LENDING CLUB CASE STUDY

SreeLakshmiKanthi   & Nivedita

INTRODUCTION

In this case study, apart from applying the techniques that we have learnt in EDA, we will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers.

PROBLEM STATEMENT

You work for a **consumer finance company** which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company chose to decide for loan approval based on the applicant's profile.
Two **types of risks** are associated with the bank's decision:

• If the applicant is **likely to repay the loan**, then not approving the loan results in a **loss of business** to the company

• If the applicant is **not likely to repay the loan,** i.e., he/she is likely to default, then approving the loan may lead to a **financial loss** for the company

The data given contains information about past loan applicants and whether they 'defaulted' or not. The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

BUSINESS OBJECTIVE

EDA to understand how consumer attributes and loan attributes influence the tendency of default.

Lending loans to 'risky' applicants is the largest source of financial loss (called credit loss).

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 this case, the customers labelled as 'charged-off' are the 'defaulters'.

If one can identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants' using EDA is the aim of this case study.

DATA UNDERSTANDING

Loan data set has the complete loan data for all loans issued through the time 2007 to 2011. Our EDA is based on following attributes.

Consumer attributes

Employment length

Employee title

Annual income

Home ownership

Purpose

Loan attributes

Loan amount

Interest rate

Loan status

Loan term

ANALYSIS APPROACH

Import data → Data cleaning → Univariate analysis with metrics → Bivariate analysis with derived metrics

DATA CLEANING

- ✓ Check for unnecessary header and footer rows
- ✓ Remove large null or missing value columns
- ✓ Impute missing values or remove rows with missing values
- ✓ Drop columns which are not useful for analysis
- ✓ Remove duplicate data
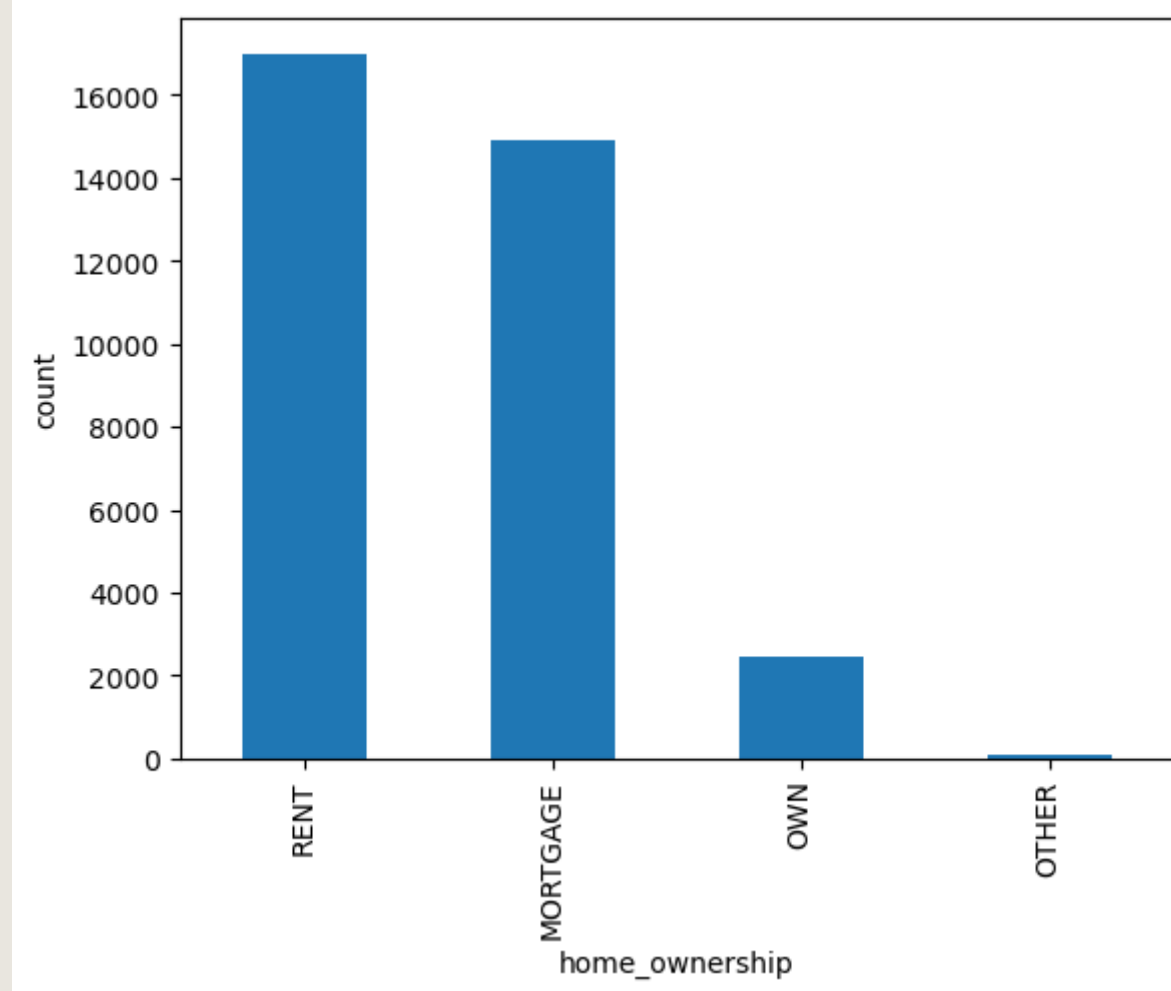- ✓ Standardize data
- ✓ Identify outliers

