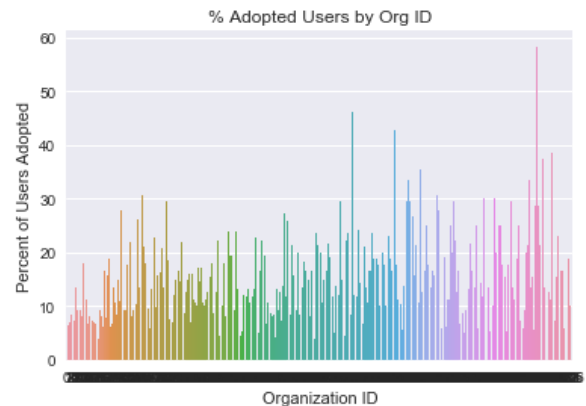
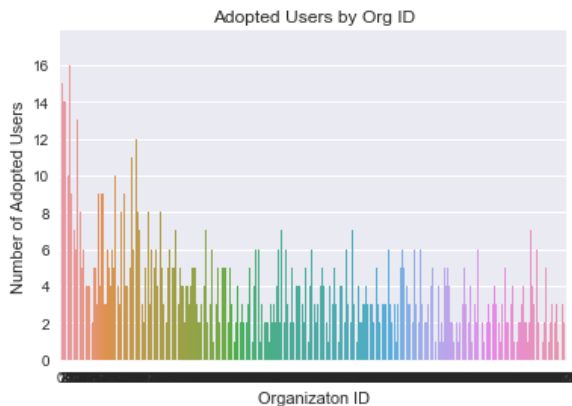


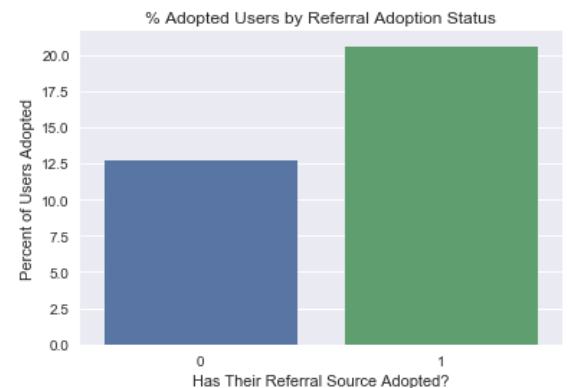
Relax Data Science Challenge: Solutions

The two factors that are the highest predictors of future user adoption are the organization (group of users) they belong to (`org_id`) and which user invited them to join (`invited_by_user_id`).



As can be seen in the plot on the right, there are certain organization IDs with far higher conversion rates: that is, users from these groups are far more likely to become adopted users. The groups with the highest conversion rates are 235, 270, and 387.

I chose to do a random forest classifier on the data to help classify and predict who is an active user. This classification model had an 82.6% accuracy rate. The most important features in the random forest classifier were clearly `org_id` and `invited_by_user_id`. The fact that `invited_by_user_id` is such an important feature in predicting user adoption suggests that there may be certain influencers. This is further demonstrated by higher adoption rates by users with referrals that were also adopted (see fig on right). Relax may want to connect with its adopted users, and in particular potential influencers, in order to harness the power of their influence and attract more regular users.



Additional useful findings include:

The percent of users adopting the product is similar among people who receive certain marketing efforts (mail, marketing drip). These are clearly not effective in upping adoption rates. I recommend adjusting strategy or reducing funds to these efforts.

