

# LENDING CLUB CASE STUDY SUBMISSION

GROUP MEMBERS :

ANIRUDDHA TRIKOLIKER

PAVAN RAGHAMWAR

# OBJECTIVE

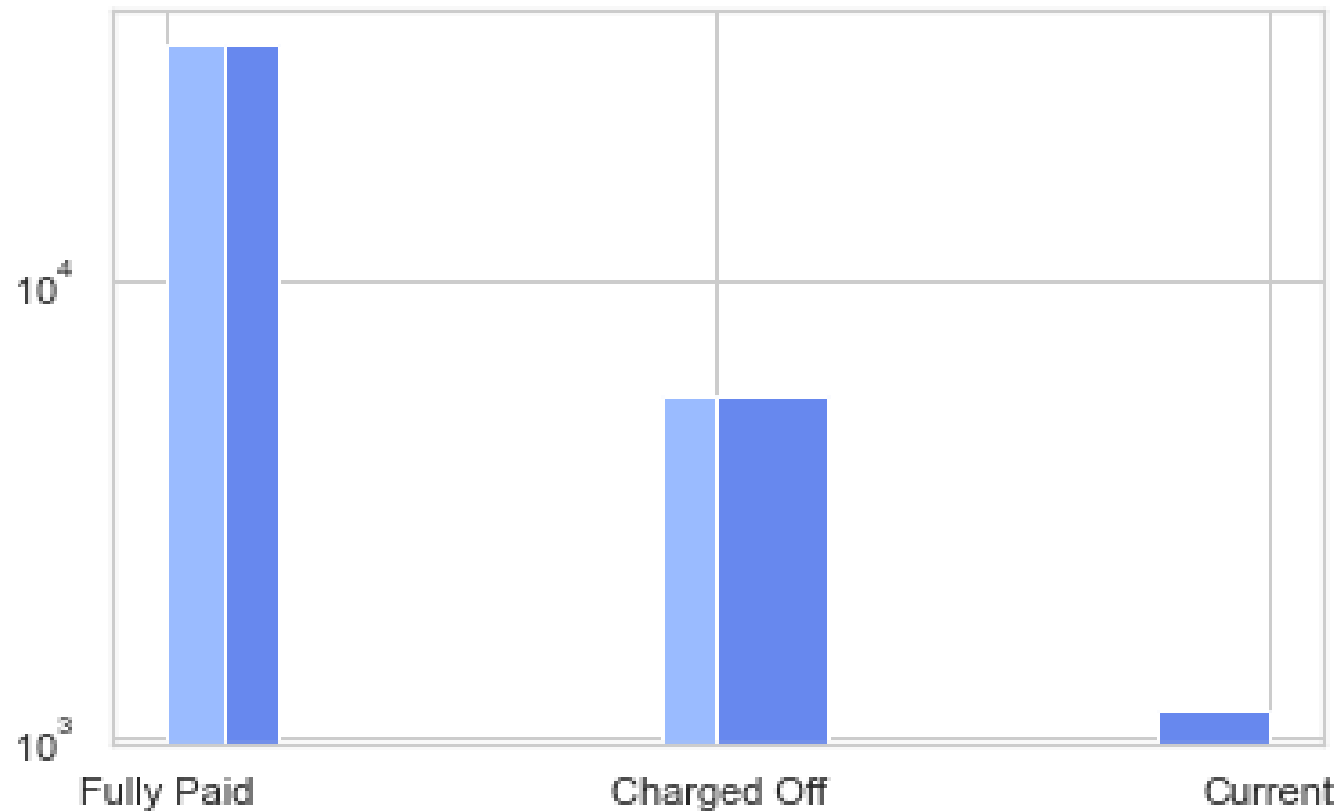
- Use EDA techniques to understand how consumer attributes and loan attributes influence the tendency of default/identification of risk applicant.

