MVP Update

At this point I have created several models and a fairly automated process to create, test, and evaluate new models. The models I have created thus far are overall not very good at predicting. However, my code is well optimized to create, test, and evaluate new models with different features and model types. Therefore, I should now be able to focus on creating many more models with different features and finding the best one.

I have experimented with three different sets of features and for each unique feature set I ran a logistic regressor and a random forest classifier. Here are the results from the best model:

Cross Validation Means:

accuracy	0.64
recall	0.64
roc_auc	0.64
f1	0.42
precision	0.31

Test Data Results:

Accuracy: 0.66
Recall: 0.59

Confusion Matrix:
[[110013 52187]
 [16863 24059]]

ROC AUC score = 0.68