Title Cover

Tracking the progression of the new omicron COVID-19 variant

A Project Report

Submitted in the partial fulfillment of the requirements for

the award of the degree of

Bachelor of Technology

in

Department of Computer Science and Engineering

By

C.Koushik Reddy   (2010030027)

D.Chaitanya   (2010030045)

MTV Soumith   (2010030097)

TSS Subramanyam (2010030167)

under the supervision of

Dr. Anal Paul

Department of Computer Science and Engineering

K L University Hyderabad,

Aziz Nagar, Moinabad Road, Hyderabad – 500 075, Telangana, India.

March, 2022

Declaration

The Project Report entitled "Tracking the progression of the new omicron COVID-19 variant" is a record of bonafide work of 2010030027,2010030045,2010030097, 2010030167 , submitted in partial fulfillment for the award of B.Tech in the Department of Computer Science and Engineering to the K L University, Hyderabad. The results embodied in this report have not been copied from any other Departments/University/Institute.

Signature of the Students

Certificate

This is to certify that the Project Report entitled "Tracking the progression of the new omicron COVID-19 variant " is being submitted by Koushik,Chaitanya,Soumith, Subramanyam submitted in partial fulfillment for the award of B.Tech in Computer Science and Engineering to the K L University, Hyderabad is a record of bonafide work carried out under our guidance and supervision.

The results embodied in this report have not been copied from any otherdepartments/ University/Institute.

Signature of theSupervisor

Name andDesignation

Signature oftheHOD **Signature of the ExternalExamine**

ACKNOWLEDGEMENT

It is great pleasure for me to express my gratitude to our honorable President
Sri. Koneru Satyanarayana, for giving the opportunity and platform with
facilities in accomplishing the project based laboratory report.

I express the sincere gratitude to our Principal Dr. L. Koteswara Rao for
his administration towards our academic growth.

I express sincere gratitude to our Coordinator Anal Paul sir
for her leadership and constant motivation provided in successful completion of
our academic semester. I record it as my privilege to deeply thank for providing
us the efficient faculty and facilities to make our ideas into reality.

I express my sincere thanks to our project supervisor <Mr. Anal Paul> for
his/her novel association of ideas, encouragement, appreciation and intellectual
zeal which motivated us to venture this project successfully.

Finally, it is pleased to acknowledge the indebtedness to all those who
devoted themselves directly or indirectly to make this project report success.

Abstract

A WHO report released Sunday said the omicron variant contains a number of genetic mutations, and the organization said the variant has a high likelihood of global spread. But the organization said it's still too early to determine whether omicron is more contagious, more deadly or more resistant to vaccines than previous Covid variants.

The United States, Japan, Israel and several countries have enacted international travel restrictions in response to the variant's spread, though the WHO has cautioned against such bans.

Omicron was first identified in Botswana and South Africa in November—although later reports showed earlier cases in the Netherlands. The first case in the U.S. was reported on Dec. 1 of last year in California. By late December, it became the predominant strain in the U.S. Both the WHO and the Centres for Disease Control and Prevention (CDC) have classified Omicron as a "variant of concern." BA.2 was discovered by scientists in November 2021, and the first case was reported in the U.S. in January.

The Omicron variant caused an alarming spike in COVID-19 cases in South Africa—cases went from 300 a day in mid-November to 3,000 a day at the end of that month. It's also been identified in at least 185 countries, according to the New York Times COVID-19 tracker.

Early reports from South Africa indicated that most cases were mild—and that symptoms for this variant seemed to be different. "The reports show that patients in South Africa—many of whom were young—have had severe fatigue, but no loss of taste or smell," says Lauren Ferrante, MD, a Yale Medicine pulmonologist.

Soon after Omicron's appearance, scientists were concerned about a key distinguishing factor in the variant. Unlike Delta and other coronavirus variants, this one carries an abundance of mutations—about 50 in all, including 26 that are unique to the variant—and more than 30 on the

spike protein, which is the viral protein that vaccines train the immune system to recognize and attack. Grub augh said some of Omicron's enhanced transmissibility could come from its ability to evade some immune responses, especially in people who were previously infected, but not vaccinated. Scientists also wanted to know if these mutations indicated a possible reduction in the effectiveness of the COVID-19 vaccines and certain monoclonal antibody treatments.

