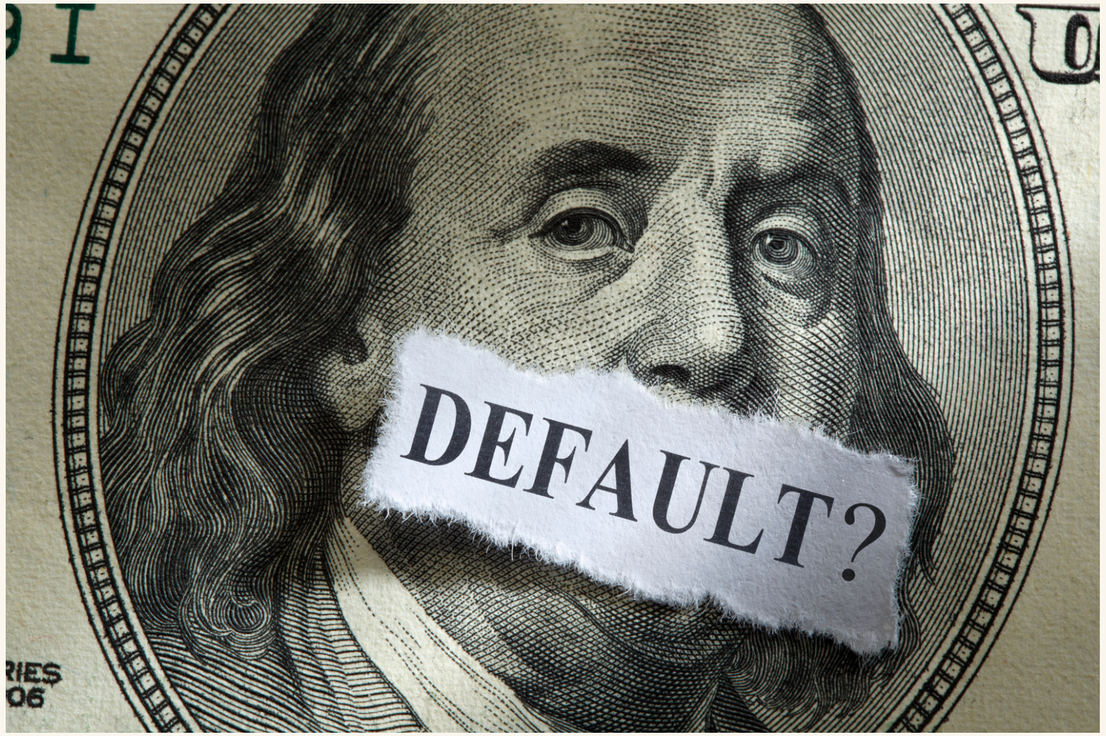
Lending Club
Case Study

Purpose of the Case Study

Understanding the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default



The Data

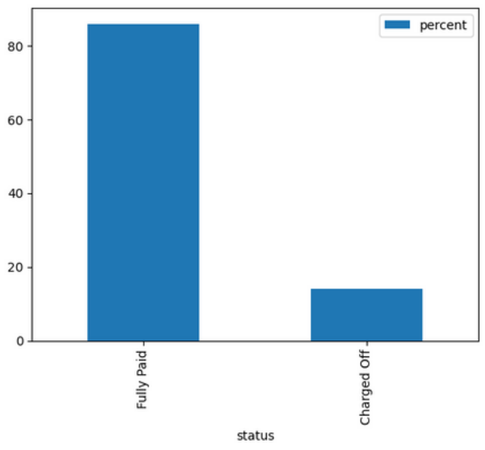
Raw Data

We were given Loan data of 4 years. From 2007 to 2011
The Data contained all the details of the loans which were approved and disbursed.
We were dealing with a huge data of 111 different variables.

