Drone Strikes in the Middle East

Pakistan, Somalia, Yemen

Ethan Allavarpu

3 June 2022

Contents

Abstract

In this report, we looked at drone strike data from 2004 onward (Pakistan) and 2007 onward (Somalia and Yemen). We considered the distribution of civilian casualties by country and president in addition to the frequency of strikes by president. Our analyses indicated that Pakistan had a higher percentage of civilian casualties, while Somalia had a civilian casualty rate (as a percent of the total persons killed) distinctly less than expected. Regarding presidents, we found evidence that the lethality of drone strikes differed by president (with Bush having the largest average fatality per strike) but did not find evidence of a significant difference in drone frequency by term length.

Background

Context

The United States has had a heavy military presence in the Middle East for the past 40 years–especially after the terrorist attacks on September 11, 2001. With improved technology, the United States expanded its attacks in size and scale. Specifically, we considered the impact of *drones* in the region. These drones provided airstrikes from above without the need to risk the lives of soldiers on the ground, potentially allowing for greater devastation in the region. Thus, we chose to investigate the impact of these strikes by the United States in three countries of interest–Pakistan, Somalia, and Yemen–and across three presidents: George W. Bush, Barack Obama, and Donald Trump.

Various projects have utilized this dataset, including Pakistan-specific drone strike visualizations. A student-created website goes into detail about the breakdown of deaths by demographic (e.g., civilian, military, child) and places the deaths on a timeline. Similarly, an infographic created by The Economist places the number of deaths from US drone strikes in the context of terrorist attacks; the visual also highlights major events over the past two decades related to these attacks and strikes.

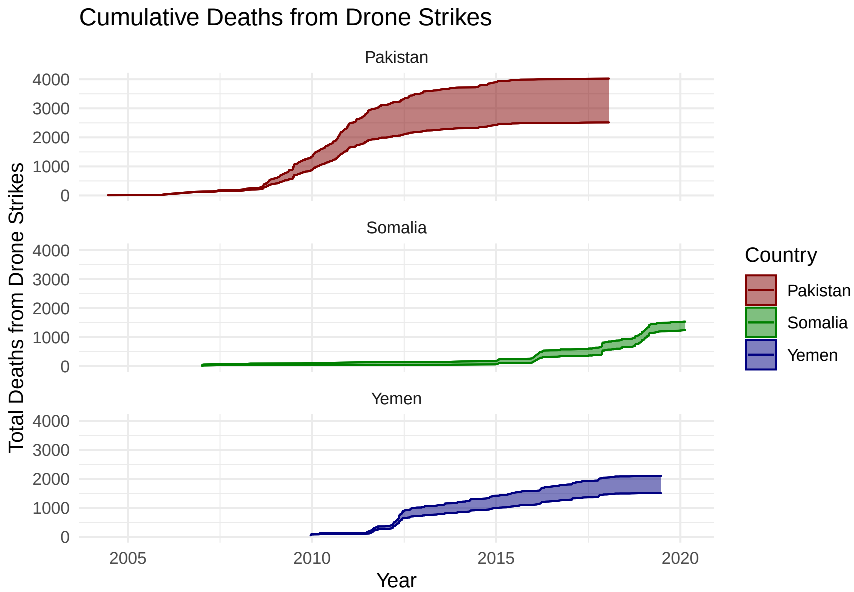
Earlier projects focus on relaying the data to users visually and placing the data in context. While useful for providing easy-to-digest results and background about the findings, these projects fail to dive deeper into any statistical analysis. In our project, we hope to provide statistical results to determine the *significance* of these results, rather than solely conducting exploratory data analysis. Additionally, our dashboard aims to provide an interactive experience for users by allowing them to filter the data at their discretion to gain a specific insight from a potentially overwhelming dataset.

