Aakash Tandel
General Assembly DC
Data Science Immersive 5
August 27th 2017

Global Terrorism Analysis

**Objective**:

The objective of this project was to gain an understanding of terrorism using Bayesian inference. This analysis specifically looked at the Phillipines in Southeast Asia. My working knowledge on terrorism in Southeast Asia is limited. But this line of inquiry was sparked by the current chaotic situation in Marawi City, Philippines where Philippine Government security forces fight ISIS affiliated militants known as Abu Sayyaf and Maute.

The Global Terrorism Database contains over 170,000 cases of terrorism from around the globe. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) maintains this database. START is headquartered at the University of Maryland. At the time of the submission of this project, 1970 to 2016 was accounted for. The database contains over one hundred and thirty variables including geographic location, type of attack, perpetrators, targets, outcomes, number of fatalities, and motivation of perpetrators. This was a very robust dataset but it does contain a fair amount of missing information. In particular, 1993 data was lost. One additional goal of this analysis was to impute the number of bombings that occurred in 1993.

**Exploratory Analysis**:

The majority of exploratory data analysis was done through the use of visualizations. These visualizations can be found in the appendix of this document. To begin, the number of armed assault and bombing terror attacks dwarf the other methods of attacks (Graph 1). Hijacking (the method used by al-Qaeda terrorists in the September 11th 2001 attacks), assassinations, and hostage situations are far less numerous. Secondly, the number of people killed, grouped by region, showed the Middle East and South Asia as the two hot zones for terror activity from 1970-2016 (Graph 2). Media coverage of the Iraq and Afghan Wars has offered citizens a deeper look into the terrorist activities in the Middle East. But South Asia terrorist groups rarely make news in the United States. I decided to take a closer look into Southeast Asia (the Philippines specifically) because of my lack of awareness on the terrorism issues of that region. Bombings resulted in a high number of fatalities and Graph 3 shows a drastic increase in the number of bombs used by terrorist groups since 2010. The Philippines have experienced this global trend as well. This noticeable spike in bombings in 2010 sparked the Bayesian analysis.

**Bayesian Analysis**:

The goal of the Bayesian analysis was to determine if there was a statistical between the number of terrorist bombings in the Philippines from pre-2010 and post-2010. From a global view, bombings increased drastically since 2010. The prior probability of terror attacks in the Philippines was a simple average the number of bombings between 1970 and 2009. This turned out to be around 29 bombings per year.

The average of bombings per year between 1970 and 2009 was a logical prior assumption for Bayesian statistics because this would have been the only information available at the onset of 2009. If we were making the analysis in 2010, we would want to see whether or not one year of increased terror attacks was statistically different than the previous forty years.

The full breakdown of the Markov Chain Monte Carlo method can be found in the Jupyter Notebook associated with this project. The MCMC method revealed there was a statistical difference between pre-2010 bombings in the Philippines and post-2010 bombings. The credible intervals between the two groups did not overlap and the difference in means was statistically different than zero. Overall, this analysis isn't very surprising because of the global trend in increased terrorism starting in the early 2000's. An industry expert with prior knowledge of terrorism could more readily utilize MCMC methods in order to derive key insights into the world of terrorism analysis.

