Slovenská štatistická a demografická spoločnosť
Slovak Statistical and Demographic Society

FORUM STATISTICUM SLOVACUM

Recenzovaný vedecký časopis
Scientific peer-reviewed journal

Číslo/Issue: 2/2022
Ročník/Volume: XVIII

FORUM STATISTICUM SLOVACUM

recenzovaný vedecký časopis Slovenskej štatistickej a demografickej spoločnosti
scientific peer-reviewed journal of the Slovak Statistical and Demographic Society

Ročník/Volume: XVIII (2022) **Číslo/Issue:** 2

Editori / Editors

Iveta Stankovičová (Fakulta managementu Univerzity Komenského v Bratislave, Slovensko)
Martin Boďa (Ekonomická fakulta Univerzity Mateja Bela v Banskej Bystrici, Slovensko)

Redakčná rada / Editorial Board

Branislav Bleha (Prírodovedecká fakulta Univerzity Komenského v Bratislave, Slovensko)
Boris Burcin (Prírodovedecká fakulta Univerzity Karlovej v Prahe, Česko)
Joanna Dębicka (Fakulta manažmentu, informatiky a financií Ekonomickej univerzity vo Wrocławi, Poľsko)
Estefanía Mourelle Espasandín (Fakulta ekonómie a podnikania Univerzity v La Coruña, Španielsko)
Stanislav Katina (Ústav matematiky a štatistiky Masarykovej univerzity v Brne, Česko)
Jana Kubanová (Fakulta ekonomicko-správna Univerzity Pardubice, Česko)
Dagmar Kusendová (Prírodovedecká fakulta Univerzity Komenského v Bratislave, Slovensko)
Viera Labudová (Fakulta hospodárskej informatiky Ekonomickej univerzity v Bratislave, Slovensko)
Jitka Langhamrová (Fakulta informatiky a štatistiky Vysokej školy ekonomickej v Prahe, Česko)
Ivan Lichner (Ekonomický ústav Slovenskej akadémie vied, Slovensko)
Tomáš Löster (Fakulta informatiky a štatistiky Vysokej školy ekonomickej v Prahe, Česko)
Janka Medová (Fakulta prírodných vied Univerzity Konštantína Filozofa v Nitre, Slovensko)
Silvia Megyesiová (Podnikovohospodárska fakulta v Košiciach Ekonomickej univerzity v Bratislave, Slovensko)
Oľga Nanásiová (Fakulta elektrotechniky a informatiky Slovenskej technickej univerzity v Bratislave, Slovensko)
Viliam Páleník (Ekonomický ústav Slovenskej akadémie vied, Slovensko)
Marek Radvanský (Ekonomický ústav Slovenskej akadémie vied, Slovensko)
Hana Řezanková (Fakulta informatiky a štatistiky Vysokej školy ekonomickej v Prahe, Česko)
Ľubica Sipková (Fakulta hospodárskej informatiky Ekonomickej univerzity v Bratislave, Slovensko)
Mária Stachová (Ekonomická fakulta Univerzity Mateja Bela v Banskej Bystrici, Slovensko)
Anna Tirpáková (Fakulta prírodných vied Univerzity Konštantína Filozofa v Nitre, Slovensko)
✝ Vladimír Úradníček (Ekonomická fakulta Univerzity Mateja Bela v Banskej Bystrici, Slovensko)
Mária Vojtková (Fakulta hospodárskej informatiky Ekonomickej univerzity v Bratislave, Slovensko)
Tomáš Želinský (Ekonomická fakulta Technickej univerzity v Košiciach, Slovensko)

Vydavateľ: Slovenská štatistická a demografická spoločnosť, Miletičova 3, 824 67 Bratislava, Slovensko. **Publisher:** Slovak Statistical and Demographic Society, Miletičova 3, 824 67 Bratislava, Slovakia**. Adresa redakcie/Editorial office:** Miletičova 3, 824 67 Bratislava, Slovakia. **IČO/Company ID:** 00178764. **DIČ/Tax ID:** 2021504276**. Mailový kontakt/E-mail contact:** adm.ssds@ssds.sk **Web site/Webové sídlo:** http://www.ssds.sk/

Registráciu vykonalo Ministerstvo kultúry Slovenskej republiky. **Dátum registrácie:** 27. júla 2005. **Evidenčné číslo:** EV 3287/09. **Tematická skupina:** B1. **Periodicita:** minimálne dvakrát ročne. **ISSN:** 1336-7420.
Registered by the Ministry of Culture of the Slovak Republic. **Date of registration:** 27 July 2005. **Registration No:** EV 3287/09. **Topic group:** B1. **Periodicity:** at least two issues per year. **ISSN:** 1336-7420.

FORUM STATISTICUM SLOVACUM Vol XVIII (2022) No 2 pp 1-12

Bayesovské modelování dat v prostředí programového systému SAS
Bayesian data modelling in the SAS software system environment

Roman Pavelka
Štatistický úrad Slovenskej republiky, Odbor metód štatistických zisťovaní, Lamačská cesta 3, 840 05 Bratislava, Slovenská republika
Statistical Office of the Slovak Republic, Statistical Surveys and Methodology Department, Lamačská cesta 3, 840 05 Bratislava, Slovak Republic
roman.pavelka@statistics.sk

Abstrakt: *Bayesovské metody se stávají stále populárnějšími v moderní statistické analýze a jsou aplikovány na širokou škálu spektra vědeckých oborů a oblastí výzkumu. Předkládaný příspěvek představuje principy postupů procedur v modulu STAT programového sytému SAS. Je zde uvedena procedura MCMC, která je určena pro obecně použitelná pro bayesovské výpočty. Procedura MCMC poskytuje výzkumníkům provádění analýzy dat na široké škále komplexních bayesovských statistických modelů. Procedura využívá Markovský řetězec Monte Carlo (MCMC) algoritmus pro náhodné výběry vzorků z libovolného aposteriorního rozdělení, které je definováno apriorními rozděleními parametrů a pravděpodobnostními funkcemi pro sledovaná zadaná data. Tento článek popisuje, jak používat proceduru MCMC v postupech pro odhad, inferenci a predikci a okrajově se zmiňuje i o dalších procedurách s možností bayesovského přístupu k analýze dat.*

Abstract: *Bayesian methods are becoming increasingly popular in modern statistical analysis and they are applied to a wide range of scientific disciplines and fields of research. The submitted contribution presents the principles of this procedure in the STAT module of the SAS programming system. There MCMC procedure is designed for generally applicable Bayesian calculations. The MCMC procedure provides researchers the data analysis on a wide range of complex Bayesian statistical models. The procedure uses the Monte Carlo Markov Chains (MCMC) algorithm for random sampling of any posterior distribution, which is defined by any prior distributions with parameters and probability functions for the data to be analysed. This article describes how to use the MCMC procedure in estimations, inferences and predictions, and marginally mentions other procedures with the possibility of a Bayesian approach to data analysis.*

Klíčové slova: *apriorní rozdělení, aposteriorní rozdělení, MCMC, pravděpodobnostní funkce.*

Key words: *prior distribution, posterior distribution, MCMC, probability function.*

1 Úvod

Nejčastěji používané statistické metody jsou známé jako četnostní (nebo-li tzv. klasické) metody. Tyto metody předpokládají, že neznámé parametry jsou fixními konstantami, a definují pravděpodobnost pomocí omezujících relativních frekvencí. Z těchto předpokladů vyplývá, že pravděpodobnosti jsou objektivní a že nelze dělat pravděpodobnostní úsudky o parametrech, protože jsou pevně

dané. Bayesovské metody nabízejí alternativní přístup; zacházejí s parametry jako s náhodnými proměnnými a definují pravděpodobnost jako „stupně víry“[1]. Z těchto postulátů vyplývá, že pravděpodobnosti jsou subjektivní a že je možné o parametrech uvádět pravděpodobnostní výroky. Termín „bayesovský“ pochází z převládajícího používání Bayesovy věty, která byla pojmenována po reverendu Thomasi Bayesovi, presbyteriánském duchovním z osmnáctého století. Bayes se zajímal o řešení otázky inverzní pravděpodobnosti: jaká je po pozorování množiny událostí pravděpodobnost jedné události?

2 Bayesův vzorec pro náhodné jevy a hypotézy

Předpokládejme, že se má realizovat odhad parametru θ z dat $\boldsymbol{y} = \{y_1, \cdots y_n\}$ pomocí statistického modelu popsaného hustotou pravděpodobnosti $p(\boldsymbol{y}|\theta)$ za platnosti parametru θ. Bayesovská statistika postuluje, že tento odhad přesně určit nelze, přičemž nejistota ohledně parametru θ je vyjádřena pomocí pravděpodobnostních úsudků a rozdělení. Například, pokud pravděpodobnostní rozdělení parametru θ se řídí normálním rozdělením s průměrem *0* a rozptylem *1*, v duchu bayesovské statistiky se má za to, že toto rozdělení nejlépe vystihuje nejistotu spojenou s parametrem θ. Následující kroky popisují základní prvky bayesovské dedukce:

- Rozdělení pravděpodobnosti pro parametr θ je formulováno jako $\pi(\theta)$ a je známé jako tzv. rozdělení apriorní[2].
- Na základě pozorovaných dat $\boldsymbol{y}$ se zvolí statistický model $p(\boldsymbol{y}|\theta)$ pro popis rozdělení daného $\boldsymbol{y}$ za předpokládané platnosti parametru θ.
- Víra (přesvědčení) o parametru θ se upraví kombinací informace z apriorního rozdělení $\pi(\theta)$ a dat $\boldsymbol{y}$ výpočtem aposteriorního rozdělení $p(\theta|\boldsymbol{y})$.

Úpravy víry (přesvědčení) o parametru θ se provádí pomocí tzv. Bayesovy věty, která odpovídá kombinaci apriorního rozdělení $\pi(\theta)$ a statistického modelu $p(\boldsymbol{y}|\theta)$ následujícím způsobem:

$$p(\theta|\boldsymbol{y}) = \frac{p(\theta,\boldsymbol{y})}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\theta)\pi(\theta)}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\theta)\pi(\theta)}{\int p(\boldsymbol{y}|\theta)\pi(\theta)\mathrm{d}(\theta)}\,, \tag{1}$$

kde

$$p(\boldsymbol{y}) = \int p(\boldsymbol{y}|\theta)\pi(\theta)\mathrm{d}(\theta). \tag{2}$$

Veličina $p(\boldsymbol{y})$ daná výrazem (2) vyjadřuje marginální (okrajové) rozdělení pozorovaných dat $\boldsymbol{y}$ a sehrává úlohu normalizační konstanty aposteriorního rozdělení. Pokud existuje marginální rozdělení ve tvaru konečného (vlastního)

[1] Pravděpodobnost události je míra, s jakou je možné se domnívat, že je událost pravdivá.

[2] Vyjadřuje víru (přesvědčení) o parametru θ ještě před analýzou dat $\boldsymbol{y}$.

integrálu, neposkytuje žádnou dodatečnou informaci o aposteriorním rozdělení $p(\theta|\boldsymbol{y})$. Proto je aposteriorní rozdělení $p(\theta|\boldsymbol{y})$ úměrné statistickému modelu $p(\boldsymbol{y}|\theta)$, který specifikuje podmíněné rozdělení pozorovaných dat za daných parametrů, a apriorní pravděpodobnosti $\pi(\theta)$ pro parametr θ. Způsob zápisu Bayesovy věty (1) pomocí věrohodnostní funkce $L(\theta)$ je následující:

$$p(\theta|\boldsymbol{y}) = \frac{L(\Theta)\pi(\Theta)}{\int L(\Theta)\pi(\Theta)\mathrm{d}(\Theta)}. \tag{3}$$

V závislosti na vlivu apriorního rozdělení může vaše předchozí přesvědčení ovlivnit generované aposteriorní rozdělení buď silně (subjektivní nebo informativní apriorní rozdělení), nebo minimálně (objektivní nebo neinformativní apriorní rozdělení). Na apriorní hustotu se lze v jistém smyslu dívat jako na míru našeho přesvědčení, že θ nabývá právě této hodnoty. Tato víra může být podložena historickými daty či pouze statistikovým subjektivním názorem (Vávra, 2018). Bayesovu vzorci (1) se také říká zákon o inverzní pravděpodobnosti (Hebák, 2012). Pojem inverzní pravděpodobnosti se především používá v souvislosti s úsudky o neznámém modelu na základě známých dat, na rozdíl od přímých pravděpodobností, které se naopak používají pro úsudky o neznámých skutečnostech na základě známého modelu.

Bayesovská analýza dat se zabývá výlučně aposteriorním rozdělením $p(\theta|\boldsymbol{y})$. Na rozdíl od klasické analýzy dat jsou parametry považovány za náhodné veličiny mající svoje pravděpodobnostní rozdělení. Veškerá statistická inference v bayesovském pojetí vychází ze souhrnných charakteristik aposteriorního rozdělení $p(\theta|\boldsymbol{y})$ včetně bodových a intervalových odhadů. Například průměr nebo medián aposteriorního rozdělení slouží jako bodové odhady a kvantily rozdělení vedou k bayesovskému intervalu spolehlivosti, tzv. "credibility" intervalu. Podobně jako u klasické statistiky tyto intervaly mohou být obou- nebo jednostranné. Každý interval (d, h) podle (Samaniego, 2010), pro který platí

$$P(\theta_d \leq \theta \leq \theta_h) = \int_{\theta_d}^{\theta_h} p(\theta|\boldsymbol{y})d\theta = 1 - \alpha. \tag{4}$$

je $(1-\alpha)100\%$ bayesovským intervalem spolehlivosti odhadovaného parametru. Interval s vyhovujícími vlastnostmi je znám jako „*highest posteriori density*“ (HPD) interval, který pro jednovrcholová rozdělení je nejkratším, neboli má nejmenší rozdíl $(\theta_h - \theta_d)$.

3 Ilustrativní příklad použití bayesovské analýzy dat

Nechť urna obsahuje *n* míčků, z nichž n_1 je bílých a n_2 černých. Z urny se náhodně vytáhne 1 míček a poté se vrací do urny nazpět; jedná se tedy o výběr s vracením. Každý míček má stejnou pravděpodobnost vytažení rovnou $1/n$. Pravděpodobnost vytažení bílého míčku je θ , černého pak $1-\theta$. Tento proces

se bude opakovat *n*-krát. Náhodná veličina udávající počet vytažených bílých míčků se označí jako *X*. Ze všech *n* pokusů je možné vybrat přesně *k* bílých míčků právě $C(k,n)$ způsoby (kde $C(k,n)$ je počet kombinací k–té třídy spomezi n prvků). Náhodná veličina *X* podléhá binomickému rozdělení, a proto statistický model (věrohodnostní funkce) bude vyjádřen ve tvaru

$$L(\theta) = C(k,n)\,\theta^k(1-\theta)^{n-k}. \tag{5}$$

Na rozdíl od četnostní statistiky, bayesovská statistika předpokládá, že parametr θ může nabývat více hodnot s apriorní vírou (přesvědčením) o pravděpodobnostech uvedených hodnot parametru θ. Příklad apriorního rozdělení parametru udávajícího pravděpodobnost θ je uveden v tabulce 1. Pro *n = 20* a *k = 2* jsou v prvním, resp. druhém sloupci všechny možné hodnoty parametru θ, resp. jejich apriorní pravděpodobnosti (rozdělení). Na základě hodnot parametru θ v prvním a hodnot apriorního rozdělení ve druhém sloupci Tabulky 1 lze spočítat hodnotu věrohodnosti do třetího sloupce pro každou z hodnot parametru θ podle vztahu (5). Výsledné aposteriorní rozdělení parametru θ (pátý sloupec tabulky 1) je určeno pomocí Bayesovy věty (3), přičemž hodnota normalizační konstanty představuje součet hodnot věrohodnostní funkce vážených apriorními pravděpodobnostmi parametru θ (ve čtvrtém sloupci tabulky 1).

Tab. 1 Empirické aposteriorní rozdělení *(Zdroj: Dohnálek, 2016)*

Hodnoty θ	Apriorní rozdělení	Funkce věrohodnosti	Funkce věrohodnosti vážená apriórním rozdělením	Aposteriorní rozdělení
0,10	0,125	0,2852	0,0356	0,1600
0,12	0,150	0,2740	0,0411	0,1845
0,14	0,225	0,2466	0,0555	0,2491
0,16	0,225	0,2109	0,0474	0,2130
0,18	0,175	0,1730	0,0259	0,1165
0,20	0,125	0,1369	0,0171	0,0768
Celkem	**1,000**		**0,2501**	**1,0000**

Takovým způsobem dochází k úpravám apriorních pravděpodobností využitím informace získané z dat na hodnoty aposteriorních pravděpodobností. Hodnoty věrohodnostní funkce pro pozorovaná data vážené apriorním rozdělením určují rozdělení aposteriorní pravděpodobnosti, které představují zkorigované prvotní, počáteční přesvědčení (víru) o parametru θ pomocí informace ze získaných dat.

4 Funkcionality bayesovské statistiky v systému SAS

Programový systém SAS nabízí 2 způsoby analýzy dat bayesovskou metodou. Ve vybraných procedurách pro statistické modelování se bayesovská analýza volá příkazem BAYES a univerzální proceduru MCMC pro obecné statistické

modelování pomocí Bayesovy věty (SAS Institute Inc., 2018a). Statistická inference – bodové i intervalové odhady - vychází z aposteriorního rozdělení $p(\theta|\boldsymbol{y})$.

Software SAS/STAT poskytuje bayesovské možnosti v sedmi procedurách: BCHOICE, BGLIMM, FMM, GENMOD, LIFEREG, PHREG a MCMC. Procedury FMM, GENMOD, LIFEREG a PHREG umožňují bayesovské analýzy jako doplněk ke standardním četnostním analýzám (Chen, 2014). Procedury tak poskytují pohodlný přístup k bayesovskému modelování a inferenci pro modely konečných směsí rozdělení, zobecněné lineární modely, modely dožití, Coxovy regresní modely a po částech exponenciální modely. Procedura BCHOICE poskytuje Bayesovu analýzu pro diskrétní modely výběru. Procedura BGLIMM dovoluje bayesovskou analýzu pro generalizované lineární modely se smíšenými efekty.

Bayesovské modelování také nabízí alternativní modelové řešení v analýze chybějících hodnot procedurou MI, ve kterém se s chybějícími hodnotami zachází jako s neznámými parametry a podle toho se odhadují. Tyto chybějící hodnoty jsou jednoduše imputovány přijatelnými hodnotami prostřednictvím simulací aposteriorního rozdělení neúplných údajů.

Jádro procedury MCMC pro generování pseudonáhodných čísel z pravděpodobnostních rozdělení tvoří metody založené na simulaci pomocí Markovských řetězců. Podle (SAS Institute Inc., 2018b) je markovský řetězec sekvence náhodných veličin, u nichž rozdělení každého prvku závisí výlučně na hodnotě prvku předchozím. Pro účely simulace metodou Markovských řetěců jsou nejpopulárnějšími algoritmy Gibbsovské vzorkování (Gibbs Sampling) a Metropolis-Hastings algoritmus. Metropolis-Hastings algoritmus simuluje výběr z aposteriorního pravděpodobnostního rozdělení za účelem aproximace požadovaného rozdělení a provádí rozhodování o přijetí nebo o odmítnutí vygenerovaného náhodného vzorku. Pokud je nově generovaný vzorek akceptován, algoritmus generuje vzorek nový, pokud je náhodně generovaný vzorek odmítnut, simulace stávajícího výběru se opakuje. Generování náhodných vzorků končí v okamžiku dosažení konvergence k cílovému aposteriornímu rozdělení. Po dosažení konvergence se vybere náhodně z několika nasimulovaných řad a vybírá se řada podle spolehlivosti konvergence k cílovému rozdělení. Gibbsovo vzorkování (jako speciální případ předchozího algoritmu) rozděluje simultánní aposteriorní rozdělení vícerozměrné náhodné veličiny na podmíněná rozdělení pro každý parametr, ze kterých náhodně generuje řetězce a sleduje konvergenci k požadovanému aposteriornímu rozdělení.

