



# **STA 199 Project: Gun Violence Trends Specific to Mass Shootings in 2023**

---

**Bella A., Sophie C., Rebecca L., Gabriela P.**

# Background of the Study

---

## Topic: Analyzing Gun Violence Trends Specific to Mass Shootings

- **How do mass shooting patterns, including incident frequency, total victims, and arrest rates, vary across U.S. regions in 2023?**
- Understanding whether fatalities differ by state and arrest outcome may reveal differences in response effectiveness or enforcement patterns. This has implications for public safety policy and resource allocation.

## Motivation: Current Events, Everytown Research Mass Shooting Statistics

- “Firearms are the **#1** leading cause of death for children and teens in America”
- “In the eight years between 2015 and 2022, over 19,000 people were shot and killed or wounded in the United States in a mass shooting”
- “Reach of each mass shooting stretches far beyond those killed and wounded, harming the well-being of survivors, their families, and entire communities”
- Duke’s Recent Mandated Active Shooter Training

“Firearms Are the Leading Cause of Death for Children and Teens in the US.” Everytown Research & Policy, 16 Jan. 2026, [everytownresearch.org/graph/firearms-are-the-leading-cause-of-death-for-american-children-and-teens/](https://everytownresearch.org/graph/firearms-are-the-leading-cause-of-death-for-american-children-and-teens/).

“Mass Shootings Archives.” Everytown, [www.everytown.org/issues/mass-shootings/](https://www.everytown.org/issues/mass-shootings/). Accessed 22 Mar. 2026.

“Mass Shooting Factsheet.” Rockefeller Institute of Government, 10 Mar. 2026, [rockinst.org/gun-violence/mass-shooting-factsheet/](https://rockinst.org/gun-violence/mass-shooting-factsheet/).

# Data Introduction

---

## Cleaning the Data:

- Loaded the dataset using `read_csv` and coded "N/A" values as missing (*NA*)
- Cleaned column names using `janitor::clean_names()` for consistency
- Prepared the dataset for analysis using tidyverse functions
- Created new variables

## Defining Variables:

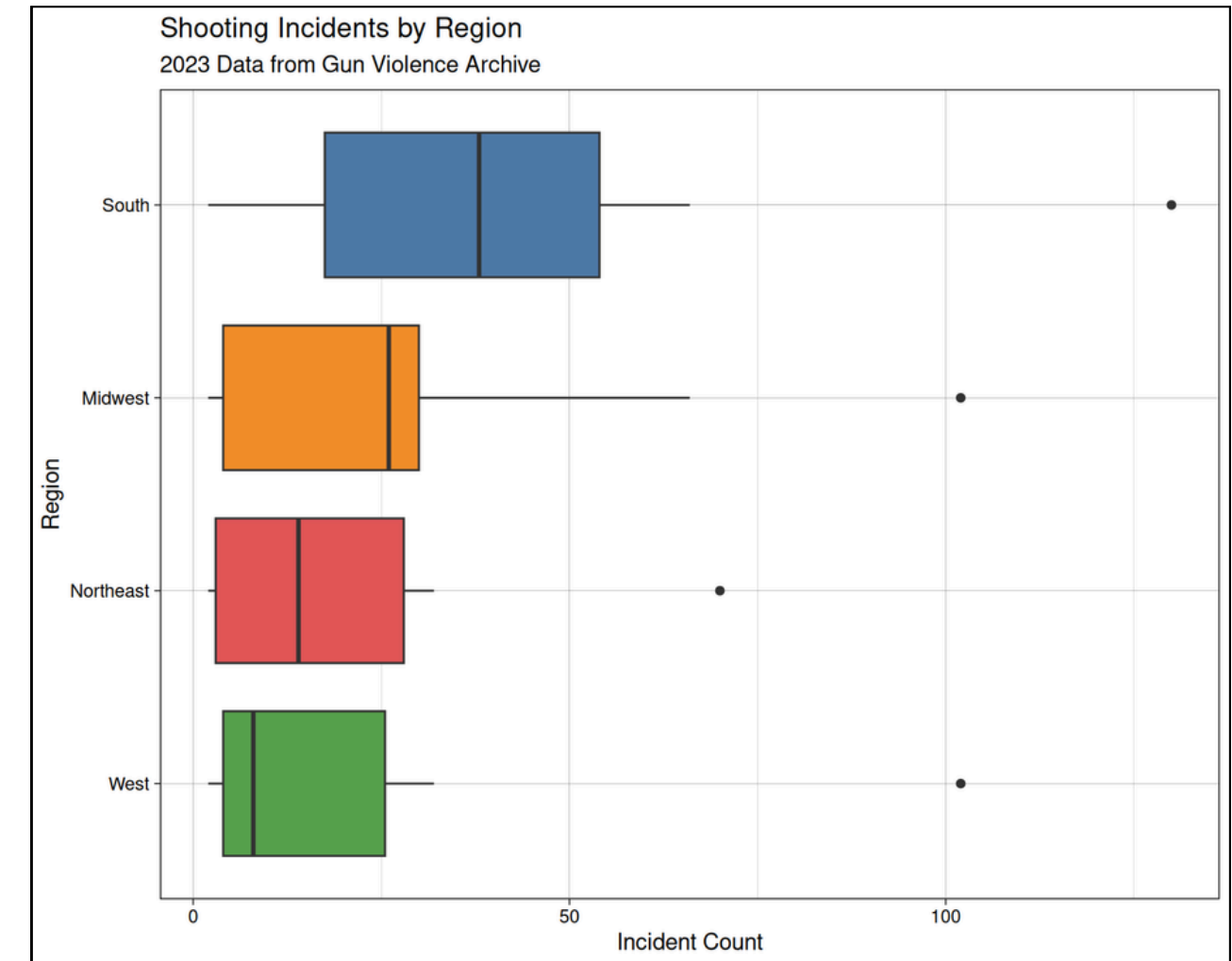
