Descriptive Analysis Report

Ebola Outbreak in Sierra Leone

Introduction

According to WHO, the 2014–2016 outbreak in West Africa was the largest Ebola outbreak since the virus was first discovered in 1976 and it is the seventh outbreak of Ebola Virus Disease since its discovery. There were more cases and deaths in this outbreak than all others combined. It started in Guinea then quickly spread to neighbouring countries Sierra Leone and Liberia. By July 2014, it had reached the capital cities of these three countries and in August 2014, WHO declared the outbreak a Public Health Emergency of International Concern.

Methodology

This report presents a descriptive analysis of the dataset which reported the state of the Ebola outbreak in Sierra Loene, within a 3-month period from the 18th of May 2014 to 17th July 2014, obtained from ekoanalytics. The data set consist of 7 columns and 201 observations. I worked on cleaning the data set, by replacing null value in the age column with its median value, also ensuring that the columns with wrong format are converted to appropriate format.

Analysis and Result

Starting with understanding the age demography of this dataset, this data reveals that the youngest person affected is female, aged 1.8 from the Kailahun district. It was a suspected case.

ID	Age	Sex	Status	Date Symptoms Started	Date of Sample	District

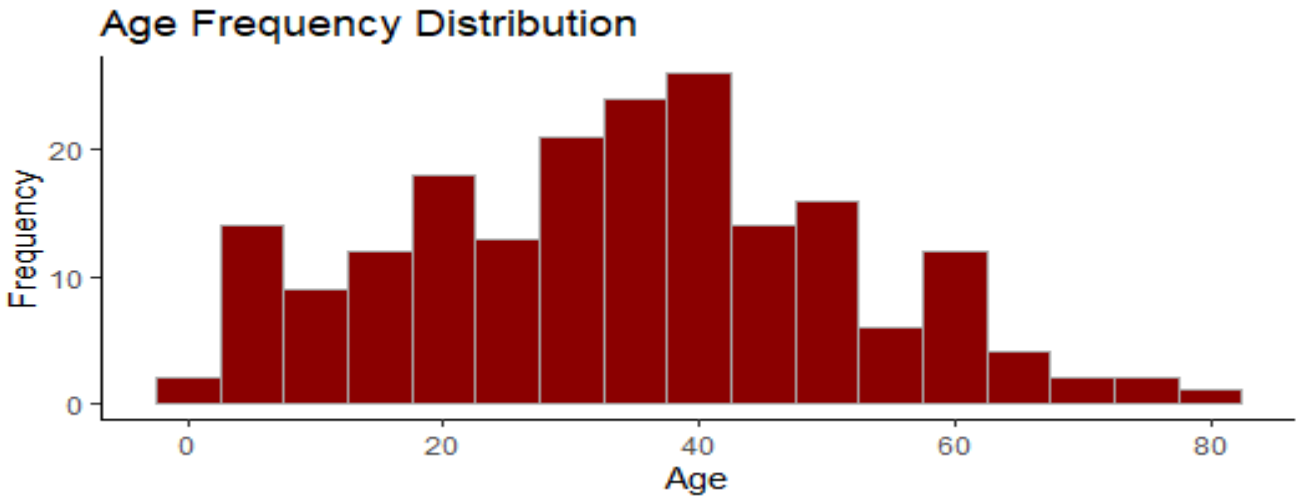
31	1.8	F	Suspected	2014-05-27	2014-06-01	Kailahum
82	80	F	Confirmed	2014-06-08	2014-06-13	Kailahum

While the oldest person aged 80, female was a confirmed case from the same district of Kailahun.

The earliest discovery of a symptomatic patient, by this data set was on the 18th of May, 2014, also from the Kailahum district.

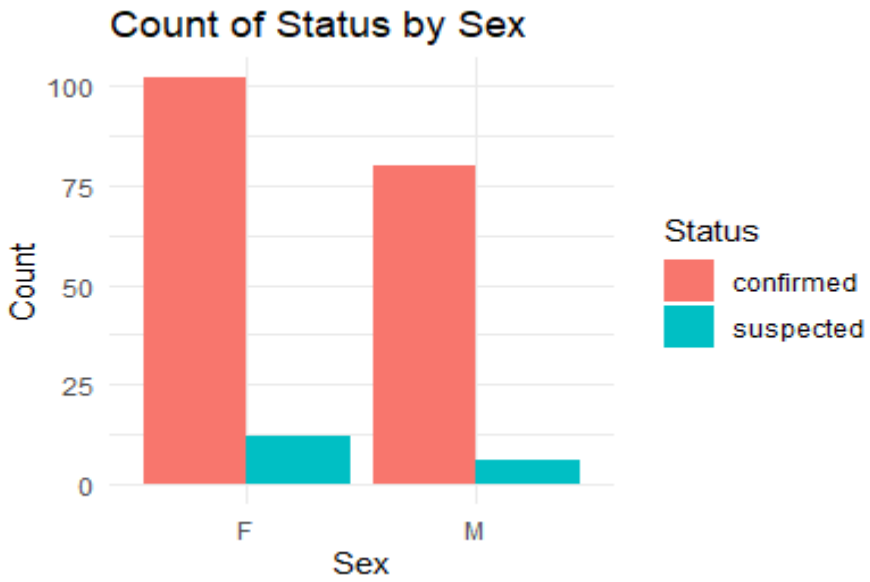
ID	Age	Sex	Status	Date Symptoms Started	Date of Sample	District
1	20	F	Suspected	2014-05-18	2014-05-27	Kailahum

The graph below shows the age distribution frequency of these sample data.



