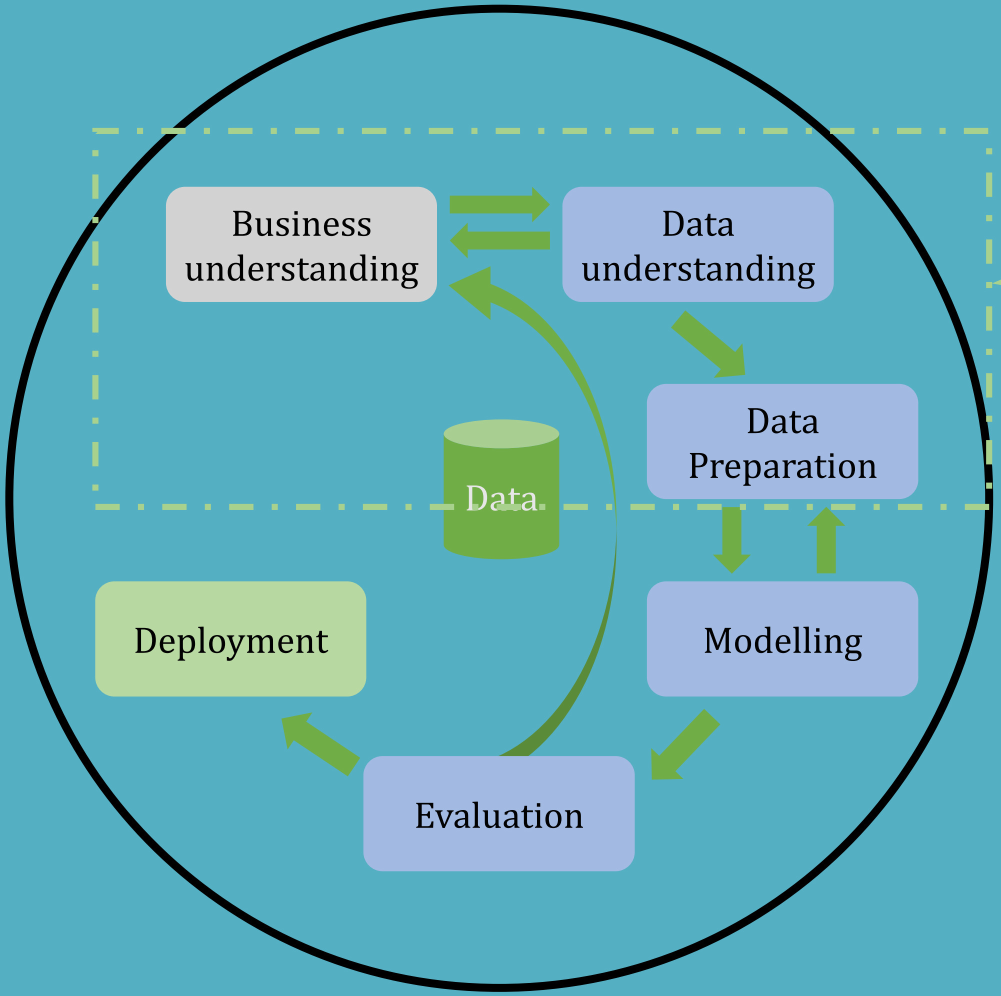
Predicting Loan Default

- Yatin Kode

CRISP DM Framework for our analysis

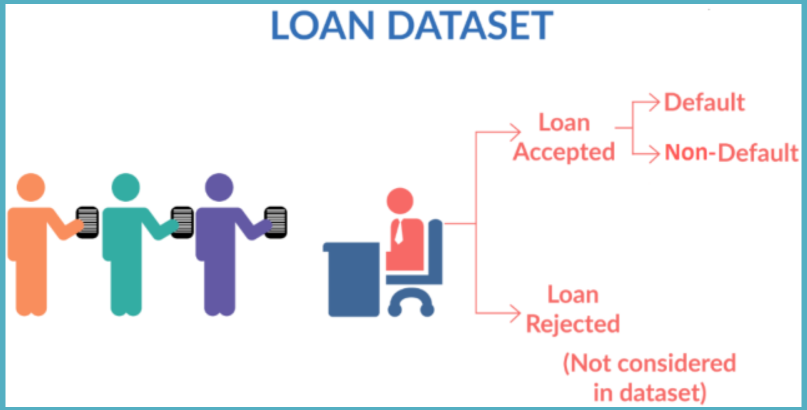


- In our studies, we have covered these aspects of CRISP-DM framework
- Our analysis will be restricted to these only.
- We will draw insights from the data based on EDA

Background and objective of analysis

Background

- Consumer finance company specializing in lending various types of loans to urban customers
- Loans approved or rejected basis **consumer attributes**



Objective

- Identify patterns to indicate if a person is likely to default, which may be used for taking actions such as denying the loan or deciding on Loan Attributes
- Identify and understand the driving factors (or driver variables) behind loan default

