Exploring effect on child's education due to school closure due to COVID-19 pandemic

Biswajit Kumar| Pankaj Kumar Singh| Pierre LeBlanc

MOTIVATION

The potential losses that may accrue in learning for today's young generation, and for the development of their human capital, are hard to fathom. Schools with more than **168 million children** globally have been completely closed for almost an entire year due to COVID-19 lockdowns in 2020

More than 1 billion children were impacted due to school closures aimed at containing the spread of COVID-19. Numerous factors disrupted kids' education during pandemic e.g., School closures, limited access to resources to continue studies etc. Disruption in education is a long-term burden on society and individuals. 188 countries imposed countrywide school closures during the pandemic, **affecting more than 1.6 billion children and youth**. Even prior to the pandemic, however, children's learning was in crisis, and the pandemic has only sharpened these inequities, hitting school children in poorer countries particularly hard. Globally, many schools lack the resources to invest in digital learning, and many children from poorer households do not have internet access. Organizations like UNICEF have been closely monitoring the impact of COVID-19 on parameters like school closures to education (Source here).

This paper attempts to explore the data available at UNICEF sites and **recreate/create the visualization for top ten mostly impacted countries.**

Definitions of school closures caused by the COVID-19 Pandemic (Source, here) :

1. **Closed due to COVID-19**: Government-mandated closures of educational institutions affecting most or all (or at least 80%) of the student population enrolled from pre-primary through to upper secondary levels [ISCED levels 0 to 3]. In most cases, various distance learning strategies are deployed to ensure educational continuity. This is termed as "Fully Closed" school closures status.

2. **Academic break**: Most schools across the country are on scheduled academic breaks for periods of at least one week. All study during this period is suspended.

3. **Fully open**: For most schools (at least 80%), classes are being held exclusively in person, noting that measures to ensure safety and hygiene in schools vary considerably from context to context and/or by level of education.

4. **Partially open**: Schools are (a) open/closed in certain areas only, and/or (b) open/closed for some grade levels/age groups only; and/or (c) open but with reduced in-person class time, combined with distance learning (hybrid approach). It also includes the countries where national governments have deferred decisions on re-opening to other administrative units (e.g., region, municipality, or individual schools), and where a variety of re-opening modalities are being used.

The goals of this paper are:

1. Determine longest and shortest school closures period using time series analysis.
2. Determine the total number of children-days-closed impacted due to school closure during COVID-19. (children days fully closed impacted by shutdowns)
3. Determine correlation between school closure vs relative GDP.

DATA SOURCES

Dataset	UNESCO school closures databases (Primary dataset)	GDP (Secondary dataset)
Link	Access to data here	Access to data here
Description	The dataset details the status of school closures by region and country.	The datasets have information about the country's GDP over time.
Source	UNESCO website	World bank
Format	Excel	CSV
Access Method	Download	Download
Variables	Date, Country ID, Country, status	Country Name, Country Code
Size	14.1 MB	301 KB

DATA MANIPULATION METHODS

Python Libraries: Pandas, missingno

Generic steps:

1. Load the excel and CSV file into a pandas dataframe
2. Explored missing values in the datasets by missingno library (Figure 1) and pandas function
3. Converting the date column to pandas' date time format
4. Clean the dataframe by filtering out specific countries of interest in the Primary dataset.
5. Renaming the Country ID column to "Status Count"
6. Performed Split, Apply and Combine techniques to create appropriate result set for top10 analysis.
7. Created dataframes with status = "Closed due to COVID-19", Partially open, Closed due to COVID-19, Partially open. Restricted the list of countries to top 10.
8. Slice the main GDP dataframe with only "Country Name", "Country Code", "2020"
9. Slice the UNESCO school closure dataset to keep columns 'Date', 'Country ID','Country','Status
10. Merge the UNESCO and GDP dataframe how='inner', left_on='Country ID', right_on='Country Code'

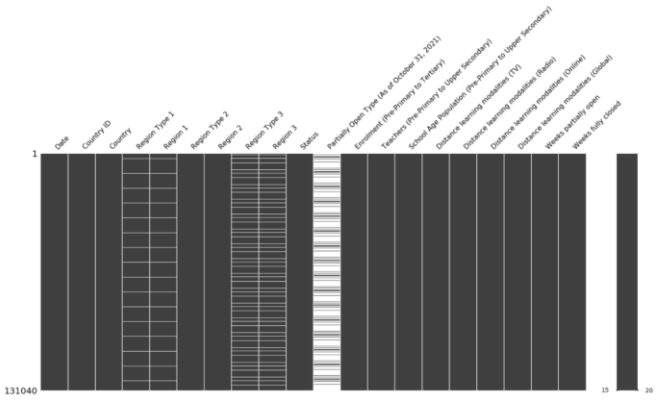

Figure 1: missingno (missing number) matrix plot showing data sparsity across all dataframe columns

Challenges and Solutions:

1. We peer reviewed the notebook however encountered issues replicating the codes into different versions of the libraries. We fixed the issue by updating the functions and renaming columns.

