

LENDING CLUB CASE STUDY PRESENTATION

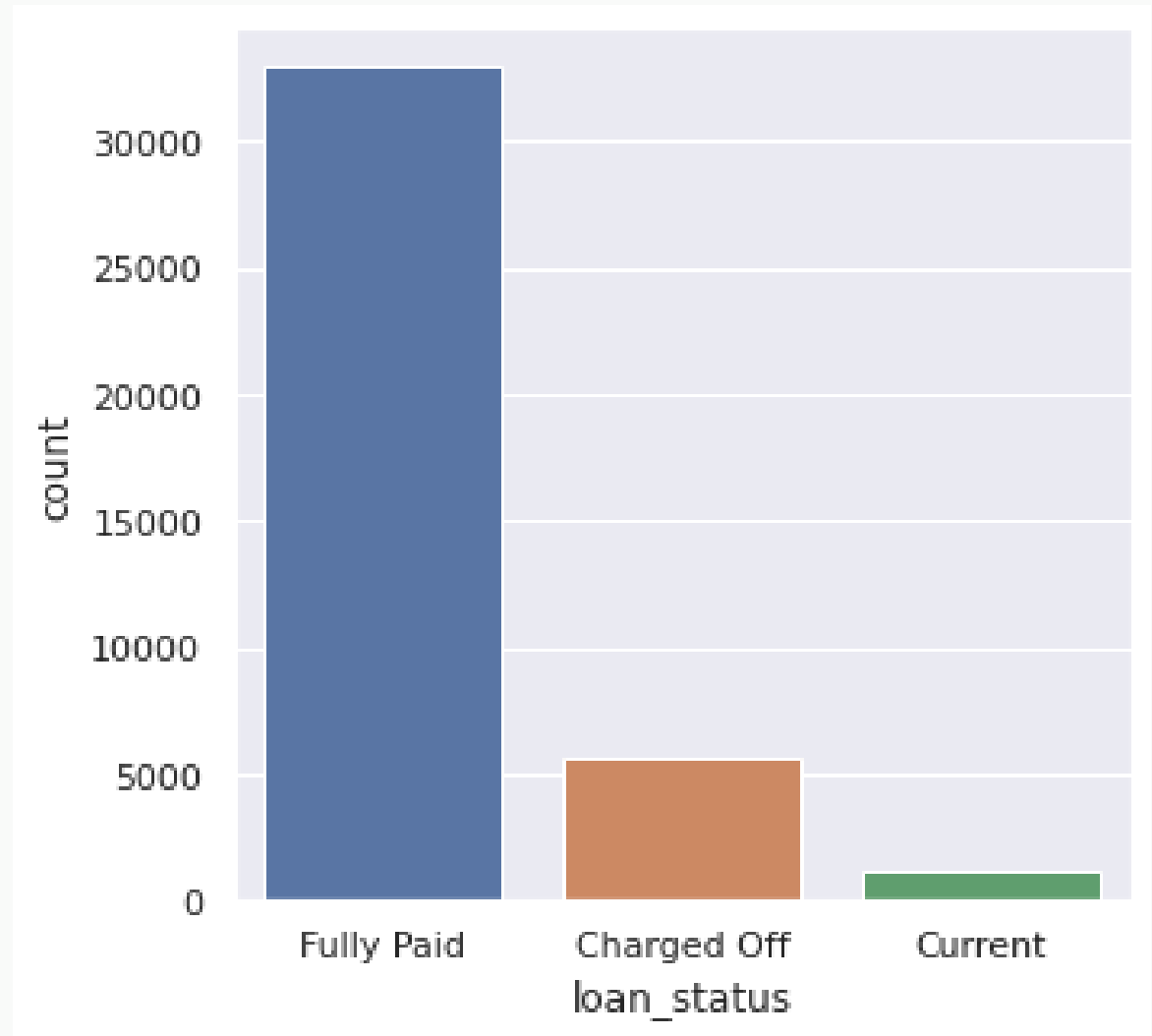
By Angat Sitoula

Problem Statement

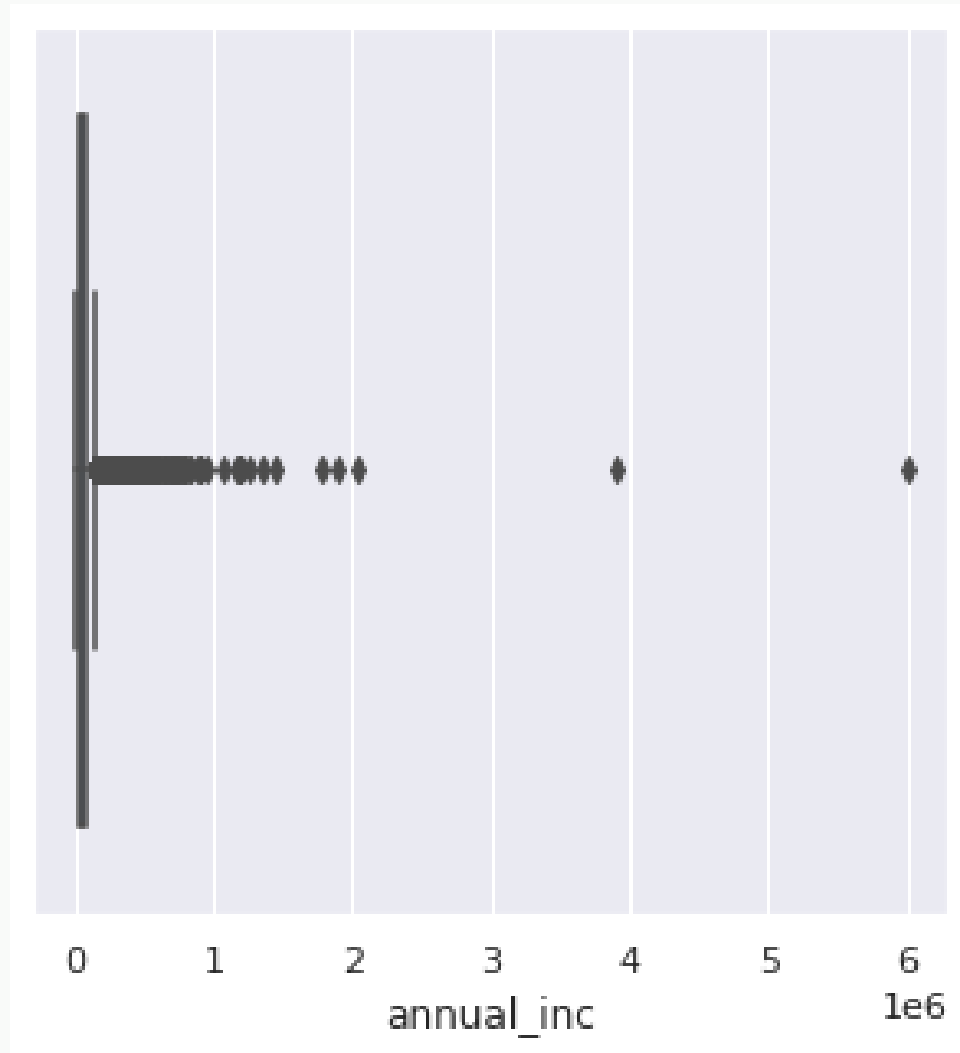
Lending Club is a consumer finance company that provides finance to urban people. They have lots of data from past consumers who have