**Problem Statement:** Relax Inc has an app that they created with data on 12,000 different users. The company wants to find out what features were correlated to clients becoming an adopted user which is defined as a user that has used the app on at least 3 different occasions in at least one 7 day period. They will use this data to make changes to their marketing and app to improve client retention.

Data Wrangling: For data wrangling, I wanted to make sure that the date was in date time format so that it could be more easily manipulated in the future. After this, I resampled the data to weekly instead of daily and grouped by User ID. This was so that I could see weekly trends in user logins instead of daily as an adopted user is defined by their weekly activity on the app and weekly sampling would be a more useful format. Each user with more than 3 logins on any week were given the value of adopted user and a new column was added to the users data frame indicating if each user was an adopted user or not. Lastly, the creation source feature was one hot encoded and split into a new column for each creation source for easier processing In the modeling section.

EDA: First, I wanted to see the distribution of logins per user to see what some of the more common total logins were most common. The chart shows that between 0 and about 80 logins were most common. Next I graphed the same distribution but for logins on a weekly basis since weekly login frequency was key in our definition of an adopted user.
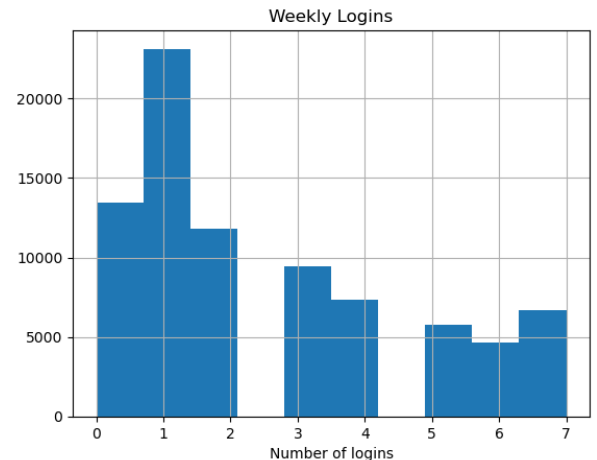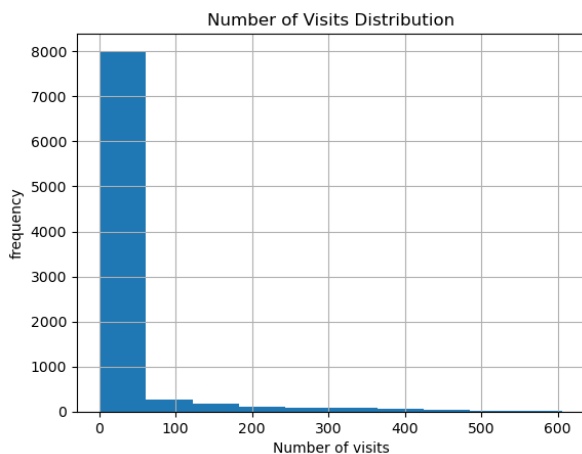
Both the logins by week and day both showed that over time, logins were increasing and capping out at nearly 5000 logins per week as compared to several hundred at the beginning.
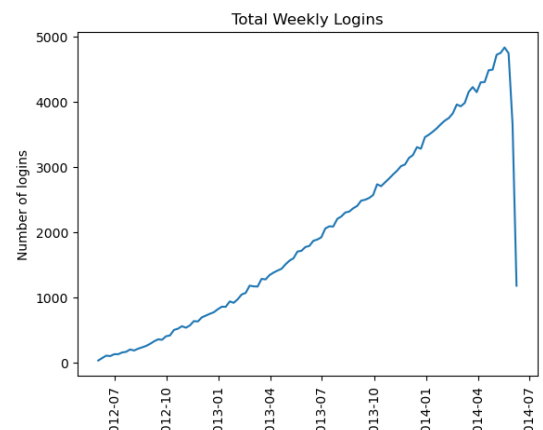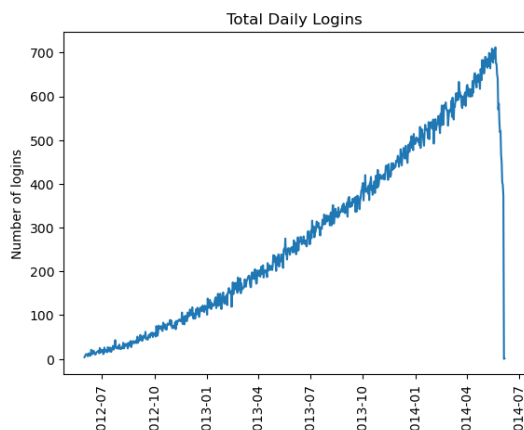
Creation did not seem to have a large impact on whether or not clients would be adopted or not. The only value considerably lower than the others in terms of percentage of cases was personal projects creation source with only about 7% of clients in this category later being considered an adopted customer with the rest falling between 11.5% and 15% with guest invite having the highest of 15%

Adopted Users by Creation Source
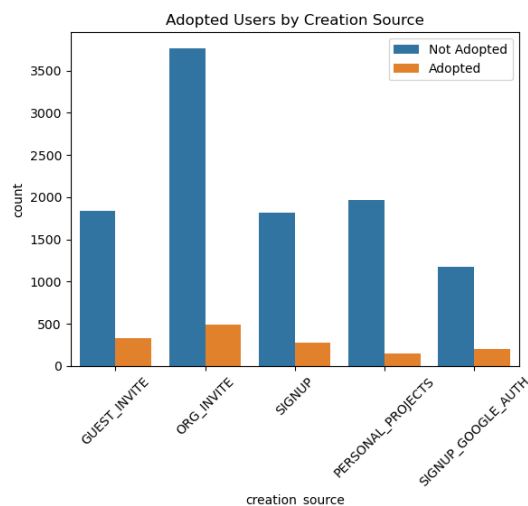
Recommendations: I would recommend changing the definition of what an adopted user is and then seeing if the features that suggest to lead to adopted users changes, or stays the same and recording the findings. Some suggestions are to see if users remained active over a time period rather than just using the app 3 times or more in any one week. You could search for users who used the app at least twice per week on three or more occasions or perhaps keeping the criteria for 3 times per week and not having a gap of usage of more than two weeks. There are many combinations that can be tested but those are just a few suggestions. I would also recommend using a grid search on the hyperparameters of the model to further tune the model and yield slightly better performance. This step will take minimal effort and is very likely to improve model accuracy.