LendingClub

Lending Clubs Loans Analytics Case Study

- Presented By –

Selva Sankari (Facilitor)
LakshmiKanth

ABOUT THE COMPANY

LendingClub

This company is the **largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.** Borrowers can easily access lower interest rate loans through a fast online interface.

Like most other lending companies, lending loans to **'risky'** applicants is the largest source of financial loss (called credit loss). **Credit loss** is the amount of money lost by the lender **when the borrower refuses to pay or runs away with the money owed.** In other words, **borrowers who default cause the largest amount of loss** to the lenders. In this case, the customers labelled as **'charged-off' are the 'defaulters'**.

Our commitment

Reinvent banking
We're the only full-spectrum fintech marketplace bank built on the belief that innovative, creative solutions deliver more value and a better experience.

Do what's right
We're leading the governance of a new industry, developing ethical, responsible ways to bring greater value and better opportunities to customers.

Deliver world-class experiences
Accessing money should be seamless. We're committed to making borrowing and saving simple and easy for everyone.

Level the playing field
Everyone deserves a better financial future if they strive for it. We've built a marketplace that keeps costs low and opportunity high.

OBJECTIVES

Business Understanding

When the company receives a loan application, ***the company has to make a decision 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 **not 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.

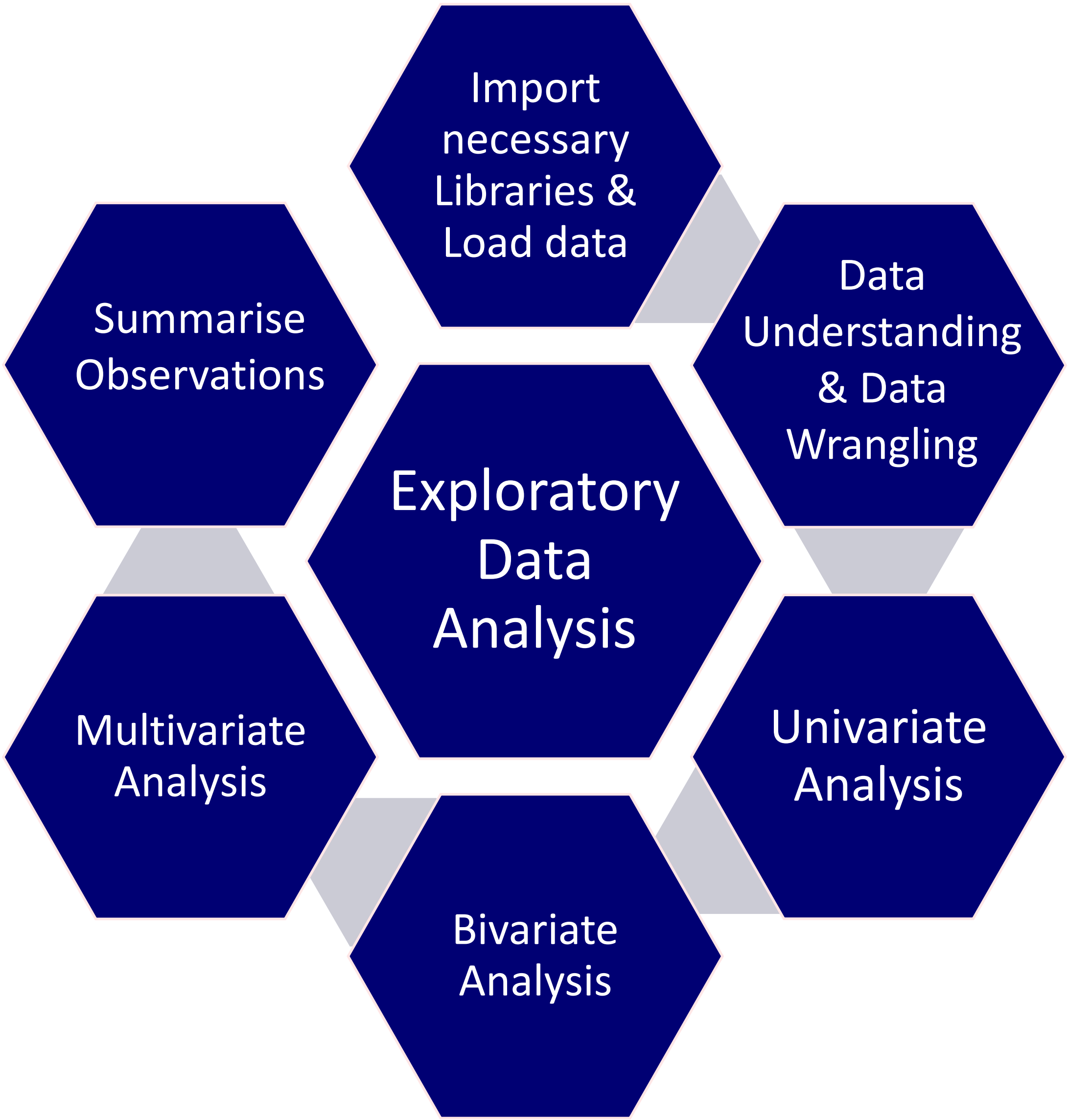
Business Requirements (Objective of Case Study)

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.** If risky loan applicants can be identified early, **then such loans can be reduced thereby cutting down the amount of credit loss.**

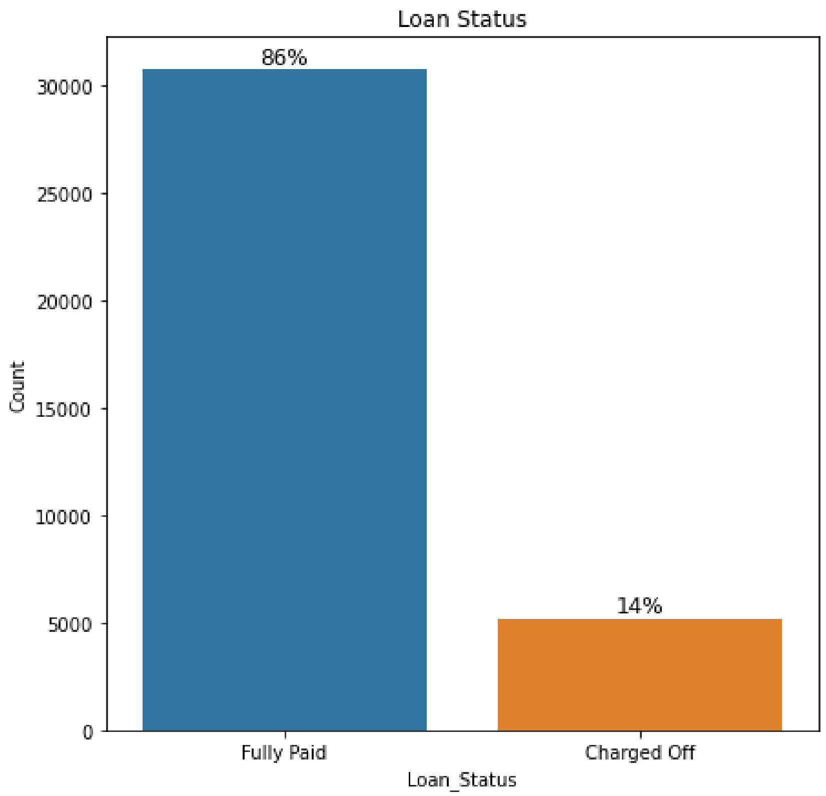
Therefore, the **need is to Perform a Detailed Exploratory Data Analysis on the Lending Club Loans Data**:

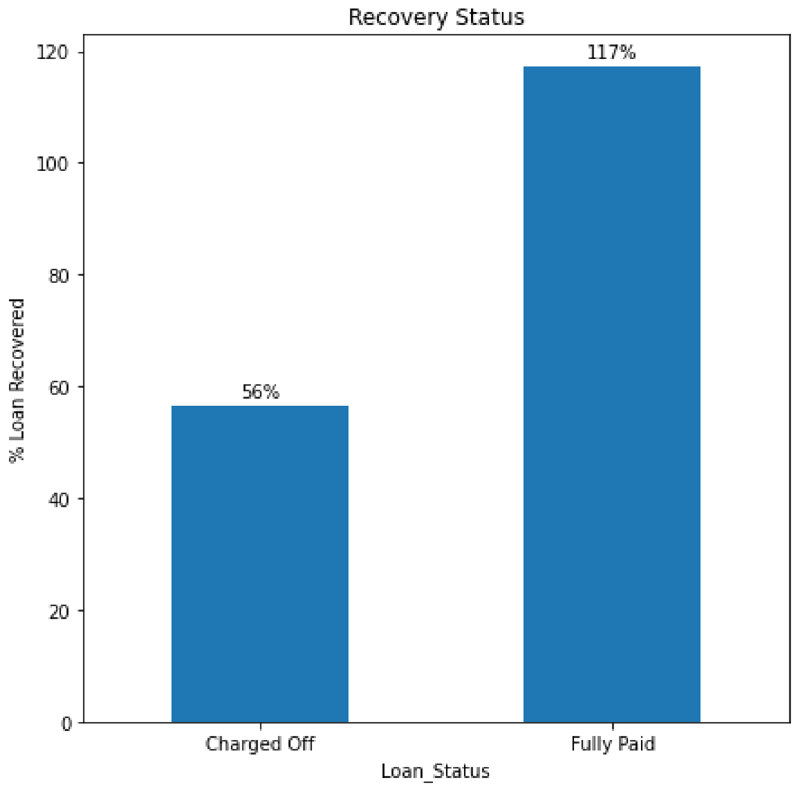
- **To identify patterns** which indicate if a **person is likely to default**.
- To understand how **consumer attributes** and **loan attributes** influence the **tendency of default**.
- **Recommend actions** such as denying the loan, reducing the loan amount, lending at a higher interest rate, etc.

APPROACH FOLLOWED FOR EDA



LOAN STATUS AND RECOVERY

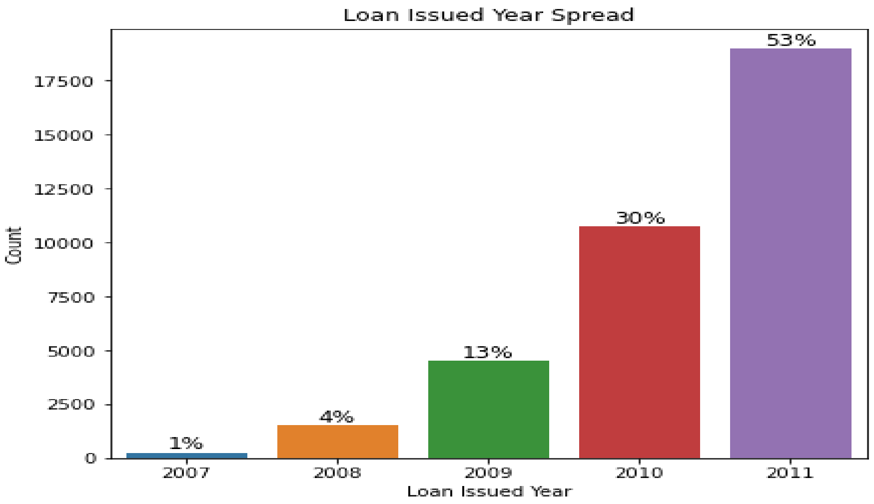


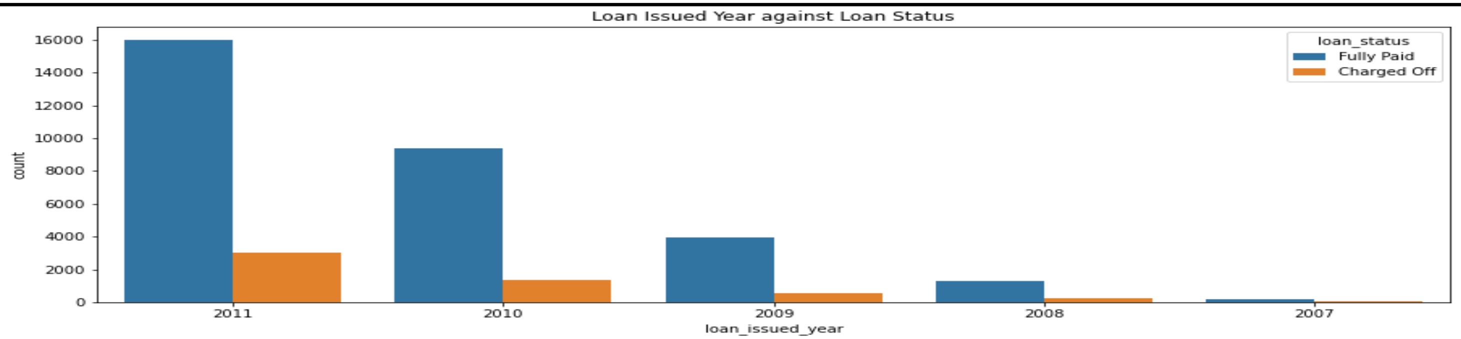


Observation:

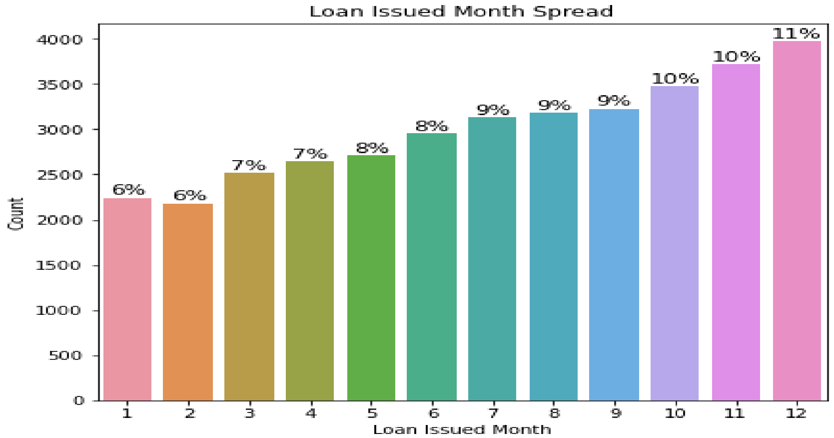
- **Off the Total Loans approved and issued, only 14% have defaulted.**
- **We can observe that only 56% of the Charged Off loans get recovered. However, in the Fully Paid loans recoveries are made more than the issued amount. This might be due to interests and other fees. This shows that the Fully Paid loans generate a profit of around 17%.**

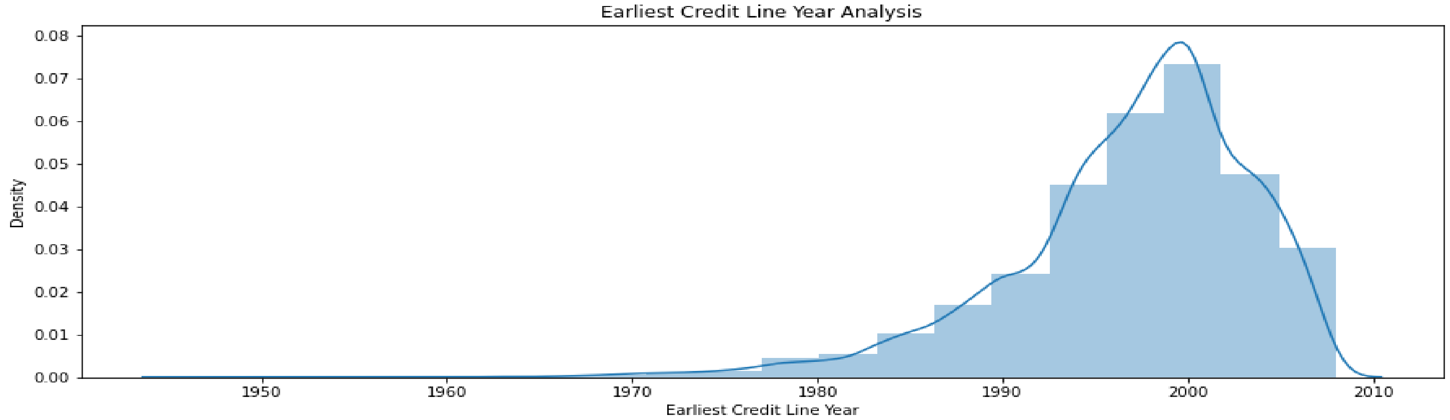
LOAN ISSUED YEAR ANALYSIS





loan_issued_year	Charged off %	Record Count
2007	17.15	239
2011	15.80	18984
2008	15.41	1493
2009	12.82	4502
2010	12.57	10719

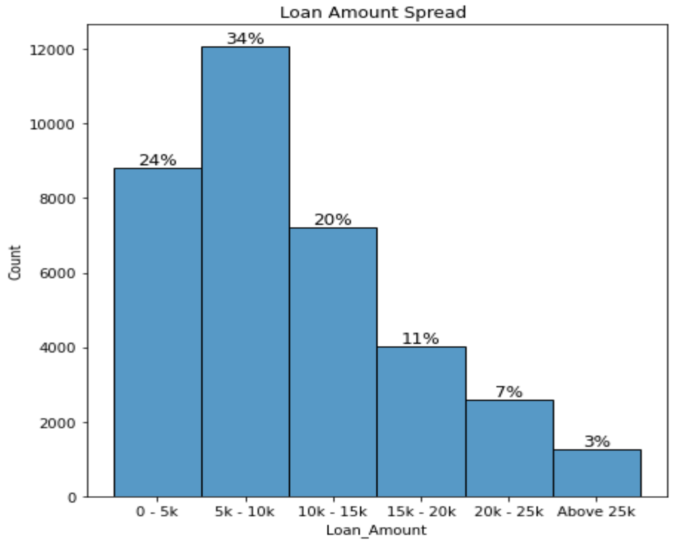


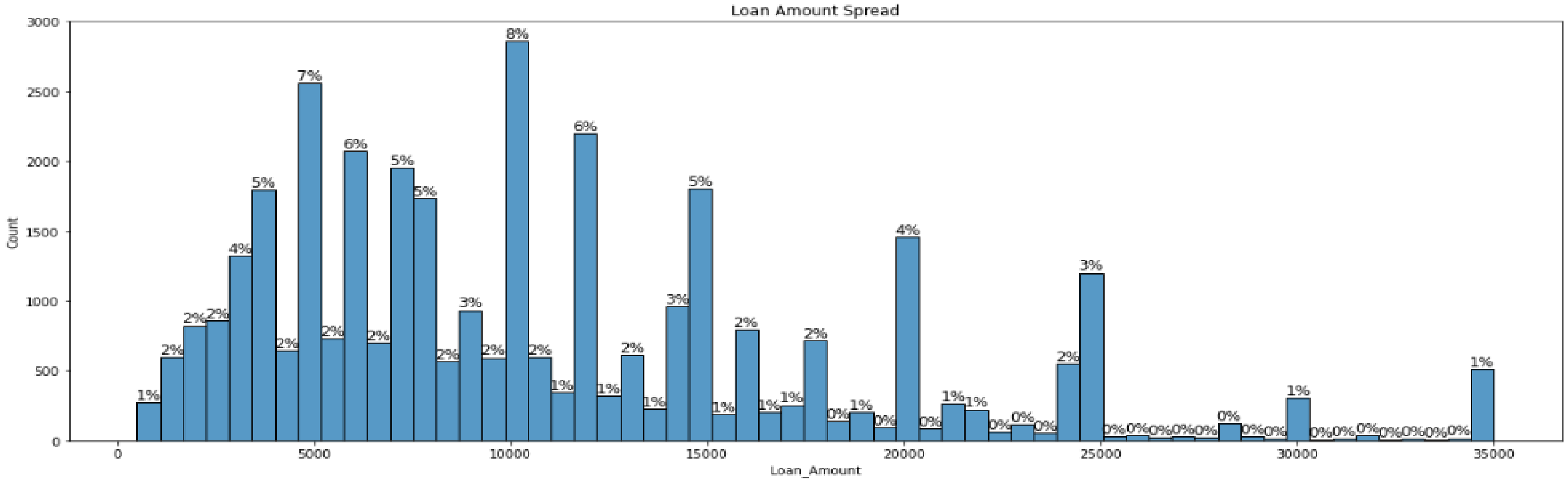


Observation:

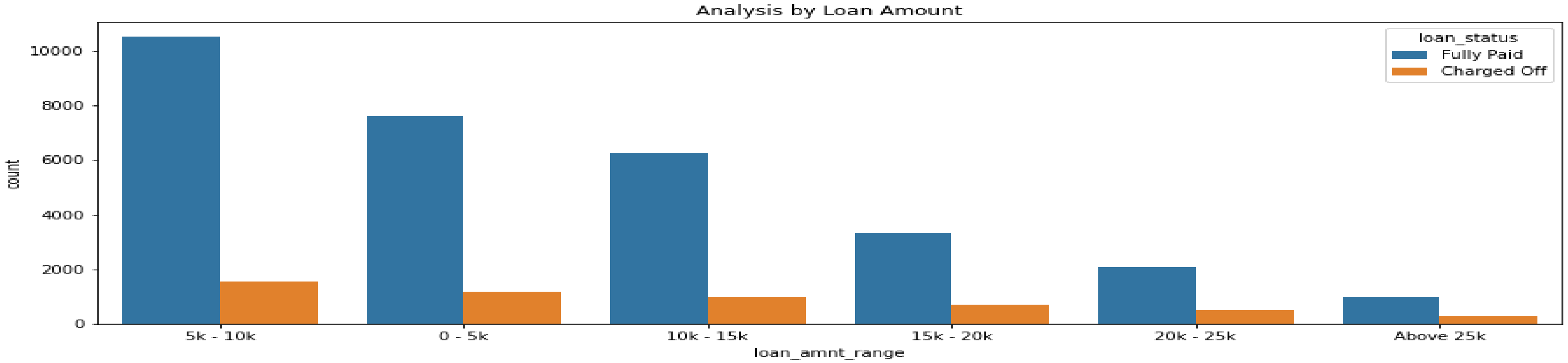
We can observe that the number of loans issued has been exponentially increasing year after year. It is also clearly seen that the majority of the loans are applied during the Q4 (Quarter 4) of every year. The Default rate is even spread across the years and the months and hence no significant observation can be made there.

