

Evanjelické lýceum
Vranovská 2, 851 01 Bratislava

Faktory ovplyvňujúce študijné výsledky a identifikácia rizikových skupín

Stredoškolská odborná činnosť

Č. odboru: 17

Bratislava
2024

Riešiteľ
Daniel Svitaň

Ročník štúdia: štvrtý

Evanjelické lýceum
Vranovská 2, 851 01 Bratislava

Faktory ovplyvňujúce študijné výsledky a identifikácia rizikových skupín

Stredoškolská odborná činnosť

Č. odboru: 17

Bratislava

2024

Riešiteľ

Daniel Svitaň

Ročník štúdia: štvrtý

Školiteľka: Martina Šandor

Čestné vyhlásenie

Vyhlasujem, že prácu stredoškolskej odbornej činnosti na tému „Faktory ovplyvňujúce študijné výsledky a identifikácia rizikových skupín“ som vypracoval samostatne, s použitím uvedených literárnych zdrojov. Prácu som neprihlásil a ani neprezentoval v žiadnej inej súťaži, ktorá je pod gestorstvom MŠM VVaŠ SR. Som si vedomý dôsledkov, ak uvedené údaje nie sú pravdivé.

Pod'akovanie

Týmto chcem poďakovať Mgr. Martine Šandor, Ing. Martinovi Berkimu, Ing. Márii Dvorskej, Mgr. Marcelovi Sokolovičovi a Georgie Polymenakou za odborné podklady, informácie a pomoc pri vypracovaní tejto práce.

Obsah

| | |
|---|-----------|
| Úvod | 7 |
| 1 Problematika a prehľad literatúry | 8 |
| 1.1 Spánok | 8 |
| 1.2 Absencia | 8 |
| 1.3 Dochádzanie | 9 |
| 1.4 Životná situácia | 10 |
| 1.5 Práca a profesionálne alebo polo-profesionálne aktivity | 12 |
| 1.6 Socio-ekonomická trieda | 13 |
| 2 Ciele práce | 15 |
| 3 Materiál a metodika | 16 |
| 3.1 Zbieranie dát | 16 |
| 3.1.1 Úvod | 16 |
| 3.1.2 Školské výsledky | 16 |
| 3.1.3 Osobné otázky | 16 |
| 3.2 „Čistenie” (spracovanie) dát | 17 |
| 3.3 Analýza a zostrojenie grafov dát | 18 |
| 3.4 Trénovanie neurónovej siete | 19 |
| 4 Výsledky | 20 |
| 4.1 Distribúcia | 20 |
| 4.2 Pohlavie | 20 |
| 4.3 Socio-ekonomický status | 21 |
| 4.4 Práca alebo aktivita | 21 |
| 4.5 Životná situácia | 21 |
| 4.6 Dochádzanie | 21 |
| 4.7 Spánok | 22 |
| 4.8 Absencia | 22 |
| 4.9 Neurónová sieť | 22 |
| 5 Diskusia | 23 |
| 5.1 Pohlavie | 23 |
| 5.2 Socio-ekonomický status | 23 |
| 5.3 Práca alebo aktivita | 24 |
| 5.4 Životná situácia | 24 |
| 5.5 Dochádzanie | 24 |
| 5.6 Spánok | 24 |
| 5.7 Absencia | 25 |
| 5.8 Neurónová sieť | 25 |
| 6 Závery práce | 26 |
| 7 Zhrnutie | 28 |
| Resume | 29 |

| | |
|-----------------------------------|-----------|
| Zoznam použitej literatúry | 30 |
| Prílohy | 32 |
| Graf 1 | 33 |
| Graf 2 | 34 |
| Graf 3 | 35 |
| Graf 4 | 36 |
| Graf 5 | 37 |
| Graf 6 | 38 |
| Graf 7 | 39 |
| Graf 8 | 40 |
| Graf 9 | 41 |
| Graf 10 | 42 |
| Graf 11 | 43 |
| Graf 12 | 44 |
| Graf 13 | 45 |
| Graf 14 | 46 |
| Graf 15 | 47 |
| Graf 16 | 48 |
| Graf 17 | 49 |
| Graf 18 | 50 |
| Graf 19 | 51 |
| Graf 20 | 52 |
| Tabuľka 1 | 53 |
| Tabuľka 2 | 53 |
| Tabuľka 3 | 53 |
| Tabuľka 4 | 54 |
| Tabuľka 5 | 54 |
| Tabuľka 6 | 55 |
| Tabuľka 7 | 57 |

Zoznam skratiek, značiek a symbolov

Zoznam skratiek

GPA - „grade point average”, mierka školských výsledkov používaná v Spojených Štátoch Amerických, ktorá sa určuje aritmetickým priemerom známok a je zväčša v rozmedzí 1,0 a 4,0, ale môže byť vo výške až 5,0 v špeciálnych prípadoch

SES - „socio-economic status”, mierka socio-ekonomického statusu, ktorá rozdeľuje populáciu na tri skupiny podľa mzdy a bohatstva: nižšia, stredná a vyššia trieda, ďalej aj ako socio-ekonomická trieda

Hviezdičková syntax - zjednodušenie písania plurálu, po slovotvornom základe slova bez meniacej sa prípony nasleduje hviezdička a potom koncovka v ženskom rode plurálu, napríklad „študent*ky” - študenti a študentky, „žia*čky” - žiaci a žiačky, „bez učiteľ*iek” - bez učiteľov a učiteľiek a ak sa k tejto forme pripája prídavné meno alebo zámeno, bude v ženskom rode, tiež je dôležité zmieniť, že túto formu slov sme prebrali z nemčiny a nie je vymyslená nami, po nemecky sa to volá Genderstern a v hovorenej reči je hviezdička znázornená glotálnou zarážkou

RBK (tiež ako rbk) - rank-biseriálna korelácia, tiež ako „veľkosť účinku”

rbk&dm&dM - rank-biseriálna korelácia, rozdiel priemerov a rozdiel mediánov (nasledujúce tri čísla budú reprezentovať tieto tri mierky v tomto poradí)

švv - štatisticky významný vzťah

n-ses - nižšia socio-ekonomická trieda

s-ses - stredná socio-ekonomická trieda

v-ses - vyššia socio-ekonomická trieda

Úvod

V predloženej práci sme sa rozhodli hľadať vzťah medzi rôznymi faktormi a výsledkami v škole. Tieto faktory zahŕňajú dĺžku spánku, dochádzku, dĺžku cesty do školy, životnú situáciu, pracovnú alebo inú časovo náročnú aktivitu a socio-ekonomickú triedu domácnosti. Ako mierku výsledkov sme vybrali priemer známok za minulý polrok a známky z angličtiny, slovenčiny a matematiky za minulý polrok. Dôvodom pre vypracovanie práce je určiť rizikové skupiny, ktoré majú predpoklad na horšie výsledky v škole, aby sa týmto skupinám mohla poskytnúť pomoc.

Cieľom práce je určiť reprezentáciu v každom faktore a určiť vzťah medzi tým faktorom a výsledkami v škole a určiť rizikové skupiny, ktoré disponujú vlastnosťami spojenými s horšími výsledkami.

V časti Problematika a prehľad literatúry poskytujeme teoretický základ pre ďalšiu analýzu. Opisujeme faktory, ktoré môžu ovplyvniť školské výsledky, ako sú spánok, absencia, dochádzanie, životná situácia, pracovné aktivity a socio-ekonomický status. V časti Ciele práce formulujeme hlavný cieľ aj čiastkové ciele, ako napríklad zber dát, ich analýzu a identifikáciu rizikových skupín. V časti Materiál a metodika vysvetľujeme, ako sme zbierali údaje prostredníctvom dotazníkov a ako sme ich následne spracovali. Popisujeme aj metódy analýzy, zostrojenie grafov a implementáciu neurónovej siete, ktorá pomáha identifikovať vzory v dátach. V časti Výsledky prezentujeme distribúciu dát a analyzujeme vzťahy medzi jednotlivými faktormi a školskými výsledkami. Výsledky sú vizualizované v podobe grafov a tabuliek. V časti Diskusia porovnávame získané výsledky s poznatkami z literatúry a analyzujeme ich význam. Venujeme sa jednotlivým faktorom a ich vplyvu na študijné výsledky, pričom hodnotíme aj výkon neurónovej siete. V časti Závery práce sumarizujeme hlavné zistenia a ich význam pre ďalší výskum či potenciálne praktické využitie. V časti Zhrnutie ponúkame stručný prehľad výsledkov a hlavných bodov práce. Nasleduje Resume, ktoré poskytuje krátky prehľad o práci v inom jazyku. Nakoniec uvádzame Zoznam použitej literatúry, ktorý obsahuje citované zdroje, a Prílohy, kde sa nachádzajú podrobné grafy a tabuľky dopĺňajúce analýzu.

1 Problematika a prehľad literatúry

1.1 Spánok

Spánok je dôležitý pre správne fungovanie a kogníciu. Nedostatok spánku môže byť spojený so zhoršeným logickým myslením, slabšou sústredenosťou a únavou, čo môže mať veľké následky na denné aktivity. V škole je potrebné sa sústrediť a dávať pozor, takisto ako používať kritické myslenie a zapamätať si informácie. Nedostatok spánku teda môže mať negatívny vplyv na žiakov a žiačky, čo by sa mohlo odraziť na ich známkach a výsledkoch.

Vzťah medzi spánkom a GPA vysokoškolských študent*iek bol študovaný Williamom E. Kellym, Kathryn E. Kelly a Robertom C. Clantonom [1]. Medzi účastníkmi ich štúdie zapadalo 15,5% ľudí klasifikovaných krátkym spánkom, 72,3% ľudí klasifikovaných stredne dlhým spánkom a 12,2% ľudí klasifikovaných dlhým spánkom s tým, že počet účastníkov bol 148. Žiadny významný vzťah nebol nájdený medzi vekom a GPA. Takisto nebol nájdený žiadny významný vzťah medzi pohlavím a dĺžkou spánku alebo GPA. GPA nahlásilo 76% účastníkov, z čoho iba 61% účastníkov s dlhým spánkom a 78% ostatných účastníkov nahlásilo svoje GPA. Jednosmerná analýza rozptylu ukázala, že účastníci s dlhým spánkom majú priemerné GPA 3,24, čo je významne vyššie ako priemerné GPA účastníkov s krátkym spánkom, čo bolo 2,74. Účastníci so stredne dlhým spánkom neboli významne rozdielni od ostatných a ich priemerné GPA bolo 3,01. Priemery GPA boli určené aritmetickým priemerom. Autor*ky štúdie uznali, že keďže menej účastníkov s dlhým spánkom nahlásilo svoje GPA, ich štúdia môže byť výrazne ovplyvnená zaujatosťou tejto skupiny pri samohlásení, keďže je možné, že účastníci s dlhým spánkom a nižšou GPA sa rozhodli ich GPA nenahlásiť.

Ďalej budeme pokračovať s hypotézou, že medzi spánkom a výsledkami existuje vzťah. Samozrejme, tento vzťah sa môže prezentovať inak, keďže sa naša štúdia neodohráva na vysokej škole. Predpoklad ale je, že žiačky, ktoré spia menej, budú mať horšie výsledky.

1.2 Absencia

Keď hovoríme o dochádzke, hovoríme o tom, koľko hodín žiačka nevymešká. Absencia je teda opak, koľko hodín žiačka vymešká. Samozrejme, prítomnosť v škole môže mať dopad nielen na známky, ale aj na socializáciu žiačky. Môžeme predpokladať, že väčšina žiačiek na

strednej škole alebo gymnáziu sa učia lepšie, keď sú fyzicky prítomní. Absencia teda môže mať negatívny vplyv na výsledky.

Vzťah medzi absenciou a výsledkami bol študovaný Sheniz Moonie, Davidom A. Sterlingom, Larry A. Figgsom a Mario Castrom [2]. Veľkosť vzorky bola 3812, z čoho 50% boli muži. Dáta boli zbierané zo základných a stredných škôl z ročníkov 3, 4, 7, 8, 10 a 11. Výsledky boli zbierané pomocou MAP testu - Missouri Assessment Program (doslova „Missouri hodnotiaci program“). Výsledky boli rozdelené do piatich skupín: menej ako napredujúci, napredujúci, skoro zdatní, zdatní a pokročilí. Do skupiny „menej ako napredujúci“ zapadalo 967 (25%) študent*iek, do skupiny napredujúci 1196 (31%) študent*iek, do skupiny skoro zdatní 1102 (28%) študent*iek, do skupiny zdatní 470 (12%) študent*iek a do skupiny pokročilí 95 (2%) študent*iek. Viacrozmerná lineárna regresia ukázala inverzný vzťah medzi absenciou a výsledkami MAP testu, teda čím menej žiak chýba, tým lepšie boli výsledky na MAP teste.

Ďalšia štúdia od Kristena J. Hancocka, Dávida Lawrence, Carringtona C. J. Shepherda, Francisa Mitroua a Stephena R. Zubricka ukázala, že študent*ky z výhodnejšieho zázemia nie sú „imúnne“ voči tomuto fenoménu [3].

Vďaka obojmu štúdiu vieme, že medzi absenciou a výsledkami existuje vzťah. Môžeme teda pokračovať s hypotézou, že žiačky, ktoré majú lepšiu dochádzku, budú mať aj lepšie výsledky.

1.3 Dochádzanie

Samozrejme, žiačky bývajú v rôznych vzdialenostiach od školy, čo má vplyv na ich dĺžku cesty do školy. Toto ich môže donútiť vstávať skôr pretože im trvá dlhšie dostať sa do školy. Dĺžka cesty môže mať teda dopad na iné faktory, ktoré môžu mať dopad na výsledky. Napríklad, žiačky, ktoré bývajú ďalej od školy, spia menej, pretože musia skôr vstávať, a tým pádom majú horšie známky. Toto je ale len predpoklad a treba ho overiť.

Martijn B.W. Kobus, Jos N. Van Ommeren a Piet Rietveld študovali vzťah medzi dĺžkou cesty na univerzitu, dochádzku a akademickými výsledkami [4]. Dáta boli zbierané na Univerzite VU v Amsterdame. V Holandsku vyžadujú univerzity iba z „druhej školy“ - na Slovensku tento koncept nemáme, ale dá sa to prirovnať ku ročníkom 5-9 základnej školy. Študent*ky majú dobrý prístup k štipendiám a verejná doprava je pre študent*ky zadarmo. Kvôli tomuto a ďalším faktorom sú univerzity v Holandsku na viac-menej tej istej úrovni a dá sa predpokladať, že na

iných Holandských univerzitách, teda štúdia dopadla veľmi podobne kvôli veľmi podobným podmienkam. Potvrdilo sa, že študent*ky, ktoré bývajú ďalej od univerzity dochádzajú na menej dní, ale keď prídu, zostanú dlhšie, tým pádom je iba minimálny alebo žiadny dopad na počet absolvovaných hodín týždenne oproti študent*kám, ktoré bývajú bližšie. Bolo dokázané, že existuje inverzný vzťah medzi dĺžkou cesty a akademickými výsledkami. V tejto štúdii bola metrika akademických výsledkov aritmetický priemer známok. Samozrejme, naša práca sa neodohráva na univerzite, takže si nemôžeme vytvoriť pevný predpoklad, ktorý by táto štúdia podporovala. Otázkou teda zostáva, či existuje vzťah medzi dĺžkou cesty a výsledkami aj na strednej škole alebo či je tento fenomén špecifický univerzitám. Predsa len, vyučovanie funguje inak na univerzitách ako na stredných školách, kde si žiačky nemôžu úplne vyberať predmety (okrem seminárov v posledných rokoch).

