

11th Scientific Meeting of the SIS Group
"Statistics for the Evaluation and Quality in Services"

BOOK OF **SHORT PAPERS**

Editors

Andrea Bucci

Alfredo Cartone

Adelia Evangelista

Andrea Marletta



**STATISTICAL METHODS
FOR EVALUATION AND QUALITY:
TECHNIQUES, TECHNOLOGIES AND TRENDS (T³)**

**IES 2023 - Statistical Methods for Evaluation and Quality:
Techniques, Technologies and Trends (T³)**

BOOK OF SHORT PAPERS

Editors: Andrea Bucci, Alfredo Cartone, Adelia Evangelista and Andrea Marletta

Book of Short papers
11th International Conference **IES 2023**
Statistical Methods for Evaluation and Quality: Techniques, Technologies and
Trends (T³)

University 'G. d'Annunzio' of Chieti-Pescara



Scientific Committee of the group of the Italian Statistical Society on Statistics for the Evaluation and Quality of Services - SVQS

Maurizio Carpita	University of Brescia
Pietro Amenta	University of Sannio
Matilde Bini	European University of Rome
Antonello D'Ambra	University of Campania 'L. Vanvitelli'
Luigi D'Ambra	University of Naples 'Federico II'
Paolo Mariani	University of Milan 'Bicocca'
Marica Manisera	University of Brescia
Monica Palma	University of Salento
Pasquale Sarnacchiaro	University of Naples 'Federico II'

Program Committee of the conference IES 2023

Chair: Paolo Mariani	University of Milan 'Bicocca'
Matilde Bini	European University of Rome
Mario Bolzan	University of Padova
Giuseppe Bove	University of Rome 3
Eugenio Brentari	University of Brescia
Ida Camminatiello	University of Campania 'L. Vanvitelli'
Maurizio Carpita	University of Brescia
Carlo Cavicchia	Erasmus University Rotterdam
Roy Cerqueti	University of Rome 'La Sapienza'
Enrico Ciavolino	University of Salento
Corrado Crocetta	University of Bari 'Aldo Moro'
Antonello D'Ambra	University of Campania 'L. Vanvitelli'
Pierpaolo D'Urso	University of Rome 'La Sapienza'
Tonio Di Battista	University 'G. d'Annunzio' of Chieti-Pescara
Simone di Zio	University 'G. d'Annunzio' of Chieti-Pescara
Mauro Ferrante	University of Palermo
Francesca Fortuna	University of Rome 3
Michele Gallo	University of Naples 'Orientale'
Claudiu Herteliu	Bucharest University of Economic Studies
Salvatore Ingrassia	University of Catania
Ron Kenett	Samuel Neaman Institute & KPA Group
Michele La Rocca	University of Salerno
Rosaria Lombardo	University of Campania 'L. Vanvitelli'
Antonio Lucadamo	University of Sannio
Marica Manisera	University of Brescia
Andrea Marletta	University of Milan 'Bicocca'
Lucio Masserini	University of Pisa
Fabrizio Maturo	University Mercatorum
Matteo Mazziotta	ISTAT
Rodolfo Metulini	University of Bergamo
Stefania Mignani	University of Bologna
Enrico Ripamonti	University of Brescia

Silvia Salini	University of Milano
Pasquale Sarnacchiaro	University of Naples 'Federico II'
Violetta Simonacci	University of Naples 'Federico II'
Maurizio Vichi	University of Rome 'La Sapienza'
Mariangela Zenga	University of Milan 'Bicocca'

Organizing Committee

Chair: Tonio Di Battista	University 'G. d'Annunzio' of Chieti-Pescara
Andrea Bucci	University 'G. d'Annunzio' of Chieti-Pescara
Alfredo Cartone	University 'G. d'Annunzio' of Chieti-Pescara
Marco Di Marzio	University 'G. d'Annunzio' of Chieti-Pescara
Marta Di Nicola	University 'G. d'Annunzio' of Chieti-Pescara
Simone Di Zio	University 'G. d'Annunzio' of Chieti-Pescara
Stefania Fensore	University 'G. d'Annunzio' of Chieti-Pescara
Lara Fontanella	University 'G. d'Annunzio' of Chieti-Pescara
Stefano Antonio Gattone	University 'G. d'Annunzio' of Chieti-Pescara
Luigi Ippoliti	University 'G. d'Annunzio' of Chieti-Pescara
Eugenia Nissi	University 'G. d'Annunzio' of Chieti-Pescara
Paolo Postiglione	University 'G. d'Annunzio' of Chieti-Pescara
Annalina Sarra	University 'G. d'Annunzio' of Chieti-Pescara
Pasquale Valentini	University 'G. d'Annunzio' of Chieti-Pescara
Adelia Evangelista (Secretariat)	University 'G. d'Annunzio' of Chieti-Pescara

Editors

Andrea Bucci University of Macerata
Alfredo Cartone University 'G. d'Annunzio' of Chieti-Pescara
Adelia Evangelista University 'G. d'Annunzio' of Chieti-Pescara
Andrea Marletta University of Milan 'Bicocca'



Il Viandante - Copyright © 2023 Edizioni BACME S.r.l.s.
Sede legale: Via Silvino Olivieri, 111 - 66100 Chieti (CH)

ISBN 979-12-803-3369-8

DOI 10.60984/978-88-94593-36-5-IES2023

<https://doi.org/10.60984/978-88-94593-36-5-IES2023>

All rights reserved.

