**Introduction**

Year 2020 was undoubtedly dominated by the pandemic caused by the new SARS-CoV-2 coronavirus. The pandemic has created many challenges for governments, enterprises and society globally. All countries in the world have been affected by the illness. Some of them have been hit harder than the others. As a healthcare worker who works in the hospital, I had great interest in all information about development of this disease, and I was especially concerned about situation in Europe.

learning of different approaches taken by countries in order to suppress the virus.

**Problem description**

In every country in the world governments had to make difficult decision how to save public health from COVID illness. Were they successful? Could school closures be avoided?

Did high numbers of hospital beds and healthcare workers helped in keeping low number of patient’s deaths? Is there any correlation between countries with low GDP and higher number of COVID cases? Were more population densified countries in higher risk of COVID?

There are a lot of questions and I would like to find answers for some of them.

**Project Scope:**

The scope of “Covid cases in Europe” Project is to identify patterns, common elements among European Union countries with high and low number of Covid new cases and new deaths. This project will target health system elements, selected demographic, economical characteristics and internal countries policies.

In project scope:

* Project will target only current European Union countries.
* Project will target period of time between 2020-01-03 and 2021-05-03.
* Project will target only the following elements of demographics characteristics**:**
* GDP, population density, median age.
* Project will target only the following elements of healthcare system characteristics**:**
  + hospital beds, nurses and midwives, physicians.
* Project will target only two specified internal restrictions:
  + School closures, internal movement restrictions.

**Dataset sources:**

* **COVID cases and deaths dataset:**

Dataset have been provided by World Health Organization website (https…) in .csv file type. Original dataset contains number of new cases, cumulative cases, and new deaths and cumulative deaths. Dataset has a minimum amount of missing data and contains worldwide daily data. In my project I use number of new cases and deaths.

* **Economical and demographic datasets**:
* Economical dataset of “GDP” , “population” and “population density” have been sourced from the website <https://data.worldbank.org/indicator/> and imported as .csv files.
* Demographic data of “Median Age” has been scrapped from the website to show ability to find different HTML components, to use Beautifull soup module and to import elements into .csv file for data analysis.
* **Healthcare dataset**:

Healthcare dataset of “Hospital beds”, “Nurses and Midwives” and “Physicians” have been sourced from the website <https://data.worldbank.org/indicator/> and imported as .csv files.

* **Internal policies dataset**:

Healthcare dataset of “Schools closures” and “internal movement restrictions” have been sourced from the website <https://ourworldindata.org/policy-responses-covid> and imported as .csv files.

**Datasets overview:**

* **COVID cases and deaths dataset:**

Issues: Negative values ---

Findings:

**Economical dataset** :

|  |  |  |  |
| --- | --- | --- | --- |
| **count** | **Population** | **GDP per capita (USD)** | **Density per m sq.** |
| **mean** | 16,574,520 | 35,299 | 182 |
| **std** | 22,282,340 | 23,217 | 288 |
| **min** | 502,653 | 9,828 | 18 |
| **25%** | 3,427,172 | 19,424 | 71 |
| **50%** | 8,877,067 | 27,858 | 107 |
| **75%** | 14,408,450 | 47,614 | 141 |
| **max** | 83,132,800 | 114,705 | 1,514 |

Top 8 richest country in Europe produces higher GDP per capita than other 19 countires.

**Healthcare dataset:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Physicians\* | Nurses\* | hospital\_beds\* |
| mean | 3.586241 | 8.461981 | 4.998148 |
| std | 0.842889 | 3.247767 | 1.711239 |
| min | 1.9509 | 3.5534 | 2.21 |
| 25% | 3.0343 | 5.9515 | 3.335 |
| 50% | 3.4664 | 7.9665 | 4.69 |
| 75% | 4.01465 | 10.90885 | 6.59 |
| max | 5.4036 | 15.5735 | 8 |

\* per 1000 population

Top 3 counties with highest number of physicians, nurses and hospital beds per 1000 populatio

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TOP 3 Highest** | | | **TOP 3 Lowest** | | |
| **Physicians** | **Nurses** | **Hospital beds** | **Physicians** | **Nurses** | **Hospital beds** |
| 1. Greece | 1. Ireland | 1. Germany | 1. Cyprus | 1. Greece | 1. Sweden |
| 2. Austria | 2. Finland | 2. Bulgaria | 2. Romania | 2. Latvia | 2.Denmark |
| 3. Portugal | 3. Germany | 3. Austria | 3. Poland | 3. Bulgaria | 3. Ireland |
| 4. Lithuania | 4. Luxemburg | 4. Hungary | 4. Malta | 4. Cyprus | 2.Spain |
| 5. Germany | 3. Belgium | 5. Romania | 5. Luxemburg | 5. Spain | 3. Italy |

**Outliers Number of physicians to GDP per Capita:**

I have noticed that countries with low GDP per Capita like Greece , Lithuania , Portugal have proportionally higher number of physicians comparing to Romania, Poland , Bulgaria seems to be one of the poorest countries with higher number of

1.The wealthier country ( high GDP per Capita) could stayed longer in full lockdown in order to supress the virus?

2.Ratio (counties with strictest lockdown) had less number of cases .

3. Counties Most liberal and stricked – consequences ?

**Country restrictions dataset:**

School closures restrictions have been divided into 4 levels, where 0 –“No measures”, 1-“ Recommended”, 2- “Required only at some levels” and 3- “Required”.

Italy (261 days), Germany (205 days), Romania (201) were among countries with highest number of closed schools.

In this part of my project I would like find out how all European union counties have managed Covid in analysed period of time by comparing total number of Covid cases and deaths per 10,000 population with countries own characteristics.

**Total Covid cases vs GDP.**

Fist of all I wanted to check if there is correlation between number of cases in each country and theirs GDP per capita. In this investigation I haven’t noticed any obvious visual correlations. The only thing that is noticeable is 4 out of 5 the richest counties (highest GDP per capita, please see Economical Dataset analysis) has the total lowest number of COVID cases. After making more research I have come across on the article that confirm this observation .Countries with much higher GDP per capita are the ones that have the highest share of employees working from home in Europe.

<https://cdn.statcdn.com/Infographic/images/normal/20743.jpeg>

**Total Covid deaths vs Median Age.**

In this pair of features you can notice not definite but noticeable trend. Number of Covid deaths is slightly falling as falling the median age of population in the country. The higher median age the country could expect higher number of Covid cases.

looks like GDP didn’t have any influence on severity of Covis.