

Lending Club Case Study

Prepared By – Shashi Bhushan Singh

Topics Being Covered

- Project Name
- General Information
- Conclusions
- Technologies Used
- Acknowledgements

Project Name – Lending Club Case Study

- General Information –

Lending Club Company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers

who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

Conclusions

- We performed following steps –
 - a. Data Cleaning
 - b. Data Analysis
 - c. Observations
 - d. Recommendations

Data Cleaning

- There are some columns with all null values, so we won't need those columns in analysis.
- Some columns has unique value that was of not much use in analysis, so we removed that as well.
- Some columns had descriptive values that we didn't utilise in our analysis as we were not analysing customer behaviour.
- Some of the columns were mostly for the post loan approval, so we dropped those columns as well
- Dropped the rows which had loan status as 'Current'

Data Analysis

We performed following types of Analysis-

1. Univariate Analysis
2. Bivariate Analysis
3. Multivariate Analysis

Observations

Following are the observations are of high level -

- People having Rented Home had higher default rate
- Customer with Consolidated debt had higher chances of defaulting
- Higher number of loan defaults are for the income range 0K to 5K
- We can see that higher defaults are happening When Loan term is shorter in this case we can see that more defaulters are for 36 months loan term.
- We can see that higher defaults are happening in the case when verification status is not verified or in other words we can say when applicant status is not verified.

- Higher defaults are for the month of December when the loan was approved/granted in the month of December.
- Also we can see one observation in second graph that year 2011 is showing highest number of defaults.
- Highest chances of defaulting is when loan amount is between 5K to 10K.
- Highest chances of defaulting is when loan instalments are between 145K to 274K.
- When Debt to income ratio is between 12 to 18; chances of defaulting is higher.

Recommendation

- I. - Customer with lower annual income should be approved lower amount of loan with longer duration.
- II. - Applicant should always be validated

- III. - There were higher default rates for month of December that indicates Validation is not thoroughly performed, so more scrutiny should kept in place while approving loans during year end.
- IV. - Applicant with past history of defaults should be given loan on higher interest rate.

- **Technologies Used**

- Python Programming Language

- Python Libraries

- ✚ - Pandas

- ✚ - numpy

- ✚ - matplotlib ✚ - seaborn

- **Acknowledgements**

This project has been created as part of UPGRAD Lending Club Case Study module.

- **Contact**

Created by – Shashi Bhushan Singh