Loan profile – Approx 15% loans defaulting

Fig 1: Loan amount values

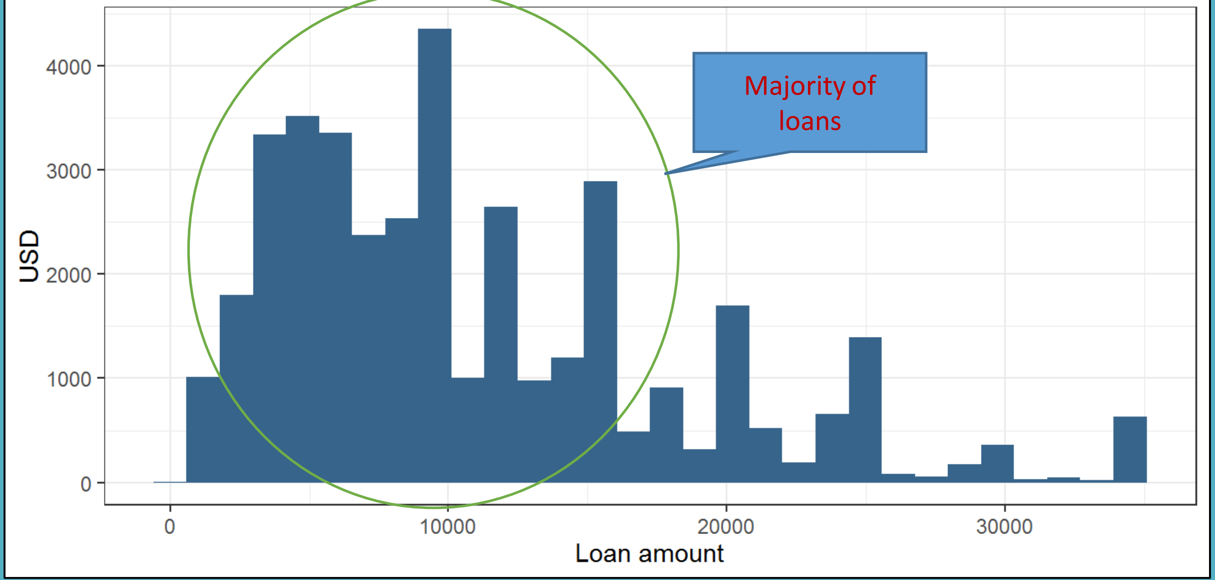
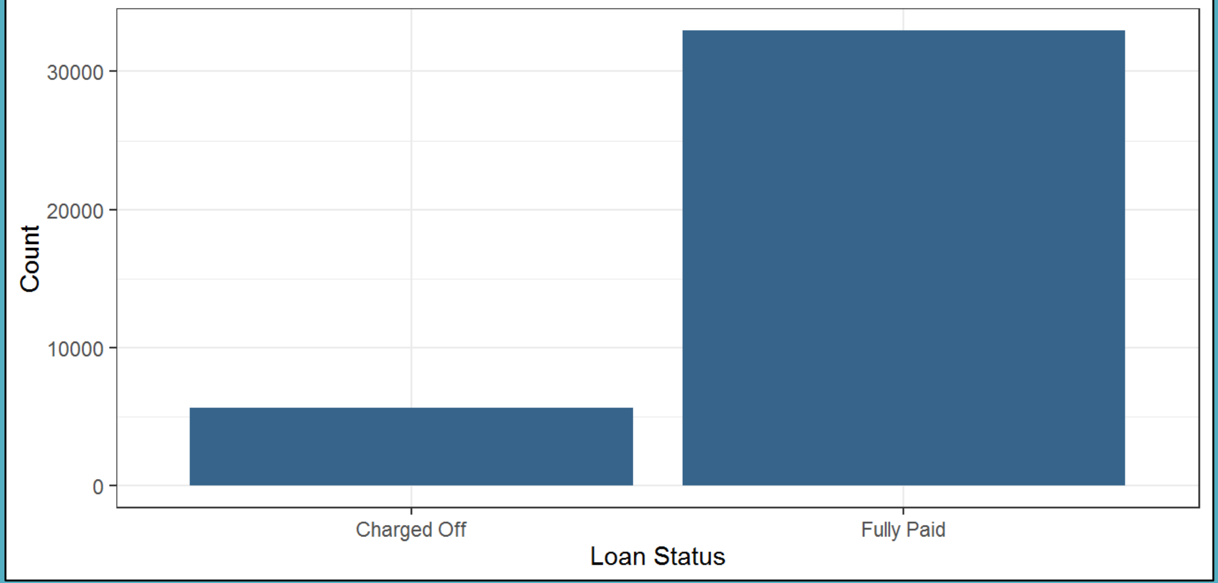


