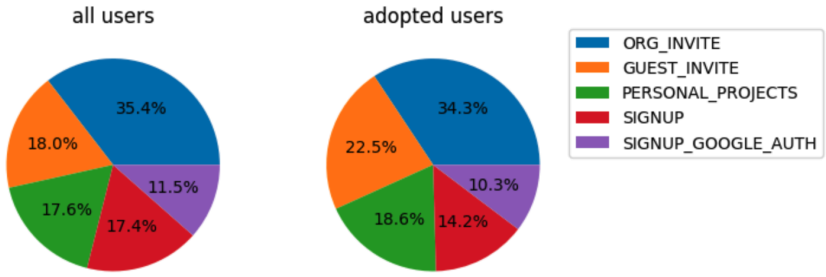
Relax Takehome Project Summary, Kealie Pretzlav

Data Exploration

The most easily identifiable differences in the data show up primarily in how the users signed up for the product. Guest invites appear to produce 4.5% more adopted users than general users. Interestingly, the general user population generally try the product via direct signup 3.2% more than adopted users, suggesting more people who find the product on their own don't end up adopted. Overall, 43% of all users were invited by another user, where 46% of adopted users were invited by another user.



There were no major differences in user adoption when invited by specific organizations or users. Marketing drip and opt-in mailing list user populations also did not vary significantly.

Random Forest Classifier Model

- Target: adopted user
- Features: user_id, creation_source (encoded), last_session_creation_time, opted_in_to_mailing_list, enabled_for_marketing_drip, org_id, invited_by_user_id
- Train/test Split: 0.75/0.25
- Train R^2 = 0.9998; Test R^2 = 0.91866

Results: Model prediction is pretty good, with very high accuracy of 92%, but recall is only 0.589 suggesting the model performs significantly better for predicted users that are not adopted. Feature importance suggests that the time since the product was last used is the most important predictor of whether a user is an adopted user. Figure on the right shows that adopted users tended to log in more recently, even if they signed up early in the product lifecycle.