SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

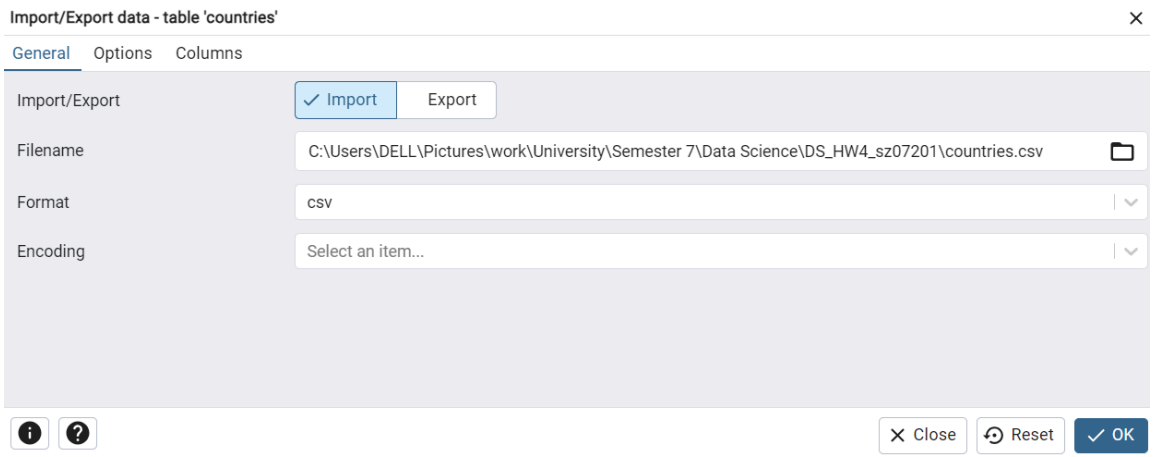
Creating Countries' Relation code:

Query | Query History

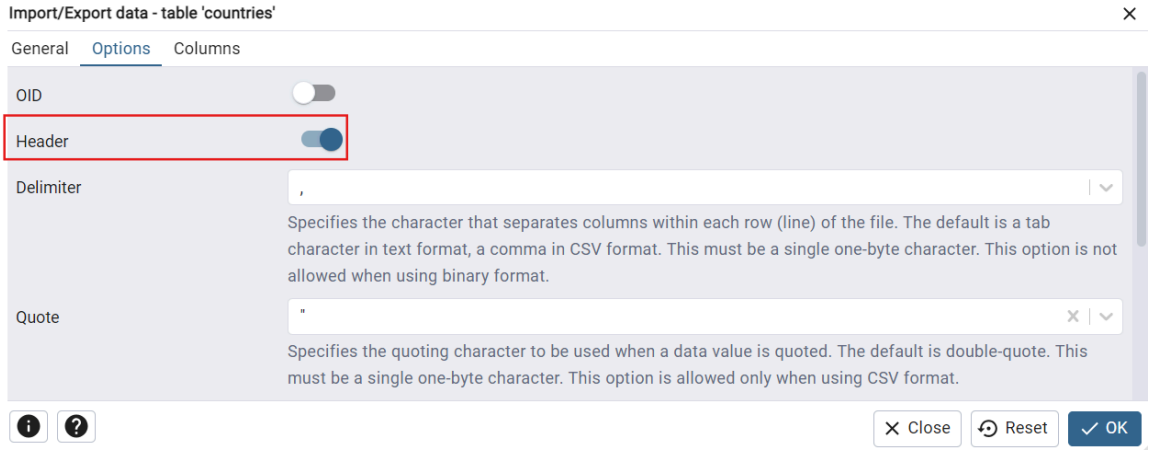
```
CREATE TABLE countries(
fact_id INT,
country_code VARCHAR(3),
country_name VARCHAR(50),
country_region VARCHAR(50),
country_sub_region VARCHAR(50),
country_area BIGINT,
country_area_land BIGINT,
country_area_water BIGINT,
country_population BIGINT,
country_population_growth FLOAT,
country_birth_rate FLOAT,
country_death_rate FLOAT,
country_migration_rate FLOAT);
```

I later "imported" the csv file into this table. It was initially failing to import due to my varchar being small or int being small so I kept them large just in case for any side-case, and import succeeded. Proof of which I have attached below

