Open Data Lab

Workshop #2:

Študijný materiál

*Toto podujatie bolo podporené z dotačnej schémy Ministerstva školstva, vedy, výskumu a športu SR Programy pre mládež na roky 2014 – 2021, ktorú administruje IUVENTA – Slovenský inštitút mládeže.*



Autor: Lukáš Jankovič  
[lukas.jankovic@alvaria.sk](mailto:lukas.jankovic@alvaria.sk)

Obsah

[Vizualizácia dát v PowerBI 3](#_Toc85551446)

[Používateľské prostredie 3](#_Toc85551447)

[Načítanie datasetu 4](#_Toc85551448)

[Práca so stĺpcami a mierkami 5](#_Toc85551449)

[Vlastné stĺpce 6](#_Toc85551450)

[Vytvorenie grafu 7](#_Toc85551451)

[Vytvorenie mapy 9](#_Toc85551452)

[Formátovanie 10](#_Toc85551453)

[Filtre 11](#_Toc85551454)

[Power Query editor 13](#_Toc85551455)

[Spájanie datasetov 14](#_Toc85551456)

[Publikovanie 16](#_Toc85551457)

**Poznámka:**

Ak je v texte spomenuté stlačenie tlačidla myši, myslí sa tým ľavé tlačidlo. V prípade ak ide o pravé tlačidlo, je to v texte uvedené.

**Pojmy**

Prvok – graf, tabuľka, mapa

Vizualizácia – celý súbor (môže obsahovať viac prvkov a viac strán)

# Vizualizácia dát v PowerBI

**Microsoft Power BI** je nástroj na **spracovanie, editovanie a vizualizáciu dát**.

Nástroj je vhodný na spracovanie otvorených dát, rovnako ako iných druhov dát a datasetov.

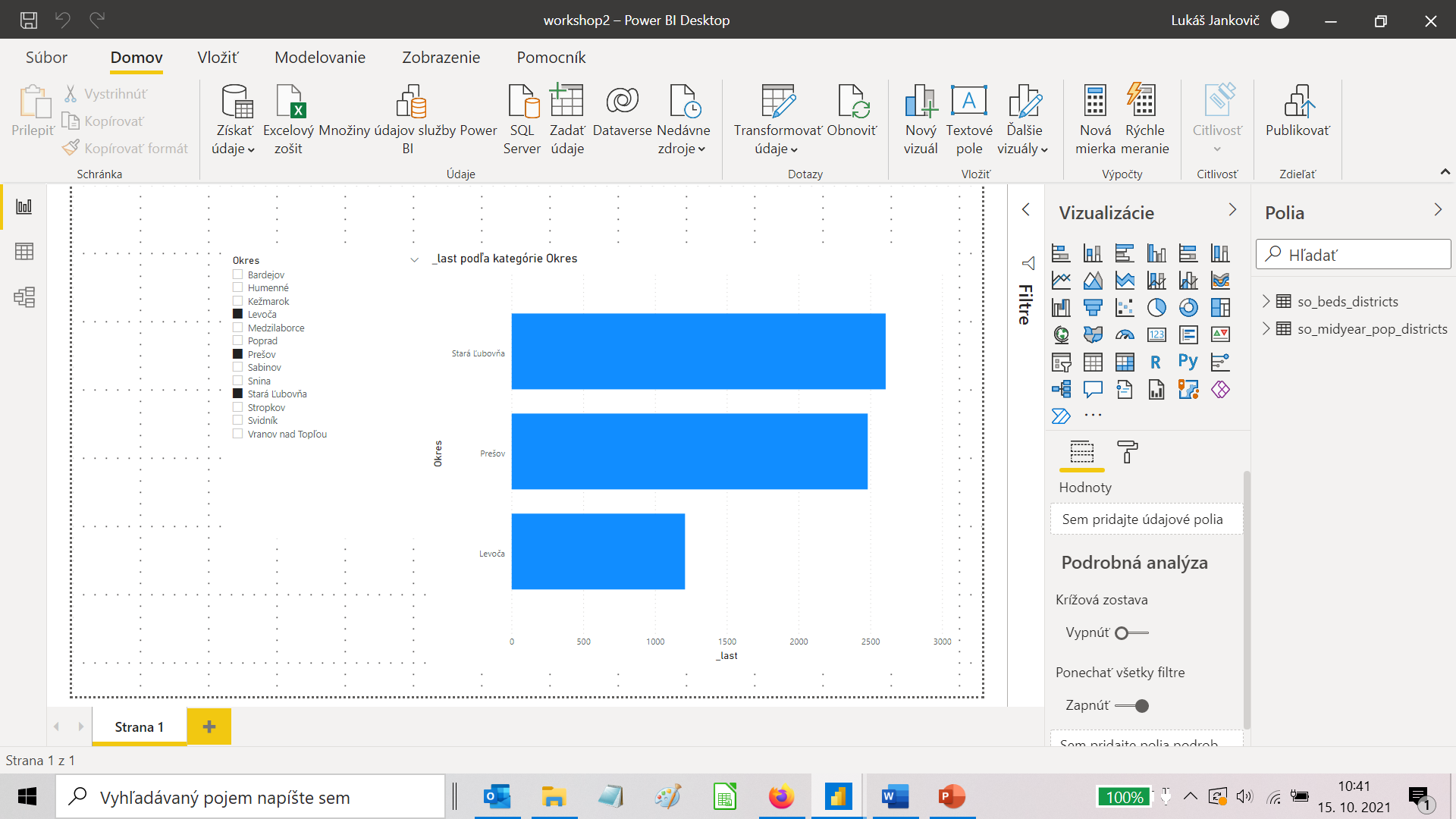
Je dostupný bezplatne pre MS Windows (Windows 8, 8.1, 10).

