Research Questions

(Temporal dimension) A timeline visualization of sustained allegations of most notorious officers (90th percentile) in allegations made by officers and civilians, horizontal axis displays personal trajectory while vertical axis should display lines representing multi-officer events, when vertical event lines are hovered upon, display allegation properties, such as victim info.

What we ended up with for this visualization, is a multilevel timeline of sustained allegations for the top 10 officers, in terms of their sustained allegation percentiles (Figure 1).

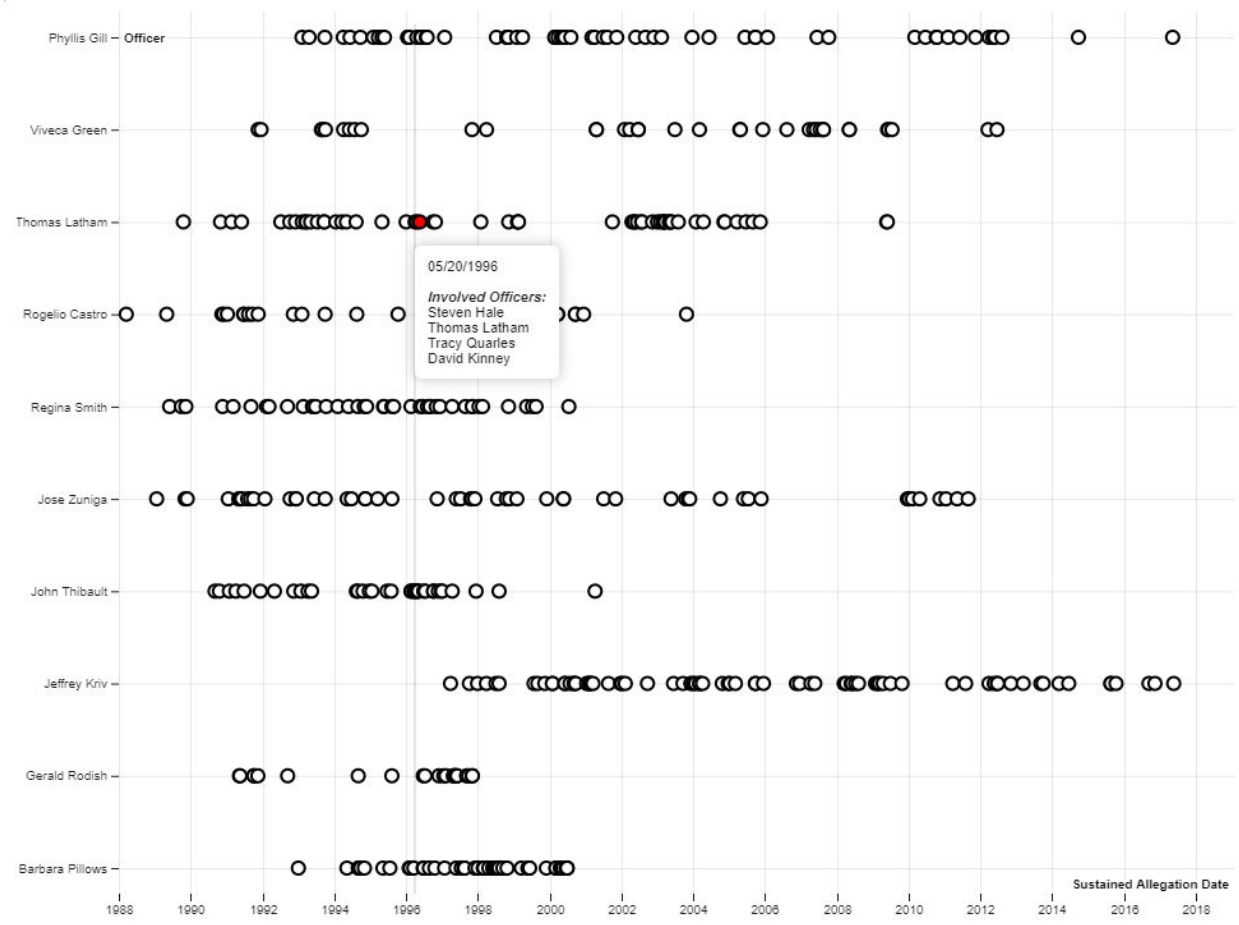

Figure 1

From the visualization, it can be seen that it is often the case that multiple officers in this top ten list often appear together on the same complaints. More generally, this visualization is also helpful for showing officers that commonly appear on the same allegations, as for any given officer above, there are usually a handful of other officers that consistently appear on the same complaints. Finally, this visualization is also useful for its temporal component. For example, Figure 1 indicates a bimodal temporal sustained allegation distribution for several officers

