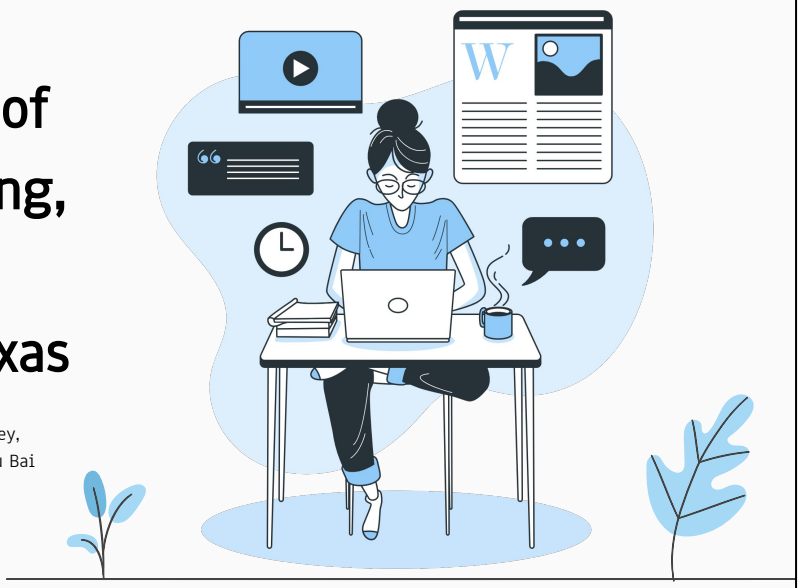
The Viability of Remote Learning, as told by Covid-19 in Texas

W201 Fall 2020 Section
Aline Jaimes, Heather Pieszala, Laurie Cuffney,
Natali Ojeda Meneses, Rakesh Walisheter, Yu Bai

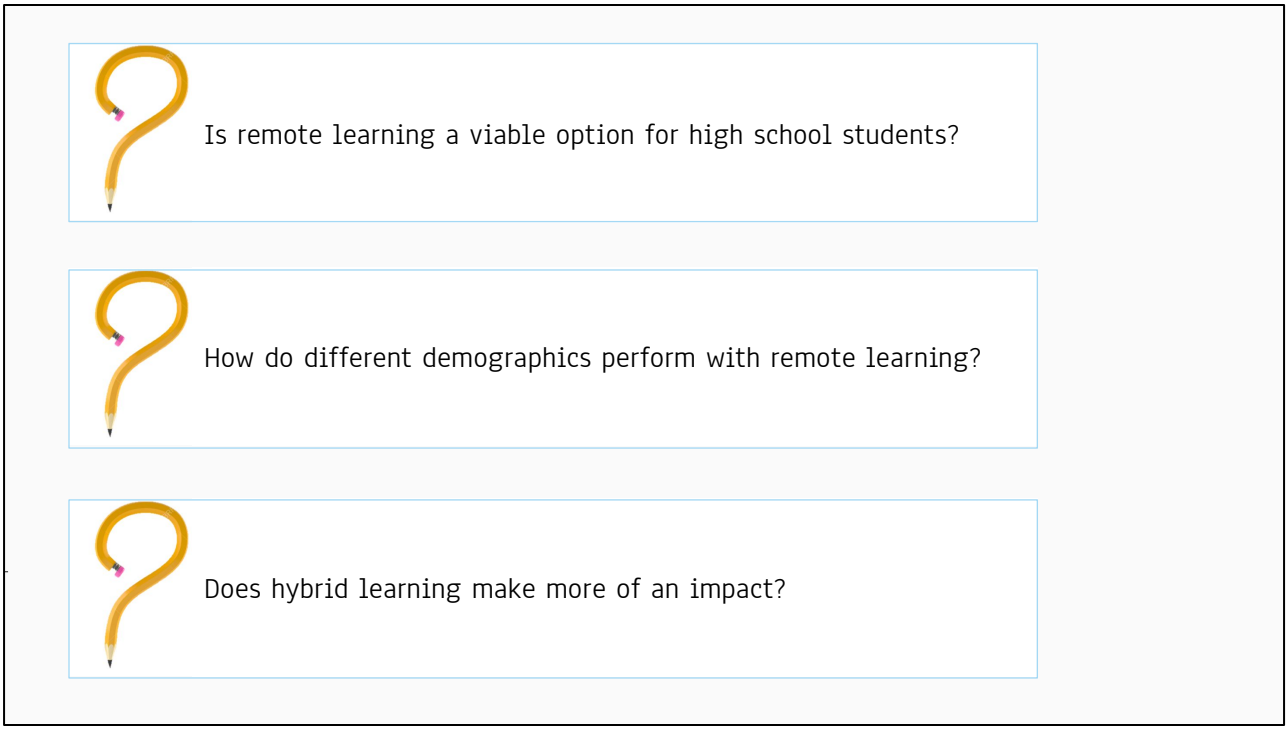
Prepared for: The Texas Education Agency



Welcome school board and Texas Education Agency representatives. Today we are here to talk about the viability of remote learning, as told by covid 19 in texas.

