DATA SCIENCE SEMINAR

HAWKINS GAY, ALEX LEIDNER, RAMSEY WEHBE

THEME

- Our team was interested in learning how police use of force was related to injury in the community.

- Specifically, are there any identifiable and predictable patterns that have a higher likelihood of resulting in injury, and if so, perhaps these can utilized to reduce injuries in the future.

- Specific Questions for Analysis:

  → What percentage of use of force events result in injury?

  → Are events that to lead injury related in anyway to demographic classes or associated with specific neighborhoods?

  → Do certain types of use of force events (physical, verbal, weapon choice) lead to higher rates of injury?

  → Are there any individual officers that are frequently involved in injury events?

- These topics were addressed with descriptive and predictive statistics.

- Visualizations were created to help highlight specific patterns.
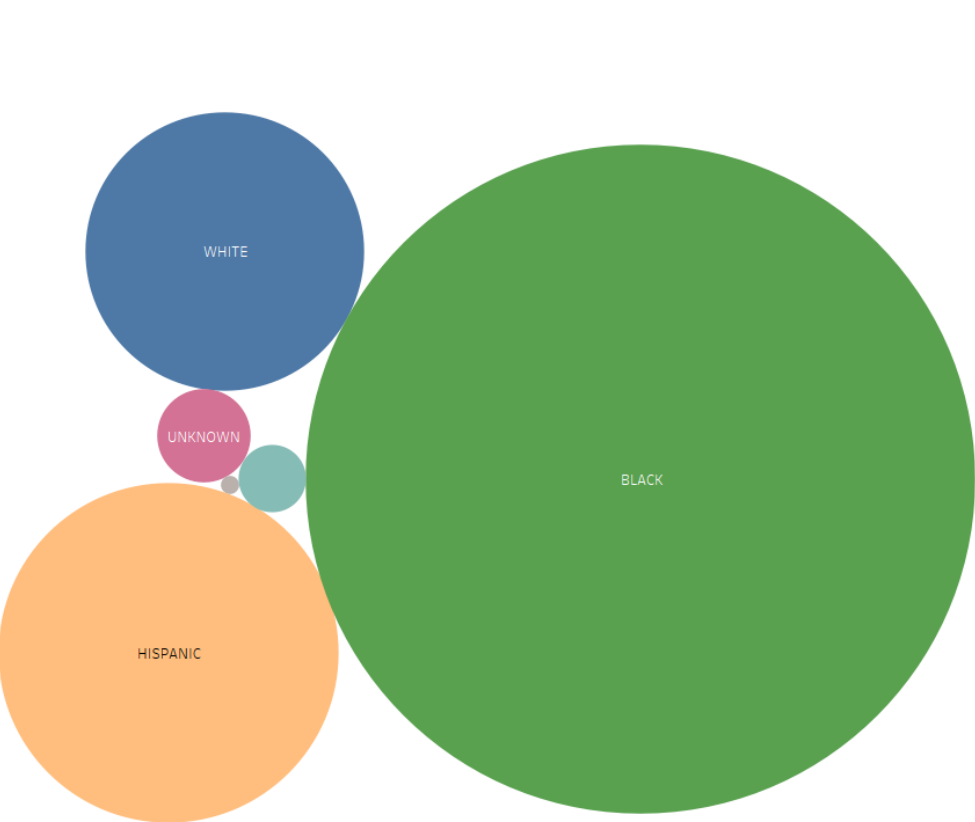
DEMOGRAPHIC FINDINGS FROM DATABASE

- CPDB contains 67,819 use of force events
- 26% resulted in subject injury and 22% resulted in officer injury
- Largest total injuries to black individuals (12,217, 24%)
- Largest percentage injury to Hispanic (33%) and white individuals (32%)
  - → Asian officers had the highest percentage of injury: 27%
  - → The race / ethnicity of the officer was not related to injury events, all officer groups fell between 22-24%
- Age range 18-40 had the most injuries(14,815)
- Older citizens were more likely to injured 40-65 (28%), >65 (28%)
- Injuries were more common among males citizens and female officers
- Firearm or Taser were more likely to result in injury, at 64% and 40%
  - → Firearms or Taser involvement similar across citizen and officer race
- Officers varied heavily in propensity for injury, some as high as 50-70%

INJURY BY RACE / ETHNICITY

Use of Force Events Resulting in Injury

Percent of Force Report Resulting in Subject Injury [Race / Ethnicity]

