Lending Club Case Study

TEAM MEMBERS:

NITIN JAIN

MRITUNJAY SINGH

Problem Statement

In this case study, you will use EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of default.

If the company approves the loan, there are 3 possible scenarios described below else it will be rejected:

1. **Fully paid**: Applicant has fully paid the loan (the principal and the interest rate)

2. **Current**: Applicant is in the process of paying the instalments, i.e., the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.

3. **Charged-off**: Applicant has not paid the instalments in due time for a long period of time, i.e., he/she has **defaulted** on the loan
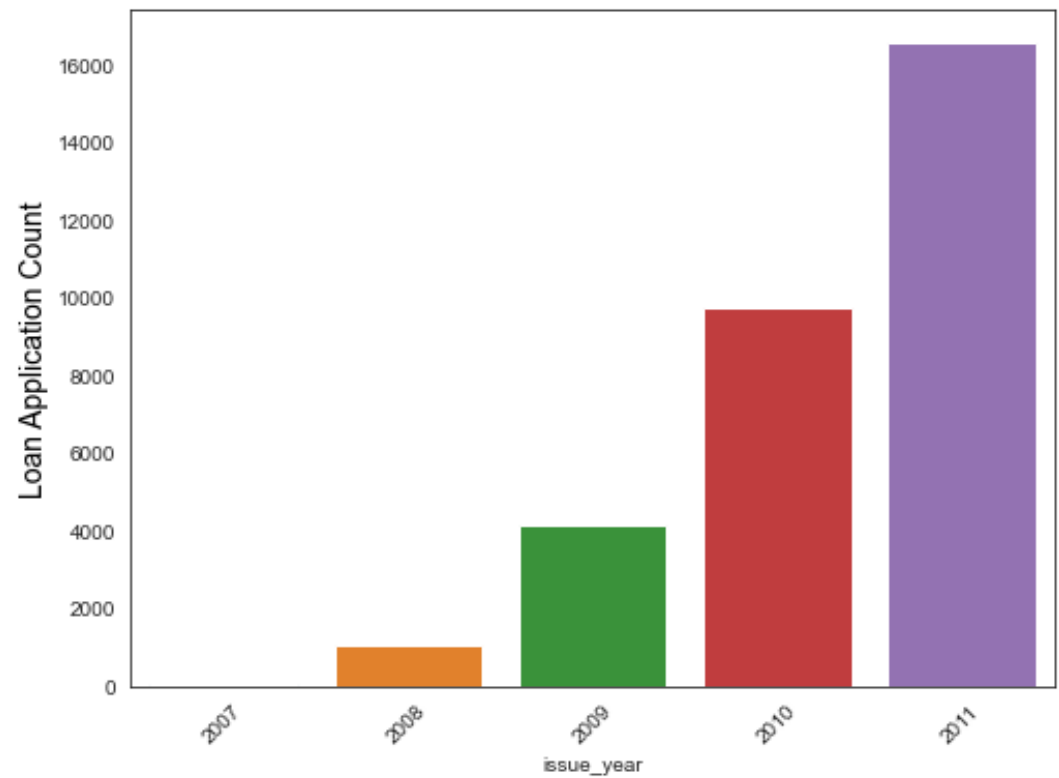
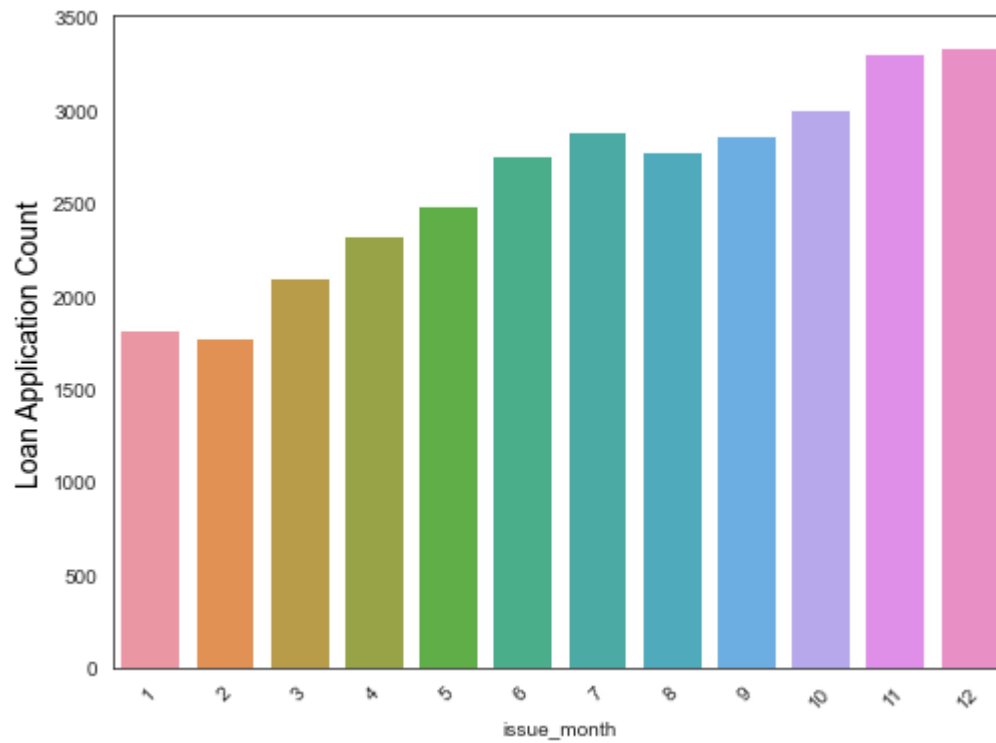# Analysis for categorical data
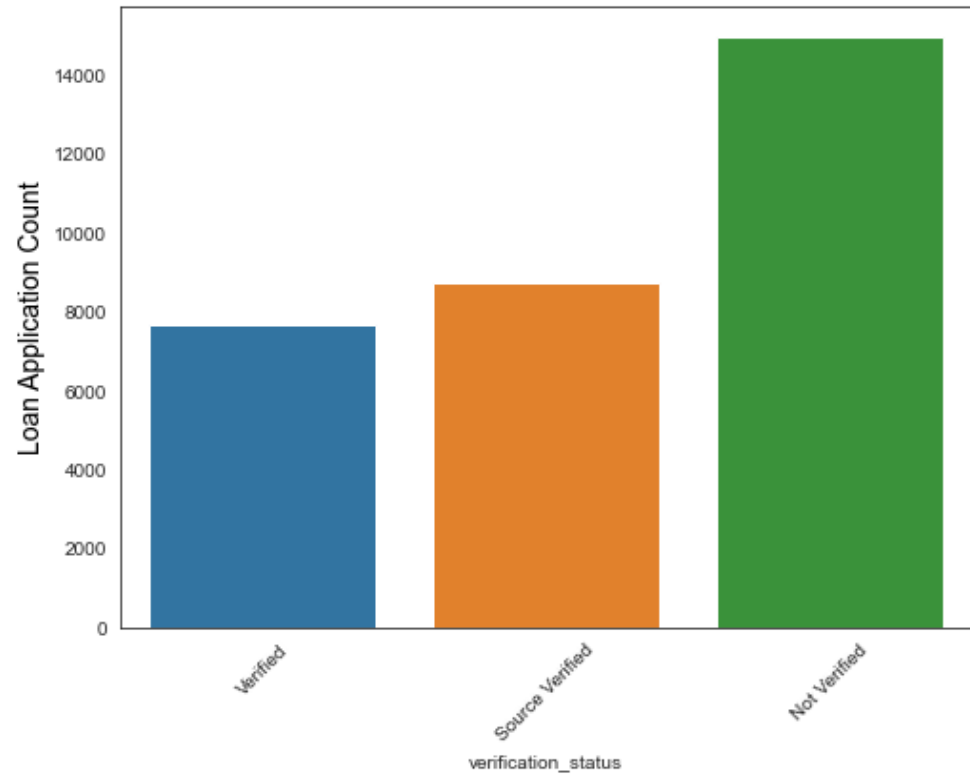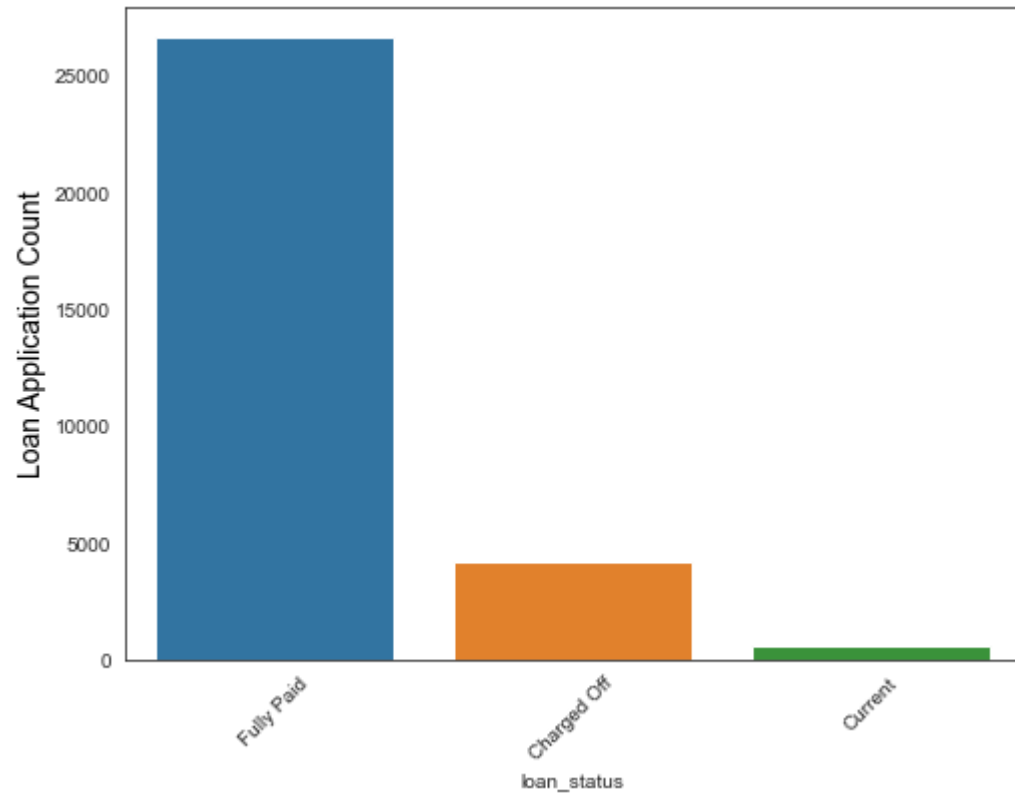
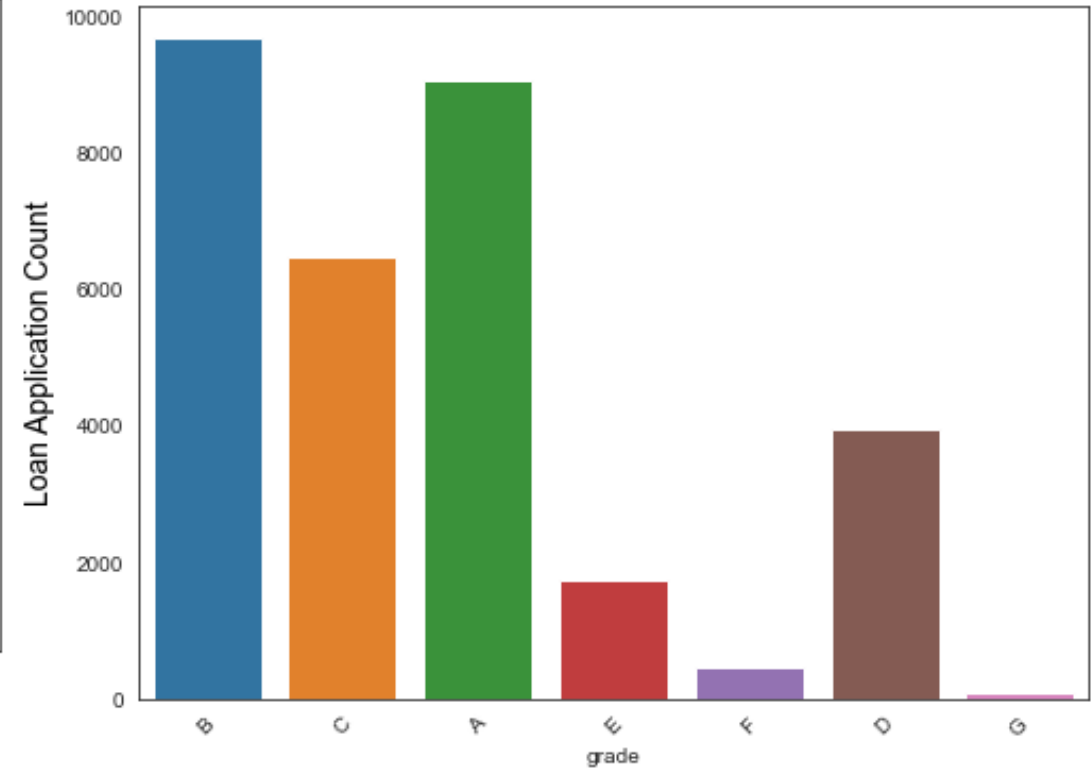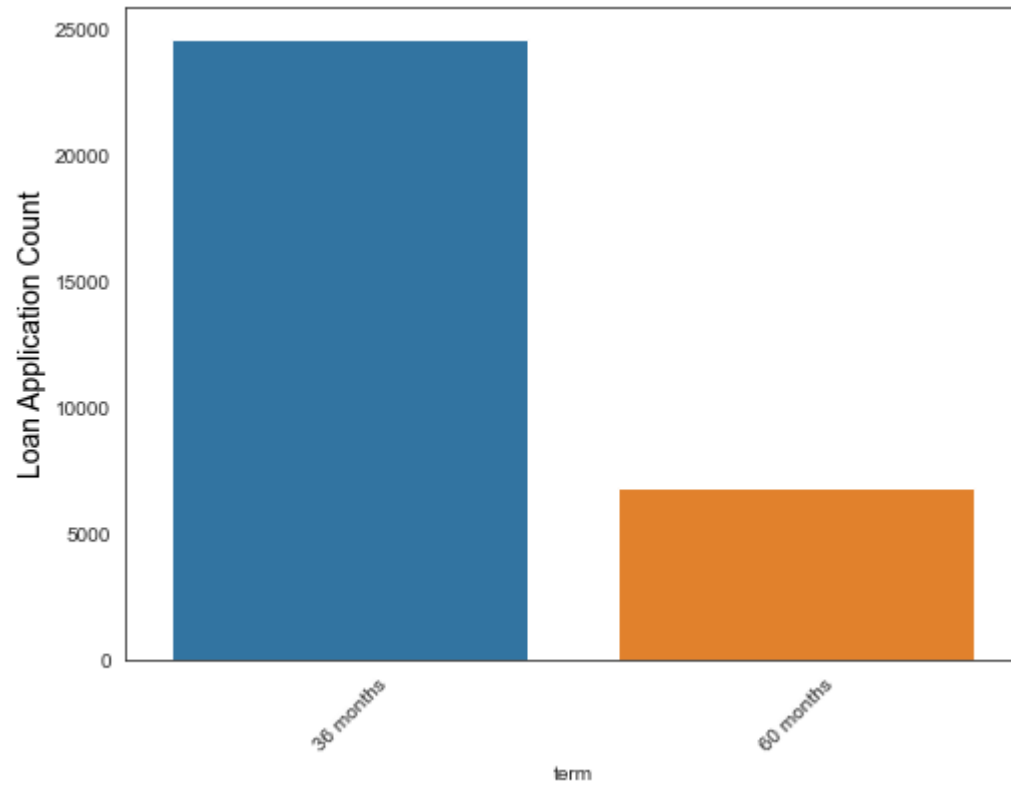# Analysis for categorical data
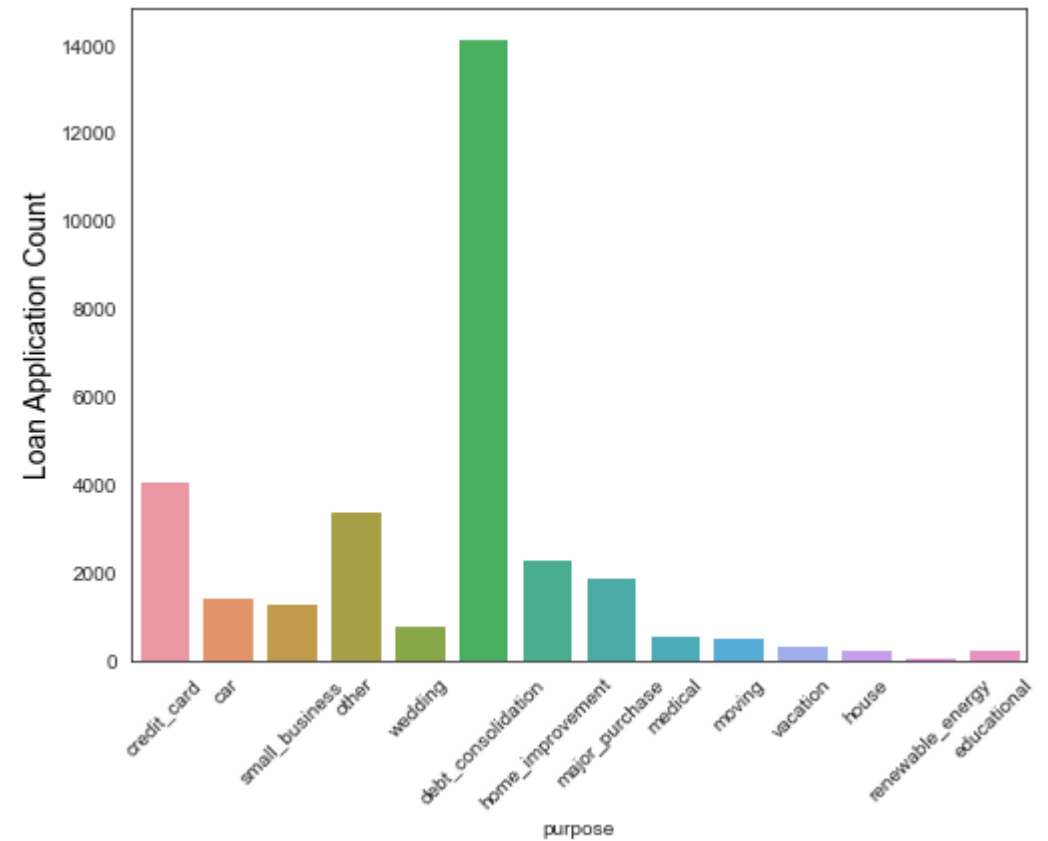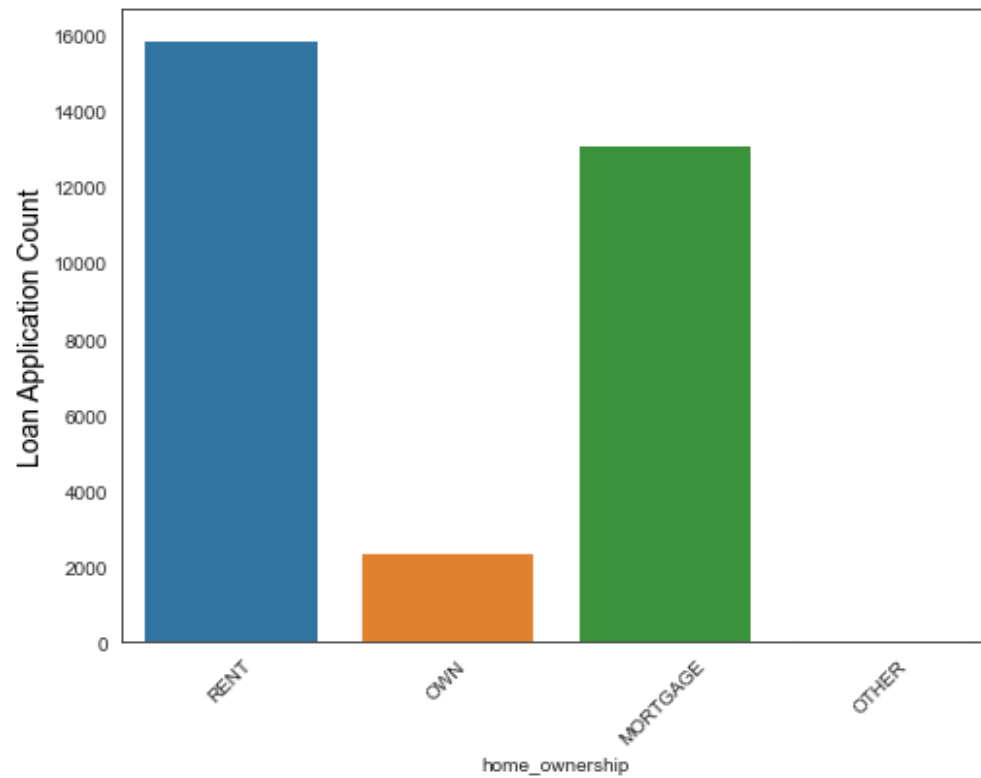
Observation

