Loan Case Study

Data Spark:

1.Vinay Dharra

2.Vijay Krishnan

3. Shubham Pali

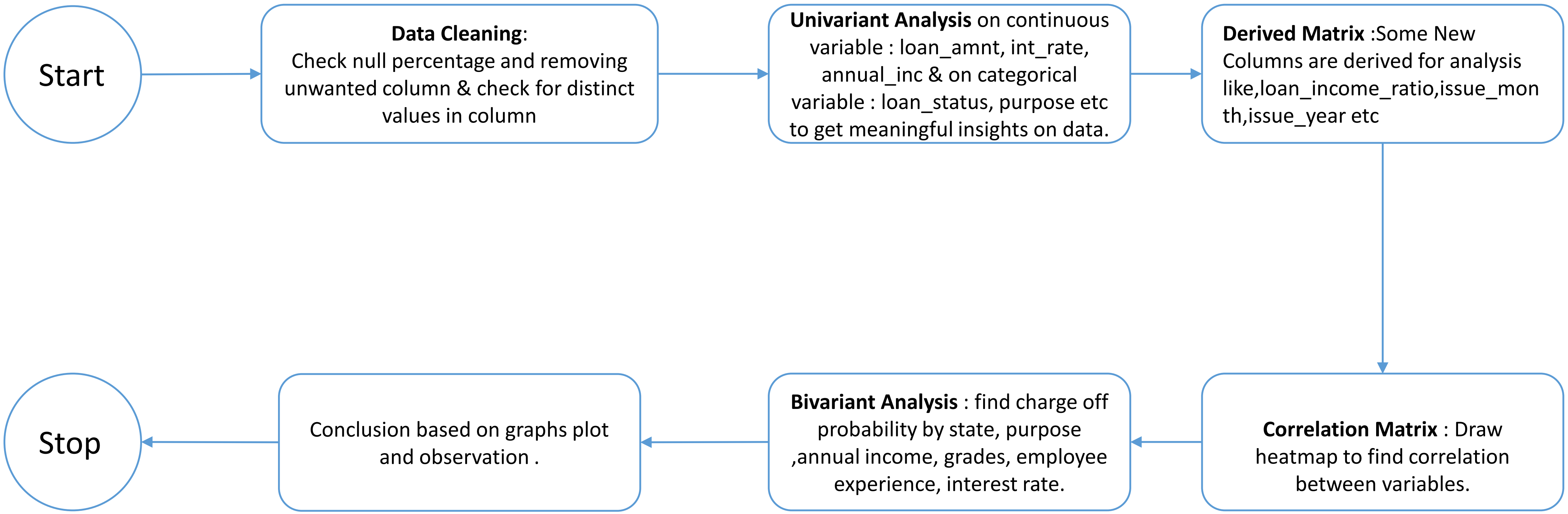
4. VENKATA NITHEESH IMMADI

<Abstract>

Based on detailed analysis we performed on dataset provided, we have concluded following results

- NV, NM , SD and NJ areas have high rates of Charged Off Loans. Hence it is prudent to discover more why default is happening in these areas and company must take care of this in future disbursals
- The small business loans, renewable energy loans and educational loans had the higher risk of loan defaults
- The loan applicants belonging to grades E to F defaulted on loans
- The applicants belonging to the annual income group which is less than 25000 USD defaulted on loans.
- Most of the applicant's loan interest rates fall between 7% to 10% but the risk is also high due to higher charged off numbers.
- The company is not advised to give loans to any applicants who have shorter duration of service (less than a year) as part of previous employment record. Maybe the applicant has joined a new company or in probation carries a higher risk of loan defaults
- The probability of higher defaults lies in the 15000-20000 (funded amount) category. Hence more analysis can be done on for what purpose this amount is used and accordingly take the action as needed

<Problem solving methodology>



<Analysis>

- Business Understanding :

A consumer finance company which specialises in lending various types of loans to urban customers. 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

- Data Understanding:

Dataset contains the information about past loan applicants and whether they 'defaulted' or not.