School overcrowding continues to worsen as the population grows. As students progress from elementary, to middle, to high school, class sizes increase and the problem continues to worsen. There did not seem to be a solution to the problem, besides building more schools, which required millions of dollars of capital, at a time when education funding was down.

Then, COVID-19 hit. In early 2020, a novel virus known as COVID-19 ravaged the world, slowed down our lives, and changed the way most individuals live. The phase we have entered, often referred to as “the new normal,” made its way into high school education in the United States in the form of remote learning. Schools had to find the resources to transition to online learning, and students had to adjust to a new way of life and new methods of development. We realized that students could learn outside the classroom, as long as they had access to the internet and a device.



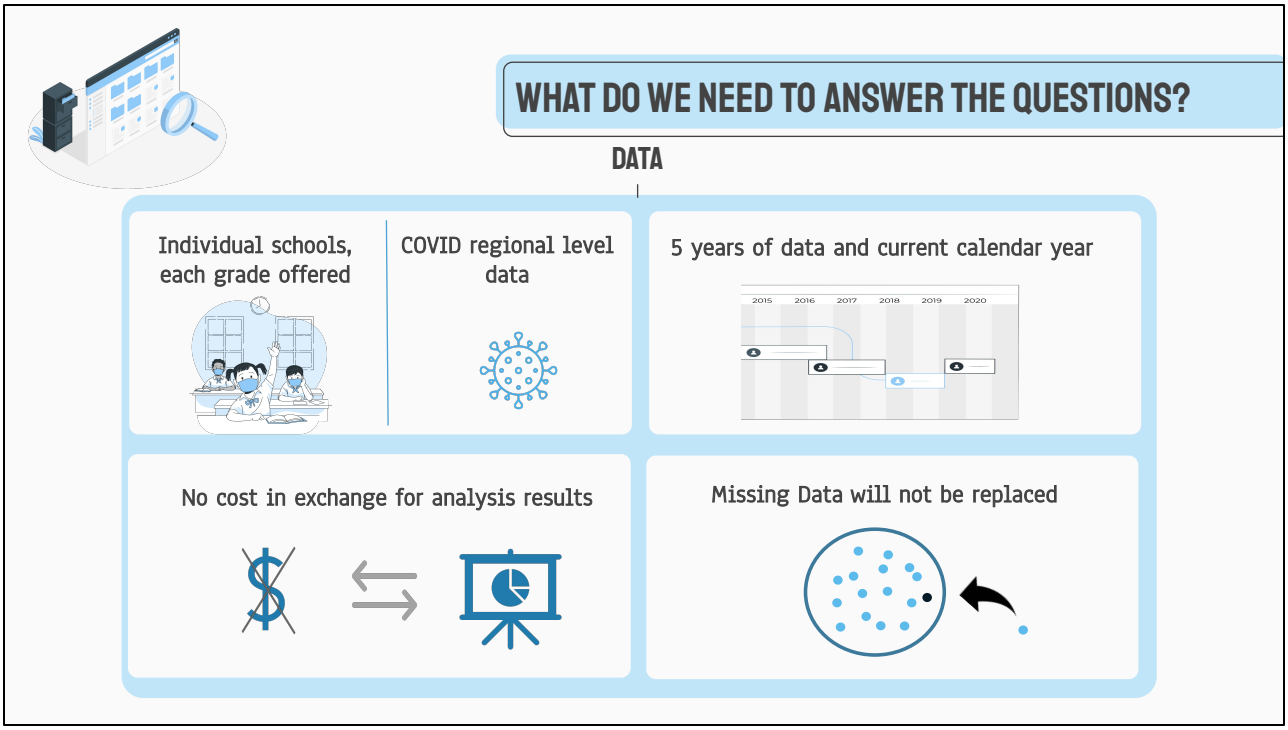
COVID-19 gave us the opportunity to see if this form of learning is viable, but the question is, could this be an option going forward? This study should continue post COVID, to validate any results and see what remote learning looks like when we are no longer in the midst of the ongoing crisis that sparked it. However, what we should and can start looking at now is, is remote learning a viable alternative to education for high school students?