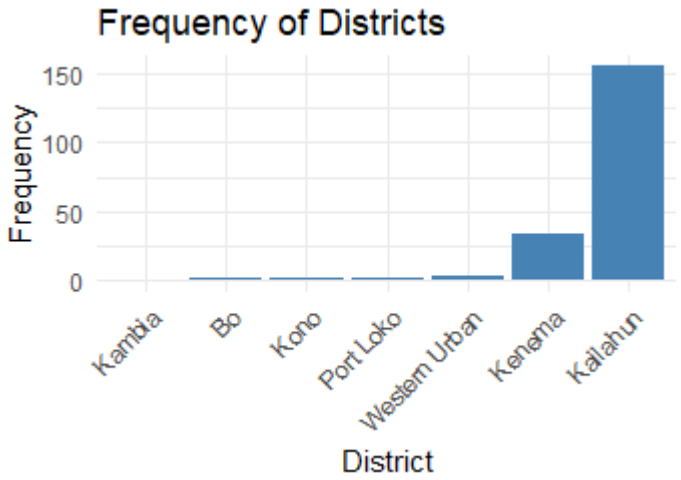
The average age of this dataset is 33.8, median age is 35 and age with the highest frequency, that is the mode is 35. The standard deviation is 17.3. However, the average age of the female segment is 34.07 while that of male is 33.5.

Looking at how the genders are affected by the disease; this sample data reveals that there are 114 (57%) female cases and 86(43%) male cases. Total confirmed cases in this dataset are 82(91%) while 18 (9%) are suspected cases, with the female gender representing the largest cases of both confirmed and suspected cases. This is confirmed by the graph and table below.



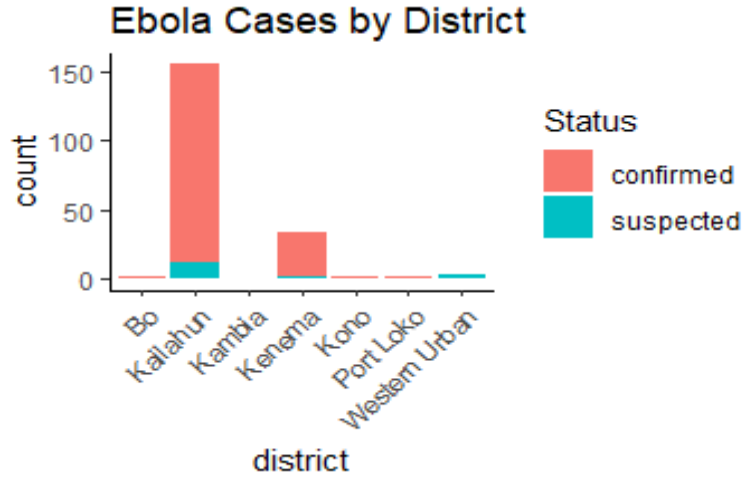
SEX	CONFIRMED	SUSPECTED	TOTAL
F	102 (51%)	12 (6%)	114 (57%)
M	80 (40%)	6 (3%)	86 (43%)
TOTAL	182 (91%)	18 (9%)	200 (100%)

Let's look at how the districts were affected by the Ebola disease. The graph below shows that the Kalihaun district had the highest cases of occurrence, leading at 78% of cases.



S/n	District	Count	Percent
1	Kailahun	155	78%
2	Kenema	34	17%
3	Western Urban	4	2%
4	Bo	2	1%
5	Kono	2	1%
6	Port Loko	2	1%
7	Kambia	1	0%

