

Comparison of Ebola Virus Disease (EVD) and Covid-19 by data visualization techniques of Machine Learning.

Karanam Santoshachandra Rao , Asst Professor, Dept of CSE
Centurion University of Technology and Management, Odhisa,India

Alugolu Avinash , Asst Professor, Dept of CSE
Centurion University of Technology and Management, Andhra Pradesh, India

Ch. Rani, Regno: 170101120070 , Dept of CSE
Centurion University of Technology and Management, Odhisa, India

Andhavarapu Anusha ,Regno: 170101120051, Dept of CSE
Centurion University of Technology and Management, Odhisa, India

Monalisa Pradhan, Regno: 170101120024, Dept of CSE
Centurion University of Technology and Management, Odhisa, India

Ayushi Mishra,Regno: 170101120034, Dept of CSE
Centurion University of Technology and Management, Odhisa, India

Abstract:

With new cases of the new coronavirus disease, **COVID-19**, growing day by day, it's natural to compare this disease to other outbreaks occurred earlier. In this paper authors have compared and analyzed the recent pandemic Covid-19 with the **Ebola Virus Disease (EVD)** using data visualization techniques in python by taking into consideration of certain constraints as 1) Confirmed cases 2) Confirmed deaths 3) Recovery cases 4) Date of observation 5) Country 6) Suspected cases 7) Probable cases etc. In this paper, the authors have taken the largest Ebola Virus Disease (EVD) outbreak from 29th August 2014 to 23rd March 2016 and recent Covid-19 outbreak data from 22nd January 2020 to 20th March 2020. By the help of this data of Ebola Virus Disease (EVD) and Covid-19 the authors visualized into different graphs in with the help of Data visualization techniques of Machine learning in python programming language.

Keywords: Ebola Virus Disease (EVD), Covid-19, Corona Virus, Pandemic, World Health Organization (WHO), Python, Data Visualization, Dataset, Matplotlib, NumPy, Pandas, Seaborn, Plotly, Packages, Library.

Introduction:

Amongst the major five global public health emergencies in the last six years declared by World Health Organization (WHO)^[1] since 2014 Ebola to recent 2019 Covid-19^[2]. The 2019 coronavirus outbreak (2019-nCoV) is the latest among the five major pandemics that occurred in a decade. So, authors in the paper compared the fatality, contiguousness, number of deaths and number of people recovered of both Ebola Virus Disease (EVD)^[3] and Covid-19.

Covid-19 (2019-nCoV):

World Health Organization (WHO) declared **Covid-19 (2019-nCoV)** as Global health emergency declared on **31 January 2020** and global pandemic^[4] declared on **11 March 2020**. The 2019 novel coronavirus outbreak from the city of Wuhan, Hubei Province, China, in December 2019 spread to more than 160 countries in less than two months. Like Ebola, Covid-19 is a family disease, in the sense that

many infections occur in the home. The WHO officially renamed the disease as Covid-19 on 11 February and the virus that causes the disease as severe acute respiratory syndrome coronavirus^[5] 2 (SARS-CoV-2). The WHO declared the disease as a public health emergency of local concern in China, due to its fast spread to other countries led to the declaration of an emergency of global concern on 31 January. The novel coronavirus outbreak caused more than 14,000 deaths as of 21 March, while more than 330,000 confirmed cases have been recorded.

Transmission:

COVID-19 is a new disease, researchers are still learning about its transmission:

As, of now it has been observed by the scientists that this disease can be transmitted through:

- direct or indirect contact
- droplet spray in short range of transmission
- long-range transmission by aerosol.

COVID-19 affects in different ways. Most of the infected people will develop mild to moderate illness and recover without hospitalization.

Symptoms:

- fever
- dry cough
- tiredness
- aches and pains
- sore throat
- diarrhea
- conjunctivitis
- headache
- loss of taste or smell
- rashes skin, may be discoloration of fingers and toes

Fatal symptoms:

- difficulty breathing or shortness of breath
- chest pain
- loss of movement or speech

2018 Ebola:

Ebola Virus Disease is viral hemorrhagic disease, which can be caused by members of the Flaviviridae, Bunyviridae, and Arenaviridae. They are characterized by fever and bleeding diathesis, followed by circulatory collapse and death. The Flaviviridae family viruses are mainly responsible for yellow fever and dengue fever. Ebola virus consists of genus Ebolavirus with 5 viral species within this genus. The virus family Filoviridae includes three genera: Cueva virus, Marburgvirus, and Ebolavirus. The virus caused outbreak the 2014–2016 West African outbreak belongs to the Zaire ebolavirus species.

The outbreak from 2014–2016 West Africa was the largest Ebola epidemic to occur where nearly 28,610 people were infected and almost 11,308 died. Ebola virus disease (EVD), formerly known as Ebola hemorrhagic fever, is a rare but severe, often fatal illness in humans. The virus is transmitted in the human

population through human-to-human transmission. The average Ebola Virus Disease (EVD) case fatality rate is approximately 50%.

The Ebola virus Disease (EVD) causes serious illness which is often fatal if untreated. Ebola Virus Disease first appeared in 1976 in 2 simultaneous outbreaks, one in what is now Nzara, South Sudan, and the other in Yambuku, DRC. It latter occurred in a village near the Ebola River, from which the disease takes its name as Ebola Virus Disease. The outbreak in West Africa from 2014 to 2016 was the largest Ebola outbreak since the virus was first discovered in 1976. The outbreak first started in Guinea and then moved to Sierra Leone, Liberia and other nearby areas.

Transmission

It is assumed that fruit bats are the main natural Ebola virus hosts. This virus is spread to the humans through close contact with the blood, secretions, or organs of infected animals such as fruit bats, chimpanzees, gorillas and monkeys which are hosts of this virus.

Ebola spreads through human-to-human transmission by direct contact with blood or body fluids of a person who is sick with or has died from Ebola Virus.

Health-care workers were frequently been infected while treating patients with suspected or confirmed Ebola Virus Disease as they will be in close contact with patients. Even Burial ceremonies which involve direct contact with the body of the dead person can also contribute in the transmission of the virus.

Symptoms

The incubation period of the virus to onset of symptoms, is from 2 to 21 days. A person infected with Ebola virus cannot spread the disease until they develop symptoms.

Symptoms of EVD include:

- Fever
- Muscle pain
- Headache
- Sore throat and Vomiting
- Diarrhea
- Rashes on the body
- Impaired kidney and liver function
- In some severe cases it may also cause both internal and external bleeding

In this paper we have tried to compare the contagiousness and fatality rate of both Ebola Virus Disease (EVD) and Covid-19 using data visualization techniques of Machine Learning by the help of python.

Following python code is required to import libraries primarily to analyze the data with the help of various functions and methods.

```
import os
import matplotlib.pyplot as plt
%matplotlib inline
from matplotlib import style
```

```

style.use("fivethirtyeight")
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
import pandas as pd
import numpy as np
from mpl_toolkits.mplot3d import Axes3D
import plotly.express as px

```

Understanding the datasets:

Here, we have used two datasets, one of them is “ebola.csv” contains data from 28th August, 2014 to 23rd March 2018 and another is the “recent_covid_19_data.csv” contains data from 22nd January, 2020 to 20th of March, 2020. The basic idea of dataset can be understood by using the below python code:

Step 1: Reading the data from .csv files by using pandas library:

```

ebola=pd.read_csv("ebola.csv")
ebola

```

The above code `ebola=pd.read_csv("ebola.csv")` reads the data from the ebola.csv dataset using `read_csv()` function from pandas^[6] library.

Fig 1 shows the output of the above python^[7] code by using code `ebola`.

Out[2]:

	Country	Date	No. of suspected cases	No. of probable cases	No. of confirmed cases	No. of confirmed, probable and suspected cases	No. of suspected deaths	No. of probable deaths	No. of confirmed deaths	No. of confirmed, probable and suspected deaths
0	Guinea	2014-08-29	25.0	141.0	482.0	648.0	2.0	141.0	287.0	430.0
1	Nigeria	2014-08-29	3.0	1.0	15.0	19.0	0.0	1.0	6.0	7.0
2	Sierra Leone	2014-08-29	54.0	37.0	935.0	1026.0	8.0	34.0	380.0	422.0
3	Liberia	2014-08-29	382.0	674.0	322.0	1378.0	168.0	301.0	225.0	694.0
4	Sierra Leone	2014-09-05	78.0	37.0	1146.0	1261.0	11.0	37.0	443.0	491.0
...
2480	Liberia	2016-03-23	5636.0	1879.0	3151.0	10666.0	NaN	NaN	NaN	4806.0
2481	Italy	2016-03-23	0.0	0.0	1.0	1.0	NaN	NaN	NaN	0.0
2482	Liberia	2016-03-23	0.0	3.0	2.0	5.0	NaN	3.0	1.0	4.0
2483	Nigeria	2016-03-23	0.0	1.0	19.0	20.0	0.0	1.0	7.0	8.0
2484	United States of America	2016-03-23	0.0	0.0	4.0	4.0	0.0	0.0	1.0	1.0

2485 rows x 10 columns

Fig 1: Displaying whole data from the ebola.csv dataset.

```

covid=pd.read_csv("recent_covid_19_data.csv")
covid

```

The above code read data from `recent_covid_19_data.csv` using `read_csv()` function from pandas library and “`covid`” prints the whole data from that dataset.

