

Portál otevřených dat ČSSZ

Tento dokument detailně popisuje strukturu a obsluhu Portálu otevřených dat ČSSZ.

Úvodní stránka

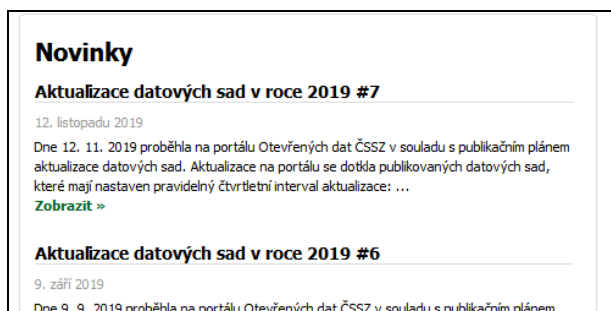
Po spuštění portálu na adrese <https://data.cssz.cz> se zobrazí úvodní stránka, kde v horní části jsou uvedeny v menu tři základní volby (viz Obrázek 1):



Obrázek 1 Hlavní menu Portálu otevřených dat ČSSZ

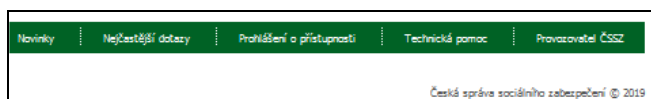
1. Volba **Základní informace** vrací na úvodní stránku portálu.
2. Volba **Katalog otevřených dat** přepíná do katalogu, který obsahuje seznam všech publikovaných datových sad (podrobněji viz níže).
3. Volba **Podmínky užití** definuje pravidla a podmínky, za kterých lze provádět zpracování publikovaných dat.

V pravé části obrazovky se nachází sloupec **Novinky**, ve kterém jsou uvedeny informace formou krátkých zpráv o provedených aktualizacích a změnách, které byly na portále provedeny.



Obrázek 2 Novinky

V zápatí stránky (viz Obrázek 3) jsou umístěny odkazy na obecné zdroje informací k portálu otevřených dat ČSSZ.



Obrázek 3 Zápatí úvodní stránky



Volba **Novinky** obsahuje odkaz na úplný seznam (historii) publikovaných novinek, kde můžete vidět průběh aktualizace stávajících a publikace nových datových sad. Volba **Nejčastější dotazy** vás dovede k seznamu dotazů a našich odpovědí na ně. Zde můžete nalézt vysvětlení k některým vašim otázkám. **Prohlášení o přístupnosti** obsahuje specifikaci základních požadavků na hardware, internetové připojení, operační systémy a webové prohlížeče. V neposlední řadě je zde uvedena i informace o cookies, jejich využívání a možnostech nastavení v jednotlivých prohlížečích. V případě, že zjistíte nějaké technické potíže při přístupu k datovým sadám, jejich stahování a podobně (technické dotazy, nikoliv věcné k obsahu dat), pak využijte odkaz **Technická pomoc**, který vás přesměruje na stránku technické podpory ke službám ČSSZ, kde také najdete veškeré potřebné informace o způsobu kontaktování Call centra technické podpory. Pro úplnost je na závěr uveden odkaz na webové stránky ČSSZ (<https://www.cssz.cz>) – **Provozovatel ČSSZ**.

Katalog otevřených dat

Tato stránka obsahuje seznam všech datových sad, které jsou na portále otevřených dat publikované (viz Obrázek 4).

