Lending Club Case Study EDA

10-Nov-2021

Sindhu Sivaraj

Business Case

Process Overview:

- Lending club is a consumer finance company, specializing in lending various type of loans to urban customers
    - If the applicant is likely to repay the loan, then not approving the loan results in a loss of business
    - If the applicant is not likely to repay the loan (likely to default), then approving the loan may lead to financial loss

Objective:

- To find the patterns which influence the applicants to default their loan based on historical transactions

Scope:

- Accepted loan transaction from the period of 2007 to 2011

EDA - Method Adopted

Understanding the data → Data Cleaning → Univariate Analysis → Multivariate Analysis → Segmented Analysis → Insights

- Understanding the shape of data (rows & columns) & the relevant data types
- Finding the missing values, outliers & treating them
- Fixing the datatypes & creating derived columns for segmented analysis
- Analysis using single variable to derive insights
- Analysis of multiple variables against the target variable & correlation analysis
- Segmented analysis using derived columns against the target variable
- Deriving meaningful insights from the analysis & recommendations
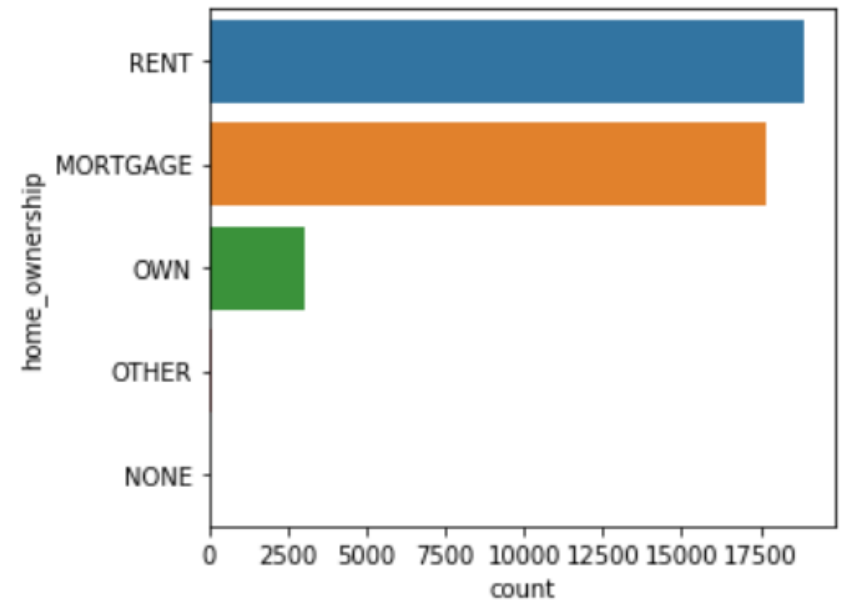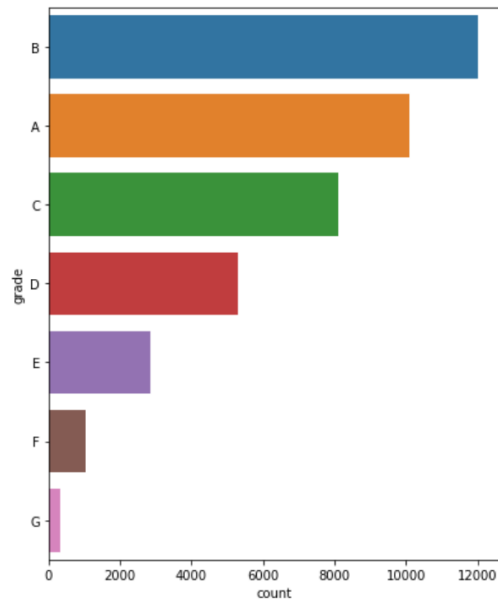
Univariate Analysis