taken the loan. So we are performing exploratory data analysis to understand the factors that impact the loan default rate.

UNIVARIANT ANALYSIS

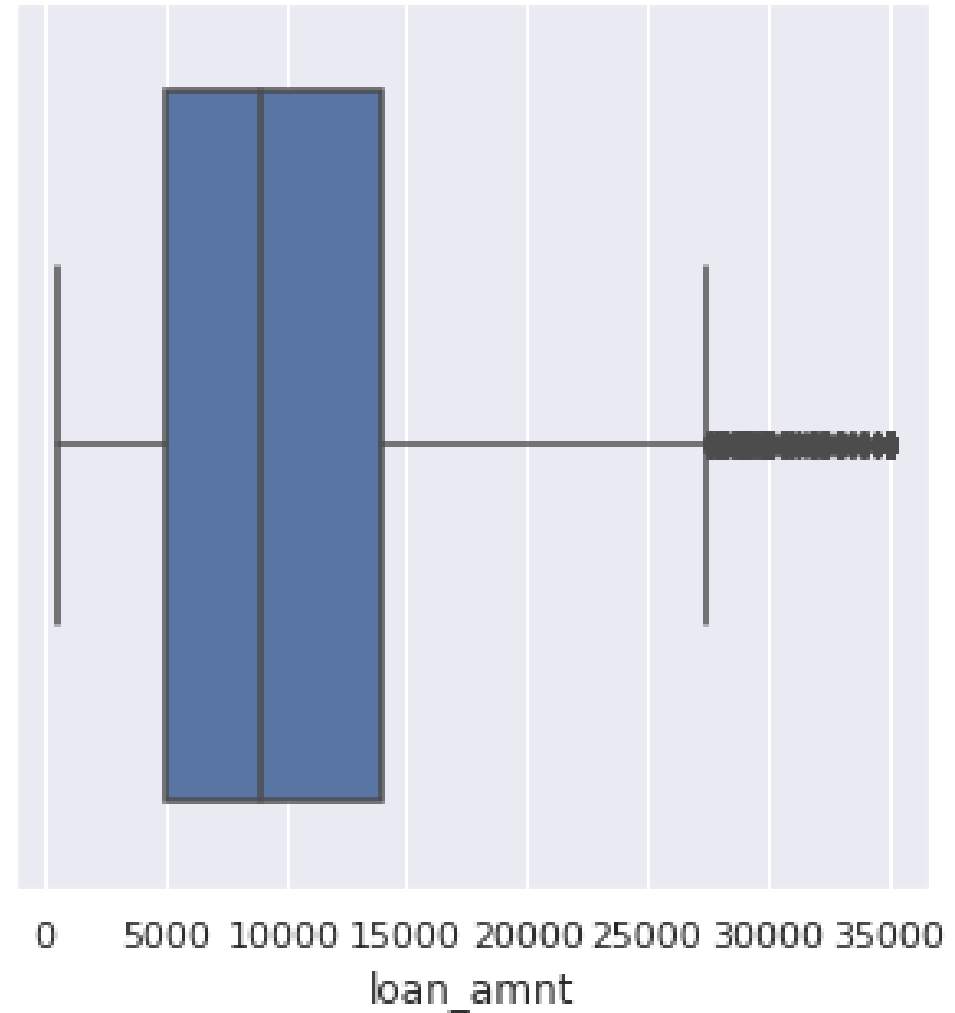
MOST OF THE LOAN HAS BEEN FULLY PAID ON THE GIVEN DATASET. AROUND 5000 OF THE LOAN ARE CHARGED OFF.