* Link na stiahnutie: <https://powerbi.microsoft.com/cs-cz/getting-started-with-power-bi/>

Podporuje viacero jazykov, okrem iného aj **slovenčinu**. Pre plnohodnotné použitie je **potrebné sa zaregistrovať** a vytvoriť si profil na platforme PowerBI.

* Link na registráciu: [<https://powerbi.microsoft.com/cs-cz/pricing/>](https://powerbi.microsoft.com/cs-cz/pricing/)

# Používateľské prostredie



Používateľské rozhranie používa Ribbon (karty) známe z MS Office. V ponuke sa nachádzajú karty **Domov, Vložiť, Modelovanie, Zobrazenie a Pomocník**. Karty sa prispôsobujú podľa toho, čo aktuálne používateľ edituje.

Okrem kariet, ktoré obsahujú panely nástrojov tvoria používateľské rozhranie:

* **3 tlačidlá pre voľbu zobrazenia** 
* **hlavný panel – panel zobrazuje obsah, napr. grafy, mapy, iné prvky**,
* panel **Vizualizácie** spolu s nastaveniami,
* panel **Polia**,
* panel **Filtre**.

# Načítanie datasetu

Základné formáty, ktoré vieme spracovať v PowerBI sú:

* CSV,
* JSON,
* XML,
* XLX/XLSX.

Načítanie (import) datasetu je možné cez voľbu **Domov > Získať údaje**.

Pri importe datasetu vo formáte CSV je dôležité nastaviť **správne kódovanie**. Mnoho súborov používa kódovanie **Unicode UTF-8**. Takisto je nutné správne nastaviť oddeľovač, pri CSV súboroch sa štandardne používa čiarka, niekedy sa vyskytuje aj bodkočiarka ;.

Obrázok, na ktorom je text

Automaticky generovaný popis

Načítanie súboru do PowerBI spustíme tlačidlom **Načítať.**

# Práca so stĺpcami a mierkami

**Náhľad údajov** si zobrazíme voľbou Údaje .

Zobrazí sa nám tabuľka s údajmi načítaného datasetu. **Kliknutím pravým tlačidlom myši na záhlavie stĺpcu môžeme stĺpec premenovať alebo vymazať**.

Obrázok, na ktorom je text

Automaticky generovaný popis

Pokiaľ **klikneme na záhlavie stĺpca** môžeme ďalej nastaviť **Typ údajov a Formát**. Pokiaľ vyberieme ako typ údajov číslo, môžeme upraviť počet desatinných miest.

Obrázok, na ktorom je stôl

Automaticky generovaný popis

# Vlastné stĺpce

Vlastné stĺpce sú užitočné v prípade, ak potrebujeme vytvoriť pomocné stĺpce alebo potrebujeme údaje v stĺpcoch upraviť. Vlastné stĺpce môžeme tvoriť v paneli **Polia**.

Panel **Polia zobrazuje načítané datasety a názvy stĺpcov**. Ak sa pri názve stĺpcu zobrazuje symbol Suma , ide o číselný údaj.

Okrem stĺpcov prítomných v datasete si vieme vytvoriť aj **vlastné stĺpce**. **Kliknutím pravým tlačidlom myši na ľubovoľný názov stĺpca a zvolením voľby Nový stĺpec** sa otvorí riadok, do ktorého **môžeme vpísať vzorec** pre nový stĺpec a takisto aj upraviť jeho názov.

Obrázok, na ktorom je stôl

Automaticky generovaný popis

Obrázok, na ktorom je stôl

Automaticky generovaný popis

Pri vytváraní nového stĺpca môžeme pracovať s existujúcimi stĺpcami ale **môžeme používať aj tradičné matematické operátory ako sú zátvorky alebo + - / \*.** Ak chceme do výpočtu zahrnúť existujúci stĺpec, začneme písať jeho názov, PowerBI **nám samo ponúkne názvy stĺpcov**, ktoré na nami napísaný text začínajú.

