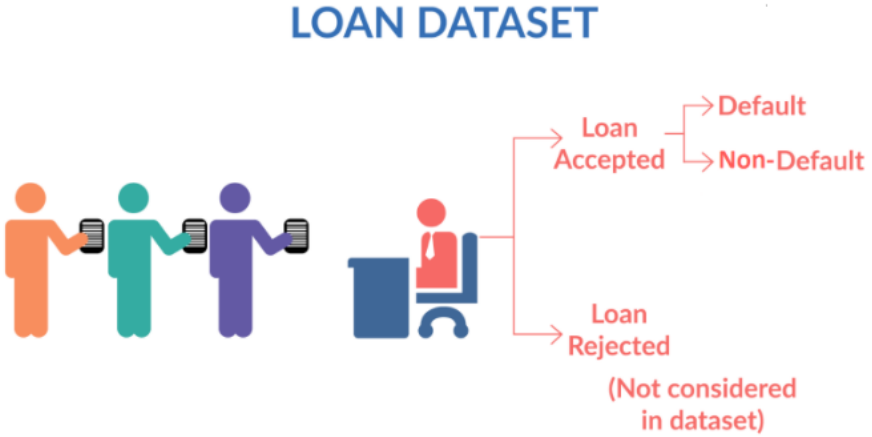
Lending Club Case Study

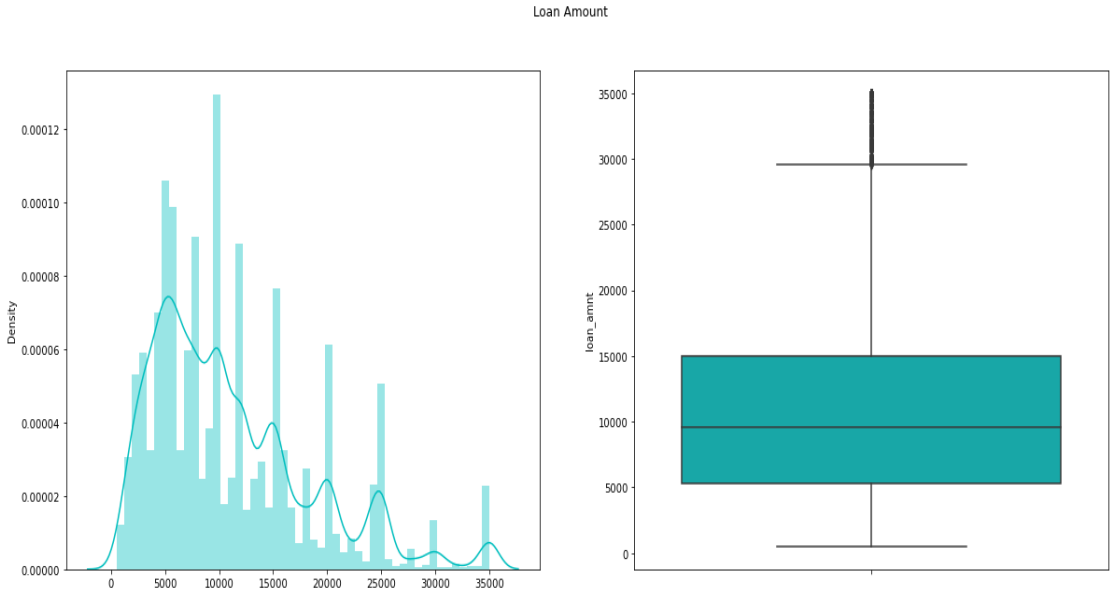
Problem Statement

A Consumer Finance company is looking to identify Driving Factors that are strong determinants of loan defaults. This analysis will help organization to assess risk and portfolio health. As per following flow, company's main objective is to increase number of Non-Default under Loan Accepted.



Data Cleaning Exercise

- ✓ There were a total of 39,717 records, across multiple years, and 111 columns representing various features of an account.
- ✓ Most of the columns – 68 – were Null and were dropped to stay focused on 43
- ✓ There were records for currently active customers, and were removed
- ✓ Some data cleaning, such as removing % from rate column, and data addition, such as creating Year & Month of Loan, was also performed



- ✓ Post Data Cleaning, we worked with 38,577 records and 45 features
- ✓ Cleaned data had Mode = Median = 10000
- ✓ 50% Population has applied for loans between 5000 and 15000
- ✓ There are very few applicants beyond loan amount 30000

Recommendations

After through analysis, refer Appendix for in depth analysis, we have inferred that there will be high probability of “Loan Default” in case loan is sanctioned as per following:

