

**EKONOMICKÁ UNIVERZITA V BRATISLAVE**  
**FAKULTA HOSPODÁRSKEJ INFORMATIKY**

Evidenčné číslo: 103005/B/2022/36122176493346052

**ŠTATISTICKÁ ANALÝZA VPLYVU VYBRANÝCH**  
**FAKTOROV NA EKVIVALENTNÝ DISPONIBILNÝ**  
**PRÍJEM SLOVENSKÝCH DOMÁCNOSTÍ**

**BAKALÁRSKA PRÁCA**

**EKONOMICKÁ UNIVERZITA V BRATISLAVE**  
**FAKULTA HOSPODÁRSKEJ INFORMATIKY**

**ŠTATISTICKÁ ANALÝZA VPLYVU VYBRANÝCH**  
**FAKTOROV NA EKVIVALENTNÝ DISPONIBILNÝ**  
**PRÍJEM SLOVENSKÝCH DOMÁCNOSTÍ**

**BAKALÁRSKA PRÁCA**

Študijný program:

Manažérske rozhodovanie

Študijný odbor:

Ekonómia a manažment

Školiace pracovisko:

Katedra štatistiky

Vedúci záverečnej práce:

Ing. Martina Košíková

## **Zadanie záverečnej práce**

## Čestné vyhlásenie

Čestne vyhlasujem, že predkladanú záverečnú prácu som vypracovala samostatne, pod odborným dohľadom vedúcej práce, a že som uviedla všetku použitú literatúru.

Dátum: .....

Podpis: .....

## **Pod'akovanie**

Za cenné pripomienky, odborné rady a konštruktívnu kritiku by som rada poďakovala vedúcej mojej bakalárskej práce p. Ing. Martine Košíkovej. Ďakujem tiež p. prof. Mgr. Erikovi Šoltésovi, PhD, ktorý prispel k vypracovaniu tejto práce. Nakoniec by som rada poďakovala svojej rodine za podporu a trpezlivosť.

## **Abstrakt**

OKASOVÁ, Karin Štatistická analýza vplyvu vybraných faktorov na ekvivalentný disponibilný príjem slovenských domácností. Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky, Katedra štatistiky – Vedúci záverečnej práce: Ing. Martina Košíková – Bratislava: FHI EU, 2022, s.

Práca je rozdelená do niekoľkých kapitol:

Prvá kapitola bude zameraná na oboznámenie sa s chudobou a s jedným z ukazovateľov chudoby, ktorým je ekvivalentný disponibilný príjem. Priblížime jej aktuálny stav doma a v zahraničí a popíšeme jej hlavné ukazovatele.

Ďalšia kapitola bude definovať cieľ práce. Na to, aby sme naplnili hlavný cieľ si určíme jednotlivé čiastkové ciele.

Tretia kapitola bude zameraná na metodiku práce a metódy skúmania, kde si priblížime všeobecný lineárny model, analýzu rozptylu, otestujeme štatistickú významnosť modelu, ale aj jednotlivých parametrov a tiež sa oboznámime so softvérom SAS Enterprise Guide.

V poslednej štvrtej kapitole práce využijeme získané poznatky a vytvoríme všeobecný lineárny model ekvivalentného disponibilného príjmu osôb slovenských domácností. Z výsledných analýz popíšeme, aký vplyv mali vybrané relevantné faktory na závisle premennú EDP a kvantifikujeme hlbšie vzťahy medzi jednotlivými faktormi zaradenými do modelu s využitím príkazu LSMEANS.

### **Kľúčové slová:**

všeobecný lineárny model, ekvivalentný disponibilný príjem, LSMEANS

## **Abstract**

OKASOVÁ, Karin Statistical analysis of the influence of selected factors on the equivalent disposable income of Slovak households. University of Economics in Bratislava, Faculty of Economic Informatics, Department of Statistics - Thesis supervisor: Ing. Martina Košíková - Bratislava: FEI UE, 2022, p.

The work is divided into several chapters:

The first chapter will focus on getting to know poverty and one of the indicators of poverty, which is the equivalent disposable income. We will present its current situation at home and abroad and describe its main indicators.

The next chapter will define the goal of the work. In order to fulfill the main goal, we will determine the individual sub-goals.

The third chapter will focus on work methodology and research methods, where we will approach the general linear model, analysis of variance, test the statistical significance of the model, but also the individual parameters, and also get acquainted with the SAS Enterprise Guide software.

In the last fourth chapter we will use the acquired knowledge and create a general linear model of the equivalent disposable income of Slovak households. From the resulting analyzes, we describe the impact of selected relevant factors on the dependent variable EDP and quantify the deeper relationships between the individual factors included in the model using the LSMEANS command.

### **Key words:**

general linear model, equivalent disposable income, LSMEANS

# Obsah

Úvod.....	9
1. Súčasný stav riešenej problematiky doma a v zahraničí.....	10
1.1. Štatistické zisťovanie EU-SILC .....	11
1.2. Ukazovatele chudoby .....	12
1.3. Príjmová chudoba.....	14
1.3.1. Ekvivalentný disponibilný príjem domácností .....	16
2. Cieľ práce.....	17
3. Metodika práce a metódy skúmania .....	18
3.1. Všeobecný lineárny model – GLM .....	18
3.2. Popis prostredia a ovládanie SAS Enterprise Guide 7.1 .....	19
3.3. Test štatistickej významnosti regresného modelu.....	20
3.4. Test štatistickej významnosti vplyvu faktorov.....	22
3.5. Procedúra GLM.....	22
4. Výsledky práce .....	25
4.1. Postup aplikácie procedúry GLM v SAS .....	25
4.2. Model EDP s vybranými premennými.....	28
Záver .....	48
Zoznam použitej literatúry .....	50
Prílohy.....	54

## Úvod

V roku 2020 Štatistický úrad Slovenskej republiky vykonal zisťovanie EU-SILC, ktorého predmetom bola analýza životnej úrovne obyvateľov SR. Zisťovanie v Slovenskej republike za rok 2020 tvorilo 5 543 domácností a 13 800 osôb vo veku minimálne 16 rokov. (Vlačuha a Kováčová, 2021)

Posudzovať kvalitu životnej úrovne obyvateľov môžeme prostredníctvom rôznych indikátorov, pričom v našej práci sa zameriame na príjmovú chudobu, konkrétne na ukazovateľ ekvivalentný disponibilný príjem osôb slovenských domácností. Vstupnú databázu na posúdenie chudoby a sociálneho vylúčenia nám budú tvoriť údaje získané zo štatistického zisťovania EU-SILC<sup>1</sup> 2020. Využitím softvéru SAS Enterprise Guide analyzujeme, aký vplyv majú vybrané ukazovatele na výšku ekvivalentného disponibilného príjmu (EDP) jednotlivých osôb domácností.

Stratégia Európa 2020, ktorá bola v roku 2010 prijatá Európskou komisiou pre Európske krajiny mala stanovených 5 hlavných cieľov, ktoré mali byť splnené do roku 2020. Jedným z hlavných cieľov mal byť boj proti chudobe a sociálnemu vylúčeniu a to tak, že sa mal znížiť počet ohrozených ľudí o 20 miliónov. Keďže tento cieľ nebol do roku 2020 úspešne dosiahnutý, Európska komisia schválila akčný plán na realizáciu Európskeho piliera sociálnych vecí, v ktorom sú zahrnuté nové ciele, ktoré by mali byť splnené do roku 2030. Jedným z hlavných cieľov je zníženie počtu obyvateľov, ktorí sú ohrození chudobou alebo sociálnym vylúčením minimálne o 15 miliónov, z toho aspoň 5 miliónov by mali byť deti.

Bližšie priblíženie problematiky chudoby alebo sociálneho vylúčenia nájdeme napríklad v Eurostat (2021), Kennedy (2021), Európska komisia (2021).

---

<sup>1</sup> EU-SILC slúži na zhromaždenie a poskytnutie aktuálnych porovnateľných údajov o chudobe, príjmoch, životných podmienkach a sociálnom vylúčení na úrovni Európskej únie (EÚ).

# 1. Súčasný stav riešenej problematiky doma a v zahraničí

V dnešnej dobe si už veľa ľudí myslí, že výrazy ako „chudobní pracujúci“ alebo príjmová chudoba, by mali byť oxymoron, pretože z logického ponímania by ľudia, ktorí tvrdo pracujú, nemali byť chudobní. V súčasnosti žije v chudobe takmer polovica obyvateľov sveta, pričom najväčšie problémy spojené s chudobou boli a aj sú v rozvojových krajinách. Ľudia častokrát nemajú prístup ani k základným ľudským potrebám, chudoba sa tu prejavuje v extrémnych podobách. (Andreß a Lohmann 2008)

Väčšina definícií o chudobe je založená na príjmoch ľudí alebo domácností. Tieto definície sú dôležitejšie pre krajiny, kde sa priemerný príjem domácností stáva vyšší. Naopak v rozvojových krajinách sa stávajú menej dôležité, pretože dôležitejšie sú pre nich základné potreby ako napríklad jedlo. (Ravallion, 1998)

Aj keď samotné rozvojové krajiny nesú na svojej zaostalosti vlastný podiel viny, vplyvajú na ňu aj vyspelé krajiny sveta. Osvedčeným receptom v boji proti extrémnej biede je ekonomický rast. A aj keď hospodárstvo v týchto krajinách zaznamenáva nárast, prudký populačný rast v rozvojových krajinách marí šance na definitívne východisko z biedy. Ako povedal António Guterres, generálny tajomník Organizácie Spojených národov, rozvojové krajiny by mali investovať do sektorov na vytvorenie pracovných miest, zabezpečenie vzdelania a kvalifikovania ľudí, dostatok potravín, vybudovanie infraštruktúr a dopravných sietí, aby znížili chudobu vo svojej krajine. Taktiež uviedol, že by sa mala investovať časť svetového bohatstva do krajín, ktoré to najviac potrebujú, aby sa vyhli chudobe. (OSN, 2022)

Chudoba v Európskej únii nepredstavuje také veľké riziko ako v rozvojových krajinách. Na základe výsledkov z Eurostatu (2021) v roku 2020 bolo 21,9 % populácie (čo predstavuje približne 96,5 miliónov ľudí) ohrozených chudobou alebo sociálnym vylúčením. Jedným z najhlavnejších aspektov chudoby, ktorý najviac postihuje ľudí je nedostatočná výška príjmu.

Chudoba nie je vo všeobecnosti chápaná len na základe jednej definície, ale je možné ju posudzovať z viacerých hľadísk, ako je napríklad výška dosiahnutého vzdelania, nedostatočný zdravotný stav, nedostatok zručností, príjmov alebo materiálneho vlastníctva

a podobne. Dlhoročné analyzovanie chudoby odhalilo tri hlavné dimenzie chudoby a sociálneho vylúčenia – príjmová chudoba, materiálna deprivácia a veľmi nízka pracovná intenzita. Zjednotením uvedených troch aspektov vznikol jeden agregovaný ukazovateľ AROPE – ľudia ohrození chudobou alebo sociálnym vylúčením. V minulosti tento ukazovateľ slúžil na sledovanie stratégie cieľa EÚ 2020, teraz slúži na kontrolovanie cieľa EÚ 2030 v zameraní na chudobu. Mieru AROPE vypočítame ako podiel celkovej populácie a ľudí ohrozených aspoň jedným z týchto ukazovateľov.

Bližšie priblíženie problematiky chudoby a ukazovateľa nájdeme napríklad v The World Bank (2001), Eurostat (2021) a pod.

Keďže chudoba patrí k celosvetovým problémom a ovplyvňuje kvalitu života obyvateľov je dôležité ju eliminovať vo všetkých jej formách a dimenziách, vrátane extrémnej chudoby do roku 2030 (stratégia Európa 2030) pre udržateľný ľudský rozvoj. (OSN, 2017)

## 1.1. Štatistické zisťovanie EU-SILC

Štatistické zisťovanie o príjmoch a životných podmienkach známe pod skratkou EU-SILC (z angl. European Union Statistics on Income and Living Conditions) je *metodikou štatistiky EÚ o príjmoch a životných podmienkach*. Výsledky zisťovania EU-SILC využívame napríklad na kvantifikáciu miery AROPE, ktorá patrí medzi jeden z najdôležitejších ukazovateľov na sledovanie cieľa stratégie EÚ 2030, a to eliminácia počtu ľudí, ktorí sú ohrození chudobou alebo sociálnym vylúčením minimálne o 15 miliónov osôb.

EU-SILC je elektronickou publikáciou Eurostatu a slúži na sledovanie a analyzovanie chudoby alebo sociálneho vylúčenia. Vznikol v roku 2003 medzi Nórskom a členskými štátmi - Belgickom, Dánskom, Gréckom, Írskom, Luxemburskom a Rakúskom na základe „džentlmenskej dohody“. Na Slovensku sa začal používať až v roku 2005, kedy sa EU-SILC rozšíril o všetky členské štáty EÚ (vtedy 25 štátov). Dnes sa už zisťovanie EU-SILC vykonáva vo všetkých krajinách Európskej únie a tiež aj v krajinách ako je Švajčiarsko, Nórsko, Island a Srbsko. (Eurostat, 2018, 2021)

## 1.2. Ukazovatele chudoby

V chudobe sa ľudia dostávajú do materiálnej deprivácie z čoho vyplýva, že majú nedostatok prostriedkov, aby si zabezpečili svoju existenciu. (Žák, 2002)

Aby bolo možné posúdiť vývoj chudoby v krajinách, v rámci Európskej únie sa chudoba meria pomocou konkrétnych ukazovateľov chudoby. Hlavným ukazovateľom na meranie chudoby je AROPE, ktorý pozostáva z troch hlavných čiastkových ukazovateľov:

- **miera závažnej materiálnej deprivácie**, ktorá vyjadruje podiel ľudí, ktorí si nemôžu dovoliť využívať služby alebo zaobstarať niektoré spotrebné tovary, prípadne predmety dlhodobej spotreby, ktoré sa bežne v spoločnosti používajú. Materiálna deprivácia je tiež jedným z ukazovateľov EU-SILC;
- **miera rizika chudoby**, ktorá slúži na meranie nízkeho príjmu ľudí v porovnaní s ostatnými ľuďmi v krajine. Vypočítame ho ako podiel ľudí s ekvivalentným disponibilným príjmom pod hranicou rizika chudoby<sup>2</sup>. Tento ukazovateľ sa nazýva aj príjmová chudoba. Ekvivalentný disponibilný príjem podrobnejšie objasníme v nasledujúcej podkapitole 1.3. Ľudia, ktorí sa nachádzajú pod hranicou rizika chudoby sa označujú ako osoby ohrozené chudobou. Riziko chudoby udávame v percentách;
- **miera veľmi nízkej intenzity práce**, ktorá predstavuje podiel počtu ľudí, ktorí žijú v takej domácnosti, v ktorej bol v rámci uplynulého roka využitý pracovný potenciál, resp. intenzita práce na menej ako 20 %. Pracovný potenciál chápeme ako celkové množstvo času v rámci uplynulého referenčného obdobia, ktoré mohli osoby vo veku 18 – 59 rokov odpracovať (v cieľoch stratégie Európa 2030 bola modifikovaná hranica produktívneho veku na 18 – 64 rokov).

