



Lending Club Case Study

ABHIROOP BHATTACHARYA

ANINDITA DEB

Introduction

Lending Club connects borrowers with investors. Borrowers can get loans for a variety of purposes whereas investors can earn interest on their loans by investing in Lending Club notes.

Problem Statement

Lending club wants to enhance it's ability to predict which borrowers are likely to default on their loans. This will help Lending Club to make better lending decisions and to protect its investors.

Data

Includes information about borrowers

EG: employment history, annual income, address.

It also includes information about loans

EG: amount, term, interest rate, installment.

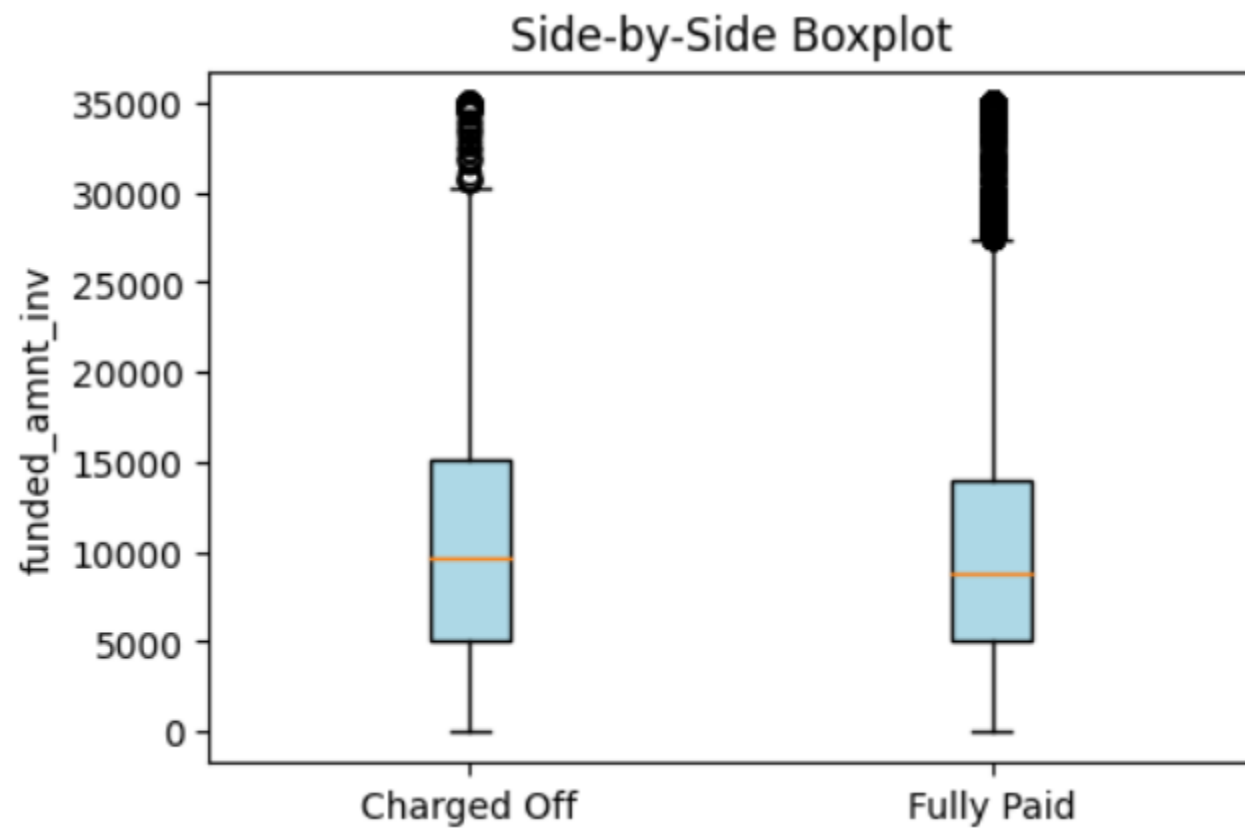
Methods

The following steps will be taken for this part of the analysis

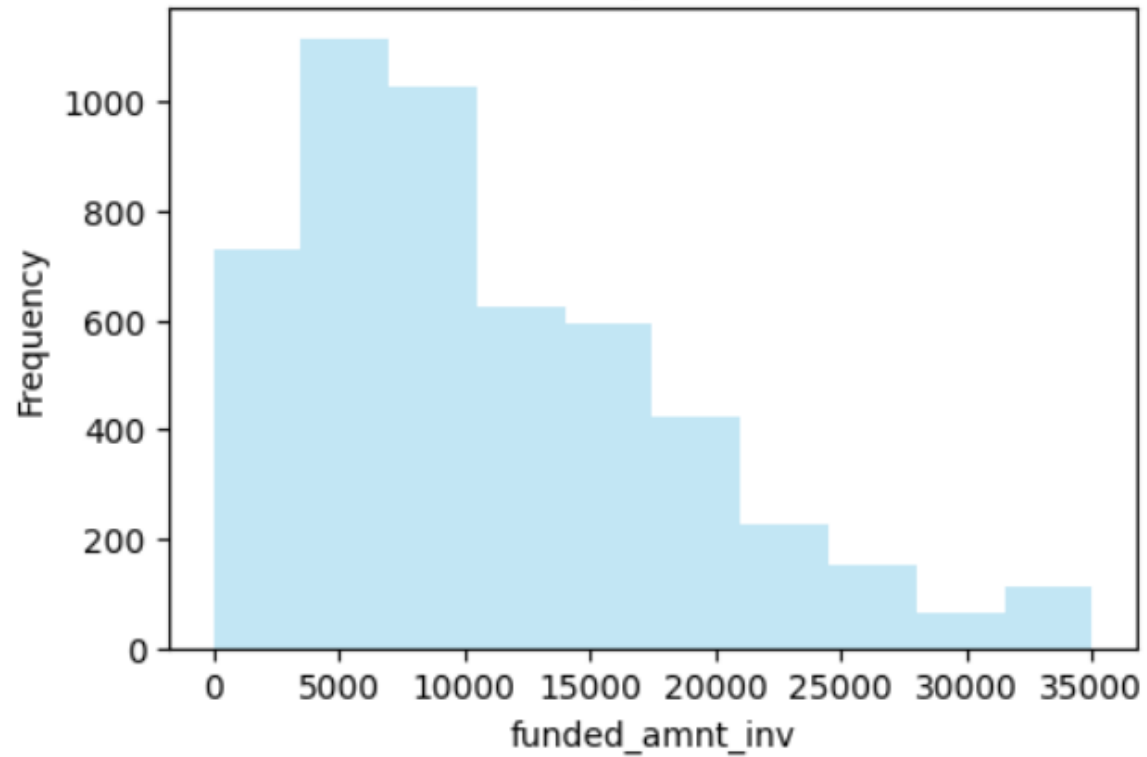
