



Team: IC22037

Project: Armed  
Conflict Location and  
Event

# Data Exploration/Preparation

- Key things missing:
  - Size of protests
  - Exact location
- Things added:
  - County FIPS codes
  - Associated actor splitting

# Questions to Answer

1. How has the number of events changed over the period from 01/01/2020 to 01/21/2022?
2. How have the types of events changed over the period?
3. Are certain types of events more prominent in specific areas?
4. Are there any significant specific locations/metro areas that have changed more than others?

## Events per month



## Events per month



## Events per month

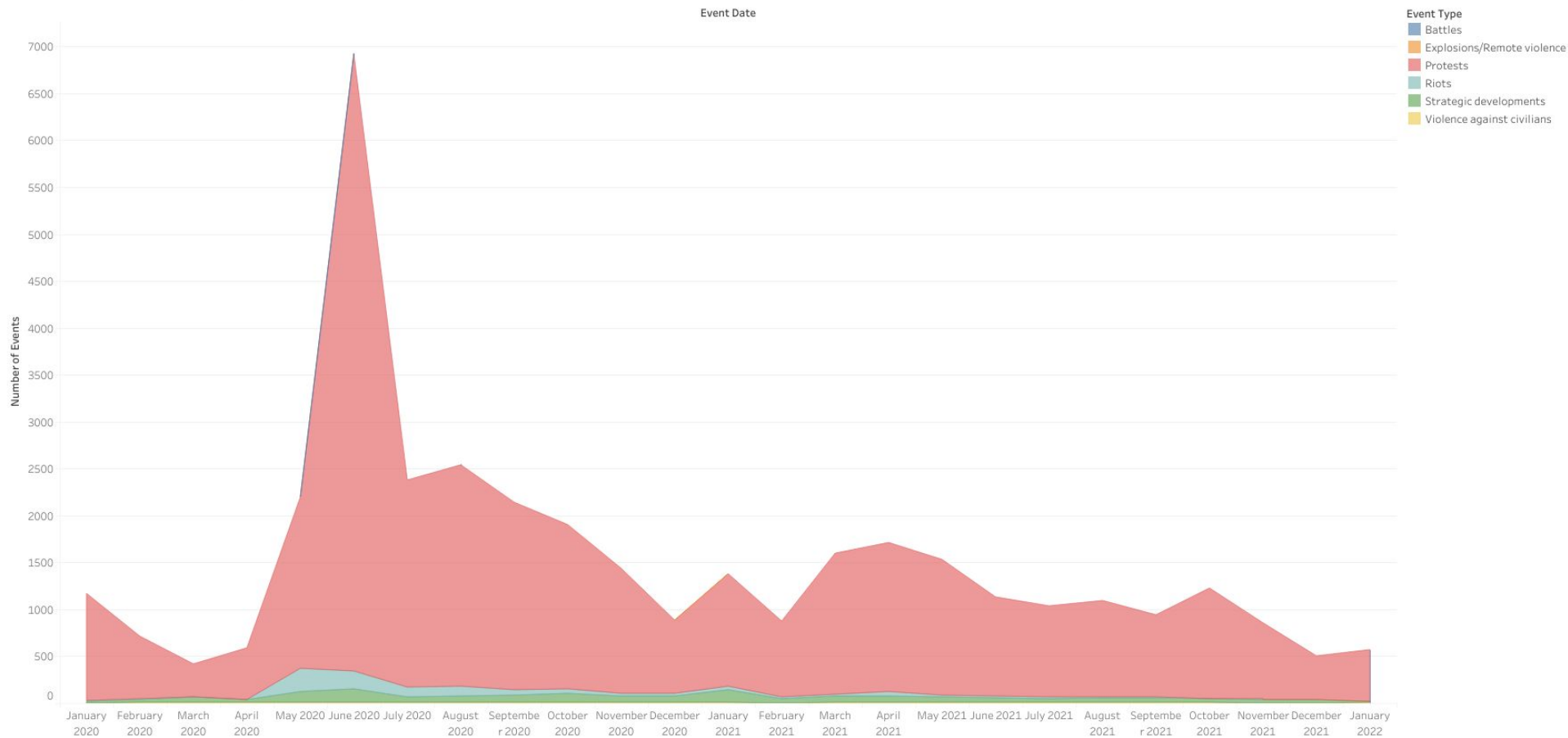




