

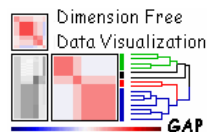
GAP Software Tutorial

Version 0.2 Build20061210

Han-Ming Wu and Chun-houh Chen
(吳漢銘，陳君厚)

**Institute of Statistical Science,
Academia Sinica, Taipei, Taiwan**

hmwu@stat.sinica.edu.tw
<http://gap.stat.sinica.edu.tw>



中央研究院 統計科學研究所
Institute of Statistical Science, Academia Sinica

2006/12/22

Outlines

2/37

■ Features

■ Two Demo Datasets

■ Four Steps of GAP

■ Presentation

■ Seriation

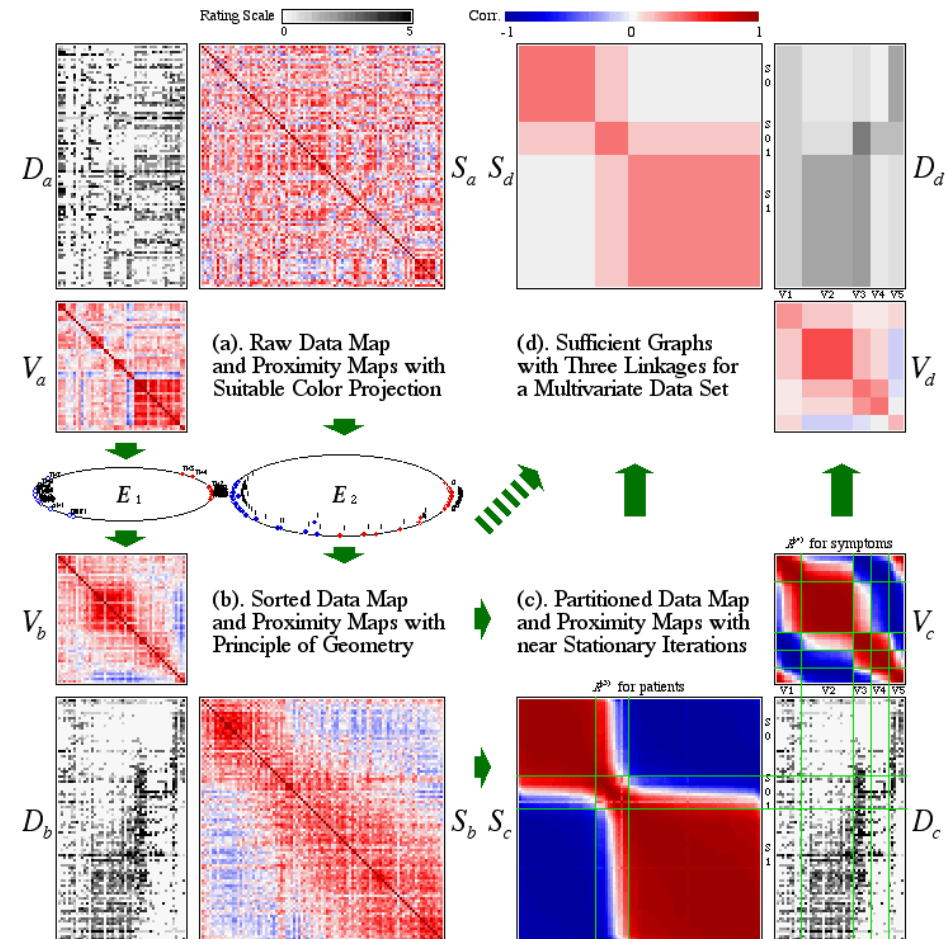
■ Partition

■ Sufficient

Raw Data Matrix and Two Proximity Matrices

■ Other Functionalities

A Complete GAP Procedure



Features version 0.2



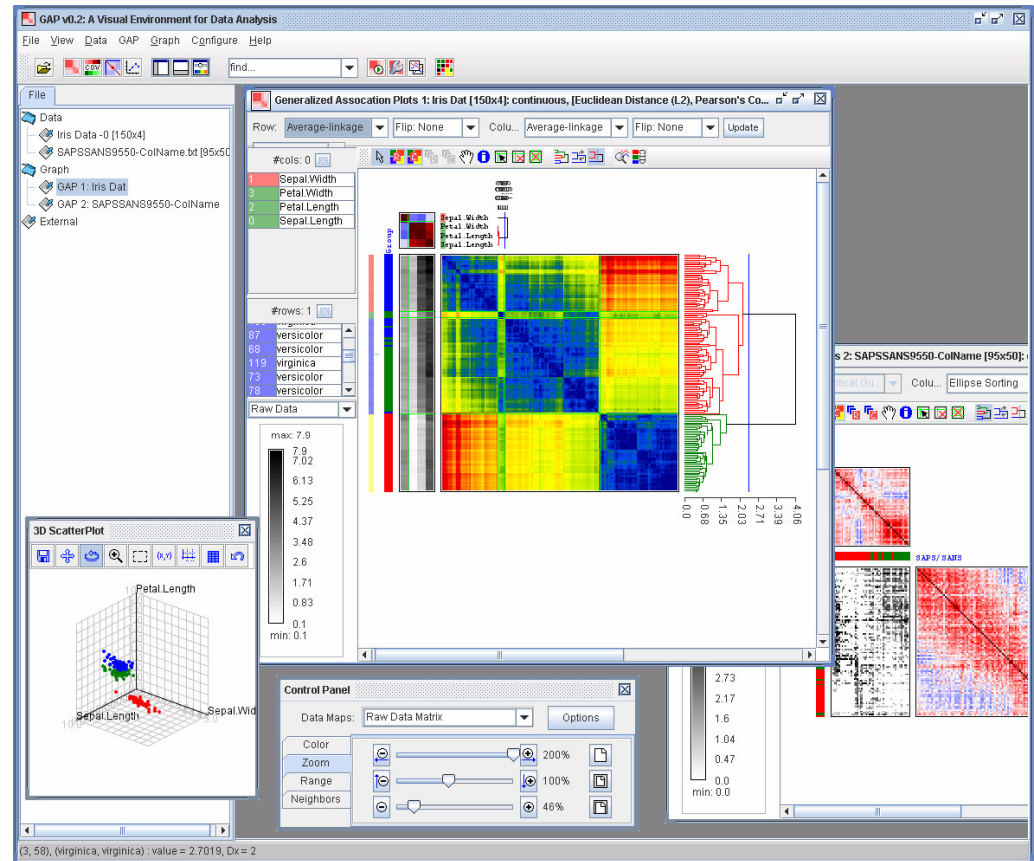
GAP

- Input Data Type: continuous or binary.
- Various seriation algorithms and **clustering analysis**.
- Various display conditions.
- Modules
(Not included in Demo Version):
GAP with Covaraite Adjusted,
Nonlinear Association Analysis,
Missing Value Imputation.



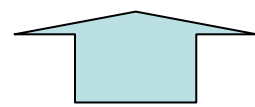
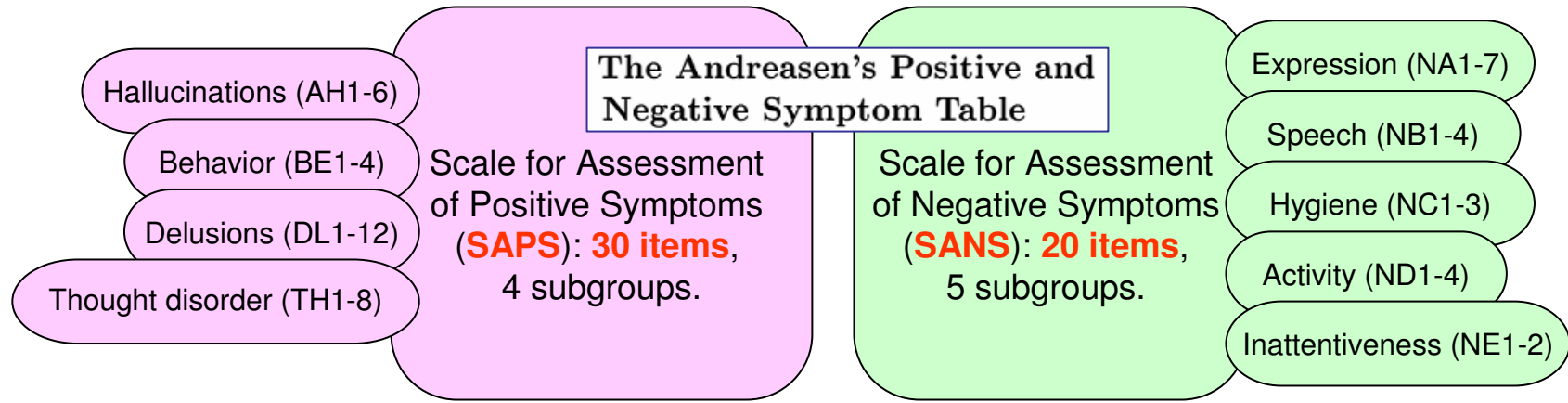
Statistical Plots

- 2D Scatterplot, 3D Scatterplot (Rotatable)



<http://gap.stat.sinica.edu.tw/Software/GAP>

Psychosis Disorder Data



50 Variables

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF
1	AH1	AH2	AH3	AH4	AH5	AH6	DL1	DL2	DL3	DL4	DL5	DL6	DL7	DL8	DL9	DL10	DL11	DL12	BE1	BE2	BE3	BE4	TH1	TH2	TH3	TH4	TH5	TH6	TH7	TH8	TH9	TH10
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	5	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	5	5	5	4	5	5	0	0	5	5	1	5	5	0	5	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	2	0	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	3	1	1	0	0	2	0	0	0	0	1	3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	5	5	1	0	0	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	2	0	2	0	0	0	0	0	0	0	3	0	1	2	2	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	5	0	1	0	0	5	0	0	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	5	0	5	0	0	0	0	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	5	5	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	5	4	0	0	0	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	4	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	3	3	3	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	2	3	1	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	4	4	4	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	3	3	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	4	4	4	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	3	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	2	2	1	1	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	4	4	1	0	0	3	4	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	4	4	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	2	2	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

All the symptoms are recorded on a six point scale (0-5).

Raw Data Matrix

69 schizophrenic

26 bipolar disorders

95 Subjects

Demo Data 2: Iris Data



The sepal length, sepal width, petal length, and petal width are measured in centimeters on 50 iris specimens from each of three species, *Iris setosa*, *I. versicolor*, and *I. virginica*.
Fisher (1936)

Iris Flowers



○ *Iris Setosa*

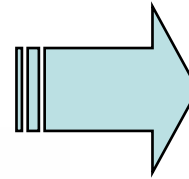


○ *Iris Versicolor*

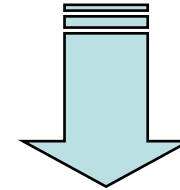


○ *Iris Virginica*

150 subjects



4 variables

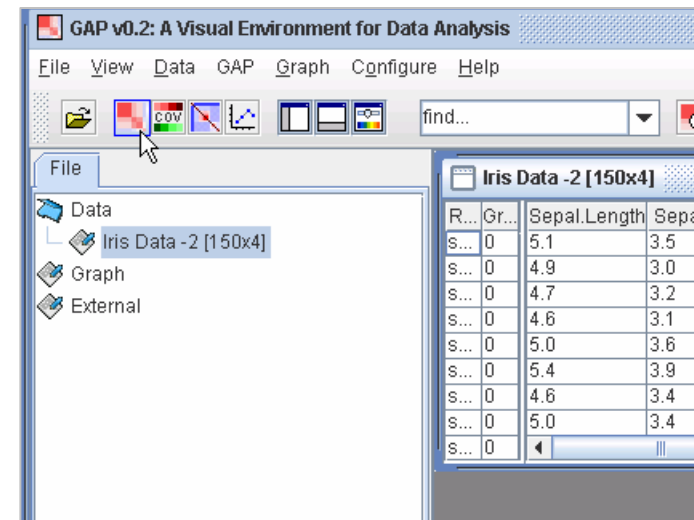
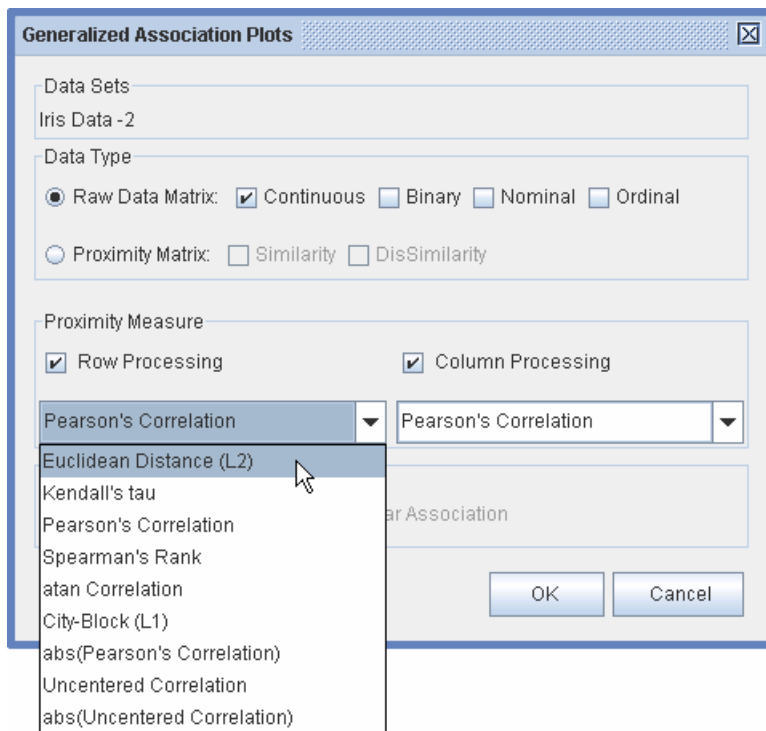
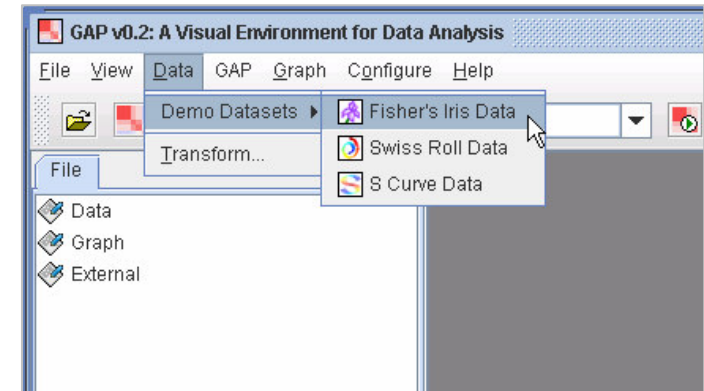
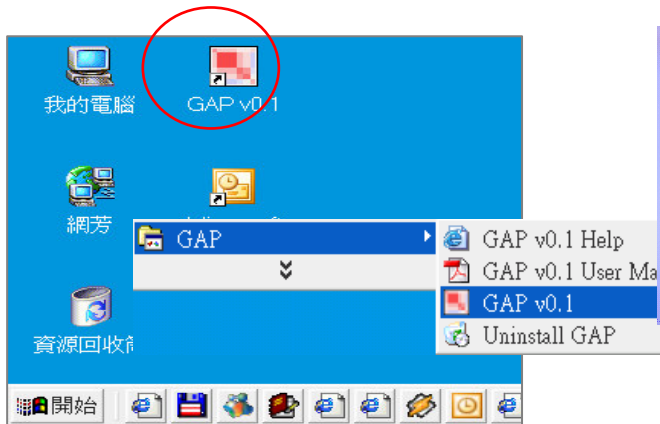


no.	Species	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
1	setosa	5.1	3.5	1.4	0.2
2	setosa	4.9	3.0	1.4	0.2
3	setosa	4.7	3.2	1.3	0.2
4	setosa	4.6	3.1	1.5	0.2
5	setosa	5.0	3.6	1.4	0.2
...					
76	versicolor	6.6	3.0	4.4	1.4
...					
150	virginica	5.9	3.0	5.1	1.8