Data Introduction

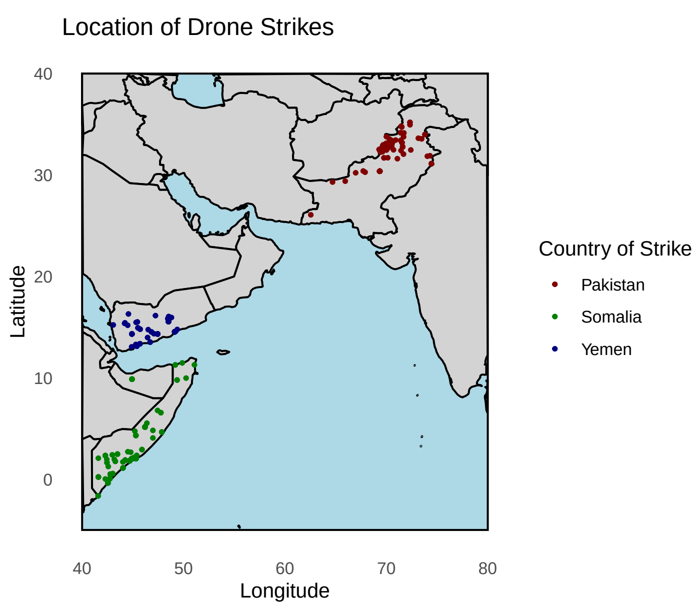
The data contains information on drone strikes in the Middle East, courtesy of The Guardian. We have drone strike data from 2004 for Pakistan and 2007 for Somalia and Yemen. Within these datasets, each observation corresponds to a single drone strike. We have information on the date and area of the strike along with the number of people, civilians, and children killed and the number of people injured. For each statistic, we had two numbers–a minimum and a maximum—which we displayed through various visualizations. For subsequent analyses, we constructed the number of deaths for a given drone strike as $\frac{\text{Minimum Deaths} + \text{Maximum Deaths}}{2}$.

From the area and country for each strike, we used a Google Maps API key (within the `ggmap` package) to geocode the areas into latitude and longitude. The latitude and longitude are for a general area–not a specific strike location–and some observations did not have geocoded locations.

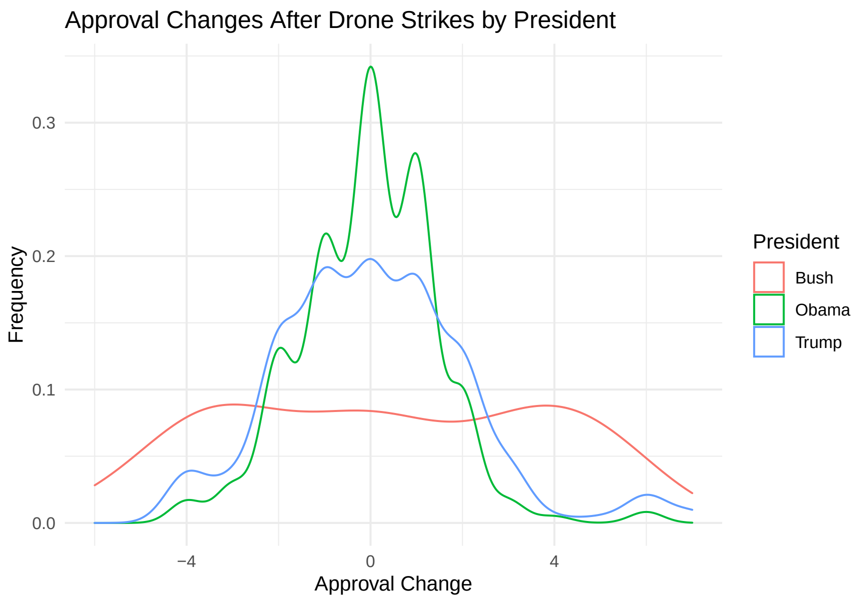
Exploratory Data Analysis



From the above graph, we saw the difference in drone deaths for each of the three countries. The upper limit for each plot represents the cumulative maximum number of deaths, while the lower limit represents the cumulative minimum number of deaths. Pakistan had the greatest number of deaths and the earliest timeline of the three countries, with data starting in 2004 and terminating around 2018. Yemen had the second-largest number of casualties, while Somalia had the fewest. Both Somalia's and Yemen's reported data began after 2005, with Somalia ranging from 2007 to 2020 and Yemen having data from 2010 to 2019. The difference in the number of deaths across the countries suggested that this may be an aspect of the data we will further explore in our statistical analyses.



