Checkpoint 3

Getting Started

To run our queries, you'll need to load the settlements data (in addition to the CPDB which you've already done by this point). To do this, follow the instructions on canvas.

And then run this (via piazza):

```
DROP TABLE IF EXISTS case_map;

CREATE TABLE case_map AS (
SELECT c.id case_id, a.id allegation_id
FROM data_allegation a, cases_ipracase c
WHERE TRIM(LEADING 'C' FROM crid)::bigint = cr_no
);

ALTER TABLE case_map ADD CONSTRAINT  ipra_fkey FOREIGN
KEY(case_id) REFERENCES cases_ipracase(id);
ALTER TABLE case_map ADD CONSTRAINT  allegation_fkey FOREIGN
KEY(allegation_id) REFERENCES data_allegation(id);
```

Setup for our specific queries:

Also found in source/setup.sql

Part 1:

```
DROP TABLE IF EXISTS salary_analysis, highest_salary, commanders
CASCADE;

-- get salary and unit info
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);
```

```
-- find the max salary per unit
CREATE TABLE highest_salary AS (
       SELECT unit_id, max(salary) as max_unit_salary
       FROM salary_analysis
       GROUP BY unit_id);

-- get the officer with the max salary, who is probably the
commander
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));

DROP TABLE IF EXISTS supervisors;

SELECT DISTINCT ON (unit_id) unit_id, unit_name, description,
officer_id
INTO supervisors
FROM commanders;

SELECT *
FROM supervisors;

DROP TABLE IF EXISTS supervisors_and_complaint_percentiles;

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

Part 2:

```
DROP VIEW IF EXISTS allegation_officer_mapping;

CREATE VIEW allegation_officer_mapping AS
(
```

```
   SELECT DISTINCT oa.allegation_id, oa.officer_id, cm.case_id,
c.unit_id, c.description
   FROM cpdb.public.data_officerallegation oa,
cpdb.public.data_allegation a, case_map cm, commanders c
   WHERE oa.officer_id = c.officer_id AND
         oa.allegation_id = a.id AND
         a.id = cm.allegation_id
 );


SELECT * FROM allegation_officer_mapping;
```

Questions and Queries

What percentage of law enforcement supervisors who are named in settlements were above the 75th complaint percentile?

Query also found in source/q1.sql

```
SELECT s.officer_id, s.unit_id, s.unit_name, s.description,
s.complaint_percentile
FROM supervisors_and_complaint_percentiles s
WHERE s.officer_id IN (SELECT officer_id FROM
allegation_commander_mapping)
GROUP BY s.officer_id, s.unit_id, s.unit_name,
s.complaint_percentile, s.description
HAVING s.complaint_percentile > 75
ORDER BY s.complaint_percentile DESC;
```

Which supervisors cost the department the most money in settlements?

Query found in source/q2.sql

```
DROP VIEW IF EXISTS supervisors_unit_settlement;

CREATE VIEW supervisors_unit_settlement AS (
SELECT aom.officer_id as supervisor_id, aom.unit_id,
aom.description, SUM(cp.payment + cp.fees_costs) total_cost
```

```
FROM allegation_commander_mapping aom, cases_payment cp,
cops_ipracop ci, cops_cop cc, cops_casecop ccc, cases_case
casecase
WHERE aom.case_id = ci.case_id AND
      ci.cop_id = cc.id AND
      cc.id = ccc.cop_id AND
      ccc.case_id = casecase.id AND
      casecase.id = cp.case_id
GROUP BY aom.officer_id, aom.unit_id, aom.description
ORDER BY total_cost DESC);

SELECT * FROM supervisors_unit_settlement;
```

Comparing the settlement costs of law enforcement supervisor vs. the average of their subordinates, is this value greater or lesser?

Query found in source/q3.sql

```
DROP VIEW IF EXISTS officers_unit_settlement;

CREATE VIEW officers_unit_settlement AS (
SELECT aom.unit_id, aom.description, SUM(cp.payment +
cp.fees_costs) as total_cost, AVG(cp.payment + cp.fees_costs)  as
avg_cost, COUNT(aom.unit_id) as total_settlements
FROM allegation_officer_mapping aom, cases_payment cp,
cops_ipracop ci, cops_cop cc, cops_casecop ccc, cases_case
casecase
WHERE aom.case_id = ci.case_id AND
      ci.cop_id = cc.id AND
      cc.id = ccc.cop_id AND
      ccc.case_id = casecase.id AND
      casecase.id = cp.case_id
GROUP BY aom.unit_id, aom.description
ORDER BY avg_cost DESC);

SELECT * FROM officers_unit_settlement;
```

Comparing the complaint percentile to the average of their subordinates, is this value greater, or lesser?

Query found in source/q4.sql

```
SELECT sus.unit_id, sus.description, sacp.officer_id as
supervisor_id, sacp.complaint_percentile as
supervisor_complaint_percentile, AVG(o.complaint_percentile) as
avg_unit_complaint_percentile
FROM supervisors_unit_settlement sus,
supervisors_and_complaint_percentiles sacp, officers_and_units
oau, cpdb.public.data_officer o
WHERE oau.officer_id = o.id AND
      oau.unit_id = sus.unit_id
GROUP BY sus.unit_id, sus.description, sacp.officer_id,
sacp.complaint_percentile;
```