

Analýza demografických procesov v okresoch SR

Ing. Iveta Stankovičová, PhD.

Katedra informačných systémov
Fakulta managementu, Univerzita Komenského v Bratislave

Abstrakt a ciele práce:

Predložený príspevok analyzuje demografickú situáciu v jednotlivých okresoch (regiónoch) SR. Ide o syntetický pohľad na demografickú situáciu v 79-tich okresoch SR na základe 16 vybraných ukazovateľov za roky 2001-2005.

Cieľom práce je viacrozmerná analýza demografických procesov v regiónoch (okresoch) Slovenska. Pokúsili sme sa nájsť „podobné“ regióny v demografickom správaní obyvateľstva na základe vybraných demografických ukazovateľov. Na dosiahnutie vytýčeného cieľa sme použili viacrozmerné štatistické metódy, ktoré poskytuje systém SAS vo svojom grafickom užívateľskom rozhraní SAS Enterprise Guide 4.1 v ponuke *Analyze – Multivariate*, konkrétne faktorovú a zhlukovú analýzu (*Factor Analysis, Cluster Analysis*).

Na analýzu sme použili priemerné hodnoty jednotlivých ukazovateľov za obdobie rokov 2001 až 2005 pre jednotlivé okresy SR, aby sme eliminovali náhodné výkyvy hodnôt, hlavne v okresoch s nízkym počtom obyvateľov. Vybrané ukazovatele charakterizujú 4 demografické procesy: sobášnosť, rozvodovosť, plodnosť a migráciu. Údaje o vymenovaných demografických procesoch sú doplnené priemerným vekom týchto procesov a sú sledované podľa pohlavia.

Počet štatistických jednotiek: 79 okresov SR.

Počet analyzovaných premenných: 16.

(Zoznam ukazovateľov pozri v tabuľke 1).

Tabuľka 1: Zoznam 16 analyzovaných ukazovateľov

Por. číslo	Premenná	Popis
1.	us_m	úhrnná sobášnosť mužů
2.	sr_m	štandardizovaná rozvodovosť mužů
3.	us	úhrnná sobášnosť ženy
4.	sr	štandardizovaná rozvodovosť ženy
5.	up	úhrnná plodnosť žien
6.	prv_m	priemerný vek mužů (všeobecne)
7.	prv	priemerný vek ženy (všeobecne)
8.	prvs_m	priemerný vek pri prvom sobáši mužů
9.	prvr_m	priemerný vek rozvode mužů
10.	prvs	priemerný vek pri prvom sobáši ženy
11.	prvr	priemerný vek rozvode ženy
12.	prvp	priemerný vek pri prvom pôrode ženy
13.	saldo_m	migračné saldo v SR mužů
14.	saldo_	migračné saldo v SR ženy
15.	saldoz_m	migračné saldo so zahraničím mužů
16.	saldoz	migračné saldo so zahraničím ženy

Podrobnejší popis:

Zo štatistického hľadiska ide o viacrozmerný problém, preto použijeme viacrozmerné metódy. Pretože analyzovaných ukazovateľov je veľa a sú závislé, ako prvú sme použili metódu faktorovej analýzy na vytvorenie nižšieho počtu významných skrytých (latentných) ukazovateľov – faktorov. Významné faktory sme potom použili ako vstupy do zhlukovej analýzy. Metódou hierarchickej zhlukovej analýzy sme sa pokúsili vytvoriť „podobné“ regióny v demografickom správaní obyvateľstva.

Ako prvý krok sme vypočítali korelačnú maticu ukazovateľov. Na základe toho, že v korelačnej matici sa nachádzali významné koeficienty korelácie medzi ukazovateľmi a na základe vysokej celkovej KMO miery pre údaje (*Kaiser's Measure of Sampling Adequacy: Overall MSA = 0.7397*), sme mohli konštatovať, že dáta sú vhodné pre faktorovú analýzu. Pri faktorovej analýze sme na odhad faktorových váh použili rôzne metódy a zistili sme, že dávajú približne rovnaké výsledky. V príspevku prezentujeme výsledky metódou PCA.