The above graphic/map depicted drone strike locations in the three countries of interest, colored by country. As mentioned earlier, one limitation of the geocoding performed on the dataset was that some data points do not have valid latitude or longitude coordinates. For this visualization–and the RShiny maps discussed later–we omitted such points; however, we retained these points when performing coordinate-agnostic analyses. From the map, we saw that drone strikes occurred in clusters of locations, suggesting potential points of interest for the U.S. military.



Our third exploratory visualization compared the distribution of approval change after drone strikes across presidents. To determine the approval change, we scraped data gathered from UCSB to get the Gallup poll ratings. From there, we took the change in approval percentage between the poll just before and just after a drone strike, segmenting the data by the president in office. From this visualization, our results indicated that the approval shift did not appear to have a distinct trend in any specific direction; all three presidents had roughly symmetrical distributions centered around 0. Thus, it did not appear that approval changes from drone strikes would be a facet to investigate further (at least at this high level; the data may warrant a more in-depth analysis in future iterations to identify and remove potential confounding factors, like the domestic economy).

Statistical Analysis

After performing exploratory data analysis, we had three main interests in the dataset: (1) the accuracy/effectiveness of the drone strikes (in terms of casualties), (2) the strike lethality for each president, and (3) the frequency of drone strikes for each president's term.

Drone Accuracy Rate

For this statistical analysis, we compared the percentage of civilian casualties for the three countries (Pakistan, Somalia, and Yemen). For this analysis, we defined the overall and civilian number of casualties as the average of the minimum and maximum deaths for people and civilians, respectively.

We conducted a Chi-squared test to look at the distribution of civilian deaths across the three countries. Our null hypothesis would assume that the distribution of civilian deaths matched the distribution of noncivilian deaths, calculated by overall deaths – civilian deaths. The reason we set the null distribution as the number of noncivilians killed–and not just a pure equality–is because the number of deaths does not remain consistent across countries, so we wanted to account for that distribution. We can summarize our null and alternative hypotheses as follows:

$$H_0 : \text{Civilian deaths follow the same distribution as noncivilian deaths}$$
$$H_a : \text{Civilian deaths do not follow the same distribution as noncivilian deaths}$$

After conducting the Chi-squared test, we obtained a test statistic of $\chi^2 \approx 217.53$, which corresponds to a p-value of $p < 2.2 \times 10^{-16} \approx 0$. Thus, we rejected our null hypothesis that civilian deaths by country follow the same distribution as noncivilian deaths, implying that the countries differed in the accuracy of their strikes with respect to civilian deaths as a fraction of total deaths. Moreover, we conducted a posthoc analysis of the test to demonstrate the stark differences at the country level:

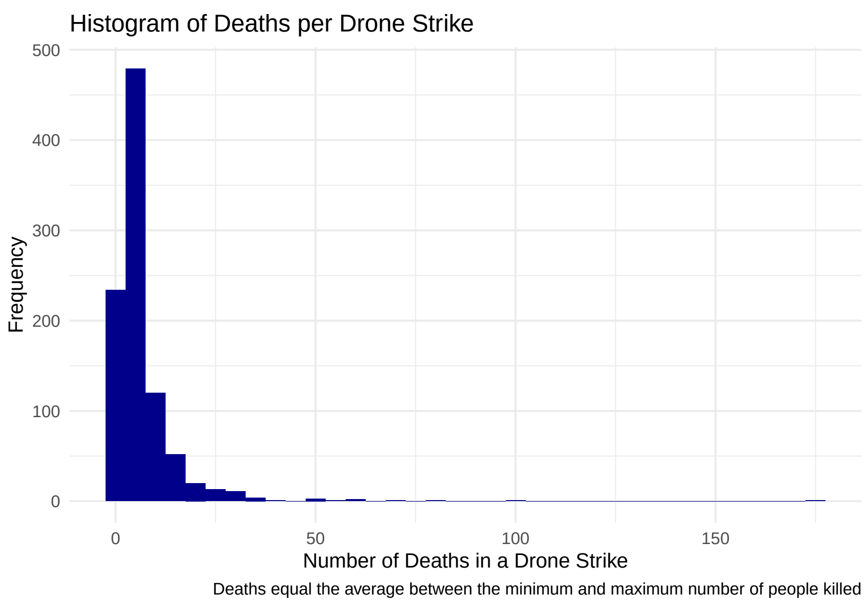
Table 1: Posthoc Summary of Civilian Deaths by Country

Country	Civilians Killed	Total Killed	Civilian Percent of Deaths	Pearson Residual
Pakistan	696	3,270	21.3%	9.3
Somalia	72	1,390	5.2%	-11.3
Yemen	266	1,811	14.7%	-1.6

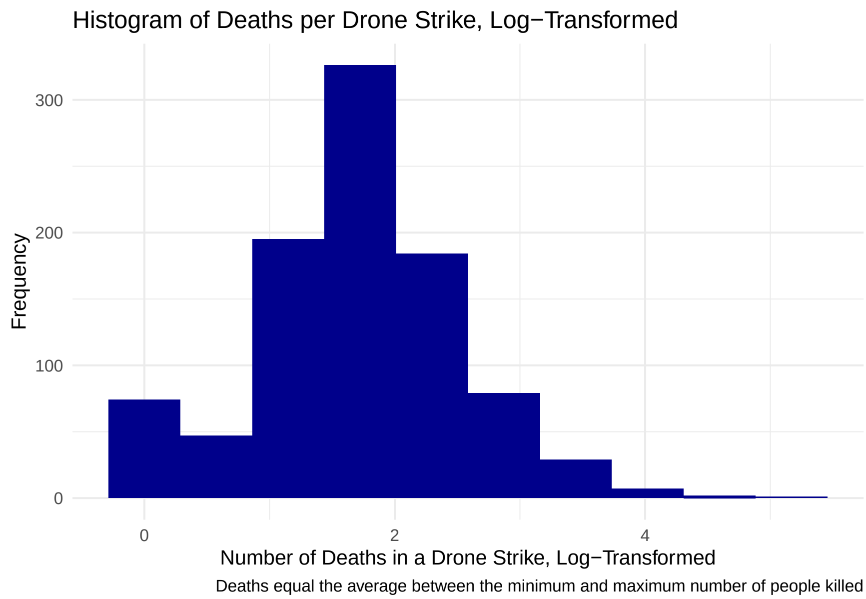
From the above posthoc analysis, we saw that Pakistan had a decidedly higher civilian death rate, while Somalia had a relatively low civilian death rate. While we could not make any causal inference from these results, it appeared that Pakistan–which does have data from an earlier time frame (since 2004)–has a less deliberate approach while Somalia had greater accuracy in eliminating hostiles.

Drone Lethality

Our second analysis considered the lethality of drone strikes, by president. Initially, when performing EDA, we noticed that the distribution of deaths from drone strikes was heavily right-skewed:



To correct for this (and meet the assumptions for the ANOVA we wished to conduct), we log-transformed the deaths after adding 1 to all deaths. This ensured that we did not have to remove strikes with 0 deaths from our analysis.



