Lending Club Case Study

Author: Kumar Krishna Chakraborty

Business Problem Statement

A consumer finance company is specialised in lending various types of loans to urban customers. When the company receives a loan application, the company must 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 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

Business Objective

The company wants to understand the driving factors (or driver variables) behind loan default, i.e., the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment

Solution Approach

Exploratory Data Analysis was used to identify the key consumer and loan attributes that influence the tendency of loan default

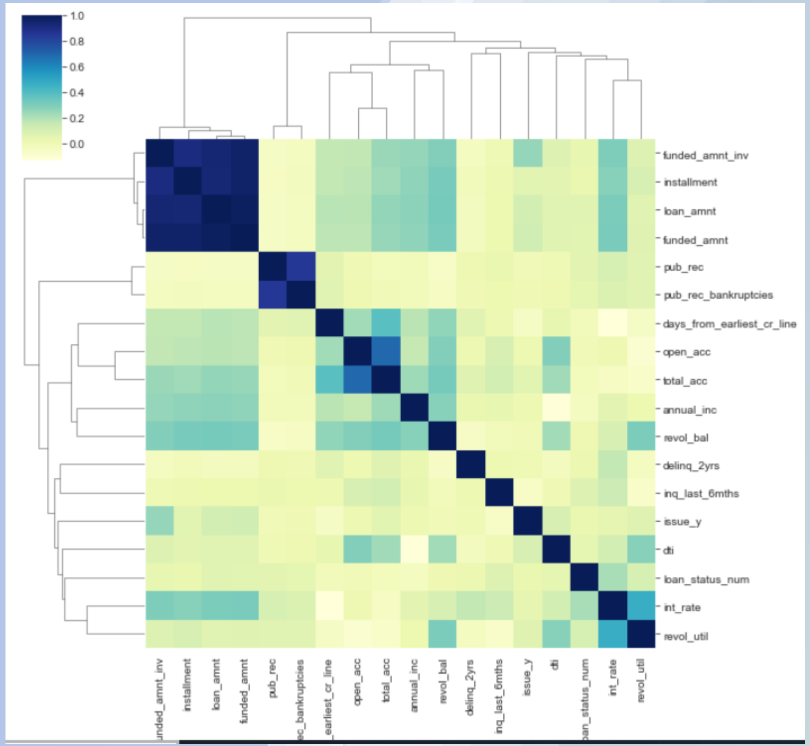
Univariate Analysis (Single Attribute impact) Key Observations

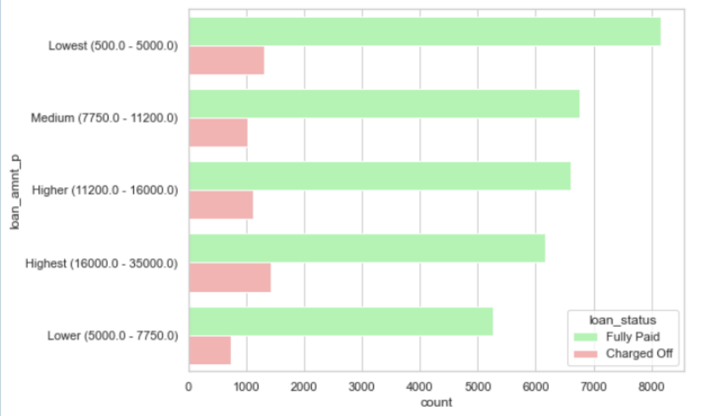
- We observed the trend that higher the loan amount, the chances of loan default is more
- We observed that higher is the interest, greater is the chance of loan default.
- We observed that higher instalment amounts show higher default percentages.
- We observed that higher annual income show lesser default percentages
- We observed that higher debt to income ratio show higher default percentages
- We observed that total credit revolving balances slightly influence the default percentage. Higher the revolving balance, greater the chance of the loan getting defaulted
- We observed that revolving line utilization rate has high impact to the loan default percentage. When this increases, the charged off percentage also increases.