Fig 2: Loan Request Status

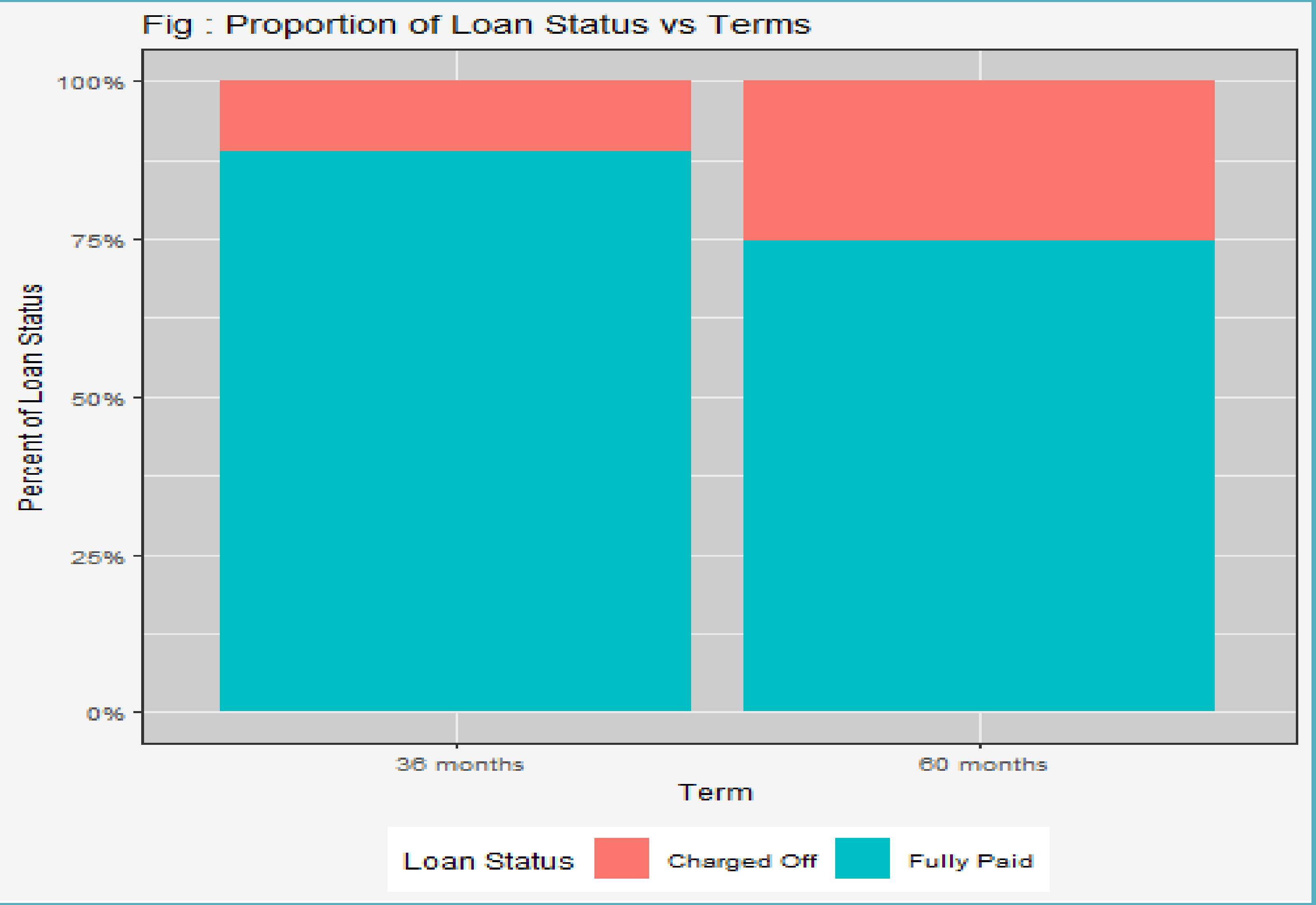


- Majority of the our loans fall in the range from 500 USD to 15,000 USD

Worrying issue

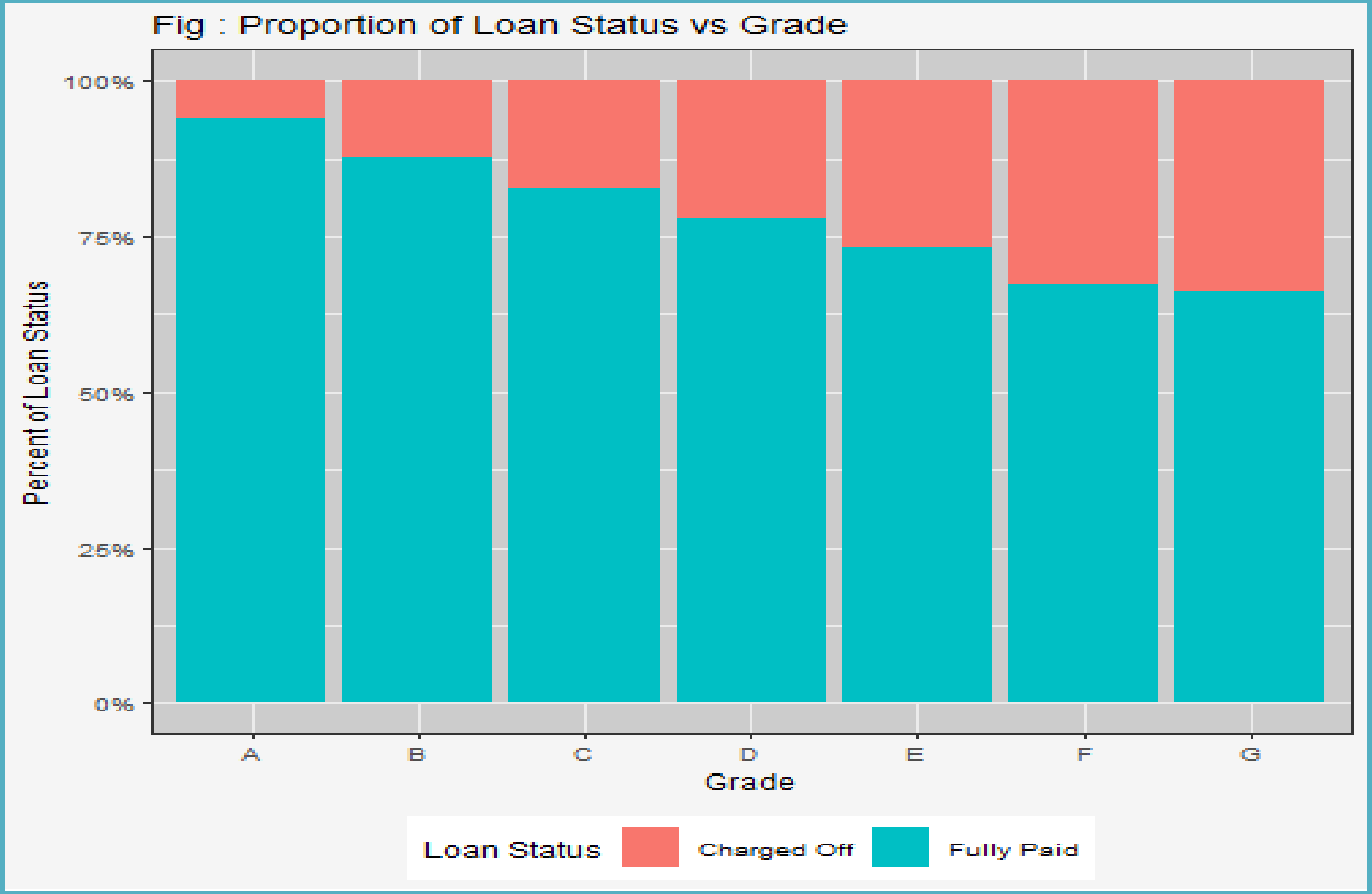
- The default on the loans is approximately 15% of total loans disbursed resulting in substantial credit loss

