

The background of the slide is a light gray gradient. It is decorated with numerous realistic water droplets of various sizes. Some droplets are at the top left, some are scattered in the middle, and a larger cluster of droplets is on the right side. The droplets have highlights and shadows, giving them a three-dimensional appearance.

LENDING CLUB CASE STUDY

BY

SABAHAT JAMIL & SOURISH ROY

OBJECTIVE

THIS CASE STUDY AIMS TO IDENTIFY PATTERNS WHICH INDICATE THE DRIVING FACTORS (OR DRIVER VARIABLES) BEHIND LOAN DEFAULT, I.E. THE VARIABLES WHICH ARE STRONG INDICATORS OF DEFAULT. THIS KNOWLEDGE CAN BE UTILIZED FOR ITS PORTFOLIO AND RISK ASSESSMENT.

DATA UNDERSTANDING

DATA SET HAS 2 FILES AS STATED BELOW:

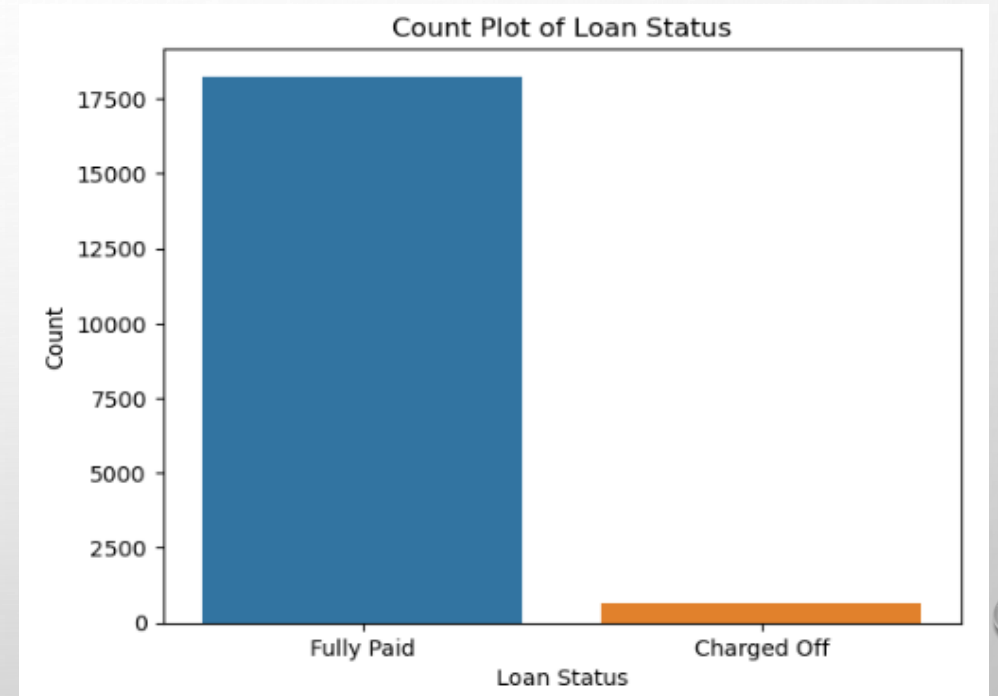
- LOAN.CSV : CONTAINS ALL THE INFORMATION OF ALL LOANS
- DATA_DICTIONARY.XLSX : DESCRIBES THE MEANING OF THESE VARIABLES PRESENT IN LOAN.CSV FILE

UNIVARIATE ANALYSIS

LOAN STATUS – FULLY PAID VS CHARGED OFF

- CLASS IMBALANCE CAN BE SEEN
- AROUND **3.5%** OF THE LOAN OFFERED BECOMES **BAD LOAN**

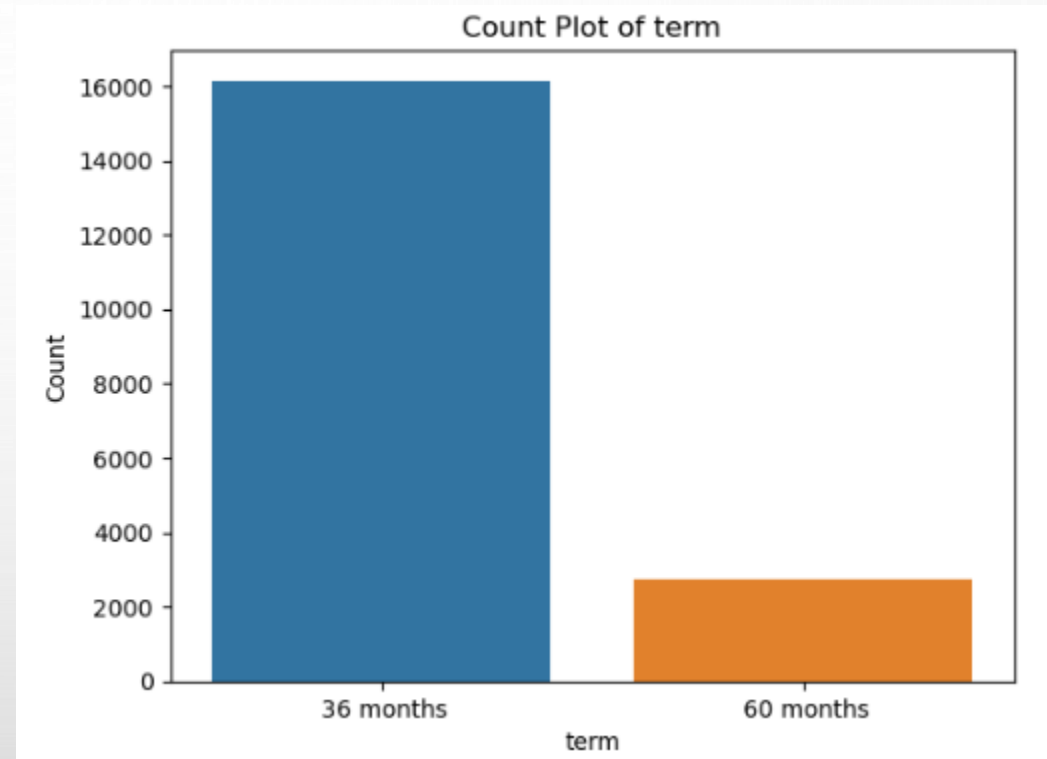
* ON X-AXIS IS LOAN STATUS AND ON Y-AXIS IS COUNT OF INDIVIDUAL



LOAN TENURE

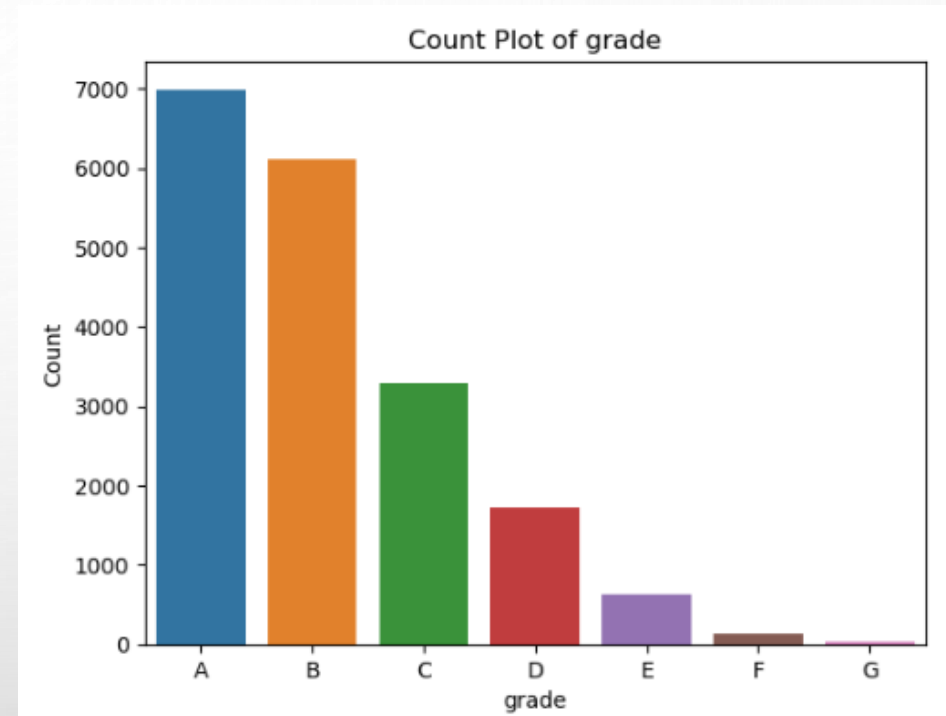
- GRAPH SHOWS MOSTLY CLIENT **PREFER LOAN TENURE 36 MONTHS OR LESS**

