**Relax Challenge**

**Findings**

Some cleaning of the data was required to as time\_stamp feature was not in datetime format and the last\_session\_creation\_time and invited\_by\_user features were missing values. The time\_stamp feature was converted, the missing last\_session\_creation\_time values were replaced with the last time the user was known to have logged in (based on the user engagement table) or the creation date if last login time could not be determined, and the missing invited\_by\_users values were replaced with 0 (corresponding to no user). I also normalized the datetime features to make analysis easier. I calculated the correlation between each feature and the target variable to determine which features would would make the best predictors of user adoption. Factors that in predict future user adoption include the last time the user logged in, when the user was created, the user organization , if the user was invited by another user, and of the user opted into receiving marketing emails.

**Factors Considered**

The time the user has been active (the time between creation date and last login) may be a factor worth exploring. Also, I noticed that there were some duplicate name and email addresses. I considered that duplicate users may be removed and the most active user retained, but I did not attempt to do so in this exploration. Duplicate accounts may be an indicator of one long term user as it is not uncommon for users to have test or play around with.

**Further**

It may be worth one-hot encoding the categorical variables and looking to see if what specific categories may be strong predictors of user adoption. I would also like to see if the amount of time a user has been active is a strong predictor.