Once we ensured that we met the assumptions of our ANOVA (the above plot is roughly symmetric), we could construct our null and alternative hypotheses and conduct the analysis. Our null hypothesis would be that the average number of deaths per strike would be the same across all presidents; our alternative would be that at least one president had a significantly different average number of deaths.

$$H_0 : \mu_{\text{Bush}} = \mu_{\text{Obama}} = \mu_{\text{Trump}}$$
$$H_a : \text{At least one } \mu \text{ is not equal}$$
$$\mu_{\text{President}} \text{ represents the average number of deaths per strike, after transformation}$$

Table 2: ANOVA Results for Strike Lethality

	df	SS	MS	F	p-value
President	2	53.89	26.94	44.93	≈ 0
Residuals	941	564.31	0.60		

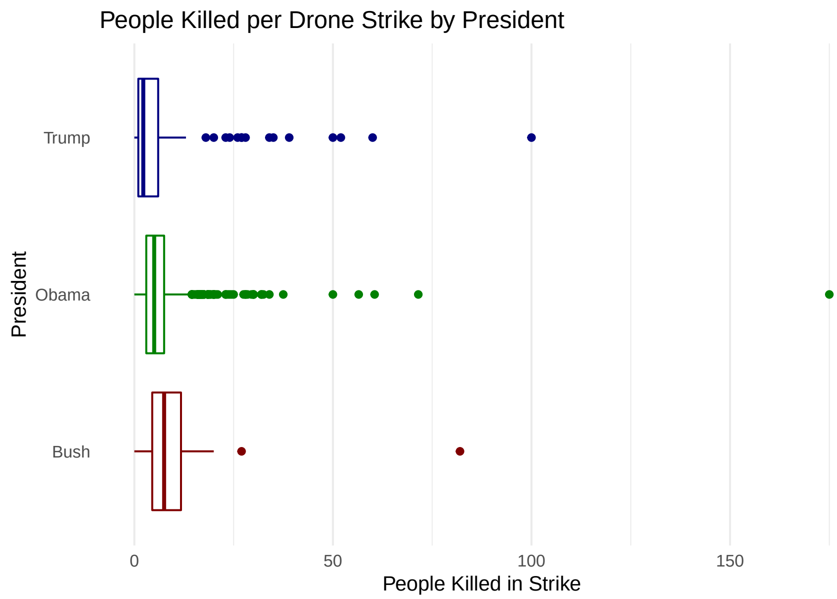
From the above ANOVA results, we saw that we have a statistically significant result. Because $p \approx 0 < 0.05 = \alpha$, we rejected the null hypothesis that $\mu_{\text{Bush}} = \mu_{\text{Obama}} = \mu_{\text{Trump}}$ and had evidence that supported the alternative that all three presidents did not have equal average death numbers per drone strike, after transformation. Thus, we decided to perform posthoc analysis. In addition to performing TukeyHSD pairwise comparisons, we calculated and plotted key summary statistics for the untransformed death numbers by president to place our findings in context.

Table 3: TukeyHSD Pairwise Posthoc

Pairwise Comparison	Adjusted p-value
Obama-Bush	0.0245
Trump-Bush	≈ 0
Trump-Obama	≈ 0

Table 4: Lethality Statistics by President

President	Average Deaths per Strike	Median Deaths per Strike
Bush	9.75	7.50
Obama	7.03	5.00
Trump	5.54	2.25



Drone Usage

Our third statistical analysis focused on the frequency of drone strike usage by each president. However, this analysis had a much greater limitation than our first analysis, due to the time frames for each country. Thus, to accommodate for data discrepancies across countries, we limited the date range to the overlapping date range across all three datasets. After filtering the data, we had strikes from December 18, 2009 (`2009-12-18`) to January 18, 2018 (`2018-01-18`). Because of this, our analysis of how frequently presidents utilized drones only applied to Presidents Obama and Trump.