Since every student does not have equal conditions to successfully perform academically, we must also ask, what types of students, or general demographics, perform well remotely?

And if fully remote is undesirable, would a hybrid approach, where students sometimes see faculty and each other, make a difference in performance?

There may be actions that the US education department should take if it is proven that remote learning slows the population's development. If remote learning proves viable, it could become a new normal for students going forward, providing them access to increased educational opportunities and better performance. On top of this, it could be a way to provide underserved populations and remote communities better access to education.

So how we will study this? I will turn it over to Natali to talk about what we need to answer the questions.

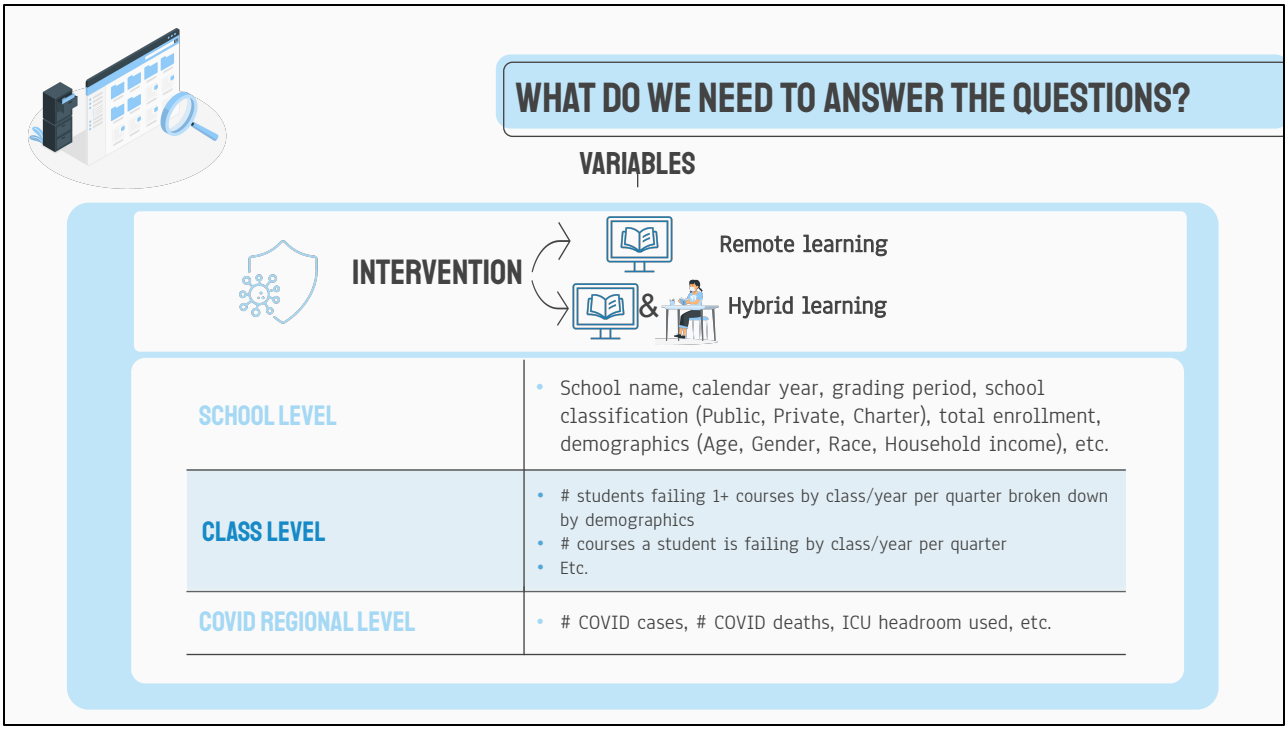


Thank you Heather.
What we need to do is to collect data from individual schools per each grade and year offered.
Help is needed from school data managers and administrators to generate relevant data for our study through their school management software. We are planning to obtain 5 years of historical data plus current calendar year. Data from different schools will be standardized into one format.

In exchange for their participation in this study, schools will be provided with the analysis results and therefore, the data will have no cost for us. Participation in this study is anticipated as this is an area of concern for many schools, however, if a school declines to participate, it won't be pulled out of the study and we may add additional schools to our sample to maintain statistical power.