Out[4]:

	SNo	ObservationDate	Province/State	Country/Region	Last Update	Confirmed	Deaths	Recovered
0	1	01/22/2020	Anhui	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
1	2	01/22/2020	Beijing	Mainland China	1/22/2020 17:00	14.0	0.0	0.0
2	3	01/22/2020	Chongqing	Mainland China	1/22/2020 17:00	6.0	0.0	0.0
3	4	01/22/2020	Fujian	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
4	5	01/22/2020	Gansu	Mainland China	1/22/2020 17:00	0.0	0.0	0.0
...
7308	7309	03/20/2020	NaN	Jersey	2020-03-17T18:33:03	0.0	0.0	0.0
7309	7310	03/20/2020	NaN	Puerto Rico	2020-03-17T16:13:14	0.0	0.0	0.0
7310	7311	03/20/2020	NaN	Republic of the Congo	2020-03-17T21:33:03	0.0	0.0	0.0
7311	7312	03/20/2020	NaN	The Bahamas	2020-03-19T12:13:38	0.0	0.0	0.0
7312	7313	03/20/2020	NaN	The Gambia	2020-03-18T14:13:56	0.0	0.0	0.0

7313 rows × 8 columns

Fig 2: Displaying whole data from the `recent_covid_19_data.csv` dataset.

Step 3: Extracting the required columns or data from the large dataset^[8] by using predefined functions in python for better visualization^[9] and analyzing.

```
e1=ebola.groupby(['Date'])['No. of confirmed cases'].sum().reset_index()
e1
```

The above python code prints the date and Number of confirmed cases of `ebola.csv` by grouping the date by using `groupby()` function and summing the No. of confirmed confirmed cases according to the date by using `sum().reset_index()` function and stored in `e1` variable. The code `e1` will display the data as given below fig 3:

Out[3]:

	Date	No. of confirmed cases
0	2014-08-29	1754.0
1	2014-09-05	2383.0
2	2014-09-08	2552.0
3	2014-09-12	2639.0
4	2014-09-16	3017.0
...
254	2015-12-17	15249.0
255	2015-12-22	15249.0
256	2015-12-23	15249.0
257	2015-12-29	15249.0
258	2016-03-23	15251.0

259 rows × 2 columns

Fig 3: Grouping and summing of Date and No. of confirmed columns of `ebola` data for better visualization.

```
c1=covid.groupby(['ObservationDate'])['Confirmed'].sum().reset_index()
c1
```

The above python code displays the ObservationDate and No. of confirmed cases from recent_covid_19_data.csv by grouping the ObservationDate using *groupby()* function and by summing the Confirmed cases according to date by using *sum().reset_index()*.

The code *c1* display the data as shown in the below fig 4:

Out[6]:

	ObservationDate	Confirmed
0	01/22/2020	555.0
1	01/23/2020	653.0
2	01/24/2020	941.0
3	01/25/2020	1438.0
4	01/26/2020	2118.0
5	01/27/2020	2927.0
6	01/28/2020	5578.0
7	01/29/2020	6165.0
8	01/30/2020	8235.0
9	01/31/2020	9925.0

Fig 4: Displaying ObservationDate and Confirmed cases of recent_Covid_19_data for better visualization.

```
e2=e1.head(21)
e2
```

The above python code *e2=e1.head(21)* takes first 21 days rows from the *e1* by using *head()* function from python and stores it in *e2* variable. The code *e2* will display the data as shown in the below fig 5.

Out[7]:

	Date	No. of confirmed cases
0	2014-08-29	1754.0
1	2014-09-05	2383.0
2	2014-09-08	2552.0
3	2014-09-12	2639.0
4	2014-09-16	3017.0
5	2014-09-18	3095.0
6	2014-09-22	3341.0
7	2014-09-24	3487.0
8	2014-09-26	3626.0
9	2014-10-01	3973.0
10	2014-10-03	4108.0
11	2014-10-08	4461.0
12	2014-10-10	4655.0
13	2014-10-15	5006.0
14	2014-10-17	4218.0
15	2014-10-22	5501.0
16	2014-10-25	5692.0
17	2014-10-29	7632.0
18	2014-10-31	7728.0
19	2014-11-05	7991.0
20	2014-11-07	8168.0

Fig 5: Displaying first 21 records Date and No. of confirmed cases from ebola dataset.

```
c2=c1.head(21)
```

```
c2
```

The python code `c2=c1.head(21)` displays the first 21 records from `c1` using predefined function `head()` from python and code `c2` displays the output as shown in the below fig 6:

```
Out[8]:
```

	ObservationDate	Confirmed
0	01/22/2020	555.0
1	01/23/2020	653.0
2	01/24/2020	941.0
3	01/25/2020	1438.0
4	01/26/2020	2118.0
5	01/27/2020	2927.0
6	01/28/2020	5578.0
7	01/29/2020	6165.0
8	01/30/2020	8235.0
9	01/31/2020	9925.0
10	02/01/2020	12038.0
11	02/02/2020	16787.0
12	02/03/2020	19881.0
13	02/04/2020	23892.0
14	02/05/2020	27636.0
15	02/06/2020	30818.0
16	02/07/2020	34392.0
17	02/08/2020	37121.0
18	02/09/2020	40151.0
19	02/10/2020	42763.0
20	02/11/2020	44803.0

Fig 6: Displaying first 21 records of ObservationDate and Confirmed cases from recent_covid_19_data dataset.

1) Comparing Confirmation cases of Ebola and Covid-19 in first 3 three weeks:

```
date=['d1','d2','d3','d4','d5','d6','d7','d8','d9','d10',
      'd11','d12','d13','d14','d15','d16','d17','d18','d19','d20',
      'd21']
import matplotlib.pyplot as plt
x=date

plt.figure(figsize=(15,15))
#plt.xticks(date,rotation=90,fontsize=30)
plt.plot(x,e2['No. of confirmed cases'],'go-')
plt.ylabel("Confirmed cases of ebola and covid-19",fontsize=22,fontweight='bold')
plt.xlabel("Days",fontsize=22,fontweight='bold')
plt.title("comparision of ebola and covid confirmed cases in 1st 3 weeks",fontsize=25,fontweight='bold')
plt.plot(x,c2['Confirmed'],'ro-')
plt.legend('ec',fontsize=18)
plt.show()
```

The above python code take date as `d1,d2,...,d21` in X-axis and plot the Confirmed cases of Ebola and Covid-19 in Y-axis.^[10]

By using `plot()` function from Matplotlib library we have visualized the data. The code `plt.show()` function we displayed the graph as shown in the below fig 7:

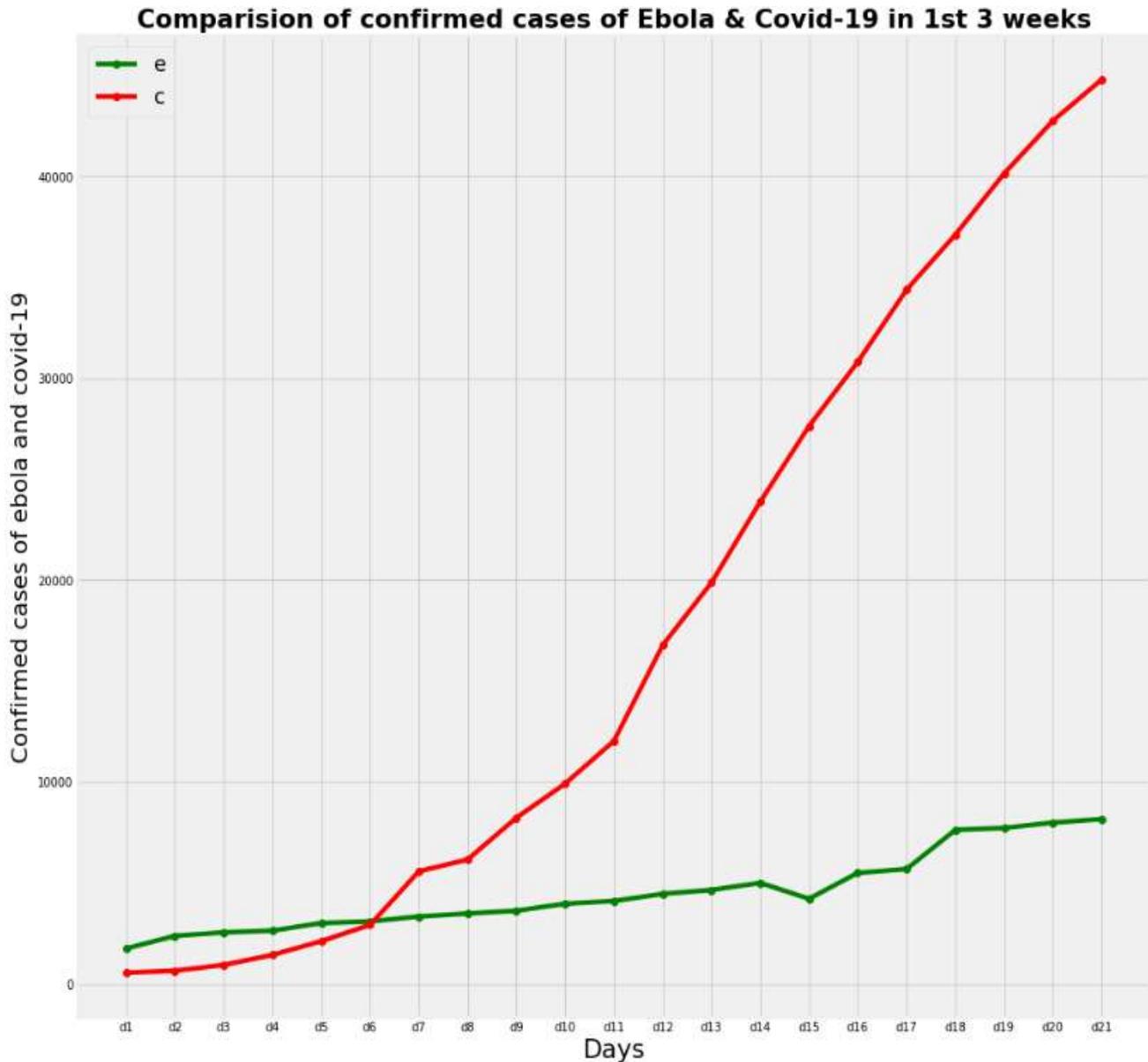


Fig 8: Visual analysis of No. of confirmed cases of ebola and covid-19 in first 3 weeks.

