UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

ŠŤASTIE A POKUSY O JEHO MATEMATICKÉ VYJADRENIE

BAKALÁRSKA PRÁCA

2014 Jozef JANOČKO

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

ŠŤASTIE A POKUSY O JEHO MATEMATICKÉ VYJADRENIE

BAKALÁRSKA PRÁCA

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	Mgr. Zuzana Bučková

Bratislava 2014 **Jozef JANOČKO**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Jozef Janočko
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský

Názov: Šťastie a pokusy o jeho matematické vyjadrenie / *Happiness and attempts to mathematical expression*

Cieľ: Prehľad rôznych prístupov výpočtu indexu šťastia, analýza dát.

Vedúci: Mgr. Zuzana Bučková
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.

Dátum zadania: 18.10.2013

Dátum schválenia: 14.11.2013

doc. RNDr. Margaréta Halická, CSc.
garant študijného programu

.....
študent

.....
vedúci práce

Poďakovanie Touto cestou sa chcem poďakovat svojej vedúcej bakalárskej práce Mgr. Zuzane Bučkovej za možnosť podieľať sa na výskume veľmi zaujímavej problematiky, taktiež za podnetné pripomienky, ktoré mi pomohli pri písaní tejto práce. Veľmi oceňujem aj odborné rady od Mgr. Jána Somorčíka, PhD. v oblasti štatistického výskumu a od Doc. RNDr. Margaréty Halickej, CSc. pri využití DEA modelovania. Ďakujem aj svojej rodine a priateľom za ich trpezlivosť a podporu.

Abstrakt

JANOČKO, Jozef: Šťastie a pokusy o jeho matematické vyjadrenie [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: Mgr. Zuzana Bučková, Bratislava, 2014, 68 s.

Predložená bakalárska práca sa zaoberá rôznymi prístupmi merania šťastia. Takéto merania sa totiž v súčasnosti stávajú kľúčovými pre smerovanie štátov, nakoľko sú komplexnejšími ukazovateľmi ako HDP. Práca podrobne analyzuje Happy Planet Index (HPI) a taktiež navrhuje vlastný index šťastia, nazývaný Index šťastia jednotlivca. Happy Planet Index je globálnym meraním zahŕňajúcim blahobyt obyvateľstva, priemernú dĺžku života obyvateľov a ekologickú stopu krajiny. V práci je zahrnutý rozbor myšlienky a kontextu tohto merania, zložiek z ktorých sa skladá a štatistických postupov použitých na jeho tvorbu. Obsahuje aj praktickú analýzu dát pomocou štatistických metód aplikovaných v programe R a alternatívu k jeho tvorbe z primárnych dát pomocou procesu DEA. Index šťastia jednotlivca je meraním zameraným na jednotlivcov a zahŕňa v sebe aktuálnu spokojnosť a potenciálne zdravotné riziko. Uvádzame popis jeho tvorby a idey stojacej za ním. Dáta na následnú štatistickú analýzu sme získali pomocou nami vytvoreného internetového dotazníka. Práca je napísaná ľahko čitateľnou formou, vhodná aj pre čitateľa bez hlbších matematických znalostí.

Kľúčové slová: Meranie šťastia, analýza dát, DEA model, Happy Planet Index (HPI), Index šťastia jednotlivca

Abstract

JANOČKO, Jozef: Happiness and attempts to mathematical expression [Bachelor Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics, Supervisor: Mgr. Zuzana Bučková, Bratislava, 2014, 68p.

The bachelor thesis is concerned with various approaches to measurement of happiness. This kind of measurements are getting crucial for leading countries as they are more complex indicators than GDP. The thesis analyzes the Happy Planet Index (HPI) in detail and also proposes our own index named the Index of Happiness of an Individual. The Happy Planet Index is global measurement involving the well-being of inhabitants, their average life expectancy and the ecological footprint of the country. The thesis comprises the analysis of the idea and context of this measurement; description of components that form the index, and statistical concepts used at its making. The thesis also includes practical analysis of data by the means of statistic methods applied in program R and an alternative to its making from its components known as DEA process. The Index of Happiness of an Individual is a measurement focused on individual person and it includes current satisfaction and potential health risk. We present the description of creating the index and also the idea which is staying behind. The data for statistical analysis were gathered from our own online questionnaire created by us. The thesis is written in an easy readable way, suitable also for a reader without deeper mathematical knowledge.

Keywords: Measurement of happiness, data analysis, DEA model,the Happy Planet Index (HPI), the Index of Happiness of an Individual

Obsah

Úvod **9**

1 História a prehľad merania šťastia **11**

2 Teória tvorby Happy Planet Indexu **15**

2.1 Úvod k HPI 15

2.1.1 Motivácia vzniku a charakteristika indexu 15

2.1.2 Celkový kontext merania progresu 16

2.2 Zložky indexu 18

2.3 Zdroje a výpočet 20

2.3.1 Výpočet 21

3 Analýza dát pre HPI **24**

3.1 Pojmológia 24

3.2 Základné ukazovatele 25

3.2.1 Zdôvodnenie rozdielu disperzií 26

3.2.2 Rozdiely v priemeroch 28

3.2.3 Korelačná analýza 30

3.3 Regionálne pozorovania 32

3.3.1 Celkové indexy 33

3.3.2 Happy Life Years 36

4 Dea-analýza údajov k HPI **40**

4.1 Načtnurie myšlienky obálkovej analýzy dát 40

4.2 Aplikácia na HPI dáta 41

5 Výskum vlastného indexu **44**

5.1 Teória tvorby 44

5.1.1 Vážené Šťastie 45

5.1.2 Upravené BMI 46

5.1.3 Informácie k dotazníku 47

5.2 DEA analýza všetkých dát 48

5.3 Analýzy VŠ a IŠJ selekcie dát 49

Záver **55**

Zoznam použitej literatúry **58**

Príloha A - tabuľky medziročných zmien HPI a HLY po subregiónoch **61**

Príloha B - matlabovský skript na proces DEA **63**

Príloha C - kód R k vlastnému výskumu **65**

Úvod

V súčasnosti vzniká čoraz väčšia potreba akýmsi spôsobom kvantifikovať šťastie. Na medzinárodnej úrovni je snaha určiť hodnotu národného šťastia v zmysle blahobytu obyvateľstva. Údaje získané interpetáciou týchto hodnôt sa totiž začínajú považovať za kľúčové pri určovaní smerovania štátov, ako sa uvádza v Svetovej Správe o Šťastí [17]. Úlohou štátu je totiž zabezpečiť spokojnosť jeho obyvateľov a tá omnoho viac súvisí s akýmsi vyjadrením ich šťastia, než vyjadrením ekonomickej veličiny akou je HDP. Spomínaný zdroj [17] uvádza, že v Amerike sa od roku 1960 faktor priemerného HDP strojnásobil, zatiaľ čo úroveň meraného šťastia obyvateľstva ostala nezmenená. Okrem toho, ak je takáto miera vhodne skonštruovaná, dokáže uspokojivým spôsobom opísať nie len hladinu šťastia, ale dokonca v sebe reflektovať viacero iných faktorov - dĺžku života, mieru dopadu na životné prostredie, mieru naplnenosti životných potrieb a mnohé iné. V porovnaní s hrubým domácim produktom, HDP, ktorý zachytáva len mieru produktivity obyvateľstva, je táto miera blahobytu vhodnejším opisom situácie v krajine, a preto môže lepšie pomôcť pri dosahovaní cieľov spoločnosti.

Problém vzniká pri určení takejto miery nakoľko je šťastie relatívny a subjektívny pojem. Relativita tohto pojmu sa pri nadnárodných meraniach zmenšuje spresnením na pojem blahobytu, pri meraniach zameraných na porovnávanie jednotlivcov ho v rámci práce spresníme na pojem spokojnosti. Subjektivita sa neodstraňuje pomocou ďalšieho vymedzenia pojmu, ale pomocou štatistických nástrojov. V štatistike všeobecne platí, že pri zvyšujúcom sa počte dát sa výsledky stávajú vierohodnejšími. Údaje s menej presnou informáciou sú pri správne nastavených meraniach v menšine a ich vplyv na výsledky sa neustále zmenšuje. Okrem toho sa za účelom zvýšenia miery objektivity do týchto meraní pridávajú iné, objektívne merané veličiny. Môžu byť na danú hodnotu šťastia aplikované akoukoľvek matematickou operáciou, závislou od toho, čo všetko má v sebe výsledná informácia zahŕňať. Takto získané údaje sa vo všeobecnosti nazývajú tzv. indexy šťastia. Napríklad Index Šťastnej Planéty (ďalej v práci používame takmer výlučne anglický originál názvu - Happy Planet Index, respektíve jeho skratku - HPI), v sebe podľa [11] zahŕňa okrem určitého vyjadrenia priemerného šťastia obyvateľov aj priemernú dĺžku života v danej krajine a veľkosť jej ekologickej stopy. Zdroj [7] tvrdí, že sa v súčasnosti považuje za jeden z najdôveryhodnejších.

Bakalárska práca sa zaoberá spôsobmi, akými sa dá merať a spracovať úroveň šťastia jednotlivca i celej spoločnosti. V práci uvádzame konkrétne rovnice a indexy zo zdrojov [7,8,14,21,23], z ktorých niektoré podrobne rozoberáme - spomínaný HPI a náš vlastný index - Index šťastia jednotlivca (skratka IŠJ). Jednak teoretickými opismi, napríklad popisom výpočtu, významu a obmedzení daných indexov, ale aj praktickými výskumnými metódami, napríklad štatistickou analýzou údajov, ako aj porovnávaním indexov s údajmi získanými alternatívou k ich tvorbe v podobe tzv. Obálkovej Analýzy Dát (v práci používame najmä anglickú skratku DEA[1]).
Štatistická analýza týkajúca sa HPI v sebe zahŕňa výsledky štatistických metód použitých v programe R na určenie významnosti nárastu alebo poklesu priemernej hodnoty skúmanej veličiny a významosť zmeny disperzie danej veličiny, v rámci porovnávania dvoch rôznych sád meraní. Pre ne skúma aj korelačné koeficienty a overuje dôveryhodnosť takto získaných koeficientov. Podobné analýzy robí aj pre údaje krajín rozdelené na subregióny. Zvažuje dôveryhodnosť získaných výsledkov, zvažuje príčiny získaných signifikantných rozdielov a vyvodzuje isté dôsledky, ktoré z nich plynú.
Práca popisuje aj tvorbu nášho vlastného indexu - IŠJ, a analýzu k nemu získaných dát. Rozoberá ho aj v súvislosti s prvým skúmaným indexom.

Cieľom práce je urobiť krátky prehľad problematikou výpočtu indexu šťastia, a následne sa zamerať na konkrétne indexy, analyzovať nimi získané údaje a ich interpretáciou poukázať na dôležitosť takýchto meraní. Je významná aj z toho dôvodu, že väčšina prác na túto tému je dostupných iba v anglickom jazyku (zdroje [3,9,10,11,12,14,17,21,23]) a zahŕňa v sebe množstvo iných údajov o šťastí ako takom [3,12], takže sa čitateľ k informáciám o jeho meraní len ťažko dostáva. Okrem toho je prínosom práce aj jej štýl, vďaka ktorému je zrozumeteľná aj pre čitateľa s minimálnymi odbornými vedomosťami, nakoľko mnohé z existujúcich prác sú písané v reči matematicko-štatistickej terminológie [17]. Dôležitosť práce demonštruje aj tvorba vlastného indexu a následnej analýzy vlastných dát.

[1]DEA - Data Envelopment Analysis

1 História a prehľad merania šťastia

Skôr, ako sa v práci zameriame na konkrétne indexy šťastia, považujeme za vhodné čitateľa aspoň v krátkosti oboznámiť s niektorými dôležitými momentami histórie merania šťastia, ako aj súčasným rozsahom riešenej problematiky, aby tak mohol plne pochopiť motiváciu práce.

Najprv by sme radi odlíšili tzv. Rovnice šťastia od indexov šťastia, pričom indexy samotné chápeme ako istú špecifickú časť týchto rovníc. Rovnica šťastia dáva do istého vzťahu rôzne faktory, ktoré podľa tvorcu takej rovnice nejakým spôsobom vplývajú na ľudské šťastie. Má skôr informatívny charakter a jej cieľom nie je ani tak niečo konkrétne kvantifikovať, ale skôr samotnou formuláciou človeka prinútiť sa zamyslieť nad vlastným šťastím a životným smerovaním. Mojou obľúbenou rovnicou šťastia je rovnica z [21]:

$$\check{S}\check{t}astie = \frac{(zdroje,\ ktorými\ človek\ disponuje)}{(jeho\ potreby)}$$

Už z takejto jednoduchej formulácie človek môže vidieť, že podľa tvorcov tohto indexu bude človek šťastný, iba ak všetky jeho zdroje - nie len finančné, budú zďaleka prevyšovať jeho potreby. Ľahko z nej teda vyvodiť myšlienku, že iba hromadením zdrojov a zvyšovaním počtu svojich potrieb človek nikdy nedosiahne „dokonalé šťastie“, pretože rovnomerne rastie čitateľ aj menovateľ zlomku. Hladina jeho šťastia teda nestúpne, možno práve naopak. Ak sa však bude schopný uskromniť vo svojich potrebách, dokáže byť šťastní aj pri chabých zdrojoch.

Teraz prejdeme k indexom šťastia, ktoré sú ústrednou témou celej práce. Tým sa rovnica šťastia stane vtedy, keď sa určia konkrétne spôsoby kvantifikácie prvkov v nej vystupujúcich, aby tak mohla istým spôsobom určiť mieru šťastia. Buď absolútne vzhľadom na nejaké vopred stanovené maximum, respektívne minimum, alebo relatívne vzhľadom na ostatné pozorované subjekty. Daný index nevyjadruje iba samotnú mieru šťastia, ale podľa jeho zložiek aj indikuje, čo a ako sa podľa tvorcov indexu podieľa na šťastí. Okrem toho môžu byť do indexu pridávané aj ďalšie činitele, ako tie explicitne súvisiace s aktuálnym šťastím, aby sa tak jeho výpovedná hodnota rozšírila

a vedel v sebe zahrnúť viac informácii – napríklad, či je toto šťastie aj udržateľné.
Tu ešte možno podotknúť, že hoci existuje množstvo rovníc šťastia, väčšina z nich je zameraná práve na šťastie jednotlivca, zatiaľ čo indexy sú skôr zamerané na meranie blahobytu celých skupín obyvateľstva, mnoho ráz dokonca krajín, nakoľko je z tejto strany skôr dopyt po konkrétnych údajoch, ako po všeobecných schémach. Je tu snaha o kvantitatívne výskumy, nakoľko pri dobre nastavených meraniach sa zvyšovaním počtu skúmaných subjektov zvyšuje celková dôveryhodnosť získaných výsledkov.
To, že táto problematika je v súčasnosti vo svete často stredobodom výskumov dosvedčuje aj fakt, že už len zdroj [23] dokumentuje 963 prieskumov sústredených na meranie šťastia a tvorbu rôznych indexov. Uvádza aj konkrétne využitia daných meraní v praxi, teda názvy a niektoré fakty ohľadom vedeckých prác, v ktorých boli použité. Väčšina meraní sa síce skladá iba z jednoduchého zaznamenávania odpovedí z viacerých možností na otázku typu „Ako sa v súčasnosti cítite šťastný?“ , z ktorých každá má svoje bodové ohodnotenie. Aj to však napovedá o snahe akosi kvantifikovať hladinu šťastia jednotlivca a na základe kvantitatívneho výskumu vyvodzovať určité závery.

Prvé kvantitatívne výskumy v oblasti merania šťastia boli motivované skutočnosťou spomínanou už v úvode k práci, a to snahou o presnejšie meranie progresu spoločnosti, než čisto ekonomickými ukazovateľmi, nakoľko sa ukázalo, že nie sú dostatočné. K príkladu z úvodu práce pridáme ešte jeden. Podľa zdroja [12] už v rokoch 1955 a 1971 sa v Deitroite uskutočnil výskum, ktorý sa zaoberal prepojením rodinných príjmov a spokojnosťou rodiny s jej životným štandardom. Zistilo sa, že hoci mediánový príjem vzrástol o 42%, priemerné hodnoty spokojnosti nezaznamenali signifikantnú zmenu. Voľný preklad zo zdroja [3] uvádza teoretické dôvody , prečo nie je dobré používať HDP ako mieru blahobytu obyvateľstva. Spomenieme tri z nich:

- HDP zahŕňa obchodované tovary a služby. To ale vylučuje veľkú časť - ak nie väčšinu - sociálnych aktivít. To znamená, že služby produkované v domácnosti a ostatné medziľudské vzťahy nie sú založené na peniazoch, a tým pádom nie sú merateľné.

- Dokonca aj negatívne udalosti sa sčasti podieľajú na zvyšovaní HDP. Dobrým príkladom sú rôzne nehody a živelné katastrofy, vďaka ktorým sa do HDP môžu

zarátať náklady na záchranné tímy, liečbu zranených a opravu poškodeného majetku.

- Ignoruje sa rozloženie príjmu v spoločnosti a zmeny tohto rozloženia, hoci je známe, že pre subjektívny blahobyt je dôležitý najmä relatívny príjem vzhľadom na okolie a spoločnosť.

Prvý index použitý celonárodnej úrovni bol podľa zdroja [17] vytvorený už v roku 1972. Zrodil sa v Bhutáne pod názvom „hrubé národné šťastie“ a je známy pod anglickou skratkou GNH[2]. Vznikol na príkaz vtedajšieho kráľa. Tento index má istým spôsobom merať spokojnosť obyvateľstva a pre krajinu je jeho zvyšovanie najvyššou možnou prioritou, tak ako pre mnohé iné krajiny HDP. Tento index sa rokmi vyvíjal a v súčasnosti, po roku 2010, je veľmi komplexný - zahŕňa v sebe až 9 oblastí, v nich dohromady 33 indikátorov a v nich spolu až 124 skúmaných premenných. Táto metóda je však vytvorená plne pre potreby a špecifický kontext krajiny Bhutan. Preto, podľa [11] , je tento index používaný iba v tejto krajine, hoci existujú snahy zaviesť ho aj v Brazílii a podobná idea merania sa šíri aj v Ekvádore a Bolívii.

V súčasnosti, tiež podľa [11][3], si potrebu kvantifikovať progres čímsi komplexnejším ako ekonomickým rastom uvedomuje množstvo politikov, akademikov a mimovládnych organizácii. Preto vznikajú nové národné a nadnárodné orgány zaoberajúce sa touto problematikou. Britský úrad pre národnú štatistiku spustil program merania Národného blahobytu. Takisto Taliansky štatisický úrad pracuje na meraní udržateľného a rovnomerného blahobytu. V Nemecku bola založená odborná komisia s cieľom preskúmať prepojenia medzi rastom, prosperitou a kvalitou života. Vedúci predstavitelia Číny tiež uznali, že je potrebné zvoliť iný prístup k rozvoju.

