**Name: Mustafa Nasikwala**

**Div: F**

**Roll No: 1**

**Experiment No.: 8**

**Title: Create a Simple Report with Inner Join and Outer Join**

**Introduction**

**In the realm of data analysis, merging tables is a crucial step that ensures comprehensive insights. Joins are operations that combine datasets based on a related key. Power BI provides several join types to accommodate different analytical needs. This report delves into the theory and practical applications of joins, using the COVID-19 World Vaccination Progress dataset and the Global COVID-19 Data by Country dataset. Both datasets contain key information about COVID-19 cases and vaccination progress, which can be combined to perform various analyses.**

**Theory of Joins**

1. **Inner Join:**
   * **Combines rows from both tables where the joining key has matching values.**
   * **Excludes non-matching rows from both tables.**
2. **Left Outer Join:**
   * **Returns all rows from the first table (left table), along with the matching rows from the second table (right table).**
   * **Non-matching rows from the left table remain, with null in the right table’s columns.**
3. **Right Outer Join:**
   * **Returns all rows from the second table (right table), along with the matching rows from the first table (left table).**
   * **Non-matching rows from the right table remain, with null in the left table’s columns.**
4. **Full Outer Join:**
   * **Combines all rows from both tables, with null where no matches exist.**

**Dataset Recap**

**COVID-19 World Vaccination Progress Dataset:**

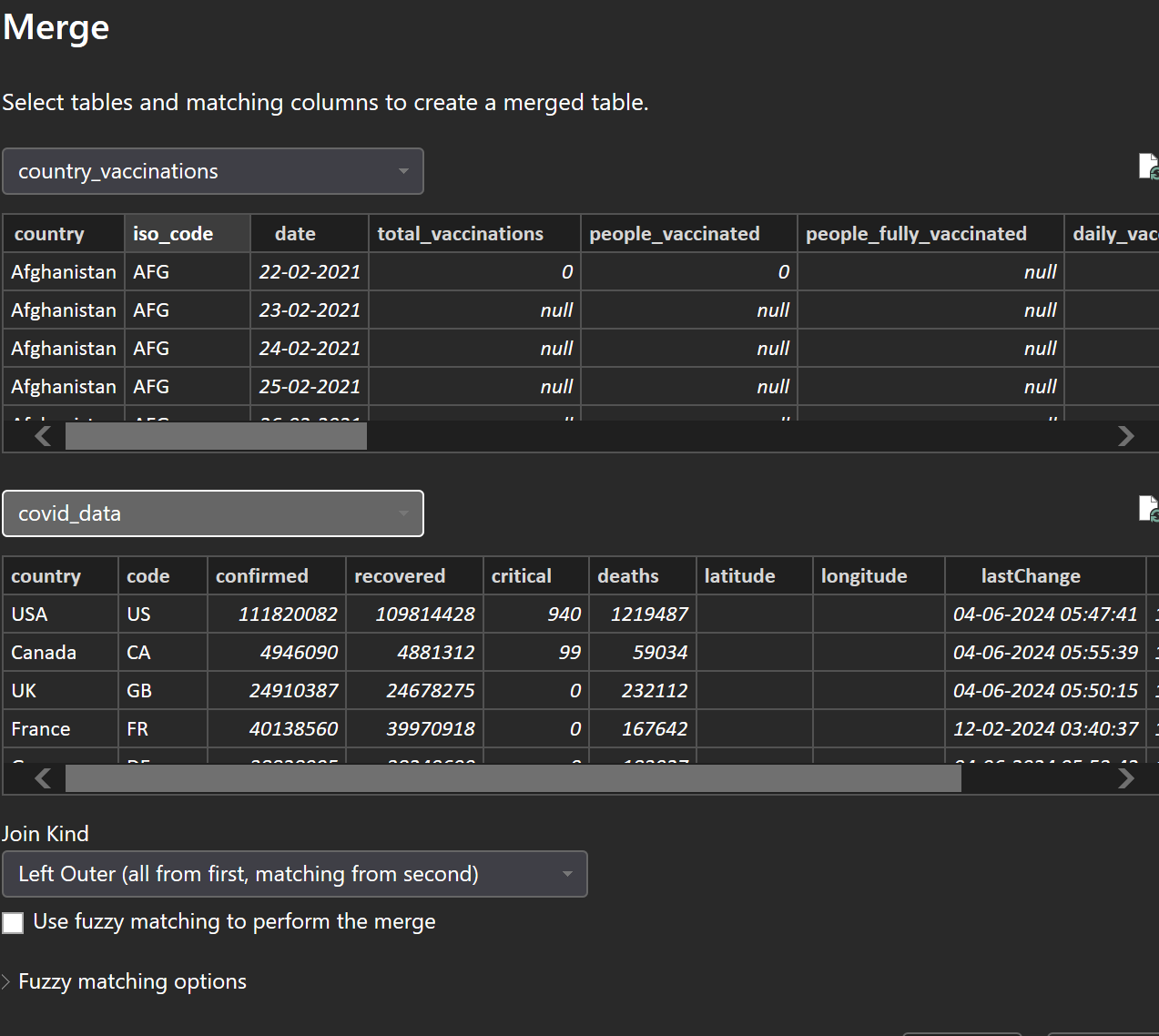
* **Country: The country name.**
* **Total Vaccinations: Total number of vaccinations administered.**
* **Date: The date of vaccination data.**