Plot – Proportion Of Term v/s Loan status



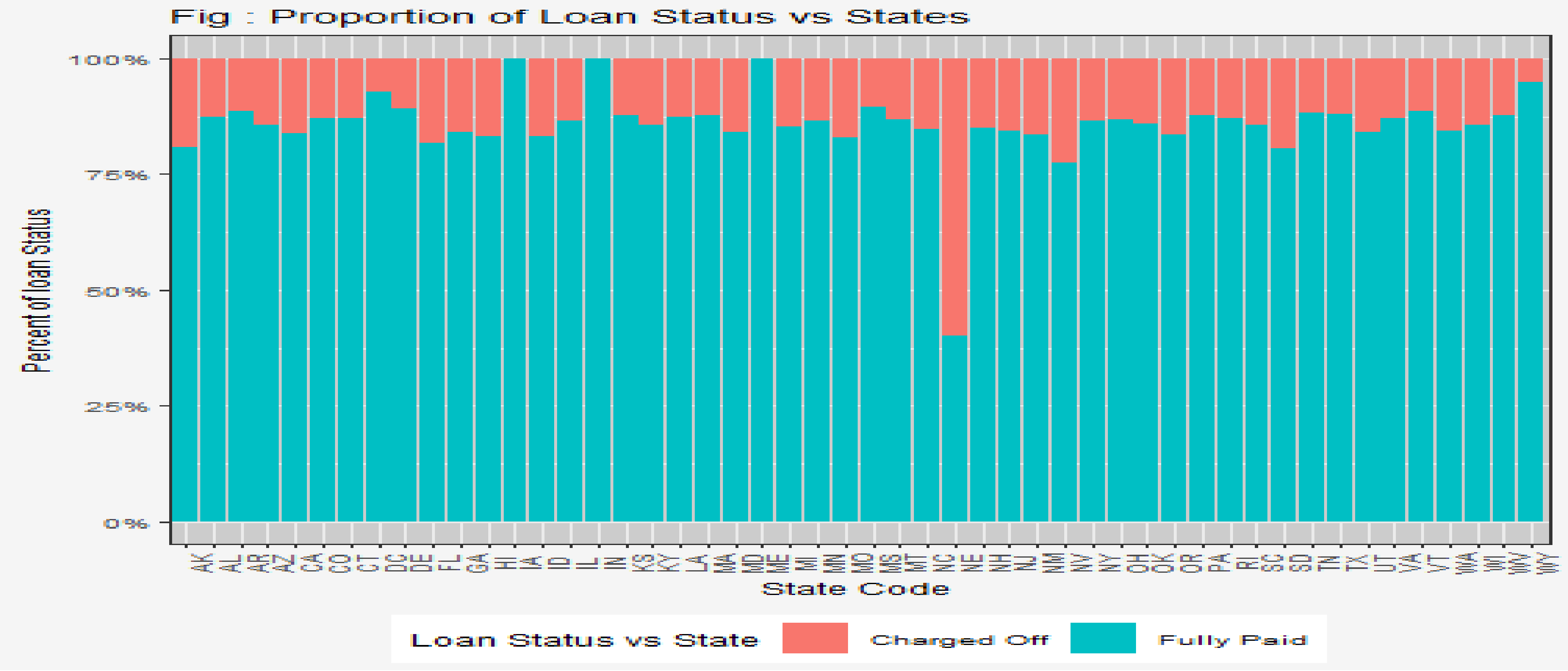
- 25% loans have been charged off with higher loan terms about 60 months ,in contrast with loan terms about 36 months.

Proportion of Grades v/s Loan status



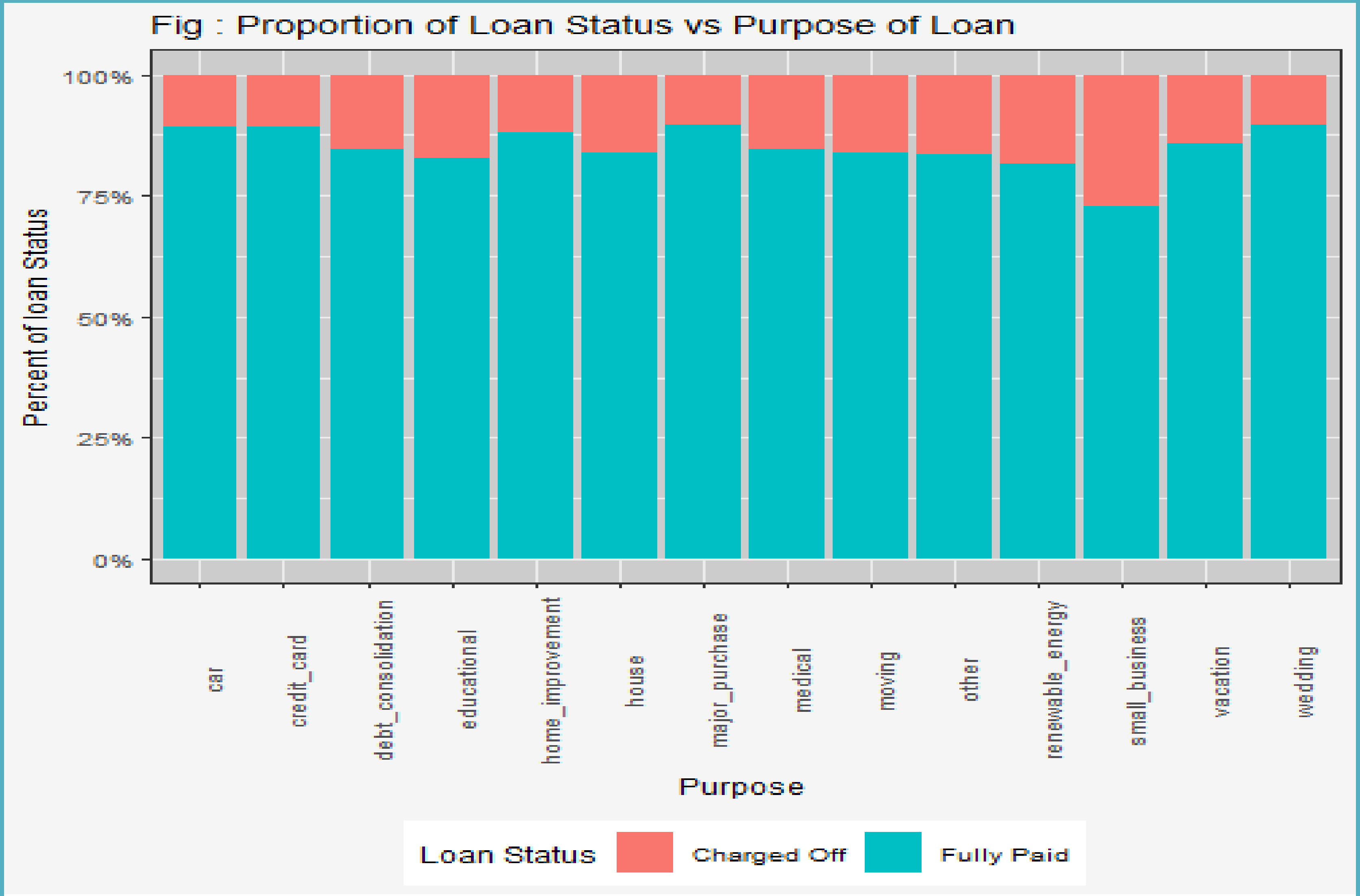
This plot shows increasing trend of defaulters along grades in ascending order from A to G respectively.

