

UPA Projekt

Zvolená téma: COVID-19

Riešitelia:

- Marián Kapišinský (xkapis00)
- Daniel Kozovský (xkozov02)
- Timotej Šurina (xsurin03)

Zvolené dotazy a formulácia vlastného dotazu:

- v grafoch zobrazte tempo zmeny počtov aktuálne chorých (absolútny aj percentuálny prírastok pozitívnych prípadov a kľzavý priemer rôznych dĺžok v rôznych časoch)
- určite vplyv epidémie COVID-19 na počet mŕtvych v porovnaní podľa počtu chorých, počtu hlásených úmrtí na ochorenie COVID-19 a v porovnaní s minulými rokmi
- v histogramoch zobrazte vývoj epidémie podľa počtu aktuálne nakazených v jednotlivých krajoch

1 Stručná charakteristika zvolenej dátovej sady

Zvolené dátové sady:

- COVID-19: Celkový (kumulatívny) počet osôb s prokázanou nákazou dle krajských hygienických stanic včetně laboratoří, počet vyléčených, počet úmrtí a provedených testů (v2)¹
- Zemřelí podle týdnů a věkových skupin v České republice²
- COVID-19: Přehled epidemiologické situace dle hlášení krajských hygienických stanic podle okresu³

Prvý dataset je potrebný pre zodpovedanie prvého aj druhého dotazu. Obsahuje:

- kumulatívny počet nakazených
- kumulatívny počet vyliečených
- kumulatívny počet úmrtí
- kumulatívny počet testov

¹<https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19/nakazeni-vyleceni-umrti-testy.csv>

²<https://czso.cz/documents/62353418/138258831/130185-20data102020.csv>

³<https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19/kraj-okres-nakazeni-vyleceni-umrti.csv>

Druhý dataset je potrebný pre zodpovedanie druhého dotazu. Dátová sada, mimo iné, obsahuje:

- dátum
- celkový počet úmrtí za týždeň

Tretí dataset je potrebný pre zodpovedanie vlastného (tretieho) dotazu. Obsahuje:

- dátum
- NUTS kód kraja
- kumulatívny počet nakazených
- kumulatívny počet vyliečených
- kumulatívny počet úmrtí

Všetky uvedené datasety sú dostupné na webe cez priložené odkazy. Ide o súbory formátu **CSV**, z ktorých bude potrebné vyextrahovať potrebné dáta a vhodne ich upraviť pre uloženie do vybranej databázy. Súbory je možné ľahko získať napríklad použitím nástroja `curl`.

2 Zvolený spôsob uloženia surových dát

InfluxDB uchováva rôzne časové rady meraní rôznych hodnôt v chronologickej postupnosti. Upravené dáta, získané z uvedených datasetov budú musieť byť upravené na *line protocol*, ktorý *InfluxDB* využíva, a následne uložené do samotnej databázy.

cumulative,status=infected	count=X	<timestamp>
cumulative,status=recovered	count=X	<timestamp>
cumulative,status=deaths	count=X	<timestamp>

Tabuľka 1: Kumulatívne dáta o COVID-19 v rámci celej Českej republiky

total_deaths	count=X	<timestamp>
--------------	---------	-------------

Tabuľka 2: Celkový počet mŕtvych v Českej republike

regional,nuts=CZ010	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ020	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ031	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ032	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ041	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ042	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ051	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ052	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ053	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ063	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ064	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ071	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ072	infected=X,recovered=X,deaths=X	<timestamp>
regional,nuts=CZ080	infected=X,recovered=X,deaths=X	<timestamp>

Tabuľka 3: Kumulatívne dáta o COVID-19 podľa krajov Českej republiky