The Lisbon ranking for smart sustainable cities in Europe

Adeoluwa Akande[a,*], Pedro Cabral[a], Paulo Gomes[a], Sven Casteleyn[b]

[a] *NOVA IMS, Universidade Nova de Lisboa, Campus de Campolide, 1070-312, Lisboa, Portugal*
[b] *GEOTEC, Institute of New Imaging Technologies, Universidad Jaime I, Avenida Sos Baynat, E-12071, Castellon de la Plana, Spain*

ARTICLE INFO	ABSTRACT
	There has recently been a conscious push for cities in Europe to be smarter and more sustainable, leading to the need to benchmark these cities' efforts using robust assessment frameworks. This paper ranks 28 European capital cities based on how smart and sustainable they are. Using hierarchical clustering and principal component analysis (PCA), we synthesized 32 indicators into 4 components and computed rank scores. The ranking of European capital cities was based on this rank score. Our results show that Berlin and other Nordic capital cities lead the ranking, while Sofia and Bucharest obtained the lowest rank scores, and are thus not yet on the path of being smart and sustainable. While our city rank scores show little correlation with city size and city population, there is a significant positive correlation with the cities' GDP per inhabitant, which is an indicator for wealth. Lastly, we detect a geographical divide: 12 of the top 14 cities are Western European; 11 of the bottom 14 cities are Eastern European. These results will help cities understand where they stand vis-à-vis other cities, giving policy makers an opportunity to identify areas for improvement while leveraging areas of strength.

1. Introduction

Cities are the hubs of innovation that drive the economic development of the world (Currid, 2006). Worldwide, the cities' population is growing, and it is projected that more than 60% of the population of the world will live in cities by 2030 (United Nations, 2014). However, the uncontrolled growth of a city can have adverse effects on the environment and its citizens (Annez & Buckley, 2008; Organisation for Economic Co-Operation and Development and China Development Research Foundation, 2010), and the anticipated growth of cities is expected to pose unprecedented sustainability challenges, both on infrastructures and the environment (David, 2017; Estevez, Lopes, & Janowski, 2016; Han et al., 2016). These in turn will affect the quality of life of citizens as well as the efficiency of a city's operations (Degbelo, Granell, et al., 2016). Some of these challenges are already being addressed through the development of intelligent technologies (Castán, Martínez, Menchaca, & Berrones, 2016; Degbelo, Bhattacharya, Granell, & Trilles, 2016; Vinod Kumar & Dahiya, 2017). However, many of these smart solutions are not aligned with sustainability targets, thereby generating the concept of smart sustainable cities (Ahvenniemi, Huovila, Pinto-Seppa, & Airaksinen, 2017): *A smart sustainable city is an innovative city that uses information and communication technologies (ICTs) and other means to improve quality of life, efficiency of urban operation and services, and competitiveness, while ensuring that it meets the*

needs of present and future generations with respect to economic, social, environmental as well as cultural aspects (UNECE, 2015).

Given the various interventions to improve liveability in cities, there is no better time to take a holistic view of the urban space, studying its sustainability using various dimensions (McDonnell & MacGregor-Fors, 2016; Phillis, Kouikoglou, & Verdugo, 2017). It is also important to be able to measure the performance of these interventions (Chourabi et al., 2011; Webb, Hawkey, & Tingey, 2016).

The demand for city rankings and assessment studies that address sustainability issues have increased over the past decade because cities are now seen as a leverage point in the quest for global sustainability due to the agglomeration of population in them (Grant & Chuang, 2012). Such studies serve as planning and evaluation tools for politicians, city administrators and urban planners to compare different project/policy alternatives. City rankings helps policy makers to understand how globalization and urbanization affect our urban spaces (Grant & Chuang, 2012). It is an important tool to help cities understand how they performed in the different dimensions of urban sustainability compared to other cities within the same region and identify areas for improvement.

In the past decade, several studies use the indicator-based approach to access various dimensions of urban smartness and sustainability, aggregate these dimensions and benchmark global cities based on them (Phillis et al., 2017). Some of these studies include the United Nation's

---

(UN) City Prosperity Index (UN-HABITAT, 2015), the Sustainable Cities Index (Batten, 2016), the Cities in Motion Index (Berrone, Ricart, Carraso, & Ricart, 2016), the Global Power City Index (Ichikawa, Yamato, & Dustan, 2017; Mori Memorial Foundation, 2016), the Mercer Quality of Living (Mercer, 2018), the Spatially Adjusted Liveability Index (The Economist Intelligence Unit, 2016), the CityCard Index (Grant & Chuang, 2012), the Cities of Opportunity index (PwC, 2016) and the Sustainable Assessment by Fuzzy Evaluation (SAFE) index (Phillis et al., 2017). These studies attempt to benchmark several global cities using indicators ranging from 17 to 77 in number with various weighting and aggregation methods.

Other studies have ranked more specific aspects of urban sustainability, such as urban mobility (Bojković, Petrović, & Parezanović, 2018), urban water management (van Leeuwen, Frijns, van Wezel, & van de Ven, 2012), urban air quality (Sheng & Tang, 2016) and urban economic development (Giffinger, Haindlmaier, & Kramar, 2010).

The European Union (EU) supports the movement of its cities to being smart and sustainable. This is exemplified by its conscious efforts to drive this by investing in various smart city initiatives. On the Market Place of the European Innovation Partnership on Smart Cities and Communities website, there are 34 EU projects in different cities focused on the various sector components of smart cities[1] (European Commission, 2016c). But beyond these smart city initiatives, the EU is also concerned with alleviating the various pressure that come along with urbanization as well as the sustainable development of its cities (European Commission, 2017).

In the past, multiple city rankings have been developed to benchmark European cities. These include the European Smart Cities ranking (Giffinger et al., 2007), the European Green Capital Award (Gudmundsson, 2015), the European Green City Index (Siemens, 2009), the European Green Leaf Award (European Commission, 2016a), European Soot-free City Ranking (Reh, Fellermann, & Duprez, 2013), Europe Quality of Life Index (Numbeo, 2016) and Urban Ecosystem Europe (Berrini & Bono, 2007). Although these studies have contributed to the developing discourse on sustainable strategies of cities within the European Union, they are still plagued by some methodological gaps (McManus, 2012; Meijering, Kern, & Tobi, 2014), which we aim to address with our research:

1. Lack of a proper definition of a ranking theme: In a bid to fuse several concepts and ideas into a single ranking study, the previous studies fail to provide a definition of their ranking theme (Meijering et al., 2014). A proper definition of the ranking theme is important because it gives potential users of the study a clear understanding of the multidimensional phenomenon being measured. This in turn determines the design of the theoretical framework and methodological characteristics of the ranking which influences the final ranking outcome (OECD, 2008). The ranking theme for this study is "smart sustainable cities" as defined by the United Nations Economic Council for Europe (UNECE, 2015).

2. Selection of Cities: Benchmarking cities involves comparing urban areas with diverse history, geography, features, population, trajectory and governance. This makes objective comparison very complex, hence requiring the need for a city selection criterion (Meijering et al., 2014). The urban benchmarking studies mentioned above have made use of either a geographic scope, population size or convenience sampling to select cities to build an index and rank. However, Aksoy et al. (2016) and KPMG (2010) recommend the use of a city typology to make benchmarking more meaningful. A city typology ensures that the cities being compared have a useful amount of homogeneity and is based on the city's population density, economic character, wealth, climate and history (KPMG, 2010). In this research, we selected cities as defined by the territorial

typologies for European cities and metropolitan areas (Eurostat, 2013). Specifically, we made use of cities categorized as "capital metro regions" within the European Union.

3. Data sourcing: The source of data for benchmarking cities determines the credibility of the index created and ranking done (Meijering et al., 2014). The urban benchmarking studies mentioned above have obtained data through various means including expert-group interviews, questionnaires and publicly available databases from national statistical offices. This can bring to question the consistency and coherence of the used data, which will inadvertently affect the results obtained. We propose to use Eurostat, which is a single open database from a credible source (Feldmann, 2008). This will ensure the consistency of the data being used and guarantee the reproducibility of our results.

4. Weighting: The creation of an index involves the appropriate weighting of variables used in its creation. While some of the urban benchmarking studies mentioned above are opaque about their weighting methodology, others make use of either an Equal Weighting (EW) approach or participatory methods. The Equal Weighting (EW) method is one where all variables are given equal weights (Debnath et al., 2014; Meijering et al., 2014). This is however not interesting because it assumes all variables contribute equally (without any empirical basis) to the phenomenon under study (Kahn, 2006; OECD, 2008). The participatory methods is one where various stakeholders are used to assign weights (Giffinger et al., 2010; Kahn, 2006; Mayer, 2008; Morse & Fraser, 2005; OECD, 2008). Although subjective, this method works well when there is a well-defined basis for evaluating the phenomenon under study, which is difficult to obtain for international comparisons (Munda, 2004). Both methods of weighting create a composite index without taking cognizance of the interrelationship between indicators. This leads to the creation of an index that is "indicator rich but information poor" which often confuses and misleads urban policy makers (OECD, 2008). It will be interesting to statistically explore the suitability, underlying nature and structure of the data set and use that information for weighting. In our approach, we will be making use of variance-based statistical techniques to determine the appropriateness of the selected indicators to describe smart sustainable cities, determine how the different indicators change in relation to each other and across European cities, and use this information to weight and aggregate our data.

The aim of this research is to use a well-motivated weighting scheme to create a properly defined ranking for smart sustainable cities for a clearly defined selection of cities in Europe based on open and credible data source(s). To achieve this, we make use of indicators jointly proposed by the United Nations Economic Commission for Europe (UNECE) and the International Telecommunications Union (ITU), data from Eurostat's Urban Audit database and principal component analysis (PCA) to rank European capital cities based on how smart and sustainable they are.

