# Lending Club Case Study (EDA)


- Suyash Kumar Singh
- Srinivasan Radhakrishnan

# Agenda

- Problem statement
- Analysis approach
- Observations
- Conclusions & Actionable Insights

# Problem statement

Lending club is a US based lending platform. Unlike banks, the loans are provided partly by the investors and partly by the lending platform.

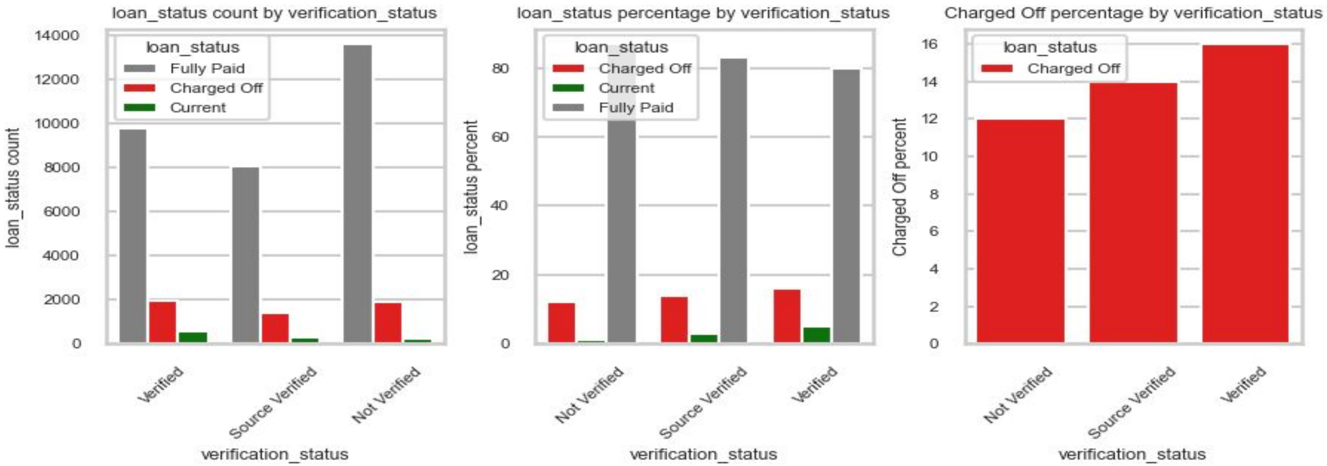
A borrower applies for a loan. The lending platform carefully vets the loan application based on the credit profile of the borrower. Even after the diligent process of vetting the loan application, a lot turn into defaults, a.k.a “Charge offs”.

The objective of this exploratory data analysis (EDA) on the Lending Club dataset is to identify various business reasons that could lead to a Charge-Off.

# Analysis Approach

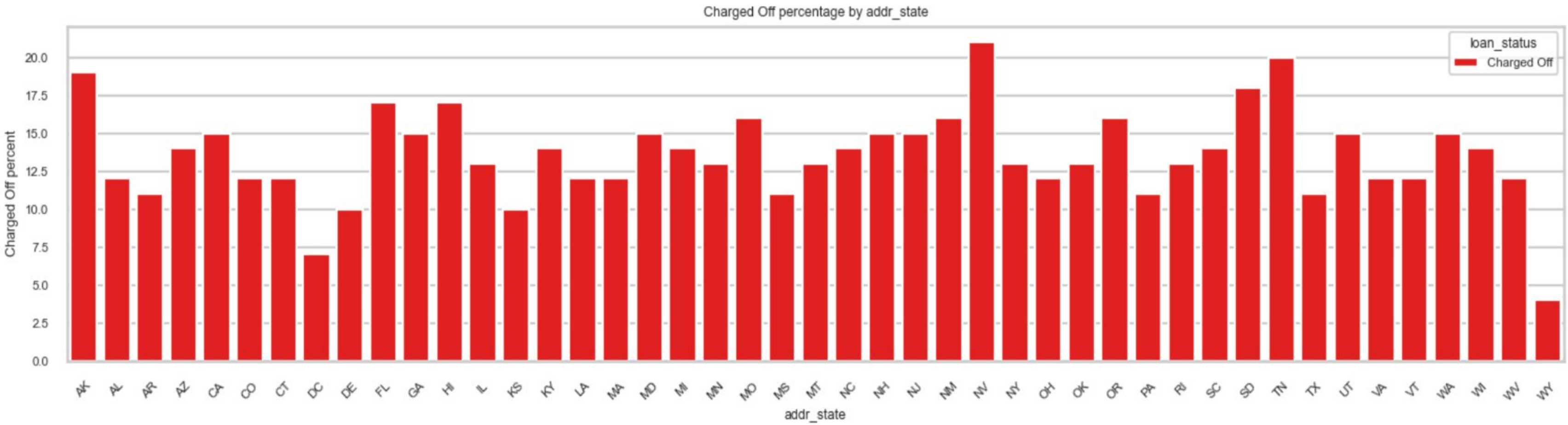
- Understand the lending domain relevant to our case study using tools like chat-gpt
- Clean and transform the dataset relevant to our case study
  - Removal of empty rows and columns
  - Removal of highly correlated columns and keeping one out of the set
  - Removal of single value columns
  - Removal of columns based on business needs
  - Expanding date columns, fixing string columns
- Analyse the data as per business needs (loan status as the central variable)
  - Segmented Univariate analysis for both categorical and binned numerical columns
  - Bivariate and Segmented Bivariate analysis for numerical columns
  - Multivariate analysis
- Record the observations and conclude with actionable insights.