(Thomas Latham and Jose Zuniga). This distribution could indicate a leave of absence from their position, a change in assignment locations, or multiple “turning points” where the officer stopped their poor behavior for a period before resuming it again. Also, evident from the temporal aspect of this visualization is that, while all of the listed officers have very high levels of police misconduct (measured by sustained allegations), the temporal density of many officers is different. For example, contrast the temporal density of the sustained allegations of officer Phyllis Gill over 16 years, compared to the much higher temporal density of Barbara Pillows over half that time. It would be interesting to know if the latter officer was terminated for such a high rate of sustained allegations, and if so, whether this rate would have continued to today.

(Geographical dimension) Extends the geo symbol map in Vis-Q2, allowing users to select the percentile range using a two-sided slider.

The findings from this visualization further corroborate those found in the second static visualization from checkpoint 2, specifically that which relates to the geographical clustering of allegations by officers in the higher officer citizen allegation percentiles. In Figure 2.1, we can see that officers with allegation percentiles of 80 and below are fairly evenly distributed

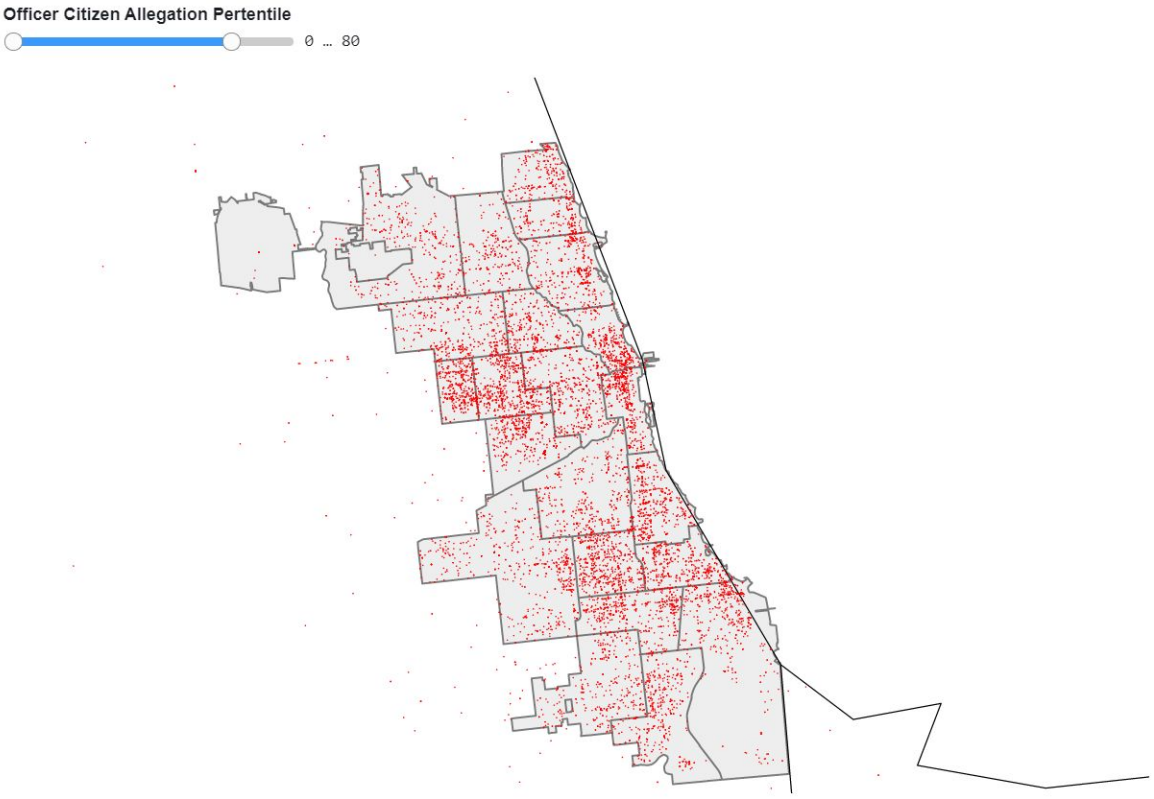

Figure 2.1

throughout Chicago, with a slight increase in allegation densities in the 1st district, likely due to the higher population density there. However, Figure 2.2 reveals higher allegation densities in the 6th, 7th, 10th, 11th, and 15th districts (district names are displayed when hovering over a given district with the mouse). Furthermore, the allegation densities in the 1st and 18th districts (middle right) appear to be fairly similar to those in Figure 2.1, confirming that this is likely just a result of high population density in this area.

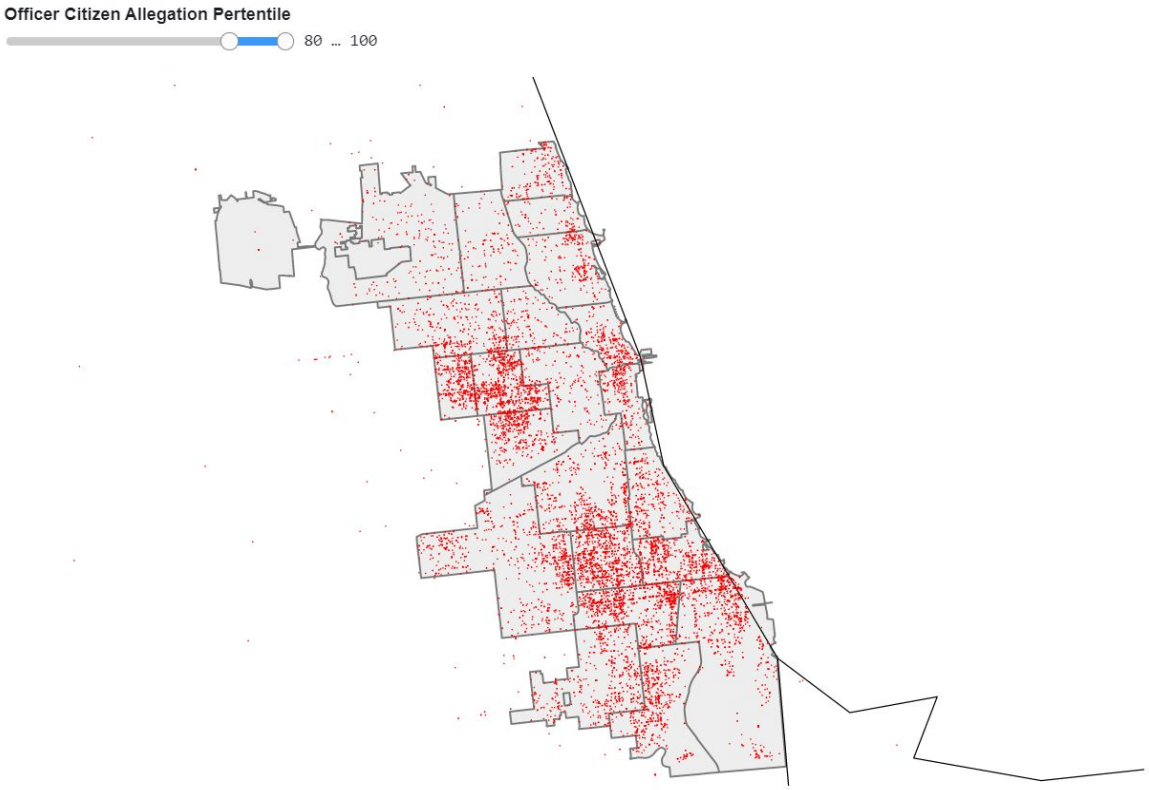

Figure 2.2

The findings from this visualization shows that the lower allegation percentile officers seem to be somewhat evenly distributed around Chicago, while the higher allegation percentile officers are grouped together geospatially, particularly in the 6th, 7th, 10th, 11th, and 15th districts. While this phenomena could simply mean that residents in these areas are for whatever reason more prone to complaining about police, it might possibly be an indicator of something deeper. For example, it might be the case that these high allegation density areas are where units of problem officer cliques are typically assigned. Under this scenario, the visualization provides insight into which officers have the greatest propensity towards misconduct.