5 Ukázky bayesovského modelování procedurou MCMC

Procedura MCMC je flexibilní, simulačně založená procedura, která je vhodná pro modelování podle široké škály bayesovských modelů. Pro úspěšné modelování je třeba zadat tvar podmíněného rozdělení pozorovaných dat za daných parametrů (statistický model) a apriorní rozdělení pro odhadované parametry. Pro modelování hierarchických modelů se vkládá také i apriorní rozdělení pro parametry náhodných efektů. Simulací procedura MCMC pak získává vzorky z požadovaného aposteriorního rozdělení, vytváří souhrnné a diagnostické statistiky a ukládá generované vzorky do výstupního datového souboru, který lze použít pro další analýzu. I když procedura MCMC podporuje celou sadu standardních rozdělení, je možné procedurou modelovat data s jakýmkoliv podmíněným rozdělením (statistickým modelem), apriorním či aposteriorním rozdělením, jsou-li jejich matematické tvary programovatelné pomocí příkazů DATA kroku. Na parametry vstupující do modelu neexistují žádná omezení, parametry do procedury mohou být zadávány buď v lineární, nebo v jakékoliv nelineární funkční podobě.

Procedura MCMC může modelovat automaticky proměnné závislé (response variable), proměnné nezávislé (kovariáty) i chybějící hodnoty. Chybějící hodnoty procedura považuje za neznámé parametry a zahrnuje chybějící hodnoty do procesu simulace. Pro každý parametr, resp. blok parametrů procedura vybírá nejvhodnější metody vzorkování. Například je-li generované rozdělení konjugované[3], vzorky jsou generovány přímo z podmíněných aposteriorních rozdělení využitím standardních číselných generátorů. V jiných případech procedura používá Metropolis-Hastings algoritmus, Gibbsovo vzorkování, případně a jiné.

Nejjednodušším příkladem bayesovského modelování procedurou MCMC je simulace vzorků z normálního (aposteriorního) rozdělení pravděpodobnosti, které se realizuje pomocí následující příkazové syntaxe:

```
data X;
     stop;
run;

ods graphics on;
ods exclude nobs;
proc mcmc data=X outpost=simout seed=23 nmc=10000 diagnostics=none
     plots=density;
     parm alpha 0;
     prior alpha ~ normal(0, sd=1);
     model general(0);
run;
```

[3] Je-li aposteriorní rozdělení členem stejné rodiny rozdělení jako rozdělení apriorní, pak obě tato rozdělení jsou konjugovaná, neboli obecně, jsou stejného typu.

Vstupem do procedury MCMC je prázdný datový soubor nazvaný X. Vstupní datový soubor X nemusí obsahovat datové věty (DATA krok s příkazem STOP), protože úlohou procedury je generování nových vzorků. Příkazem NMC=*10 000* se nastaví počet simulovaných vzorků, příkaz PARM slouží k označení parametru, jehož aposteriorní rozdělení bude procedura simulovat včetně přiřazení jeho počáteční hodnoty. Příkazem PLOTS se určí, které grafy má procedura generovat. Apriorní rozdělení simulovaného parametru se nastaví příkazem PRIOR, když v tomto případě se jedná o parametr ALPHA s aposteriorním normálním rozdělením $N(0,1)$. Příkaz MODEL, který vkládá do procedury tvar podmíněného rozdělení pozorovaných dat za daných parametrů (statistický model), musí mít každá příkazová syntaxe pro proceduru MCMC. Pro generování parametru ALPHA, kde neexistuje závislá proměnná, nabývá příkaz MODEL hodnoty GENERAL(0). Simulovaná data se ukládají do výstupního souboru SIMOUT (nastavení parametru OUTPOST). Vzhledem k jednoduchosti úlohy, není potřebná jakákoliv diagnostika - např. počet iterací do stacionarity MCMC řetězců, rychlosti konvergence generování řetězců, atd.

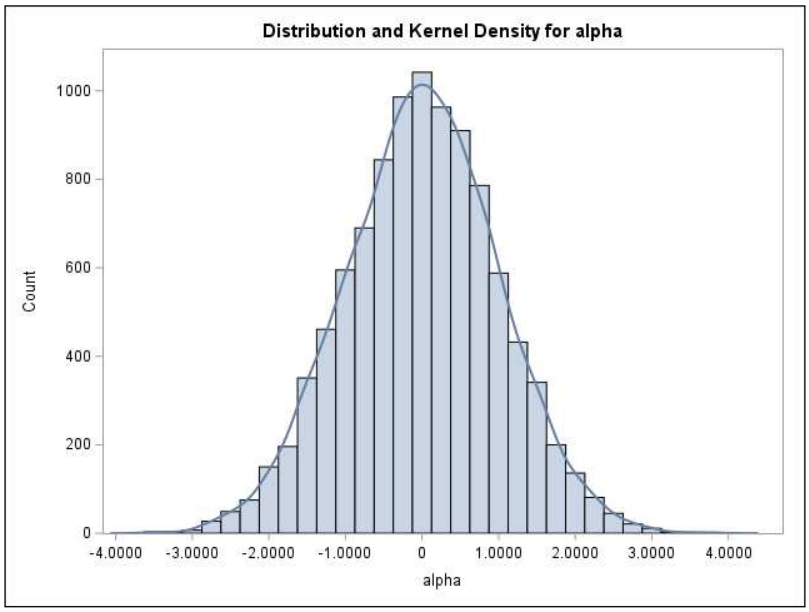
Po provedené simulaci je výstupem činnosti procedury MCMC soubor SIMOUT obsahující *10 000* vygenerovaných hodnot parametru ALPHA, jehož rozdělení pravděpodobnosti odpovídá požadovanému normálnímu rozdělení $N(0,1)$. Generování hodnot procedurou bylo implicitně realizováno přímo z podmíněného aposteriorního rozdělení (normálního) využitím standardních číselných generátorů, protože apriorní i podmíněné aposteriorní rozdělení parametru ALPHA patří do stejné rodiny rozdělení, a tedy jsou konjugované. Výstupní report v systému SAS obsahuje tabulku odhadovaných parametrů „Parameters“ s názvem použité metody vzorkování, názvem a počáteční hodnotou parametru a popisem apriorního rozdělení parametru. Sumární statistiky a inference (vycházející výlučně z aposteriorního rozdělení) jsou zachyceny v tabulce 2 „Posterior Summaries and Intervals“, kde je uveden název parametru, jeho střední hodnota, standardní odchylka a 95% HPD interval.

Tab. 2 Sumární statistiky a inference simulovaného parametru *(Zdroj: vlastní zpracování)*

Parameters				
Block	Parameter	Sampling Method	Initial Value	Prior Distribution
1	alpha	Direct	0	normal(0, sd=1)

Posterior Summaries and Intervals					
Parameter	N	Mean	Standard Deviation	95% HPD Interval	
alpha	10000	0.00195	0.9949	-1.9664	1.9302

Na obrázku 1 byl zaznamenán průběh rozdělení parametru ALPHA.



Obr. 1 Aposteriorní rozdělení parametru ALPHA *(Zdroj: vlastní zpracování)*

Následující příklad použití bayesovského modelování se bude týkat problému s chybějícími údaji. Poprvé se uvedený příklad objevil v diskusním příspěvku Murraya (1977), který představuje datový soubor s dvojrozměrným normálním rozdělením s chybějícími daty. Jedná se o uměle vytvořený problém, kterým však později zabývali i mnozí jeho následovníci jako Tanner & Wong (1987) a Tan et al. (2010).

Tab. 3 Struktura souboru s normálním rozdělením s chybějícími daty *(Zdroj: Murray, 1977)*

Proměnná	Hodnoty dat s dvojrozměrným normálním rozdělením											
x_1	1	1	-1	-1	2	2	-2	-2	.	.	.	.
x_2	1	-1	1	-1	.	.	.	.	2	2	-2	-2

Soubor dat v tabulke 3 k analýze chybějících hodnot se vytvoří pomocí příkazů:

```
data binorm;
     input x1 x2 @@;
     datalines;
1 1 1 -1 -1 1 -1 -1 2 . 2 . -2 . -2 . . 2 . 2 . -2 . -2;
run;
```

Soubor dat (nazvaný BINORM) se skládá z 12 pozorování z dvojrozměrného normálního rozdělení. U obou komponent tohoto vícerozměrného rozdělení proměnných chybí hodnoty. Strukturu souboru s chybějícími daty zobrazuje tabulka 3. Dvojrozměrné normální rozdělení má obě komponenty se stejnými nulovými průměry μ, s různými rozptyly σ_1^2 a σ_2^2 a korelačním koeficientem ρ.

Tyto předpoklady se promítají do následující kovarianční matice Σ:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{6}$$

Apriorní rozdělení kovarianční matice Σ je $\pi(\Sigma)$ a odpovídá podle Boxa & Tiaa (1973) inverznímu Wishartovu rozdělení s determinantem $\sigma_1^2\sigma_2^2(1-\rho^2)$, tj.

$$\pi(\Sigma) \propto [\Sigma]^{-\frac{p+1}{2}} = \left(\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^{-(p+1)}. \tag{7}$$

Při analýze kompletních pozorování, kdy se do úvahy berou pouze úplná pozorování, by se bayesovským modelováním vytvořil poměrně jednoduchý aposteriorní odhad: 2 datové dvojice ((1, 1) a (-1, -1)) mají korelaci rovnou 1 a 2 dvojice dat ((1, -1) a (-1, 1)) mají korelaci -1. V případě, že se budou chybějící hodnoty považovat za zcela náhodně chybějící, lze podle Chena (2009) považovat nepozorované hodnoty buď (téměř) dokonale pozitivně nebo dokonale negativně korelující s pozorovanými hodnotami 2 a –2 a že chybějící hodnoty mohou nabývat pouze hodnot, které se od nich příliš neliší. Navíc vyřazením částečně pozorovaných hodnot se ztrácí podstatná část při odhadu rozptylu obou komponent dvojrozměrného normálního rozdělení. Bayesovské modelování chybějících hodnot dvojrozměrného normálního rozdělení vykonávají následující příkazy:

```
proc mcmc data=binorm nmc=20000 seed=17 outpost=postout
     diag=none plots=none monitor=(rho sig1 sig2);
     array x[2] x1 x2;
     array mu[2] (0 0);
     array sigma[2, 2];
     parms rho 0.5 / slice;
     parms sig1 1 sig2 1;
     if (sig1 > 0 and sig2 > 0) then
          lprior = -3 * log(sig1 * sig2 * sqrt(1-rho*rho));
     else lprior = .;
     prior rho sig1 sig2 ~ general(lprior);
     sigma[1,1] = sig1*sig1;
     sigma[1,2] = rho * sig1 * sig2;
     sigma[2,1] = sigma[1,2];
     sigma[2,2] = sig2*sig2;
     model x ~ mvn(mu, sigma);
run;
```

Příkazy PARMS specifikují tři modelované parametry modelu. Volba SLICE v prvním příkazu PARMS vybere vzorkovací algoritmus, kterým se generuje parametr ρ. I když výpočetně náročnější (Neal, 2003), tento typ algoritmu je velmi vhodný pro simulaci dat z nenormálních rozdělení (kam náleží i rozdělení korelačního koeficientu ρ). Příkazy IF-ELSE zajišťují, aby simulované hodnoty parametrů směrodatné odchylky σ_1 a σ_2 byly zahrnovány do aposteriorních odhadů kovarianční matice pouze při své kladné hodnotě. Příkazem LPRIOR a příkazem PRIOR se do procedury zadávají simultánní apriorní rozdělení

parametrů σ_1, σ_2 a ρ na logaritmické stupnici[4]. Prvky kovarianční matice Σ jsou konstruovány použitím dalších programových příkazů jako funkcí parametrů modelu (v syntaxi příkazů označené jako `sigma[1,1]` až `sigma[2,2]`). Příkazem MODEL se nastavuje specifikace požadované funkce věrohodnosti nutná pro simulaci a následnou aposteriorní inferenci.

Po provedené simulaci na základě výše uvedené syntaxe je výstupem činnosti procedury MCMC soubor POSTOUT obsahující *20 000* vygenerovaných chybějících hodnot v obou proměnných x_1 a x_2 ve dvojrozměrném normálním rozdělení a vygenerované rozdělení parametrů σ_1, σ_2 a ρ podle zadání. Analýza chybějících hodnot je ilustrována v části „Missing Data Information Table" tabulky 4. Proměnná x_1 obsahuje *4* chybějící hodnoty, proměnná x_2 obsahuje také *4* chybějící hodnoty. Tabulka 4 zaznamenává indexy chybějících hodnot za jednotlivé proměnné. Aposteriorní inference o odhadovaných parametrech σ_1, σ_2 a ρ je zobrazena v části „Posterior Summaries and Intervals" tabulky 4.

Tab. 4 Sumární statistiky a inference dvojrozměrného rozdělení *(Zdroj: vlastní zpracování)*

Missing Data Information Table			
Variable	Number of Missing Obs	Observation Indices	Sampling Method
x1	4	9 10 11 12	Direct
x2	4	5 6 7 8	Direct

Posterior Summaries and Intervals					
Parameter	N	Mean	Standard Deviation	95% HPD Interval	
rho	20000	-0.00184	0.5582	-0.8591	0.8519
sig1	20000	1.6555	0.4458	0.9828	2.6014
sig2	20000	1.6306	0.4444	0.9478	2.4860

Průběh aposteriorního rozdělení parametru ρ ilustruje obrázek 2. Průměrný odhad ρ se blíží 0, se směrodatnou odchylkou 0,56. Tento odhad je nepřesný, protože aposteriorní rozdělení není unimodální. Obrázek 3 ilustruje aposteriorní rozdělení odhadovaných parametrů σ_1, σ_2.

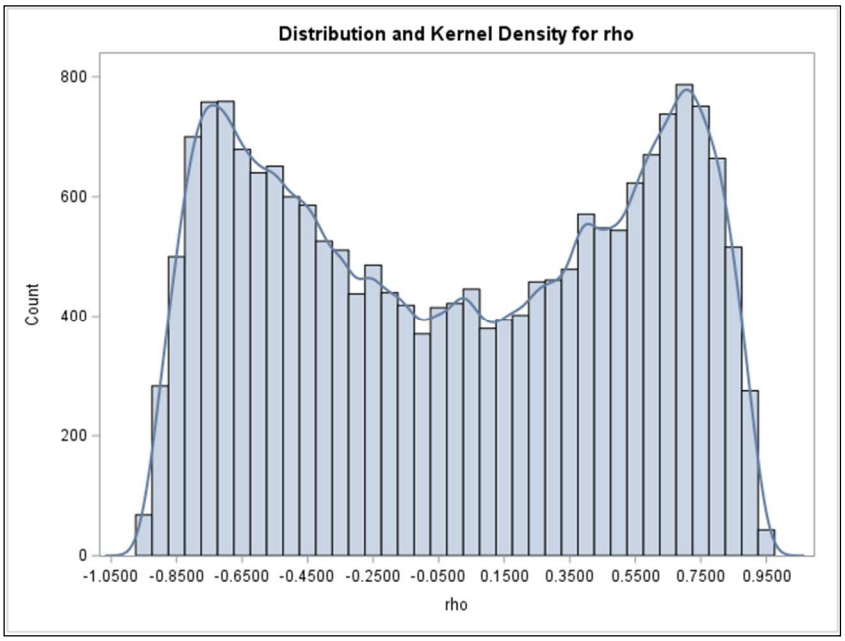
Podobně jako v předchozím případě jsou apriorní i podmíněné aposteriorní rozdělení chybějících údajů konjugované. Podmíněná aposteriorní rozdělení chybějících hodnot $\pi(x_1|.)$ a $\pi(x_2|.)$ pro jednotlivé proměnné x_1 a x_2 jsou jednorozměrná normální rozdělení a jsou dána výrazem:

$$\pi(x_1|\rho,\sigma_1,\sigma_2,x_2)\sim N\left(\rho\frac{\sigma_1}{\sigma_2}x_2,\sigma_1^2(1-\rho^2)\right). \quad (8)$$

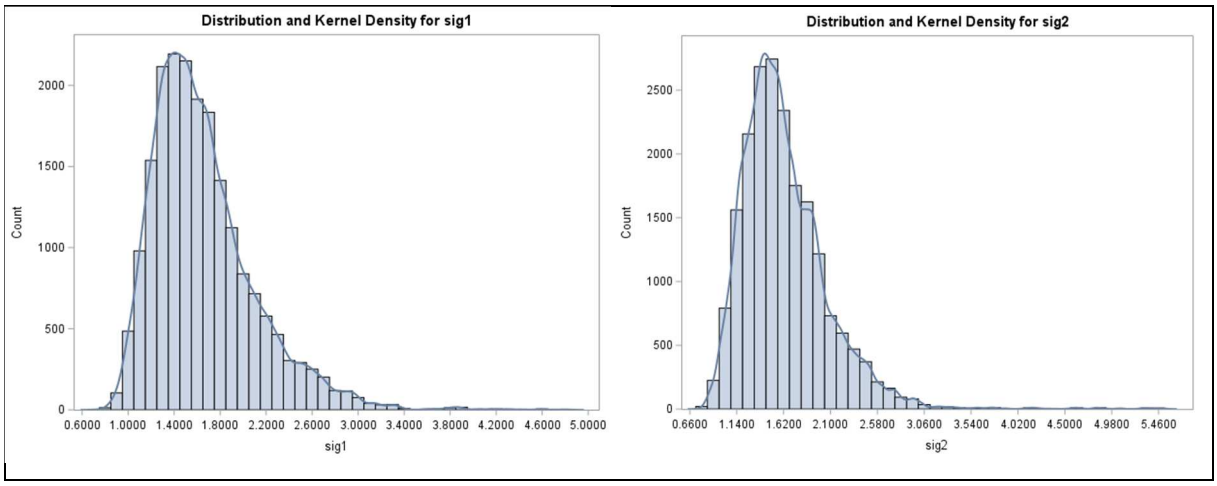
$$\pi(x_2|\rho,\sigma_1,\sigma_2,x_1)\sim N\left(\rho\frac{\sigma_2}{\sigma_1}x_1,\sigma_2^2(1-\rho^2)\right). \quad (9)$$

Podmíněná aposteriorní rozdělení chybějících hodnot (8) a (9) pro jednotlivé proměnné x_1 a x_2 jsou unimodální. Výstup zde ale není graficky zobrazen.

[4] Použitím výpočtů na logaritmické stupnici se dosahuje větší stabilita a rychlejší konvergence v číselném výpočtu.



