Lending Club Case Study

Group Member:
Abhisek Chakraborty

Problem Statement

- This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- Borrowers can easily access lower interest rate loans through a fast online interface.
- Our objective is to use the information provided and apply our learning to identify risk profiles who can default
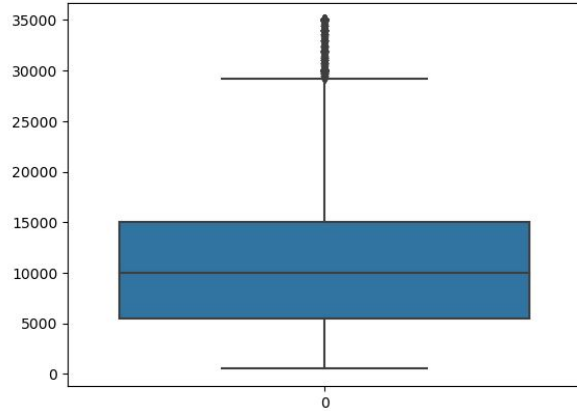
Steps for Exploratory data analysis

1. Data Cleaning
2. Data Understanding
3. Univariate Analysis
4. Segmented Univariate Analysis
5. Bivariate Analysis
6. Recommendations

Analysis of a few boxplots
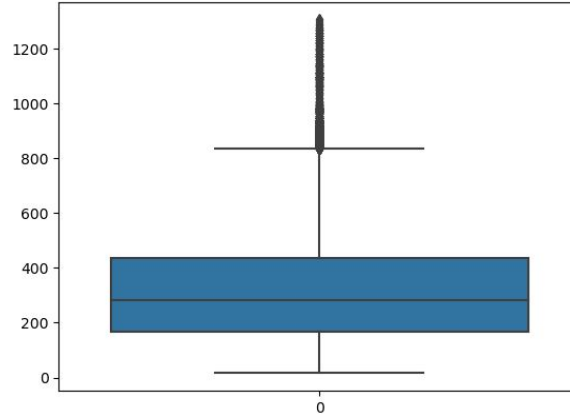


```
sns.boxplot(df['loan_amnt'])
```
<Axes: >

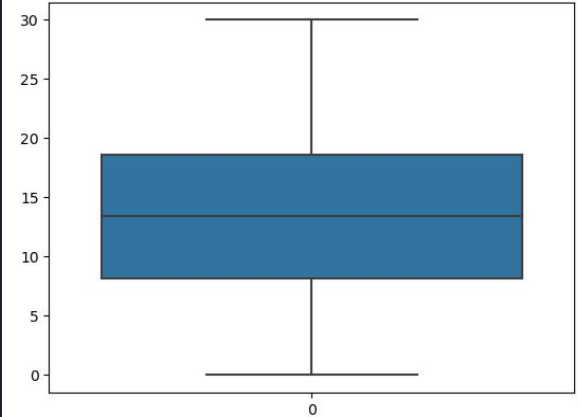```
sns.boxplot(df['installment'])
```
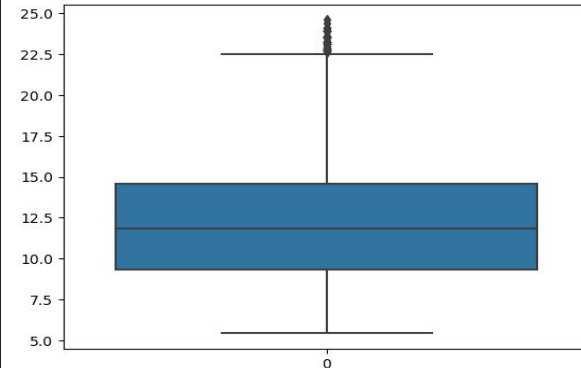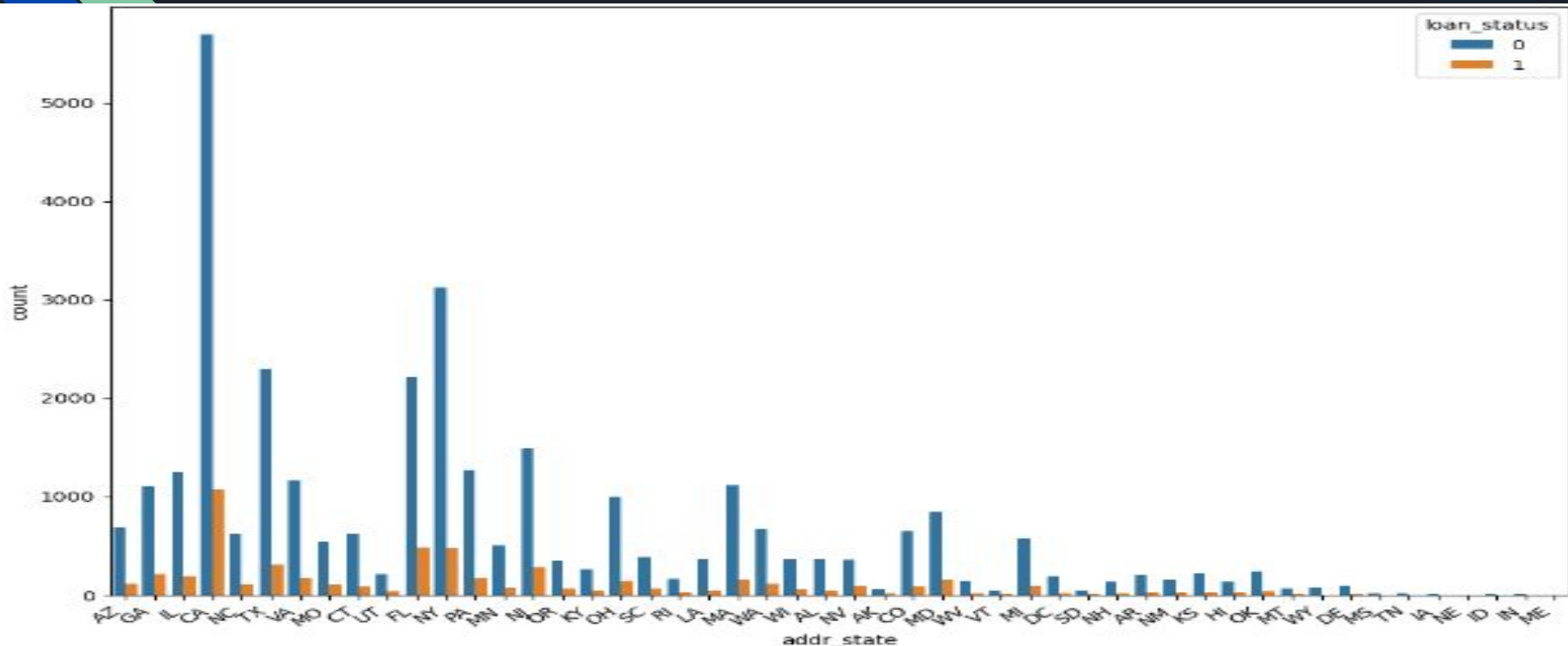<Axes: >



```
sns.boxplot(df['dti'])
```
<Axes: >



```
sns.boxplot(df['int_rate'])
```
<Axes: >

- Most loans taken are between 5000 to 15000 $
- Most installment ranges from 180 - 420 $
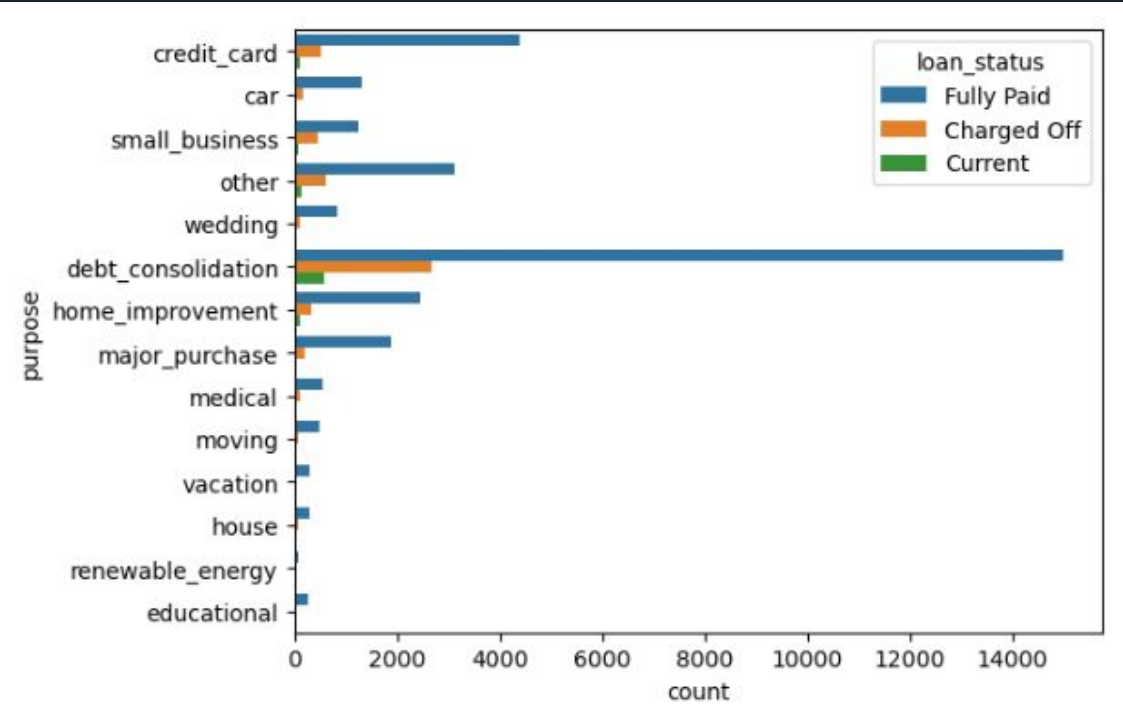- Interest rates for loan are mostly between 9%-15%
- Most dti ranges from 8 to 18