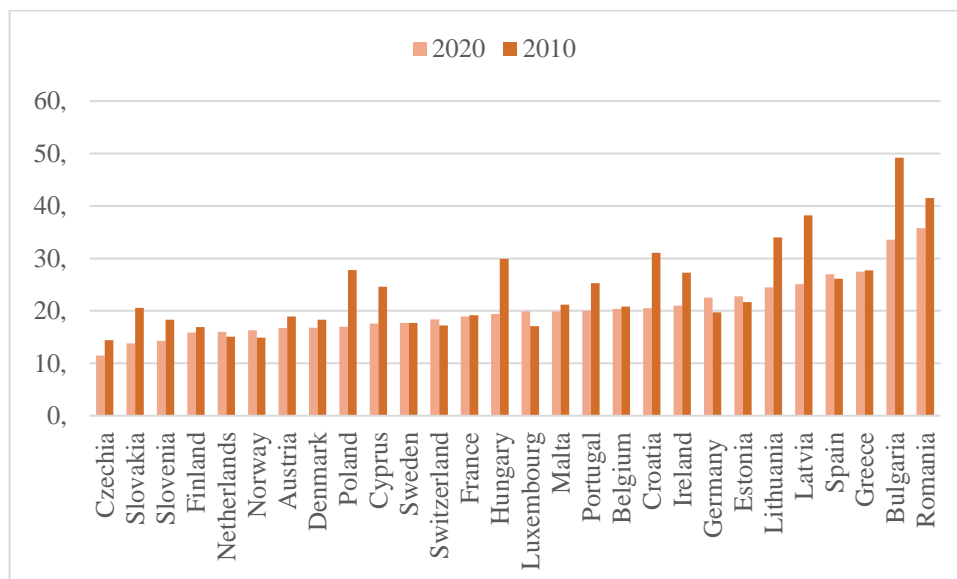
---

<sup>2</sup> Hranica rizika chudoby určená po sociálnych odvodoch na úrovni 60 % národného mediánu.

Ľudia, ktorých ovplyvňuje aspoň jeden z týchto troch ukazovateľov sa označujú ako ľudia ohrození chudobou alebo sociálnym vylúčením. Najohrozenejší sú však tí ľudia, ktorí sú ovplyvnení nielen jedným z uvedených ukazovateľov, ale súčasne dvomi alebo tromi aspektmi chudoby.

Bližšie priblíženie problematiky ukazovateľov chudoby nájdeme napríklad v OECD (2007) alebo Európska komisia (2017).

Na obrázku č. 1.1 môžeme vidieť počet ľudí v % ohrozených chudobou alebo sociálnym vylúčením za roky 2010 a 2020 v krajinách Európskej únie.



Obrázok 1.1 Ľudia ohrození chudobou alebo sociálnym vylúčením v Európskej únii v percentách za rok 2010 a 2020

Zdroj: EUROSTAT, vlastné spracovanie

Krajiny, ako sú Rumunsko (35,8 %), Bulharsko (33,6 %), Grécko (27,5 %), Španielsko (27,0 %) patrili v roku 2020 ku najviac ohrozeným krajinám Európskej únie aspoň jedným z indikátorov chudoby. Naopak v roku 2020 krajiny, ako Česká republika (11,5 %), Slovenská republika (13,8 %), Slovinsko (14,3 %), Fínsko (15,9 %), ale aj Holandsko (16,0 %) patrili ku krajinám s najnižším podielom ohrozenia chudobou.

Keďže v práci sa zameriavame predovšetkým na Slovenskú republiku, v tabuľke 1.1 môžeme vidieť ako sú ľudia ohrození chudobou alebo sociálnym vylúčením v jednotlivých krajoch.

Tabuľka 1.1 Ľudia ohrození chudobou alebo sociálnym vylúčením v jednotlivých krajoch  
SR v % za rok 2020

<b>Kraj</b>	<b>Ľudia ohrození chudobou alebo sociálnym vylúčením (v %)</b>
Bratislavský kraj	5,9
Trnavský kraj	10,8
Trenčiansky kraj	10,7
Nitriansky kraj	9,3
Žilinský kraj	14,3
Banskobystrický kraj	24,3
Prešovský kraj	20,8
Košický kraj	19,4
SR	14,8

Zdroj: Eurostat, vlastné spracovanie

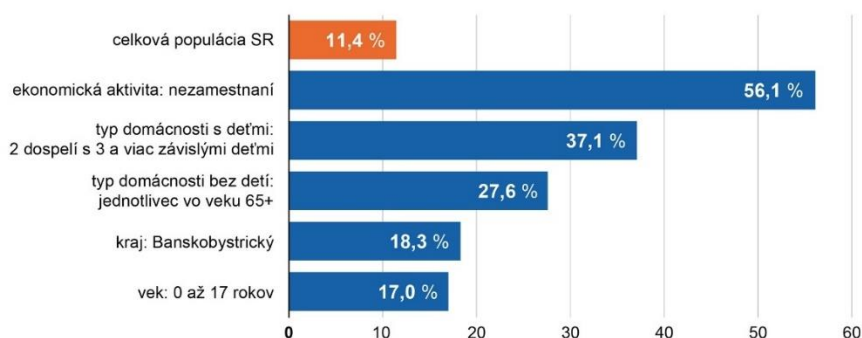
Na základe tabuľky 1.1 môžeme tvrdiť, že najmenej ľudí ohrozených chudobou alebo sociálnym vylúčením žije v Bratislavskom kraji (5,9 %), naopak najviac ohrozených ľudí žije v Banskobystrickom kraji (24,3 %). (Eurostat, 2021)

### **1.3. Príjmová chudoba**

Jednou z najpresnejších definícií o príjmovej chudobe je, že predstavuje stav, kedy jednotlivец alebo skupina ľudí nemá dostatok finančných prostriedkov na uspokojenie základných životných potrieb. (Gajdoš, 2002)

Približne 11,4 % obyvateľov, čo predstavovalo okolo 615 000 ľudí bolo v roku 2020 v Slovenskej republike ohrozených chudobou, čo môžeme vidieť aj na obrázku 1.2. Medzi najviac ohrozené skupiny osôb patrili nezamestnaní, ktorí tvorili až 56,1 %, ďalšou ohrozenou skupinou boli domácnosti s dvomi dospelými s aspoň tromi závislými deťmi (37,1 %). Medzi ohrozené skupiny patria aj neúplné domácnosti s jedným, alebo viac závislými deťmi (33,9 %) a tiež aj domácnosti dvoch dospelých, z ktorých aspoň jeden je vo veku 65 rokov a viac (27,6 %).

**Príjmová chudoba 2020\* – najohrozenejšie skupiny obyvateľstva**  
(podiel v % z danej skupiny obyvateľov)



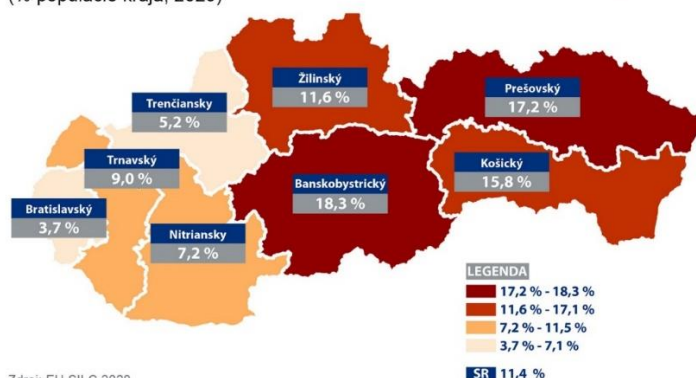
\*Miera príjmovej chudoby – prvá z troch dimenzií chudoby a sociálneho vylúčenia dotýkajúca sa najväčšieho počtu obyvateľov.  
Zdroj: EU SILC 2020

Obrázok 1.2 Príjmová chudoba najohrozenejších skupín obyvateľstva na Slovensku v roku 2020

Zdroj: ŠÚSR, 2021

Ako môžeme vidieť podľa obrázku č. 1.3 najohrozenejším krajom SR z hľadiska príjmovej chudoby bol v roku 2020 Banskobystrický kraj (18,3 %).

**Podiel obyvateľov ohrozených príjmovou chudobou v regiónoch SR**  
(% populácie kraja, 2020)



Zdroj: EU SILC 2020

Obrázok 1.3 Podiel obyvateľov SR ohrozených príjmovou chudobou v jednotlivých regiónoch

Zdroj: ŠÚSR, 2021

Druhým najohrozenejším krajom je na základe výsledkov z obrázku č. 1.3 Prešovský kraj (17,2 %). Miera rizika chudoby sa nachádza nad priemerom Slovenskej republiky (11,4 %) aj v Žilinskom a v Košickom kraji. Najmenej ohrozeným krajom je Bratislavský

kraj (3,7 %). Medzi menej ohrozené kraje patria aj ďalšie tri kraje - Trenčiansky (5,2 %), Nitriansky (7,2 %) a Trnavský kraj (9,0 %). Tieto odlišnosti v príjmovej chudobe vo všetkých krajoch spôsobuje predovšetkým nerovnomerné rozmiestnenie pracovných príležitostí na území Slovenskej republiky. (Štatistický úrad Slovenskej republiky, 2021)

### *1.3.1. Ekvivalentný disponibilný príjem domácností*

Ekvivalentný disponibilný príjem je jedným z aspektov merania chudoby, ktorý patrí pod agregovaný ukazovateľ AROPE. V štatistickom zisťovaní EU-SILC sa využíva ukazovateľ ekvivalentný disponibilný príjem domácností na meranie príjmovej chudoby. Vypočítame ho tak, že disponibilný príjem<sup>3</sup> domácnosti vydělíme počtom osôb, ktoré žijú spoločne v jednej domácnosti. Následne podľa modifikovanej škály OECD<sup>4</sup> pripíšeme každému členovi domácnosti určitú váhu:

- prvej osobe v domácnosti priradíme váhu 1,0,
- každej nasledujúcej osobe, ktorá má viac ako 14 rokov priradíme váhu 0,5 a
- ostatným osobám, ktoré sú mladšie ako 14 rokov priradíme váhu 0,3.

Konečný výsledok predstavuje ekvivalentný disponibilný príjem domácností.

Napríklad môžeme uvažovať s domácnosťou, ktorá má celkový disponibilný príjem 7 500 EUR a skladá sa z dvoch dospelých osôb, jedným dieťaťom vo veku viac ako 14 rokov a jedným dieťaťom vo veku menej ako 14 rokov. Následne priradíme jednotlivým členom domácnosti váhy:  $1,0 + 0,5 + 0,5 + 0,3 = 2,3$ . Posledným krokom je určenie ekvivalentného disponibilného príjmu domácnosti, ktorý získame podielom  $7\,500 / 2,3$ . Čiže ekvivalentný disponibilný príjem domácnosti dvoch dospelých osôb s dieťaťom vo veku viacej ako 14 rokov a dieťaťom vo veku menej ako 14 rokov je približne 3260,87 EUR.

Bližšie priblíženie problematiky ekvivalentného disponibilného príjmu nájdeme napríklad v Eurostat (2021), UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE (2011).

---

<sup>3</sup> Disponibilný príjem domácností sa skladá z hrubého príjmu všetkých členov domácností, od ktorého odpočítame pravidelné dane.

<sup>4</sup> OECD – Organisation for Economic Co-operation and Development, je medzinárodná organizácia pre hospodársku spoluprácu a rozvoj. Zameriava sa na tvorenie lepších politík na zlepšenie života.

## 2. Cieľ práce

Hlavným cieľom tejto práce je analyzovať vplyv vybraných faktorov na ekvivalentný disponibilný príjem osôb jednotlivých domácností Slovenskej republiky za rok 2020. Údaje potrebné k analýze sme získali z EU-SILC. Aby sme zabezpečili splnenie nášho hlavného cieľa, stanovili sme si čiastkové ciele:

- v rámci teoretickej časti práce sa oboznámiť so súčasným stavom riešenej problematiky na Slovensku a v krajinách Európskej únie,
- objasniť podstatu štatistického zisťovania EU-SILC,
- charakterizovať kľúčový ukazovateľ našej analýzy ekvivalentný disponibilný príjem domácností a zároveň aj ďalšie ukazovatele chudoby, ako sú materiálna deprivácia a veľmi nízka intenzita práce,
- v metodologickej časti popísať prostredie štatistického softvéru SAS Enterprise Guide 7.1, ktoré využívame pri práci na získanie výstupov a následne na ich interpretáciu,
- pochopiť a vysvetliť podstatu všeobecných lineárnych modelov a oboznámiť sa s procedúrou GLM a s príkazom LSMEANS,
- v rámci praktickej časti konštruovať všeobecný lineárny model vyjadrujúci vplyv relevantných faktorov na vysvetľovanú premennú EDP,
- overiť štatistickú významnosť odhadnutého modelu a faktorov, ktoré vstupovali do analýzy,
- použitím príkazu LSMEANS identifikovať hlbšie vzťahy medzi vybranými faktormi z pohľadu cieľovej premennej EDP,
- na záver zhodnotiť výsledky štatistických analýz a naplnenie vopred stanovených cieľov.

### 3. Metodika práce a metody skúmania

Práca využíva dáta zo zisťovania EU-SILC, pričom základom bude analýza pomocou všeobecných lineárnych modelov. V tejto kapitole sa budeme venovať oboznámeniu a priblíženiu metód a postupov, ktoré nám budú slúžiť na uvedenú analýzu.

Pri praktickej časti bakalárskej práce sme využili štatistický analytický softvér SAS, konkrétne jeho používateľské rozhranie SAS Enterprise Guide 7.1 (ďalej len „SAS EG“).

#### 3.1. Všeobecný lineárny model – GLM

Všeobecný lineárny model, z angl. General Linear Model (ďalej len „GLM“), sa používa pre lineárne modely s ľubovoľnou kombináciou číselnej vysvetľovanej premennej v závislosti od jednej alebo viacerých kvalitatívnych alebo kvantitatívnych vysvetľujúcich premenných. GLM sú základom pre rôzne štatistické metódy, z ktorých najčastejšie využívame analýzu rozptylu (ANOVA), analýzu kovariancie (ANCOVA) a lineárnu regresiu.

Všeobecný lineárny model vieme vyjadriť ako viacnásobný regresný model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_k$$

ktorý nám zobrazuje vzťah medzi závislou premennou –  $y$  a  $k$  nezávislými premennými  $x_j$ .

Kde:

$y_i$  - závisle premenná (vysvetľovaná premenná),

$\beta_0$  – lokujúca konštanta, je to hodnota  $y_i$ , ak hodnota  $x_{ik}$  sa rovná 0,

$\beta_k$  – množina regresných koeficientov,

$x_{ik}$  – množina  $k$  - nezávisle premenných.

Bližšie priblíženie problematiky všeobecného lineárneho modelu nájdeme napríklad v publikáciách: Dobson (2002), Šoltés (2019), Pacáková a kol. (2009).

Všeobecný lineárny model vieme taktiež vyjadriť aj ako model v dvojfaktorovom tvare:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

kde pre každé pozorovanie  $y_{ijk}$  predpokladáme, že:

$\mu$  bude stredná hodnota, parameter  $\alpha_i$  predstavuje účinok faktora A na  $i$  – tej úrovni,  $\beta_j$  je parameter, ktorý predstavuje účinok faktora B na  $j$  – tej úrovni. Parameter  $(\alpha\beta)_{ij}$  tvorí interakciu faktora A na  $i$  – tej úrovni a faktora B na  $j$  – tej úrovni. Posledný parameter  $\varepsilon_{ijk}$  charakterizuje náhodnú chybu. (Dodge, 2008)

### **3.2. Popis prostredia a ovládanie SAS Enterprise Guide 7.1**

SAS (*Statistical Analytical System*) je štatistický analytický softvér ktorého súčasťou je aj prostredie SAS Enterprise Guide. SAS EG je program typu „point-and-click“, ktorý nám umožňuje analyzovať údaje a zobrazovať ich výsledky aj bez písania príkazov (pri náročnejších analýzach je niekedy nutné príkazy doplniť). Základné grafické zobrazenie SAS EG je rozdelené do troch hlavných oblastí – Workspace, Project Tree a Server List.

