Lending Club Case Study

Case Study Contributors

Rupa Gattani

Shilpa PV

Business Use-case:

- In the context of loan approval at our consumer finance company, two risks are considered: declining a loan to a creditworthy applicant results in lost business opportunity, while approving a loan to a high-risk applicant may lead to financial losses due to potential defaults. Balancing these risks is crucial in making informed lending decisions.

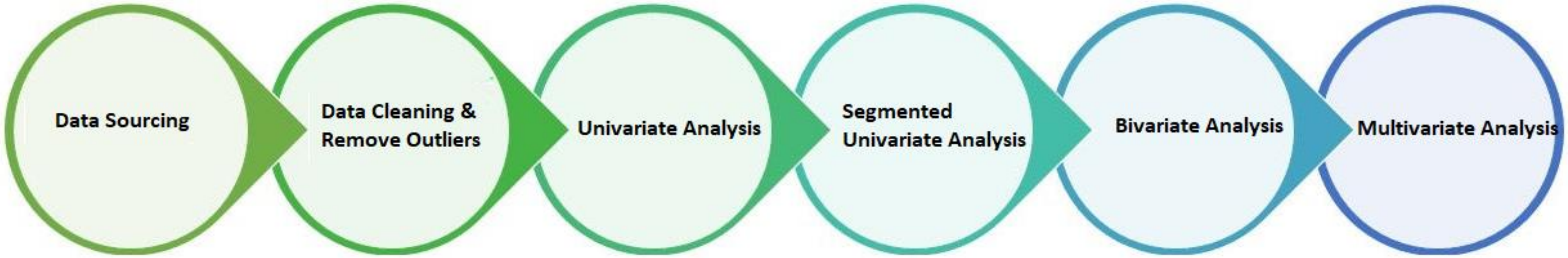
Objective:

- This case study aims to use Exploratory Data Analysis (EDA) to identify risky loan applicants and reduce credit losses. By pointing key driver variables of loan default, the company seeks actionable insights for portfolio management and risk assessment improvements.

Approach:

We have taken the below approach to perform Exploratory Data Analysis in this case study-

- Data Sourcing
- Data Cleaning
- Remove Outliers
- Univariate Analysis
- Segmented Univariate Analysis
- Bi-Variate/Multi-Variate Analysis



Data Understanding:

- From the given data set we have taken these variables for identifying the defaulters after cleaning the data.

- **Target Variable :** loan_status

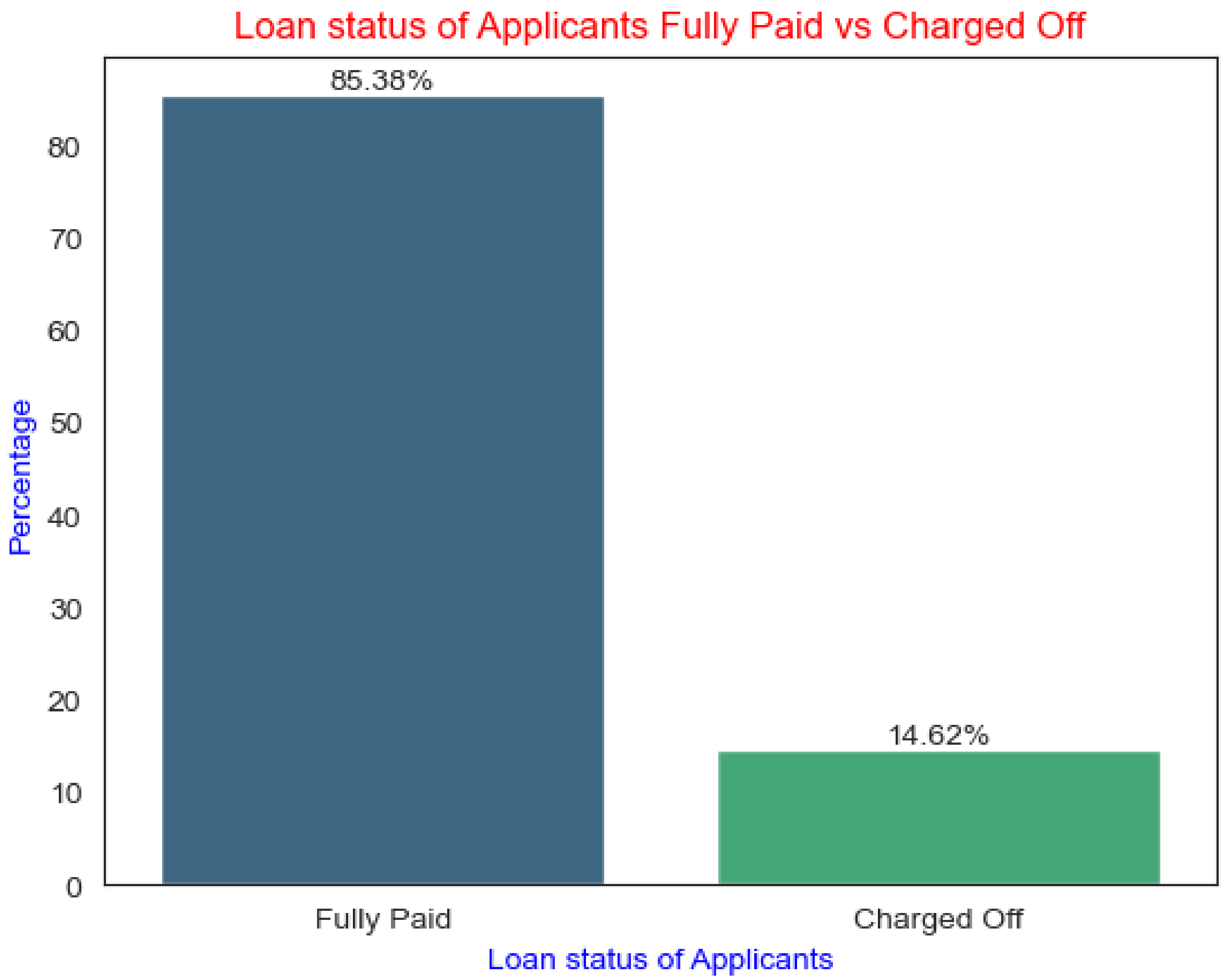
Variables considered for our Analysis :

Categorical columns	Numerical Columns
Term	loan_amnt
grade	int_rate
sub_grade	annual_inc
emp_length	pub_rec_bankruptcies
home_ownership	dti
verification_status	delinq_2years
purpose	open_acc
addr_state	public_rec
issue_month	revol_util
issue_year	

Distribution of overall Loan Status

Approximately 85.38% of applicants have fully paid their loans, indicating a strong repayment history.

