TITLE PAGE

A fuzzy approach for analysing equitable and sustainable well-being in Italian regions

Annamaria Porreca,
Department of Economics, G. d'Annunzio University of Chieti-Pescara, Pescara, Italy.
annamaria.porreca@unich.it

Salvador Cruz Rambaud (CORRESPONDING AUTHOR)
Department of Economics and Business, Calle Universidad de Almería, Almería, Spain.
scruz@ual.es

Francesca Scozzari
Department of Economics, G. d'Annunzio University of Chieti-Pescara, Pescara, Italy.
francesca.scozzari@unich.it

Marta Di Nicola
Department of Medical Oral Science and Biotechnology, G. d'Annunzio University of Chieti-Pescara Chieti, Chieti, Italy.
marta.dinicola@unich.it

A fuzzy approach for analysing equitable and sustainable well-being in Italian regions

ABSTRACT

Objective: Recently, the Italian Institute of Statistics (ISTAT) and the National Council for Economy and Labor (CNEL) proposed a measure for the equitable and sustainable well-being called the BES ("Benessere Equo e Sostenibile"). This paper aims to propose an original application of the fuzzy *k*-means approach to providing an analysis of the Italian regions according to their BES.
Methods: The fuzzy *k*-means algorithm was used for clustering the Italian regions according to BES data 2015. Afterwards, a principal component analysis was conducted to show and interpret the results.
Results: There is a clear difference between the regions of the north and the south. The only exceptions are represented by Lazio and Abruzzo, which belong to both groups with almost equal degrees of truth. Moreover, Trentino Alto-Adige and Valle d'Aosta exhibit the best condition whilst Molise is the worst region.
Conclusions: This study reveals that some Italian regions are in a state of backwardness regarding health, environment, minimum economic conditions, subjective well-being, education, employment conditions, social relationships, and working conditions. Therefore, institutions should consider local policies to address these issues.
Keywords: BES, well-being, fuzzy *k*-means, inequalities, Italian regions.

INTRODUCTION

For several decades, the Gross Domestic Product (GDP) has been considered as the most relevant measure to assess well-being and countries' development (Esterlin 1974; Kubiszewski et al. 2013, Felice 2016). In the last decades, this indicator has been used as a metric for assessing the living standards of people and, in general, their quality of life. However, a high level of GDP does not automatically imply that people are happy, satisfied, and have a good well-being condition, compared with those living in a country with a lower GDP (Marmot 2002; Vanoli 2009). In truth, well-being also depends on the freedom of choice in the political and social sphere, as well as on its perception, healthcare access, education, and many other domains (Graziani 1981; Dasgupta and Weale 1992; Burchi and Gnesi 2015). The economic condition is undoubtedly a key variable in determining people's quality of life because it affects the capacity to have a more stable and pleasant standard of living.
Moreover, greater financial opportunities favour access to better services of health, education, and social relationships, by increasing the state of global well-being (Brandolini and Vecchi 2013; Inshc and Florek 2008; Di Spalatro et al. 2017; Bellomo et al. 2017). Currently, in the industrialized countries, and particularly in Italy, there is a profound debate on the necessity of considering the development of a nation not only in merely numerical terms of GDP, but also based on much broader criteria, which are able to reflect and translate into practice a more authentic and all-encompassing index of citizens' well-being (D'Acci 2010; Oishi and Kesebir 2015). In Italy, the ISTAT has shown significant interest in the topic of well-being in several ways.
One of the reasons is certainly the economic situation of the country. "Italy is arguably the only Western country where regional imbalances still play a major role nowadays" (Felice 2017). Indeed, Italy is divided into two parts: A Center-North much more homogeneous internally, and a poorer South. Moreover, the effectiveness of state intervention in the South is almost absent due to growing political clientelism, bad industrial choices and organised crime (Felice 2017). At the same time, the regions of central Italy have always held an ambiguous position; from a historical perspective, they have been assimilated both to the North, in terms of economic development, and to the South, when dealing with socio-cultural aspects (Felice and Vasta 2015; Felice 2018).

This ambiguous position of the regions located in central Italy suggests that the gap between North and South is an erroneous convention which forces to a crisp classification (into two macro-groups) which does not correspond to the reality. Therefore, starting from the Sen's perspective of well-being (Sen 1993), we aim understanding if the fuzzy *k*-means approach may detect areas presenting nuances in the conditions of well-being.

In 2010, the ISTAT, jointly with the CNEL, has conducted a research program aimed at creating a "Steering Committee on measuring the progress of Italian society" (ISTAT 2016). The final objective of this initiative is to build a set of 134 indicators of Equitable and Sustainable Well-being (BES). This set of indices, grouped into 12 domains (or dimensions), aims to evaluate the progress of communities by considering, as well as economic aspects, some environmental factors whose definitions are based on fundamental criteria of equity, social, and collective sustainability.

Our findings can be useful for policy-makers to monitor their local-area conditions, assess well-being inequalities within Italy, and promote human health through organised efforts and informed choices of the society, with the involvement of public and private organisations, communities, and individuals.

