Some New and Easy Ways to Describe, Compare, and Evaluate Products and Assessors.

With a Dash (?) of DISTATIS

Hervé Abdi

www.utd.edu/~herve

SPISE 2007

Summer Program In Sensory Evaluation

26-27, JULY, HOCHIMINH CITY, VIETNAM

THANKS !

Colleagues, friends, accomplices, current and former students

- Dominique Valentin,
- Sylvie Chollet
- Krystel Chrea
- Betty Edelman
- Joseph Dunlop
- Nils Pénard

- Abdi H., Valentin, D., Chollet, V., & Chrea, C. (2007). DISTATIS for sorting tasks…
- Abdi H., & Valentin, D (2007). STATIS
- Abdi H., & Valentin, D (2007). Multiple correspondence analysis
- Abdi H. (2007). R_V coefficient.
- Abdi H., Valentin, D., O'Toole, A.J., & Edelman, B. (2005). DISTATIS

SOURCES: You can get them (and more) from:
www.utdallas.edu/~herve

Outline:

Standard Profile

3 New Ways

A Common Statistical Structure

One Tool Can Do it All

DISTATIS

Conventional Profiling

4-step process:

1. **Panel selection**
2. **Development of a lexicon for appearance, odor, flavor, texture, mouth feel and after-taste**
3. **Training**
4. **Ratings of the products**

SLIDE FROM DOMINIQUE VALENTIN'S PRE-SYMPOSIUM TALK

Conventional Profiling

Provide

- **Reliable, repeatable, detailled results**
- **Use very well trained assessors**

But

- **Takes a long time**
- **Cost a lot!**

Classical Profiling: Assessors

Well trained and selected

⇨ reliable, repeatable, ≈equivalent

⇨ no *strong* need to compare them

⇨ drop them from the busy maps

SO CLASSIC PROFILING

IS A GREAT TOOL BUT EXPANSIVE

CAN WE FIND
ALTERNATIVE
FASTER & CHEAPER METHODS
(UNLIKELY TO BE BETTER ON TOP!)?

New Profiling Methods: Assessors

Not trained nor selected

- ⇨ reliable, repeatable, ≈equivalent??
- ⇨ *strong* need to compare them
- ⇨ give them a map

THREE NEW TECHNIQUES GOOD FOR PROVIDING PRODUCT INFORMATION (LESS STRONG FOR DESCRIPTION)

1. SORTING TASK + VOCABULARY
2. PROJECTIVE MAPPING / NAPPING
3. FLASH PROFILING

DISCLAIMER:

I AM NOT THE AUTHOR OF ANY OF THEM

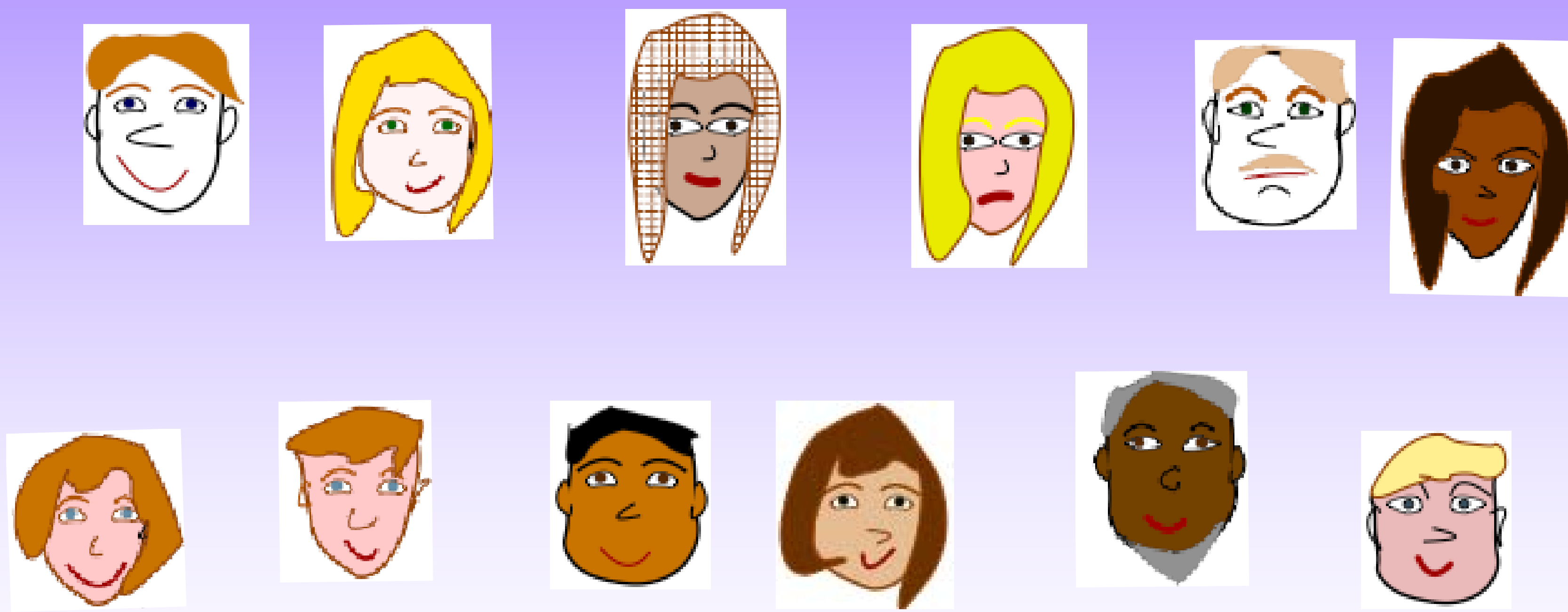
I LIKE THEM ALL!

USE THREE (COMPLETELY FAKED) EXAMPLES

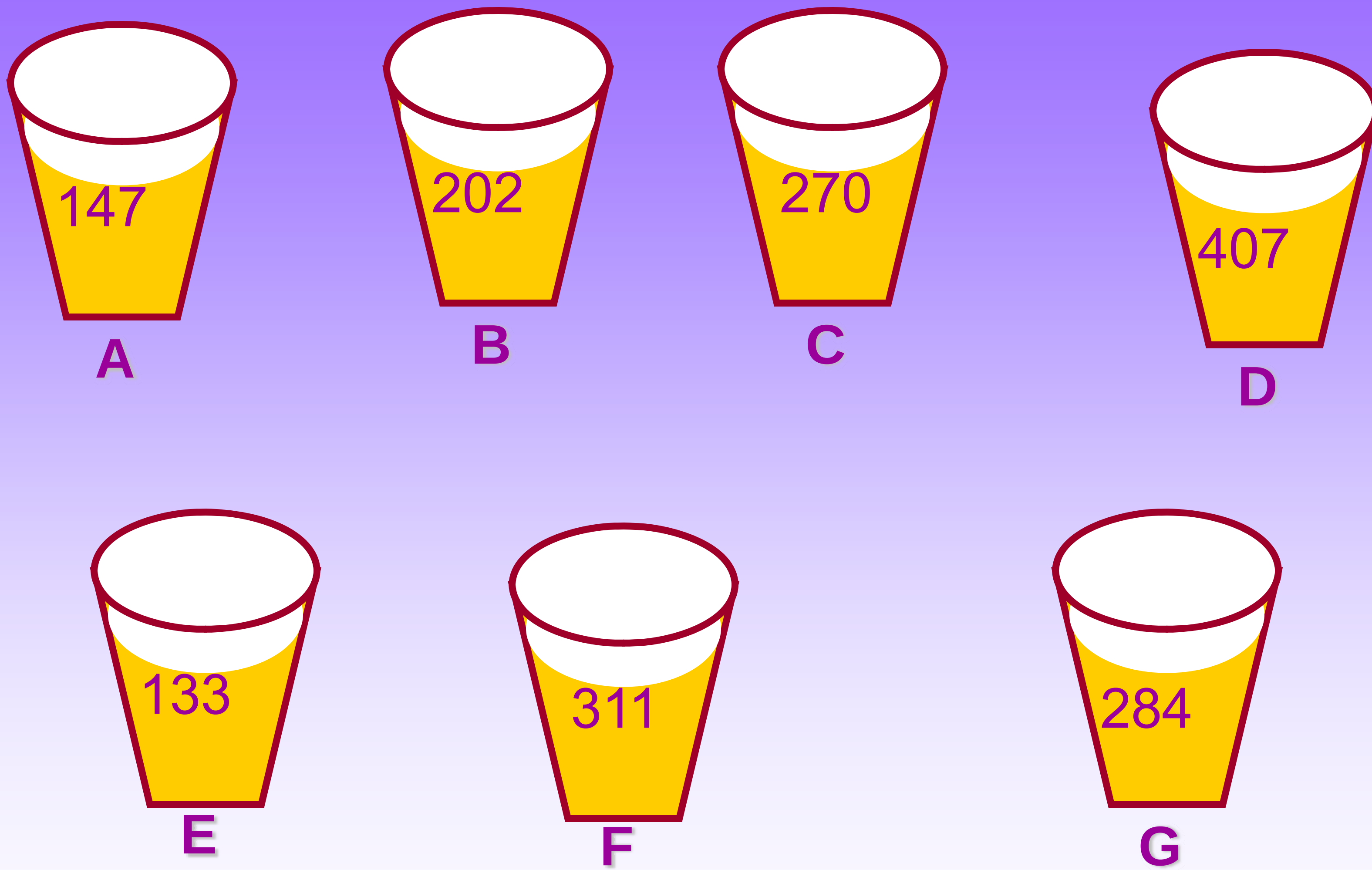
DISCLAIMER:

NO REAL ASSESSOR WAS HARMED IN THESE EXPERIMENTS!

THE BIG GROUP OF FICTITIOUS ASSESSORS



THE FICTITIOUS BEERS



FIRST TECHNIQUE
SORTING PLUS LABELING

- Later this morning: Talk of Bruno Patris *et al.* for a real example

Not Really New!

The original idea came from a *lot* of places
Some pioneers here

Psycholinguistics: G.A. Miller (1964)
Social psychology: Rosenberg (1970)
Anthropology: Ramsey (1970)
etc …

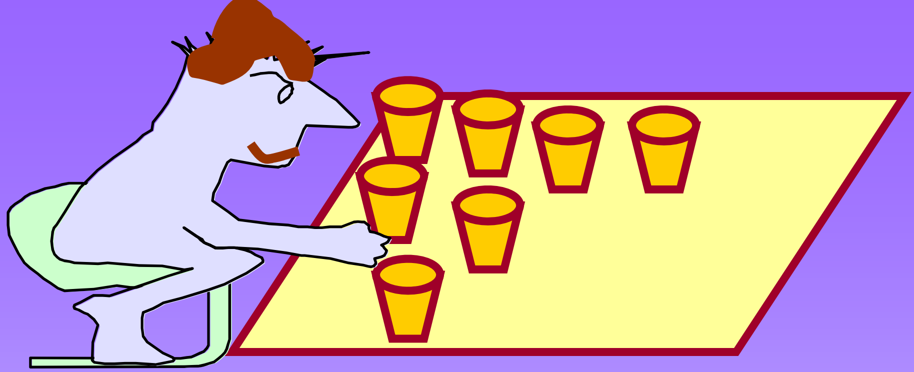
Eleven Assessors for the sorting task

1 2 3 4 5 6

7 8 9 10 11

6F/ 5M

- LET’S START WITH ASSESSOR 1
  FIRST WE SORT
  THEN WE TALK



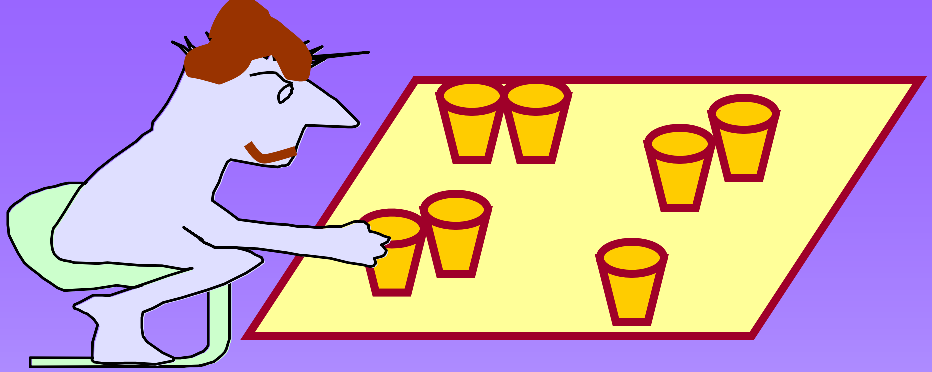
1

Assessors

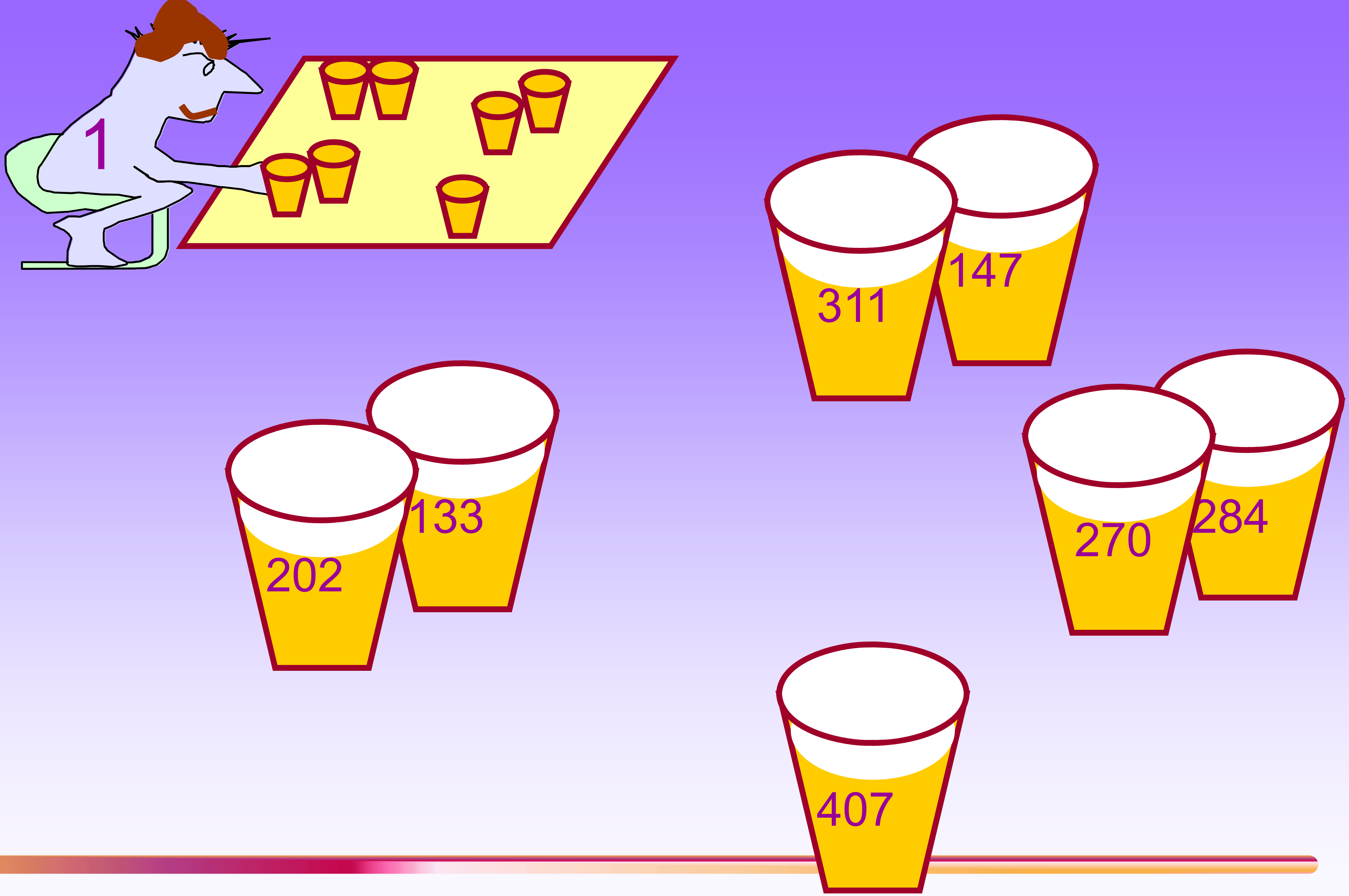
can make as many or as few

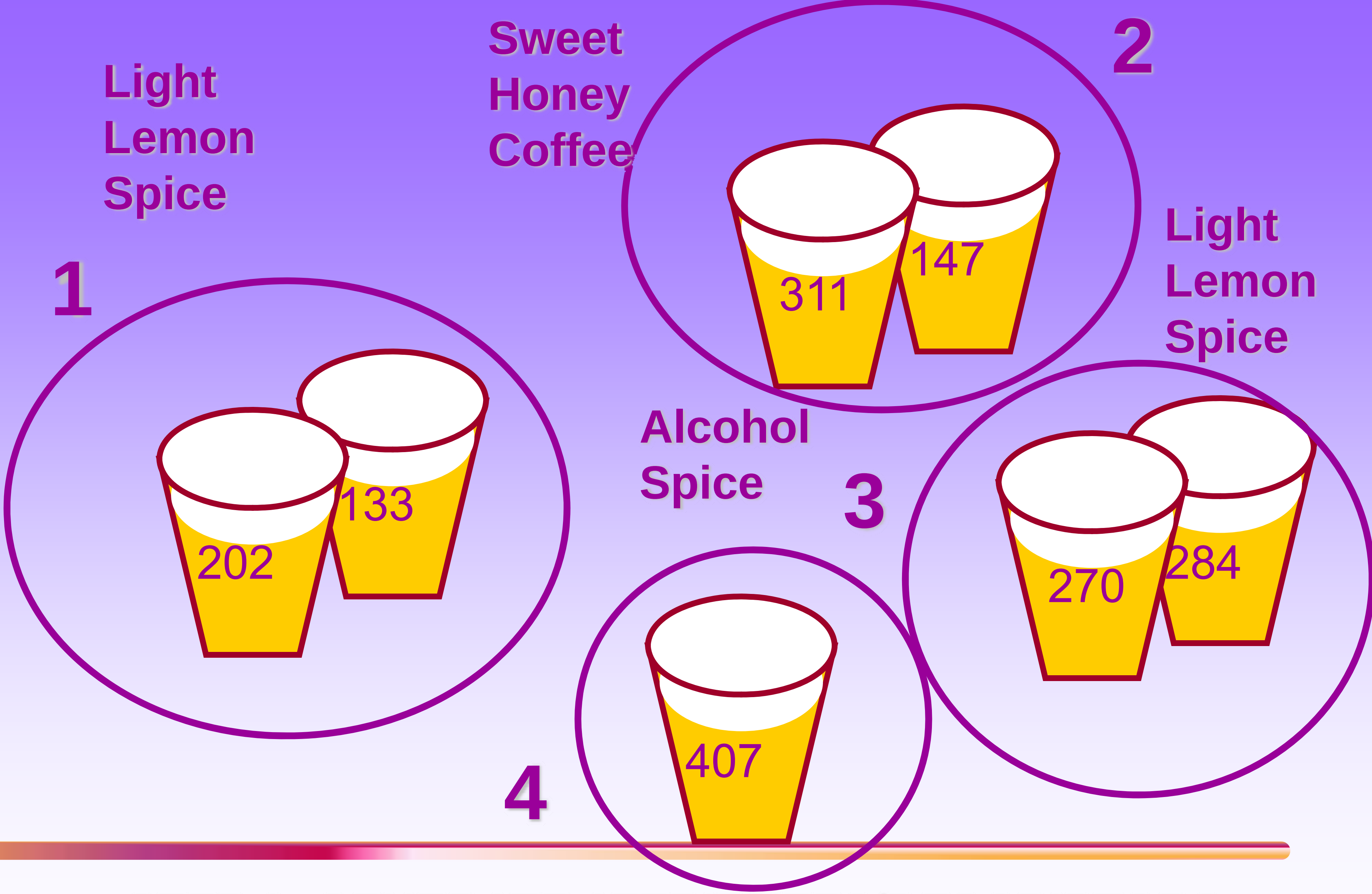
categories as they wish

STEP 1: SORTING THE BEERS



1


1
311
147
202
133
270
284
407

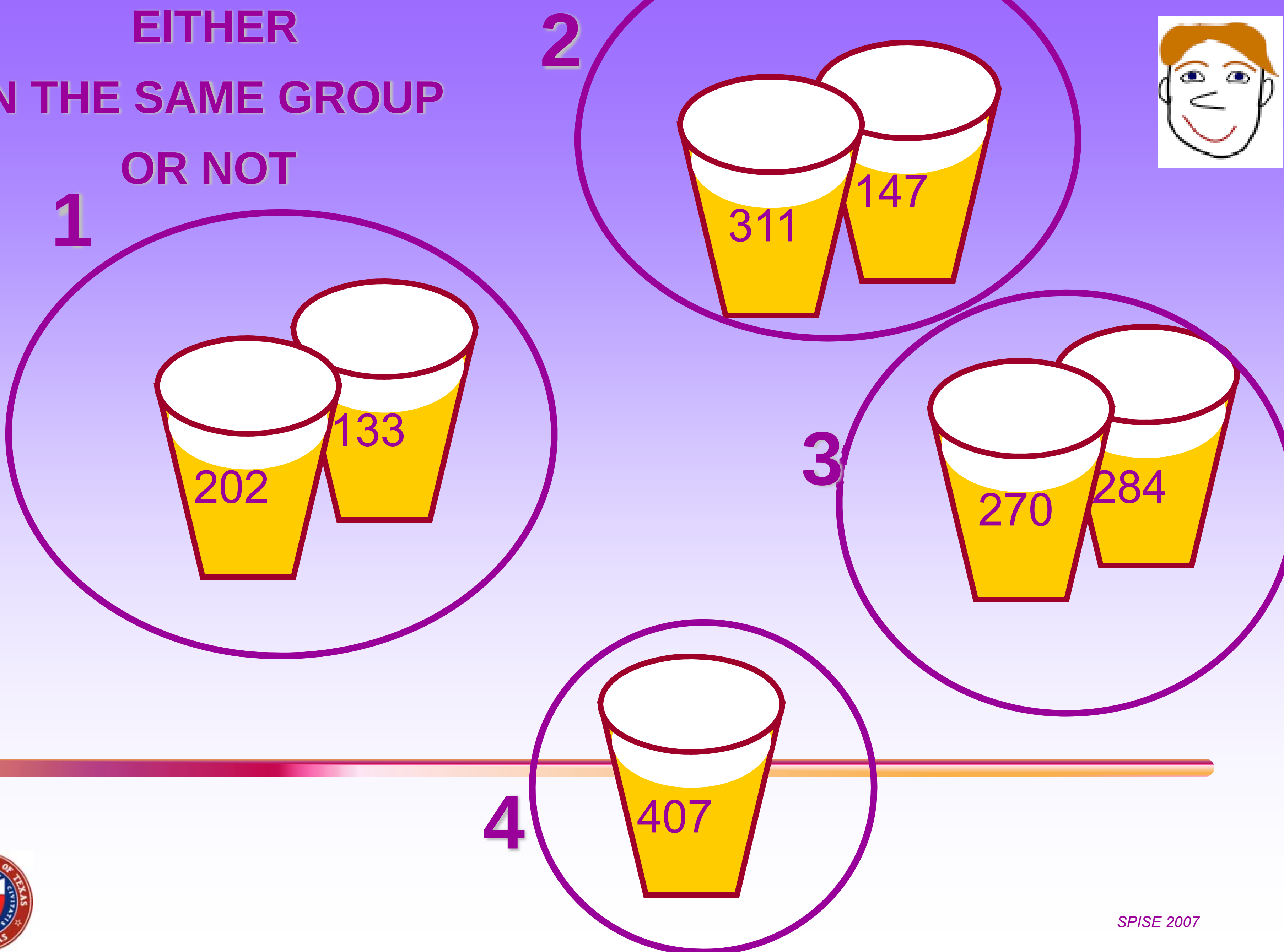


STEP 2: GENERATE ATTRIBUTES/DESCRIPTORS FOR THE GROUPS OF BEERS

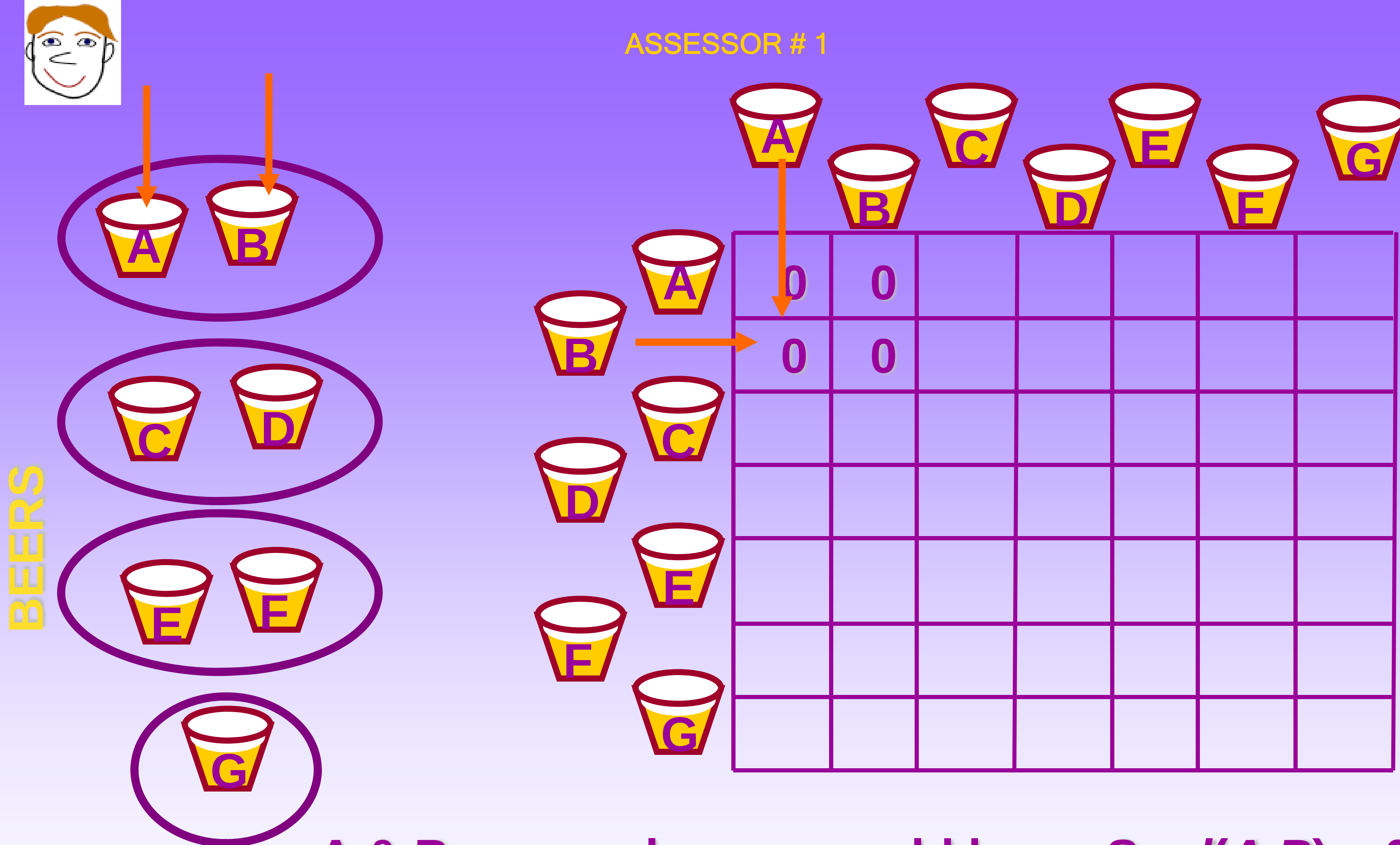
- WHAT DO WE DO WITH THE DATA NOW?

