

## Geography 418: Assignment 1

Due: September 24, 2018

Devin Hewett V00821219

The Syrian Refugee Crisis is a mass migration of people trying to flee the Syrian civil war, has resulted in the deaths of thousands. People have lost their lives while travelling or after being relocated in a different region of the world. The location of these refugees are tracked as events of people becoming missing, injured and/or deceased; this data can be analyzed to observe geospatial and statistical trends.

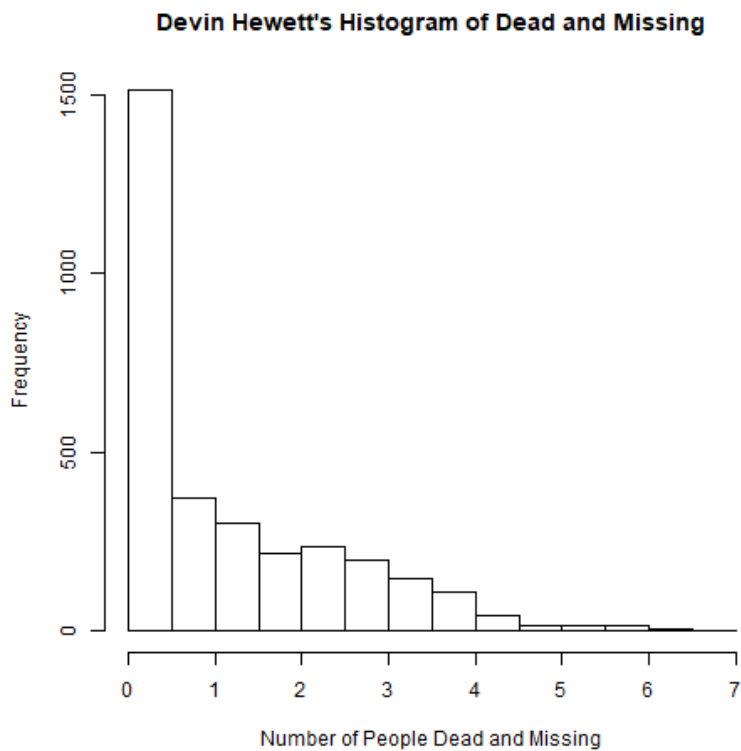
With the use of descriptive statistics, the data can be summarized to compare and describe different aspects of a large data set in order to gain a better visualization and understand of the data. Which means the large population can be summarized to understand where and when these events are happening and identify problem areas and respond to them. Also, with the descriptive statistics means, median, mode and standard deviation gives a idea of the distribution curve of the event data how it is skewed, how many people on average are involved with an event and the most common occurrence of people.

samples	means	sd	mode	medain
Population	10.90721	38.28385	1	2
2015	20.62371	75.67299	1	4

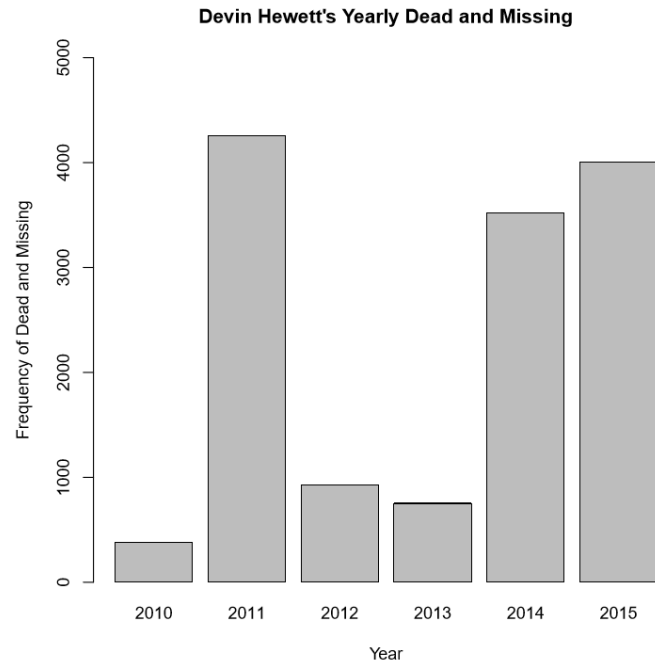
**Table 1:** Summary of the descriptive statistics of entire population of events and those that occurred in 2015.

