Bes Indicators in Italy

Alberto De Benedittis

June 9, 2022

Abstract

The fair and sustainable well-being index (**BES**) was developed by ISTAT and CNEL. It has been created to assess the development of society not only from an economic point of view but also from a social and environmental one, together with measures of inequality and sustainability. The 28^{th} of June 2016 BES entered for the first time in the Italian Government balance sheet. Thus, nowadays BES indicators help to measure Italians' quality of life and to assess the impact of the government's public policies with respect to the fundamental social measures.
However, there are still substantial differences among norther and southern ones in Italy. Hence, it is important to monitor these indexes and their trend over the years.
The goal of this paper is to present these indicators and to visualize their distribution through (choroplet)maps in order to understand how the sustainability process is evolving in Italy.

1 Introduction

BES indicators are a set of multiple indexes used by the Italian institute of statistics and the Italian government to yearly monitor the well-being of Italians and to promote a fair and sustainable development, coherently with the United Nation sustainable goals. BES indicators are grouped into twelve different domains, relevant for the measurement of the well-being and the sustainability:

- *Health* represents a central element in life and an essential condition for individual well-being and prosperity of populations (Sachs et al., 2001). Health has an impact on all dimensions of individual life in all its different phases, modifying life conditions, behaviour, social relationships, opportunities and prospects of individuals and, often, of their families. While age increases, the role played by health conditions becomes increasingly important.

- *Education and training* affect the well-being of individuals and open up opportunities otherwise precluded. Moreover, better educated people seem to be happier and more satisfied with their own life than their less educated counterparts (Möwisch, Brose, & Schmiedek, 2021).

- *Work and life balance*: A job well paid contributes significantly to the achievement of well being. However, also the balance between work and family life plays a key role for individuals well being.

- *Economic well-being*: Economic resources are the primary mean through which it is possible to have a good standard of living.

- *Social Relationships*: Relational networks to which individuals belong and in which they recognize themselves, represent a fundamental resource that allows pursuing their own ends relying on additional resources compared to the available endowments of economic and cultural capital (Rogošić & Baranović, 2016).

- *Politics and Institutions*: The trust expressed by citizens to the institutions, as well as civic and political participation, facilitate cooperation and social cohesion while allowing greater efficiency of public policies and a lower cost of transactions.

- *Security :* The subjective perception and the experience of objective safety in daily life in Italy become of paramount importance in the construction of individual and community well-being. Personal security is part of the foundation of individual well being. Being victim of a crime can result in economic loss, physical and/or psychological damage due to a suffered trauma. The most important effect of criminality on well-being is the sense of vulnerability that it determines on individuals. The fear to be victim of crime can strongly affect personal freedom, quality of life and development of territories. Also the theme of violence is closely related to personal security and quality of life.

- *Subjective well being*: Perceptions and evaluations affect the way people face life and take advantage of opportunities. Subjective indicators are useful complement to the most objectifiable indicators, because they allow evaluating the possible differences between what people report on their perceiving regarding their satisfaction and what it is captured by statistical observation of economic and social phenomena.

- *Landscape and cultural heritage*

- *Environment*: An environment which is in a vital and healthy state constitutes a prerequisite to ensure authentic well-being for all components of society.

- *Innovation, research and creativity*: Innovation and Research are an indirect determinant of well-being and the base of social and economic progress.

- *Quality of services*: The link between the availability of services and citizens' well-being is based on an interpretive approach in which high quality public investments improve the general context in which people live and work and their social and economic interconnections.

These indicators cover different aspects that involve the overall well being of individuals. The aim of this paper is to describe the distribution of these indicators in the Italian territory in order to detect whether some patterns exist. For example, one issue that is always recurring is the one of the so called *"Questione Meridionale"* (that can be effectively translated as the southern issue). This refers to the development gap that has always occurred between the northern and richer areas of Italy compared to the southern and poorer ones. Thanks to the representation of the above indicators it is possible to inspect in which domains this gap exists and if so, how big it is.

