Construction of confidence areas in sensometry

Minh Nhut HUYNH
Céline MARTIN
François HUSSON
Marine CADORET

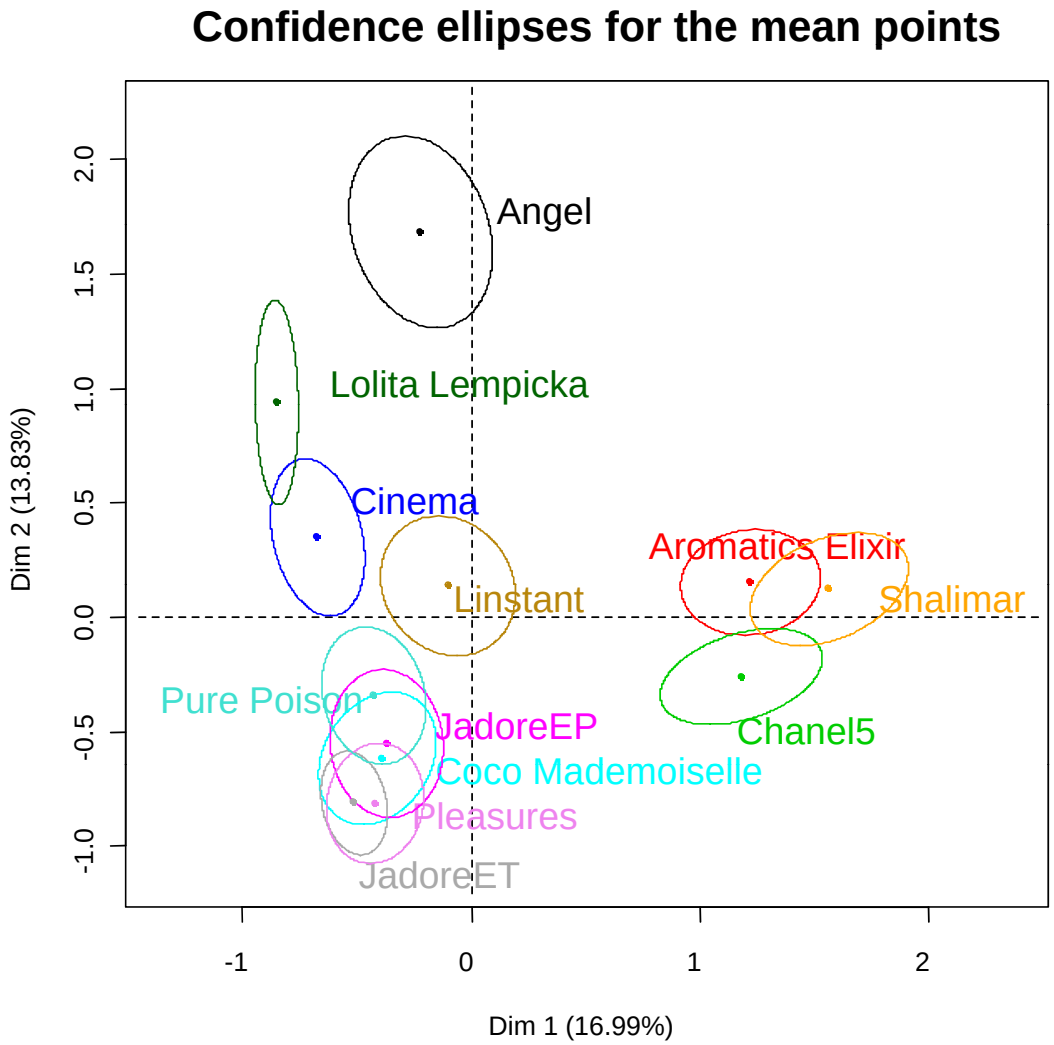
Agrocampus Ouest

Sensometrics 2012

An example of sorting task

	J.1	J.2	J.3	J.4	J.5
Angel	exotic;sweet;pleasant	flowery;soft	fruity;strong	vanilla;spicy;spirit of islands	to eat;sweet
Aromatics Elixir	strong;alcoholized	strong;man	heady	rough;strong	old
Chanel5	soap;clean	G4	heady	toilet	soap
Cinema	flowery;lilac	flowery;artificial;grass	fruity;middle	sweet	soft
Coco Mademoiselle	soap;clean	flowery;soft	fruity;middle	softness;flowery	soft
JadoreEP	soap;clean	flowery;soft	sweet;light	softness;flowery	flowery
JadoreET	soap;clean	flowery;artificial;grass	sweet;light	softness;flowery	flowery
Linstant	G3	flowery;soft	fruity;strong	sweet	old
Lolita Lempicka	exotic;sweet;pleasant	flowery;soft	fruity;middle	vanilla;spicy;spirit of islands	to eat;sweet
Pleasures	G3	strong;man	fruity;strong	sweet	flowery
Pure Poison	flowery;lilac	flowery;soft	tangy;deodorant	softness;flowery	soft
Shalimar	strong;alcoholized	flowery;artificial;grass	strong;lavander	mustiness;aggressive	old

An example of sorting task



Bootstrap technique

Real jury

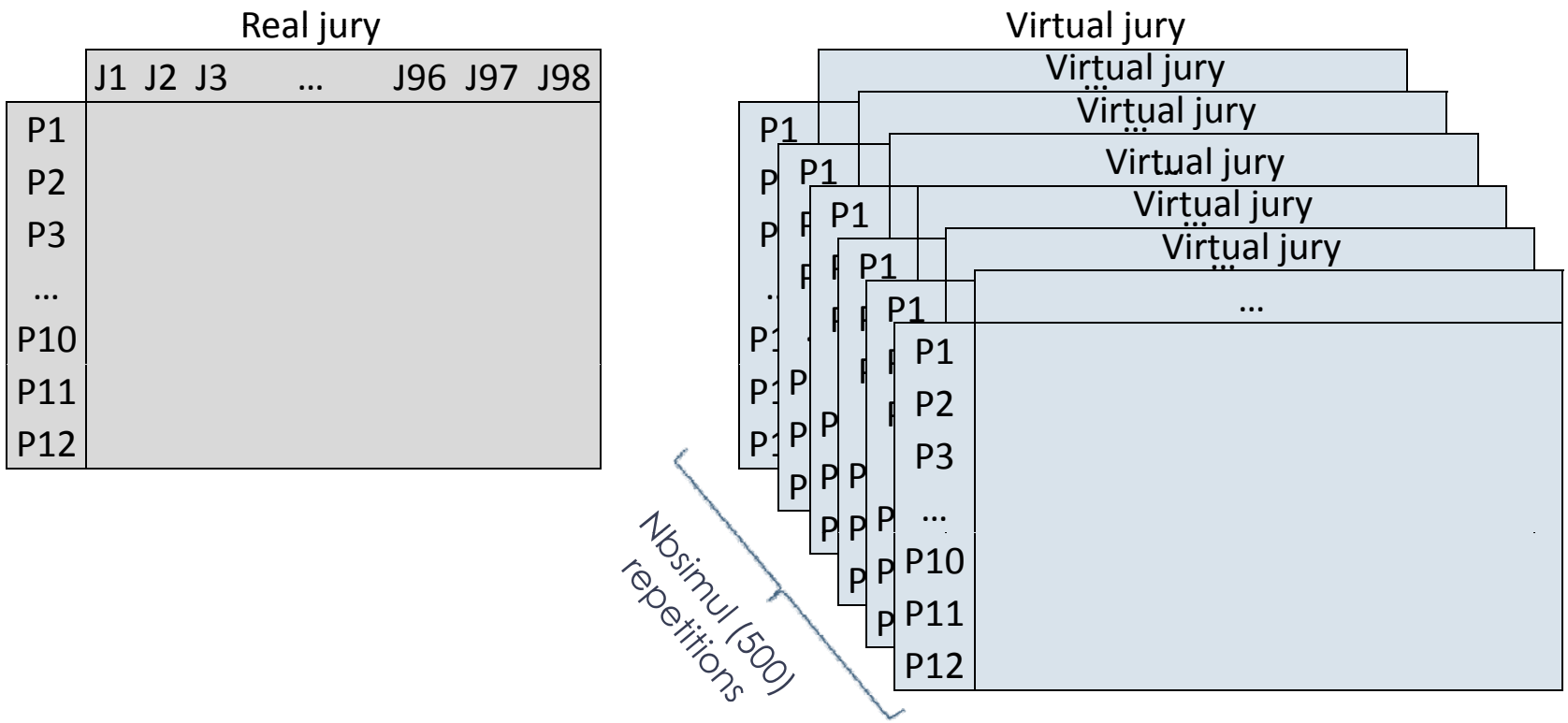
	J1	J2	J3	…	J96	J97	J98
P1							
P2							
P3							
…							
P10							
P11							
P12							