This work is protected by copyright law.

All rights, in particular those relating to translation, citation, reproduction in any form, to the use of illustrations, tables and the software material accompanying the radio or television broadcast, the analogue or digital recording, to publication and dissemination through the internet are reserved, even in the case of partial use. The reproduction of this work, even if partial or in digital copy, is admitted only and exclusively within the limits of the law and is subject with the authorization of the publisher. Violation of the rules involves the penalties provided for by the law.

Sponsored by



Preface

Statistical thinking, design and analysis play a crucial role in social life and are useful to society at large. Besides, promoting advanced methodological research is useful to facilitate the dissemination of ideas related to various fields of interest. For this purpose, experts in statistics, data analysis, data mining, statistical methods for decision making, machine learning and related methods come together to understand and analyse phenomena through data.

In line with this objective, the Statistics Group for the Evaluation and Quality of Services (SVQS; www.svqs.it) of the Italian Statistical Society (SIS) has been organizing the Innovation and Society (IeS) conference biennially since 2009, focusing on new developments and ideas in statistics applied to the evaluation and quality of public and private services, attracting national and international statisticians and data scientists. The meeting contributes to spot light on the main statistical approaches and methodologies for the evaluation of public services currently in use in different contexts, as well as to facilitate discussion on the impact of innovative statistical evaluation systems for these services, involving various economic and social policy actors.

The conference “Statistical Methods for Evaluation and Quality: Techniques, Technologies and Trends (T³)” recorded valuable contributions that are reported in this volume. The papers underscore how the growing availability of data has tasked social and economic actors, organizations, and researchers with the management and analysis of large volumes of unstructured and heterogeneous data. In recent years, many tools for both qualitative and quantitative models have been developed to better describe and understand complex systems and their underlying behaviors, and the papers reported in this volume bear witness to this.

Techniques, technologies and trends: the study of data complexity presents the potential to provide analyses with increased frequency and timeliness, accuracy and objectivity, and to define sustainable models. Traditional quantitative methods for capturing socioeconomic data have often shown limitations in their ability to examine underlying systems, and with the three ‘T’ just mentioned, the outlines of future developments are starting to emerge.

The volume reports 127 contributions in the following areas:

- Advanced statistical methods for pattern recognition
- Advances in statistical learning from high-dimensional data
- Data analysis for web sources
- Distance and depth-based statistical learning methods for robust data analysis

- Economics and environment
- Education and labour
- Inequalities in the labour market
- Innovations and challenges in official statistics
- Labour market: trends, perspectives and new challenges
- Methodological and applicative contributions for evaluating sustainable development
- Methodological developments and applications for the assessment of student competencies
- Networks data analysis: new perspectives and applications
- New advanced statistical methods for data science
- Recent advances in statistical learning and data analysis
- Statistical analysis and modeling of environmental pollution data
- Statistical methods and complexity for evaluation in finance
- Statistical methods and composite indicators for healthcare
- Statistical methods and models for land monitoring with spatio-temporal data
- Statistical methods for environmental monitoring and sustainability
- Statistical methods for the analysis of university student choices and academic performance
- Statistical methods for the assessment of transport services and sustainable emissions
- Statistical methods for education and educational services
- Statistics in sports
- Tourism and territory.

The Conference event attracted many contributions as well as numerous Authors, not just from Italy but also from abroad. Over the three-day meeting, the Community has the opportunity to witness some of the state-of-the arts, new trajectories, and methodological challenges in 24 solicited sessions, 7 sessions of free contributes, two round tables - organized by Maurizio Vichi and Matilde Bini respectively - and three keynotes sessions with Ron S. Kennet of Samuel Neaman Institute of Israel, Luigi D'Ambra of Federico II University of Naples, and the former Minister Enrico Giovannini from University of Tor Vergata.

Organizers

Paolo Mariani

Chair of the Program Committee

Tonio Di Battista

Chair of Local Organizing Committee

Maurizio Carpita

Coordinator of the SVQS Group

Contents

Plenary Session	1
Befitting Cross Validation with Three Case Studies	2
<i>Kenett R.S.</i>	
Solicited Session SS1 - <i>Statistical analysis and modeling of environmental pollution data</i>	3
Assessing environmental quality by clustering a structural equation model based index: An application to European cities air pollution	4
<i>Bottazzi Schenone M., Grimaccia E. and Vichi M.</i>	
Evaluating the nonlinear association between PM ₁₀ and emergency department visits	10
<i>Bucci A., Sanmarchi F., Santi L., Giostra F., Tubertini E., Rosa S., Nante N. and Golinelli D.</i>	
Estimating spatially varying Gaussian Graphical Models to unveil relationships among pollutants in the Red River Delta in Vietnam	16
<i>Pronello N., Cucco A., Ignaccolo R. and Ippoliti L.</i>	
Solicited Session SS2 - <i>Statistics in sports</i>	22
Clustering Athlete Performances in Track and Field Sports	23
<i>Argiento R., Colombi A., Modotti L. and Montagna S.</i>	
A Cross-Country Analysis of Engagement in Physical Activity and Sport Practice Learnt from Eurobarometer Survey Data	29
<i>Simone R.</i>	
Strong eras in male professional tennis	35
<i>Breznik K., Candila V., Milekhina A. and Restaino M.</i>	
NonParametric Combination method for data analytics in basketball matches	41
<i>Barzizza E., Biasetton N., Ceccato R., Disegna M. and Vezzosi G.</i>	
Solicited Session SS3 - <i>Statistical methods for the analysis of university student choices and academic performance</i>	47
The influence of labor market conditions on students' career disruption: first insights from Italy	48
<i>Usala C., Sulis I. and Porcu M.</i>	
Socio-economic aspects that may affect South-North students' mobility in Italy	53
<i>Genova V.G. and Boscaïno G.</i>	
An analysis of student's performance in bachelor's degree	59