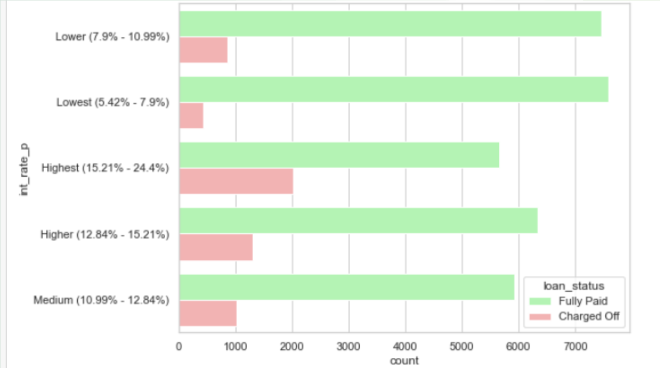
Univariate Analysis (Single Attribute impact) Key Observations

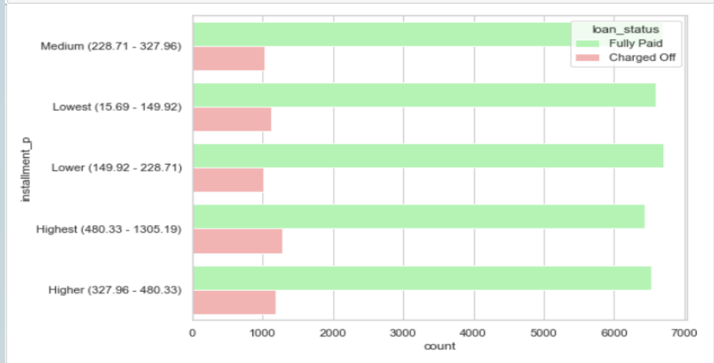
- We observed that loan repayment term plays a factor in predicting the default rate
- We observe that for G, F, E and D form grades, the default rate is much higher than others
- We observe that the G3 and F5 sub grades have default rate greater than 40%. This field is a clear indicator of the default percent
- We observe that Verified applicants have higher loan default percentage. Hence we infer that there is gap in the verification process
- We observed that loans taken for small_business, renewable_energy and educational have higher risks associated with it in terms of loan default
- We observed that NE, NV, SD, AK, FL, MO states show higher risks of loan default
- We observed that applicants with non-zero derogatory public records have greater chances of loan default
- We observed that higher is the number of public bankruptcy records, greater is the chance of loan default.
- We observed that the December month has the highest number of loan applications per year and the biggest default ratio. This is festive season. Month of May is also another one, which is during the summer break and right before the Memorial day and Independence day breaks in US

