



Data Science Seminar - MSAI 339

Checkpoint 1

October 15th, 2021

Professor

Jennie Rogers

Students

Anery Patel

Chukwueloka Obi

Ana Cheyr

Research questions

Relational Analytics

- **What is the correlation factor between the number of years in the force of an officer and the number of unsustained allegations that a police officer has?**

We see that the correlation factor is 0.56 meaning that the number of years in the force of an officer and the number of unsustained allegations that a police officer has is moderately correlated. This is moderately correlated because there are quite a few police officers that have had a long file of allegations filed against them throughout their career on the police force but there also many officers who have as many years in the force as the officers with multiple allegations and have very low amount or almost no allegations. So it really depends on who the officer is.

- **What is the average number of times a frequently accused police officer changes from one department unit to another?**

For the highest quartile of the number of unsustained allegations that an officer has, the 75th percentile, we see that on average an officer has changed from one unit to another around 6 times.

percentile_cont	avg
1	5.7791810179869881
16	

On the other hand, for the third quartile (between the 75th and 50th percentile), we see that on average an officer has changed from one unit to another 3.5 times, almost half of the highest quartile does.

percentile_cont	avg
1	3.4728434504792332
6	

This shows that the officers who have more unsustained allegations also change frequently from one district to another, which can lead to widespread misconduct.

- **Which are the top 5 neighborhoods with the most number of civilian complaints that have been unsustained? What percentage of unsustained complaints have each one from the total?**

We notice that there are some neighborhoods with a high percentage of unsustained complaints. Ideally speaking, for 99 areas, the rate is 1.01%. This gives us a direction to probe these areas in multiple ways. We can look into the functioning of the police units active there, the group of officers serving the areas and also the civilian nature: race, income, age, etc. We can identify factors that prevent the civilian complaints from being sustained.

- **What is the correlation between the number of unsustained allegations against a police officer and the number of awards that officer has?**

	corr
1	0.9845691291804248

We notice a strong correlation between the number of unsustained allegations against an officer and the number of awards he/she has. This makes sense because an officer is less likely to be blamed or disciplined by their fellow officers for misconduct if he/she has received many accolades during their career. It is an interesting yet quite simple takeaway. The system needs to have provisions in place that does not give a benefit of doubt to officers with awards because they are the ones to misuse it in the first place.

- **What is the average percentage of civilian and officer unsustained allegations an officer has?**

We can see that most of the allegations that an official has correspond to civilian unsustained allegations, with an average of 87.8%, so it is important to understand if there is any fraction of it that has been erroneously unsustained.

	percent_unsustained_civilian	percent_unsustained_officer
1	0.8787571860938076	0.033208160520121065