Okno Workspace, ktoré predstavuje najväčšie okno v rámci tohto softvéru, slúži na vlastnú prácu s dátami, ale taktiež sa tu dajú písať príkazy. Project Tree je okno, ktoré využívame na organizáciu dát a výsledkov, taktiež môžeme pomocou neho otvárať projekty a pracovať s nimi. V stromovej štruktúre sa ukladajú vstupné a aj výstupné dáta a rovnako aj výsledky analýz. Server List je posledné okno, v ktorom vieme nájsť napríklad knižnicu, pomocou ktorej sa vieme dostať ku vstupným údajom. (Delwiche a Slaughter, 2003)

### 3.3. Test štatistickej významnosti regresného modelu

Na overenie, či model je štatisticky významný alebo nie je štatisticky významný používame nulovú hypotézu:

$H_0$ : regresný model nie je štatisticky významný,

oproti alternatívnej hypotéze:

$H_1$ : regresný model je štatisticky významný.

V tabuľke 3.1 sú uvedené súčty štvorcov odchýlok, ich stupne voľnosti, priemerné štvorce odchýlok a tiež aj testovacia F-štatistika.

Tabuľka 3.1 Analýza rozptylu pre lineárny regresný model

Variabilita vysvetľovanej premennej	Súčet štvorcov odchýlok SS	Stupne voľnosti DF	Priemerný štvorec MS	Testovacia štatistika F
Variabilita vysvetlená regresným modelom <b>M</b>	$SSM = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	k	$MSM = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{k}$	$F = \frac{MSM}{MSR}$
Variabilita nevysvetlená regresným modelom <b>R</b>	$SSR = \sum_{i=1}^n (y_i - \hat{y}_i)^2$	n - k - 1	$MSR = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - k - 1}$	
Celková variabilita <b>T</b>	$SST = \sum_{i=1}^n (y_i - \bar{y})^2$	n - 1		

Zdroj: vlastné spracovanie

Odchýlkou  $i$ -tej odhadnutej hodnoty závisle premennej od jej priemernej hodnoty je rozdiel  $(\hat{y}_i - \bar{y})^2$ . Výraz  $(y_i - \hat{y}_i)^2$  je odchýlkou jednotlivých hodnôt závisle premennej od jej odhadovanej (vyrovnanej) hodnoty. Rozdielom  $(y_i - \bar{y})^2$  dostaneme odchýlku  $i$ -tej hodnoty závisle premennej od jej priemernej hodnoty.

O výsledku predpokladaných tvrdení rozhodujeme na základe testovacej  $F$ -štatistiky

$$F = \frac{MSM}{MSR} = \frac{\frac{SSM}{k}}{\frac{SSR}{n-k-1}} = \frac{(n-k-1) \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{k \cdot \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

ktorá má Fisherovo rozdelenie so stupňami voľnosti  $k$  a  $(n - k - 1)$ , kde  $n$  je rozsah pozorovaní a  $k$  je počet prvkov modelu. Ak hodnota kritickej oblasti  $F_{1-\alpha}(k; n - k - 1)$  bude väčšia ako hodnota testovacej  $F$ -štatistiky, na hladine významnosti môžeme prijať nulovú hypotézu. V prípade, že sme nezamietli nulovú hypotézu, tak sme dospeli k záveru, že model je štatisticky nevýznamný a nemá zmysel s ním v ďalších analýzach uvažovať. Ak by nastal prípad, že hodnota kritickej oblasti  $F_{1-\alpha}(k; n - k - 1)$  bude menšia oproti hodnote testovacej  $F$ -štatistiky, tak nulovú hypotézu zamietneme, čo znamená, že model môžeme považovať za štatisticky významný, pretože aspoň jedna z nezávisle premenných bude mať signifikantný vplyv na závisle premennú v modeli.

Bližšie priblíženie problematiky analýzy rozptylu nájdeme napríklad v publikáciách: Šoltés (2019), Pacáková a kol. (2009).

Pomocou koeficientu determinácie vieme posúdiť kvalitu<sup>5</sup> odhadnutého modelu a taktiež vieme kvantifikovať percento variability vysvetlenej modelom. Koeficient determinácie vypočítame ako podiel variability vysvetľovanej premennej, ktorá je vysvetlená modelom oproti celkovej variabilite a musí nadobúdať hodnoty z uzavretého intervalu 0 až 1:

$$r_{xy}^2 = \frac{SSM}{SST} \quad r_{xy}^2 \in \langle 0; 1 \rangle$$

Ako aj interpretácia veľkosti koeficientu determinácie, tak aj jeho skutočný odhad sú čiastočne závislé od teórie. (Ozer, 1985)

---

<sup>5</sup> Vtedy bude model kvalitnejší, čím jeho podiel bude mať vyššiu hodnotu (bližšiu k 1).

### 3.4. Test štatistickej významnosti vplyvu faktorov

Ak sme pri testovaní štatistickej významnosti modelu ako celku zamietli nulovú hypotézu, čiže sme zistili, že model je štatisticky významný, pokračujeme testovaním štatistickej významnosti vplyvu nezávisle premenných na závisle premennú. Definujeme nulovú hypotézu:

*H<sub>0</sub>: vplyv k-tej nezávisle premennej nie je štatisticky významný,*

oproti alternatívnej hypotéze:

*H<sub>1</sub>: vplyv k-tej nezávisle premennej je štatisticky významný.*

V ďalšom kroku overíme platnosť nulovej hypotézy, pomocou testovacej štatistiky, ktorá bude mať Fisherovo rozdelenie pravdepodobnosti:

$$F = \frac{\frac{\Delta var}{1}}{\frac{SSR^{(1)}}{n - k - 1}} = \frac{\Delta var}{MSR^{(1)}}$$

a kritickú hodnotu  $F_{1-\alpha}(1; n - k - 1)$ . Výraz  $\Delta var$ <sup>6</sup> označujeme ako prínos  $k$ -tej nezávislej premennej. Ak hodnota testovacej štatistiky bude mať väčšiu hodnotu v porovnaní s kritickou hodnotou, tak zamietneme nulovú hypotézu, z čoho vyplýva, že nezávisle premenná štatisticky významne ovplyvňuje závisle premennú. Ak by sme prijali nulovú hypotézu, tak nezávisle premennú by sme následne vylúčili z modelu. (Šoltés, 2019)

### 3.5. Procedúra GLM

Na analýzu rozptylu má softvér SAS k dispozícii niekoľko postupov – procedúr. Jednou z nich je procedúra GLM, ktorou sa budeme v práci zaoberať. Okrem analýzy

---

<sup>6</sup>  $\Delta var$  vypočítame  $SSM^{(1)} - SSM^{(2)}$ , kde  $SSM^{(1)}$  je premenlivosť  $y$  vysvetlená úplným modelom a  $SSM^{(2)}$  je premenlivosť  $y$ , ktorá je vysvetlená redukovaným modelom. Redukovaný model dostaneme tak, že vynecháme nezávislú premennú (ktorej vplyv testujeme na závislú premennú) z úplného modelu. (Šoltés, 2019)

rozptylu môžeme procedúru GLM použiť aj na iné analýzy, a to napríklad na multivariačnú analýzu rozptylu, alebo na analýzu kovariancie. Na to, aby sme získali všeobecné lineárne modely a ich parametre, je v procedúre GLM využívaná metóda najmenších štvorcov.

Procedúra GLM obsahu rôzne príkazy:

```
PROC GLM <options>;  
CLASS variable <(REF= option)> ...<variable <(REF= option)>> </ global-  
options>;  
MODEL dependent-variables = independent-effects </ options>;  
ABSORB variables;  
BY variables;  
CODE <options>;  
FREQ variable;  
ID variables;  
WEIGHT variable;  
CONTRAST 'label' effect values <...effect values> </ options>;  
ESTIMATE 'label' effect values <...effect values> </ options>;  
LSMEANS effects </ options>;  
MANOVA <test-options> </ detail-options>;  
MEANS effects </ options>;  
OUTPUT <OUT=SAS-data-set> keyword=names <...keyword=names> </  
option> ;  
RANDOM effects </ options>;  
REPEATED factor-specification </ options>;  
STORE <OUT=>item-store-name </ LABEL='label'>;  
TEST <H=effects> E=effect </ options>;
```

Príkaz CLASS sa využíva na označenie klasifikačných premenných, ktoré budú použité v modeli. Ďalším príkazom je príkaz MODEL, ktorý slúži na definovanie závisle a nezávisle premenných. Príkaz ABSORB sa v modeli používa na zníženie klasifikačných účinkov, pomocou príkazu BY a zadefinovaním premenných získame samostatné analýzy v skupinách. Na výpočet predpokladaných hodnôt prispôbeného modelu využívame príkaz CODE. Príkaz FREQ využívame na pomenovanie frekvenčnej premennej, ID je

príkaz na identifikovanie pozorovanej hodnoty. Ďalší príkaz WEIGHT slúži na stanovenie premennej, ktorá bude väčšia ako nula a pomocou nej sa pridelia váhy pozorovaniam. Príkaz MANOVA je určený na viacrozmernú analýzu rozptylu v tom prípade, že model obsahuje viac ako jednu závisle premennú. V príkaze MEANS získame porovnanie stredných hodnôt. Príkaz OUTPUT umožňuje vytvoriť nový súbor údajov. Pomocou špecifikácie niektorých náhodných účinkov modelu, získame očakávané stredné hodnoty, ak použijeme príkaz RANDOM. REPEATED je príkaz, ktorý slúži na jednorozmerné a viacrozmerné testy. Predposledný príkaz STORE zabezpečuje, aby sa správne uložil postup a výsledky štatistickej analýzy. Posledný z príkazov je príkaz TEST, prostredníctvom ktorého môžeme vytvoriť dodatočné testy, ktoré využívajú iné efekty.

Na posúdenie zhody medzi marginálnymi strednými hodnotami dvoch kategórií konkrétnej nezávisle premennej v procedúre GLM sa využíva príkaz LSMEANS.

**LSMEANS** effects </ options>;

Aby sme mohli posúdiť zhodu marginálnych stredných hodnôt, musíme pracovať s nezávisle premennými, ktoré sú kategoriálne a majú viac ako dve obmeny. Vďaka príkazu LSMEANS budeme vedieť či je významný, alebo nie je významný rozdiel medzi dvojicou úrovní kategoriálnej nezávisle premennej stredných hodnôt cieľovej premennej. Pomocou príkazu LSMEANS je možné posúdiť zhodu marginálnych stredných hodnôt len medzi dvojicou kategórií, pričom pri analýze viac ako dvoch úrovní faktora, môžeme použiť príkaz CONTRAST. Príkaz ESTIMATE umožňuje odhad strednej hodnoty cieľovej premennej súhrnne za viacej kategórii nezávisle premennej. (SAS Support, 2020)

## 4. Výsledky práce

Práca je zameraná na analýzu vplyvu relevantných faktorov na ekvivalentný disponibilný príjem jednotlivých osôb domácností. Vybrané faktory, s ktorými pracujeme sú získané zo štatistického zisťovania EU-SILC 2020 a na ich analýzu využívame štatisticko-analytický systém SAS Enterprise Guide 7.1.

Aby sme mohli analyzovať vplyv premenných, najprv sme si museli určiť, ktoré nezávisle premenné budeme skúmať:

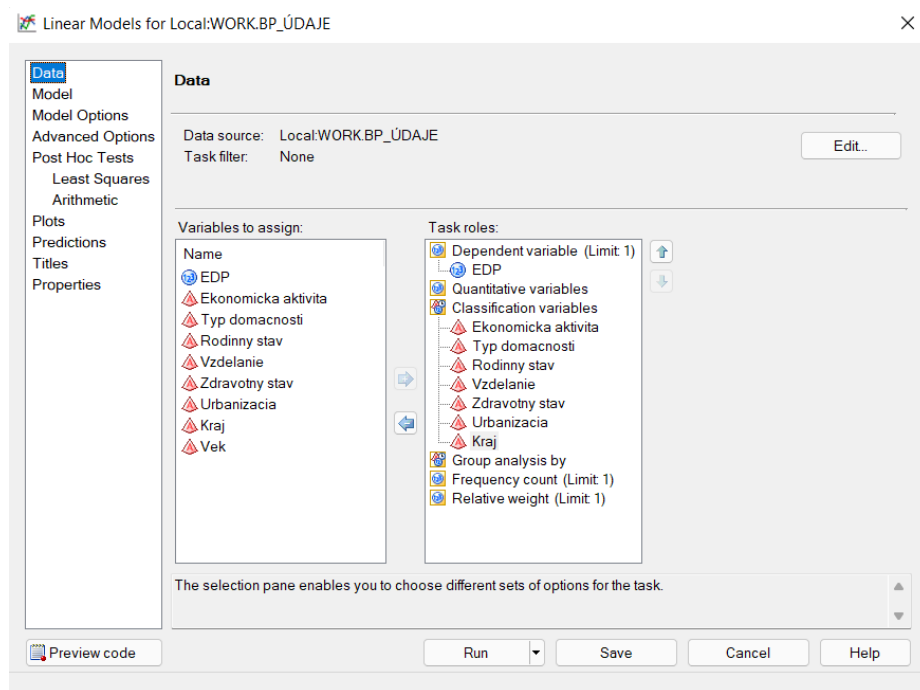
- ekonomickú aktivitu,
- typ domácností,
- rodinný stav,
- vzdelanie,
- zdravotný stav,
- stupeň urbanizácie,
- kraj,
- vek.

V tabuľke 4.1, ktorú nájdeme v prílohách sú uvedené údaje o premenných (a to ich názov, označenie a taktiež aj kategórie jednotlivých premenných), ktoré budeme analyzovať. Všetky nezávisle premenné sú kategoriálne, pričom pri každej premennej sme museli špecifikovať referenčnú kategóriu, ktorá bude v rámci analýz slúžiť ako porovnávací kategória. Referenčné kategórie sme vybrali podľa logického uváženia, pričom ich zastúpenie medzi ostatnými kategóriami premennej je dostatočne početné na to, aby mohli slúžiť ako porovnávač.

