From Wikipedia.org:

Kiva Microfunds (commonly known by its domain name, Kiva.org) is a 501(c)(3) non-profit organization that allows people to lend money via the Internet to low-income entrepreneurs and students in over 80 countries. Since 2005, Kiva has crowd-funded more than a million loans, totaling over $1 billion, with a repayment rate of between 98 and 99 percent. As of November 2013, Kiva was raising about $1 million every three days. Over a million lenders worldwide use the Kiva platform. Kiva relies on a network of field partners to administer the loans on the ground. These field partners can be microfinance institutions, social businesses, schools or non-profit organizations.

- **Problem statement:**

Understand what may cause a higher delay between the 'posted time' (the time at which the loan is posted on Kiva by the field agent), the 'funded time' (the time at which the loan posted to Kiva gets 100% funded by lenders), the 'disbursed time' (the time at which the loan is disbursed by the field agent to the borrower), and where. Subsequently, have a sense of its urgency and possibly discover hints on how they could resolve the issue.

- **Dataset**

The data from:

https://www.kaggle.com/lucian18/mpi-on-regions

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

https://www.kaggle.com/dcbaugher/selected-un-data-kiva#kiva_UN_country_data_selected.csv

Possibly, too:

https://www.kaggle.com/gaborfodor/additional-kiva-snapshot

https://www.kaggle.com/rgupta09/kiva-additional-data

- **How to solve the problem:**

Some possible steps:

1) Clean the data
2) Do an exploratory data analysis.
3) Make a predictive model of what may cause higher delays of proposed loans, if necessary also for different activities, sectors, or other characteristics that may derive from previous data exploration.
4) Build localized models to estimate the poverty levels of residents or needs in the regions where Kiva has these delays (part of the mission of Kiva is to “meet people’s diverse lending needs”) and also active loans so that Kiva can have a better understanding of the problem and a sense for the urgency of it.
5) Build predictive models of shared characteristics of major lenders.

- **Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis that they wouldn’t have done otherwise?**

The client is Kiva.org. The problem is important for them since this could be a way of understanding a fundamental step in its business process, and take actions wherever needed, and it could be a more efficient organization. Lenders would get their borrowed money faster and it could possibly be an incentive to lend again more times, and/or with more money. Depending on what insights the analysis of the data provides, Kiva can then think about what they will do. For example, changes could be made in e-mails they send to clients, create challenge proposals on teams who could be better qualified to tackle the specific delays or even create new teams since this has been proven to be effective (https://www.pnas.org/content/113/52/14944). It will depend on what the data ‘says’.

- **Deliverables**

All the code will be presented as well as an explanation of each step of the analysis.