## Events per month

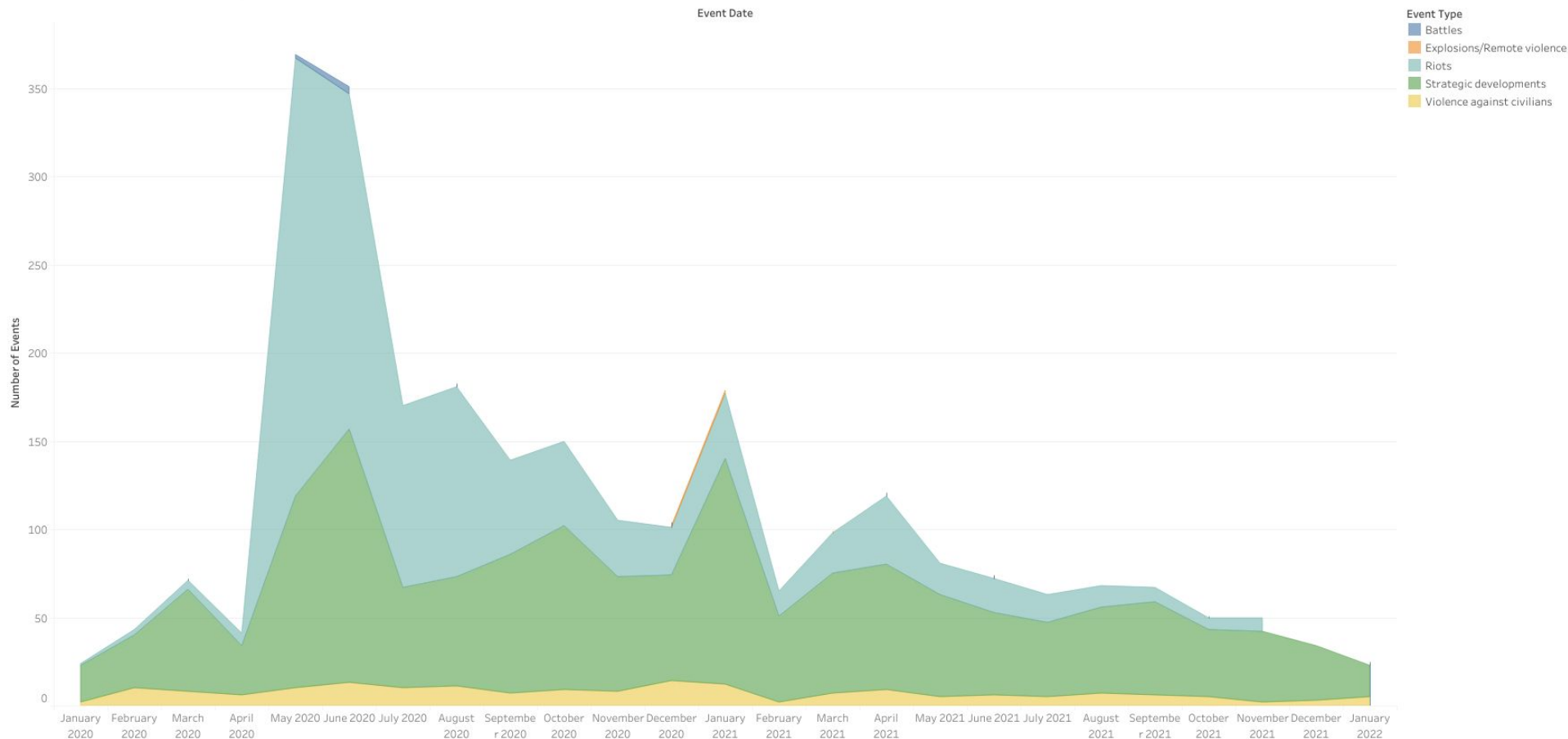


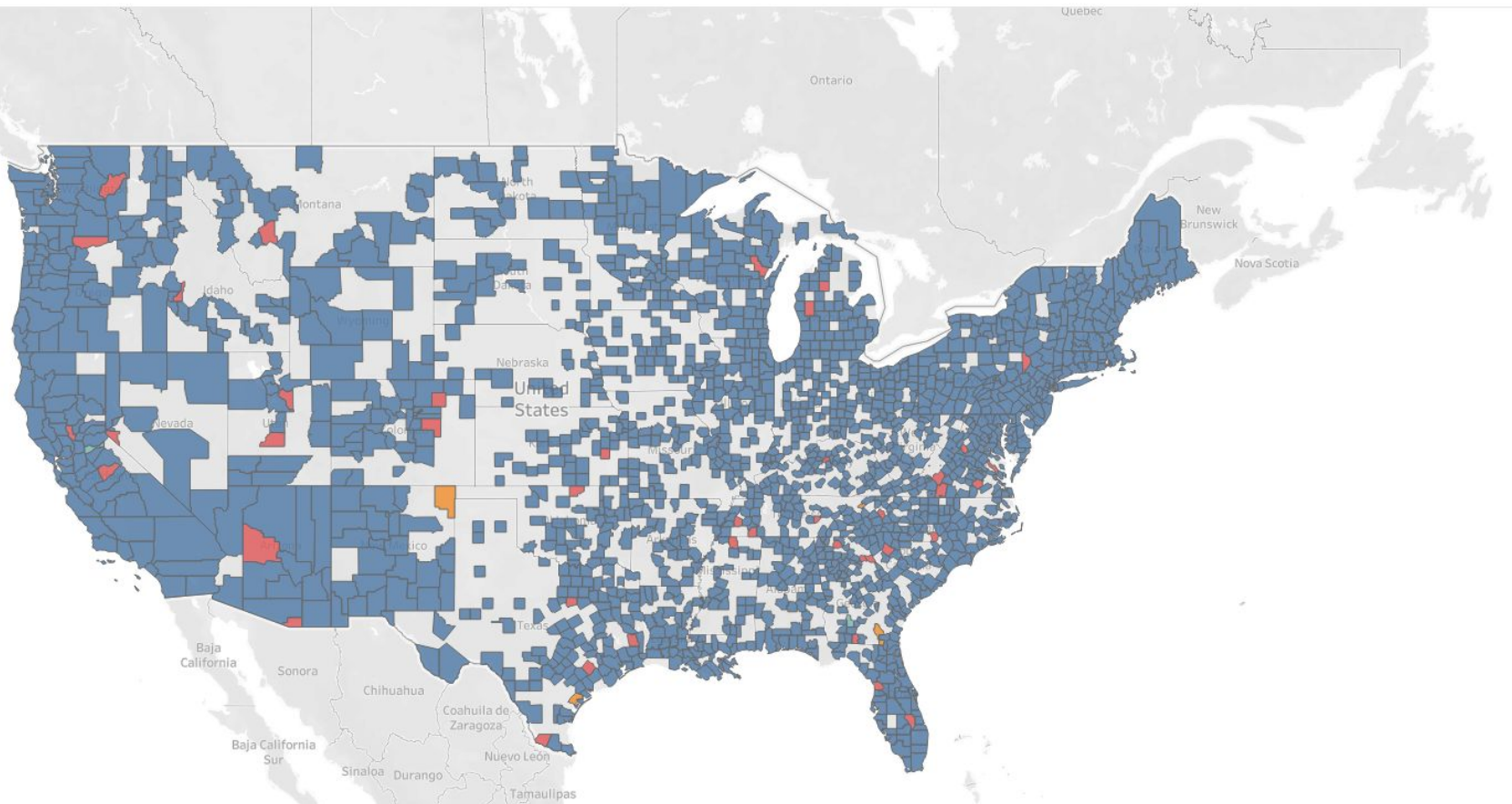
Events per month by event type



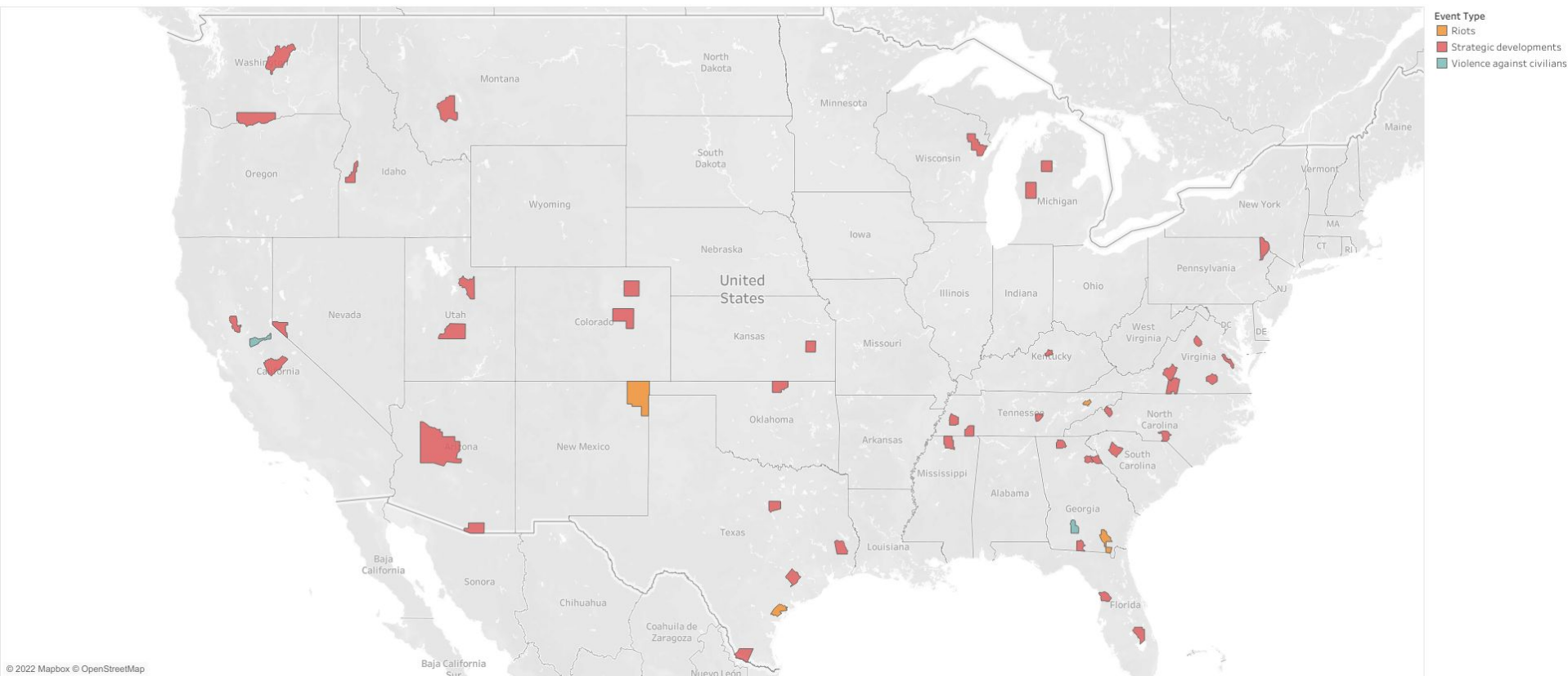


Events per month by event type excluding protests





- Event Type**
- Protests
  - Riots
  - Strategic developments
  - Violence against civilians

