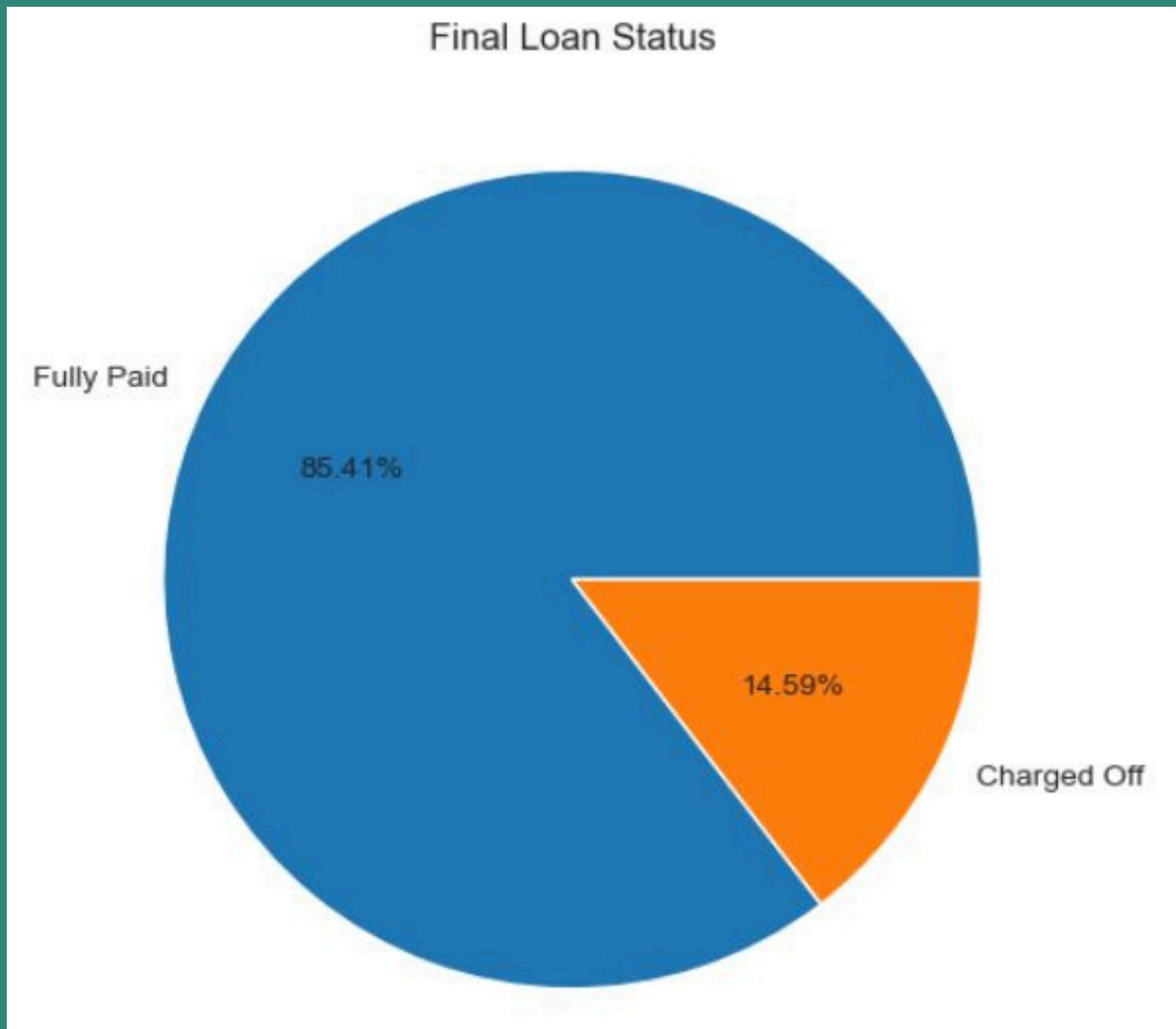# Problem Statement

To predict loan defaults from the loan applications received by the customers.

Significant Variables Considered for Analysis are :

- Loan Status- Charged off and Fully Paid
- Loan Amount
- Interest Rate
- Grade and Sub grade
- DTI (Debt to Income Ratio)
- State of Residence
- Purpose

Bonafides Vs Defaulters

Final Loan Status

Analysing data set we can conclude that 85.41% of total customers fully paid their debt , while 14.59 % were the defaulters

LOANS DISTRIBUTION BY STATE

Loans distributed by State
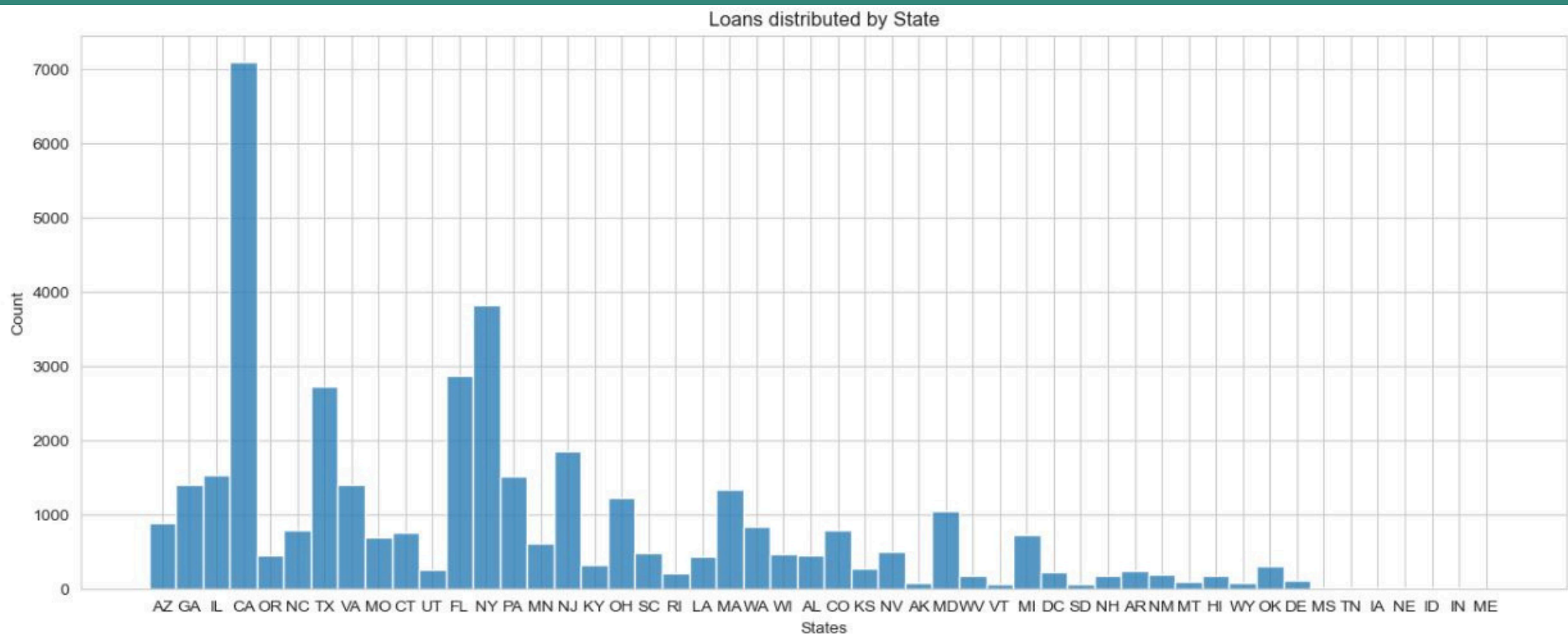
CALIFORNIA STATE HAS THE MAXIMUM BORROWERS , WHILE A INCREASED PRESENCE IS SUGGESTED IN STATES LIKE MS,NE,TN, ETC.