Table of contents

Introduction

Omicron is a new variant of SARS-CoV-2 which was recently reported from South Africa on 24th November 2021. It is also known called as B.1.1.529. It has a very large number of mutations, almost more than 30 on the viral spike protein, which is the key target of the immune response . It is detected by Covid-19 RT-PCR METHOD. Omicron may be 2 to 3 times more transmissible than Delta, and other variants.

The variant B.1.1.529 spreads more easily than the original virus that causes COVID-19 and the delta variant. However, omicron appears to cause less severe disease. People who are fully vaccinated can get breakthrough infections and spread the virus to others. But the COVID-19 vaccines are effective at preventing severe illness. This variant also reduces the effectiveness of some monoclonal antibody treatments. Omicron has a few major offshoots (sublineages), including BA.1, BA.2 and BA.3. BA.2 made up about 55% of COVID-19 infections that had genetic sequencing in the U.S. during a week in late March, according to the CDC.

So, Tracking the number of Cases of the new variants is really important as we can know how fast the variant is spreading.
During this pandemic due to some false miscommunication of omicron variant in social media , so many people are getting panic and tensed when they saw such type of news and people are getting information about the omicron variant only about our country in social media or news channels.

 Our goal is to take all the information related to the omicron variant like No of People died , Current status etc.. throughout the world and representing those data in the visualization form for getting awareness about the omicron variant.
As this Project is about tracking the spread of new covid-19 variants , we are going to collect the data from trusted website and display that information in graphical manner. We are going to display the collected data in the form of many graphs.

Literature Survey

SNO	TITLE	AUTHOR	INFORMATION WE GATHERED	Conclusion
1	Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic	Quarraisha Abdool Karim	The first sequenced omicron case was reported from Botswana on Nov 11, 2021, and a few days later another sequenced case was reported from Hong Kong in a traveller from South Africa. Several sequences from South Africa followed, after initial identification that the new variant was associated with an S-gene target failure on a specific PCR	Since the effect is more but the no is deaths is less.
2	Omicron variant and booster COVID-19 vaccines	Talha Khan Burki	70% of the UK population have had two doses of the COVID-19 vaccine. vaccination rates are considerably lower in South Africa. "It certainly looks like a three-shot vaccination schedule will be needed against omicron", commented Susanna Dunachie, professor in infectious diseases at the University of Oxford (Oxford, UK)	Vaccination is needed to overcome such variants.
3	Convulsions in children with COVID-19 during the Omicron wave	Jonas F. Ludvigsson	Two boys aged 3 and 21 months tested positive for the virus and a 14-year-old boy. The teenager had a history of urinary tract infections and a young children had a fever. The 3-month-old child had repeated convulsions for several hours, the 21-month-old had continuous convulsions for 15–20 min, and the teenager had a convulsion for 30–60s.	Convulsions may be a sign of the Omicron variant in children with COVID-19.

Hardware & Software requirements

Component Minimum requirement

Processor :64-bit, four-core, 2.5 GHz minimum per core (If your dataset size is significantly larger than the medium dataset, we recommend 8 cores.)

RAM :4 GB

Hard disk :80 GB

Windows 10 (recommended)

Software:

- We will be using Python Language to develop this project.

- The Libraries we are going to use are NumPy , pandas, seaborn, matplotlib

Functional & Non-functional requirements

Functional Requirements

- In this project we have used different visualization techniques like barplot,scatter plot to easily understand for the users.
- Here we took large dataset which represents the covid variants data not only in india but it represents the data throughout the world.
- For the survey people by seeing the plots they can easily understood how many cases are confirmed and how many deaths throughout the world.
- By the plotting techniques the one who is not educated only understand the data very easily.

Non Functional Requirements

- Performance of the system should be fast and accurate.
- Should be able to handle large amount of data.
- The system hardware and software should meet the minimum requirements in order to execute the project.
- The basic Plotting techniques should know to plot the graphs.

Proposed System

Here in our project different plotting techniques are used. In some projects they use only one plotting technique to represent the data. But here we used multiple plotting techniques to represent the covid data. We used different locations of the dataset to represent the data. And also at last we represented total world covid data in order to make users very easily. If the user want to see particular country data then he just scroll the cursor into the map and if we put a cursor on a country then it will show total confirmed omicron cases and deaths in that country. And also we represented highest affecting country wise list. Like that we used different techniques to extract the data in the dataset and representing those data in the form of graphs.

