

Fakulta riadenia a informatiky, Žilinská univerzita v Žiline

Algoritmy a údajové štruktúry 1
Semestrálna práca

2021/2022

Ján Holubčík

5ZYI36

Návrh aplikácie z pohľadu použitých údajových štruktúr

V aplikácii používam na načítanie dát sorted sequence table. Je to kvôli tomu lebo dáta v súboroch nie sú až tak pekne zoradené a pri načítaní okresov a krajov môžem použiť metódu find, ktorou nájdem ku akým okresom alebo krajom patrí.

Po načítaní používam len unsorted sequence table, do ktorej priradím prvky z ssq, ktorá mi umožňuje triediť prvky pomocou jej metódy swap.

Sorted sequence table využívam kvôli jej vyhľadávaniu čo je $\log(n)$ čo sa pri načítavaní hodí.

Na grafické spracovanie som použil Windows forms.

Code	Official title	bez ukončeného vzdelania	základné vzdelanie	stredné odborné vzdelanie	úplné stredné vzdelanie	vyššie odborné vzdelanie	vysokoškolské vzdelanie	bez školského vzdelania	nezistené	Predproduktívny	Produktívny	Postproduktívny
------	----------------	--------------------------	--------------------	---------------------------	-------------------------	--------------------------	-------------------------	-------------------------	-----------	-----------------	-------------	-----------------

UML diagramy

Loader

- služi na načítanie dát, obsahuje 4 sorted sequence table pre každý typ územnej jednotky a jej getre
- obsahuje metódu loader ktorá načítava zo súborov

Sorter

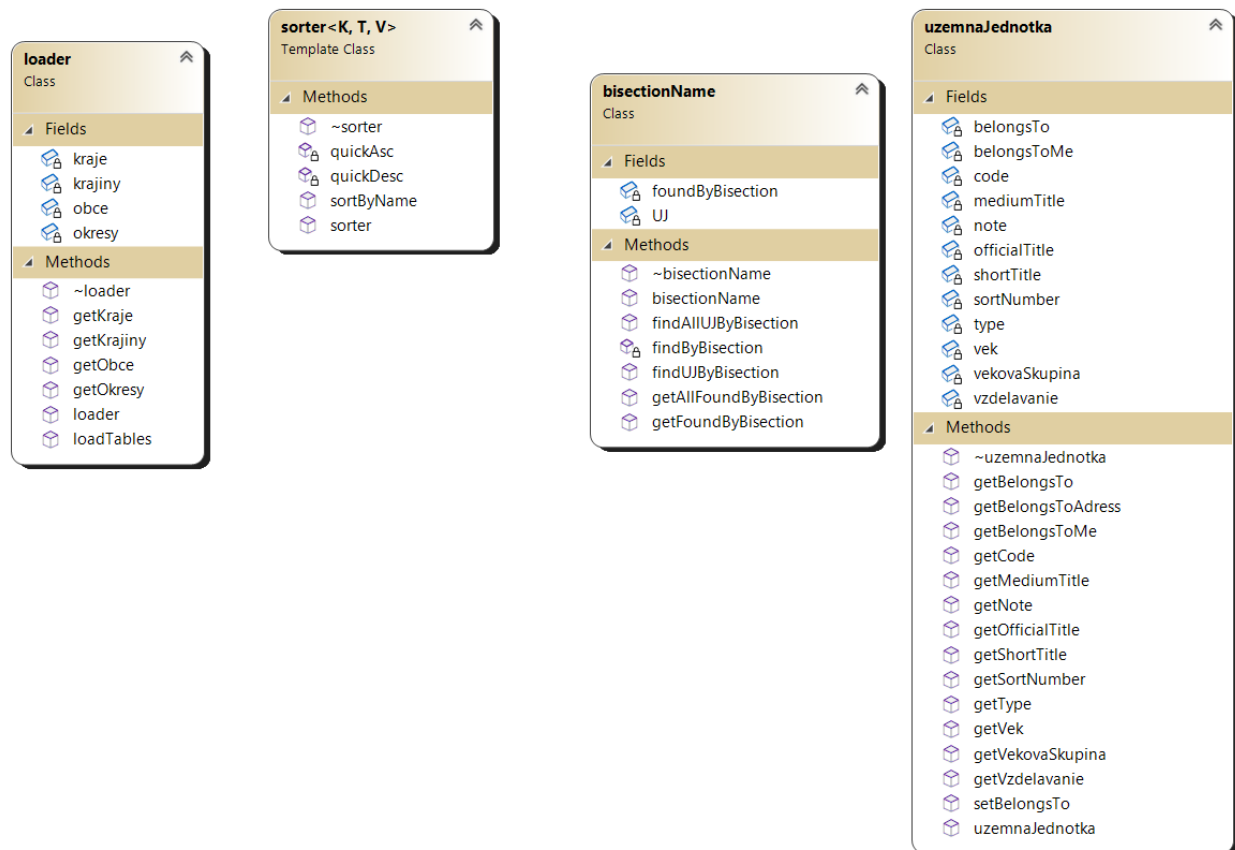
- služi na triedenie, triedi podľa metódy ktorú si zvolí, pre triedenie sa volá metóda sortByName

