

CSIT 553 Project 2: Map & Aggregation Visualization

Gurpreet Kaur





About the DataSet: Hate Crimes between 1991 and 2020

This dataset consists of 219,577 hate crimes committed between 1991 and 2020. It is collected by The Hate Crime Statistics Program of the FBI Uniform Crime Reporting (UCR) Program which collects the data regarding criminal offenses that were motivated by offender's bias against the victim's race/ethnicity/ancestry, gender, gender identity, religion, disability, or sexual orientation, and were committed against the person, property, or society.

This data contains: incident id, year, ORI, Public agency name, public agency unit, agency type name, state, division, region, population group code, population group description, incident date, adult victim count, juvenile victim count, total offender count, adult offender count, juvenile offender count, offender race, offender ethnicity, victim count, offence name, total individual victims, location, Bias description, victim types, multiple offense, multiple bias.

Objective:

- Find which areas of the United States the crimes are occurring the most.
- Find trends between the years
- Which group is the most targeted based on ethnicity, race, religion, sexual orientation, etc,

INCIDENT_ID	DATA_YEAR	ORI	PUB_AGENCY_NAME	PUB_AGENCY_UNIT	AGENCY_TYPE_NAME	STATE_ABBR	STATE_NAME	DIVISION_NAME	REGION_NAME	POPULATION_GROUP_CODE	POPULATION_GROUP_DESC	INCIDENT_DATE	ADULT_VICTIM_C
3015	1991	AR0040200	Rogers		City	AR	Arkansas	West South Central	South	5	Cities from 10,000 thru 24,999	31-AUG-91	
3016	1991	AR0290100	Hope		City	AR	Arkansas	West South Central	South	6	Cities from 2,500 thru 9,999	19-SEP-91	
43	1991	AR0350100	Pine Bluff		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	04-JUL-91	
44	1991	AR0350100	Pine Bluff		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	24-DEC-91	
3017	1991	AR0350100	Pine Bluff		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	23-DEC-91	
3018	1991	AR0350100	Pine Bluff		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	27-JUL-91	
3019	1991	AR0600200	Little Rock		City	AR	Arkansas	West South Central	South	2	Cities from 100,000 thru 249,999	14-NOV-91	
45	1991	AR0600300	North Little Rock		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	10-JUL-91	
46	1991	AR0600300	North Little Rock		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	06-OCT-91	
3020	1991	AR0600300	North Little Rock		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	04-OCT-91	
3021	1991	AR0600300	North Little Rock		City	AR	Arkansas	West South Central	South	3	Cities from 50,000 thru 99,999	06-OCT-91	
47	1991	AR0670000	Sevier		County	AR	Arkansas	West South Central	South	8D	Non-MSA counties under 10,000	14-OCT-91	
2	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	15-JAN-91	
3	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	22-JAN-91	
4	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	23-JAN-91	
5	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	04-FEB-91	
6	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	14-FEB-91	
7	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	20-FEB-91	
8	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	21-FEB-91	
9	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	13-MAR-91	
10	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	01-APR-91	
11	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	29-APR-91	
12	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	06-MAY-91	
13	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	09-JUN-91	
14	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	30-JUN-91	
15	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	11-JUL-91	
16	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	30-JUL-91	
17	1991	AZ0072300	Phoenix		City	AZ	Arizona	Mountain	West	1B	Cities from 500,000 thru 999,999	31-JUL-91	

Data Loading

```
8]: # Data Loading
df = pd.read_csv("Hate_crime.csv")

/Users/gkaur/opt/anaconda3/lib/python3.8/site-packages/IPython/core/interactiveshell.py:3165: DtypeWarning: Columns
(19) have mixed types.Specify dtype option on import or set low_memory=False.
has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