- FIRST: THE BEERS

TWO BEERS ARE EITHER IN THE SAME GROUP OR NOT

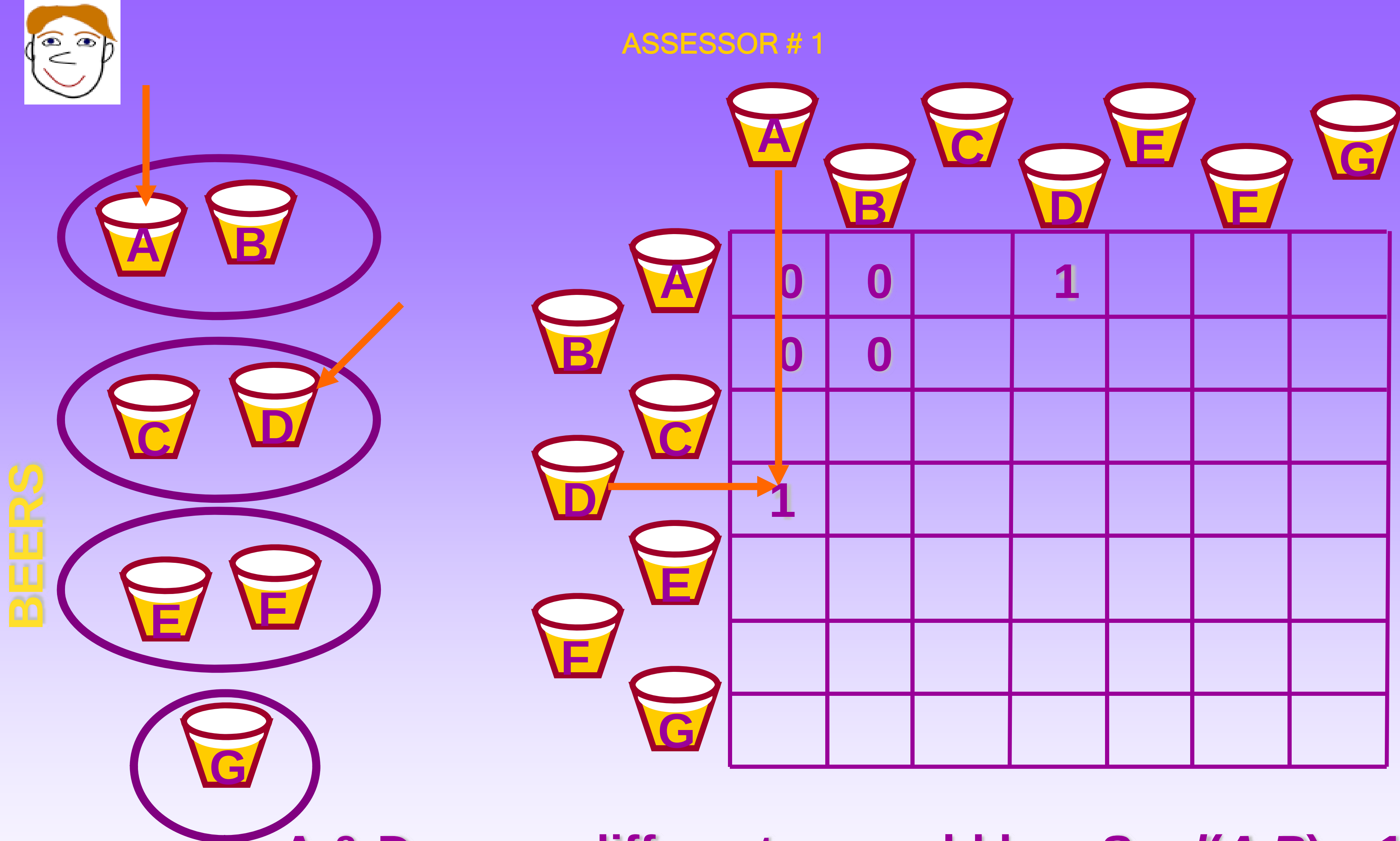


1



A & B are as close as could be. So $d(A,B) = 0$

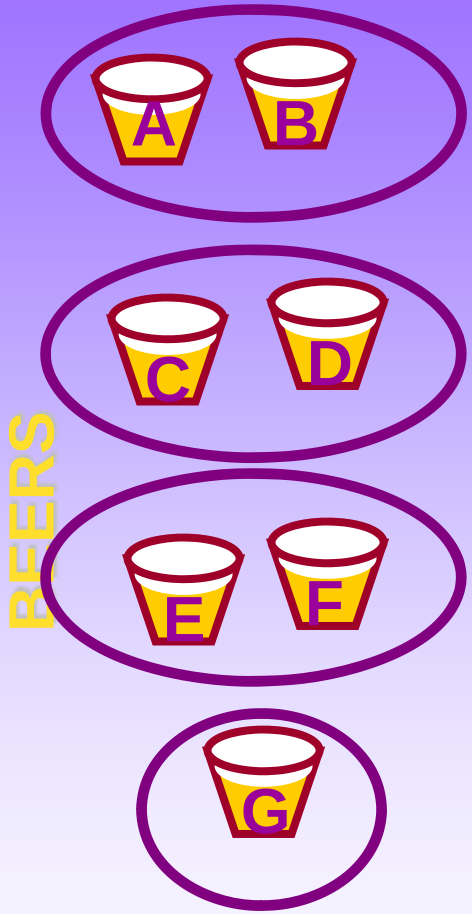
CODING SORTING: FROM PILES TO NUMBERS



A & D are as different as could be. So $d(A,B) = 1$

CODING SORTING: FROM PILES TO NUMBERS

ASSESSOR # 1



	A	B	C	D	E	F	G
A	0	0	1	1	1	1	1
B	0	0	1	1	1	1	1
C	1	1	0	0	1	1	1
D	1	1	0	0	1	1	1
E	1	1	1	1	0	0	1
F	1	1	1	1	0	0	1
G	1	1	1	1	1	1	0

Come on Everybody!

A Distance?
(cultural interlude)

Definition:

$d(i,j) \geq 0$

$d(i,i) = 0$

$d(i,j) \geq 0 \; d(i,k) + d(k,j)$ [triangle inequality]

G.A. Miller (1964):

Number of times

two things are *not* together

is a distance.

- WHAT ABOUT THE ATTRIBUTES

- PUT THEM INTO A TABLE TOO

ASSESSOR # 1 ATTRIBUTE TABLE

ATTRIBUTE

BEERS

	Alcohol	Coffee	Heavy	Honey	Lemon	Light	Spice	Sweet
A					1	1	1	
B					1	1	1	
C		1		1				1
D		1		1				1
E	1		1	1				1
F	1		1	1				1
G	1						1	

- TWO DATA TABLES

BEERS

BEERS

Distance 0/1

D

ATTRIBUTES

BEERS

Association 0/1

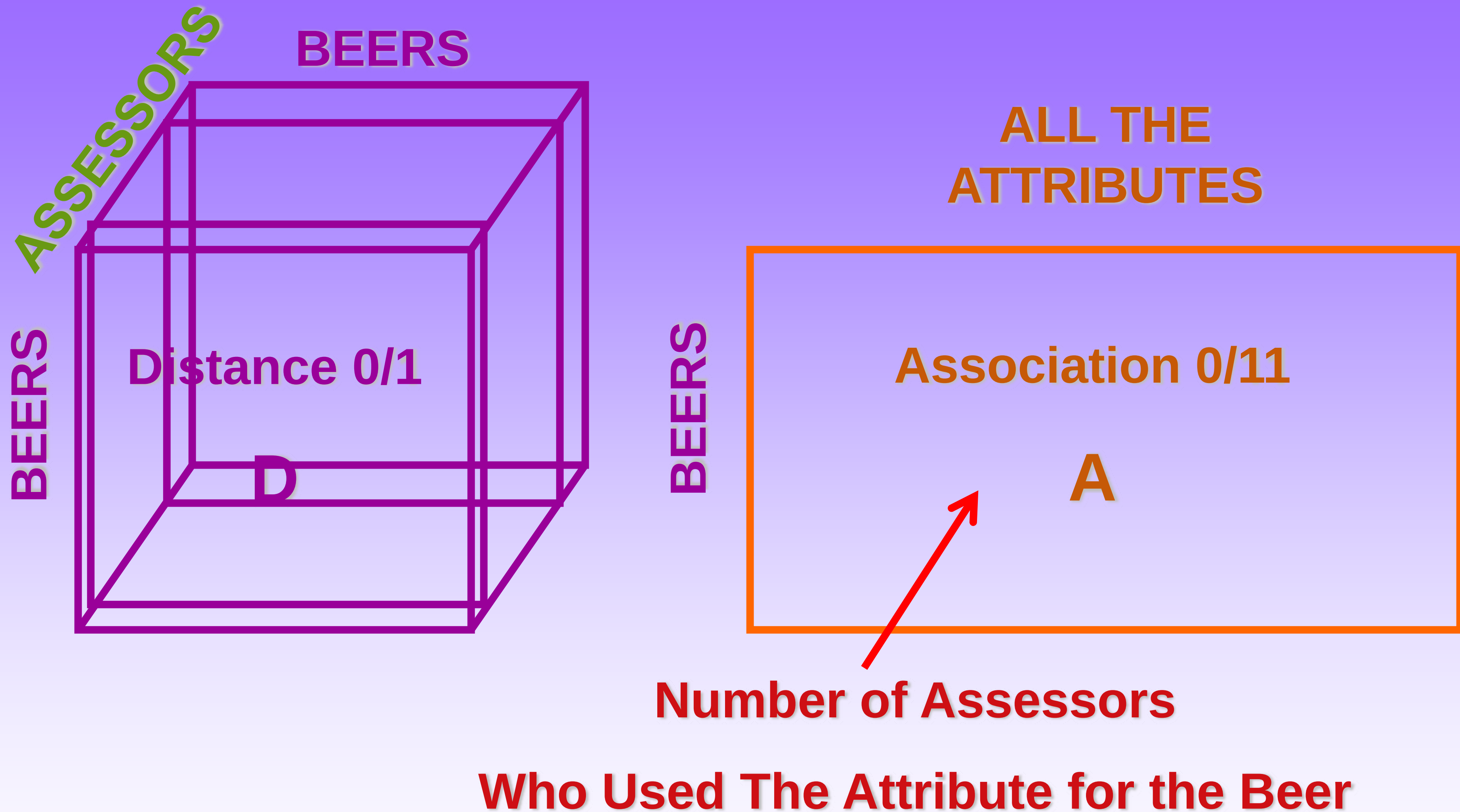
A

SO FOR ASSESSOR 1. WE HAVE:

- CUBES AND RECTANGLES

PUT ALL THE ASSESSORS TOGETHER

ONE DATA CUBE & ONE DATA TABLE



SO FOR 11 ASSESSORS. WE HAVE:

BUT THE ORIGINAL DATA TABLES LOOKED LIKE THIS

	Assessors										
	1	2	3	4	5	6	7	8	9	10	11
Beer											
A	1	1	1	1	1	1	1	1	1	1	1
B	1	2	1	1	1	1	1	1	1	1	2
C	2	3	2	2	2	3	2	2	2	2	3
D	2	3	2	2	2	3	3	2	3	3	3
E	3	4	2	3	2	3	4	3	3	3	4
F	3	4	3	3	2	3	4	4	3	3	4
G	4	4	1	3	3	2	5	4	1	1	5

AND THIS

	Assessors										
Group	1	2	3	4	5	6	7	8	9	10	11
1	{AB} light lemon spice	{A} light lemon spice	{ABG} spice citrus dentist	{AB} floral citrus	{AB} lemon light	{AB} lemon light	{AB} dentist lemon light	{AB} spice acid light	{ABG} fruity spicy citrus	{ABG} coriander lemon light	{A} cilantro light
2	{CD} sweet honey coffee	{B} coriander lemon light	{CDE} bitter fruity toasted	{CD} toffee toasted	{CDEF} alcohol bitter spice	{G} alcohol dentist fruity	{C} sweet toffee bitter	{CD} coffee toasted	{C} toffee bitter sweet	{C} dentist acid light	{B} dentist acid lemon
3	{EF} sweet alcohol heavy heavy	{CD} sweet coffee	{F} alcohol fruity	{EFG} fruity alcohol spice	{G} sweet coffee bitter	{CDEF} sweet coffee alcohol alcohol	{D} strong fruity	{E} alcohol fruity	{DEF} alcohol sweet sweet bitter	{DEF} honey toasted	{CD} honey
4	{G} alcohol spice	{EFG} sweet alcohol fruity					{EF} alcohol sweet fruity bitter	{FG} dentist strong fruity			{EF} alcohol
5							{G}				{G} strong

SECOND TECHNIQUE
PROJECTIVE MAPPING
TABLECLOTH ANALYSIS
NAPPING®

➘Later this morning: Talk of
Lucie Perrin *et al.*
for a real example

Quite New!

Two Independent Origins

Projective mapping: Risvik *et al.* (1994, 1997)

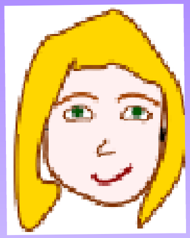
Tablecloth a.k.a. Napping: Pagès (2005)

etc …

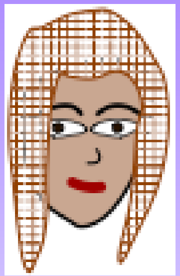
Seven Assessors for projective mapping / napping

1

2



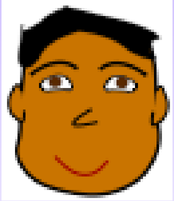
3



4

5

6



7

4F/ 3M

PLACE THE BEERS

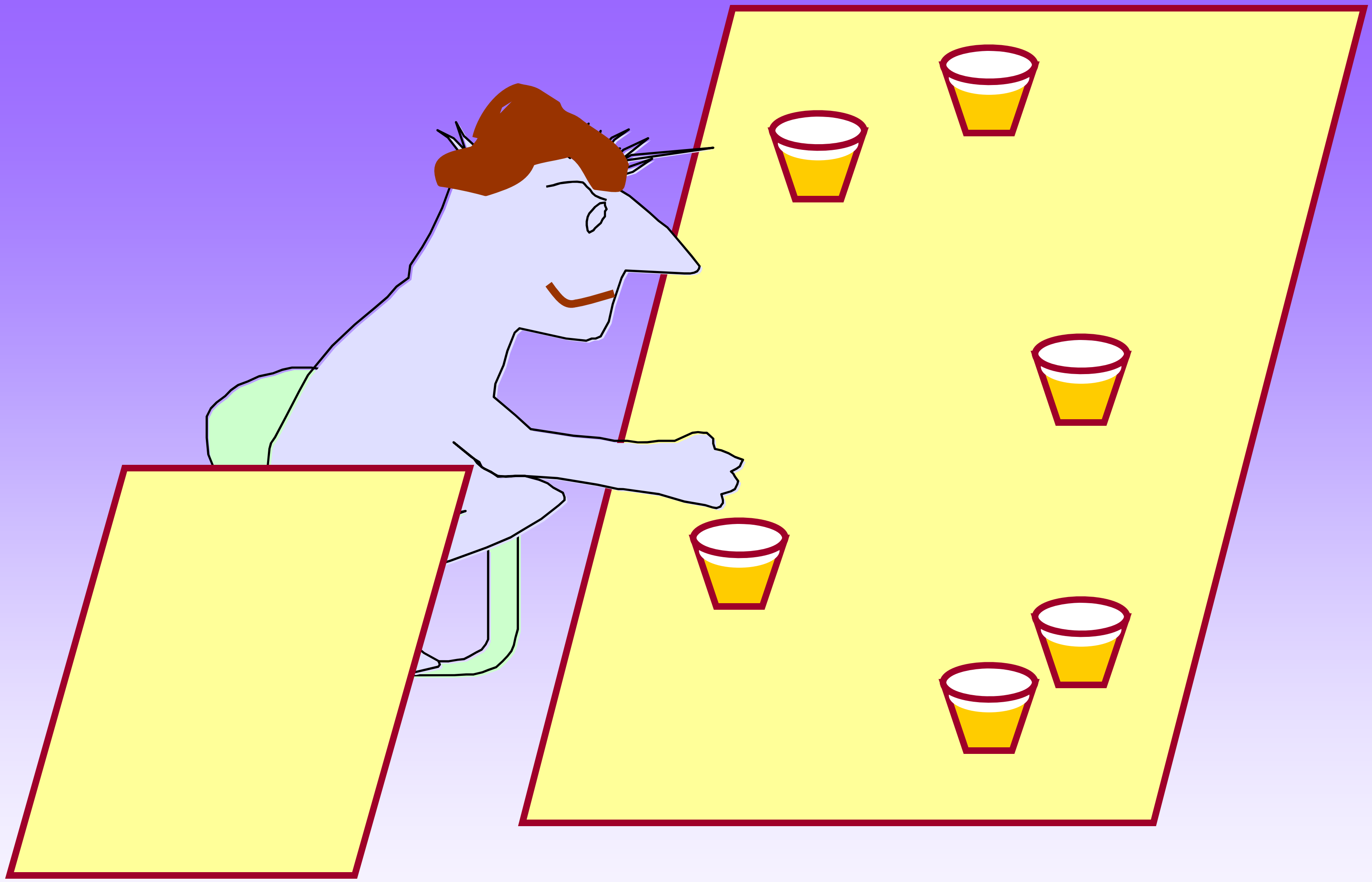
ON THE TABLECOTH

TO SHOW HOW THE

BEERS ARE RELATED

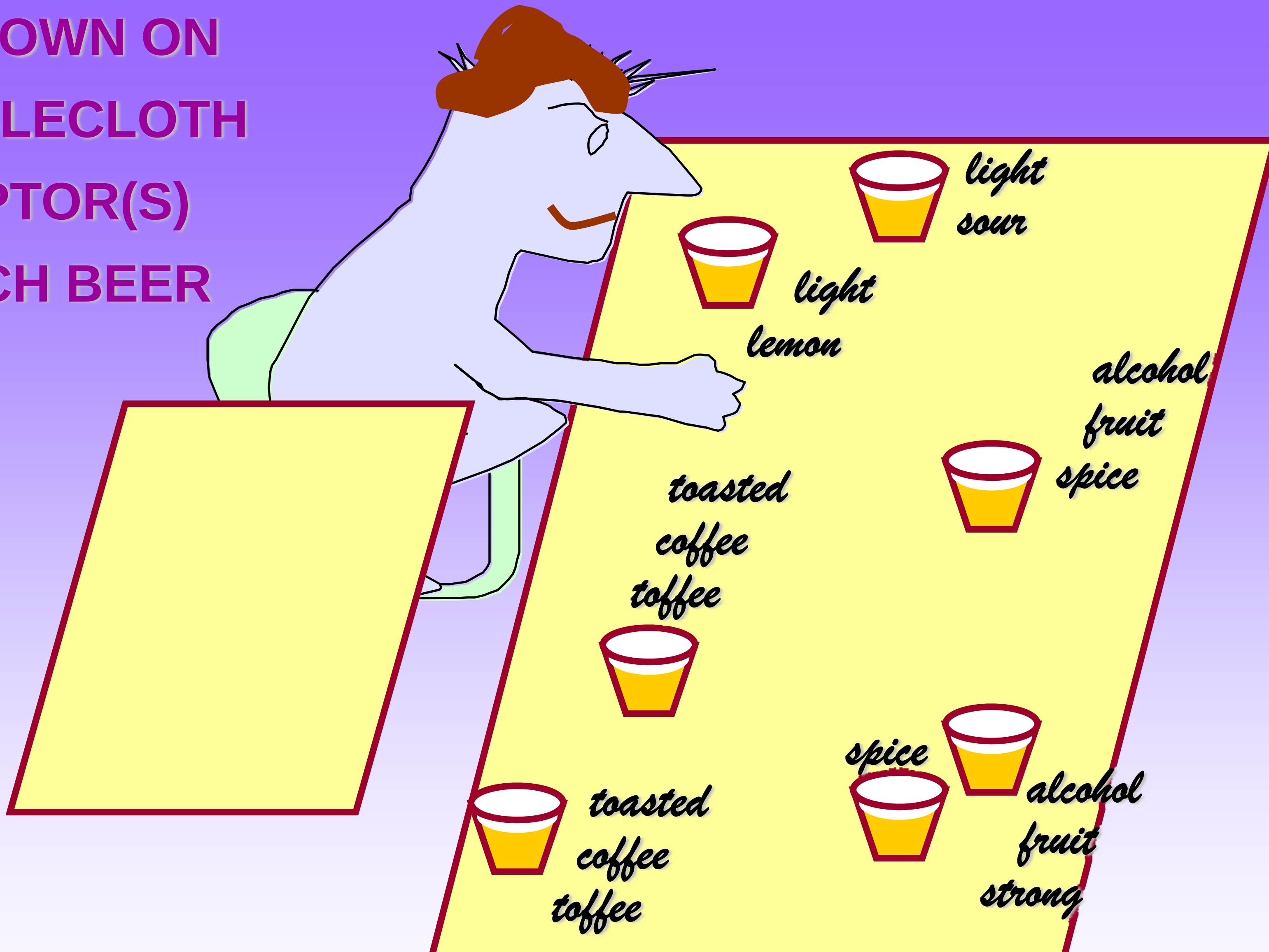
TO EACH OTHER

STEP 1: POSITIONING THE BEERS



STEP 1: POSITIONING THE BEERS

WRITE DOWN ON THE TABLECLOTH DESCRIPTOR(S) FOR EACH BEER



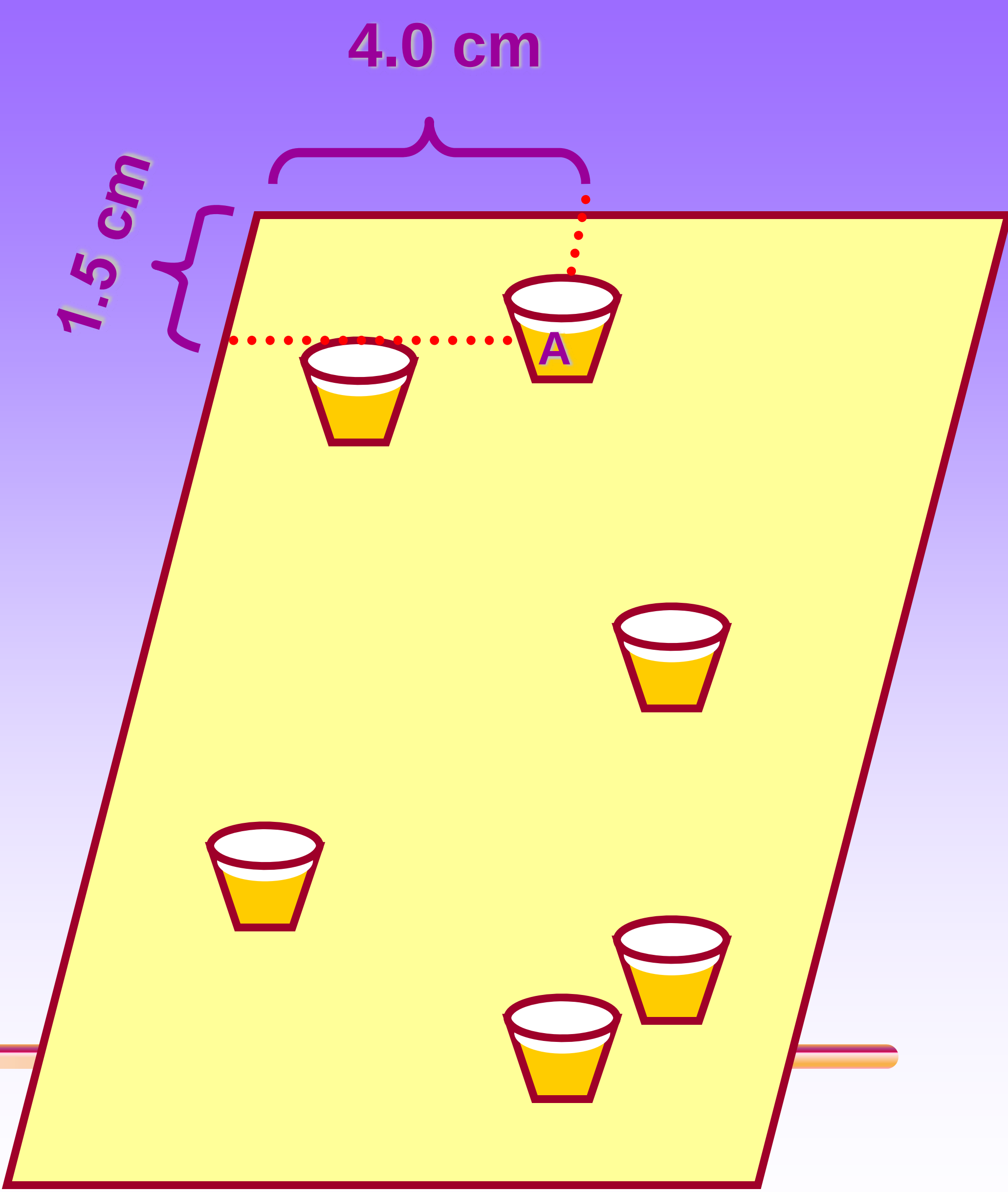
STEP 2: COLLECTING & POSITIONING ATTRIBUTES

- WHAT DO WE DO
  WITH THE DATA NOW?