Virtual jury

	J1	J1	J19	…	J98
P1					
P2					
P3					
…					
P10					
P11					
P12					

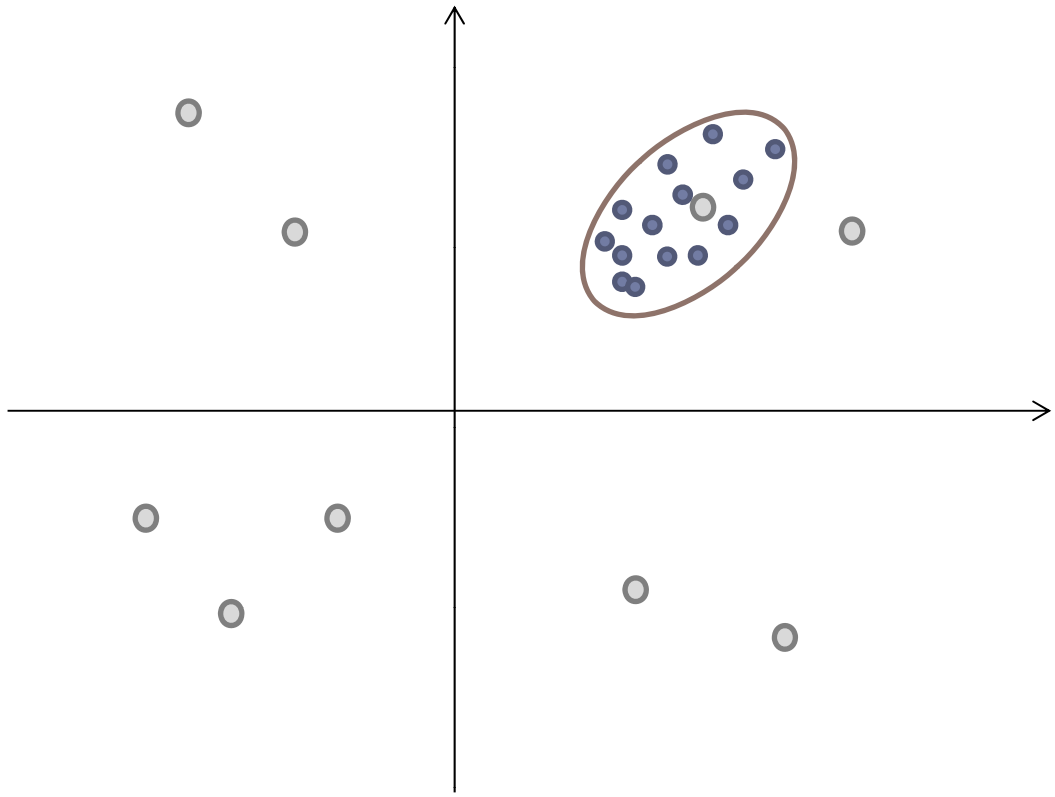
=> Random re-sampling with replacement

Bootstrap technique



- 2 ways to use bootstrapped virtual juries :
  - By projection (partial bootstrap)
  - By procrustean rotation (total bootstrap)

Partial bootstrap



1. Principal Component method on the data of the real jury

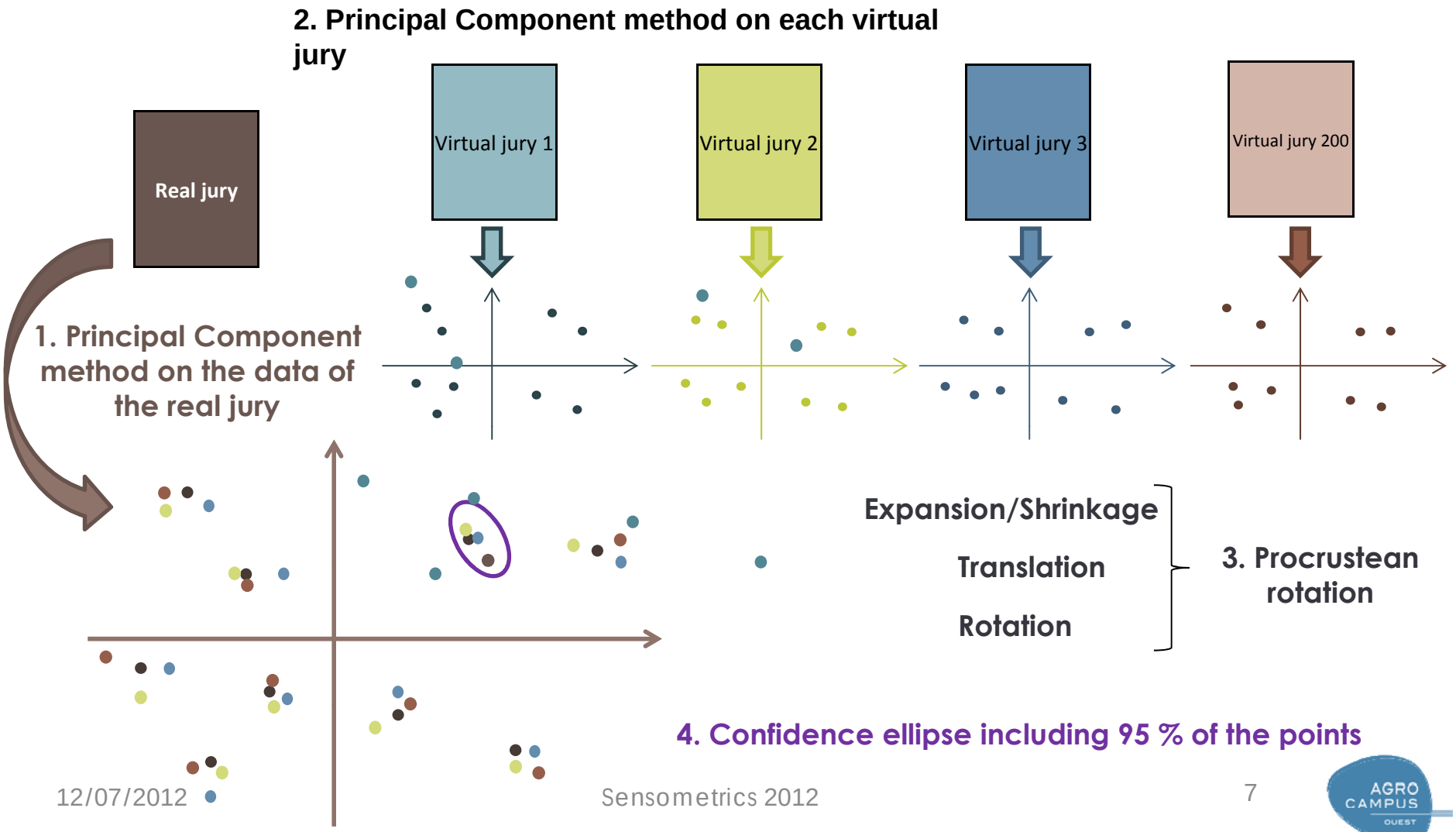
2. Projection of the judges of a virtual jury