ČESKÁ SPRÁVA SOCIÁLNÍHO ZABEZPEČENÍ
Otevřená data

Základní informace | **Katalog otevřených dat** | Podmínky užití

Katalogové záznamy datových sad

Filtrovat podle klíčového slova: [call centrum \(1\)](#) [diagnózy \(2\)](#) [doba řízení \(1\)](#) [důchodová ročenka \(17\)](#) [důchodové pojištění \(23\)](#) [exekeuce \(1\)](#) [exekeuce srážka \(1\)](#) [fluktuaace \(1\)](#) [informativní osobní list důchodového pojištění \(1\)](#) [insolvenční řízení \(1\)](#) [invalidita \(4\)](#) [kontrola \(1\)](#) [kraje \(6\)](#) [lékařská posudková služba \(1\)](#) [námitky \(2\)](#) [nemocenské pojištění \(6\)](#) [okresy \(4\)](#) [osobní účet \(1\)](#) [osvě \(4\)](#) [podání \(1\)](#) [pohledávky \(2\)](#) [pojištění \(3\)](#) [pracovní neschopnost \(2\)](#) [průměrný plat \(1\)](#) [příslušnost k českým právním předpisům \(1\)](#) [sociální zabezpečení \(1\)](#) [svobodný přístup k informacím \(1\)](#) [systemizovaná místa \(1\)](#) [věk \(4\)](#) [výdaje \(1\)](#) [výměňovací základ \(3\)](#) [vyplacené dávky \(7\)](#) [výplata důchodů \(4\)](#) [výše dávek \(5\)](#) [zaměstnanci \(5\)](#) [zaměstnavatelé \(4\)](#) [zánik výplaty \(2\)](#)

Číslo v závorce udává počet publikovaných datových sad, které obsahují dané klíčové slovo.

statistika x

Datová sada	Vizualizace
Doba řízení o námitkách	<input checked="" type="checkbox"/> Doba řízení
Kontroly plnění povinností zaměstnavatelů v oblasti sociálního zabezpečení	<input checked="" type="checkbox"/> Počty kontrol plnění povinností zaměstnavatele
Měsíční výše důchodů	<input checked="" type="checkbox"/> Měsíční výše důchodů podle roku, kvantilu a druhu důchodu
Nejčastější příčiny vzniku invalidity	<input checked="" type="checkbox"/> Nejčastější diagnózy při posuzování invalidity ve sledovaném roce
Počet důchodců s výplatou do ciziny v členění dle států	<input checked="" type="checkbox"/> Počet důchodců s výplatou do zahraničí <input checked="" type="checkbox"/> Počet důchodců s výplatou do zahraničí v daném státě <input checked="" type="checkbox"/> Státy s největším počtem důchodců s výplatou do zahraničí v roce <input checked="" type="checkbox"/> Počet důchodců s výplatou do zahraničí v daném státě - mapa světa

Obrázek 4 Katalog otevřených dat

Hlavní část stránky tvoří dva sloupce - v prvním je abecedně setříděný seznam datových sad, ve druhém jsou pak uvedeny vizualizace pro datovou sadu. Až na výjimky je vizualizace u všech datových sad a některé jich obsahují i více. Ikona před názvem vizualizace naznačuje v jaké podobě je provedena. Může se jednat o spojnicové, sloupcové nebo koláčové grafy, mapové vizualizace (okresy, kraje, státy) nebo tabulky. Detailnější popis vizualizací je uveden dále v textu.

V horní části je pak uveden seznam klíčových slov, podle kterých lze v seznamu filtrovat. U každého slova je v závorce uveden počet datových sad, ke kterým bylo zvolené klíčové slovo přiřazeno. Na uvedeném obrázku byl nastaven filtr na klíčové slovo „statistika“ (je uvedeno v buňce nad seznamem, kliknutím na křížek v buňce se vybraný filtr zruší). Lze zde provádět i vícenásobný výběr klíčových slov. Chcete-li například vyhledat, ve kterých datových sadách se nachází statistické údaje o dávkách na invaliditu, můžete postupně vybrat slova „statistika“, „vyplacené dávky“ a „invalidita“ a do seznamu se vybere jediná datová sada „Počet vyplacených invalidních důchodů v ČR dle skupin diagnóz“.

Na konci stránky je pak uveden počet zobrazených datových sad – nebude-li proveden žádný výběr, zobrazí se seznam všech datových sad.