The paper is structured as follows. Section 2 includes the description of the dataset and a summary of the fuzzy *k*-means algorithm. Part 3 displays our results, whereas Section 4 presents the discussion. The paper ends with the conclusions regarding our proposed research.

METHODS

Data

The variables used in this study are collected from the censuses carried out by the Italian National Statistical Institute and explicitly organised in the BES 2016 report (ISTAT 2016). This report provides an integrated picture of the main economic, social and environmental phenomena which have characterised the recent evolution of Italy through the analysis of a broad set of indicators, divided into 12 domains. Starting from 2010, these domains have been implemented and developed by the ISTAT. Moreover, 134 indicators have been identified to find those aspects with a direct impact on human and environmental well-being and those measuring functional elements to improve the well-being of the community and environment which surrounds it. From 2015 onwards, the BES report also proposes synthetic measures for assessing the overall trend of several domains (Oishi and Kesebir 2015; Davino et al. 2016; Calcagnini and Perugini 2018). These allow the aggregation of some indicators, summarising a domain into a single value. The method used for the calculation of these composite indicators guarantees territorial and temporal comparability. The composite indicators were developed only for nine outcome domains, i.e. those having a direct impact on well-being. For these reasons, entire domains (policy and institutions, research and innovation, and quality of services) or individual indicators were excluded from the calculation. For the interpretation of the data, the ISTAT assumes that the starting values on the composite domain indicators, at 2010, are equal to 100. To guarantee the comparability of the results at a spatial and temporal level, ISTAT uses COMIC, a generalized software for the synthesis of indicators, which provides fair values, that is purified by the measurement unit (ISTAT 2016). Table 1 shows the nine domains considered by the BES 2016 report and their respective description.

[INSERT TABLE 1 ABOUT HERE]

The fuzzy *k*-means algorithm

Cluster analysis techniques are often used in applied statistical studies to search for hidden structures and groups within large datasets. Recently, this methodology has also been widely applied in health studies (Lefèvre et al. 2014; Liao et al. 2016); however, all studies focus on a classic multivariate approach to clustering. In summary, clustering methods can be classified as hard clustering (or exclusive clustering) or soft clustering (overlapping clustering). The main limit of the former technique is that data is grouped exclusively so that if a unit belongs to a defined group, it cannot be included in another cluster (Bezdek 1981; Bora and Gupta 2014). On the contrary, in the latter method, the groupings are made so that a statistical unit can belong to several groups with a different degree of belonging. In many situations, this type of technique appears more natural. In fact, the objects on the borders between several groups are not forced to belong to one of the groups fully but rather degrees of belonging in the interval [0, 1] are assigned.

In details, the fuzzy k-means method is an unsupervised classification algorithm proposed by Bezdek (Bezdek 1981). It is the most widely used fuzzy classification criteria and is an extension of the crisp k-means method. In the initial phase of the procedure, similar to the crisp version of the algorithm, the researcher decides the number of clusters c in which he aims to classify the n units with p variables. A validation criterion based on the use of specific tests can be used to select the optimal number of groups. The algorithm proceeds iteratively through the minimisation of an objective function, from which a fuzzy classification is obtained where, for each unit, it is determined the degree of membership to the c groups. The degree of membership of the i units to the k group, denoted by u_{ik}, satisfies the following constraints:

$$0 \leq u_{ik} \leq 1$$

and

$$\sum_{k=1}^{c} u_{ik} = 1,$$

where $i = 1,\ldots,n$ and $k = 1,\ldots,c$. We denote as J_m the objective function to minimise used to calculate the optimal values of the degrees of membership. It depends on both the distance, d_{ik}, between the i-th unit and the centroid of the k-th group, and the parameter m which adjusts the level of fuzziness. Thus,

$$J_m(U, v) = \sum_{k=1}^{c} \sum_{i=1}^{n} u_{ik}^{m} d_{ik}^{2},$$

where $d_{ik} = |x_i - v_k|$, $x_i \in R^p$is the i-th component of units vector, $v_k \in R^p$ is the k-th component of the centroid vector, U is the matrix of the degree of membership of dimension $n \times c$, and $m \in [1, +\infty)$. Thus, the variables on which the minimization is performed are the cluster centers and the degrees of membership. The objective function J_m measures the quadratic error by which the n units are represented within the centroids; of course, it depends on how the units are arranged in the groups and measures the dispersion of them around the centers. The optimal partition is that minimizing J_m. The value of the parameter m to be chosen represents the degree of fuzziness, i.e. how the resulting partition must be blurred. Since this method represents a generalization of the classical k-means method, it generally presents the same type of problems which mainly consist in the difficulty of choosing the initial number of groups. Another main question concerns the choice of the value of the parameter m which regulates the level of fuzziness. The empirical applications implemented with the fuzzy k-means method have shown that there is no optimal value for the m parameter, but it can vary according to the different contexts.

The Silhouette plot is used for the interpretation and validation of the consistency within the clusters of data while the principal component analysis (PCA) is proposed to obtain a graphical representation of our results.