BORROWERS EXPERIENCE

Borrower's Experience
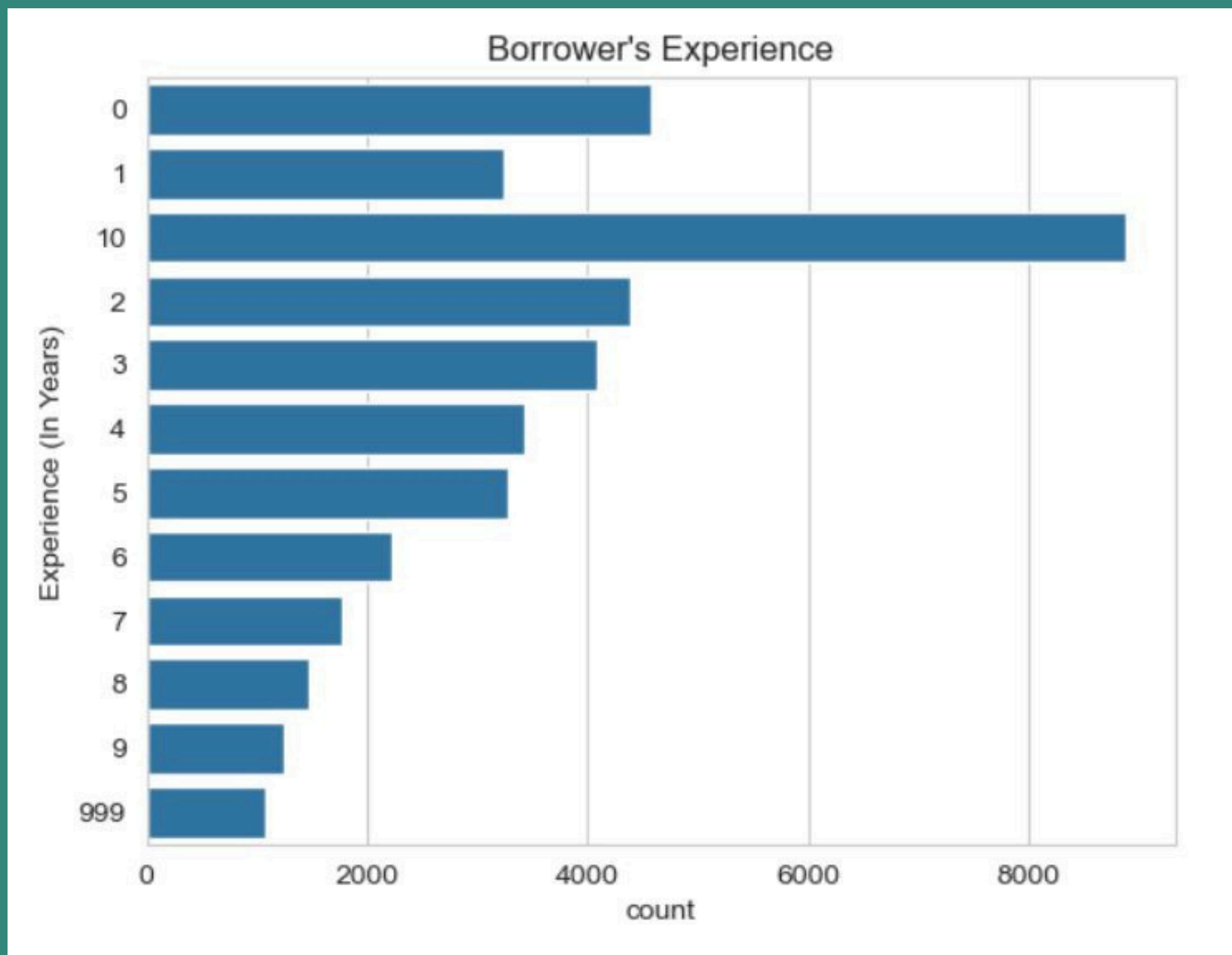
WE CAN SEE THAT MAXIMUM PERCENTAGE OF LOANS ARE GIVEN TO BORROWERS WITH 10+ YEARS OF EXPERIENCE

