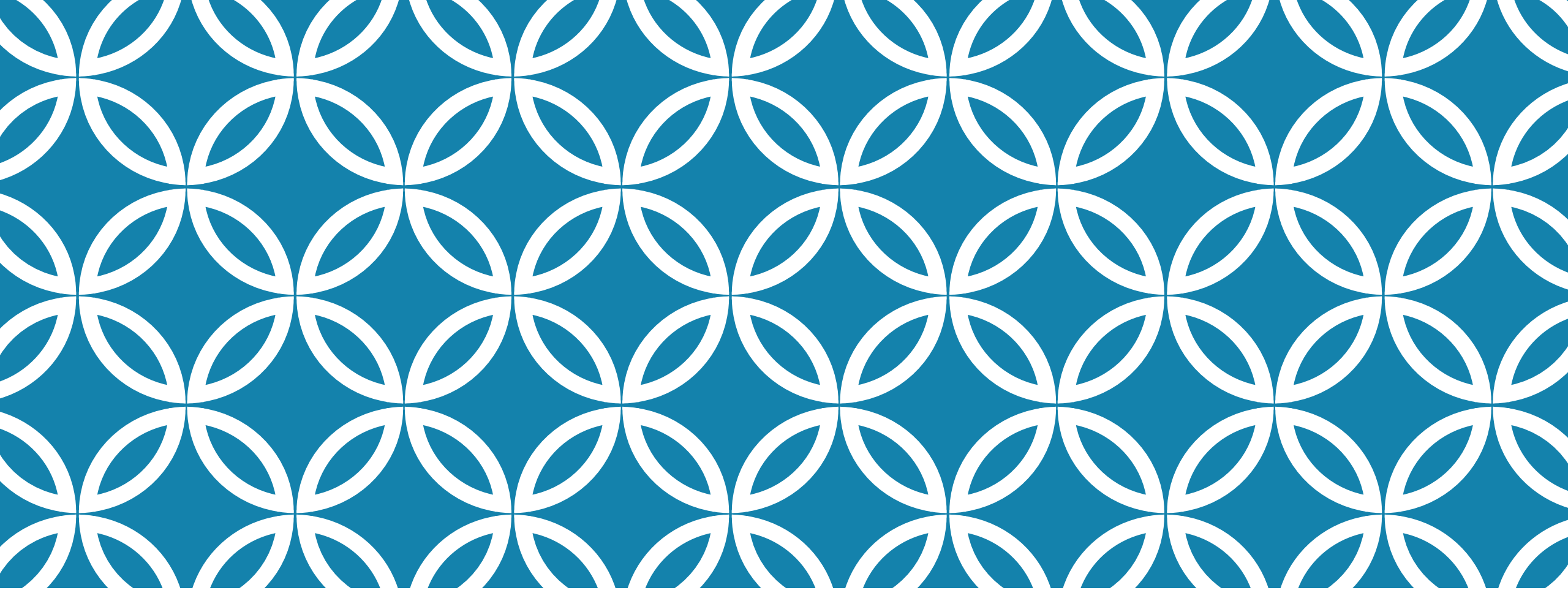


LENDING CLUB CASE STUDY

Sarada Prasad Parida &
Swapnil Dakshinkar

OBJECTIVE

- Identifying applicants where the loan default can be high, so that the company can avoid giving loans and reduces losses.

- Understand the applicant and its loan attributes where company can provide loans so that company can make profit through interest earned from this loan so that it can be profitable.

THE DATA SUMMARY

	Before Cleanup	After Cleanup
Rows	39717	34118
Columns	111	22

Assumptions: We have removed some columns like 'last_credit_pull_d','earliest_cr_line','out_prncp','out_prncp_inv','total_pymnt','total_pymnt_inv','total_rec_prncp','total_rec_int','total_rec_late_fee','recoveries' etc

Applicant Related Columns	Loan Related Columns
annual_inc	funded_amnt
purpose	int_rate
grade & sub grade	installment
Employment length	dti
Home ownership	term

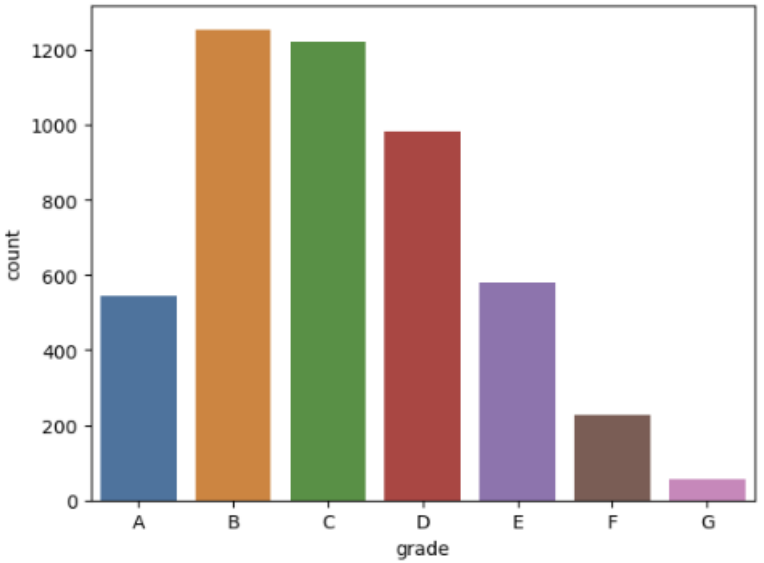
During analysis we found above columns more responsible for the outcome