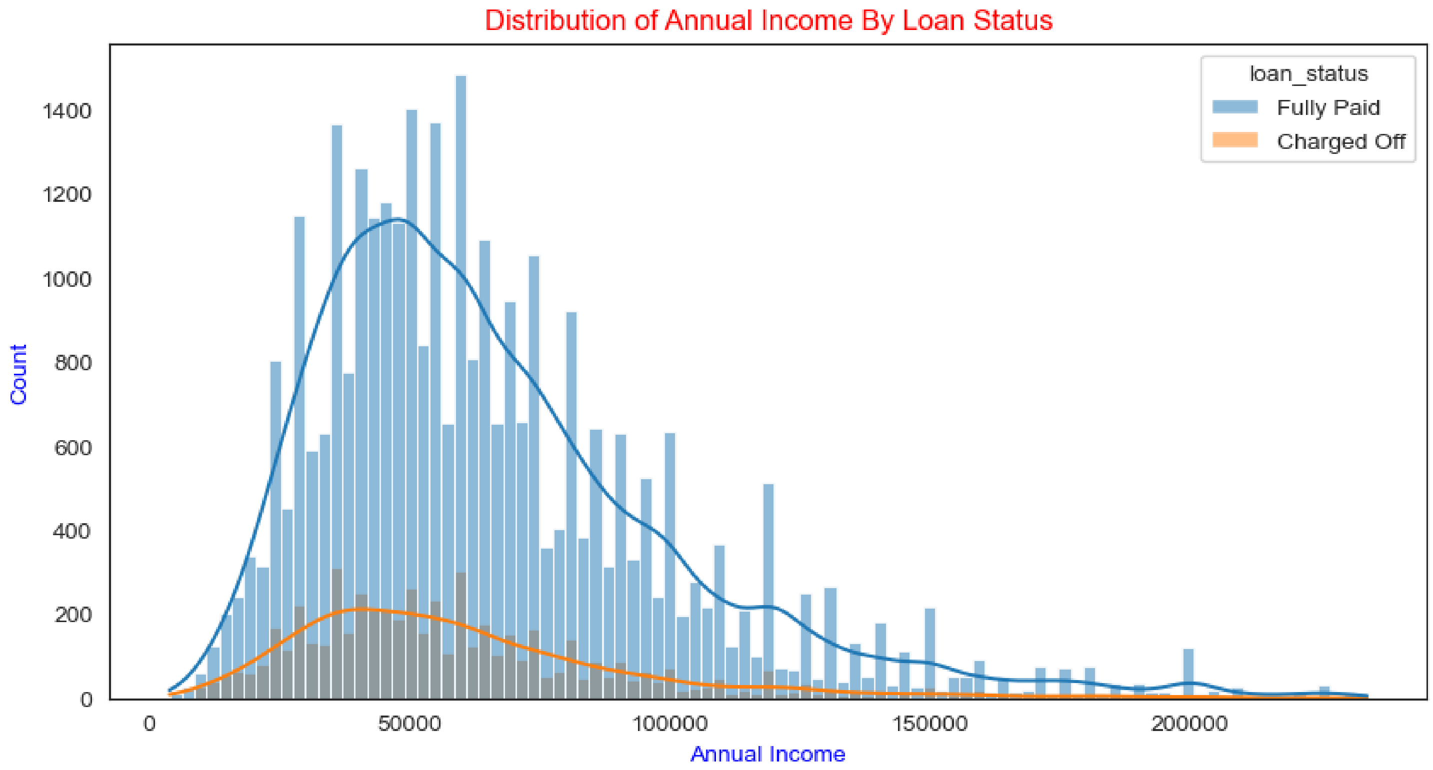
Overall default rate: Only **14.62%** of applicants have defaulted on their loans, suggesting a relatively low risk of default for the lender.



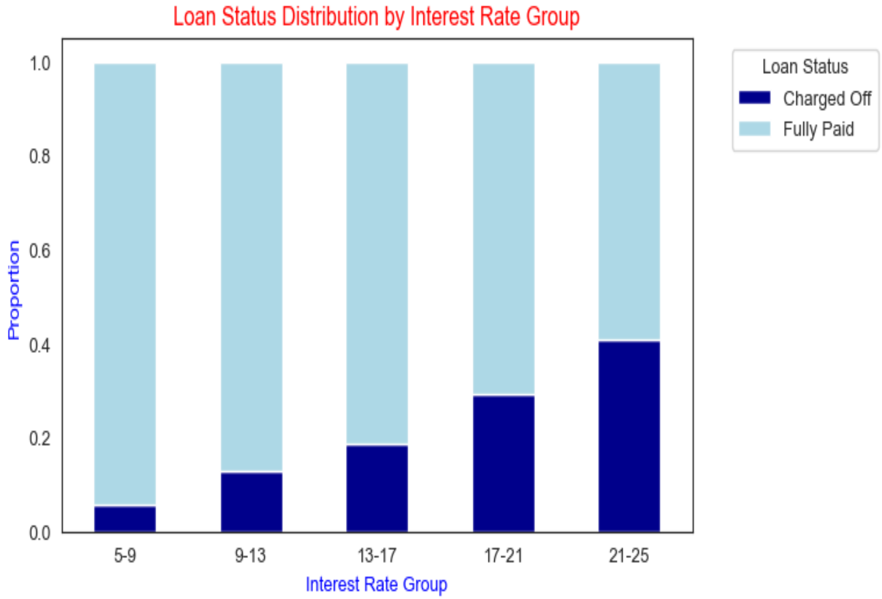
Distribution of Annual Income by Loan Status

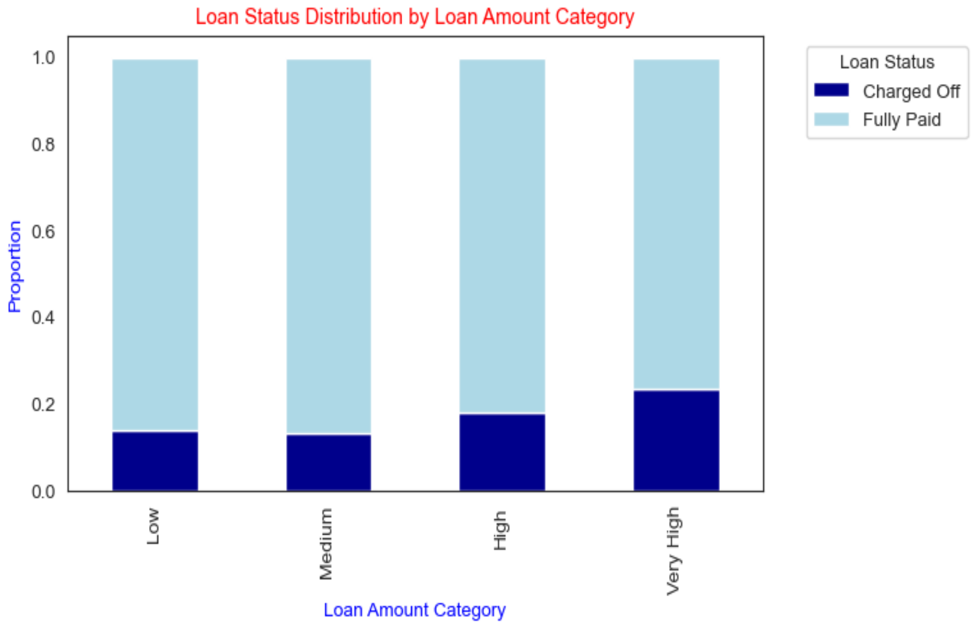
- The distribution of annual income is right-skewed for both "Fully Paid" and "Charged Off" loan statuses, indicating that most applicants have lower to moderate incomes, Very few applicants having high incomes.

