**Report on Fatal Police Shootings Dataset Analysis**

**-Nishant Rajora 23csu220**

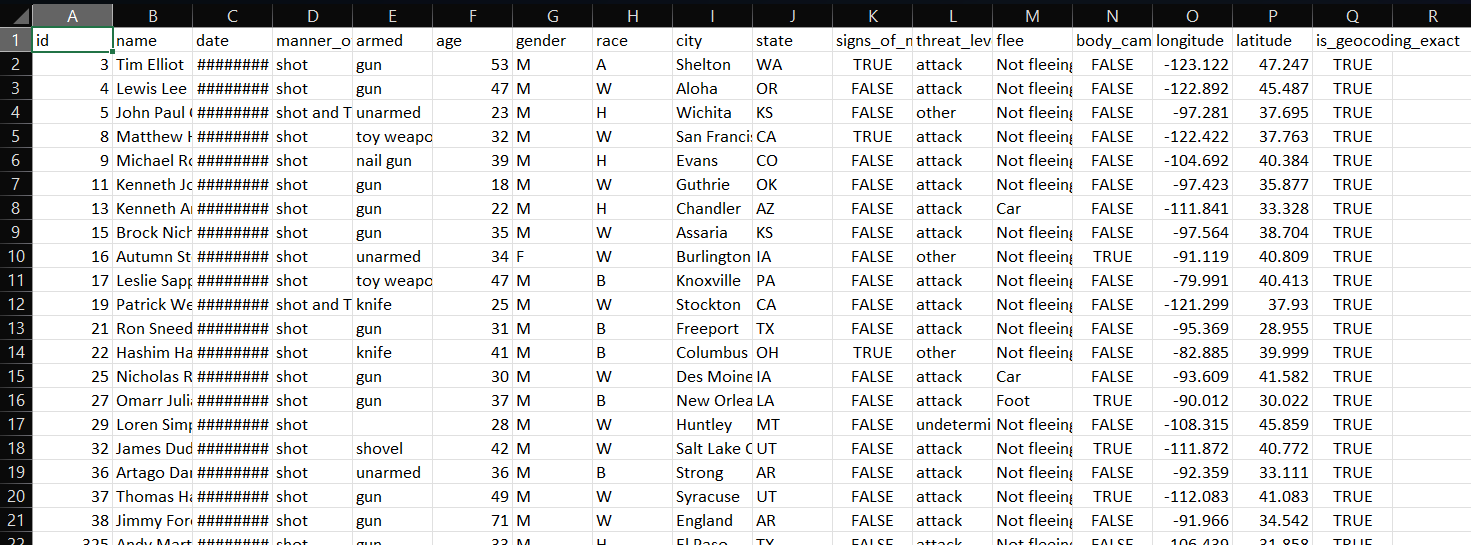
**1. Data Preprocessing**

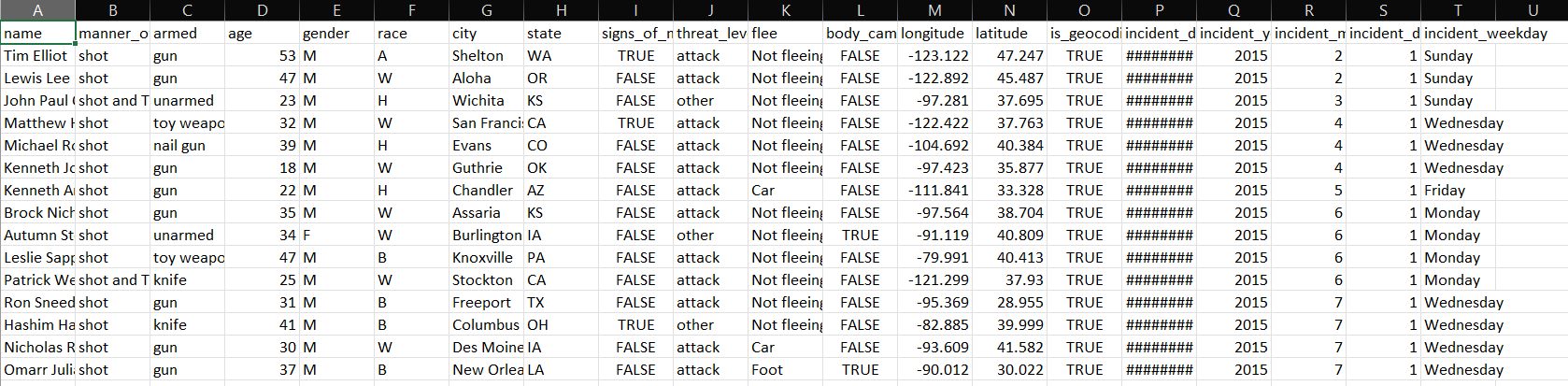
Before starting the analysis, the dataset required cleaning and preprocessing to ensure accuracy and consistency in results. The raw dataset contained missing values, inconsistent column names, and unnecessary fields that could affect further analysis. The following preprocessing steps were applied:

**Steps Performed**

1. **Loading Data**  
   The dataset (fatal-police-shootings-data.csv) was loaded using the pandas library.
2. **Handling Missing Values**
   * Rows with missing values were removed to maintain data integrity.
   * A full copy of the cleaned dataset was created to avoid altering the original file.
3. **Standardizing Column Names**
   * Column names were converted to lowercase.
   * Dots (.) were replaced with underscores (\_) for consistency.
4. **Dropping Unnecessary Columns**
   * The id column was dropped since it does not contribute to analysis.
5. **Date Conversion**
   * The date column was converted into a datetime object.
   * Renamed to **incident\_date** for clarity.
   * Invalid or missing date entries were removed.
6. **Feature Engineering (New Columns)**  
   From the cleaned incident\_date column, new features were extracted:
   * **incident\_year** (year of the incident)
   * **incident\_month** (month of the incident)
   * **incident\_day** (day of the incident)
   * **incident\_weekday** (day of the week of the incident)
7. **Saving Cleaned Data**  
   The cleaned dataset was saved as **fatal-police-shootings-cleaned.csv** for further analysis.

**Screenshots**





**2. Shootings per Year**

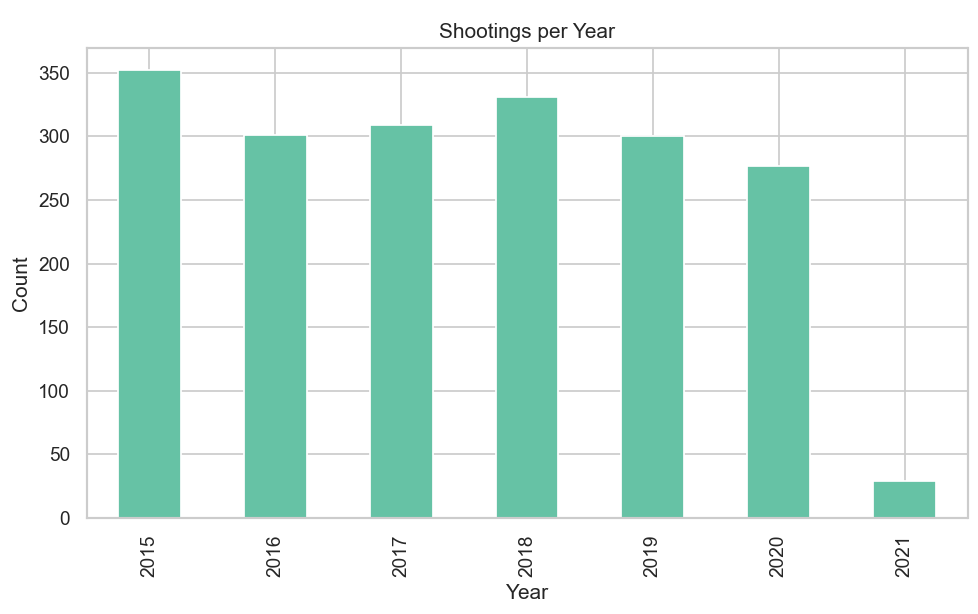
**Observation**

The bar chart presents the number of fatal police shootings recorded each year.

* **2015** witnessed the **highest count** (over 350 incidents).
* From **2016 to 2018**, shootings remained **consistently high**, reflecting little improvement in the trend.
* A **gradual decline** is noticeable after **2018**, suggesting some reduction in incidents.
* **2020** still reported a high number, though slightly lower than the earlier peak years.
* **2021** shows a **sharp decrease**, but this may be due to **incomplete data collection** or reporting delays, rather than a true decline.