DEFAULTERS BY STATE

Defaulters by State
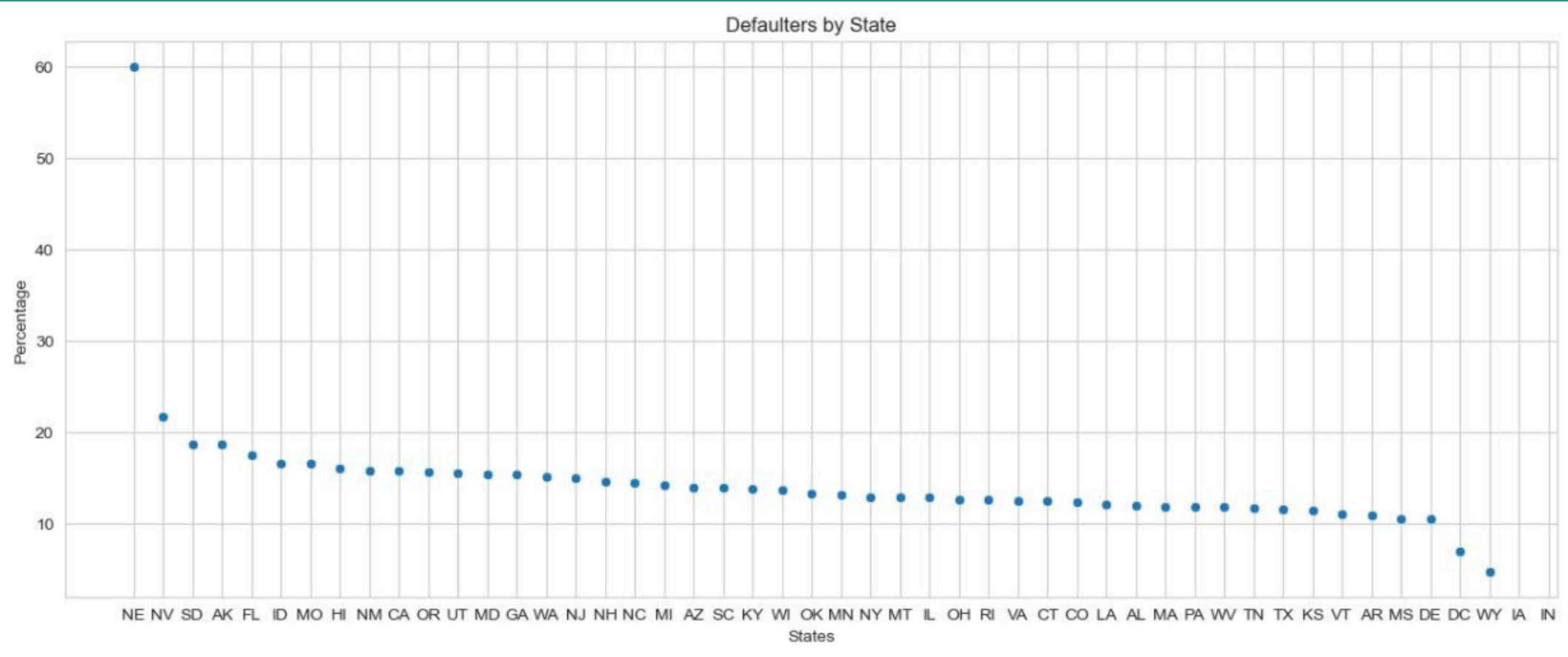
COST OF LIVING CAN VARY BY STATE , SO UNDERSTANDING DEFAULTERS BY STATE CAN BE HELPFUL

Comparision by Loan Amount
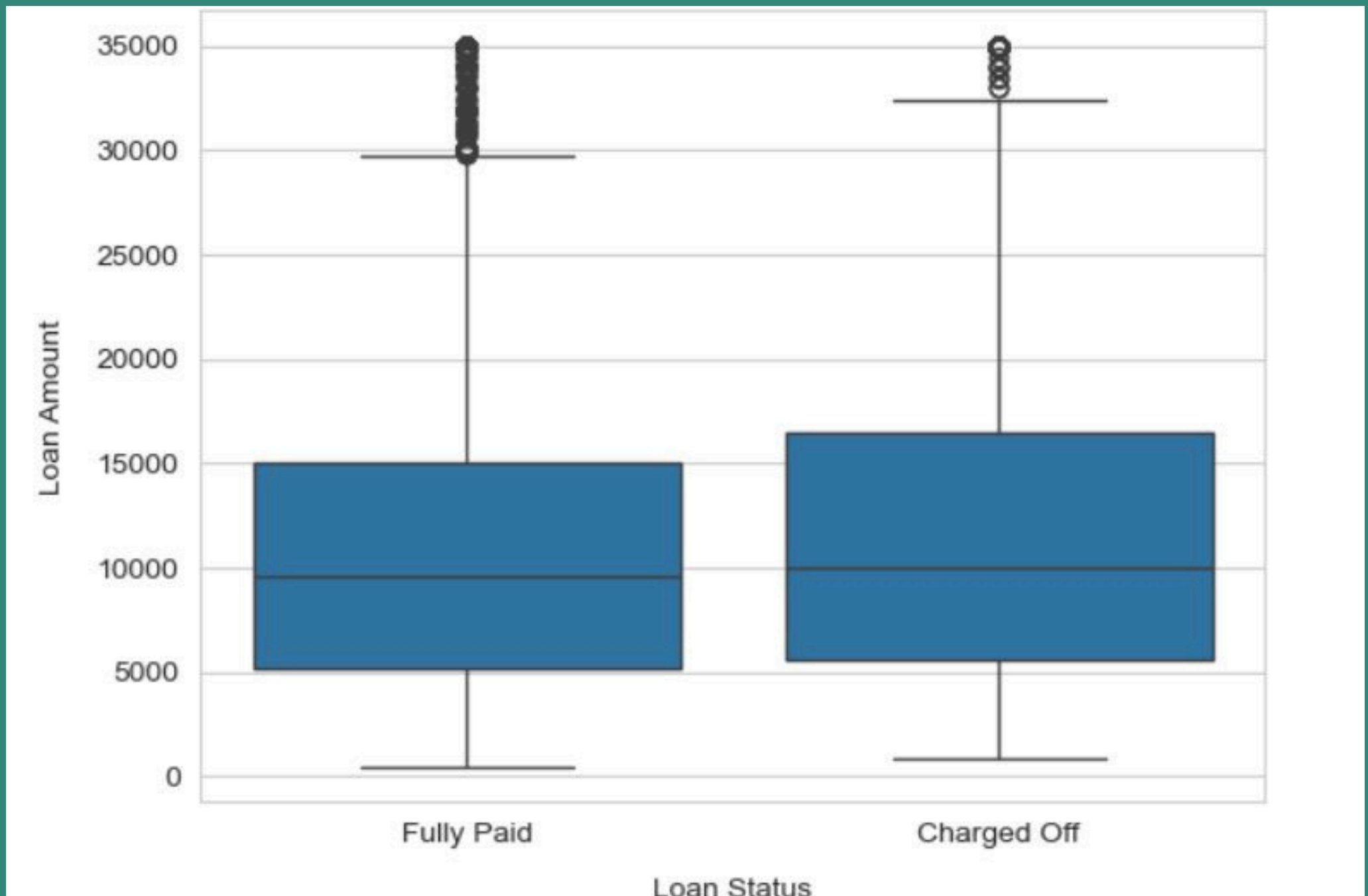
1st quartile and median are nearly same for both Bonafides and Defaulters , while 3rd quartile for defaulters is slightly above from that of bonafides. Thus, reducing upper limit loan amount is advisable.

