Title: **Predicting Borrower Risk using Lending Club's loan data Augmented with Alternative data.**

Team Members

Full Name	Net ID
Neville Jos	njos2
Subhash Talluri	subhash2
Hamdan Siddiqui	abulhs2

Background

The current US system of lending has some flaws. Mainly, it is not very good at assessing the risk of borrowers (look at the 2008 recession for example). Most major banks in the US rely regularly on a singular value to assess the risk of a borrower, known as a credit score. A credit score is a number assigned to a person based on their ability to make payments on previous credit cards, loans, mortgages etc. Though that may seem like a decent indicator of risk, it has been notoriously erroneous in actually predicting the chances of default for a borrower, since it is extremely biased against poor people who have had one late payment or people with no credit history at all [1]. Banks, knowing that credit scores are bad indicators (while still using them), will then set a relatively high interest rate to cover the uncertainty of borrower risk. This leads to people having to pay for high interest loans even though they aren't risky enough to deserve that high interest.

Mission

Lending club[2] provides public access to data on their loans with attributes such as customer income, credit score, state, credit card utilization, home ownership etc. as well as information about the loan (loan amount, term, and whether paid back or defaulted). The dataset categorizes the risk of each borrower by their credit score and sets interest rates for them appropriately.

Our goal is to challenge Lending Club's approach by assessing borrower risk without taking a look at credit score. We will be using outside datasets and APIs to augment our existing dataset. For example, we will be pulling in Zillow to analyze average house prices for a zip code, consumer price index to measure cost of goods at the time of borrowing, unemployment rates from the US bureau of Labor Statistics, population density by county from US Census bureau and any other datasets we can get our hands on. Through this process of trial and error, we hope to find attributes about a borrower that show their true risk, which can be used to give out better, low-interest loans to those who actually deserve it the most.

References

[1] Gillian B. White (2017, January 10). *Can the flaws in credit Scoring be fixed? Not easily.* Retrieved from
http://www.theatlantic.com/amp/article/512702
[2] Lending Club. (2006). Lending club statistics. Retrieved February 25, 2017, from
https://www.lendingclub.com/info/download-data.action