

## Aggregation

The following query selects the average population for each country between the years 2000 and 2018 from the 'midyear\_population' table in the 'census\_bureau\_international' dataset. The resulting table is arranged in descending order.

```
bq query --use_legacy_sql=false 'SELECT
  country_name AS country,
  AVG(midyear_population) AS average_population
FROM
  `bigquery-public-data.census_bureau_international.midyear_population`
WHERE
  year >= 2000 AND year <= 2018
GROUP BY
  country
ORDER BY
  average_population DESC
LIMIT
  20'
```

Waiting on bqjob\_r95be3d17e726415\_000001662890a68f\_1 ... (1s) Current status: DONE

country	average_population
China	1.3285399873157892E9
India	1.154912377105263E9
United States	3.0594302226315784E8
Indonesia	2.3984691394736844E8
Brazil	1.930978929473684E8
Pakistan	1.8112083526315784E8
Nigeria	1.6255564478947365E8
Bangladesh	1.447749475789474E8
Russia	1.4330035963157892E8
Japan	1.2727527184210527E8
Mexico	1.1269223210526317E8
Philippines	9.1357295E7

Vietnam	8.83786184736842E7
Ethiopia	8.460339989473683E7
Germany	8.168817173684208E7
Egypt	8.064017099999999E7
Iran	7.427240431578948E7
Turkey	7.389499394736844E7
Congo (Kinshasa)	6.82958565263158E7
Thailand	6.619103463157895E7
+-----+	

In the preceding query, the fields retrieved using the SELECT command are passed through an aggregation function to give the average of the mid-year population for the years between 2000 and 2018 inclusive. In order to mix aggregated field and non-aggregated fields, we need the GROUP BY command to group the result by one or more columns, or else only a single result will be returned because of the aggregated function.

## Joins

The following query selects the average population for each country and their life expectancy for the year 2018. The data is joined from the ‘midyear\_population’ table and the ‘mortality\_life\_expectancy’ table in the ‘census\_bureau\_international’ dataset. The resulting table is grouped by country name and year and arranged in descending order.

```
bq query --use_legacy_sql=false 'SELECT
  midyearpop.country_name AS country,
  midyearpop.year AS year,
  AVG(midyearpop.midyear_population) AS population,
  AVG(mortality.life_expectancy) AS life_expectancy
FROM
  `bigquery-public-data.census_bureau_international.midyear_population` AS
  midyearpop
JOIN
  `bigquery-public-data.census_bureau_international.mortality_life_
  expectancy` AS mortality
ON
  midyearpop.country_name = mortality.country_name
```