Credit Risk Analysis

Who defaults and why…

Comprehending Data

- Credit details of customers from around the country
- Included features like annual income, zip code, loan amount, status ranging to 111 properties *(Not same as the data dictionary)*
- Employment history is also specified for most of the customers to probably estimate job guarantee
- Lot of fields are nulls (or) 0 through-out.
- ZIP Code is scrubbed and cannot be used.
- he link https://www.lendingclub.com/foliofn/rateDetail.action helps in understanding more about the purpose of the grades & their impact to interest rates.

Purpose: Find the parameters that indicate that borrower is at a higher risk to default on the credit received.

Data Preparation

Cleansing

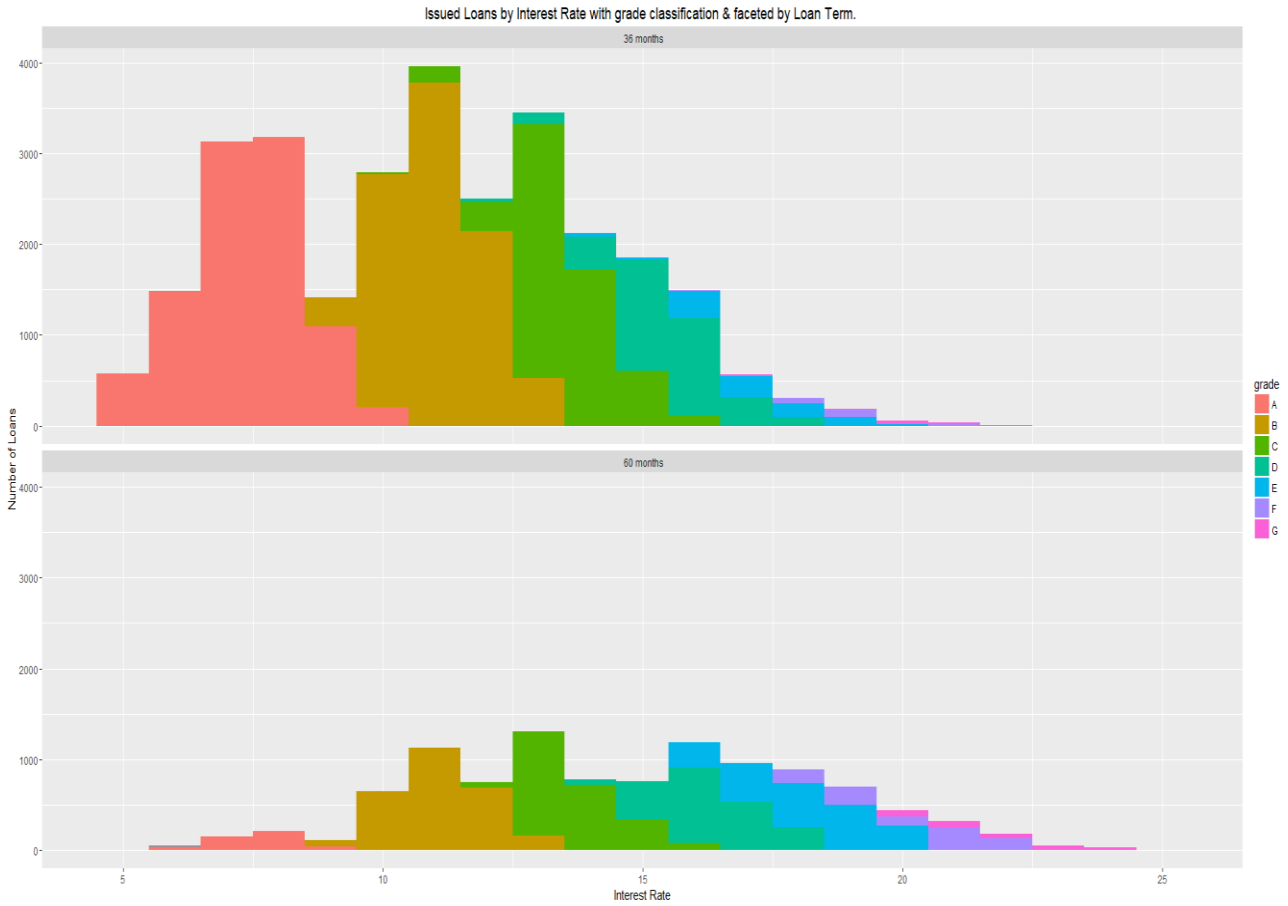
- Remove the fields that contain only NA Values.
- Remove fields that contain only 0s (or) only one value in all rows (or) not used in further analysis (e.g. ID, Zip code).
- Convert some of the fields to factors that helps in segmentation.

