Comparison of Incidences of Serious Diseases within Regions in the Czech Republic[1]

*Lucie KOPECKÁ – David ZAPLETAL – Viera PACÁKOVÁ**

Abstract

The incidence of serious diseases in most European countries has a growing trend compared to mortality due to these illness. There are significant differences in the health status of the population not only among countries but also within them. The goal of the health policy of the European Union and its member states is not only to reduce the incidence of serious diseases but also mitigate regional inequalities of their incidence. The object of this article is to compare incidences of serious diseases within the regions in the Czech Republic by using hybrid approach which combine multidimensional scaling with linear ordering of the objects. Hybrid approach is suitable for visualization of objects and determination of distances from ideal object according to indicators used. The factors of regional inequalities in the incidence of serious diseases are identified and quantified based on the results of hybrid method.

Keywords: *incidences, serious diseases, multidimensional scaling, linear ordering, hybrid approach, rank correlation, factor analysis*

JEL Classification: C38, C43, C63

Introduction

Health is central in people's lives and needs to support by effective policies and actions at EU level and in EU Member States. Health is important for the wellbeing of individuals and society, but a healthy population is also a prerequisite

* Lucie KOPECKÁ – David ZAPLETAL – Viera PACÁKOVÁ, University of Pardubice, Faculty of Economics and Administration, Institute of Mathematics and Quantitative Methods, Studentská 84, 532 10 Pardubice, Czech Republic; e-mail: lucie.kopecka1@student.upce.cz; david.zapletal@upce.cz; viera.pacakova@upce.cz

[1] This work was supported by SGS University of Pardubice, Faculty of Economics and Administration, No. SGS_2019_018.

for economic productivity and prosperity. EU Member States have the main responsibility for health policy and provision of healthcare to European citizens. Values relating to improving health must include reducing inequities in health (European Commission, 2007; Eurostat, 2017).

The main causes of death in EU countries are circulatory diseases and various types of cancer, followed by respiratory diseases and external causes of death. Diseases of the circulatory system were the most frequent causes of death also in the Czech Republic in 2015 with an age-standardised mortality rate 629 per 100 000 population, that is 65% higher than the EU average. Ischaemic heart diseases, which include heart attack and other diseases, and stroke are the most common causes of death from circulatory diseases. Deaths from coronary heart diseases have reduced considerably in Czech Republic since 1991, and are slowly closing the gap with the best performing countries in the EU. The drop is attributed to changes in therapy (43%) and risk factors (52%). Malignant neoplasms were the second most common cause of death in the Czech Republic in 2015 with an age-standardised mortality rate 279 per 100 000 population, which is about 1% less than the EU average. In 2018, 3 million new cases of cancer are expected to be diagnosed in the 28 EU member states, age-standardised incidence rate per 100 000 population in Czech Republic is estimated by 583 new cases (OECD/EU, 2018).

Mortality rates are declining despite the growing incidence of cancer due to increasing quality of medical care, due to improved organization of cancer treatment (e.g. the formation of comprehensive cancer in 2005), the availability of new diagnostic and therapeutic agents, earlier diagnosis of cancer and due to the aforementioned screening programs (ÚZIS ČR, 2016; ÚZIS ČR, 2018; Fall and Glocker, 2018). The trends of incidences and mortalities caused by serious cancers show in Figure 1 and Figure 2.

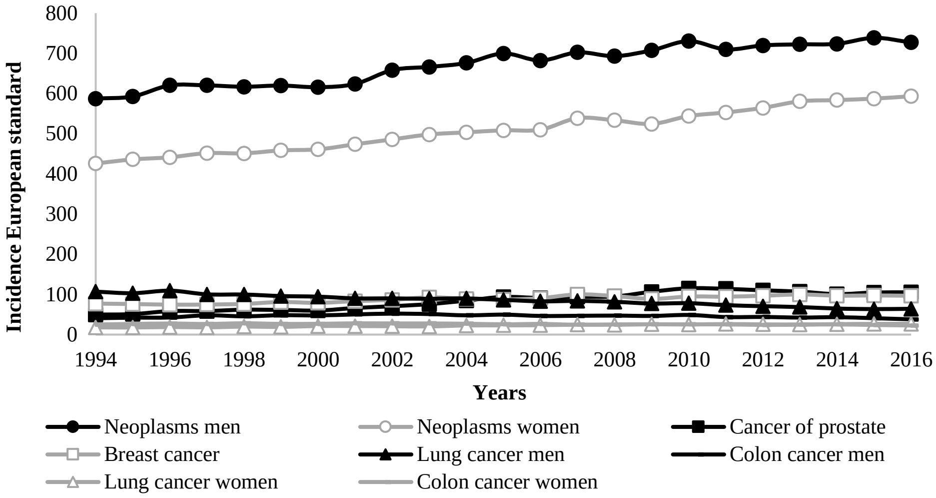
During the reporting period incidences of neoplasms for men increased by 0,975% per year on average and neoplasms for women by 1,517%. Mortality rates at the same diagnosis decreased by 2,017% per year on average in case of men and 1,530% in case of women.

Large inequalities exist in incidence and mortality due to serious disease across and within EU Member States. To reduce health inequalities is a fundamental objective for the EU policy and it can contribute to higher economic and social cohesion. Health in all policies requires health systems to build up multi-sectoral collaboration with other policy fields (Eurostat, 2017).

Policy aimed at reducing inequalities in health status and health care in each country can be effective only if these inequalities thoroughly examined. There are collected, regularly updated and on-line published a large number of databases and enormous number of indicators about health status, health care and health expenditures at regional, national, EU member countries, OECD countries and on the world level. These indicators provide very useful information on some of the weaknesses and strengths of each country's health care system and health of the inhabitants. Advanced statistical methods aimed at reducing the dimension and quantifying causal relationships can provide significant information for health policy aimed at reducing various inequalities in health.

Figure 1

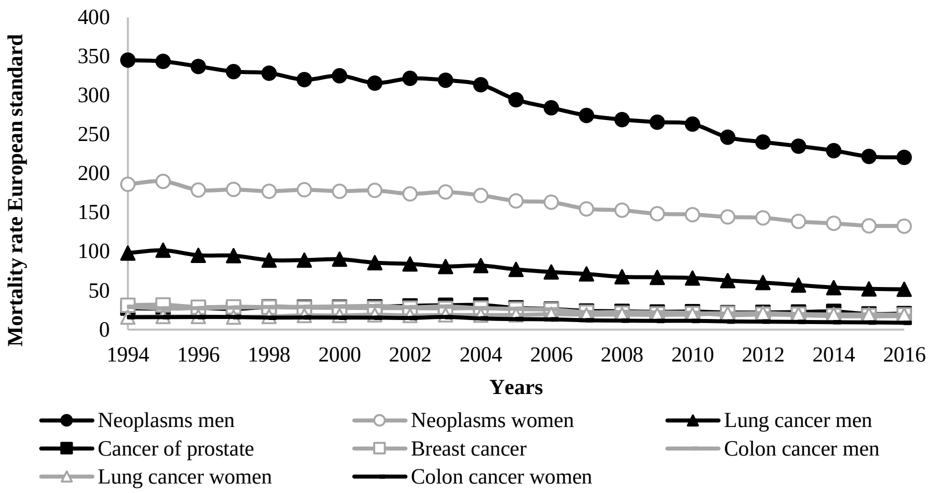
Trends of Incidences Caused by Cancers in the CR



Source: ÚZIS ČR, 2016.