*Príklady stĺpcov:*

Mojstlpec=so\_beds\_districts[\_2020]+1000

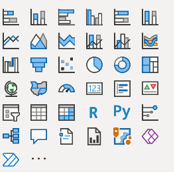
Mojstlpec=so\_beds\_districts[\_2020]+so\_beds\_districts[\_2019]

Mojstplec= so\_beds\_districts[\_2020]/100

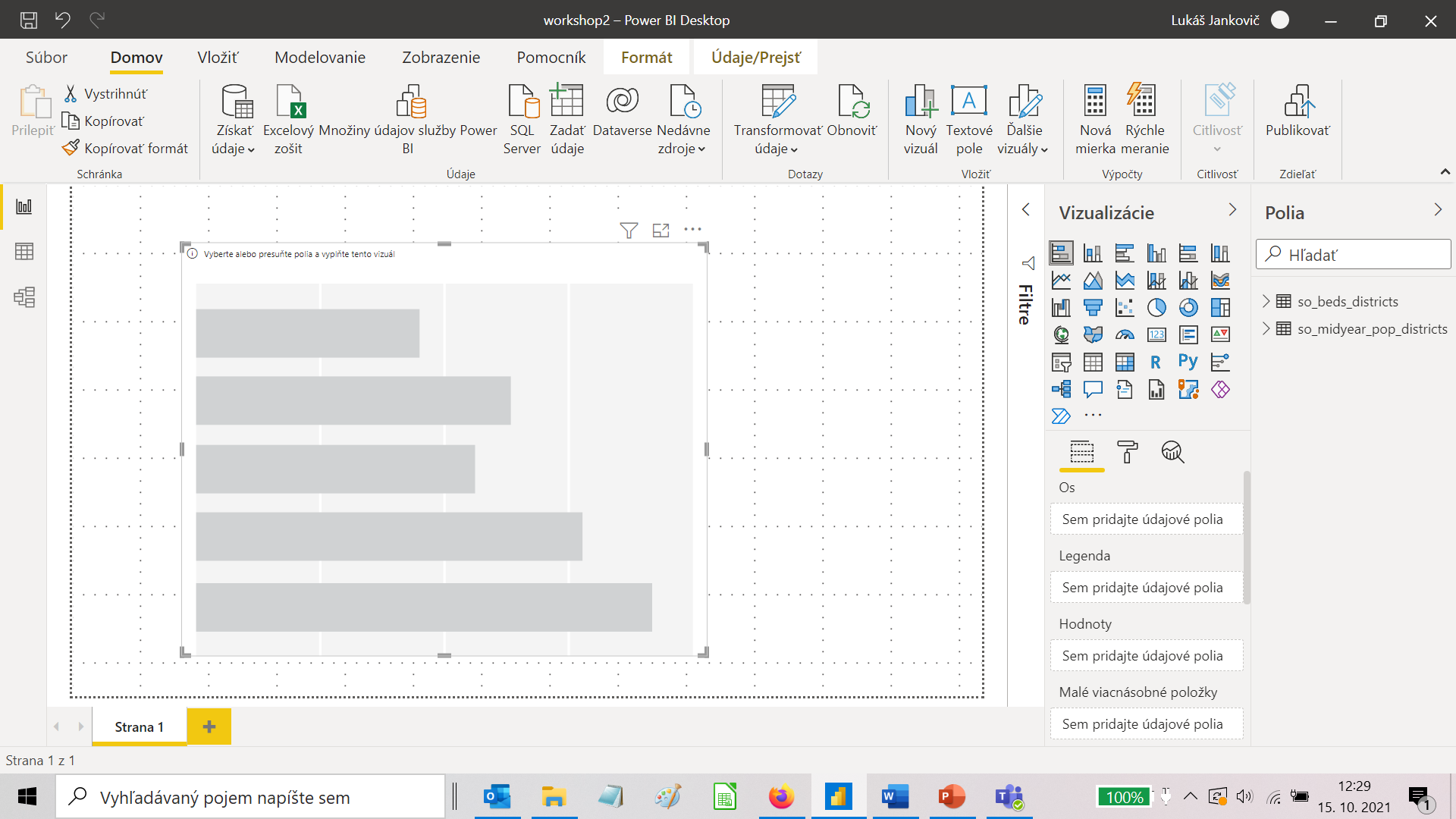
Mojstplec= (so\_beds\_districts[\_2020]-20)/100

# Vytvorenie grafu

Rôzne typy grafov je možné vytvoriť cez **tlačidlá v paneli Vizualizácie**.

****

Klikneme na tlačidlo grafu (na obrázku máme príklad pruhového grafu), graf sa nám vytvorí na hlavnom paneli. Môžeme ho zväčšovať, zmenšovať alebo ľubovoľne presúvať.



