

Abstract Book

16th Conference of the International Federation of Classification Societies

26-29 August 2019
Thessaloniki Concert Hall
Thessaloniki, Greece
#IFCS2019

ORGANISED BY



LOCAL ORGANISER



IN COOPERATION WITH

**THE GREEK
SOCIETY
OF DATA ANALYSIS**

DIRECTOR OF THE CONFERENCE

Theodore Chadjipadelis

T. +30 2310 997912, E. chadji@polsci.auth.gr



Assessment of recent social attitudes in Japan: a latent class item response theory model for web survey data	228
Miki Nakai	
POSTERS	229
The relationship of the apolipoprotein E genotype gene to the Alzheimer's disease: a meta-analysis	230
Sofia D. Anastasiadou	
Bayesian analysis for chromosomal interactions in hi-c data using hidden Markov random field model ..	231
Osuntoki G. Itunu	
New financial instruments: Pollution emission rights and their trading on the stock exchange	232
Argiro Dimitoglou	
Econometric assessment of the relation between the situation of youth on the labour market and macroeconomic situation among the EU countries	233
Beata Bal-Domańska	
Comparison of patterning methods: Clustering of variables, Implicative Statistical Analysis and Analyse Factorielle des Correspondances	234
Sofia D. Anastasiadou	
Framing coworking spaces digital marketing strategy via social media analytics	235
Dimitrios Vagianos	
Sales performance measure: A systematic review and typology of research studies	236
Tor Korneliussen	
Document clustering via multiple correspondence, term and metadata analysis in R	237
Nikos Koutsoupas	
Comparison of multivariate methods in group/cluster identification: PCA vs discriminant analysis and K-Means clustering	238
Sofia D. Anastasiadou	
Asymptotic cumulants of the minimum phi-divergence estimator for categorical data under possible model misspecification	239
Haruhiko Ogasawara	
Multidimensional data analysis in perception of European Union by different generations	240
Agnieszka Stanimir	
INDEX	242

The 2019 conference of the International Federation of Classification Societies (IFCS) will be held in Thessaloniki, Greece from August 26 to 29, 2019. The conference theme will be 'Data Analysis and Rationality in a Complex World'. The conference opening will take place on August 26 late afternoon and pre-conference workshops will be held. The conference itself will start on August 27 in the morning, and will close on August 29 with a full day conference program and a conference dinner. The conference will include a President's invited session and a Presidential address, invited presentations, invited and contributed sessions, and poster sessions. The following awards and medals ceremony will be scheduled at the conference.

Helga and Wolfgang Gaul Stiftung Award (age < 30)

Chikio Hayashi Award (age 30 – 35)

IFCS Research Medal for outstanding research achievements (age > 35)

Student/Postdoctoral Fellow Paper Competition and Travel Award (age 22 – 29)

Cluster Benchmarking Competition Award

COMMITTEES

Scientific Program Committee

Theodoros Chadjipadelis	Chair
Berthold Lausen	Vice-chair, IFCS President
Tae Rim Lee	Vice-chair
Angela Montanari	IFCS President Elect
Akinori Okada	IFCS Past President
Christian Hennig	IFCS Scientific Secretary
Sugnet Lubbe	IFCS Treasurer
Katrijn van Deun	Publication Officer
Eva Boj	SEIO-AMyC
Paula Brito	CLAD
Sanjeena Dang	CS
Giannoula Florou	GSDA
Krzysztof Jajuga	SKAD
Hans Kestler	GfKI
Simona Korenjak Černe	SSS
Peter Kovacs	HSA-CMSG
Koji Kurihara	JCS
Angelos Markos	GSDA
Paul McNicholas	CS
Fionn Murtagh	BCS
Mauricio Vichi	CLADAG
Roberto Rocci	CLADAG
Mark de Rooij	VOC
Niel le Roux	SASA-MDAG

Local Organizing Committee

Theodore Chadjipadelis	LOC Chair, Aristotle University Thessaloniki
Giannoula Florou	President GSDA
Sofia Anastasiadou	University of Western Macedonia
Dimitris Karapistolis	Alexandrian Technological Institute
Nikos Koutsoupas	University of Macedonia
Angelos Markos	Democritus University of Thrace
George Menexes	Aristotle University Thessaloniki
Odysseas Moschidis	University of Macedonia
Thanos Thanopoulos	Hellenic Statistical Authority
Christos Tzimos	Hellenic Statistical Authority
Despoina Amarantidou	ARTION conferences & events
Vicky Bouranta	Secretary
Marina Sotiropoulou	Secretary

On the use and reporting of cluster analysis in educational research: A systematic review
Hanneke van der Hoef, Matthijs Warrens, Marieke Timmerman