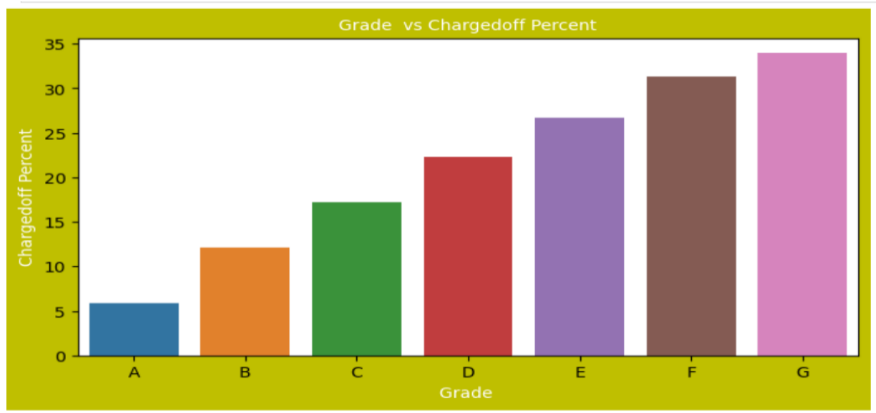
SEGMENTED UNIVARIATE ANALYSIS

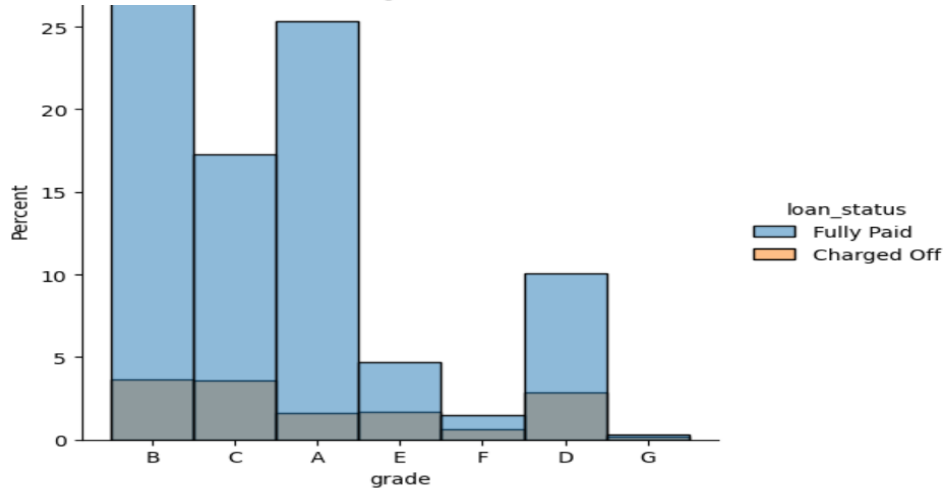
GRADE

Assumption: default percentage > 10% as very high

Observation :
%wise grade C and onwards have very defaults rates





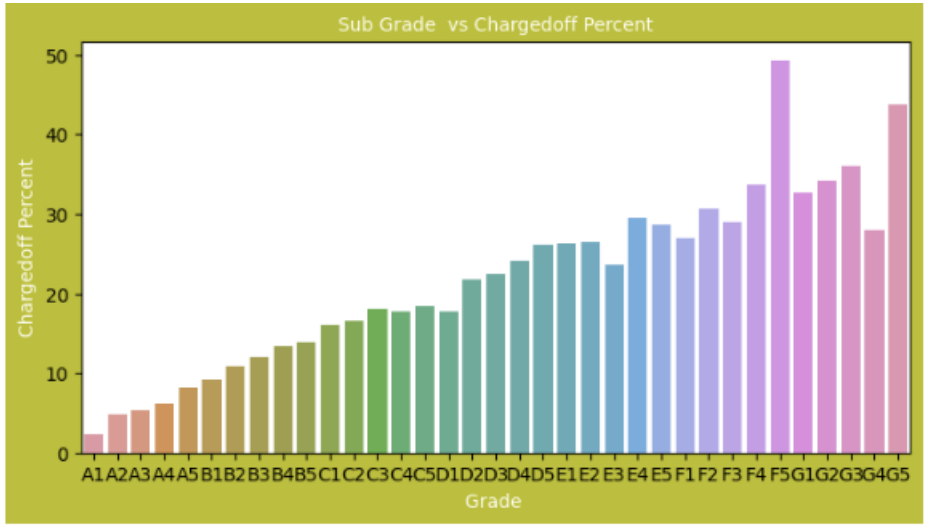


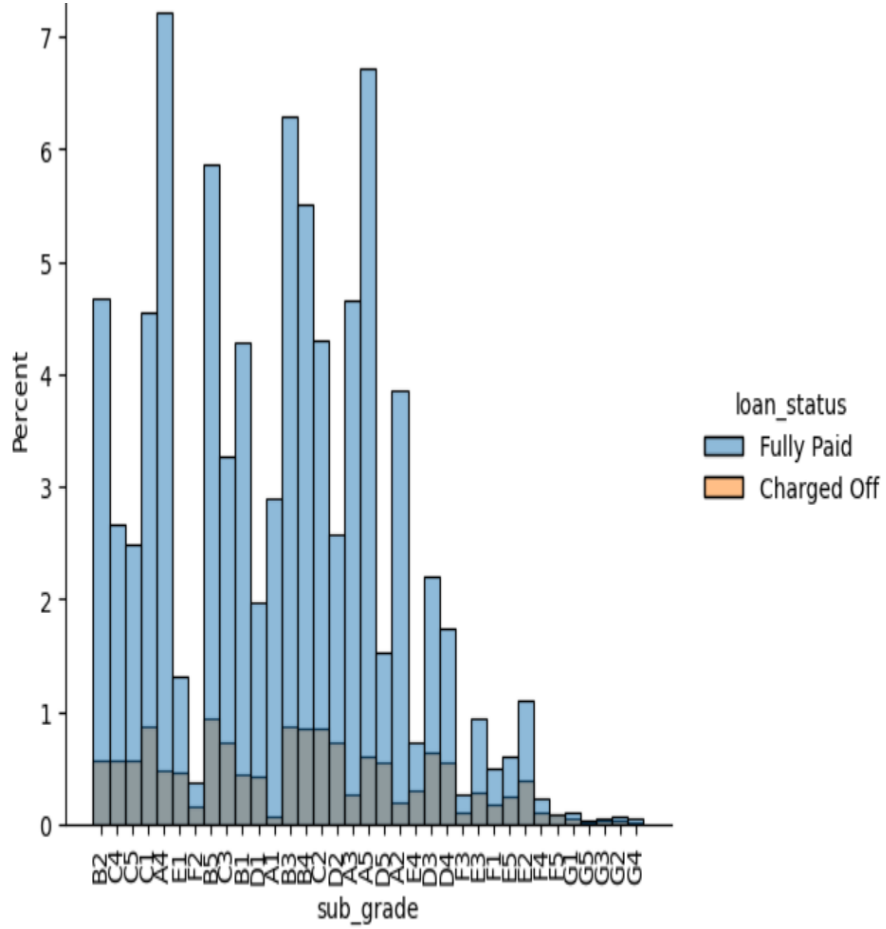
SEGMENTED UNIVARIATE ANALYSIS

SUB-GRADE

Assumption: default percentage > 10% as very high

Observation :
%wise grade B3 and onwards have very high default rates



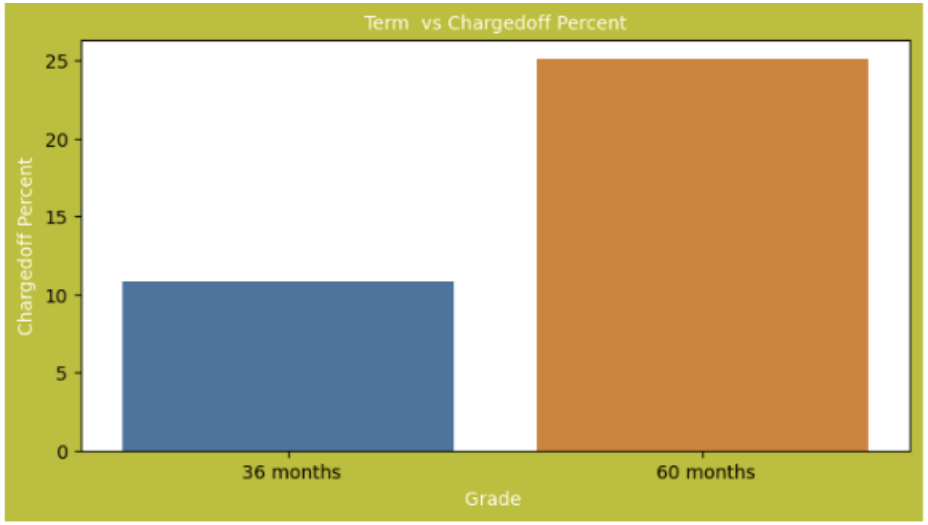


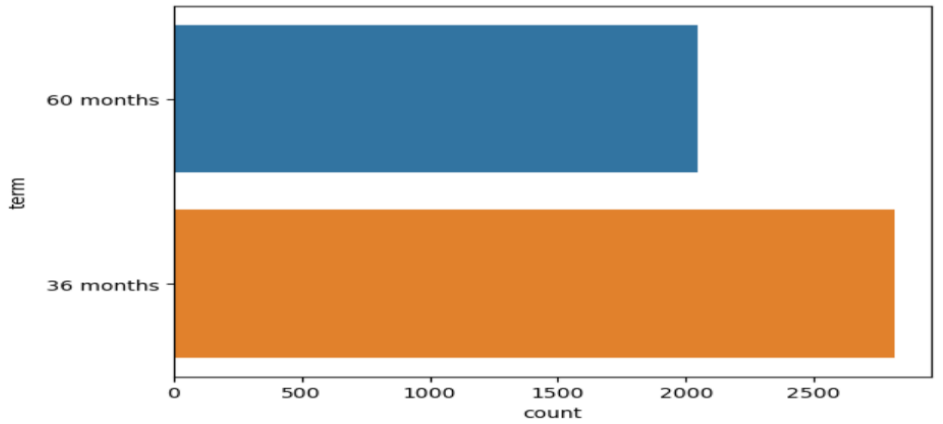
SEGMENTED UNIVARIATE ANALYSIS

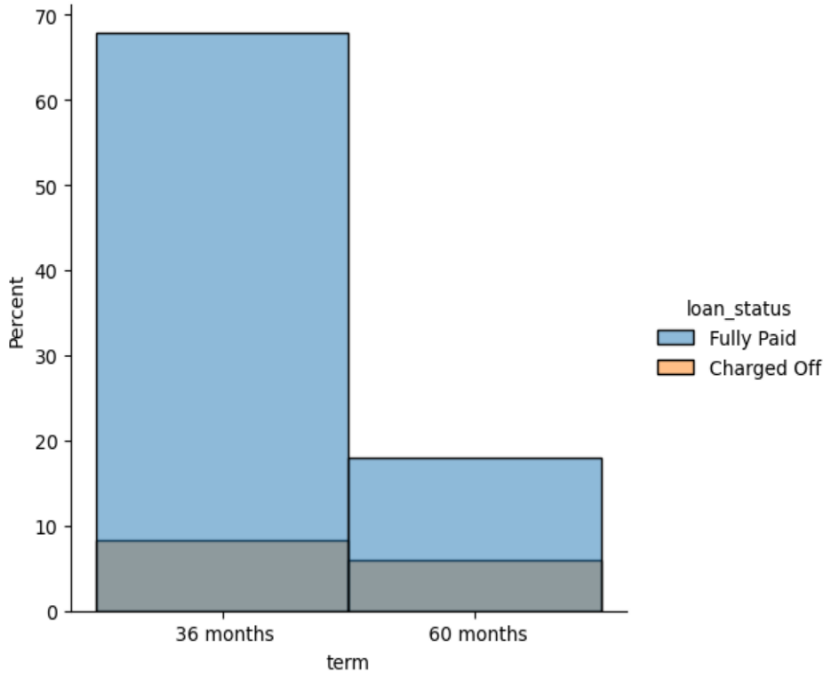
TERM

Assumption: default percentage > 10% as very high

Observation:
loan for term of '60 months' has 25% defaults, thus very high default rate





