# FIT3162: Advanced Computer Science Project
# End User and Technical Guides
Project Supervisor: Dr. Soon Lay Ki

Group: MCS15
Ooi Yi Sen – 30720699
Chan Wai Han - 31555373
Nawwaf Ali - 31261949
Yeonsoo Kim - 29584612

# Table of contents

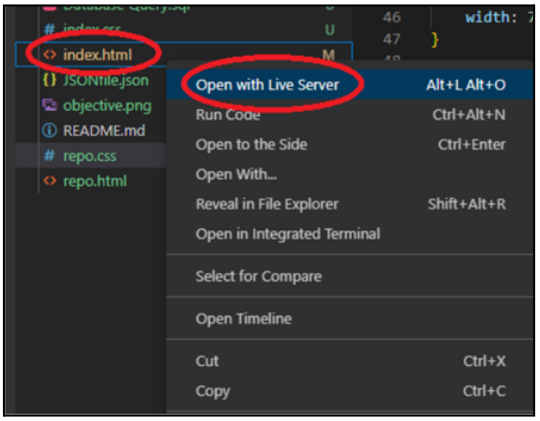
# End User Guide

## 1. Opening the software

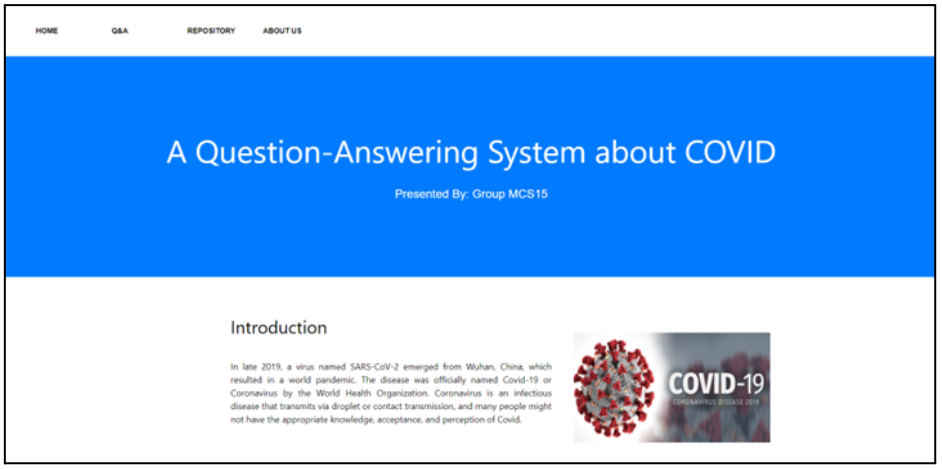
1.1. First, users are required to own a GitHub account.

1.2. Download the source code from GitHub.

1.3. Run index.html with the live server.




1.4. The user will see the Home page.



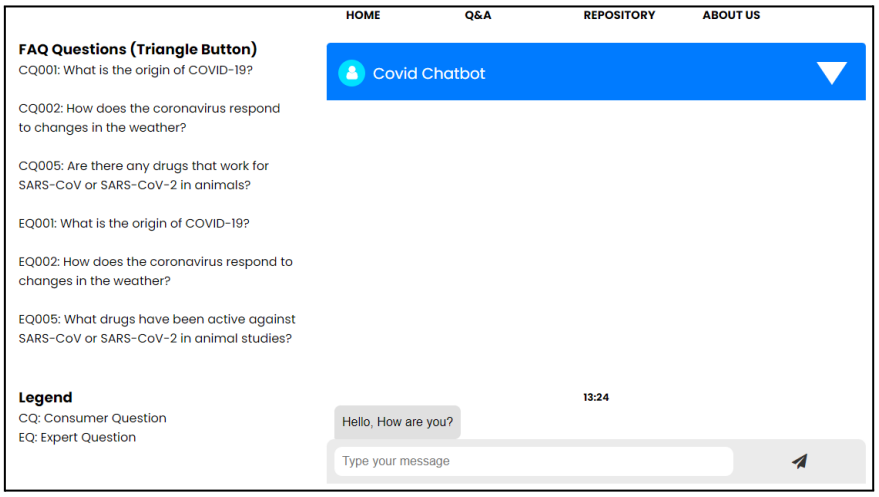

1.5. There will be a navigation bar on the top left corner.

a. Home: Home page
b. Q&A: Question and answering page
c. Repository: Self-browsing page
d. About us: Information on our project team

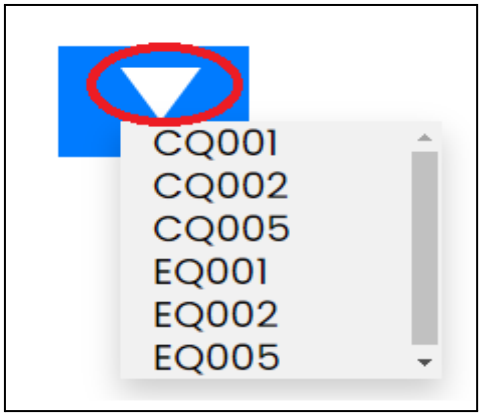
## 2. Using the Q&A page

2.1. Use the navigation bar to navigate to the Q&A page.


2.2. The user should see a page displaying the Covid Chatbot and an FAQ legend on the left side.



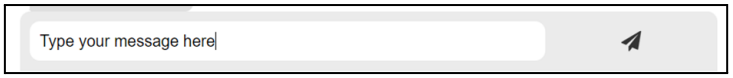

2.3. To use our Frequently Asked Questions (FAQs) functionality, click on the white triangle button at the top right corner.




2.4. The users will be allowed to choose 6 questions, 3 consumer questions, and 3 expert questions.

2.5. Each question's title is located on the left side of the page.

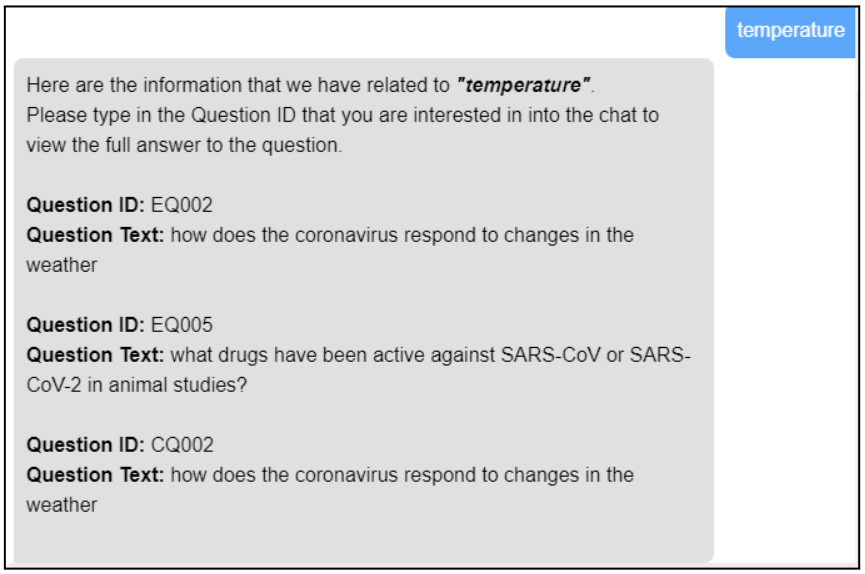
2.6. To use our chat function, type your message into the chat box.



