

# **“Statistical Analysis of Ebola Disease 2014-2016 Outbreak”**

**A Project Report**

*Submitted by*

**DIVYANSH CHOUBISA**

**17BCON393**

*In partial fulfilment for the award of the degree*

*of*

**BACHELORS OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**



**At**

**JECRC UNIVERSITY, JAIPUR**

### **Candidate's Declaration**

I, DIVYANSH CHOUBISA hereby declare that the work presented in this project entitled **“Statistical Analysis of Ebola Disease 2014-2016 Outbreak”** in partial fulfilment for the award of Degree of Bachelor of Technology, submitted in the **Department of Computer Science** at JECRC University, Jaipur, is an authentic record of my own research work under the supervision of ANUJA MISHRA

I also declare that the work embodied in the present Project Report is my original work and has not been copied from any Journal/Project Report/book, and has not been submitted by me for any other Degree/Diploma.

**Divyansh Choubisa**  
**Enrollment No. – 17BCON393**  
**Date- 17<sup>th</sup> April, 2020**

### **Certificate of the Supervisor(s)**

This is to certify that the Project entitled “**Statistical Analysis of Ebola Disease 2014-2016 Outbreak**” submitted by DIVYANSH CHOUBISA bearing roll number 17BCON393 in partial fulfilment for award of Degree of Bachelor of Technology in the Department of COMPUTER SCIENCE of JECRC University, Jaipur, is a record of authentic work carried out by him under my supervision.

The matter embodied in this Project Report is the original work of the candidate and has not been submitted for the award of any other degree or diploma. It is further certified that he has worked with me for the required period in the Department of COMPUTER SCIENCE, JECRC University, Jaipur.

**ANUJA MISHRA**

**Date: 17<sup>th</sup> April, 2020**

## **APPENDIX 3**

### **Acknowledgements**

I would like to express my sincere gratitude to my project guide “**ANUJA MISHRA**” for giving me the opportunity to work on this topic. It would never be possible for us to take this project to this level without his innovative ideas and his relentless support and encouragement.

DIVYANSH CHOUBISA  
17BCON393

## **APPENDIX 4**

### **Abstract**

In this project, there is statistical analysis of Ebola outbreak 2014-2016, the 2014-2016 outbreak in West Africa was the largest and most complex Ebola outbreak than all others since the virus was first discovered in 1976.

This project is the statistical analysis, providing graphs and numerical figures of what this outbreak was, the damage and what was the fatality rate in different regions of world, according to the data collected.

**Keywords:** Statistical Analysis, Outbreak, Fatality Rate.

# Contents

List of figures.....	<i>i</i>
List of Tables.....	<i>ii</i>
Chapter 1 INTRODUCTION.....	1
1.1 What is Ebola Disease and What was Ebola Outbreak (2014-2016) ?	
1.2 Motivation and Objective	
1.3 Project Outline	
Chapter 2 Statistical Analysis.....	3
2.1 About Data	
Chapter 3 About Hotspots.....	6
Chapter 4 Ebola Spread Analysis.....	9
4.1 Year Wise	
4.2 Total Sum	
4.3 Fatality Rate Calculation results	
Chapter 5 Guinea.....	17
5.1 The climate of Guinea	
5.2 Analysis	
Chapter 6 Sierra Leone.....	22
6.1 The climate of Sierra Leone	
6.2 Analysis	
Bibliography.....	27

## APPENDIX 6

# List of Tables

Table 1: The main dataset..... 3

Table 2: Year-wise data.....9-10

Table 3: Total Sum Data.....12

---

# List of Figures

Figure 1: World Map.....	7
Figure 2: Western Africa Map.....	8
Figure 3: Year Wise Ebola Cases in Different Countries.....	10
Figure 4: Year Wise Ebola Deaths in Different Countries.....	11
Figure 5: All Kinds of Cases in Different Countries.....	12
Figure 6: All Kinds of Deaths in Different Countries.....	13
Figure 7: Month Wise Confirmed Cases in Guinea.....	18
Figure 8: No. of Confirmed Cases in a Day (Guinea)- Letter Value Plot.....	19
Figure 9: No. of Confirmed Deaths in a Day (Guinea)- Letter Value Plot.....	20
Figure 10: Guinea- Line Graph for Different Kind of Cases and Deaths.....	21
Figure 11: Month Wise Confirmed Cases in Sierra Leone.....	23
Figure 12: No. of Confirmed Cases in a Day (Sierra Leone)- Letter Value Plot...	24
Figure 13: No. of Confirmed Deaths in a Day (Sierra Leone)- Letter Value Plot.....	25
Figure 14: Sierra Leone -Line Graph for Different Kind of Cases and Deaths....	26

---



# Chapter 1

## Introduction

### What is Ebola Disease and What was Ebola Outbreak (2014-2016)?

Ebola is a dangerous virus that causes severe bleeding, organ failure and can lead to death.

There is no cure or specific treatment for the Ebola virus disease that is currently approved for market, although various experimental treatments are being developed. For past and current Ebola epidemics, treatment has been primarily supportive in nature.

From 2014 Ebola starts to spread in many countries of Western Africa and world causing a tremendous loss of lives of people. Many people died due to this deadly disease and there is still no perfect vaccine for this disease (April,2020).

### Motivation and Objective

- My purpose of this research was that I wanted to know about what was the damage caused by the Ebola to lives of humans.
- How much deadly this virus is?
- The Fatality Rate of Virus in different regions of world.
- I also wanted to know in which time of year this disease spreads the most?

With such mottos I did this project to really get into the details of Ebola with the help of data I got.

This project will definitely tell about what Ebola did to the countries and people of countries where it was spread most and the amount of loss and damage they suffered.

## PROJECT OUTLINE

The following list is containing of the phases of the project:

### **Data Collection Phase**

As the project is on detailed data analysis of Ebola (2014-2016) outbreak the discovery phase consist of the collection of data.

### **Data cleaning and Data Manipulation Phase**

The collected data was cleaned and it was manipulated to different kinds of data frames so I can work on them to bring out different insights and answers form data.

### **Data Exploration and Visualization Phase**

From the manipulated data different graphical visualization have been done in this stage.

### **Interpretation Phase**

Finally, with the help of different graphs and different facts and figures from data frames and understandable interpretation is done and important numerical figures have been brought out for the understanding.

---

# Chapter 2

## Statistical Analysis

It's the science of collecting, exploring and presenting large amounts of data to discover underlying patterns and trends. Statistics are applied every day – in research, industry and government – to become more scientific about decisions that need to be made.

For this project I have used the Ebola Outbreak (2014-2016) data for the statistical analysis purpose.

The technology used in this project is Python as a programming language and different packages of python for different tasks such as graphical visualization, getting coordinated of different places, generating maps, mathematical computation library, package that helps in manipulating different data frames.

## About Data

The data was collected from this website:

<https://www.kaggle.com/imdevskp/ebola-outbreak-20142016-complete-dataset>

The data collected looks like this: Table1 below

1 df.head()										
	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

It contains the information about

- Countries
- Date
- No. of suspected cases
- No. of probable cases
- No. of confirmed cases
- No. of suspected, probable and confirmed cases.
- No. of suspected deaths
- No. of probable deaths
- No. of confirmed deaths
- No. of Suspected, probable and confirmed deaths.

Now the main question here is what is the difference between probable, suspected and confirmed cases (In the context of Ebola)?

**Case Definitions for Ebola (SOURCE: WHO)**

**CONFIRMED CASE** - A person with laboratory confirmation of virus is a confirmed case.

**SUSPECTED CASE** – Any person alive or dead, suffering or having suffered from a sudden onset of high fever and had contact with:

- A suspected, probable or confirmed case of Ebola case.
- A dead or sick animal (for Ebola).

**PROBABLE CASE** – Any suspected case evaluated by a clinical OR

Any deceased suspected case (where it has not been possible to collect specimen for laboratory confirmation) having an epidemiological link with a confirmed case.

The data consist of 10 Columns and 2485 rows.

There were missing values / unavailable information in data which was essential and cannot be removed, also we can't take mean of the values because we can't fill the mean deaths of whole data containing every country to a particular country's missing value of deaths etc.

But the data was sufficiently good for the analysis of confirmed cases and deaths as there were fewer missing data.