We will also acquire COVID-19 data from publicly available data sources to draw conclusions on the impact of COVID within the regions and time period studied.

Data collected for this study contains information about minors.To protect privacy, we will ensure the school community signs a consent form detailing the type and usage of data to be collected. Data does not contain identifying information and will not be collected at individual student level.

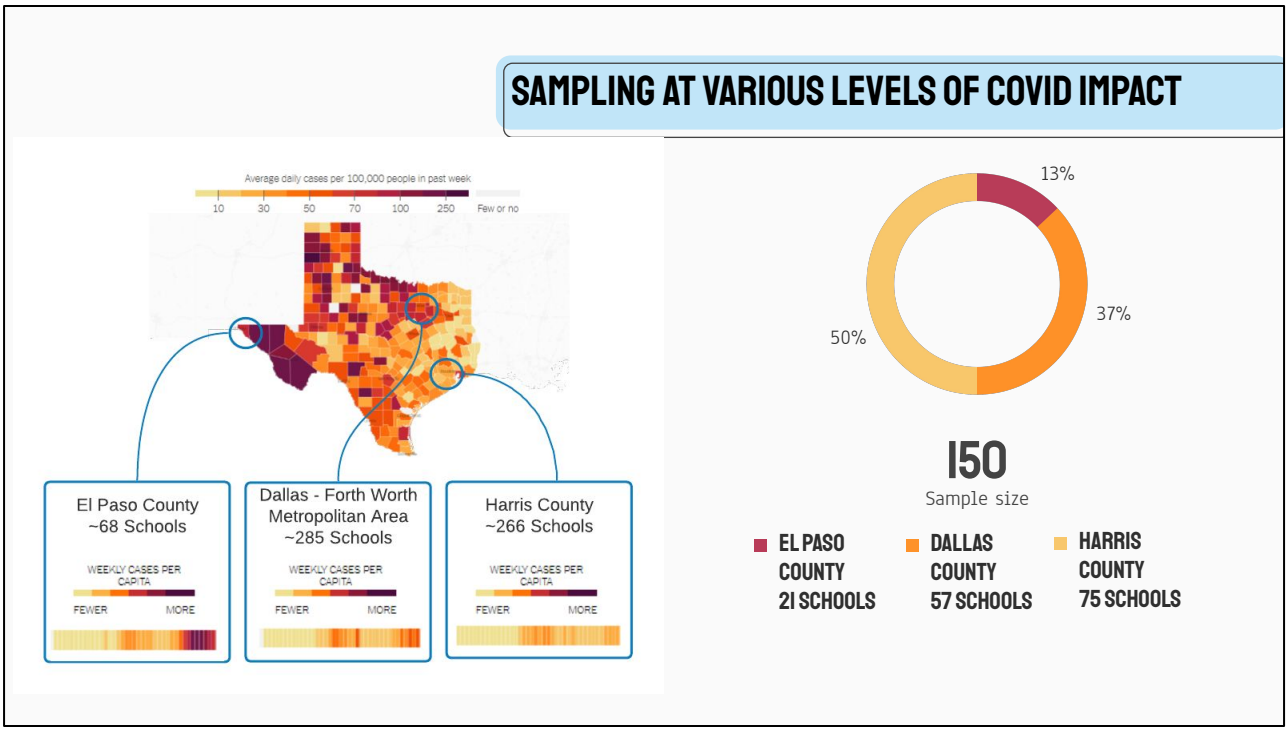


The interventions for this study are assigned by a random natural event, the Covid-Pandemic, forcing schools to switch to different instruction methods:

- Remote Learning: Online only method
- Hybrid Learning: Any mix of in-person and remote learning

The main Comparison variables (highlighted in blue) are: The number of students failing at least one course by class/year and how many courses the student is failing. Supplemental variables include but are not limited to: School level data such as school name, calendar year, school classification, total enrollment, demographic (Age, Gender, Race, Household income) and COVID regional level related variables.

Now, Aline will start telling us how we should go to answer the questions



To help us answer these questions we will be looking at three distinct regions within the state of Texas at different levels of COVID impact. We choose a regional stratification sampling method of the schools in the state according to the total student enrollment in metropolitan areas that are experiencing different levels of COVID impact.

We based our selection of metropolitan areas by following the latest Texas coronavirus map and case count. Since the beginning of the pandemic there have been over 1,300,000 cases and 23,000 deaths in Texas. From which Harris county, Dallas-Fort Worth and El Paso are the top three counties in number of total cases, however, as you can see in the heat map, the average daily cases per 100,000 shows these three areas represent different levels of impact.

