**Ranganadh Narayanam. Submitted on: 10-04-2022. GUVI D20 Master Data Science.**

**TASK\_5: PANDAS.**

**Question:**

***Identify which factors predict future user adoption:***

As per the definition of user adaption I expected that “highest number of occurrences of same user Id” can cause more number of visits(that most frequent user\_id may correspond to adapted user definition:3 times per week). I developed pandas and python based programs for finding the frequency of each user ID from “takehome\_user\_engagement.CSV” and identified the list of most frequent user\_id s. Then basing on that I identified what is their “creation\_source”. I did it for the top 10 frequent users. For those most frequent “user\_ids” I found their most often creation\_source is “ORG\_INVITE”. This also matches with most number of creation\_source occurances “ORG\_INVITE” as per the following.

2163- GUEST\_INVITE

4254- ORG\_INVITE

2111- PERSONAL\_PROJECTS

2087- SIGNUP\_GOOGLE\_AUTH

Then also we need to verify the time stamp (but i am getting error for time stamp difference of 3 days, after taking their time string).

Then basing on this intuitive idea the frequency of “user\_id” along with “creation\_source” can be parameters fo future predictions. As there are many NAN in “**invited\_by\_user\_id”** so “Org\_ID” can also play a role in the future predictions. **Other than “frequency of user\_id, Creation\_source”, “Org\_ID” no other things can be a possible factors for future predictions.**