So, to handle the missing data I just fill 0 in place of nan values i.e. missing values.

And the results were good to visualize and analyse and it affected dataset in a neutral way.

All results, conclusions, graphs and calculations are made only from this data no external data point was added to it from different source so, all results are based according to this data only.

---

# Chapter 3

## About Hotspots

According to this data I analysed I found that the Ebola disease was found in the following countries:

*Guinea*  
*Nigeria*  
*Sierra Leone*  
*Liberia*  
*Senegal*  
*United States of America*  
*Spain*  
*Mali*  
*United Kingdom*  
*Italy*

With the help of some technology I gathered the coordinates of the following countries which are as follows:

Latitude and longitude of Guinea are 10.7226226, -10.7083587.  
Latitude and longitude of Nigeria are 9.6000359, 7.9999721.  
Latitude and longitude of Sierra Leone are 8.6400349, -11.8400269.  
Latitude and longitude of Liberia are 5.7499721, -9.3658524.  
Latitude and longitude of Senegal are 14.4750607, -14.4529612.  
Latitude and longitude of United States of America are 39.7837304, -100.4458825.  
Latitude and longitude of Spain are 39.3262345, -4.8380649.  
Latitude and longitude of Mali are 16.3700359, -2.2900239.  
Latitude and longitude of United Kingdom are 54.7023545, -3.2765753.  
Latitude and longitude of Italy are 42.6384261, 12.674297.

[10.7226226, 9.6000359, 8.6400349, 5.7499721, 14.4750607, 39.7837304, 39.3262345, 16.3700359, 54.7023545, 42.6384261]  
[-10.7083587, 7.9999721, -11.8400269, -9.3658524, -14.4529612, -100.4458825, -4.8380649, -2.2900239, -3.2765753, 12.674297]

Then I located these coordinates into world map and marked Ebola hotspots i.e. where Ebola virus/disease was found.



Figure 1 showing the countries having cases of Ebola (2014 - 2016)



The above Figure 2 showing the zoomed in picture of Western African Countries where Ebola cases were found.

According to the data Ebola was mainly found in these Western African Countries.

In Europe the cases were found in United Kingdom, Spain and Italy.

Apart from these some cases were found in United States of America.

---



# Chapter 4

## Ebola Spread Analysis

In the data I further added 3 more columns they are:

Year

Day

Month

Adding these columns helps in the detailed year and month wise analysis of data.

### Year Wise Analysis

By some manipulation in original data frame I created a data frame representing the data year and country wise.

This data is as follows: Table 2 below

	Country	Year	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	2101.0	7794.0	56295.0	66190.0	45.0	7388.0	31014.0	40462.0
17	Sierra Leone	2014	31274.0	4230.0	156787.0	192291.0	2136.0	4818.0	41981.0	51945.0
14	Senegal	2014	4.0	0.0	37.0	41.0	0.0	0.0	0.0	0.0
11	Nigeria	2014	10.0	38.0	717.0	765.0	0.0	38.0	265.0	303.0
8	Mali	2014	0.0	18.0	111.0	129.0	0.0	5.0	21.0	101.0
20	Spain	2014	0.0	0.0	26.0	26.0	0.0	0.0	0.0	0.0
25	United States of America	2014	0.0	0.0	99.0	99.0	0.0	0.0	25.0	25.0
5	Liberia	2014	74602.0	59688.0	72254.0	210796.0	5929.0	8673.0	11385.0	95625.0
6	Liberia	2015	1131795.0	412533.0	693322.0	2237646.0	4.0	0.0	15548.0	999395.0
23	United Kingdom	2015	0.0	0.0	220.0	214.0	0.0	0.0	0.0	0.0
18	Sierra Leone	2015	857351.0	63140.0	1876380.0	2797015.0	34760.0	45760.0	752414.0	832934.0
3	Italy	2015	0.0	0.0	140.0	140.0	0.0	0.0	0.0	0.0
12	Nigeria	2015	0.0	216.0	4104.0	4320.0	0.0	216.0	1512.0	1728.0
26	United States of America	2015	0.0	0.0	864.0	864.0	0.0	0.0	216.0	216.0
15	Senegal	2015	0.0	0.0	216.0	216.0	0.0	0.0	0.0	0.0
21	Spain	2015	0.0	0.0	216.0	216.0	0.0	0.0	0.0	0.0
1	Guinea	2015	2192.0	91954.0	689976.0	784124.0	0.0	91958.0	428114.0	520071.0
9	Mali	2015	1.0	220.0	1540.0	1728.0	0.0	0.0	0.0	1320.0

22	Spain	2016	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
13	Nigeria	2016	0.0	1.0	19.0	20.0	0.0	1.0	7.0	8.0
16	Senegal	2016	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
10	Mali	2016	0.0	1.0	7.0	8.0	0.0	0.0	0.0	6.0
7	Liberia	2016	5636.0	1882.0	3162.0	10680.0	0.0	3.0	4.0	4813.0
4	Italy	2016	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
2	Guinea	2016	0.0	453.0	3351.0	3804.0	0.0	453.0	2083.0	2536.0
19	Sierra Leone	2016	5131.0	287.0	8704.0	14122.0	158.0	208.0	3589.0	3955.0
27	United States of America	2016	0.0	0.0	4.0	4.0	0.0	0.0	1.0	1.0

I have mainly worked with number of confirmed cases and number of confirmed deaths data because it will give most important results which all are in confirmed categories because at last what matters most is confirmed category.

To have more clear vision from this data here are some visualizations from it:

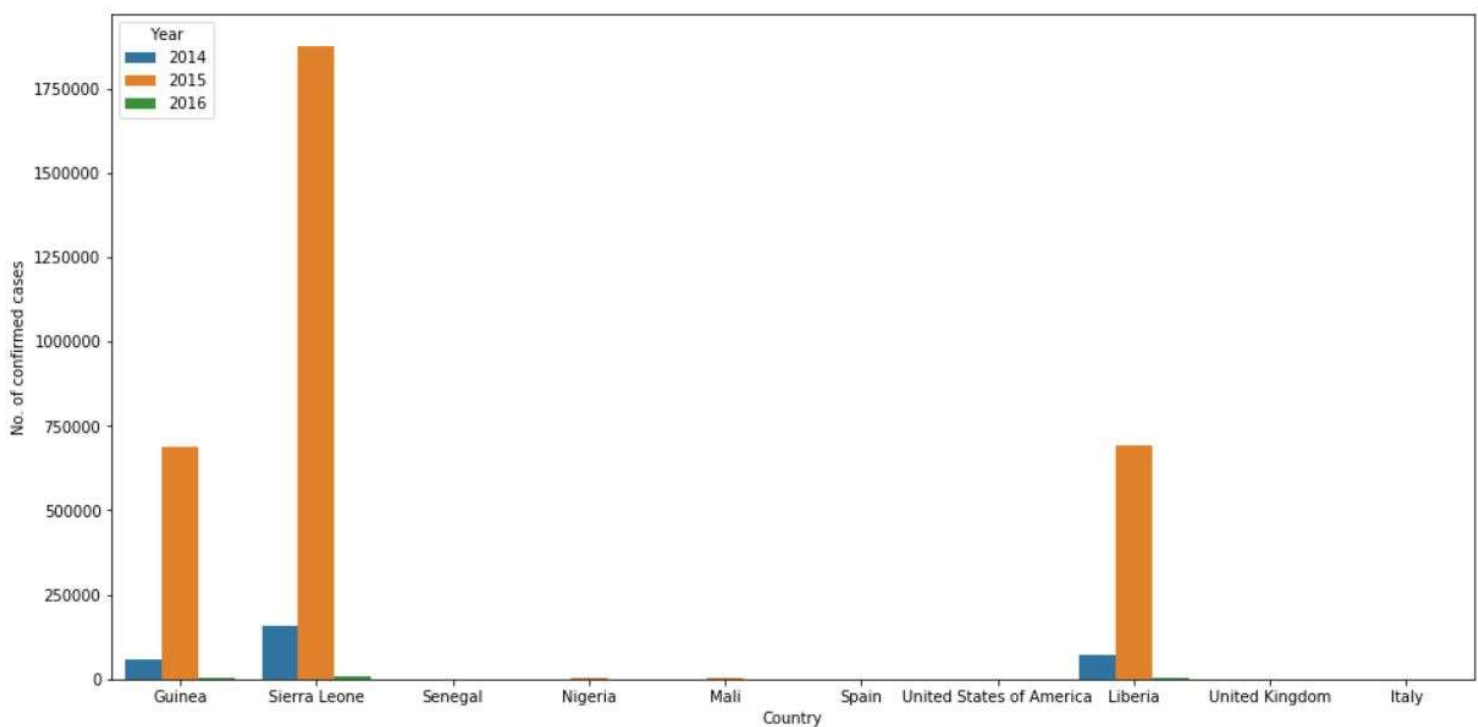


Figure 3

According to figure 3:

- The highest number of Ebola confirmed cases were found in mainly 3 countries and these countries were Guinea, Sierra Leone and Liberia.
- 2015 was the year when there were most cases confirmed in all countries.

