CS145 Programming Assignment #2

Due Friday May 7*

NOTE: *Before beginning this assignment please skim through* Oracle Version 8 SQL.† *In particular, be sure to understand the sections on "Basic SQL Features", "Indexes", and "Timing SQL Commands". We also recommend doing Written Assignment #4 before doing this programming assignment, since the written assignment will familiarize you with most constructs of SQL.*

1. **Query and Modification.** For this problem, develop and test:

   (a) At least five SQL `SELECT` statements.

   (b) At least two each of the four types of SQL data modification commands: `INSERT` a single row, `INSERT` a subquery, `DELETE`, and `UPDATE`.

   Please note:

   - For this assignment you will be invoking your SQL commands interactively through `sqlplus`, as described in *Getting Started With Oracle*.‡ Of course you should certainly build a `.sql` file, rather than typing the queries in each time you run them.
   - Please write "interesting" queries. You should try to use most or all of the SQL constructs discussed in class and in the textbook (subqueries, aggregates, set/bag operators, etc.). You will not receive full credit if your queries and modifications are extremely simple.
   - We suggest that you experiment with your SQL commands on your small hand-created database before running them on the large database for which you generated data. Initial debugging is much easier when you're operating on small amounts of data. Once you're confident that your commands are working, run them on your complete database.
   - If you discover that most or all of your "interesting" queries return an empty answer on your large database, then you probably didn't follow the instructions in Programming Assignment #1 for generating data values that join properly. You will need to modify your data generator accordingly.
   - Turn in a copy of all of your SQL commands, along with a script illustrating their execution. Your script should be sufficient to convince us that your commands run successfully. Please do not, however, turn in query results that are thousands (or hundreds of thousands) of lines long!

2. **Indexes.** In Problem #1 you probably discovered that some queries run very slowly over your large database. As discussed in class, one principal technique for improving the performance of queries is to create *indexes*. An index on an attribute A of relation R allows the database to quickly find all tuples in R with a given value for attribute A (which is useful when evaluating selection or join conditions involving attribute A). An index can be created on any attribute of any relation, or on several attributes

*Please refer to CS145 Course Information Page (`http://www.stanford.class/cs145/info.html`) for submission instructions and late policy.

†`http://www.stanford.edu/class/cs145/or-nonsql2.html`

‡`http://www.stanford.edu/class/cs145/or-intro.html`

combined. The syntax for creating indexes in Oracle is given in *Oracle Version 8 SQL*. Create at least three useful indexes for your PDA. Run your queries from Problem #1 on your large database with the indexes and without the indexes. Turn in a script showing your commands to create indexes, and showing the relative times of query execution with and without indexes.

Please note:

- As mentioned in *Oracle Version 8 SQL*, Oracle automatically creates indexes for attributes declared as keys. Since you have not yet declared keys in your schema, you may consider creating indexes on key attributes as appropriate, and such indexes should be declared as `UNIQUE`.
- As described in *Oracle Version 8 SQL*, in order to set the system to show query execution times you must issue the command "`SET TIMING ON`" at the `sqlplus` prompt.
- Your timings will be affected by external factors such as system load. However, for some of your queries, with appropriate indexes you should see a consistent dramatic difference between the execution times with indexes and the times without. If others of your queries do not show performance improvement even when relevant indexes are created, please include a short note suggesting why this might be the case.

3. **Views.**

   (a) Create two useful views on top of your database schema. For each view, write a query involving the view. Turn in a copy of your commands along with a script showing their successful execution.

   (b) View updates are extremely tricky to deal with. For this problem, you will perform some sleuth work in order to find out whether Oracle handles view updates correctly.

   - First, use the commands in `/usr/class/cs145/src/PA2/setup.sql` to set up a few tables and views for your experiment. Be sure to read and understand the view definitions in this file.
   - For each view created in `setup.sql` (except `cs145_v1`, which serves as an example below), find a view modification statement that is allowed by Oracle but produces a rather unexpected result when executed. This modification can be `INSERT`, `DELETE`, or `UPDATE`. For example, `cs145_v1` is defined as a selection view over the table `cs145_r1(a,b)`. Initially, both the table and the view are empty. Let's try the following modification:

     ```
     INSERT INTO cs145_v1 VALUES(1,1);
     ```

     Oracle happily reports "1 row created", so you'd expect (1, 1) to appear in `cs145_v1`. However, when you try:

     ```
     SELECT * FROM cs145_v1;
     ```

     you will find that `cs145_v1` is still empty!
     Turn in a script showing the modifications you came up with for views `cs145_v2` and `cs145_v3`, and show the view contents (using `SELECT * ...`) both before and after the modifications.
   - When you are done, use `/usr/class/cs145/src/PA2/cleanup.sql` to clean up the tables and views created by `setup.sql`.