The indicators used were developed by UNECE-ITU after consultations with member states and various stakeholders worldwide (UNECE, 2015). The framework proposed by UNECE-ITU uses a tripartite approach under the broad areas of economy, environment, and society (Fig. 1). Each of these three broad areas are further broken down into six topics, with a number of indicators characterizing each topic. In a bid to operationalize the theoretical concept of "smart sustainable cities", we selected 32 indicators for which publicly available data could be found from the UNECE-ITU framework. These 32 indicators have been selected as a suitable balance between the depth and width of our research (Cruz-Jesus, Oliveira, Bacao, & Irani, 2017) and also to include all the thematic areas in the framework. These 32 indicators are also contained in similar frameworks such as the International Standards Organization Indicators for city services and quality of life (ISO, 2014) and the EU sustainable development strategy (European Commission,

---

[1] https://eu-smartcities.eu/eu-projects.

Fig. 1. Visual representation of the UNECE-ITU Smart Sustainable Cities Framework.

2016b).

Data on the selected indicators were obtained from Eurostat. Eurostat is the authority on statistics for the European Union, providing statistics to enable comparison between countries, regions and cities in the EU. The data used in this study are public and can be accessed and freely downloaded online[2]. Eurostat ensures that the quality and integrity of its data are not compromised by following an encompassing quality management approach and making its data suitable for research purposes (Angeloni, 2016; Eurostat, 2017; Jacinto & Soares, 2008).

The data obtained was analysed using a dimension reduction algorithm called the principal component analysis (PCA). PCA is a multivariate statistical procedure used to synthesize multiple variables by transforming the original variables into a new set of orthogonal variables in such a way that variation is emphasized and strong patterns become noticeable (Xiao, Lu, & Xu, 2017). These new sets of variables are fewer than or equal to the number of original variables and have been transformed so that a small number of principal components will account for a large part of the original data variation (Vidal, Ma, & Sastry, 2016). In doing this, PCA makes the exploration and visualization of high dimensional datasets easier.

PCA have been applied and found useful in many fields including archaeology (Jolliffe & Cadima, 2016), atmospheric science (Hannachi, Jolliffe, Stephenson, & Trendafilov, 2006), neuroscience (Hyvärinen, 2013), data mining (Metsalu & Vilo, 2015; Witten, Frank, Hall, & Pal, 2016), finance (Liao, Huang, & Wu, 2012), taxonomy (Kucharczyk, Kucharczyk, Stanislawek, & Fedor, 2012), medicine (Caprihan, Pearlson, & Calhoun, 2008; Omucheni, Kaduki, Bulimo, & Angeyo, 2014) etc. PCA has also been used in several aspects of urban studies such as local economic development (Wong, 2002), urban economics (Chen, Ding, & Liu, 2008), quality of residential environment (Tu & Lin, 2008), life expectancy (Takano, Nakamura, & Watanabe, 2002), urban heat island (Weng, Liu, Liang, & Lu, 2008), and urban remote sensing (Li and Yeh, 1998). PCA serves as an effective tool for synthesizing multidimensional data and creating new indices, which can be used for ranking (Marsal-Llacuna, Colomer-Llinàs, & Meléndez-Frigola, 2015; Wei, Huang, Li, & Xie, 2016). However, none of the studies listed above sought to rank cities using PCA based on a standardized framework.

Specifically, the objectives of this study are to:

1. Systematically reduce the number of indicators required to characterize a smart and sustainable city using open data and multivariate statistics;
2. Develop a single quantitative index to measure and rank European capital cities, based on a synthesis of the reduced set of indicators obtained in objective 1; and
3. Find the possible association of the cities' rank score with GDP and other variables.
4. Identify specific indicators which cities can leverage to significantly improve how smart and sustainable they are in relation with other European cities.

2. Data and methods

2.1. Study area

Our focus in this research is on the capital cities of the 28 member nations of the EU shown in Fig. 2, because of the unique role they play in the EU serving as hubs of innovation, growth, and diversity (European Commission, 2016; United Nations, 2014).

2.2. Data

Data for the indicators, used to define all topics under the three thematic areas of the UNECE-ITU smart sustainable cities indicators, were obtained for the EU-28 capital cities from Eurostat (European Comission, 2016). Data from the Eurostat's general and Urban Audit database have been used for similar studies including European Cities' green performance evaluation (Serbanica & Constantin, 2017), urban mobility indicator creation (Bojković et al., 2018), EU sustainable development assessment (Szopik-Depczyńska et al., 2018) and European cities smart and sustainable urban regeneration modelling (García-Fuentes et al., 2017).

As shown in Table 1, the data used for this research are made up of 78% local data for individual cities under study, 16% regional data, and 6% national data. The list of 32 indicators for which data were obtained can be found in Appendix A. The inclusion of regional and national data was necessary because of the lack of local data for certain indicators needed to achieve full characterization of all topics under the three broad areas as outlined by UNECE-ITU. However, these data have been adequately denominated to make comparison across different sized cities adequate.

2.3. Methods

In order to rank European capital cities, it is necessary to obtain a single measure of their smartness and sustainability. Since we are working with 32 variables, we made use of a two-pronged approach of feature selection to obtain a smaller number of variables to represent the larger group of 32 variables, and then feature extraction to build a new set of variables while reducing noise and redundancy in the process.

2.3.1. Data processing

Various summary statistics including the mean, median, standard deviation, and correlation matrix of all 32 variables for the 28 selected European capital cities selected were calculated using the analysis ToolPak of Microsoft excel as one of the steps to understand the underlying structure of our data (Berk & Carey, 2009).

City rank values are usually influenced by the presence of outliers in their variables, and these must therefore be taken care of. Variables containing outliers were identified as those having a distribution with absolute skewness greater or less than one (Aesaert, Voogt, Kuiper, & van Braak, 2017; Groeneveld & Meeden, 1984). Boxplots and histograms were plotted for each variable to further understand outliers. Variables identified to be skewed were transformed using the

---

[2] http://ec.europa.eu/eurostat/web/main/home.

EU-28 Capital Cities

Fig. 2. Map showing EU 28 with their respective capital cities.

Table 1
Data used.

Database	Spatial Level	Number of indicators
Urban Audit	Local – NUTS3	25
Eurostat Database	Regional – NUTS2	5
Eurostat Database	National	2
Total		32

powerTransform function in R. This uses the maximum likelihood approach of Box & Cox (1964) to select the appropriate transformation power, which was applied on the relevant variables.

Thereafter we normalized our data, scaling down values of the indicators. Normalization is crucial in PCA because it is a variance maximizing exercise and sensitive to the relative scaling of the original variables. This step prevents one variable from dominating all others, thus enabling the data analysis method to treat the data "fairly" (Kotsiantis, Kanellopoulos, & Pintelas, 2006). To normalize our data we made use of the minimum and maximum value of each variable. In this way, we ensured that the values of each variable range between 0 and 1 (Bannerjee, Bone, & Finger, 2016).

2.3.2. Feature selection

Clustering was used as a robust method to identify homogenous group of variables called "clusters" sharing similar characteristics across all cities under study such that these clusters have "maximum internal homogeneity (withing the cluster) and maximum external heterogeneity (between clusters)" (Cruz-Jesus et al., 2017). We made use of the FactoMineR package in R to perform a hierarchical clustering on all 32 variables grouped according to the EU28 capital cities to obtain an optimal number of clusters (Husson, Lê, & Pagès, 2010). The basic algorithm for the hierarchical clustering as applied to our study can be described as follows:

1. There are 32 objects (points) to classify.
2. We find the closest two points and merge them into a new point.
3. We compute the similarity (distance) between this new point and the reminaing points.
4. We repeat steps 2 and 3 until there is only one point remaining.

We then made use of the variance based feature reduction technique called "low variance filter" (Kouser, Lavanya, Rangarajan, & Acharya Kshitish, 2016). Here, we calculated the variance of each variable in a cluster and removed those parameters with variance below a certain threshold. This is achieved by arranging the variance of all variables in a cluster in descending order and adding the variance of each variable (starting with the variable with the largest variance) until a specified threshold Z% is reached. Thereafter, the remaining variables are discarded. This step ensures that we retain only variables that hold sufficient information in each cluster.

2.3.3. Feature extraction

PCA was applied to selected variables to transform the data from a high-dimensional space to a low-dimensional space. To test the suitability of our data for reduction, we made use of the Bartlett's test of sphericity. The Bartlett's measure tests for our correlation matrix being an identity matrix, which indicates that there is some relationship between our variables (Bartlett, 1937; Doyle et al., 2017). We also performed the Kaiser-Meyer-Olkin (KMO) test to measure the sampling adequacy of our data for PCA (Cruz-Jesus et al., 2017). We thereafter did a PCA using the following steps:

1. We calculated the correlation matrix of all variables obtained from the low variance filter in the previous step.
2. Using the correlation matrix, we deduced the eigenvector and eigenvalue. The eigenvector indicates the direction of our new axis and the eigenvalue indicates the magnitude of variability in the new axis.

3. We multiplied our original data with the eigenvectors to rotate our data to align with the new axis (principal components).

The feature selection and feature extraction was implemented in R (Husson et al., 2010; Lê, Josse, & Husson, 2008).

For this research the Kaiser's criterion was used to select appropriate principal components for further analysis (Friesen, Seliske, & Papadopoulos, 2016; Scariano, 2013). Kaiser (1960) proposed selecting principal components with eigenvalues greater than one (1) and Humphreys & Montanelli (1975) recommended the Kaiser's criterion for large correlation matrices such as ours.

To determine contributing variables to each principal component under consideration, variable eigenvectors were investigated (Friesen et al., 2016). We made use of the contribution of each variable to the inertia explained by axis (CTA) and the part of variance of each variable explained by each axis (CTR) (Isnard & Sautory, 1994; Koch, 2013). We made use of CTA and CTR because they can be used in interpreting our results statistically and geometrically (Abdi & Williams, 2010).

Each component selected was weighted based on its variance in proportion to the total variance of all selected components. Variable loadings were multiplied by each selected principal component's weight and summed up to obtain a factor score. From the normalized table, we then calculated the coordinate of each capital city in relation to this factor score, multiplying the coordinate of vectors representing these cities by their respective factor score and summing them together (Friesen et al., 2016). This is more fully explained in Appendix B.