- Highest number of cases were recorded in 2015 Sierra Leone over 17,50,000 i.e. 17 lakh 50 thousand.
- So, the main conclusion we can draw here is that Guinea, Sierra Leone and Liberia were most affected by the Ebola and were most dangerous to live.

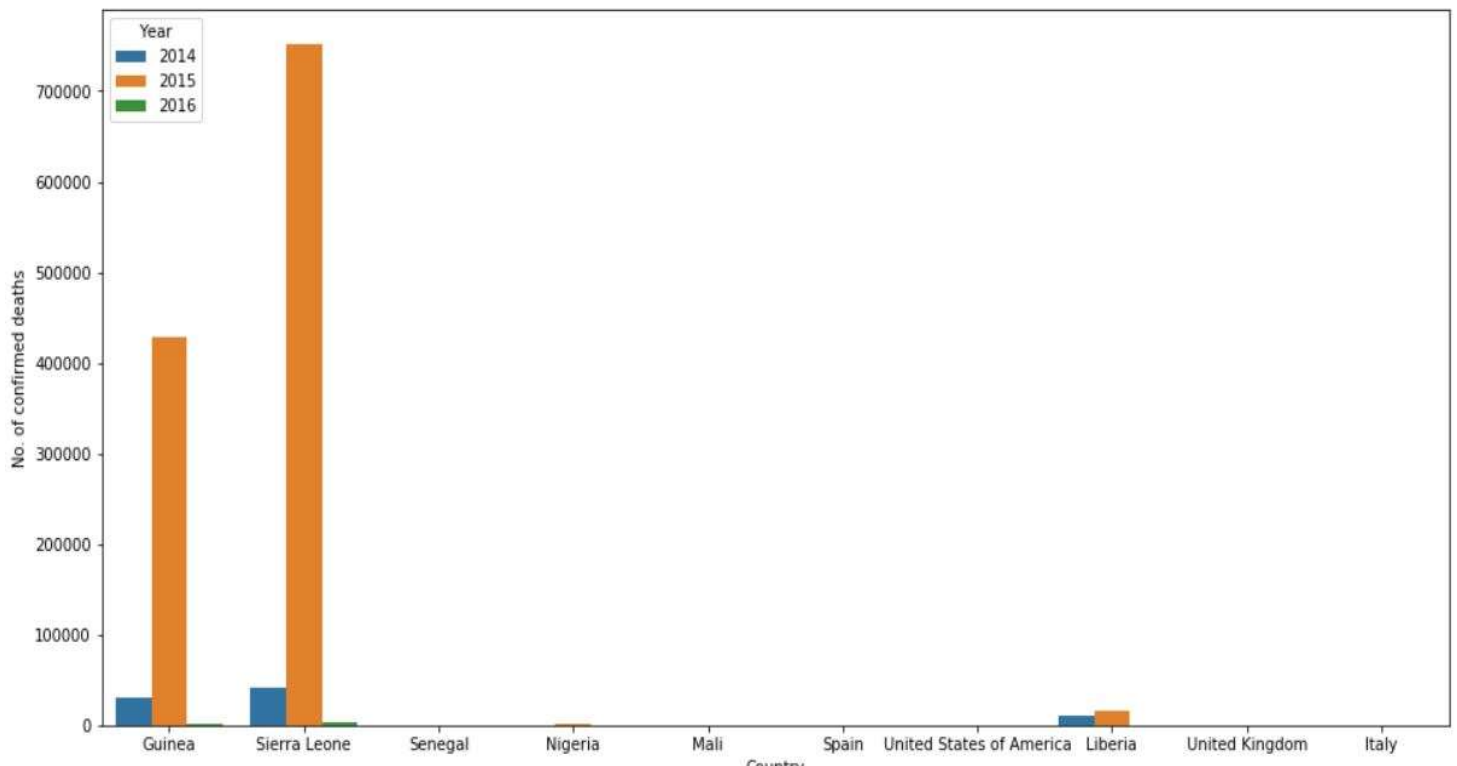


Figure 4

The main conclusions we can draw from figure 4 are that:

- As there were highest number of cases recorded in 2015 in mainly 3 countries Guinea, Sierra Leone and Liberia there were most confirmed deaths occurred in these countries in same year.
- In 2015 there were more than 7,00,000 confirmed deaths because of Ebola in Sierra Leone.
- The greatest number of deaths occurred in Guinea, Sierra Leone and Liberia.

## Total Sum Analysis

Now let's consider all kind of cases and deaths and understand using the graphs.

Again, I have created data frame which has all attributes with their final sum from 2014-2016

	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
0	Guinea	4293.0	100201.0	749622.0	854118.0	45.0	99799.0	461211.0	563069.0
1	Italy	0.0	0.0	141.0	141.0	0.0	0.0	0.0	0.0
2	Liberia	1212033.0	474103.0	768738.0	2459122.0	5933.0	8676.0	26937.0	1099833.0
3	Mali	1.0	239.0	1658.0	1865.0	0.0	5.0	21.0	1427.0
4	Nigeria	10.0	255.0	4840.0	5105.0	0.0	255.0	1784.0	2039.0
5	Senegal	4.0	0.0	254.0	258.0	0.0	0.0	0.0	0.0
6	Sierra Leone	893756.0	67657.0	2041871.0	3003428.0	37054.0	50786.0	797984.0	888834.0
7	Spain	0.0	0.0	243.0	243.0	0.0	0.0	0.0	0.0
8	United Kingdom	0.0	0.0	221.0	215.0	0.0	0.0	0.0	0.0
9	United States of America	0.0	0.0	967.0	967.0	0.0	0.0	242.0	242.0

Table 3

Now with the help of this data frame we can understand the different numerical figures.