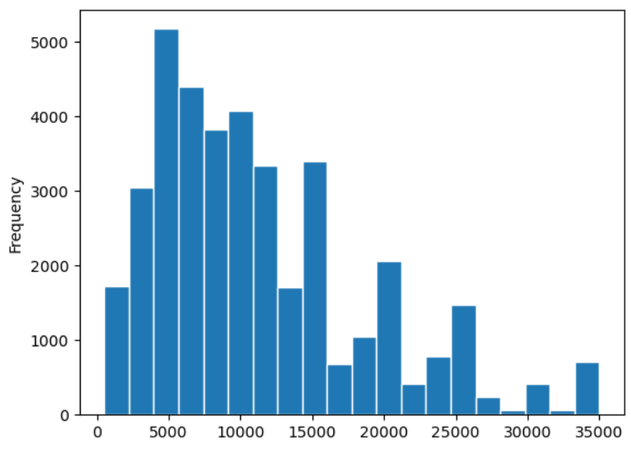
Clean Data

We Cleaned the data by removing all the unwanted variables for this case study. For eg. all the details which we get after the loan is approved is not needed for this case study as we want to predict the default before the approval. Also the records which were current were also ignored as we dont know how those would end.

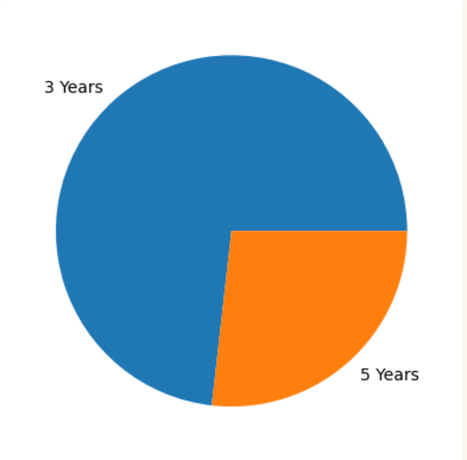
Insights of Univariate Analysis



14% of the closed loans were closed as 'Charged Off'