UNIVARIATE ANALYSIS
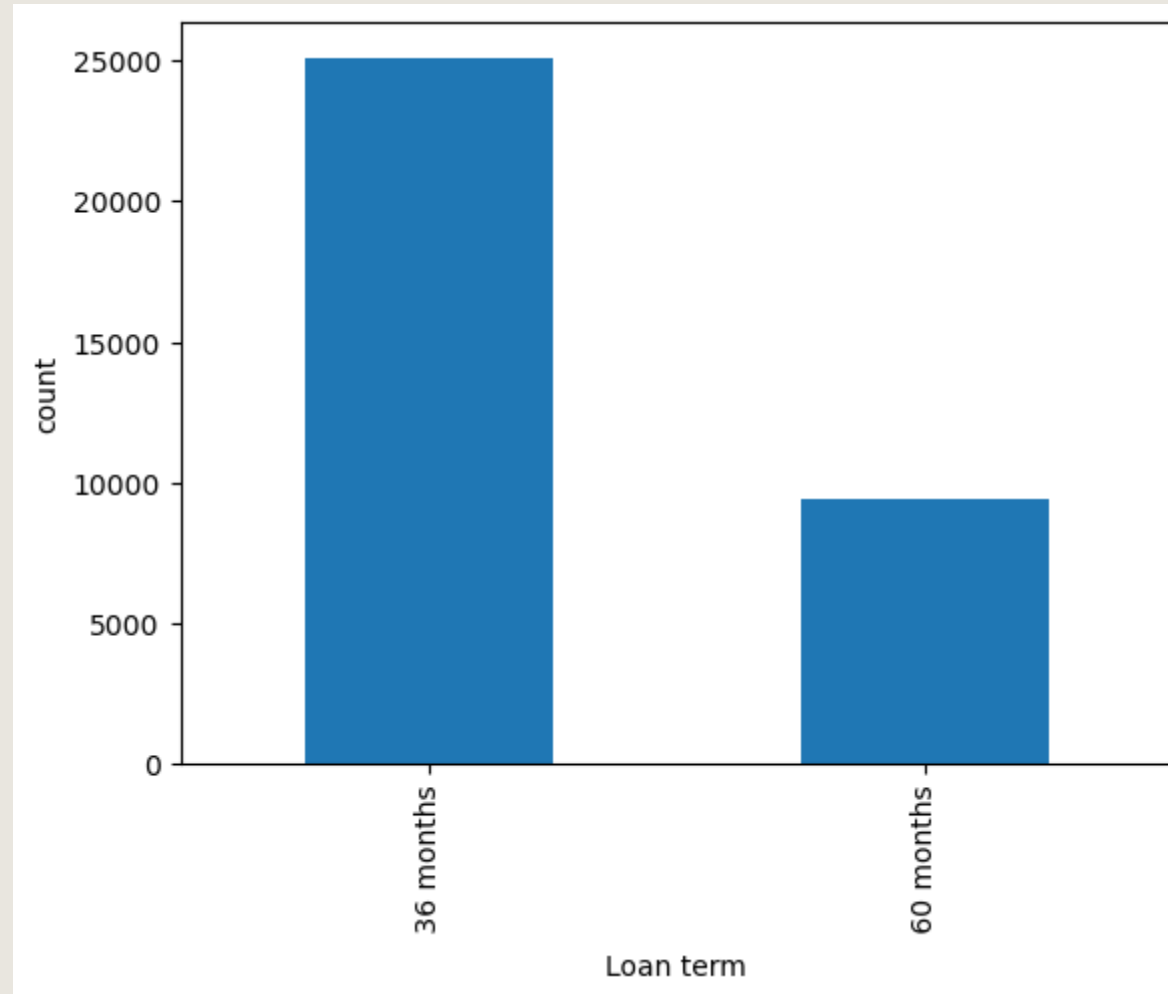
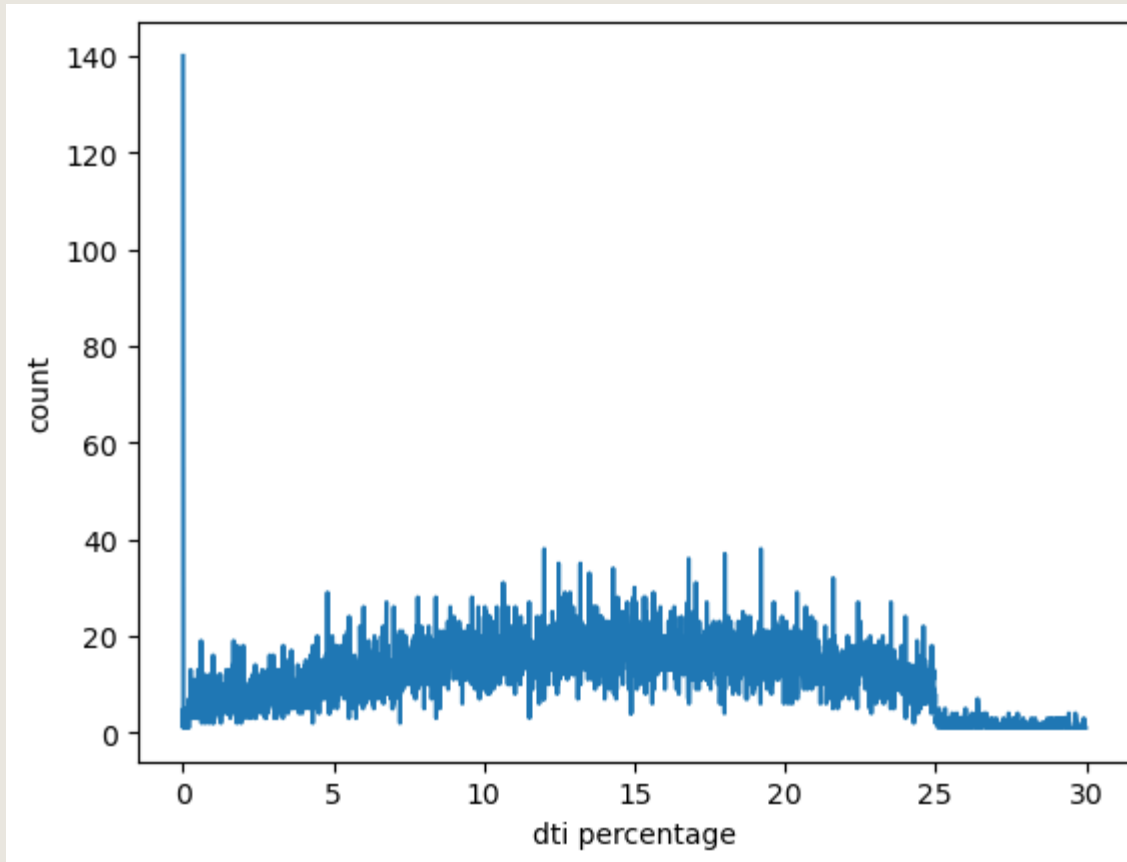Almost 75% applicants chose loan term of 36months

