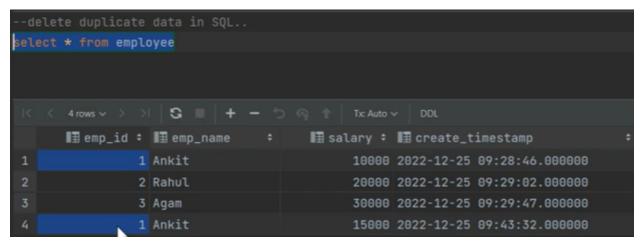
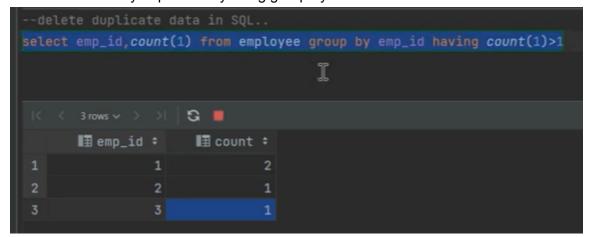
Duplicates

1.



- 2. Here we have duplicate of emp_id = 1
- 3. Also we can identify duplicates by using group by and count



4. One way of removing duplicates is to make use of timestamp columns...and if we want to keep latest timestamp id..then we use max(time stamp) ..or if we want to store the first entry then we use min(time stamp)

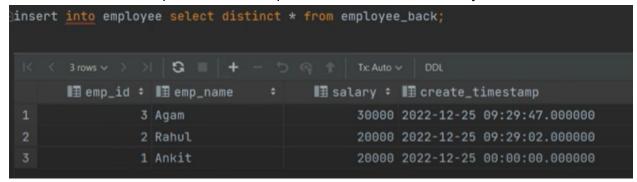
```
select emp_id,min(create_timestamp) as create_timestamp from employee group by emp_id
```

5. Now we can use delete command to delete the unwanted records

```
delete from employee where (emp_id,create_timestamp) in (select emp_id,min(create_timestamp) as create_timestamp from employee group by ema_
```

- 6. Also in realtime we'll always have the create timestamp column
- 7. What if we have exactly the same duplicates?

8. Here to eliminate the duplicates in the output...we have used distinct keyword



9.