# Observations - Verification status



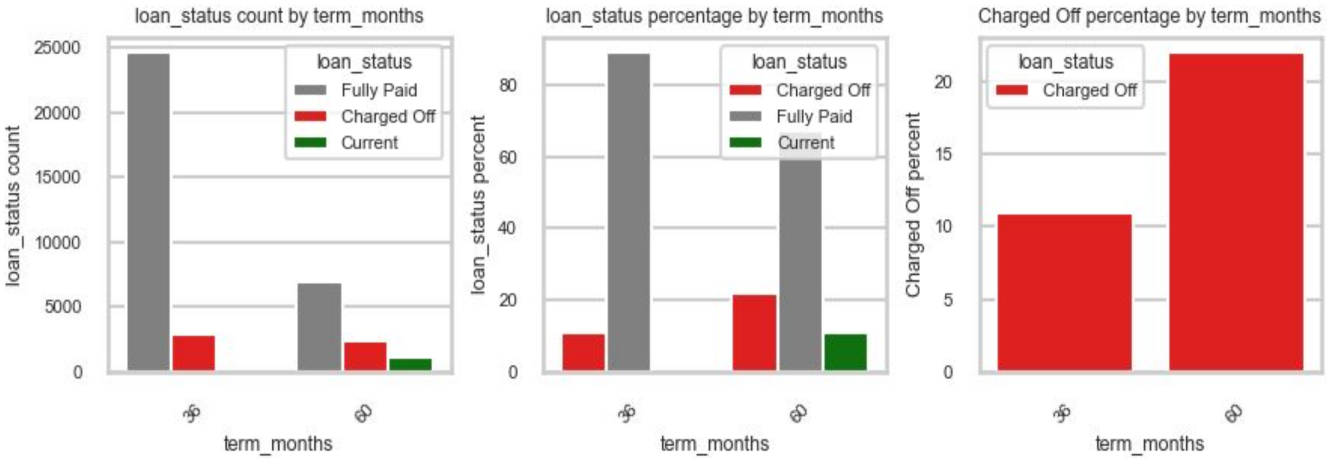

Surprisingly, the borrowers whose annual income was verified turned out to be the majority defaulters

