LENDING CLUB CASE STUDY

Group Members:

1. **Saksham Gupta**
2. **Sreehari G Varma**

The Problem

You work for a consumer finance company **Lending Club** which specializes in lending various types of loans to urban customers. 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. When the company receives a loan application, the company must **decide for loan approval based on the applicant's profile.** When a person applies for a loan, there are two types of decisions that could be taken by the company:

1. **Loan accepted:** If the company approves the loan, there are 3 possible scenarios described below:

   Fully paid: Applicant has fully paid the loan (the principal and the interest rate)

   Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.

   Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

2. **Loan rejected:** The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset)

Analysis Approach

Data Cleaning

- Drop columns with null values ,all random values or single category value
- Convert values to proper int, float, date representations

Univariate analysis

- Check distributions and frequencies of various numerical and categorical variables
- Create derived variables

Segmented univariate analysis

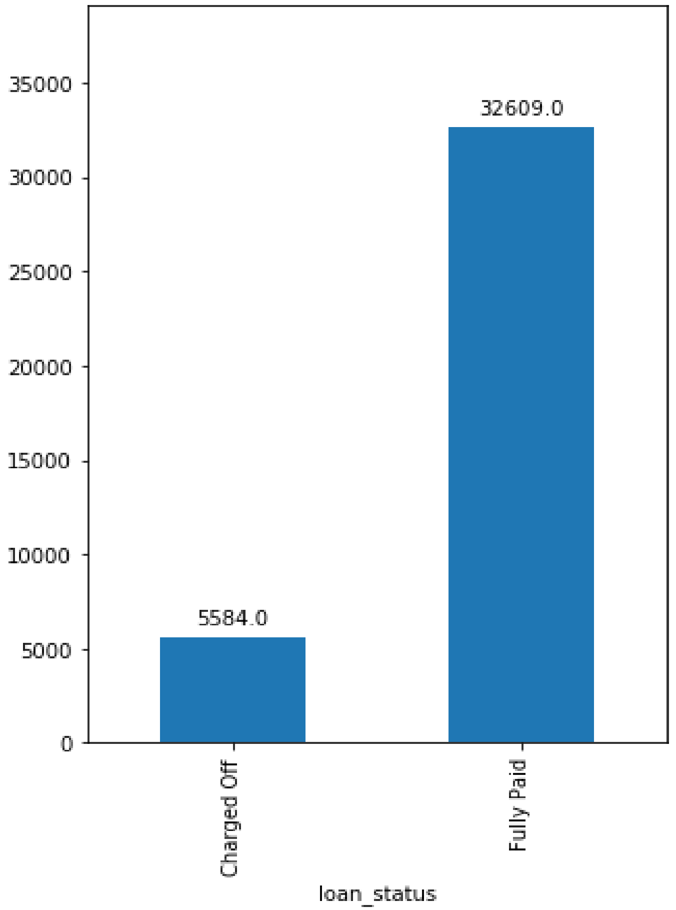
- Analyze variables against segments of other variables
- Create derived variables

Bivariate analysis

- Do correlation analysis
- Check how two variables affect each other or a third variable
- Analyze joint distributions