Derived Variables

- Remove % symbol from Interest rate & credit utilization.
- Calculate lost amount & loss percentage for analysis specific to Charged off loans.
- Segment the quartiles of funded amount by creating a variable.
- Split the credit utilization into a range (0-10, 11-21 etc.,)

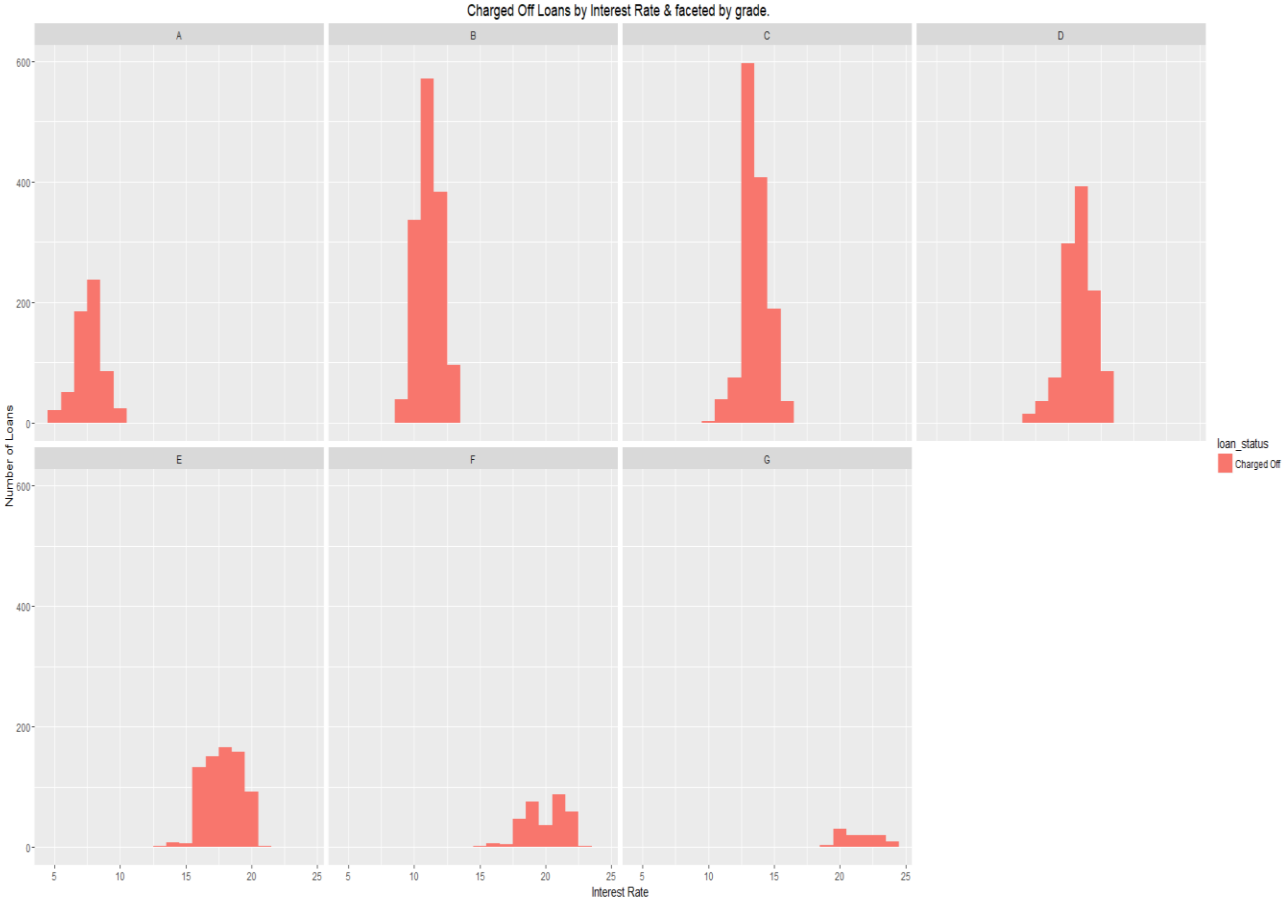
Plot: Issued Loans by Interest Rate with grade classification & faceted by Loan Term

- The Plot shows overall higher percentage of loans issues in 36 month term.
- However, for the grade level E, F & G, we see comparatively higher % of loans issued in 60 month term.



