Lending Club Case Study

Problem Statement

A consumer finance company which specializes in lending various types of loans to urban customers receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. If the applicant is not likely to repay the loan, he/she is likely to default, then approving the loan may lead to a financial loss for the company. The aim is to reduce such type of loans.

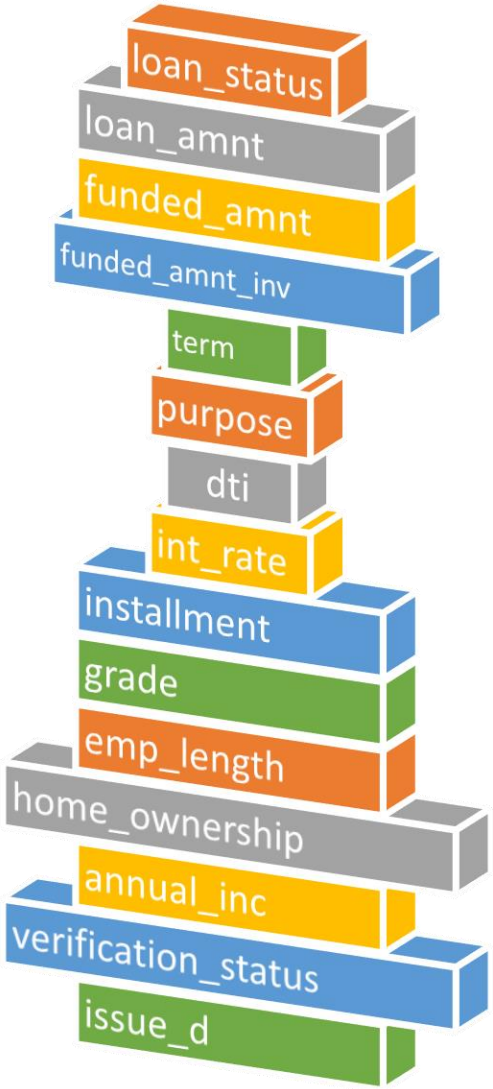
Objective

To 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. Basically the company wants to understand the **driving factors (or driver variables)** behind loan default.

Approach

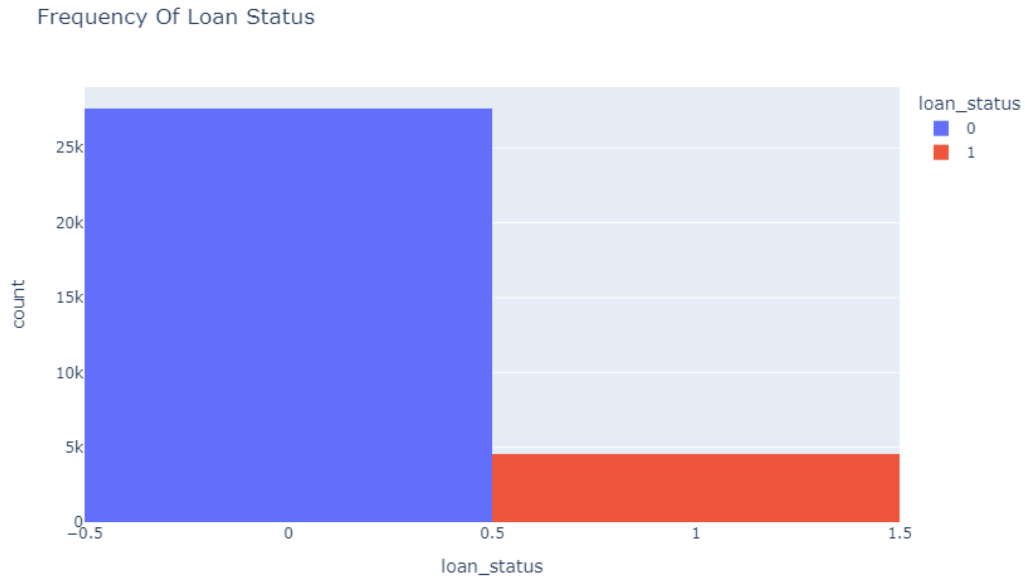
Understand The Problem And Objective
Get Some Business Context Of Risk Analytics
Understand the data dictionary attributes
Find out the customer behavioural, customer demographic and Loan Characteristics
Source the loan data and high level data check
Data Cleaning – Remove missing values, null value columns
Data Cleaning – Data Type Correction e.g interest rate
Data Cleaning – Remove behavioural attributes
Data Cleaning - Add derived Columns e.g Year and Month
Data Cleaning- Data Sanity Check by business rules e.g Removed records having Loan Amount > Funded Amount
Data Cleaning - Outlier Detection and Removal using BoxPlot e.g annual_inc
Data Cleaning – Remove Records having current loan status
Univariate Analysis e.g Grade , loan amount using Binning
Segmented Univariate Analysis e.g Purpose
Bivariate Analysis – Correlation Matrix
Conclusion

