The theme of the project is to learn about the general behavior of the officers who are charged with misconduct and how different factors may affect their behavior. For example, does the rank of officer, department of officer have any effect? Also, we want to analyze the incidents in more depth. For example, what time of the day do incidents usually occur, or what is the average age of victims, how many of the incidents were done by off-duty officers?

In this paper, we have analysed the results of the questions and presented interesting findings of the same. We have also suggested some future lines of investigation.

CP1- Relational Analytics:

1. **Officers of which rank had most complaints against them?**

	normalized_count numeric	rank character varying (100)
1	94.0000000000000000	Director Of Caps
2	22.0000000000000000	Assistant Superintendent
3	11.4164222873900293	Field Training Officer
4	10.4117647058823529	Commander
5	10.3699570815450644	Sergeant
6	9.7586206896551724	Chief
7	8.9239583333333333	Lieutenant
8	8.7868852459016393	Deputy Chief
9	8.3750000000000000	First Deputy Superintendent
10	8.2422562764916857	Detective
11	7.2173212575111034	Police Officer
12	5.1345565749235474	Captain
13	3.5000000000000000	Superintendent Of Police
14	2.6086956521739130	Deputy Superintendent
15	1.5200000000000000	Assistant Deputy Superintendent

Here, the problem that we are investigating is whether the rank of the officer is having any impact on the number of complaints getting generated against them.
We have normalized the number of complaints to get a better estimate. That is, we have divided the number of complaints by the number of officers of that particular rank. For example, higher ranked officers would be few but even few complaints amongst higher ranked officers needs to be highlighted.

We can see the number of officers in each rank using the following query:

```
select count(d_o_2.id) as count_2, d_o_2.rank from data_officer as d_o_2
        group by d_o_2.rank
```

	count_2 bigint	rank character varying (100)
1	1	Director Of Caps
2	1	Assistant Superintendent
3	3	Superintendent'S Chief Of Staff
4	6	Superintendent Of Police
5	8	First Deputy Superintendent
6	23	Deputy Superintendent
7	24	Other
8	25	Assistant Deputy Superintendent
9	29	Chief
10	61	Deputy Chief
11	136	Commander
12	327	Captain
13	341	Field Training Officer
14	960	Lieutenant
15	2198	
16	3067	Detective
17	3495	Sergeant
18	22966	Police Officer

The intuition here is that, higher ranks are positions of higher responsibility and hence, the complaints against such officers would be comparatively less. However, from the results we can see that higher ranked officers are having significant number of complaints against them. An interesting observation is that even though there is only one Director of Caps and Assistant Superintendent they still have a significant number of complaints against them.

2. **Analyzing whether officers who are charged with misconduct use weapons or beat up victims bare handed.**

	count bigint	w_type text
1	1030	Gun
2	4399	Taser
3	3754	Chemical Weapon
4	14	Hand

Another problem that we're looking into is to find out whether officers who are charged with misconduct use weapons or beat up victims bare handed. Here, we have determined the count for each weapon used. Here, we can see that the maximum count is for Taser, followed by chemical weapon, Gun and Hand. Note: We have combined the use of hands and baton under hands category for simplicity.

As we can see, guns were used in 1030 cases which is still significant. Weapons like Taser are considered "less lethal" along with Chemical Weapons (like tear gas) and hence they have a high number of usage. The usage of bare hands/baton is very limited.

Further we can analyze the cases in which guns were used by seeing how many shots were fired, the area in which maximum gun incidents occur. Another interesting thing to see would be in what kind of scenarios were Tasers and Chemical Weapons were used and even though they are less lethal, was their use justified.

3. **How many of the incidents were carried out by off duty officers?**

	count bigint	officer_on_duty boolean
1	1600	false
2	65419	true

We have used the TRR data to find the results. It required the analysis of the working hours of the officer and the time at which the incident took place.Interestingly, only a limited number of allegations were due to off-duty officers. However, the next step would be to look at the severity of these allegations. What we want to check is whether off-duty officers are more prone to taking drastic steps like using guns.

