

This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Supporting the Analysis of Inner Areas of a Territory

Paolo Buono[1][0000−0002−1421−3686], Maria Francesca Costabile[1][0000−0001−8554−0273], Palmalisa Marra[2], Valentino Moretto[2], Antonio Piccinno[1][0000−0003−1561−7073], and Luca Tedesco[2]

[1] IVU Lab, Dept. of Computer Science, University of Bari Aldo Moro, Bari, Italy
{paolo.buono, maria.costabile, antonio.piccinno}@uniba.it
[2] Links Management and Technology S.p.A., Lecce, Italy
{palmalisa.marra, valentino.moretto, luca.tedesco}@linksmt.it
http://www.springer.com/gp/computer-science/lncs

Abstract. This paper addresses the problem of supporting public decision makers in the data analysis of the territory they are responsible for. Most of the Italian territory consists of the so-called inner areas, which are primarily rural areas usually far from main centers that provide education, health and mobility services. A project that analyzes inner areas is reported; it has two main objectives. A first one is to identify proper clusters of such areas so that smaller areas can be grouped together in order to share important services in the best and more economic way; different algorithms are used to identify relevant parameters and to cluster inner areas. The second objective is to create proper visualizations of data of such inner areas, in order to better support decision makers in understanding the data and taking more informed decisions. It is remarked that using visualization techniques without adequate knowledge of their possibilities greatly limits the support that analysts may get from them.

Keywords: Smart territories · Urban Development Planning · Clustering Algorithms · Data Visualization

1 Introduction and motivations

The emphasis on smart cities keep increasing with a major focus on large metropolitan cities. Other important components of the European territory are the so-called inner areas. In Italy, inner areas are rural areas characterized by their distance from main service centers; such centers are municipalities able to provide education, health and mobility services that inner areas may lack, for example schools with a full range of secondary education, emergency care hospitals and silver category railway stations. In Italy, inner areas refer to more than 50% of Italian municipalities, are home of about 25% of the Italian population and cover about 60% of the national territory. Thus, they require specific attention by public administration managers and politicians.

The models considered for the analysis of smart cities cannot be directly used for analyzing inner areas since they have different characteristics. Such models could be used only after selecting the right indicators for inner areas. The data resulting from the analysis can then be presented to the decision makers so that they can identify proper strategies for the development of such marginalized territories [8]. Several attempts have been developed to categorize inner areas [2, 1, 10]). For example, data engineering, data visualization and intelligent data analytics methods are used to create a decision support system to be used by technical and/or political decision makers [6]. In Italy, a national strategy for inner areas was introduced with the objective of improving the quality of life of people living in inner areas and inverting the demographic trend of such areas, see e.g. [14]. Several indicators describing the social, economic and territorial environment have been proposed [13, 5, 7]. The census of inner areas performed in 2013 classifies the peripheral level according to two main criteria: the offered services and the distance of an inner area from the centers that provide the essential services.

This paper briefly reports about a project that Links Management and Technology S.p.A. is carrying out in the South of Apulia region, the Salento area, with the municipality of Galatone, a town in that area. The project involves researchers of the IVU Lab of the Computer Science Department of the University of Bari Aldo Moro, since one of the objectives is to analyze data with the support of visualization techniques, of which these researchers are expert. The overall goal of the project is to improve the potentialities of inner areas, based on the idea that inner areas can join to provide together more essential services. Some activities of the project are briefly described in Sections 2 and 3. Section 4 concludes the paper.

2 Clustering inner areas

One of the objectives of the project is to analyze how inner areas can be joined so that, together, can provide a set of essential services. For this analysis, several aspects have been considered, such as geo-demography, economy, innovation and sustainable development [11]. For each aspect, several indicators have been defined, reaching a total of about 50. Using such indicators, several clustering algorithms have been applied to cluster inner areas in order to identify those that can be grouped together. The results showed that the most useful algorithms are K-means++, Skater and an ensemble algorithm; however, they provide different classifications of the inner areas. Specifically, K-means++ reveals more irregular distribution of municipalities and classifies them in four clusters, while Skater takes more into account the distance and creates 5 clusters. The ensemble method shows the most similar municipalities to a reference one, so that all similar municipalities are selected at once and can be classified more easily.

