

Vliv věkové struktury obyvatelstva na výskyt a léčbu chronických onemocnění v rámci Evropy

Lucie Kopecská

Univerzita Pardubice, Fakulta ekonomicko-správní, Katedra matematiky a kvantitativních metod, Studentská 95, 532 10 Pardubice, Česká republika, 2. ročník, doktorské studium
Lucie.Kopeccka1@student.upce.cz

Abstrakt: Chronická onemocnění postihují evropskou populaci skrze všechny generace stále častěji. Tato onemocnění se vyznačují dlouhodobou léčbou, nedají se zcela vyléčit a zanechávají následky na zbytek života. Dlouhodobá léčba je nákladná a tím ovlivňuje výdaje do zdravotnictví. Cílem článku je odhalit vztah mezi věkovou strukturou obyvatelstva a výskytem a léčbou chronických onemocnění ve vybraných zemích Evropy pomocí vícerozměrných statistických metod za využití programů EXCEL, STATISTICA a R. Data byla získána z databází Eurostat, OECD a UNITED NATIONS.

Abstract: Chronic diseases affect European population through all generations more and more. These diseases are characterized by long-term treatment, they cannot be completely cured and they leave consequences for the rest of life. Long-term treatment is expensive and influences health expenditure. The main aim of this article is to reveal influence of the age structure of the population on incidences and treatment of chronic diseases in the selected European countries by multivariate statistical methods and EXCEL, STATISTICA and R programs. Data were obtained from databases of Eurostat, OECD and UNITED NATIONS.

Klíčová slova: Věková struktura obyvatelstva, chronická onemocnění, výdaje na zdravotnictví, vícerozměrné statistické metody.

Key words: Age structure of population, chronic diseases, health expenditure, multivariate statistical methods.

1 Úvod

Evropská populace je ohrožována různými chronickými onemocněními v různých etapách života. Věková struktura obyvatelstva je proto významným prediktorem výskytu těchto onemocnění. Některá onemocnění trápí evropskou populaci dlouhodobě, některá nově vznikají a postihují už i mladší generace, čímž se vytváří problém do budoucnosti, jelikož tato generace bude nadále stárnout a mohou se u ní kumulovat další závažná onemocnění. Náklady na léčbu chronických onemocnění jsou již dnes výrazné a v budoucnosti budou představovat důležitou součást výdajů na zdravotnictví, pro podrobnosti viz. (Jindrová & Slavíček, 2012); (Kubanová & Linda, 2014); (Linda et. al., 2014).

Eurostat databáze a databáze OECD poskytují data týkající se výskytu závažných chronických onemocnění v jednotlivých evropských zemích. Tyto databáze poskytují data nejen o výskytech onemocnění, ale také o úmrtnostech na ně, faktorech ovlivňujících zdraví (konzumace tabáku, alkoholu a obezita), výdajích na zdravotnictví a dalších (Eurostat, 2017); (OECD, 2016); (OECD, 2017).

2 Použité data a metody

2.1 Data

Cílem tohoto článku je stanovení vztahu mezi věkovou strukturou obyvatelstva a výskytem a léčbou chronických onemocnění v rámci vybraných evropských zemí.

Data, se kterými je pracováno, jsou získána z databází Eurostat a OECD. Tyto databáze poskytují informace týkající se nahlášených chronických onemocnění a vládních výdajů na zdravotnictví pro rok 2015 nebo nejbližší možné. Dále jsou data získána ze stránek UNITED NATIONS, kde jsou k dispozici informace o počtu obyvatel v jednotlivých evropských zemích, a to vždy k 1. červenci jednotlivých let podle pětiletých věkových skupin. Pro analýzy v tomto článku je vybráno následujících 11 proměnných získaných ze zdrojů (Eurostat, 2017); (OECD, 2016); (OECD, 2017); (UNITED NATIONS, 2017):