Proportion of Address States v/s Loan status



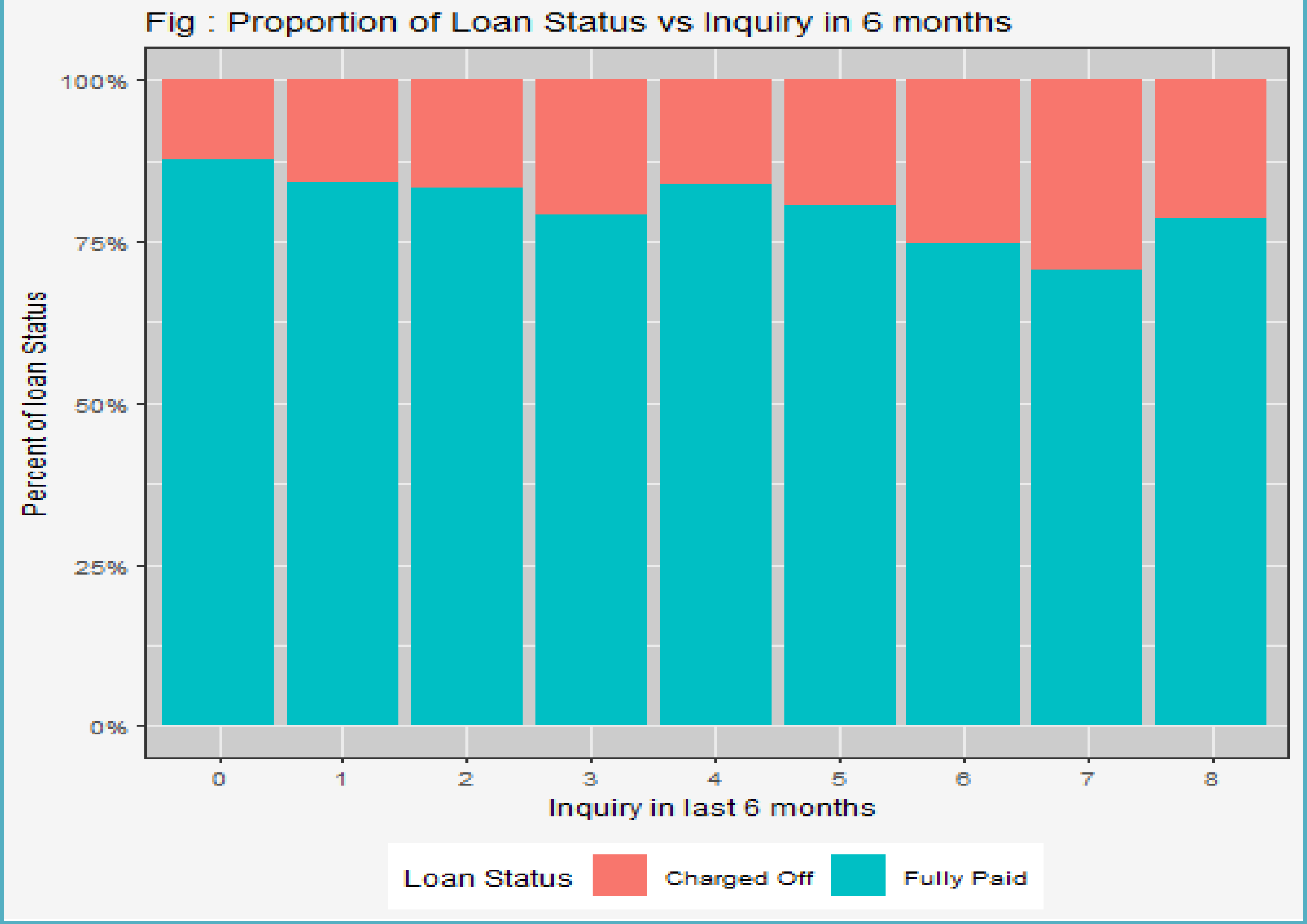
Nebraska state has clearly more defaulters so you can use it as a factor in reducing loan sanctions in that state.

