SQL ASSIGNMENT

Qn: 1

What is the midyear population for both males and females, along with the overall population, based on the census data from the **bigquery-public-data.census_bureau_international** dataset for the years 2020, 2025, and 2030?

Query:

```sql
SELECT
  A.midyear_population_female,
  A.midyear_population_female,
  B.population
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_5yr_age_sex` AS A
LEFT JOIN
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific` AS B
ON
  A.year=B.year
WHERE
  A.year IN (2020,
    2025,
    2030)
```

Result:

Query results

Row	midyear_population	midyear_population	population
1	455800	455800	396771
2	455800	455800	664
3	455800	455800	22
4	455800	455800	28258
5	455800	455800	18452
6	455800	455800	16

Qn: 2

Retrieve the common records of midyear population for each country and year from both datasets.

Query:

```sql
SELECT
```

```
  a.country_name,
  a.year,
  a.age,
  b.midyear_population_male,
  b.midyear_population_female
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific` a
INNER JOIN
  `bigquery-public-data.census_bureau_international.midyear_population_5yr_age_sex` b
ON
  a.country_name = b.country_name AND a.year = b.year;
```

Result:

Query results

Row	country_name ▼	year ▼	age ▼	midyear_population	midyear_population
1	Nauru	2001	0	638	556
2	Nauru	2001	1	445	407
3	Nauru	2001	1	360	365
4	Nauru	2001	2	638	556
5	Nauru	2001	2	167	166
6	Nauru	2001	2	1	1

Qn: 3

Sort down the average age for each country and year, excluding ages with low population counts.

Query:

```
SELECT
  country_name,
  year,
  AVG(age) AS average_age
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
WHERE
  population > 1000
GROUP BY
  country_name, year;
```

Result:

Qn:4

Sort down the average age for each country and year, excluding ages with low population counts.

Query:

```sql
SELECT
  country_name,
  year,
  AVG(age) AS average_age
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
WHERE
  population > 1000
GROUP BY
  country_name, year;
```

Result:

Qn:5

Find the countries where the percentage of the population in the age group 25-34 is higher than the global average.

Query:

```sql
WITH GlobalAverage AS (
  SELECT
    AVG(population) AS global_avg_25_34
  FROM
    `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
  WHERE
    age >= 25 AND age <= 34
)
SELECT
  country_name,
  AVG(population) AS avg_25_34
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
WHERE
  age >= 25 AND age <= 34
GROUP BY
  country_name
HAVING
  avg_25_34 > (SELECT global_avg_25_34 FROM GlobalAverage);
```

Results:

Query results

	JOB INFORMATION	RESULTS	CHART	PREVIEW

Row	country_name	avg_25_34
1	Pakistan	1539276.542857...
2	Poland	242394.5774193...
3	Philippines	787212.9922535...
4	Russia	959686.3766129...
5	Saudi Arabia	287779.7042372...
6	South Africa	439124.9446969...

Qn:6

Retrieve the total midyear population for each country, including data from the midyear_population_5yr_age_sex dataset where available. If not available, display "N/A".

Query:

```sql
SELECT
  a.country_name,
  a.year,
  (SUM(a.population), 'N/A') AS total_population_agespecific,
```

```
  (SUM(b.midyear_population), 'N/A') AS total_population_5yr_age_sex
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific` a
LEFT JOIN
  `bigquery-public-data.census_bureau_international.midyear_population_5yr_age_sex` b
ON
  a.country_name = b.country_name AND a.year = b.year
GROUP BY
  a.country_name, a.year;
```

Results:

Row	country_name	year	total..._field_1	total_population_a..._field_2	total..._field_1	total_population_5..._field_2
1	Nauru	2001	217580	N/A	3995560	N/A
2	Nauru	2002	218152	N/A	4006064	N/A
3	Nauru	2003	218372	N/A	4010104	N/A
4	Nauru	2004	219318	N/A	4027476	N/A
5	Nauru	2005	220308	N/A	4045656	N/A
6	Nauru	2006	210430	N/A	3864260	N/A

Qn:7

Find the countries with the highest total midyear population for a specific age group by grouping by country name and arrange them in descending order according to the population.

Query:

```
SELECT
  country_name,
  SUM(midyear_population) AS total_population
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_5yr_age_sex`
GROUP BY
  country_name
ORDER BY
  total_population DESC
LIMIT 10;
```

Results:

Qn:8

Identify the countries with the highest total midyear population across all years.

Query:

```sql
SELECT
  country_name,
  SUM(population) AS total_population
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
GROUP BY
  country_name
ORDER BY
  total_population DESC
LIMIT 10;
```

Results:

Qn:9

Find the average age for each country and year where the age is specified.

Query:

```sql
SELECT
  country_name,
  year,
  AVG(age) AS average_age
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
WHERE
  age IS NOT NULL
GROUP BY
  country_name, year;
```

Results:

Query results

	JOB INFORMATION	RESULTS	CHART PREVIEW	JSON

Row	country_name ▼	year ▼	average_age ▼
1	Nauru	2001	50.00000000000…
2	Nauru	2002	50.00000000000…
3	Nauru	2003	50.00000000000…
4	Nauru	2004	50.00000000000…
5	Nauru	2005	50.00000000000…
6	Nauru	2006	50.00000000000…

Qn:10

Retrieve the total midyear population for each country in a specific year.

Query:

```sql
SELECT
  country_name,
  year,
  SUM(population) AS total_population
FROM
  `bigquery-public-data.census_bureau_international.midyear_population_agespecific`
GROUP BY
  country_name, year;
```

Results:

Query results

Row	country_name	year	total_population
1	Nauru	2001	9890
2	Nauru	2002	9916
3	Nauru	2003	9926
4	Nauru	2004	9969
5	Nauru	2005	10014
6	Nauru	2006	9565