- **Loans are more likely to be "Charged Off" among applicants with lower annual incomes.**



Interest Rate and Loan Amount by Loan Status





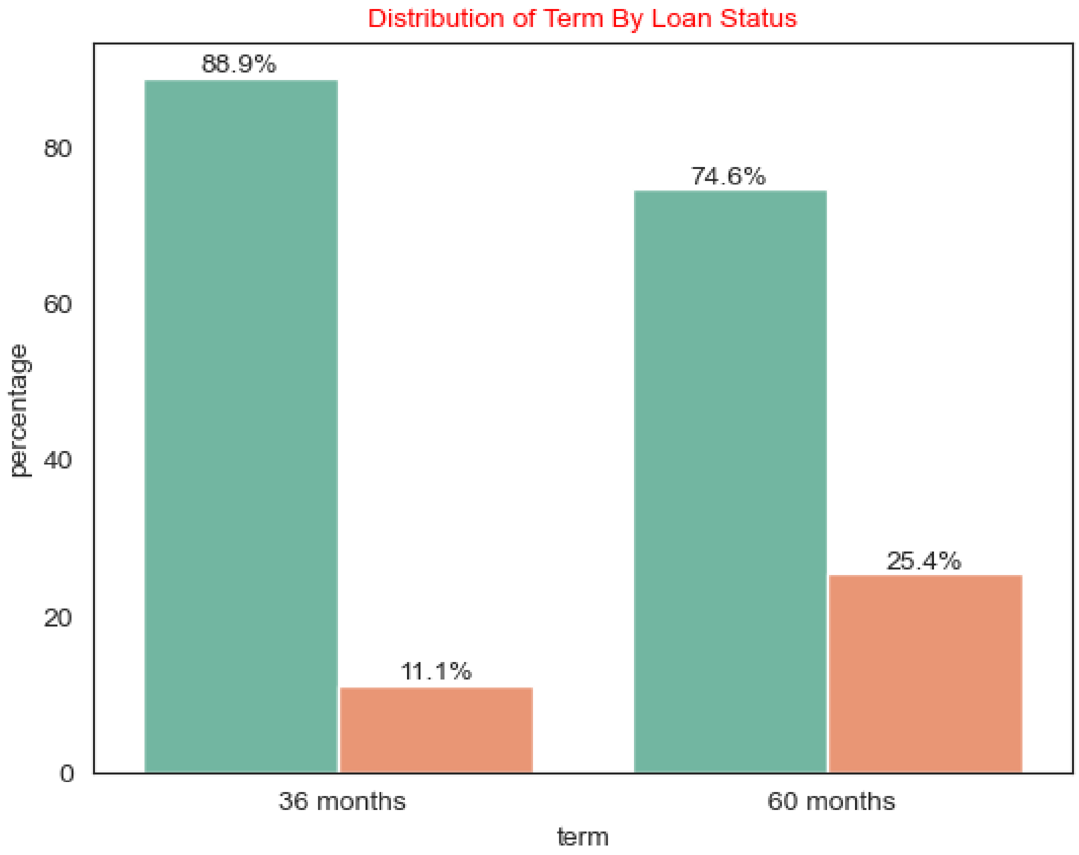
- **As the interest rate increases, the proportion of "Charged Off" loans increases significantly**.
- In the higher interest rate groups (17-21% and 21-25%), the applicants for "Charged Off" loans increases dramatically which indicates that higher the interest rate more the probability of loan default.

- **For Very High loan amounts the charged off are more and increasing gradually.**

Impact of Term on Loan Status

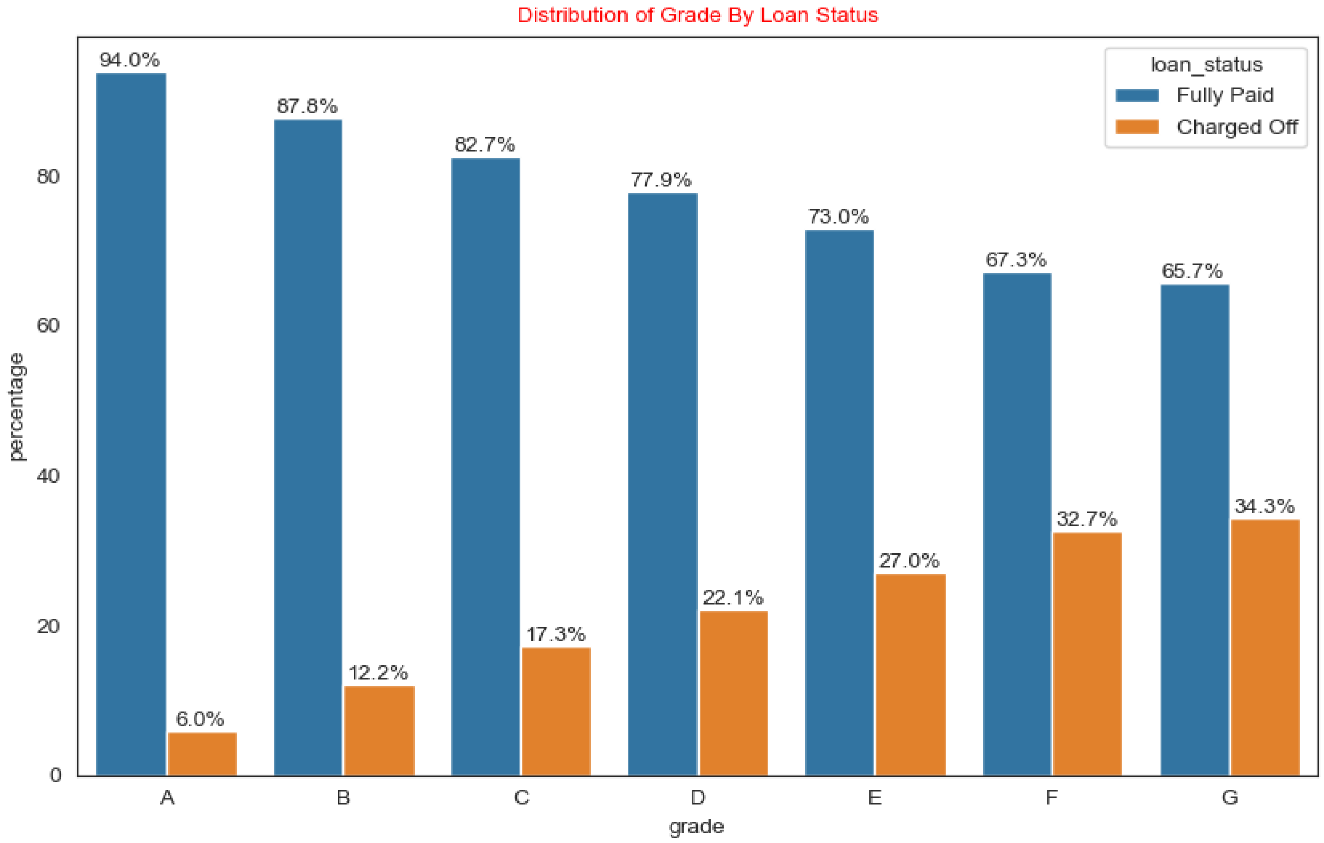
- **Charged Off loans show a higher default rate for 60-month terms compared to 36-month terms.**

- For Fully Paid loans, people who have taken 36-month term are likely to repay earlier than 60-month term loan.