Ďalšia štúdia od Erica A. Morrisa a Ying Zhou našla vzťah medzi dĺžkou cesty do práce a výškou platu [5]. Tiež sa našiel vzťah k miere vlastníctva bývania, kde 78% pracovníčiek dochádzajúcich z predmestí vlastní svoj byt alebo dom, ale len 58% pracovníčiek z metropolitnej oblasti vlastní ich bývanie. Tiež sa našiel inverzný vzťah k miere zamestnanosti a ako dôvod sa predpokladá, že bývanie vzdialenejšie od centra je častokrát lacnejšie, tým pádom musí menej členov domácnosti pracovať, ale tento predpoklad nie je podporený dátami. Môžeme teda ďalej pokračovať s hypotézou, že domácnosti študent*iek, ktoré bývajú ďalej od školy zarábajú viac a tým pádom môžu v niektorých prípadoch patriť k vyššej ekonomickej triede. Samozrejme, hypotézu o vzťahu medzi ekonomicou triedou a školskými výsledkami sme si ešte nevytvorili, takže si nevieme vytvoriť nejaký konkrétny podložený predpoklad.

1.4 Životná situácia

V našej práci sa berie životná situácia ako podmienky domácnosti, alebo presnejšie s kým alebo kde býva študent*ka. Samozrejme, životná situácia môže mať veľký vplyv na výsledky, keďže už len priestor a podmienky na učenie nie sú samozrejmé a sú určované životnou situáciou. V našej práci ale nehľadáme vplyvy, ale vzťahy. Životná situácia je komplikovaná a nedá sa jednoducho stlačiť do jedného čísla, takže si nevieme vytvoriť jasný predpoklad.

Na túto tému bolo zhotovených viacero štúdií, vrátane vzťahu medzi životnou situáciou a niektorými škodlivými zvykmi, ako konzumácia alkoholu. Prvá štúdia, na ktorú sa podľa nás oplatí pozrieť, je štúdia o dopade životnej situácie na integračný proces a progres v štúdiu od S.

Beekhovena, U. De jonga a H. Van houta [6]. Dáta boli zbierané v Amsterdame, hlavne na Univerzite Amsterdamu, kde v rokoch 1999 a 2000 sa uskutočnil prieskum študent*iek prvého ročníka denného štúdia. Aby sa zabezpečila rozmanitosť dát z vedeckých, spoločenských a humanitných vied, vzorka zahŕňala 26 rôznych odborov. 40% oslovených študent*iek vyplnilo dotazník, čo bolo 782 študent*iek. Vzorka bola mierne neobjektívna, keďže študent*ky ktoré sa zapojili, mali o niečo lepši priemer známok ako tie, ktoré sa nezapojili. 52% študent*iek na tradičných univerzitách bývalo samostatne. Tí, ktorí bývali samostatne, mali lepšie vzdelaných rodičov. Študent*ky ktoré sa odsťahovali od rodičov boli zväčša v spoločenských vedách a humanitných odboroch a tiež sa cítili menej integrovaní. Študent*ky, ktoré bývajú doma, mali celkovo lepšie podmienky ako tie, ktoré bývali samostatne. Ďalej budeme pokračovať s hypotézou, že tí, ktorí bývajú s rodičmi alebo inými rodinnými príslušníkmi, budú mať lepšie podmienky na štúdium a tým pádom aj lepšie známky.

Posledná štúdia k tejto téme tiež pochádza z Amsterdamu. Cécile RL Boot, Chris BT Rietmeijer, Peter Vonk a Frans J Meijman študovali vzťah medzi zdravím a životnou situáciou, ktorú rozdelili na tri kategórie: bývanie s rodičmi alebo rodinou, bývanie samostatne a bývanie s rovesníkmi [7]. Hypotéza zahŕňala predpoklad, že študent*ky, ktoré bývajú s rodičmi alebo rodinou budú zdravší, keďže majú doma lepšie podmienky. Z 8258 oslovených študent*iek, iba 44% sa zúčastnilo štúdie, čo je 3664. Zo zúčastnených, štvrtina bývala s rovesníčkami, 36% s rodičmi alebo rodinou a 39% samostatne alebo s partner*kou. Študent*ky bývajúce samostatne alebo s partner*kou častejšie uvádzali, že sa cítia zdraví. Títo študent*ky zároveň hlásili väčšiu spotrebu tabaku, alkoholu a marihuany. Tiež uvádzali menej častú účasť na rizikových športoch. Študent*ky bývajúce s rovesníkmi tiež hlásili viac kontaktu so zdravotníčkou ako tie, ktoré bývajú s rodinou. Študent*ky bývajúce s rovesníkmi dosahovali lepšie výsledky v logickom myslení. Ďalej budeme pokračovať s hypotézou, že študent*ky, ktoré nebývajú s rodinou, môžu mať lepšie logické myslenie a tým pádom aj lepšie výsledky v matematike, ale toto tvrdenie zostáva neúplne podložené.

Samozrejme, treba brať na vedomie, že všetci zúčastnení študujú na univerzite a nie na gymnáziu a tým pádom nevieme povedať, či tieto vzťahy platia aj na gymnáziách.

1.5 Práca a profesionálne alebo polo-profesionálne aktivity

Práca alebo profesionálne a poloprofesionálne aktivity sú častým využitím voľného času. Práca je častokrát spojená s potrebou mať peniaze, z akýchkoľvek dôvodov môže byť táto potreba vytvorená. Iné aktivity, napríklad šport alebo hudba, môžu byť vykonávané pre zábavu, ale aj iné dôvody. Samozrejme, tieto činnosti zaberajú veľa času, ale aj častokrát učia žiačky disciplínu a time management. Otázkou teda zostáva, čo má väčší dopad na výsledky. Je to nedostatok voľného času, ktorý môže byť vytvorený týmito činnosťami, alebo je to naopak a vďaka dobrej disciplíne sa môžu žiačky lepšie sústrediť na školu a efektívnejšie riešiť domáce úlohy.

Vzťah medzi prácou a akademickými výsledkami študovali Lauren Dundes a Jeff Marx na McDaniel Univerzite v Spojených Štátoch [8]. V hypotéze spomínajú ďalšie štúdiu, ktorá zistila, že 74% študentiek bakalárskeho štúdia pracujú priemerne 25,5 hodín týždenne. 85% oslovených vyplnilo dotazník, čo bolo 256. 96% zúčastnených študujú bakalárske štúdium, 98% má denné štúdium, 80% boli bieli a 57% boli ženy. Nebol rozdiel v rasovom rozložení študentiek na danej univerzite a účastníčiek prieskumu. 64% študentiek pracovali, pri čom 62% mužov a 68% žien pracovali. Tiež pracovali viac študentky vo vyšších ročníkoch, v prvom pracovalo 56%, v druhom 58%, v treťom 66% a v štvrtom 76%. Nebol nijaký významný rozdiel medzi odbormi. Výsledky študentiek sa rozdeľovali na 3 skupiny pomocou GPA: 31% malo GPA 3,5 a vyššie, 36% malo GPA 3,0 až 3,4 a 33% malo GPA medzi 2,0 až 2,9. Na dokončenie bakalárskeho štúdia si treba udržiavať GPA 2,0 a vyššie a preto posledná skupina nebola od 0, ale od 2,0, aj keď GPA môže byť medzi 0 a 4. Reprezentácia pohlavia v najvyššej skupine bola 15% mužov a 43% žien. Hodiny práce boli rozdelené na tri skupiny: málo (pod 10 hodín týždenne), stredne (10 až 19 hodín týždenne) a veľa (20 a viac hodín týždenne). Tí, ktorí pracujú stredne veľa, mali viditeľne lepšie výsledky ako ostatní. V najlepšej skupine podľa GPA patrilo 37% stredne pracujúcich, pričom iba 29% málo pracujúcich a 26% veľa pracujúcich. V strednej skupine podľa GPA patrilo 44% stredne pracujúcich, pričom iba 29% málo pracujúcich a 35% veľa pracujúcich. Toto ukazuje, že tí, ktorí pracujú od 10 po 19 hodín týždenne majú najlepšie akademické výsledky. Ďalej budeme teda predpokladať, že pracujúci 10 a viac hodín týždenne budú mať lepšie výsledky ako tí, ktorí pracujú menej alebo nepracujú. Tiež treba overiť rozloženie výsledkov podľa pohlavia.

Joy Gaston Gayles a Shouping Hu vykonali ďalšiu štúdiu, v ktorej pozorovali vzťah medzi športovou účasťou a akademickými výsledkami [9]. Ako mierku akademických výsledkov použili hodnotenie „Progress in College”. Záver bol, že rôzne typy športu majú rôzne priemery akademických výsledkov, a teda nemožno povedať, že existuje vzťah medzi športovaním celkovo a výsledkami v škole. Samozrejme, táto štúdia robila prieskum na americkej univerzite, kde športovci dostávajú vysoké štipendiá, čo je niekedy jediný spôsob ako si niektoré študent*ky môžu dovoliť cenu univerzity. Nemôžeme si však vytvoriť žiadny predpoklad, pretože šport na univerzitách v Spojených štátoch je úplne inak chápaný. Treba zistiť, či existuje vzťah medzi športovaním a výsledkami v škole a či je to kvôli nejakej vlastnosti špecifickej športu alebo či všetky profesionálne alebo polo-profesionálne činnosti majú podobný efekt.

1.6 Socio-ekonomická trieda

Socio-ekonomická trieda a teda ekonomické podmienky, v ktorých žiačka žije a pripravuje sa do školy, môžu mať veľký dopad na výsledky. Samozrejme, socio-ekonomická trieda môže byť úzko spojená so životnou situáciou a prácou, čo môže mať na výsledky zosilnený vplyv. Vplyv sa ale nesnažíme nájsť alebo dokázať, hľadáme iba vzťahy.

Efekty socio-ekonomickej triedy na akademické výsledky už boli podrobne skúmané viacerými inštitúciami. Prvá štúdia, ktorú by sme chceli predstaviť, je štúdia od Farkhanda Ahmara a Ehteshama Anwara [10]. Z päť rôznych univerzít z Lucknowu, India bolo vybraných 200 študent*iek, z čoho 102 boli muži a 98 boli ženy vo veku od 15 do 19. Ako mierka socio-ekonomickej triedy sa použila mierka SES vytvorená R. L. Bharaadwajom. Aritmetický priemer akademických výsledkov pre mužov je 59,4216 a pre ženy 58,6531 a teda neexistuje významný rozdiel. Tiež nie je významný rozdiel v akademických výsledkoch medzi mužmi a ženami vyššej vrstvy, kde priemer mužov je 72,4419 a pre ženy 72,4324. V nižšej vrstve, ženy majú priemer 50,3462 a muži iba 47,9459, čo nie je významný rozdiel. Zistilo sa teda, že medzi pohlaviami nie je významný rozdiel, ale medzi ekonomickými triedami je. Budeme teda pokračovať s hypotézou, že čím vyššia ekonomická trieda, tým lepšie budú výsledky. Je ale tiež potrebné brať do úvahy ekonomickú situáciu krajín, keďže rozdiely medzi ekonomickými triedami v Indii a na Slovensku môžu byť rôzne a tým pádom by tento výsledok nemusel platiť alebo by sa ukazoval inak.

Ďalšia práca od Jennifer Barry študuje túto problematiku v Spojených štátoch [11]. Bolo nazbieraných 7976 vzoriek. Dáta boli rozdelené na 3 skupiny podľa SES. Nižšia trieda zodpovedala najnižšiemu 33. percentilu a zahŕňala 2263 vzoriek. Stredná trieda zodpovedala strednému 33. percentilu a zahŕňala 2695 vzoriek. Vyššia trieda zodpovedala najvyššiemu 33. percentilu a zahŕňala 2838 vzoriek. Celkový priemer akademických výsledkov bol 52,39, pričom nižšia trieda mala priemer 47,23, stredná trieda 51,85 a vyššia trieda 57. Tiež treba podotknúť, že v nižšej triede bolo viac menšín, okolo 42,84% tejto triedy vytvárali menšiny, ale iba 24,94% strednej triedy a 15,70% vyššej triedy vytvárali menšiny. Tiež bolo v nižšej triede viac študent*iek so slabými akademickými výsledkami, okolo 31,69%, ale iba 24,66% strednej triedy a 20,26% vyššej triedy. Bolo zistené, že študent*ky nižšej triedy trávili viac času na sociálnych sieťach ako iné triedy. Študent*ky z nižšej triedy mali vyššiu reprezentáciu ľudí s angličtinou ako druhým jazykom (26,10%) oproti strednej (13,29%) a vyššej triede (10,47%). Tak isto mala nižšia trieda väčšie zastúpenie ľudí s postihnutím (6,96%) oproti vyššej triede (4,15%). Študent*ky z vyššej triedy tiež trávili viac času robením úloh (vyššia = 11,72, stredná = 10,09, nižšia = 9,72). Tiež majú viac vyznamenaní (vyššia = 71,93%, stredná = 62,72%, nižšia = 54,34%), trávia viac času na mimoškolských činnostiach (vyššia = 6,35, stredná = 5,03, nižšia = 3,54) a viac plánujú pokračovať vo vzdelávaní (vyššia = 95%, stredná = 85%, nižšia = 75%). Nižšia trieda má tiež viac rovesníkov, ktorí predčasne ukončili školskú dochádzku (nižšia = 1,32, stredná = 1,21, vyššia = 1,13). Táto štúdia spomína ešte viac výhod vyššej triedy, aj v časti akademie. Je teda zrejmé, že vyššia trieda má oveľa lepšie podmienky na štúdium a život, a preto budeme pokračovať s predpokladom, že študent*ky z vyššej triedy budú mať lepšie výsledky v škole oproti strednej a nižšej triede.

2 Ciele práce

Cieľom práce je nájsť vzťah medzi týmito faktormi a výsledkami.

Pri tvorbe práce sme si stanovili nasledovné čiastkové ciele:

1. Nazbierať literatúru na vzťahy medzi faktormi a výsledkami.
2. Vytvoriť anonymný dotazník založený na nazbieranej literatúre.
3. Nazbierať dáta cez osobné návštevy rôznych tried.
4. Pripraviť program na analýzu dát.
5. Analyzovať dáta pomocou programu.
6. Vyhotoviť grafy pomocou grafického programu.
7. Pripraviť program na tréovanie neurónovej siete.
8. Tréovať neurónovú sieť na dátach.
9. Vymedziť rizikové skupiny na základe výsledkov.

3 Materiál a metodika

3.1 Zbieranie dát

Na zber dát sme vytvorili anonymný formulár cez Google Forms, ktorý bol rozposlaný študent*kám školy Evanjelického lýcea v Bratislave. Povolenie spracovávať údaje z dotazníka schválilo vedenie školy a samozrejme v rámci regulácie GDPR. Sekretariát následne poskytol zoznamy s priemernými známami z minulého polroku (2. polrok 2023/2024), známami zo slovenčiny, angličtiny a matematiky a počtom vymeškaných hodín. K týmto zoznamom sme ale nemali prístup a vždy s nimi narábala iba určená učiteľ. Následne sme v priebehu dvoch týždňov osobne navštívili každú triedu od druhého až po piaty ročník spolu s určeným učiteľom prideleným v daný deň, ktorý mal zoznamy a študent*kám poskytoval ich osobné údaje, ktoré si väčšina nepamätala.

Nižšie preberieme presnú formu dotazníka a kategorizáciu odpovedí.

3.1.1 Úvod

Študent*ka bola oboznámená so zámerom práce a s narábaním osobných dát. Otázky v tejto časti sa venovali len ročníku (1 až 5, ale dáta neboli zbierané u prvého ročníku) a pohlaviu (žena alebo muž).

3.1.2 Školské výsledky

Táto časť obsahovala štyri otázky, všetky na výsledky v škole z minulého školského polroku (2. polrok 2023/2024). Prvá otázka sa pýtala na priemer známok, ktorý bol uvedený ako reálne číslo medzi 1 až 5. Ďalšie tri sa pýtali na známku zo slovenčiny, angličtiny a matematiky, ktorá bola uvedená ako celé číslo medzi 1 až 5.

3.1.3 Osobné otázky

V tejto časti sme sa pýtali na faktory, ktoré chceme analyzovať.

Prvá otázka sa týkala socio-ekonomickej triedy a odpoveď bola jedna z troch nasledujúcich možností: nižšia trieda, stredná trieda a vyššia trieda. V literatúre sme prebrali prečo rozdeľujeme populáciu na tieto tri triedy a presné hranice týchto tried sme určili pripočítaním a odpočítaním 40% priemerného platu k priemernému platu.