LOAN AMOUNT ANALYSIS





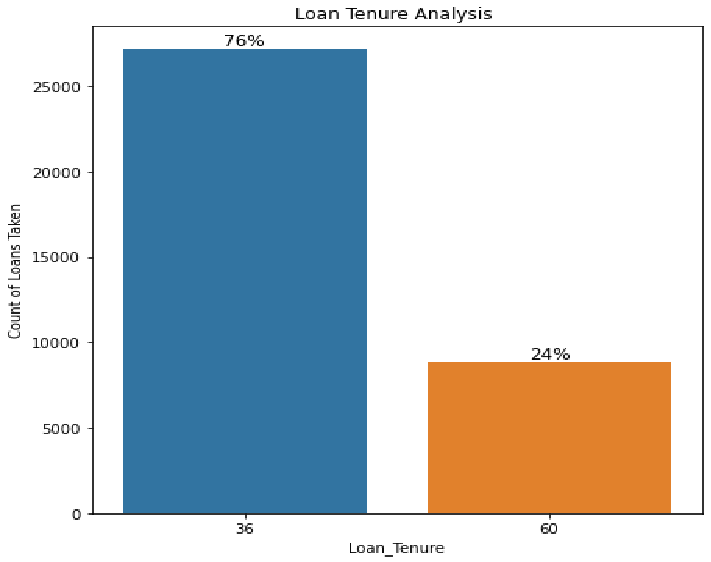
loan_amnt_range	Charged off %	Record Count
Above 25k	23.62	1249
20k - 25k	19.29	2577
15k - 20k	17.38	4028
10k - 15k	13.61	7225
0 - 5k	13.52	8804
5k - 10k	12.69	12054



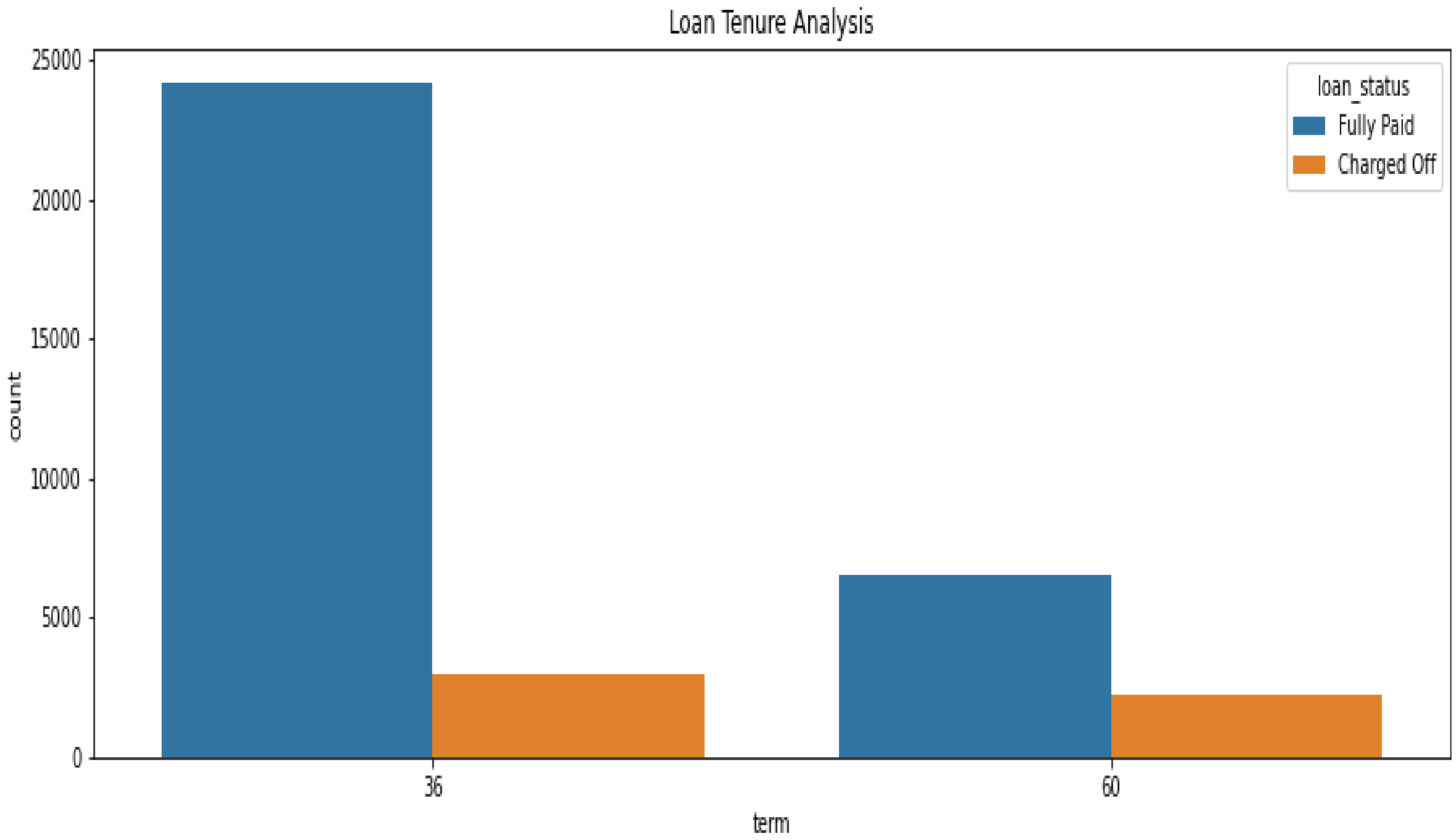
Observation:

Loan Amounts vary from 500 to 35000. Loans of 10k are the most commonly sought loan. 78% of the loans are under 15k and only 10% of the loans are greater than 20k. We can also observe that as the Loan Amount increases the Charge off % also increases. This shows that the higher loans, substantially carry a higher risk of default.

LOAN TENURE ANALYSIS



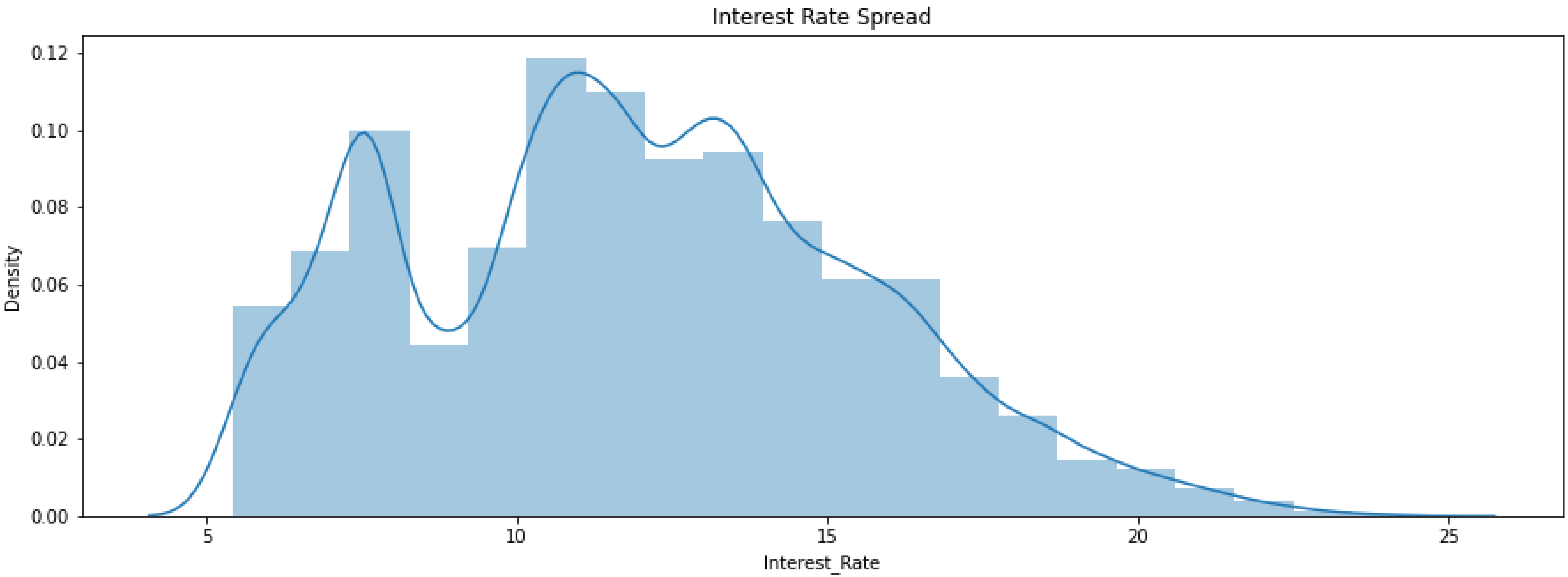
term	Charged off %	Record Count
60	25.45	8796
36	10.89	27141



Observation:

There are only 2 Loan Tenures - 36 months (3 yrs) and 60 months (5 yrs). 36 months is the most preferred choice for almost 76% of loan applicants. It is also the beneficial option for the business, since only 10% of this group ends in default. Whereas, the 60 months tenure has a default rate of 25%. This means that 1 out of every 4 60 month tenured loans has not been paid back.

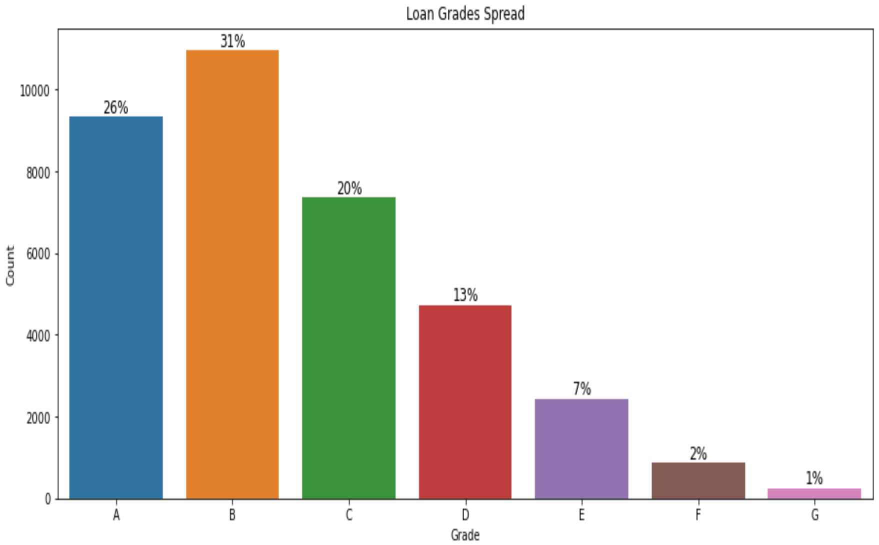
INTEREST RATE ANALYSIS

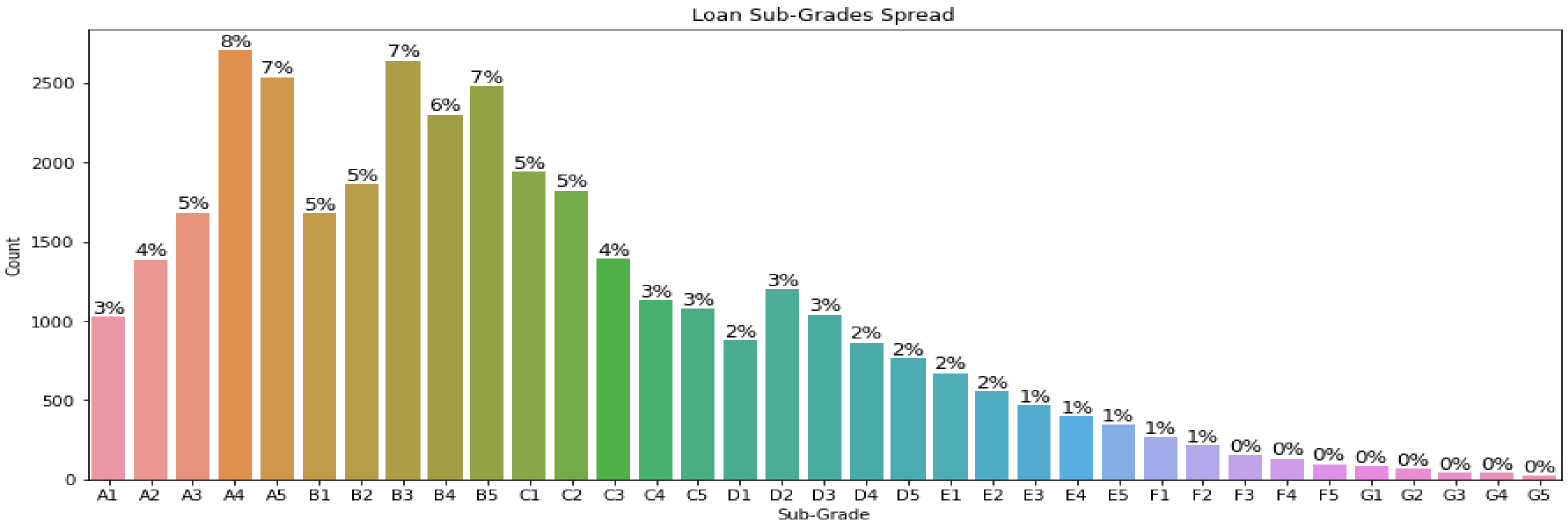


Observation:

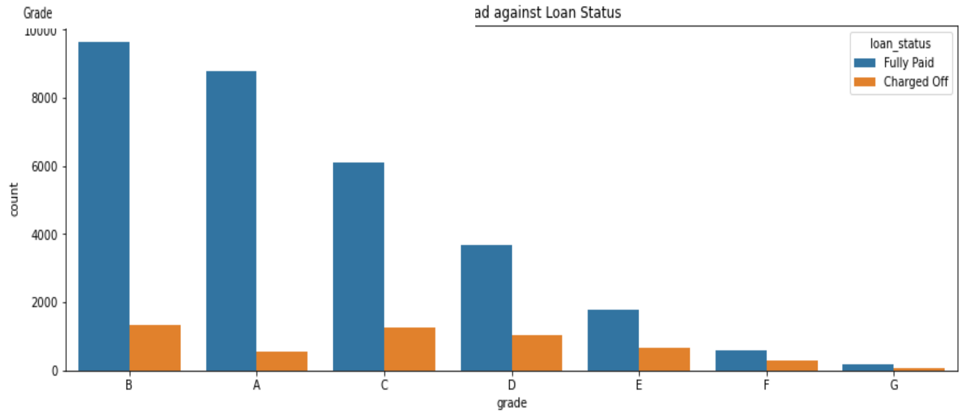
Almost 75% of the loans have 15% of lesser interest rates.

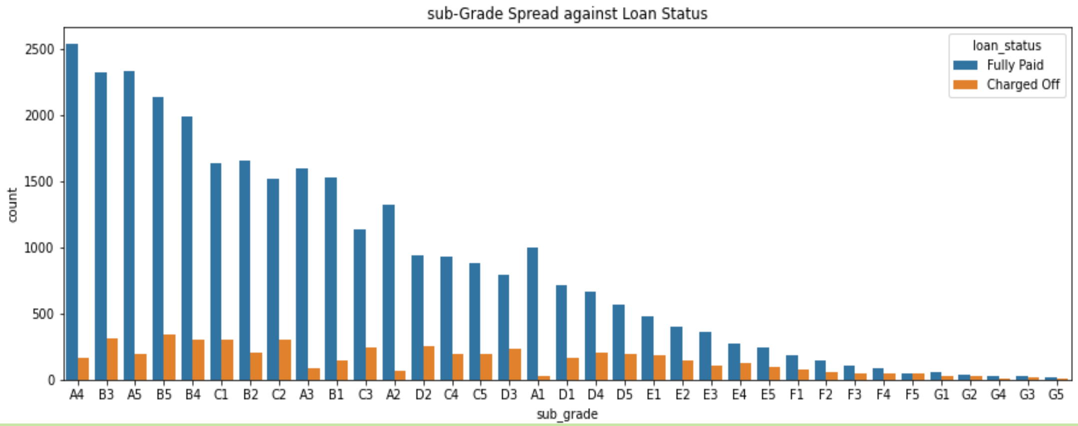
LOAN GRADES ANALYSIS





grade	Charged off %	Record Count
F	33.10	861
G	32.94	255
E	27.41	2437
D	22.14	4733
C	16.95	7356
B	12.04	10962
A	5.82	9333

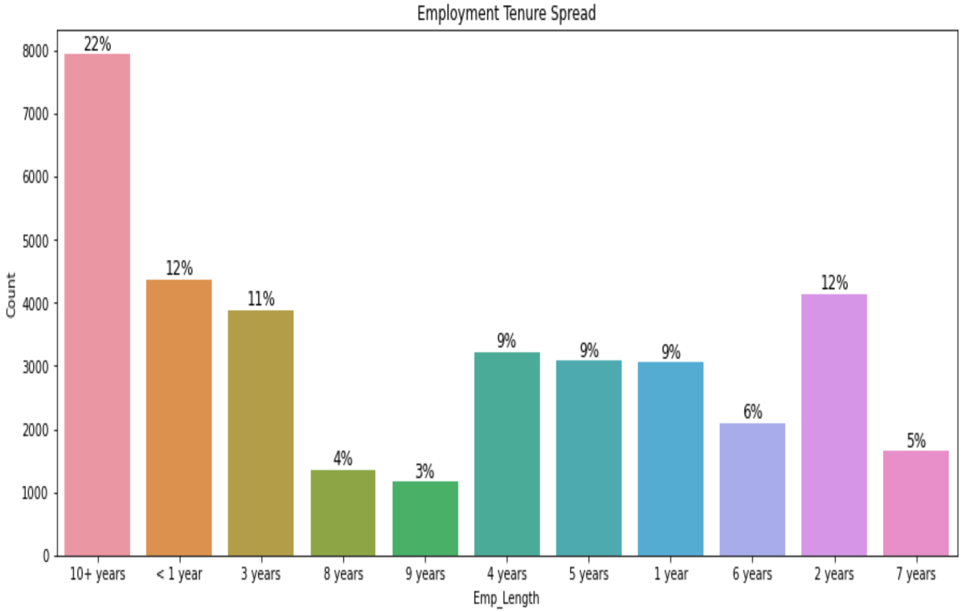




Observation:

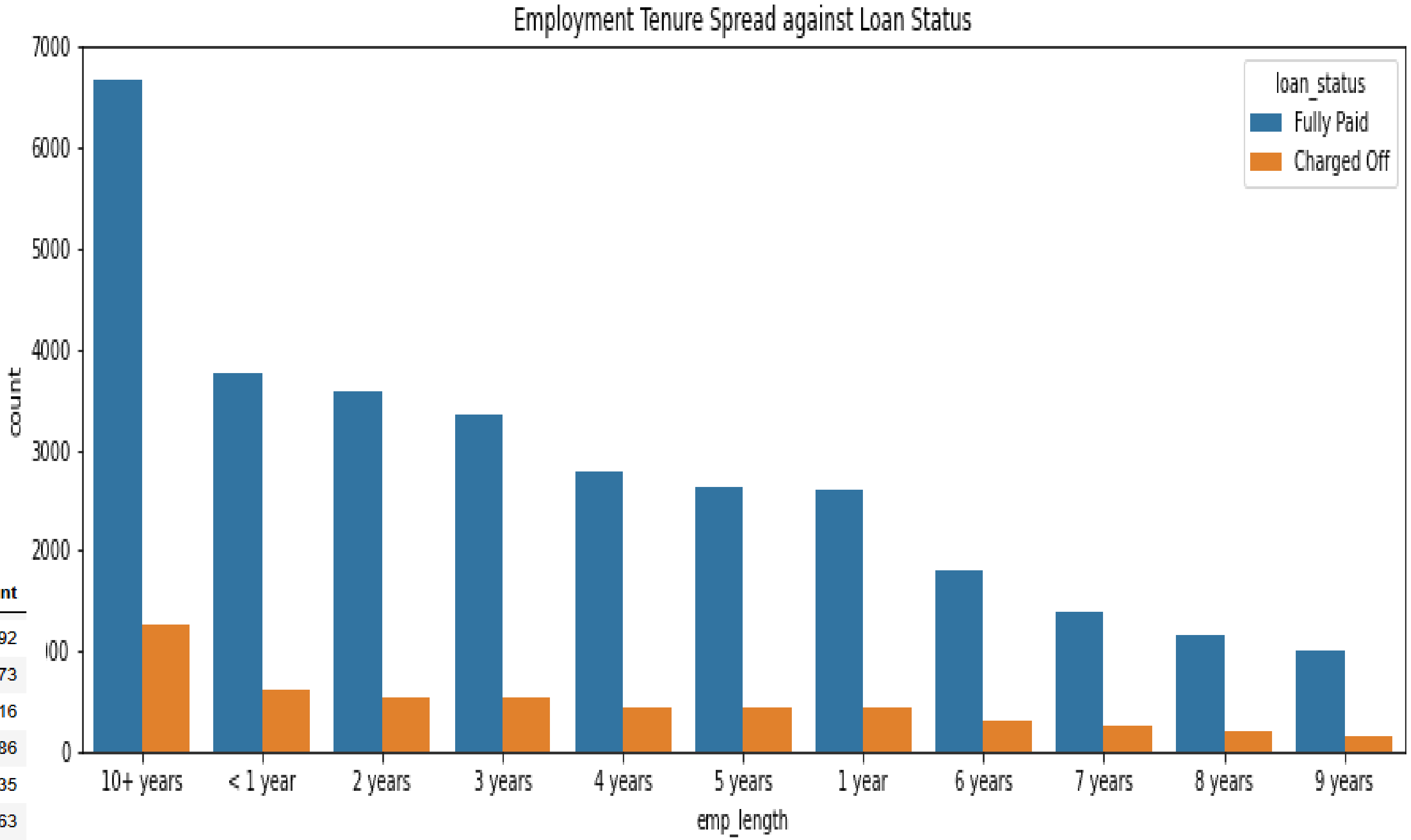
Grades A and B together account for 57% of the total loans. The Sub-Grades spread also confirms the same thing, that larger proportion of loans fall within higher grades or sub-grades. This indicates a good shift toward higher-quality borrowers, it suggests a healthier loan portfolio with potential for better financial performance, lower risk, and stronger investor confidence. This positive trend should be monitored and leveraged in strategic planning. We can also observe that the Default Rate has an inverse proportion against the grades i.e, Higher the Grade, Lower the Default Rate. This clearly indicates that the existing Grade System is beneficial to the business.