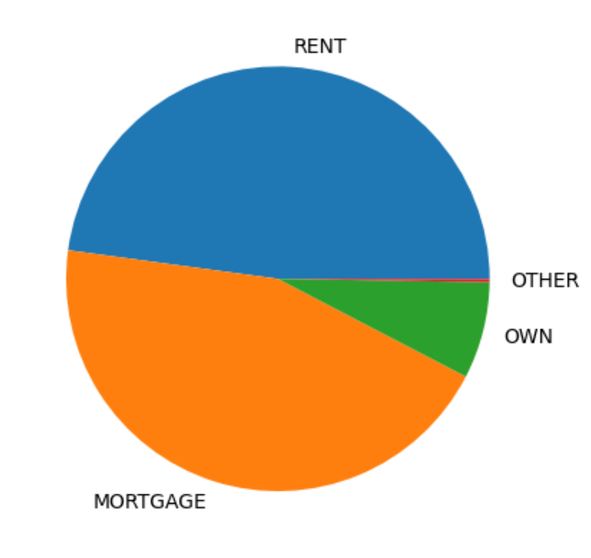


Most of the loans were between $5000 and $1000

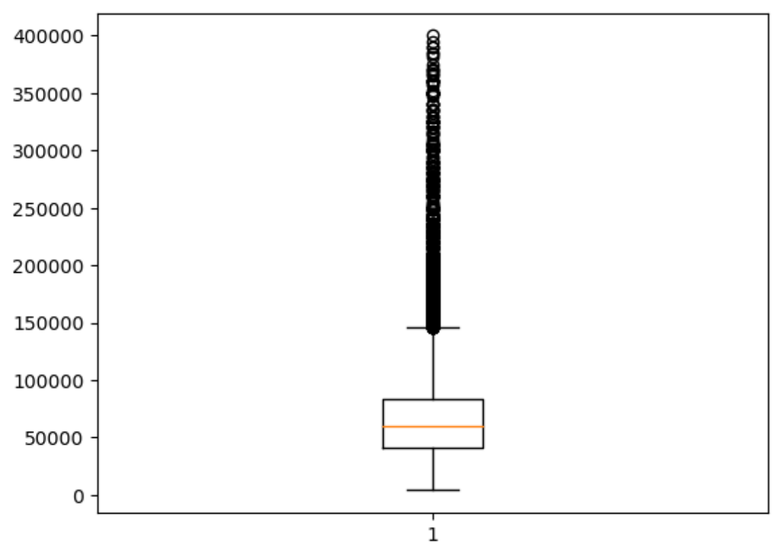


Approximately 1/3rd of the loans were of 3 years term period

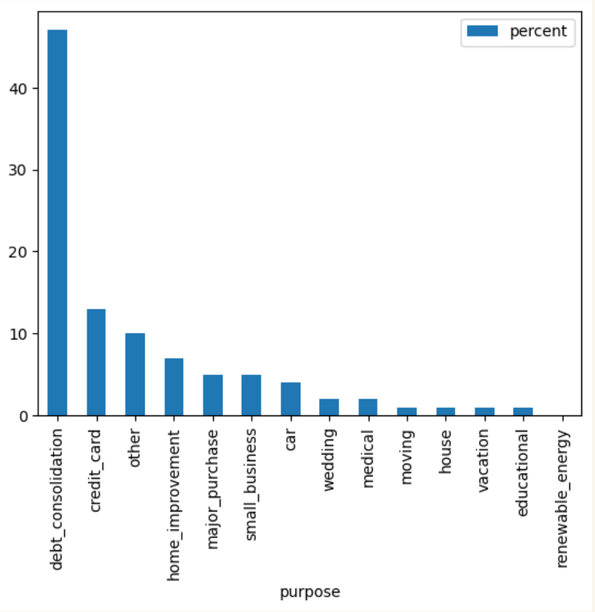
Insights of Univariate Analysis



85% to 90% of loans were taken by those renting or paying mortgage



About 50% of the loans were taken by those with annual income approx. between $40,000 to $80,000