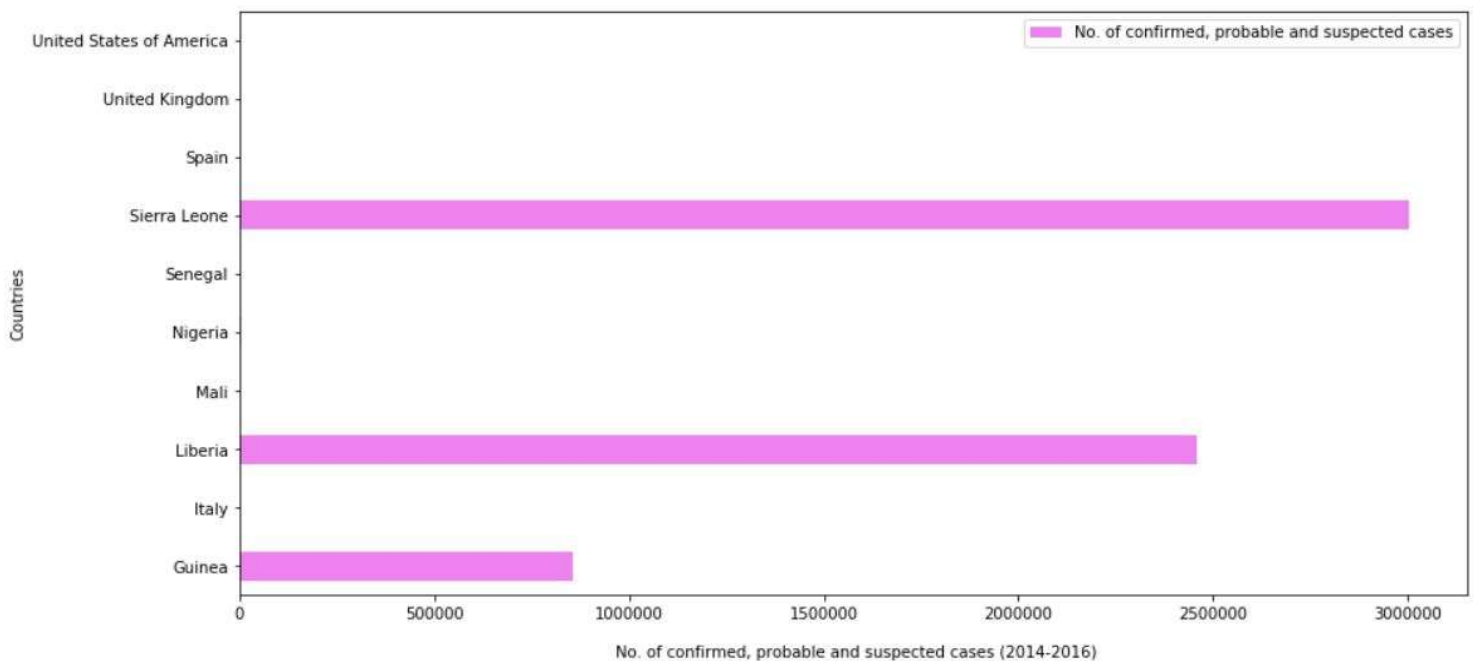


Figure 5

Now this graph tells us about No. of confirmed, probable and suspected cases from 2014-2016 in all the different countries.

Here Sierra Leone has maximum number of cases which are 30,03,428 i.e. 30 lakh, 3 thousand 4 Hundred, twenty-eight cases.

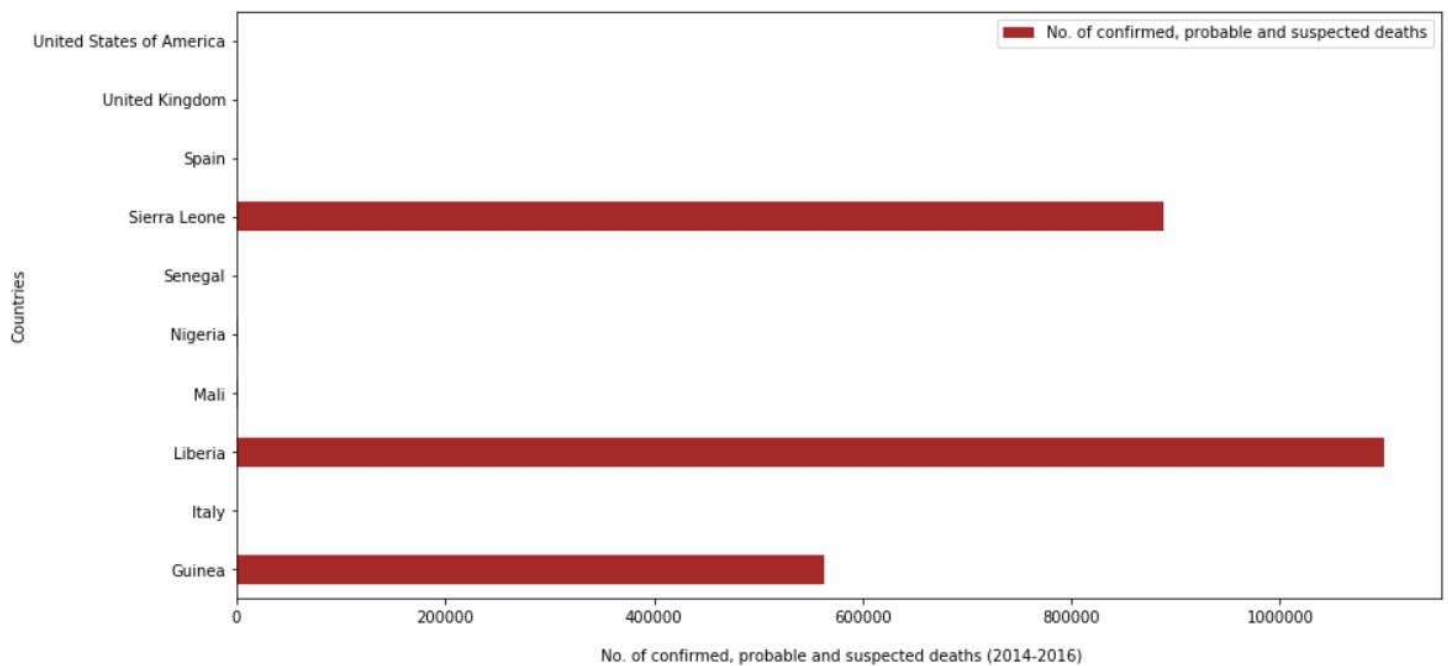


Figure 6

Now with the help of this data, the graph generated is telling that Liberia has highest amount of No. of probable, suspected and confirmed deaths from 2014-2016 i.e. 10,99,833 i.e. 10 Lakh ,99 Thousand ,8 Hundred, thirty-three deaths.

### Fatality Rate Calculation

Now comes the most important part of analysis which is called Fatality Rate.

But what is fatality rate?

The ratio of total people died among the total number of cases is called fatality rate.

More the fatality rate more the dangerous disease is. For example, in an area 10 people were having a particular disease and out of which 5 died then the fatality rate would be  $(5/10) * 100 = 50\%$ . Meaning only 50% people can/have survived the disease.

Ebola outbreak causes a very large amount of life loss and a huge amount of fear too. Ebola is such a deadly disease that the chance of death is very high.

The data I collected was having many missing /unavailable data. Because of this this deadly disease is showing very less fatality rate in some of the regions it might be correct or may be inaccurate but it can give idea of how much the deadly disease was in many regions whose data was sufficiently available to have an idea of the fatality rate (that in most cases is very high).

I created a program that has returned the fatality rate of Ebola in all the regions.

The fatality rates are as follows:

ACCORDING TO THIS DATA :

The Fatality rate in different countries might not be exactly true because there was much data missing/unavailable data in the dataset(due to some missing/unavailable data of confirmed deaths and confirmed cases)  
But it can give the an idea of fatality rate in different countries(which have some sufficient amount of data).

From 2014-2016 there were total number of 749622.0 cases of Ebola in Guinea out of which 461211.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Guinea was :  
61.525809007739895 %

From 2014-2016 there were total number of 141.0 cases of Ebola in Italy out of which 0.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Italy was :  
0.0 % (the fatality rate is low might be because of missing/unavailable data)

From 2014-2016 there were total number of 768738.0 cases of Ebola in Liberia out of which 26937.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Liberia was :  
3.5040546974391793 % (the fatality rate is low might be because of missing/unavailable data)

From 2014-2016 there were total number of 1658.0 cases of Ebola in Mali out of which 21.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Mali was :  
1.2665862484921593 % (the fatality rate is low might be because of missing/unavailable data)

From 2014-2016 there were total number of 4840.0 cases of Ebola in Nigeria out of which 1784.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Nigeria was :  
36.85950413223141 %

From 2014-2016 there were total number of 254.0 cases of Ebola in Senegal out of which 0.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Senegal was :  
0.0 % (the fatality rate is low might be because of missing/unavailable data)



From 2014-2016 there were total number of 2041871.0 cases of Ebola in Sierra Leone out of which 797984.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Sierra Leone was :

39.081019320025604 %

From 2014-2016 there were total number of 243.0 cases of Ebola in Spain out of which 0.0 number of people died due to Ebola

The Fatality rate of Ebola disease in Spain was :

0.0 % (the fatality rate is low might be because of missing/unavailable data)

From 2014-2016 there were total number of 221.0 cases of Ebola in United Kingdom out of which 0.0 number of people died due to Ebola

The Fatality rate of Ebola disease in United Kingdom was :

0.0 % (the fatality rate is low might be because of missing/unavailable data)

From 2014-2016 there were total number of 967.0 cases of Ebola in United States of America out of which 242.0 number of people died due to Ebola

The Fatality rate of Ebola disease in United States of America was :

25.0258531540848 %

The fatality rate of Ebola disease in world from 2014-2016 according to this data is: 36.09805649625689

From the above results we can say that in most of the regions where Ebola was widespread the fatality rate was very much high.

