Details of datasets in this folder are as follows:

1. Box Plots: a small exercise to represent variation of number of lessons completed in different domains across time.
2. Module Time Calculations: a collection of three approaches to calculate the time spent on a module by a user.
3. 2Pager: these are graphs to represent implementation of CLIx in terms of platform usage, for all the four states

Data Sources for 2Pager:

1. Raw data is syncthing data from July 2018 to Dec 2018
2. Processed data is modules data from [here](https://drive.google.com/drive/folders/1gM08yYhCtGifJIY5WRjsmABMT6aGKDtp)

Methodology or steps involved:

Platform data part of implementation report was generated based on the discussions between research team and implementation team specifically in the context of Mizoram (later extended to other states). Please refer to the python script in the same folder for more details of calculations.

Following are some of the points to be noted :

- Inputs from 7th Aug, 2019 call with Implementation team:

Need to show in the graph these many schools were involved, in each of these schools these many students are involved and these many days they are engaged. In these days, students together did so many modules(broken down by domain) and spent on an average so much time on modules(broken down by domain).

- Following is approach to represent information required:

We will have a grouped bar chart with grouping based on month.

In each group of charts, we will have:

1. number of schools engaged during that month

2. number of days engaged (breakdown based on server idle and active)

3. number of students engaged during that month (number of user logins)

4. number of modules attempted overall during the month with breakdown by domain.

To calculate the number of modules per domain: In a given month,

1. we get number of unique modules per domain attempted by students in a school

2. Average this number across all schools for each domain

5. Number of modules attempted by a student in a day with breakdown by domain

To calculate the number of modules per day per student per domain:

Similar procedure to that of finding time spent in a day is followed.

6. Time spent on modules in a day by a student with breakdown by domain

Following are the steps involved: For a given Month,

1. calculate time spent on module by a student in a day, based on timestamps and approximation that they might spend 0.5hr, if there is only one record.

2. Add the time spent on all modules of a domain to get time spent on domain by a student in a day

3. Average this across all students in a day to get typical time spent by students in a day

4. Average this across all days for a school to get typical time spent per day in a school

4. Average this across all schools to get typical time spent by students in a day in a state

This might look like a lot of averaging but average is the robust statistical estimate of the typical number of population at hand. Here we have averaging at higher dimensions. It is perfectly interpretable in statistics. Imagine three dimensional gauss distribution and mean of that distribution.