**Possible Reasons**

1. **Increased Public Awareness & Protests** – Events such as the *Black Lives Matter* movement (particularly post-2016) could have pressured law enforcement agencies to adopt reforms.
2. **Policy & Training Changes** – Some states may have introduced stricter guidelines on use of force, de-escalation training, and mandatory reporting.
3. **Data Collection Lag in 2021** – The significant dip in 2021 is more likely due to **reporting gaps** rather than an actual fall in incidents.
4. **External Factors** – The COVID-19 pandemic in 2020–21 may have influenced both police activity patterns and the way incidents were reported.



**3. Shootings per Month**

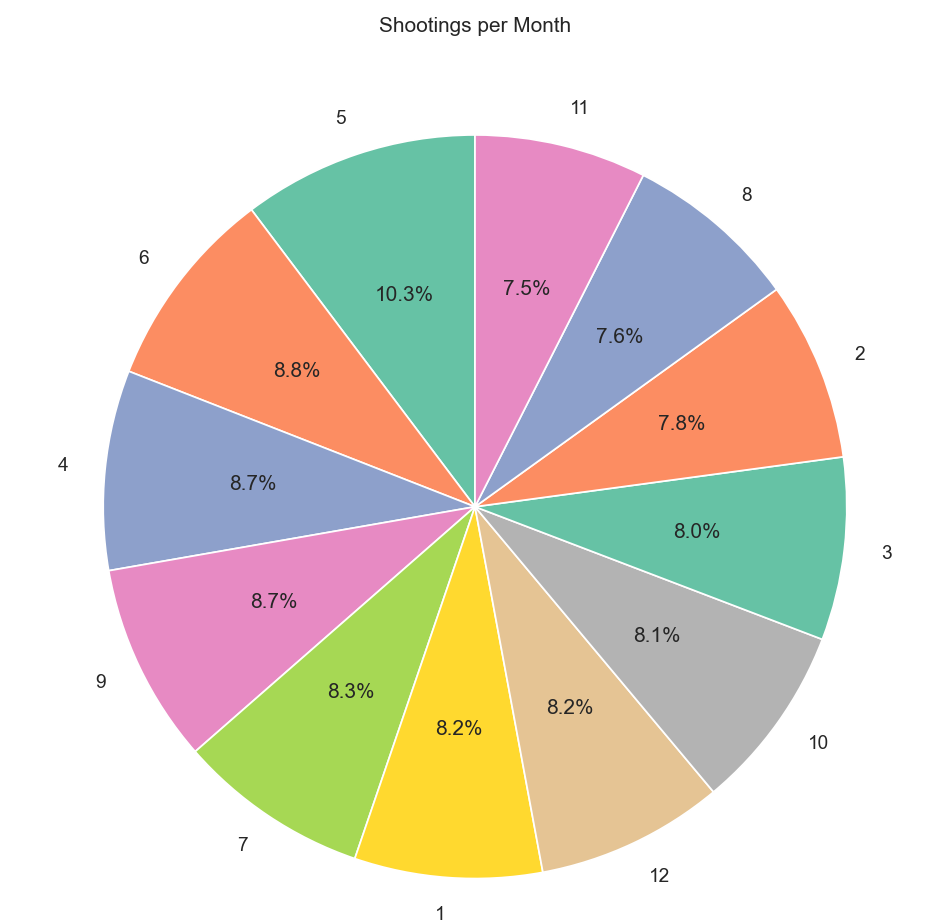
**Observation**

The pie chart shows the distribution of police shootings across different months.

* **May (10.3%)** records the **highest proportion**, indicating more incidents occur during this period.
* Other months like **June (8.8%)**, **April (8.7%)**, and **September (8.7%)** also report slightly higher shares compared to the average.
* **November (7.5%)** has the **lowest proportion** of shootings.
* Overall, the distribution is relatively **even across months**, with no single month dominating the trend.

**Possible Reasons**

1. **Seasonal Factors** – Warmer months (spring/summer) such as **May–September** often see higher outdoor activities, social gatherings, and crime rates, which may increase police interactions.
2. **Holiday & Festive Periods** – Months with fewer public events (e.g., November) might naturally report lower incidents.