Obr. 2 Aposteriorní rozdělení parametru ρ *(Zdroj: vlastní zpracování)*



Obr. 3 Aposteriorní rozdělení parametrů σ_1, σ_2 *(Zdroj: vlastní zpracování)*

6 Závěr

Při řešení analytických problémů by výzkumníci měli zvažovat alternativní statistické přístupy. I když bayesovské metody ve srovnání s klasickými postupy vyžadují hlubší znalosti teorie současně s určitými výpočetními dovednostmi, jejich správná aplikace ve výzkumné činnosti se rozhodně vyplatí. Bayesovská statistika poskytuje obecnější přístup pro statistickou inferenci. Na rozdíl od četnostní statistiky bayesovský přístup považuje odhadované parametry jako náhodné, jejichž rozdělení je možné upřesňovat informacemi získaných z dat.

Navíc v rámci bayesovského přístupu jsou znalosti zkoumaného fenoménu, které lze efektivně využít, zpravidla vždy dostupné.

V současné době je aplikace bayesovských metod mnohem přístupnější, protože už existují vhodné volně přístupné programy i placené programové systémy, například právě SAS.

7 Literatura

Box, G. E. P.,& Tiao, G. C. (1973). Bayesian inference in statistical analysis. New York: Wiley, 1973.

Dohnálek, R. (2016). Rozdělení pravděpodobnosti odvozená z urnových modelů. Bakalářská práce obhájena na Přirodovědecké univerzity Masarykovy univerzity v Brně.

Hebák, P. (2012). Srovnání klasické a bayesovské pravděpodobnosti a statistiky (1). Acta Oeconomica Pragensia 20(1):69-87.

Chen, F. (2014). Practical Bayesian computation using SAS. ASA Conference on Statistical Practices. February 20, 2014.

Murray, G. D. (1977). Contribution to discussion of paper by Dempster, Laird, and Rubin. Journal of the Royal Statistical Society, Series B, 39, 27–28.

Neal, R. M. (2003). Slice sampling. Technical report No. 2005. Department of Statistics and Department of Computer Science. University of Toronto, Toronto, Ontario, Canada, https://www.cs.toronto.edu/~radford/ftp/slc-samp.pdf.

Samaniego, F. J. (2010). A comparison of the Bayesian and frequentist approaches to estimation. New York: Springer, 2010.

SAS Institute Inc. (2018a). SAS/STAT® 15.1 user's guide. Chapter 7: Introduction to Bayesian analysis procedures. Cary, NC: SAS Institute Inc., 2018, s.129.

SAS Institute Inc. (2018b). SAS/STAT® 15.1 user's guide. Chapter 77: The MCMC procedure. Cary, NC: SAS Institute Inc., 2018, s.6011.

Tan, M. T., Tian, G. L., & Ng, K. W. (2010). Bayesian missing data problems: EM, data augmentation, and non-iterative computation. New York: Chapman & Hall/CRC, 2010.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82, 528–540.

Vávra, Jan. (2018). Bayesovská faktorová analýza. Univerzita Karlova. Matematicko-fyzikální fakulta. Praha, 2018.

FORUM STATISTICUM SLOVACUM Vol XVIII (2022) No 2 pp 13-34

Viacrozmerná analýza vybraných ukazovateľov udržateľného rozvoja v krajinách EÚ
Multivariate analysis of selected indicators of sustainable development in EU countries

Zuzana Moscoso Borsík, Iveta Stankovičová
Univerzita Komenského v Bratislave, Fakulta managementu, Odbojárov 10, 820 05 Bratislava, Slovenská republika
Comenius University in Bratislava, Faculty of Management, Odbojárov 10, 820 05 Bratislava, Slovak Republic
borsikova6@uniba.sk, iveta.stankovicova@fm.uniba.sk

Abstrakt: *Predkladaný článok sa zaoberá hodnotením 27 krajín Európskej únie na základe využitia viacrozmernej analýzy vybraných indikátorov. Indikátory opisujú plnenie strategického cieľa udržateľného rozvoja č. 10: Znížiť nerovnosť v rámci krajín a medzi nimi. Cieľom práce je vybrať relevantné indikátory a analyzovať ich pomocou viacrozmerných štatistických metód: faktorová analýza a zhluková analýza. Bolo vybratých 5 indikátorov, konkrétne 3 súvisiace s oblasťou distribúcie príjmov a 2 súvisiace s úrovňou zamestnanosti v krajine. Analyzované boli údaje za dve obdobia, roky 2010 a 2020. Krajiny boli zaradené do skupín podľa charakteristických čŕt, ktoré vykazujú. V oboch sledovaných rokoch bolo vytvorených 5 zhlukov. Charakteristické črty každého zhluku boli opísané pomocou mier popisnej štatistiky vstupných indikátorov. Na záver boli vymenované krajiny, ktoré dosiahli najlepšie a najhoršie výsledky, ako aj pozícia Slovenska.*

Abstract: *The presented article deals with the evaluation of 27 European Union countries based on the use of multidimensional analysis of selected indicators. The indicators describe the fulfilment of the strategic goal of sustainable development no. 10: Reduce inequality within and between countries. The goal of the thesis is to select relevant indicators and analyse them using multivariate statistical methods: factor analysis and cluster analysis. Five indicators were selected, namely three related to the area of income distribution and two related to the level of employment in the country. Data were analysed for two periods, the years 2010 and 2020. The countries were grouped according to the characteristic features they show. In both monitored years, five clusters were created. The characteristic features of each cluster were described using descriptive statistics measures of the input indicators. At the end, the countries that achieved the best and worst results were listed, as well as the position of Slovakia.*

Kľúčové slová: *krajiny EÚ, udržateľný rozvoj, distribúcia príjmov, zamestnanosť, faktorová analýza, zhluková analýza.*

Key words: *EU countries, sustainable development, income distribution, employment, factor analysis, cluster analysis.*

1 Úvod

Od 60. rokov 20. storočia sa rozširuje poznanie, že neobmedzený, resp. nekontrolovaný rast akéhokoľvek typu (ľudskej populácie, spotreby, znečistenia a pod.) nie je udržateľný v prostredí reálne existujúcich obmedzených zdrojov.

Model industriálnej civilizácie je preto potrebné nahradiť iným modelom, trvalejšou a spravodlivejšou koncepciou. Za takýto koncept sa v súčasnosti považuje udržateľný rozvoj (*angl. výraz sustainable development – SD*) ako možné východisko riešenia nepriaznivých dôsledkov globálnych trendov vývoja spoločnosti a ich negatívnych vplyvov na prírodu. Zdôrazňuje sa potreba založiť túto koncepciu na zdravých ekosystémoch, silnej ekonomike a fungujúcej sociálnej sfére.

Trvalo udržateľný rozvoj teda znamená uspokojovanie potrieb súčasnej generácie, hlavne kvality ľudského života, prežívajúceho v rámci únosnosti podporných ekosystémov bez toho, aby ohrozoval schopnosť budúcich generácií uspokojovať svoje potreby.

V súčasnosti sa problematika trvalej udržateľnosti rozšírila do všetkých odvetví a jej vnímanie je komplexné, pretože je potrebné zohľadniť ekonomické, sociálne aj kultúrne aspekty. Medzinárodná únia na ochranu prírody a prírodných zdrojov v roku 1980 publikovala správu „Stratégia ochrany sveta“, ktorá ponúka jednu z prvých definícií trvalo udržateľného rozvoja. Za trvalo udržateľný rozvoj vymedzuje taký rozvoj, ktorý berie do úvahy sociálne a ekologické faktory, ekonomické aspekty, živé a neživé zdroje a v neposlednom rade zohľadňuje výhody a nevýhody opatrení z krátkodobého aj dlhodobého hľadiska.

Na konferencii Organizácie spojených národov (OSN), ktorá sa konala v roku 2002 v meste Rio de Janeiro, sa členské štáty rozhodli začať proces na vypracovanie súboru cieľov trvalo udržateľného rozvoja. Dokument je známy pod anglickým názvom *Sustainable Development Goals* (ďalej len ako SDG, resp. SDGs) a vychádza z rozvojových cieľov tisícročia *Millennium Development Goals*. Dohodlo sa na integrovaní týchto cieľov s rozvojovou agendou po roku 2015. Celkovo bolo formulovaných 17 SDGs cieľov, ktoré pokrývajú ekonomické, ekologické, sociálne a aj inštitucionálne aspekty.

Na efektívne vyhodnocovanie progresu v rámci dosahovania vytýčených cieľov bolo potrebné implementovať kontrolný mechanizmus. Na úrovni Európskej únie (EÚ) je na ročnej báze publikovaná správa, ktorá sumarizuje dosiahnuté úspechy. Správa podáva informáciu o plnení SDGs v celej EÚ, ale aj v jednotlivých členských krajinách.

V článku sa venujeme viacrozmernej analýze piatich vybraných ukazovateľov cieľa 10 udržateľného rozvoja v krajinách EÚ (SDG_10), ktoré sme získali z databázy Eurostatu SDI dostupnej na https://ec.europa.eu/eurostat/web/sdi /database. Zamerali sme sa na porovnanie rokov 2010 a 2020.

2 Použité dáta a metodológia

SDG cieľ 10: *Zníženie nerovností* má za úlohu monitorovať znižovanie nerovností v rámci krajín a aj medzi nim. Dotýka sa oblastí príjmov domácností, začlenenia všetkých obyvateľov do sociálneho, ekonomického a politického diania, ale aj oblasti medzištátnej migrácie. Jedná sa teda o sociálno-ekonomické sféry, z pohľadu rozdielov medzi sociálnymi vrstvami, či regiónmi, nie z pohľadu rovnosti medzi pohlaviami (problematike rodovej rovnosti sa venuje SDG cieľ 5).

Hodnoty vybratých ukazovateľov (indikátorov) súvisiacich s cieľmi udržateľného rozvoja Agendy 2030, sú vyhodnocované na ročných bázach. Avšak pri niektorých ukazovateľoch údaje za rok 2021 ešte neboli dostupné pre všetky krajiny EÚ. Ucelený súbor dát bol dostupný za rok 2020 a pre porovnanie vývoja indikátorov za posledných 10 rokov sme vybrali aj údaje z roku 2010. Do analýzy boli zaradené všetky krajiny, ktoré sú aktuálne členmi EÚ (Poznámka: Spojené kráľovstvo do súboru nebolo zaradené, nakoľko vystúpilo z EÚ 31. januára 2020.) Celkovo tak do štatistického súboru bolo zaradených 27 krajín EÚ (EU27), čiže štatistických jednotiek.

Vzhľadom na rozsah štatistického súboru (n = 27) sme pre viacrozmernú analýzu cieľa SDG 10 vybrali len 5 ukazovateľov z celkového počtu 8. Vybraté indikátory súvisia s nerovnosťou príjmov a s úrovňou zamestnanosti.

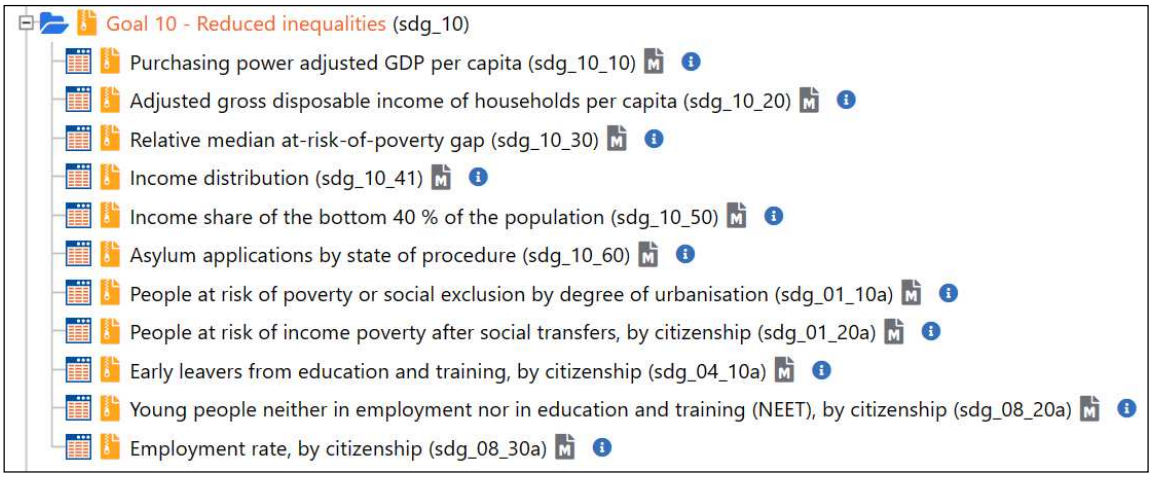
Cieľom analýzy je pomocou viacrozmerných metód zhodnotiť situáciu v krajinách EÚ v roku 2010 a 2020 v oblasti SDG cieľa 10. Výsledky porovnať a zistiť aký vývoj krajiny EÚ zaznamenali za 10 rokov v skúmanej oblasti.

Na dosiahnutie vytýčeného cieľa boli použité viacrozmerné metódy, konkrétne faktorová analýza (FA) a zhluková analýza (CA). Použité metódy sú všeobecne známe a teoreticky podrobne popísané v knihe Stankovičová & Vojtková (2007), resp. v druhom vydaní z roku 2019.

Všetky výpočty boli uskutočnené v softvéri SAS Viya dostupnom voľne a bezplatne pre pracovníkov školstva v SAS cloude na https://www.sas.com/en_us/software/viya-for-learners.html.

3 Vybrané ukazovatele SDG cieľa 10

Jednotlivé ukazovatele boli zvolené na základe oficiálneho dokumentu Globálny indikátorový rámec pre ciele trvalo udržateľného rozvoja a ciele Agendy 2030 pre trvalo udržateľný rozvoj (SDG). Pre monitorovanie SDG cieľa 10 je stanovených celkovo 11 ukazovateľov. Uvádza ich obrázok 1.



Obr. 1 Ukazovatele pre monitorovanie SDG cieľa 10
(Zdroj: https://ec.europa.eu/eurostat/web/sdi/database)

Definície a popisy vybraných ukazovateľov sme získali z metadát dostupných na stránke Eurostatu a sú nasledovné:

1. **Relatívna mediánová medzera rizika chudoby** (*relative median at-risk-of-poverty gap* – **RCHu**) je ukazovateľ vypočítaný ako rozdiel medzi mediánom ekvivalentného celkového čistého príjmu osôb pod prahovou hodnotou rizika chudoby a prahom rizika chudoby vyjadrený ako percentuálny podiel prahovej hodnoty rizika chudoby. Hraničný bod je 60% mediánu ekvivalentného príjmu v krajine (nie v celej EÚ).
2. **Rozdelenie príjmov** (*income distribution* – **RP**) predstavuje miera nerovnosti rozdelenia príjmov. Ukazovateľ je vypočítaný ako pomer celkového príjmu 20% obyvateľstva s najvyšším príjmom (horný kvintil) k príjmu 20% obyvateľstva s najnižším príjmom (dolný kvintil).
3. **Podiel príjmov spodných 40% populácie** (*income share of the bottom 40% of the population* – **PP40**) reprezentuje podiel príjmu, ktorý dostáva spodných 40% populácie. Použitým základom ukazovateľa je celkový disponibilný príjem domácností.
4. **Miera zamestnanosti** (*employment rate* – **MZ**) tvorí podiel zamestnanej populácie vo veku 20 až 64 rokov, na celkovej populácii v produktívnom veku (20 až 64 rokov). Zamestnanci sú definovaní ako všetky osoby, ktoré počas referenčného týždňa odpracovali aspoň jednu hodinu za mzdu alebo zisk alebo boli dočasne neprítomné v takejto práci.
5. **Podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave** (*young people neither in employment nor in* education *and training* – **PMl**) reprezentuje podiel populácie vo veku 15 až 29 rokov, ktorá nie je zamestnaná a nezúčastňuje sa vzdelávania alebo odbornej prípravy na celkovej populácii v tomto veku. Čitateľ ukazovateľa sa vzťahuje na osoby, ktoré spĺňajú tieto dve podmienky: (a) nie sú zamestnané (t. j. nezamestnané

alebo neaktívne podľa definície Medzinárodnej organizácie práce) a (b) nezískali žiadne vzdelanie alebo odbornú prípravu (t. j. formálne ani neformálne) počas štyroch týždňov pred výberovým zisťovaním pracovných síl (LFS). Menovateľ zahŕňa celkovú populáciu vo veku 15 až 29 rokov (okrem tých, ktorí neodpovedali na otázky týkajúce sa „účasti na riadnom (formálnom) vzdelávaní a odbornej príprave“).

Vybrané ukazovatele sú typu „+“, ale aj typu „-“.

- Pri type „+“ žiadúce sú čo najvyššie hodnoty ukazovateľa. Do tejto skupiny patria dva vybrané indikátory: podiel príjmov spodných 40% populácie (PP40) a miera zamestnanosti (MZ).
- Pri type „-“ žiadúce sú čo najnižšie hodnoty ukazovateľa. Do tejto skupiny patria tri vybrané indikátory: relatívna mediánová medzera rizika chudoby (RCHu), rozdelenie príjmov (RP) a podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl).

Pre niektoré vybrané kvantitatívne ukazovatele sú v dokumente Agendy 2030 pre trvalo udržateľný rozvoj (SDG) stanovené aj konkrétne cieľové hodnoty pre rok 2030 (k dispozícii napr. na https://ec.europa.eu/eurostat/cache/website/sdg/sdg_key/sdg_key_2022/img/target_indicators_en.pdf). Napríklad pre mieru zamestnanosti (MZ) je cieľovou hodnotou 78% a pre podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl) je to 9%. Pre ukazovatele nerovnosti v príjmoch je cieľ vyjadrený súhrne tak, že je potrebné znížiť počet ľudí v ohrození chudobou a sociálnym vylúčením o 15 miliónov osôb celkovo v krajinách EU.

4 Popis ukazovateľov

V nasledujúcich tabuľkách 1 a 2 postupne uvádzame hodnoty vybratých piatich ukazovateľov pre 27 krajín EÚ (EU27) v roku 2010 a v roku 2020.

Podľa tabuľky 1 roku 2010 najviac extrémnych hodnôt mala Litva. Dosiahla najhoršie výsledky indikátora relatívna mediánová medzera rizika chudoby (RCHu, 32,6%), ktorá v porovnaní s rokom 2009 vzrástla až o 8,8 percentuálneho bodu (p. b.). Aj indikátor rozdelenie príjmov (RP) dosiahol v Litve najvyššiu hodnotu (7,35), pričom oproti roku 2009 mierne vzrástol. Najhorší výsledok mala Litva aj pre podiel príjmov spodných 40% populácie (PP40, 17,7%).

Ukazovateľ relatívna mediánová medzera rizika chudoby (RCHu) dosahuje najlepší výsledok vo Fínsku (13,8%), čo o 8 p. b. nižšia hodnota ako priemer sledovaných krajín. Za ukazovateľ rozdelenie príjmov (RP) najlepší výsledok dosiahlo Maďarsko (3,41). Podiel príjmov spodných 40% populácie (PP40) dosiahol najvyššiu hodnotu v Slovinsku (24,9%), čiže najlepší výsledok v roku 2010 a minimálna hodnota bola v Litve (17,7%), čiže najhorší výsledok.

