Springboard Take Home Challenge: Relax Inc.

The goal of this assignment was to identify the factors which predict future user adoption, given a table of users and user activity. An “adopted user” was defined as a user who logs in to the product on three separate days within a one week period. Two tables were provided, the “users” table and the “engagement” table.

Adopted users were not given in the raw dataset. To determine which users were adopted, I used the following code:

```
adoption = engagement.sort_index().groupby('user_id').rolling('7D').visited.sum()
adopted_users = adoption[adoption >= 3].index.get_level_values(0).unique().to_numpy()
```

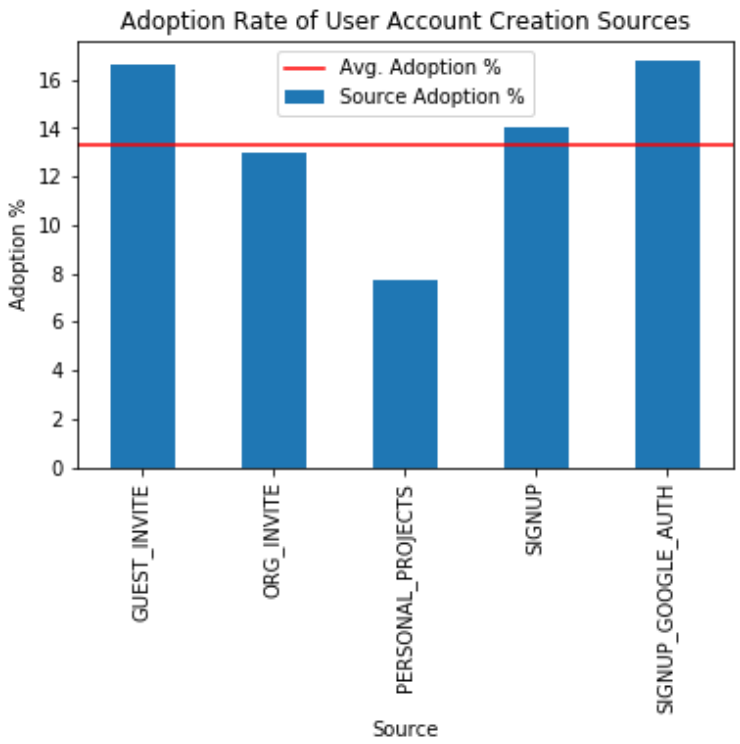
By grouping the engagement table by user, and then creating a rolling window of seven days, I could see the number of visits made by the user in that seven day period. I then extracted the users with three or more visits in a seven day period. These users were considered “adopted users.”

I then calculated the base adoption rate for all users. The percentage of adopted users among all users turns out to be roughly 13.3%. If any feature has a higher rate, then it may be worthwhile to investigate its effect on user adoption.

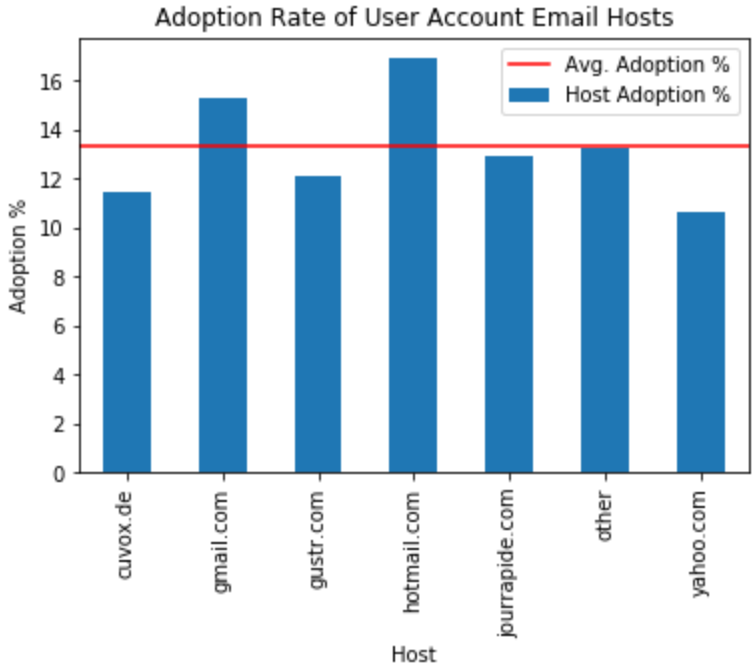
First, I looked at two seemingly obvious choices for boosting user adoption. If the user has opted into the mailing list, or has enabled the marketing drip, they might be more likely to become an adopted user. Surprisingly, the adoption rate for these groups is only 13.8% for both groups. This rate is barely above the base rate.

Some users were invited by other users. When invited by an adopted user, the adoption rate jumps to 20.6%. It would be worthwhile to focus on such users.

The users table provides a number of account creation sources, such as a guest invite or an invitation to a personal project. Looking at the rates for all sources, only the guest invite and Google Auth signup had rates significantly above the base rate. Both were above a 16% adoption rate.



For fun, I looked at which email host had the highest adoption rate. After discarding uncommon hosts, I looked at the adoption rate for the six most common hosts. Surprisingly, google.com and hotmail.com hosts had adoption rates higher than the base rate. It may be worthwhile to look into this connection with more depth.



Some areas that could use further research:

- Time since last session
- Organization adoption rates
- Confidence intervals for the above features
- A machine learning model that could combine features