**4. Age Distribution by Weekday**

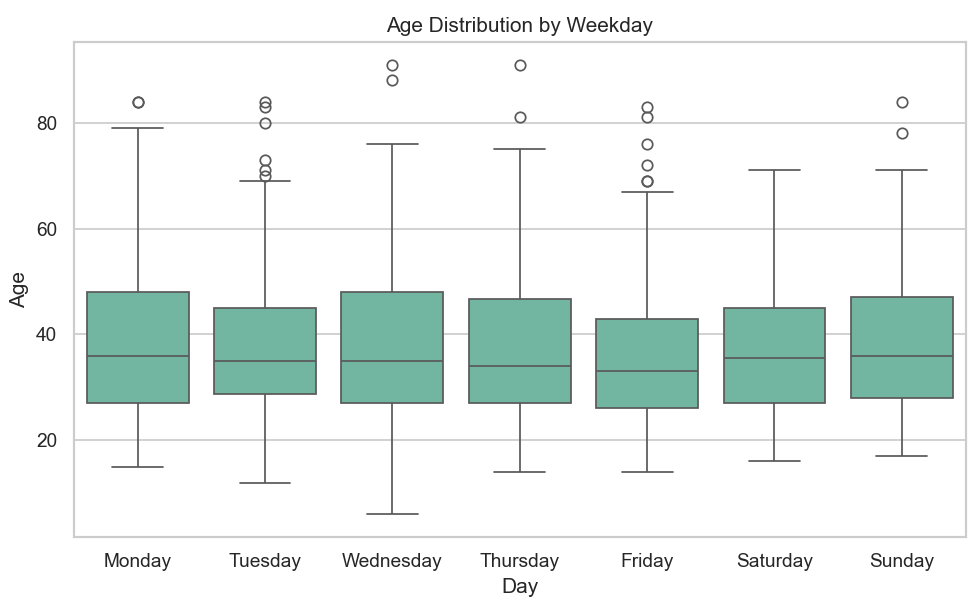
**Observation**

The boxplot illustrates the **age distribution of individuals involved in police shootings across weekdays**.

* The **median age** across all days is consistently in the **early-to-mid 30s**.
* Most cases fall within the **25–45 age range**, with interquartile ranges (IQR) overlapping across all days.
* Outliers exist in all weekdays, with some individuals aged **70–90+**.
* **Friday** appears to have slightly younger median cases compared to other days.
* No significant variation between weekdays, indicating that **age distribution is fairly stable throughout the week**.

**Possible Reasons**

1. **Demographics of Police Encounters** – Younger adults (20s–40s) are more likely to be involved in situations that escalate into police shootings due to higher mobility and interactions with law enforcement.
2. **Consistency Across Weekdays** – Since daily routines and policing intensity remain similar throughout the week, no weekday-specific trends emerge.
3. **Presence of Outliers** – Older individuals involved may represent exceptional cases (e.g., mental health crises, isolated incidents).