Zo 16 faktorov sme pre následnú zhlukovú analýzu okresov SR vybrali len 4 faktory. Aplikovali sme Kaiserovo pravidlo, že ďalej použijeme len toľko faktorov, koľko je vlastných čísel väčších ako ich priemer, čiže 1 (pri výpočtoch faktorových skóre z korelačnej matice, Tabuľka 2). Tieto 4 významné faktory vysvetľujú

až **87,2%** z celkovej variability údajov. Na faktory sme použili ortogonálnu varimax rotáciu, aby sme mohli faktory lepšie interpretovať. Výsledné faktorové váhy (Tabuľka 3) sú vlastne koeficienty korelácie medzi príslušným faktorom a pôvodnými ukazovateľmi. Na základe ich veľkosti môžeme faktory (skryté premenné) interpretovať. Naša interpretácia 4 významných faktorov je nasledovná:

- 1. faktor** je významne korelovaný s ukazovateľmi priemerný vek pri prvom sobáši a priemerný vek pri prvom pôrode ženy (prvs_m, prvs, prvp). Je to faktor veku pri prvom sobáši a pôrode.
- 2. faktor** je významne kladne korelovaný s ukazovateľmi úhrnnej sobášnosti (us_m, us) a záporne s ukazovateľmi štandardizovanej rozvodovosti (sr_m, sr). Je to faktor sobášnosti v protiklade s rozvodovosťou.
- 3. faktor** je faktorom priemerného veku pri rozvode (prvr_m, prvr).
- 4. faktor** je faktorom migrácie, ale len vnútroštátnej (saldo_m, saldo).

Tabuľka 2: Vlastné čísla korelačnej matice (metóda PCA)

Eigenvalues of the Correlation Matrix: Total=16, Average=1				
	Eigenvalue	Difference	Proportion	Cumulative
1	7.817	4.816	0.489	0.489
2	3.000	0.972	0.188	0.676
3	2.028	0.928	0.127	0.803
4	1.100	0.378	0.069	0.872
5	0.722	0.281	0.045	0.917
6	0.440	0.148	0.028	0.944
7	0.293	0.021	0.018	0.963
8	0.272	0.099	0.017	0.980
9	0.173	0.113	0.011	0.990
10	0.059	0.028	0.004	0.994
11	0.031	0.006	0.002	0.996
12	0.026	0.008	0.002	0.998
13	0.018	0.009	0.001	0.999
14	0.010	0.000	0.001	0.999
15	0.009	0.007	0.001	1.000
16	0.003		0.000	1.000

Tabuľka 3: Faktorové váhy pre 4 významné faktory

Rotated Factor Pattern				
Variable	Factor1	Factor2	Factor3	Factor4
us_m	0.007	0.962	0.038	-0.096
us	-0.069	0.943	-0.057	-0.126
sr_m	0.391	-0.514	0.283	0.009
sr	0.326	-0.577	0.210	0.038
up	-0.433	0.422	-0.063	-0.044
prv_m	0.440	-0.360	0.150	0.074
prv	0.381	-0.344	0.130	0.088
prvs_m	0.828	-0.203	0.297	0.056
prvs	0.868	-0.088	0.368	0.007
prvr_m	0.370	-0.056	0.882	-0.186
prvr	0.391	-0.029	0.860	-0.261
prvp	0.880	0.036	0.296	0.053
saldo_m	-0.025	-0.110	-0.176	0.970
saldo	0.101	-0.090	-0.123	0.970
saldoz_m	0.489	-0.025	0.242	0.096
saldoz	0.074	-0.040	-0.023	0.148

Štyri významné faktory sme použili ako vstupné premenné do hierarchickej zhlukovej analýzy. Ako zhlukovací postup sme zvolili Wardovu metódu. Na základe analýzy dendrogramu sme sa rozhodli vytvoriť 9 výsledných zhlukov, t.j. regionálnych typov demografického správania v SR.

Záver – interpretácia zhlukov:

Na interpretáciu výsledných 9-tich zhlukov je vhodné použiť mapu (obr. 1), Box-ploty (obr. 2) pre jednotlivé faktory a popisné štatistiky pôvodných demografických ukazovateľov. Zoznam okresov podľa jednotlivých zhlukov uvádzame v tabuľke 4.

Okresy zaradené do zhlukov 7 a 8 patria medzi okresy s najvyšším faktorom 1, t.j. najvyšším vekom pri prvom sobáši (zhluk 7: 29,6 muži a 27,3 ženy; 8. zhluk: 28,7 muži a 26,4 ženy) a najvyšším vekom pri prvom pôrode (27,5 pre zhluk 7 a 26,4 pre zhluk 8). Patria sem všetky mestské časti Bratislavy a Košíc. V okresoch zhluku 8 je tiež najvyšší priemerný vek pri rozvodoch (41 muži a 39,1 ženy – t.j. faktor 3) a vysoké záporné saldo vnútornej migrácie (muži -6,5 a ženy -5,9 - t.j. faktor 4).

Do zhluku 6 patria južné okresy stredného Slovenska. Tento zhluk sa vyznačuje najnižšou úrovňou faktora 1, t.j. najnižším vekom pri prvom sobáši (muži 26,9 a ženy 24,2) a pri prvom pôrode (23,3) a tiež pomerne nízkym priemerným vekom pri rozvodoch (38,4 muži a 36,4 ženy – t.j. faktor 3). Je tu aj veľký rozpor medzi sobášnosťou, ktorá je na úrovni priemeru SR (cca 0,4) a rozvodovosťou, ktorá je vysoká (priemer SR je 3,9 a priemer zhluku 6 je 4,5).

