HACKATHON

COVID-19 DATA ANALYSIS FOR EDUCATION PURP0SES

GROUP MEMBERS: -

1. YASHARTH DUBEY (19BCS114)

2.VIKASH SINGH (19BCS113)

3. SYED AFEEF HASAN KAZMI (19BCS106)

4.ASHWIN MISHRA (19BEC006)

PROBLEMS SELECTED:

1. Identify the countries as HIGH-RISK TRAVEL destination countries for Internship or Project work for the next two years.

Draw a trend graph of death growing rate for that country. The chart may

change as the data set is dynamic. Collect the death increasing rate (fatality

or mortality).

1. Find:
2. The average number of cases confirmed week wise.
3. The average number of instances cured week wise.
4. The average number of death cases reported week wise.

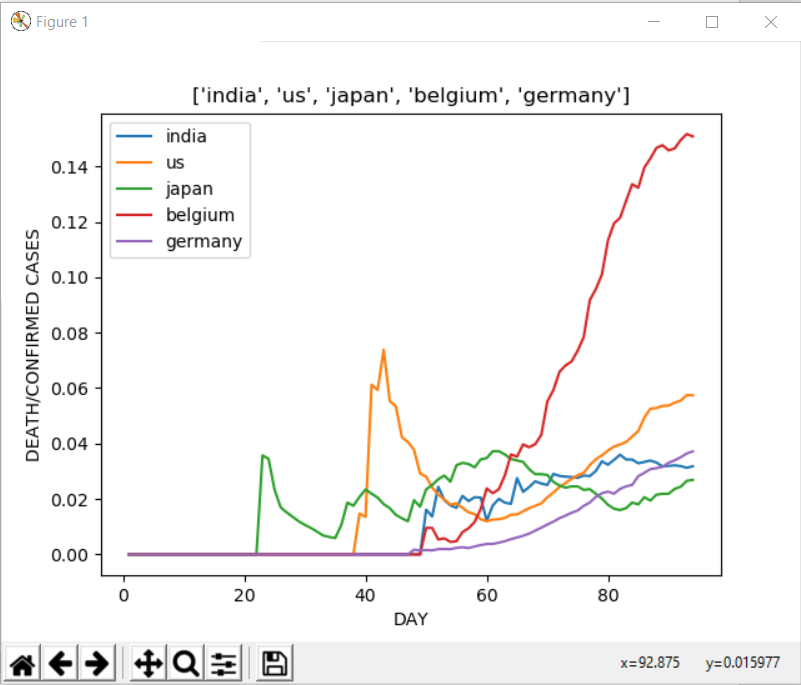
3. To estimate the chances of infection according to the answered set of questions.

PURPOSE OF RESEARCH: -

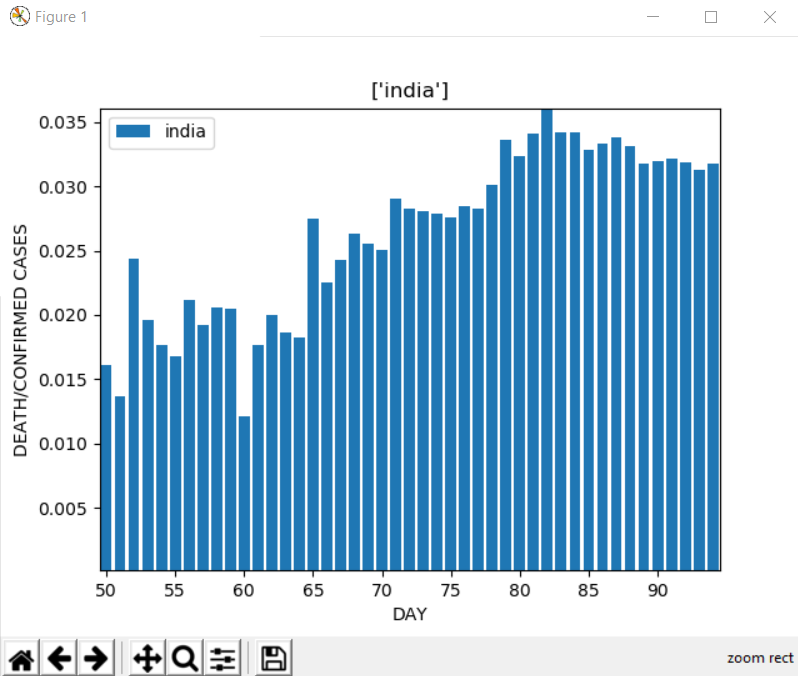
1. To find out those countries which are at high risk for a travel destination for students for internships. So, they can avoid travelling to those countries for the next two years.
2. To find an average number of cases confirmed every week. To check how active virus spreading is in the world.
3. To find out the Average number of cases cured every week, so that we can know how to control its spreading.
4. To find out the average number of deaths cases reported every week.