# EDA METHODOLOGY FOR PROBLEM SOLVING

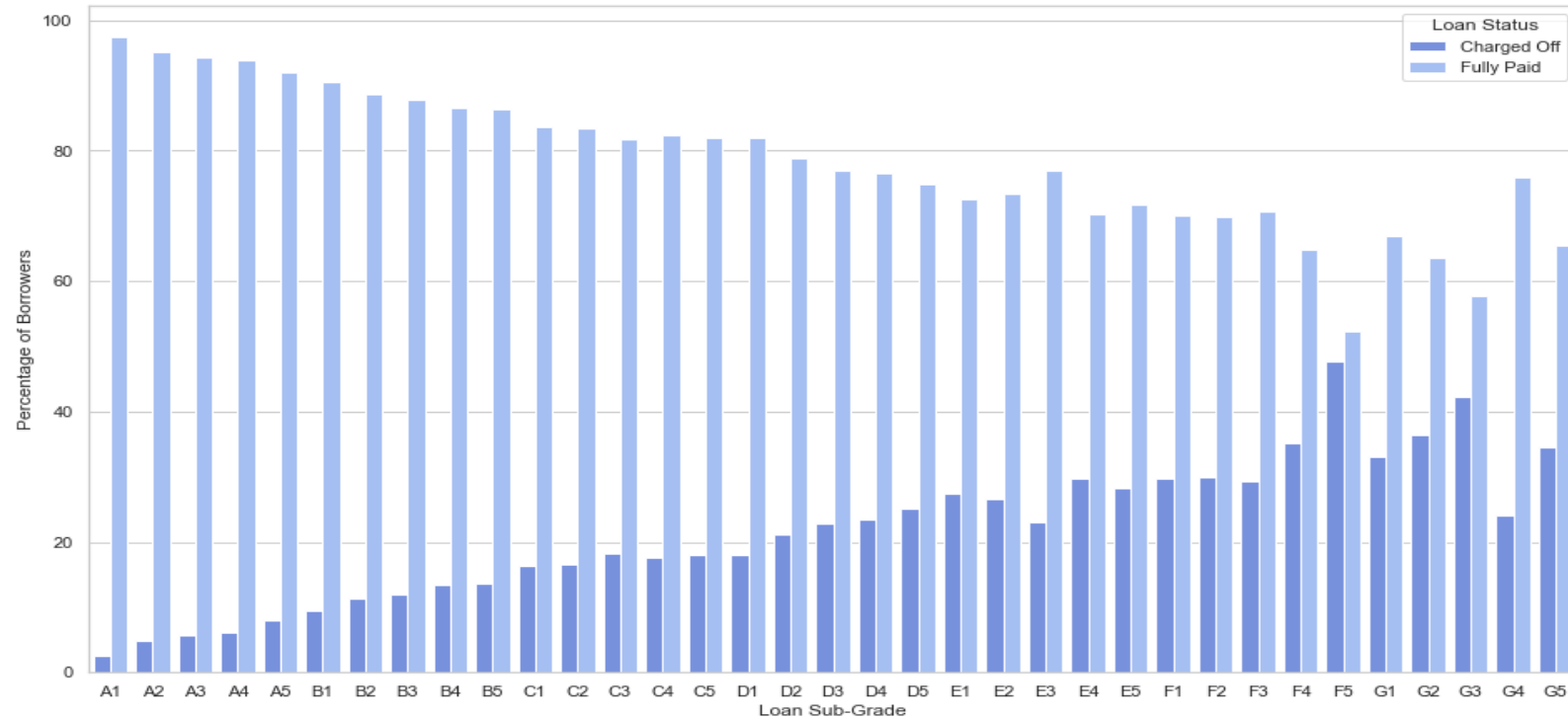
- Understanding Of Dataset - (Data Cleaning)
- Understanding Of Dataset - (Data Cleaning)
  - Data Preparation
- Exploratory Data Analysis - (Performed Univariate, Bivariate, Multivariate analysis)

# UNIVARIATE ANALYSIS

Histogram plots for loan status before removing current accounts and after removing current accounts

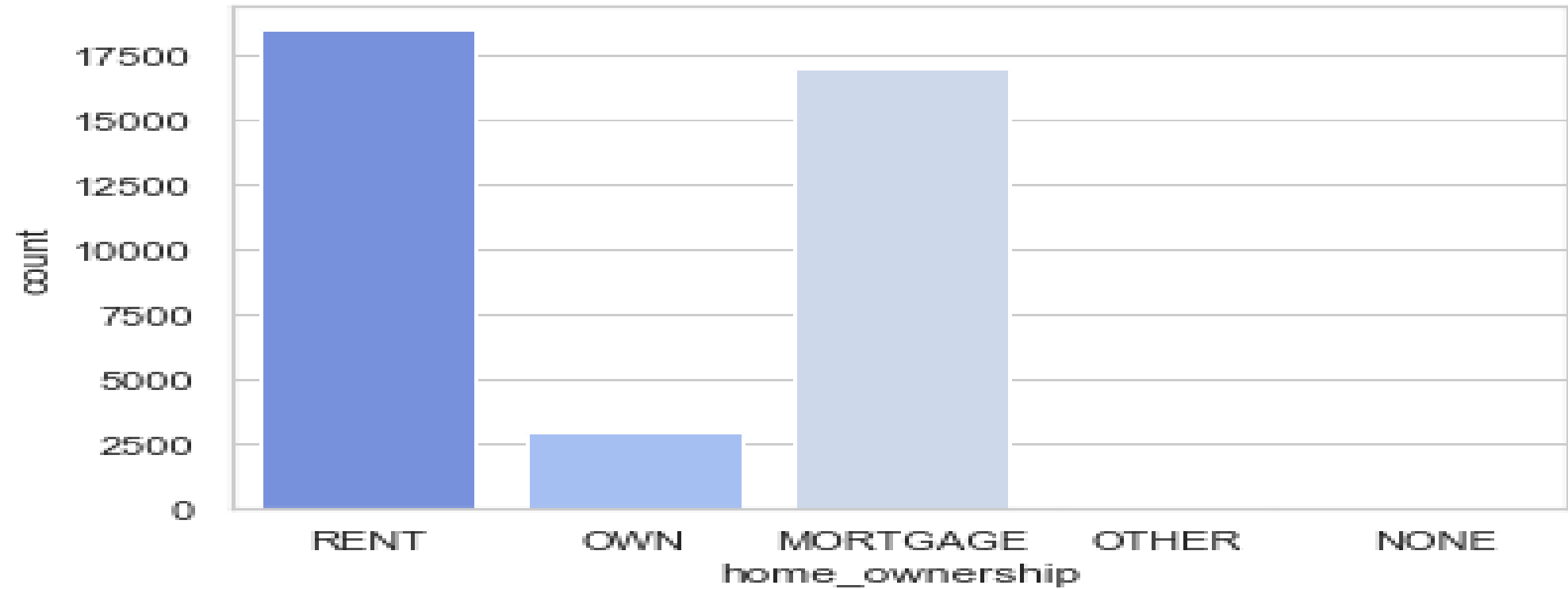


# BAR PLOT WITH SUBGRADE & LOAN STATUS



- Grade D2 and above tend to default as per above table
- Appliers with Grade A or B tend to pay off the loan