Implementation

```python
import pandas as pd
import matplotlib.pyplot as plt
data=pd.read_csv(r"C:\Users\DELL\PycharmProjects\mp-2\covid-
cases-omicron.csv")
print(data.shape)
print(data.head(5))
x=data['Entity']
y=data['Omicron_percentage']
plt.xlabel('Entity')
plt.ylabel('Omicron_percentage')
plt.bar(x,y)
plt.show()
```

```python
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt
import seaborn as sns
import pandas_profiling
import plotly.graph_objs as go
import os


        Reading dataset
data = pd.read_csv("/content/sample_data/ covid_19_data.csv")

data.head()
data.info()
data.describe()
data['Region'].unique()

Scatter Plot

data1 = data[data["Date"] == "4/29/2020"]

y = np.array(data1.loc[:,"Confirmed"])
x = np.array(data1.loc[:,"Deaths"])

#scatter

plt.figure(figsize = [20,20])
plt.scatter(x = x,y = y)
plt.xlabel("Deaths")
```

```python
plt.ylabel("Confirmed")
plt.show()
```

```python
sns.countplot(x = "Confirmed",data = data)
data.loc[:,"Confirmed"].value_counts()
```

```python
fig = go.Figure()

Region = list(data['Region'].unique())

for Region in Region:
    fig.add_trace(
        go.Scatter(
            x = data['Confirmed'][data['Region']==Region],
            y = data['Deaths'][data['Region']==Region],
            name = Region,
            mode = 'lines+markers',
            visible = True))

buttons = []

for i, location in enumerate(location):
    args = [False] * len(location)


    button = dict(label = location,
                method = "update",
                args=[{"visible": args}])

    buttons.append(button)
    fig.update_layout(
    updatemenus=[dict(direction="down",
                pad={"r": 10, "t": 10},
                showactive=True,
                x=0.15,
                xanchor="left",
                y=1.12,
                yanchor="top",
                buttons=buttons)])
```

```python
fig.update_layout(
    annotations=[
        dict(text="<b>Choose:<b>",
            showarrow=False,
            x=0,
            y=1.08,
            xref= 'paper',
            yref="paper",
            font=dict(size=15,color = 'DarkSlateBlue'))])
fig.update_layout(title ="<b>Choose Location which you want to understand<b>",
            title_x = 0.5,
            title_font = dict(size = 18, color = 'Darkred'),
            autosize=False,
            width=800,
            height=600)
fig.show()
```

<mark>Displaying highest death Country</mark>

```python
data_total_seq= data[data["Confirmed"] > 0.0].groupby('Region').agg({'Confirmed':'sum'}).reset_index().sort_values("Confirmed", ascending=False)
plt.figure(figsize=(8,12))
plt.ticklabel_format(style = 'plain')
sns.barplot(x= data_total_seq[data_total_seq["Confirmed"] > 10000].Confirmed, y= data_total_seq[data_total_seq["Confirmed"] > 10000].Region)
plt.title("All variants")
plt.show()
print("\nThe highest death
rate country is "+ str(data_total_seq.set_index("Region").index[0]) + f" have {data_total_seq.Confirmed.max()} people.\n")
```

Results Discussion

Figure 1 — □ ×

x=Indonesia y=89.8

It will display the omicron percentage in the particular region in a bar plot

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252
1	4/29/2020	NaN	Albania	766	30	455
2	4/29/2020	NaN	Algeria	3848	444	1702
3	4/29/2020	NaN	Andorra	743	42	423
4	4/29/2020	NaN	Angola	27	2	7

It will display the first 5 of the data in the dataset

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 321 entries, 0 to 320
Data columns (total 6 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Date       321 non-null    object
 1   State      140 non-null    object
 2   Region     321 non-null    object
 3   Confirmed  321 non-null    int64
 4   Deaths     321 non-null    int64
 5   Recovered  321 non-null    int64
dtypes: int64(3), object(3)
memory usage: 15.2+ KB
```

**It will display the non null values
and fill missing values**