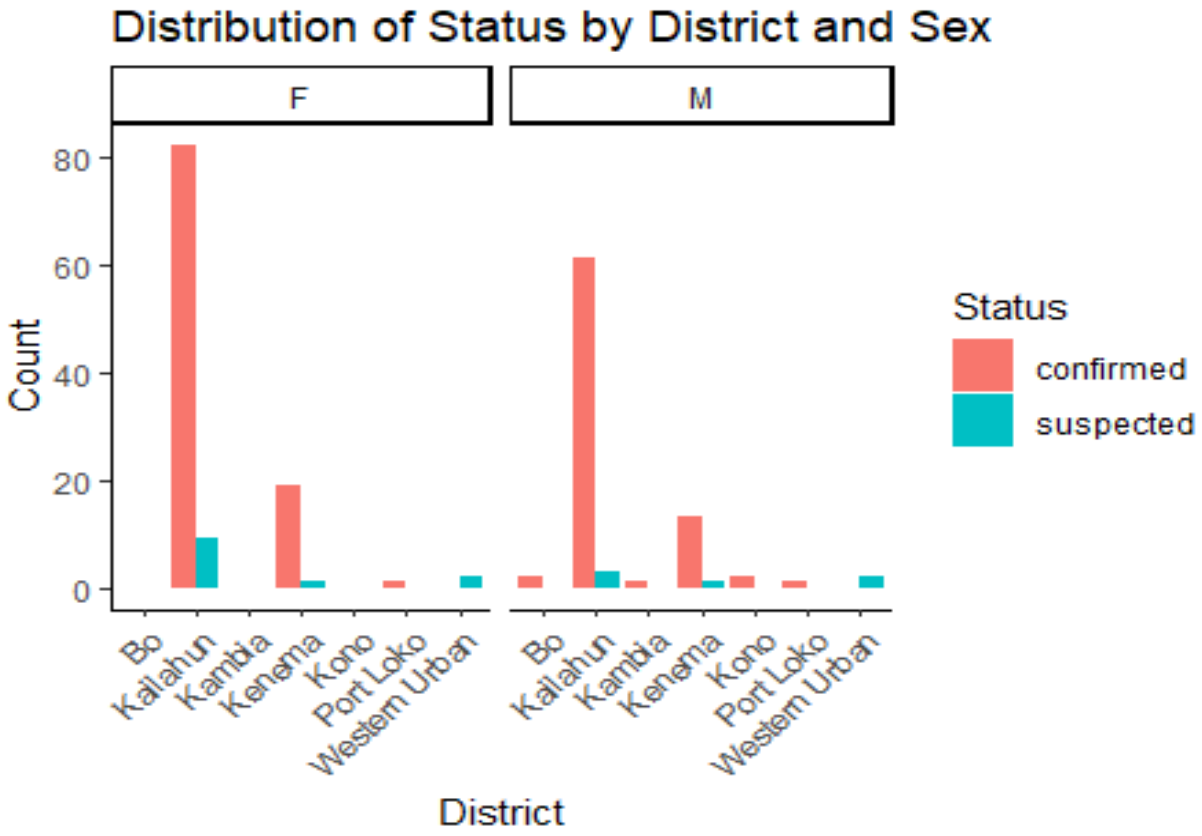
Consequently, the Kailahun district had the highest number of confirmed cases representing 71.5% (143) of the total confirmed cases, followed by Kenema at 16% (32). The remaining districts had around 1% cases each except, Western Urban who did not report any confirmed case, only 4 suspected cases as show in the graph below.



S/N	DISTRICT	CONFIRMED	SUSPECTED
1	Bo	2 (1%)	0 (0%)
2	Kailahun	143 (72%)	12 (6%)
3	Kambia	1 (0%)	0 (0%)
4	Kenema	32 (16%)	2 (1%)
5	Kono	2 (1%)	0 (0%)
6	Port loko	2 (1%)	0 (0%)
7	Western urban	0 (0%)	4 (2%)
	Total	182 (91%)	18 (9%)

A further breakdown of the cases reported from the districts according to gender and status of the cases is seen in the graph below.

The graph shows that Kailahun, which had the highest number of cases, is the district with the highest number of male and female confirmed cases, followed by Kenema district. In both districts, the female was the most affected.



Summary of Findings

- Geographical Impact:
  The Kailahun district had the highest number of confirmed and suspected Ebola cases, followed by Kenema. This suggests that these regions were the epicentres of the outbreak during the early stages.
- Gender Disparity:
  The female gender was more affected by the Ebola outbreak than the males, both in the confirmed and suspected cases. This could be due to various socio-cultural factors such as caregiving roles that women often assume, thereby increasing their exposure to the virus
- Age Demographics:
  The Ebola outbreak disproportionately affected young people. This trend could be because of more social interactions and activities found among the young populace, which must have increased their risk exposure.
- Data limitation on Mortality Rate:
  Considering these were the early stages of the outbreak, data did not contain if the confirmed cases led to death incidences, posing the challenge of assessing the full impact of the outbreak on mortality rate.

Recommendation

- The government, local authorities and community leaders have to intensity efforts at promoting hygiene and sanitation practices such as handwashing to reduce the risk of quick transmission of future occurrences
- Vaccination and treatment should be prioritized for the mostly affect regions and venerable populations
- Develop and promote public health campaigns that address the specific risks faced by women.
- Create educational programs targeting young people to raise awareness about the disease and preventive measures, encouraging adherence to safety protocols

References

Data source: https://ekoanalytics.net/data-catalogue.html

World Health Organisation (WHO)

https://github.com/aladbukky/Data-Science-Projects-at-SAIL