Miera zamestnanosti (MZ) bola najvyššia vo Švédsku (77,5%) a naopak najnižšia v Rumunsku (len 56,3%). Ukazovateľ podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl) bol najlepší v Holandsku (len 6,1%), čo predstavuje viac ako dvojnásobne nižšiu mieru ako bol priemer krajín EU27. V porovnaní s Bulharskom, pri ktorom registrujeme najhorší výsledok (až 23,5%), sa jedná o takmer štvornásobne nižšiu hodnotu.

Tab. 2 Hodnoty vybratých ukazovateľov, rok 2010 *(Zdroj: vlastné spracovanie na základe dát z Eurostatu)*

Krajina	RCHu	RP	MZ	PMl	PP40
Belgicko	18,00	3,92	67,60	13,00	23,20
Bulharsko	29,60	5,86	64,10	23,50	19,50
Česko	21,10	3,47	70,40	12,90	24,90
Chorvátsko	27,60	5,54	62,10	17,60	20,00
Cyprus	18,00	4,54	75,00	12,90	21,90
Dánsko	21,60	4,41	74,90	8,20	23,00
Estónsko	23,20	5,01	67,60	18,10	20,60
Fínsko	13,80	3,61	71,90	10,50	24,20
Francúzsko	19,50	4,43	69,60	14,80	22,20
Grécko	23,40	5,61	63,50	18,60	19,80
Holandsko	16,20	3,65	77,00	6,10	24,20
Írsko	15,50	4,70	65,50	21,70	21,20
Litva	32,60	7,35	64,30	17,00	17,70
Lotyšsko	28,90	6,84	63,90	20,70	17,90
Luxembursko	18,60	4,10	70,70	6,10	22,50
Maďarsko	16,50	3,41	62,20	17,70	24,80
Malta	17,30	4,33	60,10	12,20	22,20
Nemecko	20,70	4,49	74,00	10,80	22,00
Poľsko	22,20	4,98	62,20	14,80	20,90
Portugalsko	22,70	5,56	66,90	13,60	19,70
Rakúsko	21,80	4,34	73,90	9,10	22,60
Rumunsko	31,30	6,11	56,30	18,90	19,00
Slovensko	25,70	3,80	66,50	19,00	23,90
Slovinsko	20,20	3,42	69,70	9,40	24,90
Španielsko	26,80	6,16	62,80	20,00	19,20
Švédsko	19,90	3,85	77,50	8,30	23,60
Taliansko	24,80	5,38	60,50	22,00	20,20
Priemer	**22,13**	**4,77**	**67,43**	**14,72**	**21,70**
Štandardná odchýlka	**4,96**	**1,07**	**5,63**	**5,08**	**2,15**

Poznámka: Zelené podfarbenie označuje minimálnu hodnotu a červené maximálnu hodnotu ukazovateľa v stĺpci.

Podľa tabuľky 2 v roku 2020 najviac extrémnych hodnôt dosiahli 3 krajiny: Bulharsko, Holandsko a Slovensko. V Holandsku a Slovensku boli dosiahnuté extrémne hodnoty v pozitívnom zmysle. Holandsko vykazuje najvyššiu hodnotu miery zamestnanosti (MZ, 80,8%). Ukazovateľ podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl, 5,7%) je najnižší a je už teraz pod vytýčeným cieľom (9%) pre rok 2030.

Slovensko vyniká v oblasti distribúcie príjmov. V rozdelení príjmov (RP) sme dosiahli minimálnu hodnotu (len 3,03) a ukazovateľ podiel príjmov spodných

40% populácie (PP40) dosiahol zase maximálnu hodnotu (až 26,1%) medzi krajinami EU27 v roku 2020, čo sú najlepšie výsledky.

Naopak, Bulharsko dosahuje extrémne hodnoty v negatívnom zmysle. Oblasť, v ktorej zaostáva, je distribúcia príjmov. Rozdelenie príjmov (RP) dosiahlo maximálnu hodnotu medzi krajinami (až 8,01) a podiel príjmov spodných 40% populácie (PP40) bol minimálny, len 16,6%.

Tab. 3 Hodnoty vybratých ukazovateľov, rok 2020 *(Zdroj: vlastné spracovanie na základe dát z Eurostatu)*

Krajina	RCHu	RP	MZ	PMI	PP40
Belgicko	16,20	3,65	69,70	12,00	23,80
Bulharsko	28,30	8,01	72,70	18,10	16,60
Česko	14,80	3,34	79,70	11,00	24,90
Chorvátsko	28,00	4,61	66,90	14,60	21,80
Cyprus	16,90	4,31	74,90	15,30	22,20
Dánsko	19,30	4,00	77,80	10,20	23,40
Estónsko	21,80	5,03	79,10	11,90	20,40
Fínsko	13,70	3,72	75,50	10,30	23,70
Francúzsko	20,70	4,42	72,10	14,00	22,30
Grécko	27,30	5,23	58,30	18,70	20,50
Holandsko	17,60	4,15	80,80	5,70	22,90
Írsko	15,20	4,07	72,10	14,20	22,70
Litva	23,20	6,14	76,70	13,00	18,70
Lotyšsko	28,60	6,27	76,90	11,90	18,80
Luxembursko	17,70	4,99	72,10	7,70	20,60
Maďarsko	27,90	4,16	77,50	14,70	22,60
Malta	16,00	4,69	77,30	9,50	21,60
Nemecko	24,50	4,87	78,20	9,10	21,30
Poľsko	21,00	4,07	72,70	12,90	22,90
Portugalsko	24,40	4,99	74,20	11,00	21,00
Rakúsko	22,70	4,11	74,80	9,50	23,20
Rumunsko	31,90	6,62	65,20	16,60	18,30
Slovensko	23,70	3,03	74,60	15,20	26,10
Slovinsko	16,90	3,32	74,80	9,20	25,00
Španielsko	29,40	5,77	65,70	17,30	19,50
Švédsko	19,90	4,12	80,10	7,20	22,80
Taliansko	27,60	5,75	61,90	23,30	19,70
Priemer	**22,04**	**4,72**	**73,42**	**12,74**	**21,75**
Štandardná odchýlka	**5,28**	**1,13**	**5,65**	**3,96**	**2,23**

Poznámka: Zelené podfarbenie označuje minimálnu hodnotu a červené maximálnu hodnotu ukazovateľa v stĺpci.

Pri ukazovateli relatívna mediánová medzera rizika chudoby (RCHu) pozorujeme opäť najnižšiu hodnotu vo Fínsku 13,8%. Naopak najvyššiu hodnotu dosahuje Rumunsko (až 31,9%), čo je o 9,86 p. b. vyššia hodnota ako priemer sledovaných krajín. Ako bolo uvedené vyššie, indikátor rozdelenie príjmov (RP) dosahuje najlepší výsledok na Slovensku (3,03) a naopak najhorší v Bulharsku (8.01). Pri indikátore podiel príjmov spodných 40% populácie (PP40), pozorujeme najvyššiu a teda najpriaznivejšiu hodnotu na Slovensku (26,1%) a naopak najhorší podiel vykazuje Bulharsko (len 16,6%).

Ukazovateľ miera zamestnanosti (MZ) dosahuje najvyššiu hodnotu v Holandsku (80,8%) a najnižšiu evidujeme v Grécku (58,3%), čo je o 15,12 p. b. nižšia hodnota ako priemer analyzovaných krajín. Ukazovateľ podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl) dosahuje najnižšiu mieru v Holandsku (5,7%) a opačne najhorší výsledok malo Taliansko s hodnotou 23,3%, čo predstavuje rozdiel vyšší ako 10 p. b. oproti priemeru krajín EU27. Je to viac ako štvornásobok hodnoty vykazovanej za Holandsko.

Pri porovnaní priemerných hodnôt vybraných indikátorov za sledované obdobie 10-tich rokov, vidíme, že najvýraznejšia zmena bola pri ukazovateli miera zamestnanosti (MZ). S odstupom desiatich rokov pozorujeme nárast takmer o 6 p. b. Ďalší ukazovateľ, kde evidujeme výraznejšiu zmenu je podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl). Tu registrujeme pokles o 2 p. b., čo pri povahe ukazovateľa je pozitívny výsledok. Pri ostatných ukazovateľoch zmeny v priemerných hodnotách nepresiahli 1 p. b.

Pozorujeme tiež, že Litva, ktorá v roku 2010, dosahovala extrémne hodnoty v negatívnom zmysle, tak až pri troch ukazovateľoch sa výrazne za 10 rokov zlepšila. Pri ukazovateli relatívna mediánová medzera rizika chudoby (RCHu) sa jedná o pokles o 9,42 p. b., hodnota ukazovateľa rozdelenie príjmov (RP) poklesla o 1,21 p. b. a ukazovateľ podiel príjmov spodných 40% populácie (PP40) sa zvýšil o 1 p. b., čiže zlepšil.

5 Výsledky viacrozmernej analýzy

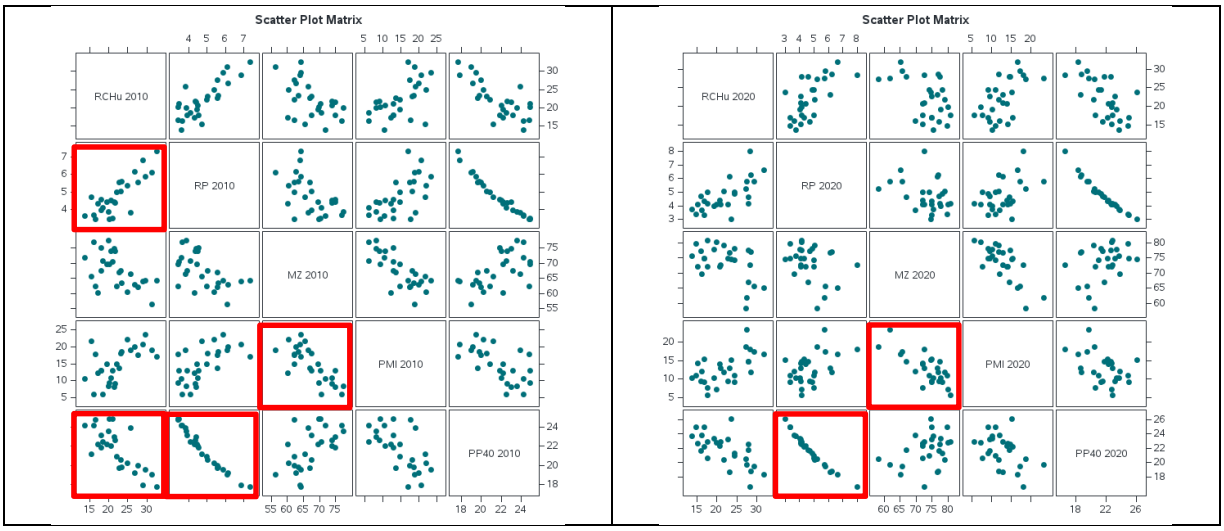
Pri viacrozmerných analýzach štatistických jednotiek sa vychádza zo súvislostí medzi jednotlivými ukazovateľmi. Základný predpoklad je, že analyzované ukazovatele sú závislé, čiže korelované. Vzťahy závislosti sme zmapovali pomocou korelačnej matice, ktorá obsahuje Pearsonove korelačné koeficienty aj s ich p-hodnotami. Korelačnú maticu znázorňuje tabuľka 3. Tučným zvýraznené koeficienty korelácie dosahujú hodnoty vyššie ako 0,5 v absolútnej hodnote a ich p-hodnoty sú blízke číslu 0, čiže sú štatisticky významné pre ľubovoľne zvolenú hladinu významnosti alfa.

Na základe hodnôt z roku 2010, vidíme že všetky koeficienty korelácie sú štatisticky významné a vyššie ako v absolútnej hodnote ako 0,5. Ale v roku 2020 až päť koeficientov korelácie dosiahlo absolútnu hodnotu nižšiu ako 0,5 a sú významné až na hladine významnosti vyššej ako 0,10 (10%).

Tab. 4 Korelačná matica ukazovateľov, rok 2010 a 2020 *(Zdroj: vlastné spracovanie v SAS Viya)*

	RRCHu 2010	RP 2010	MZ 2010	PMl 2010	PP40 2010		RCHu 2020	RP 2020	MZ 2020	PMl 2020	PP40 2020
RCHu 2010	1					**RCHu 2020**	1				
RP 2010	**0,82**	1				**RP 2020**	**0,68**	1			
	p=,000						**p=,000**				
MZ 2010	**-0,51**	**-0,55**	1			**MZ 2020**	-0,49	-0,34	1		
	p=,006	**p=,003**					p=,009	p=,082			
PMl 2010	**0,56**	**0,61**	**-0,75**	1		**PMl 2020**	**0,60**	0,46	**-0,75**	1	
	p=,002	**p=,001**	**p=,000**				**p=,001**	p=,015	**p=,000**		
PP40 2010	**-0,75**	**-0,98**	**0,57**	**-0,63**	1	**PP40 2020**	**-0,64**	**-0,98**	0,35	-0,43	1
	p=,000	**p=,000**	**p=,002**	**p=,001**			**p=,000**	**p=,000**	p=,076	p=,026	

Na nasledujúciom grafe 1 sú zobrazené rozptylové matice. Vyznačené sú tie časti, kde sú korelácie najsilnejšie. Sú tam hodnoty koeficientov korelácie absolútne nad 0,7.



Graf 1 Grafická korelačná matica analyzovaných ukazovateľov, rok 2010 a 2020 *(Zdroj: Vlastné spracovanie v SAS Viya)*

Aby sme posúdili vhodnosť údajov pre viacrozmerné metódy, tak sme vypočítali aj KMO mieru za obidva roky. Na základe KMO mier konštatujeme, že zvolené ukazovatele v obidvoch sledovaných obdobiach dosahujú celkové miery hodnotu nad 0,5 a teda sú vhodné na viacrozmerné analýzy pomocou metód FA a CA. Tieto výsledky sú prezentované v tabuľke 2. Celková hodnota KMO miery v roku 2010 je na úrovni až 0,705 a v roku 2020 sa mierne znížila (0,667). Taktiež individuálne hodnoty týchto mier za oba roky pri všetkých vybraných ukazovateľoch sú vyššie ako 0,5.

Údaje sme ďalej analyzovali metódou faktorovej analýzy (FA). Na odhad faktorových váh sme použili metódu PCA z kovariačnej matice, lebo sa jedná o relatívne ukazovatele s rozdielnou úrovňou stredných hodnôt. V tabuľke 5 uvádzame hodnoty vlastných čísel s podielom rozptylu, ktorý faktory vysvetľujú.

Tab. 5 Hodnoty KMO miery, roky 2010 a 2020 *(Zdroj: Vlastné spracovanie v SAS Viya)*

KMO štatistika 2010 Celková miera = 0,705						KMO štatistika 2020 Celková miera = 0,667				
RCHu 2010	**RP 2010**	**MZ 2010**	**PMl 2010**	**PP40 2010**		**RCHu 2020**	**RP 2020**	**MZ 2020**	**PMl 2020**	**PP40 2020**
0,735	0,632	0,782	0,811	0,655		0,916	0,605	0,627	0,672	0,609

Pre voľbu počtu významných spoločných faktorov sme použili Kaiserovo pravidlo. Z tabuľky 5 vidno, že priemerná hodnota vlastných čísel v roku 2010 bola 17,57 a v roku 2020 klesla na hodnotu 16,35. Kaiserovo pravidlo hovorí, že treba vybrať toľko spoločných faktorov z maximálneho počtu odhadnutých faktorov, koľko z nich má hodnoty vlastných čísiel vyšších ako je priemer. V oboch rokoch to platí len pre prvý faktor, ktorý vysvetľuje viac ako 70% celkovej variability údajov. Ak chceme vysvetliť vyšší podiel celkovej variability, tak je potrebné vybrať viac faktorov. Pri použití empirického pravidla, že vyberieme taký počet spoločných faktorov, ktoré vysvetľujú viac ako 5% variability, by sme vybrali 3 spoločné faktory z piatich. Takto dosiahneme takmer 100% vysvetlenie variability.

Tab. 6 Hodnoty vlastných čísel z kovariačnej matice, rok 2010 a 2020 *(Zdroj: vlastné spracovanie v SAS Viya)*

Vlastné čísla kovariančnej matice (2010) priemer: 17,57					Vlastné čísla kovariančnej matice (2020) priemer: 16,35			
	Vlastné číslo	**vysvetlený rozptyl**	**vysv. rozptyl kumulatívne**			**Vlastné číslo**	**vysvetlený rozptyl**	**vysv. rozptyl kumulatívne**
1	**64,72**	**0,74**	**0,74**		**1**	**57,64**	**0,71**	**0,71**
2	**14,45**	**0,16**	**0,90**		**2**	**16,29**	**0,20**	**0,90**
3	**6,81**	**0,08**	**0,98**		**3**	**4,57**	**0,06**	**0,96**
4	1,84	0,02	1,00		**4**	3,22	0,04	1,00
5	0,02	0,00	1,00		**5**	0,03	0,00	1,00

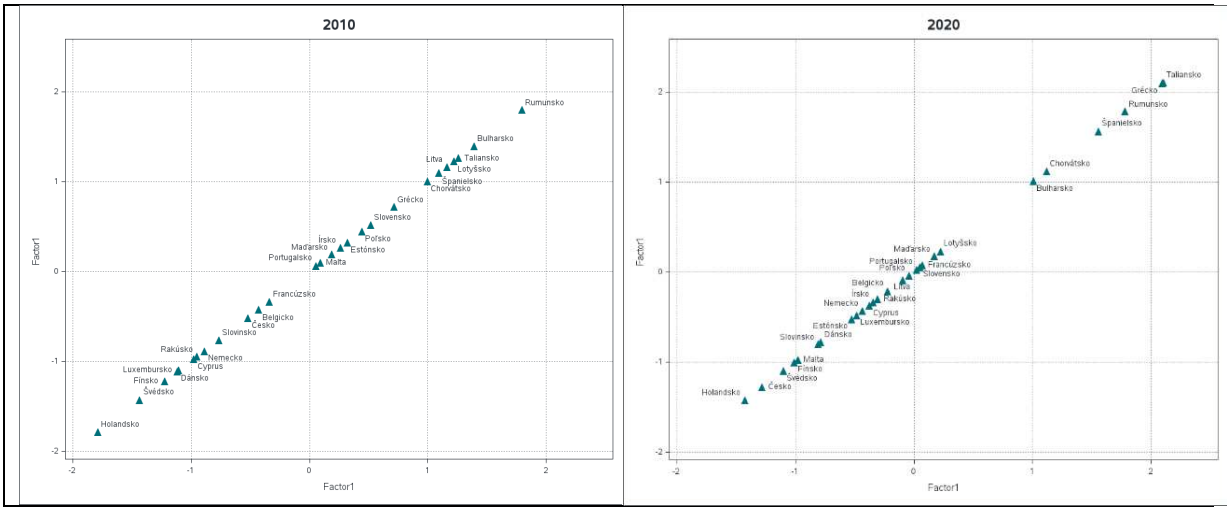
Faktorové váhy zvolených troch faktorov bez rotácie sú uvedené v tabuľke 6. Vidíme, že prvý faktor (Faktor1; na grafickom znázornení Factor1) v oboch obdobiach silno koreluje so všetkými ukazovateľmi, čiže je možné s ním uvažovať ako s integrálnym ukazovateľom. Všetky faktorové váhy pre Faktor1 sú vyššie absolútne ako 0,5. Silnú pozitívnu koreláciu vykazuje s ukazovateľmi typu „-“: relatívna mediánová medzera rizika chudoby (RCHu), rozdelenie príjmov (RP) a podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl). Naopak silnú negatívnu koreláciu má s ukazovateľmi typu „+“: miera zamestnanosti (MZ) a podiel príjmov spodných 40% populácie (PP40). Vysoké faktorové skóre pre jednotlivé krajiny tohto faktora znamená celkovo negatívne postavenie krajiny a naopak nízke hodnoty znamenajú pozitívne postavenie krajiny v rebríčku tohto viacrozmerného hodnotenia SDG cieľa 10.

