

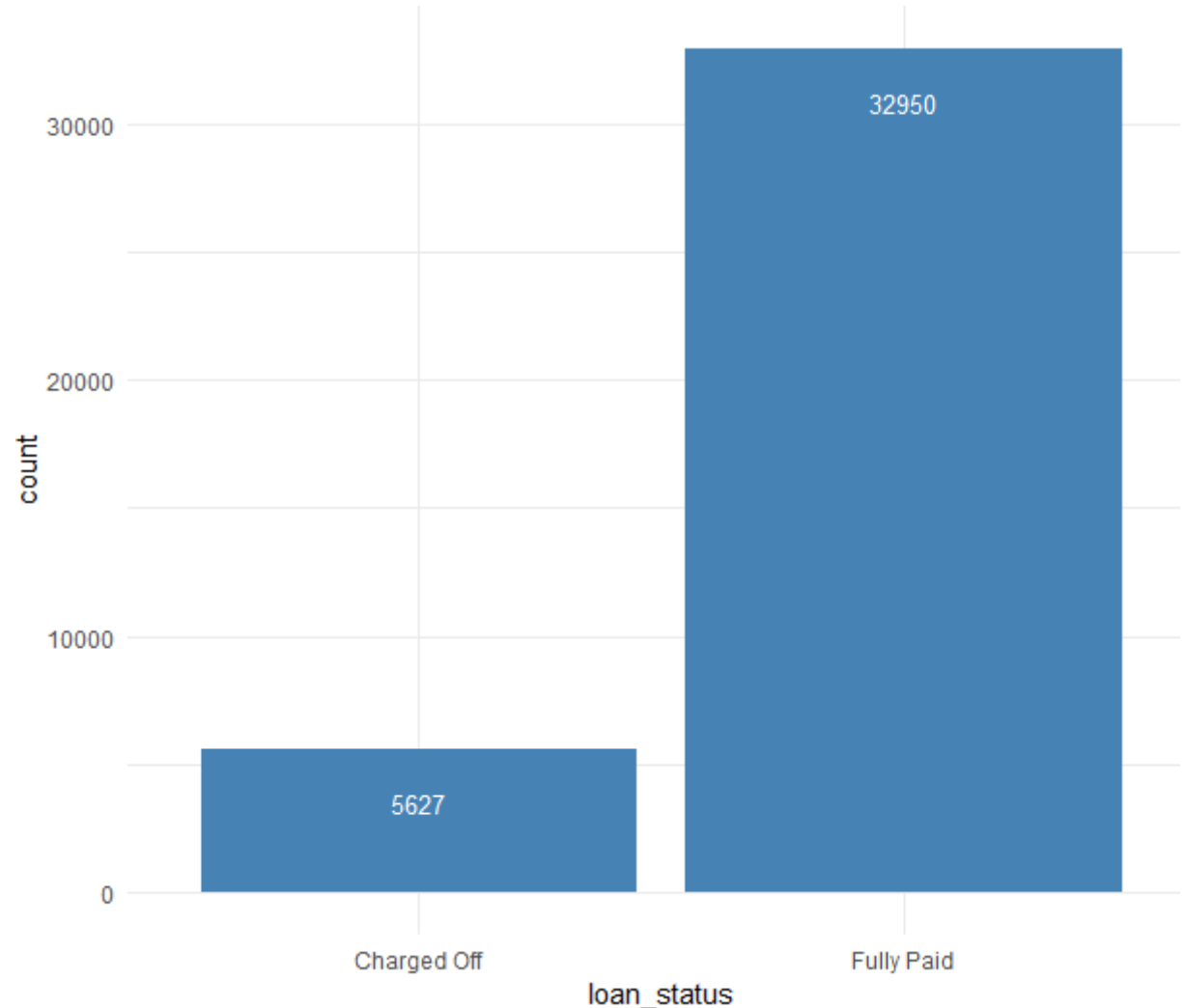
EDA Case Study

SUBMISSION

- Bharat Saxena
- Mohit Singh
- Sunil Hegde
- Saurabh Ghosh

Objectives – Summary

- Use EDA to understand the driving factors behind charged off loans.
- The data contains 3 types of loan status.
 - Charged Off
 - Fully Paid
 - Current
- Since loan with status “current” can end up either charged off or fully paid, they have not been included in this analysis.