Čo sa týka nadnárodných orgánov, snáď najznámejším je Organizácia pre Ekonomickú spoluprácu a Rozvoj, o ktorej tu máme informácie zo zdroja [15]. Je známa pod anglickou skratkou OECD a existuje už od roku 1960. Na základe analýz štatistických údajov radí vládam členských krajín a to v mnohých oblastiach, najmä ekonomických. Podľa [14] v roku 2011 spustila projekt tvorby Indexu lepšieho života. Robí tak glo-

[2]GNH - Gross National Happiness

[3]Vlastne zvyšok kapitoly zameraný na o rôzne iné národné a nadnárodné úsilie v oblasto v merania šťastia je voľným prekladom z tohto zdroja - [11], ak pri danom fakte nie je explicitne uvedené inak

bálne komplexné meranie pre 36 krajín, zahŕňajúce 11 indikátorov.

Nezávisle od toho podľa z [9] Nová ekonomická nadácia (ďalej len NEF[4]), v roku 2006 spustila prvý ročník tvorby Happy Planet Indexu spomínaného už úvode k práci až pre 178 krajín, teda väčšinu sveta.

Odvtedy sa objavilo mnoho ďalších iniciatív vyzývajúcich na nové merania progresu. V roku 2011 Valné Zhromaždenie OSN jednomyseľne prijalo rezolúciu 65/309, ktorá pozýva členské štáty, aby dohliadali na vypracovanie dodatočných meraní, ktoré lepšie zachytávajú význam realizácie zvýšenia hladiny šťastia a blahobytu v rámci rozvoja spoločnosti so zreteľom na vedenie verejnej politky.

[4]NEF - New Economic Foundation

2 Teória tvorby Happy Planet Indexu

V tejto kapitole podrobne opisujeme v súčasnosti jeden z najcitovanejších indexov šťastia spomínaný ako v úvode k práci, tak aj v predošlej kapitole 1. Jedná sa o Happy Planet Index, čoho skratku v tvare HPI budeme v práci ďalej používať. Uvádzame ako a prečo vznikol, spôsob jeho výpočtu a i to, ako sa vo svete dosiaľ používa.
Čerpáme zo zdroja [11], čo je správa vzťahujúca sa k aktuálnej tretej tvorbe tohto indexu, pokiaľ nie je pri danom fakte explicitne uvedené inak.

2.1 Úvod k HPI

2.1.1 Motivácia vzniku a charakteristika indexu

Motivácia

HPI vznikol na základe potreby lepšie kvantifikovať pokrok v spoločnosti, nakoľko už v úvode k práci ako aj v úvodnej kapitole bolo jasne ukázané, že doterajšie merania zamerané najmä na ekonomickú výkonnosť, ako napríklad HDP, nedokážu dostatočne presne zachytiť neekonomické potreby obyvateľstva, ako ani pokrok v daných oblastiach. Každý pokrok v spoločnosti je zameraný na blaho obyvateľov, preto vyvstáva otázka, či by nebolo vhodné istým spôsobom merať priamo blahobyt.
HPI sa s touto otázkou vysporiadava tak, že kvantifikuje tzv. trvalo udržateľný blahobyt, ktorý definuje ako možnosť súčasných obyvateľov žiť dlhý šťastný život s takou spotrebou prírodných zdrojov, aby rovnakú možnosť mali aj budúce generácie.
Snahou indexu je teda zachytiť akúkoľvek pozitívnu zmenu v spoločnosti, ktorá by sa mala prejaviť na zvýšení blahobytu, alebo aj dĺžky života obyvateľstva. Za pozitívnu zmenu sa považuje aj zníženie spotreby prírodných zdrojov, nakoľko sa index usiluje merať aj udržateľnosť tejto spokojnosti obyvateľstva. Tvorcom indexu je NEF, čo už bolo spomenuté v 1. To je podľa [13] britský výskumný inštitút, ktorého snahou je podporovať sociálnu, ekonomickú and environmentálnu rovnováhu vo svete.

Charakteristika indexu

HPI sa charakterizuje ako meranie efektivity krajín, pričom snahou jeho tvorcov je poukázať na potrebu maximalizácie šťastia, respektívne blahobytu obyvateľstva a jeho kvality života, zároveň s minimalizáciou ekologickej stopy. Kvôli takejto intepretácii

môžeme neskôr v kapitole 4 túto efektivitu počítať alternatívnym spôsobom.
Má iba tri zložky, preto je ľahko interpetovateľný, pričom dokáže podať celkový prehľad toho, ako si jednotlivé krajiny vedú, čo je podľa tvorcov indexu jeho kľúčová vlastnosť.

Index pre danú krajinu sa počíta z dát ohľadom priemerného blahobytu obyvateľstva danej krajiny, priemernej dĺžky života a ekologickej stopy krajiny na obyvateľa. Zjednodušene by sa dal charakterizovať rovnicou:

$$HPI = \frac{(blahobyt)(dĺžka\ života)}{(ekologická\ stopa)} \tag{1}$$

Čo presne jednotlivé zložky znamenajú je uvedené v podkapitole 2.2 a presná rovnica s spolu so zdrojmi dát je uvedená ešte neskôr v podkapitole 2.3. Táto úvodná časť má však poskytnúť všeobecný prehľad ohľadom histórie, účelu a celkového kontextu, v ktorom bol index vytvorený.

2.1.2 Celkový kontext merania progresu

Čo HPI nemeria a ako ho doplniť

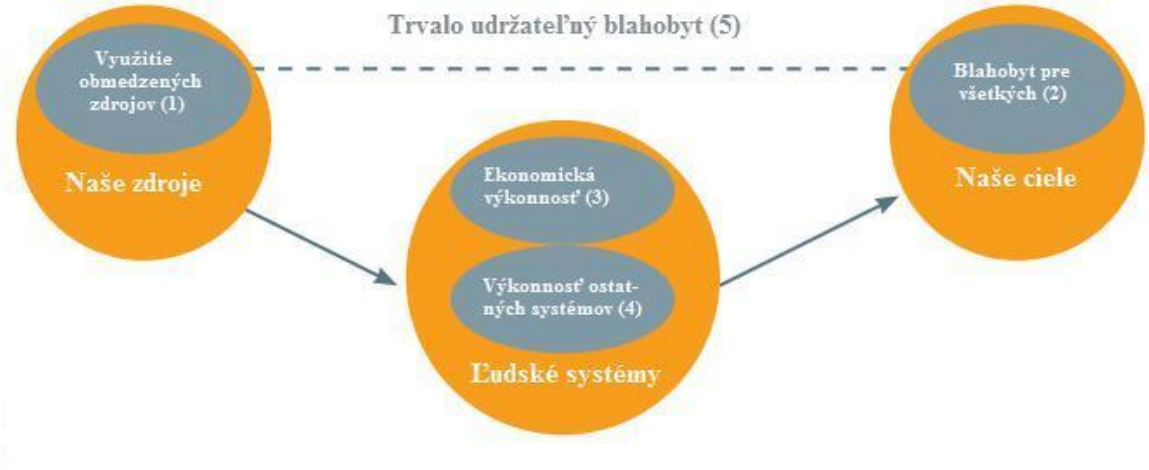
Ani HPI v sebe samozrejme nemôže zachytávať všetko. Krajiny, ktoré majú vysoké HPI skóre môžu mať veľa iných problémov, napríklad zlú situáciu v oblasti ľudských práv. Dotknutá je nimi zväčša menšina, preto sa to nemusí výrazne prejaviť na priemernej hodnote blahobytu, či dĺžky života v danej krajine.
Hoci je v HPI vrámci ekologickej stopy zahrnutých veľmi veľa druhov znečistenia, konkrétne o nej viac v podkapitole 2.2, predsa len nemeria niektoré druhy nepriameho znečisťovania prostrenia, akými sú degradácia ekosystému spojená s odlesňovaním a stratou biodiverzity.

Kvôli nedokonalostiam merania tvorcovia HPI neodporúčajú, aby bol tento index jediný indikátor, na základe ktorého sa bude riadiť verejná politika. Sú nevyhnutné aj iné indikátory, ako aj celkový kontext, vrámci ktorého má byť zasadené aj meranie HPI.

HPI vrámci celkového kontextu pre meranie progresu

Hoci HPI dodáva celkový náhľad smerovania, existuje potreba ďalších indikátorov, ktoré majú ukázať konkrétne detaily na dosiahnutie dobrého skóre v HPI, teda kvalitné životy bez zbytočných nárokov na našu planétu.

Vzhľadom na to spoločnosť NEF, ktorá je tvorcom indexu, buduje novú koalíciu organizácii, aby vyvinula pracovný rámec sústredený na meranie spoločenského progresu. Na nasledujúcom obrázku, zobrazujeme daný pracovný rámec:



Obr. 1: Pracovný rámec na meranie spoločenského progresu, preklad do slovenčiny zo zdroja [11, str.17]

V tomto pracovnom rámci sa rozlišujú tri sféry:

- naše ciele (v zmysle blahobytu pre všetkých)
- naše vzácne (limitované) zdroje
- procesy a systémy, ktoré by mali byť navrhnuté tak aby dosiahli maximálne „výstupy“ blahobytu pri minimálnych „vstupoch“ zdrojov; skrátene ich označíme ľudské systémy

Z poslednej sféry sa zvlášť vyberá ekonomický systém, na ktorý sa doteraz politiky najviac zameriavali a potrebuje najväčšie zmeny, aby umožnil trvalo udržateľný blahobyt. Vlády môžu najviac ovplyvňovať práve tieto ľudské systémy, ale mali by to robiť so zreteľom na konečný cieľ stability a blahobytu.

Na zodpovedné politické rozhodnutia sú nutné hĺbkové merania vo všetkých týchto troch sférach, preto NEF navrhuje na vytvorenie celkového obrazu päť druhov indikátorov číslovaných podľa obrázku [1], zameraných na:

1. meranie enviromentálnej záťaže, normovanej na obyvateľa - týka sa sféry zdrojov
2. meranie percenta populácie, ktoré prospreruje - týka sa cieľovej sféry
3. meranie ekonomickej výkonnosti – ale vzhľadom na trvalo udržateľný blahobyt pre všetkých - týka sa ekonomickej časti ľudskej sféry
4. meranie alebo sada meraní pre iné, neekonomické faktory, ktoré sú ovplyvniteľné prostredníctvom verejnej politiky - týka sa zvyšku ľudskej sféry
5. meranie blahobytu na jednotku enviromentálnej záťaže - jednalo by sa o HPI, alebo meranie podobného typu, prepájajúce zrojovú a cieľovú sféru

Charta šťastnej planéty

Tvorcovia HPI si uvedomujú potrebu nových meraní ľudského progresu a ponúkajú HPI, ako príklad takéhoto merania v praxi. Hoci sa jednotlivé krajiny líšia v tom, čo vrámci jeho zlepšenia musia urobiť, všetkým ponúka jednotný cieľ: vytvorenie ekonomiky, ktorá zabezpečí trvalo udržateľný blahobyt pre všetkých.

Preto podpisujú chartu šťastnej planéty, ktorou:

- Volajú na vlády, aby si osvojili nové merania progresu, ktoré budú mať za cieľ udržateľný blahobyt, a aby boli v centre rozhodovacieho procesu.
- Dávajú na známosť, že chcú vzbudiť vôľu vrámci politiky spoločností, aby sa tieto vhodnejšie merania ľudského progresu plne ujali. Činia tak prácou s partnerskými organizáciami.
- Vyzývajú OSN, aby sa výroba indikátora podobného HPI, ktorý bude merať progres vzhľadom na kľúčový cieľ ohľadom lepšej budúcnosti - trvalo udržateľný blahobyt pre všetkých, stala časťou pracovného rámca po 2015.

2.2 Zložky indexu

Podrobne rozoberáme zložky indexu, akým spôsobom sa získavajú údaje z ktorých sa počítajú, prečo sú zložky v takom vzájomnou vzťahu ako je to uvedené v rovnci základnej charakteristiky spomenutej v rovnici (1).

Okúsený blahobyt

Hodnota blahobytu pre daný národ je dodávaná do HPI ako priemer respondentov z danej krajiny, ktorí sa zúčastnili prieskumu. Prieskum je založený na priamej otázke, ktorá sa volá Rebrík života (ang. Ladder of life). V tomto prieskume je respondent vyzvaný, aby si predstavil rebrík, kde na desiatej, najvyššej priečke, je najlepší možný život a na nultej, najnižšej, ten najhorší. Následne má vyjadriť svoj názor, na ktorej priečke tohto rebríka sa jeho život aktuálne nachádza.

Podľa tvorcov indexu je na tomto prístupe dôležité to, že necháva ľudí posúdiť podľa vlastných kritérii, čo je v ich živote dôležité a takýmto spôsobom zistiť čo najvšeobecnejšiu odpoveď na otázku ohľadom subjektívneho šťastia. Nespolieha sa na expertov, ktorý by rozhodovali namiesto opýtaných, čo je pre nich najlepšie.

Meria sa veličina, ktorá je univerzálna. Každý sa chce cítiť dobre. To platí v priebehu celej histórie ľudstva vo všetkých kultúrach rovnako.

Dĺžka života

Tvorcovia indexu sa do neho rozhodli zahrnúť údaj o priemernej dĺžke života v danej krajine, nakoľko istým spôsobom reflektuje hladinu zdravia obyvateľstva a to sa všeobecne považuje za dôležitú hodnotu. Výhoda ponímať hladinu zdravia obyvateľstva prostredníctvom priemernej dĺžky života spočíva v tom, že údaje o nej sú veľmi ľahko dostupné a s jej meraním majú ľudia už dlhé skúsenosti, nakoľko sa meria už od 19.-teho storočia. Je to očakávaná dĺžka života dieťaťa v danej krajine, ak budú počas celého jeho života zachované odpozorované vzory v úmrtnosti, aké boli počas jeho narodenia.

Happy Live Years

Kombináciou predošlých dvoch zložiek indexu, konkrétne upravením priemernej dĺžky života priemerným blahobytom tak, aby zahŕňal počet šťastne prežitých rokov, vzniká nový indikátor. Nazýva sa "Štastne prežité roky", my sa však na neho budeme odvolávať skratkou anglického názvu HLY.

Zmienka o ňom je dôležitá najmä z toho dôvodu, že pri viacerých štatistických analýzach berieme do úvahy priamo hodnotu tohto indikátora a nie hodnotu blahobytu či dĺžky života osobitne. Konkrétny popis výpočtu je v práci zahrnutý neskôr.

Ekologická stopa

Ekologická stopa je meraná ako plošná miera vyjadrujúca, koľko metrov štvorcových

stredne úrodnej pôdy by potreboval priemerný obyvateľ danej krajiny na udržanie svojich spotrebných návykov a životného štandardu. Konkrétne sa teda vezme plocha vypočítaná pre celú krajinu a táto hodnota sa vydelí počtom jej obyvateľov. Zaujímavé na tomto meraní je, že hoci zďaleka nie je dokonalé, zohľadňuje veľmi veľa dôležitých faktorov. Zachytáva v sebe množstvo pôdy potrebnej na získanie obnoviteľných zdrojov, hlavne jedlo a drevo. Obsahuje v sebe priestor, ktorý zaberá infraštruktúra, aj zalesnenú oblasť potrebnú na zachytávanie CO_2 emisií. Napokon v sebe zahŕňa informáciu aj o "požičanej zemi" a emisiách "z dovozu". To znamená napríklad to, že ak sa kvôli americkým spotrebným návykom vyrobí v Číne mobilný telefón, tak sa to prejaví na ekologickej stope USA a nie Číny. Dokonca sa k nej navyše pripočíta aj ekologická stopa vzniknutá prepravou tohto mobilného telefónu.

2.3 Zdroje dát a výpočet indexu

V tejto podkapitole uvedieme zdroje dát použité pri aktuálnom reporte z roku 2012, ktorý je v poradí už tretí, ako aj poslednú známu metodiku výpočtu indexu. Tak ako pri každom serióznom viacperiódovom výskume, aj tu sa vedeckí pracovníci snažia z obdobia na obdobie svoje metódy kalibrovať, poučení predošlými chybami a nedostatkami, aby stále čím vernejšie zachytávali skutočnosť. Bude sa jednať o pomerne technické záležitosti a od čitateľa sa budú vyžadovať elementárne poznatky z oblasti štatistiky. Tie sme do tejto práce vzhľadom na jej rozsah a celkový charakter nezaradili.

Zdroje dát - dĺžka života

Čo sa týka zdrojov, boli na údaje ohľadom priemernej dĺžky života použité dáta z „2011 UNDP Human Develompent Report", je to vlastne správa z roku 2011 zaoberajúca sa pokrokom ľudstva, vypracovaná organizáciou s doslovne preloženým názvom „zjednotený národný program pre rozvoj".

Zdroje dát - zakúsený blahobyt

Dáta boli zhromažďované a spracované cez Gallup word poll, o ktorom čerpáme informácie z [4]. Je to vlastne jedna z najväčších databáz rôznych globálnych údajov týkajúcich sa až 160 krajín sveta, ktoré sú zisťované na národne reprezentatívnych vzorkách. Konkrétne tieto údaje ohľadom Rebríka života boli získané od vzorky približne

tisíc osôb nad pätnásť rokov z každej viac ako 150 krajín. Hoci sa používali najaktuálnejšie možné dáta, pre rôzne krajiny bol Gallup World Poll sprístupňovaný v rôznom čase. Pre väčšinu krajín sú dáta z roku 2010 alebo 2011. Pre dvadsať krajín sú najnovšie dáta iba z rokov 2008 alebo 2009 a pre sedem krajín sú aktuálne dáta dokonca iba z rokov 2006 alebo 2007. V predošlých dvoch správach ohľadom tohto indexu (zdroje [9] a [10]) bola použitá aj iná metóda na zistenie subjektívneho šťastia, no vzhľadom na fakt, že aktuálne zdroje týchto dát už neboli k dispozícii a toto meranie bolo s aktuálne použitým vysoko korelované, rozhodli sa tvorcovia indexu ďalej pokračovať už len s údajmi získanými pomocou otázky Rebrík života z Gallup word poll.

Zdroje dát - ekologická stopa

Pre 142 zo 151 krajín pre ktoré boli dostupné dáta o blahobyte, bol použitý ako zdroj dát „2008 Ecological Footprint data“, v preklade „2008 údaje o ekologickej stope“, ktoré boli z „2011 Edition of the Global Footprint Networks National Footprint accounts“, v preklade edícia z roku 2011, zahŕňajúca svetovú siet dát ohľadom ekologickej stopy zostavenej z národných vyhlásení ohľadom ekologickej stopy. Pre ostatných 9 krajín sa používali rôzne prediktívne modely, ktoré však nie sú predmetom nášho ďalšieho skúmania a preto ich nebudeme rozoberať do hĺbky. Vyslovujeme však predpoklad, že všetky získané údaje ohľadom ekologickej stopy krajín sú určené dostatočne presne pre potreby nášho skúmania indexu.

