Quality of life in cities: how to guide municipal governance?

Angela S. Bergantino*, Mario Intini*, Michele Vitagliano*

*Department of Economics, Management and Business Law, University of Bari Aldo Moro, Italy, Laboratory of Applied Economics, University of Bari Aldo Moro, Italy

EXTENDED ABSTRACT

The concept of quality of life has evolved significantly over time, moving beyond economic well-being and life expectancy to encompass a broader range of factors. It is now defined as the degree of satisfaction individuals derive from their social and physical environments (Mulligan, Carruthers, & Cahill, 2004). This comprehensive approach emphasizes both individual and societal well-being, integrating quality of life indicators into public policy to assess the holistic effects of decisions on daily life. Quality of life is viewed as a location-specific economic good that varies by region (Gillingham & Reece, 1979; Wingo, 1973), and it includes urban amenities—natural elements like favorable weather and human-made features such as cultural institutions and public services. Additionally, social factors, including safety and tolerance, are crucial in quality of life assessments (Diamond & Tolley, 1982; Rosen, 1979). This multidimensional view influences household and business location decisions, shaping urban development (Mulligan et al., 2004; Mulligan & Carruthers, 2011; Shapiro, 2006). The incorporation of overlooked factors into analysis has fostered deeper insights, and advancements in machine learning are enhancing the study of quality of life, particularly regarding health-related elements. A World Health Organization paper utilized a decision tree framework to identify key variables impacting life expectancy and quality of life, based on 25 factors related to health, environment, economy, and demographics observed in 2013. Similar studies have explored the relationship between clinical conditions and health issues, considering factors like socioeconomic status and infrastructure. Some researchers have expanded their focus to include aspects like urban development and climate action, suggesting they also affect quality of life. However, there is limited literature in this broader area, underscoring the need for further investigation into the relationship between quality of life and less obvious influencing factors, especially given the constraints of using a limited number of variables.

The absence of more granular indicators, both in terms of geographical scope and specific characteristics, has led most existing studies on quality of life to concentrate on multivariate analyses of aggregate spatial dimensions, such as NUTS-1 and NUTS-2 regions. Conversely, some research tends to be overly generalized, failing to offer a nuanced implementation in both characteristic and spatial terms. Additionally, other studies remain broad, merely outlining the methodology for applying machine learning within a spatial context. The focus on quality of life is particularly pertinent today, as numerous public policies utilize this concept as a benchmark to monitor the existence of uniform quality standards. Identifying the factors that significantly impact this indicator is crucial for addressing territorial disparities, which have become especially pronounced in the current historical context due to a convergence of global economic crises, rising political and economic tensions, the need to manage migration flows and promote integration, and the repercussions of the recent COVID-19 pandemic.

This paper is designed to furnish decision-makers with the pertinent information required to intervene effectively in specific territories, promoting greater alignment across the area or its particular characteristics while addressing the limitations identified in the existing literature. It seeks to explore the relationships between quality of life and a diverse, more granular dataset, employing machine learning (ML) models to identify the key drivers contributing to increasing territorial disparities and preventing the establishment of uniform quality standards. Unlike the previously discussed literature, which tends to isolate territorial disparities based on a single feature, this paper collects a comprehensive set of heterogeneous variables across a more detailed spatial dimension (NUTS-3 level, or Italian provinces). To achieve this, a combination of data from certified institutional sources and web scraping techniques has been utilized. Furthermore, this paper aims to develop a decision-making tool for policymakers through decision tree representations, ensuring enhanced explainability. More specifically, the approach trains machine learning models to identify the key characteristics relevant for distinguishing differences in quality of life. This is represented by a composite indicator known as the Life Quality Index (LQI), which, by design, aggregates several macro-categories and serves as a benchmark to address territorial disparities across various aspects of life. The proposed methodology is neutral, as it avoids any assumptions of functional dependence between the collected objective characteristics and the target variable, LQI. Furthermore, the interpretability of the machine learning model and the representation of results using decision trees allow these findings to be applicable across diverse territorial contexts. The insights gained can guide decision-makers in developing strategies to enhance the quality of life in similar settings, even if they are situated in different geographical areas. The immediate interpretability of the results makes them compelling and hard to dispute, as they can be easily recognized and compared with citizens' perceptions in specific provinces. In this paper, we have focused on the LQI (Life Quality Index) across Italian provinces, as derived from Il Sole 24 Ore's annual survey. The LQI measures the quality of life as perceived by citizens and establishes a ranking of various Italian provinces. Its primary purpose is to provide researchers and policymakers with insights into how perceived quality of life varies spatially, both within and among provinces, and over time. The annual updates to the LQI ranking often spark significant debate across the country, as they highlight the noticeable disparities between regions. This paper provides to planners and decision-makers with easily interpretable results. Our research seeks to identify the primary drivers contributing to increasing territorial disparities and inconsistent quality standards across Italian provinces. To achieve this objective, we examine the interplay between perception, which underpins the LQI (an index that aggregates distinct macro-categories), and the objective, disaggregated data drawn from diverse geographical, thematic, and political contexts. By employing machine learning approaches, we assess the impact of variables related to territorial and societal development on the perceived quality reflected by the LQI. Additionally, we compare the hierarchies of the most significant indicators across various provinces.

