

# Right Wing Extremist Violence and Police Use of Force

*Maya Van Nuys*

*12/9/2019*

## Background and Goal

My research interests currently center on ideologically-motivated violence. Recently, I have been focusing on white supremacist and other right-wing extremist violence in the United States. I decided to use geospatial mapping to compare incidents of right wing terrorism to data on police use of deadly force. The motivation for this project comes from a history of the Klu Klux Klan (KKK) and other white supremacist groups infiltrating police departments in the United States. As such, one might expect cities where police departments have a history of white supremacist infiltration to (1) have higher rates of deadly use of force, as white supremacist influences would be permissive of violence, and (2) have higher rates of far right extremist violence, as the police departments would create permissive atmospheres for these types of crimes. My goal is to visualize whether areas with higher rates of police use of force also have higher rates of far right extremist violence.

## Collecting data

For my data, I utilized the Terrorism and Extremist Violence in the United States (TEVUS) dataset, compiled by the University of Maryland's National Consortium for the Study of Terrorism and Responses to Terrorism (START). For police use of force data, I used VICE News' dataset on police use of force in the United States' 50 largest police departments between 2010 and 2016. This was the police violence data set that covered the longest period of time. I also restricted the TEVUS dataset to the same time period for consistency.

## Cleaning and Pre-processing data

Manipulating the data was the most time intensive process of the project. The main problem was that the TEVUS dataset had only latitude and longitude coordinates for each event in the dataset, and the VICE dataset had only city names and states. To map the dataset, I needed to translate the city names in the VICE dataset to latitude and longitude. To do this, I created a Google API in order to use the Google Maps service to translate the cities into coordinates. Once I accessed the API, I was able to create a vector of the unique coordinates for each city in the VICE dataset. To add the new coordinates for each of the 4,400 observations in VICE, I created two functions that translated the city names into longitude then latitude individually, then added these values into the original dataframe as a new column for each observation.

```
require(tidyverse)
require(ggmap)

#Google API
register_google(key = "redacted")
getOption("ggmap")

#Longitude:
lon_function <- function(x){
  long_lat <- as.character(x)
  geocode(long_lat) %>%
    select(lon) %>%
    mutate(lon = sapply(lon, toString)) %>%
    unlist() %>%
```

```

    as.numeric()
}

#Latitude
lat_function <- function(x){
  long_lat <- as.character(x)
  geocode(long_lat) %>%
    select(lat) %>%
    mutate(lat = sapply(lat, toString))%>%
    unlist() %>%
    as.numeric()
}

#Adding latitude into the original dataset
vice_latitude <- vice_dat %>%
  mutate(latitude = lat_function(City))

#Adding longitude into the new dataset
adjusted_vice <- vice_latitude %>%
  mutate(longitude = lon_function(City))

```

The next challenge was mapping both datasets onto an interactive map. After trying several different methods of mapping, I settled on the “mapview” function from the mapview package. While I was able to map the locations of police use of force, I found the map to be uninformative, since it only mapped the fifty cities that comprised the fifty largest police departments in the US. As such, I wanted to map the frequencies of police use of force to visualize where it was most prevalent. To do this, I added a dummy variable column to the data set consisting of “1,” then summed the dummy variable for each city. I did the same for the TEVUS data, then mapped both data sets on the same interactive map. I had to add another dummy variable column to the TEVUS data with a character to prevent the map from creating a legend with coloring by frequency. I found that points with size-based frequency was sufficient and that also changing colors of the locations by frequency made the map more confusing.

```

require(ggplot2)
require(dplyr)
require(sf)
require(mapview)
require(leaflet)
require(viridis)

#Getting frequencies for VICE
test_freq_vice <- adjusted_vice %>%
  mutate(Dummy = 1) %>%
  group_by(City, longitude, latitude) %>%
  summarize(sum(Dummy))

#Map function for VICE
Vice <- st_as_sf(test_freq_vice, coords = c("longitude", "latitude"), crs = 4326)

#Frequencies for TEVUS
test_freq_tevus <- tevus_dat %>%
  mutate(Dummy = 1) %>%
  group_by(Longitude, Latitude) %>%