Subject Race

- ASIAN/PACIFIC ISLANDER
- BLACK
- HISPANIC
- NATIVE AMERICAN/ALASKAN NATIVE
- UNKNOWN
- WHITE

Subject race (Sheet1). Color shows details about subject race (Sheet1). Size shows sum of Subject Injuries. The marks are labeled by subject race (Sheet1).
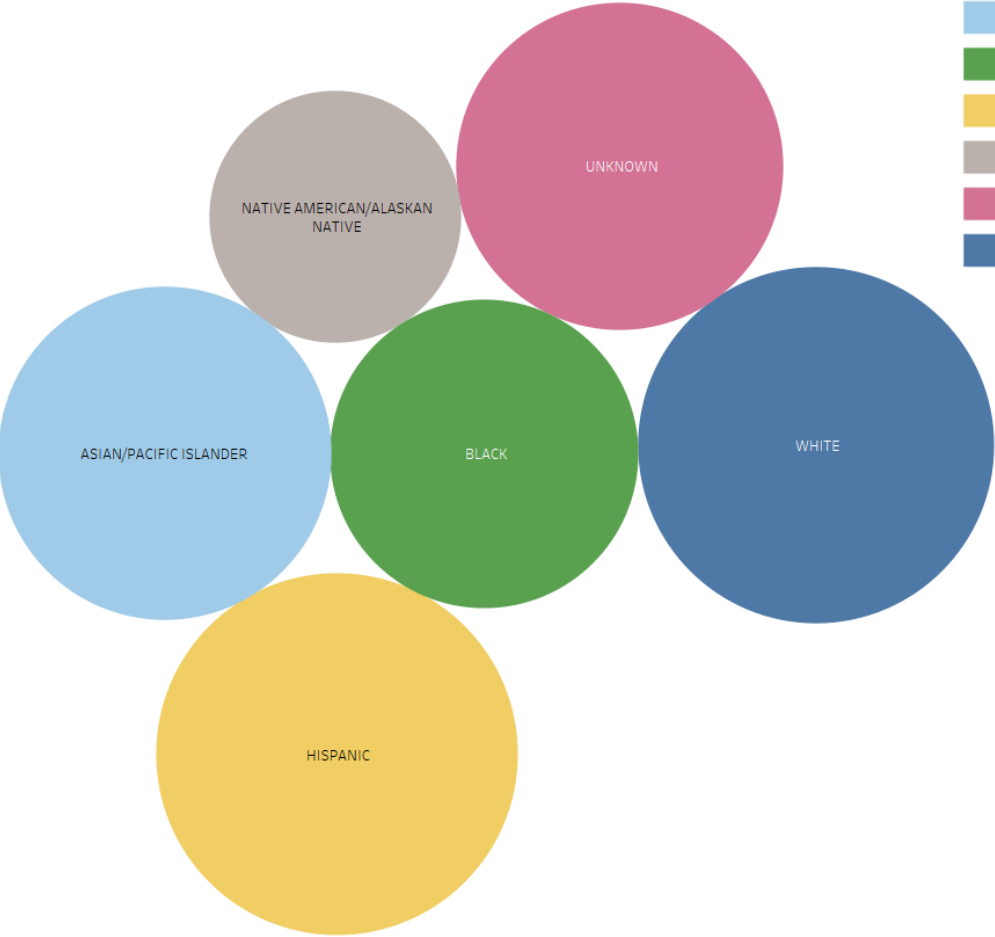
Subject Race. Color shows details about Subject Race. Size shows sum of Percent Subject Injuries. The marks are labeled by Subject Race.

INJURY GENDER AND RACE

Total Injuries by Gender and Race
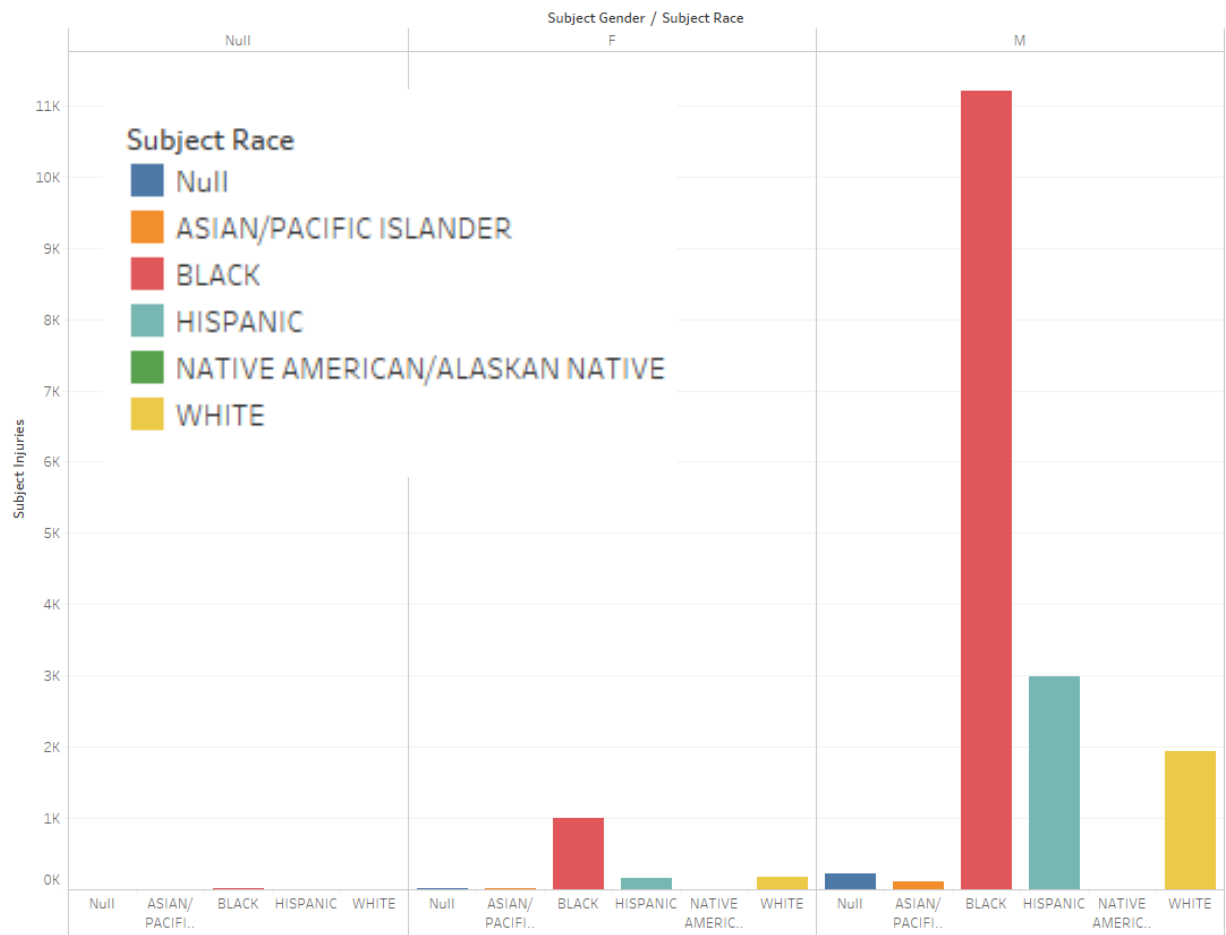
Percent of Force Resulting in Subject Injury