Therefore, as a sampling frame we selected Harris-Houston county, Dallas–Fort Worth metropolitan area, and El Paso county as metropolitan sampling areas with low, medium, high COVID levels respectively.

Roughly the breakdown for the area:

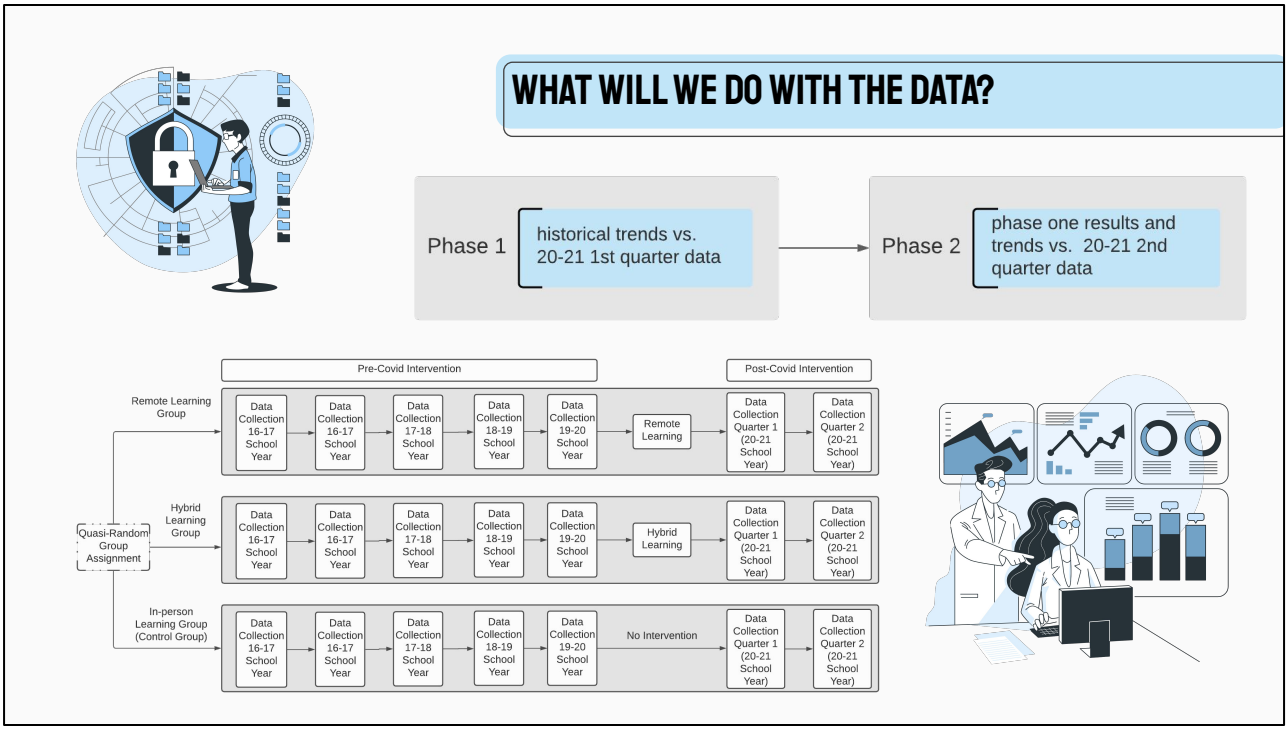
Harris County has a population of 4.7 million and we considerig about 270 schools. 50% of schools in the sampling frame will correspond to Harris county. This will represent around 58% of the population of sample.

The Dallas–Fort Worth metropolitan area represents a population of 6,3 million and has 285 schools. Around 37% of schools in the sampling frame are in Dallas county and represent ~32% of the population of sample.

El Paso County has a population of about 800000, with around 68 schools. Around 13% of schools in the sampling frame will correspond to El Paso county. This represents about 10% of the population of sample.

A list of schools and their quarter 1 learning mode will be generated for each county. A proportionate number of schools will be selected from each county to represent the portion of schools in each county as you can see in the donut chart. 21 schools from El Paso county, 57 from Dallas county, and 75 from Harris county. Within each county's sample a proportionate number of schools from each learning mode will be selected.

Texas Covid heat map and county breakdowns used to create image from https://www.nytimes.com/interactive/2020/us/texas-coronavirus-cases.html
Number of schools estimates from www.niche.com



As Heather mentioned earlier, COVID-19 has created a natural experiment in which we can study the viability of remote education. Leveraging the approximately random assignment of COVID-19 impact on communities and implementation of learning modes in schools we can study the changes in student performance within a school, within a COVID-19 community, and compare learning modes.