EMPLOYMENT TENURE ANALYSIS



emp_length	Charged off %	Record Count
10+ years	15.86	7931
7 years	15.22	1649
1 year	14.50	3056
5 years	14.46	3077
8 years	14.42	1359

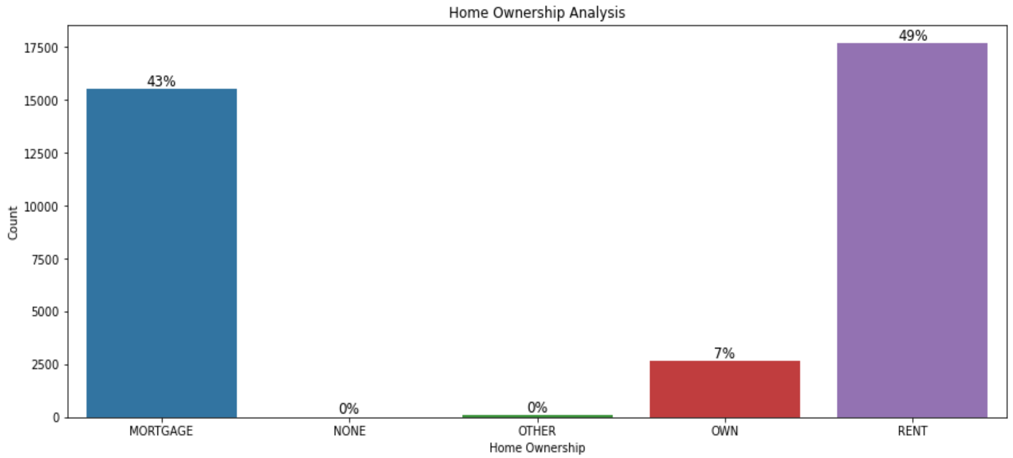
emp_length	Charged off %	Record Count
6 years	14.29	2092
< 1 year	14.20	4373
4 years	13.81	3216
3 years	13.77	3886
2 years	13.33	4135
9 years	13.07	1163

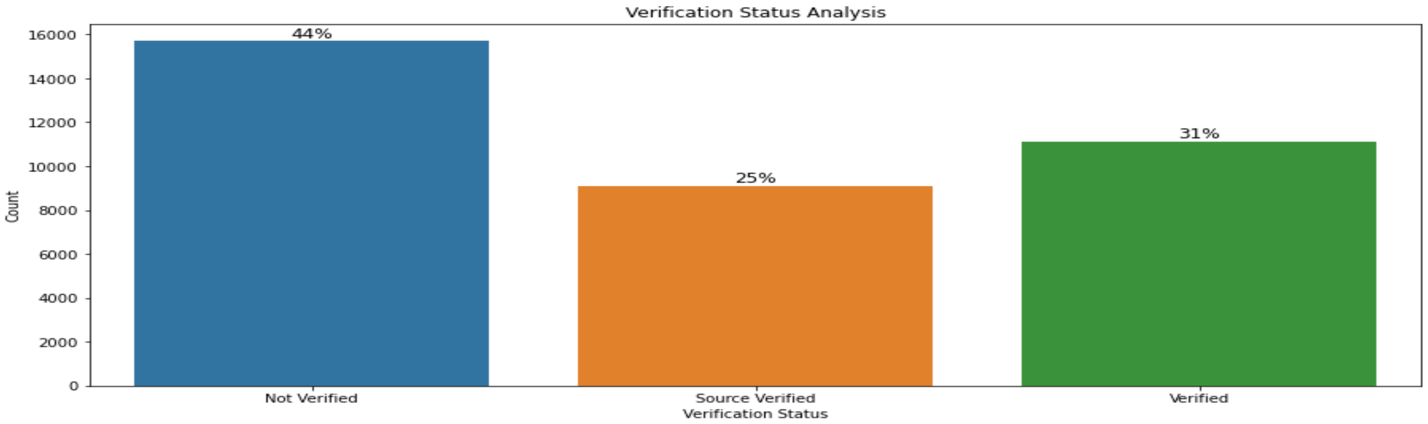


Observation:

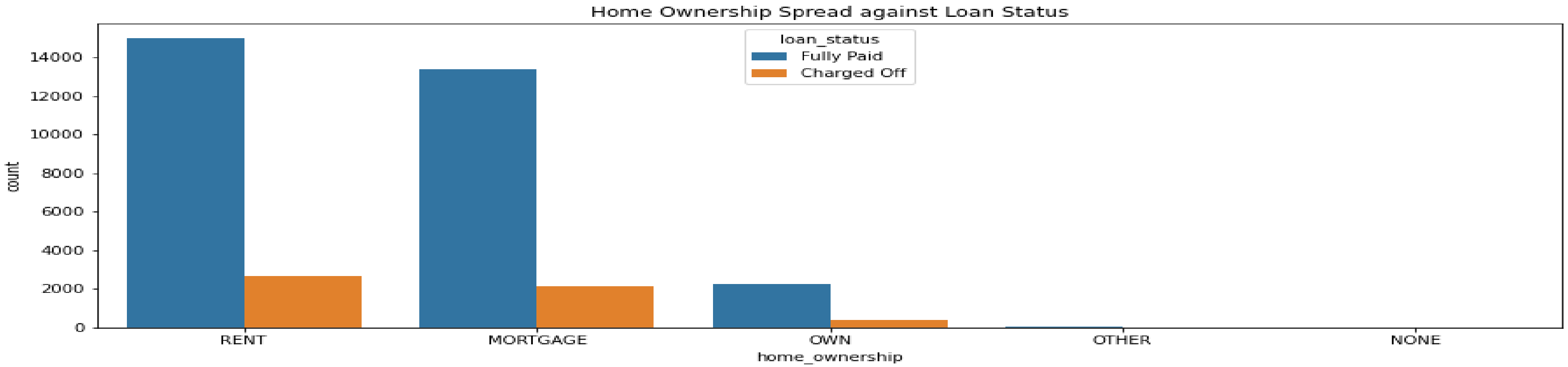
The highest percentage of loan applicants (22%) have been employed for ‘10+ years,’ indicating that a significant portion of applicants have stable, long-term employment. This clearly indicates that business should prioritize applicants with longer employment histories, as they are often seen as lower risk. However, there is still a notable percentage of applicants with shorter tenures, indicating a diverse applicant pool.

HOME OWNERSHIP & VERIFICATION STATUS ANALYSIS





home_ownership	Charged off %	Record Count
OTHER	18.75	96
RENT	15.12	17675
OWN	14.47	2661
MORTGAGE	13.67	15502

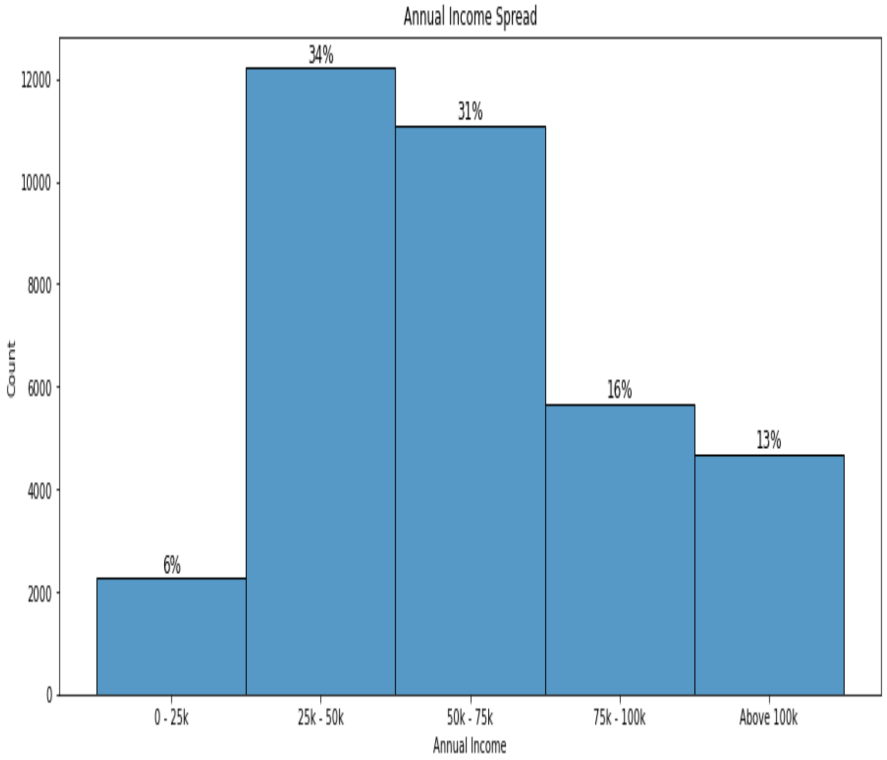


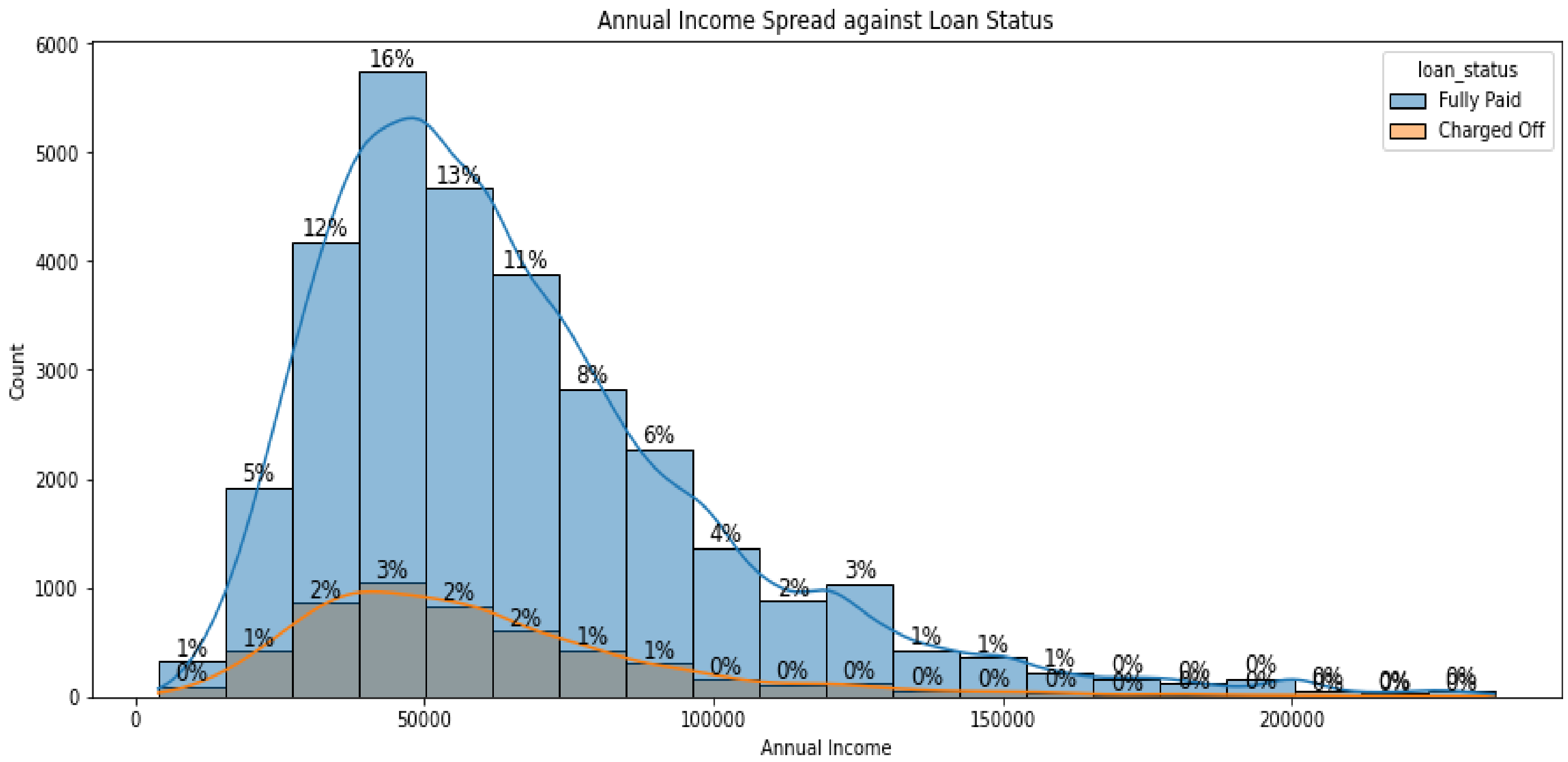
Observation:

- **Almost 92% of the applicants are either in a rental home or have a mortgage on their homes. This means that there is a high risk of loan repayment from these types of applicants. This can also be confirmed by analysing the spread against the loan status. There is a clear indication that the Default Rate is minimal in Home Owned Applicants when compared to the Applicants who do not own a house.**