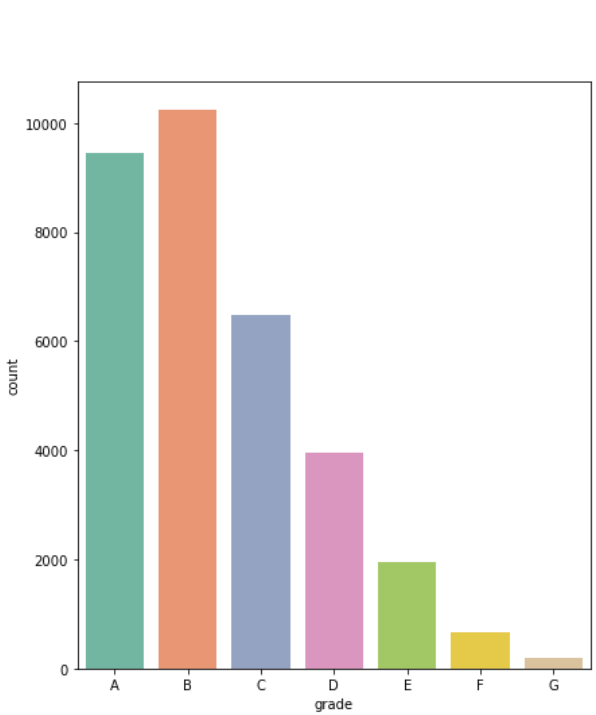
- ✓ People in Grade G
- ✓ Purpose of loan is Debt Consolidation, small businesses, or Others
- ✓ Loan distribution month is between May to October
- ✓ Loan distribution address state is CA
- ✓ Debtors are renting than own a property
- ✓ loan purpose is house and interest rate is between 14% and 16%
- ✓ Address state is AK and salary Annual Income between 60k to 175k
- ✓ People with verified status, having annual income between 68k & 72K, and loan amount > 16000

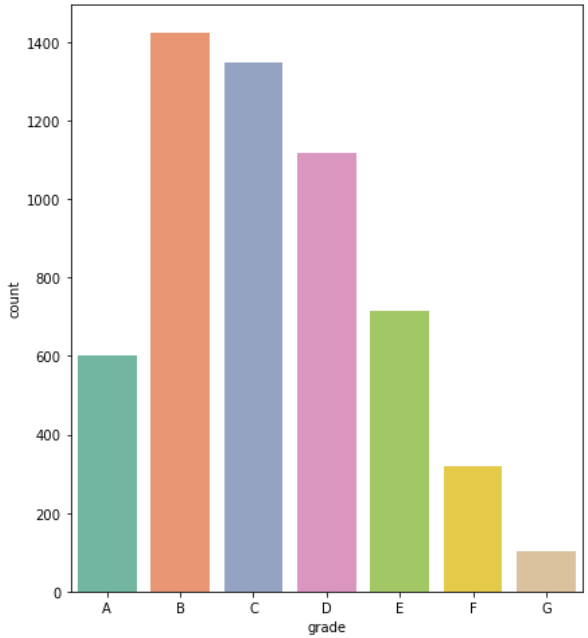
Business should diligently consider these features before proceeding further with a loan application.

Appendix

Data Analysis

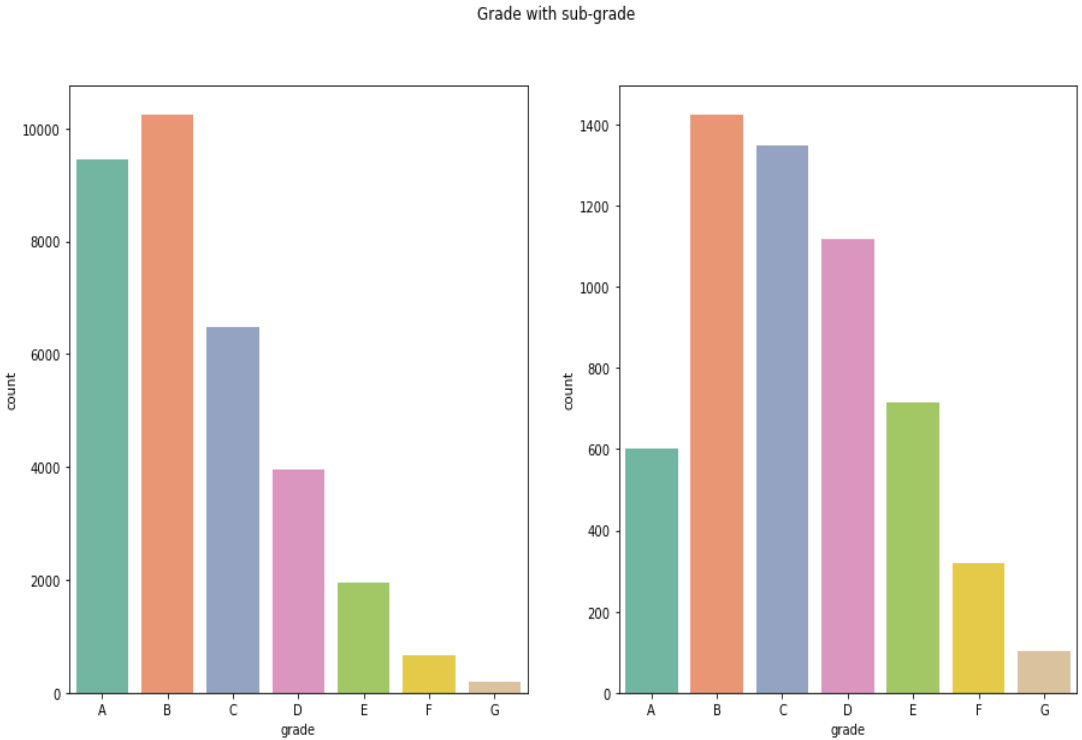
- ✓ Paid to Default ratio is decreasing from Grade A to Grade G
- ✓ Grade A Paid to Default ratio is appx 15:1
- ✓ Grade B Paid to Default ratio is appx 7:1
- ✓ Grade G its appx. 1:2





