<u>Ethical Issues Regarding Data Collection</u>

For our project, we chose to analyze a data set from the FBI's National Incident-Based Reporting System regarding the presence of hate crimes. We selected data detailing all reported incidents of hate crime in the U.S. from 2000 to 2017; columns included the state, city, location of crime, race/ethnicity of offender, number of victims, race of victims, and a description of the bias (anti-Asian, anti-Gay, etc.), among other variables. Additionally, we merged this data with another data set which contained the populations of each city, so that we could obtain the hate crime rate with respect to population (the FBI data set was structured such that each hate crime received one row).

Naturally, the collection of our data set raises many ethical issues. First, whether a hate crime is an act of discrimination or an arbitrary act irrespective of race, gender, sexuality, etc. can be subjective and prey to the bias of the law enforcement officer who fills out the report. Reasonable people often disagree as to the extent an incident is due to discrimination or chance. There could be false positives or false negatives (where an ordinary crime was mistakenly reported as a hate crime, or vice versa) depending on the person who submitted the report. Additionally, what type of crime a hate crime is classified as is also subjective to the reporter. For example, depending on the officer's bias or identity(s), a hate crime that was classified as an assault by one officer could have been reported as merely an intimidation by another.

Therefore, there are many ethical issues regarding the collection of this dataset. We do not know the biases behind those who filled out and submitted their reports- their choice to submit and what they submitted is all subject to individual biases.

<u>Ethical Concerns about Findings</u>

The integrity of our analysis of hate crimes is dependent on the integrity and reliability of the dataset. As mentioned previously, how a hate crime is reported may be affected by the reporter's individual bias. Furthermore, there was not hate crime data for every city or county in the United States. Is this because those cities had no hate crimes to report or did they choose not to report it? The FBI has a National Incident-Based Reporting System (NIBRS) to keep track of crime incidents across the United States. For NIBRS, only some of local agencies actually report information back. While Arkansas has 100% of agencies reporting to NIBRS, Texas only has 17% of agencies reporting back. While the NIBRS data is not the same as our hate crime data, we wonder if the reporting rate varies across geographical region, affecting the conclusions that we obtain.

In our project, we analyzed the mean number of hate crimes in cities of various sizes based on geographical area - Northeast, South, West, and Midwest. If the hate crime reporting rate of agencies across the United States really does vary, then our conclusions may not be entirely representative. One of our concerns is that certain areas with the worst problems may actually be the least inclined to report. In addition, while areas across the U.S. should apply the exact same standards to evaluating and reporting hate crimes, local biases may cause the hate crime reporting to seem better than it actually is. This would cause our data analysis to give a biased hate crime analysis of an area.

We found that there was a positive correlation between anti-White and anti-Black hate crimes in a city. However, that result should not be mistaken for causation. Some might obtain the conclusion that an increase in anti-White hate crimes cause an increase in anti-Black hate crimes and vice versa. Instead, all we can say is that perhaps there is causation between the two

factors or perhaps there are outside factors (such as higher racial tensions in an area) that affect anti-White and anti-Black hate crime rates.

Source:

https://www.fbi.gov/services/cjis/ucr/nibrs