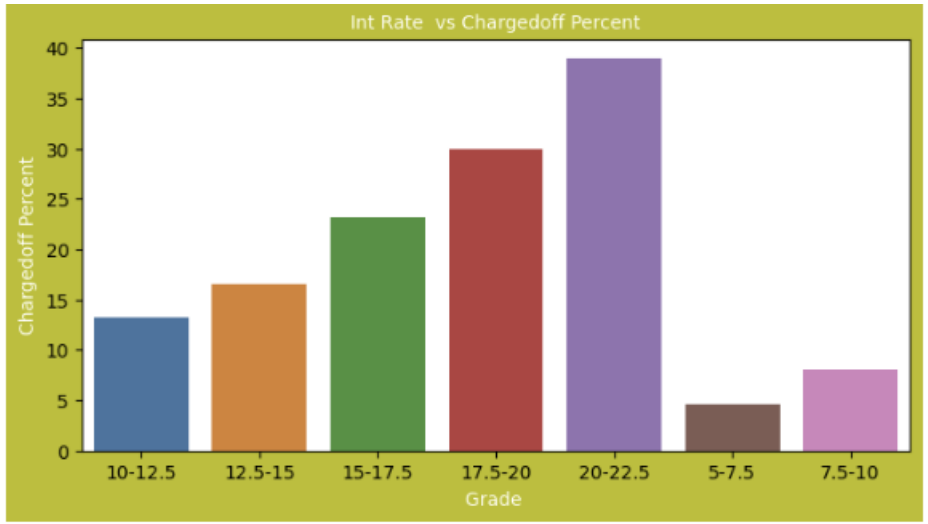


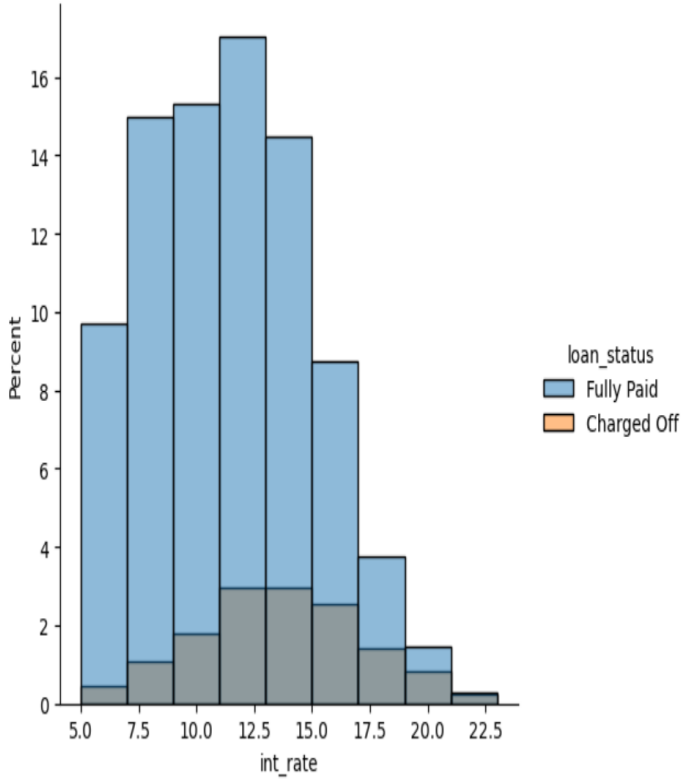
SEGMENTED UNIVARIATE ANALYSIS

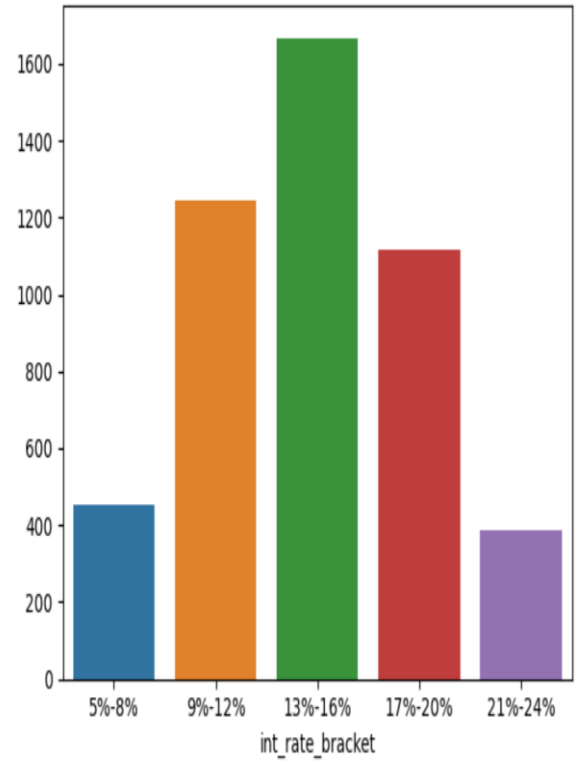
Interest Rate

Assumption: default percentage > 10% as very high

Observation:
Int rate of 10% and higher have more than 13%defaults, so very high default rates