Druhá otázka bola na výkon práce alebo nejakej profesionálnej, alebo polo-profesionálnej aktivity. Možnosti zahŕňali nasledujúce: pracujem 10 hodín a viac týždenne, pracujem menej ako 10 hodín týždenne, športujem na profesionálnej alebo polo-profesionálnej úrovni, robím muziku na profesionálnej alebo polo-profesionálnej úrovni, robím inú profesionálnu alebo polo-profesionálnu aktivitu a nerobím žiadnu takúto aktivitu. Dôvody prečo sledujeme prácu a šport boli rozobraté v prehľade literatúry. Ostatné kategórie slúžia iba na porovnávanie športu s inou časovo podobne náročnou aktivitou, aby sme zistili či má šport nejakú špeciálnu vlastnosť, alebo či vzťah platí aj pre ostatné aktivity.

Tretia otázka sa dotýkala životnej situácie. Možné odpovede zahŕňajú nasledujúce: bývam s rodičmi, bývam s iným rodinným príslušníčkou, bývam sám*a alebo so spolubývajúc*ou, bývam na internáte a mám to inak. Samozrejme, prvé dve kategórie budú pred analýzou spojené do jednej a rozdelili sme ich len preto, aby sme zistili presnejšiu distribúciu respondent*iek. Posledné dve kategórie sú problematické a podľa počtu študent*iek v týchto kategóriách budú buď vynechané alebo zapojené do analýzy.

Štvrtá otázka bola na priemernú dĺžku trvania cesty do školy, odpovede zahŕňali: bývam na internáte, menej ako 15 minút, menej ako 30 minút, menej ako hodinu a viac ako hodinu. Kategórie boli takto vybrané pre lepšie pochopenie distribúcie respondent*iek, aj keď toto rozdelenie nie je podporené žiadnou literatúrou. Prvá kategória môže byť problematická a podľa počtu študent*iek v tejto kategórii môže byť pri analýze vynechaná.

Piata otázka sa týkala priemernej dĺžky spánku a možné odpovede boli nasledujúce: 9 hodín a viac, 7 až 9 hodín a 6 hodín a menej. Toto rozdelenie je presne zobrahané z literatúry.

Šiesta a posledná otázka sa pýtala na počet vymeškaných hodín za minulý polrok. Odpoveď bola v tomto prípade nulové alebo pozitívne celé číslo.

3.2 „Čistenie” (spracovanie) dát

Keďže dáta zapisovali účastníčky štúdie, sú „špinavé” a vyžadujú čistenie. Napríklad namiesto „1,5” mohlo byť zadané „asi 1,5”, čo software nevie spracovať. Dáta sú uložené v CSV a načítané pomocou Pandas, pričom každý vstup sa testuje, či je číslo. Nečíselné hodnoty sa manuálne opravujú. Kategorické odpovede sa prevádzajú na číselné hodnoty (enum). Napr. „žena” → 0, „muž” → 1; „6 hodín a menej” → 0, „7 až 9 hodín” → 1, „9 hodín a viac” → 2. Číselné premenné ako priemer známok či absencia sa ponechávajú, ale musia byť očistené.

Vyčistené dáta sa ukladajú do 2D matice vo formáte NPY (napr. 50×12, ak je 50 odpovedí). Tento formát umožňuje rýchle načítanie a efektívne spracovanie, vrátane použitia v neurónových sieťach. Každý riadok obsahuje výlučne čísla, napr. „5.0, 0.0, 1.81, 3.0, 2.0, 1.0, 0.0, 1.0, 2.0, 1.0, 1.0, 101.0”, kde desatinné čísla používajú bodku. Poradie stĺpcov je: ročník (1-5), pohlavie (0-1), priemer známok (1-5), známky z matematiky, slovenčiny a angličtiny (1-5), socio-ekonomický status (0-2), práca/aktivita (0-5), životná situácia (0-4), dochádzanie (0-4), spánok (0-2) a absencia (0+). Napr. „5.0, 0.0, 1.81, 3.0, ...” znamená piateho ročníka, ženu, priemer 1,81, známku z matematiky 3, atď.

3.3 Analýza a zostrojenie grafov dát

Keď už máme „očistené” dáta, môžeme začať s analýzou a zostrojením grafov. Na načítanie a prácu s dátami budeme používať knižnicu Numpy, čo je skratka Numerical Python. Štatistické testy budeme robiť pomocou knižnice Scipy, čo je skratka Scientific Python. Grafy budeme zostrojovať pomocou knižnice Matplotlib. Pri každej kategorickej premennej si najprv urobíme test normality, ak ho väčšina kategórií spĺňa, použijeme ANOVU, ak nie, použijeme Kruskal-Wallis test, čo je neparametrická verzia ANOVY. Jediná nezávislá premenná, ktorá nie je kategorická, je absencia. Pre túto premennú budeme používať Kendallov Tau test typu b. Ak nám pri ANOVE výjde štatistický významný výsledok, použijeme Tukeyho HSD post hoc test a pri Kruskal-Wallis teste použijeme Dunnov post hoc test a rozdiely medzi skupinami budeme pozorovať pomocou rank-biserálnej koreácie (rbk), rozdielu medzi priermi a rozdielu medzi mediánmi.

Na zostrojenie grafov používame knižnicu Matplotlib. Pre každú kategorickú nezávislú premennú (okrem ročníka) vytvárame „violin grafy“ v štyroch podgrafoch, pričom každý podgraf zobrazuje jednu závislú premennú: priemer známok a známky z matematiky, slovenčiny a angličtiny. Napríklad pri vzťahu medzi pohlavím a výsledkami budeme mať štyri podgrafy (2x2 mriežka), každý s dvoma violin grafmi – pre ženy a mužov, spolu 8 grafov. Vzťah absencie a priemeru známok znázorňujeme „scatter grafom“, kde os x reprezentuje počet vymeškaných hodín a os y priemer známok. Ostatné vzťahy absencie zobrazujeme „box grafmi“. Distribúciu znázorňujeme prevažne „pie grafmi“, napríklad pre rozdelenie odpovedí podľa ročníkov. Pre premenné priemer známok a absenciu používame histogramy. Všetky grafy sú zahrnuté v prílohe a vo výsledkoch na ne odkazujeme.

3.4 Trénovanie neurónovej siete

Neurónovú sieť budeme trénovať pomocou ďalšieho Python programu a knižnice PyTorch. „Čisté” dáta uložené vo formáte NPY načítame pomocou knižnice NumPy. Treba si pamätať, že čisté dáta sú uložené ako 2D matica čísiel, keďže neurónová sieť sa dá trénovať iba cez homogénne matice a vektory. Vstupné dáta (x) pre trénovanie, teda nezávislé premenné, sú faktory, pohlavie a ročník. Výstupné dáta (y) pre trénovanie, teda závislé premenné, sú výsledky. Dáta rozdelíme na trénovacie (80% dát) a testovacie (20% dát). Používame veľkosť dávky 32 a rýchlosť učenia 0,001, pri čom trénujeme model cez 200 epoch. Štruktúra modelu je znázornená na obrázku 1. Aby sme zistili, aký dobrý je model, používame knižnicu scikit-learn. Toto budeme zisťovať a merať pomocou štyroch metrík: presnosť, precíznosť, odvolanie a skóre f1 - po anglicky accuracy, precision, recall, and f1 score.

4 Výsledky

4.1 Distribúcia

Na formulár odpovedalo 253 študent*iek. Prvý ročník nebol oslovený, druhý ročník predstavoval 27,66% vzorky, tretí ročník 24,50%, štvrtý ročník 24,11% a piaty ročník 23,71% (v Prílohe Graf 1). Na vedomie treba brať, že percentá boli zaokrúhlené na dve desatinné čísla a preto sa nemusia sčítavať na 100%, v tomto prípade je to iba 99,98%, ale znova, je to iba kvôli zaokrúhľovaniu a výsledok je tak či tak veľmi blízko 100%. Ženy predstavovali 62,45% a muži 37,54% (v Prílohe Graf 2). Histogram priemeru známok je zobrazený v Prílohe Graf 3. Pri známkach z matematiky, jednotky predstavovali 28,45%, dvojky 40,31%, trojky 25,29% a štvorky 5,92% (v Prílohe Graf 4). Pri známkach zo slovenčiny, jednotky predstavovali 40,31%, dvojky 45,45%, trojky 13,43%, štvorky 0,39% a päťky 0,39% (v Prílohe Graf 5). Pri známkach z angličtiny, jednotky predstavovali 63,24%, dvojky 29,64% a trojky 7,11% (v Prílohe Graf 6). Nižšia socio-ekonomická trieda predstavovala 10,27%, stredná 55,33% a vyššia 34,38% (v Prílohe Graf 7). Študent*ky pracujúce 10 a viac hodín týždenne predstavovali 16,99%, študent*ky pracujúce menej ako 10 hodín týždenne 13,83%, športujúce študent*ky 18,97%, študent*ky robiace hudbu 5,53%, študent*ky robiace inú aktivitu 9,09% a študent*ky, ktoré nerobia žiadnu aktivitu 35,57% (v Prílohe Graf 8). Študent*ky žijúce s rodinou predstavovali 83,00%, študent*ky žijúce s rodinnou príslušníčkou 2,76%, študent*ky žijúce sami alebo so spolubývajúc*ou 3,16%, študent*ky bývajúce na internáte 8,69% a iné životné situácie 2,37% (v Prílohe Graf 9). Pre dochádzanie, študent*ky z intráku predstavovali 9,09% a dĺžka dochádzky menšia ako 15 minút 11,46%, menšia ako 30 minút 22,52%, menšia ako hodinu 40,71% a viac ako hodinu 16,20% (v Prílohe Graf 10). Krátky spánok predstavoval 39,92%, stredne dlhý spánok 58,89% a dlhý spánok 1,18% (v Prílohe Graf 11). Histogram absencie je zobrazený v Prílohe Graf 12.

4.2 Pohlavie

Dáta nespĺňali normalitu, použité boli neparametrické testy. Grafy pre pohlavie sú zobrazené v Prílohe Graf 13. Pri priemere známok (švv, $p = 0,0017$) bol rozdiel medzi ženami a mužmi podľa rbk&dm&dM -0,0153, -0,1732 a -0,2950 ($p = 0,0016$, tabuľka 1). Pri známke zo slovenčiny (švv, $p < 0,0001$) bol rozdiel medzi ženami a mužmi podľa rbk&dm&dM -0,0188,

-0,3650 a 0 ($p < 0,0001$, tabuľka 2). Pri známke z angličtiny (švv, $p < 0,0039$) bol rozdiel medzi ženami a mužmi podľa rbk&dm&dM -0,0119, -0,2414 a 0 ($p = 0,0037$, tabuľka 3). Pre matematiku sa nenašiel švv.

4.3 Socio-ekonomický status

Dáta nespĺňali normalitu, použité boli neparametrické testy. Grafy pre socio-ekonomický status sú zobrazené v Prílohe Graf 14. Pre priemer známok a známku z angličtiny sa nenašiel švv. Pri známke z matematiky (švv, $p = 0,0005$) bol rozdiel medzi n-ses a s-ses podľa rbk&dm&dM 0,0658, 0,6423 a 1 ($p = 0,0014$, tabuľka 4), medzi n-ses a v-ses podľa rbk&dm&dM 0,1179, 0,7268 a 1 ($p = 0,0003$, tabuľka 4) a medzi s-ses a v-ses nebol štatisticky významný rozdiel ($p = 1,000$, tabuľka 4). Pri známke zo slovenčiny (švv, $p = 0,0357$) bol jediný štatisticky významný rozdiel medzi n-ses a s-ses, ktorý bol podľa rbk&dm&dM 0,0440, 0,4198 a 0 ($p = 0,0497$, tabuľka 5).

4.4 Práca alebo aktivita

Dáta nespĺňali normalitu, použité boli neparametrické testy. Grafy pre prácu a aktivitu sú zobrazené v Prílohe Graf 15. Pre priemer známok a známky a matematiky a slovenčiny sa nenašiel švv. Pre známku z angličtiny sa švv našiel, ale nebol nájdený jasný štatisticky významný rozdiel medzi skupinami ($p > 0,05$, tabuľka 6).

4.5 Životná situácia

Dáta nespĺňali normalitu, použité boli neparametrické testy. Grafy pre životnú situáciu sú zobrazené v Prílohe Graf 16. Nebol nájdený žiadny štatisticky významný vzťah medzi životnou situáciou a výsledkami ($p > 0,05$).

4.6 Dochádzanie

Dáta nespĺňali normalitu, použité boli neparametrické testy. Grafy pre životnú situáciu sú zobrazené v Prílohe Graf 17. Nebol nájdený žiadny štatisticky významný vzťah medzi dĺžkou dochádzania a výsledkami ($p > 0,05$).

4.7 Spánok

Dáta nespĺňali normalitu, použité boli neparametrické testy. Grafy pre spánok sú zobrazené v Prílohe Graf 18. Pre priemer známok a známku zo slovenčiny a angličtiny sa nenašiel švv. Pri známke z matematiky (švv, $p = 0,0048$) bol rozdiel medzi študent*kami, ktoré spia krátko, a študent*kami, ktoré spia stredne-dlho, podľa rbk&dm&dM 0,0128, 0,3242 a 0 ($p = 0,0046$, tabuľka 7). Tretia skupina, teda študent*ky, ktoré spia dlho, musela byť vyradená kvôli nedostatočnej veľkosti pre analýzu (3).

4.8 Absencia

Pre analýzu sme používali Kendallov Tau test, ktorý nepredpokladá normalitu. Grafy pre absenciu sú zobrazené v Prílohe Graf 19. Okrem známky z angličtiny sa pre ostatné výsledky našiel švv. Pri priemere známok sa našiel švv o sile $\tau = 0,2193$ ($p < 0,0001$). Pri známke z matematiky sa našiel švv o sile $\tau = 0,1958$ ($p < 0,0001$). Pri známke zo slovenčiny sa našiel švv o sile $\tau = 0,2124$ ($p < 0,0001$).

4.9 Neurónová sieť

Neurónová sieť bola natrénovaná a štruktúru je možné vidieť na obrázku 1. Propagáciu tréovania, čo sa týka straty tréovania a straty testovania, je možné vidieť v Prílohe Graf 20. Priemerná strata tréovania bola 1,0442 a strata testovania 0,9945. V poslednej epoche dosiahla strata tréovania 0,7700 a strata testovania 0,7330. Natrénovaný model dosiahol nasledujúce metriky: presnosť 0,7059, precíznosť 0,6176, odvolanie 0,7059 a skóre f1 0,5995.

5 Diskusia

Pred pokračovaním sa treba rýchlo vyjadriť k rozdielu medzi počtami žia*čiek z internátu pod otázkou o životnej situácii (8,69% - 22) a dochádzaní (9,09% - 23). Tento rozdiel nastal kvôli ľudskej chybe, vyššie sme presne popisovali „čistotu“ a spoľahlivosť dát a tento rozdiel je jedným z viacerých príznakov nečistých dát zbieraných pomocou dotazníku.

5.1 Pohlavie

Našli sme významný rozdiel medzi ženami a mužmi, kde okrem matematiky mali ženy lepšie výsledky ako muži. Toto viac menej zodpovedá stereotypu, že ženy sa lepšie pripravujú do školy lebo im na škole viac záleží. Naše zistenia ale nezodpovedajú predošlej štúdii [10], kde nebol medzi mužmi a ženami žiadny významný rozdiel. Samozrejme, predošlá štúdia sa odohrávala v Indii, kde je iná kultúra a školstvo tiež funguje inak. Tento rozdiel teda mohol vzniknúť kvôli rôznym podmienkam v každej krajine, napríklad sociálny dôraz na výsledky v Indii, na rozdiel od Slovenska, kde tento dôraz nemusí byť taký silný, čo môže vytvoriť rozdiel medzi ženami a mužmi, keďže študent*ky nie sú tlačené k dobrým výsledkom a teda výsledky podľa pohlavia sa líšia kvôli rôznym iným faktorom a rozdielom medzi pohlaviami. Tiež môže byť rozdiel medzi príležitosťami alebo očakávaniami. Takisto môže mať dopad aj školský systém.

