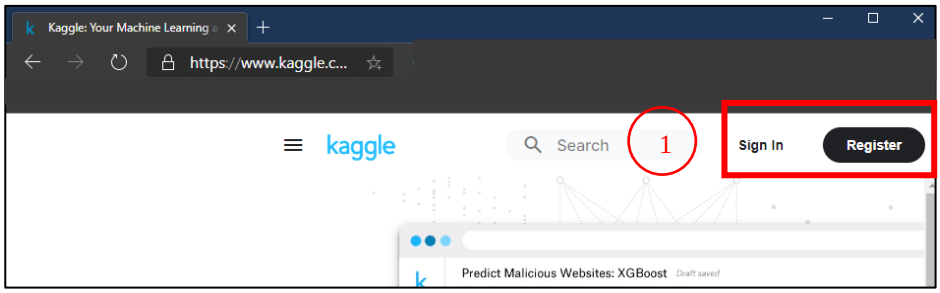
Contents



## I. Joining the Competition

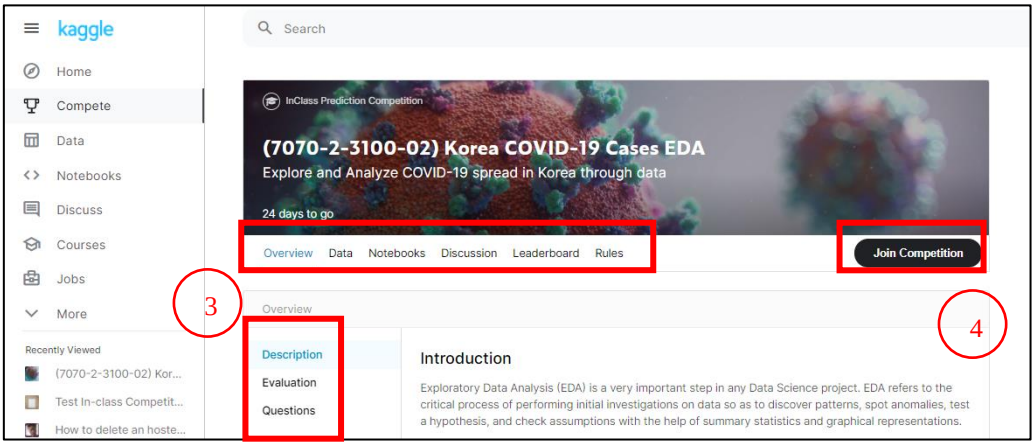
Step 1: Create or Log in to your Kaggle Account.



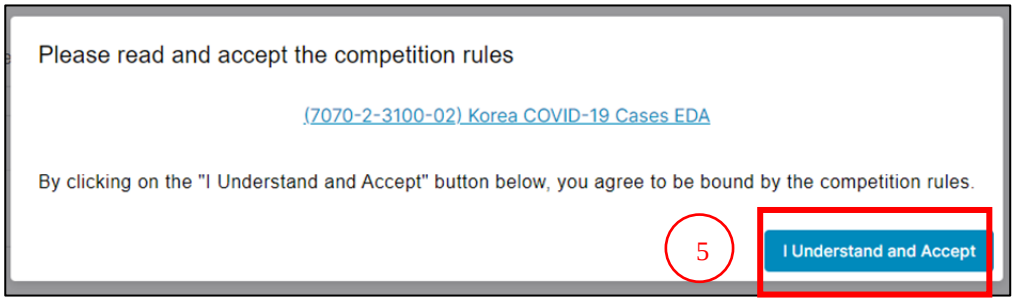

Step 2: Go to this link or on the link that will be posted in the Announcements (공지사항).

Step 3: Carefully read and understand the general instructions that are provided in the competition overview (Description, Evaluation and Questions).

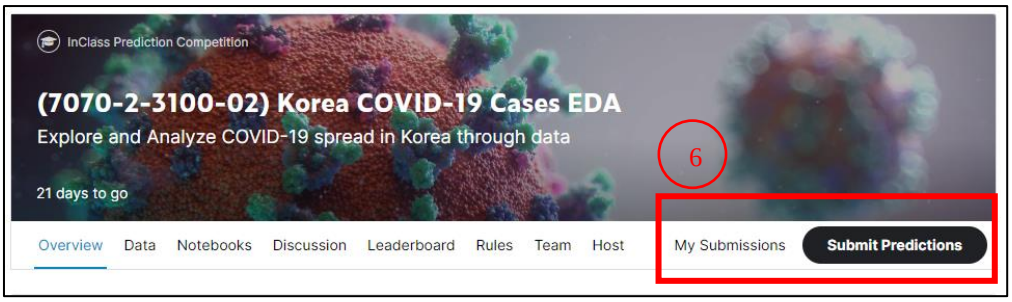
Step 4: Click on Join Competition.



