# Relax Inc. Soft Challenge

The datasets contained 8,823 user information and each row. There were 1656 adopted users out of 12000 users. The task was to investigate which factors are most important in determining user adoption. I found the length of user history was by far the most important factor. User history is defined as the number of days from account creation to the latest session. If a user was still using the service over >150 days after the creation of their account, it would be most likely that the user is an adopted, user.
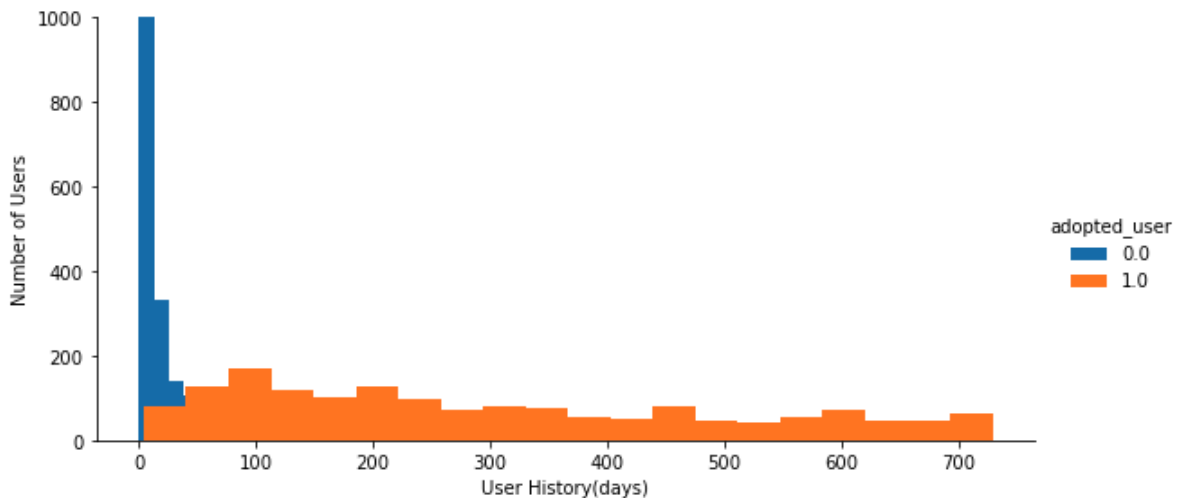
I found whether users opted into an email list or marketing drip had no impact on whether the use was an adopted user. A random forest model based on the dataset can predict user adoption with more than 94% accuracy. First of all, the most important factor is active days. It is an obvious result since the adopted user will have more active times compared to nonadopted users. The next important factor was whether or not the user opted into the mailing list. The next is the enabling for marketing drip.

```
active_days                  0.985807
opted_in_to_mailing_list     0.003000
enabled_for_marketing_drip   0.002596
ORG_INVITE                   0.002491
SIGNUP                       0.002215
invited_by_user_id           0.001771
PERSONAL_PROJECTS            0.001160
SIGNUP_GOOGLE_AUTH           0.000959
```

Based on the finding, I recommend that an effective way of growing adopted users might be encouraging existing users to log in and use the software after they had accounted for a while. The best method for making users stay online would be handing them over invitations for some events or rewards for making accounts which would cause curiosity users and want to

make them visit more. Encouraging users to invite others, and offering incentives for people to work on their personal projects using the software might also be effective.

I think it would be better for the data to contain more information about the organization's information such as demographics instead of organization ID. If there isn't any single information about the organization excluding the ID, it would not be a useful dataset in growing adopted users.