COMPARING GRADES
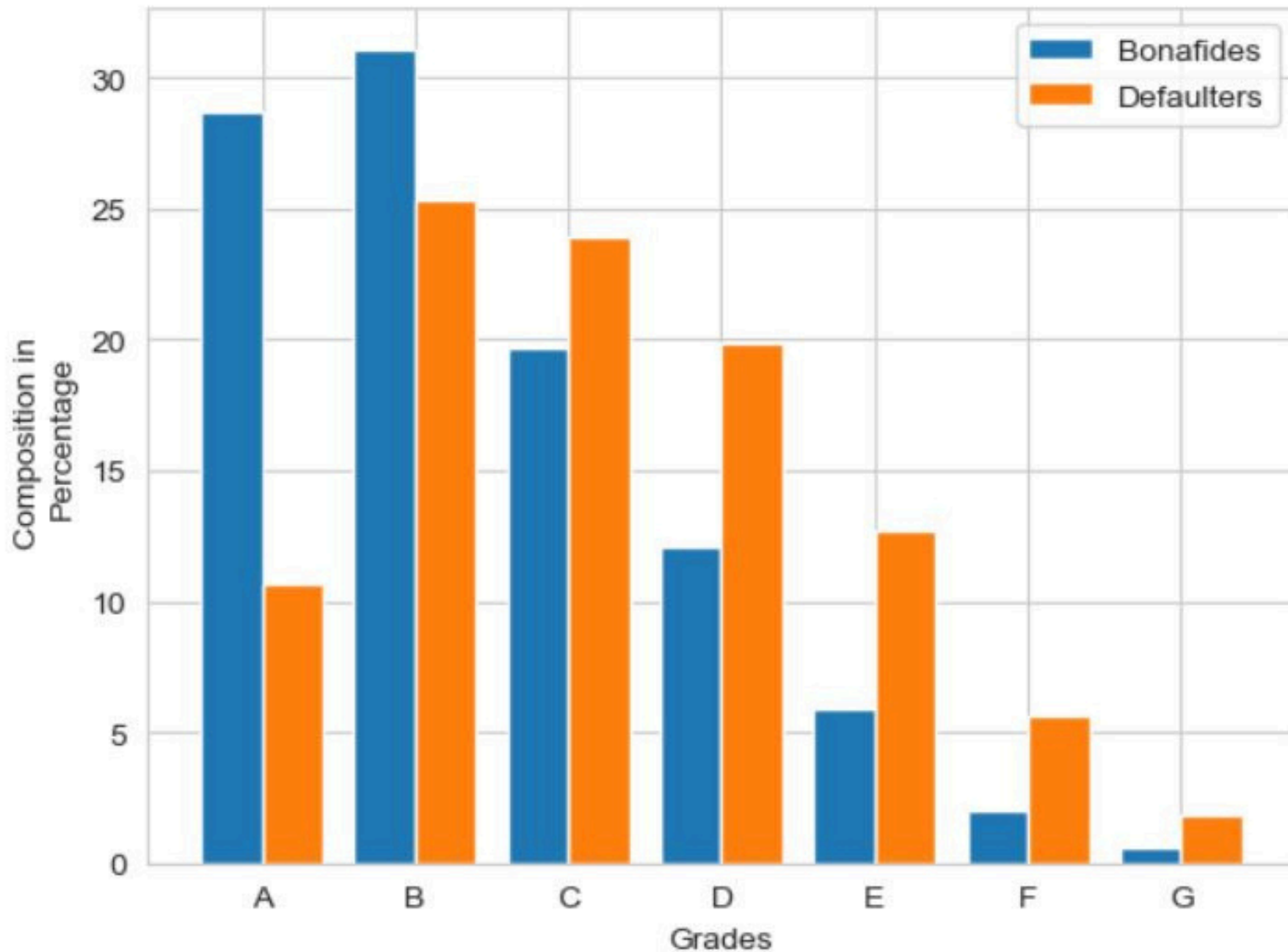
It is clear that customers with grades A, B, C, and D have a greater chance of obtaining a loan. An in-depth analysis of these grades can significantly assist in forecasting potential defaulters.

INTEREST RATES
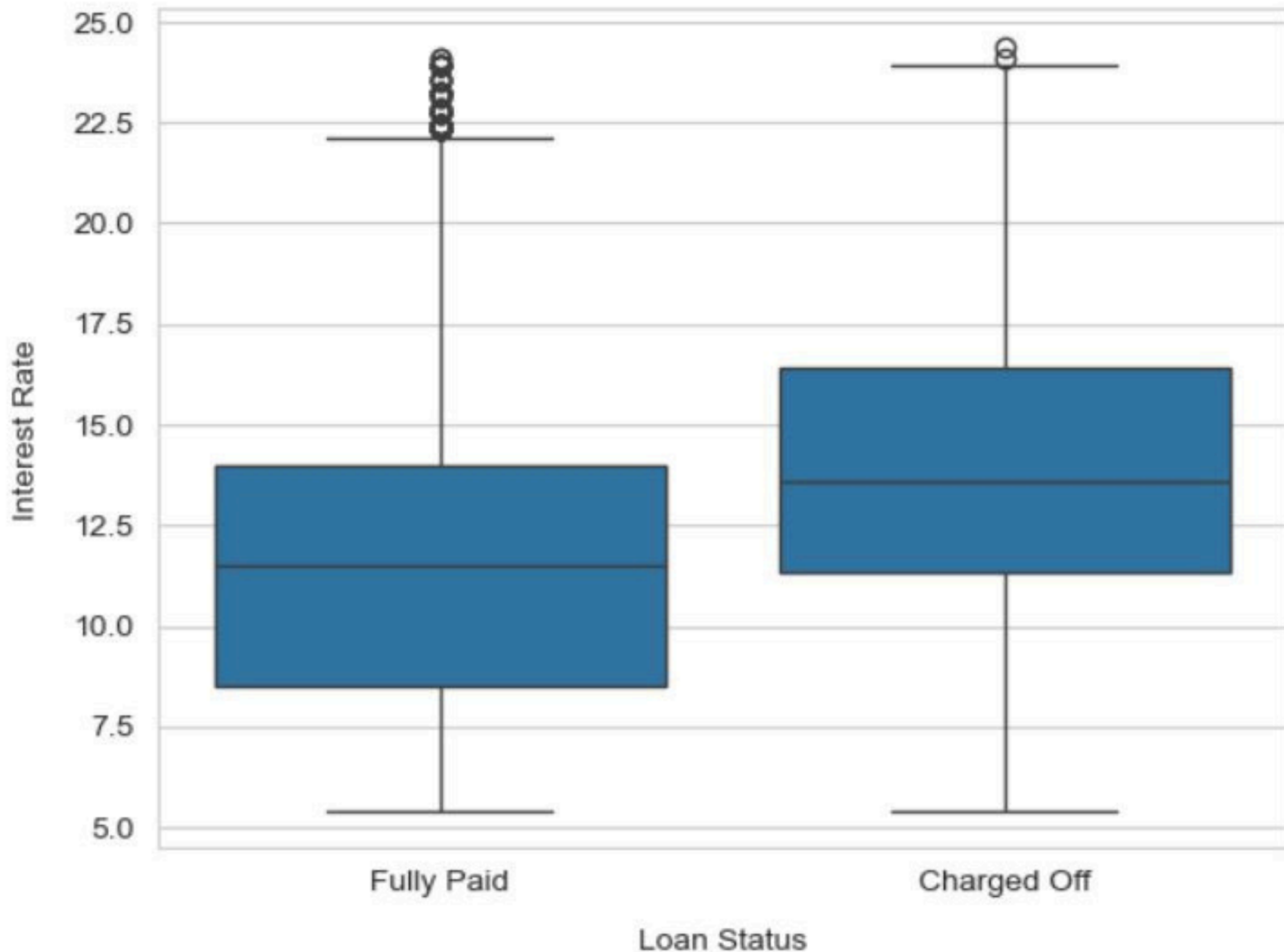
The boxplot indicates that the median of defaulters and the third quartile of bonafides are almost identical, implying that higher interest rates are correlated with an increased likelihood of defaulting.