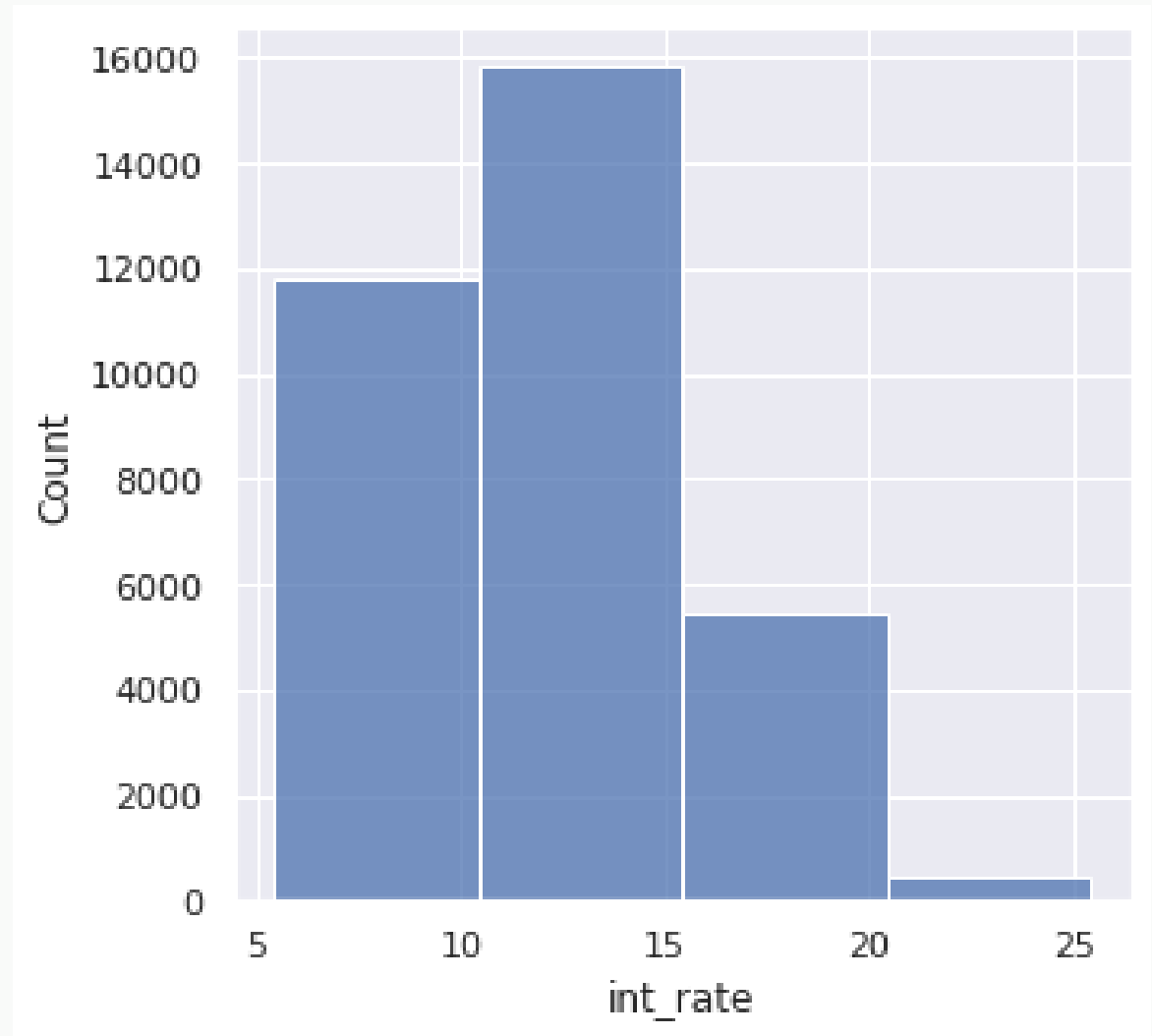
THERE HAVE BEEN
SIGNIFICANT OUTLIERS
IN THE ANNUAL
INCOME. SO IT IS
NECESSARY TO HANDLE
THE OUTLIER.



THERE IS OUTLIER
IN LOAN AMOUNT. SO
IT IS NECESSARY TO
REMOVE THOSE
OUTLIER.



MOST PEOPLE WHO TOOK
LOANS HAVE AN INTEREST
RATE BETWEEN 10-15%.
ONLY A FEW PEOPLE HAS
ABOVE 20% INTEREST
RATE



MOST OF THE PEOPLE
TOOK LOANS FOR
DEBT_CONSOLIDATION.
VERY FEW PEOPLE TOOK
LOAN FOR
RENEWABLE_ENERGY.

debt_consolidation	16012
credit_card	4236
other	3466
home_improvement	2148
major_purchase	1872
small_business	1463
car	1361
wedding	813
medical	585
moving	500
vacation	341
house	308
educational	282
renewable_energy	82

Name: purpose, dtype: int64

GRADE B HAS THE
MOST NUMBER OF
LOAN AND G HAS
LEAST ONE.