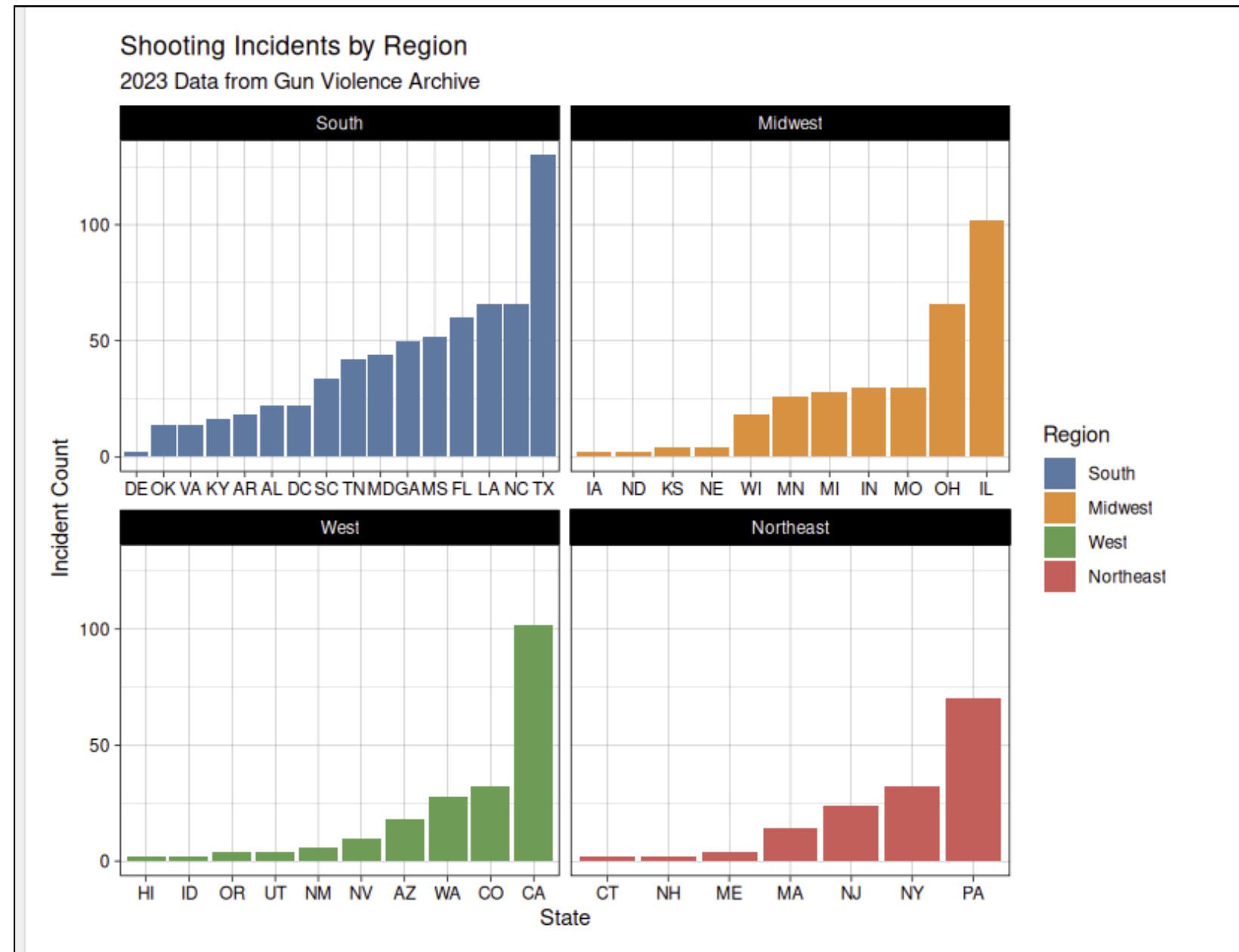
### Input (rows)