From the above fig 8 it has been observed that in the first 3 weeks the number of confirmed of Covid-19 are higher than that of Ebola Virus Disease (EVD).

It is so because a person infected with Ebola cannot spread the disease until they develop symptoms so when the symptoms are visible then the symptomatic person is isolated and tested for the confirmation of the Ebola Virus Disease (EVD) so the transmission of the disease is low. Whereas, Covid-19 is a lot more contagious than Ebola Virus Disease (EVD). It is so because of its asymptomatic transmission.

2) Comparing the Deaths of Ebola and Covid-19 in first 3 weeks.

Step 1:

```
ed=ebola.groupby(['Date'])['No. of confirmed deaths'].sum().reset_index()
ed
```

The above python code group the date by using `groupby()` function and sum the No. of confirmed deaths according to date by using `sum().reset_index()` function and stores in variable `ed`. The code `ed` displays the output as shown in the below fig 9:

Out[10]:

	Date	No. of confirmed deaths
0	2014-08-29	898.0
1	2014-09-05	1243.0
2	2014-09-08	1376.0
3	2014-09-12	1386.0
4	2014-09-16	1513.0
...
254	2015-12-17	5683.0
255	2015-12-22	5683.0
256	2015-12-23	5683.0
257	2015-12-29	5683.0
258	2016-03-23	5684.0

259 rows × 2 columns

Fig 9: Displaying the Date and No. of confirmed deaths from ebola dataset.

Step 2:

```
cd=covid.groupby(['ObservationDate'])['Deaths'].sum().reset_index()
cd
```

The above python code groups the ObservationDate by using `groupby()` function and sum the Deaths according to ObservationDate by using `sum().reset_index()` function and stores in `cd` variable. The code `cd` displays the output as shown in the below fig 10:

Out[11]:

	ObservationDate	Deaths
0	01/22/2020	17.0
1	01/23/2020	18.0
2	01/24/2020	26.0
3	01/25/2020	42.0
4	01/26/2020	56.0
5	01/27/2020	82.0
6	01/28/2020	131.0
7	01/29/2020	133.0
8	01/30/2020	171.0
9	01/31/2020	213.0
10	02/01/2020	259.0

Fig 10: Displaying records of ObservationDate and Deaths from recent_covid_19_data.

Step 3:

```
ed1=ed.head(21)
ed1
```

The above python code will read the first 21 records from *ed* and stores in *ed1* variable and the code *ed1* will display the data as shown in the below fig 11:

Out[12]:

	Date	No. of confirmed deaths
0	2014-08-29	898.0
1	2014-09-05	1243.0
2	2014-09-08	1376.0
3	2014-09-12	1386.0
4	2014-09-16	1513.0
5	2014-09-18	1590.0
6	2014-09-22	1687.0
7	2014-09-24	1698.0
8	2014-09-26	1837.0
9	2014-10-01	2006.0
10	2014-10-03	2078.0
11	2014-10-08	2337.0
12	2014-10-10	2430.0
13	2014-10-15	1587.0
14	2014-10-17	1611.0
15	2014-10-22	2945.0
16	2014-10-25	2990.0
17	2014-10-29	9.0
18	2014-10-31	9.0
19	2014-11-05	1739.0
20	2014-11-07	1780.0

Fig 11: Displaying the first 21 records of Date and No. of confirmed deaths from ebola dataset.

Step 4:

```
cd1=cd.head(21)
cd1
```

The above python code will read the first 21 records from *cd* and stores in *cd1* variable and the code *cd1* will display the data as shown in the below fig 12:

Out[13]:

	ObservationDate	Deaths
0	01/22/2020	17.0
1	01/23/2020	18.0
2	01/24/2020	26.0
3	01/25/2020	42.0
4	01/26/2020	56.0
5	01/27/2020	82.0
6	01/28/2020	131.0
7	01/29/2020	133.0
8	01/30/2020	171.0
9	01/31/2020	213.0
10	02/01/2020	259.0
11	02/02/2020	362.0
12	02/03/2020	426.0
13	02/04/2020	492.0
14	02/05/2020	564.0
15	02/06/2020	634.0
16	02/07/2020	719.0
17	02/08/2020	806.0
18	02/09/2020	906.0
19	02/10/2020	1013.0
20	02/11/2020	1113.0

Fig 12: Displaying the first 21 records of ObservationDate and Deaths from recent_covid_19_data.
Step 5:

```

date=['d1','d2','d3','d4','d5','d6','d7','d8','d9','d10',
      'd11','d12','d13','d14','d15','d16','d17','d18','d19','d20',
      'd21']
import matplotlib.pyplot as plt
x=date
plt.figure(figsize=(15,10))
rects2=plt.bar(x,ed1['No. of confirmed deaths'],color='magenta')
rects1=plt.bar(x,cd1['Deaths'],color='b')
plt.ylabel("deaths of ebola and covid-29",fontsize=22,fontweight='bold')
plt.xlabel("Days",fontsize=22,fontweight='bold')
plt.title("Three weeks deaths of ebola and covi19",fontsize=25,fontweight='bold')
plt.legend('ec',fontsize=18)
def autolabel(rects):
    for rect in rects:
        height = rect.get_height()
        plt.text(rect.get_x() + rect.get_width()/2., height,
                 '%d' % int(height),
                 ha='center', va='bottom',fontweight='bold',fontsize=15)
autolabel(rects1)
for rect in rects1:
    patterns = ['*']
    rect.set_hatch(patterns.pop(0))
def autolabel(rects3):

```

```

for rect in rects3:
    height = rect.get_height()
    plt.text(rect.get_x() + rect.get_width()/2., height,
            '%d' % int(height),
            ha='center', va='bottom',fontweight='bold',fontsize=10)
autolabel(rects2)
for rect in rects2:
    patterns = ['']
    rect.set_hatch(patterns.pop(0))
plt.savefig('barchart.png')
plt.show()
    
```

The above python code is used the data that we extracted in the above steps and visualized it by using matplotlib library. We have plotted days in X-axis and Deaths in Y-axis and compared the death rate of Ebola Virus Disease (EVD) and Covid-19.^[12]

The code `plt.show()` is used to display the output as shown in the below fig 13:

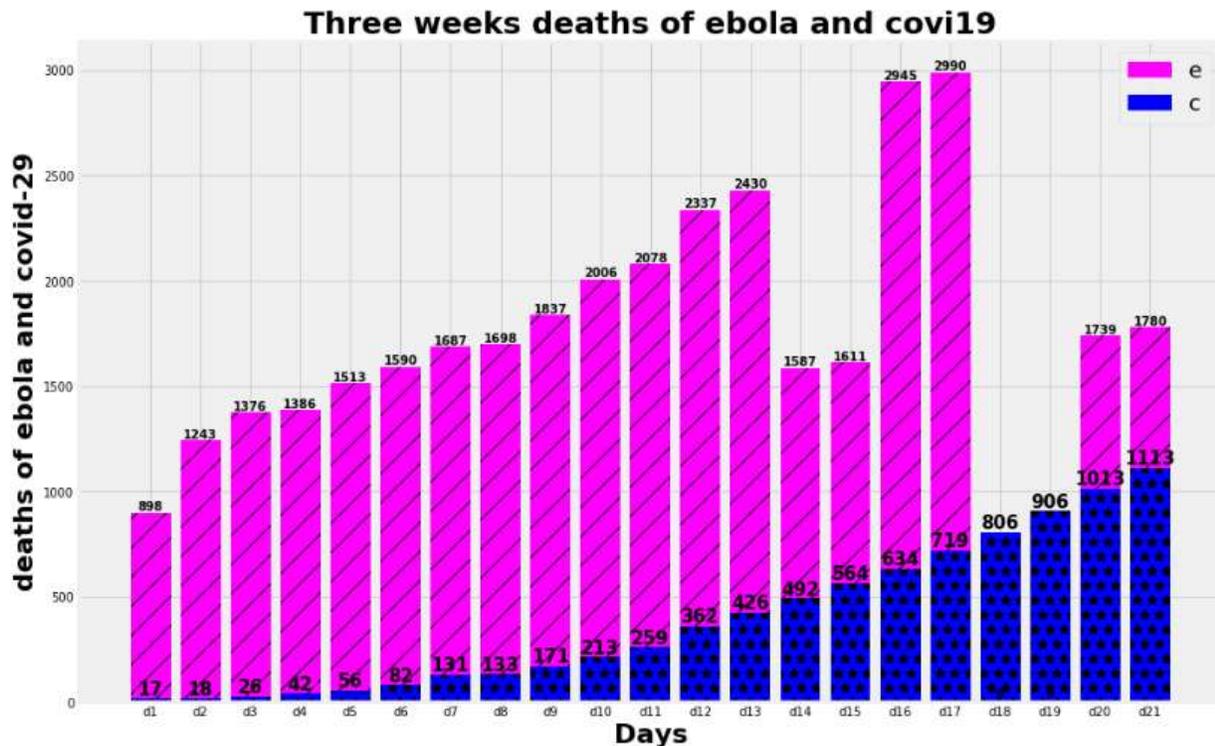


Fig 13: Visual analysis of deaths of ebola and covid-19 in first 3 weeks.

