CS 6630 Visualization for Data Science

**Process Book for**

Analysis of Hate Crimes in United States

horizontal line

# 

# Project Resource Link

[**https://github.com/varunadd2712/VisualizationProject**](https://github.com/varunadd2712/VisualizationProject)

# Team Members

Maitrey Mehta Sai Varun Shalin Parikh

U1201309 U1210409 U1206312

[maitrey@cs.utah.edu](mailto:maitrey@cs.utah.edu) [sai.varun@cs.utah.edu](mailto:sai.varun@cs.utah.edu) [shalin.parikh@cs.utah.edu](mailto:shalin.parikh@cs.utah.edu)

# Introduction

The current rise in the amount of hate crimes in the United States motivates us as students to use the power of data to present a clean and a correct visualization. We hope that this visualization would enable users to analyze the data and infer vital insights. We believe that there aren’t enough sources currently present that provide an exhaustive analysis into the type of hate crimes that span over a varied number of attributes. We plan to present data over a long period of time and aim to observe some patterns. We also want to inspect the crimes occurring both geographically and demographically over the entire country. This would enable the user to have a holistic view of the data and also focus on the specifics.

## Project Objectives

1. To see what kind of **hate crimes** are prevalent over the **United States**.
2. To see which places are unsafe, both **cities and state wise.**
3. To see **trends in the crimes** over the years.
4. Provide a **metric for comparison** between various states based on the crimes committed.
5. To analyze a **particular crime** and how it is **dependent on other factors**.

### Data and Data Processing

The data is a **12** table dataset, obtained from <https://ucr.fbi.gov/hate-crime/2016/topic-pages/incidentsandoffenses>.

We would be using the years from **2008 - 2016** for our trend analysis.

The only missing data in the dataset is certain blank rows for crime values in a city. They can be safely set to 0 and the aggregate values of the other tables aren’t affected.