Raw Data Matrix

Quick Start

6/37



GAP Main Window



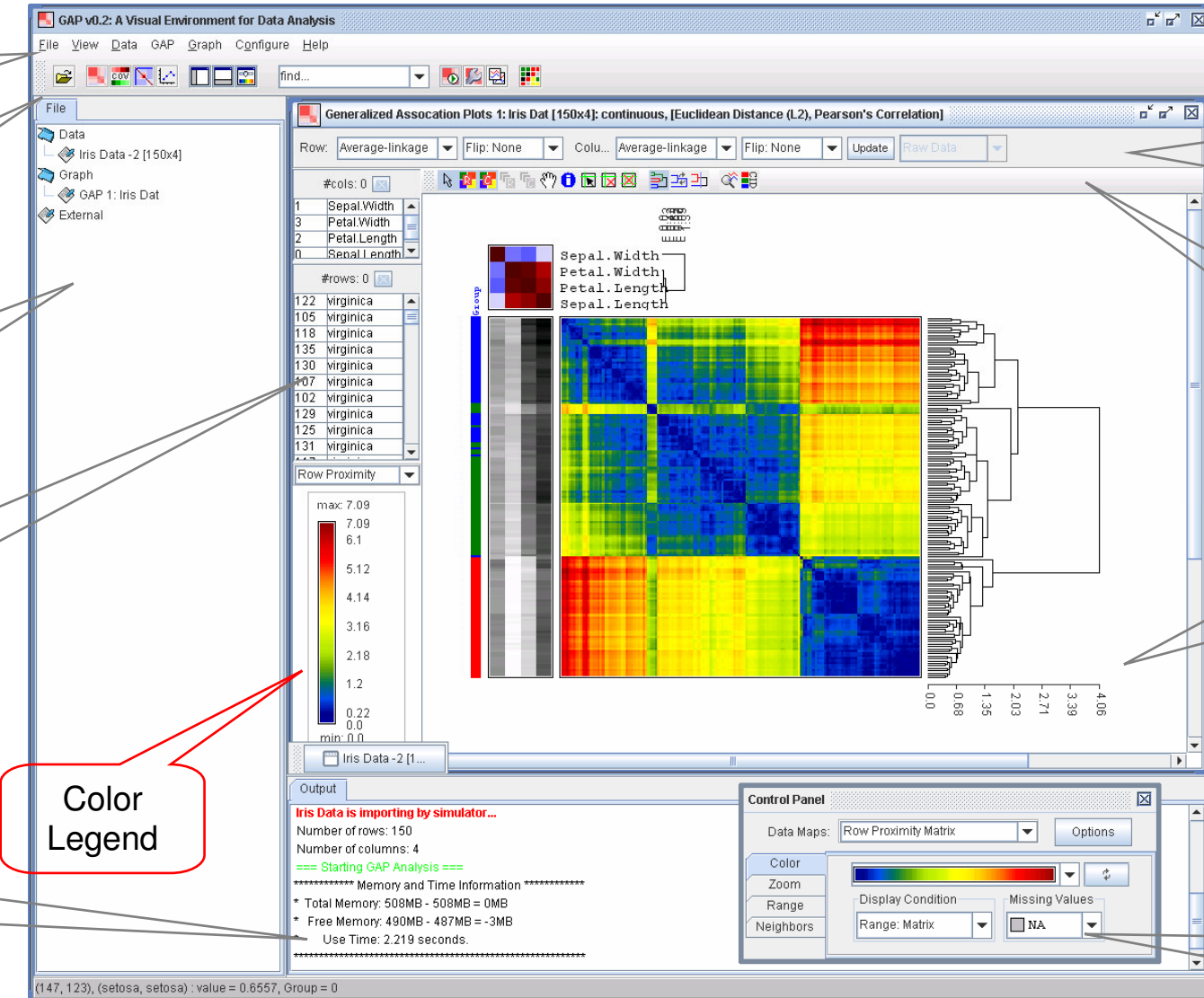
Menu Bar

Tool Bar

File/Data View

Variables and Subjects

Output Window



Serialization and Tree Flip

Mouse Tools

GAP Maps Window

Control Panel

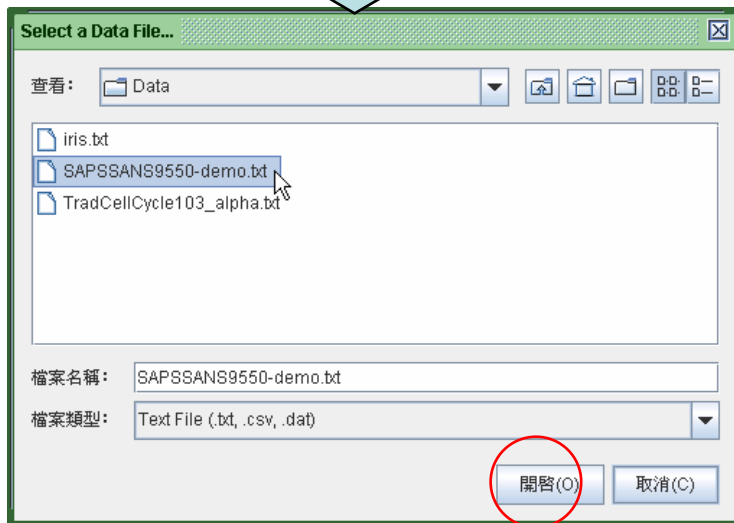
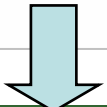
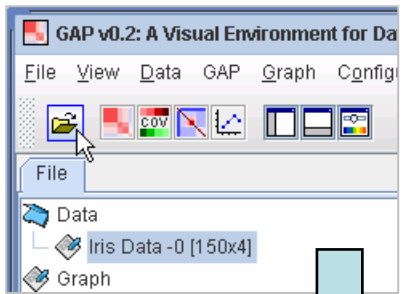
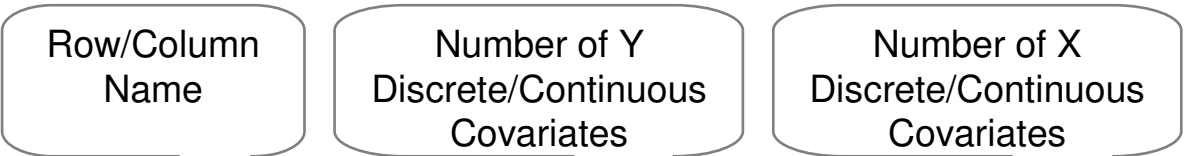
Color Legend

(147, 123), (setosa, setosa) : value = 0.6557, Group = 0

Open Data

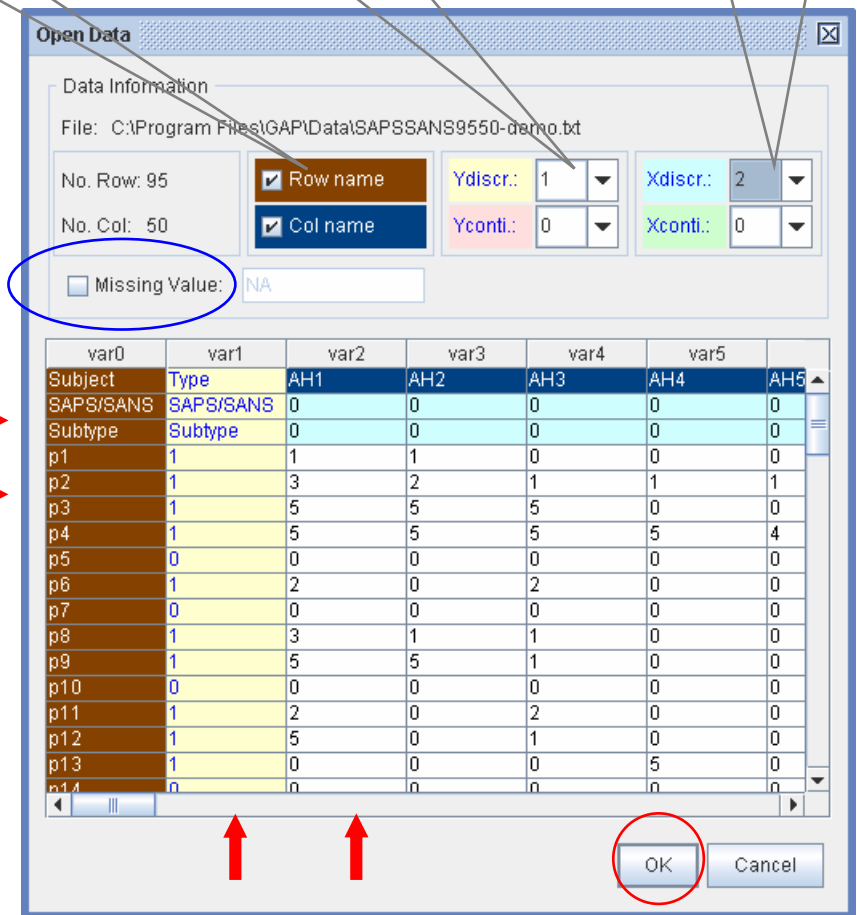
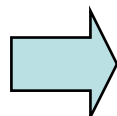
Data Format: plain text with tab separator.

Covariates: Code from 0,1,2,...



X_d →

X_c →



Open Data

Data Information

File: C:\Program Files\GAP\Data\SAPSSANS9550-demo.txt

No. Row: 95

No. Col: 50

Row name

Col name

Ydiscr.: 1

Yconti.: 0

Xdiscr.: 2

Xconti.: 0

Missing Value: NA

var0	var1	var2	var3	var4	var5	var6
Subject	Type	AH1	AH2	AH3	AH4	AH5
SAPS/SANS	SAPS/SANS	0	0	0	0	0
Subtype	Subtype	0	0	0	0	0
p1	1	1	1	0	0	0
p2	1	3	2	1	1	1
p3	1	5	5	5	0	0
p4	1	5	5	5	5	4
p5	0	0	0	0	0	0
p6	1	2	0	2	0	0
p7	0	0	0	0	0	0
p8	1	3	1	1	0	0
p9	1	5	5	1	0	0
p10	0	0	0	0	0	0
p11	1	2	0	2	0	0
p12	1	5	0	1	0	0
p13	1	0	0	0	5	0
p14	0	0	0	0	0	0

Y_d ↑

Y_c ↑

OK Cancel

Y_d **Y_c**

GAP Analysis Dialog

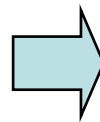
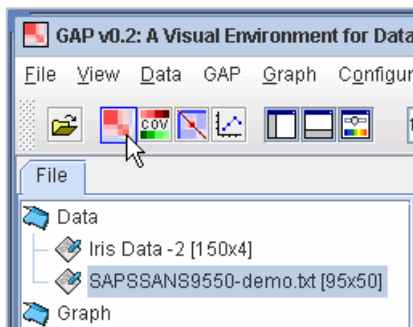
- Euclidean Distance (L2)
- Kendall's tau
- Pearson's Correlation
- Spearman's Rank
- atan Correlation
- City-Block (L1)
- abs(Pearson's Correlation)
- Uncentered Correlation
- abs(Uncentered Correlation)

- Hamman: $a/(b+c)$
- Jaccard: $a/(a+b+c)$
- Phi: $(ad-bc)/\sqrt{(a+b)(a+c)(b+d)(c+d)}$
- Rao: $a/(a+b+c+d)$
- Rogers: $(a+d)/(a+2b+2c+d)$
- Simple Match: $(a+d)/(a+b+c+d)$
- Sneath: $a/(a+2b+2c)$
- Yule: $(ad-bc)/(ad+bc)$

Not implemented Yet!

- Pearson Conti. Coeff.
- Goodman-Kruskal tau
- Cohen's kappa (I=J)

- Kendall's tau_b
- Kendaall-Stuart's tau_c
- Goodman-Kruskal's gamma
- Somers's d
- Wilson's e



Generalized Association Plots

Data Sets: SAPSSANS9550-demo.txt

Data Type: Raw Data Matrix: Continuous Binary Nominal Ordinal

Proximity Matrix: Similarity DisSimilarity

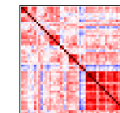
Proximity Measure: Row Processing Column Processing

Pearson's Correlation

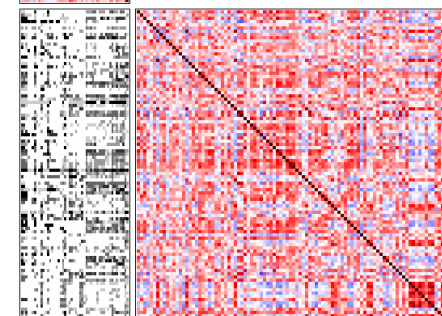
Extensions: Covariate Adjustment Nonlinear Association

OK Cancel

Column Processing

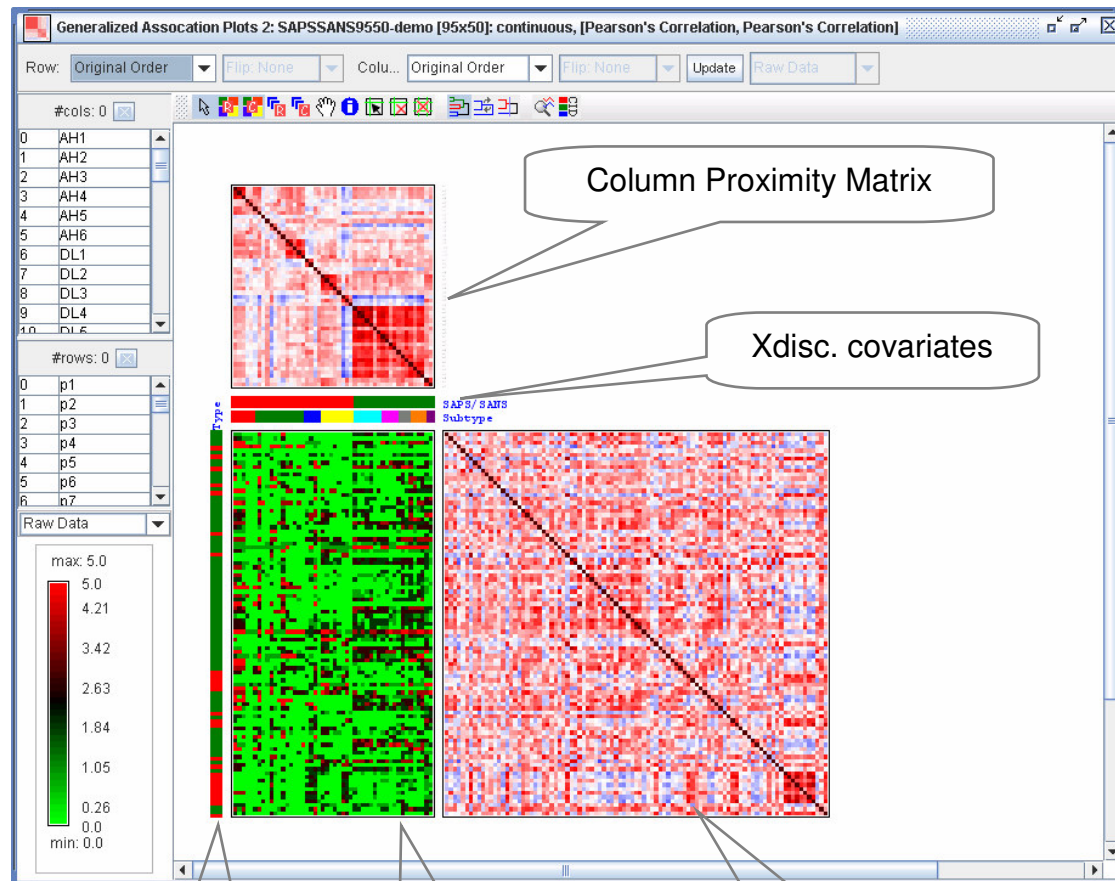


Row Processing



GAP Step 1:

Presentation of Raw Data Matrix



Column Proximity Matrix

Xdisc. covariates

Ydisc. covariates

Raw Data Matrix

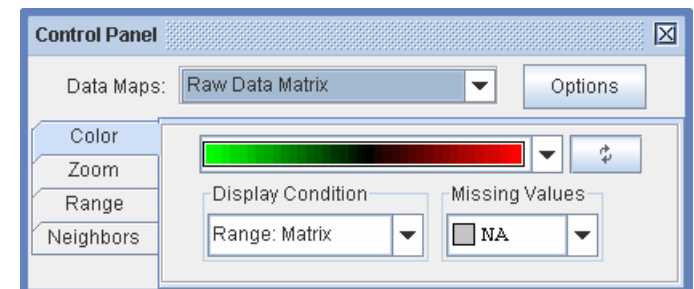
Row Proximity Matrix

Show/Hide File/Data View

Show/Hide Output Panel



Show/Hide Control Panel



Without Row/Column Processing



Generalized Association Plots

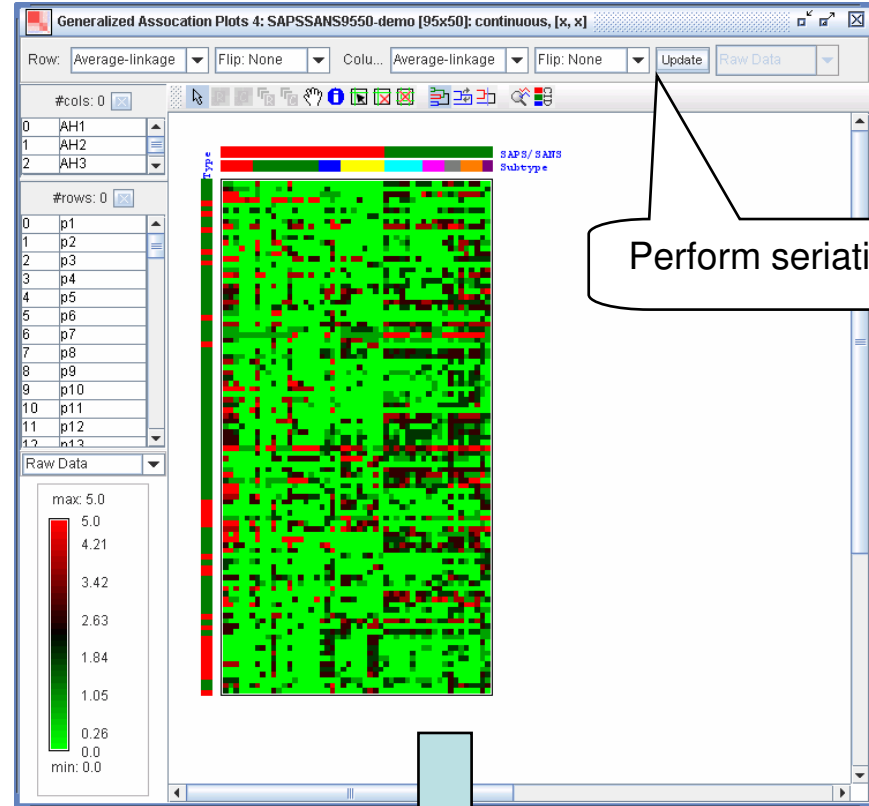
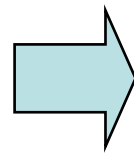
Data Sets
SAPSSANS9550-demo.bt

Data Type
 Raw Data Matrix: Continuous Binary Nominal Ordinal
 Proximity Matrix: Similarity DisSimilarity

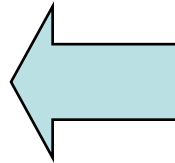
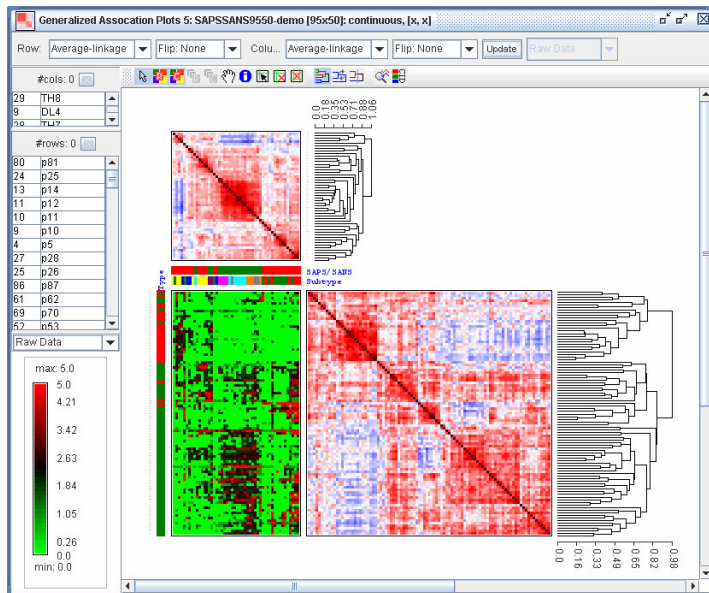
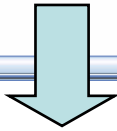
Proximity Measure
 Row Processing Column Processing
Pearson's Correlation

Extensions
 Covariate Adjustment Nonlinear Association

OK Cancel



Perform seriation



Compute Proximity








Row Processing Column Processing

Pearson's Correlation

OK Cancel

Control Panel: Color

Color Spectra

- Rainbow 130**
 Distance
- Blue-white-Red 200**
 Correlation
- Green-Black-Red 38**
 Gene Expression
- Grey 256**
 Range
- Discrete 16**

- Biallel Marker**

- ACGT**


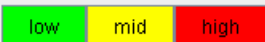
Data Maps

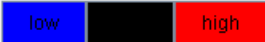
- Raw Data Matrix
- Row Proximity Matrix
- Column Proximity Matrix
- Ydisc. covaiates
- Yconti. covaiates
- Xdisc. covaiates
- Xconti. covaiates

Control Panel: Options


Color Range


User Define Color

Color Spectrum 1:  use mid

Color Spectrum 2:  use mid

Preview

Color Spectrum 1: 

Color Spectrum 2: 

Cancel

Display Conditions (only for Matrix)

- Range: Matrix
- Range: Row
- Range: Column
- Center: Matrix
- Center: Row
- Center: Column
- Std: Matrix
- Std: Row
- Std: Column
- Rank: Matrix
- Rank: Row
- Rank: Column

Control Panel

Data Maps: Raw Data Matrix

Options

Color

Zoom

Range

Neighbors

Display Condition: Range: Matrix

Missing Values: NA

Reverse

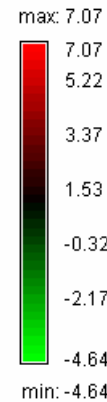
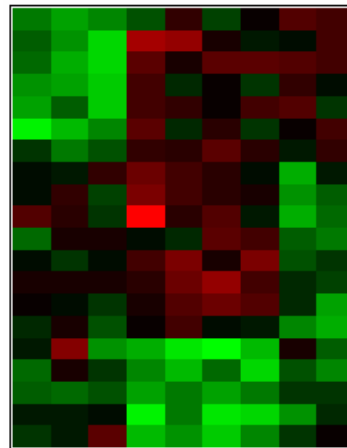
Missing Value

- NA
- NA
- NA

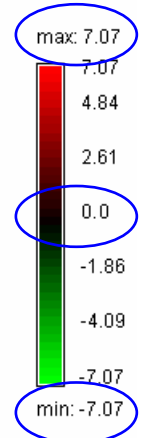
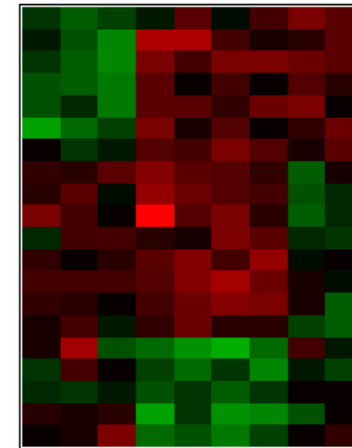
Control Panel: Color: Display Conditions



	A	B	C	D	E	F	G	H	I
1	-1.37	-2.30	-1.80	-0.55	2.45	-0.13	1.49	3.03	2.48
2	-0.68	-2.11	-3.42	4.67	4.57	1.75	0.61	0.92	2.52
3	-1.19	-2.49	-3.66	3.14	1.70	3.29	3.33	2.92	2.48
4	-1.93	-2.28	-3.16	2.51	0.32	1.49	0.21	2.20	1.03
5	-2.21	-0.79	-3.29	2.55	2.44	1.45	2.68	3.03	0.19
6	-4.14	-2.91	-1.64	3.21	0.37	1.93	0.14	1.27	2.67
7	0.21	-1.36	-0.44	2.22	1.85	3.11	2.03	0.67	2.40
8	1.13	0.79	2.25	3.65	2.52	2.09	1.13	-2.59	0.67
9	0.95	2.33	-0.07	3.89	2.72	2.13	1.75	-2.17	-0.90
10	3.04	1.85	0.21	7.07	2.01	3.05	0.76	-2.58	-1.04
11	-1.02	1.65	1.53	0.95	0.60	3.12	2.52	-0.77	-1.40
12	1.21	0.24	1.04	2.50	3.69	1.81	3.98	-0.33	0.11
13	1.74	1.60	1.70	2.02	3.45	4.46	2.69	0.41	-0.09
14	1.34	1.06	0.06	1.81	2.90	3.64	3.04	0.49	-2.33
15	0.57	1.81	-0.47	1.40	2.70	0.99	0.82	-1.61	-2.56
16	0.61	4.22	-2.03	-2.61	-4.00	-4.64	-2.92	1.55	-0.71
17	-1.13	1.64	0.01	-1.77	-2.85	-1.24	-3.41	-0.59	-1.64
18	-0.86	-1.17	-0.41	-2.20	-1.30	-2.37	-1.41	0.08	0.25
19	0.75	0.66	1.04	-4.26	-1.41	-3.99	-3.53	-2.17	0.34
20	0.15	0.68	3.18	-2.86	-2.01	-3.18	-1.58	0.10	1.28



Range Matrix Condition



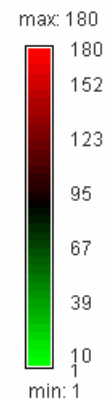
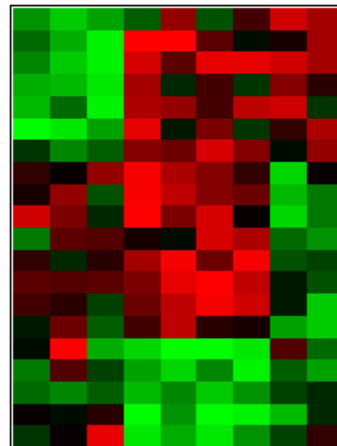
Center Matrix Condition

Bidirectional Color Spectrum

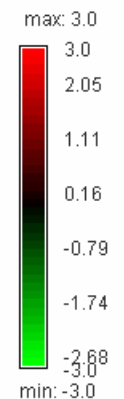
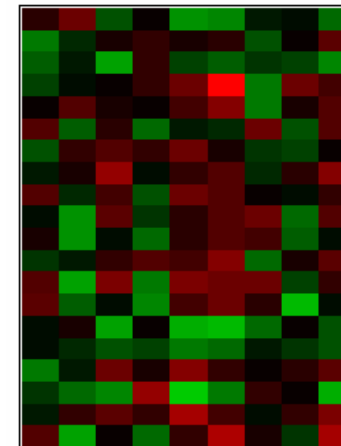
Blue-white-Red 200



Green-Black-Red 38

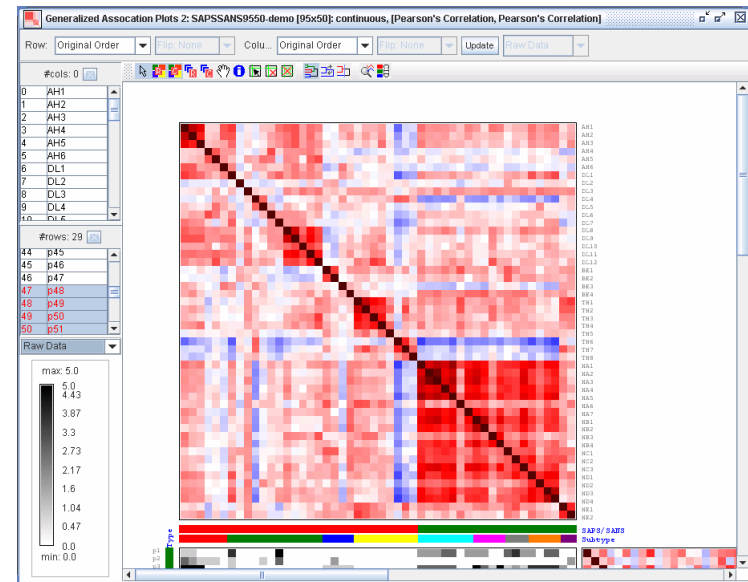
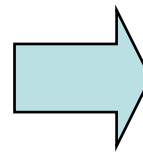
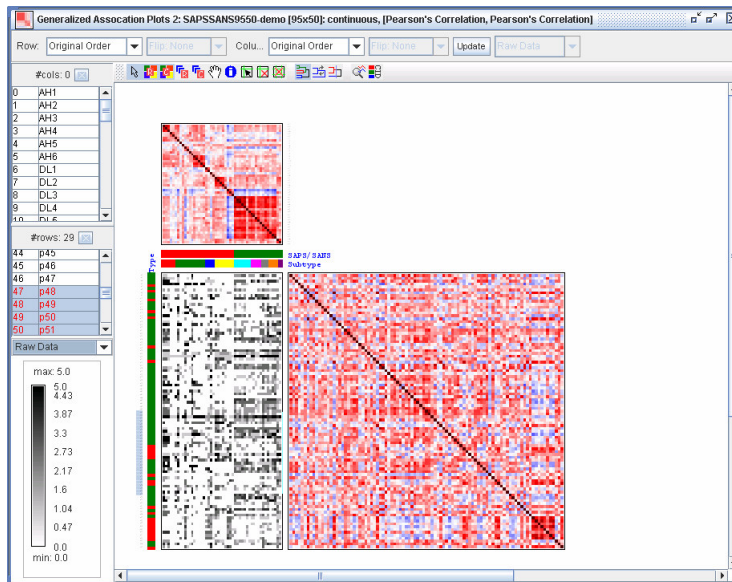


Rank Matrix Condition



Standardize Matrix Condition

Control Panel: Zoom



Control Panel

Data Maps: Options

Color				100%	
Zoom				100%	
Range				84%	
Neighbors					

Actual Size

Fit Window

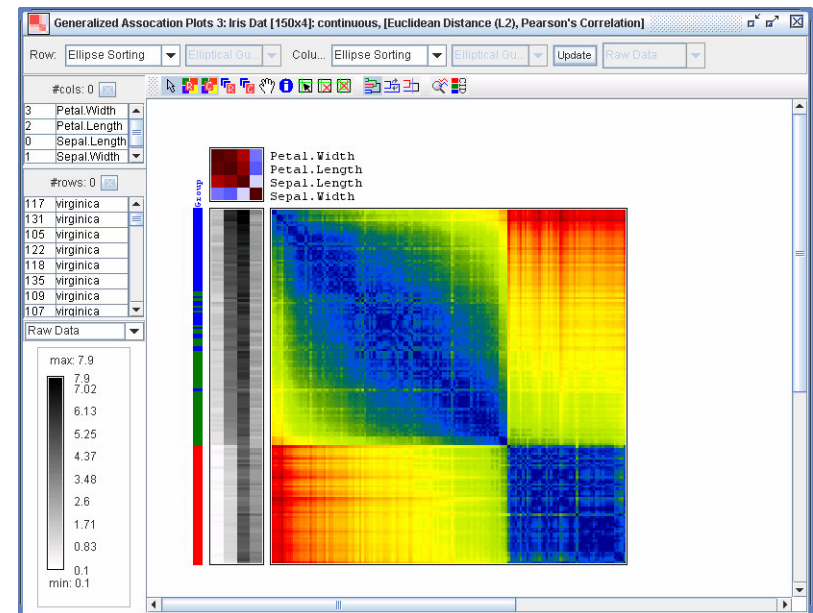
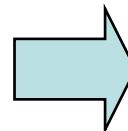
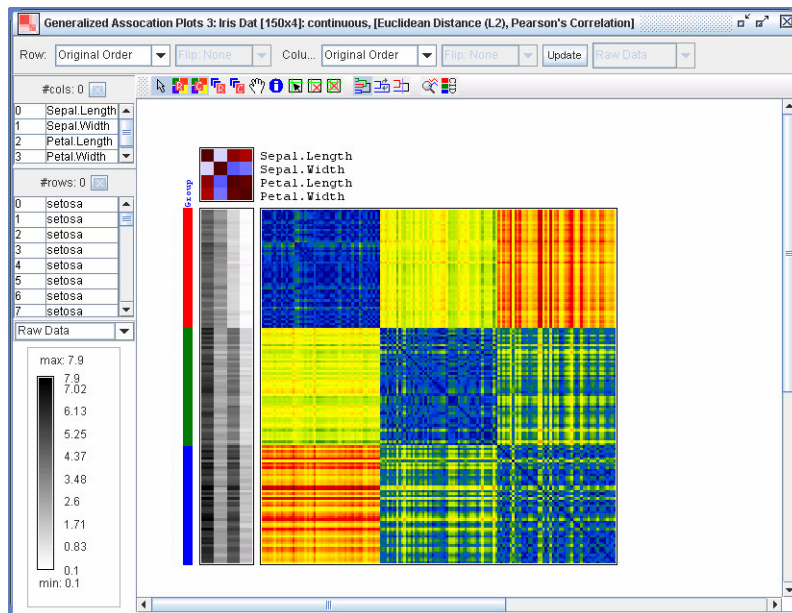
Fit Width