- **incident:** singular instance of a mass shooting

### Output (columns):

- **incident\_id:** mass shooting ID as labeled by Data is Plural
- **incident\_date:** date mass shooting occurred
- **state:** state in which the mass shooting occurred
- **city\_or\_county:** city or county in which the mass shooting occurred
- **address:** geographic location where the mass shooting occurred
- **suspects\_killed:** # of mass shooter suspects killed
- **suspects\_injured:** # of mass shooter suspects injured
- **total\_victims:** # of total victims injured or killed as a result of the mass shooting
- **region:** US geographical region (North, East, South, West) in which the mass shooting occurred

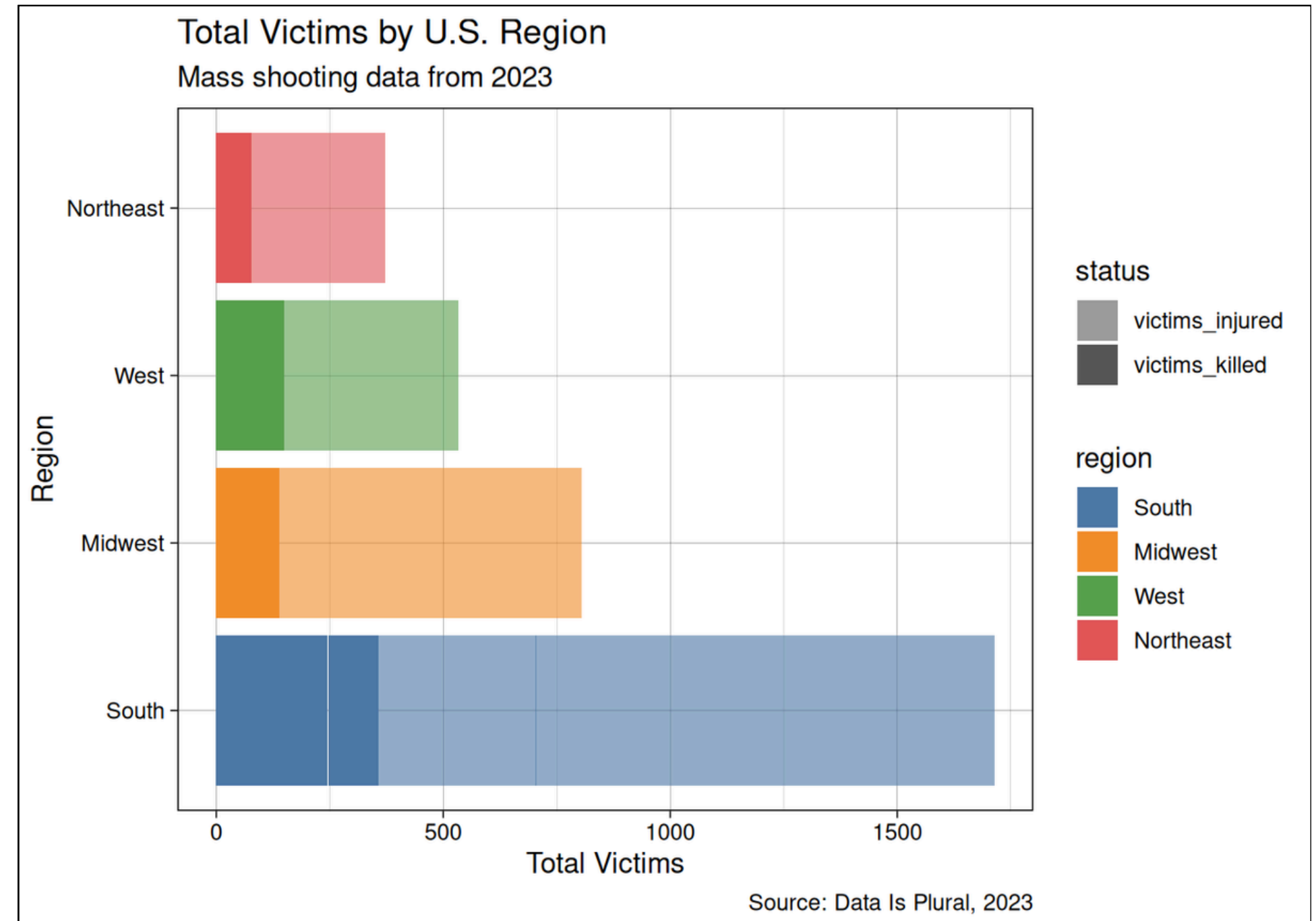
# Incident Count by State



- The **South** had the highest median overall incident counts, with TX at a notable upper outlier of 130
- Each had an upper outlier
- Northeast and West had lower ranges

