**Analysis of Fatal Police Shootings in the United States**

**Date of Report:** August 7, 2025

**Author:** Nipun Juneja (23csu215), Kunal Sharma (23csu396)

**Tools Used:** Python, pandas, Matplotlib

**Executive Summary**

This report presents a detailed analysis of the "Fatal Police Shootings in the United States" dataset, covering incidents from 2015 to early 2021. The primary dataset was subjected to a rigorous preprocessing protocol to handle missing values, outliers, and data type inconsistencies, resulting in a cleaned dataset of **5,635 records** for analysis. Key findings indicate a significant geographic concentration of incidents in states like California, Texas, and Florida. A demographic review of unarmed victims reveals they are predominantly male, with White and Black individuals being the most affected racial groups. The analysis also highlights a downward trend in incidents occurring without body camera footage. A separate analysis on the original, unprocessed data was conducted to examine trends related to victims showing signs of mental illness, revealing that White individuals constitute the largest group across most age brackets.

**Introduction and Methodology**

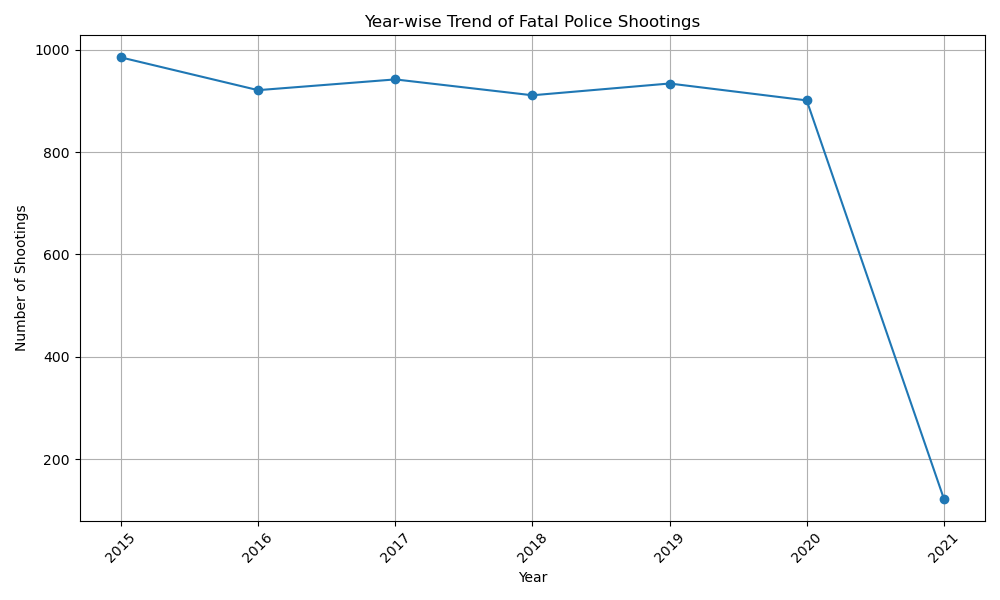
The objective of this analysis is to identify and visualize key trends and patterns within the fatal police shootings dataset. The primary dataset was cleaned to ensure the reliability of the findings. Preprocessing steps included:

* **Handling Null Values:** Missing data in categorical columns (name, armed, race, flee) were imputed using the mode or a placeholder. Missing numerical data (age) was imputed with the median.
* **Outlier Removal:** Outliers in the age column were identified and removed using the Interquartile Range (IQR) method.
* **Data Type Conversion:** The date column was converted to a datetime object for time-series analysis, and age was converted to an integer.
* **Data Integrity:** Rows with missing geographic coordinates (longitude, latitude) were dropped.

A specific analysis on victims with signs of mental illness was conducted on the **original, unprocessed dataset** to provide insights without the influence of the primary cleaning protocol.

**Overall Trends**

After cleaning, the dataset contains **5,635 records** of fatal police shootings. The year-wise distribution of these incidents shows a relatively consistent number of shootings from 2015 to 2020. The noticeable drop in early 2021 is likely due to the dataset's end date and not indicative of a long-term trend.