<i>La Rocca M., Niglio M. and Restaino M.</i>	
An exploratory strategy for analyzing students' mobility data	65
<i>Primerano I. and Giordano G.</i>	
Solicited Session SS4 - <i>Statistics in football</i>	71
Community Detection in Sport Market Networks: The Case of Italian Professional Football	72
<i>Rondinelli R. and Ievoli R.</i>	
An Original Application to Football of PLS-SEM for the xG Model	78
<i>Cefis M. and Carpita M.</i>	
Performance Assessment of Football Players and Combined Permutation Tests with application to Home-Field Advantage	84
<i>Bonnini S.</i>	
A First Proposal of the Triad Census for Weighted Networks: an Application to Football	90
<i>Rondinelli R. and Palazzo L.</i>	
Solicited Session SS5 - <i>Advances in statistical learning from high-dimensional data</i>	96
PCA approaches for vector functional time series	97
<i>Aguilera A.M., Alonso F.J. and Acal C.</i>	
Conformal Prediction for Functional Kriging Models	102
<i>Diana A., Romano E. and Adzic J.</i>	
Measuring Public-Private Connectedness in Financial Markets	108
<i>Sánchez García J. and Cruz Rambaud S.</i>	
An original approach to anomalies in intertemporal choices through Functional Data Analysis: Theory and application for the study of Hikikomori syndrome	112
<i>Martino R., Ventre V., Cruz Rambaud S. and Maturo F.</i>	
Solicited Session SS6 - <i>Labour market: trends, perspectives and new challenges</i>	118
Enriching Job Vacancy Official Information with Online Job Advertisements: Chances and Limits	119
<i>Lucarelli A. and Righi A.</i>	
Innovation in Management: towards the Open Manager	126
<i>Bruttini P., Gallo M., Mariani P. and Menini T.</i>	
Solicited Session SS7 - <i>Data analysis for web sources</i>	132
Enhancing SMEs default prediction with web-scraped data	133
<i>Crosato L., Domenech J. and Liberati C.</i>	
Web data as enabler for informed decisions in Labour Market	137
<i>Maggioni G.</i>	
The metaverse & luxury fashion brands: strategic communication exercise	141
<i>Forciniti A. and Zavarrone E.</i>	
Increasing the Geographical Granularity of Economic Indicators with Google Trends	147

Domenech J. and Marletta A.

Solicited Session SS8 - <i>Methodological and applicative contributions for evaluating sustainable development</i>	153
Evaluating sustainable development in EU countries through synthetic indicators	154
<i>Alaimo L.S. and Cucci M.</i>	
Naples and tourism sustainability: A survey of citizens' perceptions	160
<i>Aria M., Pagliara F., D'Aniello L. and Della Corte V.</i>	
Modelling inequalities for sustainable development in Italy countries	166
<i>Musella M., Borrata G., Camminatiello I. and Lombardo R.</i>	
Food Security and Sustainability: A Science Mapping Analysis	172
<i>Piscitelli A.</i>	
Solicited Session SS9 - <i>Inequalities in the labour market</i>	178
Skill similarities across Italian regions: an analysis based on the online job advertisements	179
<i>Kahlawi A., Buzzigoli L., Grassini L., Martelli C. and Giambona F.</i>	
Italian Labour Market reform and gender inequalities	185
<i>Marini C. and Nicolardi V.</i>	
Intergenerational transmission of disadvantages in the Italian labour market: evidence from AD-SILC data	191
<i>Busetta A., Fabrizi E., Ragozini G. and Sulis I.</i>	
Solicited Session SS10 - <i>Statistical methods and complexity for evaluation in finance</i>	197
Financial networks resilience and shocks propagation	198
<i>Cerqueti R., Cinelli M., Ferraro G. and Iovanella A.</i>	
How the choice of one parameter impacts the numerical stability of the efficient frontier	204
<i>Fassino C. and Uberti P.</i>	
Dynamic shrinkage for minimum variance combination of forecasts	210
<i>Mattera R.</i>	
Exploring the perception of the gender issue of Italian female entrepreneurs	217
<i>Castellano R., Riccioni J. and Rinaldi A.</i>	
Solicited Session SS11 - <i>Networks data analysis: new perspectives and applications</i>	223
Describing Italian mobility trajectories in higher education	224
<i>Genova V.G., Giordano G., Ragozini G. and Vitale M.P.</i>	
Collaboration networks: methodological issues and updated empirical evidence on Italian statisticians	230
<i>De Stefano D., Fabbrucci Barbagli A.G., Santelli F. and Zaccarin S.</i>	
Mapping Ashtma Complexity with Graph Theory: an Integrative Approach	236
<i>Cucco A., Simpson A., Murray C., Fontanella S. and Custovic A.</i>	
Investigating the patterns of Italian internal mobility: a network analysis at provincial level	242

Sarra A., D'Ingiullo D., Evangelista A., Nissi E., Quaglione D. and Di Battista T.