# COUNT IN EACH CATEGORY BY PLOTTING COUNT PLOT



# PLOT FOR HOME OWNERSHIP WITH LOAN STATUS



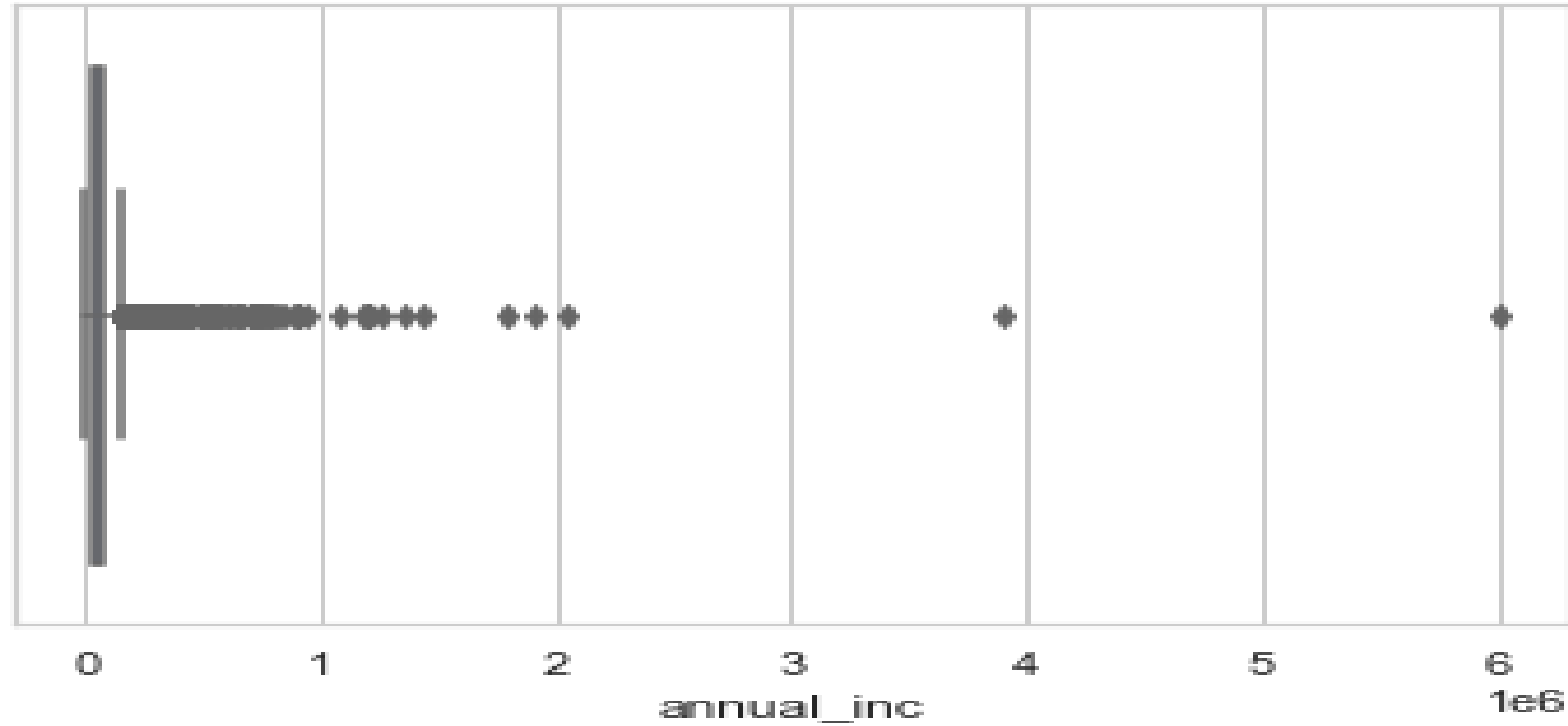
- Rented or Other categories in home ownership tend to default the loan
- Applicants living in own houses or mortgage tend to pay off the loan fully

		purpose
car	loan_status	
	Charged Off	10.67%
credit_card	Fully Paid	89.33%
	Charged Off	10.78%
debt_consolidation	Fully Paid	89.22%
	Charged Off	15.33%
educational	Fully Paid	84.67%
	Charged Off	17.23%
home_improvement	Fully Paid	82.77%
	Charged Off	12.07%
house	Fully Paid	87.93%
	Charged Off	16.08%
major_purchase	Fully Paid	83.92%
	Charged Off	10.33%
medical	Fully Paid	89.67%
	Charged Off	15.57%
moving	Fully Paid	84.43%
	Charged Off	15.97%
other	Fully Paid	84.03%
	Charged Off	16.38%
renewable_energy	Fully Paid	83.62%
	Charged Off	18.63%
small_business	Fully Paid	81.37%
	Charged Off	27.08%
vacation	Fully Paid	72.92%
	Charged Off	14.13%
wedding	Fully Paid	85.87%
	Charged Off	10.37%
	Fully Paid	89.63%