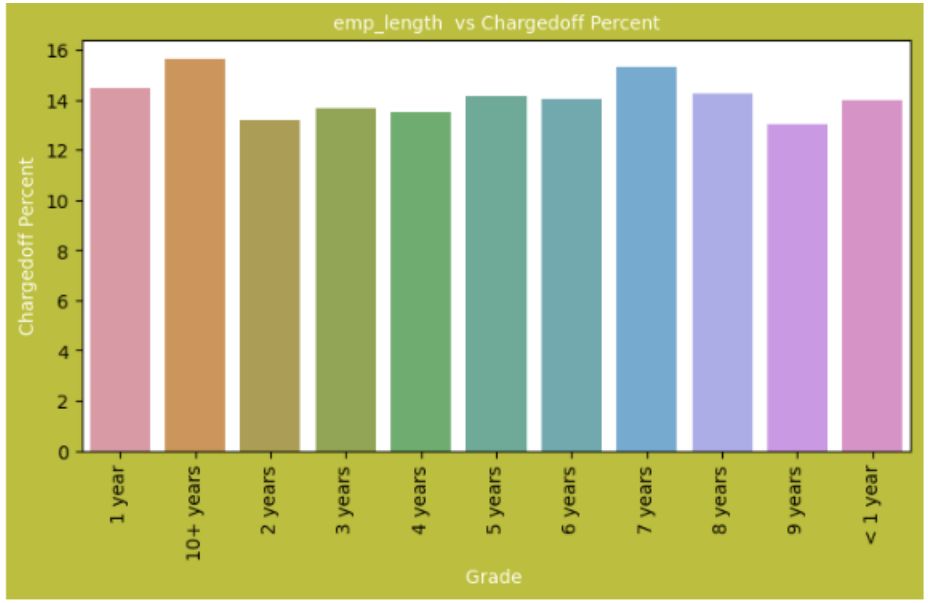


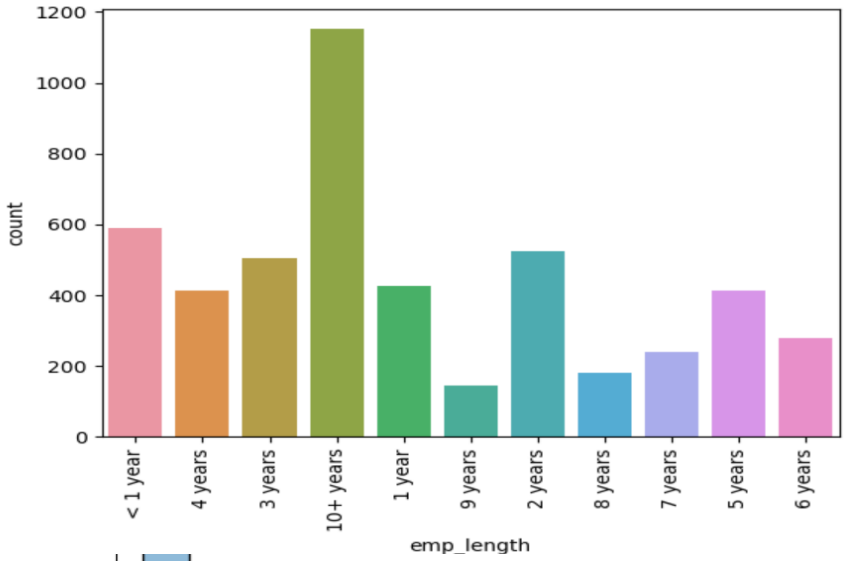
SEGMENTED UNIVARIATE ANALYSIS

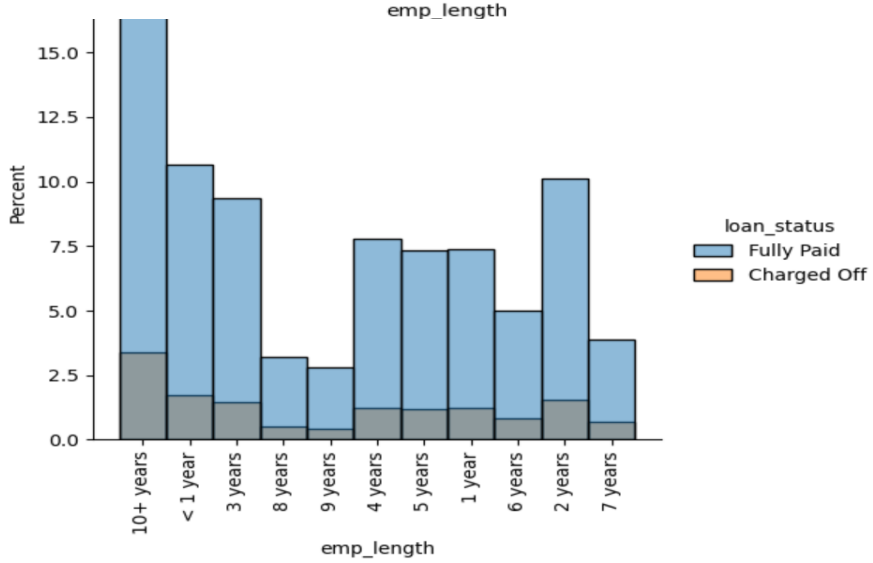
EMP LENGTH

Observation:

No significant difference in percentages of defaults for every emp_length category



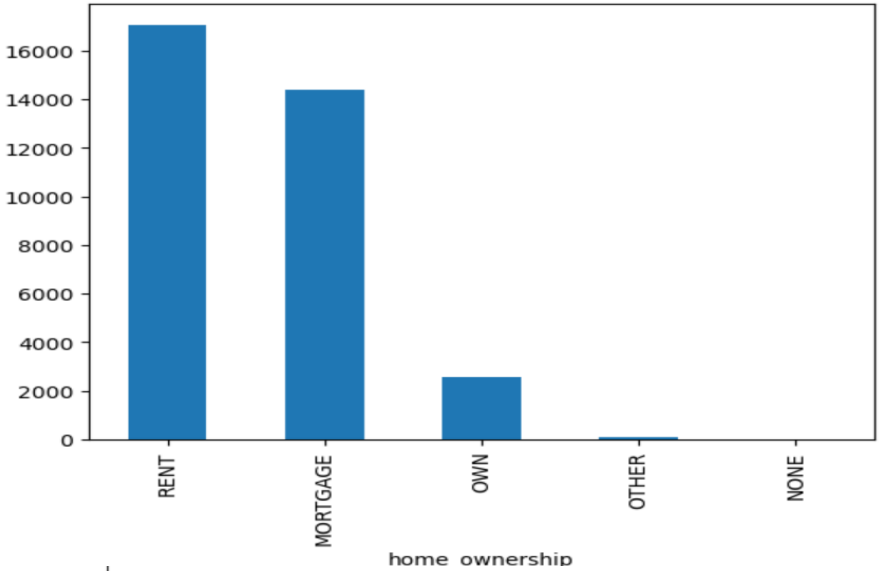


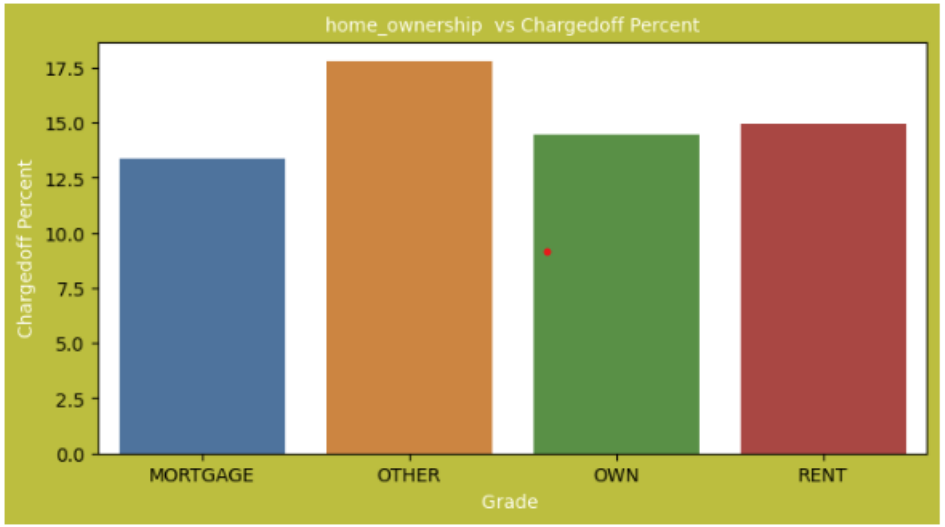


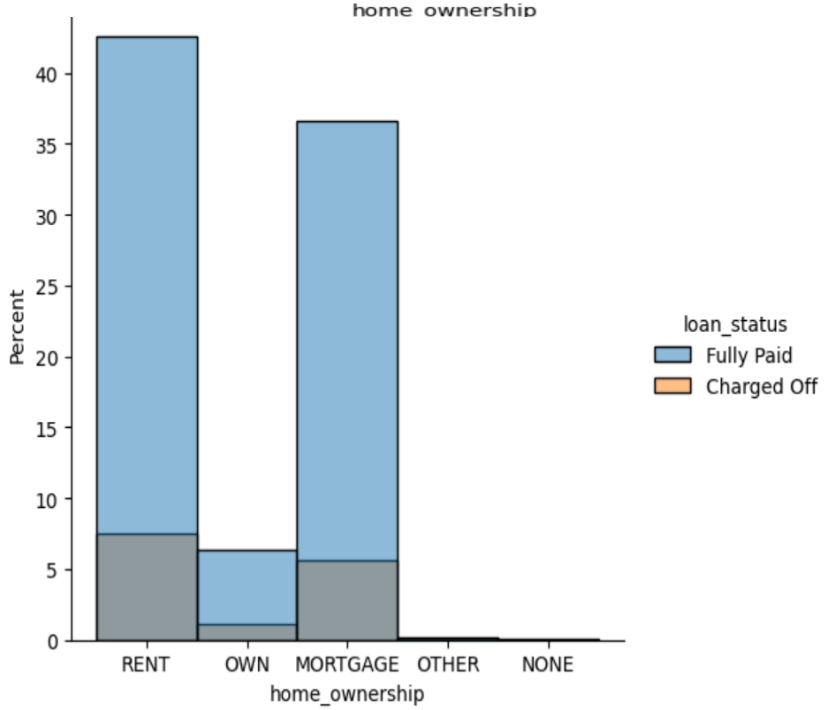
SEGMENTED UNIVARIATE ANALYSIS

home ownership

observation:
'Other' have very high default rate of 17.5 as compared to remaining categories , but number wise very less such rows with that category, also no significant difference in the default rates of remaining categories , so we conclude , home ownership has no effect on rate of defaults







