

Democrat vs Republican: A Study on the Magnitude of U.S. Drone Strikes in Foreign Countries

Statistics 140XP (Group 7)

JJ Svenson, Charlie Hoose, Mackenzie Lindholm,
Kendall Keely, Claire Nabours, Blair Warren

2025-12-03

Table of contents

1	Abstract	2
2	Introduction	2
3	Dataset	3
4	Research Questions	4
5	Exploratory Data Analysis	4
6	Methods	7
6.1	Lethality Analysis Based on Presidential Party	7
6.2	Accuracy Analysis Based on Presidential Party	9
6.3	Analysis of Temporal Trends	11
7	Conclusions	14
8	Limitations	15
9	Potential Impacts/Future Research	16
10	References	16
11	Appendix	16

1 Abstract

This study investigates whether U.S. political party control is associated with the lethality, accuracy, and overall magnitude of American drone strikes in Pakistan, Yemen, and Somalia. Using annual data on drone strike outcomes, such as total deaths, civilian, child, and military deaths, combined with political information, we apply a series of non-parametric statistical tests to evaluate partisan differences. Due to several outcome variables violating normality assumptions, we used Wilcoxon-Mann-Whitney tests to first compare lethality and accuracy across different political affiliations. Our results indicate that drone strikes under Democratic presidents were significantly more lethal on average, while accuracy did not differ significantly by party. Additionally, analyses of Obama’s “near-certainty” policy show a statistically significant decrease in lethality and increase in accuracy following the policy’s implementation. Although this observational data limits causal claims, these findings highlight important connections between political leadership and international drone strike outcomes.

2 Introduction

Drone strikes, although not a new military strategy, have become one of the most defining and controversial tools of modern-day warfare. The twenty-first century has seen a dramatic expansion in the United States’ use of drone strikes as a central tool of combat and foreign policy. As the use of drones has increased, debates have emerged around the ethical, political, and strategic implications of these strikes, particularly regarding civilian casualties and the decision-making process that guides target selection. Since drone programs operate under presidential administrations with different political affiliations and priorities, we aim to discover whether political leadership meaningfully shapes drone strike outcomes.

In this report, we investigate whether the political party in power is correlated with the lethality, accuracy, and overall magnitude of U.S. drone strikes in Pakistan, Somalia, and Yemen. This can be measured by looking at political control, such as the presidential party, congressional majorities, or the unification (or lack thereof) of government. Drone strike outcomes too can be measured in several different ways, such as looking at lethality, accuracy, and magnitude. Lethality refers to the number of individuals killed per strike event, broken up into civilians, children, and militants. Accuracy refers to the proportion of military targets among all casualties and the magnitude of the strikes can be observed through the number of strikes conducted each year.

Studying these patterns can help us better understand how drone strike outcomes shift across political administrations. It can also reveal how broader policy changes correspond to measurable changes in lethality and accuracy, such as Obama’s 2011-2013 “near-certainty” civilian protection standard discussed in the Brookings article “Biden can reduce civilian casualties during U.S. drone strikes. Here’s how.” (Kreps et al., 2022). Our analysis can provide helpful insight to the ongoing conversations about the role of politics in shaping military decision making, revealing how partisan control may influence the lethality and accuracy of U.S. drone strikes.

3 Dataset

Our analysis was based on the raw drone-strike dataset provided by *The Guardian* (The Guardian, n.d.), which we prepared and transformed in R. All the data-cleaning procedures and code used to construct our analytic datasets will be included in the appendix of this report. The original dataset compiles yearly information on U.S. drone activity in Pakistan, Yemen, and Somalia, including several key variables:

- Country
- Civilians killed (Min/Max Confidence Interval values)
- Children killed (Min/Max Confidence Interval values)
- Military personnel killed (Min/Max Confidence Interval values)
- Total number of strikes per event

For our exploratory data analysis, we focused on creating visualizations that captured overall trends in drone strike magnitude, accuracy, and lethality across the three countries. An example of this is our lethality visualizations which use a median based estimate (`med_people_killed`) derived from the average of the minimum and maximum casualty values. Similarly, our accuracy metric, defined as the proportion of military deaths to total casualties, required complete information on both civilian and military fatalities, which all three countries provided. On the other hand, to analyze lethality on a per-bomb basis, we had to focus on only Yemen and Somalia because Pakistan was missing values relating to the number of bombs dropped per strike event.

For our statistical tests, we relied on analyses that required complete data across the different casualty measures. For example, analyses on accuracy required complete information on civilian and military deaths. All three countries included complete data on these variables, contributing to our choice of these analyses.

To prepare the data, we removed any observations with missing or incomplete entries. Next, we created new columns of the dataset using the original variables, such as the aforementioned `med_people_killed` variable. We also manually added variables relating to the Presidential party, House majority, and Senate majority based on information from the House of Representatives digital archives (History, Art & Archives, U.S. House of Representatives, 2025). Cleaning the dataset in this way ensured that our comparisons across years, countries, and political administrations were consistent. All together, these datasets allowed us to examine how political party control relates to drone strike lethality, accuracy, and overall magnitude across Yemen, Somalia, and Pakistan.