Solicited Session SS12 - <i>Innovations and challenges in official statistics</i>	248
Formal and informal networks of care for the elderly: regional profiles compared	249
<i>Sicuro L., Tucci D. and Coniglio R.</i>	
Gender Gap: a multidimensional approach	255
<i>Acampora C., Fusco D., Liguori M.A. and Pagliuca M.M.</i>	
Using Whatsapp in Official Statistics: a New tool for managing the Agriculture Census	259
<i>Fabi C.</i>	
Solicited Session SS13 - <i>Statistical methods and composite indicators for healthcare</i>	265
Longitudinal composite indicators to measure the quality of health services	266
<i>Crocetta C., Antonucci L., Cataldo R. and Mazza R.</i>	
Past and Future of Doctor-Patient Communication	272
<i>Tedesco N., Zavarrone E. and Forciniti A.</i>	
Network Analysis approach to customer satisfaction and service quality detection: an application to health-care services	277
<i>Crocetta C., Grassia M.G., Marino M., Mazza R., Simonacci V. and Stavolo A.</i>	
A project evaluation study on multiset Likert scale data	283
<i>Simonacci V., Marino M., Grassia M.G. and Gallo M.</i>	
Solicited Session SS14 - <i>Distance and depth-based statistical learning methods for robust data analysis</i>	289
Robust distance-based predictive models	290
<i>Boj E., Grané A. and Parron D.</i>	
Data depth for mixed-type data through multidimensional scaling. An application to biological age imputation	294
<i>Cascos I., Grané A. and Qian J.</i>	
A compared protocol to improve clustering procedures	298
<i>Grané A., Riani M. and Salini S.</i>	
Robust diagnostics for Linear Mixed Models with the Forward Search . . .	304
<i>Corbellini A., Grossi L. and Laurini F.</i>	
Solicited Session SS15 - <i>Advanced statistical methods for pattern recognition</i>	310
Unsupervised classification of NPLs recovery curves	311
<i>Carleo A. and Rocci R.</i>	
Living alone in Italian municipalities	317
<i>Vellucci P., Benassi F., Naccarato A. and Gallo G.</i>	
Supervised learning from high-dimensional data through dynamic updating of functional classification rules	323

<i>Maturo F., Fortuna F. and Di Battista T.</i>	
Assessing the effectiveness of coordination among public authorities in cohesion expenditure	329
<i>Coco G., Monturano G. and Resce G.</i>	
Solicited Session SS16 - <i>Recent advances in statistical learning and data analysis</i>	335
A Predictive Functional Principal Component Analysis of Well-Being Data	336
<i>Marcis L., Pagliarella M.C. and Salvatore R.</i>	
Detecting the partition in the extended hierarchy of a dendrogram: an application on biomedical data	343
<i>PolICASTRO V., Palazzo L. and VISTOCCO D.</i>	
Concordance measure for rankings	350
<i>Bissiri P.G. and Nai Ruscone, M.</i>	
Quadratic discriminant scoring for selecting clustering solutions	355
<i>Coraggio L. and Coretto P.</i>	
Solicited Session SS17 - <i>Statistical methods for education and educational services</i>	361
Association between INVALSI scores and students' mobility in Italy: a preliminary analysis	362
<i>Bacci S., Bertaccini B., Lombardi G. and Tocchioni V.</i>	
Modelling Responses and Response Times: an application to Mathematics INVALSI data	368
<i>Bungaro L., Desiderati R. and Mignani S.</i>	
Latent potential outcomes: An analysis of the effects of programs aimed at improving students' non-cognitive skills	374
<i>Pennoni F., Bartolucci F. and Vittadini G.</i>	
Cognitive Skills and Non Cognitive Skills to Analyze School and Students Performances	380
<i>Vittadini G.</i>	
Solicited Session SS18 - <i>Statistical methods for the assessment of transport services and sustainable emissions</i>	386
Sustainability assessment of urban transport by an LCA comparison on different technologies vehicles	387
<i>Della Ragione L. and Meccariello G.</i>	
Passenger comfort prediction via time-series classification	393
<i>Vanacore A., Pellegrino M.S. and Ciardiello A.</i>	
A statistical model to analyse driving behavior: a case study	399
<i>Rodia G., Sarnacchiaro P. and Acciarino V.</i>	
Aggregating judgments in non negotiable group decisions in transport system	403
<i>Amenta P. and Lucadamo A.</i>	
Solicited Session SS19 - <i>New advanced statistical methods for data science</i>	408