From the graphical analysis we drew a result that Guinea, Liberia and Sierra Leone were having highest number of confirmed cases and confirmed deaths compared to other countries.

So, the fatality rate of Ebola in Sierra Leone is 39.0810 %

Of Guinea is 61.5258 %

And of Liberia is 3.5040 % which is quite odd as we compare it to other countries as there is no perfect vaccine/cure for Ebola which simply mean that there was some missing/less data of confirmed deaths as we have seen in figure 6 has highest number of probable, suspected and confirmed deaths. There is also a possibility that the data recorded was more of probable and suspected deaths than confirmed deaths. It might happen because of the unstable situation that is created in a country when it is facing such dangerous situation created by a disease.

Anyways we can say that the fatality rate of Ebola is generally very high it can range from 25% - 61% or more (According to data I have worked on)

The Fatality rate of Ebola in Nigeria was also 36.8595%.

The fatality rate of whole world (2014-2016) according to this dataset is 36.0980 %. Definitely there was some missing data which can generate a strong possibility that the fatality rate can higher of this deadly disease than just 36.0980 %.

---



# Chapter 5

## GUINEA

Guinea is a country in West Africa, bordered on the west by the Atlantic Ocean. Its also know for the Mount Nimba strict nature reserve.

Now from the previous analysis we saw that the fatality rate of Ebola in Guinea was 61.52% and there were around 7,49,622 confirmed Ebola cases out of which 4,61,211 confirmed deaths occurred.

So, we need to study what happened in Guinea and because of Ebola and different kind of graphs.

### The Climate of Guinea

As we are studying the Ebola outbreak 2014-2016 in Guinea it is better to know about climate of a country throughout the year because weather can have effects on infectious diseases.

And also, in my analysis I have done analysis according to months which will help to understand what kind of weather is in which month when the Ebola disease is having its highest effect.

The climate of Guinea is tropical and humid with a wet and a dry season. Guinea is one of the wettest countries in West Africa.

July-November: Monsoon with a south westerly wind.

December-May: Dry Season with north easterly harmattan.

### Analysis

I have done different data analysis and manipulation techniques and generated different graphs on data frame for Guinea.

Here are some most useful insights:

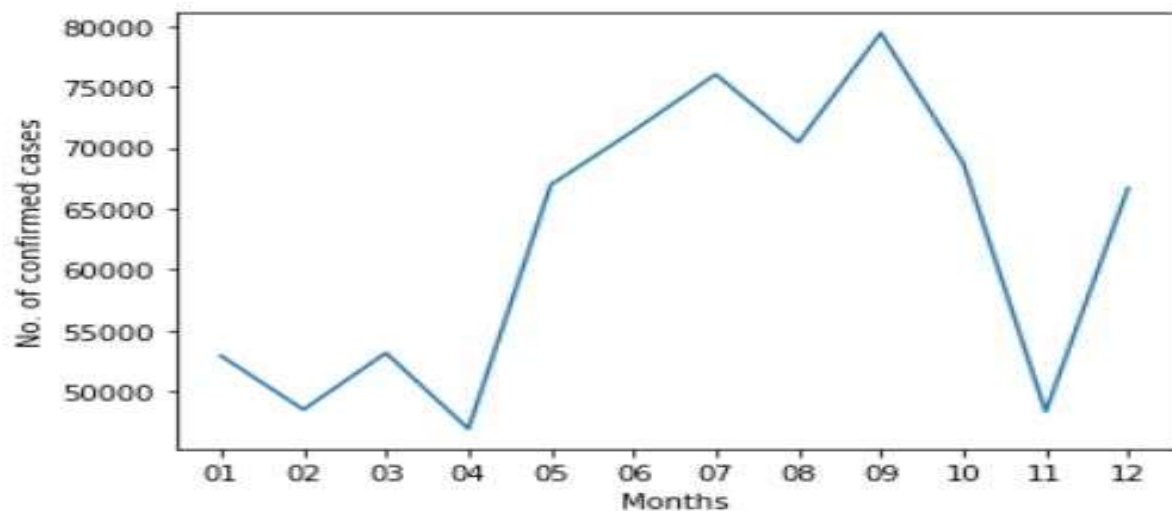


Figure 7

This graph is showing No. of confirmed cases of Ebola in different months of a year.

From this graph we can observe that there are tremendous number of confirmed cases of Ebola from May-October and we know that from July to November the weather of Guinea is Monsoon with a south westerly wind and from December to May it is Dry Season with north easterly harmattan. So, we can say that in Guinea Ebola spreads very much in the months of July – October/November when there is monsoon and cases decreases when dry season came from December to May.

Now Let's analyse some another interesting plot that is boxen plot (it is similar to box plot).

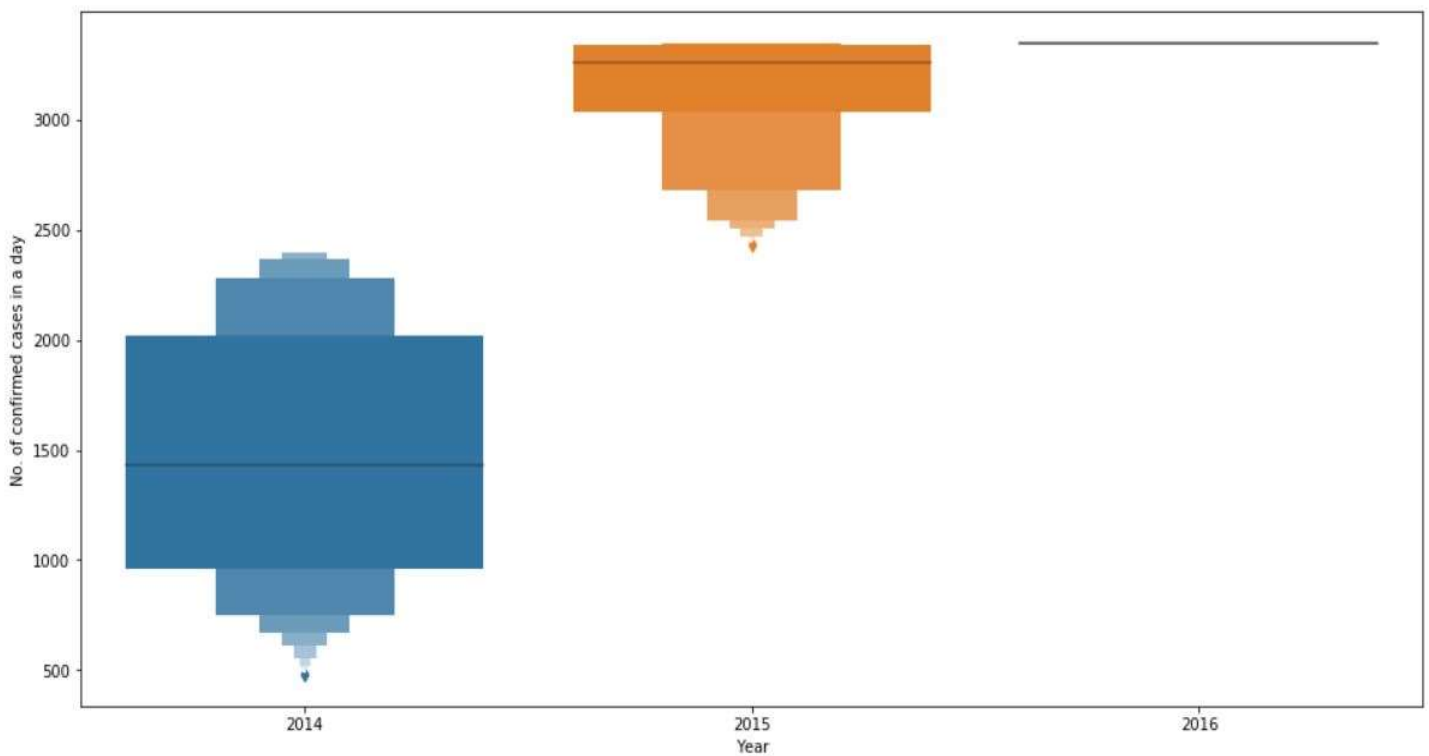


Figure 8

The above figure 8 is showing the number of confirmed cases in a day by Ebola in different years in Guinea.

