

Bankrate Case Prompt

Bankrate was founded in 1976 as Bank Rate Monitor, a print publisher for the banking industry. Our experience has fueled our reputation as the premier financial authority. When you visit Bankrate.com, the reviews, guides, and educational content have been developed by leading personal finance experts. Bankrate's product comparison tools, calculators, and educational content help over 100 million consumers make smarter financial decisions each year. No matter where you are in your financial journey, Bankrate can help you reach your goals.

One area Bankrate specializes in is personal loans. Bankrate's goal is to help match the best lending partner for each individual customer. To do that, we collect information about the customer to understand who would be the best fit. We want to maximize the likelihood someone will be approved for the loan at the best "deal" possible.

In this case, let's imagine we have three lending partners, "A", "B", and "C". In this scenario, let's imagine that each lender is providing the same offer. However, for each approval, lender A provides us \$250, lender B pays us \$350, and lender C pays us \$150.

For 100K of our customers, we have collected a set of data in hopes to understand what variables are most important in determining approvability. We believe if we collect the right information, we will be able to show the appropriate lender to each customer to maximize approval rate and revenue. The team knows there are many possible data points that we could collect to help us better understand the likelihood of someone approving for the loan, but they started with this list for now.

This data has been recently collected from the website and the leadership team has asked you to help make sense of it. They want to know if they can increase revenue per application by matching certain groups of customers to specific lenders.

There are three broad categories that they would like you to investigate. These questions should be used as thought starters - please be creative with how you analyze and explore each question, leveraging relevant statistics and visualizations where applicable.

1. Explore the variables relationship with approvability:

a. *Possible things to consider:* Which variables are the most helpful in understanding if a customer is going to be approved or denied for a loan? Are there certain variables that are not useful to collect? Are there any feature modifications or transformations that would improve the predictive power of a variable?

2. Tell us about the lenders approval rates:

a. *Possible things to consider:* What is each Lender's average approval rate? Are there any clear differences between the three different lenders on what type of customers they approve? Are there variables that reliably predict a user's approval likelihood for a particular lender?

3. Evaluate which customers we should match to each lender to maximize Revenue Per Application:

a. *Possible things to consider:* Are there groups of customers that would be a better fit for a different lender? How much incremental revenue could we make if we matched lenders to certain groups of customers more appropriately? What considerations should we have in mind if we planned to match customers with lenders in real time?