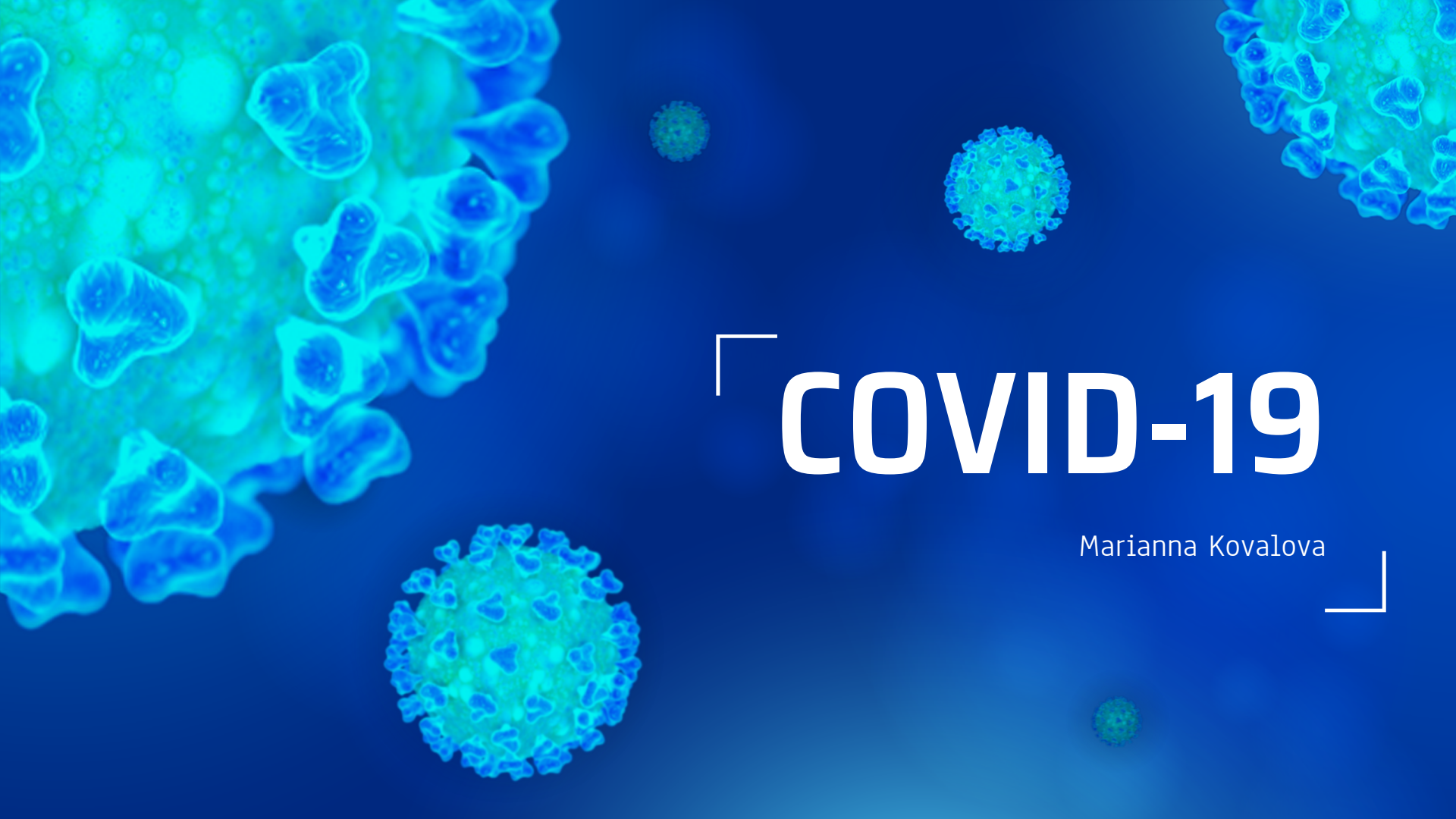

COVID-19
Marianna Kovalova

01.

ABOUT CORONAVIRUS

02.