In 2014 the 25<sup>th</sup> percentile of data is very near to 1000 and the 75<sup>th</sup> percentile is near around 2000 in 2014 which means in the throughout the 2014 more than 1000 or more than 1000 confirmed cases were getting in a day 75% of the time. And similarly

25 % of the time there were 2000 or more than 2000 confirmed cases were getting in a day in Guinea.

These numbers are so much high in 2015 that 50% of the time there were more than 3000 confirmed cases in a day.

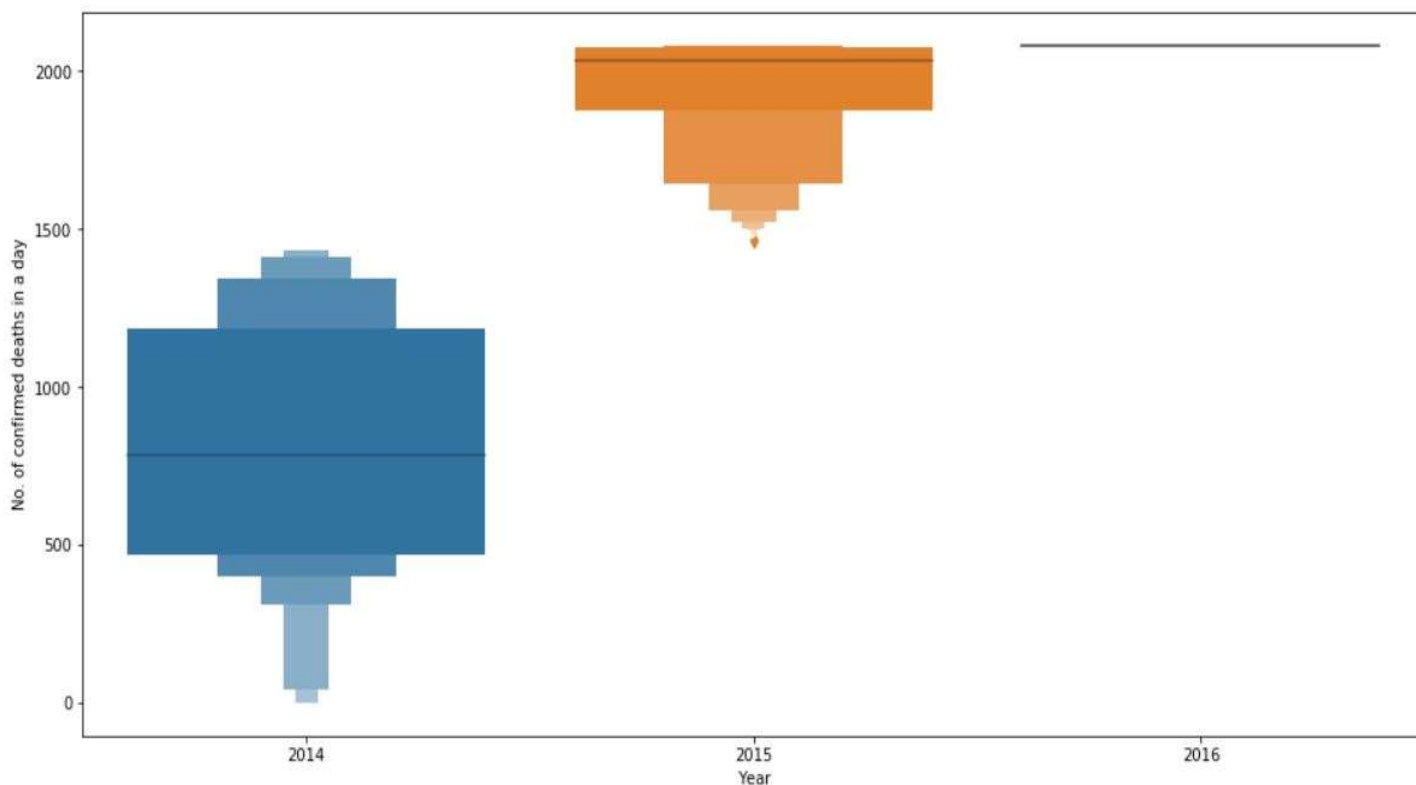


Figure 9

The above figure 9 is showing the number of confirmed deaths in a day by Ebola in different years in Guinea.

In 2014 the 25<sup>th</sup> percentile of data is around 500 and the 75<sup>th</sup> percentile is more than 1100 in 2014 which means in the throughout the 2014, 500 or more than 500 confirmed deaths were occurring in a day 75% of the time. And similarly 25 % of the time there were more than 1100 confirmed deaths were occurring in a day in Guinea.

These numbers are so much high in 2015 that 50% of the time there were near around 2000 confirmed deaths in a day.

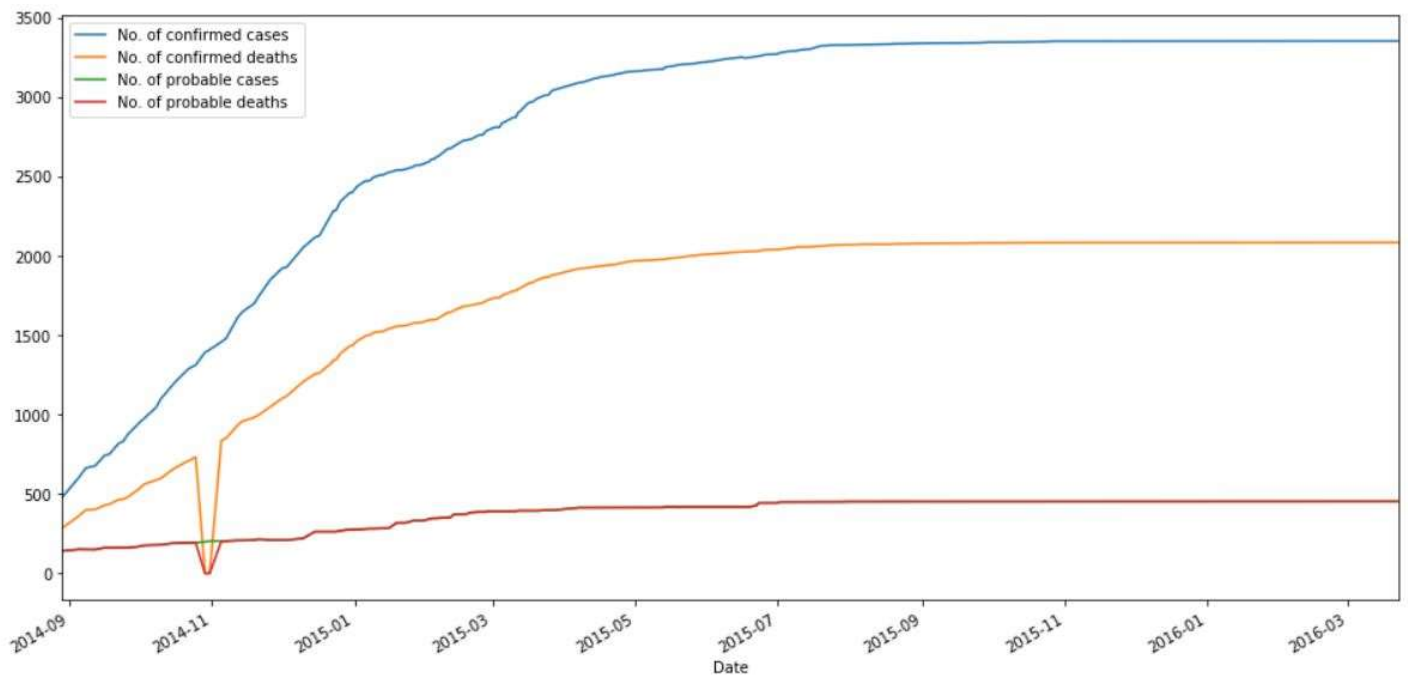


Figure 10

This above graph is showing the time line from 2014-2016 with different lines showing number of confirmed cases and deaths and number of probable cases and deaths.

In the line graph of number of confirmed deaths there is an outlier on the point near 2014-11 because of missing data being replaced with 0 value. Else whole graph is very clear.

We can say that number of confirmed cases were increasing from 2014 to 2016 and takes a constant rate of growth. We also know that the fatality rate of Ebola in Guinea was 61.525 % and that difference can be seen in the line of number of confirmed cases and number of confirmed deaths.