(X_1) podíl počtu osob ve věkové kategorii 0-29 let na celkové populaci, (X_2) podíl počtu osob ve věkové kategorii 30-49 let na celkové populaci, (X_3) podíl počtu osob ve věkové kategorii 50-69 let na celkové populaci, (X_4) podíl počtu osob ve věkové kategorii 70-89 let na celkové populaci, (C_1) nahlášené astma v %, (C_2) nahlášené infarkty v %, (C_3) nahlášené koronární srdeční choroby a angína pectoris v %, (C_4) nahlášený vysoký krevní tlak v %, (C_5) nahlášené chronické deprese v %, (C_6) nahlášená demence v %, (V_1) vládní výdaje do zdravotnictví jako podíl na HDP v %.

2.2 Korelační analýza

Vzhledem k využití pořadí zemí podle hodnot zvolených proměnných používáme v článku Spearmanův korelační koeficient. Tento koeficient udává statistickou závislost mezi dvojicemi proměnných. Výhodou Spearmanova korelačního koeficientu je to, že není citlivý na porušení předpokladu normality. Jeho výpočet je založený na pořadí hodnot, jak uvádí (Kubanová, 2008); (Řezanková et al., 2009).

Další možnou mírou pro měření závislostí mezi proměnnými je Kayser-Meyer-Olkin (KMO) index. Tento index měří závislost v rámci celé skupiny proměnných. Zjednodušenou a analogickou mírou pro KMO index je Measure of sampling adequacy (MSA) index, který je možné aplikovat na každou proměnou zvlášť. KMO index je založený na parciálních a jednoduchých korelačních koeficientech, a to podle (Hebák et al., 2007); (Stankovičová & Vojtková, 2007).

2.3 Faktorová analýza

S cílem zavedení menšího počtu proměnných využíváme faktorovou analýzu (FA), která je doplněním analýzy hlavních komponent a vychází z korelační matice. Důvodem pro aplikaci faktorové analýzy je zjištění společných faktorů.

Na určení počtu společných faktorů využíváme kritérium, při kterém bereme v úvahu vlastní čísla z korelační matice, která jsou větší než 1. Takto získaný počet faktorů můžeme použít v další analýze, jak uvádí (Hebák et al., 2007); (Pacáková & Papoušková, 2016); (Stankovičová & Vojtková, 2007).

Důležitým výstupem faktorové analýzy jsou faktorové zátěže, které určují vztah mezi původními proměnnými a nově zavedenými faktory. Ke zjištění těchto zátěží využíváme tzv. rotace faktorů, kde k nejznámější a nejpoužívanější patří Varimax rotace. Finálním výstupem faktorové analýzy jsou faktorová skóre, která zjišťujeme pro jednotlivé objekty, jak uvádí (Pacáková et al., 2016); (Pacáková & Papoušková, 2016).

2.4 Shluková analýza

Hlavním cílem shlukové analýzy je zařazení objektů do homogenních skupin a to tak, že objekty uvnitř skupin jsou si co nejvíce podobné a objekty mezi skupinami jsou co nejvíce rozdílné. Před samotnou aplikací shlukové analýzy nejprve musíme sestavit datovou matici, kde řádky této matice představují objekty, které budou shlukovány a sloupce představují jednotlivé proměnné, podle kterých jsou objekty zařazovány do shluků. Datová matice má rozměry $m \times n$, jak uvádí (Hebák et al., 2007); (Řezanková et al., 2009); (Staničková, 2015).

Existuje více měr sloužících k měření vzdáleností mezi objekty. Nejznámější z nich je Euklidovská vzdálenost, kterou zde využíváme.

Pro prezentaci výsledků shlukové analýzy využíváme jak hierarchickou, tak nehierarchickou metodu shlukování. Postup aplikace těchto metod může být následující: nejprve použijeme hierarchickou shlukovací metodu (Wardovu metodu) a následně metodu nehierarchickou (K-průměrů), která na začátku shlukování potřebuje apriorní informaci o počtu shluků, jak uvádí (Hebák et al., 2007); (Petr et al., 2010); (Řezanková et al., 2009).