1. Most of the loan application are completed.

2. Most of the application have not verified their application.

3. Most of the application have home ownership as rented.

4. Most of the loan applicant are taking loan for the purpose of debt consolidation and least for the renewable energy.
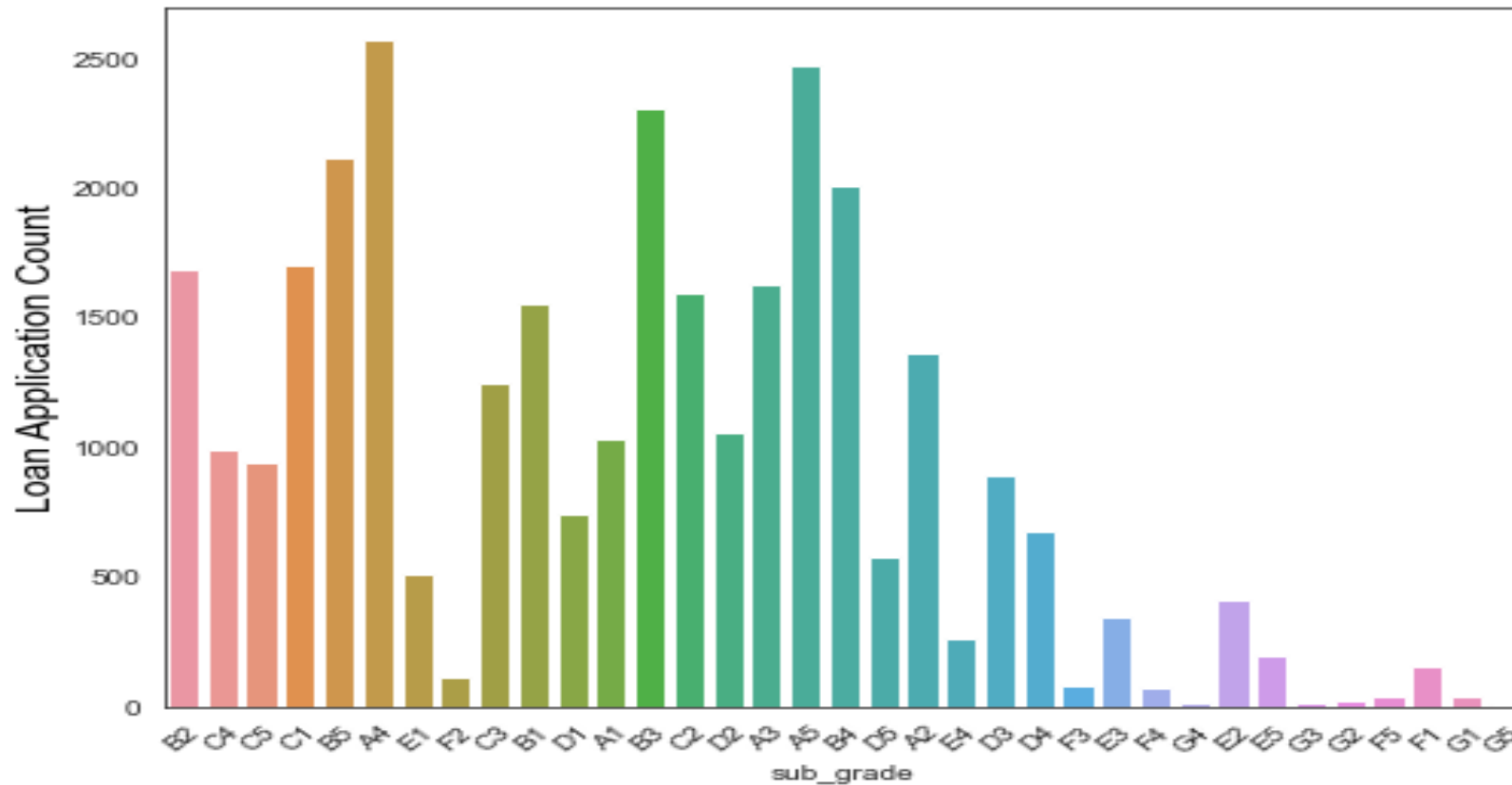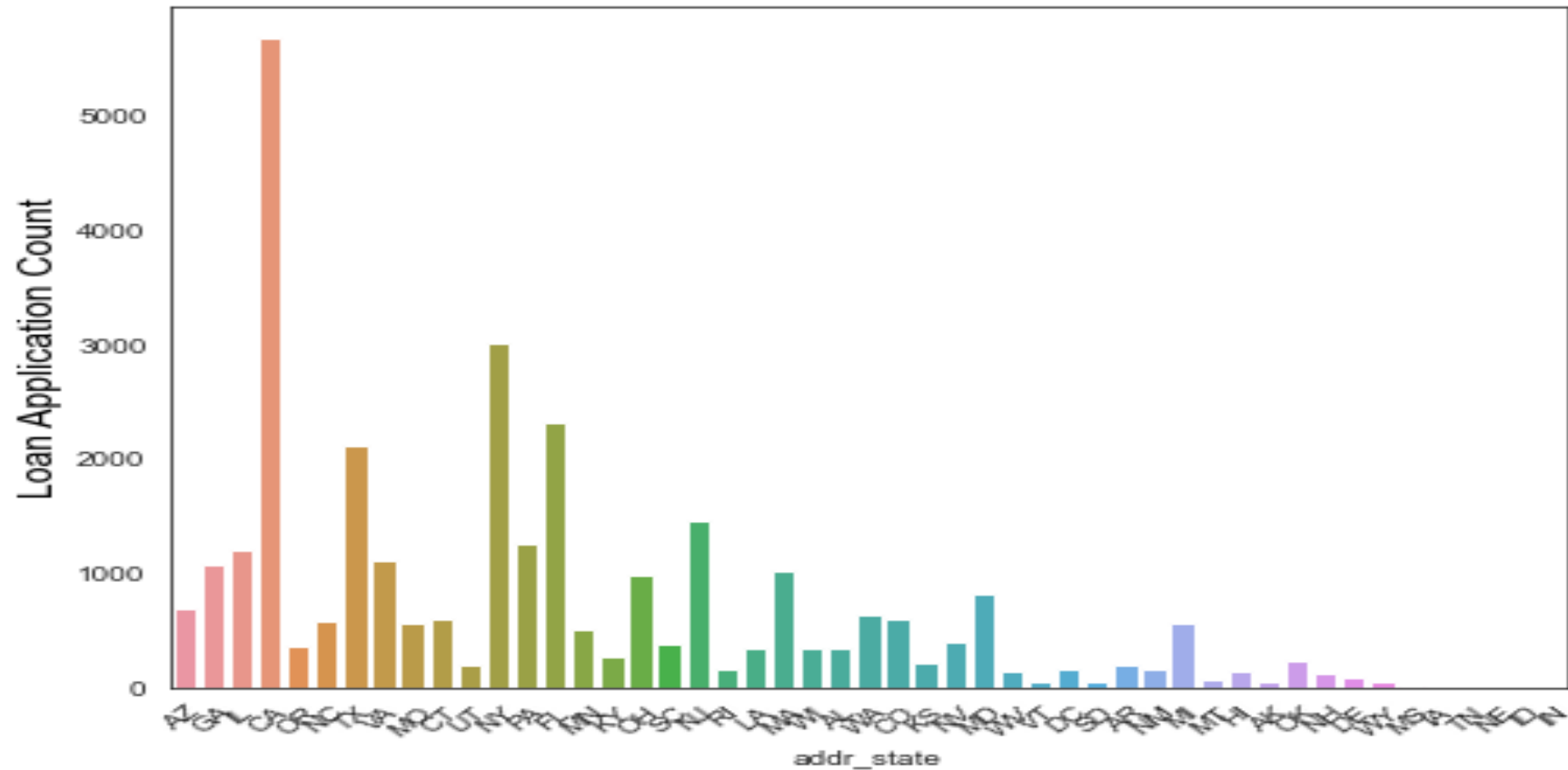
Analysis for numerical data

