LENDING CLUB CASE STUDY

PROBLEM STATEMENT

You work for a **consumer finance company** which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two **types of risks** are associated with the bank's decision:

• If the applicant is **likely to repay the loan**, then not approving the loan results in a **loss of business** to the company

• If the applicant is **not likely to repay the loan,** i.e. he/she is likely to default, then approving the loan may lead to a **financial loss** for the company

OBJECTIVE

Using Exploratory Data Analysis, find out the driving factors behind loan default.
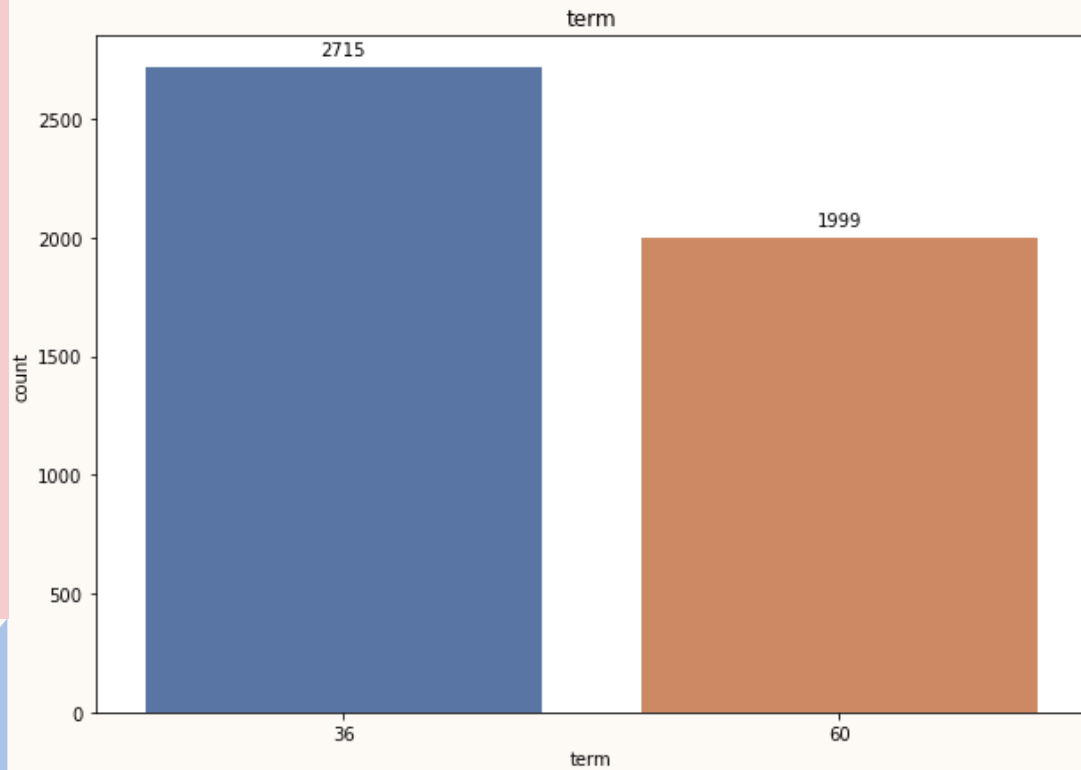
DATASET INFORMATION

- This dataset contains the complete loan data for all loans issued through the time period 2007 to 2011.

- We have conducted univariate and bivariate analysis specifically focusing on loans where the loan status is Charged Off.

- For this data set we have already performed the below steps:

  1. Data Cleaning

  2. Correcting Data Types

  3. Handling Null values

  4. Managing outliers

  5. Deriving new columns

  6. Univariate Analysis for Categorical (Ordered and Unordered) Variables and Quantitative Variables

  7. Bivariate Analysis

UNIVARIATE ANALYSIS

Ordered Categorical Variables

- Term
- Emp Length
- Grade
- Sub Grade
- Issue Year
- Issue Month

Employee Length

Applicants with more than 10 years of experience indicate that better experience doesn't highly correlate with loan repayment.
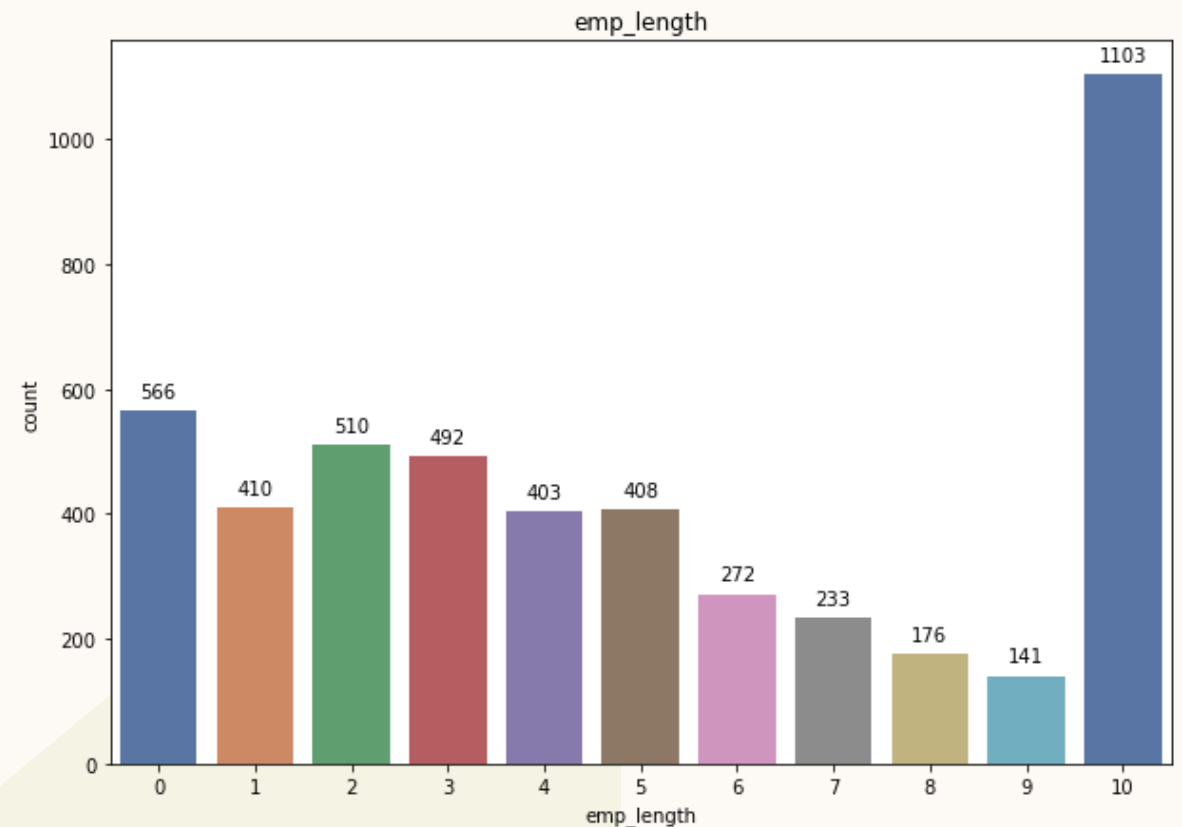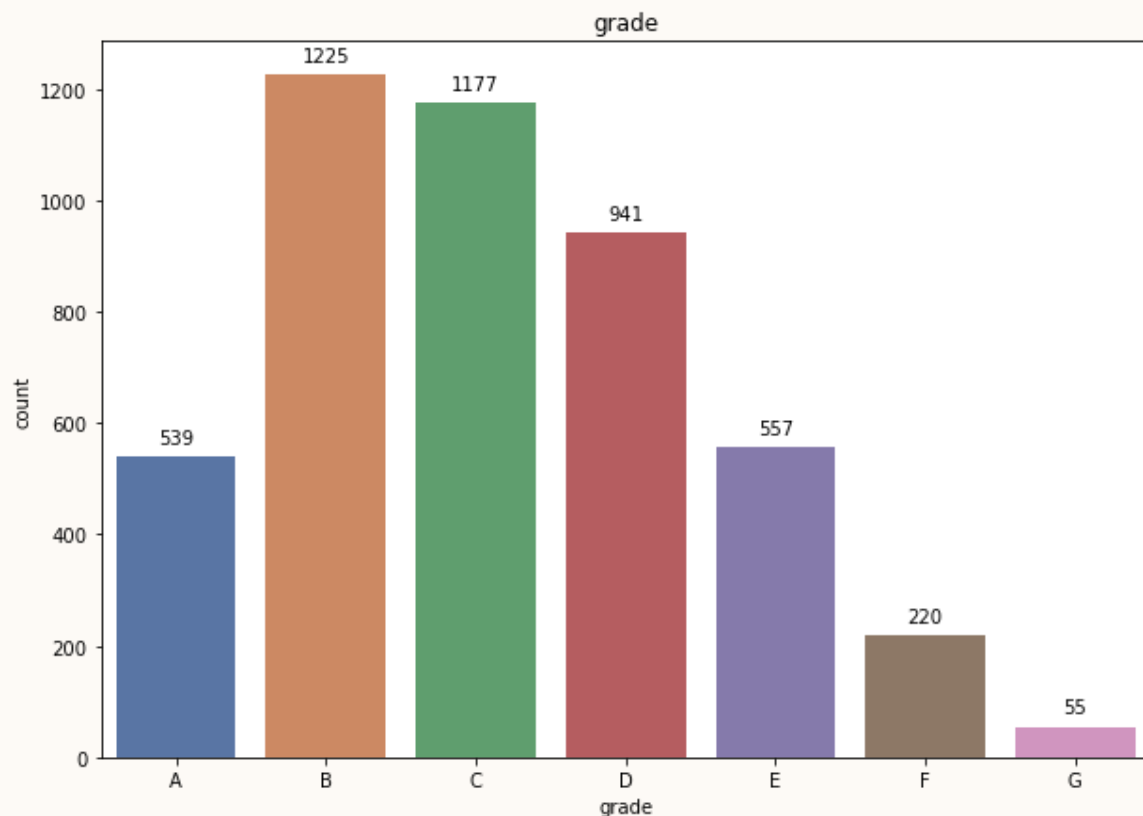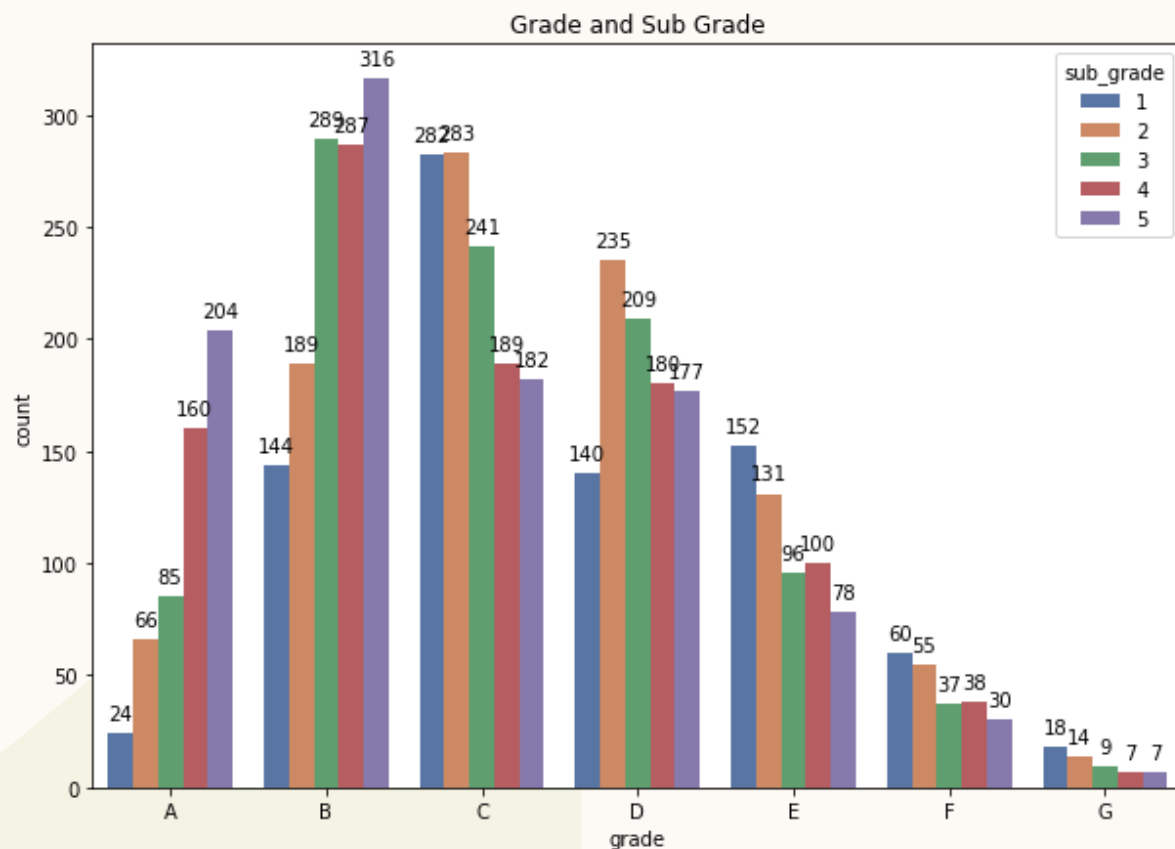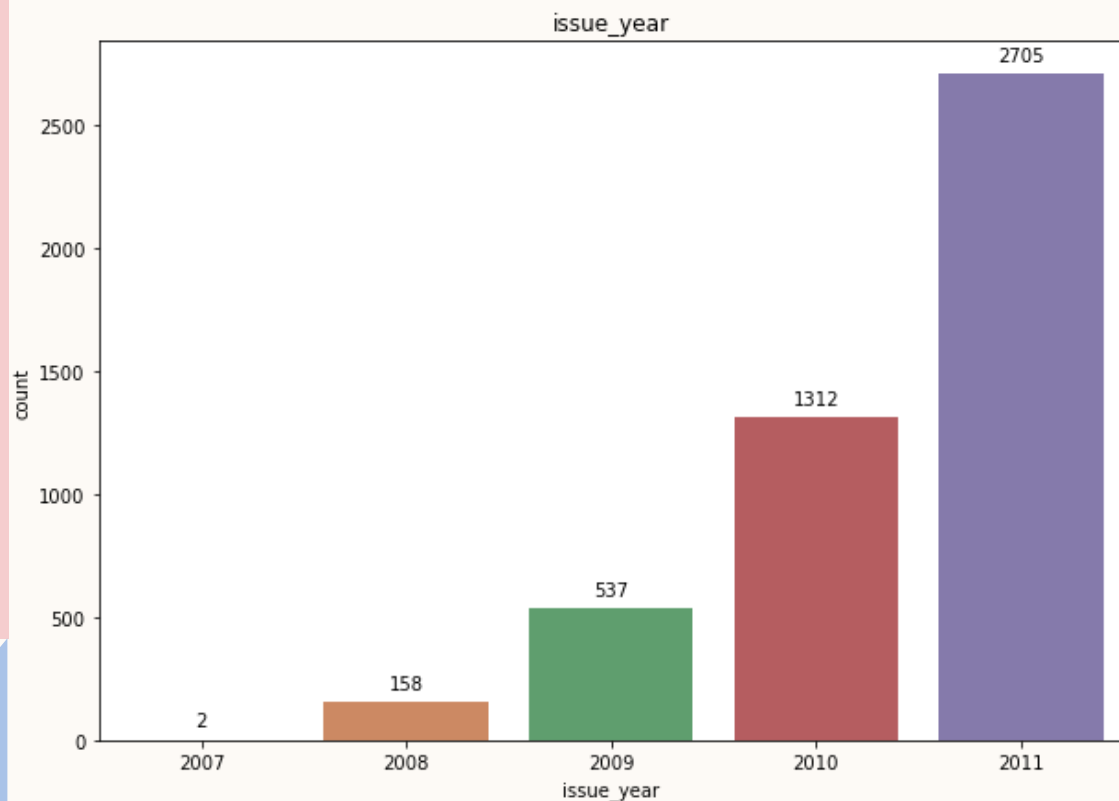
Term

