Homework Assignment

SQL: Please refer the slides.

Database Description: Please refer to the slides.

Database installation : Please refer "MySQL_Installation.pdf".

Database creation : Please refer "Tennis_Database_Creation.pdf".

Please do not modify the database after it's complete creation, use queries only for data fetching.

The database you will be writing queries about contains information concerning the Association of Tennis Professionals. It contains information about players, matches and tournaments concerning the ATP . If you are unfamiliar with tennis terminology (especially concerning scoring), a quick description can be found on Wikipedia at

https://en.wikipedia.org/wiki/Tennis#Manner

Several assumptions about the database are included below:

• The source of the data used for this database only contained information about the ATP (men's tennis) and not the WTA (women's tennis). All players involved are male; we have no information pertaining to women's singles or doubles, or mixed doubles.

•While it should be clear that Player's PID and Tournament's TID field should be unique, this is also true for Match's MID field; no IDs are re-used throughout different tournaments. Thus, a Match's MID field uniquely represents it across all Matches known to exist. This avoids having to compare Tournament IDs as well as Match IDs for several of the queries.

•When a player participates in a tournament, he first has to register. For the purposes of our database, this includes receiving a registration number for the tournament. Because the ATP manages all of the tournaments in our database, we assume that a registration number is unique throughout all of the tournaments. This means that given a only a registration number, we can determine the tournament it corresponds to 1.

•A Player's ID (Player.PID) is only used in our database for registration in tournaments. Allother locations the Player is referred to by the RegistrNum (notably in the Winner attribute of MatchResults).

•Two players participating in doubles tennis as a team share the same registration number. Note that several players register for both singles and doubles tennis in the same tournament, players in this category get two separate registration numbers.

•Both a player registered for singles play or two people registered for doubles play under the same number (i.e. are playing with each other) can both be referred to as a team: a team of one or a team of two.

•In most tournaments, the top fraction of the players registering are seeded, to help spread out the players expected to perform well (this makes the later matches more interesting!). Seed values are given on a per-tournament basis. Players who do not receive a seed in a particular tournament have a NULL value instead.

•The Tournament relation stores the number of rounds the tournament has as an attribute, which is an integer. The Round attribute of the Match relation is similar. A value of 1 corresponds to the first round, 2 to the second, and so on. Match.Round does not say "Quarterfinals," "Semifinals," or "Finals." You will have to figure out how to compute what rounds correspond to these. Don't try to hard-code values in (i.e. if you look at the database and realize that the US open singles had a total of 7 rounds, so the Semifinals is round 6). Rather, use a nested query to help.

•Each entity in the Set relation contains information about a particular set of some match. The attributes involving the "Winner" refer to the winner of the winner of the match, not necessarily the set (similarly for the "Loser" attributes). In cases where the set was determined by a tie-breaker, the TieBreaker relation has WinnerTB and LoserTB, which contain the points won in the tie-breaker game. Again, the "Winner" column of the TieBreaker refers to the winner of the match, not necessarily the set.

•Sometimes players will need to retire from a match, which counts as a forfeit. If this happens, we keep records for all completed sets of the match. For the set in which the player retired, the WinnerGames and LoserGames attributes hold the number of games won and lost. The RetiredMatch relation holds all Match IDs where the loser retired.

SQL Queries for practice to familiar with tennis database and SQL:
[No need to submit below queries, only for practice]

1. Show the names of all the players in the database.

2.Find the names, tournament types, and lengths (in days) of all tournaments that were longer than one week.

3.Find the names of all the players who have ever been assigned a seed for   any tournament [single or doubles].

4.Print the tournament name, the tournament type, the start and end dates, and the number of rounds for tournaments having more than 5 rounds

5.   Find the name of the player who has registered for the most tournaments, get the number of tournaments he has registered for. [Registration for both singles and doubles in a same tournament counts as two tournaments.]

6.Find the names of all pairs of players that have played against each other in both singles and doubles matches.

7.For all singles quarterfinal, semifinal, and final round matches that only took 3 sets, find list of the Name of the Tournament, Year, Match Winner and the score of the match [For score of the match, print two columns per set, displaying the number of games each player won. Try to call these columns something meaningful, and ignore any tiebreaker results.].