### 4.1. Postup aplikácie procedúry GLM v SAS

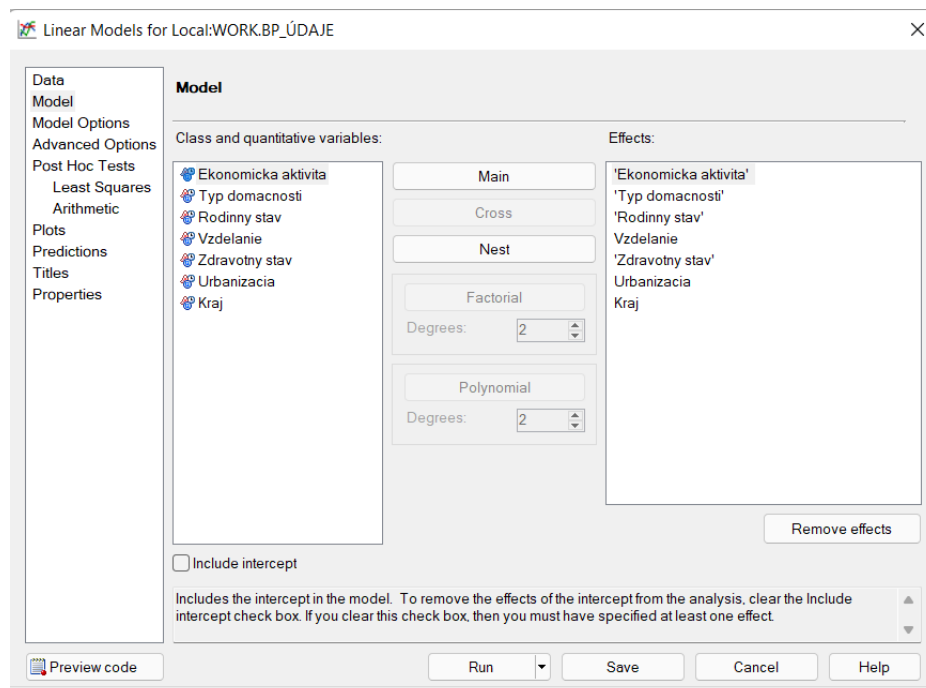
V tejto kapitole priblížime jednotlivé kroky aplikácie procedúry GLM v SAS EG 7.1. Postup je nasledovný: ako prvé nahráme do softvéru pripravené dáta, s ktorými budeme pracovať. Potom pokračujeme podľa krokov: Tasks → ANOVA → Linear Models. Následne sa dostaneme do okna Linear Models, kde v záložke Data vložíme závisle

premennú do *Dependent variable* a nezávislé premenné vložíme do *Classification variables*, ako môžeme vidieť na obrázku 4.1.



Obrázok 4.1 Určenie vstupných premenných do príslušných kategórií

Zdroj: vlastné spracovanie v SAS Enterprise Guide

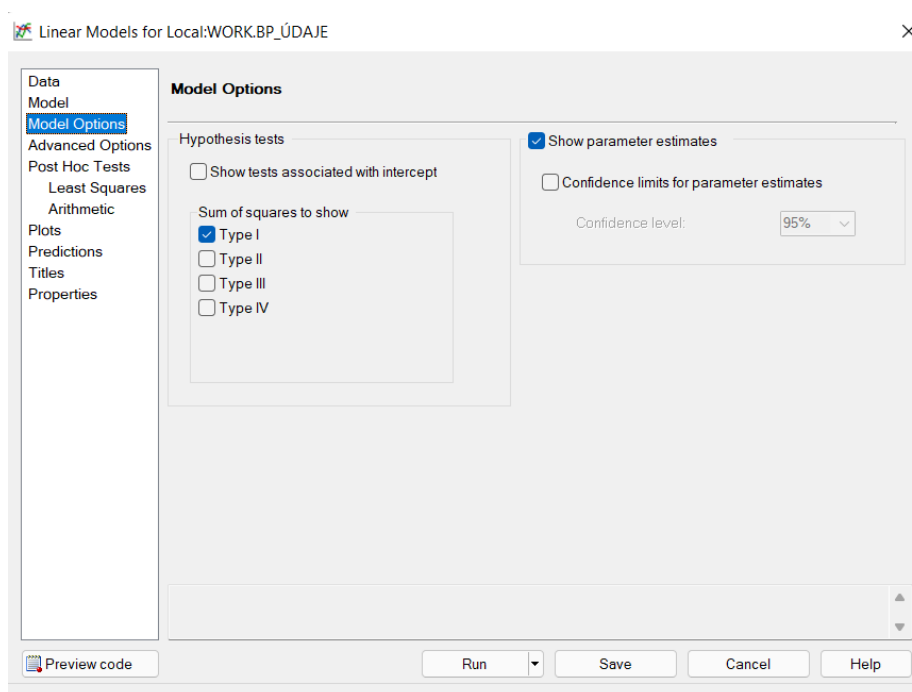


Obrázok 4.2 Určenie vplyvu jednotlivých premenných

Zdroj: vlastné spracovanie v SAS Enterprise Guide

V ďalšom kroku sa dostaneme do záložky Model na obrázku 4.2, kde určíme pomocou možnosti *Main* vplyv nezávisle premenných na závisle premennú (určíme buď vplyv každého z faktorov na závisle premennú zvlášť, alebo môžeme zvoliť interakčný vplyv niekoľkých faktorov na závisle premennú súčasne).

Na základe obrázku 4.3 v predposlednej záložke Model Options si vyberieme možnosť zobrazenia súčtu štvorcov Type I<sup>7</sup>, vďaka ktorému budeme môcť testovať hypotézy o štatistickej významnosti jednotlivých nezávisle premenných (v ponuke sú aj ďalšie varianty súčtu štvorcov, ktoré sa používajú napríklad pri nevyváženom súbore).



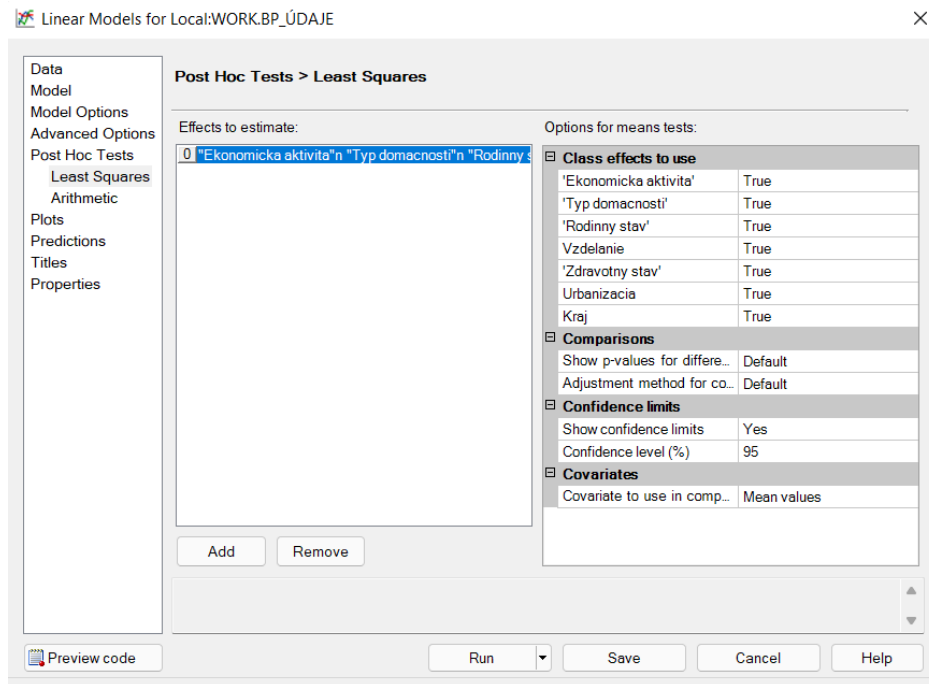
Obrázok 4.3 Zadefinovanie zobrazenia súčtu štvorcov

Zdroj: vlastné spracovanie v SAS Enterprise Guide

V poslednej záložke Post Hoc Tests, ktorú vidíme na obrázku 4.4 zvolíme, pre ktoré faktory chceme realizovať viacnásobné párové porovnávanie kategórií jednotlivých nezávisle premenných, čím odhalíme hlbšie vzťahy medzi úrovňami faktorov. Posledným krokom je spustenie tlačidla Run, vďaka ktorému získame výstup.

---

<sup>7</sup> Type I – využívame ak pracujeme s vyváženými údajmi. Ak by sme pri práci využívali nevyvážené údaje, tak použijeme Type II alebo Type III.



Obrázok 4.4 Určenie faktorov na realizáciu viacnásobného párového porovnávania

Zdroj: vlastné spracovanie v SAS Enterprise Guide

Bližšie priblíženie problematiky aplikácie procedúry GLM v SAS nájdeme napríklad v SAS SUPPORT (2013), SAS SUPPORT (2021).

## 4.2. Model EDP s vybranými premennými

Pomocou tabuľky 4.2 overíme štatistickú významnosť modelu testovaním nulovej hypotézy:

$H_0$ : model EDP domácností nie je štatisticky významný,

oproti alternatívnej hypotéze:

$H_1$ : model EDP domácností je štatisticky významný.

Tabuľka 4.2 Otestovanie štatistickej významnosti modelu pomocou analýzy rozptylu

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
<b>Model</b>	28	50957429369	1819908191.8	233.57	<.0001
<b>Error</b>	11985	93383964193	7791736.687		
<b>Corrected Total</b>	12013	144341393563			

R-Square	Coeff Var	Root MSE	EDP Mean
0.353034	31.50217	2791.368	8860.877

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Štatistickú významnosť modelu EDP domácností posúdime pomocou výslednej  $p$  – hodnoty<sup>8</sup>, pričom platí, že ak  $p$  – hodnota prekročí stanovenú hladinu významnosti, zamietneme nulovú hypotézu. Z tabuľky 4.2 vidíme, že  $p$  – hodnota je menšia ako bežne používaná hladina významnosti 0,05, čo znamená, že zamietame nulovú hypotézu a model EDP osôb jednotlivých domácností môžeme považovať za štatisticky významný.

Stĺpec DF slúži na zobrazenie vypočítaných stupňov voľnosti, kde môžeme vidieť, že celkový rozsah výberového súboru je  $n = 12\,014$  ( $n - 1$ ).

Na základe tabuľky 4.2 vidíme, že koeficient determinácie má hodnotu 0,359860. Napriek tomu, že hodnota koeficientu determinácie sa neblíži k 1 neznamená, že model nie je kvalitný. V práci pracujeme s veľkým množstvom údajov a len s vybranými faktormi, pričom existuje množstvo ďalších merateľných alebo nemerateľných faktorov, ktoré by taktiež mohli mať vplyv na ekvivalentný disponibilný príjem. Z toho dôvodu je hodnota koeficientu determinácie nízka. K zvýšeniu hodnoty koeficienta determinácie by mohli napomôcť aj iné faktory, ktoré by mohli byť zaradené do modelu, ale momentálne sme s nimi v modeli neuvažovali.

Aby sme zistili, aký vplyv majú vybrané faktory na závislú premennú, musíme otestovať ich štatistickú významnosť v tabuľke 4.2 prostredníctvom nulovej hypotézy

<sup>8</sup>  $p$  – hodnota ( $p$  – value) patrí medzi najčastejšie používaný nástroj pri štatistickej analýze. Používame ju na rozhodnutie, či zamietame nulovú hypotézu na zvolenej hladine významnosti (zvyčajne  $\alpha = 0,05$ ).

$H_0$ : vplyv vybraného faktora na závisle premennú nie je štatisticky významný,

oproti alternatívnej hypotéze

$H_1$ : vplyv vybraného faktora na závisle premennú je štatisticky významný.

Tabuľka 4.3 Testovanie vplyvu štatistickej významnosti vybraných faktorov na závislú premennú

Source	DF	Type I SS	Mean Square	F Value	Pr > F
<b>Ekonomická aktivita</b>	3	22612078527	7537359509	967.35	<.0001
<b>Vzdelanie</b>	4	7364450974	1841112744	236.29	<.0001
<b>Typ domácnosti</b>	8	14388382016	1798547752	230.83	<.0001
<b>Urbanizácia</b>	2	2332218118	1166109059	149.66	<.0001
<b>Zdravotný stav</b>	2	1624821038	812410519	104.27	<.0001
<b>Kraj</b>	7	2452713412	350387630	44.97	<.0001
<b>Rodinný stav</b>	2	182765284	91382642	11.73	<.0001

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Aby sme mohli zistiť, ktorá premenná má najväčší vplyv na ekvivalentný disponibilný príjem domácností, tak sme v tabuľke 4.3 zoradili nezávislé premenné podľa výslednej hodnoty testovacej  $F$ -štatistiky<sup>9</sup> od ich najväčšieho vplyvu po najmenší. Môžeme tvrdiť, že na závisle premennú – ekvivalentný disponibilný príjem má najväčší vplyv ekonomická aktivita ( $F = 967,35$ ), potom nasleduje vzdelanie ( $F = 236,29$ ), typ domácnosti ( $230,83$ ), urbanizácia ( $F = 149,66$ ), zdravotný stav ( $F = 104,27$ ), kraj ( $F = 44,97$ ) a posledná nezávisle premenná, ktorá má najnižší vplyv je rodinný stav ( $F = 11,73$ ).

Porovnaním  $p$  – hodnôt s hladinou významnosti  $\alpha = 0,05$  môže tvrdiť, že všetky nezávislé premenné, ktoré sme zaviedli do modelu sú štatisticky významné. To znamená, že uvedené nezávislé premenné štatisticky významne ovplyvňujú závisle premennú – ekvivalentný disponibilný príjem.

<sup>9</sup> Hodnoty testovacej štatistiky môžu nadobúdať len kladné hodnoty.

Tabuľka 4.4 Odhadnutie faktorov všeobecného lineárneho modelu EDP

Parameter	Estimate		Standard Error	t Value	Pr >  t
Intercept	11397.33366	B	120.9316983	94.25	<.0001
Ekonomická aktivita Neaktívna_osoba	-1270.14406	B	89.6179489	-14.17	<.0001
Ekonomická aktivita Nezamestnaná_osoba	-3615.59440	B	132.4362192	-27.30	<.0001
Ekonomická aktivita Osoba_na_dôchodku	-1436.70187	B	82.0863938	-17.50	<.0001
Ekonomická aktivita z_zamestnaná_osoba	0.00000	B	.	.	.
Typ domácnosti 1_dospely	242.01182	B	130.1172158	1.86	0.0629
Typ domácnosti 1_dospely_aspon_1_dieta	-915.82228	B	211.3535518	-4.33	<.0001
Typ domácnosti 2_dospeli	2575.26066	B	113.8508375	22.62	<.0001
Typ domácnosti 2_dospeli_1_dieta	538.20793	B	125.6623523	4.28	<.0001
Typ domácnosti 2_dospeli_1_dochodca	1408.96532	B	124.2819903	11.34	<.0001
Typ domácnosti 2_dospeli_aspon_3_deti	-1494.52203	B	188.5380825	-7.93	<.0001
Typ domácnosti ostatné_bezdetne_domácnosti	3033.30001	B	109.9908542	27.58	<.0001
Typ domácnosti ostatné_domácnosti_s_detmi	1270.00876	B	110.1670897	11.53	<.0001
Typ domácnosti z_2_dospeli_2_deti	0.00000	B	.	.	.
Rodinný stav slobodná_osoba	-245.91006	B	69.0696623	-3.56	0.0004
Rodinný stav vdovec_vdova	-660.18278	B	100.6447880	-6.56	<.0001
Rodinný stav z_osoba_v_manželstve	0.00000	B	.	.	.
Vzdelanie nižšie_ako_sekundárne	-2307.44018	B	100.8483335	-22.88	<.0001
Vzdelanie post-sekundárne	-1446.41665	B	200.3904724	-7.22	<.0001
Vzdelanie vysokoskolske_1_stupen	-664.80140	B	157.0764096	-4.23	<.0001
Vzdelanie vyššie_sekundárne	-1561.10014	B	72.6067333	-21.50	<.0001
Vzdelanie z_vysokoskolske_2_3_stupen	0.00000	B	.	.	.
Zdravotný stav priemerný	-623.01023	B	68.4537569	-9.10	<.0001
Zdravotný stav zlý	-943.78108	B	83.7217657	-11.27	<.0001
Zdravotný stav zz_dobrý	0.00000	B	.	.	.
Urbanizácia priemerne_osídlenie	-640.73263	B	77.2852408	-8.29	<.0001
Urbanizácia riedke_osídlenie	-928.99038	B	77.2792780	-12.02	<.0001
Urbanizácia z_husté_osídlenie	0.00000	B	.	.	.
Kraj BB	-773.26935	B	107.6281661	-7.18	<.0001
Kraj KE	-1078.00321	B	105.8658997	-10.18	<.0001
Kraj NR	-453.70685	B	108.6304826	-4.18	<.0001
Kraj PO	-1603.09392	B	106.6766961	-15.03	<.0001
Kraj TN	-544.03040	B	112.0774722	-4.85	<.0001
Kraj TT	-359.74474	B	115.1592564	-3.12	0.0018
Kraj ZA	-890.77627	B	108.5354395	-8.21	<.0001
Kraj z_BA	0.00000	B	.	.	.