**Global COVID-19 Data by Country Dataset:**

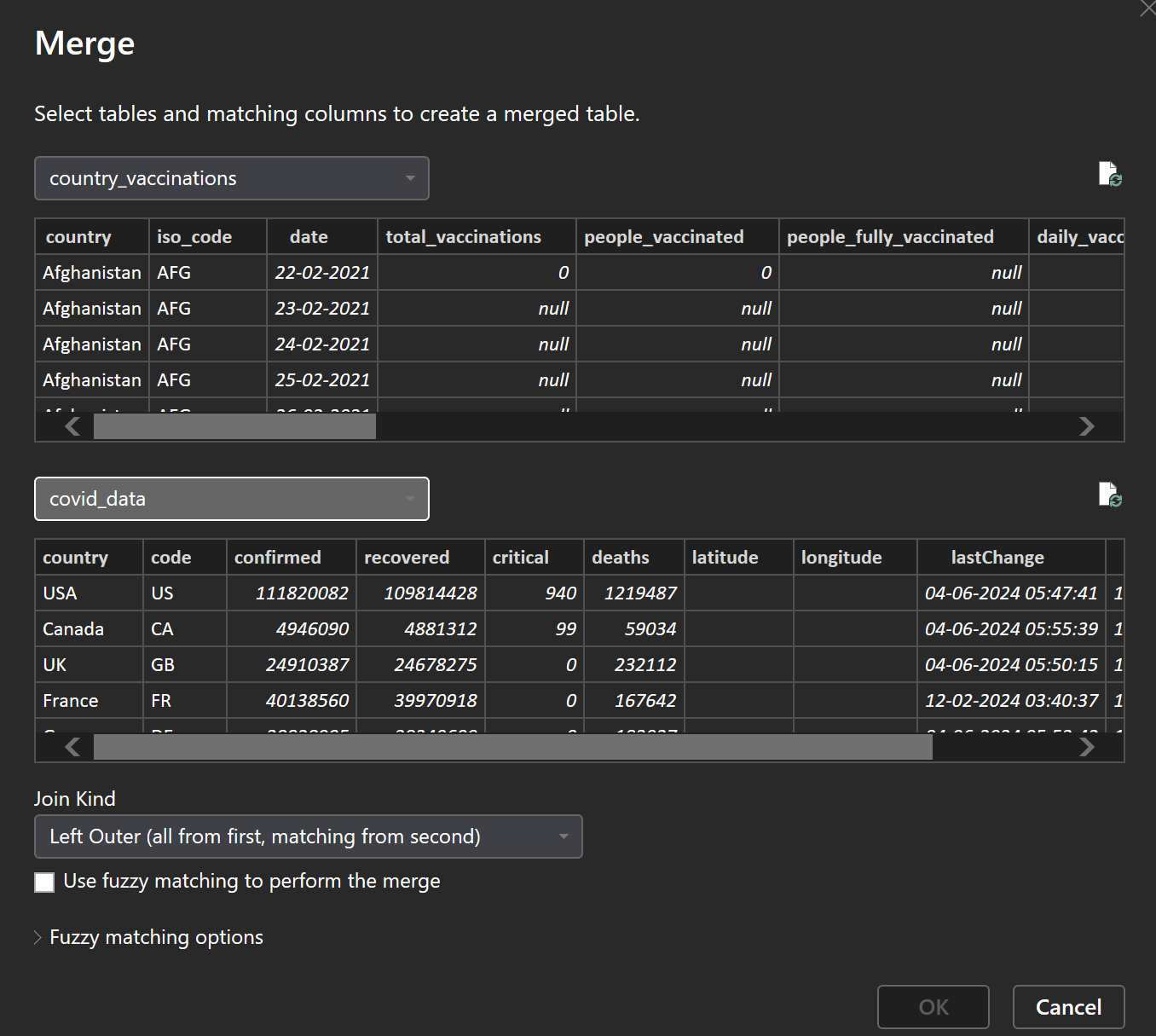
* **Country: The country name.**
* **Total Cases: The total number of COVID-19 cases.**
* **Total Deaths: The total number of deaths due to COVID-19.**
* **Date: The date of COVID-19 data.**