Step 5: Click on I Understand and Accept.



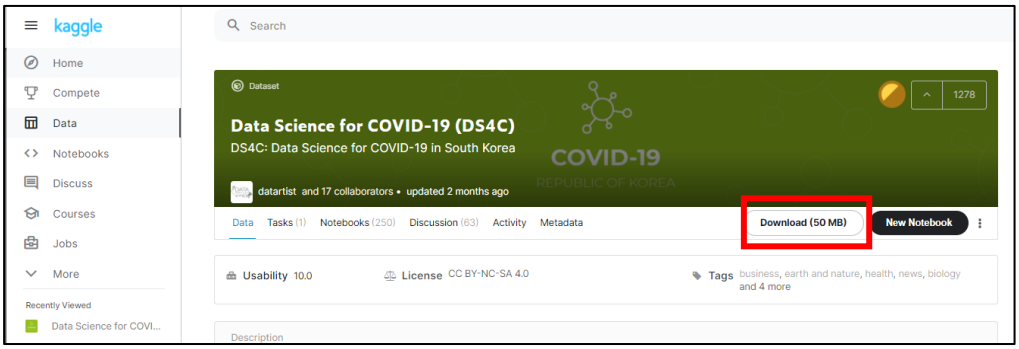

Step 6: You have officially joined the competition when you see the Submit Predictions button.




## II. About the Data

### A. For Exploration and Analysis

In this In-class competition, the students will be exploring and analyzing the DS4C Dataset. This dataset is hosted in Kaggle and it contains the partial documented cases of COVID-19 in South Korea as of June 2020. You can get further details about the dataset by clicking on this link or going to the link indicated in the In-class competition overview. You can download the data by clicking on the Download button.

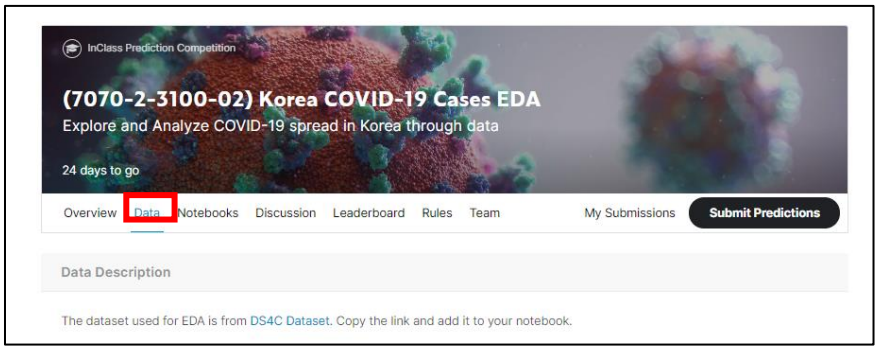



We will only use this data for this in-class competition.

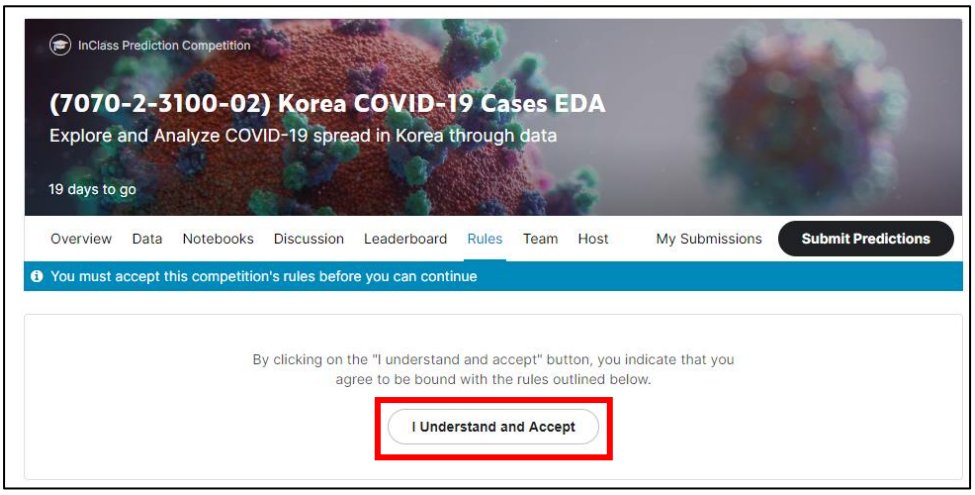
### B. For Submission of Results

Once data exploration and analysis are done, the students must answer the questions related to the tasks as indicated the Questions page of the In-class competition. The answers must be placed in the test.csv file, which serves as the