3a. Calculation of the virtual jury's barycenter

3b. Calculation of the barycenters

4. Ellipse including 95% of the points

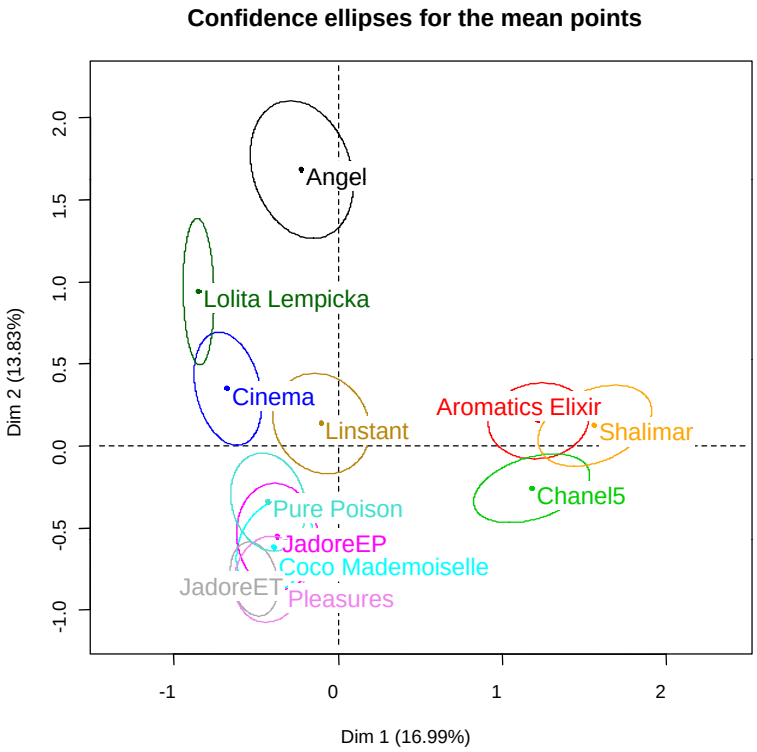
Total bootstrap



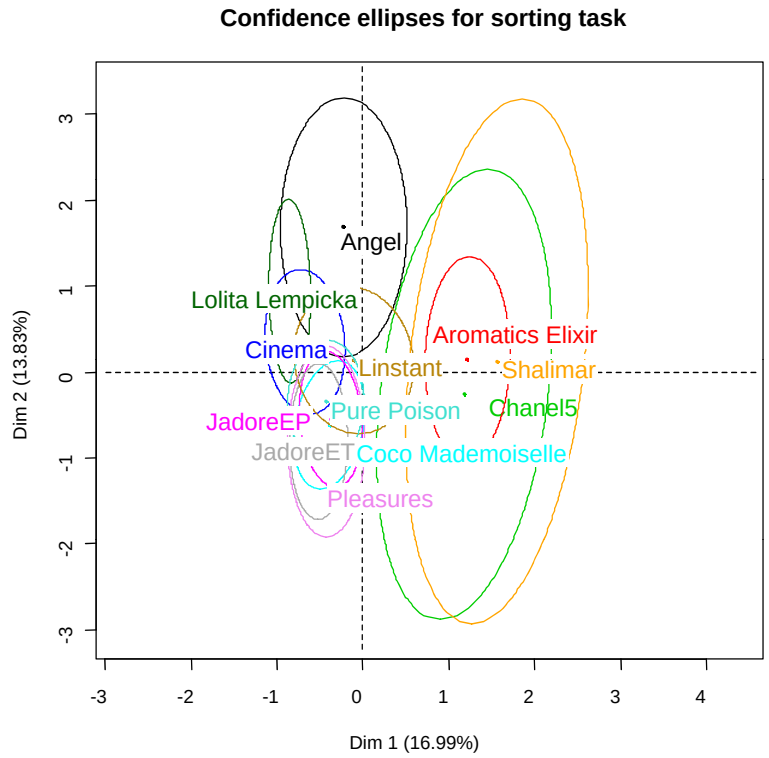
Comparison of partial and total bootstrap