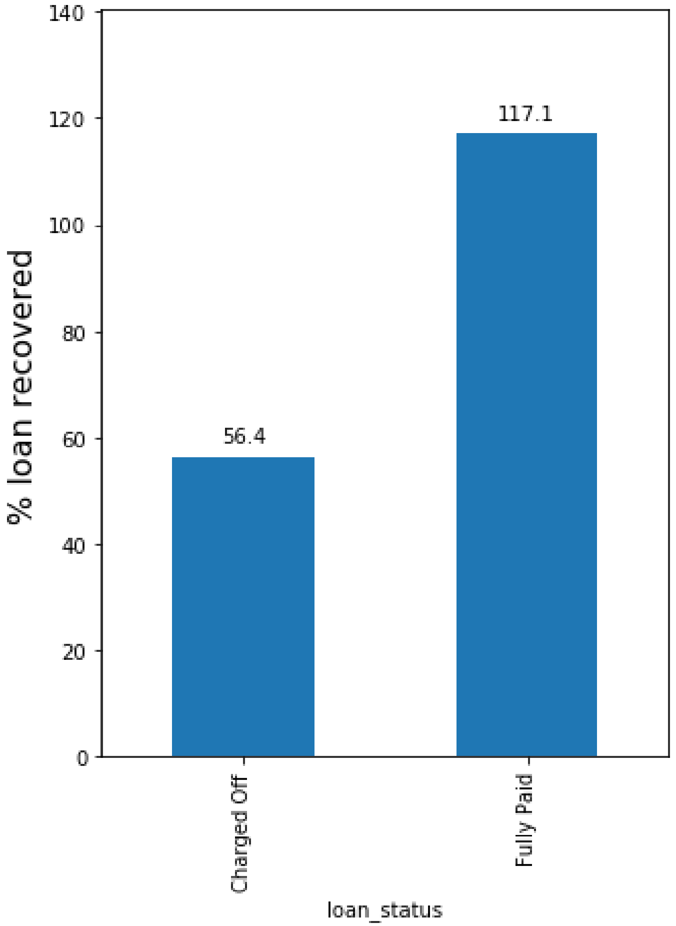
Findings

Analysis - Overall Loan Status



Total Loans

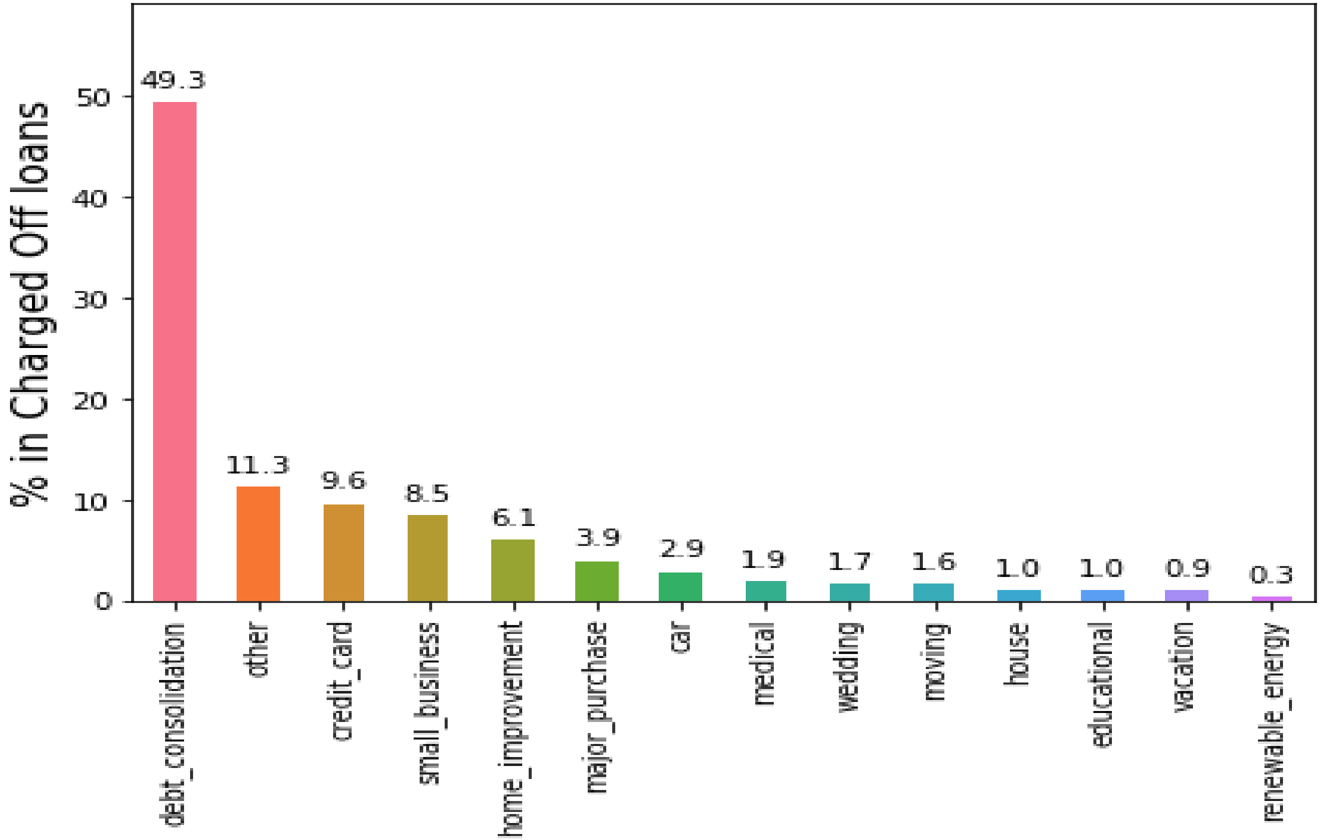
Approximately 15% of loans are defaulted