However, beyond simply looking at one or two allegation percentile ranges, as was the extent of the functionality for the corresponding visualization in checkpoint 2, this visualization allows the user to select any range for which to view geospatial allegation data. For example, Figure 2.3 shows geospatial allegation data for officers in the medium allegation percentile range (20% -

80%). Here, the geospatial distribution uniformity of medium allegation percentile officers is even clearer.

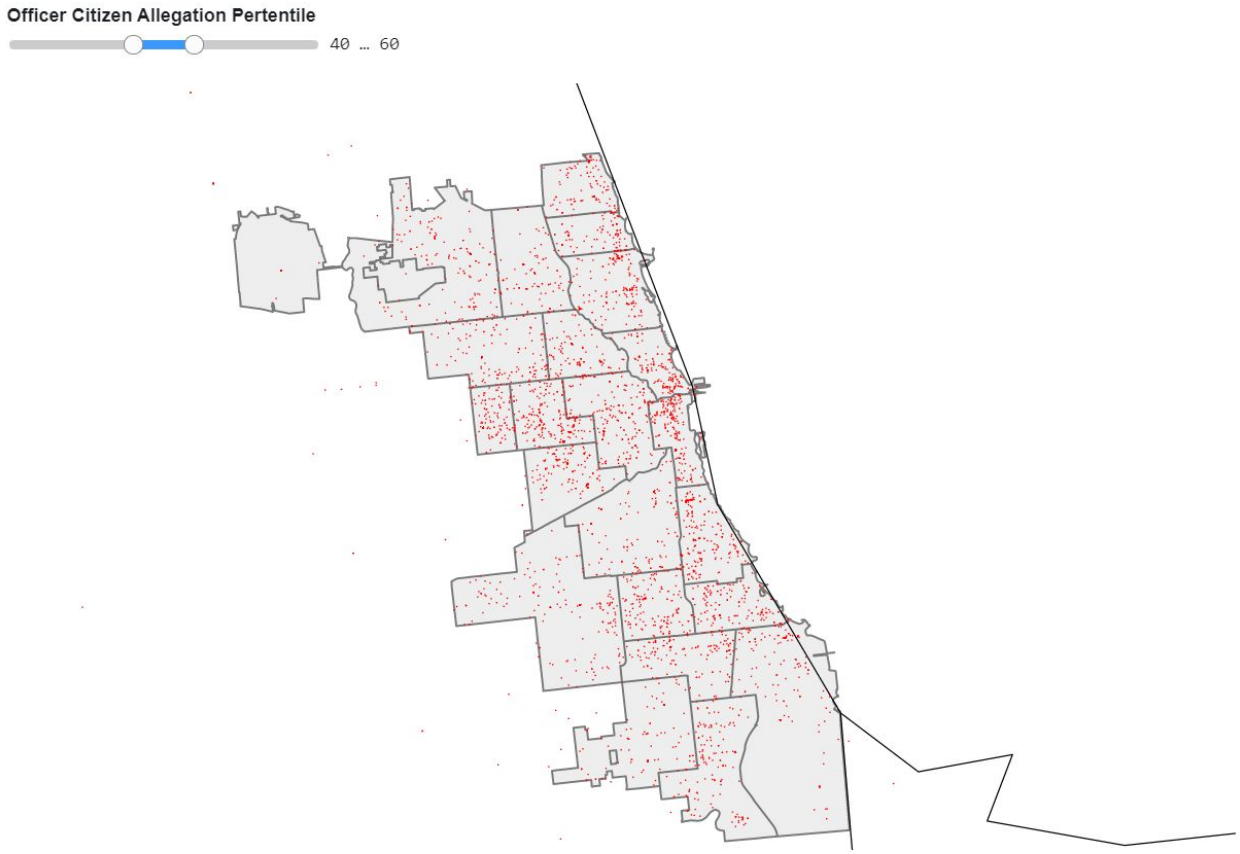

Figure 2.3

(Personal and interpersonal dimension) Using data from Vis-Q3, create a force directed graph of relation between highly correlated officers (usually occurring together in the co-accused section), highlight their correlation link (while darkens other officers), and show their most common properties (either raw data clustering such as race, etc or embedded information from topic modeling, etc).

Graph link: visualization 3

