**Free-Text**

*Please use Python or R to answer this question. Use the attachment icon in the tools menu to upload your response. If possible, please upload a single file containing your code and written answers, like a Jupyter notebook with markdown cells, or an R markdown file, or a PDF or html file. It is also acceptable to upload your code or notebook as a standalone file and to submit your written answers in the free text response space or as a separate upload.*. If you do provide your Python or R notebook or file, please also include a PDF of the file as well.

In order to improve user retention and lower churn, the growth team at Robinhood is interested in understanding why and which users withdraw money from their Robinhood account. A user is considered *churned* when their equity value (amount of money in Robinhood account) falls below `$10` for a period of `28` consecutive calendar days or longer after having previously been at least `$10`.

Using the datasets given below answer the following questions:

**a)** What percentage of users have churned in the data provided? **[4 points]**

**b)** Build a classifier that given a user with their features assigns a churn probability for every user and predicts which users will churn. How well does your classifier perform? State any metrics you deem important here. Based on the classifier output classify each user in the dataset as churned or not churned. **[5 points]**

**c)** List the most important features that correlate to user churn. **[3 points]**

Please provide the code and any explanation of your assumptions and methodology.

Datasets:

- features_data.csv - contains user level data such as:
    - *user_id* - unique id for every user
    - *risk_tolerance* - self-reported risk tolerance of the user
    - *investment_experience* - self-reported investment experience of the user
    - *liquidity_needs* - self-reported liquidity needs of the user
    - *time_horizon* - self-reported investment time horizon of the user
    - *platform* - which platform (iOS or Android) the user is on
    - *time_spent* - amount of time spent on the app
    - *first_deposit_amount* - $ value of the amount first deposited
    - *instrument_type_first_traded* - type of instrument first traded
- equity_value_data.csv - contains *user_id* and *equity_value* for user along with timestamps for days when the user's equity value is greater than or equal to `$10`. You may assume no user created their Robinhood account prior to 2016-08-16 and data is collected until 2017-08-18.