GAP Step 2-1: Seriation: Ellipse Sorting

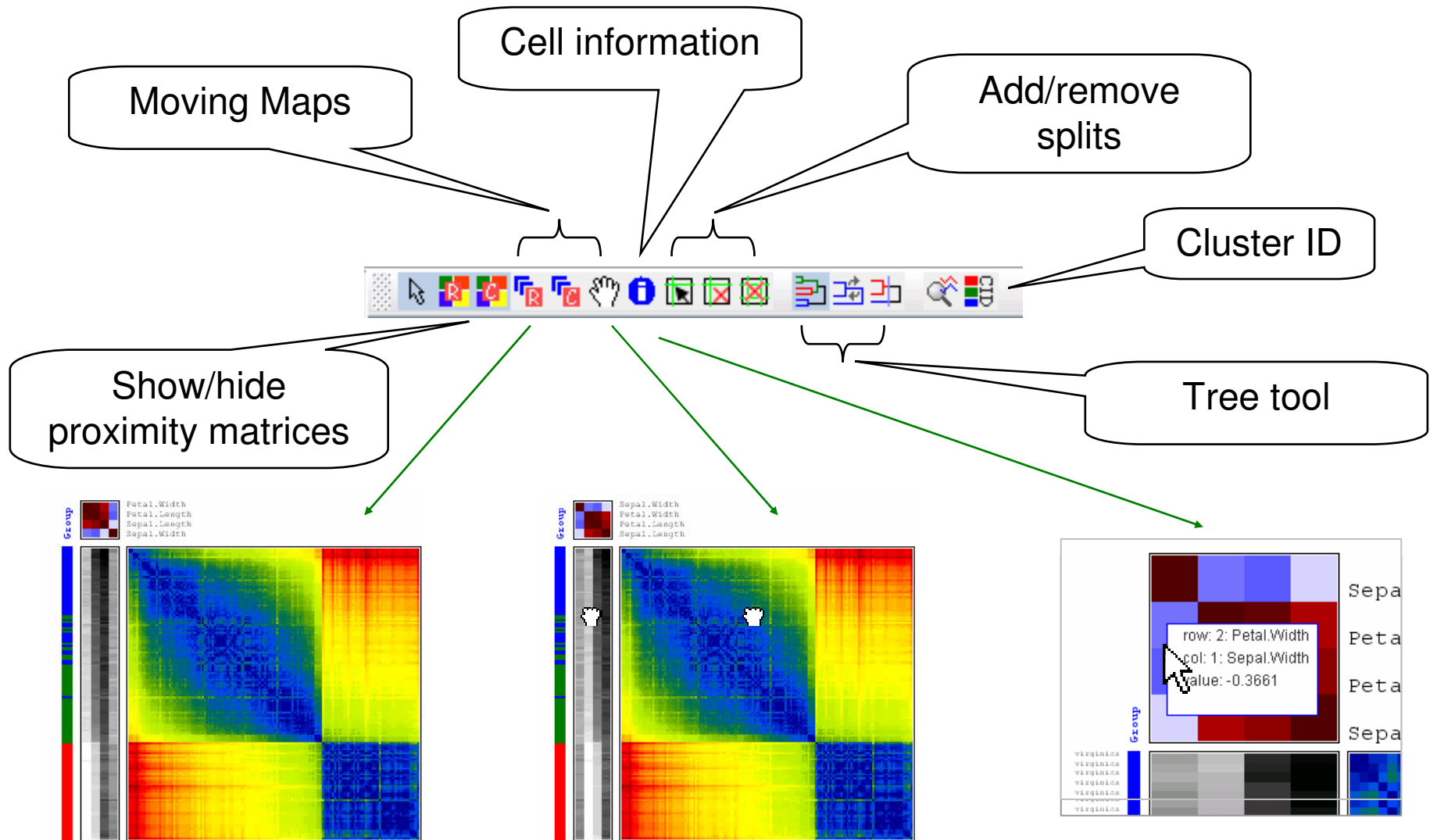
Generalized Association Plots 3: Iris Dat [150x4]: continuous, [Euclidean Distance (L2), Pearson's Correlation]

Row: Ellipse Sorting Elliptical Gu... Colu... Ellipse Sorting Elliptical Gu... Update Raw Data

- Average-linkage
- Centroid-linkage
- Complete-linkage
- Single-linkage
- Ellipse Sorting
- Original Order
- External
- Customized



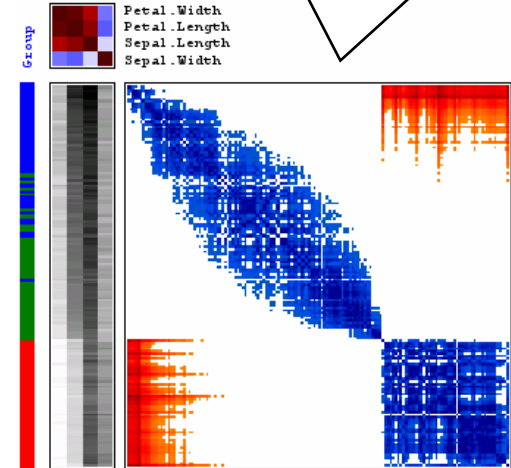
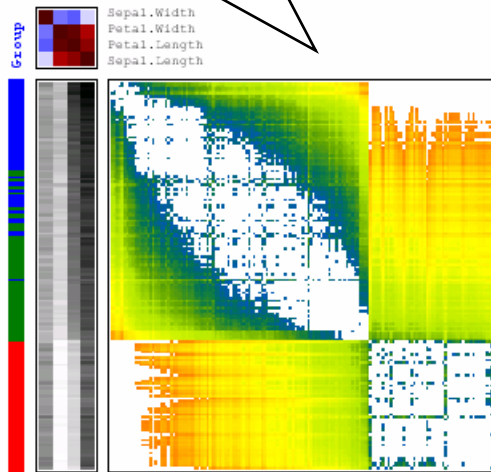
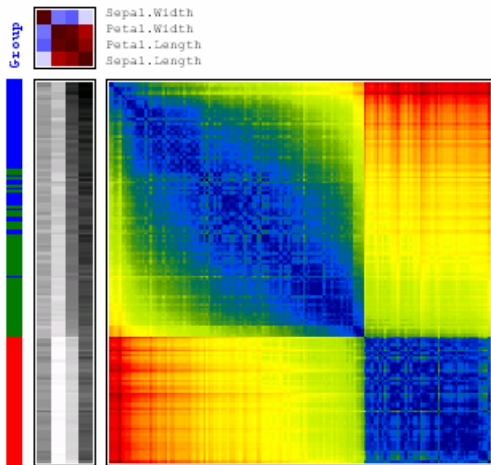
Mouse Tool



Control Panel: Range

Display within the range

Display without of the range



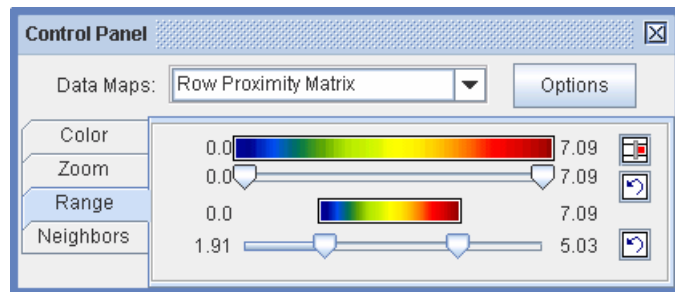
Sectional Display

The screenshot shows a 'Control Panel' window with the following settings:

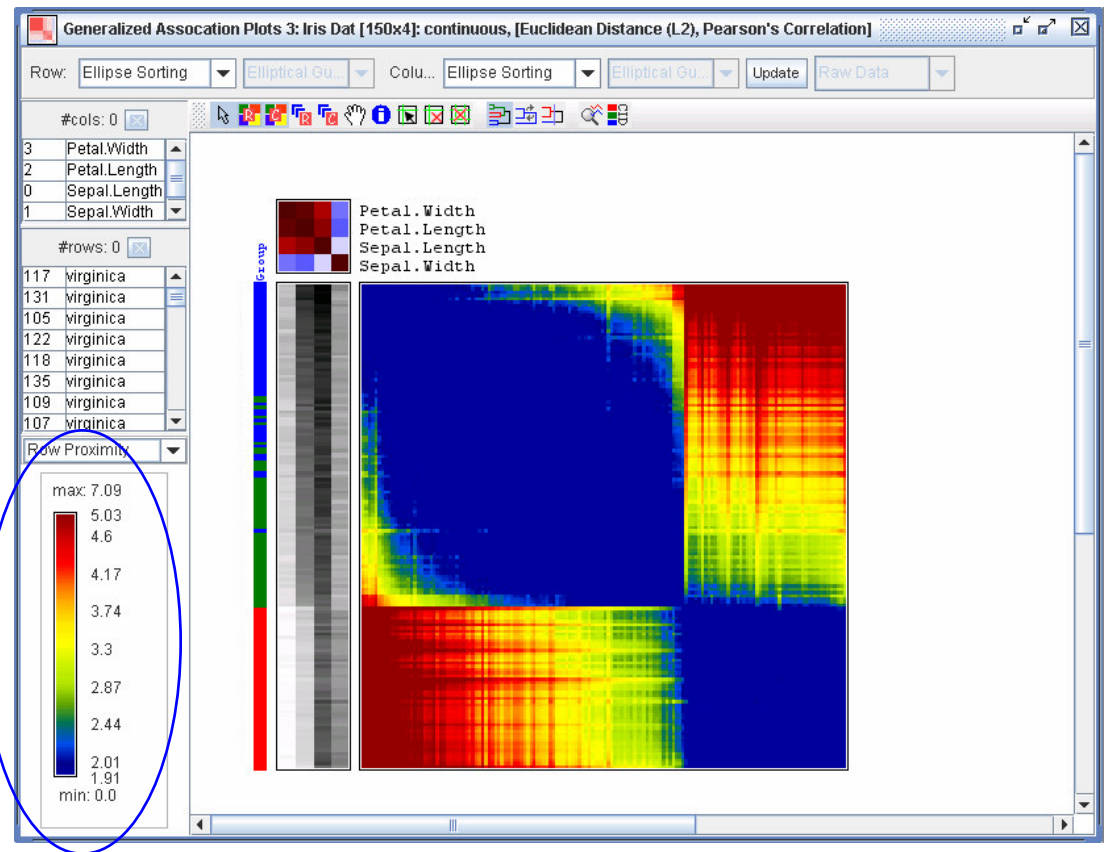
- Data Maps: Row Proximity Matrix
- Color: A color scale from 0.0 (blue) to 7.09 (red).
- Zoom: A slider set to 1.06, with a range from 1.06 to 5.03.
- Range: A color scale from 0.0 (blue) to 7.09 (red).
- Neighbors: A slider set to 'min', with a range from 'min' to 'max'.

Two green arrows point from the 'Sectional Display' text to the 'Color' and 'Range' sliders.

Control Panel: Range



Restricted Display



Control Panel: Neighbors

Control Panel [X]

Data Maps: Row Proximity Matrix [v] Options

Color

Zoom

Range

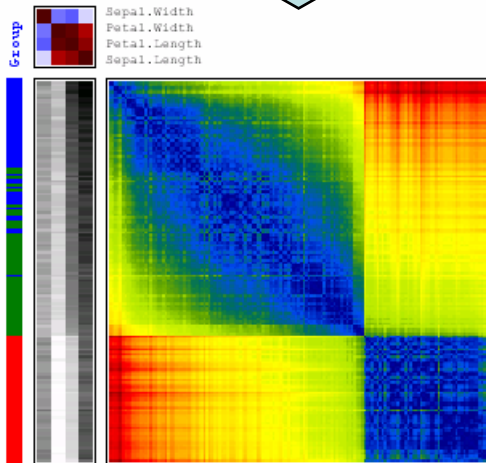
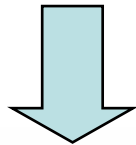
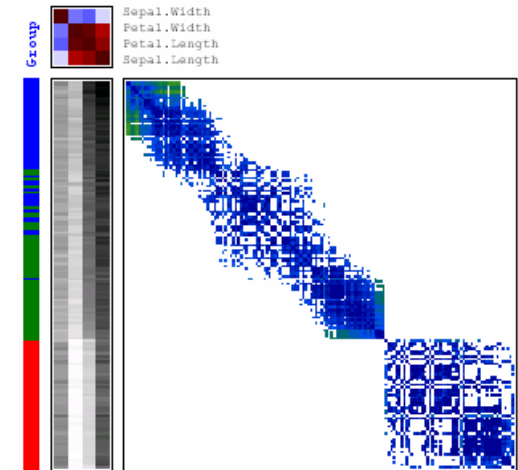
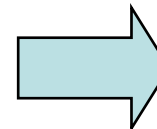
Neighbors

k nearest neighbors: 149

0 [slider] 150 [icon]

epsilon nearest neighbors: 7.085

0.0 [slider] 7.085 [icon]



Control Panel [X]

Data Maps: Row Proximity Matrix [v] Options

Color

Zoom

Range

Neighbors

k nearest neighbors: 20

0 [slider] 150 [icon]

epsilon nearest neighbors: 7.085

0.0 [slider] 7.085 [icon]

Control Panel [X]

Data Maps: Row Proximity Matrix [v] Options

Color

Zoom

Range

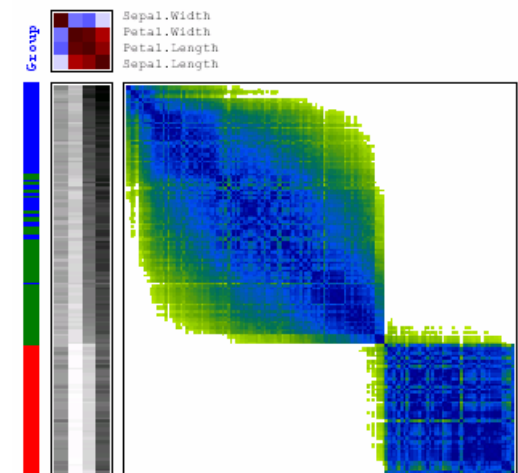
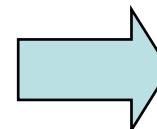
Neighbors

k nearest neighbors: 149

0 [slider] 150 [icon]

epsilon nearest neighbors: 2.48

0.0 [slider] 7.085 [icon]