From the above graph it has been observed that the number of deaths of Ebola Virus Disease (EVD) is higher than the deaths of Covid-19 in first 3 weeks.

Though the contagiousness rate of Covid-19 is higher than Ebola Virus Disease (EVD) but the death rate is higher in case of EVD, as EVD kills between 50-80% of people affected. . In fact, if we take into account most of cases have not been detected because of the lack of symptoms or testing, the final death rate may be worse than that of Covid-19. The low lethality of the Covid-19 as compared to Ebola Virus Disease is the reason for its wide pandemic spread. If the virus is very much lethal, it would have died along with its victims, preventing further spread of virus to others.

3) 3D Visualization of all the types of cases of Covid-19 and Ebola Virus Disease.

```
import plotly.express as px
df=covid
fd=ebola
fig=px.line_3d(df,x="Confirmed",y="Deaths",z="Recovered",color='Country/Region')
fig.update_layout(
    showlegend=True,
    height=800,
    width=800,
    title_text="3D model of covid cases")
fig.show()
fig=px.line_3d(fd,x="No. of confirmed cases",y="No. of confirmed deaths",z="No. of probable cases",color='Country')
fig.update_layout(
    showlegend=True,
    height=800,
    width=800,
    title_text="3D model of ebola cases" )
fig.show()
```

The above python code visualizes all the types of cases of Ebola Virus Disease (EVD) and Covid-19 according to the country wise in a 3D plot using plotly library for interactive graphs. The code `fig.show()` is used to display the graph as shown in the below fig 14:

3D model of covid cases

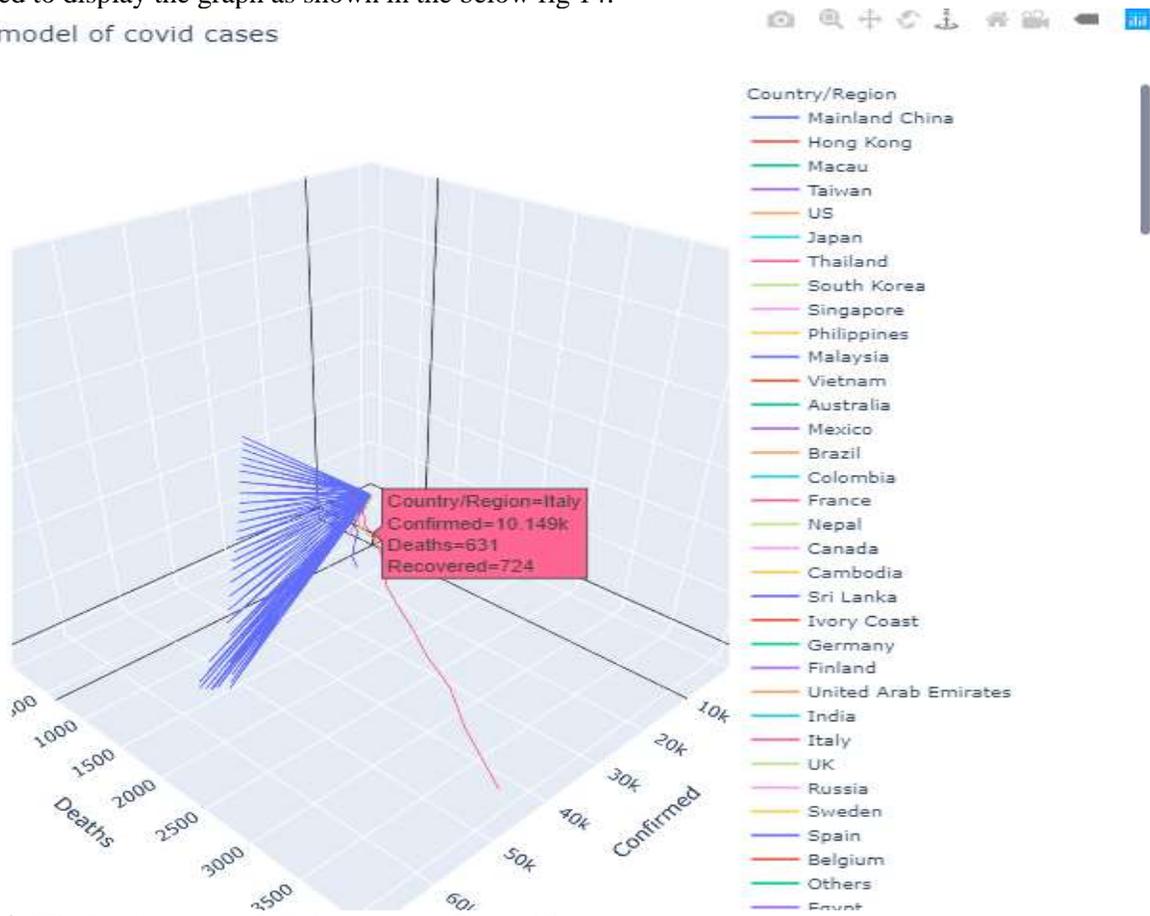


Fig14: 3D Visualization of all the cases of Covid-19 cases.

From the above fig 14 it is observed that from 22nd January, 2020 to 18th March, 2020 the cases in Mainland China are higher as compared to other countries.

3D model of ebola cases

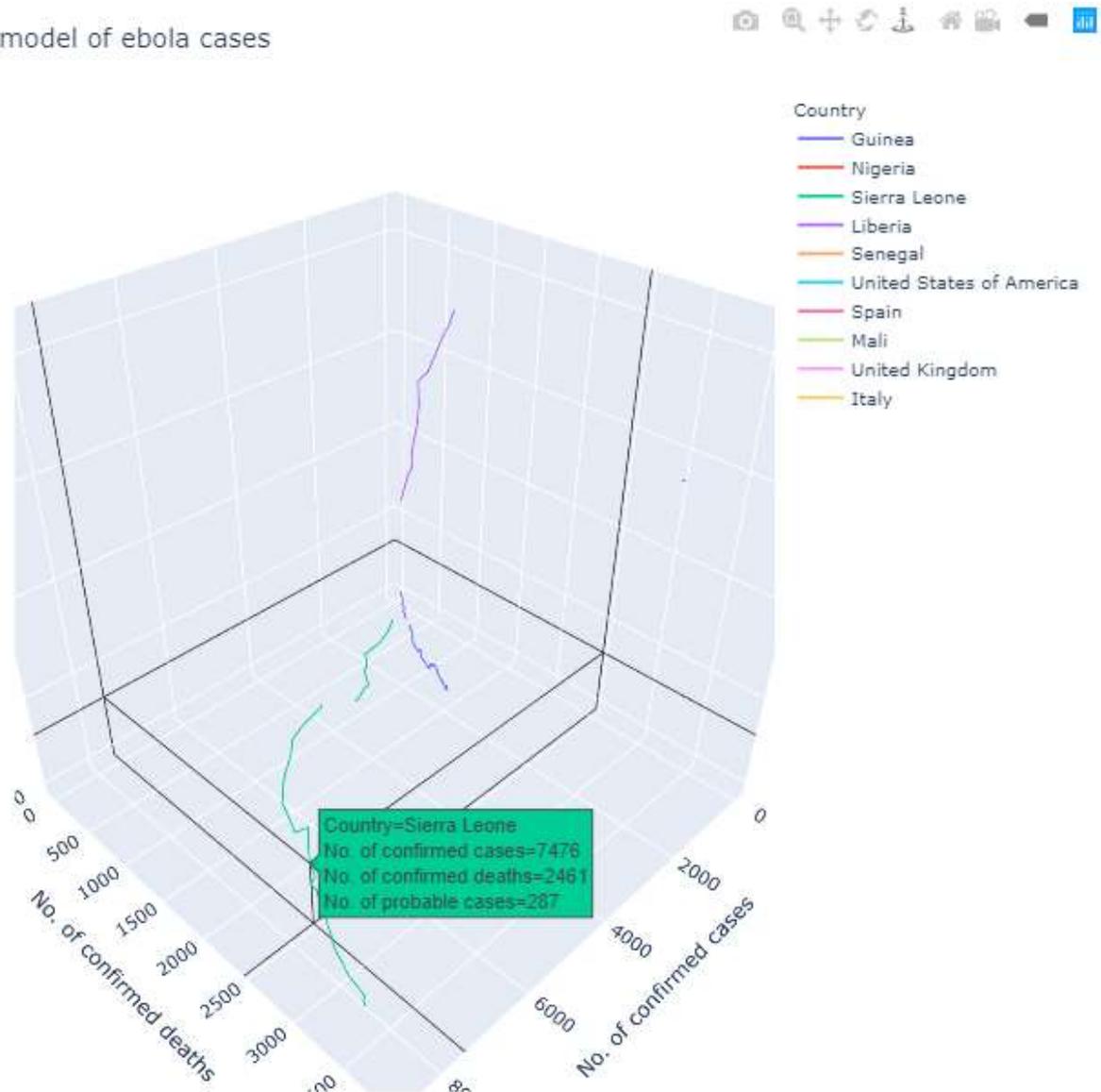


Fig 15: Visual representation of Ebola Virus Disease (EVD) cases.