---

# Chapter 6

## SIERRA LEONE

Sierra Leone is a country in West Africa, on the Atlantic Ocean. Its known for white sand beaches lining the Freetown Peninsula.

During our analysis we found that the fatality rate of Ebola in Sierra Leone 39.0810 %. There were total 2041871 confirmed cases of Ebola and 797984 confirmed deaths occurred.

So, we need to study what happened in Sierra Leone and because of Ebola and different kind of graphs.

### The climate of Sierra Leone

As we are studying the Ebola outbreak 2014-2016 in Sierra Leone it is better to know about climate of a country throughout the year because weather can have effects on infectious diseases.

And also, in my analysis I have done analysis according to months which will help to understand what kind of weather is in which month when the Ebola disease is having its highest effect.

The climate Sierra Leone is tropical and humid all the year.

November-June: Very hot and dry but costal areas are cooled by sea breezes.

December-February: Dry, Dusty and Harmattan wind blows from Sahara.

June-November: Rapid rainfall

From March there are sporadic rainfalls with strong winds and August has most rainfall.

## Analysis

Here are some most useful insights:

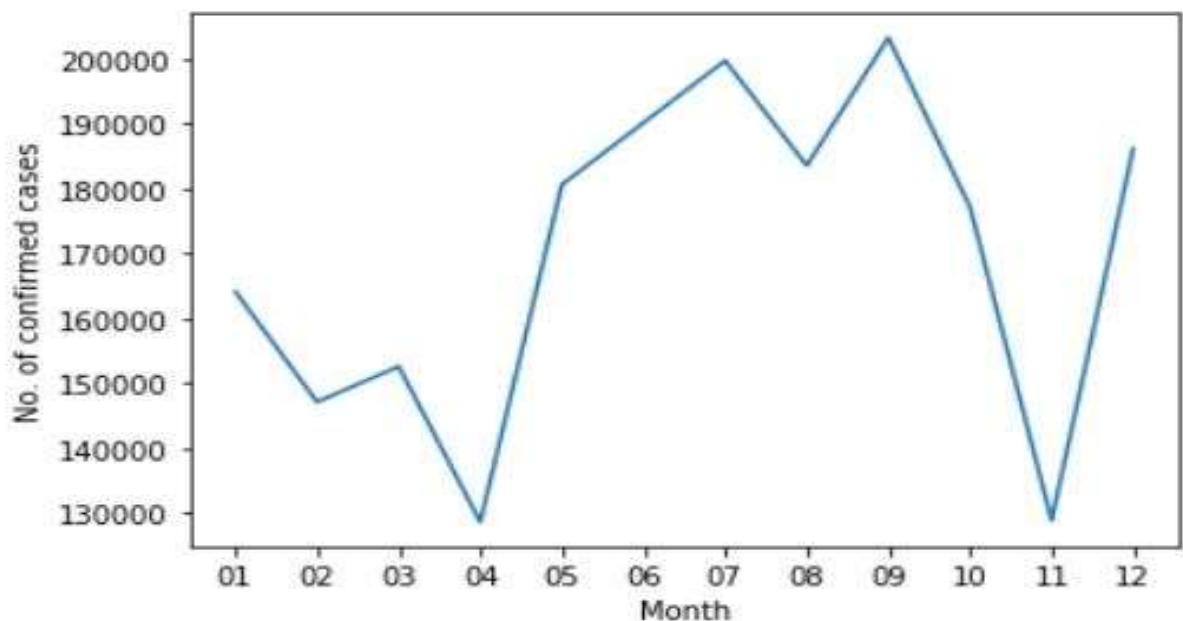


Figure 11

This graph is showing No. of confirmed cases of Ebola in different months of a year.

From this graph we can understand that from May to October there are tremendous amount of confirmed cases of Ebola. These cases have decreasing behaviour from the months of January to April.

In Sierra Leone there is rainfall from June to November which might be the cause of high number of confirmed Ebola cases while from December to February there is Dry, Dusty and Harmattan wind blows from Sahara and in March there are sporadic rainfalls (occasional).

So, we can say that during rainfall and wet humid environment the probability of Ebola to spread is very high in Sierra Leone.

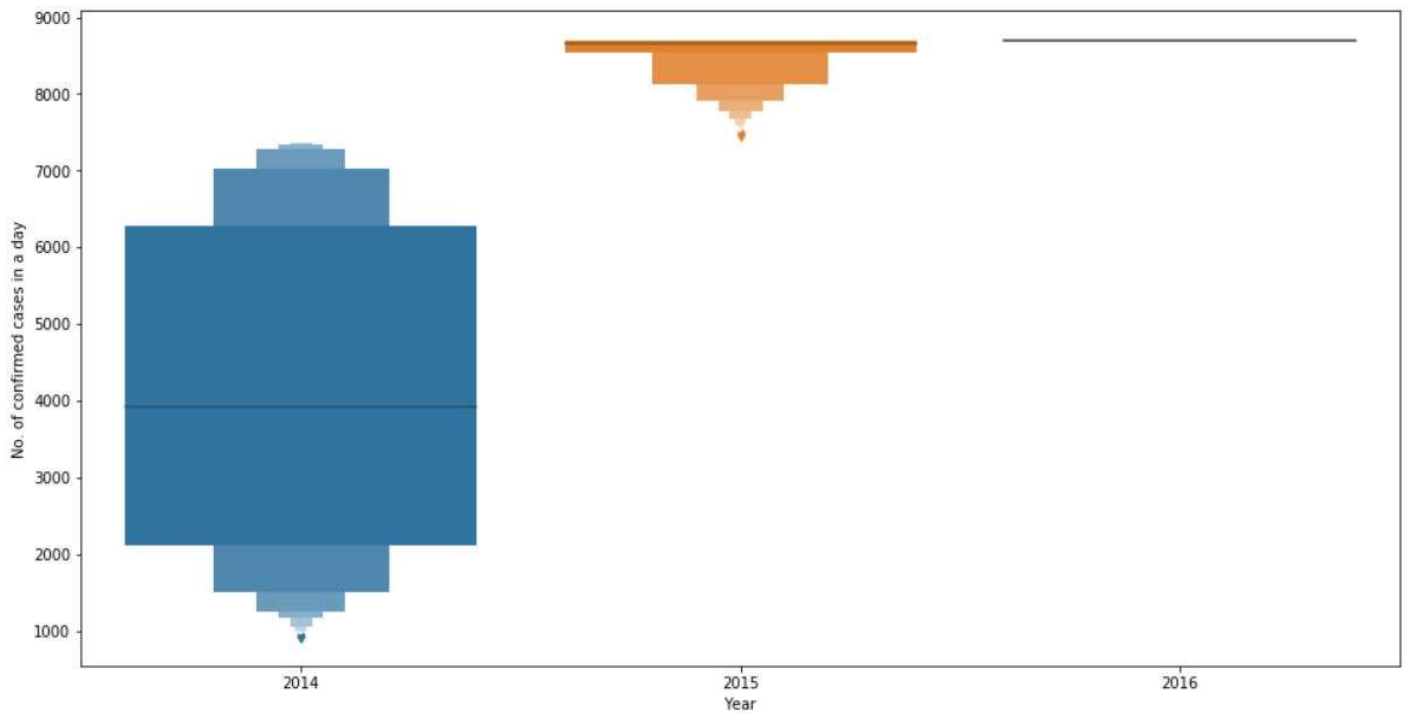


Figure 12

The above figure 12 is showing the number of confirmed cases in a day by Ebola in different years in Sierra Leone.

In 2014 the 50<sup>th</sup> percentile of data is around 4000 and the 75<sup>th</sup> percentile is more than 6000 in 2014 which means in the throughout the 2014, near around 4000 or more than 4000 confirmed cases were occurring in a day 50% of the time. And similarly

25 % of the time there were more than 6000 confirmed cases were getting in a day Sierra Leone.

These numbers are so much high in 2015 that 50% of the time there were near around 8500 confirmed cases in a day.



