

# COVID-19 Data Platform Documentation

(Note: the application is designed for windows system only)

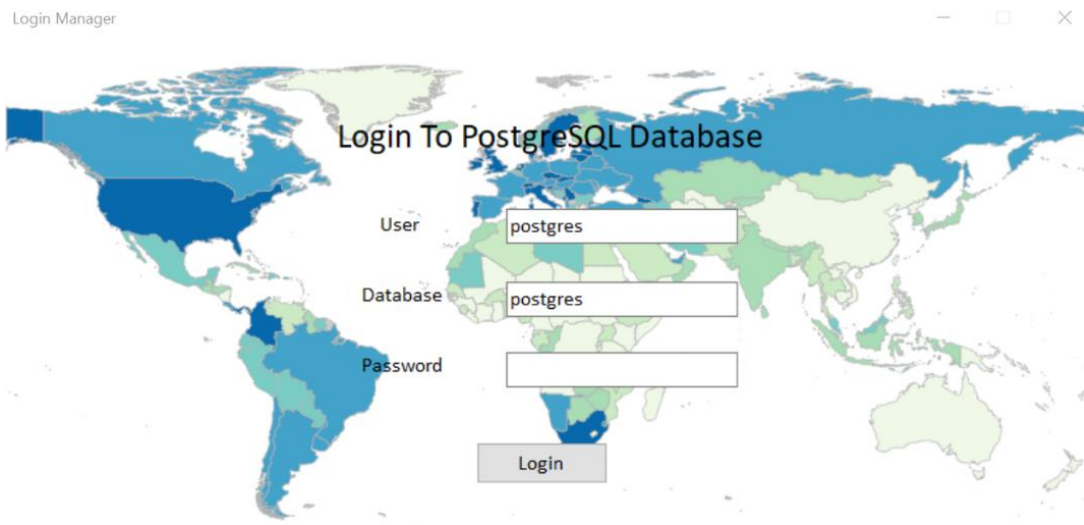
## Design

### 1. User interface

An easy-to-use graphical user interface.

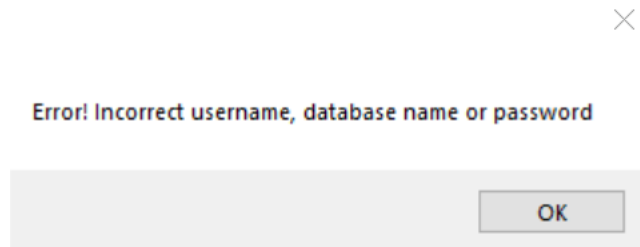
First, with PostgreSQL Server open, just a double click on “COVID-19 Data Platform.exe” in the “(Project)\COVID-19 Data Platform\bin\Release”. (“.bat” may cause serious route problem because of different windows interpretations, which has been proved in some user computers. To make sure that the application can be run successfully, the “.bat” file will only be used in the demo.)

The application contains two forms: The Login Manager and the Query Interface. Once the application is executed, the “Program.cs” in “COVID-19 Data Platform” directory will run the Login Manager form.