# Observations - States



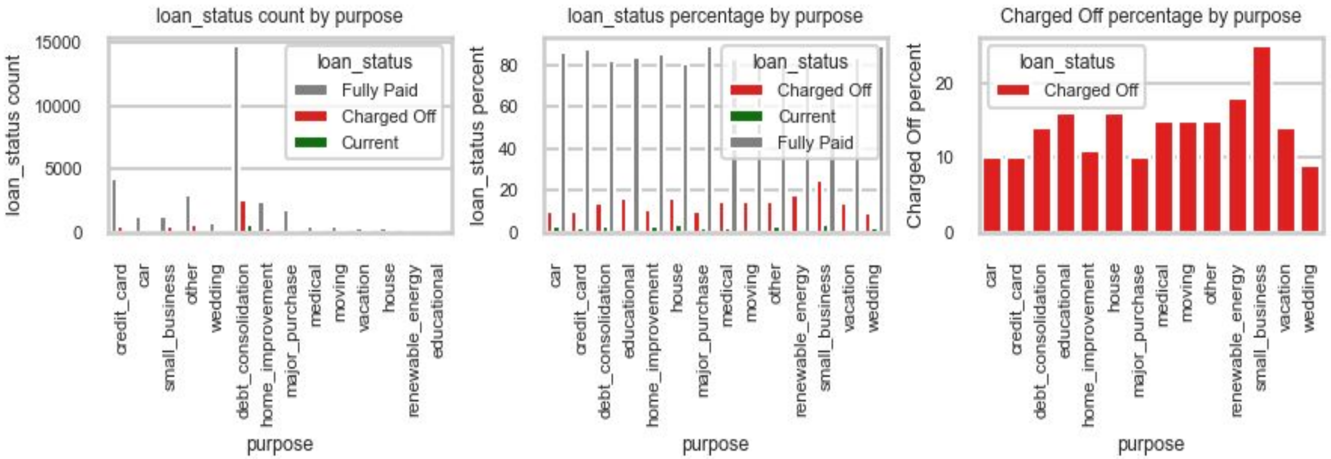

The state of Nevada has proportionally more defaulters

# Observations - Term duration



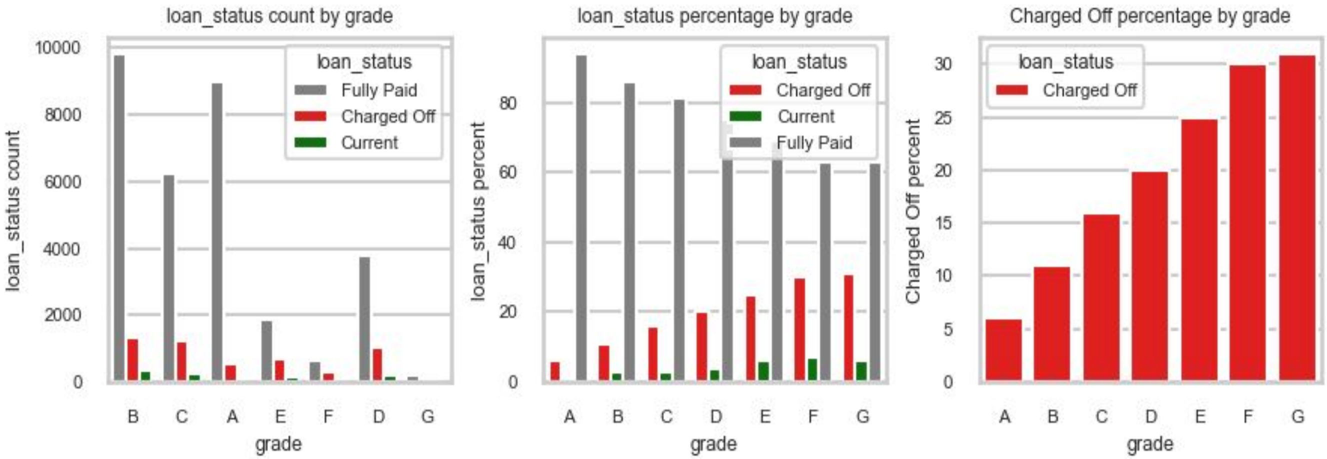

Longer term loans have higher chances of being defaulted

# Observations - Purpose of the loan



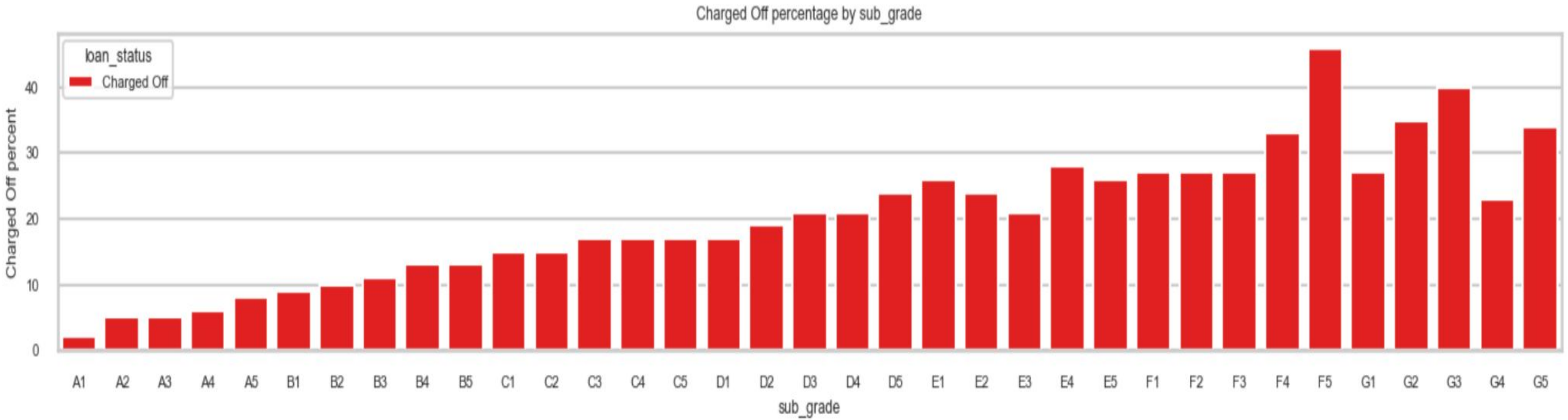

Loan towards small businesses have higher chances of being defaulted as compared to other purposes

# Observations - LC Grade



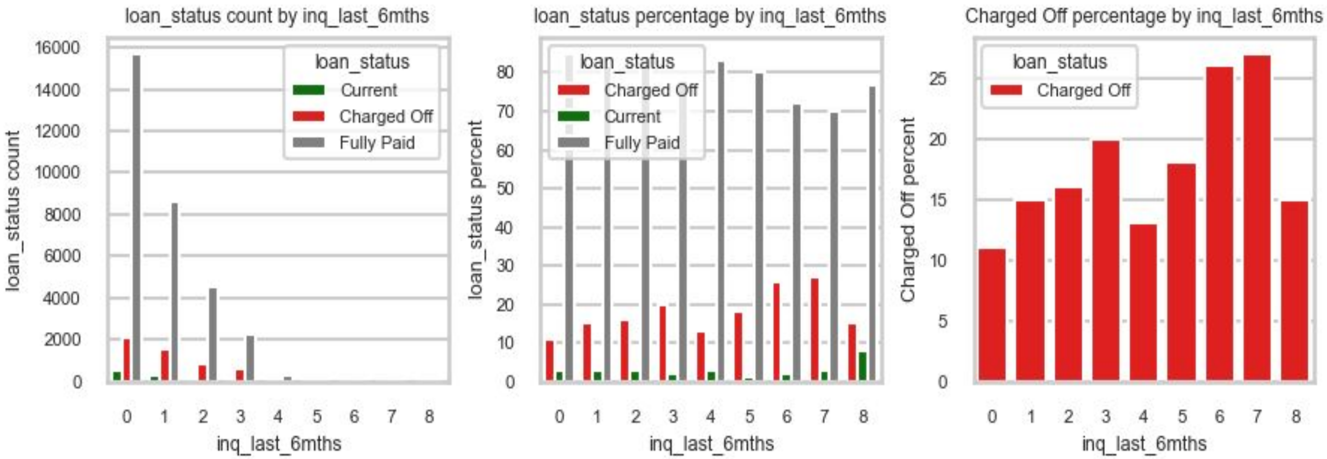

Lending Club was able to appropriately profile the loan listings according to the risk