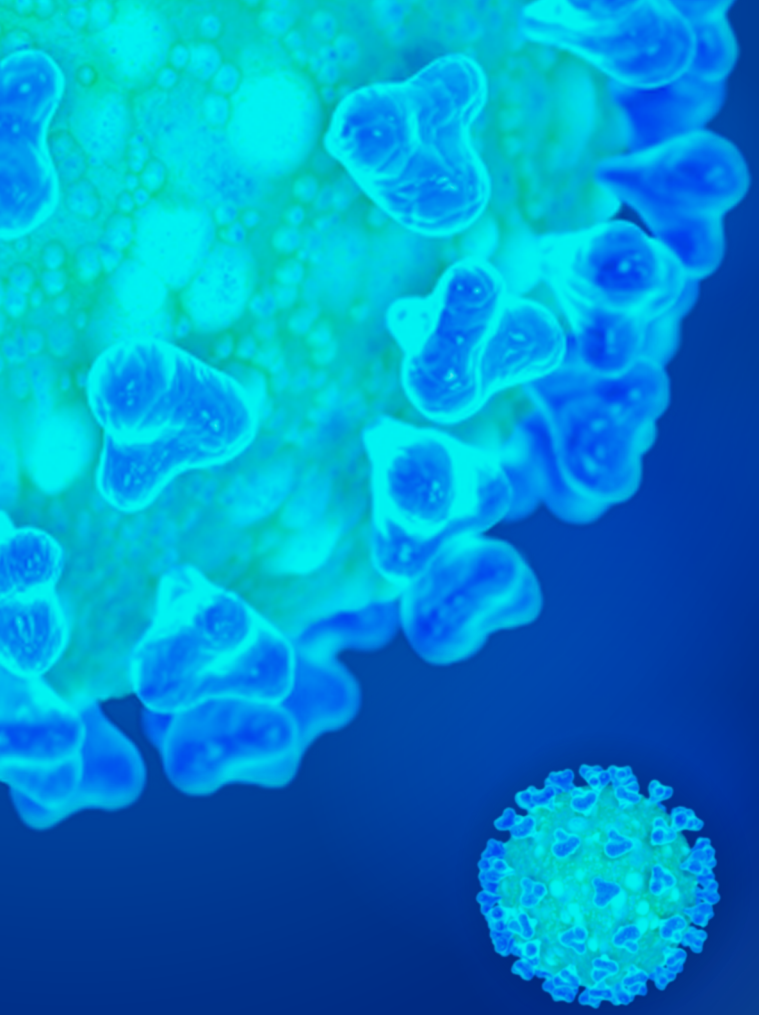
Data

03.

Methodology & Analysis

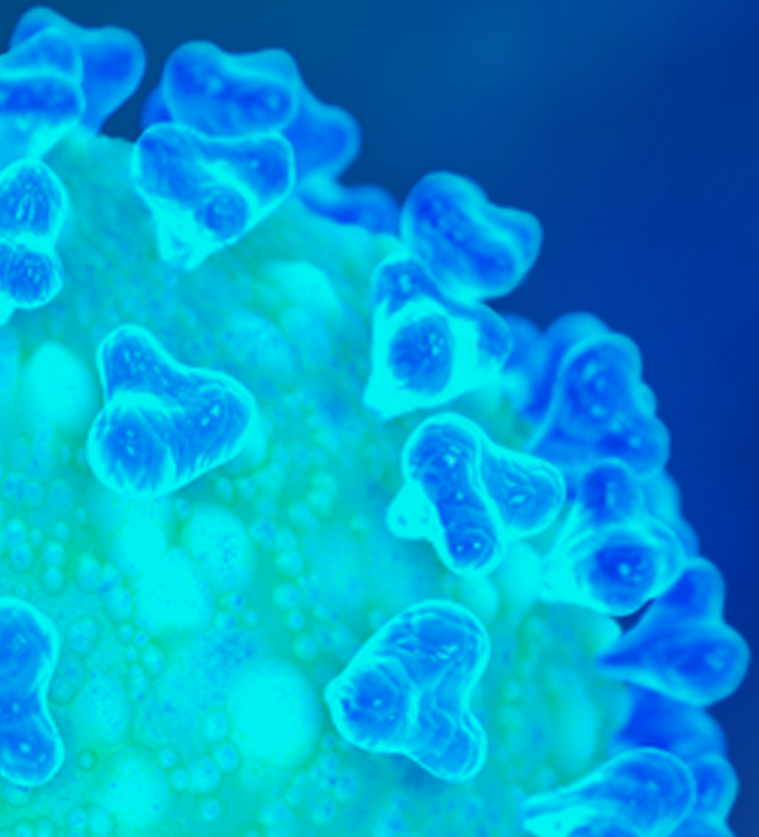
04.