**5. Top 10 Types of Weapons Used**

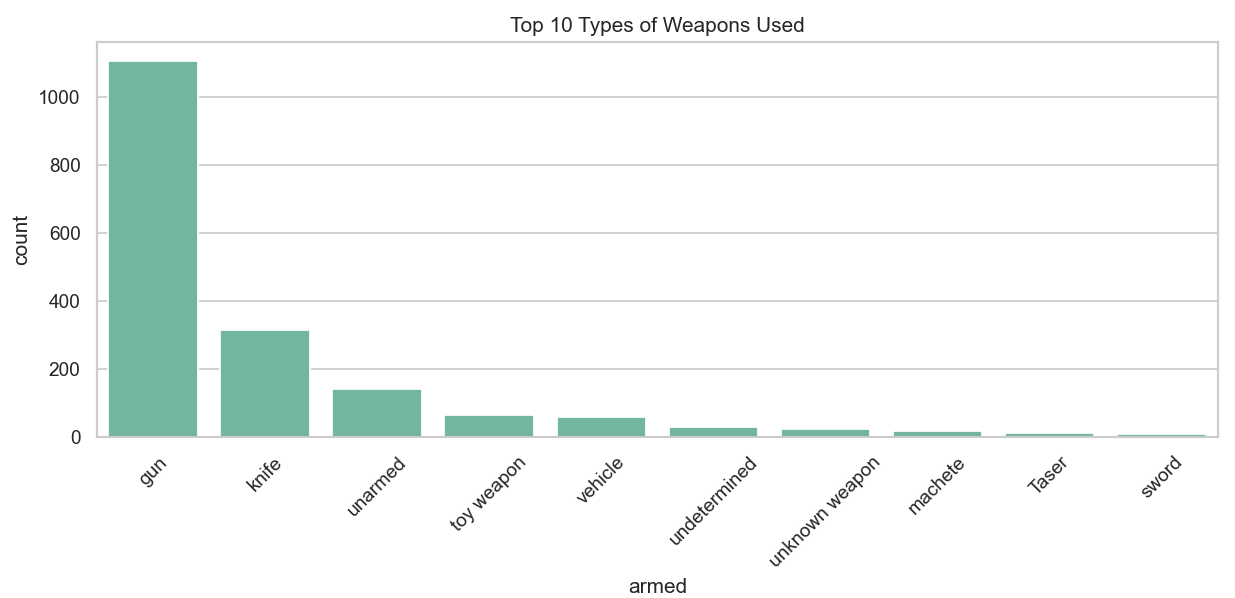
**Observation**

The bar chart highlights the **distribution of weapons involved in police shootings**:

* **Guns** dominate overwhelmingly, with more than **1,000 cases**, making them the most common weapon used.
* **Knives** are the second most frequent, but far less common than guns (approx. 300 cases).
* A significant number of incidents involve **unarmed individuals** (150+ cases), showing that not all shootings occur against armed suspects.

**Possible Reasons**

1. **High Prevalence of Firearms in the U.S.** – The widespread accessibility of guns directly explains why most police shootings involve them.
2. **Knives as Secondary Threat** – Knives are common household items and often present in domestic or street confrontations, leading to their ranking second.
3. **Unarmed Cases** – These incidents could reflect misidentification of threats, high-stress decision-making, or systemic issues in policing.
4. **Toy Weapons / Vehicles** – Misinterpretations (toy guns resembling real ones) or confrontations involving vehicles (used aggressively) can escalate into shootings.
5. **Rare Weapons** – Items like swords, machetes, or tasers are uncommon, representing isolated or unusual encounters.



**6. Fleeing Status at Time of Incident**

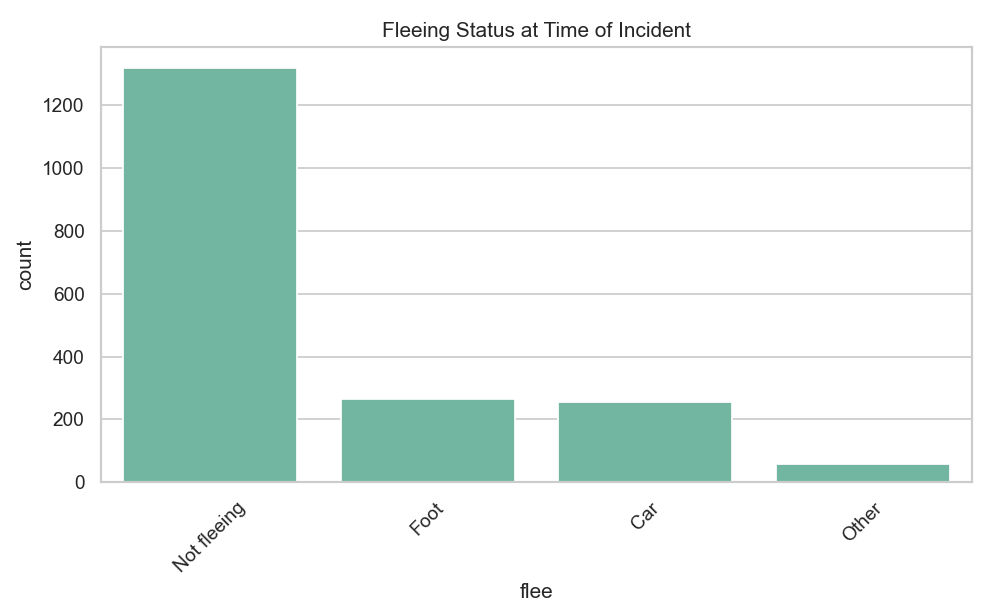
**Observation**

The bar chart illustrates the **fleeing status of individuals during police shootings**:

* A vast majority of cases (**over 1,200 incidents**) involve suspects who were **not fleeing** at the time of the incident.
* Both **foot pursuits** and **vehicle-related flights** account for a similar number of cases (around 250 each).
* A small category labeled **“Other”** represents a minimal share (< 50 cases).