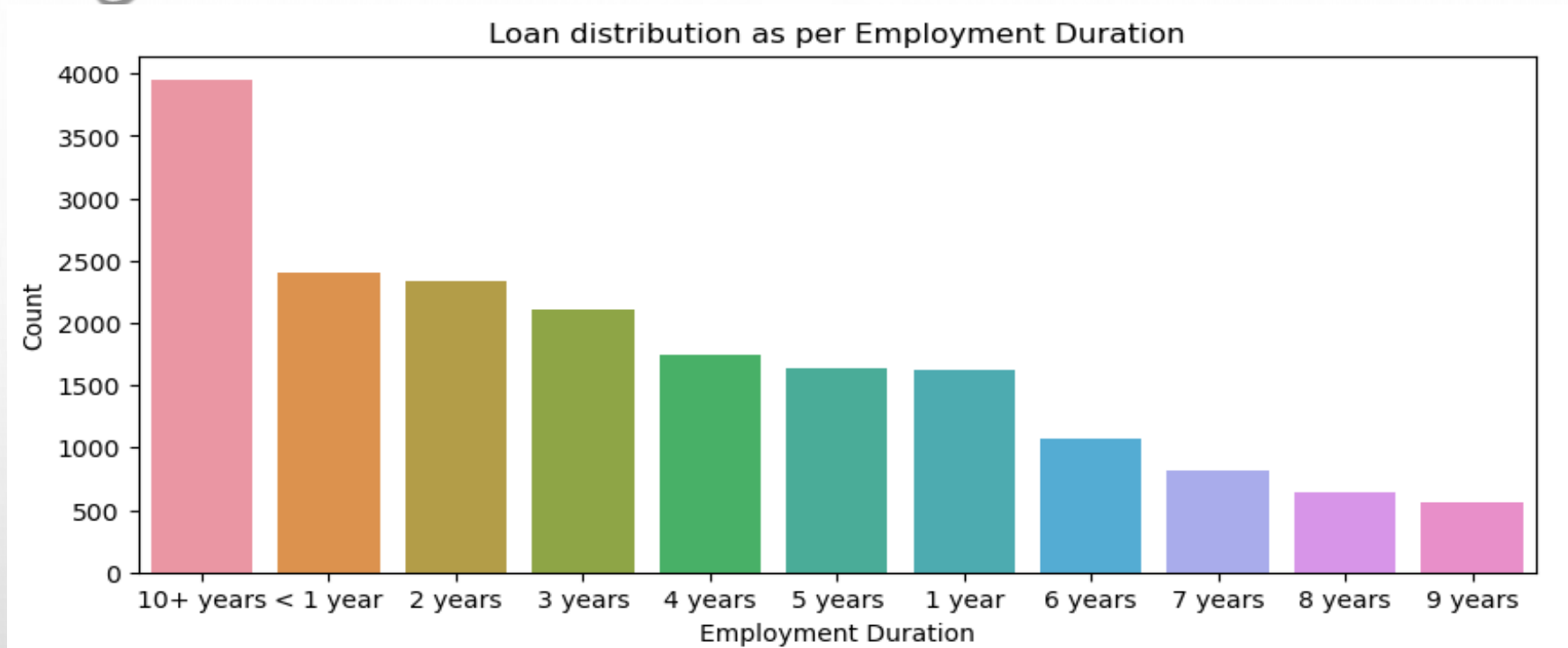
* ON X-AXIS IS LOAN TENURE AND ON Y-AXIS IS COUNT OF INDIVIDUAL



GRADATION OF LOAN BASED ON RISK

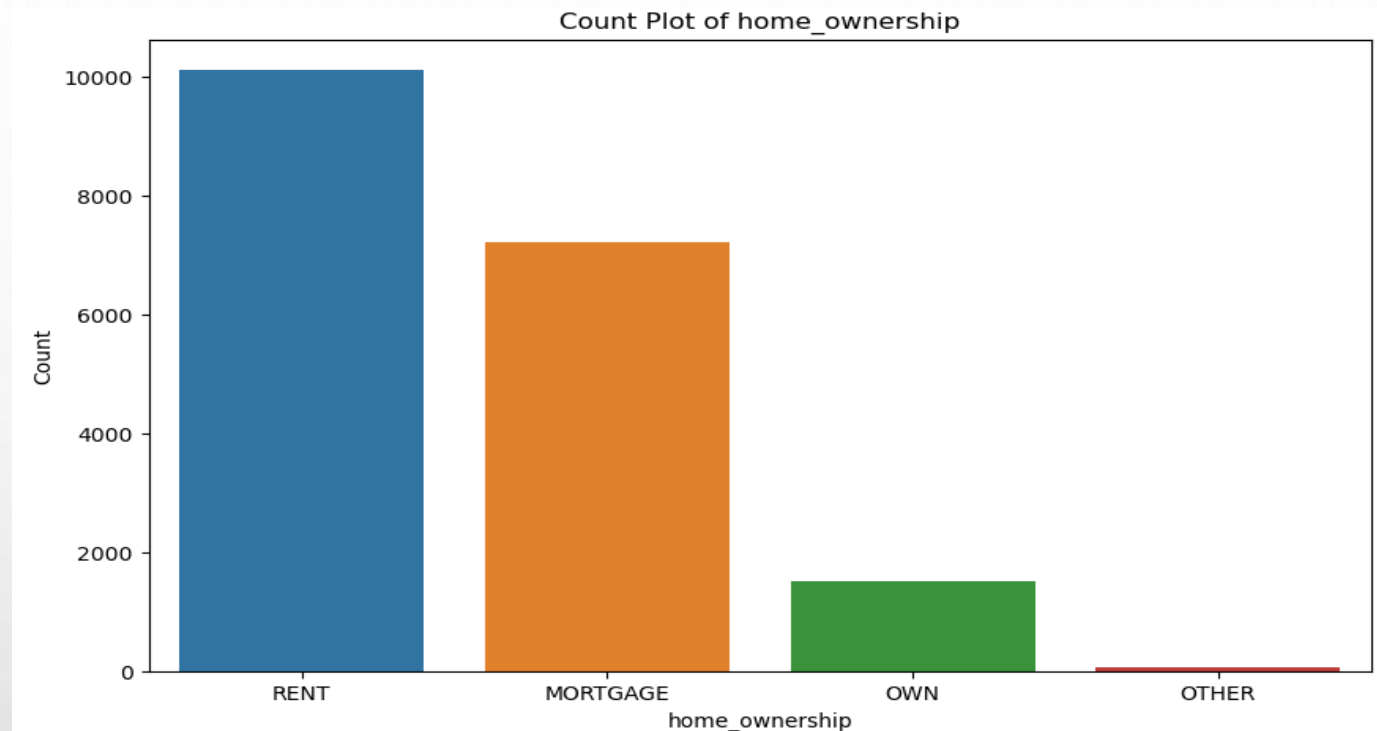
- THE LETTER GRADES THAT ARE ASSIGNED TO LOANS INDICATE THE EXPECTED RISK OF THE LOAN. **LOANS GRADED A HAVE THE LOWEST EXPECTED RISK OF LOSS** AND THEREFORE PAY THE LENDER THE LOWEST INTEREST RATE. ON THE OTHER HAND, **G LOANS HAVE THE HIGHEST EXPECTED RISK OF LOSS**
- AROUND **15%** LOANS FALL INTO **MEDIUM TO HIGH-RISK ZONE**





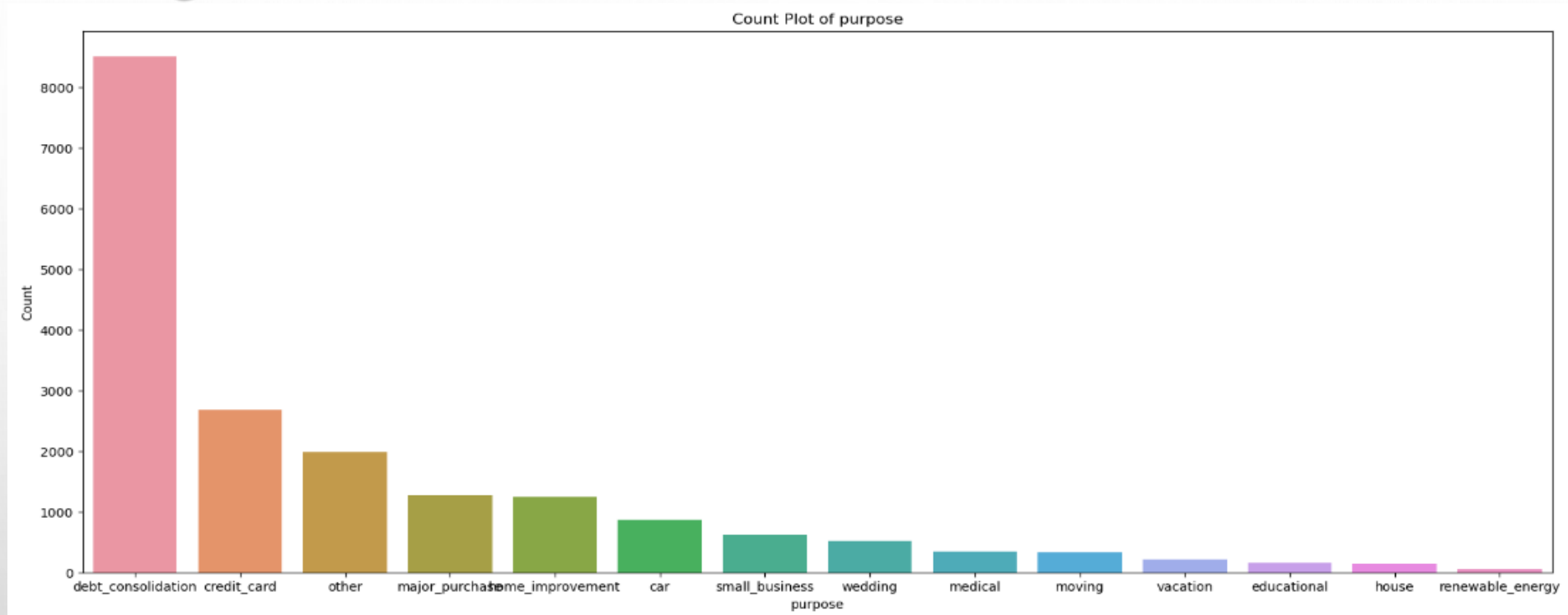
DISTRIBUTION OF LOAN AS PER EMPLOYMENT DURATION

