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

IES PESCARA 2023

STATISTICAL METHODS FOR EVALUATION AND QUALITY: TECHNIQUES, TECHNOLOGIES AND TRENDS (T^3)

EDIZIONI Il Viandante

IES 2023 - Statistical Methods for Evaluation and Quality: Techniques, Technologies and Trends (T^3)

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 (T^3)

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'

EDIZIONI IlViandante

IlViandante - 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

Sponsored by

Farabella GLUTEN FREE

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 (T^3)" 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

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

municipalitieas having percentage grater 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:

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

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 categorial 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 that 1 in the spatial units in which a population group is underrepresented or conversely,

overrepresented (LQ>1). In our analysis the population group is the typology of single member households in each municipality. The identification of the cut offs in the distribution of LQ is an open issue. In our exercise we defined 3 classes: a first (LQ $\leq$ 1.0) in which there is absence of over representation; a second (1.0 $<$ LQ $\leq$1.5) in which there is a moderate over representation; and a third (LQ $>$ 1.5) characterised by a high over representation. Maps have been done by using QGIS Desktop version 3.20.2.

4 Results and discussion

In Figure 1 we depict the visual representation of the decision tree illustrating the classification of municipalities into 'critical' or 'non critical' categories based on key predictor variables. From Figure 1 we can spot the presence of 11 classes of

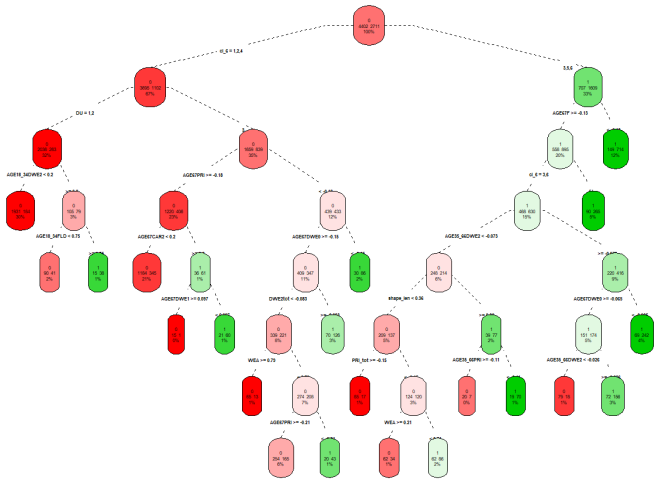

Fig. 1: Decision Tree for the full set of features according to the Gini index

critical municipalities. They can be summarized as follows[1].

Due to brevity constraints, we cannot provide a detailed description of the classes, but we can highlight some key findings. With the exception of one class that includes young individuals with a university education and home ownership, single-person households are predominantly made up of elderly individuals. A smaller number of municipalities are characterized by economically and socially stable sin-

[1] Please refer to the following legend. Degree of urbanization classifies municipalities into Cities (1), Small towns or suburbs (2) and Rural areas (3). The attractiveness variable refers to the systematic depopulation of Italian municipalities as described by [1].

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

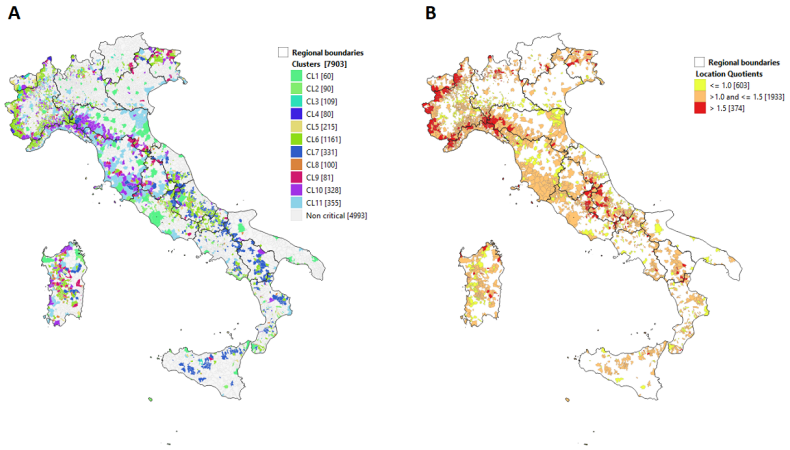

Fig. 2: Categorial map of clusters' typologies (panel A) and LQs map (panel B)

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 indeces.

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)