

Fatal Police Shootings – Report

-Piyush Chandra

About Data: The different columns in dataset are as follows:

- **id:** a unique identifier for each victim
- **name:** the name of the victim
- **date:** the date of the fatal shooting
- **manner_of_death:** shot, shot and Tasered
- **armed:** had something that could inflict harm
 - undetermined: armed or not, not confirmed
 - unknown: armed, object unknown
 - unarmed: not armed
- **age:** the age of the victim
- **gender:** the gender of the victim.
 - M: Male
 - F: Female
 - None: unknown
- **Race:**
 - W: White, non-Hispanic
 - B: Black, non-Hispanic
 - A: Asian
 - N: Native American
 - H: Hispanic
 - O: Other
 - None: unknown
- **city:** the municipality where the incident took place.
- **state:** two-letter postal code abbreviation. the state in which the incident took place.
- **signs of mental illness:** whether the victim had signs of mental illness at the time of the shooting or before it (chronic)
- **threat_level:** The attack category is meant to flag the highest level of threat. The other and undetermined categories represent all remaining cases. Other includes many incidents where officers or others faced significant threats.
- **flee:** whether victim was moving away from officers
 - Foot
 - Car
 - Not fleeing
- **body_camera:** officer wearing body camera or not, at the time of incident.

Data Cleaning:

Steps involved in data cleaning process:

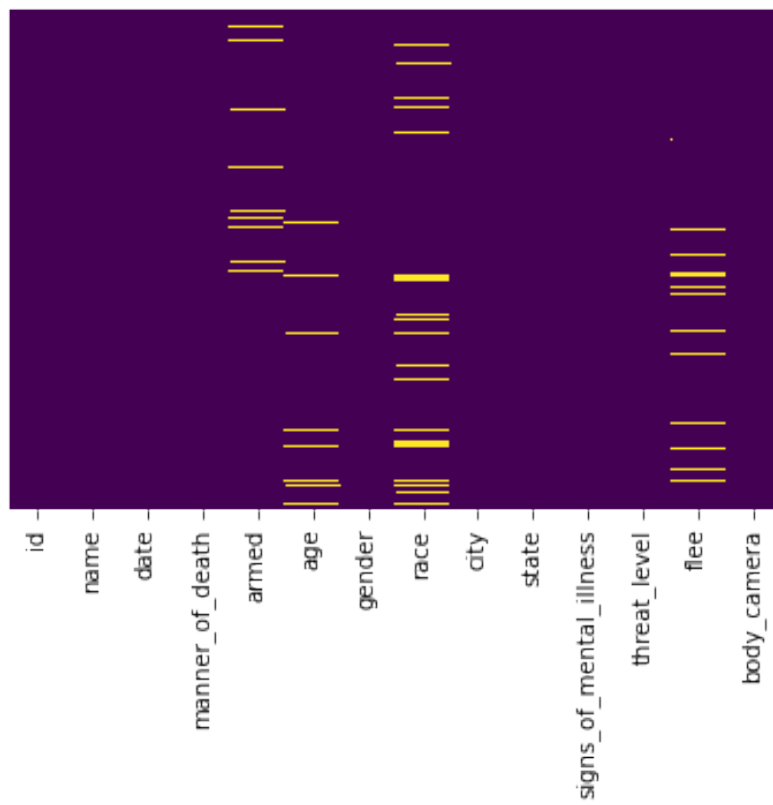
1. Dropped 'name' column. Was not useful in analysis.
2. Null values found in five columns:
 - a. armed - 227
 - b. age - 235
 - c. gender - 2
 - d. race - 521
 - e. flee - 249
3. Checked for duplicate values; found none.
4. Value entries were starting from index 3, sorted them in appropriate index starting from 0.
5. Date column dtype parsed into datetime from object.
6. Finding mode of data columns 'flee' and 'armed'.
7. Replaced null values in 'flee' and 'armed' with the mode of respective columns.
8. Null rows in 'flee', 'gender' and 'armed' are dropped.
9. Check for null values again; none found.
10. Dataset has 4809 rows and 14 columns.

Inferences:

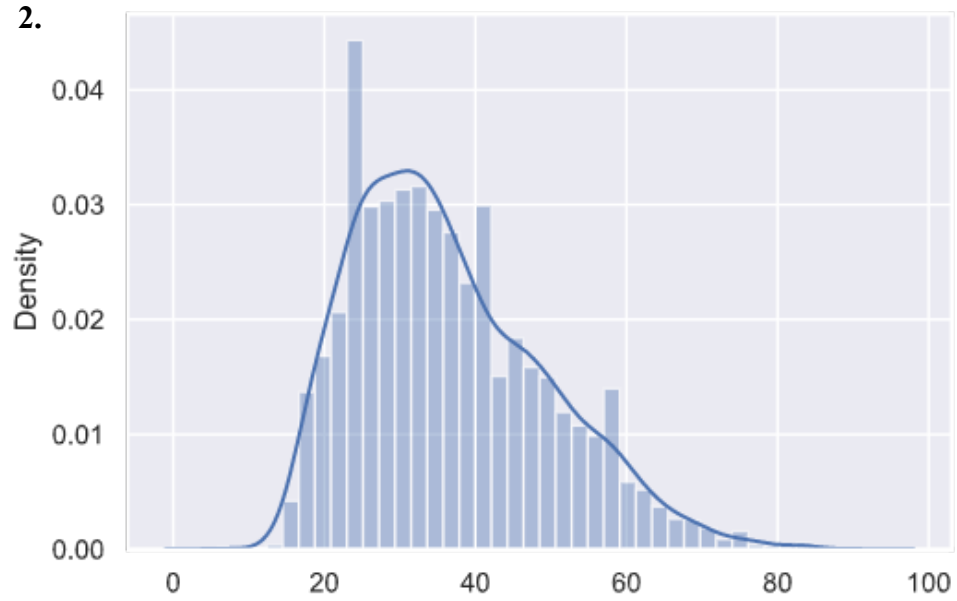
1. Mean age of all the people that died in police shootings is 37.
2. Majority of people who died were armed with a 'gun'.
3. The top city with most incidents is Los Angeles and the top state is California(CA).
4. Race 'W' were the people who had majority encounters.
5. More than 50% of all the shootings have happened with Race 'W' while more than 26% have happened with Race 'B'. And Race 'H' were involved 18% encounters.
6. A big majority of 43% of ALL the encounters were marked attacked where the victim was not trying to flee.
7. California(CA) is the state with most encounters at 14%. And Texas next with 8.6%.
8. Less than 100 incidents happened where the victim had undetermined threat with mental illness where as major victims possessing threat had no mental illness.

Visualisations:

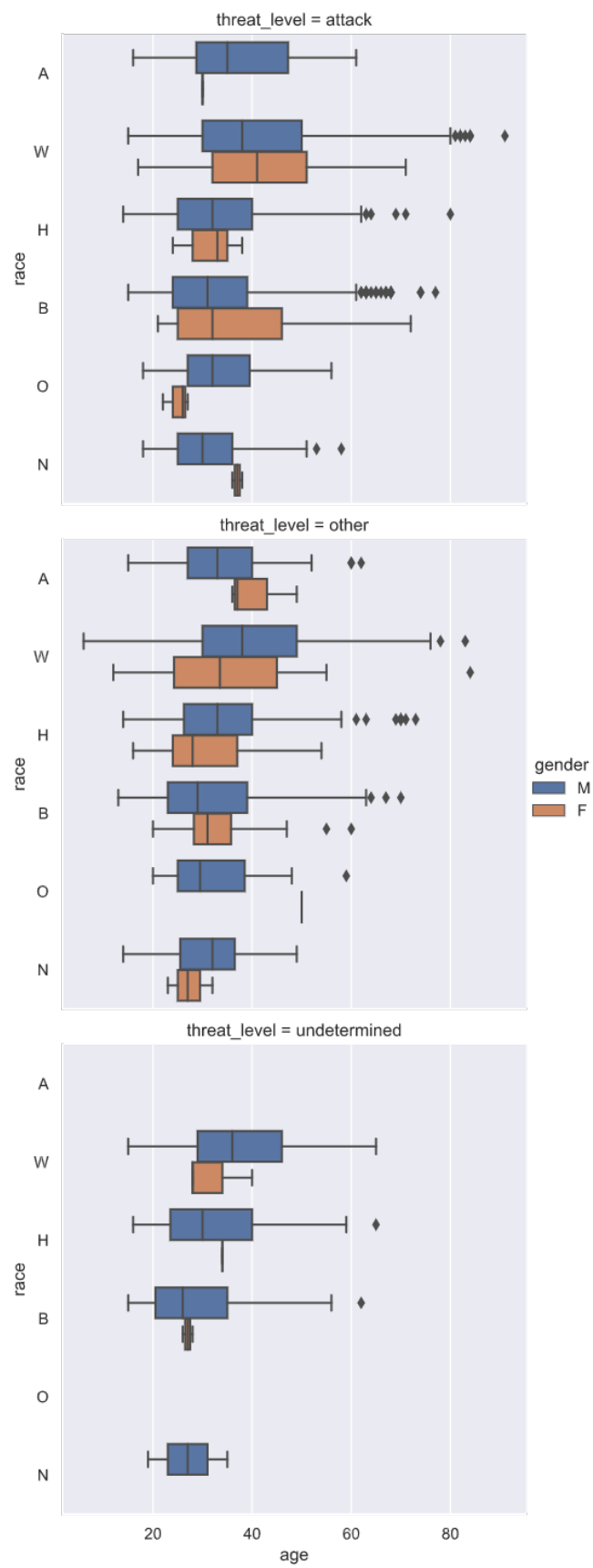
1.



2.



3.



4.