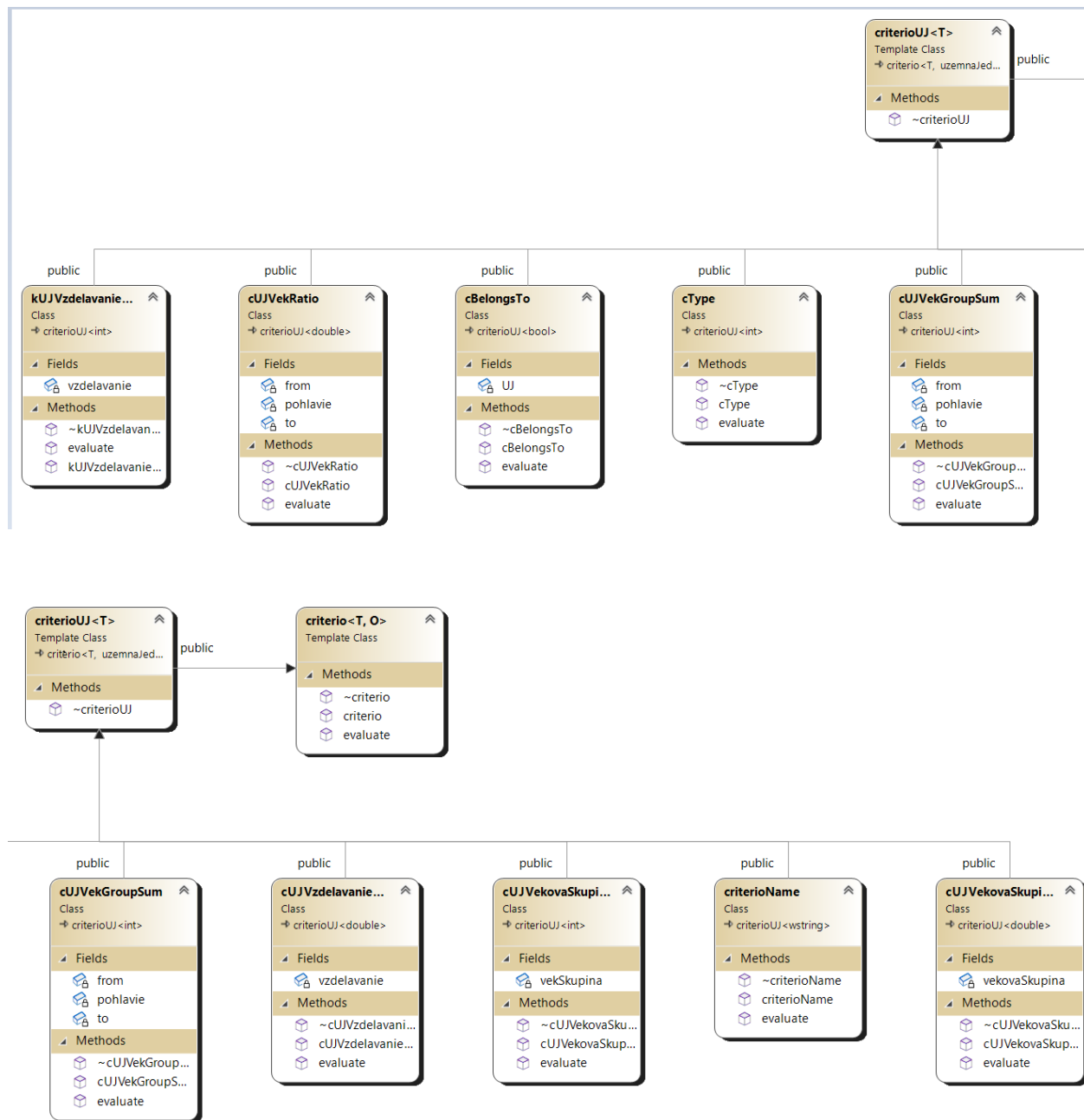
Bisection

- služi na nájdenie UJ s názvom, vracia buď všetky UJ ktoré našiel alebo len jednu