**Missing Values**:

The information from 1993 was lost. In order to develop an estimate for the number of terror attacks (attacktype1 in the Codebook), time series analysis was utilized. I used the Dicky-Fuller test to determine whether or not a unit root was present in the data. This test for stationarity concluded the data was non-stationary. Graph 4 shows a rolling mean of 12 months and Graph 5 shows a quarterly rolling mean. Both of these graphs illustrate non-stationary processes.
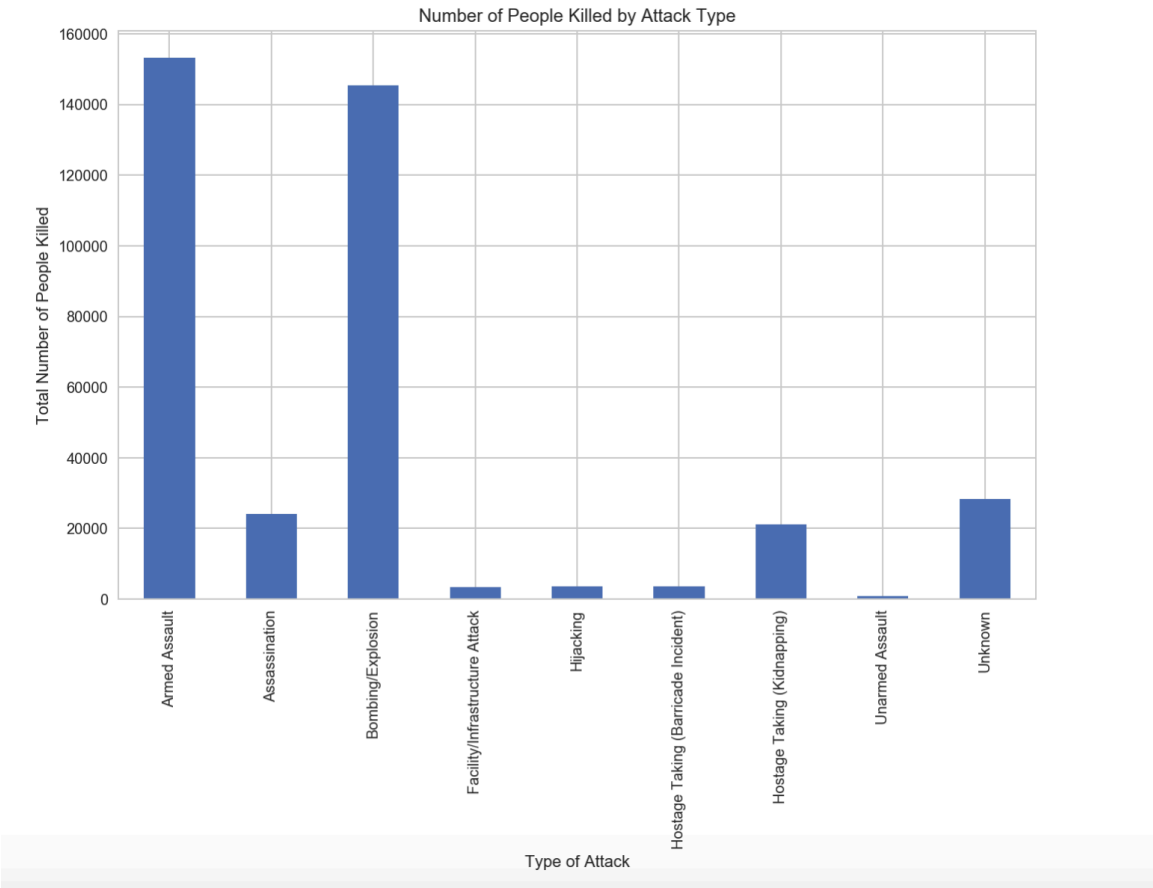
I made the decision to simply take the average number of bombings between 1992 and 1994 in order to estimate the number of bombings for 1993. The quarterly rolling mean for 1992 was 1,819 total bombings and 1,626 bombings in 1994. I estimated around 1,720 bombings for the year of 1993. There may be more elegant ways of estimating the number of total bombings in 1993, but this option seemed most effective out of the box.
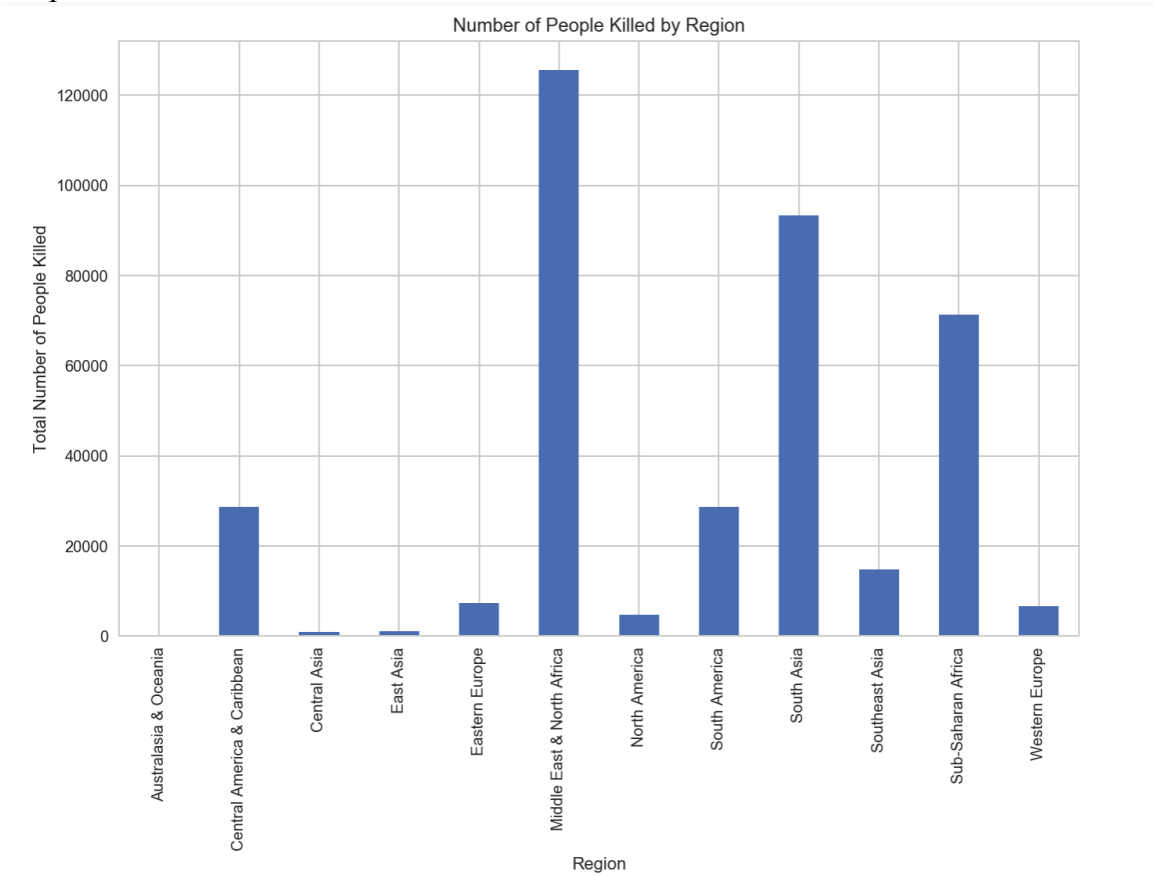
**Conclusion**:

The world has seen an increase in terrorist bombings since 2010. It fills our news and plagues many of the worlds' citizenry. The terrorist groups Abu Sayyaf (who pledged allegiance to ISIS in 2014) and Mauta group currently fight Philippine Government security forces in Marawi City, Philippines. This analysis sought to shed a little light into the crisis brewing in the Philippines. It seems that the number of terror bombings has increased across the globe since 2010 and increased markedly in the Philippines. Overall, it is with the help of START that we can learn more about terrorist activities and develop proper strategies in order to counteract their threat.

**Appendix**:

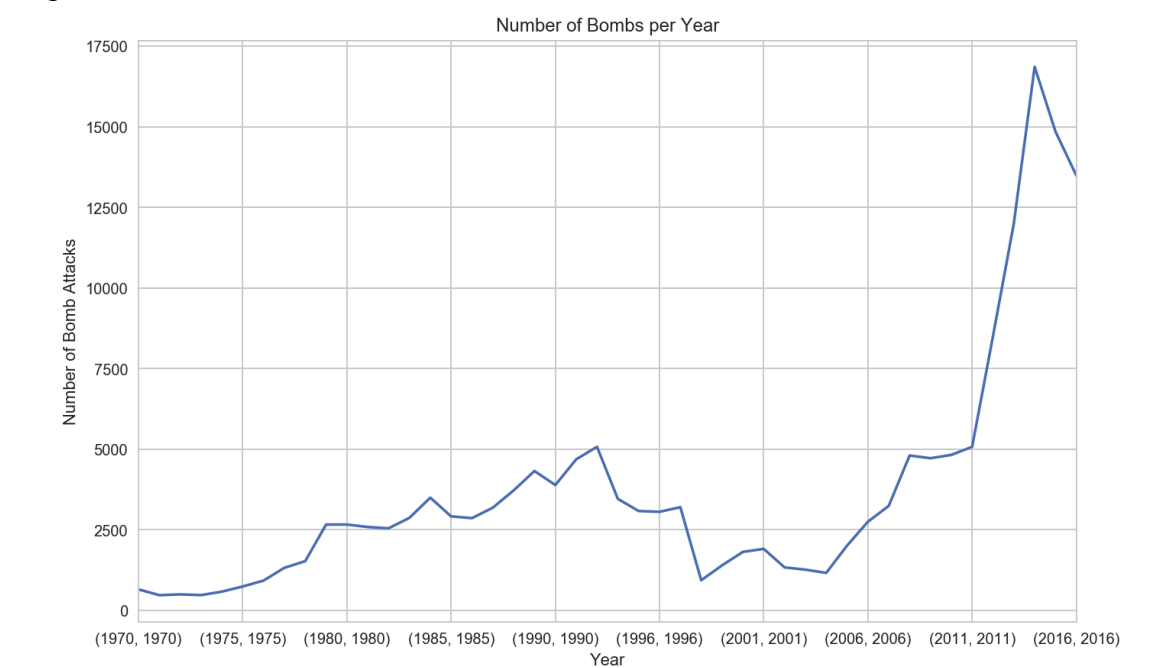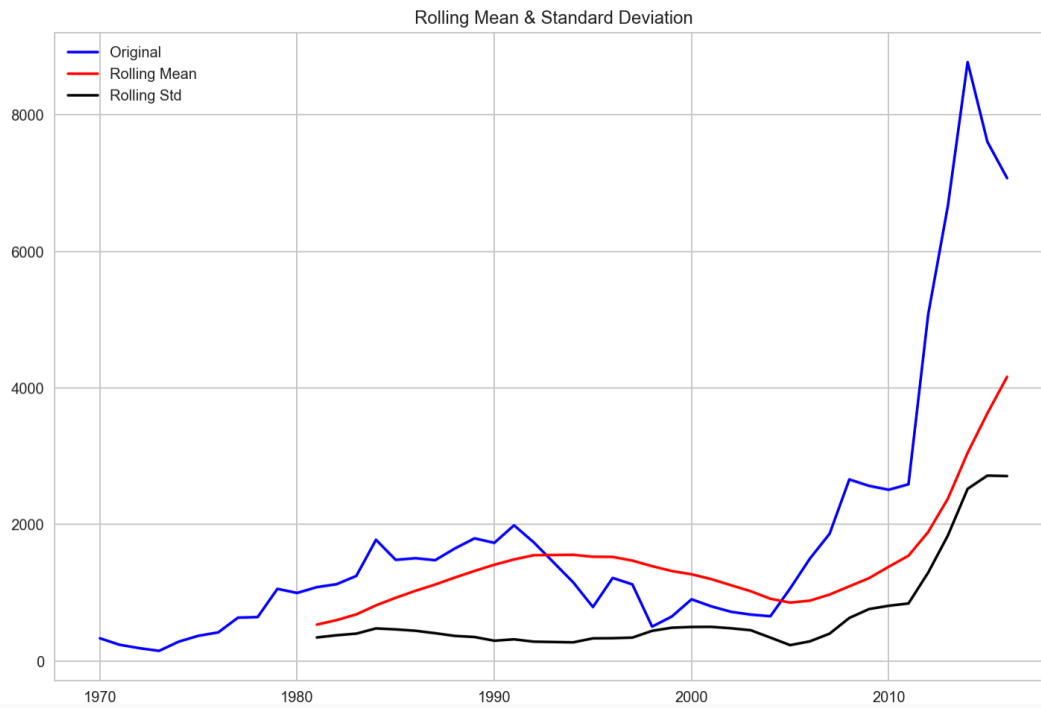Graph 1:

Number of People Killed by Attack Type

Graph 2:

**Number of People Killed by Region**



Graph 3:

**Number of Bombs per Year**



Graph 4:

Rolling Mean & Standard Deviation

Graph 5:



Rolling Mean & Standard Deviation