



LANDING CLUB CASE STUDY

LOAN DEFAULT PREDICTION

BY : VIKRANT DABAS

INTRODUCTION

We are a financial loan marketplace provider. A crucial business decision is whether to approve a loan application, balancing the risk of financial loss (if the borrower defaults) against the potential loss of business (if a creditworthy applicant is denied). Analysis of our historical loan data reveals that significant loans result in charge-offs, highlighting the significant financial impact of loan defaults. In this presentation, we will leverage Exploratory Data Analysis (EDA) to uncover patterns and insights within consumer and loan attributes that influence the likelihood of default. By understanding these factors, we can develop data-driven strategies to mitigate risk, optimize loan approval processes, and ensure the long-term success of our lending business.

DATA EXPLORATION

The historical data of the customer load has the following important fields are:

Fields	Description
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
chargeoff_within_12_mths	Number of charge-offs within 12 months
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.
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
installment	The monthly payment owed by the borrower if the loan originates.
int_rate	Interest Rate on the loan
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

DATA EXPLORATION

The historical data of the customer load has the following important fields are:

Fields	Description
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
chargeoff_within_12_mths	Number of charge-offs within 12 months
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.
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
installment	The monthly payment owed by the borrower if the loan originates.
int_rate	Interest Rate on the loan
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

DATA CLEAN UP

1. Deleted the columns with 'NA' only value
2. Rounded off the funded_amt_inv
3. Removed months suffix from the term column.
4. Dropped the rows with blank data.

but past loan applicants and whether they defaulted or not. The columns `loan_status` appears to be the target variable. We will start with univariate analysis.

PLOT