A unified framework for two-dimensional clustering on preference-approvals: an analysis of Eurobarometer data	409
<i>Albano A., Sciandra M. and Plaia A.</i>	
Pandemic Data Quality Modelling: A Bayesian Approach	415
<i>Ferrari L., Manzi G., Micheletti A., Nicolussi F. and Salini S.</i>	
Explainable AI for Peer-to-Peer Credit Risk Management	421
<i>Babaei G., Pagnottoni P. and Do T. T.</i>	
Tackling misclassification in surveys about undeclared work via the EM algorithm	427
<i>Arezzo M.F., Guagnano G. and Vitale D.</i>	
Solicited Session SS20 - <i>Tourism, territory and data analysis</i>	432
Tourism, sustainability, and territorial impact: an input-output analysis . .	433
<i>Garau G. and El Meligi A.K.</i>	
The Impact of Big Data in Tourism	439
<i>Ciuffreda R., Choedon C. and Simonetti B.</i>	
Neural network-based prediction of domestic tourists' length of stay in Italy	443
<i>Antolini F. and Cesarini S.</i>	
The management of cultural heritage in contexts of undertourism: a model for assessing the economic sustainability of public-private partnerships	450
<i>Calabrò F.</i>	
Solicited Session SS21 - <i>Statistical methods and models for land monitoring with spatio-temporal data</i>	457
Geo-referenced data and complex networks for measuring road accident risk	458
<i>Cantaluppi G., Clemente C., Della Corte F. and Zappa D.</i>	
A comparison of geospatial models for car crash risk	464
<i>Cantaluppi G., Giardino G. and Zappa D.</i>	
Geostatistical modelling of livestock-related PM _{2.5} pollution and scenario analysis for policymakers - Work in progress	470
<i>Fassò A., Rodeschini J., Fusta Moro A. and Finazzi F.</i>	
Functional clustering methods for space-time big data from mobile phone networks	476
<i>Perazzini S., Metulini R. and Carpita M.</i>	
Solicited Session SS22 - <i>Methodological developments and applica- tions for the assessment of student competencies</i>	483
Modeling the main drivers of mathematical literacy of school-leaving stu- dents. Some evidence from the Invalsi tests	484
<i>Davino C., Palumbo F., Romano R. and Vistocco D.</i>	
Educational Data Mining: clustering students' performance over time . . .	490
<i>Taraborrelli G. and Farnè M.</i>	
The nexus of cultural capital with participation in early childhood education	496
<i>Ripamonti E.</i>	
High- and Low-Performing students and future career: a gender and social issue	502
<i>Falzetti P. and Ricci R.</i>	

Solicited Session SS23 - <i>Statistical methods for environmental monitoring and sustainability</i>	508
Clustering spatial data through optimal transports	509
<i>Balzanella A. and Verde R.</i>	
New interpretative insights for environmental air quality by means of FDA	514
<i>Terzi S., Naccarato A. and Fortuna F.</i>	
A Bayesian State-Space Model to Mitigate Unmeasured Confounding . . .	520
<i>Zaccardi C., Valentini P. and Ippoliti L.</i>	
Mining social media data for damage assessment in environmental disasters	526
<i>del Gobbo E., Cafarelli B., Ippoliti L. and Fontanella L.</i>	
Solicited Session SS24 - <i>Satisfaction and behavior in tourism</i>	531
The evaluation of the hotel stay through a new development of correspondence analysis coping with ordinal variables	532
<i>D'Ambra A. and Amenta P.</i>	
Assessing the role of knowledge and authenticity in the formation of attendee loyalty at cultural festivals	536
<i>Rivetti F., Lucadamo A. and Rossi C.</i>	
Residents' Opinions and Perceptions of Tourism Development in the Historic City of Matera	542
<i>Sarnacchiaro P., Di Gennaro R. and Di Taranto E.</i>	
Exploring tourism at religious sites: The case of Assisi	547
<i>Rivetti F., Dini M. and Splendiani S.</i>	
Session of free contributes SFC1 - <i>Education and labour</i>	552
Local concordance among the items of questionnaires on student's opinion (OPIS)	553
<i>Terzi S. and Petrarca F.</i>	
High School Proficiency of Future University Students: An Analysis based on INVALSI Data	560
<i>Santelli F., Di Credico G. and Di Caterina C.</i>	
Employment vulnerability of immigrants in the labour market – Does origin matter?	566
<i>Bittaye M.</i>	
The effect of pricing policies on students' use of university canteens	572
<i>Masserini L., Bini M. and Lorenzoni V.</i>	
Gig workers' identikit	576
<i>Zavarrone E. and Forciniti A.</i>	
Session of free contributes SFC2 - <i>Tourism and territory</i>	582
Multi-source approach for statistics in tourism sector	583
<i>Bianchino A., Fusco D., Giordano P., Liguori M.A. and Summa D.</i>	
Statistical analysis of tourism sustainability in Campania: post Covid-19 review	589
<i>Giacalone M., Basile V. and Bellucci M.</i>	
Investigating recent changes in dietary behavior	596
<i>D'Uggento A., d'Ovidio F.D., Toma E. and Onorati M.G.</i>	