3 Vliv věkové struktury obyvatelstva na výskyt a léčbu chronických onemocnění

Na základě dat popsaných výše pomocí vícerozměrných statistických metod a programů EXCEL, STATISTICA a R je odhalován vliv věkové struktury obyvatelstva na výskyt a léčbu vybraných chronických onemocnění, a to v rámci vybraných zemí Evropy.

Nejprve je zjištěna statistická závislost mezi jednotlivými proměnnými pomocí Spearmanova korelačního koeficientu. V korelační matici je odhaleno více silných pozitivních i negativních závislostí mezi proměnnými. Co se týká věkové struktury obyvatelstva a výskytu chronických onemocnění, objevují se zde značné pozitivní korelace, a to konkrétně mezi relativním počtem obyvatel

ve věkové kategorii 0-29 let a výskytem astmatu ($r_s=0,40$; $p=0,047$) a relativním počtem obyvatel ve věkové kategorii 50-69 let a nemocí postihujících srdce ($r_s>0,42$; $p<0,035$). Silné závislosti existují mezi výskyty jednotlivých chronických onemocnění. Pro příklad lze uvést silné pozitivní korelace mezi výskytem astmatu, chronickými depresemi a demencí, dále pak mezi výskytem infarktů, koronárních srdečních chorob a angíny pectoris a vysokým krevním tlakem ($r_s>0,63$; $p<0,0007$). Naopak značné negativní korelace jsou mezi proměnnými týkající se astmatu a chorob postihujících srdce a dále mezi proměnnými týkající se duševního stavu (deprese, demence) a nemocí postihujících srdce ($r_s<-0,40$; $p<0,049$). Vládní výdaje do zdravotnictví silně pozitivně korelují s výskytem astmatu ($r_s=0,76$; $p=0,00001$) a duševními nemocemi ($r_s>0,55$; $p<0,005$) a negativně s výskytem chorob ohrožujících srdce ($r_s<-0,52$; $p<0,008$). Na základě takto zjištěných statistických závislostí je možné říci, že zvyšováním počtu obyvatel ve věkové kategorii 0-29 let se výskyt astmatu zvyšuje a zvyšováním počtu obyvatel ve věkové kategorii 50-69 let se zvyšuje počet nemocí postihujících srdce. Co se týká závislostí mezi jednotlivými chronickými onemocněními, je možné si všimnout, že s růstem výskytu astmatu a chronických depresí roste i výskyt obyvatel postižených demencí. Naopak s růstem astmatu klesá počet chorob postihujících srdce a s růstem počtu duševních nemocí počet chorob postihujících srdce také klesá. Dále bylo prokázáno, že s růstem vládních výdajů do zdravotnictví roste i výskyt astmatu a duševní choroby, nicméně snižuje se počet chorob postihujících srdce. Na základě korelační matice pořadových korelací, však nebyla prokázána silná statistická závislost mezi počtem obyvatel v nejvyšší věkové kategorii a duševními chorobami, která by se ovšem dala očekávat, a to na hladině významnosti 5 %.