# Observations - LC Sub-Grade



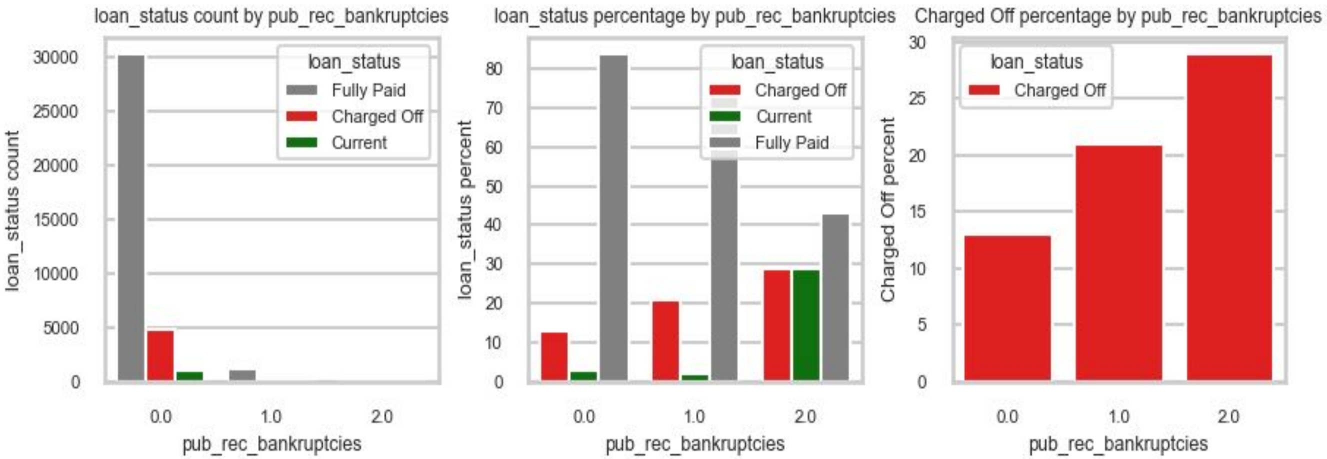

Sub-Grade assignment needs to be re-looked at for “G” sub-grade

# Observations - Inquiries in last 6 months



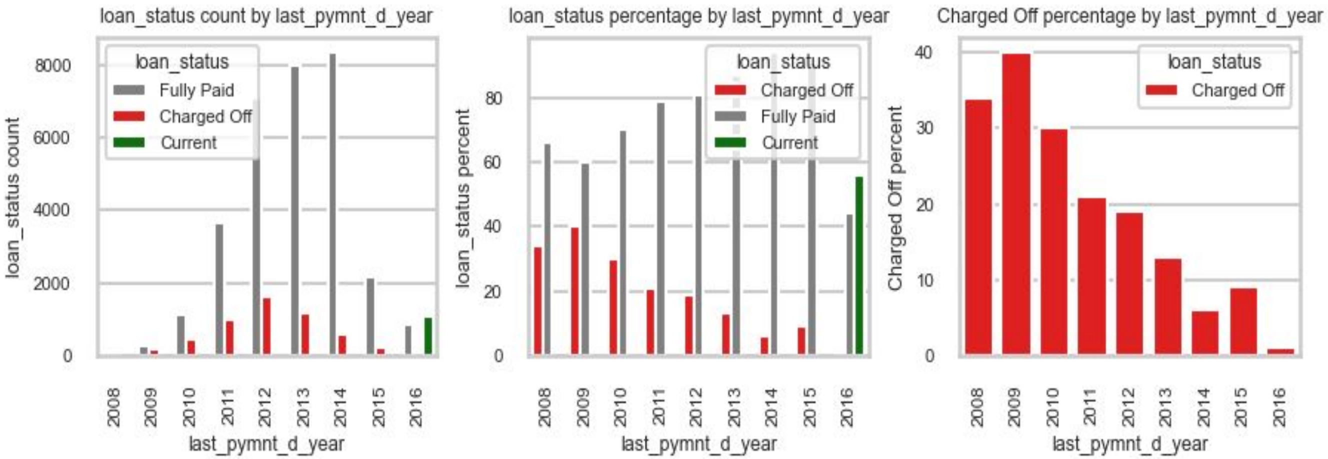

Proportionally, the more enquiries that happen on the credit profile in the last 6 months, tend to default more