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

V tabuľke 4.4 sú zobrazené kategórie nezávisle premenných. V ďalšom kroku sme vytvorili umelé premenné zo všetkých kategórií a jednu referenčnú<sup>10</sup> kategóriu. Tieto umelé premenné nám nahrádzajú nezávisle premenné v modeli. Ako príklad môžeme uviesť premennú Kraj, ktorá je tvorená ôsmymi kategóriami: Bratislavský kraj, Trnavský kraj, Trenčiansky kraj, Nitriansky kraj, Žilinský kraj, Banskobystrický kraj, Prešovský kraj a Košický kraj. Premennú Bratislavský kraj sme zvolili ako referenčnú kategóriu označovanú z\_BA.

Aby sme mohli overiť štatistickú významnosť koeficientov, musíme otestovať nulovú hypotézu:

*H<sub>0</sub>: regresný koeficient nie je štatisticky významný,*

oproti alternatívnej hypotéze

*H<sub>1</sub>: regresný koeficient je štatisticky významný.*

Na základe výsledkov t - testu môžeme stanoviť záver predpokladaných tvrdení, t. j. prijať respektíve zamietnuť nulovú hypotézu.

Na posúdenie štatistickej významnosti regresných koeficientov, využijeme p - hodnotu, ktorú porovnáme s hladinou významnosti ( $\alpha = 0,05$ ). V prípade, že p - hodnota bude menšia ako určená hladina významnosti, môžeme tvrdiť, že regresné koeficienty sú štatisticky významné.

Pri faktore ekonomická aktivita majú všetky umelé premenné p - hodnotu menšiu ako hladinu významnosti ( $\alpha = 0,05$ ), z čoho vyplýva, že regresné koeficienty sú štatisticky významné. Taktiež môžeme z tabuľky 4.4 vidieť, že najväčší rozdiel v ekvivalentom disponibilnom príjme je v porovnaní s ostatnými kategóriami medzi nezamestnanou a zamestnanou osobou. To znamená, že ročný ekvivalentný disponibilný

---

<sup>10</sup> Referenčná kategória – je spravidla najpočetnejšia alebo najporovnateľnejšia kategória podľa nášho uváženia. V práci sme referenčné kategórie označili písmenom „z“, čo môžeme vidieť v prílohe alebo v tabuľke 4.4.

príjem nezamestnanej osoby je v priemere o 3615,59 eur nižší, ako ekvivalentný disponibilný príjem zamestnanej osoby.

Umelá premenná domácnosť jedného člena pri premennej typ domácnosti má vyššiu p - hodnotu ako hladina významnosti. Môžeme teda povedať, že umelá premenná domácnosť jedného dospelého člena nie je štatisticky významná a ostatné umelé premenné môžeme považovať za štatisticky významné. Ročný EDP osoby žijúcej v domácnosti s dvomi dospelými aspoň s 3 závislými deťmi je v priemere o 1494,52 eur nižší, ako EDP osoby žijúcej v domácnosti dvoch dospelých osôb s dvomi závislými deťmi. Ročný EDP osoby žijúcej v domácnosti s dvomi dospelými bez závislých detí je v priemere o 2575,26 eur vyšší, ako EDP osoby žijúcej v domácnosti s dvomi dospelými osobami s dvomi závislými deťmi.

Pri premennej rodinný stav sú všetky premenné štatisticky významne, pretože ich p - hodnota je menšia ako hladina významnosti. V porovnaní s referenčnou kategóriou môžeme uviesť, že napríklad ročný EDP ovdovelej osoby je v priemere o 660,18 eur nižší, ako EDP osoby, ktorá je v manželskom zväzku.

Ďalšou premennou je vzdelanie, pri ktorom sú všetky umelé premenné taktiež štatisticky významné. Najväčší rozdiel v EDP je medzi osobou s nižším ako sekundárnym vzdelaním a osobou s vysokoškolským vzdelaním 2. a 3. stupňa. Ročný EDP osoby s nižším ako sekundárnym vzdelaním je v priemere o 2307,44 eur nižší, ako EDP osoby s vysokoškolským vzdelaním druhého a tretieho stupňa.

Štatisticky významné umelé premenné sú aj pri faktore zdravotný stav. Najväčší rozdiel v porovnaní s referenčnou kategóriou je medzi zlým zdravotným stavom a dobrým zdravotným stavom. Ročný EDP osoby so zlým zdravotným stavom je v priemere o 943,78 eur nižší, ako EDP osoby s dobrým zdravotným stavom.

Pri premennej urbanizácia je najväčší rozdiel medzi osobami žijúcimi na území s riedkym osídlením a referenčnou kategóriou (osobami žijúcimi na území s hustým osídlením). Ročný EDP osoby žijúcej na území s riedkym osídlením je v priemere o 928,99 eur nižší, ako EDP osoby žijúcej na území s hustým osídlením.

V prípade premennej kraj sú všetky umelé premenné štatisticky významné. Najväčší rozdiel v EDP je medzi Prešovským krajom a Bratislavským krajom. Naopak najmenší rozdiel v EDP sa potvrdil medzi Trnavským a Bratislavským krajom. Ročný EDP osôb v Prešovskom kraji je v priemere o 1603,09 eur nižší, ako EDP osôb v Bratislavskom kraji. Ročný EDP osoby žijúcej v Trnavskom kraji je v priemere o 359,74 eur nižší, ako EDP osoby žijúcej v Bratislavskom kraji.

Na overenie významnosti rozdielu medzi dvojicou kategórii premennej využijeme príkaz LSMEANS, prostredníctvom ktorého overujeme nulovú hypotézu:

*H<sub>0</sub>: medzi dvojicou kategórií rozdiel marginálnych stredných hodnôt nie je štatisticky významný,*

oproti alternatívnej hypotéze:

*H<sub>1</sub>: medzi dvojicou kategórií rozdiel marginálnych stredných hodnôt je štatisticky významný.*

O výsledku predpokladaných tvrdení sa rozhodneme na základe výsledných p - hodnôt.

Tabuľka 4.5 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od ekonomickej aktivity

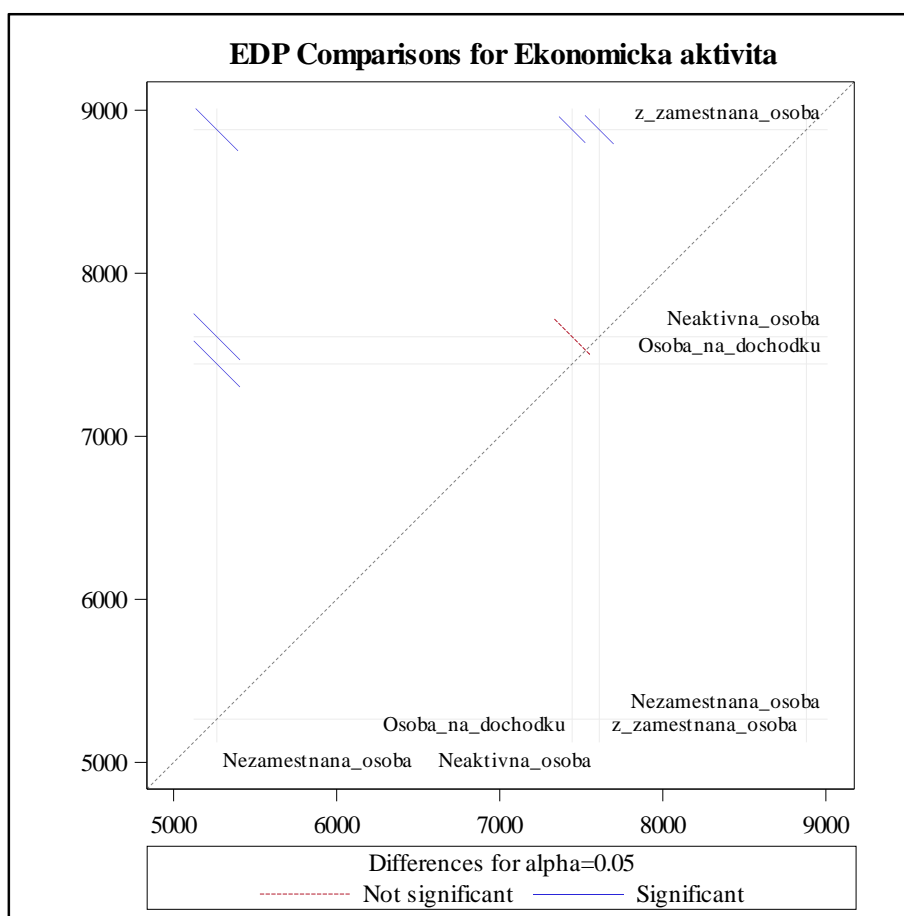
Least Squares Means for effect Ekonomicka aktivita Pr >  t  for H <sub>0</sub> : LSMean(i)=LSMean(j) Dependent Variable: EDP				
i/j	1	2	3	4
1		<.0001	0.1333	<.0001
2	<.0001		<.0001	<.0001
3	0.1333	<.0001		<.0001
4	<.0001	<.0001	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.6 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od ekonomickej aktivity

Ekonomická aktivita	Označenie kategórií	EDP LSMEAN	95% Confidence Limits	
Neaktívna_osoba	1	7610.586358	7431.615547	7789.557168
Nezamestnaná_osoba	2	5265.136014	4990.358247	5539.913781
Osoba_na_dôchodku	3	7444.028548	7268.715385	7619.341711
z_zamestnaná_osoba	4	8880.730414	8733.189957	9028.270872

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide



Obrázok 4.5 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od ekonomickej aktivity

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Štatisticky nevýznamný rozdiel medzi marginálnymi strednými hodnotami ekvivalentného disponibilného príjmu je medzi osobami na dôchodku a neaktívnymi osobami, čomu nasvedčujú výsledky v tabuľke 4.5, kde vidíme, že p - hodnota ( $p = 0,1333$ ) je pri tejto dvojici vyššia ako hladina významnosti. Taktiež aj obrázok 4.5 nám poukazuje

na štatisticky nevýznamný rozdiel dvojice, pri ktorom platí, že ak jednotlivé úsečky presiahnu uhlopriečku, tak v tom prípade medzi dvojicami existuje štatisticky nevýznamný rozdiel (červená farba).

S 95 % - nou spoľahlivosťou sa podľa tabuľky 4.6 najväčší rozdiel marginálnej hodnoty pohybuje pri nezamestnanej osobe a to v intervale od 4990,36 eur do 5539,91 eur. Naopak najmenší rozdiel s 95 % - nou spoľahlivosťou marginálnej hodnoty v intervale od 8733,19 eur do 9028,27 eur sa pohybuje pri zamestnanej osobe. Z tabuľky 4.6 taktiež vidíme EDP LSMEAN, čo predstavujú bodové odhady marginálnych stredných hodnôt EDP v prípade jednotlivých kategórií. Najnižšiu marginálnu strednú hodnotu EDP 5265,14 eur reprezentuje nezamestnaná osoba, naopak najvyššia marginálna stredná hodnota EDP je u zamestnanej osoby 8880,73 eur. Na základe týchto hodnôt, môžeme taktiež vidieť, že hodnoty sa navzájom odlišujú.

Tabuľka 4.7 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od typu domácnosti

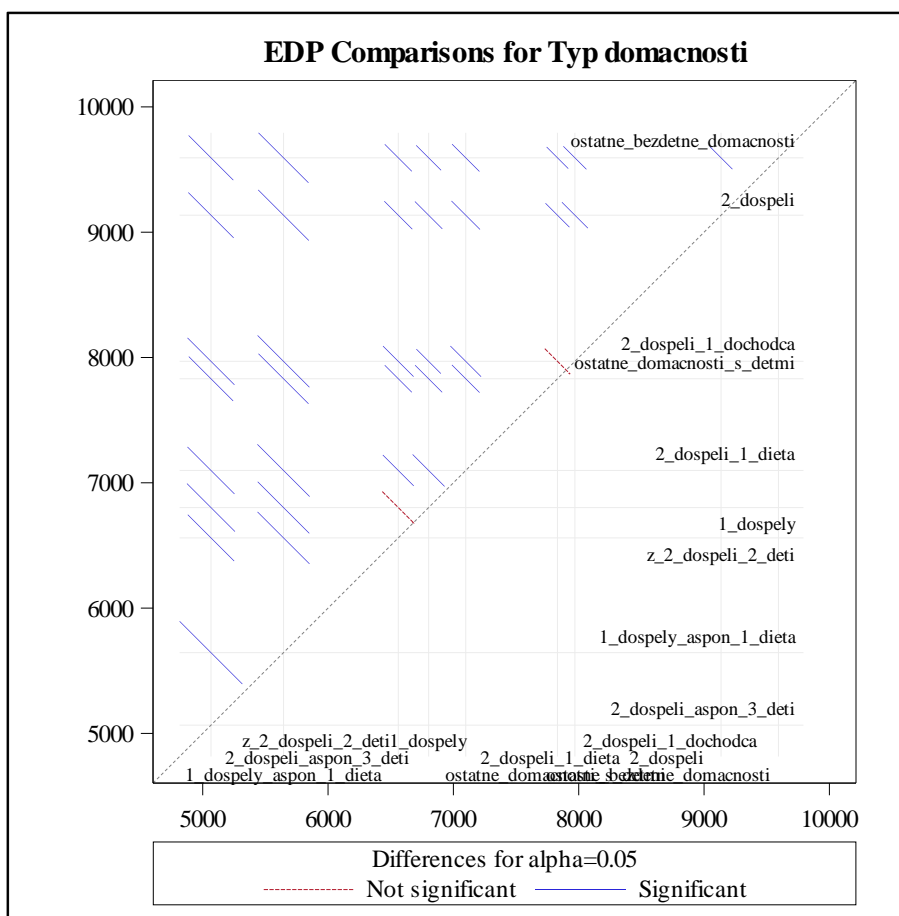
Least Squares Means for effect Typ domácnosti Pr >  t  for H0: LSMean(i)=LSMean(j)									
Dependent Variable: EDP									
i/j	1	2	3	4	5	6	7	8	9
1		<.0001	<.0001	0.0227	<.0001	<.0001	<.0001	<.0001	0.0629
2	<.0001		<.0001	<.0001	<.0001	0.0234	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
4	0.0227	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.1779	<.0001
6	<.0001	0.0234	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
7	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
8	<.0001	<.0001	<.0001	<.0001	0.1779	<.0001	<.0001		<.0001
9	0.0629	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.8 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od typu domácností

Typ domácností	Označenie kategórií	EDP LSMEAN	95% Confidence Limits	
1_dospely	1	6802.419910	6612.575581	6992.264240
1_dospely_aspon_1_dieta	2	5644.585814	5254.922229	6034.249399
2_dospeli	3	9135.668746	8948.299284	9323.038207
2_dospeli_1_dieta	4	7098.616019	6878.293276	7318.938762
2_dospeli_1_dochodca	5	7969.373413	7780.538483	8158.208342
2_dospeli_aspon_3_deti	6	5065.886059	4706.567571	5425.204546
ostatne_bezdetne_domacnosti	7	9593.708100	9423.961546	9763.454654
ostatne_domacnosti_s_detmi	8	7830.416852	7650.887922	8009.945782
z_2_dospeli_2_deti	9	6560.408090	6340.747102	6780.069078

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide



Obrázok 4.6 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od typu domácností

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Najmenší rozdiel medzi marginálnymi strednými hodnotami ekvivalentného disponibilného príjmu podľa tabuľky 4.7 je medzi osobami žijúcimi v domácnostiach dvoch dospelých, z ktorých aspoň jeden je vo veku 65 rokov a viac a medzi osobami žijúcimi v ostatných domácnostiach so závislými deťmi. P - hodnota ( $p = 0,1779$ ) je pri tejto dvojici vyššia ako hladina významnosti, čiže je medzi nimi štatisticky nevýznamný rozdiel.