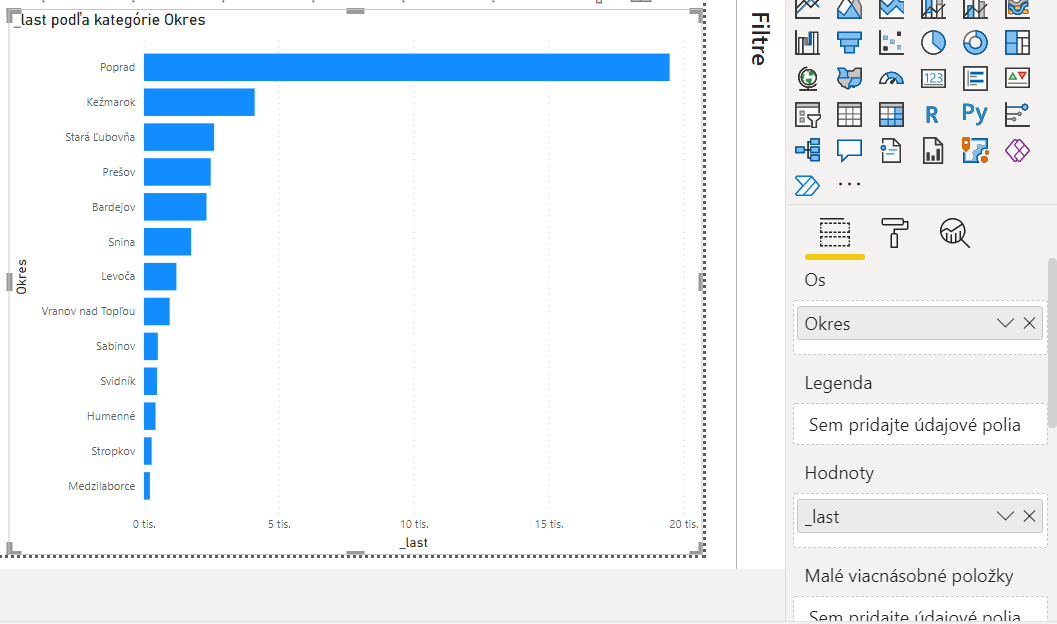
Aby sme určili, čo sa má v grafe zobraziť použijeme nastavenia v paneli Vizualizácia, konkrétne najmä polia:

* **Os** – kategórie,
* **Legenda** – farby,
* **Hodnoty** – číselné údaje.

Povinná položka je **„hodnoty“,** kde zadáme **stĺpec s číselnými údajmi**. Môžeme tiež zadať **„os“**, kde hodnoty zadaného stĺpca určujú pre každý riadok, **do akej kategórie patrí**. Kategórie sa potom zobrazia jedna pod druhou ako jednotlivé pruhy grafu. Pri **„legende“** tiež určujú hodnoty zadaného stĺpca, do akej kategórie legendy patrí, **kategórie sa však zobrazia ako rozdielne farby**.

Stĺpce pridávame do panela **metódou chyť a potiahni** z panela **Polia**.

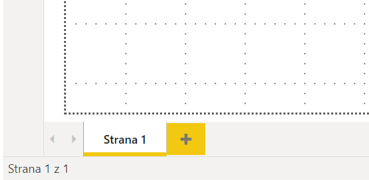
Ak klikneme **na šípku pri názve stĺpca**, môžeme si nastaviť či sa má zobraziť **suma, priemer, medián alebo počet** **položiek daného stĺpca**.



Obdobným spôsobom fungujú nastavenia pri všetkých typoch grafov.

Ak chceme do vizualizácie **pridať samostatné číslo** (napríklad zobrazenie sumy, celového počtu, atď.), môžeme tak urobiť cez prvok **Karta**  na paneli Vizualizácie. Do hlavného panelu potiahneme stĺpec, ktorý chceme zobraziť.

Ak si chceme **pridať viacero stránok** do vizualizácie, môžeme tak urobiť dole, pomocou **tlačidla plus**. Stránky **môžeme premenovať alebo vymazať ak klikneme na ich názov pravým tlačidlom myši**. Takisto ich môžeme aj **skryť** (nezobrazia sa používateľovi).