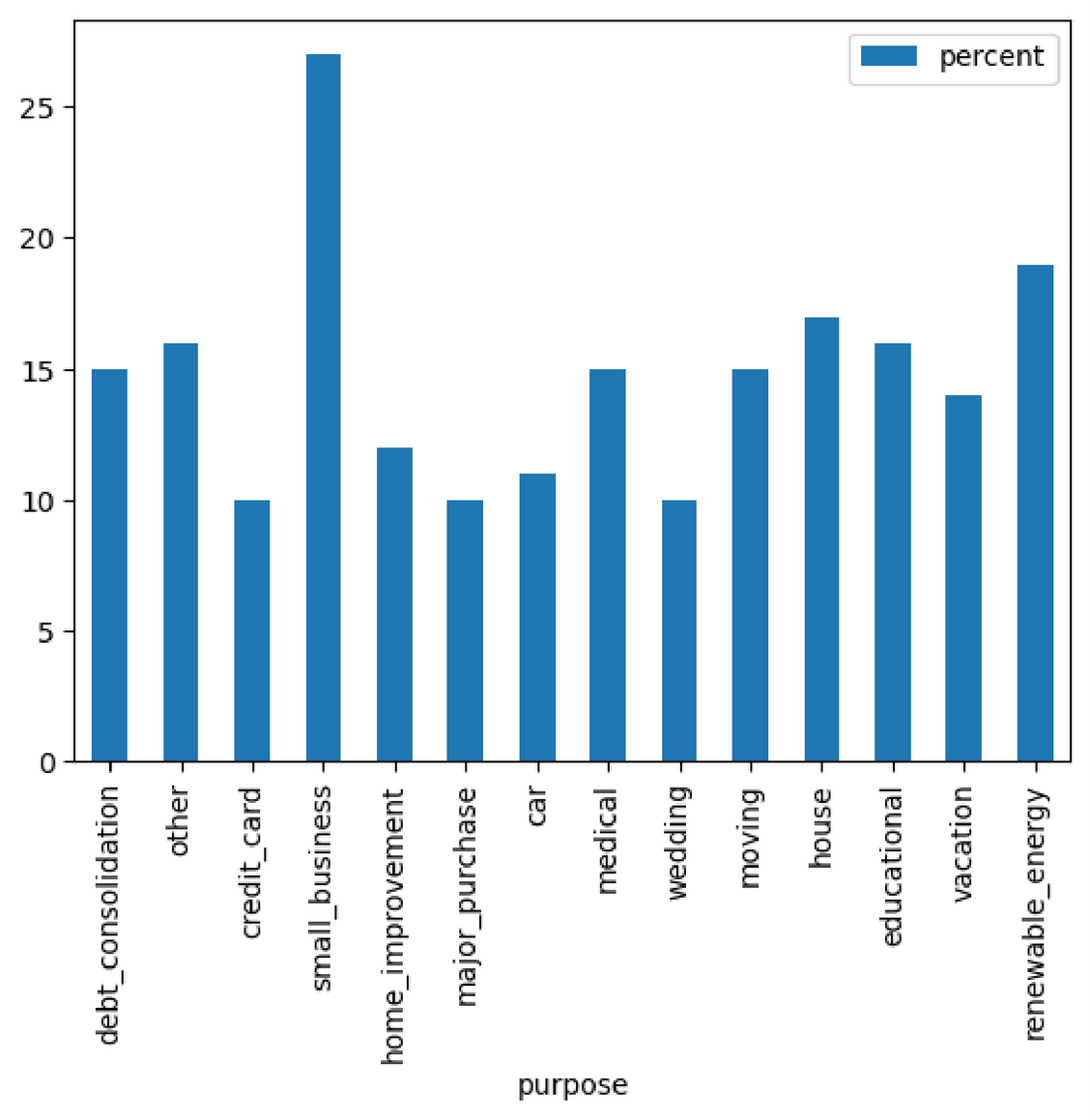


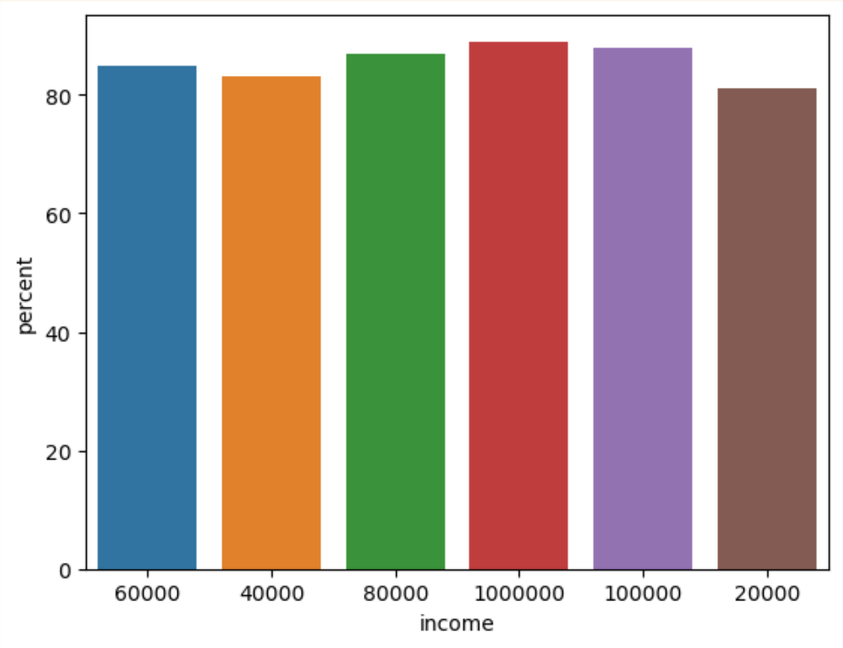
'debt consolidation' loans took a major chunk of all the approved loans

SEGMENTED UNIVARIATE ANALYSIS

Purpose of Loan

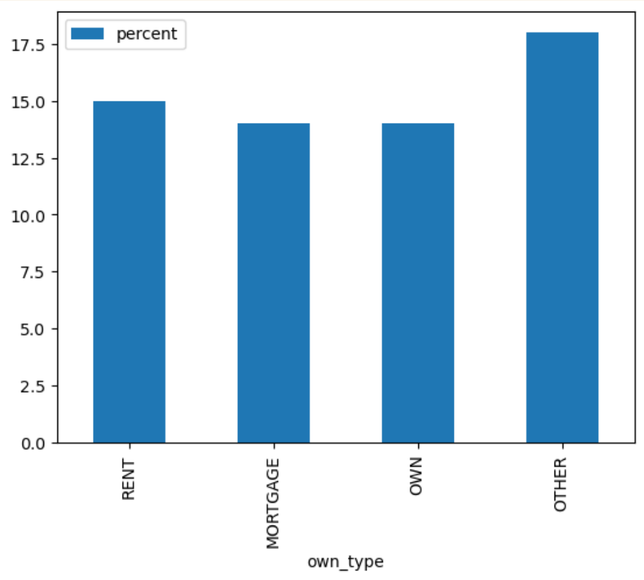
According to the data, the riskiest loan is 'Small Business loans'. A whopping 27% of these loans defaulted. The Safest were the 'Credit Card', 'Major Purchase' and 'Wedding' with a default percentage of 10%. Making them the safest loan.





