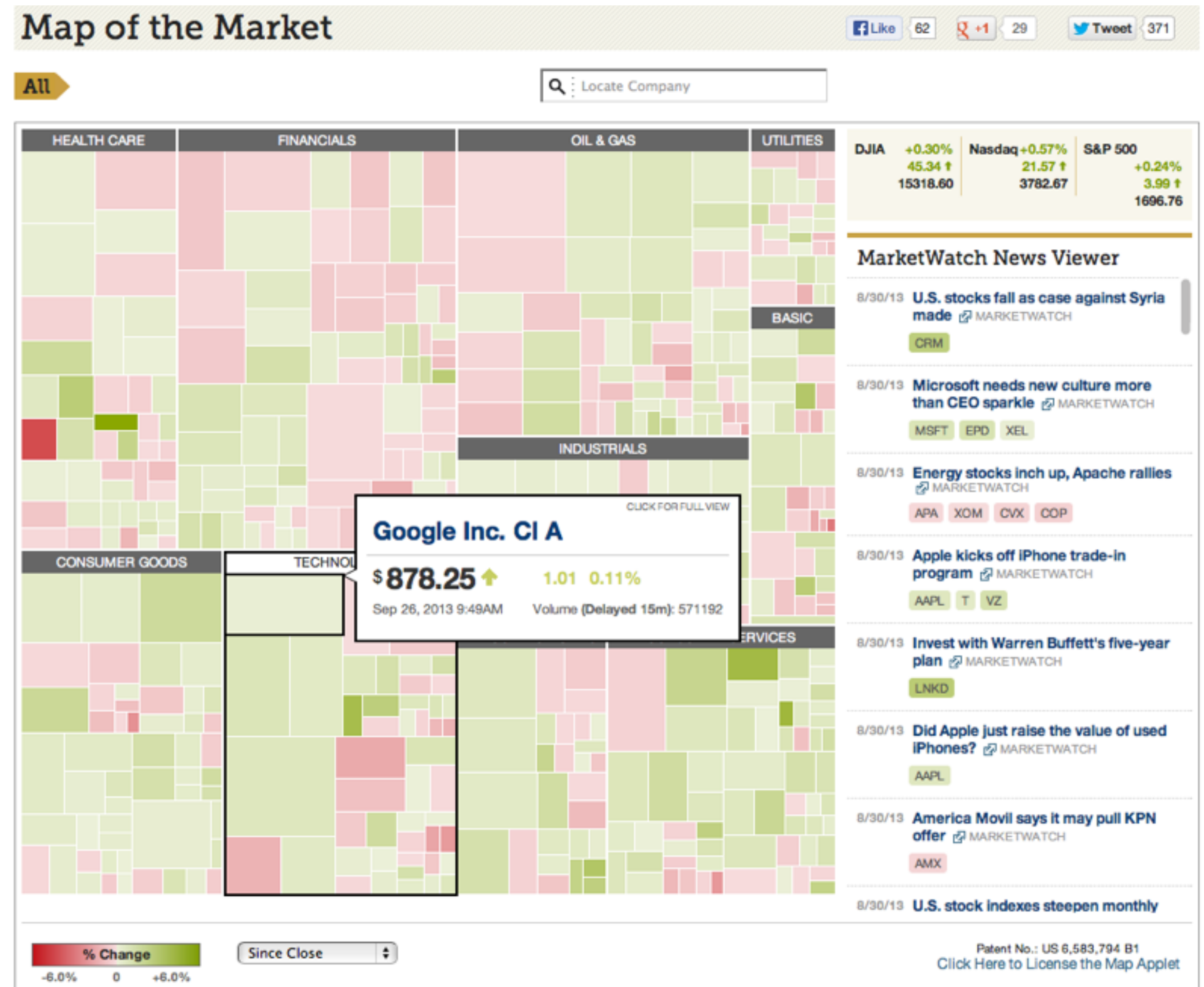
## Second Stage



# 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



# HiVE example: London property

## partitioning attributes

neighborhood location

neighborhood

house type

sale time (year)

sale time (month)