Acknowledgments

This study was carried out within GRINS project– Growing Resilient, INclusive and Sustainable and received funding from the European Union Next-Generation EU (NATIONAL RECOVERY AND RESILIENCE PLAN (NRRP), MISSION 4, COMPONENT 2, INVESTMENT 1.3 – D.D. 1558 11/10/2022, PE00000018, CUP: H93C22000650001, Spoke 7 Territorial sustainability). This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

References

1. R Pfeiffer, *et al.* The COVID-19 Recession on Both Sides of the Atlantic: A Model-Based Comparison, *European Commission*, discussion paper 191 | JULY 2023 https://economy-finance.ec.europa.eu/publications/covid-19-recession-both-sides-atlantic-model-based-comparison_en

2. NextGenerationEU and the National Recovery and Resilience Plan, *Agenzia per la Coesione Territoriale*, https://www.agenziacoesione.gov.it/dossier_tematici/nextgenerationeu-and-the-national-recovery-and-resilience-plan/

3., Responses to Coronavirus (COVID-19), Social economy and the COVID-19 crisis: current and future roles, *OECD Policy*, 30 July 2020 https://www.oecd.org/coronavirus/policy-responses/social-economy-and-the-covid-19-crisis-current-and-future-roles-f904b89f/

4. M. Habib, Z. Wang, Machine Learning Based Healthcare System for Investigating the Association Between Depression and Quality of Life, *Journal of Biomedical*, 5 January 2022, https://www.semanticscholar.org/paper/Machine-Learning-Based-Healthcare-System-for-the-of-Habib-Wang/5ff1df7a1ebfa20423c73e0864859559573e0ef7

5. I. Karacan,, *et al.* Analysis of life expectancy across countries using a decision tree, Research article EMHJ – Vol. 26 No. 2 – 2020, Eastern Mediterranean Health Journal, *World Health Organization* https://www.emro.who.int/emhj-volume-26-2020/volume-26-issue-2/analysis-of-life-expectancy-across-countries-using-a-decision-tree.html

6. M. Seok 1, *et al.* Machine Learning for Sarcopenia Prediction in the Elderly Using Socioeconomic, Infrastructure, and Quality-of-Life Data, Healthcare (Basel), 2023 Nov 1; DOI: 10.3390/healthcare11212881 https://pubmed.ncbi.nlm.nih.gov/37958025/

7. L. Salvati, *et al.* In-between regional disparities and spatial heterogeneity: a multivariate analysis of territorial divides in Italy, Published 3 June 2017, *Economics Journal of Environmental Planning and Management* https://www.semanticscholar.org/paper/In-between-regional-disparities-and-spatial-a-of-in-Salvati-Zitti/566910db03f10176b35ba0c772831d3743540554