Derivations from Univariate Analysis

Following few slides will layout factors which show an impact on the loan status outcome during univariate analysis.

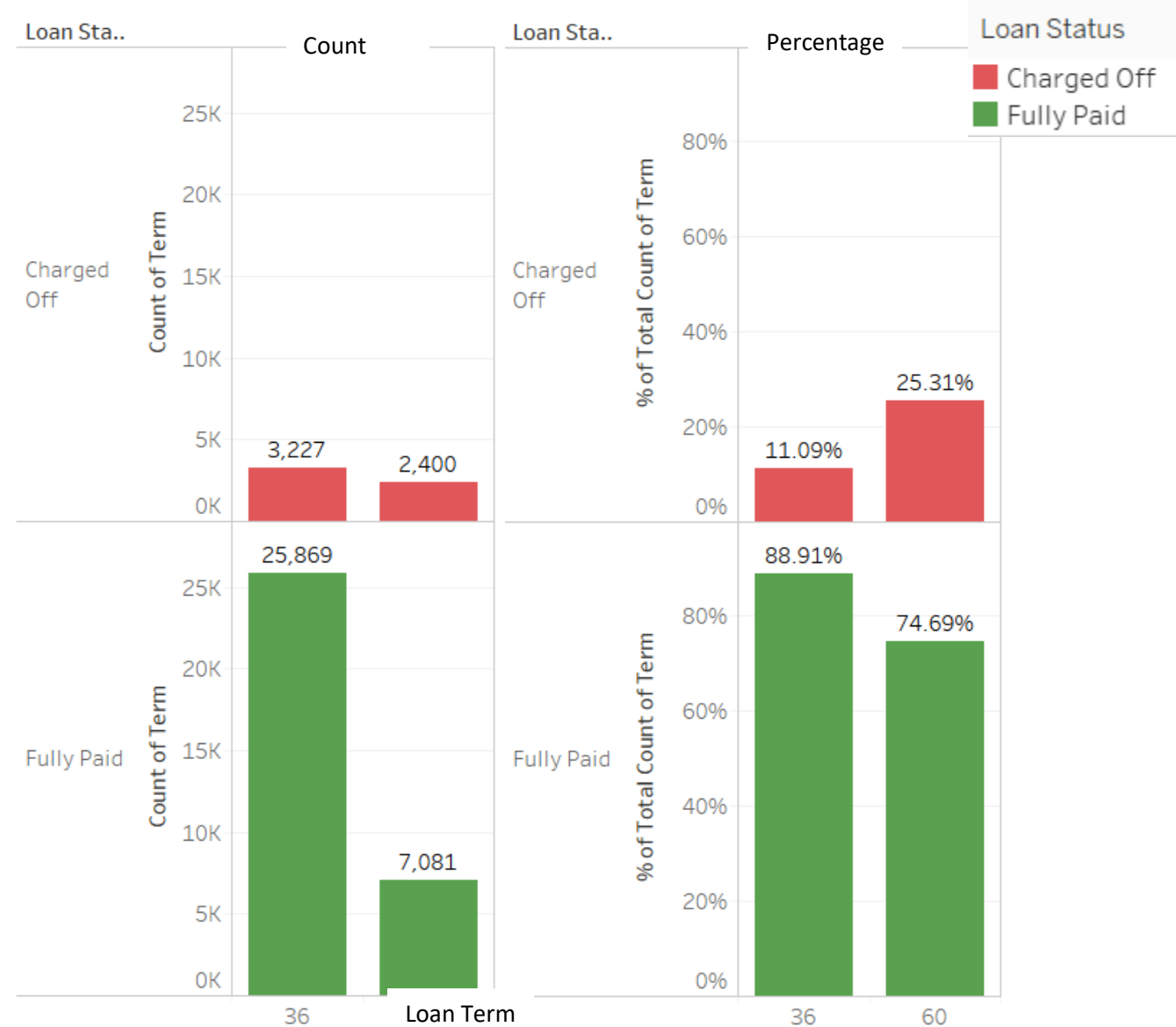
- UniA1 – Term of the loan
- UniA2 – Grade of the loan
- UniA3 – Home ownership
- UniA4 – Loan purpose reported in the application
- UniA5 – Interest Rate and DTI
- UniA6 – loan_amount, funded_amnt, and funded_amnt_inv

NOTE #1: All factors were analysed individually, and this list contains only those which were found to have an affect.