Sum of Subject Injuries for each Subject Race broken down by Subject Gender. Color shows details about Subject Race.
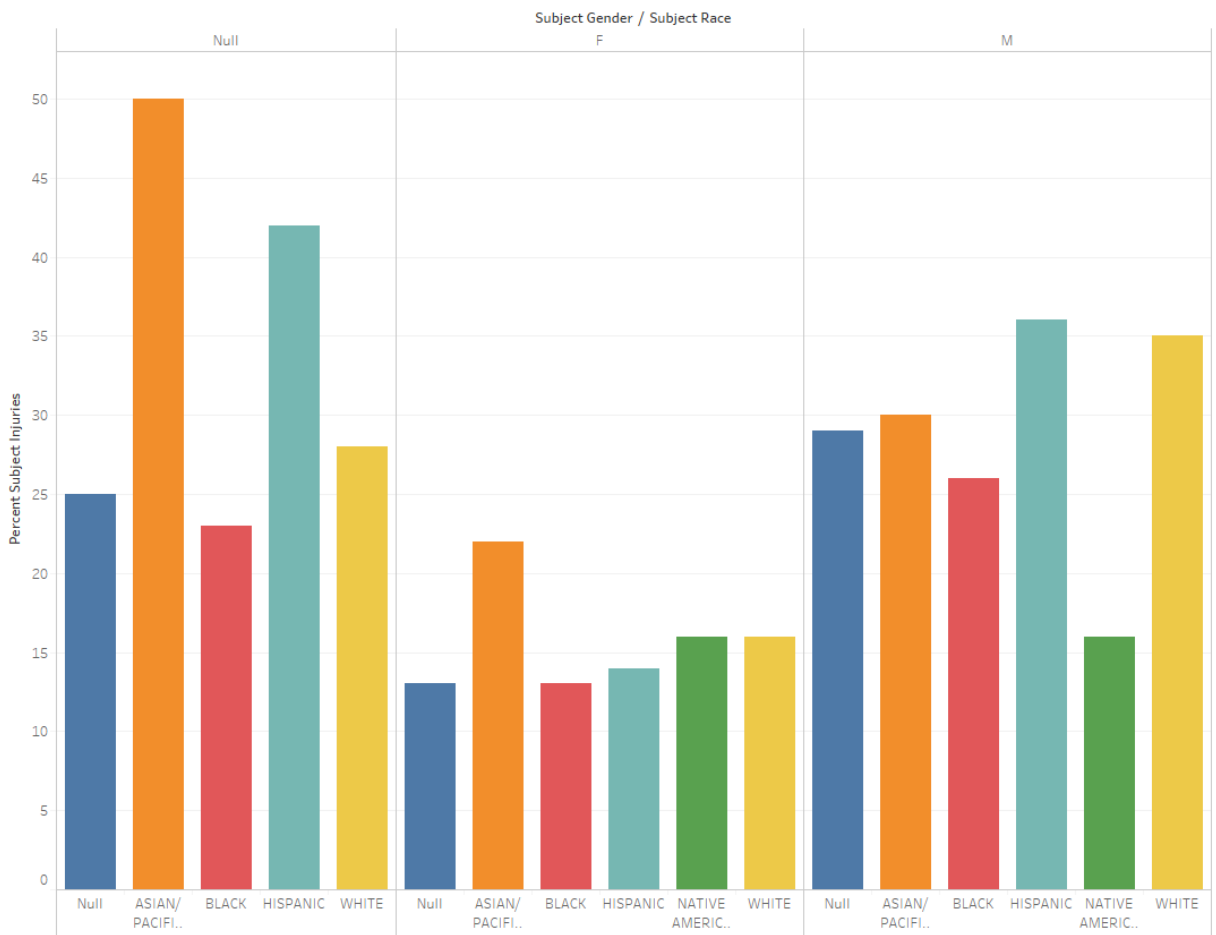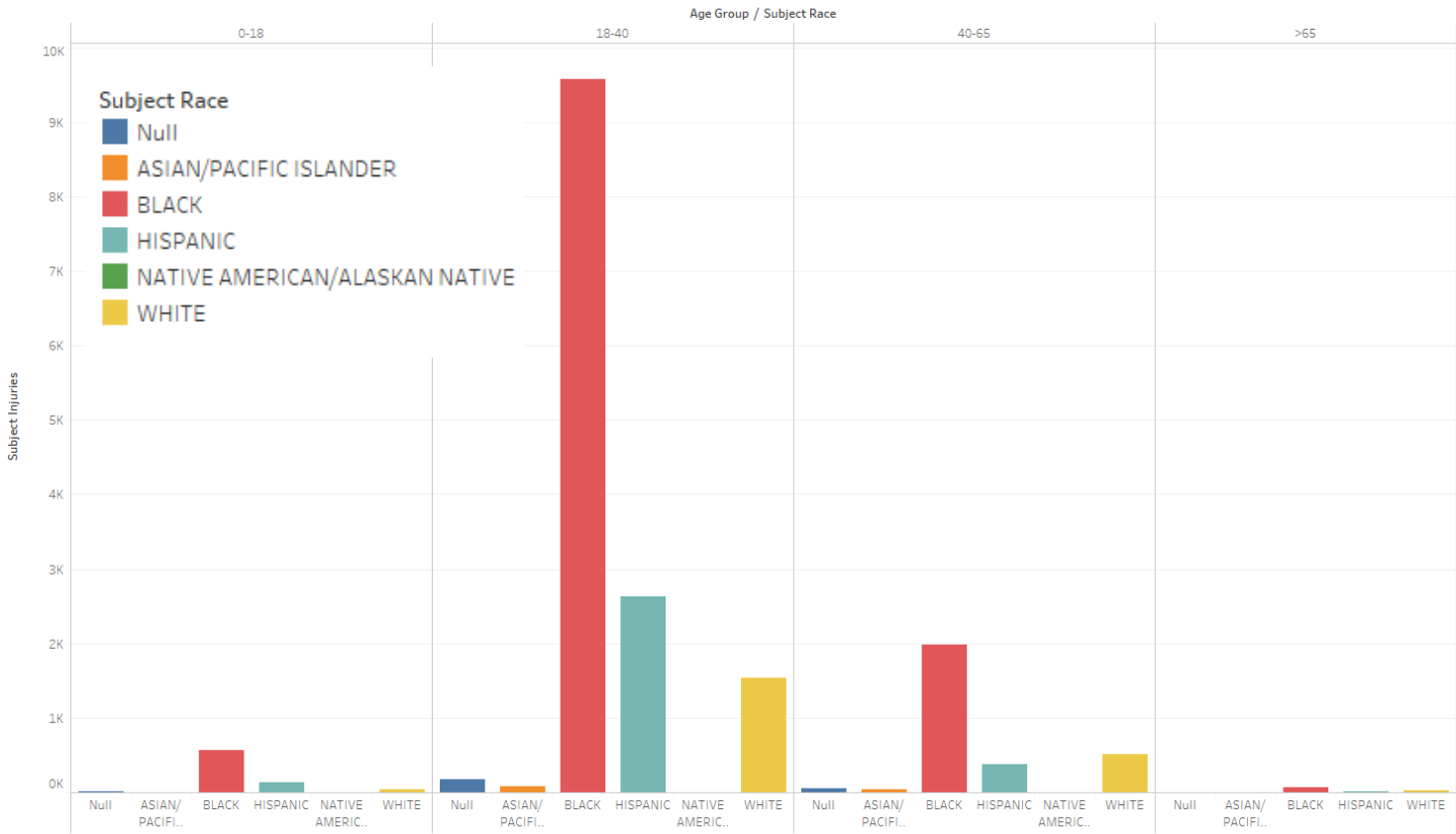