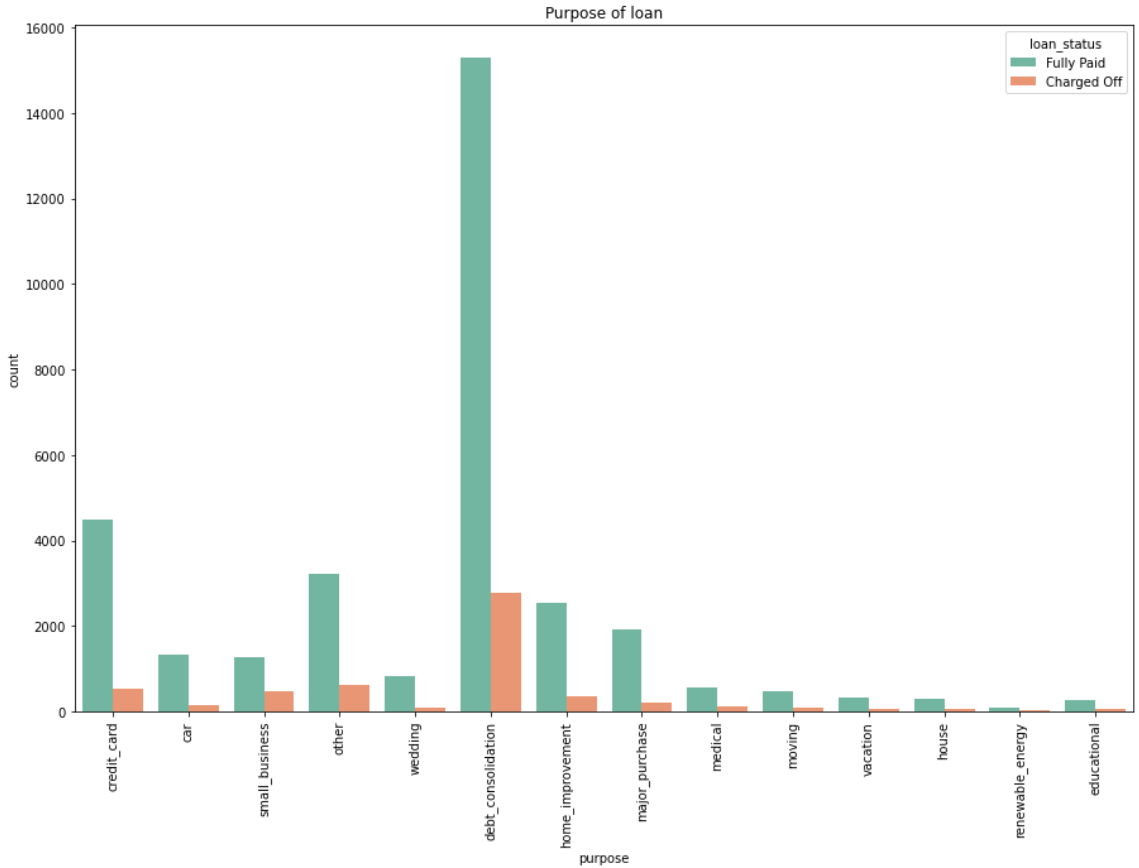
Data Analysis

- ✓ Paid to Default ratio is decreasing across Sub Grades, within Grades
- ✓ Sub Grade A1 is outlier, Paid to Default ratio is 44:1 compared to overall group that hovers around 15:1
- ✓ For all other groups, all sub-Grade's ratio hover around their respective overall grade ratio