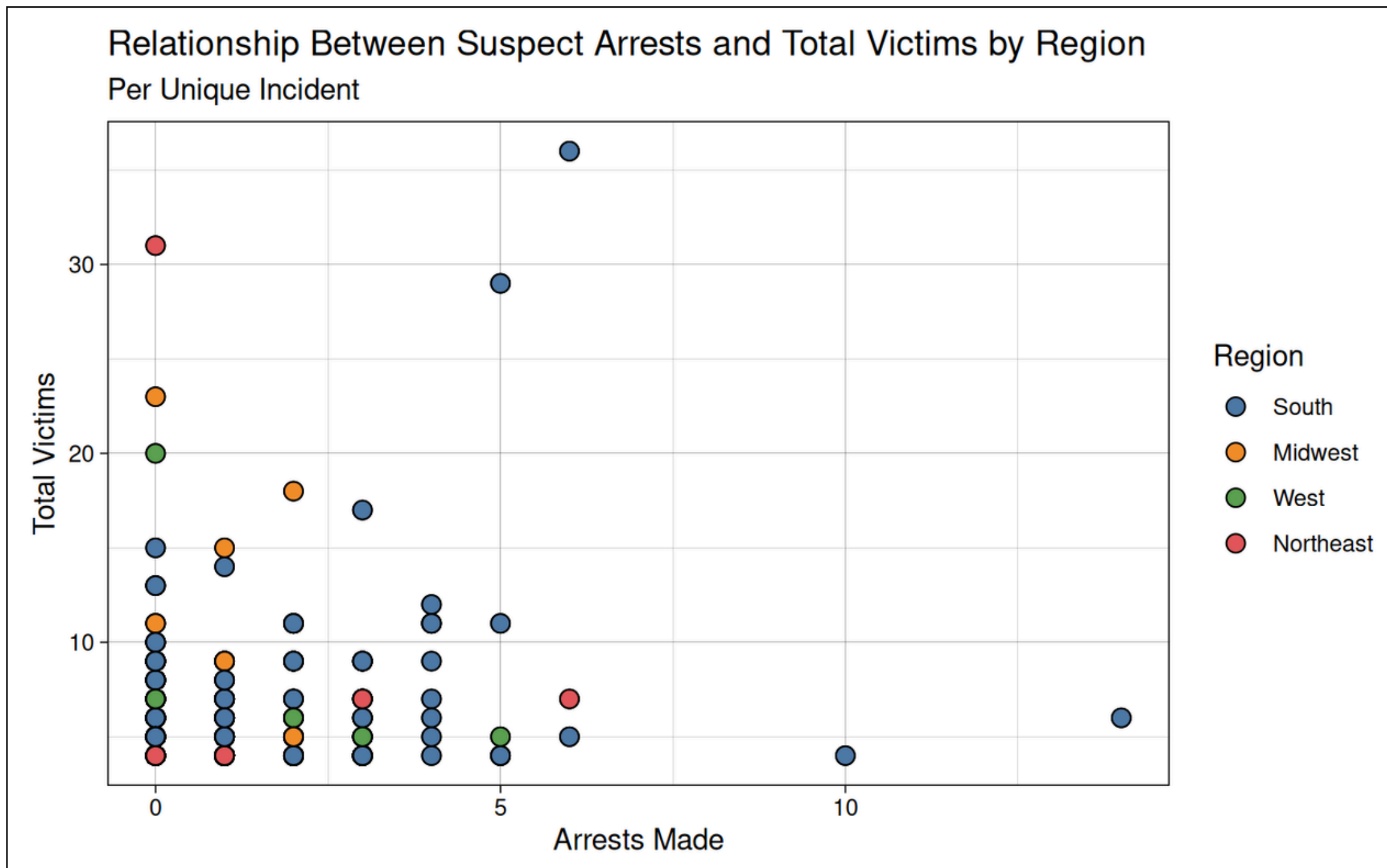
# Victims by Region

- The South has the highest number of victims, followed by the Midwest
- West and Northeast report lower victim totals
- Clear regional variability
- Highlights geographic disparities
- Even though the West had less shooting incidents than the Northeast, they had a higher victim count



# Arrests vs. Victims by Region

A tibble: 1 × 1	
	r
	<dbl>
	0.1965169
1 row	



- Points are very clustered across the graph; no clear trend line
- Correlation coefficient: 0.1965169
- Weak relationship between the number of total victims (victims injured + victims killed) and the number of arrests by region
- Implication: proactive action (policy) should be emphasized instead to minimize the number of total victims per mass shooting

# Conclusions + Discussion + Critique

---

## Limitations:

- Only 1 year of data
- Arrest rate may have confounding variables (if the suspect survived, etc. )

## More Research Possibilities:

- Multiple variable regression to find strongest predictor of number of victims between arrest rate, region, state, or more

# Takeaways

---

- Incidents and victims affected varies by region
- Arrest rates do not strongly correlate with victims affected
- Variations in state-to-state **preventative** policy, not arrest rate effectiveness, may explain the differences in mass shooting frequency
- The numbers are high everywhere, so more needs to be done to prevent these victimhoods