More than 2500 applicants choose the term as **36 months** for the repayment of loans, indicating this is the most popular term among the borrowers.

Sub Grade

Within Grade B, loan applicants with sub grade B5 had the highest number of charged-off loan applicants.

Grade

1225 Applicants with credit score B have the highest number of charged-off loan applicants indicating financial issues.

Issue Month

December has the highest number of loan applicants that were charged off indicating that during the holiday season, many of them took loans that were charged off.
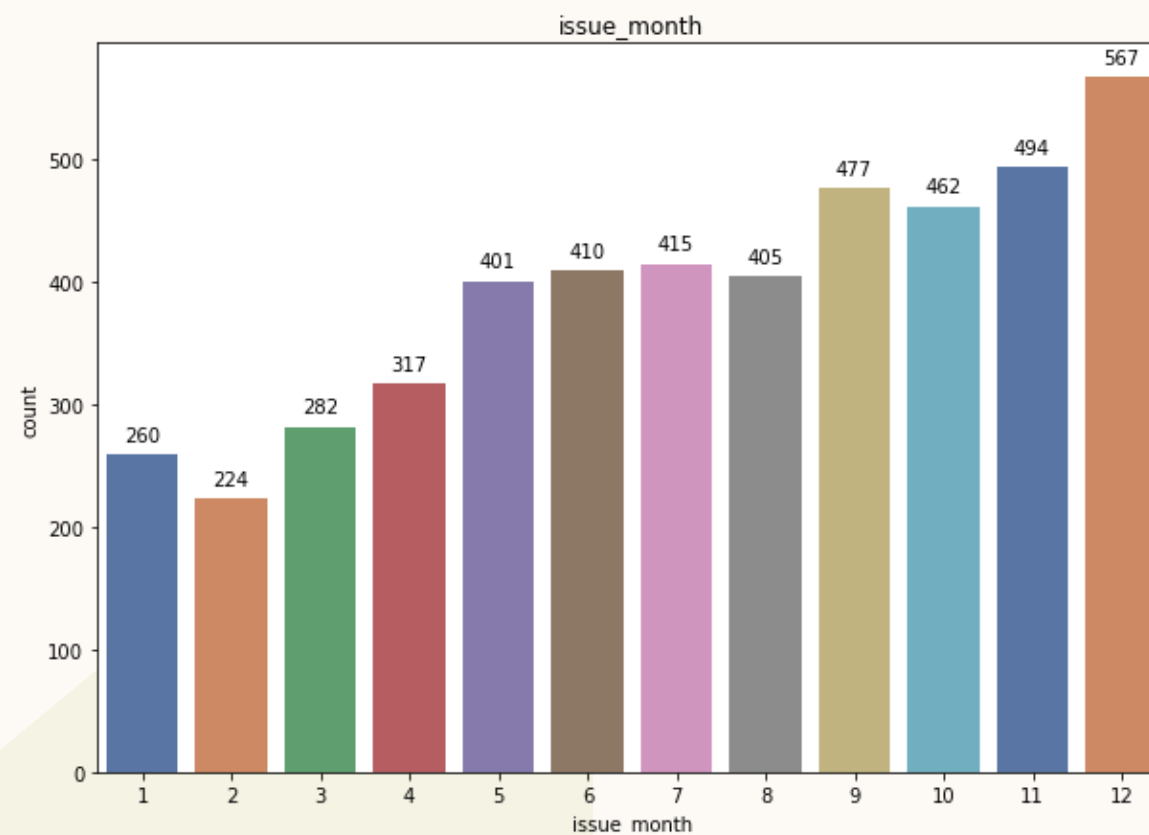
Issue Year

In the year 2011, the highest number of loan applicants were charged off indicating many of them were suffering from financial issues.
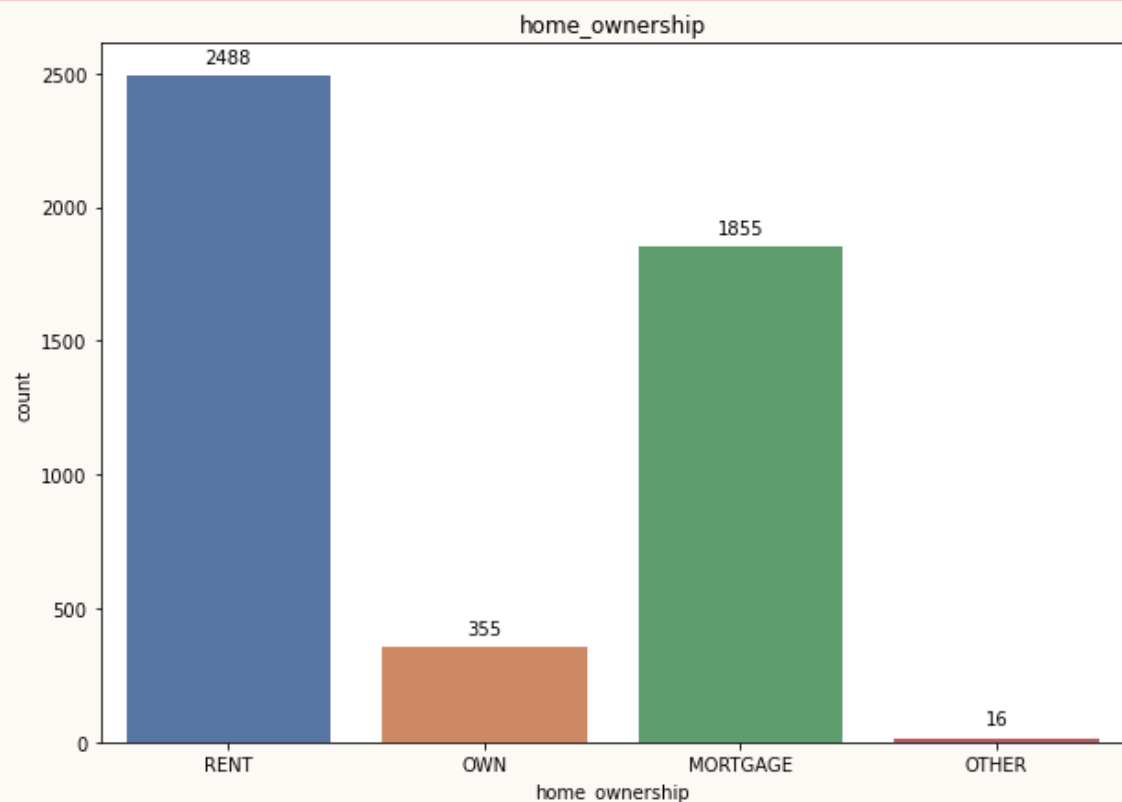
UNIVARIATE ANALYSIS