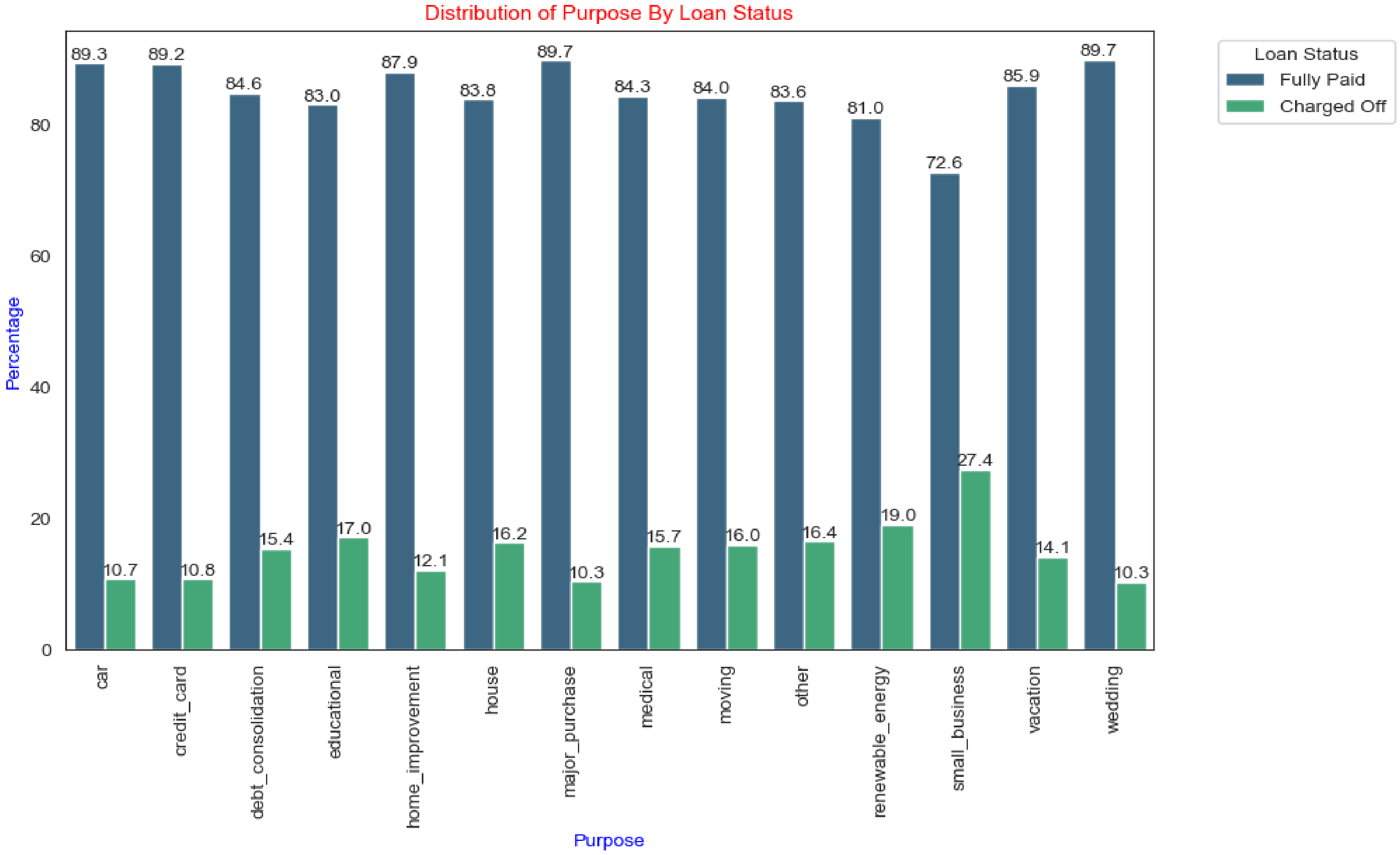
Relationship between Grade and Loan Status

- **Lower-grade loans (D, E, F, G) have a higher likelihood of defaulting.**
- As the grade decreases, the likelihood of getting "Charged Off" increases.
- Loans with higher grades (A, B) are more likely to be fully paid.



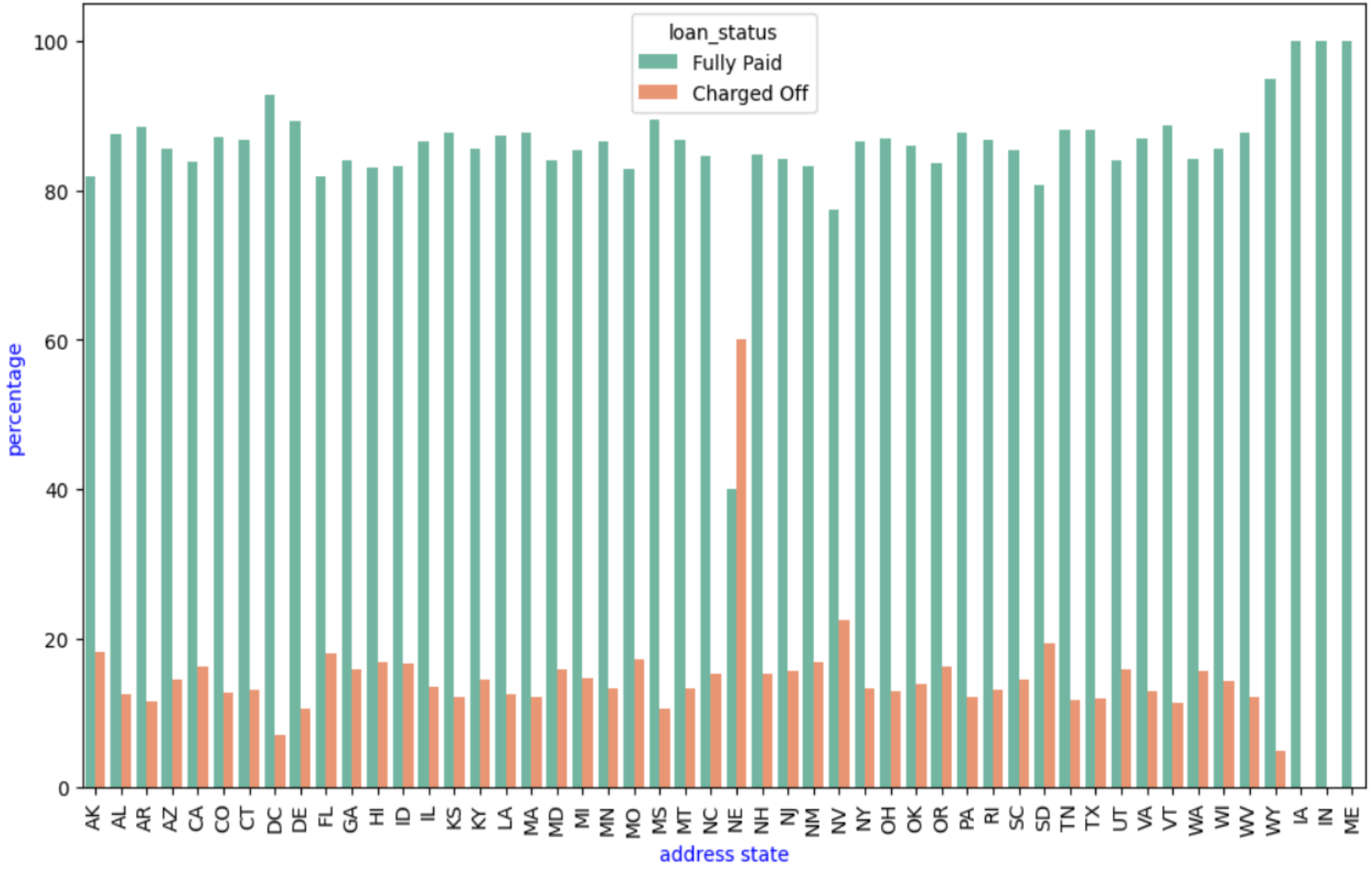
Distribution of Purpose by Loan Status

- Higher % of loans borrowed for purposes such as 'small_business', renewable_energy, and educational reasons defaulted