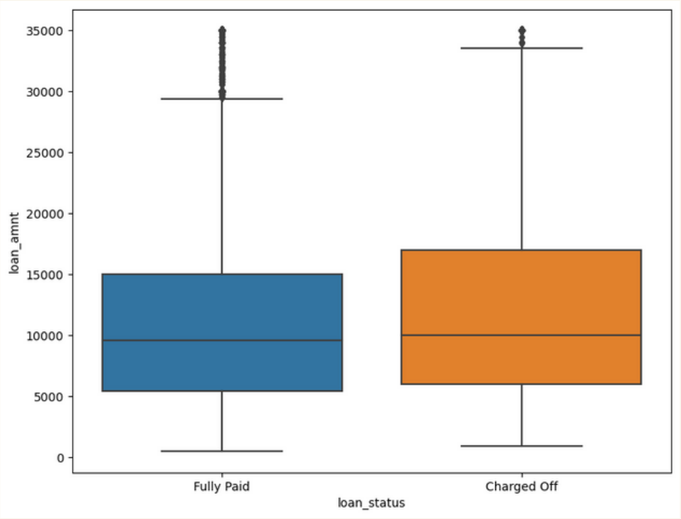
Annual Income

There seems to be a positive co-relation with Annual Income and the Loan Payment. Those with higher income show better chances of loan payment. Only 83% people with annual income lower than $40k Fully paid against 88% of those with annual income of about $100000



Home Ownership

The relationship of Home Ownership with Loan Default was neutral. About 14% to 15% of those who own, rent or mortgage default on loan. Those in the other category defaulted 18%, making it the riskiest and hence would need to be looked at more carefully

