

Asana: Predicting Adopted User Factors

Lauren Hanlon

Introduction

The goal of this assignment was to analyze data on users from Asana to identify which factors predict future user adoption. An “adopted user” is defined as a user who has logged into the product on three separate days in at least one seven-day period.

Analysis

To predict what factors lead to user adoption I created a multi variable linear regression model. I used an OLS Regression model and Figure 1 displays the table of coefficients of each predictor when regressed against `adopted_user`, which was a 1 if the user met the criteria for being considered adopted and 0 otherwise. Results are as follows.

Creation source

Since the creation source variable was a categorical variable, I dummied it out and matched each source against `GUEST_INVITE` since that creation source type had the highest correlation with `adopted_user`. We see that the other creation sources have negative coefficient estimates with our response variable, indicating that `GUEST_INVITE` is more indicative of adoption than these other sources. When analyzing this further, I created a separate regression of `adopted_user` against these creation sources (Table 2). The coefficient estimate was most negative for personal projects at -0.1005334 and was minimally positively correlated with Google signups, but exhibited a p-value of close to 1, so we would not want to use this as a positive indicator.

Emails

User signup for mailing lists exhibited a positive correlation with user adoption. In analyzing further (Table 3), `enabled_for_marketing_drip` exhibited an estimate of 0.0068191, only slightly higher than `opted_in_to_mailing_list`. But again, we see high p-values with these estimates. Thus from this model we cannot make any clear assumptions about the relationship between email marketing and user adoption.

Time Deltas

Time deltas are defined as follows: `creation_delta` = time from account creation to last login and `last_login_delta`: time from last login to when this report was created (12/9/16).

From the model, we see the coefficient estimates with the time delta variables. The `creation_delta` exhibited the highest value, with a p-value approximately 0 indicating almost perfect estimation. Exploring this relationship further, I regressed only the time deltas (Figure 4) against `adopted_user` and found that `creation_delta` has an estimate of 0.8385773 and that `last_login_delta` exhibits an estimate of -0.1223327, indicating it has a slightly negative relationship with becoming an adopted user.

Conclusions & Recommendations

In conclusion, from this data the best predictors of future user adoption is our `creation_delta` variable and the creation source to be `GUEST_INVITE`. Again, the creation delta variable was constructed as *the time from account creation to the user's last login*. Intuitively, this makes sense as users who have been using Asana for a longer span of time are more likely to use the product frequently. As users login more, they are likely to

import more information and rely on the platform as a larger part of their day to day tasks, and thus are more likely to be “adopted.” The other predictor variable was creation source as guest invite. Surprisingly personal projects as a creation source did not perform as well as I would have initially thought, but there may be something about the exclusivity and excitement of a friend or colleague inviting you to join the platform that is enticing.

Thus, conclusions would be as follows: get more guest invites sent and keep users logging in frequently (whether this be via push notification, email, or marketing advertisement).

Further Exploration

Given these results, my next step in analyzing this data would be to explore the relationship of invites/creation source further as well as organization usage. Some questions I have from the data are, 1) *If a user was invited by another Asana user, what is the likelihood that a) that initial Asana user is an adopted user and b) that the new user becomes an adopted user?* Another question I have is 2) *Does the number of people in your organization make a difference in user adoption?* (i.e. is there a tipping point of X amount of people in your organization are using Asana that a user will be default become an adopted user?)

Figures

Least Squares Coefficient Estimates

Table 1: Coefficient estimates: ALL

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0	0.006	0	1
creation_sourceORG_INVITE	-0.013	0.008	-1.688	0.091
creation_sourcePERSONAL_PROJECTS	-0.016	0.007	-2.227	0.026
creation_sourceSIGNUP	-0.017	0.007	-2.313	0.021
creation_sourceSIGNUP_GOOGLE_AUTH	-0.013	0.007	-1.863	0.063
opted_in_to_mailing_list	0.001	0.006	0.165	0.869
enabled_for_marketing_drip	0.001	0.006	0.151	0.880
creation_delta	0.789	0.006	141.316	0
last_login_delta	0.039	0.006	6.401	0

Table 2: Coefficient estimates: Source

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0	0.009	0	1
creation_sourceORG_INVITE	-0.053	0.013	-4.185	0.00003
creation_sourcePERSONAL_PROJECTS	-0.101	0.012	-8.662	0
creation_sourceSIGNUP	-0.033	0.012	-2.810	0.005
creation_sourceSIGNUP_GOOGLE_AUTH	0.001	0.011	0.079	0.937

Creation

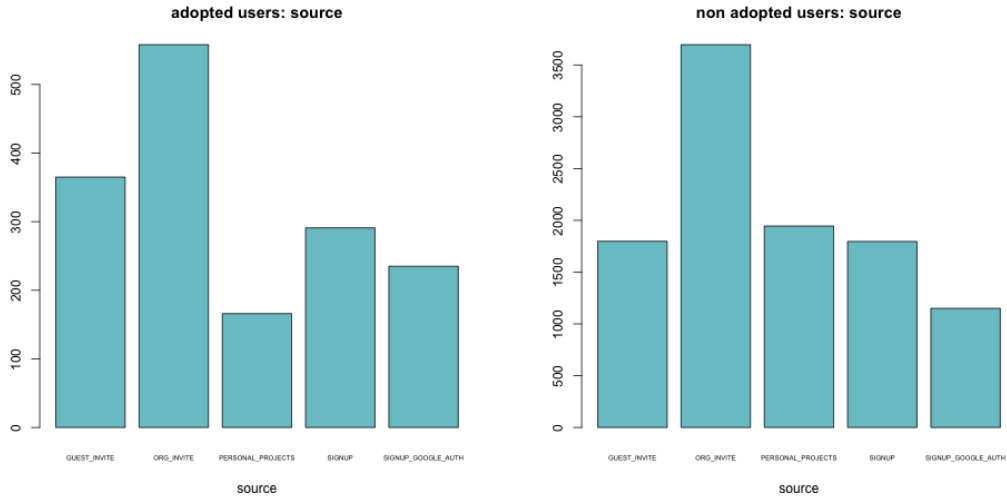
These first figures represent barplots of the relative frequencies of `creation_source`, broken up by adopted users versus non-adopted users.

Table 3: Coefficient estimates: Email

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0	0.009	0	1
opted_in_to_mailing_list	0.004	0.010	0.390	0.696
enabled_for_marketing_drip	0.007	0.010	0.654	0.513

Table 4: Coefficient estimates: Time Deltas

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.142	0.007	19.123	0
creation_delta	0.839	0.008	104.122	0
last_login_delta	-0.122	0.008	-15.189	0



Time Deltas

The figures below are boxplots of time deltas. Time deltas are defined as follows: **creation_delta** = time from account creation to last login, **lifespan_delta** = time from signup to when this report was created (12/9/16), and **last_login_delta**: time from last login to when this report was created (12/9/16).