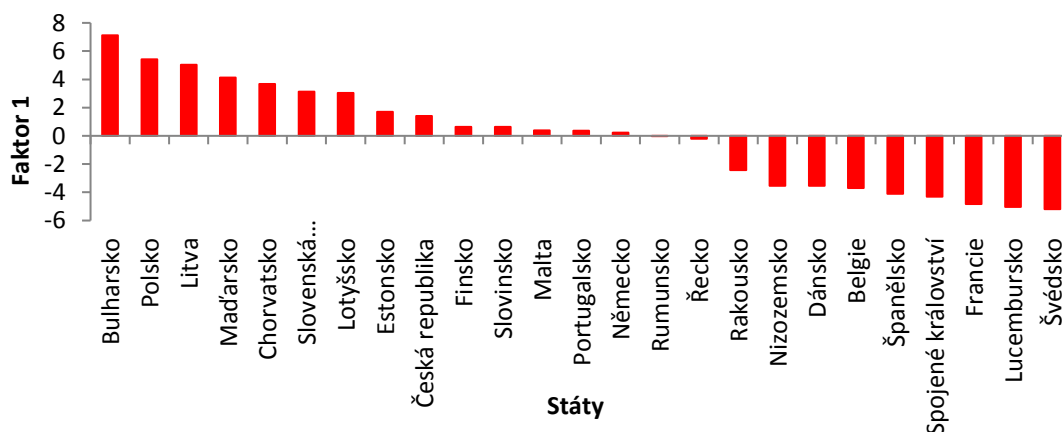
Aby byla statistická závislost zjištěna také pro celou skupinou proměnných, je použita MSA míra popsaná výše, která v tomto případě nabývá hodnoty 0,51, což je dostačující výsledek pro aplikaci faktorové analýzy na tato data. Tato míra je získána pomocí programu R.

Jak již bylo výše zmíněno, cílem faktorové analýzy je stanovení menšího počtu společných faktorů, kde každý z faktorů je vysvětlen pomocí původních proměnných. Nejprve je nutné stanovit vhodný počet faktorů, tzn. vhodný počet vlastních čísel z korelační matice. V tomto případě jsou vybrány první 3 vlastní čísla, která jsou větší než 1, a při kterých je vysvětleno téměř 78 % variability původních dat. V tabulce 1 je možné vidět faktorové zátěže po rotaci faktorů, které ukazují vztah společných faktorů a původních proměnných, pro podrobnosti.

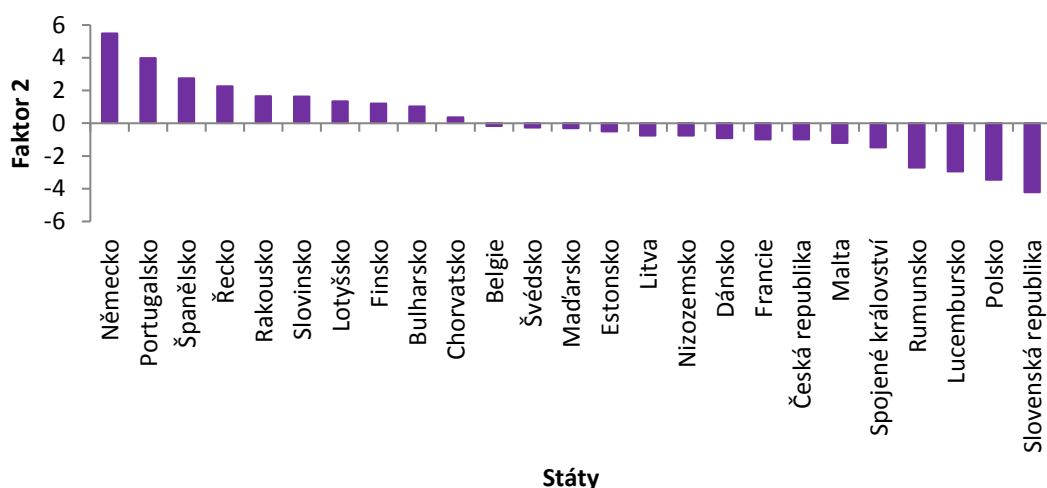
Tab. 1 Faktorové zátěže po rotaci faktorů (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)