4 Research Questions

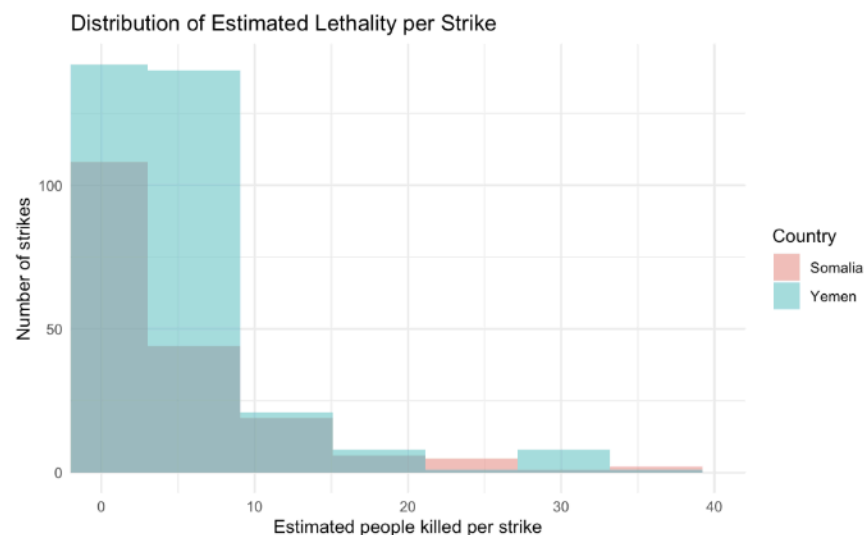
For our report, we are focusing on three main questions relating to U.S. drone strike data in Pakistan, Somalia, and Yemen.

1. Is there a significant difference in the lethality of U.S. drone strikes under Republican and Democratic presidents? For the point of this question, “lethality” is defined as the number of deaths estimated per drone strike event.
2. Is there a significant difference in the accuracy of U.S. drone strikes under Republican and Democratic presidents? We are defining accuracy to be the proportion of military targets versus civilian targets for each drone strike event, with a higher accuracy being a higher proportion of military targets.
3. How does the number of U.S. drone strikes change over time? The goal of this question is to better understand and contextualize the pattern of drone strikes across the timeline of the data. Even more, we can focus particularly on Obama’s presidency, as his number of drone strikes changed drastically over his years in office due to policy change.

Through these questions, we aim to have a better understanding of U.S. drone strikes in the twenty-first century.

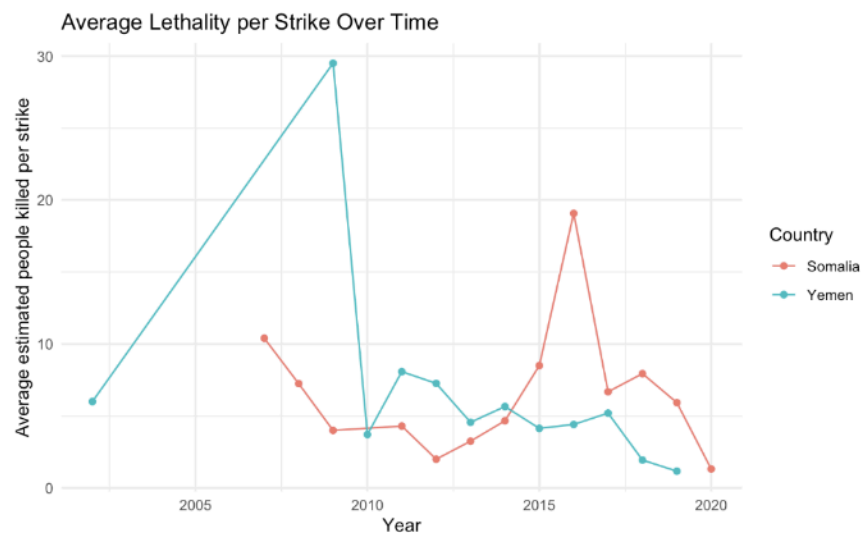
5 Exploratory Data Analysis

Before we begin our statistical analysis and testing, we first need to look at the data through a series of visualizations. This will help better frame our questions in the context of our data and provide a more practical understanding of U.S. drone strikes. First, we can look at how we can measure the lethality of drone strikes.