SEGMENTED UNIVARIATE ANALYSIS

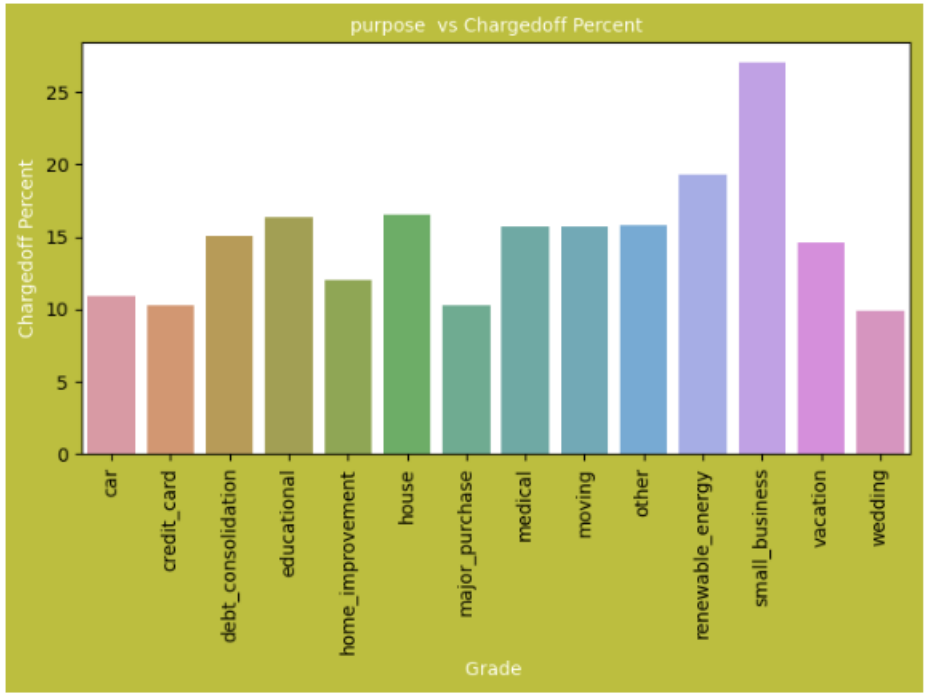
purpose

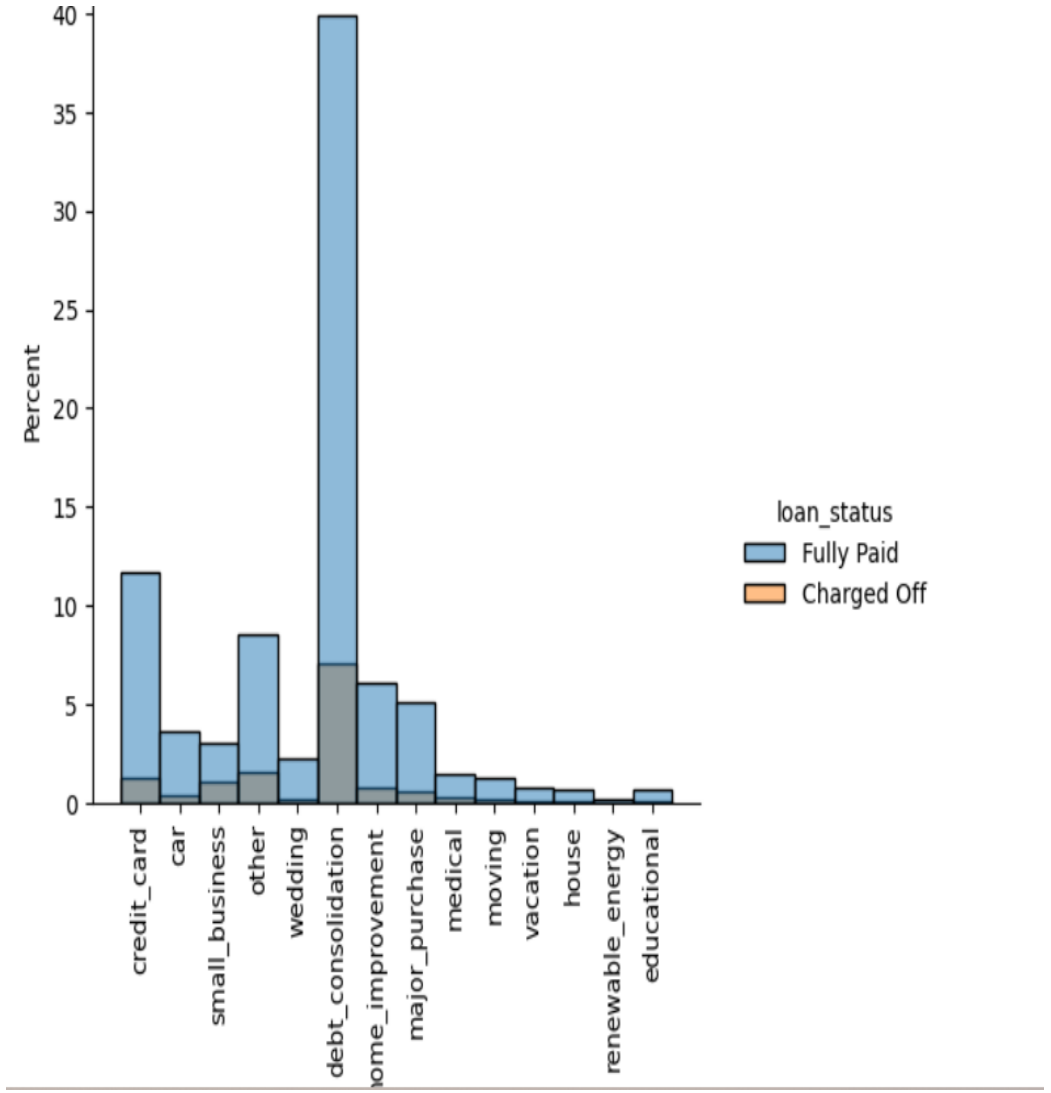
Observation:

1.wedding,major_purchase, credit_card, car, home_improvement have lower rates of defaulting

2.vacation,debt_consolidation, moving, other,educational, house have medium rates of default

3. renewable_energy, small_business have very higher rates of default



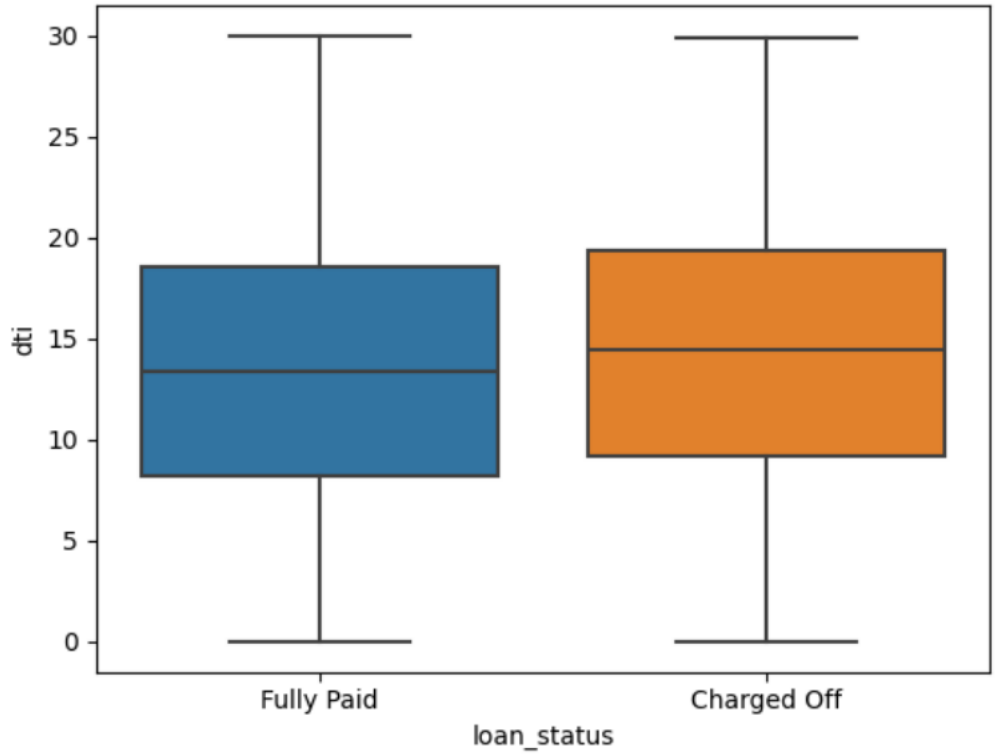


SEGMENTED UNIVARIATE ANALYSIS

dti

Observation:

No significant difference of spread of defaulted/paid loans wrt dti.