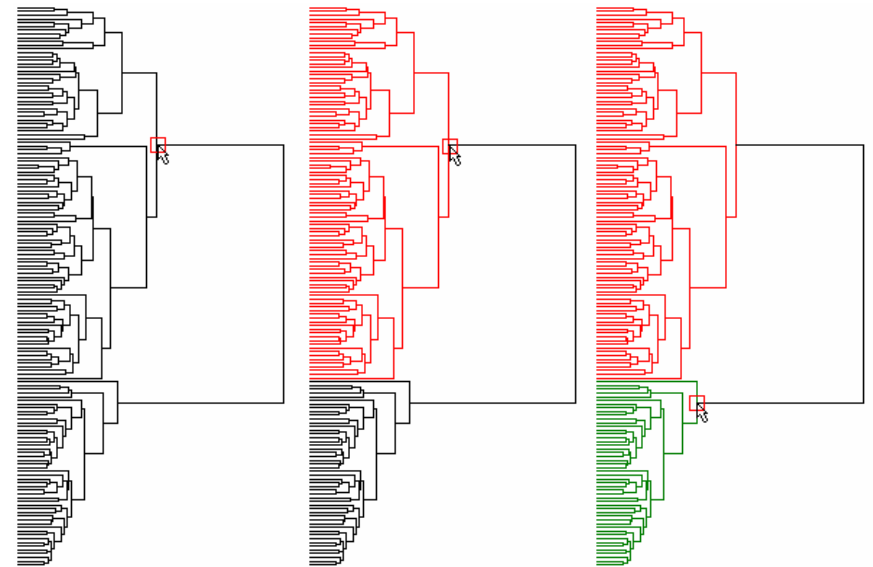
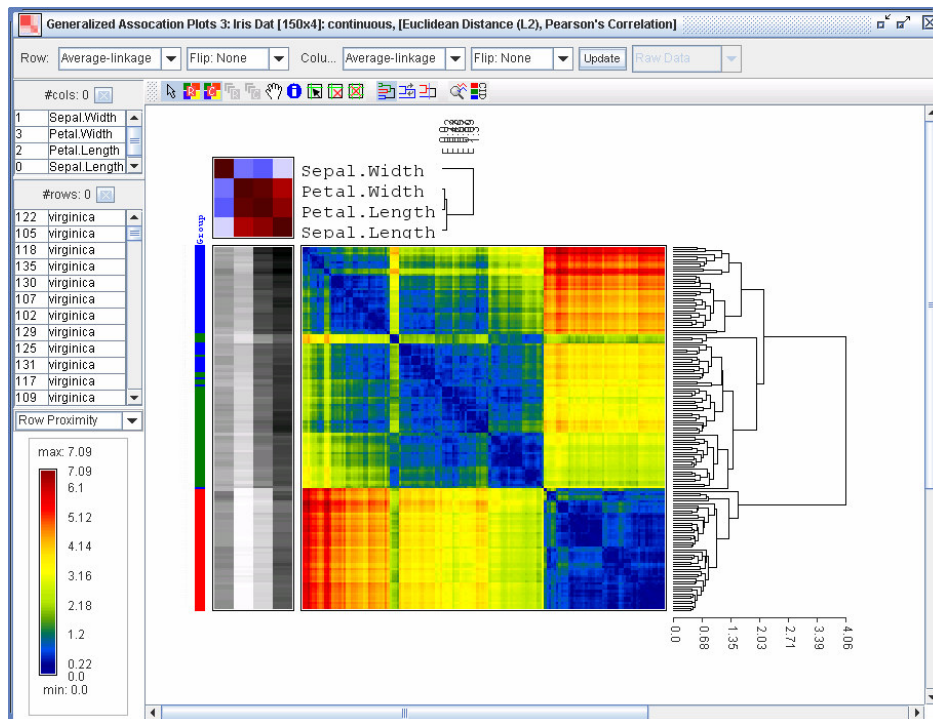
GAP Step 2-2:

Seriation: Hierarchical Clustering

- Average-linkage
- Centroid-linkage
- Complete-linkage
- Single-linkage
- Ellipse Sorting
- Original Order
- External
- Customized

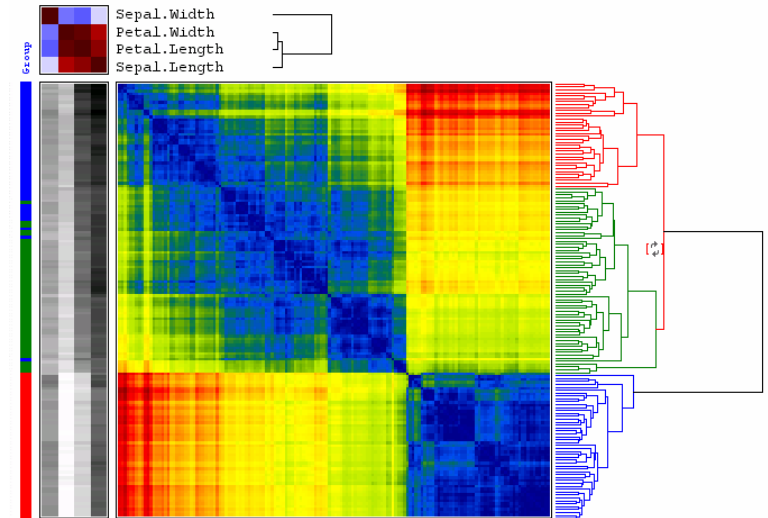
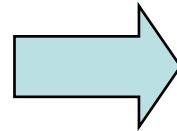
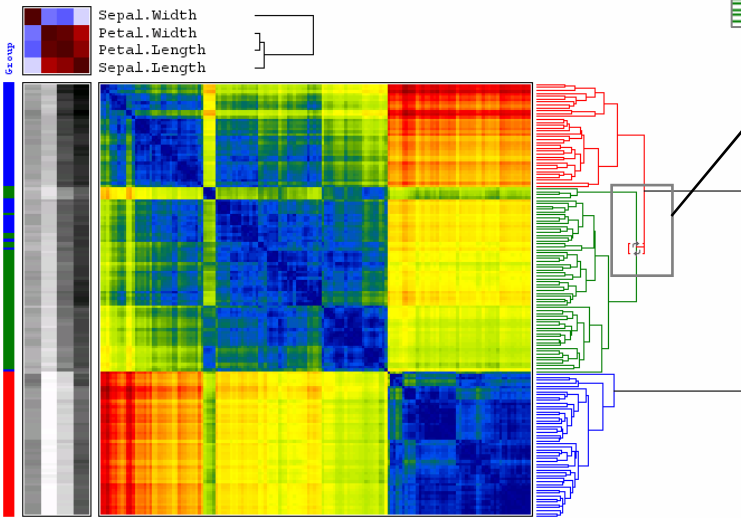
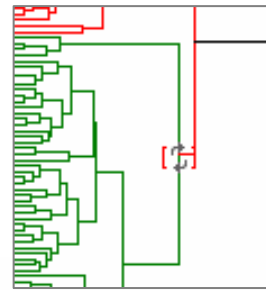


Subtree Selection

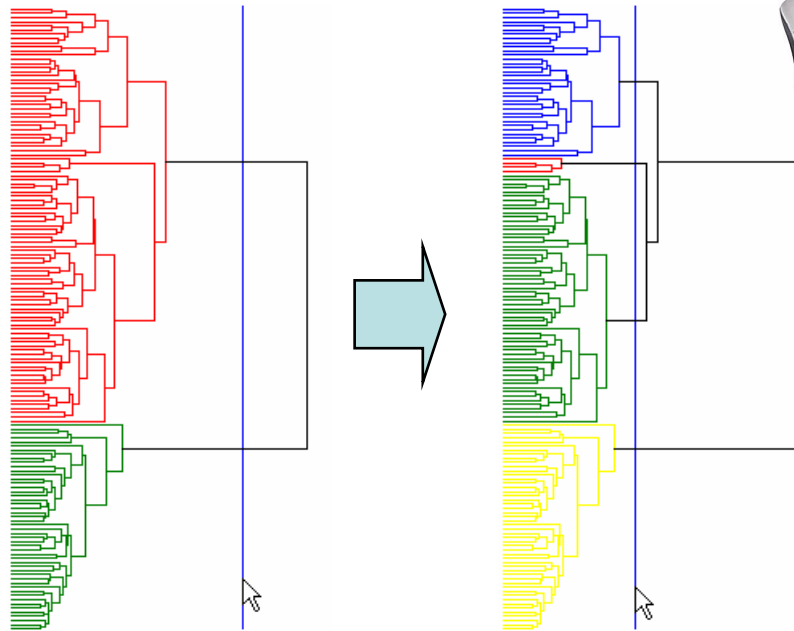


Tree Flip

- Flip: None
- Flip: Uncle
- Flip: GrandPa
- External Referenc
- Elliptical Guide

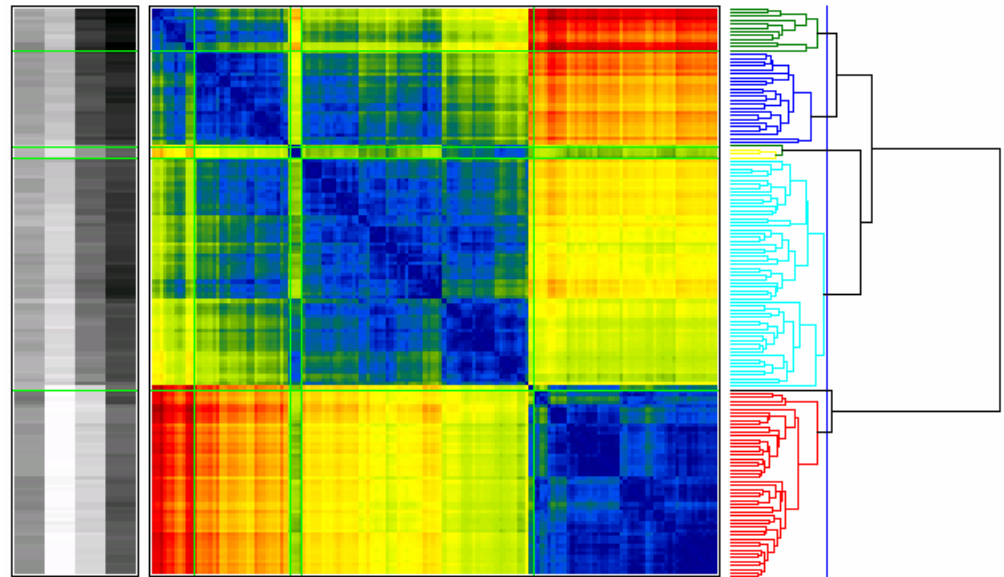


Tree Cut

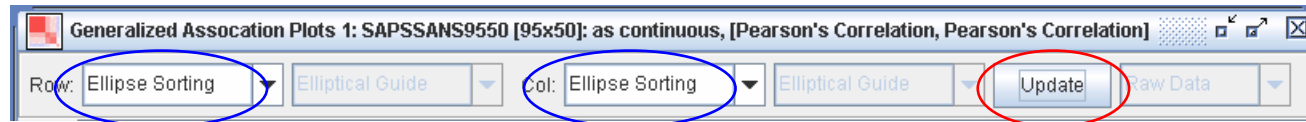


- Clear selection
- Clear cut
- Clear splits
- Random flip
- Reverse
- Get subtree
- Show splits

Make clusters



Pop-Up Menu: Sorting



Sort
Export data
Export image

Restore
Randomize
Reverse
Sort by...
Sort by this row
Sort by this col
Sedimented
Profile plot for...
Export...

Restore

Randomize

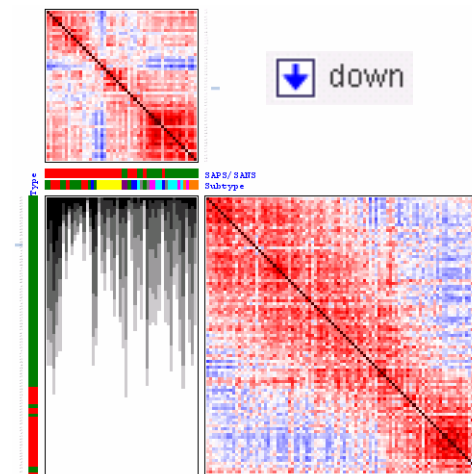
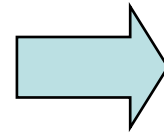
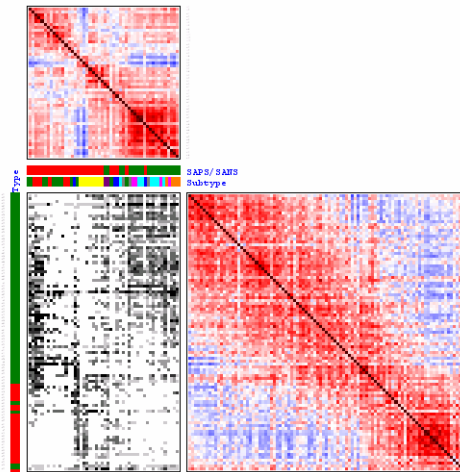
Reverse

Sort by...
Means

Pop-Up Menu: Sediment Plot

Generalized Association Plots 2: SAPSSANS9550-demo [95x50]: continuous, [Pearson's Correlation, Pearson's Correlat... □ ↗ ↘

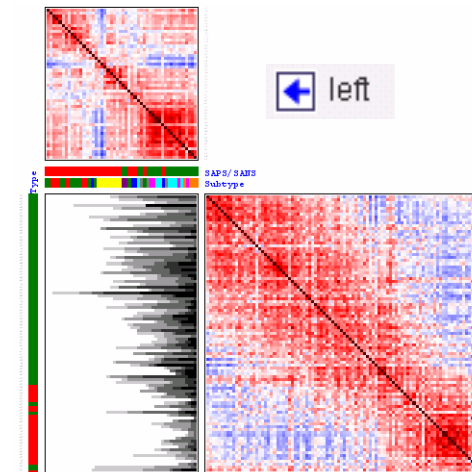
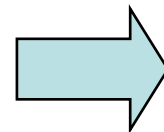
Row: Ellipse Sorting ▾ Elliptical Gu... ▾ Colu... Ellipse Sorting ▾ Elliptical Gu... **Update** Ra Data ▾



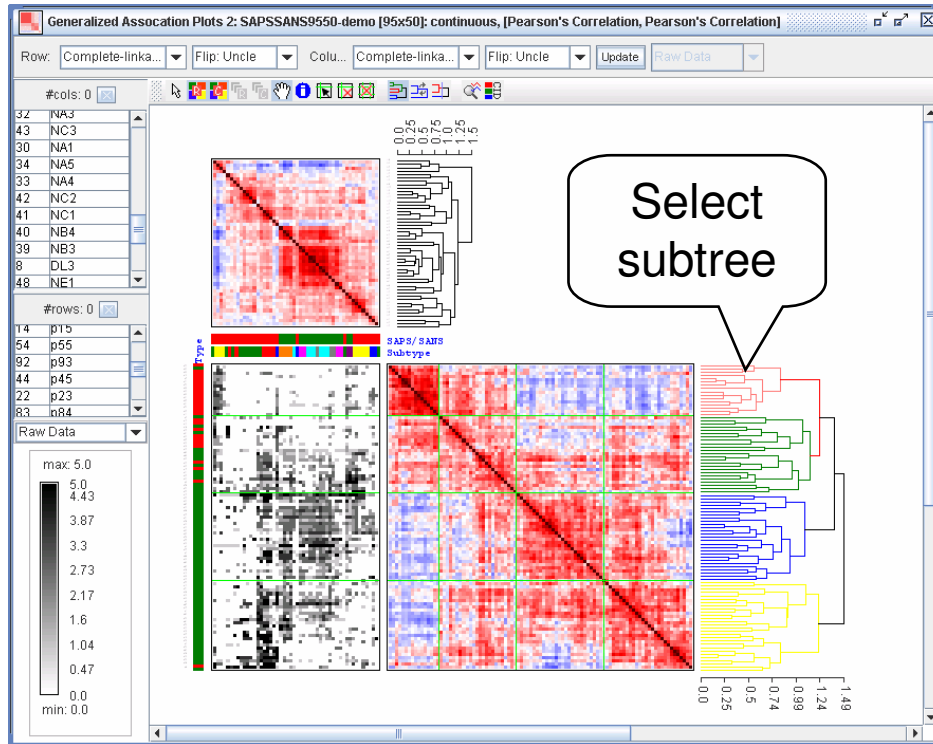
Sediment Display

- Restore
- Randomize
- Reverse
- Sort by... ▸
- Sort by this row
- Sort by this col
- Sedimented ▸
- Profile plot for... ▸
- Export... ▸

