Dokumentace datové sady (DS)

(Aktualizováno ke dni 23. 3. 2023)

Název DS:

***Průměrná hrubá měsíční mzda a medián mezd v krajích***

## Popis datové sady:

Datová sada obsahuje časovou řadu průměrných hrubých měsíčních mezd zaměstnanců (přepočtených na plný úvazek) a mediánu průměrných měsíčních mezd v České republice podle krajů a podle pohlaví. Jedná se o roční údaje od roku 2011.

*Metodické poznámky*

**Průměrná hrubá měsíční mzda** představuje podíl mezd bez ostatních osobních nákladů připadající na jednoho zaměstnance evidenčního počtu za měsíc. Do mezd se zahrnují základní mzdy a platy, příplatky a doplatky ke mzdě nebo platu, odměny, náhrady mezd a platů, odměny za pracovní pohotovost a jiné složky mzdy nebo platu, které byly v daném období zaměstnancům zúčtovány k výplatě. Nezahrnují se náhrady mzdy nebo platu za dobu trvání dočasné pracovní neschopnosti nebo karantény placené zaměstnavatelem.

Jedná se o hrubé mzdy, tj. před snížením o pojistné na veřejné zdravotní pojištění a sociální zabezpečení, zálohové splátky daně z příjmů fyzických osob a další zákonné nebo se zaměstnancem dohodnuté srážky.

**Průměrná hrubá měsíční mzda za celé národní hospodářství** zahrnuje údaje za všechny zaměstnance. Je uvedena na přepočtené počty zaměstnanců, zohledňuje tak délku pracovních úvazků zaměstnanců.

**Medián mezd**

Medián mezd představuje hodnotu mzdy zaměstnance uprostřed mzdového rozdělení; to znamená, že polovina hodnot mezd je nižší a druhá polovina je vyšší než medián. Na rozdíl od průměrné mzdy, která je vypočtena na základě podkladů z podnikového výkaznictví, je medián nutné odvozovat ze statisticko-matematického modelu na podkladě výběrového šetření, protože podnikové výkazy obsahují jen agregované údaje za celý podnik nebo organizaci.

Údaje v datové sadě jsou výstupem statistických úloh strukturálních statistik.

Výsledky strukturální statistiky produkuje Český statistický úřad ve spolupráci s Ministerstvem práce a sociálních věcí ČR (MPSV) od roku 1996. Jsou jí zjišťovány mzdy jednotlivých zaměstnanců a nikoli celkové objemy na úrovni podniků či organizací. Jsou podrobně zjišťovány složky hrubého výdělku a také důležité personální údaje o zaměstnanci jako např. pohlaví, vzdělání, věk. Získává se tak statistika velmi detailní, která slouží podrobným analýzám trhu práce a jeho vývoje. Do hrubých mezd se ve strukturální statistice počítají všechny mzdy za práci včetně  odměn a dalších platů, dále veškeré náhrady mzdy za neodpracovanou dobu (dovolenou, svátky, překážky v práci apod.) a odměny za pracovní pohotovost za celý rok. Průměrná mzda zaměstnance v daném roce je vypočtena poměřením s jeho placenou dobou, tedy počtem měsíců, za které mzdu či náhradu mzdy skutečně pobíral, odečtena je tedy doba nemocí a dalších neplacených nepřítomností v práci za daný rok. Vypočtená průměrná hrubá měsíční mzda (v Kč) tak co nejpřesněji vypovídá o srovnatelných mzdových úrovních v různých zaměstnáních (pracovních místech) při přesně zjištěném objemu placené doby. Takto vypočtená průměrná mzda však není a nemůže být shodná s průměrnou mzdou zjišťovanou z podnikového výkaznictví ČSÚ, kde je celkový objem mezd bez ostatních osobních nákladů poměřován evidenčním počtem zaměstnanců podniku, v němž jsou však zahrnuti i zaměstnanci nemocní nebo s neplacenou nepřítomností kratší než 4 týdny. Další rozdíly mezi mzdovou úrovní ve srovnání s jinými statistickými zdroji mohou plynout (kromě vlivu neplacených absencí a odlišného základního souboru šetření) z faktu, že do výsledků strukturální statistiky se nezahrnují zaměstnanci s týdenním úvazkem kratším než 30 hodin. Výsledky strukturální statistiky, protože plynou z částečně výběrového šetření, jsou zatíženy výběrovou chybou. Dále je faktem, že některé oslovené jednotky neposkytly požadované údaje nebo že některé záznamy musely být vyřazeny ze zpracování pro chybovost, a tudíž mohlo dojít k drobným zkreslením. Nakonec je třeba si uvědomit, že výsledky jsou tak kvalitní, jak kvalitní jsou podkladové databáze podniků a organizací, ze kterých je statistika čerpána, to se týká především detailních klasifikací zaměstnání či dosaženého stupně vzdělání zaměstnance.