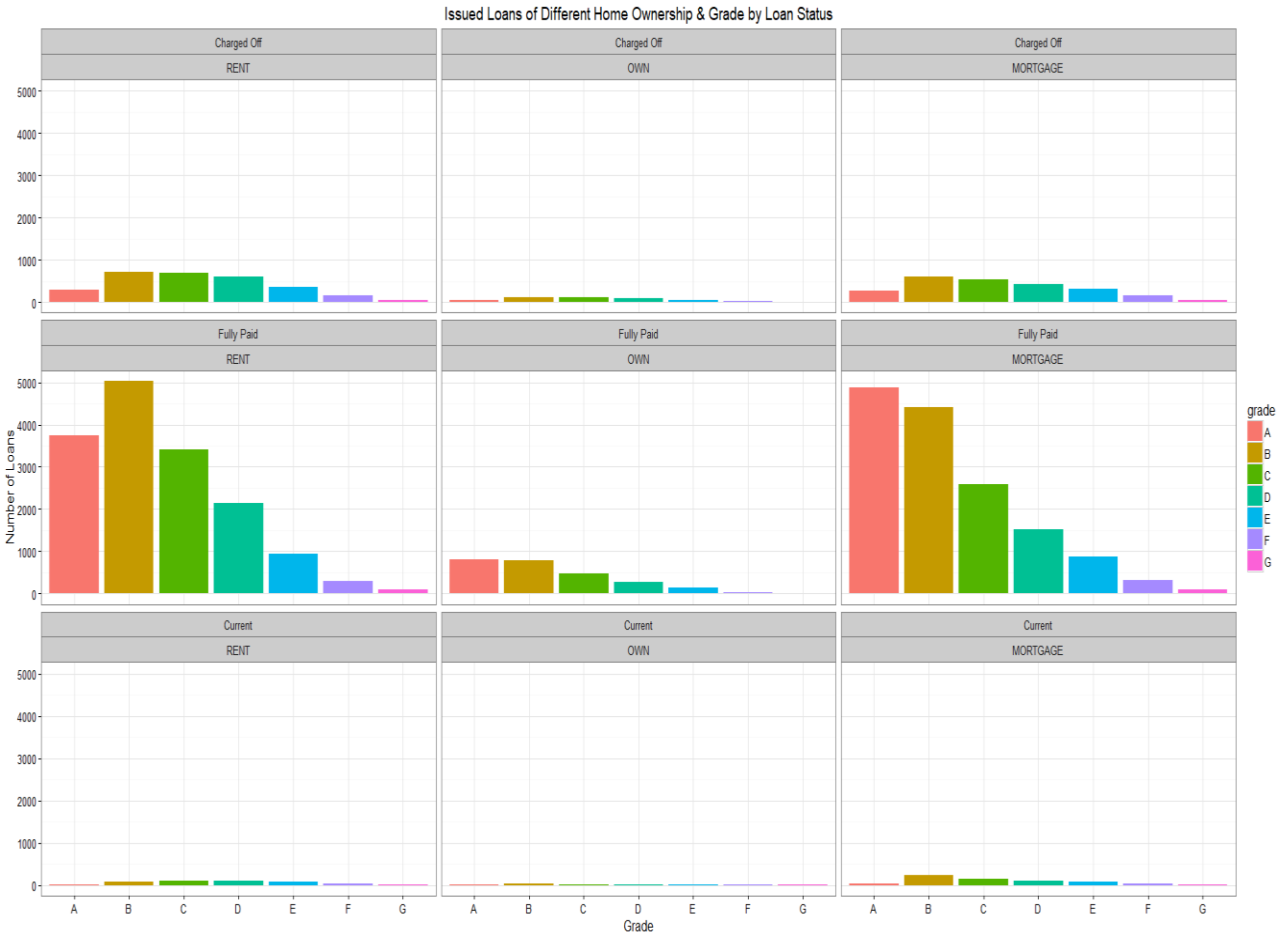
Plot: Charged Off Loans by Interest Rate & faceted by grade

- In this Plot we can see We see that the number of charged off loans are high even for grades B, C, D & E.
- This may be due to higher number of loans issues on those grades as we saw in the previous plot.