Kliknutím na název datové sady (v prvním sloupci seznamu) se přesunete na tzv. katalogizační záznam, který popisuje základní metadata o každé sadě (viz Obrázek 5). Metadata rozumíme informace o údajích, které jsou v datové sadě obsaženy.

**ČESKÁ SPRÁVA SOCIÁLNÍHO ZABEZPEČENÍ**
Otevřená data

Základní informace | Katalog otevřených dat | Podmínky užití

Počet nově přiznaných důchodů v České republice dle měsíční výše důchodu

Počet nově přiznaných důchodů v České republice podle roku, druhu důchodu, měsíční výše důchodu a pohlaví

Periodicita aktualizace	Roční
Územní pokrytí	Česká republika
Téma datové sady	Hospodářství a finance
Klasifikace dle EuroVoc	důchodové pojištění podmínky pro odchod do důchodu
Jméno a e-mail kurátora dat	Ing. Jiří Šunka
Odkaz na dokumentaci datové sady	https://data.cssz.cz/documentation/nove-priznane-duchody-dle-vyse-duchodu
Dotčené časové období od-do	01.01.2008 - 31.12.2018
Název a URL poskytovatele dat	Česká správa sociálního zabezpečení

Datové zdroje

Data	Schéma	Podmínky užití
Data (CSV) 1,5 MB	Schéma (JSON) 1,5 kB	podmínky užití
Data (RDF/TriG) 17,2 MB	Schéma (RDF/TriG) 24,7 kB	podmínky užití

Data v RDF jsou přístupná i přes SPARQL endpoint, který je na URL <https://data.cssz.cz/sparql>

Pro získání výběrů dat z datových sad je připraven jednoduchý **formulář**, který umožňuje zadávání dotazů ve SPARQL jazyku.

Metadata o datové sadě v RDF (TriG) 12,2 kB

Vizualizace

 **Počet nově přiznaných důchodů dle měsíční výše důchodu**

[důchodová ročenka](#) [důchodové pojištění](#) [statistika](#) [výše dávek](#)

Obrázek 5 Katalogizační záznam

Všechny záznamy mají jednotnou strukturu a obsahují následující informace:

- **Název datové sady** (příklad z obrázku: *Počet nově přiznaných důchodů v České republice dle měsíční výše důchodu*)
- **Popis datové sady** – podrobnější popis (příklad z obrázku: *Počet nově přiznaných důchodů v České republice podle roku, druhu důchodu, měsíční výše důchodu a pohlaví*)
- **Periodicita aktualizace** – určuje v jakém intervalu je prováděna aktualizace datové sady, provádí se buď ročně nebo čtvrtletně, případně ad hoc (příklad z obrázku: *Jednou ročně*)
- **Územní pokrytí** – vyjadřuje územní jednotku, ke které se publikovaná data vztahují – stát, kraj, okres (příklad z obrázku: *Česká republika*), odkaz obsahuje prolink na číselník ČUZK
- **Téma datové sady** – definuje oblast, které se publikovaná data týkají
- **Klasifikace dle EUROVOC** – slovně uvádí hlavní kategorie podle mezinárodního katalogu činností EU - EUROVOC (příklad z obrázku: *důchodové pojištění, podmínky pro odchod do důchodu*)
- **Jméno a e-mail kurátora dat** – kliknutím na jméno se otevře nastavený poštovní klient a připraví se prázdný e-mail s předvyplněnou adresou pro komunikaci v oblasti otevřených dat ČSSZ (příklad z obrázku: *Ing. Jiří Šunka*)
- **Odkaz na dokumentaci datové sady** – dokumentace je určena pro uživatele a má sloužit k seznámení se s charakterem dat a jejich významem, případně s upřesněním obsahu; každá



dokumentace datové sady se skládá z několika částí: obsahuje podrobný popis publikovaných dat, jejich strukturu, význam jednotlivých atributů, seznam použitých číselníků a zkratk (příklad z obrázku: <https://data.cssz.cz/documentation/nove-priznane-duchody-dle-vyse-duchodu>)