The total population compared with 2015 descriptive statistics of events where refugees are dead or missing, the data can be observed to be positively skewed as the median is left of the mean in both cases. The standard deviation in both cases are large, meaning that the data has a very large spread and the mode shows that 1 is the most common value. This shows that a lot of frequency is in the lower range of number of people, however because of the large standard deviation there are large values which pull out the distribution. But comparing the

total population verse 2015; 2015 the mean is 20.6 almost twice as high as than the total, and median is double as well as the SD. This shows that 2015, the events that occurred involved a lot more people, with an even larger spread of data than the total.

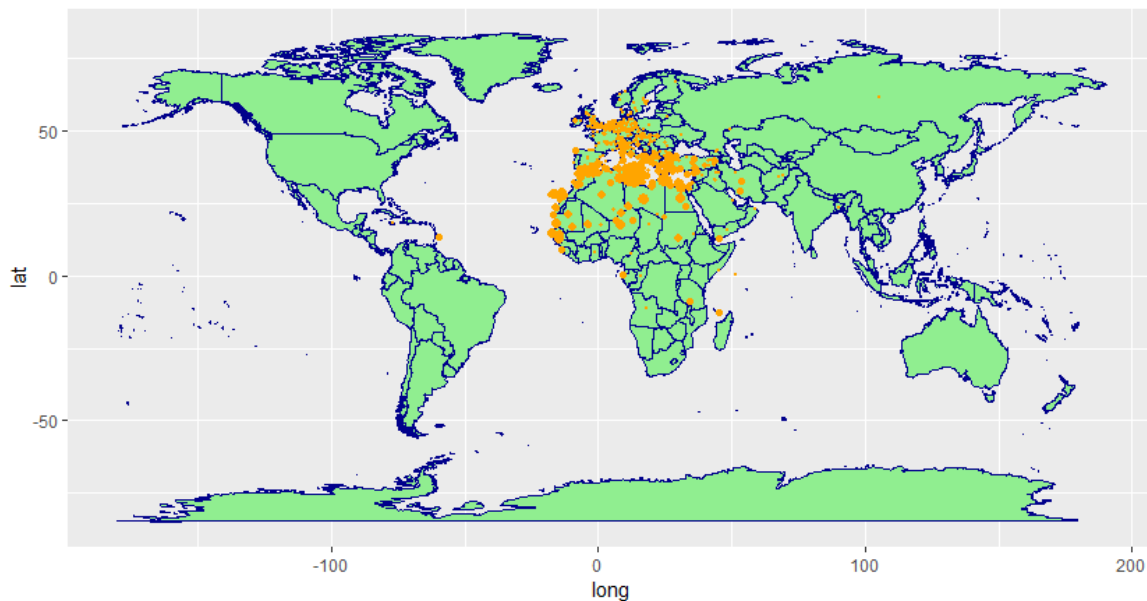


**Graph 1:** Histogram of Syrian Refugees events where people have gone dead and missing.



**Graph 2:** Bar graph of the number of people that went dead or missing by year.

The two graphs above show a visual representation of the distribution of the data as graph 1 shows the large frequency at low values as the distribution curve behaves asymptotically. Graph 2 shows the total of the frequency dead and missing by year, from this graph it is apparent that the worst years are by far 2011, 2015, 2014 respectively.



**Image 1:** Distribution of location of where missing and death events occur relative to the world.

Image 1 is a geo-spatial representation of the data as a cluster of data points mapped on the globe centred around the Mediterranean the points appear most dense, which makes sense as Syria is located on the east side of the sea. Areas of problems can be easily located on image 1, with only a handful of outliers lay outside of Europe, Africa and the Middle East.

To conclude, viewing descriptive statistics gives a good understanding of the distribution of the Syrian refugee data, then to understand further a graph of the data dead and missing frequency and then sorting by year to see a yearly trend. Lastly, adding a map representation best represents the data geo-spatially, to see where the data is distributed globally. Syrian refugees are being displaced and spread on a global scale, it is ideal to keep track of which areas are safe and where these tragic events occur to better understand how to deal with this crisis and future crises like it.