- AS PER THE GRAPH, MOSTLY LOANS ARE PREFERRED BY CLIENTS WITH OVER **10 YEARS INTO EMPLOYMENT**
- ALSO, NOTICED THAT CLIENTS ARE LESS LIKELY TO TAKE UP LOAN WITH AROUND **6-9 YEARS INTO EMPLOYMENT**



LOAN APPLIED BY CLIENT BASED ON HOME OWNERSHIP STATUS

- CLIENT WITH HOME OWNERSHIP AS '**RENT**' OR '**MORTGAGE**' TEND TO TAKE LOAN MORE COMPARED TO '**OWN**' HOME OWNERSHIP
- AROUND 15% LOANS FALL INTO **MEDIUM TO HIGH-RISK ZONE**



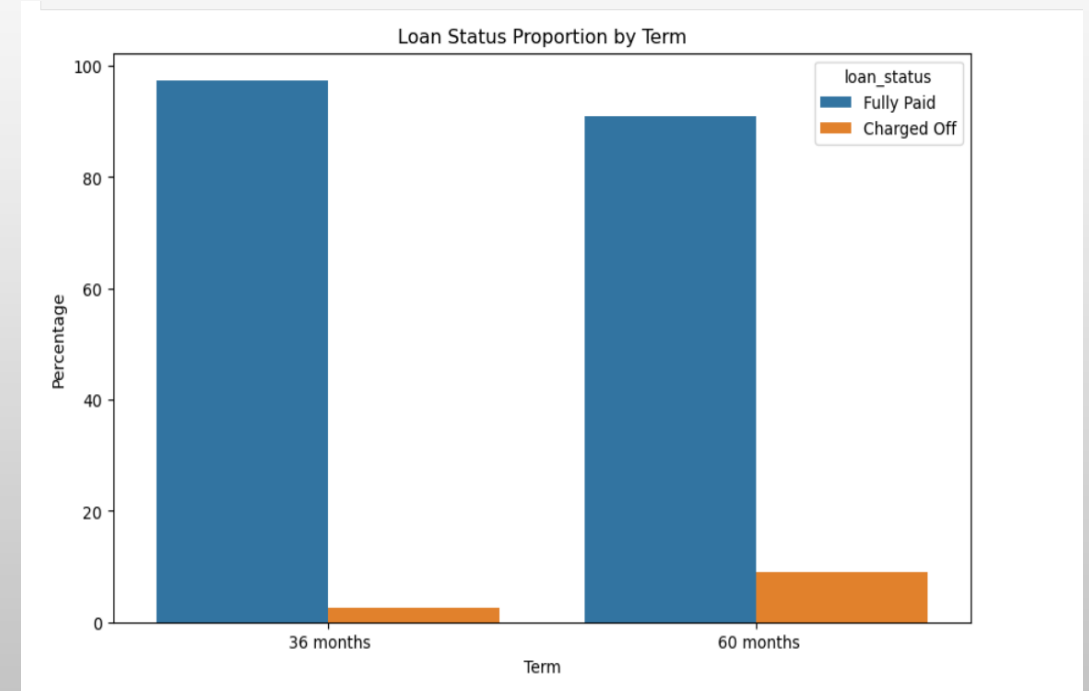
DISTRIBUTION OF LOAN APPLICANT BASED ON PURPOSE

- CLIENT MOSTLY TAKE UP LOAN FOR **DEBT CONSOLIDATION** FOLLOWED BY **CREDIT CARD**
- **RENEWABLE ENERGY** CONTRIBUTES TO LEAST FOR BORROWER

BIVARIATE ANALYSIS

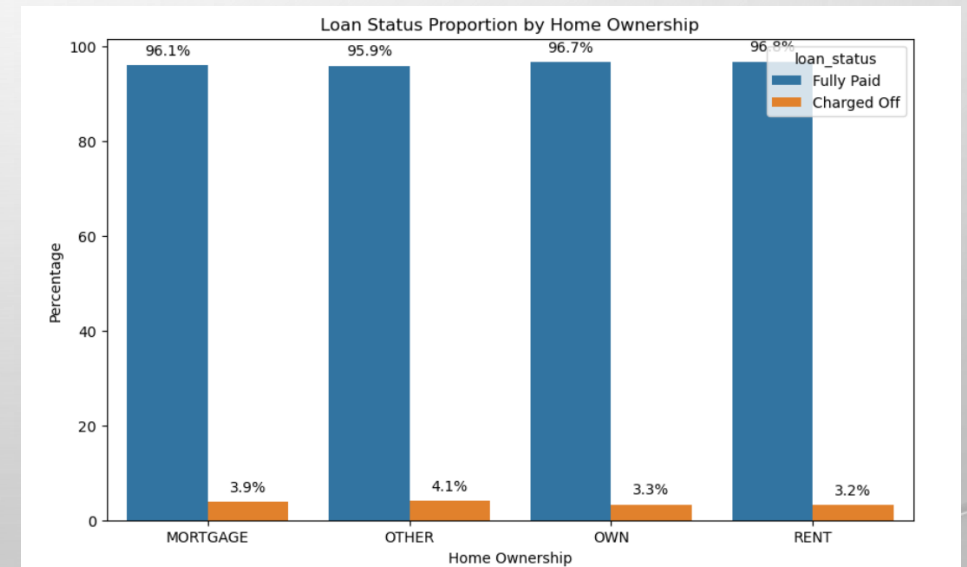
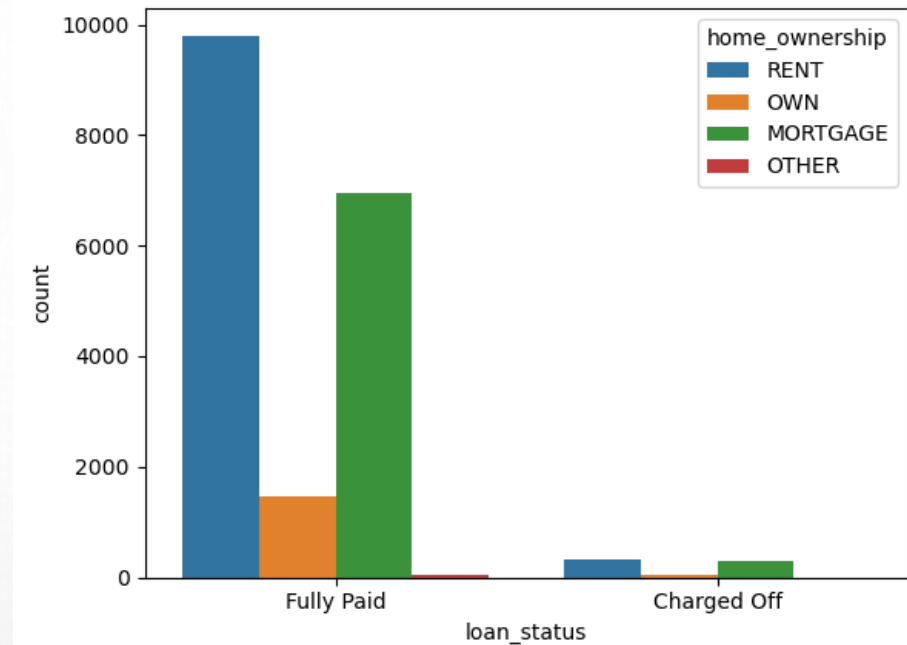
TERM VS PERCENTAGE OF PEOPLE WHO FULLY PAY OR CHARGED OFF

