Lending Club Case Study

Facilitator : Sindhu Soundararajan

Case Study:

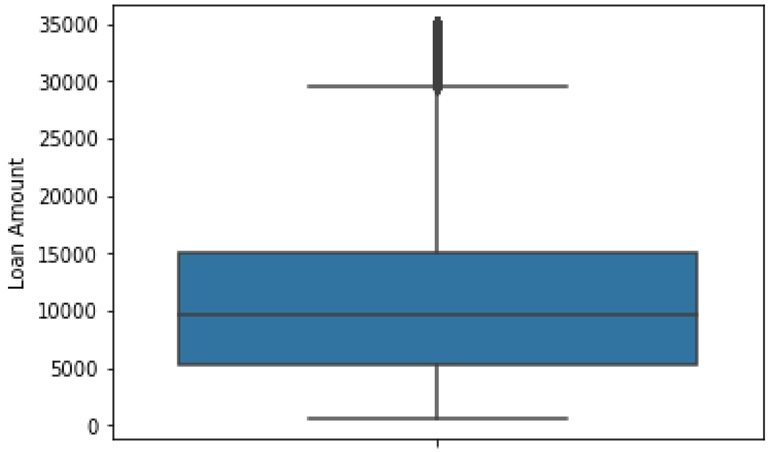
- Identify risky loan applicants to reduce credit loss.
- Use EDA understand the driving factors that are strong indicators of loan default
- This can be used for portfolio and risk assessment.

Data Cleaning

- Drop columns having all missing or null values.
- Drop column not adding any importance in analysis
- Deleting column having all same values so nothing can be inferred from that.
- Format the columns to remove unwanted elements like ‘%’, ‘-’ etc.
- Format the columns to the correct datatype.

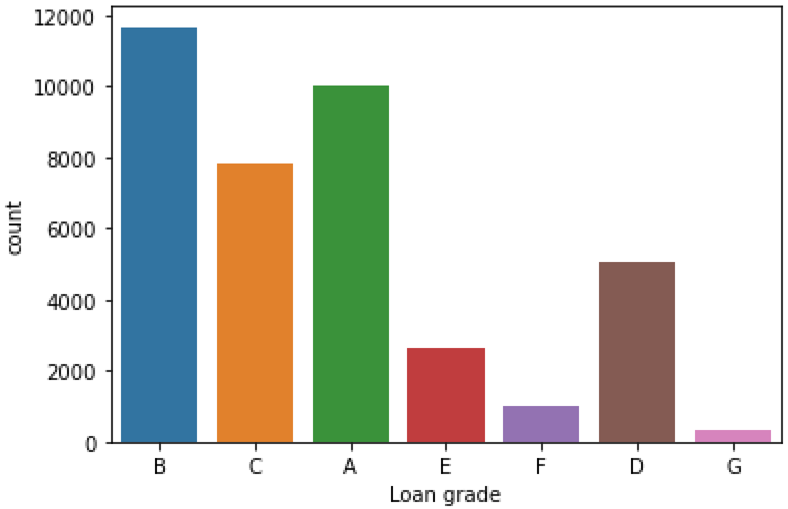
Analysis and Inferences.

Loan amount



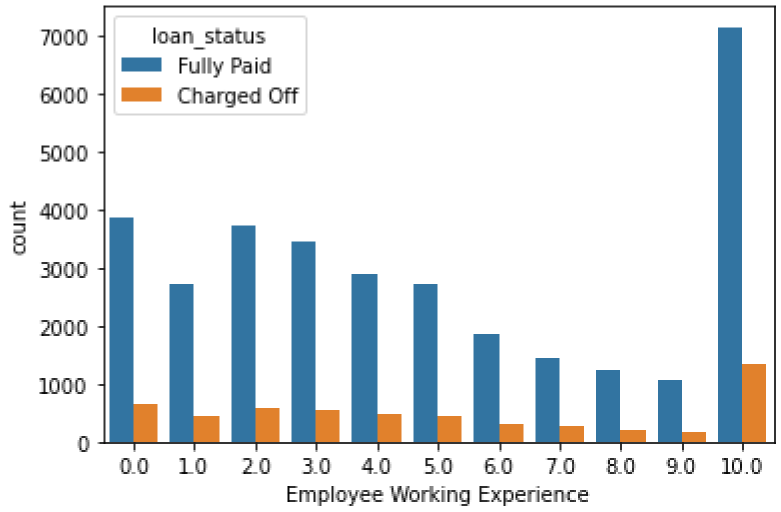
- ❖ Maximum no of loans are disbursed with loan amount between 5K and 15K.
- ❖ Maximum loan amount range from 500 to 30000
- ❖ There are outliers between 30K to 35K

