

ENGETO

Datová analýza šíření nákazy COVID19

První výsledky



Pavel Fryblík

Data Analyst Lead @ MADFINGER Games

Agenda

- Zase trochu jiná vizualizace v Tableau public
- Kdy se nám zastaví přírůstky nakažených? Odhad budoucího vývoje
- Jak jednotlivé kraje bojují s nákazou a jak bojuje svět? Trend za posledních 14 dní

COLAB - <https://colab.research.google.com/drive/1ussoV2e5L1TDVTx1gkpZVDjFleZr44b5#scrollTo=lpFQrQbmRTO3>

Zdroje dat

- Data za Českou republiku: <https://onemocneni-aktualne.mzcr.cz/covid-19>
- Kódy krajů: https://www.czso.cz/documents/10180/131222971/cz_nuts_systematicka_cast.xlsx/ba4a914d-74ec-48cb-b875-470098520a26?version=1.1
- Data ze světa: <https://opendata.ecdc.europa.eu/covid19/casedistribution/csv/>
- Kódy států: <https://datahub.io/JohnSnowLabs/country-and-continent-codes-list#python>

Zase trochu jiná vizualizace

Zase trochu jiná vizualizace

- Vizualizace dat Česká republika

<https://public.tableau.com/profile/pavel.frybl.k#!/vizhome/COVID-19Czechrepublicdata/Pokrajich>

<https://public.tableau.com/profile/patrik.ma.enka?fbclid=IwAR1IGcgfwUI4womGlioESNGwXvtzz0ALeQyXpe4ReOxJfvBsVOvev92ugQ8#!/vizhome/Covid-19PehledesckRepublika/Covid-19Pehledesko?publish=yes>

- Vizualizace dat ze světa

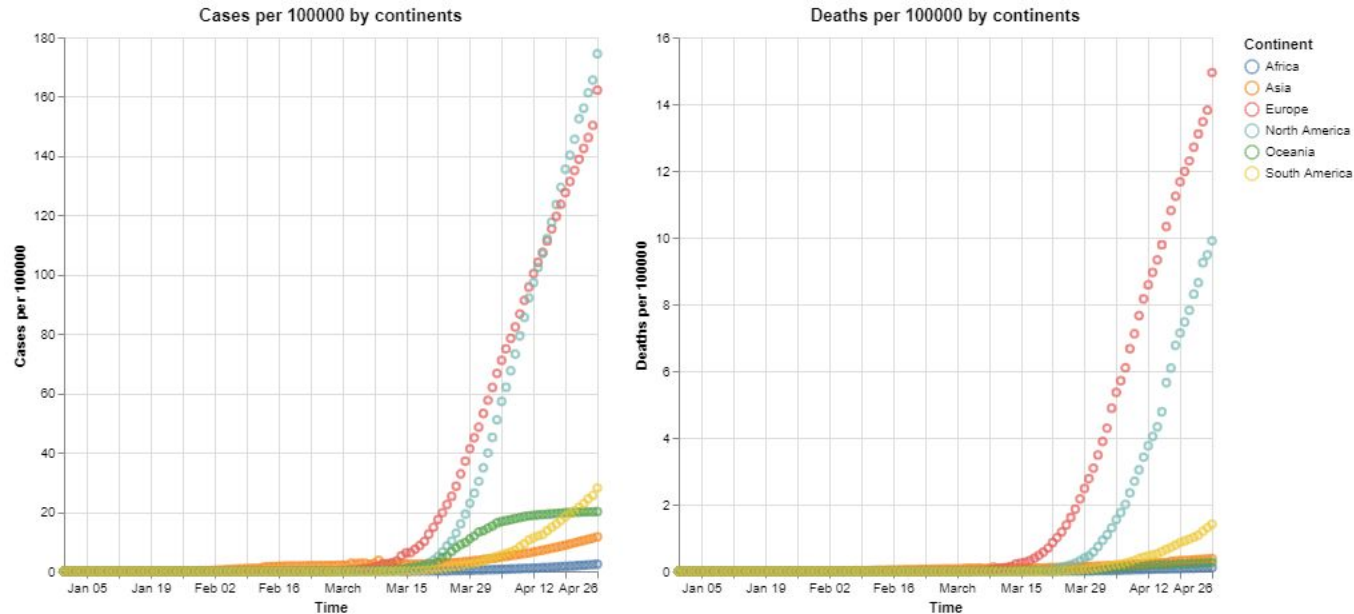
https://public.tableau.com/profile/pavel.frybl.k#!/vizhome/CovidWorld_15874834091100/World-COVID19