```
9]: df
```

```
9]:
```

	INCIDENT_ID	DATA_YEAR	ORI	PUB_AGENCY_NAME	PUB_AGENCY_UNIT	AGENCY_TYPE_NAME	STATE_ABBR	STATE_NAME	DIVISION_NAM
0	3015	1991	AR0040200	Rogers	NaN	City	AR	Arkansas	West Sout Centr.
1	3016	1991	AR0290100	Hope	NaN	City	AR	Arkansas	West Sout Centr.
2	43	1991	AR0350100	Pine Bluff	NaN	City	AR	Arkansas	West Sout Centr.
3	44	1991	AR0350100	Pine Bluff	NaN	City	AR	Arkansas	West Sout Centr.
4	3017	1991	AR0350100	Pine Bluff	NaN	City	AR	Arkansas	West Sout Centr.
...
219572	1318026	2020	WY0170100	Sheridan	NaN	City	WY	Wyoming	Mountai

Data Cleaning

- Check to see if there are any null values in each column.
- Replaced the NaN values in columns 'OFFENDER_RACE' and 'OFFENDER_ETHNICITY' with "Unknown"
- Removed column 'TOTAL_INDIVIDUAL_VICTIMS'

```
[72]: #Check to see if there are any Null values
df1.isnull().values.any()

In [72]: True

[73]: # See where are the null values occur
print(df1.isnull().sum())

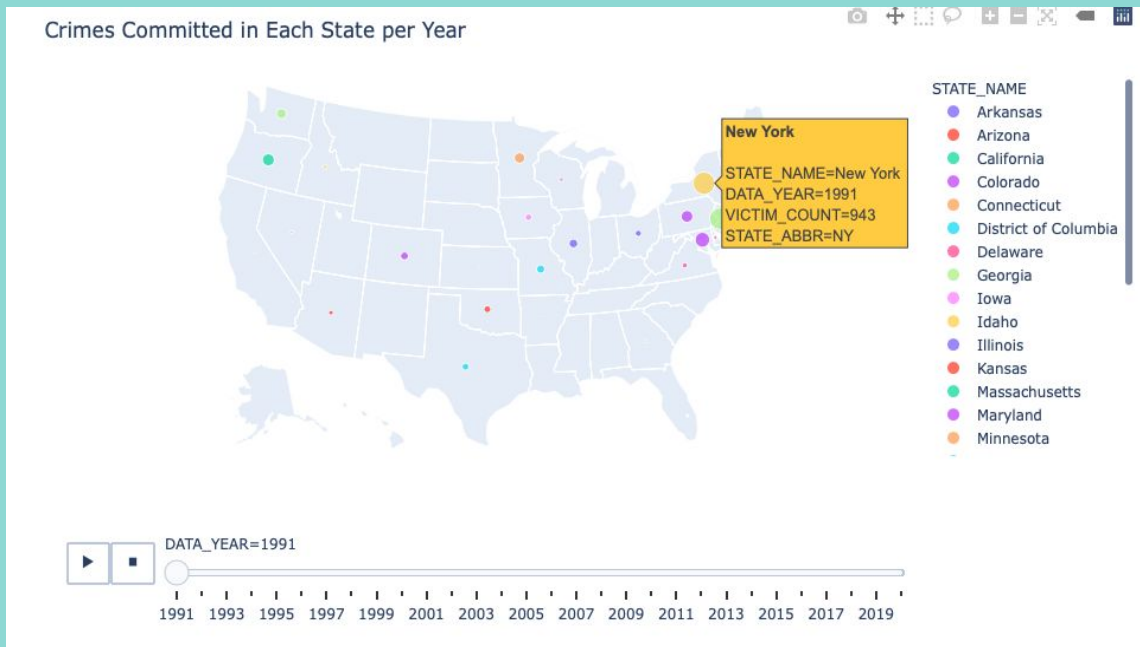
DATA_YEAR                0
PUB_AGENCY_NAME          0
AGENCY_TYPE_NAME         0
STATE_ABBR               0
STATE_NAME               0
DIVISION_NAME            0
REGION_NAME              0
POPULATION_GROUP_CODE    0
POPULATION_GROUP_DESC    0
INCIDENT_DATE            0
TOTAL_OFFENDER_COUNT      0
OFFENDER_RACE             20
OFFENDER_ETHNICITY        184704
VICTIM_COUNT              0
OFFENSE_NAME              0
TOTAL_INDIVIDUAL_VICTIMS  3479
LOCATION_NAME              0
BIAS_DESC                 0
VICTIM_TYPES              0
dtype: int64