Unordered Categorical Variables

- Loan Status
- Home Ownership
- Verification Status
- Purpose
- Address State

Loan Status

Loan Applicants that have defaulted are very less as compared to Fully paid applicants.
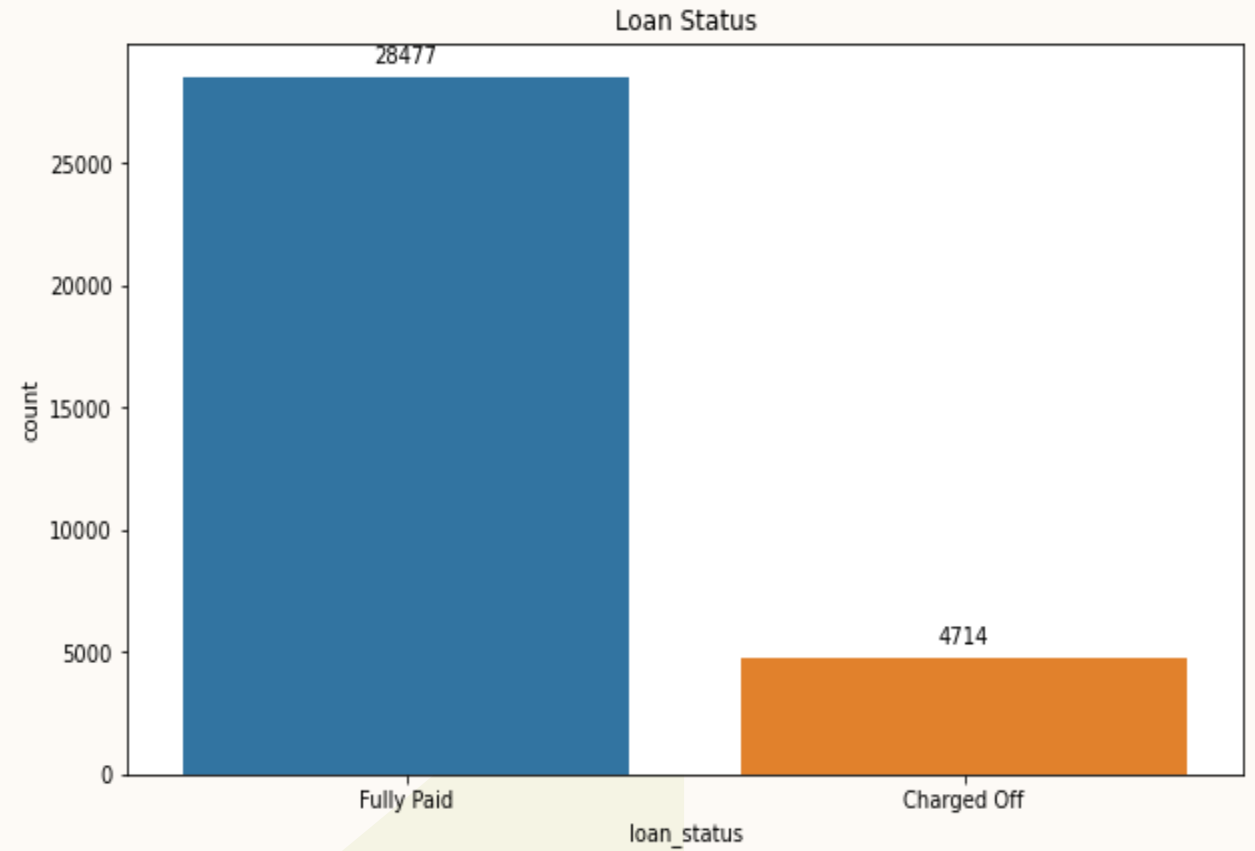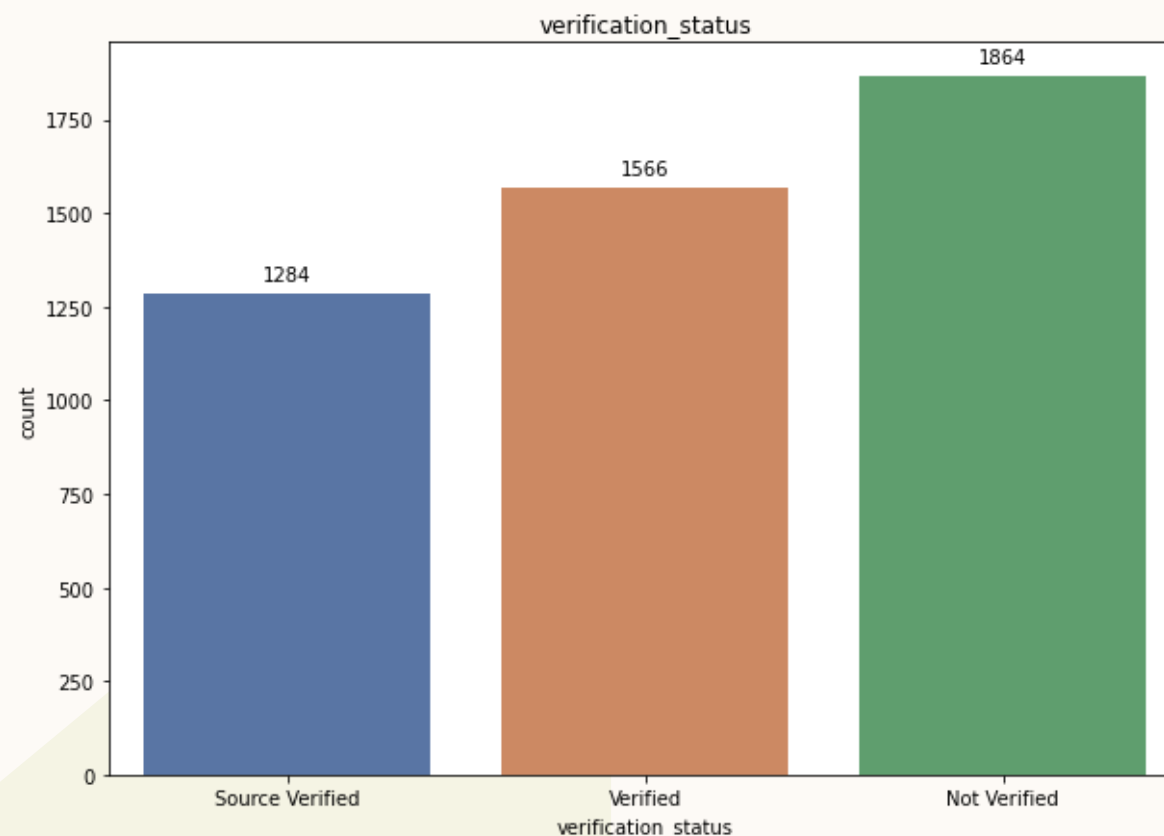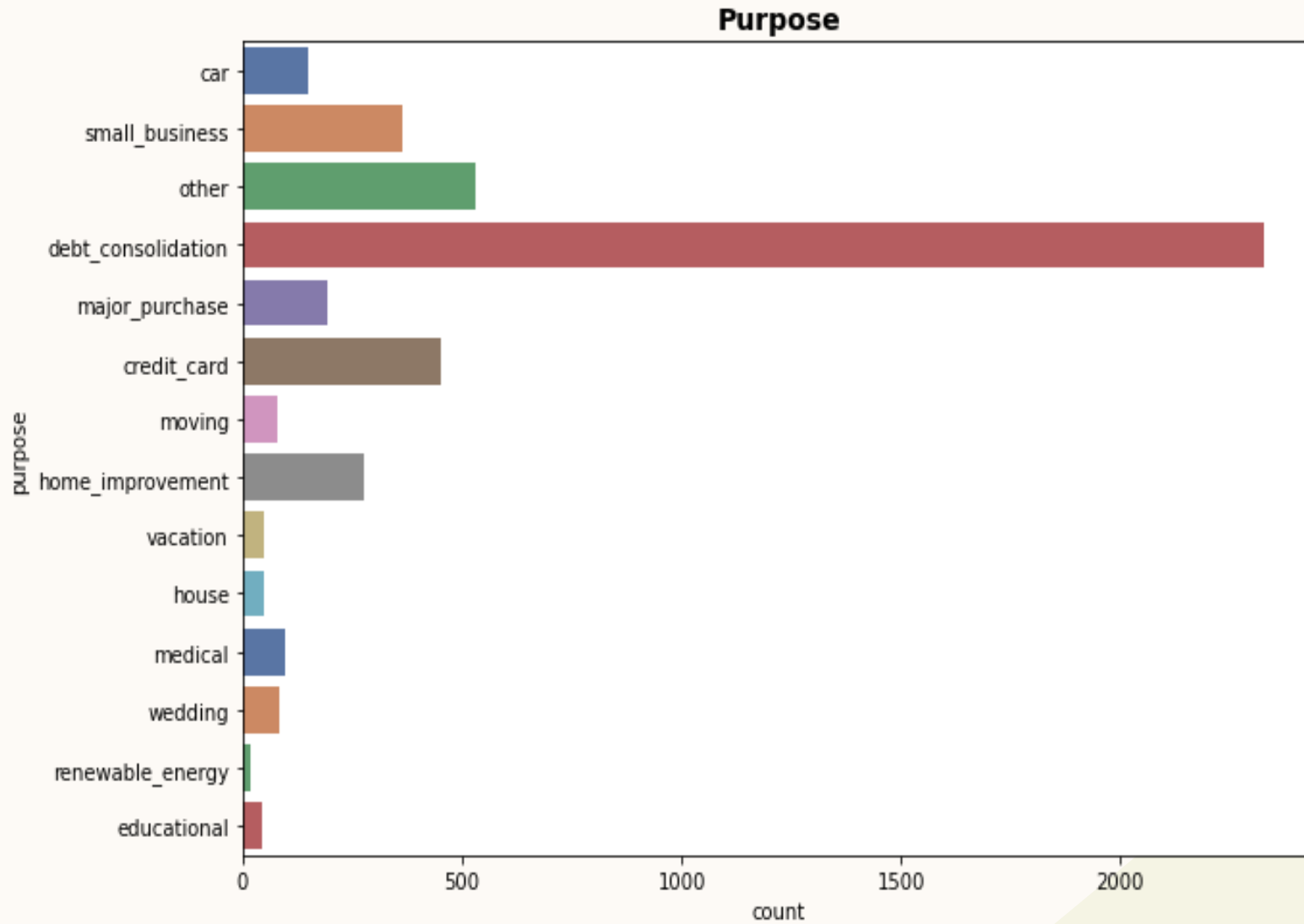
home_ownership

Home Ownership

Loan Applicants who don't own their home and are staying on rent tend to have a higher possibility of not repaying the charged off loan.

Verification Status

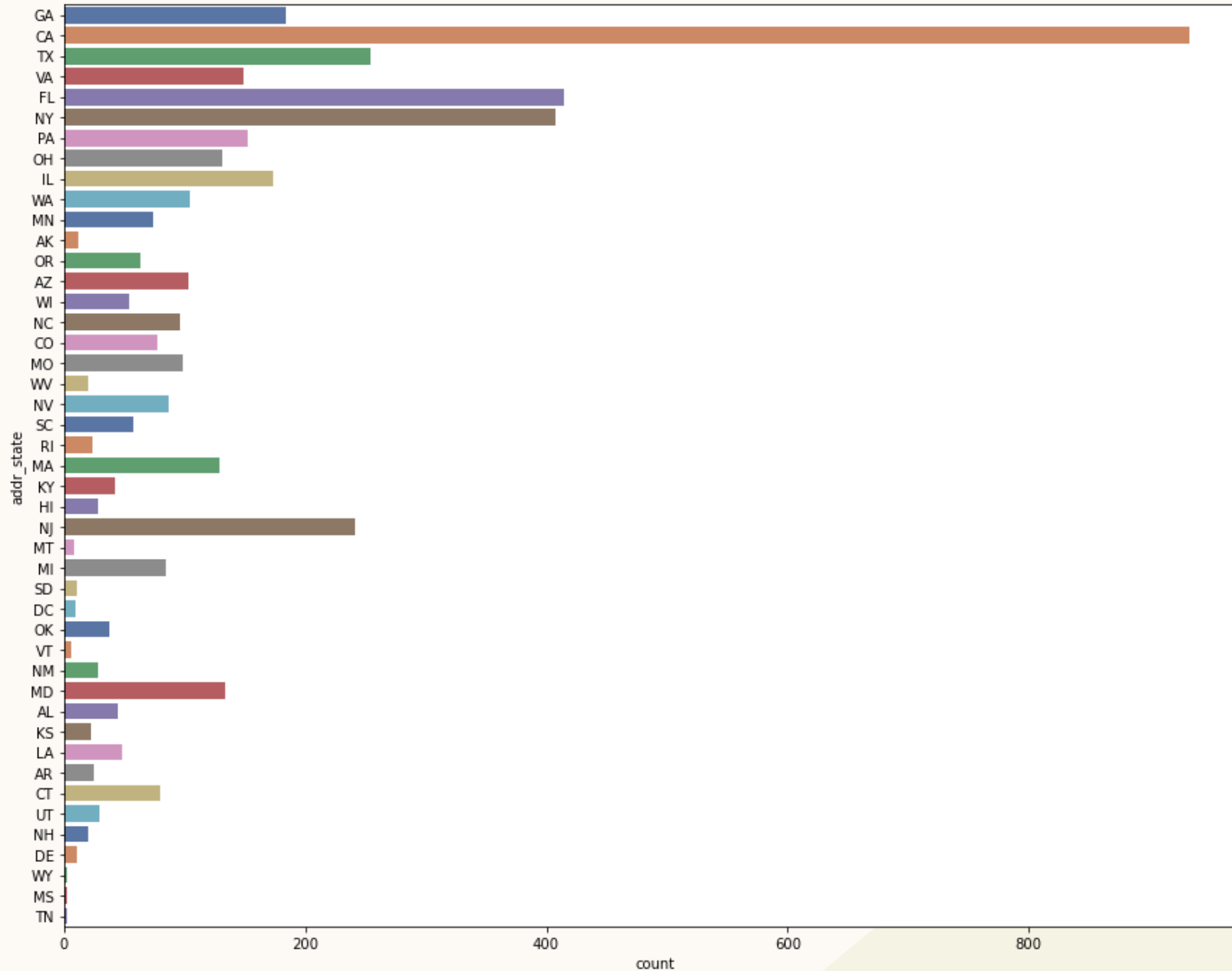More than 1700 loan applicants who were not verified took loans that were charged off.

verification_status

Purpose

Purpose

Loan Applicants who applied for loans to consolidate their debt have the highest rate of charged off loans.
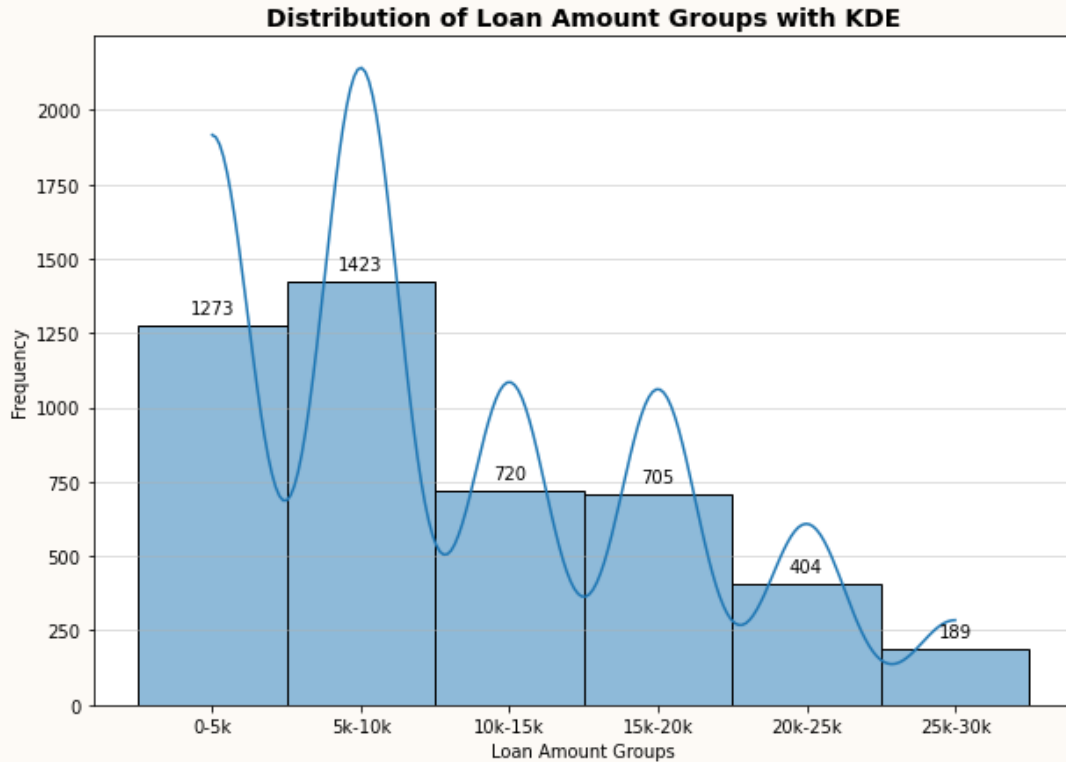
Address States

Address State

Address state 'CA' or California has the highest number of loan applicants that were charged off.

