Quiz 4 Answers, Fall 2019
CSCI 4380 DB Sys

Question (12+13 points). You are given the following data model for bird watchers and birds (keys are underlined):

`BirdWatchers(bwid, name, homestate, education, email, password)`
`Birds(birdname, scientificName, family, genus, ismigratory)`
`Habitat(birdname, state, howcommon)`
`Observations(oid, birdname, quantity, behavior, odate, otime, latitute, longitude, city, state, bwid, prev_oid)`

Birds have scientific names, family genus and the migratory status (`ismigratory` values are True/False). Birds habitats are `state`s that they are native in with a `howcommon` percentage value. Observations are by a bird watcher (`bwid`) and describe which bird was seen (`birdname`), where they were observed (`latitute, longitude, city, state`), when (`odate, otime`), in which `quantity` and the `behavior` of the birds (such as healthy, aggressive, indistress).

Write the following queries using SQL, using simplest possible expressions:

(a) Return name of birds with the highest number of observations in year `2019`, city `Troy` and state `NY`.

```
select
   o1.birdname
   , count(*) numobservations
from
   observations o1
where
   extract(year from o1.odate) = 2019
   and o1.city = 'Troy'
   and o1.state = 'NY'
group by
   o1.birdname
having
   count(*) >= ALL (select count(*) from observations o2
                    where extract(year from o2.odate) = 2019
                    and o2.city = 'Troy' and o2.state = 'NY'
    group by o2.birdname)
```

ALTERNATE SOLUTION 1:

```
with counts as (
    select count(*) as numo
    from observations o2
    where extract(year from o2.odate) = 2019
                and o2.city = 'Troy' and o2.state = 'NY'
    group by o2.birdname
)
select
   o1.birdname
   , count(*) numobservations
from
   observations o1
where
   extract(year from o1.odate) = 2019
   and o1.city = 'Troy'
   and o1.state = 'NY'
group by
   o1.birdname
having
   count(*) = (select max(numo) from counts);
```

ALTERNATE SOLUTION 2:

```
with counts as (
    select count(*) as numo
    from observations o2
    where extract(year from o2.odate) = 2019
                and o2.city = 'Troy' and o2.state = 'NY'
    group by o2.birdname
)
select
    c1.birdname
    , c1.numo
from
    counts c1
    left join counts c2
    on c1.numo < c2.numo
where
    c2.birdname is null
```

(b) For each birdwatcher, find the average time between two consecutive observations. Return id, name, homestate of each birdwatcher and the average time between two consecutive observations.

Note that the attribute `Observations.prev_oid` is the `oid` of the previous observation by the same watcher, which is NULL if there is no such previous observation. You can find the difference in time between two date and time values as follows: `(odate1+otime1) - (odate2+otime2)` This converts the attributes to a datetime attribute and find the time between them as an interval. In practice, you might need to type cast the result, but no need for it in the quiz.

```
select
   bd.bwid
   , bd.name
   , db.homestate
   , avg( (o2.odate+o2.otime) - (o1.odate+o1.odate) ) as avgobs
from
   birdwatchers bd
   , observations o1
   , observations o2
where
   bd.bwid = o1.bwid
   and bd.bwid = o2.bwid
   and o2.prev_oid = o1.oid  --o2 is the later observation
group by
   bd.bwid
   , bd.name
   , bd.homestate
```