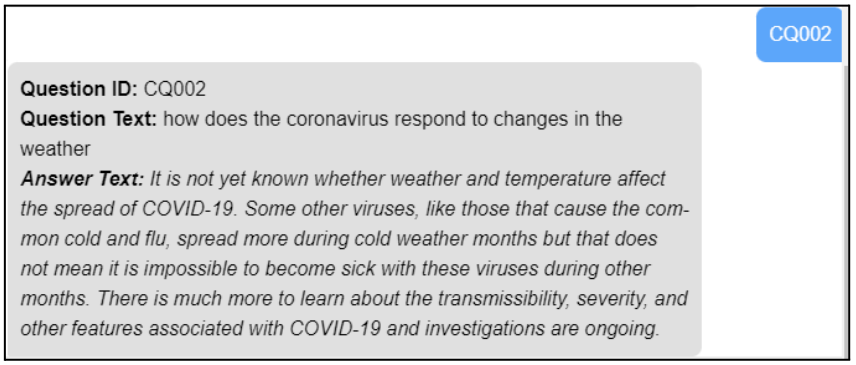

2.7. To send the message, the user can either press the enter button or click the paper plane logo on the right side.

2.8. Users are allowed to input either the keywords or the question ID.

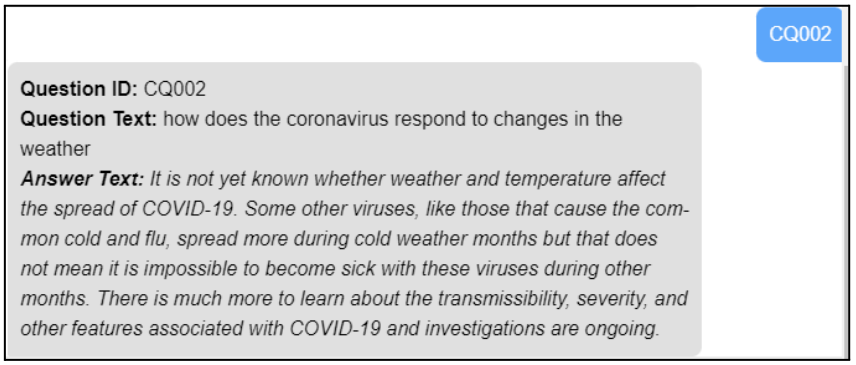
2.9. If the user chooses to input keywords, our system will then return all the possible questions.