- Univariant Analysis:

On univariant analysis on continuous and categorical variable we have observed following results.

- distribution of the loan amounts is in the range of 5000 - 2000 USD.
- distribution of the Interest Rate for the loans is around 10% - 15%
- majority of the applications applying for the loan earns around 40000 to 90000 annually.
- around 5210 applicants or 14% are charged off when they defaulted on the loan
- the majority of the loans has been taken was for debt consolidation
- It is observed that with respect to the loan status fully paid, around 41% has rented their homes and 35% percent has mortgaged their homes
- The desired loan tenure is 36 months where the loans are getting fully paid off

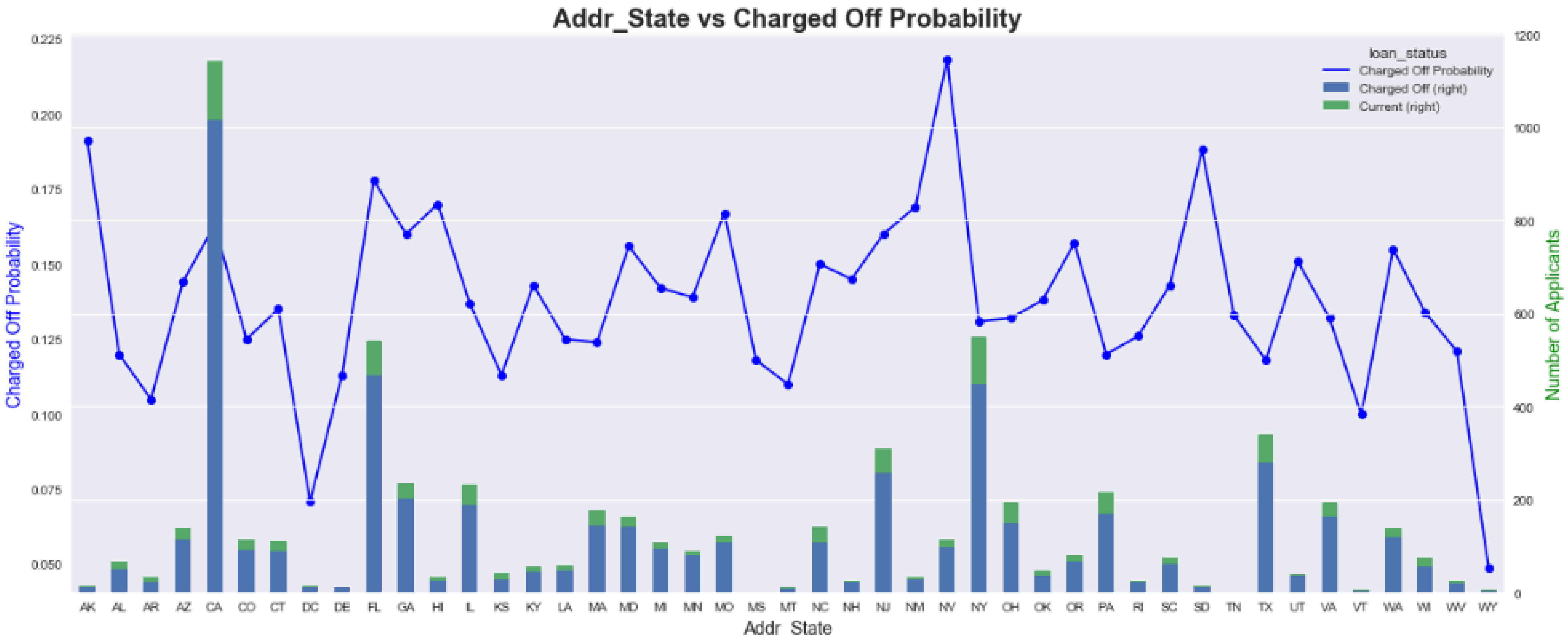
<Analysis>

- Bivariate/Multivariate Analysis:

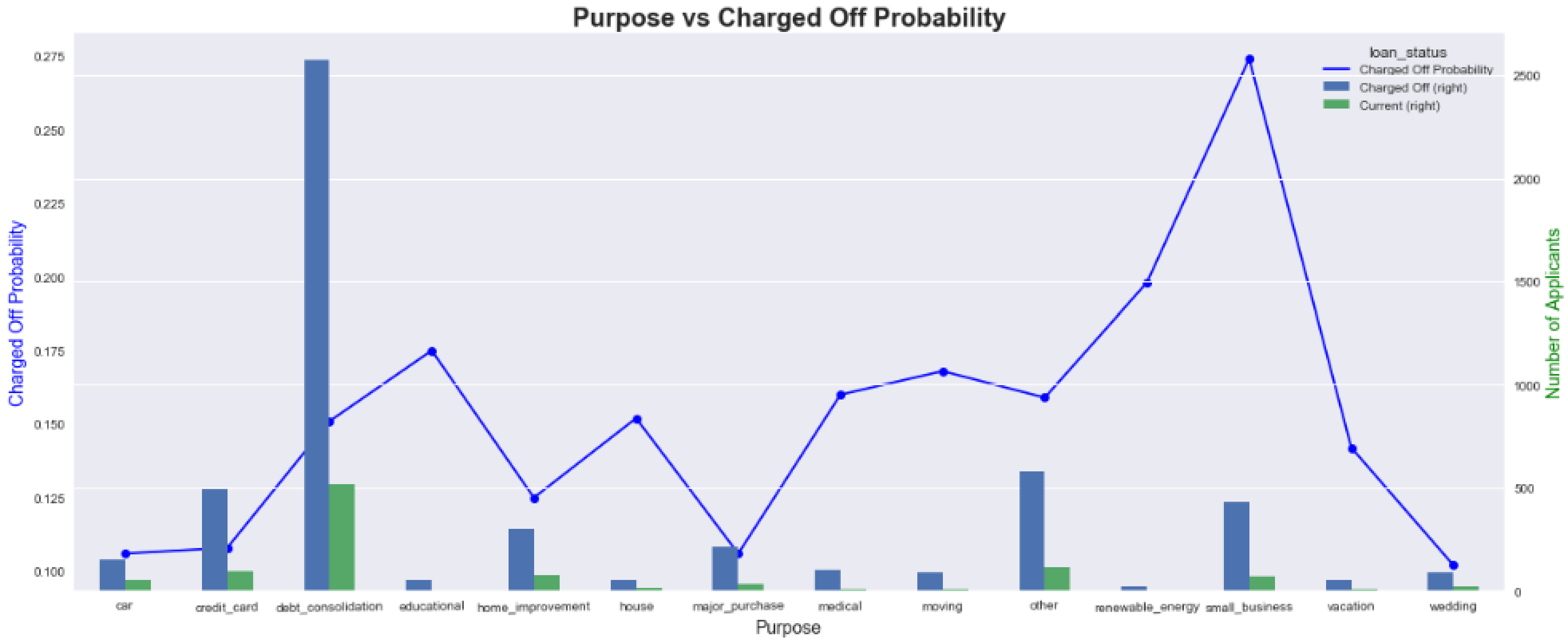
  On Multivariant analysis we have observed following results,

  - It is observed that NV, NM , SD and NJ areas have high rates of Charged Off Loans. Hence it is prudent to discover more why default is happening in these areas and company must take care of this in future disbursals
  - The small business loans, renewable energy loans and educational loans had the higher risk of loan defaults
  - The loan applicants belonging to grades E to F defaulted on loans
  - The applicants belonging to the annual income group which is less than 25000 USD defaulted on loans.
  - Most of the applicant's loan interest rates fall between 7% to 10% but the risk is also high due to higher charged off numbers.
  - The company is not advised to give loans to any applicants who have shorter duration of service (less than a year) as part of previous employment record. Maybe the applicant has joined a new company or in probation carries a higher risk of loan defaults
  - The probability of higher defaults lies in the 15000-20000 (funded amount) category. Hence more analysis can be done on for what purpose this amount is used and accordingly take the action as needed