Results and Discussion



01.

ABOUT CORONAVIRUS

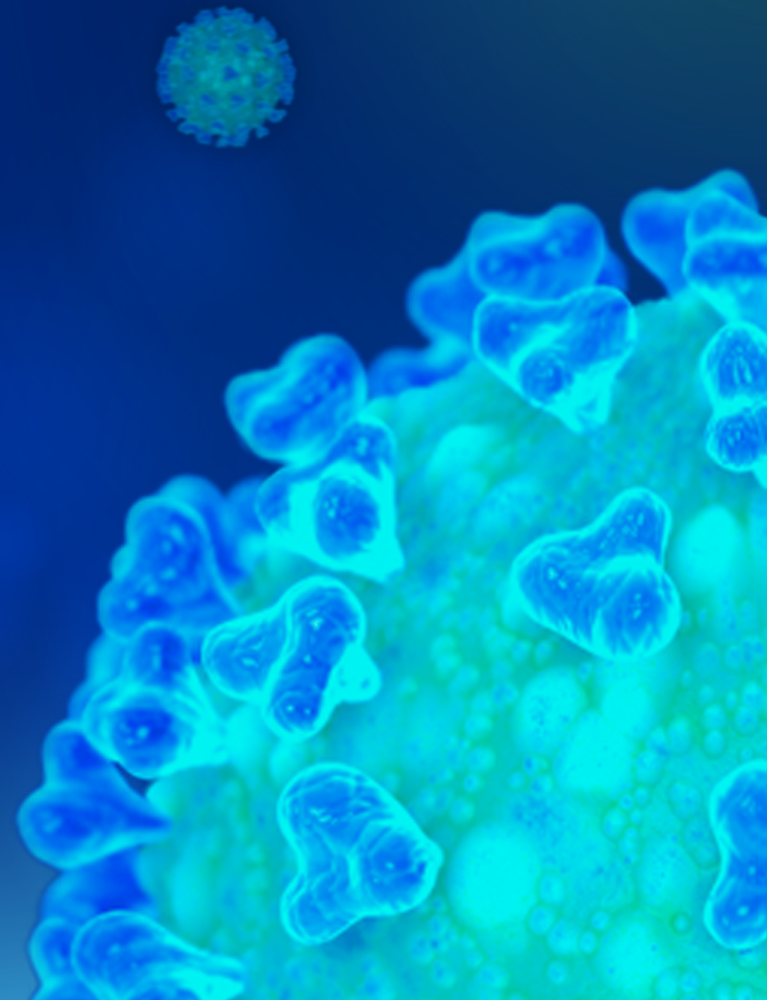


COVID-19

COVID-19 is an infectious disease caused by the recently found virus known as SARS-CoV-2 (or coronavirus). Before the outbreak originated in Wuhan, China on December 2019, there was no information about this virus

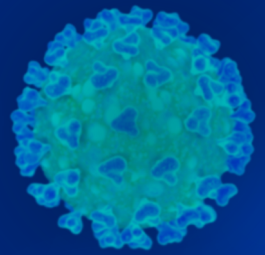
“We cannot say this loudly enough or clearly enough or often enough: All countries can still change the course of this pandemic”

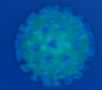
—DR. TEDROS ADHANOM GHEBREYESUS, WORLD HEALTH ORGANIZATION’S DIRECTOR GENERAL

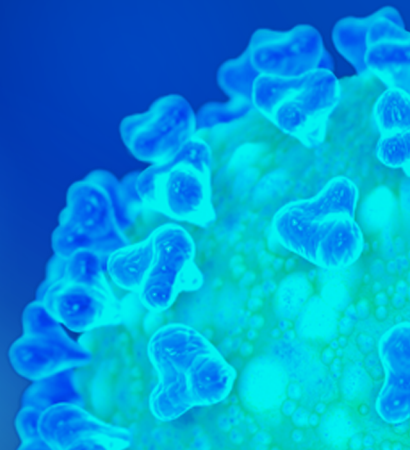


02.

