Impact of Covid-19 on various aspects of society

Author Name
Email Id

***Abstract* — In this paper we have studied the impact of COVID-19 on various aspects of society such as comparison of mortality and cases in various countries. We have also checked the working hours lost during the and after the COVID-19 period and the impact of covid on stock prices of technical companies which will reflect how the performance of these companies were during the Covid phase.**

Keywords : Covid19, Working Hours Lost, Employment Problems, Stock Prices

I. INTRODUCTION

For people all over the world Covid-19 came as a shocking disease which not only took the lives of many people but also created many long term problems such as loss of jobs, closing of markets, sharp reduction in stock prices of various companies.

There have been various studies about covid-19 related to their impact and assessment and both long term and short term impact [1] and various aspects of society such as financial aspects[2]. These papers have effectively studied the impact which Covid made especially during Covid-19 phase of 2019 to 2020. They have highlighted some good statistical results for study. So it has become interesting to study the effect on various walks of life by this type of virus. We will also be studying on similar lines and checking impact of Covid-19 through various open source datasets.

For our study we have studied 4 types of data which are respectively Covid-19 dataset [3], Tweet Dataset for Covid [4], Job Market Dataset [5] and stocks dataset of tech companies which we will be computing using finance open source api.

II. Dataset Description

```
Data columns (total 10 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Date            35156 non-null  object
 1   Country/Region  35156 non-null  object
 2   Confirmed       35156 non-null  int64
 3   Deaths          35156 non-null  int64
 4   Recovered       35156 non-null  int64
 5   Active          35156 non-null  int64
 6   New cases       35156 non-null  int64
 7   New deaths      35156 non-null  int64
 8   New recovered   35156 non-null  int64
 9   WHO Region      35156 non-null  object
```

Fig 1. Covid 19 Dataset

We can see the snippet of Covid-19 dataset where we can see features like Confirmed Cases, Deaths, Recovered ,etc which are linked to different countries on a given date. We have also observed that the 187 unique categories are present for country features and 6 for WHO Region.

For the tweets dataset we can see following data information, where most of the features are object type as they are carrying english words/tweets. This dataset has 12 columns which are related to tweet related information from the users.

```
 #   Column            Non-Null Count   Dtype
---  ------            --------------   -----
 0   user_name         179108 non-null  object
 1   user_location     142337 non-null  object
 2   user_description  168822 non-null  object
 3   user_created      179108 non-null  object
 4   user_followers    179108 non-null  int64
 5   user_friends      179108 non-null  int64
 6   user_favourites   179108 non-null  int64
 7   user_verified     179108 non-null  bool
 8   date              179108 non-null  object
 9   text              179108 non-null  object
 10  hashtags          127774 non-null  object
 11  source            179031 non-null  object
 12  is_retweet        179108 non-null  bool
```

Fig 2. Covid Tweets Dataset

For the job market dataset we have a total 9 columns which represents the amount of working hours lost by people from various

countries due to covid.

```
Data columns (total 9 columns):
 #   Column
---  ------
 0   country
 1   total_weekly_hours_worked(estimates_in_thousands)
 2   percentage_of_working_hrs_lost
 3   percent_hours_lost_40hrs_per_week
 4   percent_hours_lost_48hrs_per_week
 5   labour_dependency_ratio
 6   employed_female_25+_2019
 7   employed_male_25+_2019
 8   ratio_of_weekly_hours_worked_by_population_age_15-64
dtypes: float64(8), object(1)
```

Fig 3. Job Market Dataset

III. Exploring Dataset

1. For Covid-19 Dataset we have created the box plots for the continuous variables to see the variation of various features such as Number of Confirmed, Deaths, Recovered and Active Cases.