Depopulation in the Abruzzo municipalities	603
<i>Carulli A.L., Di Spalatro D. and Valentini A.</i>	
The Productions System of Inland Areas	607
<i>Madia Carucci A.M. and Regano A.</i>	
Session of free contributes SFC3 - <i>Social issues</i>	613
Modelling the gender gap in youth mortality with an Age-Period-Cohort analysis	614
<i>Lanfiuti Baldi G. and Nigri A.</i>	
Random forest for classifying odor emission sources	620
<i>Distefano V., Palma M., De Iaco S. and Mazuruse G.</i>	
An Experimental Annotation Task Investigating Annotator Agreement Within a Misogynistic Dictionary and Corpus	626
<i>Tontodimamma A., Ignazzi E., Anzani S., Fontanella L. and Di Zio S.</i>	
Statistical analysis of COVID19 impact on Italian mortality	632
<i>Franchetti G. and Politano M.</i>	
Measuring multidimensional deprivation using objective and subjective data: an application of the Voronoi ranking method	638
<i>Ciommi M., Mariani F., Polinesi G. and Recchioni M.C.</i>	
Session of free contributes SFC4 - <i>Trends</i>	644
The role of big data analytics in circular supply chains: A bibliometric analysis	645
<i>Khan F. and Rapposelli A.</i>	
Estimation of the ranking of incentive policies for the adoption of 4.0 tech- nologies	652
<i>Bonnini S. and Borghesi M.</i>	
Risk Management and Future Scenarios. A proposal based on a mixed- method approach	658
<i>Di Zio S., Bolzan M., Marozzi M. and Scioni M.</i>	
Explainable artificial intelligence (XAI) through artificial intelligence from a human in the loop (HITL) perspective: an interview with ChatGPT	664
<i>Santarcangelo V., Lamacchia A., Vitullo S., Di Lecce M. and Giacalone M.</i>	
Relevance in official statistics: information needs, satisfaction with data quality, some results and future perspectives	670
<i>Loporcaro M.F.</i>	
Session of free contributes SFC5 - <i>Economic issues</i>	676
Evolutionary trends of start-ups in Italy: a case study	677
<i>Dutillo P., Caruso G., Iannone B. and Gattone S.A.</i>	
Permanent establishments and efficiency analysis with global enterprises .	682
<i>Frenda A. and Sepe E.</i>	
Techniques and constructs in some recent market and organizational research	685
<i>Sciascia I.A.</i>	
The value of buildings in the Italian general government balance sheet . . .	689
<i>Santoro P. and Regano A.</i>	

Linear and nonlinear factors affecting default risk in the peer-to-peer lending market	695
<i>Giordano F., Milito S. and Parrella M.L.</i>	
Session of free contributes SFC6 - <i>Methodological issues</i>	701
Macroeconomic Time Series Classification by Nonparametric Trend Estimation	702
<i>Feo G., Giordano F., Niglio M. and Parrella M.L.</i>	
A Normalization Method for Space-time Analysis of Evaluation and Quality Indicators	709
<i>Mazziotta M. and Pareto A.</i>	
Unveiling Latent Structures: exploring Multidimensional IRT Models using Dirichlet Process Mixtures	713
<i>Valentini P., Fontanella S. and Fontanella L.</i>	
On a technique to detect accounting data manipulation	718
<i>Passamonti C.</i>	
Session of free contributes SFC7 - <i>Economics and environment</i>	724
Determinants of Water Conservation Behaviour and Spatial Heterogeneity in their Coefficients	725
<i>Mammadli R. and Gigliarano C.</i>	
Modeling the economic burden of grass pollen allergoid immunotherapy . .	731
<i>Bilancia M. and Di Bona D.</i>	
Swine fever in Liguria: who does pay for economic losses? A causal analysis	735
<i>Baggetta C., Cavalletti B. and Corsi M.</i>	
Measuring Multidimensional Poverty of the Italian Regions in the era of COVID-19	741
<i>Chelli F.M.C., Ciommi M., Gigliarano C. and Polinesi G.</i>	
Drivers of inflation: relationships changing over time	747
<i>Sokolenko O., Palumbo A., Fortuna F., Naccarato A. and Marie J.</i>	
Environmental accounting and sustainable cities: an explorative bibliometric-based literature analysis	753
<i>Kaur A., Javed A., D'Andreamatteo A. and Rapposelli A.</i>	

Living alone in Italian municipalities

Vivere da soli nei comuni italiani

Pierluigi Vellucci, Federico Benassi, Alessia Naccarato and Gerardo Gallo

Abstract New family structures have emerged in Italy in recent decades, with a trend towards smaller nuclear families due to demographic and social changes. An aging population, marital instability, declining fertility, and later marriage have contributed to this trend. It is important to understand the changing needs of families, especially the vulnerable, from both an economic and social perspective. Vulnerability is often related to economic factors, but single-person households, are often at risk. The goal of this study is to classify Italian municipalities based on single-member household characteristics, identifying areas of greater or lesser fragility.