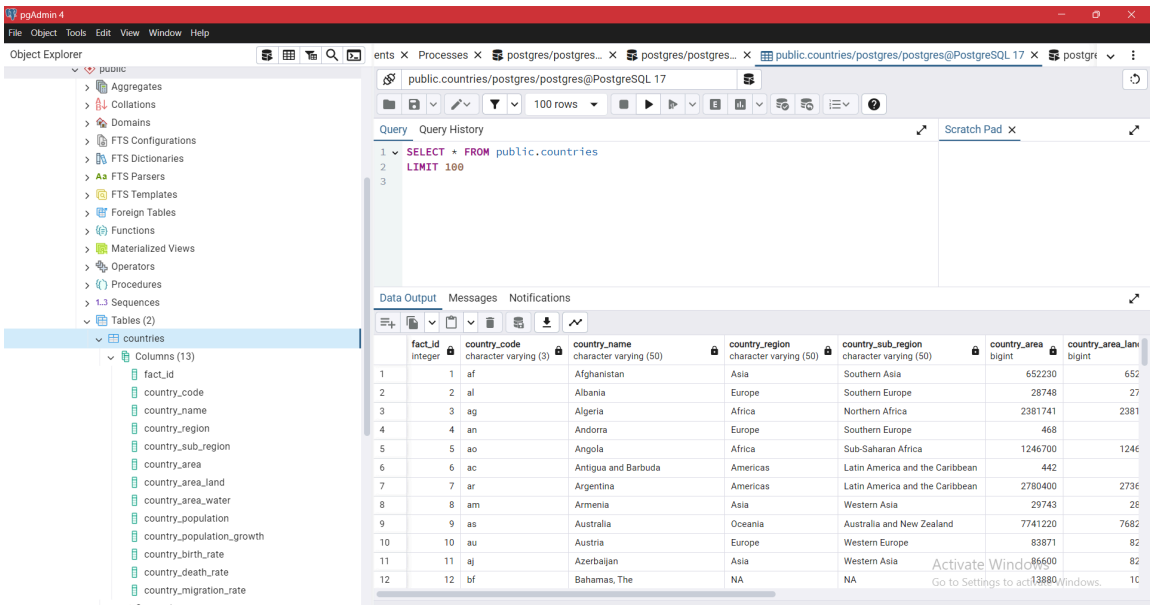
Importing:



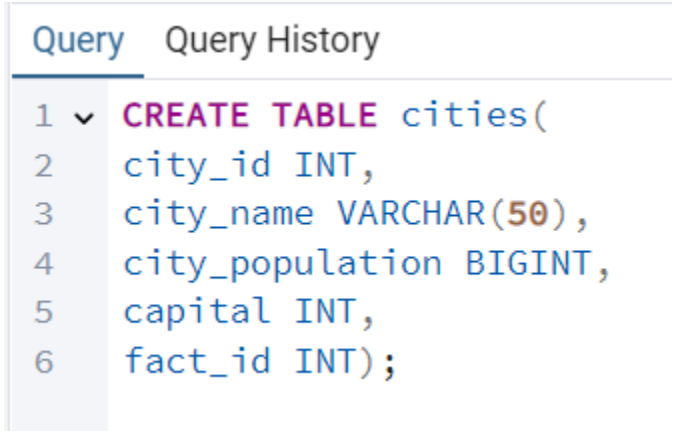
SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1



After succeeding, showcasing the first 100 rows



Doing the same for Cities!



SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

We successfully imported the data the same way, and used the same settings.