- **Dotčené časové období** – určuje časový interval, ke kterému se vztahují publikovaná data (příklad z obrázku: 01. 01. 2008 - 31. 12. 2018)
- **Název a URL poskytovatele dat** – kliknutím se přepne na webové stránky ČSSZ

Ke každé datové sadě je rovněž uveden seznam tzv. datových zdrojů, které obsahují vlastní data datové sady v různých formátech a popis jejich struktury. Popis každého zdroje je uveden ve třech sloupcích:


- **Data** – obsahuje odkaz na stažení souboru a informaci o jeho formátu; otevřená data ČSSZ jsou vždy publikována ve dvou formátech – RDF/Trig nebo CSV (u tohoto formátu je nutno upozornit na to, že je publikován podle mezinárodní definice typu souboru CSV, kdy je jako oddělovač jednotlivých dat použita čárka, nikoliv středník, jak je zvykem v českém prostředí; tato skutečnost se projeví zejména při nahrávání souboru do Excelu a popis postupu je uveden v části Nejčastější dotazy)
- **Schéma** – obsahuje odkaz na popis struktury datové sady daného formátu opět buď ve formátu RDF/Trig nebo JSON pro CVS, tyto soubory jsou určeny zejména pro strojové zpracování
- **Podmínky užití** – obsahuje odkaz na dokument stanovující podmínky, které je nutno při zpracovávání a dalším využívání publikovaných otevřených dat dodržovat.

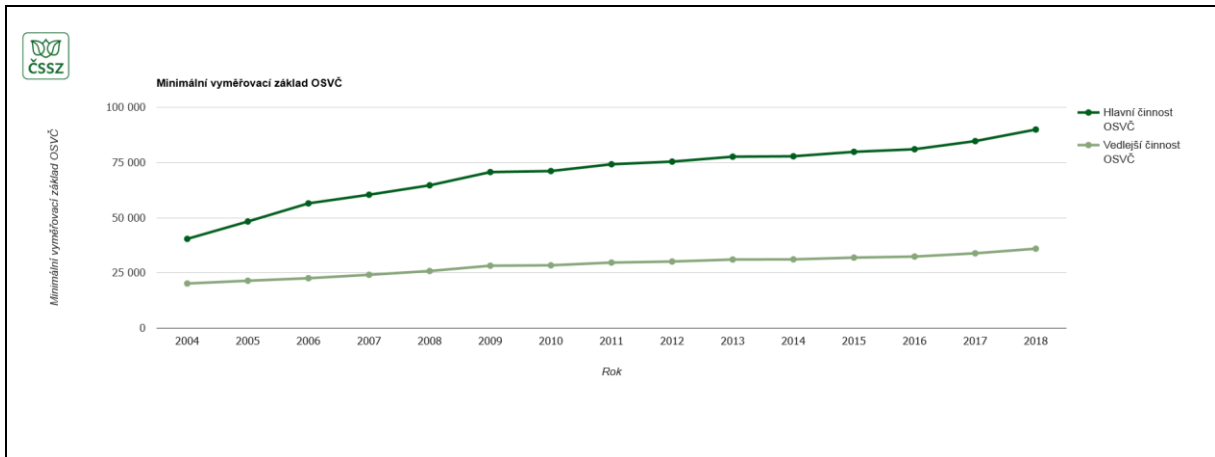
Pro datové specialisty je zde uveden i popis metadat o datové sadě v RDF formátu a odkazem na formulář pro zadávání výběrových SPARQL dotazů přímo do databáze, obsahující data všech publikovaných datových sad. Dostávají tím možnost vytvářet vlastní dotazy a vazby mezi publikovanými daty a provádět nad nimi sofistikované analýzy, zejména po propojení s daty jiných poskytovatelů (jako jsou např. Český statistický úřad, Český úřad zeměměřický a katastrální).