11.10 – 14.00 Poster Session

Poster Area

Chairs: Sofia Anastasiadou, Odysseas Moschidis

The relationship of the apolipoprotein E genotype gene to the Alzheimer's Disease: A meta-analysis

Sofia Anastasiadou

Bayesian analysis for chromosomal interactions in hi-c data using hidden Markov random field model

Osuntoki Intunu, Andrew Harrison, Hongsheng Dai, Yanchun Bao, Nicolae Zabet

New financial instruments: Pollution emission rights and their trading on the stock exchange

Argiro Dimitoglou

Econometric assessment of the relation between the situation of youth on the labour market and macroeconomic situation among the EU countries

Beata Bał-Domańska, Elżbieta Sobczak

Comparison of pattering methods: Clustering of variables, Implicative Statistical Analysis and Correspondence Analysis

Sofia Anastasiadou

Framing coworking spaces digital marketing strategy via social media analytics

Dimitrios Vagianos, Nikos Koutsoupas

Sales performance measure: A systematic review and typology of research studies

Tor Korneliussen, Per Ivar Seljeseth, Michael Greenacre

Document clustering via multiple correspondence, term and metadata analysis in R

Nikos Koutsoupas, Kyriakos Mikelis

Comparison of multivariate methods in group/cluster identification: PCA vs Discriminant Analysis and K-Means clustering

Sofia Anastasiadou

Asymptotic cumulants of the minimum phi-divergence estimator for categorical data under possible model misspecification

Haruhiko Ogasawara

Multidimensional data analysis in perception of European Union by different generations

Agnieszka Stanimir

13.00 – 14.00 Mid-day Break

14.00 – 14.50 Presidential Address: Predictive ensemble methods for event time data

Berthold Lausen

Chair: Angela Montanari

14.50 – 15.40 Plenary Invited: On the consistency of supervised learning with missing values

Julie Josse

Chair: Angelos Markos

Plenary Hall

15.40 – 16.10 Coffee Break

Parallel Sessions

16.10 – 17.10 SP12: New developments in clustering and scaling data

organized by J. F. Vera & E. Boj del Val

Chair: Jose Luis Vicente-Villardón

Room I



Document clustering via multiple correspondence, term and metadata analysis in R

Nikos Koutsoupias, and Kyriakos Mikelis

Abstract We introduce the combined use of multiple correspondence analysis, metadata and term frequencies for clustering articles of a scientific journal. A period of five years (2010-2014) is covered, with approximately 125 articles. Through specific R packages for multidimensional data analysis and text mining, the approach links quantitative analysis of discourse to clustering documents considering both metadata and frequent terms.

Keywords document clustering; hierarchical clustering; multiple correspondence analysis; document metadata; text mining

Nikos Koutsoupias

University of Macedonia, Greece, e-mail: nk@uom.gr

Kyriakos Mikelis

University of Macedonia, Greece, e-mail: mikelis@uom.gr

Document Clustering via Multiple Correspondence, Term and Metadata Analysis in R

Nikos Koutsoupias / Kyriakos Mikelis | {nk/mikelis}@uom.gr | Department of International & European Studies | University of Macedonia, Thessaloniki, GR

PROBLEM

Divide a set of documents based on metadata and term frequencies

DATA SOURCE

- Focus on articles from “*Perceptions: Journal of International Affairs*”
- Published in English under the auspices of the Centre for Strategic Research (SAM), Turkey
- A period of 5 years (2010-2014) is covered.
- Nearly 125 full articles examined
- Editorials and book reviews were left out



