

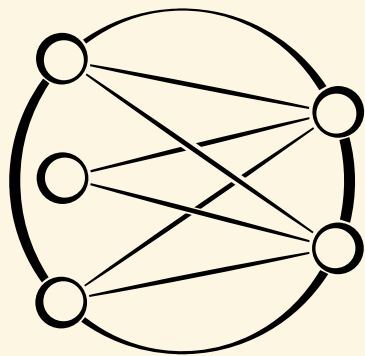
# ZARAŻLIWOŚĆ KORONAWIRUSA

Analiza rozprzestrzeniania się koronawirusa w  
początkowym okresie pandemii

Patryk Gronkiewicz

Piotr Krawiec

**O NAS**



KOŁO NAUKOWE  
**MACHINE LEARNING**  
POLITECHNIKI RZESZOWSKIEJ

Logo koła naukowego

# NASZ CEL

Znalezienie reguł kierujących rozwojem pandemii

**DANE**

Do głównych źródeł danych zaliczyć możemy:

Do głównych źródeł danych zaliczyć możemy:

- IHME

Do głównych źródeł danych zaliczyć możemy:

- IHME
- [koronawirusunas.pl](https://koronawirusunas.pl)



Do głównych źródeł danych zaliczyć możemy:

- IHME
- [koronawirusunas.pl](https://koronawirusunas.pl)
- [policja.pl](https://policja.pl)

**DLACZEGO NIE DANE  
RZĄDOWE?**

Dwa główne problemy:

## Dwa główne problemy:

- Na stronie ministerstwa publikowane były wyłącznie dane z danego dnia.

## Dwa główne problemy:

- Na stronie ministerstwa publikowane były wyłącznie dane z danego dnia.
- Publicznie dostępne dane dotyczące wielu krajów nie zawierały podziału na województwa.

[← Powrót](#)

## Mapa zarażeń koronawirusem (SARS-CoV-2)

MAPA ZACHOROWAŃ

TABELA ZACHOROWAŃ

Dane pochodzą z Ministerstwa Zdrowia, aktualne na : 07.04.2020 10:00



[← Powrót](#)

## Mapa zarażeń koronawirusem (SARS-CoV-2)

[MAPA ZACHOROWAŃ](#)[TABELA ZACHOROWAŃ](#)

Dane pochodzą z Ministerstwa Zdrowia, aktualne na : 07.04.2020 10:00

Liczba elementów: **16**

Województwo	Liczba	Liczba zgonów	Akcje
dolnośląskie	507	13	<a href="#">zobacz szczegóły</a>
kujawsko-pomorskie	245	1	<a href="#">zobacz szczegóły</a>
lubelskie	191	7	<a href="#">zobacz szczegóły</a>
lubuskie	62		<a href="#">zobacz szczegóły</a>
łódzkie	340	3	<a href="#">zobacz szczegóły</a>
małopolskie	363	8	<a href="#">zobacz szczegóły</a>
mazowieckie	1065	23	<a href="#">zobacz szczegóły</a>
opolskie	101	2	<a href="#">zobacz szczegóły</a>
podkarpackie	159	7	<a href="#">zobacz szczegóły</a>
podlaskie	114	1	<a href="#">zobacz szczegóły</a>
pomorskie	127		<a href="#">zobacz szczegóły</a>
śląskie	594	26	<a href="#">zobacz szczegóły</a>
świętokrzyskie	110	1	<a href="#">zobacz</a>

# **TECHNIKI ZBIERANIA DANYCH**





Python 3



BeautifulSoup4

**API**

API używamy do pobierania danych pogodowych.  
Pochodzą one z [Meteostatu](#), a pobieramy je dla  
Warszawy.

**JAK PRACOWALIŚMY Z  
DANYMI?**



File Edit View Run Kernel Tabs Settings Help

+

Filter files by name

/ notebooks /

Name	Last Modified
data	25 days ago
Analiza.ipynb	2 minutes ago
ImportDan...	25 days ago

Launcher

Analiza.ipynb

Python 3 (ipykernel)

from corona\_analysis.scrapers import healthdata, koronawirusunas, meteostat, policjapl, rmf24, wiki\_demo, wiki\_urban

## Wczytanie danych

[45]:

df = pd.read\_csv(Path(paths.get\_path(), "df.csv"), index\_col=0, parse\_dates=True, dayfirst=True)

Sprawdzam czy w danych nie ma braków, a jeżeli są to uzupełniam je 0.

