

# Popis dátového súboru o absolventoch stredných škôl

**Matúš Baliak (ISP)**

**Matěj Bělín (ISP)**

**Štefan Domonkos (ISP)**

**Tadeáš Chujac (ISP)**

**Dávid Martinák (IVP)**

**Marec 2022**

## 1 Úvod

Cieľom tohto dokumentu je poskytnutie podrobnejšieho opisu databázového súboru zverejneného spolu s komentárom *Absolventi v čase pandémie* z dielne Inštitútu sociálnej politiky a Inštitútu vzdelávacej politiky. Dátový súbor poskytuje základné informácie o postavení na trhu práce absolventov stredných škôl, ktorí ukončili svoje stredoškolské štúdium v školskom roku 2019/2020. Dáta pokrývajú prvý polrok roka 2021 (január 2021 – jún 2021). Štruktúra zverejnených tabuliek pozostáva z troch úrovní detailu v delení na študijné a učebné odbory. Najdetailnejšiu analýzu nástupu absolventov na trh práce umožňujú údaje na úrovni školy. Vyšší stupeň agregácie nájdeme na krajskej úrovni. Najmenej granulózne výsledky poskytujú údaje na úrovni celého Slovenska. V nasledujúcich častiach dokumentu sa oboznámime so štruktúrou a obsahom jednotlivých dátových súborov. Zverejnené údaje na rôznych úrovniach detailu zdieľajú totožné základné charakteristiky a na výpočet sa uplatňujú rovnaké zásady, avšak štruktúra premenných je na každom stupni agregácie mierne odlišná. Používateľ by mal pri voľbe medzi dátami na rôznej úrovni agregácie vybrať tú úroveň, ktorá je najbližšia ku skúmanej vecnej problematike (napr. pri vyhodnocovaní priemerných miezd za odbory na celoštátnej úrovni je vhodné vychádzať z datasetu *Úroveň detailu – SR*). Dôvodom pre opatrnosť pri tejto voľbe je nutnosť anonymizovať dáta, ktorá vedie najmä pri vyššej úrovni detailu k rastúcemu počtu buniek, ktoré nemohli byť zverejnené kvôli riziku deanonymizácie.

## 2 Úroveň detailu – škola

### 2.1 Opis datasetu

Údaje na najdetailnejšej úrovni poskytujú prehľad o priemernom stave postavenia absolventov konkrétnej školy. Vďaka kombinácii jednoznačného identifikátora školy a študijného a učebného odboru sa nám podarilo správne zaradiť absolventov a sledovať ich nástup na trh práce. Hodnoty v databáze sú vypočítané ako priemer za január – jún 2021. Z údajov dokážeme identifikovať počet absolventov konkrétnej kombinácie škola - študijný alebo učebný odbor. Ďalej je možné určiť status absolventov (pracujúci, materská/rodičovská, pokračovanie v štúdiu, nezamestnaní, mimo evidencie). Súčasťou zverejnených údajov na úrovni školy je priemerná a mediánová hrubá mzda. Oba ukazovatele vychádzajú z údajov za celé sledované obdobie a do výpočtu vstupovali iba zamestnanecké poisťovne, ku ktorým prislúchala hrubá mesačná mzda aspoň na úrovni minimálnej mzdy (623 eur za rok 2021). Ukazovatele TOP1\_podiel až TOP3\_nazov prezentujú podiel a názov troch hlavných tried zo štatistickej klasifikácie SK NACE Rev.2, kde absolventi pracovali najčastejšie. Z dôvodu zabezpečenia ochrany osobných údajov podliehajú zverejnené dáta anonymizácii. V prípade, ak je celkový počet spracovaných osôb na úrovni školy a študijného a učebného odboru nižší ako 5, údaje o postavení absolventov na trhu práce nebudú zverejnené. Pre takýto postup sme sa rozhodli, aby nebolo možné identifikovať stav konkrétnych absolventov. Anonymizácii podliehajú aj mzdové ukazovatele spolu s podielmi odvetví, kde sa absolventi najčastejšie uplatnili. V prípade, ak je priemerný počet zamestnaných osôb menší ako 3, mzdové ani odvetvové ukazovatele nebudú zverejnené.