Loan Grade



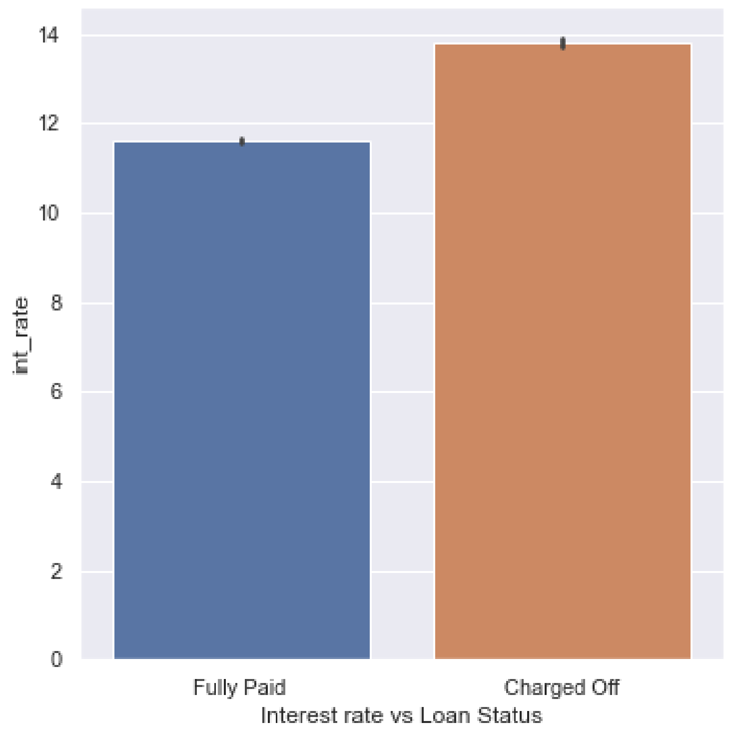
If loan grade is from highest A to lowest B, then from the above plot maximum loans availed are high grade loans.

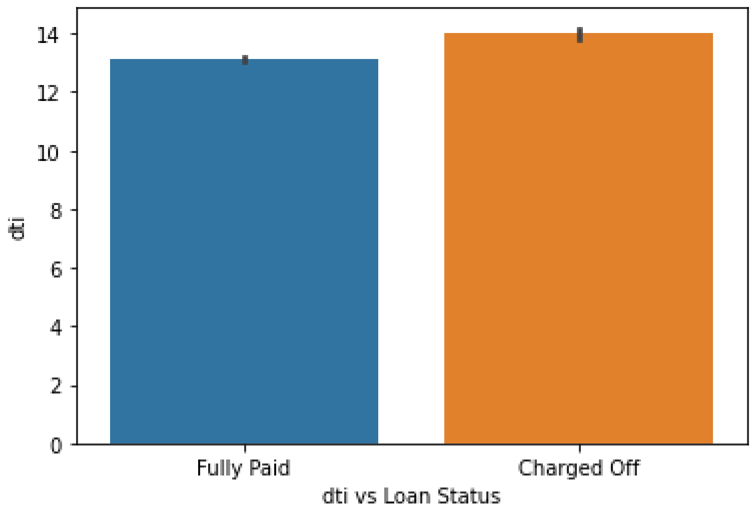
Employee Working Experience



- Maximum loans are availed by people having working experience GT or EQ 10 years.
- Maximum charged off and fully paid loans belong to this category of people.