Process completed

Copying table data 'public.cities' on database 'postgres' and server 'PostgreSQL 17 (localhost:5432)'

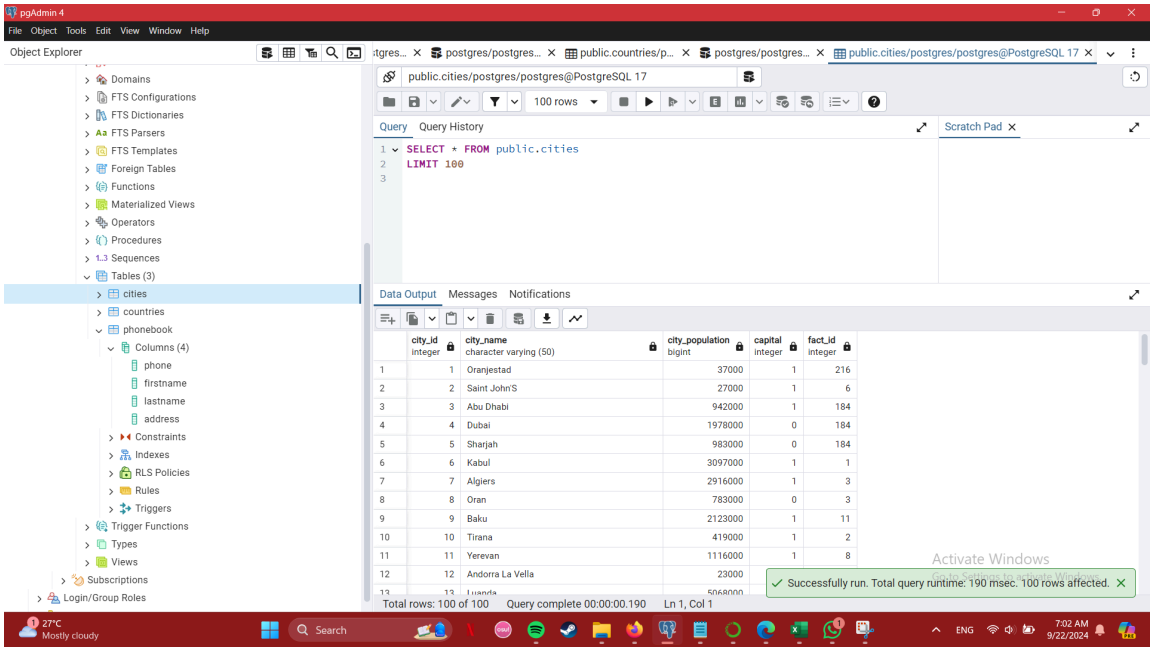
View Processes

Process started

Copying table data 'public.cities' on database 'postgres' and server 'PostgreSQL 17 (localhost:5432)'

View Processes

Activate Windows
Go to Settings to activate Windows.



SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

Query the database tables and interpret the results, displaying:

1. the count of the total number of records in each table.

Query | Query History

```
1 SELECT COUNT(*) AS total_records
2 FROM countries;
```

Data Output | Messages | Notifications

	total_records bigint
1	261

Query | Query History

```
1 SELECT COUNT(*) AS total_records
2 FROM cities;
```

Data Output | Messages | Notifications

	total_records bigint
1	397

The dataset contains 261 countries and 391 cities, indicating an average of approximately 1.5 cities per country.

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

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

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

Data Output Messages Notifications

	country_name character varying (50)	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
16	Brazil	6
17	Japan	6
18	Pakistan	6
19	Italy	6
20	United Kingdom	6
21	Turkey	6

Data Output Messages Notifications

	country_name character varying (50)	city_count bigint
23	Australia	6
24	Mexico	6
25	Argentina	6
26	Saudi Arabia	5
27	Congo, Democratic Republic of the	5
28	Germany	4
29	Syria	4
30	Philippines	4
31	Bangladesh	4
32	Vietnam	4
33	Malaysia	3
34	Spain	3
35	United Arab Emirates	3
36	Chile	3
37	Israel	3
38	Bolivia	3
39	Burma	3
40	Netherlands	3
41	Thailand	2
42	Switzerland	2
43	Mozambique	2

The data presents the count of cities associated with various countries, highlighting significant disparities in urban density and distribution. Notably, South Africa leads with seven cities, showcasing a relatively high urban count. A group of countries, including China, Russia, and several others, have six cities each, indicating robust urbanization. Conversely, many nations, such as Saudi Arabia and the Democratic Republic of the

Congo, have slightly fewer, with five cities. A large number of countries, including those with smaller populations or geographic areas, have only one city listed, *which reflects either a concentration of urbanisation in a single location or a limited number of significant urban centres.*