UNIVARIATE ANALYSIS
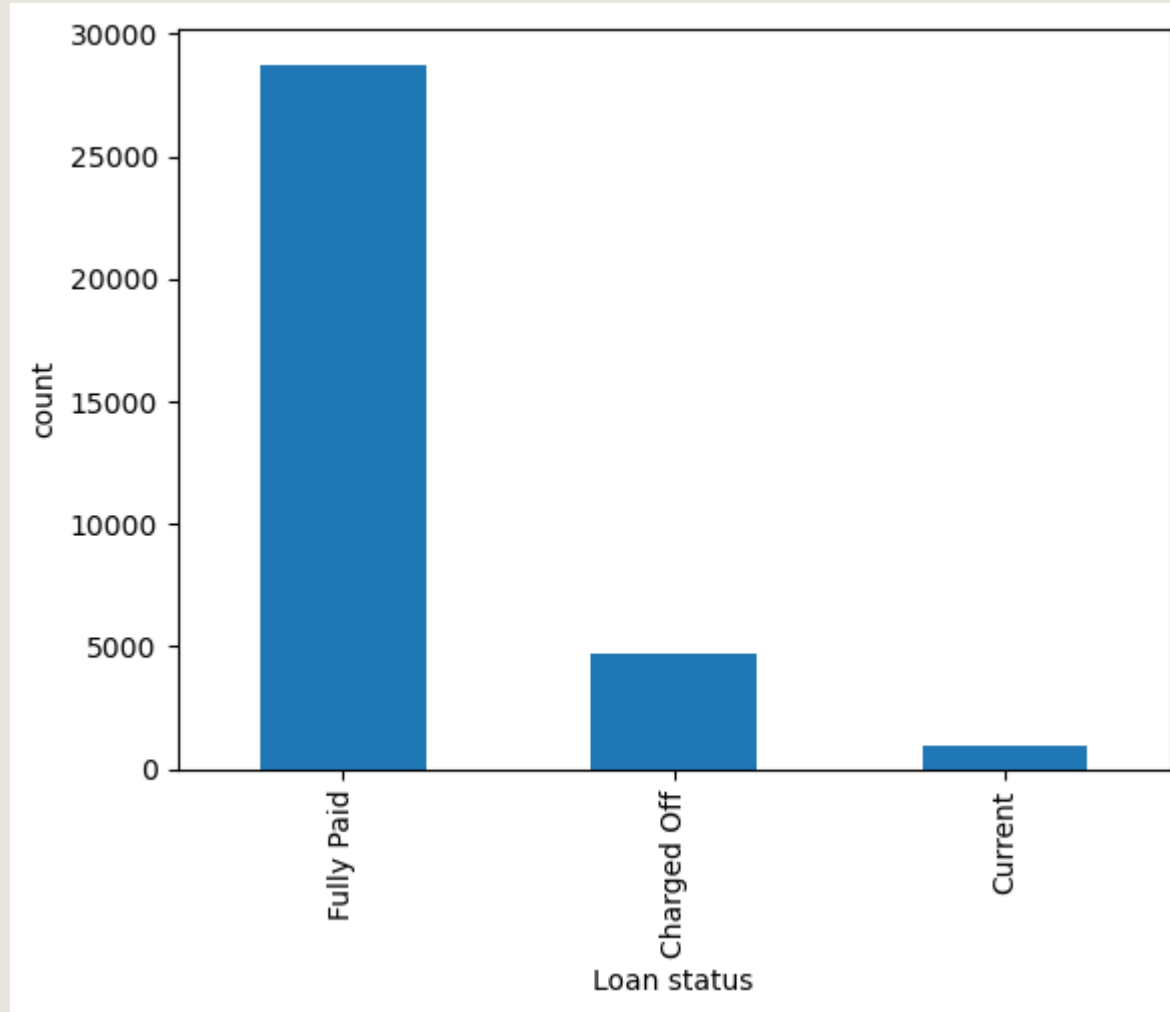
Around 14% of loan applicants are charged off

