



Data Analytics Bootcamp

Project 1:

Hate Crimes in the United States

Group 3

Chris Johnson, Andy Grimm, Daniela Castellon, Fekadu Kassa, Tymber Hamilton

Introduction

```
# Clean up the data
# Drop columns of data not needed
clean_crime = crime_data_df.drop(columns=["ORI", "PUB_AGENCY_UNIT", "JUVENILE_VICTIM_COUNT", "ADULT_VICTIM_COUNT", "ADULT_OFFENDER_COUNT",

# Dropping "Guam", "Federal", "District of Columbia" from clean_crime to keep only the named continental states
state_drop = ["Guam", "Federal", "District of Columbia"]
clean_crime = clean_crime[~clean_crime["STATE_NAME"].isin(state_drop)]

#Split Offense names into main and sub categories to only look at the main offense
clean_crime[["MAIN_OFFENSE", "SUB_OFFENSE"]] = clean_crime["OFFENSE_NAME"].str.split('/', n=1, expand= True)
clean_crime["MAIN_OFFENSE"] = clean_crime["MAIN_OFFENSE"].str.split(';', n=1, expand= True)

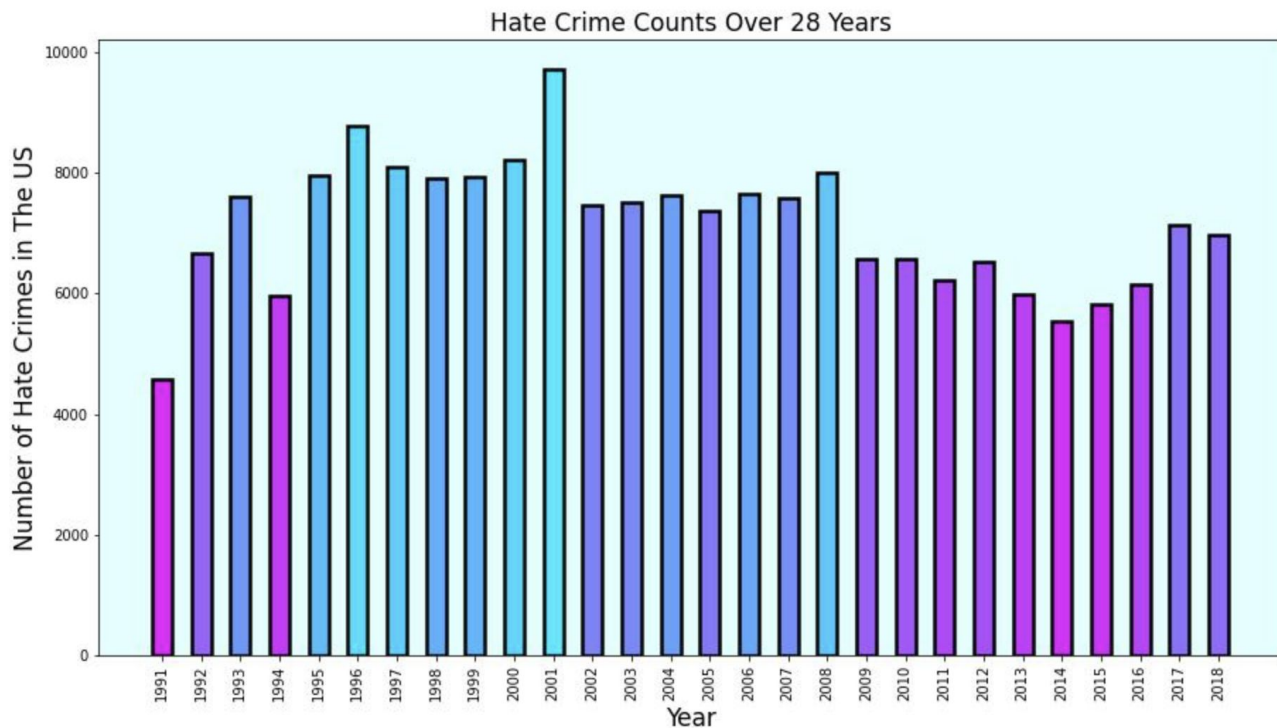
# Split the incident date into year, month, and day
dates = clean_crime['INCIDENT_DATE'].str.split("-", n = 2, expand = True)