From here, we conducted another Chi-squared test to see if the frequency of strikes aligned with the term length for each president over the restricted time frame. Therefore, we can summarize our null and alternative hypotheses as:

$$H_0 : \text{The number of drone strikes follows the distribution of term length}$$
$$H_a : \text{The number of drone strikes does not follow the distribution of term length}$$

After running the analysis, we obtained a Chi-squared statistic of 0.1316, which had a p-value of 0.7168. Thus, at the $\alpha = 0.05$ significance level, we failed to reject the null hypothesis–we did not have sufficient evidence that the frequency of drone strikes differed between Presidents Obama and Trump. However, as we noted earlier, this analysis was severely constrained due to the shorter time frame that overlapped across all three countries.

Dashboard

We expanded upon the basic graphics made in our exploratory data analysis and added more complex visualizations for the RShiny dashboard.

Specifically, we incorporated the visualizations for the cumulative deaths over time overall and by country from the EDA. From these base plots, we added filters for each country–and each president–so that an end-user could select which countries and presidential terms to highlight.

We also created a tab with an interactive timeline to visualize the strike locations over time. This visualization shows strikes over time at their impact point (found through geocoding), with the size and color correlating to the number of deaths caused by that strike.

The interactive timeline map presented itself as the most intensive element of the dashboard–and the project. In the Sources section, under Help Websites, we have linked two websites that proved useful for the construction of our RShiny app. While rudimentary, we hope it provides a more interactive high-level look at various statistics than a static exploratory data analysis, as is present in this report.

Conclusions

Our results provided a high-level insight into drone strike implementation in the Middle East countries of Pakistan, Somalia, and Yemen. These analyses provide great insight into *what* happened, but more research should be conducted to extract the *why*. That said, our investigation yielded the following results:

Regarding the civilian death rate as a percentage of total deaths, Pakistan had a higher civilian death rate at around 21.3%, while Somalia had a relatively low civilian death rate of only 5.2% (Yemen had a death rate of 14.7%). From these results, it seemed that Pakistan followed a blanket attack strategy, with little precision regarding sparing civilian lives. Somalia, by contrast, focused on targeted attacks, as most deaths were among non-civilians.

As for the presidents, the results are a little murkier. When looking at the average fatality of a strike (represented by the average number of deaths per drone strike), we discovered that President Bush had a significantly higher death rate than Presidents Obama and Trump, with Bush having a median of around 7.5 deaths per strike (Obama and Trump had a median of approximately 5 and 2.25 deaths per strike, respectively). While the median difference between Presidents Bush and Obama may not seem large, we found significance on the logarithmic (base-e) scale, which accounted for the right-skew of the initial data set when calculating the mean. Our results, though statistically significant, could be interpreted in two lenses: some might view the high death rates of Bush's strikes as an indicator of success in killing "the enemy," but others might interpret the same findings as Bush being brazen and careless–the killing may not have been necessary. Ultimately, this is where additional research by domain experts can help determine whether the high death toll was necessary (and therefore Bush was successful), or if it was a misuse of power by him.

The last analysis had a limited scope due to the constraints of our data source. We looked at whether presidents differed in their drone strike usage–more usage might indicate that a president could have abused his powers as Commander in Chief. Ultimately, our results provided no statistical significance, but we had severe date constraints (as described in the Limitations section). We could only compare Presidents Obama and Trump because of the overlapping time frames for strikes in all three countries–we limited the analysis to the overlap because we wanted to be able to get results from all three sets to avoid bias. Therefore, we did not consider the drone strike usage of President Bush, who had a significantly higher death rate per drone strike in our prior analysis.