Proměnné	Faktor 1 (F_1)	Faktor 2 (F_2)	Faktor 3 (F_3)
X_1	-0,20	-0,82	0,47
X_2	-0,26	0,07	-0,86
X_3	0,73	0,24	0,25
X_4	0,11	0,87	0,13
C_1	-0,37	-0,08	0,76
C_2	0,74	-0,15	-0,28
C_3	0,88	-0,12	-0,26
C_4	0,83	0,35	-0,22
C_5	-0,30	0,46	0,57
C_6	-0,49	0,59	0,53
V_1	-0,49	0,09	0,73

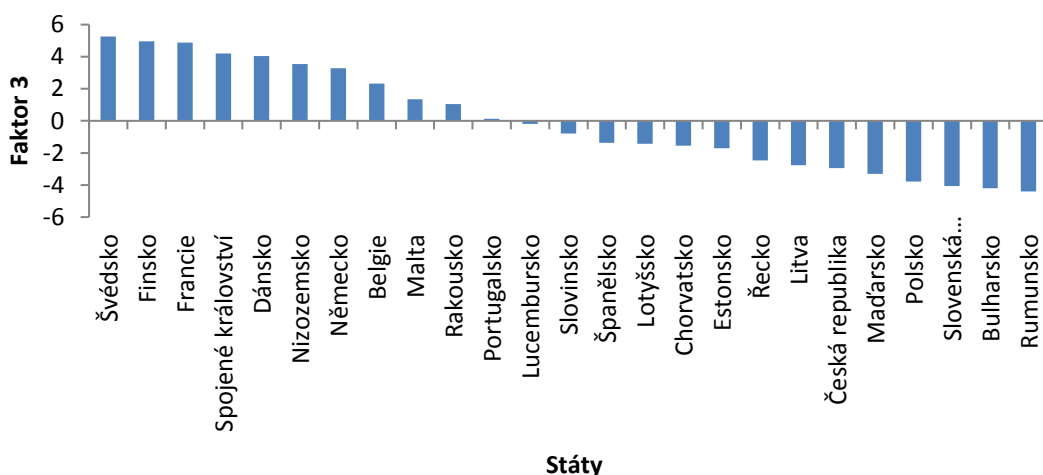
Na základě tabulky 1 je možné říci, že faktor F_1 silně koreluje s proměnnými týkající se relativního počtu obyvatel ve věkové kategorii 50-69 let a výskytu chorob postihujících srdce (C_2 - C_4). F_2 je vyjádřen pomocí relativního počtu obyvatel ve věkové kategorii 70-89 let, demence a relativním počtem obyvatel ve věkové kategorii 0-29 let, kde existuje pozitivní korelace mezi F_2 a X_4 , C_6 a negativní korelace mezi F_2 a X_1 . V případě F_3 existují pozitivní korelace s proměnnými týkající se astmatu, chronických depresí a vládních výdajů do zdravotnictví a negativní korelace s proměnnou vyjadřující relativní počet obyvatel ve věkové kategorii 30-49 let. Faktorové zátěže poukazují na to, že evropská populace ve věkové kategorii 50-69 let má největší problémy s kardiovaskulárními nemocemi, nejstarší osoby jsou postiženy demencí a evropské země, které vynakládají vysoké vládní výdaje do zdravotnictví, jsou nejvíce zasaženy výskytem astmatu a chronických depresí. V grafech 1, 2, 3 jsou znázorněna standardizovaná faktorová skóre pro jednotlivé evropské státy.



Graf 1 Pořadí evropských států podle Faktoru 1 (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)



Graf 2 Pořadí evropských států podle Faktoru 2 (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)



Graf 3 Pořadí evropských států podle Faktoru 3 (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)

V grafu 1 jsou státy s nejvyšším standardizovaným faktorovým skóre nejvíce postiženy onemocněním srdce a vysokým krevním tlakem a je zde i vysoké zastoupení obyvatel ve věkové kategorii 50-69 let. V grafu 2 evropské země s nejvyšším standardizovaným faktorovým skóre patří k zemím, kde jsou lidé ohrožováni demencí, je zde vysoké zastoupení obyvatel v nejvyšší věkové kategorii a nízké zastoupení obyvatel v nejnižší věkové kategorii. V posledním grafu 3 nejvyšším standardizovaným faktorovým skóre jsou ohodnoceny státy, které jsou nejvíce postiženy výskytem astmatu, chronickými depresemi, a ve kterých jsou nejvyšší vládní výdaje do zdravotnictví. Na druhé straně tyto

státy patří k těm, kde je nízký relativní počet obyvatel ve věkové kategorii 30-49 let.