Driving Variables


loan_status
loan_amnt
funded_amnt
funded_amnt_inv
term
purpose
dti
int_rate
installment
grade
emp_length
home_ownership
annual_inc
verification_status
issue_d

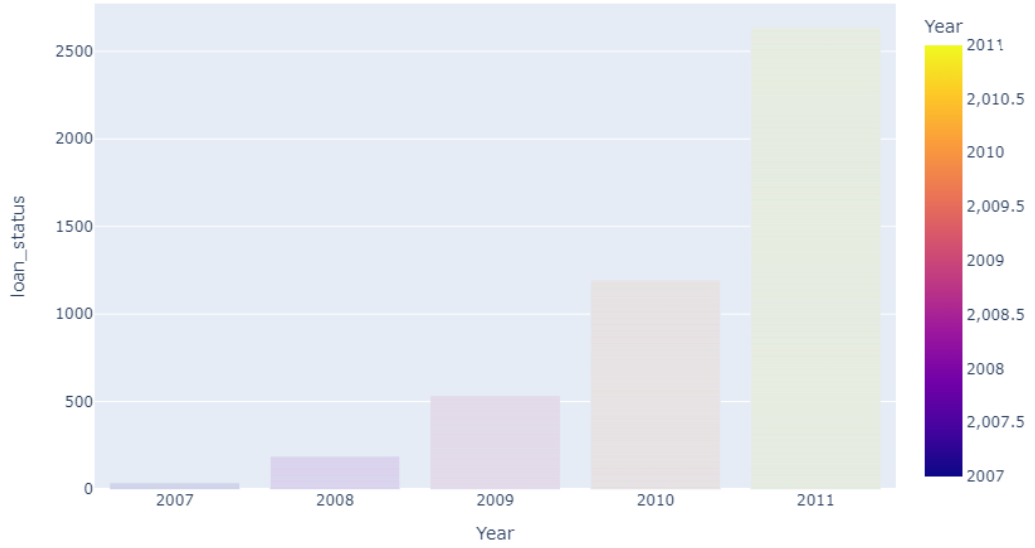
Analysis & Findings – Loan Status

- Around 14% of loan is defaulted



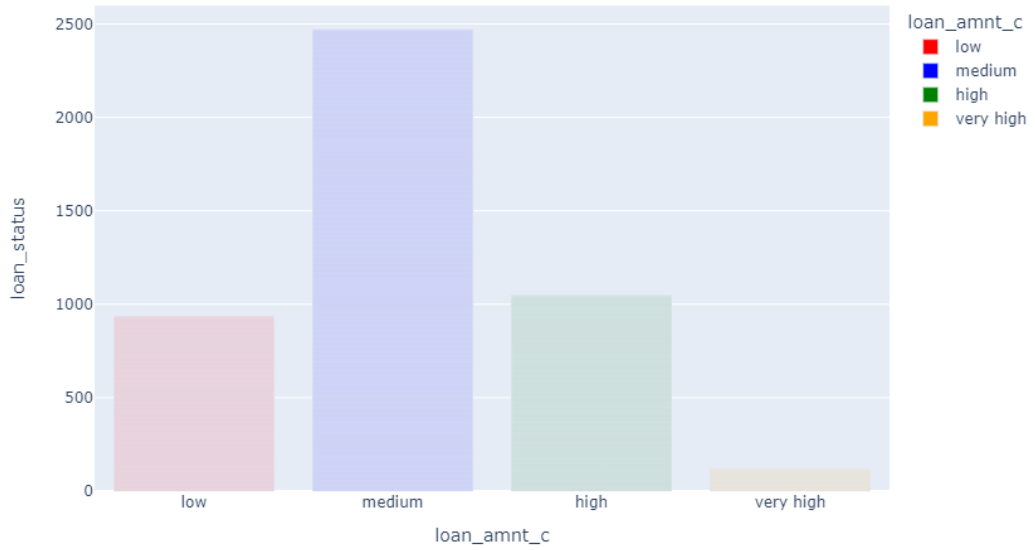
Analysis & Findings – Year

- Default rate increases year by year