Analysis and Visualization

Goal # 1: Determine longest and shortest school closures period from February 2020 to October 2021 using time series analysis.

Analysis:

The school closure analysis has been performed for the top ten countries for four different school closure statuses (status as described in the motivation section).

- Bangladesh has been highly impacted with 439 days of "fully closed" school (Figure 2). In terms of duration, Panama (Figure 3) had longest continued fully closed schools for more than 3 quarters from March 2020 to January 2021.
- The United States has been highly impacted with 465 days of "Partially Opened" school (Figure 4). Duration wise the United States, Argentina, and Colombia (Figure 5) had longest continued partially opened status for close to 3 quarters.
- Solomon Islands and Singapore with 27 days of fully closed school status (Figure 6) were in the top for the "shortest fully closed" status. The majority of school closures happened between March-29 to April-12 in 2020 (Figure 7) for this category.
- Algeria and Croatia with 14 days of partially opened school status (Figure 8) were in the top for the "shortest partially opened" status. The shortest partially opened duration was observed in June 2020 and Nov 2020 for Croatia and Algeria (Figure 9).
- Belarus, Burundi, Nauru and Tajikistan did not have school closures due to Covid-19 (Figure 10).

Visualizations:

All the graphs below (except Figure11) are newly created as part of this goal and do not exist on the Unicef site. The source for the data used in all the graphs is the primary dataset as described under the data sources table above.