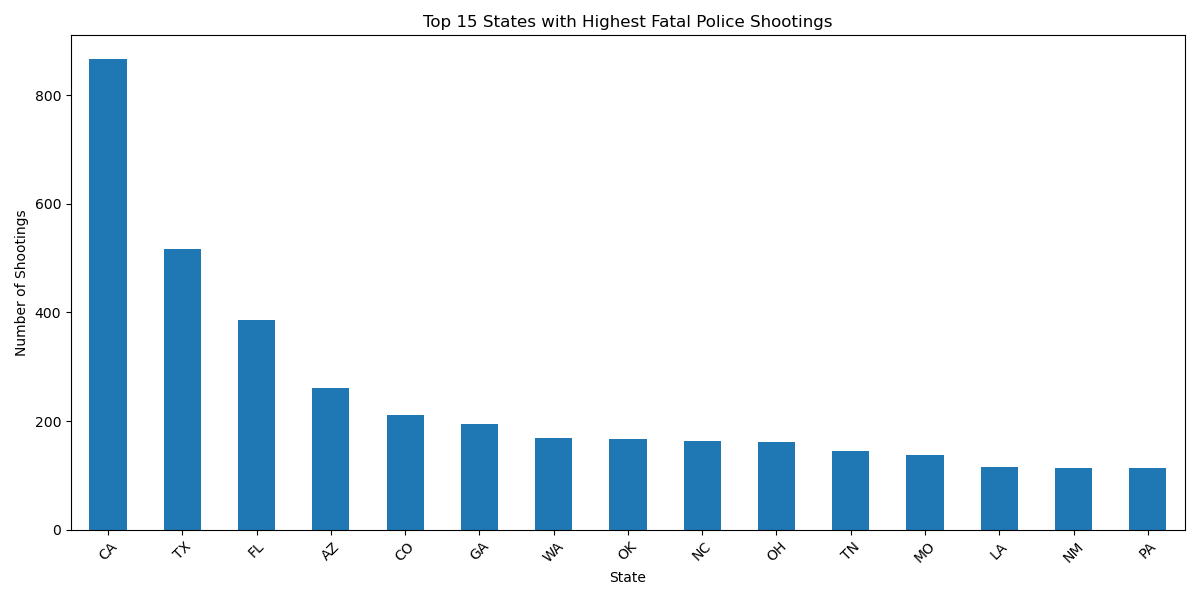


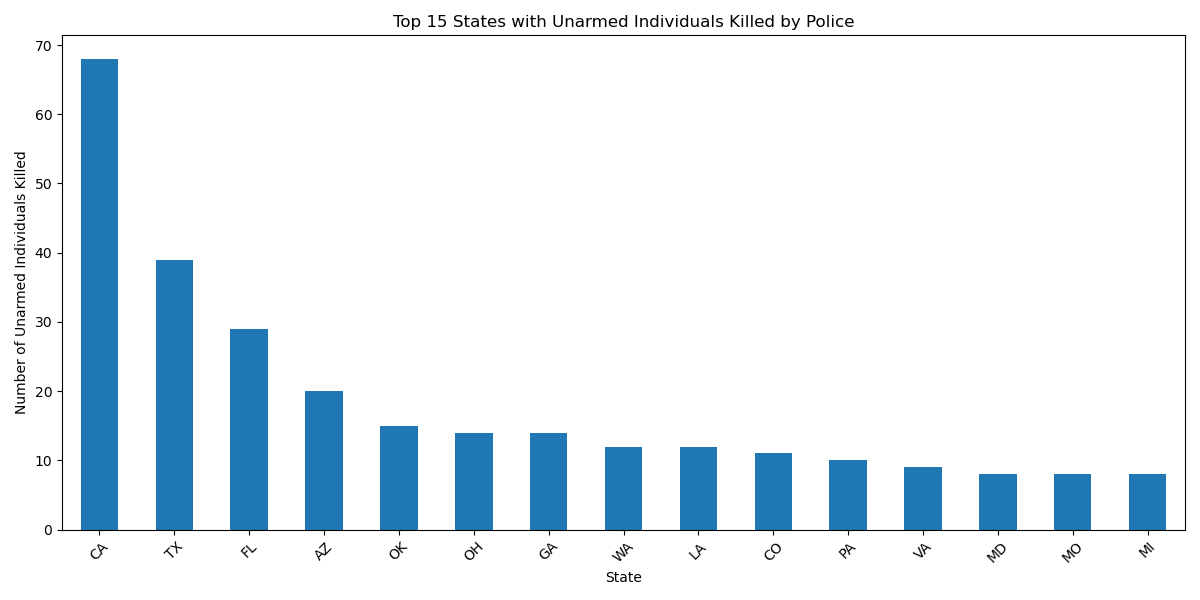
**Geographic Distribution of Incidents**