**Explanation of Each Join's Output**

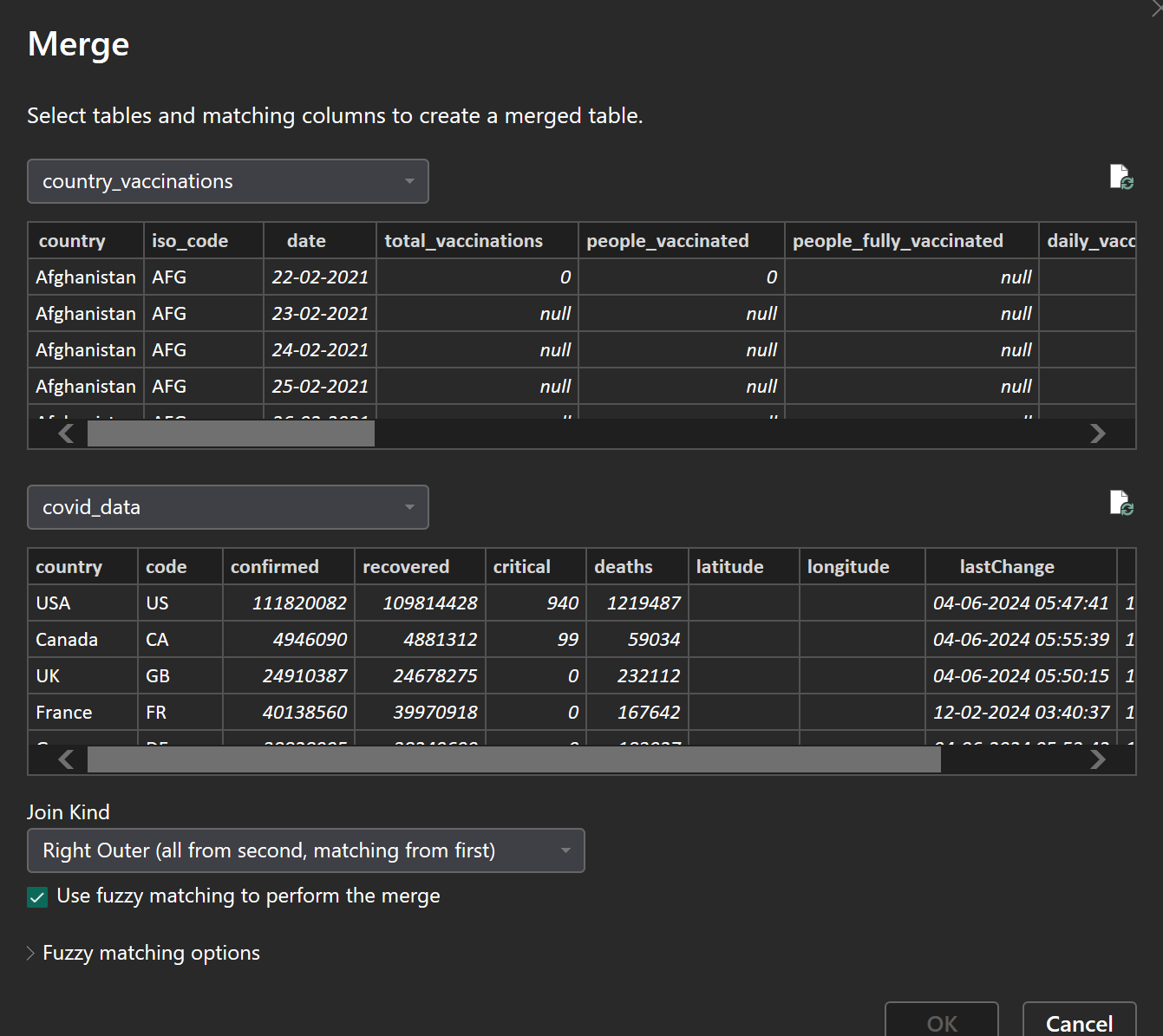
**1. Inner Join:**

* **Logic: Only rows with matching Country values in both tables are included. If there’s no data for a country in both datasets, that country is excluded from the result.**
* **Output:** 
  + **Only countries with both vaccination data and COVID-19 case data will be included.**
  + **Columns: Country, Total Cases, Total Deaths, Total Vaccinations, Date.**
* **Interpretation: The inner join provides a dataset of countries where both vaccination progress and COVID-19 data exist. This is useful for analyzing how vaccination rates correspond with cases and deaths.**