NOTE #2: These factors alone can not be used to predict the loan status outcome. However they seem influential.

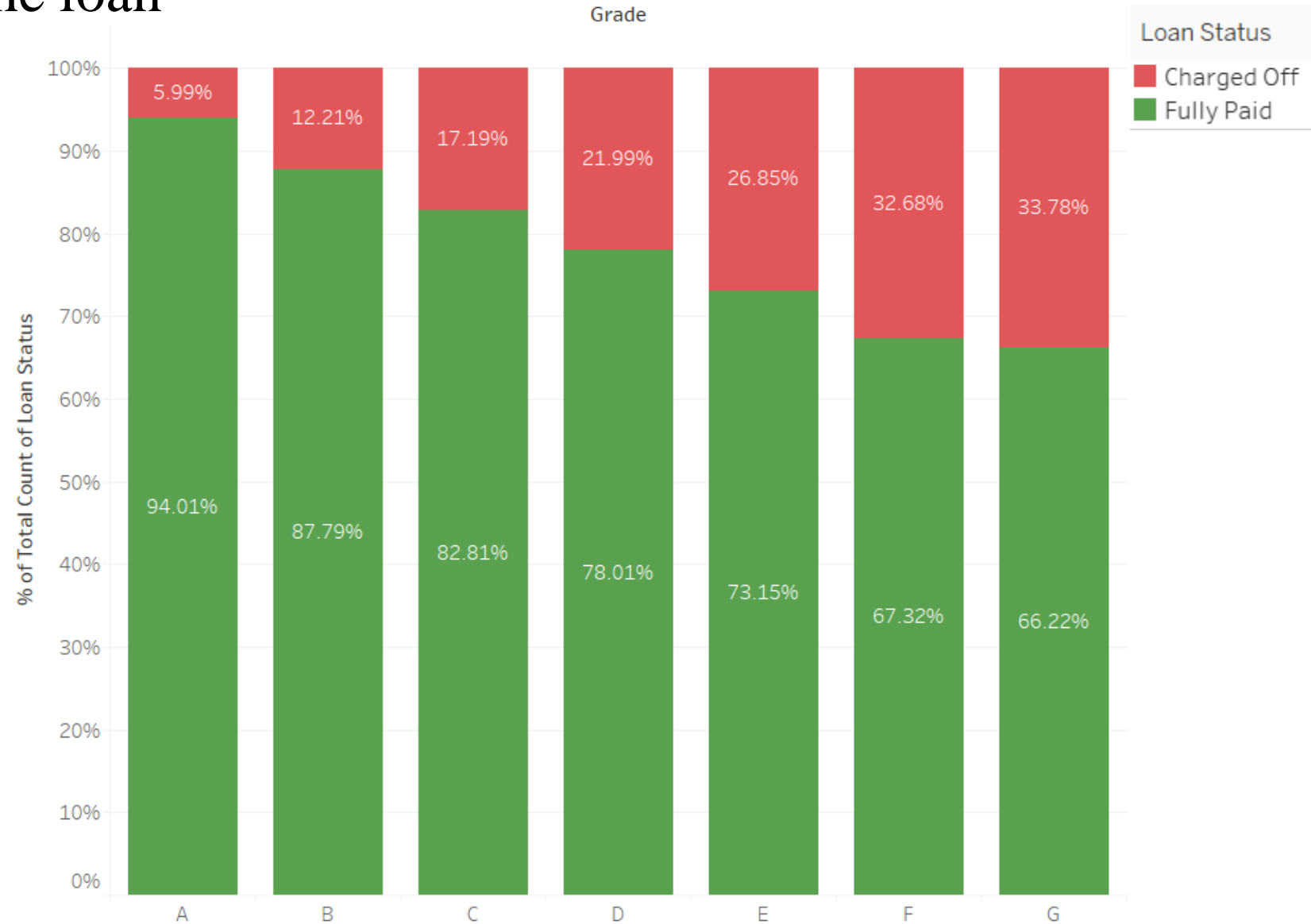
UniA1 – Term of the loan

- 36 month's term loan have higher number of charged off loans.
- But that's only 11% of the total 36 month term loans.
- 60 month's term loan have 25% charged off ratio.
- **60 month term loans have higher % of default rate.**