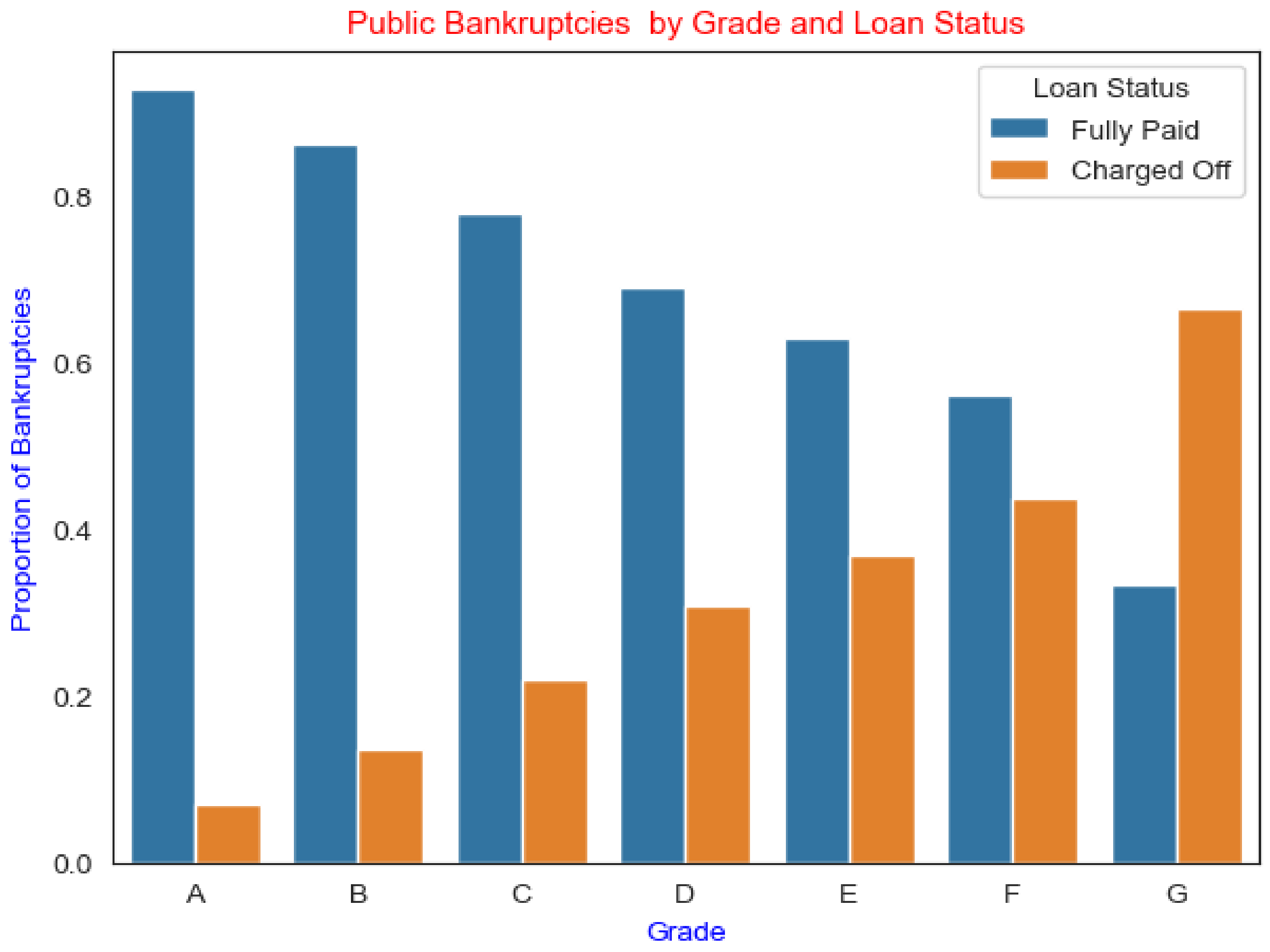
Loan Status spread across Address States