UNIVARIATE ANALYSIS

Numerical Or Quantitative Variables

- Loan Amount
- Annual Income
- DTI
- Interest Rate
- Installment
- Public Record Bankruptcies

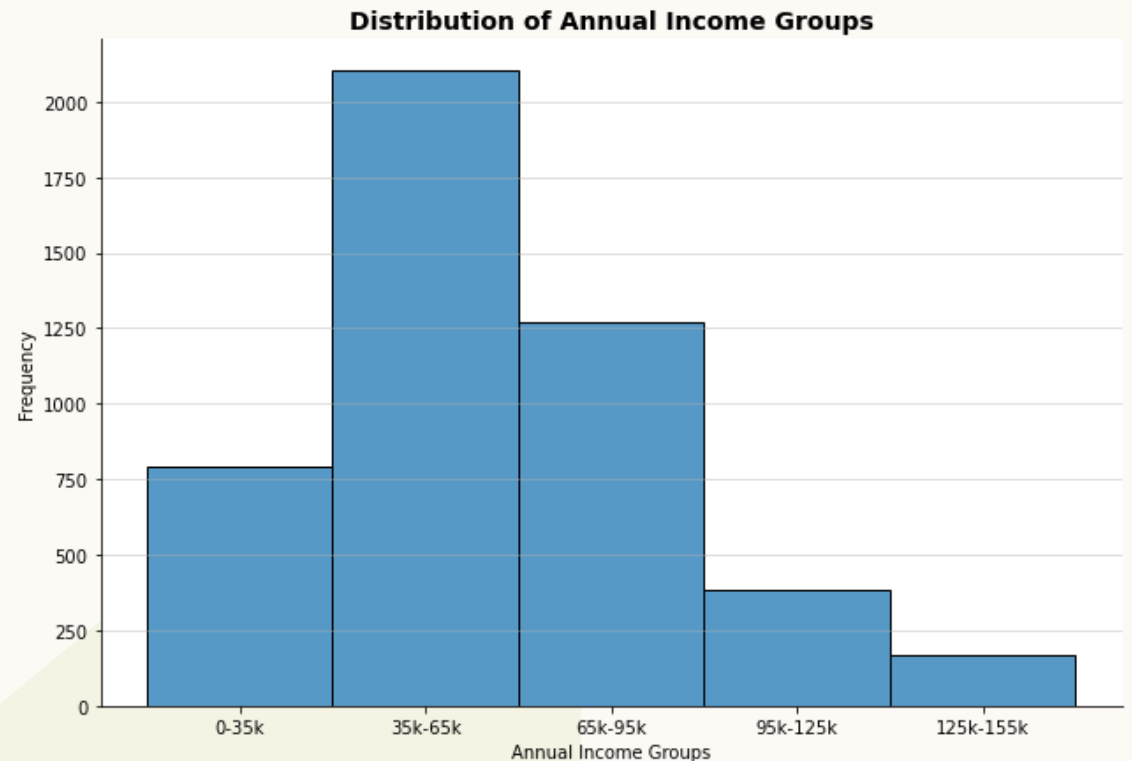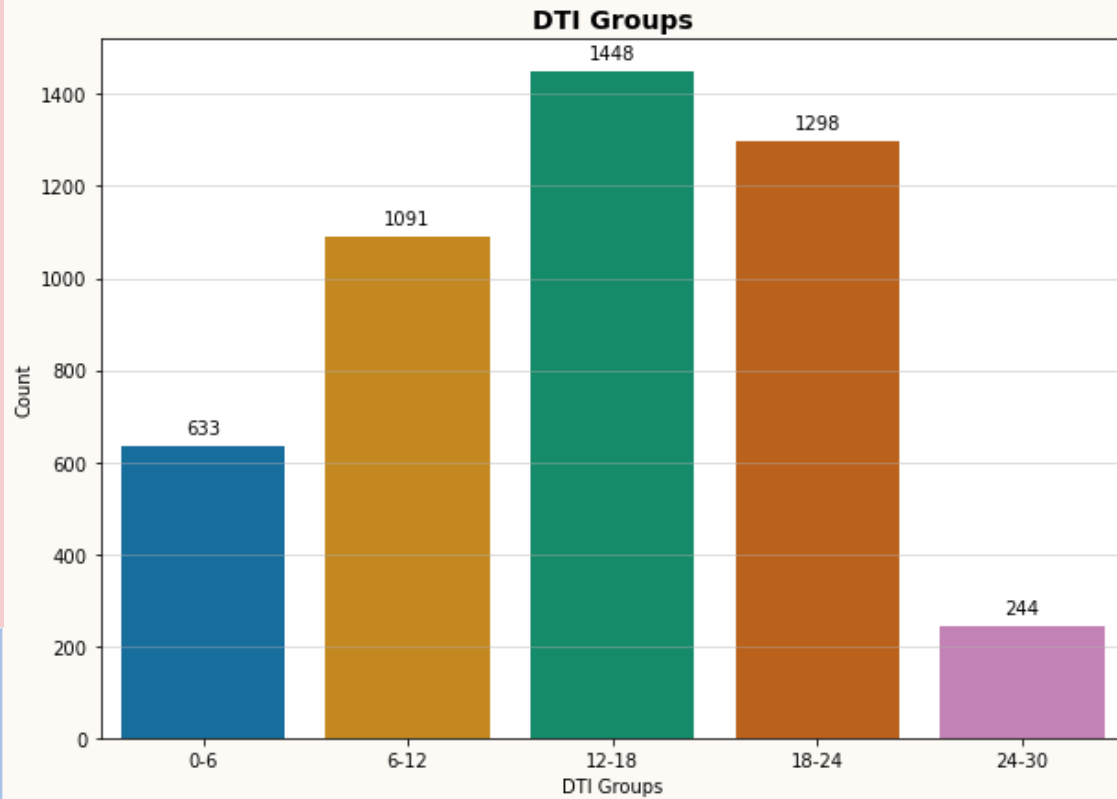**Distribution of Loan Amount Groups with KDE**

Loan Amount

Applicants who have loan amounts ranging from 5k to 10k USD do not tend to repay the loan that was charged off indicating potential risk among this group

Annual Income

Applicants who have an annual income ranging from 35k to 65k USD display a higher tendency to default on their loans.
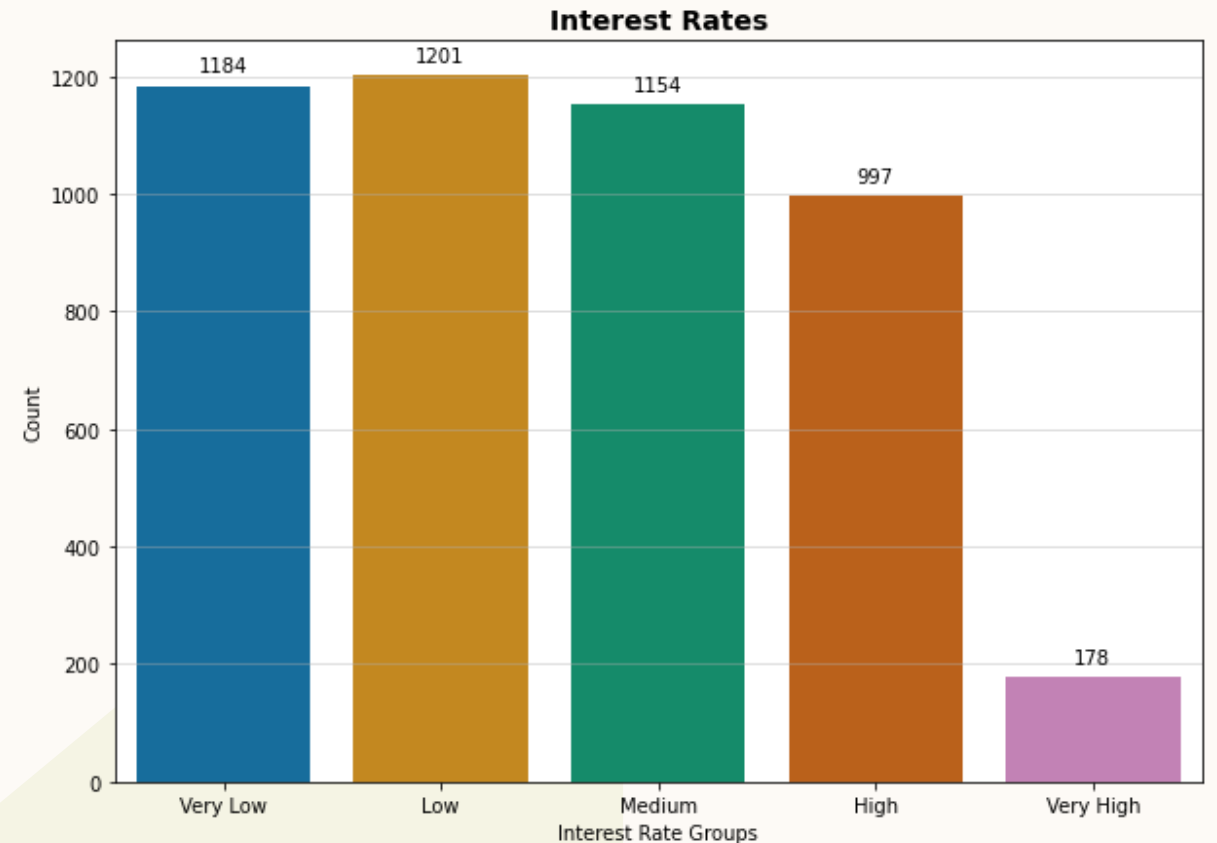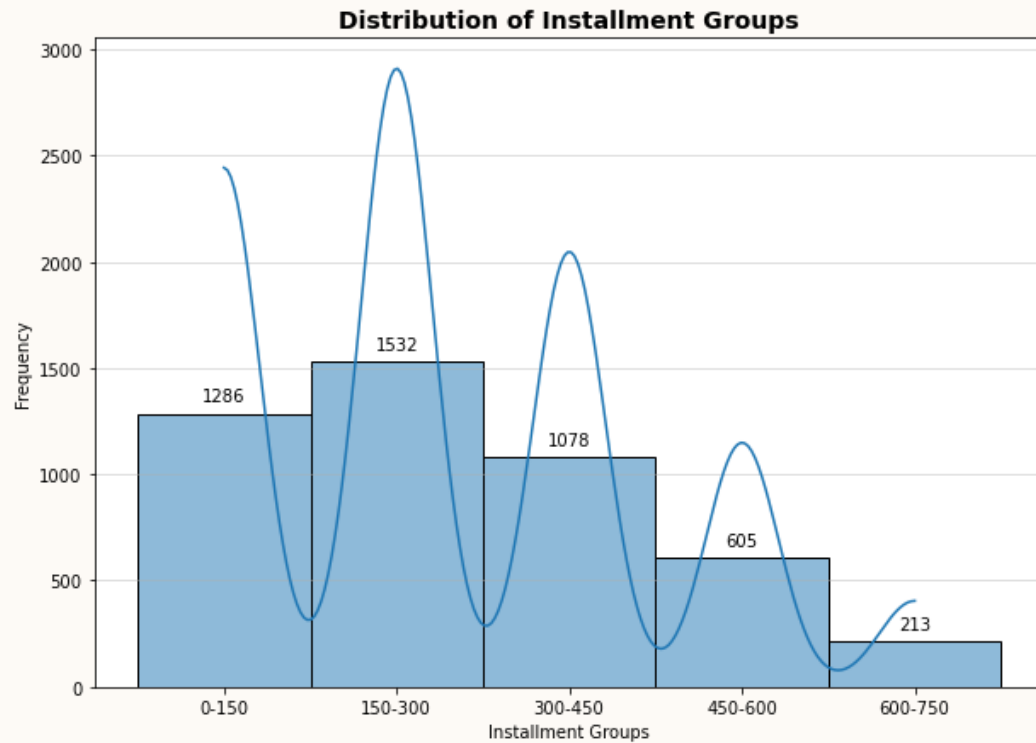
Distribution of Annual Income Groups

Interest Rate

Applicants with very low, low and medium interest rates have a higher tendency in defaulting their loans.