### 2.2 Popis premenných súboru na úrovni školy a učebného / študijného odboru

Dátový súbor (Úroveň detailu – škola) pozostáva z nasledovných 28 premenných:

#### Identifikácia školy a študijného a učebného odboru:

eduid\_sasz – EDUID kód strednej školy (jednoznačný identifikátor strednej školy)

nazov\_skoly – názov strednej školy

kraj – kraj sídla školy

obec – obec sídla školy

sauo\_kod – kód študijného a učebného odboru

nazov\_sauo – názov študijného a učebného odboru

**Stav absolventov:**

spracované osoby – celkový počet absolventov

muz – celkový počet absolventov mužského pohlavia

zena – celkový počet absolventov ženského pohlavia

pracujuce\_osoby – počet pracujúcich absolventov, výpočet = zamestnane\_osoby + osoby\_na\_dohode + osoby\_szco.

zamestnane\_osoby – počet zamestnaných absolventov (rozhodujúci je zamestnanecký poistný vzťah v SP)

osoby\_na\_dohode – počet absolventov pracujúcich podľa dohody o prácach vykonávaných mimo pracovného pomeru

osoby\_szco – počet absolventov vykonávajúcich samostatne zárobkovú činnosť (identifikovaní z Registra právnických osôb)

osoby\_mater\_rodic – počet absolventov na materskej/rodičovskej dovolenke

pokrac\_studium\_SS – počet absolventov pokračujúcich v štúdiu na strednej škole na Slovensku

pokrac\_studium\_VS\_SR – počet absolventov pokračujúcich v štúdiu na vysokej škole na Slovensku

pokrac\_studium\_zahr – počet absolventov pokračujúcich v štúdiu na vysokej škole v zahraničí

nezamestnane\_osoby – celkový počet absolventov zaregistrovaných ako uchádzači o zamestnanie

z\_toho\_nezamestnane\_osoby\_na\_dohode – počet absolventov zaregistrovaných ako uchádzač o zamestnanie a súčasne pracujúcich podľa dohody o prácach vykonávaných mimo pracovného pomeru

mimo\_evidencie – počet osôb mimo evidencie (napr. pracujúci/ neidentifikovaní študujúci v zahraničí, dobrovoľne nezamestnaní)

**Mzdové ukazovatele a odvetvové uplatnenie:**

priem\_hmesmzda – priemerná hrubá mesačná mzda (do úvahy sa berú iba zamestnanecké poistné vzťahy s vymeriavacím základom aspoň vo výške minimálnej mzdy)

medianova\_mzda – mediánová hrubá mesačná mzda (do úvahy sa berú iba zamestnanecké poistné vzťahy s vymeriavacím základom aspoň vo výške minimálnej mzdy)

TOP1\_podiel – podiel sekcie SK NACE, v ktorej bolo zamestnaných v priemere za sledované obdobie najviac absolventov

TOP1\_nazov – názov sekcie SK NACE, v ktorej bolo zamestnaných v priemere za sledované obdobie najviac absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP2\_podiel – podiel sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie druhý najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP2\_nazov – názov sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie druhý najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP3\_podiel – podiel sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie\* tretí najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP3\_nazov – názov sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie\* tretí najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