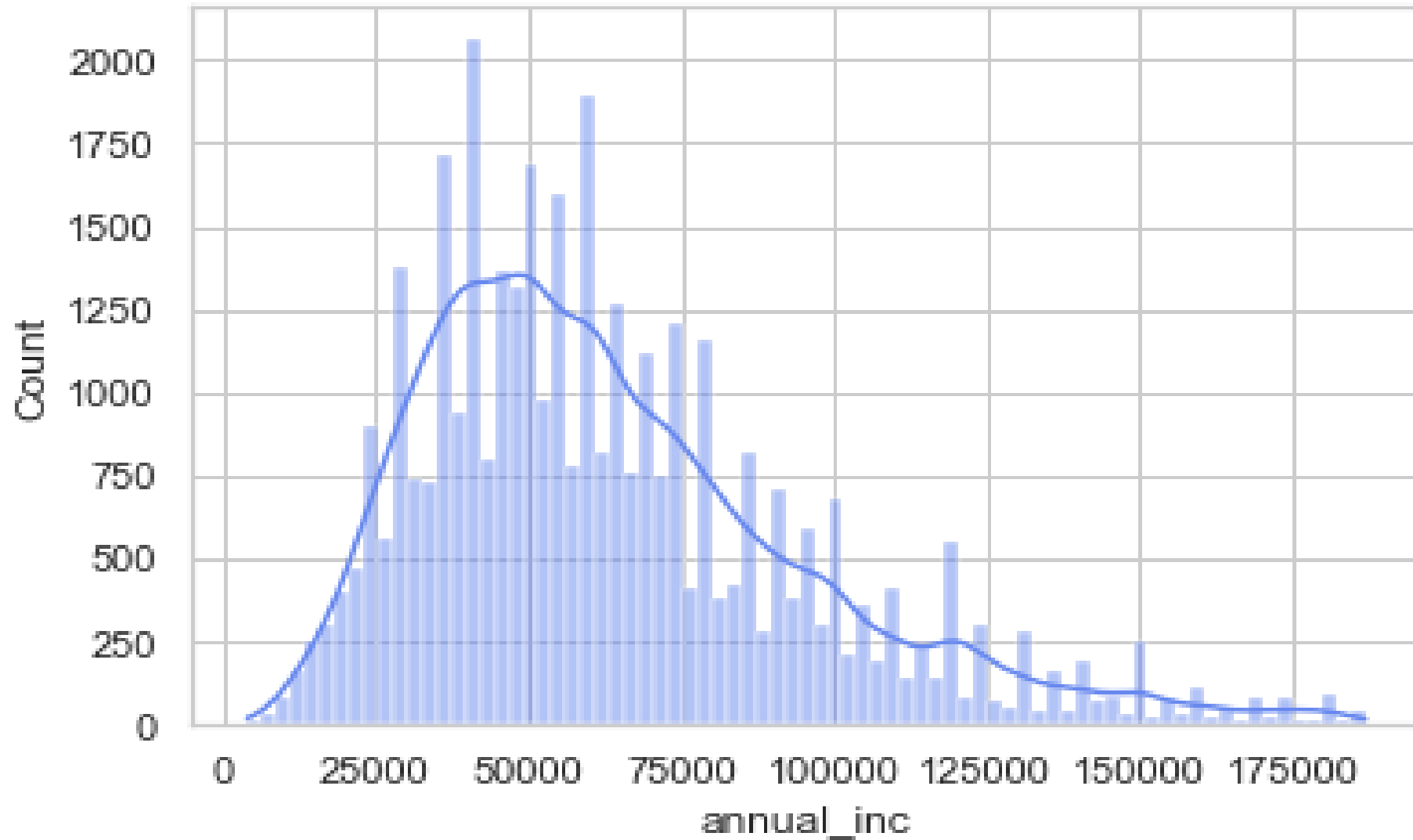
- - Borrowers who bought the loan for small business purpose tend to default the highest
- - Borrowers who bought the loan for renewable energy purpose tend to default more
- - Borrowers who bought the loan for major purchase and wedding purposes tend to fully pay the loan more



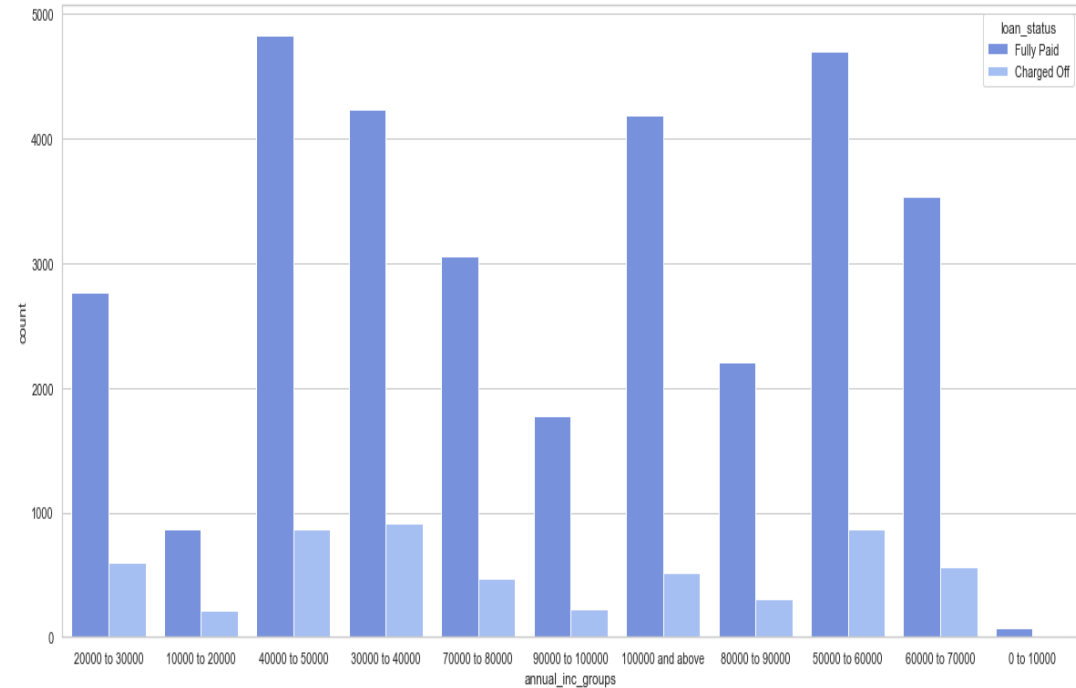
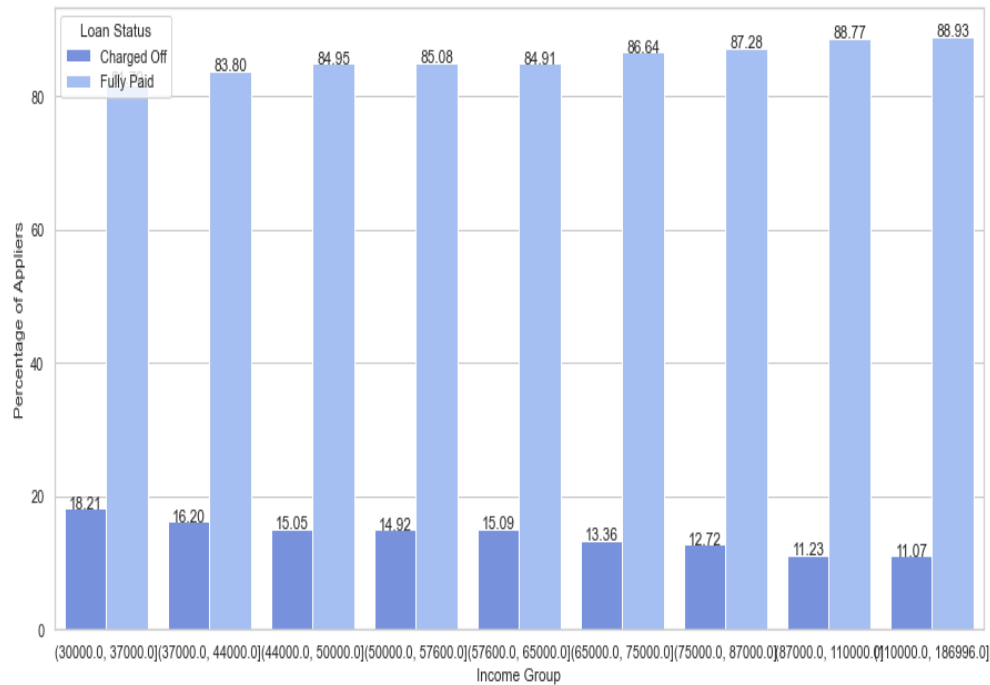
# OUTLIERS IN ANNUAL INCOME USING BOXPLOT



