# Kiva Funding Poverty Predictions

Team Members: Arsen Akishev, Quanhao Li

## Problem Description

The Kiva crowdfunding platform provides loans to multiple people throughout the world. Kiva understands that knowing the poverty level of the borrower is important in making decisions when it comes to their business. Thus, the issue that arises is how does a provider obtain poverty levels and welfare statuses without a rigid system that will inform the provider. Kiva is proposing a data science challenge that will implement modern techniques in order to predict these poverty levels.

## Approach and Methodology

Our method to solve this problem is focused around the online competition that Kiva is hosting on “Kaggle.com”. The competition revolves around a given main dataset and external datasets. External datasets are encouraged in the competition as paired predictors of poverty levels. Our approach in solving this prediction issue is to initially analyze external datasets provided by other competitors followed by seeking our own external datasets for analysis.

## Course Relevance

Since the project may involve using multiple large datasets, the course relevance comes into play when implementing heavy analysis that requires high performance through Python.

Aside from performance issues several other things come to mind:

- Prediction submissions must be done through jupyter notebook
- Data must be wrangled using numpy and pandas

## Expected Results

We expect to make a final model submission on **kaggle** that will use multiple datasets in order predict poverty levels. There are no expectations on the level of accuracy of our predictions however we expect to find multiple models with varying levels of accuracy.

## Schedule

Final submissions are due May 15. Since we are presenting to the class then our final submissions must be ready by the last lecture(02/05).

## Reference

https://www.kaggle.com/kiva/data-science-for-good-kiva-crowdfunding