2 Problem and Data

2.1 Problem

The aim of this analysis is to create maps to describe the distribution of BES indicators, to identify the best areas for each domain, and to assess meaningful association through spatial auto-correlation. Eventually, the results of the analysis are represented on a dashboard that allows to exploit web mapping tools.

For instance, the BES reports issued by ISTAT every year represent the trend of each indicator trough line charts or histograms. In addition to these classical representations, it could be useful to add choropleth maps. This is a type of statistical thematic map that uses intensity of color to correspond with an aggregate summary of a geographic characteristic within spatial enumeration units, such as population density or per-capita income (Dent, Torguson, & Hodler, 2008). Choropleth maps provide an easy way to visualize how a variable varies across a geographic area or show the level of variability within a region.

The potential advantages of using web mapping tools to integrate socio-economic data into global and national platforms could be substantial in terms of facilitating research and allowing the public to compare and contrast locations across the world using a range of indicators (Smith, 2016).

In addition to this tool, it is also interesting to perform *spatial auto-correlation*. This term refers to the presence of systematic spatial variation in a mapped variable where:

- adjacent observations have similar data values then the map shows positive spatial auto-correlation

- adjacent observations tend to have very contrasting values then the map shows negative spatial auto-correlation.

(Haining & Haining, 2003)

2.2 Data

The data used for the analysis come from two main sources:

- The Italian Institute of Statistics (ISTAT)
- Open Street Map (OSM), a collaborative project to create a free editable geographic database of the world.

It has been possible to download both the data about the twelve BES indicators and the ones about the actual administrative borders of Italy from the ISTAT portal.

The BES data is organized as a huge spreadsheet file with almost twelve thousand rows. Each observation has information about its domain e.g. Health, and its indicator e.g. car accidents' mortality. The data covers the period between 2004 and 2019. However, there are many missing observations. This is due to the fact that each year some indicators have been added or removed. Moreover, there are missing data for some of provinces of Sardinia (Ogliastra,Olbia-Tempio,Medio Campidano,Carbonia-Iglesias). The data about the administrative boarders is instead an ESRI shapefile with WGS84 reference system. It contains four files: the geometry (.shp), the index(.shx), the projection (.prj), and the table(.dbf). There are three different shapefiles: one for regions, one for provinces, and one for macro areas (North, Center, South). Lastly, three data-sets have been downloaded, containing the reference point of factories, no profit organizations and the points of interest, such as theater and cinemas, in the Italian soil. This data has been collected through the python library pyrosm which allows to read OpenStreetMap from Protocolbuffer Binary Format -files (*.osm.pbf) into Geopandas GeoDataFrames. Pyrosm makes it easy to extract various data-sets from OpenStreetMap pbf-dumps including e.g. road networks, buildings, Points of Interest (POI), landuse, natural elements, administrative boundaries and much more.
All the data that have been used are Open Data. More precisely, the data is under the creative commons license. Thus, it is possible to freely reproduce, distribute, share and adapt data and analysis of ISTAT and OSM even for commercial purposes with the only constraint of citing the source.

3 Methodology

The main goal of this analysis is to present in an easy and understandable way the statistics related to the BES indicators in the Italian territories. The statistics cover the period from 2004 and 2018 (with some exceptions due to missing data) and are made on three level of analysis: Macro-areas, Regions and Provinces. To do so, a dashboard has been created which allows users to play with the data and to find meaningful associations. The dashboard provides a general introduction to what the BES is. Then, thanks to a side menu it is possible to choose the domain of interest, the year of analysis, the level of analysis and whether the users wants to look at dynamic or static maps: static ones (Figure 1) uses quintiles (the population is divided into 5 different groups) to decide the color of each territory. Dynamic ones (Figure 2) present a colorbar that associates the intensity of the color with the values of the indicator. Furthermore, when an user moves the cursor on a territory a pop up message displays the indicator unit of measure and the corresponding value. In addition to the maps, a line chart is displayed (Figure 3). This representation helps the user to visualize the trend of the chosen indicators along the years. By default in the line chart are shown the data of the macro-areas. However, the user can interact with the line chart. Indeed, by clicking on the right scroll-down menu he/she can select one or more territories he/she is interested about.
However, when dealing with spatial data it is important to remember that geo-referenced data are subjected to Tobler's first law of geography: "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Hence, the user can be interested in understanding whether there exist spatial auto-correlation (more details on spatial auto-correlation are provided in the Appendix). If the null hypothesis of spatial randomness is rejected, it means that there exist spatial auto correlation. Then, a choropleth maps is displayed and it shows the different territories colored based according to the Moran scatter plot . According to this criterion, territories are classified in one out of four categories:

- Hot spot: this represents the territories with high spatial correlation characterized by high values of the measure of analysis

- Cold spot : this represents the territories with high spatial correlation characterized by low values of the measure of analysis
- Doughnut: that can be interpreted as an outlier compared to close territories
- Ns: Represent the territories that are not characterized by a specific trend.

The last feature of the dashboard is an interactive map that shows the best three territories relatively to the chosen domain. The ranking system used for determining the best territories is a simple one: for each indicator territories are ranked according to their score. This operation is repeated for each indicator. Then, the rankings are summed and the territories with the lowest values, namely, the ones that have been in the first positions more frequently, are labelled as the best ones.
Obviously, this is an over simplistic model. In order to create a better model an ad hoc study should be conducted, that tells how much each indicator influences the overall quality of life.
Nevertheless, this simple model still gives useful insights to understand whether there exist inequalities among Italian regions.

The above holds for all domains with the exception of the Subjective well-being, as there is no data available for this domain. However, since a description of the indicators is provided, it has been possible to approximate one of the them: Leisure time satisfaction. To do so, an isochrone map has been created (Figure 6) that shows for some of the biggest cities in Italy, the places classified as entertainment, art and culture according to the OpenStreetMap schema. More precisely, this map shows all the places that can be reached in 10-20 minutes of walking from the city centre.

4 Conclusion

To conclude, the dashboard is a good tool to integrate the BES report that is published yearly by ISTAT. As a matter of fact, the choropleth maps are quite intuitive to interpret and they allow also non expert users to understand the distribution of well being in Italy. Indeed, by looking at the best territories (summarized in Table 1) for each domain, one can easily notice that the best territories are almost all in the north or in the center of Italy. For instance, by looking at the macro-areas the south is never the best territory. Considering the regional level, Lomabardia is the region with the highest number of 1st classified (4 times over 11 domains), while the one that occurs most in the first three positions is Emilia-Romagna (7 over 11 domains). The best southern region, according to these indicators is Abruzzo. It has been classified as best territory for two domains: Politics and Institutions and Education and training. Considering the provinces the situation is even more dramatic. Over 23 provinces just 2 belong to the south of Italy: Catania relatively to Politics and Institutions and Nuoro relatively to security. The best provinces seem to be Brescia which has been classified as best territory and as third best territory twice, followed by Bologna which has been classified best territory twice.
This impressive result is not unexpected. Indeed, the imbalance between the quality of well-being among the north and the south of Italy is a well known issue. Hence, the Italian government should consider the BES statistics more seriously and should implement policies that try to mitigate this problem.

5 Open Issues

The analysis has two main limits:

- The dashboard is extremely slow. For example, when it needs to compute the spatial auto-correlation plot. This is due the fact that it was not possible to cache the computations so every time the user load the dashboard the user has to wait that all computations are executed. A possible solution would have been storing all the maps and the graphs on a cloud storage system such as Google cloud platform. Unfortunately, in order to save all the maps as an HTML file there is the need to have more than 100 gigabyte. So, it would have been necessary to subscribe and pay for the service.
- The model that shows the best areas is an over approximation. For instance, the same weight is given to all the indicators. It seems obvious that some indicators are more important than others.

