

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

Breeha Qasim

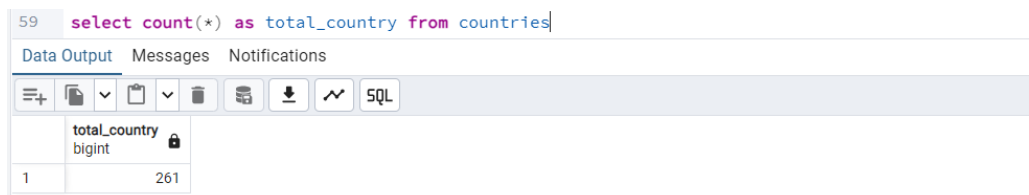
bq08283@st.habib.edu.pk

Part 1

- Create a SQL database tables for both datasets countries.csv and cities.csv separately, using a RDBMS (PostgreSQL preferred).

```
CREATE TABLE countries (  
    fact_id SERIAL PRIMARY KEY,  
    country_code VARCHAR(2),  
    country_name VARCHAR(255),  
    country_region VARCHAR(255),  
    country_sub_region VARCHAR(255),  
    country_area INTEGER,  
    country_area_land INTEGER,  
    country_area_water INTEGER,  
    country_population BIGINT,  
    country_population_growth DECIMAL(4, 2),  
    country_birth_rate DECIMAL(4, 2),  
    country_death_rate DECIMAL(4, 2),  
    country_migration_rate DECIMAL(4, 2)  
);  
  
CREATE TABLE cities (  
    city_id SERIAL PRIMARY KEY,  
    city_name VARCHAR(255),  
    city_population INTEGER,  
    capital BOOLEAN,  
    fact_id INTEGER REFERENCES countries(fact_id)  
);
```

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



59 `select count(*) as total_country from countries`

Data Output Messages Notifications

	total_country bigint
1	261

59 `select count(*) as total_cities from cities`

Data Output Messages Notifications

	total_cities bigint
1	397

- the count of number of cities for each country in descending order of count (use group by)

35 `select country_name, count(*) as city_count`
 36 `from cities join countries on cities.fact_id = countries.fact_id`
 37 `group by country_name`
 38 `order by city_count desc;`

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

Total rows: 210 of 210 Query complete 00:00:00.188 Ln 30, Col 1

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

40 `select country_name, count(country_region) as count_region, count(country_sub_region) as count_sub_region`
 41 `from countries`
 42 `group by country_name, country_region, country_sub_region`
 43 `order by country_name asc;`

Data Output Messages Notifications

	country_name character varying (255)	count_region bigint	count_sub_region 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

Total rows: 261 of 261 Query complete 00:00:00.125 Ln 40, Col 106

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

```
45 SELECT c.country_name, ci.city_name, ci.city_population
46 FROM cities ci
47 JOIN countries c ON ci.fact_id = c.fact_id
48 WHERE ci.capital = 'TRUE'
49 ORDER BY ci.city_population DESC
50 LIMIT 10;
```

Data Output Messages Notifications

	country_name character varying (255)	city_name character varying (255)	city_population integer
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)

```
51 select capital, avg(city_population) as average_population
52 from cities
53 group by capital;
```

Data Output Messages Notifications

	capital boolean	average_population numeric
1	false	2891392.063492063492
2	true	2140442.307692307692

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

```
55 select country_region, country_sub_region, avg(country_birth_rate) as average_birth_rate
56 from countries
57 group by country_region, country_sub_region;
```

Data Output Messages Notifications

	country_region character varying (255)	country_sub_region character varying (255)	average_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
Total rows: 18 of 18 Query complete 00:00:00.309 Ln 56, Col 15			