UNIVARIATE ANALYSIS

Around 49% of loan applicants are having rental accommodation followed by 43% are mortgage
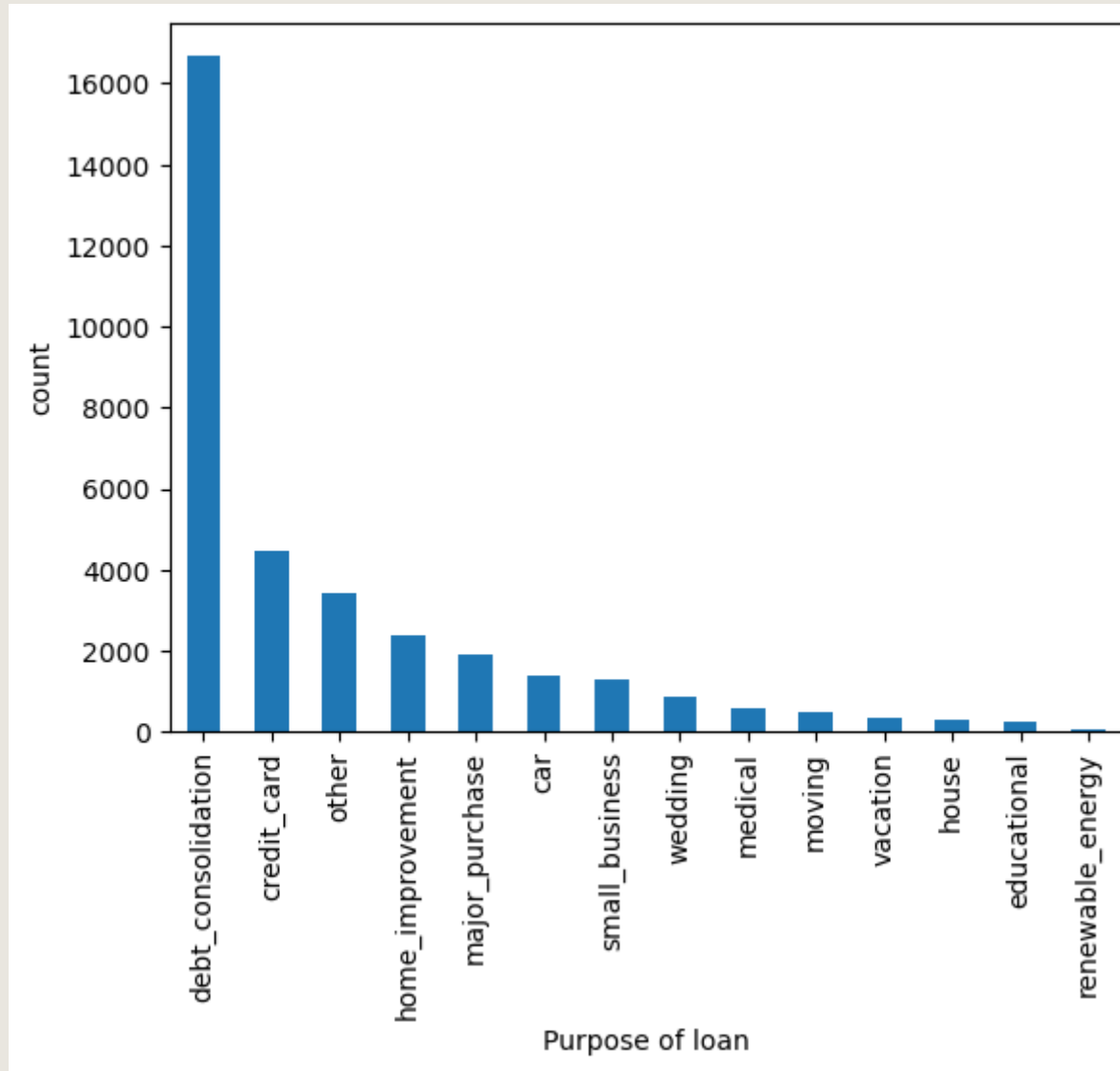
UNIVARIATE ANALYSIS

