Project Outline

Start EDA

* Both working on cleaning data and understanding of data

We decided that the rejected didn’t add value, unless we figured out how to calculate their risk score (this seemed like a project in itself)

Thoughts of attacking project as an entrepreneur that wanted to take this over as Lending Club gets out of the business

Current plan:

Sophia – Working on Machine learning to find a model to maximize return

Paul – working on a simulator that will allow us to test those models on the actual data and on simulated data using a Monte Carlo simulation

Progress Report:

We have done our EDA and have cleaned the data, but with any project, we are still cleaning as we go.

Accepted loans are broken have field that indicates GOOD or BAD based on, Good = current, paid off. BAD is all the rest. About 86% of loans are good. We also know which are good based on year and sub\_group and can see the percent change as risk is increased

Working on how to find the mis-classified or safer loans in the riskier loan groups (D, E, F)

Classifier of borrower (risk assessment)

Risk\_score - - - -