Columbia University
STAT GR5293-004
Fall 2020
Financial Technology and Data-Driven Innovation

Unit 2
Loans: Finance Foundations & The Structural Model of Default

Problem Set 2: Introduction to Loan Products

Dr. Margaret Holen, Instructor

September 18, 2020

Instructions:

- **Tools:** This problem set will be accompanied by course notes that provide context, along with required data files and Python and LaTeX "helper" notebooks.
- **Submissions:** Please submit in a single type-set pdf, following carefully the steps below
    1. Every new problem begins at the top of a page with problem numbers clearly labeled.
    2. Solutions should span three components:
        (a) A summary in English, clear sentences or bullet points,
        (b) A description of the technical work in equations, with excerpts of code or pseudo-code *as relevant*.
        (c) Numerical and graphical output or results *as relevant*

    * Please do NOT submit code, Jupyter notebooks or uncompiled LaTex or Tex!
- **Collaboration:** Collaboration is encouraged, but write-ups should be your own individual work, unless otherwise indicated in the assignment.
- **Advice:** The problem statements invite you to consider the application context for each problem, which makes them longer than they would be otherwise. We encourage you to spend more time thinking through your approach before starting problems so you can focus your efforts and answers.

Problem 1: Installment Loan Terms and Trade-offs

- This problem relates to the difficulty consumers have in making choices among loans, and the trade-offs between different loan terms. It also reflects choices you and/or members of your family may face....
- Suppose a cousin of yours just moved from New York City to the suburbs and is shopping for a car, using a car loan to finance the purchase. You have told them about this class, and they turn to you for help.
- Your cousin wants to buy a $\$25,000$ new car, a Subaru outback [1]. You have helped identify different loan offers[2], and you found the following rates.

Loan Term	Rate from Lender 1	Rate from Lender 2
36 months	3.00%	3.50%
60 months	5.00%	4.50%
72 months	3.00%	2.00%

(a) For the lowest rates available in each maturity, you calculate the following for your cousin. (Please specify the formula you are using for each item, and provide a table of values.)

(i) Monthly payments

(ii) Total interest paid over the life of the loan

(iii) DTI (debt-to-income) change for your cousin assuming an income of $\$50,000$ annually

(b) Based on your calculations for part (a), what loan would you recommend and why?

(c) Your cousin heads to a car dealer to purchase the car and finalize the loan. The dealer will be eager to not only sell the car but also arrange the loan. You want to prepare your cousin to negotiate with the dealer. To do that you show your cousin some specific examples, and then discuss the results. (Answers here will include equations used to compute, the result of computations, and a general conclusion you would want your cousin to make.)

(i) **Rate choices** *[Fixed maturity, Fixed principal]*: Suppose the dealer offers to beat the interest rate on the best loan you found by 0.10%. What is the savings your cousin will get each month, and over the life of the loan, for each of the maturities?

(ii) **Principal & monthly payment choices** *[Fixed loan rate, Fixed maturity]*: Suppose the dealer offers to round up monthly payments in return for extra features on the car. What would the principal amount be for your cousin if they agree to pay $\$25$ monthly for a 36-month loan vs. for a 72-month loan?

(iii) **Principal amount choices** *[Fixed monthly payment, Fixed loan rate]*: The dealer wants to sell your cousin a more expensive car and offers to extend the repayment term to 84 months, at the same rate as the best offer you found for 72 months. How much more principal would your cousin be borrowing under those circumstances?

[1] See: https://www.subaru.com/vehicles/outback/index.html

[2] You used a search engine and found sites like https://www.primefinancialcu.org/loans/auto-loans/ and https://www.easyautolenders.com/

Answer guidance:
Tools: This problem involves evaluating formulas over multiple different input values, and you should aim to use either Python or a spreadsheet tool to facilitate this work.

Numerical Answer Templates:
For part (a), please present the equations you are using, calculate numerical answers and share them in a table form, with a column for each part as shown below, and finally discuss your observations.

Loan Term	Lowest Rate	(i) Monthly Payment	(ii) Total Interest	(iii) DTI change
36 months				
60 months				
72 months				

For part (c)(i),numerical answers will correspond to entries in the following table:

Loan Term	New/Lower Loan Rate	Monthly Payment Savings	Total Interest Savings
36 months			
60 months			
72 months			

For part (c)(ii) numerical answers will correspond to entries in this table:

Loan Term	Loan Rate	Original Monthly Payment	New Monthly Payment	New Principal
36 months				
72 months				

Part (c)(iii) involves only a single answer, no table!

Problem 2: Lending Club Loan Examples

- In this problem, we ask you to look examples of LendingClub loans from one another data set, this one directly from LendingClub, in a file provided, "LendingClub_2016Q1.csv", consisting of 36-month loans. Unlike the data set considered in the previous problem, this one does not contain FICO scores, but it contains a wider range of loans categorized by LendingClub's own 'Grade' system.
- NOTE: All the loans in this data set are 36 months, i.e. There are 36 installment payments.
- This problem asks you to examine several individual rows of loans and to examine trends over loan grades.