# ANALYSE ANNUAL INCOME USING HISTPLOT

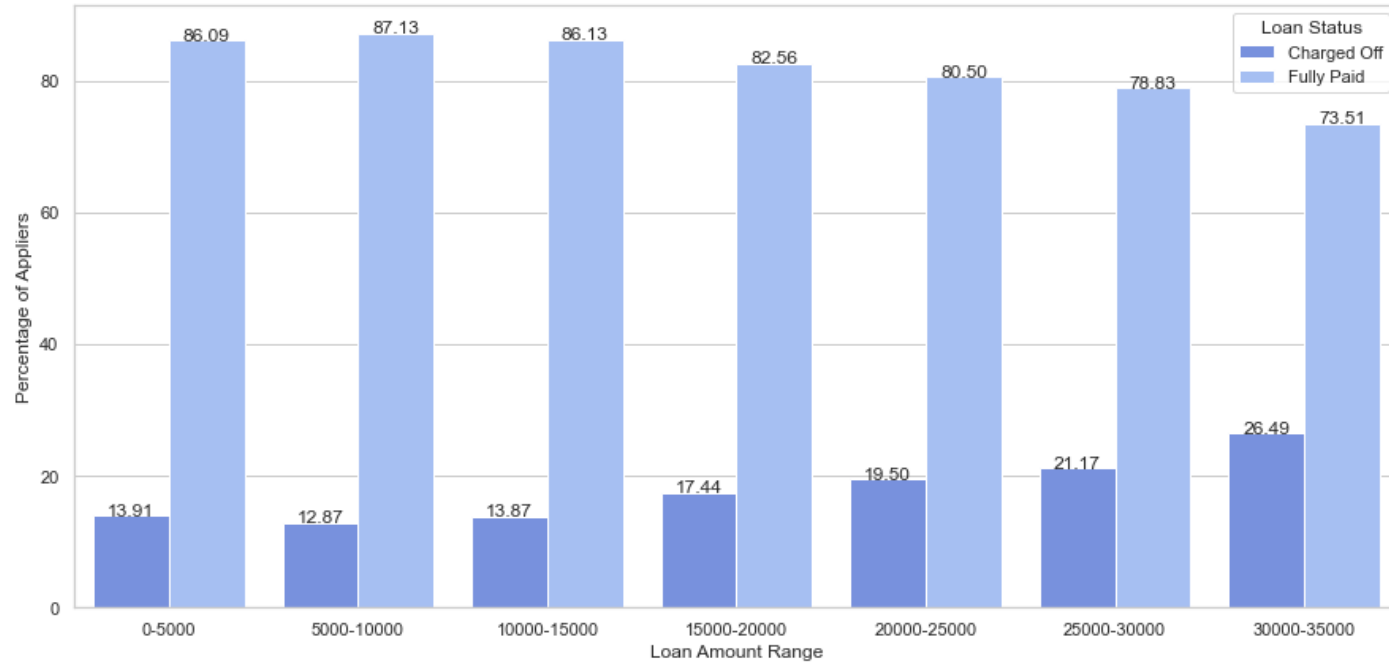


## INCOME LEVEL LABELS TO THE DATAFRAME AND GET THE COUNT OF IT IN EACH LOAN STATUS TYPE



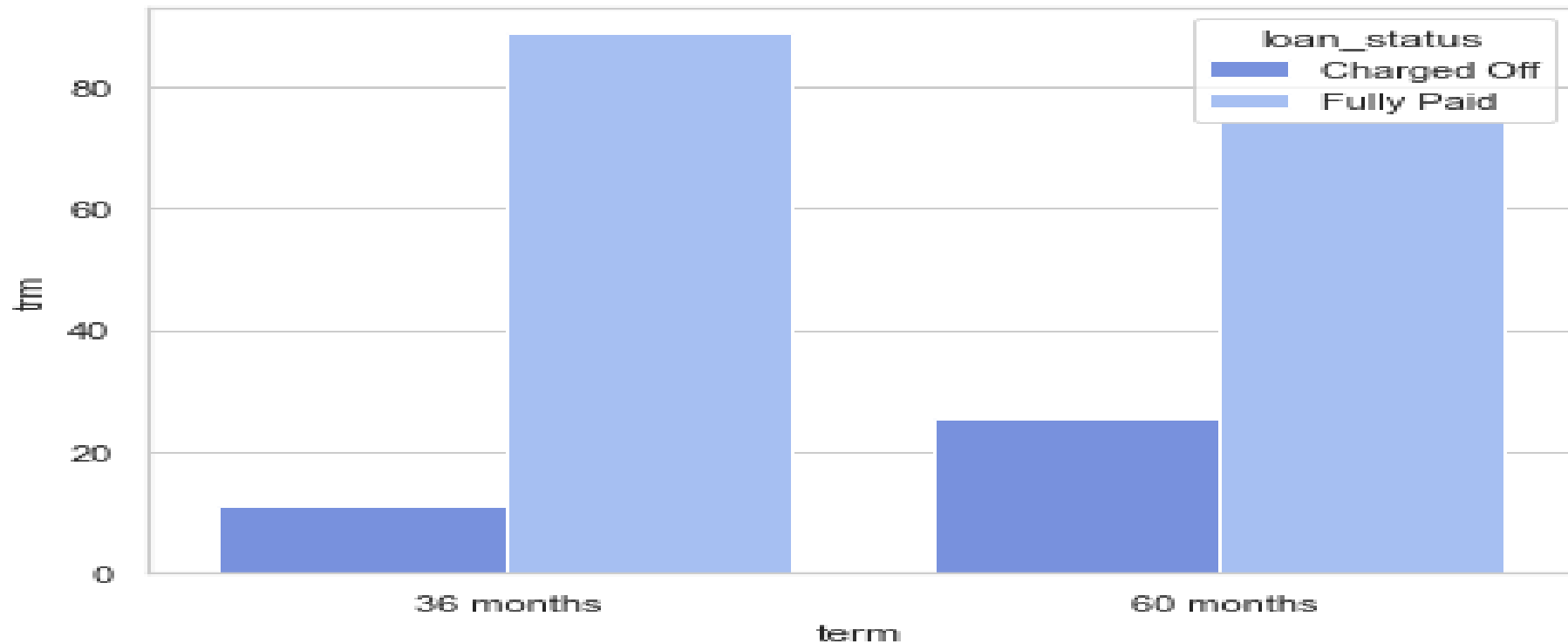
- Based on above information we can clearly see that appliers with annual income < 55,000 tend to default and with > 65,000 tend to repay the loan

