# Идеја за практично истраживање

|  |
| --- |
| **Истраживачи: име, презиме, број индекса (највише четири особе, ред по особи)** |
| Miloš Sirar IN 3/2020 |

|  |
| --- |
| **Шифра истраживања: кодни назив (једна или две речи)** |
| **Vreme\_Šangaj** |

|  |
| --- |
| **Наслов истраживања: сажето исказана суштина истраживања (до 15 речи)** |
| Eksplorativna analiza podataka o vremenskim prilikama (parametrima) i kvalitetu vazduha u Šangaju u periodu 2014-2021 |

|  |
| --- |
| **Предмет, циљ, методи и очекивани резултати истраживања (бар 100 речи)** |
| **Istraživanje** obuhvata učitavanje podataka o vremenskim prilikama i kvalitetu vazduha u Šangaju za naznačeni period, analizu podataka koji su učitani, predstavljanje podataka pomoću tabela i grafikona, traženje povezanosti između različitih faktora i donošenje zaključaka na osnovu analize.  **Cilj** je da se kroz ovo istraživanje stekne šira slika o oblasti analize podataka i zaključivanja na osnovu dobijenih rezultata. Pored toga bitna je i tematika: razumevanje podataka o vremenu i kvalitetu vazduha i njihov međusobni uticaj.  Što se tiče **metoda**, planirano je da se koristi programski jezik Python (i potrebne biblioteke). Koristiće se metode za obradu podataka i tekstualnih fajlova, kao i metode za kreiranje grafikona i tabela.  **Rezultati istaživanja** bi trebalo da pokažu kako su se vreme i vremenski uslovi menjali tokom perioda i njihovu potencijalnu povezanost sa kvalitetom vazduha. Sve ovo će biti predstavljeno grafički da bi se lakše uočile zavisnosti i dali zaključci. |

|  |
| --- |
| **Опис скупова података (најмање један скуп података, бар 30 речи по скупу података)** |
| **Skup podataka 1** – 2502 podatka o vremenskim prilikama tokom perioda 2014-2021, sadrži 22 kolone, podaci su u csv formatu. Izvor: Kaggle  **Skup podataka 2** – 2503 podatka o kvalitetu vazduha tokom perioda 2014-2021, sadrži 7 kolona, podaci su u csv formatu. Izvor: Kaggle  **Skup podataka 3** – 2587 podataka, još dodatnih informacija koje će po potrebi biti uvrštene u istraživanje, sadrži 25 kolona, podaci su u csv formatu. Izvor: Kaggle  Link do sva 3 skupa podataka: https://www.kaggle.com/datasets/erhankul/shanghai-air-pollution-and-wheather-20142021?select=Shanghai+AQI+and+Wheather+2014-2021.csv |

|  |
| --- |
| **Опис и план активности (бар 100 речи)** |
| **Prikupljanje i analiza skupova podataka**  Prikuputi i pregledati sva tri skupa podataka, otkloniti nevalidne ili nedostajuće vrednosti, uočiti ponavljanje podataka i povezati podatke iz različitih csv fajlova.  **Početak i podela istraživanja**  Pregled svih učitanih i filtriranih podataka. Raspodela podataka po redosledu obrade, odnosno krenuće se prvo od nekih opštih analiza i ispitivanja, a zatim će se preći na konkretniju analizu. Prvo će se gledati promena vremena i vremenskih uslova tokom vremena, a zatim će se preći na analizu kvaliteta vazduha na osnovu podataka. Zatim će se tražiti zavisnosti između dobijenih rezultata.  **Spisak nekih od analiza koje će biti rađene**  Prvo će se raditi deskriptivna statistika, a zatim analize pojedinačnih obeležja: prosečna mesečna temperatura, prosečna godišnja temperatura, maksimalne i minimalne temperature po mesecima, predstavljanje padavina, predstavljanje vidljivosti, predstavljanje indeksa kvaliteta vazduha i parametara koji dovode do takvog indeksa. Zatim će se uraditi neke komplikovanije analize koji će prikazati povezanost između udara vetra i kvaliteta vazduha, padavina i kvaliteta vazduha, vidljivosti i kvaliteta vazduha. Ovo su neki od primarnih ciljeva istraživanja, a pored njih će po potrebi biti urađene dodatne analize koje će pomoći u shvatanju problematike. Sve ove analize će biti propraćene grafikonima: stubičastim, tačkastim, linijskim, histrogram, heatmap. Ovo su grafikoni koji su planirani da se koriste, uz mogućnost dodavanja drugih vrsta ukoliko budu prikladniji u odnosu na podatke.  **Prikaz dobijenih rezultata analiza**  Nakon analize podataka rezultati će biti predstavljeni u formi grafikona. Preko njih će se prikazati kakvo je bilo vreme i kvalitet vazduha tokom ovog perioda preko određenih parametara, kao i zavisnosti između parametara.  **Zaključak**  Nakon svih analiza i prikaza biće sumirano šta je tokom ovog istraživanja zaključeno i šta nije moglo da se zaključi. Onda će zaključci će biti prezentovani pomoću PowerPoint prezentacije gde će se sumirati istraživanje i predstaviti ceo postupak kroz bitne faktore. |

|  |
| --- |
| **Расподела активности међу истраживачима (ако више особа, на равномеран начин)** |
|  |