**Lending Club - Insightful Loan Default EDA**

Lending Club is a peer to peer lending company based in the United Statesis 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.

Investors provide funds for potential borrowers and investors earn a profit depending on the risk they take (the borrowers credit score). Lending Club provides the "bridge" between investors and borrowers

It services people that need personal loans between $1,000 and $40,000. Borrowers receive the full amount of the issued loan minus the origination fee, which is paid to the company. Investors purchase notes backed by the personal loans and pay Lending Club a service fee. The company shares data about all loans issued through its platform during certain time periods.

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.

Therefore, we will be working on exploring and crunching out the driving factors behind the loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

**About The Dataset**  
These files contain complete loan data for all loans issued through the 2007-2017, including the current loan status (Current, Charged-off, Fully Paid, etc.) and latest payment information. The file containing loan data through the "present" contains complete loan data for all loans issued through the previous completed calendar quarter. Additional features include credit scores, number of finance inquiries, and collections among others. The file is a matrix of about 39 thousand observations and 111 variables. A data dictionary is provided in a separate file.

**Questions**

* What type of data are we working with?
* What types of features do we have?
* Do we need to treat missing values?
* What is the distribution of Loan Status?
* What is the distribution of Loan Default with other features?
* What all plots we can draw for inferring the relation with Loan Default?
* Majorly, what all features best describes the Loan Default?

**Feature Information**

* Loan Variables such as loan amount, term, interest rate, etc which shows the information about the loan that will help us in finding loan default.
* Profile Variables such as employment status, relationship status, etc. which shows the information about the borrower profile which is not useful for us.
* Miscellaneous Variables such as geographic, EMI, delinquency, etc. which shows the information which is updated after providing the loan which in our case is not useful as we need to decide whether to provide loan or not by default analysis.