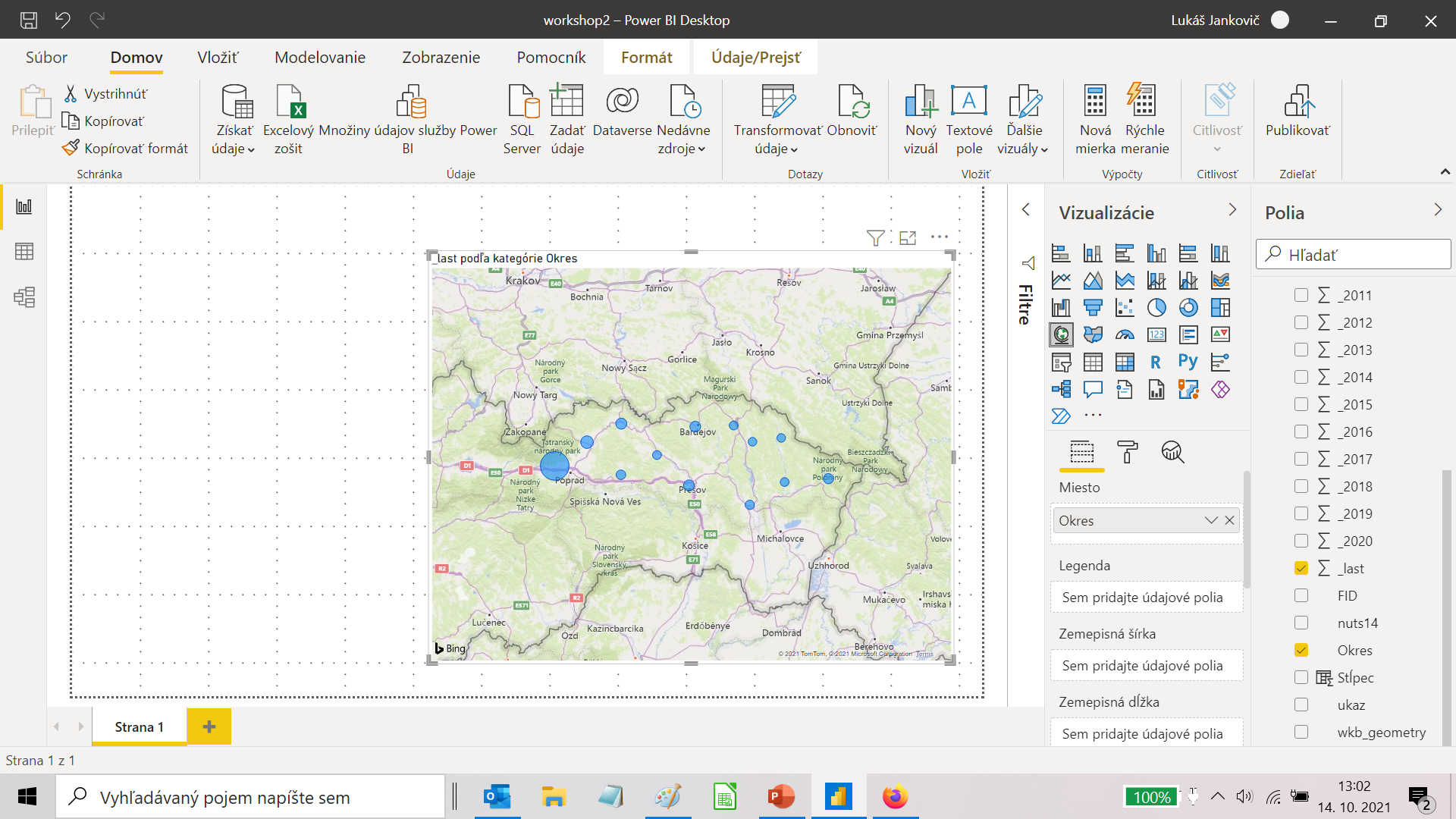
# Vytvorenie mapy

Typ grafu môžeme jednoducho zmeniť. Klikneme **na graf na hlavnom paneli a vyberieme iný typ grafu z panelu Vizualizácie, ktorý chceme nastaviť**. **Údaje sa nám zachovajú** aj v novom grafe. Jednoducho tiež môžeme vytvoriť mapu, kliknutím na tlačidlo **Mapa** .

Pri mape máme k dispozícii tieto polia:

* **Miesto** – mesto, obec, kraj, adresa
* **Legenda** – farba údajov
* **Zemepisná dĺžka** – GPS súradnica
* **Zemepisná šírka** – GPS súradnica
* **Veľkosť** – veľkosť bodu na mape (číslo)

Môžeme použiť určenie polohy bodov na mape **buď pomocou „miesta“ alebo pomocou zemepisnej dĺžky a šírky**.

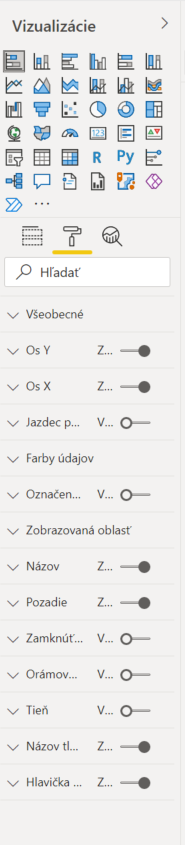
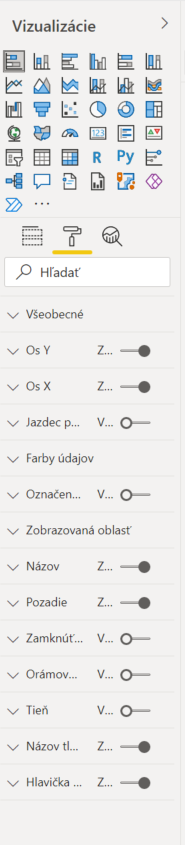


Ak by sme chceli použiť namiesto bodov na mape plochu, vyberieme v paneli Vizualizácie **kartogram.**

# Formátovanie

Nastavenie formátovania umožňuje meniť vzhľad prvkov – grafov, máp, atď. (farby, veľkosť a typ písma, pozadie atď.)

Možnosť formátovania máme pri každom prvku (mapa, graf, tabuľka, iné) a aj pri vzhľade hlavného panelu.