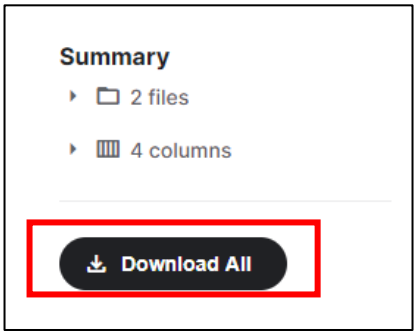
students' answer sheet. This file can be found under the Data page of the In-class competition website.




You must agree to the competition rules to be able to download the submission files.




You can download test.csv together with the sample submission file by scrolling down the Data page and clicking on Download All.




## III. Notebook Format

After having officially joined the competition, the students can then proceed on coding to solve the tasks and answer the questions described in the In-class competition.

To facilitate ease in checking the codes, a sample notebook is provided for the students to fill in with their own codes. This notebook also contains examples and further instructions regarding the in-class competition.

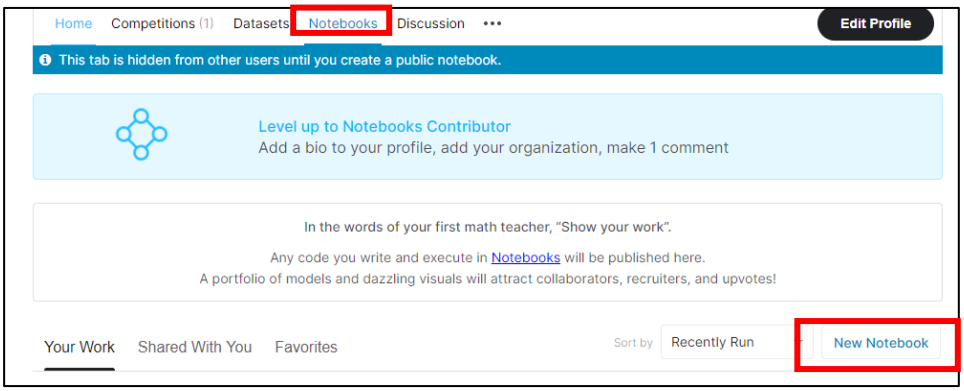
You can download the sample notebook by going to this link or by going to the Lecture Archives (강의 자료실).

## IV. Working on Kaggle Environment

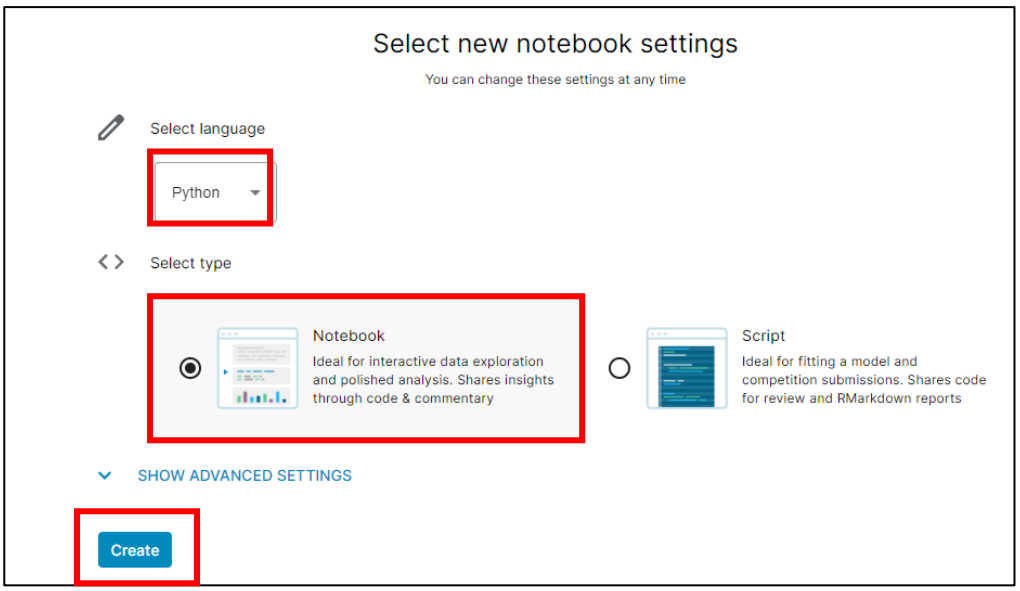
The students have several environments to choose from to work on their solution codes. First is on Kaggle. Kaggle provides free but limited access to accelerators such as GPUs and TPUs. The maximum runtime for a Kaggle kernel is 9 hours. This means that after 9 hours, your session will disconnect, or refresh and you will lose all the progress prior made. Keep this in mind while working on Kaggle.