|  |  |
| --- | --- |
| acc\_now\_delinq | The number of accounts on which the borrower is now delinquent. |
| acc\_open\_past\_24mths | Number of trades opened in past 24 months. |
| addr\_state | The state provided by the borrower in the loan application |
| all\_util | Balance to credit limit on all trades |
| annual\_inc | The self-reported annual income provided by the borrower during registration. |
| annual\_inc\_joint | The combined self-reported annual income provided by the co-borrowers during registration |
| application\_type | Indicates whether the loan is an individual application or a joint application with two co-borrowers |
| avg\_cur\_bal | Average current balance of all accounts |
| bc\_open\_to\_buy | Total open to buy on revolving bankcards. |
| bc\_util | Ratio of total current balance to high credit/credit limit for all bankcard accounts. |
| chargeoff\_within\_12\_mths | Number of charge-offs within 12 months |
| collection\_recovery\_fee | post charge off collection fee |
| collections\_12\_mths\_ex\_med | Number of collections in 12 months excluding medical collections |
| delinq\_2yrs | The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years |
| delinq\_amnt | The past-due amount owed for the accounts on which the borrower is now delinquent. |
| dti | A ratio calculated using the borrower’s total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower’s self-reported monthly income. |
| dti\_joint | A ratio calculated using the co-borrowers' total monthly payments on the total debt obligations, excluding mortgages and the requested LC loan, divided by the co-borrowers' combined self-reported monthly income |
| earliest\_cr\_line | The month the borrower's earliest reported credit line was opened |
| emp\_length | Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years. |
| emp\_title | The job title supplied by the Borrower when applying for the loan.\* |
| funded\_amnt | The total amount committed to that loan at that point in time. |
| funded\_amnt\_inv | The total amount committed by investors for that loan at that point in time. |
| grade | LC assigned loan grade |
| home\_ownership | The home ownership status provided by the borrower during registration or obtained from the credit report. Our values are: RENT, OWN, MORTGAGE, OTHER |
| il\_util | Ratio of total current balance to high credit/credit limit on all install acct |
| initial\_list\_status | The initial listing status of the loan. Possible values are – W, F |
| inq\_fi | Number of personal finance inquiries |
| inq\_last\_12m | Number of credit inquiries in past 12 months |
| inq\_last\_6mths | The number of inquiries in past 6 months (excluding auto and mortgage inquiries) |
| installment | The monthly payment owed by the borrower if the loan originates. |
| int\_rate | Interest Rate on the loan |
| issue\_d | The month which the loan was funded |
| last\_credit\_pull\_d | The most recent month LC pulled credit for this loan |
| last\_pymnt\_amnt | Last total payment amount received |
| last\_pymnt\_d | Last month payment was received |
| loan\_amnt | The listed amount of the loan applied for by the borrower. If at some point in time, the credit department reduces the loan amount, then it will be reflected in this value. |
| loan\_status | Current status of the loan |
| max\_bal\_bc | Maximum current balance owed on all revolving accounts |
| mo\_sin\_old\_il\_acct | Months since oldest bank installment account opened |
| mo\_sin\_old\_rev\_tl\_op | Months since oldest revolving account opened |
| mo\_sin\_rcnt\_rev\_tl\_op | Months since most recent revolving account opened |
| mo\_sin\_rcnt\_tl | Months since most recent account opened |
| mort\_acc | Number of mortgage accounts. |
| mths\_since\_last\_delinq | The number of months since the borrower's last delinquency. |
| mths\_since\_last\_major\_derog | Months since most recent 90-day or worse rating |
| mths\_since\_last\_record | The number of months since the last public record. |
| mths\_since\_rcnt\_il | Months since most recent installment accounts opened |
| mths\_since\_recent\_bc | Months since most recent bankcard account opened. |
| mths\_since\_recent\_bc\_dlq | Months since most recent bankcard delinquency |
| mths\_since\_recent\_inq | Months since most recent inquiry. |
| mths\_since\_recent\_revol\_delinq | Months since most recent revolving delinquency. |
| next\_pymnt\_d | Next scheduled payment date |
| num\_accts\_ever\_120\_pd | Number of accounts ever 120 or more days past due |
| num\_actv\_bc\_tl | Number of currently active bankcard accounts |
| num\_actv\_rev\_tl | Number of currently active revolving trades |
| num\_bc\_sats | Number of satisfactory bankcard accounts |
| num\_bc\_tl | Number of bankcard accounts |
| num\_il\_tl | Number of installment accounts |
| num\_op\_rev\_tl | Number of open revolving accounts |
| num\_rev\_accts | Number of revolving accounts |
| num\_rev\_tl\_bal\_gt\_0 | Number of revolving trades with balance >0 |
| num\_sats | Number of satisfactory accounts |
| num\_tl\_120dpd\_2m | Number of accounts currently 120 days past due (updated in past 2 months) |
| num\_tl\_30dpd | Number of accounts currently 30 days past due (updated in past 2 months) |
| num\_tl\_90g\_dpd\_24m | Number of accounts 90 or more days past due in last 24 months |
| num\_tl\_op\_past\_12m | Number of accounts opened in past 12 months |
| open\_acc | The number of open credit lines in the borrower's credit file. |
| open\_acc\_6m | Number of open trades in last 6 months |
| open\_il\_12m | Number of installment accounts opened in past 12 months |
| open\_il\_24m | Number of installment accounts opened in past 24 months |
| open\_il\_6m | Number of currently active installment trades |
| open\_rv\_12m | Number of revolving trades opened in past 12 months |
| open\_rv\_24m | Number of revolving trades opened in past 24 months |
| out\_prncp | Remaining outstanding principal for total amount funded |
| out\_prncp\_inv | Remaining outstanding principal for portion of total amount funded by investors |
| pct\_tl\_nvr\_dlq | Percent of trades never delinquent |
| percent\_bc\_gt\_75 | Percentage of all bankcard accounts > 75% of limit. |
| pub\_rec | Number of derogatory public records |
| pub\_rec\_bankruptcies | Number of public record bankruptcies |
| purpose | A category provided by the borrower for the loan request. |
| pymnt\_plan | Indicates if a payment plan has been put in place for the loan |
| recoveries | post charge off gross recovery |
| revol\_bal | Total credit revolving balance |
| revol\_bal\_joint | Sum of revolving credit balance of the co-borrowers, net of duplicate balances |
| revol\_util | Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit. |
| sec\_app\_chargeoff\_within\_12\_mths | Number of charge-offs within last 12 months at time of application for the secondary applicant |
| sec\_app\_collections\_12\_mths\_ex\_med | Number of collections within last 12 months excluding medical collections at time of application for the secondary applicant |
| sec\_app\_earliest\_cr\_line | Earliest credit line at time of application for the secondary applicant |
| sec\_app\_inq\_last\_6mths | Credit inquiries in the last 6 months at time of application for the secondary applicant |
| sec\_app\_mort\_acc | Number of mortgage accounts at time of application for the secondary applicant |
| sec\_app\_mths\_since\_last\_major\_derog | Months since most recent 90-day or worse rating at time of application for the secondary applicant |
| sec\_app\_num\_rev\_accts | Number of revolving accounts at time of application for the secondary applicant |
| sec\_app\_open\_acc | Number of open trades at time of application for the secondary applicant |
| sec\_app\_open\_il\_6m | Number of currently active installment trades at time of application for the secondary applicant |
| sec\_app\_revol\_util | Ratio of total current balance to high credit/credit limit for all revolving accounts |
| sub\_grade | LC assigned loan subgrade |
| tax\_liens | Number of tax liens |
| term | The number of payments on the loan. Values are in months and can be either 36 or 60. |
| title | The loan title provided by the borrower |
| tot\_coll\_amt | Total collection amounts ever owed |

|  |
| --- |
|  |
| tot\_cur\_bal | Total current balance of all accounts |
| tot\_hi\_cred\_lim | Total high credit/credit limit |
| total\_acc | The total number of credit lines currently in the borrower's credit file |
| total\_bal\_ex\_mort | Total credit balance excluding mortgage |
| total\_bal\_il | Total current balance of all installment accounts |
| total\_bc\_limit | Total bankcard high credit/credit limit |
| total\_cu\_tl | Number of finance trades |
| total\_il\_high\_credit\_limit | Total installment high credit/credit limit |
| total\_pymnt | Payments received to date for total amount funded |
| total\_pymnt\_inv | Payments received to date for portion of total amount funded by investors |
| total\_rec\_int | Interest received to date |
| total\_rec\_late\_fee | Late fees received to date |
| total\_rec\_prncp | Principal received to date |
| total\_rev\_hi\_lim |  |
| verification\_status | Indicates if income was verified by LC, not verified, or if the income source was verified |
| verification\_status\_joint |  |