Ultimately, the analyses we performed proved interesting. When we consider military action in the Middle East, we tend to group the Middle East as a whole and assume general effectiveness; however, the country-level analysis proved that each country had a statistically significant different civilian death rate as a percent of total deaths, hinting at different levels of effectiveness, precision, or discretion. When looking at the presidential performance, the average death toll per drone strike proved greatest for President Bush, which goes contrary to the recency bias which compels people to consider President Trump to be the most military-heavy president. In fact, Trump had the smallest death toll per strike of the three presidents–less than half of Obama's. These findings allow us to reassess our misconceptions about military action by presidents and the United States in the Middle East and evaluate our positions and decisions in the region.

Suggestions for Future Research

Additional research into these drone strikes includes looking at the relationship to terrorism attacks, including whether or not U.S. citizens or affiliates were affected by the attack. One strong data source would be the Global Terrorism Database (GTD); we chose not to use this dataset because we wanted to make our analysis and dashboard public, which would have required payment (the data is only free for private research).

Another sector to investigate would be the impact of drone strikes–either count, death rates, or injuries–on presidential approval rating. From our high-level exploration, there appeared to be limited-to-no relationship between approval change and various drone statistics, but a deeper dive into these two elements could glean previously undiscovered insights.

Outside of additional statistical analyses, we might further investigate potential causes or reasons behind our findings. While statistical analyses present great insights into the significance of certain trends, we must place our findings into context: the statistics alone cannot paint to complete picture.

Limitations

In the analyses, we encountered a big limitation in the date range for the individual countries. The original dataset contained data on drone strikes from Pakistan, Somalia, and Yemen. For Pakistan, the data begins in 2004; for Somalia and Yemen, we do not have drone strike data before 2007. To avoid issues that may arise due to missing/unaccounted data across the three countries, we decided to bound the range of interest. We limited the date range of interest between December 17, 2009 and January 24, 2018 because this was the range of dates for which all three countries had drone strikes. Therefore, we could only compare the strike frequency between Presidents Obama and Trump and ended up retaining 703 of 944 drone strikes (74.47%) from this analysis.

Additionally, for geocoding the areas and countries to derive latitude and longitude, we used a Google Maps API. While useful, we remained unable to geocode some locations due to undefined areas or incorrectly mapped locations. To verify that the geocoded location was correct, we redetermined the spatial country location from the latitude and longitude. If the spatial country and the original country did not match, we assigned the latitude and longitude to `NA` and dropped those strikes from map visualizations. In total, we removed 109 of 944 drone strikes (11.55%) for the map visualizations. However, we still included these strikes in subsequent analyses.

Sources

Data

- The Guardian (via Professor Vivian Lew, UCLA Department of Statistics)
  - `us-pakistan-strikes-from-2004.xlsx`
  - `us-somalia-strikes-from-2007.xlsx`
  - `us-yemen-strikes-from-2007.xlsx`
- Gallup (via UCSB)
  - `bush-ratings.csv`
  - `obama-ratings.csv`
  - `trump-ratings.csv`

Help Websites

- Creating spatial mapping in R from coordinates
- Creating an RShiny map timeline

R Packages

- `ggmap`: D. Kahle and H. Wickham. ggmap: Spatial Visualization with ggplot2. The R Journal, 5(1), 144-161. URL http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf
- `lubridate`: Garrett Grolemund, Hadley Wickham (2011). Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3), 1-25. URL https://www.jstatsoft.org/v40/i03/.
- `readxl`: Hadley Wickham and Jennifer Bryan (2022). readxl: Read Excel Files. R package version 1.4.0. https://CRAN.R-project.org/package=readxl
- `rvest`: Hadley Wickham (2021). rvest: Easily Harvest (Scrape) Web Pages. R package version 1.0.2. https://CRAN.R-project.org/package=rvest
- `rworldxtra`: Andy South (2012). rworldxtra: Country boundaries at high resolution.. R package version 1.01. https://CRAN.R-project.org/package=rworldxtra
- `sp`: Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), https://cran.r-project.org/doc/Rnews/.
- `tidyverse`: Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

GitHub Repository: https://github.com/ethan-allavarpu/drone-strikes