- FIRST: THE BEERS

DIMENSIONS

	D1	D2
A	1.5	4.0



- Positions of the Beers for Assessor 1

Beer	D1	D2
A	1.5	4.0
B	2.5	2.0
C	8.0	1.5
D	6.0	3.0
E	7.5	7.5
F	5.5	8.5
G	4.5	6.5

KEEP ON TILL WE GET

ASSESSOR # 1

Beer	D1	D2
A	1.5	4.0
B	2.5	2.0
C	8.0	1.5
D	6.0	3.0
E	7.5	7.5
F	5.5	8.5
G	4.5	6.5

	A	B	C	D	E	F	G
A	0	5	49	21	48	36	15
B		0	31	12	55	51	24
C			0	6	36	55	37
D				0	22	31	15
E					0	5	10
F						0	5
G							0

From the coordinates (and Pythagorean theorem) we get an Euclidean Distance (could have measured directly)

- WHAT ABOUT THE ATTRIBUTES

- PUT THEM INTO A TABLE TOO

ASSESSOR # 1 ATTRIBUTE TABLE

ATTRIBUTE

BEERS

	alcohol	coffee	fruity	lemon	light	sour	spicy	strong	sweet	toasted	toffee
A					1	1					
B				1	1						
C		1								1	1
D		1								1	1
E	1		1					1	1		
F	1		1					1	1		
G	1		1				1				

- TWO DATA TABLES

BEERS × **BEERS**

Distance

D

BEERS × **ATTRIBUTES**

Association 0/1

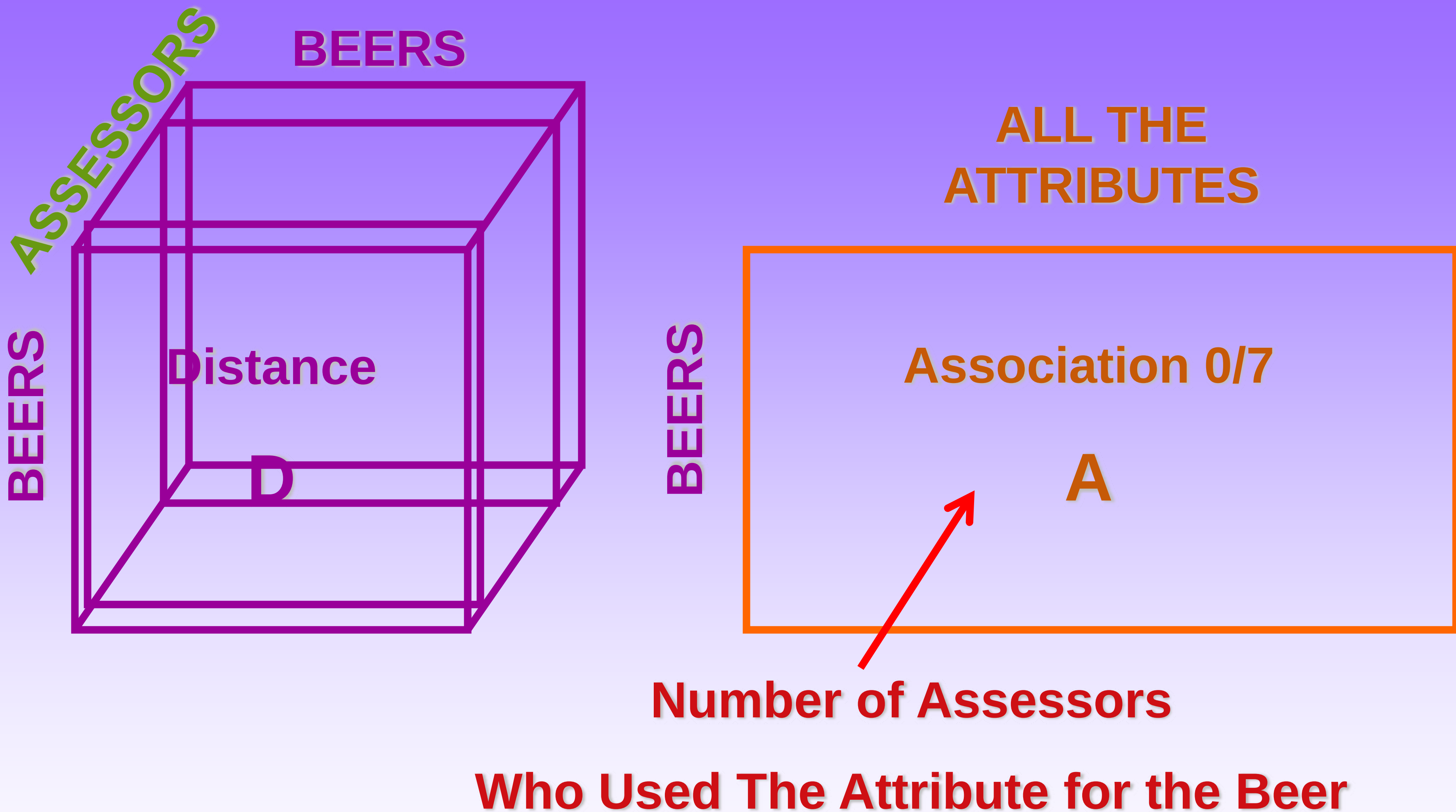
A

SO FOR ASSESSOR 1. WE HAVE:

- CUBES AND RECTANGLES

PUT ALL THE ASSESSORS TOGETHER

ONE DATA CUBE & ONE DATA TABLE



SO FOR 7 ASSESSORS. WE HAVE:

BUT THE ORIGINAL DATA TABLES LOOKED LIKE THIS

	Assessor 1		Assessor 2		Assessor 3		Assessor 4		Assessor 5		Assessor 6		Assessor 7	
	Dimension		Dimension		Dimension		Dimension		Dimension		Dimension		Dimension	
Beer	D1	D2	D1	D2	D1	D2	D1	D2	D1	D2	D1	D2	D1	D2
A	1.5	4.0	1.0	5.0	6.5	2.0	1.0	2.0	8.5	5.0	1.5	5.0	5.0	1.5
B	2.5	2.0	2.5	3.0	8.5	1.0	2.0	3.0	6.5	3.5	1.0	3.0	5.5	1.5
C	8.0	1.5	6.5	5.5	6.0	5.5	7.0	2.5	2.5	3.0	8.0	1.5	0.5	5.5
D	6.0	3.0	8.0	1.0	9.0	6.5	8.0	3.0	1.0	1.5	7.0	2.0	1.0	6.0
E	7.5	7.5	7.0	9.0	3.0	8.0	7.0	5.5	3.5	6.5	8.5	5.5	9.0	8.5
F	5.5	8.5	8.0	8.5	3.5	6.0	9.0	9.0	1.5	8.5	7.0	4.5	8.0	8.0
G	4.5	6.5	4.0	8.0	1.5	3.5	2.0	7.0	6.0	9.0	5.5	6.0	9.0	2.5

AND THIS

Beer	Assessor 1	Assessor 2	Assessor 3	Assessor 4	Assessor 5	Assessor 6	Assessor 7
A	light sour	citrus	light sour	floral spicy lemon	dentist citrus light	light floral	sour
B	light lemon	floral	light spicy	floral spicy lemon	citrus light	light floral	sour
C	toffee coffee toasted	sweet toasted	toasted honey coffee	sweet toasted	coffee	sweet strong honey	toffee
D	toffee coffee toasted	sweet honey	toasted honey	sweet toasted	toffee	sweet strong honey	toffee
E	alcohol fruity strong sweet	heavy sweet alcohol	alcohol sweet strong	fruity sweet	alcohol sweet	sweet strong fruity	alcohol sweet
F	alcohol fruity strong sweet	strong fruity alcoho l	alcohol sweet strong	fruity alcohol	alcohol sweet fruity	sweet strong fruity	alcohol sweet
G	alcohol fruity spicy	spice alcohol	alcohol spicy	dentist alcohol	alcohol dentist	sweet strong dentist	alcohol spicy

THIRD
FLASH PROFILING

➘For a real example

with Vietnamese data:

Blancher *et al.* (2007). FQP.

Quite New Also!

Dairou & Siefferman (2002)

etc …

Definition

Flash Profile (FP) is a sensory descriptive method derived from Free Choice Profiling where each subject chooses and uses his/her own words to evaluate the whole product set comparatively.

→ Describe rapidly the sensory properties of a product set

Dairou V. and Siefferman J.M (2002) A comparison of 14 jams characterized by conventional profile and a quick original method, the Flash Profile. *Journal of food science*, 67,826-834

Methodology: Three Steps

Preliminary session (30 mn to 1h): each assessor write down a first list of descriptors

Inter-session: each assessor read the descriptors written by other assessors

Main Session (2 to 3 h): comparative evaluation of the set of products for each generated descriptor

SLIDE FROM DOMINIQUE VALENTIN'S PRE-SYMPOSIUM TALK

Principle: Rank the beers for each attribute

Descriptor:

\- → +

Descriptor: Fruity

A

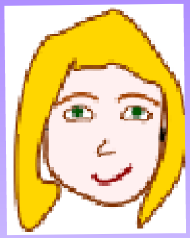
\- B D E G F C → +

SLIDE FROM DOMINIQUE VALENTIN'S PRE-SYMPOSIUM TALK

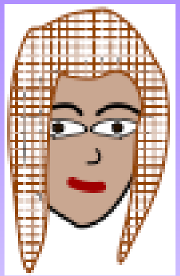
Six (trained) Assessors for the Flash Profiling Task

1

2



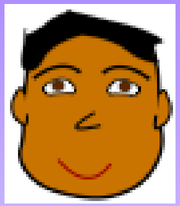
3



4

5

6



3F/ 3M

- Acid
- Acidic
- Alcohol
- Bitter
- cereal
- Citrus
- Clove
- Coriander
- Honey
- Hop
- Floral
- Fruity
- Lemon
- Malt
- Ripe fruit
- Sour
- Spice
- Sweet

FLASH: THE LIST OF ATTRIBUTES

ASSESSOR # 1'S ATTRIBUTE LIST

1

ALCOHOL

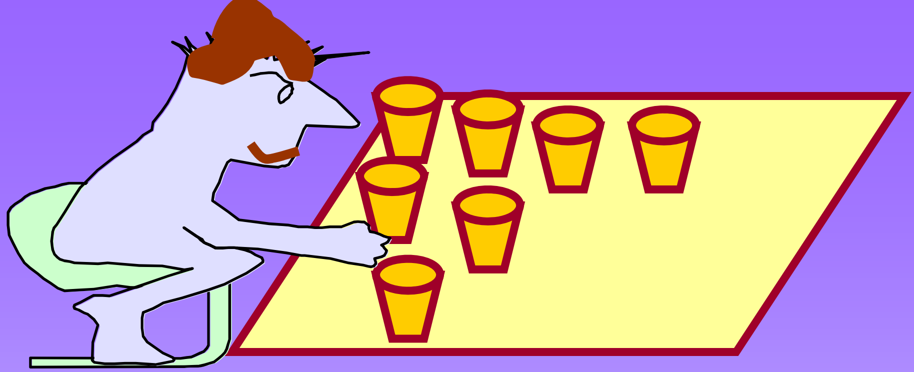
MALT

FRUITY

BITTER

SPICE

HOP



Assessor 1.

Rank the beers

According to each attribute

STEP 1: RANKING THE BEERS

Assessor 1.

Descriptor: Alcohol

\- A B C G D F +

Descriptor: Hop

\- D C E A G B +

SLIDE FROM DOMINIQUE VALENTIN'S PRE-SYMPOSIUM TALK

- WHAT DO WE DO WITH THE DATA NOW?

- FIRST: THE BEERS

ASSESSOR 1 RANKED THE BEERS

	Attributes					
Attribute #	1	2	3	4	5	6
A	1	1	3	1	5	4
B	2	5	1	4	7	7
C	3	6	4	6	1	2
D	6	7	5	2	2	1
E	5	2	7	7	4	3
F	7	4	6	5	3	6
G	4	3	3	3	6	5

ASSESSOR # 1

	Attributes					
Attribute #	1	2	3	4	5	6
A	1	1	3	1	5	4
B	2	5	1	4	7	7
C	3	6	4	6	1	2
D	6	7	5	2	2	1
E	5	2	7	7	4	3
F	7	4	6	5	3	6
G	4	3	3	3	6	5

	A	B	C	D	E	F	G
A	0	43	75	84	71	78	19
B		0	76	101	88	69	18
C			0	29	40	45	54
D				0	63	46	57
E					0	23	42
F						0	33
G							0

From the ranks (and Pythagorean theorem) we get an Euclidean Distance

- WHAT ABOUT THE ATTRIBUTES

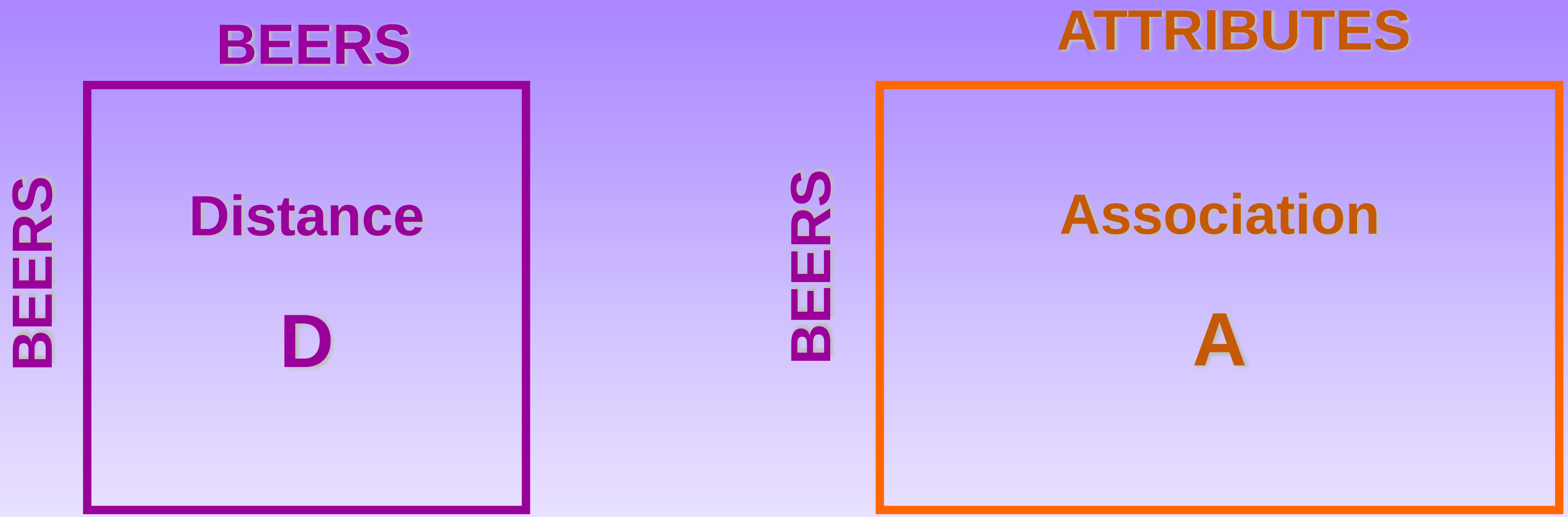
- PUT THEM INTO A TABLE TOO

ASSESSOR 1 RANKED 7 BEERS FOR SIX ATTRIBUTES

	Attributes					
Attribute #	1	2	3	4	5	6
A	1	1	3	1	5	4
B	2	5	1	4	7	7
C	3	6	4	6	1	2
D	6	7	5	2	2	1
E	5	2	7	7	4	3
F	7	4	6	5	3	6
G	4	3	3	3	6	5

The rank matrix is an Association Matrix

- TWO DATA TABLES

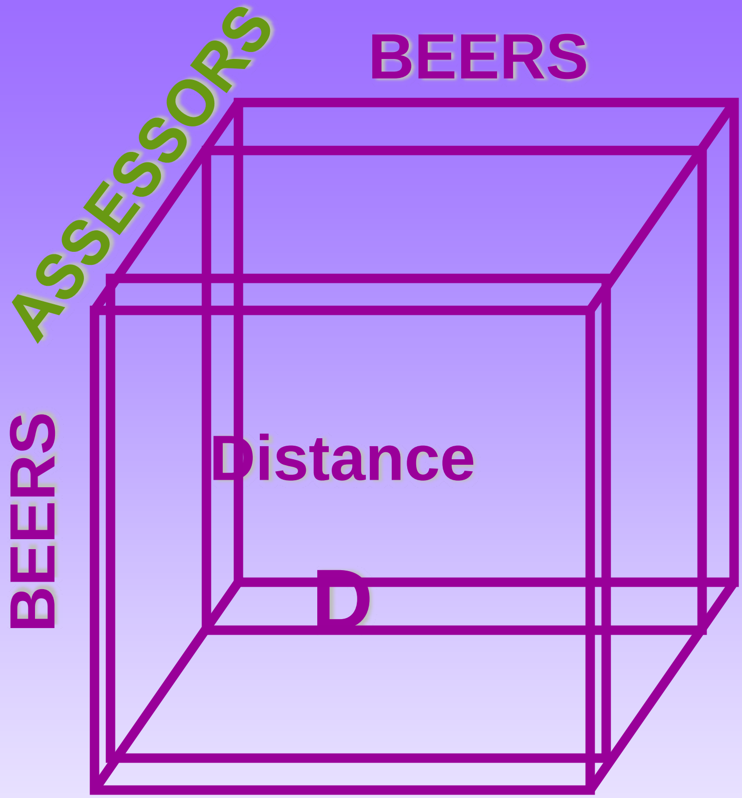


SO FOR ASSESSOR 1. WE HAVE:

- CUBES AND RECTANGLES

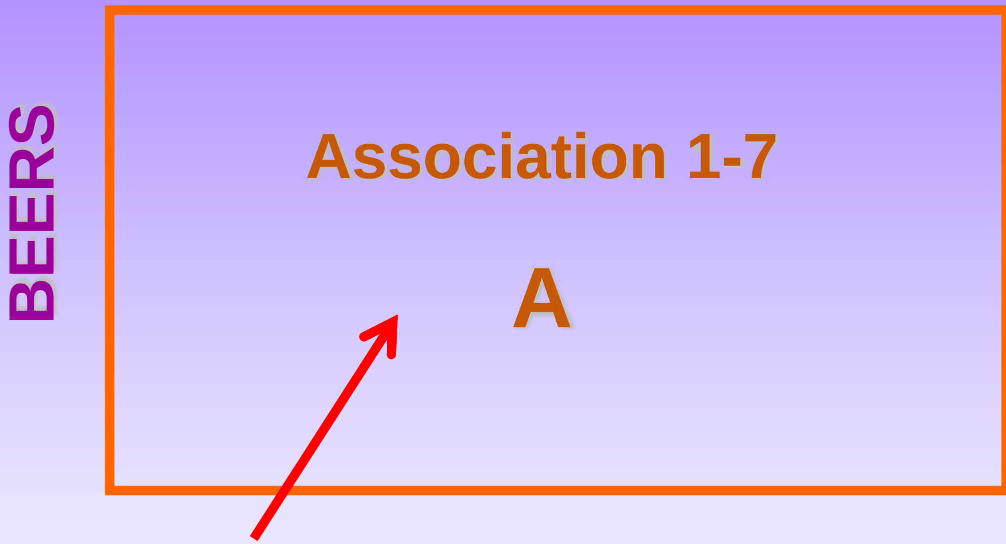
PUT ALL THE ASSESSORS TOGETHER

ONE DATA CUBE & ONE DATA TABLE



ALL THE ATTRIBUTES OF

ALL THE ASSESSORS



Rank Given For a Beer/Attribute

By an Assessor

SO FOR 6 ASSESSORS. WE HAVE:

BUT THE ORIGINAL DATA TABLES LOOKED LIKE THIS

Beer Ranking According To Attributes

	Assessors																
	1						2					3					
	Attributes						Attributes					Attributes					
Attribute #	1	2	3	4	5	6	1	2	3	4	5	1	2	3	4	5	6
A	1	1	3	1	5	4	2	7	4	5	7	6	3	6	1	7	1
B	2	5	1	4	7	7	1	6	5	2	6	7	7	7	2	6	3
C	3	6	4	6	1	2	5	1	7	7	1	2	1	3	3	1	6
D	6	7	5	2	2	1	6	2	6	1	3	1	2	5	5	4	7
E	5	2	7	7	4	3	4	3	1	6	2	3	5	2	6	2	6
F	7	4	6	5	3	6	7	5	2	4	4	4	6	1	7	6	5
G	4	3	3	3	6	5	3	4	3	2	5	5	4	4	4	3	2

	Assessors																		
	4							5					6						
	Attributes							Attributes					Attributes						
Attribute #	1	2	3	4	5	6	7	1	2	3	4	5	1	2	3	4	5	6	7
A	2	1	1	1	6	7	4	1	3	1	1	5	3	7	1	6	5	3	1
B	1	5	3	2	7	6	7	5	7	2	2	1	5	6	4	7	7	2	2
C	4	6	7	5	1	3	2	7	2	3	7	7	7	1	6	1	1	7	5
D	3	3	6	6	2	4	1	6	3	5	6	2	6	2	7	2	2	6	3
E	7	7	4	3	5	1	3	4	5	6	5	6	1	3	2	5	3	5	6
F	6	4	5	7	3	2	6	3	6	7	4	4	4	4	5	3	4	4	7
G	5	2	2	4	4	5	5	2	4	4	3	3	2	5	3	4	6	1	4

AND THIS

Attributes Chosen By The Assessors

	Assessors					
	1	2	3	4	5	6
Attribute #						
1	alcohol	alcohol	coriander	ripe fruit	toasted cereal	malt
2	malt	citrus	lemon	bitter	clove	citrus
3	fruity	cereal	honey	sweet	alcohol	coriander
4	bitter	bitter	alcohol	alcohol	sweet	coriander
5	spice	floral	acid	coriander	bitter	hop
6	hop		sweet	honey		sweet
7				hop		alcohol

LOOK LIKE

SORTING

NAPPING

FLASH PROFILING

CREATE

THE SAME STRUCTURE

A CUBE

AND A RECTANGLE

HOW TO ANALYZE THAT

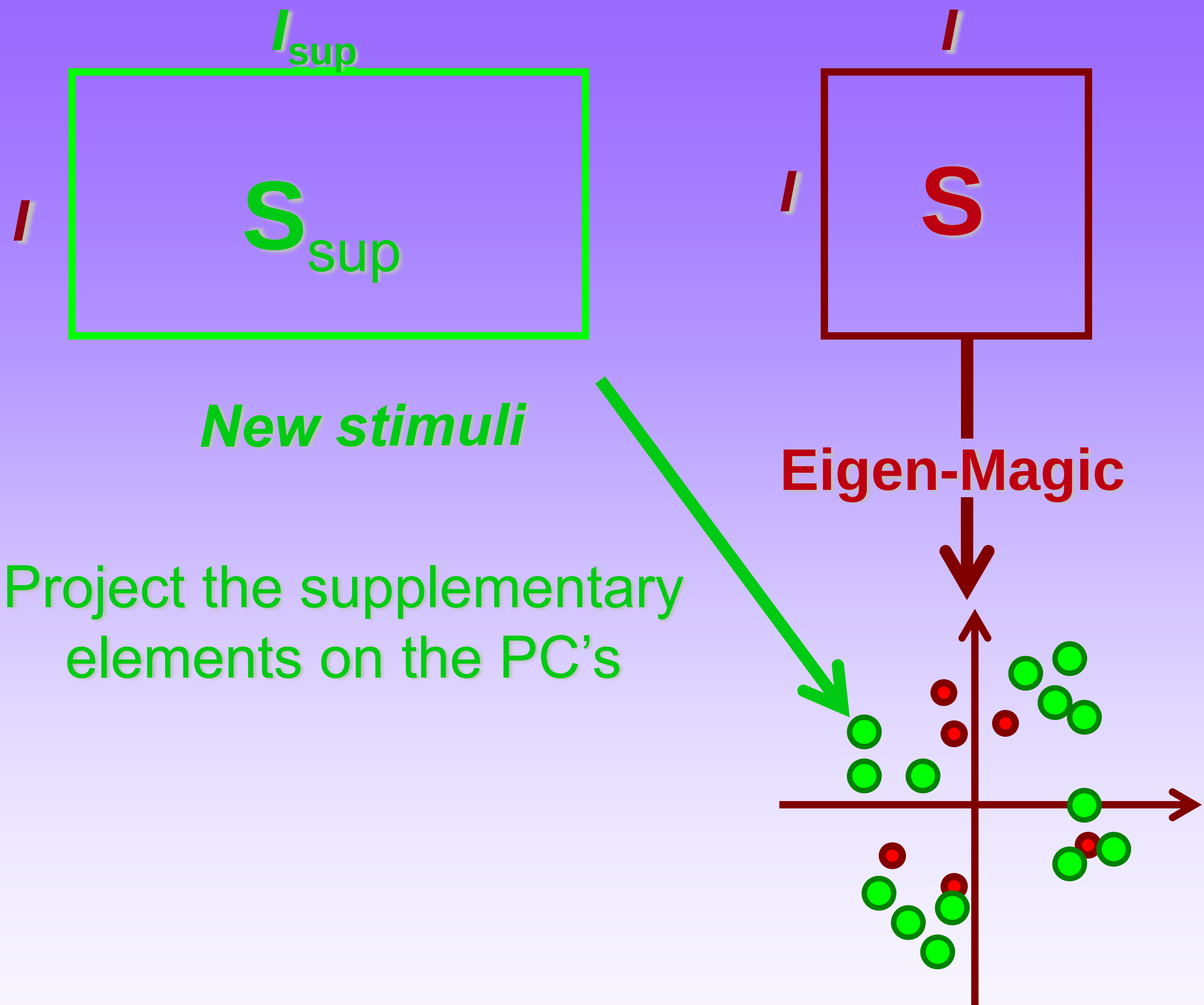
LET'S GO FOR AN
EUCLIDEAN TOUR TO GET
IDEAS

THE EUCLIDEAN WORLD *IS* BEAUTIFUL

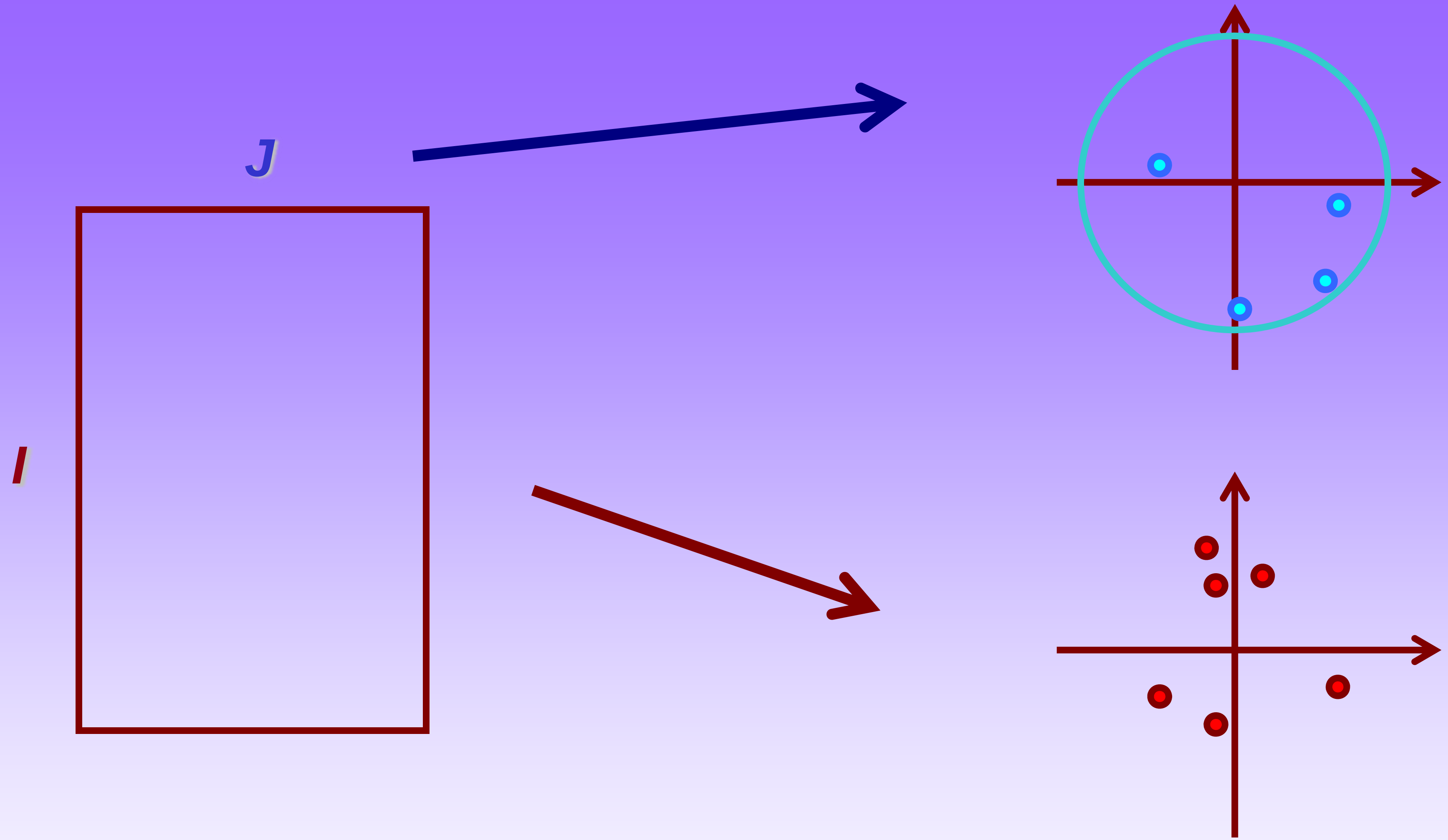
IT WORKS WITH THE PYTHAGOREAN THEOREM!