[74]: # Replace the Null values in these columns to 'Unknown'
df1[['OFFENDER_ETHNICITY', 'OFFENDER_RACE']] = df1[['OFFENDER_ETHNICITY', 'OFFENDER_RACE']].fillna('Unknown')
```

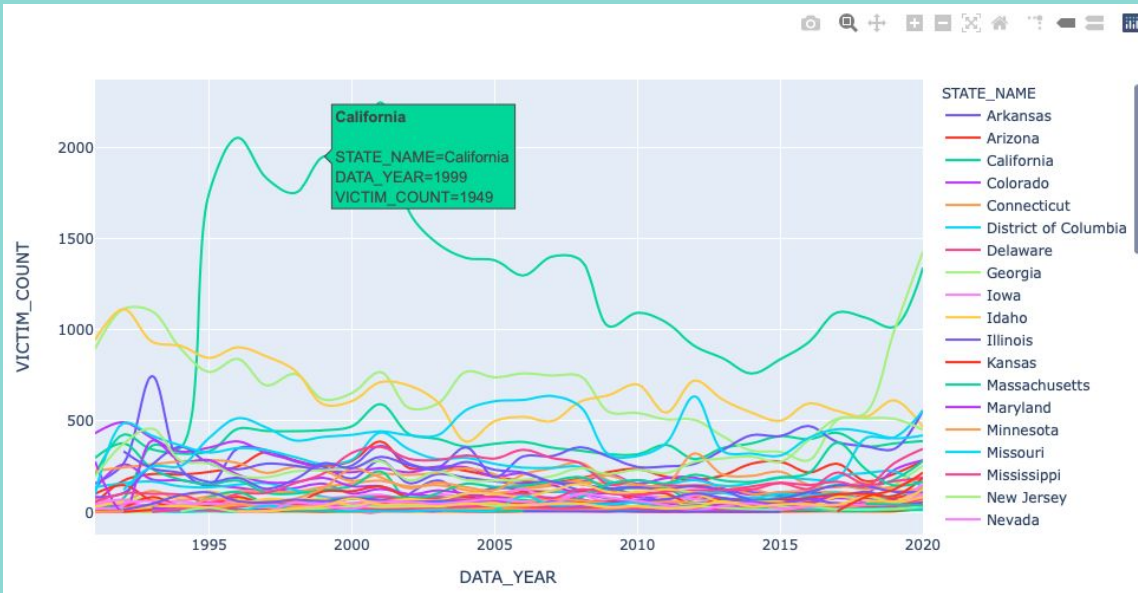
OFFENDER_RACE	OFFENDER_ETHNICITY	VICTIM_COUNT	
White	Unknown	1	
Black or African American	Unknown	1	
Black or African American	Unknown	1	
Black or African American	Unknown	2	Assault;C
Black or African American	Unknown	1	
...	
White	Not Hispanic or Latino	2	
Unknown	Unknown	1	Des
Black or African American	Not Hispanic or Latino	1	
White	Not Hispanic or Latino	1	
White	Not Hispanic or Latino	1	

Hate Crimes Committed in Each State per Year

- Idiom: Bubble Map
- Marks: Points
- Channels:
 - Color: hue
 - Vertical and horizontal position
 - Size(area)
- Contains a slider
- Observations:
 - Larger circles = more crimes
 - 2020 has more marks with greater area than the previous years.



Hate Crimes Committed in Each State per Year



- Graph: Line Graph
- Marks: Points and Lines
- Channels:
 - Color: hue
 - Vertical and horizontal position
 - Aligned lengths to express quant value
- Task: find trends
- Observations
 - Overall California has the most crimes throughout the years
 - NJ (lime green) has increased exponentially recently

Hate Crimes per Region 1991-2020

Total Hate Crimes per Region



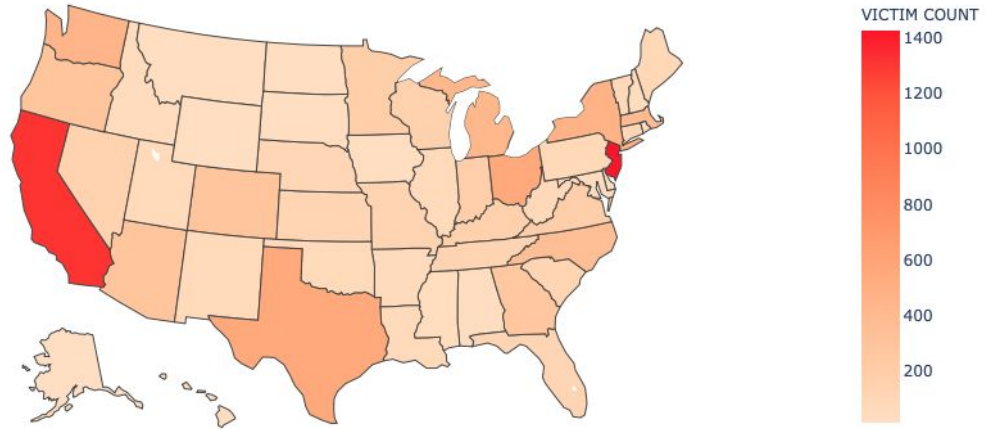
```
hfig = px.histogram(df1, x="REGION_NAME", color_discrete_sequence=['skyblue'],  