To set up a Kaggle notebook, do the following:

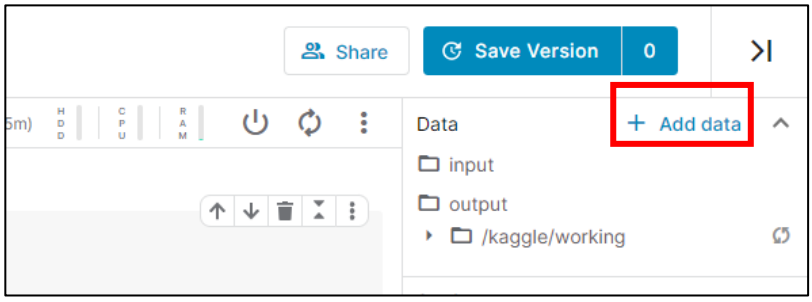
Step 1: Go to your profile and create a New Notebook under the Notebooks tab.



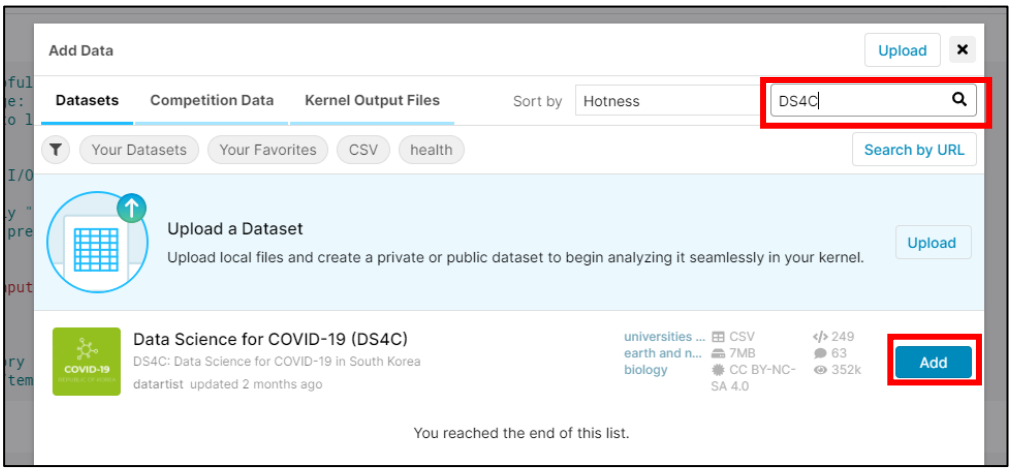

Step 2: Choose Python as Language and Notebook as the Type. Then click on Create.

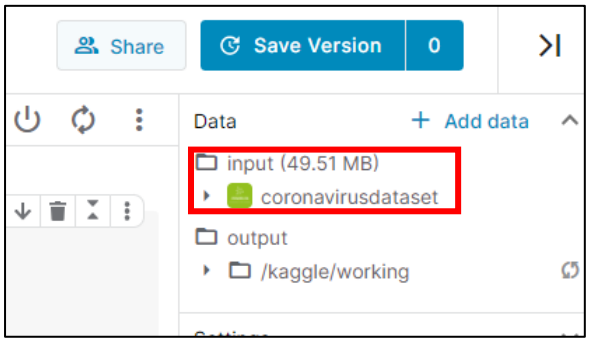


Step 3: Add data to your notebook by clicking on Add data.