[46]:

df.head(3)

	Unnamed: 0.1	Interwencje	Zatrzymani na gorącym uczynku	Zatrzymani poszukiwani	Zatrzymani nietrzeźwi kierujący	Wypadki drogowe	Zabici w wypadkach	Ranni w wypadkach	temperature	temperature_min	...	chor	zgo	wyl	pieszo	pojazdem	hosp	kwar	kwar_z
2020-01-01	12.0	19145.0	734.0	188.0	332.0	42.0	1.0	46.0	1.4	0.8	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2020-01-02	11.0	14654.0	513.0	333.0	166.0	60.0	5.0	67.0	0.8	-1.7	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2020-01-03	10.0	15507.0	468.0	405.0	122.0	62.0	3.0	69.0	0.5	-4.1	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

3 rows × 39 columns

[47]:

df.tail(3)

	Unnamed: 0.1	Interwencje	Zatrzymani na gorącym uczynku	Zatrzymani poszukiwani	Zatrzymani nietrzeźwi kierujący	Wypadki drogowe	Zabici w wypadkach	Ranni w wypadkach	temperature	temperature_min	...	chor	zgo	wyl	pieszo	pojazdem	hosp	kwa
2020-05-26	1.0	15783.0	562.0	281.0	144.0	48.0	3.0	54.0	13.9	8.0	...	11030.0	1024.0	10020.0	-31.65	-5.06	2171.0	78864.0
2020-05-27	0.0	16226.0	596.0	311.0	186.0	60.0	6.0	65.0	14.8	8.0	...	11115.0	1028.0	10330.0	-31.65	-5.06	2320.0	79007.0
2020-05-28	0.0	16226.0	596.0	311.0	186.0	60.0	6.0	65.0	15.9	8.0	...	11227.0	1038.0	10560.0	-31.65	-5.06	2248.0	79818.0

3 rows × 39 columns

[48]:

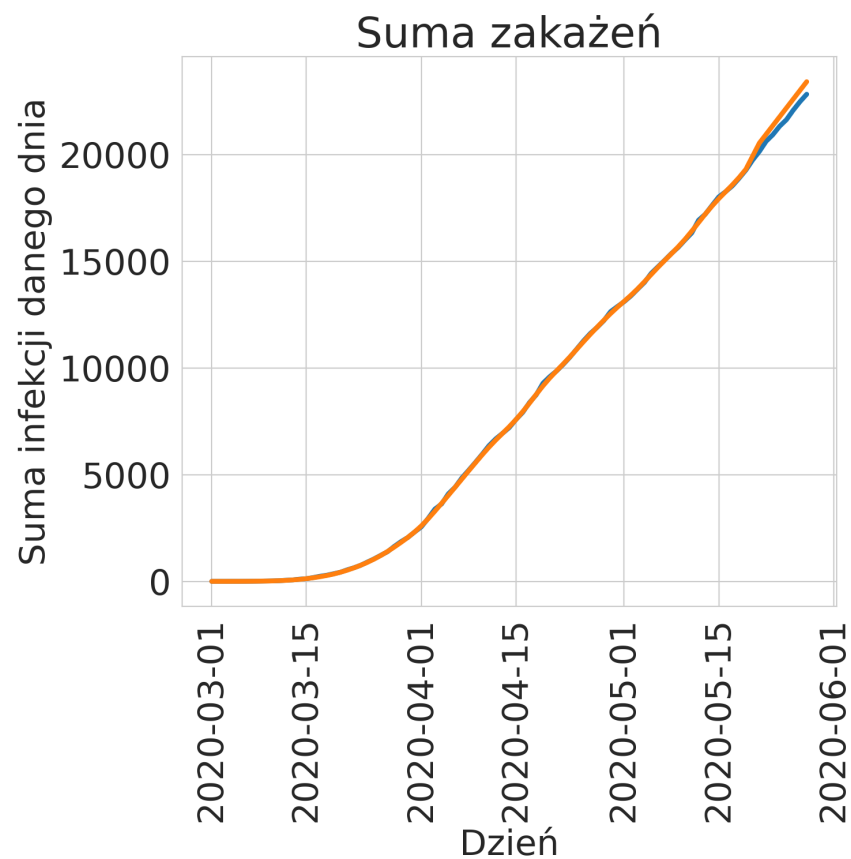
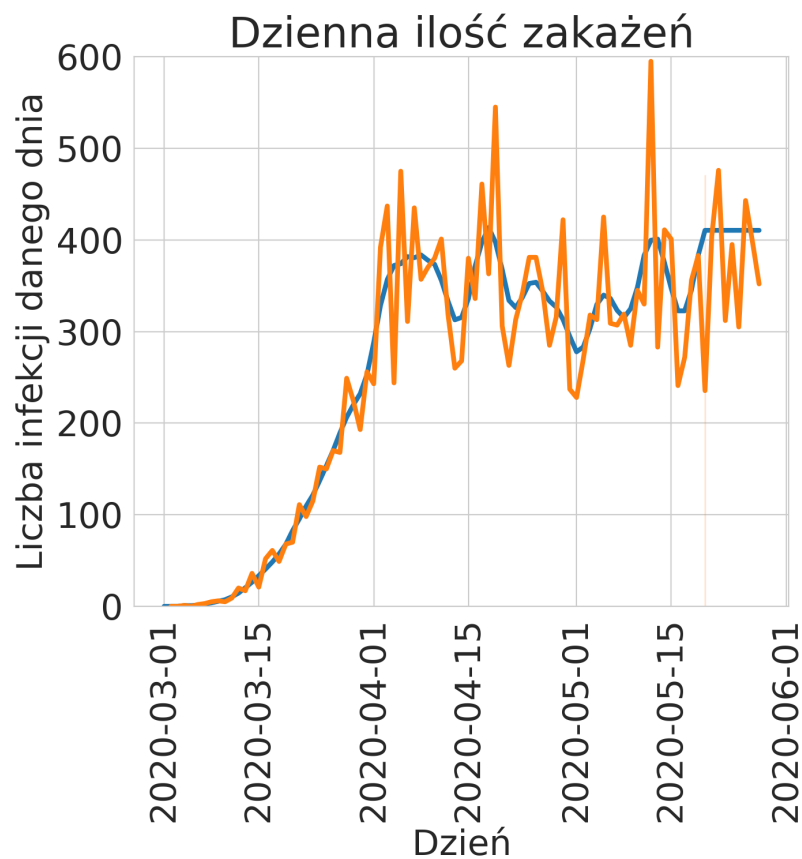
woj = koronawirusunas.SUB\_SITES  
województwa = pd.DataFrame()  
for i in range(1, len(woj)):  
 dat = pd.read\_csv(Path(paths.get\_path(), "woj[%d].1.csv") % i, index\_col=0, parse\_dates=True, dayfirst=True)

