**CS 157A Spring 2020 Homework 4**

Point: 10 (each question is 1 point)

**Part 1** (Listen to the posted Video lecture on Nested/Hash Join and answer following questions):

1. What is the time complexity of nested join algorithm?
   1. (a)  O(m+n)
   2. (b)  O(m\*n)
   3. (c)  O(m)

\*m and n are number of rows of tables being joined.

1. If we use hash join algorithm to join takes and section tables, then which table should

be used to build hash table?  
(a) takes and section (both)  
(b) either of them (takes or section)  
(c) pick a table with fewer number of rows to build hash table (**video said to choose a table that all rows fit in memory)**

**Part 2**: SQL questions  
The following queries are based on the campaign data (campaign-ca-2016.sql)

1. how many contributions are contained in the data?
   1. select count(\*) from campaign;
   2. 180478
2. show the min and maximum amounts of all contributions?
   1. select min(contb\_receipt\_amt),max(contb\_receipt\_amt) from campaign;
   2. -10000|10800
3. list the (distinct) ids and names of all the candidates in the data. order by name?

select distinct cand\_id,cand\_nm from campaign order by cand\_nm;

1. show the candidate name and number of contributions, for each candidate? order by number of contributions in descending order?

select cand\_nm,count(contbr\_nm) from campaign group by cand\_nm order by count(contbr\_nm) desc;

1. show the candidate name and average contribution amount for each candidate, looking at positive contributions only? Order by average amount in descending order?

select cand\_nm,avg(contb\_receipt\_amt) from (select \* from campaign where contb\_receipt\_amt > 0) group by cand\_nm order by avg(contb\_receipt\_amt) desc;

1. show the candidate name and the total amount received by each candidate. Order the output by total amount received.

select cand\_nm,sum(contb\_receipt\_amt) from campaign group by cand\_nm order by sum(contb\_receipt\_amt) desc;

9. how do you set the SQL output so that it no longer goes to a file?

sqlite>

.out stdout

OR

.output (no file name, defaults to standard output)

**Part 3**: Use the courses data (read the files courses-ddl.sql and courses-small.sql).

10. write an SQL query that gives the number of courses taken for every student in the student table?

select name,count(sec\_id) from student natural join takes group by STUDENT.ID;

Zhang|2

Shankar|4

Brandt|1

Chavez|1

Peltier|1

Levy|3

Williams|2

Sanchez|1

Brown|2

Aoi|1

Bourikas|2

Tanaka|2