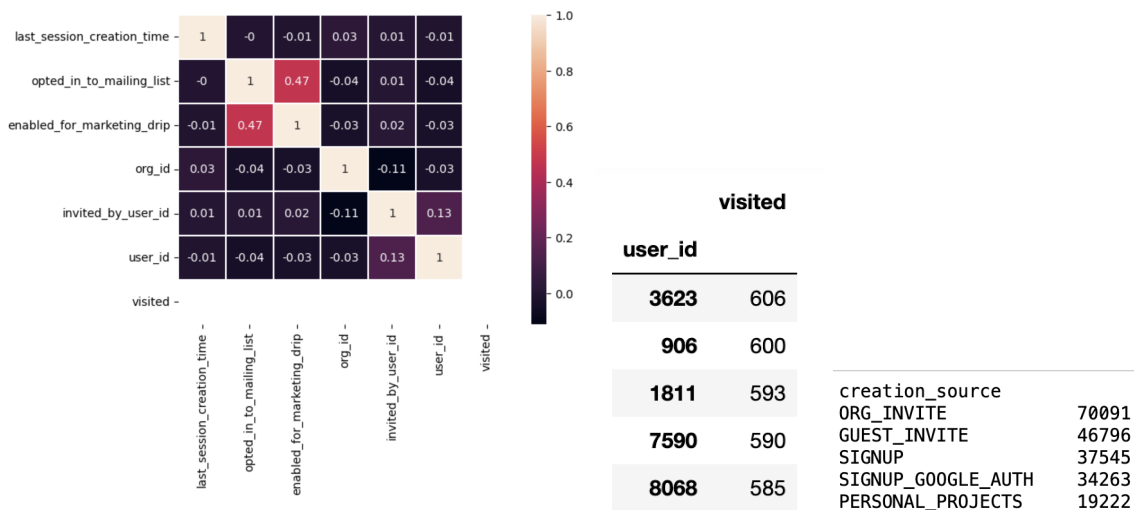


## Factors That Predict Future User Adoption

The purpose of this challenge is to analyze data on 12,000 unique users who have created a Relax account between 2012 and 2014 to identify which factors predict future user adoption. Of the 12k unique users, 8,823 logged in and navigated Relax's website. Adoption is defined as a user who has logged into the product on three separate days in at least one 7-day period. There are 2 data files provided (one containing the user information and the other containing user login information).

Upon merging both datasets together, the collective features are not too strongly positive to leverage for predictive analysis but it can still be done (see the heatmap below). As a result, predictive accuracy will take a hit understandably. The most significant features are object\_id, invited\_by\_user\_id, opted\_in\_to\_mailing\_list, and enabled\_for\_marketing\_drip.



After conducting data wrangling and EDA, it was found that the user adoption rate is 0.19 or 19% (2,248/12k) which is quite decent. with gmail as the dominant email domain at 35% followed by Yahoo at 16% (collectively 51%!). Of the 8,823 users who have logged in, the vast majority (6,235) did so only once. Meaning, a smaller subset of users represent the majority of users which is operationally fine in the short-term but user adoption growth avenues must be found in the long-term. The table above shows the top 5 users. By digging into the source of users created accounts, organic invites pace the group at a rate of 34% followed by guest invites at 23% (collectively 57%!). Both findings clearly articulate the value of targeting people who use gmail and targeting potential adopters via organic and guest invites. This information is valuable for the marketing team.

Going forward, future work could include more extensive EDA and modeling as I kept my work to the recommended time allocation of 1-2 hours. Tuning the top model or two also would enhance the value of the end product.