```
data.describe()
```

|	Confirmed	Deaths	Recovered
count	321.000000	321.000000	321.000000
mean	9949.800623	709.152648	3030.277259
std	31923.853086	3236.162817	14364.870365
min	0.000000	0.000000	0.000000
25%	104.000000	2.000000	2.000000
50%	653.000000	12.000000	73.000000
75%	4655.000000	144.000000	587.000000
max	299691.000000	27682.000000	132929.000000

**It will display the
count of total cases,recovered,Deaths
and will count mean values,training
percentage and testing percentage.**

```
data['Region'].unique()
```

```
array(['Afghanistan', 'Albania', 'Algeria', 'Andorra', 'Angola',
       'Antigua and Barbuda', 'Argentina', 'Armenia', 'Austria',
       'Azerbaijan', 'Bahamas', 'Bahrain', 'Bangladesh', 'Barbados',
       'Belarus', 'Belgium', 'Belize', 'Benin', 'Bhutan', 'Bolivia',
       'Bosnia and Herzegovina', 'Botswana', 'Brazil', 'Brunei',
       'Bulgaria', 'Burkina Faso', 'Burma', 'Burundi', 'Cabo Verde',
       'Cambodia', 'Cameroon', 'Central African Republic', 'Chad',
       'Chile', 'Colombia', 'Congo (Brazzaville)', 'Congo (Kinshasa)',
       'Costa Rica', 'Croatia', 'Cuba', 'Cyprus', 'Czech Republic',
       'Denmark', 'Diamond Princess', 'Djibouti', 'Dominica',
       'Dominican Republic', 'Ecuador', 'Egypt', 'El Salvador',
       'Equatorial Guinea', 'Eritrea', 'Estonia', 'Eswatini', 'Ethiopia',
       'Fiji', 'Finland', 'France', 'Gabon', 'Gambia', 'Georgia',
       'Germany', 'Ghana', 'Greece', 'Grenada', 'Guatemala', 'Guinea',
       'Guinea-Bissau', 'Guyana', 'Haiti', 'Holy See', 'Honduras',