Tab. 7 Faktorové váhy bez rotácie, rok 2010 a 2020 *(Zdroj: vlastné spracovanie v SAS Viya)*

Faktorové váhy bez rotácie (2010)			
	Faktor1	**Faktor2**	**Faktor3**
RCHu 2010	**0,78**	**0,62**	-0,08
RP 2010	**0,77**	0,42	0,00
MZ 2010	**-0,90**	0,34	0,28
PMl 2010	**0,90**	-0,16	0,41
PP40 2010	**-0,77**	-0,33	-0,02

Faktorové váhy bez rotácie (2020)			
	Faktor1	**Faktor2**	**Faktor3**
RCHu 2020	**0,82**	**0,56**	-0,08
RP 2020	**0,61**	0,45	0,02
MZ 2020	**-0,88**	0,44	0,16
PMl 2020	**0,87**	-0,14	0,47
PP40 2020	**-0,59**	-0,42	0,03

Krajiny sme zoradili na základe hodnôt faktorového skóre prvého spoločného faktora v oboch sledovaných rokoch, čo vidno na grafe 2. Najlepšie faktorové skóre v oboch obdobiach dosiahlo Holandsko, v roku 2010 s hodnotou -1,78 a v roku 2020 s hodnotou -1,43. Naopak poslednú priečku v roku 2010 obsadilo Rumunsko s hodnotou skóre 1,8 a v roku 2020 zase Taliansko s hodnotou 2,1.



Graf 2 Zoradenie krajín podľa Faktor1 bez rotácie, rok 2010 a 2020)
(Zdroj: vlastné spracovanie v SAS Viya)

Keď porovnáme poradie krajín v sledovaných obdobiach, najvýraznejšie zmeny v pozitívnom zmysle pozorujeme v týchto krajinách: Litva (v 2010 Faktor1 = 1,16; v 2020 Faktor1 = -0,10), Malta (v 2010 Faktor1 = 0,09; v 2020 Faktor1 = -0,98) a Lotyšsko (v 2010 Faktor1 = 1,22; v 2020 Faktor1 = 0,22). Naopak, najviac si faktorové skóre zhoršili krajiny: Grécko (v 2010 Faktor1 = 1,71; v 2020 Faktor1 = 2,09), Taliansko (v 2010 Faktor1 = 1,26; v 2020 Faktor1 = 2,10) a Rakúsko (v 2010 Faktor1 = -0,95; v 2020 Faktor1 = -0,30). V roku 2010 Slovensko dosiahlo hodnotu faktorového skóre 0,52 a skončilo na 19. mieste, v roku 2020 hodnota faktorové-ho skóre bola 0,05 a umiestnilo sa na 18. mieste.

Ak sa rozhodneme použiť až 3 spoločné faktory, ktoré vysvetľujú takmer celú variabilitu údajov v rokoch 2010 (98%) a 2020 (96%), tak je potrebné faktory rotovať. Na lepšiu interpretáciu faktorov, umelo vytvorených premenných, sme

použili formu transformácie faktorov, tzn. rotáciu faktorov. Zabezpečili sme tým výhodnejšie prerozdelenie vysvetľovanej variability medzi jednotlivé faktory. Konkrétne sme použili metódu ortogonálnej rotácie varimax, pričom sa jedná pravouhlú rotáciu s ortomaxnou váhou rovnej číslu 1. Použitím ortogonálnej rotácie dostaneme faktory, ktoré nie sú korelované, čo je jednou z podmienok na využitie premenných v zhlukovej analýze (ďalej CA).

V tabuľke 7 sú uvedené faktorové váhy prvých troch faktorov po rotácii varimax. V oboch sledovaných obdobiach je prvý faktor v silnej miere pozitívne korelovaný s ukazovateľom rozdelenie príjmov (RP), naopak s ukazovateľom podiel príjmov spodných 40% populácie (PP40) pozorujeme silnú negatívnu koreláciu. Pri faktore 2 pozorujeme taktiež v oboch sledovaných obdobiach silnú pozitívnu koreláciu s mierou zamestnanosti (MZ) a strednú negatívnu koreláciu s ukazovateľom podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl) (v roku 2010 je to 0,4 a v roku 2020 to je 0,45). Tretí z faktorov silne pozitívne koreluje s ukazovateľom relatívna mediánová medzera rizika chudoby (RCHu), pričom žiadne ďalšie korelácie faktor nevykazuje. Ide teda o triviálny faktor.

Rotáciou sa okrem faktorových váh zmenil aj podiel vysvetlenej variability jednotlivými faktormi. V roku 2010 prvý faktor vysvetlil iba 14,1% celkovej variability, na druhý faktor pripadá 30,55% a tretí faktor 19,81%. Pri dátach z roku 2020 na prvý faktor pripadá podiel 11,62% z celkového rozptylu, pri druhom faktore je to 31,58% a tretí faktor vysvetľuje 22,65% variability.

Tab. 8 Faktorové váhy po rotácii Varimax, rok 2010 a 2020 *(Zdroj: Vlastné spracovanie v SAS Viya)*

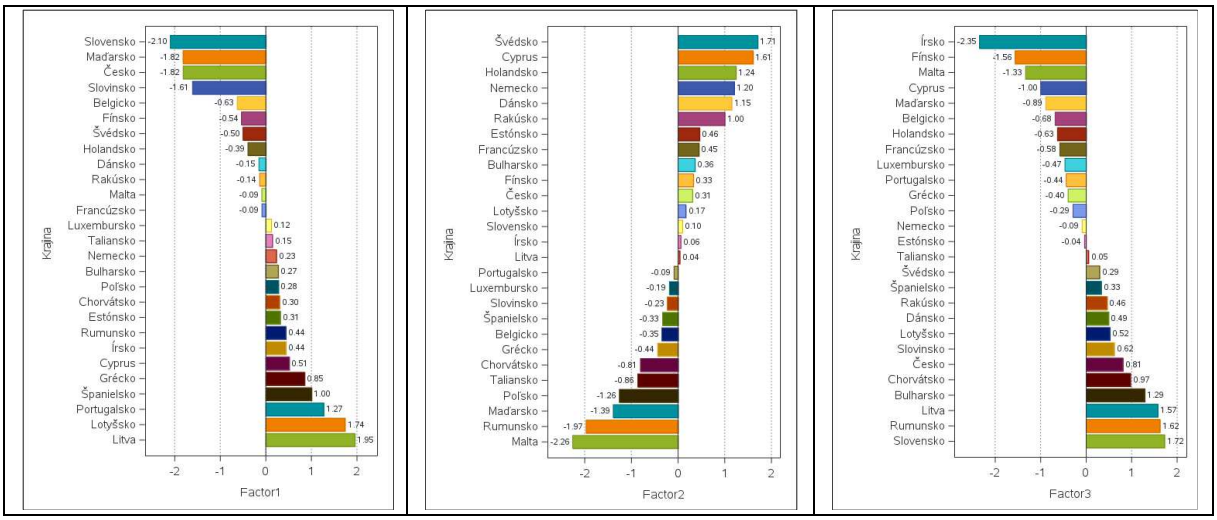
Faktorové váhy po rotácii Varimax (2010)					Faktorové váhy po rotácii Varimax (2020)			
	Faktor1	**Faktor2**	**Faktor3**			**Faktor1**	**Faktor2**	**Faktor3**
RCHu 2010	0,47	-0,20	0,84		**RCHu 2020**	0,41	-0,22	0,85
RP 2010	0,86	-0,21	0,39		**RP 2020**	0,94	-0,11	0,27
MZ 2010	-0,25	0,89	-0,18		**MZ 2020**	-0,14	0,92	-0,18
PMl 2010	0,29	**-0,40**	0,21		**PMl 2020**	0,21	**-0,45**	0,25
PP40 2010	-0,89	0,24	-0,28		**PP40 2020**	-0,95	0,14	-0,23

Získané prvé 3 spoločné faktory v tabuľke 7 je možné na základe rotovaných faktorových váh interpretovať nasledovne.

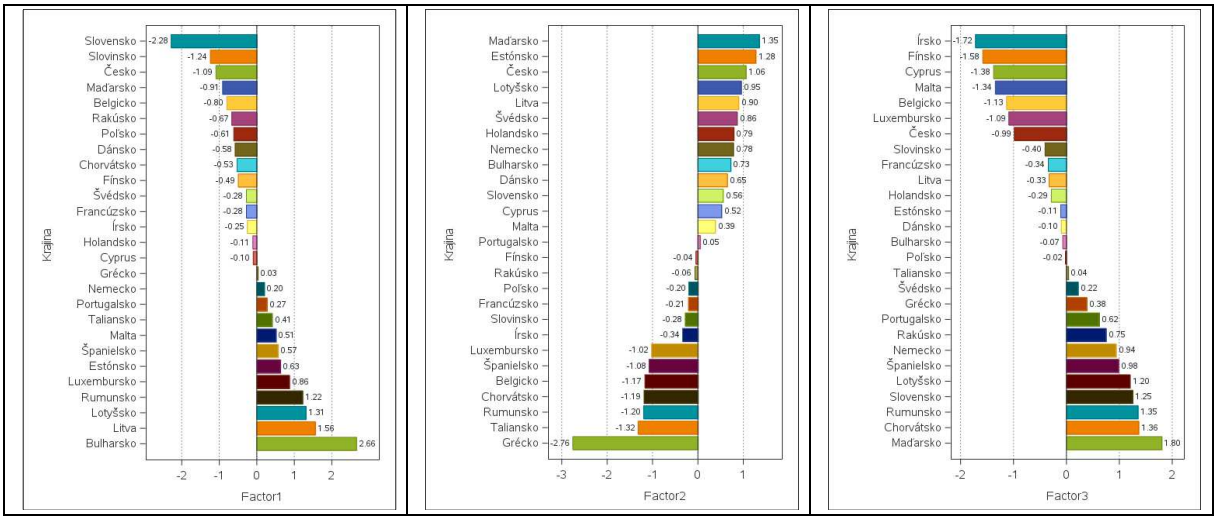
- Faktor1 je miera nerovnosti distribúcie príjmov v krajine (čím vyššia hodnota skóre tohto faktora, tým je rozdelenie príjmov nerovnomernejšie).
- Faktor2 je miera celkovej zamestnanosti v krajine, prihliadajúc na začlenenie mladých ľudí v krajine (čím vyššia hodnota skóre tohto faktora, tým je celková zamestnanosť v krajine vyššia a nezamestnanosť alebo neúčasť na odborných tréningoch mladých ľudí nižšia).

- Faktor3 je miera rizika chudoby v krajine (čím vyššia hodnota skóre tohto faktora, tým je v krajine vyššia miera chudoby a zároveň nerovnomernejšie rozdelenie príjmov). Je to triviálny faktor, lebo je vysoko korelovaný len s 1 indikátorom, s RCHu.

Na základe hodnôt faktorových skóre týchto 3 spoločných faktorov sme krajiny zoradili. Zoradenie je prezentované na grafoch 3 a 4. Vzostupné poradie sme použili pri prvom a treťom faktore, keďže nižšia hodnota tohto skóre predstavuje rovnomernejšie rozdelenie príjmov v rámci krajiny, nižšiu mieru rizika chudoby a teda lepší hodnotenie krajiny. Pri druhom faktore sme naopak zvolili zostupné poradie, nakoľko vyššia hodnota predstavuje vyššiu zamestnanosť v krajine a tak lepšie hodnotenie.



Graf 3 Zoradenie krajín podľa faktorov, rok 2010 *(Zdroj: vlastné spracovanie v SAS Viya)*



Graf 4 Zoradenie krajín podľa faktorov, rok 2020 *(Zdroj: vlastné spracovanie v SAS Viya)*

Čo sa týka prvého faktora, pozorujeme, že v obidvoch sledovaných obdobiach dosahuje najlepšie faktorové skóre Slovensko, v roku 2010 je to -2,1 a v roku

2020 je to -2,28. Taktiež v obidvoch sledovaných obdobiach sú na prvých piatich priečkach rovnaké krajiny: Slovensko, Maďarsko, Česko, Slovinsko a Belgicko. Naopak na poslednej priečke v roku 2010 skončila Litva s hodnotou 1,95 a v roku 2020 Bulharsko, ktoré dosiahlo skóre až 2,66. Najvýraznejšie zmeny hodnôt pri tomto faktore pozorujeme pri krajinách ako Portugalsko, kde sa jedná o pozitívnu zmenu (v 2010 Faktor1 = 1,27; v 2020 Faktor1 = 0,27) a Bulharsko, kde naopak konštatujeme negatívnu zmenu (v 2010 Faktor1 = 0,27; v 2020 Faktor1 = 2,66).

Pri druhom faktore v roku 2010 dosahuje najlepšie faktorové skóre Švédsko (1,71) a naopak najhoršiu hodnotu tohto skóre mala Malta (-2,26). V roku 2020 sa poradie výrazne zmenilo. Najlepšiu hodnotu skóre dosiahlo Maďarsko (1,35) a najhoršiu Grécko (-2,76). Najvýraznejšie vývojové zmeny pozorujeme pri Maďarsku, pričom sa jedná o pozitívnu zmenu (v 2010 Faktor2 = -1,39; v 2020 Faktor2 = 1,35) a pri Grécku, kde ide o negatívnu zmenu (v 2010 Faktor2 = -0,44; v 2020 Faktor2 = -2,76). Slovensko dosiahlo v roku 2010 faktorové skóre 0,1 a umiestnilo sa na 13. mieste. V roku 2020 hodnota tohto faktorového skóre sa zvýšila (0,56) a tak SR skončila na 11. mieste.

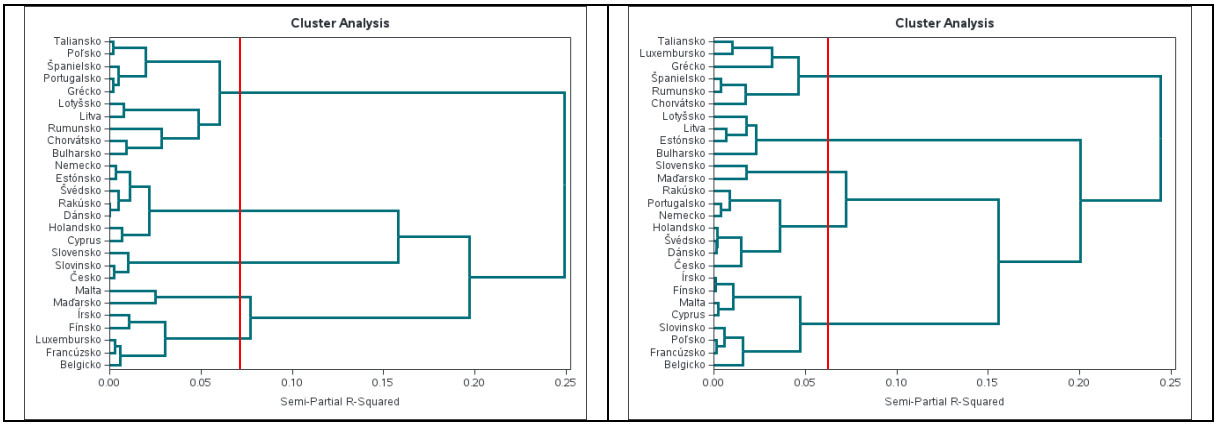
Pri treťom faktore pozorujeme najlepšie faktorové skóre v oboch sledovaných obdobiach v Írsku (v roku 2010 Faktor3 = -2,35 a v roku 2020 Faktor3 = -1,72). Naopak najhoršie faktorové skóre dosiahlo v roku 2010 Slovensko s hodnotou 1,72 a v roku 2020 Maďarsko s hodnotou 1,8. V roku 2020 pozorujeme mierne zlepšenie pozície Slovenska z poslednej priečky na 24. miesto, hodnota faktorového skóre je 1,25. Čo sa týka najvýraznejších zmien vo faktorom skóre podľa tretieho faktora, tak zlepšenie vykazuje Litva (v 2010 Faktor3 = 1,57; v 2020 Faktor2 = -0,33) a na druhej strane najvýraznejšie zhoršenie sme zaznamenali pri Maďarsku (v 2010 Faktor3 = -0,89; v 2020 Faktor2 = 1,8).

Dimenziu pôvodných piatich ukazovateľov sme teda znížili o dve úrovne. Ako bolo vyššie uvedené, použitím FA s pravouhlou rotáciou sme zabezpečili nekorelovanosť umelo vytvorených ukazovateľov (tzv. latentných premenných), troch faktorov. Použijeme ich následne v zhlukovej analýze (CA), aby sme vytvorili zhluky (skupiny) krajín s podobnými charakteristikami podľa jednotlivých faktorov.

Pri zhlukovej analýze (CA) existuje niekoľko spôsobov, na základne ktorých sa vytvárajú zhluky. Zvolili sme hierarchický postup zhlukovania Wardovou metódou, ktorá je založená na minimalizácii vnútrozhlukovej variability. Pri použití nie je potrebné dopredu zadať konečný počet zhlukov.

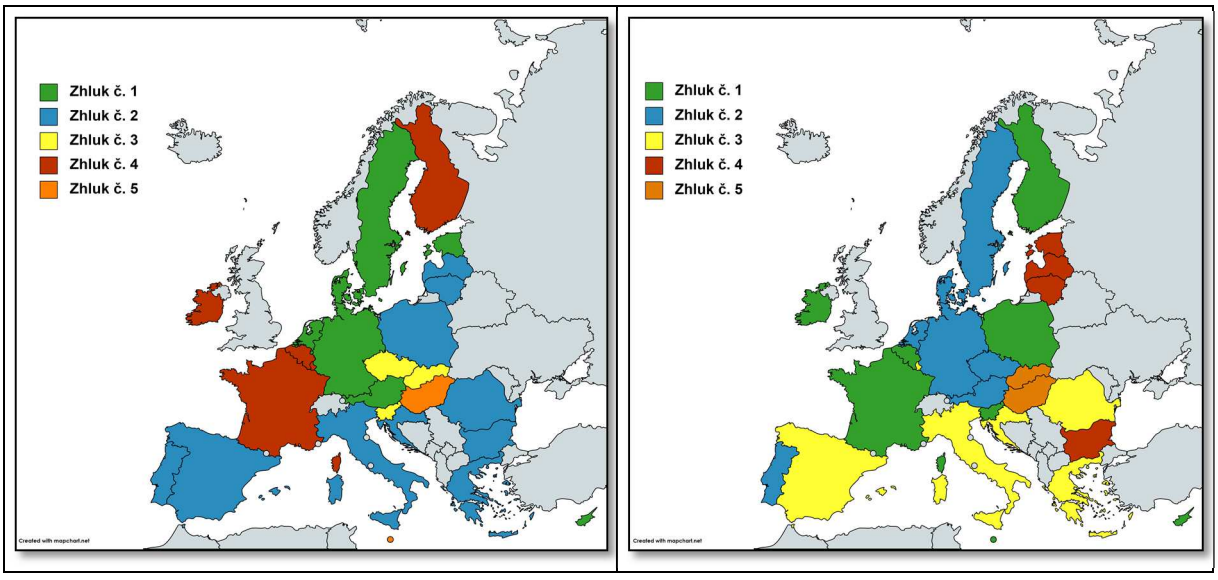
Konečný počet zhlukov sa dá určiť na základe priebehu zhlukovania znázorneného na diagrame zhlukovania, tzv. dendrograme , ktorý je súčasťou grafu 5. Zároveň sme si v SAS Viya vypočítali semiparciálny koeficient

determinácie, ktorý vyjadruje pomer medzi rozdielom vnútroskupinovej variability novovytvoreného zhluku a súčtu vnútroskupinových variabilít tvoriacich nový zhluk voči celkovej variabilite (ďalej len ako SPRSQ). Na základe hodnôt tohto koeficientu sme určili vhodný počet zhlukov. Pre obe sledované obdobia je to 5 zhlukov (hodnota SPRSQ pre rok 2010 = 0,06 a pre rok 2020 = 0,05).