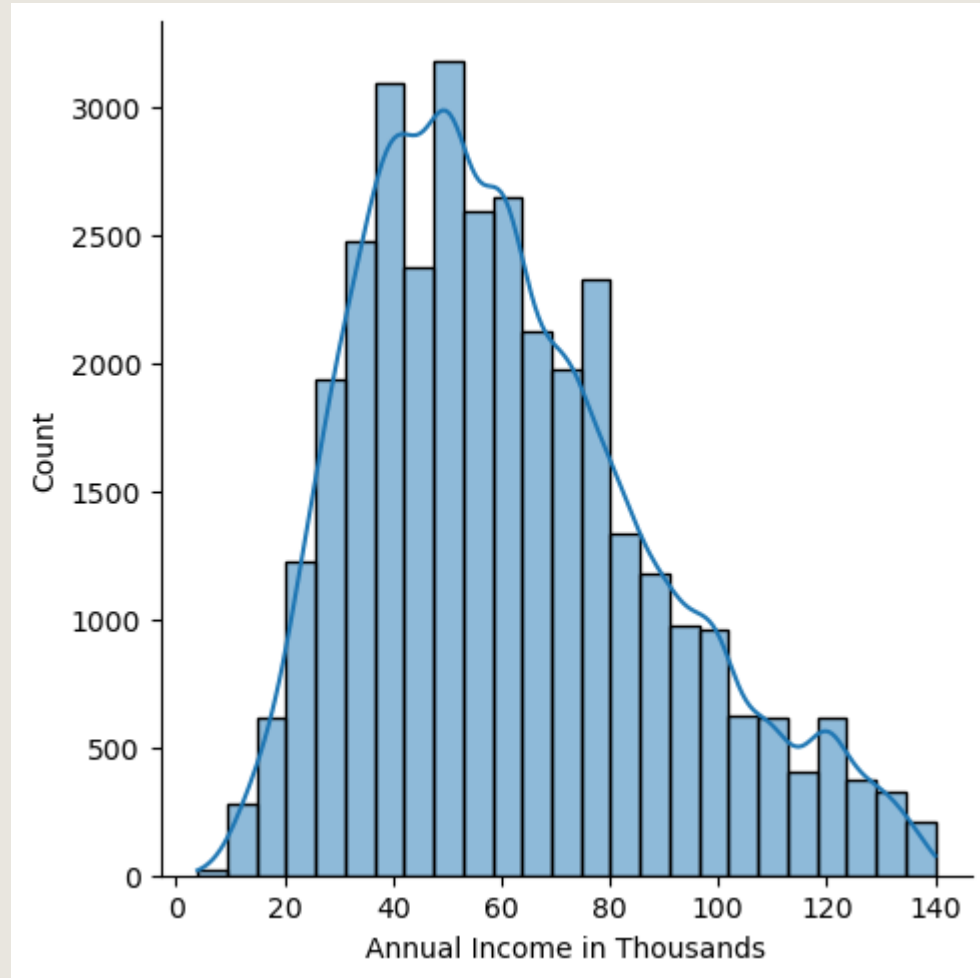Lender can allow max dti up to 30%

A high DTI means that more of applicant's money already goes towards debt repayment.
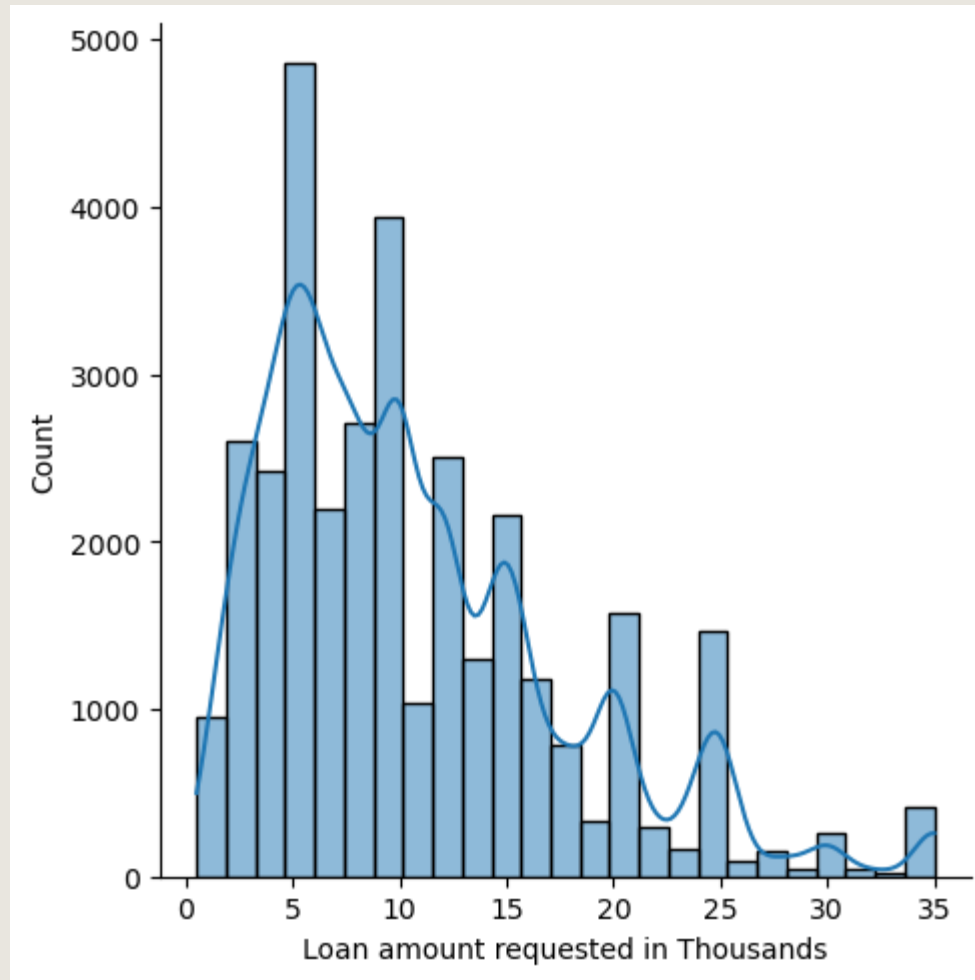
UNIVARIATE ANALYSIS