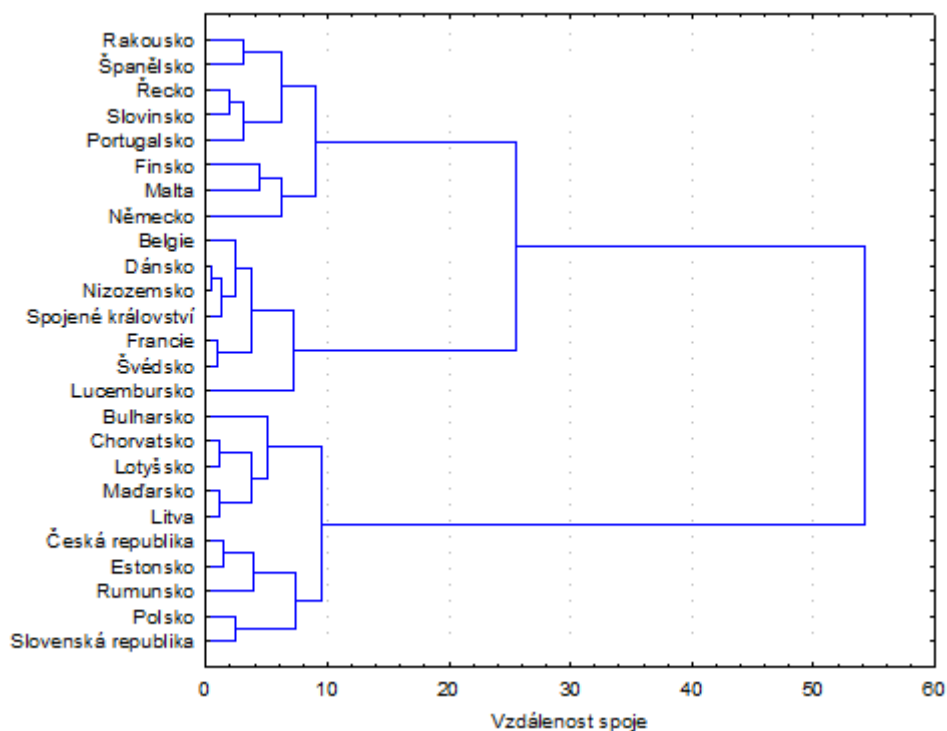
V následující tabulce 2 jsou evropské země zařazeny do 8 skupin, které reprezentují všechny kombinace hodnot faktorových skóre.

Tab. 2 Pozitivní a negativní hodnoty faktorových skóre pro jednotlivé evropské země (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)

Hodnoty faktorových skóre	Státy
+++	Finsko, Portugalsko, Německo
++-	Bulharsko, Chorvatsko, Lotyšsko, Slovinsko
+--	Polsko, Litva, Maďarsko, Slovenská republika, Estonsko, Česká republika
-++	Rakousko
---+	Nizozemsko, Dánsko, Belgie, Spojené království, Francie, Švédsko
---	Lucembursko, Rumunsko
-+-	Řecko, Španělsko
+--+	Malta

Na nově zavedené 3 faktory, které vysvětlily vztahy mezi původními 11 proměnnými, je aplikována shluková analýza. Konkrétně se jedná o Wardovu metodu a metodu K-průměrů, které jsou uvedeny výše, protože každá z těchto metod má jak své výhody, tak nevýhody, a je proto vhodné výsledky těchto dvou metod porovnat. Nejprve je použita hierarchická metoda shlukování, a to Wardova metoda a pro měření vzdáleností mezi objekty je použita Euklidovská vzdálenost. Následně je použita nehierarchická metoda shlukování, a to metoda K-průměrů, která získá apriorní informaci o počtu shluků na základě metody hierarchické.