# LOAN AMOUNT IN EACH LOAN STATUS TYPE



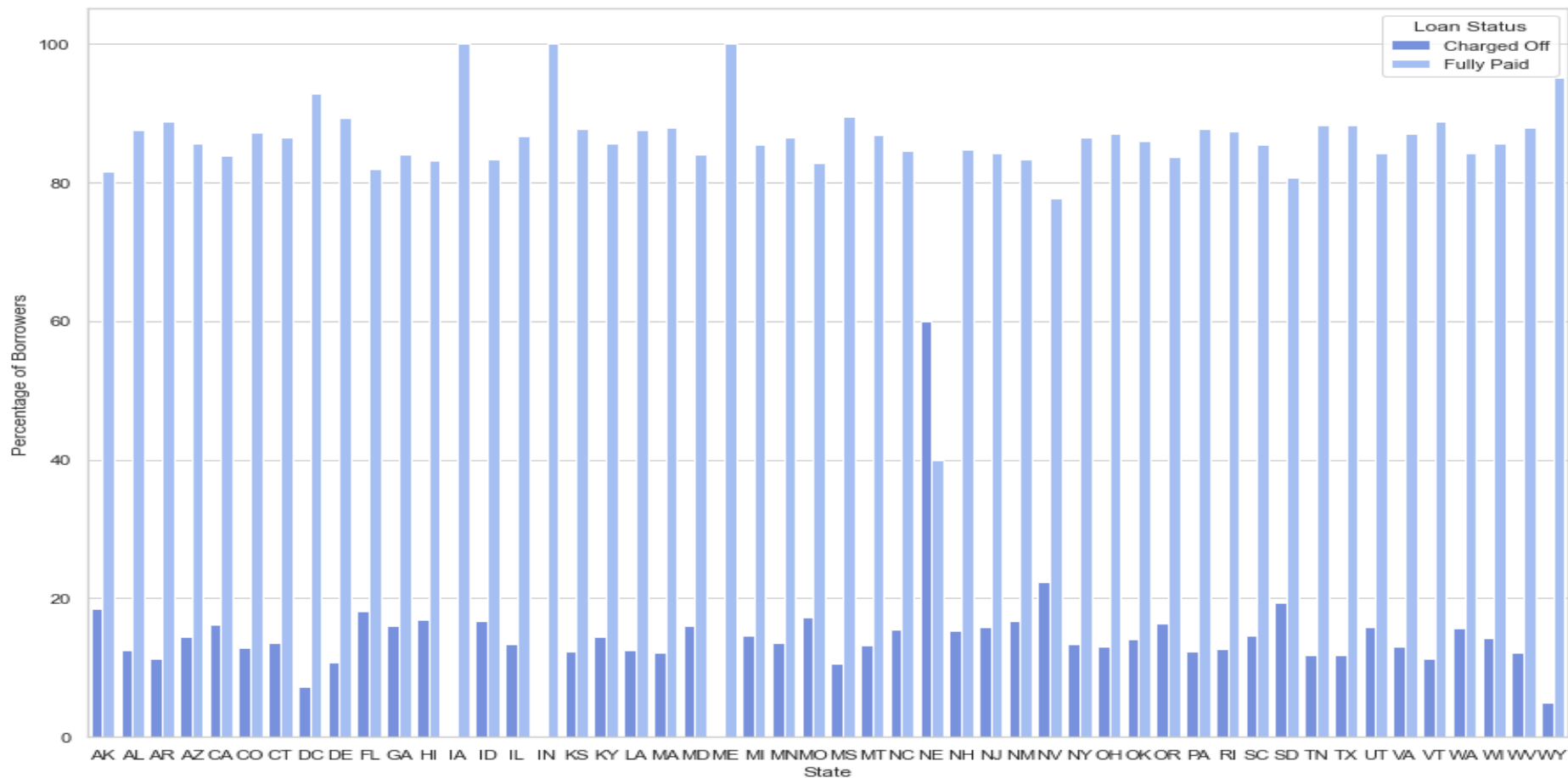
- Applicants with loan amount greater than 15,000 tend to default and with less than 15,000 tend to repay

# BAR PLOT TO UNDERSTAND THE RELATIONSHIP WITH TERM AND LOAN STATUS



- Appliers who have taken the loan for 5 years are more likely to default the loan
- Borrowers who have taken the loan for 3 years are more likely to fully pay the loan