In this visualization, we are using a dataset made up of officers co-occurring equal to or more than twice in the same set of allegations with officers that are above the 99th civilian allegation percentile. The dataset contains 1597 officers, with 3109 co-accusal relations, below is a figure of the global view of this relation network.

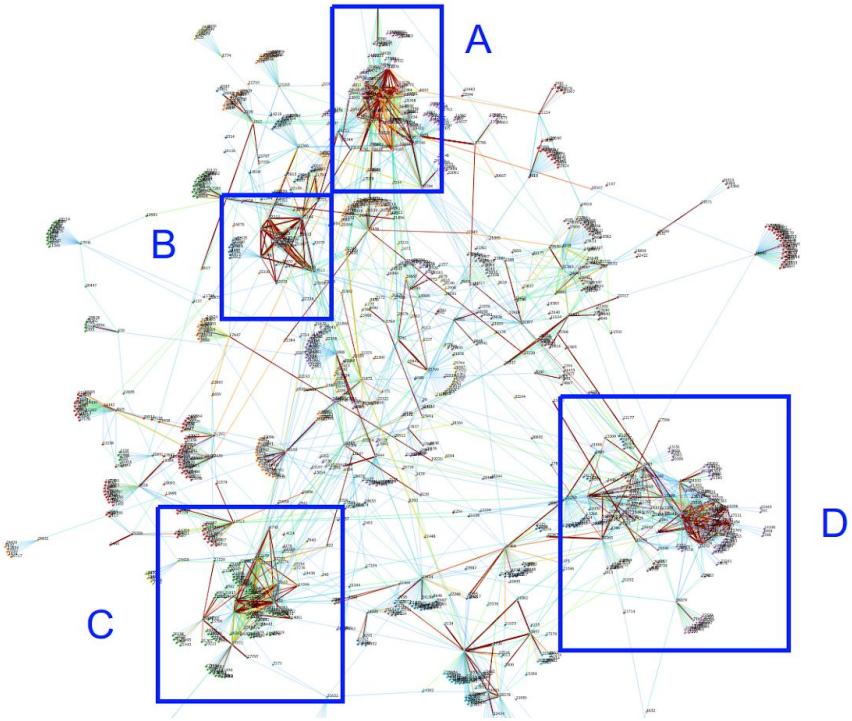

Figure 3.1

We represent officers as nodes, and co-accusal relationships as edges. The allegation count of a co-accusal relationship is represented by the thickness, length and color of the edge. Reder, shorter and thicker edges mean stronger relationships.

For better crime cluster visualization, we also apply the clustering algorithm described in "Finding community structure in very large networks." written by Aaron Clauset, M. E. J. Newman, Cristopher Moore et al. to the relation network, aiming to discover crime communities . When users hover on some node, a highlighted hull of the cluster this node belongs to will be shown, and information about this cluster will be displayed in the table placed above this graph.