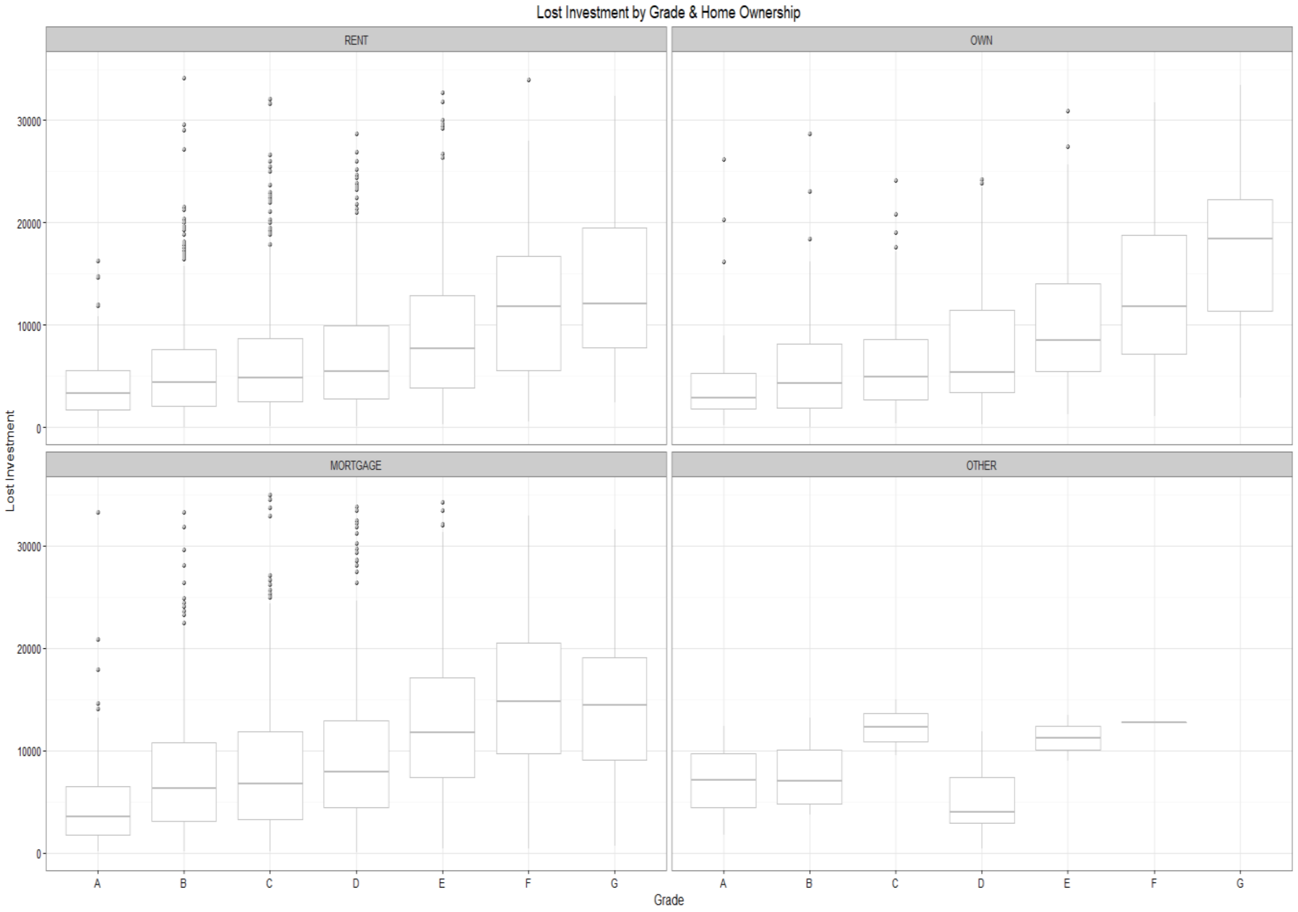
Plot: Issued Loans of Different Home Ownership & Grade by Loan Status

- We can see that higher number of loans are issued for borrowers who Rent or Mortgage.
- Also there is a higher number of borrowers in Grade A in Mortgage compared to rent.
- The borrowers who rent a home are more concentrated in Grade B & C when compared to mortgage.
- Thus borrowers who mortgage are having a higher chances to be in better grade A.