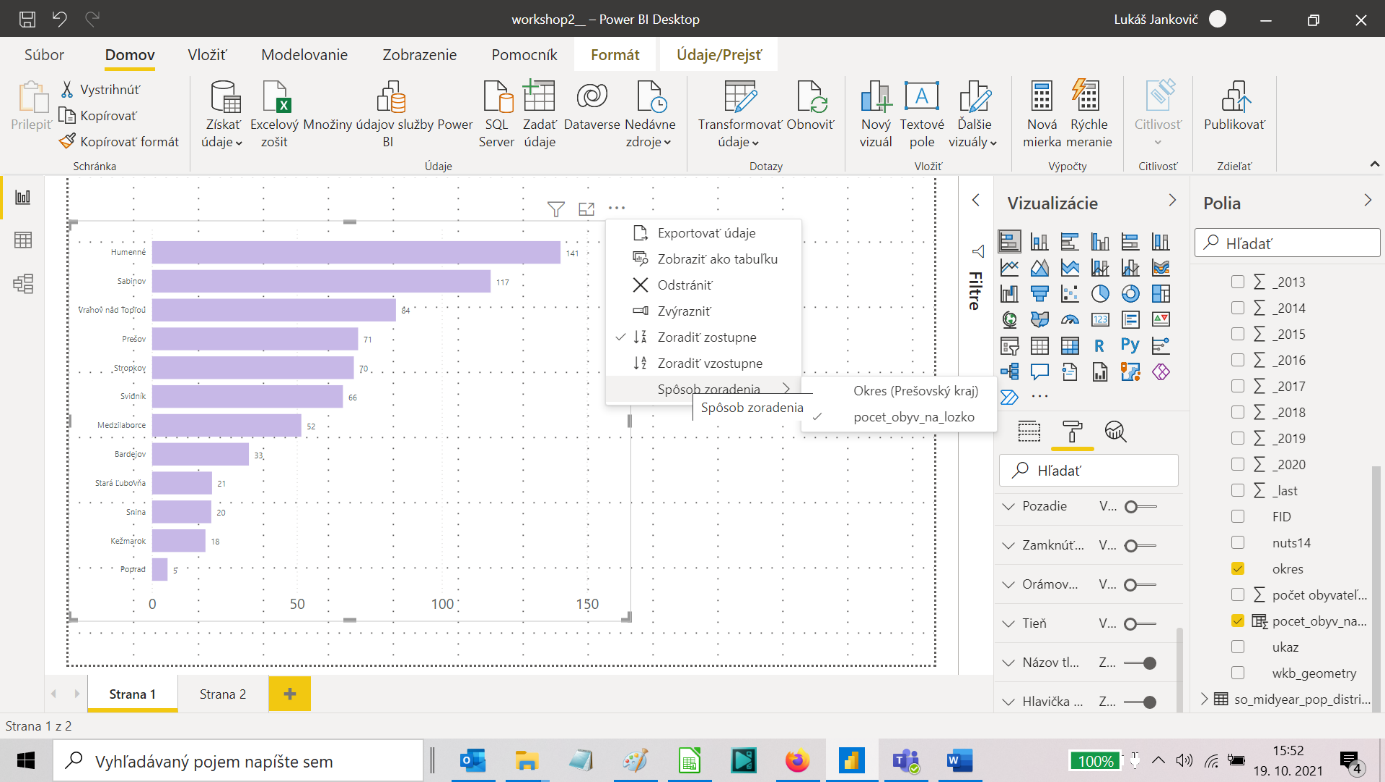


Ak chceme formátovať prvok, **najskôr si vyberieme tento prvok**, **tým že na neho klikneme**. To neplatí pri hlavnom paneli, ktorý formátujeme tak, že klikneme hocikde, kde sa nenachádza prvok. **Možnosti formátovania** sú dostupné po stlačení **tlačidla Formát** , ktoré sa **nachádza pod panelom Vizualizácie**. Niektoré možnosti majú zobrazený prepínač, ktorý zapína alebo vypína danú možnosť (napr. zobrazuje alebo nezobrazuje názov).

**Základné možnosti formátovania**

* **Názov** – názov prvku (hore)
* **Farby** – vieme tu nastaviť farby stĺpcov, farby bodov na mape, atď.
* **Osi** – možnosť zapínať alebo vypínať osy, ich názvy, formátovanie ich hodnôt
* **Označenie údajov** – hodnoty údajov
* **Pozadie** – možnosť nastaviť farbu/obrázok a priesvitnosť

**Zoradenie údajov** prvku je možné cez voľbu **vpravo hore – tri bodky**. Je nutné nastaviť stĺpec, podľa ktorého sa majú údaje zoradiť a spôsob zoradenia (vzostupne alebo zostupne).



**Dôležitá môže byť tiež voľba Flexibilné v možnosti Všeobecné. Táto voľba zapína alebo vypína automatické prispôsobenie prvku, takže môže pomôcť v prípade, ak prvok obsahuje viac položiek ako sa do neho zmestí.**

# Filtre

**Filtrovanie umožňuje zobraziť iba vybranú časť údajov**, ktoré sa majú zobraziť vo vizualizácii. To môže byť užitočné v prípade ak potrebujeme z vizualizácie **niektoré dáta vylúčiť** alebo aj v prípade, ak chceme používateľovi našej vizualizácie ponechať **možnosť vybrať si z položiek alebo si vyhľadať položky**.