V grafu 4, který je získán z programu STATISTICA, je vyobrazen dendrogram znázorňující podobnosti mezi jednotlivými evropskými státy. Tento graf odhaluje 3 významné shluky.



Graf 4 Dendrogram zobrazující 3 významné shluky evropských států (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)

Z grafu 4 je možné vyčíst, že první shluk je tvořen státy Rakousko, Španělsko, Řecko, Slovinsko, Portugalsko, Finsko, Maltu a Německo. Další shluk zahrnuje Belgii, Dánsko, Nizozemsko, Spojené království, Francii, Švédsko, a Lucembursko. Poslední shluk je tvořen státy Bulharsko, Chorvatsko, Lotyšsko, Maďarsko, Litva, Česká republika (ČR), Estonsko, Polsko a Slovenská republika (SR).

Při následném shlukování metodou K-průměrů program STATISTICA nabízí analýzu rozptylu, pomocí níž je možné otestovat, zda jsou jednotlivé proměnné vstupující do shlukové analýzy statisticky významné. Tabulka 3 ukazuje výsledky analýzy rozptylu, které vypovídají o tom, že všechny 3 nově zavedené faktory jsou statisticky významné a je vhodné je použít pro shlukování.

Tab. 3 Analýza rozptylu (Zdroj: vlastní zpracování podle Eurostat, 2017; OECD, 2016; OECD, 2017; UNITED NATIONS, 2017)

Proměnné	Meziskupinový SČ	sv	Vnitřní SČ	sv	F	významnost p
Faktor 1	251,85	2	64,26	22	43,11	0,000000
Faktor 2	57,78	2	64,16	22	9,91	0,000855
Faktor 3	178,28	2	74,61	22	26,29	0,000001

Výstupy shlukování metodou K-průměrů pro 3 shluky, které byly určeny v rámci Wardovy metody, se v tomto případě rovnají výstupům shlukování Wardovou metodou při použití Euklidovské vzdálenosti, avšak ne vždy tomu tak musí být, a to kvůli citlivosti průměrů na odlehlé objekty v případě metody K-průměrů.

Do prvního shluku jsou zařazeny státy jako Rakousko, Finsko, Německo atd., které jsou na tom nejhůře s výskytem demence, hůře s výskytem astmatu, chronických depresí a mají vysoké vládní výdaje do zdravotnictví, a to podle FA. Druhý shluk tvoří státy jako Belgie, Dánsko, Nizozemsko atd., kde se podle FA nejméně vyskytují nemoci srdce a méně duševní nemoci. Poslední shluk tvoří státy, jako je ČR, SR, Polsko atd., které jsou na tom nejlépe s astmatem, chronickými depresemi, mají nízké vládní výdaje do zdravotnictví a jsou na tom nejhůře s nemocemi srdce, a to opět podle výsledků FA. Nicméně je si možné všimnout i rozdílů ve výsledcích shlukové a faktorové analýzy. Například v případě Lucemburska a Rumunska, které FA zařadila do stejné skupiny na základě hodnot faktorových skóre, viz. tabulka 2. Obě země v rámci FA byly zařazeny do skupiny, kde podle všech faktorů vykazují negativní hodnoty faktorových skóre. Nicméně Rumunsko je průměrným státem, co se týká onemocnění srdce a naopak Lucembursko je průměrným státem v případě výskytu astmatu a chronických depresí, což může vést k odlišným výsledkům v rámci shlukové analýzy.