Graf 5 Dendrogramy procesu zhlukovania, rok 2010 a 2020
(Zdroj: vlastné spracovanie v SAS Viya)

Na základe zvoleného počtu zhlukov 5, boli krajiny rozdelené do skupín tak ako to znázorňujú mapy na obrázku 2. Označenia zhlukov číslami a farbami nie je podstatné, lebo ich zhlukom priradil softvér. Dôležité je sledovať a interpretovať hodnoty popisných štatistík vstupných ukazovateľov pre vytvorené zhluky krajín. Pomáhajú nám pri tom tabuľky a grafy uvedené v prílohách článku (Príloha 1 až 3).



Obr. 2 Mapa rozdelenia krajín do zhlukov, rok 2010 a 2020
(Zdroj: Vlastné spracovanie na mapchart.net)

Podľa základných charakteristík pôvodných ukazovateľov v jednotlivých zhlukoch, môžeme opísať spoločné znaky krajín v danom zhluku. Pre ***zhluky vytvorené pre rok 2010*** konštatujeme nasledovné skutočnosti:

- **Zhluk č. 1** tvorí 7 krajín: ***Dánsko, Rakúsko, Nemecko, Estónsko, Švédsko, Holandsko, Cyprus***. Krajiny spája vysoká miera zamestnanosti (MZ), ktorá dosahuje najlepšie výsledky práve v tomto zhluku (priemer = 74,24%). Rovnako aj ukazovateľ podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl), dosahuje v tomto zhluku najlepšie výsledky (priemer = 10,5%). Vzhľadom na vysokú variabilitu tohto ukazovateľa (variačný koeficient Vk = 37,91%), však musíme konštatovať, že medzi krajinami sú značné rozdiely. Čo sa týka rozdelenia príjmov (RP, PP40), tak krajiny zhluku dosahujú dobré výsledky s nízkou variabilitou.
- **Zhluk č. 2** je tvorený 10 krajinami: ***Poľsko, Taliansko, Portugalsko, Grécko, Španielsko, Lotyšsko, Litva, Chorvátsko, Bulharsko, Rumunsko***. Jedná sa o najpočetnejší zhluk, v ktorom až 4 indikátory dosahujú maximálne hodnoty. Krajiny spája predovšetkým pomerne vysoká miera podielu mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl, 18,67%) a zároveň pomerne nízka miera zamestnanosti (MZ, 62,66%). Táto skupina vykazuje najvyššiu priemernú hodnotu relatívnej mediánovej medzery rizika chudoby (RCHu, 26,99%) s pomerne vysokou variabilitou (Vk = 13,53%). Ukazovatele distribúcie príjmov (RP a PP40) dosahujú najhor šie výsledky medzi vytvorenými zhlukmi.
- **Zhluk č. 3** tvoria 3 krajiny: ***Česko, Slovinsko, Slovensko***. Tento zhluk dosahuje veľmi dobré hodnoty ukazovateľov distribúcie príjmov. Konkrétne rozdelenie príjmov (RP, 3,56) a podiel príjmov spodných 40% populácie (PP40, 24,57%) boli v tomto zhluku najnižšie a aj s nízkou variabilitou. Pri porovnaní priemerných hodnôt ukazovateľov miery zamestnanosti (MZ, 68,87%) a PMl (13,77%) môžeme tiež konštatovať, že dosahujú lepšie výsledky ako je priemer EU27.
- **Zhluk č. 4** pozostáva z 5 krajín: ***Francúzsko, Luxembursko, Belgicko, Fínsko, Írsko***. Krajiny v tomto zhluku takmer pre všetky ukazovatele dosiahli priemerné hodnoty. Výnimku predstavuje relatívna mediánová medzera rizika chudoby (RCHu, 17,08%), ktorá je vyššia ako priemer a aj variabilita ukazovateľa je pomerne vysoká.
- **Zhluk č. 5** predstavujú 2 krajiny: ***Malta, Maďarsko.*** Tento zhluk spája iba dve krajiny a zaznamenávame tu až dve extrémne hodnoty. Bola tu najnižšia hodnota relatívnej mediánovej medzery rizika chudoby (RCHu, 16,9%) a zároveň najvyššia miera zamestnanosti (MZ, 61,15%). Ostatné ukazovatele dosahujú mierne lepšie výsledky ako je priemer EU27. Čo sa týka ukazovateľa

podielu mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl), tak nadobudol priaznivejšiu hodnotu ako bol priemer EU27 v roku 2010.

V roku 2020 sa rozdelenie krajín do skupín výrazne zmenilo v porovnaní s rokom 2010. Pre ***zhluky vytvorené pre rok 2020*** na základe opisných štatistík pozorujeme nasledovné charakteristiky:

- **Zhluk č. 1** je zostavený z 8 krajín: ***Fínsko, Írsko, Poľsko, Francúzsko, Cyprus, Malta, Slovinsko, Belgicko.*** Tento zhluk je najpočetnejší a bolo sem zaradených takmer 30% krajín EU27. Výraznou črtou tohto zhluku je najnižšia priemerná hodnota ukazovateľa relatívna mediánová medzera rizika chudoby (RCHu, 17,8%). V ostatných ukazovateľoch dosahuje zhluk mierne lepšie hodnoty ako je priemer EU27.
- **Zhluk č. 2** tvorí 7 krajín: ***Dánsko, Švédsko, Holandsko, Nemecko, Portugalsko, Rakúsko, Česko.*** Tento zhluk vykazuje najlepšie výsledky v oblasti zamestnanosti. Miera zamestnanosti je maximálna (MZ, 77,94%) a podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl, 9,1%) je minimálny. Čo sa týka ukazovateľov distribúcie príjmov, tak pozorujeme tiež pomerne dobré výsledky. Ich hodnoty sú mierne nad priemerom celého súboru.
- **Zhluk č. 3** reprezentuje 6 krajín: ***Španielsko, Rumunsko, Taliansko, Luxembursko, Chorvátsko, Grécko.*** Krajiny v tomto zhluku dosahujú najhoršie priemerné hodnoty až v troch zo sledovaných piatich ukazovateľov. Jedná sa o indikátory ako miera zamestnanosti (MZ, 65,02%) a podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl, 16,37%). Taktiež relatívna mediánová medzera rizika chudoby je maximálna (RCHu, 26,98%). Vzhľadom na vysokú variabilitu ukazovateľov v zhluku konštatujeme, že medzi krajinami sú významné rozdiely. Pri ukazovateľoch distribúcie príjmov ide o podpriemerné hodnoty v porovnaní s priemermi všetkých sledovaných krajín.
- **Zhluk č. 4** obsahuje 4 krajiny: ***Estónsko, Litva, Lotyšsko, Bulharsko***. Tento zhluk v negatívnom zmysle „vyniká“ v distribúcii príjmov. Oba ukazovatele, rozdelenie príjmov (RP, 6,36) aj podiel príjmov spodných 40% populácie (PP40, 18,63%), dosahujú najhoršie hodnoty. Pri miere zamestnanosti (MZ, 76,35%) pozorujeme mierne lepšiu hodnotu ako je priemer všetkých sledovaných krajín. Naopak pri podiele mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl, 13,73%) tento zhluk vykazuje mierne horší výsledok ako priemer.
- **Zhluk č. 5** –je tvorený 2 krajinami: ***Slovensko, Maďarsko.*** Jedná sa o najmenší zo zhlukov vytvorených v danom období. Výraznou črtou sú výborné výsledky

v oblasti distribúcie príjmov. Zhluk dosahuje najpriaznivejšie, čiže minimálne priemerné hodnoty v oboch ukazovateľoch: rozdelenie príjmov (RP, 3,6) a podiel príjmov spodných 40% populácie (PP40, 24,35%). Miera zamestnanosti (MZ, 76,05%) je na pomerne vysokej úrovni. No naopak, krajiny vykazujú vysoký podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave (PMl, 14,95%). Relatívna mediánová medzera rizika chudoby (RCHu, 25,8%) je tiež pomerne vysoká.

6 Záver

Na základe výsledkov viacrozmernej analýzy možno konštatovať, že krajiny EU27 je možné rozdeliť podľa sledovaných ukazovateľov do pomerme homogénnych piatich skupín (zhlukov).

V roku 2010 najlepšie výsledky dosiahol zhluk č. 1, do ktorého bolo zaradených 7 krajín: Dánsko, Rakúsko, Nemecko, Estónsko, Švédsko, Holandsko a Cyprus. Tieto krajiny vykazujú najpriaznivejšie, resp. najvyššie hodnoty celkovej zamestnanosti. Podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave je nízky. Pri ukazovateľoch distribúcie príjmov a rizika chudoby dosahujú nadpriemerné hodnoty v porovnaní s priemermi všetkých sledovaných krajín. Naopak najslabšie výsledky dosahuje zhluk č. 2, kde bolo zaradených desať krajín: Poľsko, Taliansko, Portugalsko, Grécko, Španielsko, Lotyšsko, Litva, Chorvátsko, Bulharsko a Rumunsko. Pre tieto krajiny predstavuje najproblematickejšiu oblasť riziko chudoby ako aj vysoký podiel mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave. Miera zamestnanosti v zhluku 2 dosahuje podpriemerné hodnoty. Slovensko bolo v tomto období (rok 2010) zaradené do zhluku č. 3, ktorý predovšetkým vyniká v oblasti distribúcie príjmov.

V roku 2020 dosiahol najlepšie výsledky zhluk č. 2, ktorý tvorí 7 krajín: Dánsko, Švédsko, Holandsko, Nemecko, Portugalsko, Rakúsko a Česko. V týchto krajinách je najlepšie situácia v podiele mladých ľudí, ktorí nie sú zamestnaní, ani nie sú vo vzdelávaní a odbornej príprave ako aj v celkovej zamestnanosti. Tento zhluk dosahuje priaznivé hodnoty v ukazovateľoch distribúcie príjmov a rizika chudoby. Pri zhluku č. 3 pozorujeme naopak najslabšie výsledky. Patria tam krajiny: Španielsko, Rumunsko, Taliansko, Luxembursko, Chorvátsko a Grécko. Vykazujú najhoršie výsledky v rámci zamestnanosti ako aj rizika chudoby. Pri distribúcii prímov taktiež pozorujeme horšie výsledky ako je priemer všetkých sledovaných krajín. Čo sa týka Slovenska, bolo zaradené do zhluku č. 5, ktorý sa pýši najlepšími výsledkami v rámci distribúcie príjmov. Za vyzdvihnutie stojí fakt, že Slovensko v roku 2020 dosiahlo najlepšie výsledky v distribúcii príjmov spomedzi všetkých sledovaných krajín.

Záverom konštatujeme, že 5 krajín EU27 si udržalo svoje prvenstvo v obidvoch sledovaných obdobiach. Sú to krajiny: Dánsko, Rakúsko, Nemecko, Švédsko a Holandsko. No žiadna z analyzovaných krajín nevykazuje nadpriemerné výsledky pri všetkých indikátoroch súčasne. Každá krajina má teda oblasť, na zlepšení ktorej je potrebné ešte pracovať, tak aby sa znížili rozdiely a v konečnom dôsledku, aby sa podarilo naplniť strategický cieľ 10 udržateľného rozvoja.

Literatúra

Brundtland, G. H. et al. (1991). Naše společná budoucnost. Praha: Academia, 1991.

Lafortune, G. et al. (2022). Achieving the SDGs: Europe's compass in a multipolar world. Europe Sustainable Development Report 2022. SDSN and SDSN Europe. Paris, 2022

Mederly, P. (2001). Indikátory trvalo udržateľného rozvoja. Bratislava: REC Slovensko, 2001.

Norström, A.V. et al. (2014). Three necessary conditions for establishing effective sustainable development goals in the Antropocene. *Ecology and Society* 19(3): 8.

Stankovičová, I., Vojtková, M. (2007). Viacrozmerné štatistické metódy s aplikáciami. Iura Edition. Bratislava, 2007.

Švihlová, D. (2004). Trvalo udržateľný rozvoj na miestnej úrovni. Banská Bystrica: Ekonomická fakulta Univerzity Mateja Bela v Banskej Bystrici, 2004.

Príloha 1

Tab. 9 Popisné štatistiky zhlukov, rok 2010 (*Zdroj: vlastné spracovanie v SAS Viya)*

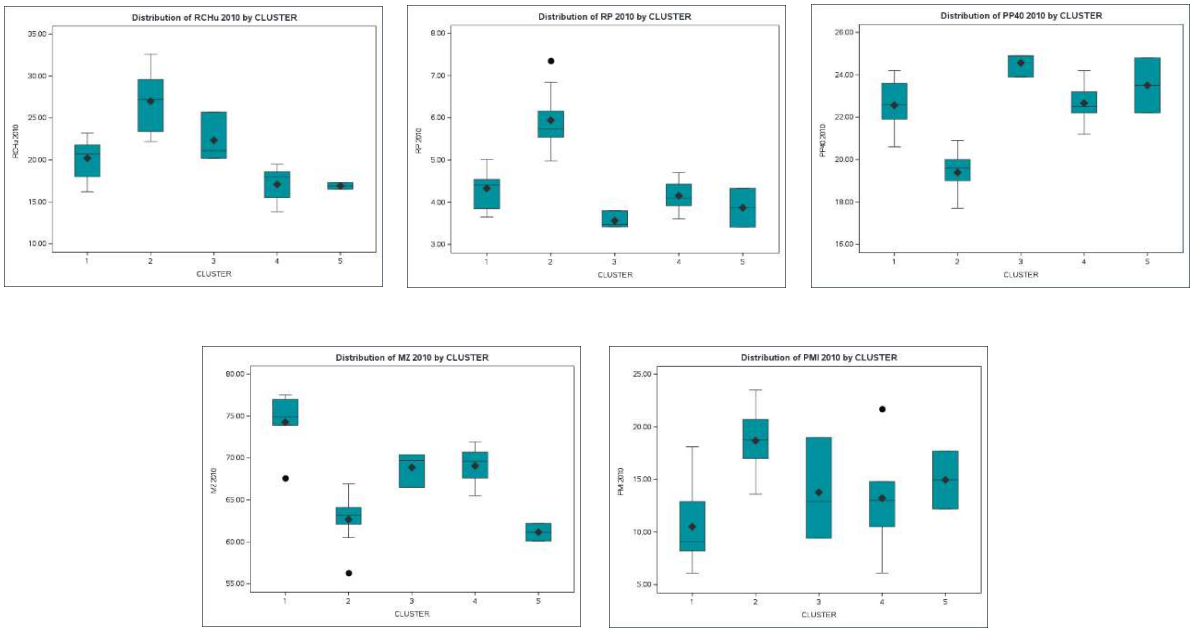
Zhluk	Počet krajín	Ukazovateľ	Priemer	Štandardná odchýlka	Minimum	Maximum	Medián	Variačný koeficient
1	7	RCHu 2010	20,20	2,40	16,20	23,20	20,70	11,91
		RP 2010	4,33	0,45	3,65	5,01	4,41	10,47
		MZ 2010	74,27	3,25	67,60	77,50	74,90	4,38
		PMl 2010	10,50	3,98	6,10	18,10	9,10	37,91
		PP40 2010	22,56	1,19	20,60	24,20	22,60	5,29
2	10	RCHu 2010	26,99	3,65	22,20	32,60	27,20	13,53
		RP 2010	5,94	0,71	4,98	7,35	5,74	11,95
		MZ 2010	62,66	2,80	56,30	66,90	63,15	4,47
		PMl 2010	18,67	3,07	13,60	23,50	18,75	16,46
		PP40 2010	19,39	0,99	17,70	20,90	19,60	5,11
3	3	RCHu 2010	22,33	2,95	20,20	25,70	21,10	13,21
		RP 2010	3,56	0,21	3,42	3,80	3,47	5,79
		MZ 2010	68,87	2,08	66,50	70,40	69,70	3,02
		PMl 2010	13,77	4,86	9,40	19,00	12,90	35,29
		PP40 2010	24,57	0,58	23,90	24,90	24,90	2,35
4	5	RCHu 2010	17,08	2,36	13,80	19,50	18,00	13,81
		RP 2010	4,15	0,43	3,61	4,70	4,10	10,27
		MZ 2010	69,06	2,54	65,50	71,90	69,60	3,68
		PMl 2010	13,22	5,76	6,10	21,70	13,00	43,54
		PP40 2010	22,66	1,12	21,20	24,20	22,50	4,95
5	2	RCHu 2010	16,90	0,57	16,50	17,30	16,90	3,35
		RP 2010	3,87	0,65	3,41	4,33	3,87	16,81
		MZ 2010	61,15	1,48	60,10	62,20	61,15	2,43
		PMl 2010	14,95	3,89	12,20	17,70	14,95	26,01
		PP40 2010	23,50	1,84	22,20	24,80	23,50	7,82

Príloha 2

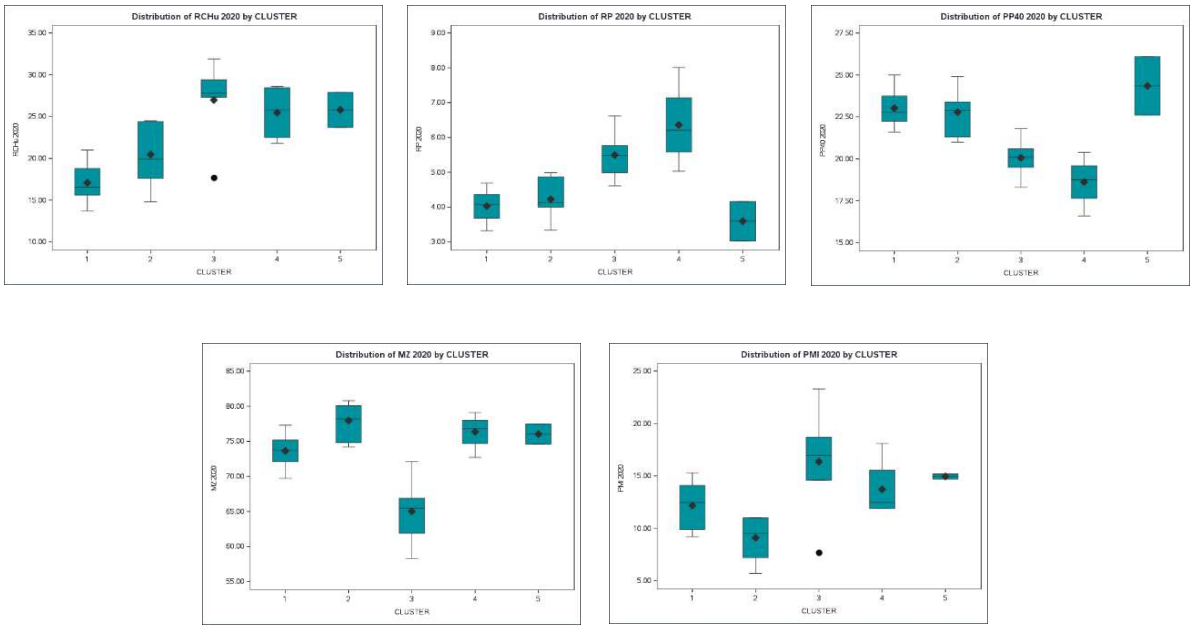
Tab. 10 Popisné štatistiky zhlukov, rok 2020 (*Zdroj: vlastné spracovanie v SAS Viya)*