### 3 Úroveň detailu – kraj

#### 3.1 Opis datasetu

Vyšší stupeň agregácie úvodných krokov absolventov na trhu práce poskytujú údaje na úrovni krajov. Pri tomto dátovom súbore vytvárame jedinečné kombinácie kraj - študijný a učebný odbor. Zvolený prístup umožňuje porovnať nástup absolventov na trh práce podľa regionálnej príslušnosti absolvovanej školy a z pohľadu študijného a učebného odboru. Základné princípy použité na úrovni školy sa uplatňujú rovnako aj na úrovni krajov. Naďalej platí, že hodnoty v databáze sú vypočítané ako priemer za január – jún 2021. Z údajov dokážeme identifikovať počet absolventov konkrétnej kombinácie kraj - študijný odbor. Ďalej je možné určiť status absolventov (pracujúci, materská/rodičovská, pokračovanie v štúdiu, nezamestnaní, prípadne mimo evidencie). Súčasťou zverejnených údajov je priemerná a mediánová hrubá mzda za kraje. Oba ukazovatele vychádzajú z údajov za celé sledované obdobie a do výpočtu vstupovali iba zamestnanecké poisťovne vzťahy, ku ktorým prislúchala hrubá mesačná mzda aspoň na úrovni minimálnej mzdy (623 eur za rok 2021). Ukazovatele TOP1\_podiel až TOP3\_nazov prezentujú podiel a názov troch hlavných tried zo štatistickej klasifikácie SK NACE Rev.2, kde absolventi pracovali najčastejšie. Z dôvodu zabezpečenia ochrany osobných údajov podliehajú zverejnené dáta anonymizácii. V prípade, ak je celkový počet spracovaných osôb na úrovni študijného a učebného odboru za kraj nižší ako 5, údaje o stave absolventov budú skryté. Pre takýto postup sme sa rozhodli, aby nebolo možné identifikovať stav konkrétnych absolventov. Anonymizácii podliehajú aj mzdové ukazovatele spolu s podielmi odvetví, kde sa absolventi najčastejšie uplatnili. V prípade, ak je priemerný počet zamestnaných osôb menší ako 3, mzdové ani odvetvové ukazovatele nebudú zverejnené.

#### 3.2 Popis premenných súboru na úrovni kraja (podľa umiestnenia školy) a študijného / učebného odboru

Dátový súbor (Úroveň detailu – kraj) pozostáva z nasledovných 25 premenných:

##### Identifikácia kraja a študijného a učebného odboru:

kraj – kraj sídla školy

sauo\_kod – kód študijného a učebného odboru

nazov\_sauo – názov študijného a učebného odboru

##### Stav absolventov:

spracovane\_osoby – celkový počet absolventov

muz – celkový počet absolventov mužského pohlavia

zena – celkový počet absolventov ženského pohlavia

pracujuce\_osoby – počet pracujúcich absolventov, výpočet = zamestnane\_osoby + osoby\_na\_dohode + osoby\_szco.

zamestnane\_osoby – počet zamestnaných absolventov (rozhodujúci je zamestnanecký poisťový vzťah v SP)

osoby\_na\_dohode –počet absolventov pracujúcich podľa dohody o prácach vykonávaných mimo pracovného pomeru

osoby\_szco –počet absolventov vykonávajúcich samostatne zárobkovú činnosť (identifikovaní z Registra právnických osôb)

osoby\_mater\_rodic –počet absolventov na materskej/rodičovskej dovolenke

pokrac\_studium\_SS –počet absolventov pokračujúcich v štúdiu na strednej škole na Slovensku

pokrac\_studium\_VS\_SR –počet absolventov pokračujúcich v štúdiu na vysokej škole na Slovensku

pokrac\_studium\_zahr –počet absolventov pokračujúcich v štúdiu na vysokej škole v zahraničí

nezamestnane\_osoby – celkový počet absolventov zaregistrovaných ako uchádzači o zamestnanie

z\_toho\_nezamestnane\_osoby\_na\_dohode – počet absolventov zaregistrovaných ako uchádzač o zamestnanie a súčasne pracujúcich podľa dohody o prácach vykonávaných mimo pracovného pomeru

mimo\_evidencie – počet osôb mimo evidencie (napr. pracujúci/ neidentifikovaní študujúci v zahraničí, dobrovoľne nezamestnaní)