UniA2 – Grade of the loan

- Grade G loans have the most share of the loans which get charged-off.
- This general trend is seen in the sub-grade categories as well. With sub-grade F5 being the exception as it has highest ratio of charged off loans.
- **Loan Default proportion increases from Grade A to G.**



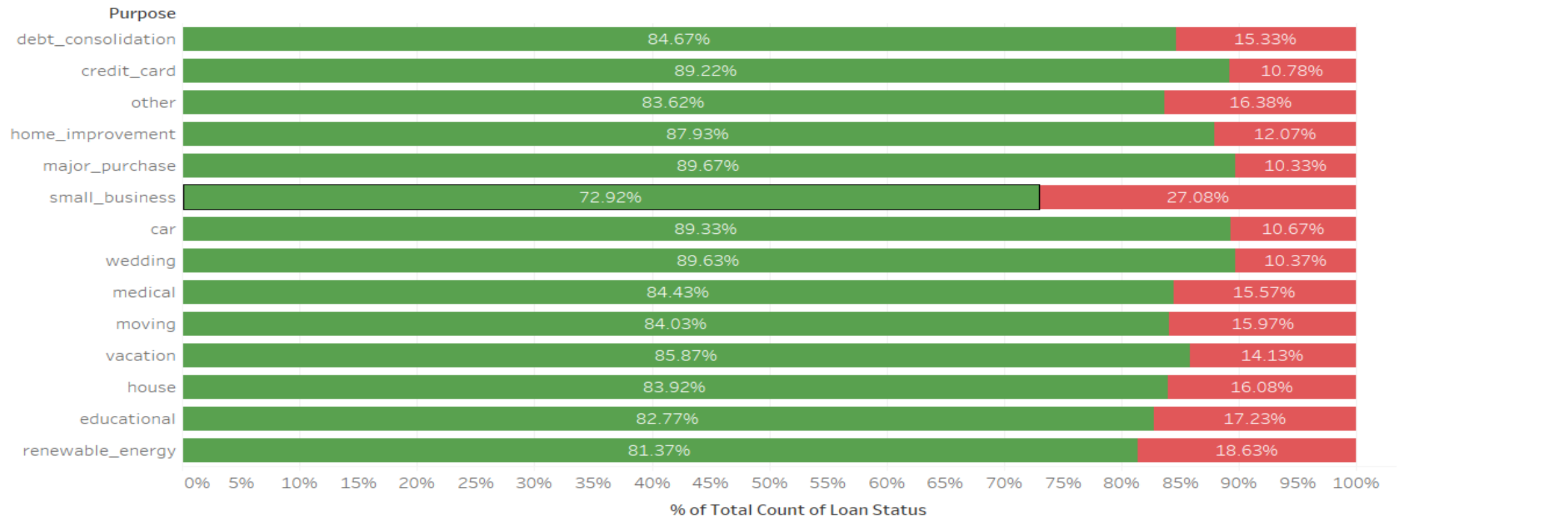
UniA3 – Home ownership

- Home ownership status of “others” show the most percentage of loans being charged off followed by status “rent”.
- But the charged off ratio is almost similar across the board.
- Home ownership doesn't really have much impact on loan status.



UniA4 – Loan purpose reported

- Loan purpose of “Debt consolidation” accounts for about 47% of all the loans.
- Loan purpose of “small business” sees most number of charge off. 27% of all the loan issued with the purpose “small business” end up being charged off.