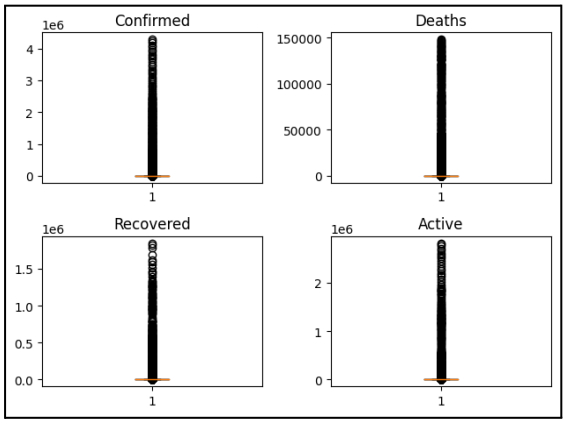

Fig 4: Box plots of Continuous Variables

We also studied to find the outlier by using Interquartile Range Method and got the following results:

```
Percentage of Outliers in  Confirmed  Data is 17.96 %
Percentage of Outliers in  Deaths  Data is 18.43 %
Percentage of Outliers in  Recovered  Data is 17.68 %
Percentage of Outliers in  Active  Data is 17.93 %
Percentage of Outliers in  New cases  Data is 18.06 %
Percentage of Outliers in  New deaths  Data is 20.96 %
Percentage of Outliers in  New recovered  Data is 18.98 %
```

Fig 5: Statistics of outliers in the data

We also grouped the cases by Countries to find the most infected ones and we found the top 10 as below:

Country/Region	
US	224345948
Brazil	89524967
Russia	45408411
India	40883464
Spain	27404045
United Kingdom	26748587
Italy	26745145
France	21210926
Germany	21059152
Iran	19339267

Fig 6: Top 10 County list with covid cases

We also studied the mortality rate by computing the number of death cases divided by there current population and we found following top countries having highest mortality rates:

Country/Region	
San Marino	0.123755
Belgium	0.085030
UK	0.068333
Andorra	0.067290
Peru	0.061860
Spain	0.060954
Italy	0.058206
Sweden	0.057057
Chile	0.051687
USA	0.049156

Fig 7: Top 10 County list with covid mortality rates

2. For analysis related to the tweet dataset, we created top_july_tweets, top_august_tweets and top_all_tweets which are tweets of July and August month respectively. We used them to study the active users and what their country is for each of the months.

The above features helped us to find the most number of tweets per country. We can see that

countries like India and the USA saw the most tweets.

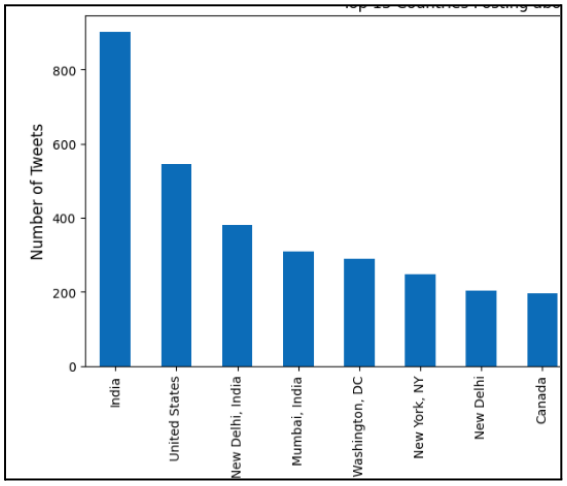

Fig 8: Countries with maximum tweets

We also created the word cloud which has frequency of words as well from the overall aggregated tweets. This helps to check in one sight what is being discussed during these covid periods.

Fig 9: Word Cloud for tweets made during covid time

We also computed the list of users who are creating the most number of tweets during Covid times. The result can be seen in the Figure where the top user is shown in the leftmost part of image and then decreasing consequentially.

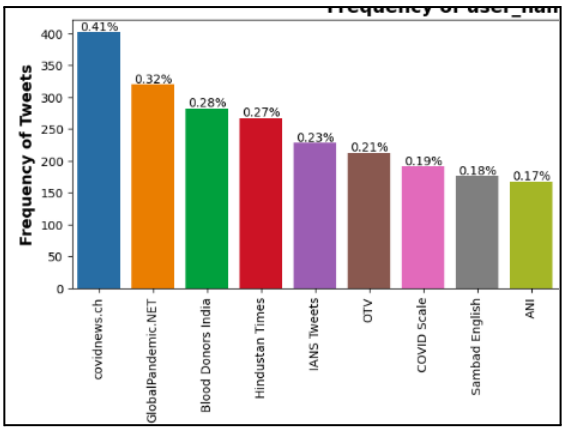

Fig 10: Users with maximum number of tweets

Similarly we also computed the treemap for finding most tweets with maximum tweets.

