

Geometric Data Analysis (GDA)

I — Introduction

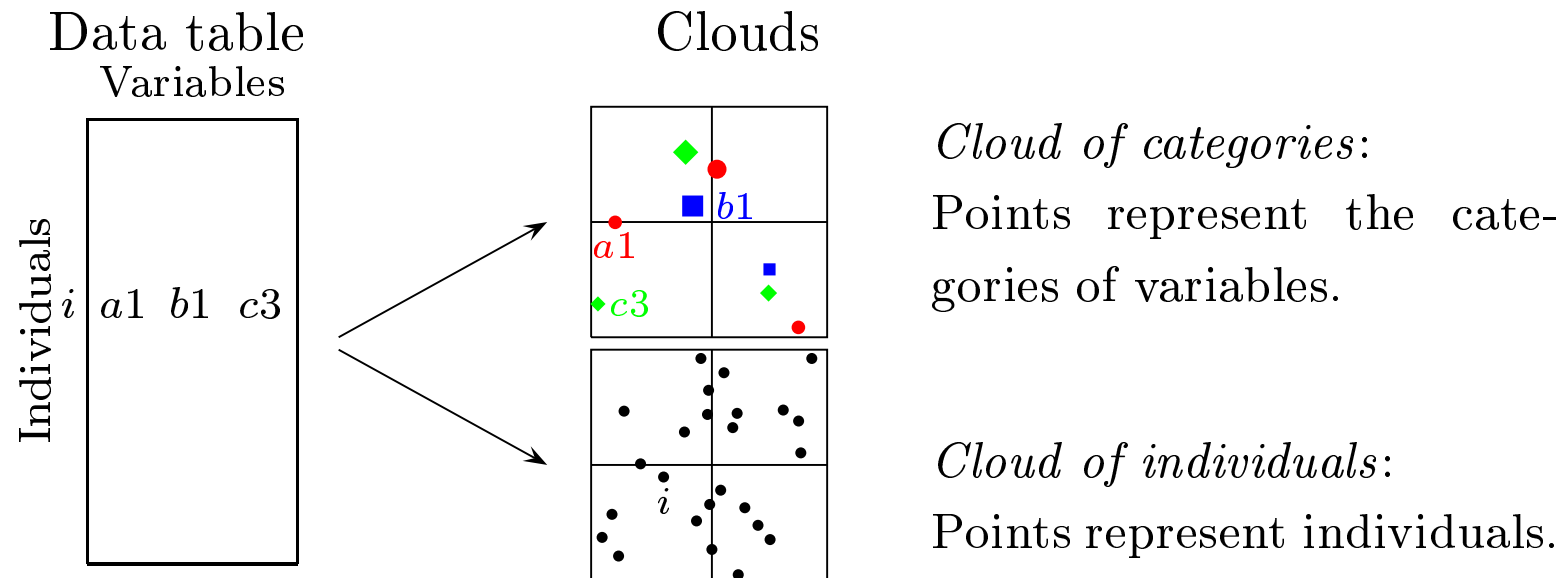
Brigitte.LeRoux@mi.parisdescartes.fr

www.mi.parisdescartes.fr/~lerb/



I.1 Three Key Ideas

- *Geometric modeling*



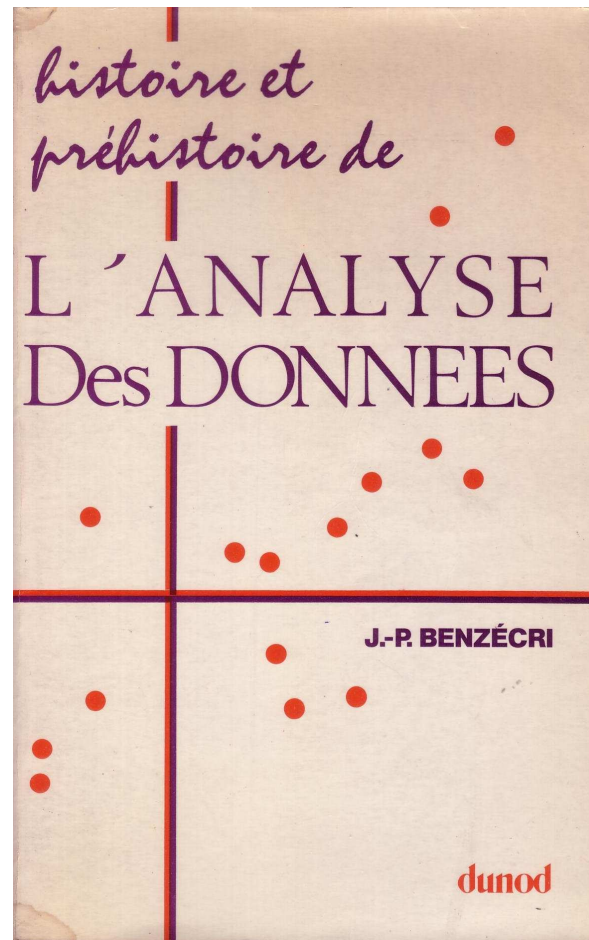
- *Formal approach.*

- *Description first!:* “The model should follow the data, not the reverse!”.

I.2 Three paradigms

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

