

# I — Overview of Geometric Data Analysis (GDA)

Brigitte.LeRoux@math-info.univ-paris5.fr  
rouanet@math-info.univ-paris5.fr

[www.math-info.univ-paris5.fr/~lerb/](http://www.math-info.univ-paris5.fr/~lerb/)  
[www.math-info.univ-paris5.fr/~rouanet/](http://www.math-info.univ-paris5.fr/~rouanet/)

# 1 What is GDA?

Data are represented by clouds of points.

The statistical interpretation is based on clouds.

Data Table



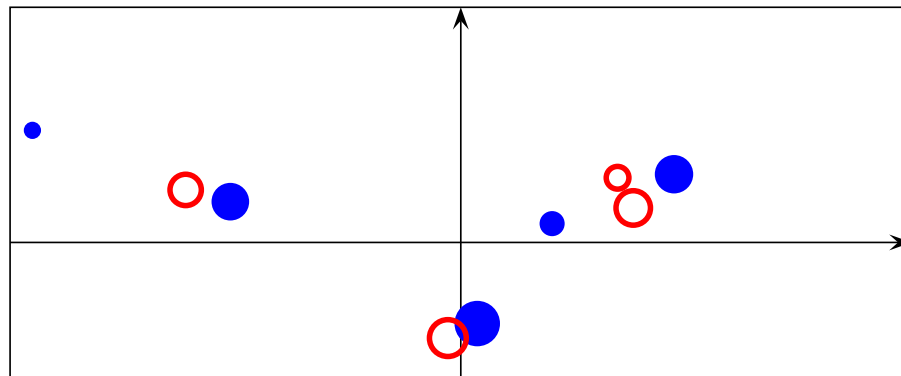
Clouds of points

Table

326	688	343	98
38	116	84	48
241	584	909	403
110	188	412	681
3	4	26	85



Clouds (• & ○)



## 2 Three paradigms of GDA

GDA methods were introduced by Benzécri around the leading case of Correspondence Analysis (CA).

They are well known in France as “Analyse des données”

- Contingency table  $\longrightarrow$  Correspondence Analysis
- Individuals  $\times$  Variables table  $\longrightarrow$ 
  - Principal Component Analysis (PCA): numerical variables
  - Multiple Correspondence Analysis (MCA): categorized variables



### 3 Three Key Ideas

	Main stream of Statistics	GDA
1.	Quantitative	Geometry (multidimensional)
2.	Matrix approach	Linear Algebra
3.	Sampling-oriented	Inductive

Remark for 3.

*Sampling-oriented*: a dataset is supposed to be a **random sample**

vs *Inductive approach*: **Description comes first.**

## 4 Historical sketch

- *Precursors*

Karl Pearson (1901): Geometric Display of Data

Hirschfeld (1935), Fisher (1940): equations of CA without geometric developments

Guttman (1941): equations of MCA

- *Landmarks for GDA*

1963-73: Emergence. Benzécri & al 1973: Analyse des données (I. Taxinomie, II. Analyse des correspondances)

1973-80: Golden age in France and splendid isolation

Since 1981: Bounded international recognition: Greenacre (1984), Lebart & al (1984), Benzécri (1992).

*Where do we stand now?* (2006)

## 5 Methodological Strong Points

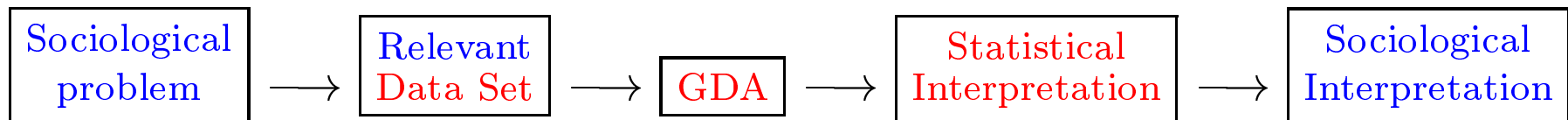
- *GDA as a Frame Model*

Homogeneity and Exhaustiveness

Construction of social spaces: the case of Bourdieu's sociology

Emphasizing individuals.

Explanatory schemes; geometric analysis of questionnaires



## 6 Structured Data Analysis

*Structuring factors*

*Structured data* = data tables with structuring factors



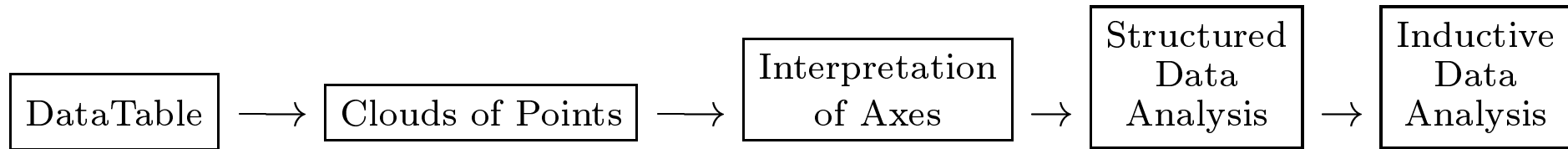
## 7 Statistical inference in GDA

Statistical inference should be conducted as an *Inductive Data analysis*, that is, in the line of the Data Analysis Philosophy, as a natural extension of Descriptive Data Analysis.



*“The model should follow the data, not the reverse!”*

## 8 Summarizing: What Is GDA?



## 9 About Software

The case for SPAD

### References

Le Roux B. & Rouanet H. (2004), *Geometric Data Analysis: From Correspondence Analysis to Structured Data Analysis*, chapter 1, p.1-22.

For software, see

[www.math-info.univ-paris5.fr/~lerb/Logiciels/logiciels.html](http://www.math-info.univ-paris5.fr/~lerb/Logiciels/logiciels.html)