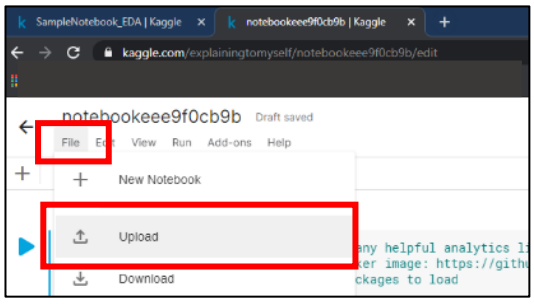



Since we are using a Kaggle-hosted data in this In-class competition, just search for the name of the dataset, in our case the name is “DS4C”, then click Add. The data is attached to your kernel if it appears in the input directory. Note that any data in the input directory will not be lost even if the session disconnects or restarts. This is not the case for data in the output directory. You cannot save output data or edit the files in the input folder.

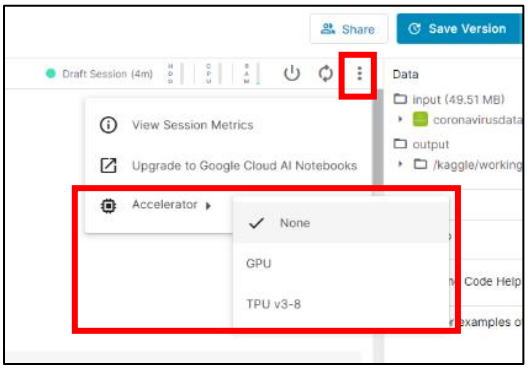


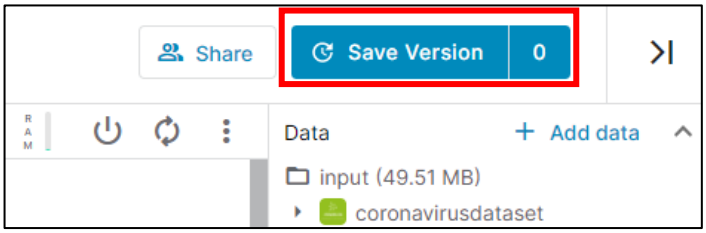

Step 4: Upload the sample notebook format by going to File → Upload.



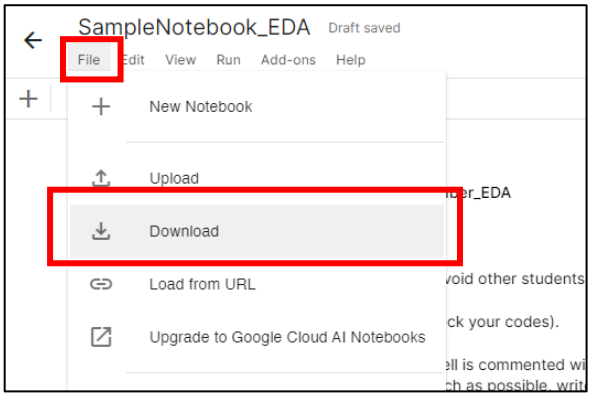
Step 5: Select your runtime/accelerator. For this task, accelerator is not necessary. You can now start editing your solution code.




Step 5: After doing all the required tasks, save your notebook by clicking on Save Version.




Step 6: You can download your notebook by going to File → Download.




## V. Working on Jupyter Notebook Environment

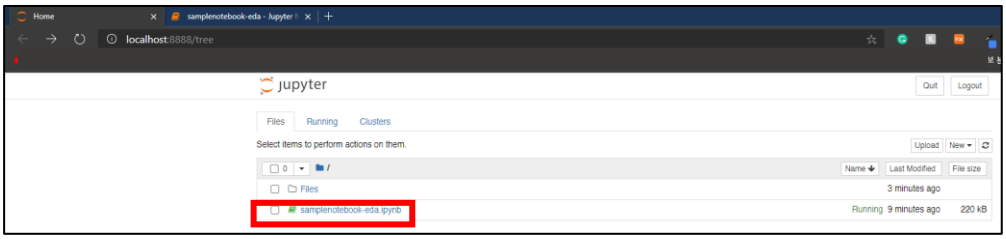
The students can also work on their local machines by using Jupyter Notebook. This environment uses whatever compute capabilities are available in the local machine. The advantage of using this environment is that the users do not have to worry about runtime disconnections, hence, it can be assured that the progress will not be lost, unless there are other issues such as power interruption.