Total Money Earned

Lending Club only recovers 57% of the loan amount when loans are **defaulted**.

Analysis- Defaults by Loan-Purpose



Charged Off Rate - Top 3 Borrower's Purpose:

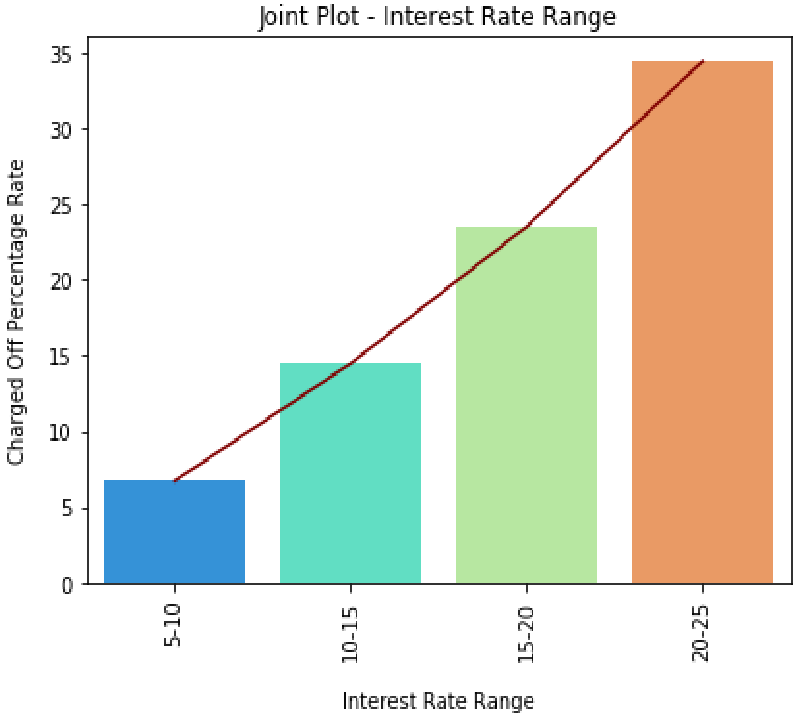
1. small business = 26.27%
2. renewable energy = 18.81%
3. educational = 17.03%

Analysis- Defaults by Employment Length

Top 4 employment length belonging to Charged Off category :

1. 1 year
2. 0 year (< 1 year)
3. 7 year
4. 10 Years (> 10 years)

