Data Analysis Interview Challenge – Relax Inc.

Task: Determine the factors or features that would help predict user adoption, in order to build a predictive model for user adoption.

Identify criteria for “adoption”
My definition of “adoption”
The service/software is a productivity or project management related, and a user login multiple times indicates the user is actively using the platform or actively used it in past. Therefore, a user who has logged in for more than 3 times a week, over a period of 2 recent months, or did use the platform at least 3 times in a week from the date of signup is considered an active user.

1. Initial setup

Load data
Load both data sets into the Jupyter Notebook and reviewed
Merge the two tables (users: object_id and engagement:user_id)

Groupe the rows by user_id
Explore the counts of the number of logins and the patterns – mainly time periods.

Create the target variable – “adopted” based on the criteria

2. Visualizations

Once I have obtained the login counts I may enhance the user table by adding the count column to the user table - takehome_users. This will enable me to perform analysis based on the user’s interactions.

3. Identify important features
    - creation time
    - creation source
    - opted into mailing list
    - enabled for marking drip
    - any associations with an organization – org id
    - invited by

Once I have determined the important features, I will use those features to identify patterns and insights.
Identify the last logins – to identify currently active users based on the last login:
Group users based on the important features, active time-frames.

4. Feature engineering

Derive new features based on user behavior such as usage patterns, user attributes such as creation source, marketing drip selection, mailing list.

5. Modeling

Build a predictive model, train and evaluate the model
Final model performance evaluation

6. Improve model.