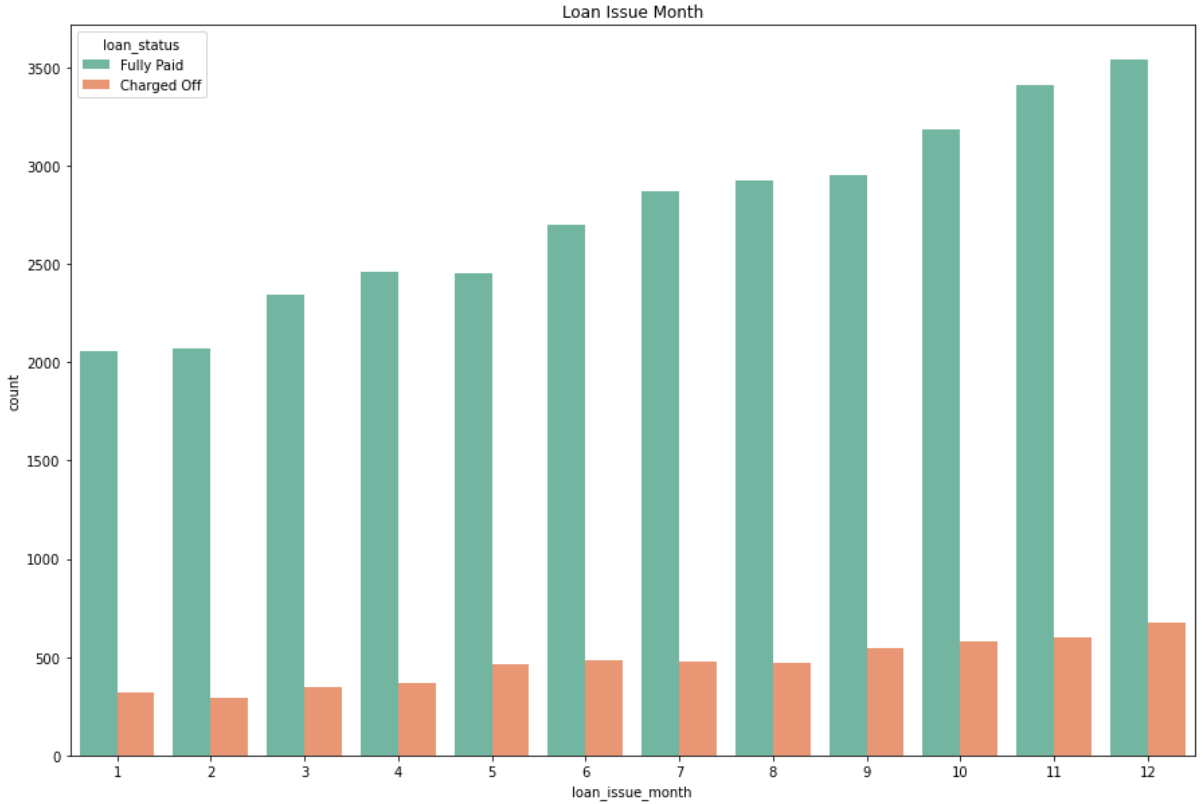
Data Analysis

- ✓ Paid to Default ration is lower for purposes such as Debt Consolidation & small businesses, appx 5:1
- ✓ Paid to Default ratio is higher for purposes such as major Purchase & Cars



