We are interested in the adopted users who logged into the product on 3 separate days in at least one 7-day period. We first start by merging the user table and the user usage table in order to find these users and perform Principal Component Analysis to find the features which give rise to adoption. The following steps were taken to produce the Principal Component Analysis needed:

1. Searching for 7-day periods, we found users that logged on in those periods.
2. Dropped personal data that was not relevant and filled in missing data with a placeholder.
3. Used one hot encoding to convert categorical data to numerical.
4. Performed Principal Component Analysis on the feature set by identifying the most relevant features.

Below are listed the sum correlations of the most important features according to the PCA.

creation_source_SIGNUP_GOOGLE_AUTH	2.375500
creation_source_SIGNUP	2.270659
creation_source_PERSONAL_PROJECTS	2.019568
creation_source_ORG_INVITE	1.959578
org_id	1.792770
creation_source_GUEST_INVITE	1.782990
last_session_creation_time	1.700941
enabled_for_marketing_drip	1.405251
opted_in_to_mailing_list	1.401398
invited_by_user_id	1.400076

After one hot encoding, the different categories of “creation source” were divided into their own features. Of the top 5 most important features, only one was not a part of the creation source was the organization that the user was a part of (‘org_id’). It seems like a drawback that all 5 different creation sources are listed so high up in feature important and it is difficult to draw a conclusion from why that would play a role in user adoption.