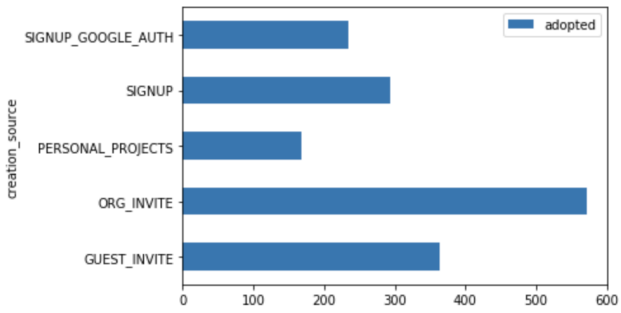
Relax Inc Take Home Challenge

by: Lauren Broussard | Springboard DSC

Problem Statement: *...Identify which factors predict future user adoption.*

Adopted Users: Using a rolling 7 day window to count user sign-ins (with the .rolling() method in Python), I was able to establish that the number of adopted users is 1631, or 14% of the dataset.

Analysis: I looked at adoption rates and counts of adopters by a number of variables, including whether or not an adopter was invited by a user, if they joined the mailing list, what group/organization they are a part of, and the day/week/month etc. A few things that stood out were that **of adopters, there were more that were invited by another user than not (22%), and more that were brought there by an invite from an organization/group (35%).**



Looking further, I used a Random Forest Classifier to train two models against the data to look further into important features for user adoption. In both models, **the org_id appeared to be the largest predictor** of adopters (up to 60% in one model).

Key Factors: The features that seem to be most important to predicting user adoption are:

- **Group/Organization:** Group/organizations 1,7, and 4 have the most adopted users. It would be important to look further into these groups - i.e. what is their function? How are they using the product? How are they communicating invites to the product?
- **Invited By User:** More users who adopted were invited by another user. This may speak to the power of word of mouth and referrals for the product. In a future analysis, it may be worth looking at the most influential users.

Further Research/Future Considerations: In the future, it might be useful to figure out if a certain email address is more likely to adopt (i.e. gmail accounts vs yahoo, etc.). **Additionally, when looking at date related signups, there was a dip in user adoptions in the month of May. It would be worth looking further into this to see if it is a seasonal trend or anomaly.** Lastly, there was also class imbalance (there are fewer adopted users than non-adopted users) in the model, so it would be important to look at this again when the adoption rate goes up, to see if the results hold.