

Tennis – Grand Slam Data

Please submit the queries you run for each question in a single plain text file (e.g. in notepad).

Instructions

Open the GrandSlamData2.csv data set in Excel. Clean the data by removing blank rows and cleaning out unwanted cells if necessary.

Create a suitable database and table and import the csv data using column headers from spreadsheet. Note that you may have to clean the file to remove foreign language characters.

Using suitable queries:

1. Produce a list of years and winners of Wimbledon since 1960.
2. Find all the players from USA who have won a tournament and how many each has won.
3. Find the average seeding of the winner and the runner up in the French open.
4. Find the number of players from outside the top 5 who have won the US Open.
5. Find the number of distinct players who have won Wimbledon.
6. Produce a list of the top ten tournament winners of all time.
7. Find the number of tournaments won by players from each country since 1950.
8. Find the proportion of times the French Open has been won by a non-seeded player.
9. Using an appropriate string function find the proportion of finals where at least one set went to a tie break.
10. Using a subquery find the list of tournaments, years and winners where the ranking of the winner was lower than the average ranking of the winner across all tournaments.
11. Find the most common combination of winner and runner up in a tournament. You may first need to use the CONCAT() string function to combine the winner and runner up names.
12. List the players who have won a grand slam (all four tournaments in the same year).

All questions carry equal marks.