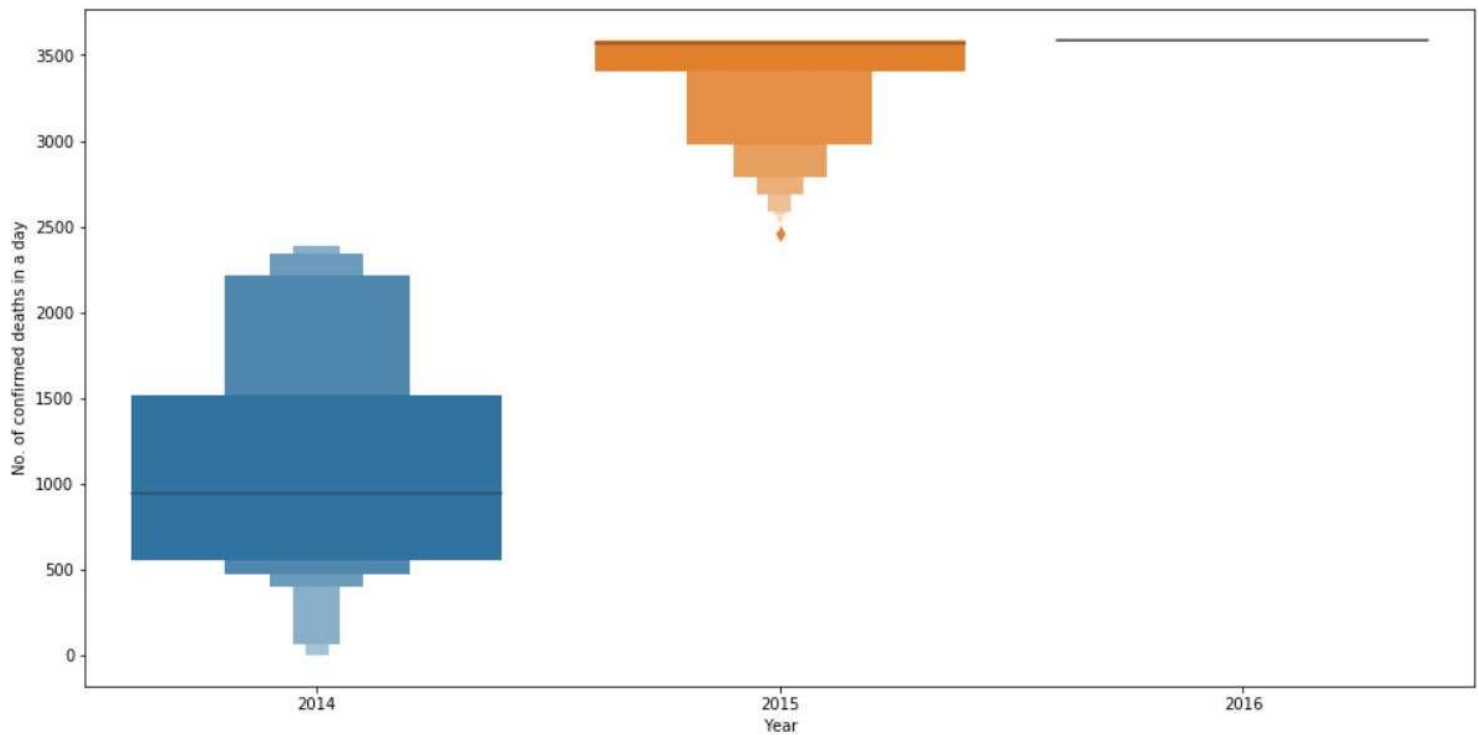


Figure 13

The above figure 13 is showing the number of confirmed deaths in a day by Ebola in different years in Sierra Leone.

In 2014 the 50<sup>th</sup> percentile of data is more than 900 and the 75<sup>th</sup> percentile is around 1500 in 2014 which means in the throughout the 2014, near around 1000 or more than 900 confirmed deaths were occurring in a day 50% of the time. And similarly

25 % of the time there were more than or equal to 1500 confirmed deaths were getting in a day Sierra Leone.

These numbers are so much high in 2015 that 50% of the time there were near around 3500 confirmed deaths in a day.

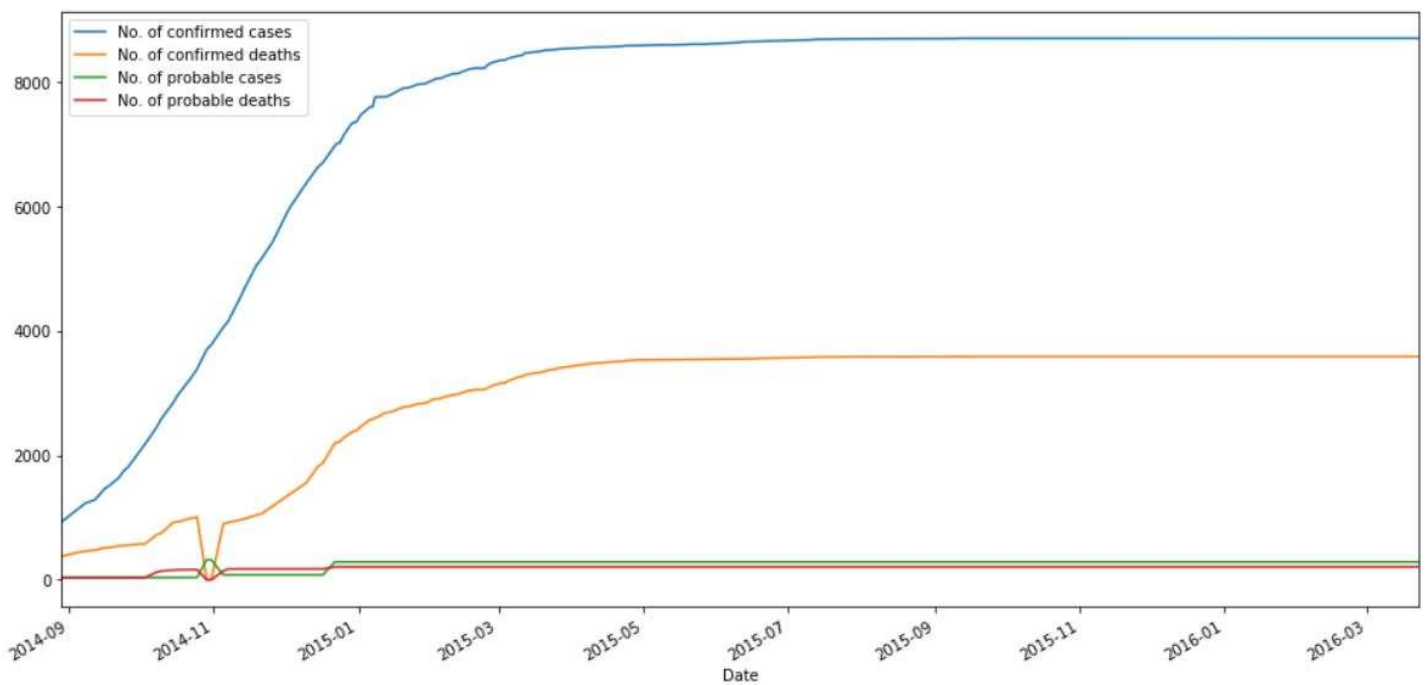


Figure 14

This above graph is showing the time line from 2014-2016 with different lines showing number of confirmed cases and deaths and number of probable cases and deaths.

In the line graph of number of confirmed deaths there is an outlier on the point near 2014-11 because of missing data being replaced with 0 value. Else whole graph is very clear.

We can say that number of confirmed cases were increasing from 2014 to 2016 and takes a constant rate of growth. We also know that the fatality rate of Ebola in Sierra Leone was 39.0810 % and that difference can be seen in the line of number of confirmed cases and number of confirmed deaths.

---

# BIBLIOGRAPHY

- The Ebola 2014-2016 Dataset was collected from Kaggle.

The website address is:

<https://www.kaggle.com/imdevskp/ebola-outbreak-20142016-complete-dataset>

- World Health Organization - “Case definition recommendations for Ebola or Marburg Virus Diseases”

<https://www.who.int/csr/resources/publications/ebola/ebola-case-definition-contact-en.pdf>

- Information on Guinea’s climate:

<https://www.worldtravelguide.net/guides/africa/guinea/weather-climate-geography/>

- Information on Sierra Leone’s climate:

<https://www.worldtravelguide.net/guides/africa/sierra-leone/weather-climate-geography/>

---

