Loan Repayment Challenge

1 Introduction

At MoneyLion we are constantly working to assess the risk of our applicants more accurately. Being successful in this objective enables us to better price customers and control losses on our portfolio of loans.

This challenge is intentionally meant to be open ended. The most important thing to remember is that we are evaluating your thought process and ideas! The more you explain your thinking, in a clear and precise manner, the better. If you get stuck, describe what additional information or data you might look to collect, and trying a different idea is highly encouraged.

2 Problem

We would like to predict the loan risk or quality (loan repayment) on a given applicant.

Remember to consider your ideas and work in a business context. Think about what you are trying to solve and why? You could start with more simple analyses as it may inform your direction. Perhaps you may get a few insights from the data that could be useful to the problem at hand?

3. Guidance and tips

The Data Scientist position at Moneylion is extremely competitive and we receive many applications, so do consider how you could **make yourself stand out** from the rest of the candidates. Here are some skills that we're looking for:

- Data preparation, exploration and visualization.
- Telling a story using data
- Clean coding style
- Clearly communicate your thought process
- Back your assumptions with evidence.
- Sensible modeling

4 Data

You are provided with 3 files: `loan.csv`, `payment.csv` and `clarity_underwriting_variables.csv`.

Detailed descriptions of the columns are available in the dictionaries.zip file.

5 Rules

You may use any language, packages, or external libraries for the challenge, though Python or R are preferred. An IPython notebook might be the best way to communicate your thoughts to follow along with the code.

You may refer online for technical information but **DO NOT** directly use any help from other people, sources, online forums, etc., your submission should be solely your ideas and work.

There's no expectation on the amount of time you could or should spend on the challenge. That said, do share how much time you spent on it.

6 Deliverables

A zip file with contents grouped into the following sub-directories (you may omit empty directories):

- data (only if you add any new ones; DO NOT send back the original assessment data!)
- notebooks (we appreciate if you **include a html file of your notebook** as well as the raw file)
- code (if you use any custom new helper functions)