B	10261
A	8642
C	6896
D	4486
E	2229
F	757
G	198

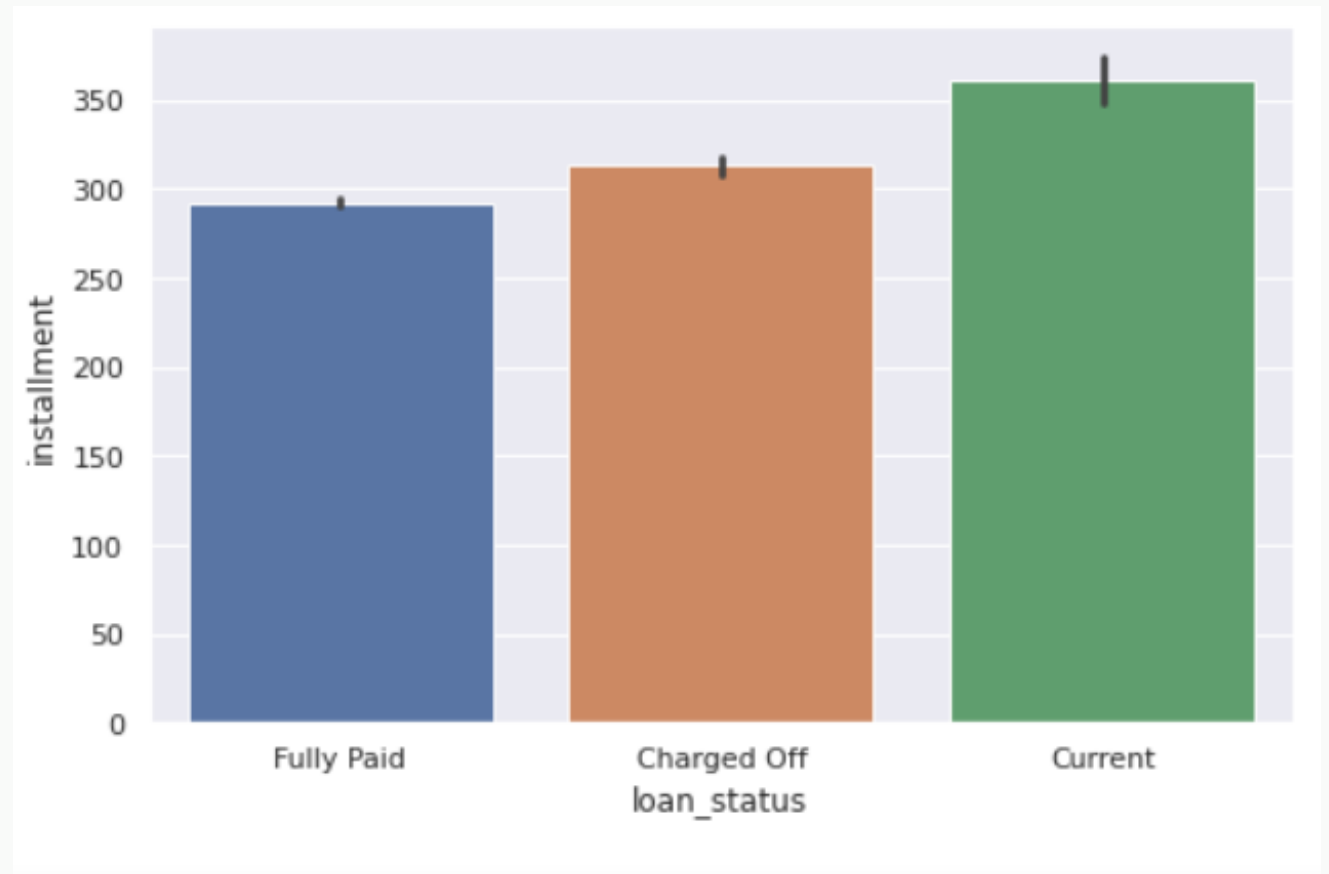
Name: grade, dtype: int64

BIVARIATE ANALYSIS

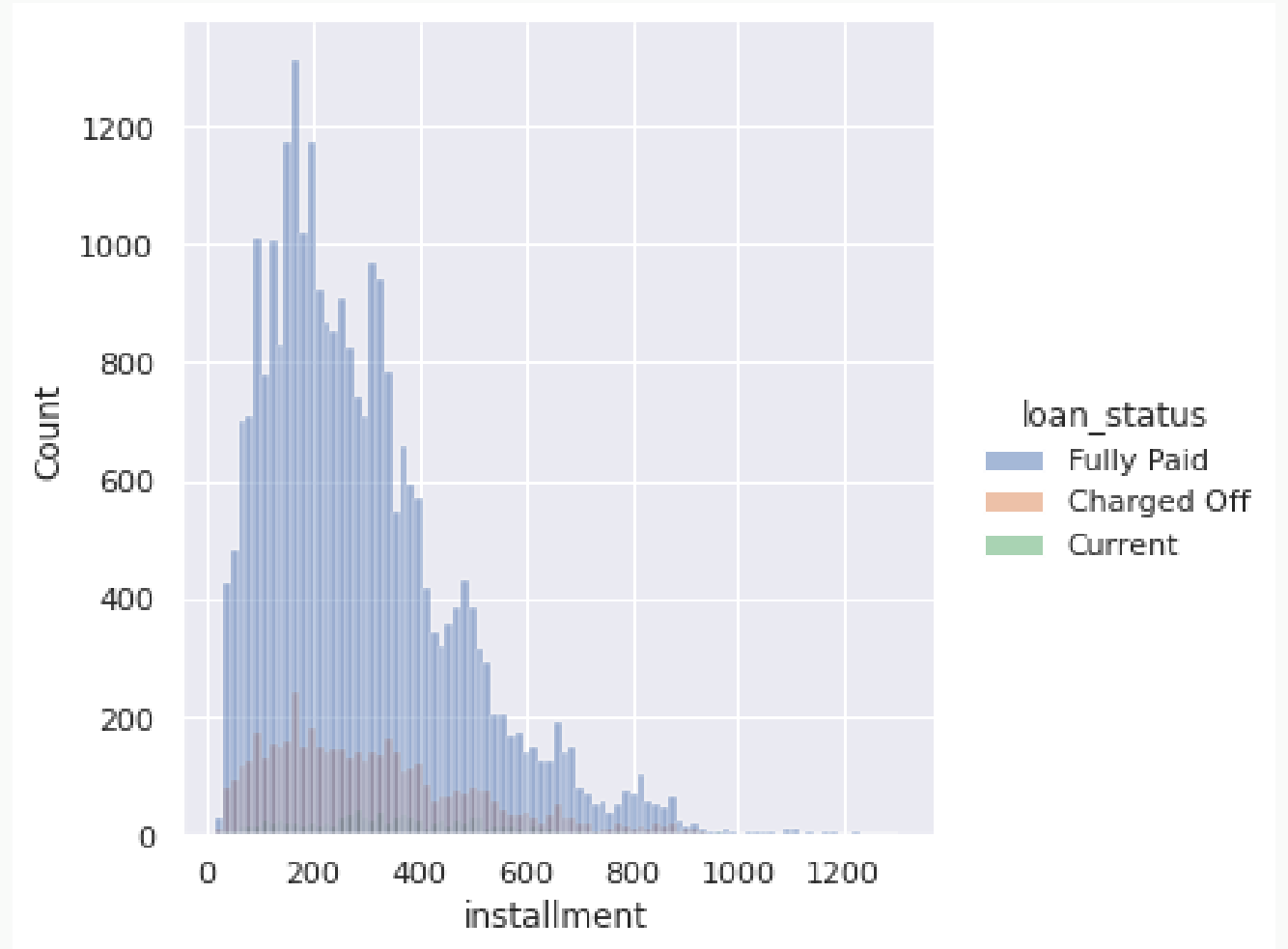
CORRELATION CHART
SHOWS THAT LOAN
AMOUNT AND
INSTALLMENT HAS THE
HIGHEST CORRELATION.

	loan_amnt	int_rate	installment	annual_inc	dti	open_acc	pub_rec	revol_bal	total_acc	revol_util
loan_amnt	1.000000	0.279631	0.928961	0.382285	0.114187	0.155467	-0.043981	0.316971	0.225317	0.070047
int_rate	0.279631	1.000000	0.257743	0.048576	0.100838	-0.023610	0.105345	0.065992	-0.083376	0.466805
installment	0.928961	0.257743	1.000000	0.375451	0.100189	0.147493	-0.039664	0.307235	0.194804	0.103396
annual_inc	0.382285	0.048576	0.375451	1.000000	-0.024832	0.257383	0.003901	0.351744	0.364442	0.044993
dti	0.114187	0.100838	0.100189	-0.024832	1.000000	0.303535	-0.006606	0.299241	0.259278	0.267325
open_acc	0.155467	-0.023610	0.147493	0.257383	0.303535	1.000000	0.007955	0.278128	0.680481	-0.110505
pub_rec	-0.043981	0.105345	-0.039664	0.003901	-0.006606	0.007955	1.000000	-0.057176	-0.010081	0.060248
revol_bal	0.316971	0.065992	0.307235	0.351744	0.299241	0.278128	-0.057176	1.000000	0.297842	0.322703
total_acc	0.225317	-0.083376	0.194804	0.364442	0.259278	0.680481	-0.010081	0.297842	1.000000	-0.094541
revol_util	0.070047	0.466805	0.103396	0.044993	0.267325	-0.110505	0.060248	0.322703	-0.094541	1.000000