2.9.1. Users can then choose questions they are interested in, by querying the question ID.



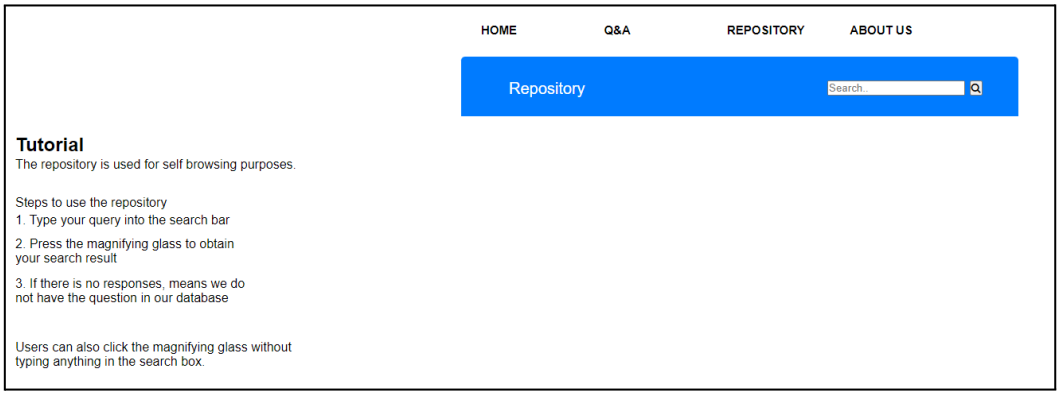

2.10. If the user chooses to input the question ID, our system will then straight return the answer.



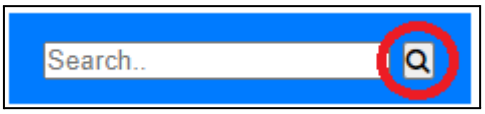
## 3. Using the Repository page

3.1. Use the navigation bar to navigate to the Repository page.

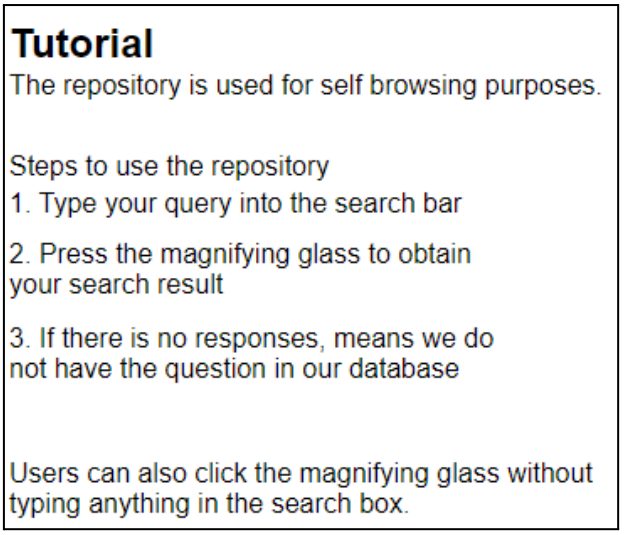

3.2. The user should see a page displaying the Repository search bar and a Repository guide on the left side.




3.3. To use the Repository, click on the magnifying glass located under the navigation bar.

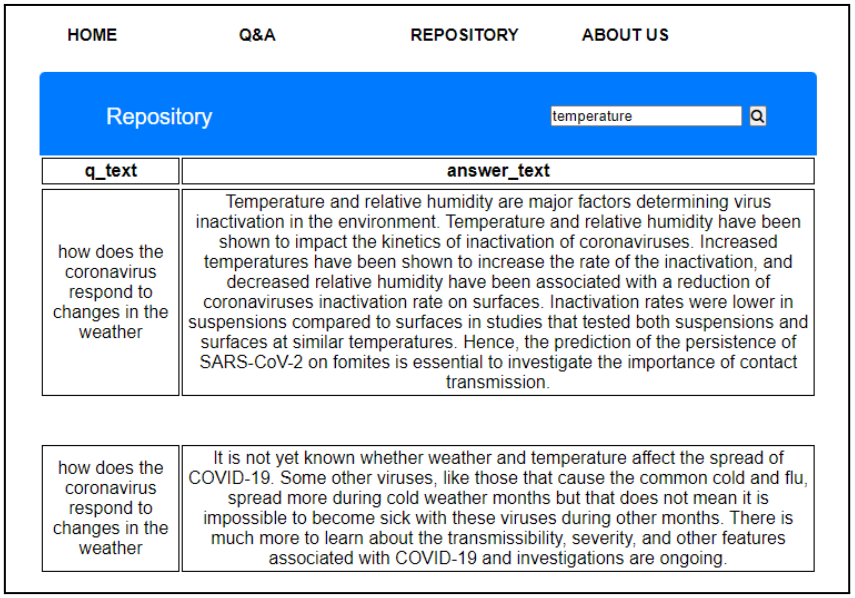



3.4. The instruction to use the Repository page is located on the left side of the page.



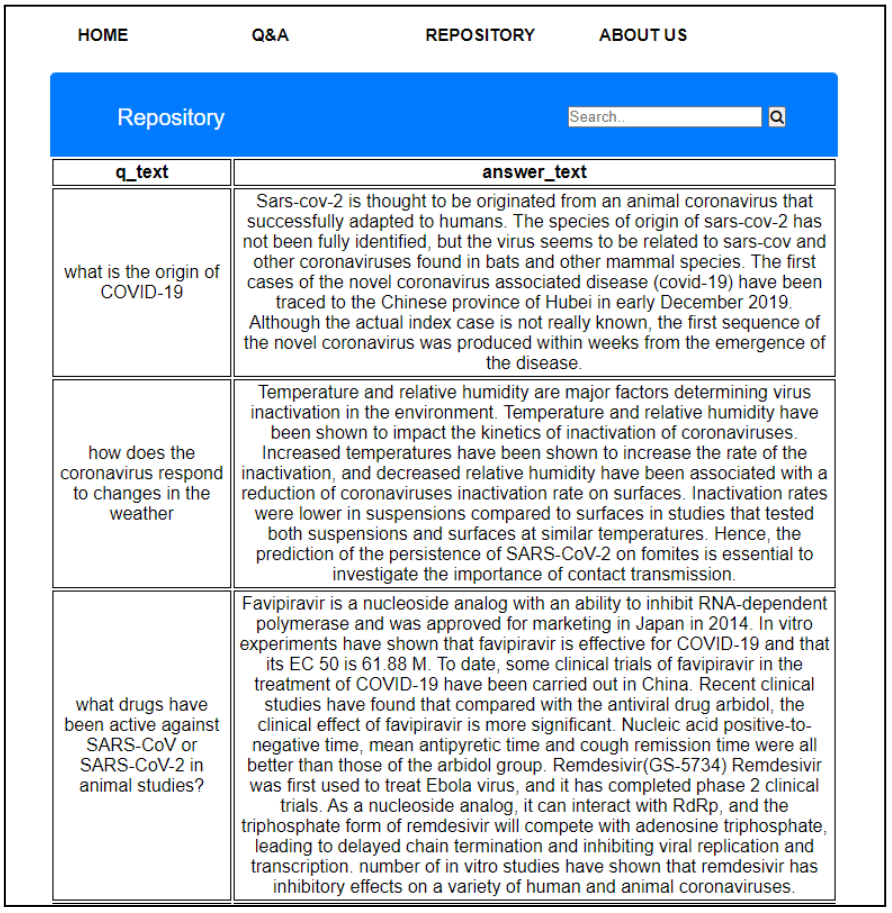
3.5. The user can click on the search icon without typing anything in the search box to get a full view of our dataset, else the user can type in keywords to query for specific data.

a. Query with keywords.




| q_text | answer_text |
|---|---|
| how does the coronavirus respond to changes in the weather | Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission. |
| how does the coronavirus respond to changes in the weather | It is not yet known whether weather and temperature affect the spread of COVID-19. Some other viruses, like those that cause the common cold and flu, spread more during cold weather months but that does not mean it is impossible to become sick with these viruses during other months. There is much more to learn about the transmissibility, severity, and other features associated with COVID-19 and investigations are ongoing. |

b. Query without keywords.

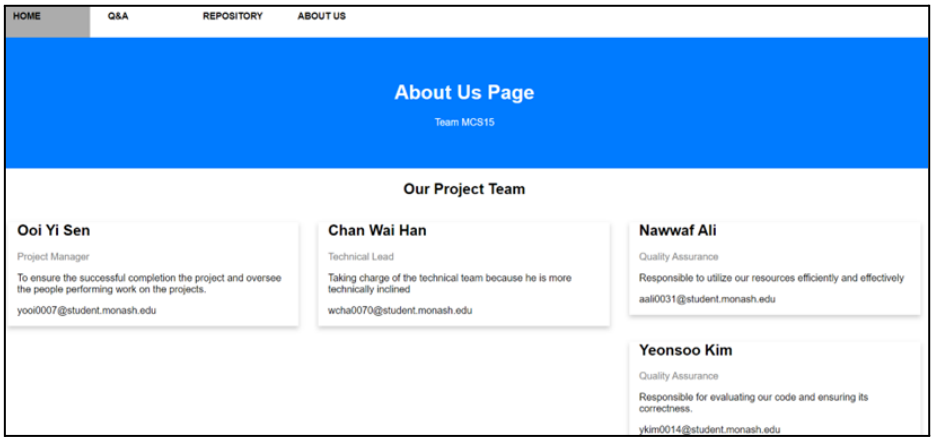



| q_text | answer_text |
|---|---|
| what is the origin of COVID-19 | Sars-cov-2 is thought to be originated from an animal coronavirus that successfully adapted to humans. The species of origin of sars-cov-2 has not been fully identified, but the virus seems to be related to sars-cov and other coronaviruses found in bats and other mammal species. The first cases of the novel coronavirus associated disease (covid-19) have been traced to the Chinese province of Hubei in early December 2019. Although the actual index case is not really known, the first sequence of the novel coronavirus was produced within weeks from the emergence of the disease. |
| how does the coronavirus respond to changes in the weather | Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission. |
| what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies? | Favipiravir is a nucleoside analog with an ability to inhibit RNA-dependent polymerase and was approved for marketing in Japan in 2014. In vitro experiments have shown that favipiravir is effective for COVID-19 and that its EC 50 is 61.88 M. To date, some clinical trials of favipiravir in the treatment of COVID-19 have been carried out in China. Recent clinical studies have found that compared with the antiviral drug arbidol, the clinical effect of favipiravir is more significant. Nucleic acid positive-to-negative time, mean antipyretic time and cough remission time were all better than those of the arbidol group. Remdesivir(GS-5734) Remdesivir was first used to treat Ebola virus, and it has completed phase 2 clinical trials. As a nucleoside analog, it can interact with RdRp, and the triphosphate form of remdesivir will compete with adenosine triphosphate, leading to delayed chain termination and inhibiting viral replication and transcription. number of in vitro studies have shown that remdesivir has inhibitory effects on a variety of human and animal coronaviruses. |