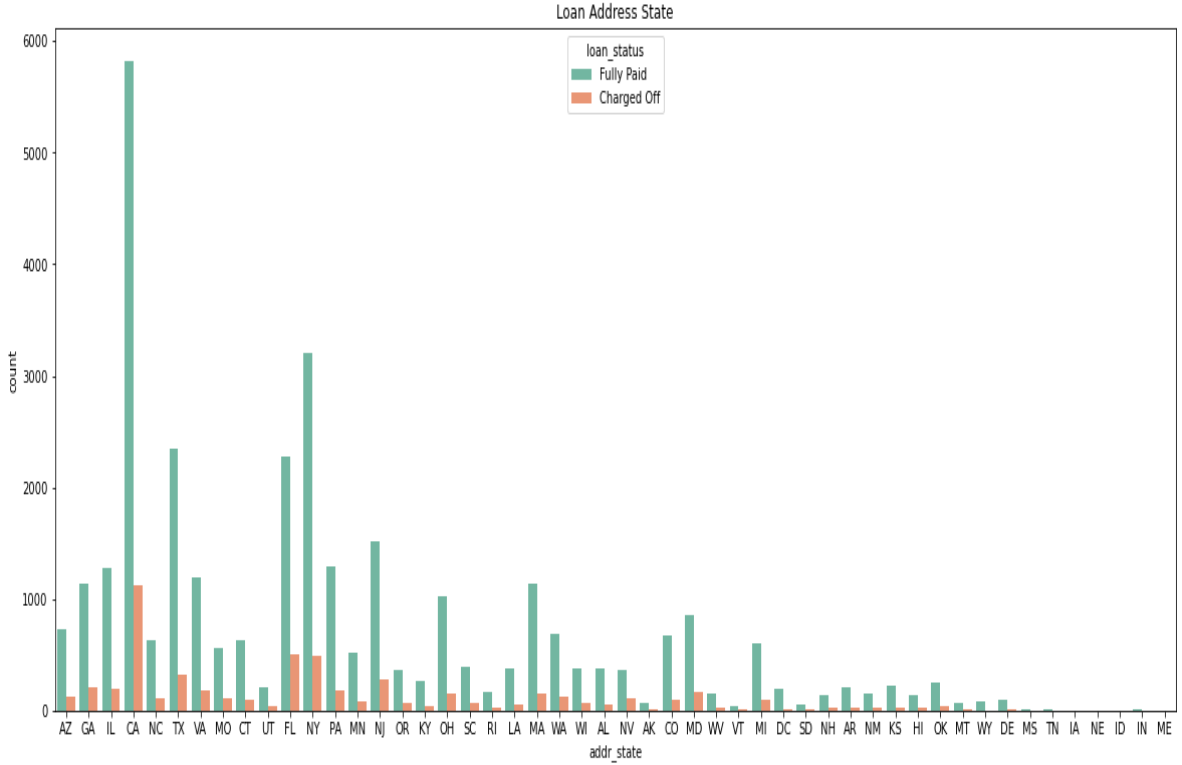
Data Analysis

- ✓ Ratio of Paid to Default is lower for purposes such as Debt Consolidation & small businesses, appx 5:1
- ✓ Ratio of Paid to Default is higher for purposes such as major Purchase & Cars