Štatisticky nevýznamný rozdiel ( $p = 0,0629$ ) sa potvrdil aj v prípade osoby žijúcej v domácnosti len jedného dospelého a osoby žijúcej v domácnosti dvoch dospelých s dvomi závislými deťmi. To znamená, že v prípade tejto dvojice môžeme považovať marginálne stredné hodnoty EDP za zhodné.

V ostatných prípadoch medzi jednotlivými dvojicami je rozdiel štatisticky významný, kde to potvrdili aj p - hodnoty, ktoré sú menšie ako hladina významnosti alebo aj úsečky na obrázku 4.6, ktoré charakterizuje modrá farba.

Domácnosť s 2 dospelými a aspoň s 3 deťmi má podľa tabuľky 4.8 najnižšiu marginálnu strednú hodnotu 5065,89 eur. Na druhej strane najvyššiu marginálnu strednú hodnotu 9593,71 eur predstavujú ostatné bezdetné domácnosti. Medzi týmito hodnotami môžeme vidieť, že sa navzájom odlišujú, čo znamená, že predstavujú štatisticky významný rozdiel.

Tabuľka 4.9 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od rodinného stavu

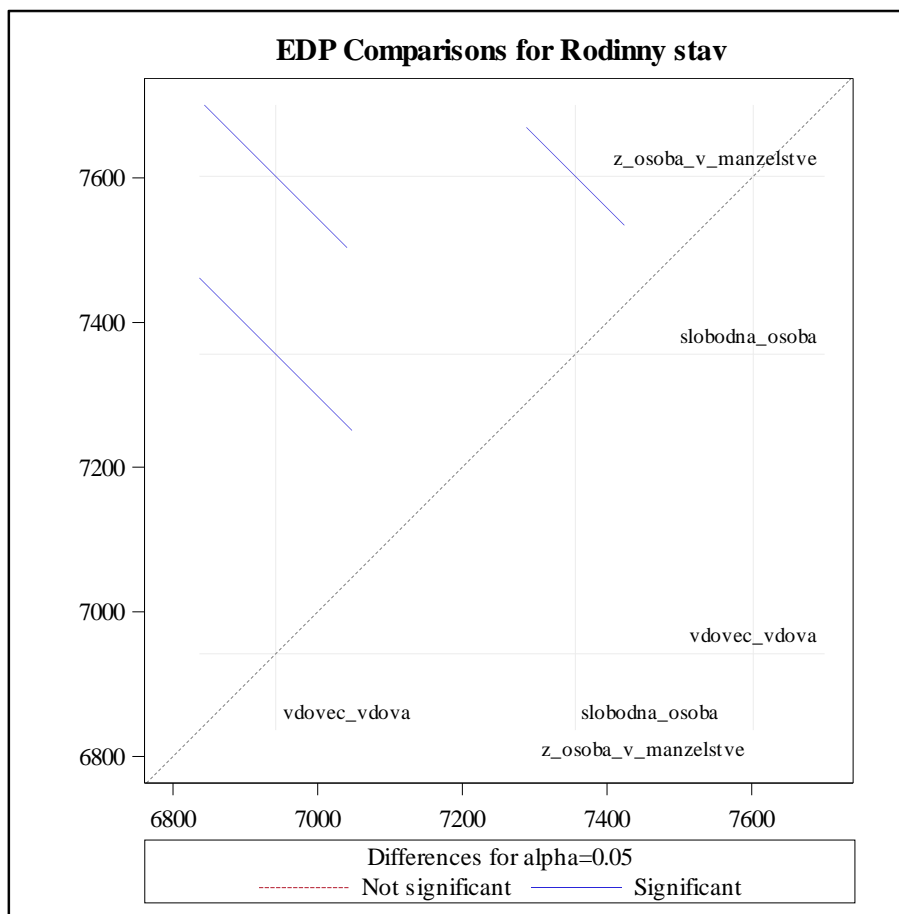
<b>Least Squares Means for effect Rodinný stav</b>			
<b>Pr &gt;  t  for H0: LSMean(i)=LSMean(j)</b>			
<b>Dependent Variable: EDP</b>			
<b>i/j</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>1</b>		0.0001	0.0004
<b>2</b>	0.0001		<.0001
<b>3</b>	0.0004	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.10 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od rodinného stavu

Rodinný stav	Označenie kategórií	EDP LSMEAN	95% Confidence Limits	
slobodná_osoba	1	7356.241219	7196.909387	7515.573052
vdovec_vdova	2	6941.968501	6727.036503	7156.900499
z_osoba_v_manželstve	3	7602.151280	7460.550897	7743.751662

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide



Obrázok 4.7 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od rodinného stavu

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Na základe tabuľky 4.9 môžeme tvrdiť, že medzi všetkými dvojicami kategórií je rozdiel v marginálnych stredných hodnotách ekvivalentného disponibilného príjmu štatisticky významný. Na štatistickú významnosť poukazuje aj obrázok 4.7, pretože platí, že pokiaľ jednotlivé úsečky nepresiahnu uhlopriečku, tak v tom prípade medzi dvojicami existuje štatisticky významný rozdiel (modrá farba). Taktiež vidíme, že najmenší avšak stále

štatisticky významný rozdiel je medzi osobou v manželstve a slobodnou osobou ( $p = 0.0004$ ).

Najnižšia marginálna stredná hodnota EDP podľa tabuľky 4.10 je u ovdovelej osoby 6941,97 eur, ďalej nasleduje slobodná osoba, ktorej marginálna stredná hodnota EDP je 7356,24 eur a posledná je osoba v manželstve, ktorej marginálna stredná hodnota EDP predstavuje zároveň aj najvyššiu hodnotu 7602,15 eur. Z uvedených výsledkov môžeme vidieť, že sa jednotlivé hodnoty odlišujú, čiže medzi nimi existuje štatisticky významný rozdiel.

Tabuľka 4.11 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od vzdelania

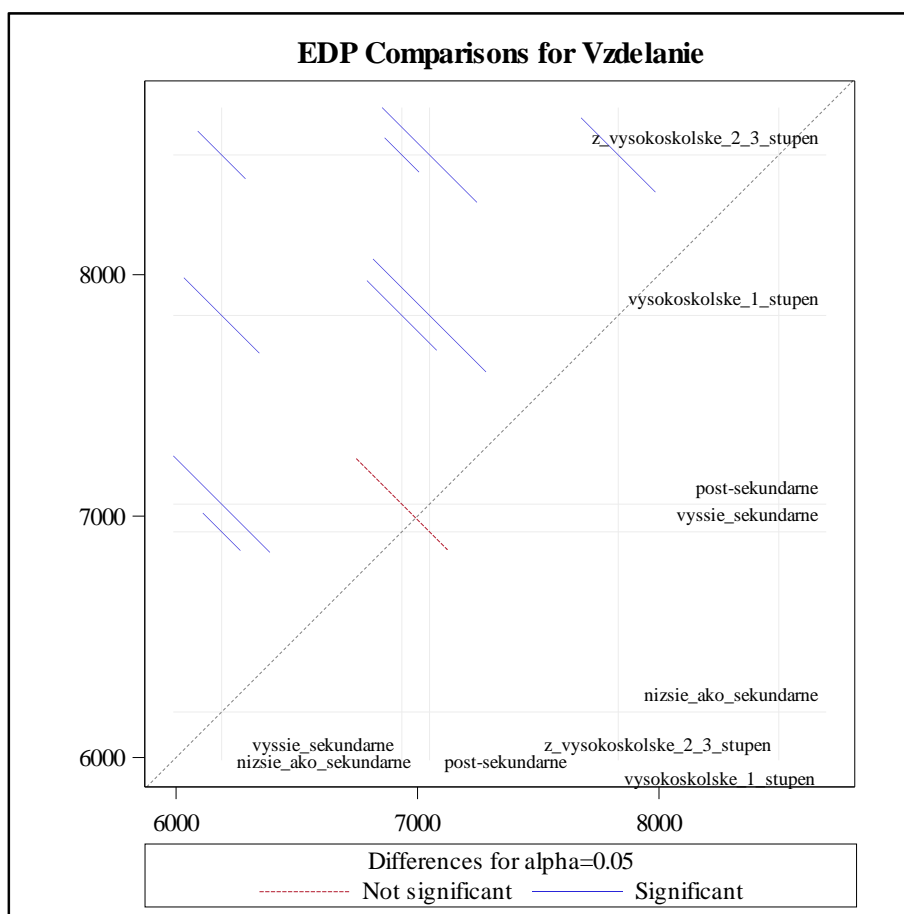
<b>Least Squares Means for effect Vzdelanie</b> <b>Pr &gt;  t  for H0: LSMean(i)=LSMean(j)</b>					
<b>Dependent Variable: EDP</b>					
<b>i/j</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>1</b>		<.0001	<.0001	<.0001	<.0001
<b>2</b>	<.0001		0.0011	0.5515	<.0001
<b>3</b>	<.0001	0.0011		<.0001	<.0001
<b>4</b>	<.0001	0.5515	<.0001		<.0001
<b>5</b>	<.0001	<.0001	<.0001	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.12 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od vzdelania

<b>Vzdelanie</b>	<b>Označenie kategórii</b>	<b>EDP LSMEAN</b>	<b>95% Confidence Limits</b>	
<b>nižšie_ako_sekundárne</b>	<b>1</b>	6188.631832	6031.530715	6345.732949
<b>post-sekundárne</b>	<b>2</b>	7049.655355	6663.658212	7435.652498
<b>vysokoskolske_1_stupen</b>	<b>3</b>	7831.270602	7533.691614	8128.849590
<b>vyššie_sekundárne</b>	<b>4</b>	6934.971871	6817.925096	7052.018645
<b>z_vysokoskolske_2_3_stupen</b>	<b>5</b>	8496.072007	8330.756484	8661.387531

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide



Obrázok 4.8 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od vzdelania

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Najmenší rozdiel medzi marginálnymi strednými hodnotami ekvivalentného disponibilného príjmu je viditeľný medzi osobou s post-sekundárnym vzdelaním a osobou s vyšším sekundárnym vzdelaním. Pri tejto dvojici je p - hodnota ( $p = 0,5515$ ) vyššia ako hladina významnosti, z toho vyplýva, že medzi uvedenou dvojicou nie je štatistická významnosť. Preto môžeme uvažovať nad zlúčením týchto dvoch kategórií do jednej novej kategórie. V prípade ďalších dvojíc je p - hodnota menšia ako hladina významnosti, čo môžeme vidieť na základe tabuľky 4.11, alebo obrázku 4.8, kde ostatné úsečky nepresahujú uhlopriečku, čo poukazuje na existenciu štatisticky významného rozdielu marginálnych stredných hodnôt EDP.

Z tabuľky 4.12 vidíme, že najnižšiu marginálnu strednú hodnotu EDP má kategória osoba s nižším ako sekundárnym vzdelaním (6188,63 eur), na druhej strane najvyššiu marginálnu strednú hodnotu vidíme v prípade kategórie osoba s vysokoškolským vzdelaním

2. a 3. stupňa, ktorej bodový odhad predstavuje hodnotu 8496,07 eur. Môžeme taktiež vidieť, že tieto hodnoty sa od seba vzájomne líšia. Naopak marginálne stredné hodnoty EDP osôb s post-sekundárnym vzdelaním (7049,66 eur) a osôb s vyšším sekundárnym vzdelaním (6934,97 eur) sa od seba výrazne neodlišujú, čo predstavuje ich štatisticky nevýznamný rozdiel.

Tabuľka 4.13 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od zdravotného stavu

Least Squares Means for effect Zdravotný stav Pr >  t  for H0: LSMean(i)=LSMean(j)			
Dependent Variable: EDP			
i/j	1	2	3
1		0.0001	<.0001
2	0.0001		<.0001
3	<.0001	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.14 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od zdravotného stavu

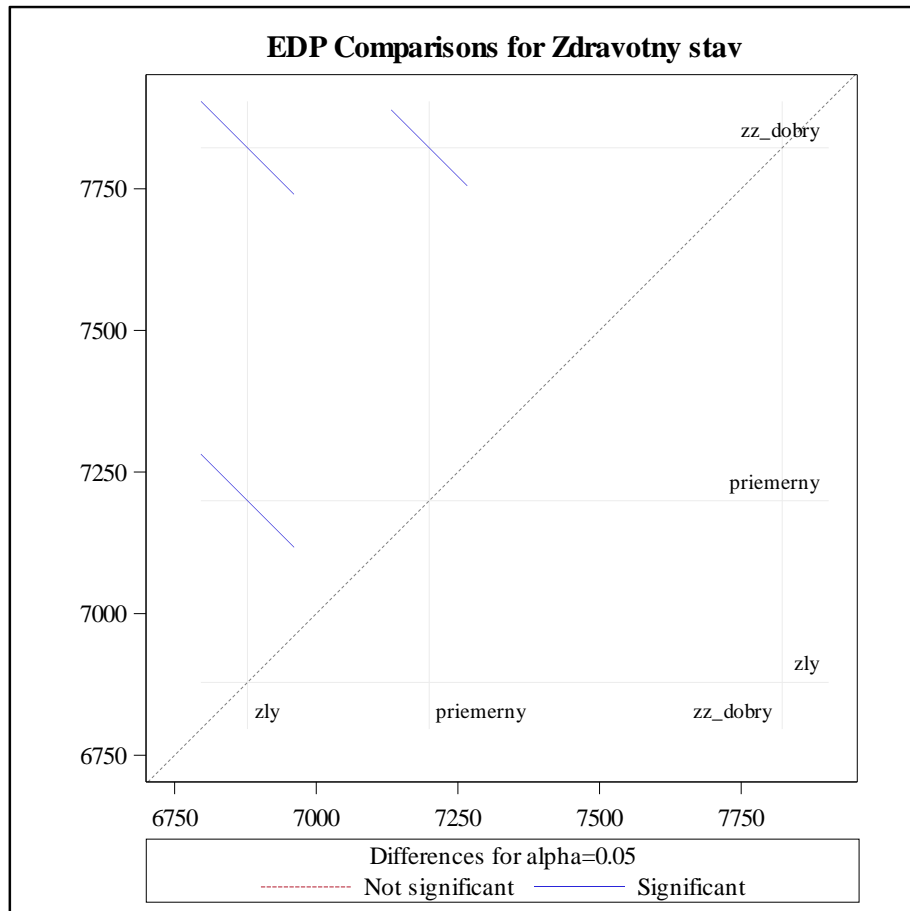
Zdravotný stav	Označenie kategórií	EDP LSMEAN	95% Confidence Limits	
priemerný	1	7199.373875	7034.237747	7364.510003
zlý	2	6878.603022	6690.759397	7066.446646
zz_dobry	3	7822.384104	7683.924964	7960.843243

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Na základe tabuľky 4.13 vidíme, že medzi jednotlivými dvojicami existuje štatisticky významný rozdiel v marginálnych stredných hodnotách ekvivalentného disponibilného príjmu. Z uvedeného vyplýva, že všetky marginálne stredné hodnoty EDP medzi dvojicami sú rozdielne. Najmenší, ale napriek tomu štatisticky významný rozdiel medzi marginálnymi strednými hodnotami ekvivalentného disponibilného príjmu je medzi priemerným a zlým zdravotným stavom osôb.