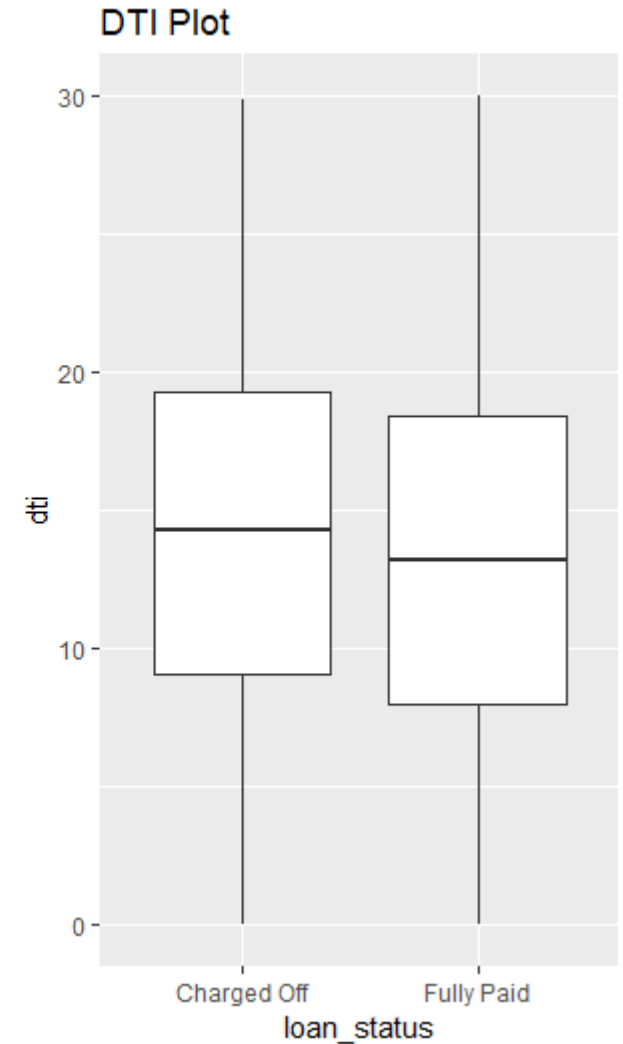
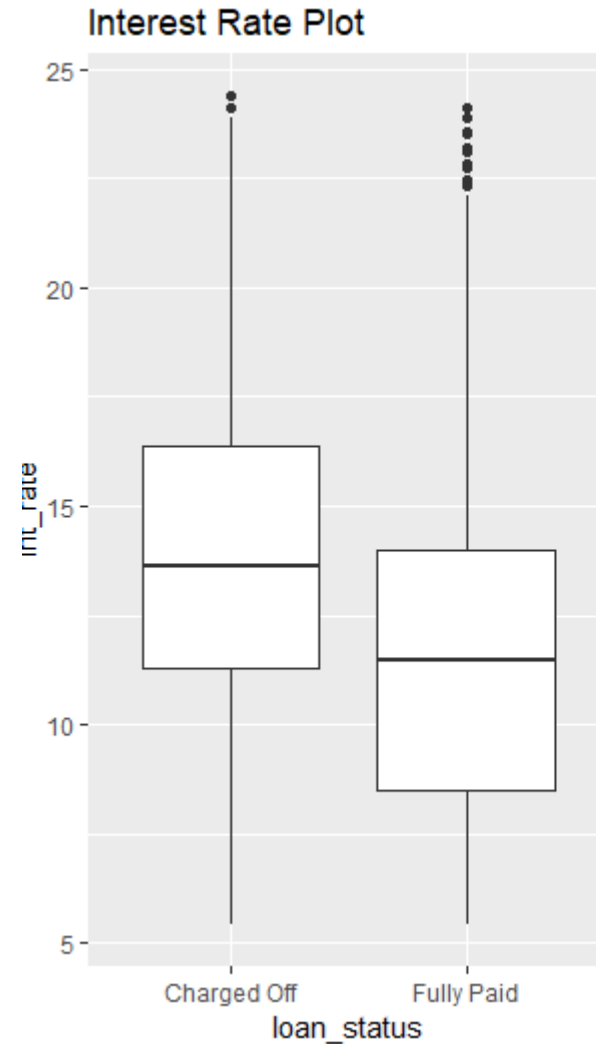
Derivations from Bivariate Analysis

Following few slides will layout factors show an impact on the loan status outcome during univariate analysis.