- MOSTLY CLIENTS WHO TENDS TO BE DEFAULTERS TAKE 60 MONTHS TERM



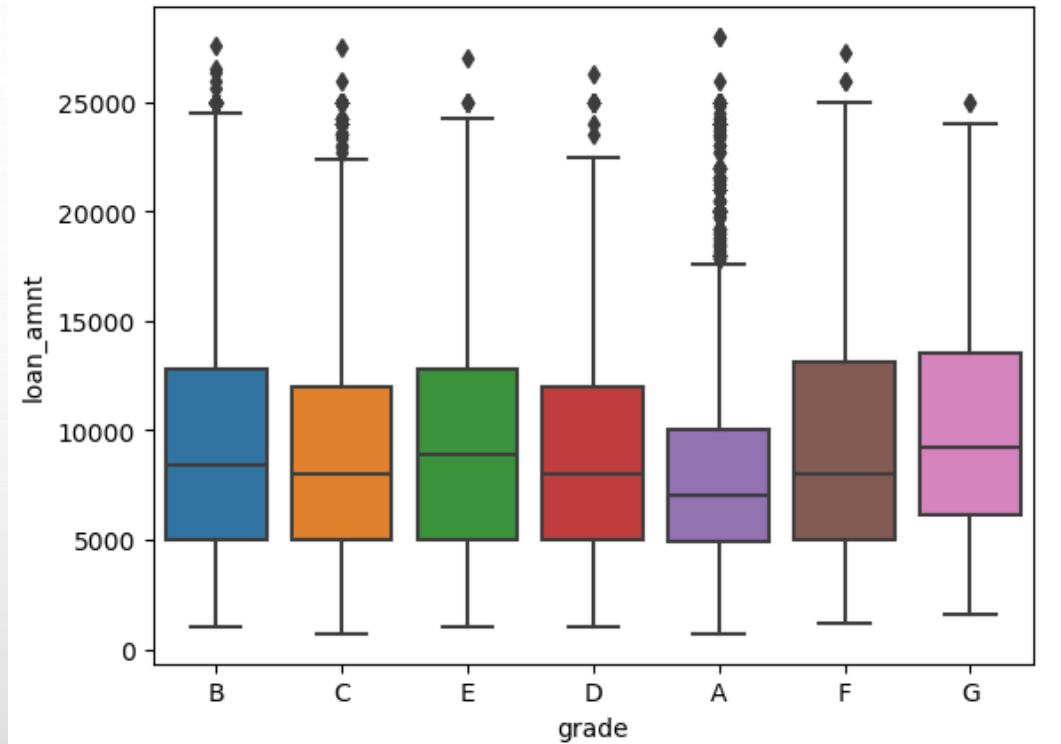
LOAN STATUS VS HOME OWNERSHIP

- MOSTLY CLIENTS WHO LIVE IN **RENTED** OR **MORTGAGE** HOUSES THE PAY THE LOAN FULLY
- CLIENT IN **OTHER** CATEGORY TEND TO DEFAULT



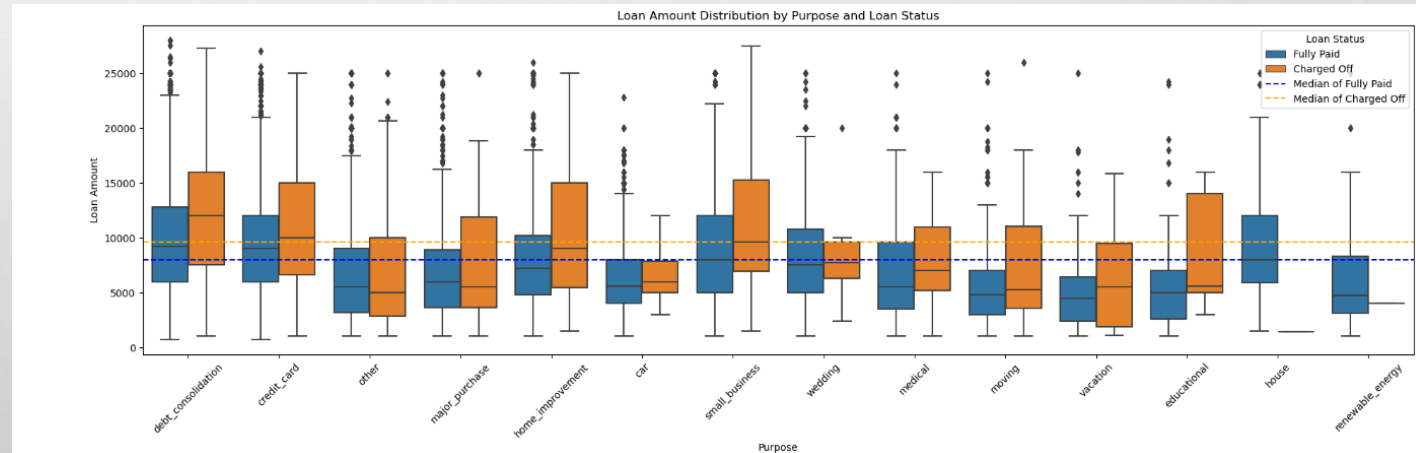
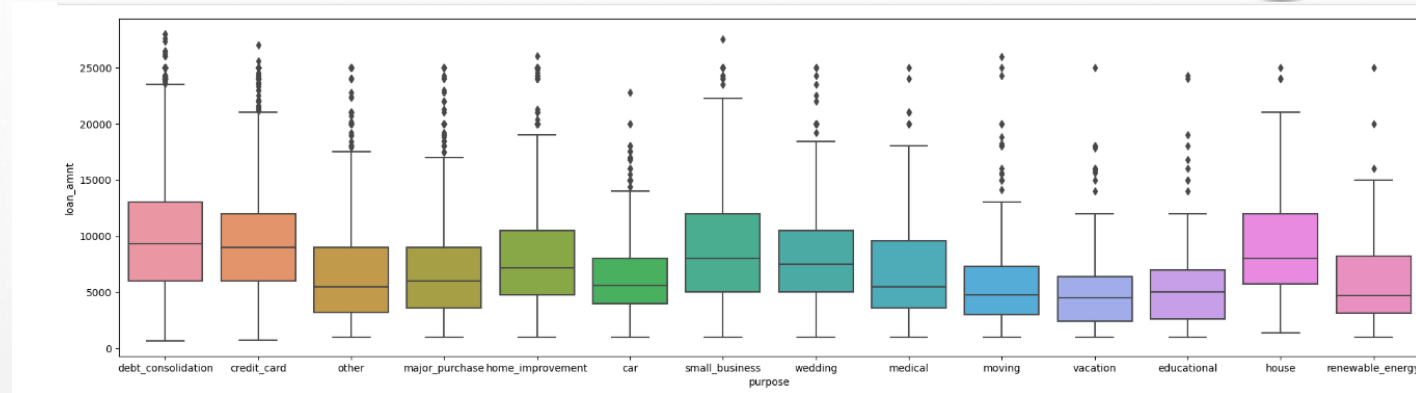
LOAN RISK GRADE VS LOAN AMOUNT

- GRADE **G, E** TAKES MORE LOAN
- GRADE **A** TAKES LEAST AMOUNT
- HIGHEST LOAN AMOUNT IS TAKEN BY GRADE **B** AND **G**
- GRADE **A** HAS MOST OUTLIERS



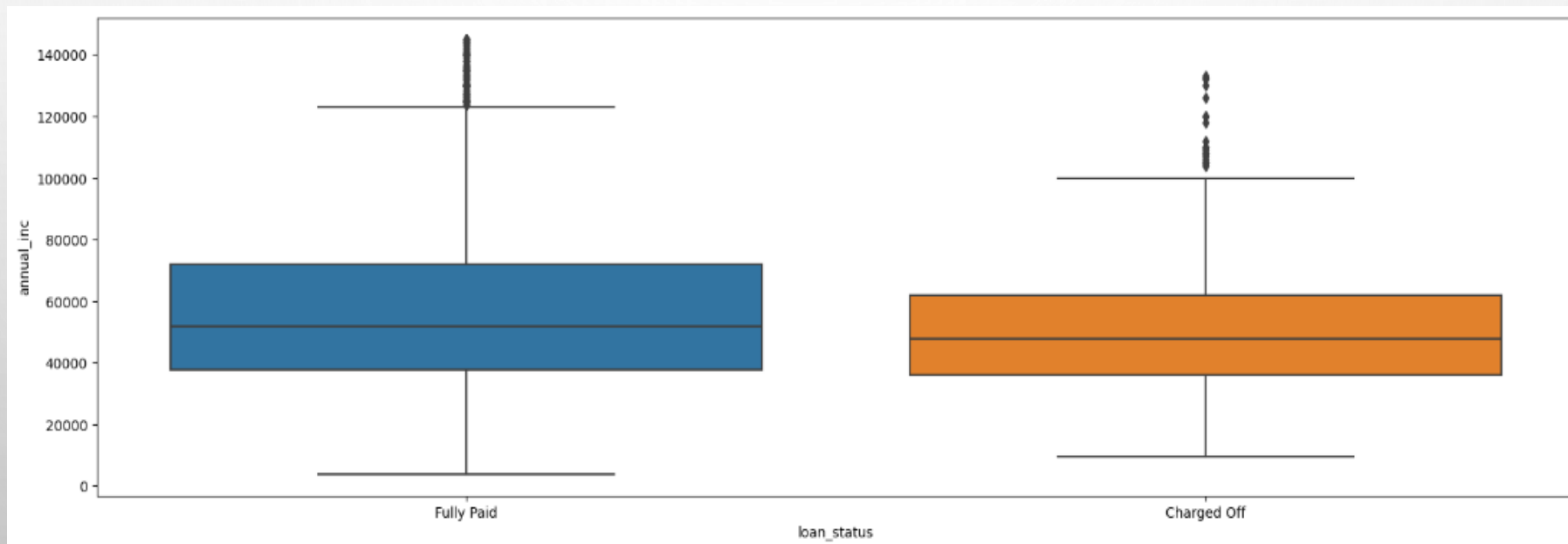
• DISTRIBUTION OF LOAN AMOUNT BASED ON PURPOSE