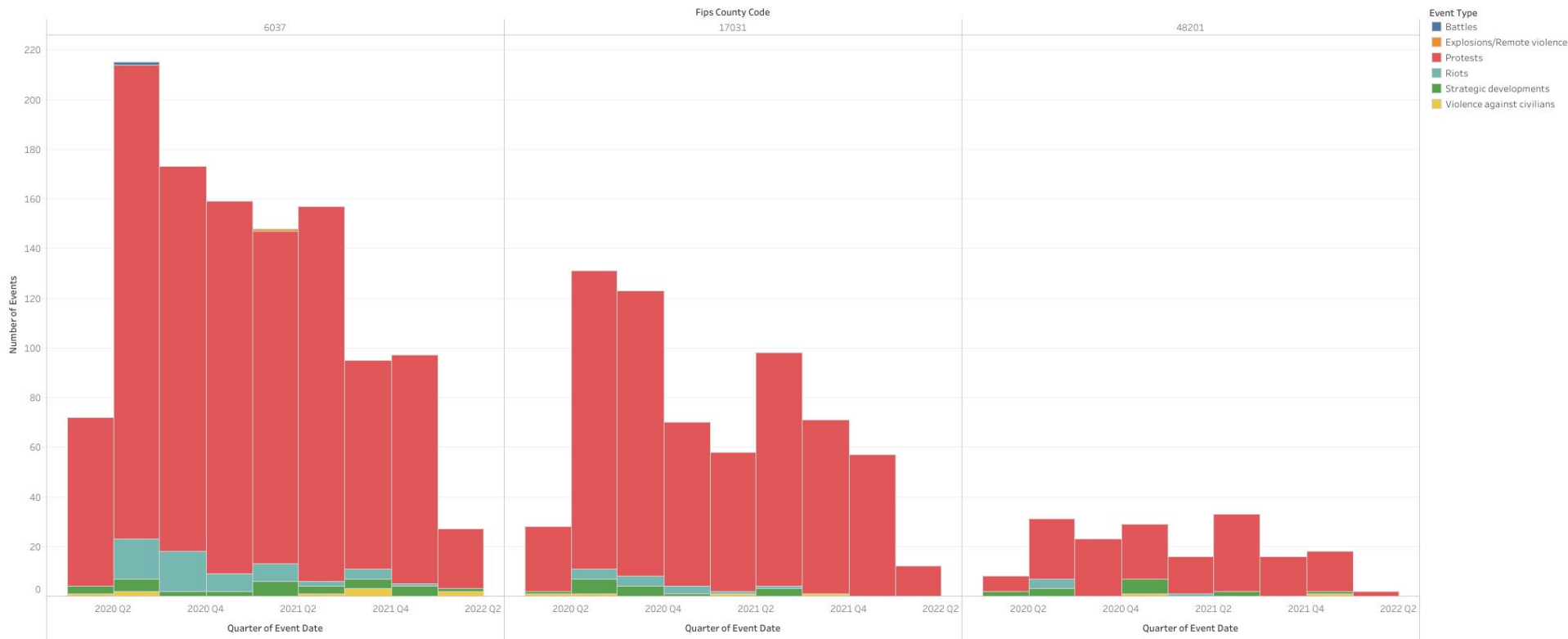


## Los Angeles County, CA

## Cook County, IL

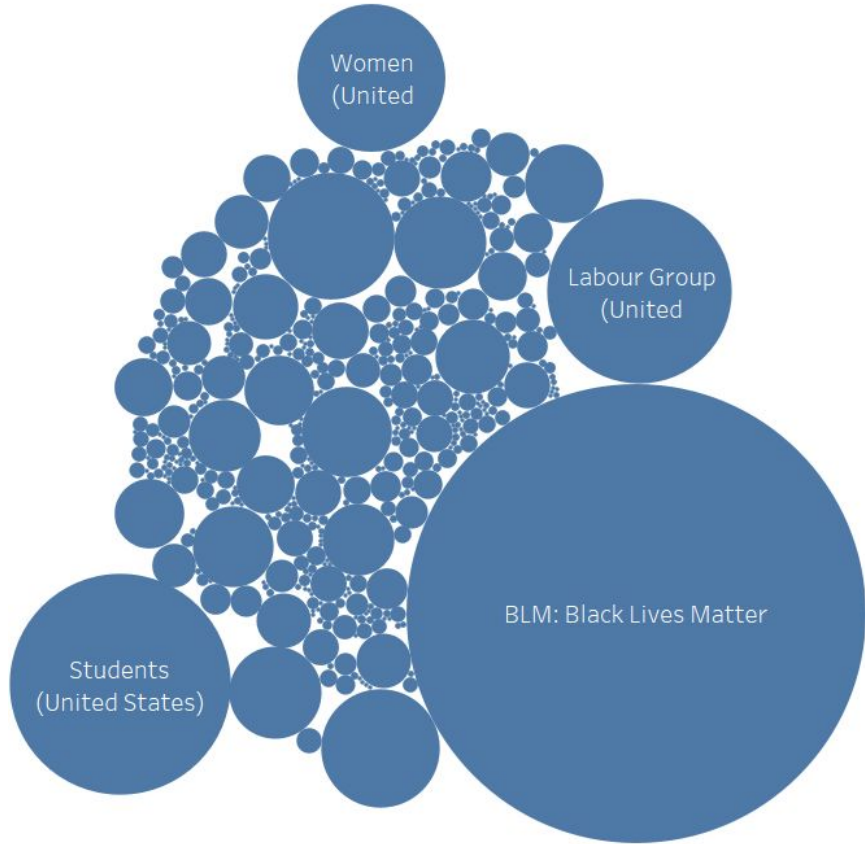
## Harris County, TX

Events in 3 most populated counties

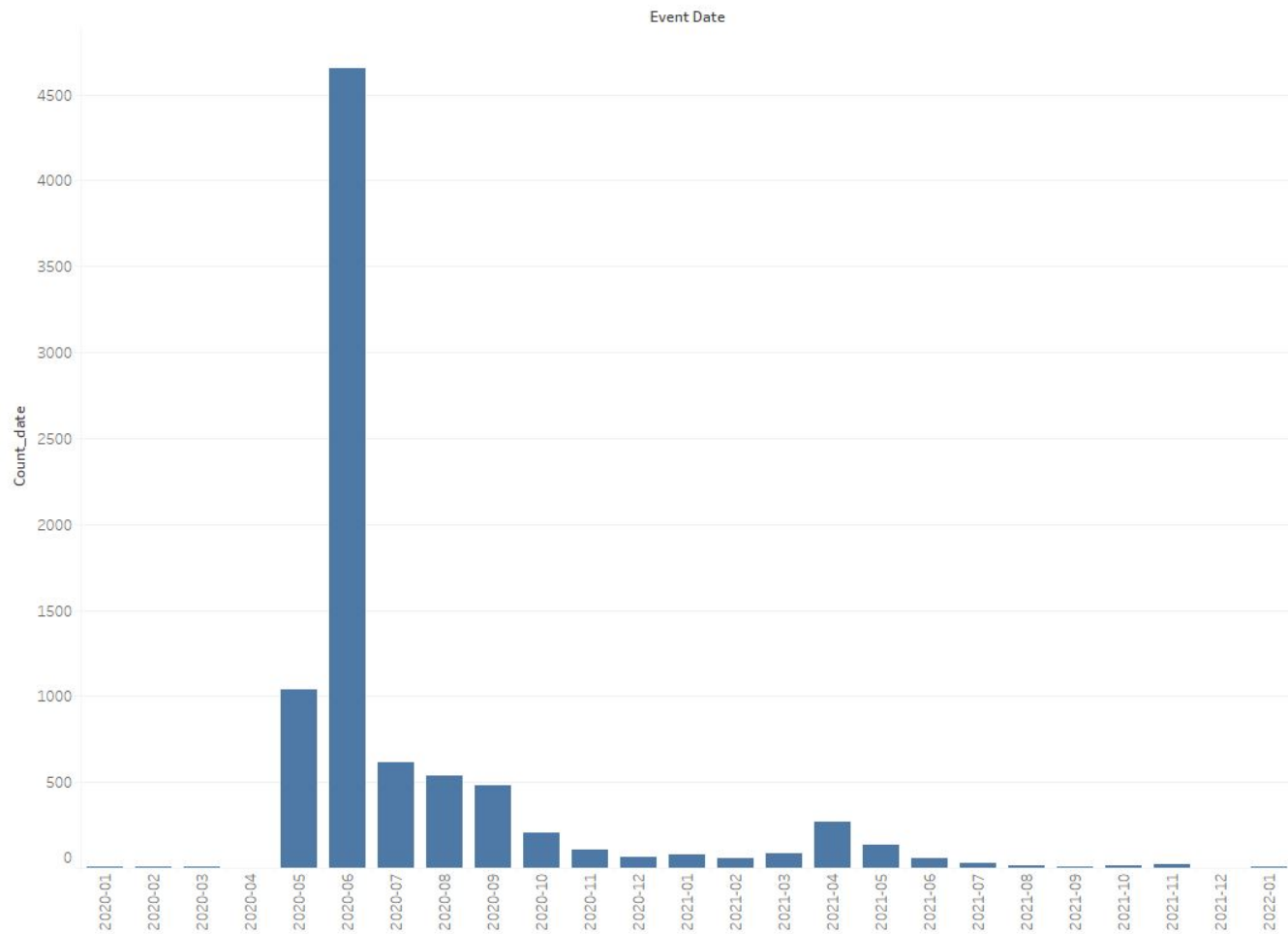


# Digging Deeper...

## Most Common Actors

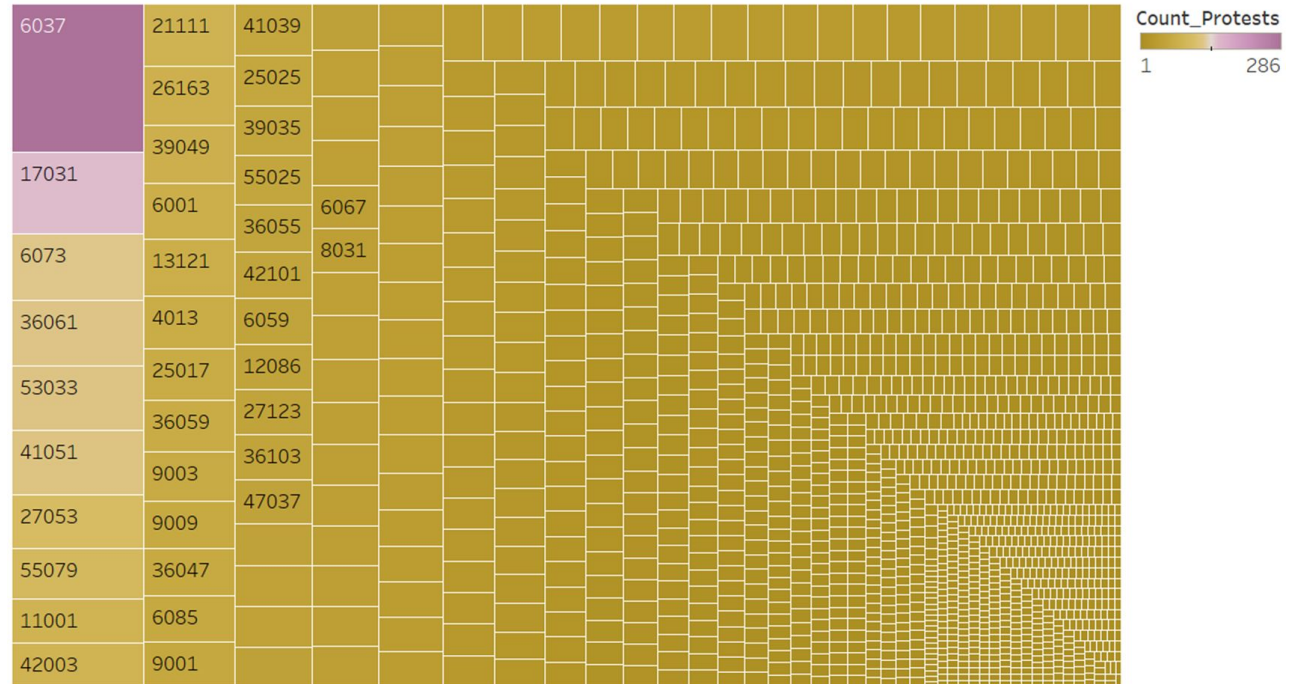


BLM NumEvents Per Date



# Counties with the Most BLM Protests

County Codes with Most BLM Protests