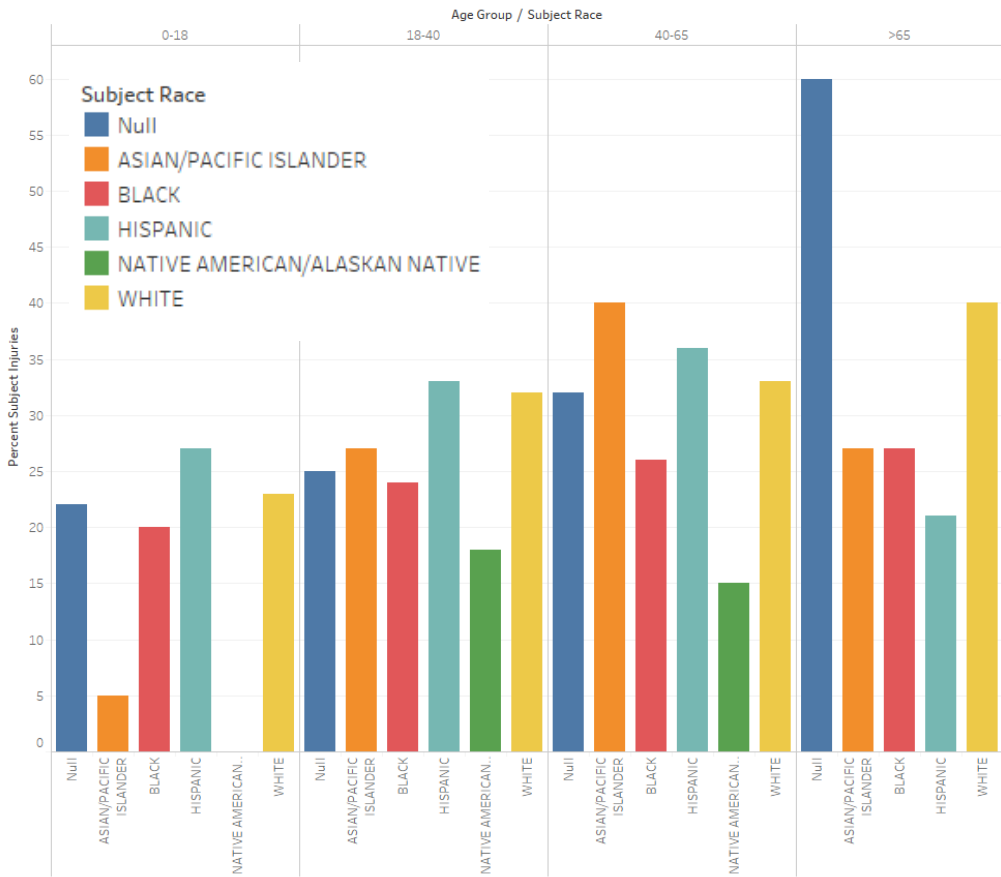
Sum of Percent Subject Injuries for each Subject Race broken down by Subject Gender. Color shows details about Subject Race.