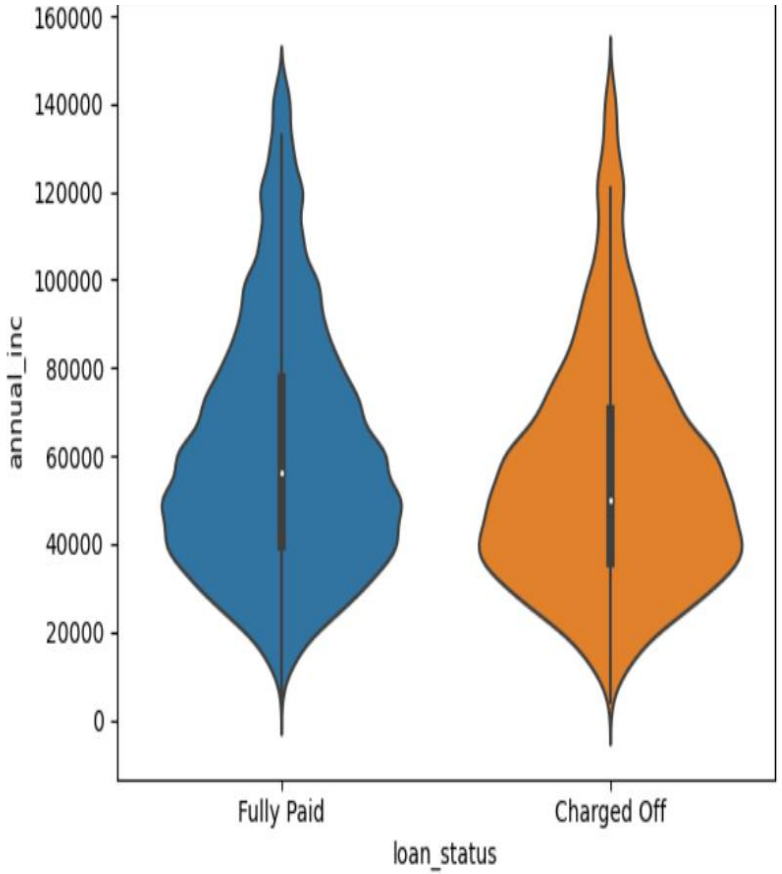


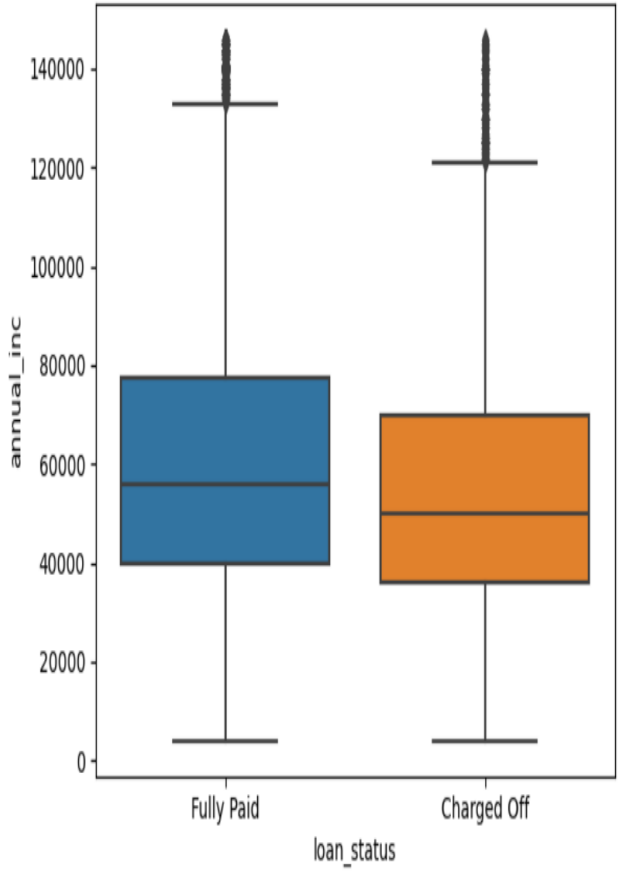
SEGMENTED UNIVARIATE ANALYSIS

annual_inc

Observation:

No significant difference of spread of defaulted/paid loans wrt annual_income



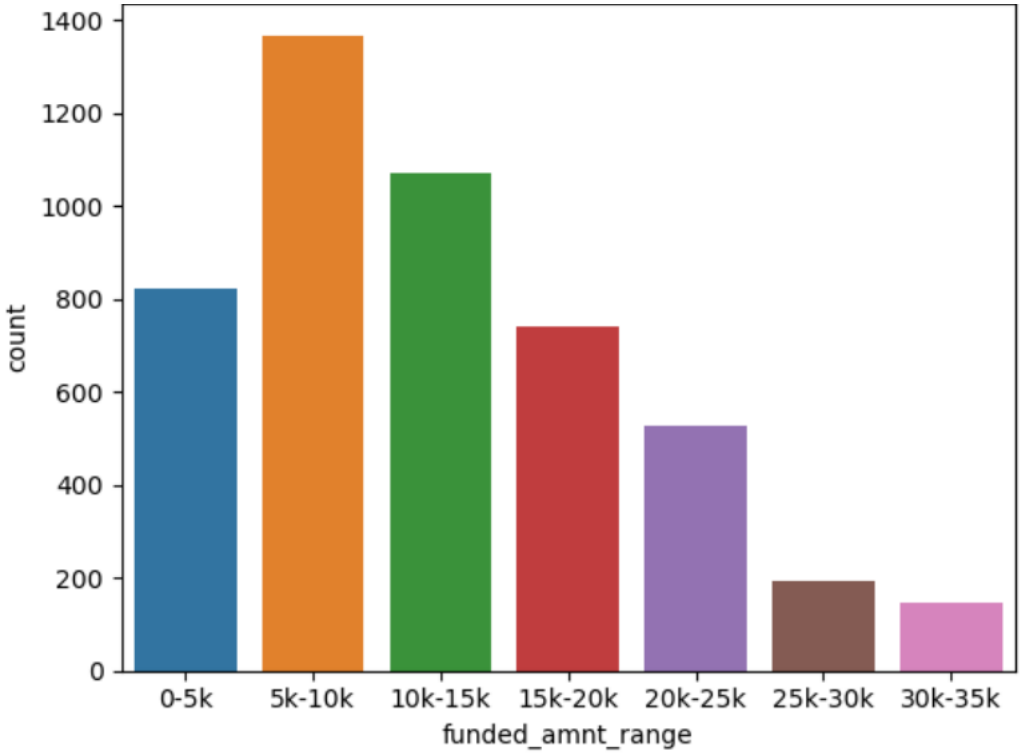


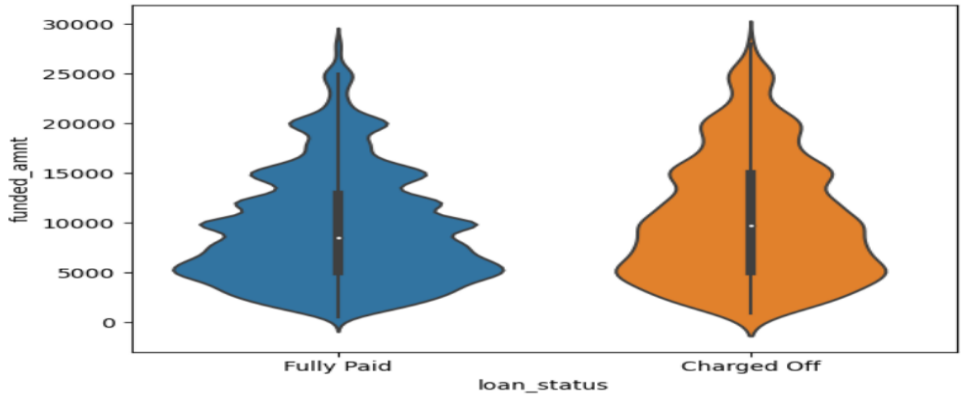
SEGMENTED UNIVARIATE ANALYSIS

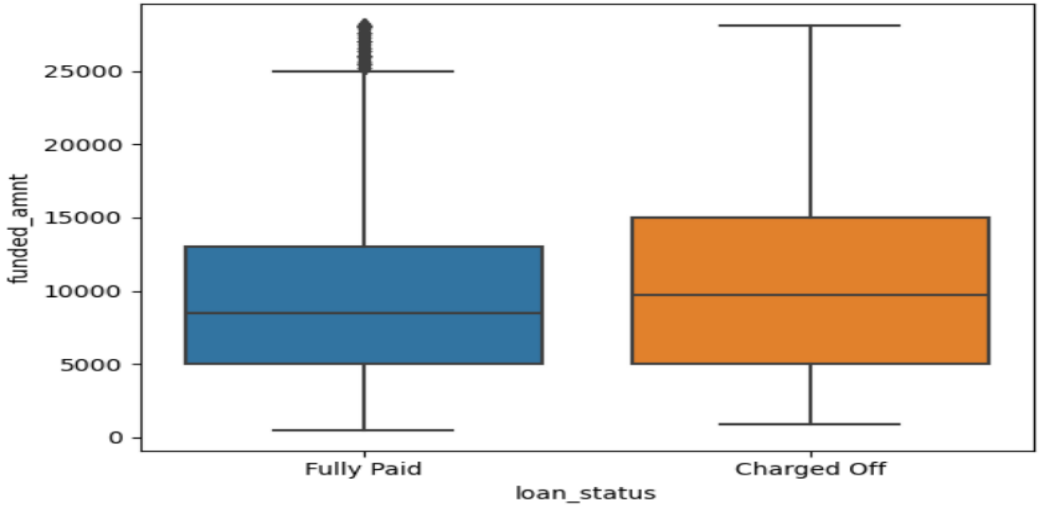
funded_amnt

Observation:

Slightly more chances of default if the funded amount is greater than 13000





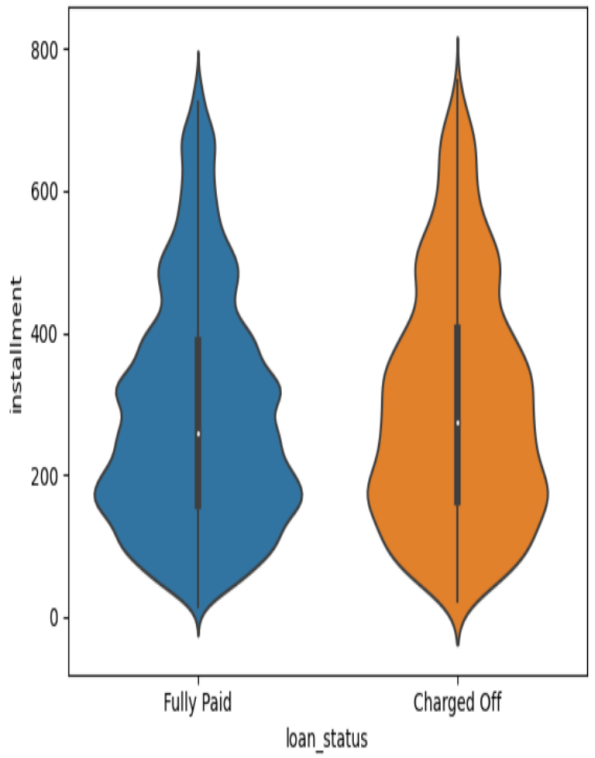


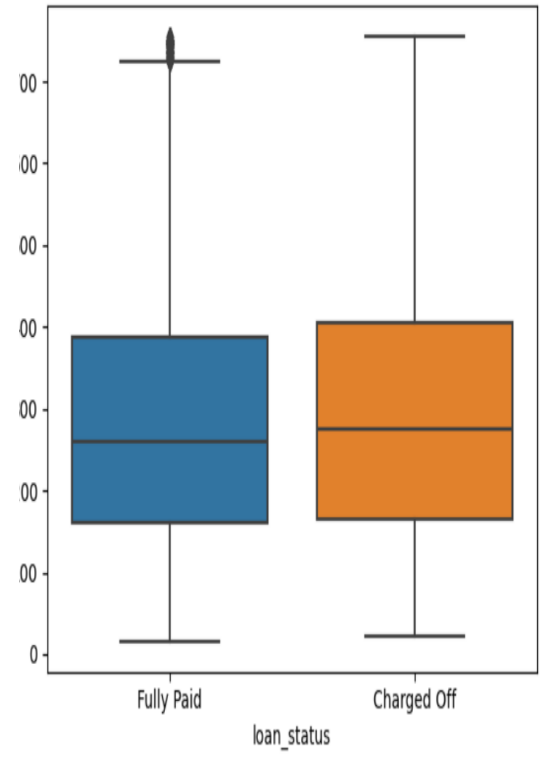
SEGMENTED UNIVARIATE ANALYSIS

installment

Observation:

No significant difference of spread of defaulted/paid loans wrt installment



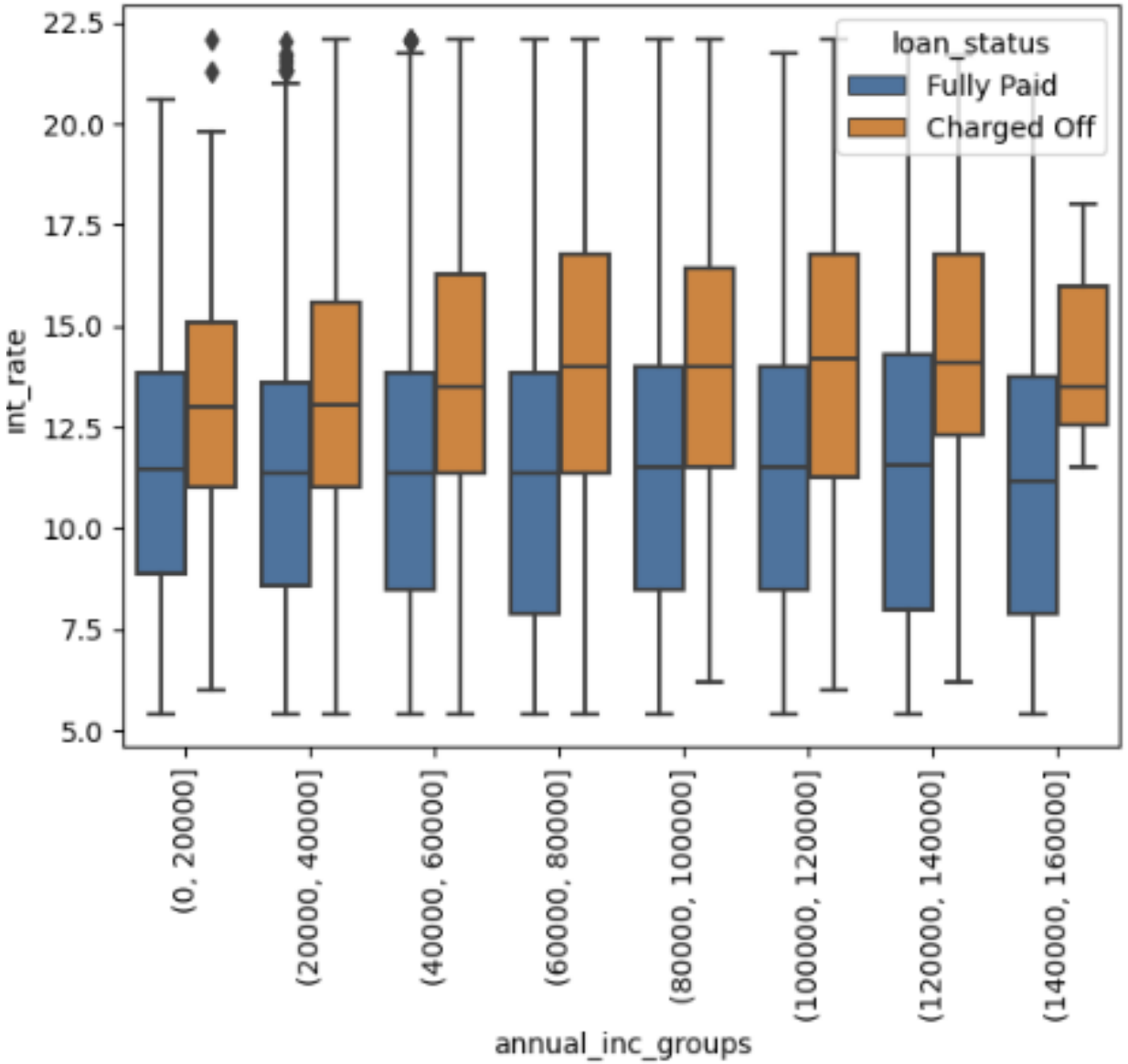


BIVARIATE ANALYSIS

interest vs annual_inc

Observation:

apparently more chances of defaults across all salary groups if rate of interest is greater than 13.85



BIVARIATE ANALYSIS

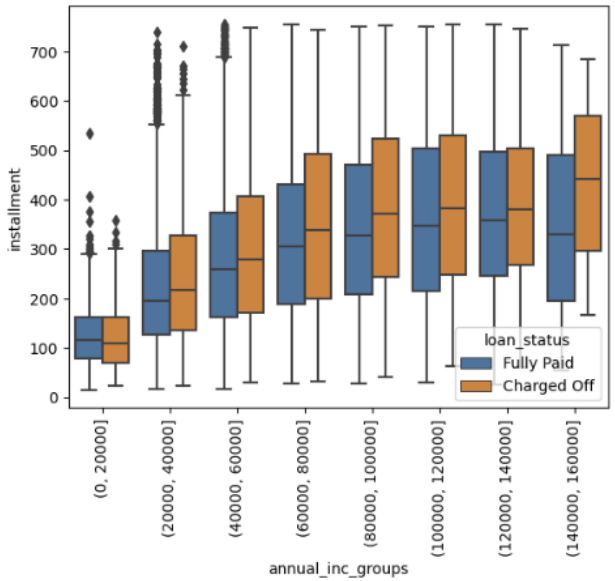
installment vs annual_inc

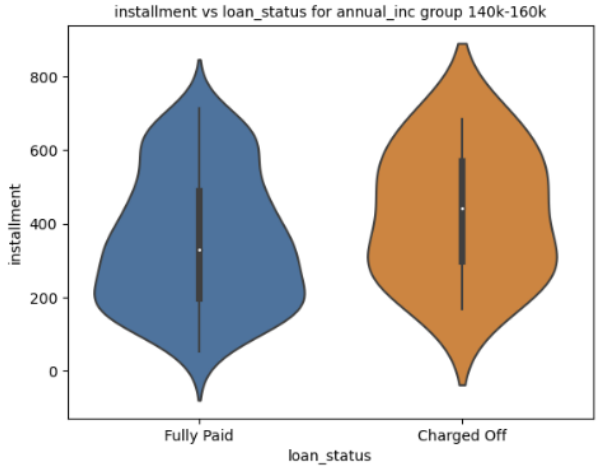
Observation:

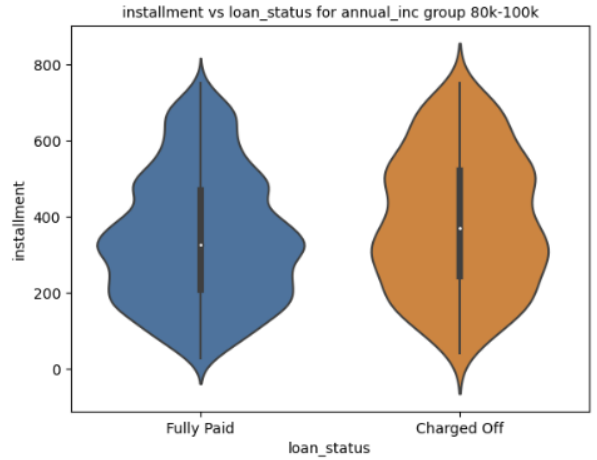
For annual_inc 140k-160k , installment>490 then there are more chances of defaulting