Zhluk	Počet krajín	Ukazovateľ	Priemer	Štandardná odchýlka	Minimum	Maximum	Medián	Variačný koeficient
1	8	RCHu 2020	17,08	2,55	13,70	21,00	16,55	14,92
		RP 2020	4,03	0,45	3,32	4,69	4,07	11,15
		MZ 2020	73,64	2,42	69,70	77,30	73,75	3,28
		PMl 2020	12,18	2,31	9,20	15,30	12,45	18,95
		PP40 2020	23,03	1,09	21,60	25,00	22,80	4,73
2	7	RCHu 2020	20,46	3,62	14,80	24,50	19,90	17,71
		RP 2020	4,23	0,56	3,34	4,99	4,12	13,19
		MZ 2020	77,94	2,58	74,20	80,80	78,20	3,31
		PMl 2020	9,10	1,99	5,70	11,00	9,50	21,87
		PP40 2020	22,79	1,32	21,00	24,90	22,90	5,79
3	6	RCHu 2020	26,98	4,85	17,70	31,90	27,80	17,98
		RP 2020	5,50	0,71	4,61	6,62	5,49	12,91
		MZ 2020	65,02	4,67	58,30	72,10	65,45	7,18
		PMl 2020	16,37	5,15	7,70	23,30	16,95	31,47
		PP40 2020	20,07	1,19	18,30	21,80	20,10	5,92
4	4	RCHu 2020	25,48	3,48	21,80	28,60	25,75	13,68
		RP 2020	6,36	1,23	5,03	8,01	6,21	19,35
		MZ 2020	76,35	2,67	72,70	79,10	76,80	3,49
		PMl 2020	13,73	2,96	11,90	18,10	12,45	21,58
		PP40 2020	18,63	1,56	16,60	20,40	18,75	8,37
5	2	RCHu 2020	25,80	2,97	23,70	27,90	25,80	11,51
		RP 2020	3,60	0,80	3,03	4,16	3,60	22,23
		MZ 2020	76,05	2,05	74,60	77,50	76,05	2,70
		PMl 2020	14,95	0,35	14,70	15,20	14,95	2,36
		PP40 2020	24,35	2,47	22,60	26,10	24,35	10,16

Príloha 3



Graf 6 Krabicové grafy pre vstupné indikátory podľa zhlukov, rok 2010
(*Zdroj: vlastné spracovanie v SAS Viya)*



Graf 7 Krabicové grafy pre vstupné indikátory podľa zhlukov, rok 2020
(*Zdroj: vlastné spracovanie v SAS Viya)*

Odchod prof. Ing. Hedvigy Bakytovej, PhD.
(* 12-08-1932 † 18-08-2022)
Smútočný oznam

Departure of prof. Ing. Hedviga Bakytová, PhD.
(* 12-08-1932 † 18-08-2022)
Obituary

V lete tohto roku sme sa dozvedeli smutnú správu. Tento svet opustila osobnosť, ktorá sa významnou mierou zaslúžila o pozdvihnutie úrovne výučby a výskumu v oblasti štatistiky na Slovensku. Dlhoročná členka katedry štatistiky a dekanka Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave (FHI EU v Bratislave) prof. Ing. Hedviga Bakytová, PhD., zomrela 18. augusta 2022 v Bratislave, tesne po dožití svojich 90-tich narodenín.



Narodila sa 12. augusta 1932 v Podhradí, okres Prievidza, kde prežila celé svoje detské a mládežnícke roky. Vysokoškolské štúdium ukončila na Vysokej škole ekonomickej v Bratislave v roku 1954. V rokoch 1954 – 1964 pracovala v rôznych funkiách vo sfére obchodu a tiež ako vedúca personalistiky. Na katedru štatistiky nastúpila v roku 1964 ako odborná asistentka s orientáciou na štatistiku obchodu a spotreby. Neskôr, od roku 1979, už ako docentka sa vo väčsej miere zaoberala matematickou štatistikou a viacrozmernými metódami. Patrila k iniciátorom vzniku študijného odboru ekonomická štatistika na fakulte. V roku 1991 bola vymenovaná za profesorku, pričom

v tejto funkcii pôsobila na katedre štatistiky až do roku 2000. V rokoch 1990 – 1997 pôsobila vo funkcii dekanky FHI EU v Bratislave (presnejšie do roku 1992 Fakulty riadenia Vysokej školy ekonomickej v Bratislave) zatiaľ ako jediná žena. Ako dekanka FHI sa významne podieľala na rozpracovaní študijných programov pre novovzniknuté študijné odbory, ako aj na zavedení bakalárskeho štúdia. Z jej iniciatívy sa veľmi výrazne rozvinula medzinárodná spolupráca fakulty s celým radom významných zahraničných univerzít. Bola iniciátorkou vzniku germanofónneho štúdia v spolupráci s univerzitami v Nemecku a Rakúsku a taktiež iniciovala rozvoj spolupráce s Institute of Actuary v Oxforde a City University v Londýne. Vypracovala a koordinovala projekty medzinárodnej spolupráce programu TEMPUS, v rámci ktorého sa zúčastnili výmenných študijných pobytov mnohí učitelia a študenti fakulty.

Pani profesorka pôsobila v pedagogickom procese v rokoch 1964 až 2000, čiže 36 rokov. Bola autorkou resp. spoluautorkou celého radu učebníc a publikácií. Medzi najvýznamnejšie patrí učebnica Základy štatistiky (1975, 1979), ktorá patrila k prvým tohto druhu na Slovensku a využívali ju študenti ekonomických a technických odborov vysokých škôl, ako i odborníci z praxe.

Pani profesorka Bakytová bola jednou z osobností, ktorá sa počas svojho takmer 40-ročného pôsobenia zaslúžila nielen o rozvoj Katedry štatistiky na Fakulte riadenia Vysokej školy ekonomickej v Bratislave, dnešnej Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave, ale aj celej univerzity. Ako dekanka všetku svoju prácu sústredila na dosiahnutie medzinárodnej úrovne štúdia na fakulte a v rámci svojich kompetencií prispela významnou mierou k rozvoju štúdia na FHI EU vo všetkých jeho formách. Za svoju prácu dostala v roku 2017 medailu Rudolfa Briška ako ocenenie, že sa významným spôsobom zaslúžila o rozvoj FHI EU v Bratislave a významne prispela k rozvoju ekonomickej vzdelanosti a vedy na Slovensku.

Pani profesorku Hedvigu Bakytovú som mala česť poznať nielen ako prísnu vyučujúcu, pri ktorej sme spolu s ostatnými spolužiakmi mali strach, či sa nám podarí úspešne urobiť skúšku, ale tiež ako kolegyňu resp. pani dekanku, ktorá mladých aj starších kolegov výrazne podporovala, aby sme vycestovali a zbierali skúsenosti na partnerských zahraničných univerzitách.

Česť jej pamiatke!

Mária Vojtková
Katedra štatistiky FHI EU v Bratislave

21. slovenská štatistická konferencia
Informácia o konaní udalosti SŠDS

21st Slovak Statistical Conference
Information about an event by SŠDS

Slovenská štatistická a demografická spoločnosť (SŠDS) a Štatistický úrad Slovenskej republiky (ŠÚ SR) v roku 2022 usporiadali 21. slovenskú štatistickú konferenciu. Konferencia sa uskutočnila v dňoch 13. – 14. septembra 2022 v Žiline, konkrétne v areáli účelového zariadenia Žilinskej univerzity v Žiline. Po organizačnej stránke konferenciu výborne zvládli pracovníčky Pracoviska ŠÚ SR v Žiline na čele s pani riaditeľkou Ing. Oľgou Chovanovou.

Na podujatí sa stretli zástupcovia štátnej správy, akademickej obce i súkromného sektora v celkovom počte 54 účastníkov. Veľký podiel účastníkov tvorili pracovníci ŠÚ SR (35 účastníkov). Program konferencie sa skladal z 15 vystúpení a bol rozdelený do 4 blokov. Pracovníci ŠÚ SR dodali a predniesli celkovo 5 príspevkov, a tak zabezpečili tretinu programu konferencie.

Na úvod ako tradične, predstavenie kraja v ktorom sa koná konferencia, predniesla riaditeľka Pracoviska ŠÚ SR v Žiline, pani Oľga Chovanová. Venovala sa histórii kraja, zloženiu kraja z okresov, prírodným a kultúrnym zaujímavostiam. Zloženie obyvateľstva, priemysel a poľnohospodárstvo predstavila pomocou rôznych štatistických ukazovateľov.

Vedecká časť konferencie bola rozdelená do štyroch častí. Prvú sekciu príspevkov otvoril V. Páleník zástupca Ekonomického ústavu SAV, s príspevkom Analýza financovania dlhodobej starostlivosti na Slovensku. Využitie kompozitného indexu v oblasti poskytovania sociálnych služieb predstavila A. Kaščáková z Ekonomickej fakulty Univerzity Mateja Bela v Banskej Bystrici. Analýze aktuálnej situácie v odmeňovaní mužov a žien na Slovensku sa venoval F. Foltán zo spoločnosti Trexima. P. Knížat zo ŠÚ SR priblížil využitie web scrapovania dát v súvislosti s výpočtami spotrebiteľského cenového indexu.

Vystúpenia prezentujúcich v druhom bloku súviseli s projektom Sčítania obyvateľov, domov a bytov 2021 (SODB) a odzneli tri príspevky. Hovorkyňa SODB 2021 J. Stauder priblížila propagačnú kampaň realizovanú v súvislosti s prebiehajúcim šírením výsledkov sčítania. Metódu „signs of life“ aplikovanú pri spracovaní výsledkov sčítania obyvateľov odprezentoval M. Kočiš zo ŠÚ SR. Porovnaniu vekovo-pohlavnej štruktúry obyvateľstva v slovenských obciach

s vekovo-pohlavnou štruktúrou štátov sveta sa vo svojom príspevku venoval A. Chromeček zo ŠÚ SR.

Druhý deň konferencie otvoril hosť z Českej republiky, predseda České statistické společnosti pôsobiaci na Univerzitě Palackého v Olomouci, O. Vencálek. V príspevku sa venoval efektivite očkovania proti covid-19. V medicínskej tematike pokračovala K. Uchaľ z Fakulty managementu Univerzity Komenského v Bratislave (UK BA). Zaoberala sa chorobami dýchacích ciest v slovenskej ambulantnej praxi v rokoch 2017 – 2022 v kontexte covid-19. I. Waczulíková, zástupkyňa Fakulty matematiky, fyziky a informatiky UK BA, analyzovala prínos Team-Based Learning a výučby štatistiky pre uplatnenie absolventov biomedicínskej fyziky v praxi. T. Chujac, doktorand z Fakulty managementu UK BA, predniesol príspevok na tému dolovanie asociačných pravidiel z transakčných dát obchodnej spoločnosti pomocou softvéru R.



Obr. 1 Zábery z konferencie

V poslednej 4. časti konferencie odzneli štyri príspevky. Prvý predniesol R. Vlačuha zo ŠÚ SR, kde informoval o dopadoch legislatívy Európskej únie na zmeny vo výberových zisťovaniach v domácnostiach, ktoré realizuje ŠÚ SR (EU-SILC a Rodinné účty). Analýze príjmovej nerovnosti v domácnostiach na základe údajov EU-SILC pomocou kvantilovej regresie sa venovala I. Stankovičová z Fakulty managementu UK BA. R. Herasimau z Ekonomického ústavu SAV sa zaoberal nákladmi na deti na Slovenku. Geografické a demografické determinanty rozvodovosti na Slovensku a v Česku priblížila A. Garajová z Prírodovedeckej fakulty UK BA.

Uskutočnená 21. slovenská štatistická konferencia priniesla zaujímavé príspevky z rôznych oblastí výskumu a opäť zdôraznila význam štatistiky a demografie vo vedeckom bádaní a v praxi. Slovenskí odborníci z tejto oblasti mali príležitosť na opätovné stretnutie, na výmenu poznatkov, informácií a skúseností, na prezentáciu výsledkov svojho výskumu a práce.

Predstavitelia Slovenskej štatistickej a demografickej spoločnosti avizovali, že o dva roky sa v Nitre uskutoční spojená štatistická a demografická konferencia.

Martin Boďa, Iveta Stankovičová
za vedecký a organizačný výbor

Nové orgány SŠDS vo funkčnom období rokov 2023 – 2026 a plán akcií SŠDS na rok 2023

Informácia o konaní udalosti SŠDS

New bodies of SŠDS for the term of office 2023 – 2026, and the plan of events hosted by SŠDS for 2023

Information about an event by SŠDS

Orgány spoločnosti, ktoré vykonávajú svoju činnosť v rokoch 2019 – 2022, boli zvolené na Valnom zhromaždení Slovenskej štatistickej a demografickej spoločnosti (SŠDS) konanom 14. júna 2018 v Častej – Papierničke a ďalej kreované na Výbore SŠDS konanom 6. decembra 2018 v Bratislave. Končiace sa funkčné obdobie v roku 2022 si vyžiadalo voľby do Výboru SŠDS aj Revíznej komisie SŠDS, ktoré sa uskutočnili v septembri 2022. Samotné hlasovanie sa realizovalo elektronicky cez funkcionalitu Google Forms od štvrtku 8. septembra 2022 0:00 h do pondelka 12. septembra 2022 12:00 h. Avízo o hlasovaní bolo cirkulované vnútri SŠDS mailom 7. septembra 2022 a pripomenuté 11. septembra 2022.

Na Valnom zhromaždení SŠDS, ktoré sa konalo prezenčným spôsobom v Žiline 13. septembra 2023, boli prezentované výsledky hlasovania a technické a procedurálne zabezpečenie bolo osvedčené dvoma skrutátorkami zvolenými priamo na Valnom zhromaždení SŠDS. Do volieb sa zapojilo 67 členov SŠDS spôsobom, že každý elektronický hlas bol platný. Výsledky prezentuje tabuľka 1, ktorá udáva tiež počet získaných hlasov. Podľa stanov SŠDS sa členom Výboru SŠDS stáva na základe svojej funkcie aj predseda Štatistického úradu Slovenskej republiky.

Za členov revíznej komisie boli podľa pravidiel hlasovania zvolení Mária Kanderová, Jana Špirková a Alena Mojsejová.

Zloženie Výboru SŠDS a Revíznej komisie SŠDS na funkčné obdobie rokov 2023 – 2026 bolo kreované na zasadnutí Výboru SŠDS, ktoré sa uskutočnilo elektronickým spôsobom cez Microsoft Teams 13. decembra 2022. Na zasadnutí členovia nového Výboru SŠDS podľa spôsobu určeného stanovami spomedzi seba zvolili členov užšieho riadiaceho ramena Výboru SŠDS a určili funkcie predsedu, podpredsedov a hospodára. Rovnako bolo predpísaným spôsobom určená funkcia predsedu Revíznej komisie SŠDS.

Zloženie Výboru SŠDS a Revíznej komisie SŠDS vo funkčnom období rokov 2023 – 2026 aj s určením funkcií predostierajú tabuľky 2 a 3. Mandát členom nového Výboru SŠDS a Revíznej komisie SŠDS vzniká až od 1. januára 2023.

Tab. 1 Výsledky volieb členov Výboru a Revíznej komisie SŠDS na obdobie rokov 2023 – 2026

Výsledky online hlasovania k 12.09.2022 (12:00 h)		
	Počet hlasujúcich:	67
	Počet platných hlasov:	67
	Počet neplatných hlasov:	0
P. č.	Meno a priezvisko kandidáta (pracovisko)	Počet získaných hlasov
1	Branislav Bleha (PriF UK Bratislava)	47
2	Martin Boďa (EF UMB B. Bystrica)	61
3	Renáta Dušová (ŠÚ SR Nitra)	41
4	Helena Glaser-Opitzová (ŠÚ SR ústredie Bratislava)	37
5	Ľudmila Ivančíková (ŠÚ SR ústredie Bratislava)	46
6	Mária Kanderová (EF UMB B. Bystrica)	25
7	Alena Kaščáková (EF UMB B. Bystrica)	27
8	Libuša Kolesárová (ŠÚ SR ústredie Bratislava)	33
9	Ján Kozoň (ŠÚ SR Prešov)	39
10	Dagmar Kusendová (PriF UK Bratislava)	28
11	Ivan Lichner (EÚ SAV Bratislava)	28
12	Janka Medová (UKF Nitra)	32
13	Silvia Megyesiová (PHF EU Košice)	30
14	Alena Mojsejová (EkF TU Košice)	26
15	Viliam Páleník (EÚ SAV Bratislava)	35
16	Marek Radvanský (EÚ SAV Bratislava)	28
17	Mária Stachová (EF UMB B. Bystrica)	28
18	Iveta Stankovičová (FM UK Bratislava)	60
19	Iveta Syneková (ŠÚ SR ústredie Bratislava)	33
20	Jana Špirková (EF UMB B. Bystrica)	25
21	Branislav Šprocha (Infostat VDC Bratislava)	45
22	Anna Tirpáková (UKF Nitra)	30
23	Róbert Vlačuha (ŠÚ SR ústredie Bratislava)	30
24	Mária Vojtková (FHI EU Bratislava)	36
25	Iveta Waczulíková (FMFI UK Bratislava)	27
26	Tomáš Želinský (EkF TU Košice)	42

Tab. 2 Výbor SŠDS na obdobie rokov 2023 – 2026

P.č.	Priezvisko a meno	Pracovisko	Funkcia
1	Bleha Branislav	PriF UK Bratislava	podpredseda SŠDS pre demografiu
2	Boďa Martin	EF UMB B. Bystrica	vedecký tajomník, podpredseda SŠDS pre medzinárodné vzťahy
3	Dušová Renáta	ŠÚ SR Nitra	podpredsedníčka SŠDS pre štátnu štatistiku
4	Glaser-Opitzová Helena	ŠÚ SR ústredie BA	
5	Ivančíková Ľudmila	ŠÚ SR ústredie BA	
6	Kaščáková Alena	EF UMB B. Bystrica	
7	Kolesárová Libuša	ŠÚ SR ústredie BA	
8	Kozoň Ján	ŠÚ SR Prešov	
9	Kusendová Dagmar	PriF UK Bratislava	
10	Lichner Ivan	EU SAV Bratislava	
11	Medová Janka	UKF Nitra	
12	Megyesiová Silvia	PHF EU Košice	
13	Páleník Viliam	EU SAV Bratislava	
14	Radvanský Marek	EU SAV Bratislava	
15	Stachová Mária	EF UMB B. Bystrica	
16	Stankovičová Iveta	FM UK Bratislava	predsedníčka SŠDS
17	Syneková Iveta	ŠÚ SR ústredie BA	hospodárka
18	Šprocha Branislav	Infostat VDC BA	
19	Tirpáková Anna	UKF Nitra	
20	Vlačuha Róbert	ŠÚ SR ústredie BA	
21	Vojtková Mária	FHI EU Bratislava	podpredsedníčka SŠDS pre akademickú štatistiku
22	Waczulíková Iveta	FMFI UK Bratislava	podpredsedníčka SŠDS pre medicínsku štatistiku
23	Želinský Tomáš	EkF TU Košice	
24	Peťko Peter	ŠÚ SR ústredie BA	člen výboru z titulu funkcie

Tab. 3 Revízna komisia SŠDS na obdobie rokov 2023 – 2026

P.č.	Priezvisko a meno	Pracovisko	Funkcia
1	Kanderová Mária	EF UMB B. Bystrica	predsedníčka RK SŠDS
2	Špirková Jana	EF UMB B. Bystrica	členka RK SŠDS
3	Mojsejová Alena	EkF TU Košice	členka RK SŠDS

Plán akcií SŠDS v roku 2023 bol koncipovaný na základe konzultácií s tradičnými garantmi organizovaných podujatí. Implementuje už víziu predostrenú na Valnom zhromaždení SŠDS 13. septembra 2022 o spojení Slovenskej štatistickej konferencie a Slovenskej demografickej konferencie, ktoré sú obvykle organizované rotačne na bienálnej báze. Podľa tohto zámeru sa budú obe konferencie konať ten istý rok v štýle dvojkonferencií, ako to bolo

v roku 2021. Druhá veľká dvojkonferencia, ktorá bude 22. slovenskou štatistickou konferenciou a 19. slovenskou demografickou konferenciou (22SSK&19SDK), sa uskutoční až v roku 2024. Predpokladá sa, že sa uskutoční v Nitre s pomocou Pracoviska Štatistického úradu Slovenskej republiky v Nitre. Plán akcií SŠDS na rok 2023 je súčasťou tabuľky 3.