Figure 2

Trends of Mortalities Caused by Cancers in the CR



Source: ÚZIS ČR, 2016.

1. Background of the Study

The main objective of the article is to compare incidences of serious diseases within the regions in the Czech Republic by using hybrid approach, which combines multidimensional scaling with linear ordering of data. The data were obtained from the Institute of Health Information and Statistics of the Czech Republic (ÚZIS ČR, 2016) and Regional statistics of the Czech Statistical Office (CZSO, 2015).

Several publications confirm the suitability of multivariate statistical methods such as correlation analysis, component analysis, factor analysis, cluster analysis or multidimensional comparative analysis for reducing the dimension of health indicators and assess of various causal relationships between them. For details see (Jindrová and Kopecká, 2017; Kopecká, 2018a; 2018b; Kopecká and Jindrová, 2017; Pacáková et al., 2016; Pacáková and Kopecká, 2018; Pacáková and Papoušková, 2016). There is also a lot of literature that deals with theory of multivariate statistical methods, like (Hair et al., 1992; Hebák et al., 2007; Stankovičová and Vojtková, 2007; Schiffman et al., 1981).

A key policy challenge in most EU countries is to improve outcomes of the health care system while containing cost pressures. Public spending on health care is one of the largest government spending items, representing on average 6% of GDP. Furthermore, health care costs are rising rapidly, driven by population ageing, rising relative prices and costly developments in medical technology.

Efficiency estimates allow the spotting of strengths and weaknesses for each country and identifying areas where achieving greater consistency in policy settings could yield efficiency gains. Efficiency measures are often used and mainly focus on hospital care (Pilyavskyy et al., 2006; Pilyavskyy and Staat, 2006; Hussey et al., 2009; Joumard, 2010, Pilyavskyy and Kopecká, 2018). Visualization of the objects by using hybrid approach which combines multidimensional scaling and linear ordering is described e.g. in articles (Walesiak, 2016; Walesiak and Dehnel, 2018). These approach can be used for comparing objects (regions) according to health indicators (incidences of serious diseases), ordering regions based on aggregate measure, distinguishing groups of regions sharing a similar or the same level of health and identifying regions with the similar level of health but with different location on isoquant of development, as describes Walesiak and Dehnel, 2018.

2. Data and Methodology

Data were obtained from database of ÚZIS ČR, 2016. This database provides data related to incidences caused by serious diseases, specifically cancers (C), infectious diseases (V), cardiovascular diseases (I), diabetes mellitus (DM) and asthma (A). Exactly thirteen variables were selected for our calculations (see Table 1) and fourteen territorial administration units, called regions, of the Czech Republic were evaluated by these health indicators. The regions are: South Bohemia (JHC), South Moravia (JHM), Karlovy Vary (KVK), Hradec Králové (HKK), Liberec (LBK), Moravia-Silesia (MSK), Olomouc (OLK), Pardubice (PAK), Plzeň (PLK), Prague (PHA), Central Bohemia (STC), Ústí nad Labem (ULK), Vysočina (VYS) and Zlín (ZLK). Incidences of serious diseases affecting Czech population are displayed in Table 1. These indicators related to incidences of serious diseases are calculated per 100 000 inhabitants for 2015 or the latest available year.

T a b l e 1

Selected 13 health indicators (incidences)

Variables	**Description of variables**
C1	colon and rectum cancer
C2	trachea, bronchus and lung cancer
C3	leukemia
C4	malignant neoplasms
V1	tick-borne encephalitis
V2	acute hepatitis A
V3	acute hepatitis B
V4	bacterial meningitis
V5	tuberculosis
I1	acute myocardial infarction
I2	stroke
DM1	diabetes mellitus
A1	asthma

Source: ÚZIS ČR, 2016.

Selected method, *hybrid approach* is procedure allowing the visualization of linear ordering results for the set of objects, as describes Walesiak (2016). It means visualisation where the *multidimensional scaling* and *linear ordering* of multidimensional objects are combined. Multidimensional Scaling (MDS) can be considered to be an alternative to factor analysis. In general, the goal of the analysis is to detect meaningful underlying dimensions that allow us to explain observed similarities or dissimilarities (distances) between the investigated objects. In factor analysis, the similarities between objects are displayed in the correlation matrix. The basis for MDS can be any kind of similarity or dissimilarity matrix, in addition to correlation matrices.

Suppose that a matrix of distances between objects is available (the regions of CR in our case). We then analyse this matrix, specifying that we want to reproduce the distances based on two dimensions. As a result of the MDS analysis, we would most likely obtain a two-dimensional representation of the locations of the regions. In general then, MDS attempts to arrange objects in a space with a particular number of dimensions (two-dimensional is the most common) so as to reproduce the observed distances. As a result, we can „explain“ the distances in terms of underlying dimensions. As in factor analysis, the actual orientation of axes in the final solution is arbitrary. We could rotate the map in any way we want, the distances between objects remain the same. Thus, the final orientation of axes in the plane or space is mostly the result of a subjective decision by the researcher, who will choose an orientation that can be most easily explained.

Multidimensional scaling is a way to „rearrange“ objects in an efficient manner, so as to arrive at a configuration that best approximates the observed distances. The procedure actually moves objects around in the space defined by the requested number of dimensions, and checks how well the distances between objects can be reproduced by the new configuration. In more technical terms, the procedure uses a function minimization algorithm that evaluates different configurations with the goal of maximizing the goodness-of-fit (or minimizing „lack of fit“). The most common measure that is used to evaluate how well (or poorly) a particular configuration reproduces the observed distance matrix is the stress measure.

The STATISTICA software package was used for calculations where the MDS procedure is an implementation of nonmetric multidimensional scaling. After determining the starting configuration STATISTICA will begin iterations under steepest descent. The goal of these iterations is to minimize the so-called *raw stress* (see, for example, Schiffman, 1981). The raw stress is defined as

$$S_0 = \sum_{i>j} \left(d_{ij} - f\left(\delta_{ij}\right) \right)^2 \tag{1}$$

In this formula, d_{ij} stands for the reproduced distances, given the respective number of dimensions, and δ_{ij} stands for the input data (i.e., observed distances). The expression $f\left(\delta_{ij}\right)$ indicates a nonmetric, monotone transformation of the observed input data (distances). In general, the nonmetric MDS attempt to minimize the differences between the reproduced distances and a monotone transformation of the input data, that is, the procedure attempt to reproduce the rank-ordering of the input distances or similarities (hence, also the name nonmetric multidimensional scaling).