```

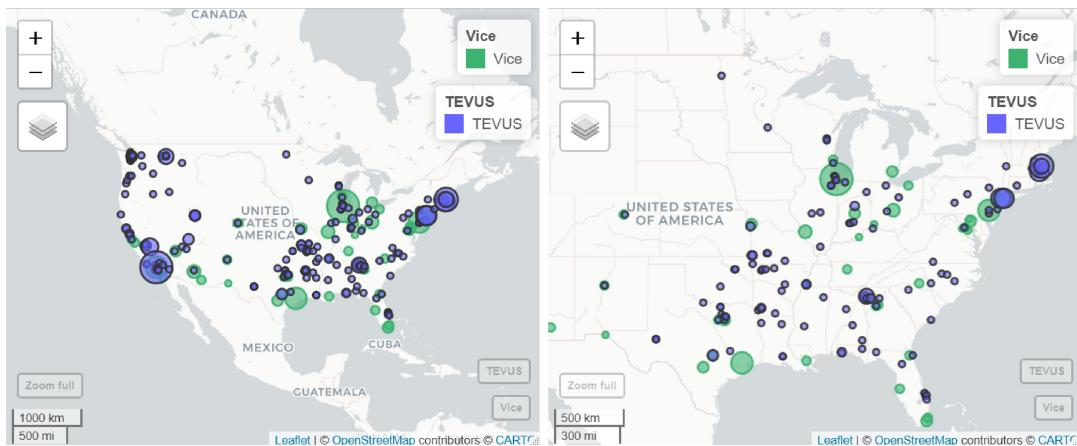
```

summarize(sum(Dummy))

#Map function for TEVUS
TEVUS <- st_as_sf(test_freq_tevus, coords = c("Longitude", "Latitude"), crs = 4326) %>%
  mutate(Dummy = "a")

#Mapping both datasets
mapview(Vice, color = "mediumseagreen", col.regions = "mediumseagreen", cex = "sum(Dummy)") +
  mapview(TEVUS, color = viridis, col.regions = viridis, cex = "sum(Dummy)", Legend = FALSE)

```



Next, I wanted to merge the datasets to see which locations had both police violence and right wing extremist violence. To do this, I needed to merge based on a unit of location more general than latitude and longitude. However, as I noted earlier, the TEVUS dataset only has coordinates. As such, I used the “revgeo” function to convert the coordinate data into city, state, and zipcode for each observation. This process introduced a host of new problems because I decided to also gather zip codes for the VICE dataset. However, the Google Maps API crashed everytime I tried to gather zip codes for the entire VICE dataset. To get around this issue, I created a dataframe of the unique locations in the VICE dataset, passed this set into the revgeo function, and then added the new location information back into the adjusted VICE dataframe.

```

require(revgeo)

#Unique latitude and longitude data for TEVUS
tevus_lat_long <- tevus_dat[, 2:3]

#Unique latitude and longitude data for VICE
vice_lat_long <- adjusted_vice[, 18:19]

#Converting TEVUS coordinates to city, state, and zipcode
county_tevus <- revgeo(tevus_lat_long$Longitude, tevus_lat_long$Latitude,
                         provider= 'google', API =
                           "redacted", output =
                           'frame')

#Finding unique latitude and longitude values for VICE
vice_lat_long_unique <- vice_lat_long %>%
  group_by(latitude, longitude) %>%
  unique()

county_vice <- revgeo(vice_lat_long_unique$longitude,

```

```

vice_lat_long_unique$latitude, provider= 'google', API =
  "redacted", output =
  'frame')

#Creating a dataframe to be added back to full unique dataset
total <- vice_lat_long_unique %>%
  mutate("ID" = 1:47)

names(county_vice)[1] <- "ID"

#Creating VICE data set with county and zip for unique coordinates
merged_vice <- county_vice %>%
  left_join(total, by = "ID")

names(merged_vice)[9] <- "latitude"
names(merged_vice)[10] <- "longitude"

#Adding to full adjusted vice data
full_vice_dat <- merged_vice %>%
  left_join(adjusted_vice, by = "latitude", "longitude")

#Removing a redundant longitude row
full_vice_dat <- full_vice_dat[, -28]
names(full_vice_dat)[10] <- "longitude"