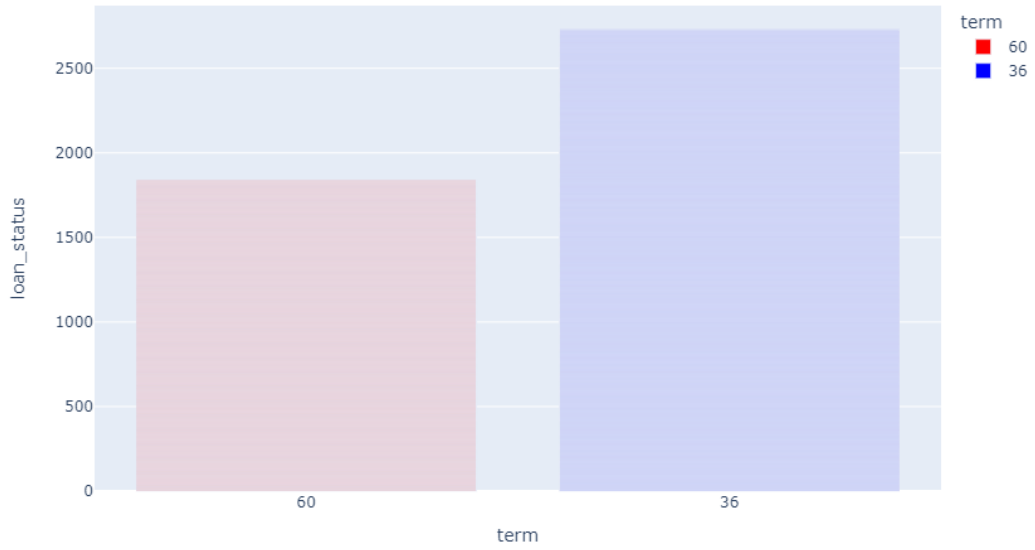
Analysis & Findings – Loan Amount

- Higher the loan amount is more the risk of default. And medium and high ranged loans are more prone to default.



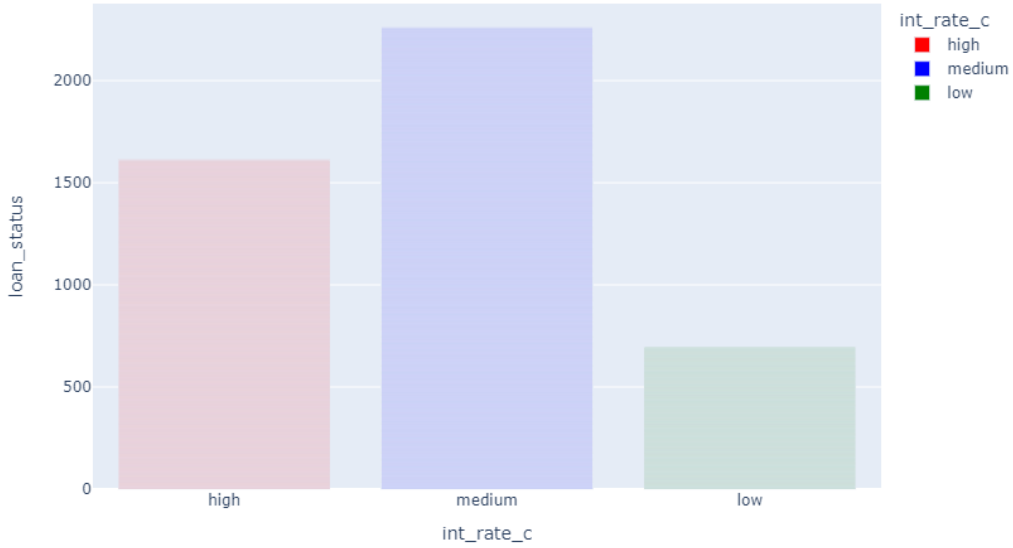
Analysis & Findings - Term

- Higher the term is lesser the default rate.
- 36 months term default rate is higher than 60 months term



