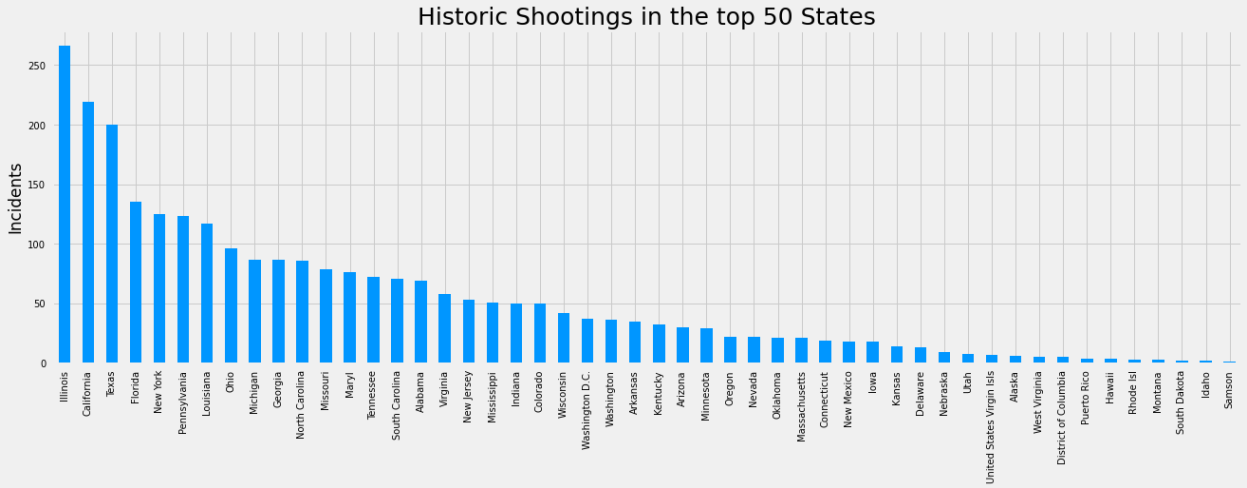
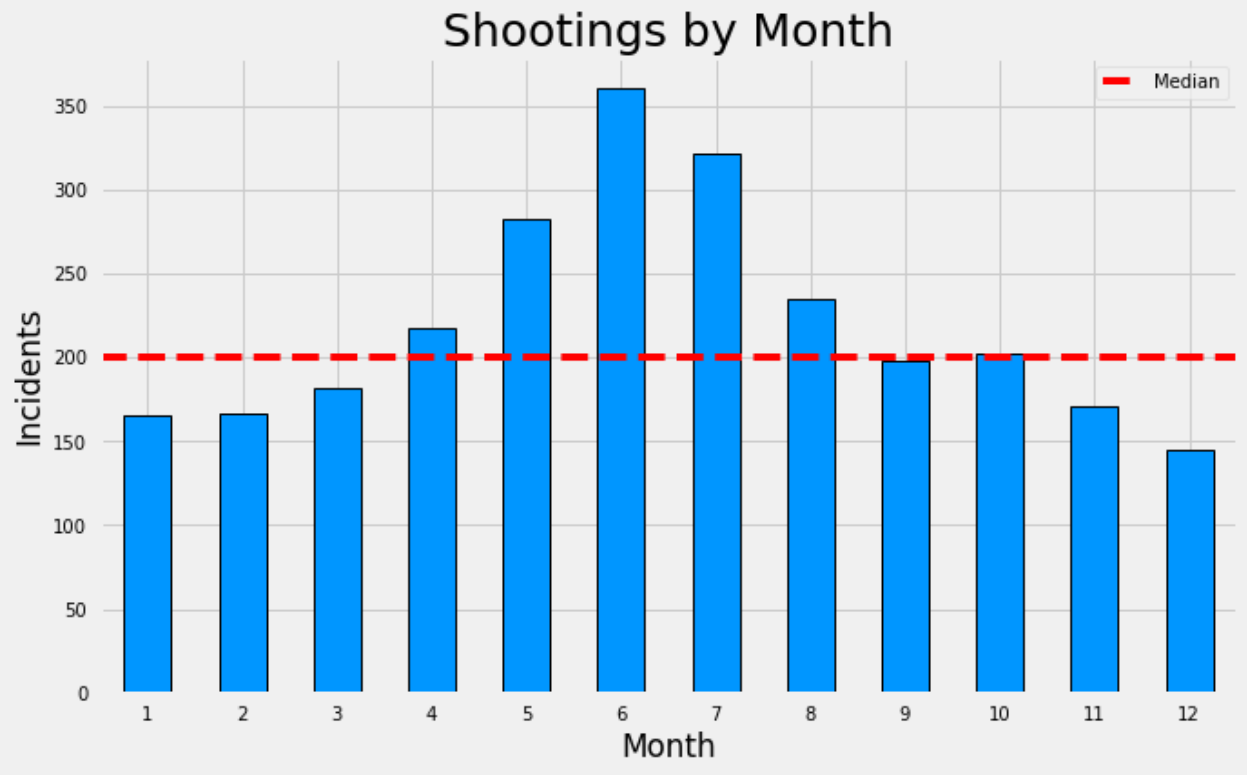
Gun control has been and will continue to be a profoundly divisive issue in American public discourse. My exploratory data analysis of a dataset reporting 2647 shootings in the US between 1924 and 2022 reveals peculiar patterns of American gun pandemic as well as sobering prospects for the threat of gun violence in the States. These discoveries, however, are likely to be ill-founded due to the poor quality of the dataset