4. **Impact of inspecting officer's race on the result of the investigation.**
   If an officer with allegation and the investigating officer has the same race, does it affect the outcome of the investigation?

	count bigint	final_outcome character varying (32)
1	543	1 Day Suspension
2	62	10 Day Suspension
3	4	12 Day Suspension
4	1	13 Day Suspension
5	1	14 Day Suspension
6	51	15 Day Suspension
7	2	16 Day Suspension
8	198	2 Day Suspension
9	22	20 Day Suspension
10	1	22 Day Suspension
11	1	23 Day Suspension
12	14	25 Day Suspension
13	1	27 Day Suspension
14	166	3 Day Suspension
15	56	30 Day Suspension
16	1	326 Day Suspension
17	1	35 Day Suspension
18	3	365 Day Suspension
19	17	4 Day Suspension
20	4	45 Day Suspension
21	133	5 Day Suspension
22	4	6 Day Suspension
23	2	60 Day Suspension
24	14	7 Day Suspension
25	1	8 Day Suspension
26	3	90 Day Suspension
27	5	99 Day Suspension
28	19	Administrative Termination
29	37292	No Action Taken
30	98	Penalty Not Served
31	1	Reinstated By Court Action
32	8	Reinstated By Police Board
33	831	Reprimand
34	233	Resigned
35	31	Separation
36	4	Suspended For 180 Days
37	11	Suspended Over 30 Days
38	2610	Unknown
39	274	Violation Noted

	count bigint	final_outcome character varying (32)
1	736	1 Day Suspension
2	87	10 Day Suspension
3	4	12 Day Suspension
4	2	13 Day Suspension
5	1	14 Day Suspension
6	72	15 Day Suspension
7	4	16 Day Suspension
8	1	17 Day Suspension
9	307	2 Day Suspension
10	39	20 Day Suspension
11	19	25 Day Suspension
12	2	28 Day Suspension
13	214	3 Day Suspension
14	69	30 Day Suspension
15	4	365 Day Suspension
16	29	4 Day Suspension
17	1	40 Day Suspension
18	8	45 Day Suspension
19	199	5 Day Suspension
20	12	6 Day Suspension
21	5	60 Day Suspension
22	16	7 Day Suspension
23	3	8 Day Suspension
24	1	9 Day Suspension
25	1	90 Day Suspension

26	4	99 Day Suspension
27	37	Administrative Termination
28	42963	No Action Taken
29	124	Penalty Not Served
30	2	Reinstated By Court Action
31	10	Reinstated By Police Board
32	1077	Reprimand
33	304	Resigned
34	55	Separation
35	2	Suspended For 180 Days
36	27	Suspended Over 30 Days
37	2916	Unknown
38	401	Violation Noted

Same Race:

```
select count(tb2.final_outcome), tb2.final_outcome
from
(
        select d_i_a.allegation_id as allegation_id, d_o.race
        from data_officer as d_o,
        data_investigator as d_i,
        data_investigatorallegation as d_i_a
        where d_i.officer_id = d_o.id and d_i.id = d_i_a.investigator_id
) as tb1,
(
        select d_o2.race as race, d_o_a2.allegation_id as allegation_id,
d_o_a2.final_outcome as final_outcome from
        data_officer as d_o2,
        data_officerallegation as d_o_a2
        where d_o2.id = d_o_a2.officer_id
) as tb2
where tb1.allegation_id = tb2.allegation_id and tb1.race == tb2.race
group by tb2.final_outcome
```

Different race:

```
select count(tb2.final_outcome), tb2.final_outcome
from
(
        select d_i_a.allegation_id as allegation_id, d_o.race
```

```
        from data_officer as d_o,
        data_investigator as d_i,
        data_investigatorallegation as d_i_a
        where d_i.officer_id = d_o.id and d_i.id = d_i_a.investigator_id
) as tb1,
(
        select d_o2.race as race, d_o_a2.allegation_id as allegation_id,
d_o_a2.final_outcome as final_outcome from
        data_officer as d_o2,
        data_officerallegation as d_o_a2
        where d_o2.id = d_o_a2.officer_id
) as tb2
where tb1.allegation_id = tb2.allegation_id and tb1.race != tb2.race
group by tb2.final_outcome
```

As we can see in the above two results (Investigating officer's race same as allegated officer and not the same as allegated officer respectively), the number of no action taken were 37292 for same and 42963 for different races. So it seems the race is not an important factor in this.

Normalized result:
For same race - 37292/42723 = 0.872
For different race - 42963/49756 = 0.863