- **Means, Variance, Inertia are *natural* (cosines and contributions)*.***
- **We love to minimize sum of squares (and the eigenworld is *magic*).**
- **Good routines means very large data sets.**
- **Easy Duality**
- **Generalizes beautifully (masses, weights, …)**
- **And if it is multivariate normal we are *in Paradise!***

- *I* by *I* (similarity) data sets: MDS

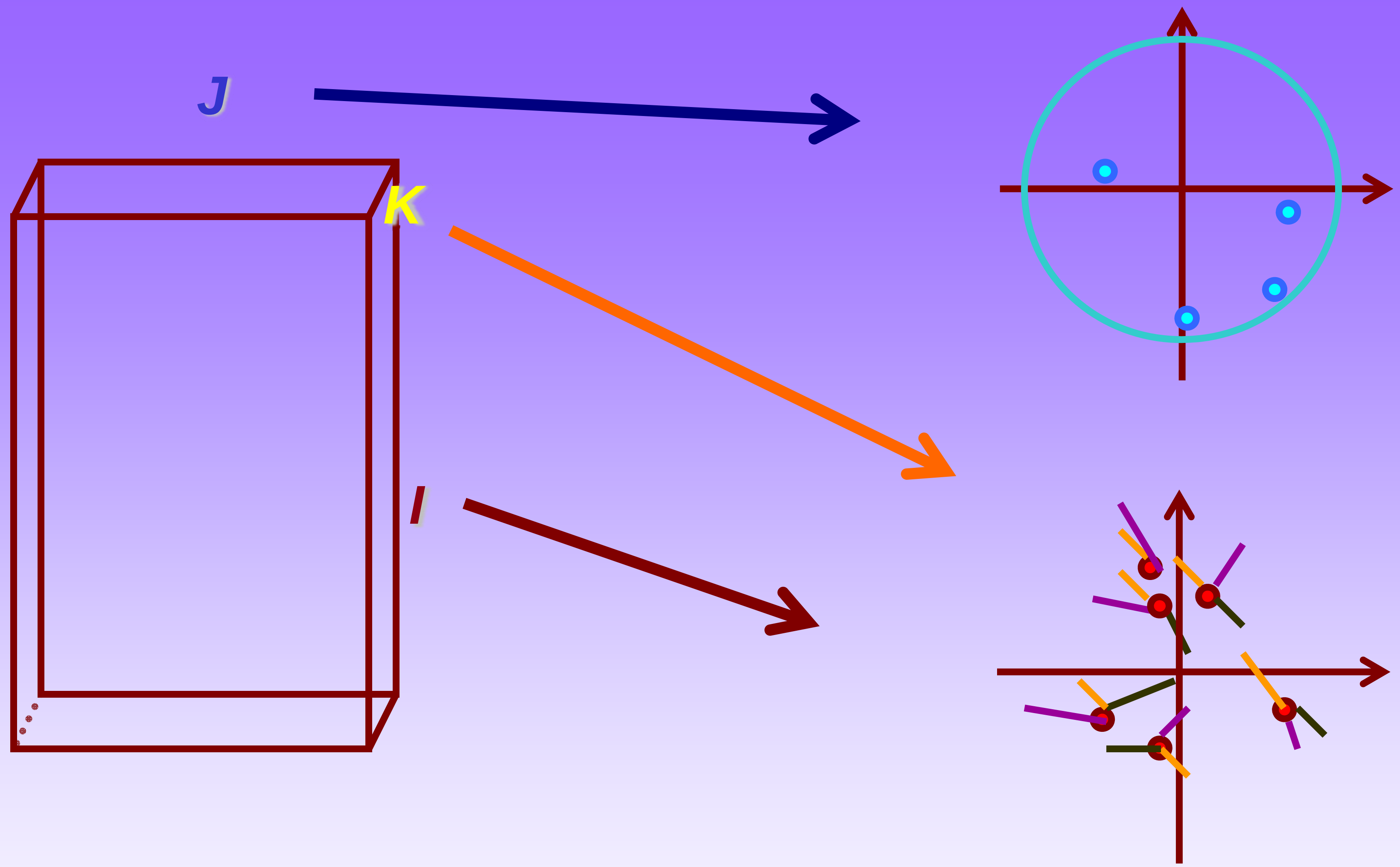


Refresher: Supplementary (illustrative) elements: *e.g.,* "testing set"



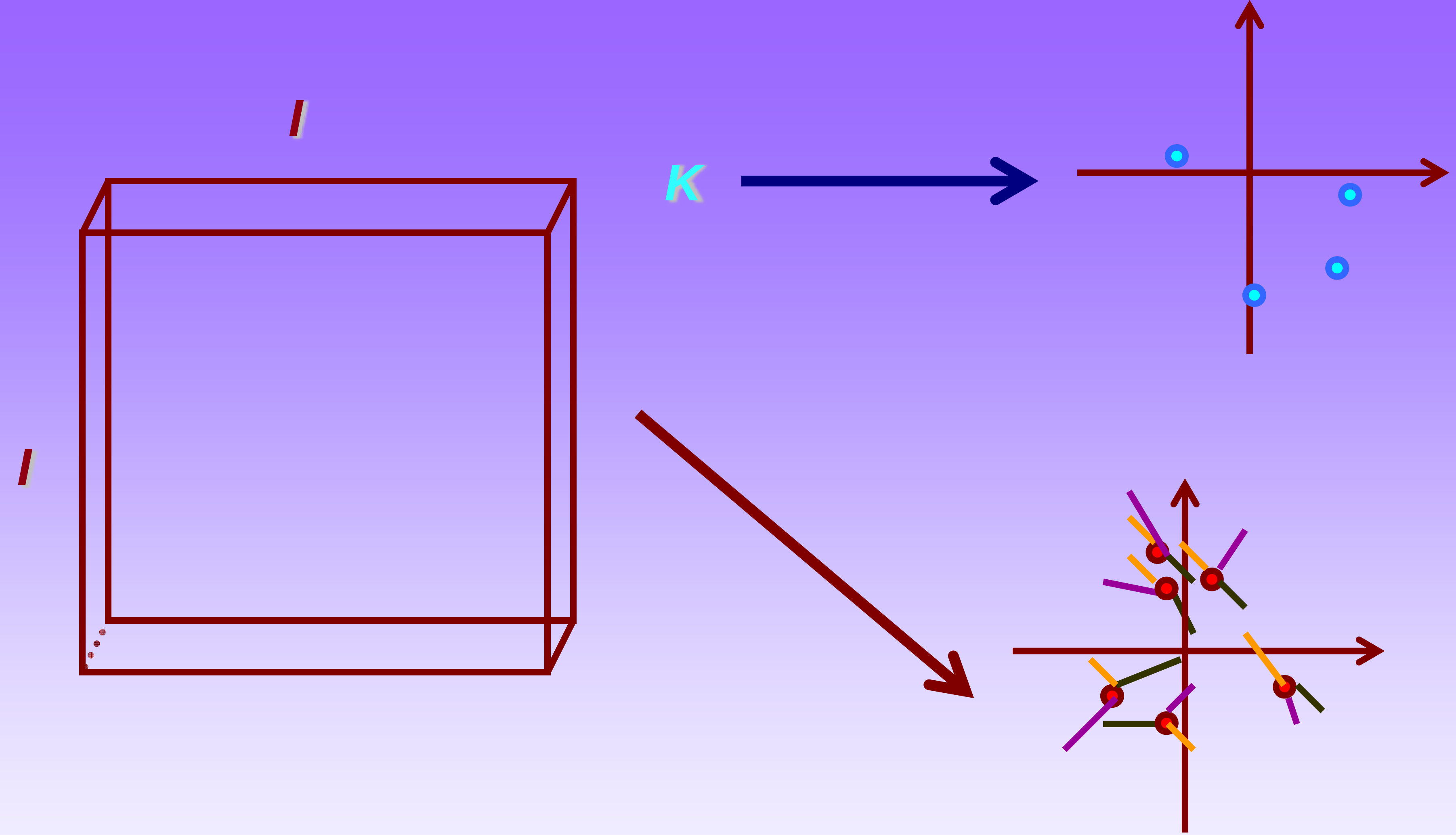
- *I* by *J* data sets: PCA, CA, Biplots, *etc.*

The beauty of Euclide …



- *I* by *J* by *K* data sets: MFA, STATIS Tucker-*n* *etc.*

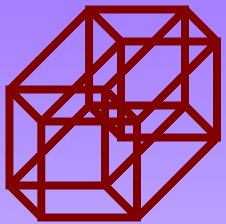
The beauty of Euclide ...



- *I* by *I* by *K* data sets:
  Procrustean family, **DISTATIS** .

The beauty of Euclide …

THE EUCLIDEAN WORLD IS *NOT ALWAYS* BEAUTIFUL



It works with the Pythagorean theorem!

- **It is made of *dimensions,***
- **And we are good only for 2, maybe 3 D.**
- **Sum of squares are sensitive to outliers (not that big of a deal: Robust approaches)..**
- **May not be a good *model* of human similarity**

Refresher PCA and Distance

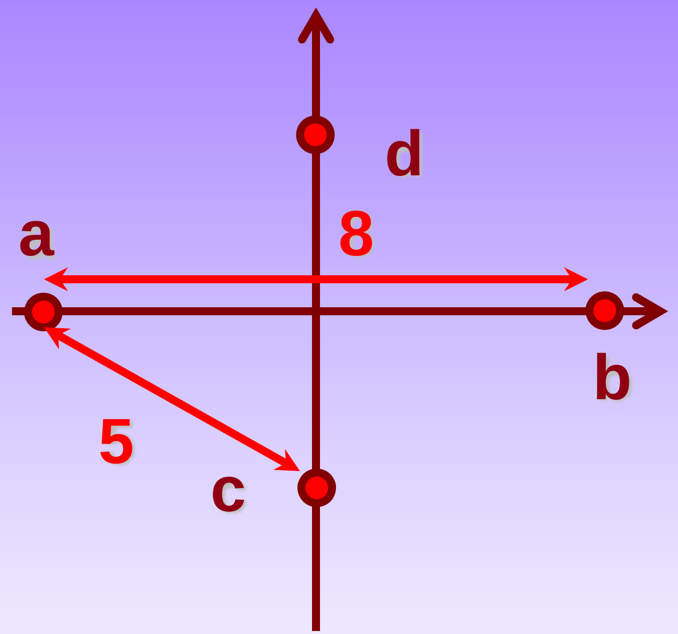
Metric Multidimensional Scaling (Togerson 1958)

Representing a distance matrix as a map: MDS

	a	b	c	d
a	0	8	5	5
b	8	0	5	5
c	5	5	0	6
d	5	5	6	0

Some

Magic



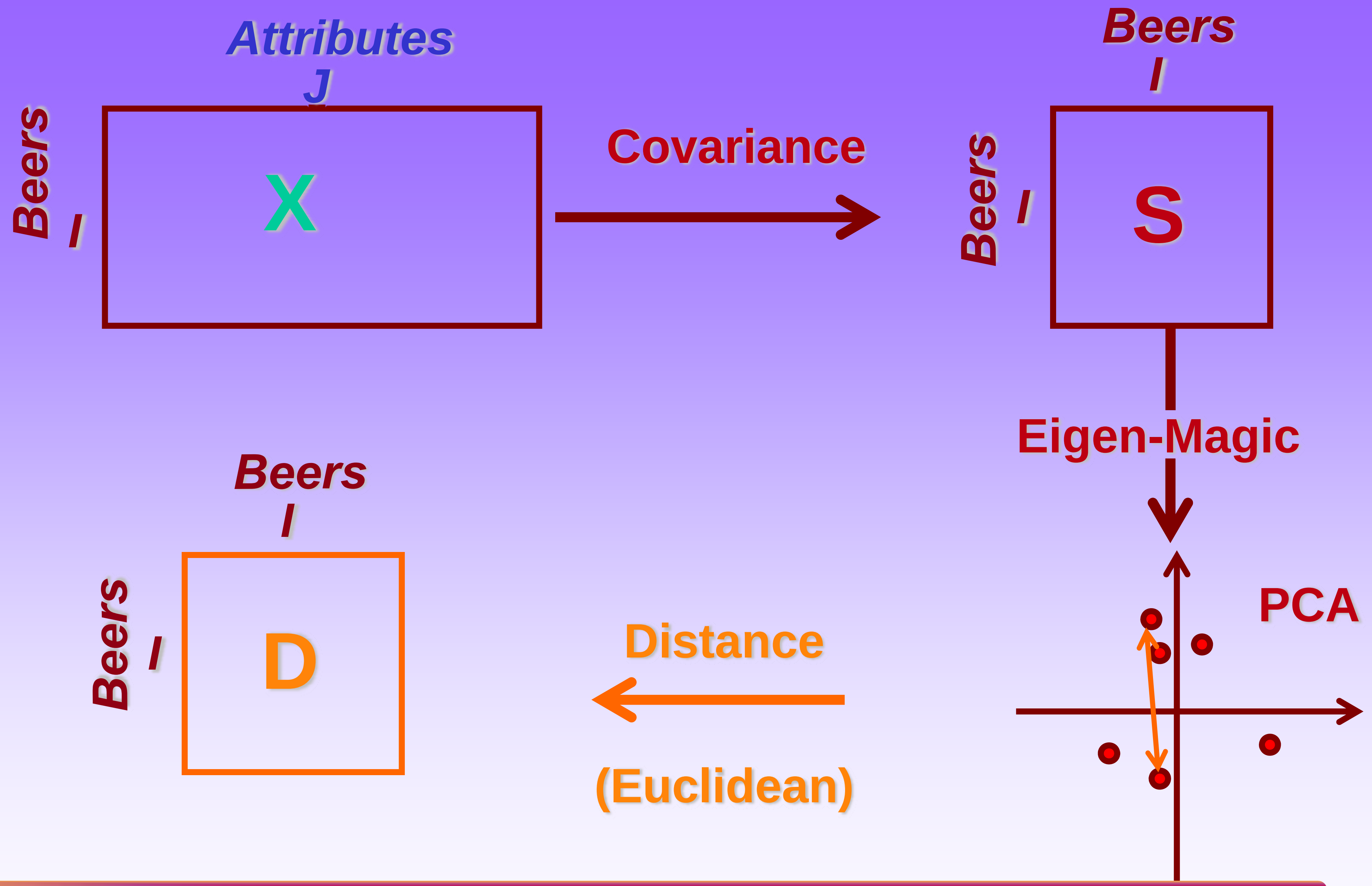
$d(\mathbf{a,a}) = 0$

$d(\mathbf{a,b}) > 0$

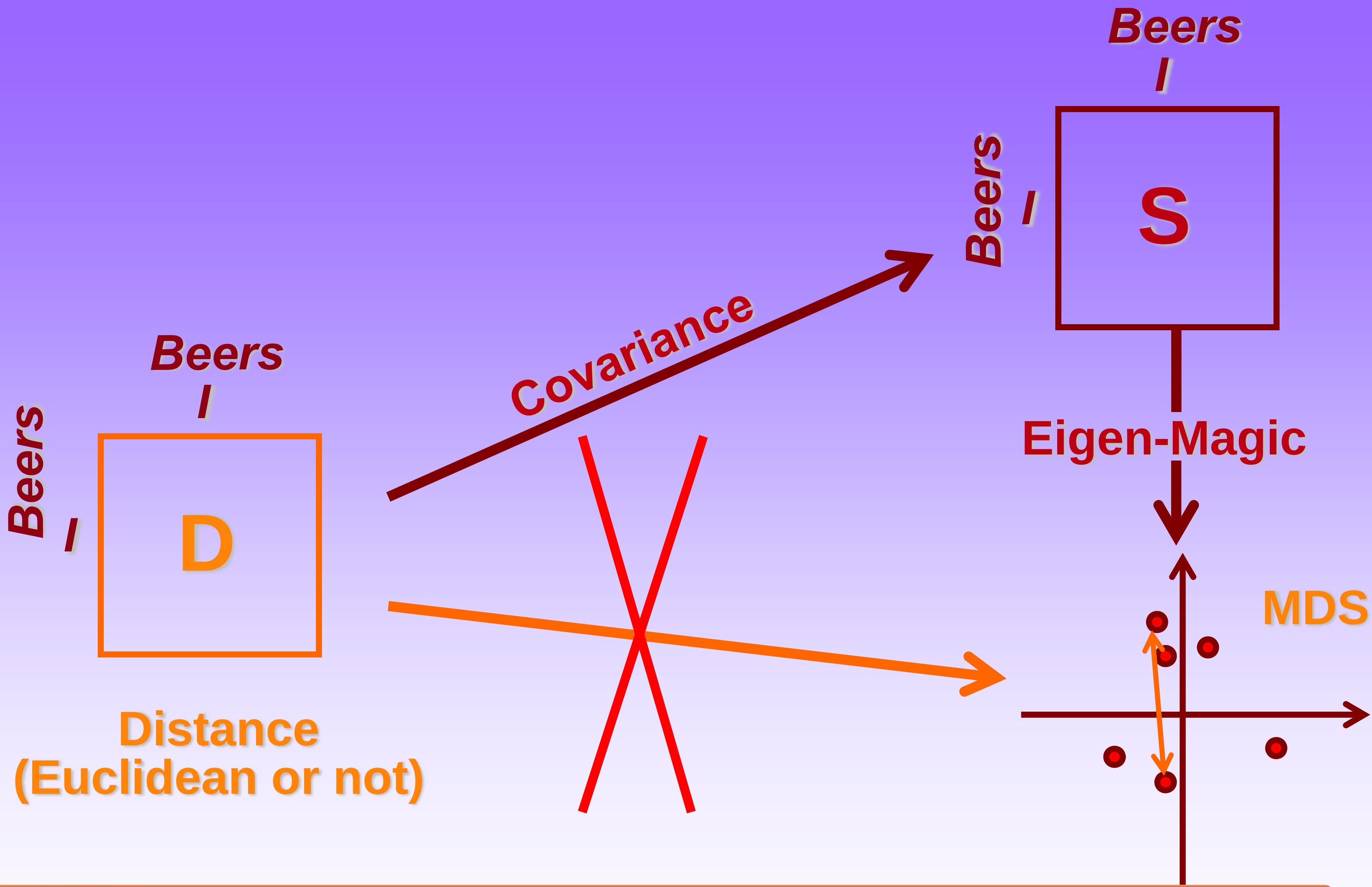
$d(\mathbf{a,c}) \geq d(\mathbf{a,b}) + d(\mathbf{b,c})$

How to do it?

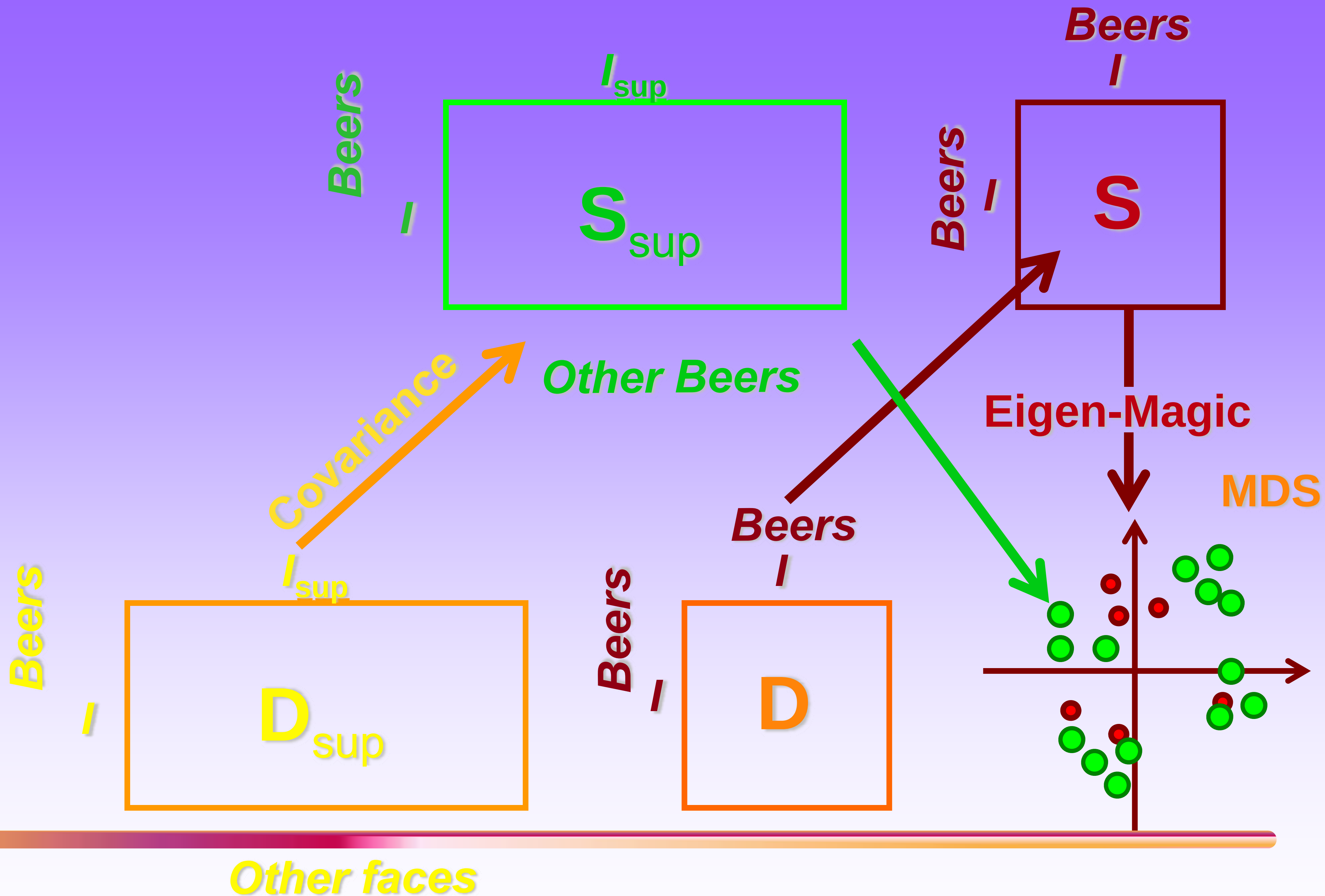
A PCA to MDS refresher



From PCA to MDS: 1 PCA ...