Example of a sorting task dataset

Partial bootstrap



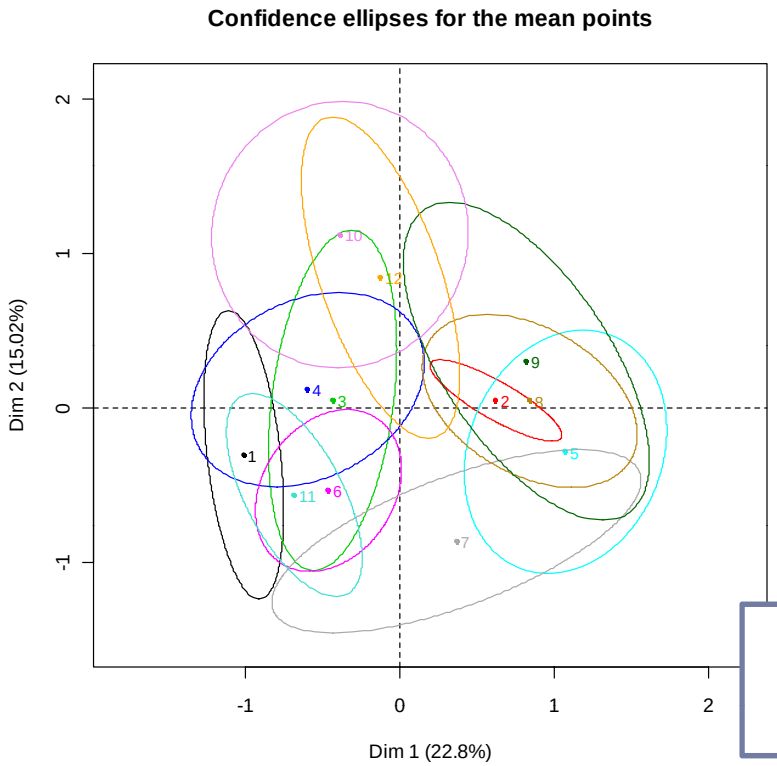
Total bootstrap



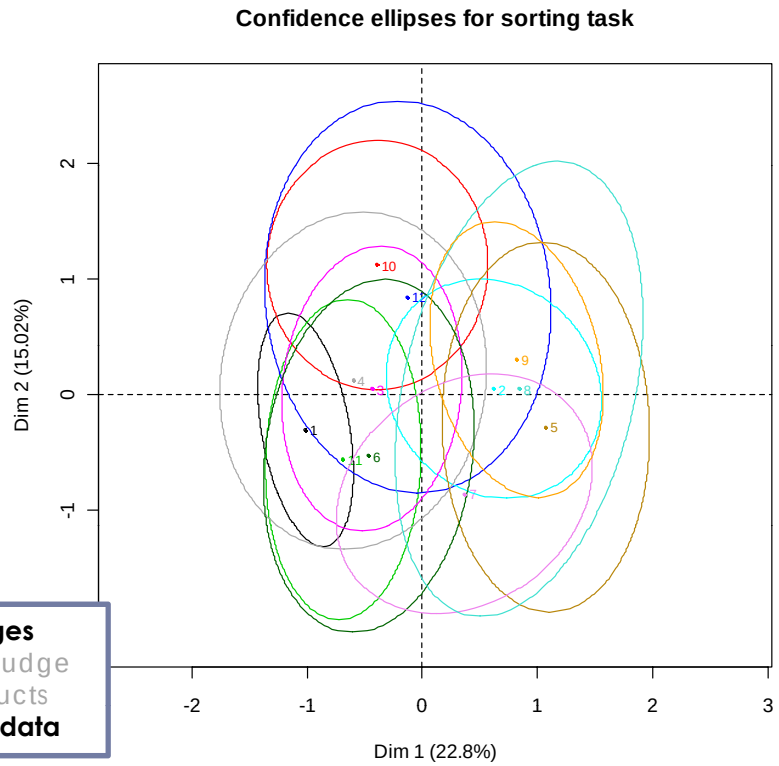
Comparison of partial and total bootstrap

*Example of a **RANDOM** sorting task dataset*

Partial bootstrap



Total bootstrap

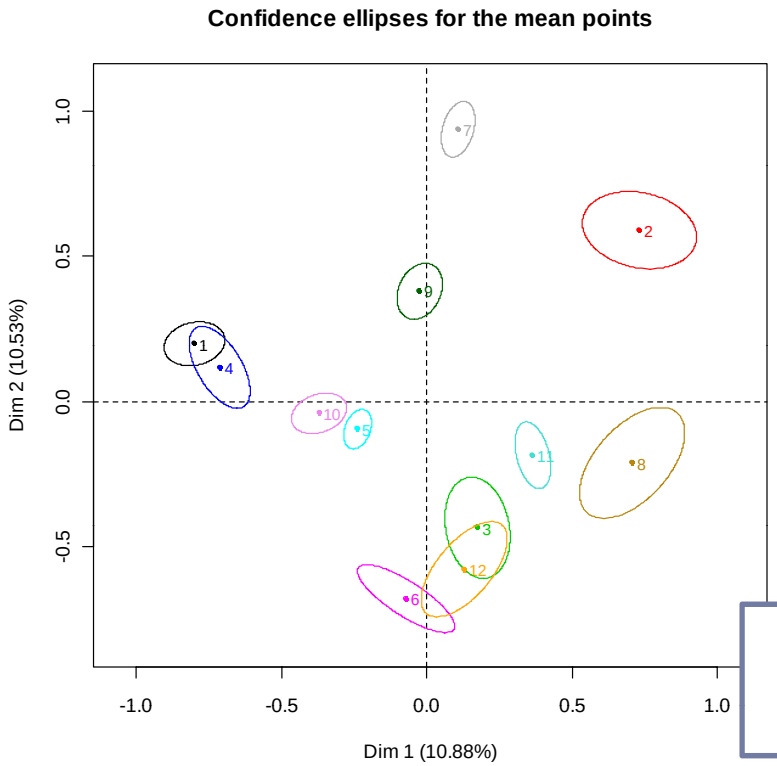


10 judges
3 groups/judge
12 products
Random data

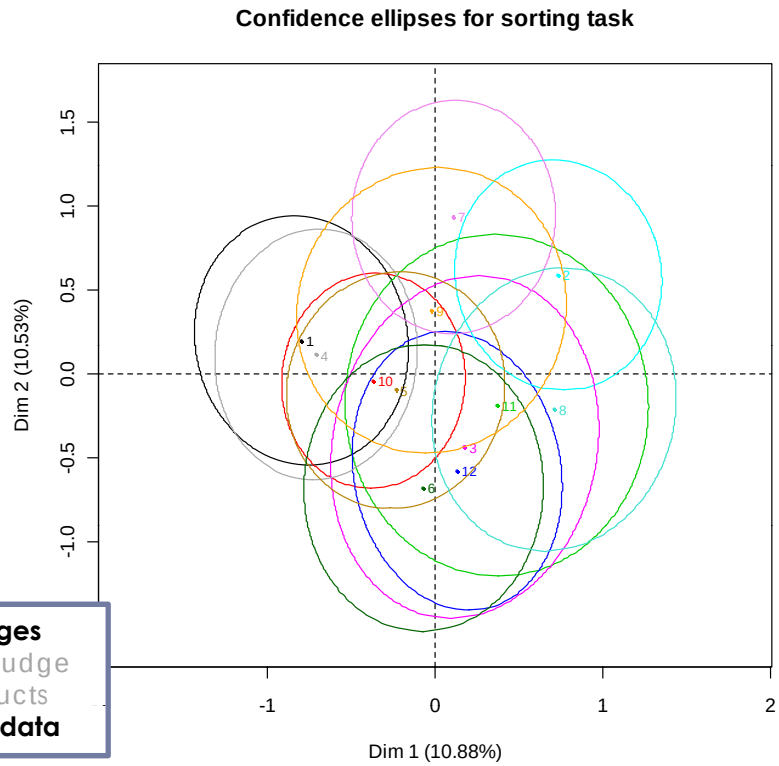
Comparison of partial and total bootstrap

