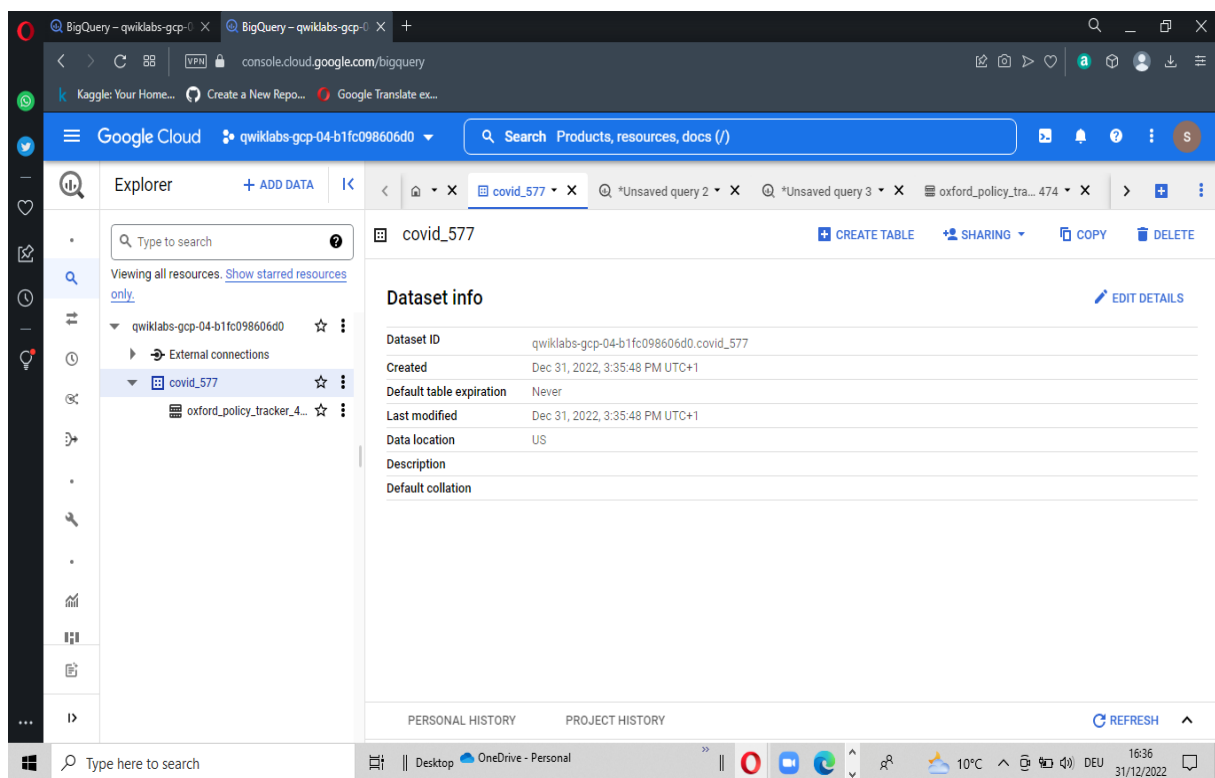


Build and Optimize Data Warehouses with BigQuery: Challenge Lab

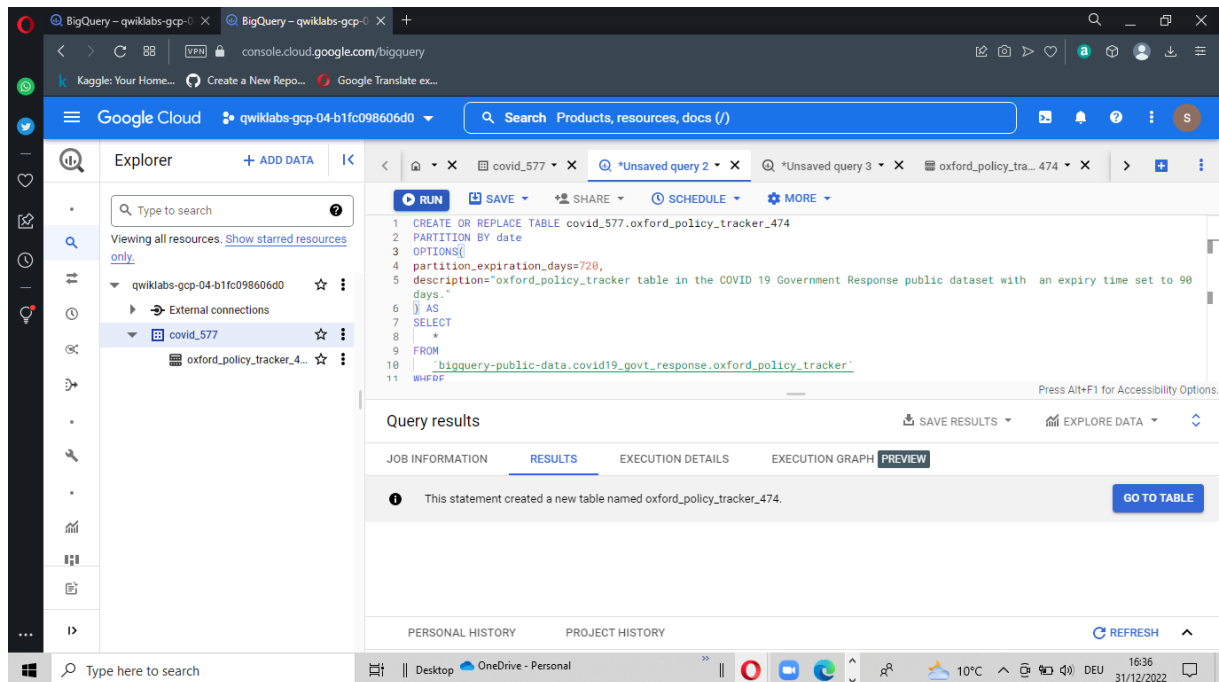
Task 1. Create a table partitioned by date

1. Create a new dataset `covid_577` and create a table `oxford_policy_tracker_474` in that dataset partitioned by date, with an expiry of 720 days. The table should initially use the schema defined for the `oxford_policy_tracker` table in the [COVID 19 Government Response public dataset](#).
2. You must also populate the table with the data from the source table for all countries except the United Kingdom (GBR), Brazil (BRA), Canada (CAN) and the United States (USA).

Solution 1.1:



Solution 1.2:

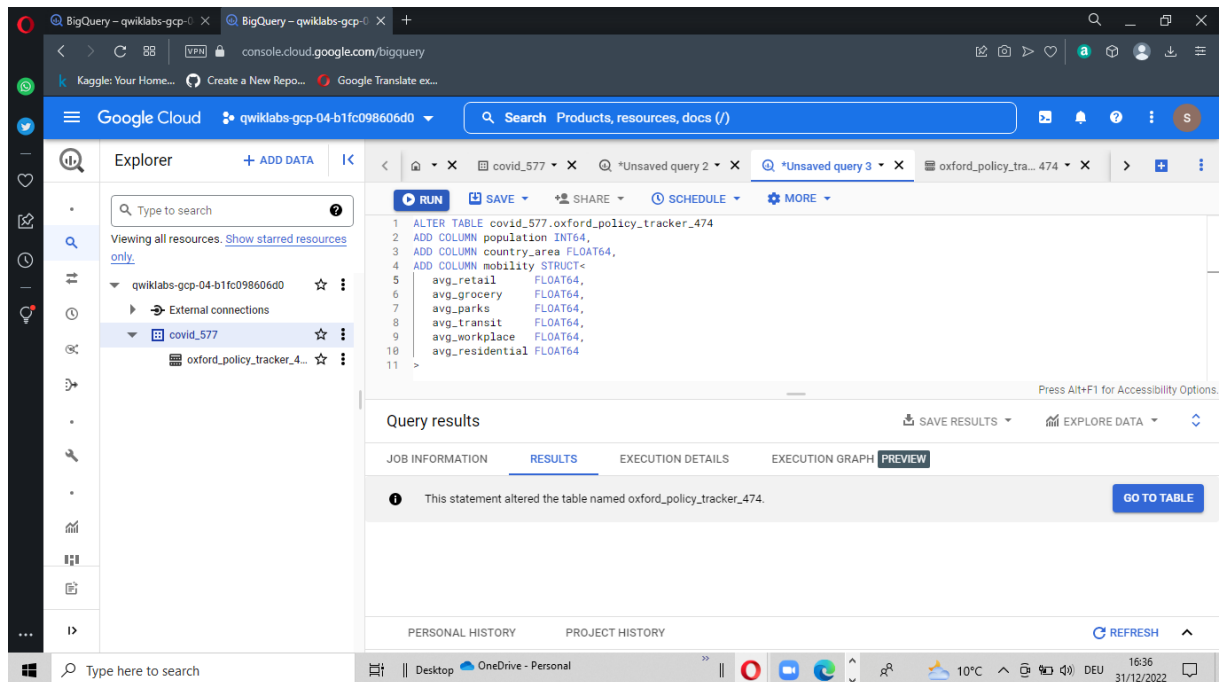


Task 2. Add new columns to your table

- Update your table to add new columns to your table with the appropriate data types to ensure alignment with the specification provided to you:

New Column Name	SQL Data Type
population	INTEGER
country_area	FLOAT
mobility	RECORD
mobility.avg_retail	FLOAT
mobility.avg_grocery	FLOAT
mobility.avg_parks	FLOAT
mobility.avg_transit	FLOAT
mobility.avg_workplace	FLOAT
mobility.avg_residential	FLOAT

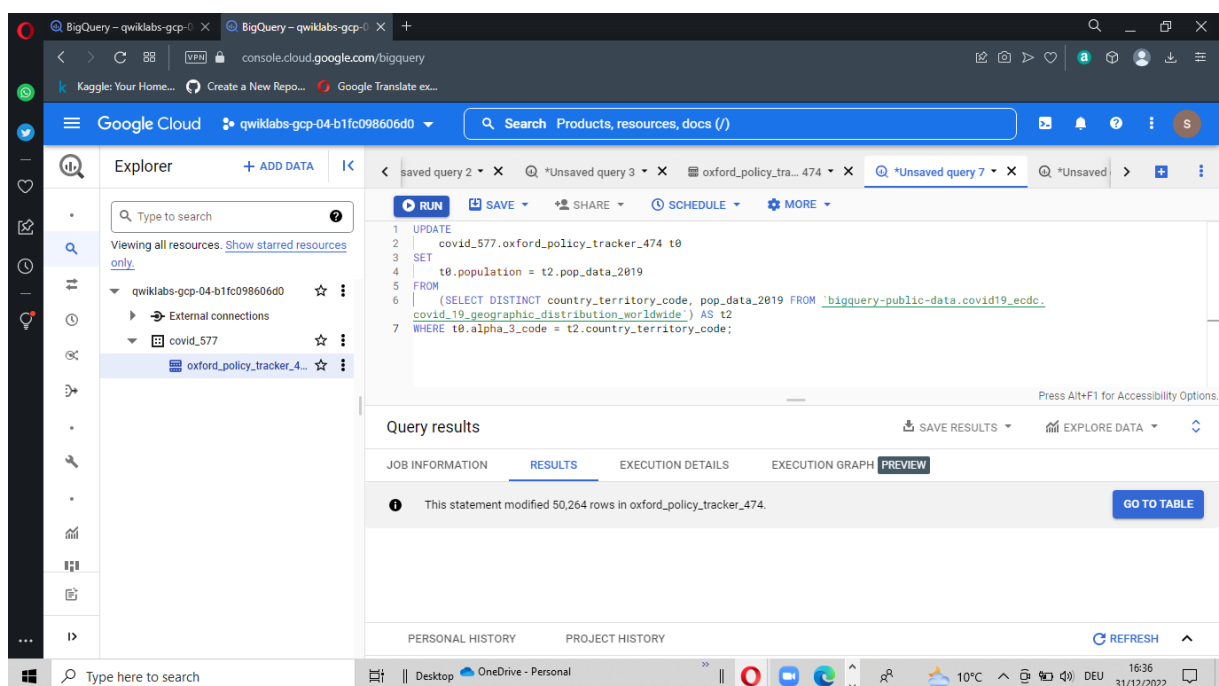
Solution:



Task 3. Add country population data to the population column

- Add the country population data to the population column in your table with covid_19_geographic_distribution_worldwide table data from the [European Center for Disease Control COVID 19 public dataset](#) table.

Solution:



Task 4. Add country area data to the country_area column

- Add the country area data to the country_area column in your table with country_names_area table data from the [Census Bureau International public dataset](#).

Solution:

The screenshot shows the Google Cloud BigQuery console interface. On the left, the Explorer pane displays the project 'qwiklabs-gcp-04-b1fc098606d0' and the table 'oxford_policy_tracker_474'. The main editor shows a SQL query:

```
1 UPDATE
2   `covid-577.oxford_policy_tracker_474` t0
3 SET
4   t0.country_area = t1.country_area
5 FROM
6   `bigquery-public-data.census_bureau_international.country_names_area` t1
7 WHERE
8   t0.country_name = t1.country_name
```

Below the query editor, the 'Query results' section shows a message: 'This statement modified 48,556 rows in oxford_policy_tracker_474.' and a 'GO TO TABLE' button. The bottom of the console shows the Windows taskbar with the date 31/12/2022 and time 16:36.

Task 5. Populate the mobility record data

- Populate the mobility record in your table with data from the [Google COVID 19 Mobility public dataset](#).

Solution:

The screenshot shows the Google Cloud BigQuery console interface. The Explorer on the left lists the project 'qwiklabs-gcp-04-b1fc098606d0' and the dataset 'covid_577', with the table 'oxford_policy_tracker_474' selected. The main editor displays an SQL query that updates the 't0' table in 'covid_577.oxford_policy_tracker_474' by setting various mobility metrics from table 't1'. The query is as follows:

```
1 UPDATE
2   `covid_577.oxford_policy_tracker_474` t0
3 SET
4   t0.mobility.avg_retail = t1.avg_retail,
5   t0.mobility.avg_grocery = t1.avg_grocery,
6   t0.mobility.avg_parks = t1.avg_parks,
7   t0.mobility.avg_transit = t1.avg_transit,
8   t0.mobility.avg_workplace = t1.avg_workplace,
9   t0.mobility.avg_residential = t1.avg_residential
10 FROM
11   / SELECT country region date
```

The 'Query results' section shows a message: 'This statement modified 29,523 rows in oxford_policy_tracker_474.' Below this, there are tabs for 'JOB INFORMATION', 'RESULTS', 'EXECUTION DETAILS', and 'EXECUTION GRAPH'. The 'RESULTS' tab is active, and a 'GO TO TABLE' button is visible. The bottom of the screen shows the Windows taskbar with the date 31/12/2022 and time 16:36.

Task 6. Query missing data in population & country_area columns

- Run a query to find the missing countries in the population and country_area data. The query should list countries that do not have any population data and countries that do not have country area information, ordered by country name. If a country has neither population or country area it must appear twice.

Solution:

The screenshot shows the Google Cloud BigQuery console interface. The Explorer on the left is the same as in the previous screenshot. The main editor displays an SQL query designed to find countries with missing population or country area data. The query is as follows:

```
1 SELECT distinct country_name
2 FROM covid_577.oxford_policy_tracker_474
3 where population is null
4 UNION ALL
5 SELECT distinct country_name
6 from covid_577.oxford_policy_tracker_474 where country_area is null
7 order by country_name asc
```

The 'Query results' section shows a table with the following data:

Row	country_name
1	Bahamas
2	Cape Verde
3	Congo
4	Czech Republic

The bottom of the screen shows the Windows taskbar with the date 31/12/2022 and time 16:37.

