SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 218

A TAXONOMIC ANALYSIS OF SELECTED REGIONS OF EUROPEAN COUNTRIES IN TERMS OF IMPLEMENTATION OF SMART CITY COMPONENTS – A DYNAMIC APPROACH

Anna JANIGA-ĆMIEL

University of Economics in Katowice; anna.janiga-cmiel@ue.katowice.pl, ORCID: 0000-0002-7668-9225

Purpose: The aim of the study is to: apply synthetic measures as a tool for quantifying the development level of selected Smart City components in European countries and determine the development level of selected Smart City components in European countries by means of the WAP (multivariate comparative analysis) tools.
Design/methodology/approach: The study examines the use of the Internet and information and communication technologies selected regions of European countries, with particular attention devoted to the period of the Covid-19 pandemic. Data were drawn from Statistics Poland and Eurostat, taking into account the thematic area of the study and data availability. The first stage of the analysis included the construction of a synthetic variable. The synthetic variable made it possible to compare the level of development of the phenomenon in selected regions in 2020, 2021, 2022, 2023 and 2024. Then, a distance matrix was determined. Taking into account the distance matrices from each year, a matrix of intensity indicators was determined for the entire period.
Findings: The study examines the use of the Internet and information and communication technologies in the regions of selected European countries. Taxonomic analysis is an effective tool for assessing the influence of factors on selected characteristics of the development of the studied phenomenon.
Research limitations/implications: The inability to create a set of comparable statistical data over many years is the main limitation of the analysis.
Originality/value: The presented taxonomic analysis of the phenomenon under consideration can be used to compare different objects (countries, provinces, municipalities) or to carry out comparative analysis of other aspects of the issue, and the results of these studies will contribute to further research in this area.
Keywords: taxonomic methods, Internet and ICT usage, COVID-19.
Category of the paper: Research paper.

1. Introduction

Modern times are characterized by extensive usage of internet of things, artificial intelligence, machine learning, big data, etc. as well as efforts to identify the needs of urban dwellers with the aim of improving their quality of life. The proper implementation of the smart city concept requires finding optimal development solutions. Cities undergo transformation due to economic and social globalization, metropolization and technological advancements. Modern cities have to continuously adjust their transformation to the ever-changing knowledge economy and knowledge society and the development of new technologies, which, in turn, are affected by the ongoing demographic, social and cultural changes, the emergence of new services, etc (Makowski, Kidyba, 2018). These processes have an impact on the development and transformation of cities, whose purpose is to fulfill complex socio-economic needs. Another significant factor driving urban transformation was the Covid-19 pandemic, which brought about social, economic and health crises. It inspired creativity and innovation in cities, contributed to the implementation of the smart city concept and the promotion of sustainable development solutions. The pandemic showed the necessity for safer digital ICT and IT solutions. It also accelerated and intensified the use of artificial intelligence in city management and e-commerce.

Smart city is one of the most common current buzz phrases. It is usually associated with the development and application of modern technologies. Big data is used on a large scale, which helps cities to develop. Technologies make cities safer, streamline their management, increase mobility, and have a positive impact on public health. On the other hand, a smart city can be a hindrance, e.g. when it lacks proper strategic management.

To sum up, the constant development of urban areas and the improvement of the quality of life of their inhabitants pose one of the major global challenges to urban policies.

The aim of the study is to:

- apply synthetic measures as a tool for quantifying the development level of selected Smart City components in the regions of European countries,
- determine the development level of selected Smart City components in the regions of European countries by means of the WAP (multivariate comparative analysis) tools.

2. The set of diagnostic characteristics of the problems under study

The Smart City issue covers many different thematic areas. Due to the lack of statistical data, it was necessary to limit the thematic scope of the data set. The analysis presented is the first stage of research taking into account data from the thematic area of the use and implementation of modern technologies and the Internet. Modern technologies help to integrate all systems within an urbanized area and improve communication (Wi-Fi routers for residents). Due to the possibility of collecting data for the analysis, regions were included for the countries compared.

The basis for the analysis was the final set of diagnostic variables in the years 2024, 2023, 2022, 2021 and 2020 with particular emphasis on the Covid-19 pandemic. In summary, the analysis transferred the smart city issue in the selected topic to the regions due to the possibility of collecting data.

