**Ranganadh Narayanam. Submitted on: 08-04-2022. 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. So I wrote a program from the "takehome\_user\_engagement.CSV” file to read. Then I wrote a program to find the number of occurrences of each user ID in that column using a for loop and forming list. From this list I found what is the max occurances and which is the user ID for that max occurrences. For that user ID I wrote a program for that user ID what is the “**creation\_source” (from shell 29 to 62), I found it to be “ORG\_INVITE” is being the most often. Then I found the next max occurring USER ID in that list. For that I found what is the “creation\_source”, it also found to be “ORG\_INVITE”. Then I found few more maximums I found first most occurances is**

**“ORG\_INVITE”.**

**Then “GUEST\_INVITE”. And so on.**

Then I found how many “PERSONAL\_PROJECTS “, “GUEST\_INVITE” , “ORG\_INVITE”, “SIGNUP”, “SIGNUP\_GOOGLE\_AUTH”. I found it to be coincidentally matching to the above result that:

2163 GUEST\_INVITE

4254 ORG\_INVITE

2111 PERSONAL\_PROJECTS

2087 SIGNUP\_GOOGLE\_AUTH

Highest : ORG\_INVITE, next : GUEST\_INVITE

The programs for these is from : Shell 5, 55, 56,57,10.

So the intuitive idea here is that more occurrences of user Ids in the file: takehome\_user\_engagement.CSV can be the solution for finding ,  identify  which  factors  predict  future  user adoption.