Objective: The goal of this analysis was to identify factors that predict future user adoption for Relax Inc users. Adopted users are defined as one who has logged into the product on three separate days in at least one seven-day period.

Data Preprocessing: We identified missing values in the invited_by_user_id and last_session_creation_time fields. These missing values were imputed by:

- invited_by_user_id as 0 since NA means no user invited them.
- Last_session_creation_time: with the creation_time, assuming NA means the user has not logged in since creation.

From analysis, we discovered:

- There has been an increase in user creation overtime with a spike in the last trimester
- almost 25% of the people do not opt in for mailing list and almost 16% are not enabled for marketing drip
- Regarding the org_id, there seems to be an imbalance on the number of people from the lower org #s (under 25 ID)
- It looks like there is no significant ID member who is inviting the most users, so we will use this column as a invited by user (bool)
- It looks like most people come by ORG_invite with an even distribution of roughly 18% for every other source.
- And finally and most importantly, it looks like 96% of the users are not adopted users.

When looking at correlation in the data, we identified a somewhat strong correlation between opted_in_to_mailing_list and enabled_for_marketing_drip. Interestingly, It looks like some users are enabled for marketing drip even if they did not opt in to mailing list, to capture this, we combined both columns into one.

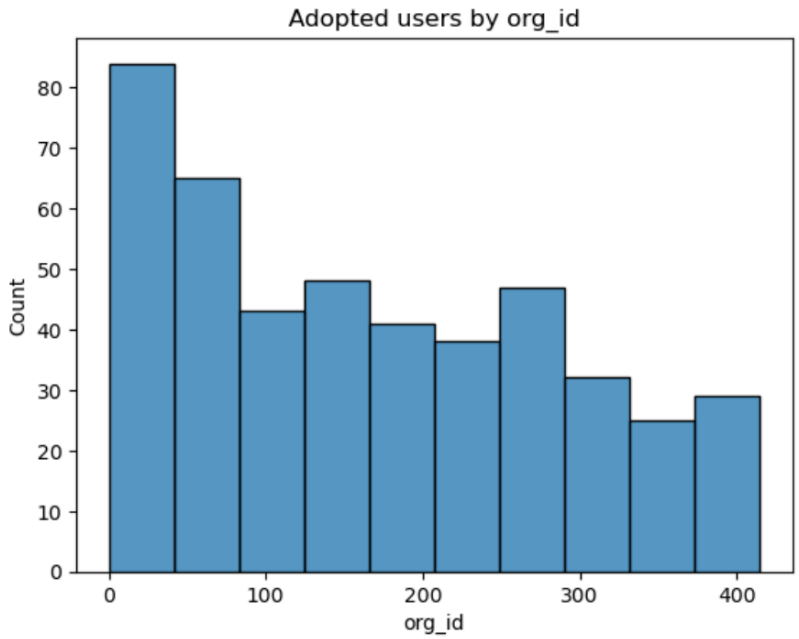
Finally, we transformed the ‘creation_time’ and ‘last_session_creation_time’ to days since the event to be able to better model the data.

Modeling: We employed logistic regression to model the probability of user adoption. The following features were identified as significant predictors of user adoption

```
Feature Importance (Coefficient and Odds Ratio):
        Feature  Coefficient  Odds Ratio
145  org_id_140     3.928593   50.835421
57    org_id_52     3.783498   43.969566
246  org_id_241     3.728318   41.609042
192  org_id_187     3.521867   33.847555
24    org_id_19     3.463408   31.925597
..          ...          ...         ...
59    org_id_54    -3.837857    0.021540
65    org_id_60    -4.216837    0.014745
395  org_id_390    -4.226955    0.014597
358  org_id_353    -4.897930    0.007462
134  org_id_129    -5.027397    0.006556
```

	Feature	Coefficient	Odds Ratio
0	days_since_creation	0.007746	1.007776
1	days_since_last_session	-1.521964	0.218283
2	creation_source_ORG_INVITE	-0.313088	0.731186
3	creation_source_PERSONAL_PROJECTS	0.082099	1.085563
4	creation_source_SIGNUP	0.328202	1.388469
5	creation_source_SIGNUP_GOOGLE_AUTH	0.053836	1.055312

It looks like the most important variable predicting the future adoption of the product is the org_id the customer belongs to, with org 140 being the most likely to adopt.



Insights and Recommendations:
We recommend the company concentrate on targeted onboarding, given that users who were invited by others, especially in an organization setting, showed higher adoption rates. This suggests that encouraging organizations to invite new users could increase engagement.

All the analysis was done on Jupyter notebooks and can be found here