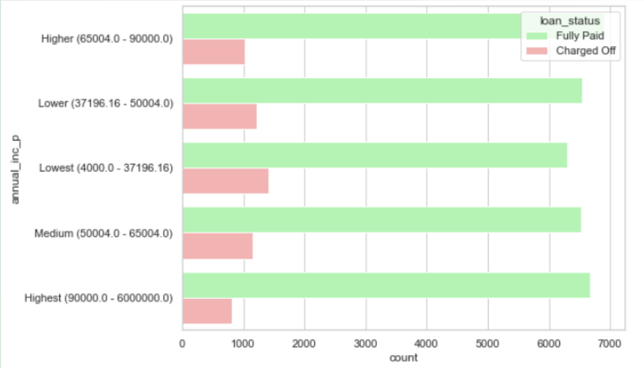
Univariate Analysis: Reference graphs

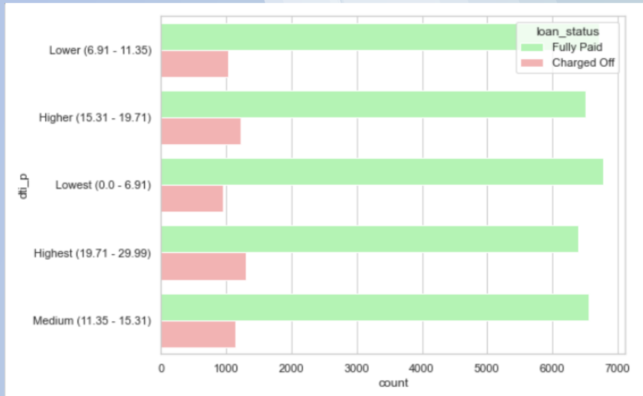


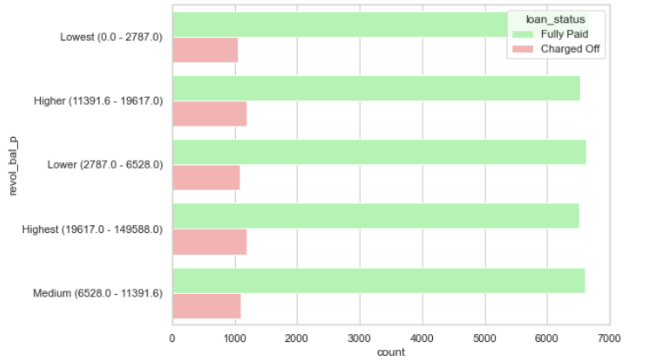


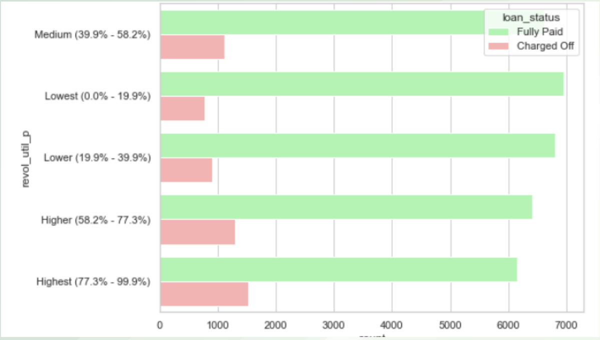