DTI

Applicants with DTI groups ranging from 12 to 18 tend to default on their loans as compared to other groups.

Distribution of Installment Groups

Public Record Bankruptcies

More than 4000 applicants have no record of public bankruptcies showing a lower tendency to default and less than 500 applicants show a tendency to default on their loans.
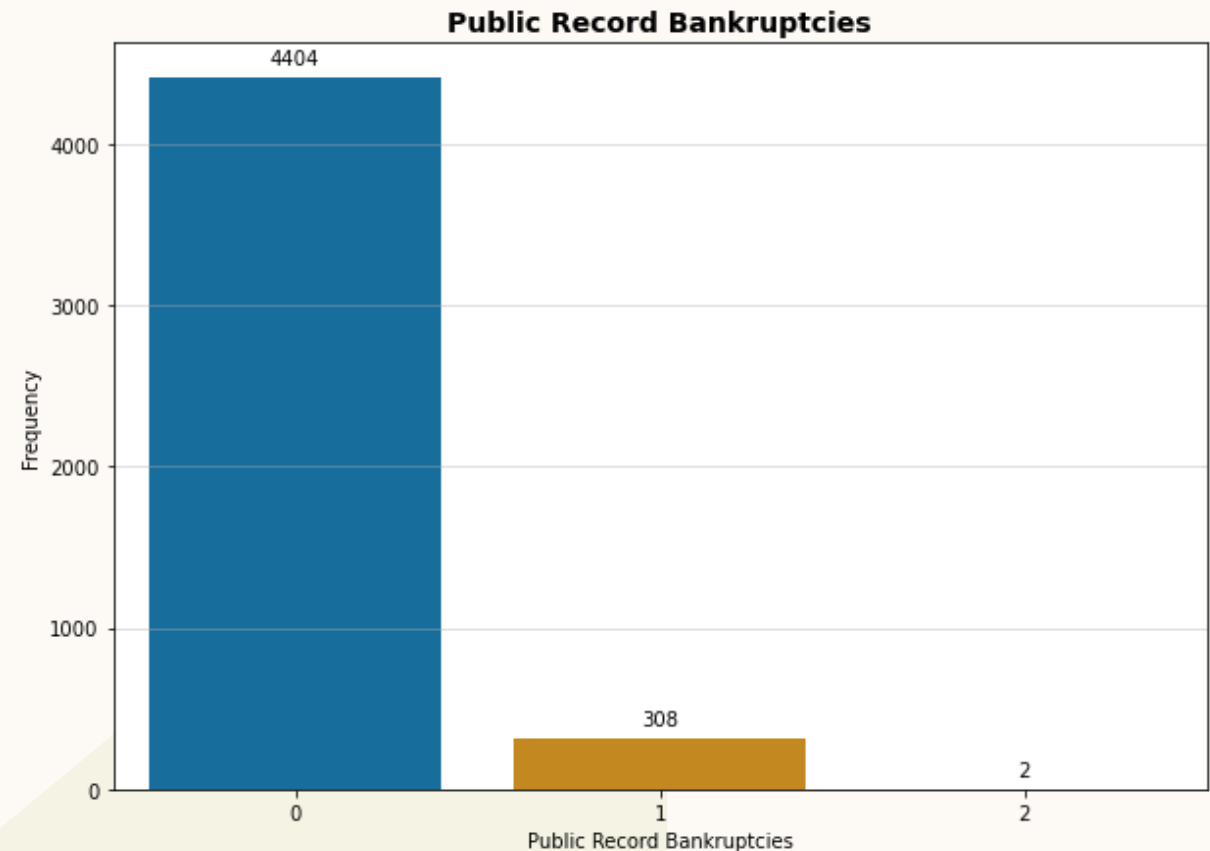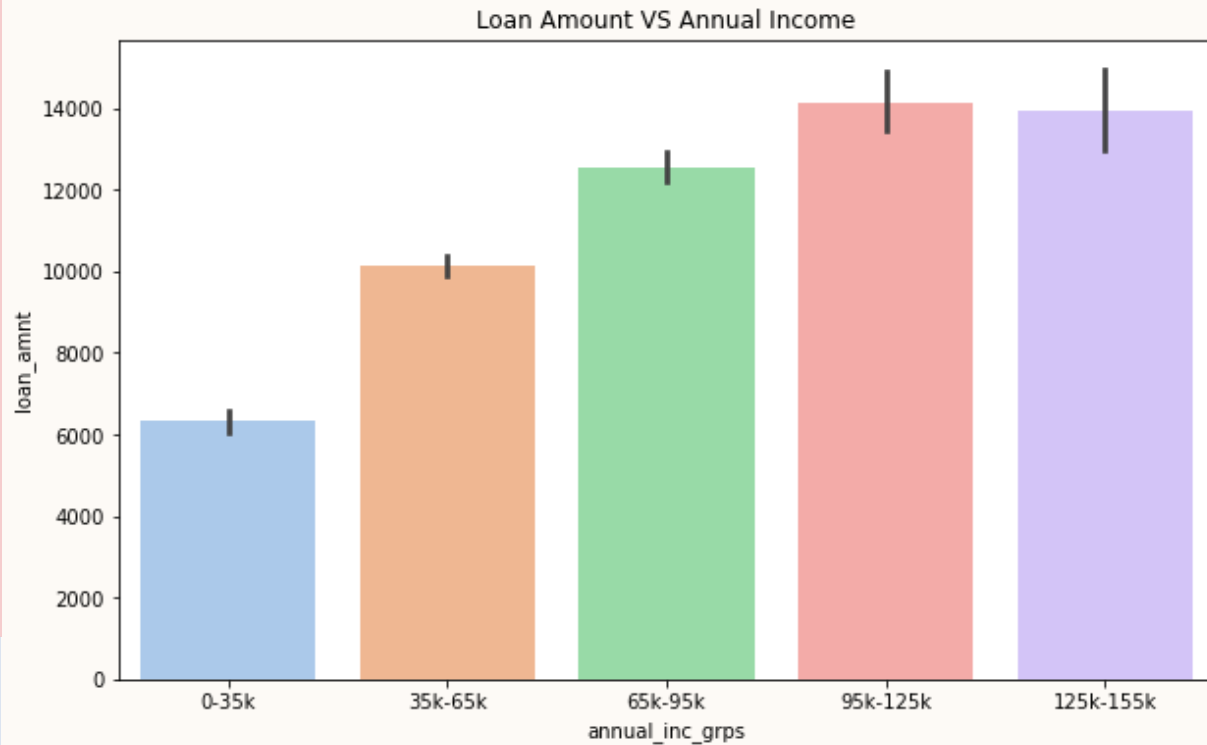

Public Record Bankruptcies

Installment

Applicants who have installments between 150 to 300 tend to face difficulties in paying their charged off loans.

BIVARIATE ANALYSIS

Loan Amount VS DTI

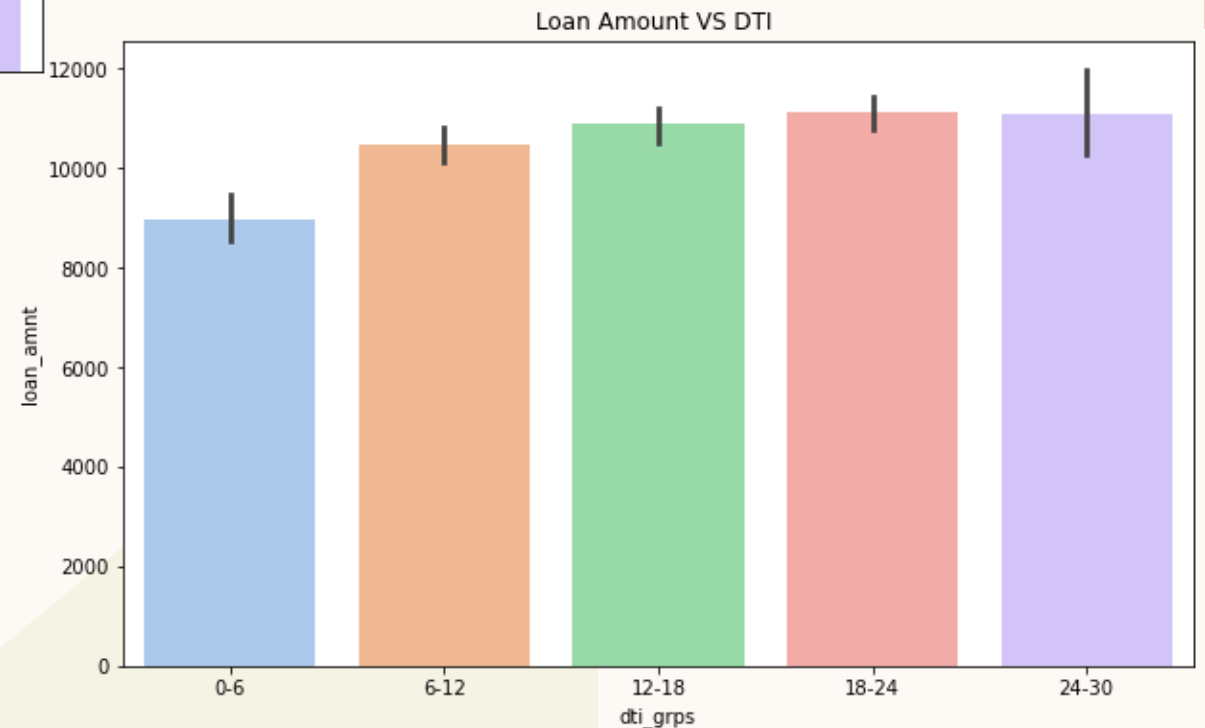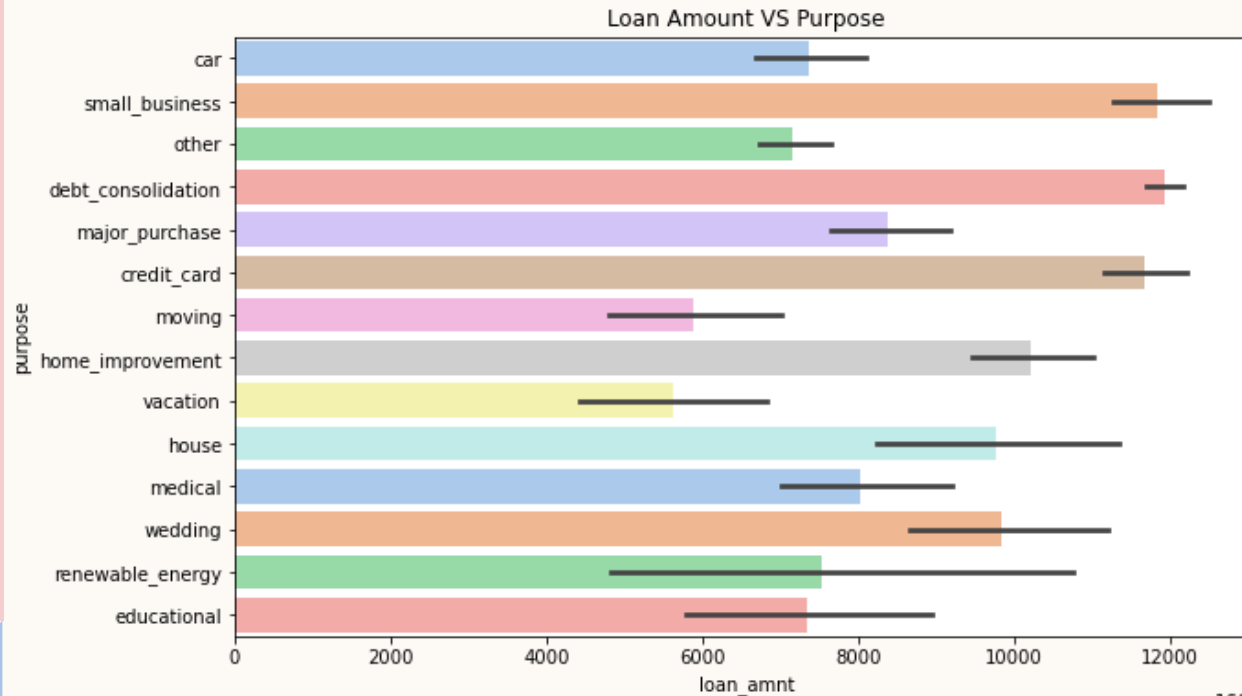As the DTI increases, loan amount also increases. Higher DTI groups require higher loan amounts

Loan Amount VS Annual Income

As the annual income increases, the loan amount also increases for the higher income groups.