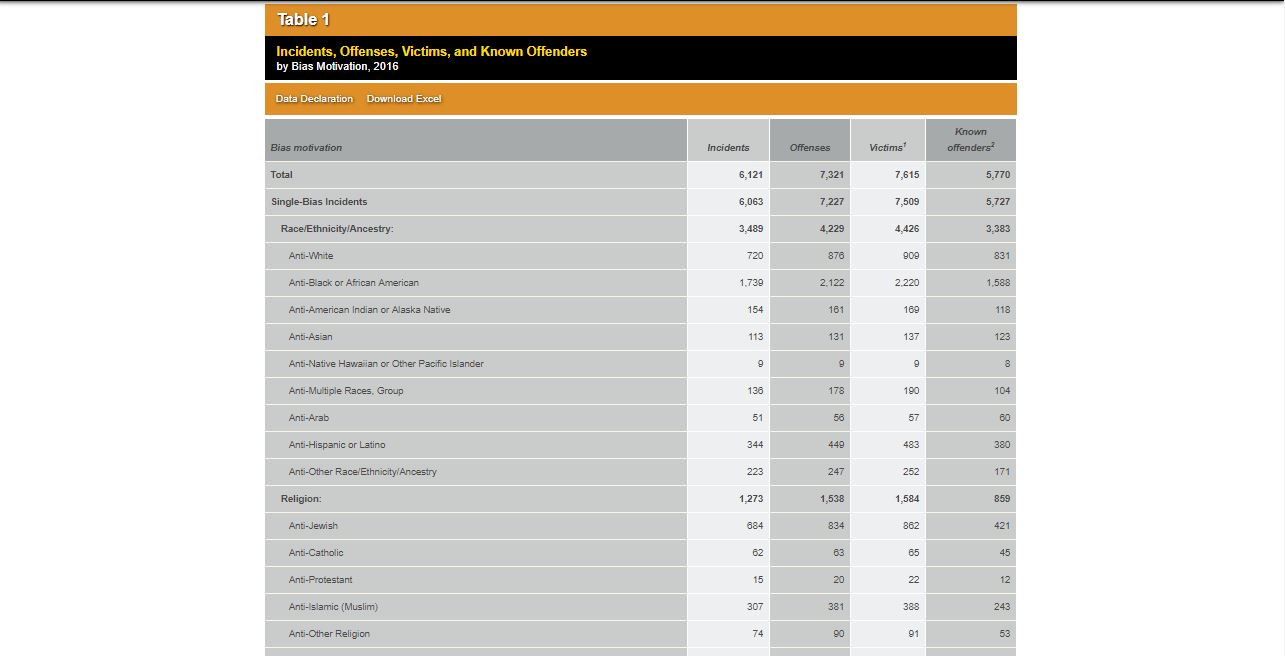
Our data can be categorized into

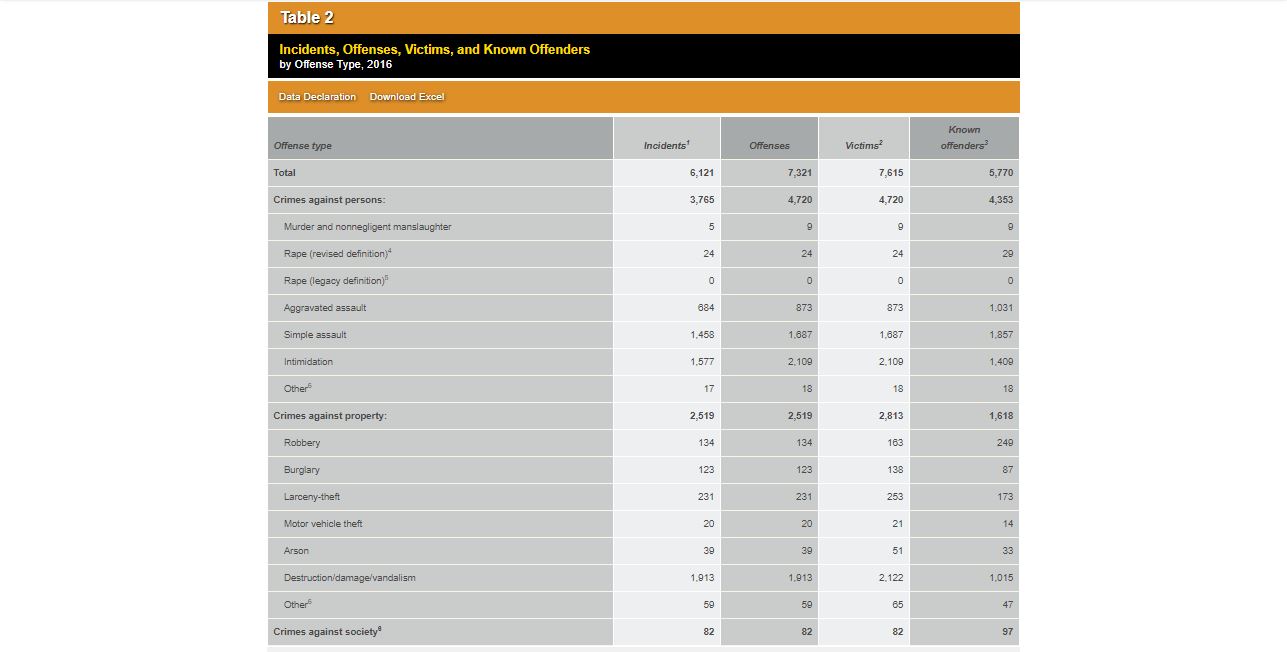
1. Statistics of the **Incidents and offenses**.
2. Statistics on the **victims.**
3. Statistics on the **offenders**.
4. Split of the above statistics based **on jurisdiction (i.e. state)**

Since we have this data over several years, we can use it to view trends or patterns in the crimes. On the whole, our visualization will be split into a visualization for the current year and a trend visualization over the years.

The different tables in the dataset give different aggregations. We’ll be using the tables individually for our charts.

Examples of tables in our dataset are as follows





# Visualization Design

We have decided to implement our design using the [Five Design Sheet Methodology](http://fds.design/).

The following images represents a sequential and a iterative process of our thinking.

We have sketched our interpretation of the data representation that according to us would be the most apt in visualizing the data.

We have thoroughly analyzed our data set to show each entity true to its data value that it encodes.

We have followed the standard design guidelines to the best of our capabilities to minimize any form of discrepancies between the data and its realization.

Each chart and or diagram has been carefully sketched, keeping in mind the larger demographic of users and using the most suitable Marks and Channels.

This is our selected design.

# 

# Must Have Features

1. An **year picker** to pick dataset from **different years.**
2. A chart to view the different proportions of **Crime-Motivations/Offense Types/Victim Types.** (Preferably a **Treemap**)
3. A **geographical view of the crimes**. (Geographical chart of US color coded by total crime values).
4. A **mechanism to compare different states** based on crime values (Preferably **Donut Chart**)
5. A chart to visualize the trends across states and crime types (Preferably **Line charts**)

# Optional Features

1. **Sorting of the bar charts** of states on the basis of **Total Crimes/ Motivation Types** etc.
2. **Semantic zooming** on a state’s geographical chart when it is selected on the map. then we color the cities instead of the states.
3. Additional **text field** showing the news articles about the crime.

**Progress Till First Milestone**

We completed a total of 5 charts for our milestone 1. Having implemented

1. An year picker.
2. A chart for total victim count.
3. Treemap for crime types.
4. Choropleth map of United States.
5. Bar charts for each state in United States.

We have finalized the data format that we required for each chart and can easily extend the above to all years.