Abstract *Negli ultimi decenni sono emerse in Italia nuove strutture familiari, la cui caratteristica principale è la riduzione del numero di componenti e l'aumento del numero di famiglie mono-personali. Fattori come l'invecchiamento della popolazione, l'instabilità matrimoniale, la diminuzione della fertilità e il matrimonio più tardivo hanno contribuito a questa evoluzione. In ottica di policies, bisogna osservare che i nuclei familiari composti da una sola persona, sono maggiormente esposti a situazioni di vulnerabilità. Lo scopo di questo studio è classificare i comuni italiani in base alle caratteristiche dei nuclei familiari monocomponenti, identificando le aree di maggiore o minore fragilità.*

Pierluigi Vellucci

Dept. of Economics, Roma Tre University, Via Silvio D'Amico 77, 00145 Rome, Italy, e-mail: pierluigi.vellucci@uniroma3.it

Federico Benassi

Dept. of Political Sciences, University of Naples Federico II, Via Leopoldo Rodinò 22, 80133 Naples, Italy, e-mail: federico.benassi@unina.it

Gerardo Gallo

Italian National Institute of Statistics, Via Cesare Balbo 16, 00184 Rome, Italy, e-mail: gegallo@istat.it

Alessia Naccarato

Dept. of Economics, Roma Tre University, Via Silvio D'Amico 77, 00145 Rome, Italy, e-mail: alessia.naccarato@uniroma3.it

Key words: living alone people, decision algorithm, geographical pattern

1 Introduction

In the last decades new forms of families spread in Italy. The combination of both social and demographic changes resulted in a families' nuclearization process, manifested by the progressive decrease in the average number of members [1]. There are many demographic factors that have contributed most to this process. The aging of the population, which has led many elderly people to live alone, and the marital instability. These drivers are compounded by declining fertility [8], longer length of stay in the family of origin by young adults [3], and higher age of first marriage [5]. From the perspective of both economic and social planning, it is necessary to understand how the needs of families have changed, especially for the most vulnerable ones. The vulnerability of families is generally related to economic dimensions [2] however, people living alone are more likely to experience fragile situations, particularly in cases of elderly person. The aim of this contribution is to classify the Italian municipalities according to the characteristics of single-member households and to study their spatial distribution. This allows us to highlight territories of greater or lesser fragility and identify in which territories it is most important to intervene.

2 Data description

In order to classify Italian municipalities according to the types of single-member households, data from the 2020 Italian Permanent Population and Housing Census (PPHC) on households residing in Italy's 7903 municipalities were used. The characteristics of the single-member households refer to age (AGE), gender (GEN), citizenship (CIT), ownership of one or more dwellings (DWE), ownership of one or more cars (CAR), educational (EDU), whether they receive citizenship income or not (CIN). Moreover, we consider six characteristics of municipalities, geographic macro-area (GA), region (REG), degree of urbanization (DU), number of beds available in elderly housing facilities (BED), local wealth (WEA), and finally a categorical variable that groups municipalities according to their demographic profiles and attractiveness (ATT).

3 Methods

We employ the labelled dataset described in Sect. 2 to identify, by means of a decision algorithm, the amount of municipalities presenting a "critical" percentage of single-person households. In this first experiment we consider as "critical" those

municipalities having percentage greater than the average percentage computed for Italy. Only municipalities so identified were classified according to the prevalent single-member household type (PSMH). The target variable, which our supervised learning engine tries to predict, is the status of being a *critical municipality*. If we denote with i the municipality of our dataset and with y_i the ratio between the number of single-person households in 2020 and the total amount of population in the same year in the municipality i , a municipality i is called *critical* if y_i is larger than the average value of y_i over the entire dataset.

3.1 The decision algorithm

Decision algorithms [6] are commonly used in machine learning and data mining tasks to build predictive models. As a criterion for the splitting we will use the Gini impurity. The Gini impurity measures the probability of incorrectly classifying a randomly chosen element from the set. For a binary classification problem, where there are two classes (0 and 1), the Gini impurity can be defined as:

$$Gini(p) = 1 - (p_0^2 + p_1^2)$$

where p_0 and p_1 represent the probabilities of each class in the subset of data being considered. The cost function quantifies the quality of a split based on the reduction in impurity achieved. A common cost function is the weighted sum of the impurities of the resulting subsets after the split.

$$Cost = \frac{N_{\text{left}}}{N} \cdot Gini_{\text{left}} + \frac{N_{\text{right}}}{N} \cdot Gini_{\text{right}}$$

where N_{left} and N_{right} are the number of instances in the left and right subsets, N is the total number of instances, and $Gini_{\text{left}}$ and $Gini_{\text{right}}$ are the Gini impurities of the left and right subsets, respectively.

3.2 Spatial analysis

The decision algorithm leads us to the identification of 11 PSMHs on which a first spatial analysis has been done. Using the 11 PSMH typologies, we built a categorical map at municipality level. This map allows us to reconstruct the (local) spatial distribution of each typology and therefore to verify the existence (or absence) of specific geographical patterns. In a second step, for each municipality, we computed and mapped the location quotient [7]. The location quotient (LQ) is used to measure and map relative distributions or relative concentrations of a character in a subarea compared to the area as a whole [4], [9]. LQ varies from 0 to ∞ . It is lower than 1 in the spatial units in which a population group is underrepresented or conversely,

gle elderly individuals, while in the majority of cases they experience situations of vulnerability.