- **The highest number of defaulters are in the states of NE,NV,SD,AK.**

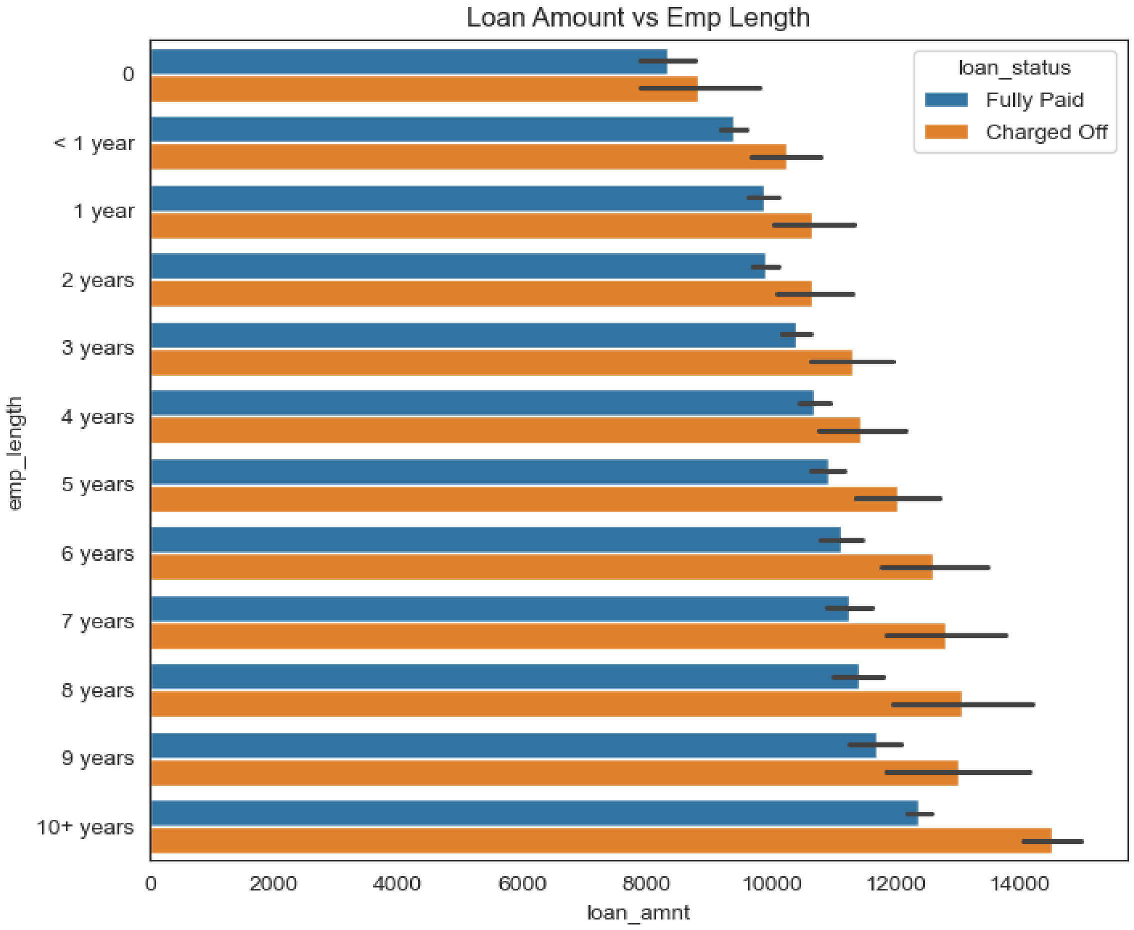


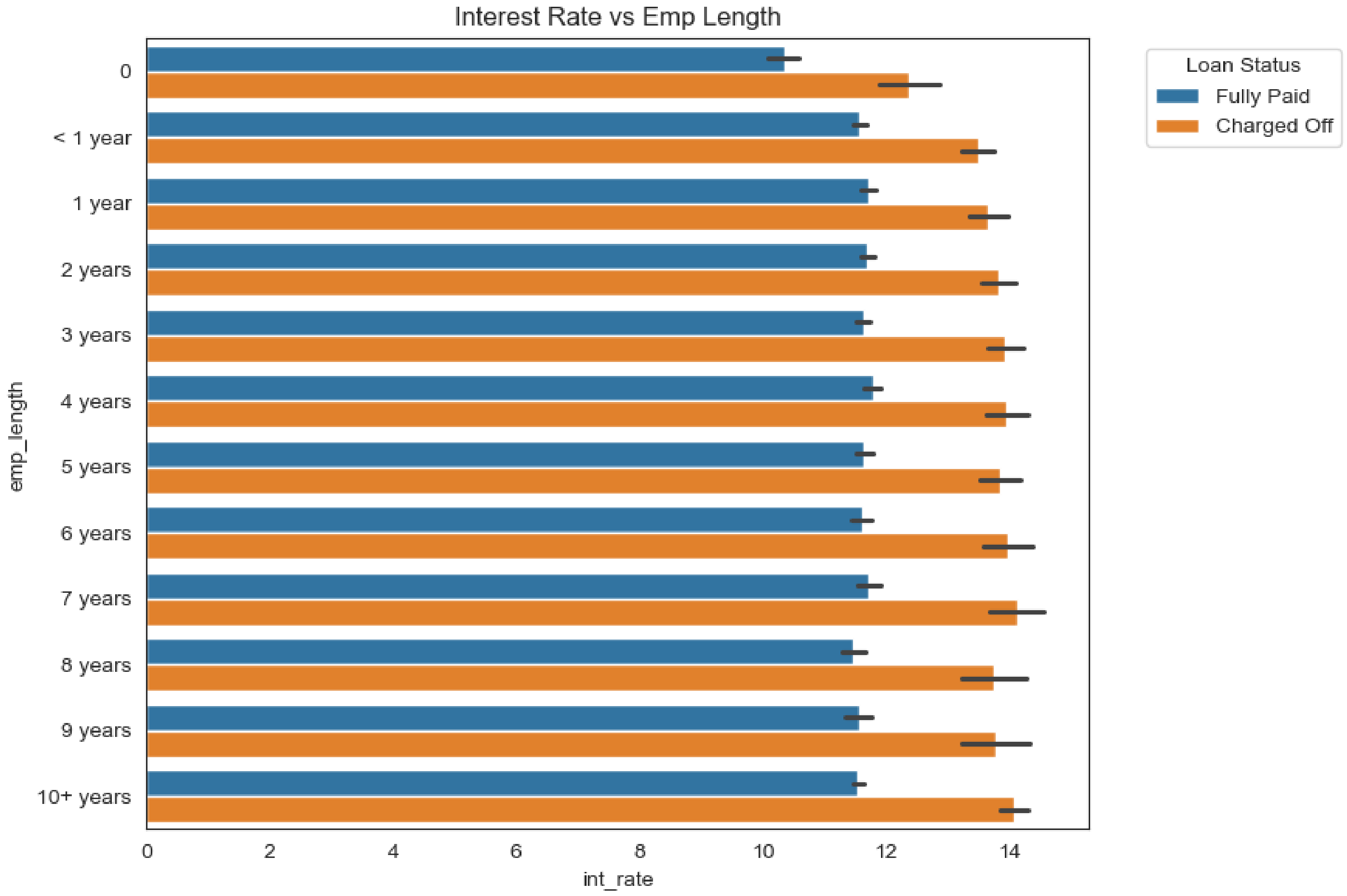
Bankruptcy trends by Grade and Loan Status

- **The percentage of getting Bankrupt increases as we move from high grade to low grade, so as the risk of ChargedOff increases.**



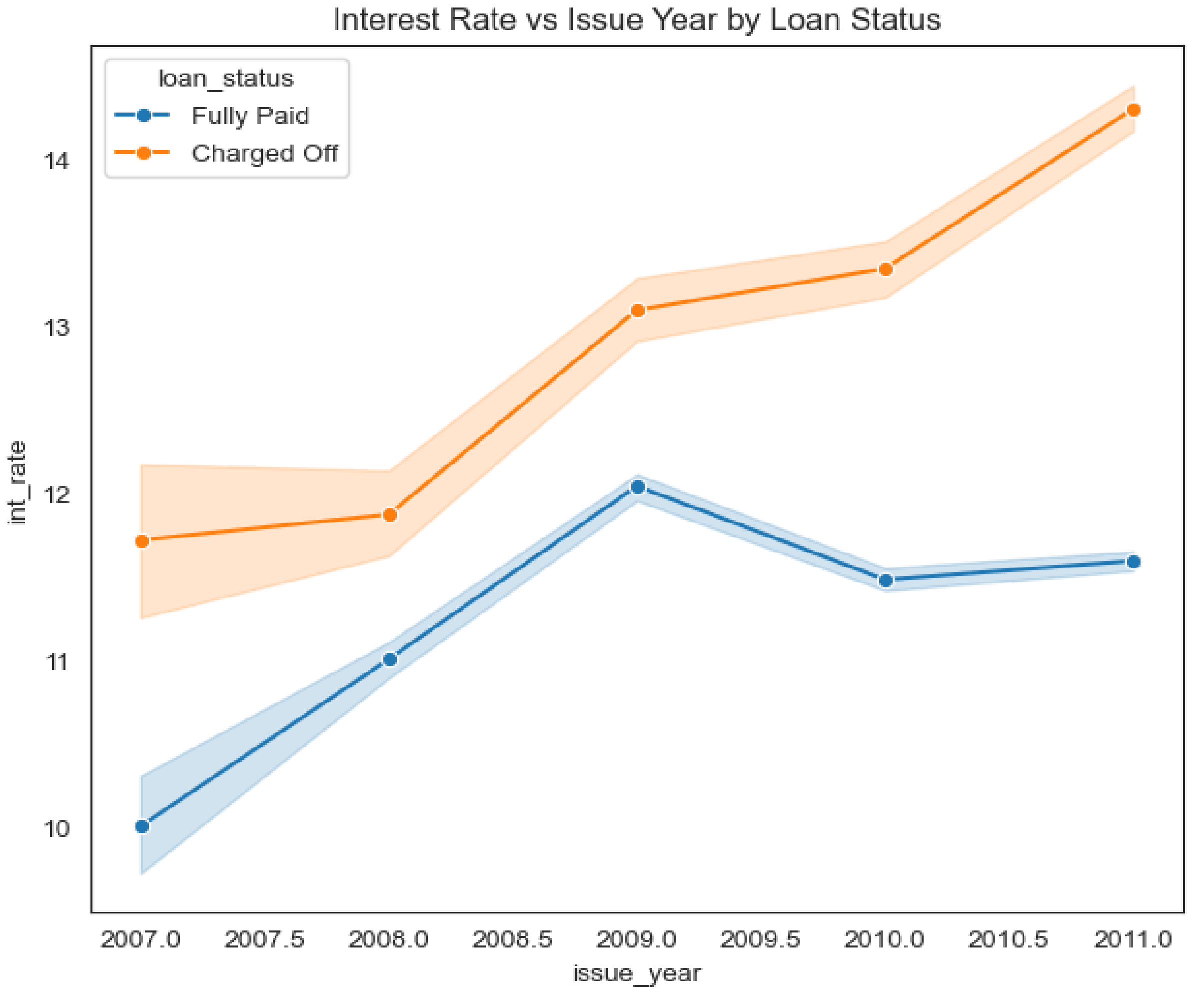
Loan Amount and Interest Rates by Emp Length, Loan Status