Simple

0 1

Python 3 (ipykernel) | Idle

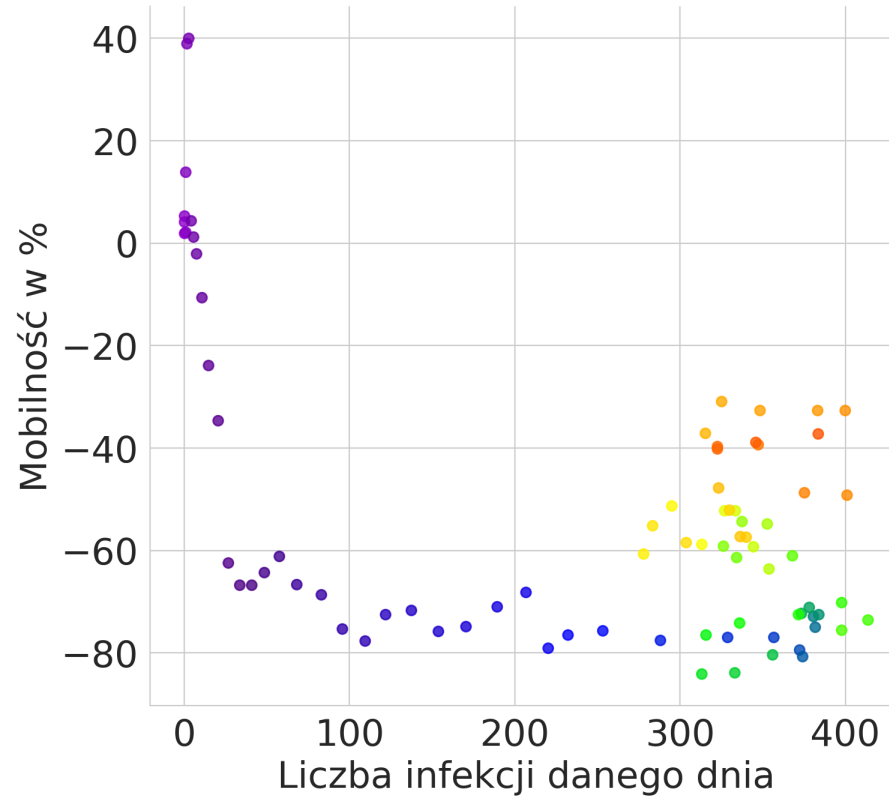
**ANALIZA DANYCH**



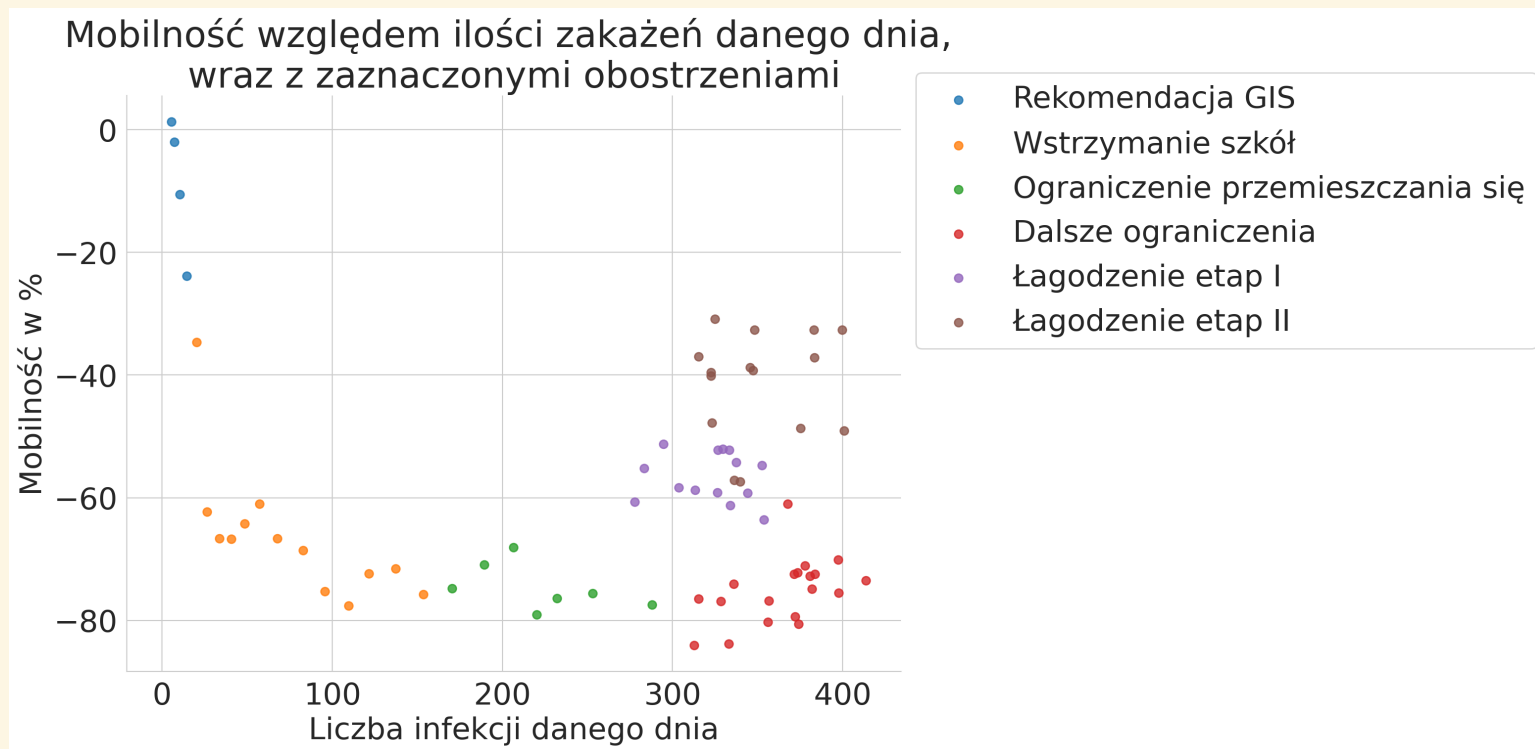
Porównanie IHME i koronawirusunas



Mobilność względem ilości zakażeń danego dnia w ujęciu czasowym  
Czas w skali koloru - od fioletu do czerwieni.