**Possible Reasons**

1. **Majority Not Fleeing** – This suggests that most police shootings occur in **stationary confrontations** (e.g., during domestic incidents, direct confrontations, or standoffs), rather than in pursuit scenarios.
2. **Flight by Foot or Car** – Pursuits are inherently high-risk, leading to escalated tensions and potential misjudgments. Individuals fleeing might be perceived as **greater threats**, especially if suspected of being armed.
3. **“Other” Cases** – Likely include less common scenarios (e.g., hiding, resisting arrest without fleeing, or ambiguous situations).
4. **Policy & Training Impact** – In many regions, police training emphasizes lethal force when escape poses a threat to public safety, which may explain why fleeing suspects still represent a notable fraction.



**7. Top 10 Cities by Shootings**

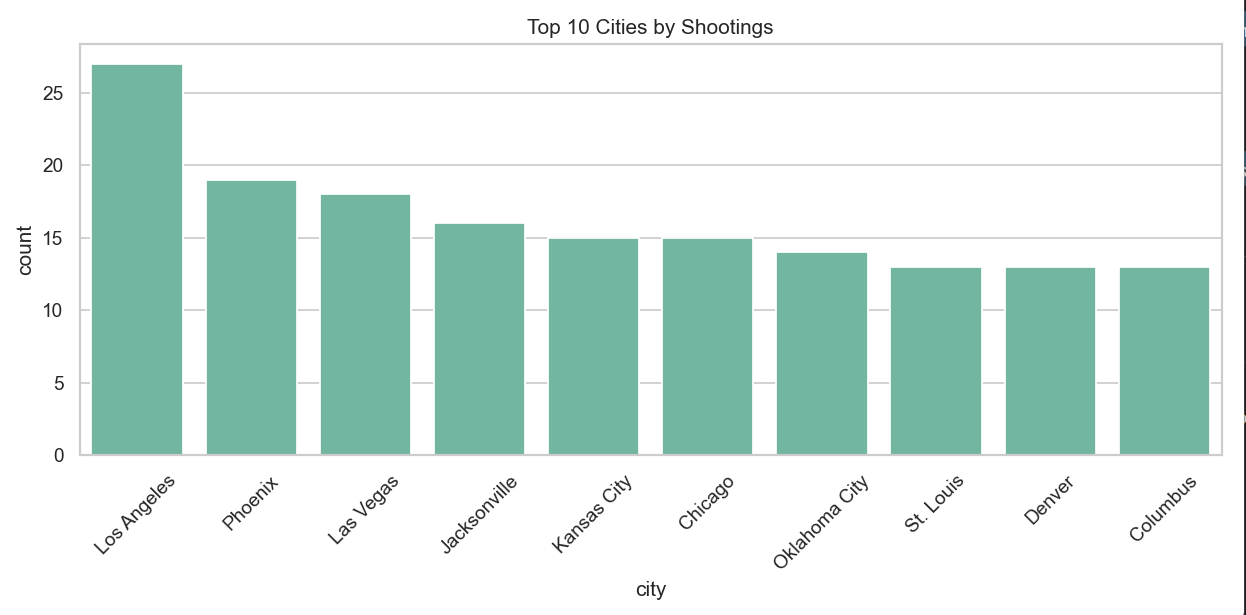
**Observation**

This bar chart highlights the **cities with the highest number of police shootings**:

* **Los Angeles** stands out significantly with **over 25 shootings**, far more than any other city.
* **Phoenix** and **Las Vegas** follow, with **~18–19 incidents** each.
* Other cities like **Jacksonville, Kansas City, and Chicago** each report around **15 incidents**.
* Cities like **Oklahoma City, St. Louis, Denver, and Columbus** round out the top 10, with around **13–14 shootings** each.