3. Results and discussion

3.1. Hierarchical clustering

As a first step in reducing the dimensionality of our data, hierarchical clustering was done to identify groups of similar variables in our dataset. We explored our hierarchical cluster solution using the dendogram shown in Fig. 3. The dendogram lists all the variables represented as numbers, which are clustered along the x-axis and the distance at which the clusters are formed at each level on the y-axis. Five clusters were selected to characterize the structure of our data (indicated using coloured boxes on the x axis in Fig. 3). Taking a closer look at variables in each cluster, it can be seen that, with the exception of a few variables, each cluster tends to contain variables that belong to the same topic and area in the UNECE-ITU smart sustainable cities framework.

The individual factor map in Fig. 4 is a plot of the principal component for variables on the first two principal components. It reveals the structural relationship between the variables, the cities, and the components. The first component accounts for 64.05% of the total variance while the second component accounts for an additional 15.63%. Together they account for a total variance of 79.68%. From it, we can see that variable 7, which is the length of dedicated bicycle lanes, has a high score on component one, while variable 23, which is the percentage of total deaths, has a low score. The five selected clusters can also be seen on the factor map (Fig. 4; same colour codes used as in Fig. 3).

3.2. Principal component analysis

After reducing the number of variables using the low variance filter approach highlighted in our methodology above, 15 variables were selected as shown in Table 2.

We performed the Bartlett's measure test and obtained a *p*-value of $2.54 * 10^{-5}$, which is less than 0.05 and is statistically significant. Hence, we can perform a PCA on our dataset. Furthermore, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy for our overall dataset yielded a value of 0.65 indicating a relative compactness of the patterns of the correlations in our dataset (Cruz-Jesus et al., 2017). Our

principal component analysis should therefore yield distinct and reliable results.

Performing a PCA on the variables in Table 2 gives a total of 15 components. In Table 3, the proportion of variance indicates how much of the total variance a principal component has. The first principal component explains the greatest amount of the total variance of our data. The total amount of variance explained by subsequent components decreases with their distance from PC1. However, only the first four principal components are of interest to us because their eigenvalues are greater than one (Kaiser, 1960). 69.1% of variance can be explained by four principal components retained. All other components are ignored.

The first principal component represents 32.5% of the total variance, as shown in Table 3. Variables with CTA values greater than the average in an axis, Table 4, are considered to contribute significantly to that axis. These CTA values have been highlighted in grey in Table 4. For the vector generating axis one, this includes the number of patent applications made to the European Patent Office per million of active population (V3), the percentage of e-commerce, customer relation management (CRM) and secure transactions in a city (V5), the length of bicycle network (V7), the number of days particulate matter (PM10) concentrations exceed $50\,\mu g/m^3$ (V8), the share of urban waste water load treated to applicable standards (V12), share of protected terrestrial area (V17), share of deaths (V23) and Gini coefficient of disposable income (V32). A further analysis of CTR values shows that the percentage of voter turnout in national and EU parliamentary elections (V31) is a variable to include in the explanation of the first axis because its CTA is quite close to the average CTA and the first axis explains the main part of the variance associated to this variable. The second principal component represents an additional 14.5% of the total variance with percentage of persons employed between the ages of 20 and 64 (V4), number of days of particulate matter concentration exceeding $50\,\mu g/m$ (V8), number of theatres (V29), and gender pay gap (V30) contributing significantly to this component. PC3 represents 12.7% of the total variance of the greenhouse gas emissions from transport (V10), electricity generated from renewable energy (V19), and the percentage of voter turnout in national and EU parliamentary elections (V31) contribute significantly to the inertia associated with this component. PC4 represents 9.4% of the total variance and is highly influenced by the percentage of persons employed between the ages of 20 and 64 (V4), the amount of greenhouse gas emissions (V10), protected terrestrial areas (V17), and Gini coefficients of disposable income (V32). Overall, all variables contribute to the four selected components.

The factor scores are a weighted summation of the four principal loadings (Table 4).

Finally, we multiplied the factor scores of each variable by their respective normalized values of indicators for each city and summed in order to obtain a rank score. This was arranged in descending order to give a ranking of the capital cities as shown in Figs. 5 and 6.

3.3. Sensitivity analysis

Beyond ranking cities based on how smart and sustainable they are, this research aims to help decision makers by pinpointing those indicators that have a huge effect in determining the ranking. This was done by eliminating the factor score of each of the final variables and comparing the resulting ranking of cities with the original ranking (Saisana & Saltelli, 2011). Table 5 shows the top three indicators that affects the ranking of each city. These indicators serve as leverage points which city councils and urban planners can use to either maintain or improve the overall sustainability of their city.

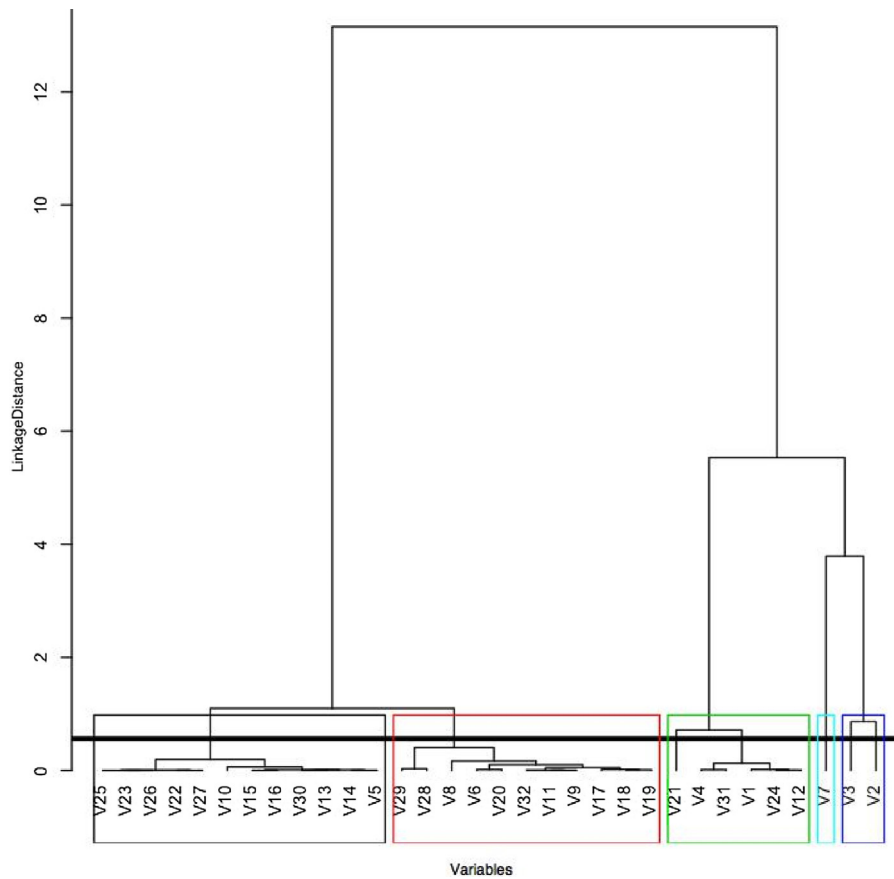We discuss the ranking and its relation to cities' geographies and demographics in the following subsections.

Fig. 3. Hierarchical Clustering Dendogram.

3.4. Rank score comparison with Gross Domestic Product (GDP)

A scatterplot of the rank score of each capital city and its Gross Domestic Product (GDP) obtained from Eurostat, shown in Fig. 7, reveals a relationship between how smart and sustainable a city is, according to our calculated rank scores, and its wealth. Except for Berlin

and Luxemburg, which are outliers, the strength of this relationship is measured using a correlation coefficient of 0.80, which is statistically significant at 5%. Wealthier cities primarily located in Western and Northern Europe tend to be ranked better than other cities. This is because initiatives and projects that drive the competitiveness of a city in terms of how smart and sustainable the city is are usually capital
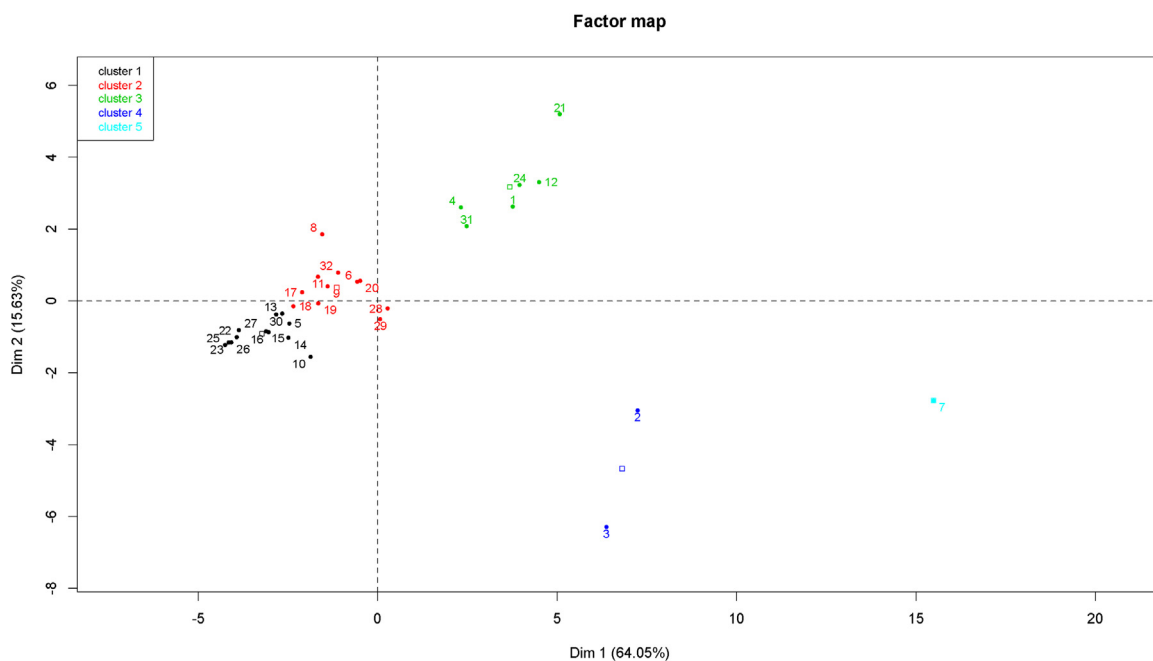

Fig. 4. Individuals factor map.