Data were drawn from Statistics Poland and Eurostat. The dataset includes the scope of the study and the availability of data. The diagnostic variables adopted in the present study are measurable and best describe the development of the analysed phenomenon. Based on the calculated values of the coefficients of variation and the results of verifying correlation analysis conducted by means of an inverted correlation matrix, the final set of diagnostic characteristics which describes the phenomenon was obtained (Młodak, 2006; Chomątowski, Sokołowski, 1978; Pociecha et al., 1988; Panek, 2009; Zeliaś, 2004; Strahl, 1990; Malina, 2008; Mika, 1995) (in the description of the data are used "S" - stimulant):

x_1 – Number of households with access to the Internet at home – [%] (S),

x_2 – Number of individuals who ordered goods or services over the internet for private use – [%] (S),

x_3 – Number of Individuals who used the internet, frequency of use and activities [%] (S).

3. Measure of development

We consider the set of stimulants Q at the time t = 1, …, T. The development of the phenomenon is described in the form of a block matrix (Strahl, 1990):

$$\left[y_{kj}^{1t}; y_{kl}^{2t}\right]_{(m+n)xTK} \tag{1}$$

where:

t = 1, …, T, k = 1, …, K, j = 1, …, m, i = 1, …, n,

y_{kj}^{1t} – the value of j-this variable,

y_{kl}^{2t}– the value of l-this variable.

The normalization is determined according to the formula (Strahl, 1990):

$$y_{kj}'^{t} = \frac{y_{kj}^{t}}{S_j} \tag{2}$$

where

$$S_j = \sqrt{\frac{1}{KT}\sum_{k=1}^{K}\sum_{t=1}^{T}\left(y_{kj}^{t} - \bar{y}\right)^2} \tag{3}$$

The absolute measure at the moment t proposed by M. Cieślak is defined according to the formula (Strahl, 1990):

$$m_k^t = \sum_{j=1}^{m} y_{kj}^{\prime t} \tag{4}$$

While the measure in all the study periods is determined by the formula:

$$m_k = \sum_{t=1}^{T} m_k^t. \tag{5}$$

4. Determination of a synthetic variable - an empirical example

After normalizing the variables, we synthesize each of the selected groups of measures and calculate the synthetic variable. The analysis covered 177 regions for selected European countries (in the analysis for selected countries, smaller territorial units were also left, e.g. Spain), the time interval was 5 years (2024, 2023, 2022, 2021 and 2020), and the number of variables was 3 (k = 3, variables listed in the previous chapter and two research topics). Table 1 presents the calculated values of the synthetic variable for the selected regions in the analyzed years.

Table 1.
Determined values of the synthetic variable - 2020, 2021, 2022, 2023, 2024

	2020		2021	2022	2023	2024
	Region designation	Region				
1	NO06	Trøndelag	NL34	NO02	NL31	NL11
2	NO07	Nord-Norge	NO06	NO07	NL21	NL34
3	NL23	Flevoland	NO02	NO0	NL32	NL33
4	FI1B	Helsinki-Uusimaa	IE04	ES64	NO07	NL23
5	DK01	Hovedstaden	NO07	NL13	NO02	NL3
6	NL21	Overijssel	NO0	NL32	NL3	NL31
7	NL2	Oost-Nederland	NL31	NL34	NL23	NL32
8	DK04	Midtjylland	FI1B	NO06	NL34	NL41
9	DK03	Syddanmark	NL21	NL31	NL42	NL2
10	NL31	Utrecht	DK01	FI1B	NL4	NL22
11	NL22	Gelderland	DK04	NL11	NL41	NL4
12	NL32	Noord-Holland	IE06	NL41	NL2	NL21
13	NO0	Norge	NL2	NL21	NO0	NL1
14	NL13	Drenthe	NL3	DK01	NL13	NL42
15	DEE	Sachsen-Anhalt	NL32	NL3	NL33	NO07
16	SE1	Östra Sverige	NL23	NL4	NL11	NL12
17	NL12	Friesland (NL)	NL22	NL22	NL1	NL13
18	NL1	Noord-Nederland	NL13	NL2	NO06	NO02
160	BG4	Yugozapadna i Yuzhna tsentralna Bulgaria	RO2	RO4	HU32	PT3
161	PT11	Norte	BG41	RO41	ITF2	PT30
162	RO31	Sud-Muntenia	ITF5	RO22	BG4	BG4
163	ITF3	Campania	RO31	RO2	PT1	BG32
164	RO21	Nord-Est	RO41	ITF3	RO31	RO22
165	RO2	Macroregiunea Doi	PT11	RO31	RO22	RO41
166	ITF2	Molise	ITF	PT11	ITF4	ITF4

Cont. table 1.