5.2 Socio-ekonomický status

Rozdiely medzi socio-ekonomickými triedami sme našli iba pri známke z matematiky a slovenčiny. V prvom prípade bol jasný rozdiel medzi nižšou triedou a ostatnými, s tým, že stredná a vyššia trieda sa skoro vôbec nelíšili. V druhom prípade bol jasný rozdiel medzi nižšou a strednou triedou, ale nejasný rozdiel medzi nižšou a vyššou a strednou a vyššou triedou. V oboch prípadoch bola ale nižšia trieda znevýhodnená. Tieto výsledky z časti zodpovedajú predošlým štúdiám, až na to, že nebol jasný vzťah pre priemer alebo známku z angličtiny a zo slovenčiny nebol jasný rozdiel medzi strednou a vyššou triedou. Samozrejme, záver v oboch prípadoch je, že nižšia trieda je znevýhodnená. Na toto môže mať vplyv viacero faktorov, vrátane nedostatku materiálnych potrieb pre učenie, ako aj nižšia kvalita života spôsobená nedostatkom financií.

5.3 Práca alebo aktivita

Jediný významný rozdiel, ktorý sme našli, bol v známke z angličtiny, ale nebol jasný rozdiel medzi samostatnými skupinami. Nemôžeme teda porovnávať s predošlými štúdiami, keďže nevieme presné rozdiely medzi skupinami. Samozrejme, kultúra a ekonomika tiež môžu mať vplyv na tento faktor, napríklad práca v jednej krajine môže pomáhať školským výsledkom, keďže sa študent*ka učí disciplíne, ale v niektorých ubližovať, keďže pracovné podmienky a očakávania viac študent*ke škodia. Šport môže byť tiež menlivý faktor, keďže v niektorých kultúrach je na šport veľký dôraz a v niektorých nie. Tiež môžu v iných oblastiach prevládať iné športy s inými požiadavkami a vplyvmi.

5.4 Životná situácia

Pri životnej situácii sme žiadny vzťah nenašli. Znova môže hrať rolu krajina štúdie a lokálna kultúra ako aj to, že sa predošlé štúdie odohrávali na vysokej škole, kde je konzumácia alkoholu normalizovaná a legálna. Naša distribúcia tiež nebola rozmanitá ako sme predpokladali, väčšina študent*iek bývala s rodičmi. Tento faktor nám hovorí teda málo a bolo by ho treba viac študovať.

5.5 Dochádzanie

Dochádzanie tiež neukázalo žiadny významný vzťah. Predošlé štúdie sa ale odohrali na univerzite a teda nemôžeme úplne porovnávať výsledky. Dve pätiny študent*iek majú ale dĺžku dochádzania medzi 30 minútami a hodinou, čo ukazuje, že veľa študent*iek dlho dochádza, čo môže hovoriť o vysokej kvalite školy, keďže pre ňu sú ochotné študent*ky dochádzať aj z ďaleka.

5.6 Spánok

Vzhľadom na to, že sme nemali dosť vzoriek pre každú kategóriu, museli sme kategóriu dlhého spánku vyradiť z analýzy. Toto poukazuje na veľký problém spolu s faktom, že skoro dve pätiny študent*iek spia šesť hodín a menej a teda študent*ky spia málo. Toto môže byť spôsobené počtom domácich úloh a učenia, teda školou ovplyvniteľnými faktormi. Významné rozdiely sme ale našli iba pri známke z matematiky, kde bolo málo spánku spojené s horšími

známkami. Toto je z časti v súlade s predošlými štúdiami, ktoré naznačovali, že menej spánku je spojené s horšími výsledkami. My sme tento fenomén pozorovali iba pri známke z matematiky. Kvôli nedostatku dát v kategórií dlhý spánok, ale nevieme úplne porovnávať a teda treba presnejšiu analýzu s lepšími dátami.

5.7 Absencia

Významný rozdiel sme našli v priemere známok a známke z matematiky a slovenčiny. Všetky mali pozitívne tau, čo znamená, že čím viac bola žiačka neprítomná, tým vyššiu známku mala, teda čím rastie absencia, tým rastie známka. Tento výsledok je celkom intuitívny a tiež zapadá do stereotypu, kde žiačky ktoré chýbajú viac, majú horšie známky. Naše výsledky teda aj súhlasia s predošlou štúdiou, samozrejme, okrem známky z angličtiny.

5.8 Neurónová sieť

F1 skóre nám vyšlo 0,6, čo znamená úplne zlý model, teda, nepodarilo sa nám natrénovať presný a spoľahlivý model. Toto je pravdepodobne spôsobené dvoma faktami. Za prvé, nedostatok dát, teda nebol dostatok vzoriek. Za druhé, nedostatočná kvalita dát, keďže dáta zadávali študent*ky, naše dáta nie sú úplne spoľahlivé, čo pri analýze nie je problém, ale pri trénovaní neurónovej siete je, keďže model je citlivejší a potrebuje spoľahlivejšie dáta. Pre lepší model by bolo teda treba nazbierať oveľa viac dát, ideálne desiatky alebo stovky tisíc vzoriek a tiež zabezpečiť lepšiu kvalitu a spoľahlivosť dát.

6 Závěry práce

Úlohou tejto práce bolo určiť rizikové skupiny alebo kategórie, v ktorých môžu mať žiačky problémy so známami, aby sa im mohla ponúknuť ďalšia pomoc. Dáta sme zbierali prostredníctvom dotazníka, čo mohlo viesť k menším chybám v dátach. Preto nemusia byť dáta úplne spoľahlivé a objektívne. Lepšie dáta by sa mohli zbierať výskumníčkami namiesto dotazníka, keďže môžu zaistiť ich vyššiu spoľahlivosť. Jedna z rizikových skupín boli muži, ktorí mali horšie známky ako ženy. Toto poukazuje na významný rozdiel medzi pohlaviami, ktorý treba riešiť. Ďalšia riziková skupina sú študentky z nižšej socio-ekonomickej triedy. Toto opäť poukazuje na rozdiel medzi triedami, predovšetkým medzi nižšou triedou a ostatnými. Problém nerovnosti medzi pohlaviami a triedami treba riešiť, aby sa zaistila rovnosť a inklúzia. Ďalšie dve rizikové skupiny sú ovplyvniteľné a prvá z nich sú študentky, ktoré spia šesť hodín a menej. Toto môže byť dôsledok množstva školskej práce, ako aj iných faktorov. Samozrejme, aj tento problém treba riešiť a treba zaistiť aby študentky mali dostatok spánku. Ak škola a školské povinnosti prispievajú k tomuto nedostatku spánku, treba zvážiť ich počet a dôraz na ne, ako aj ich implementáciu a dôležitosť. Najprv však treba vykonať ďalší prieskum a zistiť, čo konkrétne prispieva k tomuto nedostatku spánku. Až potom sa môže navrhnúť riešenie. Druhá ovplyvniteľná skupina zahŕňa študentky, ktoré často chýbajú, teda, sú často absenčné. Toto môže byť spôsobené rôznymi inými príčinami, napríklad fyzické zdravie a mimoškolské aktivity, ale aj duševné zdravie a kriminalita. Tento problém je komplikovanejší ako predošlý, pričom riešenia sú rôznorodé a často zložité, keďže niekedy je potrebné zasahovať do osobného života žiačky. Tiež je zrejmé, že absencia môže byť veľmi individuálna a preto nemusí existovať jedno všeobecné riešenie a treba sa zamerať na individuálny prístup, čo prispieva k zložitosti, ako aj zaberá viac zdrojov na rozdiel od riešenia na predošlé problémy, ktoré sa častokrát dajú riešiť plošne a jedno všeobecné riešenie pomôže väčšine. Vytýčili sme štyri rizikové skupiny, ktoré môžu mať problémy v škole a dosahovať horšie výsledky. Prvé dve skupiny sú založené na neovplyvniteľnom faktore, teda pohlavie a socio-ekonomická vrstva. Druhé dve skupiny sú založené na ovplyvniteľnom faktore, teda spánok a absencia. V týchto druhých dvoch skupinách treba tiež dávať dôraz na obmedzenie rizikového faktoru a limitovanie počtu študentiek v rizikovej skupine, teda napríklad zabezpečenie lepšieho a dlhšieho spánku, čo by odstránilo študentku z rizikovej skupiny nedostatočného spánku. Vo všetkých štyroch prípadoch je potrebné ponúknuť rizikovým skupinám pomoc, napríklad vo forme rovesníckeho doučovania.

Projekt rovesníckeho doučovania nie je zdrojovo náročný a môže pomôcť študent*kám v núdzi zlepšiť si výsledky. Takýto projekt by mal byť zameraný hlavne na rizikové skupiny a žia*čky ktoré patria do viacero rizikových skupín, napríklad žiakov z nižšej triedy, ktorí spia málo a majú veľa absencií. Samozrejme, tento projekt nemusí byť zameraný iba na rizikové skupiny, ale aj na všetky žia*čky so slabými výsledkami. Prioritou by však malo byť zabezpečenie rovnosti a zaručenie rovnakých príležitostí, predovšetkým v prípade neovplyvniteľných faktorov, ako sú pohlavie a socio-ekonomická trieda. Projekt rovesníckeho doučovania je ale len jeden z rôznych riešení na túto problematiku. V každom prípade, treba zistiť viac o dopadoch rôznych faktorov a vytvoriť konkrétne riešenia na základe zistení. Nanešťastie, nepodarilo sa nám vytvoriť spoľahlivú neurónovú sieť na identifikáciu rizikových študent*iek. Tento proces teda nemožno automatizovať a bude vyžadovať ľudský prístup.

7 Zhrnutie

Úlohou tejto práce bolo nájsť vzťah medzi rôznymi faktormi a výsledkami a následne určiť rizikové skupiny. Najprv sme si vymedzili faktory a ich kategorizáciu, pričom faktory zahŕňali nasledovné: pohlavie, socio-ekonomický status, práca alebo aktivita, životná situácia, dochádzanie, spánok a absencia. Ako mierku výsledkov sme si zobrali nasledovné z minulého polroku: priemer známok a známky z matematiky, slovenčiny a angličtiny. Dáta sme zbierali pomocou dotazníku, na ktorý odpovedalo 253 žiakov. Dáta boli vo väčšine prípadov dostatočne diverzné, okrem napríklad v prípade spánku s kategóriou dlhého spánku. Väčšina distribúcií v rôznych skupinách ale nespĺňala normalitu, takže sme používali neparametrický Kruskal-Wallis test a rozdieli medzi skupinami sme pozorovali podľa rank-biserálnej korelácie a rozdielu v priemeroch a mediánoch, ako aj neparametrický Dunn test na overenie p-hodnoty medzi dvoma skupinami. Pri štyroch faktoroch sme našli významný vzťah, tie boli nasledovné: pohlavie, socio-ekonomický status, spánok a absencia. Jedno z viacerých možných riešení je rovesnícke doučovanie pre žiakov z rizikových skupín. Tiež treba ale limitovať počet študentov v ovplyvniteľných rizikových skupinách, ako nedostatok spánku a veľká absencia.

Resume

The task of this work was to find the relationship between various factors and outcomes and then to identify groups at risk. First, we defined the factors and their categorization, and the factors included the following: gender, socioeconomic status, occupation, living situation, commute, sleep, and absence. We took the following from last semester as a measure of achievement: grade point average and grades in math, English, and Slovak. We collected the data using a questionnaire to which 253 pupils responded. The data were sufficiently divergent in most cases, except for example in the case of sleep with the category of long sleep. However, most of the distributions in the different groups did not meet normality, so we used the nonparametric Kruskal-Wallis test and observed differences between groups by rank-biserial correlation and difference in means and medians, as well as the nonparametric Dunn test to test the p-value between the two groups. For four factors we found a significant relationship, these were as follows: gender, socioeconomic status, sleep and absenteeism. Peer tutoring for at-risk pupils is one of several possible solutions. However, we also need to limit the number of students in influential at-risk groups, such as sleep deprivation and high absenteeism.

Zoznam použitej literatúry

1 Kelly, W. E., Kelly, K. E., Clanton, R. C. (2001) 'The relationship between sleep length and grade-point average among college students', *College Student Journal*, 35(1), 84-86. Dostupné na:

https://www.researchgate.net/publication/285664611_The_relationship_between_sleep_length_and_grade-point_average_among_college_students

2 Moonie, S., Sterling, D. A., Figgs, L. W., & Castro, M. (2008) 'The Relationship Between School Absence, Academic Performance, and Asthma Status', *Journal of School Health*, 78(3), 140-148. doi:10.1111/j.1746-1561.2007.00276.x

3 Hancock, K. J., Lawrence, D., Shepherd, C. C. J., Mitrou, F., & Zubrick, S. R. (2017) 'Associations between school absence and academic achievement: Do socioeconomics matter?' *British Educational Research Journal*, 43(3), 415–440. doi:10.1002/berj.3267

4 Kobus, M. B. W., Van Ommeren, J. N., & Rietveld, P. (2015) 'Student commute time, university presence and academic achievement', *Regional Science and Urban Economics*, 52, 129–140. doi:10.1016/j.regsciurbeco.2015.03.001

5 Morris, E. A., & Zhou, Y. (2018) 'Are long commutes short on benefits? Commute duration and various manifestations of well-being', *Travel Behaviour and Society*, 11, 101–110. doi:10.1016/j.tbs.2018.02.001

6 Beekhoven *, S., De jong U., & Van hout H. (2004) 'The impact of first-year students' living situation on the integration process and study progress', *Educational Studies*, 30(3), 277–290. doi:10.1080/0305569042000224224

7 Boot, C. R., Rietmeijer, C. B., Vonk, P., & Meijman, F. J. (2009) 'Perceived health profiles of Dutch university students living with their parents, alone or with peers', *International Journal of Adolescent Medicine and Health*, 21(1). doi:10.1515/ijamh.2009.21.1.41

8 Dundes, L., & Marx, J. (2006) 'Balancing Work and Academics in College: Why Do Students Working 10 to 19 Hours per Week Excel?', *Journal of College Student Retention: Research, Theory & Practice*, 8(1), 107–120. doi:10.2190/7ucu-8f9m-94qg-5wwq

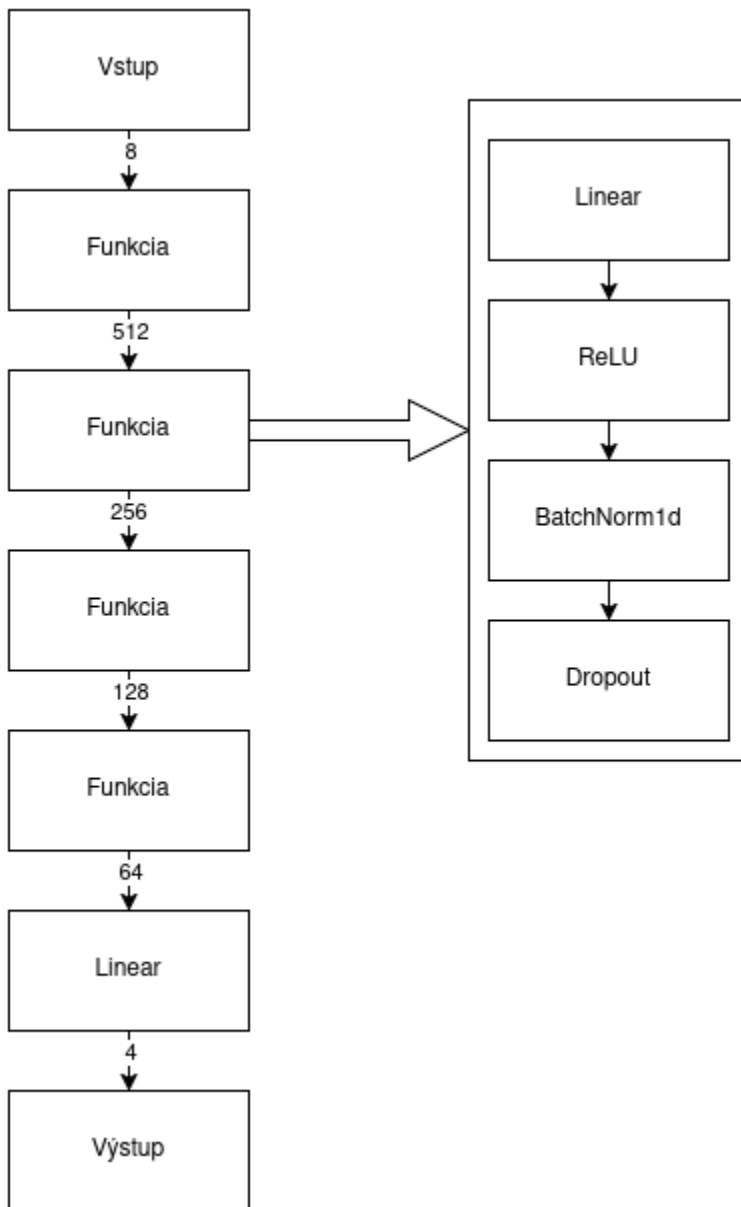
9 Gayles, J. G., & Hu, S. (2009) 'The Influence of Student Engagment and Sport Participation on College Outcomes Among Division I Student Athletes', *The Journal of Higher Education*, 80(3), 315–333. doi:10.1080/00221546.2009.11779015