All the analyses are performed by using the R statistical environment (version 3.4.3). The descriptive statistics are achieved using the "sjPlot" package (Lüdecke 2018). To choose the optimum number of groups, the "PC" (partition coefficient), "MPC" (modified partition coefficient) and "SIL.F" (fuzzy silhouette index) functions contained in the R package "Fclust" are used (Bezdek 1981; Ferraro and Giordani 2015). The "fanny" (Maechler et al. 2018) function, which is in the R package "cluster" is used for the fuzzy clustering, and the "PCA" function, located in the "FactorMineR" R package, is adopted (Le et al. 2008).

RESULTS

Table 2 shows the descriptive statistics of the nine BES domains assessed for 20 Italian regions by considering the 2015 data. The reference year is 2010, i.e. the indicators for this year are all equal to 100. First, we can observe that the indicators of education, health, and environment have had a noticeable improvement over the last five years. On the contrary, we can note that the subjective well-being perceived by individuals has worsened considerably in recent years. Moreover, the unemployment domain is characterised by a very high variability (18.45) which indicates that there are several regions where the unemployment rate is very high.

[INSERT TABLE 2 ABOUT HERE]

Figure 1 shows the value of the three clustering indices (PC: partition coefficient, MPC: modified partition coefficient, and SIL.F: Fuzzy silhouette) according to the number of groups. Precisely, these three are calculated from 2 to 10 groups, confirming that the classification with fuzzy clustering must be performed into two groups. Indeed, the three statistics to identify the optimal number of groups reach their maximum in correspondence of a number of groups equal to 2.

[INSERT FIGURE 1 ABOUT HERE]

Table 3 indicates the degrees of membership of each region to the two clusters. Naturally, in harmony with the logic of fuzzy clustering, each region belongs to both groups with a certain degree of truth. From a first analysis of the table, the areas of Northern Italy have a very high degree of membership to cluster 1 while the regions of Southern Italy and the islands are better represented by cluster 2. Lazio and Abruzzo have a similar degree of membership to both clusters. Table 3 also shows the Normalized Dunn's Coefficient. This index ranges from 0 to 1 and, in case of high values, we can state the existence of a very crisp clustering. In this case, we obtain a value equal to 0.36 from which we can conclude that the use of this fuzzy clustering well represents the observed phenomenon.

[INSERT TABLE 3 ABOUT HERE]

Figure 2 illustrates the Silhouette values for each region. The amplitude of the Silhouette is used to study the goodness of the adaptation of the fuzzy cluster. It provides a measure of the closeness of each unit to units in the near group which can vary between −1 and 1. A coefficient close to 1 indicates that the unit is very far from the neighbouring cluster, while a value of 0 means that the unit is very close to the next cluster and, therefore, we are in a very nuanced situation. Lombardia, Veneto, Emilia Romagna, Piemonte, Toscana and Friuli Venezia-Giulia have a very high value of the silhouettes demonstrating the fact that they are well classified in group 1. For group 2, the highest values are found for Puglia, Campania, and Calabria. We can observe that all the silhouettes have positive values except Abruzzo. At the same time, even the silhouette of Lazio has a very low value if compared to the other regions. This result emphasises that these two regions are in an intermediate position. In any case, the preliminary statistics confirm that the optimal number of groups is two. This circumstance justifies the non-insertion of a third group.

[INSERT FIGURE 2 ABOUT HERE]

Figure 3 shows the results of the clustering by using the first two principal components (PCs) and the relationship between the PCs and the original dimensions. Figure 3b) shows that all dimensions are positively correlated with the first PC where the second PC is positively correlated with the environment, subjective well-being, and social relationships. On the contrary, it is negatively

correlated with all the other domains. Therefore, the second PC can be substantially seen as the perception of people in evaluating the quality of the environment in which they live and partly as subjective well-being. Instead, the first PC can be seen as the satisfaction of people for all the dimensions regarding the economy, education, and health.
Figure 3a) shows that Trentino Alto Adige and Valle d'Aosta are the regions with the highest well-being. Abruzzo and Lazio are located in an intermediate position, but Abruzzo is clearly above the mean concerning the perception of the environment. Molise and Tuscany are characterized by a low quality of the environment while Sicilia, Campania, and Calabria are characterized by low perceptions of economic well-being, education, and health. The regions of the Northern Italy, except Valle d'Aosta and Trentino Alto Adige, are very close to each other considering both PCs.

[INSERT FIGURE 3 ABOUT HERE]

