Questions:

1. What are the most common allegation categories for the allegations with arrests associated with them. How has this trend changed since the dataset has gone public?
2. What are the normalized settlement amount per rank. To check whether settlement amount has gone down, with the rank i.e to see if more amount was spent in settlements for higher ranked officers. For cases with officers with different ranks, the highest ranked officer in that settlement is responsible for that settlement. Has this trend changed since the data has gone public?
3. Average settlement amount in cases where the subject (victim) was armed. If a subject was armed, there should be less chances of settlement or lesser settlement amount offered as this could be a possible explanation for the officer using force of weapon.
4. What is the average duration of the cases per rank of officers. Are the cases involving higher rank officer settled comparatively faster? Have the average numbers been affected by the data going public?

To begin with, make sure that case\_map is properly populated and settlement\_officer\_map.sql is executed. The instructions for setting up these two relations can be found here: <https://piazza.com/class/k0y01tc35wg2oi?cid=198>

**The queries were executed using pgAdmin and not psql so we recommend that the .sql files are loaded and executed through pgAdmin to avoid any issues.**

To execute the files:

Open the query tool by clicking on the schema (cpdb) – Then go to tools – Query tools

Load the .sql files question wise into the query tool of pgAdmin. **The order of questions is important as we have created temp tables and used them in later questions.**

There are multiple select statements in each .sql file. Only the final select will be printed. So to check all the results provided in the findings.pdf, we recommend that those select statements should be executed manually.