       'Hungary', 'Iceland', 'India', 'Indonesia', 'Iran', 'Iraq',
       'Ireland', 'Israel', 'Italy', 'Ivory Coast', 'Jamaica', 'Japan',
       'Jordan', 'Kazakhstan', 'Kenya', 'Kosovo', 'Kuwait', 'Kyrgyzstan',
       'Laos', 'Latvia', 'Lebanon', 'Liberia', 'Libya', 'Liechtenstein',
       'Lithuania', 'Luxembourg', 'MS Zaandam', 'Madagascar', 'Malawi',
       'Malaysia', 'Maldives', 'Mali', 'Malta', 'Mauritania', 'Mauritius',
       'Mexico', 'Moldova', 'Monaco', 'Mongolia', 'Montenegro', 'Morocco',
       'Mozambique', 'Namibia', 'Nepal', 'Netherlands', 'New Zealand',
       'Nicaragua', 'Niger', 'Nigeria', 'North Macedonia', 'Norway',
       'Oman', 'Pakistan', 'Panama', 'Papua New Guinea', 'Paraguay',
       'Peru', 'Philippines', 'Poland', 'Portugal', 'Qatar', 'Romania',
       'Russia', 'Rwanda', 'Saint Kitts and Nevis', 'Saint Lucia',
       'Saint Vincent and the Grenadines', 'San Marino',
       'Sao Tome and Principe', 'Saudi Arabia', 'Senegal', 'Serbia',
       'Seychelles', 'Sierra Leone', 'Singapore', 'Slovakia', 'Slovenia',
       'Somalia', 'South Africa', 'South Korea', 'South Sudan', 'Spain',
       'Sri Lanka', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syria',
```
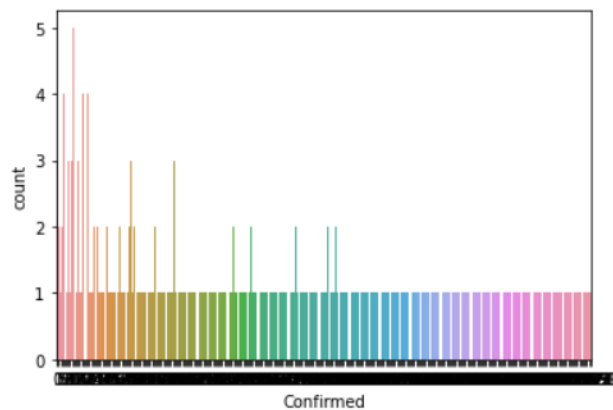
It will display the all the regions in the world.
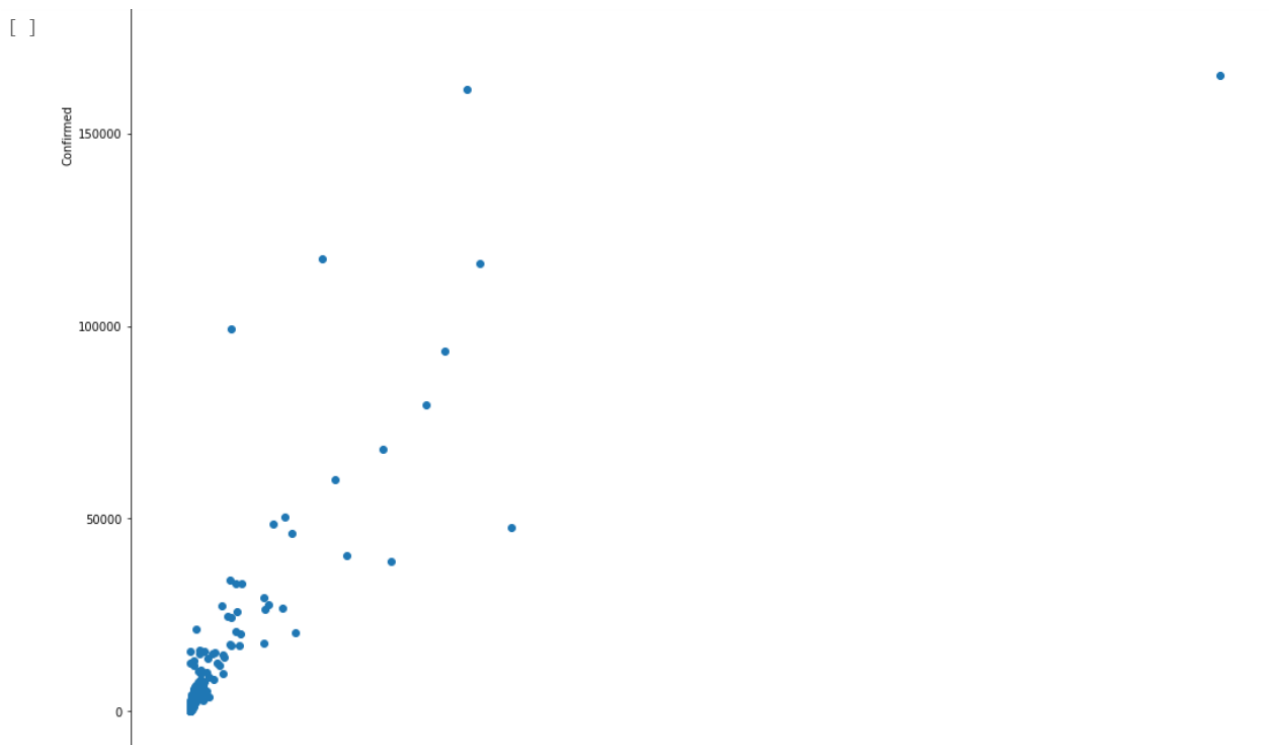
It displays the total data in scatter plot.
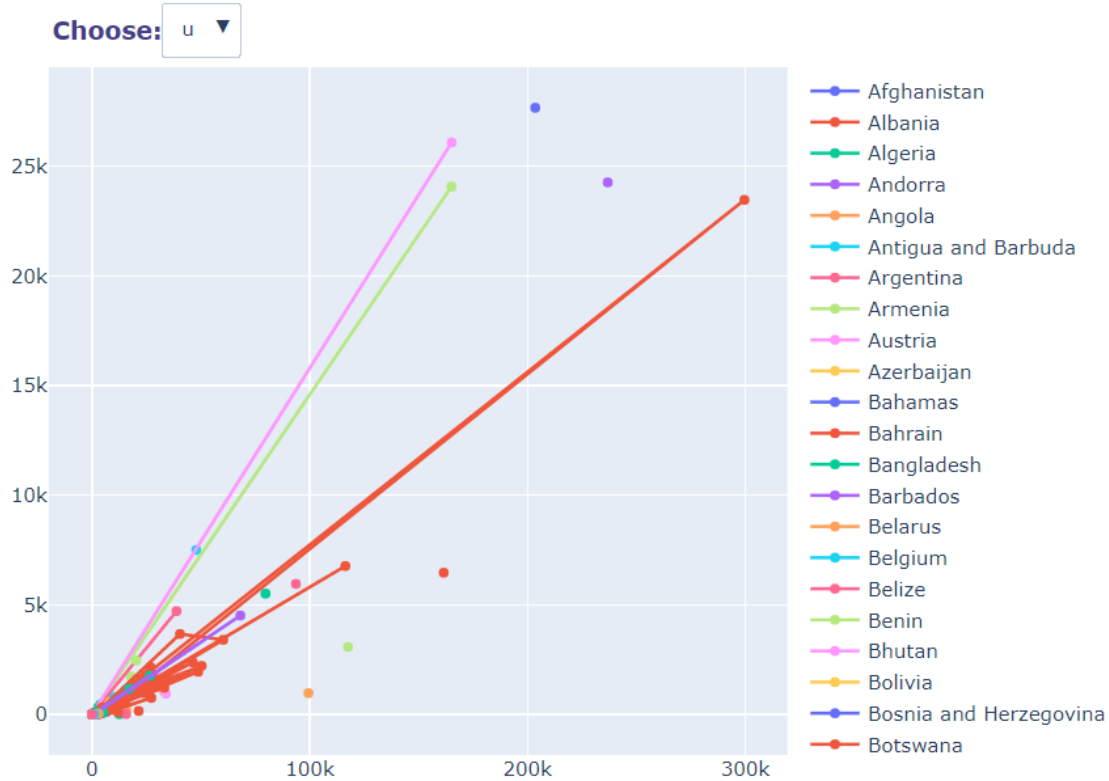
```
[ ]  sns.countplot(x = "Confirmed",data = data)
     data.loc[:,"Confirmed"].value_counts()
```

