<div align="center">

CSC110 Project Proposal
Educational Crisis - A Closer Examination on the Correlations Between Covid-19 and School Closures Around the Globe

</div>

<div align="center">

Shouyi (Ray) Hung, Yuhan (Charlotte) Chen, Mengyuan (Alyssa) Li, and Fuyang (Scott) Cui

November 5, 2021

</div>

Problem Description and Research Question

COVID-19 profoundly impacted students' learning environment and strategies. **Therefore, we are curious about how this global pandemic correlates with school closures all around the world as time passes, which is one of the main influencing factors that entirely changed our way of learning and living.**

As a group of students, COVID-19 changed our way of learning from face-to-face to online for quite a long time. However, after COVID-19 eased a little bit, some of our schools reverted to the traditional in-person learning classes. Frequently switching between different learning environments and methods is unhealthy for our personal development because we need time to get accustomed to new things, and it is hard for us to normally keep up with the pace of our teachers in this condition.

Besides, as international students, we were energetic and excited about future university life. However, everything became harsh and unpredictable after the emergence of COVID-19. We are now bothered by expensive flight tickets, personal safety issues, and potential school closures as a result of the pandemic.

Therefore, we aim to discover a general relationship between COVID-19 and school closures. With the help of the observed correlation, we could be more prepared in countering the impacts of COVID-19 as individuals. For example, we could reasonably predict the next virus outbreak based on our project and switch to online classes beforehand.

Additionally, from a broader scope, our project could provide intuitions to educational institutions about the trend of school closures and COVID-19 cases. Therefore, they could identify whether they made a correct decision of closing/opening schools during the pandemic, and draft plans to minimize the impacts in the future. In other words, we could utilize our project as a guide to help prevent future impacts that could rain onto the educational sectors that suffered during the current pandemic.

Dataset Description

We have identified two main datasets that will be relevant for our project's implementation.

These are:

1. Global School Closures for COVID-19 – Obtained from Kaggle, compiled by Saleh Ahmed Rony, sourced from UNESCO

2. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University – Obtained from GitHub, compiled by JHU, sourced from WHO, ECDC, DXY, US CDC, etc.

Both datasets will be stored in a Comma Separated Value file, which will allow us to read from them easily through Python's csv library.

Furthermore, both datasets are very credible as they are sourced from multiple sites, including but not limited to WHO, ECDC, and US CDC. Furthermore, these datasets are also licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0), which allows us to utilize these data for our own needs.

The datasets that we have downloaded and utilized in this project were chosen because they were compiled in a way that allows easy access and manipulation. By using datasets that are already organized could improve the efficiency and robustness of our program.

The Global COVID-19 Dataset (Time series) has the following structure:

Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	...
	Afghanistan	33.93911	67.70995	0	0	0	...
	Albania	41.1533	20.1683	0	0	0	...
...	...	...	...	...	...	...	...

The headers extend up until today.

The US Data set for COVID-19 (Time series) will have a little bit of variation, but it is generally the same.

The School Closure Dataset has the following structure:

Date	ISO	Country	Status
17/02/2020	CHN	China	Partially open
17/02/2020	MNG	Mongolia	Closed due to COVID-19
...	...	...	...

The data is organized by entries of different country each day.

Some of the country names in our datasets contain characters that are not part of the standard ASCII table. An example would be "Curaçao", where "ç" is a French (Latin script) letter.

Since these letters are not in the standard ASCII Table, displaying and encoding issues may arise. So for convenience, we will ignore them.

For the purpose of this project, we will simplify the standard ASCII Table to only contain characters from "a-z", "A-Z", "0-9", "!-~".

These letters also correspond to "string.ascii_lowercase", "string.ascii_uppercase", "string.digits", and "string.punctutation" respectively, from the Python string library.

Computational Plan

Our project will be separated into 2 parts: data processing and Graphical User Interface (GUI) implementation.

Processing Data

We will firstly read the data into Python with help of the Python *csv* library.

Then, we will convert raw data into many data classes so we could easily manipulate those data in practice.

GUI Implementation

Currently, we plan to use *PyQt5* to generate an interactive user interface, *Matplotlib* to plot graphs, and some other helper tools like Qt Designer and other libraries.

We will implement the following functionalities, or parts:

- Interactive COVID-19 cases and deaths and school closures visualizations supporting filtering, grouping, sorting functionalities, etc.

  We will implement most algorithms by ourselves.

- Interactive table view of the visualizations above.

With a GUI, users of our application could quickly scan the data that we have provided and create their own plots by filtering the data by their will. Also, they could check which point on the graph may correspond to which observation.

With the help of our program, people could view, compare, and even predict the trend of the COVID-19 pandemic. Furthermore, we can also answer our research question by quickly filtering countries around the world and attempt to generalize a relationship between the trend of COVID-19 and school closures.

With these technical implementations, we believe that the impacts brought by the COVID-19 Pandemic would be conspicuous and people could formulate solutions towards the impacts illustrated.

References

"Coronavirus." World Health Organization, WHO, 10 Jan. 2020, www.who.int/health-topics/coronavirus#tab=tab_1.

"CSV File Reading and Writing — Python 3.10.0 Documentation." Python Documentation, docs.python.org/3/library/csv.html. Accessed 30 Oct. 2021.

Johns Hopkins University. "GitHub - CSSEGISandData/COVID-19: Novel Coronavirus (COVID-19) Cases, Provided by JHU CSSE." GitHub, CSSEGISandData, github.com/CSSEGISandData/COVID-19. Accessed 30 Oct. 2021.

Kaggle — Global School Closures for COVID-19. Saleh Ahmed Rony, www.kaggle.com/salehahmedrony/global-school-closures-covid19. Accessed 30 Oct. 2021.

"Pygame v2.0.1.Dev1 Documentation." Pygame, www.pygame.org/docs. Accessed 30 Oct. 2021.

"Scikit-Learn 1.0.1 Documentation." Scikit-Learn, scikit-learn.org/stable/index.html. Accessed 30 Oct. 2021.

"Sustainable Development Goal 4 (SDG 4) — Education within the 2030 Agenda for Sustainable Development." SDG 4 Education 2030, sdg4education2030.org/the-goal. Accessed 30 Oct. 2021.