Data Report

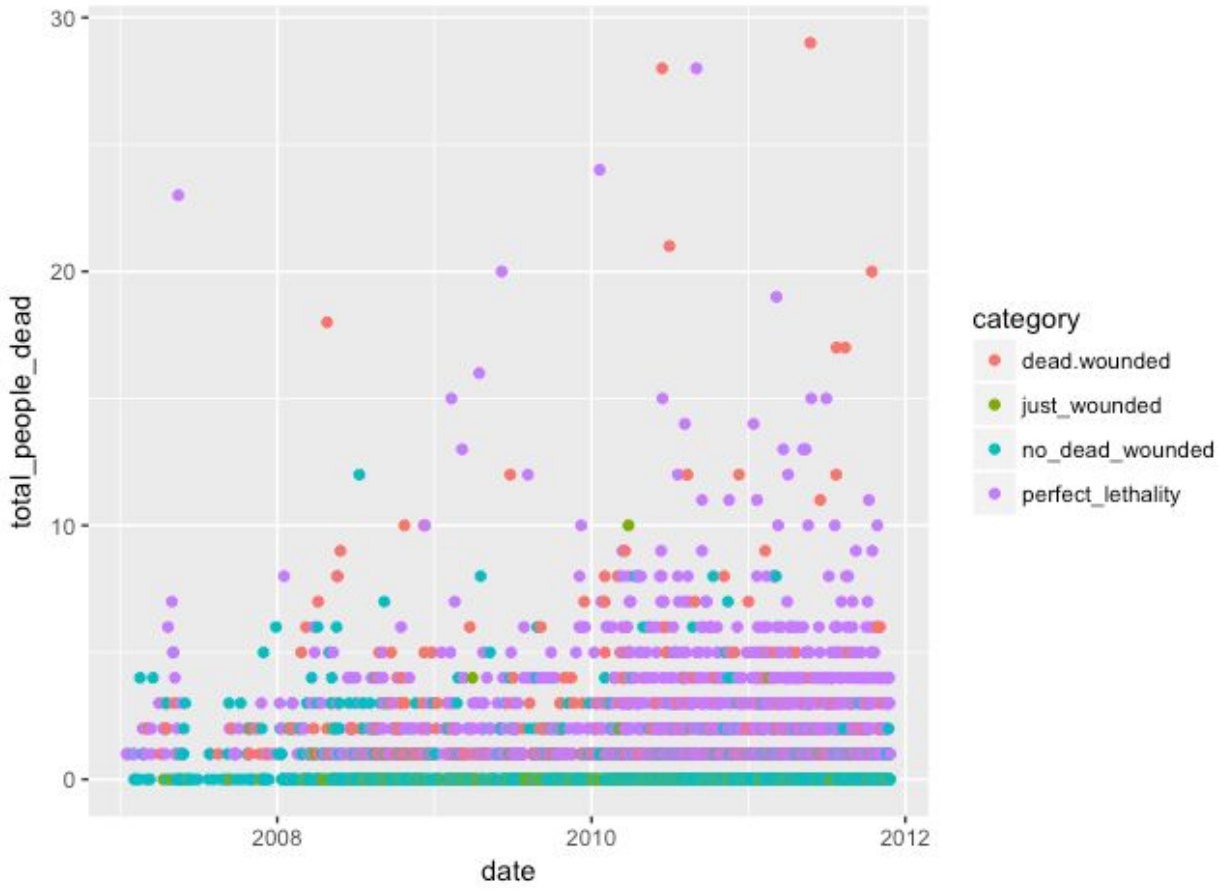
I calculated the percent of perfect lethalities for each group listed by calculating the sum of those who died in perfect lethality events over the sum of all who died in the designated events.

To get the sum of those who died in perfect lethality events, I had to filter for perfect lethality == 1 and then sum the GROUP_dead column. To calculate the total deaths, I added the numbers in the GROUP_dead column from the larger data set.

Percentage of Perfect Lethality

	Civilians	**Army**	**Navy**	**Police**	**All**
Confrontations	42.29%	Inf	47.37%	18.52%	64.92%
All Events	60.85%	Inf	50%	22.76%	57.83%

I could not replicate the 86.1 number. According to my calculations described above, the number was 42.29%. It's possible that I miscalculated, so I will further review my analysis to ensure that it is correct.



- Is this the right metric to look at? Why or why not?

I think there are a lot of metrics that need to be used in tandem to better understand the problem. Using the % of deaths that were perfect lethality is limited because we don't know how many deaths took place total. Was it 10,000 or was it 100? Those are very different numbers. It might be helpful to know the volume, frequency, timeline, etc on top of this number. Ideally the index would be very helpful too, but it is hard to calculate.

- What is the "lethality index" showing explicitly? What is it not showing? What is the definition assuming?

The lethality index is showing how many deaths happen for every wounded. It is not showing how difficult it is to calculate accurately. It also lacks the metrics mentioned above. The definition is assuming that it is a good indicator of what happens in every incidence.

- With the same available data, can you think of an alternative way to capture the same construct? Is it "better"?

- What additional information would you need to better understand the data?

I would like to better understand the location of these incidences. I imagine if these events are highly concentrated in a single state, then it changes the implications of the data. In addition, I'm curious what kind of crime is most connected with the killings. Is it robbery, drugs, gangs, etc? These metrics will help determine root cause, which will help with solutions for the data. We can analyze all we want but without actions based on the data, it's somewhat useless to know.

- What additional information could help you better capture the construct behind the "lethality index"

It would be helpful to compare to other countries to better understand the average or what is common. It would also be helpful to know the source of the killings. Are the criminals connected? Like I mentioned earlier, are they in the same area geographically? More demographic information would be helpful too. The data as is may reveal a problem but doesn't have other variables that would help solve it.

Overall, this challenge helped me see some gaps in my ability to apply information I've obtained by reading and watching in my data set. I've read a lot on charts and graphs, but struggled to find the right one's and the right metrics to place on the graph in a timely manner. I also struggled with applying the format of projects such as writing the readme, etc. To fill those gaps, I plan to take the weekend to work on creating charts and graphs and solidifying my knowledge of project format/git.