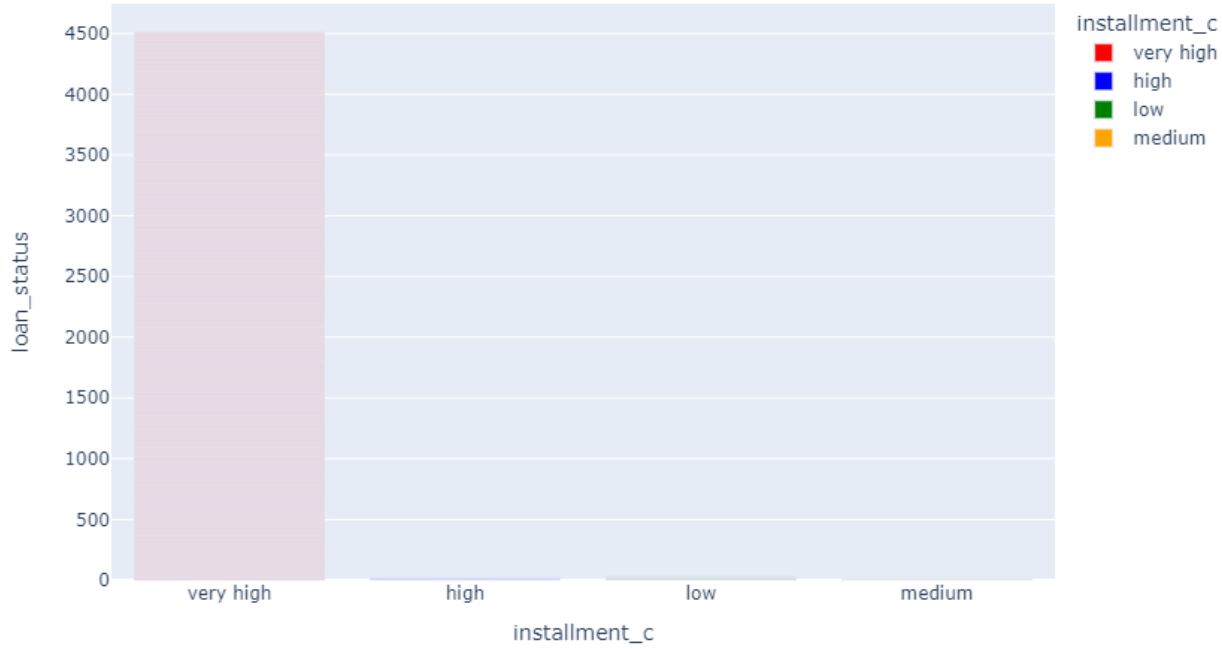
Analysis & Findings – Interest Rate

Default rate increases with interest rate



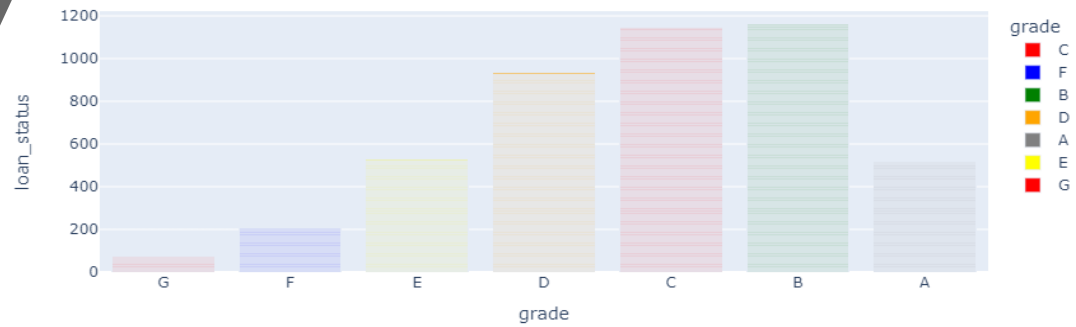
Analysis & Findings – Installment

Higher the installment amount leads to higher number of defaults



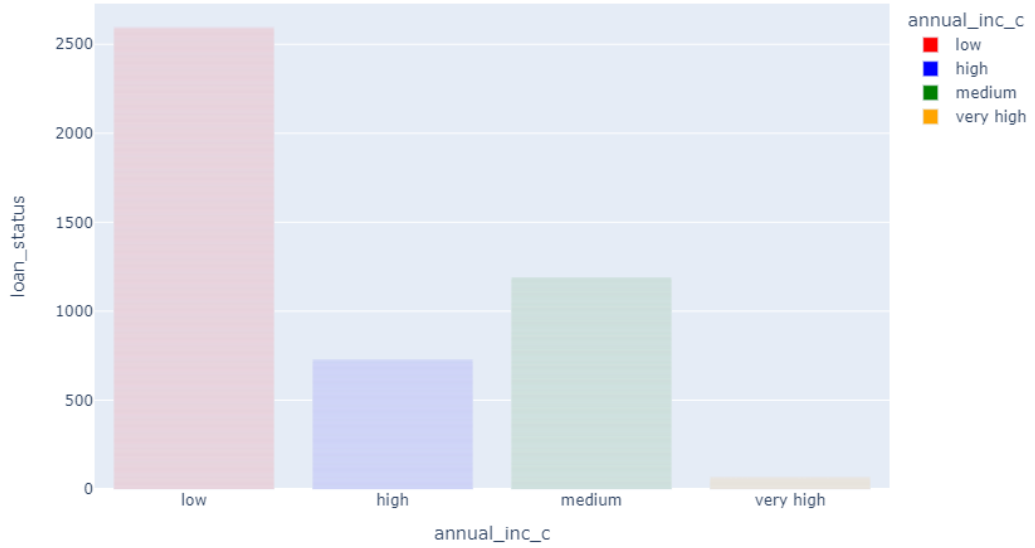
Analysis & Findings – Grade