#### **Mzdové ukazovatele a odvetvové uplatnenie:**

priem\_hmesmzda – priemerná hrubá mesačná mzda (do úvahy sa berú iba zamestnanecké poisťovne vzťahy s vymeriavacím základom aspoň vo výške minimálnej mzdy)

medianova\_mzda – mediánová hrubá mesačná mzda (do úvahy sa berú iba zamestnanecké poisťovne vzťahy s vymeriavacím základom aspoň vo výške minimálnej mzdy)

TOP1\_podiel – podiel sekcie SK NACE, v ktorej bolo zamestnaných v priemere za sledované obdobie najviac absolventov

TOP1\_nazov – názov sekcie SK NACE, v ktorej bolo zamestnaných v priemere za sledované obdobie najviac absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP2\_podiel – podiel sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie druhý najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP2\_nazov – názov sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie druhý najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP3\_podiel – podiel sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie\* tretí najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP3\_nazov – názov sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie\* tretí najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

## **4 Úroveň detailu – SR**

### **4.1 Opis datasetu**

Najvyšší stupeň agregácie v podobe dát na úrovni Slovenskej republiky v delení podľa študijných a učebných odborov zabezpečuje odlišný pohľad na prvé kroky na trhu práce. Keďže ide o hrubšiu granularitu údajov, môže to znamenať, že niektoré regionálne špecifiká zostanú nebadané, ale naďalej agregované údaje v takejto podobe poskytujú vhodný zdroj na vytvorenie si základného prehľadu o štruktúre absolventov a ich celkovej

úspešnosti pri vstupe na trh práce na celoštátnej úrovni. Základné princípy pre výpočet ukazovateľov sa uplatňujú rovnako aj v tomto prípade. Naďalej platí, že hodnoty v databáze sú vypočítané ako priemer za január – jún 2021. Z údajov dokážeme identifikovať počet absolventov konkrétneho študijného odboru na úrovni celej Slovenskej republiky. Ďalej je možné určiť status absolventov (pracujúci, materská/rodičovská, pokračovanie v štúdiu, nezamestnaní, prípadne mimo evidencie). Súčasťou zverejnených údajov je priemerná a mediánová hrubá mzda. Oba ukazovatele vychádzajú z údajov za celé sledované obdobie a do výpočtu vstupovali iba zamestnanecké poisťovne vzťahy, ku ktorým prislúchala hrubá mesačná mzda aspoň na úrovni minimálnej mzdy (623 eur za rok 2021). Ukazovatele TOP1\_podiel až TOP3\_nazov prezentujú podiel a názov troch hlavných tried zo štatistickej klasifikácie SK NACE Rev.2, kde absolventi pracovali najčastejšie. Na zabezpečenie ochrany osobných údajov podliehajú zverejnené dáta anonymizácii. V prípade, ak je celkový počet spracovaných osôb na úrovni študijného a učebného odboru nižší ako 5, údaje o stave absolventov budú skryté. Pre takýto postup sme sa rozhodli, aby nebolo možné identifikovať stav konkrétnych absolventov. Anonymizácii podliehajú aj mzdové ukazovatele spolu s podielmi odvetví, kde sa absolventi najčastejšie uplatnili. V prípade, ak je priemerný počet zamestnaných osôb menší ako 3, tak mzdové ani odvetvové ukazovatele nebudú zverejnené.

#### **4.2 Popis premenných súboru o študijných / učebných odboroch na úrovni SR**

Dátový súbor (Úroveň detailu –SR) pozostáva z nasledovných 24 premenných:

##### **Identifikácia študijného a učebného odboru:**

sauo\_kod – kód študijného a učebného odboru

nazov\_sauo – názov študijného a učebného odboru

##### **Stav absolventov:**

spracované osoby – celkový počet absolventov

muz – celkový počet absolventov mužského pohlavia

zena – celkový počet absolventov ženského pohlavia

pracujuce\_osoby – počet pracujúcich absolventov, výpočet = zamestnane\_osoby + osoby\_na\_dohode + osoby\_szco.