10 Ahmar, F., Anwar, E. (2013) 'Socio Economic Status and its Relation to Academic Achievement of Higher Secondary School Students', *IOSR Journal Of Humanities And Social Science*, 13(6), 13-20. doi:10.9790/0837-1361320

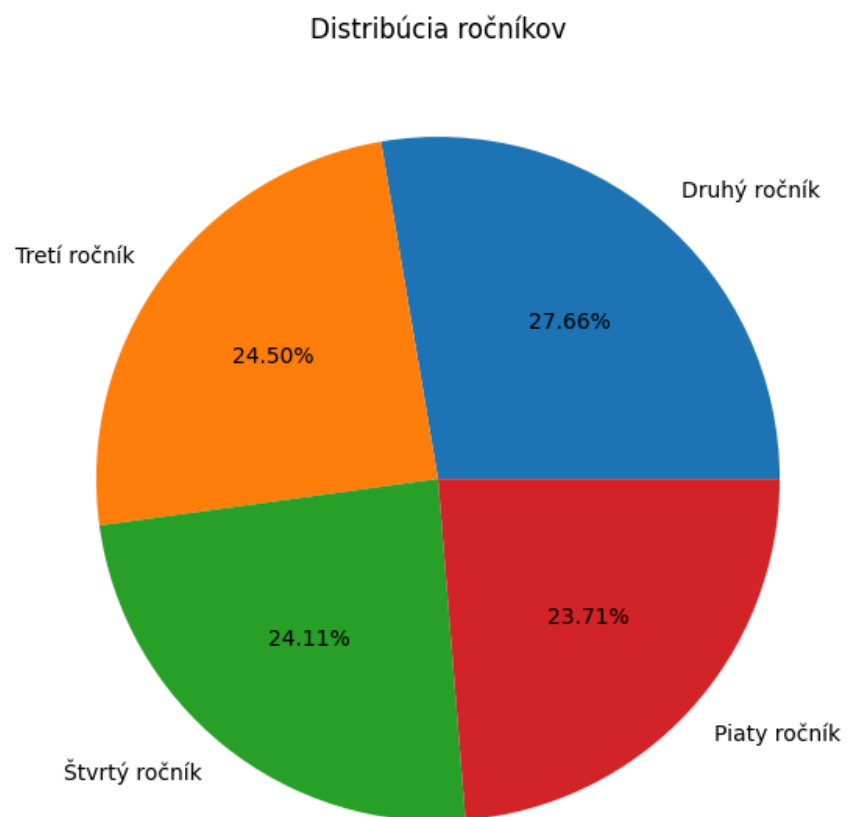
11 Barry, J. (2005) 'The effects of socio-economic status on academic achievement', *Wichita State University*, Wichita, USA. Dostupné na: <https://soar.wichita.edu/server/api/core/bitstreams/2d8e70a4-fbc7-44de-abf9-0a695b62efa8/content>