<https://colab.research.google.com/drive/1ussoV2e5L1TDVTx1qkpZVDjFleZr44b5#scrollTo=lpFQrQbmRTO3>

Zase trochu jiná vizualizace

- Vizualizace dat ze světa - vývoj v čase po kontinentech na 10000 obyvatel

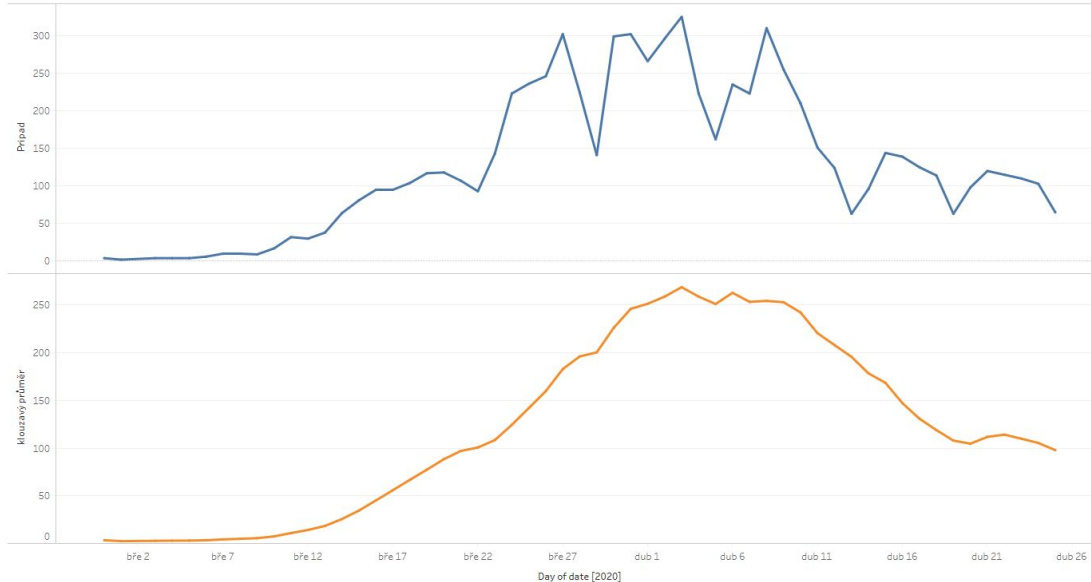
Zdroj: colab <https://colab.research.google.com/drive/1ussoV2e5L1TDVTx1qkpZVDjFleZr44b5#scrollTo=lpFQrQbmRTO3>



Kdy se zastaví přírůstky nakažených v ČR?

Kdy se zastaví přírůstky nakažených v ČR?

Vyvoj + klouzavý průměr



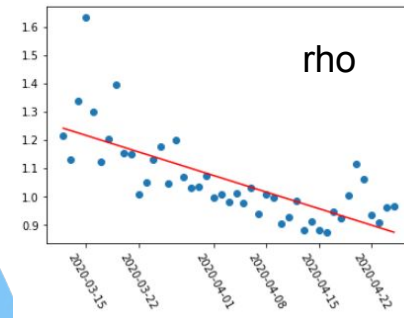
Obvyklá časová řada s konstantním procesem

$$Y_t = \rho \cdot Y_{t-1} + \varepsilon$$

Naše časová řada, proces se mění v čase - opatření apod.