## 4. Using the About Us page

4.1. Use the navigation bar to navigate to the About Us page.

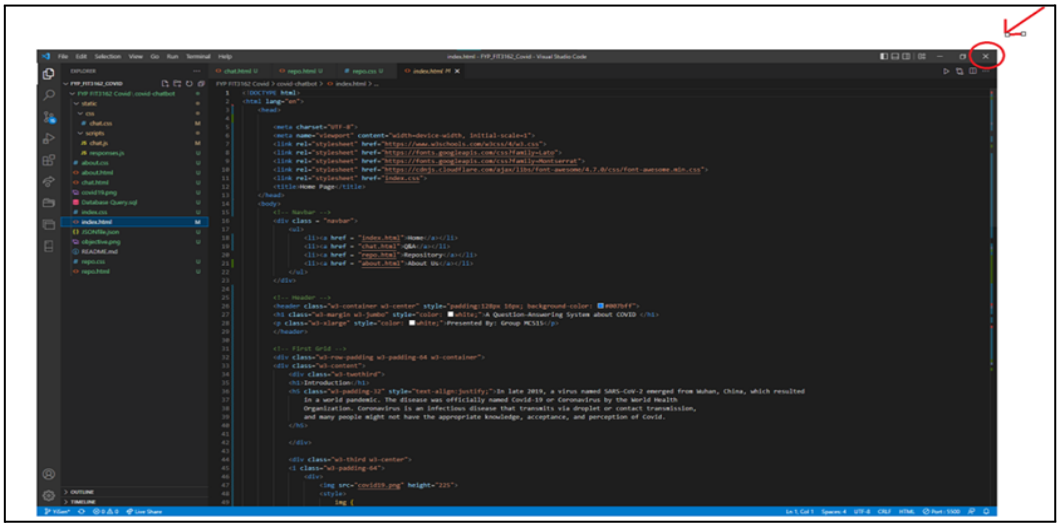

4.2. The user will see a page displaying the information of our project team.



## 5. Exiting the software

5.1. To exit the webpage just press the x button on the top right corner of the webpage.


5.2. To exit visual studio code just press the x button on the top right corner of the webpage



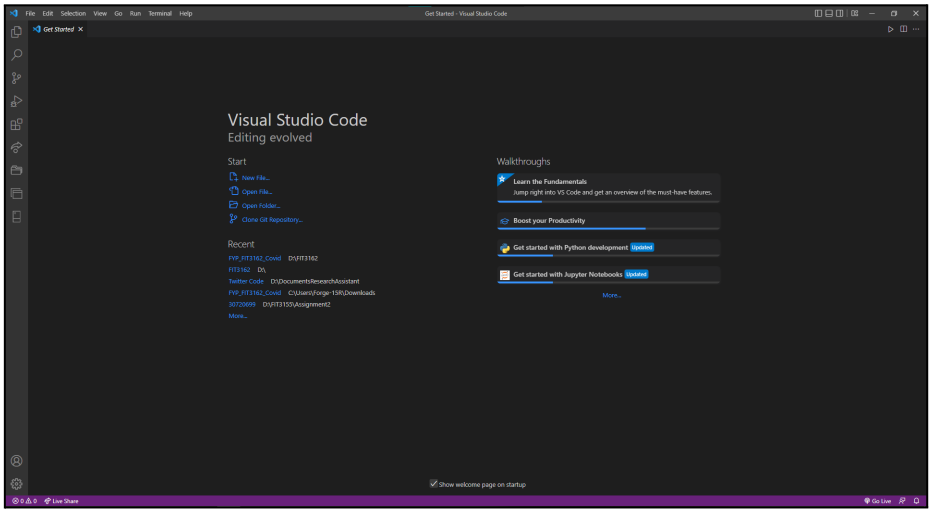
## 6. Limitations

6.1. User best query for keywords only.

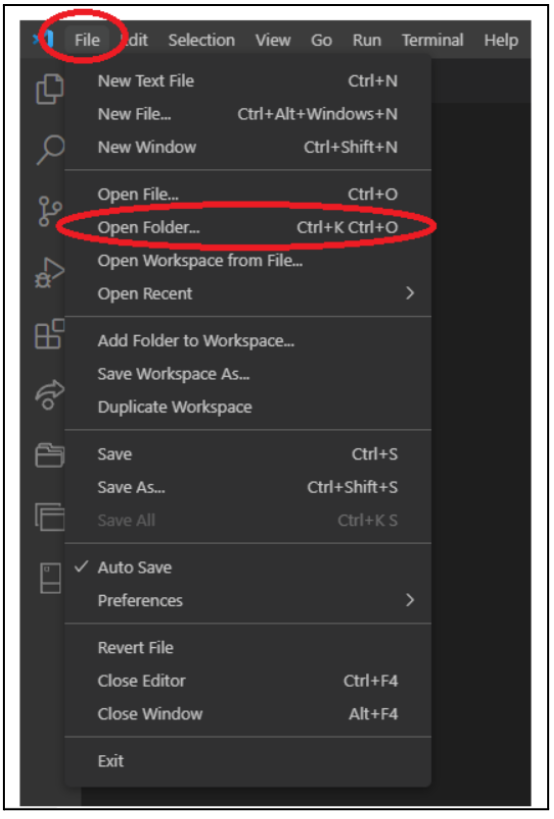
# Technical Guide

## 1. Setting up the front-end

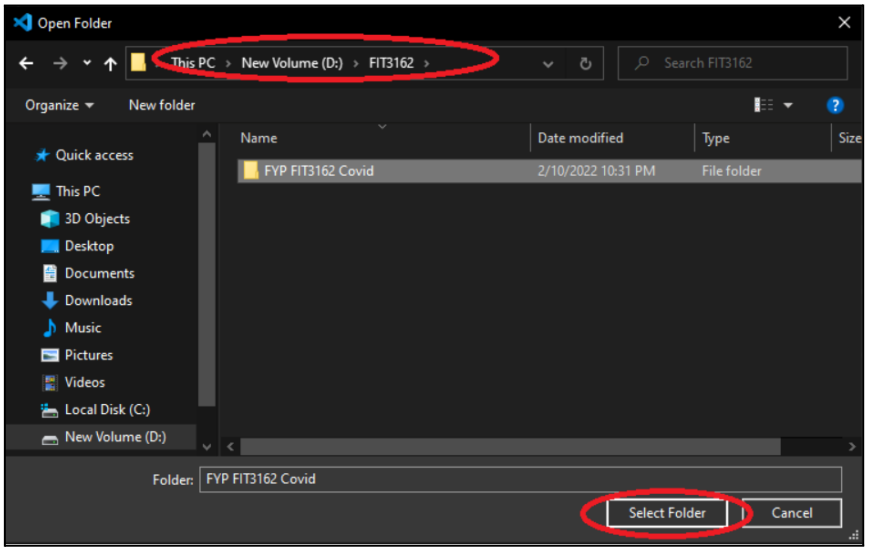
1. Download Visual Studio Code.
    1.1. Open Visual Studio Code. The user will be brought to a welcome screen.



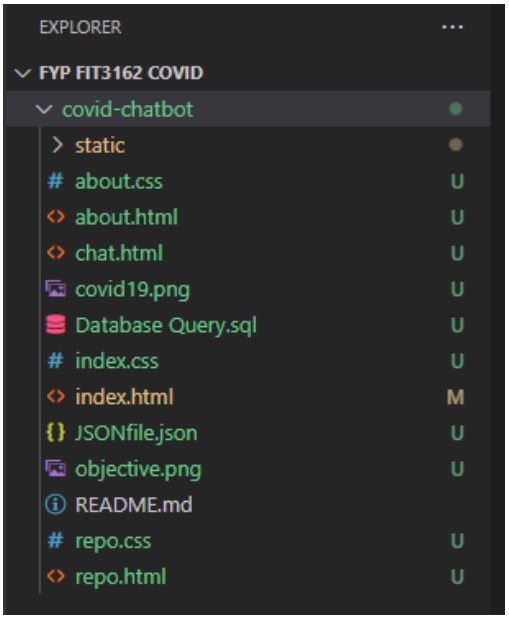

    1.2. Open the project folder by navigating to the top left corner.



1.3. Go to the directory that the folder is stored in and click select folder.



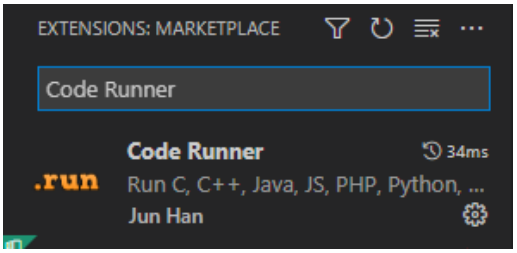

1.4. The user will see a list of files under the explorer section.



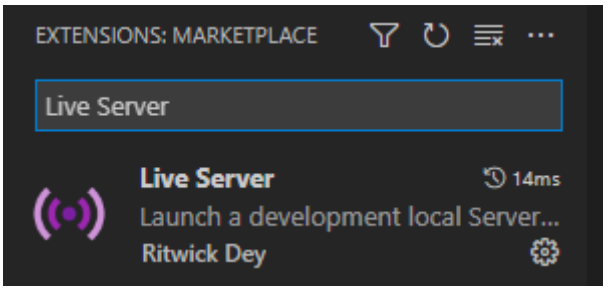

2. Install the required extensions in Visual Studio Code.

2.1. Go to the navigation bar on the left side of the screen and click on the cubes as shown below.

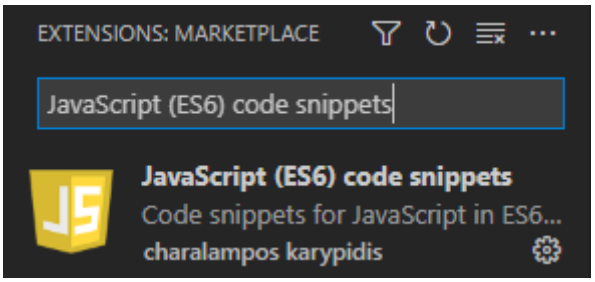
2.2. Search for Code Runner, and click install. By downloading Code Runner it will allow the user to run JavaScript codes.



