CS/CE 457/464 - Homework Assignment 4: SQL

Owais Aijaz – oa07610

- Create a SQL database tables for both datasets countries.csv and cities.csv separately, using a RDBMS (PostgreSQL preferred). You need to submit a create table query in the final document.

```sql
CREATE TABLE cities(
    city_id SERIAL PRIMARY KEY,
    city_name VARCHAR(255) NOT NULL,
    city_population BIGINT,
    capital BOOLEAN,
    fact_id INT,
    CONSTRAINT fk_country_fact_id FOREIGN KEY(fact_id) REFERENCES countries(fact_id)
);
```

Query   Query History   ↗

```sql
1  ∨  CREATE TABLE countries (
2          fact_id SERIAL PRIMARY KEY,
3          country_code VARCHAR(10) NOT NULL,
4          country_name VARCHAR(255) NOT NULL,
5          country_region VARCHAR(255),
6          country_sub_region VARCHAR(255),
7          country_area NUMERIC,
8          country_area_land NUMERIC,
9          country_area_water NUMERIC,
10         country_population BIGINT,
11         country_population_growth DECIMAL,
12         country_birth_rate DECIMAL,
13         country_death_rate DECIMAL,
14         country_migration_rate DECIMAL
15     );
16
```

Data Output   Messages   Notifications

```
CREATE TABLE

Query returned successfully in 67 msec.
```

- Load/Import the dataset into the table.

Done via pgAdmin itself.

	PID	Type	Server	Object	Start Time ∨	Status
📄	1824	Import Data	PostgreSQL 16 (loca…	homework4/public.cities	9/23/2024, 8:49:33 …	Finished
📄	7304	Import Data	PostgreSQL 16 (loca…	homework4/public.countries	9/23/2024, 8:49:00 …	Finished

- Query the database tables and interpret the results, displaying:
1. the count of total number of records in each table.

```
1    select count(*) from cities;
```
Data Output   Messages   Notifications

	count bigint
1	397

```
3    select count(*) from countries;
```
Data Output   Messages   Notifications

	count bigint
1	261

2. the count of number of cities for each country in descending order of count

```
5 ∨  SELECT c.country_name, COUNT(ci.city_id) AS city_count
6    FROM countries c JOIN cities ci ON c.fact_id = ci.fact_id
7    GROUP BY c.country_name
8    ORDER BY city_count DESC;
9
```

Data Output   Messages   Notifications

	country_name character varying (255)	city_count bigint
1	South Africa	7
2	China	6
3	Russia	6
4	France	6
5	Venezuela	6
6	Colombia	6
7	Canada	6
8	Korea, South	6
9	Morocco	6
10	Indonesia	6
11	Iraq	6
12	India	6
13	Iran	6
14	Ukraine	6
15	United States	6

Total rows: 210 of 210   Query complete 00:00:00.107   Ln 8, Col 26

3. the count of regions and sub-regions in each country. Sort them by ascending order of country name. (use group by)

```sql
SELECT country_name, COUNT(country_region) AS region_count, COUNT(country_sub_region)
AS subregion_count FROM countries GROUP BY country_name ORDER BY country_name;
```

Data Output | Messages | Notifications

	country_name character varying (255)	region_count bigint	subregion_count bigint
1	Afghanistan	1	1
2	Akrotiri	1	1
3	Albania	1	1
4	Algeria	1	1
5	American Samoa	1	1
6	Andorra	1	1
7	Angola	1	1
8	Anguilla	1	1
9	Antarctica	1	1
10	Antigua and Barbuda	1	1
11	Arctic Ocean	1	1
12	Argentina	1	1
13	Armenia	1	1
14	Aruba	1	1
15	Ashmore and Cartier Islands	1	1
16	Atlantic Ocean	1	1

Total rows: 261 of 261   Query complete 00:00:00.092   Ln 11, Col 20

4. Top 10 most populous capital cities. Display country, city and population in descending order.

```sql
SELECT c.country_name, ci.city_name, ci.city_population
FROM countries c JOIN cities ci ON c.fact_id = ci.fact_id
WHERE ci.capital IS true ORDER BY ci.city_population DESC LIMIT 10;
```

Data Output | Messages | Notifications

	country_name character varying (255)	city_name character varying (255)	city_population bigint
1	Japan	Tokyo	37217000
2	India	New Delhi	22654000
3	Mexico	Mexico City	20446000
4	China	Beijing	15594000
5	Bangladesh	Dhaka	15391000
6	Argentina	Buenos Aires	13528000
7	Philippines	Manila	11862000
8	Russia	Moscow	11621000
9	Egypt	Cairo	11169000
10	Indonesia	Jakarta	9769000

5. Average city population of capital and non-capital cities. (use group by)

```sql
17    SELECT capital, AVG(city_population) AS avg_city_pop FROM cities GROUP BY capital;
```

Data Output   Messages   Notifications

	capital boolean	avg_city_pop numeric
1	false	2891392.063492063492
2	true	2140442.307692307692

6. Average country birth rate for each region and sub-region (use group by)

```sql
19    SELECT country_region, country_sub_region, AVG(country_birth_rate) AS avg_birth_rate
20    FROM countries GROUP BY country_region, country_sub_region;
```

Data Output   Messages   Notifications

	country_region character varying (255)	country_sub_region character varying (255)	avg_birth_rate numeric
1	Europe	Northern Europe	11.6621428571428571
2	Asia	Southern Asia	21.0944444444444444
3	Oceania	Micronesia	19.7000000000000000
4	Africa	Sub-Saharan Africa	32.4055813953488372
5	Americas	Northern America	11.2000000000000000
6	Europe	Western Europe	10.1633333333333333
7	NA	NA	9.2364583333333333
8	Americas	Latin America and the Caribbean	15.8915384615384615
9	Europe	Southern Europe	9.8693333333333333
10	Asia	South-eastern Asia	19.8177777777777778
11	Europe	Eastern Europe	10.1520000000000000
12	Asia	Central Asia	20.5820000000000000
13	Oceania	Melanesia	21.9900000000000000
14	Oceania	Australia and New Zealand	4.2466666666666667
15	Asia	Western Asia	18.5035294117647059
16	Africa	Northern Africa	22.6057142857142857

Total rows: 18 of 18   Query complete 00:00:00.101   Ln 20, Col 60