# Analyzing the states which sees the maximum defaults



In the above bar graph 1 is the charge off loans and 0 is the fully paid loans

Observation: States like CA TX FL NY NJ has the maximum charge off
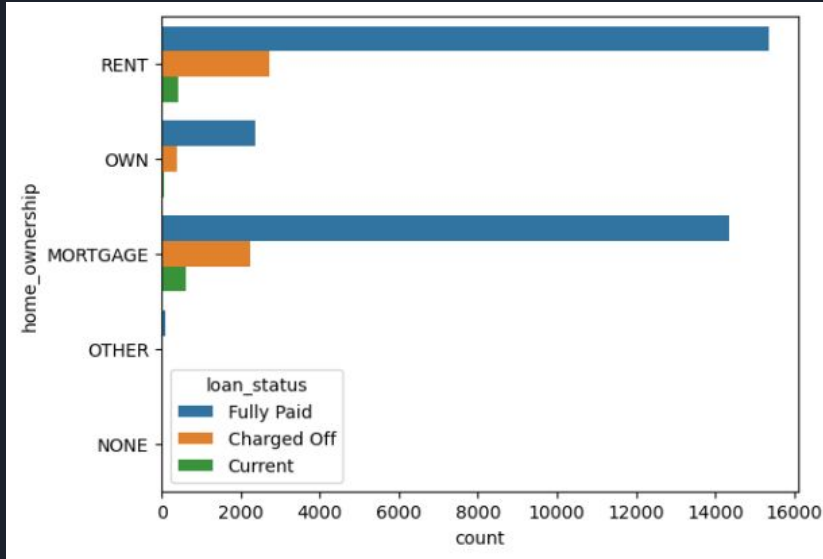
# Analyzing purpose of loan data



Observation:
Most of the loans were taken for debt consolidation & credit card bills. Charge off is high for them as well.

# Analyzing home ownership data



Observation: Most of the applicants are either staying in rented property or mortgaged their home. Since applicants are high in these segments so is the Charge off.