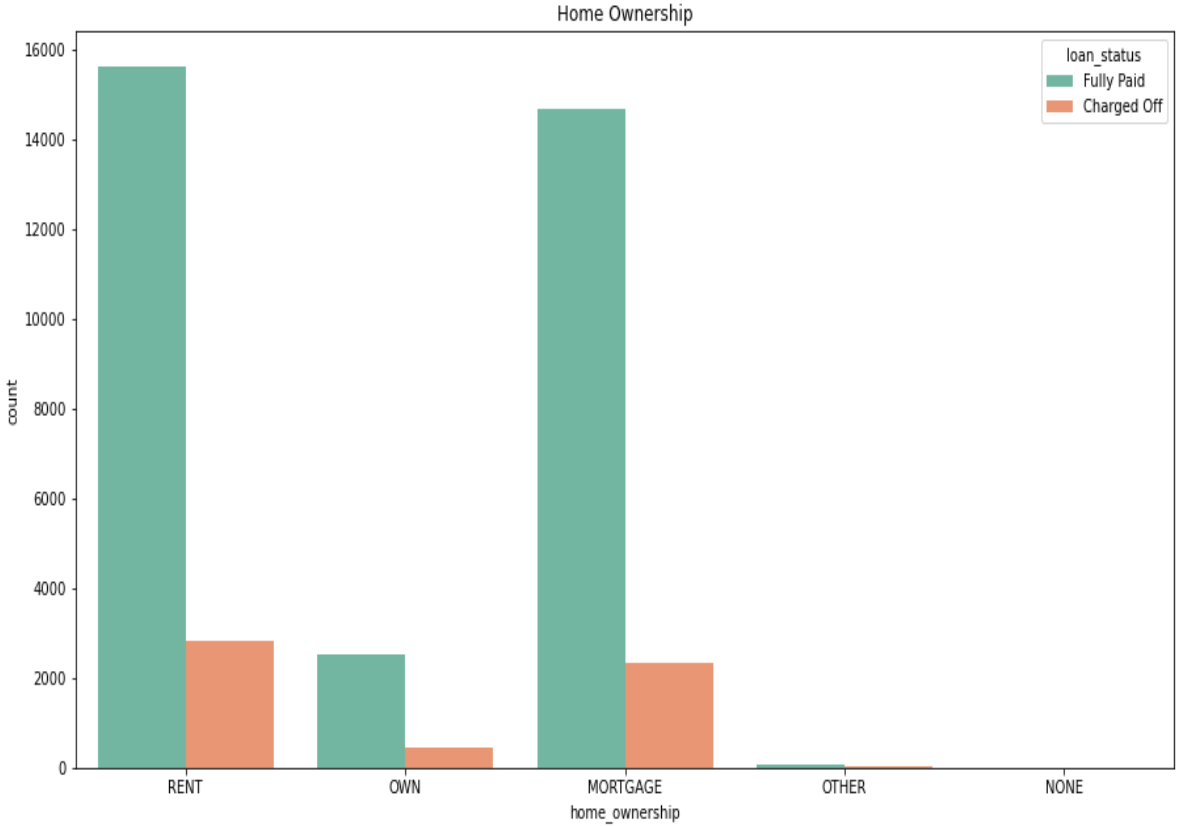
Data Analysis

- ✓ Ratio of Paid to Default is lower for purposes such as Debt Consolidation & small businesses, appx 5:1
- ✓ Ratio of Paid to Default is higher for purposes such as major Purchase & Cars



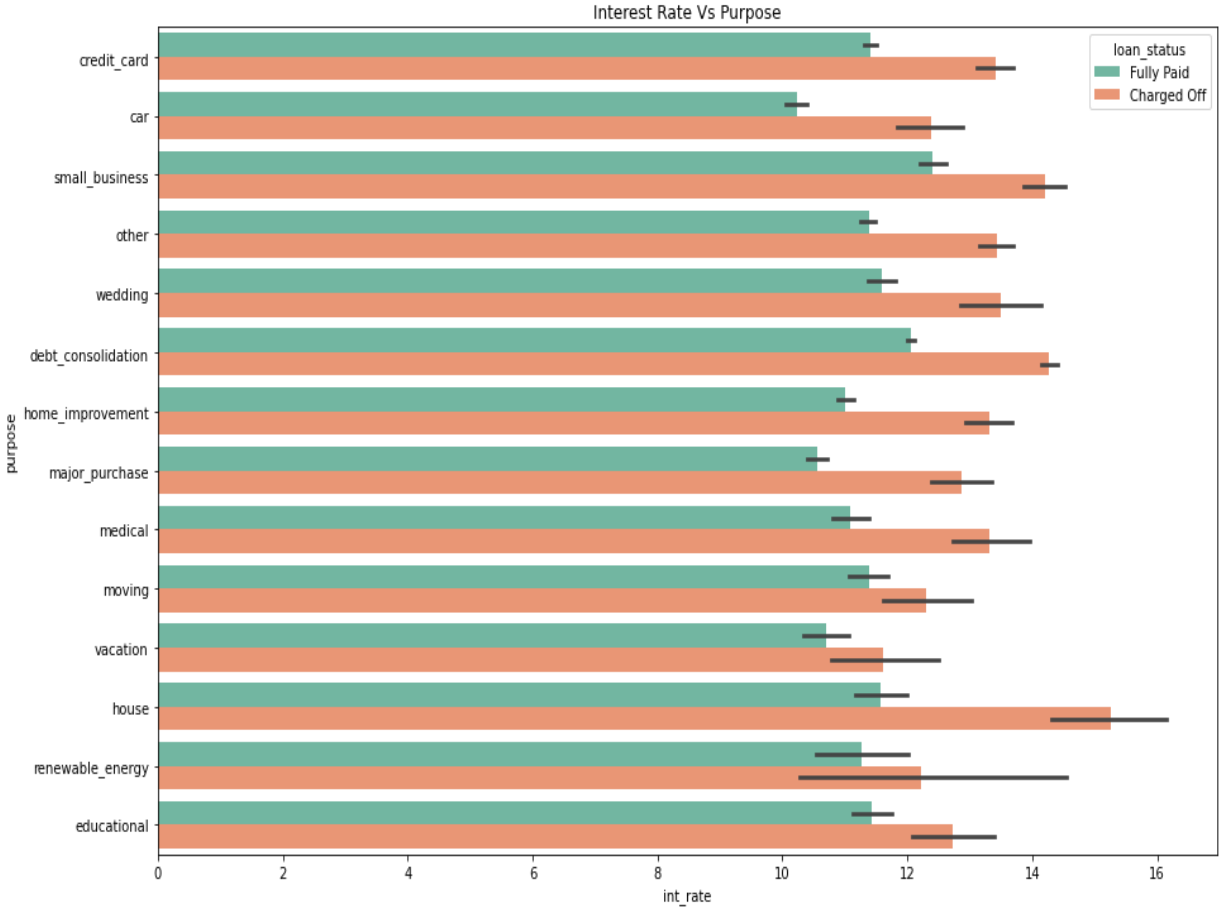
Data Analysis

- ✓ Maximum debtors Rent their houses, Paid to Default ratio for such debtors is appx. 5:1, which is almost same for Mortgage
- ✓ Paid to Default ratio for such homeowners is 21:1