The quality of MDS solution can also be asses by so-called *Shepard diagram*. It is a scatterplot of the reproduced distances for a particular number of dimensions against the observed input data (distances). This plot shows the reproduced distances plotted on the vertical (y) axis versus the original similarities plotted on the horizontal (x) axis. This plot also shows a step-function. This line represents the so-called D-hat (D^) values, that is, the result of the monotone transformation $f\left(\delta_{ij}\right)$ of the input data. If all reproduced distances fall onto the step-line, then the rank-ordering of distances (or similarities) would be perfectly reproduced by the respective solution (dimensional model). Deviations from the step-line indicate lack of fit.

Next, the results of multidimensional scaling (the coordinate system for the first and second dimension) are used for linear ordering of the multidimensional objects. The object are ordered according to *aggregate measure* d_i, which is given by

$$d_i = 1 - \sqrt{\sum\nolimits_{j=1}^{2}\left(v_{ij} - v_{+j}\right)^2} \,/\, \sqrt{\sum\nolimits_{j=1}^{2}\left(v_{+j} - v_{-j}\right)^2},\ i = 1,\ldots,n \qquad (2)$$

where

v_{ij} – j-th coordinate for the i-th object,

v_{+j} – j-th coordinate for the Pattern object,

v_{-j} – j-th coordinate for the Anti-pattern object.

The aggregate measure d_i lies between 0 and 1. The high values of aggregate measure indicate low level of incidences caused by serious diseases (Pattern object equals to 1) in contrast with low values of this measure (Anti-pattern object equals to 0). Pattern object can be explained as the best hypothetic object which is evaluated by desirable values of variables evaluating real objects (regions) in contrast with Anti-pattern object. For great details see (Walesiak, 2016; Walesiak and Dehnel, 2018).

3. Results an Discussion of Hybrid Methods

As mentioned above, hybrid approach which combine multidimensional scaling with linear ordering multidimensional object is used for comparing regions according to health indicators for instance incidences of serious diseases, which were selected (cardiovascular, oncological, infectious, diabetes and asthma). Before analysis, it is necessary to reveal associations among original variables which are presented in Table 1. Two possibilities of revealing associations among variables are used here (Spearman correlation coefficient and Kaiser-Meyer-Olkin index).

For details see Hebák et al., 2005; Hebák et al., 2007; Stankovičová and Vojtková, 2007. *Spearman correlation coefficients* for each pair of original variables are shown in Table 2. Correlation coefficients greater than 0.5 are shaded.

In addition, *Kaiser-Meyer-Olkin* (KMO) index was calculated. This index measures association among whole group of original variables. In this case, KMO index equals to 0.26. Both low value of KMO index and values of Spearman correlation coefficients (see Table 2) indicate poor correlations among original variables. KMO index was also calculated for each variable (see Table 3). Such index then measure the association of the individual variable with the rest of the variables.

T a b l e 2

Spearman Correlation Coefficients

Varia-bles	C1	C2	C3	C4	V1	V2	V3	V4	V5	I1	I2	DM1	A1
C1	1.000	–0.112	0.486	0.591	0.347	–0.236	0.029	0.445	–0.433	0.235	0.292	–0.055	–0.314
C2	–0.112	1.000	–0.543	0.037	–0.075	0.756	0.757	–0.290	0.053	0.270	0.011	0.576	0.108
C3	0.486	–0.543	1.000	0.486	–0.110	–0.297	–0.345	0.429	0.009	–0.169	–0.015	–0.046	–0.020
C4	0.591	0.037	0.486	1.000	0.066	0.207	0.073	0.334	–0.130	–0.068	–0.134	0.191	–0.393
V1	0.347	–0.075	–0.110	0.066	1.000	–0.164	0.016	–0.036	–0.257	0.024	–0.009	–0.451	–0.128
V2	–0.236	0.756	–0.297	0.207	–0.164	1.000	0.484	–0.367	0.315	–0.185	–0.319	0.438	0.090
V3	0.029	0.757	–0.345	0.073	0.016	0.484	1.000	0.149	0.185	0.263	–0.117	0.529	0.398
V4	0.445	–0.290	0.429	0.334	–0.036	–0.367	0.149	1.000	–0.164	0.066	0.168	0.376	0.002
V5	–0.433	0.053	0.009	–0.130	–0.257	0.315	0.185	–0.164	1.000	–0.680	–0.117	0.117	0.378
I1	0.235	0.270	–0.169	–0.068	0.024	–0.185	0.263	0.066	–0.680	1.000	0.310	–0.015	–0.125
I2	0.292	0.011	–0.015	–0.134	–0.009	–0.319	–0.117	0.168	–0.117	0.310	1.000	0.099	–0.547
DM1	–0.055	0.576	–0.046	0.191	–0.451	0.438	0.529	0.376	0.117	–0.015	0.099	1.000	0.002
A1	–0.314	0.108	–0.020	–0.393	–0.128	0.090	0.398	0.002	0.378	–0.125	–0.547	0.002	1.000

Source: Authors' calculations (ÚZIS ČR, 2016).

The results of correlation analysis (Table 2 and Table 3) point to impossibility of using component or factor analysis to investigate associations among original variables. It means that results of component or factor analysis would not be correct and useful.

T a b l e 3

Values of KMO index for 13 variables

Varia-bles	C1	C2	C3	C4	V1	V2	V3	V4	V5	I1	I2	DM1	A1
KMO	0.23	0.30	0.25	0.37	0.49	0.24	0.19	0.25	0.25	0.23	0.19	0.22	0.54

Source: Authors' calculations (ÚZIS ČR, 2016).

As mentioned above, multidimensional scaling is useful for finding a suitable coordinate system. It visualises multidimensional objects (regions) which are evaluated by a few variables (incidences of serious diseases) in two dimensions. A sufficient input for multidimensional scaling is distance matrix. In Table 4, the distance matrix of Euclidean distances is shown. The table contain fourteen original regions of the Czech Republic mentioned above and two *„artificial, hypothetic"* regions, namely *Pattern* (P) and *Anti-pattern* (AP). Pattern object has been created by the minimum of all original variables and Anti-pattern object has been constructed by the maximum. The reason for determination of Pattern and Anti-pattern objects in this manner is that the indicators of incidences of serious diseases are considered as destimulants. It means that low values of variables are desirable.