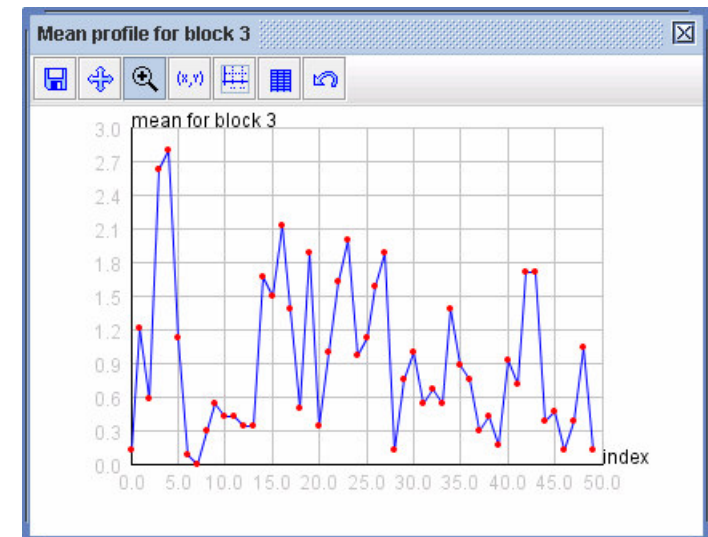
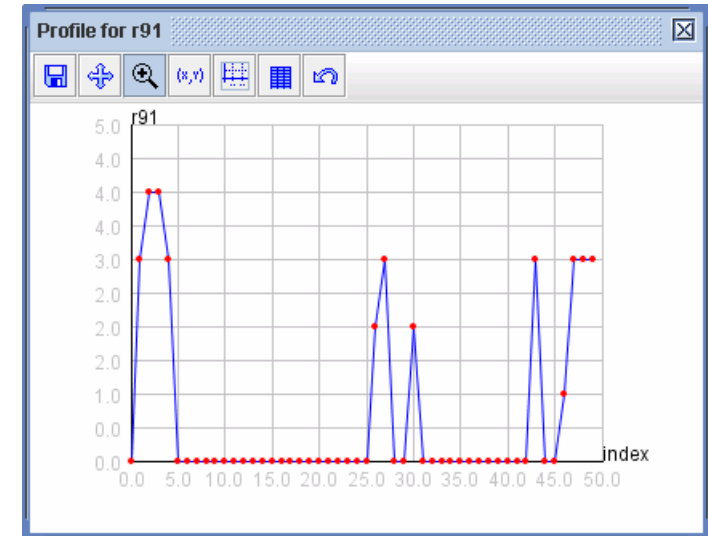
- ↑ up
- ↓ down
- ← left
- right



Pop-Up Menu: Profile Plot



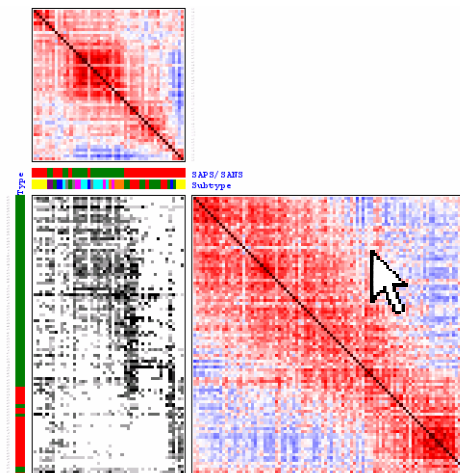
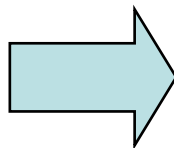
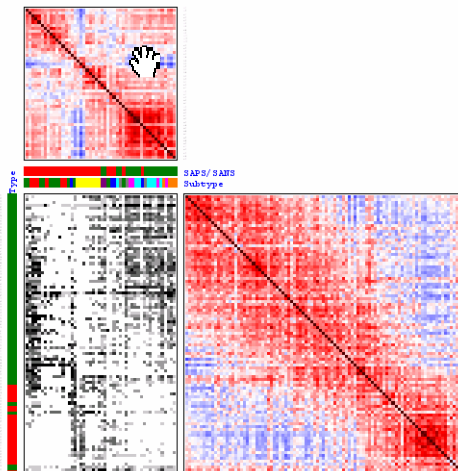
- Restore
- Randomize
- Reverse
- Sort by... ▶
- Sort by this row
- Sort by this col
- Sedimented ▶
- Profile plot for... ▶
 - this row
 - this block (mean)
 - all blocks (mean)
- Export...



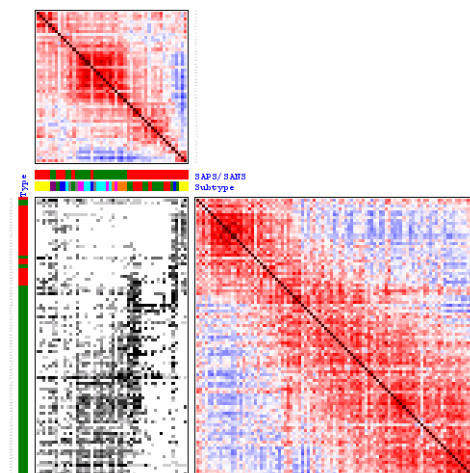
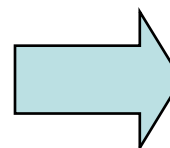
Exercise

Generalized Association Plots 2: SAPSSANS9550-demo [95x50]: continuous, [Pearson's Correlation, Pearson's Correlat...

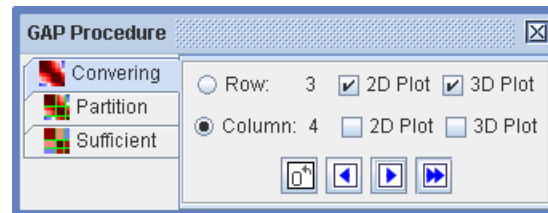
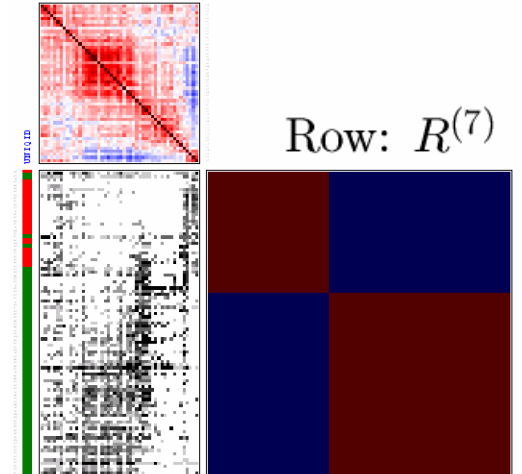
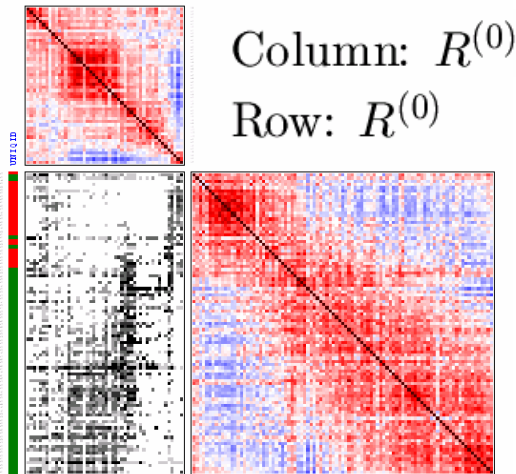
Row: Colu...



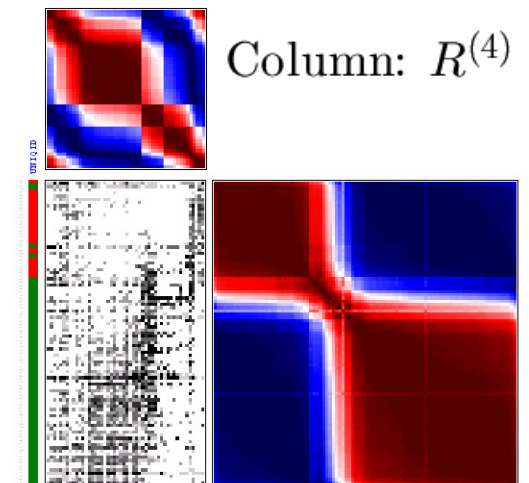
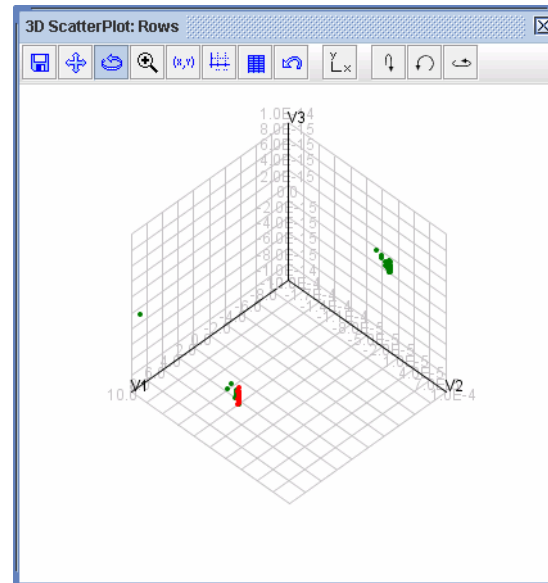
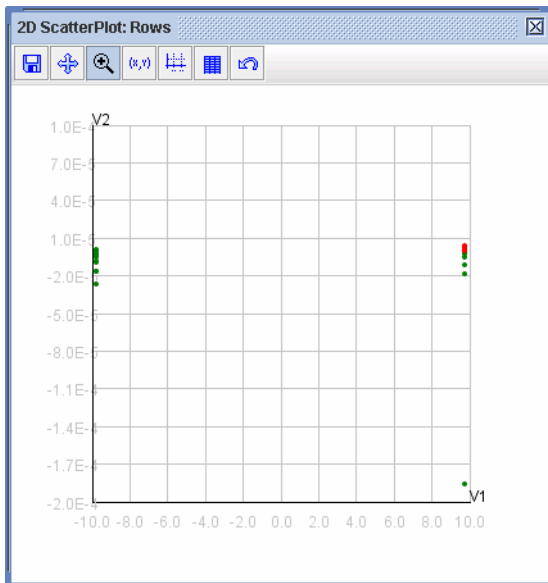
Reverse



Converging Correlation Matrices



Row: $R^{(7)}$



GAP Step 3:

Partitions of Permuted Matrix Maps

GAP Procedure: Finding Blocks

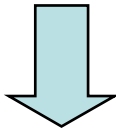
Converging Row Column

Partition Simple Simple

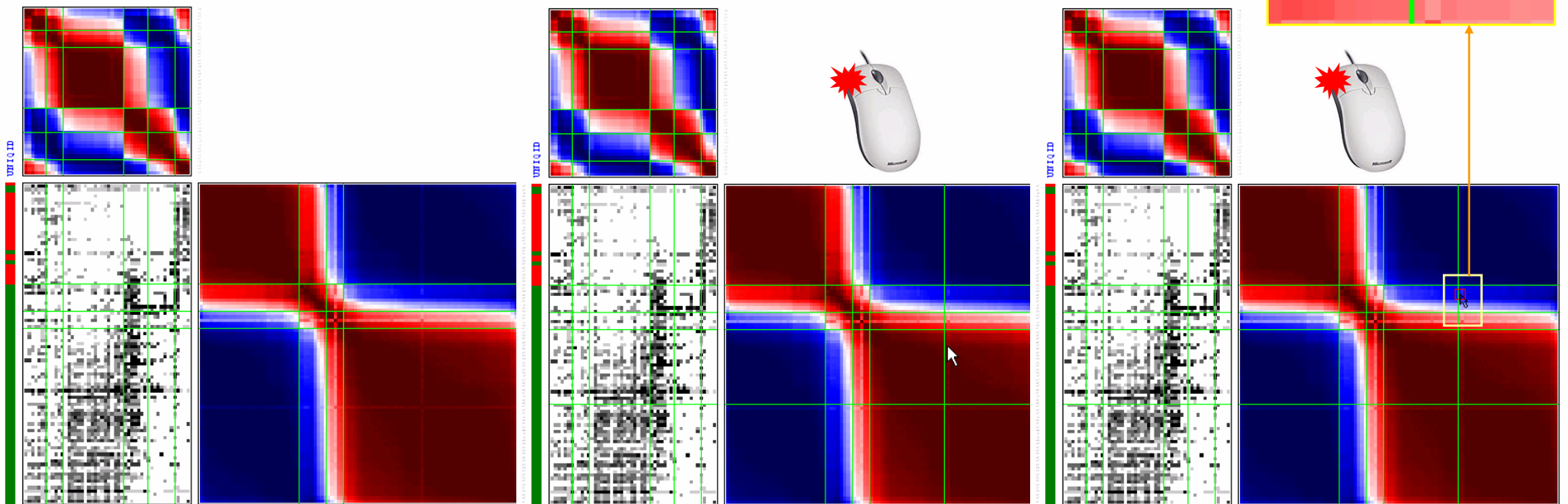
Sufficient

Update Clear

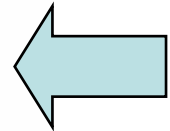
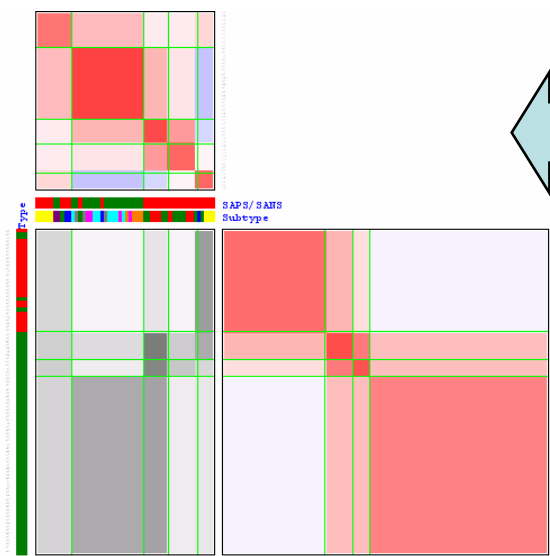
Simple
Dendrogram
max(BSSWSS)



add/remove splits



GAP Step 4: Sufficient Graph



GAP Procedure: Sufficient Display

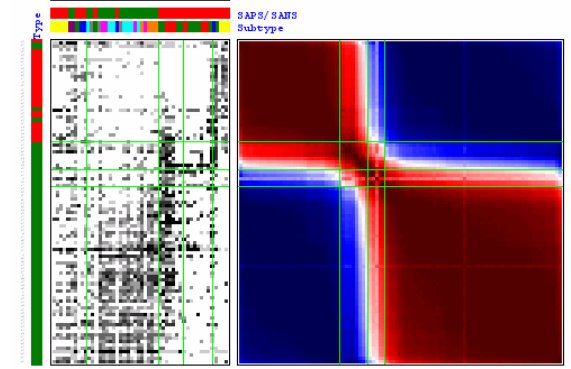
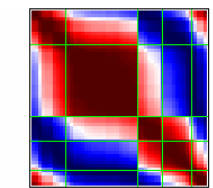
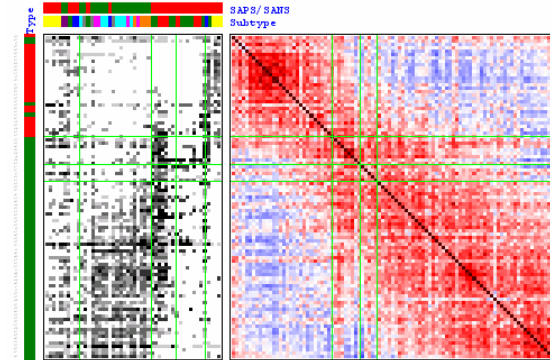
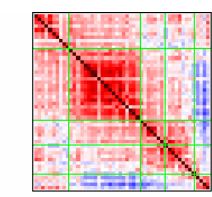
Converging Matrix Row Column