V PowerBI poznáme 2 základné spôsoby filtrovania:

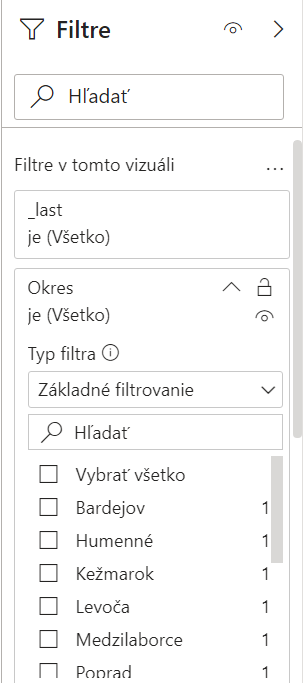
**Pevný filter**

Pevný filter sa nastavuje cez panel **Filtre** (ak nie je zobrazený, zobrazí sa po kliknutí na šípku). Tu môžeme definovať 3 druhy filtrov:

* **Filter vizuálu** – filter sa aplikuje len na vybraný prvok – graf, mapu alebo tabuľku
* **Filter stránky** – filter sa aplikuje na všetky prvky na tejto stránke
* **Filter na všetkých stránkach** – filter sa aplikuje na všetky stránky vizualizácie

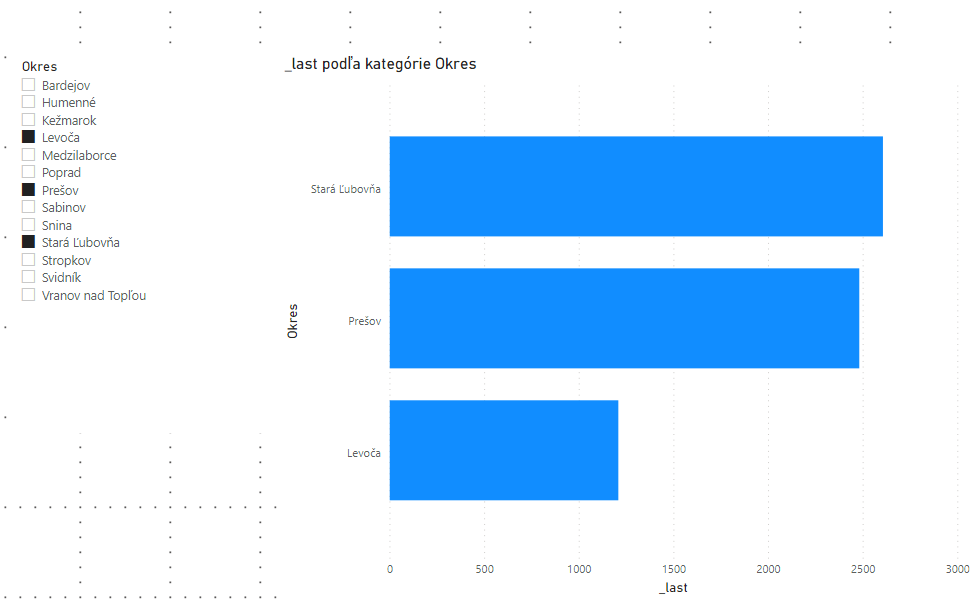
Filter nastavíme tak, že **potiahneme z panelu Polia stĺpec, ktorý chceme filtrovať**. Potom nastavíme či má byť filter **základný** alebo **pokročilý**. Základný filter zobrazí dostupné hodnoty, vybrať z nich môžeme kliknutím, prípadne si môžeme hodnotu vyhľadať. Pokročilý filter určuje rozmedzie hodnôt.

**Filtre, ktoré takto nastavíme nemôže používateľ zmeniť.**

****

**Používateľský filter**

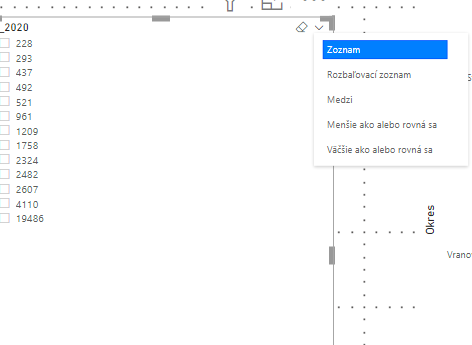
Ak chceme aby používateľ **mohol filter zmeniť**, použijeme z panelu Vizualizácie položku **Rýchly filter** . Pri tomto prvku musíme nastaviť jedinú hodnotu a to „Pole“. Pole obdobne ako pri pevnom filtri určuje, **na základe akého stĺpca sa bude filtrovať**.



Ako „pole“ môžeme použiť stĺpce obsahujúce text ale aj čísla.

Používateľovi sa zobrazí filter, v ktorom môže filtrovať potrebné hodnoty. **Filter môžeme rovnako ako ostatné prvky formátovať** a taktiež môžeme nastaviť **pomocou šípky v pravom** **hornom rohu či má zobraziť zoznam, rozbaľovací zoznam alebo iný typ filtrovania.** Tieto možnosti sú **závislé od typu údajov v stĺpci**, či ide o text alebo čísla.

Nastavenie filtra **ovplyvní všetky ostatné prvky** na celej strane.



# Power Query editor

Power Query editor slúži na **úpravu datasetov** (napr. pridanie stĺpcov, odoberanie) **a ich spájanie**.

**Používateľské rozhranie je tvorené tabuľkou, panelom nástrojov a postupnosťou krokov (vpravo).** **Postupnosť krokov je postup**, ktorý obsahuje jednotlivé kroky, **ktoré sa vykonajú s datasetom po jeho načítaní.** Bez ohľadu na to či sa dataset zmení, **vždy sa vykonajú tie isté kroky**. Kroky môžeme pridávať, mazať (kliknutím na symbol „x“) alebo meniť ich poradie (potiahnutím).