Data







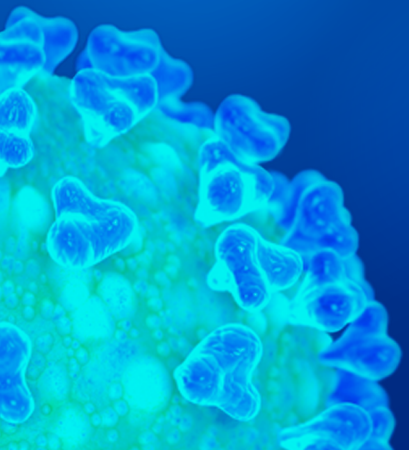
DATA

We need data for every country about:

- the number of confirmed cases
- the number of deaths
- the number of recovery cases.

Also for better analyze we need: demographic data:

- number of inhabitants for each country



JHU DATA

The main reasons cited for using the JHU data are:

- JHU is already a trusted and respected institution,
- They cite many sources, which are themselves reputable,
- The data is updated daily, and
- It is provided in an easily digestible format (.csv in a github repository). This dataset has daily level information on the number of affected cases, deaths and recovery from 2019 novel coronavirus. The number of cases on any given day is the cumulative number.

EXAMPLE OF DATA

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	...	4/17/20	4/18/20	4/19/20	4/20/20	4
0	NaN	Afghanistan	33.0000	65.0000	0	0	0	0	0	0	...	906	933	996	1026	1
1	NaN	Albania	41.1533	20.1683	0	0	0	0	0	0	...	539	548	562	584	6
2	NaN	Algeria	28.0339	1.6596	0	0	0	0	0	0	...	2418	2534	2629	2718	2
3	NaN	Andorra	42.5063	1.5218	0	0	0	0	0	0	...	696	704	713	717	7
4	NaN	Angola	-11.2027	17.8739	0	0	0	0	0	0	...	19	24	24	24	2

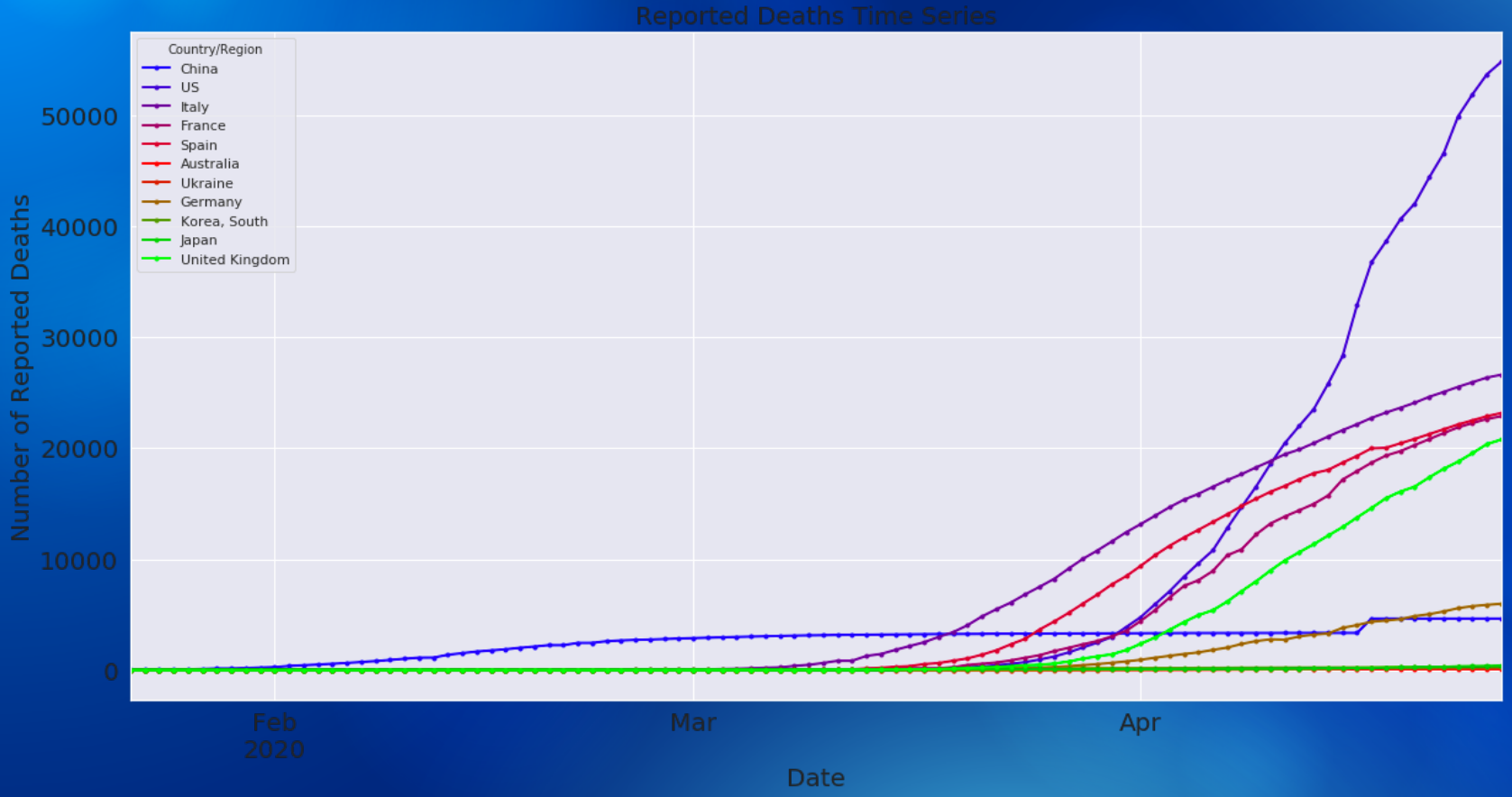
LINK: https://github.com/CSSEGISandData/COVID-19