Table 4

Euclidean Distance (input matrix)

Reg.	**JHC**	**JHM**	**KVK**	**HKK**	**LBK**	**MSK**	**OLK**	**PAK**	**PLK**	**PHA**	**STC**	**ULK**	**VYS**	**ZLK**	**P**	**AP**
JHC	0.0	3.4	4.4	4.7	4.5	4.4	2.8	3.6	4.1	6.7	4.9	5.2	2.5	3.8	5.4	7.9
JHM	3.4	0.0	4.4	3.0	4.1	2.8	2.7	3.2	4.0	5.5	4.3	4.7	3.1	2.2	4.9	7.8
KVK	4.4	4.4	0.0	4.3	4.7	4.6	4.1	4.4	4.0	5.4	4.1	3.7	4.8	5.3	6.2	6.8
HKK	4.7	3.0	4.3	0.0	4.4	3.8	3.3	3.6	3.5	4.8	4.5	4.9	4.2	3.4	6.0	7.0
LBK	4.5	4.1	4.7	4.4	0.0	4.2	3.6	3.7	5.2	7.1	4.2	3.9	4.7	4.6	5.8	7.9
MSK	4.4	2.8	4.6	3.8	4.2	0.0	3.1	4.3	4.1	6.0	5.1	4.5	4.0	2.9	6.7	6.1
OLK	2.8	2.7	4.1	3.3	3.6	3.1	0.0	2.3	3.7	5.8	4.1	4.6	2.8	2.5	5.4	7.2
PAK	3.6	3.2	4.4	3.6	3.7	4.3	2.3	0.0	5.1	5.3	2.5	4.8	2.7	2.9	3.9	8.5
PLK	4.1	4.0	4.0	3.5	5.2	4.1	3.7	5.1	0.0	5.7	5.8	4.5	5.0	4.4	7.3	5.7
PHA	6.7	5.5	5.4	4.8	7.1	6.0	5.8	5.3	5.7	0.0	4.8	5.8	6.0	5.7	6.9	7.3
STC	4.9	4.3	4.1	4.5	4.2	5.1	4.1	2.5	5.8	4.8	0.0	4.0	4.1	4.5	3.7	8.8
ULK	5.2	4.7	3.7	4.9	3.9	4.5	4.6	4.8	4.5	5.8	4.0	0.0	5.6	5.4	6.3	7.0
VYS	2.5	3.1	4.8	4.2	4.7	4.0	2.8	2.7	5.0	6.0	4.1	5.6	0.0	2.9	5.0	8.0
ZLK	3.8	2.2	5.3	3.4	4.6	2.9	2.5	2.9	4.4	5.7	4.5	5.4	2.9	0.0	5.0	8.0
P	5.4	4.9	6.2	6.0	5.8	6.7	5.4	3.9	7.3	6.9	3.7	6.3	5.0	5.0	0.0	11.6
AP	7.9	7.8	6.8	7.0	7.9	6.1	7.2	8.5	5.7	7.3	8.8	7.0	8.0	8.0	11.6	0.0

Source: Authors' calculations (ÚZIS ČR, 2016).

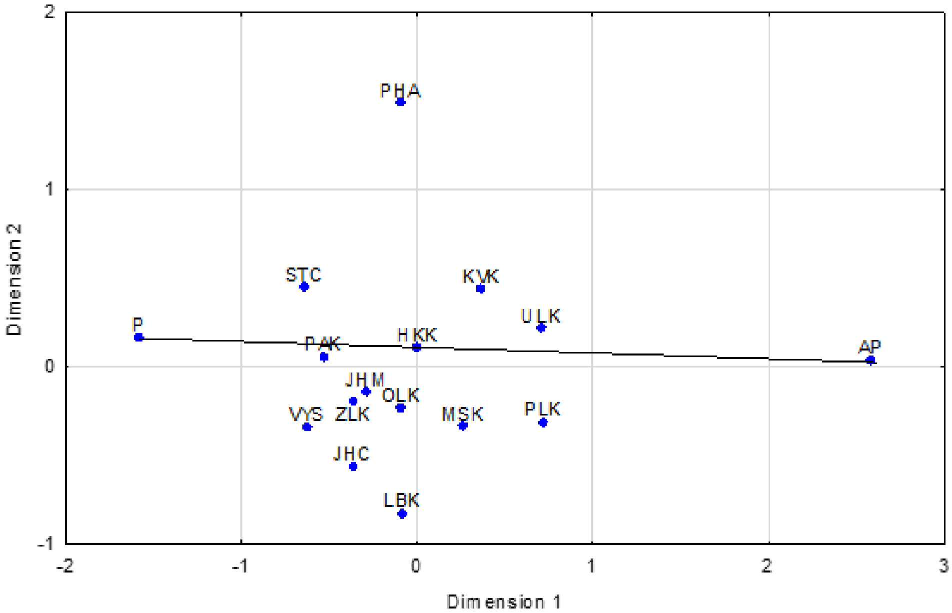
Two dimensions were chosen for visualisation of regions by employing multidimensional scaling. The original results of MDS for 16 objects are displayed in Figure 3. The largest distance is of course between hypothetic *„best"* Pattern and hypothetic *„worst"* Anti-pattern object which are connected by straight line (so-called axis of the set). The real regions of the Czech Republic occur between these two artificial regions. Most regions create one large group except of Prague. Prague as the capital city of the CR represents an outlier. Capital cities are often outliers in case of many indicators such as socio-economic, demographic or health and so on. It means that Prague is different in comparison with other regions. It has a wide range of reasons. There are the highest earnings, the largest focus on services, the widest offer of work, tourism and, of course, the best access to the best health care.

Situation in incidences of serious diseases are closer to Pattern object than Anti-pattern object (see Fig. 3). This could be a good message because these regions are closer to the minimum level of values of serious diseases indicators than to maximum level of these indicators. The two dimensions in Figure 3 did not specifically named because the original variables are not highly correlated as mentioned above.

This is the reason why the results of component analysis or factor analysis was not presented here for explanation of used two dimensions. Only this is known that *Dimension 1* together with *Dimension 2* visualise similarities or differences between objects base on indicators describing incidences caused by serious diseases.