For this experiment, we look at two time periods, Pre-COVID-19 intervention, and Post-COVID-19 intervention. As shown in the bottom diagram, the Pre-COVID-19 intervention time period covers the 5 years of historical data from quarter 1 2016 to quarter 3 2020. The Post-COVID-19 intervention period includes data from the 20-21 school year after the learning modes were implemented within schools.

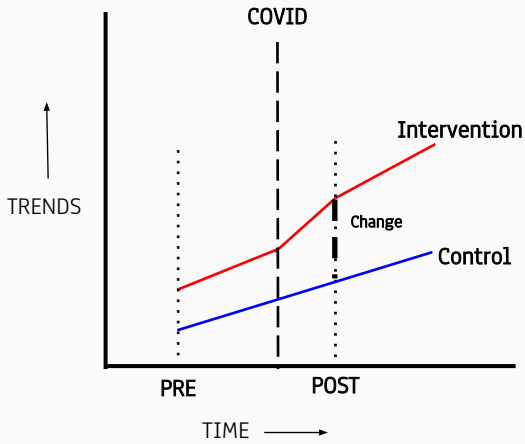
Quarter 4 of the 19-20 school year will be assessed to determine if it should be included in the study. We anticipate that this quarter will be an outlier and should not be included since the implementation of learning modes was inconsistent and a wide array of grading policies will impact data comparisons.

This study is broken into two phases. Phase 1 will focus on the difference in data trends between pre-COVID-19 intervention historical trends and data from the first quarter of the 20-21 school year. Phase 2 will expand on the results from phase 1 and looks at continuing or diverging trends starting in quarter 2. Due to schools

transitioning between learning modes an additional variable will be added during phase 2 to track schools as they transition between learning modes.

I will turn it over to Rakesh who will explain how we will analyze trends in the data.

GET THE DATA TO SPEAK FOR US



- Longitudinal Time-Series Study
- Analyse trends in Control Group and Intervention Group.
- Difference in Differences Technique - Highlights change and evolution in trends.

- Thank you Laurie,
- (slide check)
- What we saw so far
  - Study Design establishes processes**.**
  - How Data is Collected from sample.
- It also details a **careful choice of Control and Intervention groups**. This **isolates the effect of Intervention**.
- **Intervention here refers to a school adopting Remote or Hybrid Learning** in response to the pandemic.
- We now allow the curated **data to speak for us through statistical methods**.
- Each study group's data is structured as **longitudinal time-series data** and exhibits the trends for that group over time.
- For each research question, a **pair of longitudinal time-series are**

- **considered to draw comparisons**. One series for each group.

- This study intends to use a statistical-technique called **Difference in Differences -** this technique

- highlights how the **trends in the intervention group evolved after the intervention** when compared to the control group.