**IMPORTANT OBSERVATION FOUND IN DATA STUDY: -**

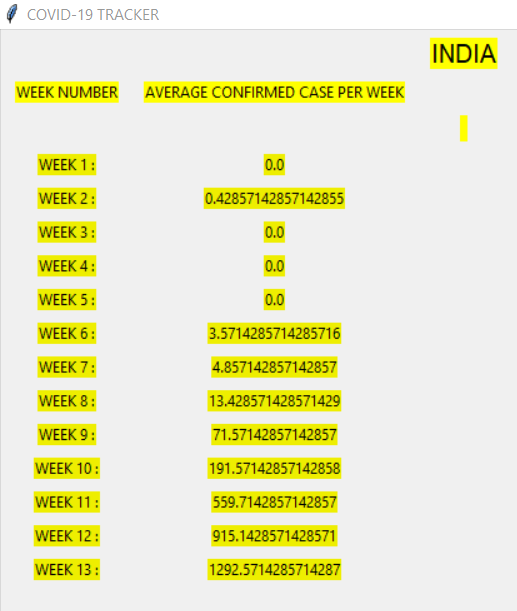
1. After looking at data for counting confirmed, death cases country-wise. We observed that Nowadays U.S.A. is at a very high risk having more than seven lakhs of active instances. And having more than 50 thousand of deaths. And it has become an unsafe country
2. So, our second observation was about the day-wise data for deaths per confirmed cases. As you can see the graph below which is



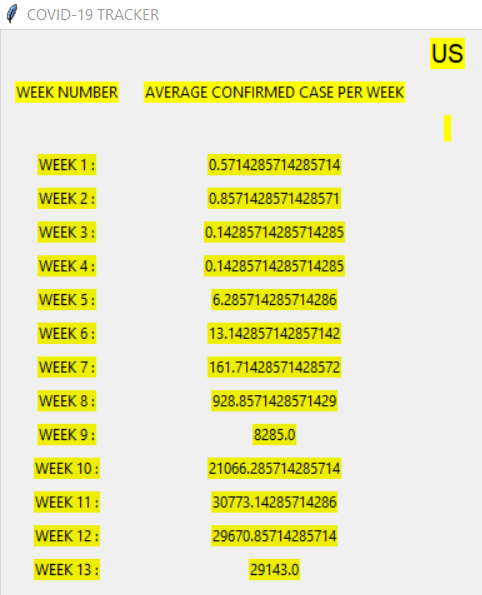
We can compare the data from different countries with these features. Its death/uphold rate per day and what we can see is the graph for all the countries are showing an upward trend except three countries which are South Korea, Japan and China. Now let us talk about our country, which is India, the graph of India is below to study the data more efficiently.



So, for a detailed study, we have plotted the bar graph what we see here is that there is very little growth. **The Nazimuddin jamaat incident on 27/03/2020. Five days tremendous growth in confirmed cases were recorded. Seven days prior, it started going upward as casualties increased**. As we can see, the people did not obey the lockdown properly, which increased in confirmed cases, we can see this trend after seeing the week wise data below.

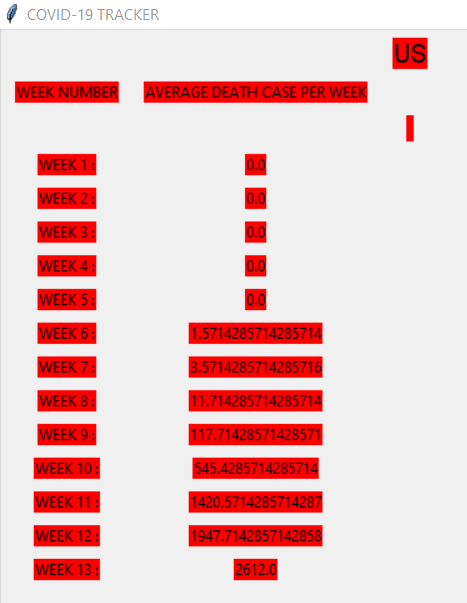


The data is showing us the trend of confirmed cases in the country**. We can see sharp growth in the data in week nine.** Now we should see the confirmed cases of the worst-hit country, which is the U.S.



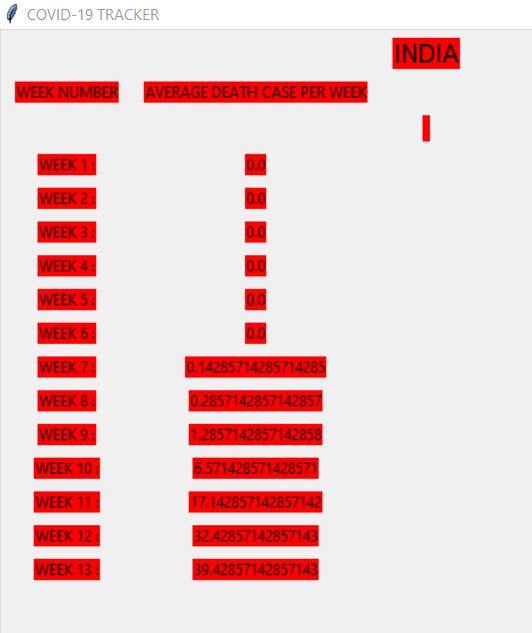
What we can see is the week wise confirmed cases in us were not showing a regular trend but rather a high low pattern.