The geographic analysis of the cleaned data reveals a stark concentration of fatal police shootings in a few states. California leads with a significantly higher number of incidents, followed by Texas and Florida. This pattern remains consistent when analyzing the subset of shootings involving unarmed individuals.

**Top 10 States with Highest Fatal Shootings (Cleaned Data):**

1. **California:** 836
2. **Texas:** 508
3. **Florida:** 375
4. **Arizona:** 274
5. **Colorado:** 204
6. **Georgia:** 192
7. **Oklahoma:** 174
8. **Washington:** 168
9. **North Carolina:** 165
10. **Ohio:** 159

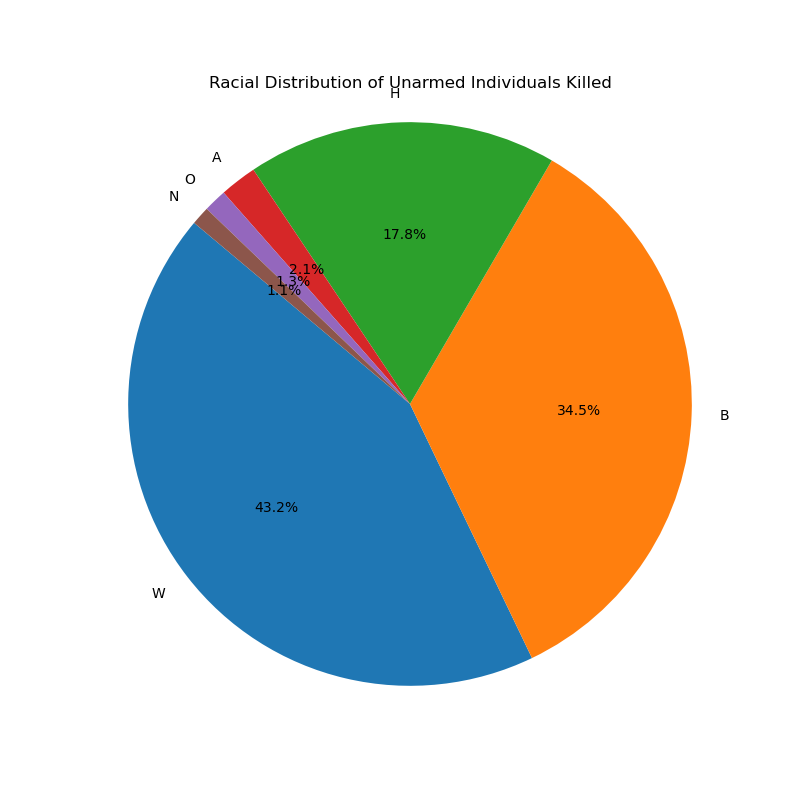


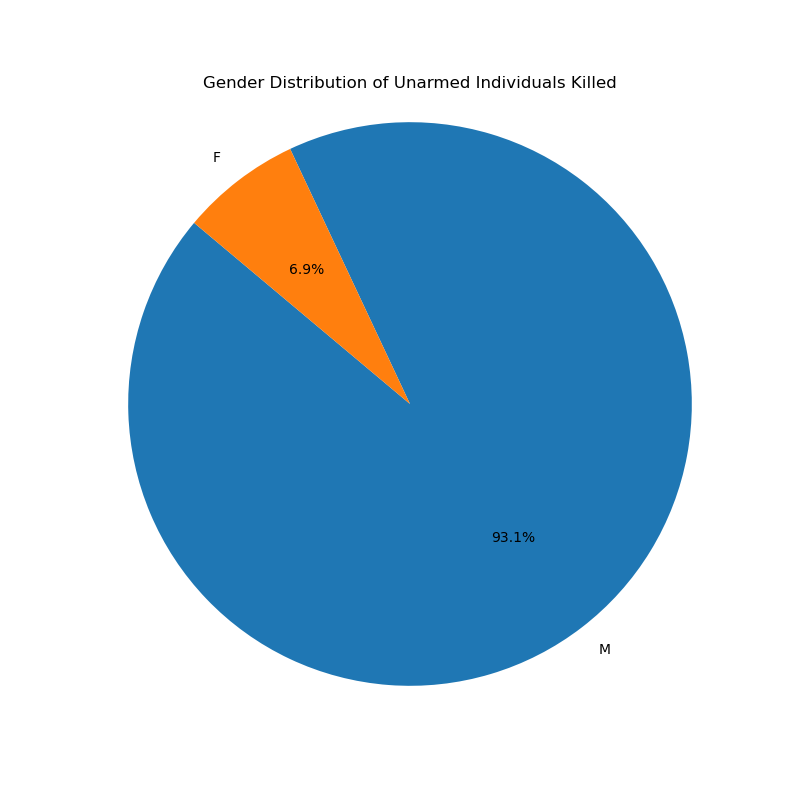


**Victim Demographics (Cleaned Data)**

From the cleaned dataset of 5,635 incidents, **368** individuals were identified as unarmed. The demographic profile of these unarmed victims is as follows:

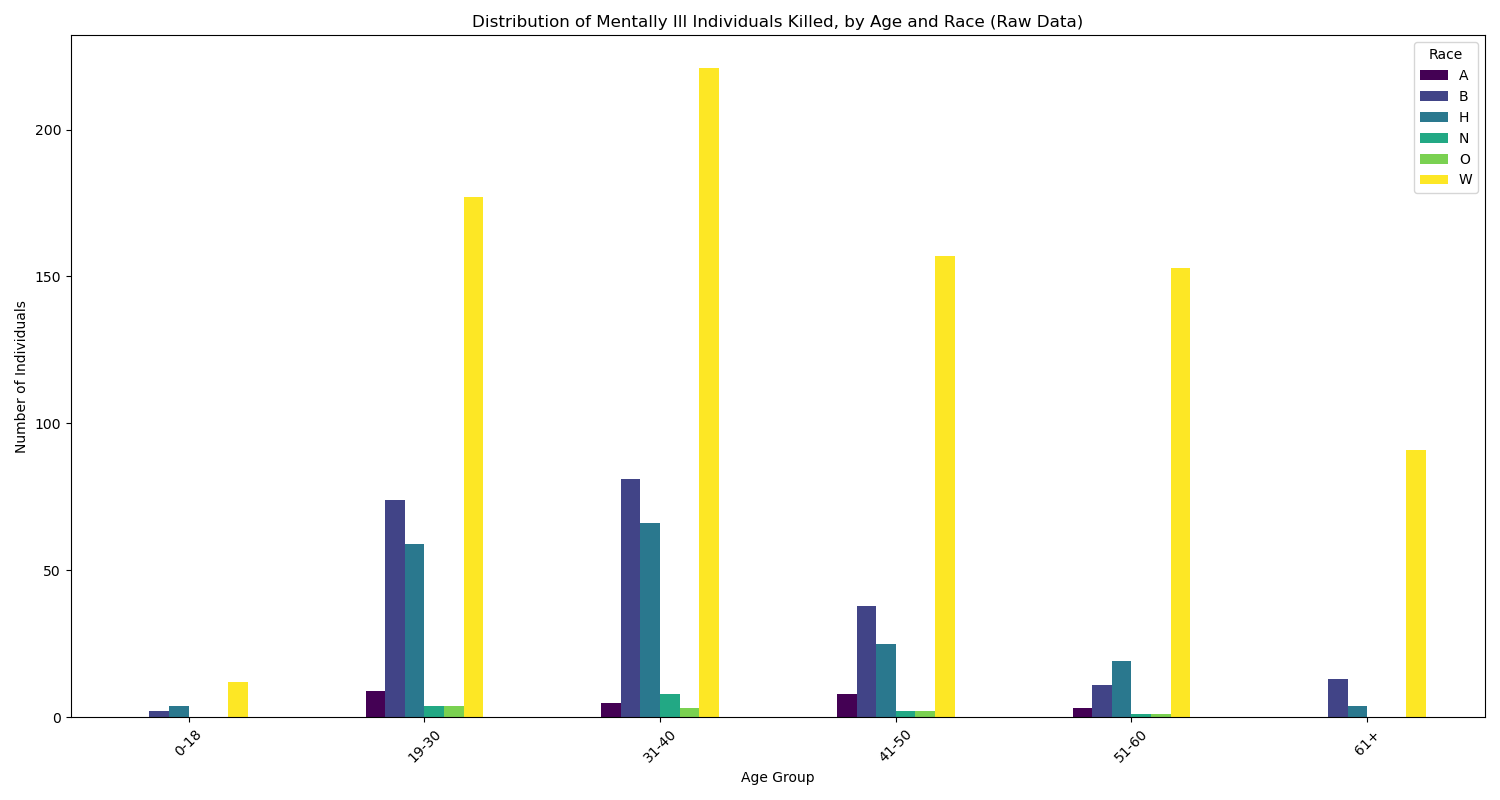
* **Race:** The majority of unarmed victims were White (152), followed by Black (131) and Hispanic (65) individuals.
* **Gender:** The victims were overwhelmingly male (341), compared to 27 female victims.