Loan Amount VS Grade

Applicants with lower credit card grades are taking larger loan amounts which can lead them to not pay the loan amount
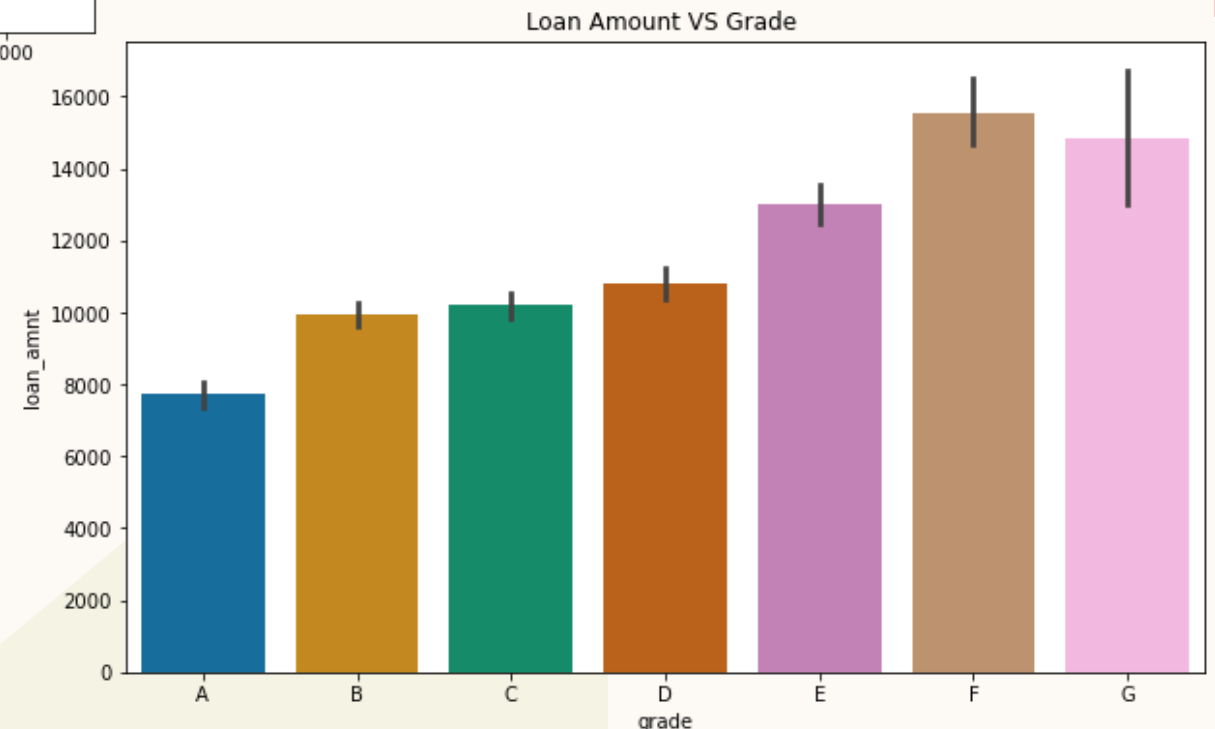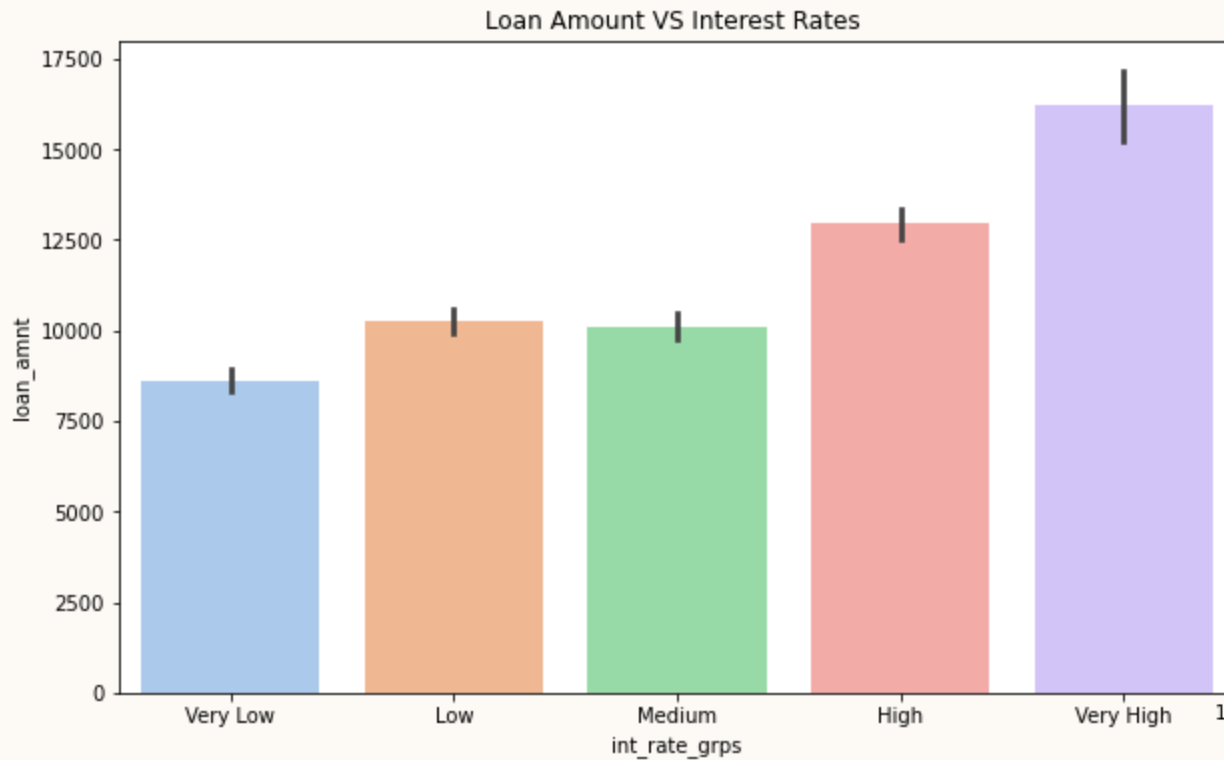
Loan Amount VS Purpose

Loan amounts that are utilized for purposes like 'debt consolidation' and 'small business' have a higher default rate.

Loan Amount VS Employee Length

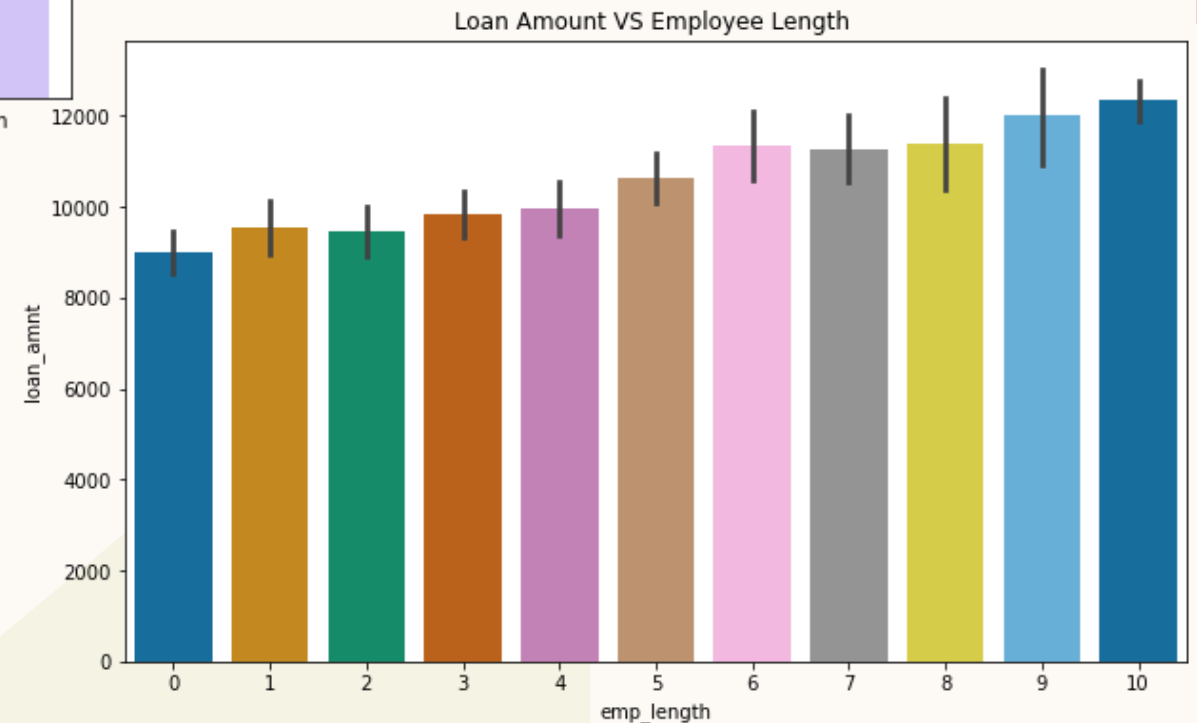Applicants who have more than 10+ years of experience tend to take higher amount of loans.
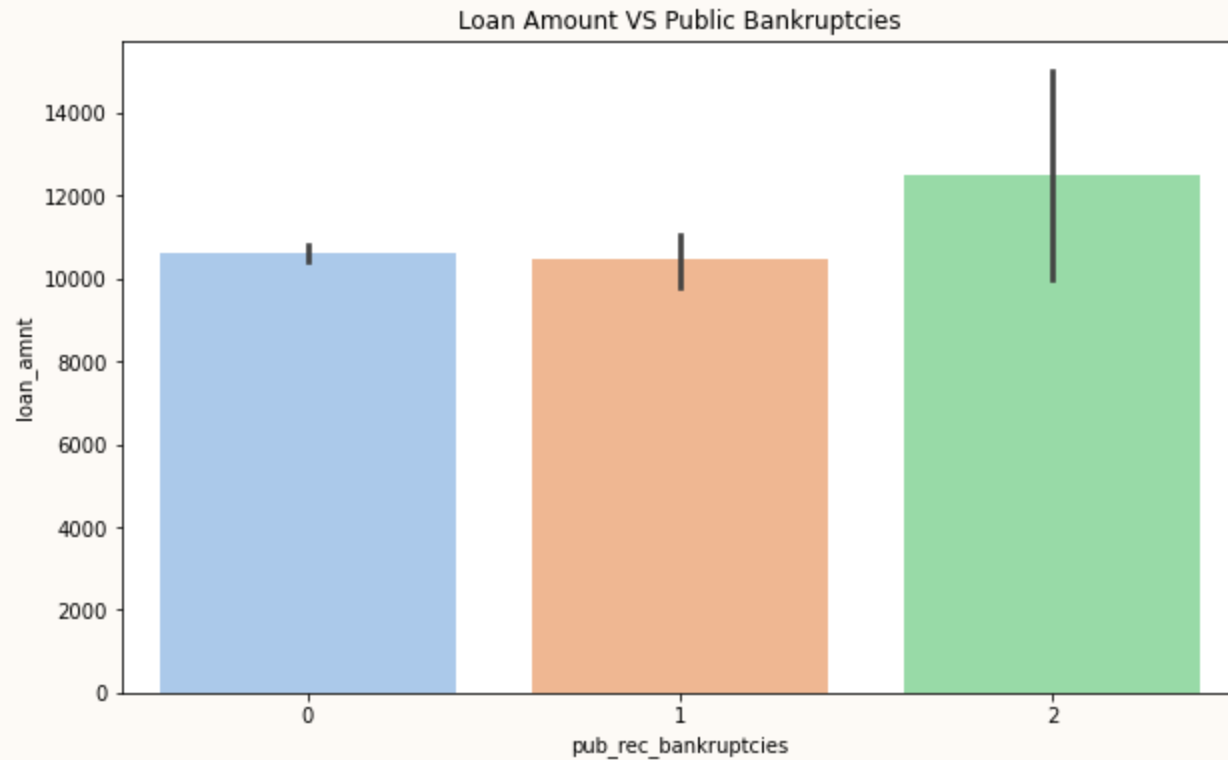
Loan Amount VS Interest Rate

Applicants with very high interest rates tend to take higher loan amounts increasing the possibility of defaulting.