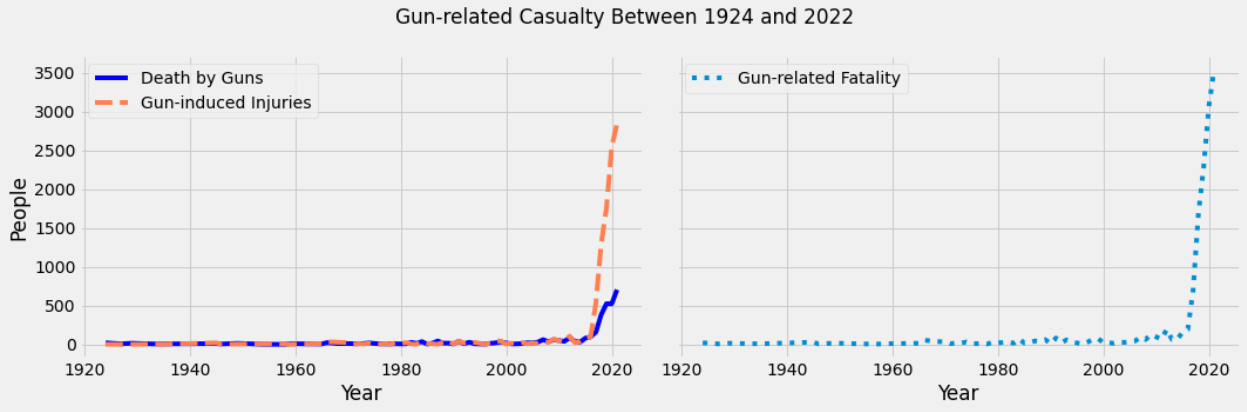
First, a quick examination of the dataset informs us of the ten states that, historically, have had the most mass shootings, which includes Illinois and California at the top, followed by Texas, Florida, New York and Pennsylvania with Lousisiana, Ohio, Michigan and Georgia at the bottom. These ten states form a much peculiar, unlikely collection as states with reportedly excellent gun control policy such as Illinois, California and New York top the chart. Nevertheless, the explanation for their abnormally high level of gun violence can easily be found in their high level of urban concentrations with metropolitan melting-pots such as Chicago, Las Vegas and New York City incurring disproportionately more violence than other states. The remaining states in the top 10, however, are not as surprising as these are the states that have the least gun control (Texas, Florida, Louisiana, Ohio, Georgia[1]). Therefore, a reasonable conjecture on the state of gun violence in the United States is that gun crimes concentrate in big cities, urban centers and states with weak gun laws.