DISCUSSION

The literature displays a comprehensive discussion about the capacity of GDP to represent people well-being (Stewart 2005; Mohammed and Ghebreyesus 2018). This is also emphasised by the Sen's perspective (Sen 1993; Bérenger and Verdier-Chouchane 2007; Chakravarty 2017; De Rosa 2017), which stresses that well-being is a multidimensional concept and encompasses many aspects besides the traditional economic indicators.
In this paper, we focused on the Italian context because previous studies have provided a crisp image of Italy, i.e. a nation divided into two parts, the North as an area rich in well-being, and the South as a backward area (Cerqueti and Ausloos 2014; Felice 2017; Felice 2018). For this reason, the ISTAT has created a nationwide survey, called the BES, for the measurement of some dimensions which make up the broader concept of well-being (Roche 2008). The goal of this research is to consider a multidimensional approach to subjective well-being, and understand if the Italian regions should really be interpreted as two broad groups (Franzini and Giannoni 2010).
Because well-being is a vague concept, we propose a novel approach based on a fuzzy classification. Effectively, offering a study of Italian regions through a crisp approach would make little sense because one cannot consider only the extreme cases of well-being or the lack of well-being in an absolute sense since intermediate situations are possible. The results confirm this idea because there are Italian regions that, according to the perceived well-being, are in an intermediate position between the two main groups. The objective of our analysis is to provide useful indications to policy-makers which are differentiated according to each considered dimension. In this regard, we propose a classification to identify similar patterns between Italian regions based on the 2016 BES report.
Our research highlights that there are some regions in an intermediate position despite their geographical location. For example, Abruzzo geographically belongs to southern Italy, but partly fits the first group with a high degree of truth. This circumstance is because a high quality of the environment characterises Abruzzo. Lazio is another region in an intermediate position between the two groups. However, differently, from Abruzzo, this is due to other dimensions such as health, income, and occupation. The location of these two regions justifies the use of this type of approach without which they would have been forced to belong to the South or North group without highlighting their peculiarities. Trentino-Alto Adige and Valle d'Aosta have a high quality of life. Abruzzo has a high quality of well-being regarding the environment. Even in Calabria, the perception of well-being related to the environment is above the national mean, but it is not accompanied by the good quality of the other dimensions. Very similar patterns of behaviour characterize Liguria, Umbria, Piemonte, Friuli Venezia-Giulia, Marche, and Lombardia. Emilia Romagna and Toscana are characterized by poorly perceived well-being related to the environment. Molise, Basilicata, and Sicily are below the average for both the quality of the environment and the rest of the dimensions. We stress that also a classical *k*-means approach has been made to compare

our results. According to the traditional technique, Lazio and Abruzzo should be forced to belong to group 1 whereas the other regions would be assigned as shown in Table 3.
We believe that our approach could have public health policy implications and help to look for options that would potentially aid to reduce the observed inequalities in well-being. However, research on the determinants on regional inequalities in Italy still has a long road ahead (Felice 2016). Mainly, neither geography (e.g., market size) nor human and social capital seem to be the crucial determinants in the long-run. Clearly, the problem of detecting the most important predictors of regional disparities is strongly connected to the identification of a proper measure to replace GDP and to better represent people well-being.
In our opinion, the search for the perfect measure of well-being is trivial because the latter is a multidimensional concept and cannot be synthesized by a single index. Therefore, we propose the fuzzy *k*-means approach as a complementary indicator to help policy-makers in assessing groups with similar patterns of disparities. According to this perspective, public health policy could identify sets of regions with similar drawbacks and needs, and thus it should be possible to program possible joint intervention in broader areas (not only at regional level). In effect, some aspects connected with the quality of life are not always limited within regional boundaries, and it does make sense to think about macro-areas of interventions according to specific clustering results.

CONCLUSIONS

Well-being is by nature a fuzzy concept, and thus the proposed approach does not find its foundations in the Italian context but seems the most natural approach to the phenomenon under study. Indeed, in economic and social science, fuzzy set theory has been proposed in many studies with different purposes (see, e.g., Maturo 2016; Eze and Onasanya 2018). We believe that the "fuzzy" information can be beneficial for institutions to promote regional development policies for those regions showing below-average well-being conditions. In effect, Sicilia, Calabria, Basilicata, Molise, and Sardegna need interventions aimed at improving their economic conditions and health. On the other hand, Molise, Toscana, Emilia Romagna, Lombardia, and Basilicata certainly need some interventions to improve the perceived quality of the environment and the subjective well-being. The final aim of our study is to provide policy-makers with useful indications to program targeted interventions for the improvement of well-being conditions within the Italian regions and greater awareness of where to carry out these interventions to reduce costs and waste.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. The authors also declare that they have read and approved the journal's regulation regarding the ethical responsibilities and this research is carried out in compliance with the journal's ethical standards.

Funding: This study was funded by the Spanish Ministry of Economy and Competitiveness (grant number DER2016-76053R). We are very grateful for the comments and suggestions offered by two anonymous referees and the Associate Editor.