... From (metric) MDS to PCA ... to MDS



… Projecting Supplementary Elements for MDS

I

I **D** $\xrightarrow{\text{Double Centering}}$ I **S**

Distance

With a formula:

$$s_{i,j} = d_{i,j} - (d_{i,+} - d_{+,+}) - (d_{+,j} - d_{+,+})$$

With matrices:

$$\mathbf{S} = -.5\boldsymbol{\Xi}\mathbf{D}\boldsymbol{\Xi}^{\mathsf{T}} \text{ with } \boldsymbol{\Xi} = \mathbf{I} - \mathbf{1}\mathbf{m}^{\mathsf{T}} \text{ and } \mathbf{m}^{\mathsf{T}}\mathbf{1} = 1$$

Transforms the Distances into Covariance

Whites, Rosés, and Reds
An example of MDS plus projection

Cultural Interlude and Wine Break:

Connaisseurs & Novices

(Data courtoisie of Dominique Valentin)

TASK:

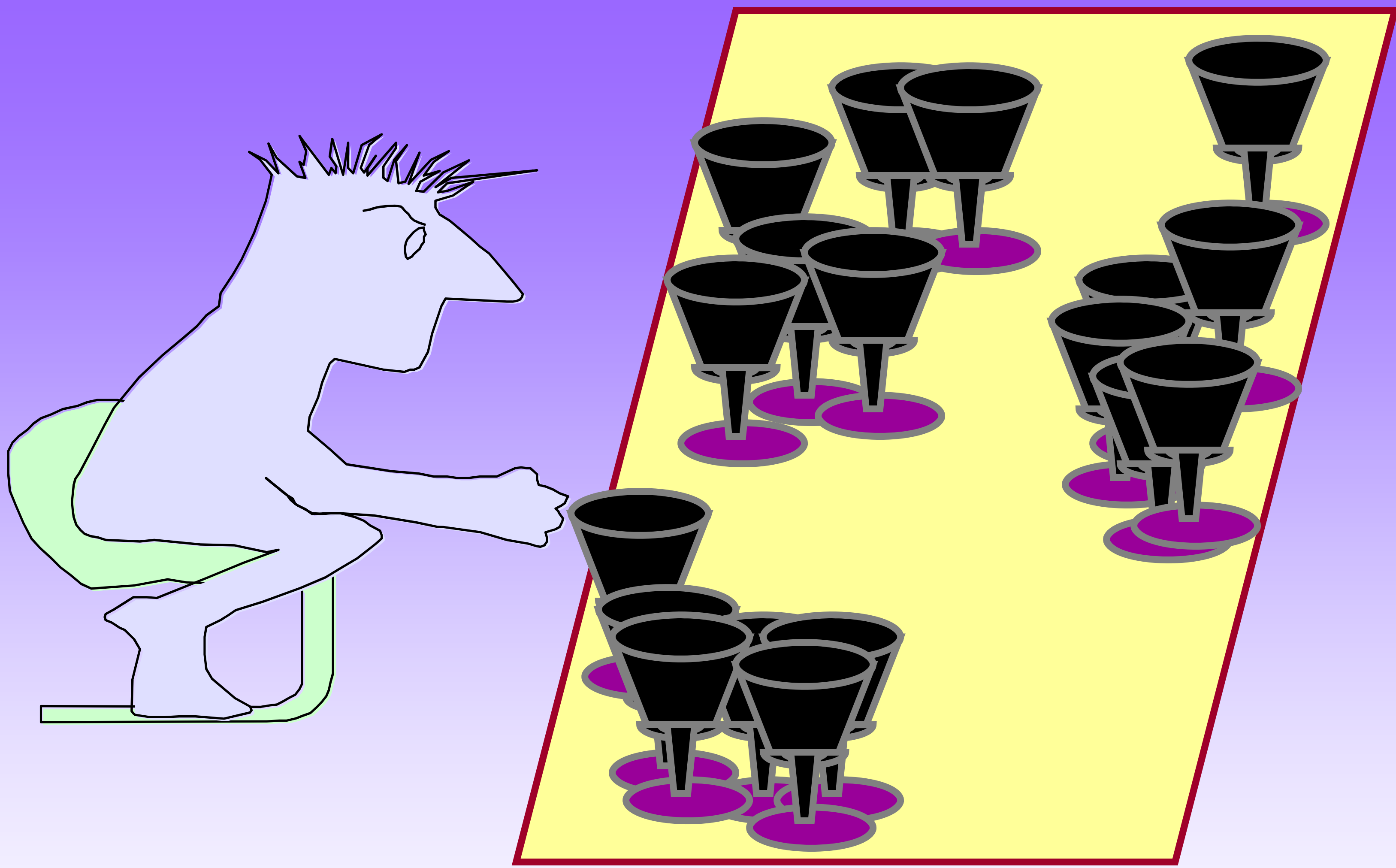
Sort

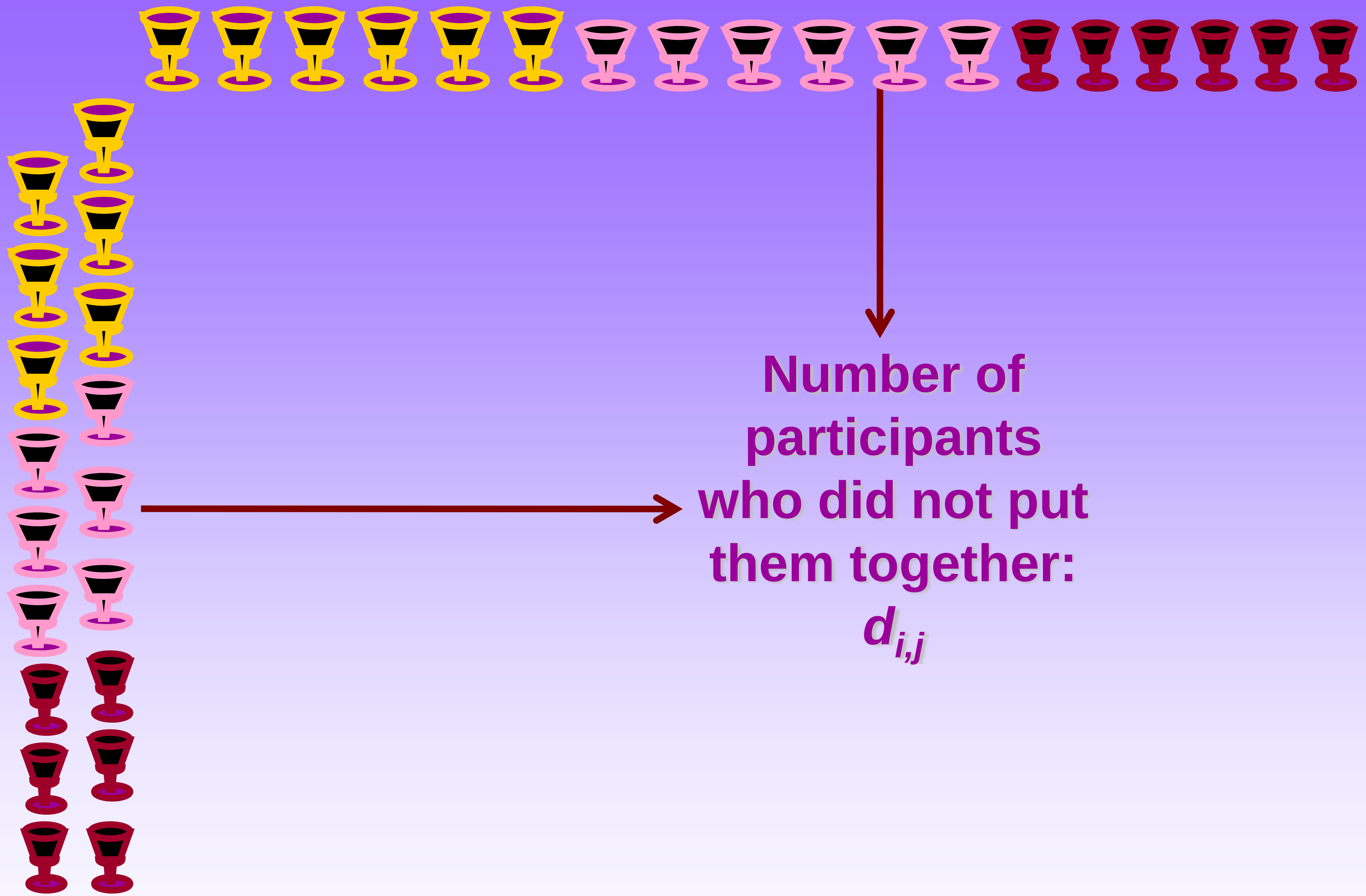
18 Wines

into

3 groups

(Red, White, Rosé)

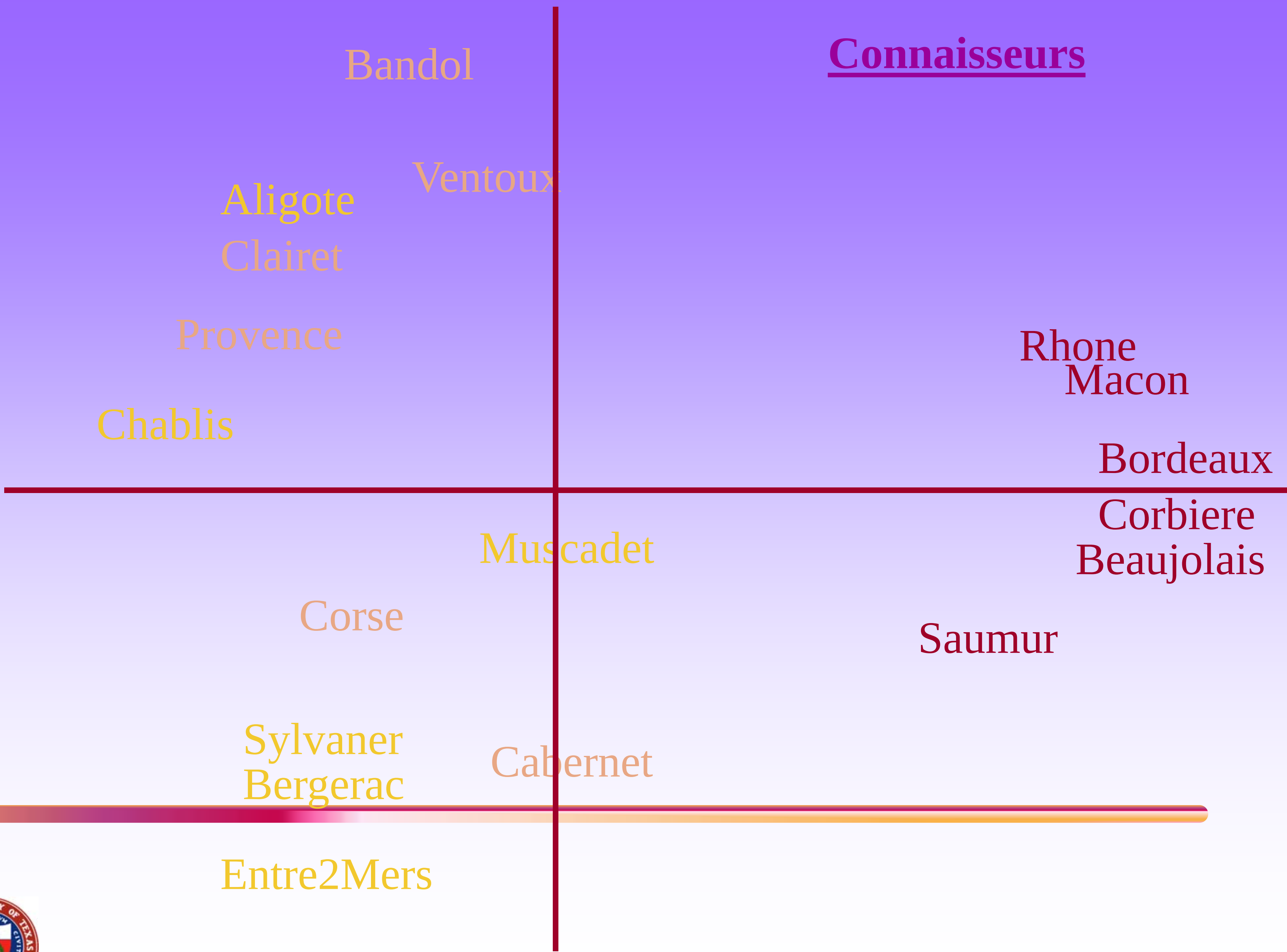




Remember: This is a distance matrix,

A Sensory Evaluation Break

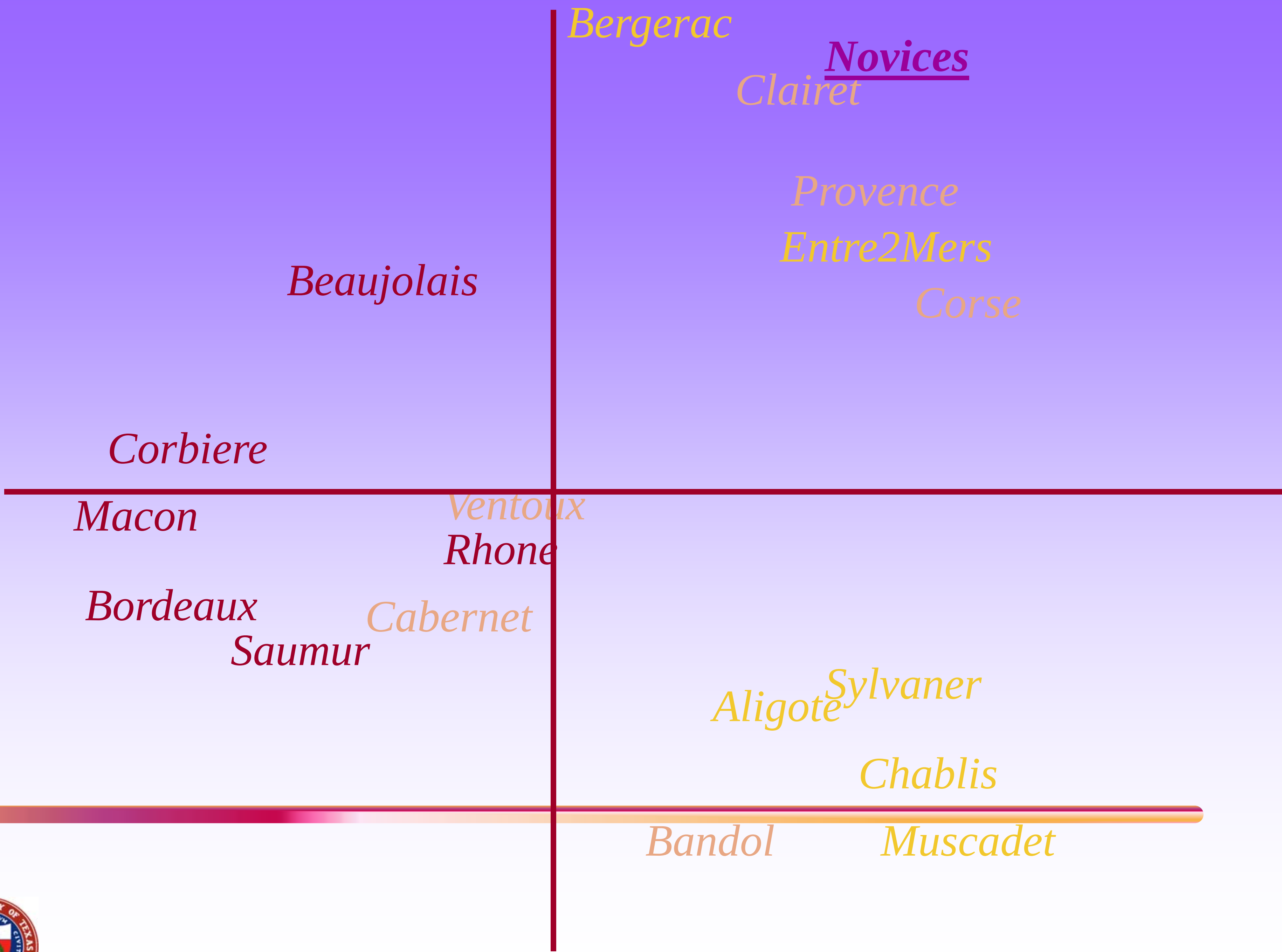
15 “Connaisseurs” (from a questionnaire)


Connaisseurs
Bandol
Ventoux
Aligote
Clairet
Provence
Chablis
Rhone
Macon
Bordeaux
Corbiere
Beaujolais
Muscadet
Corse
Saumur
Sylvaner
Bergerac
Cabernet
Entre2Mers

A Sensory Evaluation Break

9 “Novices”

(from the same questionnaire)

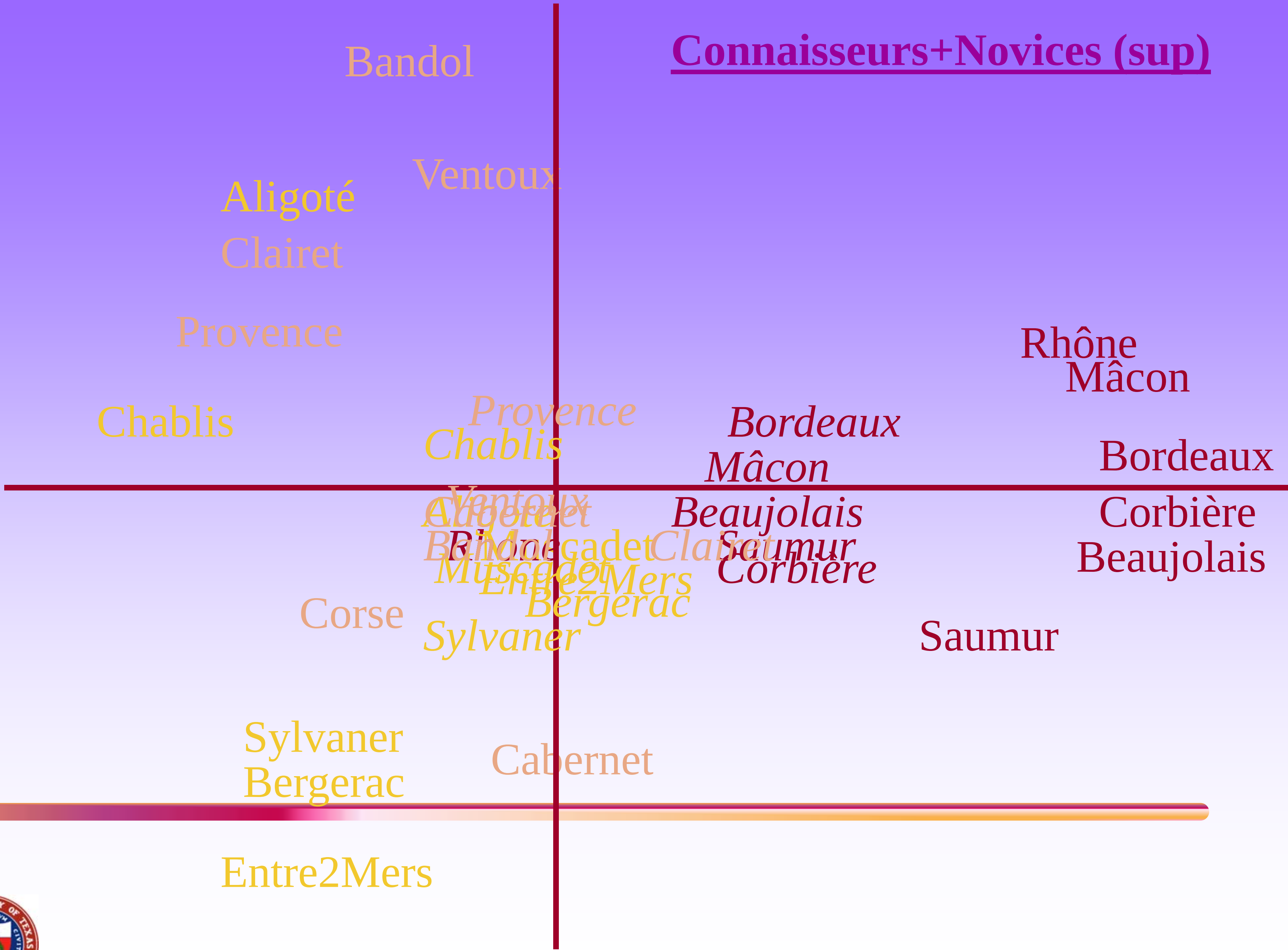

Novices
Bergerac
Clairet
Provence
Entre2Mers
Corse
Beaujolais
Corbiere
Macon
Ventoux
Rhone
Bordeaux
Cabernet
Saumur
Sylvaner
Aligote
Chablis
Bandol
Muscadet

Connaisseurs and Novices

Can we predict the novice space from the connaisseur space?

Projection …

Connaisseurs+Novices (sup)



Connaisseurs and Novices

Unbalanced design!

The novice space seems to collapse

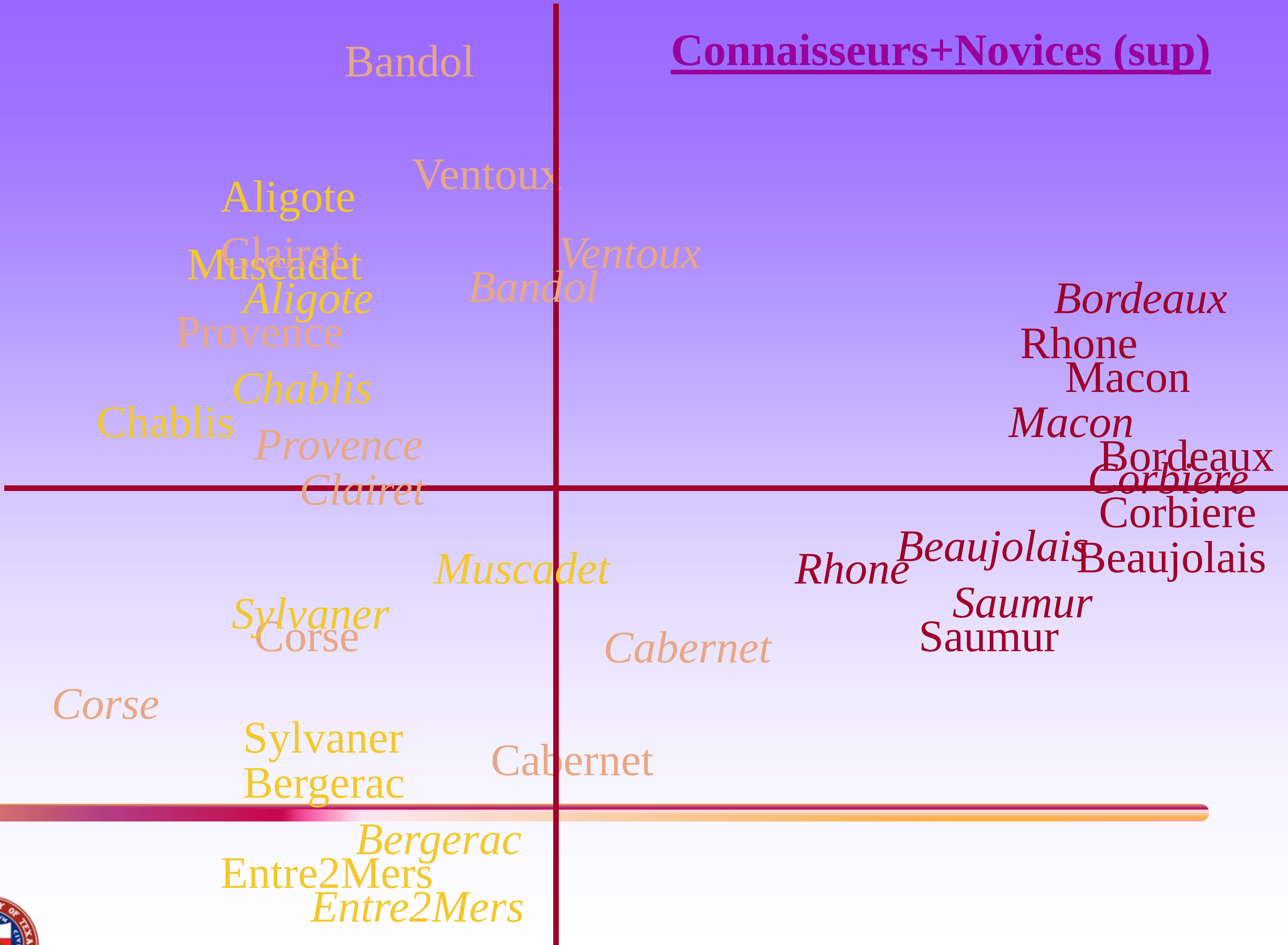
Connaisseurs and Novices

Problem!

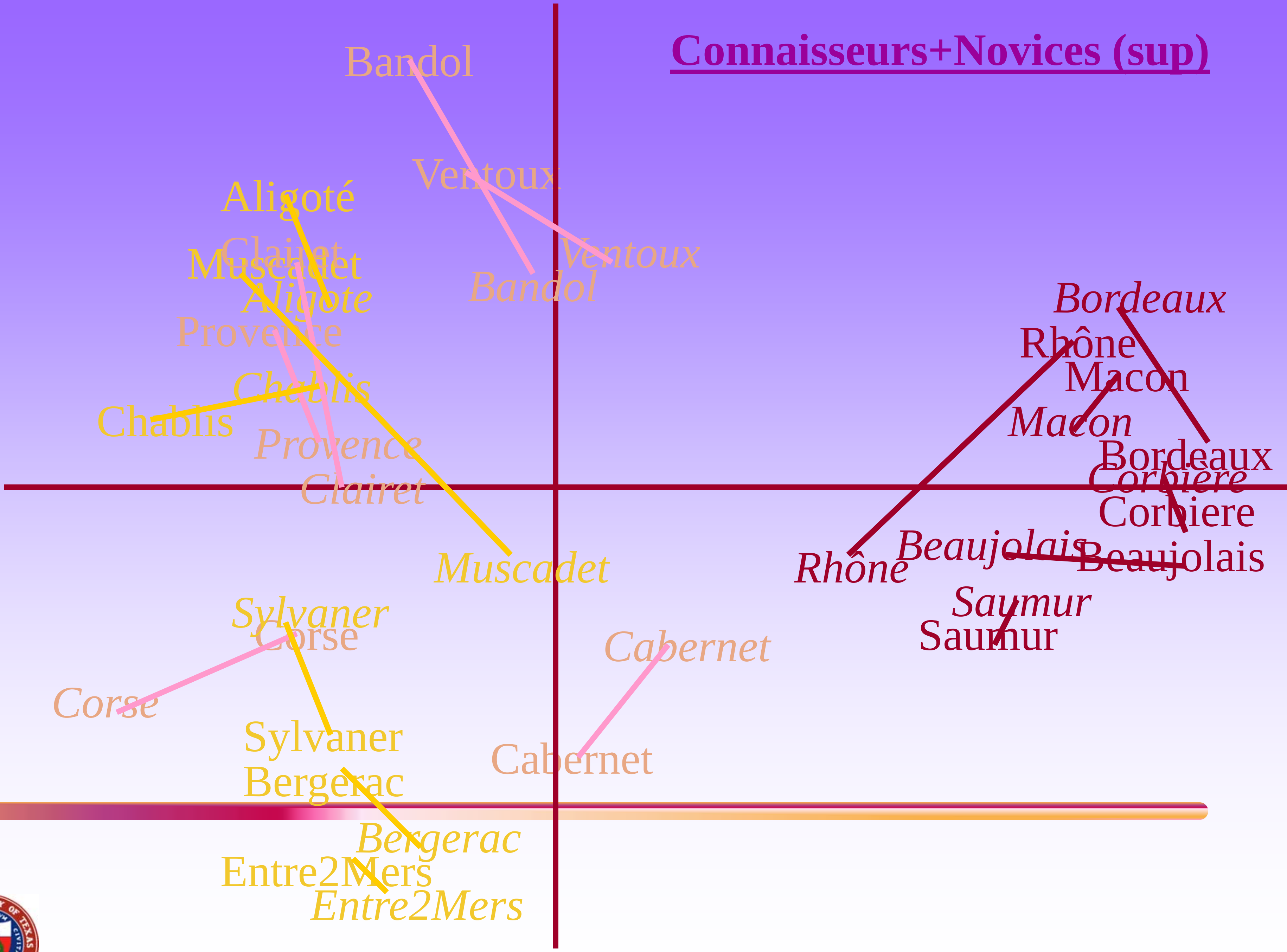
Solution: Normalize the space
(kind of *Z*-scores).