- **The verification status reveals that 44% of borrowers are unverified, indicating potential risk. While 31% have verified status, suggesting creditworthiness, the low verification rates highlight opportunities for improved processes. Enhanced verification could reduce default rates, ensure compliance, and inform targeted lending strategies for better portfolio performance.**

ANNUAL INCOME SPREAD ANALYSIS

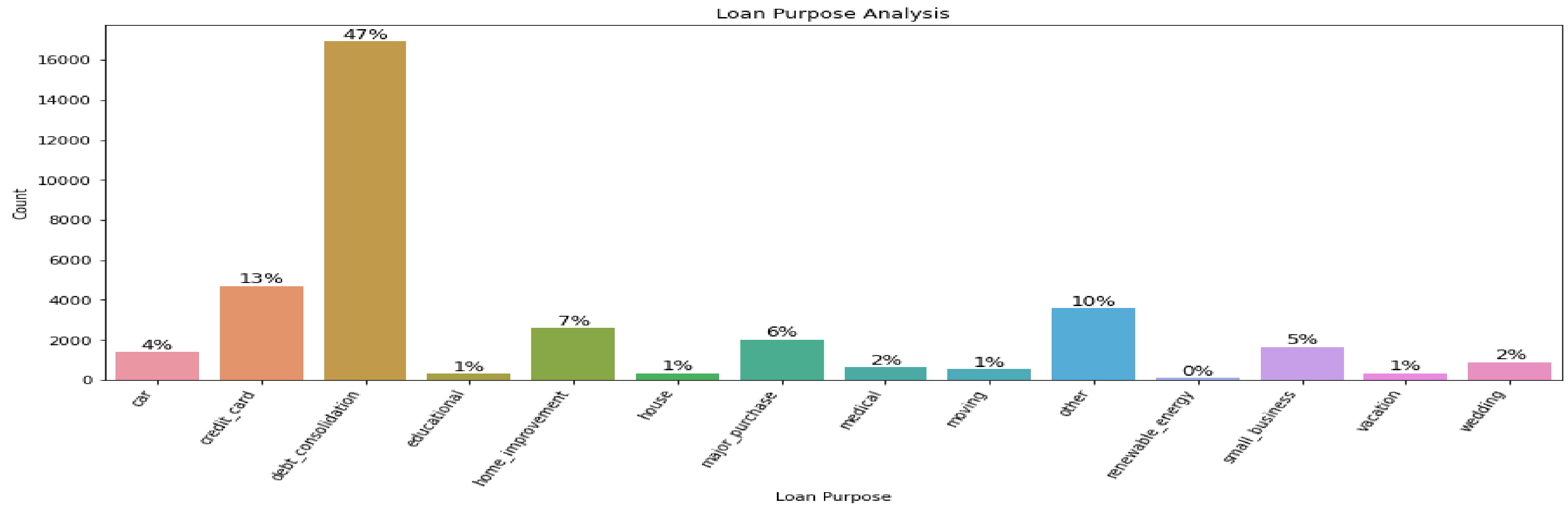


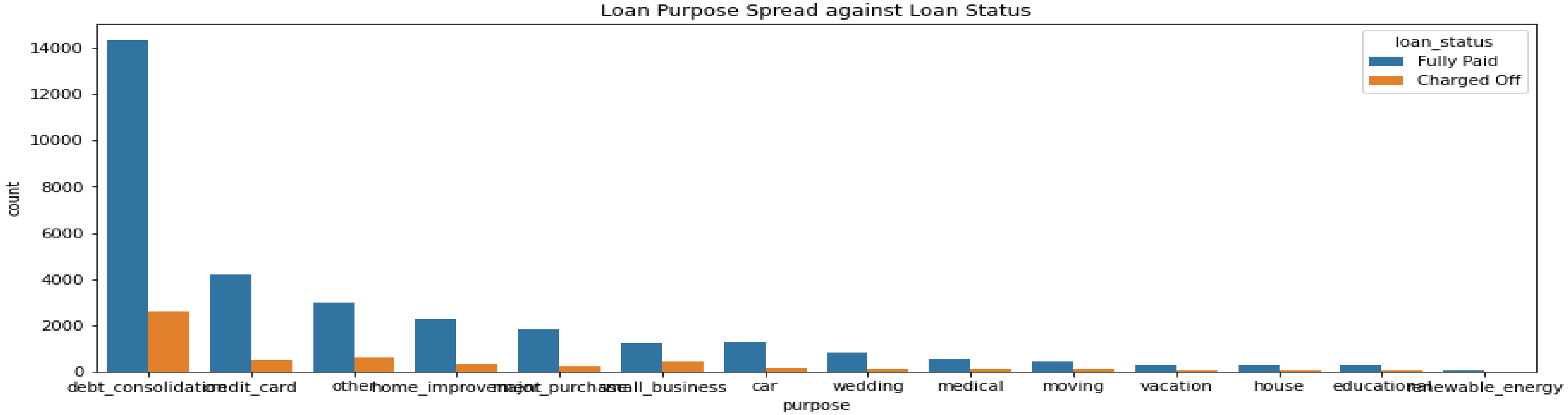


Observation:

- **The income distribution shows that 34% of borrowers earn 25k-50k, with 65% in the 25k-75k range, indicating a focus on middle-income earners. Only 6% earn 0-25k, suggesting higher risk.**
- **There is potential to tailor loan products for these demographics while engaging higher-income borrowers effectively.**
- **Analysing the Income Spread against the Loan Statuses, we can see that mostly fully paid loans are concentrated around lower annual incomes; higher incomes have fewer defaults (charged off).**
- **Currently loans are evenly distributed across all income levels.**
- **Applicants with an annual income of 50k or lesser are most likely to default and applicants with higher annual income are less likely to default.**

LOAN PURPOSE SPREAD ANALYSIS



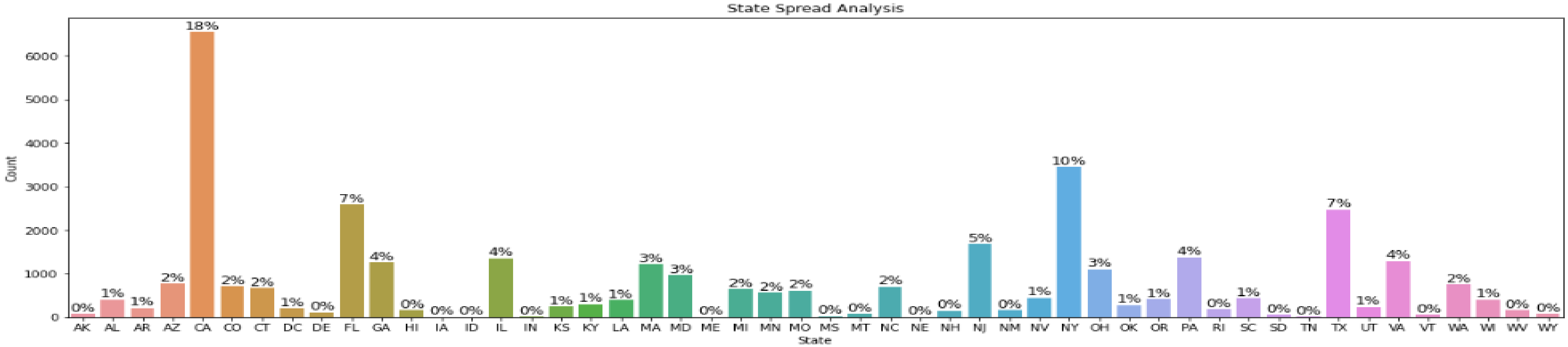


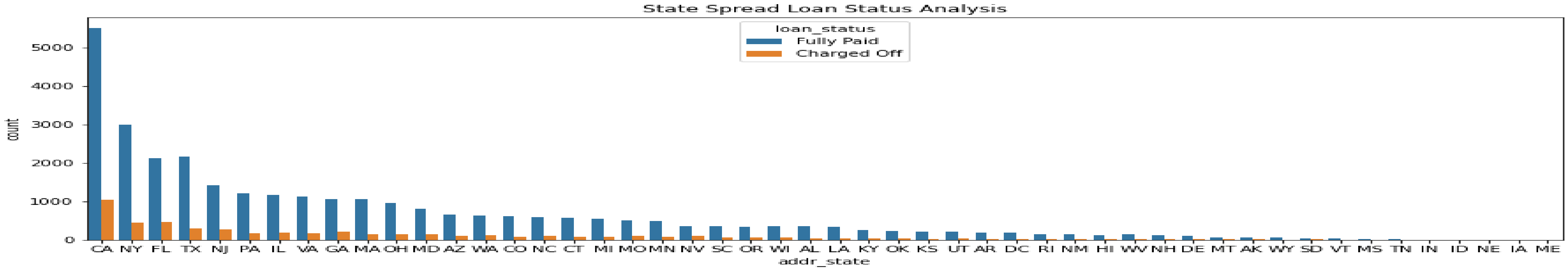
purpose	Charged off %	Record Count
small_business	27.06	1630
renewable_energy	20.69	87
house	17.01	335
educational	16.34	306
other	16.09	3579
medical	15.34	626
debt_consolidation	15.24	16912
moving	15.20	533
vacation	14.58	336
home_improvement	11.95	2603
car	10.89	1396
credit_card	10.44	4692
major_purchase	10.24	2011
wedding	10.10	891

Observation:

- The most popular loan purpose is "debt consolidation" with 47% contribution followed by the "credit cards" with 13% contribution.
- Together they form 60% of the loans, indicating that the loan applicants predominantly use the loans for debt management and handling personal expenses.
- On analysing the spread against the loan status, we can see that 27% of loans for small business are Charged off making them the most risky.