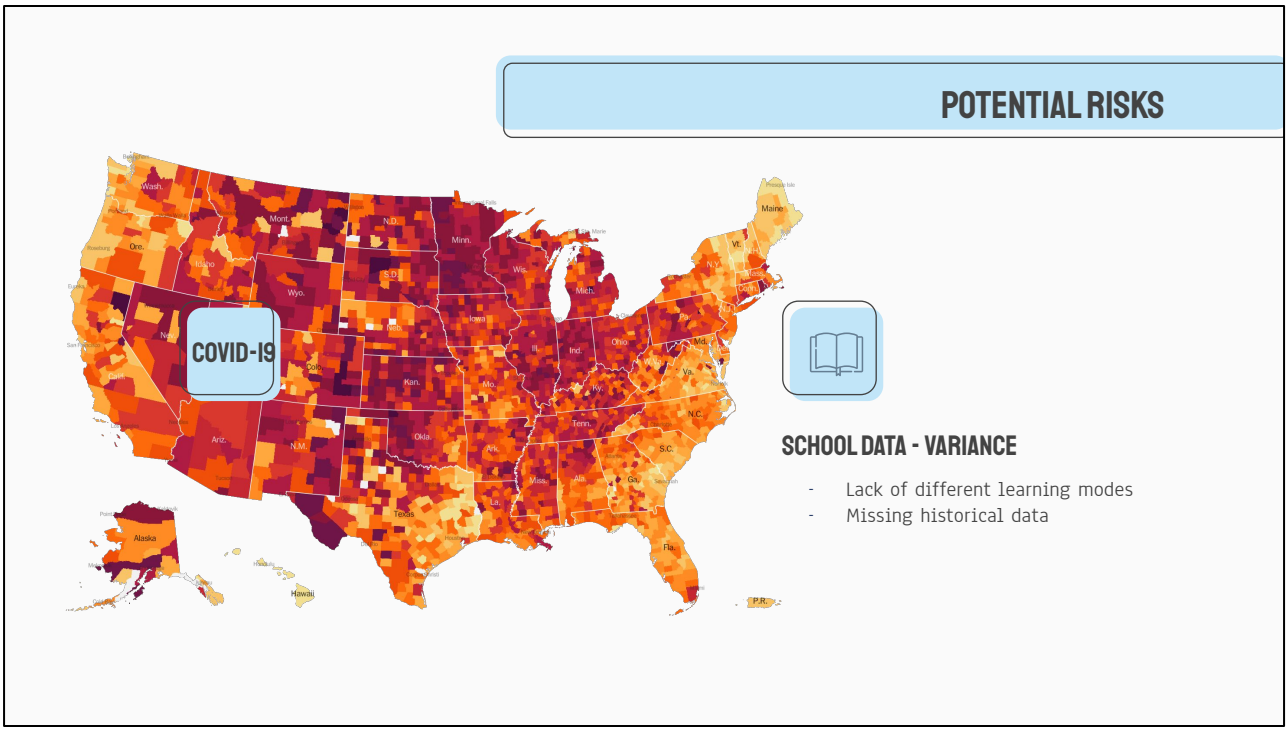
- The study also **analyses the discontinuity in the longitudinal data model**:
    - This gives an **idea of how the trends evolved since the intervention**.

    - The effect of this intervention could be a **constant shift effect**.

    - OR it could show a **temporal change** following the intervention which hints to a **continually scaling effect**.

- This study strives to communicate the statistical observations in **appropriate language so as to motivate action from the Texas Education Agency and other representatives**.

- I will now request Abby to talk about **potential risks identified for this project**. Thank you.



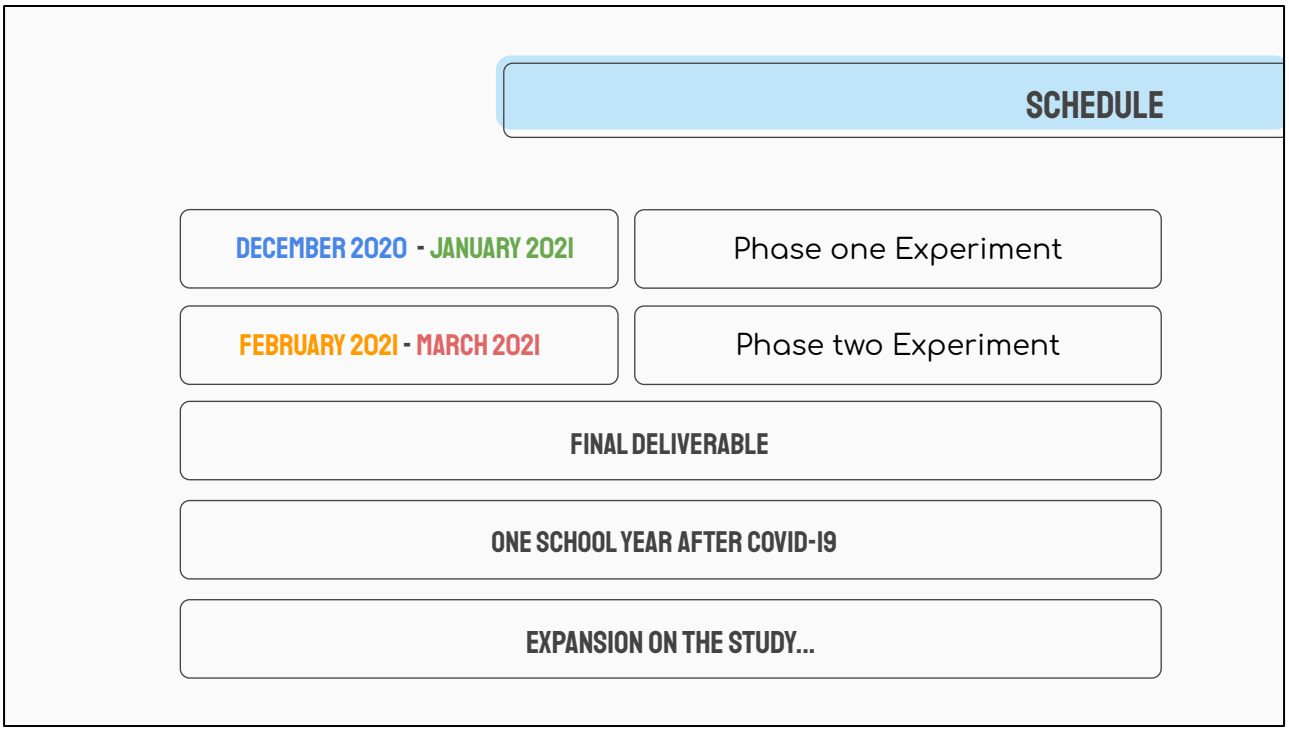
About potential risks associated with this study, I will include two main concerns here.