From the above fig 15 it is observed that from 29th August 2020 to 23rd March 2016 the Number cases in Sierra Leone are higher than other countries.

The above the graphs are interactive graphs i.e. by moving the cursor on the graph we can get the data of the cases of respective countries.^[13]

4) Visualization of Ebola and covid-19 cases on the basis of Confirmed cases using Bubble plot.

Step 1:

```
import plotly.express as px
fig = px.scatter(covid.query("ObservationDate>='01/22/2020'"), x="Confirmed",
y="Deaths",size="Confirmed",
color="Country/Region",
hover_name="Country/Region", log_x=True, size_max=100)
fig.update_layout(
showlegend=True,
height=800,
width=800,
title_text="Covid-19 cases "
)
fig.show()
```

The above python code takes the size as confirmed cases, and color as the Country/Region and confirmed cases and Deaths in X-axis and Y-axis respectively and displays a bubble plot as shown in the below fig 16 by using code `fig.show()`:

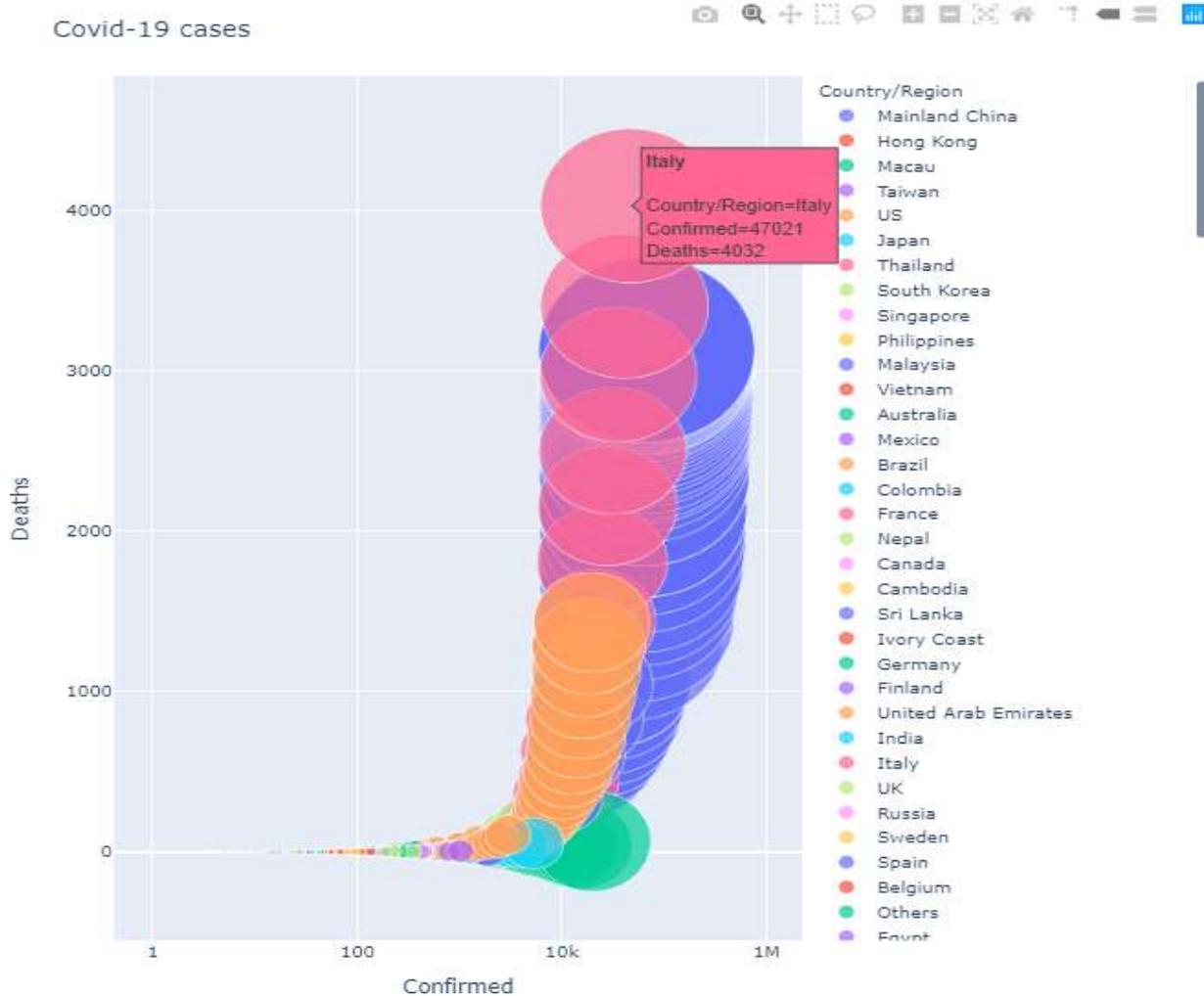


Fig 16: Bubble plot visualization of Covid-19 cases.

From the above bubble plot it is observed that Mainland China and Italy are having the higher number of Covid-19 confirmed cases and the death toll is higher in Italy as per the confirmed cases from 22nd January, 2020 to 18th March, 2020.

Step 2:

import plotly.express as px

```
fig = px.scatter(ebola.query("Date>='2014-08-29'"), x="No. of confirmed cases", y="No. of confirmed deaths",
```

```
size='No. of confirmed, probable and suspected deaths',color="Country",  
hover_name="Country", log_x=True, size_max=150)
```

```
fig.update_layout(
```

```
showlegend=True,
```

```
height=800,
```

```
width=800,
```

```
title_text="Ebola virus cases ")
```

```
fig.show()
```

The above python code takes the size as No. of Confirmed, probable and suspected deaths, and color as the Country and No. of Confirmed cases and No. of Confirmed deaths in X-axis and Y-axis respectively. Displays a bubble plot using plotly for interactive graph as shown in the below fig 17 by using code *fig.show()*:

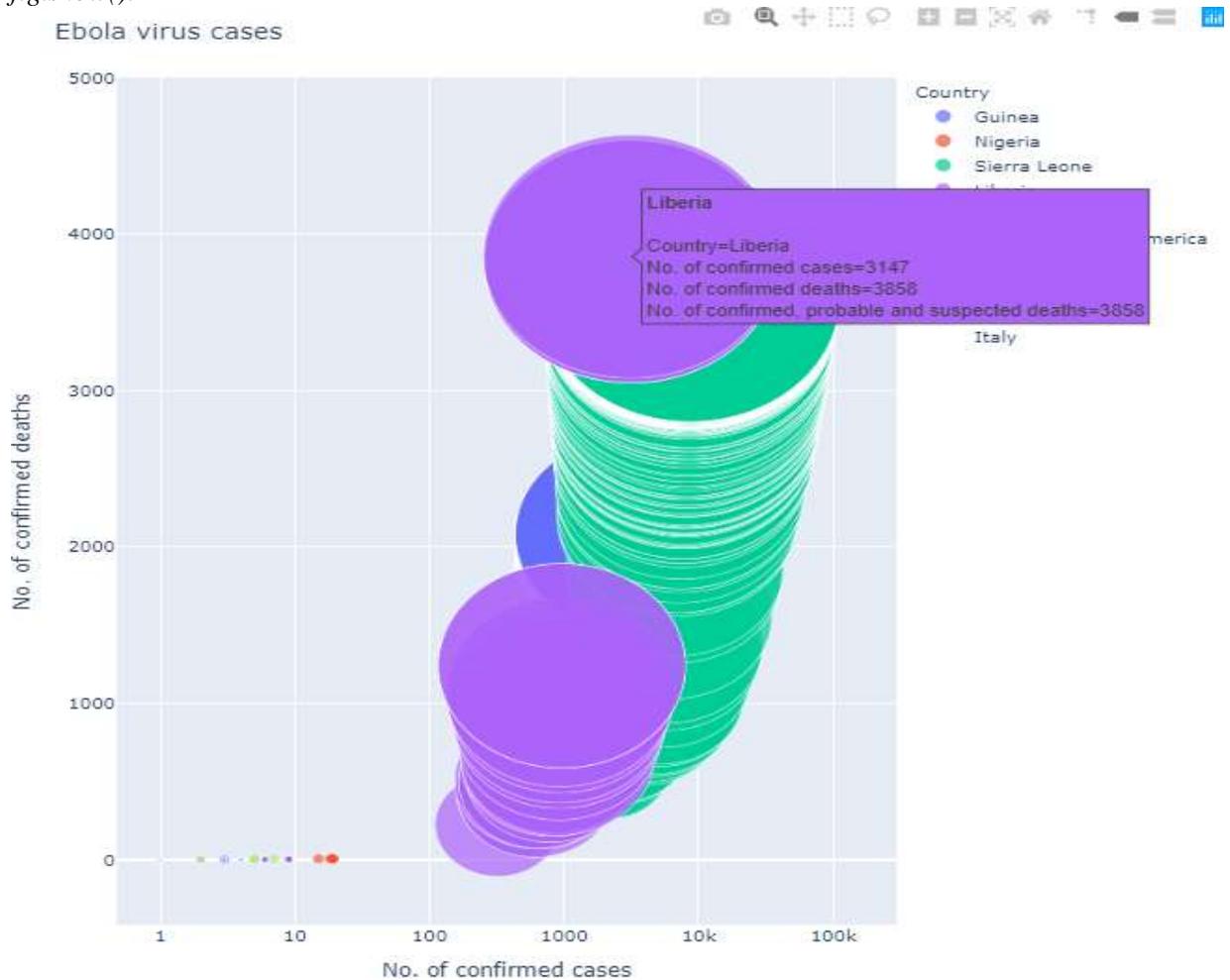


Fig 17: Bubble plot visualization of Covid-19 cases.