Najnižšiu hodnotu bodového odhadu podľa tabuľky 4.14 má osoba so zlým zdravotným stavom (6878,60 eur), za ňou nasleduje osoba s priemerným zdravotným

stavom (7199,37 eur) a posledná je osoba s dobrým zdravotným stavom, ktorej marginálna stredná hodnota EDP je 7822,38 eur.



Obrázok 4.9 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od zdravotného stavu

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.15 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od urbanizácie

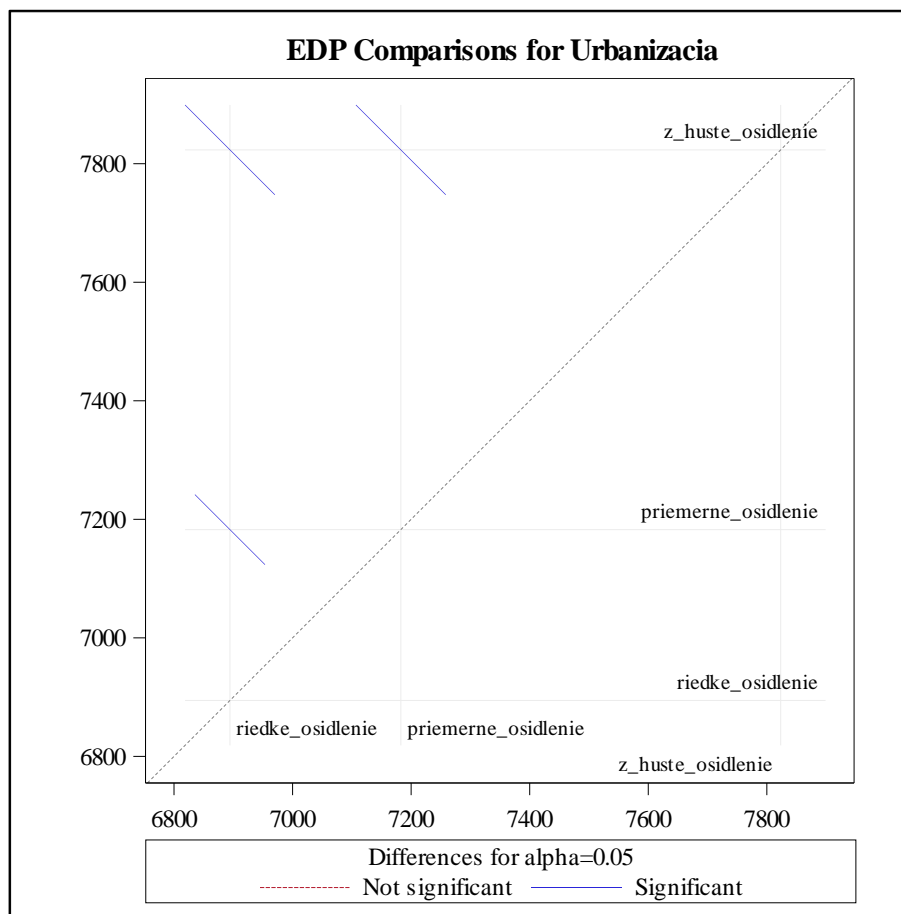
Least Squares Means for effect Urbanizácia			
Pr >  t  for H0: LSMean(i)=LSMean(j)			
Dependent Variable: EDP			
i/j	1	2	3
1		<.0001	<.0001
2	<.0001		<.0001
3	<.0001	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.16 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od urbanizácie

Urbanizácia	Označenie kategórií	EDP LSMEAN	95% Confidence Limits	
priemerne_osídlenie	1	7182.628706	7029.858735	7335.398677
riedke_osídlenie	2	6894.370958	6739.328848	7049.413069
z_husté_osídlenie	3	7823.361336	7649.305366	7997.417307

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide



Obrázok 4.10 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od urbanizácie

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Hodnoty v tabuľke 4.15 nám nasvedčujú, že medzi všetkými dvojicami kategórií je štatisticky významný rozdiel v marginálnych stredných hodnotách EDP, pretože p - hodnota je v prípade všetkých dvojíc kategórií menšia ako hladina významnosti. Taktiež sa môžeme o tvrdení presvedčiť aj na základe obrázku 4.10, kde vidíme, že ani jedna z úsečiek nepresiahla uhlopriečku, čo značí, že medzi dvojicami existuje štatisticky významný rozdiel,

tým pádom marginálne hodnoty ekvivalentného disponibilného príjmu osôb žijúcich na území s riedkym osídlením a osôb žijúcich na území s priemerne hustým osídlením sú rozdielne.

Štatisticky významný rozdiel hodnôt môžeme sledovať aj na základe tabuľky 4.16. Vidíme, že najnižšia hodnota bodového odhadu EDP je v prípade osoby žijúcej na území s riedkym osídlením (6894,37 eur), pri osobe žijúcej na území s priemerne hustým osídlením má bodový odhad EDP hodnotu 7182,63 eur a nakoniec bodový odhad ekvivalentného disponibilného príjmu osoby žijúcej na území s hustým osídlením má hodnotu 7823,36 eur. Podľa výsledkov vidíme, že jednotlivé hodnoty sa od seba odlišujú, čo nám potvrdzuje štatisticky významný rozdiel medzi dvojicami kategórií faktora.

Tabuľka 4.17 Overenie štatistickej významnosti zhody marginálnych stredných hodnôt EDP osôb domácnosti v závislosti od kraju

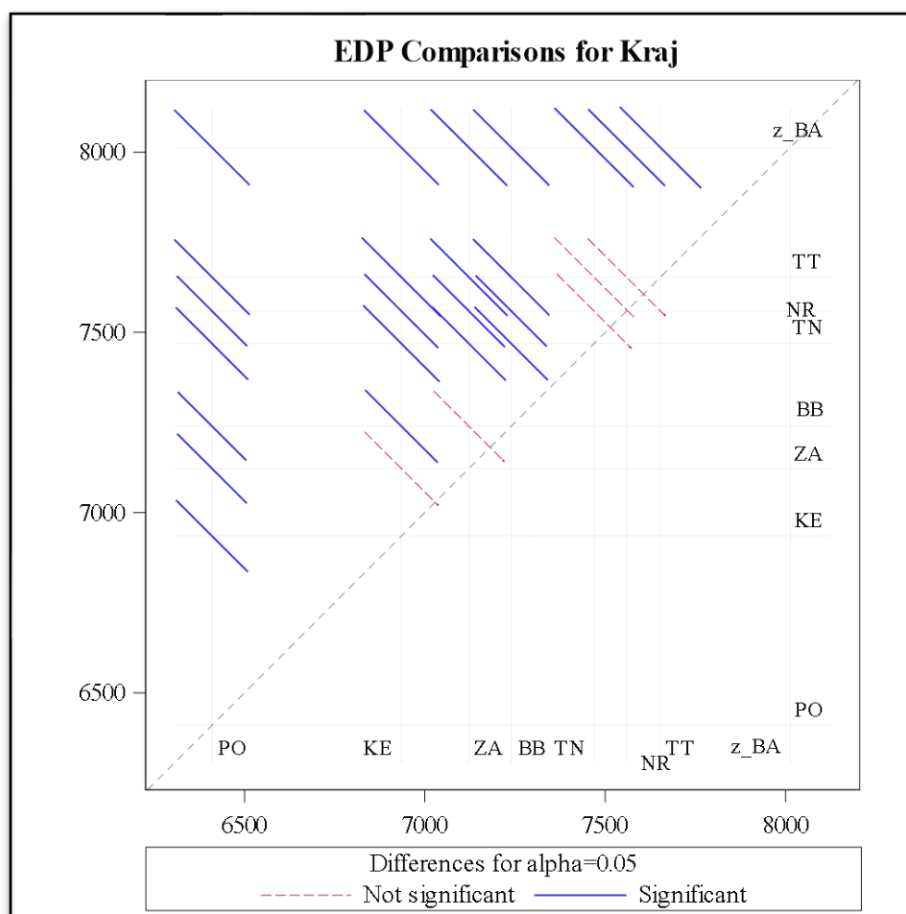
Least Squares Means for effect Kraj Pr >  t  for H0: LSMean(i)=LSMean(j)								
Dependent Variable: EDP								
i/j	1	2	3	4	5	6	7	8
1		0.0031	0.0015	<.0001	0.0270	0.0001	0.2397	<.0001
2	0.0031		<.0001	<.0001	<.0001	<.0001	0.0725	<.0001
3	0.0015	<.0001		<.0001	0.3908	0.3915	<.0001	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001
5	0.0270	<.0001	0.3908	<.0001		0.1007	0.0009	<.0001
6	0.0001	<.0001	0.3915	<.0001	0.1007		<.0001	0.0018
7	0.2397	0.0725	<.0001	<.0001	0.0009	<.0001		<.0001
8	<.0001	<.0001	<.0001	<.0001	<.0001	0.0018	<.0001	

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Tabuľka 4.18 Bodové a intervalové odhady marginálnych stredných hodnôt pre EDP osôb domácností v závislosti od kraju

Kraj	Označenie kategórií	EDP LSMEAN	95% Confidence Limits	
BB	1	7239.679073	7051.760085	7427.598062
KE	2	6934.945213	6741.660307	7128.230118
NR	3	7559.241572	7365.236883	7753.246262
PO	4	6409.854511	6224.654305	6595.054717
TN	5	7468.918027	7269.869117	7667.966938
TT	6	7653.203687	7443.580842	7862.826531
ZA	7	7122.172159	6928.835735	7315.508582
z_BA	8	8012.948426	7816.566332	8209.330521

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide



Obrázok 4.11 Interval spoľahlivosti marginálnych stredných hodnôt pre EDP domácností v závislosti od kraju

Zdroj: EU-SILC 2020, vlastné spracovanie v SAS Enterprise Guide

Podľa tabuľky 4.17 vidíme, že najmenší rozdiel medzi marginálnymi strednými hodnotami EDP je medzi osobami žijúcimi v Košickom kraji a osobami žijúcimi v Žilinskom kraji, čomu nasvedčuje aj výsledná p - hodnota ( $p = 0,0725$ ), vďaka ktorej môžeme potvrdiť, že medzi uvedenou dvojicou nie je štatisticky významný rozdiel v marginálnych stredných hodnotách EDP. To isté platí aj pre dvojice kategórií – osoby žijúce v Trenčianskom kraji a osoby žijúce v Trnavskom kraji ( $p = 0,1099$ ), osoby žijúce v Žilinskom kraji a osoby žijúce v Banskobystrickom kraji ( $p = 0,2397$ ), osoby žijúce v Trenčianskom kraji a osoby žijúce v Nitrianskom kraji ( $p = 0,3908$ ), osoby žijúce v Trnavskom kraji a osoby žijúce v Nitrianskom kraji ( $p = 0,4064$ ). Z toho dôvodu môžeme uvažovať nad zlúčením jednotlivých dvojíc do jednej novej kategórie. V prípade ostatných dvojíc nám analýzy potvrdili štatisticky významný rozdiel v marginálnych stredných hodnotách EDP.

Najnižšiu marginálnu strednú hodnotu EDP podľa tabuľky 4.18 predstavuje kategória Prešovský kraj (6409,85 eur) a naopak najvyššiu marginálnu strednú hodnotu EDP reprezentuje kategória Bratislavský kraj s hodnotou 8012,95 eur.

## Záver

Chudoba je veľmi aktuálnou témou súčasnosti a zároveň aj jedným z najväčších problémov, ktorým svet čelí. Je preto nutné vyvinúť maximálne úsilie v boji proti nej. Základom boja je poznanie, čo problém chudoby spôsobuje. V práci sme sa zamerali na analýzu faktorov, ktoré môžu ovplyvňovať chudobu na Slovensku.

Cieľom práce bolo analyzovanie životnej úrovne domácností Slovenskej republiky, na ktorú sme využili všeobecné lineárne modely. V práci sme analyzovali vplyv premenných ako sú ekonomická aktivita, typ domácnosti, rodinný stav, vzdelanie, zdravotný stav, urbanizácia a kraj na závisle premennú, ktorú predstavoval ekvivalentný disponibilný príjem slovenských domácností v roku 2020.

Na základe výsledkov analýz sme zistili, že najvýznamnejší vplyv na ekvivalentný disponibilný príjem domácností má práve ekonomická aktivita, kde sa prejavil najväčší signifikantný rozdiel medzi nezamestnanou osobou a zamestnanou osobou. Ďalší veľký signifikantný rozdiel sa prejavil medzi osobou s nižším ako sekundárnym vzdelaním a osobou s vysokoškolským vzdelaním 2. a 3. stupňa. Na základe toho môžeme konštatovať, že najohrozenejšie skupiny osôb, čeliace príjmovej chudobe sú predovšetkým tí, ktorí sú nezamestnaní, resp. majú primárne vzdelanie.

Použitím príkazu LSMEANS sme identifikovali hlbšie vzťahy medzi kategóriami vybraných faktorov z pohľadu cieľovej premennej EDP. Zistili sme, že existencia signifikantného rozdielu marginálnych stredných hodnôt EDP medzi niektorými kategóriálnymi premenných sa nepotvrdila. Pri premennej ekonomická aktivita sme zistili, že medzi kategóriou osoba na dôchodku a neaktívna osoba je štatisticky nevýznamný rozdiel. Pri premennej typ domácnosti je štatisticky nevýznamný rozdiel medzi osobami žijúcimi v domácnostiach dvoch dospelých, z ktorých aspoň jeden je vo veku 65 rokov a viac a medzi osobami žijúcimi v ostatných domácnostiach so závislými deťmi a taktiež medzi osobou žijúcou v domácnosti len jedného dospelého a osobou žijúcou v domácnosti dvoch dospelých s dvomi závislými deťmi. Pri ďalšej premennej vzdelanie je štatisticky nevýznamný rozdiel v marginálnych stredných hodnotách cieľovej premennej EDP medzi osobou s post-sekundárnym vzdelaním a osobou s vyšším sekundárnym vzdelaním. Pri poslednej premennej, ktorou je kraj, je štatisticky nevýznamný rozdiel medzi dvojicami osôb

žijúcich v Košickom a Žilinskom kraji, osôb žijúcimi v Trenčianskom a Trnavskom kraji, osôb žijúcich v Žilinskom a Banskobystrickom kraji, osôb žijúcimi v Trenčianskom a Nitrianskom kraji a poslednú dvojicu tvoria osoby žijúce v Trnavskom a Nitrianskom kraji. Pri všetkých uvedených dvojiciach môžeme považovať marginálne stredné hodnoty EDP za zhodné a v prípade ďalších analýz môžeme uvažovať nad zlúčením jednotlivých dvojíc kategórií do jednej novej kategórie.

Bakalárska práca poskytuje stručný návod na aplikáciu všeobecných lineárnych modelov (procedúra GLM) pri posúdení vplyvu relevantných faktorov na cieľovú premennú využitím štatistického analytického softvéru SAS Enterprise Guide. Zároveň sme priblížili použitie príkazu LSMEANS pri kvantifikácii marginálnych stredných hodnôt a posúdení existencie signifikantného rozdielu medzi dvojicami kategórií faktora, čo je predovšetkým prínosné v prípade ďalších iných analýz, nielen pri analýze ekvivalentného disponibilného príjmu. Okrem spomínaného postupu aplikácie GLM je prínosom našej práce aj poskytnutie výsledkov analýz, ktoré vychádzajú z reálnych údajov, ktoré sme získali zo štatistického zisťovania EU-SILC 2020 a tým pádom pohľad na sociálno-ekonomickú situáciu obyvateľstva Slovenskej republiky.