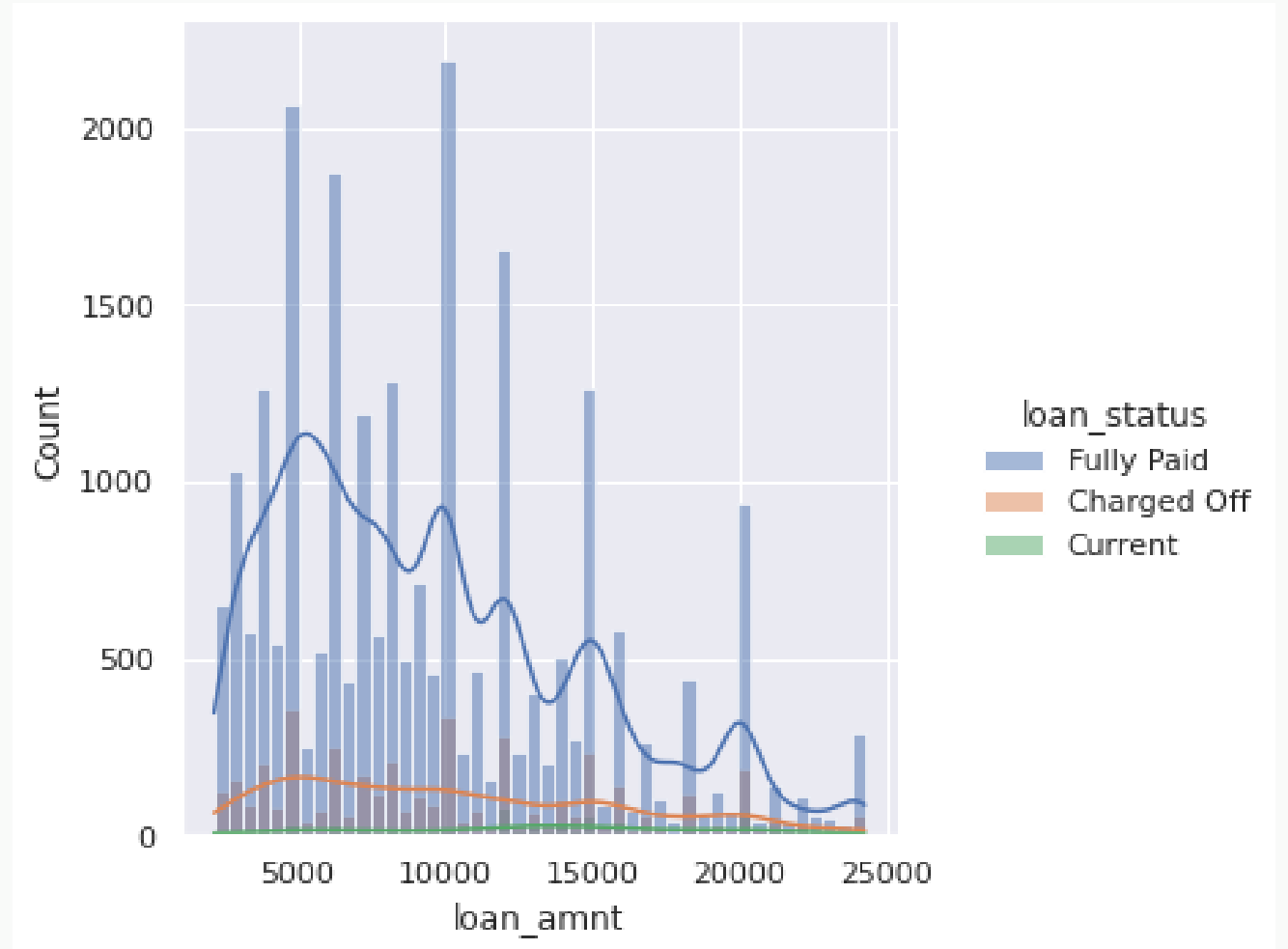
AS THE INSTALMENT
INCREASES, CHARGED-
OFF LOAN STATUS
INCREASES, AND THE
FULLY PAID RATE
DECREASES.