Proportion of Purpose of Loan vs Loan status



We can see that small businesses tend to default more maybe to loss in the business are they might be closing up

Proportion of Inquiry in last 6 months v/s Loan status



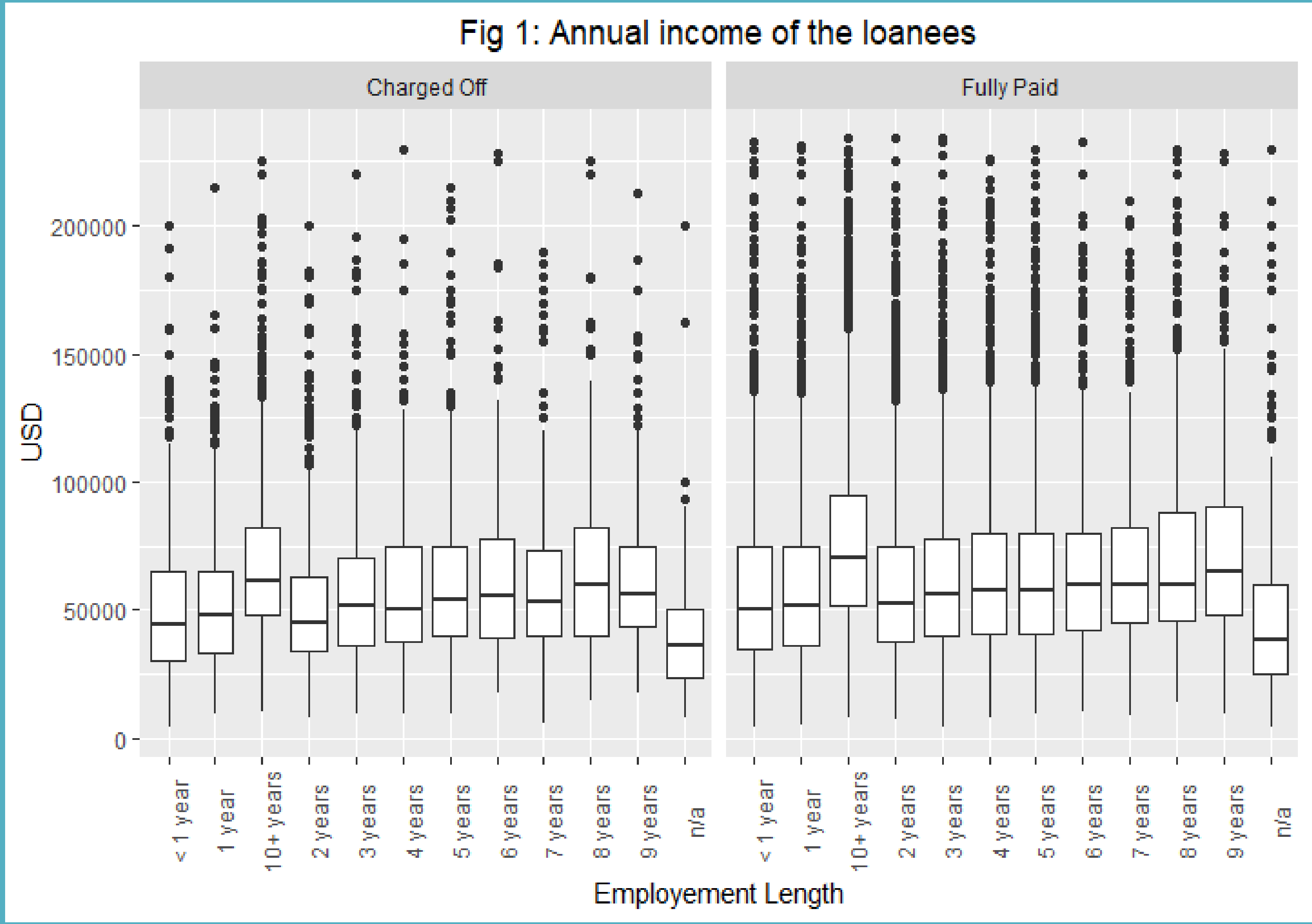
- As the Inquiry increases so as the defaulting
- This trend is a little less but still considerable

Customer profile – table

Customer attribute	Inference on loan default
Employment length	Does not have impact.
Home ownership	Does not have impact.
Annual income	Low income higher rate of defaulting.
Purpose of the loan	Small business default rate is high
Address state	Nebraska states defaults more.
Earliest credit line	Does not have impact.