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

1. Bellomo RG, Tripodi D, D'Ercole S et al. (2017) Quality of life: Effects of physical activity in an anthropometric, cognitive and psychosocial background, and variation of odontoiatric parameters. Science & Philosophy 5(2):3-18. doi:10.23756/sp.v5i2.371
2. Bérenger V, Verdier-Chouchane A (2007) Multidimensional measures of well-being: Standard of living and quality of life across countries. World Development 35(7):1259–1276. doi:10.1016/j.worlddev.2006.10.011
3. Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Springer, Boston, MA. doi:10.1007/978-1-4757-0450-1
4. Bora DJ, Gupta AK (2014) A Comparative study between fuzzy clustering algorithm and hard clustering algorithm. International Journal of Computer Trends and Technology 10(2):108–113. doi:10.14445/22312803/ijctt-v10p119
5. Brandolini A, Vecchi G (2013) Standards of living. In: Toniolo G (ed) The Oxford handbook of the Italian economy since unification. Oxford University Press, Oxford, pp. 227–48
6. Burchi F, Gnesi C (2015) A review of the literature on well-being in Italy: A human development perspective. Forum for Social Economics 45(2-3):170–192. doi:10.1080/07360932.2014.995197
7. Calcagnini G, Perugini F (2018) A well-being indicator for the Italian provinces. Social Indicators Research 142(1):149–177. doi:10.1007/s11205-018-1888-1
8. Cerqueti R, Ausloos M (2014) Statistical assessment of regional wealth inequalities: the Italian case. Quality and Quantity 49(6):2307–2323. doi:10.1007/s11135-014-0111-y
9. Chakravarty SR (2017) Analyzing multidimensional well-being. John Wiley & Sons, Inc, Hoboken, New Jersey, United States. doi:10.1002/9781119257424
10. D'Acci L (2010) Measuring well-being and progress. Social Indicators Research 104(1):47–65. doi:10.1007/s11205-010-9717-1
11. Dasgupta P, Weale M (1992) On measuring the quality of life. World Development 20(1):119–131. doi:10.1016/0305-750x(92)90141-h
12. Davino C, Dolce P, Taralli S, Vinzi VE (2016) A quantile composite-indicator approach for the measurement of equitable and sustainable well-being: A case study of the Italian provinces. Social Indicators Research 136(3):999–1029. doi:10.1007/s11205-016-1453-8
13. De Rosa D (2017) Capability approach and multidimensional well-being: The Italian case of BES. Social Indicators Research 140(1):125–155. doi:10.1007/s11205-017-1750-x
14. Di Spalatro D, Maturo F, Sicuro L (2017) Inequalities in the provinces of Abruzzo: A comparative study through the indices of deprivation and principal component analysis. Studies in Systems, Decision and Control 104:219–231. doi:10.1007/978-3-319-54819-7_15
15. Easterlin R (1974) Does economic growth improve the human lot? Some empirical evidence. In: David, R. and Reder, R., Eds., Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz, Academic Press, New York. 89–125. doi:10.1016/b978-0-12-205050-3.50008-7
16. Eze M, Onasanya B (2018) Case studies on the application of fuzzy linear programming in decision-making. Ratio Mathematica 35:29–45. doi:http://dx.doi.org/10.23755/rm.v35i0.424
17. Felice E (2016) The misty grail: The search for a comprehensive measure of development and the reasons for GDP primacy. Development and Change 47(5):967–994. doi:10.1111/dech.12257
18. Felice E (2017) The roots of a dual equilibrium: GDP, productivity and structural change in the Italian regions in the long-run (1871–2011). SSRN Electronic Journal 40. doi:10.2139/ssrn.3082184
19. Felice E (2018) The socio-institutional divide: explaining Italy's long-term regional differences. The Journal of Interdisciplinary History 49(01):43–70. doi:10.1162/jinh_a_01231

20. Felice E, Vasta M (2015) Passive modernization? The new human development index and its components in Italy's regions (1871–2007). European Review of Economic History 19(1):44–66. doi:10.1093/ereh/heu018
21. Ferraro MB, Giordani P (2015) A toolbox for fuzzy clustering using the R programming language. Fuzzy Sets and Systems 279:1–16. http://dx.doi.org/10.1016/j.fss.2015.05.001
22. Franzini L, Giannoni M (2010) Determinants of health disparities between Italian regions. BMC Public Health 10(1):296. doi:10.1186/1471-2458-10-296
23. Graziani A (1981) Regional inequalities in Italy. Disparities in economic development since the industrial revolution. In: Bairoch P, Lévy-Leboyer M (eds) Disparities in economic development since the industrial revolution. Palgrave Macmillan, London, pp. 319–330. doi:10.1007/978-1-349-04707-9_29
24. Insch A, Florek M (2008) A great place to live, work and play. Journal of Place Management and Development 1(2):138–149. doi:10.1108/17538330810889970
25. ISTAT (2016) Il benessere equo e sostenibile in Italia. Accessed 25 July 2018. https://www.istat.it/it/files/2016/12/BES-2016.pdf
26. Kubiszewski I, Costanza R, Franco C et al (2013) Beyond GDP: Measuring and achieving global genuine progress. Ecological Economics 93:57–68. doi:10.1016/j.ecolecon.2013.04.019
27. Le S, Josse J, Husson F (2008) FactoMineR: An R package for multivariate analysis. Journal of Statistical Software 25(1):1-18. doi: 10.18637/jss.v025.i01
28. Lefèvre T, Rondet C, Parizot I, Chauvin P (2014) Applying multivariate clustering techniques to health data: the 4 types of healthcare utilization in the Paris metropolitan area. PLoS ONE 9(12):1–20. doi:10.1371/journal.pone.0115064
29. Liao M, Li Y, Kianifard F, Obi E, Arcona S (2016) Cluster analysis and its application to healthcare claims data: A study of end-stage renal disease patients who initiated hemodialysis. BMC Nephrology 17(1). doi:10.1186/s12882-016-0238-2
30. Lüdecke D (2018) sjPlot: Data visualization for statistics in social science. doi: 10.5281/zenodo.1308157, R package version 2.6.1, https://CRAN.R-project.org/package=sjPlot.
31. Maechler M, Rousseeuw P, Struyf A et al (2018) Cluster: Cluster analysis basics and extensions. R package version 2.0.7-1.
32. Marmot M (2002) The influence of income on health: Views of an epidemiologist. Health Affairs 21(2):31–46. doi:10.1377/hlthaff.21.2.31
33. Maturo F (2016) Dealing with randomness and vagueness in business and management sciences: the fuzzy-probabilistic approach as a tool for the study of statistical relationships between imprecise variables. Ratio Mathematica 30(1):45-58. doi:http://dx.doi.org/10.23755/rm.v30i1.8
34. Mohammed AJ, Ghebreyesus TA (2018) Healthy living, well-being and the sustainable development goals. Bulletin of the World Health Organization 96(9):590–590A. doi:10.2471/blt.18.222042
35. OECD (2015) Beyond GDP per capita: Other policy objectives: European Union. Organisation for Economic Co-Operation and Development. http://dx.doi.org/10.1787/growth-2015-graph75-en
36. Oishi S, Kesebir S (2015) Income inequality explains why economic growth does not always translate to an increase in happiness. Psychological Science 26(10):1630–1638. doi:10.1177/0956797615596713
37. Roche JM (2008) Monitoring inequality among social groups: A methodology combining fuzzy set theory and principal component analysis. Journal of Human Development 9(3):427–452. doi:10.1080/14649880802236706
38. Sen A (1993) Capability and well-being. In: Nussbaum M, Sen A (eds) The quality of life. Clarendon Press, Oxford, pp. 30–53. doi:10.1093/0198287976.003.0003