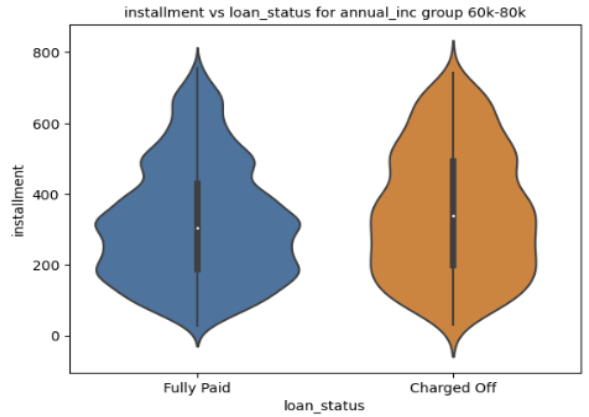
For annual_inc 80k-100k , installment>470 then there are more chances of defaulting

For annual_inc 60k-80k , installment>430 then there are more chances of defaulting







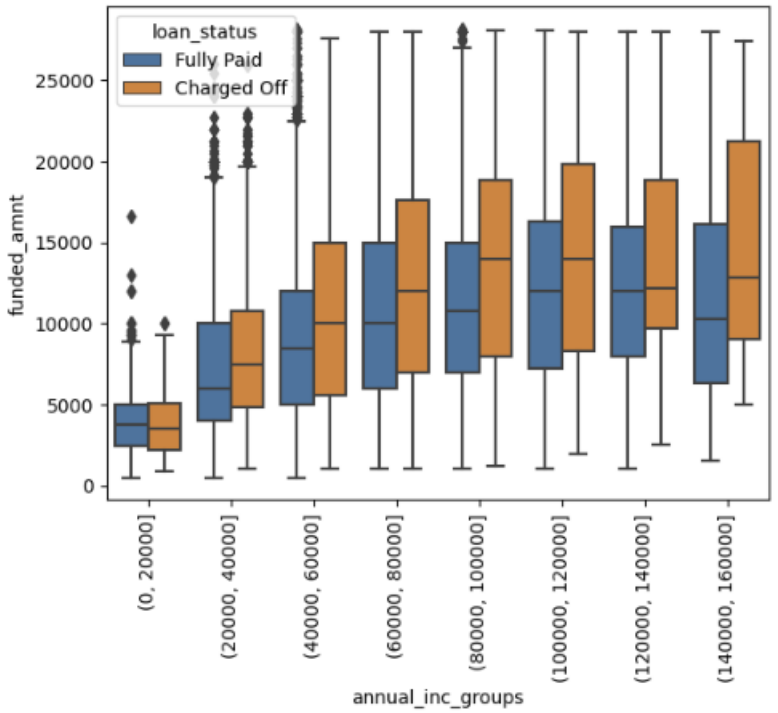


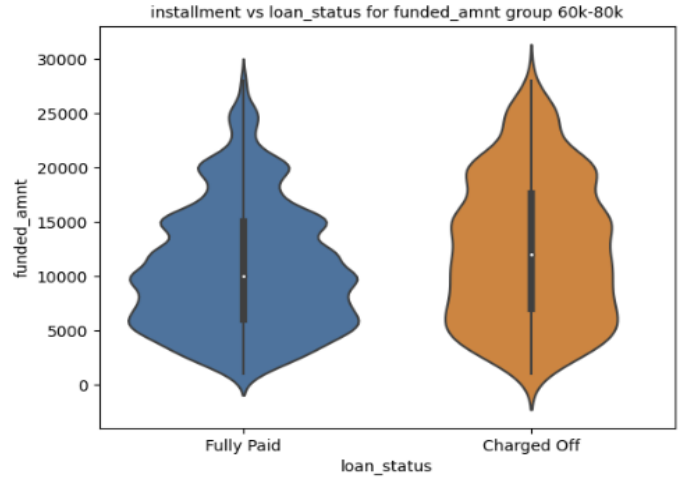
BIVARIATE ANALYSIS

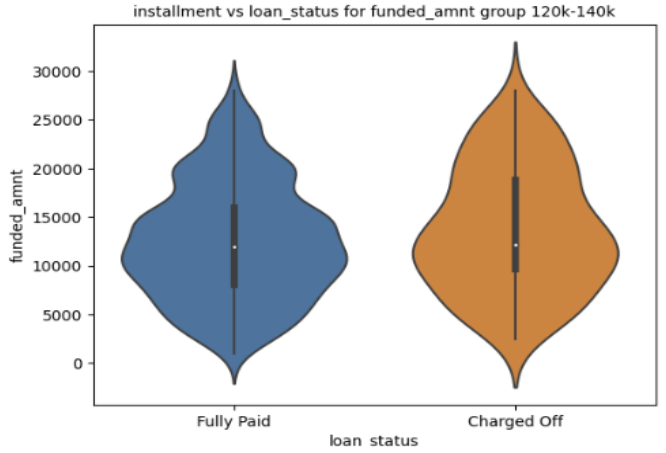
funded_amnt vs annual_inc

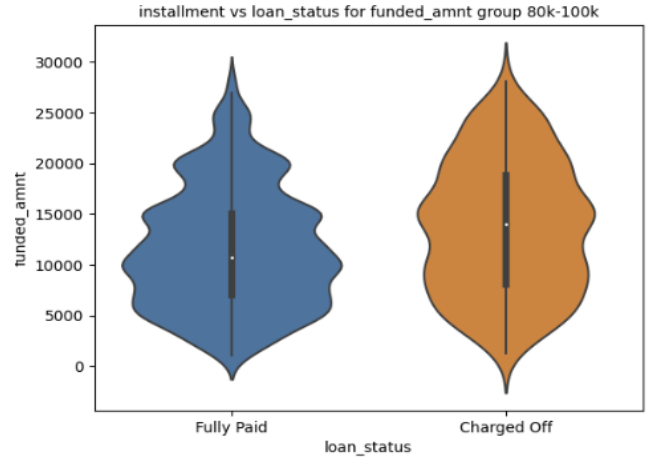
Observations:

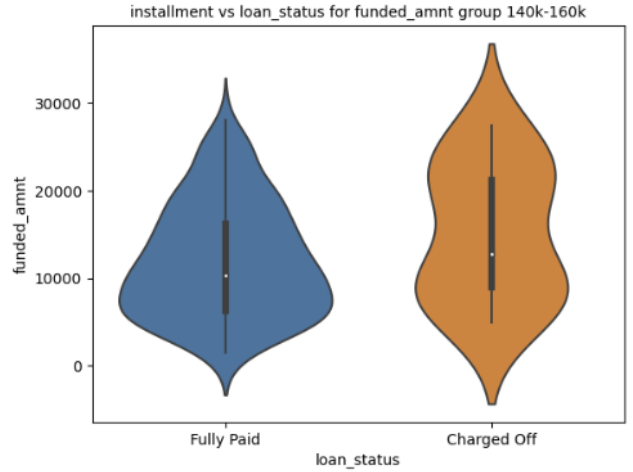
1. for salary groups of 40k-60k apparently more chances of failure if funded_amnt is greater than 12000
2. for salary groups of 60k-80k,80k-100k apparently more chances of failure if funded_amnt is greater than 15000
3. for salagy group of 120k-140k apparently more chances of failure if funded_amnt is greater than 16275
4. for salary group of 140k-160k apparently more chances of failure if funded_amnt is greater than 16500











IMPORTANT CONCLUSIONS

1. Percentagewise grade B3 to G4 and onwards have very high defaults rates (Assuming >10% default rate is very high)
2. Loan for term of '60 months' has 25% defaults, thus very high default rate(Assuming >10% default rate is very high)
3. Interest rate of 10% and higher have more than 13% defaults, so very high default rates
4. Apparently more chances of defaults across all salary groups if rate of interest is greater than 13.85
5. Comparatively, purpose "wedding,major_purchase, credit_card, car, home_improvement" have lower rates of defaulting ,"vacation,debt_consolidation, moving, other,educational, house" have medium rates of default and "renewable_energy, small_business" have very higher rates of default
6. For annual_inc 140k-160k , installment>490 then there are more chances of defaulting,for annual_inc 80k-100k , installment>470 then there are more chances of defaulting,for annual_inc 60k-80k , installment>430 then there are more chances of defaulting
7. For salary groups of 40k-60k apparently more chances of defaults if funded_amnt is greater than 12000, for salary groups of 60k-80k,80k-100k apparently more chances of failure if funded_amnt is greater than 15000, for salagy group of 120k-140k apparently more chances of failure if funded_amnt is greater than 16275, for salary group of 140k-160k apparently more chances of failure if funded_amnt is greater than 16500

Thank You..!!