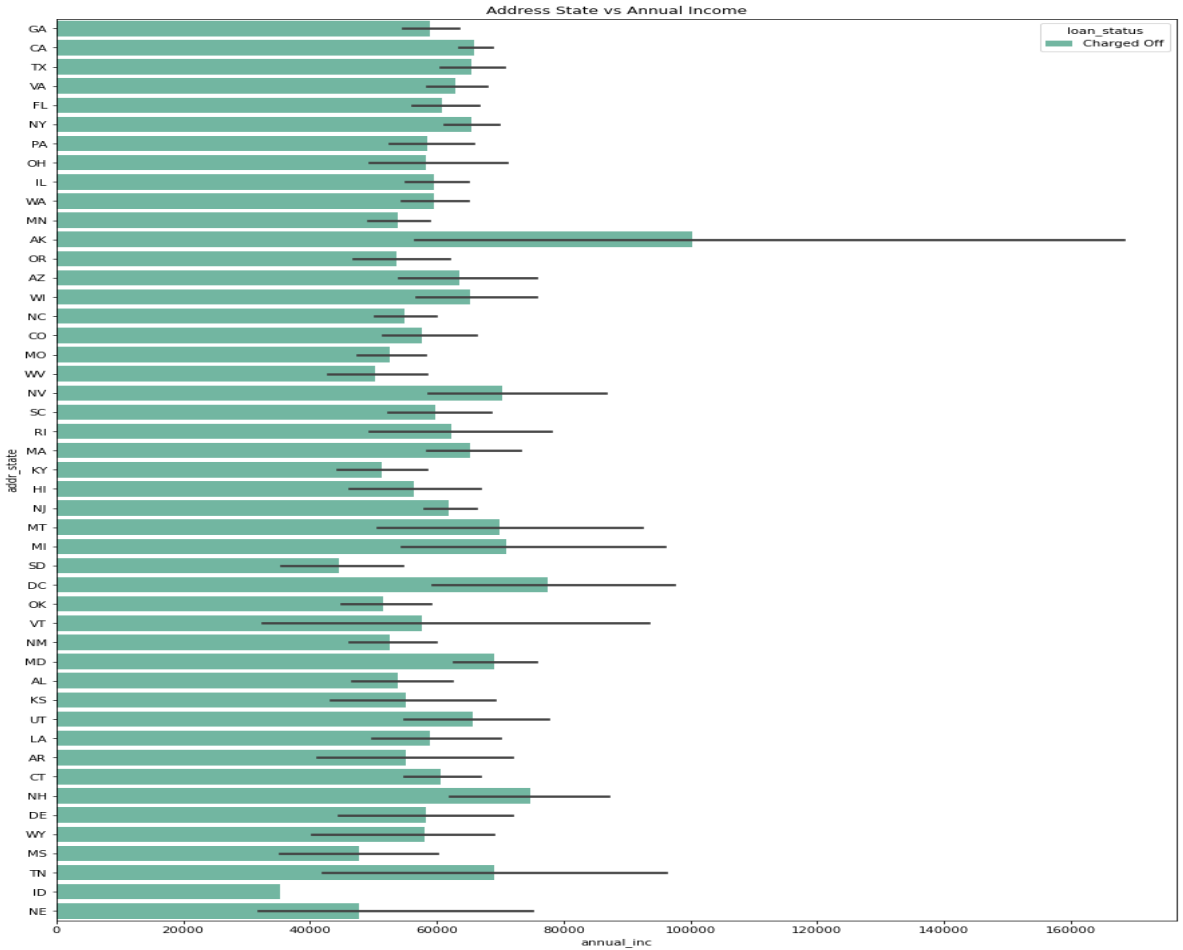
Data Analysis

- ✓ People with loan purpose “house” default the most with “interest rates” between 14% to 16%