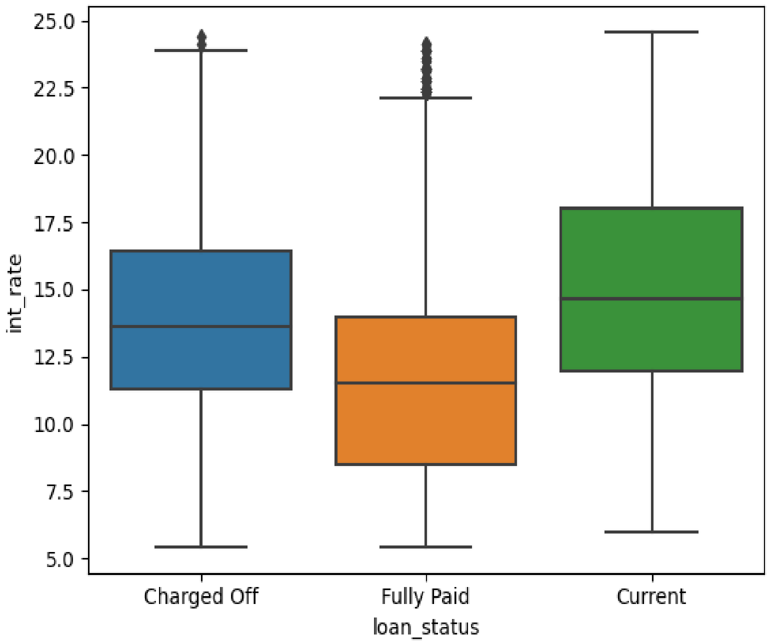
Analysis- Defaults by Interest Rate



- Charged Off Rate - Top 2 Interest Rate Range:

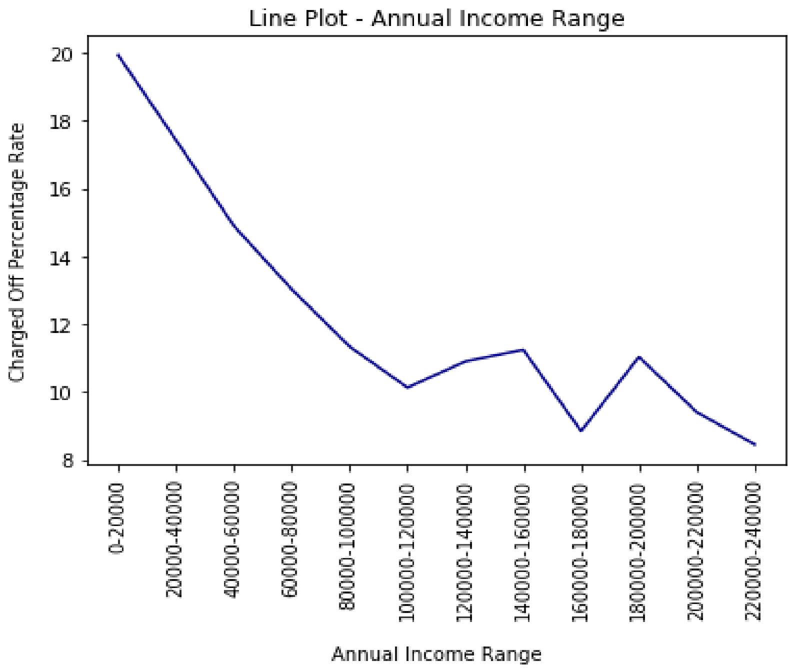
    1. 20-25 = 34.44%
    2. 15-10 = 23.51%
    3. 10-15 = 14.46%

- There is a clear indication that higher the 'interest rate’ , higher the chances to fall into charged off category



- The 25% quartile of 'charged-off' is almost same to the median of 'fully paid' status. This indicates a similar composition (~25%) to fall into charged off category if they have higher interest rates.