Analysis for numerical data

Analysis for numerical data

Analysis for numerical data

Observation

1. In recent years, the loan application have been increased

2. In the month of oct, nov and dec most of the loans are being issued.

3. This is due to decrease in the interest rate

4. There are more applicant who took 36 months duration loan as compared to 60 months.

5. Most of the applicants belongs to grade A and B

6. For subgrade same as most of the applicants belong to Grade A and B.
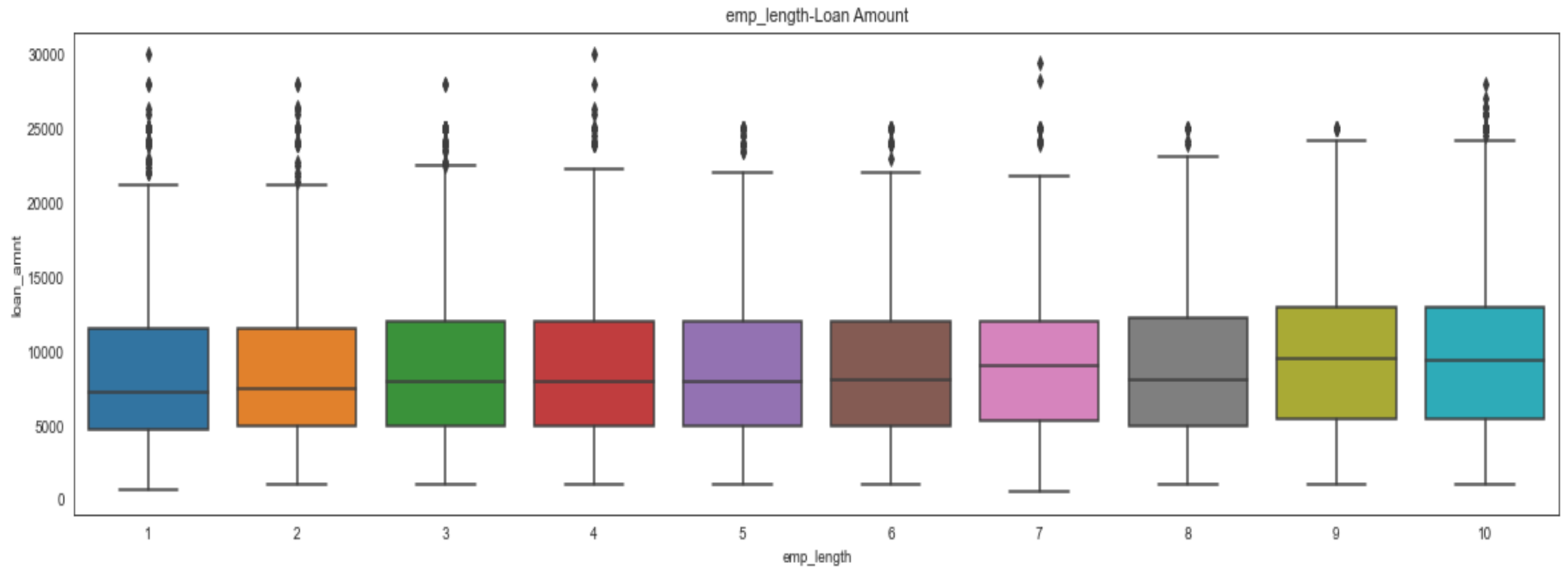
7. Most of the applicants are from NY and CA.
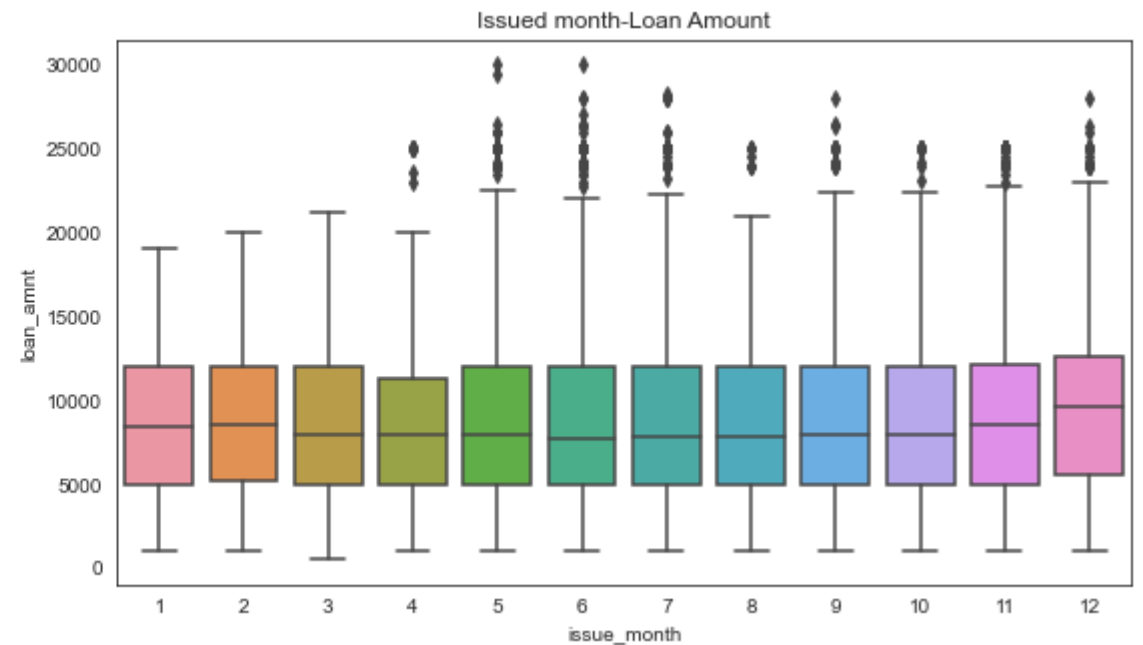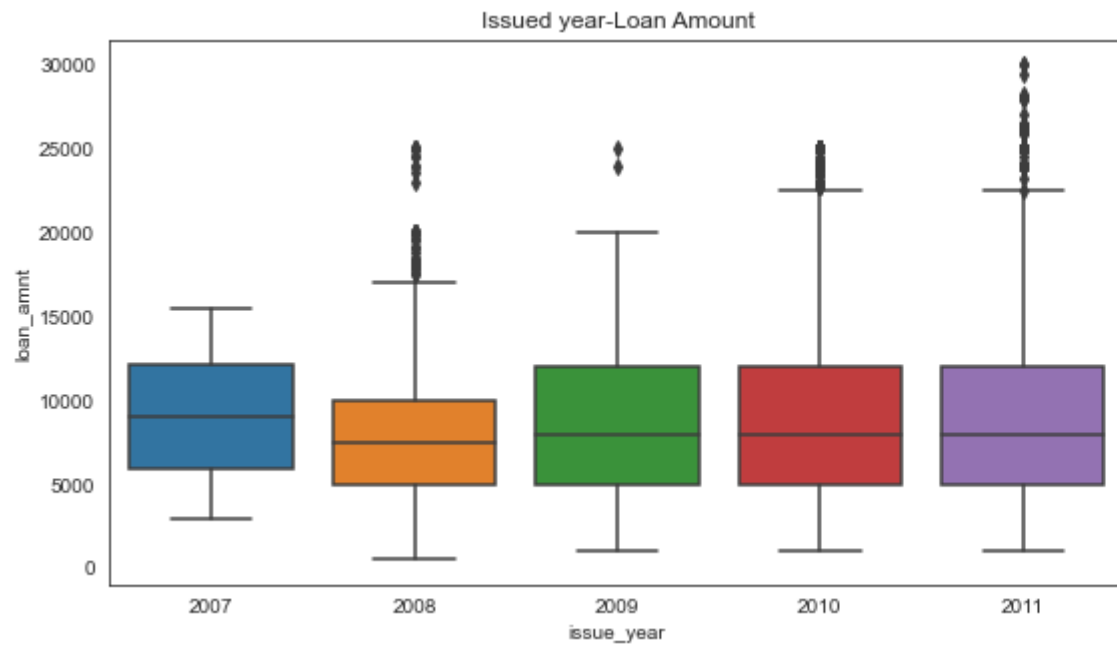
Segmented Analysis
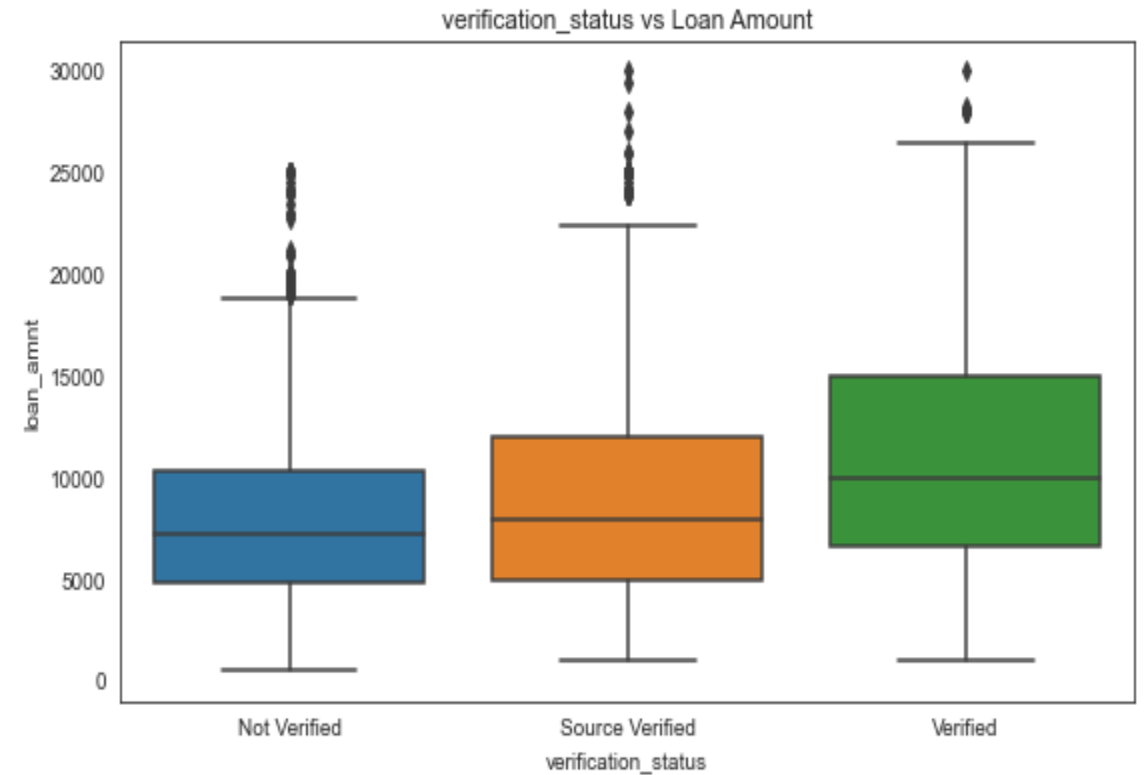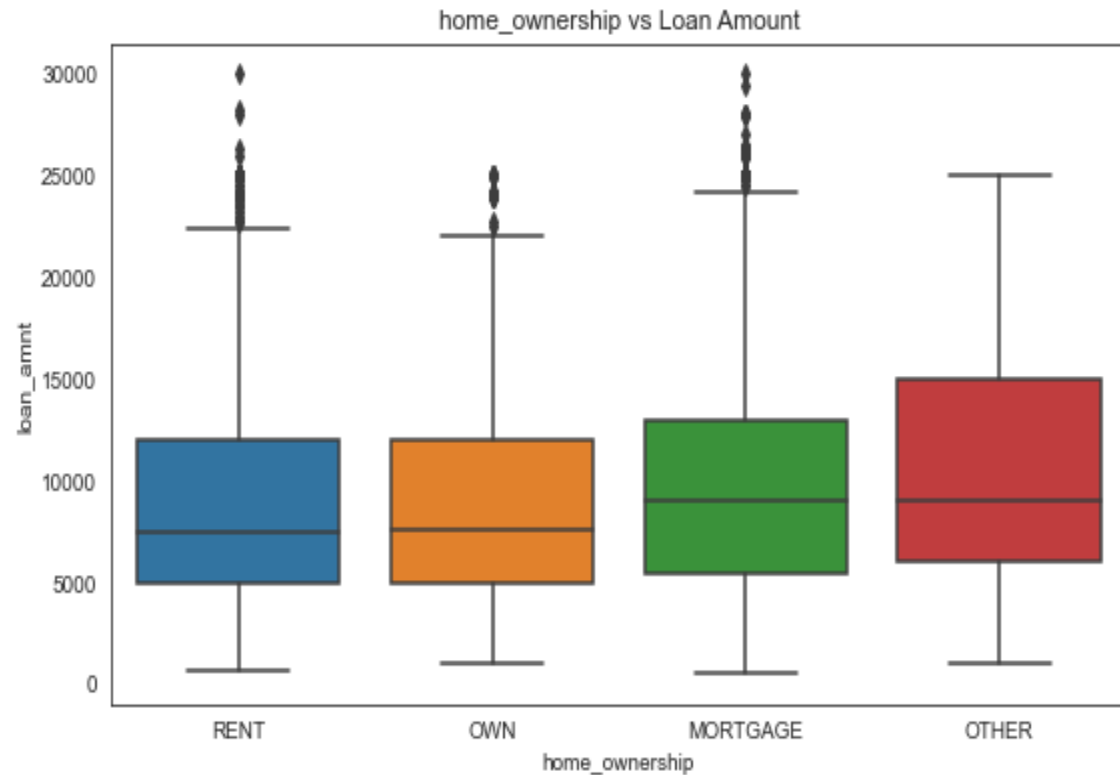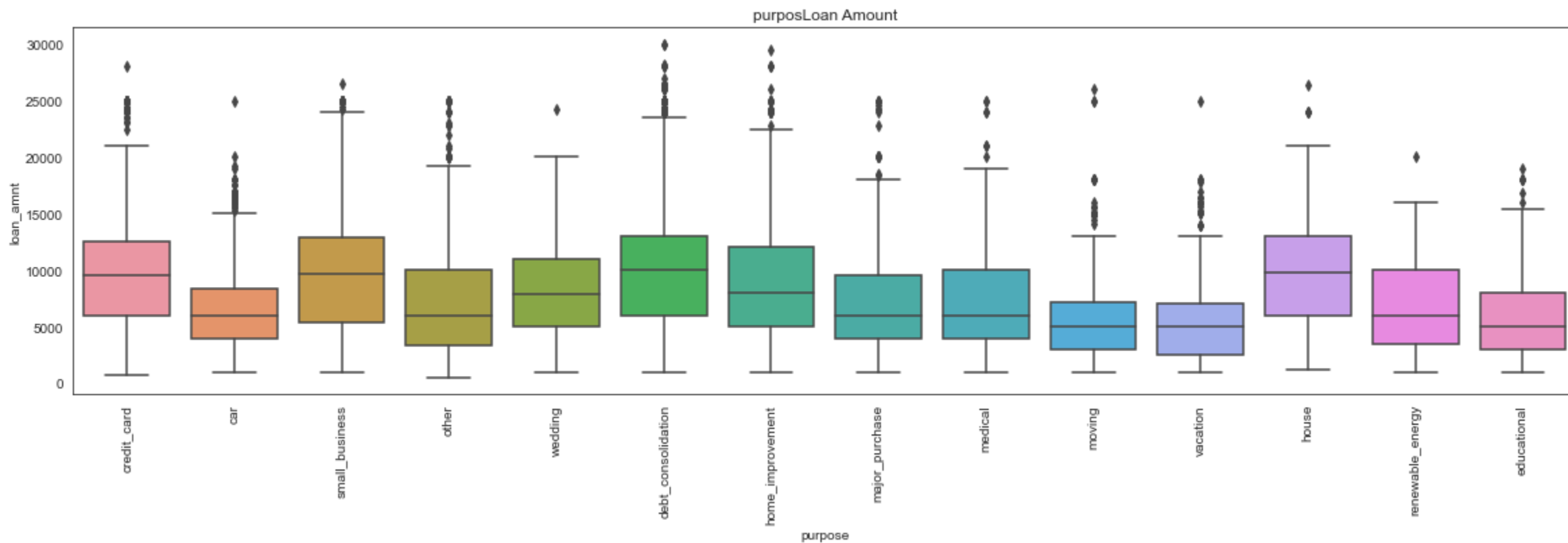
home_ownership vs Loan Amount