2.3.1 Výpočet

V tejto časti práce uvádzame metodiku výpočtu, ktorá bola použitá na výpočet aktuálnych hodnôt indexu. Jedná sa o modifikovanú metódu, ktorá pozostáva z dvoch fáz:

1. výpočet HLY pre každú krajinu zvlášť pomocou účelovo modifikovaného súčinu hodnôt blahobytu a dĺžky života

2. výpočet samotného HPI pre každú krajinu osobitne pomocou účelovo modifikovaného podielu hodnôt HLY a ekologickej stopy.

Pojem „účelová modifikácia“, ktorý je spomínaná vyššie, v sebe zahŕňa dve nevyhnutné štatistické úpravy výrazov, či už HLY, alebo celkového indexu.

- Prvá štatistická úprava sa robí za účelom toho, aby význam žiadneho komponentu vo výraze nedominoval nad ostatnými, či už v HLY alebo celom HPI.

  Ak by bolo HPI počítané iba násobením čísel Dĺžky života a Blahobytu a výsledný produkt následne vydelený ekologickou stopou, posledná spomínaná veličina by dominovala celému indexu. Dôvodom je fakt, že má omnoho väčšiu varianciu ako zvyšné zložky. To sa týka každej zo zvyšných zložiek zvlášť a rovnako aj oboch spolu v HLY.

  Je to nežiaduce, nakoľko sa nepredpokladá, že variancia niektorej zložky indexu je dôležitejšia, než variancie ostatných. Vo výpočte HPI sa preto používa štatistická úprava na ovplyvnenie stupňa variancie jednotlivých komponentov.

  Smerodajnou sa stala variancia dĺžky života a následné úpravy boli robené s hodnotami blahobytu a ekologickej stopy tak, aby sa varianca všetkých zložiek indexu rovnala.

- Druhá štatistická úprava je nevyhnutá z dôvodu škálovania, nakoľko HPI skoré má byť v rozmedzí medzi 0 a 100. Určia sa rozumné hranice, za akých podmienok je už HPI nulové a za akých je rovné maximálnej hodnote, teda číslu sto.

Prvá fáza - vytvorenie HLY

Najprv sa k hodnote blahobytu pridáva konštanta α, pričom je konštruovaná tak, aby následne zvyknutá veličina mala vzhľadom na celú databázu varianciu zhodnú s tou pre údaje týkajúce sa dĺžky života.

Následne je tento upravený blahobyt násobený hodnotou dĺžky života a vzniknutý produkt je vydelený číslom $(10 + \alpha)$. Jedná sa vlastne o určité normovanie, pretože chceme, aby hodnota indikátora HLY bola kdesi medzi nulou a priemernou dĺžkou života v danej krajine. HLY sa teda počíta podľa nasledovnej rovnice:

$$HLY = \frac{(blahobyt + \alpha)(dĺžka\ života)}{10 + \alpha} \tag{2}$$

kde $\alpha = 2,93$.

Druhá fáza

Od údajov HLY sa odpočíta konštanta γ, ktorá zaručí, že krajina s priemerným blahobytom nula a dĺžkou života pod 25 rokov dosiahne napokon celkové HPI skóre

rovné nule. Túto hodnotu pre potreby ďalšieho výpočtu nazveme „upravené HLY“.

Určí sa konštanta β, ktorá sa pridá k hodnote k ekologickej stopy za účelom toho, aby sa variancia týchto údajov rovnala variancii HLY. Túto hodnotu pre potreby ďalšieho výpočtu nazveme „upravená ekologická stopa“.

Napokon sa vypočíta samotný index podielom hodnoty upravených HLY a upravenej ekologickej stopy. Výsledok sa ešte násobí konštantou δ. Táto konštanta je zvolená tak, že krajina s priemerným blahobytom 10, dĺžkou života 85 rokov a ekologickou stopou 1,78 g ha na hlavu, čo predstavuje optimálne dáta, dosiahne maximálne HPI skóre, teda hodnotu sto.
HPI sa teda cez HLY a ekologickú stopu počíta podľa nasledovnej rovnice:

$$HPI = \delta \frac{(HLY - \gamma)}{(ekologická\ stopa + \beta)} \tag{3}$$

kde $\beta = 5,67$; $\gamma = 4,38$; $\delta = 7,77$.

Takisto môžeme vyjadriť výpočet HPI prostredníctvom pôvodných veličín a použitím štyroch konštánt, ktorých úloha nemusí byť nutne zhodná s tou v predošlých rovniciach:

$$HPI = \omega \frac{(blahobyt + \alpha)(dĺžka\ života) - \pi}{(ekologická\ stopa + \beta)} \tag{4}$$

kde $\alpha = 2,93$; $\beta = 4,38$; $\pi = 73,35$; $\omega = 0,60$.

3 Analýza dát pre HPI - medziročná aj subregionálna

V tejto kapitole prechádzame od teórie k praxi. Pomocou štatistických analýz v nej porovnávame údaje z aktuálnej tretej tvorby HPI, indikátora HLY, a merania zložiek HPI, ktoré sme podrobnejšie opisovali v predošlej kapitole 2, s údajmi použitými pri predchádzajúcej druhej tvorbe indexu. Teda prechádzame od okolností a popisu výskumu, k výskumu samotnému.

Údaje, ktoré používame z aktuálnej tvorby indexu, sme čerpali zo zdroja [2]. V práci ich označovujeme ako aktuálne alebo nové dáta. Údaje, ktoré používame z predošlej tvorby indexu, sme získali zo zdroja [1]. V práci vystupujú ako staré alebo staršie dáta. Kvôli korektnosti porovnávaní medzi starým a novým súborom dát sme z oboch vylúčili údaje tých krajín, ktoré sú v zaznamenané iba v jednom súbore, teda vybrali len údaje pre tie krajiny, ktoré sa nachádzajú v oboch. Aj po tejto úprave však máme údaje zo 141 krajin, čo považujeme za dostatočne veľkú databázu, nakoľko podľa zdroja [16] bol v roku 2013 celkový počet suverénnych štátov sveta 195, teda do nášho výskumu zahŕňame približne 72% všetkých krajín sveta.

Odborné štatistické poznatky explicitne spomenuté alebo implicitne zahrnuté v celom texte tejto kapitoly čerpáme zo zdrojov [18] a [19]. Poznatky z ich aplikácie v štatistickom softvéri R zasa zo zdroja [20]. Pre lepšiu vizualizáciu s ním pracujeme v prostredí programu R studio.

Vďaka tomu, že máme k dispozícii výskumné správy týkajúce sa oboch týchto meraní (zdroje [10] a [11]) môžeme si mnohé naše zistenia zdôvodniť konkrétnymi úmyslami tvorcov indexu.

3.1 Pojmológia

V nasledujúcich odsekoch definujeme nové významy bežným slovným spojeniam tak, aby v nich matematicky vzdelaný čitateľ dokázal rozoznať konkrétne odborné štatistické pojmy. Odbornému čitateľovi sú venované aj poznámky pod čiarou uvádzané v tejto kapitole.

Definujeme ***Požadované alebo vhodné vlastnosti dát ohľadom normality*** tak, že nebola zamietnutá hypotéza o pôvode týchto dát z normálneho rozdelenia a dajú sa

preto použiť parametrické testy na zisťovanie signifikantnosti rozdielu výberových disperzií, výberových priemerov alebo výberových korelačných koeficientov. Pre sady dát, ktoré mali viac ako 40 údajov sme za týmto účelom používali Kolmogorov-Smirnovov test a pre sady dát, ktoré mali menej údajov sme používali Shapiro-Wilkov test.

Silné testy nazývame parametrické testy, , nakoľko ich ochota zamietať nulovú hypotézu a teda pripúšťať štatistickú signifikantnosť nejakého rozdielu je vyššia ako pri neparametrických testoch, ktoré budeme v prípade potreby používať namiesto nich. Konkrétne paramtertické testy budú Studentove t-testy, na ktoré sa v danom kontexne odvolávame ako na ***klasické testovanie***, a Welchove t-testy, ktoré v danom kontextne vystupujú ako ***alternatívne testovanie***. Z neparametrických testov budeme používať Wilcoxonove testy pracujúce so súčtami pozícii dát po ich zoradení vzostupne (angl. ranks) a označovať ich budeme ako slabé alebo ***slabšie testy***.

Disperziou vždy označujeme výberovú disperziu priemeru náhodného výberu z normálneho alebo nezisteného rozdelenia. ***Testom alebo testovaním na zistenie rovnosti disperzií*** nazývame vždy, okrem predtestu pre metódu ANOVA, test známy ako F-test, ktorý využíva F rozdelenie podielu výberových disperzií, pričom predpoklada normalitu testovaných dát.

Významný alebo signifikantný ***rozdiel***, je štatisticky významná zmena potvrdená niektorým z hore definovaných testov na hladine významnosti 5%. Týmito testami sme zamietli konzervatívnu hypotézu žiadnej alebo znamienkovo opačnej zmeny a prijali alternatívnu hypotézu.

20% údajov je zväčša myslených ako údaje z 1/5 medzikvantilového rozpätia, alebo ich miernou úpravou, napríklad pridaním istého nízkeho počtu dát aby „slabé“ testy boli ochotné zamietať konzervatívnu hypotézu.

Korelačným koeficientom sa má na mysli podľa kontextu buď Pearsonov alebo Spearmanov výberový korelačný koeficient.

3.2 Základné ukazovatele oboch sád dát a ich analýza

Na začiatok je dobré zistiť základné štatistické ukazovatele oboch sád meraní a následne analyzovať dané výsledky, prípadne vysloviť nejakú hypotézu a tú overiť.

V nasledovnej tabuľke budeme pre úsporu miesta jednorázovo používať nasledovné

skratky:

- „ekologická stopa“ má skratku „ek. st.“
- „dĺžka života“ má skratku „dl. živ.“
- „okúsený blahobyt“ má skratku „ok. bl.“

a uvedieme si v nej dva najzákladnejšie štatistické ukazovatele pre tieto dáta, teda priemer a veličinu vyjadrujúcu rozptyl údajov, v matematickom svete známu ako varianciu alebo disperziu.

Tabuľka 1: Základné štatistiky

	Priemer HPI	Disperzia HPI	Priemer HLY	Disperzia HLY
Aktuálne dáta	42,70851	83,57978	45,93191	136,0296
Staré dáta	43,21929	155,1141	41,43333	213,0554
	Priemer ek. st.	Disperzia ek. st.	Priemer dl. živ.	Disperzia dl. živ.
Aktuálne dáta	3,030213	4,234345	70,14255	95,19618
Staré dáta	2,902837	4,458706	67,91064	123,2365
	Priemer ok. bl.	Disperzia ok. bl.		
Aktuálne dáta	5,422695	1,399053		
Staré dáta	5,91773	1,912469		

3.2.1 Zdôvodnenie rozdielu disperzií

Pozorný čitateľ si už pri nahliadnutí do tabuľky môže všimnúť, že s výnimkou údajov ohľadom ekologickej stopy sa disperzia nových a starých údajov výrazne líši.
Táto hypotéza bola overovaná štatistickou metódou, konkrétne testom na zistenie rovnosti disperzií dvoch sád údajov. Signifikantný rozdiel medzi starými a novými údajmi bol zaznamenaný iba u HPI a HLY, pričom tieto dáta dobre spĺňali predpoklady tohto testu, ktoré sa týkajú požadovaných vlastností stĺpcov dát ohľadom ich pôvodu z normálneho rozdelenia. Pri dátach okúseného blahobytu nebol zistený signifikatný rozdiel, čo však mohlo byť spôsobené aj faktom, že údaje takmer nespĺňali tieto požadované

vlastnosti ohľadom normality[5]. Preto tento výsledok nepovažujeme za dôveryhodný. Dáta ohľadom ekologickej stopy a dĺžky života neboli vôbec vhodné na tento druh testovania, takže tu naše podozrenie nebolo možné overiť. V nasledujúcom texte sa však pokúsime túto hypotézu podporiť pomocou empirických faktov týkajúcich sa jednotlivých údajov, ako aj štatistických postupov, ktorými boli získané.

Dĺžka života a okúsený blahobyt

V prípade dĺžky života a okúseného blahobytu budeme tvrdiť, že tento pokles disperzie dát nastal prirodzeným spôsobom. Rozoberieme najprv hypotézu ohľadom dĺžky života. Pri zlepšovaní životných podmienok sa zvýšenie dĺžky života omnoho výraznejšie prejaví pri krajinách, u ktorých boli tieto podmienky katastrofálne, než v krajinách, kde bol i predtým dosiahnutý vyšší životný štandard a podmienky pre život nastavené tak, aby jeho dĺžka dosahovala aktuálne možné maximum. Preto sa jeho dĺžka už výrazne nezvyšovala. Týmto spôsobom sa prirodzene zmenšujú rozdiely v priemernej dĺžke života jednotlivých krajín sveta a klesá teda rozmanitosť údajov.

Hypotézu, ktorú sme predostreli v predošlom odseku, podložíme nasledujúcim faktami: Najnižšia priemerná dĺžka života bola v starých údajoch 40,5 pre krajinu Zambia a v súčasných až 47,8 pre krajinu Sierra Leone. Zatiaľ čo najvyššia priemerná hodnota v starých údajoch bola 82,3 a v súčasnosti stúpla iba na 83,4, pričom obe tieto hodnoty prislúchajú Japonsku. Medziročná zmena v minimálnych údajoch, ktoré obe prislúchajú chudobným a zaostalým krajinám, je teda zhruba 6,6-krát vyššia ako zmena v maximálnych hodnotách prislúchajúcim rozvinutej krajine Japonsku.

Našu hypotézu podporíme aj štatisticky korektným tvrdením, že disperzia približne 20-tich percent krajín z najnižšou priemernou dĺžkou života je signifikatne vyššia ako disperzia približne 20-tich percent krajín s tou najvyššou. Na rozdiel od celkových udajov pre dĺžku života tento výber dát už mal požadované vlastnosti ohľadom normality potrebné na testovanie rovnosti disperzií. Tento fakt bol zistený rovnako pre aktuálne, ako aj pre staré údaje.

Teraz rozoberieme hypotézu ohľadom okúseného blahobytu. Predpokladáme, že na zmenu disperzie pôsobili hneď dva prirodzené faktory. Prvým je zlepšovanie podmienok

[5]p-hodnota príslušného Kolmogorov-Smirnovovho testu na overenie normality dát bola len tesne nad 5%.

v rozvojových krajinách, ktoré už bolo spomínané vyššie. To sa prejaví na výraznom raste hodnoty blahobytu. Na druhej strane krajiny s vysokým priemerným okúseným blahobytom sú zväčša rozvinuté krajiny a ľudia sú informovanejší o zlej situácii v krajine i v zahraničí. Rozvíja sa konzumný spôsob života a karierizmus, menej pozornosti sa venuje budovaniu psychickej, duševnej a emocionálnej rovnováhy, čo v konečnom dôsledku vedie k zníženiu hodnoty subjektívneho blahobytu jedinca a následne zníženiu hodnoty pre celú krajinu. Túto hypotézu však nevieme overiť faktami, pretože vzhľadom zmenu spôsobu merania a aj meracej škály tejto hodnoty medzi starými a novými údajmi by bolo zdôvodnenie konkrétnymi číslami irlevatné. Tvorcovia indexu tieto zmeny odôvodnili nedostatkom predošlého typu údajov, ako aj vysokou koreláciou oboch týchto typov meraní navzájom. Aj preto si možno tieto zmeny v disperzii dát ohľadom blahobytu vysvetliť aj pričinením tohto umelého faktora.

Ostatné zmeny

Aj v prípade ostatných zmien v disperziách, teda v údajoch pre HLY a HPI, ktoré sú ešte výraznejšie, nakoľko sa ukázali signifikantné pomocou štatistických metód, sa podpísali umelé kroky tvorcov indexu. Tentoraz však aj tie zahrnuté do tvorby týchto veličín, ktoré vznikajú z pôvodných údajov. Najvýraznejšia zmena v nových údajoch je oproti predošlým je v tvorbe HLY, nakoľko sa v nich predtým vôbec neuvažovalo pridanie konštanty α ako je to zahrnuté v jeho aktuálnej tvorbe v rovnici (2). Jeden z dôvodov pridania tejto konštanty je aj vyrovnať vplyv disperzií blahobytu a dĺžky života pri tvorbe HLY, pričom sa za smerodajnú považuje práve disperzia údajov pre dĺžku života. Druhým je akési normovanie tohto indikátora. Tieto inovácie vo výpočte teda výrazne znižujú disperziu HLY a má to následne vplyv aj na významný pokles disperzie celkového indexu.

3.2.2 Rozdiely v priemeroch

Začneme zisťovaním významnosti a následným zdôvodňovaním týchto rozdielov pre jednotlivé zložky indexu a neskôr pristúpime týmto procedúram pre indikátor HLY a celkový index.

Významnosť zmien v zložkách indexu

Vzhľadom na zmenu spôsobu merania okúseného blahobytu a jeho škálovania je

vcelku očakávané, že sa preukáže významná zmena v priemeroch tejto veličiny. Konkrétne, keďže priemer z aktuálnych meraní je nižší ako ten starý, budeme sledovať signifikantnosť jeho poklesu. Ten sa ukázal nie len pre celkový priemer a pre vzorku približne 20% navyšších údajov, kde by mohol nastať aj celkom prirodzeným spôsobom popísaným v predošlej podkapitole, ale aj pre vzorku približne 20% najnižších údajov, kde sme predtým racionálne predpokladali, že by sa malo jednať skôr o signifikatný nárast tohto priemeru vzhľadom na postupné zlepšovať životných podmienok v zaostalých častiach sveta, kde sú hodnoty tohto indikátora najnižšie.

Čo sa týka priemernej dĺžky života, vychádza významne vyššia a to aj keď berieme do úvahy fakt, že pre nevhodné vlastnosti týchto údajov ohľadom normality bolo treba použiť slabé testy, ktorých ochota pripustiť významnosť rozdielu je veľmi nízka. Avšak tento významný nárast bol zistený aj pre horných, resp. dolných približne 20% údajov, kde už dáta mali požadované vlastnosti a mohli sme teda použiť silnejšie testy.

V dátach pre ekologickú stopu nám podľa očakávaní z relatívnej blízkosti starého a nového priemeru nebol potvrdený významný rozdiel, hoci sme kvôli nevhodným vlastnostiam ohľadom normality dát znova museli použiť slabé testy, teda metódy málo ochotné tento rozdiel pripúšťať, a existuje možnosť, že ak by sme mohli použiť silné testy, aj takýto malý rozdiel by sa ukázal ako signifikantný.