# Add new columns to the cleaned dataframe and dropped the original date column
clean_crime["DAY"] = dates[0]
clean_crime["MONTH"] = dates[1]
clean_crime["LAT"] = ""
clean_crime["LNG"] = ""
clean_crime.drop(columns = ["INCIDENT_DATE"], inplace = True)

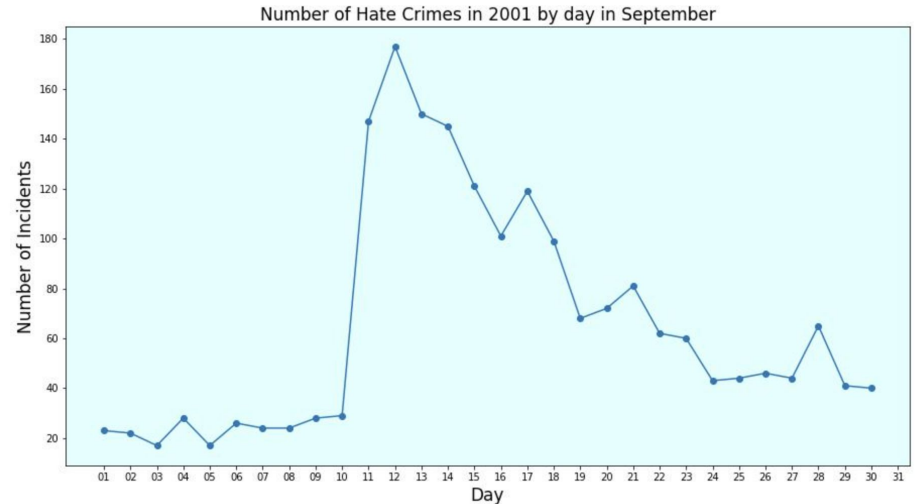
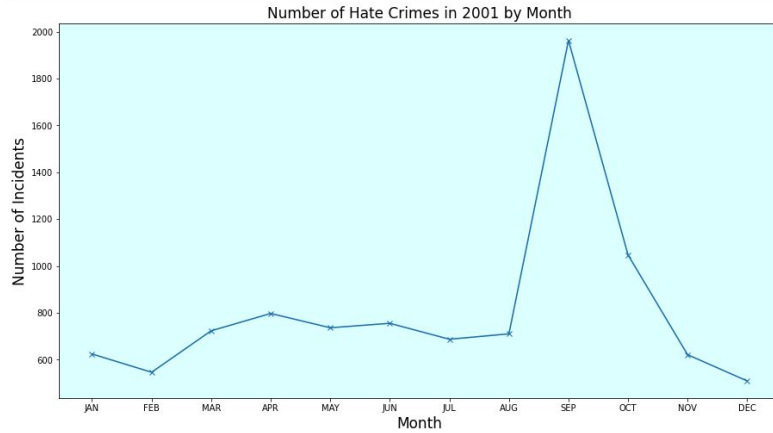
# Rename the year column and replace offender race to only the main category and change the order of the columns to have all 3 date column
clean_crime.rename(columns={'DATA_YEAR': 'YEAR'}, inplace=True)
clean_crime['OFFENDER_RACE'] = clean_crime['OFFENDER_RACE'].replace(['American Indian or Alaska Native', 'Asian', 'Native Hawaiian or Other Pacific Islander', 'Black or African American', 'White'], ['American Indian or Alaska Native', 'Asian', 'Native Hawaiian or Other Pacific Islander', 'Black or African American', 'White'])
clean_crime = clean_crime[['STATE_NAME', 'STATE_ABBR', 'PUB_AGENCY_NAME', 'LAT', 'LNG', 'YEAR', 'MONTH', 'DAY', 'MAIN_OFFENSE', 'BIAS_DESC', 'OFFENDER_RACE', 'OFFENSE_NAME', 'SUB_OFFENSE']]

#display new dataframe
# clean_crime
```