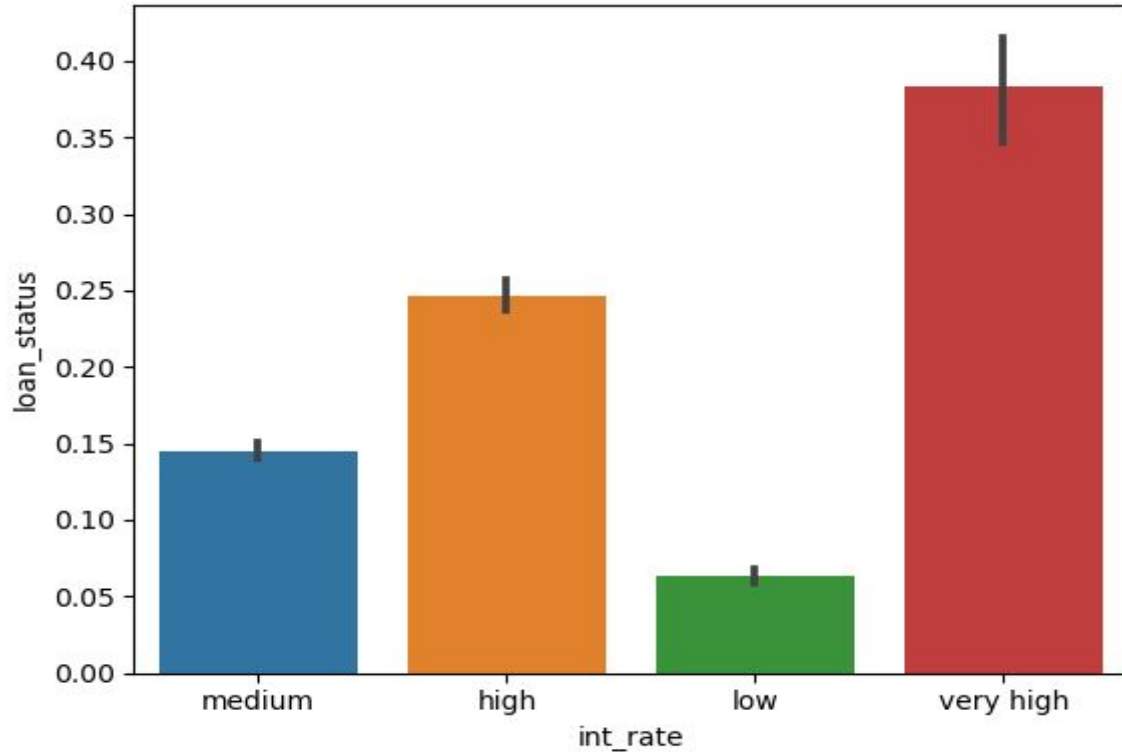
# Analyzing the term of loan and loan status of the data
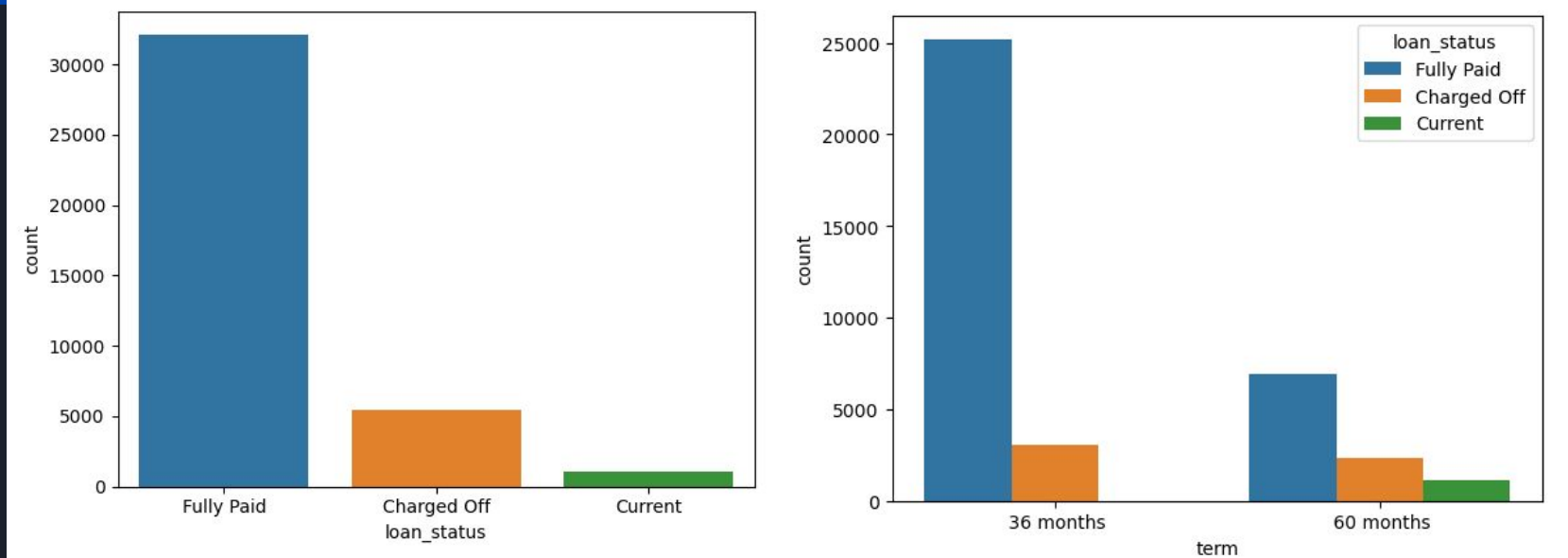


- Loan terms are 36 and 60 months. Over 70% of the loans were taken for 36 months term.
- Ignoring the current loans, 15% of the loans are Charged off and 85% are completed paid
- Charge off in 60 months loan is higher for the number of applicants
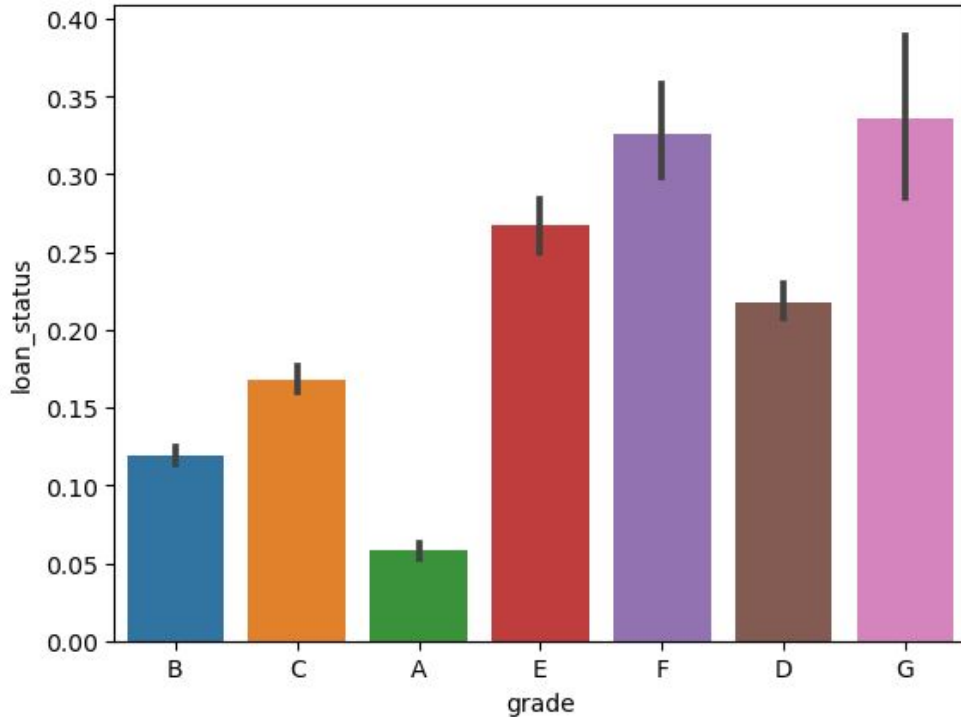
```
Fully Paid   32143
Charged Off   5399
```

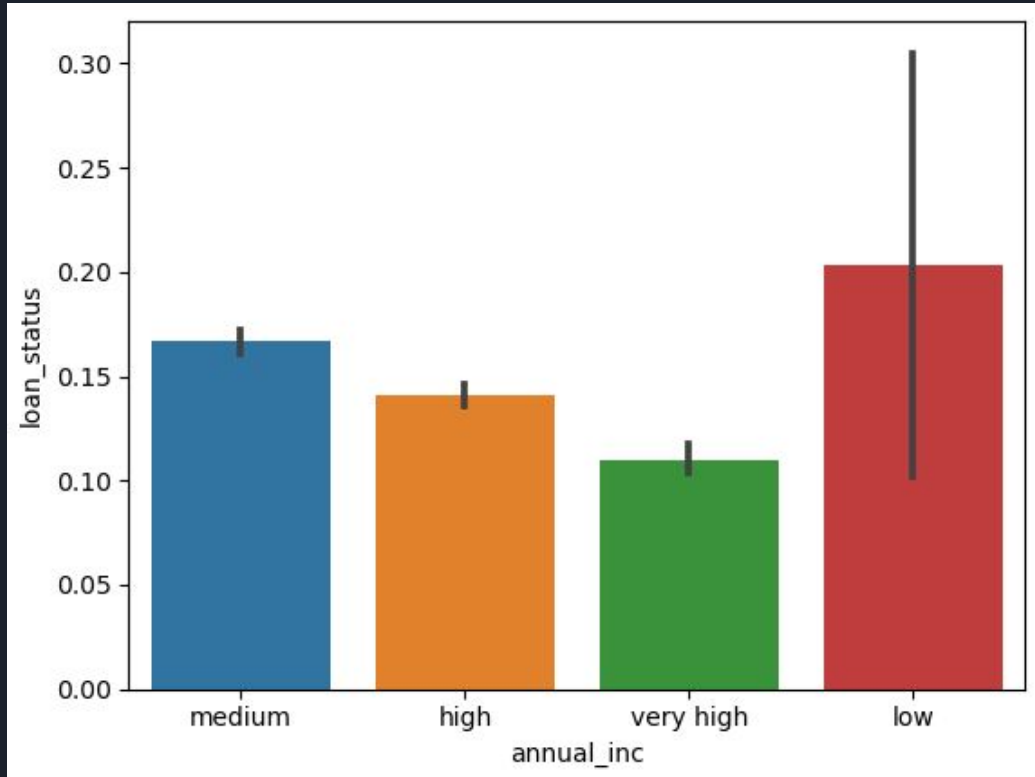# Analyzing the loan status vs interest rate



Observation:
Higher the interest rate higher the chance of default (Charge off) mainly above 15%

# Analyzing the loan status vs grade



Observation:
Grade D-G tend to default

# Analyzing the loan status vs annual income



```python
def fun2(n):
    if n<10000:
        return 'low'
    elif n>=10000 and n<50000:
        return 'medium'
    elif n>=50000 and n<90000:
        return 'high'
    elif n>=90000:
        return 'very high'
```
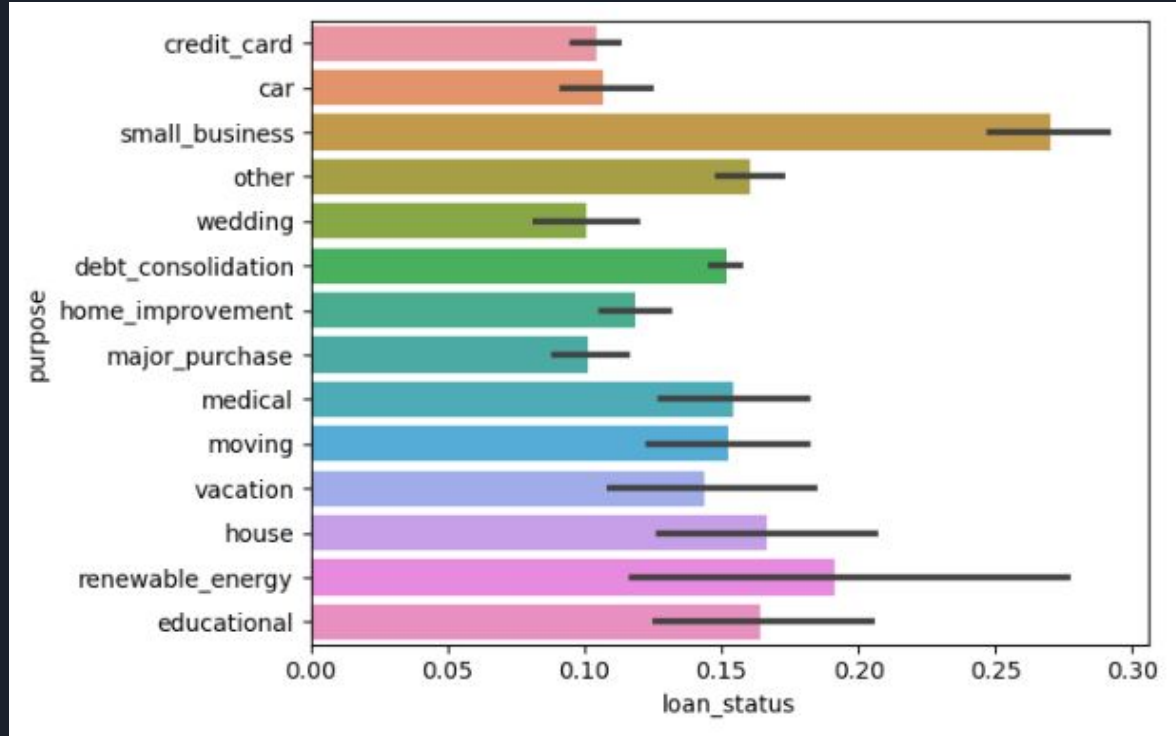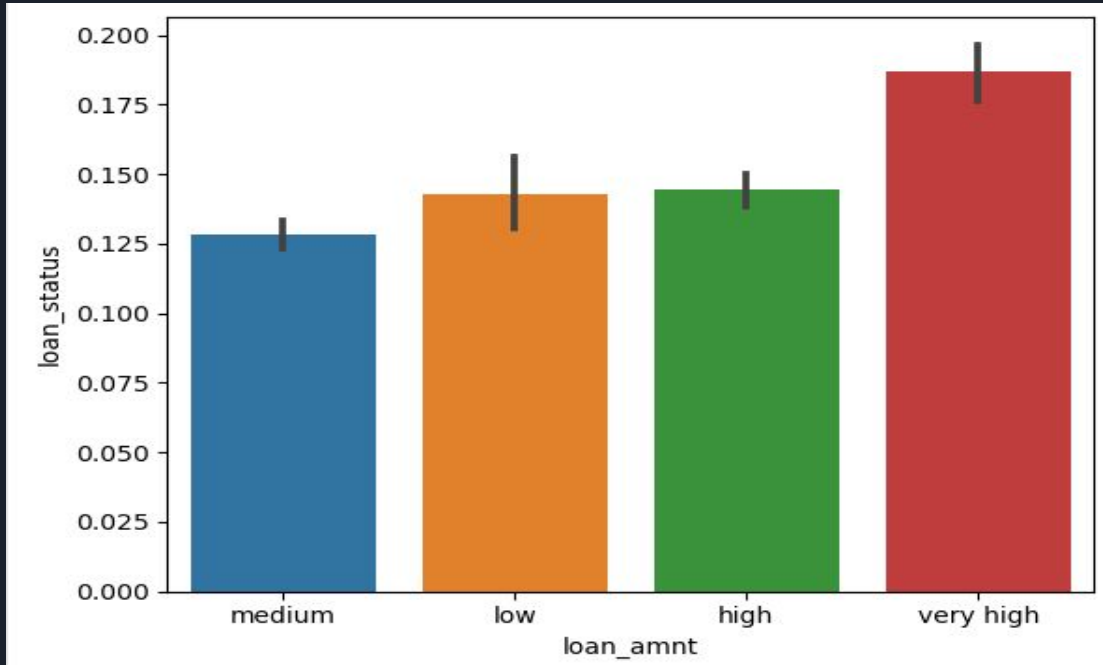
Observation: Higher the income lower the chance of default (Charge off)

Analyzing the purpose of loan vs loan status