References

Anselin, L. (1996). The moran scatterplot as an esda tool to assess local instability in spatial. *Spatial Analytical*, *4*(111), 9780203739051–8.

Cliff, A. D., & Ord, K. (1970). Spatial autocorrelation: a review of existing and new measures with applications. *Economic Geography*, *46*(sup1), 269–292.

Dent, B., Torguson, J., & Hodler, T. (2008). *Thematic map design*. McGraw-Hill New York, New York, NY.

Haining, R. P., & Haining, R. (2003). *Spatial data analysis: theory and practice*. Cambridge university press.

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, *10*(2), 243–251.

Möwisch, D., Brose, A., & Schmiedek, F. (2021). Do higher educated people feel better in everyday life? insights from a day reconstruction method study. *Social Indicators Research*, *153*(1), 227–250.

Rogošić, S., & Baranović, B. (2016). Social capital and educational achievements: Coleman vs. bourdieu. *Center for Educational Policy Studies Journal*, *6*(2), 81–100.

Sachs, J., et al. (2001). Macroeconomics and health. *WHO, Geneva*.

Smith, D. A. (2016). Online interactive thematic mapping: Applications and techniques for socio-economic research. *Computers, Environment and Urban Systems*, *57*, 106-117. Retrieved from https://www.sciencedirect.com/science/article/pii/S0198971516300023 doi: https://doi.org/10.1016/j.compenvurbsys.2016.01.002

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography*, *46*(sup1), 234–240.

6 Appendix

6.1 GitHub

Dashboard
Notebook

6.2 Spatial Auto-Correlation

The term spatial auto-correlation refers to the presence of systematic spatial variation in a mapped variable. Where adjacent observations have similar data values, the map shows positive spatial auto-correlation. Where adjacent observations tend to have very contrasting values then the map shows negative spatial auto-correlation.

6.2.1 Moran's I Statistic

In order to assess whether spatial auto-correlation is statistically significant, it is necessary to perform the Moran I statistic (Moran, 1948). In essence, it is a cross-product statistic between a variable and its spatial lag, with the variable expressed in deviations from its mean. For an observation at location i, this is expressed as $z_i = x_i \bar{x}$, where $\bar{x}$ is the mean of variable x.

Moran's I statistic is then: $I = \frac{\sum_i \sum_j w_{ij} * z_i * z_j / S_0}{\sum_i z_i^2 / n}$

with w_{ij} as the elements of the spatial weights matrix, $S_0 = \sum_i \sum_j w_{ij}$ as the sum of all the weights, and n as the number of observations.

Inference for Moran's I is based on a null hypothesis of spatial randomness. The distribution of the statistic under the null can be derived using either an assumption of normality (independent normal random variates), or so-called randomization (i.e., each value is equally likely to occur at any location). While the analytical derivations provide easy to interpret expressions for the mean and the variance of the statistic under the null hypothesis, inference based on them employs an approximation to a standard normal distribution, which may be inappropriate when the underlying assumptions are not satisfied (Cliff & Ord, 1970).

6.2.2 Moran scatter plot

The Moran scatter plot (Anselin, 1996) consists of a plot with the spatially lagged variable on the y-axis and the original variable on the x-axis. The slope of the linear fit to the scatter plot equal Moran's I. A variable z is considered, given in deviations from the mean. With row-standardized weights, the sum of all the weights (S_0) equals the number of observations (n). As a result, the expression for Moran's I simplifies to: $I = \frac{\sum_i \sum_j w_{ij} * z_i * z_j / S_0}{\sum_i z_i^2 / n} = \frac{\sum_i (z_i X \sum_j w_{ij} z_j)}{\sum_i z_i^2}$ This turns out to be the slope of a regression of $\sum_j w_{ij} z_j$ on z_i. This is the principle underlying the Moran scatter plot.