# Observations - Publicly recorded bankruptcies



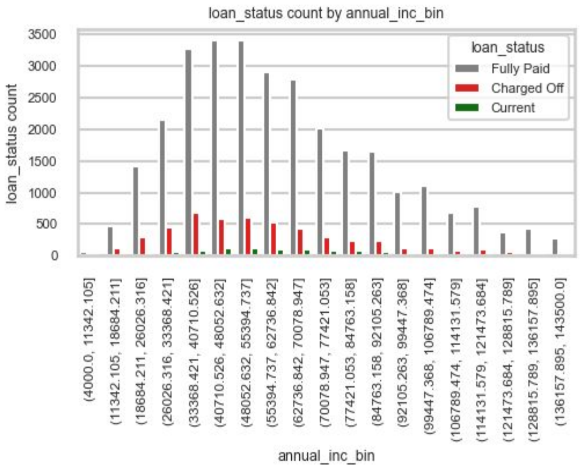

Not a lot of borrowers have publicly declared bankruptcies, but those who have, are more likely to default

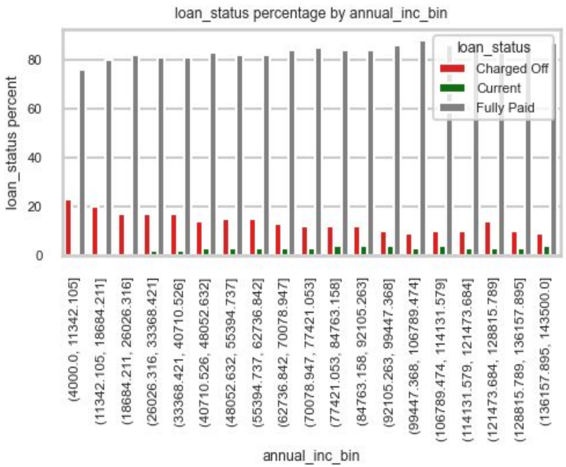
# Observations - Last payments during recession years

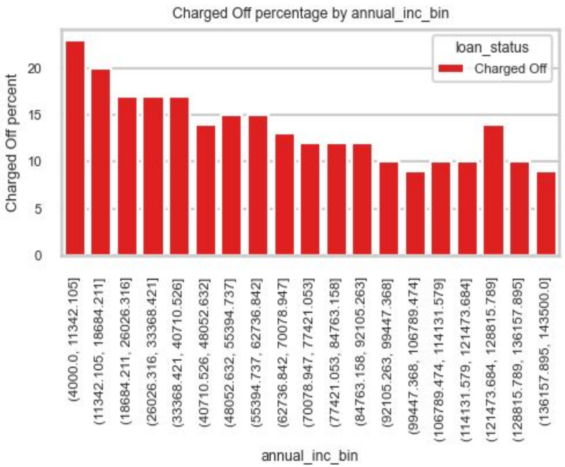



A lot of last payments happened in the year of 2008 and 2009 due to recession and were declared charged off since they couldn't pay after 2009