Around 48% of loan applicants have mentioned purpose as debt consolidation followed by 13% credit card
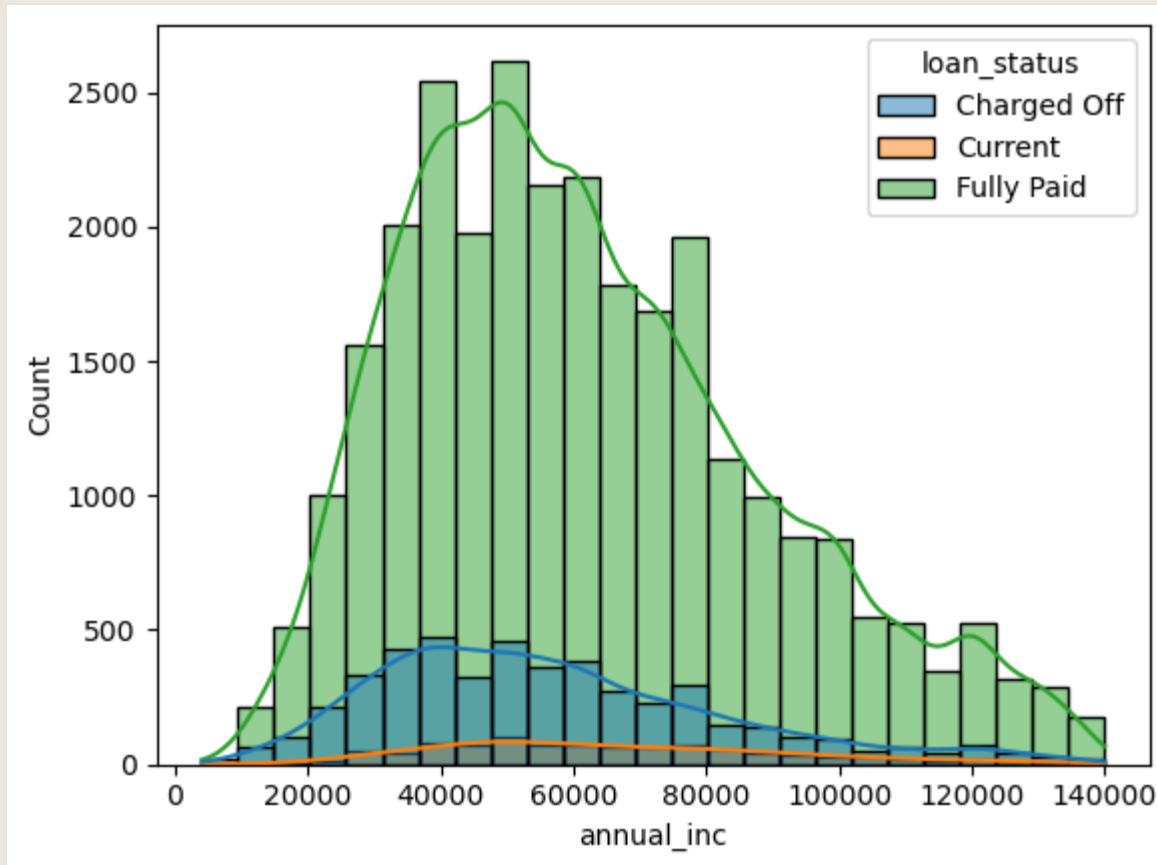
UNIVARIATE ANALYSIS

Majority of loan applicants have annual income between 40K to 80K
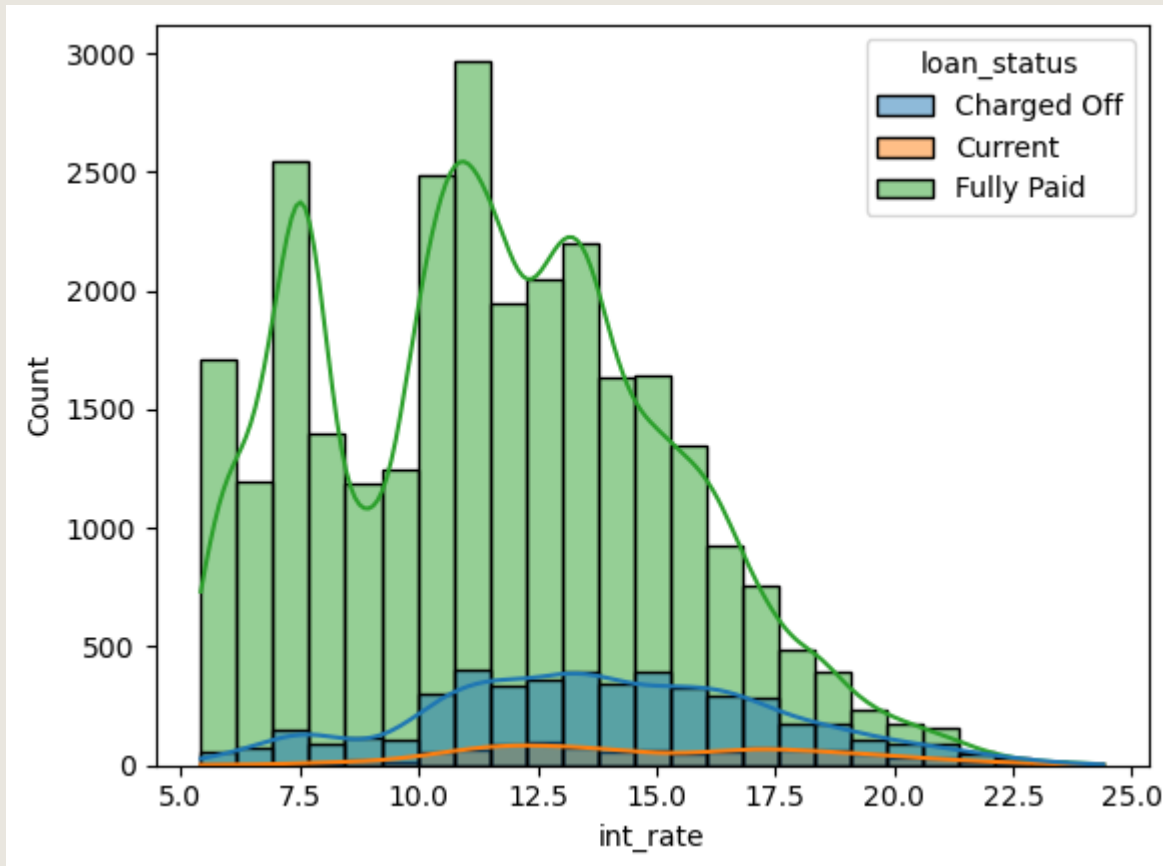
UNIVARIATE ANALYSIS

Min of 500 and Max up to 35K loan amount is being requested. Central tendency of loan amount request is 5200 to 15K