INJURY AGE AND RACE

Total Injuries by Age and Race

Percent of Force resulting in Subject Injury

subject_injured

255   1,881

percent_subject_injured

22   35

Total Injuries Per District

Percent Force Resulting Injuries Per District

© 2020 Mapbox © OpenStreetMap
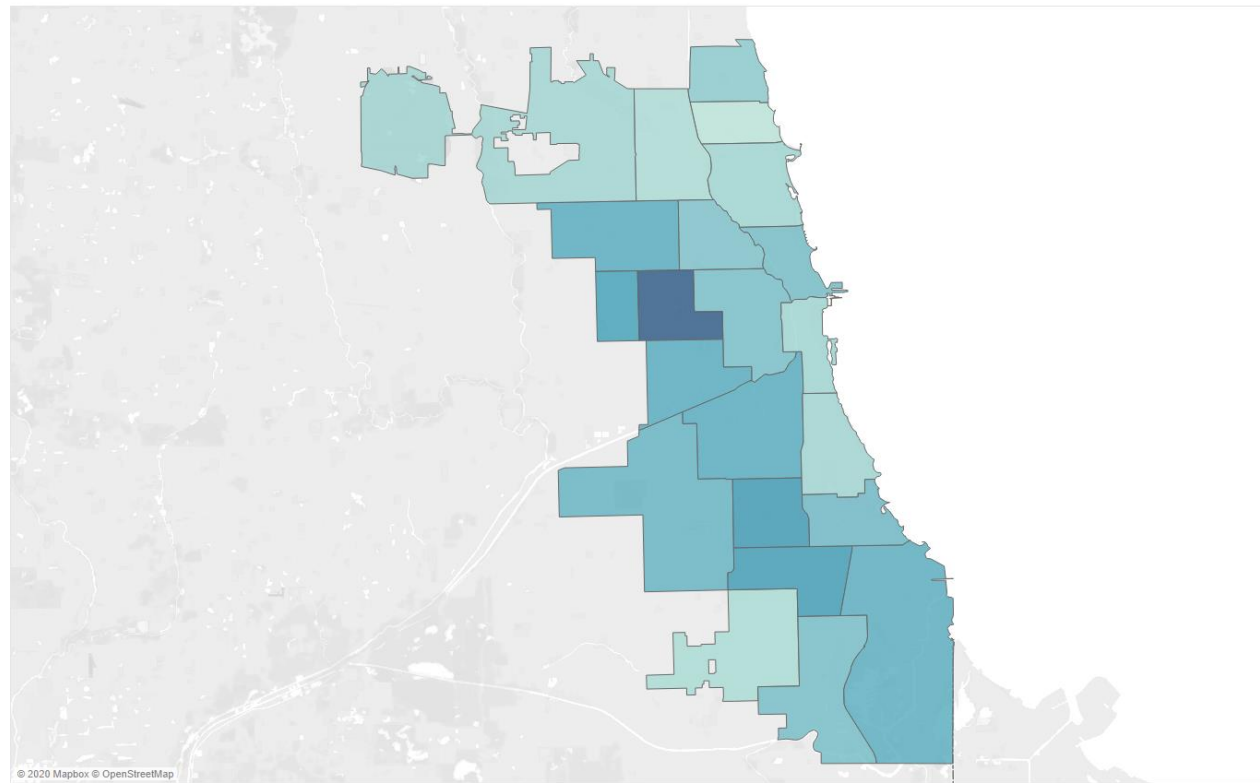
© 2020 Mapbox © OpenStreetMap