Vizualizace

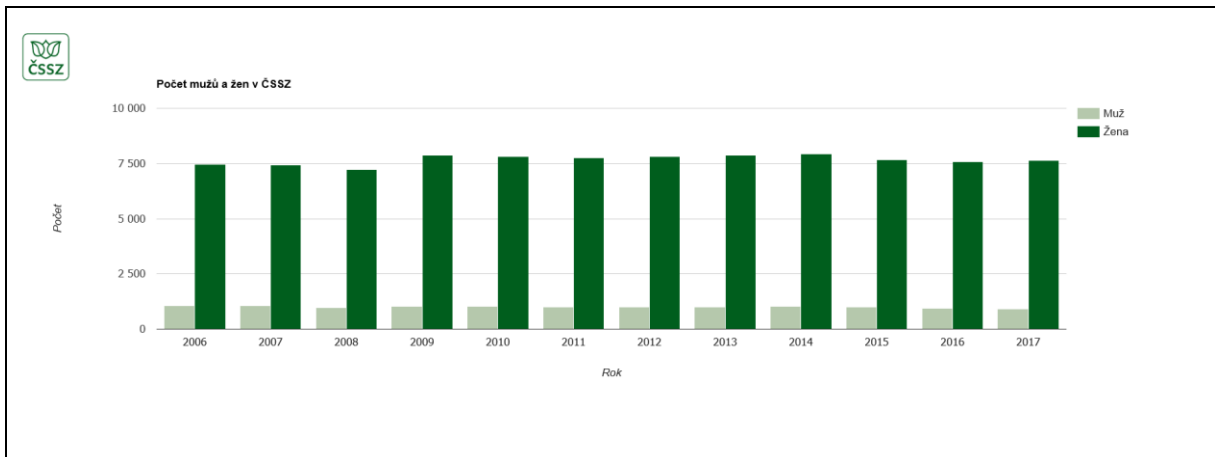
Dříve než je budeme detailně popisovat, objasníme si zjednodušeně dva pojmy z teorie grafů – *měření* a *dimenze*. První představuje zjišťovanou hodnotu a druhý parametry (podmínky), pro které je měření zjišťováno. Například pro počet vyplacených starobních důchodů ženám v roce 2018 je měřením počet důchodů a dimenzemi jsou typ důchodu (starobní), pohlaví (žena) a rok výplaty (2018). Při práci s daty se často dimenze ztotožňuje s parametry výběru (filtry), které se definují pro výběr dat k zobrazení.

V katalogu otevřených dat ČSSZ (nebo přímo na katalogizačním záznamu) je u každé datové sady uveden seznam vizualizací, které má tato datová sada k dispozici. Pro prezentaci dat bylo vybráno několik základních typů grafů, jejichž použití vyplývá z typu dat a ze zvoleného pohledu na ně. Ikona před názvem vizualizace naznačuje v jaké podobě je vizualizace provedena. Jedná se o tyto typy:

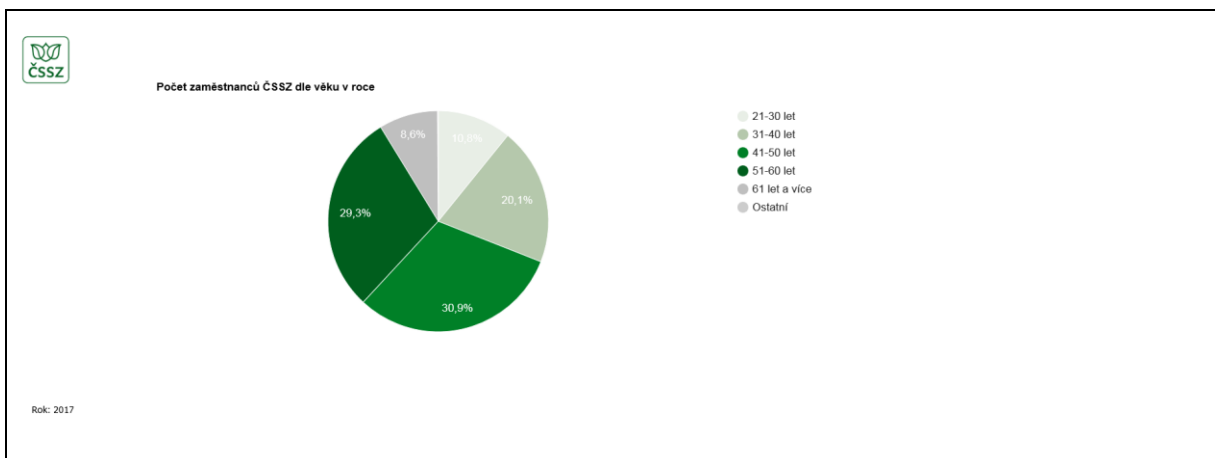
 **spojnicové** – tento typ se používá k prezentaci časové řady a jednoduchou formou zobrazuje vývoj sledované veličiny v čase; typicky je na ose x časová veličina (tedy dimenze, např. rok, čtvrtletí, měsíc) a na ose y měřená hodnota (např. počet); v grafu může být uvedeno několik spojnic, které odpovídají další dimenzi; příkladem může být například graf „*Minimální vyměřovací základ OSVČ*“, kde kromě roku je další dimenzí typ činnosti OSVČ a pro její obě hodnoty (hlavní činnost a vedlejší činnost) jsou pak vykresleny spojnice.



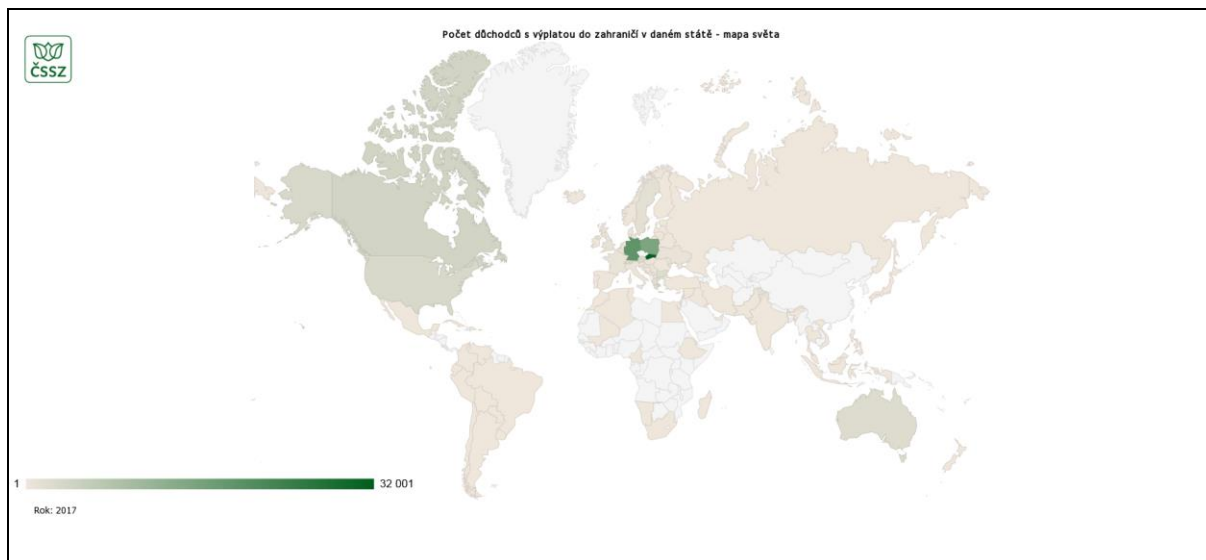
sloupcové – umožňují vzájemné srovnání výsledků měření několika dimenzí současně ve formě sloupců; příkladem může být například graf „Počet mužů a žen v ČSSZ“