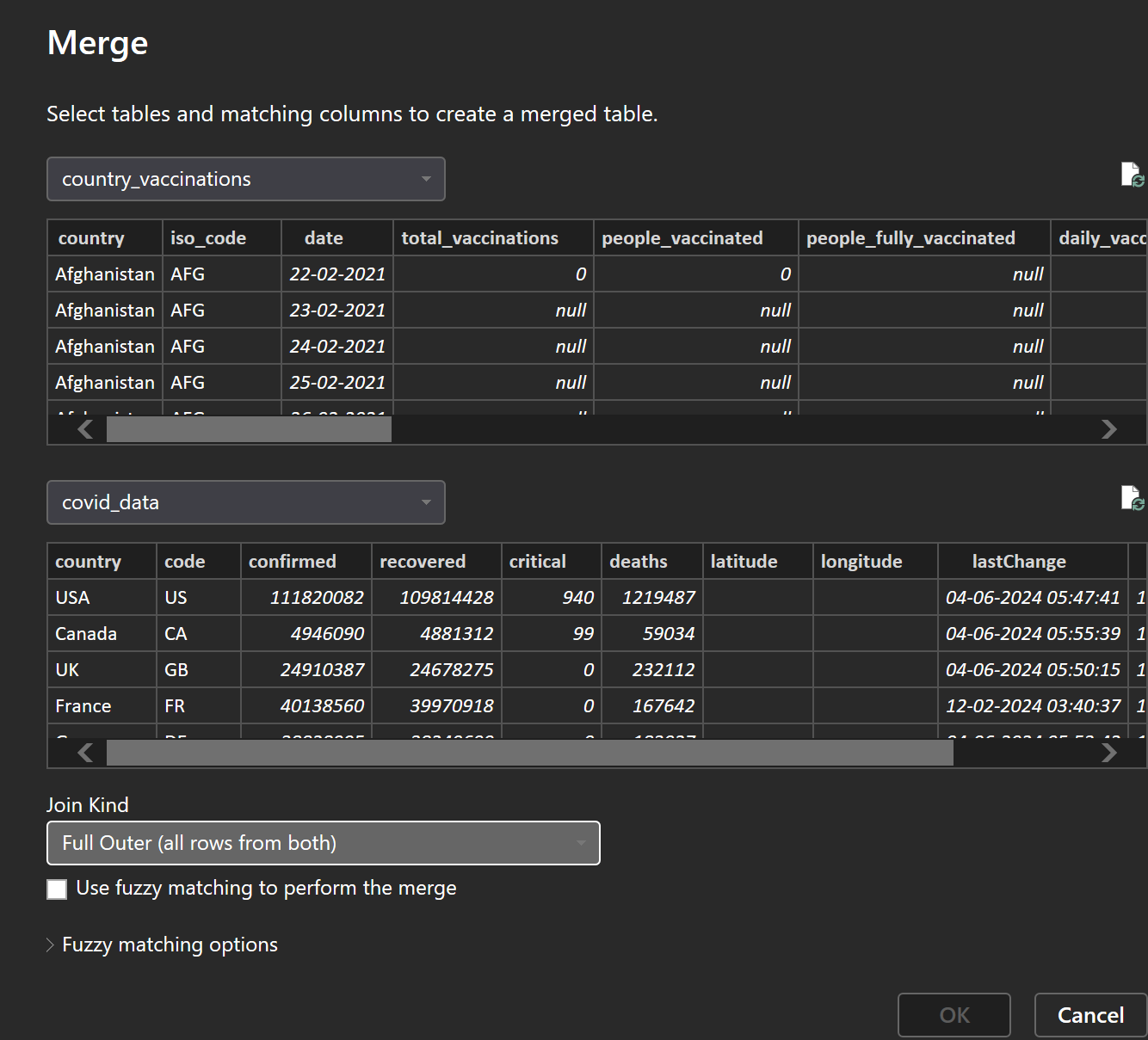
**2. Left Outer Join:**

* **Logic: All rows from the Global COVID-19 Data by Country table (left table) will be included, with matching rows from the COVID-19 World Vaccination Progress table (right table). If there’s no matching data in the vaccination table, the vaccination columns will show null.**
* **Output:**
* 
  + **All countries from the COVID-19 data will appear, including those without vaccination data.**
  + **Columns: Country, Total Cases, Total Deaths, Total Vaccinations (null if missing), Date.**
* **Interpretation: This join allows you to include all countries for which you have COVID-19 case and death data, even if vaccination data is missing. This could be helpful when you want to track COVID cases in countries that haven’t reported vaccination progress.**