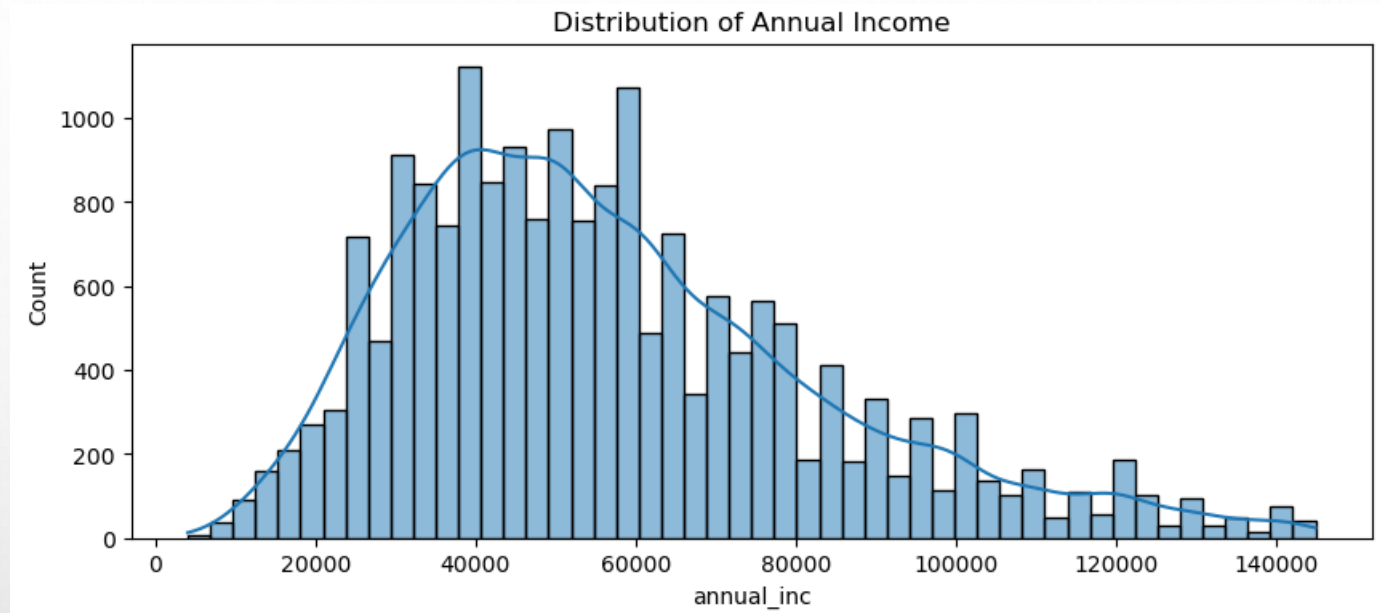
- LOAN AMOUNT IS HIGHER FOR THE PURPOSE OF **DEBT CONSOLIDATION, SMALL BUSINESS AND HOUSE PURCHASE**
- LOAN AMOUNT TAKEN IS LEAST FOR THE PURPOSE OF **VACATION**
- **MAJOR PURCHASE** HAS THE MOST OUTLIERS
- MOST CLIENT (BASED ON MEDIAN LINE) TEND TO DEFAULT WHO TAKES BIGGER LOAN FOR DEBT CONSOLIDATION ,SMALL BUSINESSES , HOME IMPROVEMENT AND EDUCATION



LOAN STATUS AGAINST ANNUAL INCOME

- CLIENT WITH **HIGH INCOME** GENERALLY **PAYS OFF THE LOAN** MORE OFTEN



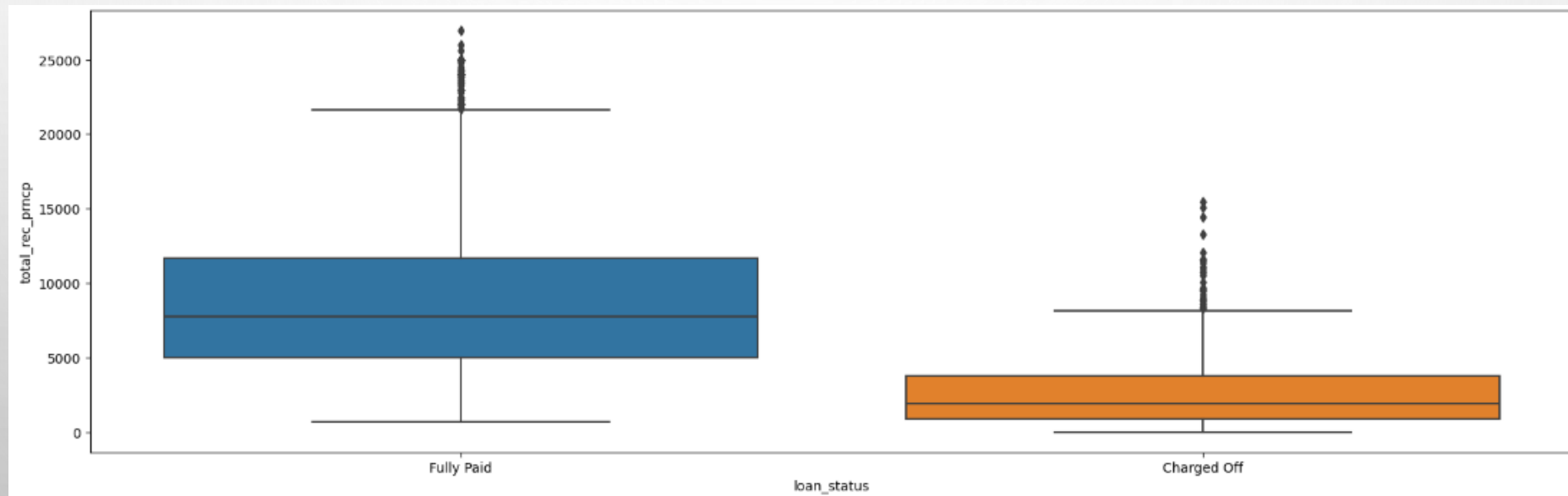


DISTRIBUTION OF ANNUAL INCOME

- CLIENTS WITH INCOME BETWEEN 35000 AND 60000 TENDS TO TAKE UP LOAN MORE OFTEN
- AS THE INCOME INCREASES, CLIENT IS LESS LIKELY TO TAKE UP LOAN