zamestnane\_osoby – počet zamestnaných absolventov (rozhodujúci je zamestnanecký poisťový vzťah v SP)

osoby\_na\_dohode – počet absolventov pracujúcich podľa dohody o prácach vykonávaných mimo pracovného pomeru

osoby\_szco – počet absolventov vykonávajúcich samostatne zárobkovú činnosť (identifikovaní z Registra právnických osôb)

osoby\_mater\_rodic – počet absolventov na materskej/rodičovskej dovolenke

pokrac\_studium\_SS – počet absolventov pokračujúcich v štúdiu na strednej škole na Slovensku

pokrac\_studium\_VS\_SR – počet absolventov pokračujúcich v štúdiu na vysokej škole na Slovensku

pokrac\_studium\_zahr – počet absolventov pokračujúcich v štúdiu na vysokej škole v zahraničí

nezamestnane\_osoby – celkový počet absolventov zaregistrovaných ako uchádzači o zamestnanie

z\_toho\_nezamestnane\_osoby\_na\_dohode – počet absolventov zaregistrovaných ako uchádzač o zamestnanie a súčasne pracujúcich podľa dohody o prácach vykonávaných mimo pracovného pomeru

mimo\_evidencie – počet osôb mimo evidencie (napr. pracujúci/ neidentifikovaní študujúci v zahraničí, dobrovoľne nezamestnaní)

#### **Mzdové ukazovatele a odvetvové uplatnenie:**

priem\_hmesmzda – priemerná hrubá mesačná mzda (do úvahy sa berú iba zamestnanecké poisťovacie vzťahy s vymeriavacím základom aspoň vo výške minimálnej mzdy)

medianova\_mzda – mediánová hrubá mesačná mzda (do úvahy sa berú iba zamestnanecké poisťovacie vzťahy s vymeriavacím základom aspoň vo výške minimálnej mzdy)

TOP1\_podiel – podiel sekcie SK NACE, v ktorej bolo zamestnaných v priemere za sledované obdobie najviac absolventov

TOP1\_nazov – názov sekcie SK NACE, v ktorej bolo zamestnaných v priemere za sledované obdobie najviac absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP2\_podiel – podiel sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie druhý najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP2\_nazov – názov sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie druhý najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP3\_podiel – podiel sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie\* tretí najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

TOP3\_nazov – názov sekcie SK NACE, v ktorej bol zamestnaný v priemere za sledované obdobie\* tretí najvyšší počet absolventov (do úvahy sa brali všetky zamestnané osoby)

## **5 Záver**

Tento dokument poskytol základný rámec o zverejnených dátových súboroch ku komentáru Absolventi v čase pandémie. Záber dátových súborov je kohorta absolventov školského roku 2019/2020 a ich prvé kroky na trhu práce. Údaje sú prezentované na troch stupňoch agregácie. Najdetailnejšie údaje sú k dispozícii na úrovni školy, kde identifikujeme kombináciu škola a študijný a učebný odbor. Medzistupeň predstavuje agregácia na úrovni krajov, kde vytvárame kombinácie kraj a študijný a učebný odbor. Tretí súbor obsahuje dáta na úrovni Slovenskej republiky, kde identifikujeme celkový počet absolventov v delení podľa študijného a učebného odboru. Každý z prezentovaných dátových súborov disponuje určitými výhodami, resp. nevýhodami voči ostatným datasetom. Plusy, prípadne mínusy využitia dátových súborov s vyššou, resp. nižšou agregáciou priamo súvisia s procesom anonymizácie, t.j. nezverejňovania údajov za také príliš malé skupiny študentov, v prípade ktorých by hrozila deanonymizácia dát. Čím sa dostávame na nižšiu úroveň detailu sledovania absolventov (kraj, resp. krajina), tým menej buniek je potrebné ponechať nevyplnených z dôvodu anonymizácie. S tým úzko súvisí aj nárast schopnosti vytvoriť si lepšiu mozaiku o nástupe absolventov na trh práce. Súčasne si treba uvedomiť, že čím vyšší stupeň agregácie používame, tým viac dochádza k potláčaniu špecifik na úrovni škôl, resp. krajov.