Table 2
Variables used in ranking.

Patent applications to the EPO per million of active population	V3
Persons employed between the ages of 20 and 64 (%)	V4
E-commerce, Customer Relation Management (CRM), and secure transactions (%)	V5
Length of bicycle network (dedicated cycle paths and lanes) (km)	V7
Number of days particulate matter PM10 concentrations exceed 50 μg/m³	V8
Greenhouse gas emissions from transport (million tonnes)	V10
Share of the urban waste water load (in population equivalents) treated according to the applicable standard (%)	V12
Proportion of population living in households considering that they suffer from noise (%)	V13
Protected terrestrial area (%)	V17
Electricity generated from renewable sources (%)	V19
Share of total deaths per year (%)	V23
Number of theatres	V29
Gender pay gap in unadjusted form (%)	V30
Voter turnout in national and EU parliamentary elections (%)	V31
Gini coefficient of equalized disposable income	V32

Table 3
Principal Component Eigenvalues.

Component	Eigenvalue	Difference	Proportion	Cumulative	PCw (%)
PC1	4.87	2.71	32.50	32.50	47.05
PC2	2.17	0.27	14.46	46.95	20.93
PC3	1.89	0.48	12.66	59.61	18.33
PC4	1.42	0.43	9.45	69.06	13.68

intensive, giving wealthier governments an advantage (Siemens, 2009). In contrast, poorer governments are usually more focused on more basic developmental issues.

A temporal analysis of GDP per capita shows that although cities in central and Eastern Europe have an average yearly growth of 3.7%, this growth rate is still less than 75% of the EU average. This deprives these cities of the much-needed funds needed for research and development (R&D) to foster innovation and pivot towards a more sustainable future. R&D is the heart of smart sustainable development and requires a lot of money (Serbanica & Constantin, 2017).

Furthermore, being wealthier gives governments leverage in setting more ambitious policy goals for their cities compared to less wealthy governments. For example, Copenhagen has an ambitious climate plan to be the first carbon-neutral capital city by 2025 (Bodansky, 2010). Stockholm also aims to have vehicle fleets completely rid of fossil fuel by 2030. Berlin, the only capital city in Europe with its GDP per inhabitant lower than the national average, benefits from well-crafted polices in air quality, energy, and environmental governance (European Commission, 2016). The GDP of Luxemburg is extremely high because

the country has an unusual financial and tax system and serves as a host to many international companies (Annaert, 2004; European Commission, 2014a). Wealth and government policies can also be thought of in a feedback loop in which money (wealth) is needed to be able to set ambitious policy goals and craft carefully designed policies, which in turn help the government to save more money. For example, using policies to drive energy efficiency in buildings and vehicles can save money and cut emissions (Hughes, Chu, & Mason, 2018; Zhou et al., 2016). Hence, there needs to be a balance of maximizing the performance of a city with as little money as possible.

3.5. Rank score comparison with geographical location

An important insight is gained from the spatial pattern of the ranking results. A visual inspection of Fig. 6 reveals that similar ranking values are clustered together in the map. A test of spatial autocorrelation using Global Moran I reveals that there is a positive spatial autocorrelation (0.31) among the city ranks. We can therefore infer that the performance of each city in our ranking is not randomly spatially distributed but each city influences its neighbours making cities with similar behaviours clustered together. This is responsible for a geographical divide between cities that are well ranked and those that are not. Twelve of the top 14 cities are in Western Europe while 11 of the bottom 14 cities are in Eastern Europe. This result corroborates findings by the European Union which identified a developmental gap between western and Eastern European cities (European Commission, 2014b). It is interesting to note that Budapest, Prague, Bratislava, Warsaw, Ljubljana, Vilnius, Riga, Tallinn, Sofia and Bucharest all belonged to the former Communist Bloc until 1990 and went through several years of transitions (Roaf, Atoyan, Joshi, & Krogulski, 2014; Serbanica & Constantin, 2017). Although, the communist laws did not entirely neglect the environment, industries were not adequately incentivised to adopt more efficient processes and adhere to the laws (Constantin, 1999; Hirt & Stanilov, 2009). The collapse of the communist bloc led to a change in the existing urban patterns with an increase in private car usage, a decrease in open and green spaces and a conversion of garages and ground floors of buildings into shops and offices (European Commission, 2016b; Hirt & Stanilov, 2009). We hypothesize that joining the European Union played a role in driving cities in central and Eastern Europe towards a smart and sustainable path. This is evidenced by the fact that Valletta, Riga, Budapest, Vilnius, Warsaw, Bratislava, Ljubljana are cities in countries that joined the EU in 2004 while the two least ranked cities (Sofia and Bucharest) are in countries joined the EU in 2007. Prague and Tallinn are exceptions because even though they were a part of the communist bloc, they are ranked in the top 14 cities. These two cities are located in countries categorised as "fast-track reforming states" because of their high exposure to globalization,
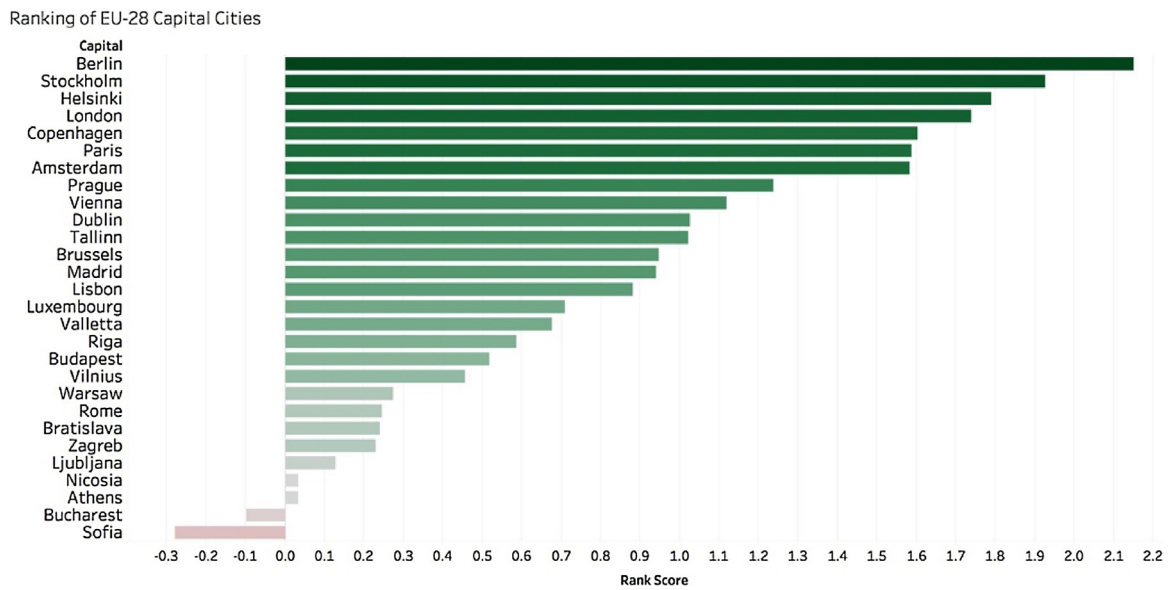
Table 4
Indicator Variable Loadings*1000, CTA, CTR*1000 and factor score*1000.

Variable	μ1	CTA	CTR	μ2	CTA	CTR	μ3	CTA	CTR	μ4	CTA	CTR	Factor Score
V3	0.853	14.9	728	−0.089	0.36	8	0.034	0.06	1	−0.084	0.50	007	378
V4	0.207	0.88	43	0.716	23.6	512	−0.119	0.75	14	0.381	10.2	145	278
V5	0.829	14.1	687	0.070	0.23	5	−0.014	0.01	0	−0.125	1.09	15	385
V7	0.804	13.3	647	−0.040	0.08	2	0.252	3.34	63	−0.104	0.76	11	402
V8	−0.587	7.07	345	−0.520	12.4	270	−0.037	0.07	1	0.305	6.56	93	−350
V10	0.420	3.62	176	−0.352	5.7	124	0.636	21.3	404	0.356	8.92	126	289
V12	0.693	9.86	481	0.302	4.20	91	−0.175	1.61	31	0.184	2.38	34	382
V13	0.211	0.92	45	−0.302	4.22	91	−0.346	6.30	120	0.696	34.2	484	068
V17	−0.699	10	490	−0.129	0.77	17	0.015	0.01	0	−0.263	4.90	69	−390
V19	0.094	0.18	9	0.072	0.24	5	0.746	29.3	556	−0.144	1.46	21	176
V23	−0.641	8.42	411	0.285	3.74	81	0.277	4.04	77	0.193	2.63	37	−165
V29	0.305	1.91	93	−0.554	14.1	306	0.493	12.8	243	0.254	4.55	65	153
V30	0.267	1.47	72	0.694	22.2	482	0.252	3.33	63	0.188	2.49	35	343
V31	0.565	6.54	319	−0.337	5.22	113	−0.553	16.1	306	0.098	0.68	10	107
V32	−0.575	6.79	331	0.248	2.83	61	0.136	1.0	19	0.515	18.7	265	−123

Fig. 5. Bar chart showing rank scores of European capital cities.

extreme "EU-ization" influences and creative deployment of technology to foster a sustainable and inclusive society (Ian, Kaliopa, & Natasa, 2003; Nam & Pardo, 2011). Furthermore, 3 of the top ranked 5 cities are cities from the Scandinavian region of Northern Europe. Stockholm, Helsinki, and Copenhagen have very strong environmental policies and are focused on improving the quality of life of their citizens (Lindström & Eriksson, 1993).