Tab. 3 Plán akcií SŠDS v roku 2023

Termín konania	Názov podujatia
apríl 2023	Pohľady na ekonomiku Slovenska 2023 (online: 3 – 4 prednášky)
máj 2023	EKOMSTAT 2023, Trenčianske Teplice
jún 2023	PRASTAN (JMFS + SŠDS), Kočovce
07. - 08.12.2023	30. medzinárodný vedecký seminár Výpočtová štatistika, Bratislava
07. - 08.12.2023	Prehliadka prác mladých štatistikov a demografov, Bratislava
08.12.2023	Analytika očami profesionálov, Bratislava
priebežne	regionálne akcie: diskusné popoludnia, prednášky

Martin Boďa
vedecký tajomník SŠDS

Nitrianske štatistické dni 2022 v novom formáte
Informácia o konaní udalosti SŠDS

Nitra Statistical Days 2022 in a new format
Information about an event by SŠDS

Na jeseň 2022 sa po trojročnej prestávke spôsobenej opatreniami v súvislosti s pandémiou COVID19 v Nitre uskutočnilo už tradičné podujatie s názvom Nitrianske štatistické dni, po prvýkrát v novom formáte. Tradičnú konferenciu nahradilo pásmo prednášok a workshopov. Niektoré z nich sa realizovali v hybridnej, iné v on-line forme. Podujatie, tradične organizované Katedrou matematiky Fakulty prírodných vied a informatiky Univerzity Konštantína Filozofa v Nitre (FPVaI UKF) v spolupráci so Slovenskou štatistickou a demografickou spoločnosťou a Pracoviskom Štatistického úradu Slovenskej republiky v Nitre, sa otvorilo v pondelok 14. novembra 2022 na Katedre matematiky. Prezenčne boli prítomní a konferenciu otvorili dekan FPVaI UKF v Nitre prof. RNDr. František Petrovič, PhD., vedúca Katedry matematiky FPVaI UKF v Nitre doc. PaedDr. Gabriela Pavlovičová, PhD., vedúca programového výboru NŠD 2022 doc. PaedDr. Janka Medová, PhD., ako aj zástupkyňa Pracoviska Štatistického úradu Slovenskej republiky v Nitre Mgr. Mária Rybanská. Za Slovenskú štatistickú a demografickú spoločnosť privítala účastníkov on-line jej predsedníčka doc. Ing. Iveta Stankovičová, PhD.

Prvou prednášajúcou podujatia bola assoc. prof. Sibel Kazak, Ph.D., podpredsedníčka medzinárodnej asociácie pre vyučovanie štatistiky International Association of Statistics Education (IASE), didaktickej sekcie International Statistical Institute (ISI), z Middle East Technical University v tureckej Ankare. V prednáške s názvom Statistical investigation process with image-based data predstavila nezvyčajnú aktivitu pre budúcich učiteľov matematiky.

Nitrianske štatistické dni 2022 pokračovali 22. novembra 2022 prednáškou RNDr. Ľubomíra Rybanského, PhD. z Katedry matematiky FPVaI UKF v Nitre s názvom Teória odpovedí na položku a jej aplikácie (nielen) vo vyučovaní matematiky. Prednášajúci predstavil aktuálne najviac využívanú alternatívu tradičnej teórie testov, analýzu odpovedí na položku (Item Response Theory, IRT) a ilustroval jej využitie na dátach z dvoch rôznych výskumov.

Dňa 24. novembra 2022 predniesol prof. RNDr. Anatolij Dvurečenskij, DrSc. tretiu prednášku cyklu s názvom Teória pravdepodobnosti: Od hazardných hier ku kvantovej mechanike, v ktorej prehľadne popísal históriu a vývoj teórie pravdepodobnosti, od prvých aplikácií v hazardných hrách, cez unikátnu prácu

Kolmogorova, ktorý sformalizoval pravdepodobnosť pomocou množín až po aktuálne výsledky v tomto odbore. Nielen o štatistickej a lingvistickej analýze príhovorov československých a českých prezidentov pojednal doc. Mgr. Ján Mačutek, PhD. z Katedry matematiky FPVaI UKF v Nitre v prednáške Aplikácie matematickej štatistiky v lingvistike.

Prvé decembrové popoludnie patrilo dvojici prednášajúcich zo Štatistického úradu Slovenskej republiky: Jasmíny Stauder, hovorkyne pre Sčítanie obyvateľov, domov a bytov 2021 a Mgr. Lucie Vanišovej, vedúcej oddelenia sčítania obyvateľov, domov a bytov a prierezových štatistík. Počas workshopu Čo nám hovoria dáta zo SOBD 2021 predstavili účastníkom, aké dáta zo SOBD 2021 sú dostupné a akým spôsobom k nim pristupovať. Súčasťou boli aj ukážky výberov z databáz Štatistického úradu SR, ako aj priestor na diskusiu, otázky a podnety účastníkov.

Dvojica workshopov plánovaná na 1. a 15. decembra 2022 pod vedením Dr. Edyty Nowinskej z Osnabrück University v Nemecku bola pre zdravotné problémy prednášajúcej zrušená.

Všetky uskutočnené prednášky rozprúdili diskusiu či už medzi prítomnými v Nitre alebo medzi on-line účastníkmi. Zmena formátu prispela k zvýšeniu počtu účastníkov. Na podujatie sa zaregistrovalo viac ako 250 účastníkov a na každej z uskutočnených prednášok či workshopov sa zúčastnilo viac ako 50 účastníkov z radov akademikov, študentov, ale aj účastníkov z praxe.

Janka Medová
predsedníčka programového výboru



Obr. 2 Otvorenie podujatia Nitrianske štatistické dni 2022

Analytika očami profesionálov
Informácia o konaní udalosti SŠDS

Analytics through the eyes of professionals
Information about an event by SŠDS

Slovenská štatistická a demografická spoločnosť (SŠDS) pravidelne už niekoľko rokov začiatkom decembra organizuje medzinárodný vedecký seminár *Výpočtová štatistika*. Zorganizovali sme do roku 2019 spolu 28 ročníkov. Tradíciu seminárov založil Jozef Chajdiak (Ekonomická univerzita v Bratislave) a potom organizáciu prevzala Iveta Stankovičová (Fakulta managementu Univerzity Komenského v Bratislave), ktorá ich organizuje do dnešných dní. Posledné roky sa však seminár neuskutočnil kvôli covidovej pandémii. V roku 2022 sme chceli pokračovať v tradícii, ale záujem bol nízky, tak sme opäť museli akciu zrušiť. V rámci seminára od roku 2010 sme poriadali aj *Prehliadku prác mladých štatistikov a demografov* pre študentov vysokých škôl zo SR a ČR, čo sa tiež nepodarilo uskutočniť pre nízky počet prihlásených študentských prác. Zorganizovali sme však pásmo prednášok pod názvom *Analytika očami profesionálov 2022*.

Tento rok SŠDS usporiadala prvýkrát pásmo prednášok pre študentov *Analytika očami profesionálov* v spolupráci s Fakultou managementu Univerzity Komenského v Bratislave. Akcia sa uskutočnila prezenčnou formou 9. decembra 2022 na tejto fakulte. Zúčastnilo sa jej 65 účastníkov, prevažne študentov Fakulty managementu, ale aj študentov z Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave. Pohľad na uskutočnené prednáškové pásmo poskytuje a ich atmosféru dokumentuje obrázok 1.

Program sa skladal z troch prednášok odborníkov na analýzu údajov z praxe. Mená prednášajúcich a názvy ich prezentácií sú:

1. Milan Schnorrer: Použitie dát v maloobchode vo svete.
2. Peter Fusek, Michal Šuran: Big dáta v každodennom živote.
3. Lucia Čekaňáková: Život data scientistu alebo načo je tá štatistika vlastne dobrá?

Milan Schnorrer je expertom na Customer Engagement & customer data. Študoval na Fakulte managementu Univerzity Komenského v Bratislave (Bc. stupeň) a v Holandsku na Univerzite v Groningene (Mgr. stupeň). Je CEO a zakladajúcim partnerom analytickej firmy Pricewise, dnes pod novým názvom Voxwise (www.voxwise.com). Má 14-ročné skúsenosti v oblasti analýzy údajov

o zákazníkoch. Podieľal sa na návrhu a vypracovaní 10 CRM riešení, stratégií zapájania zákazníkov s kvantifikovanými dopadmi na výnosy firmy. Má rozsiahle skúsenosti v oblasti dizajnu ciest zákazníka (CX). V súčasnosti je členom poradnej rady a hosťujúcim lektorom na Univerzite v Groningene, Holandsko. Podieľa sa aj na skvalitňovaní výučby na Fakulte managementu a Fakulte matematiky, fyziky a informatiky Univerzity Komenského v Bratislave.



Obr. 3: Analytika očami profesionálov - pohľady do prednáškovej sály na prednášajúcich a ich poslucháčov

Peter Fusek je podnikateľ, špecialista na go-to-market a budovanie technologických firiem. V súčasnosti pomáha v oblasti budovania produktov spoločnosti Ataccama. Pri stavbe svojich biznisov zapája mladé talenty už zo stredných škôl. Ako absolvent magisterského programu finančný manažment na

Fakulte managementu Univerzity Komenského v Bratislave, už viac ako 20 rokov pôsobí v bankovom a telekomunikačnom sektore. Najskôr pracoval vo vnútri korporácií, neskôr už ako zakladateľ a výkonný manažér v rámci vlastných startupov (Instarea, DingoDot). Viaceré produkty uspeli u globálnych korporátnych klientov v rámci EMEA regiónu, napríklad Slisor v Saudskej Arábii. B2B, SaaS služby, Market Locator aj Callinspector boli akvirované spoločnosťou Slovak Telekom patriacej do Deutsche Telekom Group a sú prevádzkované doteraz v rámci CEE regiónu.

Michal Šuran pôsobí v TELCO biznise, v ktorom sa pohybuje už od ukončenia štúdií na FM UK pred takmer 25 rokmi. Snaží sa kombinovať IT znalosti a marketing, aj keď dominanciu v doterajšej kariére mal práve marketing. Prešiel si aj marketingovou komunikáciou na strane reklamných agentúr a následne aj na strane zadávateľov v Slovak Telekome, oblasťou riadenia vzťahov s verejnosťou počas digitalizácie televízneho vysielania, biznis developmentom až po sales v Market Locatore. Aktuálne so svojou firmou Sense4 pôsobí ako biznis-technický partner Slovak Telekomu a v spoločnosti Orange v Big Data projekte Market Locator.

Lucia Čekaňáková je Business Consultant poradenskej a softvérovej spoločnosti SAS Slovakia, s.r.o. od roku 2015, kde pôsobí ako analytička a pracuje s klientami z oblasti finančného sektora, telekomunikácií, retailu, výroby a verejnej správy. Je absolventkou bakalárskeho stupňa na Ekonomickej univerzite v Bratislave.

Po každej prednáške prebehla bohatá diskusia. Študentov témy prezentácií zaujali a položili viac ako 30 otázok. Cieľom akcie bolo, je aj bude, aby sme študentov manažmentu a ekonómie motivovali k štúdiu kvantitatívnych metód, ktoré sa dnes v praxi používajú na analýzy množstva údajov, ktoré má každá firma k dispozícii. Výsledky analýz sa potom používajú v manažmente aj marketingu.

Iveta Stankovičová
predsedníčka organizačného výboru

Publikačná etika a vyhlásenie k nekalým praktikám

Redakčná rada Forum Statisticum Slovacum (ďalej len RR FSS) sa v plnom rozsahu stotožňuje s princípmi etiky publikovania deklarovanými Výborom pre pulikačnú etiku. RR FSS prijíma štandardy publikačnej etiky a zavádza opatrenia proti akejkoľvek publikácii s nekalými praktikami. Autori predkladajú svoje práce do časopisu na publikovanie s vyhlásením, že predložené práce sú príspevky autorov a neboli kopírované alebo plagizované z iných prác. Editori posudzujú rukopisy na základe ich intelektuálneho obsahu, bez ohľadu na rasu, pohlavie, sexuálnu orientáciu, náboženské vyznanie, etnický pôvod, štátnu príslušnosť alebo politickú filozofiu autorov. RR FSS očakáva, že všetky strany podieľajúce sa na FSS dodržia publikačnú etiku. RR FSS nebude tolerovať plagiátorstvo alebo iné neetické správanie a neuverejní rukopis, ktorý nespĺňa tieto normy.

Zodpovednosť autorov: Autori potvrdzujú, že ich rukopis je ich pôvodnou prácou, že nebol doteraz uverejnený v rovnakej podobe a zároveň nie je v súčasnej dobe podaný na uverejnenie inde. Autori musia bezodkladne oznámiť RR FSS všetky strety záujmov. Autori musia uviesť všetky zdroje použité pri tvorbe svojho rukopisu. Autori musia bezodkladne nahlásiť všetky chyby, ktoré objavia vo svojom rukopise RR FSS.

Zodpovednosť recenzentov: Recenzenti musia oznámiť RR FSS všetky strety záujmov. Recenzenti musia uchovávať informácie týkajúce sa rukopisu ako dôverné. Recenzent musí upozorniť predsedu RR FSS na informácie, ktoré môžu byť dôvodom na zamietnutie vydania rukopisu. Recenzenti posudzujú rukopisy len na základe ich intelektuálneho obsahu.

Zodpovednosť RR FSS: RR FSS musí uchovávať informácie o podaných rukopisoch ako dôverné. RR FSS musí posúdiť rukopisy len na základe ich intelektuálneho obsahu. RR FSS rozhoduje o zverejnení predložených rukopisov. RR FSS odmietne príspevok, ktorý nie je v súlade s požiadavkami etiky publikovania.

Publication ethics and malpractice statement

The Editorial Board of Forum Statisticum Slovacum (hereinafter abbreviated as EB/FSS) is fully associated with the principles of ethics of publication declared by the Committee of Publication Ethics. The EB/FSS is committed to upholding the highest standards of publication ethics and takes all possible measures against any publication malpractices. Authors submitting their works to the journal for publication as original articles attest that the submitted works represent their authors' contributions and have not been copied or plagiarized in whole or in part from other works. An editor at any time evaluate manuscripts for their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors. The EB/FSS expects all parties participating in the publication of Forum Statisticum Slovacum commit to these publication ethics. The EB/FSS does not tolerate plagiarism or other unethical behaviour and will remove any manuscript that does not meet these standards.

Author responsibilities: Authors certify that their manuscripts are their original work unpublished previously elsewhere in the same form and not currently being considered for publication elsewhere. Authors must notify the EB/FSS of any conflicts of interest without any delay. Authors must identify all sources used in the creation of their manuscript. Authors must report any errors they discover in their manuscript to the EB/FSS without any delay.

Reviewer responsibilities: Reviewers must notify the EB/FSS of any conflicts of interest. Reviewers must keep information pertaining to the manuscript confidential. Reviewers must bring to the attention of the head of EB/FSS any information that may be reason to reject publication of a manuscript. Reviewers must evaluate manuscripts only for their intellectual content.

Editorial responsibilities: The EB/FSS must keep information pertaining to submitted manuscripts confidential. The EB/FSS must evaluate manuscripts only for their intellectual content. The EB/FSS is responsible for making publication decisions for submitted manuscripts. EB/FSS refuses the manuscript, which is not in accordance with the requirements of publishing ethics.

Obsah / Table of contents

Odborné a vedecké články / Instructive and regular papers

Bayesovské modelování dat v prostředí programového systému SAS
Bayesian data modelling in the SAS software system environment
Roman Pavelka 1

Viacrozmerná analýza vybraných ukazovateľov udržateľné-ho rozvoja v krajinách EÚ
Multivariate analysis of selected indicators of sustainable development in EU countries
Zuzana Moscoso Borsík, Iveta Stankovičová 13

% = odborný článok / instructive paper

Zo života SŠDS / From the life of SŠDS

Odchod prof. Ing. Hedvigy Bakytovej, PhD. (* 12-08-1932 † 18-08-2022) [smútočný oznam]
Departure of prof. Ing. Hedviga Bakytová, PhD. (* 12-08-1932 † 18-08-2022) [obituary]
Mária Vojtková 35

21. slovenská štatistická konferencia [informácia o konaní udalosti SŠDS]
21st Slovak Statistical Conference [information about an event by SSDS]
Martin Boďa, Iveta Stankovičová 37

Nové orgány SŠDS vo funkčnom období rokov 2023 – 2026 a plán akcií SŠDS na rok 2023 [informácia o konaní udalosti SŠDS]
New bodies of SŠDS for the term of office 2023 – 2026, and the plan of events hosted by SŠDS for 2023 [information about an event by SSDS]
Martin Boďa 40

Nitrianske štatistické dni 2022 v novom formáte [informácia o konaní udalosti SŠDS]
Nitra Statistical Days 2022 in a new format [information about an event by SSDS]
Janka Medová 44

Analytika očami profesionálov [informácia o konaní udalosti SŠDS]
Analytics through the eyes of professionals [information about an event by SSDS]
Iveta Stankovičová 46

ISSN 1336-7420

Cena / price: 25 €
Ročné predplatné / annual subscription: 50 €
Publikované / published: December 2022