Prílohy

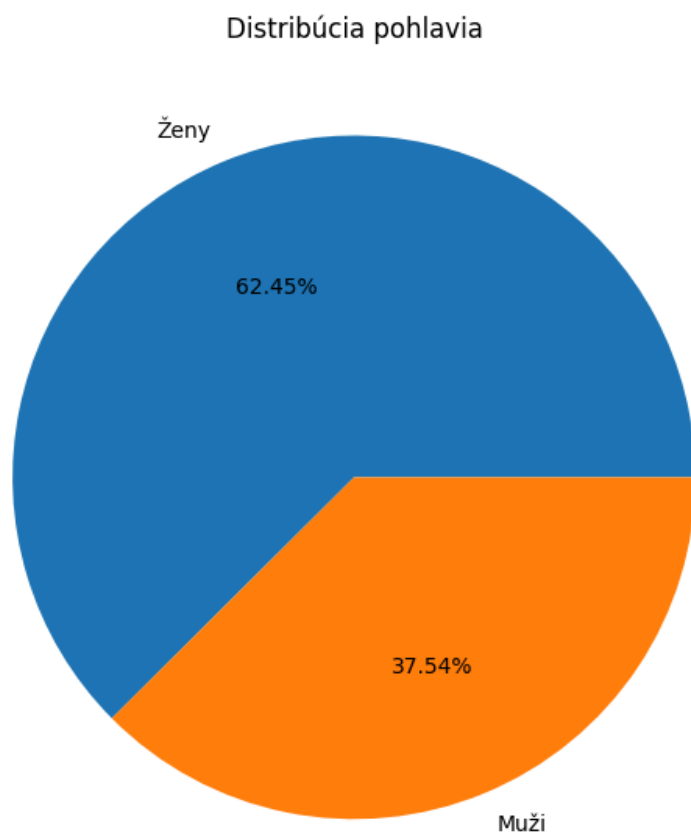
Obrázok 1



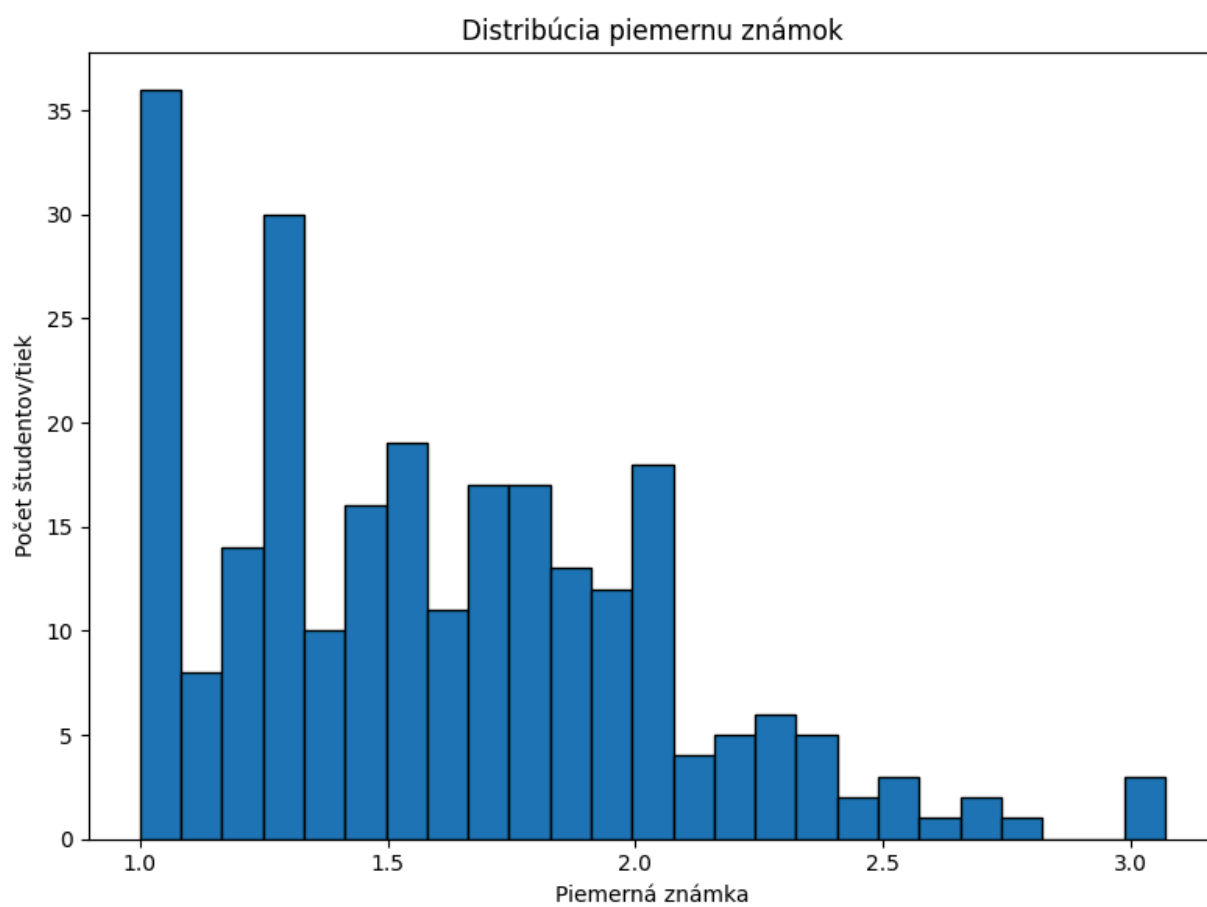
Graf 1



Graf 2

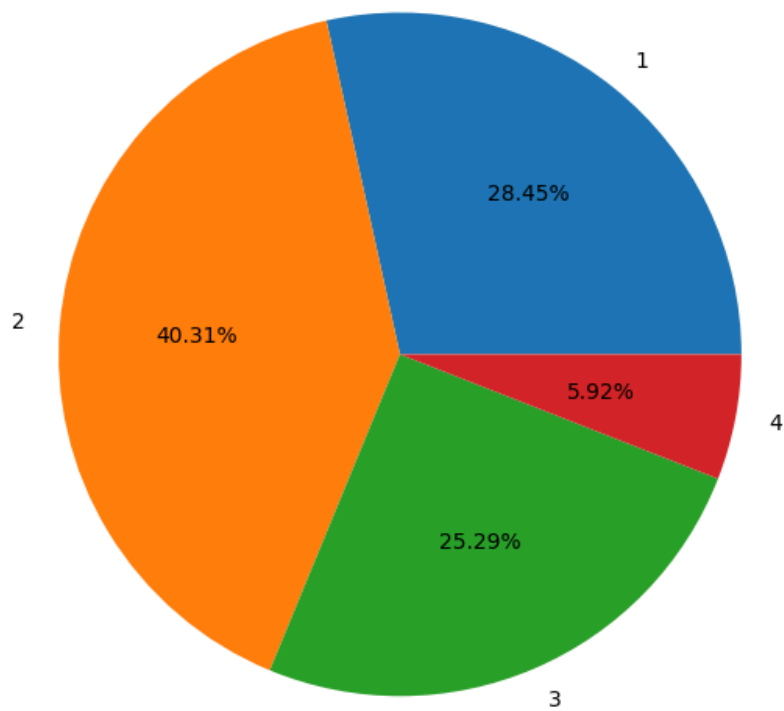


Graf 3



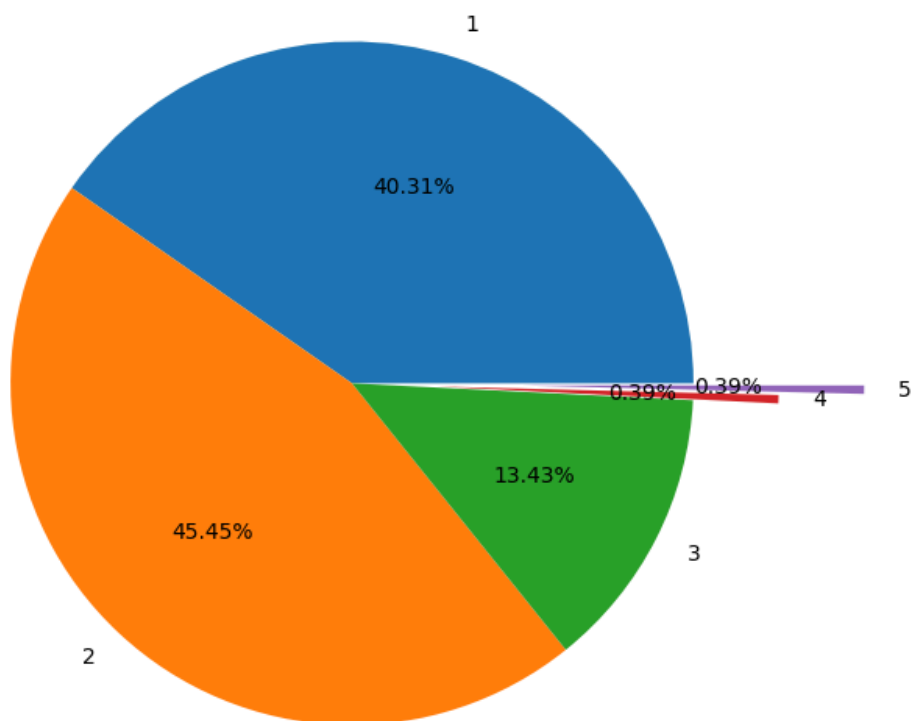
Graf 4

Distribúcia známok z matematiky



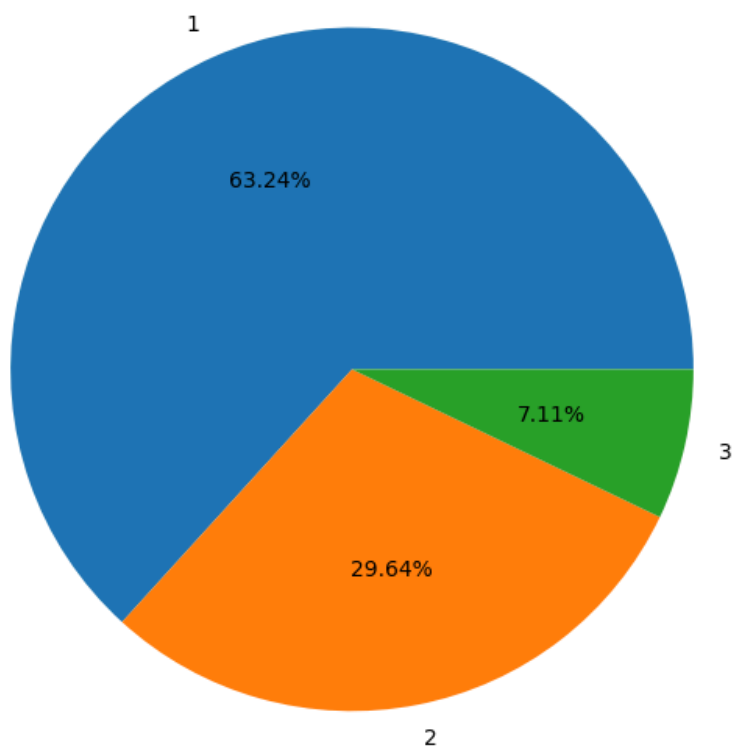
Graf 5

Distribúcia známok zo slovenčiny



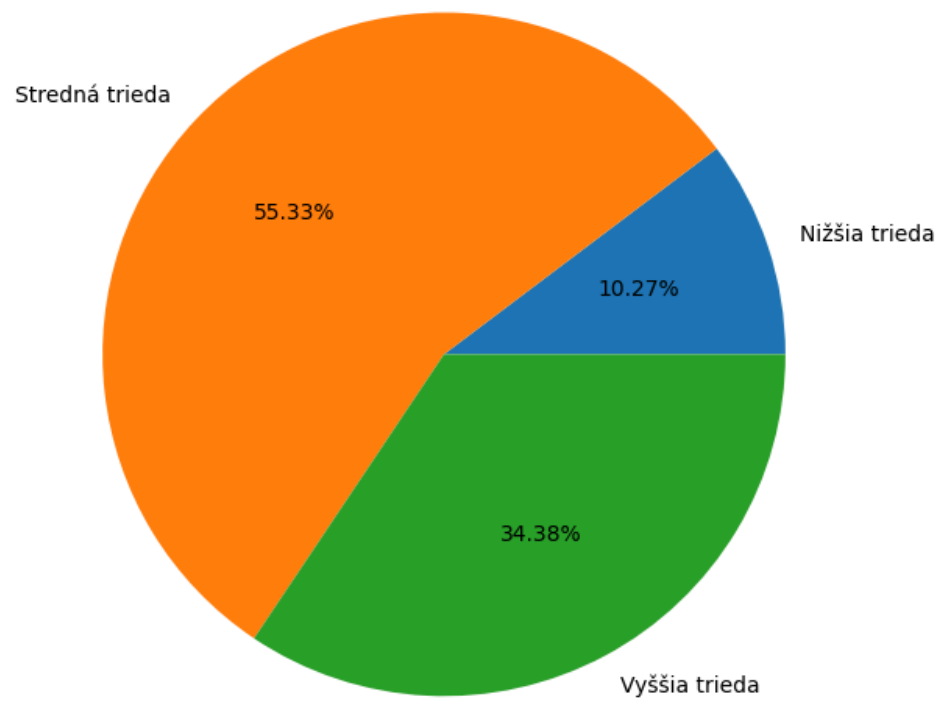
Graf 6

Distribúcia znáмок z angličtiny

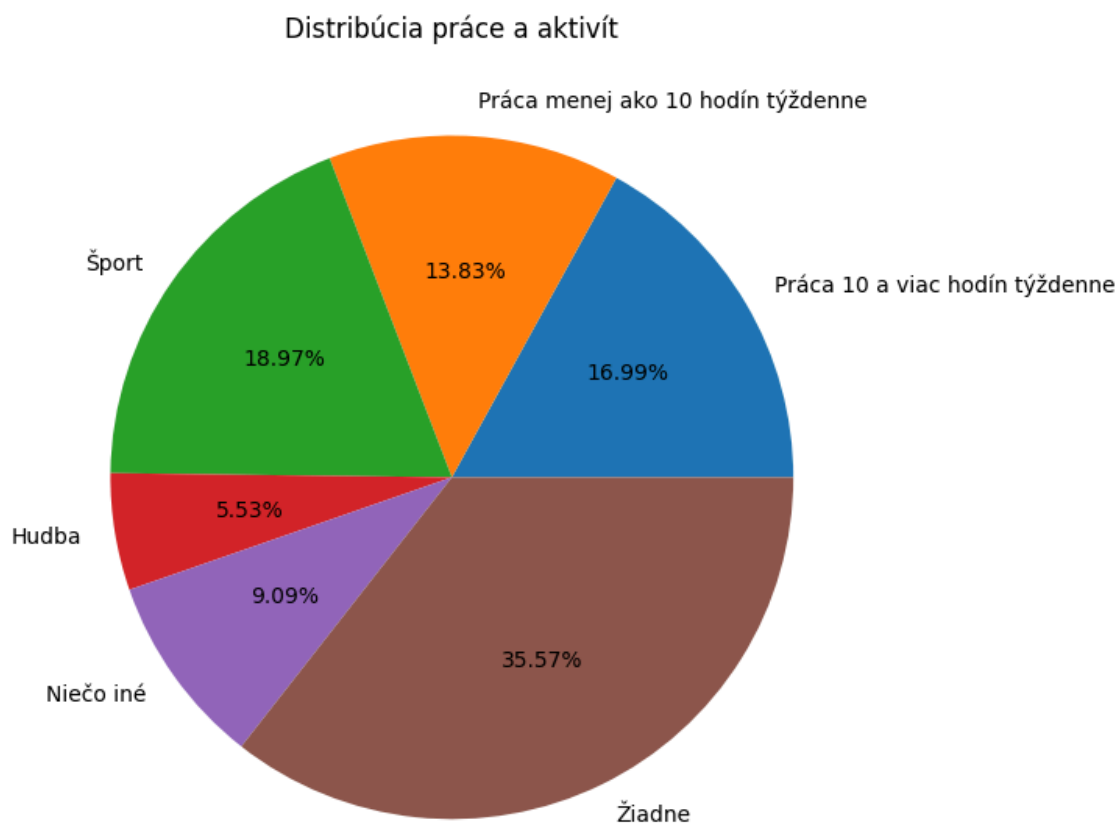


Graf 7

Distribúcia socio-ekonomických tried

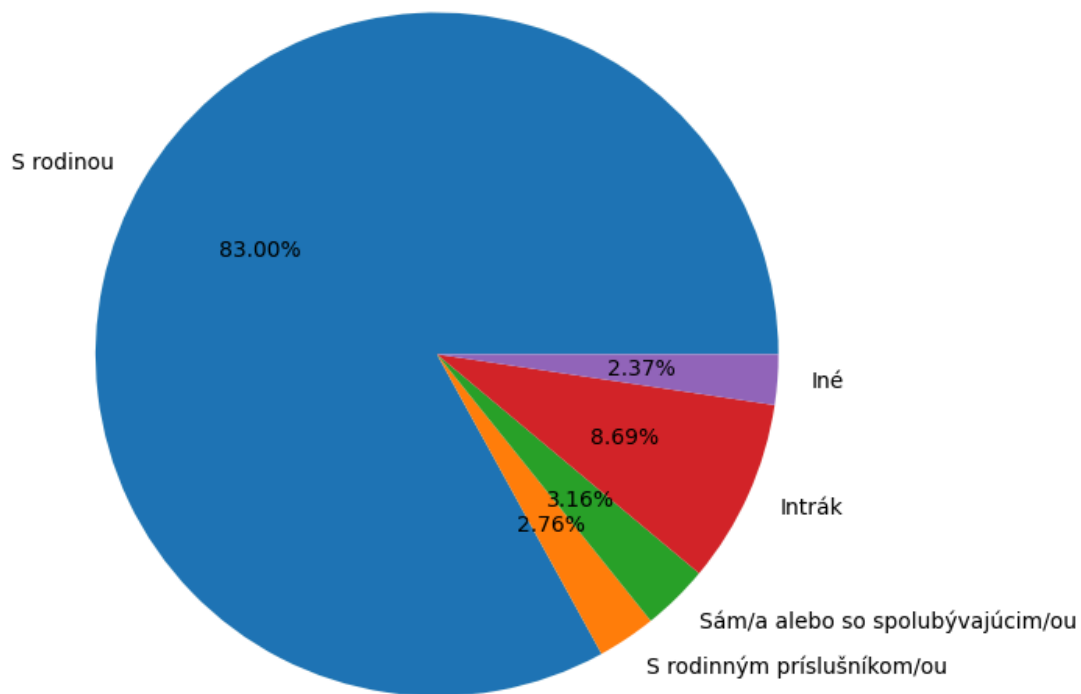


Graf 8

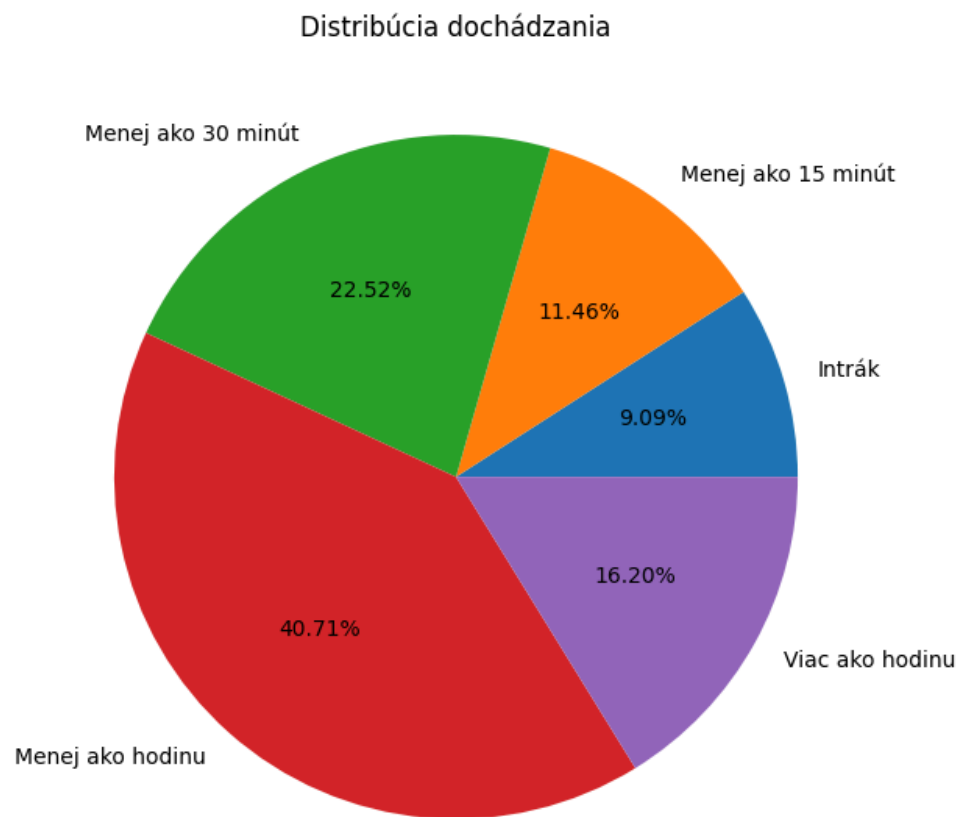


Graf 9

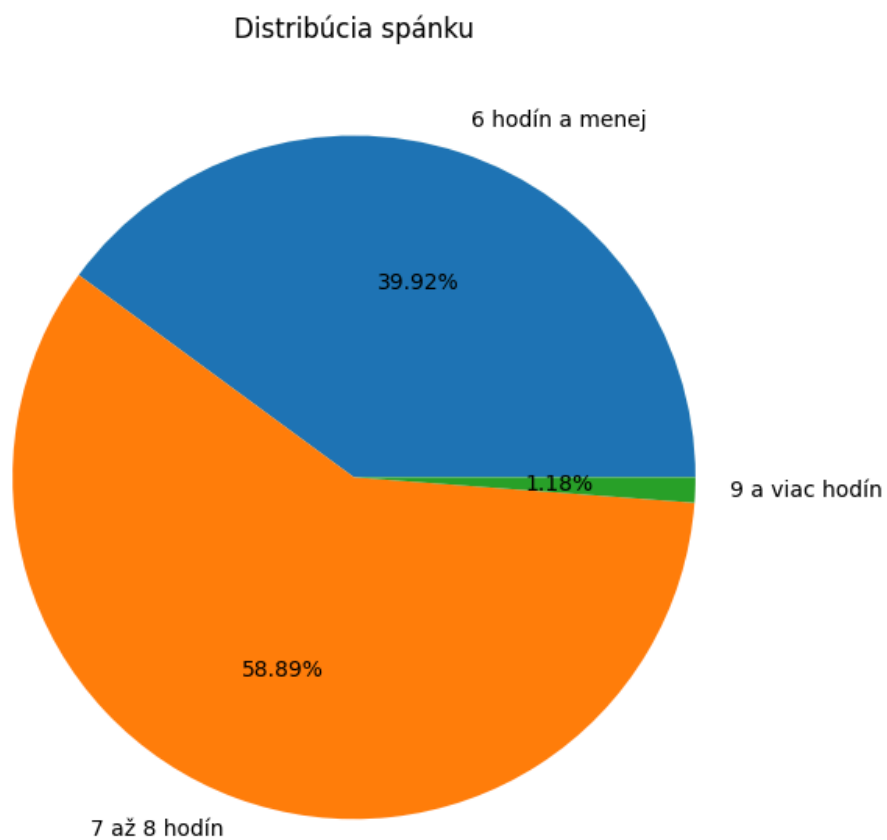
Distribúcia životných situácií



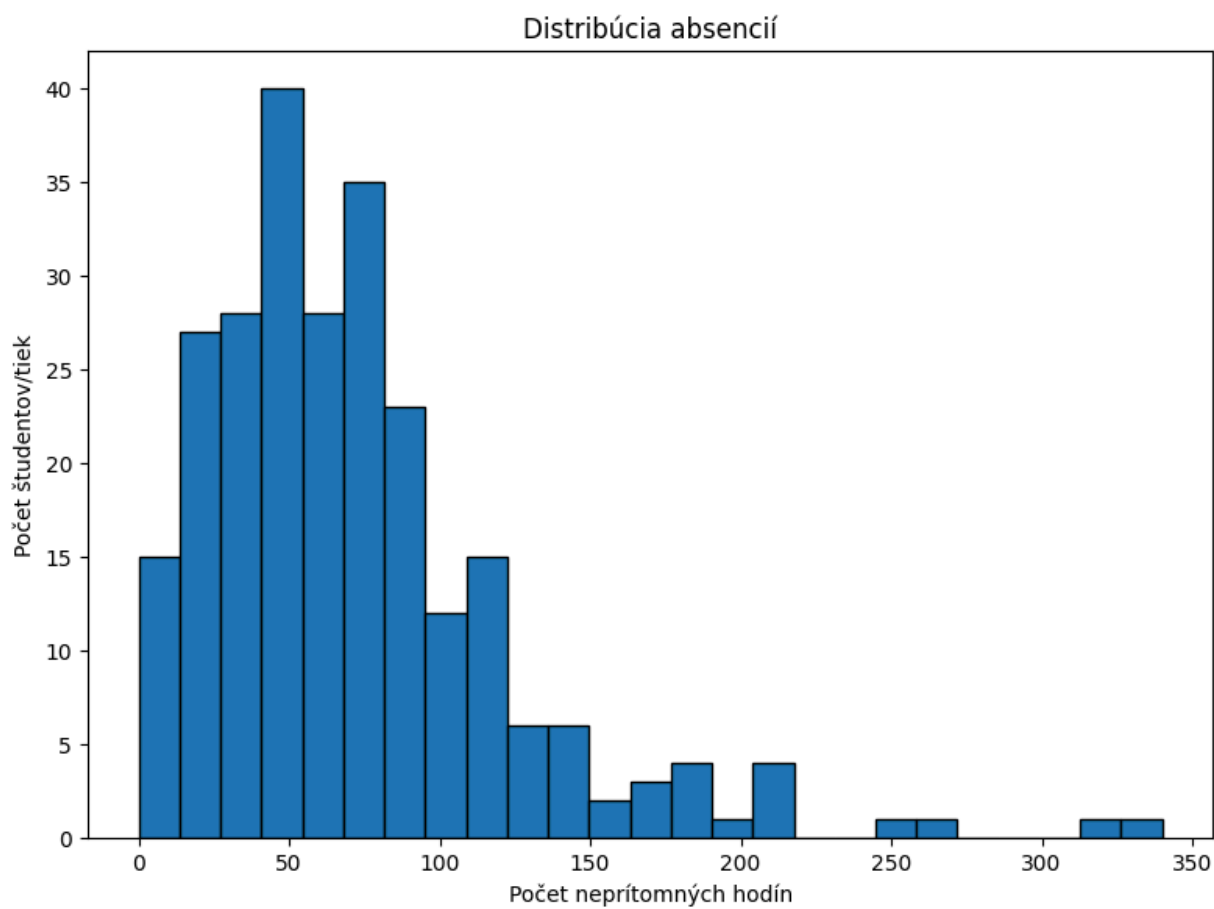
Graf 10



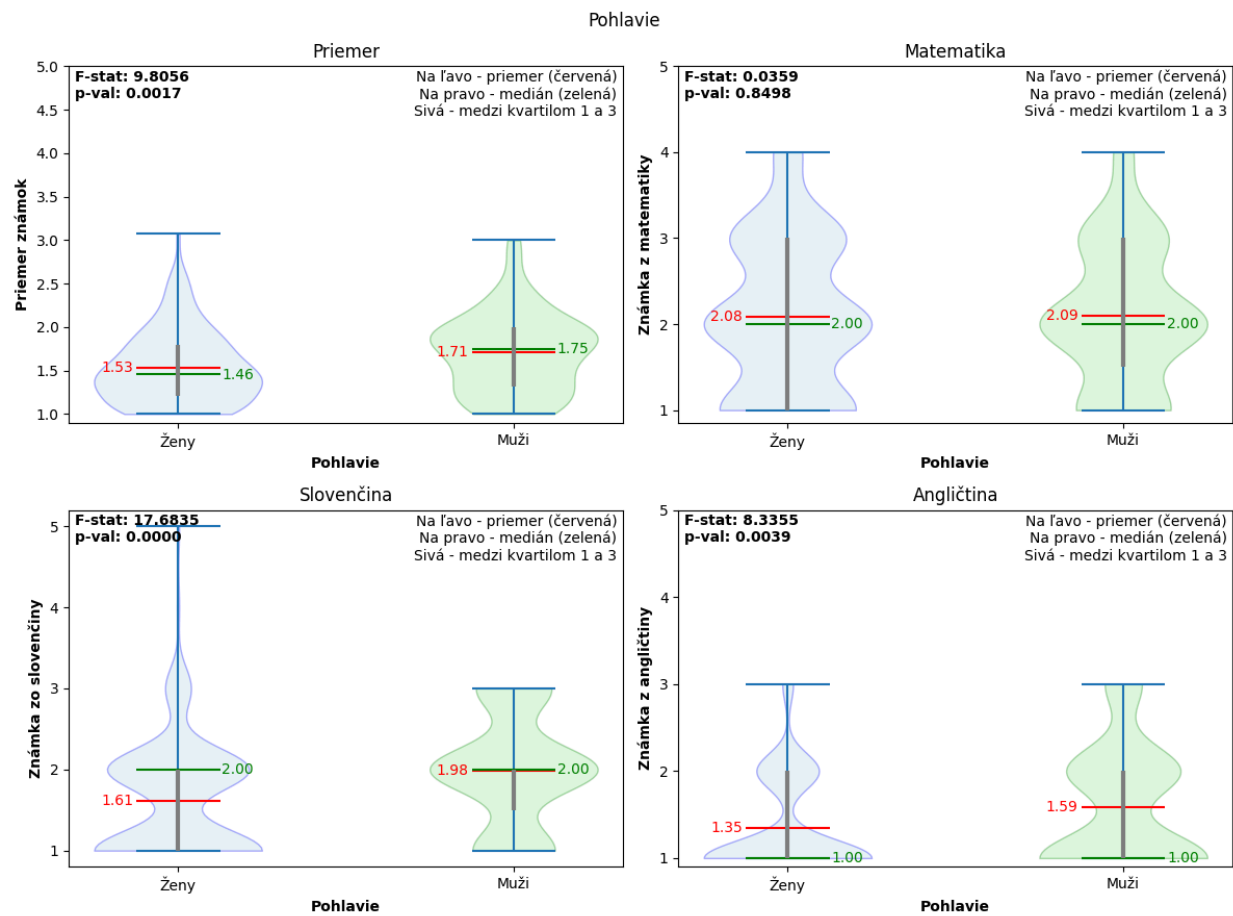
Graf 11



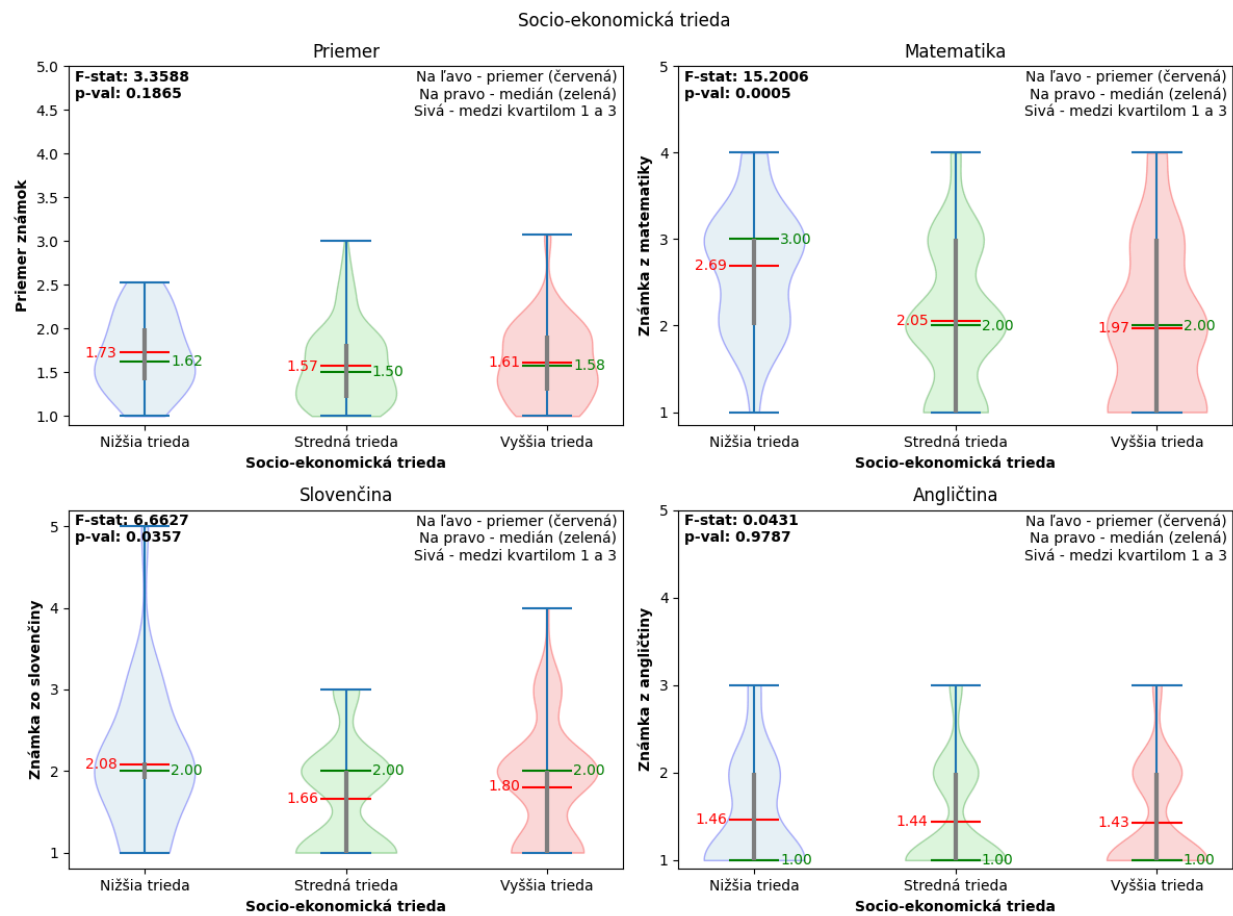
Graf 12



Graf 13

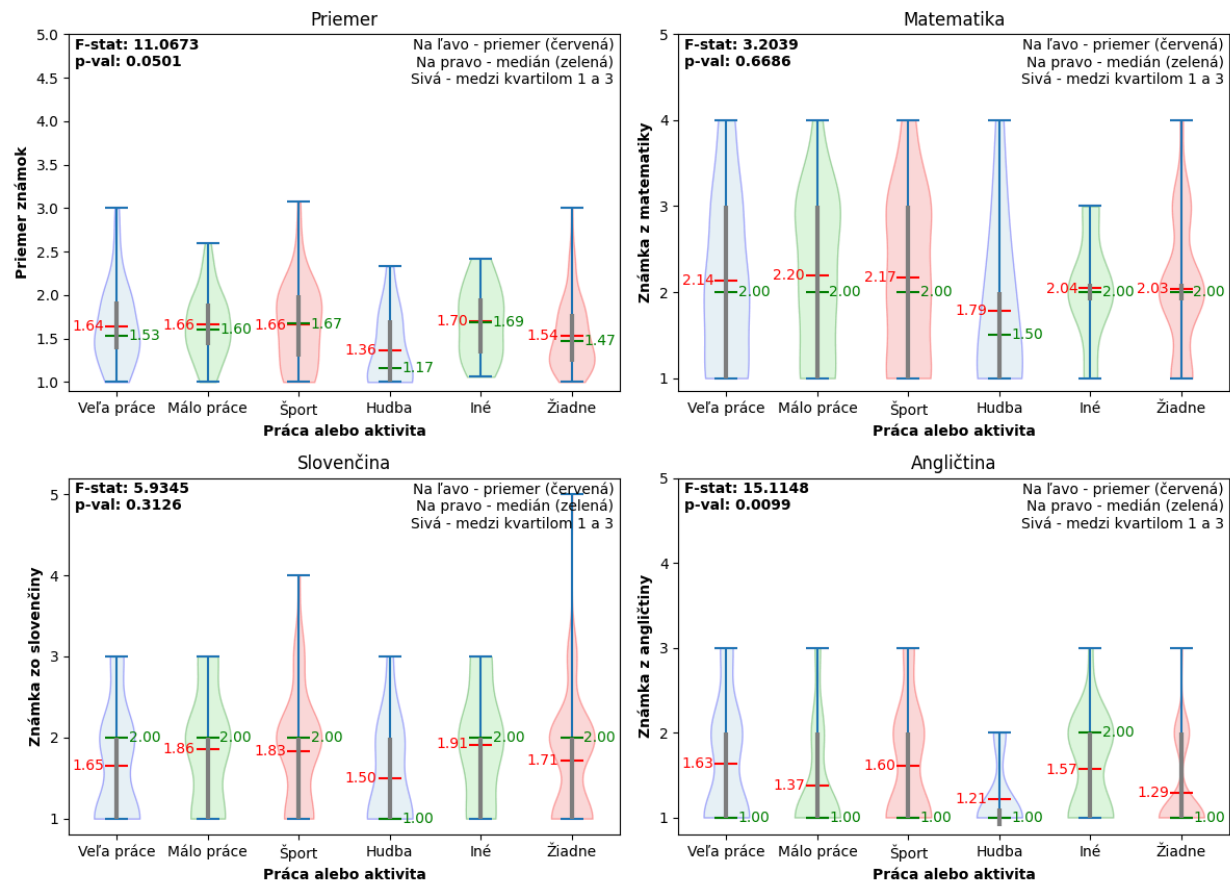


Graf 14

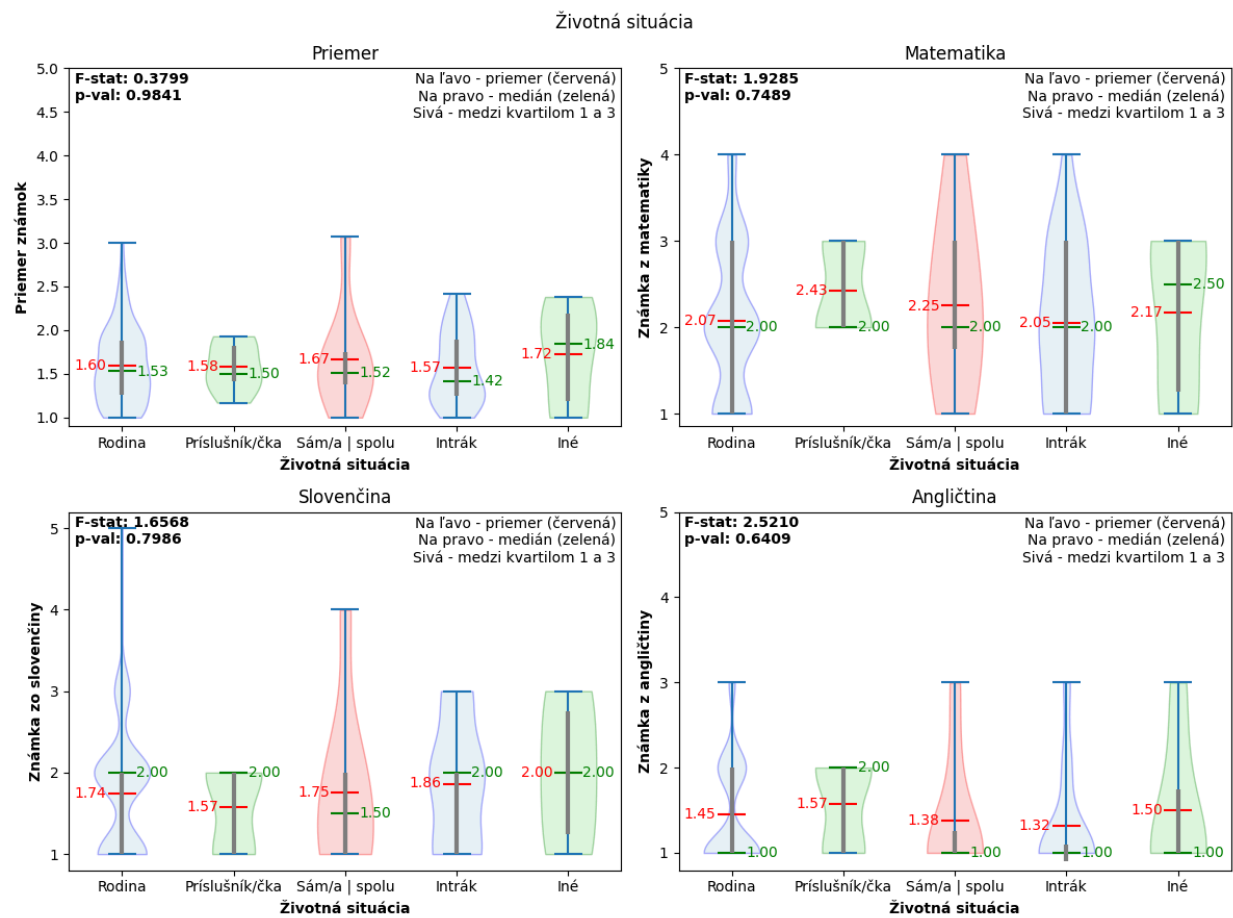


Graf 15

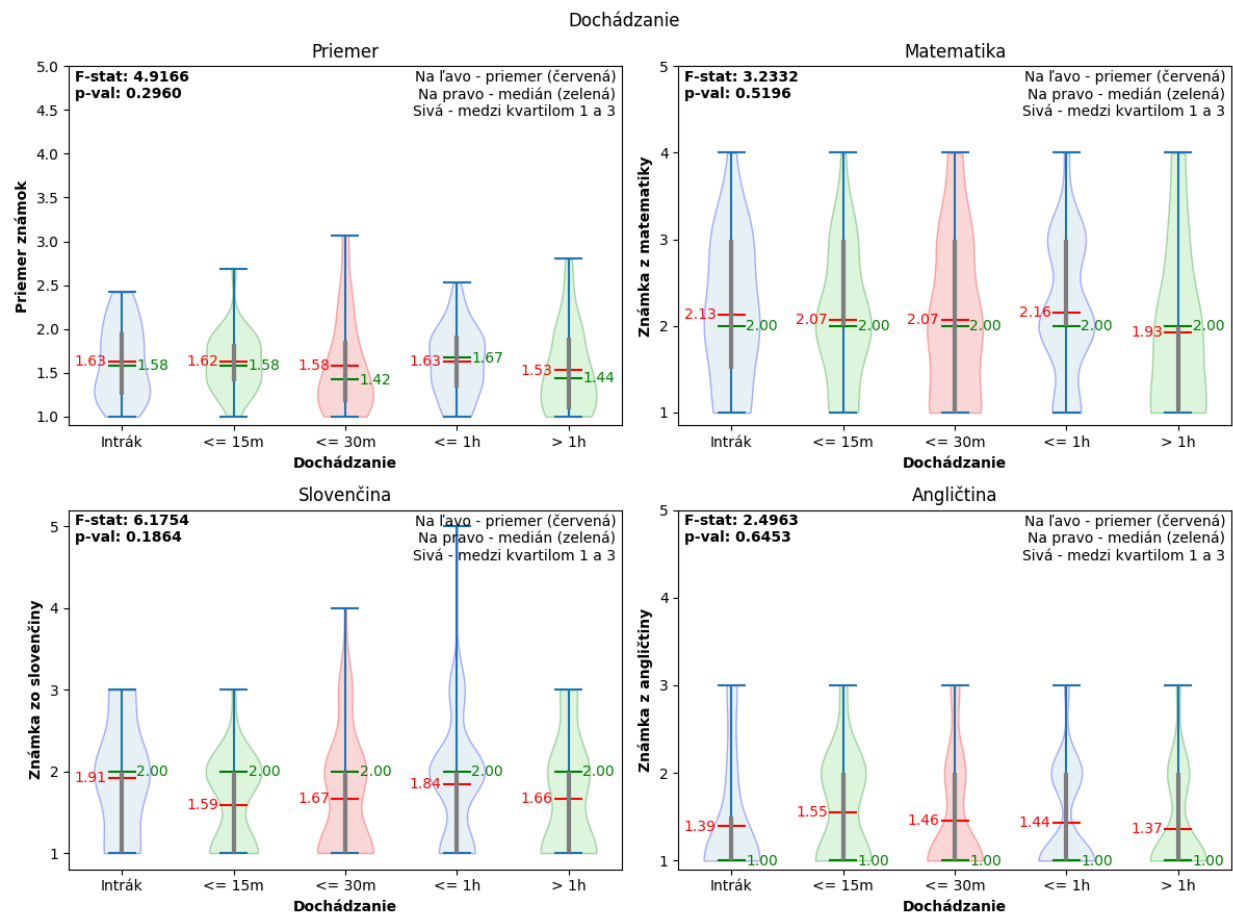
Práca alebo aktivita



Graf 16

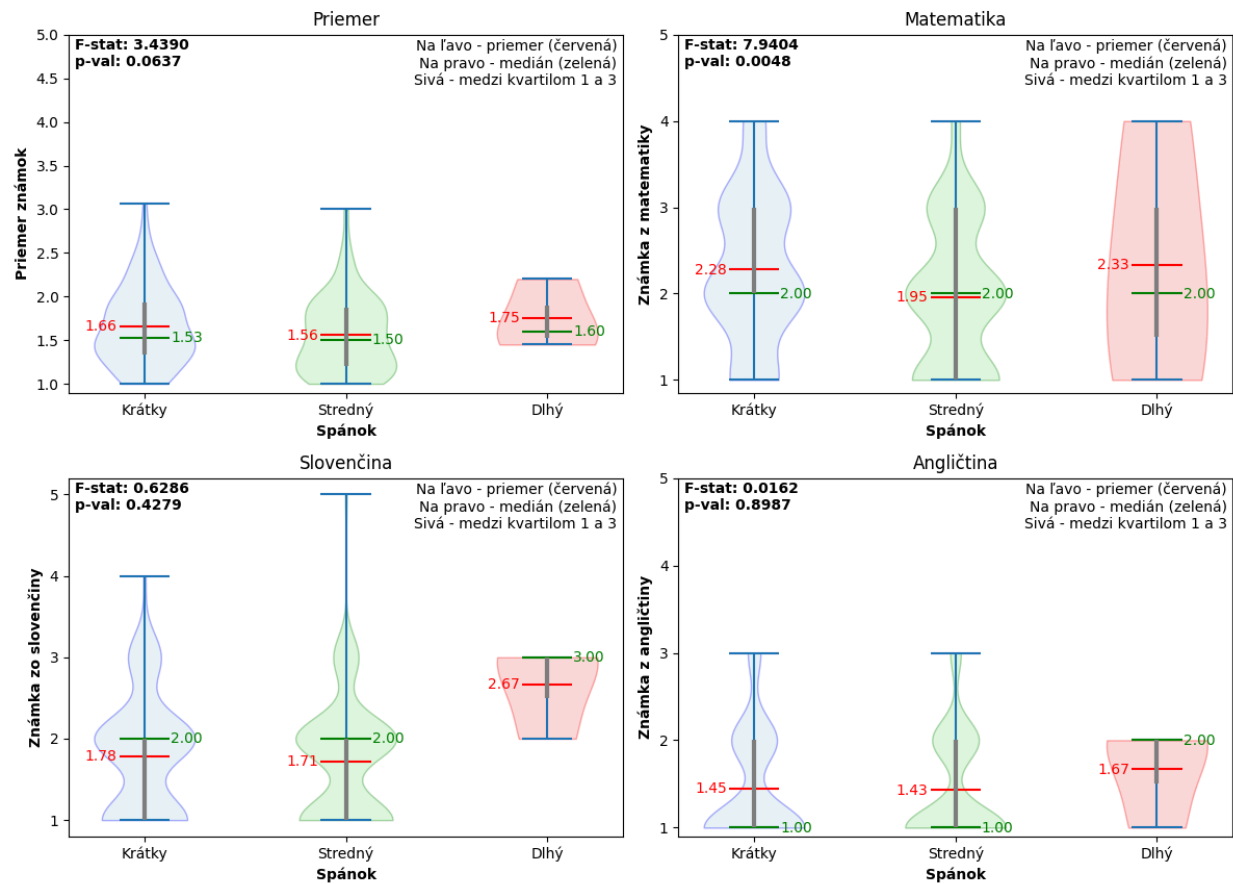


Graf 17

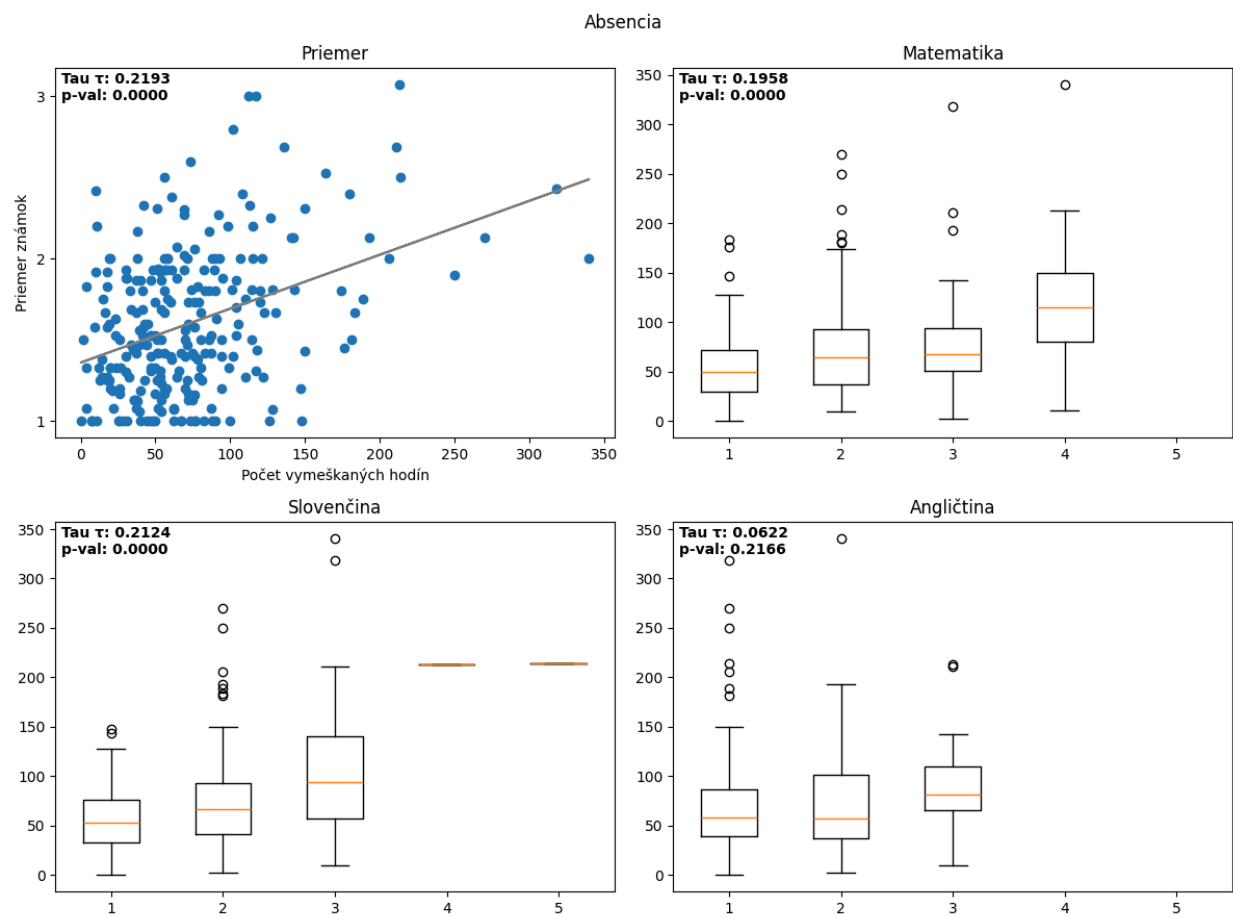


Graf 18

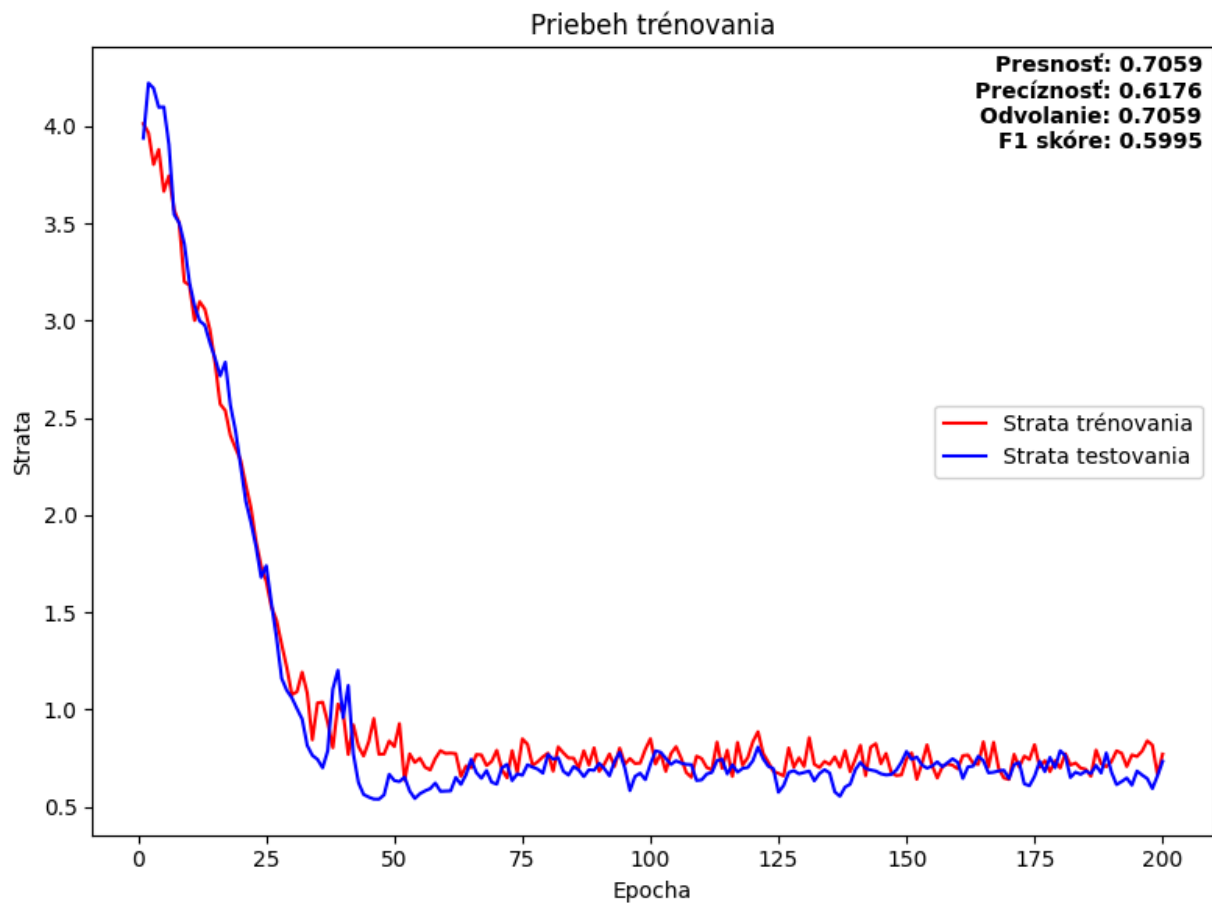
Spánok



Graf 19



Graf 20



Legenda pre tabuľky 1, 2 a 3