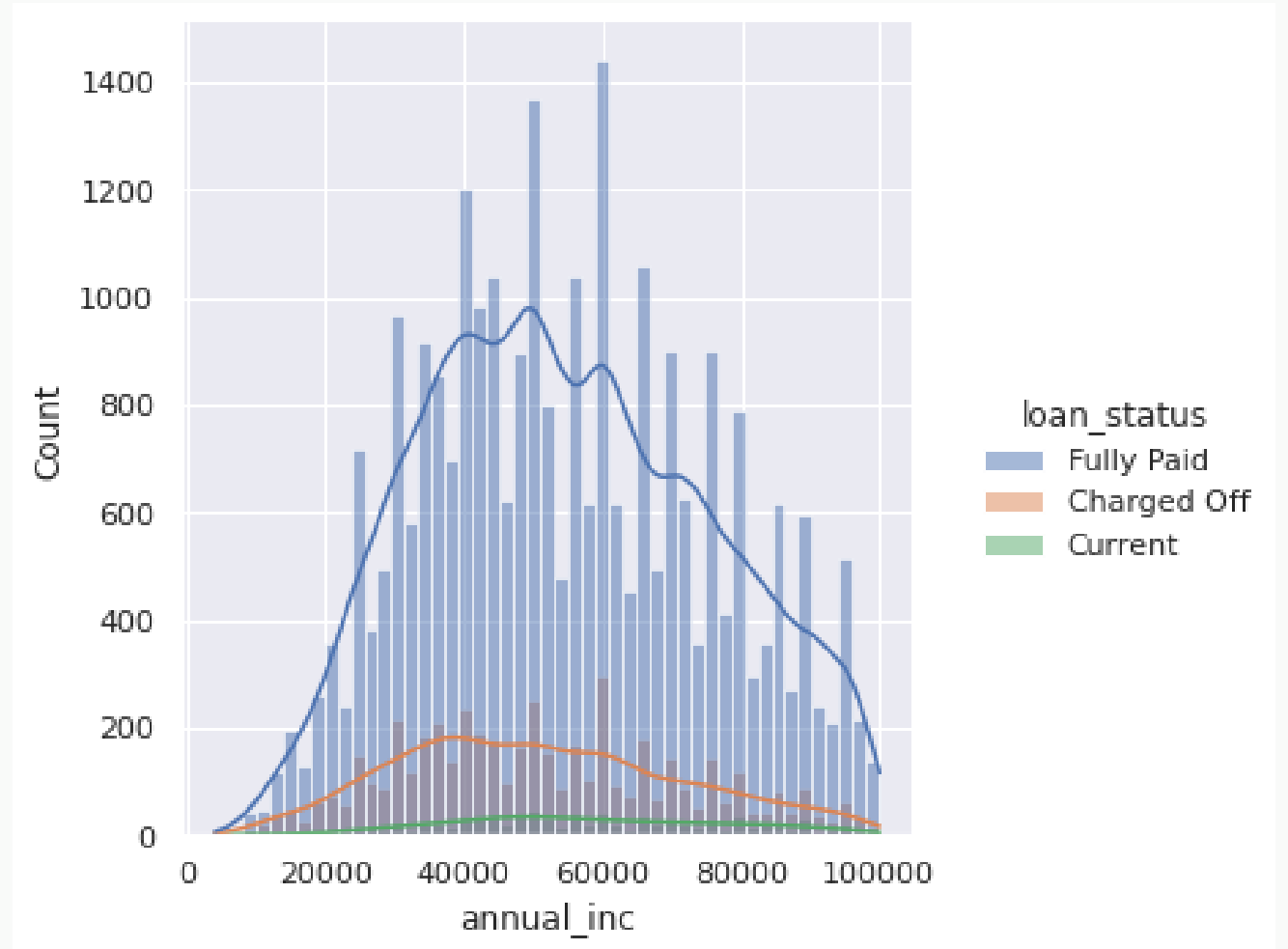
DISTRIBUTION OF
INSTALLMENT WITH
EACH
LOAN_STATUS
CLASS.