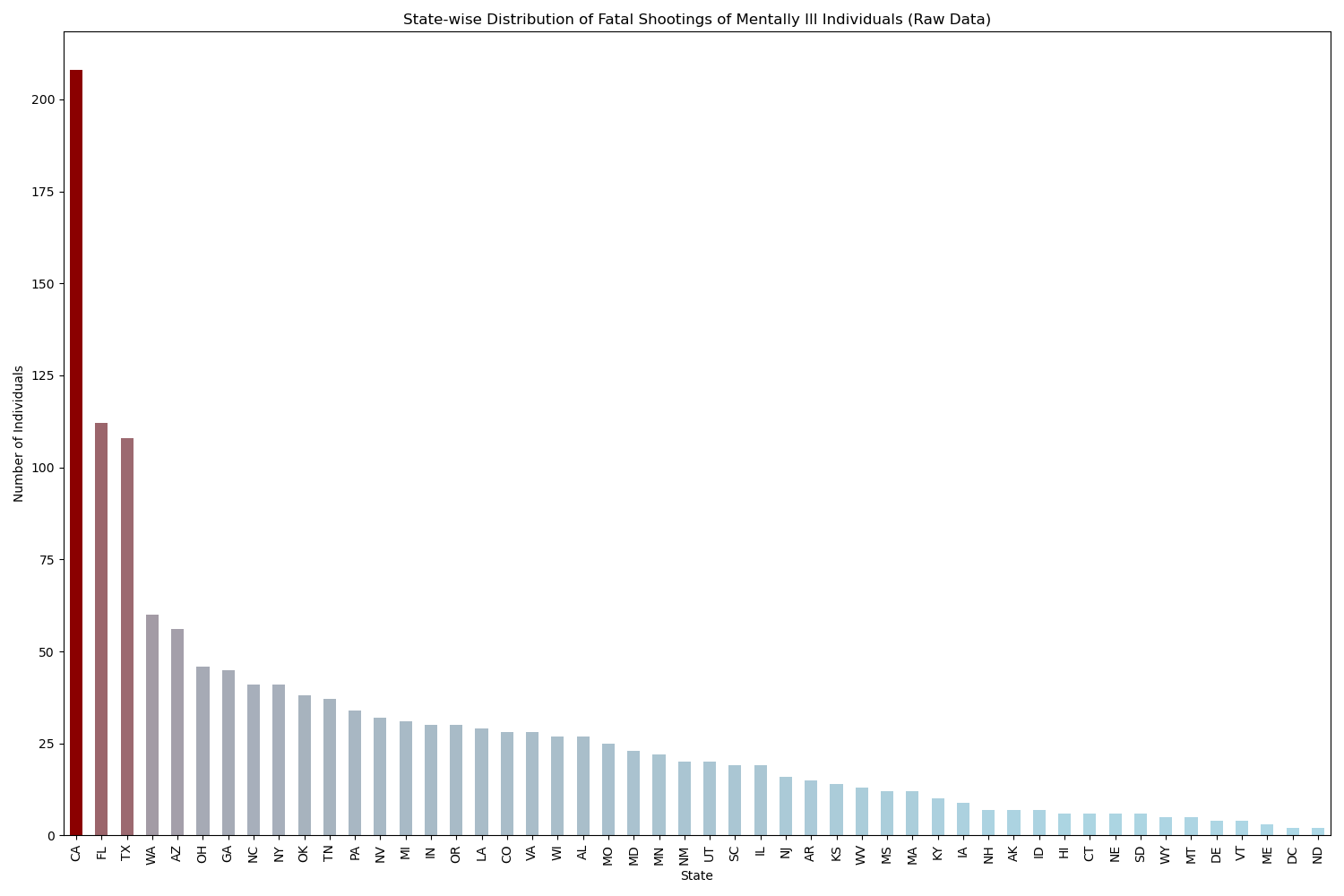
**Analysis of Victims with Signs of Mental Illness (Unprocessed Data)**