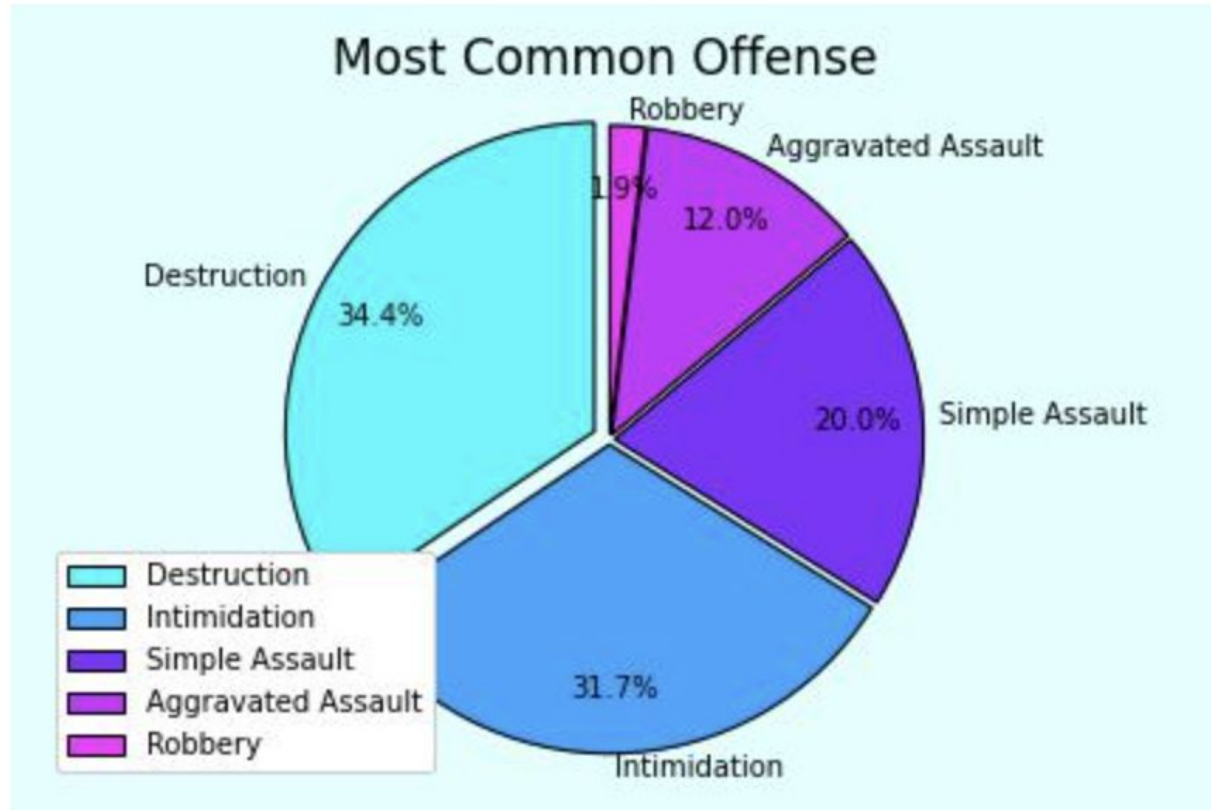
What year had the highest number of hate crimes?



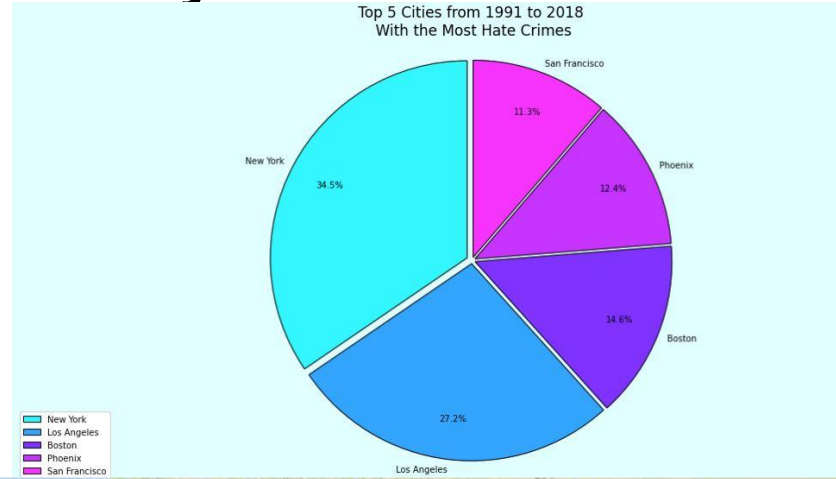
Did hate crimes spiked after 09/11/2001?