HEAT MAP

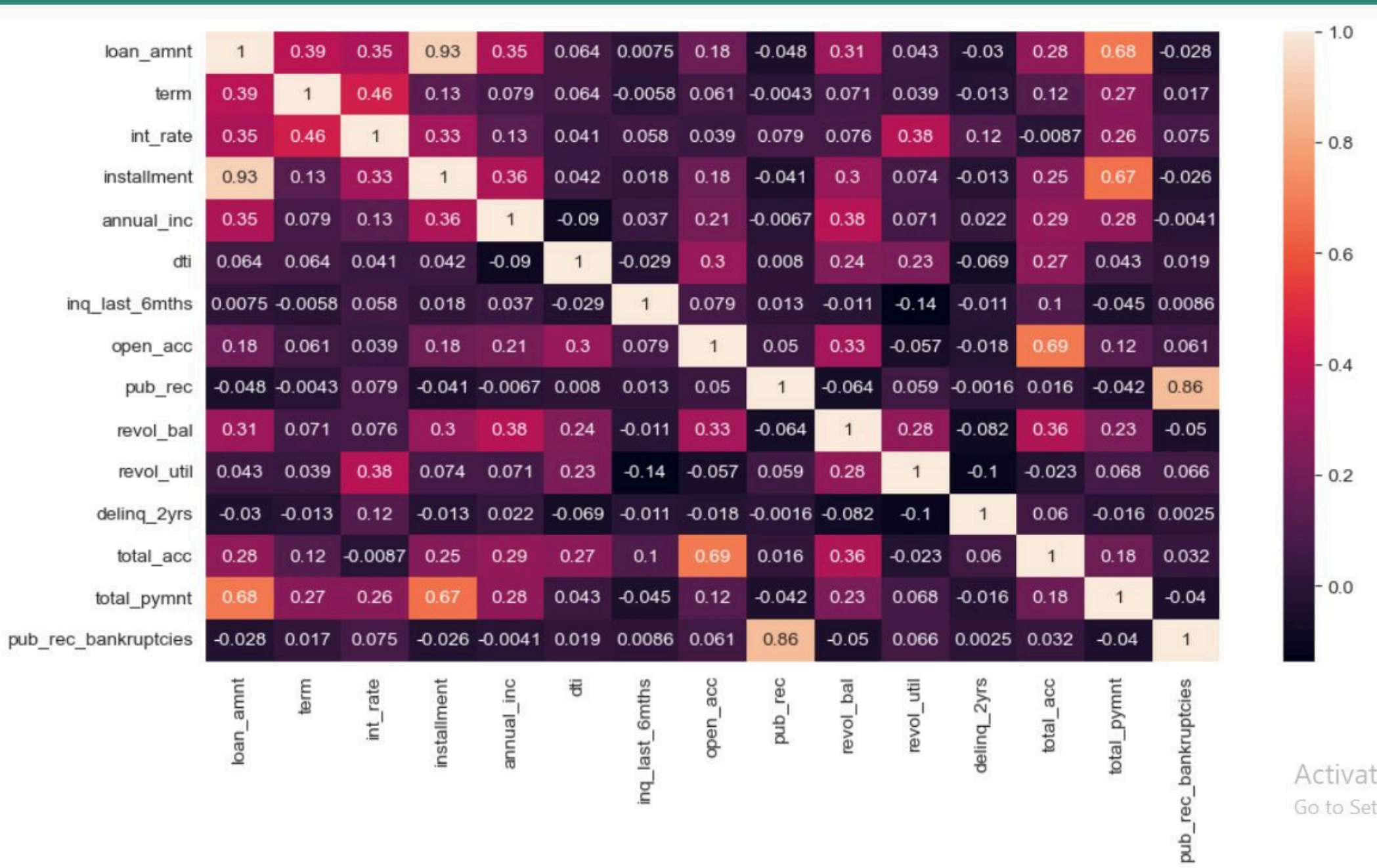|	loan_amnt	term	int_rate	installment	annual_inc	dti	inq_last_6mths	open_acc	pub_rec	revol_bal	revol_util	delinq_2yrs	total_acc	total_pymnt	pub_rec_bankruptcies
loan_amnt	1	0.39	0.35	0.93	0.35	0.064	0.0075	0.18	-0.048	0.31	0.043	-0.03	0.28	0.68	-0.028
term	0.39	1	0.46	0.13	0.079	0.064	-0.0058	0.061	-0.0043	0.071	0.039	-0.013	0.12	0.27	0.017
int_rate	0.35	0.46	1	0.33	0.13	0.041	0.058	0.039	0.079	0.076	0.38	0.12	-0.0087	0.26	0.075
installment	0.93	0.13	0.33	1	0.36	0.042	0.018	0.18	-0.041	0.3	0.074	-0.013	0.25	0.67	-0.026
annual_inc	0.35	0.079	0.13	0.36	1	-0.09	0.037	0.21	-0.0067	0.38	0.071	0.022	0.29	0.28	-0.0041
dti	0.064	0.064	0.041	0.042	-0.09	1	-0.029	0.3	0.008	0.24	0.23	-0.069	0.27	0.043	0.019
inq_last_6mths	0.0075	-0.0058	0.058	0.018	0.037	-0.029	1	0.079	0.013	-0.011	-0.14	-0.011	0.1	-0.045	0.0086
open_acc	0.18	0.061	0.039	0.18	0.21	0.3	0.079	1	0.05	0.33	-0.057	-0.018	0.69	0.12	0.061
pub_rec	-0.048	-0.0043	0.079	-0.041	-0.0067	0.008	0.013	0.05	1	-0.064	0.059	-0.0016	0.016	-0.042	0.86
revol_bal	0.31	0.071	0.076	0.3	0.38	0.24	-0.011	0.33	-0.064	1	0.28	-0.082	0.36	0.23	-0.05
revol_util	0.043	0.039	0.38	0.074	0.071	0.23	-0.14	-0.057	0.059	0.28	1	-0.1	-0.023	0.068	0.066
delinq_2yrs	-0.03	-0.013	0.12	-0.013	0.022	-0.069	-0.011	-0.018	-0.0016	-0.082	-0.1	1	0.06	-0.016	0.0025
total_acc	0.28	0.12	-0.0087	0.25	0.29	0.27	0.1	0.69	0.016	0.36	-0.023	0.06	1	0.18	0.032
total_pymnt	0.68	0.27	0.26	0.67	0.28	0.043	-0.045	0.12	-0.042	0.23	0.068	-0.016	0.18	1	-0.04
pub_rec_bankruptcies	-0.028	0.017	0.075	-0.026	-0.0041	0.019	0.0086	0.061	0.86	-0.05	0.066	0.0025	0.032	-0.04	1