Cluster info name	value
officerNum	126
gender	M: 88.89%, F: 11.11%
race	White: 53.97%, Hispanic: 29.37%, Black: 15.87%, Asian/Pacific: 0.79%
appointedYear	1990s: 52.38%, 2000s: 39.68%, 1980s: 6.35%, 1970s: 1.59%
birthYear	1940s: 1.59%, 1960s: 30.95%, 1980s: 5.56%, 1970s: 54.76%, 1950s: 7.14%
civilianAllegationPercentile	0~51.61: 5.56%, 51.61~77.42: 23.02%, 77.42~90.32: 24.60%, 90.32~96.77: 24.60%, 96.77~100: 22.22%
internalAllegationPercentile	0~51.61: 33.33%, 51.61~77.42: 33.33%, 77.42~90.32: 20.63%, 90.32~96.77: 10.32%, 96.77~100: 2.38%
salary	0~40000: 0.79%, 120000~160000: 2.38%, 160000~200000: 2.38%, 40000~80000: 5.56%, 80000~120000: 88.89%

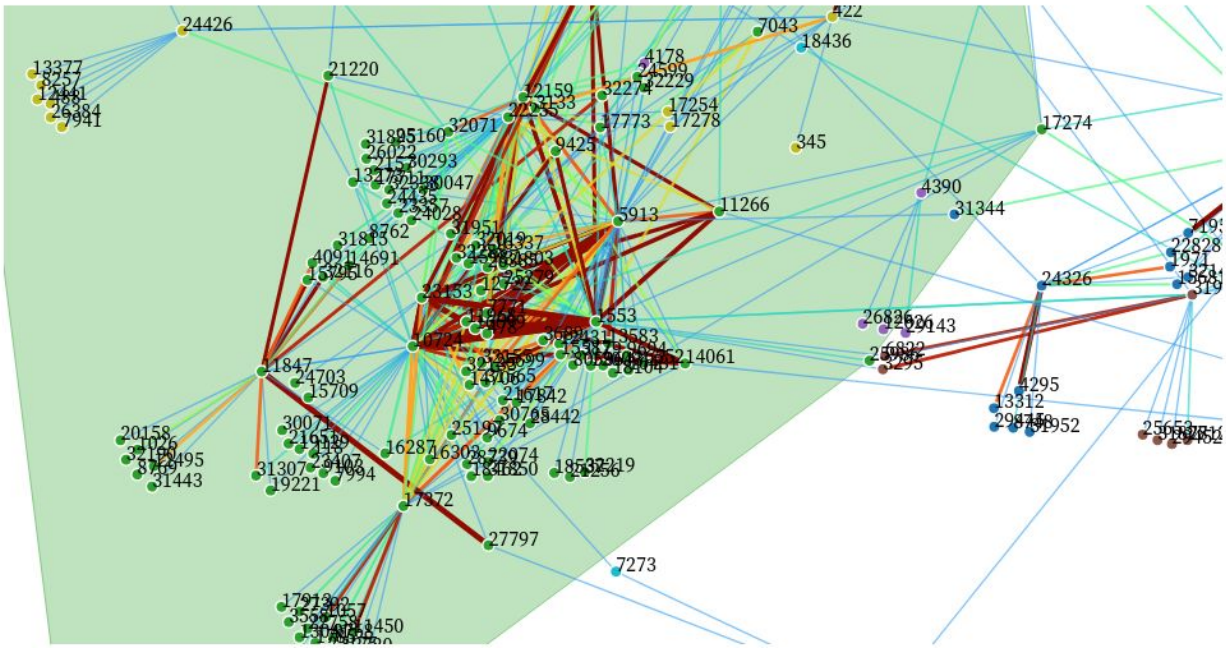

Figure 3.2

As can be seen from figure 3.1, There are mainly four obvious possible crime clusters, we label them as A, B, C and D, and list their clustered info in table 3.1.

cluster	A	B	C	D
officerNum	137	82	126	128
gender	M: 91.24%, F: 8.76%	M: 91.46%, F: 8.54%	M: 88.89%, F: 11.11%	M: 96.88%, F: 3.13%
race	Black: 41.61%,	Black: 15.85%,	Black: 15.87%,	Black: 5.47%,