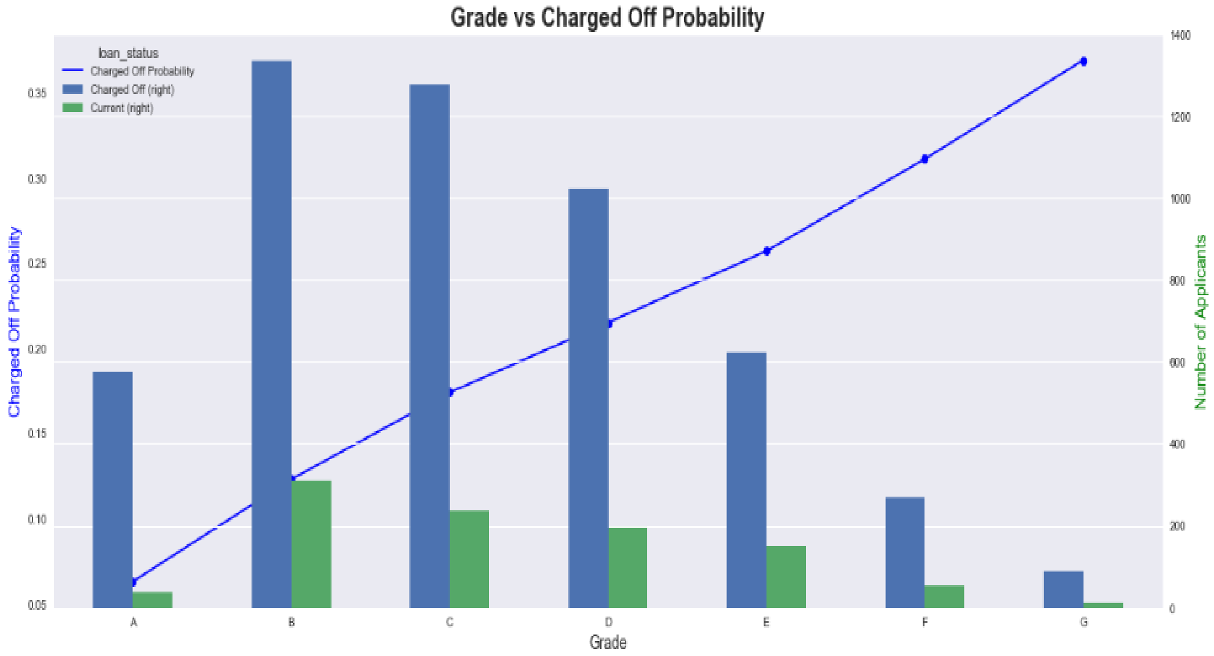
Charged Off Probability by State

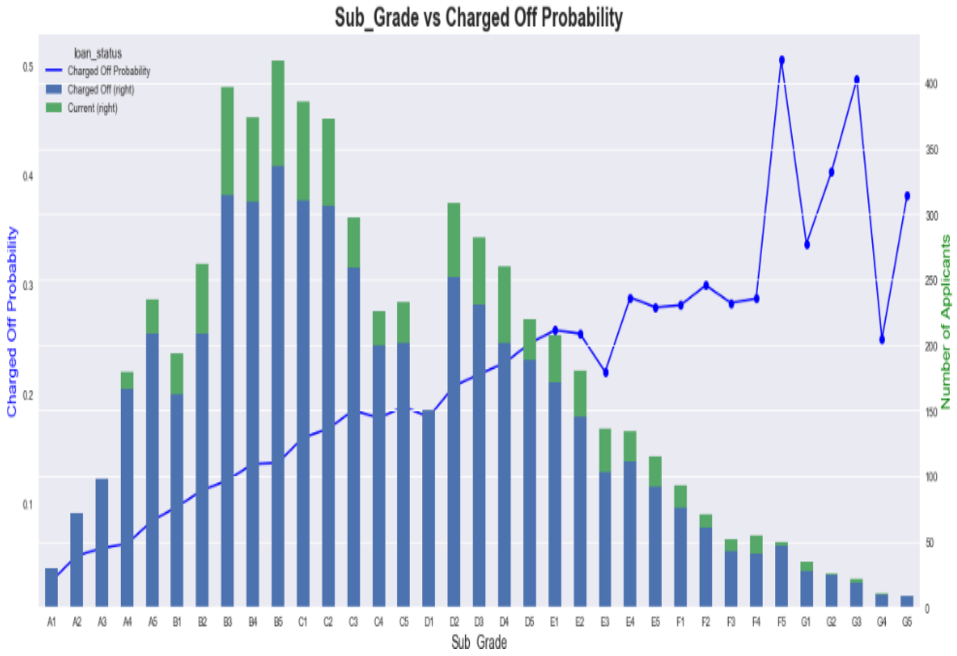


Charged Off Probability by Purpose

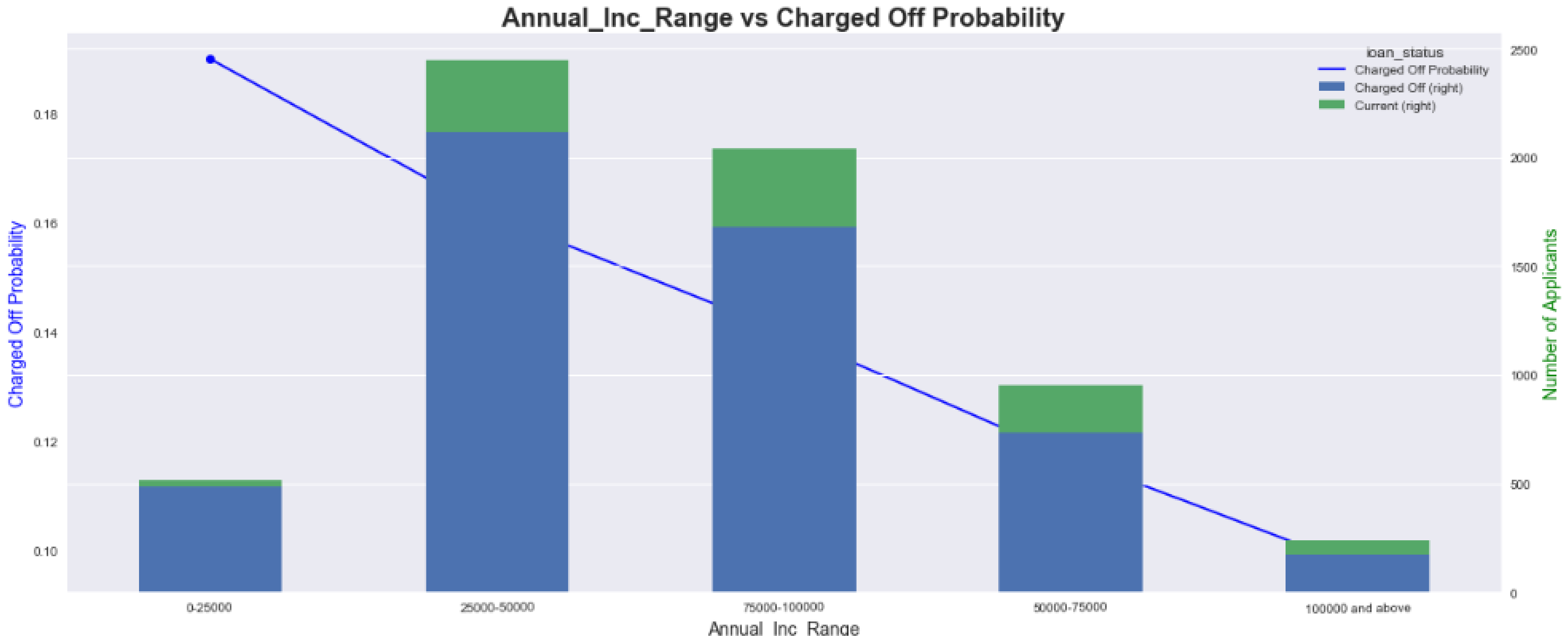