Významnosť zmien v HLY a HPI

Pre indikátor HLY bol potvrdený signifikantný nárast priemerov, pričom sme pri testovaní zohľadňovali fakt, že disperzie týchto meraní sa významne líšili[6].

Tento rozdiel však môže byť spôsobený zmenou tvorby HLY oproti predošlej tvorbe indexu, kde sa ešte neuvažovalo o konštante α, ktorá slúži na vyrovnanie hladiny disperzie medzi dátami z okúseného blahobytu a priemernej dĺžky života.

Pre celkový index nastal pokles priemerov, ten však nebol potvrdený ako signifikantný. Takisto mohol byť spôsobený zmenou tvorby tohto indexu, nakoľko sa predtým neuvažovalo o odrátali konštanty γ, ktorá slúži na určenie akéhosi minimálneho HLY na to, aby mal človek nenulové HPI skóre.

[6]Použili sme Welchow t-test na hladine významnosti 5%.

3.2.3 Korelačná analýza

Pri porovnávaní dvoch sád údajov je vždy dobré zistiť, do akej miery spolu súvisia. V našej ďaľšej analýze sa zaoberáme najmä výberovým korelačným koeficientom.
Korelačný koeficient je číslo medzi -1 a 1. Ak je kladný, zjednodušene by sa dalo povedať, že vyjadruje, do akej miery sa v druhej sade dát opakujú relatívne vzťahy dát z prvej. Napríklad na našich dátach sa dá interpretovať ako percento vyjadrené v desatinných číslach vyjadrujúce, nakoľko sa relatívne vyššie hodnoty HPI niektorých krajín v starých údajoch budú opakovať v tých nových, resp. do akej miery budú medzi údajmi podobné vzťahy, ako v prvej sade. Ak je záporný vyjadruje, do akej miery nová sada sleduje presne opačné vzťahy, než aké platili v starej. Teda v našom prípade, nakoľko platí, že krajiny, ktoré v prvej sade dosiahli relatívne vyššie skóre v novej sade dosiahli naopak relatívne nižšie hodnoty. Čím je bližší nule, tým sú tieto dve sady údajov menej závislé.
Slovo „výberový“, ktoré budeme neskôr vynechávať, znamená, že sa nejedná o presnú hodnotu tohto koeficientu, ale o odhad pomocou našich údajov. Ak by sme mali viac údajov, napríklad hodnoty aj pre jednotlivé regióny daných krajín, bolo by možné daný odhad upresniť. Úplne presnú hodnotu by sme poznali, iba ak by sme ich mali nekonečne veľa, alebo poznali niektoré konkrétne vlastnosti našich sĺpcov dát[7]. Preto je na mieste testovať hypotézu, či by táto skutočná hodnota mohla byť aj nulová. Teda, hoci to vyzerá, že medzi dátami existuje závislosť, či by sa pri zvyšovaní počtu dát mohlo ukázať, že tieto údaje sú nezávislé.

Pri porovnávaní indikátorov HLY nám tento koeficient vyšiel pomerne vysoký - 0,9558. Overovali sme hypotézu o jeho kladnosti odôvodnenú v predošlom odseku a pomocou štatisticky korektnej metódy[8] sme boli schopní takmer s istotou overiť, že bude naozaj kladný.
Tento výsledok je zaujímavý aj z toho dôvodu, že priemer indikátora HLY medziročne výrazne stúpol. Teda skoro všetky krajiny si oproti predošlým meraniam o čosi polepšili, pričom skoro presne sledovali relatívne vzťahy z minulých meraní. Teda krajiny

[7] Napríklad parametre presného štatistického rozdelenia, z ktorého pochádzajú.

[8] Bol použitý test pre overenie kladnosti Pearsonovho výberového korelačného koeficientu na hladine významnosti 5%.

príliš nezmenili svoju pozíciu v rebríčku zostavenom podľa hodnoty HLY.
Avšak, ako sme už spomínali vyššie pri analýze zmeny priemerov v podkapitole 3.2.2, tento významný nárast mohol byť spôsobený zmenou tvorby tejto veličiny a na vyvodenie nejakého záveru by bolo treba skúmať významnosť jej vplyvu na celkový rozdiel týchto priemerov.

Pri porovnávaní údajov ekologickej stopy sme dosiahli tento koeficient podobne vysoký a to 0,9168. Aj pre tieto údaje sme schopní takmer s istotou tvrdiť, že skutočná hodnota tohto koeficientu bude kladná, hoci sme pre nevhodnosť údajov vzhľadom na normalitu museli použiť zložitejšiu štatistickú metódu na toto testovanie[9]. Nakoľko sa nám však pri predošlom skúmaní nepreukázal rozdiel priemerov týchto veličín ako významný, nemôžeme tvrdiť, že si všetky krajiny polepšili alebo pohoršili, skôr to, že dosiahli podobné hodnoty v pre obe merania.

Pri porovnávaní HPI nám tento koeficient vyšiel iba 0,8213 a rovnakým testovaním ako pre HLY sme takmer s istotou overili jeho kladnosť, teda pozitívny súvis údajov predošlého a aktuálneho výpočtu indexu.
Toto číslo je o čosi nižšie aké sme zistili pre HLY a ekologickú stopu, pričom sme boli schopný dokázať významosť tohto rozdielu [10] pre korelačné koeficienty rátané pre HLY a pre HPI.
Môže to byť spôsobené napríklad faktom, že korelačný koeficient sád dát prislúchajúcim HLY a ekologickej stope ako v aktuálnom, tak aj predošlom meraní vyšiel omnoho nižší ako ten pre medziročné zmeny týchto veličín - pre staré dáta 0,6942 a nové 0,7128. Teda, keďže celkové indexy sú tvorené z týchto dvoch veličín, dôsledkom bude to, že ich medziročná zmena bude celkovo menej predpovedateľná, ako zmena v jednotlivých veličinách, čo vlastne znamená, že údaje budú spolu o čosi menej súvisieť.
Stále je však tento súvis pomerne vysoký. No podobne ako pre údaje ohľadom ekologickej stopy, ani tu nám pri predošlom skúmaní nebol zistený výzmnaný rozdiel priemerov.

[9] Test na overenie kladnosti Spearmanovho výberového korelačného koeficientu na hladine významnosti 5%.

[10] Testom na zistenie významnosti rozdielu Pearsonových korelačných koeficientov na hladine významnosti 5%, ktorý je založený na tzv. Fisherovej Z-transformácii týchto koeficientov, konkrétne operácia arcustangens hyperbolický, po ktorej má približne normálne rozdelenie a potom porovnáva dve normálne rozdelné veličiny na mieru upraveným t-testom.

3.3 Regionálne pozorovania

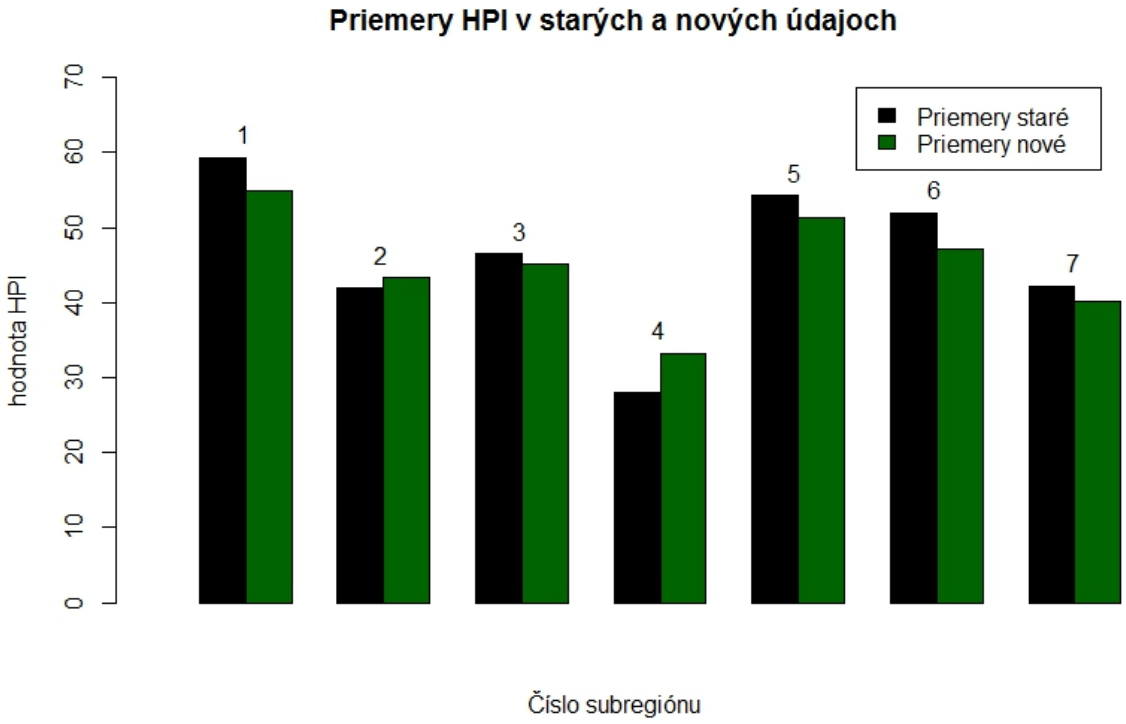
V údajoch, ktoré sme skúmali, bolo k dispozícii nasledovné delenie krajín po určitých oblastiach, tzv. subregiónoch, celkovo ich je sedem a sú to:

1. Latinská Amerika
2. Západný svet - Vyspelé krajiny z Európy, Severnej Ameriky; Austrália a Nový Zéland
3. Stredný východ a Severná Afrika
4. Subsaharská Afrika
5. Južná Ázia
6. Východná Ázia
7. Ostatné štáty - najmä nerozvinuté krajiny v rozvinutých častiach sveta - napríklad Bulharsko; do tejto kategórie tvorcovia indexu zahrnuli aj Slovensko

Nakoľko pre každý subregión máme údaje zo starej i aktuálnej tvorby indexu a všetkých jeho zložiek, celkovo sa jedná o 70 vzoriek dát. Nie je v našich silách urobiť podrobnú analýzu základných štatistík, ako tomu bolo pre všetky údaje spolu, pretože by si to zrejme vyžadovalo omnoho väčší rozsah práce. Na ilustráciu rozsahu ich uvádzame v prílohe A pre dáta HPI a HLY rozdelené po subregiónoch. V tejto časti sa preto budeme explicitne zaoberať iba niektorými testovaniami a uvedieme len tie údaje, ktoré sú nevyhnutne potrebné.
Takisto nebudeme vždy používať všetky nám dostupné údaje, ale niekedy z daných subregiónov vylúčime položky tak, aby zvyšné údaje mali požadované vlastnosti ohľadom normality ktoré sú predpokladom pre testy, ktoré budeme používať[11].

[11] Aplikovali sme na dáta proces odstraňovania „outlierov“, teda hodnôt menších ako spodná, resp. väčších ako vrchná hradba.

3.3.1 Celkové indexy



Graf 1: Porovnanie priemerov HPI po subregiónoch

Po takejto úprave za účelom analýz celkových indexov vylúčime údaje dvoch krajín z prvého subregiónu zo starých i nových údajov. Bez nich majú tieto dáta požadované vlastnosti ohľadom normality. Celkovo sa nám vzorka zmenší z 23 na 21 údajov.

Pri pohľade na graf sa nám logicky javí otázka, či sú medziročné zmeny v hodnotách priemeru HPI pre jednotlivé subregióny štatisticky významné. Samotný graf totiž môže byť zavádzajúci, nakoľko ukazuje, že pre väčšinu subregiónov priemer HPI medziročne klesol, avšak celkovo tento pokles nie je signifikantný, ako uvádzame v 3.2.2. Naša prvá hypotéza teda znie: Medzi subregiónmi preukazujúcimi výraznú zmenu, bude rovnako zastúpený významný nárast a pokles tejto hodnoty.

Vďaka úprave prvého subregiónu všetky sady dát spĺňajú požadované vlastnosti ohľadom normality a preto na overenie tejto hypotézy použijeme silnejšie testy. Tie nám potvrdili výrazný medziročný pokles priemeru HPI pre prvý subregión a významný nárast pre štvrtý subregión. Pre ostatné sa nám nepotvrdila signifikantnosť tejto zmeny[12].

[12]Hoci pre šiesty subregión vyšla p-hodnota v príslušnom t-teste iba čosi nad 6%.

Celkovo teda môžme skonštatovať, že táto hypotéza bola potvrdená.

Čo sa týka analýzy hodnôt subregiónov medzi sebou v rámci daného ročníka merania, môžeme testovať testovať hneď dve všeobecné hypotézy týkajúce sa všetkých dvojíc súčasne a to:

1. Či je medzi nejakou z nich významný rozdiel v disperziách.[13]

2. Či je medzi nejakou z nich významný rozdiel v priemeroch. [14]

Pričom výsledok prvej hypotézy ovplyvňuje možnosť testovania druhej a to tak, že ak sa pre nejakú dvojicu zistí významný rozdiel disperzií, nebude už možno takto celkovo testovať významnosť rozdielu priemerov, aspoň nie nami používanou metódou.
Pre údaje HPI z predošlých meraní nám testovanie prvej hypotézy zamietlo možnosť, že by disperzie tohto údaju pre všetky subregióny boli pomerne blízko seba, teda aspoň medzi dvoma z nich existuje významný rozdiel. To nám však zabraňuje testovať druhú zo spomínaných hypotéz.
Pre aktuálne merania HPI nám však prvé testovanie túto možnosť pripustilo. Domnievame sa, že tento rozdiel vo výsledku je spôsobený najmä zmenou výpočtu indexu oproti predošlým meraniam, pretože v tom aktuálnom sa koriguje miera disperzie pri oboch krokoch výpočtu. Vďaka tomu však môžme testovať druhú zo spomínaných hypotéz. Testovanie zamietlo možnosť, že by sa priemerné hodnoty HPI pre všetky subregióny od seba príliš nelíšili, teda potvrdil významný rozdiel aspoň pre jednu dvojicu dát.

Tvorcovia indexu sa v zdroji [11] vyjadrili, že najvyššie hodnoty HPI sa po obe roky dosahujú v Latinskej Amerike, v našom meraní je to prvý subregión. Deje sa tak z toho dôvodu, že majú relatívne nízke hodnoty ekologickej stopy a zároveň pomerné vysoké HLY. Skutočne nám priemerné hodnoty HPI vyšli pre obe sady meraní najvyššie pre tento subregión. Chceme však zistiť, od ktorých ostatných subregiónov je táto hodnota signifikantne vyššia, ako pre staré, tak aj pre nové údaje. Vďaka vyššie spomenutej príprave dát tvorenej vylúčením niektorých nevhodných údajov sme mohli na zistenie významnosti rozdielu priemerov zakaždým použiť silný test[15].

[13] Používame Barttletov test na hladine významnosti 5%.

[14] Používame metódu ANOVA na hladine významnosti 5%.

[15] Pre aktuálne dáta sme vďaka vhodnému výsledku Barlettového testu mohli všade použiť t-test na

Pre staré údaje sme dokázali overiť, že priemer HPI pre Latinskú ameriku bol významne vyšší ako vo všetkých ostatných subregiónoch. V piatom subregióne obsahujúcom krajiny z Južnej Ázie však máme iba 5 údajov, navyše pri príprave na toto testovanie sa zistil významný rozdiel medzi disperziami tohto a prvého subregiónu. Preto výsledok porovnávania priemerov subregiónov číslo 1 a 5 nepovažujeme za vierohodný[16].
Pre nové hodnoty sme boli schopný zistiť, že priemer tohto indexu pre tento subregión bol významne vyšší od všetkých, okrem piateho subregiónu, hoci i tu táto možnosť nebola pripustená pomerne tesne[17].
Ak by sme to zhrnuli tak, že vzhľadom na malý počet údajov výsledky týkajúce sa piateho subregiónu nepovažujeme za dôležité, môžeme vyhlásiť, že sme boli schopný štatistickými metódami potvrdiť túto zmienku tvorcov indexu.

Pre celkové indexy po subregiónoch ešte spravíme korelačnú analýzu medziročných údajov. Vo všetkých subregiónov vyšiel korelačný koeficient nižšií ako pri korelácií všetkých dát spolu, čo je pomerne logické, nakoľko sa vo väčšom počte dát skôr preukáže vyššia miera závislosti. Toto tvrdenie sme preukázali aj zistením, že po testovaní, ktoré bolo prvý raz uvádzané v 3.2.3, či dané korelácie môžu byť nulové, nám táto možnosť vyšla iba pri piatom subregióne. Jeho korelácia jednak vyšla najnižšia, no najmä má tento subregión najmenej údajov. Ďalej môžeme porovnávať významnosť rozdielu korelačných koeficientov jednotlivých subregiónov, tak ako sme to urobili pre korelačné koeficienty HLY a HPI v 3.2.3. Testovanie[18] preukázalo významný rozdiel medzi korelačnými koeficientami HPI iba medzi subregiónmi 4 - Subsaharská Afrika a 7 - nerozvinuté krajiny v rozvinutých častiach sveta. Počty dát v subregiónoch sú veľmi podobné v 4 je to 33 a v 7 je to 27, preto toto testovanie považujeme za dôveryhodné. Keďže korelačný koeficient pre Subsaharskú Afriku vyššiel nižší - s hodnotou 0,315, znamená to, že údaje z nových meraní pre tento subregión sledujú relatívne vzťahy z predošlého merania výrazne menej ako údaje subregiónu 4, pre ten vyšiel korelačný

hladine významnosti 5%. Pre staré dáta sme museli po dvojiciach používať F-test a podľa výsledku rozhodovať o tom, či následne použiť t-test alebo Welschov t-test, obe na hladine významnosti 5%.