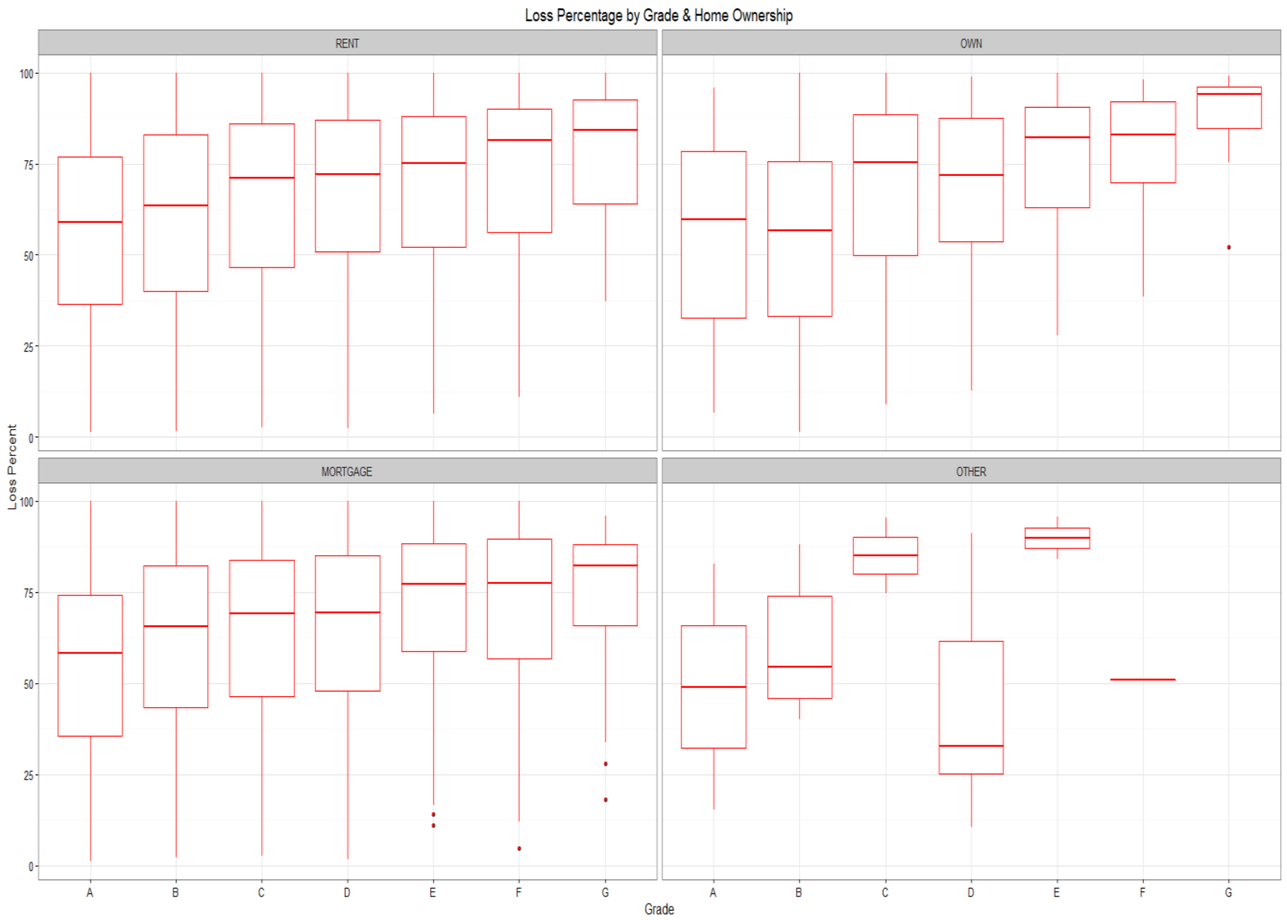
Plot: Loss amount by Different Home Ownership & Grade

- Overall median loss amount is high for borrowers who Mortgage in Grade G
- Median loss is very high in Home owners who have G grade.
- Overall, Renters are low risk across the grades as they incur less loss.



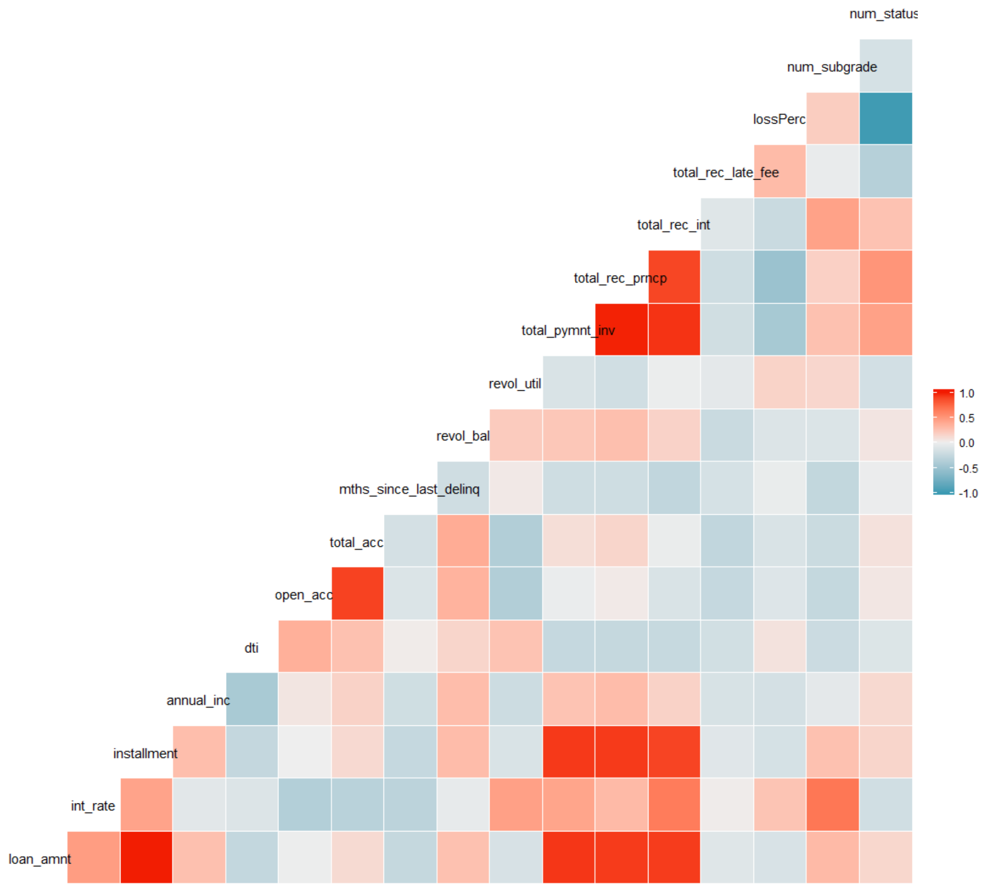
Plot: Loss Percentage by Different Home Ownership & Grade.