03.

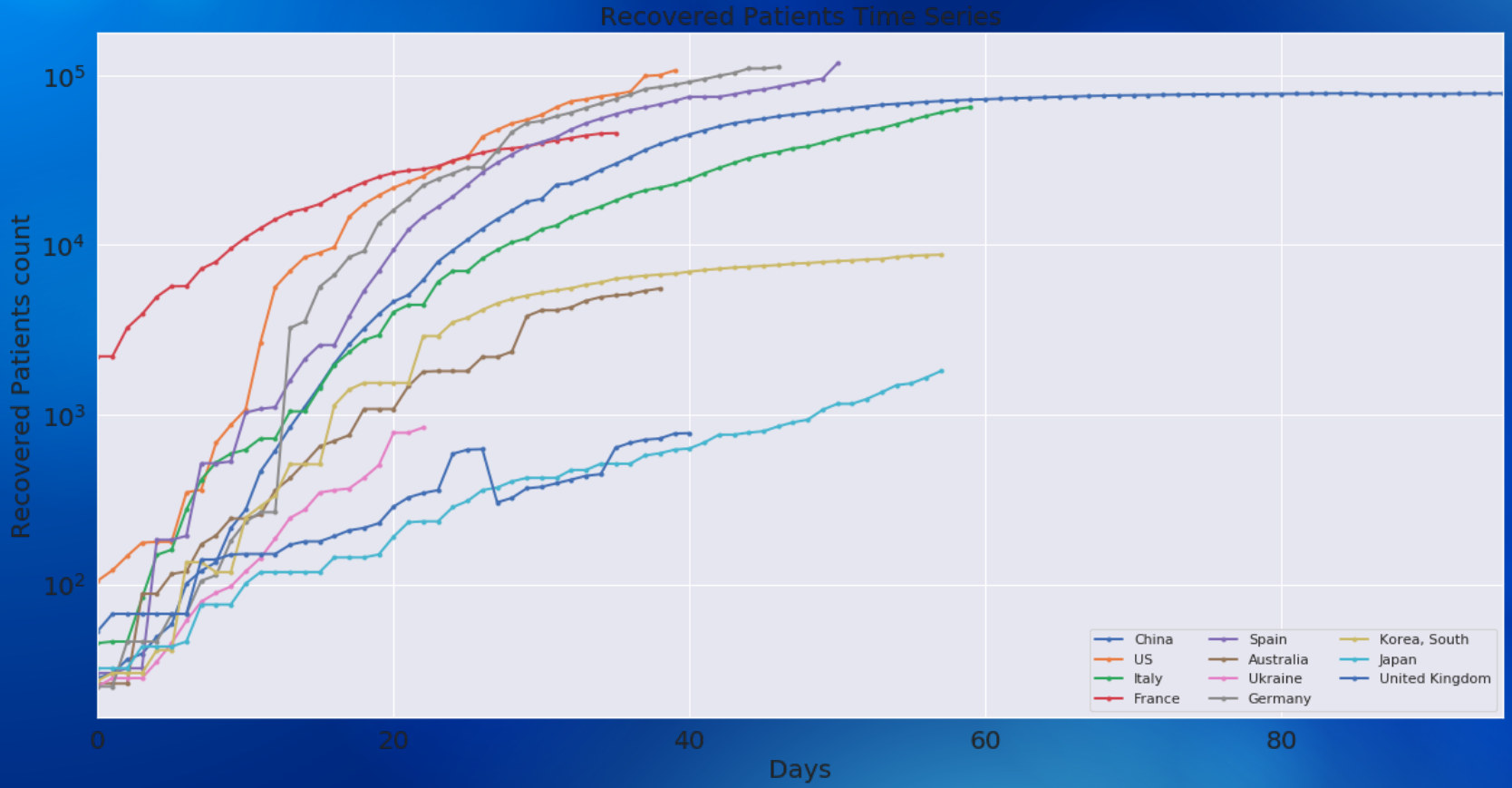
Methodology & Analysis

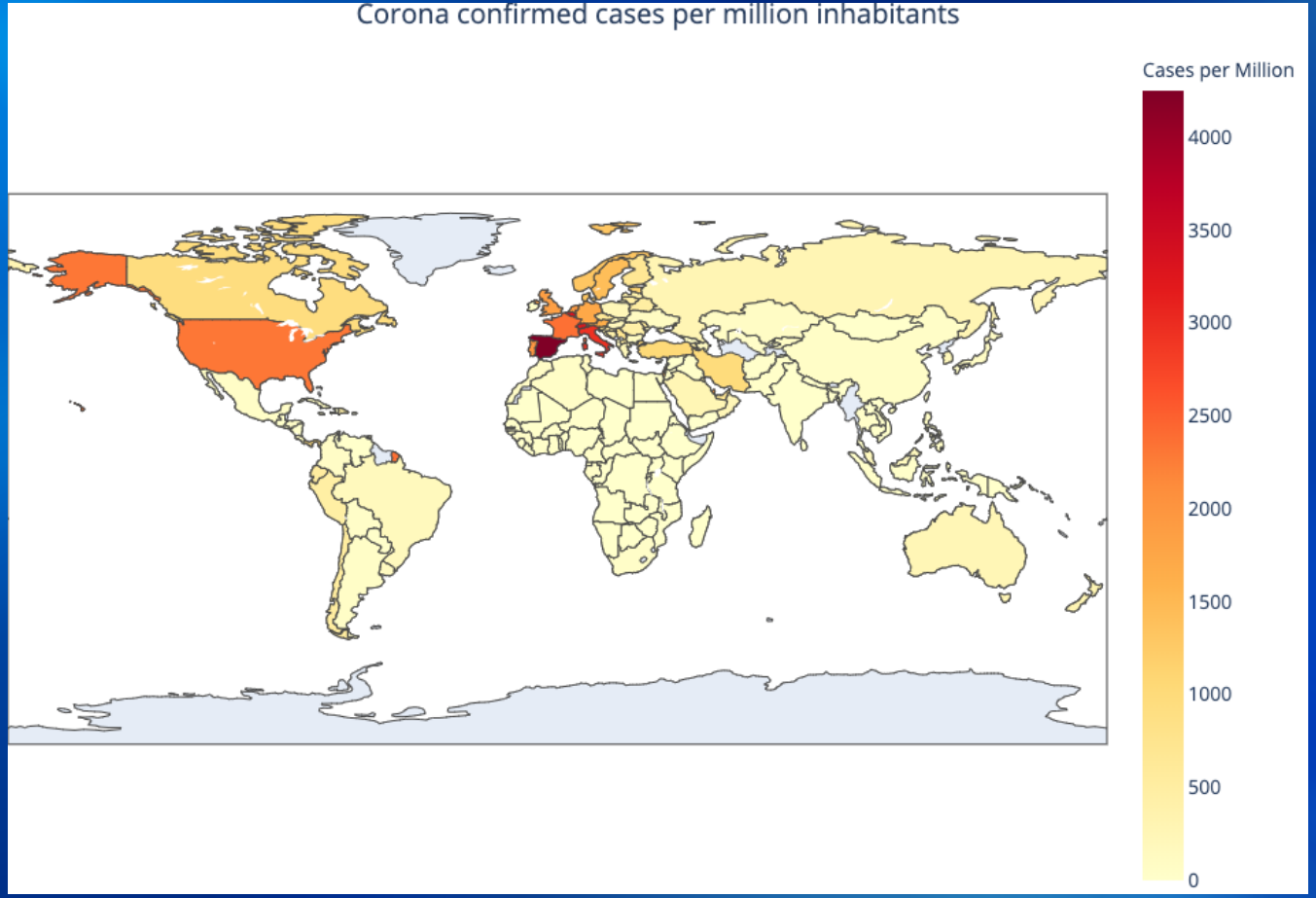
Reported Confirmed cases count

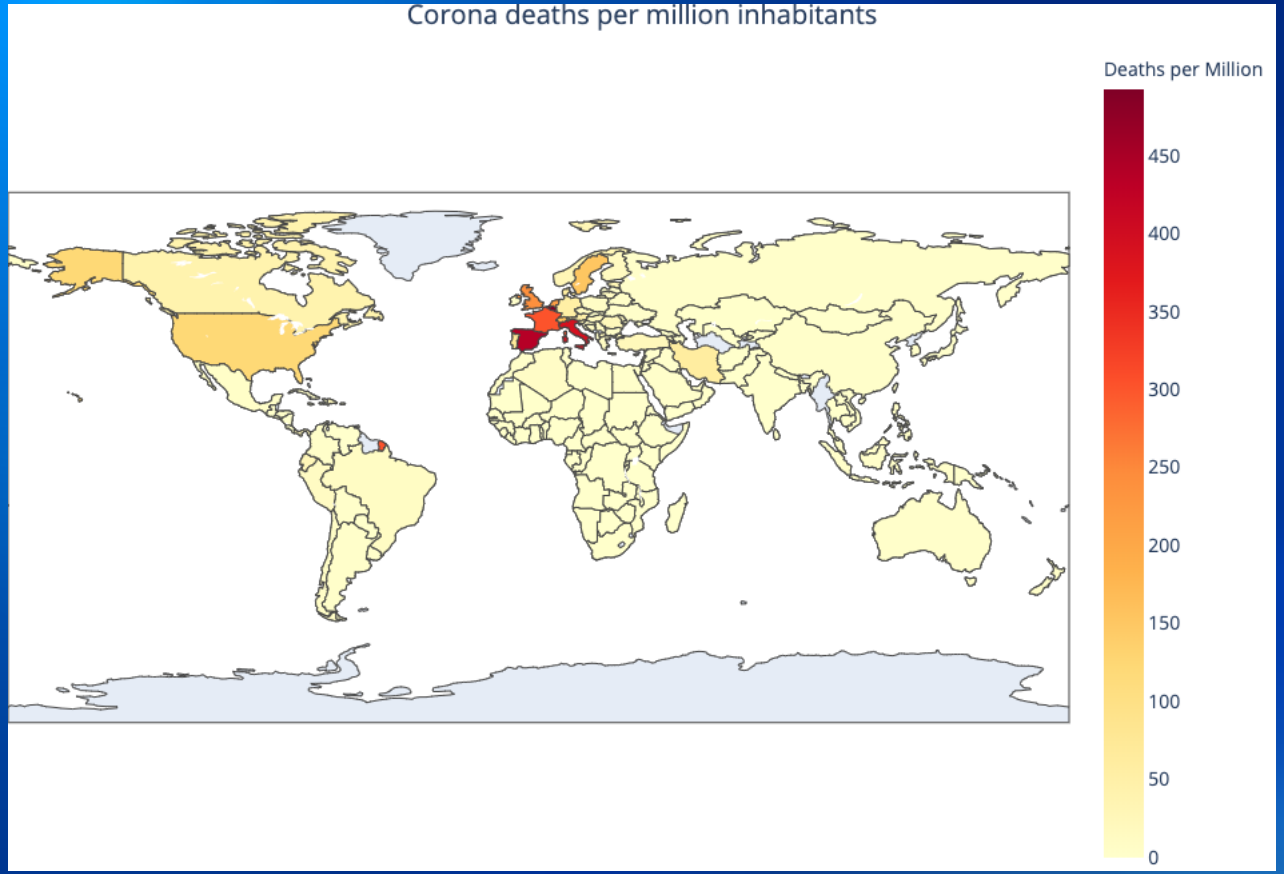
Reported Deaths Time Series



Recovered Patients Time Series

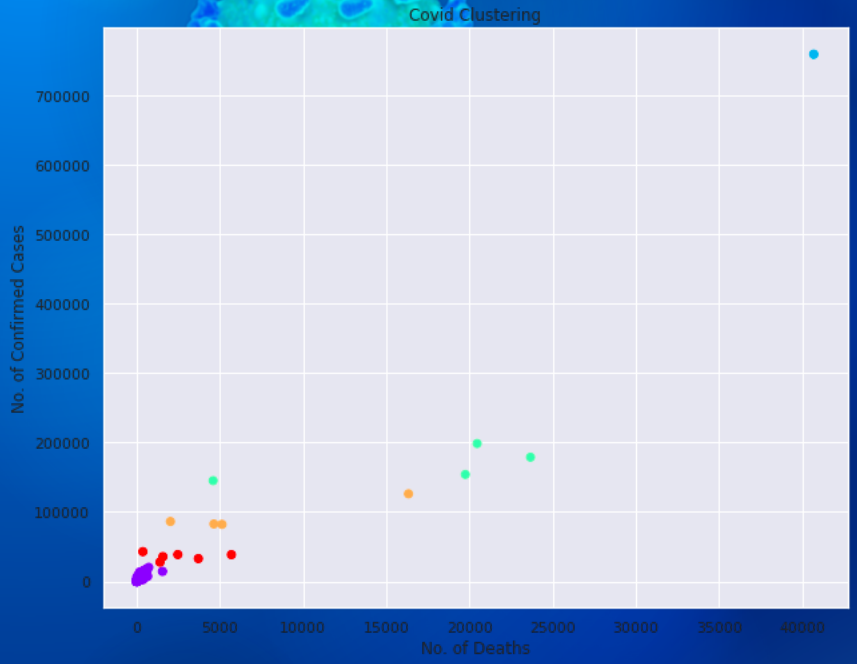



Corona confirmed cases per million inhabitants
Cases per Million
4000
3500
3000
2500
2000
1500
1000
500
0



By the number of deaths, Belgium takes the 1st place, the US - 10th.

CLUSTERIZATION



	Confirmed	Deaths	Recovered
cluster			
1	1842.6	63.1	458.0
2	759774.0	40684.0	70522.0
3	169234.2	17110.8	62398.8
4	94361.0	7023.8	36642.0
5	36041.8	2526.8	10691.8

04.

Results and Discussion

1. I analyzed data connected with the spreading of COVID-19 in the world.
2. Such countries like China, South Korea, Australia are now on a stage of a plateau
3. We saw dynamics and if we recalculate it with new data we will able to make an assumption on how far a country is from a stage of a plateau.
4. Despite a big number of cases in the USA the situation in Europe is worse if look at data through a number of population.
5. In the USA how we know from the news and how analysis shows the worse situation is in New York state and New York City inside state.
6. Also, clusterization was made. Countries were grouped on 5 clusters and we can see dependencies inside clusters (for example, there is a cluster for the US, cluster for Europ's countries with a lot number of cases).

What could be improved? Make a clusterization on weighted data by population. Also, I think, it will be great to find some data about the number of tests in each country, because if not enough tests were made the situation isn't real. Additional information could be a number of hospitals, doctors, some rating of health on the nation.

THANKS!

And Stay At Home!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik