**Vprašanja največje katastrofe 21. stoletja**

Podatkovno rudarjenje

Avtorja: Timotej Šušteršič, 63210333

Luka Podvratnik, 63210254

19.4.2023, Ljubljana

Kazalo vsebine

[**1. Uvod** 2](#_Toc132640987)

[**2. Delavno okolje** 2](#_Toc132640988)

[2.1 Podatki 2](#_Toc132640989)

[2.2 Knjižnice 2](#_Toc132640990)

[**3. Klasifikacija** 2](#_Toc132640991)

[3.1 Pridobitev podatkov 2](#_Toc132640992)

[**4. Zaključek** 3](#_Toc132640993)

[**5. Viri** 3](#_Toc132640994)

[Slika 1: Histogram števila zmag v zadnji sezoni po ekipah 2](#_Toc121089622)

[Slika 2:Hisogram največkrat uvrščenih v osmino finala 3](#_Toc121089623)

[Slika 3: Histogram razmerij vodstva v zadnji četrtini in zmage po ekipah 3](file:///C:\Users\Timotej\Downloads\Luka_Podvratnik_63210254_Seminarska_StrojnoUcenje.docx#_Toc121089624)

[Slika 4: Škatla z brki - razmerje zmage in vodstva v zadnji četrtini 3](file:///C:\Users\Timotej\Downloads\Luka_Podvratnik_63210254_Seminarska_StrojnoUcenje.docx#_Toc121089625)

[Slika 5: Povprečje doseženih točk na sezono za ekipo BKN 4](#_Toc121089626)

[Slika 6: Histogram razmerij zmag v podaljšku 4](file:///C:\Users\Timotej\Downloads\Luka_Podvratnik_63210254_Seminarska_StrojnoUcenje.docx#_Toc121089627)

[Slika 7: Škatla z brki - Razmerje zmag v podaljšku 4](file:///C:\Users\Timotej\Downloads\Luka_Podvratnik_63210254_Seminarska_StrojnoUcenje.docx#_Toc121089628)

[Slika 8: Prikaz načinov pridobivanja podatkov 5](#_Toc121089629)

[Slika 9: Drevo s ciljnim atributom "W" 6](#_Toc121089630)

[Slika 10: Drevo s ciljnim atributom "razlika točk" 7](#_Toc121089631)

# **1. Uvod**

Seminarska naloga govori pandemiji Covid v Sloveniji, ki je trajala od leta 2020 do 2022. Izbrala sva jo, ker je v času korone veliko vprašanj ostalo neodgovorjenih, ljudje so delali špekulacije in teorije zarote Sedaj sva dobila priložnost, da na podlagi resničnih podatkov pripraviva neko zanimivo statistiko, ki bo razjasnila veliko vprašanj in to negotovo obdobje pojasnila z dejstvi in grafi prikazani na zanimiv način.

# **2. Delavno okolje**

## 2.1 Podatki

Podatke črpava iz Spletne strani Covid Sledilnik oz. njegovega Git repozitorija, kjer imajo shranjene v csv datotekah. Časovno obdobje je za večino obdobje pandemije torej od leta 2020 do danes(april 23). Večinoma sva črpala podatke iz datotek o stanju v bolnišnicah, okuženosti, smrtnosti in cepljenju. Vsebujejo pa tudi druge bolj splošne podatke kot so smrtnost v zadnjih 20 letih, policijska poročila, stanje v kanalizacijah, podatke o šolah, itd.

Podatke imajo načeloma shranjene vedno skozi čas, kjer imajo shranjene dnevne podatke, večinoma pa jih hranijo v sistemu »to date« oz. vsak dan je seštevek vseh prejšnjih plus današnji.

## 2.2 Knjižnice

Za delo sva v večini uporabljala knjižnico Pandas oz. njegov dataframe za obdelovanje podatkov in kot dodatek po potrebi Numpy za sezname. Za prikazovanje podatkov, kot so grafi pa sva uporabljala matplotlib.

# **3. Študije**

## 3.1 Smrtnost

V Sloveniji je bilo veliko govora o smrtnosti, češ da vsak, ki umre okužen z covidom, je umrl zaradi njega. Tukaj preizkušava to tezo. To sva se lotila tako, da izračunava pričakovano smrtnost in nato primerjava, če se presežek ujema z številom žrtev covida.

Slika, ki vsebuje besede diagram

Opis je samodejno ustvarjen

Slika 1: Graf števila umrlih ljudi v Sloveniji po letih (modra), pričakovana smrtnost (oranžna)

Graf pokaže z modro črto koliko ljudi je umrlo od leta 2010 do 2022. Z oranžno pa smo poskušali ugotoviti trend rasti, ki znaša 0.87% letno od 2010 do 2019. Na ta način sva dobila neko pričakovano smrtnost za covidna leta ampak kot vidimo je umrlo veliko več ljudi.

Slika, ki vsebuje besede diagram

Opis je samodejno ustvarjen

Slika 2: Smrtnost z pričakovano vrednostjo plus prijavljene žrtve covida

Sedaj pa sva pa sva tej napovedi dodala število ljudi, za katere je bil vzrok smrti covid. Kot vidimo je leta 2020: Umrlo nekaj 100 več ljudi kot bi pričakovali. Torej so v prvem letu dokaj konzervativno in korektno ocenili smrtnost o covidu. Naslednje leto je umrlo največ ljudi za covidom. Razlika v grafu je 847 ljudi, kar je že kar veliko ljudi, katerih vzrok smrti je bil napačno ocenjen. Zadnje leto korone, pa očitno niso več kaj dosti komplicirali. Razlika je kar 1454 ljudi, kar predstavlja 6% vseh umrlih in je relativno velika napaka.

Torej lahko potrdimo tezo da je bilo v letu 2021 in 2022 veliko vzrokov smrti napačno ocenjenih in v realnosti za covidom ni umrlo 9201 človek kot je uradni podatek. Pravi podatek se seveda ne da izračunati, ampak glede na najino oceno o pričakovani smrtnosti je to 6843 ljudi.

# **4. Zaključek**

Seminarska naloga pokriva skoraj vsa področja epidemije in je bilo tudi za naju izredno zanimivo primerjati, kaj so mediji govorili po poročilih, kakšna je bila ulična propaganda, kako je takratna vlada reagirala z ukrepi. Meniva da so mediji velikokrat nekorektno predstavljali informacije, verjetno samo z namenom, da imajo zgodbo, realnost pa ni bila tako siva.

# **5. Viri**

* Covid Sledilnik: <https://covid-19.sledilnik.org/sl/stats>
* Covid Sledilnik (Git): <https://github.com/sledilnik/data-api/>
* Pandas: <https://pandas.pydata.org/>
* Matplotlib: <https://matplotlib.org/>
* Numpy: <https://numpy.org/>