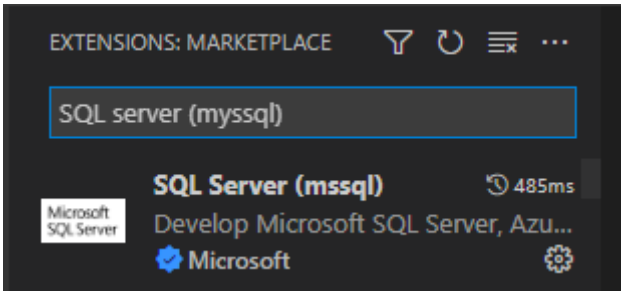
2.3. Search for Live Server, and click install. By downloading the Live Server it will allow the user to launch a development local server with live reload features and static and dynamic pages.




2.4. Search for JavaScript (ES6) code snippets, and click install. By downloading the JavaScript (ES6) code snippets it will allow visual studio code to have ES6 syntax.

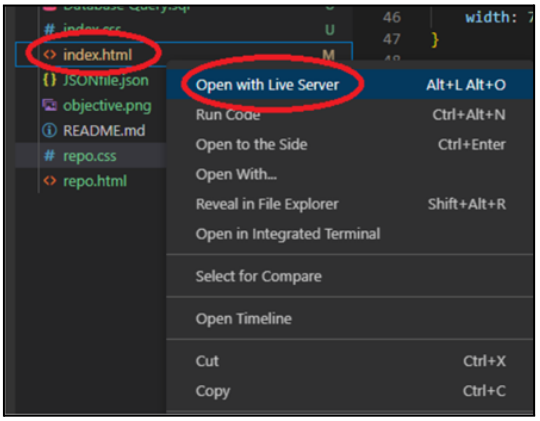



2.5. Search for SQL server (myssql), and click install. By downloading the SQL server (myssql) the user will be able to access the Azure SQL Database in Visual Studio Code.

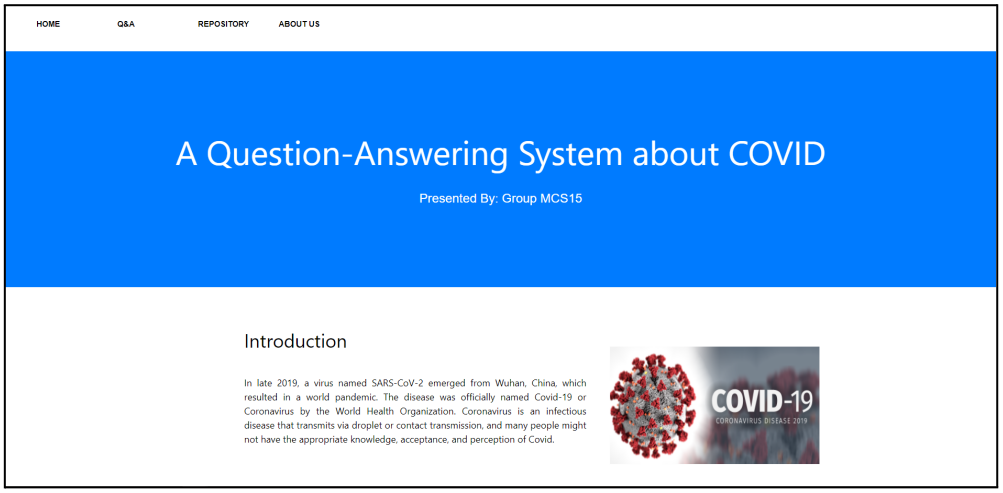



3. At this stage the user will be done setting up the software for the front-end needs.

4. Open the code in visual studio code and run index.html with the live server.




5. The page will load up and display the Home page.




5.1. At this stage the user will not have access to the database. Go to **Section 2 of connecting the front end to the back end** and setup the connection with the database.

## 2. Connecting the front end to the back end

1. Get server connection information.
   1.1. Sign in to the Azure Portal.
   1.2. Navigate to the SQL databases or SQL Managed Instances page.
   1.3. On the Overview page, review the fully qualified server name next to Server name for SQL Database or fully qualified server name next to Host for a SQL Managed Instance. To copy the server name or hostname, hover over it and select the copy icon.

2. Set language mode to SQL.
   2.1. Open a new visual studio code window.
   2.2. Press Ctrl + N. To open a new plain text file.
   2.3. Select the Plain Text in the status bar's lower right-hand corner.
   2.4. In the Select language mode drop-down menu select SQL.

3. Connect to your database.
   3.1. In visual studio code, press Ctrl+Shift+P or F1 to open the Command Palette.
   3.2. Select MS SQL:Connect and choose Enter.
   3.3. Select Create Connection Profile.
   3.4. Follow the prompts to specify the new profile's connection properties. After specifying each value, choose Enter to continue.

| Property | Suggested value | Description |
|---|---|---|
| Sever name | The qualified server name | eg: myserver2020.database.windows.net |
| Database name | mySampleDatabase | The database to connect to |
| Authentication | SQL Login | SQL Authentication |
| User name | User name | User name of the server that is used to create the server |
| Password (SQL Login) | Password | Password of the server that is used to create the server |
| Save password | Yes or No | Select yes if you do not want to enter the password each time |
| Enter a name for this profile | A profile name<br>eg: mySampleProfile | A saved profile speeds up your connection on subsequent logins |

   3.5. If successful, a notification will appear showing that the profile is created and connected.

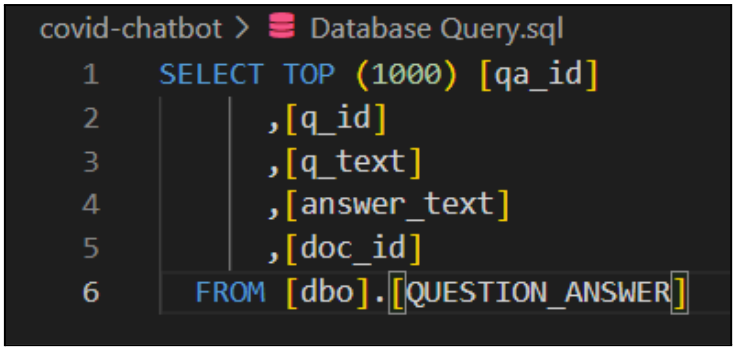
4. Query Data.
   4.1. In the editor window, past the following SQL Query to query data.

covid-chatbot > Database Query.sql

```
SELECT TOP (1000) [qa_id]
      ,[q_id]
      ,[q_text]
      ,[answer_text]
      ,[doc_id]
  FROM [dbo].[QUESTION_ANSWER]
```

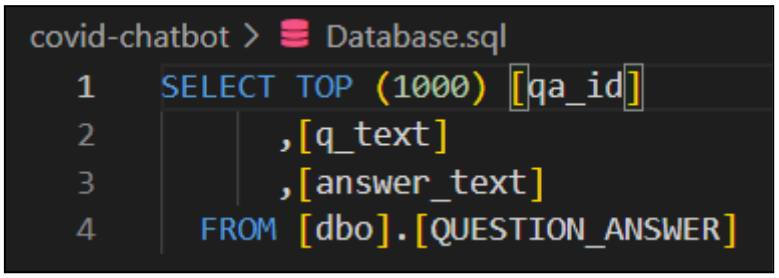
4.2. Press Ctrl+Shift+E to run the query and display the result from the question-answer table.

5. Update data for JSONfile.json.
   5.1. If there are any updates in the future where new data are added to the database, run the SQL statement.




   5.2. Download the new JSON file and name it JSONfile.json.
   5.3. Manually update the data in response.js under the variable var json.
   5.4. The JSONfile.json is for the Q&A Chatbot.