METADATA

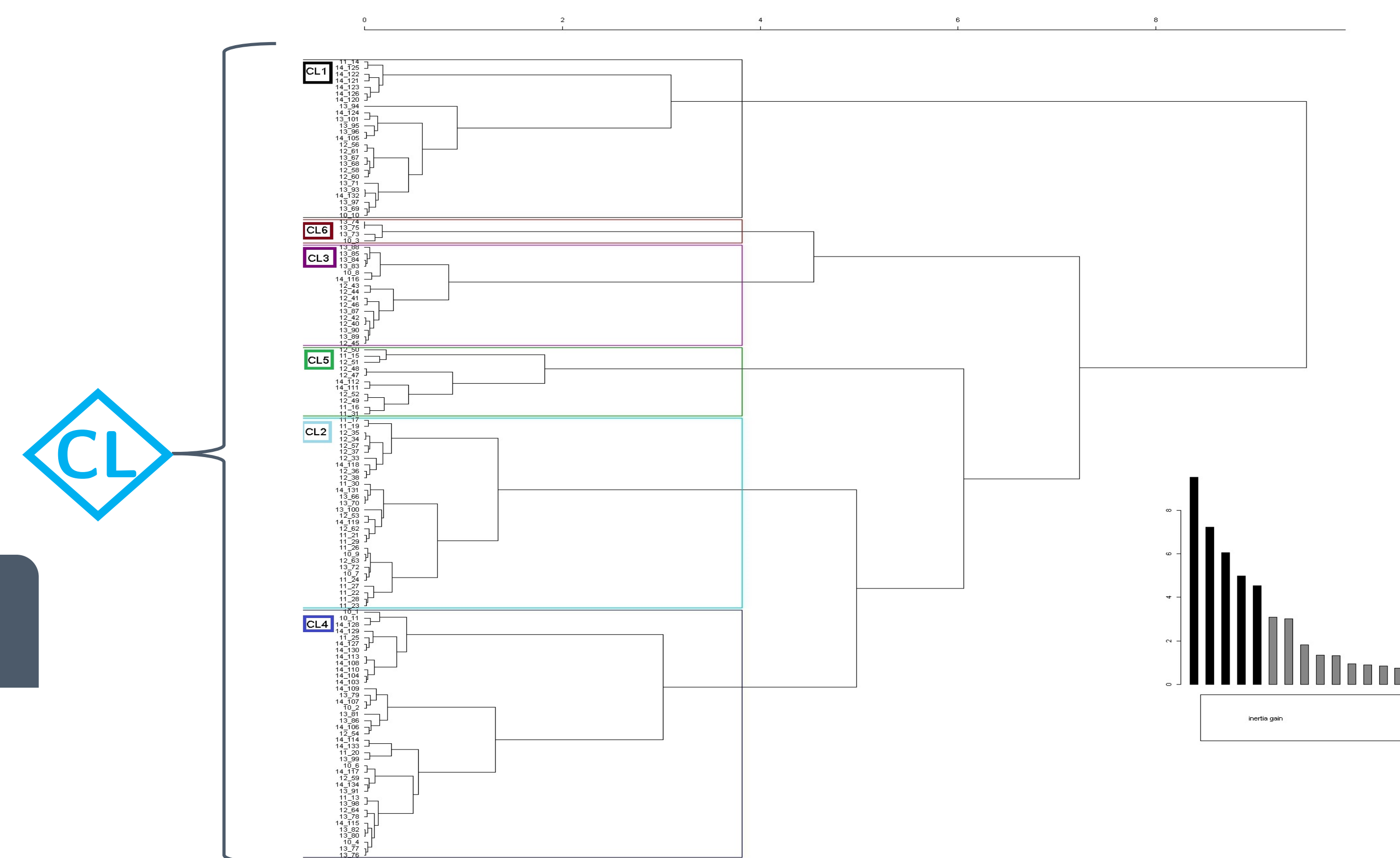
The initial meta/data file included (5) columns:

- Origin of the author (**Authr**): Au.T/Au.nT, Turkish or non-Turkish, reflected in author’s address information
- Scholarly or political orientation of the author (**Artcl**): Ar.Sc/Ar.P, based on declared affiliation
- Turkish or non-Turkish centered agenda (**Them**): ThT/Th.nT, primarily based on the inclusion of the term Turk* in the title of the article
- The adoption of the realist International Relations (IR) perspective or not (**Rl.nRI**): Re/nRe, based on the explicit or implicit acceptance (or, in reverse, critique) of IR realism or more broadly state centrism
- Orientation of the article, concerning the theory-praxis focus (theory oriented vs policy oriented – (**Thr.Plc**): Thry/Plcy, reflected in the degree of theoretical embeddedness or of the emphasis on policy-relevant issues and on empirical/historical analysis

ADDED CONTENT INFO

To incorporate IR related content in the metadata table we applied:

- Text Mining on the corpus of the articles for frequent IR terms
- Hierarchical Clustering on IR term frequencies (using the tf-idf method) producing 6 clusters of articles:



Thus, a new piece of qualitative information emerged, indicating article cluster by IR content, amending the metadata table as follows:

EXAMINED DATA

article	Authr.	Artcl.	Them.	Rl.nRI	Thr.Plc	CL
10.1	Au.nT	Ar.Sc	Th.nT	nRe	Thry	C5
10.2	Au.T	Ar.Sc	Th.nT	nRe	Thry	C4
10.3	Au.nT	Ar.Sc	Th.nT	Re	Plcy	C6
10.4	Au.nT	Ar.P	Th.nT	nRe	Thry	C4
10.6	Au.T	Ar.Sc	Th.T	Re	Plcy	C2
10.7	Au.T	Ar.Sc	Th.T	Re	Plcy	C2
:	:	:	:	:	:	:

Finally, MCA applied on the above data to produce groups of articles focusing on IR content as indicated in column CL above.

