

# Checkpoint 1 Report

Masum Patel, Sizheng Zhang, Jiawei Zhang

## Introduction

The questions in our first checkpoints aim to provide us with insights into potential attribute candidates that can serve as the commonality between the top 4000 officers who received the most allegations (“repeaters”). We used SQL queries to calculate and compare the values of attributes between the “repeaters” and the rest of the officers who received at least one allegation against them (“offenders”).

The original theme stated that we would analyze the top 5% of officers who received most allegations, but the top 5% converts to about 1000 officers, which is a relatively small sample size compared to over 23,000 officers with allegation records. Therefore, we changed the theme to analyzing the top 4000 officers with the most complaints, and they accounted for 50% of all complaints filed against officers. We believe the top 4000 officers are more representative of the officers who repeatedly receive complaints and serve as a good sample relative to the 23,000 total alleged officers.

The results of checkpoint one provided us with valuable understanding into what kind of data is available for us and the role each attribute plays in distinguishing the “repeaters” from “offenders”.

## Results of Research Questions

1. The average number of allegations per officer among the “repeaters” and the “offenders”

Repeaters	Offenders
31.07 Allegations/Officer	10.65 Allegations/Officer

2. The average salary per officer among the “repeaters” and the “offenders”

Repeaters	Offenders
\$94392.63/Officer	\$89619.11/Officer

3. The average amount awards received by “repeaters”

Repeaters
47.82 awards/Officer

4. The region/area where “repeaters” were deployed the most to

Mostly deployed regions for repeaters
7th police-districts, Englewood

5. Maximum complaints from repeaters/beat = 971  
Average complaints from repeaters/beat = 226  
(Does not include beats where complaints against any repeater were not registered)

Maximum co-accused repeaters in a complaint = 53

Average co-accused repeaters in a complaint = 3

*\*co-accused are officers who have their names together in a complaint report.*

## Significance of Results

By just analyzing the average number of allegations that the top 4000 of most alleged officers, we saw a huge gap between the number of the “repeaters” and the “offenders”, with “repeaters” having average allegations almost three times larger than that of “offenders”. The result emphasized how influential the top 4000 of the most alleged officers are. Surprisingly, while the “repeaters” received so much more allegations than the average officers in the force, their average salary is 6.8% higher than the salaries of the “offenders”. The salary is an attribute worth investigating later on in the project because we expected the “repeaters” would receive lower salary. If higher salary is indeed common among “repeaters”, then the legitimacy of the high pays of these officers becomes questionable.

Another interesting trend that we see from the data is that despite the number of allegations that “repeaters” might have against themselves, they are considered as “good” police officers, as shown by the average number of awards they receive in their police careers. A “repeater” receives roughly 48 awards throughout his or her career, which is a very high number. This might indicate either that police violence is considered to be acceptable or even encouraged among Chicago police or that only more experienced and revered officers are dared to use violence. The true reason requires further investigation of the dataset.

Moreover, by analyzing the region where most “repeaters” receive allegations, we discovered that 7th police-districts, Englewood is the place where most allegations happened. The reason for the high allegation rate indicates a high rate of using violence and local distrust towards police forces, which might be correlated to the high local crime rate.

As part of investigating commonalities between repeaters, we investigated the beats assigned to repeaters versus the complaints they get. We see that on average a beat has 226 complaints

from repeaters, and the maximum is 971. This is indicative of two things. Either repeaters are often placed in beats with high crime activity or repeaters are more prone to get a complaint if they work with each other. The other possibility can be further investigated to see if beats with higher complaints also have higher number of coaccused repeaters registered.

### **Future Investigations**

The results from our initial analysis posted many interesting questions, such as why does “repeaters” have higher average salaries than offenders and why do “repeaters” receive so many awards throughout their police careers. Now we have a much better understanding of the data and the patterns of attributes of “repeaters”, we can further investigate the reasons behind such patterns or develop more concrete relationships in these patterns in future check points.