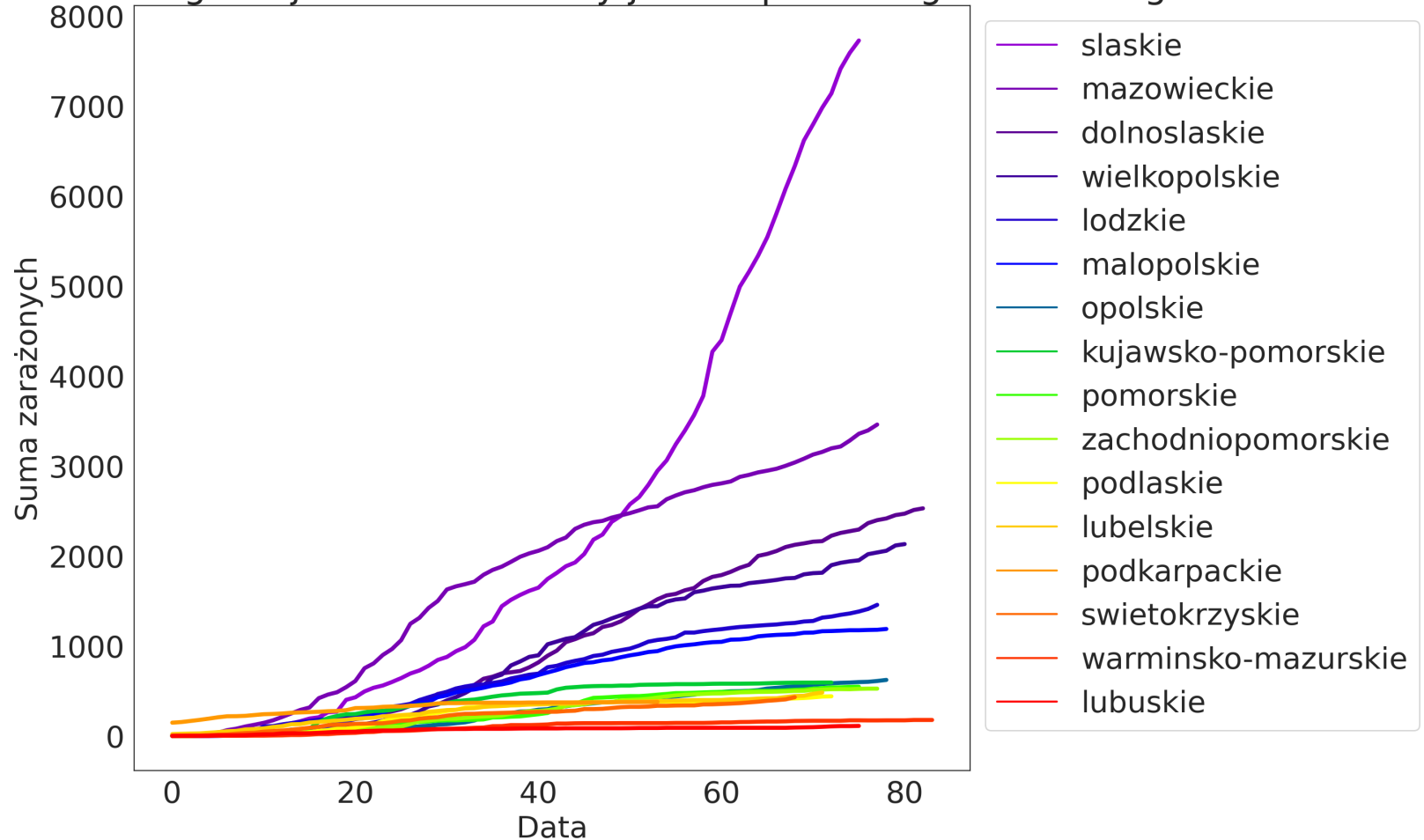
Porównanie faz



# Mobilność i infekcje, a rekomendacje rządowe

## Liczba zarażonych w czasie