For stat-buffs: Divide by the first eigenvalue
(*cf.* Multiple Factor Analysis)

Connaisseurs+Novices (sup)



Connaisseurs+Novices (sup)



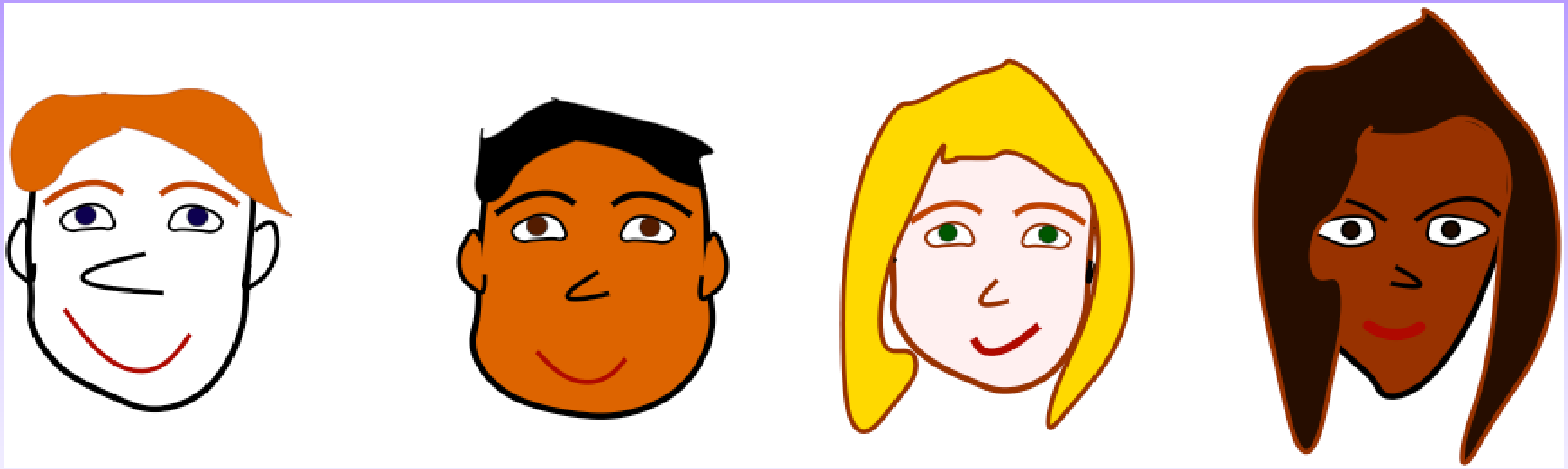
So What do we know as this point?

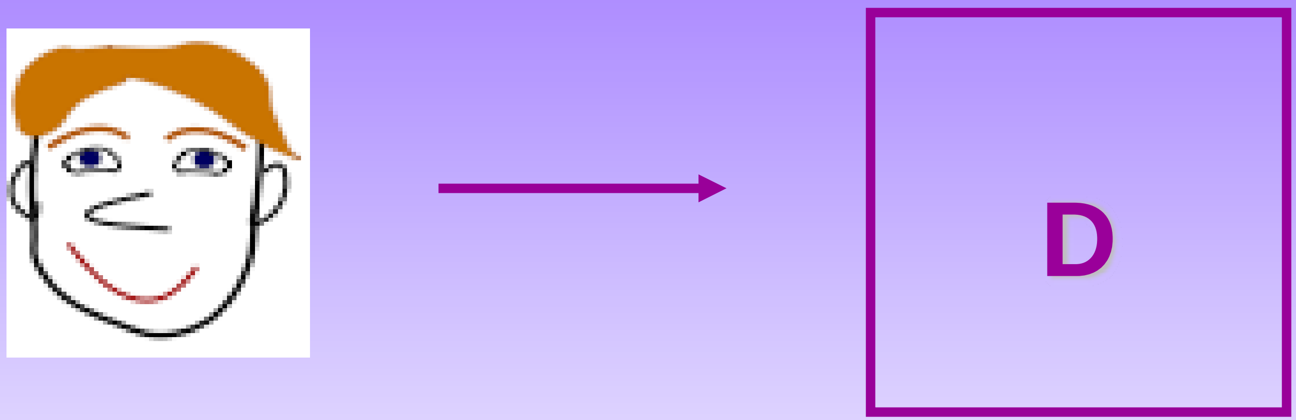
1. How to transform a distance into a cross product matrix (covariance)
2. How to normalize covariance matrices (kind of Z-scores)
3. How to project illustrative/supplementary observations

WE ARE NOW READY FOR DISTATIS!

HOW TO ANALYZE A CUBE OF DISTANCE

A PICTURE TOUR OF DISTATIS

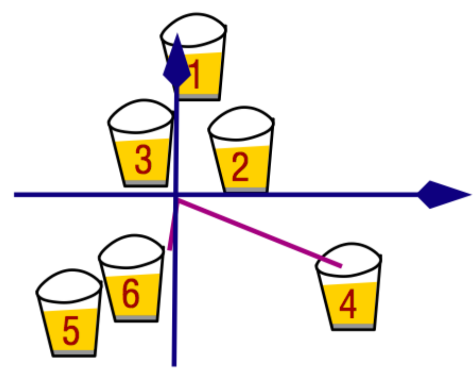




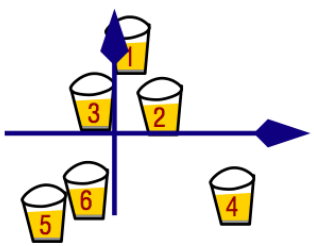
ONE ASSESSOR GIVES A DISTANCE MATRIX

DISTANCE	COVARIANCE	NORMED COVARIANCE
D	$\tilde{S}$	S

$D \rightarrow \tilde{S} \rightarrow S$

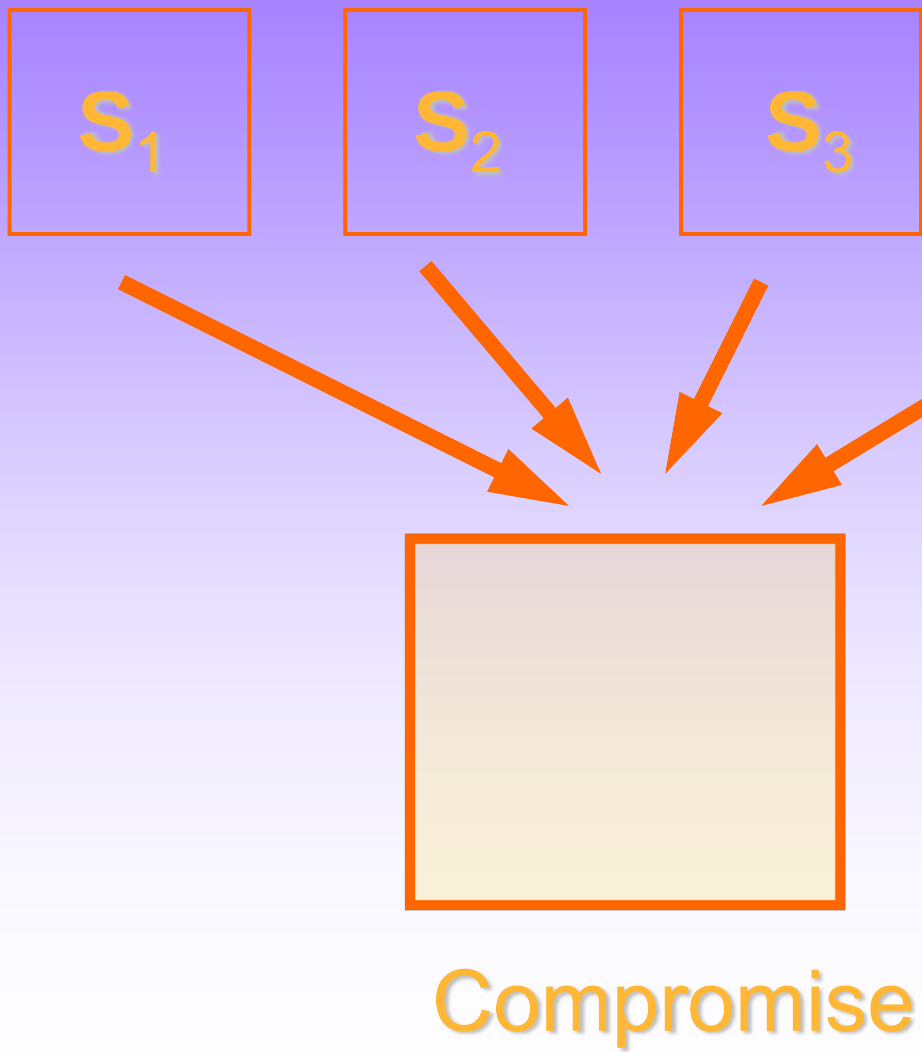


2



3

We have one S matrix per participant
What to do with these covariance matrices?
Find the best *compromise*



- To find the compromise:

  Mix them up!

- This means: find an optimal linear combination of the **S** matrices
- *Optimal*?
- Studies close to a common pattern should be weighted the most

How to find a compromise

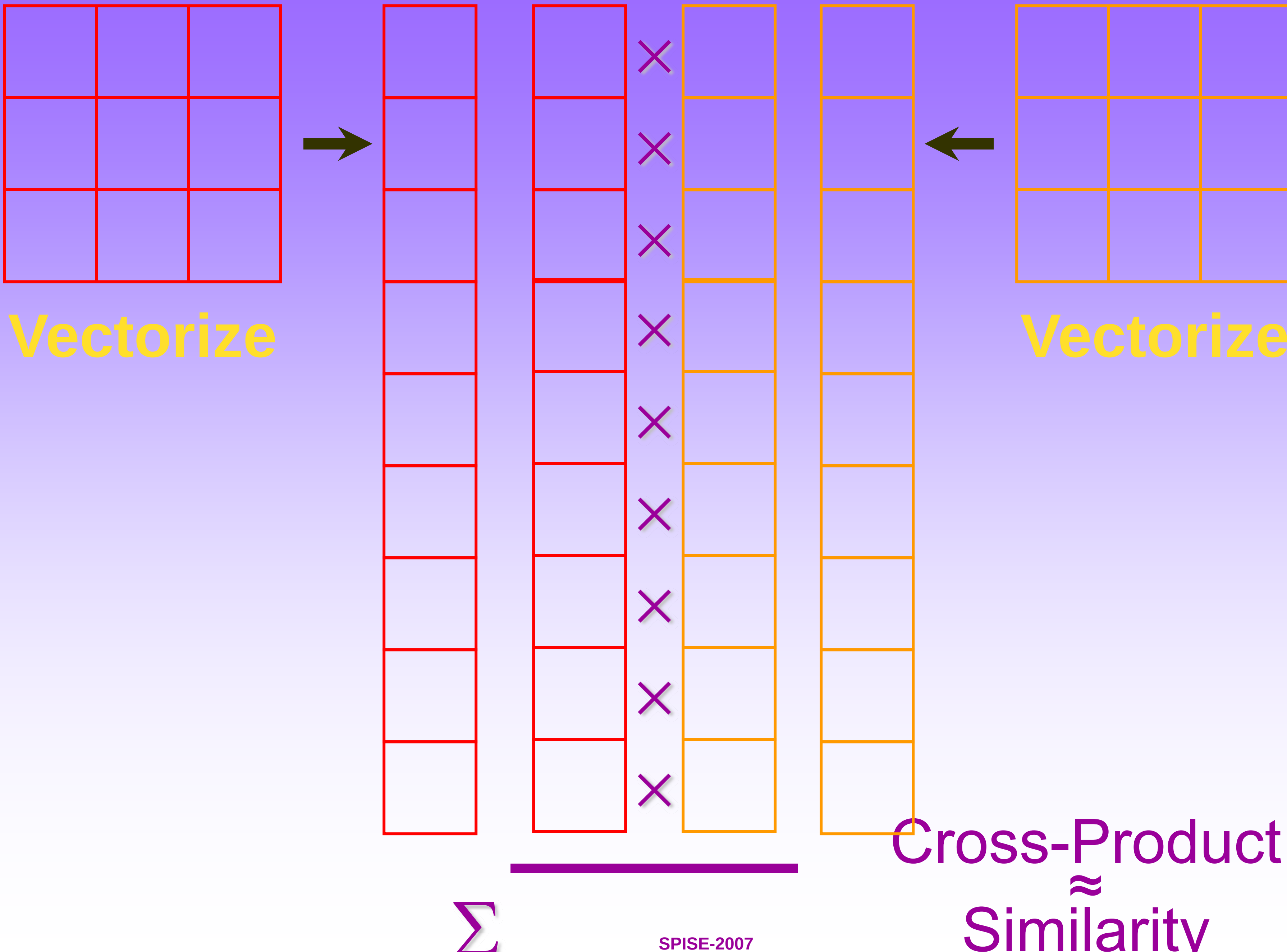
- Look at the between respondent structure:
- Need a coefficient of similarity between the **S** matrices

(for "*connaisseurs*": semi-positive matrices)

- This is the R_V coefficient

How to find the compromise?

Evaluating the Similarity of Two Matrices with R_V



$$\text{Cross-Product} = \sum x_i \times y_i$$

Problem: Cross-Product has “*x* by *y*” unit

Solution: Normalize!

$\Rightarrow$

Unitless Number: This is R_V

The R_V Coefficient:

Computation

$$R_V = \frac{\sum x_i \times y_i}{\sqrt{\left(\sum x_i^2\right) \times \left(\sum y_i^2\right)}}$$

R_V is a "cosine" between "vectorized" matrices.

Kind of correlation.

Tricky part: here it is a squared cosine $0 \le R_V \le 1$

R_V Formally:
X and Y are 2 positive definite matrices

$$R_V = \frac{\mathrm{trace}\{\mathbf{X}^T\mathbf{Y}\}}{\sqrt{[\,\mathrm{trace}\{\mathbf{X}^T\mathbf{X}\}\,] \times [\,\mathrm{trace}\{\mathbf{Y}^T\mathbf{Y}\}\,]}}$$

R_V is a "cosine" between psd matrices.

psd ⇨ $0 \leq R_V \leq 1$

R_V Formally (alternative version)
X and Y are 2 positive definite matrices

$$R_V = \frac{\langle \mathbf{X}^\mathsf{T}\mathbf{Y} \rangle}{|\mathbf{X}^\mathsf{T}\mathbf{X}| \times |\mathbf{Y}^\mathsf{T}\mathbf{Y}|}$$

$\langle , \rangle$: scalar product; $||$: norm

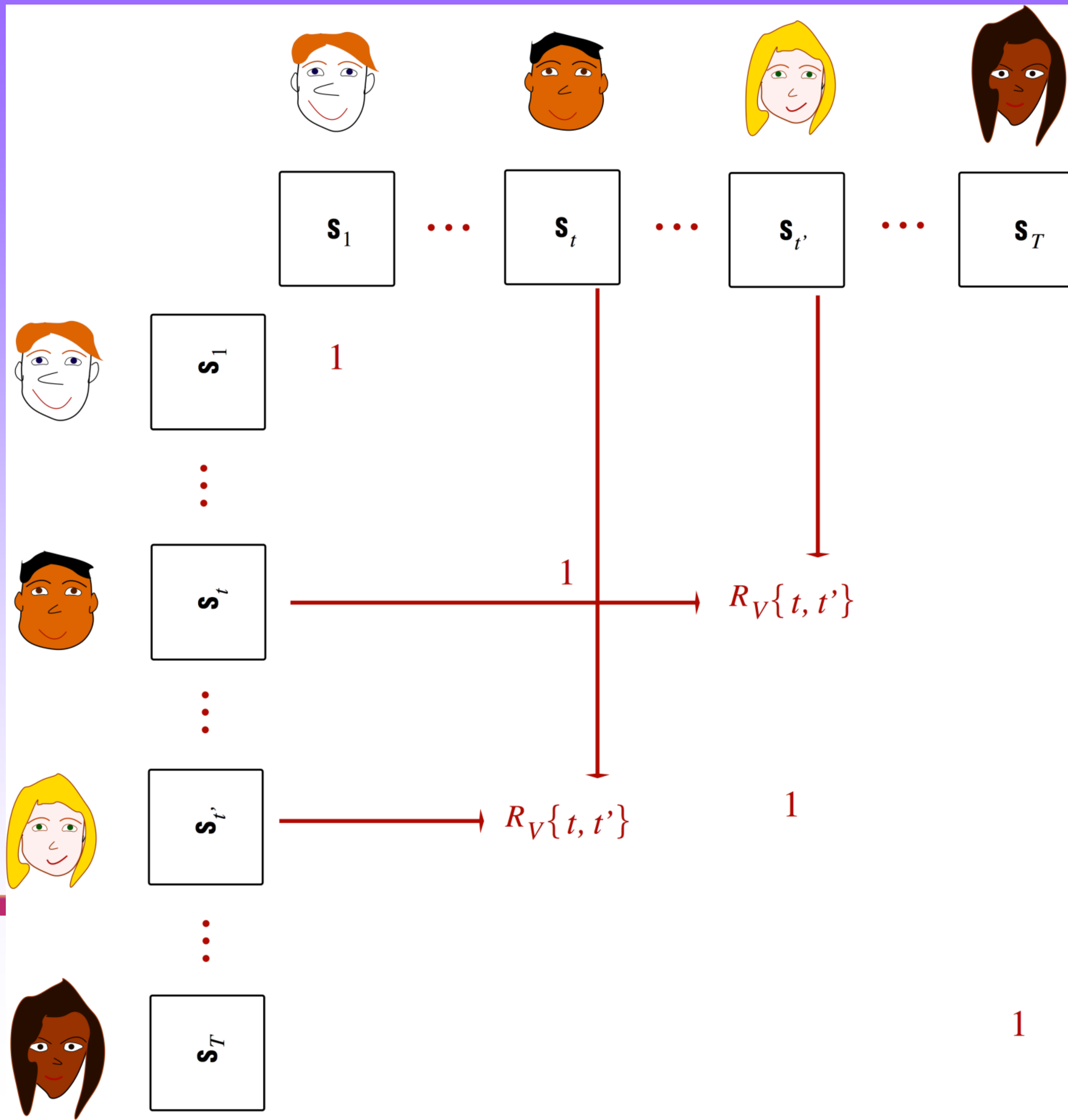
R_V is a "cosine" between psd matrices.

psd ⇨ $0 \leq R_V \leq 1$

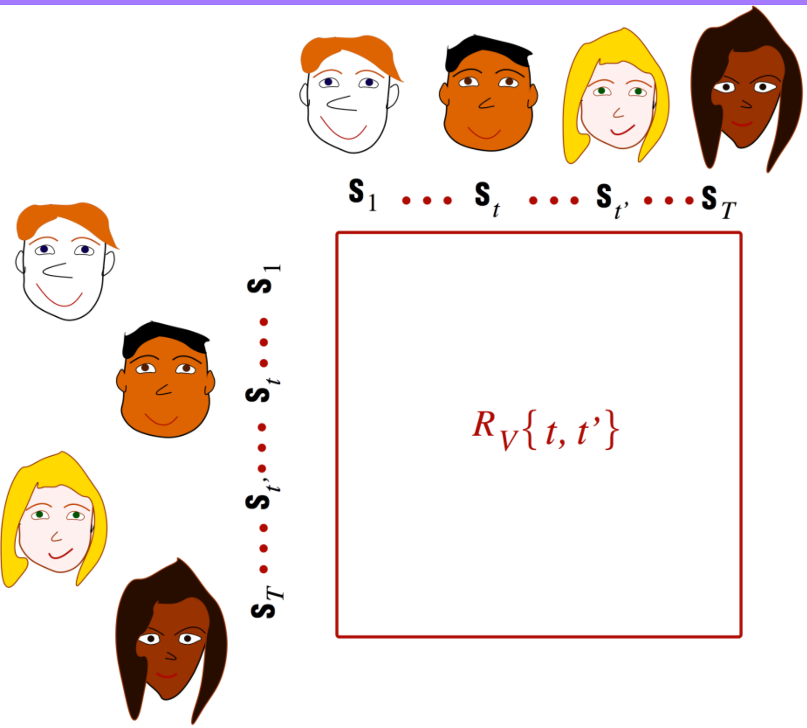
R_V Computational: X and Y are 2 matrices

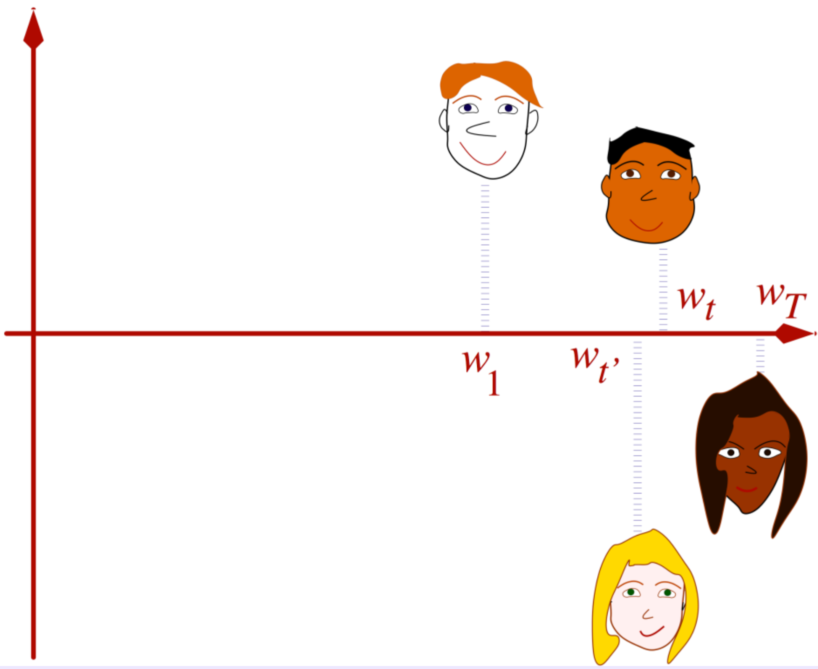
$$R_V = \frac{\mathrm{vec}(\mathbf{X})^{\mathsf{T}}\,\mathrm{vec}(\mathbf{Y})}{\sqrt{\left[\,\mathrm{vec}(\mathbf{X})^{\mathsf{T}}\,\mathrm{vec}(\mathbf{X})\,\right] \times \left[\,\mathrm{vec}(\mathbf{Y})^{\mathsf{T}}\,\mathrm{vec}(\mathbf{Y})\,\right]}}$$

R_V is a “cosine” between vectorized matrices



s_1 … s_t … $s_{t'}$ … s_T
s_1
1
s_t
1
$R_V\{t, t'\}$
$s_{t'}$
$R_V\{t, t'\}$
1
s_T
1





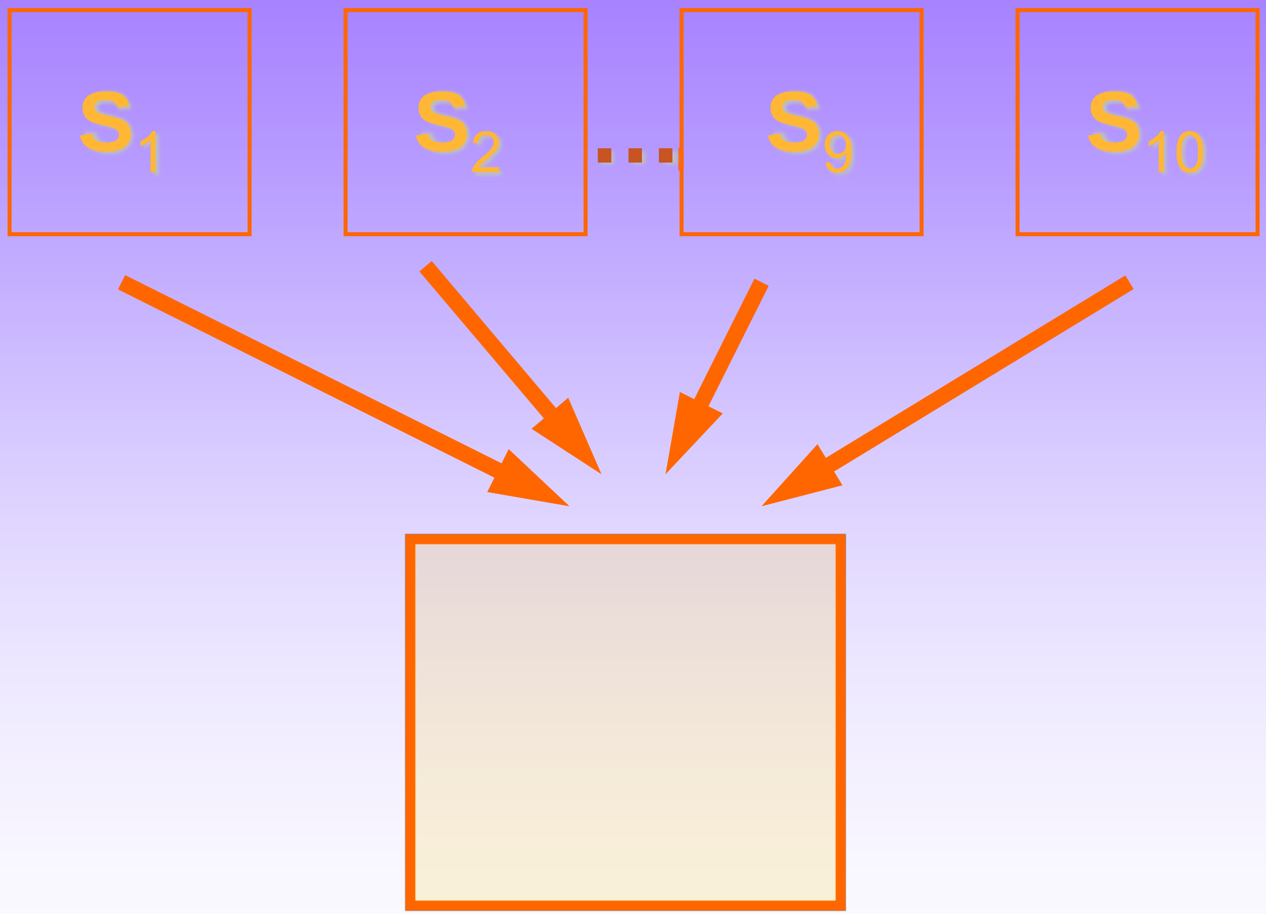
PCA of the participants (R_V matrix)