**Possible Reasons**

1. **Los Angeles as Outlier** – Being the largest city in California with a massive police force (LAPD), higher population density, and a long history of gang activity and policing issues, LA unsurprisingly tops the list.
2. **Southwest & Western Cities** – Phoenix, Las Vegas, and Los Angeles all have relatively high rates of violent crime and larger metropolitan areas, which may explain their presence at the top.
3. **Urban Crime Centers** – Chicago, St. Louis, and Kansas City are historically recognized for violent crime rates, making them hotspots for police encounters that escalate.



**8. Police Shootings per Million People by Race**

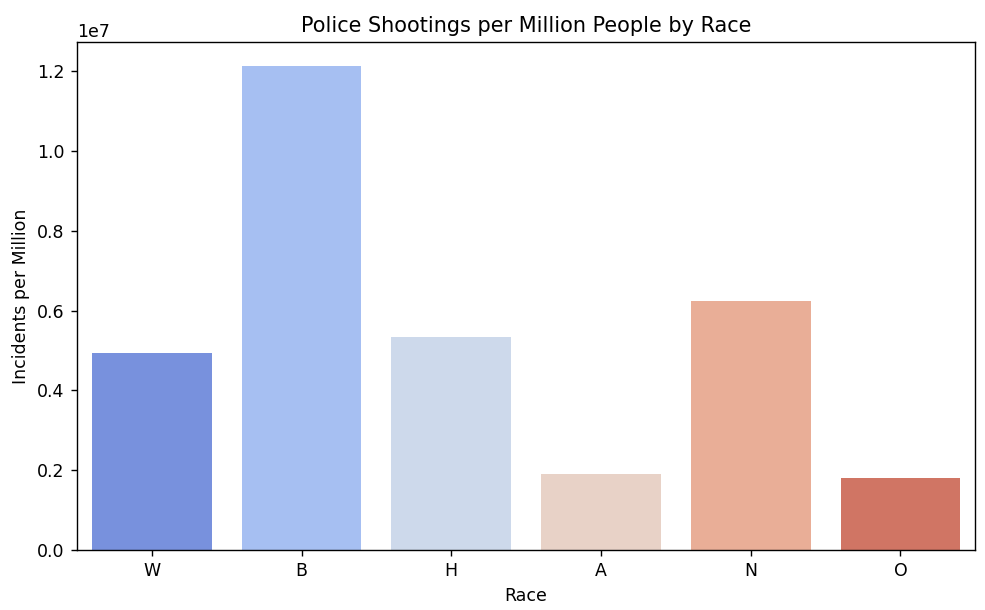
**Observation**

This chart shows the **rate of police shootings per million people across racial groups**:

* **Black (B)** individuals face the **highest rate**, exceeding **12 million per million people**, which is vastly disproportionate compared to other groups.
* **Native American (N)** populations also face a **high rate (~6.2 million per million)**, despite being a smaller population group.
* **Hispanic (H)** and **White (W)** individuals have **moderate rates** (~5–5.3 million per million).

**Possible Reasons**

1. **Systemic Inequities** – The disproportionately high rate for **Black** individuals may reflect structural racism in policing, socioeconomic disparities, and historical tensions between law enforcement and minority communities.
2. **Native American Overrepresentation** – Despite a small population, the high rate for **Native Americans** might be linked to rural policing practices, jurisdictional conflicts on reservations, and higher levels of poverty and substance abuse issues in certain areas.
3. **Population Size vs. Rate** – While more **White** individuals may be shot in absolute numbers (because they are the largest population group), the **rate per million** highlights that minorities are disproportionately affected.



**Conclusion**

The analysis of fatal police shootings highlights clear **demographic disparities**:

* **Men**, especially in the **18–44 age group**, are overwhelmingly more likely to be victims.
* **Black** and **Native American** communities face **disproportionately higher shooting rates** compared to their population share.
* **Urban centers** (like Los Angeles, Houston, and Phoenix) account for a significant portion of cases, but disparities exist nationwide.

These patterns suggest that **systemic, social, and situational factors**—such as socioeconomic inequalities, racial profiling, and concentrated policing in urban areas—play a major role in shaping outcomes.