Fig 11: Treemap for maximum users

We also studied the data and checked the number of tweets in which job related details and health related details were mentioned. We found following result:

	% of Tweets
Job Related	0.16%
Health Related	1.20%

Table 1: Job and Health Data in Total Tweets

Similarly for this dataset we computed the treemap for the channels with maximum number of followers as shown below:

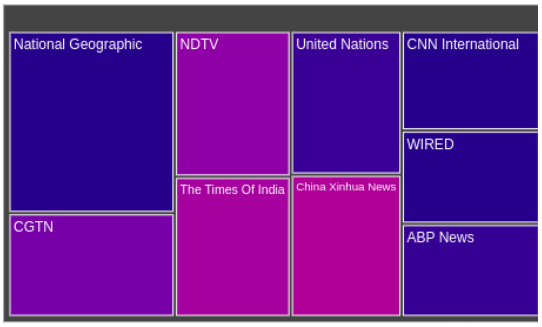

Fig 12: Treemap for channels with maximum followers

3. For the job market analysis dataset we have checked the plots of various variables to understand the flow, which can be seen in the adjoining figure.

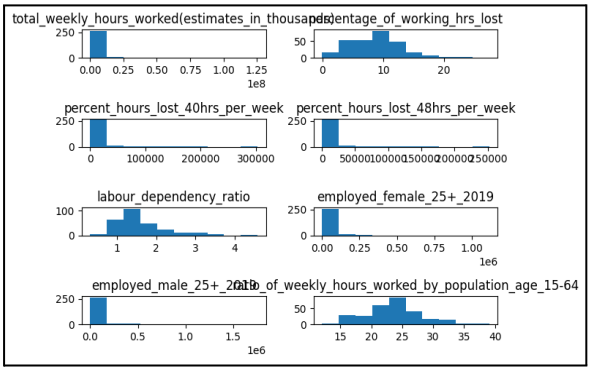

Fig 13: Plots to show maximum variation of continuous variables

We did grouping of data to find out the countries with maximum number of working hours lost and we found below top countries for the same:

country	percentage_of_working_hrs_lost
Peru	27.5
Honduras	24.3
Panama	23.5
Argentina	21.0
Colombia	20.9
Bolivia	20.5
El Salvador	19.4
Latin America and the Caribbean: Lower-middle ...	18.5
Americas: Lower-middle income	18.5
South America	17.7

Fig 14: Countries working hour lost data
We can see that Peru is at top followed by Honduras and Panama.

Similarly we checked the list of countries where maximum hours were worked during this phase and we found below results:

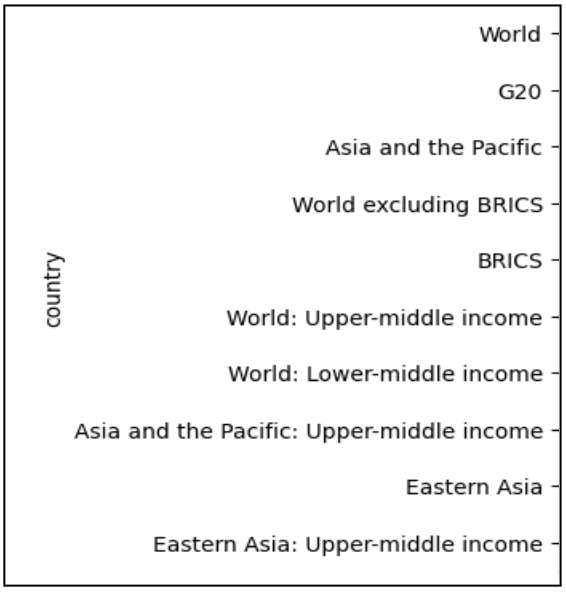

Fig 15: List of countries/groups with maximum number of working hours

We also computed countries with maximum female working hours and found below result:

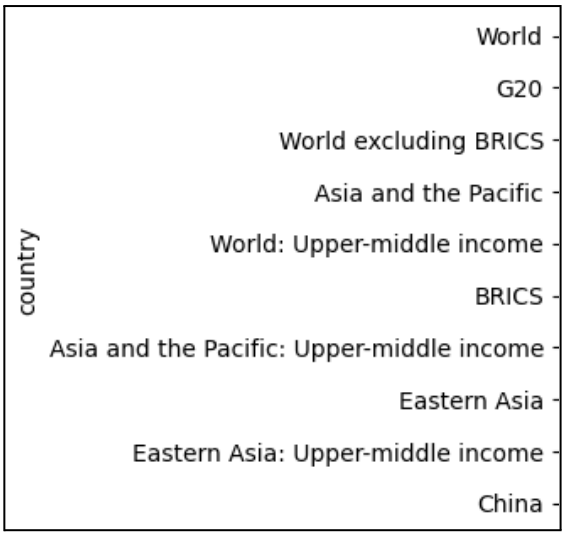

Fig 16: List of countries/groups with maximum number of female workers

So we can see country wise China is in the list which can also be attributed to the contribution of a huge percentage of its working population.
We also computed the study for male working countries as well and found similar results.