verification_status vs Loan Amount

Segmented Analysis
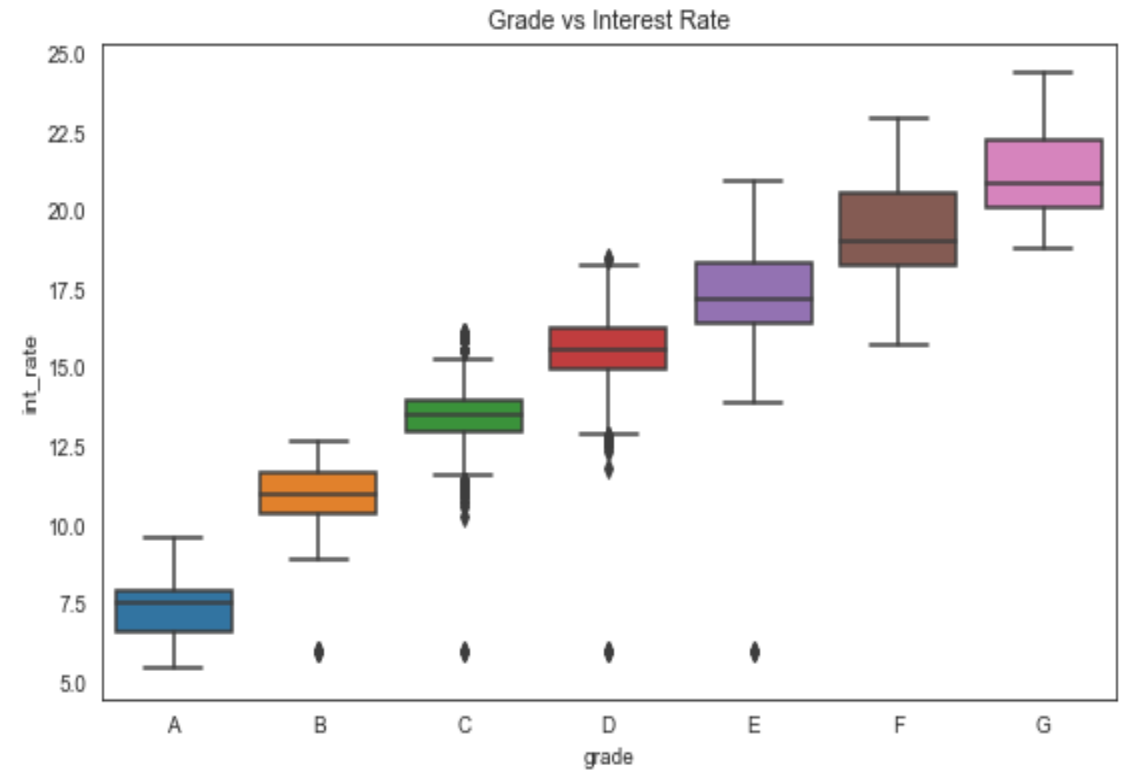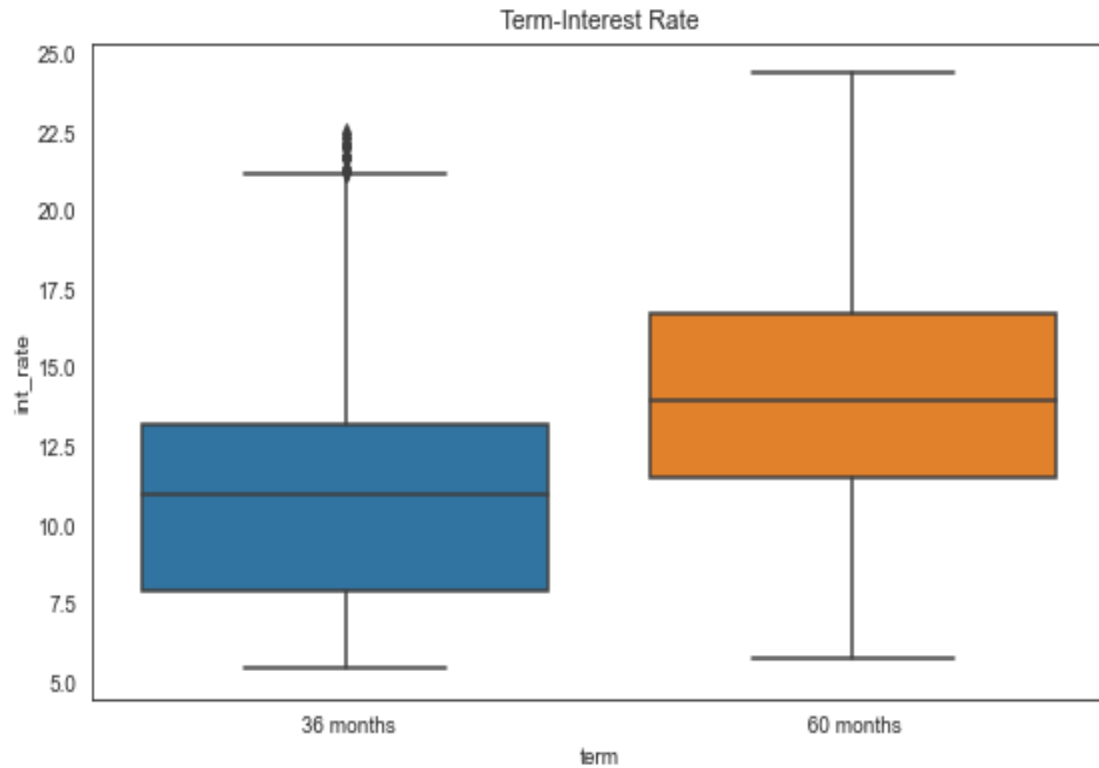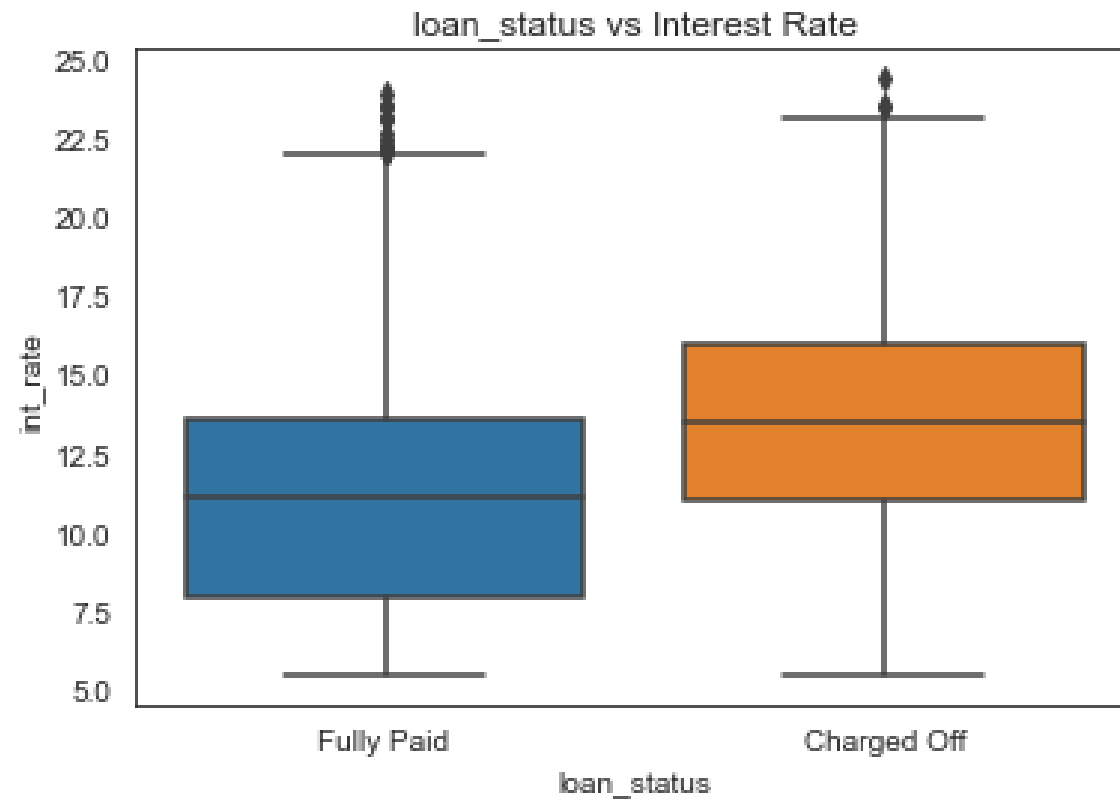
Segmented Analysis