- Comparing only Rent, Own & Mortgage homeowner categories, we see that renters perform better compared to others.
- Loss percentage in home owners are skewed across grade while comparing with Renters or Mortgaged borrowers.



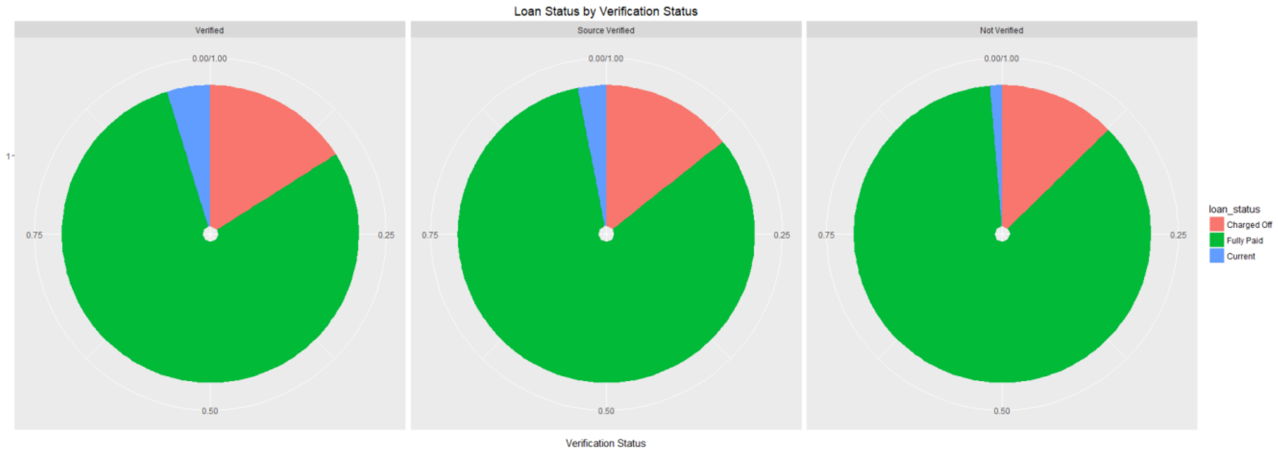
Plot: Correlation between variables

- dti is not correlated with loan *(Finding)*
- Payments received to date are dependent on instalment amount *(This is expected)*
- Principle received has a minor correlation with Loan Status *(Finding)*
- Interest Rate is correlated with Grade. *(Finding)*



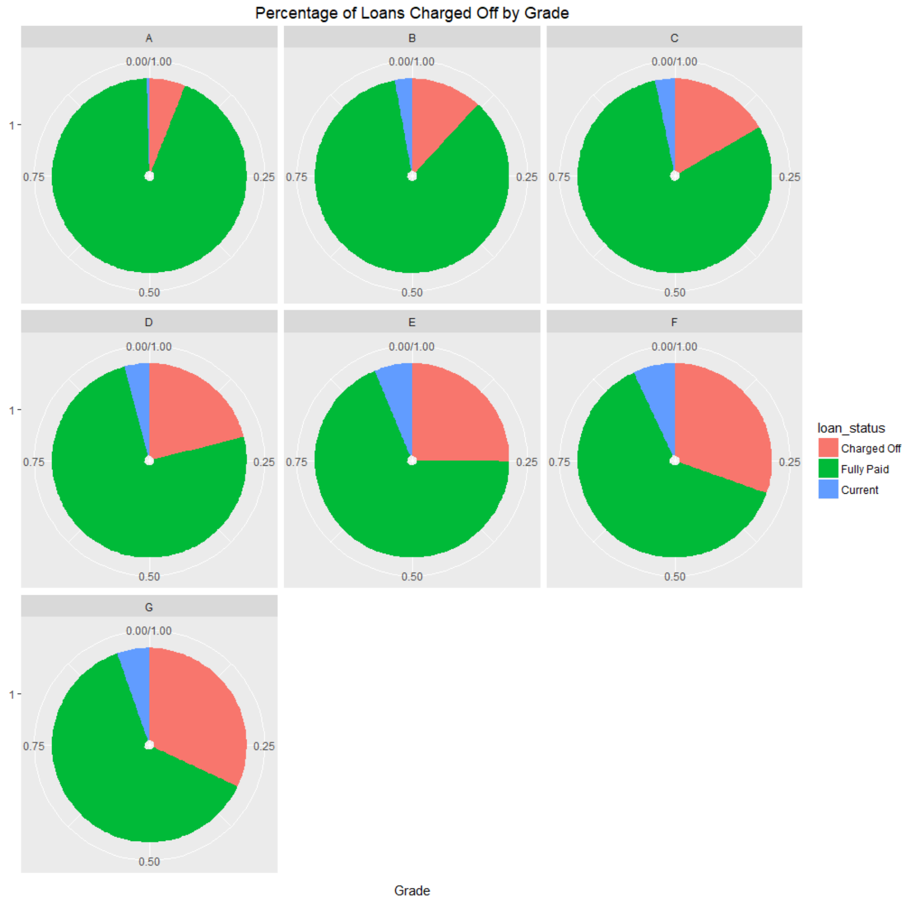
Plot: Charge Off Loans percentage by Verification Status

- Verification status didn’t make much difference in expectation of a charge-off.