Observation: Loans on Small Business, Educational and Renewable Energy higher the chance of default

Analyzing the loan status vs installments


```python
def fun(n):
    if n<3000:
        return 'low'
    elif n>=3000 and n<10000:
        return 'medium'
    elif n>=10000 and n<20000:
        return 'high'
    elif n>=20000:
        return 'very high'
```
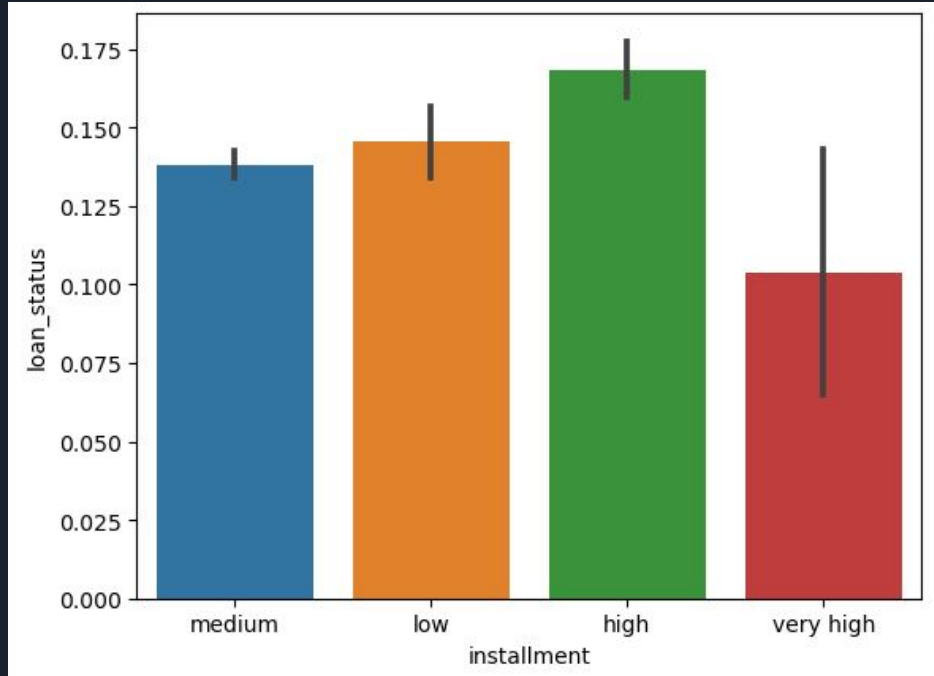
Observation: Higher the loan amount higher the chance of default for loans above 20K

Analyzing the loan status vs installments


```python
def fun3(n):
    if n<100:
        return 'low'
    elif n>=100 and n<500:
        return 'medium'
    elif n>=500 and n<1000:
        return 'high'
    elif n>=1000:
        return 'very high'
```

Observation: Installment amount between 500 - 1000 $ has a higher chance of default

Conclusions & Observations

Grade D-G tend to charged off

Longer the term increases the chance of charged off

Home ownership doesn't directly relate to charged off

Higher the loan amount higher the chance of charged off for loans above 20K

Higher the interest rate higher the chance of charged off mainly above 15%

Lower the income higher the chance of charged off

Loans on Small Business, Educational and Renewable Energy higher the chance of charged off

Installment amount between 500 - 1000 $ has a higher chance of charged off

Interest rates for loan are mostly between 10%-15%

Close to 14% of loans were charged off

Most of the loans were taken for debt consolidation & credit card bills. Charge off is high for them as well.

Most of the applicants are either staying in rented property or mortgaged their home. Since applicants are high in these segments so is the Charge off.

States like CA TX FL NY NJ has the maximum charge off

Charge off in 60 months loan is higher for the number of applicants