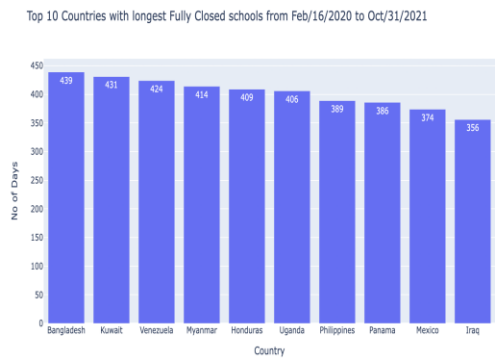

Figure 2

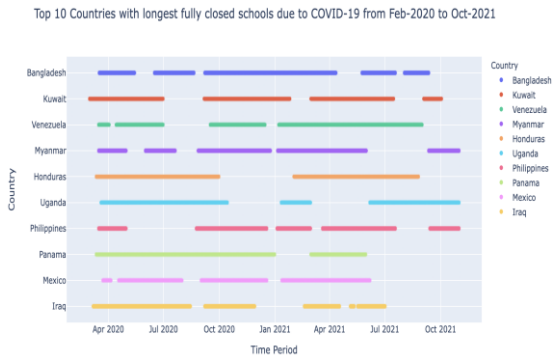

Figure 3

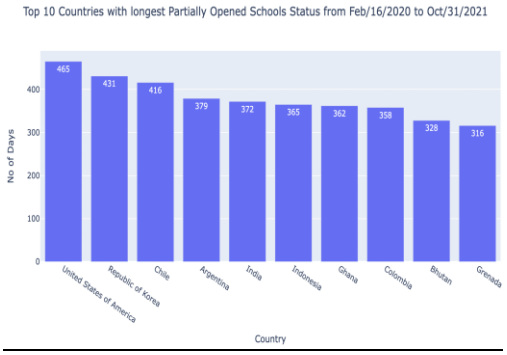

Figure 4

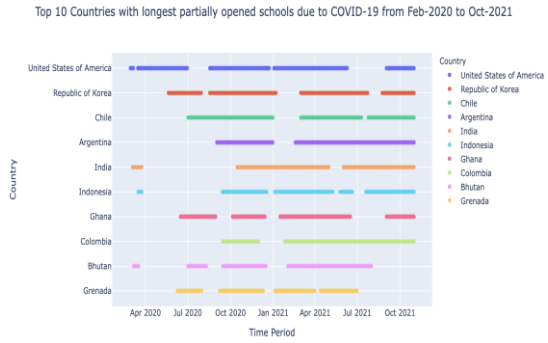

Figure 5

Figure 6

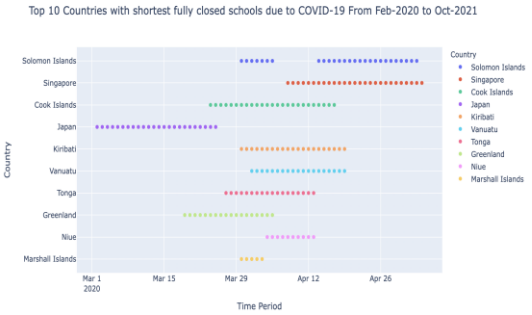

Figure 7

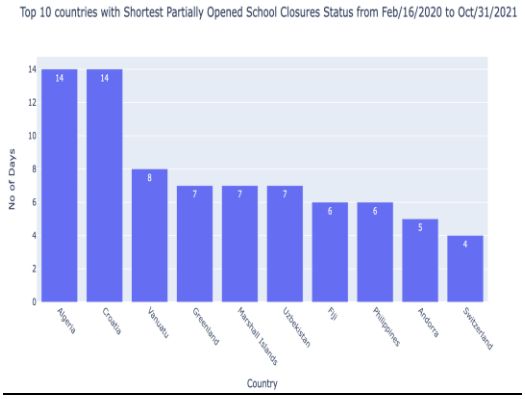

Figure 8

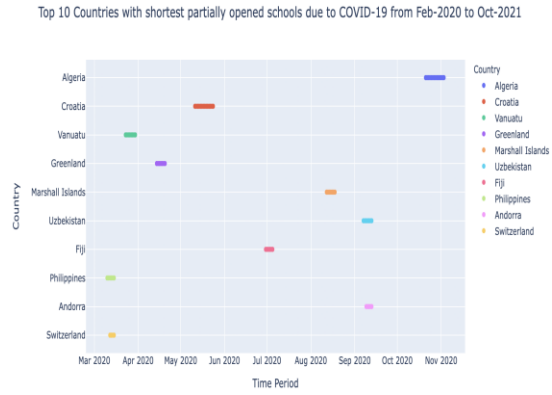

Figure 9

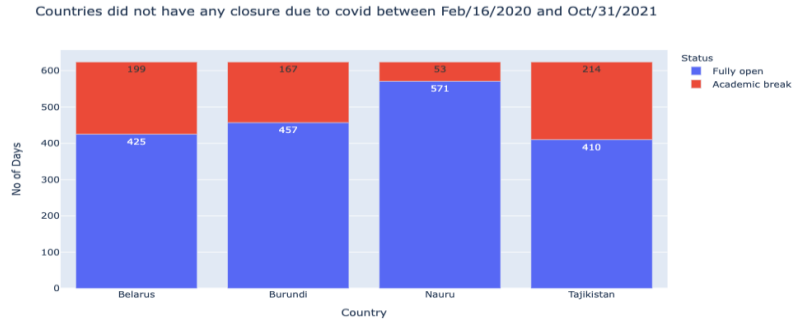

Figure 10

The trend of the global school closures status for different statuses were observed with the help of time series graphs (Figure 11 & 12). The school closures became more prevalent from March

July-2020 when pandemic was on the rise. In September, a reversal in trends emerged and around 100 countries opened schools fully, while about 50 countries opened schools partially and around 25 countries kept their schools fully shut.

Source: https://data.unicef.org/resources/one-year-of-covid-19-and-school-closures/

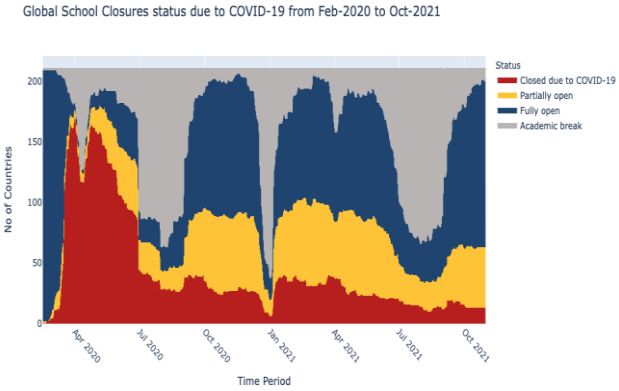

Figure 11

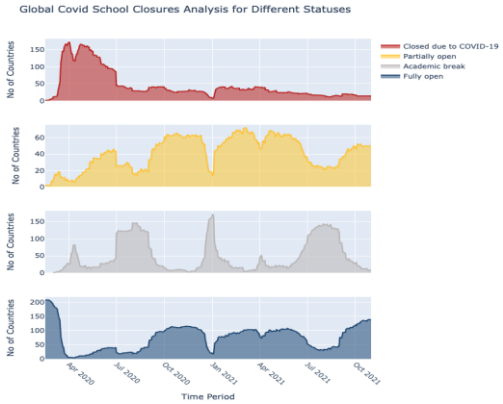

Figure 12

Goal # 2:

Determine the total number of children-days-closed impacted due to school closure during COVID-19. (children days fully closed impacted by shutdowns)

Analysis:

- India had more than triple the total impacted children-days-closed compared to the second top country Bangladesh (Figure 13). Why this is the case needs to be further analyzed, however, India has much more school children and therefore the number of children enrolled multiplied by the number of days fully closed would result in a larger number of children-days-closed.
- Interesting to note the total number of country-school days added together show the schools were open most of the time compared to the other status of closed or partially closed (Figure 13).
- Definitions from UNESCO are listed in the Motivation section at the top of this paper which explain the different status of closure: Closed due to COVID-19, Academic break, Fully open, and Partially open.
- Unfortunately, globally **289.5 Billion Student-Days** were impacted due to fully closed schools as of October 31, 2021 (Figure 13 data set). The assumptions of how this calculation was performed are explained in the Bangladesh example below. Essentially, we use the number of fully closed days multiplied by the total number of students enrolled to get the student-days-closed metric for each country.

Visualizations:

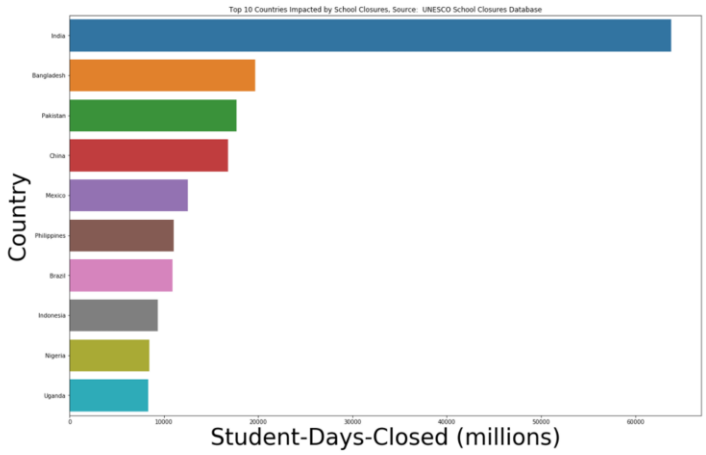

Figure 13

Source: UNESCO map on school closures (https://en.unesco.org/covid19/educationresponse)

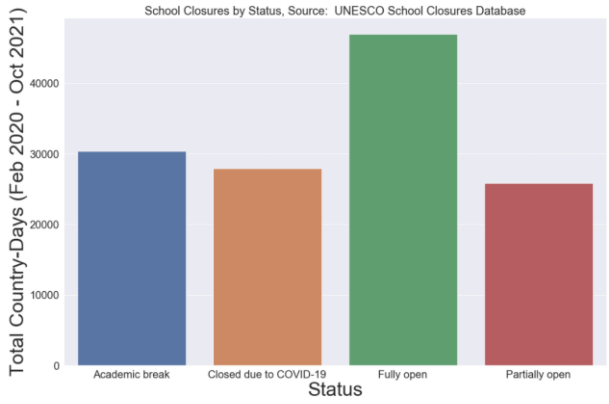

Figure 14

Source: UNESCO map on school closures (https://en.unesco.org/covid19/educationresponse), October 2021

School Closure, Number of Children-Days Impacted, Calculation Method: ***Bangladesh:***

Total Days from data set: Mar 2020 - October 2021 (624 days)
19,684 Million (19.7 Billion) Student-Days closed based on:
*(Fully Closed Days * School Age Population Enrolled)*
439 = Fully Closed Days
44,838,266 School Age Pop Enrolled

Total Days	**624**
Academic Break Days	116
Closed Days	439
Open Days	30
Partially Open Days	39

School Age Population (Pre-Primary to Upper Secondary) Enrolled:	44,838,266

Total Student-Days-Closed (in millions)	**19,684**

Source: UNESCO map on school closures (https://en.unesco.org/covid19/educationresponse) and UIS, October 2021 (http://data.uis.unesco.org)

Goal # 3: Determine correlation between school closure vs relative GDP.