From the above bubble plot it is observed that Sierra Leone, Liberia and Guinea are having the higher number of Ebola Virus Disease (EVD) confirmed, probable and suspected cases and the death toll is higher in Italy as per the confirmed cases from 22nd January, 2020 to 18th March, 2020. We can also observe from above graph that the no. of confirmed deaths are higher than no. of confirmed cases it is so because Ebola Virus Disease (EVD) is Fatal and the death rate is high and the it is not transmitted until the symptoms are visible. Almost 50% of the confirmed cases are may die.

- 5) Confirmed cases comparison of Ebola Virus Disease (EVD) and Covid-19 in first 31 days in the form of pie chart.

Step 1:

```
ec=e1.head(10)
ec
```

The above python code take the first 10 records i.e. 31 days records from 2014-08-29 to 2014-10-01 from the *e1* variable and stores in *ec* variable. Where *e1* stores the records of Date and no. of confirmed cases as we can observe it from above fig 3. The code *ec* will display the output as shown in the below fig 18:

Out[18]:

	Date	No. of confirmed cases
0	2014-08-29	1754.0
1	2014-09-05	2383.0
2	2014-09-08	2552.0
3	2014-09-12	2639.0
4	2014-09-16	3017.0
5	2014-09-18	3095.0
6	2014-09-22	3341.0
7	2014-09-24	3487.0
8	2014-09-26	3626.0
9	2014-10-01	3973.0

Fig 18: Displaying first 10 records i.e. 31 days of Date and No. of confirmed cases of ebola virus.

Step 2:

```
cc=c1.head(31)
cc
```

The above python code take the first 31 records from 01-22-2020 to 02-21-2020 from the *c1* variable and stores in *cc* variable. Where *c1* stores the records of ObservationDate and Confirmed cases as we can observe it from above fig 4. The code *cc* will display the output as shown in the below fig 19:

Out[82]:

	ObservationDate	Confirmed
0	01/22/2020	555.0
1	01/23/2020	653.0
2	01/24/2020	941.0
3	01/25/2020	1438.0
4	01/26/2020	2118.0
5	01/27/2020	2927.0
6	01/28/2020	5578.0
7	01/29/2020	6165.0
8	01/30/2020	8235.0
9	01/31/2020	9925.0
10	02/01/2020	12038.0
11	02/02/2020	16787.0
12	02/03/2020	19881.0
13	02/04/2020	23892.0
14	02/05/2020	27636.0
15	02/06/2020	30818.0
16	02/07/2020	34392.0
17	02/08/2020	37121.0
18	02/09/2020	40151.0
19	02/10/2020	42763.0
20	02/11/2020	44803.0
21	02/12/2020	45222.0
22	02/13/2020	60370.0
23	02/14/2020	66887.0
24	02/15/2020	69032.0
25	02/16/2020	71226.0
26	02/17/2020	73260.0
27	02/18/2020	75138.0
28	02/19/2020	75641.0
29	02/20/2020	76199.0
30	02/21/2020	76843.0

Fig 19: Displaying first 31 days records of ObservationDate and Confirmed cases of Covid-19.

Step 3:

```

labels = ['Confirmed', 'No. of confirmed cases']
pieData = [cc ['Confirmed'].sum(), ec ['No. of confirmed cases'].sum()]
plt.axis("equal")
colors=['green','yellow']
plt.pie(pieData, labels = labels, colors=colors,
        shadow = True, explode = (0.0, 0.2),
        radius = 1.2, autopct = '%1.3f%%')
plt.legend()
plt.title('Percentage of confirmed cases of ebola and covid-19',fontsize=15,fontweight='bold')
plt.show()

```

The above python code takes the number of confirmed cases from both Covid-19 and ebola dataset and plot the overall confirmed cases of first 31 days and compare. The output of the above code is displayed by the help of `plt.show()` as shown in the below fig 20:

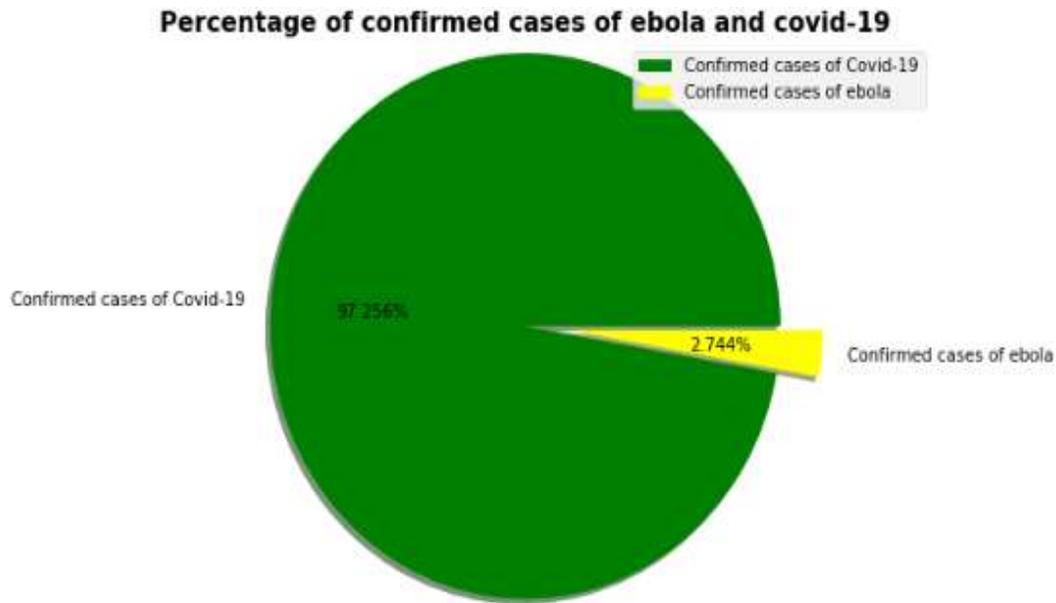


Fig 20: Pie chart representation of first 31 days confirmed cases comparison between Ebola and covid-19.

From the above fig 20 it is clearly observed that the number of confirmed cases of Covid-19 are far higher than that of Ebola Virus Disease in the first 31 days. As mentioned above the rate of contagiousness of Covid-19 is faster than that of Ebola Virus Disease (EVD) due to its asymptomatic transmission.

- 6) Confirmed death cases comparison of Ebola Virus Disease (EVD) and Covid-19 in first 31 days in the form of pie chart.

Step 1:

```
cod=cd.head(31)
cod
```

The above python code take the first 31 records from 01-22-2020 to 02-21-2020 from the *cd* variable and stores in *cod* variable. Where *cd* stores the records of ObservationDate and Deaths cases as we can observe it from above fig 10. The code *cod* will display the output as shown in the below fig 20:

Out[21]:

	ObservationDate	Deaths
0	01/22/2020	17.0
1	01/23/2020	18.0
2	01/24/2020	26.0
3	01/25/2020	42.0
4	01/26/2020	56.0
5	01/27/2020	82.0
6	01/28/2020	131.0
7	01/29/2020	133.0
8	01/30/2020	171.0
9	01/31/2020	213.0
10	02/01/2020	259.0
11	02/02/2020	362.0
12	02/03/2020	426.0
13	02/04/2020	492.0
14	02/05/2020	564.0
15	02/06/2020	634.0
16	02/07/2020	719.0
17	02/08/2020	806.0
18	02/09/2020	906.0
19	02/10/2020	1013.0
20	02/11/2020	1113.0
21	02/12/2020	1118.0
22	02/13/2020	1371.0
23	02/14/2020	1523.0
24	02/15/2020	1666.0
25	02/16/2020	1770.0
26	02/17/2020	1868.0
27	02/18/2020	2007.0
28	02/19/2020	2122.0
29	02/20/2020	2247.0
30	02/21/2020	2251.0

Fig 21: Displaying first 31 days records of ObservationDate and Deaths of Covid-19.

Step 2:

```
evd=ed.head(10)
```

```
evd
```

The above python code take the first 10 records i.e. 31 days records from 2014-08-29 to 2014-10-01 from the *ed* variable and stores in *evd* variable. Where *ed* stores the records of Date and no. of confirmed death cases as we can observe it from above fig 9. The code *evd* will display the output as shown in the below fig 22:

Out[22]:

	Date	No. of confirmed deaths
0	2014-08-29	898.0
1	2014-09-05	1243.0
2	2014-09-08	1376.0
3	2014-09-12	1386.0
4	2014-09-16	1513.0
5	2014-09-18	1590.0
6	2014-09-22	1687.0
7	2014-09-24	1698.0
8	2014-09-26	1837.0
9	2014-10-01	2006.0

Fig 22: Displaying first 10 records i.e. 31 days of Date and No. of confirmed death cases of ebola virus.