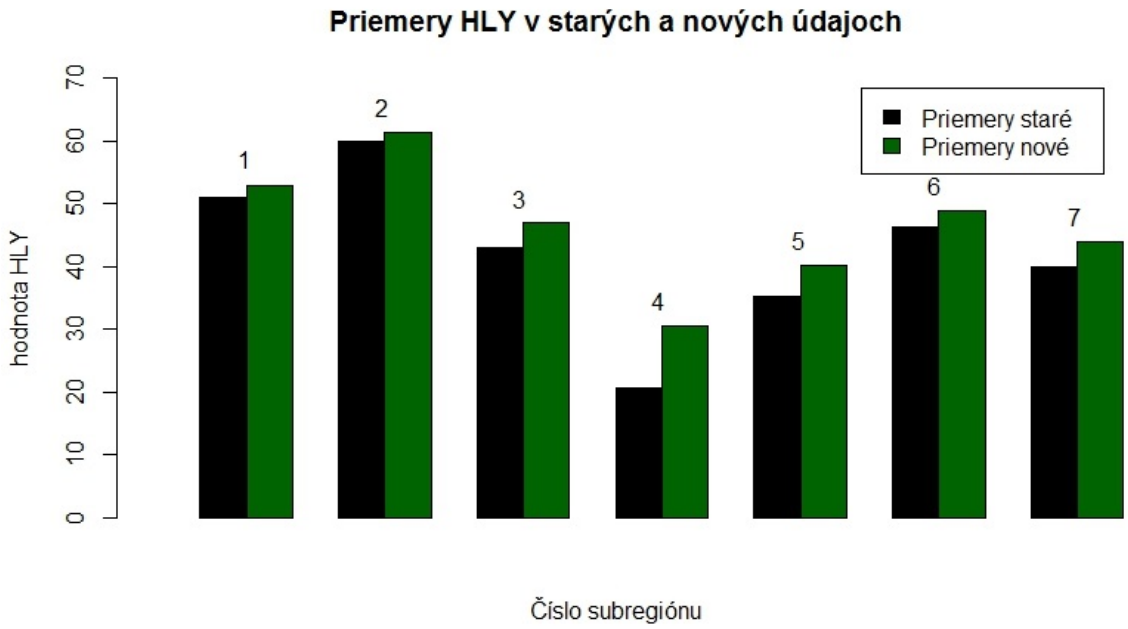
[16] Museli sme totiž použiť Welschov t-test, ktorý používa iba istú aproximáciu studentovho rozdelenia a tá je pre taký malý počet dát v tejto sade veľmi nepresná.

[17] V danom t-teste na hladine významnosti 5% vyšla p-hodnota 6,68 percenta.

[18] popísané v poznámke pod čiarou č. 10 v 3.2.3 .

koeficient 0,7967.

3.3.2 Happy Life Years



Graf 2: Porovnanie priemerov HLY po subregiónoch

Pre prvý subregión sme použili upravenú skupinu dát pochádzajúcu z predošlého pozorovania ohľadom HPI. Zároveň kvôli nevhodným vlastnostiam ohľadom normality novej sady dát HLY siedmeho subregiónu vylúčime z tejto sady údaje dvoch krajín zo starých o nových údajov metódou spomenutou v 3.3[19]. Vzorka sa zmenší z 27 na 25 údajov. Žiaľ, stará sada dát druhého subregiónu nám požadované vlastnosti nespĺňa ani po pokuse o takúto úpravu. Hovoríme o „pokuse“, nakoľko nám daný štatistický postup nevylúčil z datasetu žiadne údaje a preto nezmenil jeho nevhodné vlastnosti. Preto pri testovaní dvoch úvodných hypotéz ako ma začiatku predošlej časti práce 3.3.1 venovanej celkovým indexov v subregiónoch, budeme musieť vylúčiť práve tieto údaje. Treba však zdôrazniť, že pre individuálne testovania túto sadu dát netreba vylúčiť.

Tentoraz nám pohľad na graf skôr potvrdzuje doterazjšie zistenia, nakoľko sa nám v 3.2.2 potvrdil signifikantný nárast celkového priemeru HLY medziročne. Graf ukazuje,

[19] proces odstraňovania outlierov, popísaný v poznámke pod čiarou č. 11

že by takéto čosi mohlo nastať i v niektorých subregiónoch. Tvrdíme niektorých, pretože vzhľadom na menší počet dát v subregiónoch oproti celkovej vzorke sú bežne používané testy menej náchylné pripustiť významnosť rozdielu. Navyše, pre spomínané nevhodné vlastnosti dát druhej sady meraní pre ich zmenu použijeme slabší test, ktorého ochota pripúšťať významný rozdiel je nižšia ako pri silných testoch. Naša prvá hypotéza tejto časti práce teda znie: Aspoň pre niektoré subregióny nám príslušné testy pripustia významný nárast priemernej hodnoty HLY. Táto hypotéza sa nám aj potvrdila, pretože významná zmena bola zaznamenaná v štvrtom, piatom a siedmom subregióne a vždy sa jednalo o nárast.

Testovanie hypotézy o tom, či sú všetky hodnoty disperzií subregiónov, s výnimkou druhého, pomerne blízko seba, nám pre staré, ako aj pre nové údaje túto možnosť vylúčilo. To nám však zabraňuje testovať druhú zo spomínaných všeobecných hypotéz, či sú si ich priemery pomerne blízke.

Pri jednotlivých testovaniach významnosti rozdielu priemerov sme sa zamerali na skúmanie dvoch hypotéz. Sú to:

- Či hodnoty priemeru HLY druhého subregiónu zahŕňajúce vyspelé krajiny západného sveta sú signifikatne vyššie od ostatných regiónov.
- Či hodnoty HLY prvého subregiónu sú signifikatne vyššie od väčšiny subregiónov.

Prvá hypotéza v sebe okrem overenia dominancie HLY vyspelých krajín západného sveta zahŕňa aj štatisticky zaujímavé pozadie, nakoľko kvôli nevhodným vlastnostiam dát druhého subregiónu ohľadom normality pre novú i starú sadu meraní tento rozdiel overujeme slabším testom[20] a keďže tieto testy majú menšiu ochotu pripúšťať signifikantnosť rozdielu než testy, ktoré by sme mohli použiť ak by dáta mali tieto požadované vlastnosti.
Avšak pre súčasné i staršie pozorovania nám toto testovanie potvrdilo dominanciu HLY vyspelých krajín západného sveta nad zvyškom, teda sme boli schopní overiť signifkantnosť rozdielu HLY tohto regiónu od HLY akéhokoľvek iného regiónu.

Druhá hypotéza nadväzuje na predošlé testovanie hypotézy ohľadom celkových indexov, kde chceme overiť tvrdenie tvorcov indexu, že hodnoty HPI sa v Južnej Amerike

[20]pod týmto názvom bol v podkapitole 3.1 definovaný neparametrický test na overenie signifikantnosti rozdielu dvoch výberových priemerov, známy ako Wilcoxonov test.

dosahujú najvyššie aj preto, že jeho HLY je pomerne vysoké, čo intepretujeme v podobe tejto hypotézy. Zo štatistickej stránky zasa pre toto testovanie budeme môcť používať silné testy[21] , čo je obmena oproti testovaniu prvej hypotézy.
Najprv sme testovali možnosť rovnosti disperzie prvého a každého z ostatných subregiónov, pričom nám táto možnosť bola zakaždým pripustená, hoci v niektorých prípadoch len pomerne tesne[22]. V tých prípadoch sme však klasické testovanie signifikantnosti rozdielu priemerov overovali alernatívnym.
Priemer HLY sa nám pre nové i staré údaje južnej Ameriky sa nám ukázal ako signifikatne vyšší od priemeru HLY všetkých ostatných subregiónov sveta okrem druhého, ktorý zahŕňa vyspelé krajiny západného sveta, a šiesteho, ktorý zahŕňa krajiny východnej Ázie. Obyvatelia východnej Ázie sú však známy svojou dlhovekosťou, ktorá v podobe priemernej dĺžky života výrazne prispieva do veľkej hodnoty HLY týchto krajín, takže hoci tento výsledok nebol očakávaný, je pre nás pochopiteľný.
Aj napriek tomu však prehlasujeme, že aj naša druhá hypotéza bola potvrdená, nakoľko Latinská amerika v HLY dominuje nad štyrmi subregiónmi, ktoré zahŕňajú približne 72% zvyšných krajín zahrnutých v tomto testovaní.

Pre HLY po subregiónoch ešte urobíme korelačnú analýzu medziročných údajov, pričom používame skoro rovnaké postupy ako pri korelačnej analýze HPI po subregiónoch. Výsledky tiež vychádzali podobné - vo všetkých regiňoch vyšiel korelačný koeficient nižší ako pre celkové dáta a pre všetky subregióny okrem piateho, ktorý obsahoval namenej dát a jeho korelačný koeficient vyšiel najnižší, sme dokázali štatisticky overiť kladnosť tohto koefcientu[23].
Okrem toho porovnávame významnosť rozdielu korelačných koeficientov jednotlivých subregiónov, no z tejto analýzy musíme opäť vylúčiť nové i staré dáta krajín druhého subregiónu. Testovanie preukázalo významný rozdiel medzi korelačnými koeficientami

[21] pod týmto názvom boli v podkapitole 3.1 definované parametrické testy na overenie signifikantnosti rozdielu dvoch výberových priemerov a to t-test a Weschov t-test.

[22] napríklad P-value F-testu nám v prípade testovania rovnosti disperzie medzi v nových údajoch medzi dátami prvého a tretieho subregiónu vyšla iba 5,33%.

[23] Pre druhú sadu dát bol kvôli nevhodným vlastnostiam ohľadom normality použitý test pre overenie kladnosti Spearmanovho výberového korelačného koeficientu na hladine významnosti 5% namiesto Pearsonovho, ktorý sme v práci používali pre dáta preukazujúce požadované vlastnosti.

HLY iba medzi subregiónmi 4 (Subsaharská Afrika) a 6 (východná Ázia) hoci počty dát nie si nie sú príliš blízke - štvrtý subregión obsahuje 33 údajov a šiesty iba 18, a významnosť tohto rozdielu bola pripustená vo veľmi slabej miere[24]. Je to zaujímavé z toho dôvodu, že sa jedná o regióny, ktoré majú pomerne vysoké HLY, pretože z toho môžme vyvodiť určité závery. Korelačný koeficient novej a starej sady dát Latinskej Ameriky vyšiel 0,5459 a Východnej Ázie 0,8764, teda zatiaľ čosi vo Východnej Ázii polepšili skoro všetky krajiny v rovnakej miere, v Latinskej Amerike si zrejme niektoré krajiny polepšili výraznejšie a iné menej výrazne.

[24]p-hodnota vyšla čosi nad 4,9%, teda stačila by $\frac{1}{10}$%, aby tento rozdiel nebol uznaný za signifikantný.

4 Dea-analýza údajov k HPI

4.1 Načtnurie myšlienky obálkovej analýzy dát

V tejto kapitole budeme za účelom nového výskumného elementu pre HPI dáta, tzv. obálkovej analýzy dát (budeme ju v práci označovať anglickou skratkou DEA[25]), využívať najmä zdroje [5] a [6].

Doteraz sme vypočítaný index a daný výpočtový postup brali ako pevný fakt, preto sme iba analyzovali údaje o HPI prevzaté zo zdrojov [10] a [11]. Teraz sa pokúsime nahradiť tvorenie indexu z jeho zložiek alternatívnou metódou a získané výsledky s ním porovnať. Podnietili nás k tomu samotní tvorcovia indexu, ktorí HPI definovali ako určitú mieru efektivity krajín - nakoľko sú schopné zabezpečiť obyvateľom dlhé a šťastné životy za čo najmenšiu cenu v podobe nárokov na životné prostredie. Vyjadrili to istou úpravou výrazu, ktorý pozostával zo súčinu hodnôt priemernej dĺžky života a priemerného okúseného blahobytu, a podielu získanej hodnoty a ekologickej stopy.

Čo ak by sme však tú istú myšlienku vyjadrili inak? Nakoľko sa totiž jedná o porovnávanie efektivity rôznych útvarov, v tomto prípade krajín, núka sa nám namiesto tvorby HPI aplikovať na dáta proces DEA. Ten sa totiž zaoberá práve efektivitou rôznych útvarov. Určuje ju relatívne, teda niekoľko útvarov vyhodnotí ako maximálne efektívne a ďalšie relatívne neefektívne podľa toho, nakoľko sa líšia od efektívnych. Každý útvar má svoje vstupy, ktoré chce minimalizovať a výstupy, ktoré chce naopak maximalizovať. Efektivita je percentuálna, teda jej hodnody sú v škále medzi 0 a 1, a v rámci tejto teórie sa vo všeobecnosti definuje nasledovne:

$$Efektivita(E) = \frac{(produktivita\ daného\ útvaru)}{(maximálna\ produktivita)} \tag{5}$$

Pričom:

$$Produktivita = \frac{(vážený\ súčet\ výstupov)}{(vážený\ súčet\ vstupov)} \tag{6}$$

DEA povoľuje, aby si objekt tieto kladné váhy nastavoval tak, aby jeho efektivita bola maximálna možná. Nakoľko sa však pri takýchto úplne voľných váhach často

[25]DEA - data envelopment analysis

stáva, že váha niektorej zložky je nulová (dalo by sa to interpretovať tak, že daný vstup alebo výstup nemá žiadny význam v príspevku do efektivity útvaru), je niekedy vhodné určiť aj určité hranice na pomery váh (vyjadrujúce, nakoľko maximálne alebo minimálne má jeden vstup/výstup vyššiu prioritu ako druhý). Toto obmedzenie budeme v práci používať[26].

4.2 Aplikácia na HPI dáta

My teda chceme maximalizovať efektivitu krajín, pričom vezmeme do úvahy len údaje, ktoré už boli použité v kapitole 3, aby naše porovnávanie s výsledkami získanými v nej bolo korektné.

Bolo nám odporučené [6], aby sme všetky tri zložky indexu brali ako výstupy a maximalizovali ich. Hodnoty okúseného blahobytu a hodnoty dĺžky života sú v tomto smere v poriadku, avšak hodnoty ekologickej stopy musíme upraviť, aby pre nás bolo výhodné ich maximalizovať. Spravíme to nasledovne (v rovnici na hodnotu ekologickej stopy použijeme skratku ek.st.):

$$upravená\ ek.st. = [1 + max(ek.st.)] - (ek.st.)$$

Následne chceme, podobne ako pri tvorbe HPI, zaručiť, aby žiadna zložka nedominovala, teda aby ideálne váhy nemali tendenciu nejak významne zvýhodňovať niektoré zložky oproti iným. Spravíme to tak, že všetky naškálujeme do rozmedzia medzi 0 a 1, podľa jednoduchého vzťahu, pre každú zložku zvlášť:

$$upravená\ zložka = \frac{zložka}{maximálna\ hodnota\ zložky}$$

Kvôli tomuto pri samotnom procese DEA ešte určíme obmedzenia na váhy tak, aby ich pomer nemohol klesnúť, respektívne stúpnuť nad stanovenú hodnotu, ktoré sme spomínali už v 4.1.

[26] Teória v pozadí je samozrejme omnoho komplikovanejšia, je to aplikácia lineárnej optimalizácie. V práci neuvádzame konkrétne definície daného prístupu, ale iba spôsob jeho aplikácie na naše dáta. Podrobnosti DEA modelovania sú v zdroji [5].

Samotnú realizáciu tejto metódy vykonávame pomocou programu Matlab v nami urobenom skripte, ktorý je priložený ako príloha B. Následné štatistické spracovanie vykonávame, tak ako v kapitole 3, v programe R a využívame poznatky zo zdrojov [18],[19] a [20].

Prvú analýzu vykonáme pre tzv. jednotkové váhy, teda pomery váh zhora i zdola ohraničíme jednotkou. Robíme tak preto, nakoľko to vyjadruje, že jednotlivé zložky indexu majú úplne rovnakú dôležitosť. O to sa snažili i tvorcovia indexu, keď sa rôznymi štatistickými úpravami popísanými v 2.3.1 usilovali prispôsobiť disperzie jednotlivých zložiek indexu. No vzhľadom na fakt, že sme teraz pri príprave dát na DEA vyrovnávali disperziu inak než tvorcovia indexu a napokon i samotná DEA analýza pracuje s iným vzorcom, nemusia nám tieto váhy priniesť želaný výsledok v podobe podobnosti efektivity HPI a DEA. Spravíme teda ďalšie analýzy ešte pre hranice pomeru váh 0,5 a 2 - jedna váha bude maximálne dvojnásobkom inej, tiež pre 0,25 a 4 - jedna váha bude maximálne štvornásobkom inej.

Najprv porovnáme priemery efektivít, ktoré si znázorníme v nasledujúcej tabuľke:

Tabuľka 2: Priemery efektivít

Pomer váh/sada dát	1	0,5-2	0,25-4
Nová	0,8539598	0,9045674	0,9339359
Stará	0,8245249	0,8784737	0,9177352

Okamžite odtiaľ možno vypozorovať nasledované:

1. Pri akomkoľvek pomere dát je priemerná efektivita nových údajov cez DEA vyššia ako tých starých.

2. Pre nové i staré dáta nám priemerná efektivita stúpa spolu s uvoľňovaním požiadaviek na váhy.

Prvý fakt môže byť mätúci, nakoľko DEA určuje analýzu relatívne, takže to nemusí znamenať nejaké celkové objektívne zlepšenie v náraste hodnôt blahobytu, dĺžky života, alebo poklesu hodnoty ekologickej stopy. Vyjadruje iba to, že rozsah produktivity krajín sa zmenšil. To vlastne znamená, že dané hodnoty údajov sú si v celom datasete

bližšie, čo už ale bolo odpozorované v 3.2.1, kde sme analyzovali rozdiely disperzií medzi týmito dvoma sadami dát.
Podobne ako vtedy pre medziročnú zmenu disperzie HPI, aj táto medziročná zmena priemeru efektivít sa nám ukázala signifikantná, pričom sme vďaka vyhovujúcim vlastnostiam údajom efektivít ohľadom normality mohli zakaždým použiť silné testy[27].

Druhý fakt vyjadruje istú všeobecne platnú normu pri DEA, že hodnoty efektivít jednotlivých útvarov stúpajú s uvoľnujúcimi sa váhami. Je to preto, že daný útvar si môže po uvoľnení pomeru na váhy vo väčšej miere maximalizovať svoju efektivitu - dať zložke, ktorá má vyššiu hodnotu ešte väčšiu váhu eko predtým, a naopak zložke, ktorá má nižšiu hodnotu ešte menšiu váhu. Preto sa pri úplne voľných váhach neraz stane, že niektorá z nich je nulová.

Ukazuje sa, že pri všetkých troch druhoch obmedzení váh vykazuje umiestnenie krajín podľa HPI a DEA efektivity novej sady dát meraní menšiu zhodu, ako starej. Môže to byť spôsobené niektorými dodatočnými štatistickými úpravami, ktoré neboli vykonávané pri predošlej tvorbe indexu.
Ďalším zaujímavým pozorovaním je, že z pomedzi vybraných pomerov váh sa nám pri dátach zo starej tvorby indexu najlepšia zhoda porania neukazuje pri jednotkovom pomere, ako by sme to mohli čakať, ale pri pomere medzi 0,5 až 2. Zdôvodňujeme to jednak tým, že v tomto prípade ešte vplyv disperzie tvorcovia indexu úplne nevyrovnali a jednak tým, čo sme už spomínali vyššie, že jednotkové váhy po našej úprave disperzií pre proces DEA nemajú na dáta identický efekt ako tvorenie HPI.
Naproti tomu pri nových dátach nám vyšli pomerne podobné, no nie uspokojivé miery zhôd pri jednotkovom pomere váh a pomere medzi 0,5 až 2. Tvrdíme, že sa tak deje preto, že tu je už vplyv disperzií na tvorbu HPI vyrovnanejší, stále však platí nezhoda v metodike v normovaní. Najvoľnejšie váhy, tj pomer 0,25 až 4, nám pre nové i staré údaje priniesli najmenšiu zhodu. Je to očakávané, nakoľko sa do efektivity niektoré zložky zahŕňajú omnoho viac ako iné, čo je v protiklade s myšlienkou efektivity podľa HPI.