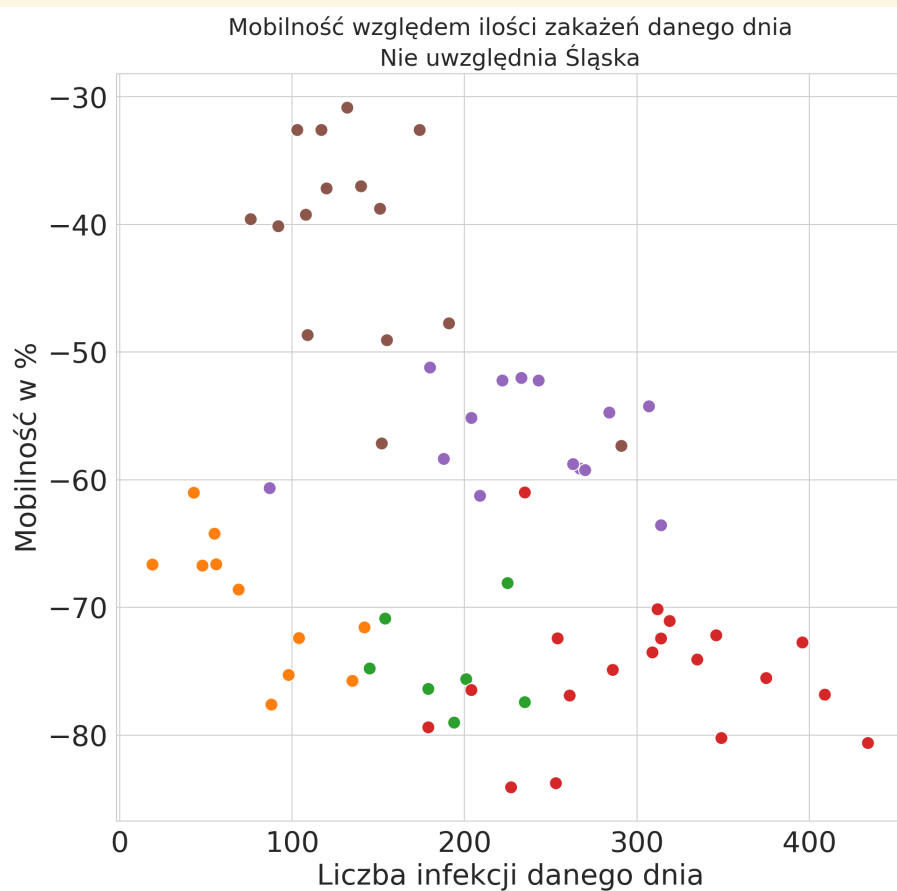
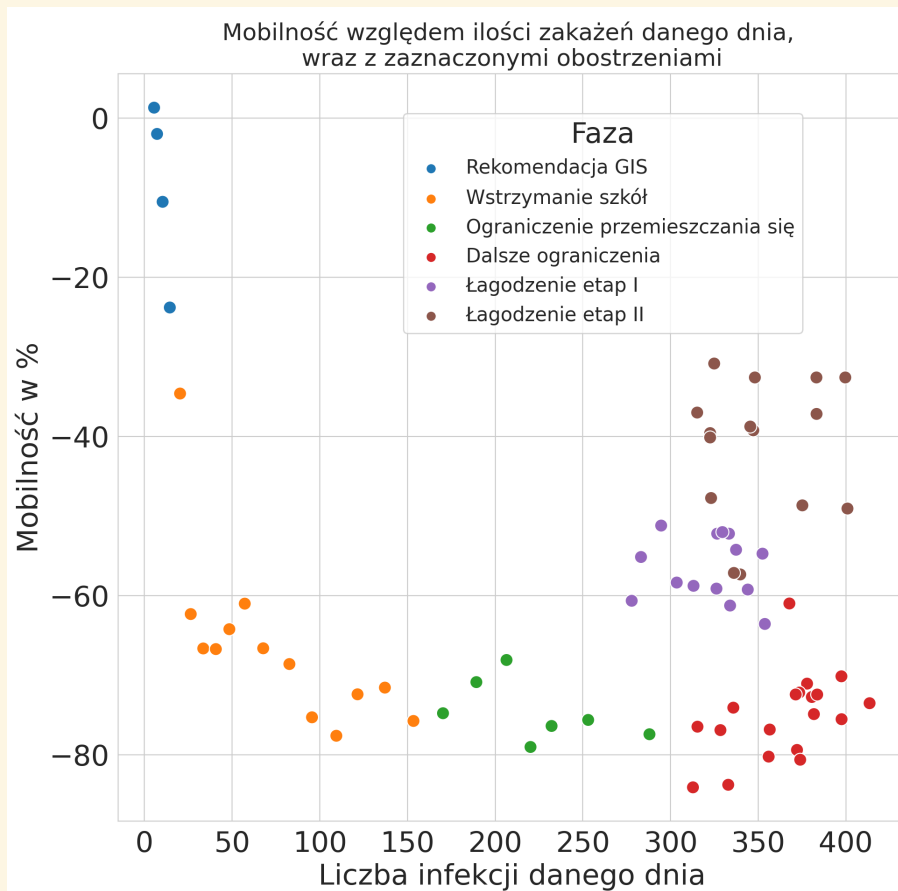
Czas dla każdego województwa liczony jest od pierwszego zakażonego