167	ITF	Sud	ITF2	RO21	PT11	ITF5
168	RO22	Sud-Est	BG33	BG33	ITF	PT11
169	ITF4	Puglia	RO22	ITF5	ITF3	BG33
170	ITF5	Basilicata	BG4	ITF	BG42	BG3
171	BG33	Severoiztochen	ITF6	ITF4	BG33	ITF2
172	ITF6	Calabria	ITF4	BG42	BG34	ITF
173	BG42	Yuzhen tsentralen	BG32	BG32	BG3	BG42
174	BG34	Yugoiztochen	BG42	BG3	BG32	ITF3
175	BG3	Severna i Yugoiztochna Bulgaria	BG3	BG34	ITF5	BG31
176	BG32	Severen tsentralen	BG34	BG31	BG31	BG34
177	BG31	Severozapaden	BG31	ITF6	ITF6	ITF6

Source: based on own research (https://ec.europa.eu/eurostat

Due to the large amount of data, the table presents the first 16 highest values of the synthetic variable and the last 16 lowest values.

In 2020, the first two places were taken by regions from Norway, followed by regions from the Netherlands, Finland and Denmark. The last five places are occupied by regions from Bulgaria. In 2021, the first place was taken by a region from the Netherlands, then two regions from Norway. The last places were the same as in 2020.

In 2022, the top three places were taken by regions from Norway. The last places were taken by the Italian region, and the remaining four places were again taken by regions from Bulgaria.

In 2023, the lead of the Dutch regions increased - the first three places. The positions of the last places remained unchanged - also in 2024.

In 2024, the first ten places were occupied by regions from the Netherlands.

5. Analysis of the intensity of distance changes over time

We construct the matrix of observations of diagnostic variables in year t and determine the standardization of variables. We determine the distances between objects in year t according to the formula (Nowak, 1990):

$$\boldsymbol{d_{ij}^t = \frac{1}{K}\sum_{k=1}^{K}\left|z_{ik}^t - z_{jk}^t\right|} \tag{6}$$

We determine the distance matrix T:

$$\boldsymbol{D^t = \begin{bmatrix} 0 & d_{12}^t & \cdots & d_{1N}^t \\ d_{21}^t & 0 & & d_{2N}^t \\ \vdots & & & \vdots \\ d_{N1}^t & \ldots & \ddots & 0 \end{bmatrix}} \tag{7}$$

To determine the intensity of changes in the level of distance between objects in the period [1,T], we use an individual indicator of the intensity of changes in distance (Nowak, 1990):

$$\boldsymbol{v_{ij} = \frac{1}{T-1}\sum_{t=2}^{T}\left|d_{ij}^t - z_{ij}^{t-1}\right|} \tag{8}$$

We calculate the assessment of the strength of changes in time in the level of distance between all objects in the period [1,T] (Nowak, 1990):

$$\boldsymbol{v} = \frac{\mathbf{2}}{\boldsymbol{N(N-1)}} \sum_{\boldsymbol{i=2}}^{\boldsymbol{N}} \sum_{\boldsymbol{j>i}} \boldsymbol{v_{ij}} \tag{9}$$

The lower the value of the v indicator, the more stable the distances between all objects in time.

6. Analysis of the intensity of distance changes - an empirical example

The analysis included as objects selected regions of European countries in 2020, 2021, 2022, 2023 and 2024 (k = 177). The diagnostic variables were presented in the previous chapter. Distance matrices were determined for each year. Due to the large dimension of the matrix (177 rows and 177 columns), it was not presented. The determined distance matrices made it possible to determine the matrix of individual change intensity indicators.

On the basis of the determined matrix of distance intensity change indicators, we can examine fluctuations in the development of the phenomenon during the analyzed period. Taking into account the values of the elements of the obtained matrix, we can conclude that the greatest fluctuations in the years 2020-2024 were the distances between the Italian region (ITF6) and the Belgian region (BE2) and between the Norwegian region (NO06) and the Belgian region (BG31), and then the German region (DED) and the Belgian region (BE1).

The most stable over time are the distances between the Norwegian region (NO07) and the Dutch region (NL2) and the Norwegian region (NO06) and the Dutch region (NL11). The results for the Polish region (PL9) were also examined. The greatest fluctuations were observed in this region with the Belgian regions (BG31, BG32, BG34), and Hungary (HU12). The distances between the Polish region and the Austrian region are the most stable. (AT34, AT33, AT31).