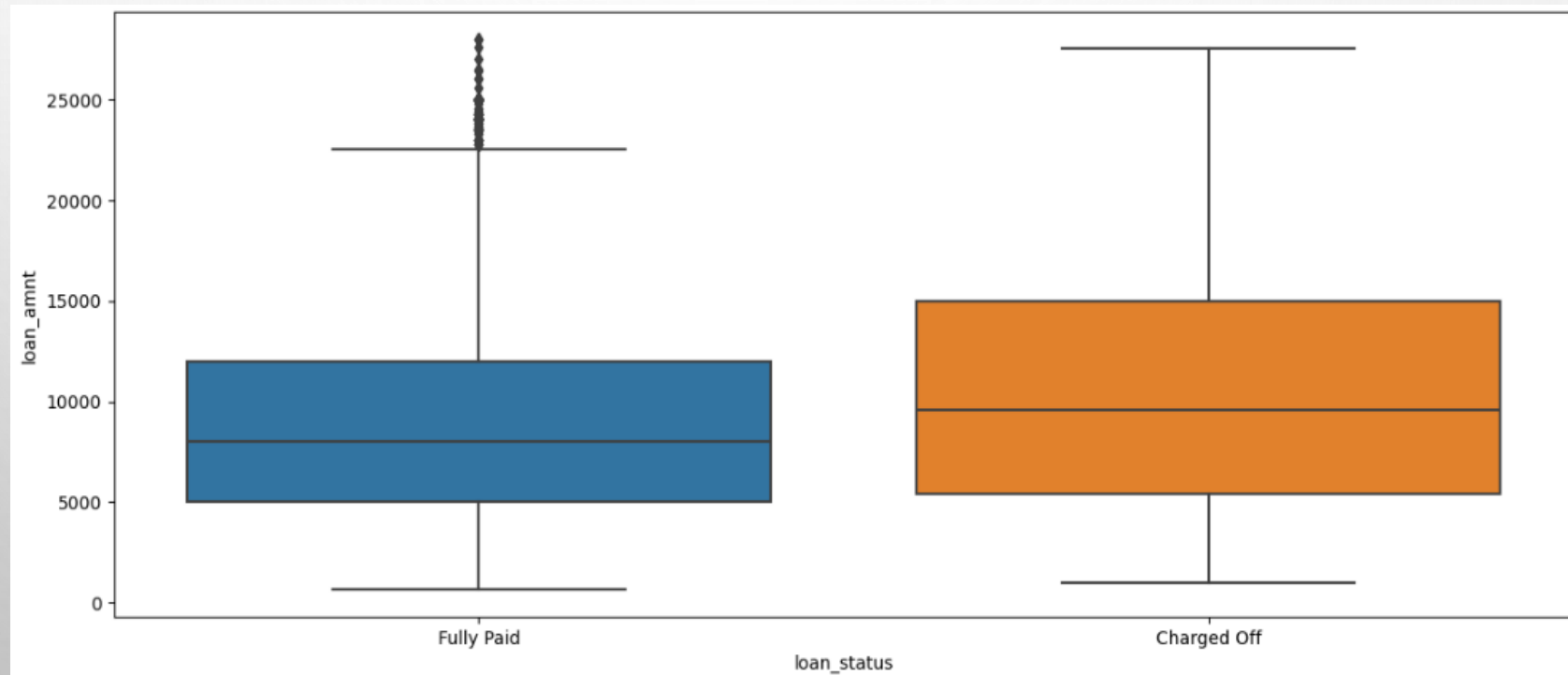
PRINCIPAL RECEIVED TO DATE VS LOAN STATUS

- CLIENTS WHO PAY THE PRINCIPAL AMOUNT ARE MORE LIKELY TO PAY THE LOAN OFF



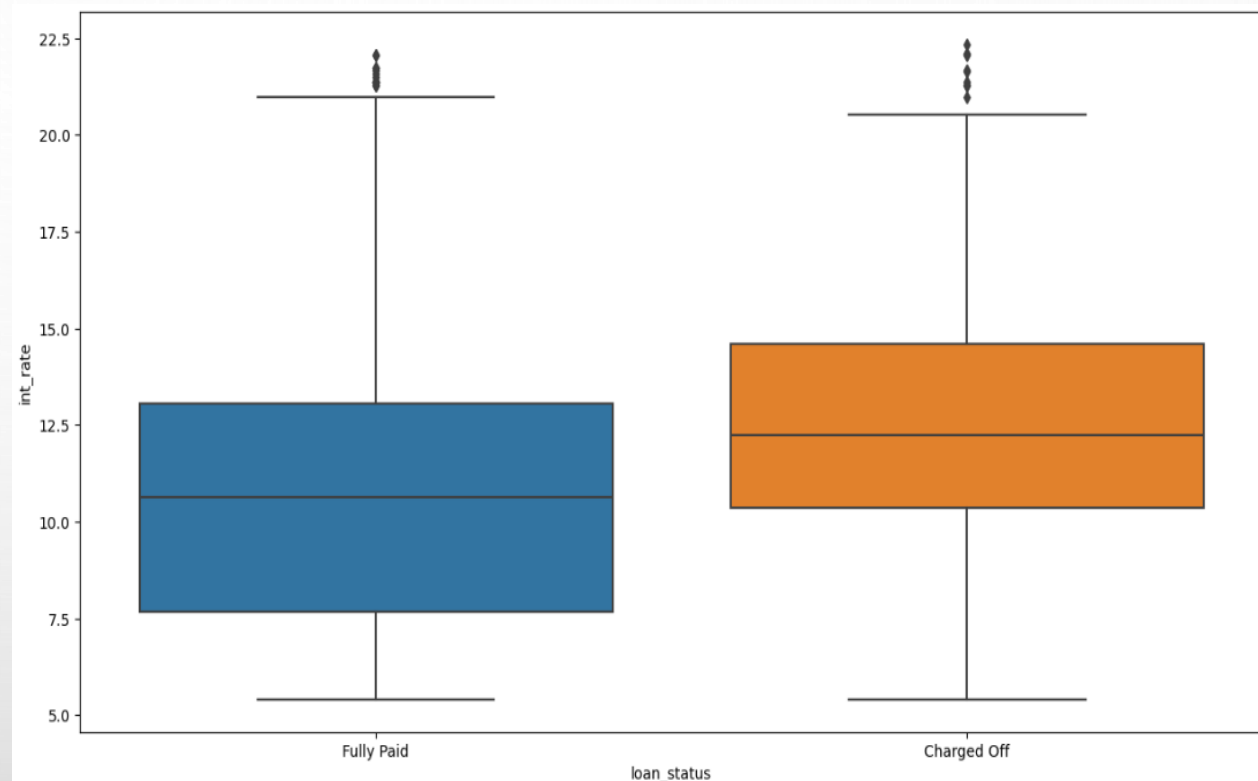
LOAN STATUS VS LOAN AMOUNT

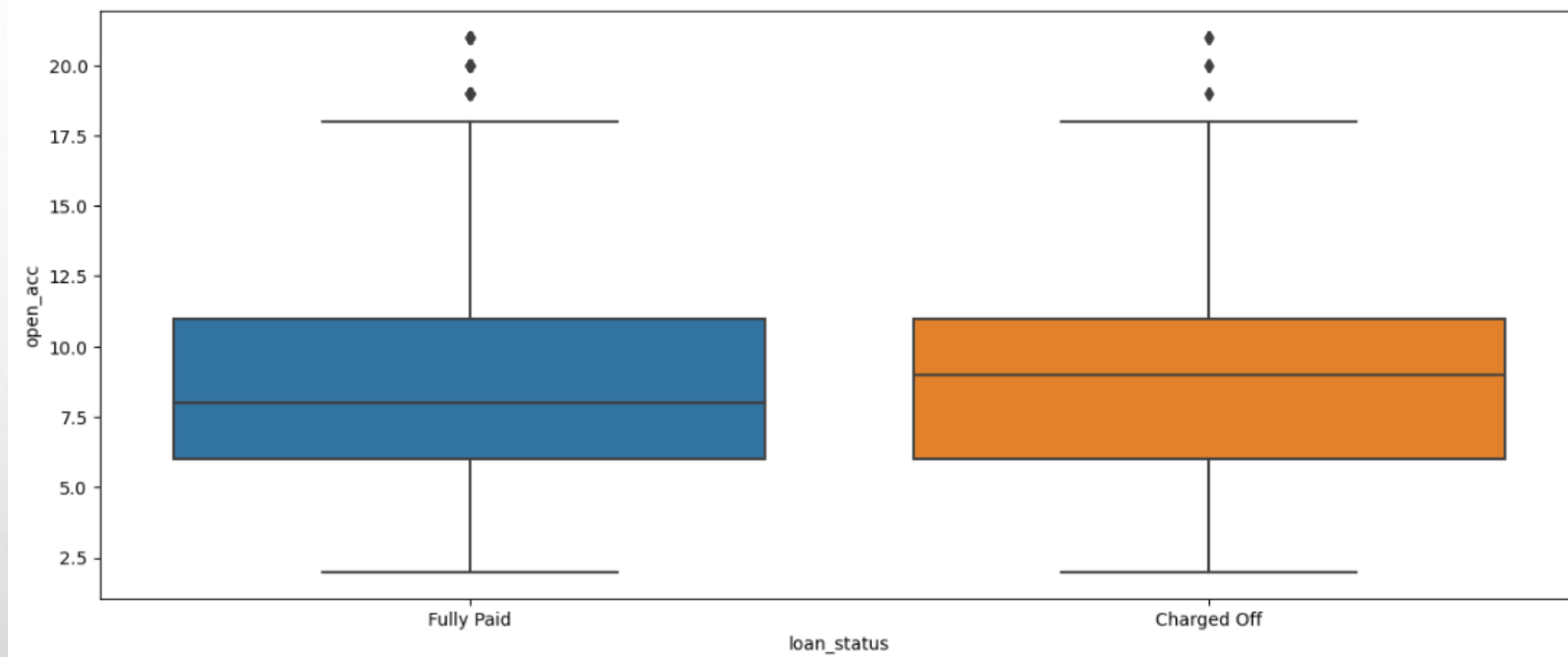
- **HIGHER** THE **LOAN** AMOUNT, **MORE LIKELY** THE **LOAN** WILL BE **CHARGED-OFF**
- LOAN WITH STATUS OF FULLY PAID HAS HIGH OUTLIERS