- Converting date columns to Pandas DateTime type and percentage columns to numerical type after removing the % symbol.
- Removing columns which contain more than 10% of missing values or only a single unique value since they will not be of use in analysis.
- Dropping missing values where imputing them with mean/median/mode will heavily skew the data.
- Imputing missing values where feasible.
- Converting numeric columns to categorical where feasible.
- Removing outliers from all the numerical columns.

Results

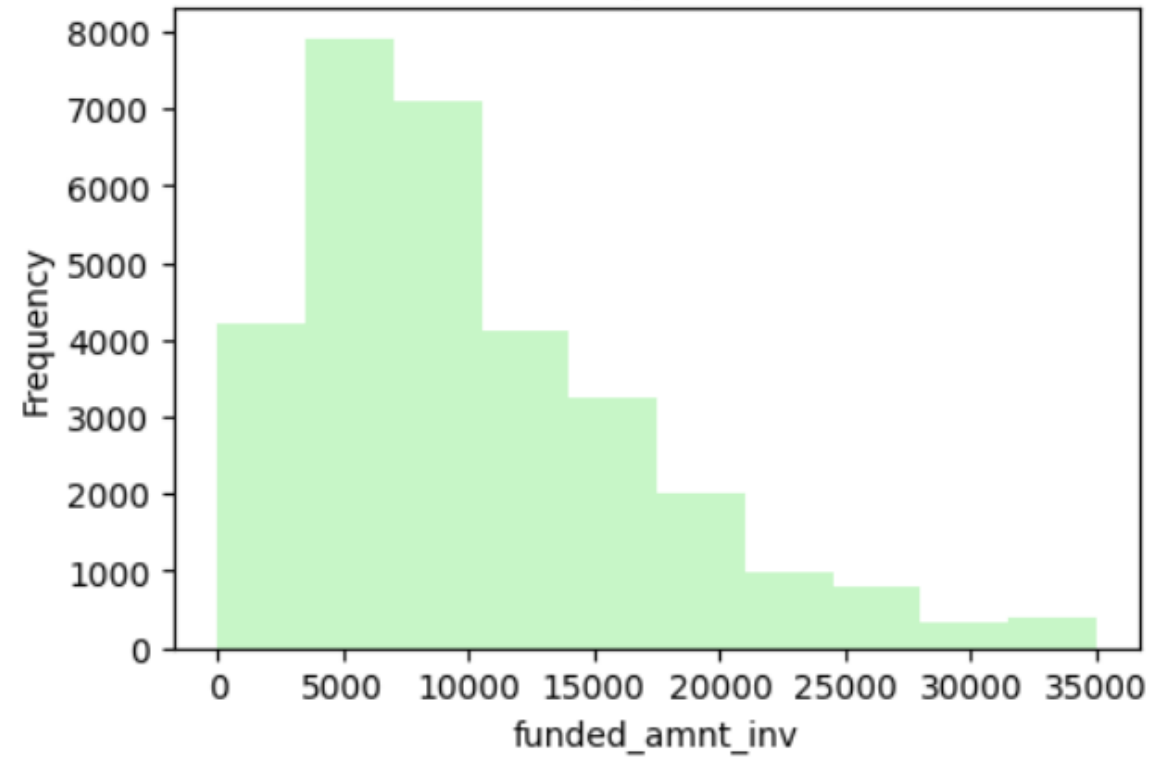
Need to write



Charged Off



Fully Paid



Conclusion

We were able to improve Lending Club's ability to predict which borrowers are likely to default on their loans to help Lending Club make better lending decisions and to protect its investors.

Thank You