8. L. S. Alaimo, *et al.* Sustainable Development Goals Indicators at Territorial Level: Conceptual and Methodological Issues - The Italian Perspective, *Social Indicators Research* 1 January 2020, Economics https://www.semanticscholar.org/paper/Sustainable-Development-Goals-Indicators-at-Level%3A-Alaimo-Maggino/858ed01786b2c311117a130d2899bca8c19db84e DOI:10.1007/S11205-019-02162-4Corpus ID: 201389763

9. Katarzyna Kopczewska , Spatial machine learning: new opportunities for regional science, Published: 24 December 2021, volume 68, pages713–755 (2022), The Annals of Regional Science, *An International Journal of Urban, Regional and Environmental Research and Policy* https://link.springer.com/article/10.1007/s00168-021-01101-x

10. E.Colombo, *et al.* La Dolce Vita: Hedonic Estimates of Quality of Life in Italian Cities, Regional Studies, 2014. Vol. 48, No. 8, 1404–1418, ISSN: 0034-3404 (Print) 1360-0591 (Online) https://doi.org/10.1080/00343404.2012.712206

11. Sarah E. Chambliss, *et al.* Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring, research article, *Environmental Sciences*, MA, July 26, 2021 https://doi.org/10.1073/pnas.2109249118

12. V. Tomaselli, *et al.* Building Well-Being Composite Indicator for Micro-Territorial Areas Through PLS-SEM and K-Means Approach, Ideas, https://ideas.repec.org/a/spr/soinre/v153y2021i2d10.1007_s11205-020-02454-0.html

13. M. Mazziotta and A. Pareto, Use and Misuse of PCA for Measuring Well-Being, Social Indicators Research, Vol. 142, No. 2 (April 2019), pp. 451-476 (26 pages), Published By: Springerhttps://www.jstor.org/stable/48704617

14. D. Teoli; A. Bhardwaj, Quality Of Life , NIH National Library of Medicine, March 27, 2023, https://www.ncbi.nlm.nih.gov/books/NBK536962/

15. S. Tekouabou, *et al.* Reviewing the application of machine learning methods to model urban form indicators in planning decision support systems: Potential, issues and challenges, *Computer and Information Sciences*, Volume 34, Issue 8, Part B, 2022, Pages 5943-5967, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2021.08.007

16. P. Cairney, *et al.* Public Policy to Reduce Inequalities across Europe, *Oxford University press*, November 2022, https://global.oup.com/academic/product/public-policy-to-reduce-inequalities-across-europe-9780192898586?cc=us&lang=en&

17. B. Mahbooba, M., *et al.* Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model, *Complexity*, vol. 2021, Article ID 6634811, 11 pages, 2021. https://doi.org/10.1155/2021/6634811

18. D. V. Pynadath,, *et al.* Explainable Reinforcement Learning in Human-Robot Teams: The Impact of Decision-Tree Explanations on Transparency, 2022 31st *International Conference on Robot and Human Interactive Communication* (RO-MAN), Aug 2022Pages 749–756 https://doi.org/10.1109/RO-MAN53752.2022.9900608

19. Zins C., Conceptual approaches for defining data, information, and knowledge, February 2007, *Journal of the American Society for Information Science and Technology* 58(4):479-493, DOI:10.1002/asi.20508, SourceDBLP, https://www.researchgate.net/publication/220432993_Conceptual_approaches_for_defining_data_information_and_knowledge

20. D. Laney, 3D Data Management: Controlling Data Volume, Velocity, and Variety,. META Group, (February 2001) https://www.bibsonomy.org/bibtex/742811cb00b303261f79a98e9b80bf49

21. V. N. G. Raju, *et al.* Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 729-735, doi: 10.1109/ICSSIT48917.2020.9214160. https://ieeexplore.ieee.org/document/9214160

22. S. Frei, *et al.* Benign Overfitting without Linearity: Neural Network Classifiers Trained by Gradient Descent for Noisy Linear Data, *Computer Science & Machine Learning*, 14 Sep 2023 https://doi.org/10.48550/arXiv.2202.05928