An important aspect of the visualization in the Moran scatter plot is the classification of the nature of spatial auto-correlation into four categories. Since the plot is centered on the mean (of zero), all points to the right of the mean have $z_i > 0$ and all points to the left have $z_i < 0$. We refer to these values respectively as high and low, in the limited sense of higher or lower than average. Similarly, we can classify the values for the spatial lag above and below the mean as high and low.

The scatter plot is then easily decomposed into four quadrants. The upper-right quadrant and the lower-left quadrant correspond with positive spatial auto-correlation (similar values at neighboring locations). We refer to them as respectively high-high and low-low spatial auto-correlation. In contrast, the lower-right and upper-left quadrant correspond to negative spatial auto-correlation (dissimilar values at neighboring locations). We refer to them as respectively high-low and low-high spatial auto-correlation.

The classification of the spatial auto-correlation into four types begins to make the connection between global and local spatial auto-correlation. However, it is important to keep in mind that the classification as such does not imply significance. This is further explored in our discussion of local indicators of spatial association (LISA).

6.3 Table of best territories

Table with the first three territories for each domain for each level of analysis

Domain	Macro-Areas	Regions	Provinces
Health	North Center South	Trentino-Alto Adige Umbria Toscana	Prato Perugia Trento
Education and Training	Center South North	Abruzzo Umbria Emilia-Romagna	Piacenza Potenza Brescia
Work and Life balance	Center South North	Lombardia Emilia-Romagna Friuli-Venezia Giulia	Brescia Como Udine
Economic and well-being	North Center South	Lombardia Emilia-Romagna Piemonte	Bologna Milano Trieste
Social Relationships	Center South North	Umbria Trentino-Alto Adige Friuli-Venezia Giulia	Udine Trento Grosseto
Politics and Institutions	Center South North	Abruzzo Umbria Emilia-Romagna	Catania Piacenza Brescia
Security	North Center South	Molise Valle d'Aosta Marche	Nuoro Verbano-Cusio-Ossola Ascoli Piceno
Landscape and Cultural Heritage	Center North South	Friuli-Venezia Giulia Toscana Veneto	Trieste Firenze Pisa
Environment	North South Center	Trentino-Alto Adige Friuli-Venezia Giulia Emilia-Romagna	Mantova Monza Ravenna
Innovation, Research and Creativity	North Center South	Lombardia Emilia-Romagna Piemonte	Bologna Milano Novara
Quality of services	Center North South	Lombardia Emilia-Romagna Friuli-Venezia Giulia	Brescia Como Ancona

6.4 Example of maps and charts

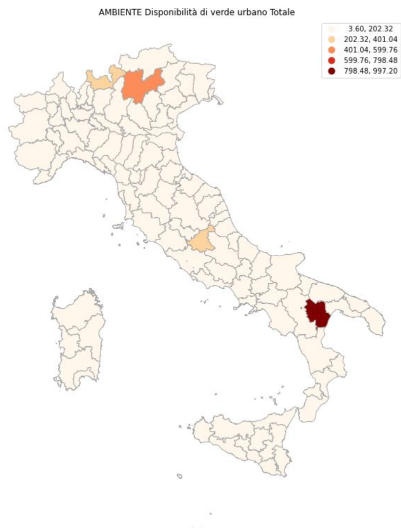

Figure 1: Static Choropleth map representing urban green areas in person per m^2

Figure 2: Dynamic Choropleth map representing urban green areas in person per m^2