Wykres województwa

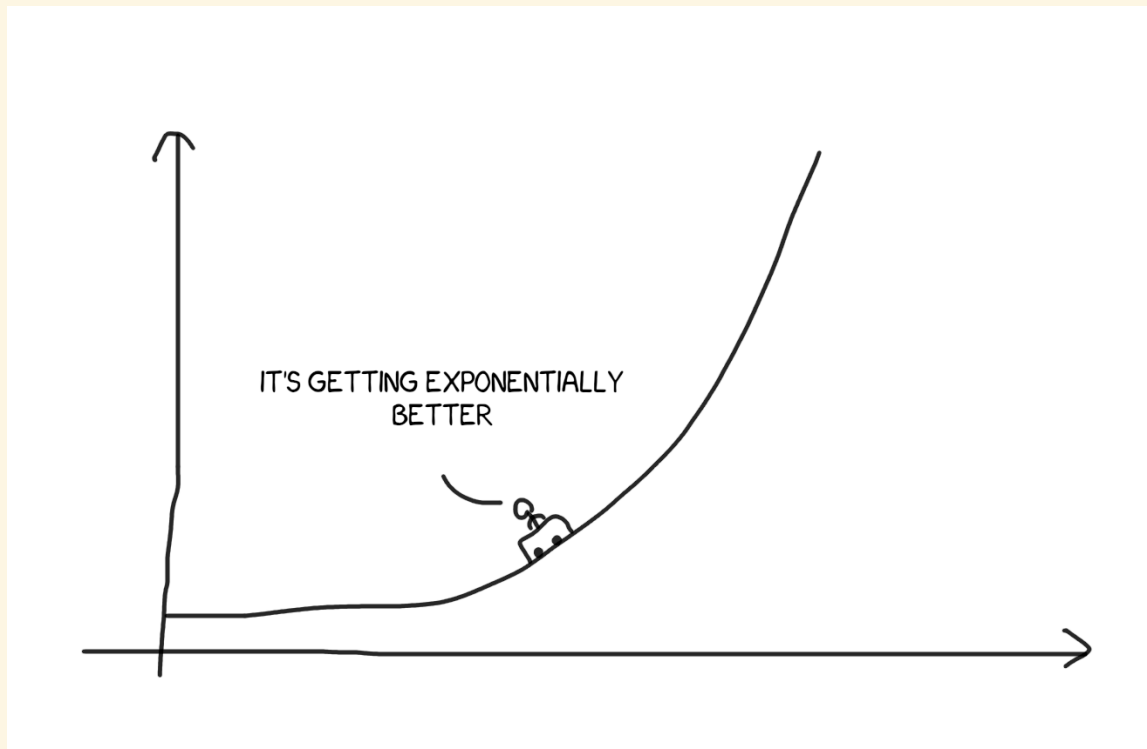
**A CO GDYBY...**

**“POZBYĆ SIĘ” ŚLĄSKA?**



Fazy w czasie bez śląska

**WNIOSKI**



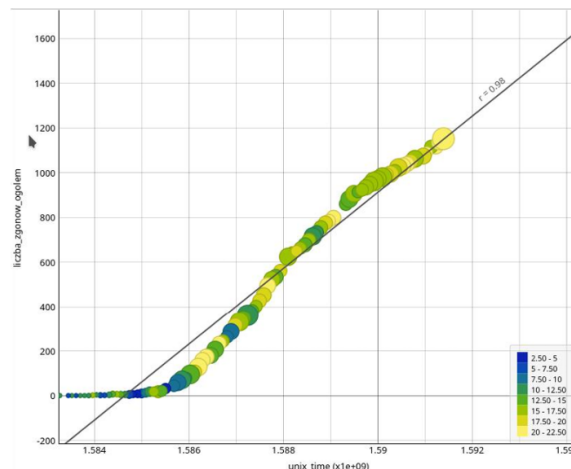
Wykres funkcji wykładniczej (inspirowane XKCD)



**CZEGO SIĘ  
NAUCZYLIŚMY?**

**CO INNEGO  
ZYSKALIŚMY?**

Wykres zachorowań w czasie (styczeń–maj) wraz z temperaturą (kolor). Wielkość okręgów to nowe zachorowania. Wynik pracy wygenerowany podczas hackathonu w maju 2020 r. Autor Piotr Gul.



## Machine Learning nie zwalnia podczas zdalnego nauczania

Patryk Gronkiewicz  
dr Michał Piętał  
dr Ewa Rejwer-  
-Kosińska

Koło Naukowe Machine Learning zostało oficjalnie utworzone w październiku ubiegłego roku, czyli w okresie największego obciążenia przez naukę zdalną. Jednak jego działalność można było już zauważyć w marcu i kwietniu tego samego roku, kiedy to studenci wzięli udział w pierwszej edycji hackathonu, który dotyczył badań nad wpływem różnych czynników na rozwój pandemii w kraju.

Koło ma głównie na celu popularyzację i rozwój wiedzy na temat Machine Learning, a także tematów ściśle z nim związanych, takich jak ogólnie pojęte Data Science, wizualizacja danych, sztuczna inteligencja czy analiza dużych zbiorów danych. Dzięki działalności koła w semestrze zimowym 2020/2021 udało się zrealizować prelekcje na różne tematy, które przedstawiali zarówno pracownicy Politechniki Rzeszowskiej, jak