23. Ahsan MM, *et al.* Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance. *Technologies*. 2021; 9(3):52. https://doi.org/10.3390/technologies9030052

24. V.Sharma, A Study on Data Scaling Methods for Machine Learning, (IJGASR) *International Journal For Global Academic & Scientific Research*, ISSN Number: 2583-3081Volume 1, Issue No. 1, 31–42, DOI: 10.55938/ijgasr.v1i1.4, journals.icapsr.com/index.php/ijgasr

25. H. Qian, *et al*. RobustScaler: QoS-Aware Autoscaling for Complex Workloads, Computer Science Distributed, Parallel, and Cluster Computing, 14 Apr 2022 https://doi.org/10.48550/arXiv.2204.07197

26. U. Mahadeo Khaire, R. Dhanalakshmi, Stability of feature selection algorithm: A review, *Computer and Information Sciences*, Volume 34, Issue 4, 2022, Pages 1060-1073, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2019.06.012.

27. Jeon H., Sejong Oh, Hybrid-Recursive Feature Elimination for Efficient Feature Selection, *Applied Sciences* https://www.mdpi.com/2076-3417/10/9/3211

28. T. I. Terintegrasi, *et al.* Impurity-Based Important Features for feature selection in Recursive Feature Elimination for Stock Price Forecasting October 2023Jurnal 6(4):1182-1194 DOI:10.31004/jutin.v6i4.17726 https://www.researchgate.net/publication/375521283_Impurity-Based_Important_Features_for_feature_selection_in_Recursive_Feature_Elimination_for_Stock_Price_Forecasting

29. Rácz A, *et al*. Effect of Dataset Size and Train/Test Split Ratios in QSAR/QSPR Multiclass Classification. Molecules. 2021 Feb 19;26(4):1111. doi: 10.3390/molecules26041111. PMID: 33669834; PMCID: PMC7922354. https://pubmed.ncbi.nlm.nih.gov/33669834/

30. Gholamy A., *et al.* Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation https://scholarworks.utep.edu/cgi/viewcontent.cgi?article=2202&context=cs_techrep

31. Dobbin, K.K., Simon, R.M. Optimally splitting cases for training and testing high dimensional classifiers. BMC *Med Genomics* 4, 31 (2011). https://doi.org/10.1186/1755-8794-4-31

32. Xu, Y., Goodacre, R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. J. Anal. Test. 2, 249–262 (2018). https://doi.org/10.1007/s41664-018-0068-2

33. Jerome H. Friedman, Greedy function approximation: A gradient boosting machine, October 2001https://projecteuclid.org/journals/annals-of-statistics/volume-29/issue-5/Greedy-function-approximation-A-gradient-boosting-machine/10.1214/aos/1013203451.full

34. Chen T., Guestrin C., XGBoost: A Scalable Tree Boosting System , *Computer Science, Machine Learning* https://doi.org/10.48550/arXiv.1603.02754

35. Max Kuhn, Kjell Johnson, Applied Predictive Modeling 1st ed. 2013, Corr. 2nd printing 2018 Edition, *Charter 8 Regression Trees and Rule-Based Models* https://vuquangnguyen2016.files.wordpress.com/2018/03/applied-predictive-modeling-max-kuhn-kjell-johnson_1518.pdf

36. Buongiorno A, Intini, M., Sustainable tourism and mobility development in natural protected areas: Evidence from Apulia. Land Use Policy 2021, DOI http://dx.doi.org/10.1016/j.landusepol.2020.105220 https://www.researchgate.net/publication/347524072_Sustainable_tourism_and_mobility_development_in_natural_protected_areas_Evidence_from_Apulia

37. Bergantino A.S., Buonarota M., Buongiorno A., Intini M, Regional multimodal accessibility: Policies and strategies for sustainable tourism destinations in coastal areas, *Research in Transportation Business & Management*,,Volume 48, 2023, 100872, ISSN 2210-5395, DOI https://doi.org/10.1016/j.rtbm.2022.100872 https://www.sciencedirect.com/science/article/pii/S2210539522000931