- BiA1 – Interest Rate and DTI
- BiA2 – Approved vs Defaulted Loans over time
- BiA3 – Annual income against loan amount approved
- BiA4 – loan_amount, funded_amnt, and funded_amnt_inv

BiA1 – Interest Rate and DTI

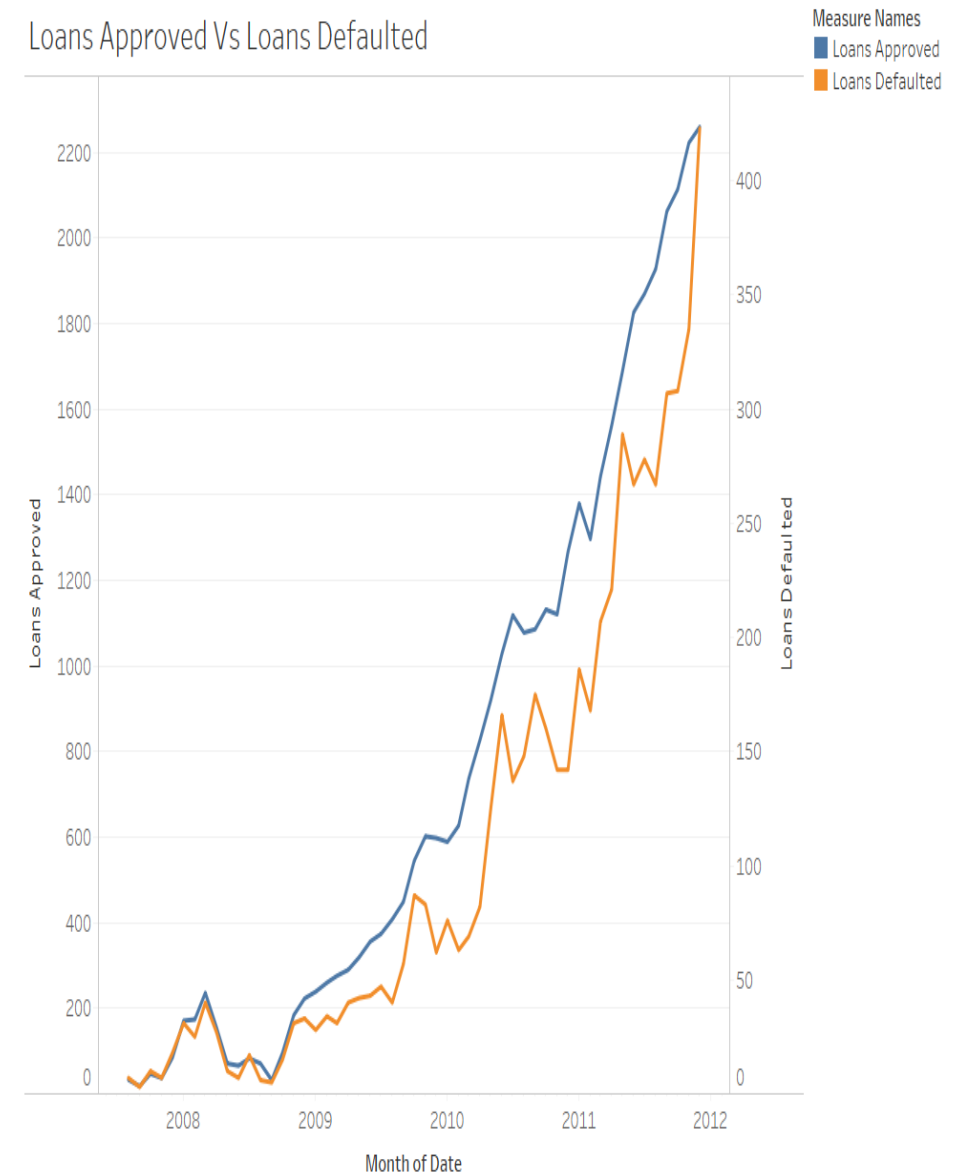
- Charged off loans on an average have **higher interest rates** were higher compared to fully paid.
- Charged off loans have people with **higher average DTI**.