i firm zewnętrznych. W zimowym cyklu spotkań zostały omówione takie tematy, jak: analiza danych bioinformatycznych, wizualizacja baz danych, teoria machine learning, podstawy sieci neuronowych, detekcja anomalii zarówno w sesjach

WWW, jak i w sieciach elektroenergetycznych, a także zostały przeprowadzone warsztaty z narzędzia MS PowerBI. Przedstawiono również platformy do konkursów data science, na których poszukiwane jest najlepsze rozwiązanie problemów związanych ze zbiorami danych, które na pierwszy rzut oka nie są proste do klasyfikacji.

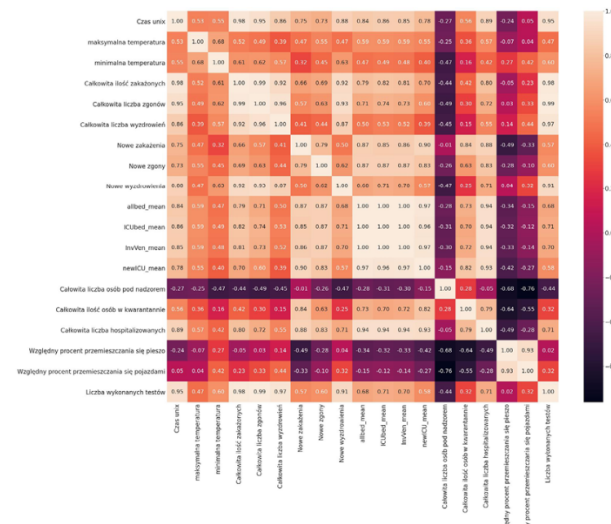
W semestrze letnim dla równowagi do zajęć teoretycznych prowadzony jest hackathon. „Maraton programistyczny oraz analiz danych w tym roku akademickim występuje pod

szłydem „Open Gov Data”, co oznacza otwarte repozytoria danych publicznych, tj. dane dostępne dla każdego, pochodzące ze źródeł publicznych, ministerstw, województw, gmin, urzędów i innych instytucji rządowych. Dane są dostępne w Polsce pod adresem: dane.gov.pl oraz pochodzące z UE: ec.europa.eu/eurostat/data/database, choć to są tylko przykłady. Początki otwartych danych publicznych datuje się od tzw. dyrektywy re-use z 2003 r. (Dyrektywa 2003/98/WE Parlamentu Europejskiego i Rady w sprawie ponownego wykorzystywania informacji sektora publicznego)” – mówi dr Michał Piętał z Wydziału Elektrotechniki i Informatyki Politechniki Rzeszowskiej.