```

Now that both the TEVUS and VICE datasets had latitude, longitude, city, state, and zip code information, I began to merge datasets of only the locations at different levels of analysis. To do this, I created columns with the unique zip codes, cities, and states and the number of events at this location for each data set. I then merged these columns together to create a new dataframe of one location (either zip code, city, or state) and two columns for each data set listing the frequency of attacks at this location. The data sets only had seven zip codes in common, so I knew I would have to focus on cities or states. The data sets shared 17 cities - still too few to do analysis on. There were 24 states with both police use of force events and right wing extremist violence, giving me the most data to work with for analysis purposes.

```

#States
Vice_state_count <- full_vice_dat %>%
  count(full_vice_dat$state)

Tevus_state_count <- county_tevus %>%
  count(county_tevus$state)

names(Vice_state_count)[1] <- "state"
names(Tevus_state_count)[1] <- "state"

state_merge <- Vice_state_count %>%
  full_join(Tevus_state_count, by = "state") %>%
  na.omit()

names(state_merge)[2] <- "vice"
names(state_merge)[3] <- "tevus"

head(state_merge)

```

```
## # A tibble: 6 x 3
```

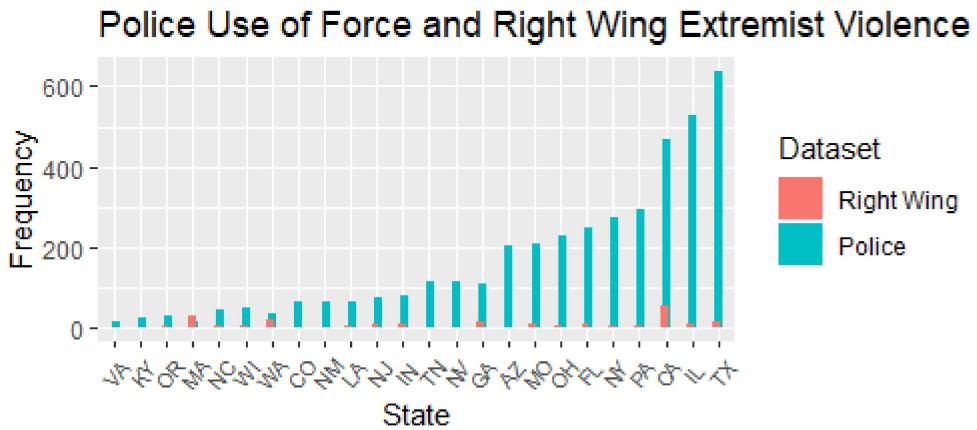
```

##   state      vice tevus
##   <chr>     <int> <int>
## 1 Arizona    205    2
## 2 California  470   55
## 3 Colorado    65    1
## 4 Florida     247   12
## 5 Georgia    110   13
## 6 Illinois   528   11

```

## Analysis and visualization

Along with the map visualization above, I decided to make a plot of the number of police use of force and right wing extremist events in each state that had both event types. Finally, I ran a simple linear regression of the VICE data set on TEVUS, producing an AJPS-style text table.



```

##
## Regression Results for States
##
## -----
## DV: VICE Police Use of Force
## -----
## TEVUS: Right Wing      5.419*
##                      (2.761)
## N                      24
## -----
## ***p < .01; **p < .05; *p < .1

```

## Future work

Ultimately, the results are inconclusive. There are not enough right wing extremist events to properly compare between the two datasets. In continuing this work, I hope to add another data set of hate crimes by location from the US Department of Justice, as these crimes are often associated with white supremacist activity. I also hope to merge several of the existing datasets on police use of force in the United States to get a more comprehensive data set over a longer period of time. With these additions, I would be able to run more robust statistical models.