# Observations - Annual Income

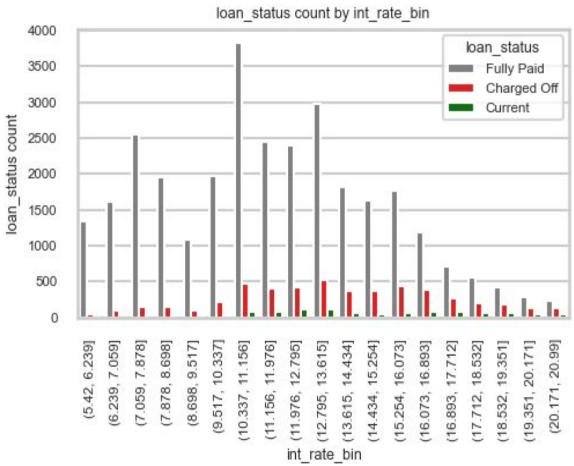


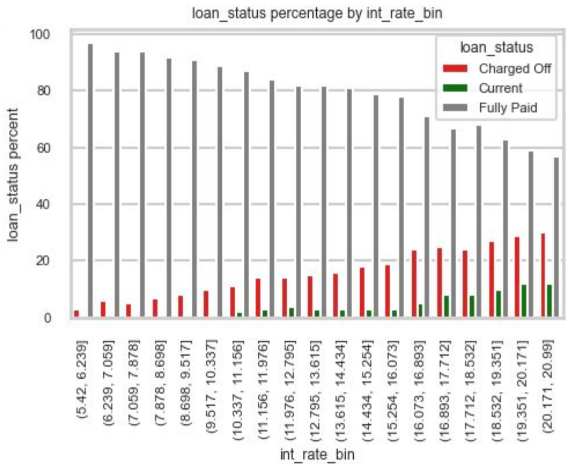



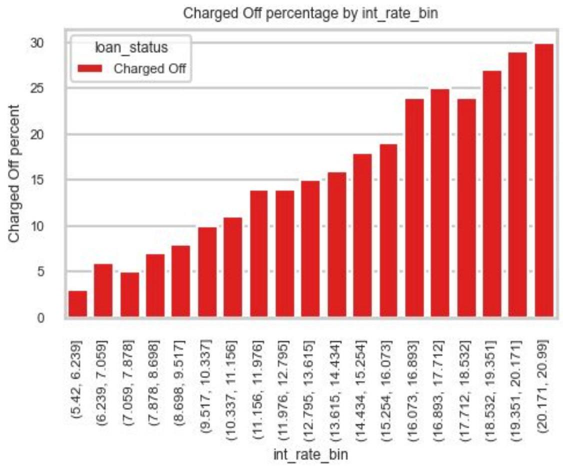




Borrowers having high annual income tend to default less

# Observations - Interest Rates

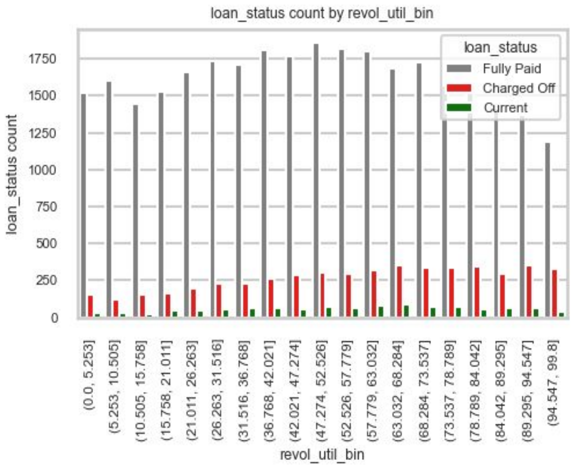


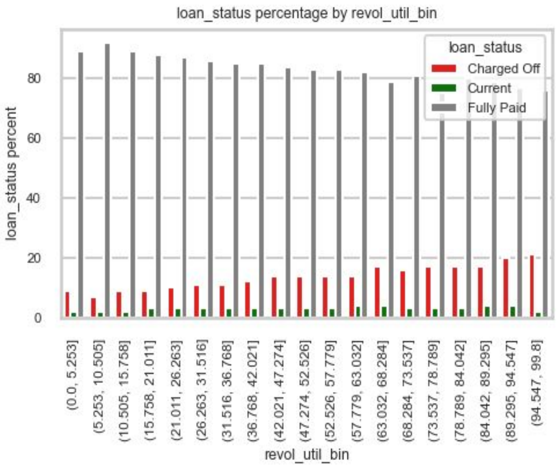



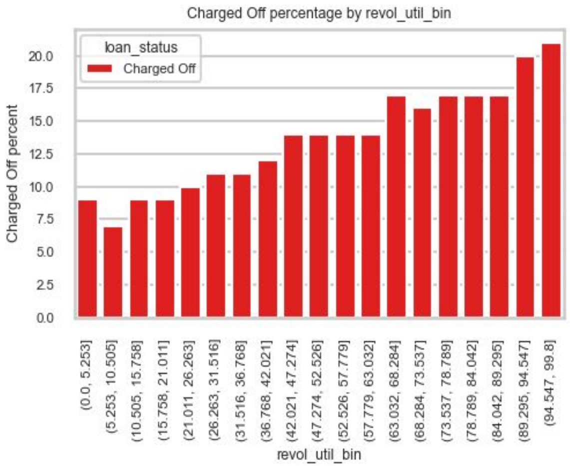




Chances of a loan default increase with increasing interest rates

# Observations - Revolving Utilization

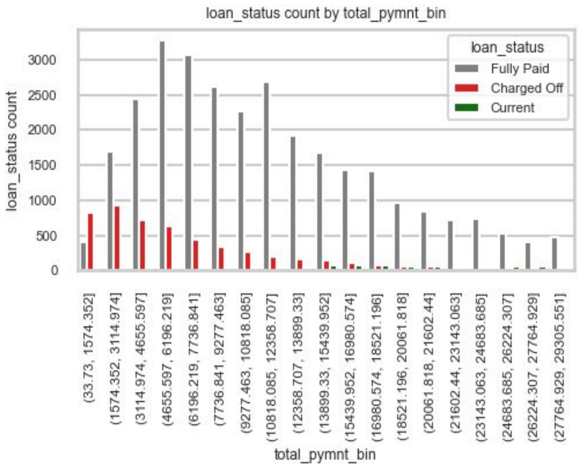


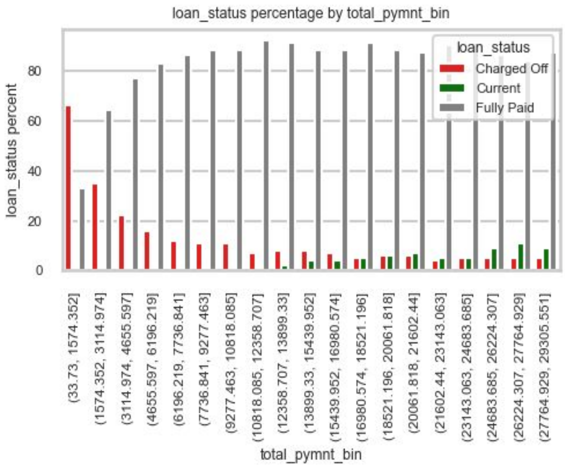



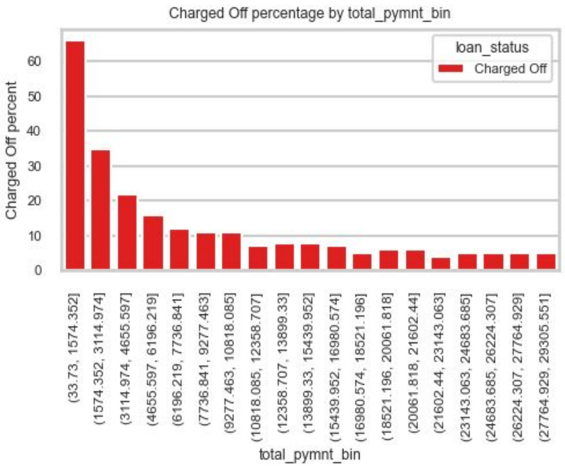




Higher utilization of credit limit (revolving balance utilization) shows signs of distress and a weak financial stability. The higher the utilization, the higher are the chances of loan defaults.