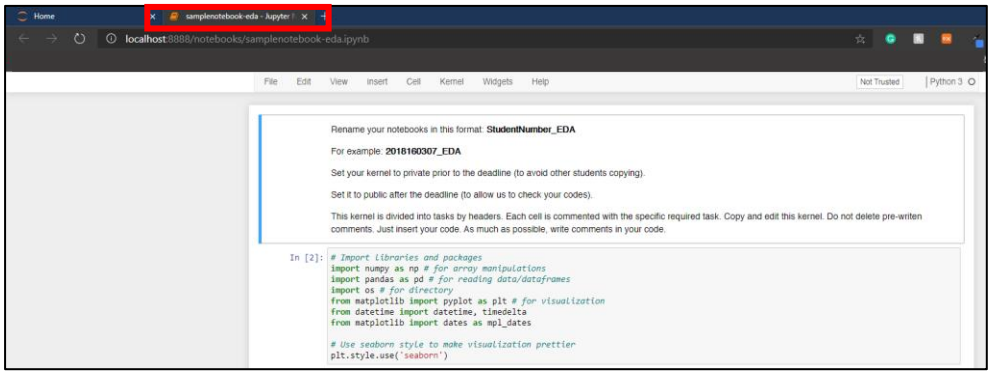
To set up a Jupyter Notebook, do the following:

**Step 1:** Download and Install Jupyter Notebook module or Anaconda Environment (which comes with pre-installed Jupyter Notebook).

**Step 2:** Download the sample notebook format and dataset to your local machine.

**Step 3:** Launch Jupyter Notebook and you can now start editing your solution code.




## VI. Working on Colab Environment

Colab is another option that offers free but limited access to accelerators such as GPU and TPU. Just like in Kaggle, whatever progress you made prior to the runtime disconnection will be lost. However, unlike Kaggle, you will also lose any data you uploaded to the environment. Keep this in mind while working on Colab.

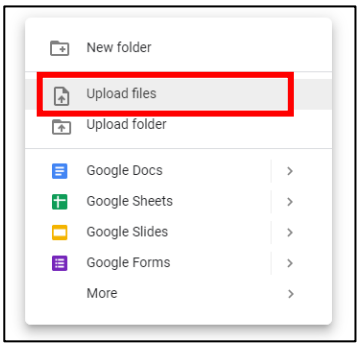
To set up a Colab Notebook, do the following:

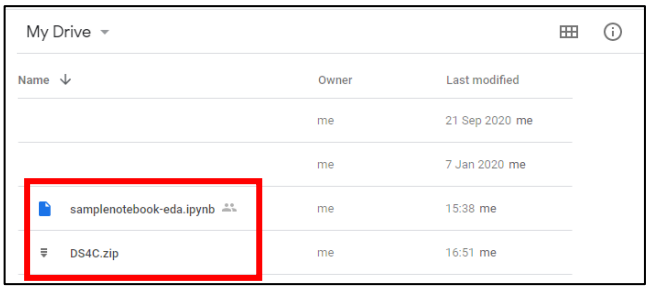
**Step 1:** Create a Google account.

Step 2: Open your Google Drive.

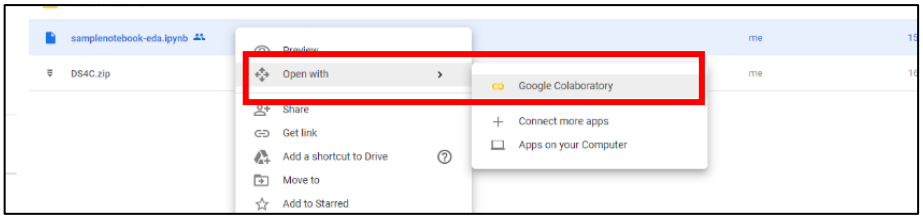
Step 3: Upload the sample notebook format to your drive.

Step 4: Upload the dataset to your drive.