emp_length-Loan Amount

Segmented Analysis

Segmented Analysis

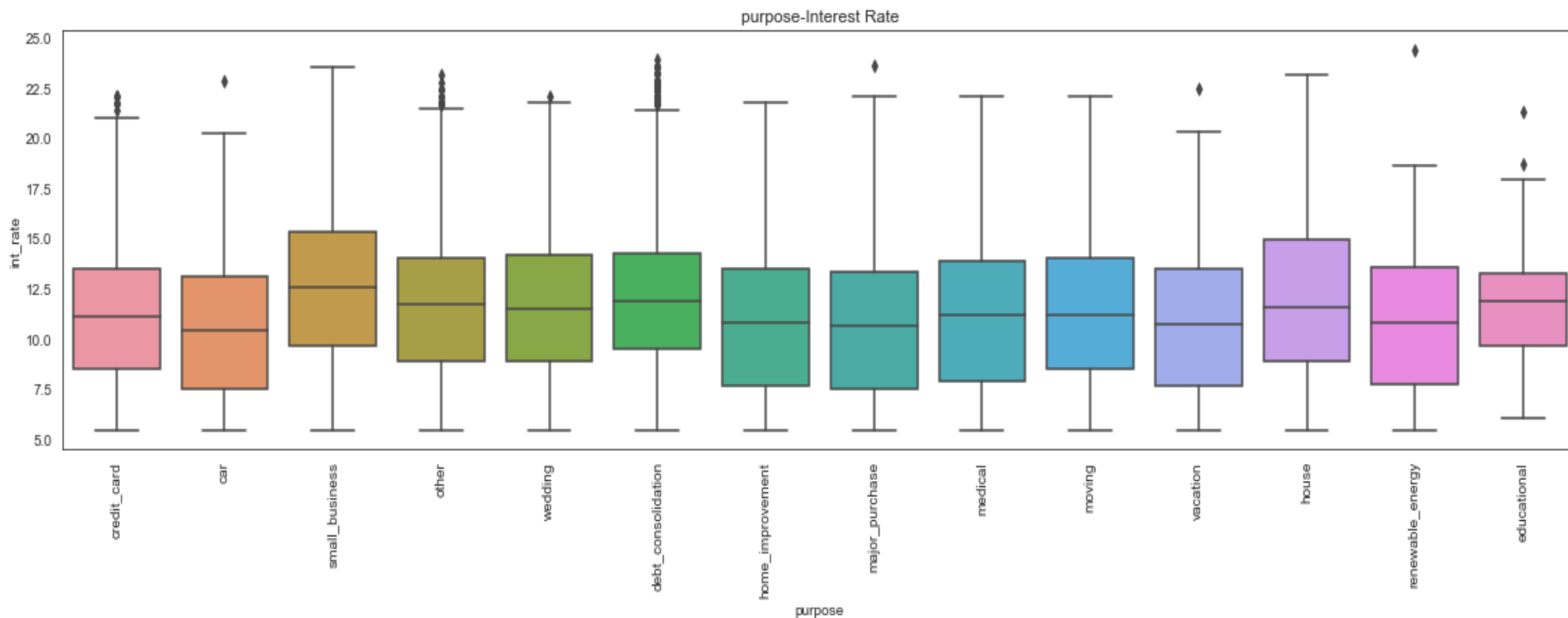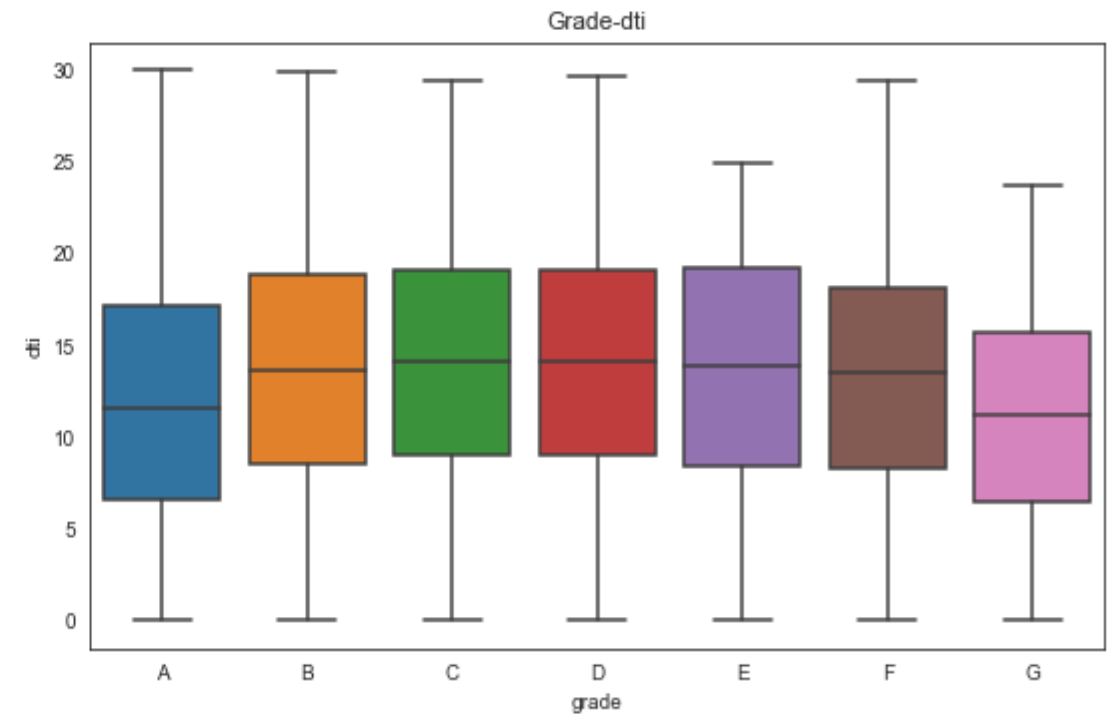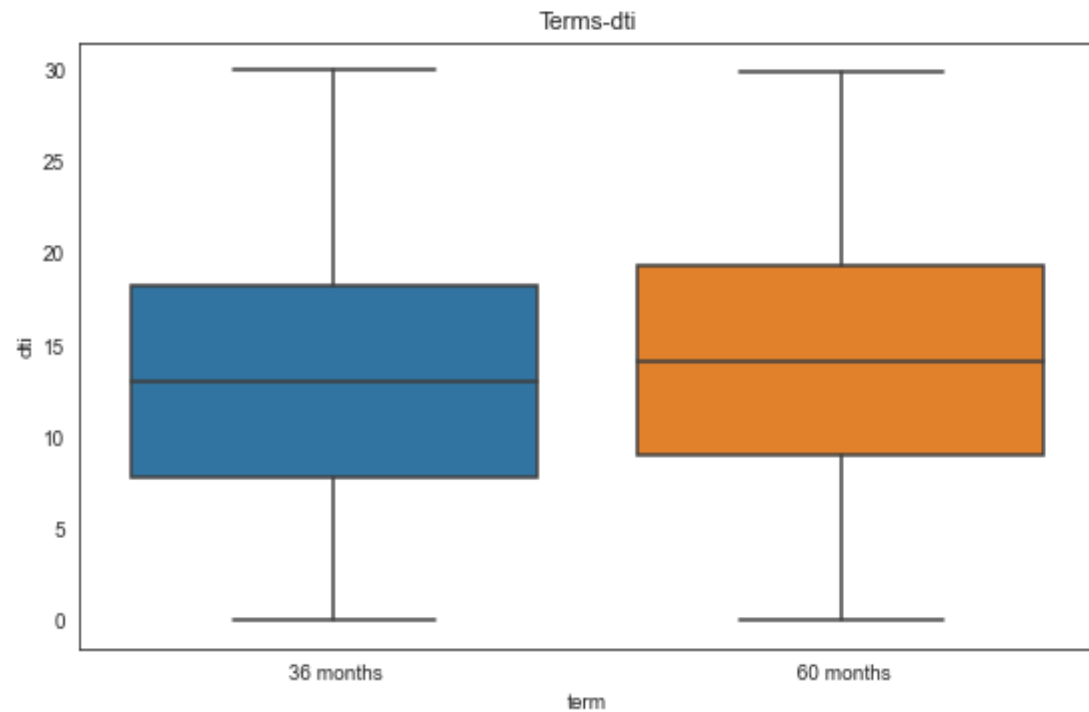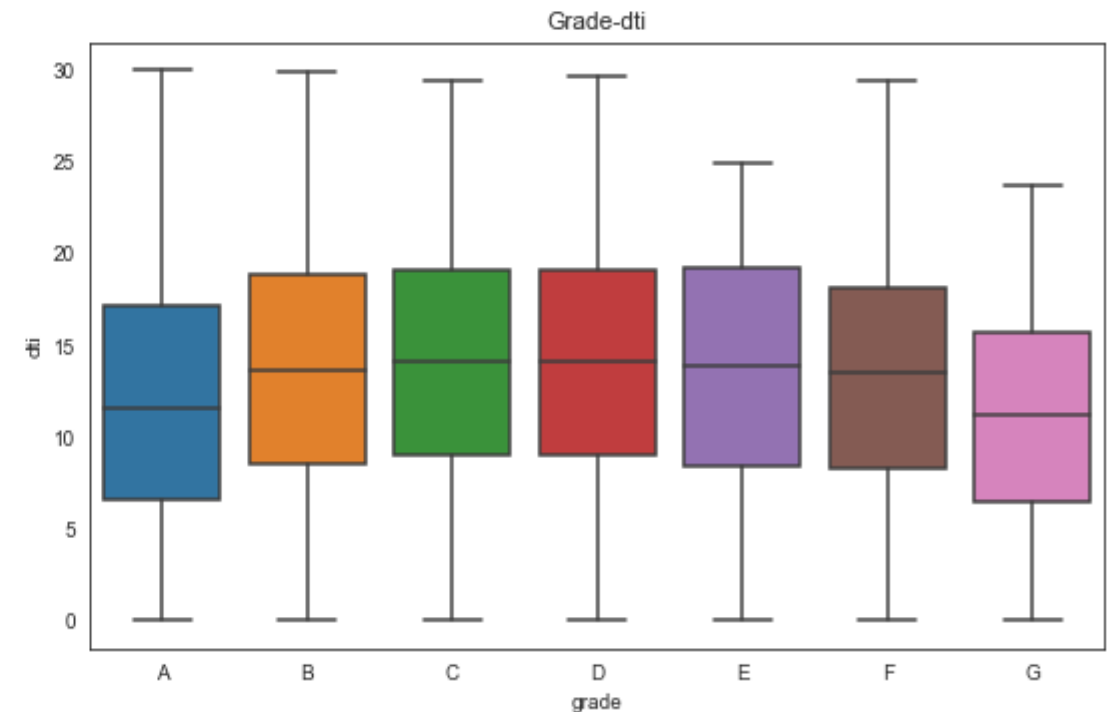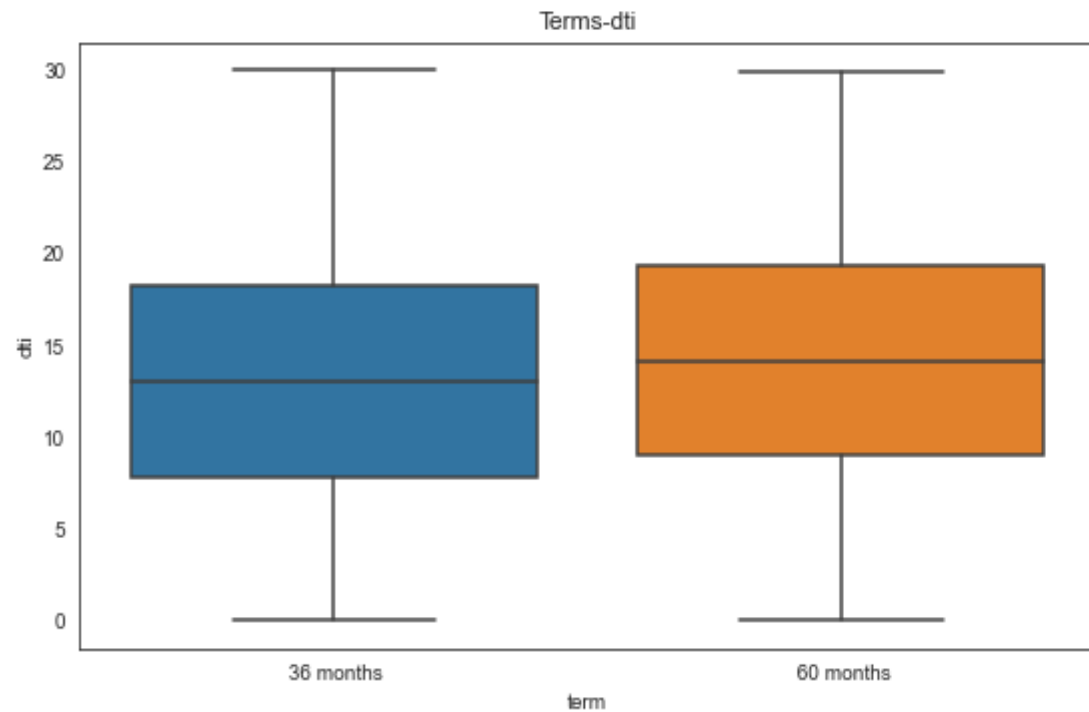# Segmented Analysis