$$Y_t = \rho(t) \cdot Y_{t-1} + \varepsilon$$

$$\rho = a \cdot t + b$$



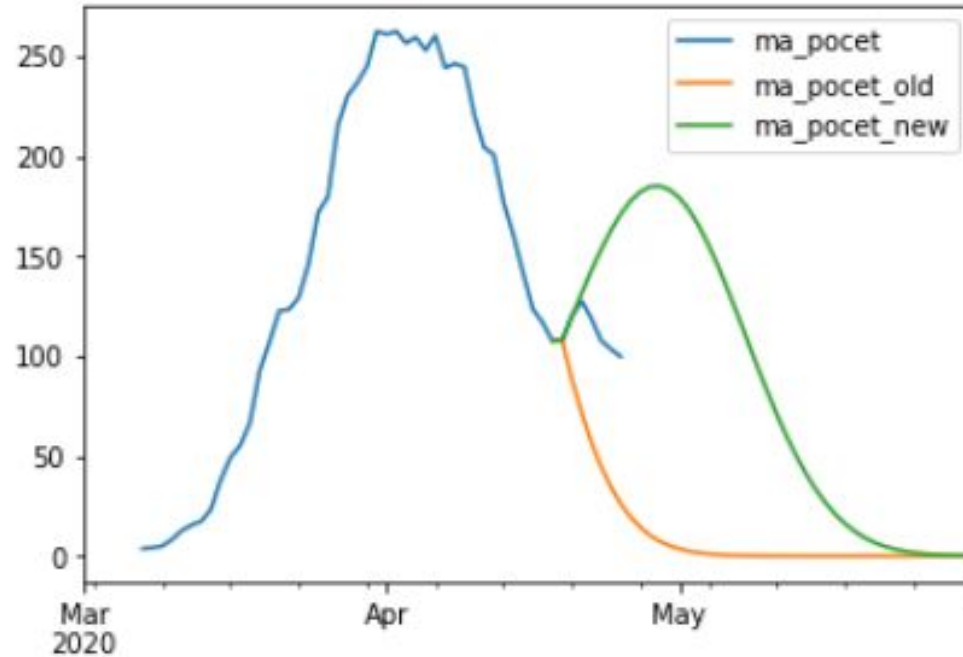
Kdy se zastaví přírůstky nakažených v ČR?

Počítáno 19.4.2020

Odhad 1: data několik dní po Velikonocích

Počet nakažených
7000

Konec přírůstků
5.5.2020



Kdy se zastaví přírůstky nakažených v ČR?

Počítáno 21.4.2020

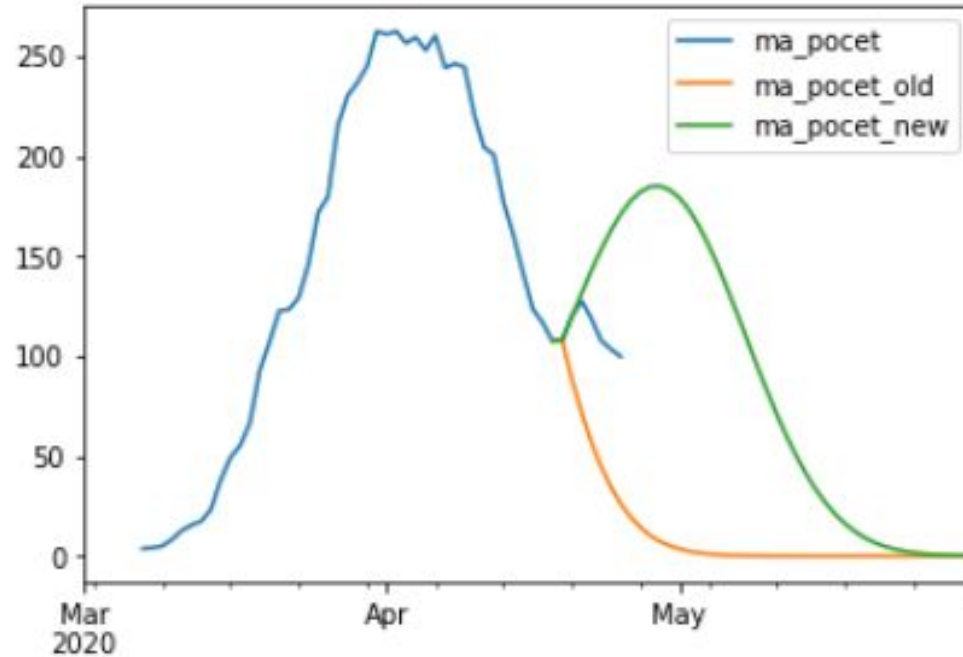
Odhad 2: Velikonoční šok - šok v rho

Konec přírůstků
27.5.2020

Šok se nakonec nepotvrdil!
Nedostatek modelu - přidat šok v počtu nakažených!