The value of the integral indicator determined over time is: v = 0.198.

This value means that there are significant fluctuations in the level of distance for all the regions analyzed. The reason may be the different development of the phenomenon in the analyzed regions (Nowak, 1990).

7. Conclusion

The aim of the paper was to:

- apply synthetic measures as a tool for quantifying the development level of selected Smart City components in the regions of European countries,
- determine the development level of selected Smart City components in the regions of European countries by means of the WAP (multivariate comparative analysis) tools.

The study covered the years 2024, 2023, 2022, 2021 and 2020, including the period of the Covid-19 pandemic. The first step in the study was to define a synthetic variable.

The synthetic variable can present a complex phenomenon by means of one variable. The synthetic variable represented the level of development of the analyzed phenomenon in the regions of the European countries in the selected years. The values of the synthetic variable also made it possible to rank the regions from the most advanced to the least advanced in terms of the phenomenon studied. The distance matrix was then determined for the years 2024, 2023, 2022, 2021 and 2020. The distance matrices also contain information on the distances between each of the regions selected for analysis.

Taking into account the distance matrix, it was possible to determine the intensity indicator matrix for the entire time period considered. To sum up, the taxonomic analysis tools selected in the analysis have made it possible to study the phenomenon both in one year and for a total of 5 years. The results obtained showed the level of development of the analyzed phenomenon in the regions in one period and in total over the entire period considered. Additionally, the information obtained made it possible to compare which regions are characterized by similar development and which are significantly different. The issue of smart cities is an important one that concerns us all. The analysis included a thematic area related to the use of modern technologies and the Internet. The results showed, among other things, that the Polish region was subject to the smallest fluctuations in distance during the period analyzed compared to the regions of Austria (after all, Vienna has been called the smartest city in the world). The results obtained make it possible to compare regions in European countries in terms of the level of use of new technologies and Internet penetration. In different countries with diverse economies and different economic, socio-cultural, technical, and administrative-legal conditions, certain regions may achieve a similar level of the analyzed phenomenon. Additionally, with the possibility of collecting a larger amount of data, and thus expanding the analysis, one can examine and identify factors and conditions that influence the achievement of a certain level of the phenomenon—while ensuring that regions and cities do not lose the uniqueness that makes them distinctive.

References

1. Chomątowski, S., Sokołowski, A. (1978). Taksonomia struktur. *Przegląd Statystyczny*, *Vol. 25, Iss. 2*, pp. 217-226.
2. Grabiński, T., Wydymus, A., Zeliaś, A. (1989). *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*. Warszawa: PWN, pp. 25-49.
3. https://ec.europa.eu/eurostat
4. https://stat.gov.pl
5. Makowski, Ł., Kidyba, M. (2018). *Smart City. Innowacyjne rozwiązania w administracji publicznej a zarządzanie inteligentnym miastem*. Poznań: Wydawnictwo Wyższej Szkoły Bankowej w Poznaniu, pp. 9-44.
6. Malina, A. (2008). *Przestrzenno czasowa analiza rynku pracy w Polsce i krajach Unii Europejskiej*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, pp. 215-280.
7. Mika, J. (1995). *Analiza statystyczna Polski na tle krajów Unii Europejskiej*. Katowice: Śląsk, pp. 10-25.
8. Młodak, A. (2006). *Analiza taksonomiczna w statystyce regionalnej*. Warszawa: Difin, pp. 25-66.
9. Nowak, E. (1990). *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych.* Warszawa: PWE, pp. 150-200.
10. Panek, T. (2009). *Statystyczne metody wielowymiarowej analizy porównawczej*. Warszawa: Szkoła Główna Handlowa w Warszawie – Oficyna Wydawnicza, pp. 58-114.
11. Pociecha, J., Podolec B., Sokołowski, A., Zając, K. (1988). *Metody taksonomiczne w badaniach społeczno-ekonomicznych.* Warszawa: PWN, pp. 20-90.
12. Strahl, D. (1990). *Metody programowania rozwoju społecznego-gospodarczego*. Warszawa: PWE, pp. 35-69.
13. Zeliaś, A. (2006). *Poziom życia w Polsce i krajach Unii Europejskiej*. Warszawa: PWE, pp. 195-212.