STATE SPREAD ANALYSIS

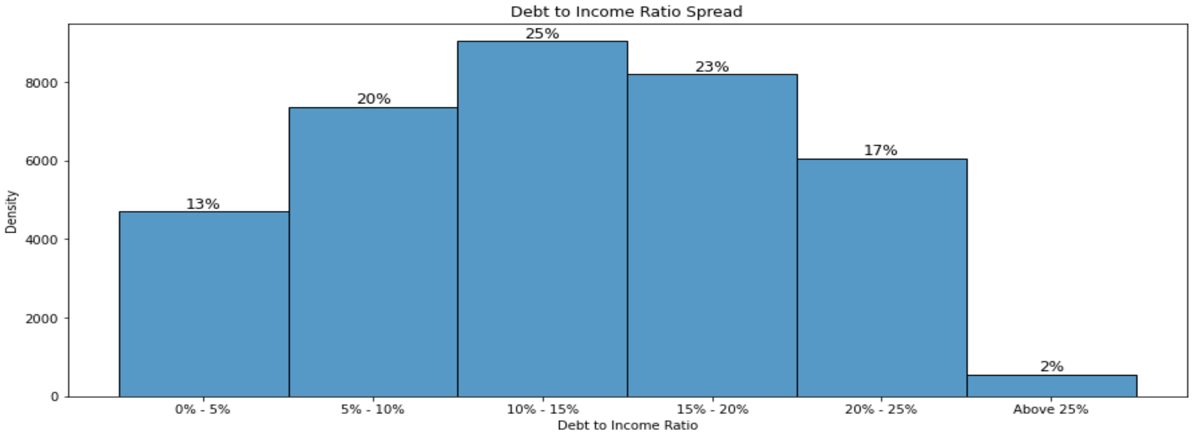


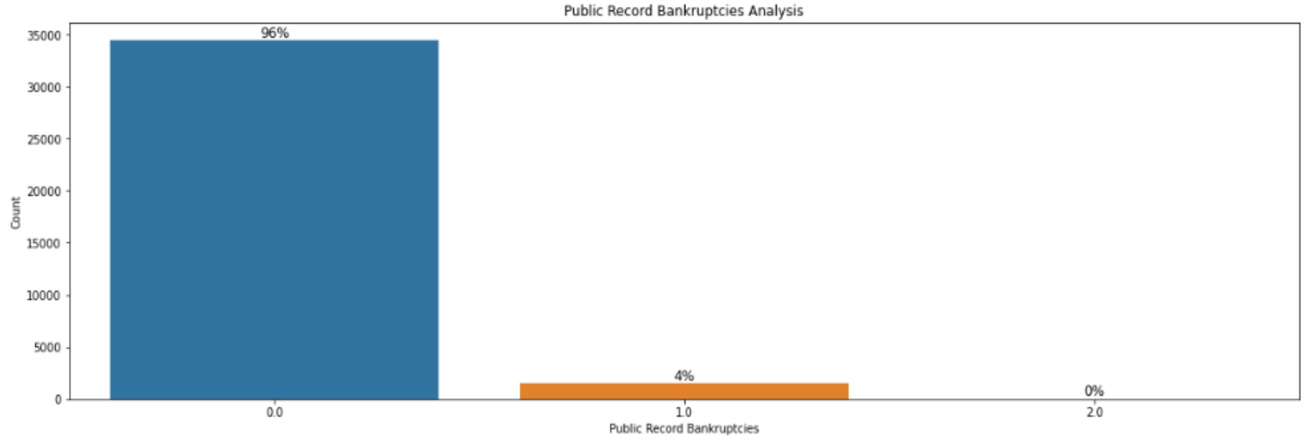


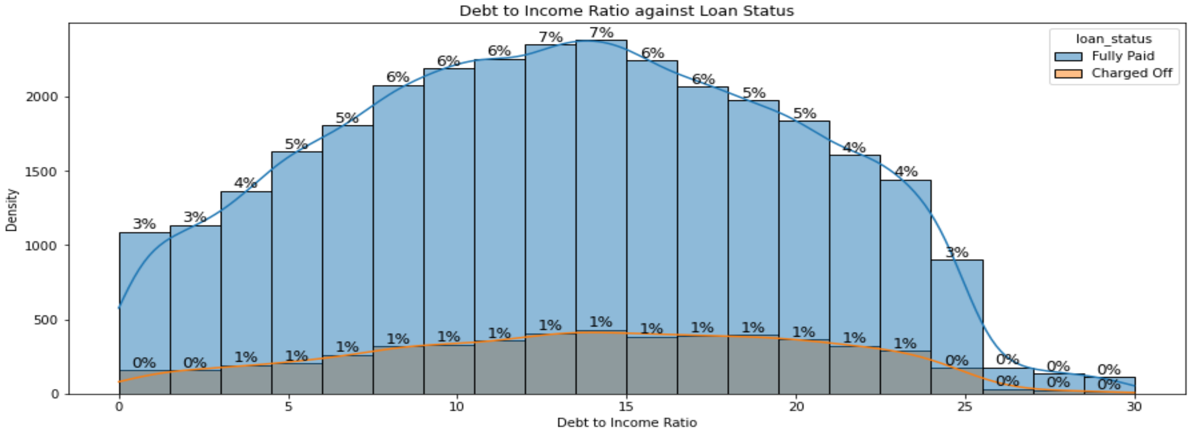
Observation:

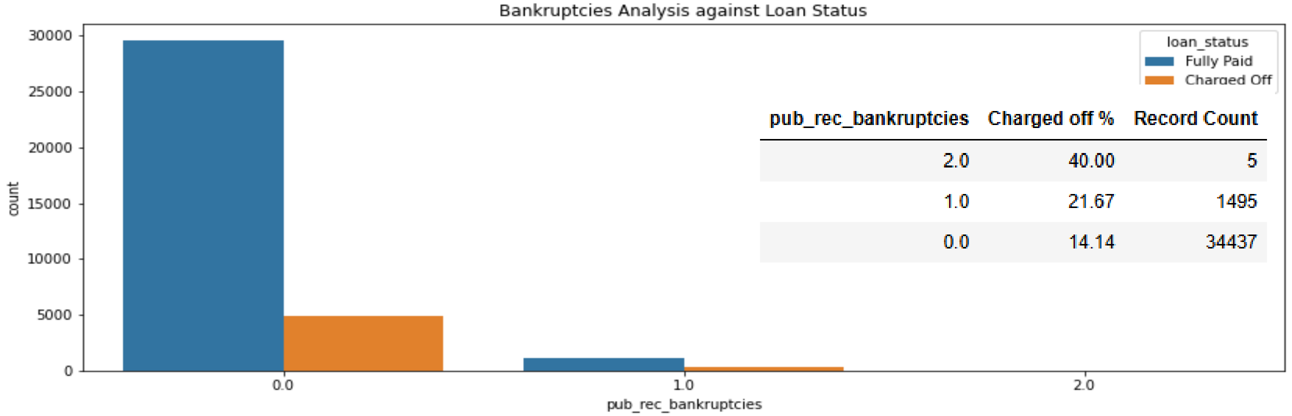
- We can observe that California has the highest number of loans with 18%, followed by New York at 10% and Florida at 7%.
- Most states have a higher proportion of Fully Paid loans compared to Charged Off loans.
- However, there are notable exceptions, such as Nevada (60% Default Rate) and Arkansas (21% Default Rate), which have a relatively high proportion of Charged Off loans.
- This suggests that while some states may have higher overall lending activity, the risk of loan default varies across different regions.

DEBT TO INCOME SPREAD & BANKRUPTCIES







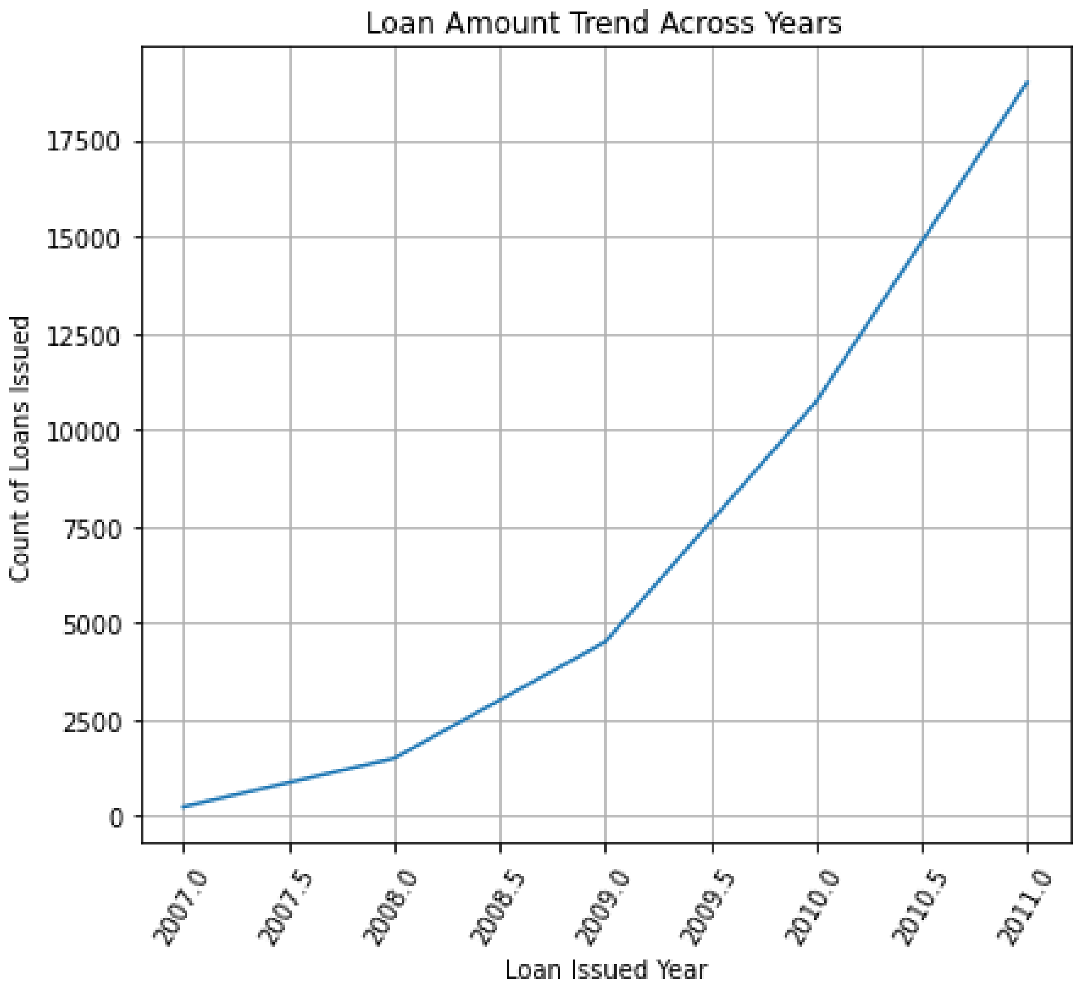


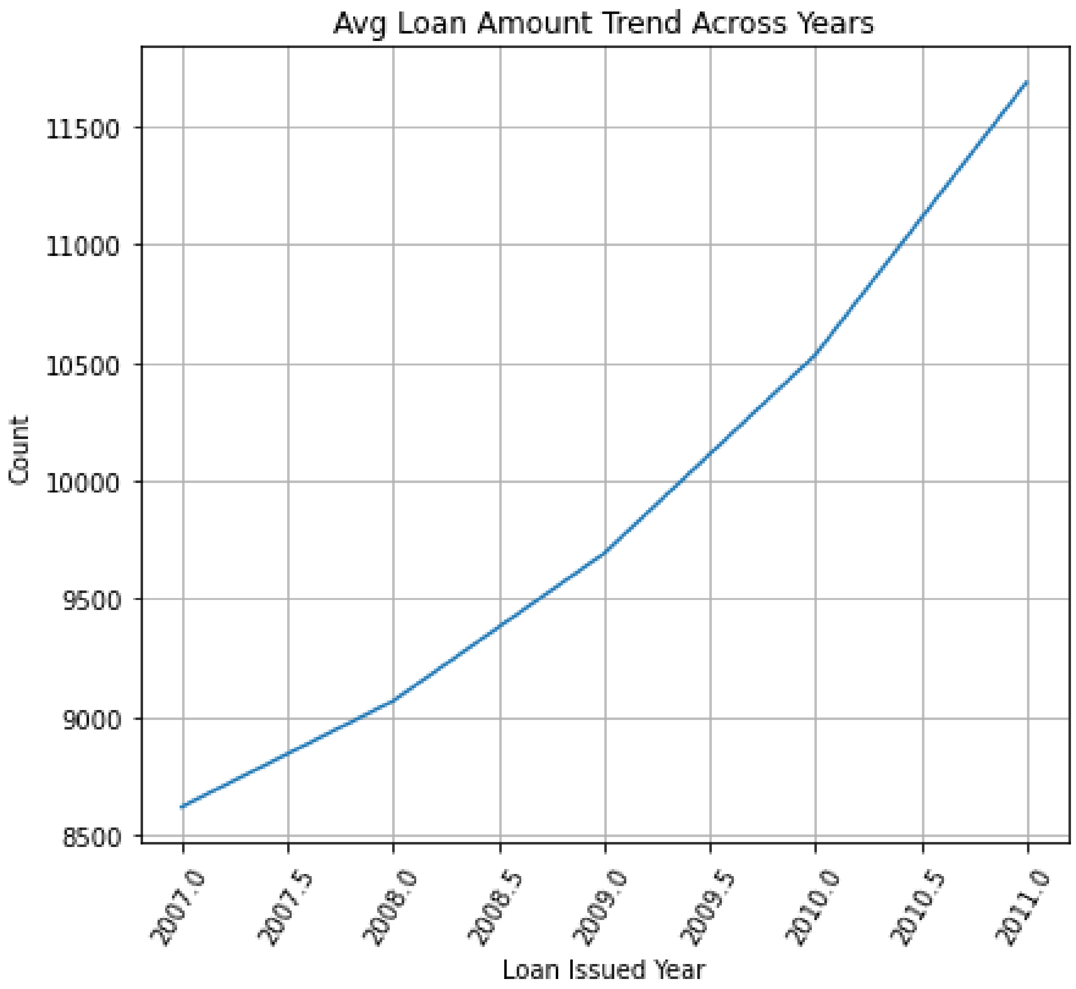
pub_rec_bankruptcies	Charged off %	Record Count
2.0	40.00	5
1.0	21.67	1495
0.0	14.14	34437

