Checkpoint 1

Getting Started

Start running the CPDP database, then run our queries from either copying and pasting from here, or opening the sql files that correspond to the question. **Our queries need to be run in order**, as some build off of one another.

Our Questions

- Which police units have average complaint_percentiles above the 75th percentile?
- What supervisors have complaint_percentiles above 75?
- Which commanders have complaint_percentiles above the 75th percentile?
- How do unit sizes relate to complaint_percentiles of units and their commanding officers?

Queries

Which police units have average complaint_percentiles above the 75th percentile?

Run question1.sql or copy and paste the queries below

```
SELECT data_officerhistory.unit_id,
data_policeunit.unit_name, data_policeunit.description,
data_officerhistory.officer_id
```

```
INTO units
FROM data_policeunit
INNER JOIN data_officerhistory on data_policeunit.id =
data_officerhistory.unit_id
WHERE active = TRUE;
SELECT units.unit_id, units.unit_name, units.description,
units.officer_id, data_officer.complaint_percentile
INTO officers_and_units
FROM data_officer
INNER JOIN units on officer_id = data_officer.id;
SELECT unit_id, unit_name, description,
AVG(complaint_percentile)
FROM officers_and_units
GROUP BY unit_id, unit_name, description
HAVING AVG(complaint_percentile) > 75;
```

What supervisors have complaint_percentiles above 75?

Run question2.sql or copy and paste the queries below

```
CREATE TEMP TABLE supervisors_and_complaint_percentiles
AS (
    SELECT officer_id, unit_id, unit_name,
complaint_percentile
    FROM supervisors s, data_officer d
    WHERE s.officer_id = d.id
);
SELECT officer_id, unit_id, unit_name,
complaint_percentile
FROM supervisors_and_complaint_percentiles s
GROUP BY s.officer_id, s.unit_id, s.unit_name,
s.complaint_percentile
```

```
HAVING s.complaint_percentile > 75;
```

What are the complaint_percentiles for the supervisors of the units from question 1?

Run question3.sql or copy and paste the queries below

```
DROP TABLE salary_analysis, highest_salary, commanders,
supervisors;
CREATE TEMP TABLE salary_analysis AS (SELECT o.id
officer_id, pu.id unit_id, pu.unit_name, salary
FROM data_officer o, data_officerhistory oh,
data_policeunit pu, data_salary s
WHERE o.id = oh.officer_id AND oh.unit_id = pu.id AND
s.officer_id = o.id
AND oh.end_date IS NULL AND pu.active AND
o.resignation_date IS NULL);
CREATE TEMP TABLE highest_salary AS (
        SELECT unit_id, max(salary) as max_unit_salary
        FROM salary_analysis
        GROUP BY unit_id);
CREATE TEMP TABLE commanders AS (SELECT officer_id,
unit_id, unit_name
FROM salary_analysis s
WHERE salary = (SELECT max_unit_salary FROM
highest_salary hs WHERE hs.unit_id = s.unit_id));
SELECT DISTINCT ON (unit_id) unit_id, unit_name,
officer_id
INTO supervisors
FROM commanders;
DROP TABLE supervisors_and_complaint_percentiles;
```

```
CREATE TEMP TABLE supervisors_and_complaint_percentiles
AS (
    SELECT officer_id, unit_id, unit_name,
complaint_percentile
    FROM supervisors s, data_officer d
    WHERE s.officer_id = d.id
);
SELECT s.officer_id, s.unit_id, s.unit_name,
s.complaint_percentile
FROM supervisors_and_complaint_percentiles s,
most_problematic_units m
WHERE s.unit_id = m.unit_id
```

How do unit sizes relate to complaint_percentiles of units and their commanding officers?

Run question4.sql or copy and paste the queries below

```
DROP TABLE salary_analysis, highest_salary,commanders;
CREATE TABLE salary_analysis AS (SELECT o.id officer_id,
o.complaint_percentile complaint_percentile, pu.id
unit_id, pu.unit_name, pu.description, salary
FROM data_officer o, data_officerhistory oh,
data_policeunit pu, data_salary s
WHERE o.id = oh.officer_id AND oh.unit_id = pu.id AND
s.officer_id = o.id
AND oh.end_date IS NULL AND pu.active AND
o.resignation_date IS NULL);
CREATE TABLE highest_salary AS (
        SELECT unit_id, max(salary) as max_unit_salary
        FROM salary_analysis
        GROUP BY unit_id);
```

```sql
CREATE TABLE commanders AS (SELECT officer_id, unit_id,
unit_name, description, complaint_percentile
FROM salary_analysis s
WHERE salary = (SELECT max_unit_salary FROM
highest_salary hs WHERE hs.unit_id = s.unit_id));
SELECT commanders.unit_id, COUNT(d.officer_id) AS
number_of_officers, unit_name, description,
commanders.complaint_percentile AS
supervisor_complaint_percentile,
        AVG(o.complaint_percentile) as
avg_unit_complaint_percentile
FROM commanders
INNER JOIN data_officerhistory d on commanders.unit_id =
d.unit_id
INNER  JOIN data_officer o on d.officer_id = o.id
WHERE active = 'Yes'
GROUP BY commanders.unit_id, commanders.unit_name,
description, commanders.complaint_percentile
ORDER BY COUNT(d.officer_id) DESC;
```