```
11       5
6        4
16       4
18       4
10       3
        ..
322      1
4321     1
38802    1
1476     1
1268     1
Name: Confirmed, Length: 282, dtype: int64
```


It displays all the confirmed cases in the region.

Choose Location which you want to understand

Choose: [u ▼]

It displays all the location here if we choose india then total cases,recovered,deaths,confirmed are showmin the graph.

```
data_total_seq= data[data["Confirmed"] > 0.0].groupby('Region').agg({'Confirmed':'sum'}).reset_index().sort_values("Confirmed", ascending=False
plt.figure(figsize=(8,12))
plt.ticklabel_format(style = 'plain')
sns.barplot(x= data_total_seq[data_total_seq["Confirmed"] > 10000].Confirmed, y= data_total_seq[data_total_seq["Confirmed"] > 10000].Region)
plt.title("All variants")
plt.show()
print("\nThe heighest country is "+ str(data_total_seq.set_index("Region").index[0]) + f" have {data_total_seq.Confirmed.max()} people.\n")
```

```
The heighest country is US have 1039909 people.
```

It displays the highest deaths country.

Conclusion and Future Work

In this project we have visualized total number of cases,deaths,positive rate and new cases rate through different graphs. So that every one can understand and get to know the number of omicron cases in the country.By visualizing this graph a data analyst can also predict the

Future positivity rate which help common people to take more measures and protect themselves from harmful virus.

In future we will try to visualize the data through more number of graphs and try to update the data in the graph on daily basis.

Specific Contribution

C.Koushik Reddy: plotting graphs code

MTV Soumith: calculating accuracy,min score of dataset

D.Chaitanya: Data visualization code and finding different libraries.

TSS Subramanyam: Acquiring dataset ,labelling,cleaning

References

1)https://www.yalemedicine.org/news/covid-19-variants-of-concern-omicron

2)https://www.who.int/en/activities/tracking-SARS-CoV-2-variants

3)https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)02758-6/fulltext

4)https://www.nejm.org/doi/full/10.1056/NEJMoa2119658