BIVARIATE ANALYSIS
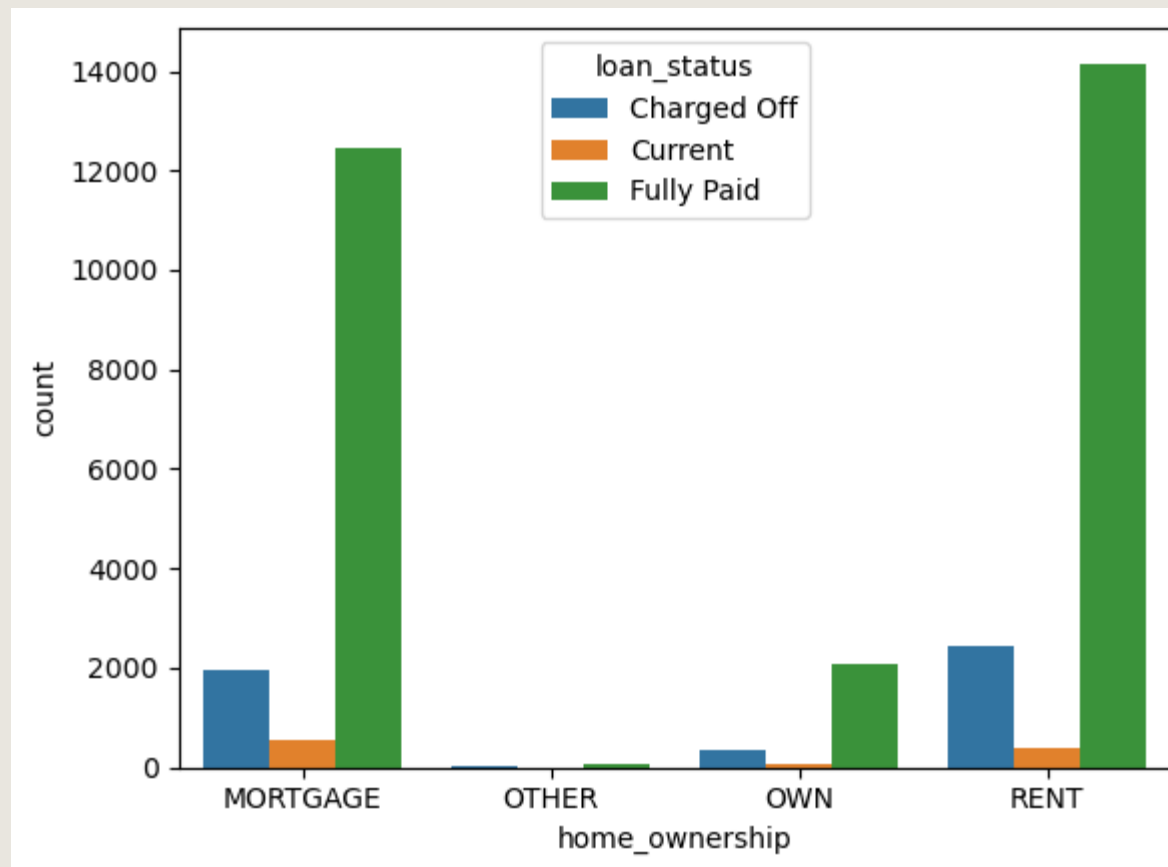

Loan applicants with annual income 20k to 80k have more defaults and annual income greater than 120k charged off count is coming down significantly
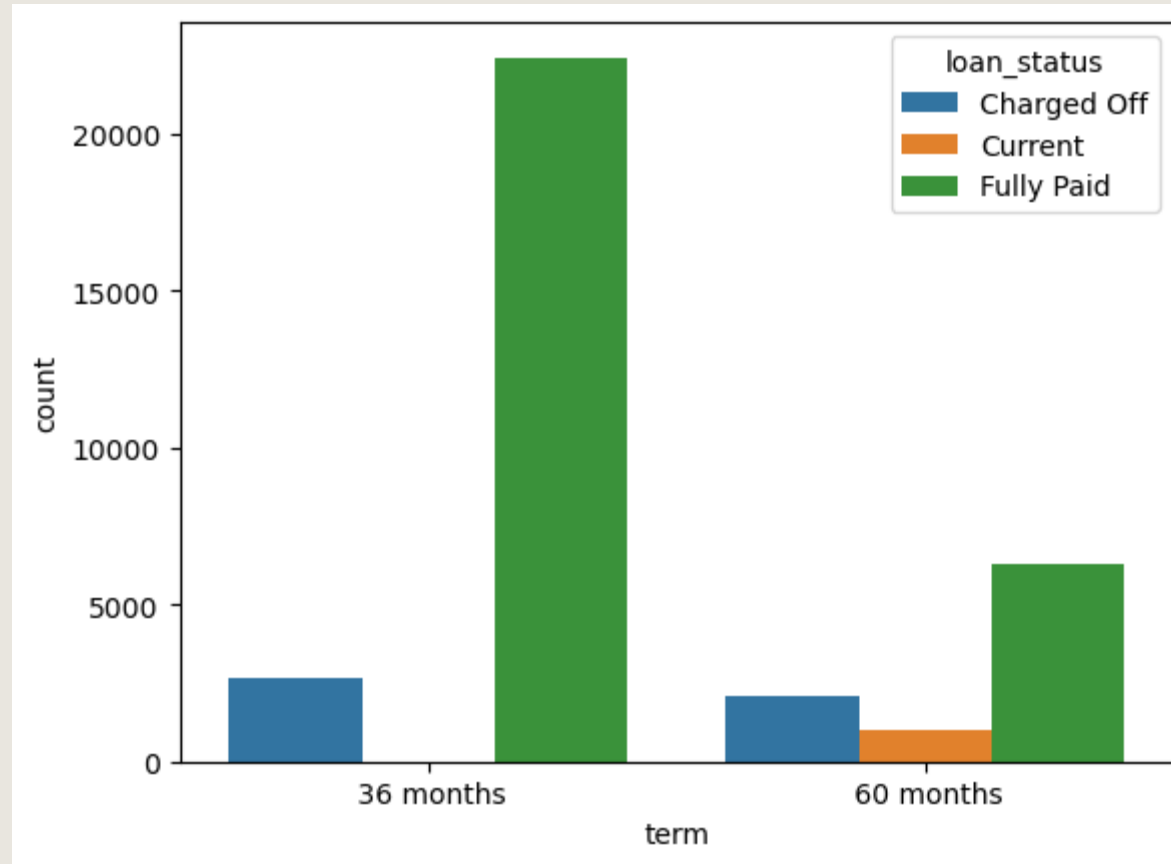
BIVARIATE ANALYSIS