[27] *„vyhovujúce vlastnosti dát ohľadom normality“* a *„silné testy“* sme ako pojmy korektne štatisticky definovali v 3.1.

5 Vlastný výskum - Index šťastia jednotlivca

V tejto kapitole uvedieme náš vlastný index šťastia, ktorý sme nazvali Index šťastia jednotlivca a budeme ho v práci označovať skratkou IŠJ. Ukážeme si techniku jeho tvorby, povieme si čosi o spôsobe zbierania dát prostredníctvom online dotazníku a následne budeme tieto dáta analyzovať. Urobíme iba niektoré vybrané analýzy, nakoľko sa nám podarilo nazbierať veľmi veľké množstvo dát. Pre niektoré analýzy môžeme totiž použiť údaje až od 586 respondentov.

Budeme používať štatistické a matematické nástroje spomínané v predošlých kapitolách, teda štatistické analýzy budeme vykonávať v programe R využívajúc najmä vedomosti z [18],[19] a [20], pričom použitý skript uvádzame ako prílohu C. Vedomosti potrebné na DEA zasa z [5] a [6], aplikujúc ich v programe matlab, využívajúc najmä skript uvedený ako príloha B. Niektoré úpravy dát sme robili aj programe Excel.

5.1 Teória tvorby

Cieľom Indexu šťastia jednotlivca je vytvoriť verziu HPI pre konkrétnu osobu. Nakoľko však pre jednotlivca neexistuje priemerná dĺžka života, keďže svoj život ešte neodžil, a nie sme schopní ani zmerať jeho ekologickú stopu, pokúsili sme sa nejak poňať ideu HPI a pretransformovať ju do podoby vhodnej pre jednotlivca. HPI definoval čosi ako celoživotné šťastie delené cenou za šťastie, pričom celoživotným šťastím bol myselný indikátor HLY a cenou za šťastie ekologická stopa. V HPI išlo o to, aby dané celoživotné šťastie mohli prežívať aj ďalšie generácie. My sme upustili od celoživotného šťastia a v rámci nášho výskumu toto šťastie - šťastie jednotlivca, definujeme ako jeho aktuálnu celkovú spokojnosť samého so sebou. To budeme vyjadrovať zložka Vážené Šťastie, ktorú niekedy nazývame skratkou VŠ a charakterizujeme ju neskôr. A cenou za šťastie máme zasa na mysli, do akej miery bude môcť toto šťastie prežívať ďalej vo svojom živote. Preto sme zvolili mieru zdravia, čo bude upravená hodnota indexu telesnej hmotnosti, ktorý ďalej spomíname len v skratke anglického názvu - BMI. Idea znázornená rovnicou teda bude:

$$IJŠ = \frac{Vyjadrenie\ Šťastia}{Vyjadrenie\ ceny\ za\ Šťastie} = \frac{(Vážené\ Šťastie)}{(upravené\ BMI)} \tag{7}$$

5.1.1 Vážené Šťastie

Za zložkou Vážené Šťastie stojí nasledovná myšlienka:
Ak má človek nejakú oblasť v živote ako prioritnú, jeho spokojnosť v nej sa viac prejaví na jeho celkovom šťastí. Naopak, ak mu na danej oblasti skoro nezáleží, jeho spokojnosť v nej sa na jeho celkovom šťastí takmer neprejaví. Uvedieme príklad:
Ak človek cíti, že športovanie by malo byť akýmsi jeho životným smerovaním, chcel by sa napríklad stať vrcholovým športovcom, tak jeho nespokojnosť v jeho športových snaženiach, napríklad s napĺňaním kondičných plánov, mu môže naozaj vážne znížiť hladinu jeho šťastia. A zasa, ak sa mu darí, môže to dlhodobo držať hladinu jeho šťastia pomerne vysoko. Na druhej strane, ak mu na športe až tak nezáleží, nemá v tejto oblasti nejaké veľké ambície, tak či už úspech alebo neúspech v napĺňaní nejakých jeho športových plánov nebude nejak výrazne ovplyvňovať hladinu jeho šťastia.
Metodiku tvorby si popíšeme neskôr. Našou ideou bolo, aby táto zložka mohla pre jednotlivca komplexnejšie charakterizovať jeho osobitnú hladinu šťastia, než odpoveď na otázku Rebrík života spomínanú v 2.2, ktorou sa určoval okúsený blahobyt pre HPI. Celkovo sme do váženého šťastia zahrnuli nasledovných deväť oblastí, ktoré je nutné najprv zoradiť podľa svojich životných priorít zostupne od 8 po 0 a potom nezávisle od toho ohodnotiť spokojnosť v každej z nich na škále od 1 do 10. Oblasti aj s ich krátkym popisom sú nasledovné:

a) Každodenné vzťahy - bežné, nie príliš osobné vzťahy, ich budovanie a udržiavanie - napr. vzťahy s spolužiakmi zo školy, spolupracovníkmi na pracovisku, známymi, susedmi, náhodnými známosťami.

b) Užšie vzťahy - blízke vzťahy, ich budovanie a udržiavanie - napr. vzťahy s rodinou, s partnerom, s blízkymi priateľmi.

c) Voľný čas: aktívne využívanie - aktívne využitie voľného času. Činnosti, ktoré sa robia spravidla v kolektive, ale nie je to podmienkou - športové aktivity, kultúrne podujatia, zážitkové akcie.

d) Voľný čas: pasívny oddych - pozeranie TV, internetových stránok, čítanie bulváru, poobedný spánok.

e) Príroda - vzťah s prírodou, zvieratami, rastlinami, turistika, ochrana prírody, chovateľstvo, pestovateľstvo.

f) Kreativita - činnosti vyžadujúce istú dávku kreativity a fantázie. Akákoľvek tvorba, akákoľvek forma umenia, ako aj sýtenie sa akoukoľvek formou umenia.

g) Intelekt - činnosti vyžadujúce istú mieru intelektu, napr. vedecký výskum, čítanie odbornej literatúry, ale aj riešenie rôznych logických úloh, napr. sudoku.

h) Duchovno - vzťah k akémusi božstvu alebo životným princípom, snaha o žitie podľa tejto viery/ týchto princípov.

i) Štúdium, Práca - štúdium a práca vzhľadom na možnosti zárobku (potešenie z danej činnosti by malo spadať do niektorej z predošlých kategórii).

Vážené šťastie sa z týchto údajov vypočítavalo podľa vzorca:

$$V\check{S} = \sum_{j=1}^{9} [Priorita\ v\ oblasti]_j\ .\ [Spokojnosť\ v\ oblasti]_j\ . \tag{8}$$

5.1.2 Upravené BMI

Informácie o BMI sa vo viacerých zdrojoch jemne líšili, pre našu prácu sa stal smerodajným zdroj [22]. Za nasledovnými úpravami pôvodných hodnôt BMI počítaných z výšky a váhy respondentov na konečný indikátor zdravia zvaný „upravené BMI“ stoja dve nasledovné požiadavky:

1. Kvantifikovať „zhoršenie“ indikátora rovnakou mierou nahor i nadol od ideálnej hodnoty

2. Naškálovať jeho hodnoty do rozmedzia od 0,8 po približne 1,25.

Prvú požiadavku sme si určili kvôli povahe BMI, kde je ideálna hodnota naozaj kdesi uprostred a zhoršenie indikátora sa prejavuje obojstranne. Preto bude v neskoršom vzorci pre upravené BMI vystupovať absolútna hodnota rozdielu BMI a konštanty 2,3: |BMI-2,3|. Konštanta 2,3 sa podĺa zdroja [BMI] nachádza v pásme pre ideálnu hodnotu a my ju budeme ďalej označovať ako ideálnu hodnotu BMI. Vzali sme hodnotu o čosi vyššiu než priemer tohto pásma z dôvodu, že ak BMI ukazuje podváhu, zdravotné riziko stúpa omnoho rýchlejšie, ako keď ukazuje nadváhu.

Druhú požiadavku čerpáme z toho istého zdroja, bola nám však odkomunikovaná aj viacerými respondentmi dotazníku. A síce, že BMI zďaleka nie je ideálna miera zdravia jednotlivca. Preto sme sa rozhodli v niektorých analýzať použiť iba hodnotu indikátora Váženého Šťastia a dať samotnému BMI v IŠJ pomerne malú váhu - pri ideálych mierach spôsobí násobenie hodnoty Váženého Šťastia faktorom 1,25 a pri najhoršej možnej miere obezity 2. stupňa spôsobí delenie hodnoty Váženého Šťastia týmto faktorom. Pripustili sme i väčšiu váhu tohto indikátora, ale len v prípade, že sa BMI nachádza v oblasti obezity 3. stupňa, ktorá zo sebou nesie veľmi vážne zdravotné riziká. Touto úpravou sa teda budú zaoberať ostatné časti vzorca pre indikátor „upravené BMI“.

$$Upravené\ BMI = 0,45\ .\ \frac{|BMI - 2,3|}{(3,99 - 2,3)} + 0,8 \tag{9}$$

Vzorec sme zámerne nijak nezjednodušovali, aby sme tak demonštrovali úpravy plynúce z hore uvedených požiadaviek. 3,99 je hodnota BMI spomínaná ako vrchná hranica obezity 2. stupňa. 0,45 je rozsah, v ktorom sa po naškálovaní nachádzajú hodnoty medzi ideálnym BMI a touto hranicou. 0,8 vyjadruje posun hodnôt, aby boli v nami požadovanom rozpätí 0,8 po približne 1,25. Zodpovedá teda tiež aj ideálnej hodnote BMI.

5.1.3 Informácie k dotazníku

Ešte pred samotným skúmaním dát napíšeme čosi o ich zhromažďovaní. Bol použitý online dotazník, uvedený ako zdroj [8], ktorý nám nazbierané informácie automaticky ukladal do prehľadnej excelovskej tabuľky. Tú sme si po uzavretí dotazníka stiahli a niektoré základné úpravy dát robili priamo v nej.

Počet respondentov zďaleka prekonal naše očakávania, a síce ku dňu uzávierky dotazník obsahoval 594 odpovedí. Po odstránení niektorých skušobných vzoriek a údajov, o ktorých sme vedeli, že sú duplicitne, nám ostalo 586 odpovedí. Mnoho respondentov však nepochopilo správne požiadavku na zoradenie oblastí podľa priorít a napísali niektoré číslo priority viac ráz, čím sa pre nás ich údaje vzhľadom na tvorbu IŠJ stali nepoužiteľné. Na tú sme mohli použiť len 337 odpovedí. Aby sme však nejak plne

využili potenciál všetkých odpovedí, rozhodli sme sa aj v rámci výskumu naších dát použiť proces DEA, na ktorý dáta nepotrebujú spĺňať podmienku skriktného zoradenia oblastí podľa priorít.

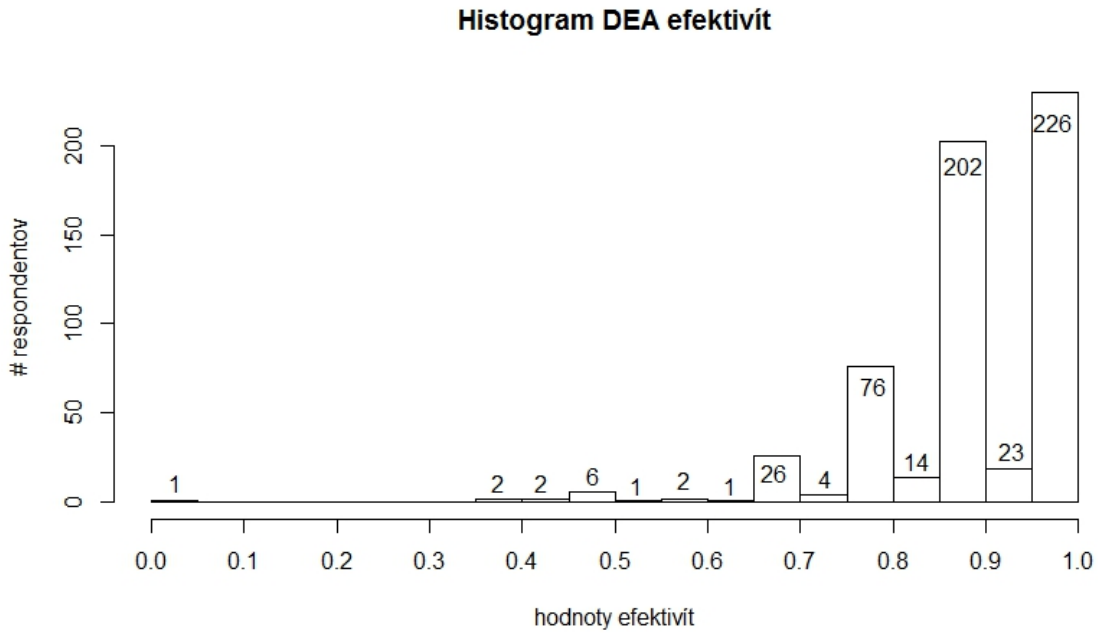
Okrem samotných informácií potrebných na Vážené Šťastie a BMI dotazník zisťoval niekoľko základných informácii o respondentoch - oblasť Slovenska, kde strávili väčšinu života (na výber Západné/Stredné/Východné Slovensko a zahraničie), či pochádzajú z dediny, mesta alebo veľkomesta, koľko majú rokov atď. Tieto údaje použijeme na rôznu formu vytvárania delení údajov na následné testovanie hypotéz. Napríklad, či sú respondenti nad istú vekovú hranicu nejak významne šťastnejší, ako tí pod ňou.
Dotazník ešte obsahoval aj nepovinný identifikačný prvok, vďaka ktorému si po mailovom vyžiadaní mohli respondenti zistiť svoje skóre.

5.2 DEA analýza všetkých dát

Výhodu tejto metódy sme avizovali už vyššie, a síce, že túto analýzu môžeme robiť pre všetky dáta, teda nie len tie spĺňajúce predpoklady presného zoradenia životných oblastí podľa priorít.
Nevýhodou však je, že pri takomto obrovskom množstve dát je proces DEA už pri voľných váhach pomerne komplikovaný[28], a isté percento výsledkov nie je dôveryhnodné. Preto nekladieme požiadavky na maximálne/minimálne pomery váh, ako tomu bolo v predošlej kapitole. A vzhľadom na to DEA model, ktorý používame, vyhodnotí efektivitu každého dotazníka, ktorý má aspoň v jednej zložke maximálnu spokojnosť - 10, ako maximálnu - 1 (100%). Je tak preto, že môže zanedbať všetky ostatné zložky (dať im nulovú váhu) a táto zložka sa bude ako jediná započítavať do efektivity s akoukoľvek veľkou váhou. Podobne by to fungovalo s maximálnou hodnotou 9 - dosiahle efektivitu 0,9 (90%) a tak ďalej. Tento výskum by sa teda dal interpretovať aj tak, že určuje, aké maximálne percentuálne šťastie vzhľadom na ostatných ľudí vypĺňajúcich dotazník by človek mohol dosiahnuť, ak by si priority určoval úplne ľubovoľne ako nezáporné celé čísla. Neplatilo by to iba v prípade, ak by v niektorej oblasti žiadny z respondentov nemal hodnotu 10, veľmi jednoducho sme však overili, že takýto prípad nenastal.
Teoretická štruktúra výsledkov je teda zrejmá, ako sme však už spomínali vyššie, pri ta-

[28]Matlab týmto skriptom rieši 586-krát úlohu LP s rozmerom matice 586*10

kýchto veľkých DEA ľahko príde k určitým nepresnostiam, preto niektoré hodnoty vyšli medzi týmito hladinami efektivít. V skutočnosti nám teda povie iba to, aké množstvo ľudí dosiahlo najvyššiu spokojnosť z daných oblastí na hodnote 10, aké 9 a podobne. Dokážeme tiež zistiť, aký je počet chybových údajov medzi každými dvoma kategóriami. To všetko uvádzame v nasledujúcom histograme, v ktorom značka # znamená počet a je uvádzaná z dôvodu nekompatibility softvéru na tvorbu grafu a niektorými znakmi slovenskej diakritiky. Bude použitá aj pri ďalších histogramoch.



Histogram1: DEA efektivita respondentov nášho výskumu.

Vždy naľavo od desatinnej hodnoty je stĺpec vyjadrujúci počet respondetov s DEA efektivitou zodpovedajúcou tejto hladine a stĺpce medzi tým vyjadrujú počet údajov efektivít medzi nimi. Osobitný je stĺpec pri nule, ktorý vyjadruje, že pre jedného respondenta proces DEA nedokázal zrátať hodnotu efektivity.

5.3 Analýzy VŠ a IŠJ selekcie dát

Medzi ďalšími analýzami a tou predošlou robenou cez DEA sú dve zásadné rozdiely, ktoré však spolu súvisia.

Prvým je, že sa teraz neusilujeme o maximalizáciu šťastia jednotlivcov, ale zachovávame

myšlienku, s ktorou sme index tvorili - teda, že prioritnejšie oblasti majú na tvorbe hodnoty subjektívneho šťastia väčšiu váhu, ako tie menej prioritné.
Druhým rozdielom je zmenšený počet dotazníkov, ktorých dáta takto môžeme skúmať, nakoľko sa nám nepodarilo softvérovo ošetriť v dotazníku zadávanie niektorej váhy duplicitne pre viaceré oblasti. Jediné, čo sme boli schopní urobiť, bolo následne zo všetkých odpovedí vyfiltrovať tie nevhodné. Po tejto úprave nám na VŠ respektíve IŠJ analýzy ostalo 337 vzoriek údajov.

Treba však podotknúť, že hoci tieto výsledky budú mať o čosi vyššiu výpovednú hodnotu ako tie získané predošlým DEA procesom, aj tak však budú mať určité nedostatky. Tie nám boli odkomunikované osobne - a síce:

a) Viacero respondentov si nebolo celkom istých, ako dané oblasti zoradiť podľa priorít a podľa vlastných slov by to možno zoradili inak, ak by dotazník vypĺňali znova

b) Niektorí respondenti si plne neuvedomili fakt, že priotity vôbec nemusia súvisieť s časom a úsilím venovaným danej oblasti, ale hlavne s voľou mať toto úsilie, a tým pádom jemne korelovali údaje priorít s údajmi spokojnosti. Teda oblasti, kde dosahovali menšiu spokojnosť, ohodnotili nižšiou prioritou.