(a) Individual loan inspection, part 1: Print the information for the first 20 loans. Looking only at the variable **grade**, pick the first four loans with different grades, i.e. the first grade A loan, the first grade B loan, etc. Then, compute the total interest paid, and note the loan amount, interest rate, monthly payments due, and loan grade.

(i) Equation for calculating the total interest paid, and Numerical values corresponding to selected loans. Present these results in a table with rows and columns like those below. ***NOTE: The table in your solution can - AND SHOULD - be output directly from your Python notebook or from whatever tool you are using to compute the answers. Table below is just a sample, and the provided notebook helps produce one with a similar form.***

Loan Attribute (field name)	Example 1	Example 2	Example 3	Example 4
Loan amount ("loan_amnt")				
Interest rate ("int_rate")				
Monthly payment ("installment")				
Lending Club Grade ("grade")				
Lending Club Sub-grade ("subgrade")				
Total interest due (*calculated*)				

(b) Individual loan inspection, part 2: Now check the Jupyter notebook. We have presented you several methods for sampling loans: from a "lazy" sample that explores the very first four loans to a more sophisticated one that investigates some of the loan features. Propose an alternative method for selecting specific loans, answering a question you may find relevant. You do not need to implement this sampling method, but feel free to do it.

(c) Overview of LendingClub Grades: Group lending club grades and for each of them, compute the percent of loans in the data set for that grade. Compute the mean for the following variables within each group: DTI, Annual Income, Installment, Loan Amount, Interest Rate. Finally, please comment on how borrowers and loan attributes vary across grades.

Grade	Pop. Density	DTI ('dti')	Annual Income ('annual_inc')	Loan Amount ('loan_amt')	Interest Rate ('int_rate')	Installment ('installment')
A						
B						
C						
D						
E						
F						
G						

Problem 3: LeandingClub - Probability of Default, and Loss Given Default

- This problem entails approximately calculating probability of default and loss given default across a LendingClub dataset.
- – It also points to how the context matters in how one answered the question: a careful financial analysis of what happened can differ from a statistical question about how what happened might translate into the future....

- Lending Club categorizes loan by grades. Loan grades represent LendingClub's estimation of borrower risk and lead to interest rate assignments. In this problem, we'll examine the returns across grades.

Instructions: For each grade, calculate the following, with your own script or with the help of the sample notebook provided:

1. Average interest rate for loans in that grade
2. Realized default rate for loans in that grade, i.e.

$$PD = \frac{\text{Total Count of Defaults}}{\text{Total Count of Loans of that grade}} = \frac{n_d}{n}$$

   i.e. n to denotes the count of loans in each grade, and n_d the number of defaulting loans....

3. For non-defaulting loans, the Average Return calculated as follows[3]:

$$\text{Avg Return if No Default} = \frac{1}{n - n_d} \sum_{i}^{n-n_d} \frac{\text{Total Loan Payment}_i}{\text{Loan Amount}_i} - 1$$

4. For defaulting loans, compute the average Loss Given Default defined as

$$\text{Avg Loss Given Default} = 1 - \frac{1}{n_d} \sum_{i}^{n_d} \frac{\text{Total Loan Payment}_i}{\text{Loan Amount}_i}$$

5. An expected return calculated as:

$$\begin{aligned}\text{Expected Return} &= (\text{Avg Return if No Default}) * (1 - PD) + (\text{Return if Default}) * PD \\ &= (\text{Avg Return if No Default}) * (1 - PD) + (1 - \text{Avg Loss Given Default}) * PD\end{aligned} \tag{1}$$

6. Return of portfolio of given grade computed as:

$$\text{Portfolio Return} = \frac{\sum_{i=1}^{n} \text{Total Payment}_i}{\sum_{i=1}^{n} \text{Total Loan Amount}_i} - 1 \tag{2}$$

Summarize the results in a table, using the format below, and *briefly* answer the following questions:

(a) Which grade offered the most favorable portfolio return overall, and which the least favorable? Now, let us focus on the probability of default and on the loss given default across grades. Which of the two quantities presents a higher variability as the loan grade changes? (Please show equations as well as numerical values in your response.)

(b) Note there is a difference between the portfolio return and the expected return for each grade.

[3]While most contexts in finance will focus on return *rates*, where the timing of payments matters, this sort of calculation also used in practice. See e.g. https://www.investopedia.com/terms/r/realization_multiple.asp

(i) Explain this difference mathematically. (Hint: if all loans had the same original loan balance/"loan_amnt", would there still be a difference?)

(ii) Check how returns on larger loans compare to returns of smaller loans for the A-grade and B-grade loans compared to D- and E- grade loans.

(c) Please explain the reason that the "Average Interest Rate" differs from the "Average Return if No Default." Compute what the ratio would be if there for each grade an individual loan in that grade with the rate equal to the average rate.

Grade	Interest Rate	Default Prevalence	Avg Return Default	Avg Return No Default	Expected Return	Portfolio Return
A						
B						
C						
D						
E						
F						
G						