Figure 3

Visualisation of Regions in Two-dimensional Space



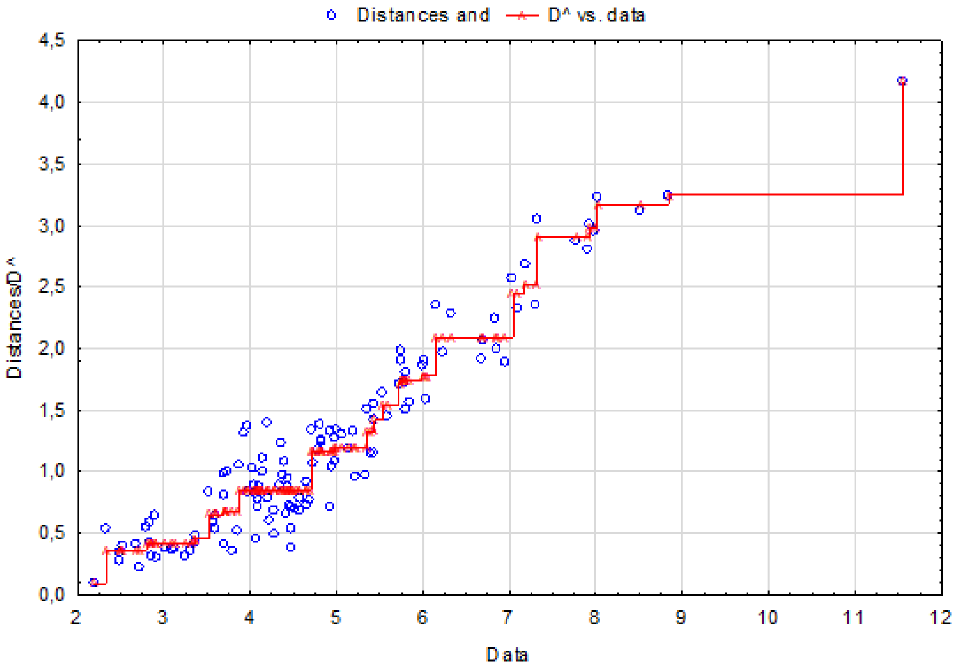
Source: Authors' calculations (ÚZIS ČR, 2016).

However, the quality of our visualisation is important as well. The more dimensions, the better results are from multidimensional scaling. On the other hand, the less dimensions, the more readable are these results. Shepard diagram can give the answer to the quality of the visualisation (see Figure 4). The dots represent distances between each pair of regions in Figure 4.

There are 120 pairs of regions. The original distances which come from original distance matrix are situated on the *x*-axis. After that the *y*-axis represents reproduced distances within multidimensional scaling. The dots show relationship between original distances and reproduced distances. Then *D-hat function* (D^) is fitted based on the dots. This function represent ideal relationship between original distances and reproduced distances within multidimensional scaling. The closer are the dots to D-hat function, the better is the visualisation in the two-dimensional space.

Figure 4

Shepard Diagram – Relationships between Original Distances and Reproduced Distances



Source: Authors' calculations (ÚZIS ČR, 2016).

The dots are situated close to the function but it is not the exact result of the quality of visualisation in two dimensions. The Stress value can give us the exact result of the quality of the model. The value of the, so-called, raw stress calculated by (1) equals to 0.12, which is *„fair"* according to Kruskal's table of rules describing quality of visualisation (Hebák et al., 2007).

Next step is to calculate aggregate rate d_i according to Eq. 2 based on the results of multidimensional scaling (specifically by the values of Dimension 1 and Dimension 2 shown in Table 5) and the distances from Pattern artificial object. The values of the aggregate rate d_i are mentioned in last column of Table 5.

Combination of multidimensional scaling and linear ordering is called as hybrid approach, as mentioned above. Based on Figure 5, it is possible to classify and to identify the regions with the similar or the same level of incidences (regions lying between two isoquants of development (circles) or in individual isoquants of development (circles)) but with different combination of the incidences of the individual serious diseases, according to Walesiak, 2016; Walesiak and Dehnel, 2018.

Table 5

Coordinate System and Aggregate Rate

Regions	Dimension 1	Dimension 2	d_i
South Bohemia (JHC)	–0.36	–0.56	0.66
South Moravia (JHM)	–0.29	–0.14	0.68
Karlovy Vary (KVK)	0.37	0.44	0.53
Hradec Králové (HKK)	0.01	0.11	0.62
Liberec (LBK)	–0.08	–0.83	0.57
Moravia-Silesia (MSK)	0.26	–0.34	0.54
Olomouc (OLK)	–0.09	–0.23	0.63
Pardubice (PAK)	–0.53	0.06	0.75
Plzeň (PLK)	0.72	–0.31	0.44
Prague (PHA)	–0.09	1.49	0.52
Central Bohemia (STC)	–0.64	0.45	0.76
Ústí nad Labem (ULK)	0.71	0.22	0.45
Vysočina (VYS)	–0.62	–0.34	0.74
Zlín (ZLK)	–0.36	–0.20	0.69
Pattern (P)	–1.58	0.17	1.00
Anti-pattern (AP)	2.58	0.04	0.00

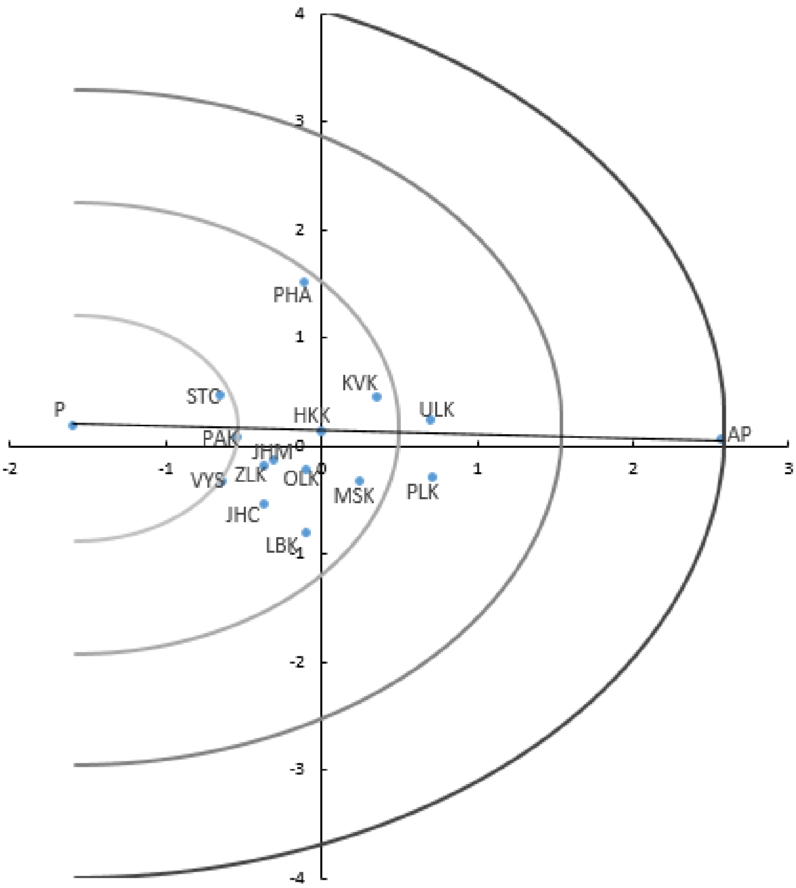
Source: Authors' calculations (ÚZIS ČR, 2016).