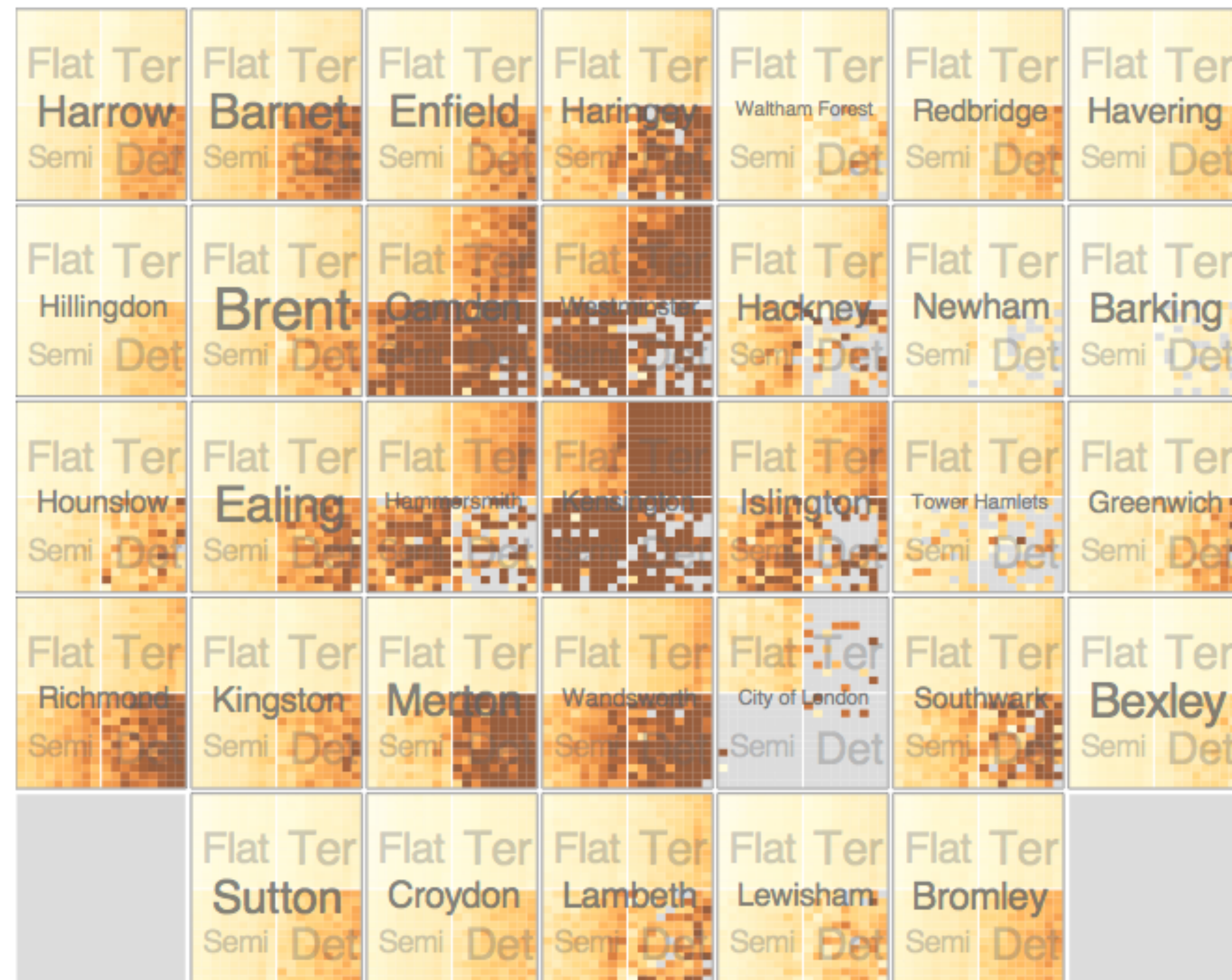
## encoding attributes

average price (color)

