CS315: Database Systems

Assignment 1 (100 marks)

Due on: 14th February, 2020, 10:00pm

Consider the three census files `multilingual-age.ods`, `multilingual-education.ods` and `age-education.ods` available in `canvas`. Please use their corresponding `.csv` versions. However, please understand the format of the data from the `.ods` files.

Q1 (30 marks)

Create a SQlite3 database file called `population.db`.
  After studying the three files, appropriately create the relations in the database.
  Please explain the choice of your relation schemas.

Q2 (30 marks)

Insert the data appropriately from the three files.
  If you have modified the files, please include the scripts of the modification, and explain the process.

Q3 (40 marks)

Answer the following queries:

1. Order the states and union territories by the percentage of trilingual population in them.

2. Which age group has the largest percentage of multi-lingual population?

3. For the whole of India, which age group depicts the most skewed gender ratio?

4. Find the total number of people in India who speaks only one language.

5. In which state or union territory is the average age the highest? (For the age-group 0-6, assume the average age to be 3, while for the age group 80+, assume the average age to be 85. For the groups 20-24, etc., assume the average age to be 22, etc.)

Submission

Submit the assignment as a `zip` file named `roll-asgn1.zip` with the procedure for each question explained clearly. Please also submit the code/script files but not the database files. Please submit at `canvas`.