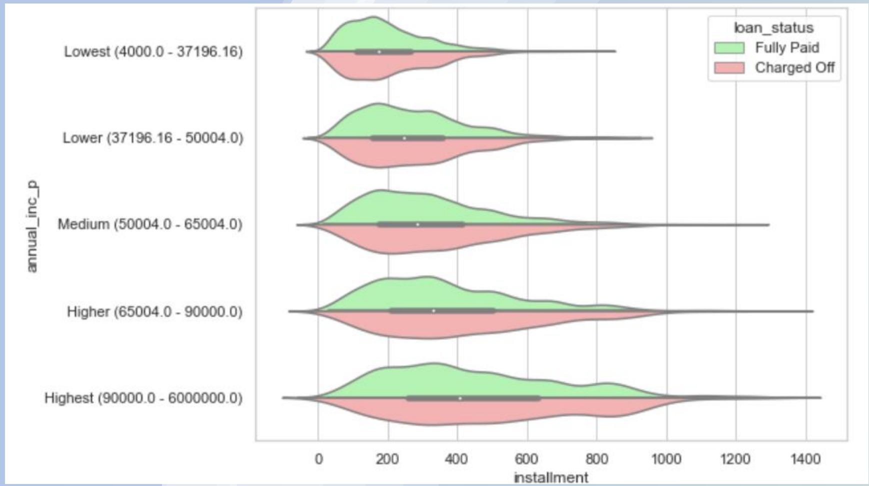


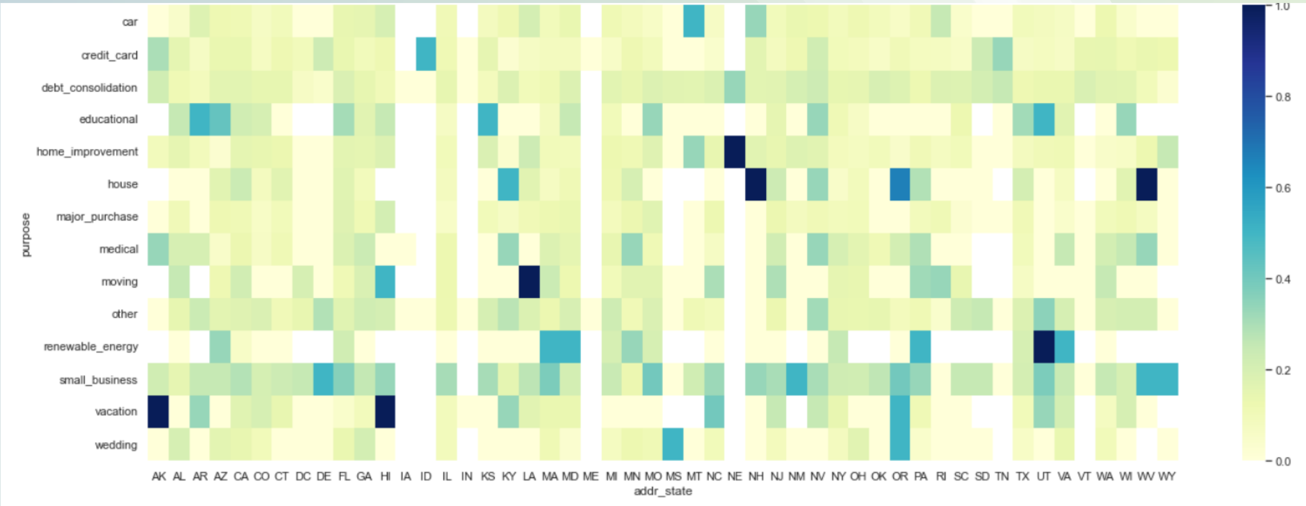


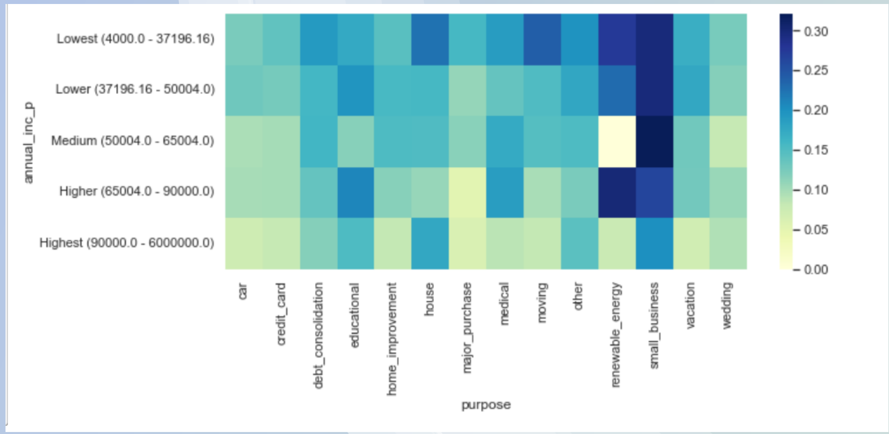
Bivariate Analysis (Combined attribute impact) Key Observations