*Example of a **RANDOM** sorting task dataset*

Partial bootstrap



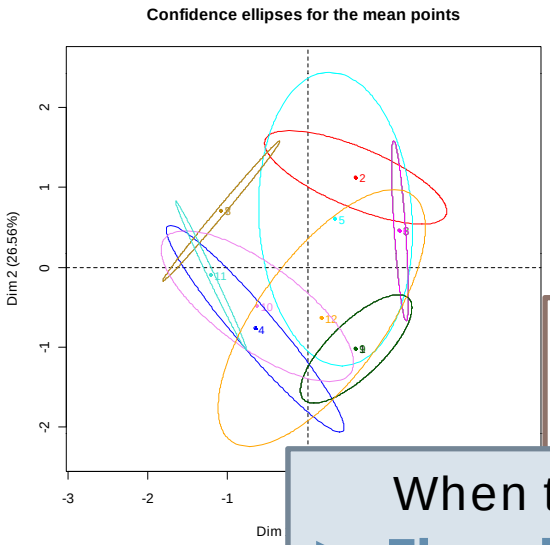
Total bootstrap



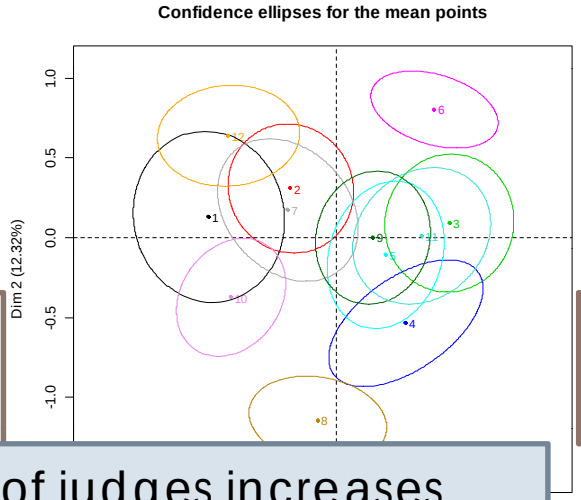
300 judges
3 groups/judge
12 products
Random data

Why partial bootstrap doesn't work ?

I. Random qualitative dataset



3 judges
3 groups/judge
12 products
Random dataset

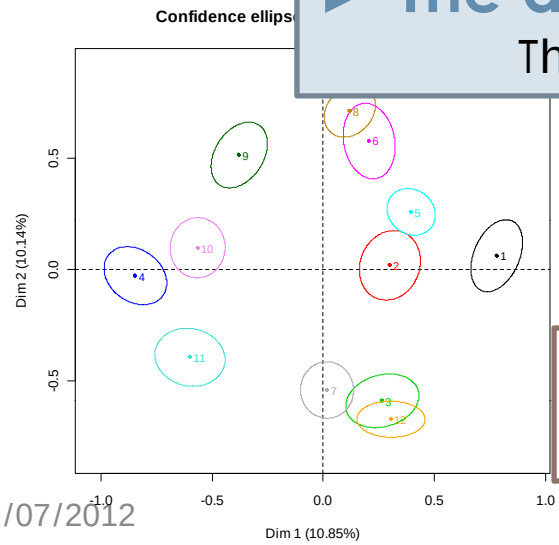


30 judges
3 groups/judge
12 products
Random dataset

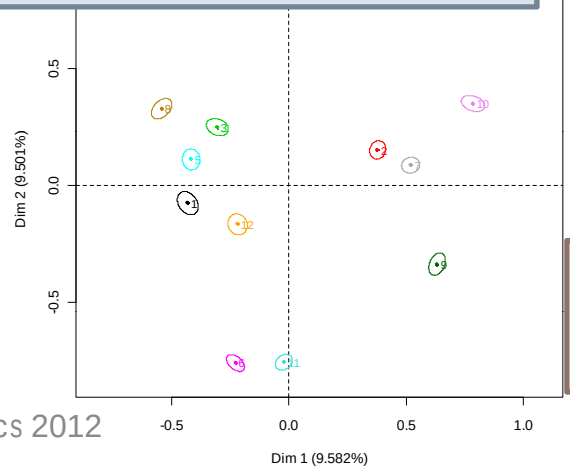
When the number of judges increases

▶ **The dimensionality increases**

The ellipses become smaller



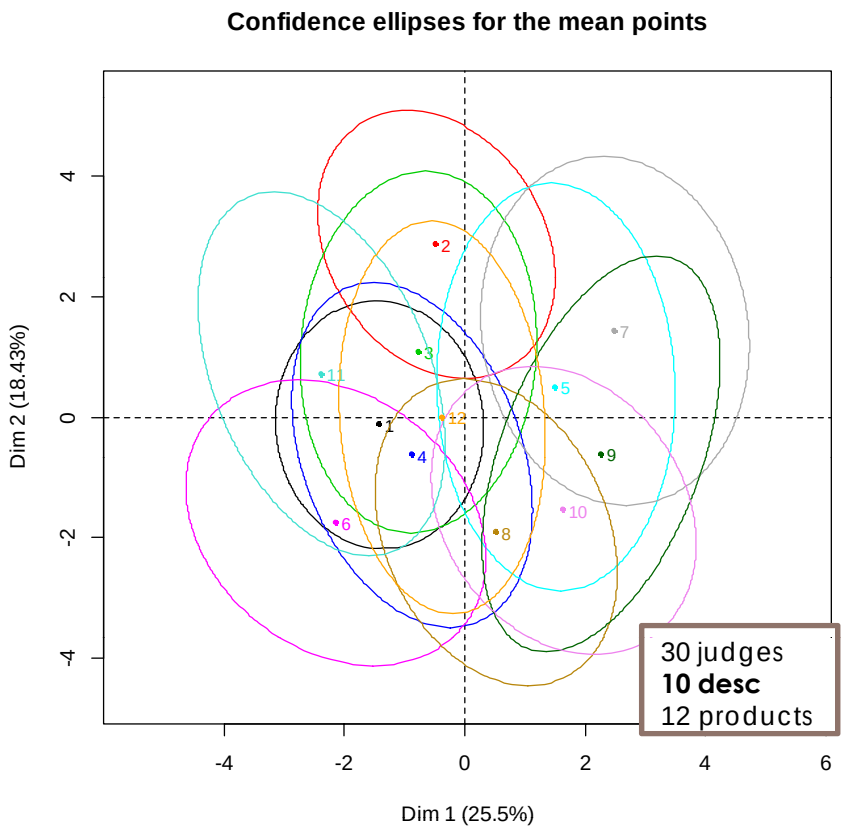
300 judges
3 groups/judge
12 products
Random dataset