Loan Amount VS Home Ownership

Applicants with 'Others' as their home ownership tend to take higher loan amounts as compared to people those who rent, own or have a mortgage.
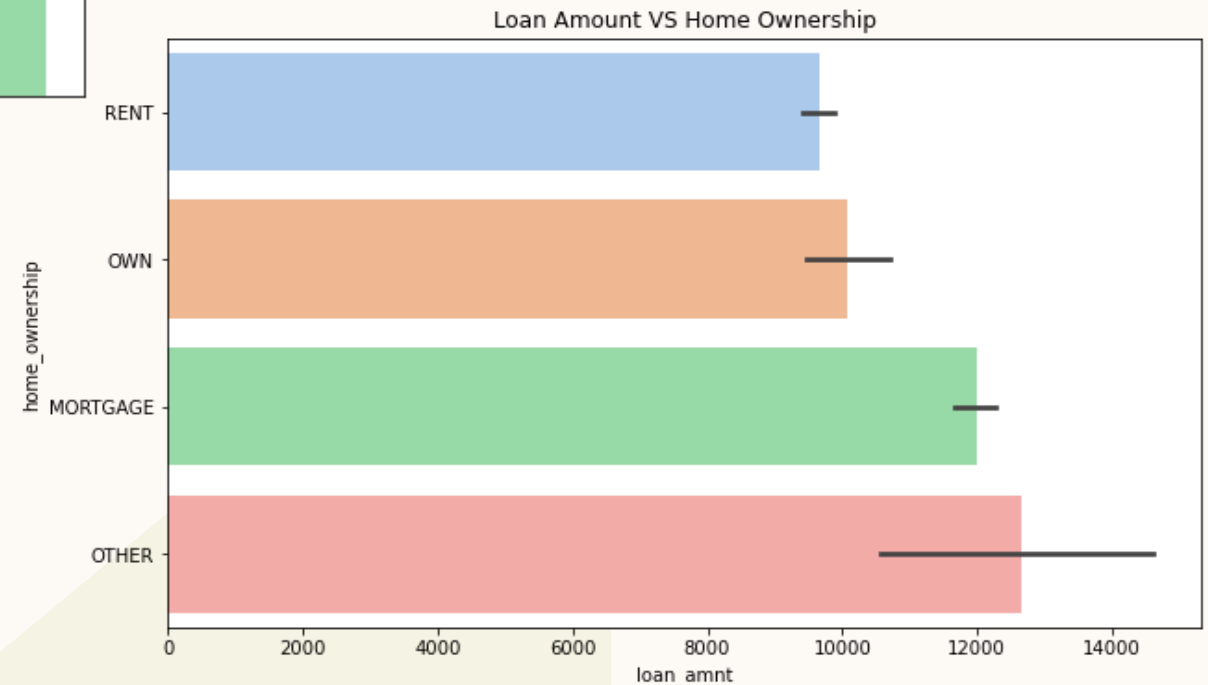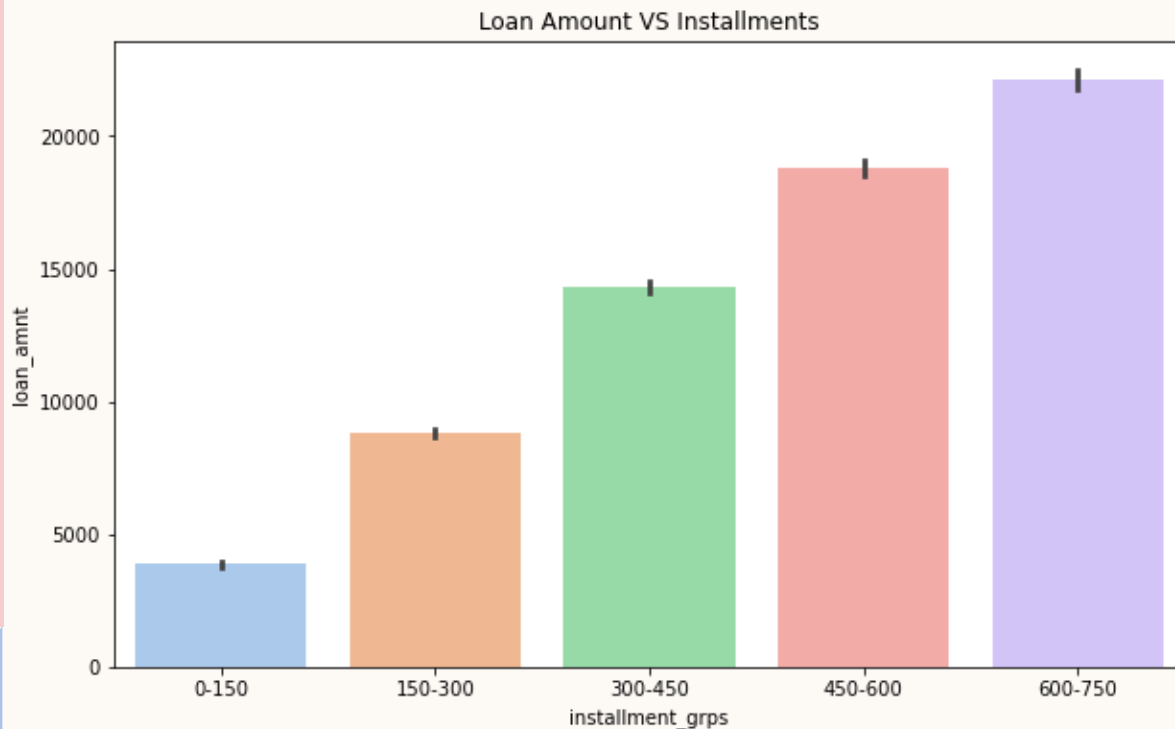
Loan Amount VS Public Bankruptcies

Applicants who have the record of higher public bankruptcies show a higher tendency in loan default.

Loan Amount VS Funded Amount

Loan applicants having higher funded amounts tend to default on their loans as compared to other funded amount groups.

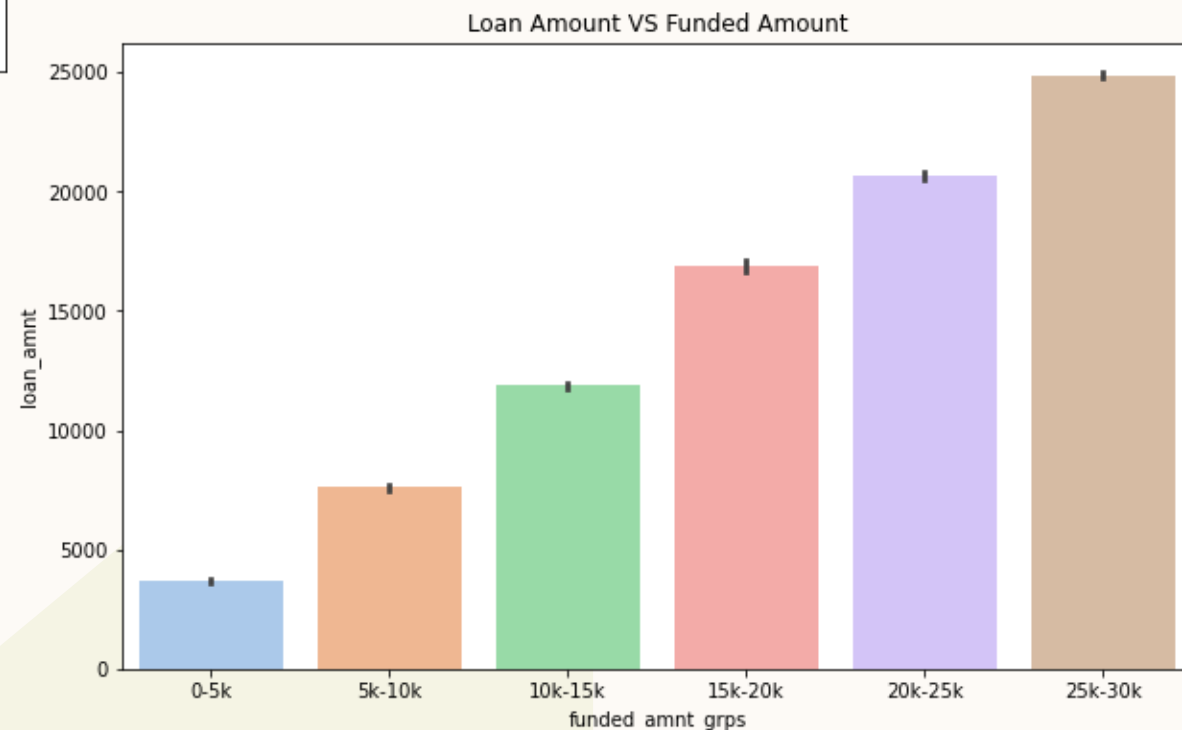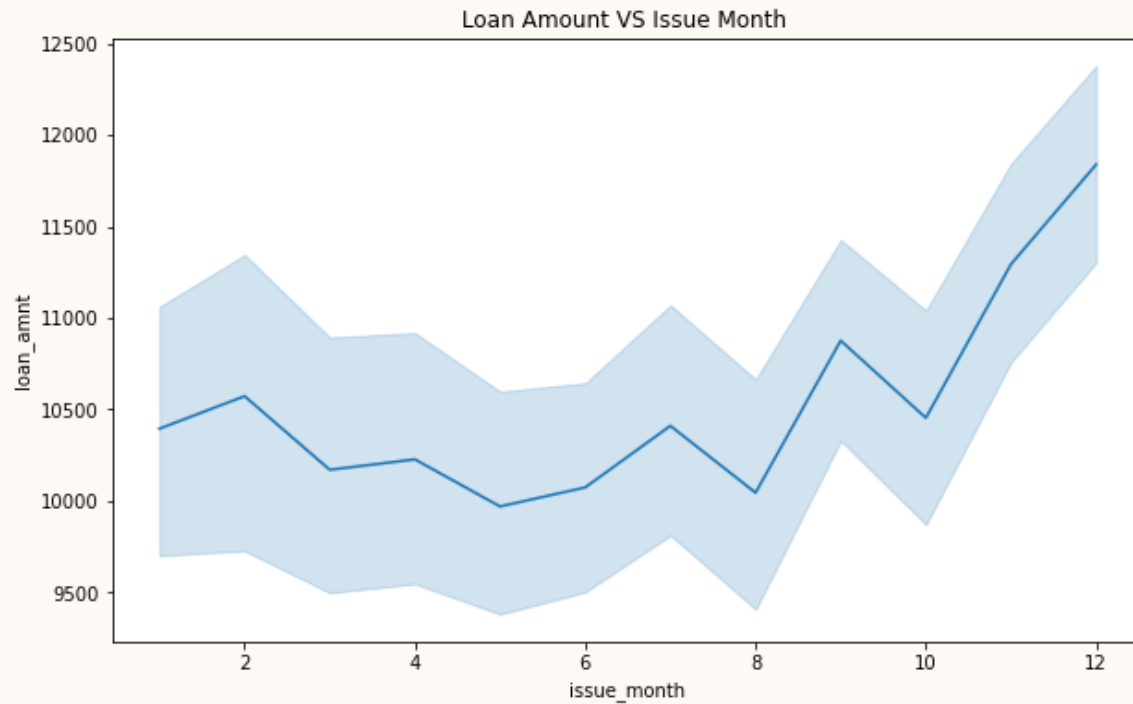Loan Amount VS Installments

Loan Applicants with higher number of installments have a higher tendency to default on their loans

