# Take Home Challenge of Relax Inc

- Define target feature
  - Method: Using defaultdict imported from collections to store user and corresponding using time;
  - Result: 8823 out of 12000 users have engagement information, of which 23.4% are adopted users.

- There are two factors that strongly or moderately correlated with target feature
  - Number of visits in the first month: [Fig 1]
  - If login in and after 2014 April: [Fig 2]
    **It looks like there were some events or changes that happened in and after 2014 April. We need more data to verify if the event of change had long-term positive impact on adoption**.



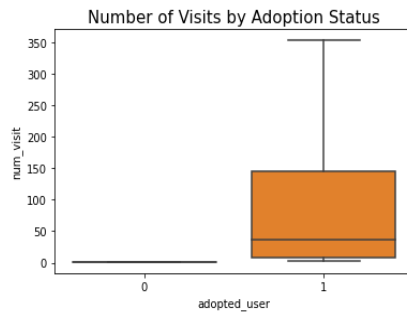Fig 1                                          Fig 2
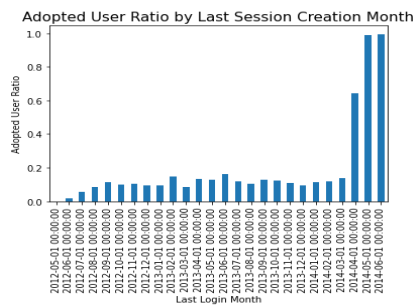
- Select important features with RFECV and RandomForestClassifer feature importance
  - Heatmap of features correlation: [Fig 3]
  - The feature selected out by RFECV: num_visit
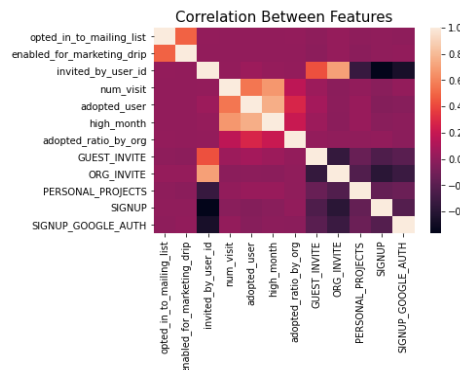  - Feature importance calculated by RandomForestClassifer: [Fig 4]
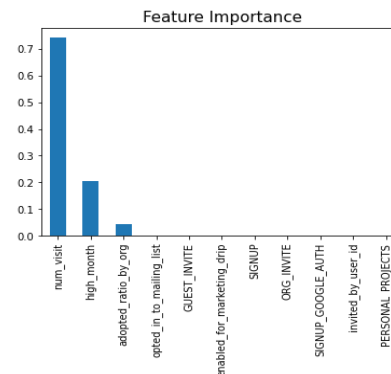


Fig 3                                          Fig 4

It's surprising to see the high_month feature is not selected as an important feature by RFECV and RandomForestClassifer.