UzemnaJednotka

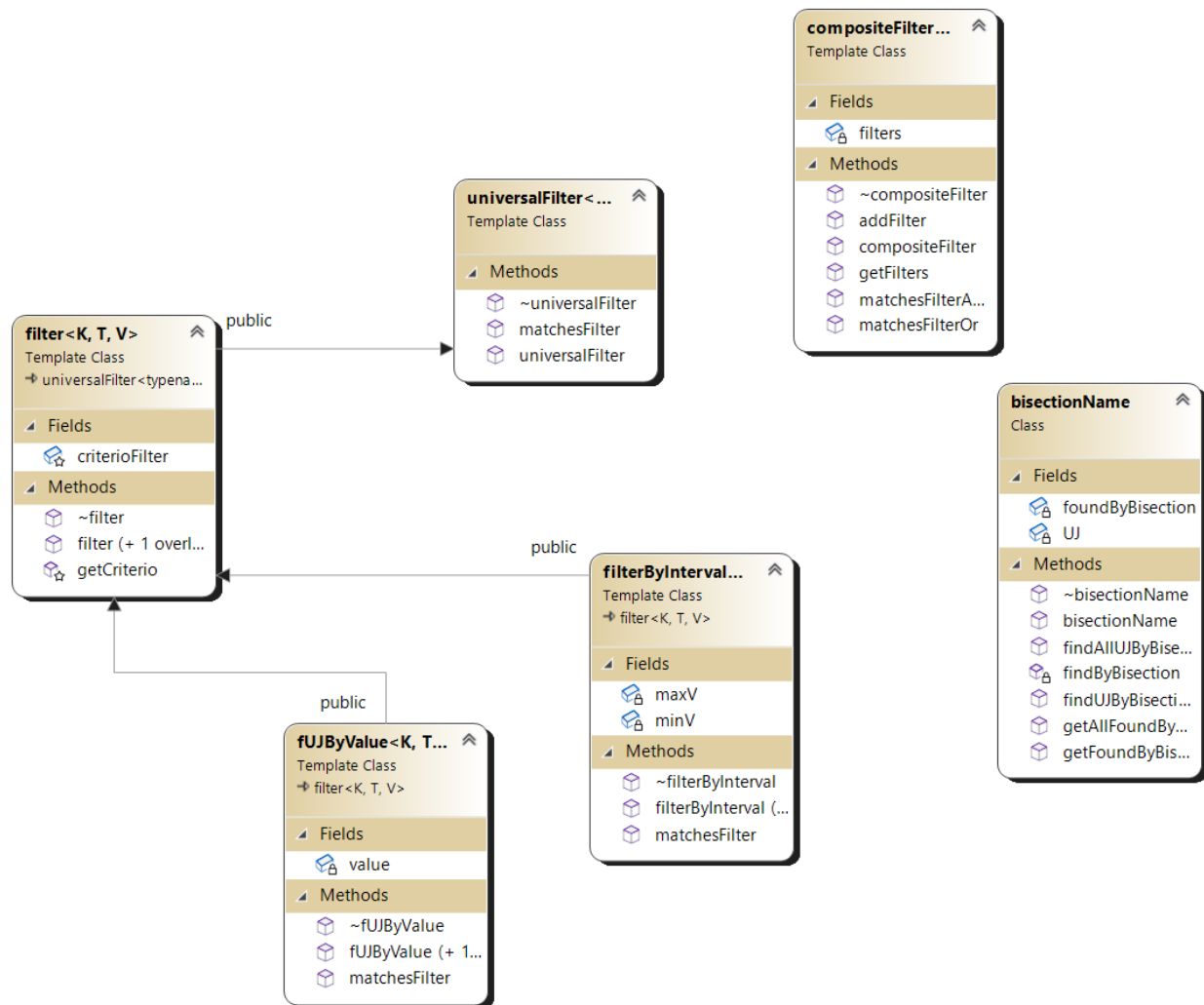
- obsahuje všetky atribúty ktoré sa nachádzajú v súboroch, ich vzdelanie, vek ale aj vekové skupiny
- obsahuje aj UJ ku ktorej a patrí a UJ ktoré patria pod ňu
- vracia všetky atribúty ale aj vzdelanie, vek a vekové skupiny ktoré sú uložené v poliach