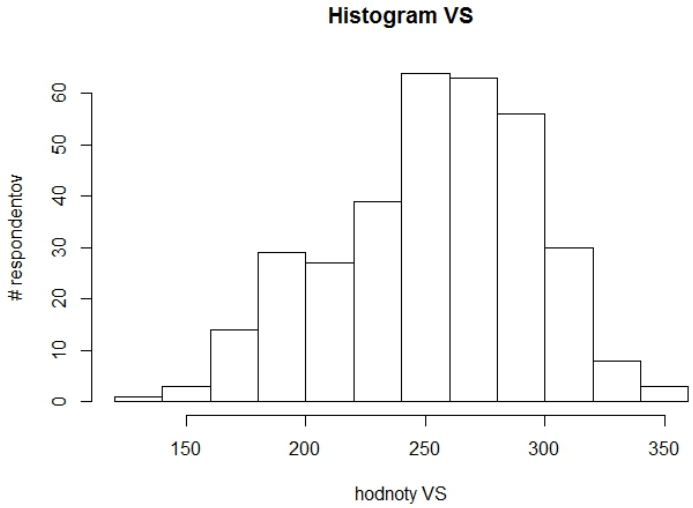
Napriek tomu však predpokladáme, že pri takom veľkom počte dotazníkov tieto údaje príliš nezavažujú a preto výsledky získané korektnými štatistickými metódami budeme považovať za hodnoverné.

Ako možno vidno na nasledujúcich histogramoch, VŠ je oproti IŠJ viac koncentrované, nakoľko v IŠJ pribúda ešte jeden rozptyľujúci prvok - upravené BMI. Teda zatiaľ čo je maximum VŠ na hodnote 360, maximum IŠJ je na hodnote 450, čo spôsobuje, že údaje VŠ sú rozptýlené len na intervale [127,351][29], a údaje IŠJ až na intervale približne (147,416)[30]. Taktiež možno na nasledujúcich histogramoch empricky overiť očakávaný jav - že rozdelenie výsledkov VŠ (v histograme VS) a IŠJ (v histograme ISJ) sa približne správa podľa normálneho rozdelenia. Teda okolo priemernej hodnoty je najviac výsledkov a smerom k maximálnym ako aj minimálnym hodnotám sa početnosti výsledkov

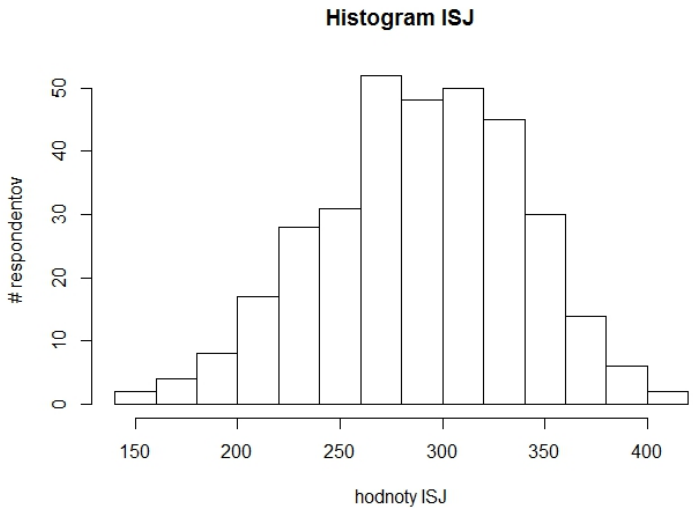
[29]Typom zátvoriek [] myslíme vrátane týchto hodnôt, nakoľko presne zodpovedajú minimálnej a maximálnej hodnote.

[30]Presné minimum 147,9037 a maximum = 415,4954.

znižujú.



Histogram2 - Histogram demonštrujúci náhodné rozdelenie výsledkov respondentov v indikátore VŠ.



Histogram3 - Histogram demonštrujúci náhodné rozdelenie výsledkov respondentov v indikátore IŠJ.

Všeobecné informácie k nasledovnému výskumu

V nasledujúcej časti budeme deliť spomínaných 337 vzoriek údajov vypočítaného VŠ a IŠJ, postupne podľa nasledovných štyroch kritérii: Vek; Prostredie, z ktorého respondent pochádza; Prostredie, v ktorom respondent aktuálne žije; región Slovenska.

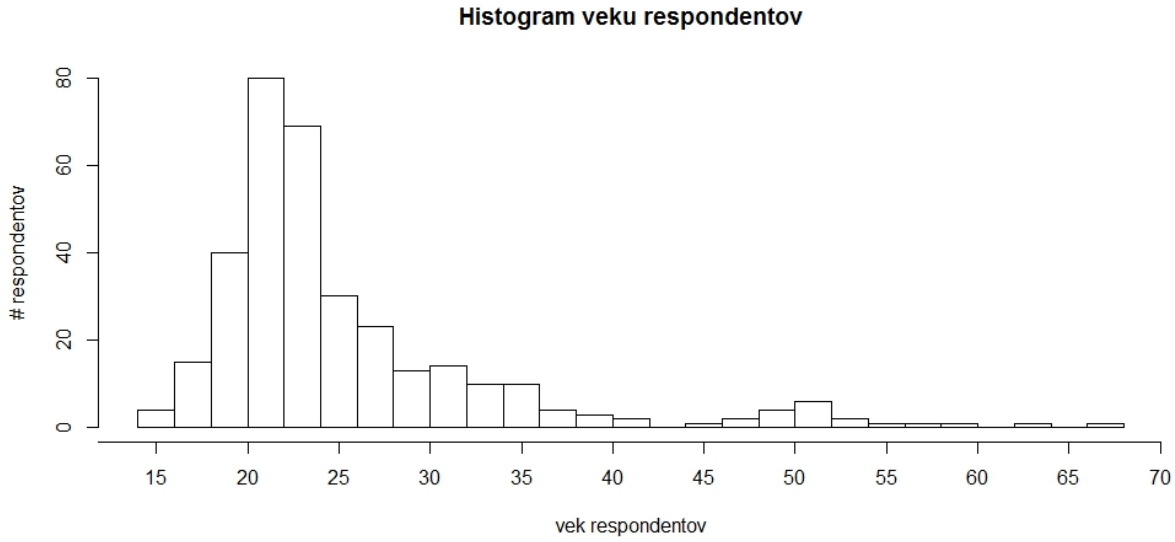
Budeme skúmať štatistickú významnosť rozdielu priemerov týchto skupín navzájom. Používame pojmy, napríklad „silné testy“ a „klasické testovanie“, ktoré boli definované v 3.1.

Teraz uvedieme niektoré špecifikácie, ktoré platia všeobecne pre všetky následné výskumy. Pri akomkoľvek delení nám vzniknuté podskupiny dát VŠ a IŠJ spĺňali požadované vlastnosti ohľadom normality a preto pri všetkých analýzach môžeme používať silné testy[31]. Navyše pre všetky porovnávané dvojice dát bola testom na zisťovanie rovnosti disperzií pripustená rovnosť a preto môžeme pristúpiť ku klasickému testovaniu významnosti rozdielu priemerov skupín dát.
Takmer vo všetkých prípadoch sú priemery podskupín VŠ a IŠJ v rovnakom vzťahu - napríklad ak je priemer VŠ jednej skupiny dát vyšší ako druhej, deje sa tak aj pri IŠJ. V IŠJ možno pozorovať väčší rozdiel, ktorý je však spôsobený väčším rozptylom dát, a zväčša sa v oboch kategóriách pripúšťa signifikantný rozdiel v rovnakých prípadoch.

Delenie podľa veku

Vzhľadom na fakt, že väčšina účastníkov je skôr v mladšom veku, ako to ukazuje nasledujúci histogram, sme sa rozhodli rozdeliť účastníkov podľa mediánu a nie podľa priemeru. Mediánový vek bol rovný 23.

[31] Ako zaujímavosť ešte možno spomenúť, že úprava VŠ na IŠJ má za následok, že požadované vlastnosti ohľadom normality dát sú omnoho lepšie splnené pre IŠJ dáta - p-hodnota príslušných Kolmogorových Smirnových testov je rádovo o niekoľko desiatok percent vyššia.



Histogram4 - Histogram ukazujúci vekové rozloženie respondentov tu používanej vyselektovanej časti dotazníku.

Vhľadom na snahu o rovnomerné rozdelenie dát, sme do jednej skupiny dali údaje osôb s vekom menším a rovným a do druhej s väčším ako je medián. Počty dát v skupinách vyšli 181 a 156. Pri VŠ aj IŠJ nám vyšli podobné hodnoty priemerov a ani pri jednom z nich nebol tento rozdiel signifikantný.

Delenie podľa pôvodu a podľa aktuálneho bydliska

V týchto dvoch kategóriách si respondenti vyberali, z akého prostredia pochádzajú a v akom prostredí žijú, pričom možnosti boli: dedina, mesto a veľkomesto (nad 100-tis. obyvateľov). Pre VŠ a IŠJ vyšli výsledky nasledovných analýz úplne totožné, preto budeme spomínať iba indikátor VŠ.

Väčšina účastníkov, v súčastnosti žije vo veľkomeste (220 údajov) a zvyšok je skoro rovnomerne rozdelený (64-mesto,53-dedina). Preto, hoci v tejto kategórii vyšiel signifikantný vyšší priemer VŠ pre respondentov z veľkomesta od respondentov z dediny, nedávame tomuto výsledku nejakú obrovskú váhu.

Naproti tomu rozdelenie respondetov podľa pôvodu bolo pomerne rovnomerné, z dediny 98, z mesta 123 a z veľkomesta 116. Pričom nám bol potvrdený signifikantne lepší priemer respondentov z dediny od respondentov z veľkomesta a pomerne tesne nesignifikantný rozdiel týchto respondetov od respondentov z mesta[32]. To sú pomerne

[32]Pre VŠ p-hodnota iba pár stojín nad 5%, pre IŠJ čosi nad 8%.

očakávané výsledky, nakoľko sa vo všeobecnosti tvrdí, že ľudia pochádzajúci z dedinského prostredia sú akýsi vyrovnanejší a šťastnejší. Práve veľká časť z nich teraz žije vo veľkomeste - napríklad kvôli štúdiu. Aj preto nepovažujeme výsledky prvých analýz v tejto kategórii za prielomové.

Delenie podľa regiónu

V tejto kategórii si respondenti vyberali, v ktorom regióne Slovenska prežili väčšinu života. Mali na výber Západné, Stredné a Východné Slovensko. Na výber bolo aj zahraničie, no tieto vzorky boli iba 4, preto sme ich nezahrnuli do tejto analýzy, porovnávali sme teda iba regióny Slovenska.

Zaujímavým výsledok tohto testovania je, že hoci priemer VŠ aj IŠJ vyšiel skoro rovnaký pre Západné a Stredné Slovensko, vyšiel signifikantne vyššií, ako ten pre Východné Slovensko. Dôvodom môžu byť ťažšie podmienky pre život na Východnom Slovensku - napríklad vysoká nezamestnanosť.

Záver

Cieľom práce bolo poskytnúť čitateľovi všeobecný prehľad o problematike merania šťastia, ako aj spomenúť niektoré vybrané indexy šťastia a pomocou ich analýzy demonštrovať význam ich tvorby a používania v súčasnom svete.
V našej práci analyzujeme Happy planet index, ktorý je podľa viacerých zdrojov ([7],[11],[13] a [17]) svetovo známy a zameriava sa na meranie blahobytu a jeho udržateľnosti vo svetovom meradle. Okrem neho analyzujeme druhý index, ktorý bol navrhnutý nami pre účely bakalárskej práce. Ten sa zasa zameriava na meranie aktuálnej spokojnosti a jej udržateľnosti, pričom je zameraný na jednotlivca.

Úvod a kapitola 1 boli zamerané na všeobecný náhľad na problematiku. Ich cieľom bolo čitateľa oboznámiť s významom a rozsahom problematiky merania šťastia. Význam sme ukázali v skupinovej až národnej úrovni, kde sme uviedli možnosť lepšie kvantifikovať pokrok v danej spoločnosti, nakoľko sme na viacerých miestach ukázali doteraz najpoužívanejšiu mieru používanú za týmto účelom HDP - ako nedostatočnú, nakoľko zachytáva iba ekonomickú výkonnosť, ktorá nemusí priamo súvisieť s blahom obyvateľov. Rozsah sme zasa ukázali na rozmanitosti a veľkosti historických i súčasných výskumných projektov - GNH pre Bhutan, OECD Better Life Index pre 36 krajín, HPI pre približne 150 krajín, mnohé iné národné orgány ustanovené za podobným účelom a napokon zmienkou o rezolúcii OSN, ktorá vyzýva členské štáty k tvorbe podobných meraní.

V nasledovných kapitolách sme za zamerali na konkrétne indexy. Prvým bol HPI. Jeho analýza bola v troch kapitolách rozdelená nasledovne:
Kapitola 2, uvádzala množstvo informácii, ktoré sa ho týkali, od úseku venovanému motivácii k jeho skúmaniu a jeho základným črtám (podkapitola 2.1), ktorý obsahoval aj základnú rovnicu jeho tvorby (1); cez časť o pracovnom rámci, v ktorom je toto meranie ukotvené v predstavách jeho tvorcov (podkapitola 2.1.2), až po konkrétny popis jeho zložiek (úsek 2.2) a konkrétne štatistické postupy, ktorými z nich vzniká (posledná časť - 2.3). Touto kapitolou sme chceli čitateľa s HPI zoznámiť, jednak aby pochopil jeho komplexnosť a jednak aby mal prehľad o tom, čo sa robí jeho následných analýzach v ďalších kapitolách.
Kapitola 3 sa venovala konkrétnej štatistickej analýze zložiek HPI, indikátora HLY

a aj HPI samotného. Bola rozdelená na dve hlavné podkapitoly. Prvá (3.2) skúmala zmeny medzi aktuálnou a staršou sadou dát slúžiacich na tvorbu indexu, a podľa typu výskumu bola rozdelená ešte na ďalšie časti. Druhá (3.3) skúmala rozdiely medzi subregiónmi vytvorenými z údajov všetkých krajín medzi sebou ako aj ich medziročné zmeny, tá bola zasa rozdelená podľa objektu záujmu. Text je písaný voľným štýlom, zrozumiteľnou rečou, pričom definície v 3.1 a poznámky pod čiarou prinášajú matematicky korektnejšie vysvetlenia.
Kapitola 4 uvádzala alternatívu k tvorbe HPI z jeho zložiek, a za tú sme zvolili proces DEA. Bola rozdelená na teoretickú časť - 4.1, ktorá veľmi všeobecne zhŕňala teóriu stojacu za touto matematickou technikou, a praktickú časť - 4.2, v ktorej sme uviedli a vysvetlili nami použitý postup na jej aplikovanie na naše dáta. Následne sme v nej ešte analyzovali získané výsledky. Cieľom tejto kapitoly bolo čitateľovi predstaviť túto alternatívu, jej klady i zápory oproti samotnej tvorbe HPI.

V poslednej kapitole 5 sme uvádzali a analyzovali náš vlastný index - tzv. Index šťastia jednotlivca. Najprv sme si uviedli motiváciu tvorby, ktorou bolo akési HPI pre jednotlivca, a potom uviedli konkrétnu metodiku, ktorú sme za týmto účelom vytvorili, kde sme uviedli aj konkrétne vzorce:

a) Vzorec na tvorbu celkového indexu - (7).

b) Vzorec na tvorbu indikátora Vážené Šťastie - (8).

c) Vzorec na tvorbu indikátora Upravené BMI - (9).

Zhnieme to teda nasledovne. Po krátkom prehľade problematiky sme komplexne rozobrali dva indexy šťastia, pričom sme dbali ako na teoretickú stránku tohto rozboru, spočívajúcu v popise pozadia a tvorby indexov, tak aj na praktickú stránku zahŕňajúcu hĺbkovú štatistickú analýzu dát a ukážku alternatívy k daných indexom. Všetky tieto analýzy sme však robili aj so zámerom, aby im bol schopný porozumieť aj človek bez hlbších matematických znalostí.

Práca je prínosná pre všetkých, ktorých zaujíma problematika merania šťastia a žiadajú si literatúru v slovenskom jazyku. Ďalej je určená tým, ktorý by sa chceli zoznámiť s ideou štatistickej analýzy hladiny či už národného, alebo subjektívneho šťastia bez hĺbkových znalostí v oblasti štatistiky.

Prínosom v oblasti merania šťastia je najmä porovnávanie HPI s alternatívnym zisťovaním efektivít krajín v podobe DEA analýzy a tiež tvorbou nového indexu šťastia, zozbieraním a analýzov vlastných dát k nemu prislúchajúcich.
Pre samotného autora boli najväčšími prínosmi možnosť oboznámenia sa danou mimoriadne zaujímavou problematikou a možnosť rozšíriť vlastné vedomosti v oblasti štatistických analýz a DEA modelovania.

Ako motiváciu na ďalší výskum v oblasti uvedieme nasledujúci fakt týkajúci sa našej práce. Totiž to, že sa totiž po krátkom úvode zamerala len na dva konkrétne indexy a hoci ich rozobrala pomerne komplexne, bolo by možné tak učiniť v omnoho hlbšej miere - deliť analyzovať aj konkrétne zložky indexu po subregiónoch, deliť dáta vo vlastnom výskume podľa ešte nevyužitých alebo kombinovaných kritérii. Už to naznačuje obrovský rozsah problematiky merania šťastia a potrebu ďalších prác, najmä v slovenskom jazyku. Námetom na bakalársku prácu by mohlo byť hĺbkové rozobratie Indexu šťastia podľa OECD a jeho súvisu s HPI, ku ktorej prvotné zdroje by pisateľ mohol čerpať z [14] a [15] . Ďalším námetom by bolo popísať a nejakým rozumným spôsobom roztriediť typy výskumov uvedených v zdroji [23]. Iná bakalárska práca alebo jej časť by mohla obsahovať návrh vlastného indexu a analýzu vlastných dát, vítaný je aj návrh na vylepšenie nášho indexu a ďaľšia hĺbková analýza nami získaných dát, ktoré v prípade záujmu ochotne poskytneme. Analýza GNH by kvôli svojej komplexnosti, čo sa týka robustnosti použitých matematicko-štatistických metód z časti opísaných v zdroji [17], sama mohla byť námetom dokonca aj na diplomovú prácu.