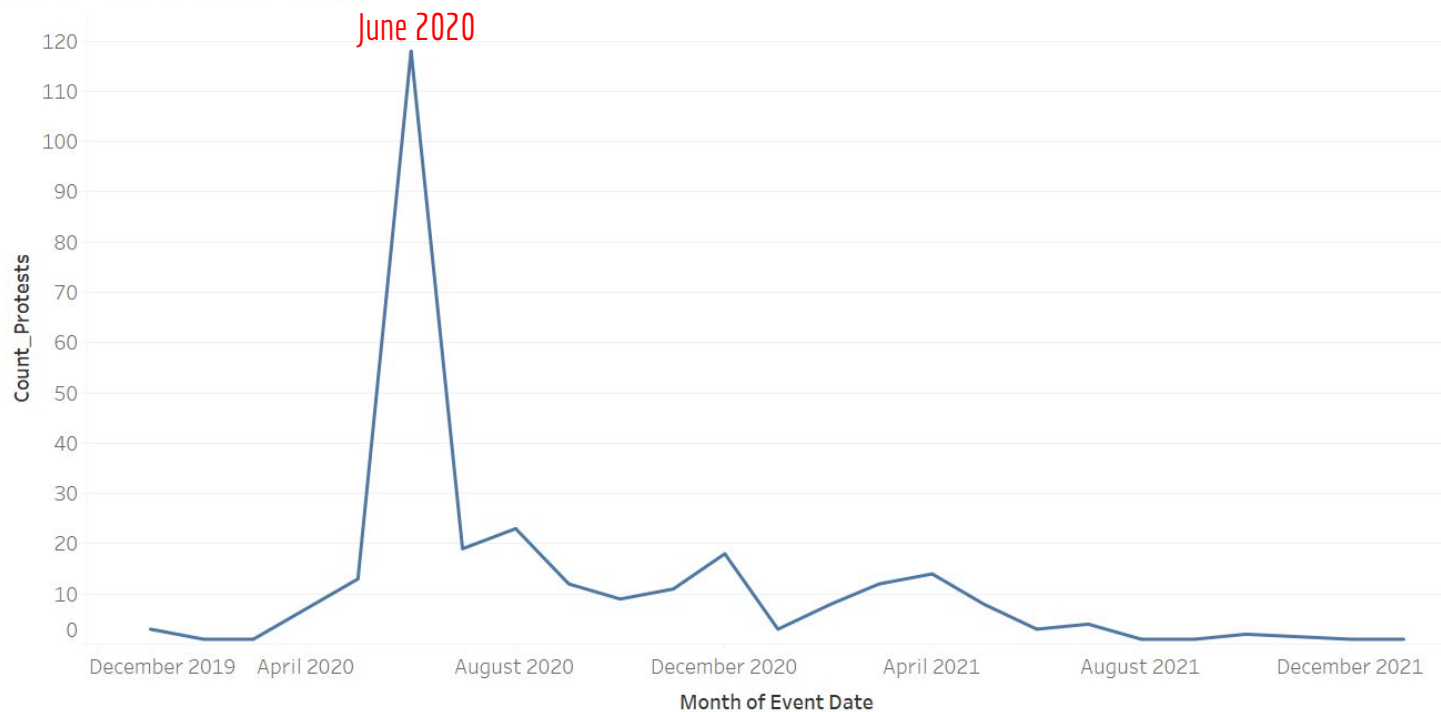


6037: Los Angeles, CA  
17031: Cook, IL



# LA's Protests over Time

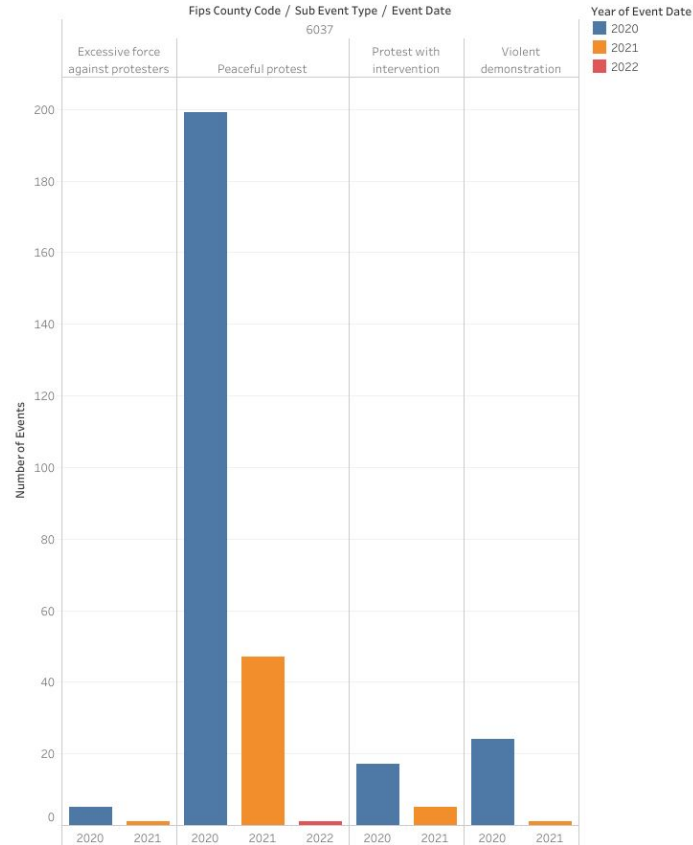
LA's Protests over Time



# Disproving the Myth

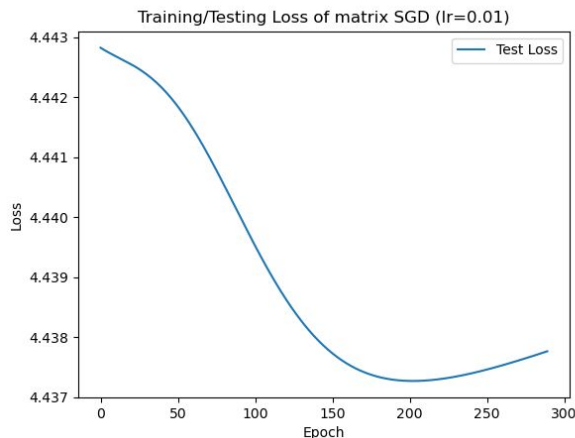
- BLM protests have become increasingly violent and dangerous in LA: NOT TRUE!
- Most are peaceful!

Subtypes of events per year in LA County