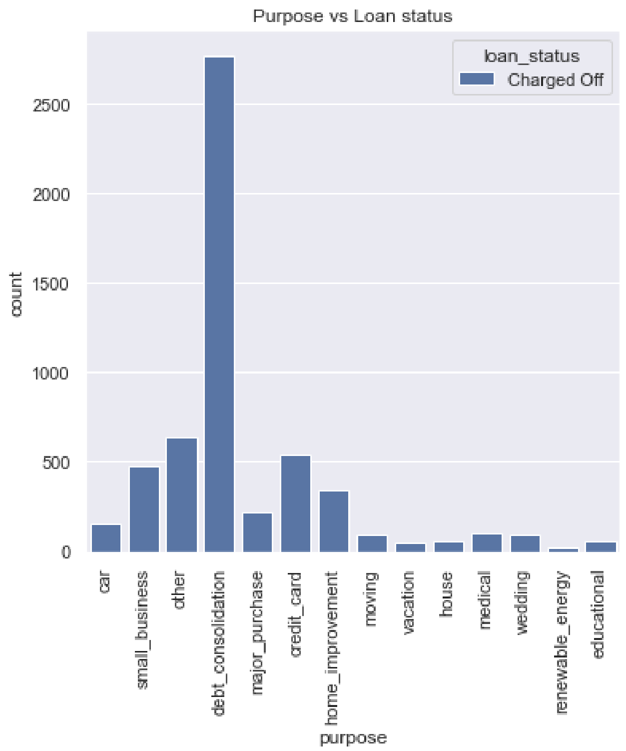
Interest rate vs Loan status

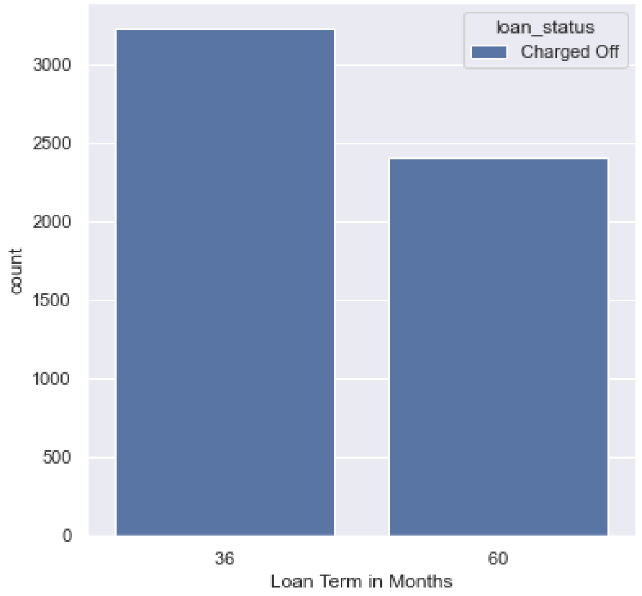


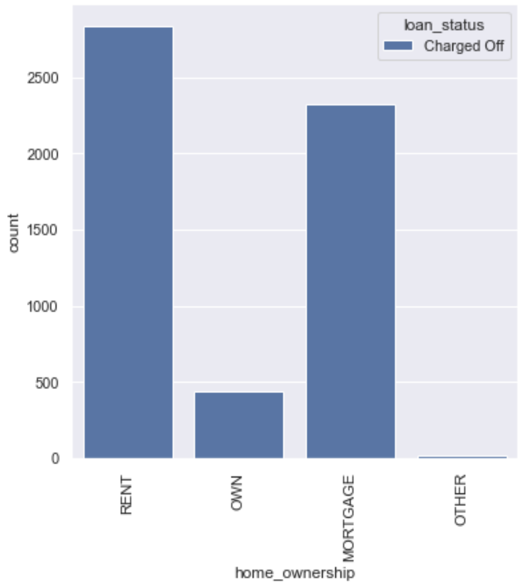


- Charged off loans have the high interest rate
- Debt to income ratio impact charged off.

Charged Off loans







- Loans with Debt consolidation is the most charged off loans.
- Charged off loans are most in 36 term policy
- Charged off loans takes either live in rent or mortgaged places.

Conclusions

- Charged off loans are most in 36 term policy
- High Debt to income ratio lead to charge off.
- Loans with high interest rate are likely to default
- 43% of loans are not verified.
- Loans taken with Debt consolidation purpose is the most charged off loans.