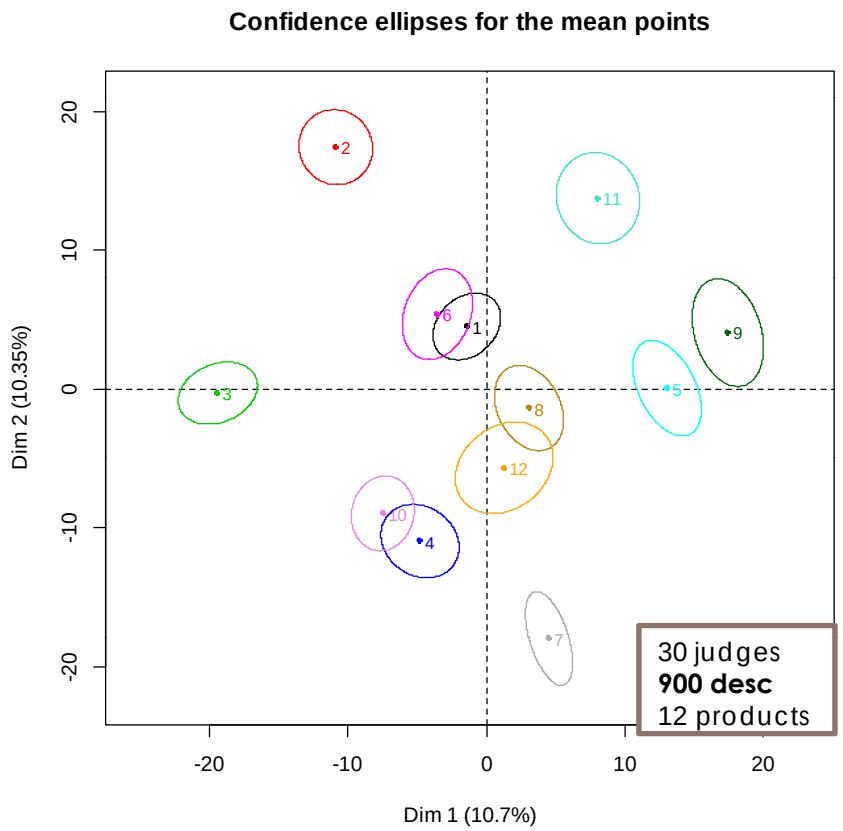


3000 judges
3 groups/judge
12 products
Random dataset

II. Quantitative dataset

*An example of QDA (**RANDOM** quantitative dataset)*

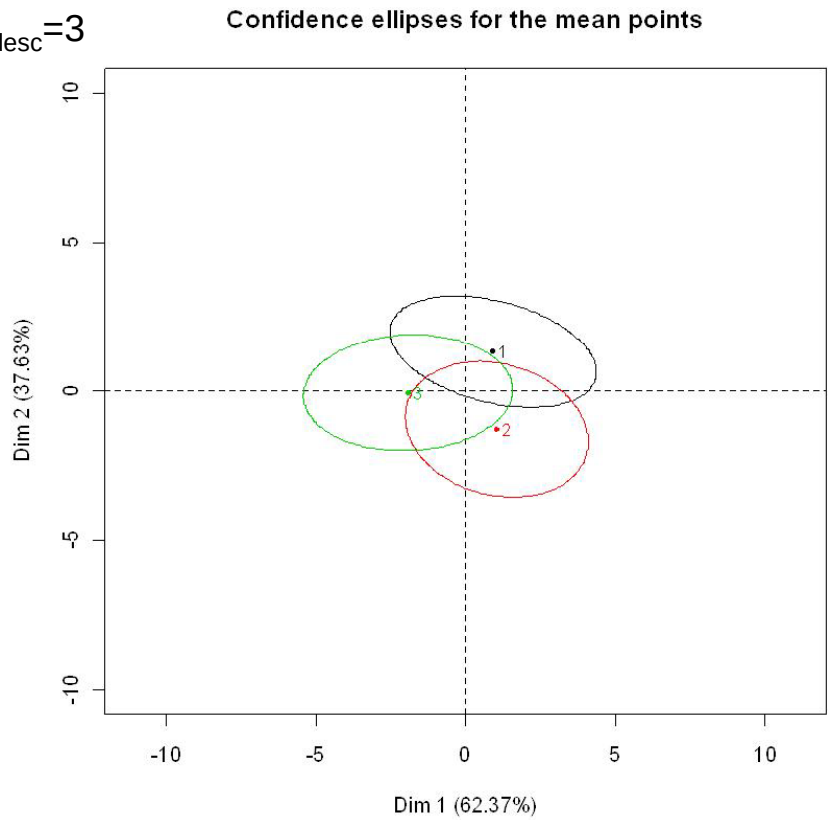




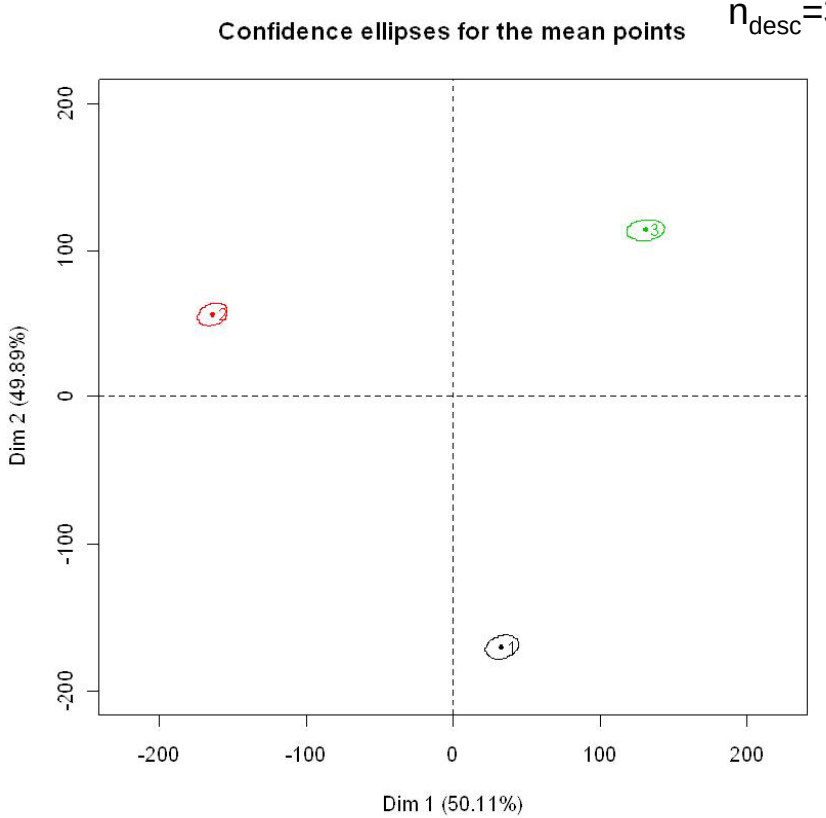
II. Quantitative dataset