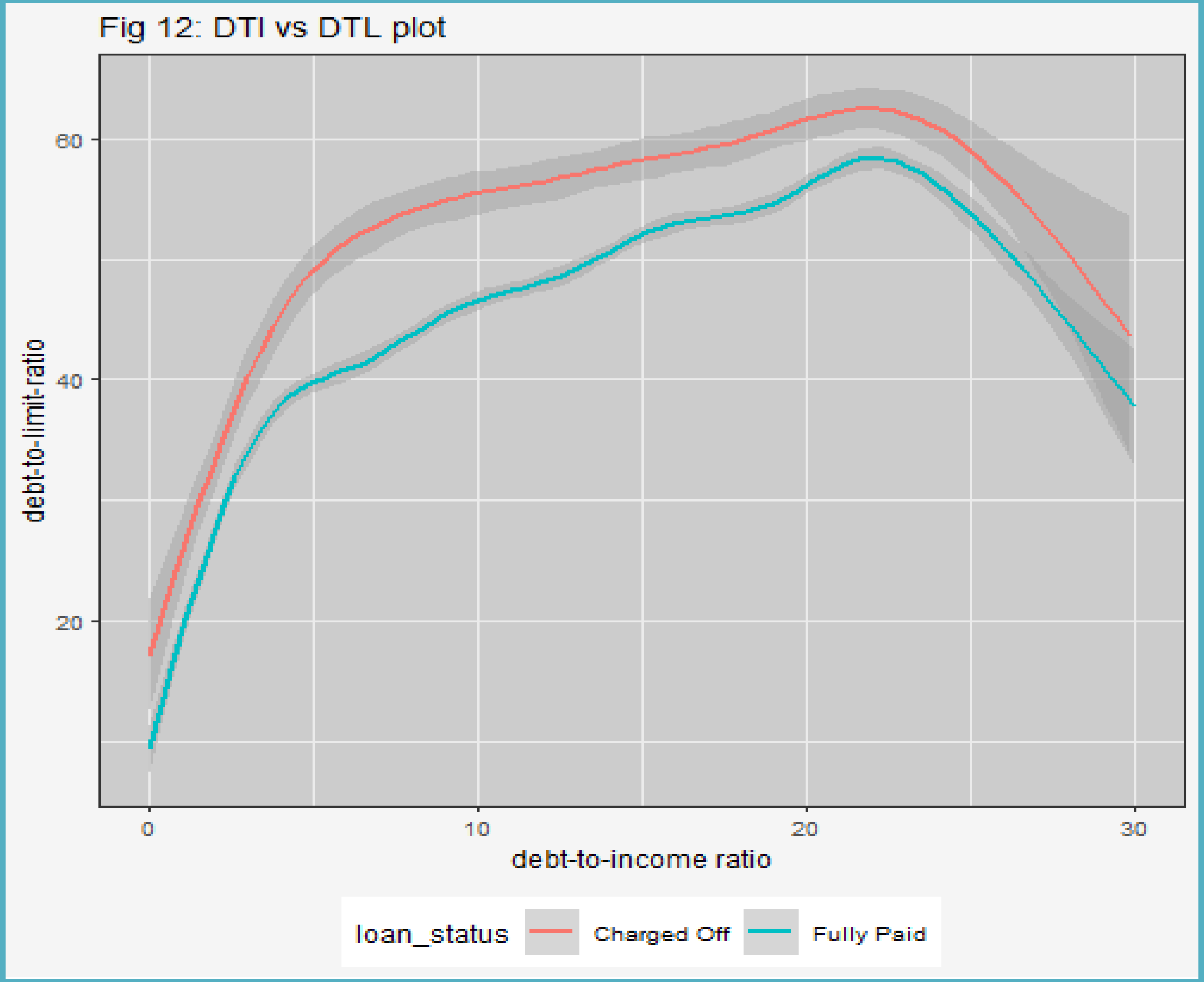
Customer attribute	Inference on loan default
Total account	Does not have impact.
Inquiry in last 6 months	Does not have impact.
Open account	Does not have impact.
Revolving balance	Does not have impact
Revolving utilization	Does not have impact.

Annual Income and Employment Length Effect



- As we see annual income of people defaulting on loans is visibly less than the one who are paying back fully.
- Annual income should be considered as one of the variable to decide on the eligibility and interest rate of the loans