I.3 Historical sketch



J.-P. Benzécri (1982)

- *Precursors*

Karl Pearson (1901), Hirschfeld (1935).

Should we need an Anglo–Saxon patronage for “Analyse des Données”, we would be pleased to turn to the great Karl Pearson.

Benzécri (1982), p. 116

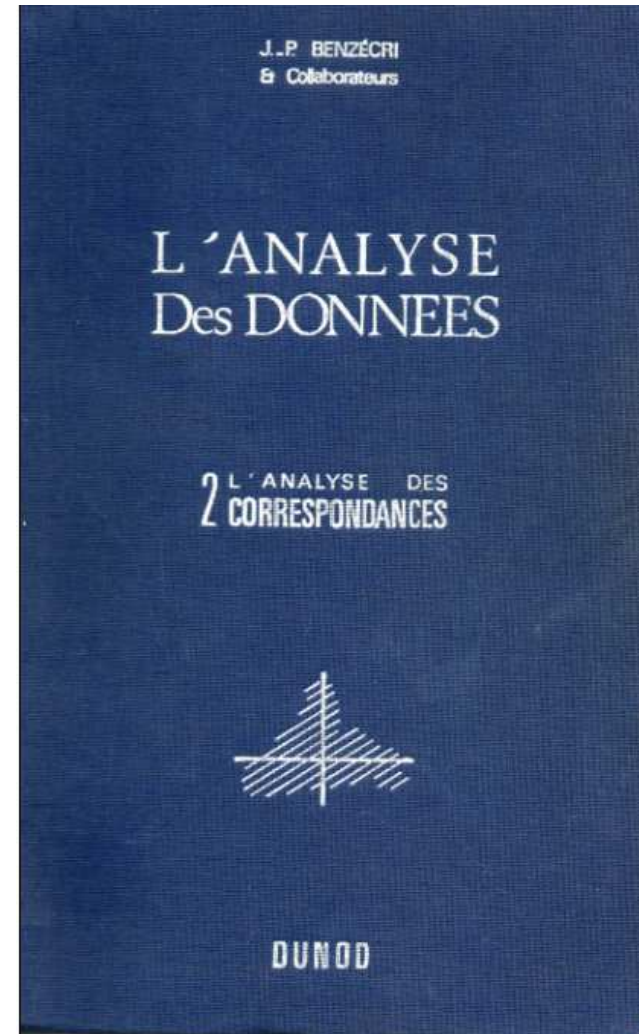
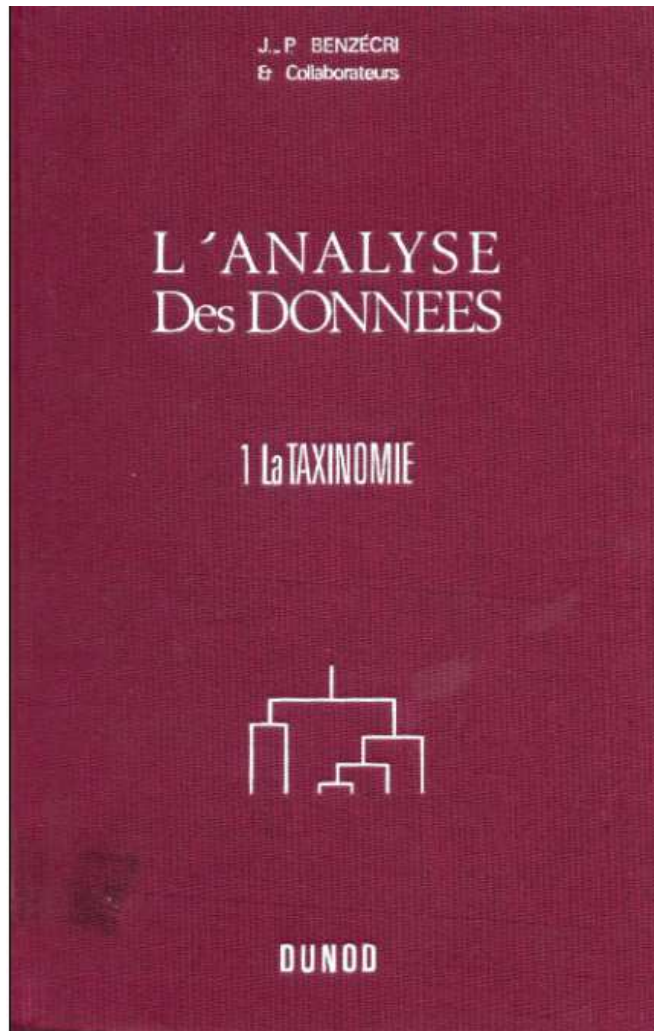
— Optimal scaling: Fisher (1940), Guttman (1942)