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

```
38  SELECT
39      country_name,
40      COUNT(DISTINCT country_region) AS region_count,
41      COUNT(DISTINCT country_sub_region) AS sub_region_count
42  FROM
43      countries
44  GROUP BY
45      country_name
46  ORDER BY
47      country_name ASC;
```

Data Output | Messages | Notifications

	country_name character varying (50)	region_count bigint	sub_region_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

Extra (Cause I believe this question wanted to ask us this): the count of countries in each region and sub-regions. Sort them by ascending order of count of countries (use group by)

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

```
26  SELECT
27      country_region,
28      country_sub_region,
29      COUNT(*) AS country_count
30  FROM
31      countries
32  GROUP BY
33      country_region,
34      country_sub_region
35  ORDER BY
36      country_count ASC;
```

Data Output | Messages | Notifications

	country_region character varying (50)	country_sub_region character varying (50)	country_count bigint
1	Asia	Eastern Asia	5
2	Oceania	Melanesia	5
3	Asia	Central Asia	5
4	Americas	Northern America	5
5	Oceania	Micronesia	6
6	Oceania	Australia and New Zealand	6
7	Africa	Northern Africa	7
8	Asia	South-eastern Asia	9
9	Asia	Southern Asia	9
10	Europe	Western Europe	9
11	Oceania	Polynesia	9
12	Europe	Eastern Europe	10
13	Europe	Northern Europe	14
14	Europe	Southern Europe	15
15	Asia	Western Asia	17
16	Americas	Latin America and the Caribbean	39
17	Africa	Sub-Saharan Africa	43
18	NA	NA	48

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

We can see countries categorized by their respective regions and sub-regions, revealing significant diversity across the globe. Sub-Saharan Africa emerges as the most populous region, encompassing 43 countries, highlighting its vast cultural and geographical variety. Latin America and the Caribbean follow closely with 39 countries, indicating a rich tapestry of nations characterised by shared historical and cultural experiences.

In Asia, Western Asia stands out with 17 countries, while Southern Asia and South-eastern Asia each have 9. This suggests a substantial concentration of countries in these areas, reflecting a complex interplay of cultures and histories. In contrast, Oceania shows moderate representation, with sub-regions like Melanesia and Micronesia having 5 to 6 countries each, indicating a more limited yet distinct diversity.

Europe is characterised by a significant clustering of countries, particularly in Southern Europe, which has 15 countries, and Northern Europe, which has 14. This distribution illustrates the rich cultural and historical connections within the continent. Additionally, the entry labelled "NA," with 48 countries, suggests the presence of territories or nations that do not fit neatly into conventional regional classifications, emphasising the complexity of global geography. Overall, the data provides a comprehensive view of how countries are distributed across different regions and sub-regions, highlighting both diversity and interconnectedness.

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

```sql
SELECT
    countries.country_name,
    cities.city_name,
    cities.city_population
FROM
    cities
JOIN
    countries
ON
    cities.fact_id = countries.fact_id
WHERE
    cities.capital = 1
ORDER BY
    cities.city_population DESC
LIMIT 10;
```

This query retrieves the top 10 most populous capital cities by joining the cities and countries tables on “fact_id”, which links each city to its country. The “WHERE” clause ensures only capital cities are selected (capital = 1), and the results are sorted in descending order of city population using “ORDER BY cities.city_population DESC”. The “LIMIT 10” clause restricts the output to the top 10 cities.

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

Data Output | Messages | Notifications

	country_name character varying (50)	city_name character varying (50)	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

The data showcases the most populous cities across different countries, with Tokyo, Japan, at the forefront, boasting around 37.2 million residents, followed by New Delhi, India, at 22.6 million. Other major cities include Mexico City (20.4 million), Beijing (15.6 million), and Dhaka (15.4 million), highlighting significant urban centers in both developed and developing nations. Buenos Aires, Manila, Moscow, Cairo, and Jakarta further illustrate the diversity of urban populations and their importance as cultural, political, and economic hubs. This information reflects global urbanization trends and the associated challenges of managing large, densely populated cities.