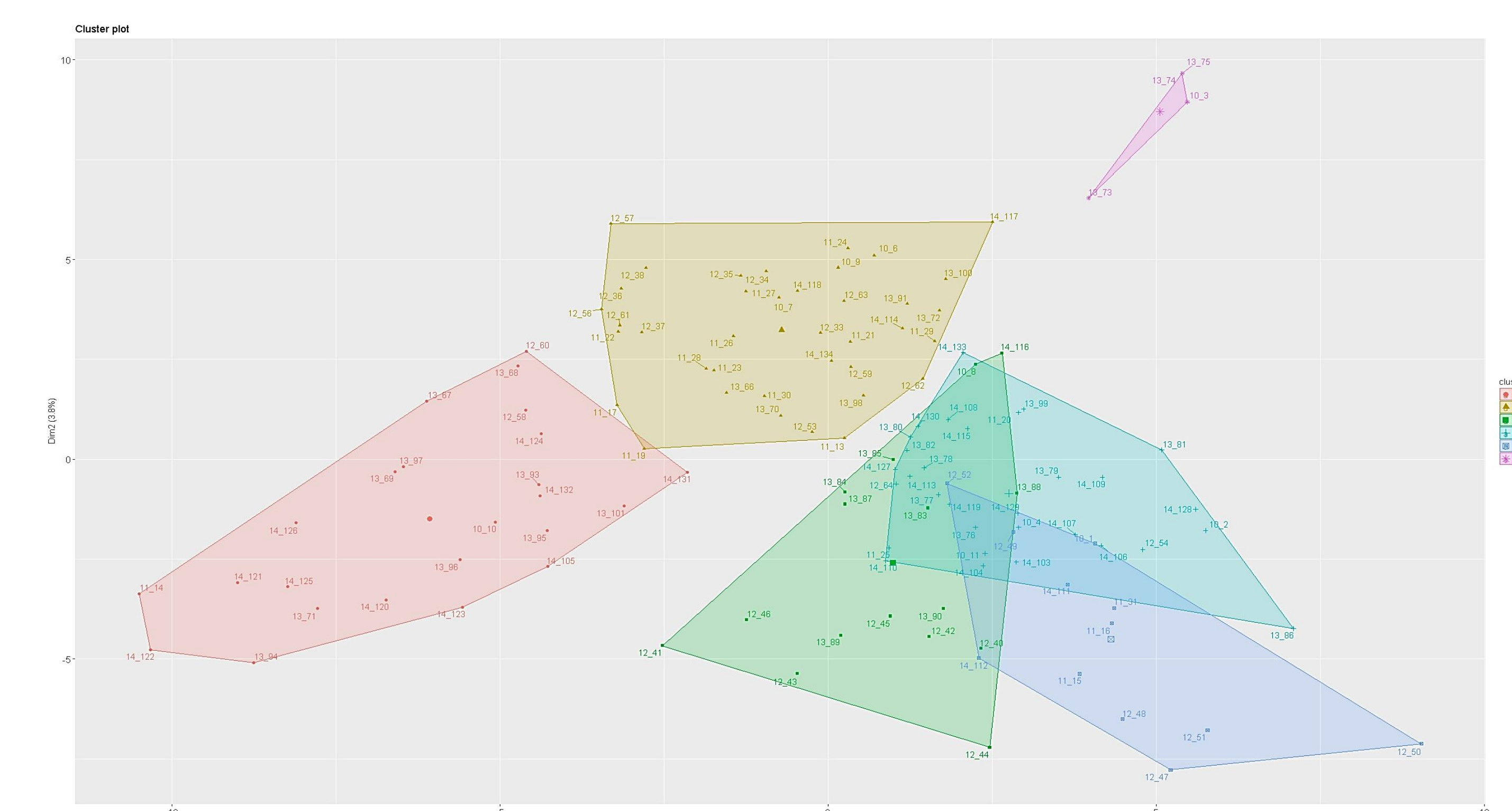
IN A NUTSHELL...

A combination of Text Mining and Multivariate Data Analysis Techniques can be a new approach to Document Clustering

For works cited and ... more, please visit:



MCA



The range of IR themes in journal articles seems more efficiently reflected when Document Clustering is combined with Multivariate Data Analysis techniques.

SOFTWARE USED

R / RStudio / R packages: FactoMineR, factoextra & tm

RELATED WORK

- Koutsoupias, N., & Mikelis, K. (2019). Exploring International Relations Journal Articles: A Multivariate Approach. SAGE Publications Ltd. DOI: <https://dx.doi.org/10.4135/9781526486189>
- Koutsoupias, N., & Mikelis K. (2019), The Peloponnesian War: Preview with N-grams and Correspondence Analysis, GSDA Conference Book of Abstracts, Thessaloniki.