— Factor analysis: Burt (1950)

— Quantification method: Hayashi (1952)

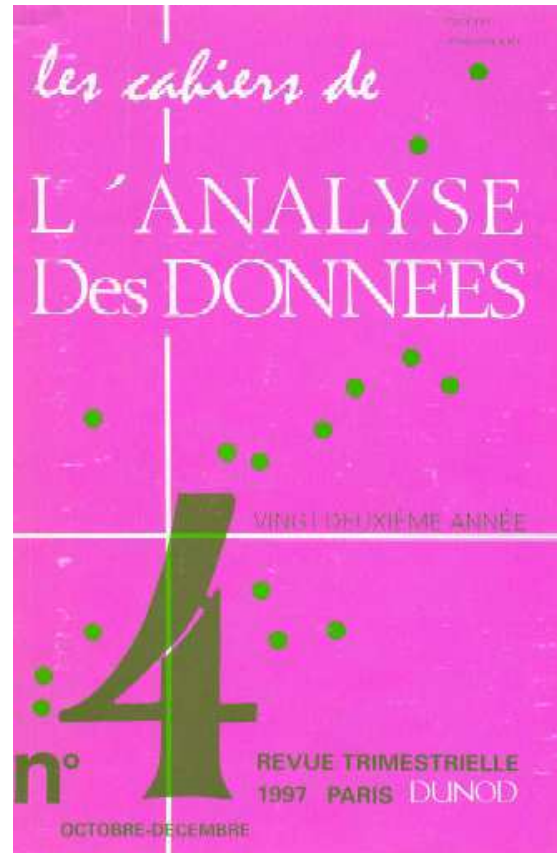
— MDS: Shepard (1962).

• *Emergence (1963-73).*



Benzécri & al (1973): L'ANALYSE Des DONNÉES
1 la TAXINOMIE 2 L'ANALYSE DES CORRESPONDANCES.