Map based on Longitude (generated) and Latitude (generated). Color shows sum of subject_injured. Details are shown for district_name.
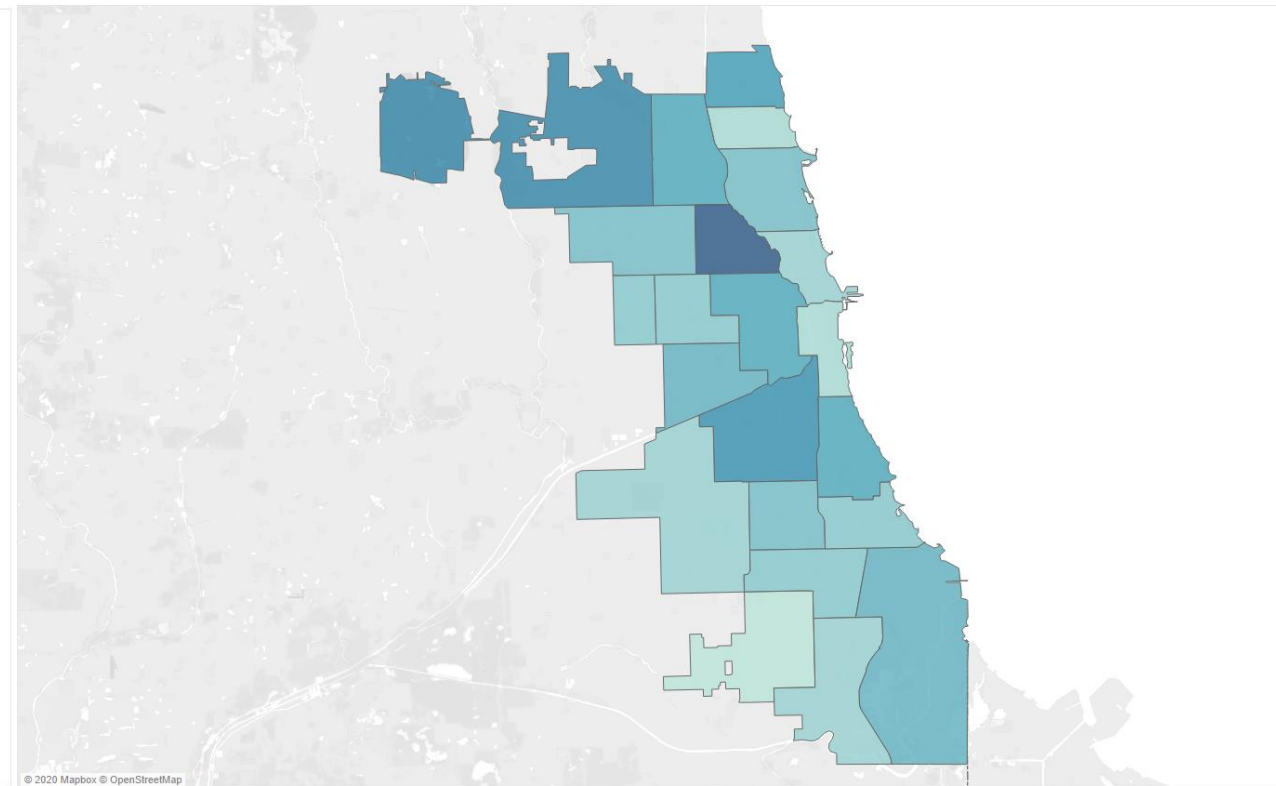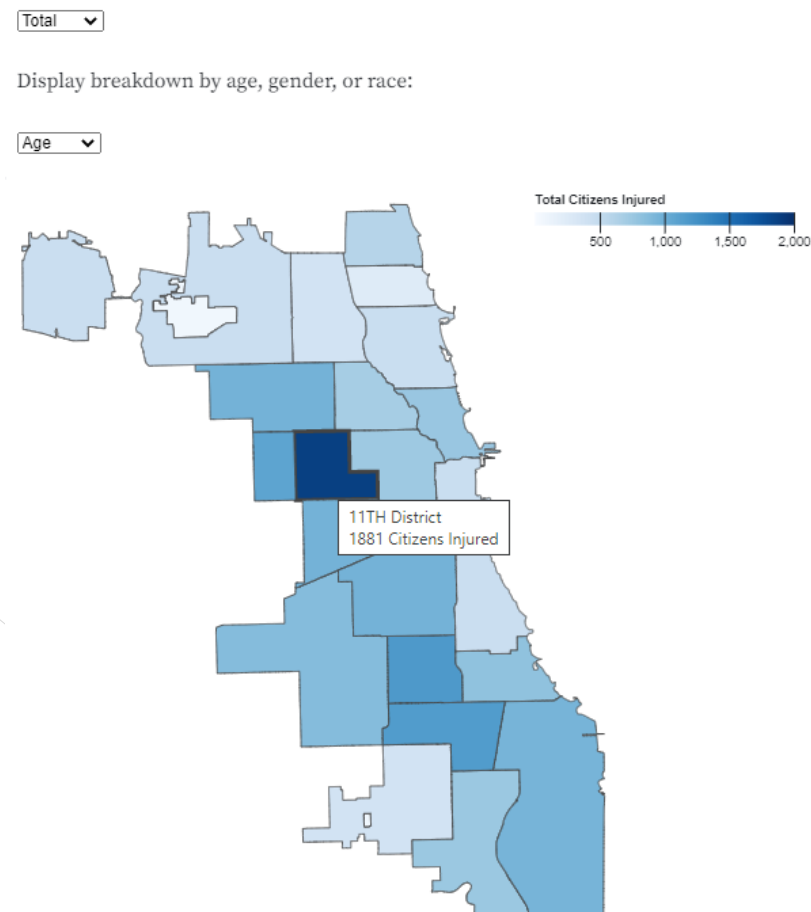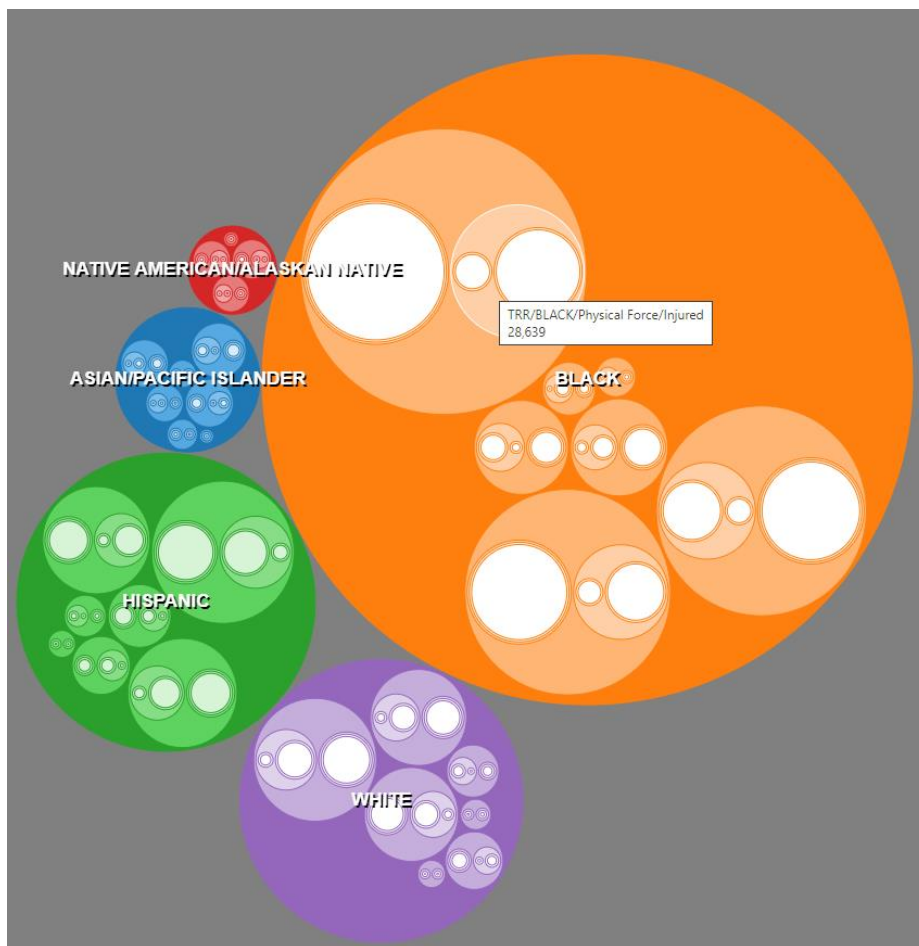
Map based on Longitude (generated) and Latitude (generated). Color shows sum of percent_subject_injured. Details are shown for district_name.

