

Loan Data from Prosper

By Mina Nabil

This data set contains 113,937 loans with 81 variables on each loan, including loan amount, borrower rate (or interest rate), current loan status, borrower income, and many others.

Main findings from the exploratory data analysis:

- The number of current and completed loans are the higher count numbers in the data set, which reflects that most of the clients that requested a loan are able to complete it.
- People with Income Range from \$25,000 to 74,999 are the ones who are taking loans most.
- Interest Rate is normally distributed with the peak between 14 and 18 percent in addition we have some increase from 30 to 35 percent.
- As we can see from the histogram that the Interest Rate is normally distributed with the peak between 15 and 20 percent, as we can see peak of APR higher than Interest rate peak that is because the APR includes not only the interest expense on the loan but also all fees and other costs involved in procuring the loan.
- Most borrowers get Prosper Rating of 'C'.
- Charged off and defaulted status has a lower credit score.
- The clients with past due greater than 120 days have the higher interest rate.
- The higher the Income Range the lower the Annual Interest rate.
- The higher the loan original amount the lower the interest rate.
- High risk and E prosper rating clients have the highest annual interest rate, and also noticed that the clients with A and AA rating have the lowest annual interest rate.
- the highest loan original amount with highest interest annual rate existed in California state.
- Borrower APR and Lender yield are directly positively correlated as when the interest borrowers pay more, the lender yield will increase. Higher the prosper score lower will lower borrower APR and then the lender yield will also be lower.
- Most of the stats has Loan original amount between 10000 and 20000\$

how you chose the results to put in your explanatory analysis?

Based on the three main questions that I choose:

- What factors affect a loan's outcome status?
- What affects the borrower's APR or interest rate?
- Are there differences between loans depending on how large the original loan amount was?

List of Resources:

Websites:

- <https://classroom.udacity.com>
- <https://seaborn.pydata.org/examples/index.html>
- https://matplotlib.org/3.3.1/tutorials/introductory/sample_plots.html#sphx-glr-tutorials-introductory-sample-plots-py
- [Charged Off Debt Must Still Be Repaid - Experian](#)
- [Average Personal Loan Interest Rates | Bankrate](#)
- [Personal Loan Interest Rates - Latest Interest Rate in 2021 \(bankbazaar.com\)](#)
- [Indexing and selecting data — pandas 1.3.4 documentation \(pydata.org\)](#)
- [Yield vs Interest Rate: What's the Difference? \(investopedia.com\)](#)
- [python - How to divide pandas column by list of integers based on row value - Stack Overflow](#)
- https://video.udacity-data.com/topher/2019/April/5ca78b26_dataset-project-communicate-data-findings/dataset-project-communicate-data-findings.pdf
- [KeyError: 'Passing list-likes to .loc or \[\] with any missing labels is no longer supported' · Issue #147 · hyperopt/hyperopt-sklearn \(github.com\)](#)

Books:

- Python for data analysis
- Naked statistics.
- Thinking with data