HATE CRIMES IN THE UNITED STATES (1991-2018)

DATA ANALYST: JULIANA LEE
01.26.2023

OVERVIEW

1. Executive Summary

> The Hate Crime Statistics Program of the FBI's Uniform Crime Reporting (UCR) Program collects data regarding criminal offenses that were motivated, in whole or in part, by the offender's bias against a race, gender, gender identity, religion, disability, sexual orientation, or ethnicity, and were committed against persons, property, or society.

2. Data Source

> Data Source: This data is publicly available. It is an open-source data downloaded from <u>United States Hate Crimes (1991-2018) | Kaggle.</u>

> Data Collection: This data is administrative data. This dataset was retrieved from the <u>FBI: Crime Data Explorer</u>. The data is collected every year.

> Data Contents: The data contains the date, state, offender race, victim count, type of offense, etc. The dataset ranges from 1991 to 2018.

> Data Relevance: This data is relevant and significant to answer the questions of the project.

3. Limitations and Ethics (not done yet)

> The data is meant to be informative. So, the likelihood of the data being biased is rare. However, because motivation is subjective, it is sometimes difficult to know with certainty whether a crime resulted from the offender's bias. Moreover, the presence of bias alone does not necessarily mean that a crime can be considered a hate crime. Only when a law enforcement investigation reveals sufficient evidence to lead a reasonable and prudent person to conclude that the offender's

actions were motivated, in whole or in part, by his or her bias, should an agency report an incident as a hate crime.

The data is not up to date because the data is collected every year, but the last data recorded was in 2018.

4. Data Cleaning and Data Consistency Checks

- ➤ **Adjusted data type.**
- ➤ **Checked for missing data.**
- ➤ **Checked for duplicate data.**
- ➤ **Checked for mixed data type.**

5. Data Profile

Original data consists of 201,403 rows and 28 columns.

New data consists of 199,797 rows and 22 columns.

6. Column Details

Column	Column Description	Data type	Time Variant (Yes? or No?)
INCIDENT_ID	Identification number assigned to incident.	Qualitative, nominal	No
DATA YEAR	The data is from 1991-2018.	Qualitative, ordinal	No
ORI	State abbreviations followed by numerical values assigned to	Qualitative. ordinal	No

	each incident id.		
PUB_ AGENCY_NAME	Name of agency.	Qualitative, nominal	No
AGENCY_TYPE_NAME	Type assigned to agency.	Qualitative, ordinal	No
STATE_ABBR	Abbreviations of the state.	Qualitative, nominal	No
STATE_NAME	Name of the state.	Qualitative, nominal	No
DIVISION_NAME	Name of the division.	Qualitative, nominal	No
REGION_NAME	Name of the region.	Qualitative, nominal	No
POPULATION_GROUP_CODE	The range of the population codes.	Qualitative, ordinal	No
POPULATION_GROUP_DESC	Population assigned to each population group code.	Qualitative, ordinal	No
INCIDENT_DATE	The format of when the incident happened is MM-DD-YYYY.	Qualitative, ordinal	No
TOTAL_OFFENDER_COUNT	How many offenders there were for the incident.	Quantitative, discrete	Yes
OFFENDER_RACE	Race of the offender.	Qualitative, nominal	No
VICTIM_COUNT	How many victims there	Quantitative, discrete	Yes

	was for the incident.		
OFFENSE_NAME	Name of the offense.	Qualitative, nominal	No
TOTAL_INDIVIDUAL_VICTIMS	Total victims of the incident.	Quantitative, discrete	Yes
LOCATION_NAME	Location of where incident happened.	Qualitative, ordinal	No
BIAS_DESC	Description of bias	Qualitative, nominal	No
VICTIM_TYPES	Type of victims: individual, organization, etc.	Qualitative, ordinal	No
MULTIPLE_OFFENSE	S or M assigned if there were single or multiple offenses.	Quanlitative, ordinal	No
MULTIPLE_BIAS	S or M assigned if there were single or multiple biases.	Qualitative, ordinal	No

7. Questions to Explore

1. What type of hate crimes are more prevalent throughout 1991-2018?
2. Which ethnic group was most affected, and which ethnic group was least affected?
3. Which states had the most hate crimes reported throughout 1991-2018?