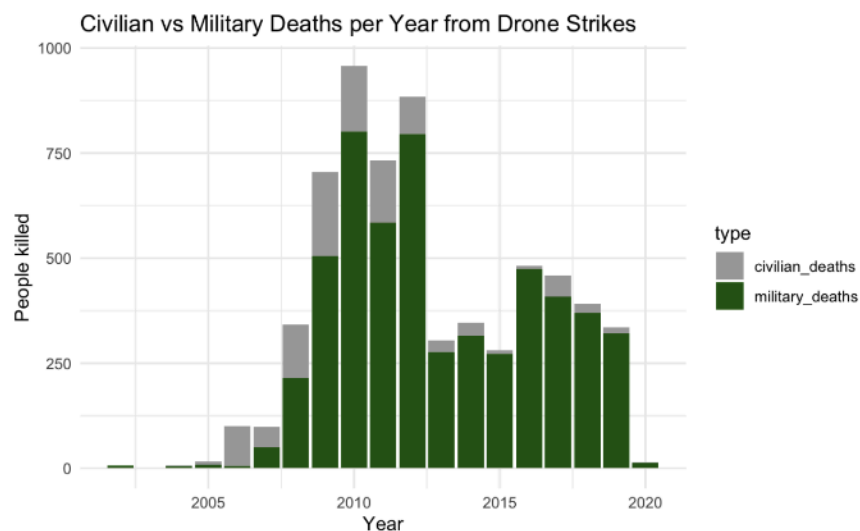


The graph above shows the distribution of people killed per strike event. This shows that deadlier strike events are much less common than smaller, more precise drone strikes that kill 10 or less people. For deadlier strikes that killed 10 or more people, Yemen and Somalia seem to have similar

distributions. However, Yemen has endured a higher number of strikes overall, making their number of low-lethality strike events higher than Somalia's. It is also important that we analyze lethality trends over time.

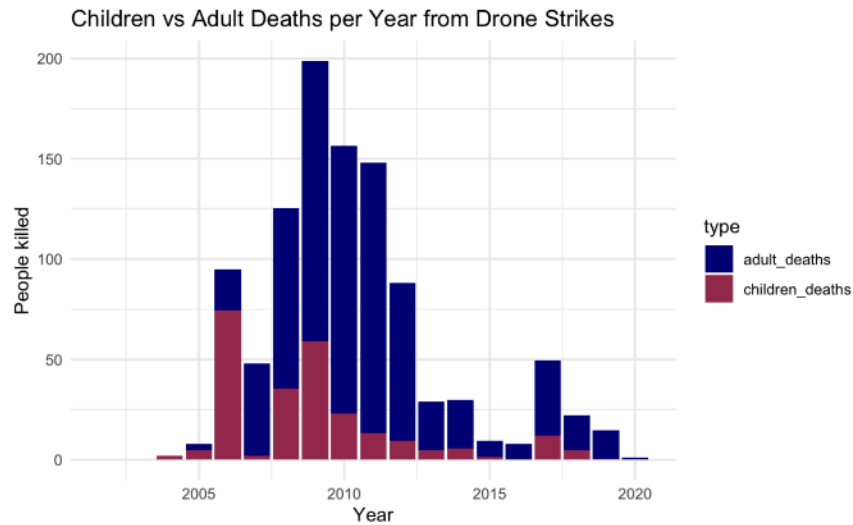


Looking at the graph above, we can see the average lethality of drone strike events over time in Yemen and Somalia. There were some missing values in the early- to mid-2000s, hence the discrepancy in data points there. Also, drone strikes in Somalia did not begin until the later part of the decade. As we can see, there are some distinct jumps in the amount of people killed per strike, primarily in the years 2008 and 2016. Both of these years were hallmarked by the transitions of one president to another, hence why their values could be so staggering. Ultimately, we can see how the lethality of drone strikes looks, but it is important to next look at accuracy.

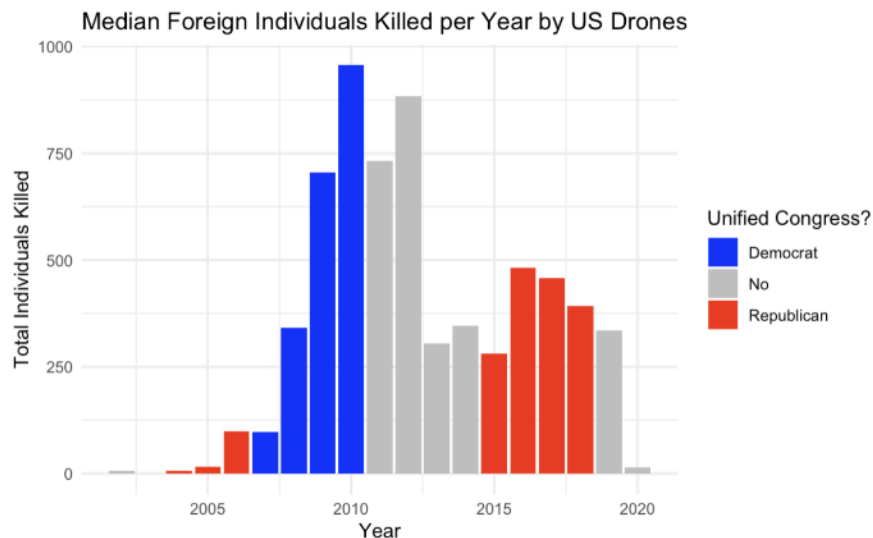


Shown above is the number of people killed per year, grouped by civilian and military individuals. We can see how there were not many deaths per year up until 2008, when things really picked up.

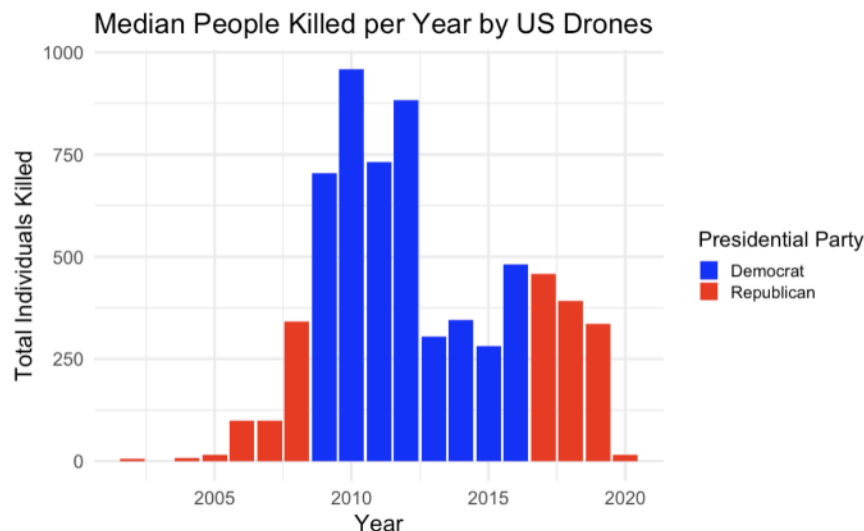
We can also see a much higher proportion of civilian deaths in the earlier years compared to later, which most likely comes from Obama's policy of "near-certainty." The goal of drone strikes would (hopefully) be to target military individuals and keep innocent civilians safe, so this graph does a good job at expressing when the U.S. lacked that accuracy. We can also analyze the age of drone strike victims.



In a similar vein, we can see the number of adults versus children killed per year as a result of U.S. drone strikes. This continues to paint a picture of the accuracy of drone strikes, we can assume that much more children than adults tend to be innocent civilians. Although most of the data shows that adults were the main target per year, 2006 is a special outlier in which over 75% of the individuals killed that year were children. This helps paint a better picture of how the accuracy of U.S. drone strikes changed over the years, which can help frame our analysis better. Moving on, it is important to look into how the political climate of America ties into the drone strike figures.



The graph above shows the status of Congress, whether unified under one political party or not, relates to the number of individuals killed from drone strikes over the years. It can help show how politics shaped the landscape of U.S. drone strikes, as we can see a large increase in strikes under Democratic control of Congress in the late 2000s.



Finally, instead of looking at the congressional makeup per year, we can examine whether the president was a Democrat or Republican. The presidents in question within the timeframe of the dataset are George W. Bush, Barack Obama, and Donald Trump (in order). We can see that there was a substantially high number of people killed per year by drone strikes in Obama's first term, however the rest are all much more similar.

Overall, through the process of EDA, we can get a greater understanding of the dataset, which can inform us to make better statistical tests and analyses. Next, we will dive into answering the questions posed above, using tests to analytically show whether there are significant differences in the accuracy and lethality of U.S. drone strikes in foreign countries under different political rules.

6 Methods

To investigate whether political party control is related to the lethality and accuracy of U.S. drone strikes in Yemen, Pakistan, and Somalia, we used a series of statistical tests. Our approach was to examine how drone-strike patterns changed across different political conditions such as the presidential party. We chose this design as the data varies year by year, allowing us to compare trends across different administrations.

6.1 Lethality Analysis Based on Presidential Party

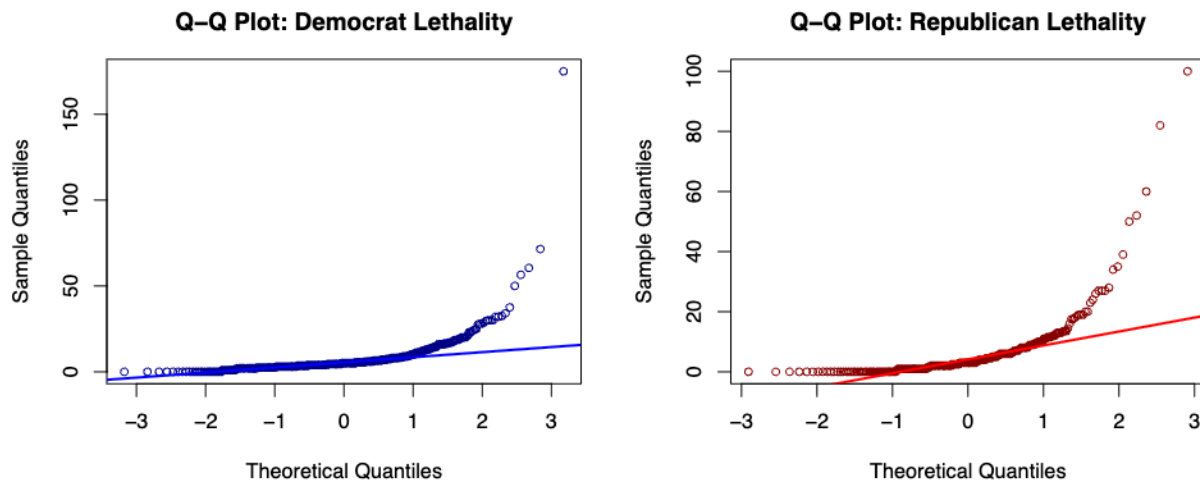
First, we aimed to analyze the effects of the presidential party on the lethality of drone strikes. To do this, we had to define a set of null and alternative hypotheses:

6.1.1 Hypotheses

- H_0 : There is no significant difference in the lethality of U.S. drone strikes between Democrat and Republican presidents.
- H_a : There is a significant difference in the lethality of U.S. drone strikes between Democrat and Republican presidents.

6.1.2 Assumption Testing

In order to select our test, we needed to analyze the normality and variance of the distribution of lethality values. To do this we chose to further visually explore the data:



Due to the non-linear nature of the Q-Q plots, we suspected that the lethality was non-normally distributed. We formally tested this using the following Shapiro-Wilk tests:

Democrat Lethality: Shapiro-Wilk: $W = 0.4377$, $p\text{-value} = 3.30e-41$

Republican Lethality: Shapiro-Wilk: $W = 0.5326$, $p\text{-value} = 1.95e-26$

The p-values of the Shapiro-Wilk tests for both Democrats and Republicans were significantly below the 0.05 threshold, suggesting that the lethality data is not normally distributed. To test the assumption of equal variances between the two parties, we employed Levene's test:

Levene's Test: $F = 3.7834$, $p\text{-value} = 0.0521$

The p-value of Levene's test was 0.0521 which is greater than 0.05, suggesting that we can assume equal variances.

6.1.3 Statistical Test

Since the normality assumption was violated but the equal variance assumption was met, we chose to use the **Wilcoxon-Mann-Whitney test**, which is the non-parametric alternative to the independent samples t-test. This test does not require the assumption of normality and is appropriate for comparing two independent groups. A key assumption made in this analysis is that the samples of drone strikes between Democratic and Republican administrations were independent of each other.

Wilcoxon rank sum test with continuity correction

```
data: lethality by presidential_party
W = 114776, p-value = 8.875e-10
alternative hypothesis: true location shift is not equal to 0
```

6.1.4 Results

The Wilcoxon-Mann-Whitney test yielded a W statistic of 114,776 with a p-value close to 0. Since $p = 8.875 \cdot 10^{-10} < 0.05$, we reject the null hypothesis. We can conclude that there is a statistically significant difference in the lethality of U.S. drone strikes between Democrat and Republican presidents.

Table 1: Lethality Statistics by Presidential Party

presidential_party	n	Mean	Median	SD	Min	Max
Democrat	666	7.04	5	9.52	0	175
Republican	275	6.43	3	11.00	0	100

Looking at the summary statistics above, we can infer that drone strikes during Democratic presidencies tend to be more lethal.

6.2 Accuracy Analysis Based on Presidential Party

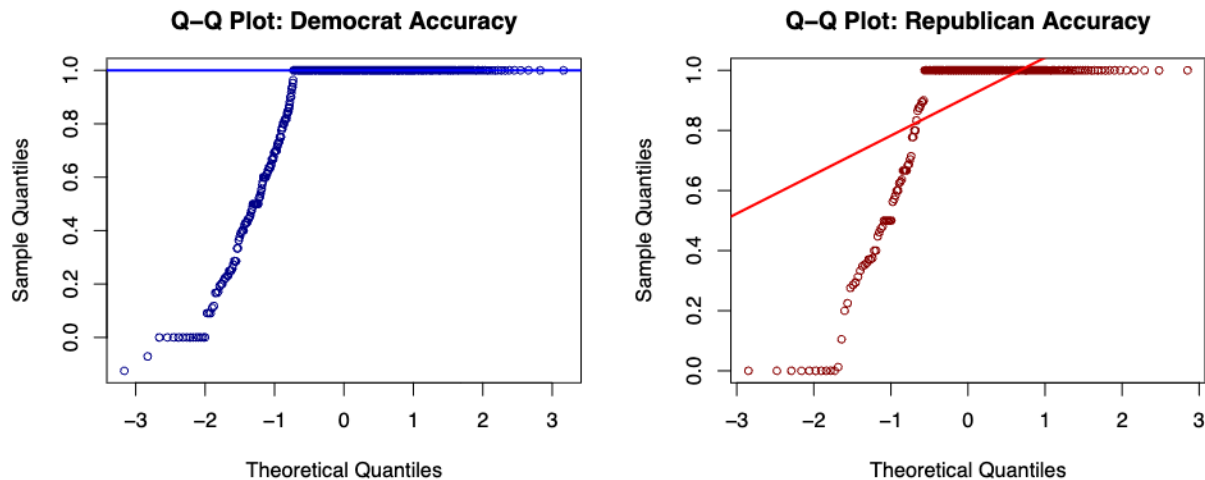
Next, we examined the effects of presidential party on the accuracy of drone strikes, defined as the proportion of military targets (military killed / total killed).

6.2.1 Hypotheses

- H_0 : There is no significant difference in the accuracy of U.S. drone strikes between Democrat and Republican presidents.
- H_a : There is a significant difference in the accuracy of U.S. drone strikes between Democrat and Republican presidents.

6.2.2 Assumption Testing

We performed the same diagnostic tests for the accuracy data:



Democrat Accuracy: Shapiro-Wilk: $W = 0.5314$, $p\text{-value} = 4.90\text{e-}38$

Republican Accuracy: Shapiro-Wilk: $W = 0.5973$, $p\text{-value} = 9.18\text{e-}23$

Levene's Test: $F = 3.7086$, $p\text{-value} = 0.0545$

6.2.3 Statistical Test

The accuracy data also violated the normality assumption (both $p\text{-values} < 0.05$) but met the equal variance assumption (Levene's $p = 0.0545 > 0.05$). Therefore, we again employed the Wilcoxon-Mann-Whitney test. Again, we assumed independence between the Democrat and Republican samples.

Wilcoxon rank sum test with continuity correction

```
data: accuracy by presidential_party
```

```
W = 77237, p-value = 0.08202
```

```
alternative hypothesis: true location shift is not equal to 0
```

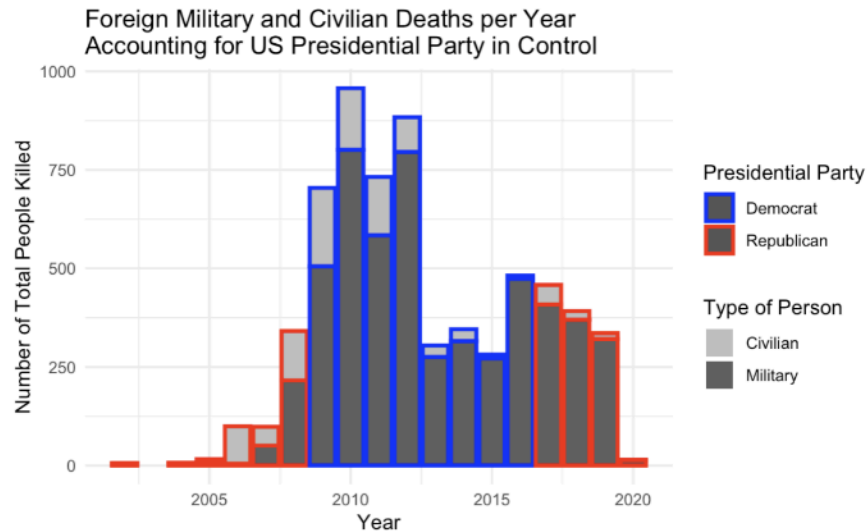
6.2.4 Results

The Wilcoxon-Mann-Whitney test resulted in a $p\text{-value}$ of 0.08202 which is greater than 0.05, so we fail to reject the null hypothesis. There is no statistically significant difference in the accuracy of U.S. drone strikes between Democrat and Republican presidents.

6.3 Analysis of Temporal Trends

Our group was also interested in exploring the general change in lethality of U.S. drone strikes over time, regardless of political party. To begin this process, we did some more visual analysis.

The graph below displays the number of individuals killed per year by U.S. drone strikes, broken down by civilian personnel (light grey) and military personnel (dark grey). At the same time, each year's bar is highlighted by the then-current U.S. president's political party, blue for Democrat and red for Republican.



Our Wilcoxon-Mann-Whitney test on the lethality between presidential parties in power proved that drone strikes with a Democratic president have generally been more lethal. However, the graph above distinctly shows a drop off in lethality around the year 2012 during Obama's administration. This trend aligns with Obama's "near-certainty" policy discussed in the Brookings article "Biden can reduce civilian casualties during U.S. drone strikes. Here's how." (Kreps et al., 2022). This article details a policy developed by Obama between 2011 and 2013 that required the U.S. to be "near-certain" that a certain strike was not going to have civilian casualties in order to carry it out. Our visual analysis along with this commentary inspired our group to employ formal statistical tests to explore if this policy truly lowered lethality and increased accuracy of U.S. drone strikes.

6.3.1 Near Certainty Policy Lethality Analysis

To explore the lethality aspect, we set up the following hypothesis structure:

6.3.1.1 Hypotheses

- H_0 : There is no difference in drone strike lethality before (2009-2011) vs after (2012+) the policy implementation.
- H_a : There is a difference in drone strike lethality before vs after the policy implementation.

6.3.1.2 Assumption Testing

We performed the same diagnostic tests as in previous sections:

Before Policy (2009–2011): Shapiro-Wilk $W = 0.6757$, $p\text{-value} = 1.77\text{e-}23$

After Policy (2012+): Shapiro-Wilk $W = 0.3098$, $p\text{-value} = 6.78\text{e-}35$

Levene's Test: $F = 1.9104$, $p\text{-value} = 0.1674$

6.3.1.3 Statistical Test

The data violates the normality assumption (both $p\text{-values} < 0.05$) but met the equal variance assumption (Levene's $p = 0.1674 > 0.05$). Therefore, we again employed the Wilcoxon-Mann-Whitney test. We assumed independence between the strikes before and after the Near Certainty Policy.

Wilcoxon rank sum test with continuity correction

data: lethality by policy_period

$W = 40870$, $p\text{-value} = 1.579\text{e-}08$

alternative hypothesis: true location shift is not equal to 0

6.3.1.4 Results

The Wilcoxon-Mann-Whitney test showed a significant difference ($p = 1.579 \cdot 10^{-8}$).

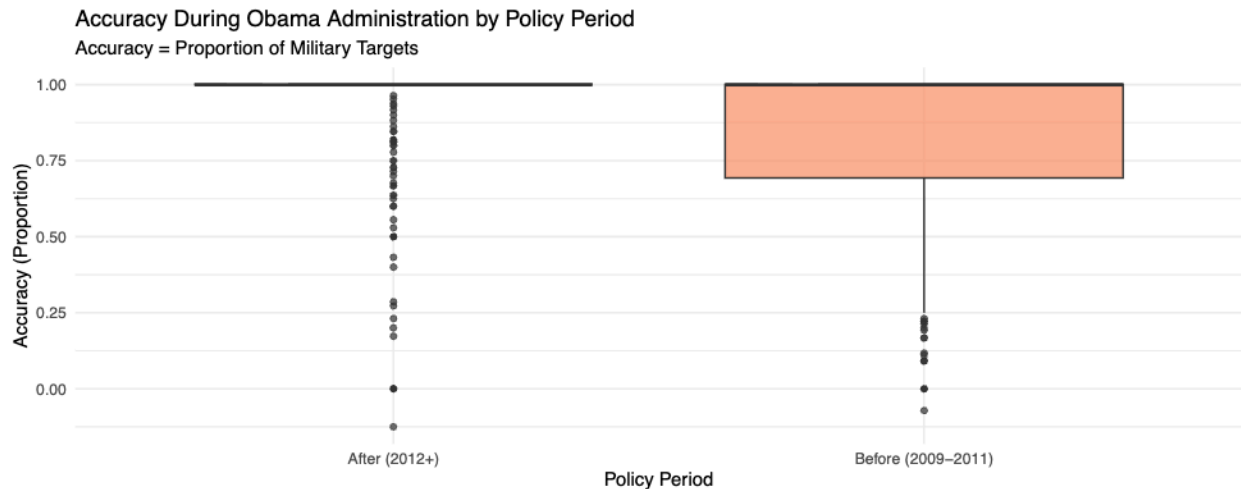
Table 2: Lethality Before vs After Near Certainty Policy

policy_period	n	Mean	Median	SD
After (2012+)	369	6.22	4.5	10.56
Before (2009-2011)	297	8.06	5.5	7.93

Looking at the summary statistics above, we can conclude that drone strike lethality decreased after the Near Certainty Policy was implemented.

6.3.2 Near Certainty Policy Accuracy Analysis

We also examined whether the accuracy of drone strikes (proportion of military targets) changed after the policy implementation. Through visual analysis there seemed to be a distinct difference before and after the policy was implemented:



To formally test this, we used the following hypotheses:

6.3.2.1 Hypotheses

- H_0 : There is no difference in drone strike accuracy before (2009-2011) vs after (2012+) the policy implementation.
- H_a : There is a difference in drone strike accuracy before vs after the policy implementation.

6.3.2.2 Assumption Testing

We tested the normality assumption using Shapiro-Wilk tests:

Before Policy (2009-2011): Shapiro-Wilk $W = 0.6567$, $p\text{-value} = 1.22e-23$

After Policy (2012+): Shapiro-Wilk $W = 0.3910$, $p\text{-value} = 9.57e-33$

Both periods showed $p\text{-values} < 0.05$, indicating non-normal distributions. We then tested equal variances:

Levene's Test: $F = 34.2775$, $p\text{-value} = 0.0000$

The Levene's test showed unequal variances ($p < 0.05$), indicating that the distributions differ in spread.

6.3.2.3 Statistical Test

The Wilcoxon-Mann-Whitney test, which does not assume equal variances, was therefore appropriate for comparing the distributions of accuracy before and after the policy implementation. However, because the variances are not equal, the Wilcoxon-Mann-Whitney test can reveal a difference in distributions rather than a difference in medians.

Wilcoxon rank sum test with continuity correction

```
data: accuracy by policy_period
W = 60397, p-value = 1.382e-08
alternative hypothesis: true location shift is not equal to 0
```

6.3.2.4 Results

The Wilcoxon-Mann-Whitney test showed a significant difference ($p = 1.382 \cdot 10^{-8}$). Since the Levene’s test indicated unequal variances ($p < 0.05$), this test is comparing the overall distributions rather than just medians.

Table 3: Accuracy Before vs After Near Certainty Policy

policy_period	n	Mean	Median	SD
After (2012+)	354	0.936	1	0.188
Before (2009-2011)	286	0.824	1	0.293

Looking at the summary statistics, the after-policy period shows both higher mean and lower standard deviation in accuracy. This suggests that the Near Certainty Policy was associated with changes in the distribution of drone strike accuracy, with generally higher and more consistent accuracy values after implementation. In other words, the strikes were more consistently accurate following the implementation of the policy. This aligns with the visual difference we observed in the boxplot above.

7 Conclusions

All analyses employed the Wilcoxon-Mann-Whitney test due to violations of the normality assumption as confirmed by Shapiro-Wilk tests (all p-values < 0.05). The Wilcoxon-Mann-Whitney test is a non-parametric test appropriate for comparing two independent groups without assuming normality. Equal variance assumptions were met in all comparisons (Levene’s test p-values > 0.05), except for the analysis on the accuracy before and after the Near Certainty Policy. Our results were as follows:

Table 4: Summary of Statistical Tests

Analysis	Test	P_Value	Significant
Lethality (Presidential Party)	Wilcoxon-Mann-Whitney	8.875e-10	Yes
Accuracy (Presidential Party)	Wilcoxon-Mann-Whitney	0.08202	No
Obama Policy: Lethality	Wilcoxon-Mann-Whitney	1.579e-08	Yes
Obama Policy: Accuracy	Wilcoxon-Mann-Whitney	1.382e-08	Yes

Through our statistical analysis, we concluded that there was a significant difference between the lethality of drone strikes based on the presidential party in power. Drone strikes under a Democratic presidential administration were significantly more lethal. On the other hand, we found that there was not a significant difference in the accuracy of drone strikes between the two political parties.

In terms of temporal trends, we used two Wilcoxon-Mann-Whitney tests to confirm the effectiveness of Obama’s Near Certainty Policy in terms of both lethality and accuracy. The tests revealed that there is statistically significant evidence that U.S. drone strikes had decreased lethality and increased accuracy after the Near Certainty Policy was implemented.

8 Limitations

There are several limitations that should be considered when interpreting the results of our study. First, the observational nature of our data prevents us from making any causal claims about how political control affects drone strike outcomes. Thus, we can identify and describe patterns and associations, but we cannot determine whether political leadership directly caused changes in lethality or accuracy. Second, the dataset suffers from missing or incomplete information, particularly for the number of bombs per strike event in Pakistan. This makes it impossible to evaluate lethality on a per-bomb basis in Pakistan.

Another potentially key limitation is the confounding variable that is intensity of conflict during each presidency. Different presidential terms mark years where the U.S. was fighting in different locations, as well as different opponents and context, such as the resurgence of the Taliban, the rise of ISIS, etc. Also, there is a very uneven and unbalanced sample size of years in which the U.S. engaged in heavy drone activity. Likewise, these events are distributed unevenly among different presidents and policy periods. Another uncontrollable but key confounding variable is the potential measurement error and misclassification/classification uncertainty. Child vs. militant vs. civilian deaths are purely based on secondary reporting and tracking and can often have errors in classifying who exactly was killed. This uncertainty can also lead to heavily under- or over-reporting the death/injury toll.

Another key limitation and call for further analysis is the heterogeneity amongst the countries. Pooling together Pakistan, Yemen, and Somalia disregards the differences in terrains, allies, and intelligence quality. By grouping these three all in one bucket, we are essentially assuming that political party effects are the same across all three countries, which of course is very unlikely. This can make statistics between the three harder to distinguish, even though the same metric can mean something completely different across different countries. Ignoring country-specific context can be dangerous and may not give as definitive of a result as we may like. We briefly analyzed the differences between the countries in our EDA and, for the purpose of this study, decided to

focus on the effects of political control rather than differences between countries. However, future research could help address some of these limitations.

9 Potential Impacts/Future Research

Our analysis can advance current knowledge by helping show how political leadership shapes the scale and accuracy of U.S. drone strikes. Drone policies often change across administrations, so identifying patterns between political control and the lethality or frequency of strikes can provide useful context for researchers and policymakers. This work also supports future research by showing how changes in policy standards may influence civilian and military casualty patterns over time. These results can inform ongoing discussions about targeting guidelines, reporting standards, and civilian protection policies. Our work highlights trends linked to political control, which can provide evidence that explains policy changes and guides future evaluations of U.S. drone practices. The broader impacts of our research relate to the social effects of drone strike policies and how they shape public understanding of military actions. Civilian casualties are an ongoing concern in media coverage and public discussions, and public trust is influenced by the level of transparency in military operations. By revealing how political conditions relate to drone strike outcomes, our findings can help people better understand the human consequences of these decisions and support accountability and public conversations about the use of drones.

Future research can build on our findings by using more detailed and frequent data to better capture the timing and effects of policy changes. Studies could examine the effects of specific policies, like Obama’s near certainty standard, on civilian and military casualties. Researchers could also examine the political motivations behind changes in drone activity by incorporating additional variables such as international relations, conflict intensity in each country, or public approval ratings. Another direction for future research could also compare the U.S. with other countries to see if political control has a similar effect.

Also, we can move from simple comparisons to multivariable models to account for differences in country, year, and intensity of conflict. This can help us differentiate party effects from differences in setting. Finally, qualitative studies of policy documents or expert interviews could help explain the reasoning behind observed patterns.

10 References

- History, Art & Archives, U.S. House of Representatives. (2025). *Party government since 1857*. <https://history.house.gov/Institution/Presidents-Coinciding/Party-Government/>
- Kreps, S. et al. (2022). Biden can reduce civilian casualties during US drone strikes. Here’s how. *Brookings*. <https://www.brookings.edu/articles/biden-can-reduce-civilian-casualties-during-us-drone-strikes-heres-how/>
- The Guardian. (n.d.). *US drone strikes data*. The Guardian.

11 Appendix

To see our project’s source code please refer to our [GitHub repository](#).