Observation:

- **Most loans are fully paid across all debt-to-income ratios, peaking around 17.5.**
- **Higher ratios correlate with more defaults (charged off), indicating increased risk.**
- **Lower ratios suggest better financial health and higher repayment likelihood. Currently loans are evenly distributed across all ratios.**
- **Almost 96% of the applicants do not have a bankruptcy record. This is very beneficial to the business and indicates that our loan approval processes is very strong.**
- **While analysing bankruptcies against the loan status, we see that the Default Rate is Directly Proportional and Positively Correlated with the Bankruptcies.**
- **That is, Higher the Bankruptcy, Higher the Chance of Default Rate.**

LOAN AMOUNT TREND

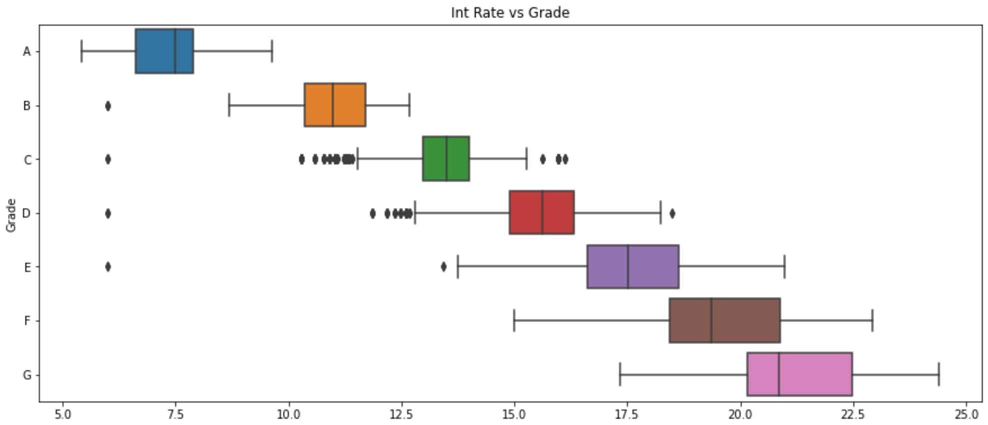


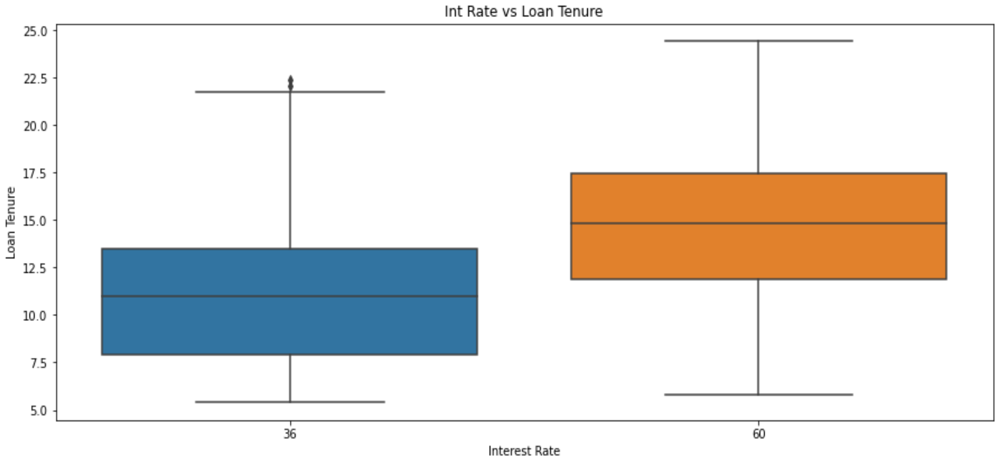


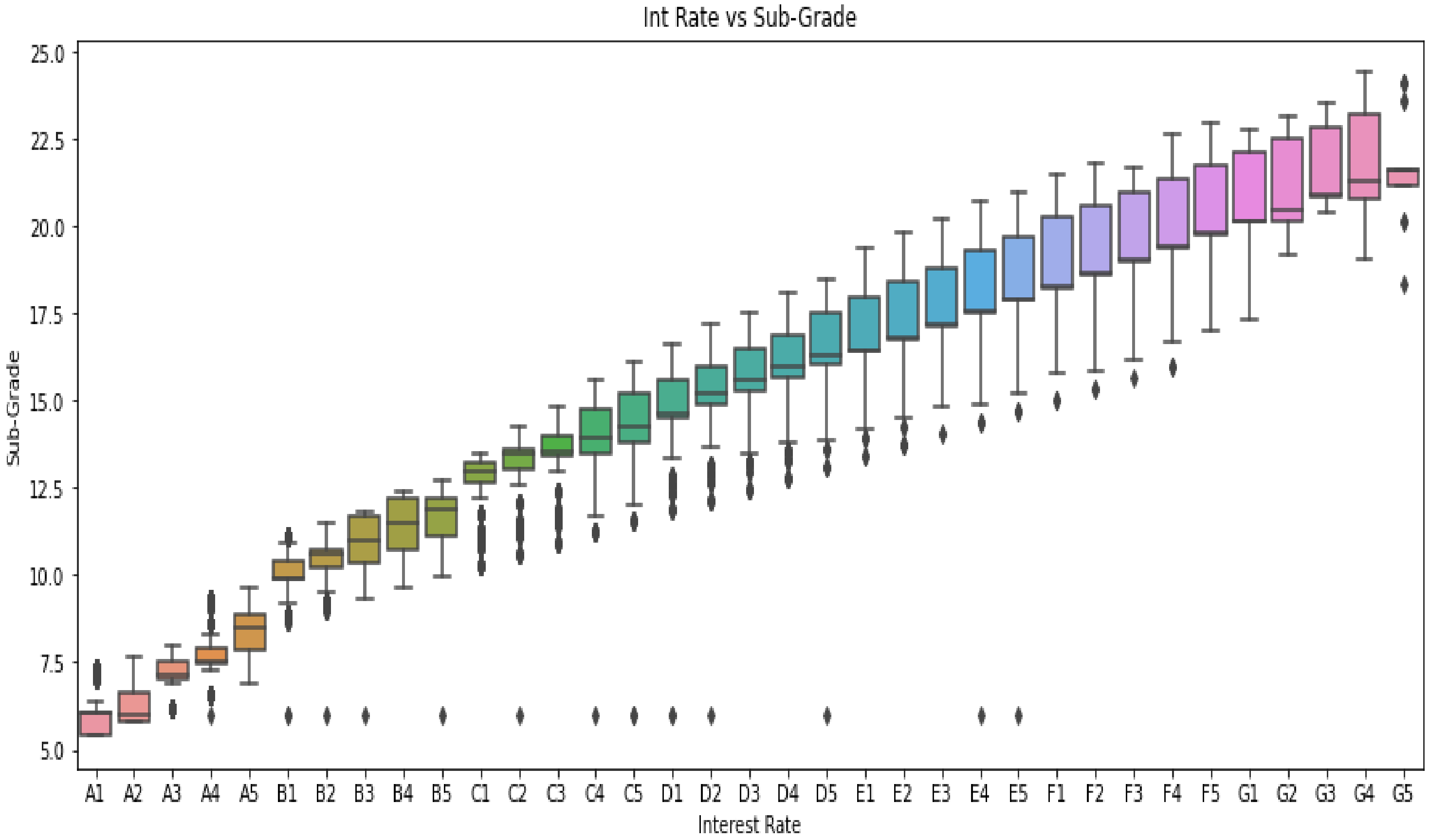
Observation:

- **We see an increasing trend across the years. This means that the business is expanding and it is a good positive indicator for growth.**

INTEREST RATE VS CATEGORICAL COLUMNS







Observation:

- **Interest Rates are directly proportional to the loan tenure. Earlier we saw that the Loan Amounts were also high for higher tenure. Hence we can understand that higher the amount, higher the tenure and hence higher the interest rate as well, which is actually beneficial to business since it will yield more revenue.**
- **Interest Rates are inversely proportional to the Grades or Sub-Grades. There is a negative correlation between them. This means that Higher the Grades, Lower the Interest Rate. Earlier, we saw that the Default Rates are quite high in the lower grades. Hence we say that the current interest rate slabs are perfectly aligned, since for Lower Grades, the Risk is higher and hence the Interest Rates are also higher ensuring loss compensation.**

OVERALL RECOMMENDATIONS TO BUSINESS

Based on the Univariate observations and Bivariate observations mentioned in each slides, below is the list of recommendations to the business related to Business Growth and Default Rate Reduction.

Business Growth

- **Target High-Quality Borrowers:** Focus marketing efforts on applicants with 10+ years of stable employment (22% of borrowers).
- **Optimize Loan Amounts:** Limit higher loan amounts (over 20k, which only account for 10% of loans) to reduce the risk of defaults.
- **Encourage Shorter Tenures:** Promote 36-month loans, which have a default rate of only 10% compared to 25% for 60-month loans.
- **Tailor Products for Income Segments:** Develop customized loan products for the 65% of borrowers earning between 25k and 75k.
- **Monitor Regional Risks:** Conduct assessments on states like Nevada (60% default rate) and Arkansas (21% default rate) to inform lending strategies.
- **Enhance Debt Management Options:** Offer financial education resources for borrowers focused on debt consolidation (47% of loans).
- **Leverage Positive Borrower Trends:** Utilize insights from higher grades (A and B, which make up 57% of total loans) to enhance portfolio quality.
- **Utilize Installment Flexibility:** Offer varied installment plans for loans under 25k, where 96% of applicants do not have bankruptcy records.
- **Track Employment Duration:** Regularly analyze employment tenure to focus on the 22% of borrowers employed for 10+ years.
- **Assess Market Timing:** Align marketing strategies with Q4, when 30% of loans are issued, the peak application period.

OVERALL RECOMMENDATIONS TO BUSINESS

Based on the Univariate observations and Bivariate observations mentioned in each slides, below is the list of recommendations to the business related to Business Growth and Default Rate Reduction.

Default Rate Reduction

- **Enhance Borrower Verification:** Implement robust verification for the 44% of borrowers currently unverified, increasing creditworthiness.
- **Adjust Interest Rate Strategies:** Review pricing models for the 75% of loans with interest rates under 15% to encourage early repayments.
- **Incorporate Alternative Data:** Explore additional metrics beyond credit history, as variables like public record bankruptcies showed weak correlations.
- **Analyze Loan Purpose Impact:** Address the 27% default rate for small business loans, the highest among loan purposes, with targeted support.
- **Review Recovery Strategies:** Strengthen recovery processes to improve the current recovery rate of 56% for charged-off loans, focusing on enhancing collection efforts.


THANK
YOU