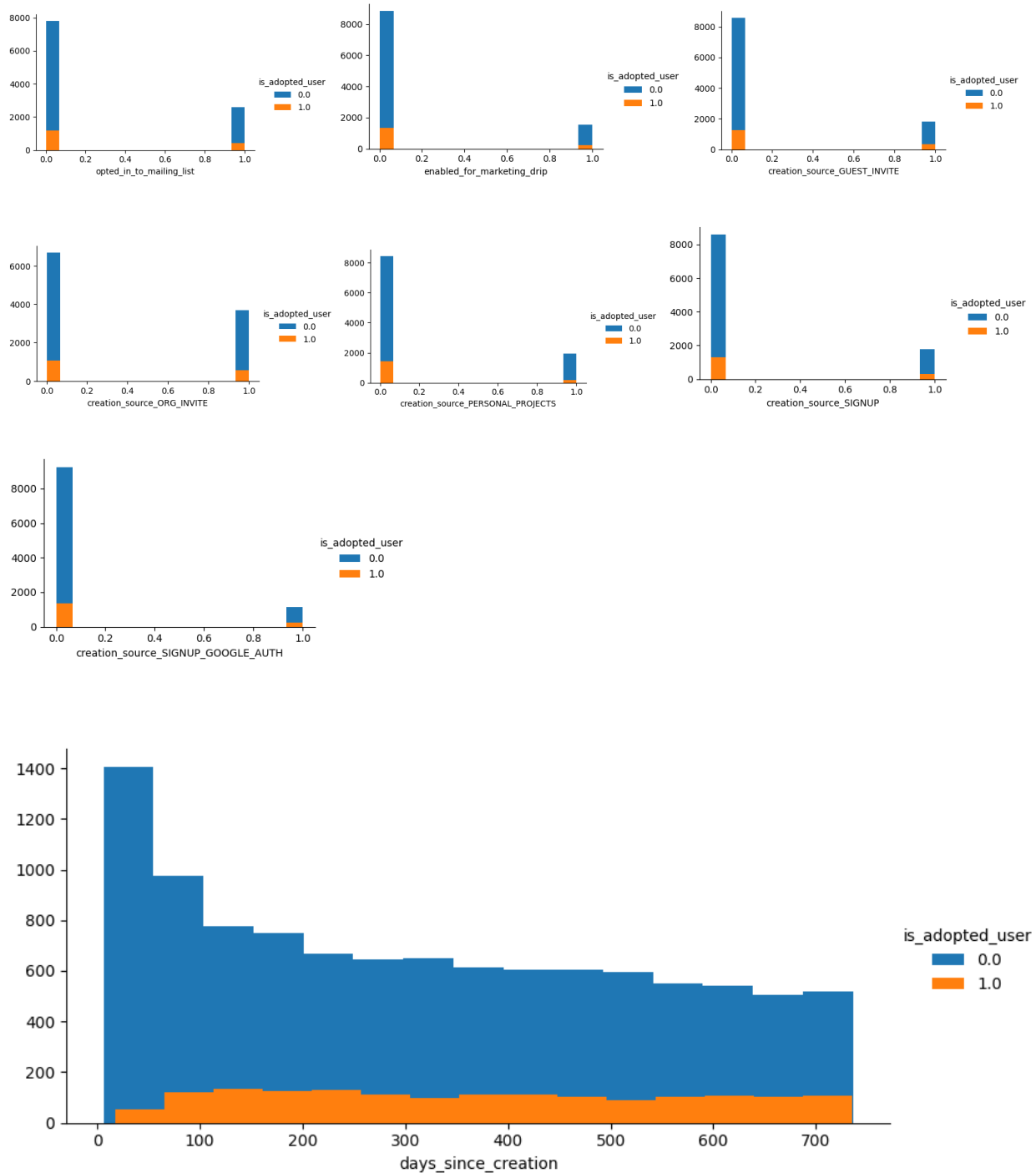


Aim: To identify which factors predict future user adoption.

My Approach and Findings: There are 2 csv files containing the data, one is user table comprises of 12,000 users and 10 columns. And another file (user engagement) has 2,07,917 rows and 3 columns.

- There are null values in '*last_session_creation_time*' and '*invited_by_user_id*' columns of user table. Converted *last_session_creation_time* column of users table to datetime format. There are no null values in the user engagement table.
- Imported the required libraries like pandas, numpy, matplotlib and seaborn to work with the data.
- Used 'groupby' method, rolling method and apply method to sort the adopted user meeting the defined criteria of an adopted user.
- Used merge function to merge the 2 dataframes. Before that dropped the unnecessary columns from the user engagement dataframe.
- The null values in the *is_adopted_user* and *last_session_creation_time* are filled in with 0 because we can assume that those users aren't adopted users.
- Tried to extract useful information from the email variable like its domain. There are 1184 email domains and most of them looked like fake domains so dropped the column entirely.
- For the *invited_by_user_id* let's convert the NULL values to 0 because the column has a Non-Null value only if the *creation_source* was a GUEST_INVITE or a ORG_INVITE anyways.
- For the *creation_time* column a column is added, which calculates how old the account is, i.e. the number of days since the account was created.
- These features are looked like important:
 - a. *days_since_creation*
 - b. *org_id*
 - c. *invited_by_user_id*
 - d. *creation_source_PERSONAL_PROJECTS*
 - e. *opted_in_to_mailing_list*



Conclusion:

We can see that the `adopted_user` class is pretty unbalanced because only about 13% of the total 12000 users are adopted.

Future possible work: We can also add a feature which calculates the difference between the creation date of the account and the first login of the user.

- D. B. Srihridai (imbhanu1996@gmail.com, dw28)