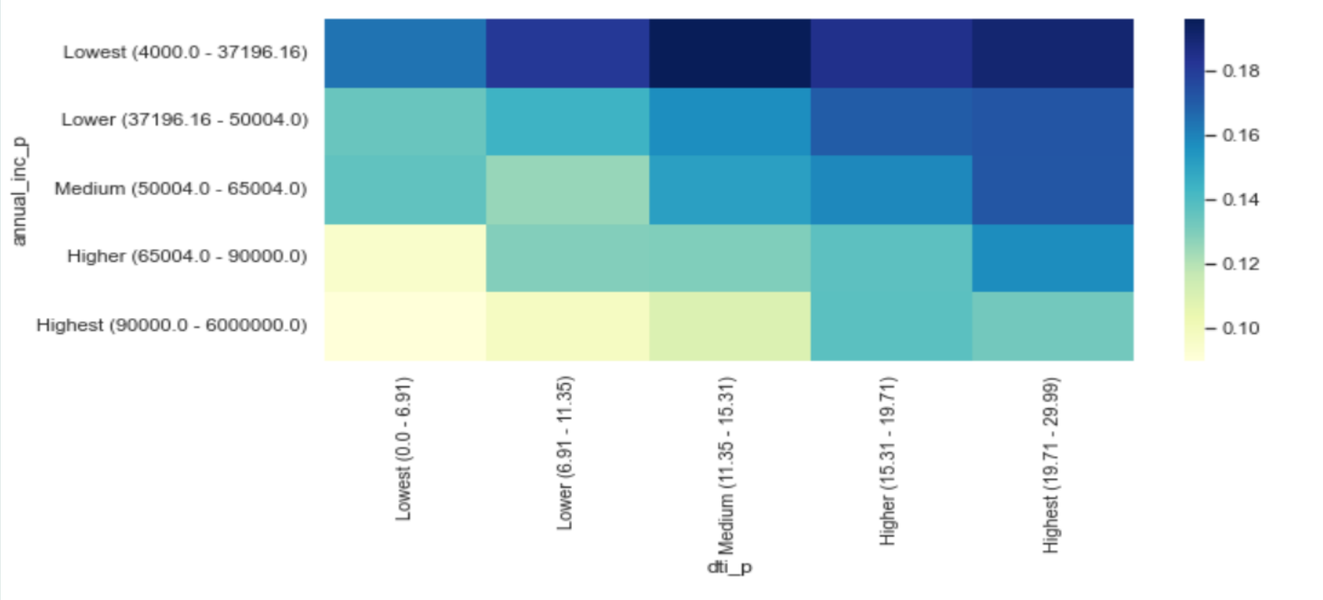
- We observed that loan repayment term plays a factor in predicting the default rate
- We observed that higher instalments for any income group have more number of loan defaults
- We observe that greater is the correlation of addr_state with the purpose of the loan, the greater is the chance of loan default
- By comparing the annual income with loan purpose we observed that small business loans for lowest and medium income groups have high chances of loan default. Also renewable energy loans for higher income group show the same trend.
- When we compare annual income with debt-to-income group, we observe that medium debt-to-income group in the lowest income range shows high trend of loan default

Bivariate Analysis: Reference graphs









Final conclusion

We have categorized our results under 2 categories

- Attributes showing lesser impact to the trend of loan default
- Attributes showing higher impact to the trend of loan default

Lesser impact

- Lower annual income (below 37000)
- Higher loan amount (above 16000)
- Higher debt to income ratio (above 15%)
- Higher instalment amount (above 327)
- Loan issue month (Dec, May)
- Applicant’s address state (NV, SD, AK, FL)

Higher impact

- Higher interest rate (above 13%)
- Repayment term (5 years)
- Loan purpose (small business, renewable energy, educational)
- Higher revolving line utilization rate (above 58%)
- Loan grade & sub-grade (D to G)
- Public bankruptcy records (1 or 2)
- Derogatory public records (1 or 2)
- Missing employment record

Final conclusion

In terms of the effects of combined attributes on the tendency of loan default, we observe that the following show higher combined impact

- High loan amount and interest rate for lower income group
- High instalment and longer repayment term
- Residential state and loan purpose
- Income group and loan purpose
- Home ownership (other) and loan purpose (car, moving or small business)

Few Recommendations

- The company should be very cautious in approving loans with high amount and interest rates to lower income group applicants
- The company can structure the loans in a manner to avoid high instalment and longer repayment terms.
- The company should be careful in approving small business loans to low and medium income groups. The same aspect applies to renewable energy loans for high income group also. The verification process should be more stringent in these cases.
- The company should not approve loans for applicants for whom there are no proper employment records.
- The company should try to avoid approving loans to applicants having Derogatory public records and Public bankruptcy records (1 or 2)
- The company should revisit applicant verification process and rectify it to make it more accurate. We had observed that verified applications have more tendency to loan default.