	White: 42.34%, Hispanic: 13.14%, Native American/Alaskan Native: 0.73%, Asian/Pacific: 2.19%	White: 62.20%, Hispanic: 21.95%	White: 53.97%, Hispanic: 29.37%, Asian/Pacific: 0.79%	White: 79.69%, Hispanic: 14.06%, Asian/Pacific: 0.78%
appointedYear	1960s: 0.73%, 1980s: 8.76%, 1990s: 62.77%, 2000s: 25.55%, 2010s: 2.19%	1970s: 2.44%, 1980s: 10.98%, 1990s: 53.66%, 2000s: 32.93%	1970s: 1.59%, 1980s: 6.35%, 1990s: 52.38%, 2000s: 39.68%	1980s: 17.97%, 1990s: 75.78%, 2000s: 6.25%
birthYear	1930s: 0.73%, 1940s: 1.46%, 1950s: 10.22%, 1960s: 41.61%, 1970s: 32.12%, 1980s: 13.87%	1950s: 12.20%, 1960s: 37.80%, 1970s: 43.90%, 1980s: 6.10%	1940s: 1.59%, 1950s: 7.14%, 1960s: 30.95%, 1970s: 54.76%, 1980s: 5.56%	1940s: 0.78%, 1950s: 9.38%, 1960s: 48.44%, 1970s: 41.41%
civilianAllegationPercentile	0~51.61: 6.57%, 51.61~77.42: 9.49%, 77.42~90.32: 22.63%, 90.32~96.77: 29.93%, 96.77~100: 31.39%	0~51.61: 7.32%, 51.61~77.42: 10.98%, 77.42~90.32: 28.05%, 90.32~96.77: 19.51%, 96.77~100: 34.15%	0~51.61: 5.56%, 51.61~77.42: 23.02%, 77.42~90.32: 24.60%, 90.32~96.77: 24.60%, 96.77~100: 22.22%	0~51.61: 15.63%, 51.61~77.42: 12.50%, 77.42~90.32: 23.44%, 90.32~96.77: 22.66%, 96.77~100: 25.78%
internalAllegationPercentile	0~51.61: 27.74%, 51.61~77.42: 24.09%, 77.42~90.32: 32.85%, 90.32~96.77: 10.22%, 96.77~100: 5.11%	0~51.61: 30.49%, 51.61~77.42: 23.17%, 77.42~90.32: 29.27%, 90.32~96.77: 13.41%, 96.77~100: 3.66%	0~51.61: 33.33%, 51.61~77.42: 33.33%, 77.42~90.32: 20.63%, 90.32~96.77: 10.32%, 96.77~100: 2.38%	0~51.61: 39.06%, 51.61~77.42: 29.69%, 77.42~90.32: 21.88%, 90.32~96.77: 6.25%, 96.77~100: 3.13%
salary	0~40000: 2.92%, 40000~80000: 8.03%, 80000~120000: 82.48%, 120000~160000: 3.65%, 160000~200000: 2.92%	0~40000: 3.66%, 40000~80000: 8.54%, 80000~120000: 85.37%, 120000~160000: 2.44%	0~40000: 0.79%, 40000~80000: 5.56%, 80000~120000: 88.89%, 120000~160000: 2.38%, 160000~200000: 2.38%	0~40000: 3.91%, 40000~80000: 14.06%, 80000~120000: 77.34%, 120000~160000: 3.13%, 160000~200000: 1.56%

Table 3.1

We color attributes of the majority (>=70%) as red, and it reveals that for all clusters, the majority of officers share very similar attributes: for example, they are mainly male officers, appointed in 1990s or 2000s, and were born in 1960s or 1970s. However, while these clusters have many attributes in common, the race composition of each cluster shows that each cluster also has their own specific characteristics. For instance, cluster A is mainly made up of white and black officers and they are close in number, while cluster B and C are mainly formed of white and hispanic officers, with white officers as the majority. Cluster D is mainly composed of white officers.

Base on this visualization, we have made the following conclusions:

1. There do exist crime clusters in the police community, with a compact center mass. And several nodes connecting this center mass to periphery officers with less allegations.
2. Crime clusters connect to each other with a bridge formed by 1 ~ 2 officers (gateway officers), such as officers with ID 32366, 14721 and 32107 connecting cluster A and B. The bridge is usually connected to many officers in each cluster, with a small count of common allegations (< 10).
3. Crime clusters share many common characteristics, like salary, age, appointed date, etc, but they also do not share specific attributes, mainly race composition.
4. Officers at a high civilian allegation percentile are usually at a low internal allegation percentile.