# Attempts at Prediction Models

- Low-Rank Matrix Completion



- KNN Classifier

```
In [22]: df['event_sub_type'].value_counts()
```

```
Out[22]: 1      33859  
         2       1041  
         3        949  
        10        945  
         6        439  
         7        173  
         4        143  
         9         82  
         8         62  
         5         38  
        14         15  
        13         10  
        11          9  
        12          4  
        15          1  
        16          1
```

1: Peaceful Protests

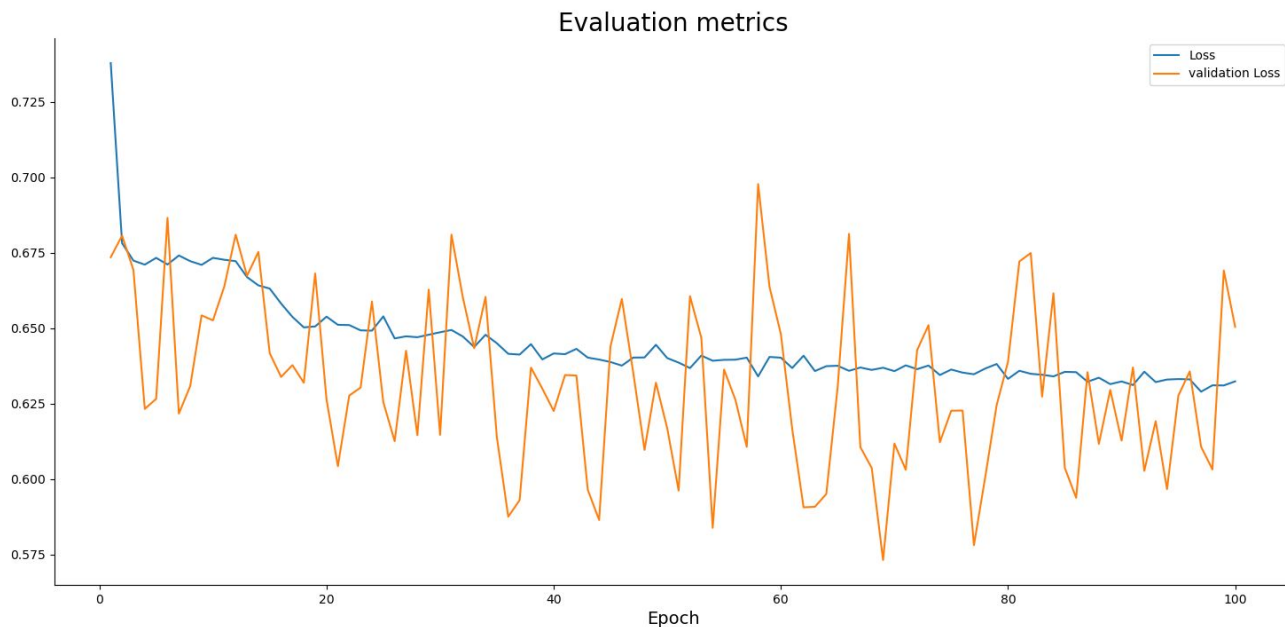
2: Protest with Intervention

```
Name: event_sub_type, dtype: int64
```

# Attempts at Prediction Models

- Multi-label Classification TensorFlow Model

65% Accuracy




# Model Conclusions

- More diverse data required
  - Over Reliance on group 1 data (i.e. Peaceful Protests)
- The model was consistently having issues learning the difference between groups 2-16, because they were all being learned as simply (not group 1) because of the high amount of group 1 data.

1	33859
2	1041
3	949
10	945
6	439
7	173
4	143
9	82
8	62
5	38
14	15
13	10
11	9
12	4
15	1
16	1

Name: event\_sub\_type, dtype: int64



Thank you!  
Questions?

