Details of datasets in this folder are as follows: :

Domain wise number of activities per **unique** user is calculated. Mean, median, SD and mode are also calculated. These are done for all states together, mz, ct, tg and rj.

Data Sources ::

raw - syncthing data from 2018 July to 2018 Dec

processed - modules data from [this](https://drive.google.com/drive/u/0/folders/1gM08yYhCtGifJIY5WRjsmABMT6aGKDtp) directory

Methodology or steps involved :

1. For each domain, do the following -
2. Find max activities\_visited for each user and unit\_name combination
3. Remove those logs which correspond to zero activities\_visited
4. Now sum up these max activities\_visited for each user to get all activities visited by a user in a given module
5. Get the total unique number of users for this domain
6. Sum up all the activities\_visited by these users
7. Divide ‘e’ by ‘d’ to get average activities\_visited per user
8. Find median, sd and mode for activities\_visited in this domain
9. Find the unique number of schools for this domain

Variables in Data:

domain - subject (e, m, s)

users - number of unique users

activities\_per\_user - average per unique user

activities\_sd - standard deviation

activities\_mode - Mode

num\_schools - number of schools considered to generate these numbers