3.6. Rank score comparison with size and population

City's size and population can be either an advantage or a drawback in determining how smart and sustainable it is (Siemens, 2009). All other things being held constant, a city should be able to coordinate the activities of a million residents better than that of ten million residents (Mori & Christodoulou, 2012). However, the city with ten million residents has a leverage resource-wise to pursue smart and sustainable
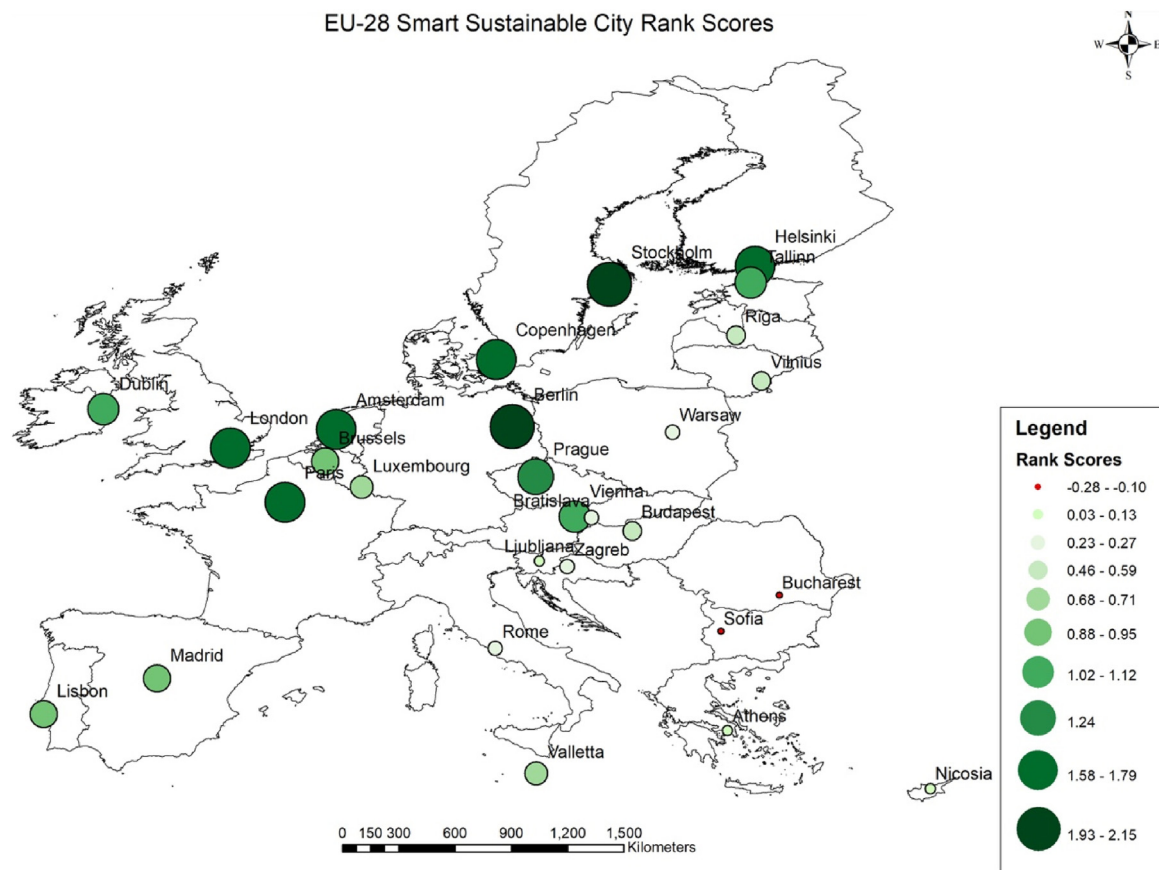

Fig. 6. Smart Sustainable ranking of European capital cities.

Table 5
Smart Sustainable City ranking and top three indicators that affects ranking.

Rank	City	Top three indicators that affects ranking
1	Berlin	Bicycle network, Wastewater treatment, E-commerce
2	Stockholm	Wastewater treatment, Patent applications, E-commerce
3	Helsinki	Wastewater treatment, Patent applications, Bicycle network
4	London	Wastewater treatment, Bicycle network, E-commerce
5	Copenhagen	Wastewater treatment, E-commerce, Patent applications
6	Paris	Wastewater treatment, Unemployment, GHG emissions
7	Amsterdam	Wastewater treatment, E-Commerce, Unemployment
8	Prague	Wastewater treatment, E-Commerce, Gender inequality
9	Vienna	Gender Inequality, Patent applications, E-Commerce
10	Dublin	E-Commerce, Wastewater treatment, Unemployment
11	Tallinn	Wastewater treatment, Gender Inequality, Unemployment
12	Brussels	E-Commerce, Wastewater treatment, Patent application
13	Madrid	Wastewater treatment, Protected terrestrial area, E-Commerce
14	Lisbon	Wastewater treatment, Unemployment, E-Commerce
15	Luxembourg	Wastewater treatment, Protected terrestrial area, Unemployment
16	Valletta	Wastewater treatment, PM10 concentration, E-commerce
17	Riga	Wastewater treatment, Unemployment, Gender inequality
18	Budapest	Wastewater treatment, Unemployment, Protected terrestrial area
19	Vilnius	Unemployment, E-commerce, Gender inequality
20	Warsaw	Wastewater treatment, PM10 concentration, Protected terrestrial area
21	Rome	GHG emissions, Protected terrestrial area, PM10 concentration
22	Bratislava	Protected terrestrial area, Gender inequality, E-commerce
23	Zagreb	Protected terrestrial area, E-commerce, Unemployment
24	Ljubljana	Protected terrestrial area, E-commerce, Waste water treatment
25	Nicosia	Protected terrestrial area, PM10 concentration, Unemployment
26	Athens	Protected terrestrial area, Wastewater treatment, PM10 concentration
27	Bucharest	PM10 concentration, Wastewater treatment, Protected terrestrial area
28	Sofia	PM10 concentration, Protected terrestrial area, Unemployment

policies and infrastructure (Munda, 2006). We found no statistically significant correlation between a city rank and its population and size. This is in line with findings by (Serbanica & Constantin, 2017) who concluded that green cities can be equally small, medium or large. In fact, contrary to expectations, the number of registered cars in smaller cities are usually more than the number of registered cars in larger cities. This can be attributed to a more developed public transport system in larger cities, reducing the need to own private vehicles (Serbanica & Constantin, 2017).

3.7. Further discussion

In terms of social development, Sofia and Bucharest, the 2 least ranked cities, are in countries with very high level of poverty when compared with the European average. Other cities like Budapest, Zagreb, Athens, Nicosia, Ljubjana have a "lower then European average" quality of healthcare systems, employment rates and amount of disposable incomes and are ranked very low. This shows the importance of the social systems in a city in enabling an inclusive society.

The environment also plays a huge role in determining the sustainability of cities. Our sensitivity analysis shows that a reduction of PM10 concentration in Sofia will significantly improve its ranking. In Sofia, the population-weighted concentration of PM10 exceeds the annual EU limit value of $40\,\mu g/m^3$. Nicosia and Warsaw also have very high values of PM10 concentration. The other cities studies have values of PM10 well below the annual EU with northern European countries recording the lowest.

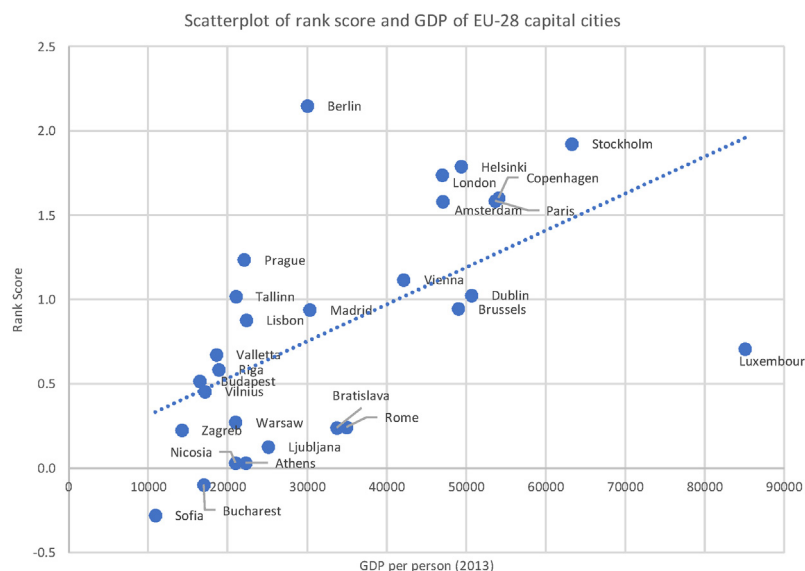Citizen engagement is another relevant issue for a smart and

Fig. 7. Scatterplot of the rank score and GDP of EU-28 capital cities.

sustainable city. Beyond government policies, the individual actions of citizens can collectively have more influence than policies in determining how smart and sustainable a city is (Berry & Portney, 2013). Such actions include the cultivation of an energy saving culture in households, sorting of waste, the decision to commute using ride-sharing rather than private vehicle, amongst others (Fellows & Pitfield, 2000; Poortinga, Steg, & Vlek, 2004; Sharholy, Ahmad, Mahmood, & Trivedi, 2008). A research project by Siemens (2009) showed a high correlation between citizen engagement and the green rank score of cities. According to this report, "about three-quarters of the existing technological changes that would help London to meet its long-term carbon reduction targets depended on the decisions of citizens or companies, not of governments" (Denig, 2011; Siemens, 2009). Citizens' actions and attitudes can be influenced through incentives and penalties that encourage a change in behaviour (Osbaldiston & Schott, 2012). Education and public awareness also go a long way in arming the public with knowledge to make good and informed decisions that affect the ranking of a city (Tilbury, 1995).