Low Grade Applicant has high default rate



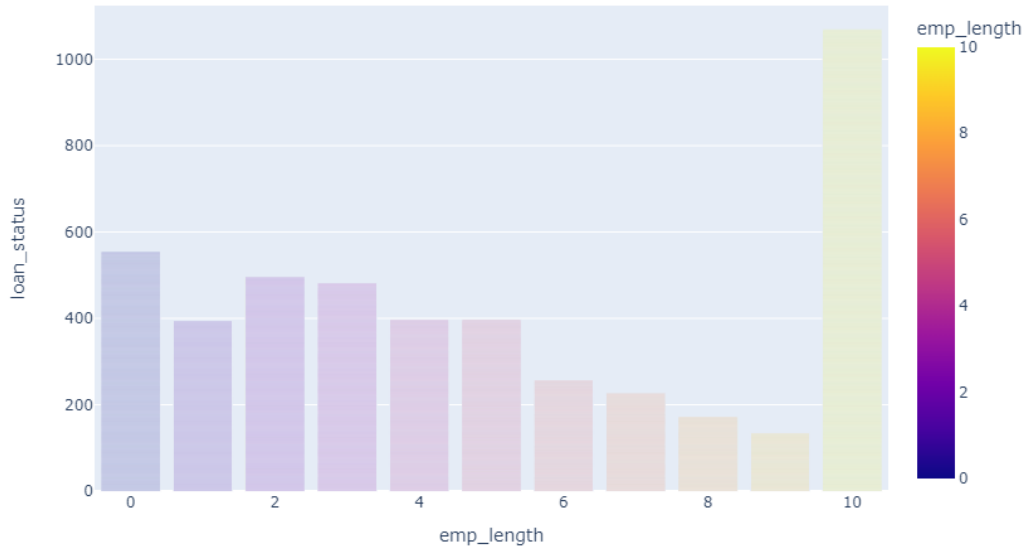
Analysis & Findings – Annual Income

Highly Income Applicant has less default rate



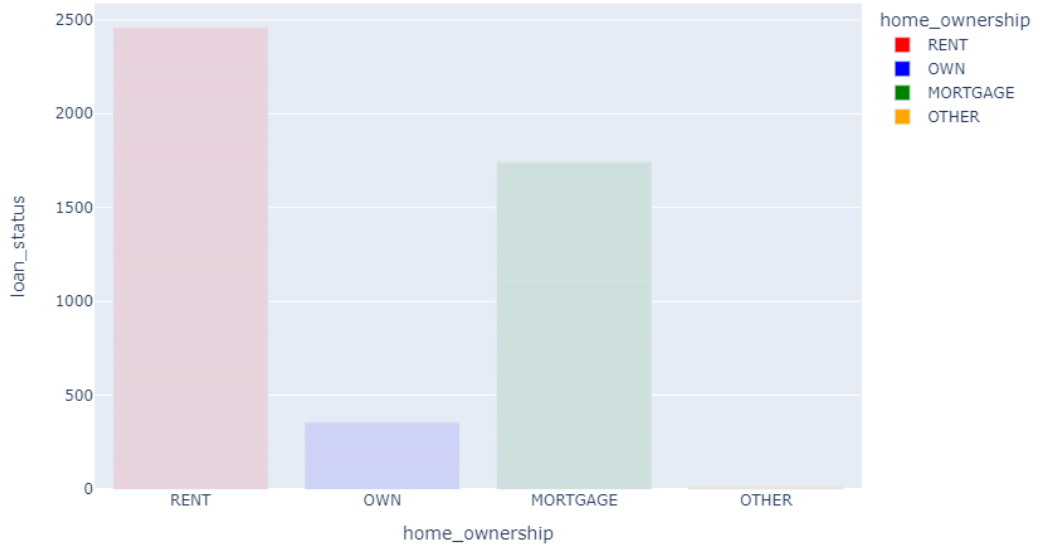
Analysis & Findings – Applicant Experience

Interestingly applicant having more than 10 years or less than 2-3 years have high default rate