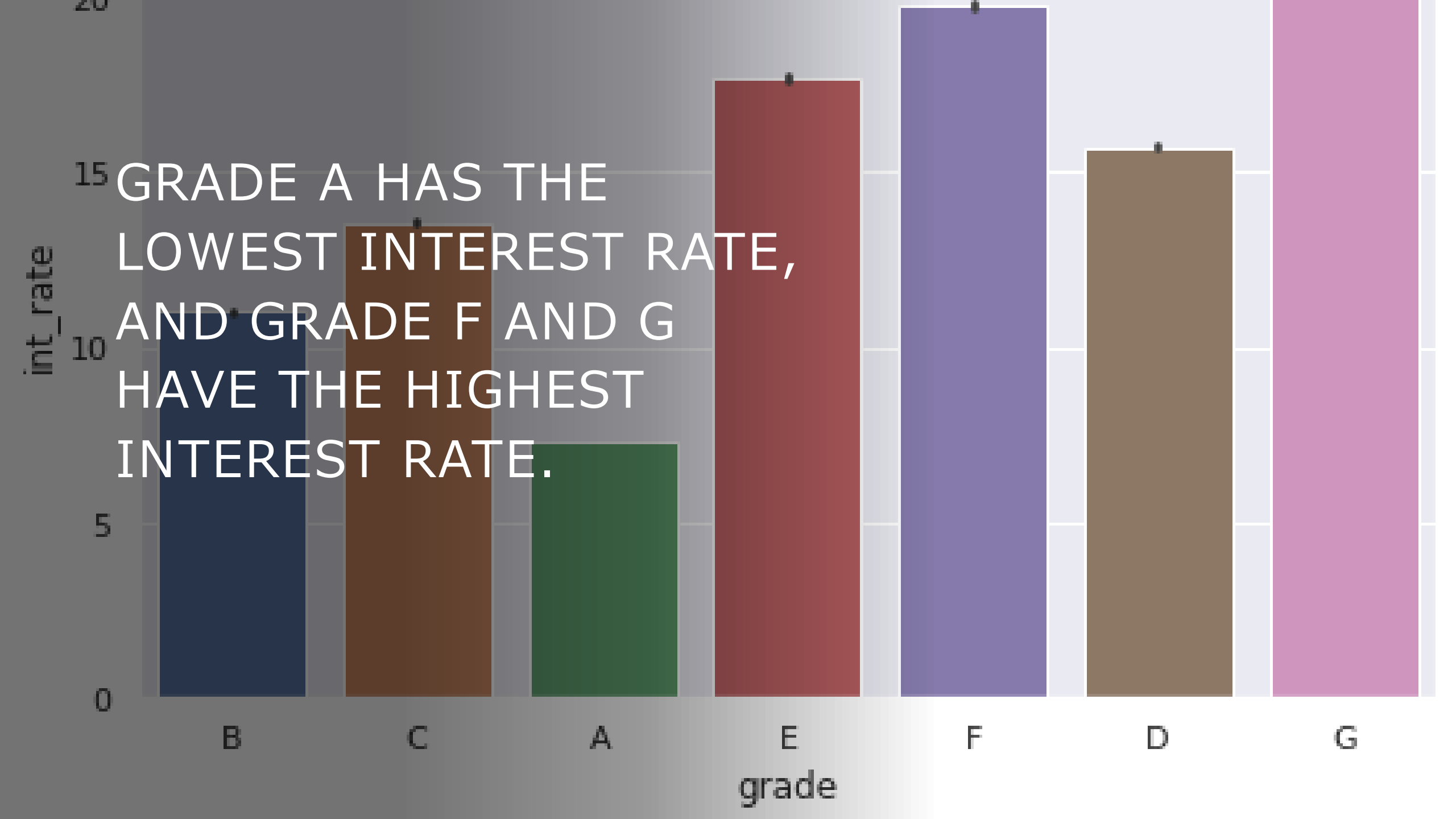


DISTRIBUTION OF LOAN AMOUNT WITH EACH LOAN _STATUS CLASS.

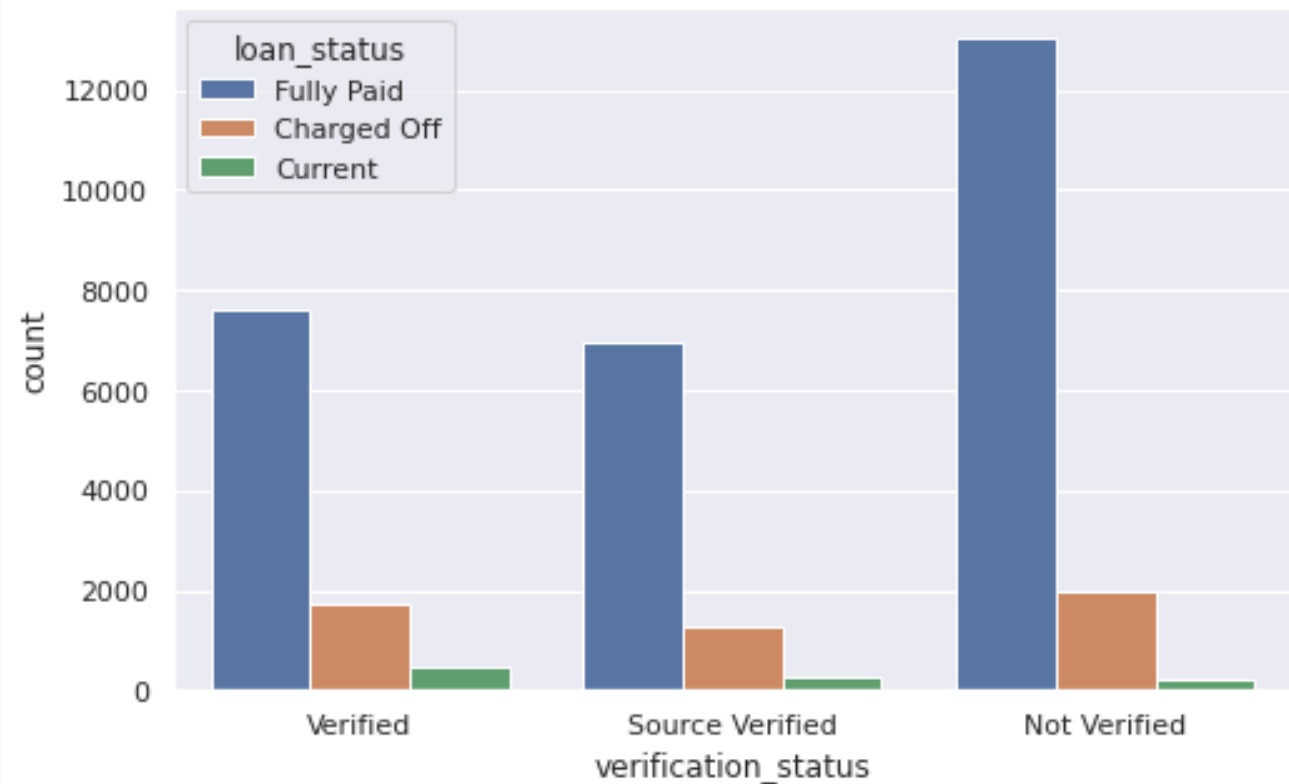


DISTRIBUTION OF
ANNUAL AMOUNT
WITH EACH LOAN_
STATUS CLASS.





COUNT PLOT ON VERIFICATION STATUS AND LOAN STATUS



Key Finding

- 1. Annual income does not have an impact on loan status.**
- 2. The instalment affects loan status. As the instalment increases there is a high chance of the loan being charged off.**
- 3. Interest rate is another factor that impacts the loan status.**
- 4. Most people have a loan with interest between 10-15%.**
- 5. If we look at the term, we find that 36 months have a high loan status charge-off.**
- 6. Verification status is another factor that impacts the loan status. Not Verified has a high charge-off compared to others.**