[1] https://giffords.org/lawcenter/gun-laws/browse-gun-laws-by-state/

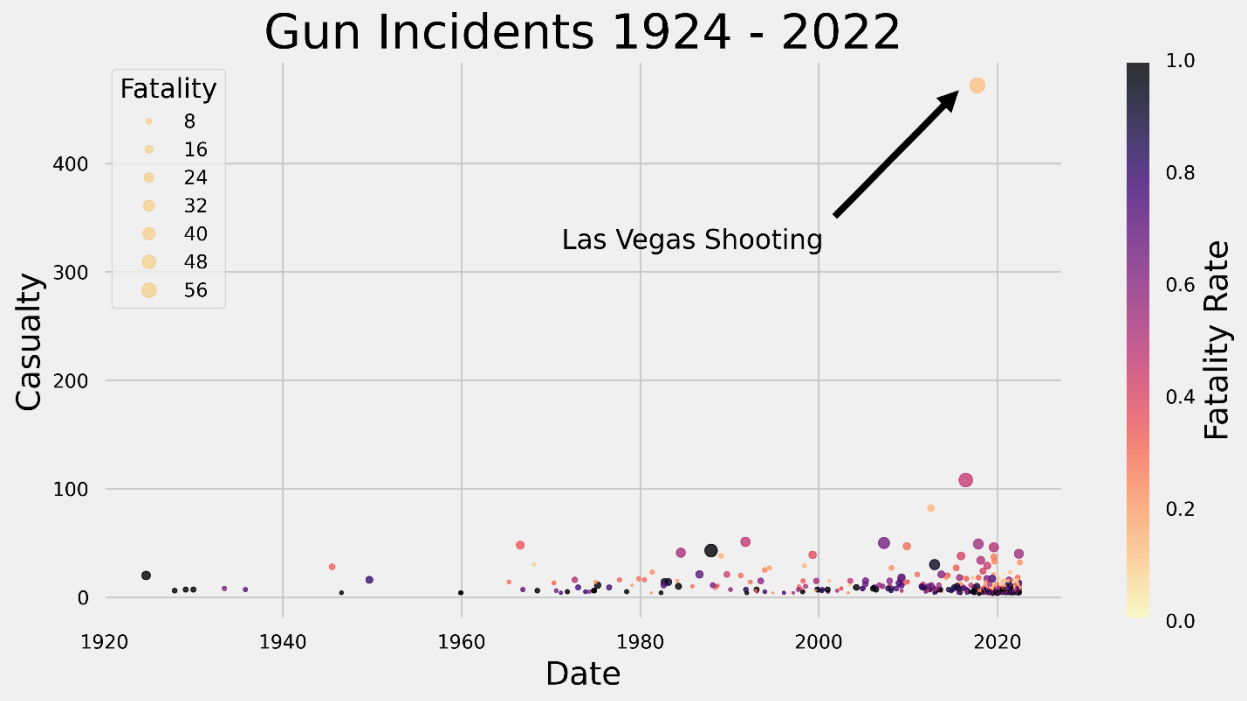


A second look at the dataset reveal an uneven distribution of shootings throughout the year. Historically, the most violent months gravitate toward the middle of the year with May, June and July reporting hundreds of cases more than the year median. This pattern aligns with a very interesting relationship between crime and heat, which allegedly arises because economically disadvantaged citizens who have no access to air-conditioned homes and buildings are forced to endure the temperature and violence of the street while the refuge for underprivileged students that the school system usually provides during the year is also taken away in summer[2]. In short, the heat of summer, coupled with the ceased operation of the school system, creates much moreS fuel and opportunities for violence.

[2] https://giffords.org/blog/2019/07/why-deadly-cycles-of-violence-spike-in-the-summer-blog/



On the other hand, a different approach to the dataset reports a steep rise in both the number of deaths and injuries, and consequently the casualty, by gun from 2017 onward.



This rise, as can be observed from this scatterplot, results from bewilderingly sharp increases in both the frequency and the deadliness of shootings with thousands of shootings cluttering together from around 2010 onward. This evidently signals a more violent future that looms ahead for the United States.

Although, as with any other data, these discoveries are inherently incomplete due to the nature of the data itself. In the case of this specific dataset, the Las Vegas outlier, I would suspect, should confounds much of our perception of the number of deaths and injuries by gun violence in and around the year of

2017. Although I try to work around it by working not with discrete numbers of casualties but instances of shootings, there should be many other outliers that my analysis have not properly handled. Moreover, the abnormally large concentration of gun crimes after 2010 (taking up 95% of the dataset) leads me to believe that the dataset is incomplete, either due to its formulation or due to the nature of gun crime reports. At the same time, however, this analysis have not fully exploited the information embedded within this data frame as information of city where the shootings take place or the description of the shootings is not employed.