Zadaniem zespołów startujących w hackathonie jest stworzenie funkcjonalnej aplikacji (choć może to być prototyp), która wykorzystuje repozytoria otwartych danych rządowych, jest innowacyjna oraz stosuje zaawansowany aparat matematyczno-informatyczny. Przesłanki te są jednocześnie kryteriami oceny projektów przez mentora hackathonu, którymi są dr Ewa-Rejwer-Kosińska, dr Michał Piętał oraz dr Marcin Kowalik. Mentorzy spotykają się z wszystki-

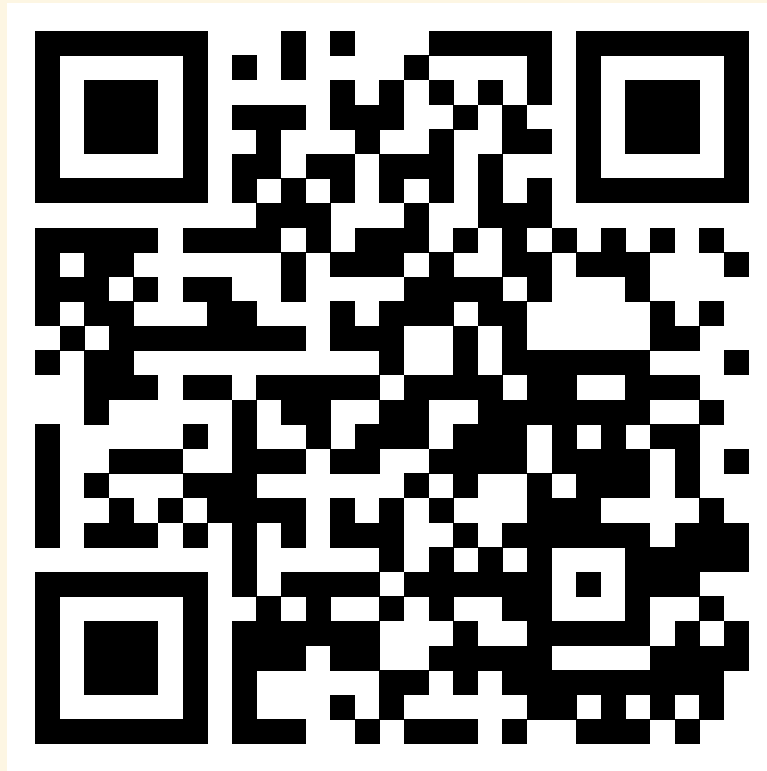
mi zespołami online, osobno, co tydzień, celem – podobnie jak to ma miejsce na innych hackathonach – udzielania odpowiedzi na nurtujące pytania członków zespołów, a dotyczące takich aspektów, jak: przetwarzanie i dostęp do danych, używane technologie teleinformatyczne czy aspekty aplikacyjne/biznesowe aplikacji. Zgodnie z „Regulaminem własności intelektualnej na PRz” powstałe w ramach hackathonu aplikacje są własnością studentów.

Hackathon trwa ponad trzy miesiące (począwszy od 15 marca br.) i jest przeznaczony dla członków Koła Naukowego Machine Learning (udział jest dobrowolny). Biorą w nim udział trzy zespoły: RTFD, Wściekle Perceptrony oraz Olimpijczycy. „Uczestnikami hackathonu jest 12 studentów studiujących na pierwszym lub drugim roku kierunku inżynieria i analiza danych. Zasady hackathonu ustanowione są przez regulamin, który został skonsultowany i zaakceptowany przez Centrum Transferu Technologii Politechniki Rzeszowskiej. Razem z pozostałymi mentorami liczymy na ciekawą i oryginalną aplikację, które powstaną w toku prac studentów w ramach Koła Naukowego ML oraz hackathonu. Nie jest wykluczona dalsza komercjalizacja tych aplikacji, które ostatecznie będą posiadały potencjał biznesowy” – podkreśla dr Michał Piętał.



Mapa korelacji pomiędzy poszczególnymi czynnikami. Wynik pracy otrzymany w trakcie hackathonu w maju 2020 r. Autor Hubert Mazur.

# REPOZYTORIUM



<https://github.com/knmlprz/corona-analysis-1>

**DZIĘKUJEMY ZA UWAGĘ**