$$Y_t = \rho(t) \cdot Y_{t-1} + \varepsilon$$

$$\rho = a \cdot t + b$$



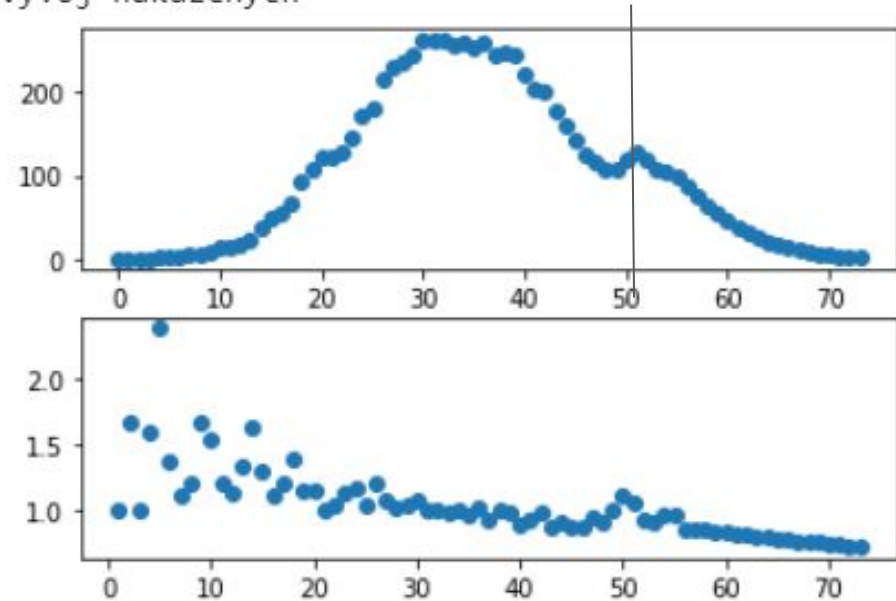
Kdy se zastaví přírůstky nakažených v ČR?

**Odhad 3: data po druhou
“vlnku” a původní model**

Počet nakažených
7860

Konec přírůstků
15.5.2020

Vyvoj nakazených



Kdy se zastaví přírůstky nakažených v ČR?

Validace modelů

Chyba 'původní rho': 34.35796877534158

Chyba 'rho se šokem': 48.1962602336551

Chyba ARIMA modelu: 43.10677374086128

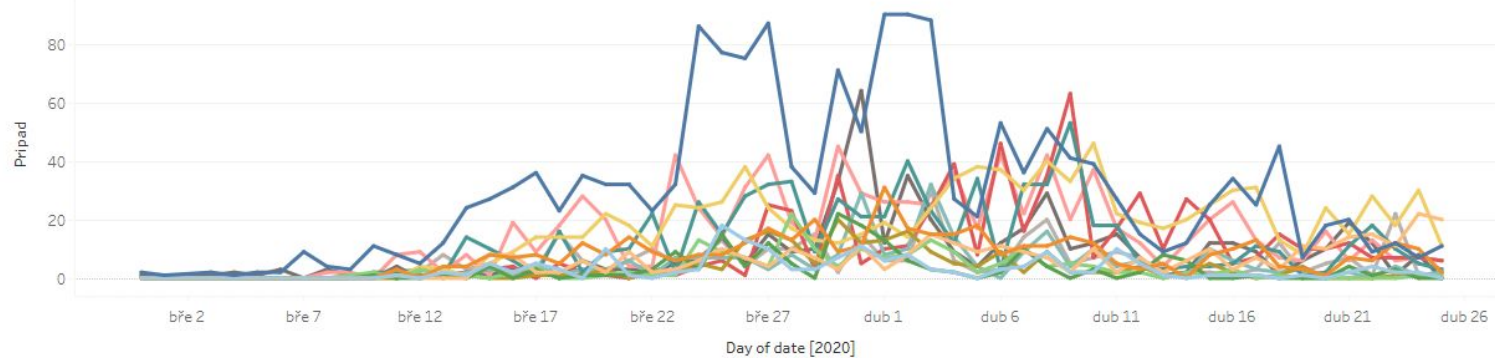


Jak s nákazou bojují jednotlivé kraje?

Jak s nákazou bojují jednotlivé kraje?

Někdy jen vizualizace nestačí. Pomůžeme si datovou analýzou.

Počet po krajích



Jak s nákazou bojují jednotlivé kraje?

DASK + posledních 14 dní + trend (regrese)

kraj	
Hlavní město Praha	0
Jihomoravský kraj	0
Jihočeský kraj	1
Karlovarský kraj	1
Kraj Vysočina	0
Královéhradecký kraj	0
Liberecký kraj	0
Moravskoslezský kraj	0
Olomoucký kraj	1
Pardubický kraj	0
Plzeňský kraj	0
Středočeský kraj	0
Zlínský kraj	1
Ústecký kraj	1



Jak s nákazou bojuje svět?

DASK + posledních 14 dní + trend (regrese)

PRACUJEME NA VIZUALIZACI

Děkuji!



Pavel Fryblík
Data Analyst Lead @ MADFINGER Games