*A simple **RANDOM** dataset*

n_{desc}=3



n_{desc}=30000

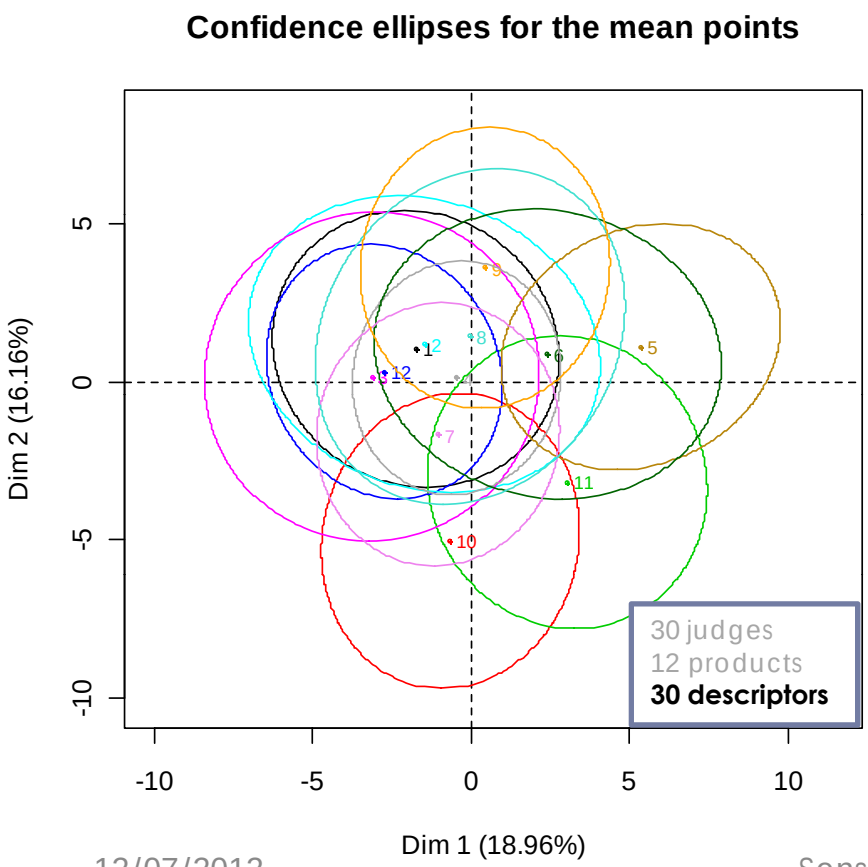


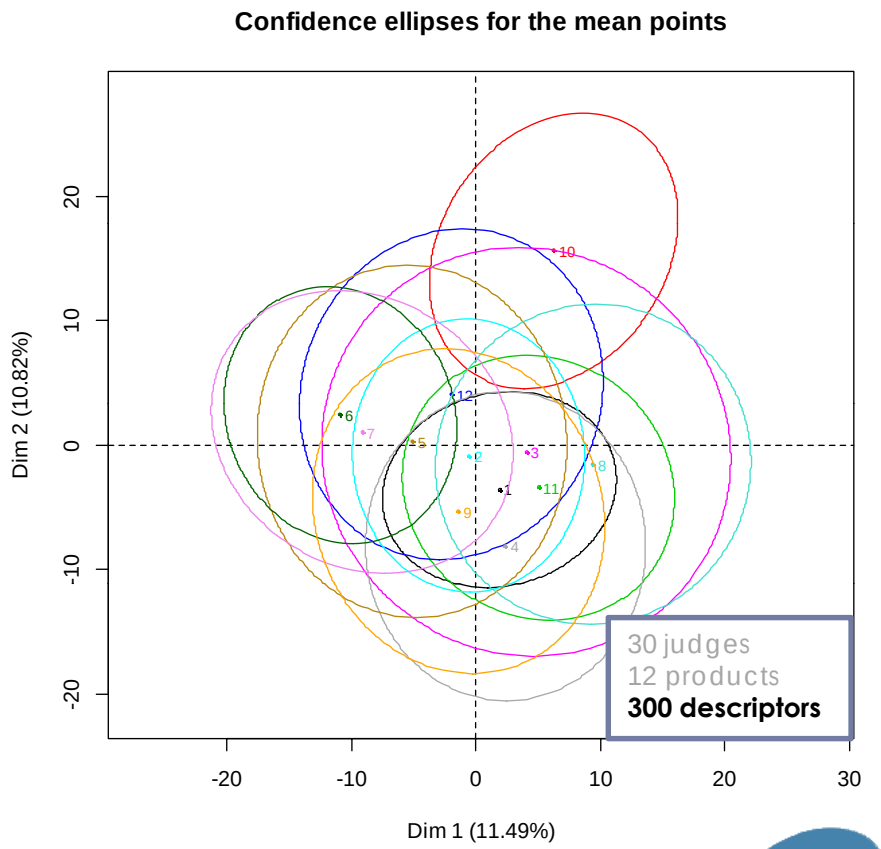
Projection of the average points of the virtual juries in illustrative
=> A point is a product as seen by a virtual jury