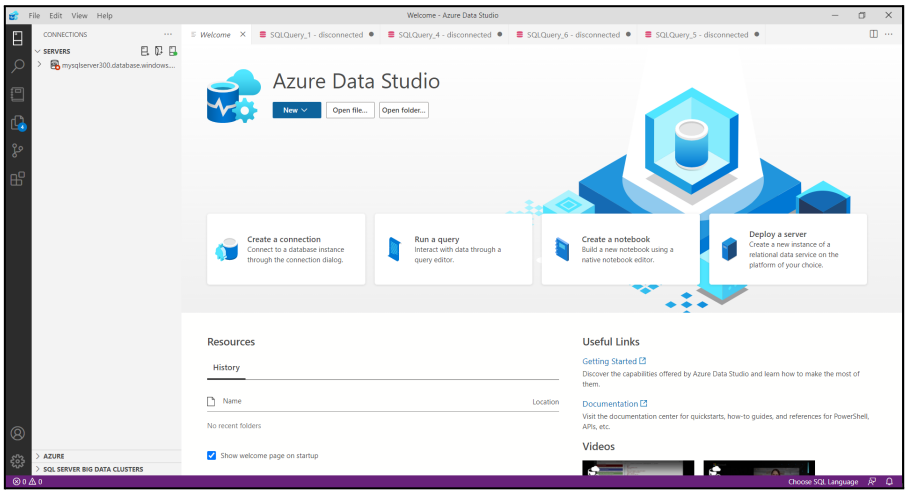
6. Update data for REPOfile.json.
   6.1. If there are any updates in the future where new data are added to the database, run the SQL statement.



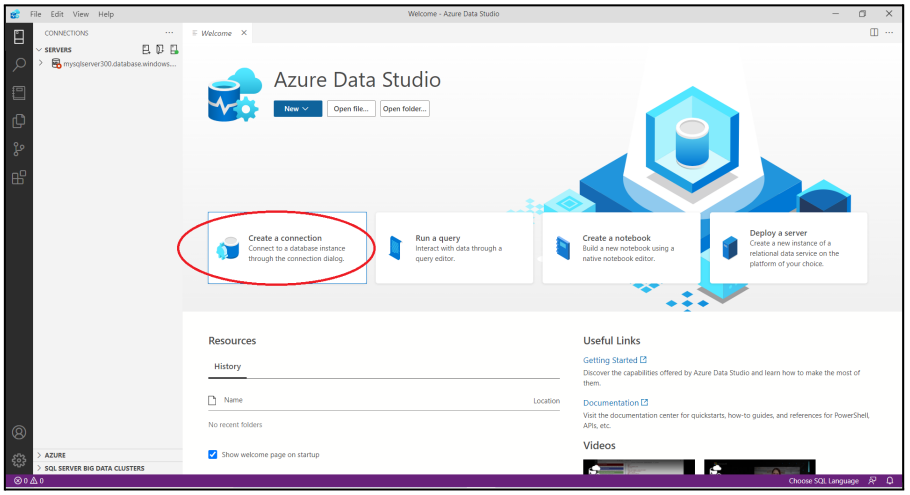

   6.2. Download the new JSON file and name it REPOfile.json.
7. After query all the data from the database, to ensure all the function works run the local host and try out the chatbot and repository.

# 3. Setting up the back-end

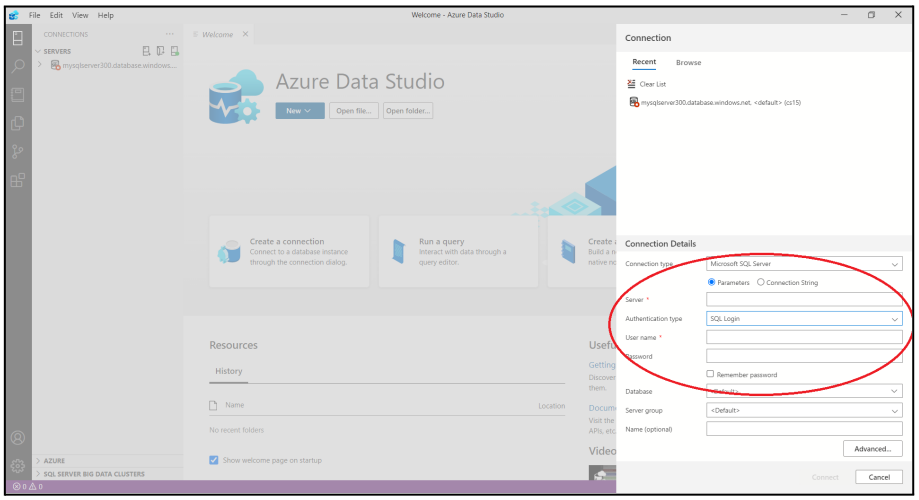
3.1. Install Azure Data Studio (assuming you have created the DB on microsoft azure portal).