- Loan Amount ranges from 500 to 35000
- Most of loan amounts are rounded off in multiples of 5000
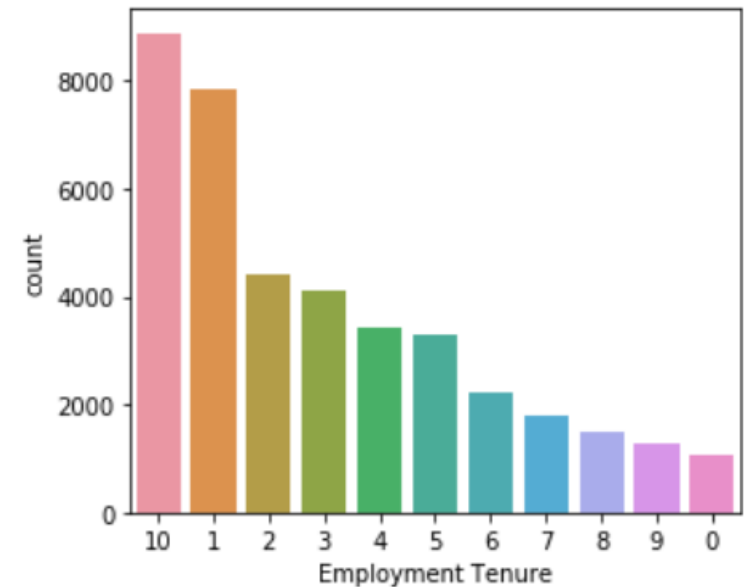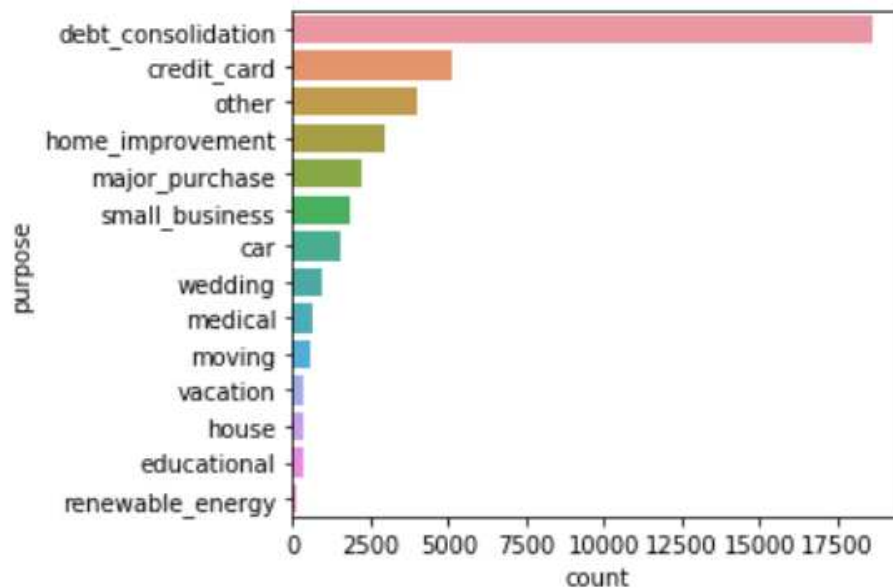- Majority of loans have Interest rate in the range of 9% to 15%
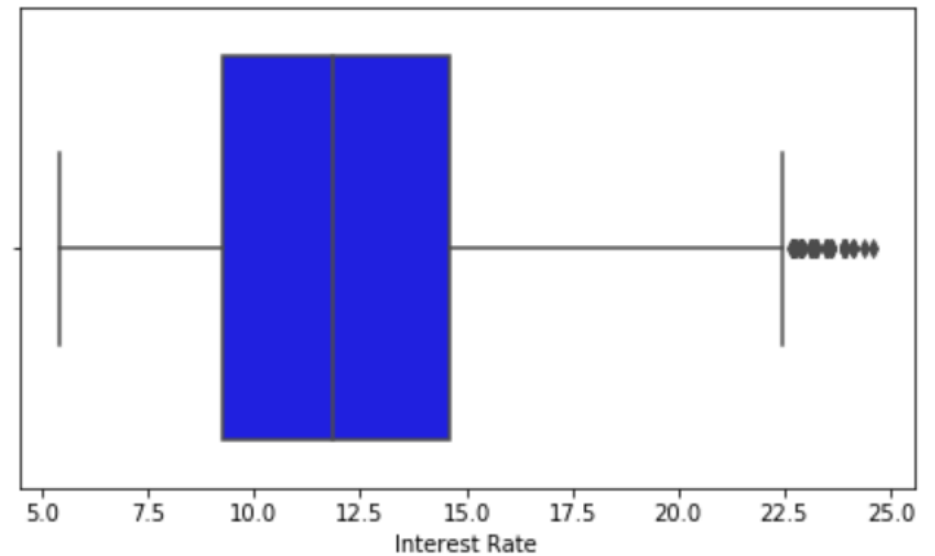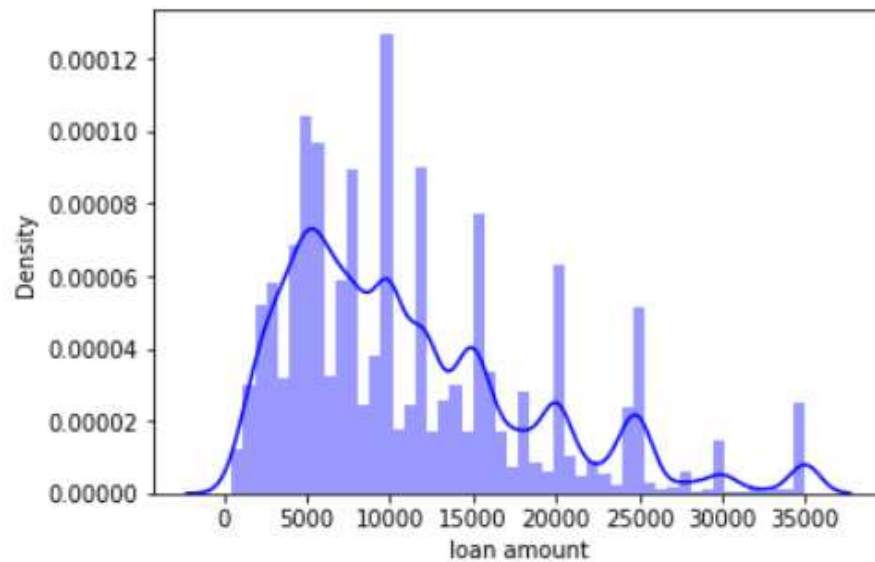
Univariate Analysis

- Major loan applicants are in grades B, A & C
- Majority of people who apply for loan stays on Rent or has mortgage
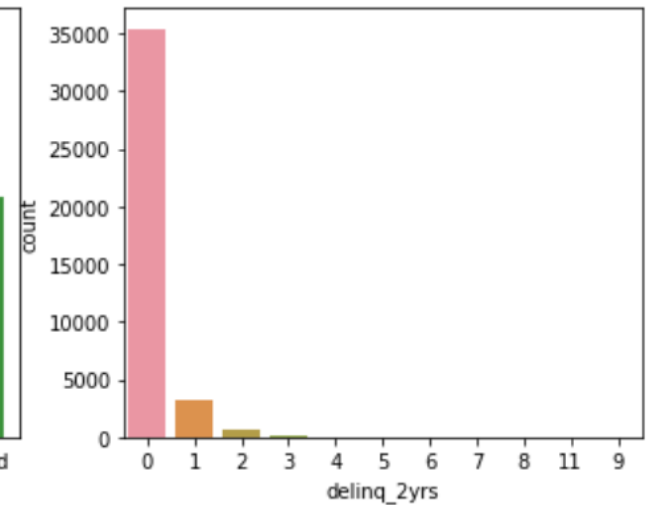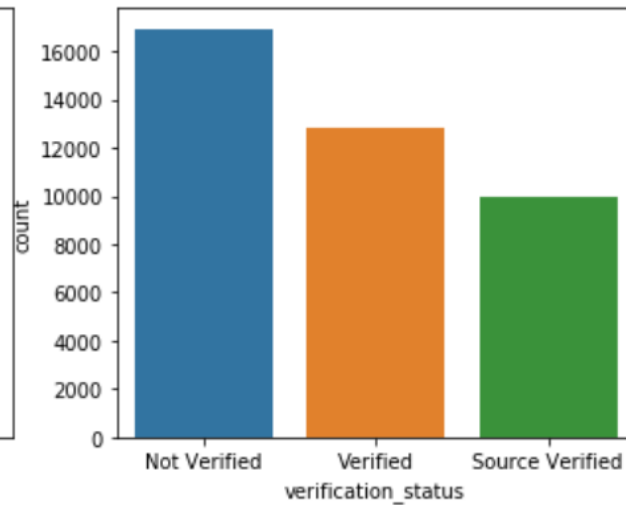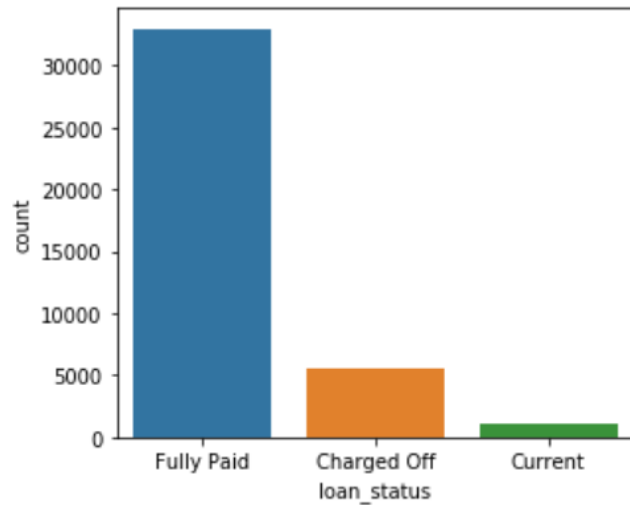
Univariate Analysis

- Top purpose of loan application is for debt consolidation
- Majority of people who apply for loan have 10+ years of employment followed by <=1 year tenure
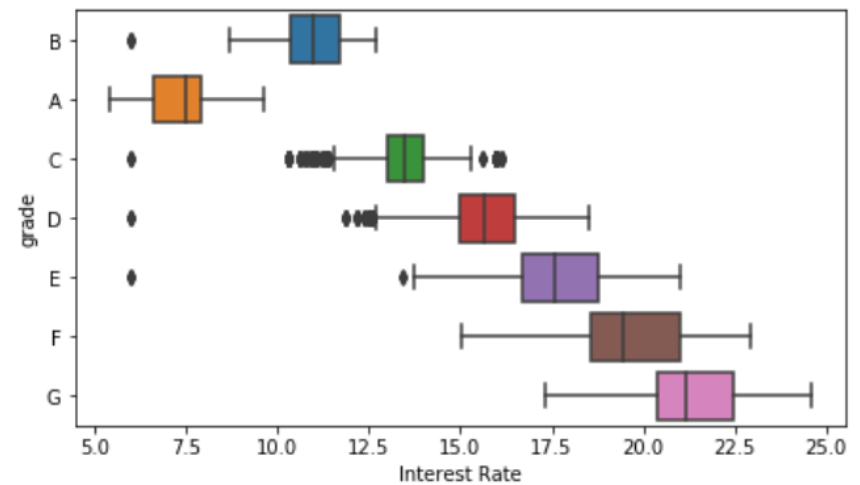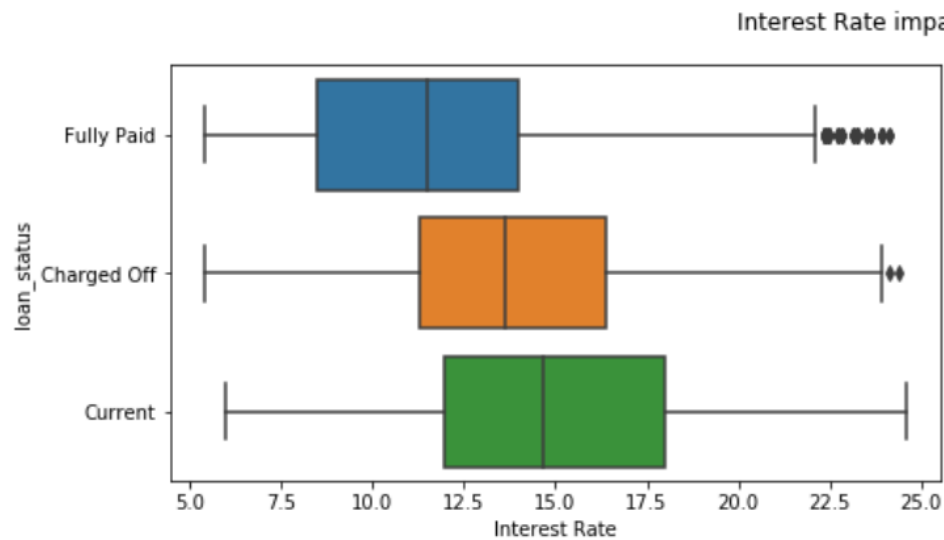
Univariate Analysis

- 14% of loans are charged off (defaulted) & 83% of loans fully paid
- The income is either verified or source verified for major of loan applications.
- Delinquency in past 2 years is mostly nil

Bivariate Analysis

○ Loans which are defaulted have higher interest rate (11% - 16%) compared to the loans that are fully paid (8% - 14%)

○ Interest rate is lowest for A grade, and it increases in order of B < C < D < E < F < G

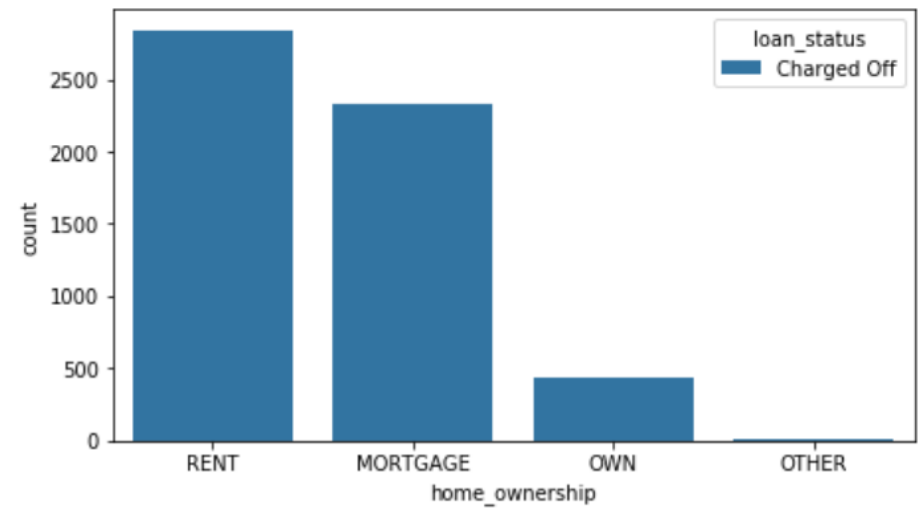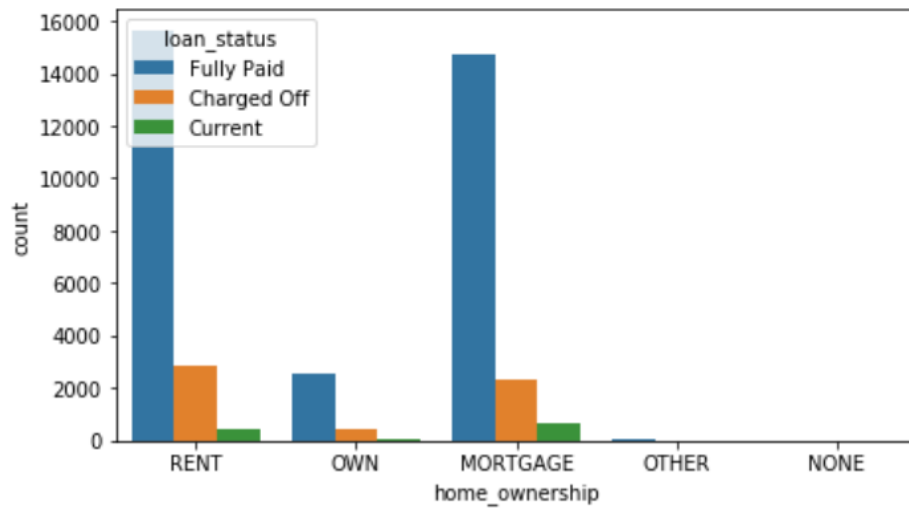Interest Rate impact on Grade & loan status

Bivariate Analysis

o People who are in rent or mortgage have defaulted their loans in higher % compared to the people who own the house
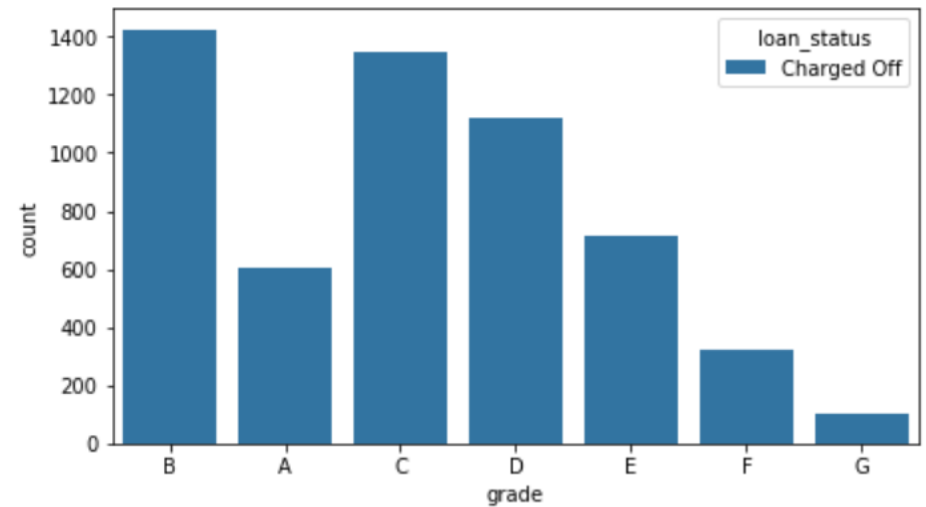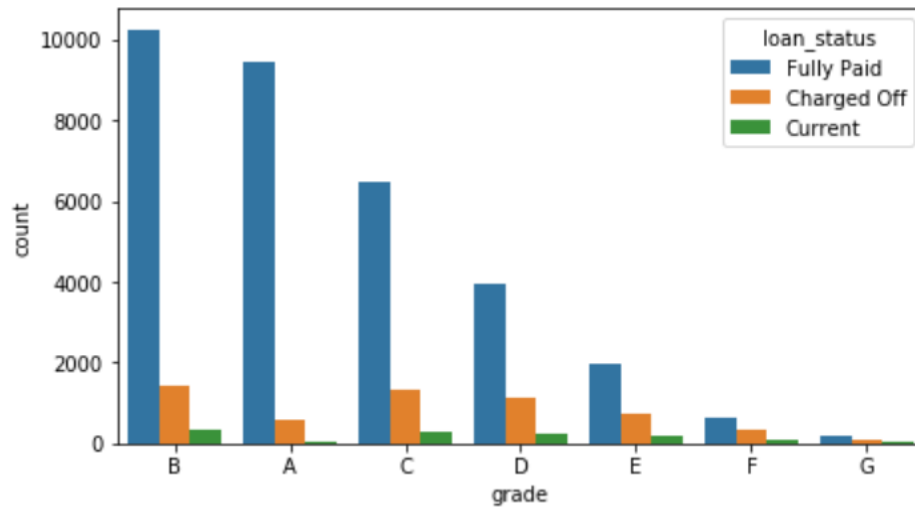

Impact of home ownership on loan status

Bivariate Analysis

o Loan Applicants with grade A has paid their loan & has low default rate compared to other grades

Impact of grade on Loan Status

Bivariate Analysis

o Loan borrowers are higher for 36 months term

o The loans with 60 months term has more default rate % compared to 36 months

Impact of loan term on loan status
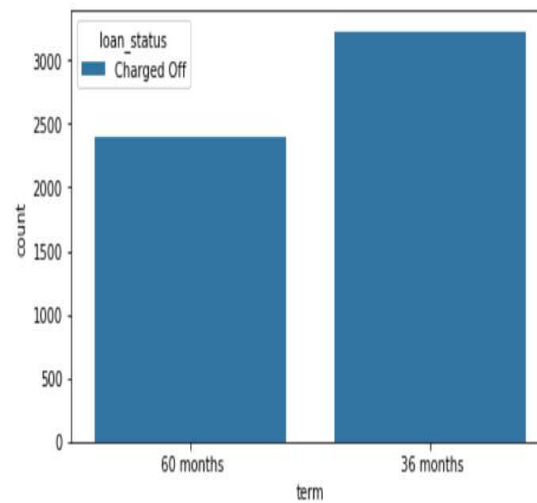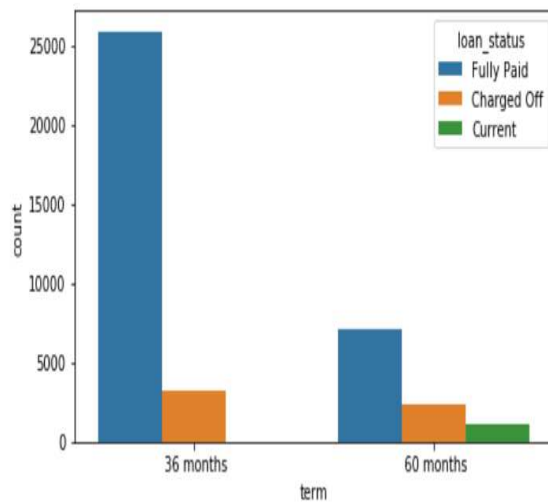
| loan_status	Charged Off	Current	Fully Paid	Total	Charged_Off %
term					
60 months	2400.0	1140.0	7081.0	10621.0	22.596742
36 months	3227.0	0.0	25869.0	29096.0	11.090872

Bivariate Analysis

o Debt consolidation, Credit card are the top purpose for the application of loan.

o Default rate is higher in the loans applied for small business, renewable energy and education
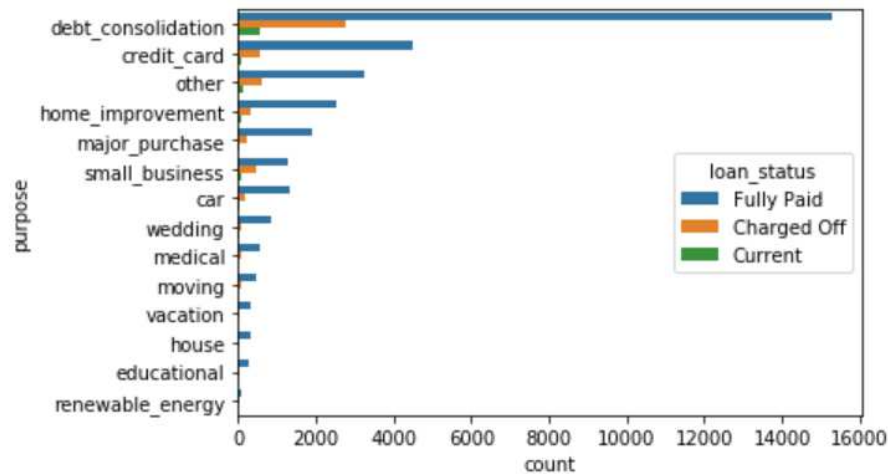
loan_status	Charged Off	Current	Fully Paid	Total	charged_off %
purpose					
small_business	475.0	74.0	1279.0	1828.0	25.984683
renewable_energy	19.0	1.0	83.0	103.0	18.446602
educational	56.0	0.0	269.0	325.0	17.230769
other	633.0	128.0	3232.0	3993.0	15.852742
moving	92.0	7.0	484.0	583.0	15.780446
house	59.0	14.0	308.0	381.0	15.485564
medical	106.0	12.0	575.0	693.0	15.295815
debt_consolidation	2767.0	586.0	15288.0	18641.0	14.843624
vacation	53.0	6.0	322.0	381.0	13.910761
home_improvement	347.0	101.0	2528.0	2976.0	11.659946
credit_card	542.0	103.0	4485.0	5130.0	10.565302
car	160.0	50.0	1339.0	1549.0	10.329245
major_purchase	222.0	37.0	1928.0	2187.0	10.150892
wedding	96.0	21.0	830.0	947.0	10.137276

Bivariate Analysis

o Revolving line utilization rate is higher for defaulted borrowers compared to fully paid

Impact of revolving utilization rate on loan status
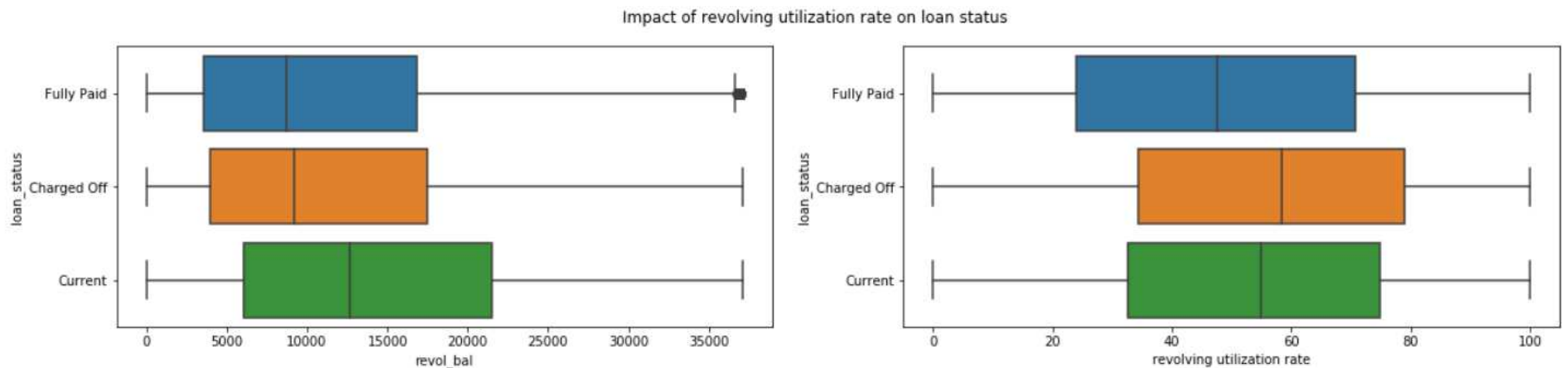
Segmented Analysis

o The loans with interest rate 11 -20% have higher default rate compared to the loans with interest rate <=10%
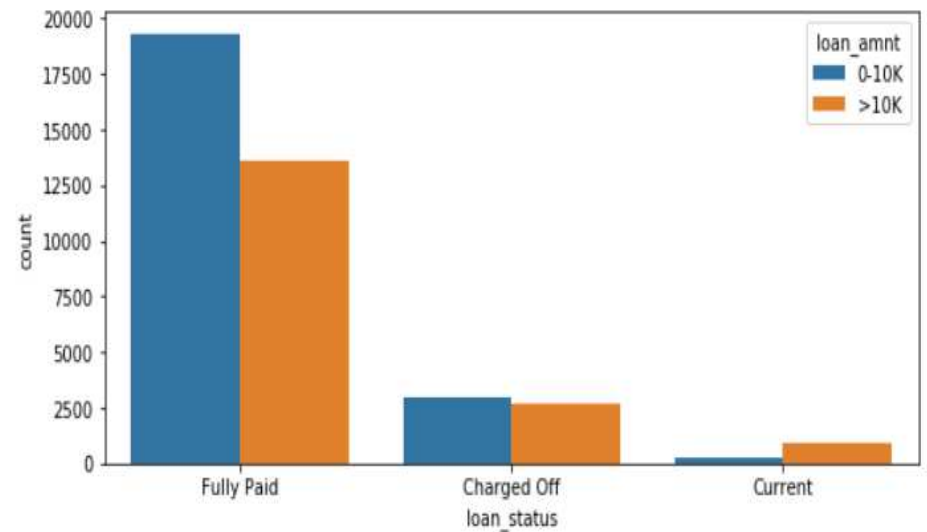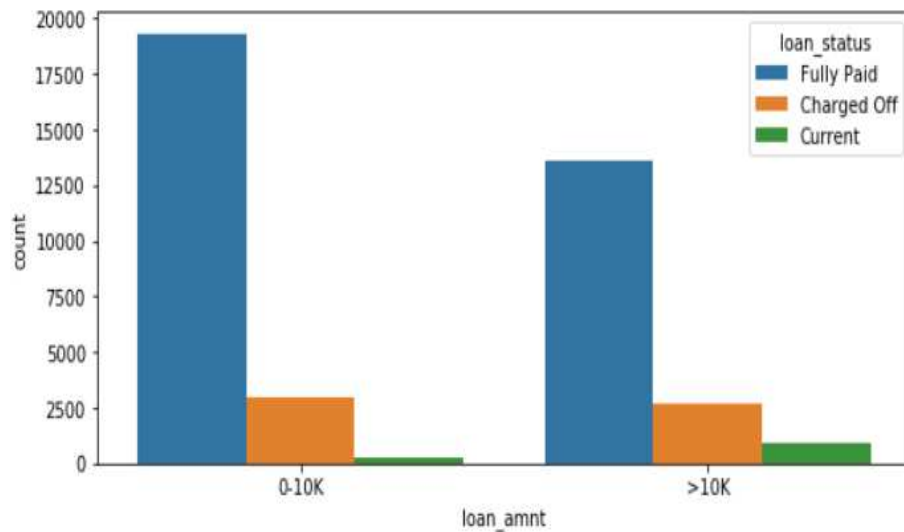