Obrázok, na ktorom je stôl

Automaticky generovaný popis

**Power Query editor otvoríme v zobrazení údajov, keď klikneme pravým tlačidlom na záhlavie ľubovoľného stĺpca a zvolíme Upraviť dotaz.**

# Spájanie datasetov

**Spájanie datasetov je veľmi užitočné, pretože pridávame do dát pridanú hodnotu.** Môžeme vizualizovať aj jeden dataset, pridaním iných údajov však **vytvárame unikátny produkt.** Veľmi často sa spájanie datasetov používa ak potrebujeme **spojiť podobné dáta z rôznych zdrojov** (napríklad dotácie z rôznych miest).

V PowerBI poznáme **2 základné typy** spájania údajov – **zlúčenie a pripojenie**. Rozdiel medzi nimi je ilustrovaný obrázkom nižšie.

Obrázok, na ktorom je stôl

Automaticky generovaný popis

Zlúčenie

Pripojenie

Pri **zlúčení** ide o dva datasety, ktoré majú **spoločný unikátny stĺpec**, na základe ktorého sa údaje spoja. Príkladom je ak máme dataset s prevádzkami a ich adresami a dataset s adresou a rokom výstavby budovy. Na základe spoločného unikátneho stĺpca, adresy, vieme do datasetu s prevádzkami doplniť rok výstavby budovy.

**Pri pripojení ide o dva datasety**, ktoré obsahujú v **zásade podobné údaje, ktoré potrebujeme spojiť do jedného datasetu s ktorým budeme ďalej pracovať.** Príkladom sú napríklad dotácie rôznych miest, kde pri každom meste budeme mať napríklad sumu, žiadateľa, názov projektu.

Ak potrebujeme spojiť viac ako dva datasety, musíme postupovať krok za krokom. **Pri zlúčení zlúčime dva datasety, s výsledkom zlúčime tretí a tak ďalej**. **Pri pripojení môžeme v PowerBI k datasetu pripojiť aj viac ako jeden dataset.**

**Pri spájaní datasetov musíme dbať na tom aby spájané stĺpce mali rovnaký typ (napr. spájame text s textom, čísla s číslami, atď.)**

**Pri zlučovaní datasetov používame na zlúčenie pole, ktoré má nastavený typ na text (aj ak obsahuje čísla).**

V Power Query editore sú obe možnosti spájania dostupné ak **klikneme pravým tlačidlom myši na ikonu v ľavom hornom rohu tabuľky.**

Obrázok, na ktorom je stôl

Automaticky generovaný popis

Pri oboch datasetoch vyberieme polia, ktoré sú spoločné.

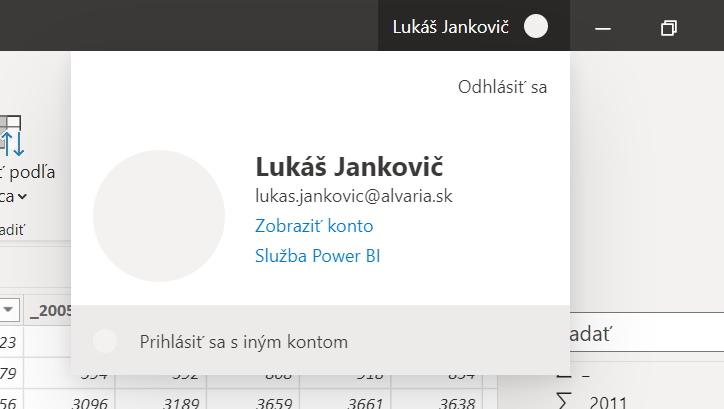
Obrázok, na ktorom je text

Automaticky generovaný popis

Vyberieme tabuľku/y (datasety), ktorá/é sa majú pripojiť k vybranému datasetu.

# Publikovanie

Pre **publikovanie** vizualizácie je nutné **byť prihlásený**. Prihlásiť sa je možné **v pravom hornom rohu aplikácie.**



Vizualizáciu publikujeme stlačením tlačidla Domov > **Publikovať.** Vizualizácia sa **nahrá na server PowerBI**.

Aby sme vizualizáciu mohli zverejniť, prihlásime sa na server. To môžeme urobiť už pri po nahraní vizualizácie na server (kliknutím na **Otvoriť položku v Power BI**).

Obrázok, na ktorom je text

Automaticky generovaný popis

Link na prihlásenie: <https://app.powerbi.com/>

Na obrazovke vyhľadáme názov vizualizácie, ktorý sme zadali a klikneme. Otvorí sa nám vizualizácia. Z menu hore vyberieme možnosť **Súbor >** **Vložiť zostavu**.

Obrázok, na ktorom je text, počítač, vnútri, prenosný počítač

Automaticky generovaný popis

Vizualizáciu môžeme vložiť:

* **Sharepoint,**
* **Webová lokalita alebo portál** – viditeľné len pre prihlásených,
* **Publikovať na webe (verejné)** – zverejnenie pre všetkých.

Následne sa nám zobrazí výzva na potvrdenie a po potvrdení uvidíme **odkaz na vizualizáciu** a **embed kód** (začína znakom „<“) na vloženie do webovej stránky.