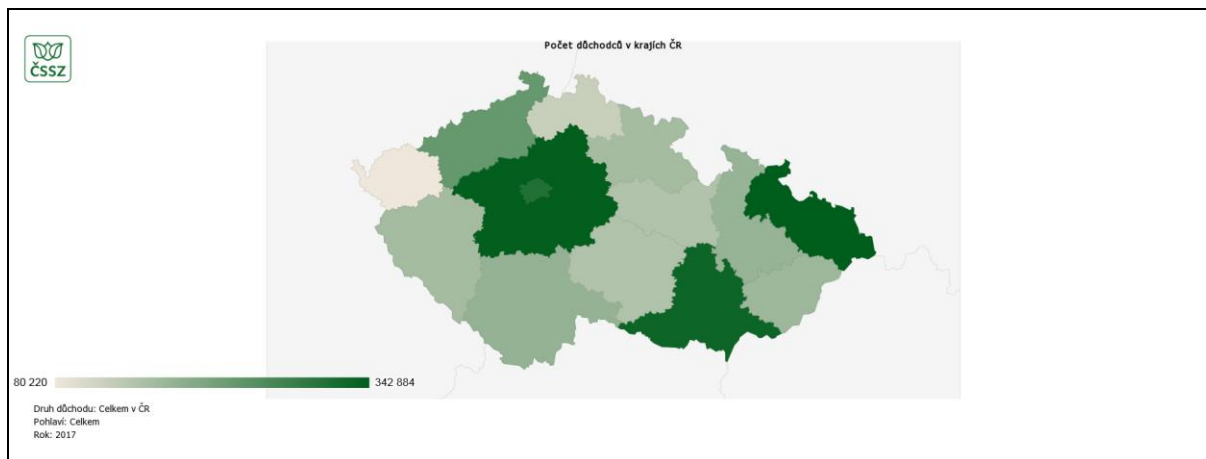
koláčové – jsou využity tam, kde je vhodné nebo nutné vyjádřit podíl výsledku měření pro jednotlivé dimenze na celku; jako příklad můžeme uvést graf „Počet zaměstnanců ČSSZ podle věku“



mapové vizualizace (státy) – jsou použity v případech, kdy jednou z dimenzí je stát, například „Počet důchodců s výplatou do zahraničí v daném státě“; kdy je měřená veličina podle hodnoty vyjádřena různou barvou pro jednotlivé státy a je znázorněna v dolní části grafu ve formě barevné škály od minimální po maximální zobrazenou hodnotu



mapové vizualizace (kraje, okresy) – tuto vizualizaci používáme na prezentaci měřených hodnot na úrovni jednotlivých krajů nebo okresů České republiky, jako příklad lze uvést „Počet důchodců v krajích ČR“;



tabulky – tato forma sice není přímo vizualizací v tom smyslu, jako ostatní a je použita tam, kde jiná forma by byla nevhodná; typicky se využívá u číselníků pro vyplňování tiskopisů, například „Číselník pohlaví“



Kód položky	Název položky	Zkrácený název položky	Platnost od	Platnost do
0	nezjištěno	0	1.1.2009	
1	mužské	M	1.1.2009	
2	ženské	Ž	1.1.2009	

Aby byla umožněna aktivní analytická práce s vizualizacemi, je většina z nich doplněna o filtry, pomocí nichž lze nastavit výběrová kritéria (dimenze) pro zobrazovaná data. Filtry jsou uvedeny v záhlaví grafu a jsou vytvořeny rozbalovacími seznamy, kdy po kliknutí na šipku v pravé části se zobrazí hodnoty dimenzí, ze kterých lze vybírat (viz Obrázek 6). Lze samozřejmě provádět nastavení libovolné kombinace předpřipravených výběrových podmínek, a v případě, že nelze graf vykreslit, protože např. pro zadanou kombinaci data neexistují, zobrazí se informační zpráva.