4 Závěr

Databáze Eurostat a OECD poskytují množství dat týkajících se závažných onemocnění. V tomto článku byl zjišťován vztah mezi věkovou strukturou obyvatelstva a výskytem a léčbou chronických onemocnění. Spearmanův korelační koeficient odhalil významné statistické závislosti mezi proměnnými, které byly zajímavé pro další zkoumání pomocí faktorové analýzy. Na základě ní byly stanoveny tři společné faktory, které vysvětlily vzájemné vztahy mezi nimi a jednotlivými proměnnými. Zajímavostí je, že v rámci Spearmanova korelačního koeficientu byla zjištěna značná korelovanost mezi výskytem astmatu a nejnižší věkovou kategorií, což však faktorová analýza nepotvrdila.

V rámci shlukové analýzy byly zjištěny 3 významné shluky, které obsahují u obou použitých metod shlukování stejné evropské státy. V některých případech se však liší od výsledků FA.

Vzhledem k tomu, že se některá chronická onemocnění nově objevují a postihují evropskou populaci v raném věku, je nutné předpokládat, že se tato onemocnění mohou s postupem času kumulovat a náklady na jejich léčbu porostou, což zvýší výdaje do zdravotnictví a zatíží rozpočty jednotlivých států.

5 Literatura

- Eurostat (2017). *Persons reporting a chronic diseases, by diseases, sex, age and educational attainment level*. [Dostupné na: http://ec.europa.eu/eurostat/web/products-datasets/product?code=hlth_ehis_cd1e].
- Hebák, P., Hustopecký, J., Pecáková, I., Průša, M., Řezánková, H., Svobodová A., Vlach, P. (2007). *Vícerozměrné statistické metody (3)*. Druhé vydání. Praha: Informatorium.
- Jindrová, P., Slavíček, O. (2012). *Life Expectancy Development and Prediction for Selected European Countries*. Managing and Modelling of Financial Risks. Vol. VI, pp. 303-312.
- Kubanová, J. (2008). *Statistické metody pro ekonomickou a technickou praxi*. Třetí vydání. Bratislava: Statist.
- Kubanová, J., Linda, B. (2014). *The process of population aging and its impact on country's economy*. Political sciences, Law, Finance, Economics and Tourism, Vol. IV, pp. 767-774.
- Linda, B., Kubanová, J., Pacáková, V. (2014). *Mortality Rate Estimate for the Old Age Population by Bootstrap Method*. Political sciences, Law, Finance, Economics and Tourism, Vol. IV, pp. 481-486.
- OECD (2016). *Health at a glance: Europe (2016)*. [Dostupné na: https://ec.europa.eu/health/state/glance_en].
- OECD (2017). *OECD health statistics 2017*. [Dostupné na: <http://www.oecd.org/els/health-systems/health-data.htm>].
- Pacáková, V., Jindrová, P., Zapletal, D. (2016). *Comparison of Health Care Results in Public Health Systems of European Countries*. European Financial Systems 2016. Vol. XIII, pp. 534-541.
- 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.
- Petr, P., Křupka, J., Provazníková R. (2010). *Statistical Approach to Analysis of the Regions*. International Conference on Applied Computer Science. 10, pp. 280-285.
- Řezánková, H., Húsek, D., Snášel, V. (2009). *Shluková analýza dat*. Druhé vydání. Praha: Professional Publishing.
- Staničková, M. (2015). *Classifying the EU Competitiveness Factors Using Multivariate Statistical Methods*. Procedia Economics and Finance. 23, pp. 313-320.
- Stankovičová, I., Vojtková, M. (2007). *Viacrozměrné statistické metody s aplikáciami*. Bratislava: IURA Edition.
- UNITED NATIONS. (2017). *Population indicators*. [Dostupné na: <https://esa.un.org/unpd/wpp/Download/Standard/Population/>].

6 Poděkování

Článek byl vypracovaný za podpory Stipendijního fondu FES, UPCE (SF420099).