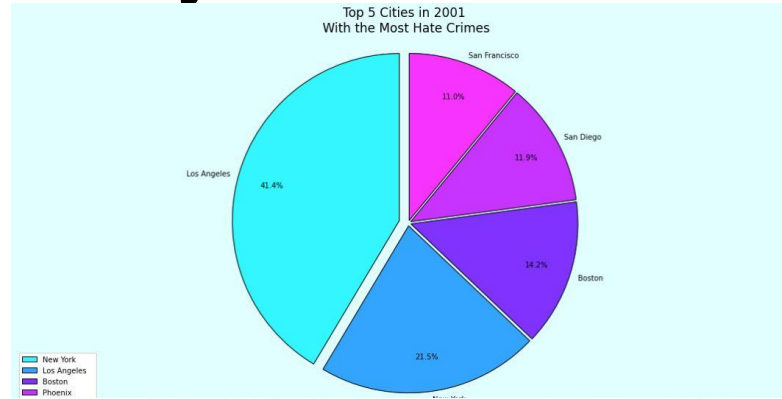
Which hate crime was most prevalent?



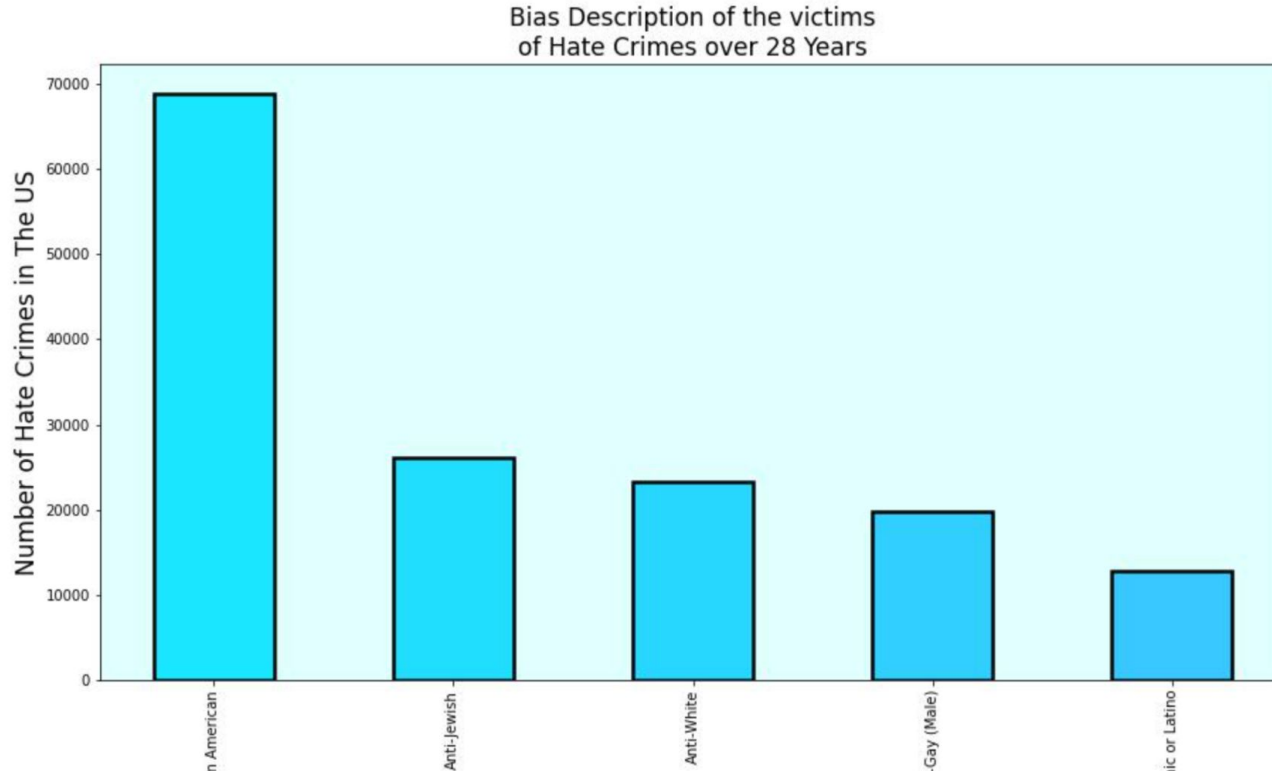
Which city has most hate crime(s)?