Step 3:

```

labels = ['Deaths of covid-19', 'Deaths of ebola']
pieData = [cod ['Deaths'].sum(), evd ['No. of confirmed deaths'].sum()]
plt.axis("equal")
colors=['green','yellow']
plt.pie(pieData, labels = labels, colors=colors,
        shadow = True, explode = (0.0, 0.2),
        radius = 1.2, autopct = '%1.3f%%')
plt.legend()
plt.title('Percentage of confirmed deaths of ebola and covid-19',fontsize=15,fontweight='bold')
plt.show()

```

The above python code takes the number of confirmed death cases from both Covid-19 and ebola dataset and plot the overall confirmed death cases of first 31 days and compare. The output of the above code is displayed by the help of `plt.show()` as shown in the below fig 23:

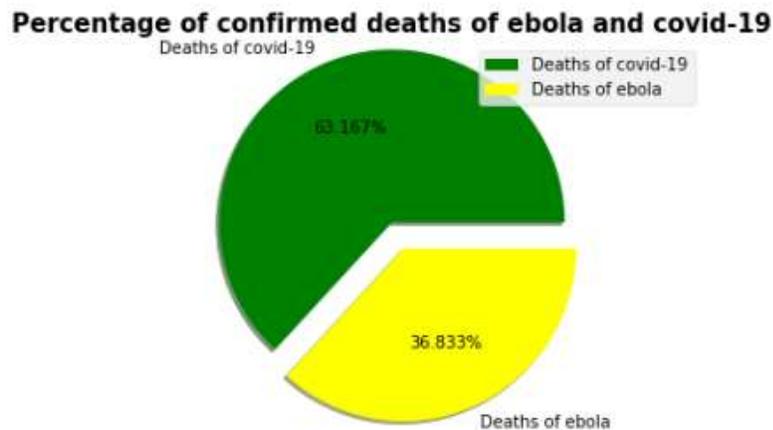


Fig 23: Pie chart representation of first 31 days confirmed death cases comparison between Ebola and covid-19.

From the above fig 23 it is clearly observed that the number of confirmed death cases of Covid-19 are far higher than that of Ebola Virus Disease in the first 31 days. As the transmission rate is slow in case of Ebola Virus Disease (EVD) as the disease cannot be transmitted until the symptoms are visible unlike Covid-19 where the transmission of disease is faster due to its asymptomatic transmission.

7) Country wise Ebola Virus Disease (EVD) cases in the form of horizontal bar graph:

Step 1:

```

SI = ebola.groupby("Country")
SI

```

The python code takes the `Country` column from ebola dataset and groups it with the help of `groupby()` predefined function present in python. It stores the grouped country data in `SI` variable.

Step 2:

```
Sta=SI.sum()
Sta
```

The code sum all the types of cases as per the *Country* column present in the ebola dataset from *SI* variable with the help of `sum()` predefined function of python and store it in the *Sta* variable. The code *Sta* is used to display the output as shown in the below fig 24:

Out[52]:

Country	No. of suspected cases	No. of probable cases	No. of confirmed cases	No. of confirmed, probable and suspected cases	No. of suspected deaths	No. of probable deaths	No. of confirmed deaths	No. of confirmed, probable and suspected deaths
Guinea	4293.0	100201.0	749622.0	854118.0	45.0	99799.0	461211.0	563069.0
Italy	0.0	0.0	141.0	141.0	0.0	0.0	0.0	0.0
Liberia	1212033.0	474103.0	768738.0	2459122.0	5933.0	8676.0	26937.0	1099833.0
Mali	1.0	239.0	1658.0	1865.0	0.0	5.0	21.0	1427.0
Nigeria	10.0	255.0	4840.0	5105.0	0.0	255.0	1784.0	2039.0
Senegal	4.0	0.0	254.0	258.0	0.0	0.0	0.0	0.0
Sierra Leone	893756.0	67657.0	2041871.0	3003428.0	37054.0	50786.0	797984.0	888834.0
Spain	0.0	0.0	243.0	243.0	0.0	0.0	0.0	0.0
United Kingdom	0.0	0.0	221.0	215.0	0.0	0.0	0.0	0.0
United States of America	0.0	0.0	967.0	967.0	0.0	0.0	242.0	242.0

Fig 24: Displaying all the types of cases of ebola virus according to Country column of ebola dataset.

Step 3:

```
Sta.plot.barh(figsize=(20,26),width=0.5, fontsize=18,alpha=1)
plt.ylabel("country",fontsize=25,fontweight='bold')
plt.xlabel("types of cases",fontsize=25,fontweight='bold')
plt.title("country wise ebola cases",fontsize=30,fontweight='bold')
plt.legend(fontsize=18)
plt.show()
```

The python code takes *Country* column in Y-axis and all the other cases count in X-axis and displays the all the types of cases of ebola virus present in ebola dataset with the help of `plt.show()` as shown in the below fig 25:

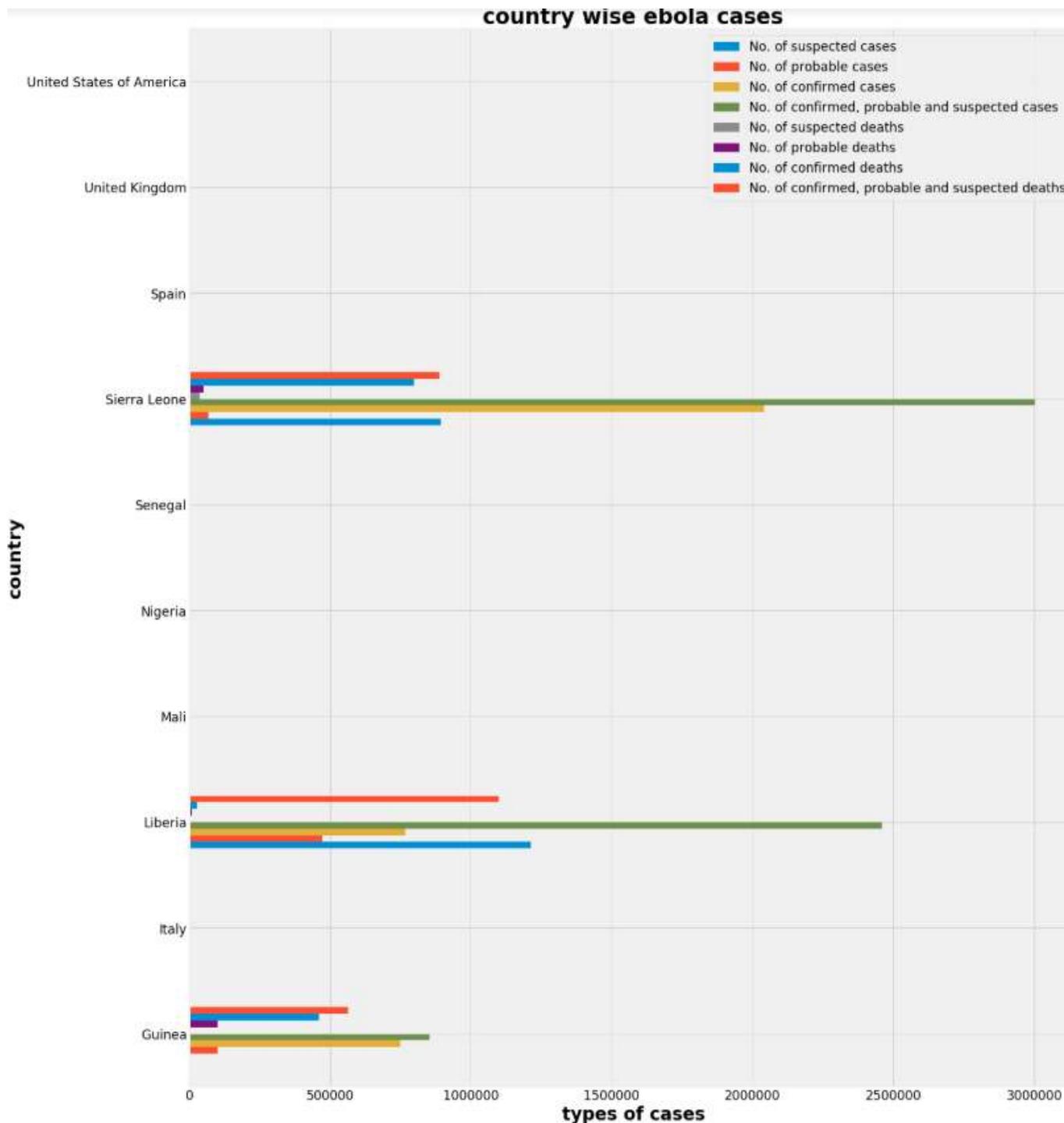


Fig 25: Displaying all the ebola virus cases in a horizontal bar graph.