LOAN STATUS AGAINST INTEREST RATE

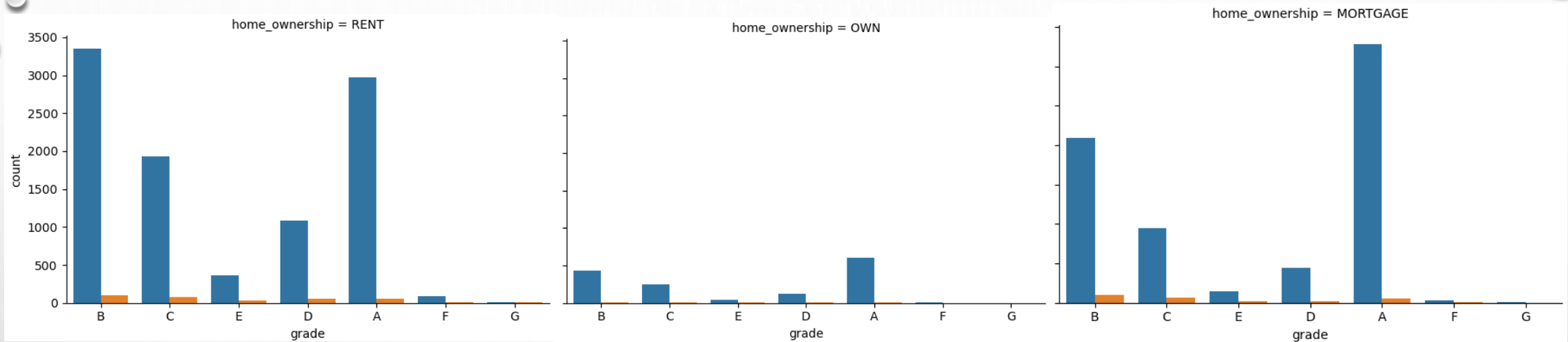
- HIGHER THE INTEREST RATE, LOAN MORE LIKELY TO DEFAULT
- MEDIAN AND UPPER QUARTILE FOR CHARGED OFF LOAN IS HIGHER





NUMBER OF OPEN CREDIT LINES (OPEN ACCOUNTS) VS LOAN STATUS

- OPEN ACCOUNTS ARE SIMILAR FOR THE FULLY PAID AND CHARGED OFF
- MEDIAN OF OPEN ACCOUNTS IS HIGHER FOR CHARGED-OFF LOAN

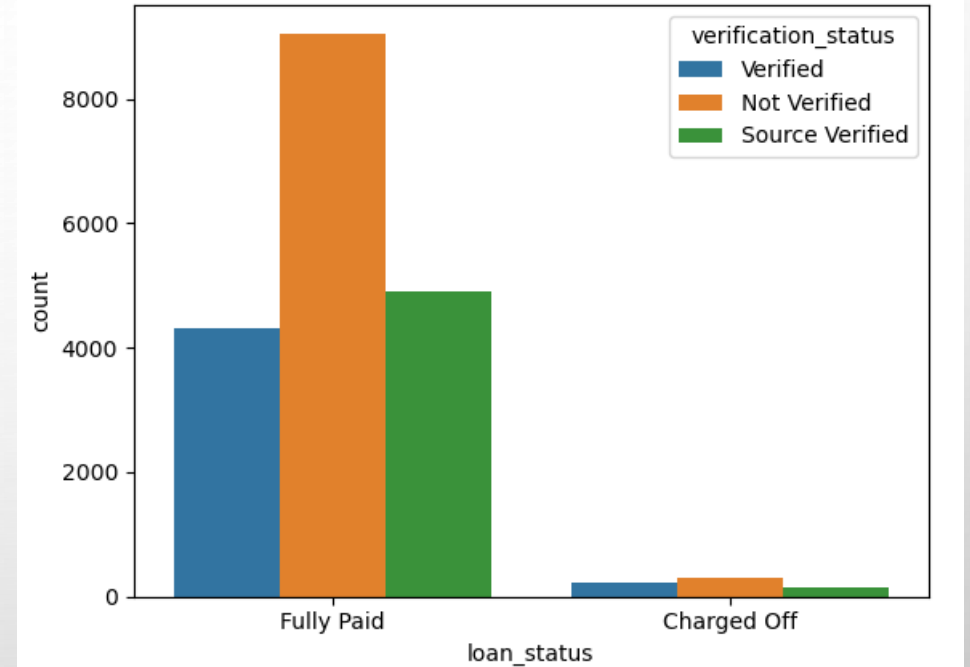


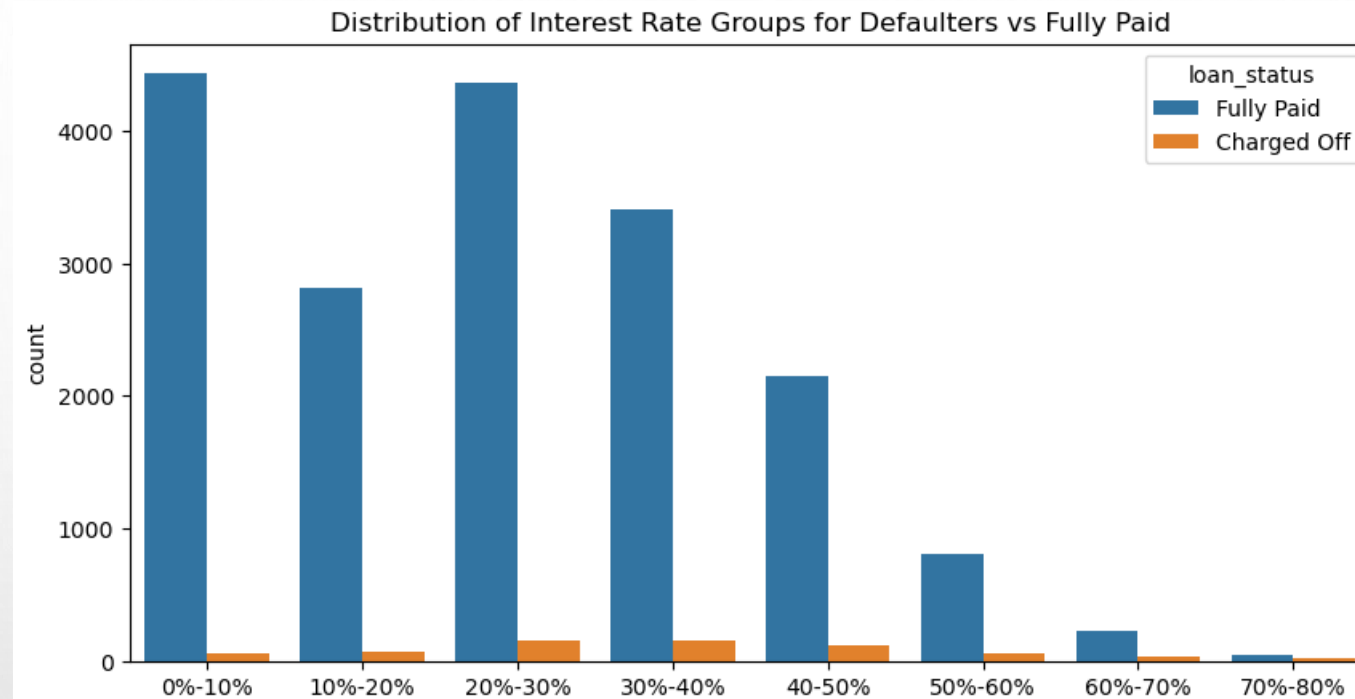
RISK-GRADE VS HOME OWNERSHIP

- ACROSS VARIOUS HOME OWNERSHIP, GRADE A AND B HAS THE MOST LOAN PAID OFF
- CLIENT WHO LIVES IN RENTED OR MORTGAGE HOUSE TENDS TO PAY FULL LOAN AMOUNT