The Figure 5 brings a good message in the sense that the most of regions are closer to Pattern object in comparison with Anti-pattern object. On the other hand, this may indicate insufficient detection of these diseases because the incidences of serious diseases can be influenced not only by lifestyle, environmental pollution, socio-economic situation, genetic predisposition but also access to health care in individual regions. The four *„circles"* or isoquants of development which have a centre in the Pattern object are shown in Figure 5. The regions which lie in the same circle have the same level of aggregate incidences but they may have different combination of values of indicators describing incidences of individual serious diseases. This situation occurs, for example, in the Pardubice (PAK) and Vysočina (VYS) regions. The distance between Pattern *„best"* hypothetic object and Anti-pattern *„worst"* hypothetic object is divided by the circles into the four parts *„annulus"*. Two regions, Ústí nad Labem (ULK) and Plzeň (PLK), from the fourteen regions lie closer to the Anti-pattern object in the third annulus from the Pattern object. This situation represents high values of indicators of incidences caused by serious diseases in these two regions.

The relationship between aggregate rate d_i and distance from the Pattern object is displayed in Figure 6. There is shown linear ordering of regions in the Czech Republic. For instance, Central Bohemia (STC) region has the lowest incidences of serious diseases in contrast with Plzeň (PLK) region where incidences of serious diseases are the highest. From the set of the regions STC is the closest to the hypothetic *„best"* object and PLK is the closest to the hypothetic *„worst"* object. In addition, differences and similarities in the levels of aggregate incidences caused by serious diseases are displayed better in Figure 6. For example, Karlovy Vary (KVK) and Prague (PHA) regions are almost overlapping (nearly the same situation in incidences) but these two regions have different combinations of values of indicators describing incidences caused by mentioned serious diseases, as describes Figure 5. Next, very similar situation in level of incidences is in case of Pardubice (PAK) and Vysočina (VYS) regions or in case of Olomouc (OLK) and Hradec Králové (HKK) regions.

Figure 5

Visualisation of the Results of Multidimensional Scaling and the Level of Incidences

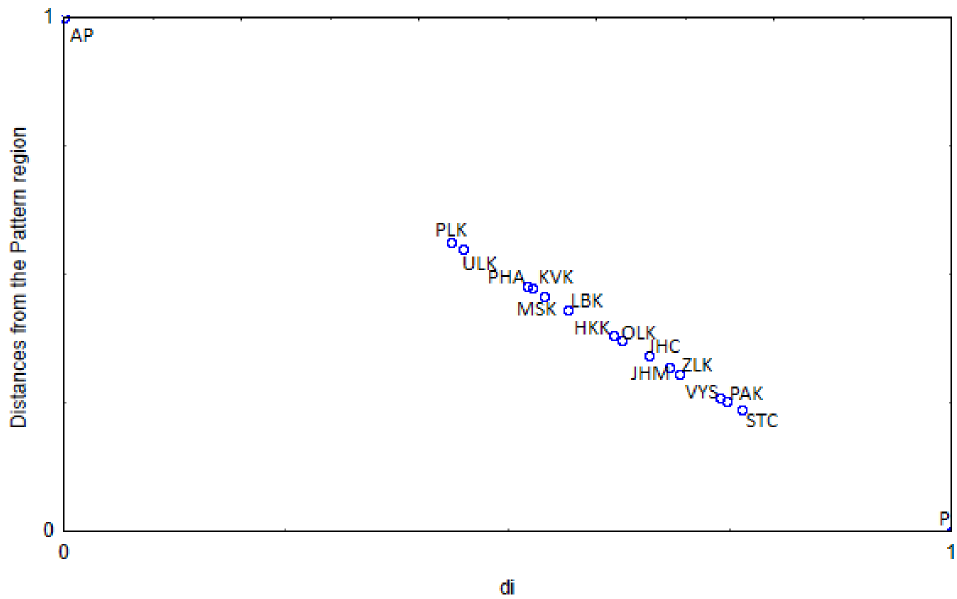


Source: Authors' calculations (ÚZIS ČR, 2016).

Based on the indicators describing incidences caused by serious diseases such as oncological, infectious, cardiovascular, diabetes and asthma hybrid approach point out the proximity of most regions to the *„best"* Pattern object (low level of values of indicators describing incidences caused by serious diseases) in both Figure 5 and Figure 6.

Figure 6

Visualisation of Linear Ordering of Regions in the Czech Republic Based on the Incidences of Serious Diseases

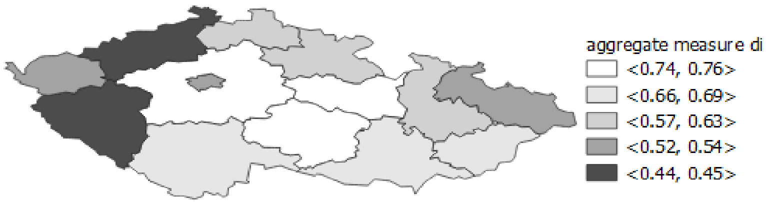


Source: Authors' calculations (ÚZIS ČR, 2016).

The Figure 7 displayed map of the Czech Republic with groups of regions based on values of aggregate measure d_i. The light colour mean high values of this measure, which means low incidences of serious diseases. The first group is created by regions Central Bohemia (STC), Pardubice (PAK) and Vysočina (VYS) where the level of incidences is the lowest. The south-eastern part of the CR represent the second the best group (South Moravia (JHM), Zlín (ZLK) and South Bohemia (JHC) regions). Next group is created by the regions Olomouc (OLK), Hradec Králové (HKK) and Liberec (LBK). Olomouc and Hradec Králové regions are well known for good access to health care thanks to University Hospitals. Liberec region which belongs to regions with higher level of incidences of serious diseases is a neighbour of Hradec Králové and represents Northern part of the Czech Republic. After that, the last two groups created by regions such as Moravia-Silesia (MSK), Karlovy Vary (KVK), Prague (PHA) regions and Ústí nad Labem (ULK) and Plzeň (PLK) regions belong to the regions with the worst situation in health status of population according to mentioned incidences. Primarily Eastern region (MSK) and North-western part of the Czech Republic are affected by bad socio-economic situation. On the other hand, the capital city Prague represents the region with the best access to health care, which shows the good quality of health care.

F i g u r e 7

Visualisation of the Situation in Individual Regions of the Czech Republic



Source: Authors' calculations (ÚZIS ČR, 2016).

The Czech Republic belongs to the states in Europe with the most accessible health care because of the highest coverage of Czech population through public health insurance. However, disparities in health status are considerable. The border areas show higher incidences than the centre of territory, which indicates the continuity of health status and socio-economic situation.

4. The Causes of Regional Differences

Confirmation of significant differences in the incidence of serious diseases in the regions of the Czech Republic is undoubtedly an important result of the presented hybrid method. However, it does not answer the question of what causes these differences. The ranking of regions according to the severity of serious diseases is the starting point for finding out which demographic, social, environmental and health indicators influence this order.