BiA2 – Approved vs Defaulted Loans

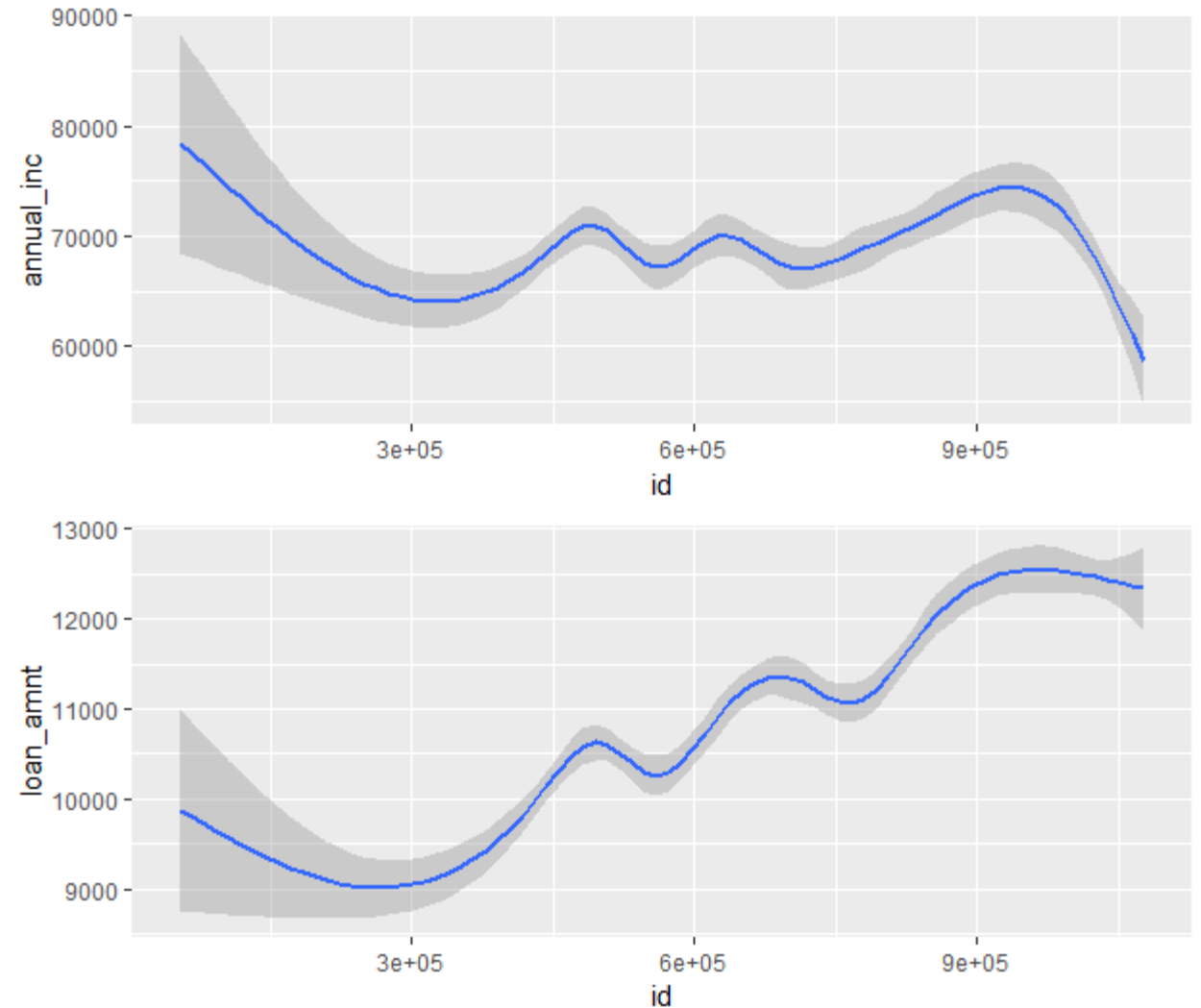
- The number of defaulters increased with the increase in number of loans over the period of time.
- **During end of 2011 default count has increased drastically.**