# ANALYSIS BASED ON ADDRESS STATE



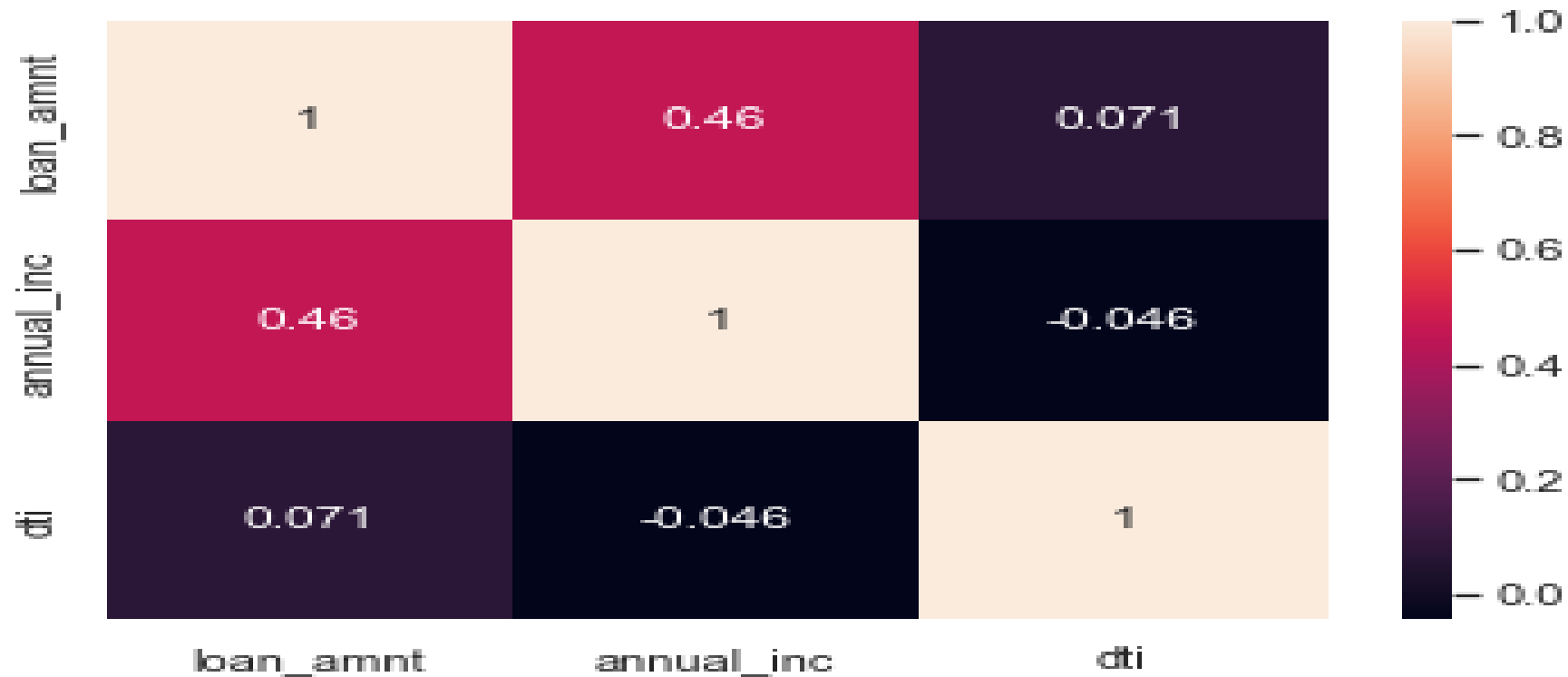
- Appliers from the states Nevada and South Dakota are likely to default the loan
- Appliers from the states Washington D.C. and Wyoming are likely to fully repay the loan

# ANALYSIS BASED ON VERIFICATION STATUS



- As per the above graph we can see that verified Applicants tend to default more than Not Verified Applicants

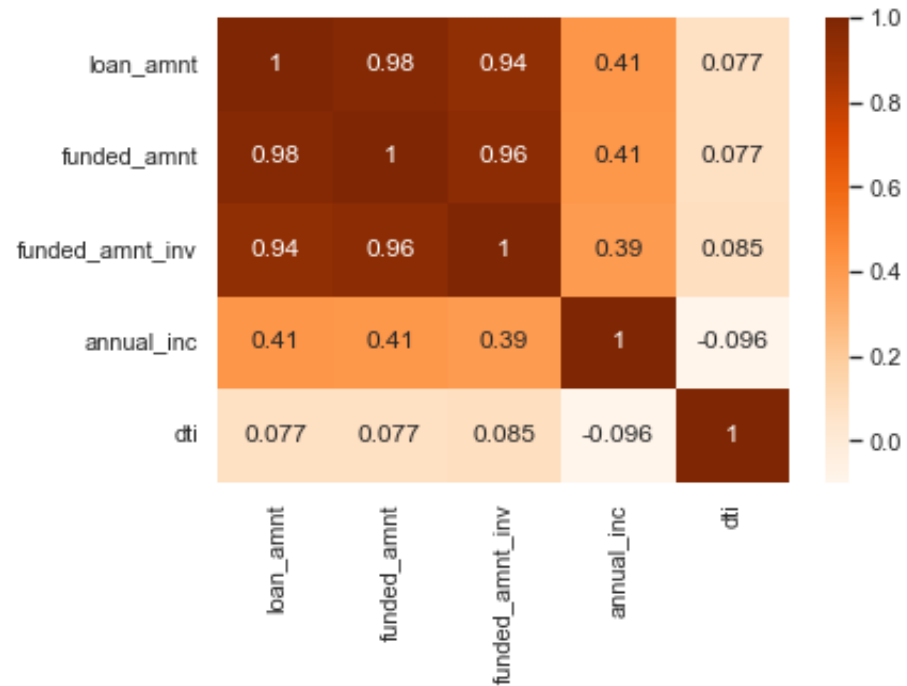
# CORRELATION BETWEEN LOAN AMOUNT, ANNUAL INCOME AND DTI USING HEAT MAP



- As per the there is less correlation (0.46) between loan amount and annual income in case of default loans, so, no relationship exists between loan default with annual income and loan amount



# Correlation of different variables



- Annual Income is negatively correlated with Debt to Income
- Loan Amount, Investor Amount and funding amount are strongly correlated