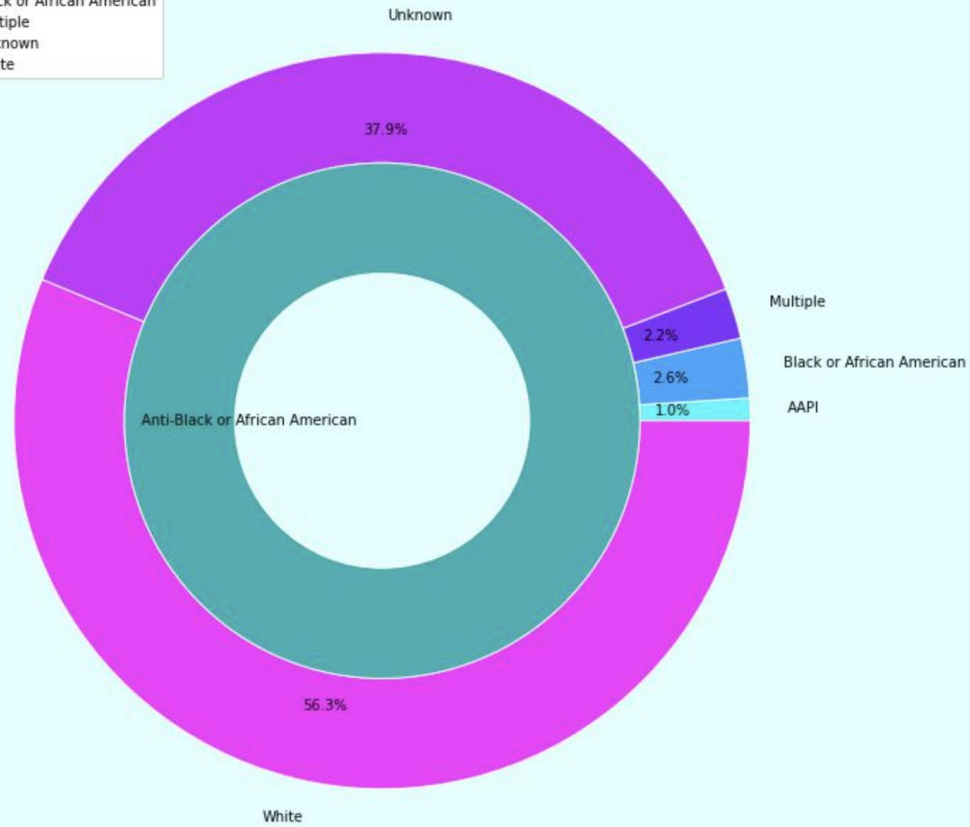
Which city has most hate crime(s)?



What targeted group is most susceptible?



Offender Totals for the most Suseptible
Bias Description (Anti-Black or African American)



Conclusion

- The most hate crimes occurred during the year of 2001, with a total of 9,730
- California was the state with the most reported hate crimes of 33,891
- The frequent occurring hate crime is Destruction/Damage/Vandalism of Property with a total of 64,534
- The most susceptible targeted group is Anti-Black with a total of 68,882

We looked at several crime datasets and we landed on the FBI's hate crime data. Hate crimes was an interesting specific type of crime to look at and the dataset was plentiful.



Questions and Feedback

Questions:

Which hate crime most prevalent?

Destruction/Damage/Vandalism of Property

What bias targeted group is most susceptible?

Anti-Black or African American

Which city/state has most hate crime(s)?

New York, New York and Los Angeles, California

What is correlation of the most common crimes throughout the years?

What is correlation of the most common crimes throughout the years?