Do zhluku 9 sa zaradil len okres Senec. Spôsobil to silný vplyv faktora 4, t.j. vnútorná migrácia. V tomto okrese je nadpriemerne vysoké kladné vnútorné migračné saldo (16,6 muži a 15,6 ženy; priemer SR: muži - 0,15 a ženy -0,25), ktoré je spôsobené výstavbou rodinných domov v lokalitách blízkyh Bratislave. Interpretácia výsledkov je skrátaná pre nedostatok priestoru a je zameraná len na extrémne hodnoty.

Tabuľka 4: Zoznam okresov SR podľa zhlukov

zhluk 1 n = 16	zhluk 2 n = 13	zhluk 3 n = 15	zhluk 4 n = 11	zhluk 5 n = 8	zhluk 6 n = 6	zhluk 7 n = 5	zhluk 8 n = 4	zhluk 9 n = 1
Banská Bystrica	Banská Štiavnica	Dunajská Streda	Bánovce nad Bebravou	Bardejov	Krupina	Bratislava II	Bratislava I	Senec
Dolný Kubín	Brezno	Galanta	Bytča	Humenné	Poltár	Bratislava III	Bratislava V	
Ilava	Detva	Košice-okolie	Čadca	Kežmarok	Revúca	Bratislava IV	Košice II	
Liptovský Mikuláš	Gelnica	Levice	Kysucké Nové Mesto	Medzilaborce	Rimavská Sobota	Košice I	Košice III	
Martin	Hlohovec	Lučenec	Levoča	Poprad	Rožňava	Košice IV		
Myjava	Komárno	Malacky	Námestovo	Stará Ľubovňa	Veľký Krtíš			
Nové Mesto nad Váhom	Nové Zámky	Michalovce	Sabinov	Stropkov				
Piešťany	Partizánske	Nitra	Senica	Svidník				
Považská Bystrica	Prievidza	Pezinok	Snina					
Prešov	Šaľa	Skalica	Topoľčany					
Púchov	Zvolen	Sobrance	Vranov nad Topľou					
Ružomberok	Žarnovica	Spišská Nová Ves						
Trenčín	Žiar nad Hronom	Trebišov						
Trnava		Turčianske Teplice						
Tvrdošín		Zlaté Moravce						

Použité SAS moduly a nástroje:

SAS/STAT, SAS Enterprise Guide 4.1 - procedúry CORR, FACTOR, CLUSTER

Zdroj údajov:

Výskumné demografické centrum, INFOSTAT Bratislava, <http://www.infostat.sk/vdc/>

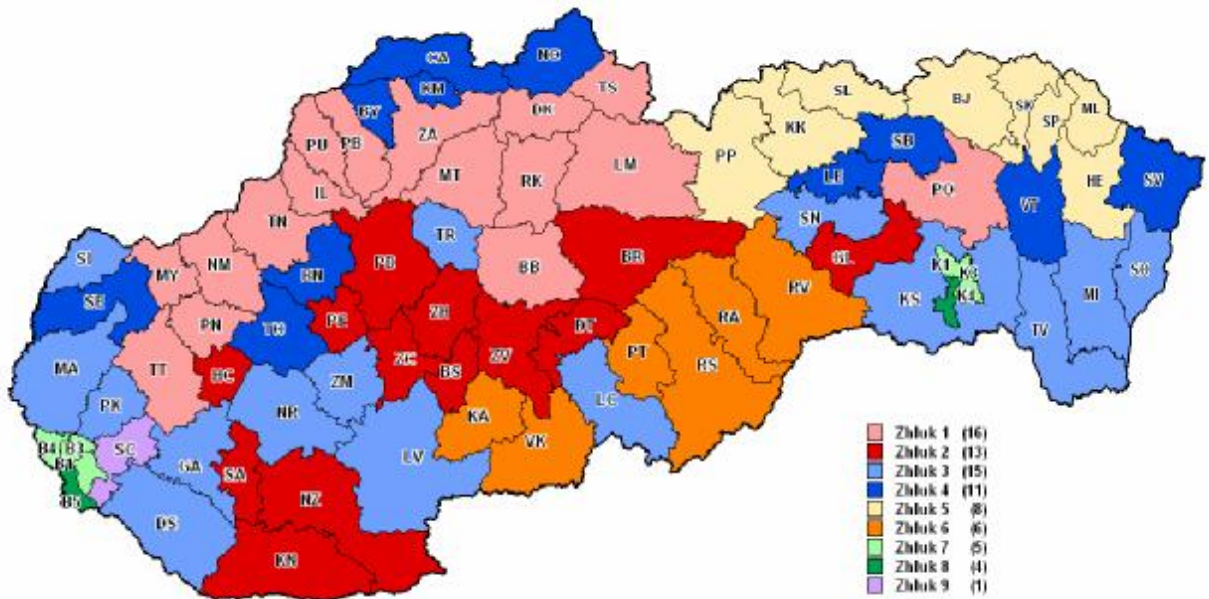
Literatúra a referencie:

- [1.] JURČOVÁ, D. – MESZÁROŠ, J. - STANKOVIČOVÁ, I.: Populačný vývoj v okresoch SR, In: Sborník príspevku XXXVII. demografické konferencie ČDS, Olomouc 23.-24.5.2007. s. 118-123. Praha: ČDS, 2007.
- [2.] VOJTKOVÁ, M.: Územná klasifikácia SR podľa ukazovateľov charakterizujúcich intenzitu pohybu obyvateľstva. In.: 10. Slovenská demografická konferencia. Naša demografia súčasnosť a perspektívy. Smolenice, s. 217-220, Bratislava: SŠDS, 2005.

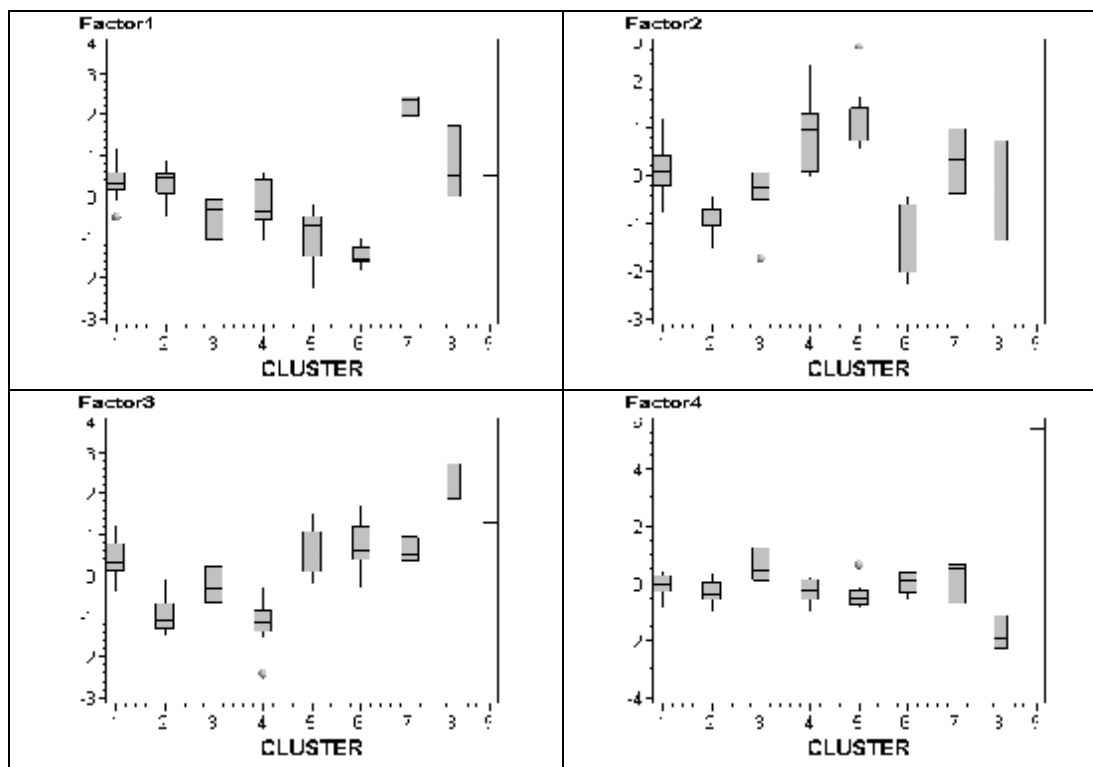
Kontakt autori:

Ing. Iveta Stankovičová, PhD., e-mail: iveta.stankovicova@fm.uniba.sk
Katedra informačných systémov, Fakulta managementu UK v Bratislave
Odbojárov 10, 820 05 Bratislava

Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/2631/05 "Analýza možností aplikácie viacrozmerných štatistických metód na skúmanie ekonomických výsledkov na príklade priemyslu SR prípadne iných oblastí ekonomiky".



Obrázok 1: Výsledná mapa pre 9 zhlukov okresov Slovenska



Obrázok 2: Box-ploty pre jednotlivé faktory podľa zhlukov