Checkpoint 1

Aristana Scourtas, (aristanascourtas2021@u.northwestern.edu), KJ Schmidt (kjschmidt2020@u.northwestern.edu), Zane Denmon (zanedenmon2020@u.northwestern.edu)

When creating and executing queries to answer our questions, we had quite a few issues. To figure out who the supervisor is of each unit, Rajiv suggested on Piazza to look at the highest paid officer of the unit. Upon looking at the outcomes of this query, we saw many officers in the same police unit had the same salary. Furthermore, we couldn't successfully link identified max salaries to an officer ID using the SQL query. This prevented us from being able to answer our Questions 2 and 3.

We also found an officer can be linked to multiple unit_ids. We were able to see that we could separate their associations with different unit_ids over time, and realized that this may be useful for future queries, but may make our query for police units with average complaint_percentiles above the 75th percentile not as precise as we would like. That said, we were able to successfully find the police units with average complaint percentiles above the 75th percentile.

Allegation sustainment was difficult to determine, so we're using the "disciplined" column instead. There were issues with calculating the percentage disciplined, so we will continue to work on that query. There was no allegation_count column as specified in the schema documentation. We had to count allegations ourselves.

More of a note than an issue, we saw when looking at units that unit 197, the "Inspector General Detail Unit", is comprised of only one person. We will make sure we discard this unit for our analysis. This suggests we should look for more units like this -- we should expand upon our questions by investigating how many officers are in each of these units.

Overall, we learned a lot more about the database and how we can search it. This information will further develop our questions and queries for this project; our questions definitely need to change based on what we were able to query. We found that we need clarification on how to identify a supervisor of a unit.