INTERACTIVE QUERY

- Can be viewed: https://observablehq.com/d/47bd6e39f1003039; https://observablehq.com/d/befde87915d165a6
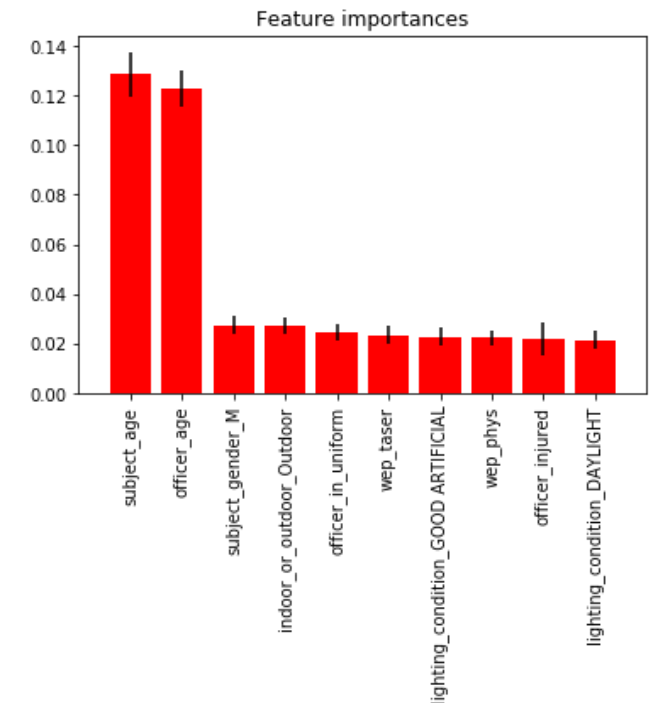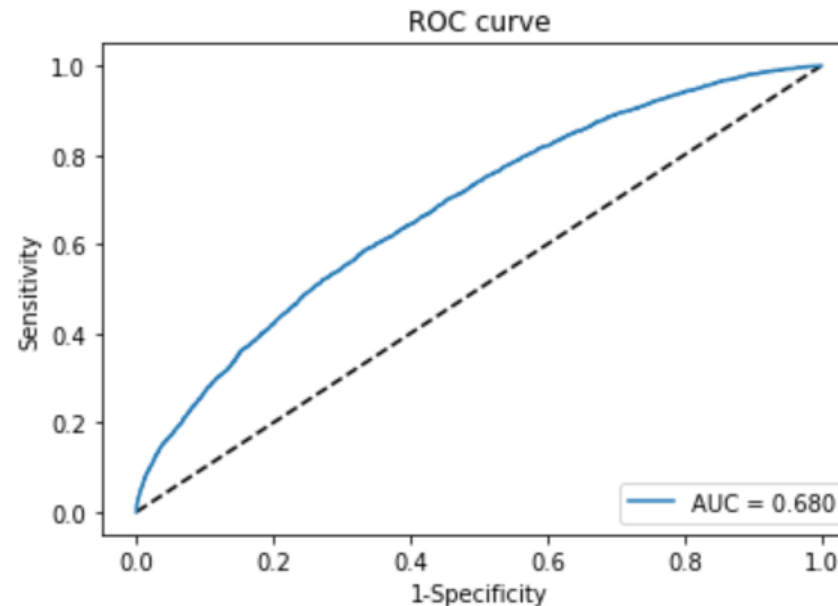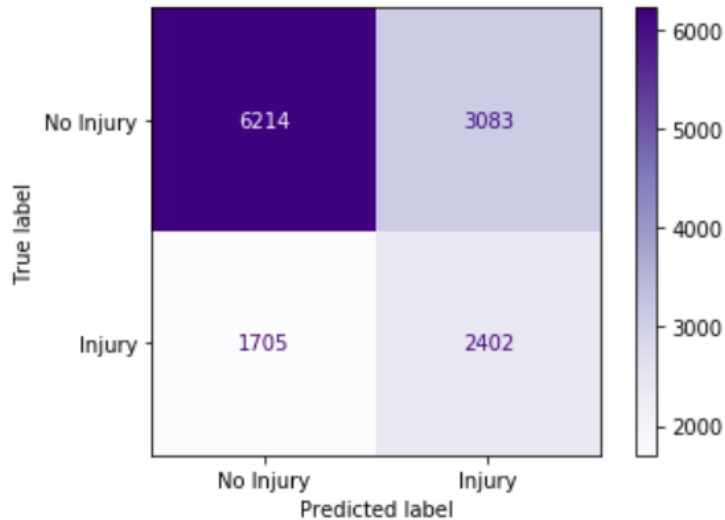
PREDICTIVE MODELING - INJURY