This section is based on the original, unprocessed dataset to provide a distinct view of this cohort. A significant portion of victims exhibited signs of mental illness. The demographic breakdown by age and race indicates that White individuals represent the largest group of fatalities across most age categories. The geographic distribution shows a nationwide spread, with notable clusters in densely populated areas.



A graph with blue dots

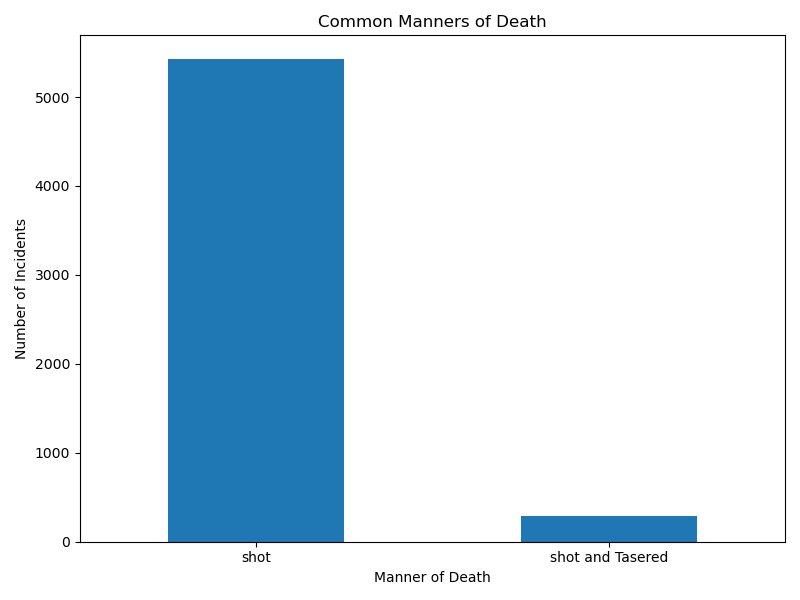
AI-generated content may be incorrect.