4. Validation

To validate our study, we compared the ranking for smart sustainable cities in Europe with other related European urban ranking systems in Table 6. The smart sustainable cities (SSC) ranking has eight cities in common with the European Soot-free city ranking in the top-ten list, seven common cities with the European Green city index ranking, four common cities with the European Quality of Life Index ranking and no common cities with the European medium-sized cities ranking. Using a Kendall's τ rank correlation test, we see that our ranking is strongly correlated with the European Soot-free city ranking and the European Green City Index but moderately correlated with the European Quality of Life Index. Although, we had a Kendall τ correlation coefficient of one (1.00) between our ranking and the European medium-sized cities ranking, this is because the two ranking systems had only 2 cities in common with similar ranks in both ranking systems. These results show that although our ranking correlates with other ranking systems of related philosophy, it still sufficiently differs because of the unique ranking attribute and methodological characteristics of our study.

Finally, it should be noted that this ranking is influenced by the set of final variables selected to characterize how smart and sustainable a city is and the year of the data used. However, as demonstrated by our methodology, the selected variables can be appropriately used as a representative sample of the indicators in the UNECE smart sustainable framework.

5. Conclusions

Cities can be viewed as the source of and the solution to many of today's economic, social, and environmental challenges. Because of this, the EU is promoting various initiatives to drive cities to be more sustainable, resource-efficient, and inclusive. In this article, we ranked European capital cities based on how smart and sustainable they are, using a selection of indicators from a framework proposed by UNECE-ITU. We did this by obtaining publicly available data, from Eurostat, on the indicators in the framework and systematically reducing the number of indicators using multivariate statistics. Using hierarchical clustering, we created five (5) homogenous groups of indicators, selected representative indicators for each group using variance as the selection criterion and then applied PCA to the selected indicators to obtain a composite index. The ranking is based on a composite index which conceals multiple subjective assessment under a veil of objectivity. Furthermore, we carried out a sensitivity analysis and validation study of our results.

Relating our ranking of European capital cities with geographical and demographic parameters, we found that Nordic cities and cities in Western Europe perform better in our ranking than cities in Eastern Europe. Using GDP per inhabitant as an indicator for wealth, we determined that wealthier cities perform better in our ranking. Finally, we detected no correlation between a city's rank score and its population and size.

Our method is an effort to simplify and summarize a very complex concept into a manageable form. It should be noted that PCA is completely non- parametric: any data set can be plugged in and an answer comes out, requiring no parameters to tweak and no regard for how the data were recorded. From one perspective, the fact that PCA is non-parametric can be considered as a positive feature because the answer is unique and independent of the user. From another perspective, the fact that PCA ignores the source of the data is also a weakness. However, we have taken steps to ensure the quality of our data to mitigate this.

Overall, this research has contributed to knowledge by using a multivariate data analysis approach to rank capital metro regions within the European Union using data from a single open database based on how smart and sustainable they are. This approach ensures that the cities being compared have a useful amount of homogeneity and uses the underlying structure of the dataset to weight and aggregate our data while guaranteeing the consistence, coherence and reproducibility of our results.

This ranking is meant to attract attention and induce competition amongst cities. By utilizing the method and result of this research, cities and their stakeholders will not only be able to objectively assess the extent to which they may be perceived as being smart and sustainable, but also be able to identify leverage points to improve their sustainability.

Conflicts of interest

None declared.

Funding

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the

Table 6
Top 10 cities in four European ranking systems and their correlation with Smart Sustainable Cities Ranking.

SSC Ranking	European Soot-free City Ranking	European Green city Index	Europe Quality of Life Index	European Smart Cities Ranking
Berlin	Zurich	Copenhagen	Zurich	Luxembourg
Stockholm	Copenhagen	Stockholm	Frankfurt	Aarhus
Helsinki	Vienna	Oslo	Munich	Turku
London	Stockholm	Vienna	Edinburgh	Aalborg
Copenhagen	Berlin	Amsterdam	Trondheim	Odense
Paris	Helsinki	Zurich	Geneva	Tampere
Amsterdam	London	Helsinki	Vienna	Oulu
Prague	Paris	Berlin	Copenhagen	Eindhoven
Vienna	Stuttgart	Brussels	Stockholm	Linz
Dublin	Amsterdam	Paris	Berlin	Salzburg
Number of cities	23	30	58	70
Common cities with SSCR	15	24	28	2
Kendall's τ rank correlation with SSCR	0.70	0.61	0.49	1.00

online version, at https://doi.org/10.1016/j.scs.2018.10.009.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433–459. https://doi.org/10.1002/wics.101.

Aesaert, K., Voogt, J., Kuiper, E., & van Braak, J. (2017). Accuracy and bias of ICT self-efficacy: An empirical study into students' over- and underestimation of their ICT competences. *Computers in Human Behavior, 75*(Supplement C), 92–102. https://doi.org/10.1016/j.chb.2017.05.010.

Ahvenniemi, H., Huovila, A., Pinto-Seppä, I., & Airaksinen, M. (2017). What are the differences between sustainable and smart cities? *Cities, 60*, 234–245. https://doi.org/10.1016/j.cities.2016.09.009.

Aksoy, E., Gregor, M., Fons, J., Garzillo, C., Cugny-Seguin, M., Löhnertz, M., et al. (2016). City typologies of Europe: A tool to support urban sustainability studies and practices. *WIT Transactions on Ecology and the Environment, 204*(Sc), 199–210. https://doi.org/10.2495/SC160171.

Angeloni, S. (2016). Data on some socio-economic parameters explaining the movement of extra-EU asylum seekers in Europe. *Data in Brief, 9*, 966–969. https://doi.org/10.1016/j.dib.2016.11.017.

Annaert, J.-L. (2004). ECFIN country focus A bright spot in the heart of Europe: What can we learn from the Luxembourg success story? *Focus, 1*(8), 1–9.

Annez, P. C., & Buckley, R. M. (2008). *Urbanization and growth: Setting the context.* https://doi.org/10.1596/978-0-8213-7573-0.

Bannerjee, S., Bone, J., & Finger, Y. (2016). *European Digital City Index – Methodology Report* https://doi.org/978-1-84875-153-8.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 160*(901), 268–282 http://www.jstor.org/stable/96803.

Batten, J. (2016). *Sustainable Cities Index 2016. Arcadis Global.*

Berk, K. N., & Carey, P. M. (2009). (3rd ed.). *Data analysis with microsoft excel: Updated for officeVol. 2007.* Boston, MA: Duxbury Press.

Berrini, M., & Bono, L. (2007). *Urban ecosystem europe.* Europe: Rome.

Berrone, P., Ricart, J. E., Carraso, C., & Ricart, R. (2016). *IESE cities in motion index, 2016,* 60. https://doi.org/10.15581/018.ST-471.

Berry, J. M., & Portney, K. E. (2013). Sustainability and interest group participation in city politics. *Sustainability (Switzerland), 5*(5), 2077–2097. https://doi.org/10.3390/su5052077.

Bodansky, D. (2010). The Copenhagen climate change conference: A postmortem. *American Journal of International Law, 104*(2), 230–240. https://doi.org/10.5305/amerjintelaw.104.2.0230.

Bojković, N., Petrović, M., & Parezanović, T. (2018). Towards indicators outlining prospects to reduce car use with an application to European cities. *Ecological Indicators, 84*, 172–182. https://doi.org/10.1016/j.ecolind.2017.08.061.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B: Statistical Methodology, 26*(2), 211–252 Retrieved from http://www.econ.illinois.edu/~econ508/Papers/boxcox64.pdf.

Caprihan, A., Pearlson, G. D., & Calhoun, V. D. (2008). Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. *NeuroImage, 42*(2), 675–682. https://doi.org/10.1016/j.neuroimage.2008.04.255.

Castán, J. A., Martínez, S. I., Menchaca, J. L., & Berrones, M. G. T. (2016). Improving vehicular mobility in urban traffic using ubiquitous computing. *Journal of Computer and Communications, 4*(August), 57–62. https://doi.org/10.4236/jcc.2016.410006.

Chen, C., Ding, Y., & Liu, X. (2008). City economical function and industrial development: Case study along the railway line in North Xinjiang in China. *Journal of Urban Planning and Development, 134*(4), 153–158. https://doi.org/10.1061/(ASCE)0733-9488(2008)134:4.(153).

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., et al. (2011). Understanding smart cities: An integrative framework. *Proceedings of the Annual Hawaii International Conference on System Sciences,* 2289–2297. https://doi.org/10.1109/HICSS.;1; 2012.615.

Constantin, D. L. (1999). Environmental preservation strategy in the New Romania: Institutional and behavioural challenges. In S. C. Lonergan (Ed.). *Environmental change adaptation, and security* (pp. 317–333). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-011-4219-9_22.

Cruz-Jesus, F., Oliveira, T., Bacao, F., & Irani, Z. (2017). Assessing the pattern between economic and digital development of countries. *Information Systems Frontiers, 19*(4), 835–854. https://doi.org/10.1007/s10796-016-9634-1.

Currid, E. (2006). New York as a global creative hub: A competitive analysis of four theories on World cities. *Economic Development Quarterly, 20*(4), 330–350. https://doi.org/10.1177/0891242406292708.

David, D. (2017). Environment and urbanization. *The International Encyclopedia of Geography, 24*(1), 31–46. https://doi.org/10.1002/9781118786352.wbieg0623.

Debnath, A. K., Chin, H. C., Haque, M. M., & Yuen, B. (2014). A methodological framework for benchmarking smart transport cities. *Cities, 37*, 47–56. https://doi.org/10.1016/j.cities.2013.11.004.

Degbelo, A., Bhattacharya, D., Granell, C., & Trilles, S. (2016a). *Toolkits for smarter cities: A brief assessment. UCAmI 2016-10th international conference on ubiquitous computing & ambient intelligence.* Las Palmas, Gran Canaria: Springer International Publishing431–436. https://doi.org/10.1007/978-3-319-48799-147.