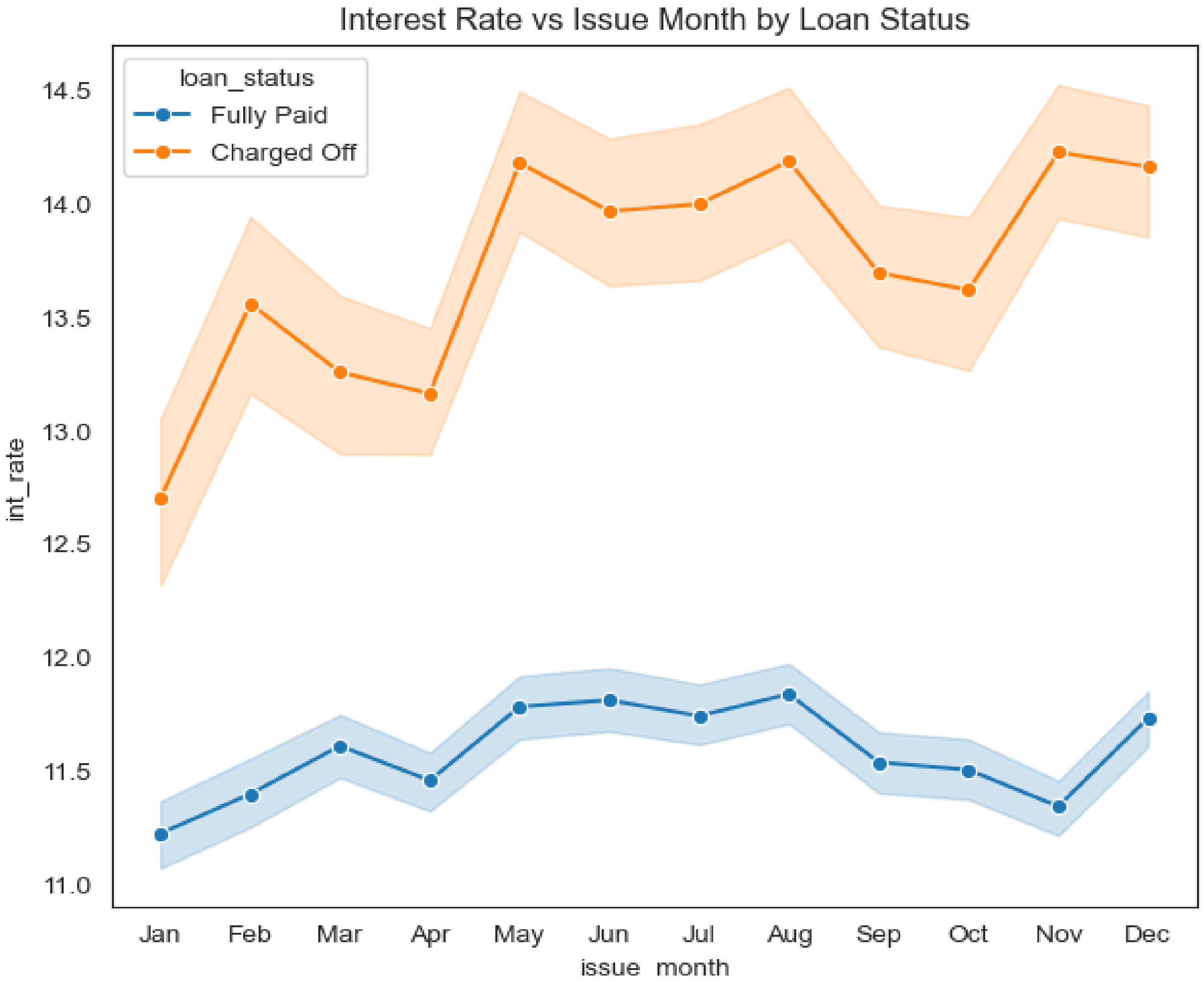




- ***Employment Length and Loan Amount:*** The graph **suggests that the likelihood of loan defaulters increases with both employment length and loan amount.** Applicants with longer employment lengths tend to have higher loan amounts.
- ***Employment Length and Interest Rates:*** **The graph suggests that the interest rate is high across the employment length for "Charged Off" loans**. Also interest rate seems to be consistent across the emp_length.