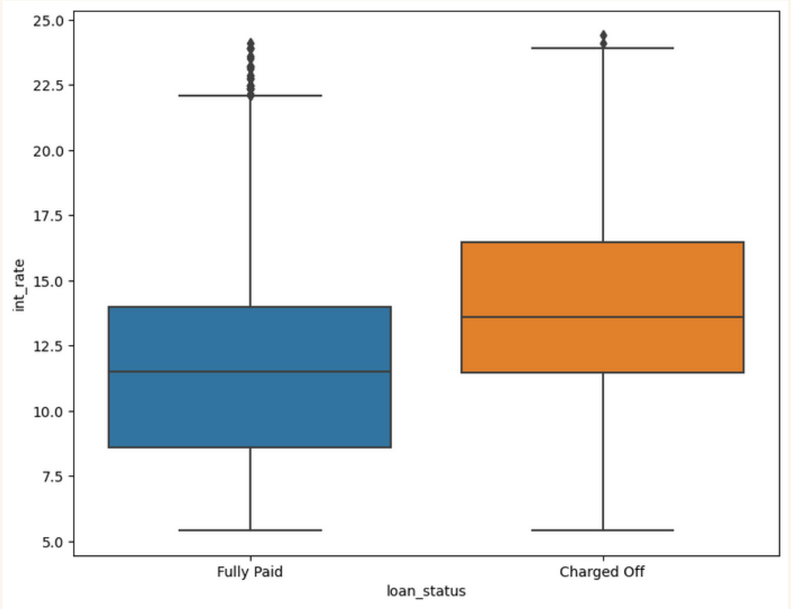


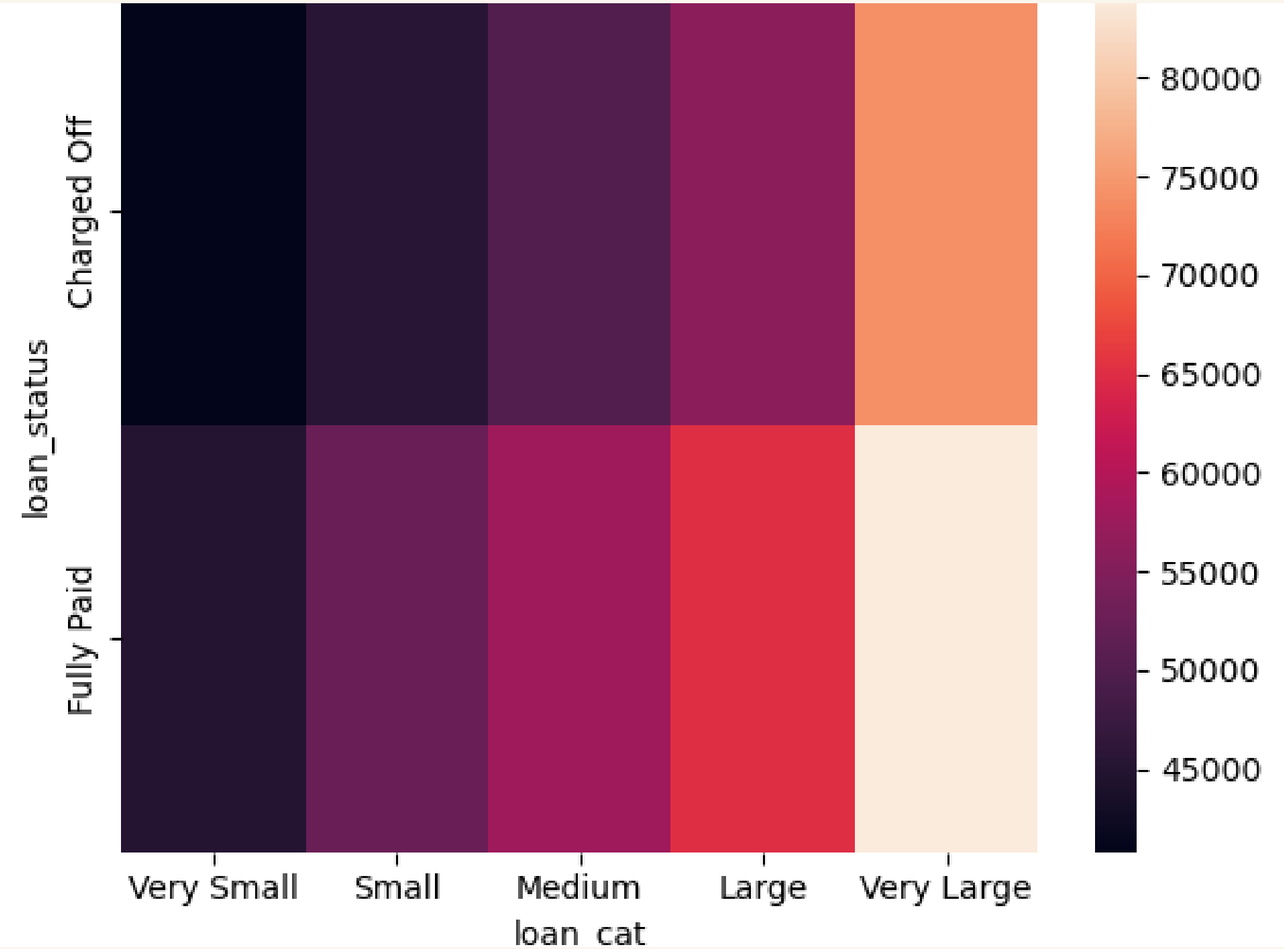
Loan Amount

As seen from the boxplot the loans which defaulted on an average are the ones which are of slightly higher value than the ones which were fully paid. The top bar which is the 100th percentile is significantly higher for the loans which defaulted than the one which were paid off.

Hight Interest Rate

When we look at the interest rate once again we see the same trend where the rates of the loans which defaulted are significantly higher than the ones which were fully paid. There is a difference of approximately 2% to 3% between the ones which defaulted and the ones which were fully paid. It seems like higher Interest rates also contribute in a way to the loans being Charged Off.





Income and Loan Amount

As seen from the Heat Map for almost all the loans that were disbursed there is a correlation with income. Better income results in more chances of the loan being fully paid. In all types of loan from Very Small to Very Large the ones that were paid have on an average higher income than the ones which were Charged Off.

Key Takeaways

01

Risky Loans

Small Business Loans are the riskiest loans and carry higher chances of default.

02

Annual Income

The higher the Annual Income the better the chances of Loan Payment. Low income relates with higher chances of Default

03

Interest Rates

Higher Interest rates also contribute to loan default. Those which were marked Charged off had significantly higher Interest Rates than the ones which were fully paid.

04

Popular Loans

Loans for amounts ranging from $5000 to $10000 are mostly in demand. Also the loans with a term of 3 years are higher than the ones with 5 year terms.

05

Home Ownership

People Renting or paying Mortgage are the ones most likely to take loans than the ones who own a house.

Thank
You