Degbelo, A., Granell, C., Trilles Oliver, S., Bhattacharya, D., Casteleyn, S., & Kray, C. (2016b). Opening up smart cities: Citizen-centric challenges and opportunities from GIScience. *ISPRS International Journal of Geo-Information, 5*(2), 16. https://doi.org/10.3390/ijgi5020016.

Denig, S. (2011). *Sustainable urban infrastructure: London edition – A view to 2025, London.* Retrieved from https://w5.siemens.com/web/si/sl/corporate/portal/raziskave/Documents/sustainable_urban_infrastructure-study_london.pdf.

Doyle, K., Sainsbury, K., Cleary, S., Parkinson, L., Vindigni, D., McGrath, I., et al. (2017). Happy to help/happy to be here: Identifying components of successful clinical placements for undergraduate nursing students. *Nurse Education Today, 49*, 27–32. https://doi.org/10.1016/j.nedt.2016.11.001.

Estevez, E., Lopes, N. V., & Janowski, T. (2016). Smart sustainable cities. *Reconnaissance Study,* 330.

European Commission (2016). Urban Europe: Statistics on cities. *Towns and Suburbs.* https://doi.org/10.2785/91120.

European Commission (2014a). *European Economy: Macroeconomic Imbalances Luxembourg, 3209.* https://doi.org/10.2765/74234.

European Commission (2014b). *The Urban Dimension of EU Policies – Key Features of an EU Urban Agenda, 12.* Retrieved from http://ec.europa.eu/regional_policy/consultation/urb_agenda/index_en.cfm.

European Commission (2016a). *European Green Capital Award and European Green Leaf Award. Brussels.* Retrieved from http://ec.europa.eu/environment/europeangreencapital/europeangreenleaf/.

European Commission (2016b). *Sustainable development in the European Union.* https://doi.org/10.1002/(SICI)1099-0976(199711)7:6<181:AID-EET134>3.0.CO;2-S.

European Commission (2016c). *The marktketplace of the European Innovation Partnership on Smart cities and Communities (EIP-SCC). Brussels.* Retrieved from http://ec.europa.eu/eip/smartcities/files/operational-implementation-plan-oip-v2_en.pdf.

European Commission (2017). *Sustainable development in the European Union: Monitoring report on progress towards the SDGs in an EU context256*https://doi.org/10.1002/(SICI)1099-0976(199711)7:6<181:AID-EET134>3.0.CO;2-S.

Eurostat (2013). *Territorial typologies for European cities and metropolitan regions. Luxembourg,* 2443–8219.

Eurostat (2017). *Final report of the quality of life indicators*Luxembourg: Publications office of the European Union https://doi.org/10.2785/021270.

Feldmann, B. (2008). The Urban Audit — Measuring the quality of life in European cities. *Eurostat – Statistics in Focus, 82* Retrieved from http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-08-082/EN/KS-SF-08-082-EN.PDF.

Fellows, N. T., & Pitfield, D. E. (2000). An economic and operational evaluation of urban car-sharing. *Transportation Research Part D: Transport and Environment, 5*(1), 1–10. https://doi.org/10.1016/S1361-9209(99)00016-4.

Friesen, C. E., Seliske, P., & Papadopoulos, A. (2016). Using principal component analysis to identify priority neighbourhoods for health services delivery by ranking socio-economic status. *Online Journal of Public Health Informatics, 8*(2), e192. https://doi.org/10.5210/ojphi.v8i2.6733.

García-Fuentes, M., Quijano, A., De Torre, C., García, R., Compere, P., Degard, C., et al. (2017). European cities characterization as basis towards the replication of a smart and sustainable urban regeneration model. *Energy Procedia, 111*(September 2016), 836–845. https://doi.org/10.1016/j.egypro.2017.03.246.

Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N., & Meijers, E. (2007). *Ranking of European medium-sized cities*Vienna Retrieved from http://www.smart-cities.eu/download/smart_cities_final_report.pdf.

Giffinger, R., Haindlmaier, G., & Kramar, H. (2010). The role of rankings in growing city competition. *Urban Research & Practice, 3*(3), 299–312. https://doi.org/10.1080/17535069.2010.524420.

Grant, K. A., & Chuang, S. (2012). An aggregating approach to ranking cities for knowledge-based development. *International Journal of Knowledge-Based Development, 3*(1), 17. https://doi.org/10.1504/IJKBD.2012.045558.

Groeneveld, R. A., & Meeden, G. (1984). Measuring Skewness and Kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician), 33*(4), 391–399 Retrieved from http://www.jstor.org/stable/2987742.

Gudmundsson, H. (2015). The European Green Capital Award. Its role, evaluation criteria and policy implications. *Toshi Keikaku, 64*(2), 15.

Han, J., Meng, X., Zhou, X., Yi, B., Liu, M., & Xiang, W.-N. (2016). A long-term analysis of urbanization process, landscape change, and carbon sources and sinks: A case study in China's Yangtze River Delta region. *Journal of Cleaner Production, 141*, 1040–1050. https://doi.org/10.1016/j.jclepro.2016.09.177.

Hannachi, A., Jolliffe, I. T., Stephenson, D. B., & Trendafilov, N. (2006). In search of simple structures in climate: Simplifying EOFS. *International Journal of Climatology, 26*(1), 7–28. https://doi.org/10.1002/joc.1243.

Hirt, S., & Stanilov, K. (2009). *Twenty years of transition: The evolution of urban planning in Eastern Europe and the Former Soviet Union 1989–2009. Human Settlements Global Dialogue Series, No. 5.* Retrieved from http://unhabitat.org/books/twenty-years-of-transition-the-evolution-of-urban-planning-in-eastern-europe-and-the-former-soviet-union-1989-2009/.

Hughes, S., Chu, E. K., & Mason, S. G. (2018). Introduction. In S. Hughes, E. K. Chu, & S. G. Mason (Eds.). *Climate change in cities: Innovations in multi-level governance* (pp. 1–15). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-65003-6_1.

Humphreys, L. G., & Montanelli, R. (1975). An investigation of the parallel analysis criteria for determining the number of common factors. *Multivariate Behavioral Research, 10*(1), 193–205.

Husson, F., Lê, S., & Pagès, J. (2010). *Exploratory multivariate analysis by example using R. Chapman & Hall/CRC computer science & data analysis, Vol. 40*https://doi.org/10.1080/02664763.2012.657409.

Hyvärinen, A. (2013). Independent component analysis of images. In D. Jaeger, & R. Jung (Eds.). *Encyclopedia of computational neuroscience* (pp. 1–5). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-7320-6_708-1.

Ian, H., Kaliopa, D.-A., & Natasa, P.-M. (2003). *Transformation of cities in central and Eastern Europe: Towards globalization* (1st ed.). Tokyo: United Nations University Press

Retrieved from https://chisineu.files.wordpress.com/2012/08/biblioteca_hamilton_transformation.pdf.

Ichikawa, H., Yamato, N., & Dustan, P. (2017). Competitiveness of global cities from the perspective of the global power city index. *Procedia Engineering, 198*(September 2016), 736–742. https://doi.org/10.1016/j.proeng.2017.07.125.

Isnard, M., & Sautory, O. (1994). Les macros SAS d' analyse des données. *Institut National de la Statistique et des études Économiques*.

ISO (2014). *ISO-37120:2014 sustainable development of communities — Indicators for city services and quality of life, Geneva* Retrieved from https://www.iso.org/standard/61885.html.

Jacinto, C., & Soares, C. G. (2008). The added value of the new ESAW/Eurostat variables in accident analysis in the mining and quarrying industry. *Journal of Safety Research, 39*(6), 631–644. https://doi.org/10.1016/j.jsr.;1; 2008.10.009.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202.

Kahn, M. E. (2006). *Green cities: Urban growth and the environment.* https://doi.org/10.1111/j.1467-9787.2006.00531_8.x.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141–151. https://doi.org/10.1177/001316446002000116.

Koch, I. (2013). *Analysis of multivariate and high-dimensional data.* Cambridge University Press https://doi.org/10.1017/CBO9781139025805.

Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). Data preprocessing for supervised learning. *International Journal of Computer Science, 1*(2), 111–117. https://doi.org/10.1080/02331931003692557.

Kouser, K., Lavanya, P. G., Rangarajan, L., & Acharya Kshitish, K. (2016). Effective feature selection for classification of promoter sequences. *PLOS ONE, 11*(12), 1–20. https://doi.org/10.1371/journal.pone.0167165.

KPMG (2010). City typology as the basis for policy. *City*.

Kucharczyk, H., Kucharczyk, M., Stanislawek, K., & Fedor, P. (2012). *Application of PCA in taxonomy research. Principal component analysis – Multidisciplinary applications*https://doi.org/10.5772/711.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software, 25*(1), 1–18. https://doi.org/10.1016/j.envint.2008.06.007.

Li, X. A., & Yeh, G. O. (1998). Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta. *International Journal of Remote Sensing, 19*(8), 1501–1518. https://doi.org/10.1080/014311698215315.

Liao, T. L., Huang, C. J., & Wu, C. Y. (2012). *Chapter 5 – Empirical Study: Do fund managers herd to counter investor sentiment? Principal component analysis – Multidisciplinary applications (2000)* Retrieved from http://cdn.intechopen.com/pdfs/30006/InTech-Empirical_study_do_fund_managers_herd_to_counter_investor_sentiment_.pdf.

Lindström, B., & Eriksson, B. (1993). Quality of life among children in the Nordic countries. *Quality of Life Research, 2*(1), 23–32. https://doi.org/10.1007/BF00642886.

Marsal-Llacuna, M. L., Colomer-Llinàs, J., & Meléndez-Frigola, J. (2015). Lessons in urban monitoring taken from sustainable and livable cities to better address the Smart Cities initiative. *Technological Forecasting and Social Change, 90*(PB), 611–622. https://doi.org/10.1016/j.techfore.2014.01.012.

