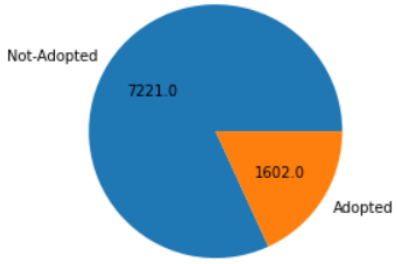
The main objective is to find the adopted users, adopted users are the ones who logged in to the platform at-least three times in one seven-day period.

The approach taken here is:

1. Filtration of unique IDs in the users_engagement table, with that all the duplicate entries will be removed.
2. All the date-time columns are converted to date-time format.
3. Since a single user have logged in multiple times, a count of each user_id login details with respect to date was generated using df.rolling() function of pandas.
4. Since we need data for seven-day window, use df.groupby() function to group the user_ids with '7D' login.
5. The main objective is to find the user who logged in at-least 3 times in a week hence use >=3.0 for the previous 7 day visit variable.
6. Create a column named adopted_users for categorical value (if user login is >=3, print '1' else print '0')
7. Count the number of adopted users. The below chart is the number of adopted users where 1602 users have adopted and 7221 users have not adopted. This analysis is for the total of 8823 unique user_ids in user_engagement file.



8. This adopted users columns was then merged to the main data file takehome_users.
9. After merging, the columns such as last_session_creation_time and adopted users have NA values, those columns were filled with '0's in the places of NAs.
10. To find the significance of each features to the adopted users conversions, a correlation was established with adopted users column and the rest of the columns in the table.

Observations:

1. Total entries in the takehome_users table was 12,000. Unique IDs in user_engagement file reduced to 8823. Adopted users count was reduced to 1602 which is just 13.36 % of the total dataset. Although the conversion rate is very low, it is required to find the influence of each feature contributing to the user adoption.
2. On establishing the correlation with adopted users and other features, the following were the results.
3. It is clearly observed that there is really no correlation with any of the features with adopted users. Particularly it is good to drop columns such as object_id, opted_in_to_mailling_lsit, invited_by_user_id columns since they make no sense to the data.

object_id	0.005750
last_session_creation_time	0.246048
opted_in_to_mailing_list	0.008097
enabled_for_marketing_drip	0.004653
org_id	0.064523
invited_by_user_id	0.000675
adopted_user	1.000000

4. The final say I have is that the number of features given here are not sufficient and really not relevant. More number of features are required to find out why the conversion rate is very low.