Partition Mean Mean Mean

Sufficient

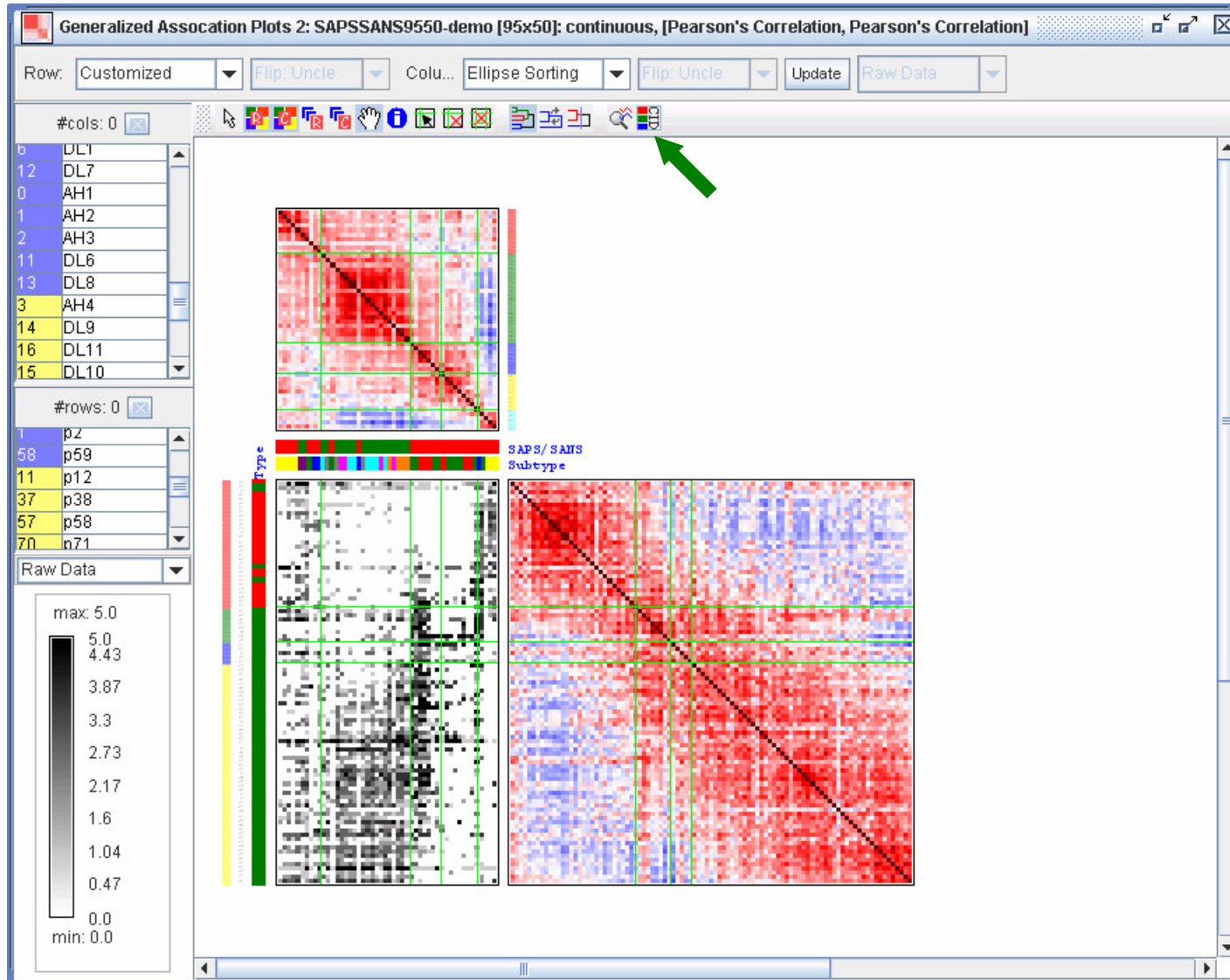
Update

- Mean
- Median
- Std.
- CV



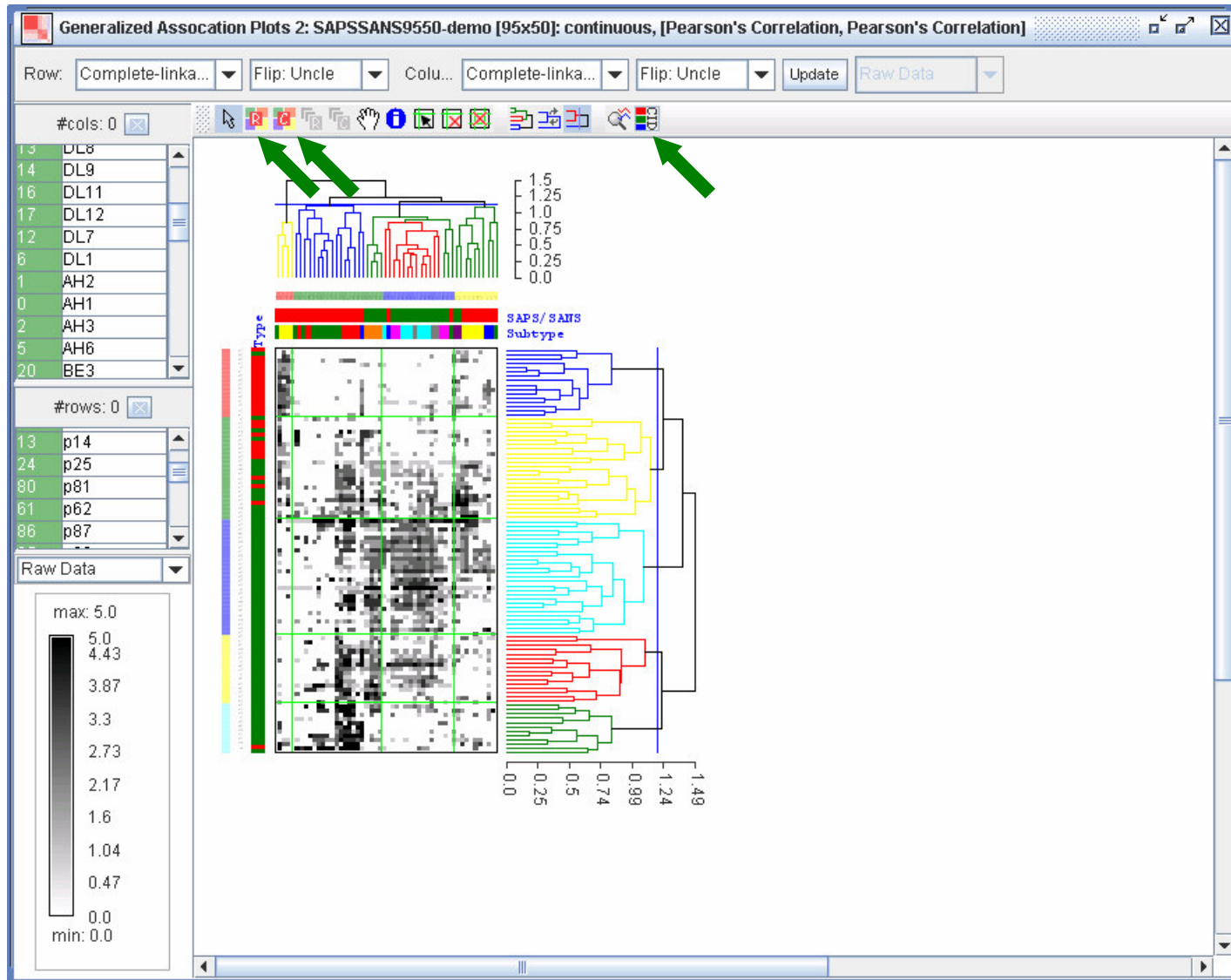
Make Cluster ID

30/37

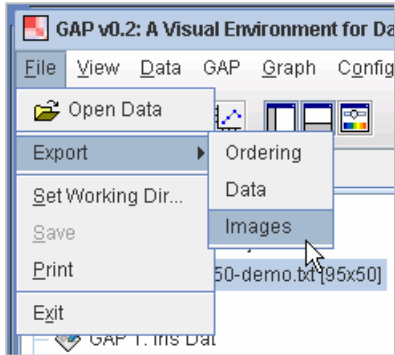


Make Cluster ID

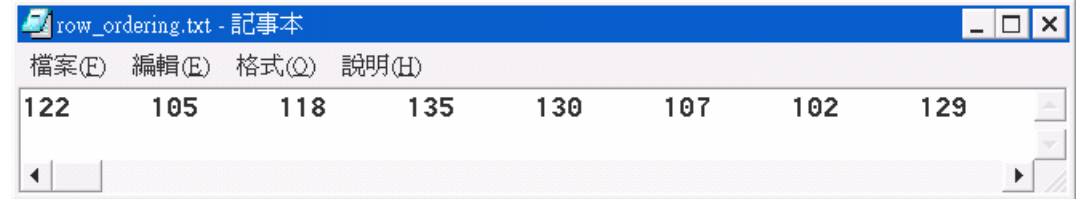
31/37



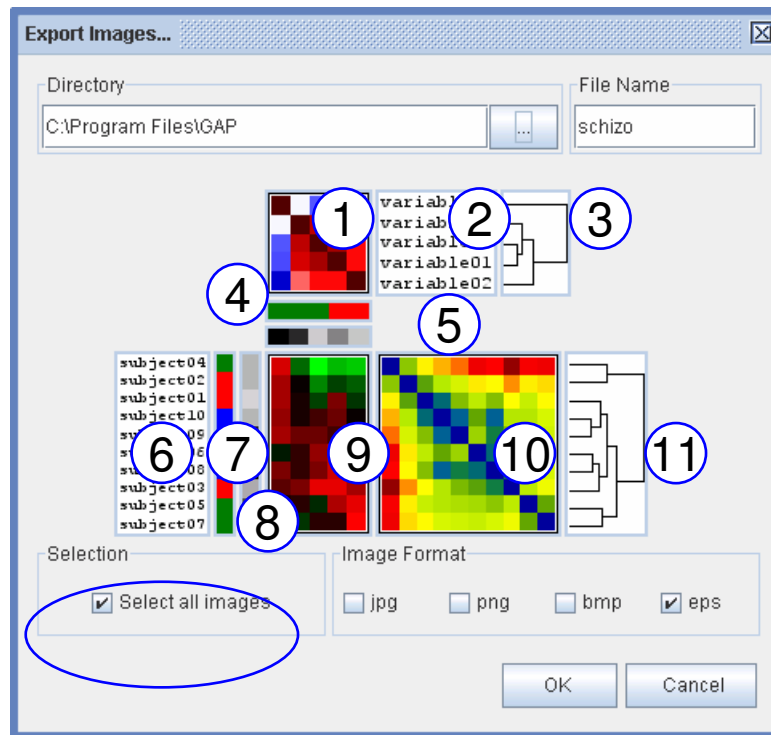
Export Ordering, Data, Images



row_ordering.txt
col_ordering.txt



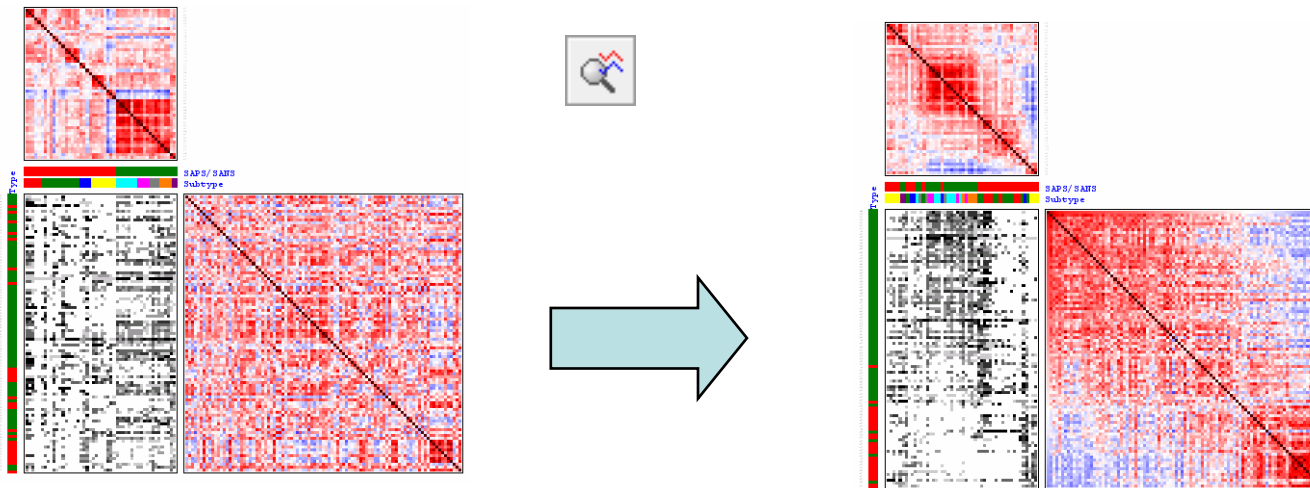
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	LIN01D	TH3	TH4	TH5	TH1	NE2	NE1	DL2	BE1	BE2	NA6	NC1	DL3	NC2	NE2	NE1	NA
2	162	1	1	5	1	5	5	5	0	2	3	1	1	4	4	2	
3	227	2	2	1	2	1	1	2	0	0	2	0	1	0	2	2	
4	152	0	0	2	0	2	0	3	0	0	0	0	0	2	0	0	
5	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
6	189	0	0	1	2	2	2	1	0	0	0	0	0	0	0	0	
7	184	0	0	1	0	2	1	1	0	0	2	1	1	1	1	1	
8	188	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	
9	160	2	0	3	1	3	3	3	0	2	0	0	0	0	3	0	
10	130	0	0	3	0	2	0	2	0	0	0	0	0	0	0	0	
11	187	0	0	0	0	0	1	1	0	2	3	2	2	0	1	0	
12	191	0	0	0	0	2	3	3	3	3	0	2	0	3	0	0	
13	186	0	0	0	0	0	0	3	0	2	3	0	0	0	0	0	
14	115	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
16	172	0	0	0	0	0	1	3	0	1	0	1	0	1	0	2	
17	190	0	0	0	0	2	1	1	0	2	2	1	0	0	0	0	
18	186	0	0	0	0	0	1	3	0	2	0	0	0	2	1	1	
19	171	0	0	0	0	0	3	4	0	0	3	0	0	0	0	0	
20	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
21	194	0	0	0	2	0	0	3	0	0	0	0	0	0	3	0	
22	151	5	5	5	0	5	0	2	0	0	2	2	0	3	2	2	
23	181	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
24	182	0	2	3	0	3	2	3	0	0	0	0	0	0	0	0	
25	124	0	0	0	0	3	2	0	0	2	0	1	0	0	0	0	
26	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
27	180	0	0	0	0	2	0	3	0	1	0	0	0	2	0	0	
28	113	0	0	0	0	0	3	3	0	0	0	0	0	1	0	0	
29	189	3	0	3	0	4	2	3	0	0	2	2	0	4	0	0	
30	183	0	3	4	4	0	0	1	0	2	2	0	0	0	0	0	
31	189	0	0	3	0	2	0	0	3	3	2	1	0	2	0	0	
32	129	4	4	3	1	4	1	1	0	0	1	1	0	0	1	1	
33	119	0	0	2	0	2	2	0	0	2	0	0	3	0	2	0	
34	164	3	4	4	2	5	5	4	0	2	0	4	2	0	4	0	
35	142	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
36	175	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	
37	133	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	
38	13	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	
39	173	2	3	2	0	3	2	2	0	2	2	0	0	4	0	0	
40	136	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	
41	110	0	0	0	0	0	0	2	0	0	0	0	0	0	3	0	
42	120	3	4	3	0	4	5	4	0	2	4	0	3	0	3	0	
43	11	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	
44	156	0	2	1	2	2	1	2	0	0	0	0	0	0	3	1	
45	111	0	0	0	0	0	1	1	0	1	3	0	3	0	2	0	
46	137	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	
47	157	0	3	0	0	0	1	1	0	0	2	0	0	0	0	3	
48	170	0	0	0	0	0	2	2	0	0	0	2	0	3	0	0	
49	140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	



- 1 _gapCol.eps
- 2 _gapColName.eps
- 3 _TreeCol.eps
- 4 _gapXd.eps
- 5 _gapXc.eps
- 6 _gapRowName.eps
- 7 _gapYd.eps
- 8 _gapYc.eps
- 9 _gapMatrix.eps
- 10 _gapRow.eps
- 11 _TreeRow.eps