VERIFIED INCOME WITH RESPECT TO LOAN STATUS

- MOST OF THE INCOME OF CLIENT IS UNVERIFIED
- CHANCES OF CHARGED-OFF LOAN INCREASES IF THE INCOME IS UNVERIFIED

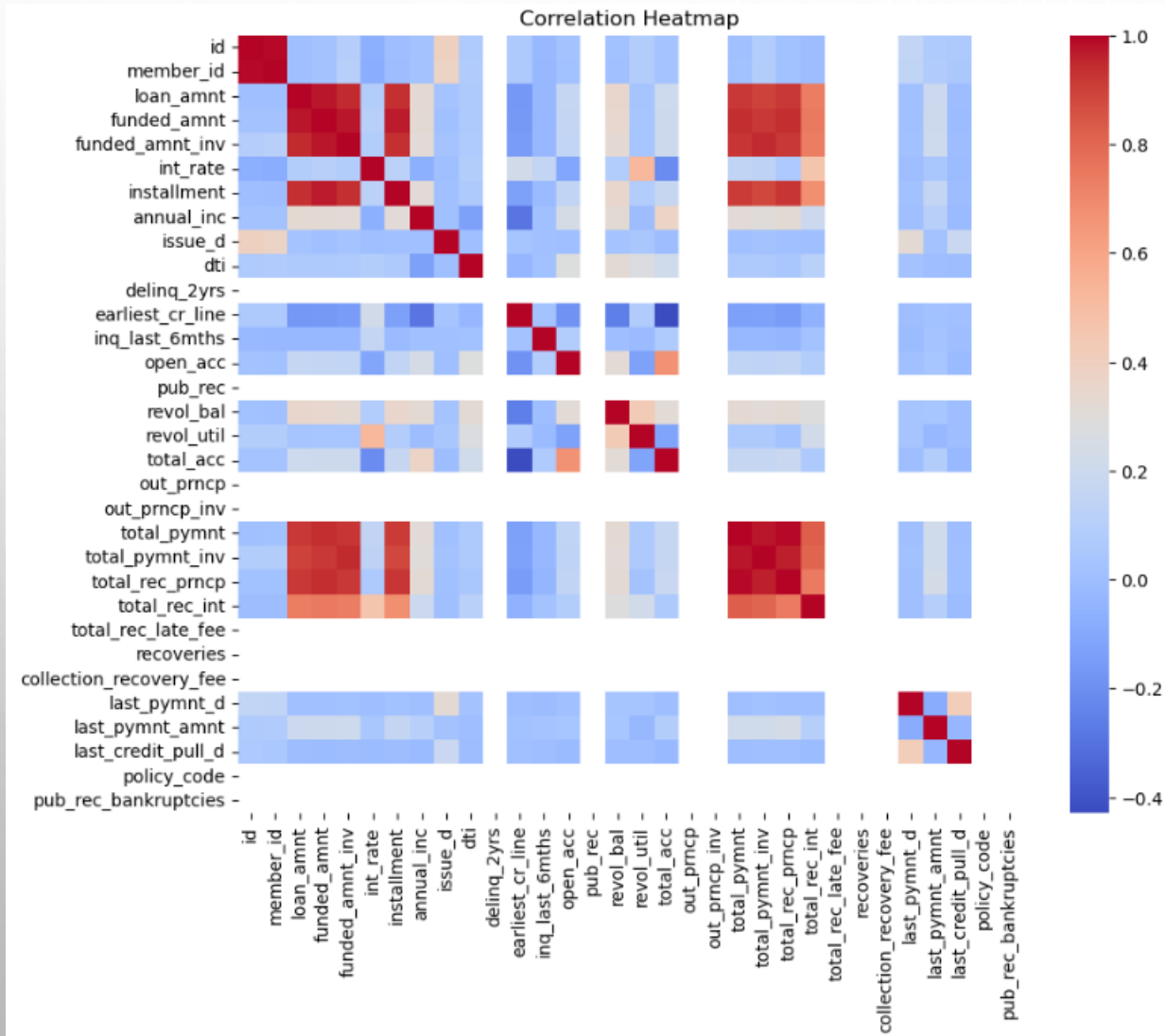




LOAN STATUS AGAINST INTEREST RATE

- MOST OF THE CLIENTS HAS TAKEN INTEREST RATE IN **RANGE 0%-10% AND 20%-30%**
- RISK OF CHARGED-OFF LOAN INCREASES WHEN THE INTEREST RATE IS IN **RANGE 20%-50%**
- RATIO OF CHARGED-OFF LOAN AGAINST FULLY PAID LOAN IS MAXIMUM IN RANGE 40-50%

CORRELATION MATRIX



CORRELATION HEATMAP

- INTEREST RATE IS INVERSELY PROPORTIONAL TO LOAN AMOUNT, DTI
- INSTALLMENT AMOUNT IS HIGHLY CORRELATED TO LOAN AMOUNT AND TOTAL PAYMENT MADE
- 'ID', 'MEMBER_ID', 'FUNDED_AMNT', 'FUNDED_AMNT_INV', 'INSTALLMENT', 'ISSUE_D', 'TOTAL_PYMNT', 'TOTAL_PYMNT_INV', 'TOTAL_REC_PRNCP', 'TOTAL_REC_INT', 'LOAN AMOUNT' ARE HIGHLY CORRELATED VARIABLES.

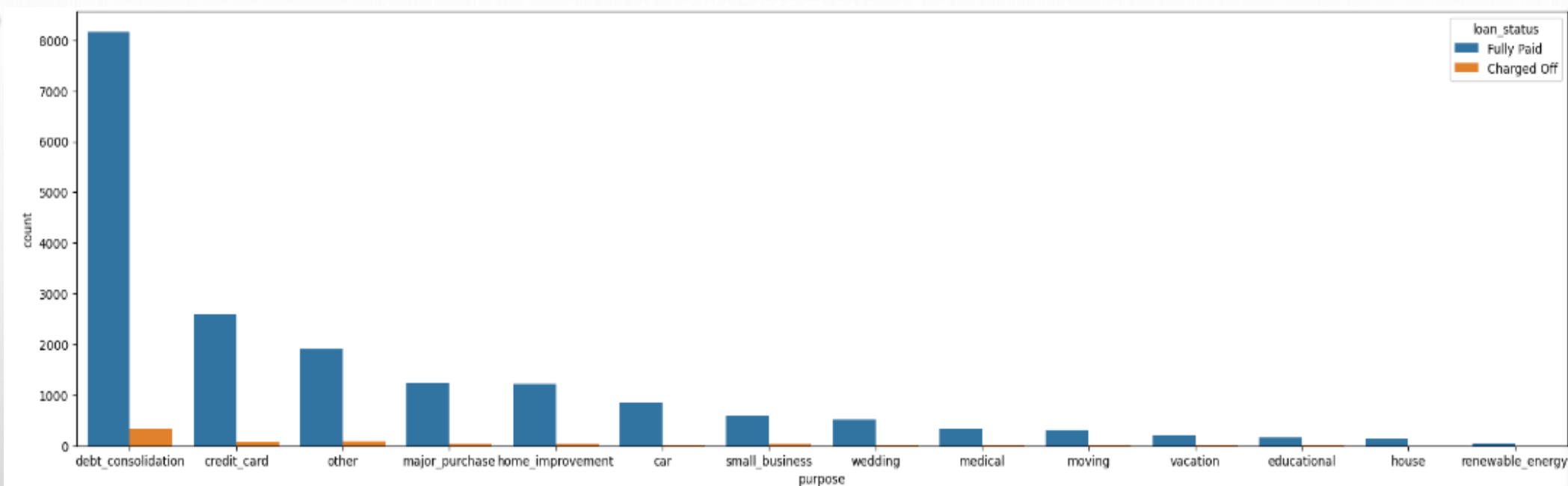
CONCLUSION

- AROUND **3.5%** OF THE LOAN OFFERED BECOMES **BAD LOAN**
- AROUND **15%** LOANS FALL INTO **MEDIUM TO HIGH-RISK ZONE**
- MOSTLY LOANS ARE PREFERRED BY CLIENTS WITH OVER **10 YEARS INTO EMPLOYMENT**
- CLIENTS WITH INCOME BETWEEN 35000 AND 60000 TENDS TO TAKE UP LOAN MORE OFTEN
- MOSTLY CLIENTS WHO LIVE IN **RENTED** OR **MORTGAGE** HOUSES THE PAY THE LOAN FULLY
- MOST CLIENT (BASED ON MEDIAN LINE) TEND TO DEFAULT WHO TAKES BIGGER LOAN FOR DEBT CONSOLIDATION ,SMALL BUSINESSES , HOME IMPROVEMENT AND EDUCATION
- CREDIT CARD LOAN ARE MORE RISKER THAN LOAN FOR DEBT CONSOLIDATION
- MOST OF THE INCOME OF CLIENT IS UNVERIFIED
- HIGHER THE INTEREST RATE, LOAN MORE LIKELY TO DEFAULT
- HIGHER THE LOAN AMOUNT, MORE LIKELY THE LOAN WILL BE CHARGED-OFF
- LOAN AMOUNT IS HIGHER FOR THE PURPOSE OF **DEBT CONSOLIDATION, SMALL BUSINESS AND HOUSE PURCHASE**



THANK YOU

APPENDIX



LOAN STATUS BASED ON PURPOSE

- **DEBT CONSOLIDATIONS HAS MORE LOAN DEFAULTERS** THAN REST OF THE CATEGORIES (PURPOSE)
- CREDIT CARD LOAN ARE MORE RISKER THAN LOAN FOR DEBT CONSOLIDATION