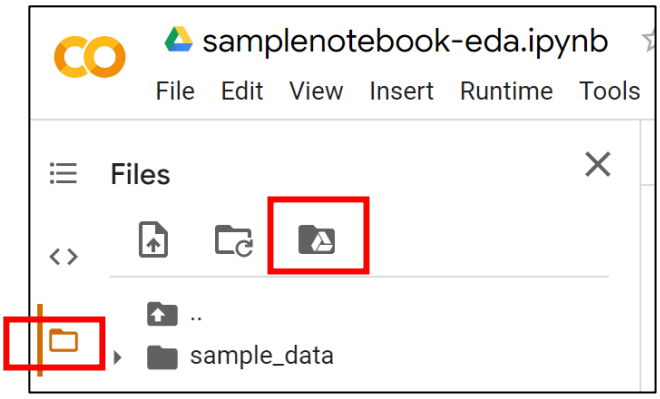


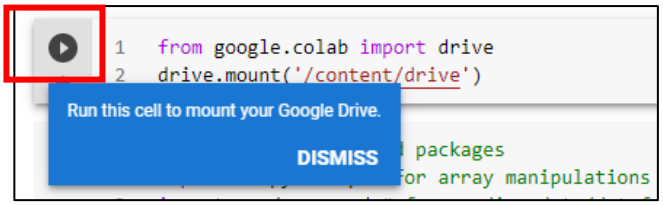

Step 5: Open the sample notebook with Google Colab. This will open a new tab that shows the Colab environment.



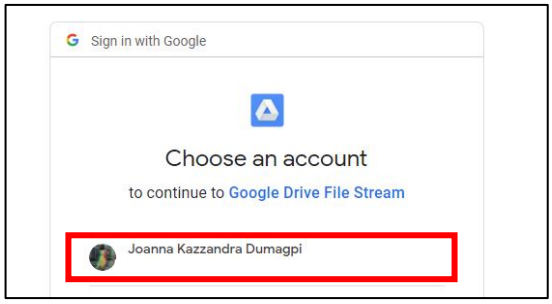

Step 6: Mount your drive to Google Colab environment to access the dataset. You can do this by going to the Files tab and clicking on the Mount Drive icon.

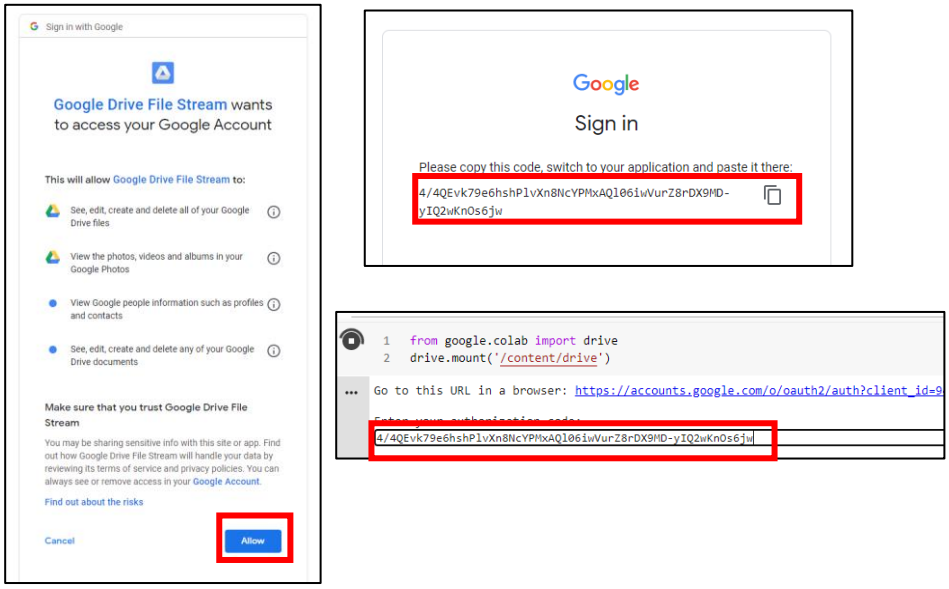



Run the code in the cell that automatically appears.

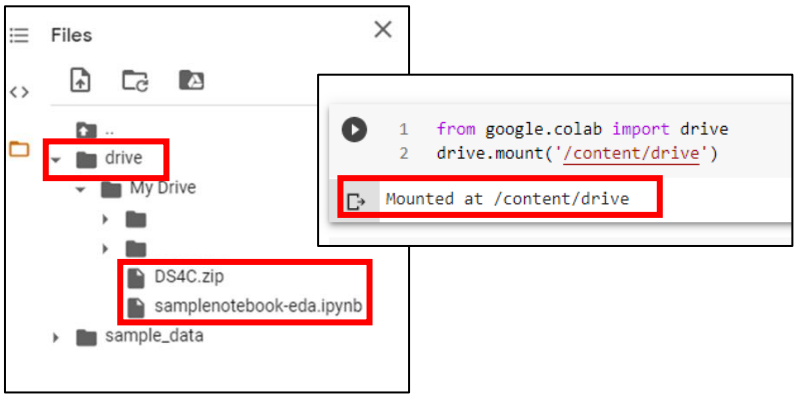



Click on the URL, select the Google account that contains your dataset, allow access to your account, and copy-paste the authorization code.