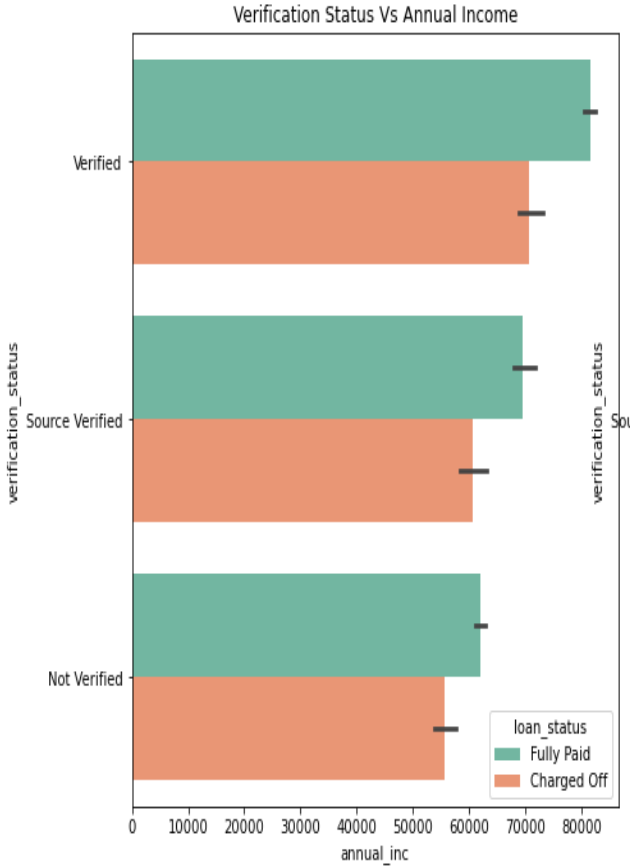
Data Analysis

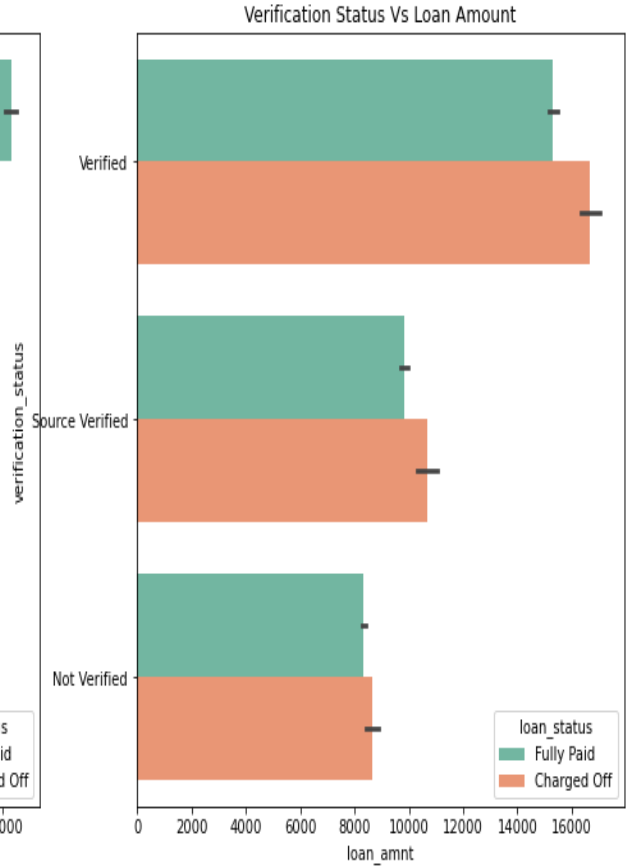
- ✓ People in AK and salary ranges between 60k and 160k, have tendency to default the most

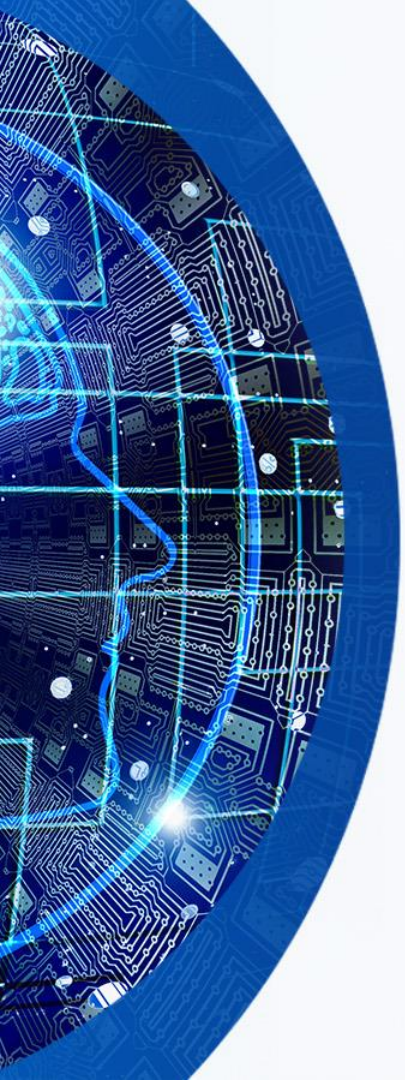


Data Analysis

- ✓ People with verified status and annual income between 68k and 72k, tend to default more
- ✓ People with verified status and loan amount > 16000, tend to default more







Thank You

Pawan Garg
Hemant Grover