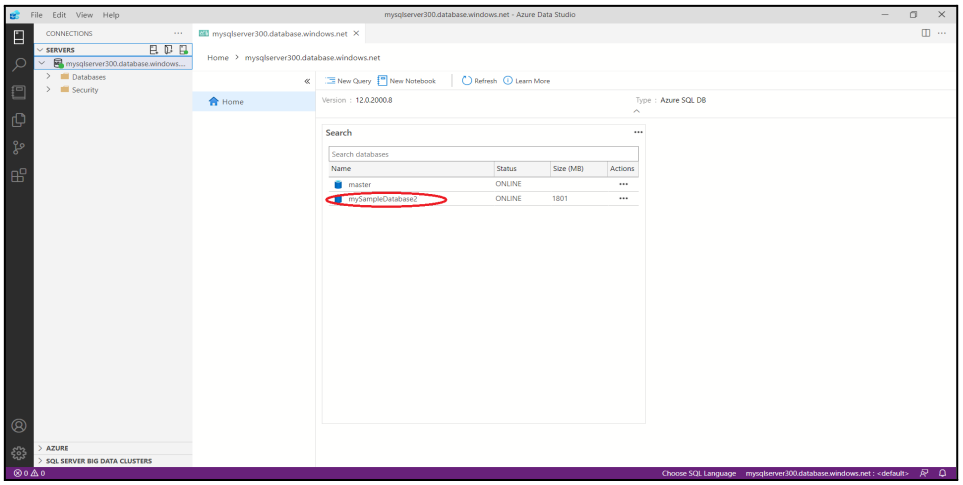

3.2. Then Create connection.




3.3. Login using your server admin information.

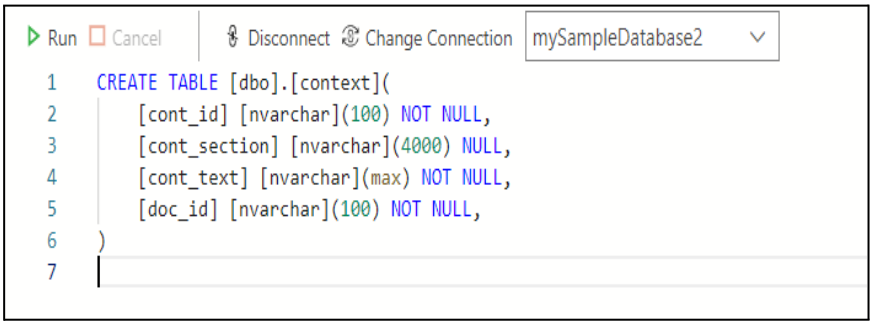


3.4. Then select the database.

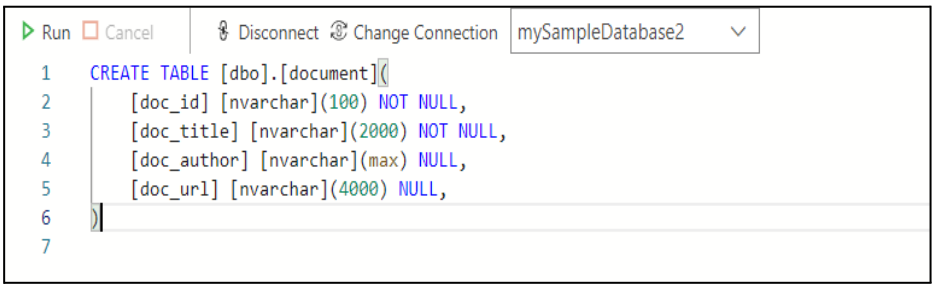



3.5. Then create a query and set up the tables. The code to set up the database tables is as follows.

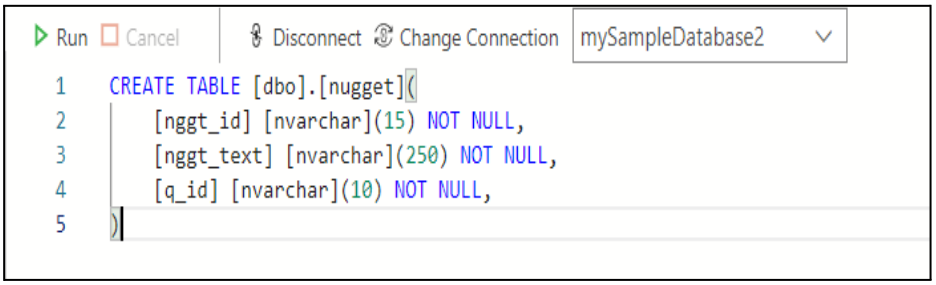
3.5.1. To set up the context table.



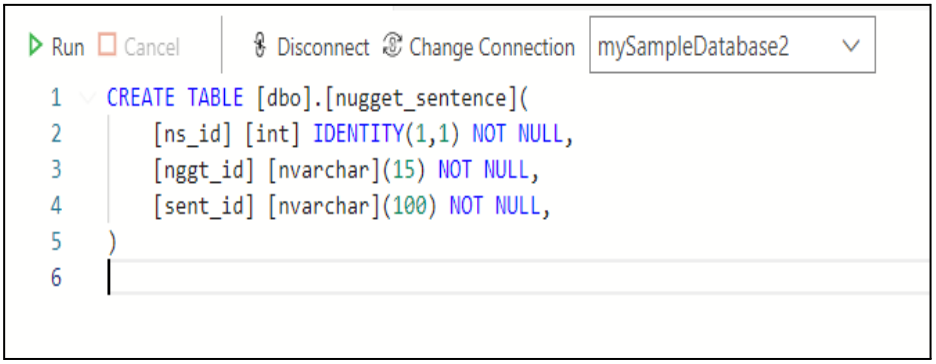

3.5.2. To set up the document table.



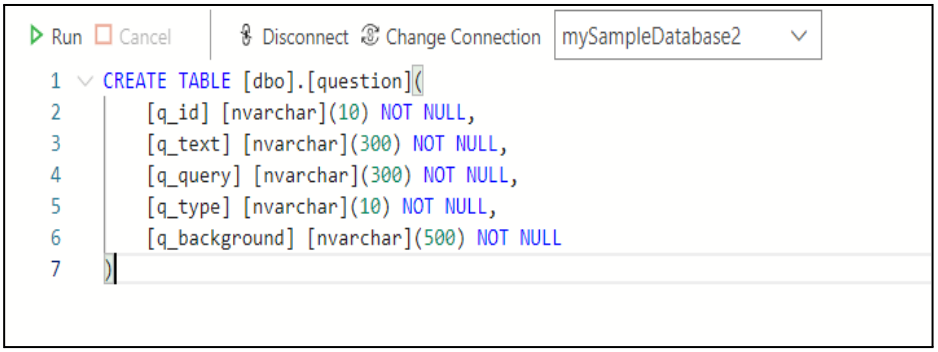

3.5.3. To set up the nugget table.



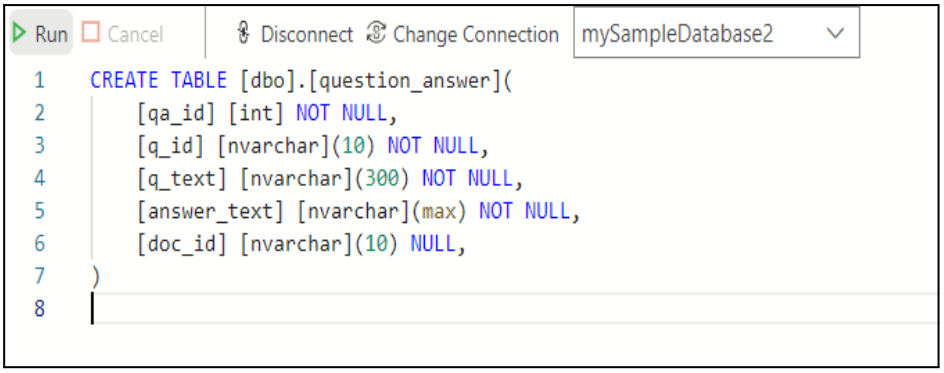
3.5.4. To set up the nugget_sentence table.