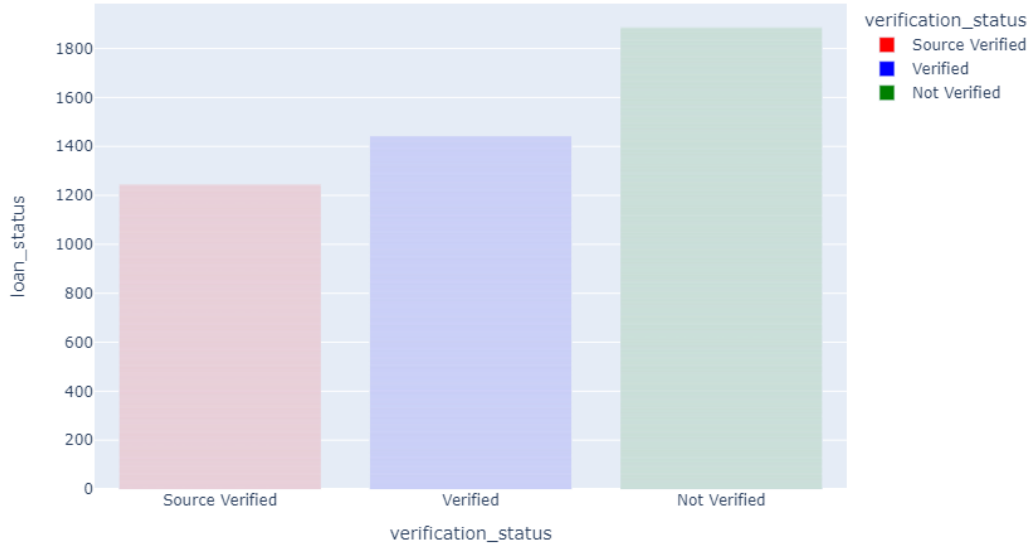
Analysis & Findings – Home Ownership

Rented applicants have high default rate



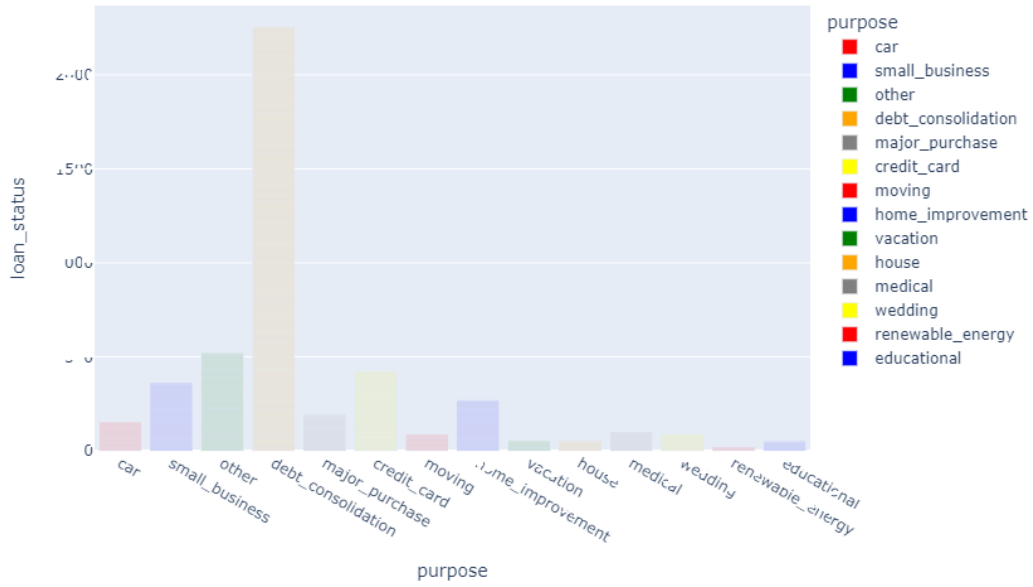
Analysis & Findings – Verification Status

Both verified applicants have higher default rate



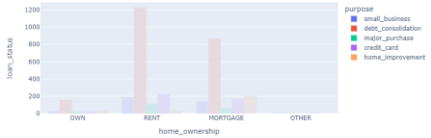
Analysis & Findings – Purpose

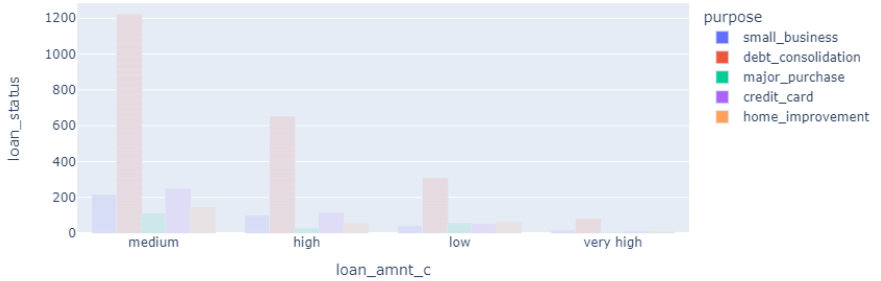
- **Debt consolidation, credit card, small business, home improvement and major purchase purpose loans are top five default loan**
