

Andrew Halpern

Ryan Campbell

Jiabin Dong

## **Ebola Data & Infographics**

*CS/INFO 3300 - Project 1*

For our project we used data from the Humanitarian Data Exchange Website. We used four different CSV files and converted them into JSON files that each helped show statistics on the impact Ebola had in the world. We used multiple datasets to illustrate the different effects of Ebola including spread by location, number of people infected over time, ratio of number of deaths to number of cases of Ebola by country, and how much relief areas were receiving in terms of food assistance. We also used an external data source to map ISO-3166 country codes to the respective country names and abbreviations for use in mapping and display of data.

### **Total Number of Cases by Region**

This map shows the total number of Ebola cases by physical location on a global map. The data source (GeneralCasesDeaths.json) provided weekly updates for the confirmed, probable, and suspected Ebola cases in all countries that may be exposed to the virus. For the map, we used the most recent data point for each country, from February 23, 2015. Since the data is very skewed, with some countries having less than 20 cases and others having thousands, we used a power scale with exponent 0.35 to map case count to a color. This allows the graph to define both ends of the data while maintaining a sufficient contrast to emphasize the wide gap between them.

By seeing the differences in colors in the map, the viewer can see the tight concentration of Ebola cases in a very small part of western Africa. Despite the fact that the media makes the issue out to be a pandemic, the map clearly shows how few cases exist outside of Sierra Leone, Guinea, and Liberia. These countries make up a very small land area, and even most of Africa is free of the disease. Though other countries have documented cases, there are most likely far fewer than the media leads us to believe.

### **Number of Cases Over Time**

This graph illustrates the number of Ebola cases by country spanning from August 29, 2014 through February 23, 2015 using the GeneralCasesDeaths.json data. It shows how the spread of Ebola has increased or decreased at various times in the countries that have been exposed to the virus. As with the map, the spread of data made it challenging to show trends in all of the countries. Thus, we used a square root scale on the Y axis. This allows the trends to be apparent in the graph without making all of the low data points look like a flat line at  $Y=0$ .

The figure helps to further demonstrate the tight concentration of Ebola cases in only three small African countries. By viewing the data over time, the observer can also see how the countries that seem to get the most media attention: the US, Spain, and the UK, have not seen any increase in Ebola cases. Each had a few cases all at once, but the virus never spread after that. On the other hand, the graph also demonstrates that the spread of the virus is slowing. By the end of 2014, the rate of increase of cumulative Ebola cases decreased to almost zero in most countries. Sierra Leone, Guinea, and Liberia are still exhibiting signs of viral spread, but at a far slower rate than before. It seems that our preventative measures and treatment have been relatively successful!

### **Death Count vs. Case Count by Country**

For this infographic it was important to only display the total death count and case count as of February 23, 2015. The dataset (GeneralCasesDeaths.json) showed each number of confirmed, probable, and suspected cases of Ebola for each week, but this infographic goal was to illustrate the total number of cases while the line graph showed how the number of cases developed over time. By breaking up the data this way, it was easier to see a difference between the ratio of death count and case count by region.

Each red circle indicates number of Ebola cases for that particular region and each black circle inside represents the number of Ebola related deaths by that region. The radius of each circle corresponds to the number of cases or deaths, respectively. The radius of each circle is obtained by using the square root of the death count and case count to make it proportional to the area of each circle. This allows users to see the proportion of deaths to number of cases by region. The labels for each country are printed line up vertically with the center of each respective circle and the calculations for the ratio are shown below in the corresponding colors. The x-axis is organized in ascending case counts and are spaced with the first  $\frac{3}{4}$  of the data go to the first  $\frac{1}{4}$  of the xScale to minimize overlapping between the extremely large case count countries and the really small case count countries. Countries with a radius of less than 5px were scaled up by 4x because the circles were too small to see.

This infographic helps visualize how the Western countries were at an incredibly less risk of an Ebola impact than Sierra Leone, Liberia, and Guinea. There was only one confirmed death in a western country and this graphic help illustrate how dangerous and lethal Ebola was in West Africa compared to the rest of the world. It's surprising to see how little impact

Ebola has on the western world given the amount of media coverage the news put on possible outbreaks in the United States.

### **Weekly World Food Program Assistance by Country**

When people get sick, food is crucial to their recovery especially in impoverished areas of the world where resources are slim. Thus, it is essential to analyze the food assistance as a gauge for our global responsiveness during the Ebola outbreak .

This graphic shows the number of people receive food by country during Ebola outbreak (from 09/23/2015 to 02/12/2015). The data is collected from "FoodAssistance.json". Since Ebola spreads mostly among impoverished west African countries, the food assistance data only includes the number of people who received food in Guinea, Liberia and Sierra Leone. To display the number more poignantly, we use a gradient green bar to indicate the growing number of people receiving food. The bright green contrasts the dark red and blacks to show that countries most affected are at least receiving a lot of aid from the rest of the world. For the timeline variable, the "Date" is mapped to the timeline using "time.scale" method. Since the report points are not exactly evenly distributed, the green rectangles' width vary according to the report time gap. For the color variable, color gradient from gray to green is mapped from the number of people receive food.

In this scaling map procedure, the linear map is used because the green color bar is able to show up the weekly data quite distinctively. In the graph, each green rectangle starts from the one report time point and ends at the next report time point horizontally, filled with the mapped color. To show the data more directly, we limit the number range from least number of people receive food (42503) at the earliest report time point (09/23/2015) to the largest number of people receive food (1644264) at the very last report time (02/12/2015).

As you can observe from the graph, Sierra Leone has the highest number of people receiving food since the number of Ebola cases and deaths also turns out to be the highest among the countries, especially when it comes to Feb 2015. But at the first report point, Liberia gets the most food assistance because people there are more likely to get Ebola at that time. In comparison to the number of cases over time graph, the food assistance data graph corresponds to an increase in food assistance with more cases of Ebola in each country.

### **Conclusion**

This data shows the viewer that the global response to the Ebola outbreak has been swift and effective. We see that food assistance begins to increase very quickly in all of the countries, as all bars beginning darkening in color before the end of 2014. In only a few months, food was provided to over 1.6 million people in Sierra Leone--up from 44,000 in August 2014. Similarly in Liberia, food assistance ramped up by a factor of ten from about 100,000 people to about 1,000,000 people. The areas plagued with Ebola the worst were provided with a tremendous amount of external food resources.

In addition, we see that Western countries do not have such an immediate risk as the media suggests; the number of cases in those countries is very low and is not increasing with low deaths rates as well. Though the Ebola virus poses a threat to the world as a potential pandemic, the data suggests that our responsiveness has slowed its spread and risk outside of the borders of a few small countries in western Africa.

## **List of Data Sources**

1. ISO-3166 Country Information (countryCodes.json)
  - a. <https://github.com/luke/ISO-3166-Countries-with-Regional-Codes/tree/master/slim-2>
2. Weekly cumulative total of Ebola cases by Country (GeneralCasesDeaths.json)
  - a. <https://data.hdx.rwlab.org/dataset/ebola-cases-2014>
3. World Map(world-50m.json)
  - a. <https://github.com/mboystock/topojson/tree/master/examples>
4. Weekly number of people receive food in Guinea, Liberia and Sierra Leone(FoodAssistance.json)
  - a. <https://data.hdx.rwlab.org/dataset/wfp-indicators>