*n/a* (size)

## results

expensive neighborhoods  
near center of city





# Configuring Hierarchical Layouts to Address Research Questions



CITY UNIVERSITY  
LONDON

Aidan Slingsby, Jason Dykes and Jo Wood

giCentre, Department of Information Science, City University London

[http://www.gicentre.org/hierarchical\\_layouts/](http://www.gicentre.org/hierarchical_layouts/)



CITY UNIVERSITY  
LONDON

# LAYERING

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

## **rational**

supports a larger, more detailed view than using multiple views

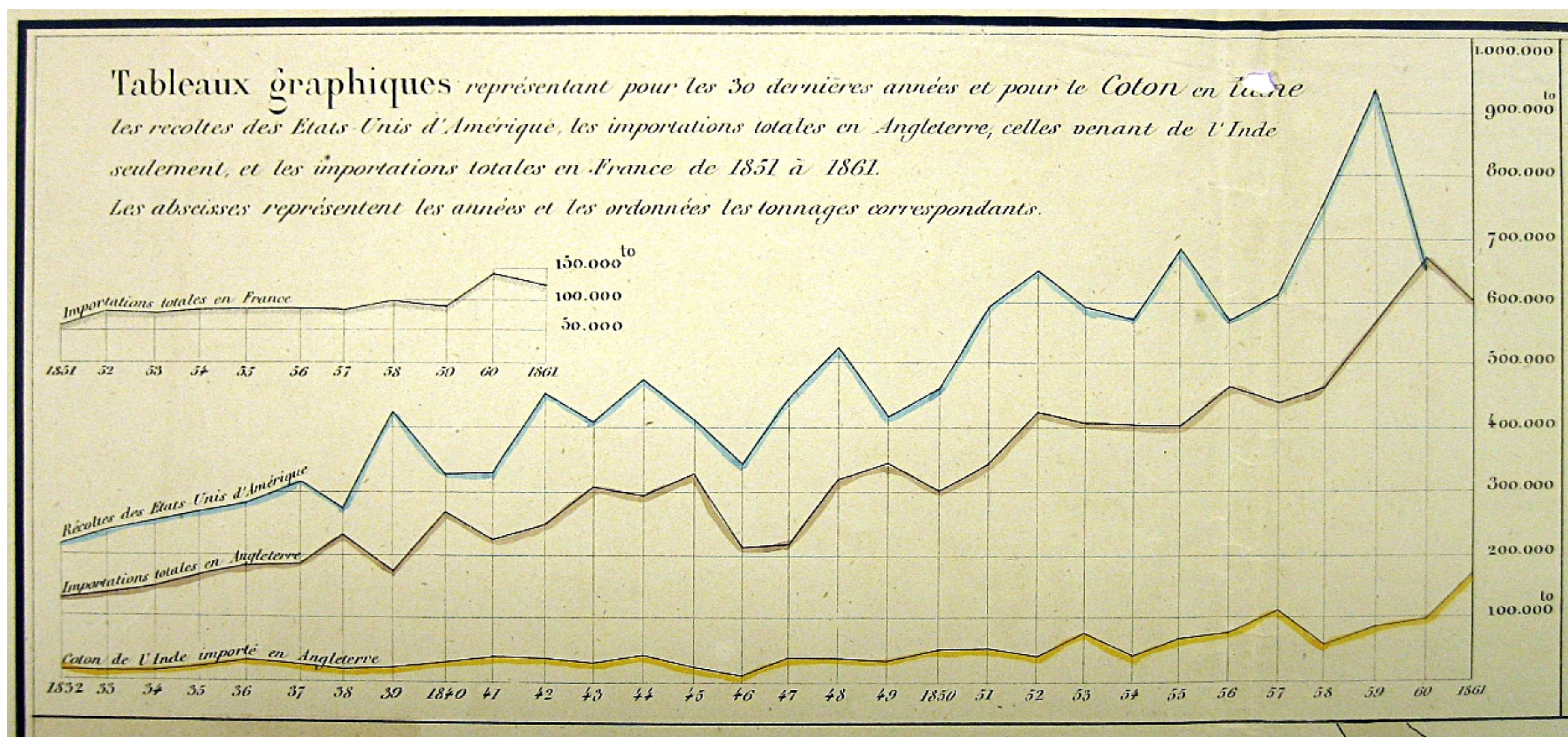
