

ANA 515 Assignment 2

Grace Zeng

6/16/2024

Description of the data:

The dataset* selected for this analysis comes from the FBI and the Southern Poverty Law Center (SPLC) and the research question is: What are the trends in hate crime rates by state, and what state-level factors are associated with these variations?

Data collection: The federal government does not systematically track hate crimes, as reporting to the FBI is voluntary. The Southern Poverty Law Center relies on media accounts and self-reports to evaluate the situation. Both sources provide publicly available data.

*The dataset is found on GitHub (<https://github.com/fivethirtyeight/data/tree/master/hate-crimes>) and it is in csv. format. More background information can be found in this article (<https://fivethirtyeight.com/features/higher-rates-of-hate-crimes-are-tied-to-income-inequality/>).

Get a glimpse of the data

```
## Rows: 51
## Columns: 12
## $ state                <chr> "Alabama", "Alaska", "Arizona...
## $ median_household_income <dbl> 42278, 67629, 49254, 44922, 6...
## $ share_unemployed_seasonal <dbl> 0.060, 0.064, 0.063, 0.052, 0...
## $ share_population_in_metro_areas <dbl> 0.64, 0.63, 0.90, 0.69, 0.97,...
## $ share_population_with_high_school_degree <dbl> 0.821, 0.914, 0.842, 0.824, 0...
## $ share_non_citizen      <dbl> 0.02, 0.04, 0.10, 0.04, 0.13,...
## $ share_white_poverty    <dbl> 0.12, 0.06, 0.09, 0.12, 0.09,...
## $ gini_index             <dbl> 0.472, 0.422, 0.455, 0.458, 0...
## $ share_non_white        <dbl> 0.35, 0.42, 0.49, 0.26, 0.61,...
## $ share_voters_voted_trump <dbl> 0.63, 0.53, 0.50, 0.60, 0.33,...
## $ hate_crimes_per_100k_splc <dbl> 0.12583893, 0.14374012, 0.225...
## $ avg_hatecrimes_per_100k_fbi <dbl> 1.8064105, 1.6567001, 3.41392...
```

Characteristics of the data

This dataframe has 51 rows and 12 columns. The names of the columns and a brief description of each are in the table below:

Description of Columns in hate_crimes_arrange_by_income Dataset

Columns	Description
state	State name

Columns	Description
median_household_income	Median household income, 2016
share_unemployed_seasonal	Share of the population that is unemployed (seasonally adjusted), Sept. 2016
share_population_in_metro_areas	Share of the population that lives in metropolitan areas, 2015
share_population_with_high_school_degree	Share of adults 25 and older with a high-school degree, 2009
share_non_citizen	Share of the population that are not U.S. citizens, 2015
share_white_poverty	Share of white residents who are living in poverty, 2015
gini_index	Gini Index, 2015
share_non_white	Share of the population that is not white, 2015
share_voters_voted_trump	Share of 2016 U.S. presidential voters who voted for Donald Trump
hate_crimes_per_100k_splc	Hate crimes per 100,000 population, Southern Poverty Law Center, Nov. 9-18, 2016
avg_hatecrimes_per_100k_fbi	Average annual hate crimes per 100,000 population, FBI, 2010-2015

Summary of data subset

Subset of the data:

```
## # A tibble: 6 × 4
##   state      median_household_inc...1 share_unemployed_sea...2 avg_hatecrimes_per_1...3
##   <chr>          <dbl>          <dbl>          <dbl>
## 1 Mississi...      35521          0.061          0.623
## 2 West Vir...      39552          0.073          2.04
## 3 Alabama          42278          0.06           1.81
## 4 Louisiana          42406          0.06           1.34
## 5 Kentucky          42786          0.05           4.21
## 6 Tennessee          43716          0.057          3.14
## # i abbreviated names: 1median_household_income, 2share_unemployed_seasonal,
## # 3avg_hatecrimes_per_100k_fbi
```

Summary of the subset:

```
##      state      median_household_income share_unemployed_seasonal
## Length:51      Min.      :35521      Min.      :0.02800
## Class :character 1st Qu.:48657      1st Qu.:0.04200
## Mode  :character Median :54916      Median :0.05100
##              Mean  :55224      Mean   :0.04957
##              3rd Qu.:60719      3rd Qu.:0.05750
##              Max.  :76165      Max.   :0.07300
##
## avg_hatecrimes_per_100k_fbi
## Min.      : 0.2669
## 1st Qu.: 1.2931
## Median : 1.9871
## Mean      : 2.3676
## 3rd Qu.: 3.1843
## Max.      :10.9535
## NA's      :1
```