Now we should talk about the week wise deaths in countries so as we can see that here also the worst-hit country in the U.S. Let us know the data for it.



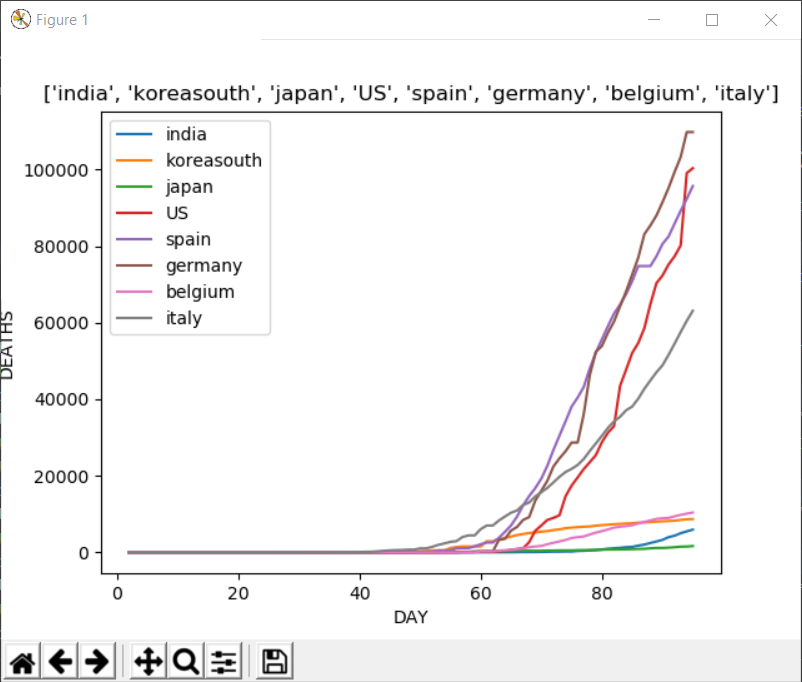
What we see in the above data is that the deaths are more than any other country. The rate by which it is going up is a very alarming one. The critical **observation which we can take out is that in us, the patients in the age group of 45-85 are highest, which is causing the death to be more. One more observation we get is the immunity of the people in the west is much lower than in the east.** That is why it is said not to eat packed food.

Now we are talking about the data of India we should take a look below



What we can see here is **the data trends show us that death in India is pretty low and not only India, but the same is for Japan and South Korea. (\*\*we do not include china as the data provided by Chinese media are questionable\*\*). What we can conclude from this week wise death data is the death depends on the age group. Immunity of the person and the Asian continent has registered a low number of casualties because of the high immunity power and having a large number of cases in the age between 15-25. Some countries in Europe are also handling the problem very well, namely France and Germany.**

**We are now going to talk about the graphical data of Recovered cases in India, Japan, South Korea, the US, Germany, Belgium, Italy and Spain.**



We can see here that the recovered patients are growing everywhere. However, the rate with which countries are improving is more in European countries.

Below is the calculated recovery rate\* of some countries:

U.S.: 2.17 recoveries per death

SPAIN: 4.71

ITALY: 2.39

GERMANY: 18.68

SOUTH KOREA: 36.020

JAPAN: 4.6\*

INDIA: 7.19\*

BELGIUM: 1.54

The countries with a higher number of active cases are hard to predict. Still, the low recovery rate of countries like Spain, Belgium, Italy, and the U.S. is showing us. That we need to invest more in developing our health care than to prepare weapons for war. **NOW WE SHOULD UNDERSTAND THAT WE CANT KILL A VIRUS BY NUCLEAR WEAPON.**

If we see the continent wise break up, then the rank would be:

1. NORTH AMERICA

2. EUROPE

3. ASIA

4. SOUTH AMERICA

5. AUSTRALIA AND OCEANIA

6. AFRICA

The trend which will amaze us is that the epicentre of the pandemic was Asia and still it is in 3rd place. As it is an airborne disease. It should spread mainly in developing countries, but Africa is still in the last this shows us that the continent which is densely populated are more affected by it.

So, these were the observation we found out from the data study.

**Data files and references: -**

1. recovered time series: <https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_recovered_global.csv>
2. death time series: <https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv>
3. confirmed time data: <https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv>
4. Count: <https://www.worldometers.info/coronavirus>

**Data types or data structures used: -**

1. List.
2. Strings
3. Excel file

We have also used bs4 for web scrapping, Tkinter for graphics and also matplot library for graphs, xlrd, openplyx, xlwt for working in excel sheets.

**RESULT DESCRIPTION: -**

The following countries are at very high risk for travel:

1. U.S.A
2. Spain
3. Italy
4. France
5. Germany
6. U.K
7. Turkey
8. Iran
9. Brazil
10. Belgium

So, these are the top 10 most risky travel destination countries for internships or project work for the next two years.

So, students are advised to avoid travelling to these countries.