                    title= 'Total Hate Crimes per Region', histnorm = 'percent')  
hfig.show()
```

- Idiom: Histogram
- Marks: Lines
- Channels:
 - Color: blue
 - Vertical position
- Data: table
- Task: find distribution
- Normalization: percent
- Observations:
 - West Region has the most percent of hate crimes, Northeast with a close second.

Number of Victims in Each State in 2020

- Idiom: Choropleth Map
- Marks: areas
- Channels:
 - Color hue
 - Spatial region
- Observations:
 - California and New Jersey has the most amount of victims and seems to be outliers in this graph.
 - Most states have a similar shade

Number of Victims in Each State In 2020



Population Distribution 2020



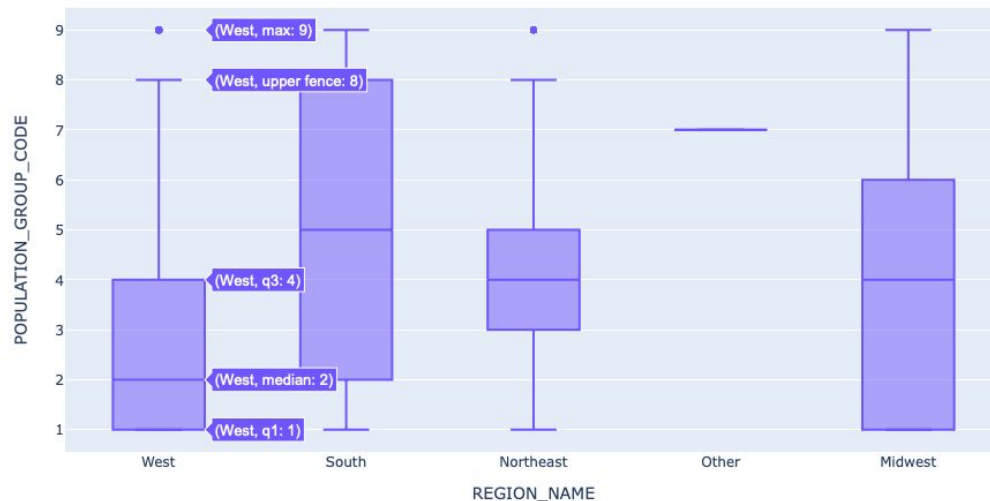
Cities are categorized into their Population Group Code based on their population.

- 1: 250,000 and more
- 2: 100,000 to 249,999
- 3: 50,000 to 99,999
- 4: 25,000 to 49,999
- 5: 10,000 to 24,999
- 6: Less than 10,000
- 7: -
- 8: N/A (Non metropolitan County)
- 9: N/A (Metropolitan County)

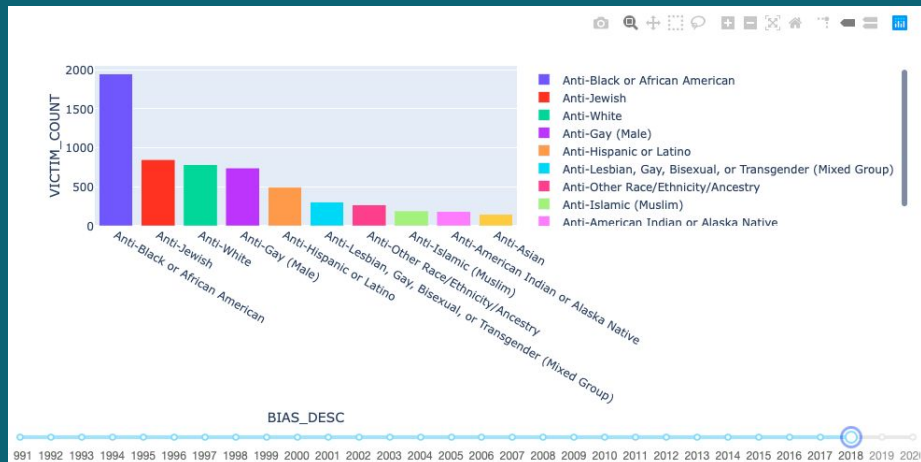
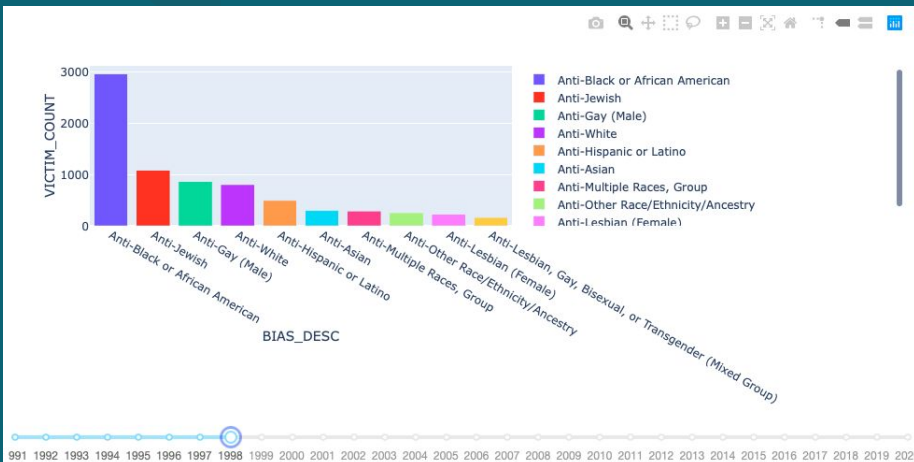
Metropolitan: 50,000 inhabitants, suburbs, Metropolitan counties
Non-Metropolitan: Cities outside of metropolitan areas, Non metropolitan Counties

Population Distribution 2020

- Idiom: Boxplot
- Task: find distribution
- Data: table
- Derived data
 - Median: central line
