**Question:** Is it possible that some visualizations can miscommunicate the actual information and be biased towards a specific community/race?

**Motivation:** The burden of fatal police violence is an urgent public crisis to be addressed in USA. Visualizations done without proper normalization can mask the racial bias happening to people of certain races. I have gone through many articles which provide mounting evidence that shows that the fatalities caused by police shootings disproportionately impact people of certain races. I have taken this point to analyze how some visualizations can mislead us into interpreting the same inferences in a different way. I also provided visualization that gives evidence into the actual insights after normalizing the data correctly with the total population.

**Types of Visualizations used:**

**Bar charts with Long Format data: (using plotly)** I have used an interactive bar chart with long format data to visualize the “Year wise distribution of victim’s race”. I have chosen this type of visualization type because it can effectively communicate both the count and distribution of all fatalities according to races for different years in a single graph.

**Bar plot: (using matplotlib)** I have used standard bar plot using matplotlib to visualize the ‘Percentage of deaths from police shootings’ compared to ‘percentage of population by race’. I used this visualization to know the percentages of race explicitly in comparison with the percentages of total population by race.

**Size, Scale and color:** Sizes of both the plots has been chosen in such a way that it is clear and legible. Both the font and scale are set in order to ensure better readability. Various colors are used to distinguish multiple races in the database and the choice of colors are chosen in such a way that the plots are pleasant on the eye.

**Inferences:** It can be observed from first visualization made using plotly that more fatalities are observed in White race across all the years. So, it is bound to understand from both the visualizations that white race is affected majorly in the fatal shootings by police. But it is not the actual scenario. If we take the recent census data and compare with percentage of total population by race, it can be observed that 21.98% of the total fatalities are corresponding to Black race where they comprise of only 12.16% of the total population. This discussion can be extended to all the races.

Following the similar observation, we can conclude that both black people and native people are disproportionately killed by the police. We can also see that for Asians and white people, the percentages are far below that of their overall percentages of the population in the United States.

**Pros:**

* It can be observed in all the plots that Tufte’s principle is strictly followed to maximize the data-ink ratio. Non data ink is erased and every pixel in frame communicates some data variations.
* The point of showing the racial bias that is hidden microscopically in the database and how wrong inferences can made by making visualization using normalization with total population is communicated effectively by comparing both kinds of visualizations.

**Cons:**

* Only insights regarding the racial bias have been used for analysis from the database. Other factors have not been exploited for better insights into the data.