- *Recognition and splendid isolation (1973-1980)*



1977–1997

Gower (1966), Good (1969), Gabriel (1971)

Ignored in Shepard, Romney, Nerlove (1972), Kruskal & Wish (1978), Shepard (1980) and in Kendall & Stuart (1976)

• *International recognition (since 1981):*

Greenacre (1984), Lebart & al (1984), Jambu (1991),
Benzécri (1992) (translation of the introductory book, Dunod,
1984);

Malinvaud (1980), Deville & Malinvaud (1983):

“Econometrics without stochastic models”;

Tenenhaus & Young (1985) : Psychometry;

Nishisato (1980): Dual Scaling;

Gifi (1981/1990): Homogeneity Analysis;

Carroll & Green (1988), Weller & Romney (1990): MDS group;

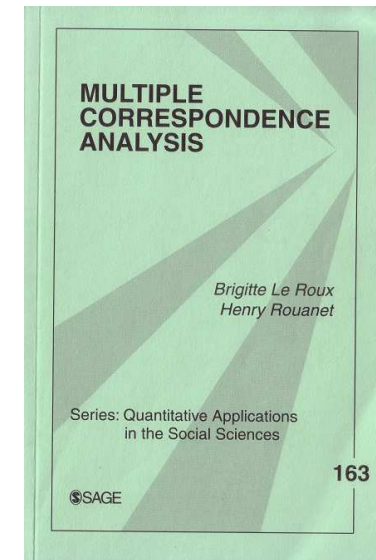
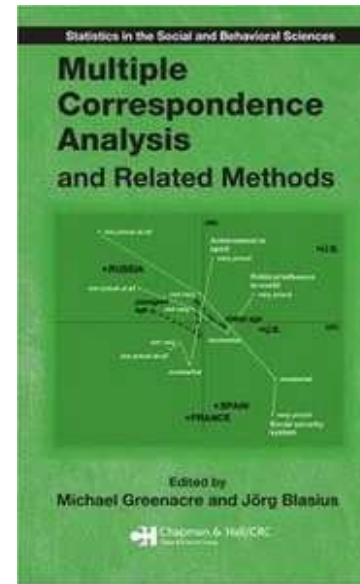
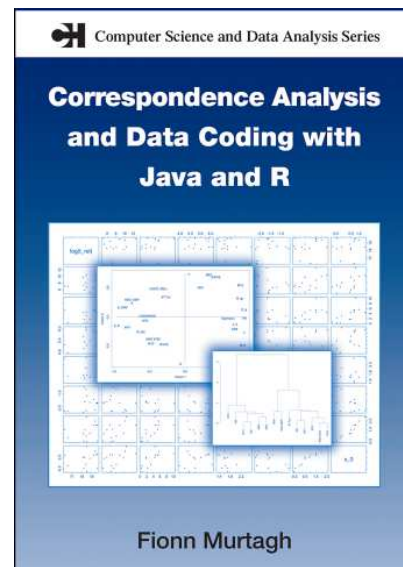
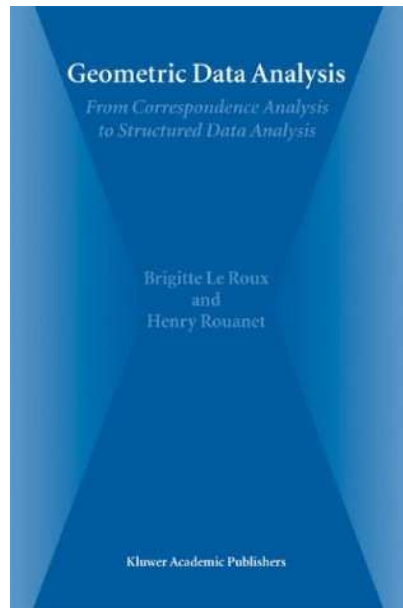
Gower & Hand (1996): biplot.

- *Where do we stand now?*

Weller & Romney (1990), Goodman (1991), Rao (1995),
CARME network: international conferences in Cologne (1991, 1995,
1999), Barcelona (2003), Rotterdam (2007) and Rennes (2011)..

Recent Books:

Le Roux & Rouanet (2004) Murtagh (2005) CARME 2003 (2006) Le Roux & Rouanet (2010)



CA is now recognized and used, but GDA as a whole methodology, and MCA in particular, are waiting to be discovered by a large audience.