 - Lower and upper quartiles: boxes
 - Lower upper fences: whiskers
 - Outliers
- Observations:
 - Other: territories
 - West Region has cities with higher population where crimes are committed the most



Top 10 Groups Targeted Each Year



- Idiom: Bar Graph
- Marks: Lines
- Channels:
 - Color hue
 - Vertical position
- Data: table

- Contains slider
- The graph is sorted from most to least.
- Observations:
 - 1998 vs 2018: similar graphs but with some differences in the group targeted
 - Slight decrease in the amount of hate crimes in the 20 yr span.

Conclusion

- Observations:
 - 2020 has the most hate crimes committed
 - Most of the crimes were committed in California in recent years
 - CA has the largest population in the USA
 - CA has the most cities with large population
 - Most of the Hate Crimes were Anti-African American
 - This group has been most targeted group every year
- About the Data Set
 - Vast amount of data
 - More can be explored
 - Not much numerical data to work with



References

- Dataset Source:
https://s3-us-gov-west-1.amazonaws.com/cg-d4b776d0-d898-4153-90c8-8336f86bdfec/hate_crime.zip
- FBI. (2016, August 11). *Area definitions*. FBI. Retrieved April 30, 2022, from <https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/resource-pages/area-definitions>
- FBI. (2018, September 14). *Hate crime*. FBI. Retrieved April 30, 2022, from <https://www.fbi.gov/services/cjis/ucr/hate-crime>