Analysis:

- The analysis is intended to observe any correlation between GDP and School closure.
- Based on correlation matrix and value of Correlation coefficient (R), there seems to be a weak positive correlation between GDP and School closure (Figure 16).
- We can deduce the information that the overall pandemic has impacted across the globe irrespective of countries GDP status with Children's education being impacted badly globally due to school closures (as detailed in goal number 1).

Visualizations:

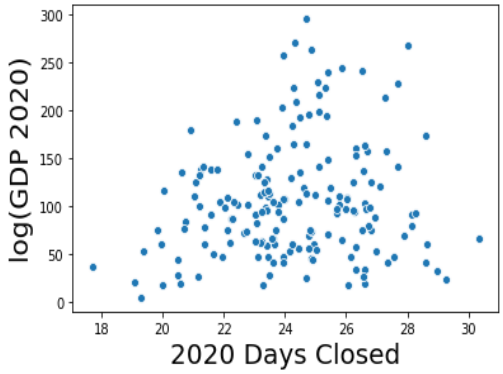

Figure 15

Figure 16

REFLECTION:

It has been an enriching experience to work on this project like forming a team, developing and refining scope, continuous guidance from the advisor, accessing public datasets, understanding the methodologies to replicate the results. First time we came for a project with common goals which helped us to break learning silos. We gathered a lot of perspective from other team projects during the peer review phase of the milestone I and II and received valuable inputs from peers which helped each team member to work on their strength. We learned in the process: nuances of programming like Plotly graphical fixes, explored various platforms like Stack overflow, GitHub to fix the codes which exposed us to the myriad of concepts and development in the field. Application of first year MADS concepts like Data wrangling, Normalizations, Time-series, pair plot etc has been quite exciting under the guidance of Dr. Chris Teplovs in the weekly meeting to whom we remain thankful for all the support. We have been chasing the whooshing noise of deadlines at the professional front along the project but as a team collaborated well for the encouraging outcome. The overall planning of the project timing like in winter break helped us to balance the commitments. We are looking forward to continuing our interactions and sharing learnings even after the course is over.

ETHICAL CONSIDERATION

All datasets were imported and collected from various public platforms or resources. After examination, we determined no data aggregation effect; therefore, we could consider no notable privacy concern. Unfortunately, we have no resources or toolkits to examine further the dataset's healthiness concerning its collection and processing. Thus, we must assume the data is sufficient to provide insights with acceptable accuracy and variances. Since the data is gathered by global organizations like UNICEF and UNESCO, we will always convey the level of uncertainty warranted by our analysis.

FUTURE DIRECTION

We started the project with an objective to replicate the visualizations and gain insights to assess the impact of school closure on kids' education during the pandemic (Goals stated in the motivation section). We found that **289.5 Billion Student-Days** were impacted with Bangladesh being the highly impacted country with 439 days of "fully closed" school in terms of duration (Figure 2). There is weak positive correlation between GDP and School closure, however as we delved more into the data, we developed significant interest to explore more insights from the UNESCO datasets beyond the scope of the project. We would like to further assess the impact of school closure on various socio-economic parameters for eg:

1. To assess correlation between school closure and impact on child marriage.

   Even short disruptions in children's schooling can have long-lasting negative impacts due to factors including the lack of structured programs for catching up. In the past, school closures have led to an increase in child marriage and child labor which often prevent children from continuing their education. The relevant information are available here

2. Existing remote learning options available and opportunities to improvise.

   Based on UNICEF assessment, almost all the countries have implemented digital and broadcast remote learning policies however only 60 percent did so for pre-primary education. A further analysis on current remote learning and comparing different nations would help us to understand best practices for improvisation of digital learnings.

STATEMENT OF WORK

Pankaj Kumar Singh	Pierre LeBlanc	Biswajit Kumar
● Basic data cleaning and manipulation. ● UNICEF School closure database for Goal #1 ● Drafted final report.	● Basic data cleaning and manipulation. ● UNICEF School closure database for Goal # 2 ● Drafted final report.	● Basic data cleaning and manipulation. ● UNICEF School closure database for Goal # 3 through peer review of codes ● Drafted final report.

References

1) COVID-19 and children - UNICEF DATA
2) https://data.unicef.org/topic/education/covid-19/
3) GDP (current LCU) | Data (worldbank.org)
4) https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help
5) https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/ca-impact-of-covid-learning-loss-fall-2020.pdf
6) https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help
7) https://www.lowellma.gov/AgendaCenter/ViewFile/Item/14803?fileID=27539
8) UMICH MADS SIADS 505,501,522,532 for Python Programming