## **trade-off**

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



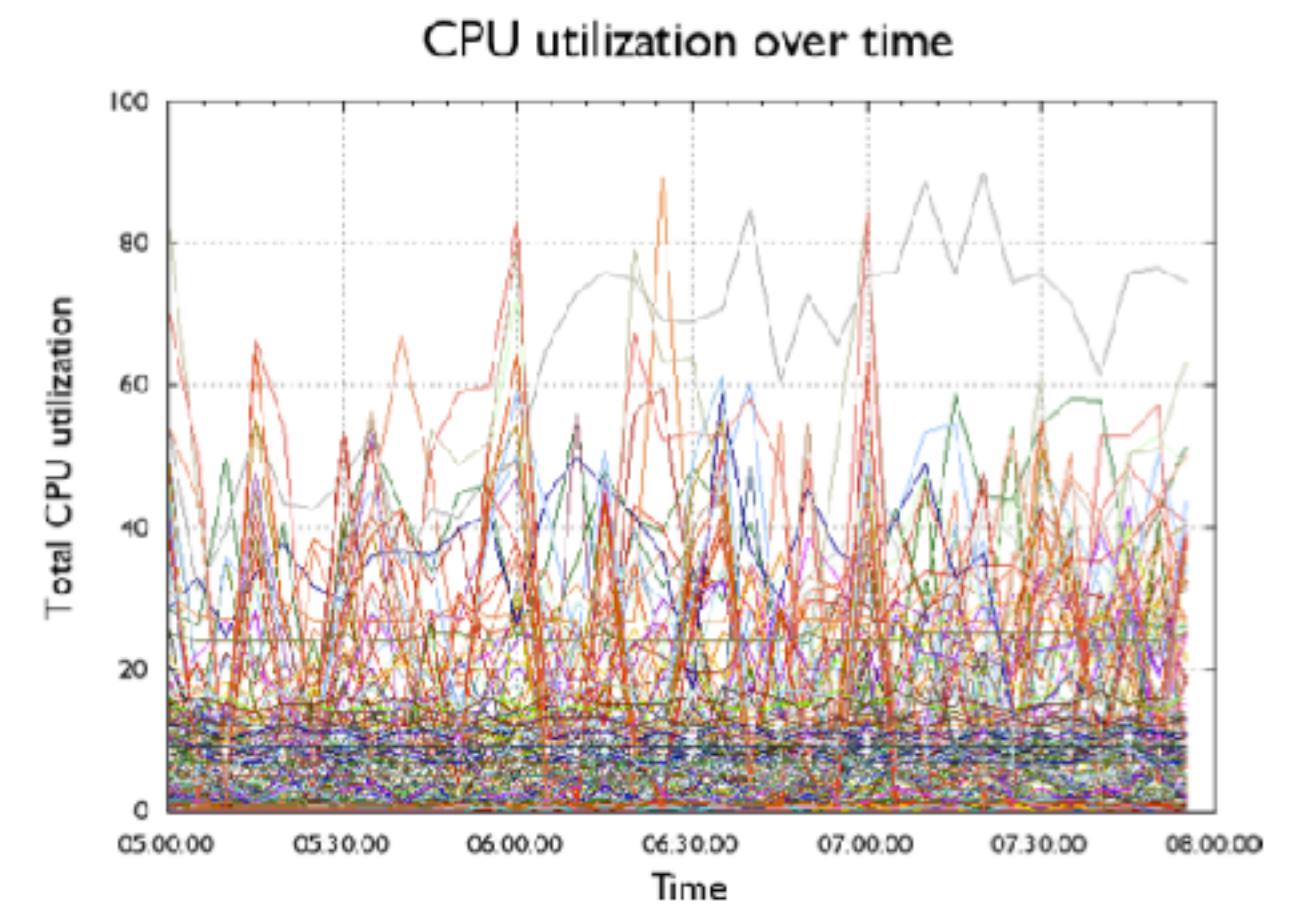
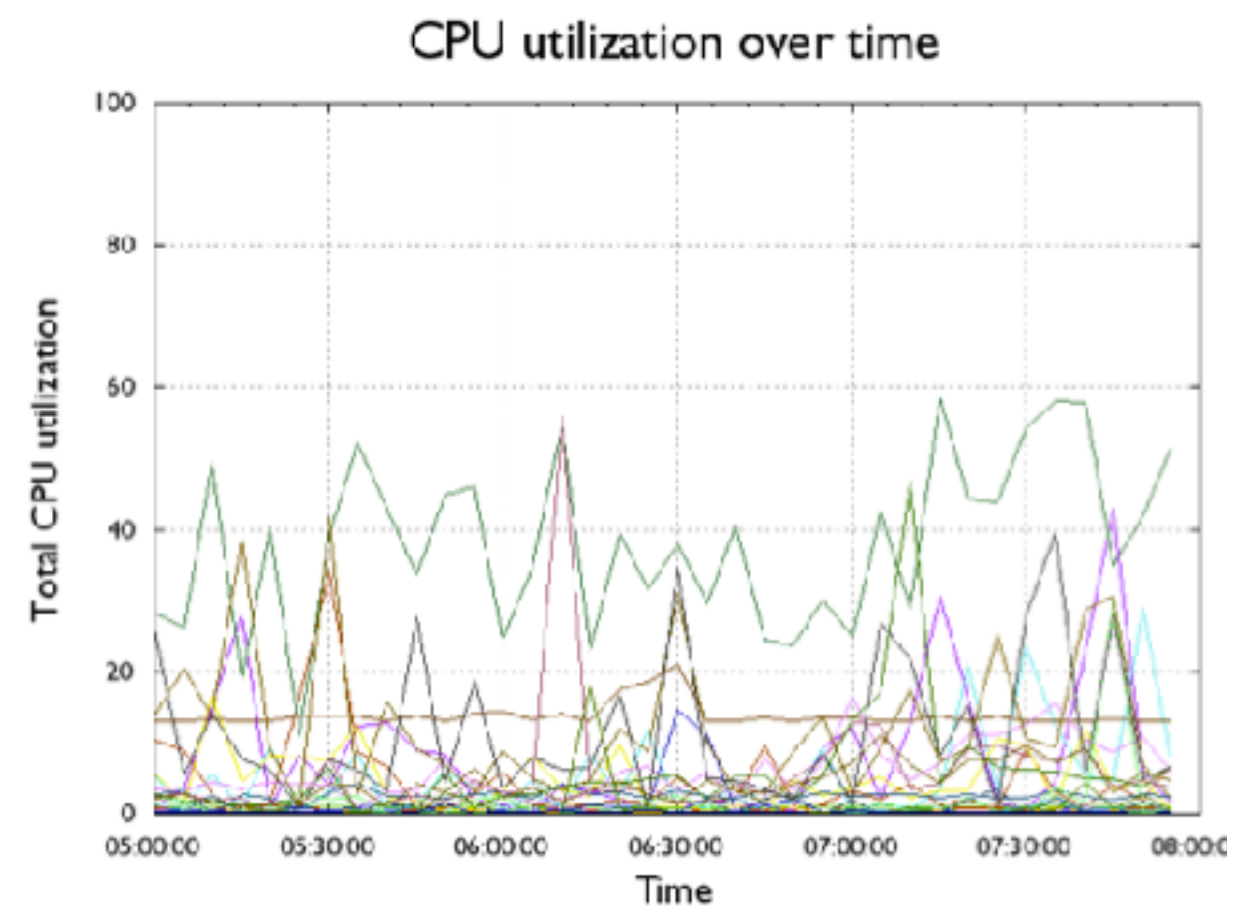
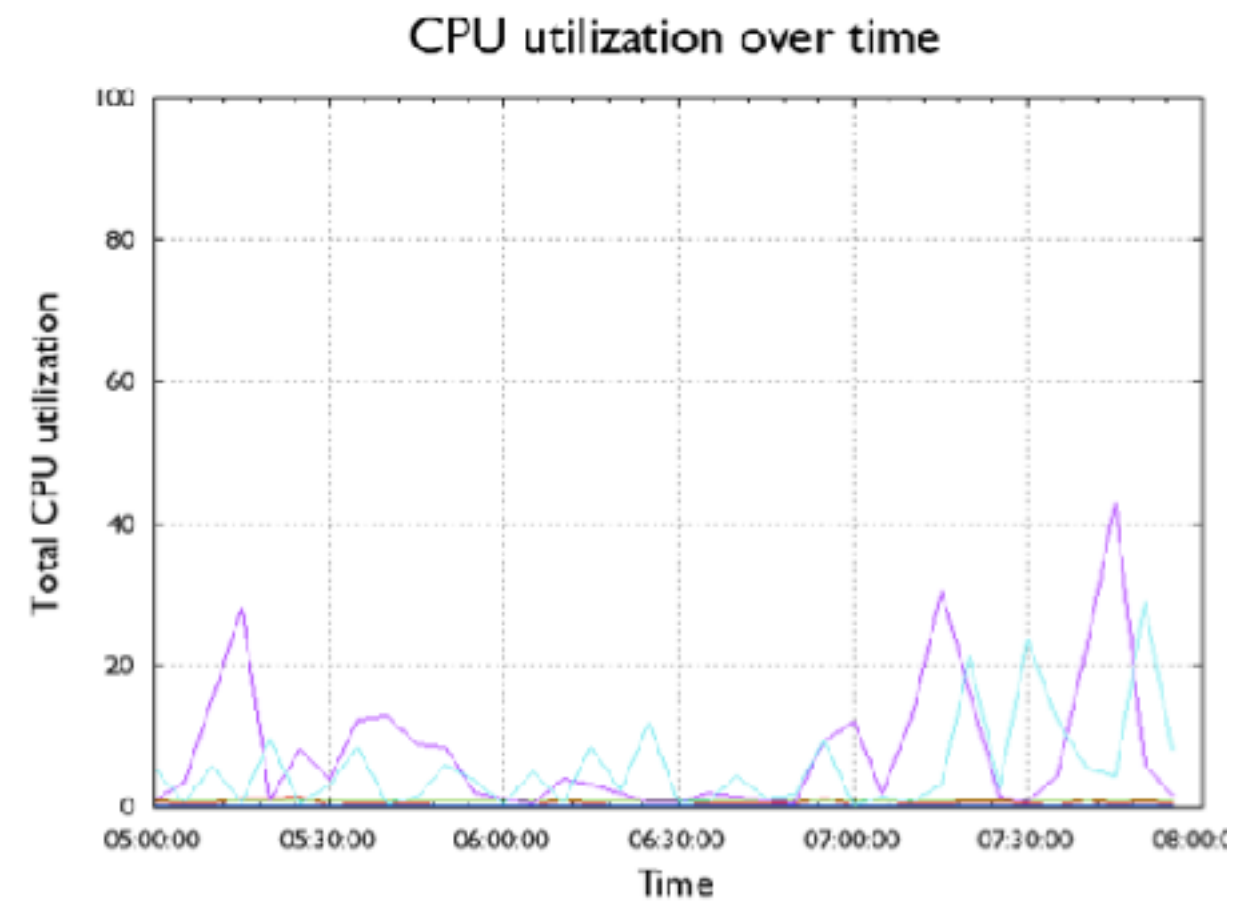
# JOSEPH MINARD