Kritéria

- criterioUJ dedí z criterio
- ostatné kritéria dedia z criterioUJ
- criterioUJ používam pri triedení



Filtre

- fUJByValue, filterByInterval universalFilter dedia z filtra
- filter obsahuje kritérium podľa ktorého filtruje tabuľku
- compositeFilter obsahuje pole v ktorom sa ukladajú filtre ktorými chcem filtrovať
- v universalFilter používam virtuálny konštruktor aby pri vymazaní filtrov mi vymazalo aj kritérium z ktorým pracujem

Hierarchia načítania

Najskôr načítam zo súboru `obce.csv` všetky obce. Keď sú obce načítané načítavam zo súboru `vzdelanie` kde načítam `code`, podľa ktorého nájdem obec ku ktorému patrí.

Pre načítanie obce je zložitosť $O(N)$.

Pre načítanie vzdelania a veku je zložitosť $O(N \cdot \log(N))$. $\log(N)$ preto lebo vyhľadávam obec ku ktorému patria.

Pre načítavanie okresov načítam kód zo súboru a prechádzam načítané obce. Ku okresu potom spočítavam vzdelanie a vek z obcí, ktoré mu patria.

Pre načítanie okresu je zložitosť $O(N*N)$.

Pre načítavanie krajov znova načítam kód zo súboru a prechádzam načítané okresy. Ku kraju potom spočítavam vzdelanie a vek z okresov, ktoré mu patria.

Pre načítanie kraja je zložitosť $O(N*N)$.

Bodové vyhľadávanie

Pre bodové vyhľadávanie si vyberiem jeden z typov UJ ten potom uloží do unsorted sequence table. Tú utriedim a pomocou bisekcie podľa názvu nájdem buď jednu alebo viacero UJ. Taktiež v aplikácii používam čiastočne vyhľadávanie ktoré sa vykoná ak nenašlo danú obec. Zoberie sa string vo vyhľadávacom poli a prejde všetky UJ a hľadá či sa tam nenáchadza.

Pre hľadanie pomocou bisekcie je zložitosť $O(\log n)$.

Pre čiastočne vyhľadávanie je zložitosť $O(n*k)$ kde k sú dĺžky stringu v ktorom vyhľadávam.

Filtrovanie

Pre filtrovanie používam filtre vo vyššie uvedenom uml diagrame.

Keďže sa majú vyfiltrovať všetky UJ nevyhnem sa zložitosti $O(n)$.

V aplikácii dokážem použiť viac ako jeden filter.

Triedenie

V aplikácii využívam quicksort na triedenie tabuľky.

Jeho zložitosť závisí podľa výberu jeho pivota pričom pri ideálnom pivotovi je $O(n\log(n))$.

V aplikácii môžem triediť podľa hodnôt ktoré sú v tabuľke ale aj pomocou vyfiltrovanej tabuľky podľa použitého kritéria.

Výber najlepších

Pre výber najlepšieho alebo najhoršieho sa tiež nevyhnem zložitosti $O(n)$, keďže potrebujem prejsť všetky UJ ktoré sú v tabuľke pre nájdenie najhoršieho/najlepšieho.

Použité kritéria

KNázov – vraciam oficiálne meno UJ ako wstring, $O(1)$

KTyp – vraciam typ jednotky (0 pre obec, 1 pre okres, 2 pre kraj, 3 pre krajiny), $O(1)$

KPríslušnosť – vraciam ku ktorej UJ patrí, v tomto kritériu aj hľadám či nepatrí aj ešte vyššej UJ. Ak vyššiu UJ nenašlo alebo nepatrí do UJ vracia false ak našlo vracia true.