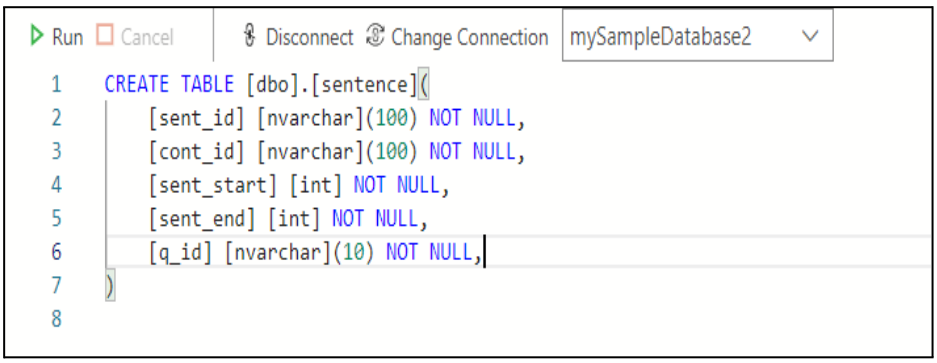

3.5.5. To set up the question table.




3.5.6. To set up the question_answer table.

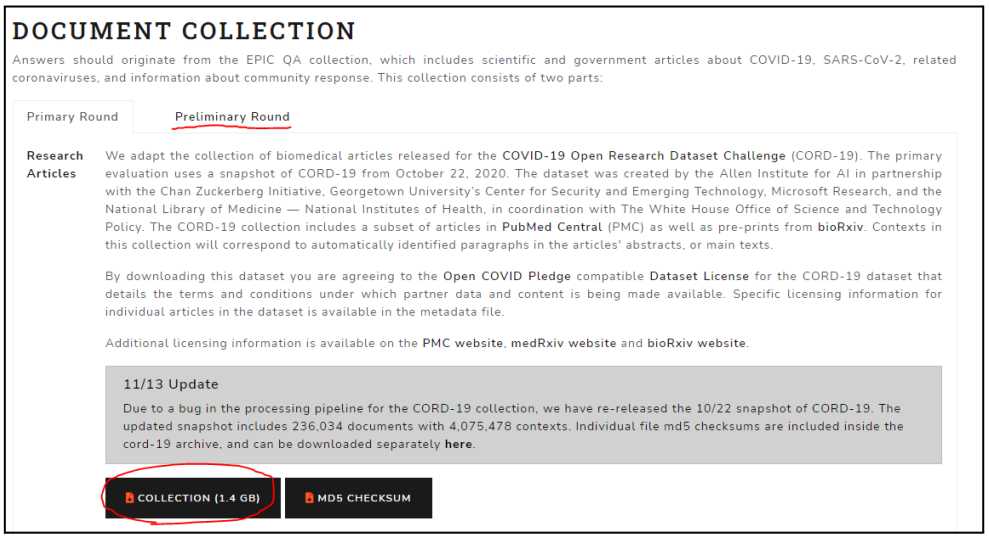



3.5.7. To set up the sentence table.



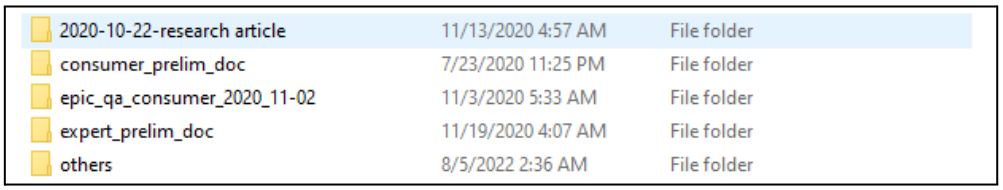
# 4. Populating the database

4.1 Download all the datasets from 'EPIDEMIC QUESTION ANSWERING' project website.

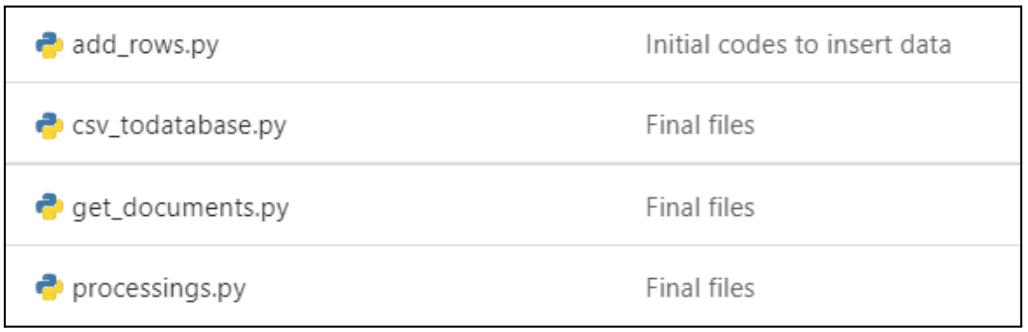



All the Preliminary Round datasets must be downloaded as well.

4.2 Unzip all the datasets and save the files to any convenient locations.




4.3 Download the python code files from our project's Gitlab repository.



4.4 Open add_rows.py, which is the main data population code, and modify variables in the codes so that it matches your database set-ups and dataset file location.

```
# DB server address
server = 'mysqlserver300.database.windows.net'
# connecting DB name
database = 'mySampleDatabase2'
# user name
username = 'cs15'
# user password
password = 'Computer@2020'
```

```
# Directory to expert articles JSON
document_expert_json_dir = "C:/Users/pckim/Desktop/University_Files/Year3-Se
# Directory to consumer articles JSON
document_consumer_json_dir = "C:/Users/pckim/Desktop/University_Files/Year3-

# Directory to expert questions prelim
prelim_question_expert_json_dir = "C:/Users/pckim/Desktop/University_Files/Y
# Directory to consumer questions prelim
prelim_question_consumer_json_dir = "C:/Users/pckim/Desktop/University_Files

# Directory to judgements file
judgements_json_dir = "C:/Users/pckim/Desktop/University_Files/Year3-Sem2/FI
```

4.5 Uncomment the lines that you want to run. This will start the actual data population process and the process will be done fully automatic.

```
# add_document_expert()
# add_context_expert()
# add_sentence_expert()
# add_prelim_question_expert()

# add_document_consumer()
# add_context_consumer()
# add_sentence_consumer()
# add_prelim_question_consumer()

# add_sentence_nugget_only()

# add_nugget()
```

(Note: Each function takes a very long time (2+ hours) and it is strongly recommended to run each function one by one in order to prevent errors).

4.6 If you experience timeout errors in running the add_rows.py, you should use processings.py and csv_todatabase.py codes, which are the alternative codes for bulk-data population to resolve the issue.
4.7 ‘processings.py’ converts datasets into a .csv file and csv_todatabase.py inserts the created .csv file into the database.

```
# Directory to expert articles JSON
document_expert_json_dir = "C:/Users/pckim/Desktop/University_Files/Year3
# Directory to consumer articles JSON
document_consumer_json_dir = "C:/Users/pckim/Desktop/University_Files/Yea

# Directory for csv
context_expert_csv = "C:/Users/pckim/Desktop/University_Files/Year3-Sem2/
context_consumer_csv = "C:/Users/pckim/Desktop/University_Files/Year3-Sem
```

Similar to 4.4, you have to modify the file directories to match your dataset file location and you have to specify the directories for the .csv output files.

4.8 After converting the dataset into .csv files, now you have to specify the directory of the .csv file in the csv_todatabase.py codes and run the python script.

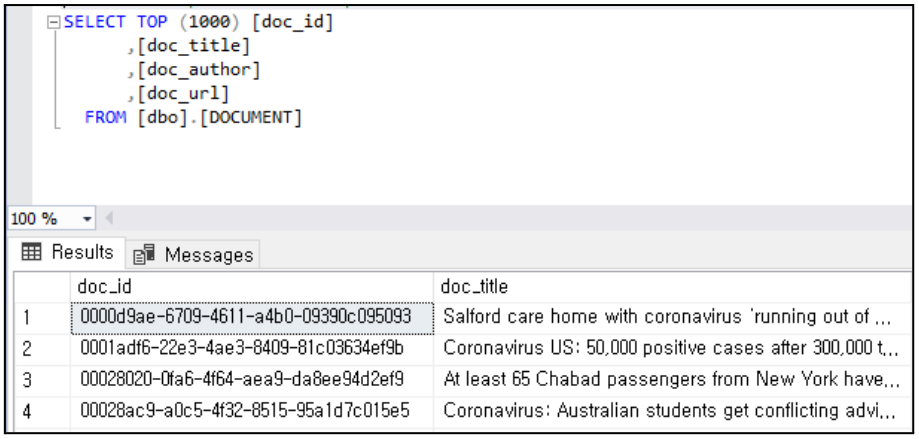
```
connection_uri = sa.engine.URL.create(
    "mssql+pyodbc",
    username='cs15',
    password='Computer@2020',
    host=server,
    database='mySampleDatabase2',
    query={"driver": "ODBC Driver 17 for SQL Server"},
)



df = pd.read_csv("C:/Users/pckim/Desktop/University_Files/Year3-Sem2/FIT3
engine = create_engine(connection_uri, fast_executemany=True)
conn = engine.connect()

df.to_sql(name="CONTEXT", con=engine, if_exists='append', index=False)
```

4.9 Modify the DB server settings and the directory of the .csv file so that it matches your current state. Then, run the codes to start inserting the dataset.
4.10 Check the database after running the scripts.
4.11 If there is no error in running the python scripts, the data will exist in the database.



```
SELECT TOP (1000) [doc_id]
      ,[doc_title]
      ,[doc_author]
      ,[doc_url]
  FROM [dbo].[DOCUMENT]
```

100 %

Results | Messages

| | doc_id | doc_title |
|---|---|---|
| 1 | 0000d9ae-6709-4611-a4b0-09390c095093 | Salford care home with coronavirus 'running out of ... |
| 2 | 0001adf6-22e3-4ae3-8409-81c03634ef9b | Coronavirus US: 50,000 positive cases after 300,000 t... |
| 3 | 00028020-0fa6-4f64-aea9-da8ee94d2ef9 | At least 65 Chabad passengers from New York have... |
| 4 | 00028ac9-a0c5-4f32-8515-95a1d7c015e5 | Coronavirus: Australian students get conflicting advi... |