Zoznam použitej literatúry

[1] Databáza údajov NEF: HPI-2.0-Results, dátový súbor typu .xls, hárok: All Data

[2] Databáza údajov NEF: hpi-data, dátový súbor typu .xlsx, hárok: Complete HPI Databases, dostupné na internete na stiahnutie (20.05.2014): http://www.happyplanetindex.org/data/

[3] Frey B., Stutzer A.: *Happiness and Economics*, Princeton University Press, Princeton, 2002

[4] Galup World Poll: Scientifically measuring attitudes and behaviors worldwide, dostupné na internete (20.05.2014): http://www.gallup.com/strategicconsulting/en-us/worldpoll.aspx

[5] Halická M.: *DEA Modely*, učebné texty, FMFI UK, Bratislava, 2014, dostupné na internete (20.05.2014): http://www.iam.fmph.uniba.sk/institute/halicka/text/TextDEA39%2826_3%29.pdf

[6] Halická M.: osobná komunikácia, FMFI UK, Bratislava, 2014

[7] Happy Planet Index: The leading global measure of sustainable well-being, dostupné na internete (20.05.2014): http://www.happyplanetindex.org

[8] Janočko J.: *Index Šťastia Jednotlivca*, dotazník, dostupné na internete (20.05.2014): https://docs.google.com/forms/d/1nedeKbkjRzQAf8PrYiqfYqUhCgeiMOPNmOuDAuLk9RQ/viewform

[9] Marks N. et al.: *The Unhappy Planet Index*, preprint, The New Economic Foundation (nef), London, 2006, dostupné na internete (20.05.2014): http://www.nuigalway.ie/dern/documents/nef_report.pdf

[10] Marks N. et al.: *The Happy Planet Index 2.0*, preprint, The New Economic Foundation (nef), London, 2009, dostupné na internete (20.05.2014): http://base.socioeco.org/docs/the_happy_planet_index_2.0_1.pdf

[11] Marks N. et al.: *Happy Planet Index:2012 report*, preprint, The New Economic Foundation (nef), London, 2012, dostupné na internete (20.05.2014): http://www.happyplanetindex.org/assets/happy-planet-index-report.pdf

[12] Michalos A.: *Citation Classics from Social Indicators Research*, Springer, Dordrecht, 2006

[13] nef: Economics as if people and planet mattered, dostupné na internete (20.05.2014): http://www.neweconomics.org/

[14] OECD Better Life Index, dostupné na internete (20.05.2014): http://www.oecdbetterlifeindex.org/

[15] OECD: Better policies for better lifes, dostupné na internete (20.05.2014): http://www.oecd.org/

[16] Political Geography: Updates on the changing world political map, dostupné na internete (20.05.2014): http://www.polgeonow.com/2011/04/how-many-countries-are-there-in-world.html

[17] Sachs J., Helliwel J., Layard R.: *World Happiness report*, preprint, The earth Institute, Columbia University, Cape Town, 2012, dostupné na internete (20.05.2014): http://www.earth.columbia.edu/sitefiles/file/Sachs%20Writing/2012/World%20Happiness%20Report.pdf

[18] Somorčík J.: mailová komunikácia, FMFI UK, Bratislava, 2014

[19] Somorčík J.: Štatistické metódy, prednášky, FMFI UK, Bratislava, 2014

[20] Somorčík J.: Počítačová štatistika, prednášky, FMFI UK, Bratislava, 2014

[21] The Times of India: Matematical equation of eternal happiness, dostupné na internete (20.05.2014):
http://timesofindia.indiatimes.com/home/opinion/edit-page/Mathematical-equation-for-eternal-happiness/articleshow/261808.cms

[22] Výpočet: Výpočet BMI, Body Mass Index, dostupné na internete (20.05.2014):
http://www.vypocet.cz/bmi?utm_source=bodymassindex.cz&utm_medium=redirect

[23] World Database of Happiness: Archive of research findings on subjective enjoyment of life, dostupné na internete (20.05.2014):
http://www1.eur.nl/fsw/happiness/hap_quer/hqi_fp.htm

Príloha A - tabuľky medziročných zmien HPI a HLY po subregiónoch

V nasledovných tabuľkách budeme používať nasledujúce skratky: no. - nové, st. - staré, Kor.Koef. - korelačný koeficient.

„Po úprave“ znamená po procese odstraňovania outlierov a všetky ďalšie hodnoty (priemer, disperzia, korelácia) sú počítané po tejto úprave sady dát. „Po predošlej úprave“ v tabuľke pre HLY znamená, že berieme sadu dát, ktorá bola už predtým upravená pre analýzy HPI.

Subregión \ údaj	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Normalita st. dát	Áno	Áno	Áno	Áno	Áno	Áno	Áno
Normalita no. dát	Áno po úprave	Áno	Áno	Áno	Áno	Áno	Áno
Počet dát	21 po úprave	24	16	33	5	12	27
Priemer st. dát	59,209	42,042	46,479	27,968	54,228	51,977	42,14
Priemer no. dát	54,829	43,354	45,156	33,133	51,26	47,083	40,07
Významný rozdiel priemerov	Áno (pokles)	Nie	Nie	Áno (nárast)	Nie	Nie	Nie
Disperzia st. dát	68,593	36,184	106,051	33,913	3,498	65,228	41,669
Disperzia no. dát	22,133	27,541	55,403	29,551	17,273	52,051	39,988
Rovnost disperzií	Nie	Áno	Áno	Áno	Áno	Áno	Áno
Kor. koef.	0,61	0,551	0,591	0,315	0,396	0,75	0,796
Pripustenie Kor. Koef. 0	Nie	Nie	Nie	Tesne nie	Áno	Nie	Nie

Tabuľka zachytávajúca medziročné zmeny v HPI po subregiónoch

Subregión \ údaj	1	2	3	4	5	6	7
Normalita st. dát	Áno	Nie	Áno	Áno	Áno	Áno	Áno
Normalita no. dát	Áno	Nie	Áno	Áno	Áno	Áno	Áno po úprave
Počet dát	21 už upraven.	18	16	33	5	12	27
Priemer st. dát	51,029	59,983	43,038	20,619	35,26	46,208	39,864
Priemer no. dát	52,895	61,417	47,038	30,388	40,18	48,883	44,032
Významný rozdiel priemerov	Nie	Nie	Nie	Áno (nárast)	Áno (nárast)	Nie	Áno (nárast)
Disperzia st. dát	33,305	19,385	70,753	23,575	5,143	76,295	19,866
Disperzia no. dát	25,598	21,499	64,969	14,693	6,297	59,772	13,947
Rovnost disperzií	Áno	Nemohlo byť testované	Áno	Áno	Áno	Áno	Áno
Kor. Koef.	0,643	0,897 pearson 0,833 spearman	0,793	0,546	0,382	0,876	0,768
Pripustenie Kor. Koef. 0	Nie	Nie	Nie	Nie	Áno	Nie	Nie

Tabuľka zachytávajúca medziročné zmeny v HLY po subregiónoch

Príloha B - matlabovský skript na proces DEA

```
function [E,U,V]=CCR_O_MM_R(X,Y,P,Q)
% E,U,V,X,Y,P,Q su matice korektne definovane v zdroji [5]

if (nargin<2 || nargout==0)
disp('Nezadali ste dostatok vstupnych alebo vystupnych argumentov')
E=[];
U=[];
V=[];
return
end

E=[];
if nargout>1
U=[];
end
if nargout>2
V=[];
end
if nargin==2
P=[];
Q=[];
end
if nargin==3
Q=[]
end

if (size(Q,1)==0 && size(P,1)==0)
A=[Y',-X'];
else if ( size(Q,1)==0 || nargin==3)
A=[Y',-X';zeros(size(P',1),size(Y',2)),P'];
else if size(P,1)==0
A=[Y',-X';Q',zeros(size(Q',1),size(X',2))];
else
A=[Y',-X';zeros(size(P',1),size(Y',2)),P';Q',zeros(size(Q',1),size(P',2))];
end
end
```

```
end

for i=1:size(X,2)
yO=Y(:,i);
xO=X(:,i);
b=zeros(size(A,1),1);
f=[ zeros(1,size(yO,1)),xO];
Aeq=[yO',zeros(1,size(xO,1))];
x=linprog(f,A,b,Aeq,1,zeros(size(A,2),1));
d=x((size(yO)+1):size(x));
e=1/d;
E=[E;e];
if nargout>1
u=x(1:size(yO));
U=[U,u];
end
if nargout>2
v=x((size(yO)+1):size(x));
V=[V,v];
end
end
end
```

Príloha C - kód R k vlastnému výskumu

```
#Na DEA analyzu

E <- scan("C:/Users/Jozef/Desktop/Bakalárka/E-vl1.txt", dec=".")
hist(E,nclass=20, xlab="hodnoty efektivít",
ylab="počet respondentov", main="Histogram DEA efektivít",
xaxt='n')
axis(side=1, at=seq(0,1, 0.1))
pocty <- c(length(E[E==1]), length(E[E<1 & E>0.9]), length(E[E==0.9]),
length(E[E<0.9 & E>0.8]), length(E[E==0.8]),length(E[E<0.8 & E>0.7]),
length(E[E==0.7]),length(E[E<0.7 & E>0.6]),length(E[E==0.6]),
length(E[E<0.6 & E>0.5]),length(E[E==0.5]),length(E[E<0.5 & E>0.4]),
length(E[E==0.4]),length(E[E<0.4]))

text(locator(length(pocty)) , labels=pocty)

#Predpriprava vl. vyskumu

ci_ok <- ("C:/Users/Jozef/Desktop/Bakalárka/ci_ok_prior.txt") #vektor obsahujuci 1 a 0 dlzky poctu respondentov
#podla toho ci maju spravne zoradene priority
vlastny_vyskum <- read.table("C:/Users/Jozef/Desktop/Bakalárka/vlastny_vyskum.txt", header=TRUE,dec=".")
View(vlastny_vyskum)
vlastny_vyskum_sel <- vlastny_vyskum[ci_ok==1,]
View(vlastny_vyskum_sel)
attach(vlastny_vyskum_sel)
ISJ <- VS/upr_BMI
View(ISJ)
maxIJS <- 360/0.8sds
hist(ISJ, nclass="Freedman-Diaconis", xlab="hodnoty ISJ",
ylab="# respondentov", main="Histogram ISJ")
hist(VS, nclass="Freedman-Diaconis", xlab="hodnoty VS",
ylab="# respondentov", main="Histogram VS")

#Analyzy skupin dat

length(vek[vek>median(vek)])
#[1] 156
length(vek[vek<=median(vek)])
#[1] 181
VS_pod <- VS[vek<=median(vek)]
VS_nad <- VS[vek>median(vek)]
length(VS_nad)
mean(VS_nad)
mean(VS_pod)
ks.test(VS_nad,"pnorm",mean=mean(VS_nad),sd=sqrt(var(VS_nad))) #OK
ks.test(VS_pod,"pnorm",mean=mean(VS_pod),sd=sqrt(var(VS_pod))) #OK
```

```
var.test(VS_nad,VS_pod)
t.test(VS_nad,VS_pod, alternative="greater",var.equal=TRUE) #nie je sign

ISJ_pod <- ISJ[vek<=median(vek)]
ISJ_nad <- ISJ[vek>median(vek)]
length(ISJ_nad)
mean(ISJ_nad)
mean(ISJ_pod)
ks.test(ISJ_nad,"pnorm",mean=mean(ISJ_nad),sd=sqrt(var(ISJ_nad))) # viac OK
ks.test(ISJ_pod,"pnorm",mean=mean(ISJ_pod),sd=sqrt(var(ISJ_pod))) # viac OK
var.test(ISJ_nad,ISJ_pod)
t.test(ISJ_nad,ISJ_pod, alternative="less",var.equal=TRUE) #nie je sign

VS_dedina <- VS[povod=="dedina"]
VS_mesto <- VS[povod=="mesto"]
VS_velkomesto <- VS[povod=="velkomesto"]
length(VS_dedina)+length(VS_mesto)+length(VS_velkomesto) #337
mean(VS_dedina)
mean(VS_mesto)
mean(VS_velkomesto)
ks.test(VS_dedina,"pnorm",mean=mean(VS_dedina),sd=sqrt(var(VS_dedina)))
ks.test(VS_mesto,"pnorm",mean=mean(VS_mesto),sd=sqrt(var(VS_mesto)))
ks.test(VS_velkomesto,"pnorm",mean=mean(VS_velkomesto),sd=sqrt(var(VS_velkomesto)))
#vsetky tri p value aspon 0.15 - ok
var.test(VS_dedina,VS_mesto)
var.test(VS_dedina,VS_velkomesto)
var.test(VS_mesto,VS_velkomesto)
t.test(VS_dedina,VS_mesto, alternative="less",var.equal=TRUE) #tesne nie je sign
t.test(VS_dedina,VS_velkomesto, alternative="less",var.equal=TRUE) #je sig
t.test(VS_mesto,VS_velkomesto, alternative="less",var.equal=TRUE) #nie je sig

ISJ_dedina <- ISJ[povod=="dedina"]
ISJ_mesto <- ISJ[povod=="mesto"]
ISJ_velkomesto <- ISJ[povod=="velkomesto"]
length(ISJ_dedina)+length(ISJ_mesto)+length(ISJ_velkomesto) #337
mean(ISJ_dedina)
mean(ISJ_mesto)
mean(ISJ_velkomesto)
ks.test(ISJ_dedina,"pnorm",mean=mean(ISJ_dedina),sd=sqrt(var(ISJ_dedina)))
ks.test(ISJ_mesto,"pnorm",mean=mean(ISJ_mesto),sd=sqrt(var(ISJ_mesto)))
ks.test(ISJ_velkomesto,"pnorm",mean=mean(ISJ_velkomesto),sd=sqrt(var(ISJ_velkomesto)))
#Vsetky tri p value aspon 0.15 - ok
var.test(ISJ_dedina,ISJ_mesto)
var.test(ISJ_dedina,ISJ_velkomesto)
var.test(ISJ_mesto,ISJ_velkomesto)
t.test(ISJ_dedina,ISJ_mesto, alternative="less",var.equal=TRUE) # nie je sign
t.test(ISJ_dedina,ISJ_velkomesto, alternative="less",var.equal=TRUE) #je sig
t.test(ISJ_mesto,ISJ_velkomesto, alternative="less",var.equal=TRUE) #nie je sign
```

```
VS_dedina <- VS[zijem=="dedina"]
VS_mesto <- VS[zijem=="mesto"]
VS_velkomesto <- VS[zijem=="velkomesto"]
length(VS_dedina)+length(VS_mesto)+length(VS_velkomesto) #337
mean(VS_dedina)
mean(VS_mesto)
mean(VS_velkomesto)
ks.test(VS_dedina,"pnorm",mean=mean(VS_dedina),sd=sqrt(var(VS_dedina)))
ks.test(VS_mesto,"pnorm",mean=mean(VS_mesto),sd=sqrt(var(VS_mesto)))
ks.test(VS_velkomesto,"pnorm",mean=mean(VS_velkomesto),sd=sqrt(var(VS_velkomesto)))
#vsetky tri p value aspon 0.15 - ok
var.test(VS_dedina,VS_mesto)
var.test(VS_dedina,VS_velkomesto)
var.test(VS_mesto,VS_velkomesto)
t.test(VS_dedina,VS_mesto, alternative="greater",var.equal=TRUE) #tesne nie je sign
t.test(VS_dedina,VS_velkomesto, alternative="greater",var.equal=TRUE) #je sig
t.test(VS_mesto,VS_velkomesto, alternative="less",var.equal=TRUE) #nie je sig

ISJ_dedina <- ISJ[zijem=="dedina"]
ISJ_mesto <- ISJ[zijem=="mesto"]
ISJ_velkomesto <- ISJ[zijem=="velkomesto"]
length(ISJ_dedina)+length(ISJ_mesto)+length(ISJ_velkomesto) #337
mean(ISJ_dedina)
mean(ISJ_mesto)
mean(ISJ_velkomesto) ks.test(ISJ_dedina,"pnorm",mean=mean(ISJ_dedina),sd=sqrt(var(ISJ_dedina)))
ks.test(ISJ_mesto,"pnorm",mean=mean(ISJ_mesto),sd=sqrt(var(ISJ_mesto)))
ks.test(ISJ_velkomesto,"pnorm",mean=mean(ISJ_velkomesto),sd=sqrt(var(ISJ_velkomesto)))
#Vsetky tri p value aspon 0.15 - ok
var.test(ISJ_dedina,ISJ_mesto)
var.test(ISJ_dedina,ISJ_velkomesto)
var.test(ISJ_mesto,ISJ_velkomesto)
t.test(ISJ_dedina,ISJ_mesto, alternative="greater",var.equal=TRUE) # tesne nie je sign
t.test(ISJ_dedina,ISJ_velkomesto, alternative="greater",var.equal=TRUE) #je sig
t.test(ISJ_mesto,ISJ_velkomesto, alternative="less",var.equal=TRUE) #nie je sign

VS_zapad <- VS[obl_slov=="zapad"]
VS_stred <- VS[obl_slov=="stred"]
VS_vychod <- VS[obl_slov=="vychod"]
VS_zahranicie <- VS[grep("Zahra",obl_slov)]
length(VS_zapad)+length(VS_stred)+length(VS_vychod) +length(VS_zahranicie) #337
mean(VS_zapad)
mean(VS_stred)
mean(VS_vychod)
ks.test(VS_zapad,"pnorm",mean=mean(VS_zapad),sd=sqrt(var(VS_zapad)))
ks.test(VS_stred,"pnorm",mean=mean(VS_stred),sd=sqrt(var(VS_stred)))
ks.test(VS_vychod,"pnorm",mean=mean(VS_vychod),sd=sqrt(var(VS_vychod)))
#vsetky tri p value aspon 0.15 - ok
```

```
var.test(VS_zapad,VS_stred)
var.test(VS_zapad,VS_vychod)
var.test(VS_stred,VS_vychod)
t.test(VS_zapad,VS_stred, alternative="greater",var.equal=TRUE) # nie je sign
t.test(VS_zapad,VS_vychod, alternative="greater",var.equal=TRUE) #je sig
t.test(VS_stred,VS_vychod, alternative="greater",var.equal=TRUE) #je sig

ISJ_zapad <- ISJ[obl_slov=="zapad"]
ISJ_stred <- ISJ[obl_slov=="stred"]
ISJ_vychod <- ISJ[obl_slov=="vychod"]
ISJ_zahranicie <- ISJ[grep("Zahra",obl_slov)]
length(ISJ_zapad)+length(ISJ_stred)+length(ISJ_vychod) +length(ISJ_zahranicie) #337
mean(ISJ_zapad)
mean(ISJ_stred)
mean(ISJ_vychod)
ks.test(ISJ_zapad,"pnorm",mean=mean(ISJ_zapad),sd=sqrt(var(ISJ_zapad)))
ks.test(ISJ_stred,"pnorm",mean=mean(ISJ_stred),sd=sqrt(var(ISJ_stred)))
ks.test(ISJ_vychod,"pnorm",mean=mean(ISJ_vychod),sd=sqrt(var(ISJ_vychod)))
#vsetky tri p value aspon 0.15 - ok
var.test(ISJ_zapad,ISJ_stred)
var.test(ISJ_zapad,ISJ_vychod)
var.test(ISJ_stred,ISJ_vychod)
t.test(ISJ_zapad,ISJ_stred, alternative="greater",var.equal=TRUE) # nie je sign
t.test(ISJ_zapad,ISJ_vychod, alternative="greater",var.equal=TRUE) #je sig
t.test(ISJ_stred,ISJ_vychod, alternative="greater",var.equal=TRUE) #je sig
```