IT IS CLEAR THAT CO RELATION DOES NOT EXISTS AMONG SIGNIFICANT VARIABLES

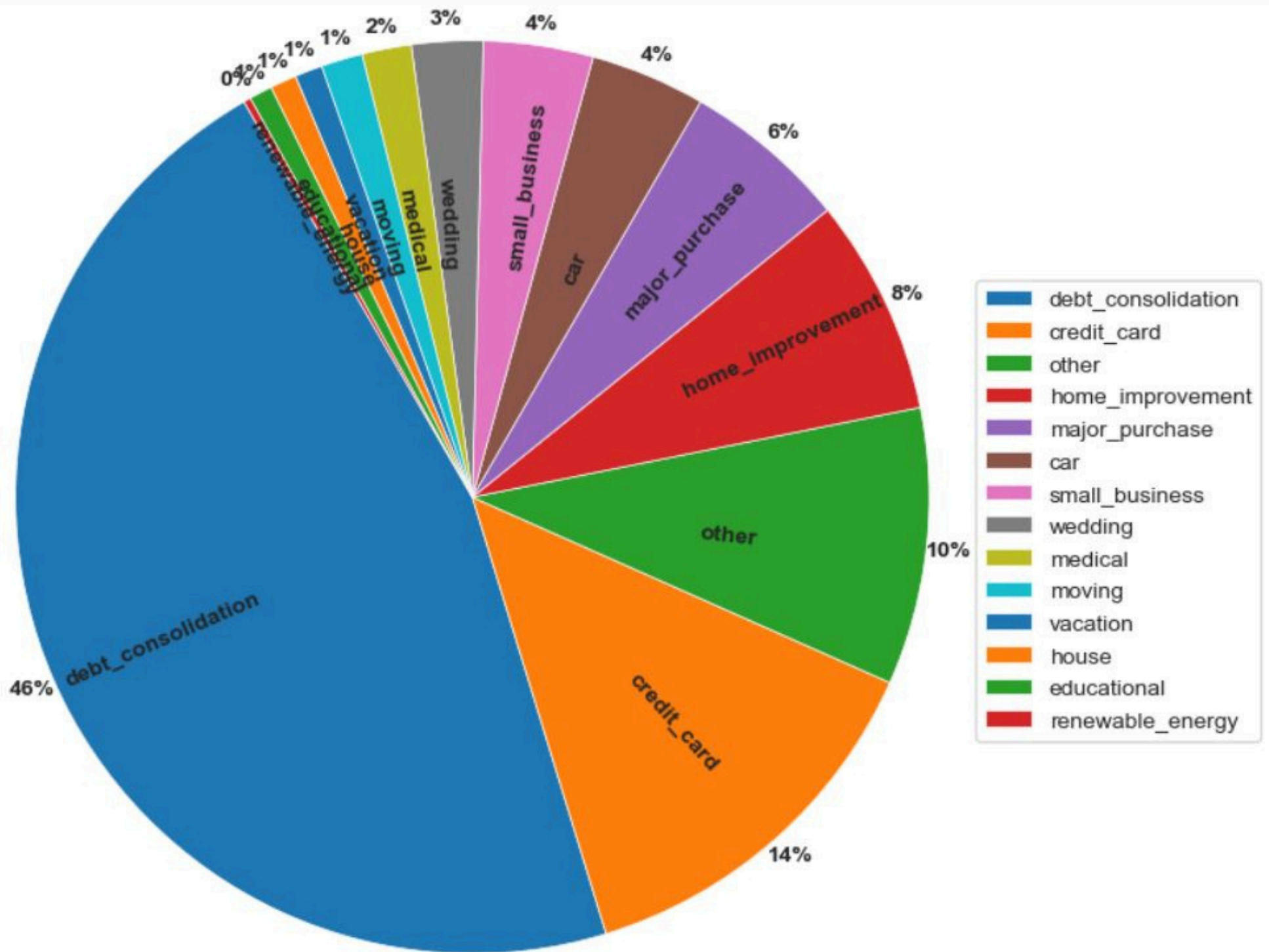# DEBT TO INCOME RATIO FOR BONAFIDES
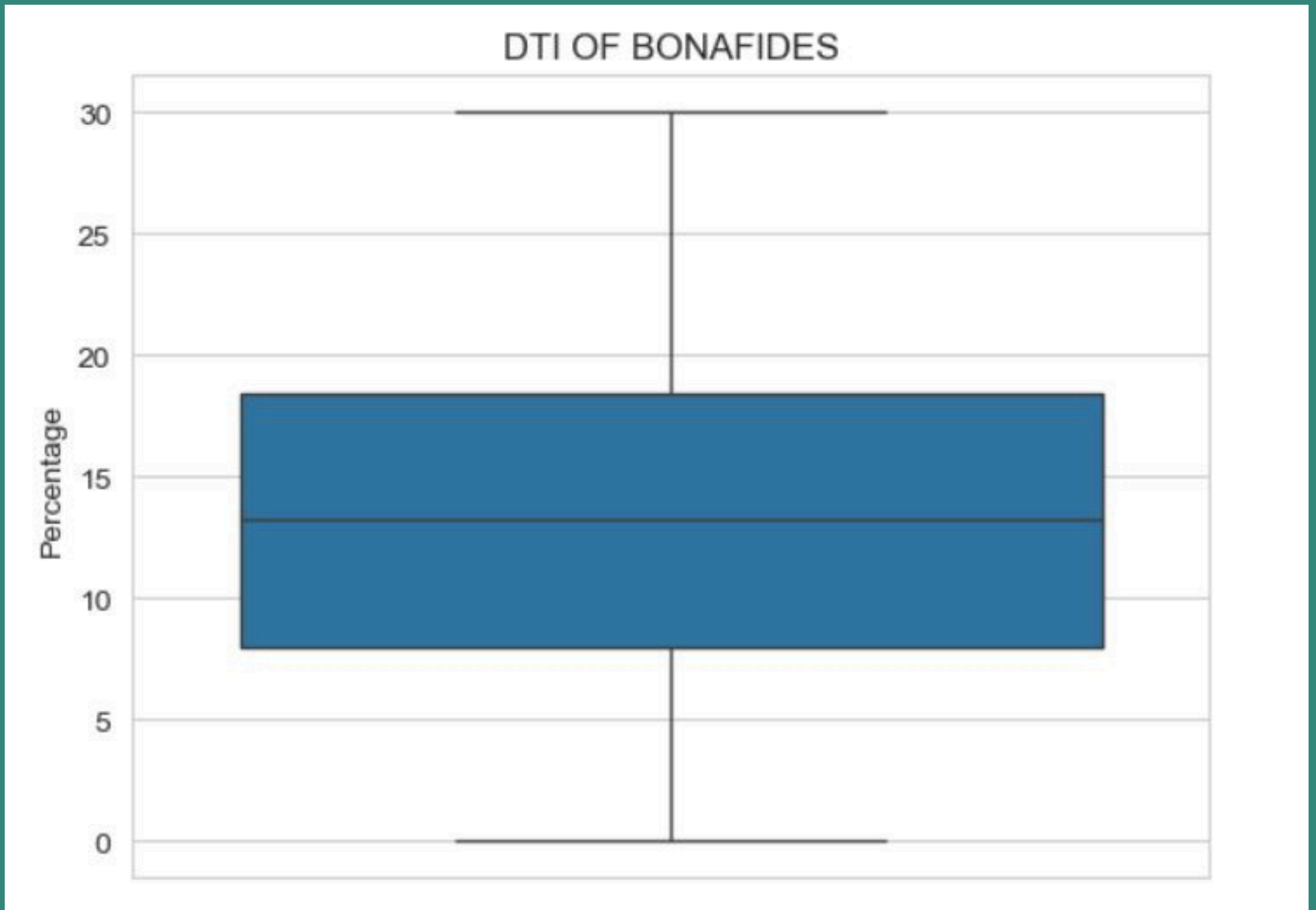
DTI OF BONAFIDES

MEDIAN OF BONAFIDES FALLS NEAR 13 PERCENT WHILE 3RD QUARTILE IS NEAR 18 PERCENTAGE.

DEBT TO INCOME RATIO FOR DEFAULTERS
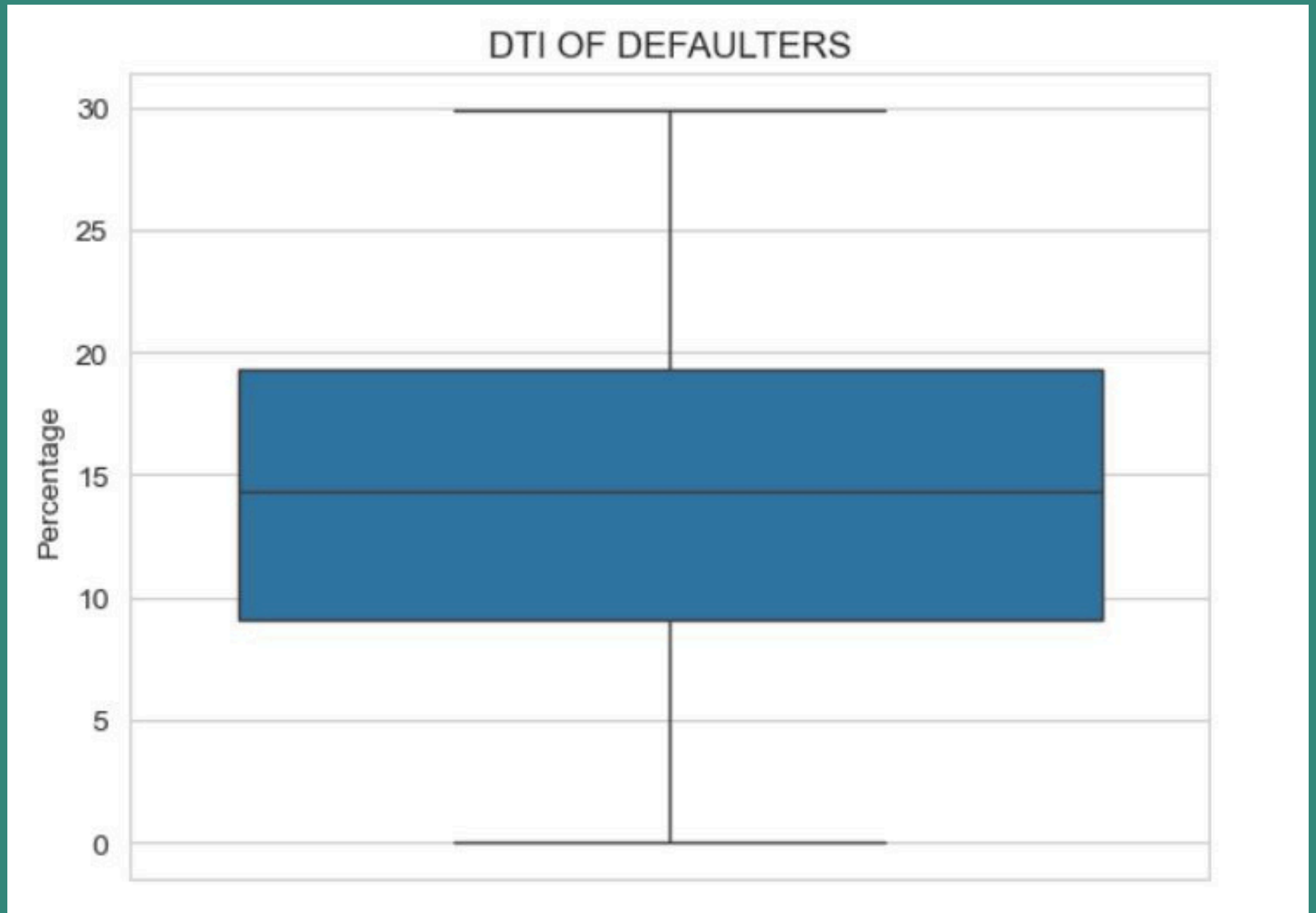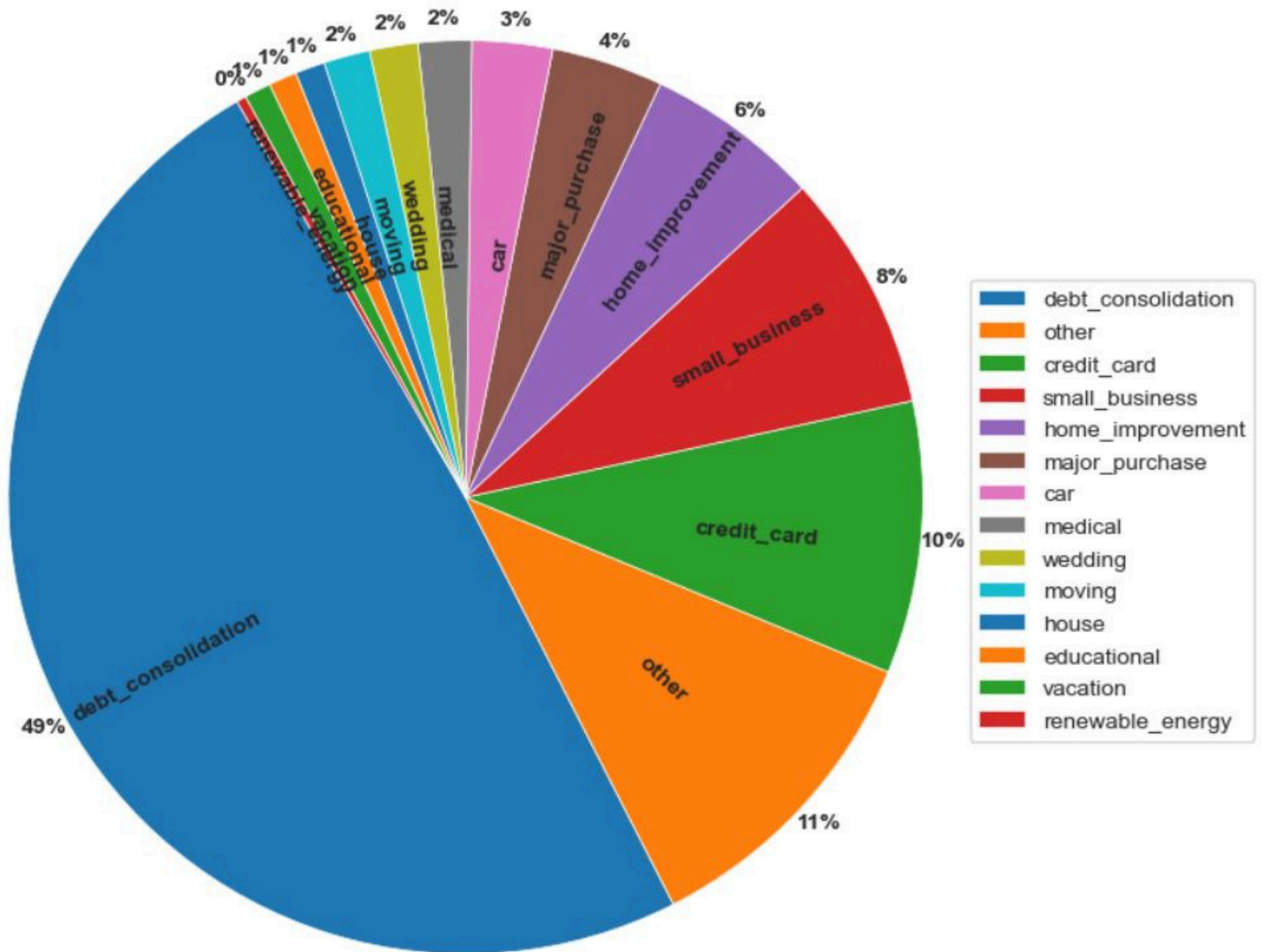

DTI OF DEFAULTERS

MEDIAN OF DEFAULTERS IS SLIGHTLY HIGHER THAN DEFAULTERS THAT IS NEAR 14 PERCENTAGE.

PURPOSE OF LOAN BY BONAFIDES

PURPOSE OF LOAN BY DEFAULTERS

Legend:
- debt_consolidation
- other
- credit_card
- small_business
- home_improvement
- major_purchase
- car
- medical
- wedding
- moving
- house
- educational
- vacation
- renewable_energy

debt_consolidation 49%
other 11%
credit_card 10%
small_business 8%
home_improvement 6%
major_purchase 4%
car 3%
medical 2%
wedding 2%
moving 2%
house 1%
educational 1%
vacation 1%
renewable_energy 0%

SUMMARY

LOAN AMOUNT AND GRADES ARE VARIABLES IDENTIFIED AS GIVING MEANINGFUL INSIGHTS FOR PREDICTION OF DEFAULTERS.

IT IS SEEN THAT MISSING ENTRIES IN FEW VARIABLES ARE ALSO IMPACTING ANALYSIS WHICH IS RECOMMENDED TO BE FILLED AS MANDATORY FOR CONCLUSIVE REPORT.