I would like to investigate the role the police unit plays in allegations. Do we find that different characteristics about units result in higher or lower rates of allegations? What role does the make up of police units play? By investigating key features of the police unit and tying them out with the allegation data, we can look for patterns that may emerge. Ideally, we will uncover characteristics about lower offending units and identify potential changes to be made to the compositions of police units. For example, do units with more ethnic diversity have lower rates of allegations? Do units with LEO’s that have been in the same unit for a long time see lower rates? By answer questions like these, we may be better able to construct a police unit in a manner that is correlated with lower rates of allegations.

Relational Analytics

* How large is each police unit?
* The average police units?
* Many background questions such as, for each unit, the average:
  + Age, years on force, % male vs. female, number of different races, etc.
* Number of complaints filed against members of each different unit, aggregating to average number of complaints per LEO in each unit
* How often are members of the same unit coaccused/how many different units are identified on an allegations?

Data Exploration

* Histograms of key characteristics (same as those listed above) for different units so we can compare the homogeneity of each unit.
* Distribution of complaints filed against member of each police unit. Are there few units with many complaints or are the complaints fairly dispersed across each unit?
* Scatterplot with key characteristics of units vs. number of complaints against the unit
* Calculation of average entropy of each characteristic of a unit, to indicate the homogeneity of a unit, vs. the number of allegations against the unit

Interactive Visualization

* Density plots to compare the different densities of demographic information between CPD units and those filing complaints. I expect to see very different density plots for the two groups if my hypothesis that differences between individuals leads to interactions results in in complaints. Interactivity allows for including multiple different demographic and non-demographic (average time on force, time in unit, etc.) values to be selected within a filter.
* Streamgraph detailing the differences over time of complaints between members of specific units. Can we see a change in the rate of complains as the make up of units changes? When new members join or old members leave a unit, what changes in complaint rates do we see?

Graph Analytics

* Is an officer working in a unit with a member of a specific ethnic group less likely to have an interaction with a citizen of that ethnic group that results in a complaint?
* Are officers that are co-accused on a complaint from the same are different demographic groups, units, age group, etc?

NLP Models

* Sentiment analysis on the available free-form text of a complaint then related to the demographics of the officers’ specific units. Do we see a more negative sentiment between complaints against an officer of a unit that is more homogenous (and different from the accuser)?