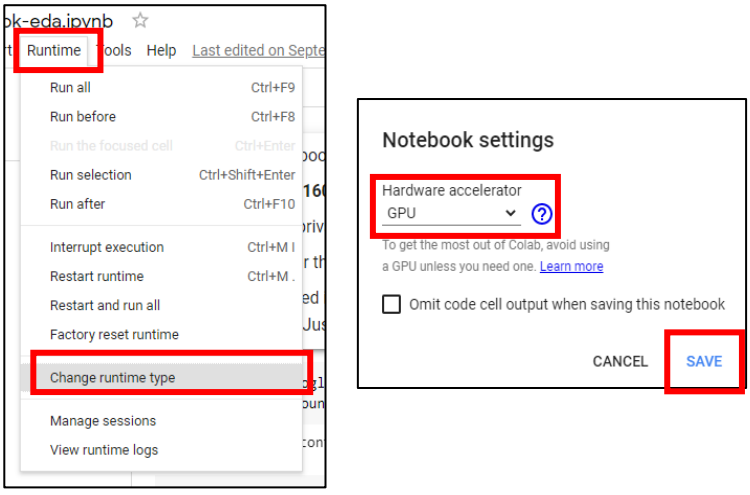






The drive is successfully mounted when the folder appears on the Files tab. Unlike in Kaggle, any data that is read and written to your drive will not get lost in case of runtime disconnection.




Step 7: Select your runtime/accelerator by going to Runtime →Change Runtime Type. For this task, accelerator is not necessary. You can now start working on the competition tasks in your Colab notebook. Any changes you make to the notebook can be saved to the notebook file in your drive.



## VII. Submitting the Results

After completing the tasks, you should be able to answer all 40 questions related to the dataset.

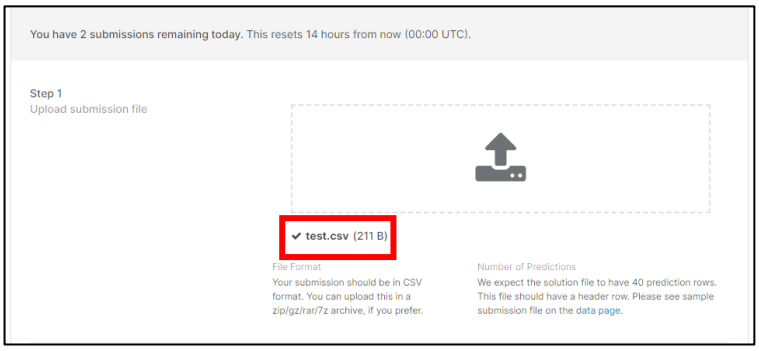
To proceed with the submission, do the following:

Step 1: Download the test.csv.

Step 2: Edit the file and insert your answers to the column named answer and to the appropriate row that corresponds to the question number. See the sampleSubmission.csv, which contains the correct answers to the first five questions, for additional guidance.

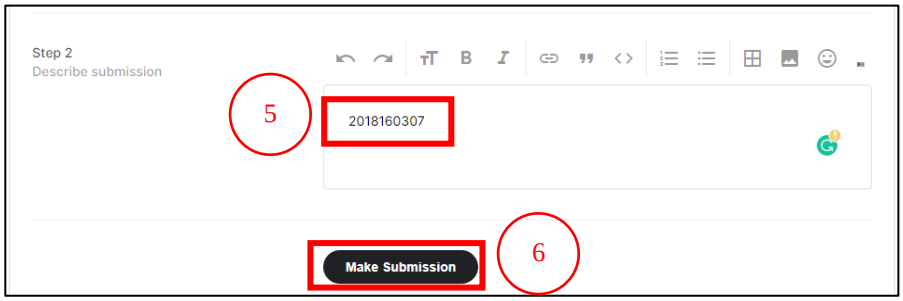
Step 3: Go to the In-class competition website and click Submit Prediction.


Step 4: Upload your filled-in test.csv file.




Step 5: In the description, write your student number.

Step 6: Click Make Submission.



Step 7: Your submission is confirmed after it has been scored based on the evaluation metric.


Step 8: You must also submit the notebook with your solution code in KLAS Submit Assignments (과제 제출).