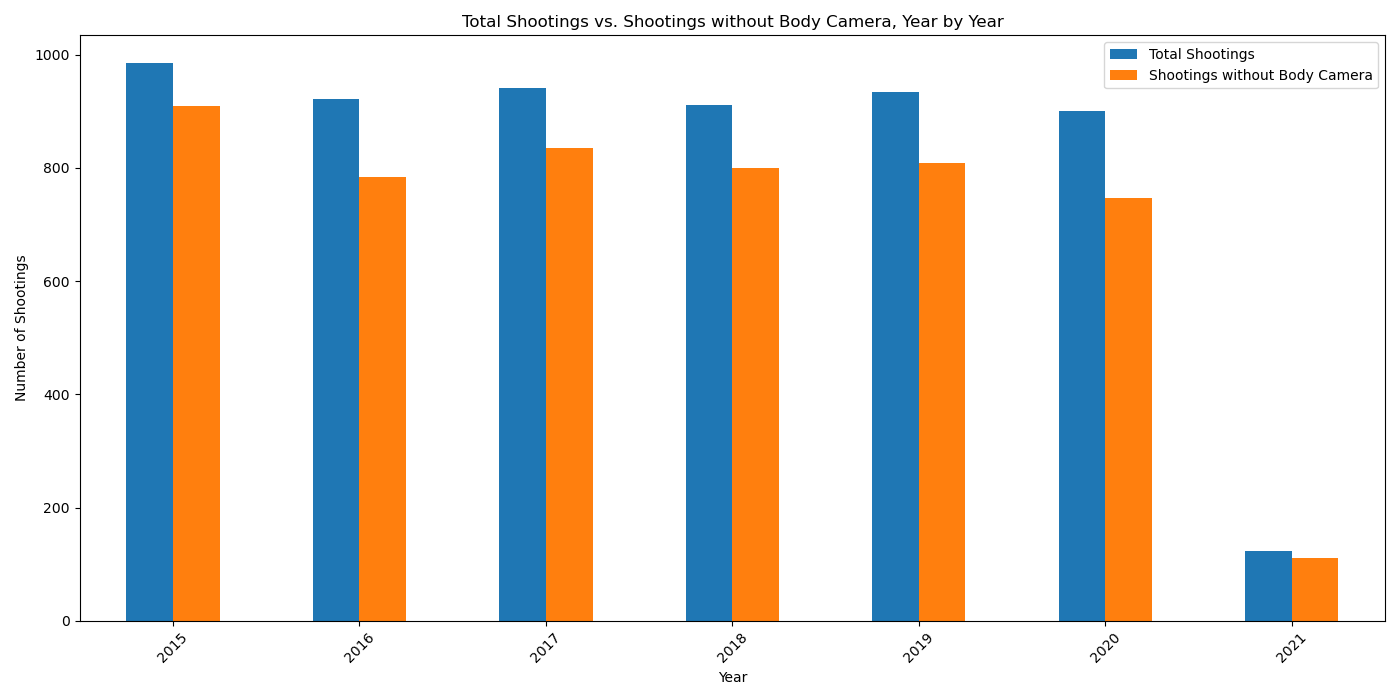
**Circumstances of Shootings (Cleaned Data)**

The analysis of the circumstances surrounding the shootings provides the following insights:

* **Manner of Death:** The predominant manner of death was by shooting. A smaller subset of incidents also involved the use of a Taser.



* **Body Camera Usage:** A substantial number of fatal shootings occurred without the presence of a body camera. However, a year-over-year analysis shows a clear and consistent decline in such incidents, pointing to the increasing adoption of body cameras by law enforcement agencies.



**Conclusion**

The analysis of the fatal police shootings dataset, both preprocessed and raw, offers several key takeaways. The incidents are not uniformly distributed across the nation but are concentrated in specific states. The demographic data of unarmed victims highlights significant racial and gender disparities. The trend in body camera usage is encouraging and suggests a move towards greater transparency. The separate analysis of victims with signs of mental illness underscores the importance of considering mental health factors in these encounters. This report provides a foundational, data-driven perspective on a complex and critical issue.