The clustering techniques have been selected empirically, taking into account the knowledge of the analyst. In order to support the analyst on the correct choice of the method or the parameters, a visual tool may help. Some authors

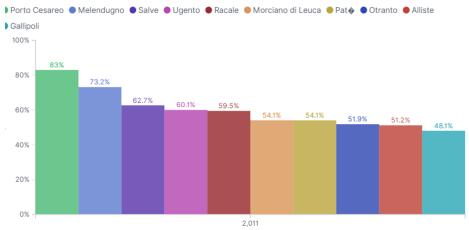


Fig. 1. Occupied houses by municipalities

use visual techniques to compare predictive models (e.g., see [4, 9]). Pister et al. [12] propose PK-clustering to support the user choosing the best method compared to the user's prior knowledge. However, visual techniques are useful not only for complementing other analytical methods, they are very powerful as visual analysis tools themselves, as described in the next section.

3 Visual analysis of inner areas

Several tools are available in the market, which provide different visualization techniques. One of these is Kibana[3], that is used in the project with Galatone municipality. However, the collaboration with researchers of the IVU Lab of the University of Bari was instrumental for a more valuable use of the techniques in Kibana, as reported in the following. In addition, IVU researchers discussed with Links researchers the importance of the validation with real users of the adopted visualizations, as for example done in [3].

In this project, a first report was produced by Links consisting of 26 dashboards (a dashboard is a set of visualizations proposed for a certain goal); such visualizations referred to data relative to various aspects of the inner areas: environment and territory; economy; infrastructures and mobility; socio-

[3] https://www.elastic.co/kibana, last visited: July 2021

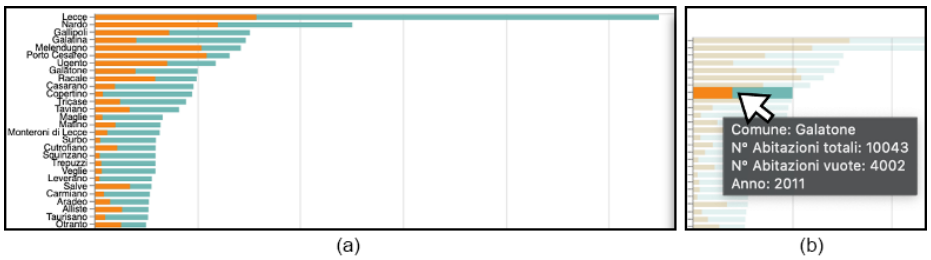


Fig. 2. (a) Salento municipalities ordered by the number of houses; orange color indicates empty houses and blue color indicates occupied houses. (b) Details of Galatone municipality when hovered with the mouse

demographic, tourism. An example of visualization using one of the technique in Kibana is in Fig. 1: it visualizes the top 10 municipalities of Salento in terms of occupied houses; each column represents the percentage of houses occupied by people; the legend at the top of the figure associates the color of each bar with the referred municipality.

Let us discuss now how this data visualization was redesigned by the IVU experts in order to allow the analyst to be faster in the data analysis and get more information, possibly discovering new and unexpected patterns.

Fig. 2(a) still uses a histogram as in Fig. 1 but, being rotated of 90 degrees, it may visualizes all 97 municipalities in Salento (for the sake of space, the image is cropped and only the top 28 municipalities are visible here). In addition, each bar presents two colors because each color codifies an attribute: the blue color indicates houses occupied in the municipality and the orange color indicates empty houses. The length of the colored bar is proportional to the number of houses. This is an interactive visualization. As shown in Fig. 2(b), the visualization changes when hovering with the mouse on one bar: all the other municipalities are grayed and a popup shows detailed information about the selected municipality (Galatone, in the case of Fig. 2(b)). The municipalities can be shown in the histogram in various orders chosen by the user. In Fig. 2, the data are ordered by the number of houses. It is immediately visible that the proportion of occupied and free houses in the municipalities is different in the municipalities. For example, the three municipalities above Galatone (reported by eighth bar from the top in Fig. 2(a)) have much more empty houses than occupied ones. By further investigating on these three municipalities, the analyst finds that they refer to the most touristic towns in Salento, located on the sea, and the empty houses are a lot since they are occupied only during summer.