From the above graph it is observed that Sierra Leone is the most effected country due to Ebola Virus Disease (EVD) followed by Liberia and Guinea as compared to others countries effected by the Ebola virus. The Ebola Virus Disease outbreaks occur mainly in tropical regions of sub-Saharan Africa. The largest outbreak to date was the epidemic in West-Africa, which occurred from December 2013 to January 2016, with 28,646 cases and 11,323 deaths which is approximately 50% of the confirmed cases.

Other outbreaks in Africa began in the Democratic Republic of the Congo in May 2017, and 2018. In July 2019, the World Health Organization (WHO) declared the Congo Ebola outbreak a world health emergency. The above graph is of outbreak from 22nd December 2014 to 23rd May 2016. The West African EVD epidemic of 2013–2016, which caused more than 28,000 cases which is larger than the recent Ebola Virus outbreak from 2017-2019. Each of the countries involved in that outbreak had more cases than the DRC outbreak.

The main route of EVD transmission, in this as in previous outbreaks, was direct personal contact with the blood or other body fluids of a person with symptomatic. A sudden spike in transmission in Guinea during March and April 2014 generated more than 100 new cases in total from there, infection spread within Guinea and across the nearby national boundaries. The disease appeared in Lofa and Margibi counties in northern Liberia before the end of March and in the Kailahun district in eastern Sierra Leone during May.

It is also observed that although the first case was reported in Guinea and initially the cases were high in this area but later Sierra Leone was most effected by the virus in this outbreak.

8) Overall effect of Covid-19 and Ebola Virus Disease (EVD):

Step 1:

```
labels = ['No. of confirmed cases', 'No. of confirmed deaths', 'No. of probable deaths']
pieData = [ebola['No. of confirmed cases'].sum(), ebola['No. of confirmed deaths'].sum(), ebola['No. of probable deaths'].sum()]
plt.axis("equal")
colors=['Green','Blue','Red']
plt.pie(pieData, labels = labels, colors=colors,
        shadow = True, explode = (0.0, 0.1, 0.1),
        radius = 1.3, autopct = '%1.3f%%')
plt.legend()
plt.title('Ebola virus effect',fontsize=15,fontweight='bold')
plt.show()
```

```
labels = ['Confirmed', 'Deaths', 'Recovered']
pieData = [covid['Confirmed'].sum(), covid['Deaths'].sum(), covid['Recovered'].sum()]
plt.axis("equal")
colors=['Green','Blue','Red']
plt.pie(pieData, labels = labels, colors=colors,
        shadow = True, explode = (0.0, 0.1, 0.1),
        radius = 1.3, autopct = '%1.3f%%')
plt.legend()
plt.title('Covid-19 effect',fontsize=15,fontweight='bold')
plt.show()
```

The above python code take all the cases of Ebola Virus Disease (EVD) and Covid-19 and plot 2 pie charts respectively and shows the effect of the Ebola Virus Disease (EVD) and Covid-19 for better understanding. The above python code displays the output with help of *plt.show()* as shown in the bellow fig 24:

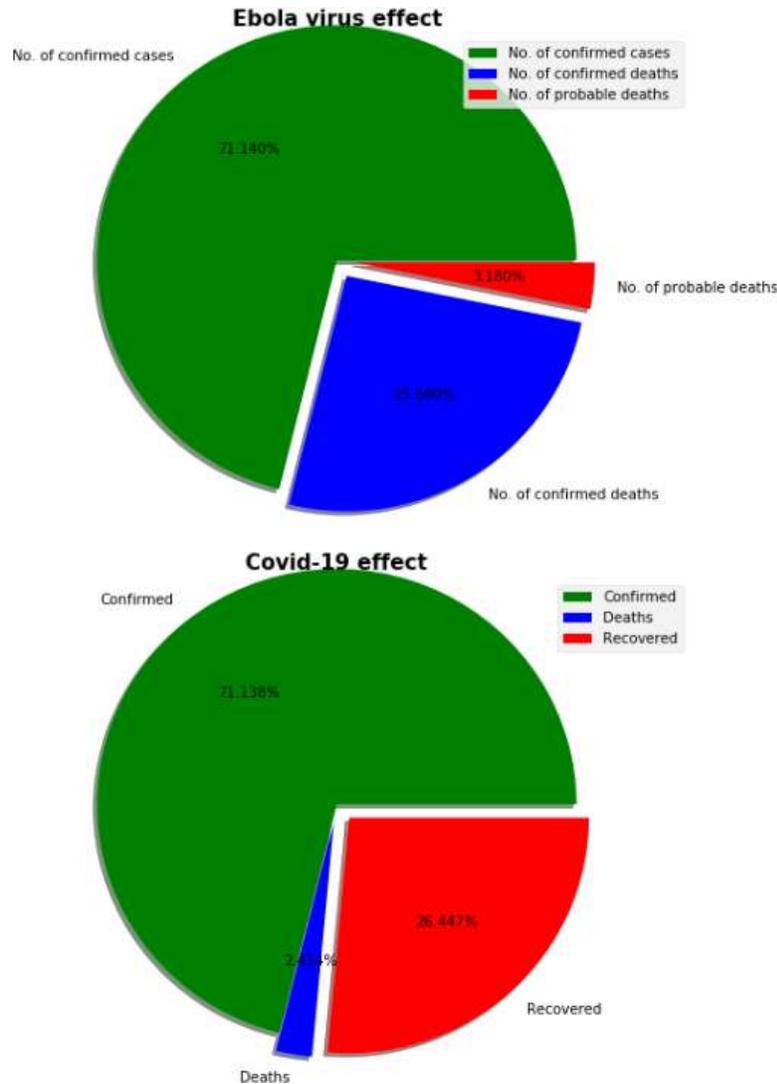


Fig 24: Displaying the overall effect of Ebola Virus Disease (EVD) and Covid-19.

From the above figure it is clearly observed the effect of Ebola Virus Disease (EVD) and Covid-19 and they are being compared to other types of cases present in the dataset. We can also observe that in Ebola effect virus effect the No. of confirmed deaths are high approximately 50% of the no. of confirmed cases which implies approximately 50% people infected from the Ebola virus have lost their lives, this clearly shows the fatality rate of the Ebola virus.

Whereas, in Covid-19 effect graph we can clearly observe that the percentage of confirmed cases are very much higher and the number of deaths as compared to the number of confirmed cases are very low. This shows the fast contagiousness and low death rate of the Covid-19 as compared to Ebola Virus Disease (EVD).

Conclusion:

From all the above illustrations it has been observed that Ebola Virus disease (EVD) is extremely deadly, killing up to 50 percent of those who got sick. But because it predominantly spread through bodily fluids like sweat and blood during the last stages of the disease, it wasn't as contagious as COVID-19. Also as symptoms were so severe in case of Ebola Virus Disease (EVD), health officials were able to quickly identify those who'd been in contact with people who had it and isolate them to stop further transmission and they were given treatment. The outbreak in West Africa from 2014 to 2016 was the largest Ebola epidemic to occur in which nearly 28,610 people were infected and 11,308 died. Ebola is rare, but it has a high mortality rate of about 50% unlike Covid-19 which is being spread rapidly.

It has been also above from above visualizations that the Ebola Virus Disease (EVD) is more fatal as compared to Covid-19. Whereas, the Covid-19 is more contagious than the Ebola Virus Disease (EVD) due to its asymptomatic transmission. People are much aware of social distancing, physical hygiene, and importance of wearing a mask etc. by the help of previous pandemics like Ebola Virus Disease (EVD). COVID-19, the disease caused by the novel coronavirus, isn't the first threatening disease that's surged around the world and nor will it be the last.

So, when will things calm down with COVID-19?

According to Mr. Mark R.Schleiss director, professor of pediatrics in Minneapolis, Minnesota, USA, it's going to take herd immunity which basically blocks out the virus when a large population already being sick and is immune from along with an effective vaccine.

"We really, really need a vaccine," he said, adding that because the Food and Drug Administration will need to prove a vaccine is safe, it could take a year or two — best case scenario.

We also have a lot more we need to learn: the prevalence of the infection along with how you contract the virus and all the different roads of transmission.

Until then, we're going to need to practice social distancing to help minimize the number of people who contract it, says Johnson.

We'll need to work together to limit exposure to one another — especially with older adults and people with underlying illness who have the greatest risk of developing severe symptoms.

We don't need to panic. Remember: The vast majority of COVID-19 cases are mild. But we do need to take action to contain the spread and protect those who are most vulnerable.

References:

1. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
2. <https://www.pharmaceutical-technology.com/features/coronavirus-ebola-latest-global-health-emergencies-who/>
3. <https://www.hindawi.com/journals/ijmicro/2015/769121/>
4. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>
5. <https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease>

6. <https://www.healthline.com/health-news/how-deadly-is-the-coronavirus-compared-to-past-outbreaks>
7. <https://www.mdlinx.com/article/5-viruses-more-dangerous-than-the-new-coronavirus/7wQzXjJudXjqIZVb8eU6LI>
8. <https://www.afro.who.int/news/drawing-ebola-readiness-tackle-covid-19>
9. <https://www.kaggle.com/imdevskp/ebola-outbreak-20142016-complete-dataset>
10. <https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge>
11. <https://www.hindawi.com/journals/ijmicro/2015/769121/>
12. <https://www.nejm.org/ebola-virus>
13. https://www.google.com/search?q=symptoms+of+covid-19&rlz=1C1GCEA_enAE791AE791&oq=symptoms+of+covid-19&aqs=chrome..69i57j0l7.5908j0j7&sourceid=chrome&ie=UTF-8