Loans Approved Vs Loans Defaulted



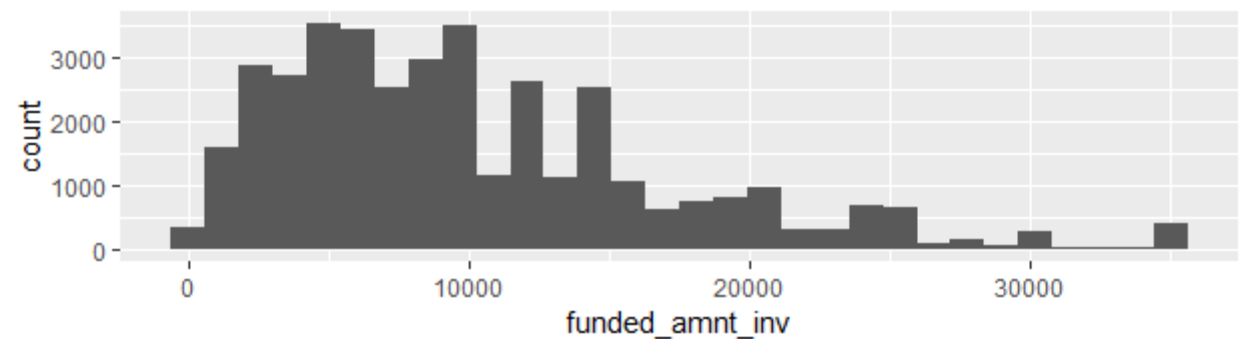
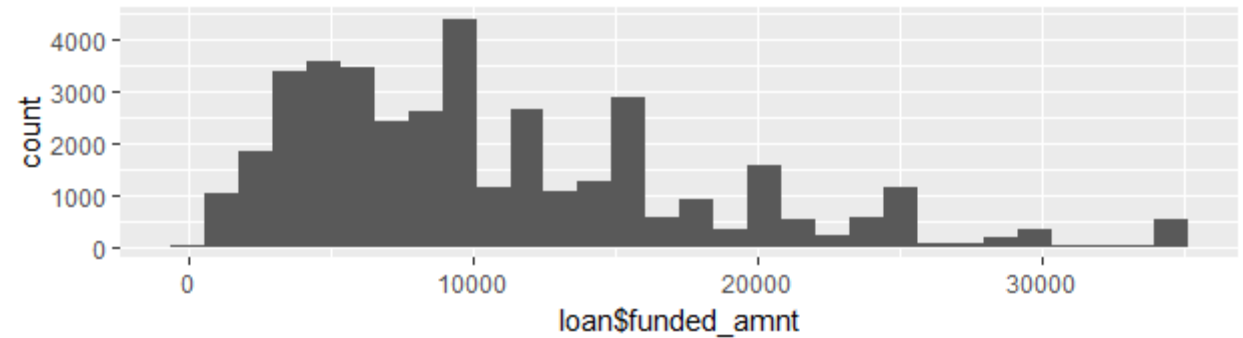
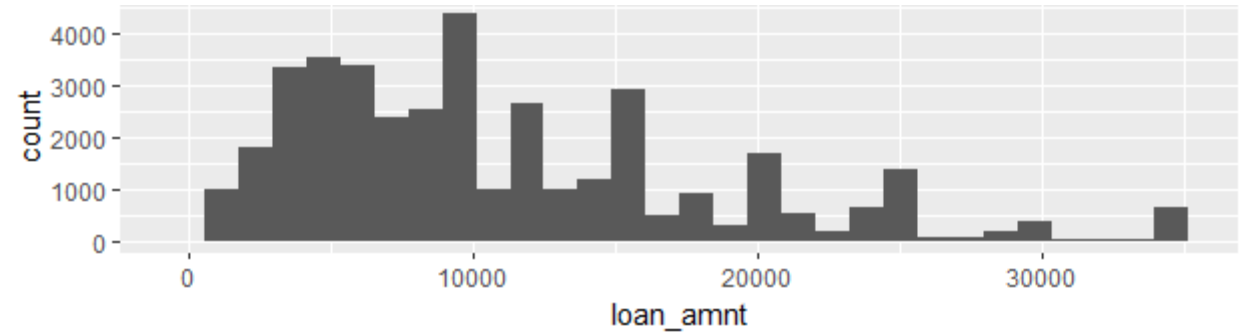
BiA3 – Annual income against loan amount approved

- As the graph suggests high loan amount has been approved for less annual income borrowers.



BiA4 – loan_amount, funded_amnt, and funded_amnt_inv

- There appears to be a balance between funding, lending and requested loan amount.



Analysis findings.

- Defaulted loan have **higher** Interest rate as compared to fully paid loans.
- Defaulted loan have **higher** DTI as compared to fully paid loans.
- People with defaulted loan, have **lesser Average annual** salary compared to people who have paid fully.

Loan Status	int_rate		DTI		annual_inc	
	Mean	Median	Mean	Median	Mean	Median
Charged Off	13.82043	13.61	14.00062	14.29	62427.30	53000.00
Fully Paid	11.60976	11.49	13.14842	13.20	69862.50	60000.00