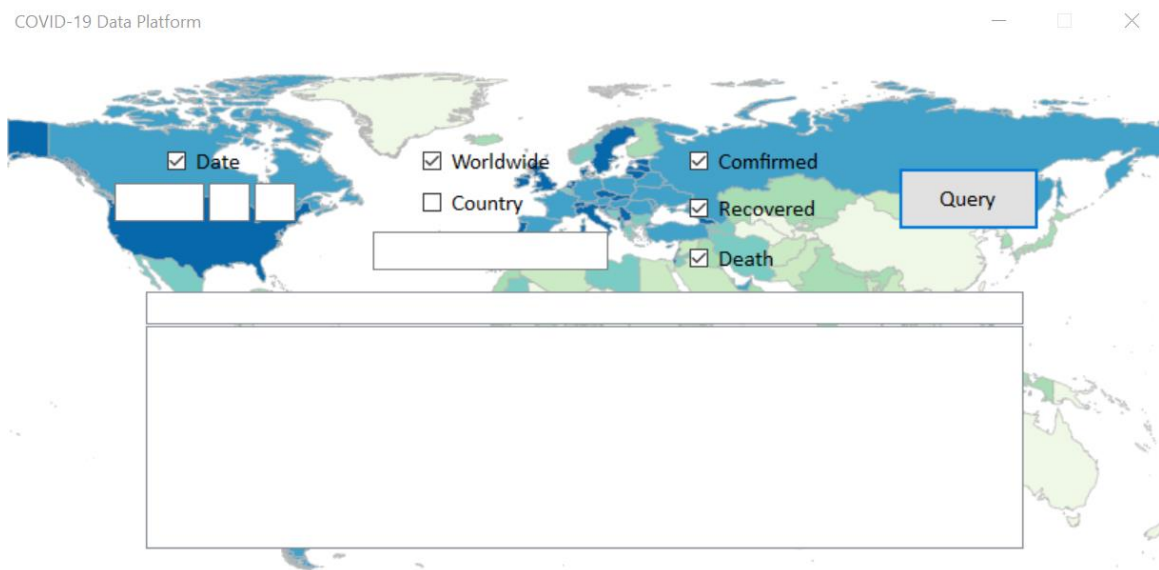


The user logs in to the PostgreSQL server by input his or her username, database name and password into the text boxes. The username and database name are “postgres” by default, which means it is better to have user “postgres” and database “postgres” in your PostgreSQL. Normally, PostgreSQL creates them during installation. The password will be hidden by characters ‘●’.

When the user clicks the “Login” button, the application tries to establish a connection with PostgreSQL (port 5432). A warning message will be shown if the username, database name or password are incorrect.



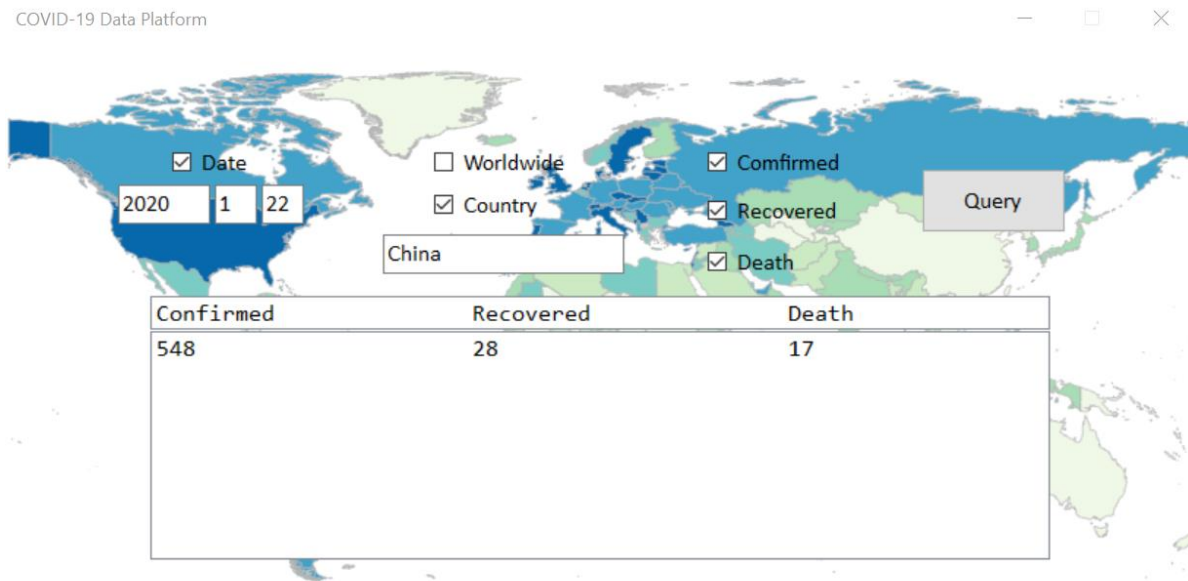
After a successful login, the Login Manager calls the Query Interface. (This process may take about 15s when the application is run on a new computer for the first time, because the application is initializing data tables at this stage.)



The Query Interface contains six check boxes, four text boxes, a button and two lists. Check boxes are for query options, text boxes are for query content, the button is for query execution and lists (one for the header and one for the data) show the results with automatic adjustment.