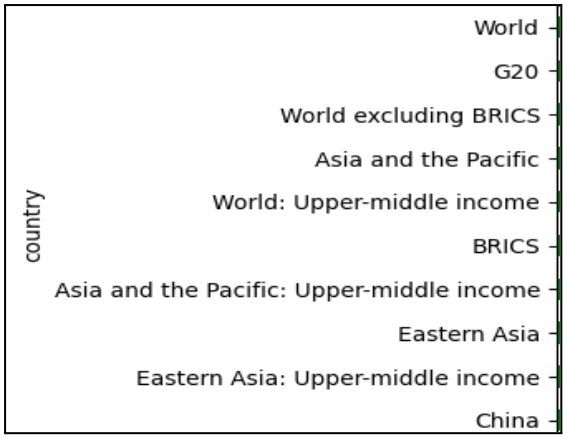

Fig 17: List of countries/groups with maximum number of male workers

4. Stocks Related Datasets : For this dataset we fetched the stock prices of various tech companies and checked whether their prices were growing during and after covid phase. We can see the plot below that the tech stock was rising continuously during and after covid.

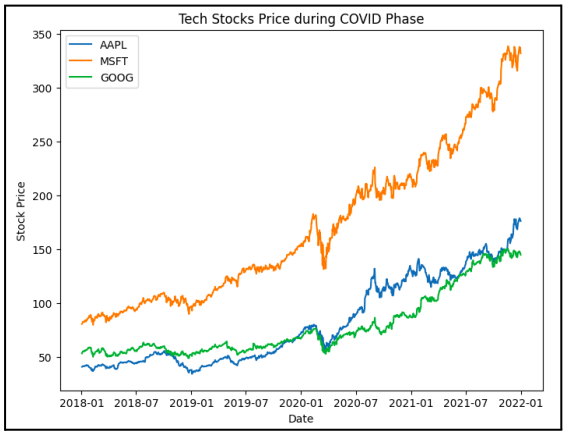

Fig 18 : Plot for tech companies stock price during and after covid period.

IV. Results

We studied and found that the maximum number of covid cases were found in U.S. Brazil and Russia from the given datasets and the maximum mortality were for San Marino, Belgium and UK. From the tweet related dataset we found that the maximum tweets were made from India and U.S. related to covid and some of the maximum tweets were related to Pandemic, Vaccines, Health Supply and other related stuff such as masks. Also we studied that the maximum number of tweets were by users such as covidnews.ch, GlobalPandemic.net.

We also studied and found that maximum job losses data were for countries such as Peru, Honduras and Panama, Argentina and Colombia.

V. Conclusion

We have studied the impact of covid19 on 4 datasets and found their impact on health, job and stock prices. There have been mixed results with some countries getting very high cases while some are effectively able to manage the cases. Similarly we checked the tweets information and made some analysis related to it through health and job status. Similarly we found that the tech stock had grown effectively during this period.

REFERENCES

1. Tan, Xiaoyu, et al. “The Impact of the COVID-19 Pandemic on the Global Dynamic Spillover of Financial Market Risk.” Frontiers, 12 July 2022, https://doi.org/10.3389/fpubh.2022.963620.
2. Zhao, Linhai, et al. “Effects of COVID-19 on Global Financial Markets: Evidence From Qualitative Research for Developed and Developing Economies.” *PubMed Central (PMC)*, 21 Jan. 2022, https://doi.org/10.1057/s41287-021-00494-x.
3. “COVID-19 Dataset.” *COVID-19 Dataset | Kaggle*, /datasets/imdevskp/corona-virus-report.
4. “COVID19 Tweets.” *COVID19 Tweets | Kaggle*, /datasets/gpreda/covid19-tweets.
5. “Impact of Covid-19 on Employment - ILOSTAT.” *Impact of Covid-19 on Employment - ILOSTAT | Kaggle*,

/datasets/vineethakkinapalli/impact-of-covid19-on-employment-ilostat.