39. Stewart K (2005) Dimensions of well-being in EU regions: Do GDP and unemployment tell us all we need to know? Social Indicators Research 73(2):221–246. doi:10.1007/s11205-005-2922-7
40. Vanoli A (2010) On the report by the commission on the measurement of economic performance and social progress (2009). SSRN Electronic Journal. doi:10.2139/ssrn.1714428

Domain	Variables' name	Description
Health	HEALTH	It describes objective, functional, and subjective health.
Education and training	EDUC	It considers the participation in kindergartens, levels of education, school dropouts and continuing education.
Employment	EMPLOY_QUAL	It considers work participation and job satisfaction.
Quality of work	WORK_QUAL	It considers labour quality indicators (e.g. fixed-term employment, lower remuneration than the median value, and irregular employment)
Income	INCOME	It is about income, wealth, and consumptions.
Minimum economic conditions	EC_CONDIT	It evaluates the material deprivation, very low intensity of family work, quality of houses, and difficulty to have enough money for personal consumptions.
Social relations	SOC_REL	It deals with the individual well-being regarding relationships.
Satisfaction for life	SUB_W-B	It describes the cognitive dimension, i.e. the process by which everyone evaluates (in terms of "satisfaction") his life in a retrospective way.
Environment	ENVIR	It regards pollution, biodiversity protection, energy, and citizens' perception.

Table 1. *The nine outcome domains of the Report on Equitable and Sustainable Well-Being (BES: "Benessere Equo e Sostenibile"). Country: Italy. Year: 2016.*

Observation	*n*	Mean	SD	Median	Min.	Max.
SUB_W-B	20	90.37	11.25	89.96	68.88	120.62
EDUC	20	108	9.77	109.52	87.39	125.92
HEALTH	20	103.49	8.75	103.83	86.33	121.31
ENVIR	20	105.74	6.98	105.66	90.23	121.25
INCOME	20	99.59	12.47	102.36	74.44	116.75
EC_CONDIT	20	94.99	10.93	98.99	70.33	106.91
WORK_QUAL	20	93.54	11.58	97.8	69.73	108.33
EMPLOY_QUAL	20	99.8	18.45	109.79	67.24	122.64
SOC_REL	20	99.63	12.01	103.34	79.14	125.77

Table 2. *Descriptive statistics of the 20 Italian regions. The values shown are purified by the unit of measurement. Data are gathered from the Report on the Equitable and Sustainable Well-being (BES: "Benessere Equo e Sostenibile"). Country: Italy. Year: 2016.*

Degree of membership - Fuzzy clustering			
Membership coefficients (in %, rounded)			
	[Cluster 1]	[Cluster 2]	Group membership according to the classical *k*-means (non-fuzzy)
Piemonte	89	11	1
Valle d'Aosta	77	23	1
Liguria	84	16	1
Lombardia	89	11	1
Trentino-Alto Adige	69	31	1
Veneto	88	12	1
Friuli-Venezia Giulia	84	16	1
Emilia-Romagna	88	12	1
Toscana	87	13	1
Umbria	79	21	1
Marche	77	23	1
Lazio	56	44	1
Abruzzo	45	55	1
Molise	29	71	0
Campania	19	81	0
Puglia	14	86	0
Basilicata	23	77	0
Calabria	19	81	0
Sicilia	22	78	0
Sardegna	26	74	0
Fuzziness coefficients			
Normalized Dunn's coefficient			
0.36			