- **Debt consolidation has the highest default rate**

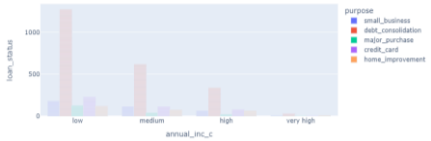


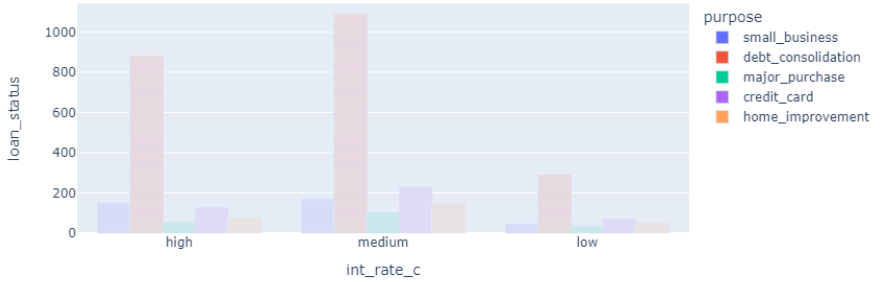
Segmented Analysis & Findings – Purpose

- **Debt consolidation - Term, Grade, Interest Rate, Loan Amount impact a lot to default rate for debt consolidation**
- **Credit card - Grade , Interest Rate, Loan Amount , Debt To Income Ratio impact a lot to default rate for Credit card**
- **small business – DTI, Interest Rate, Annual Income and Home Ownership impact a lot to default rate for small business**