Impact of interest rate segment on loan status
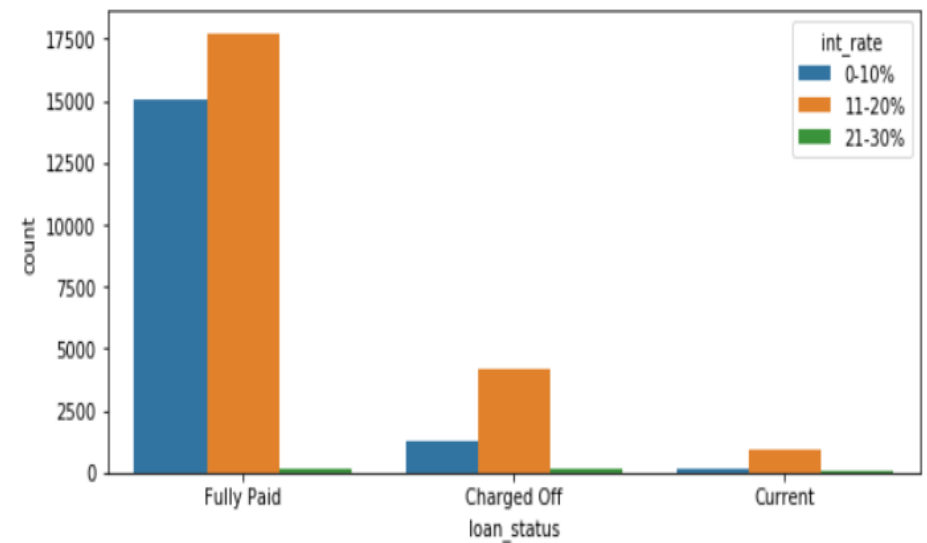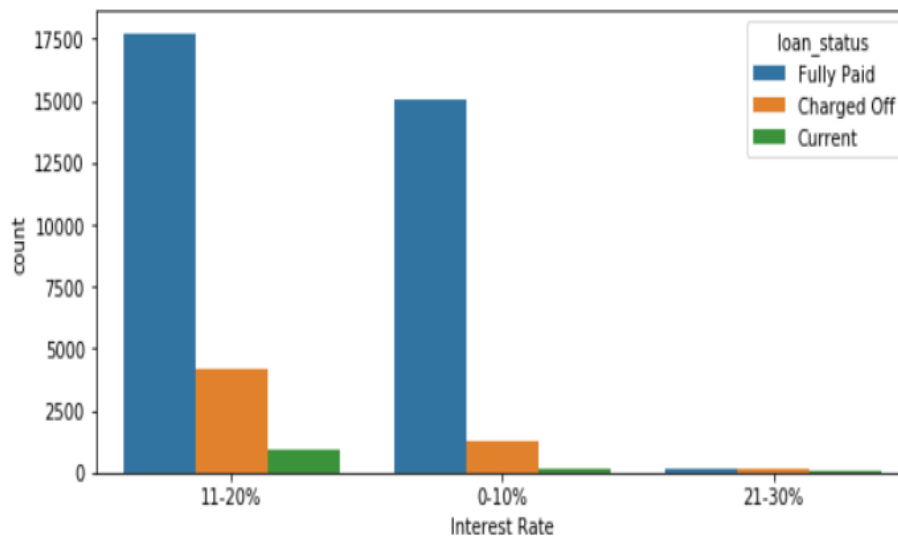
Segmented Analysis

- Most of borrowers apply for loans less than 10K.
- Considering default rate, borrowers with loan amount >10K tend to default marginally more than the borrowers with loan amount <10K

Impact of loan amount segment on loan status

Recommendations:

- Grade A borrowers have more chance of making full payments on their loan. Focusing & offering more loans for grade A borrowers can reduce the charge off instances

- Charge off % is higher for the loans borrowed for small business, renewable energy & education purposes. Extra scrutiny can be deployed on loan processing process in these segments

- Loan term of 60 months has 22% probability of being charged off against 11% for the term of 36 months. Loan tenure can be further reviewed & reduced in the possible cases

- Revolving line utilization rate is higher for defaulted borrowers. So, it can be set as an additional criteria while approving loans

- The loans with interest rate <= 10% have less probability to be not paid. Interest rate ranges can be reviewed & reduced to best fit the customers to avoid the charge offs