Table 3. *The degree of membership and normalized Dunn's coefficient. The fuzzy clustering is performed on the data gathered from the Report on the Equitable and Sustainable Well-Being (BES: "Benessere Equo e Sostenibile"). Country: Italy. Year: 2016.*

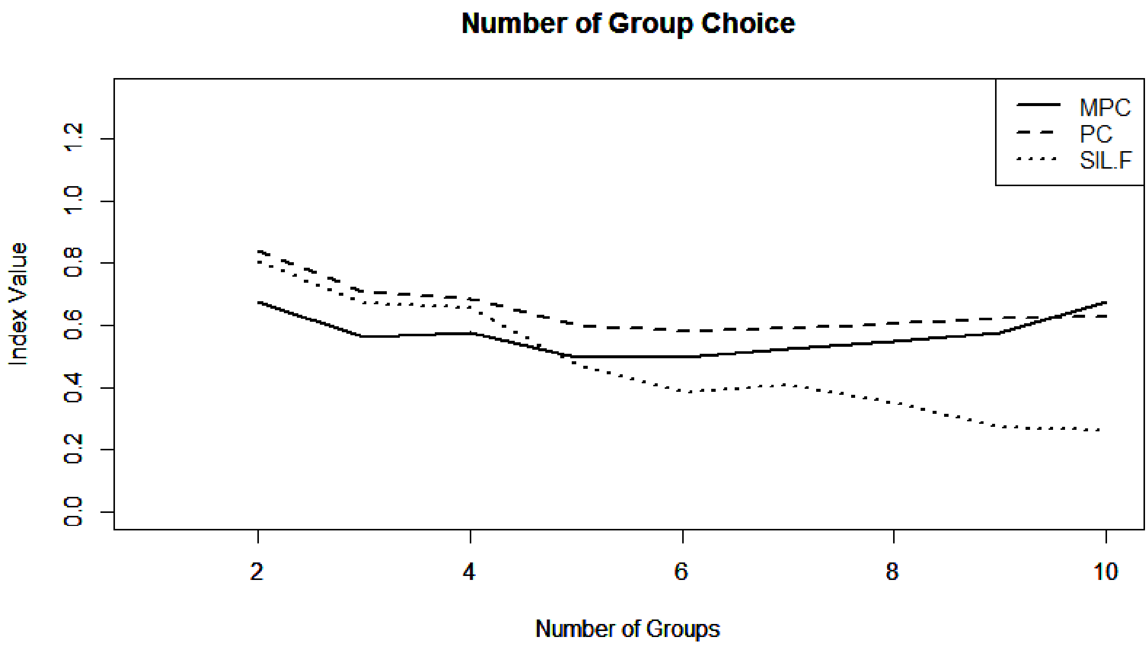

Figure 1. *PC (partition coefficient), MPC (modified partition coefficient), SIL.F (Fuzzy silhouette), according to the number of groups. The tests are performed on the data gathered from the Report on the Equitable and Sustainable Well-Being (BES: "Benessere Equo e Sostenibile"). Country: Italy. Year: 2016.*

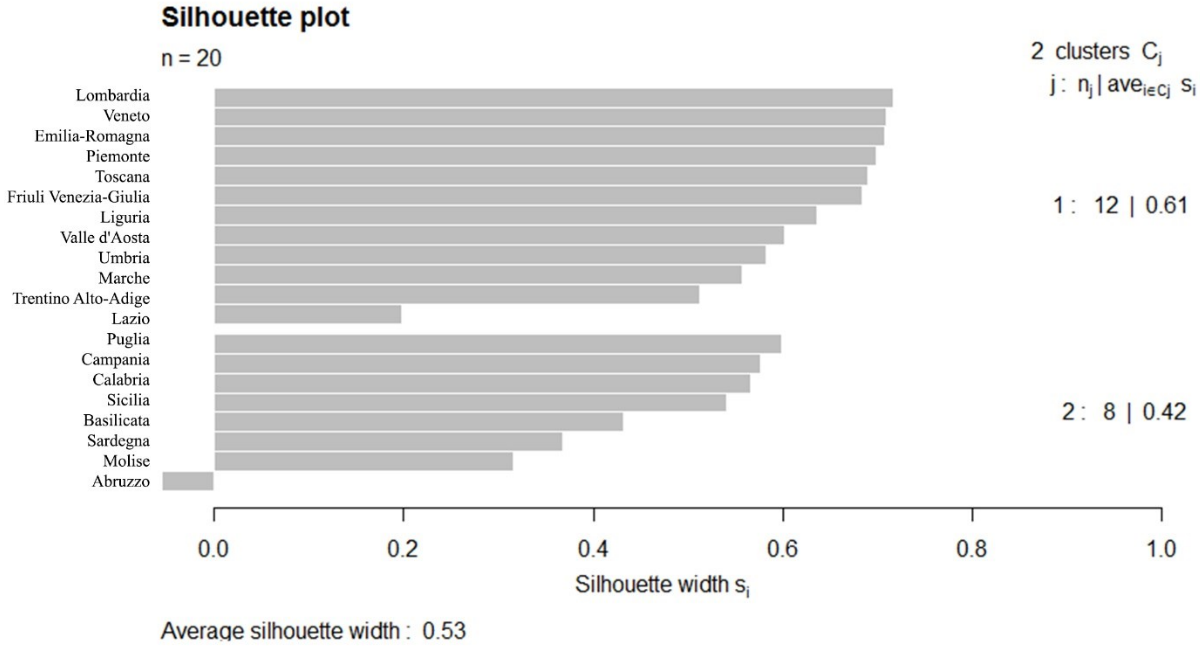

Figure 2. *The plot of Silhouette for the 20 Italian regions. The test is performed after clustering the regions using the fuzzy k-means approach and the data gathered from the Report on the Equitable and Sustainable Well-Being (BES: "Benessere Equo e Sostenibile"). Country: Italy. Year: 2016.*