(pre obec $O(1) + O(1) + O(1)$, pre okres $O(1) + O(1)$, pre kraj $O(1)$)

KVzdelaniePočet – vracia int ktoré je uvedené v poli na indexe vzdelanie, posielané int do kritéria by malo byť od 0 až 7, pretože sa nachádza v poli o 8 miestach, $O(1)$

KVzdelaniePodiel – vracia podiel vzdelania na indexe v poli vzdelavanie, tento podiel sa vyráta pomocou intu ktorý pošlem, kde zoberiem hodnotu na indexe vzdelania a vydám ho všetkými hodnotami v poli ($O(1) + O(n)$ kde n je počet prvkov v poli)

KVekPočet – vracia počet ľudí v intervale kde hodnoty v tomto intervale udávajú, od akého veku po aký vek má tieto hodnoty spočítať a vrátiť sumu z týchto hodnôt, taktiež posielam aj pohlavie (0-muž, 1-žena, -1 pre obidva), $O(n)$ kde n sú všetky hodnoty v intervale

KVekPodiel – vracia podiel ľudí v intervale kde hodnoty v tomto intervale udávajú, od akého veku po aký vek má tieto hodnoty spočítať a vrátiť podiel zo sumy hodnôt a všetkých ľudí v UJ, taktiež posielam aj pohlavie (0-muž, 1-žena, -1 pre obidva), $O(n) + O(k)$ kde n sú všetky hodnoty v intervale a k sú všetky hodnoty v poli

KVekovaSkupinaPočet- vracia počet z polia vekovej skupiny ktorú si vyberiem cez int ktorý posielam, $O(1)$

K VekovaSkupinaPodiel- vracia podiel z polia vekovej skupiny ktorú si vyberiem, tento podiel sa vyráta z podielu vekovejSkupiny a všetkých ľudí vo vekových skupinách v danej UJ, $O(1) + O(n)$

4.úroveň

- Hodnoty kritéria *K Názov* a *K UJTyp* pre územnú jednotku a pre všetky vyššie jednotky, do ktorých patrí
- Hodnoty kritéria *K UJVzdelaniePočet* pre $\forall vzdelanie \in VZDELANIE$
- Hodnoty kritéria *K UJVekováSkupinaPočet* pre $\forall skupina \in EVS$

-vypisujem názov, všetky typy, všetky hodnoty polia vo vzdelaní a všetky hodnoty polia vo vekovej skupine UJ do riadkov tabuľky, zložitosť $O(k*n)$ kde k sú všetky atribúty a n je počet územných jednotiek

- Ak bol zadán filter $F_{UJVzdelaniePočet}$ tak aj hodnoty kritéria $K_{UJVzdelaniePočet}$
- Ak bol zadán filter $K_{UJVzdelaniePodiel}$ tak aj hodnoty kritéria $F_{UJVzdelaniePodiel}$
- Ak bol zadán filter $K_{UJVekPočet}$ tak aj hodnoty kritéria $F_{UJVekPočet}$
- Ak bol zadán filter $K_{UJVekPodiel}$ tak aj hodnoty kritéria $F_{UJVekPodiel}$
- Ak bol zadán filter $K_{UJVekováSkupinaPočet}$ tak aj hodnoty kritéria $F_{UJVekováSkupinaPočet}$
- Ak bol zadán filter $K_{UJVekováSkupinaPodiel}$ tak aj hodnoty kritéria $F_{UJVekováSkupinaPodiel}$

-ak sa vo filtri nachádza kritérium ktoré sa používa tak vytvorím nový stĺpec v tabuľke kde jeho názov je názov kritéria, zložitosť pre filtrovanie $O(k*n)$, kde k je počet filtrov a n je počet UJ
 -pre výpis zoberiem vyfiltrované hodnoty ktoré vypíšem aj z hodnotou kritéria, $O((j+k)*n)$ kde j je počet kritérií, k je počet atribútov v UJ a n je počet UJ

- Hodnoty kritéria $K_{UJVekPočet}$ pre $\forall pohlavie \in POHLAVIE$ a parameter vek z kritéria $K_{UJVekPodiel}$
- Zoradíte vzostupne všetky okresy Žilinského kraja podľa podielu mužov vo veku od 20 do 40 rokov, kde je ich podiel medzi 20 – 30%
- Hodnoty kritéria $K_{UJVekováSkupinaPodiel}$ pre $\forall vekováSkupina \in EVS$
- zoradíte zostupne všetky obce podľa počtu ľudí v produktívnom veku, kde je ich počet medzi 1000 – 10000.