Defaulters are spread almost linearly across 10% to 20% interest rate. We see less defaulters at low interest rate < 10% or high interest rate > 20%
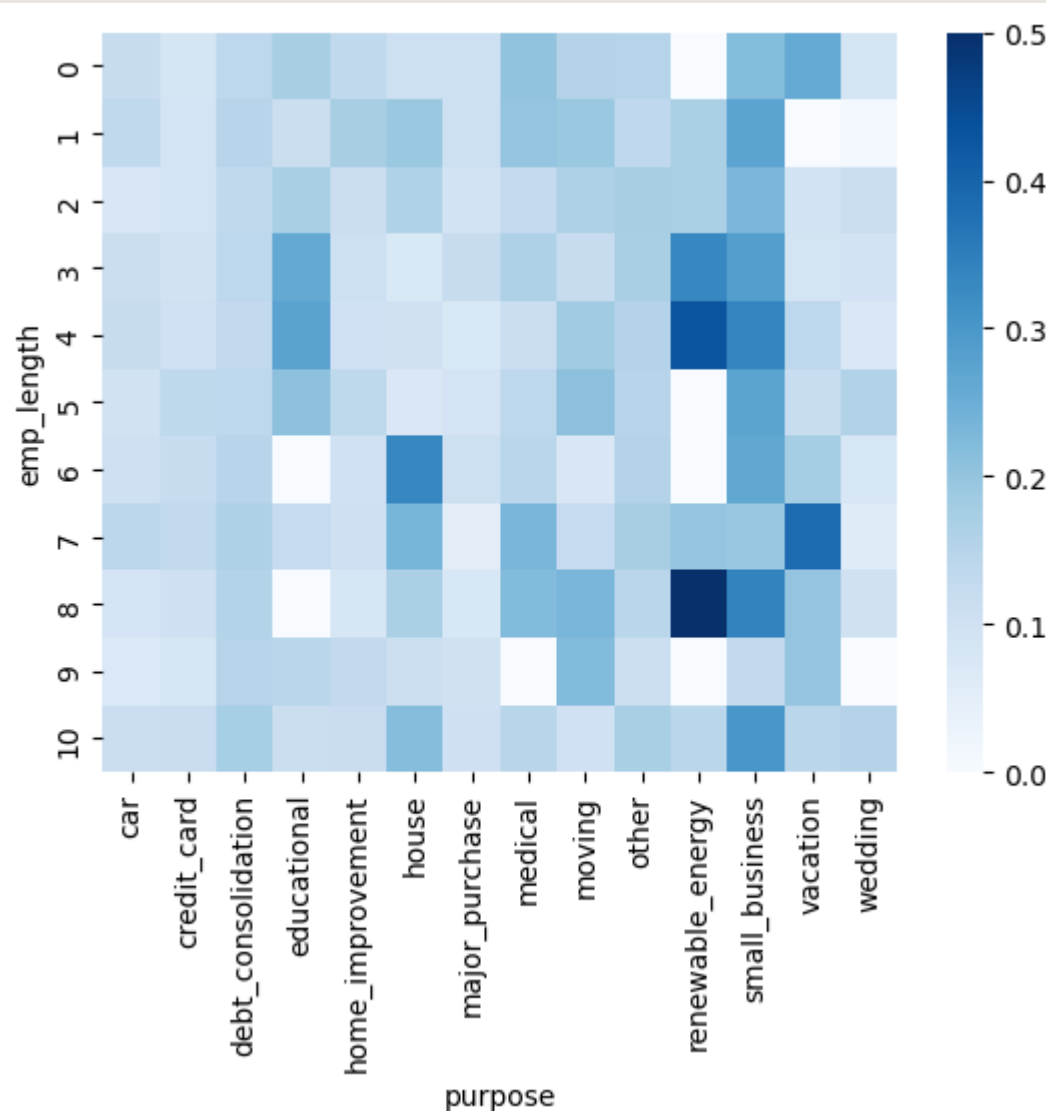
BIVARIATE ANALYSIS

People with own houses are preferable for lending loans

BIVARIATE ANALYSIS

The percentage of defaulters' vs fully paid are more for 60 months compared to 36 months

BIVARIATE ANALYSIS

Darker colors indicate stronger correlations, while lighter colors indicate weaker correlations.

Applicants who will tend to default based on purpose of loan and emp_length

- Higher educational loans for emp_length 3 to 4 years

- Renewable energy loans for emp_length 8 years

- House loans for emp_length 6 years

Most suitable purposes to consider granting loans irrespective of emp length are

- Car, credit_card home_improvement, and major_purchase

BIVARIATE ANALYSIS

Applicants who tend to default more based on home_ownership correlated to term

home_ownership as other and with loan term of 60 months

THANK YOU

Group Facilitator : Name: SreeLakshmiKanthi
Email ID: sl.kanthi@gmail.com
Phone no：9916629319

Team Member Detail: Name: Nivedita
Email ID: Niveditaganda@gmail.com
Phone no: 8447319676