38. Bergantino A.S., Intini M., Tangari L., Influencing factors for potential bike-sharing users: an empirical analysis during the COVID-19 pandemic, *Research in Transportation Economics*, Volume 86, 2021, 101028, ISSN 0739-8859, DOI https://doi.org/10.1016/j.retrec.2020.101028 https://www.sciencedirect.com/science/article/pii/S0739885920302262

39. Bergantino A.S., Buongiorno A., Intini M., Mobilità e sviluppo turistico sostenibile. Una prospettiva economica, Edizione: Marzo 2021, *Collana: Studi Superiori*, ISBN: 9788829005642

40. Ansell, B. W. The politics of housing. Annu. Rev. Politi. Sci. 22, 165–185 (2019)

41. Ayala, L., Bárcena-Martín, E., Cantó, O. & Navarro, C. COVID-19 lockdown and housing deprivation across European countries. Soc. Sci. Med. 298, 114839 (2022)

42. Keller, A. et al. Housing environment and mental health of Europeans during the COVID-19 pandemic: A cross-country comparison!. Sci. Rep. 12, 5612 (2022)

43. Amerio et al., 2020 A. Amerio, A. Brambilla, A. Morganti, A. Aguglia, D. Bianchi, F. Santi, S. Capolongo, COVID-19 lockdown: Housing built environment's effects on mental health, *International Journal of Environmental Research and Public Health*, 17 (16) (2020), https://www.mdpi.com/1660-4601/17/16/5973

44. D. Ryu, S. Sok, Prediction model of quality of life using the decision tree model in older adult single-person households: a secondary data analysis, *Original Research*, 31 August 2023, DOI 10.3389/fpubh.2023.1224018

45. C. L. Reynolds, A. L. Weinstein, Gender differences in quality of life and preferences for location-specific amenities across cities, *Regional Sci*, 2021, 61, 916–943. https://doi.org/10.1111/jors.12520

46. Fernández, Raquel, and Alessandra Fogli. 2009. Culture: An Empirical Investigation of Beliefs, Work, and Fertility American Economic Journal: *Macroeconomics*, 1 (1): 146-77. DOI: 10.1257/mac.1.1.146

47. Shapiro, Jesse M. Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital *The Review of Economics and Statistics*, vol. 88, no. 2, 2006, pp. 324–35. JSTOR, http://www.jstor.org/stable/40042998

48. F. J. Goerlich, E. Reig, Quality of life ranking of Spanish cities: A non-compensatory approach, *Cities*, Volume 109, 2021, 102979, ISSN 0264-2751,https://doi.org/10.1016/j.cities.2020.102979

49. G. R. Patil, G. Sharma, Urban Quality of Life: An assessment and ranking for Indian cities Transport Policy, Volume 124, 2022, Pages 183-191, ISSN 0967-070X, https://doi.org/10.1016/j.tranpol.2020.11.009

50. Anderson Y., *et al.* Lively social space, well-being activity, and urban design: Findings from a low-cost community-led public space intervention, *Environment and Behavior*, 49 (6) (2017), pp. 685-716, https://journals.sagepub.com/doi/10.1177/0013916516659108

51. Lyons P., *et al.* The dynamics of urban metabolism in the face of digitalization and changing lifestyles: Understanding and influencing our cities Resources, Conservation and Recycling, 132 (2018), pp. 246-257, https://linkinghub.elsevier.com/retrieve/pii/S0921344917302112

52. Z. Erdoğan, E Namlı, A living environment prediction model using ensemble machine learning techniques based on quality of life index, Journal of Ambient Intelligence and Humanized *Computing*, *Springer-Verlag GmbH Germany*, part of Springer Nature 2019, https://doi.org/10.1007/s12652-019-01432-w

53. M. Savić, V., *et al.* Analysis of Machine Learning Models Predicting Quality of Life for Cancer Patients MEDES '21: Proceedings of the 13th *International Conference on Management of Digital EcoSystems*, November 2021Pages 35–42 https://doi.org/10.1145/3444757.3485103