

1) How large is each police unit?

Since there are so many different units, I will not include the size of each in this findings document. While the results of this query themselves do not provide a lot of insight, the output will visualize nicely in a histogram that allows me to learn the distribution of some basic statistics about the police units. In the broader scope, seeing the range of police unit sizes may help me to further subset my future analysis to only units that are sufficiently large in order to be able to draw significant conclusions.

2) The average police units?

The average size of the active police units is 109.35 LEOs. This helps me to better understand the basic information about the units before digging in deeper.

3) Many background questions such as, for each unit, the average: age, years on force, % male vs. female, number of different races, etc.

Similar to the first question, since there are so many units, I will not include the output of this query. However, by aggregating the data and looking at average values, I am able to learn a lot of each unit. Going forward, this data will be extremely useful in order to determine how similar/homogenous each police unit is. By creating a metric of entropy for each of characteristics included in the output, I can create a proxy metric that averages the entropy of each included statistic. That information will be extremely useful in my goal of measuring any potential correlation between unit composition and rate of allegations and uses of force.

4) Number of complaints filed against members of each different unit, aggregating to average number of complaints per LEO in each unit.

Similar to the above question, aggregating this data allows me to related the information from the previous question to this data. In order to see differences in groups with different number of officers, I normalized for the number of officers and I was encouraged to see a wide range between different units, meaning that there are significant differences between units. This is encouraging because it means there is analysis that can be done to determine the underlying differences within these units and hopefully identify a correlation between some characteristic and the rate of allegations against officers within said units.

- 5) How often are members of the same unit coaccused/how many different units are identified on an allegations?

My findings for this questions were not as significant. I intended to determine how often members of the same unit were coaccused but ended up, instead, determining the number of different units included on the allegation. While not the same thing, this does allow me to identify allegations in which multiple members of the same unit were both accused (when the number of distinct units is less than the co-accused count). In the broader scope of the project, I will look to utilize this information for anecdotal evidence to support future findings.