Lab Sheet 3

Grouping And Aggregation

This section is based on the `countries` that we encountered in Lecture 6. The file `countries.sqlite` containing the database you will need is linked to from the Canvas page.

You will recall that this single-table DB had the following schema:

```
countries(name, region, area, population, gdp)
```

1. What is the greatest area of any country?
2. What is the largest population of any country in Africa?
3. What is the total GDP of Europe?
4. List the names an populations of all countries whose GDP is not known (NULL).
5. List the names and GDPs of all countries for which a GDP is known.
6. List the name and average GDP of each region.
7. List all the countries whose name contains the region name as a substring.
8. List the minimum and maximum per capita GDP for each region.
9. List the number of countries and total population for each of the following regions: Europe, Africa and the Middle East.
10. What is the total population, area and GDP of France, Germany and Spain (taken together)?
11. List by region the number of countries in that region with a population greater than 100 million.
12. For each letter of the alphabet, list the number countries whose names begin with that letter and the first and last country (alphabetically). [1]
13. List all the countries in the world region by region (alphabetically) and by descending order of population within each region.
14. List the number of countries and population density (area divided by population) in all regions with total population greater than one billion.

[1] SQLite supports a range of useful functions, such as `substr` that permits you to extract a portion of a string e.g. its first letter. See `https://sqlite.org/lang_corefunc.html` for details of this and other useful functions.

Most of the tasks below rely on SQL concepts we haven't seen yet such as subqueries and joins. You may wish to return to these in the weeks ahead to tackle these when you have time.

15. List the names of all the countries in the same region as Jordan.

16. How many countries are in the same region as Jordan?

17. List those countries in the same region as Spain that have a greater area than Spains's.

18. (Tricky!) List all the countries that have an area that is at least 10% of the total area of the region to which they belong.

19. List the countries in decreasing order of population band; for each band list the number of countries and the minimum and maximum area. We use 100 million as our band with, i.e. the first band consists of countries with populations less than 100 million and so on.

20. What is the minimum population of any country in the same region as China?

21. (Tricky!) Determine the name of the country with the greatest population. (Note: Bizarrely, SQLite will give you the answer with the "obvious" approach

    ```
    SELECT name, MAX(population)
    FROM countries;
    ```

    Most SQL systems e.g. MySQL will serve the maximum population number coupled with a "random" country name. SQL generally frown on mixing aggregates and non- aggregates on the SELECT line, except for non-aggregates used ro define groups in GROUP BY clause.

    To rely on such a quirky works-only-with-SQLite feature for your queries is to be avoided. What, for example, would you expect the following query to produce?

    ```
    SELECT name, MAX(population), MAX(area), MAX(dgp)
    FROM countries;
    ```

22. List all the countries whose per capita GDP is at least as great as China's. The list should appear in descending order order of per capita GDP.