Using the Spearman's rank correlation coefficient (r_s), the degree of compliance was assessed in the order of the regions of the Czech Republic according to the incidence of serious diseases (values d_i in Table 5) and 13 demographic, 9 social, 6 environmental and 8 health indicators (see Appendix 1). In line with the previous parts of the article data for 2015, published by the Statistical Office of the Czech Republic, for comparison of regions within regional statistics have been used (CZSO, 2015).

In case of some indicators, the Spearman coefficients (r_s) do not confirm the dependence between the order of regions according to the incidence of serious diseases and the order of regions according to those indicators. The dependence is weak for the following indicators: D1 – Average age of population in years (r_s = –0.103), D2 – Aging index (r_s = –0.09), S3 – Median gross monthly wages (r_s = –0.200), S8 – Regional GDP per capita (CZK) – current prices (r_s = 0.068), E1 – Specific emissions – particulate matter (t/km2) (r_s = –0.068) and E2 – Specific emissions – carbon monoxide (t /km2) (r_s = 0.103), H2 – Hospitals (r_s = –0.069) and H6 – average time of treatment (r_s = –0.068). However, a significant dependence has been detected for some indicators. The Spearman rank correlation coefficients for these indicators are contained in Table 6.

T a b l e 6

Significant Spearman Rank Correlation Coefficients

Variables	**D4**	**D6**	**D9**	**D11**	**D13**	**S1**	**S2**	**S4**	**S6**
r_s	0.508	0.472	–0.314	–0.538	–0.433	–0.292	–0.323	–0.314	0.354
Variables	**E3**	**E4**	**E5**	**E6**	**H1**	**H4**	**H5**	**H7**	**H8**
r_s	0.473	–0.332	–0.824	0.486	–0.429	–0.459	–0.429	–0.543	–0.398

Source: Authors' calculations (CZSO, 2015; ÚZIS ČR, 2016).

Based on the obtained results, the occurrence of serious diseases is stronger affected by women's demographic indicators in comparison with the same ones for men. The Spearman coefficients between d_i and the indicators D3 and D5 (relevant to men – see Appendix) are equal 0.288 and 0.121 respectively. But the coefficients for the similar indicators corresponding with women are considerably larger (see Table 6 for D4 and D6).

Notable results regarding the determinants of serious diseases in the regions of the Czech Republic provide the factor analysis methods. The purpose of this analysis is to obtain a small number of factors that account for most of the variability in the 13 variables (this is the maximum possible number of variables relative to the number of regions). In this case, four factors have been extracted, since 4 factors had eigenvalues greater than number 1. Together they account for 89.88% of the variability in the original data, Factor 1 (F1) account 44.09%, Factor 2 (F2) 22.60%, Factor 3 (F3) 13.24% and Factor 4 (F4) 9.95%. Factor Loadings (Table 7) present the correlations between the original variables and the extracted factors and they are the key to identifying and understanding of the factors.

T a b l e 7

Factor Loadings Matrix after Varimax Rotation

Variables	**F1**	**F2**	**F3**	**F4**
D4	–0.8490	0.1862	–0.4522	0.1140
D6	–0.8959	0.2236	–0.3154	0.1080
D11	0.9330	0.1810	0.0664	–0.0274
D13	0.8721	–0.2374	0.3707	–0.0464
S1	0.2691	–0.1550	0.8977	–0.1913
S2	0.3587	0.1612	0.7928	0.0351
S4	0.5257	–0.0510	0.7968	–0.0797
E3	–0.3781	–0.3493	0.5403	0.4996
E4	0.3366	–0.4152	0.0754	–0.7870
E6	0.1175	–0.3086	–0.1262	0.8696
H1	–0.1569	0.9549	–0.2151	–0.0335
H4	–0.1259	0.8800	0.0429	–0.0601
H7	0.0515	0.9266	0.0170	–0.0161

Source: Authors' calculations (CZSO, 2015; ÚZIS ČR, 2016).

Based on the values of the factor loadings, we found out that the first factor F1 has strong correlation with the demographics indicators, the second factor F2 shows strong correlation with the health care indicators, the third factor F3 with the economic indicators and the fourth factor F4 with the environmental indicators.

T a b l e 8
Spearman Rank Correlation Coefficients for Factors and Distances d_i

Factors	F1	F2	F3	F4
r_s	–0.5604	–0.3495	–0.1736	0.6571

Source: Authors' calculations (CZSO, 2015; ÚZIS ČR, 2016).

The Spearman rank correlation coefficients in Table 8 quantify the degree of agreement in the arrangement of studied regions according to the occurrence of serious diseases (represented by d_i) and the extracted factors. We can see that the highest positive correlation with d_i shows factor F4 representing the environmental indicators and the highest negative correlation has F1 characterizing the demographic situation. On the other hand, the dependence of the occurrence of serious diseases and F3 representing the economic indicators is not very strong. This is probably due to relatively small economic differences between regions of the Czech Republic.

Conclusions

The main goal of the paper was to compare incidences of serious diseases within the regions in the Czech Republic by using so-called hybrid approach that combine the multidimensional scaling with the linear ordering of objects. The results of this approach confirm significant differences in incidences of serious diseases among the regions of the Czech Republic. Linear ordering provides a ranking of regions from „the worst" to „the best" (or vice versa), according on the situation in incidences of serious diseases, even in visual form.

Based on the findings, the elimination of regional differences in the incidence of serious diseases requires the elimination especially of environmental and demographic differences between regions in the Czech Republic. Proven dependencies can serve a more effective regional policy to reduce these disparities.

The hybrid approach is useful for comparing and linear ordering of a multivariate objects. However, this method does not answer the question of what factors cause inequalities of multidimensional objects, but it can be the starting point for this finding. This can be achieved by using Spearman rank correlation coefficients, or by applying the methods of factor analysis, as presented in the article. Obviously, the presented methods have general use in comparing multivariate objects and measuring multivariate dependencies.

Acknowledgement

This article was prepared with the support of the research project SGS_2019_018 funded by the University of Pardubice.

References

CZSO (2015): Regionální statistiky. Prague: Statistický úřad České republiky. Available at: <https://www.czso.cz/csu/czso/regiony_mesta_obce_souhrn>.

EUROPEAN COMMISSION (2007): Together for Health: A Strategic Approach for the EU 2008 – 2013. [White Paper.] Brussels: EC. Available at: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52007DC0630>.

EUROSTAT (2017): Sustainable Development in the European Union. Available at: <https://ec.europa.eu/eurostat/documents/3217494/8461633/KS-04-17-780-EN-N.pdf/f7694981-6190-46fb-99d6-d092ce04083f>.