Issue Year and Issue Month by Interest Rate, Loan Status

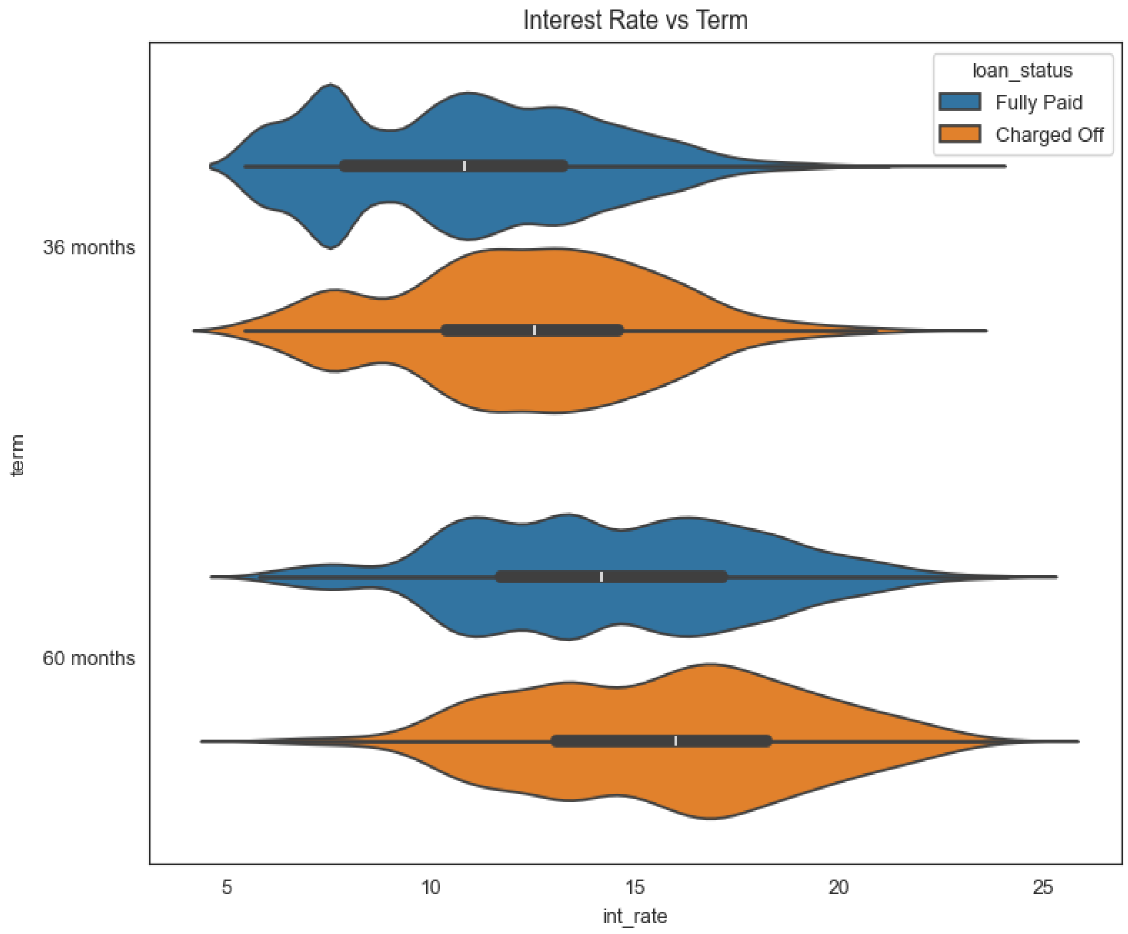


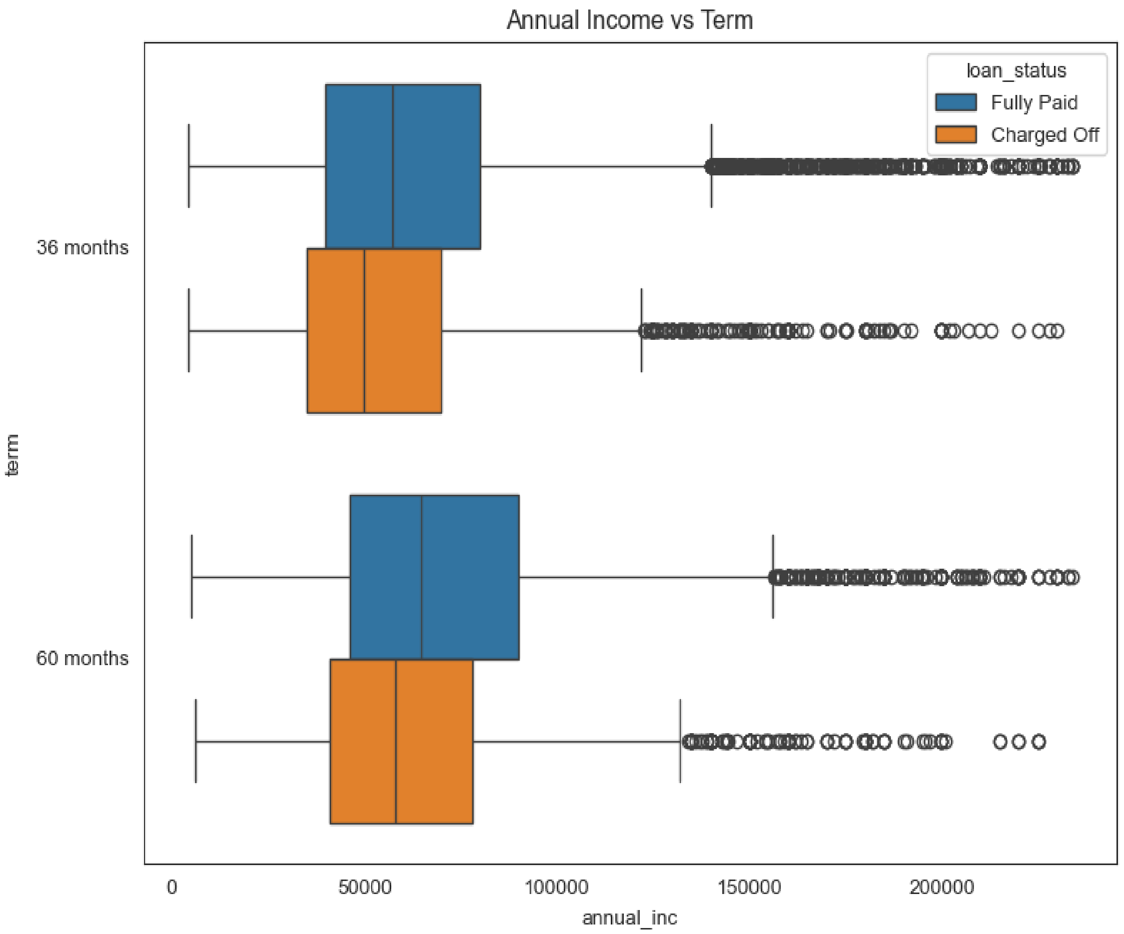


- ***Interest Rate vs Issue Year:*** **Interest rates for "Charged Off" loans increase significantly over the years, peaking around 2011**. Interest rates for "Fully Paid" loans also increase over the years but at a slower rate.

- ***Interest Rate vs Issue Month:*** **Interest rates for "Charged Off" loans are generally higher across all months, with noticeable peaks in May, Aug and November**. Interest rates for "Fully Paid" loans show less variability and generally lower values across the months.

Interest Rate and Annual Income by Term, Loan Status





- ***<u>Interest-Rate vs Term:</u>*** **The impact of Term on Interest Rate for Fully Paid and charged off is similar.** Higher interest rates are associated with a higher proportion of charged-off loans.

- ***<u>Annual Income vs Term:</u>*** **Lower the Annual Income higher the chance to get Charged Off.** The term doesn't have any significant impact on the annual_inc against loan_status.

Key Findings

- **Interest rate** is a strong predictor of loan default risk. Higher interest rates are associated with significantly increased charge-off rates.
- **Borrower income** is a critical factor in loan performance. Lower-income individuals exhibit higher default rates.
- **Loan grade** is a significant determinant of loan risk. Lower-grade loans (D, E, F, G) have a substantially higher likelihood of default compared to higher-grade loans (A, B).
- **Debt-to-income (DTI) ratio** is positively correlated with loan default risk. Higher DTI ratios indicate increased likelihood of charge-off.
- **Bankruptcy history** is a strong indicator of elevated loan default risk.

Driving Factor Variables:

- ✓ Interest rate
- ✓ Loan grade
- ✓ Annual income
- ✓ Debt-to-income ratio
- ✓ Public recorded bankruptcies

Observations

- Applicants with shorter employment lengths and higher interest rates are more likely to default on their loans.
- Loan purpose impacts loan performance. Loans for small businesses tend to have higher interest rates and default rates.
- Economic conditions can influence loan performance. The year 2011, assumed to be a period of economic downturn, saw an increase in both loan volume and default rates.
- Loan grade is a strong indicator of loan performance. Higher-grade loans have lower default rates.
- Income level is associated with loan performance. Lower-income borrowers have a higher likelihood of default compared to higher-income borrowers.
- Bankruptcy history is likely correlated with loan default, as indicated by the relationship between grade and charge-off rates.