Find Similar Patterns

33/37



Finding Similar Patterns Dialog

Option

Subject: 2:p3

Pattern: .0 4.0 4.0 4.0 4.0 2.0 0.0

Data Type: Continuous

Similarity: Pearson's Correlation

Preview

Compute Close

Finding Similar Patterns Dialog

Option

Subject: 2:p3

Pattern: 5 5 4 0 0 0 5 5 0 0 0 0 4 4

Data Type: Continuous

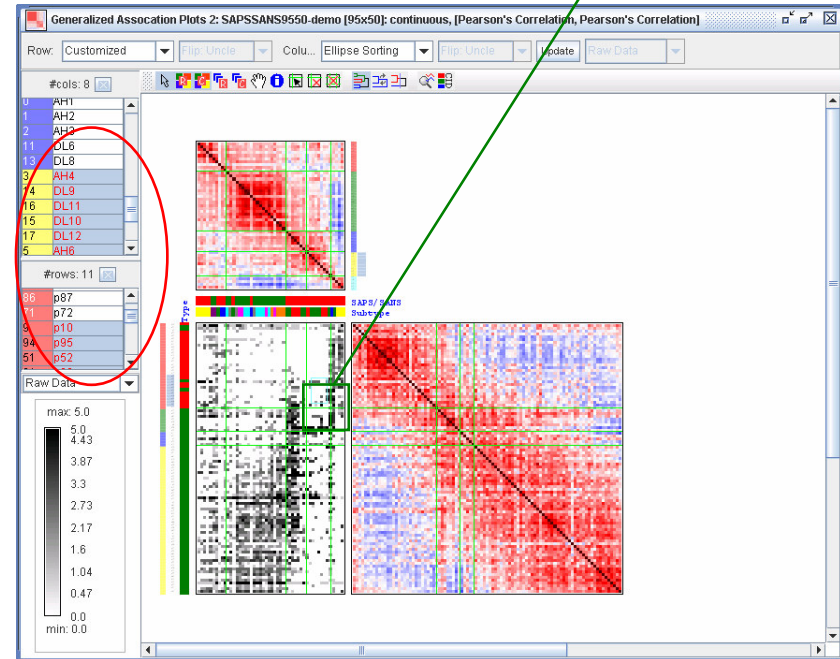
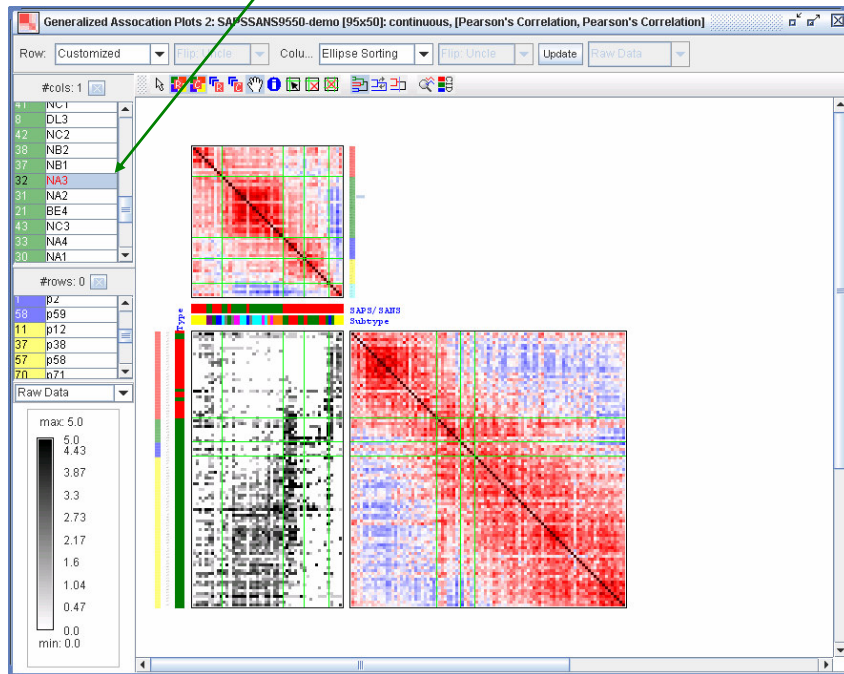
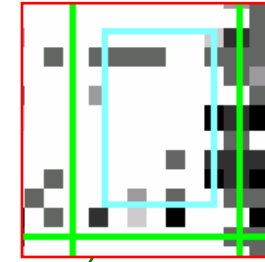
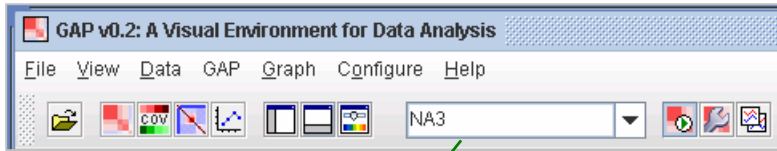
Similarity: Pearson's Correlation

- Euclidean Distance
- Kendall's tau
- Pearson's Correlation
- Spearman's Rank

Preview

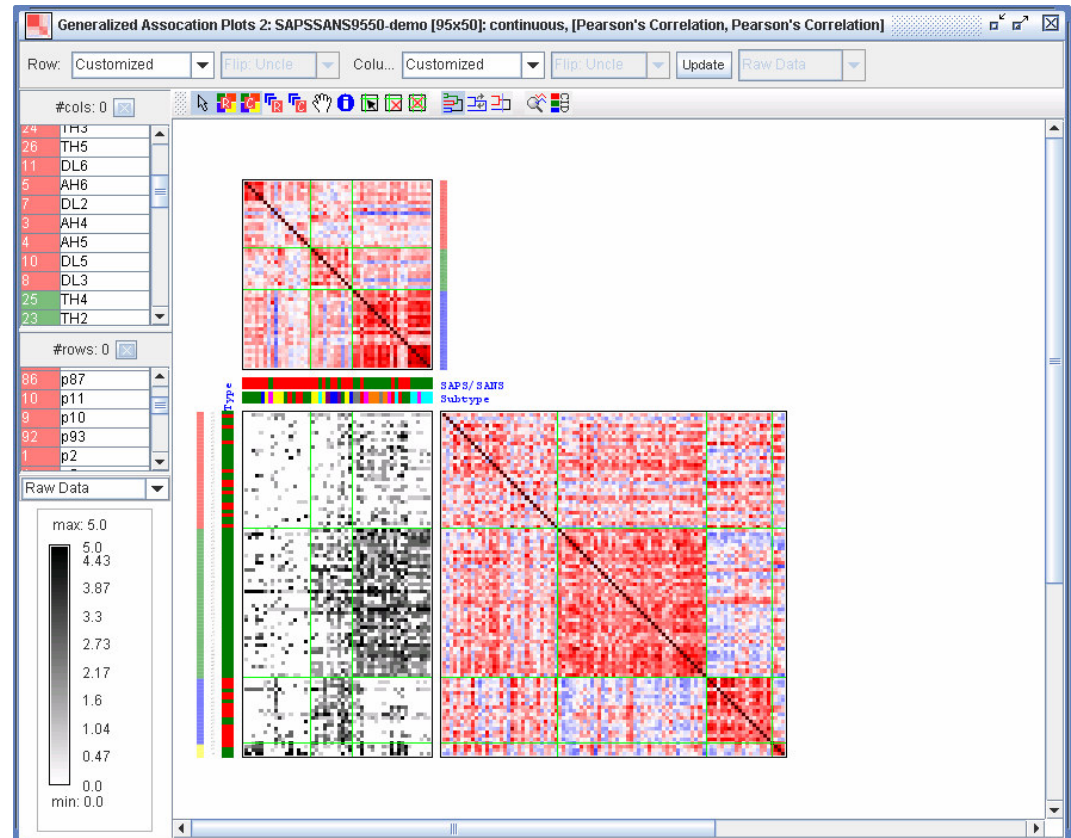
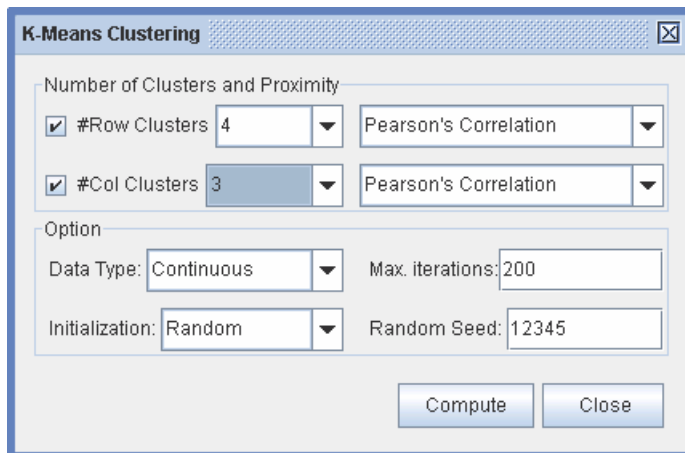
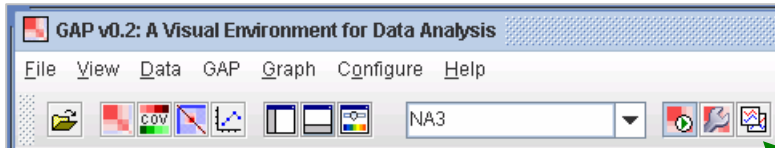
Compute Close

Find... and Select

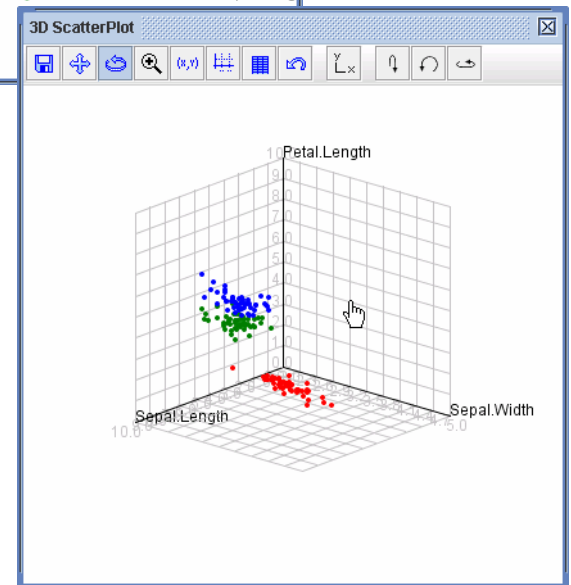
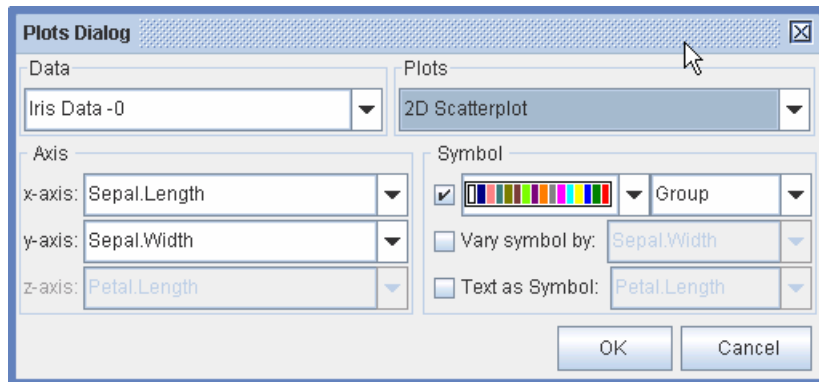
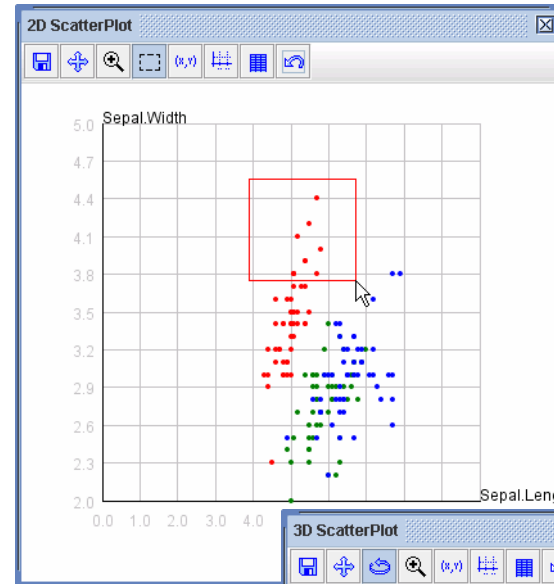
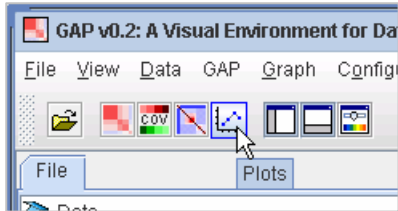


K-means Clustering

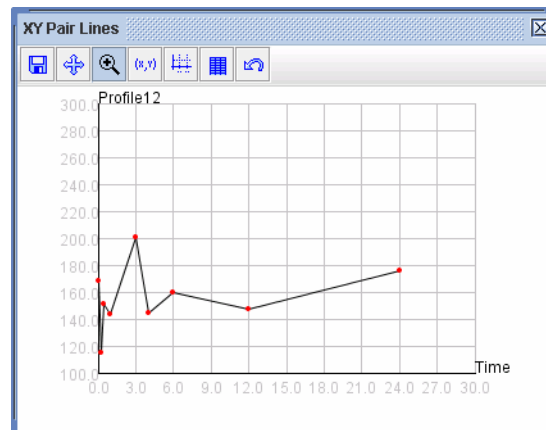
35/37



Plots



- 2D Scatterplot
- 3D Scatterplot
- XY PairLines Plot
- 2D Boxplot (X)
- 3D Boxplot (X)
- Histogram (X)



Web Site



Lab for Information Visualization

Dimension Free Data Visualization
Lab for Information Visualization
資訊視覺化研究室
Institute of Statistical Science, Academia Sinica

Home | Research | Members | Database

Chun-hou Chen
陳君厚
Associate Research Fellow
Institute of Statistical Science
Academia Sinica

- Information Visualization**
 - Generalized Association Plots (GAP)
 - Sliced Inverse Regression (SIR)
 - Multidimensional Scaling (MDS)
- Psychiatry Research**
 - Psychiatry
- Bioinformatics**
 - Microarray Data Analysis
 - SNPs
- Talks/Seminar**
 - Lecture Notes
 - Posters
- News/Conference** [more]

2006: BIBE | CAMDA | CSB | GIW | IMS/ENAR | ISMB/ECCB | InfoVis | JSM | ...

Handbook of Computational Statistics (Volume III): Data Visualization
Chun-hou Chen, Wolfgang Härdle, and Antony Unwin (eds)
Springer-Verlag, Heidelberg

<http://gap.stat.sinica.edu.tw>

GAP Software

GAP: Generalized Association Plots - Microsoft Internet Explorer

GAP is a java-designed software for generalized association plots (Chen, 2006) in a java runtime environment 1.5 (JRE version 1.5.0_04), which is available for m...

Authors:
Han-Ming Wu and Chun-hou Chen at
[Lab for Information Visualization](http://gap.stat.sinica.edu.tw)
Institute of Statistical Science,
Academia Sinica
Taipei, Taiwan, R. O. C.

Official Website of GAP Software:
<http://gap.stat.sinica.edu.tw/Software/GAP>

Current Version:
v0.2, Build 20061210

Contact:
hmwu@stat.sinica.edu.tw

Features:

- Clustering Analysis: hierarchical clustering, ...
- Various Display Conditions
- Histogram, 2D Scatterplot, 3D Scatterplot (Rotatable)
- Latest Features**

GAP Forum Generalized Association Plots (GAP) Forum.

- Please register in the GAP Forum, so

<http://gap.stat.sinica.edu.tw/Software/GAP>

GAP Forum

Generalized Association Plots (GAP) Forum - Microsoft Internet Explorer

Generalized Association Plots (GAP) Forum
GAP: a graphical environment for matrix visualization and information mining

Index | User list | Register | Login

You are not logged in.

Announcement

GAP is a Java-designed software for general purposes matrix visualization (MV) and exploratory data analysis (EDA). It provides powerful visual aids for rapid and dynamic exploration of a data matrix and the corresponding proximity matrices for variables and subjects interactively. Various clustering methods with internal/external permutation algorithms are developed. GAP also offers other unique functionalities with friendly graphical user interface. Please visit us at [Lab for Information Visualization](http://gap.stat.sinica.edu.tw)

Forum	Topics	Posts	Last post
News & updates Announcements, news and updates about GAP. (Moderated by hmwu)	1	1	2006-02-05 17:02:23 by hmwu
General Questions about using GAP (Moderated by hmwu)	0	0	
Suggestions & Feedback Give us your thoughts and comments to improve GAP. (Moderated by hmwu)	0	0	
Bug Report Please report bugs here! (Moderated by hmwu)	0	0	

Newest registered user: [hoh1](#)
Registered users online: 0
Guests online: 1

Total number of registered users: 2
Total number of topics: 1
Total number of posts: 1

Powered by [EunBR](#)
© Copyright 2002-2005 Rickard Andersson

<http://gap.stat.sinica.edu.tw/GAPforums>