FALL, F. – GLOCKER, D. (2018): Improving the Czech Health Care System. [OECD Economics Department Working Papers, No. 1522.] Paris: OECD Publishing. Available at: <https://doi.org/10.1787/9686b4f3-en>.

HAIR, J. F. et al. (1992): Multivariate Data Analysis with Regardings. New York: Macmillan Publishing Company.

HEBÁK, P. et al. (2005): Vícerozměrné statistické metody 2. Prague: Informatorium. ISBN 80-7333-036-9.

HEBÁK, P. et al. (2007): Vícerozměrné statistické metody 3. Prague: Informatorium. ISBN 978-80-7333-052-1.

HUSSEY, P. S. et al. (2009): A Systematic Review of Health Care Efficiency Measures. Health Services Research, *44*, No. 3, pp. 784 – 805.

JINDROVÁ, P. – KOPECKÁ, L. (2017): Assessment of Risk Factors of Serious Diseases in OECD Countries. In: The 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. [Proceedings.] Cracow: Foundation of the Cracow University of Economics, pp. 123 – 132.

JOUMARD, I. et al. (2010): Health Care Systems: Efficiency and Institutions. [OECD Economics Department Working Papers, No. 769.] Paris: OECD Publishing. Available at: <http://dx.doi.org/10.1787/5kmfp51f5f9t-en>.

KOPECKÁ, L. (2018a): Comparison of Mortality Caused by Serious Diseases within Regions of the Czech Republic. Scientific Papers of the University of Pardubice – Series D, Faculty of Economics and Administration, *43*, No. 2, pp. 112 – 122.

KOPECKÁ, L. (2018b): Vliv věkové struktury obyvatelstva na výskyt a léčbu chronických onemocnění v rámci Evropy. Forum Statisticum Slovacum, *14*, No. 1, pp. 19 – 28.

KOPECKÁ, L. – JINDROVÁ, P. (2017): Comparison of Mortality Due to Critical Illnesses in the EU Countries. In: The 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. [Proceedings.] Cracow: Foundation of the Cracow University of Economics, pp. 133 – 142.

OECD/EU (2018): Health at a Glance: Europe 2018: State of Health in the EU Cycle. Paris/EU, Brussels: OECD Publishing. Available at: <https://doi.org/10.1787/health_glance_eur-2018-en>.

PACAKOVA, V. et al. (2016): Comparison of Health Care Results in Public Health Systems of European Countries. In: European Financial Systems 2016: The 13th International Scientific Conference. [Proceedings.] Brno: Masaryk University, pp. 534 – 541.

PACÁKOVÁ, V. – KOPECKÁ, L. (2018): Inequalities in Health Status Depending on Socio-economic Situation in the European Countries. E+M Economics and Management, *21*, No. 2, pp. 4 – 20.

PACÁKOVÁ, V. – PAPOUŠKOVÁ, M. (2016): Multidimensional Comparisons of Health Systems Functioning in OECD Countries. International Journal of Mathematical Models and Methods in Applied Sciences, *10*, pp. 388 – 394.

PILYAVSKYY, A. I. et al. (2006): East-West: does it Makes a Difference to Hospital Efficiencies in Ukraine? Health Economics, *15*, No. 11, pp. 1173 – 1186.

PILYAVSKYY, A. I. – KOPECKÁ, L. (2018): The Efficiency of Health Care Systems in OECD Countries. Does it Make a Difference? In: European Financial Systems 2018: The 15th International Scientific Conference. [Proceedings.] Brno: Masaryk University, pp. 525 – 530.

PILYAVSKYY, A. I. – STAAT, M. (2006): Health Care in the CIS Countries: The Case of Hospitals in Ukraine. European Journal of Health Economics, *7*, No. 3, pp. 189 – 195.

SCHIFFMAN, S. S. et al. (1981): Introduction to Multidimensional Scaling. Theory, Methods and Applications. Bingley: Emerald Group Publishing Limited.

STANKOVIČOVÁ, I. – VOJTKOVÁ, M. (2007): Viacrozmerné štatistické metódy s aplikáciami. Bratislava: Iura Edition. ISBN 978-80-8078-152-1.

ÚZIS ČR (2016): Regionální zpravodajství Národního zdravotnického informačního systému. Available at: <https://reporting.uzis.cz/>.

ÚZIS ČR (2018): Novotvary 2016. Available at: <https://www.uzis.cz/katalog/zdravotnicka-statistika/novotvary>.

WALESIAK, M. (2016): Visualization of Linear Ordering Results for Metric Data with the Application of Multidimensional Scaling. Ekonometria, *2*, No. 52, pp. 9 – 21.

WALESIAK, M. – DEHNEL, G. (2018): Evaluation of Economic Efficiency of Small Manufacturing Enterprises in Districts of Wielkopolska Province Using Interval-valued Symbolic Data and the Hybrid Approach. In: The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. [Proceedings.] Cracow: Foundation of the Cracow University of Economics, pp. 563 – 572.

Appendix

D1	Average age of population in years
D2	Aging index – ratio of population aged 65+ to population aged 0–14
D3	Life expectancy at birth – men
D4	Life expectancy at birth – women
D5	Life expectancy at age 65 – men
D6	Life expectancy at age 65 – women
D7	Percentage of the population with tertiary education in % – annual averages
D8	Standardized mortality from circulatory system diseases (per 1 000 population) – males
D9	Standardized mortality from circulatory system diseases (per 1 000 population) – women
D10	Standardized mortality rate for neoplasms (per 1 000 population) – males
D11	Standardized mortality from neoplasms (per 1 000 population) –women
D12	Total mortality (per 1 000 population) – males
D13	Standardized mortality, total (per 1 000 population) – females
S1	General unemployment rate (annual average, %)
S2	Long–term unemployed of all unemployed in %
S3	Median gross monthly wages in CZK
S4	Percentage of people at risk of poverty or soc. Exclusion (2013)
S5	Percentage of Material Deprivation (2013)
S6	Dwellings completed per 1 000 mid–year population
S7	Average old – age (full) pension – solo in CZK
S8	Regional GDP per capita (CZK) – current prices
S9	Net disposable income of households per capita (CZK) – current prices
E1	Specific emissions – particulate matter (t / km2)
E2	Specific emissions – carbon monoxide (t / km2)
E3	Environmental protection investment per capita in CZK
E4	Coefficient of ecological stability
E5	Urban population in %
E6	Generation of municipal waste per capita in kg
H1	Physicians per 1 000 population
H2	Hospitals
H3	Beds per 1000 population
H4	Hospitalized patients1 per 1000 population
H5	Days of treatment per 1000population
H6	Average time of treatment (days)
H7	Paramedical workers with professional qualifications per 10 000 inhabitants
H8	General nurses and midwives

Source: CZSO, 2015; ÚZIS ČR, 2016