CS-586 DATABASE NAMAGEMENT SYSTEM

TOPIC : 2021 Olympics in Tokyo database

NAME : SAITEJA GARLAPATI

PSU ID : 985031204

DESCRIPTION :

Games are always interesting to me. This data source matches with my interest and it gives a good scope to understand different perspectives of data like, games, athletes, coaches and achievements etc. This is a good source for the scope of database understanding. The datasets taken for this project are Athletes, Coaches, Teams participating , Medals as well as the Entries by gender. These datasets contains the details of over 11,000 athletes, with 47 disciplines, along with 743 teams taking part in the 2021(2020) Tokyo Olympics. It contains their names, countries represented, discipline, gender of competitors, name of the coaches. This dataset link contains xlsx files. I will try to import these datasets into the POSTGRESQL using the import option available. I will create same datatypes as of xlsx data as it is necessary for same datatypes when importing. We will be using many datatypes like text integers enums to acquire hands on knowledge and to give preciseness of data storage.

QUESTIONS :

1) Create the tables (Athletes, Coaches, Entries, Medals, Teams) and add the appropriate primary and foreign keys.
2) Find the names of the athletes who's discipline is badminton
3) Count the total number of countries who has swimming as their discipline.
4) Find the names of the coaches and their respective countries whose event value is men.
5) Find the name and NOC of coaches that do not have football as their game.
6) Find the countries that play archery and those that do not have mixed team as their event.
7) Find the countries and their ranks whose total medals are between 30 to 100.
8) Find the coaches names from spain who plays atleast two different games.
9) Count the number of games who has males and females less than 100.
10) Find the top countries that received most gold medals and find the countries that received the least gold medals.
11)   (i)   Find the country that produces the highest number of coaches.

       (ii)   Find the countries that produces the least number of coaches.
12) Find the number of the athletes who has more participation in each game.
13) Produce a list of the number of different discipline that are had by countries.
14) Find the coach that got more medals.

15) List out the athletes with same first name as name 1 and name2.
16) List out coach, number of athletes he/she trains and medals that the team received.
17) Which countries played either basketball or karate.
18) Display the games and count of the athletes and count of the athletes that played by them.
19) What are the games that have same average of male and female
20) What are the countries that have same number of silver medals.

SOURCE :

The dataset that I intend to use for this project is   https://www.kaggle.com/arjunprasadsarkhel/2021-olympics-in-tokyo .