Identifying Crime Patterns

Liam O'Farrell

The Question

- What city characteristics are most closely related to crime rates?
  - Can a model be constructed to identify features that correlate to crime rates in specific areas?
  - Goal is to identify features that can effectively reduce crime

Data

- Crime data from Chicago, IL from 2001-2015
- All reported crime including type, date, location, county and weather or not an arrest was made
- Shape: (1041703, 23)

Data Exploration

- Removed null values from data set
  - No idea what would be appropriate replacements
  - Over 1 million data points so it shouldn't effect the results that much
- Features – first identify features that are correlated with crime
  - Neighborhood
  - **Distance from police station**
  - Weapons
  - Time of day
- Response
  - Apprehension

Distance Formula

- nearest_pt = []for i in xrange(len(crime_lat)):

  initial_dist = 1000

  for j in xrange(len(station_lat)):

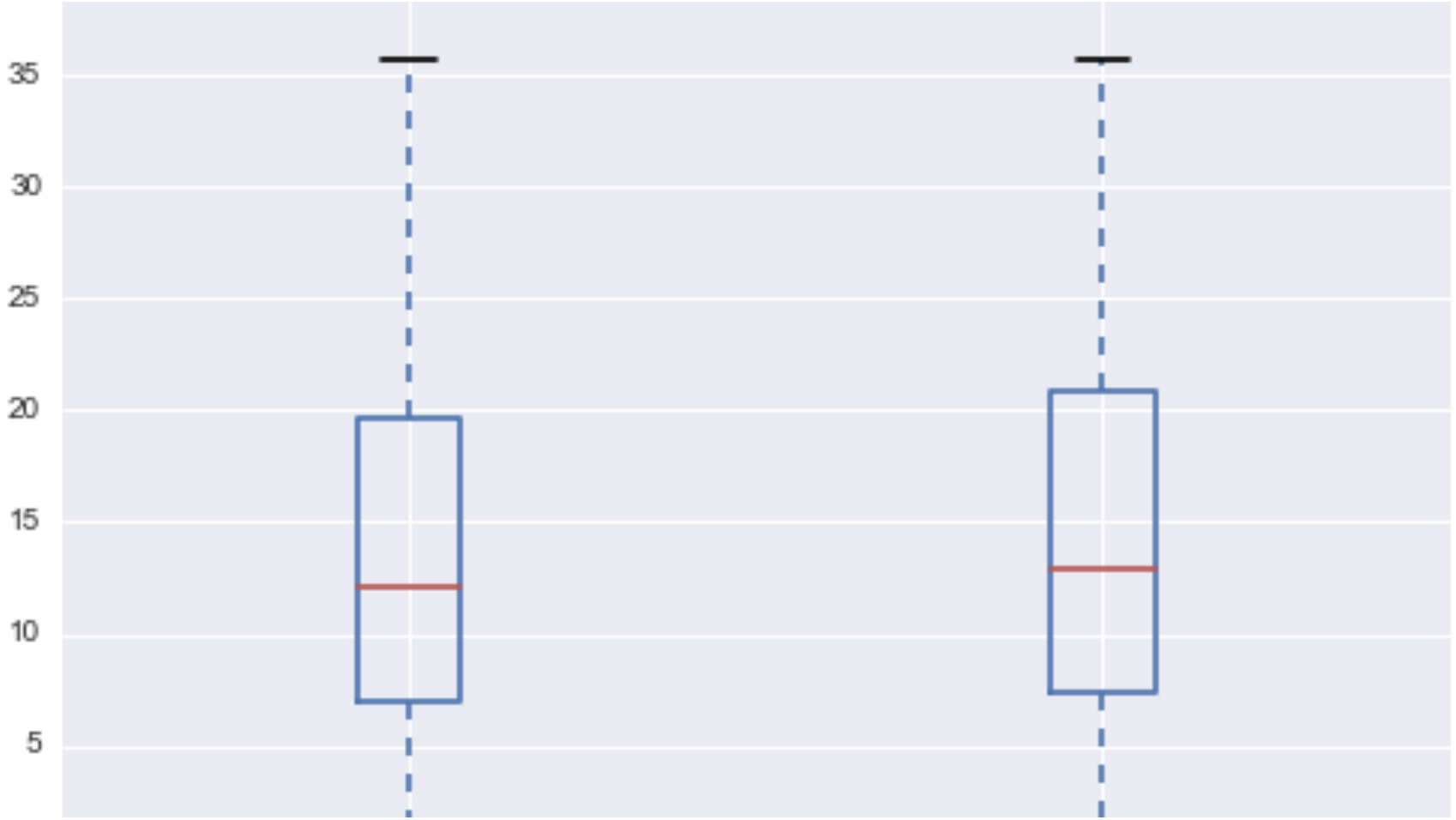
    dist = haversine(crime_lat[i], crime_long[i],station_lat[j], station_long[j])

  if dist<initial_dist:

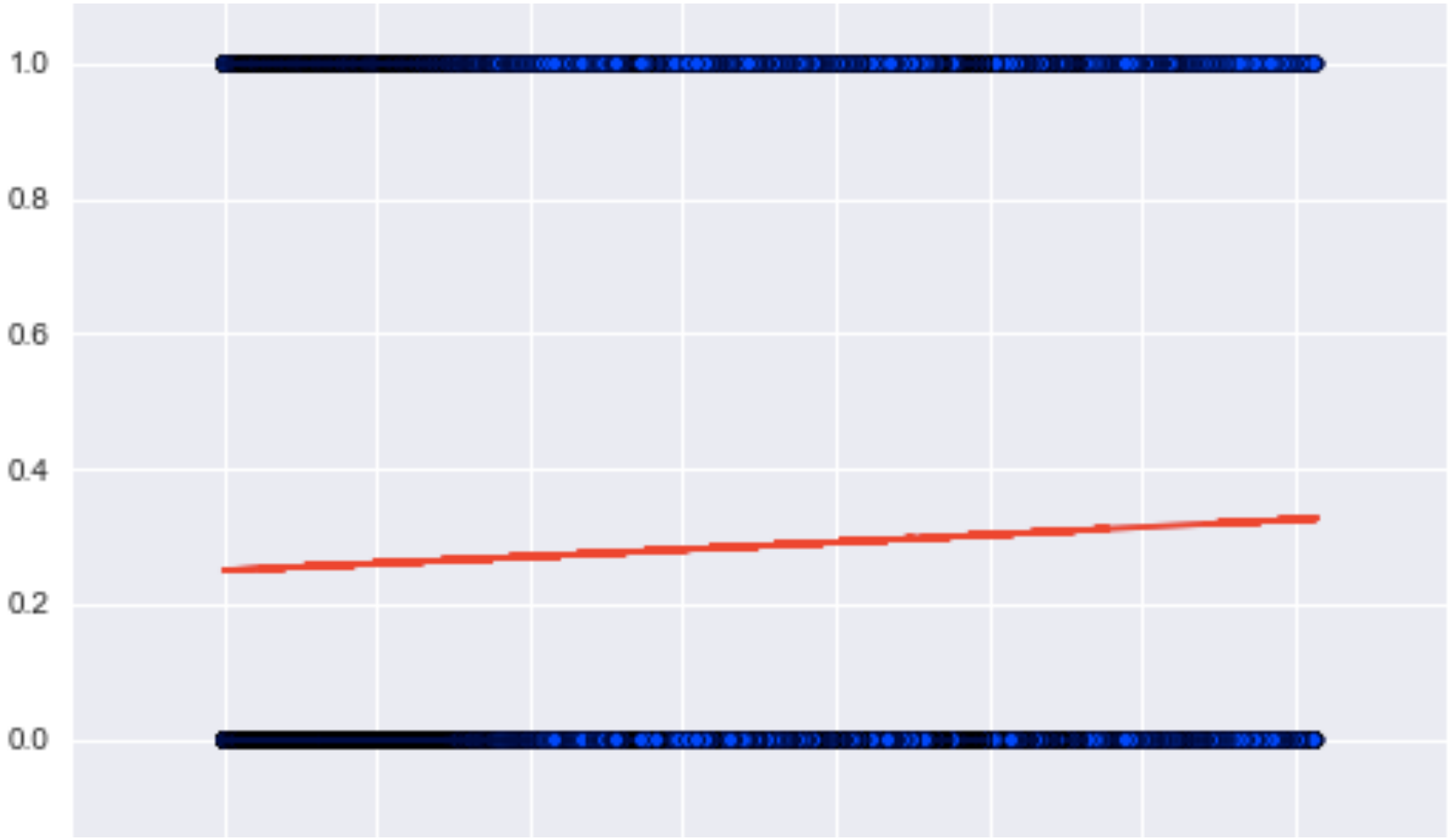
    initial_dist = dist

  – nearest_pt.append(dist)

Box plot of Arrest and Distance



Regression Arrest v. Distance



Next Steps

- Improve model
- Explore other features
- Classify crime severity
- Gather data on current and previous police programs
  - Deterrence programs
  - Public safety budget
  - Number of officers
  - Law changes