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

```
50  SELECT
51      capital,
52      AVG(city_population) AS average_population
53  FROM
54      cities
55  GROUP BY
56      capital;
```

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

	capital integer	average_population numeric
1	0	2891392.063492063492
2	1	2140442.307692307692

The results reveal that non-capital cities have a significantly larger average population (approximately 2,891,392) compared to capital cities (around 2,140,442). This suggests that non-capital cities may serve as major urban centers or economic hubs, attracting larger populations due to better job opportunities and infrastructure. Additionally, some capitals may be smaller due to historical, geographical, or urban planning factors. This data challenges the assumption that capital cities are always the most populous, highlighting the complex dynamics of urban development and population distribution across regions

The average birth rate data highlights significant regional disparities, with Sub-Saharan Africa recording the highest average at approximately 32.41, indicating high fertility levels influenced by socio-economic factors. Northern Africa follows with 22.70, while Oceania, particularly Australia and New Zealand, shows notably low rates around 4.25. In the Americas, Latin America and the Caribbean average 15.89, compared to Northern America's 11.20. Most Asian regions range from 11.67 to 21.09, with Central Asia at 20.58, reflecting varied reproductive health outcomes. European sub-regions exhibit low birth rates below 12, pointing to challenges related to aging populations. Overall, these trends emphasize the need for targeted policies addressing population growth and healthcare across different regions.

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

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

```
58  SELECT
59      country_region,
60      country_sub_region,
61      AVG(country_birth_rate) AS average_birth_rate
62  FROM
63      countries
64  GROUP BY
65      country_region,
66      country_sub_region
67  ORDER BY
68      country_region,
69      country_sub_region;
```

	country_region character varying (50)	country_sub_region character varying (50)	average_birth_rate double precision
1	Africa	Northern Africa	22.695714285714285
2	Africa	Sub-Saharan Africa	32.40558139534885
3	Americas	Latin America and the Caribbean	15.891538461538458
4	Americas	Northern America	11.2
5	Asia	Central Asia	20.582
6	Asia	Eastern Asia	11.674000000000001
7	Asia	South-eastern Asia	19.817777777777778
8	Asia	Southern Asia	21.094444444444445
9	Asia	Western Asia	18.503529411764706
10	Europe	Eastern Europe	10.151999999999997
11	Europe	Northern Europe	11.662142857142856
12	Europe	Southern Europe	9.869333333333334
13	Europe	Western Europe	10.163333333333334
14	NA	NA	9.236458333333335
15	Oceania	Australia and New Zealand	4.246666666666667
16	Oceania	Melanesia	21.990000000000002
17	Oceania	Micronesia	19.7
18	Oceania	Polynesia	14.833333333333334

The birth rate data presents a nuanced picture of global demographic trends, particularly highlighting significant disparities across regions. In Sub-Saharan Africa, the average birth rate of approximately 32.41 indicates a persistently high fertility rate, often driven by factors such as limited access to education and healthcare. This situation underscores the urgent need for enhanced reproductive health services and education initiatives to empower individuals in family planning and improve overall health outcomes.

SYED ASGHAR ABBAS ZAIDI
DATA SCIENCE L1 - sz07201
HW4 - Part 1

In contrast, regions like Northern America and Europe exhibit markedly low birth rates, frequently falling below replacement level. This trend points to an aging population and potential future labor shortages, raising concerns about economic sustainability and the need for policies that support families, such as parental leave and childcare support. The low birth rates in these areas reflect a shift towards smaller family sizes, often influenced by urbanization and changing societal norms.

Asia presents a varied landscape, with significant differences in birth rates across sub-regions. Southern and Western Asia maintain higher birth rates, which may be influenced by cultural factors and socio-economic conditions, while Eastern Asia shows a trend towards declining fertility. This divergence reflects the complexities of population dynamics in a rapidly changing socio-economic context, where urbanization and increased educational attainment are leading to smaller family sizes.

Oceania's data further illustrates these contrasts, with Australia and New Zealand displaying low birth rates indicative of effective family planning and education. In contrast, Melanesia's higher birth rate signals differing socio-economic challenges. These trends emphasize the need for tailored policies that address the unique reproductive health and economic contexts of each region, ultimately aiming for sustainable population growth and improved quality of life.

Overall, the insights drawn from these trends reveal the intricate interplay of cultural, economic, and social factors influencing birth rates globally. Understanding these dynamics is essential for policymakers seeking to implement effective strategies for population management and development.