Obrázek 6 Filtry a ovládání vizualizace

Kromě možnosti nastavení filtrů je vždy v záhlaví zobrazena trojice tlačítek:

Zobrazit graf nebo **Zobrazit mapu** – při otevření stránky s vizualizací se automaticky načte graf nebo mapa pro výchozí nastavení parametrů výběru; pokud provedete změnu parametrů, je nutno pro vykreslení grafu nebo mapy s novými parametry stisknout toto tlačítko.

Zobrazit tabulku – pokud si chcete prohlédnout data, ze kterých byl graf sestaven, tato volba vám je zpřístupní ve formě tabulky; takto získaná data můžete standardně označit, zkopírovat (Ctrl+C) a následně vložit do Excelu nebo Wordu (Ctrl+V) či jiných aplikací a zpracovat dle své potřeby; data lze rovněž stáhnout i ve formátu CSV (viz dále).

Export – tato volba nabízí tři možnosti exportu:

- **CSV** - všechna data z datové sady jsou obsažena v datovém zdroji, který je ke stažení z katalogizačního záznamu; pokud však požadujete pro svou potřebu pouze podkladová data k zobrazené vizualizaci, můžete si je stáhnout pomocí této volby; vygenerovaný CSV soubor na prvních řádcích obsahuje název vizualizace a nastavené výběrové podmínky, za kterými následuje řádek s názvy sloupců, pod kterým jsou uvedeny měřené hodnoty oddělené čárkou.
- **PNG, PDF** – jsou dvě prakticky shodné volby, které do výstupního souboru vygenerují zobrazený graf s identifikací nastavených výběrových podmínek; podle vybrané varianty je výstupní soubor uložen buď ve formátu PNG (obrázek) nebo ve formátu PDF.



Některé grafy obsahují v zápatí doplňující informace o obsahu dat, případně vysvětlení použitých zkratk, například u vizualizace krajů je uveden seznam kódů pro jednotlivé kraje, u některých je uveden např. seznam zkratk druhu důchodů apod. U spojnicových, sloupcových a koláčových grafů v případě zobrazení více měřených hodnot je v pravé části popsáno přiřazení jednotlivých barev pro prezentované dimenze.

Budete-li se nad grafy nebo mapami pohybovat kurzorem, bude se Vám zobrazovat ve vyznačených bodech spojnicového grafu, sloupcích, výsečích koláčového grafu nebo na vybraném území (stát, kraj, okres) příslušná hodnota měřené veličiny spolu s dimenzemi, které ji určují.

Dynamika a interaktivita vizualizací je zajištěna tím, že se vždy zobrazují na základě aktuálně nastavených výběrových podmínek definujících rozsah použitých dat z databáze, která obsahuje veškerá publikovaná data pravidelně aktualizovaná v souladu s publikačním plánem otevřených dat ČSSZ.

Podmínky užití

Tato stránka obsahuje podmínky užití, které jsou definovány v souladu s doporučenou metodikou ministerstva vnitra.

Pro všechny distribuce všech datových sad v rámci otevřených dat ČSSZ platí pro jednotlivé kategorie následující:

- [Distribuce datové sady neobsahuje autorská díla](#)
- [Distribuce datové sady není autorskoprávně chráněnou databází](#)
- [Distribuce datové sady není chráněna zvláštním právem pořizovatele databáze](#)
- [Distribuce datové sady neobsahuje osobní údaje](#)

Závěr

Rádi bychom Vás, uživatele otevřených dat ČSSZ, požádali o spolupráci, a to v zájmu zvyšování kvality portálu Otevřených dat ČSSZ. Velmi oceníme, když se s námi podělíte a napíšete nám informace o Vašem způsobu využití a výsledcích zpracování publikovaných dat. Uvítáme také jakékoliv další poznatky, připomínky a náměty k této dokumentaci.

Zprávu nám můžete poslat na adresu: opendata@cssz.cz .