Charged Off Probability by Grade

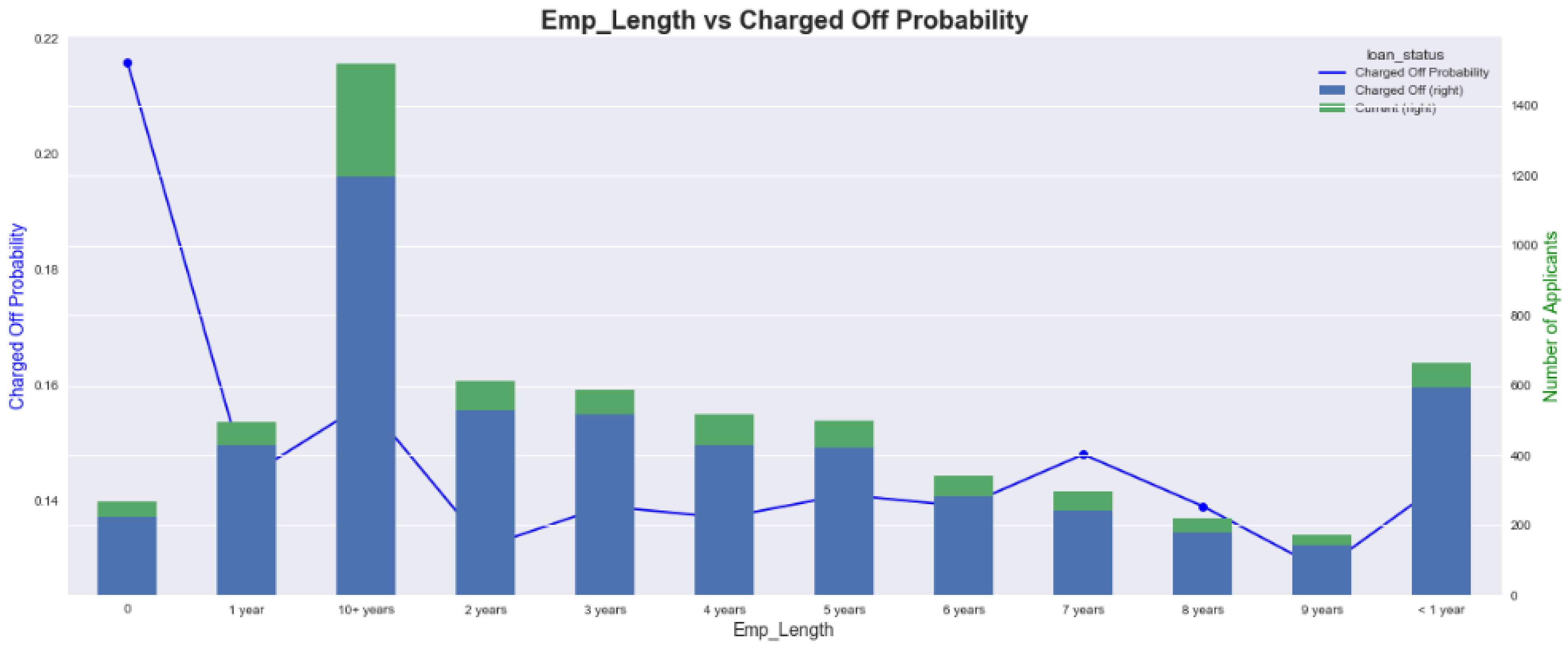




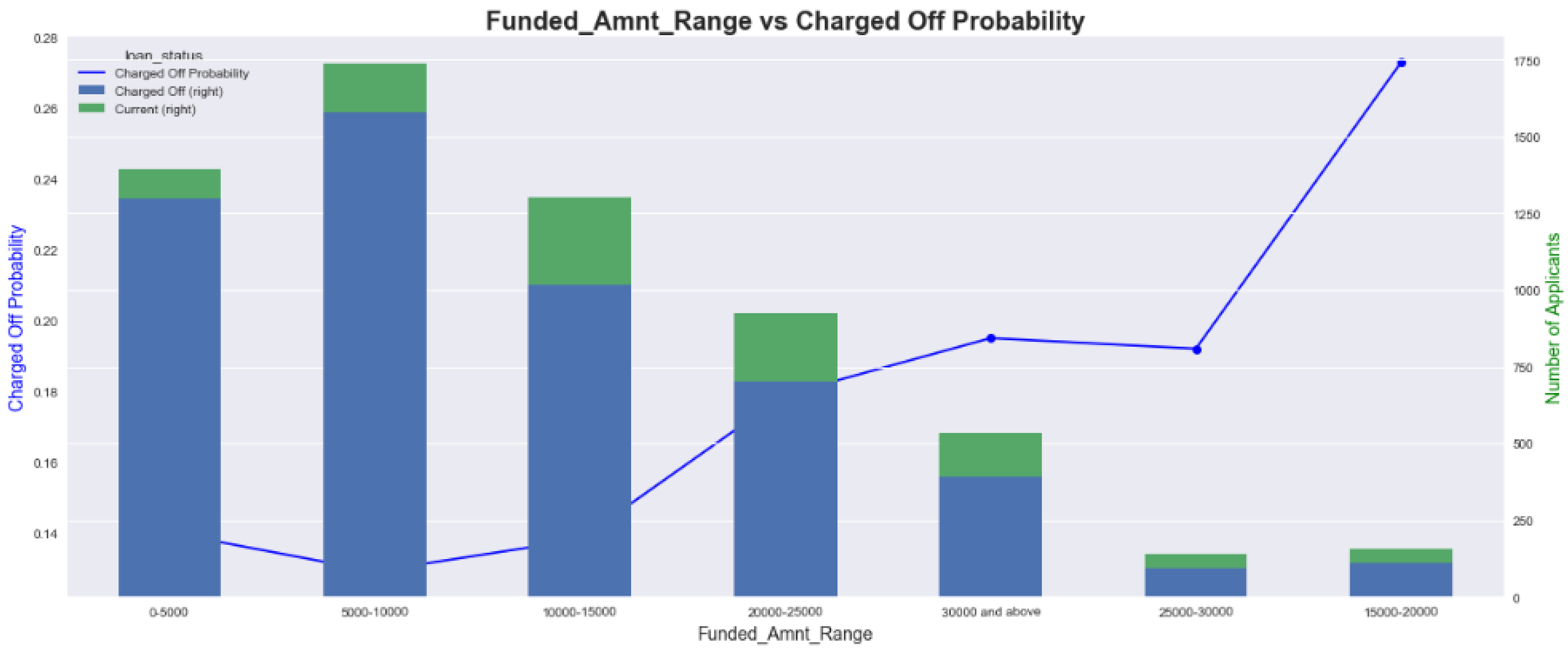
Charged Off Probability by Annual Income



Charged Off Probability by Employment Service



Charged Off Probability by Loan Funded Amount



<Conclusions>

Following driving factors are behind the loan defaults :

- Location or District or State where the Loan Applicant belongs to.
- Purpose of the Loan
- Annual Income of the Loan Applicant
- Loan Grade to which the Applicant is categorized
- Employee Term of Service i.e the Applicant Employment Record
- Loan Funded Amount