| Písmeno | Skupina |
|---------|---------|
| A | Ženy |
| B | Muži |

Tabuľka 1

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | -0,0153 | -0,1732 | -0,2950 | 0,0016 |

Tabuľka 2

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | -0,0188 | -0,3650 | 0,0000 | 0,0000 |

Tabuľka 3

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | -0,0119 | -0,2414 | 0,0000 | 0,0037 |

Legenda pre tabuľky 4 a 5

| Písmeno | Skupina |
|---------|---------------------------------|
| A | Nižšia socio-ekonomická trieda |
| B | Stredná socio-ekonomická trieda |
| C | Vyššia socio-ekonomická trieda |

Tabuľka 4

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | 0,0658 | 0,6423 | 1,0000 | 0,0014 |
| A | C | 0,1179 | 0,7268 | 1,0000 | 0,0003 |
| B | C | 0,0051 | 0,0845 | 0,0000 | 1,0000 |

Tabuľka 5

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | 0,0440 | 0,4198 | 0,0000 | 0,0497 |
| A | C | 0,0399 | 0,2723 | 0,0000 | 0,5239 |
| B | C | -0,0084 | -0,1475 | 0,0000 | 0,3706 |

Legenda pre tabuľku 6

| Písmeno | Skupina |
|---------|-----------------------------------|
| A | Práca 10 a viac hodín týždenne |
| B | Práca menej ako 10 hodín týždenne |
| C | Šport |
| D | Hudba |
| E | Iné |
| F | Žiadne |