Segmented Analysis



purpose-Interest Rate

Segmented Analysis

Segmented Analysis

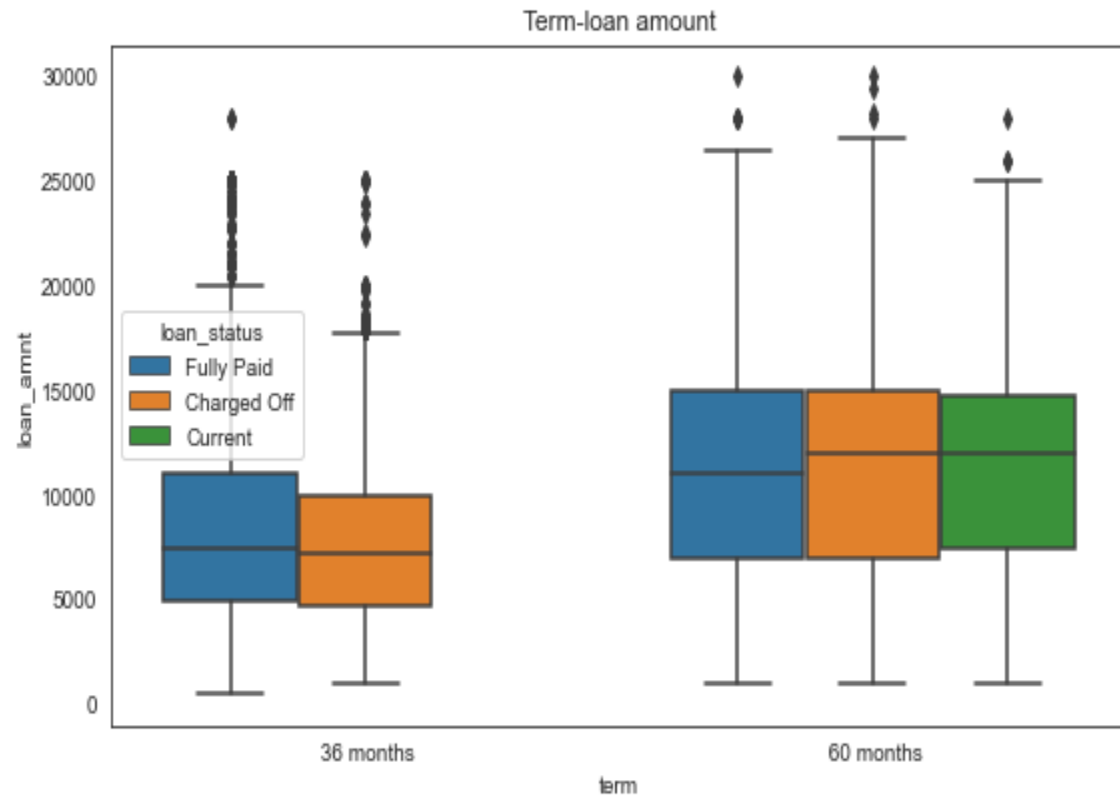# Observation

1. Most of the loan applicants are having MORTGAGE homeownership.

2. Most of the borrower have sourced.

3. Most of the loan amount is for Small business

4. More applicants who got loan approved are of 10yrs, so less age is the factor to get loan rejected.

5. More amount of loan is disbursed in year 2011 and in the last three months of the year.

6. DTI is high 60 months tenure

Bivariate Analysis
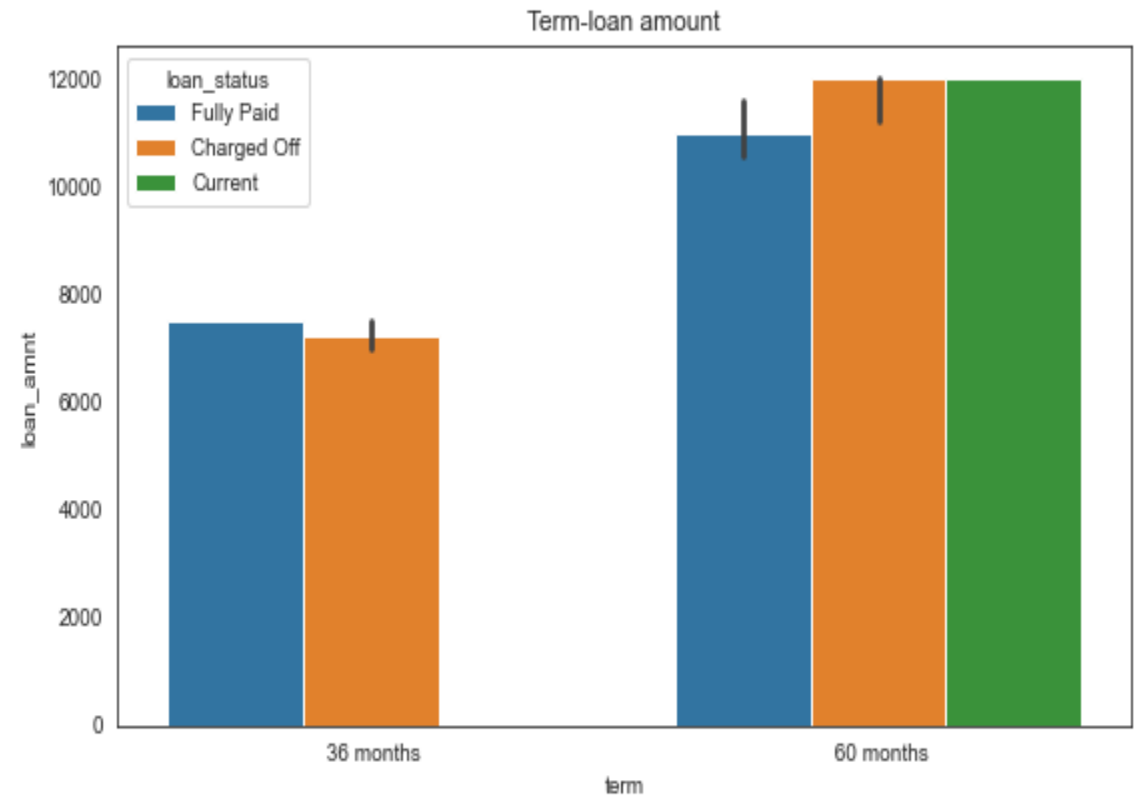
Bivariate Analysis

Verification Status vs interest rate

Bivariate Analysis

Purpose vs Loan status

Observation

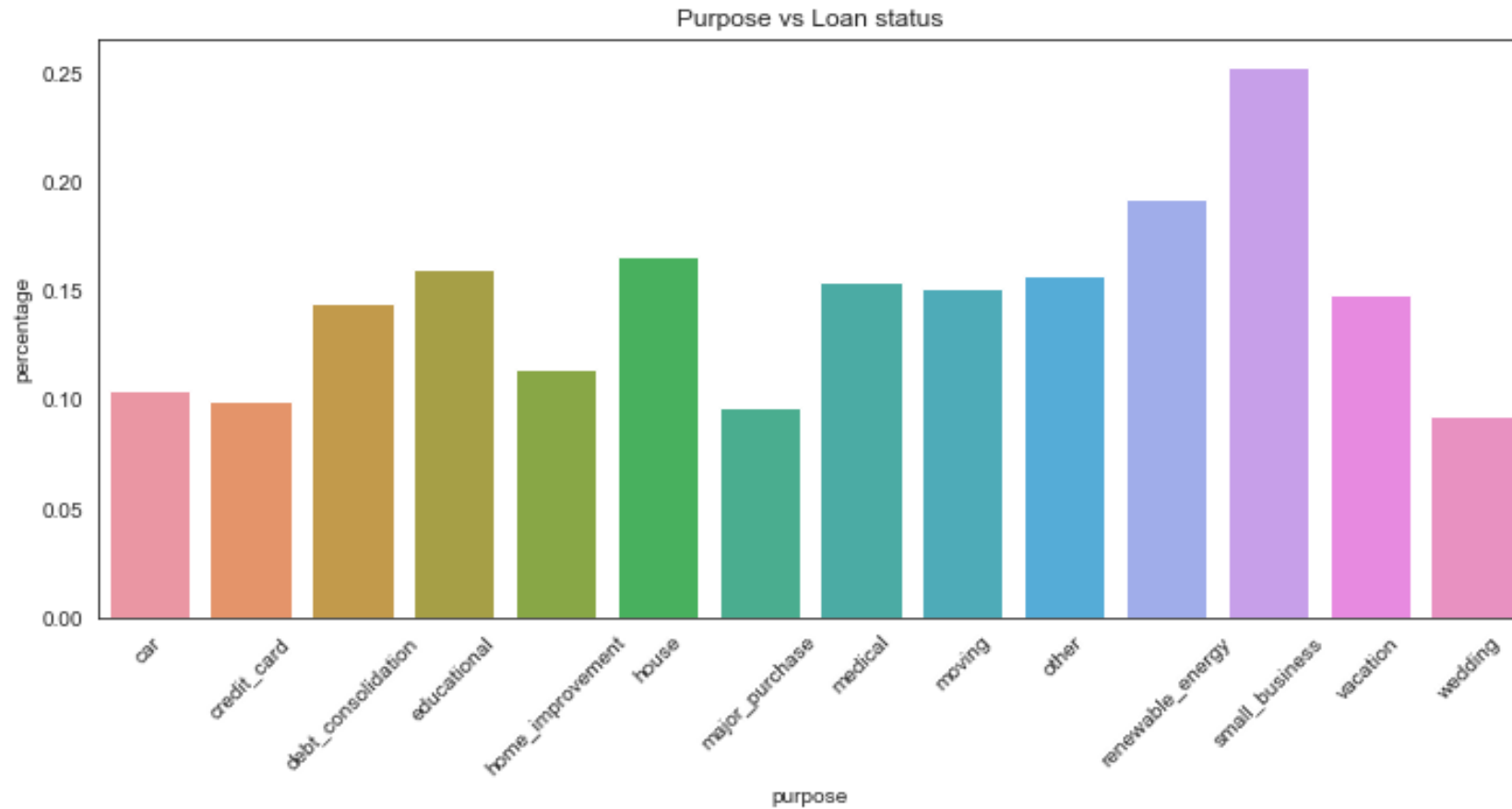1. More proportion of borrowers defaulted loan in 60 months.

2. Borrowers with higher employment lengths and took more loan amounts got more default rate.

3. The interest rate iS POSITIVELY correlated with default rate.

4. Charged-off are higher for small_business comparitively.

5. The public derogatory records correlated with public bankruptcies records.

6. Interest rates are high for people with high revol utilisation

Your best quote that reflects your approach… "It's one small step for man, one giant leap for mankind."

- NEIL ARMSTRONG