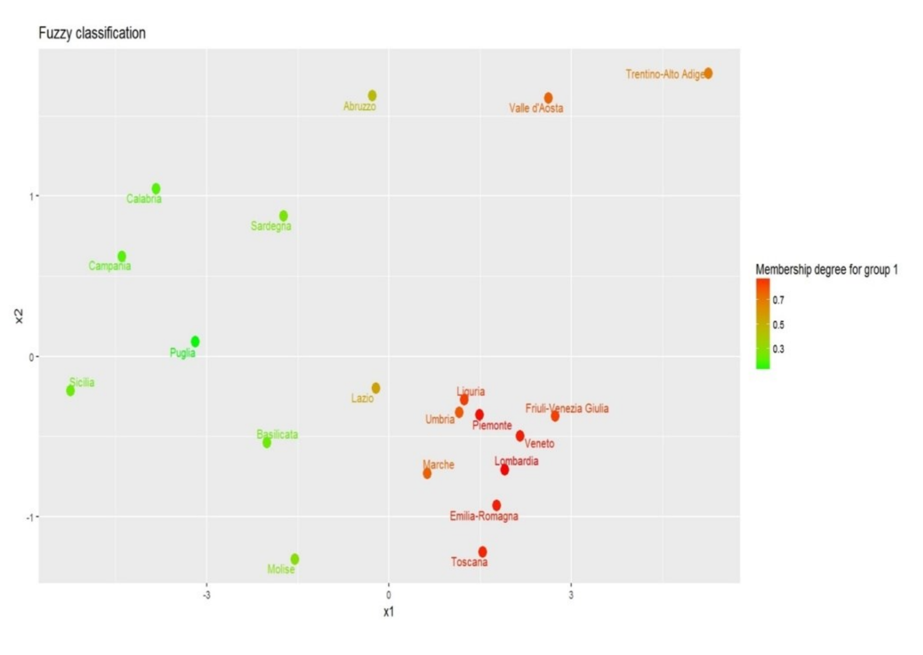

Figure 3a)

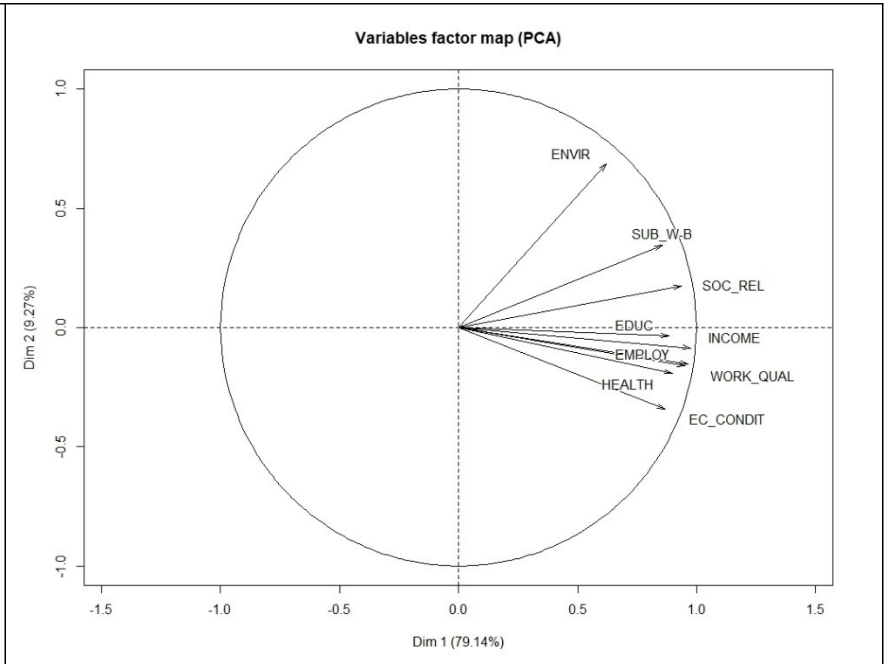

Figure 3b)

Figure 3. *Figure 3a). Fuzzy clustering of Italian regions. The fuzzy clustering is performed on the data gathered from the Report on the Equitable and Sustainable Well-Being (BES: "Benessere Equo e Sostenibile"). Figure 3b). Principal components interpretation for the nine dimensions of the Report on the Equitable and Sustainable Well-Being (BES: "Benessere Equo e Sostenibile"). Country: Italy. Year: 2016.*