## Zoznam použitej literatúry

ANDREß, H. J. a Lohmann, H.. The working poor in Europe: Employment, Poverty and Globalisation. Northampton. [elektronický zdroj]. MA: Edward Elgar, 2008. 336s. [cit. 04.02.2022]. ISBN 978-1-84720-798-2.

DELWICHE, L. D. a SLAUGHTER, S. J. The Little SAS Book: A primer THIRD EDITION [elektronický zdroj]. NC, USA 2003. [cit. 10.02.2022]. 354 s. ISBN 1-59047-333-7 Dostupné na: [https://www.dermepi.eu/wp-content/uploads/2017/04/Little.SAS\\_Book\\_A\\_Primer.Third\\_Edition.pdf](https://www.dermepi.eu/wp-content/uploads/2017/04/Little.SAS_Book_A_Primer.Third_Edition.pdf)

DOBSON, A. J. AN INTRODUCTION TO GENERALIZED LINEAR MODELS [elektronický zdroj]. 2. vyd. USA: Taylor & Francis Ltd, 2001. 240 s. [cit. 06.02.2022]. ISBN 9781584881650. Dostupné na: <https://reneues.files.wordpress.com/2010/01/an-introduction-to-generalized-linear-models-second-edition-dobson.pdf>

DODGE, Y. The Concise Encyclopedia of Statistics [elektronický zdroj]. Springer Science & Business Media, 2008. 622 s. [cit. 08.02.2022]. ISBN 978-0-387-33828-6. Dostupné na: [https://books.google.sk/books?id=k2zklGOBRDwC&pg=PA544&lpg=PA544&dq=y\\_ijk%3D%CE%BC%2B+%CE%B1\\_i%2B%CE%B2\\_j%2B%E3%80%96\(%CE%B1%CE%B2\)%E3%80%97\\_ij%2B%CE%B5\\_ijk&source=bl&ots=9n\\_4e26jp-&sig=ACfU3U2WgbIgad5jdXiABzc4Cqn5fBJO8g&hl=sk&sa=X&ved=2ahUKEwiHkey6yPf2AhVThP0HHSvDCucQ6AF6BAgWEAM#v=onepage&q&f=false](https://books.google.sk/books?id=k2zklGOBRDwC&pg=PA544&lpg=PA544&dq=y_ijk%3D%CE%BC%2B+%CE%B1_i%2B%CE%B2_j%2B%E3%80%96(%CE%B1%CE%B2)%E3%80%97_ij%2B%CE%B5_ijk&source=bl&ots=9n_4e26jp-&sig=ACfU3U2WgbIgad5jdXiABzc4Cqn5fBJO8g&hl=sk&sa=X&ved=2ahUKEwiHkey6yPf2AhVThP0HHSvDCucQ6AF6BAgWEAM#v=onepage&q&f=false)

EUROSTAT, Glossary: At risk of poverty or social exclusion (AROPE) [elektronický zdroj]. 2021. [cit. 05.02.2022]. Dostupné na: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:At\\_risk\\_of\\_poverty\\_or\\_social\\_exclusion\\_\(AROPE\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:At_risk_of_poverty_or_social_exclusion_(AROPE))

EUROSTAT, Glossary: At-risk-of-poverty rate [elektronický zdroj]. 2021. [cit. 04.02.2022]. Dostupné na: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:At-risk-of-poverty\\_rate](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:At-risk-of-poverty_rate)

EUROSTAT, Glossary: Equivalised disposable income [elektronický zdroj]. 2021. [cit. 02.02.2022]. Dostupné na: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Equivalised\\_disposable\\_income](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Equivalised_disposable_income)

EUROSTAT, Glossary: Material deprivation [elektronický zdroj]. 2021. [cit. 04.02.2022]. Dostupné na: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Material\\_deprivation](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Material_deprivation)

EUROSTAT, Glossary: The European Pillar of Social Rights Action Plan (EU 2030 targets) [elektronický zdroj]. 2021. [cit. 10.02.2022]. Dostupné na: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:The\\_European\\_Pillar\\_of\\_Social\\_Rights\\_Action\\_Plan\\_\(EU\\_2030\\_targets\)&stable=0&redirect=no](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:The_European_Pillar_of_Social_Rights_Action_Plan_(EU_2030_targets)&stable=0&redirect=no)

EUROSTAT, One in five people in the EU at risk of poverty or social exclusion [elektronický zdroj]. 2021. [cit. 28.01.2022]. Dostupné na: <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20211015-1>

EURÓPSKA KOMISIA, OZNÁMENIE KOMISIE EURÓPSKEHO PARLAMENTU, RADE, EURÓPSKEMU HOSPODÁRSKEMU A SOCIÁLNEMU VÝBORU A VÝBORU REGIÓNOV : Akčný plán na realizáciu Európskeho piliera sociálnych práv [elektronický zdroj]. 2021. [cit. 03.02. 2022]. Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:52021DC0102>

EURÓPSKA KOMISIA, Tematický informačný prehľad európskeho semestra: Sociálne začlenenie [elektronický zdroj]. 2017. 16 s. [cit. 06.02. 2022]. Dostupné na: [https://ec.europa.eu/info/sites/default/files/file\\_import/european-semester\\_thematic-factsheet\\_social\\_inclusion\\_sk.pdf](https://ec.europa.eu/info/sites/default/files/file_import/european-semester_thematic-factsheet_social_inclusion_sk.pdf)

GAJDOŠ, P. Človek, spoločnosť, prostredie, Bratislava: Sociologický ústav SAV, 2002. 374 s. ISBN 80-855444-15-6.

KENNEDY, A. BOJ PROTI CHUDOBE, SOCIÁLNEMU VYLÚČENIU A DISKRIMINÁCIÍ [elektronický zdroj]. 2021. [cit. 03.02.2022]. Dostupné na: [https://www.europarl.europa.eu/ftu/pdf/sk/FTU\\_2.3.9.pdf](https://www.europarl.europa.eu/ftu/pdf/sk/FTU_2.3.9.pdf)

PACÁKOVÁ, V. a kol. ŠTATISTICKÉ METÓDY PRE EKONÓMOV, 1. vyd. Bratislava: Iura Edition, spol. s r. o., 2009. 441 s. ISBN 978-80-8078-284-9.

RAVALLION, M. Poverty lines in theory and practice [elektronický zdroj]. Washington, DC: World Bank, 1998. 53 s. [cit. 05.02.2022]. ISBN 0821342266. Dostupné na: <https://documents1.worldbank.org/curated/en/916871468766156239/pdf/multi-page.pdf>

SAS SUPPORT, The GLM Procedure [elektronický zdroj]. SAS Institute Inc., SAS Campus Drive, Cary, NC, 2020. [cit. 12.03.2022]. 201 s. Dostupné na: [https://documentation.sas.com/api/docsets/statug/15.2/content/glm.pdf?locale=sk#nameddest=statug\\_glm\\_toc](https://documentation.sas.com/api/docsets/statug/15.2/content/glm.pdf?locale=sk#nameddest=statug_glm_toc)

SAS SUPPORT, SAS/STAT® 13.1 User's Guide [elektronický zdroj]. SAS Institute Inc., SAS Campus Drive, Cary, NC, 2013. [cit. 01.04.2022]. 9497 s. Dostupné na: <https://support.sas.com/documentation/cdl/en/statug/66859/PDF/default/statug.pdf>

SAS SUPPORT, SAS® 9.4 Guide to Software Updates and Product Changes [elektronický zdroj]. SAS Institute Inc., SAS Campus Drive, Cary, NC, 2021. [cit. 01.04.2022]. 448 s. Dostupné na: [https://documentation.sas.com/api/collections/pgmsascdc/9.4\\_3.5/docsets/whatsdiff/content/whatsdiff.pdf?locale=ja#nameddest=n1af4n56n0r8gvn1kt0j8iagcc86](https://documentation.sas.com/api/collections/pgmsascdc/9.4_3.5/docsets/whatsdiff/content/whatsdiff.pdf?locale=ja#nameddest=n1af4n56n0r8gvn1kt0j8iagcc86)

Society at a Glance: OECD Social Indicators [elektronický zdroj]. OECD: Paríž, 2006 [cit. 06.02.2022]. Dostupné na: <https://stats.oecd.org/glossary/detail.asp?ID=7326>

ŠOLTÉS, E. REGRESNÁ A KORELAČNÁ ANALÝZA s aplikáciami v softvéri SAS [elektronický zdroj]. 1. vyd. Bratislava: Letra Edu, s. r.o., 2019. 238 s. [cit. 06.02.2022]. ISBN 978-80-89962-39-6

ŠÚSR. EU SILC 2020: Podiel obyvateľov SR ohrozených príjmovou chudobou sa znížil na 11,4 % [elektronický zdroj]. Ústredie ŠÚ SR, 2021. [cit. 02.02.2022]. Dostupné na: [https://slovak.statistics.sk/wps/portal/ext/aboutus/office.activites/officeNews/vsetkyaktuality/405b6bfa-aa43-4422-829c-e2d26a3db217/lut/p/z1/tVFNU8IwFPwtHjiWvDRpG46BYWiROIJF2lyctBSIpS0fsci\\_N3U8qDMgHszhJXmzb\\_ftLBIoRqKSjVpJrepKbsw\\_Ee7z1AtYv485QHg7gIAE\\_rTPI0yxjebfAewuGkLwyO9H0ZhioA4Sl-efkEAIq\\_RWr1FSpwe5tg](https://slovak.statistics.sk/wps/portal/ext/aboutus/office.activites/officeNews/vsetkyaktuality/405b6bfa-aa43-4422-829c-e2d26a3db217/lut/p/z1/tVFNU8IwFPwtHjiWvDRpG46BYWiROIJF2lyctBSIpS0fsci_N3U8qDMgHszhJXmzb_ftLBIoRqKSjVpJrepKbsw_Ee7z1AtYv485QHg7gIAE_rTPI0yxjebfAewuGkLwyO9H0ZhioA4Sl-efkEAIq_RWr1FSpwe5tg)

6FpaqlJQvdAVNUo\_TJ3K9yYx4daA65Lr42KDipmy6lJSUIFqW2bTG711m5vbBdSRapjb  
1WZZupBUquQs9\_s926gjOHg5kXH5DBiPvUmwCwyciBgPuzqDclBDj5BFzgSMwO3tk  
dfIzmj cqPaFbV-9Ik9fBHiz6gMRIqLbvHrOxCFzPmMgoudhzq4R64bf bqZbcT3ARUVzp\_  
0yj-x4SMnL0PB-HK2JB63SrUKP7BheKruLblrGtKZBXLcEioSE9HfvMOlfXJsg!!/dz/  
d5/L2dBISEvZ0FBIS9nQSEh/

THE WORLD BANK. World Development Report 2000/2001: Attacking Poverty [elektronický zdroj]. New York: Oxford University Press 2000. 356. s. [cit. 05.02.2022]. ISBN 0-19-521129-4. Dostupné na: <https://documents1.worldbank.org/curated/en/230351468332946759/pdf/226840WDR00PUB0ng0poverty0200002001.pdf>

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE. Canberra Group handbook on household income statistics [elektronický zdroj]. 2. vyd. 2011. 208 s. [cit. 03.02.2022]. Dostupné na: [https://unece.org/fileadmin/DAM/stats/publications/Canberra\\_Group\\_Handbook\\_2nd\\_edition.pdf](https://unece.org/fileadmin/DAM/stats/publications/Canberra_Group_Handbook_2nd_edition.pdf)

UN NEWS: UN prioritizing least developed countries in plans, investments and actions. [elektronický zdroj]. 2022. [cit. 04.04.2022]. Dostupné na: <https://news.un.org/en/story/2022/03/1114122>

VLAČUHA, R. – KOVÁČOVÁ Y. EU SILC 2020: Indikátory chudoby a sociálneho vylúčenia [elektronický zdroj]. Bratislava, 2021. 74 s. [cit. 28.01.2022]. ISBN 978-80-8121-828-6.

ŽÁK, M. a kol. Velká ekonomická encyklopedie, 2. vyd. Praha: Linde Praha, a. s., 2002. 886 s. ISBN 9788072013814

## Prílohy

Tabuľka 4.1 Informácie o vstupných premenných z EU-SILC 2020

Názov premenných	Premenné	Katégoria	Označenie kategórií v SAS	Poradie kategórie
Ekonomická aktivita	RB210	Iná neaktívna osoba	Neaktívna_osoba	1
		Nezamestnaný	Nezamestnaná_osoba	2
		Starobný dôchodca, osoba v predčasnom dôchodku	Osoba_na_dochodku	3
		Zamestnaná osoba – referenčná kategória	z_zamestnaná_osoba	4
Typ domácnosti	HT	Jednočlenná domácnosť	1_dospely	1
		Domácnosť jedného rodiča s jedným, alebo viac závislými deťmi	1_dospely_aspon_1_dieta	2
		Domácnosť s 2 dospelými vo veku pod 65 rokov	2_dospeli	3
		Domácnosť dvoch dospelých s jedným závislým dieťaťom	2_dospeli_1_dieta	4
		Domácnosť s 2 dospelými, kde aspoň jeden dospelý je vo veku 65 rokov a viac	2_dospeli_1_dochodca	5
		Domácnosť dvoch dospelých s tromi, alebo viac závislými deťmi	2_dospeli_aspon_3_deti	6
		Zvyšné domácnosti bez závislých detí	ostatné_bezdetne_domácnosti	7
		Zvyšné domácnosti so závislými deťmi	ostatné_domácnosti_s_detmi	8
		Domácnosť dvoch dospelých s dvomi závislými deťmi – referenčná kategória	z_2_dospeli_2_deti	9
		Rodinný stav	PB190	Slobodný/ á
Vdovec/ vdova	vdovec_vdova			2
Ženatý/ vydatá – referenčná kategória	z_osoba_v_manželstve			3
Vzdelanie	PE040	Nižšie ako primárne, primárne a nižšie sekundárne vzdelanie	nižšie_ako_sekundárne	1
		Post-sekundárne	post-sekundárne	2
		Terciárne vzdelanie I. stupňa	vysokoskolske_1_stupen	3
		Vyššie sekundárne vzdelanie	vyššie_sekundárne	4
		Terciárne vzdelanie II. stupňa – referenčná kategória	z_vysokoskolske_2_3_stupen	5
Zdravotný stav	PH010	Priemerný	priemerný	1
		Zlý	zlý	2
		Dobrý – referenčná kategória	zz_dobry	3

<b>Stupeň urbanizácie</b>	DB100	Územie s priemerne hustým osídlením	priemerne_osídlenie	1
		Územie s riedkym osídlením	riedke_osídlenie	2
		Územie s hustým osídlením – <b>referenčná kategória</b>	z_husté_osídlenie	3
<b>Kraj</b>	Kraj	Banskobystrický kraj	BB	1
		Košický kraj	KE	2
		Nitriansky kraj	NR	3
		Prešovský kraj	PO	4
		Trenčiansky kraj	TN	5
		Trnavský kraj	TT	6
		Žilinský kraj	ZA	7
		Bratislavský kraj – <b>referenčná kategória</b>	z_BA	8

Zdroj: vlastné spracovanie, EU-SILC 2020