

CaseStudyPhase1

Samiyanur Islam, Cindy Chi, Jimena Martinez-Dominguez

2023-10-27

Contents

Questions: Problem, Data & Their Relationship	1
--	----------

Questions: Problem, Data & Their Relationship

Having been grounded in the basics of peer-to-peer lending and specifically on LendingClub as a platform, it is time to start thinking about the problem and the data at hand. Answer the following questions, which will guide your thinking to that end. Admittedly, the questions are open-ended and there might be multiple possible answers to each. The goal in Phase I is to ensure that you brainstorm and ponder over the problem, the data, and solutions as much as possible. The grading will be adjusted accordingly.

1. (i) As an investor, what are the decisions you would need to make? (ii) Which of those decisions can you make using the available data from LendingClub and which one(s) would require additional resources?
2. (i) What is your objective when making those decisions in Q1? (ii) Explain how you would be able to distinguish “better” decisions from “worse” ones using the data?
3. Note that loans are temporal entities (36 or 60 months-long term). Different loans could default at different times; some will default soon after approval, some much later. Some, on the other hand, might be repaid early, before their term ends. Would these facts affect your downstream analysis and decision-making? How/Why?
4. Based on the discussions thus far, do you think historical data would be helpful? In which ways could you use such data to help make the decisions of your interest?
5. Next you will take a look at the data.
 - (i) Write down a high-level description of the different features—that is, the variables describing the loans. How would you categorize these features? (Note that there may be multiple ways of categorizing the features; think in terms of the source of the measurements, the type, and temporal characteristics.)
 - (ii) Just based on the feature descriptions, give an example to features that are likely to be (strongly) related if plotted on a scatter plot.
 - (iii) Which do you think are most valuable to an investor like yourself?