- We evaluated a number of classification models to assess the predictability of injury based on the features included in the CPDP database
    - → Models assessed: Random Forest, Logistic Regression, k Nearest Neighbor
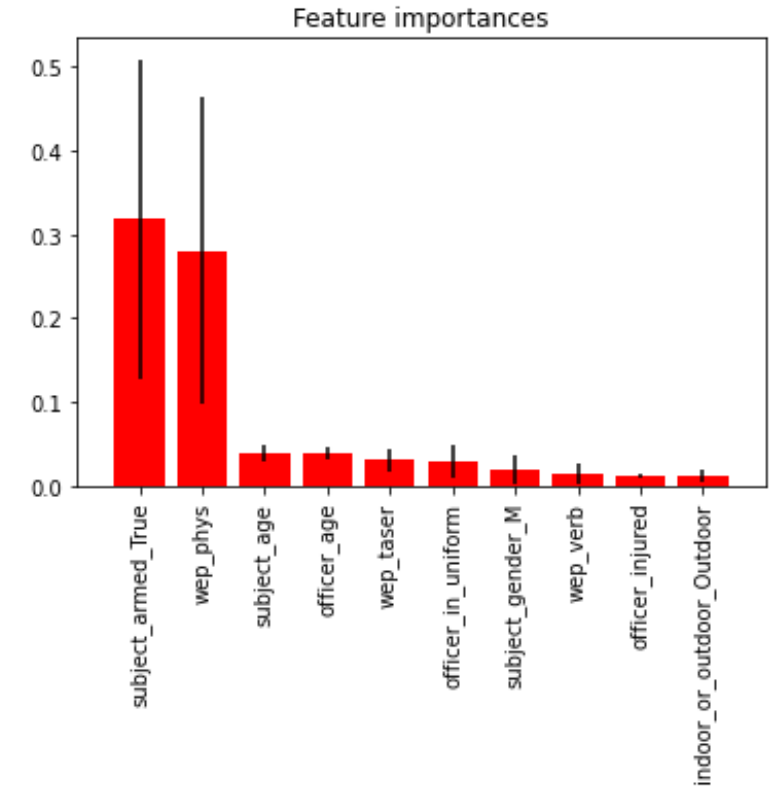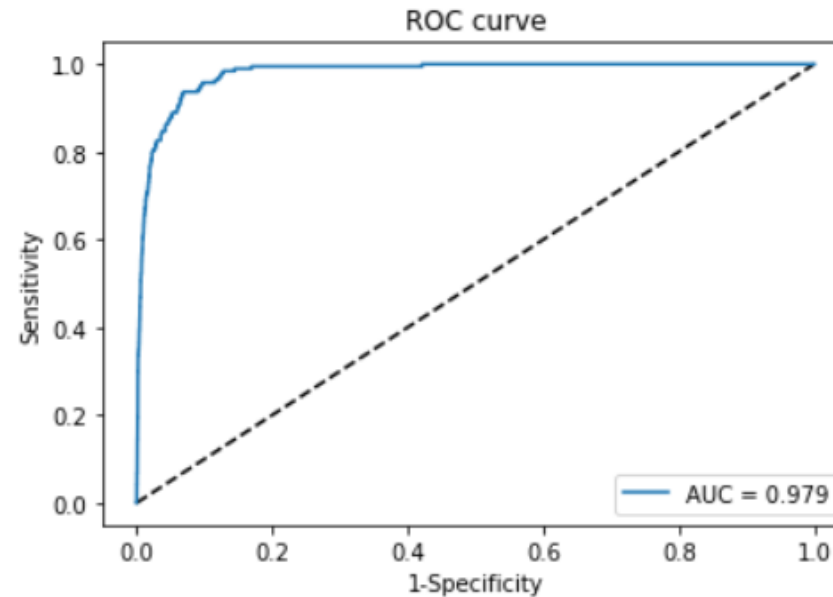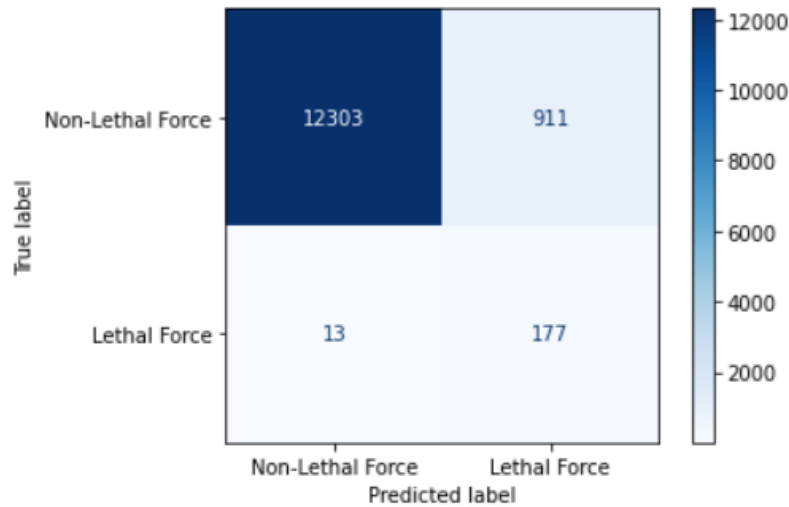   - → Best was Radom Forest – *accuracy 64%*

PREDICTIVE MODELING – LETHAL FORCE

- We utilized a random forest model to assess the predictability of lethal forces

- To fine tune the hyperparameters for best fit, we utilized the GridSearchCV function
    → Best fit model – *accuracy: 93%*

RESULTS

- Black individuals in Chicago are injured more by police use of force than other demographics, especially considering 2010 census reports

- There are geographic hot spots for total injury and injury rate

- Rates of injuries appear similar amongst many officer and subject demographics

- It is difficult to predict which use of force events result will result in injury

- Predicting lethal use of force by an officer is possible with a high degree of accuracy

QUESTIONS FOR FUTURE INVESTIGATION

- Normalize our observations about demographics by city and police district using Chicagoland census data

- Can percent use of force event resulting in injuries be used to identify outliers in both officers and police commands?

- Can we determine injury severity through NLP and analyzing complaint reports for invoking medical assistance?

  → Correlate with hospital proximity

- Can we correlate complaint text to subject alleged or officer confirmed injury?