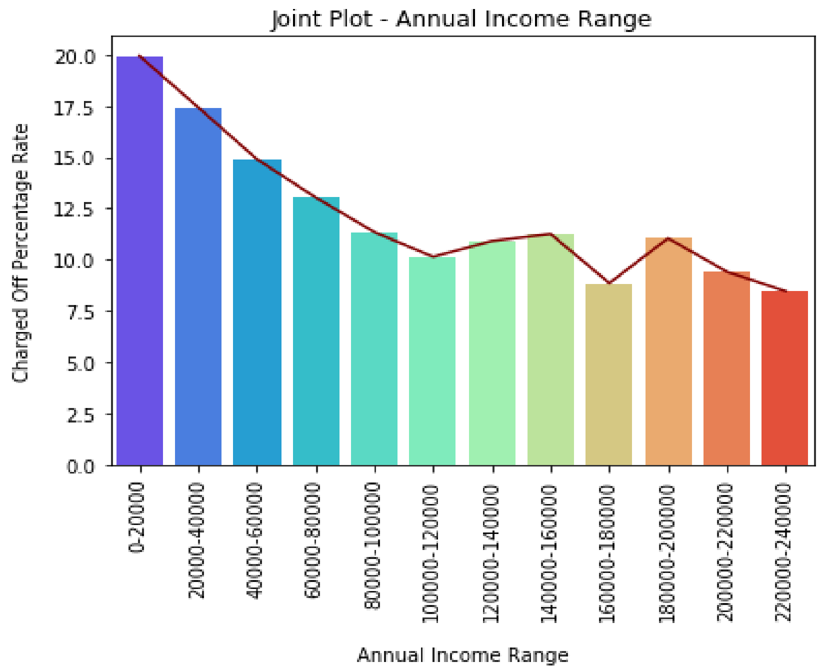
Analysis- Defaults by Annual Income



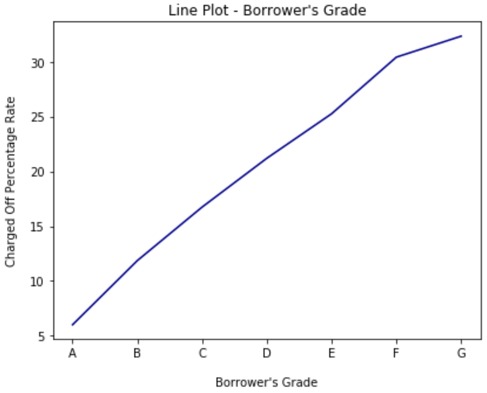
Charged Off Rate - Top 3 Annual Income Range:

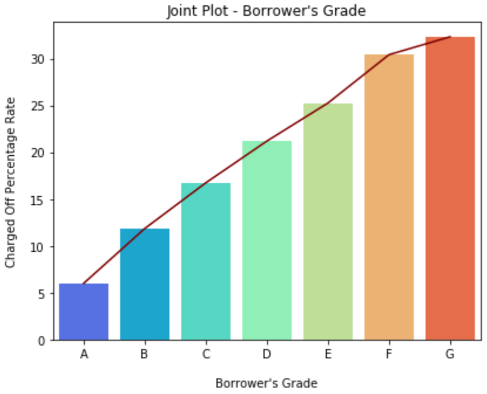
1. 0-20000 = 19.93%
2. 20000-40000 = 17.43%
3. 40000-60000 = 14.90%