4 Conclusions

This paper discussed the problem of analyzing inner areas with clustering techniques as well as visualizations techniques, in order to better support the the decision maker of public administrations to group them in larger areas that are more efficient from the point of view of the available services. An on-going project is illustrated, on which Links Management and Technologies is working on, also collaborating with researchers of the IVU Lab of the University of Bari Aldo Moro. There is still a significant work to do on different directions, including the identification of the best clustering algorithms and the investigation of additional visualization techniques, not only to improve the comprehension of the analyzed data but also to provide new perspectives or analysis dimensions. A future plan is also to provide citizens with dashboards they may directly use. This is a challenging task because the visualizations must be understandable for the mass without requiring any training or specific expertise.

5 Acknowledgments

The authors thank Andrea De Matteis for his contribution in the implementation of the dashboards.

References

1. Battaglia, M., Annesi, N., Pierantoni, I., Sargolini, M.: Future perspectives of sustainable development: An innovative planning approach to inner areas. experience of an italian alpine region. Futures **114**, 102468 (2019)
2. Bertolini, P., Pagliacci, F.: Quality of life and territorial imbalances. a focus on italian inner and rural areas. Bio-based and Applied Economics **6**(2), 183–208 (Oct 2017). https://doi.org/10.13128/BAE-18518, https://oaj.fupress.net/index.php/bae/article/view/3312
3. Buono, P., Costabile, M., Lanzilotti, R.: A circular visualization of people's activities in distributed teams. Journal of Visual Languages and Computing **25**(6), 903–911 (2014). https://doi.org/10.1016/j.jvlc.2014.10.025
4. Buono, P., Legretto, A., Bertini, E., Costabile, M.F.: Visual techniques to compare predictive models. In: Proc. of the 13th Biannual Conference of the Italian SIGCHI Chapter: Designing the next Interaction. CHItaly '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3351995.3352035, https://doi.org/10.1145/3351995.3352035
5. Cazzato, F.: Le politiche di sviluppo rurale in italia: un'analisi territoriale alla luce dei nuovi programmi comunitari. Le politiche di sviluppo rurale in Italia: un'analisi territoriale alla luce dei nuovi programmi comunitari pp. 79–106 (2007)
6. Conejero, J.M., Preciado, J.C., Prieto, A.E., Bas, M., Bolós, V.J.: Applying data driven decision making to rank vocational and educational training programs with topsis. Decision Support Systems **142**, 113470 (2021)
7. Girard, L.F.: Multidimensional evaluation processes to manage creative, resilient and sustainable city. Aestimum pp. 123–139 (2011)
8. Jones, G., Leimgruber, W., Nel, E.: Issues in geographical marginality. General and Theoretical Aspects (Grahamstown: Rhodes University) (2007)
9. Legretto, A., Buono, P.: Poster: Analyzing predictive models with the help of visualizations. In: womENcourage (September 2019)
10. Marucci, A., Fiorini, L., Di Dato, C., Zullo, F.: Marginality assessment: computational applications on italian municipalities. Sustainability **12**(8), 3250 (2020)
11. Moretto, V., Elia, G., Schirinzi, S., Vizzi, R., Ghiani, G.: A knowledge visualization approach to identify and discovery inner areas: a pilot application in the province of lecce. Management Decision (Jan 2021). https://doi.org/10.1108/MD-01-2021-0104
12. Pister, A., Buono, P., Fekete, J.D., Plaisant, C., Valdivia, P.: Integrating prior knowledge in mixed-initiative social network clustering. IEEE Transactions on Visualization and Computer Graphics **27**(2), 1775–1785 (2021). https://doi.org/10.1109/TVCG.2020.3030347
13. Tempestai, T., Bazzani, G.M., Thiene, M.: Le conseguenze della riforma della politica agricola comunitaria sul paesaggio rurale. Le conseguenze della riforma della Politica Agricola Comunitaria sul paesaggio rurale pp. 79–95 (2006)
14. Zolin, M.B., Ferretti, P., Grandi, M.: Sustainability in peripheral and ultra-peripheral rural areas through a multi-attribute analysis: the case of the italian insular region. Sustainability **12**(22), 9380 (2020)