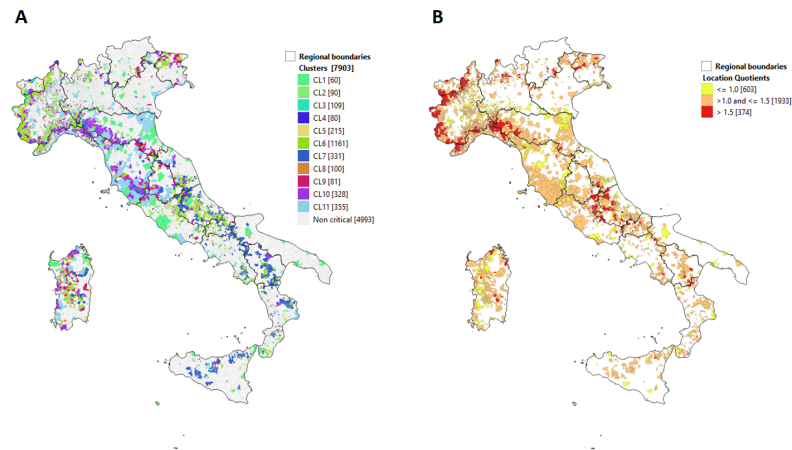


Fig. 2: Categorical map of clusters' typologies (panel A) and LQs map (panel B)

Looking at the two maps of Figure 2 we can infer some interesting (albeit partial) results. The distribution of single-member household typologies does not follow the classic south-north divide that characterised most of the socio-economic processes in Italy. Nevertheless, if we look to the distribution of the single typology some geographical patterns emerge. For a sake of brevity, we recall here just a few: the big cities are divided in two clusters (CL 1 and CL11). The first includes Rome, Bologna, and some other important cities of the Southern Italy like Palermo, Bari Naples, and Reggio di Calabria. The second holds Milan, Genoa, Venice, Turin, and some other metropolitan cities located in the Centre (Florence) and the South (Cagliari, Catania, Messina). These two urban clusters are contrasted by clusters composed of rural and more marginal municipalities like the CL6 and CL7. The first is composed by municipalities that are mainly located on the border of the Italian country especially on the North-West border (Piedmont region) and in some inner areas (mainly located in Abruzzo, Molise, and Sardinia regions). CL 7 is composed by isolated and quite marginal municipalities that are located in majority in the South part of Italy. It is interesting to note a quite different level of "mixing" of each region. It is quite evident that in some regional contexts (i.e. Sardinia and Tuscany) we find different typologies of clusters so that there isn't a typology that is clearly predominant. In other region (i.e. Lazio and Apulia) the typologies of clusters are very few. The LQs map (panel B of Figure 2) reveals some other interesting details. The condition of a severe over representation is related to a small number of municipalities (374 over 2.910). Nevertheless, their geographies is quite peculiar: they are concentrated in specific areas of the country defining some spatial patterns. One is located on the Northwest border, another is located at the intersection of three

regions (Emilia-Romagna, Liguria and Piedmont) while the last is again located at the intersection of three regions (Lazio, Marche and Abruzzi).

5 Conclusion

This is a first explorative analysis that shows from one hand the potentialities of using different data source (coming from official statistics and other sources) at very fine geographical scale. And, on the other, the usefulness in the classification of municipalities. These last allowed the identification of specific spatial patterns of certain typologies (or categories). This also results in focusing on particularly critical areas (i.e., municipalities where the level of over representation is comparatively high, $LQ > 1.5$). Results tell us that there is no clear north-south dynamic which is, in a certain manner, a novel aspect in the socio-economic and demographic literature. However, some patterns emerge: 1) a distinction between large cities and the other municipalities; 2) area of internal marginalisation; 3) a different level of heterogeneity inside of each regional context. The study needs further development such as to modify the initial classification criteria, i.e. considering the quantiles of the target variable to identify "critical" municipalities, and to measure the level of inner heterogeneity of each region using specific indices.

References

1. AISP – Associazione Italiana per gli Studi di Popolazione: Rapporto sulla popolazione, Le famiglie in Italia. Forme, ostacoli, sfide. Eds. Vignoli, D., Tomassini, C., Il Mulino, 10-34, (2023)
2. Benassi, F., Naccarato, A.: Households in potential economic distress. A geographically weighted regression model for Italy, 2001-2011, *Spat. Stat.*, 21, 362-376 (2017)
3. Billari, F. C., Manfredi, P., Valentini, A.: Macro-demographic effects of the transition to adulthood: Multistate stable population theory and an application to Italy, *Math. Popul. Stud.*, 9(1), 33-63 (2000)
4. Crawley A., Beynon M., Munday M.: Making location quotients more relevant as a policy aid in regional spatial analysis. *Urban Stud.*, 50(9), 1854-1869 (2013)
5. Gabrielli G., Vignoli D.: The Breaking-Down of Marriage in Italy: Trends and Trendsetters, *Popul. Rev.*, 52(1) (2013)
6. Hastie, T., Tibshirani, R., Friedman, J. H., Friedman, J. H.: The elements of statistical learning: data mining, inference, and prediction. Vol. 2. Springer, New York (2009)
7. Isard W.: Methods of regional analysis: an introduction to regional science. Cambridge: The MIT Press (1960)
8. Kertzer, D. I., White, M. J., Bernardi, L., Gabrielli, G.: Italy's Path to Very Low Fertility: The Adequacy of Economic and Second Demographic Transition Theories, *Eur. J. Popul.*, 25(1), 89-115 (2009)
9. Wheeler J.O.: Geography. In: Kimberly K.L. (Ed.) *Encyclopedia of Social Measurement*, Elsevier, 115-123 (2005)