

Task

Given a client's financial information, Quicken Loans offers that client a particular loan *strategy*. The client can then decide whether or not to take out the loan. More specifically, there are many different *results* that can occur if a client decides to take out a loan. Each result is considered a positive outcome, and no result is considered a negative outcome. Your goal is to predict whether or not a particular client and strategy combination will result in a positive outcome.

Since the loan strategies and outcomes are IP of Quicken Loans, they have been anonymized in this dataset. Strategies are named after Michigan Football players, and results are given numerical codes. Note that all numerical codes, including "000", are considered positive results. Each strategy has several associated results that QL bankers are expected to achieve, but since each client is unique, it does not always work out this way.

In addition to knowing which strategies were offered to each particular client (loan_table_train and loan_table_test), we also provide you with a list of strategies that each client is eligible for (play_table_train and play_table_test). You are also know which results are expected to come from each strategy (strategies_and_results_anonymized), but note that sometimes the results do not correspond to expectation.

For a concrete example of how to set up your prediction task, I recommend you take a look at our [starter code on github](#).

File descriptions

- loan_table_train.csv - list of clients, strategy offered, and outcome
- loan_table_test.csv - list of clients, strategy offered, no outcome
- play_table_train.csv - list of strategies that each client from loan_table_train is eligible for
- play_table_test.csv - list of strategies that each client from loan_table_test is eligible for
- strategies_and_results_anonymized.csv - list of results that correspond with each strategy
- schema.docx - description of data fields
- sample_submission.csv - a sample submission file in the correct format

Data fields - loan_table_train and loan_table_test

- id - anonymized unique loan identifier
- FICO - Credit score
- largest_open_mortgage_UPB - Remaining balance on the largest open mortgage
- largest_open_mortgage_open_date - Date on which the largest open mortgage was opened
- largest_open_mortgage_payment - Monthly payment of the largest open mortgage
- largest_open_mortgage_type - Mortgage type of the largest open mortgage

- largest_open_mortgage_original_term - Full term of the largest open mortgage
- total_revolving_tradeline_debt - Total amount of revolving (e.g. credit card) debt
- total_revolving_tradeline_payment - Payment made last month on revolving accounts
- total_student_loan_tradeline_debt - Total amount of student debt
- total_student_loan_tradeline_payment - Monthly payment towards student debts
- total_mortgage_tradeline_debt - Combined debt of mortgages
- total_mortgage_tradeline_payment - Combined mortgage monthly payments
- nonmortgage_past_due_amount - Late payment amount on non-mortgage related debt
- past_pulls_14days - Number of credit pulls in the last 14 days
- public_record_count - Number of public records (bankruptcies, tax liens, etc.)
- loan_purpose - Purchase or refinance
- Income - Monthly income
- state - State of property related to mortgage
- zip - Zip code of property related to mortgage
- serviced - Whether or not the client already has a mortgage with us
- DTI - Debt to Income ratio. Ratio of monthly payments to monthly income
- strategy - Strategy presented
- result - What result client had (NULL if they didn't move forward)

Data fields - play_table_train and play_table_test

- id - anonymized unique loan identifier
- strategy - strategy applicable to that client

Data fields - strategies_and_results_anonymized.csv

- strategy - strategies
- result - results tied to strategies