Revolving utilization v/s Debt to Income ratios



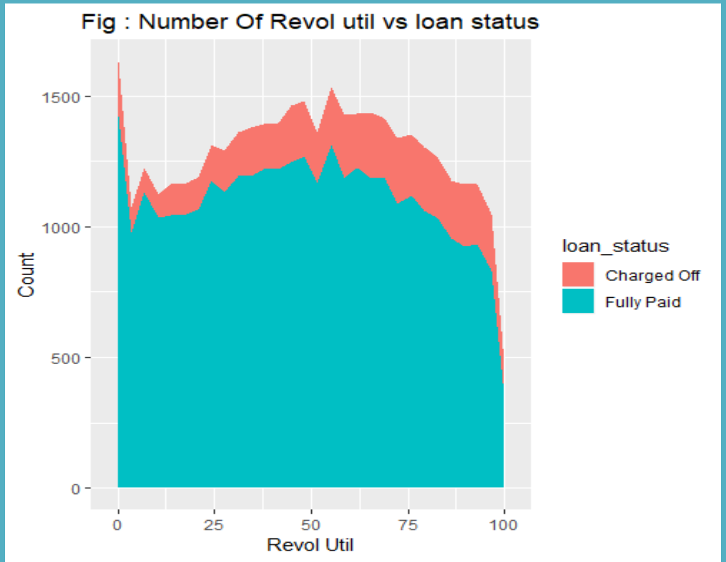
DTI must be high

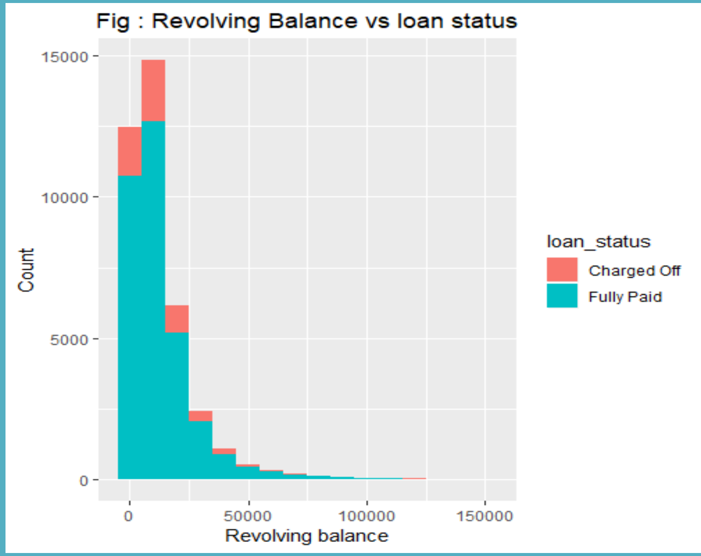
DTL(revol_util) must be low

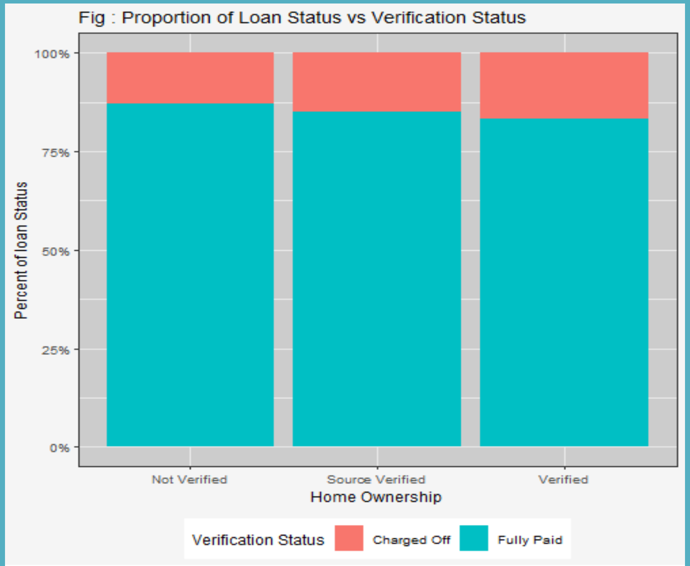
The plot clearly indicates that the entire trend for defaulters is higher than paid of trend.

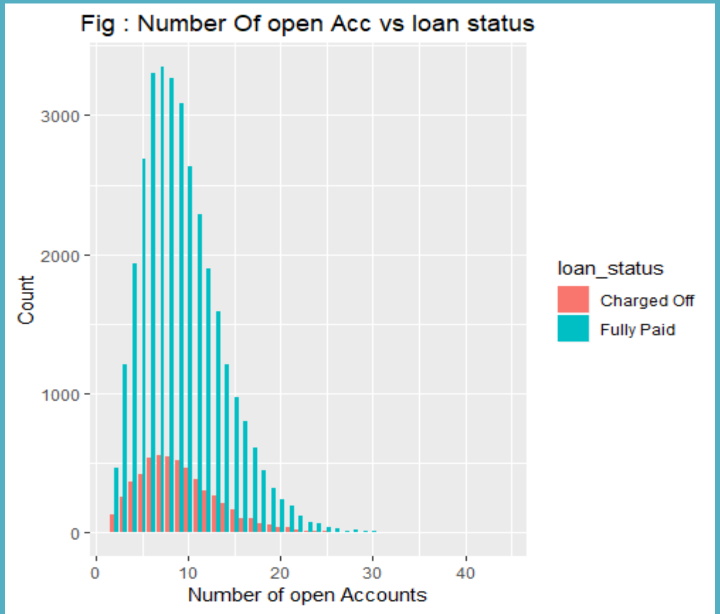
We can also infer that people within 20-25 range of DTI have a highest revolving utilization ranging between 55-65.

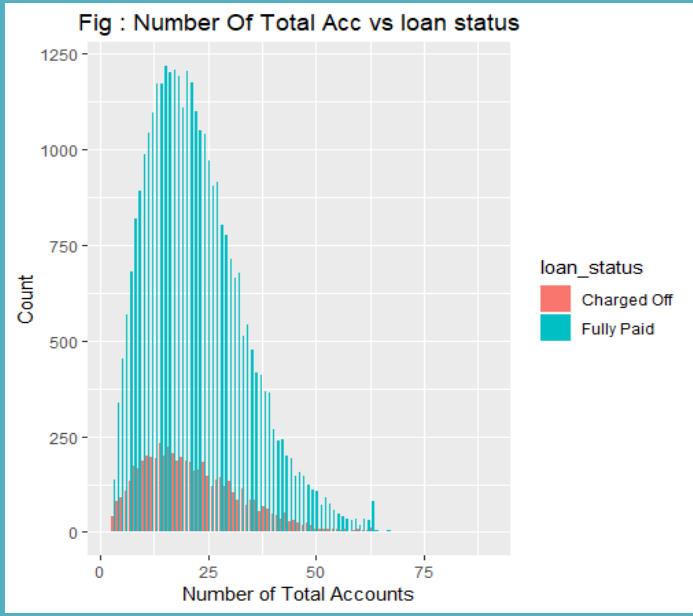
Factors Not having Effect of Loan Defaulting

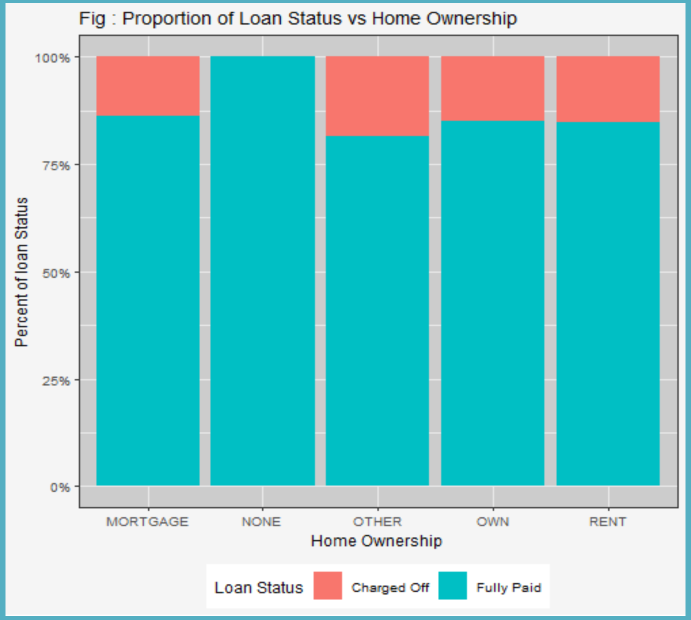












Thank You