Loan Amount VS Issue Month

Loan Amount VS Issue Year

There was a peak of loan amount in the year 2007 and again in the year 2011 we saw a growth in the loan amount indicating that in years 2007 and 2011, applicants have taken larger loan amounts
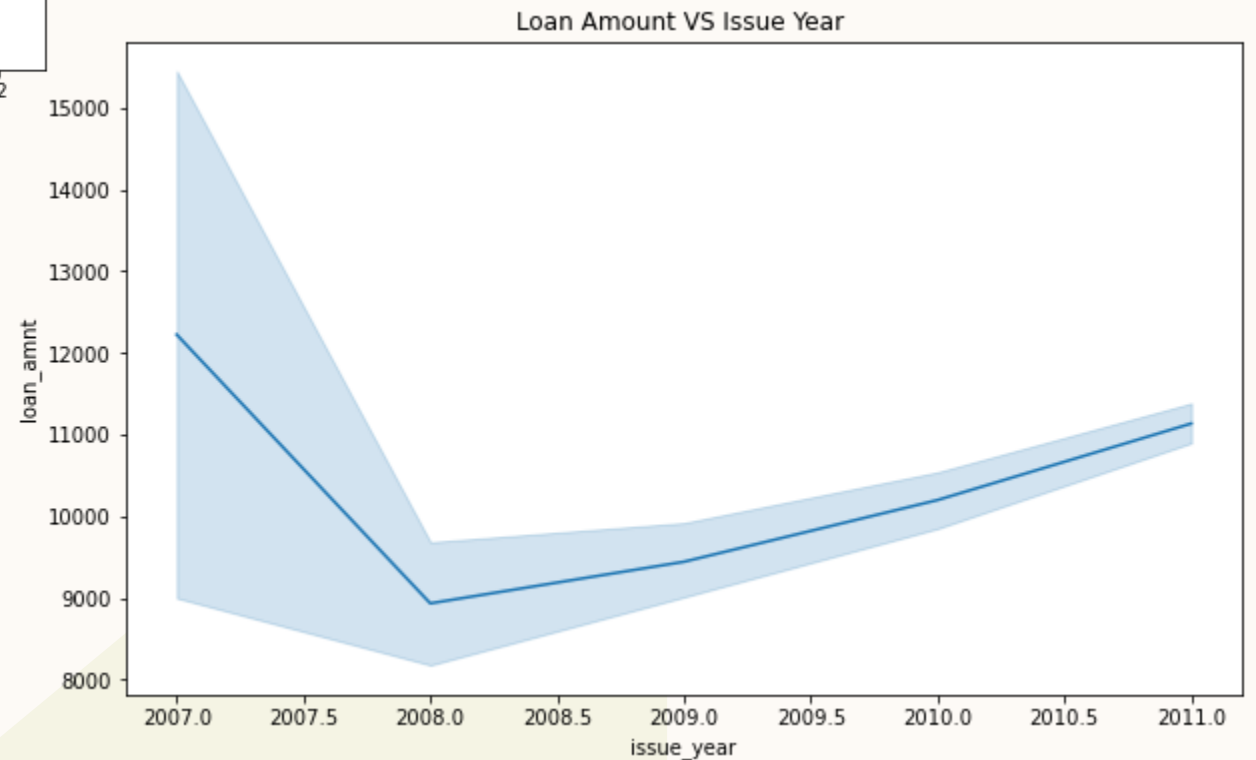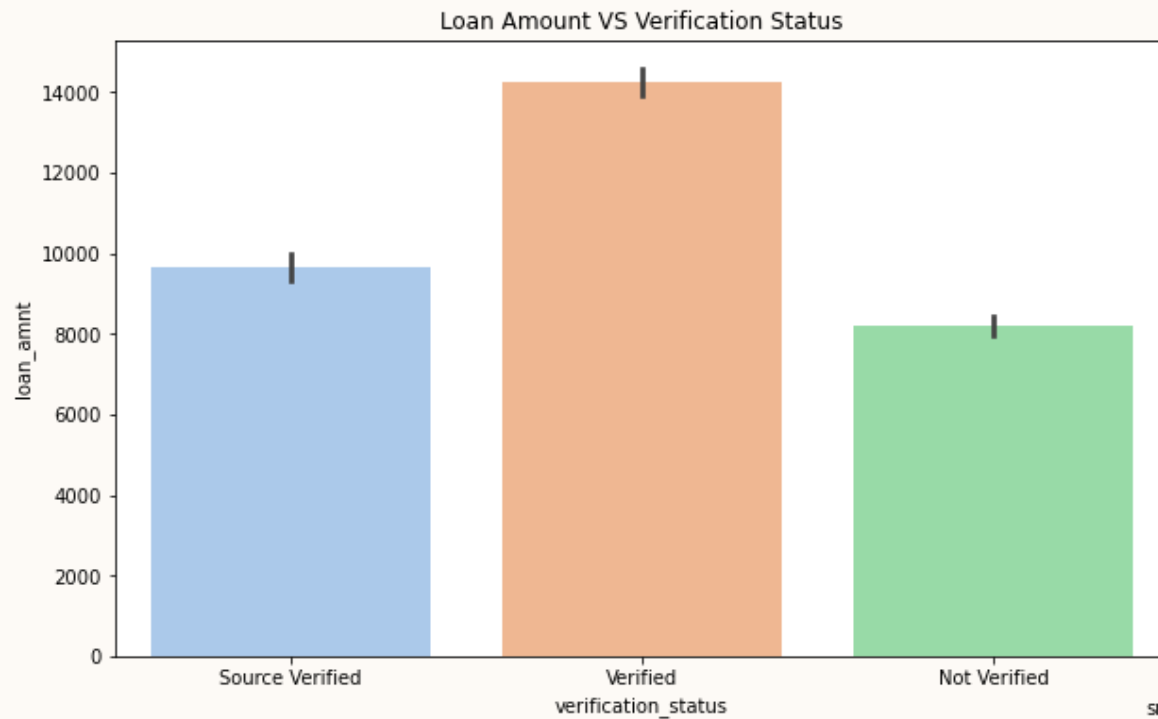
Loan Amount VS Issue Month

There was a peak in the loan amount in the month of December indicating many individuals might have taken loans during the holiday season.

Loan Amount VS Purpose

Applicants who have small business or who are consolidating their debts tend to have higher loan amounts indicating defaulting on their loan.
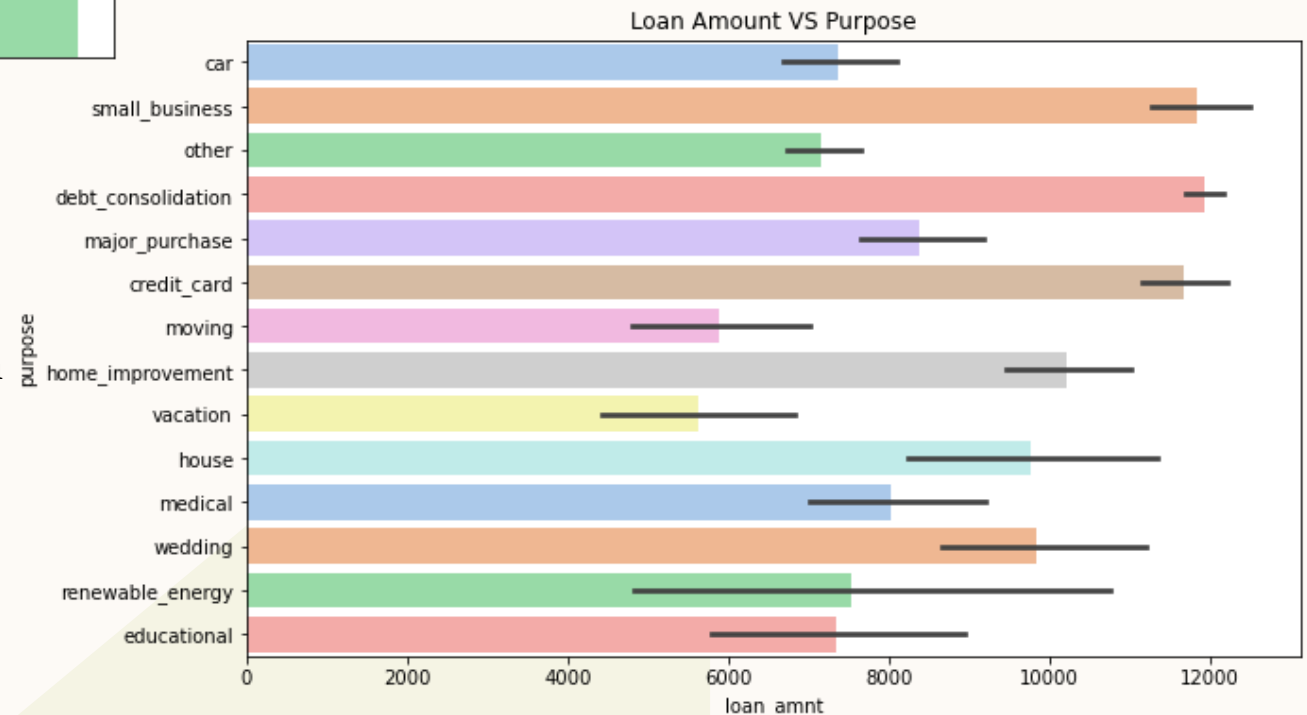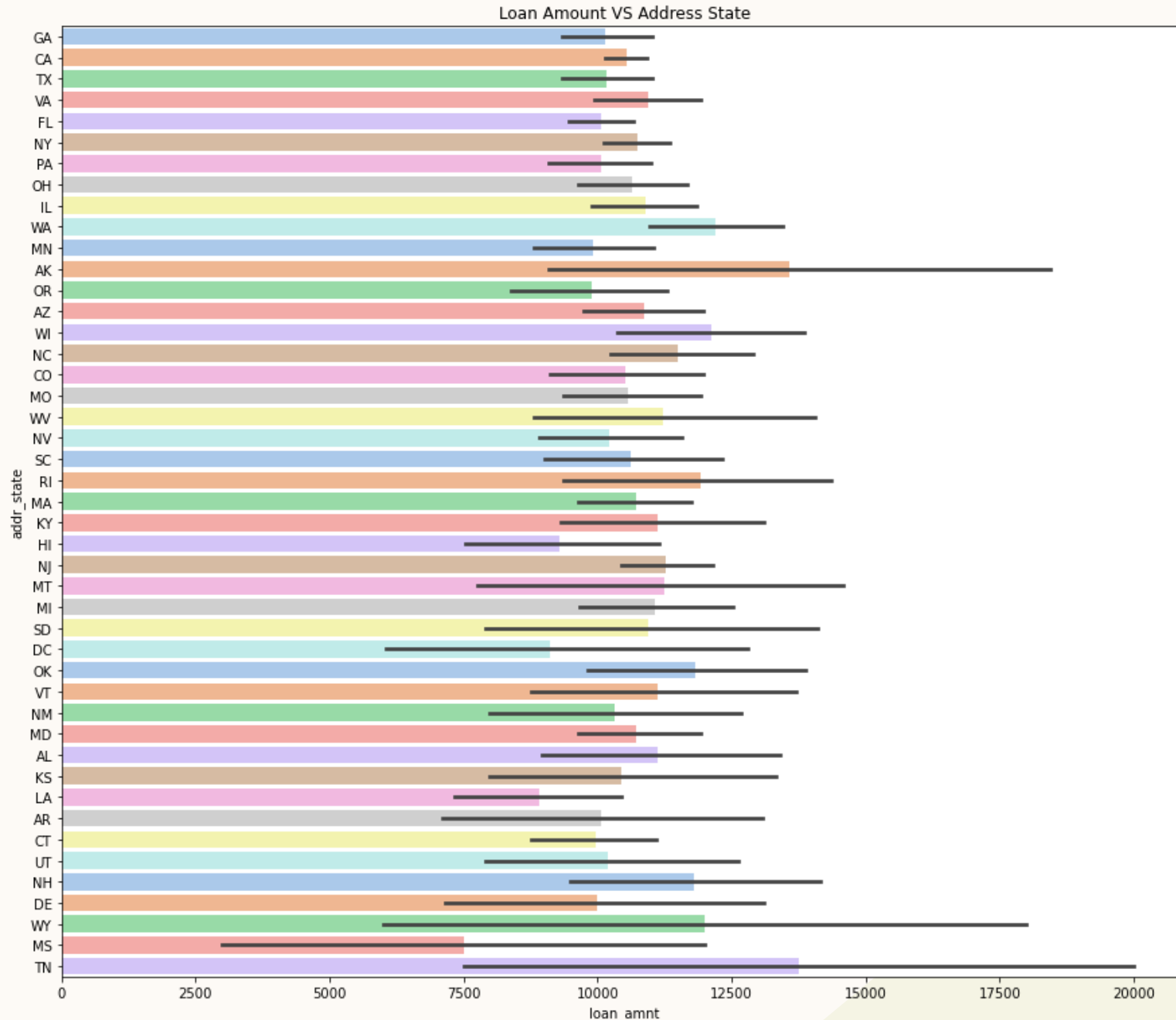
Loan Amount VS Verification Status

Surprisingly, loan applicants having verification status as 'Verified' have a higher chance in loan defaulting.

Loan Amount VS Address State

Applicants from State 'TN' and 'AK have highest number of loan amounts having a tendency to default on their loans.

SUMMARY

- Applicants with High DTI have a higher tendency to default on their loans.

- Applicants who have a higher range of annual income can also tend to have higher default rates.

- Applicants with a record of public Bankruptcies are more likely to default on their loans.

- Surprisingly, applicants having verification status as 'Verified' have higher default rates.

- Applicants with lower credit card grades like F or G can also lead to loan defaulting.

- Loans with a very high interest rate have a higher tendency in the possibility of loan defaulting

- Applicants having more than 10 years of experience also show a higher tendency in loan defaulting.

- Applicants having small business and are consolidating their debts also have higher chances of loan defaulting.

- Applicants having higher funded amounts and high installments are more likely to default.

- Address states 'TN' and 'AK' have the highest number of applicants that have defaulted on their loans.

- Additionally, in the month of December many applicants have defaulted on their loans due to holiday season.