- A good linear combination has weights summing to 1.
- So: rescale the projections

  into weights with a sum of 1

Transform the projections to get the α_t

Combine the S matrices to get the *compromise*



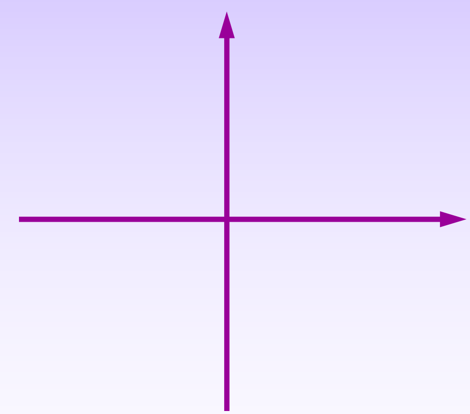
Compromise

Combine the studies: Compute the Compromise

$$\boxed{w_1\mathbf{S}_1} + \boxed{w_2\mathbf{S}_2} + \ldots + \boxed{w_k\mathbf{S}_k} + \ldots + \boxed{w_K\mathbf{S}_K}$$

$$=$$

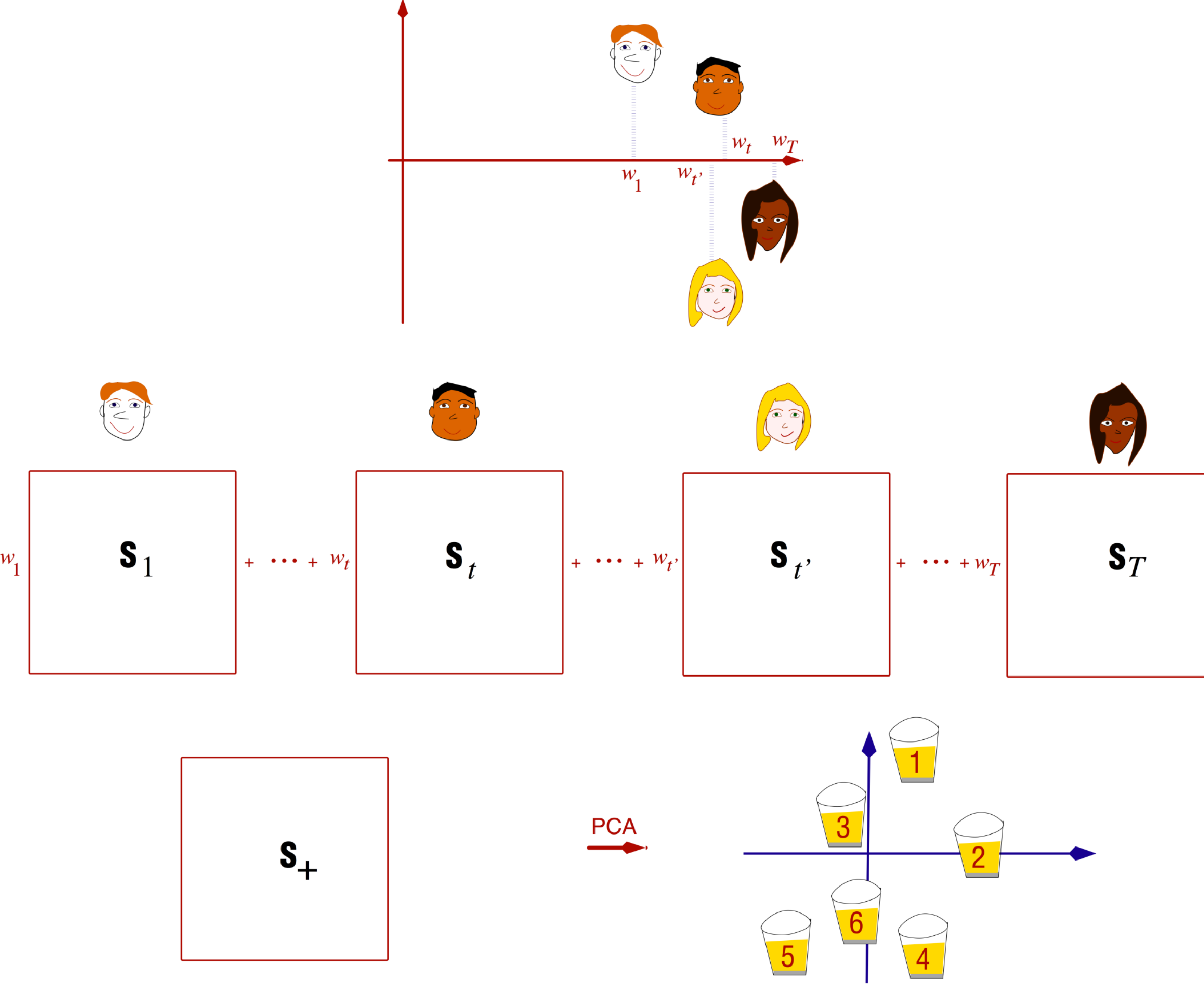
$$\boxed{\mathbf{S}_{[+]}} \rightarrow$$

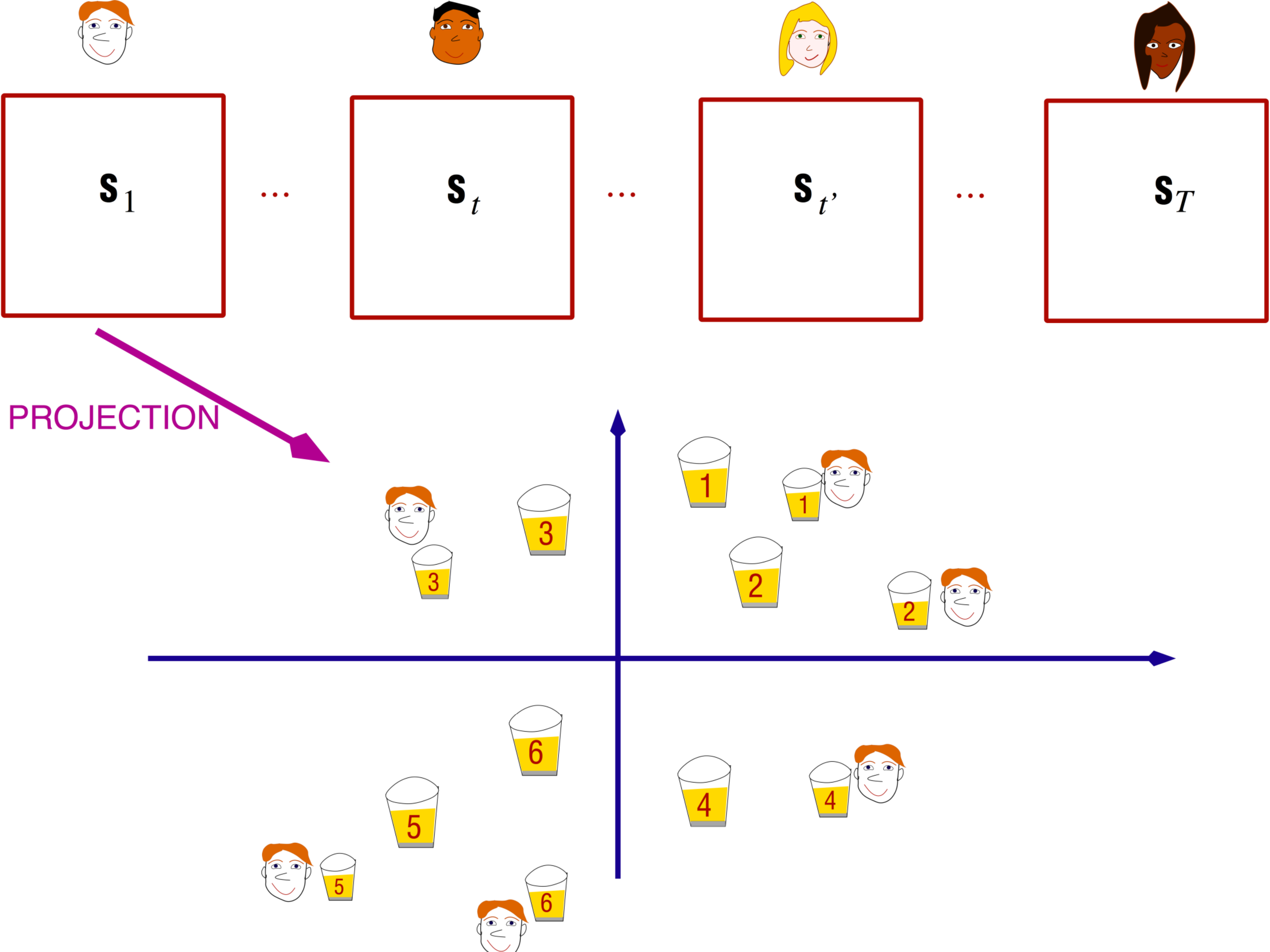


Compromise

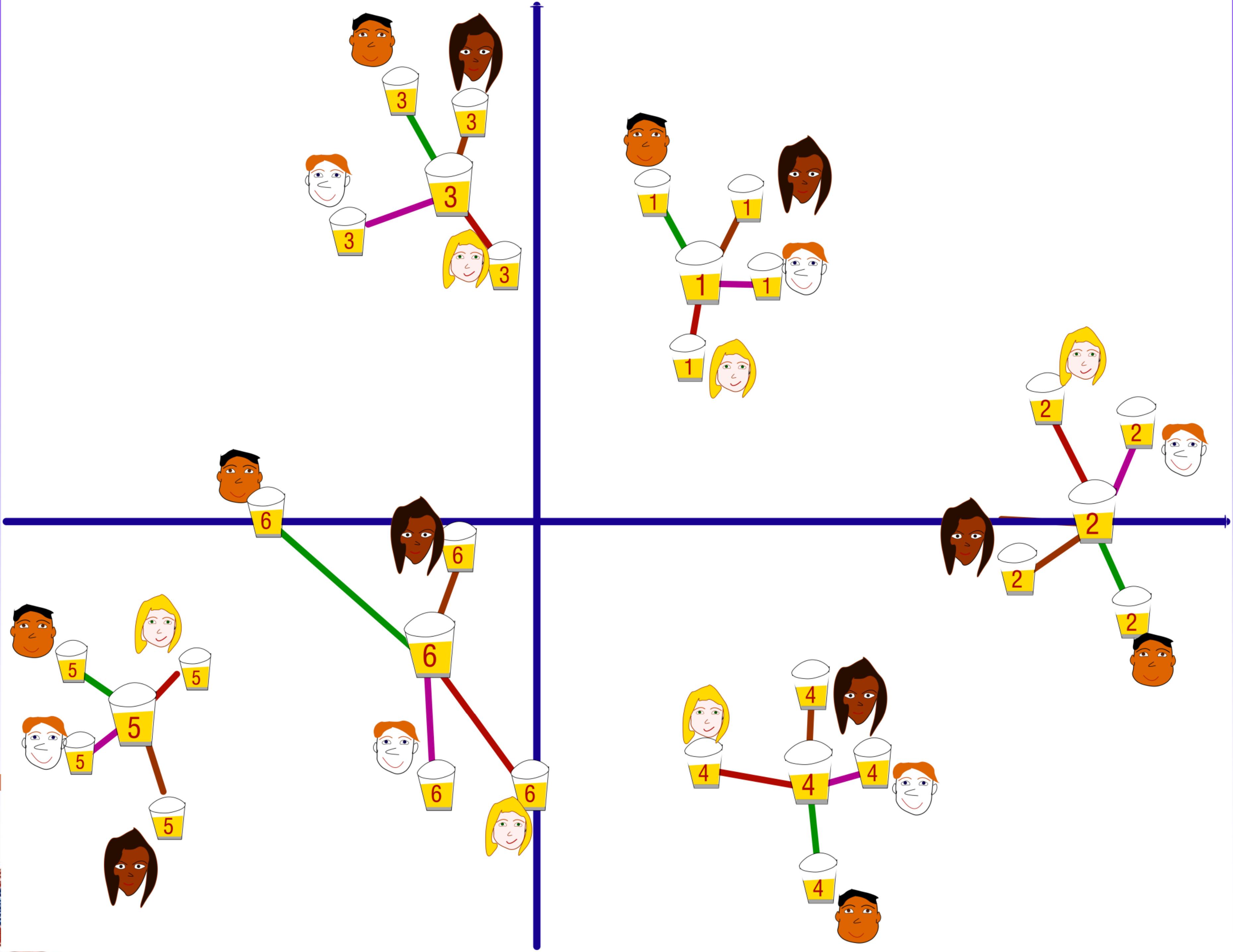
- Quality of the compromise:
- Percentage of explained variance
- Ratio of first eigenvalue to total

How Good is the compromise?


w_1
w_t
$w_{t'}$
w_T
$w_1 \mathbf{S}_1 + \cdots + w_t \mathbf{S}_t + \cdots + w_{t'} \mathbf{S}_{t'} + \cdots + w_T \mathbf{S}_T$
$\mathbf{S}_+$
PCA
1
2
3
4
5
6



$\mathbf{s}_1$
$\mathbf{s}_t$
$\mathbf{s}_{t'}$
$\mathbf{s}_T$
PROJECTION
1
1
3
3
2
2
6
5
4
4
5
6

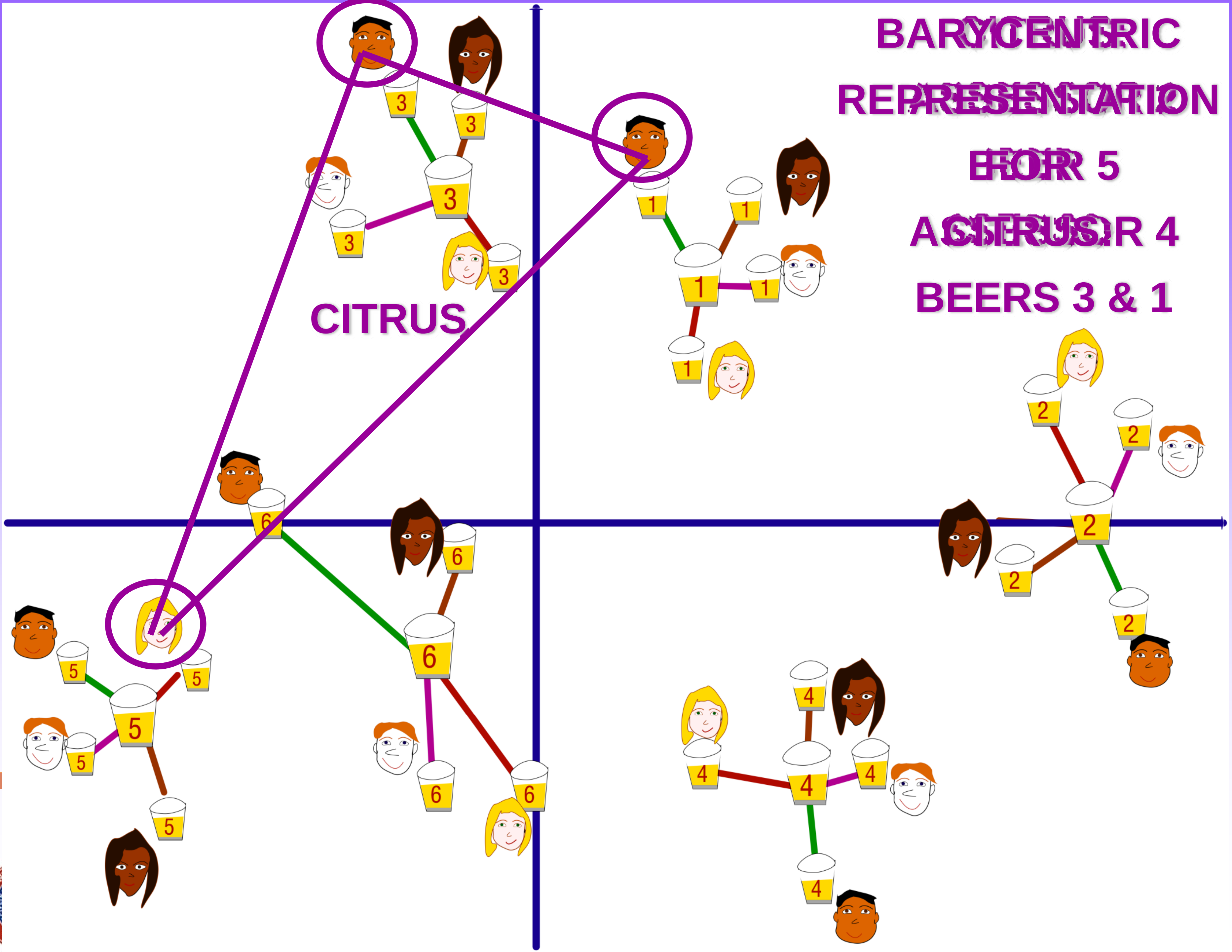

3
3
3
3
3
1
1
1
1
1
2
2
2
2
2
6
6
6
6
6
5
5
5
5
5
4
4
4
4
4

PROJECTING ATTRIBUTES

GOOD OLD MEANS

(CALL THEM BARYCENTERS)

BARYCENTRIC
REPRESENTATION
FOR 5
ASSESSOR 4
CITRUS:
BEERS 3 & 1



THE EARTH IS ROUND: $p < .05$
cf. J. Cohen (1994)

Or the magic of the significance level for publication

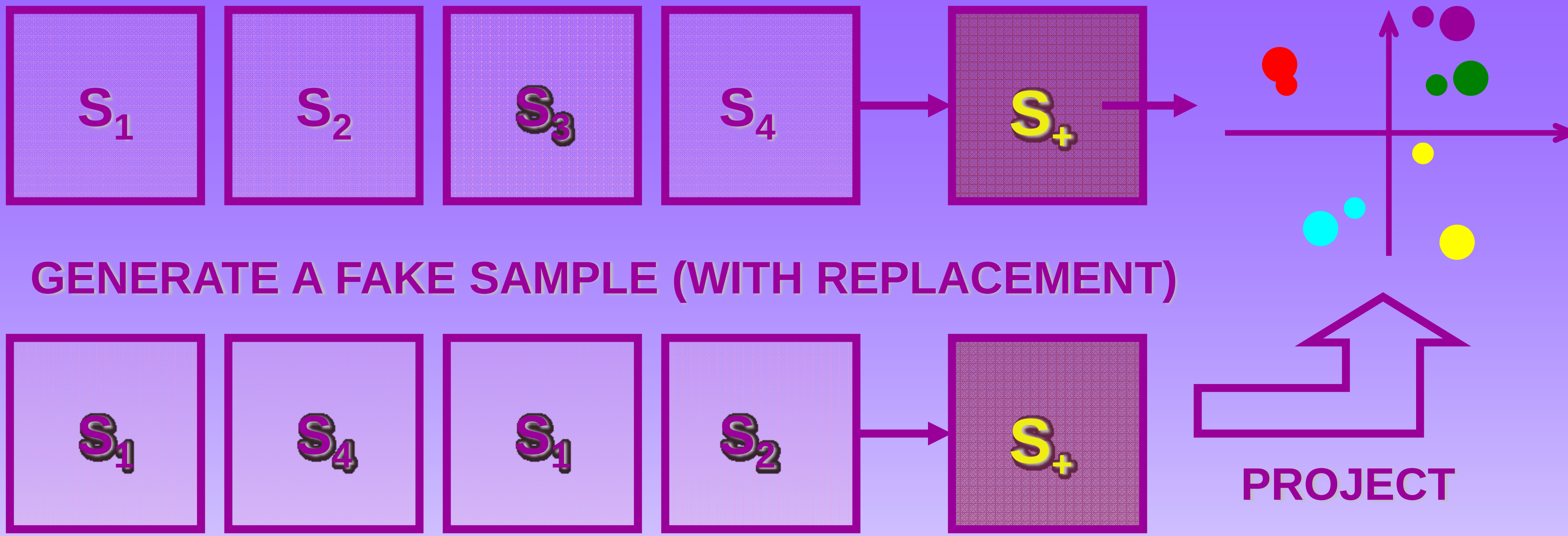
HOW STABLE ARE THE RESULTS?

CONFIDENCE INTERVALS FOR THE PRODUCTS

(SEE ALSO LÊ *et al.*, In FQP for SIMILAR IDEAS)

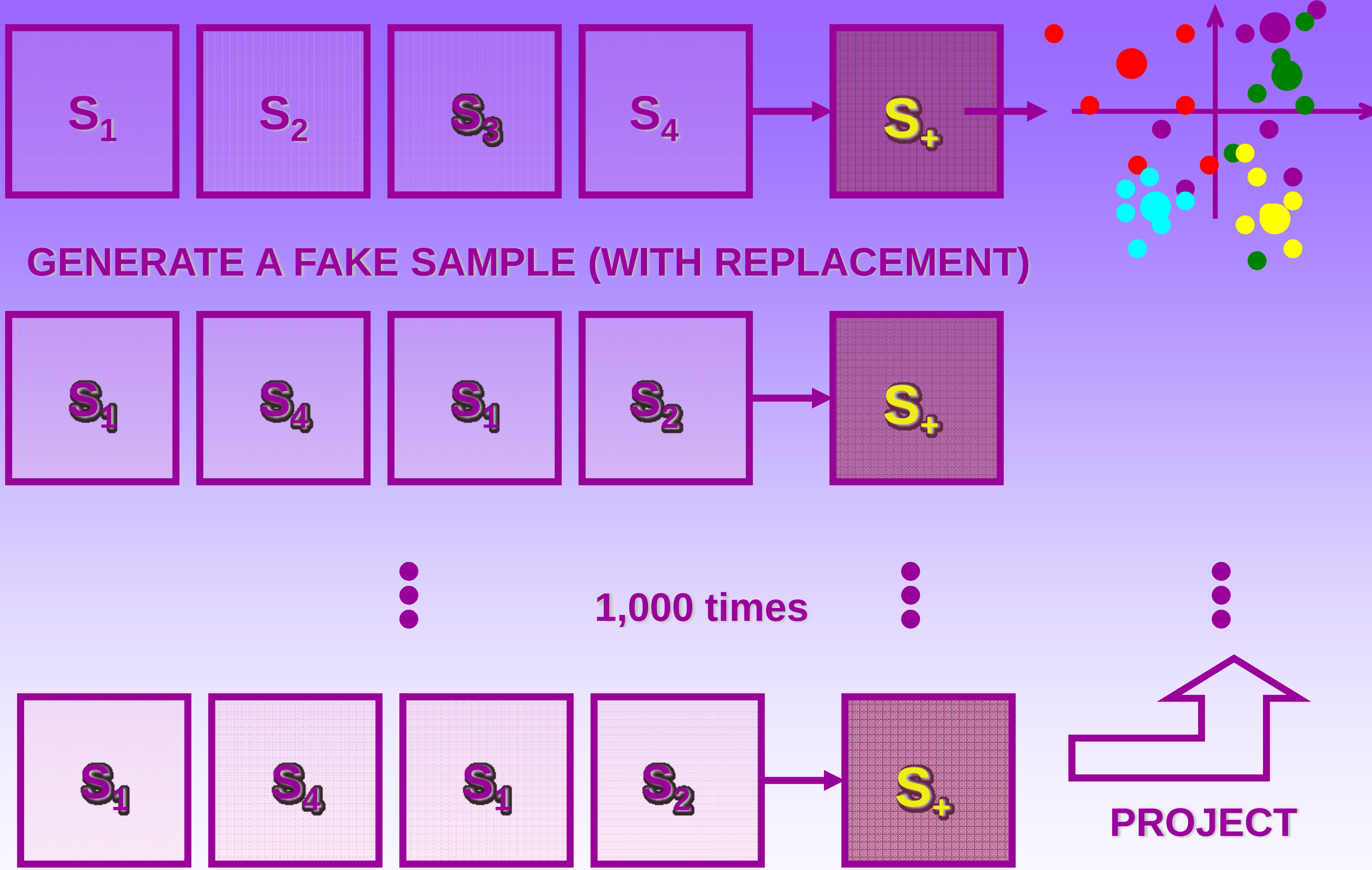
USE CROSS-VALIDATION
TO CREATE CONFIDENCE INTERVALS

BOOTSTRAP HERE

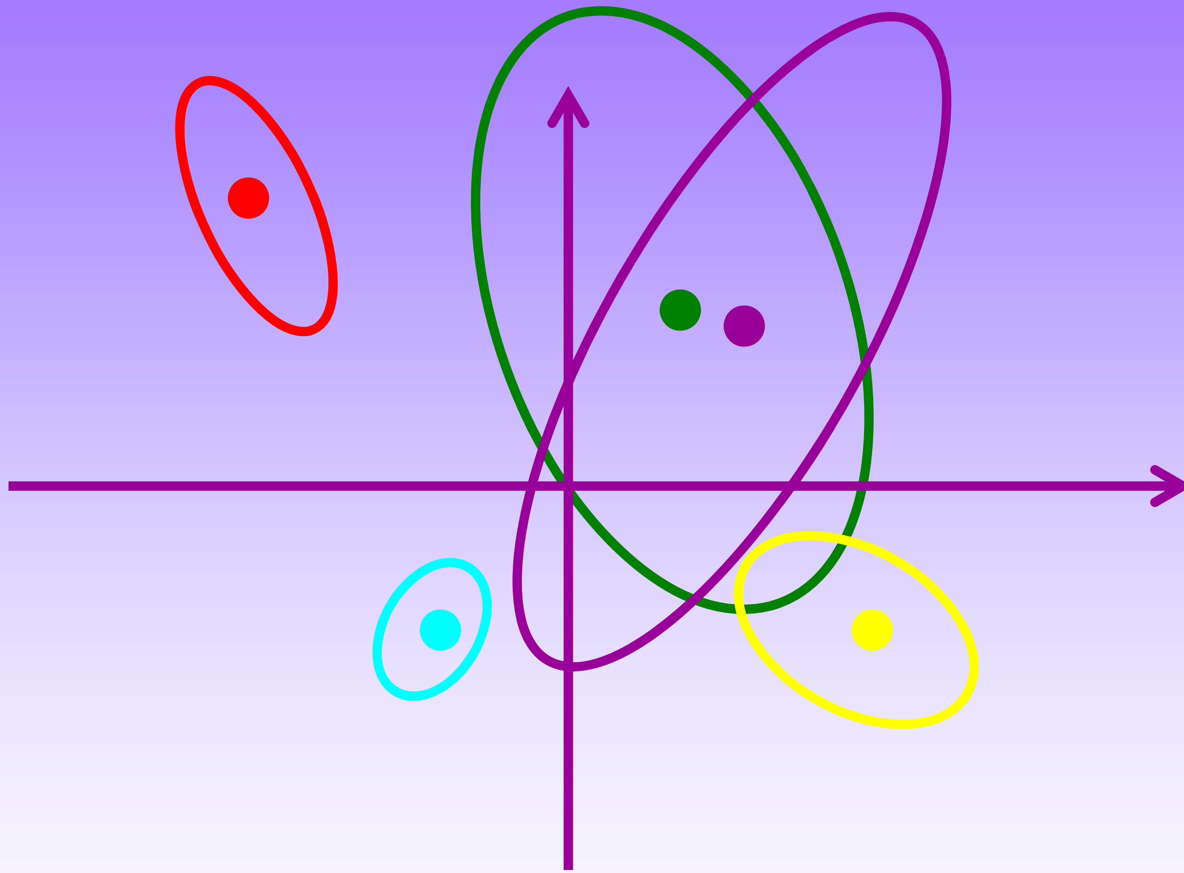


KEEP ON DOING THAT A LOT

BOOTSTRAP


S_1 S_2 S_3 S_4 → S_+
GENERATE A FAKE SAMPLE (WITH REPLACEMENT)
S_1 S_4 S_1 S_2 → S_+
1,000 times
S_1 S_4 S_1 S_2 → S_+
PROJECT

REPLACE THE DOTS BY 95% ELLIPSOIDS



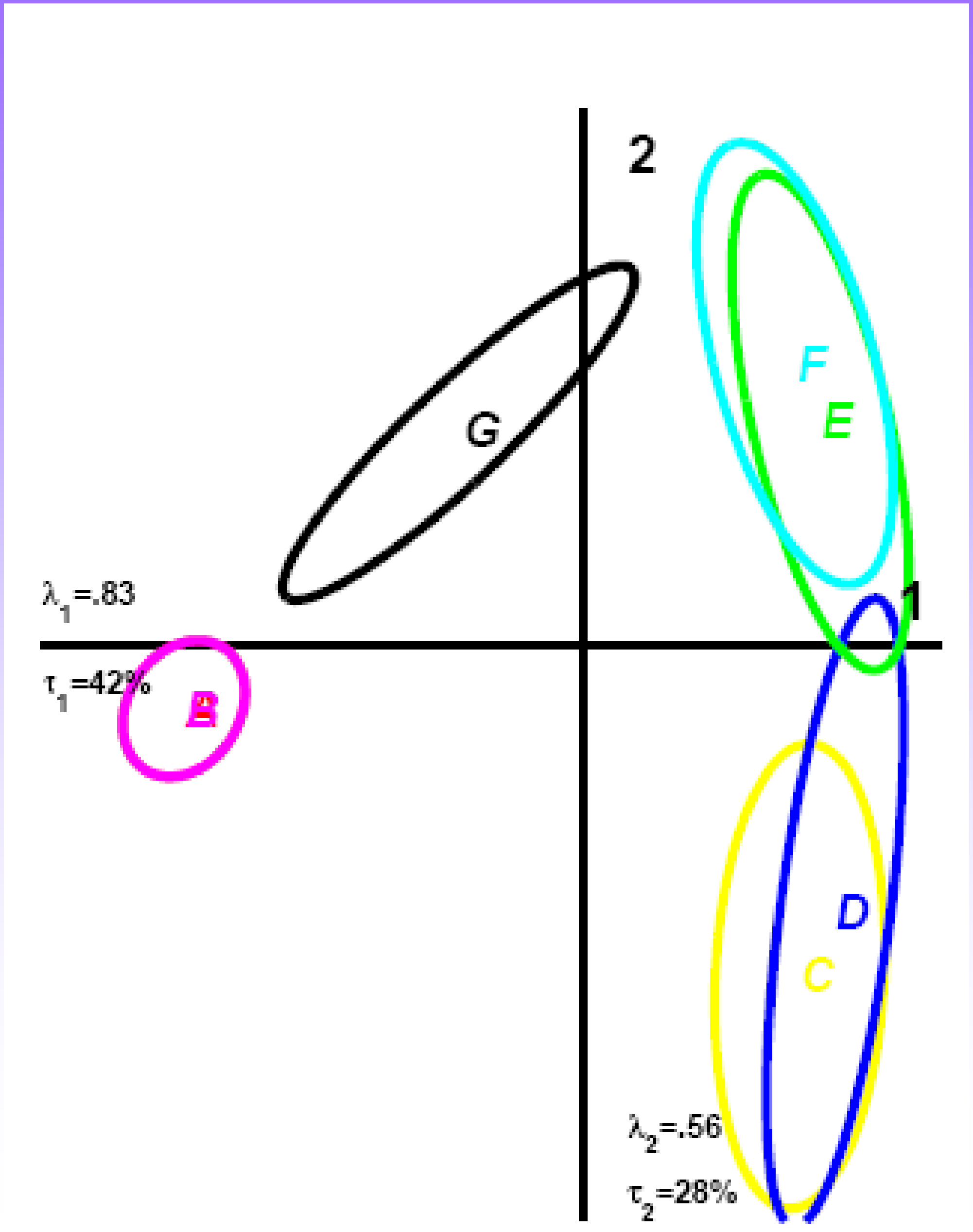
And you have your $p < .05$

ARE YOU STILL AWAKE?