# Observations - Total payments

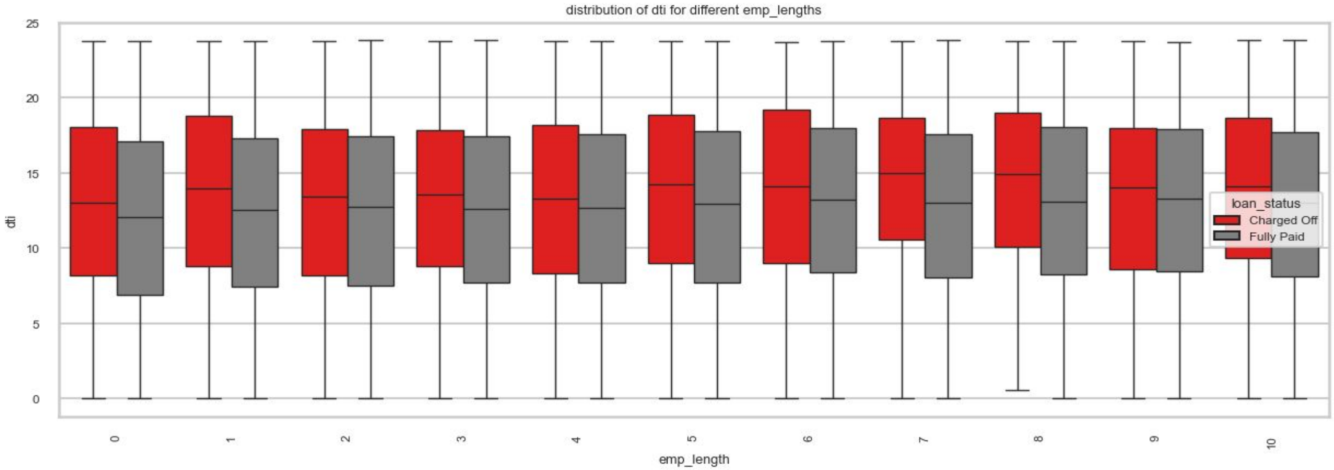


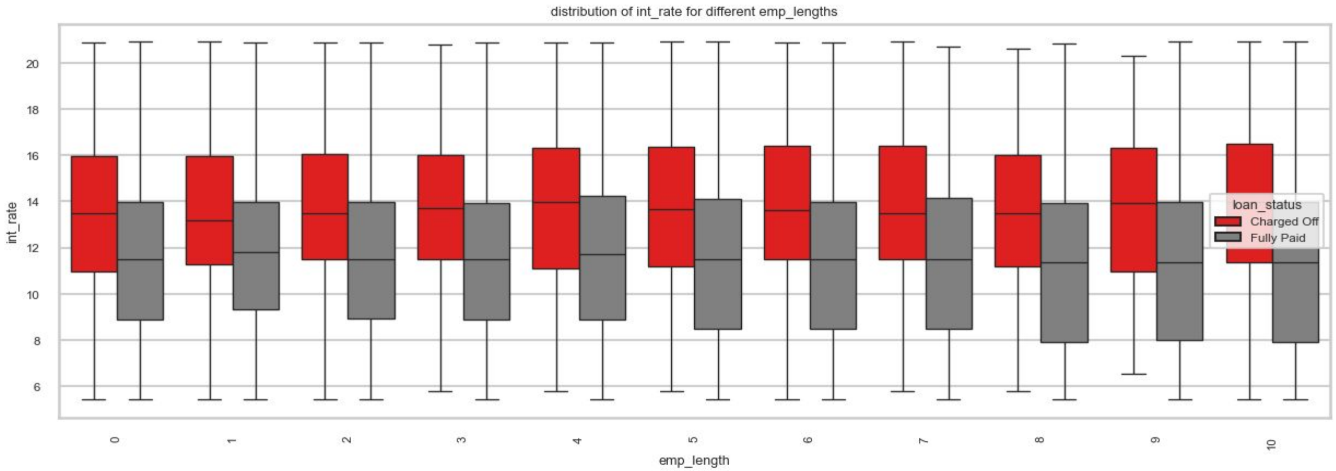

Borrowers who have done more payments show intent of fully paying the loan

# Observations - Debt to Income Ratio



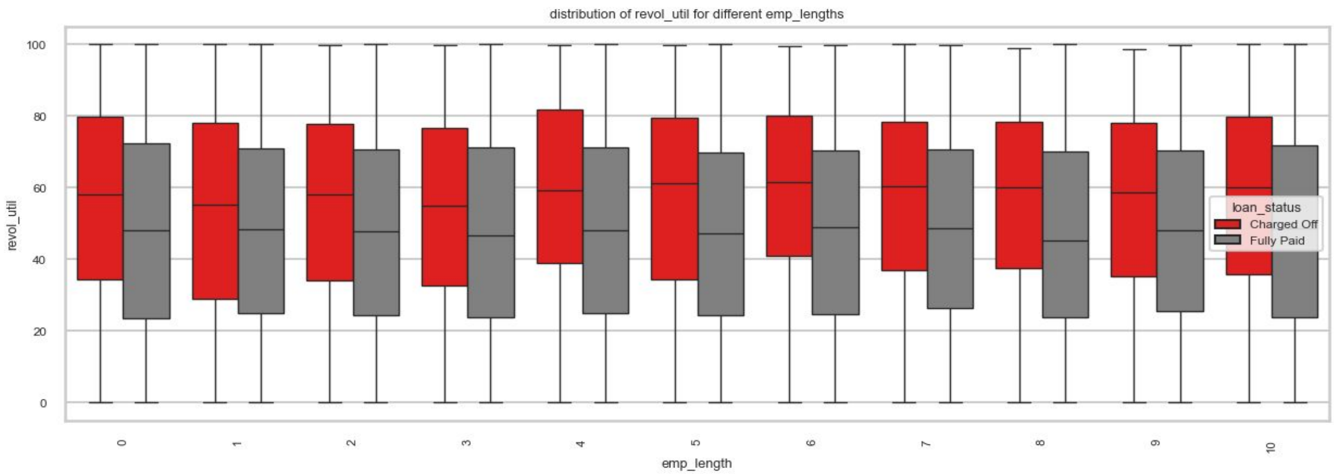

Borrowers having high debt to income ratio across all segments have higher chances of charging off

# Observations - Interest Rates



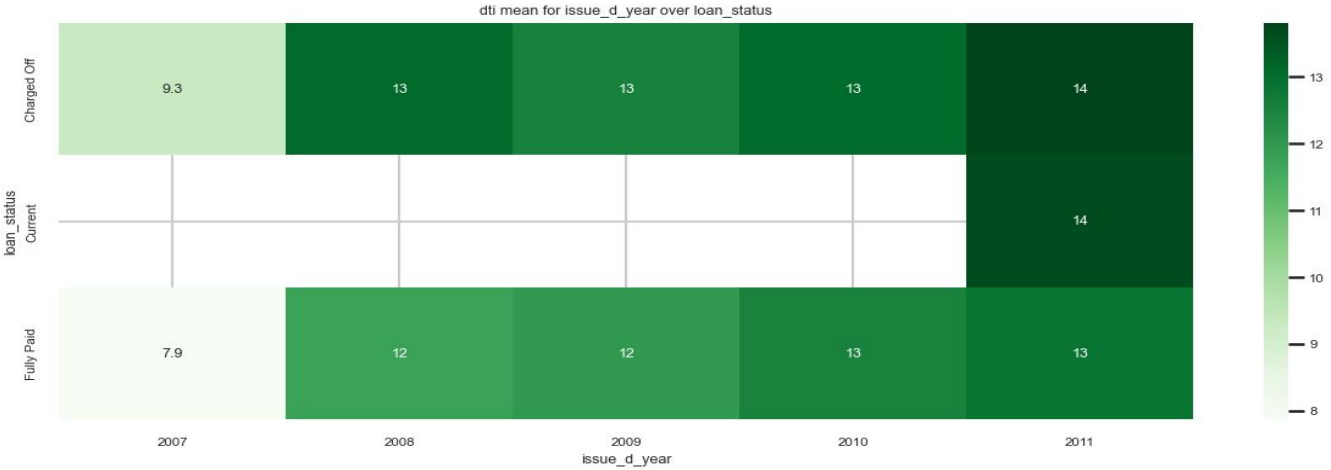

Borrowers having higher interest rates tend to default more as compared to borrowers having lower interest rates, across all categories

# Observations - Revolving Utilization




Borrowers having higher utilization of credit limit tend to charge off more across all the categories

# Observations - Debt to Income ratio




Debt to income ratio has been increasing over the years, which is also leading to higher charge offs

# Conclusions & Actionable Insights

- Borrowers with high debt to income ratio, high revolving balance utilization, borrowers who needed to have multiple inquiries in the last 6 months and borrowers who have had publicly recorded bankruptcies, need proper scrutiny before lending
- Longer term duration, higher interest rates, low annual income pose more risk of being defaulted, hence borrowers with high annual income, lower term duration and relatively lower interest rates are more promising clients
- Events like recession (2008-2009) are likely to impact the lending industry negatively and more defaulters will be seen
- There seem to be a discrepancy in the annual income verification process, as the verified borrowers seem to default more
- Lending club platform needs to re-look at sub-grade “G” assignments, but overall LC Grades properly depict the risk of the loan entry.