So, as the Annual Income decreases, Charge Off Rate Increases



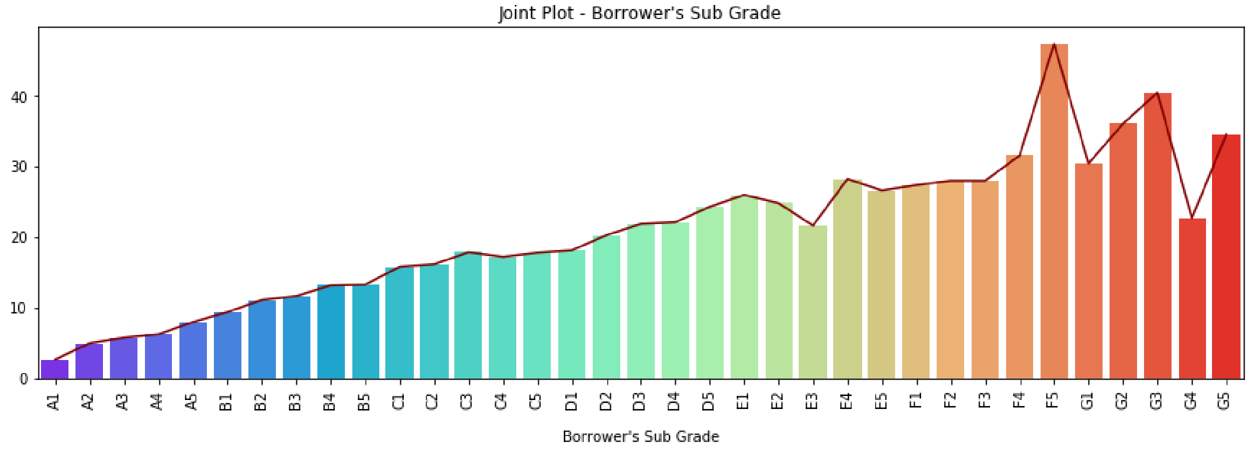
Analysis- Defaults by Grade & Sub-grade





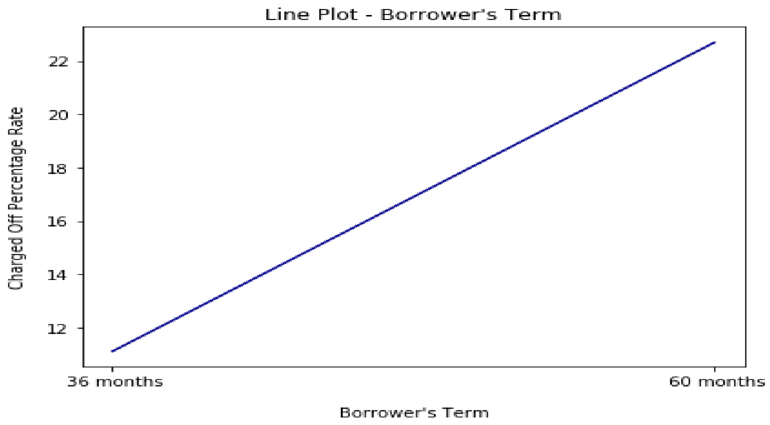
Charged Off Rate - Grade:
As the Grade increases,
Charged Off Rate increases.

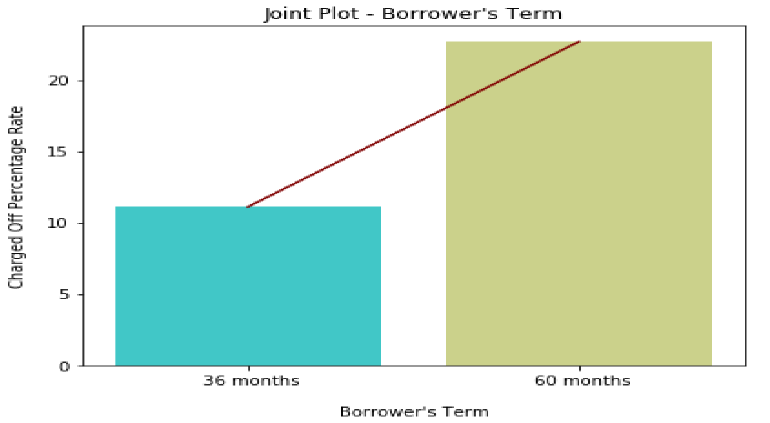
Top Order:
1. G
2. F
3. E
4. D
5. C
6. B
7. A



Charged Off Rate - Sub Grade:
As the Grade and the Sub Grades increases,
Charged Off Rate increases.

Analysis- Defaults by Term





Charged Off Rate - Borrower's Top Loan Term:

1. 60 months = 22.70%
2. 30 months = 11.10%

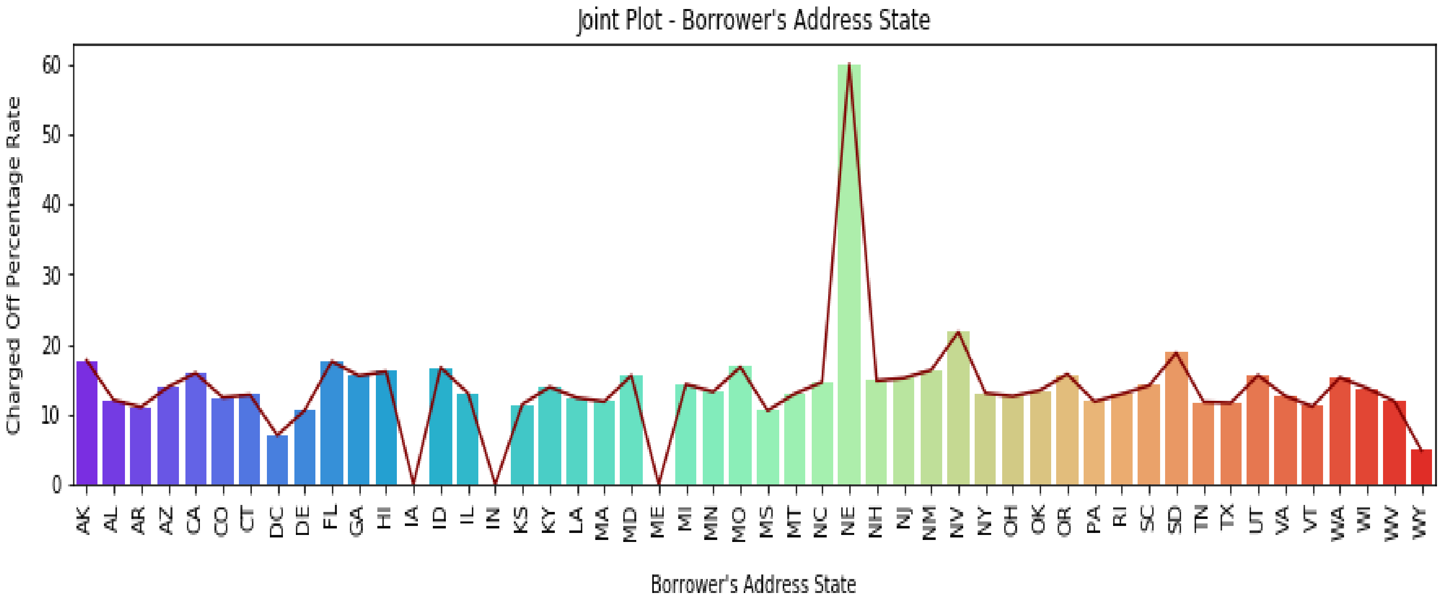
Which means 60 Months are more prone to defaults and there will be twice as chance of default than 36 months time period loan.

Analysis- Defaults by Address State

Charged Off Rate - Top 5 Borrower's Address State:

1. NE (Nebraska) = 60.00%
2. NV (Nevada) = 21.75%
3. SD (South Dakota) = 18.75%
4. AK (Alaska) = 17.72% 5. FL (Florida) = 17.54%

*NE (Nebraska) has a total of 5 loans out of which 3 were charged-off.



Conclusion:

Based on our Analysis, we produced following recommendations for Investors :

- Customers with long term loans tend to default more, recommendation is to encourage shorter term loans (36 month) as the interest rates are also higher in long term loans and customers fail to repay the loans.
- Higher LC Grade Loans have an impact on default rate as the interest rates also increase from grade A to grade G. Nearly 33% of all loans in Grades F and G see a default. Recommendation is to prefer Grade A, B, C over D, E, F, G. This is a Loan Attribute influencing tendency to default.
- Purpose of loan is also a strong indicator; recommendation is to scrutinize more for small business and renewable energy related loans. Wedding, cars, big purchases remains safe bet. Also one can provide small business loans at lower interest rates to reduce defaults. This is a consumer Attribute influencing tendency to default.
- Income source verification is not showing impact on tendency to default, recommendation is to not rely on income source verification for approving loans. But customers with lower incomes are tending to default more.
- Loan amounts of 30% of annual income or higher see a high rate of default. As long as loan amount is less than 20% of annual income, defaults are low. This could be considered if other risk factors are higher to provide large loan amounts.
- Loan Default tendency driving variables are **Loan Purpose**, **Emp length**, **Grade**, **Interest rate**, **Annual Income**, **Term of Loan**.

Other General Observations:

- The percentage of Charged Off loans is marked higher when the borrower has a prior record of bankruptcy.
- Higher the ‘dti’ higher the chances of loan being Charged Off
- Charged Off rate increases for people with less employment length.