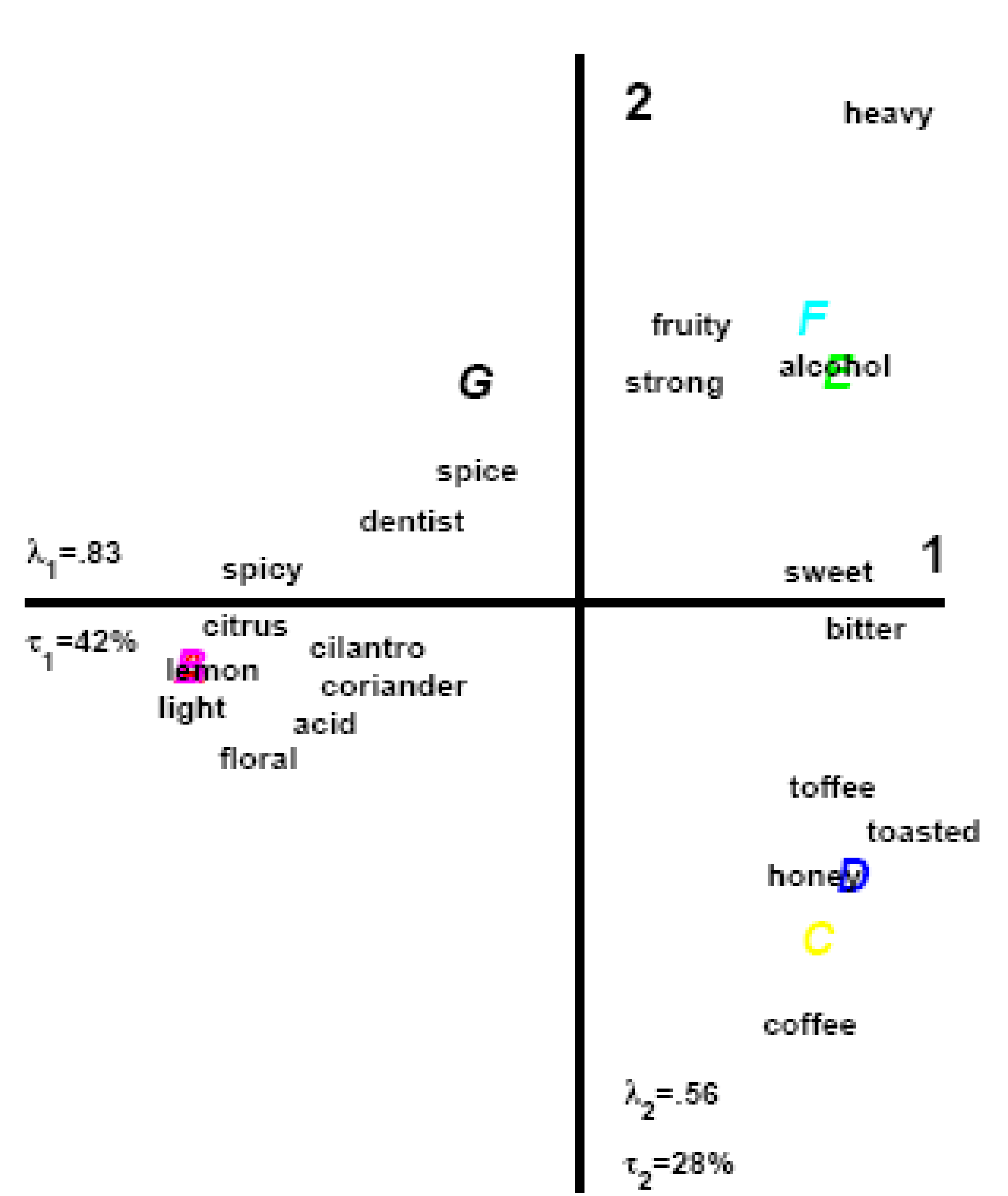
Time To Get Back To The Beers

Remember
the beers, the sorting, the napping, the flashing?

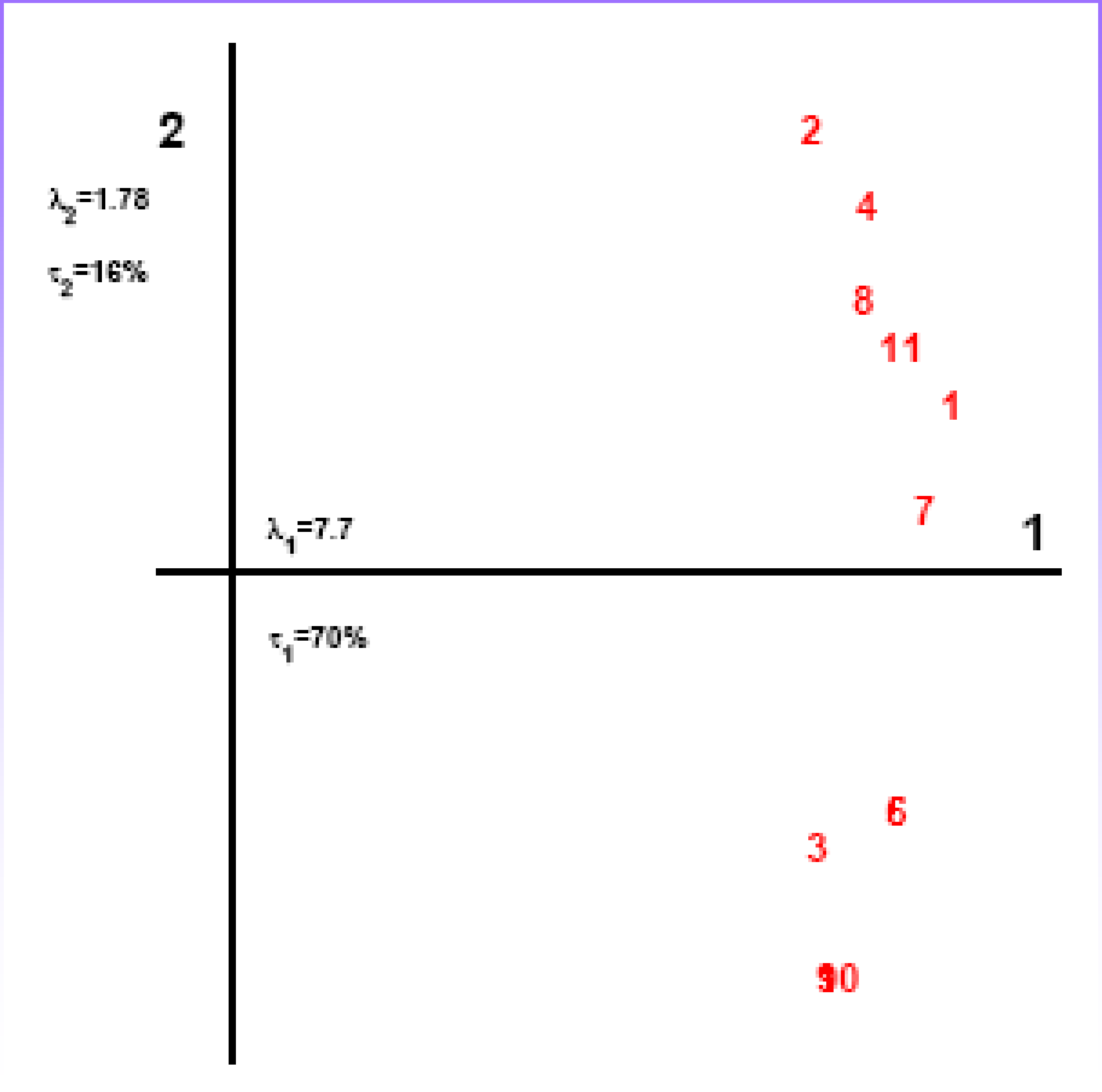
SORTING: THE BEERS WITH CONFIDENCE



SORTING: THE BEERS WITH ATTRIBUTE

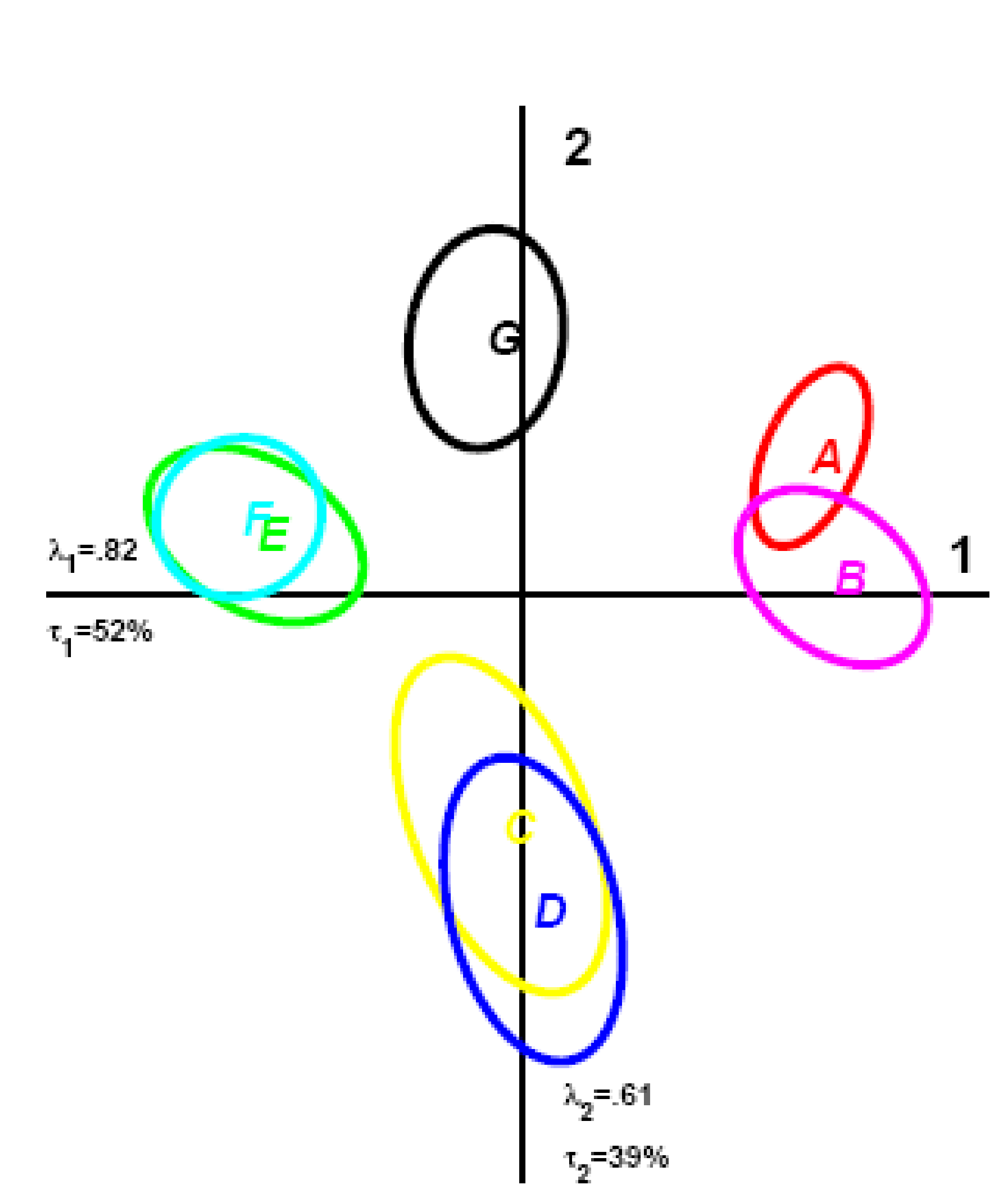


SORTING: THE ASSESSORS

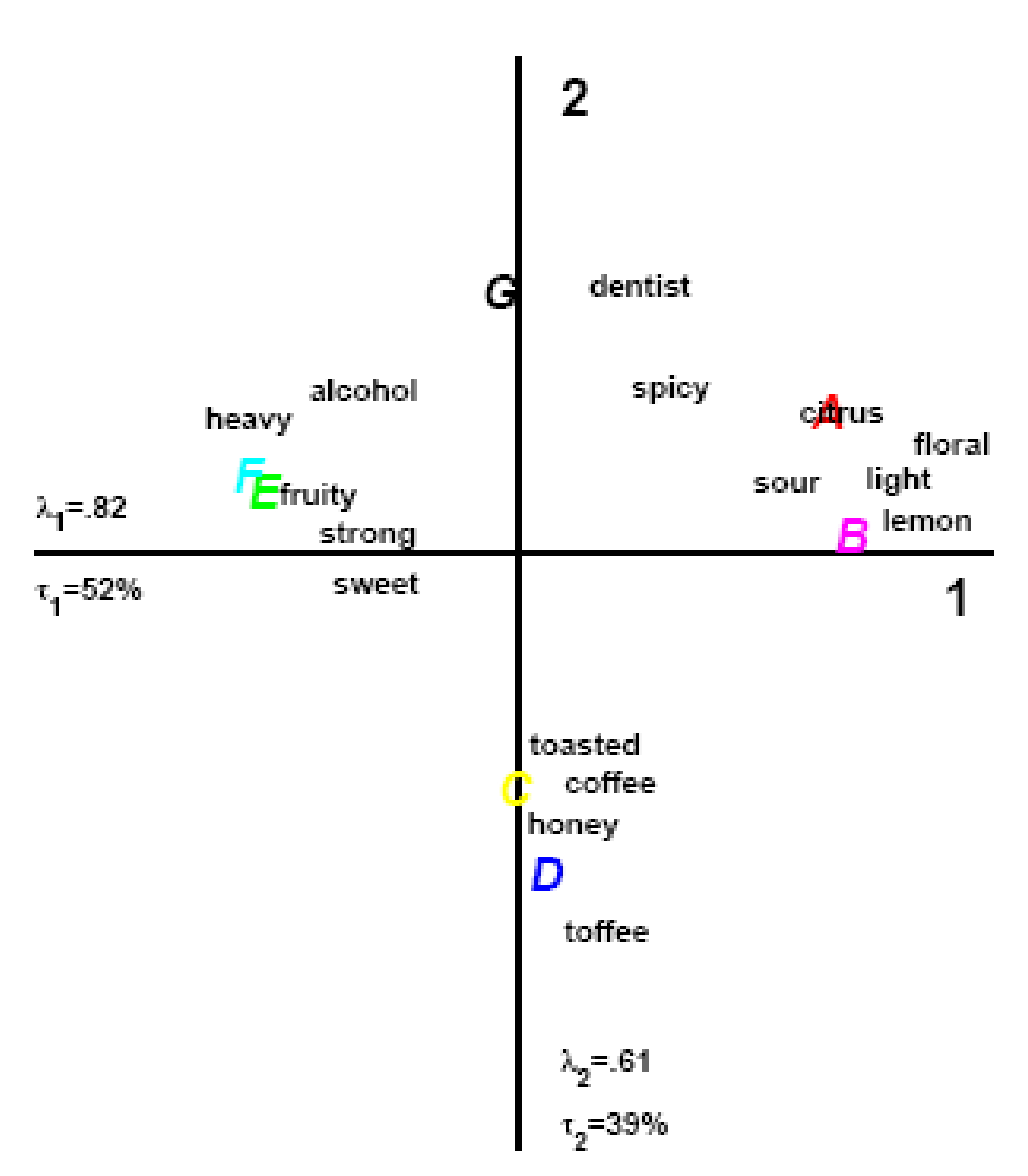


PREFERENCE MAPPING / NAPPING
(7 ASSESSORS SORTED 7 BEERS)

NAPPING: THE BEERS WITH CONFIDENCE



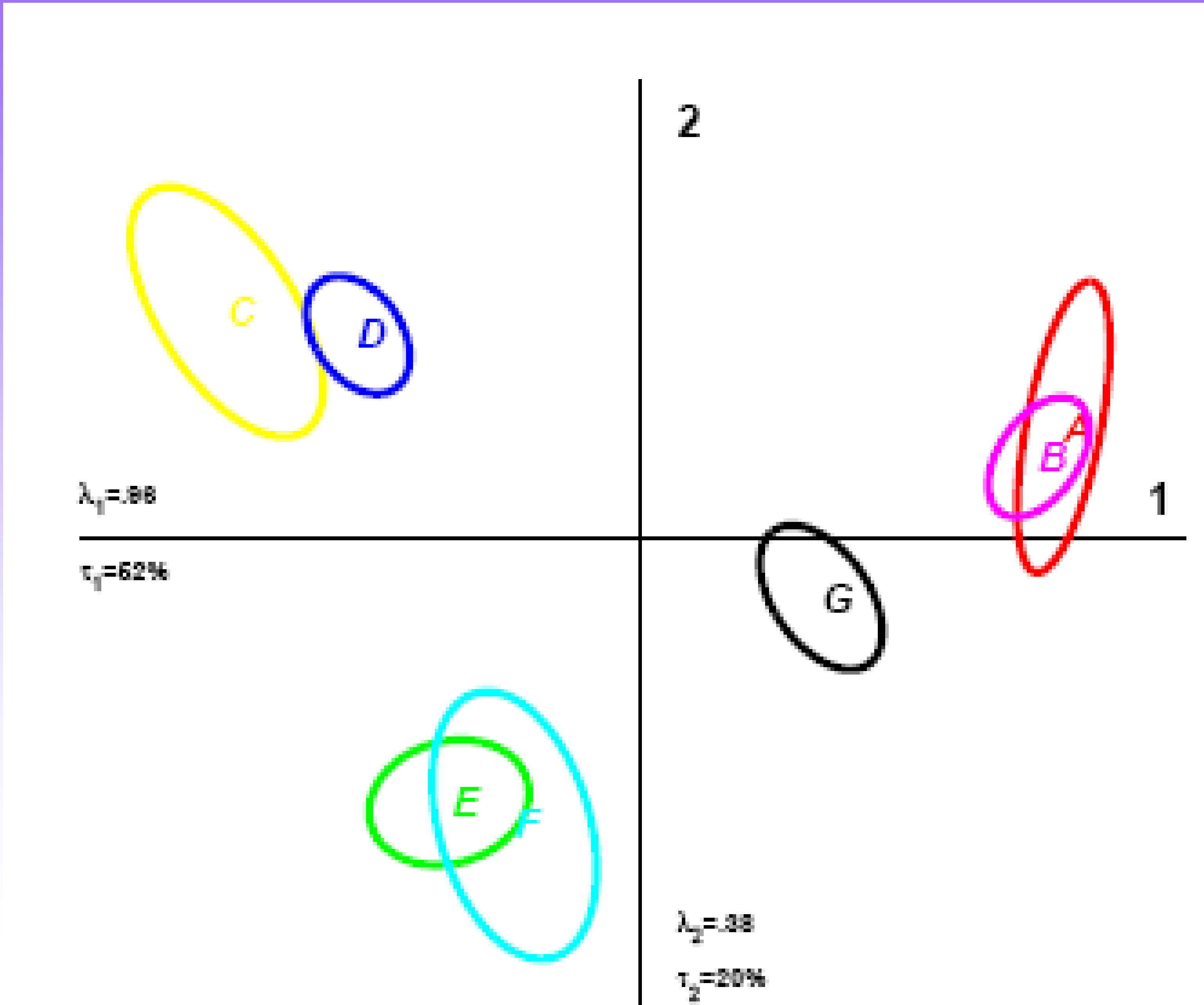
NAPPING: THE BEERS WITH ATTRIBUTE



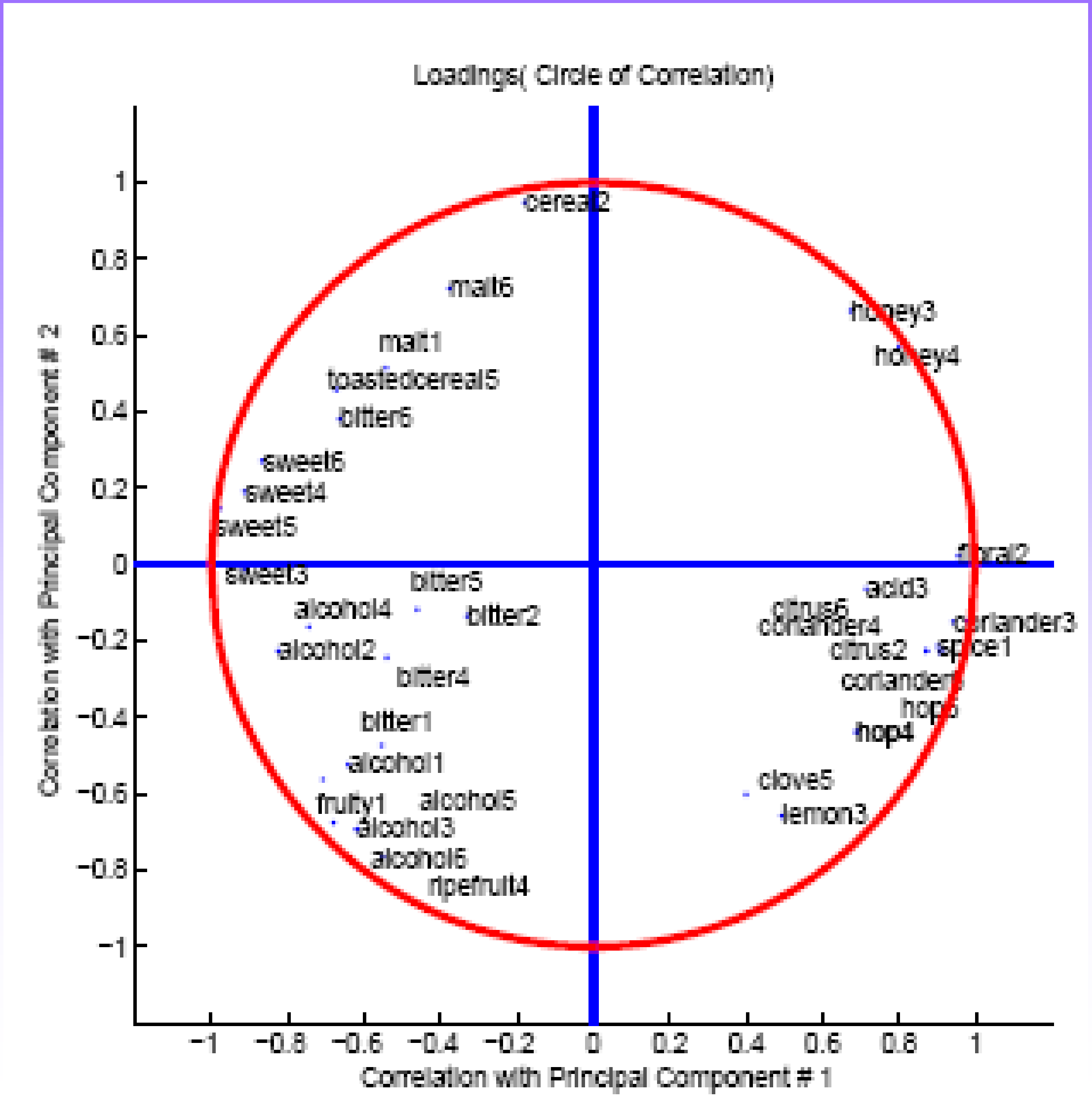
NAPPING: THE ASSESSORS

FLASH PROFILING
(6 ASSESSORS EVALUATED 7 BEERS)

FLASH: THE BEERS WITH CONFIDENCE



FLASH: THE BEERS WITH ATTRIBUTE



FLASH: THE ASSESSORS

TIME TO CONCLUDE

NEW TECHNIQUES FOR PROFILING

FOR THESE TECHNIQUES IT IS IMPORTANT TO HAVE MAPS OF THE ASSESSORS ALSO

HAVING THE SAME TOOL TO ANALYZE WILL MAKE COMPARISONS EASIER

OBVIOUS ALTERNATIVES TO DISTATIS
MULTIPLE FACTOR ANALYSIS
GENERAL PROCRUSTEAN ANALYSIS
STATIS

- Variation of STATIS
- Incorporate features of MFA.
- Close to Generalized Procrustean (but does not need iterations).
- Close to INDSCAL and ALSCAL.
- To develop more: DISTATIS and MCA
- Open question: Sampling distribution of R_V for sorting, napping & flash profiling?

THANK YOU

Your Turn:

Time for Questions and Comments