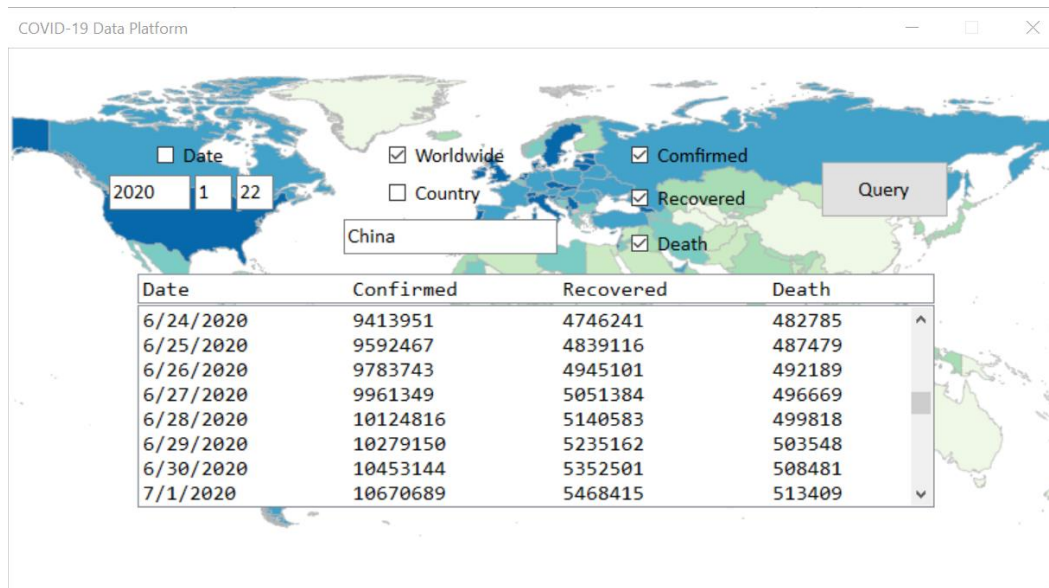
## 2. Query

The user can choose what to query by changing the check boxes' states, filling the text boxes, and finally clicking the "Query" button. Below is an example query of "how many confirmed, recovered and death cases in China on January 22<sup>nd</sup>, 2020".



Query tips:

- If the user does not fill all three date text boxes, the check box for date will be set to false will the "Query" button is clicked.
- Only one of the region options can be selected. However, the user still can choose different cases as he or she wants.
- If there are plenty of results, a scroller will be appeared at the right side of the data list. The user can scroller down to see more results with the header staying still (see the screenshot below).
- The typesetting of the query result will be arranged automatically to maintain pleasantness (see the same screenshot below, the number of columns and the space between columns are modified compared to the last query).



- e. (Bug) If there is complex data (e.g., very large data) in the text box, and the user quickly click “Query” button for several times, the “Query” button will be locked until the application restarts. This problem is caused by repeated SQL data reader operation, which can be solved by “multiple active result sets” connection. Nevertheless, current PostgreSQL does not support this function. So, range check may be the only way to partially fix the problem.

## Database

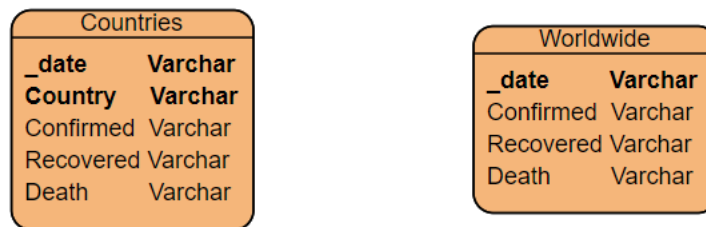
PostgreSQL is used for this windows forms application although SQL Server suits C# programs best, since we are currently learn PostgreSQL, which saves the trouble of database installation. Npgsql is installed to the application to establish PostgreSQL connection. The package can be found in the “packages” directory.

At the first-time setup, the application reads two text files modified from the “.csv” files in the “Data” directory and inserts their data into two data tables, “countries” for epidemic data for different country and “worldwide” for worldwide data, with “CREATE TABLE” and “INSERT INTO” statements. For convenient, all data types are set to “varchar”, so that the header can be included into the database. It will greatly decrease the effort needed for generating header and improve maintainability of the application while making data calculation

more difficult. In the “countries” table, the date and the country are set as composite key, and in the “worldwide” table, the date is set as primary key.

At the query stage, the application processes the input and generates a customized “SELECT” statement for each query to fetch data from the database.

#### Database's ERD



Tables “Countries” and “Worldwide” are paralleled, so there is no relationship between them, and \_date is not set as foreign key in “Countries”.

#### Input

Login Manager: username, database name, password into textboxes, and a click on the “Login” button.

Query Interface: whether to use date query, date, using worldwide or country query, (country, ) whether to include confirmed, recovered and death in the results.

For more detailed information, see “Design” part of this documentation.

#### Output

Login Manager: login failed message.

Query Interface: a header and a list of all the data satisfies the query or “No results found.” warning.

For more detailed information, see “Design” part of this documentation.