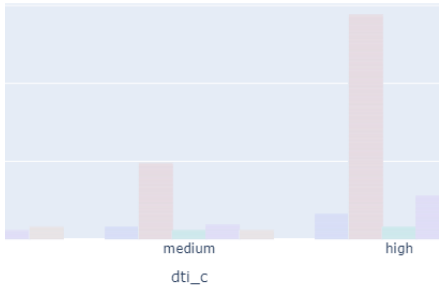


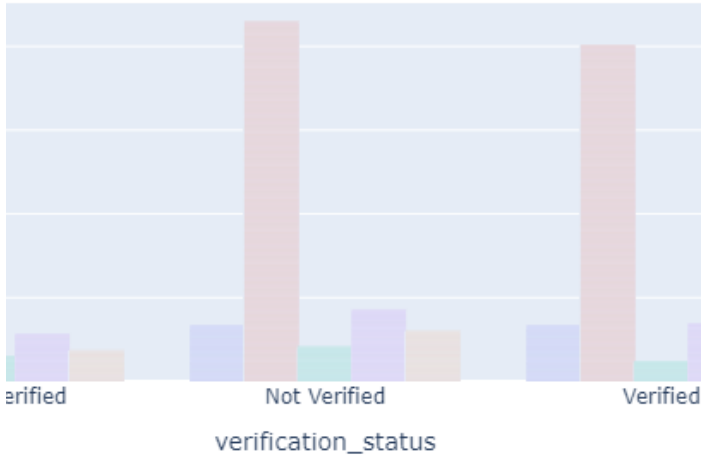


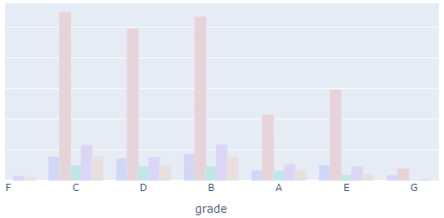


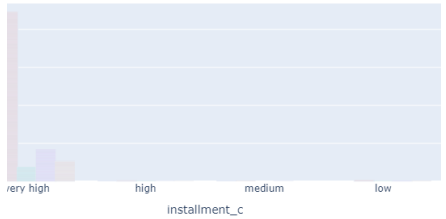
Segmented Analysis & Findings – Purpose

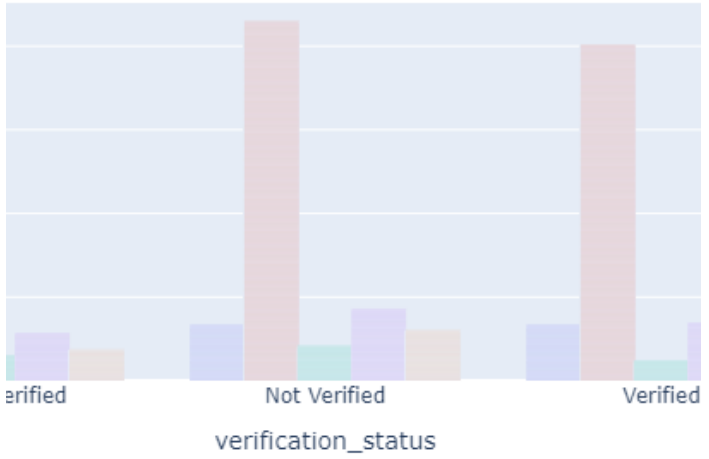
- home improvement - Home Ownership , annual income, interest rate and term impact a lot to default rate for home improvement . And interestingly mortgage home owners have high default rate
- major purchase – Interest Rate, Term and Grade are key variables for high default rate of major purchase loan











Thank You