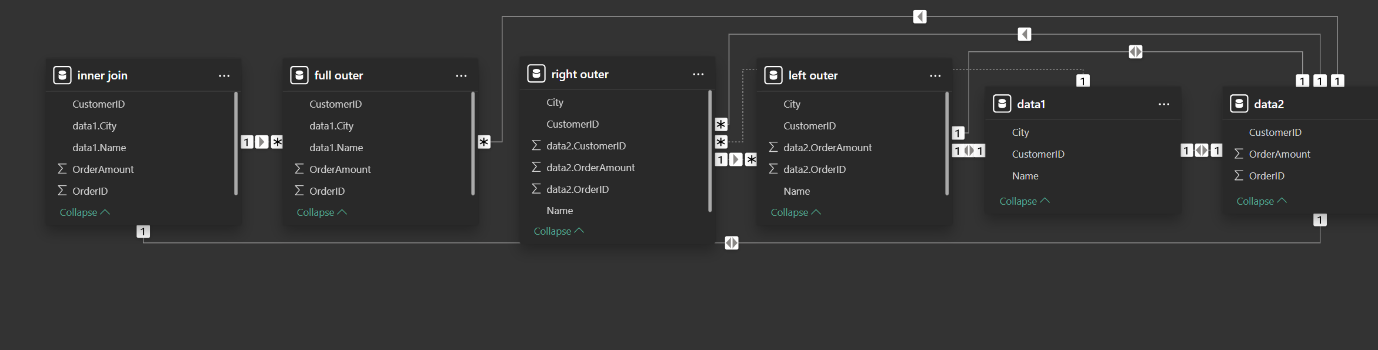
**3. Right Outer Join:**

* **Logic: All rows from the COVID-19 World Vaccination Progress table (right table) will be included, with matching rows from the Global COVID-19 Data by Country table (left table). If there’s no matching data in the COVID-19 data table, the COVID-19 columns will show null.**
* **Output:**
* 
  + **All countries from the vaccination dataset will appear, including those without COVID-19 case data.**
  + **Columns: Country, Total Cases (null if missing), Total Deaths (null if missing), Total Vaccinations, Date.**
* **Interpretation: This join ensures that all countries with vaccination data are included in the analysis, even if no COVID-19 case data is available for them. This can be useful for tracking countries that have vaccination data but no recent COVID case reporting.**

**4. Full Outer Join:**

* **Logic: Combines all rows from both tables, with null where no matches exist. This join returns all countries from both datasets, showing null for any missing data.**
* **Output:**
* 
  + **Includes all countries from both datasets. If a country is missing data from one of the tables, the missing columns will be filled with null.**
  + **Columns: Country, Total Cases (null if missing), Total Deaths (null if missing), Total Vaccinations (null if missing), Date.**
* **Interpretation: A full outer join provides the most comprehensive dataset, including every country from both datasets. If a country has data in one table but not the other, the missing data will appear as null.**

**Model View:**



**Tables:**

* **Vaccination Data (COVID-19 World Vaccination Progress):** 
  + **Contains columns such as Country, Total Vaccinations, and Date.**
* **COVID-19 Data (Global COVID-19 Data by Country):** 
  + **Contains columns such as Country, Total Cases, Total Deaths, and Date.**

**Joins:**

* **Inner Join:**
  + **Combines only matching rows between Country in both tables.**
  + **The result includes the Country, Total Cases, Total Deaths, Total Vaccinations, and Date.**
* **Left Outer Join:**
  + **Returns all rows from the COVID-19 Data table and matching rows from the Vaccination Data table. Non-matching rows in Vaccination Data appear as null.**
* **Right Outer Join:**
  + **Returns all rows from the Vaccination Data table and matching rows from the COVID-19 Data table. Non-matching rows in COVID-19 Data appear as null.**
* **Full Outer Join:**
  + **Combines all rows from both tables, showing null for missing data from either table.**

**Conclusion**

**Joins are powerful tools for combining and analyzing data from different sources. In this report, we explored how different types of joins—Inner Join, Left Outer Join, Right Outer Join, and Full Outer Join—are applied to combine the COVID-19 World Vaccination Progress and Global COVID-19 Data by Country datasets. By understanding how each join type affects the data, we can create insightful analyses that show the relationship between vaccination progress and COVID-19 cases and deaths across countries.**