CSCI 4180: Big Data Analytics
CSCI/DASC 6010: Big Data Analytics and Management
Spring 2023
Homework 2 – Relational databases and SQL
Due Sunday, February 5, at 11:59 PM

Preliminary

Download the MONDIAL database (the `mondial-schema.psql` is used to create the tables, `mondial-inputs.psql` contains the data along with the queries to load it into the database, and `mondial-drop-tables.psql` contains the queries to delete these data once this assignment is completed).

1 SQL exercises (100 pts)

Load the data in the database and list the **queries** and **their results** for each of the following questions.

1. (20 pts) How many distinct airports are in the database?

2. (20 pts) What's the longest country code and how many characters it has?

3. (20 pts) Which country has the largest number of airports?

4. (20 pts) Which country has the largest largest land area comprised of islands?

5. (20 pts) Which countries have a population between 1% and 10% of the world population?

2 Extra credit (10 pts)

For extra credit you will need to use R Studio and the Rmd template file provided. The slides for the relational databases, `03. Relational databases.pdf` contain content that is not searchable (you cannot double-click on a word and have it highlighted). Use the template provided to write the content for these slides.

- If your last name starts with A–E, create content only for slides 1 through 6 (the slide numbers are in the bottom-right corner).

- If your last name starts with F–J, create content only for slides 7 through 13 (the slide numbers are in the bottom-right corner).

- If your last name starts with K–P, create content only for slides 14 through 20 (the slide numbers are in the bottom-right corner).

- If your last name starts with R–Z, create content only for slides 21 through 27 (the slide numbers are in the bottom-right corner).

Submit

Upload to Canvas the following:

- **Required:** Your answers to questions 1 through 5 (one PDF file with the **queries** and **their corresponding results**, and one SQL file with the queries).

- **Optional:** Your extra credit files (one Rmd and one PDF file).

Compress all files (two PDF, one SQL, and one Rmd) into one ZIP file and upload this file.