Validity of total bootstrap

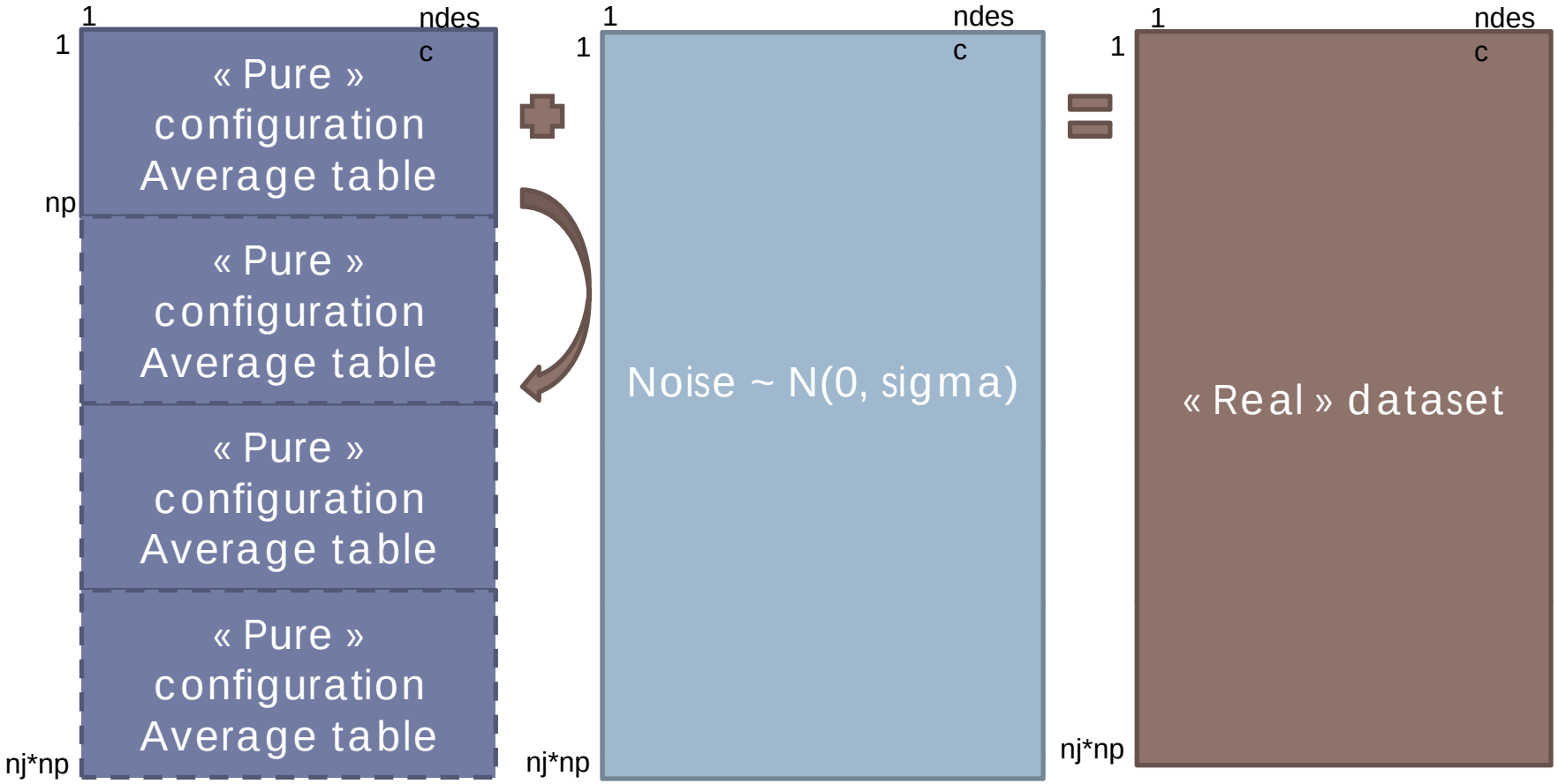
The case of completly random data

RANDOM data without structure



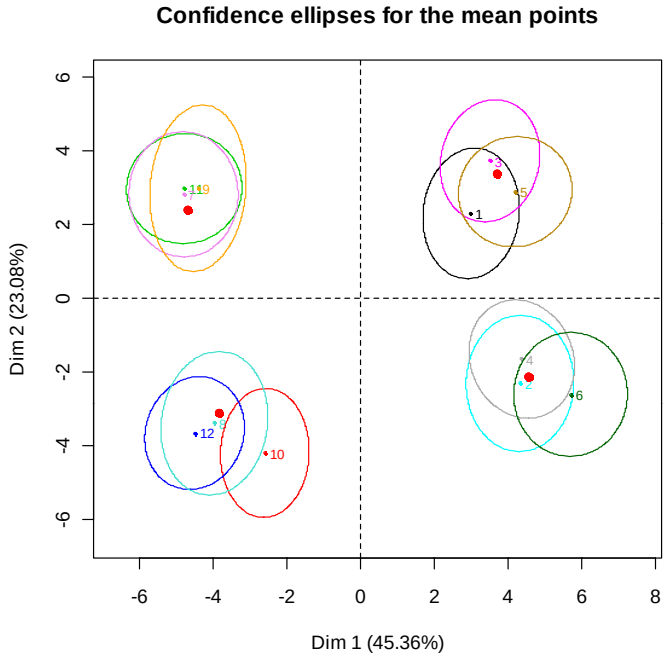


Data simulation procedure



Do the ellipses include the « real » structure of the data ?

Count the number of times that a point is included in an ellipse, given its coordinates on the factorial plan



	Sd	Frequence
Structured dataset	0.1	92,09%
	0.5	91,5%
	1	91,67%
	1.5	91,09%
	2	89,5%
	3	91,75%
	4	91,92%
	7	94,92%

	n_j	12	30	90	210
α = 5%	12	88,50	88,08	89,17	92,42
	30	91,42	90,00	91,00	93,92
	200	94,08	94,33	95,33	93,92

The proportion tends to 95% when the number of judges increases (asymptotic behavior of bootstrap techniques)

Conclusion

- One parameter must be chosen:
  - **Number of dimensions** for the procrustean rotations
    => n = 2 in many sensometry applications

- **Dimensionality problem** highlighted :
  Confidence ellipses are essential (but may be built according to total bootstrap)

- **Total bootstrap can be applied to all holistic** approaches:
  - Napping
  - Sorting task
  - Hierarchical sorting
  - Free Choice Profiling

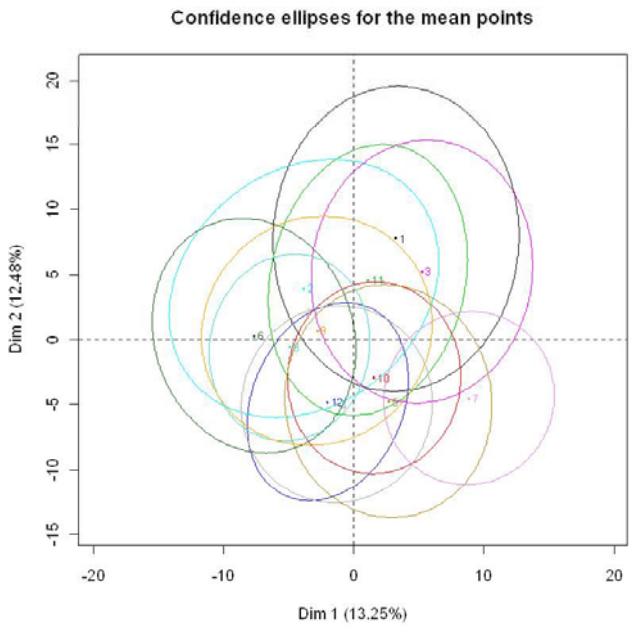
- Available into the **R package SensoMineR** through the boot function

Thank you

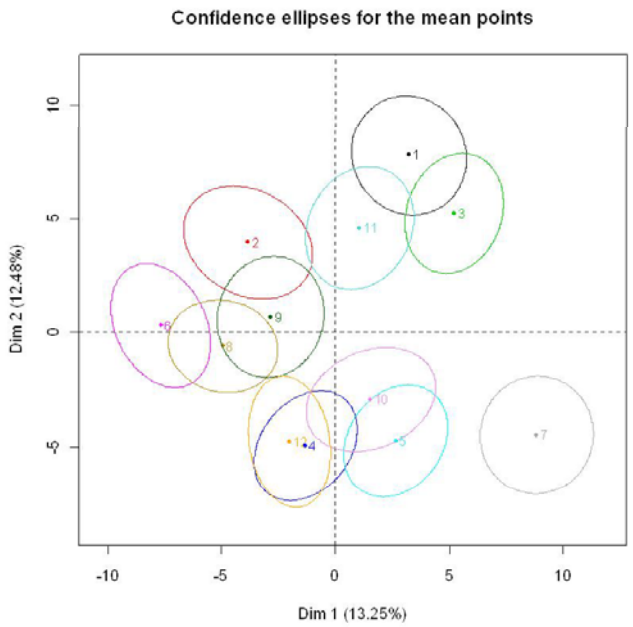
Do you have any questions ?

L'évolution de la taille des ellipses en fonction du nombre de dimensions utilisées pour les rotations procustéennes

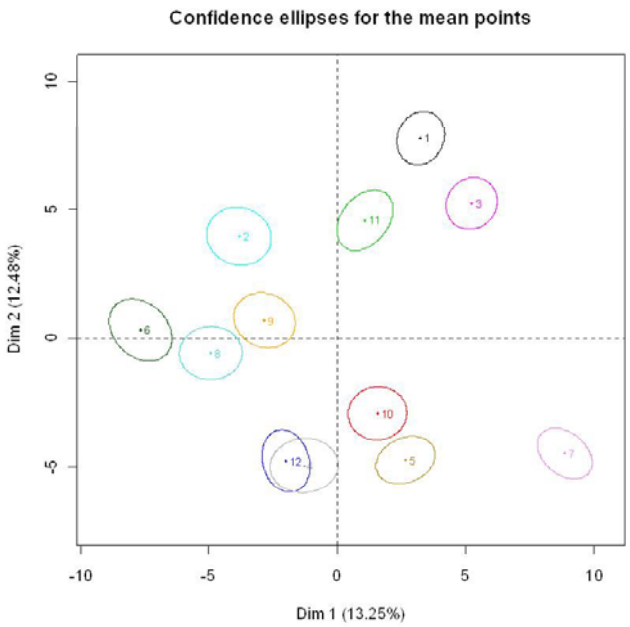
30 juges
12 produits
Sans structure



ncp=2



ncp=7



ncp=11