Tabuľka 6

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | 0,0768 | 0,2565 | 0,0000 | 0,8790 |
| A | C | 0,0041 | 0,0237 | 0,0000 | 1,0000 |
| A | D | 0,2089 | 0,4136 | 0,0000 | 0,5968 |
| A | E | 0,0021 | 0,0627 | -1,0000 | 1,0000 |
| A | F | 0,0439 | 0,3390 | 0,0000 | 0,0844 |
| B | C | -0,0658 | -0,2327 | 0,0000 | 1,0000 |
| B | D | 0,0799 | 0,1571 | 0,0000 | 1,0000 |
| B | E | -0,1193 | -0,1938 | -1,0000 | 1,0000 |
| B | F | 0,0083 | 0,0825 | 0,0000 | 1,0000 |
| C | D | 0,1810 | 0,3899 | 0,0000 | 0,7009 |
| C | E | -0,0043 | 0,0389 | -1,0000 | 1,0000 |
| C | F | 0,0385 | 0,3153 | 0,0000 | 0,1001 |
| D | E | -0,3495 | -0,3509 | -1,0000 | 0,9831 |
| D | F | -0,0199 | -0,0746 | 0,0000 | 1,0000 |

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| E | F | 0,0640 | 0,2763 | 1,0000 | 0,4553 |

Legenda pre tabuľku 7

| Písmeno | Skupina |
|---------|---------------------|
| A | Krátky spánok |
| B | Stredne-dlhý spánok |
| C | Dlhý spánok |

Tabuľka 7

| Skupina 1 | Skupina 2 | Veľkosť účinku | Rozdiel priemerov | Rozdiel mediánov | Post-Hoc p-hodnota |
|-----------|-----------|----------------|-------------------|------------------|--------------------|
| A | B | 0,0128 | 0,3242 | 0,0000 | 0,0046 |