Mayer, A. L. (2008). Strengths and weaknesses of common sustainability indices for multidimensional systems. *Environment International, 34*(2), 277–291. https://doi.org/10.1016/j.envint.2007.09.004.

McDonnell, M. J., & MacGregor-Fors, I. (2016). The ecological future of cities. *Science, 352*(6288), 936–938. https://doi.org/10.1126/science.aaf3630.

McManus, P. (2012). Measuring urban sustainability: The potential and pitfalls of city rankings. *Australian Geographer, 43*(4), 411–424. https://doi.org/10.1080/00049182.2012.731301.

Meijering, J. V., Kern, K., & Tobi, H. (2014). Identifying the methodological characteristics of European green city rankings. *Ecological Indicators, 43*, 132–142. https://doi.org/10.1016/j.ecolind.2014.02.026.

Mercer (2018). *Quality of living rankings*.

Metsalu, T., & Vilo, J. (2015). ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Research, 43*(W1), W566–W570. https://doi.org/10.1093/nar/gkv468.

Mori, K., & Christodoulou, A. (2012). Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review, 32*(1), 94–106. https://doi.org/10.1016/j.eiar.2011.06.001.

Mori Memorial Foundation (2016). Global Power City Index 2016. *Institute for Urban Strategies, 74*(4), A28–A29. https://doi.org/10.1002/ana.24042.

Morse, S., & Fraser, E. D. G. (2005). Making 'dirty' nations look clean? The nation state and the problem of selecting and weighting indices as tools for measuring progress towards sustainability. *Geoforum, 36*(5), 625–640. https://doi.org/10.1016/j.geoforum.2004.10.005.

Munda, G. (2004). Social multi-criteria evaluation: Methodological foundations and operational consequences. *European Journal of Operational Research, 158*(3), 662–677. https://doi.org/10.1016/S0377-2217(03)00369-2.

Munda, G. (2006). Social multi-criteria evaluation for urban sustainability policies. *Land Use Policy, 23*(1), 86–94. https://doi.org/10.1016/j.landusepol.2004.08.012.

Nam, T., & Pardo, T. A. (2011). *Conceptualizing smart city with dimensions of technology, people, and institutions. Proceedings of the 12th annual international digital government research conference: Digital government innovation in challenging times.* New York, NY: ACM282–291. https://doi.org/10.1145/2037556.2037602.

Numbeo (2016). *Europe: Quality of Life Index 2015.* Retrieved from https://www.numbeo.com/quality-of-life/region_rankings.jsp?.title=2015&region=150 [28.07.18].

OECD (2008). *Handbook on constructing composite indicators: Methodology and user guide.* Paris: European Commission Joint Research Centerhttps://doi.org/10.1787/9789264043466-en.

Omucheni, D. L., Kaduki, K. A., Bulimo, W. D., & Angeyo, H. K. (2014). Application of principal component analysis to multispectral-multimodal optical image analysis for malaria diagnostics. *Malaria Journal, 13*(485), 1–11. https://doi.org/10.1186/1475-2875-13-485.

Organisation for Economic Co-Operation and Development, & China Development Research Foundation (2010). *Trends in urbanisation and urban policies in OECD Countries: What lessons for China?*219. https://doi.org/10.1787/9789264092259-en.

Osbaldiston, R., & Schott, J. P. (2012). Environmental sustainability and behavioral science: Meta-analysis of proenvironmental behavior experiments. *Environment and Behavior, 44*(2), 257–299. https://doi.org/10.1177/0013916511402673.

Phillis, Y. A., Kouikoglou, V. S., & Verdugo, C. (2017). Urban sustainability assessment and ranking of cities. *Computers, Environment and Urban Systems, 64*, 254–265. https://doi.org/10.1016/j.compenvurbsys.2017.03.002.

Poortinga, W., Steg, L., & Vlek, C. (2004). Values, environmental concern, and environmental behavior: A study into household energy use. *Environment and Behavior, 36*(1), 70–93. https://doi.org/10.1177/0013916503251466.

PwC (2016). *Cities of opportunity*.

Reh, W., Fellermann, A., & Duprez, L. (2013). *Soot-free cities: A European City Ranking, Brussels.* Retrieved from http://www.sootfreecities.eu/sootfreecities.eu/public/download/city-ranking-background.pdf.

Roaf, J., Atoyan, R., Joshi, B., & Krogulski, K. (2014). *25 years of transition: Post-communist Europe and the IMF.* Retrieved from https://www.imf.org/external/pubs/ft/reo/2014/eur/eng/pdf/erei_sr_102414.pdf.

Saisana, M., & Saltelli, A. (2011). Rankings and ratings: Instructions for use. *Hague Journal on the Rule of Law, 3*(2), 247–268. https://doi.org/10.1017/S1876404511200058.

Scariano, S. M. (2013). *An Introductory Application of Principal Components to Cricket Data, 21*(3), 83–103.

Serbanica, C., & Constantin, D. L. (2017). Sustainable cities in central and eastern European countries. Moving towards smart specialization. *Habitat International, 68*, 55–63. https://doi.org/10.1016/j.habitatint.2017.03.005.

Sharholy, M., Ahmad, K., Mahmood, G., & Trivedi, R. C. (2008). Municipal solid waste management in Indian cities – A review. *Waste Management, 28*(2), 459–467. https://doi.org/10.1016/j.wasman.2007.02.008.

Sheng, N., & Tang, U. W. (2016). The first official city ranking by air quality in China – A review and analysis. *Cities, 51*, 139–149. https://doi.org/10.1016/j.cities.2015.08.012.

Siemens (2009). *European Green City Index: Assessing the environmental impact of Europe's major cities.* Munich: Siemens AG Retrieved from http://scholar.google.com/scholar?.hl=en&btnG=Search&q=intitle:European+Green+City+Index#0.

Szopik-Depczyńska, K., Cheba, K., Bąk, I., Stajniak, M., Simboli, A., & Ioppolo, G. (2018). The study of relationship in a hierarchical structure of EU sustainable development indicators. *Ecological Indicators, 90*(December 2017), 120–131. https://doi.org/10.1016/j.ecolind.2018.03.002.

Takano, T., Nakamura, K., & Watanabe, M. (2002). Urban residential environments and senior citizens' longevity in megacity areas: The importance of walkable green spaces. *Journal of Epidemiology and Community Health, 56*(12), 913–918 Retrieved from http://jech.bmj.com/content/56/12/913.Abstract.

The Economist Intelligence Unit (2016). Best cities ranking and report: A special report from the Economist Intelligence Unit. *The Economist*.

Tilbury, D. (1995). Environmental Education for Sustainability: Defining the new focus of environmental education in the 1990s. *Environmental Education Research, 1*(2), 195–212. https://doi.org/10.1080/1350462950010206.

Tu, K.-J., & Lin, L.-T. (2008). Evaluative structure of perceived residential environment quality in high-density and mixed-use urban settings: An exploratory study on Taipei City. *Landscape and Urban Planning, 87*(3), 157–171. https://doi.org/10.1016/j.landurbplan.2008.05.009.

UN-HABITAT (2015). The city prosperity initiative: 2015 global city report. *Nairobi*.

UNECE (2015). *Key performance indicators for smart sustainable cities to assess the achievement of sustainable development goals, 1603* ITU-T L.1603.

United Nations (2014). *World urbanization prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352). New York.* United States of Americahttps://doi.org/10.4054/DemRes.2005.12.9.

van Leeuwen, C. J., Frijns, J., van Wezel, A., & van de Ven, F. H. M. (2012). City Blueprints: 24 indicators to assess the sustainability of the urban water cycle. *Water Resources Management, 26*(8), 2177–2197. https://doi.org/10.1007/s11269-012-0009-1.

Vidal, R., Ma, Y., & Sastry, S. S. (2016). *Principal component analysis. Generalized principal component analysis.* New York, NY: Springer25–62. https://doi.org/10.1007/978-0-387-87811-9_2.

Vinod Kumar, T. M., & Dahiya, B. (2017). Smart Economy in smart cities. In T. M. Vinod Kumar (Ed.). *Smart economy in smart cities. International Collaborative Research: Ottawa, St.Louis, Stuttgart, Bologna, Cape Town, Nairobi, Dakar, Lagos, New Delhi, Varanasi, Vijayawada, Kozhikode, Hong Kong* (pp. 3–76). Singapore: Springer. https://doi.org/10.1007/978-981-10-1610-3_1.

Webb, J., Hawkey, D., & Tingey, M. (2016). Governing cities for sustainable energy: The UK case. *Cities, 54*, 28–35. https://doi.org/10.1016/j.cities.2015.10.014.

Wei, Y., Huang, C., Li, J., & Xie, L. (2016). An evaluation model for urban carrying capacity: A case study of China's mega-cities. *Habitat International, 53*, 87–96. https://doi.org/10.1016/j.habitatint.2015.10.025.

Weng, Q., Liu, H., Liang, B., & Lu, D. (2008). The spatial variations of urban land surface temperatures: Pertinent factors, zoning effect, and seasonal variability. *IEEE Journal*

of Selected Topics in Applied Earth Observations and Remote Sensing, 1(2), https://doi.org/10.1109/JSTARS.2008.917869.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques.* Elsevier Science Retrieved from https://books.google.pt/books?.id=1SylCgAAQBAJ.

Wong, C. (2002). Developing indicators to inform local economic development in England. *Urban Studies, 39*(10), 1833–1863. https://doi.org/10.1080/0042098022000002984.

Xiao, S., Lu, Z., & Xu, L. (2017). Multivariate sensitivity analysis based on the direction of eigen space through principal component analysis. *Reliability Engineering and System Safety, 165*(March), 1–10. https://doi.org/10.1016/j.ress.2017.03.011.

Zhou, Z., Zhang, S., Wang, C., Zuo, J., He, Q., & Rameezdeen, R. (2016). Achieving energy efficient buildings via retrofitting of existing buildings: A case study. *Journal of Cleaner Production, 112*(Part 5), 3605–3615. https://doi.org/10.1016/j.jclepro.2015.09.046.