1781-1870





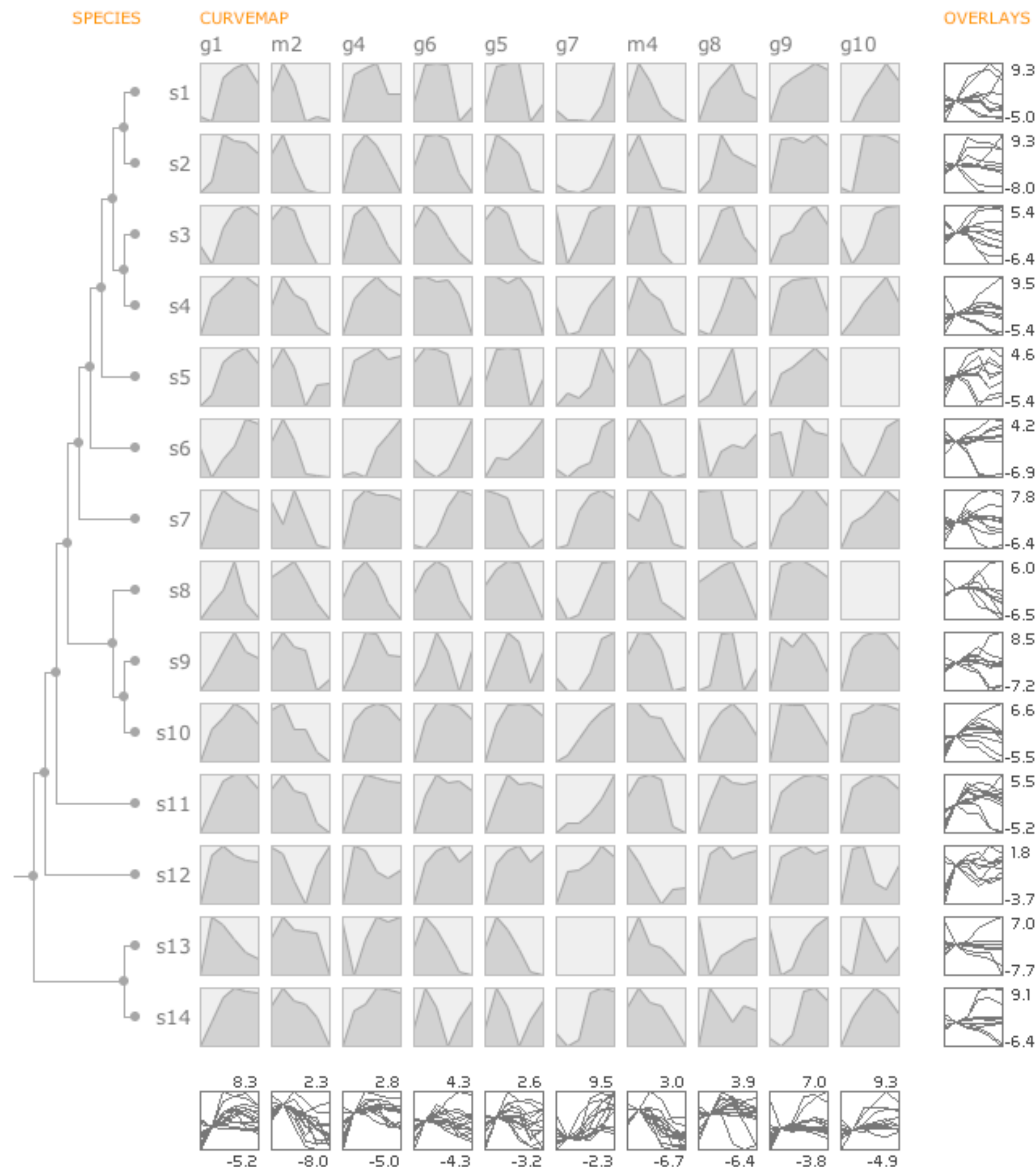
# overlays



# Combined

Partitioned + layered graph

Synchronized through  
highlighting





# MCV to the Max