## Struktura dat:

Datovou sadu představuje CSV soubor, jednotlivé položky (tvořící sloupce) jsou zprava i zleva ohraničeny znakem " (uvozovky), oddělovačem položek je čárka, oddělovačem desetinné části tečka.

Každý statistický údaj tvoří jeden řádek souboru, zjištěná hodnota je ve sloupci *hodnota*.

Statistický údaj je definován z hlediska věcného (statistická proměnná=ukazatel), časového (referenční období) a územního (území).

| SLOUPEC | VYZNAM | POZNAMKA |
| --- | --- | --- |
| idhod | unikátní identifikátor údaje Veřejné databáze ČSÚ | využije se v případě dotazu ke konkrétnímu údaji |
| hodnota | zjištěná hodnota | v numerickém formátu,  v případě, že se jedná o důvěrný údaj, je sloupec prázdný |
| stapro\_kod | kód statistické proměnné | v této datové sadě pouze 5958 |
| spkvantil\_cis | kód číselníku pro kvantil | v datové sadě použity číselník 7636 |
| spkvantil\_kod | kód položky číselníku pro kvantil | položka Q5 odpovídá mediánu,  pokud není vyplněna, jedná se o průměrnou mzdu |
| pohlavi\_cis | kód číselníku pro pohlaví | v datové sadě použit číselník 102 |
| pohlavi\_kod | kód položky číselníku pro pohlaví | pokud není vyplněn, jedná se o údaje za obě pohlaví |
| rok | rok referenčního období | ve formátu RRRR |
| uzemi\_cis | kód číselníku pro referenční území | 97 - Česká republika  100 – kraje |
| uzemi\_kod | kód položky číselníku pro referenční území | kód území v příslušném číselníku |
| stapro\_txt | text statistické proměnné (ukazatele) |  |
| uzemi\_txt | text položky číselníku pro referenční území |  |
| spkvantil\_txt | text položky číselníku pro kvantil | pokud není uveden, jedná se o průměr za všechny zaměstnance |
| pohlavi\_txt | text položky číselníku pro pohlaví | pokud není uveden, jedná se o údaje za obě pohlaví |

## Použité číselníky a referenční údaje:

Pro identifikaci (definování) každého statistického údaje se využívá Statistický metainformační systém ČSÚ (SMS), a to subsystém SMS-UKAZ (statistické proměnné) a SMS-KLAS (statistické číselníky a klasifikace).

V datové sadě jsou použity:

**Seznam statistických proměnných** - <https://www.czso.cz/csu/czso/statisticke-promenne-ukazatele>

|  |  |
| --- | --- |
| kod | TEXT |
| 5958 | Průměrná hrubá mzda zaměstnance |

**Číselník krajů** – kód číselníku ČSÚ 100

CSV: <https://apl.czso.cz/iSMS/do_cis_export?kodcis=100&typdat=0&cisjaz=203&format=2>

XML: <https://apl.czso.cz/iSMS/do_cis_export?kodcis=100&typdat=0&cisjaz=203&format=0>

**Číselník pohlaví -**  kód číselníku ČSÚ 102

CSV: <https://apl.czso.cz/iSMS/do_cis_export?kodcis=102&typdat=0&cisjaz=203&format=2>

XML: <https://apl.czso.cz/iSMS/do_cis_export?kodcis=102&typdat=0&cisjaz=203&format=0>

V datové sadě použity tyto položky:

|  |  |  |
| --- | --- | --- |
| Kód číselníku | Položka číselníku | Text |
| 102 | 1 | muž |
| 102 | 2 | žena |

**Časové vymezení**

Časová řada začíná rokem 2011 a obsahuje roční údaje.