First of all, COVID-19 is the potential threat to validity: this random natural event caused the intervention to occur. It could cause, for example, students to get sick or not to be able to continue their studies because of their parents losing their jobs or health.

- Hence, in order to increase the validity, it is recommended that the study be continued after COVID-19 subsides for a more accurate comparison and analysis.

Secondly, Variations in Schools introduce bias to the research: Public, private, charter schools may be different sizes, while don't all have the same structure, like year round schools, trimester schools.

- For schools with different structures, evaluating them as a whole school year would be fair to avoiding time differences and study workload differences



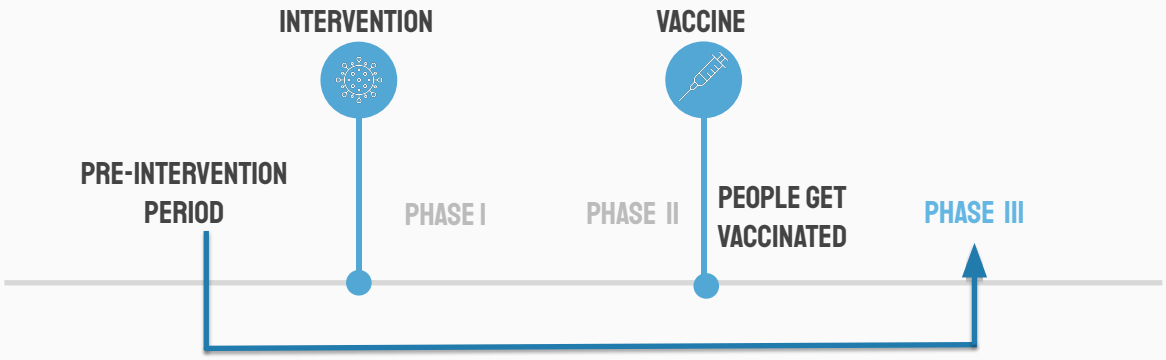
Speaking of the deliverables,
Here's the anticipated timeline for execution of research design:

For phase 1, collecting data and comparing pre-intervention to post-intervention data within each school will consume two months, from December 2020 to January 2021. Note that this analysis of trends is done within the same school groups

For phase 2, Evaluations within schools that have transitioned between learning modes will be conducted for the next two months to draw a conclusion for the final deliverable, which will be giving the result of efficiency of different learning modes, especially whether remote learning is a viable option in future education. And imply the difference that hybrid approach makes for student success.

HOW CAN WE EXPAND ON THE STUDY?

Phase III: To reduce the noise and bias that pandemic could introduce to conclusions about student's performance with different instruction methods.



Thanks Abby,
Ideally, this study should continue after COVID-19 settles in order to reduce the noise and bias that pandemic could introduce to conclusions about student's performance with different instruction methods.

Should phase I and phase II analyses show hybrid/remote learning are viable options, then schools could be more easily persuaded to continue piloting remote learning in a post COVID environment for a semester or year. The start of the study expansion would depend on when people get vaccinated in the studied regions.

Phase III analysis would be similar to phase II, but it would compare pre-intervention trends with continuing trends in a post-vaccine period. Pilots should be designed such that we have enough volume to track the distinct pathways. Since this would be a controlled experiment, bias due to outside influence would be less compared to phase II.

Depending on the results found in phase III, a strategy with different instruction methods could be designed so that education reaches more people and academic performance is optimized based among the groups analyzed.

(switch slides)

We hope that the results of the study can help you make decisions as the COVID-19 pandemic continues, as well develop new education initiatives in the future as student populations continue to grow. Thank you for listening, we appreciate your time.

THANK YOU FOR LISTENING

Does anyone have any questions?