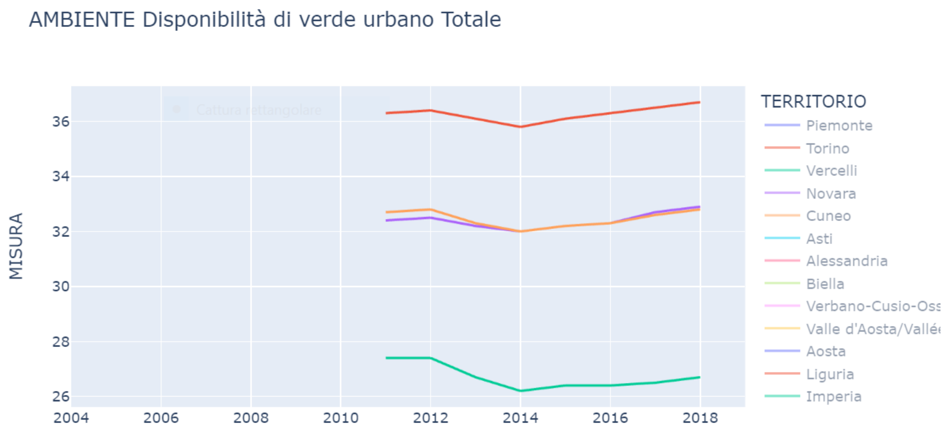

Figure 3: Line chart representing urban green areas in person per m^2

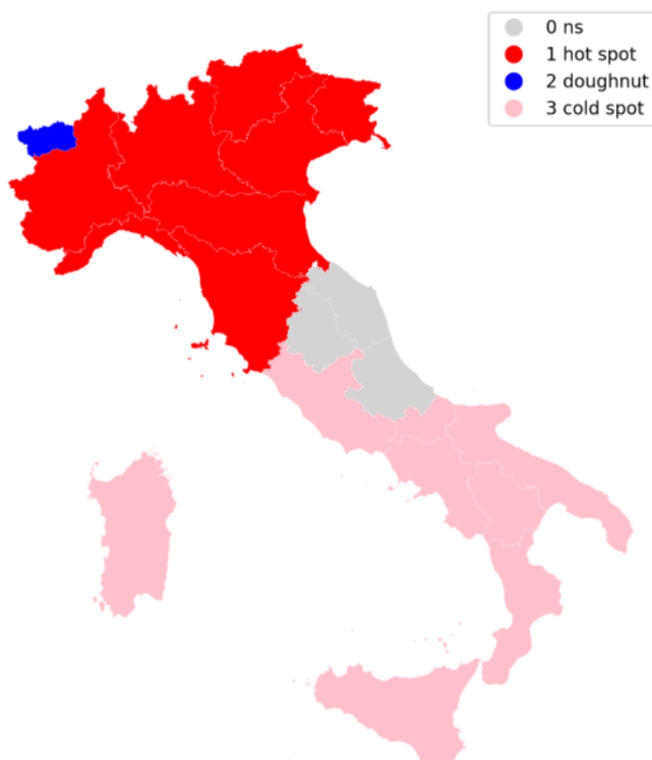

Figure 4: Choropleth map showing the Spatial auto-correlation relatively to the average salary of male employees

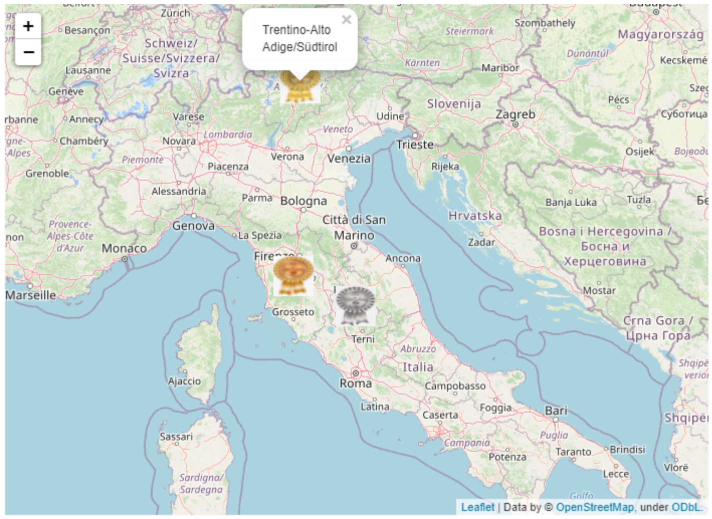

Figure 5: Map showing the best territories, at regional level, relatively to the health domain

Figure 6: Isochrone map showing the places of interest within 10-20 of walking from the Bari city centre