-pre zoradenie si vytvorím filtre ktoré potrebujem potom podľa kritéria ktoré som použil utriem tabuľku, zložitosť $O(n*k) + O(n \log(n))$, kde k je počet filtrov, n je počet UJ
 -pre výpis zoberiem utriedené hodnoty ktoré vypíšem aj z hodnotou kritéria, $O((j+k)*n)$ kde j je počet kritérií, k je počet atribútov v UJ a n je počet UJ

- Vypíšte informácie o okrese v Trnavskom kraji, ktorý má najviac obyvateľov v produktívnom veku.
- Vypíšte informácie o obci v Trenčianskom kraji, ktorý má najmenej obyvateľov v predproduktívnom veku.

-pre hľadanie najlepšieho/najhoršieho prechádzam celú tabuľku kde porovnávam hľadanú hodnotu so zatiaľ najlepšou/najhoršou, zložitosť $O(n)$

-pre výpis $O(n)$, kde n je počet atribútov v UJ

-ak hľadám maximum pri použití filtra tak $O(n+k)$, kde n je počet atribútov v UJ a k je počet kritérií

Používateľská príručka

Obr.1 : Grafické rozpracovanie aplikácie

The screenshot shows a web application window titled "MyForm". It features a search and filter interface. On the left, there's a section for "Zvol' územnú jednotku" (Select territorial unit) with a dropdown menu showing "Obec" and a text input field. Below it is a button "Vypíš územné jednotky". To the right, there are sorting options: "Triedenie bez filtra" and "Triedenie z filtrom", each with a dropdown. A "Ukáž filtre" button is also present. Further right, there are checkboxes for "vzostupne?", "Filtrovať podľa intervalu?", and "Obidva pohlavia". Below these are input fields for "Hľadanie maxima bez filtra", "Hľadanie maxima s filtrom", "od", and "do", along with a checkbox "Hľadať minimum?". A "Pridaj filter" button is at the bottom of this section. On the far right, there's a "Filter" section with a "Podľa" dropdown and a list of filter categories: "Code", "Official title", "bez ukončeného vzdelania", "základné vzdelanie", "stredné odborné vzdelanie", "úplné stredné vzdelanie", "vyššie odborné vzdelanie", "vysokoškolské vzdelanie", "bez školského vzdelania", "nezistené", "Predproduktívny", "Produktívny", and "Postproduktívny". Below the filter section are buttons for "Filtruj" and "Resetuj filtre".

Code	Official title	bez ukončeného vzdelania	základné vzdelanie	stredné odborné vzdelanie	úplné stredné vzdelanie	vyššie odborné vzdelanie	vysokoškolské vzdelanie	bez školského vzdelania	nezistené	Predproduktívny	Produktívny	Postproduktívny
------	----------------	--------------------------	--------------------	---------------------------	-------------------------	--------------------------	-------------------------	-------------------------	-----------	-----------------	-------------	-----------------

Obr.2 : Slúži na bodové vyhľadávanie. Ak sa do poľa nič nezadá vypíše všetko.

This is a close-up of the search input field from the previous screenshot. It shows the "Zvol' územnú jednotku" dropdown menu with "Obec" selected, followed by a text input field. Below the input field is a button labeled "Vypíš územné jednotky".

Obr.3 : Slúži na triedenie a hľadanie maxima. Ešte pred triedením sa musí tabuľka zaplniť inak padne program. Triedenie a hľadanie maxima bez filtra slúži na triedenie podľa názvu, vzdelania alebo vekovej skupiny alebo na hľadanie maxima v týchto prvkoch . Ak používateľ použil filtre môže triediť aj vyfiltrovanú tabuľku alebo hľadať maximum.

Triedenie bez filtra Triedenie z filtrom

▼ ▼

☐ vzostupne?

Hľadanie maxima bez filtra Hľadanie maxima s filtrom

▼ ▼

☐ Hľadať minimum?

Obr.4 : Ak používateľ správne zadá do filtrov tak tlačidlo pridaj filter pridá filter. Ak používateľ spravil niečo zle aplikácia ho upozorní.

Ukáž filtre Podľa podielu veku ▼

☐ Filtrovať podľa intervalu ?

od do

_____ _____

☐ Ženské pohlavie

☐ Obidva pohlavia

Pridaj filter

Obr.5 : Ukazuje pridané filtre. Tlačidlo filtruj tabuľku ju vyfiltruje. Tlačidlo resetuj filtre zruší filtre.

Filter	Podľa

< >

Filtruj Resetuj filtre