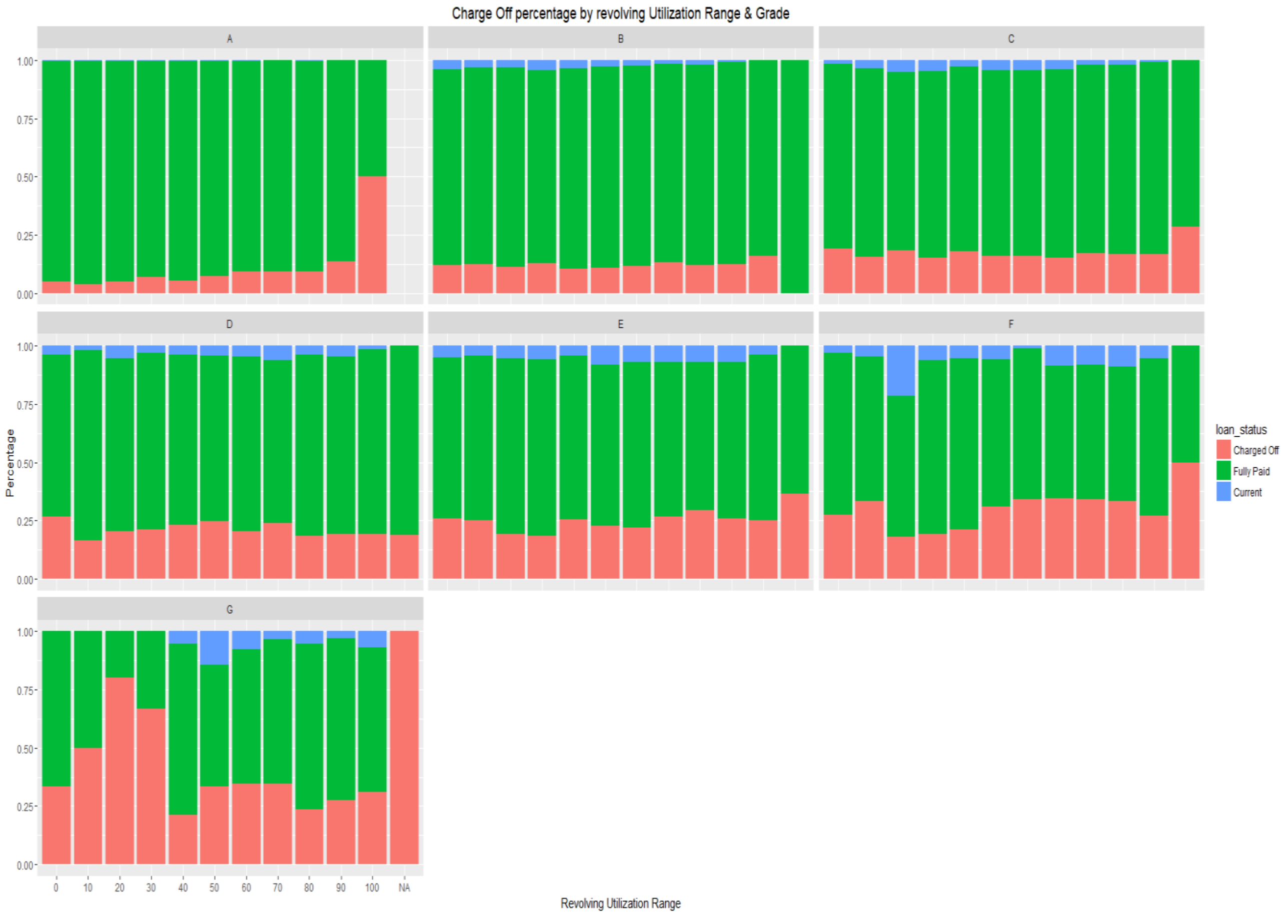
Plot: Charge Off Loans by Grade

- The Percentage of loans charged off increases with the grade of the borrower.



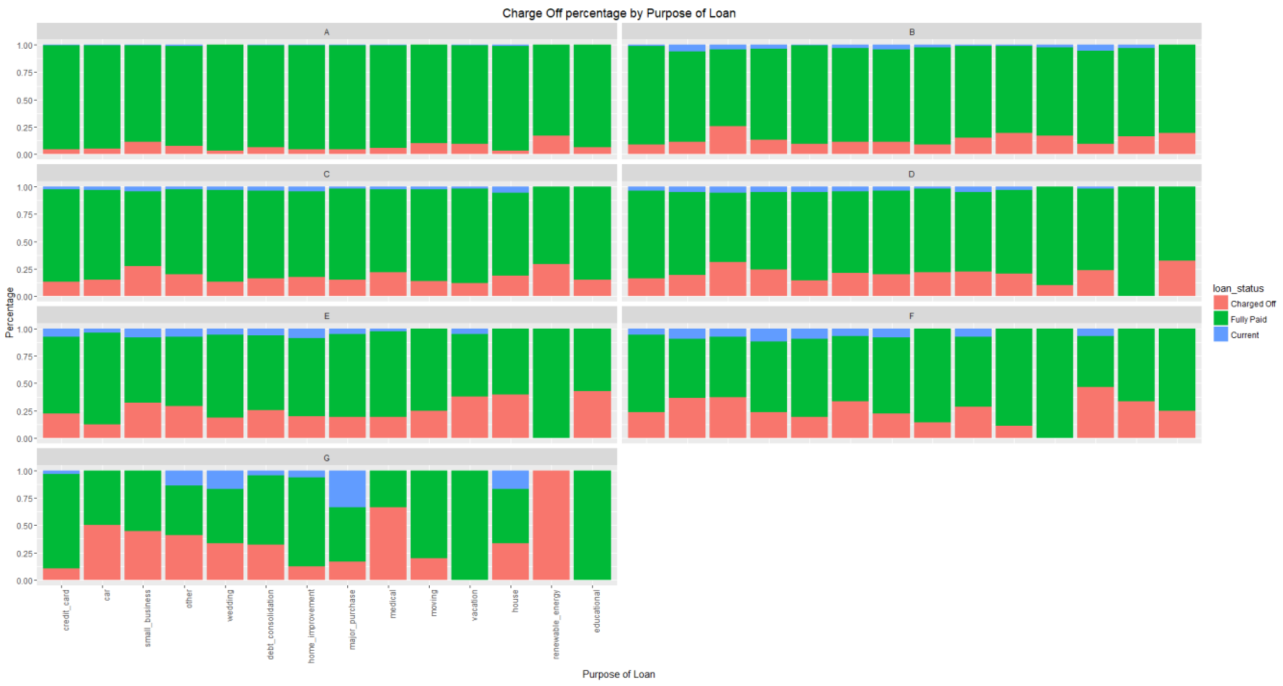
Plot: Charge off Loan Percentage by revolving utilization range & grade

- We see that in Grade A, as the revolving utilization percentage increases, the charge off percentage increases.
- For Grade F & G, the data is skewed. But the charge off percentage is highest for harvest Utilization percentage.



Plot: Charge Off Loans Purpose

- Charge off percentage of loans is higher for Small business purpose across the board.
- In Grade G, loans for renewable energy & medical are the highest risk.
- In grades A through C, loans for purpose of renewable energy are at a higher risk.



Final Summary

No Effect on Defaulting Risk

- Verification status does not impact default percentage.
- Home Ownership does not impact default percentage.
- Defaulting risk is same across the states where we have significant data. *(States like NE, IN & IA have low data and thus their findings are excluded)*

High Defaulting Risk

- Higher grade of Borrower
- Loans for small Business Investment.
- Higher net loss when investment amount is high.
- Home owners with G Grade are highly likely to default.
- Risk is high when employment data is not available.
- Most charged off percentage is for F and G grades, where as maximum charged-off loans belong to B and C

THANK YOU!