LEADING CLUB ANALYSIS

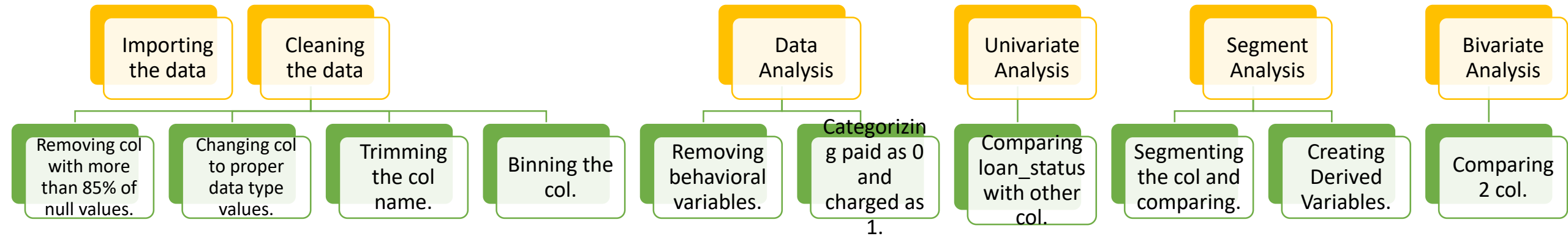Presented By : Sanjana.R and Parth Raichura

LEADING CLUB CASE STUDY

- ## OBJECTIVE -

- 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,** then approving the loan may lead to a **financial loss** for the company.

STRATEGY -

- We should find out **top factors** which lead for an applicant to become a **defaulter** and warn the company about those kind of applicants to save them from **financial loss.**

PROBLEM SOLVING METHODOLOGY

Importing the data
- Removing col with more than 85% of null values.
- Changing col to proper data type values.

Cleaning the data
- Trimming the col name.
- Binning the col.

Data Analysis
- Removing behavioral variables.
- Categorizing paid as 0 and charged as 1.

Univariate Analysis
- Comparing loan_status with other col.

Segment Analysis
- Segmenting the col and comparing.
- Creating Derived Variables.

Bivariate Analysis
- Comparing 2 col.

UNIVARIATE ANALYSIS

Plotted graph of **fully paid** and **Charged Off**. Where fully paid is denoted by 0 and charged off is denoted by 1.
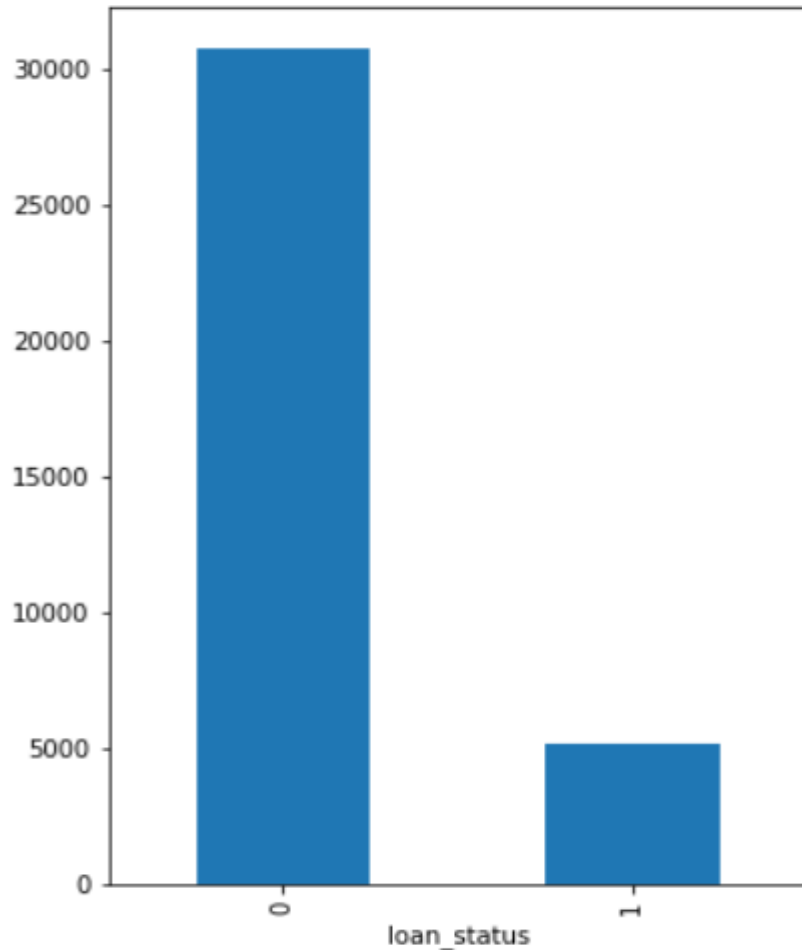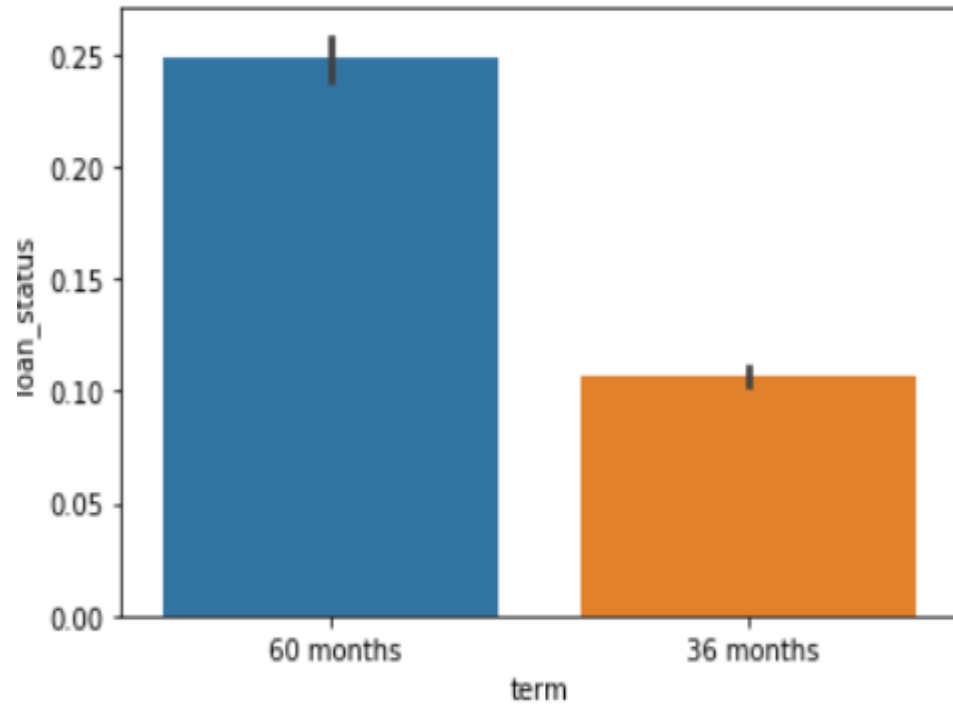

OBSERVATIONS –

- Approximately 14% of loans in the dataset are defaulted.
- Any variable that increases percentage of default to higher than 16.5% should be considered a business risk. (16.5 is 18% higher than 13.97 - a large enough increase)

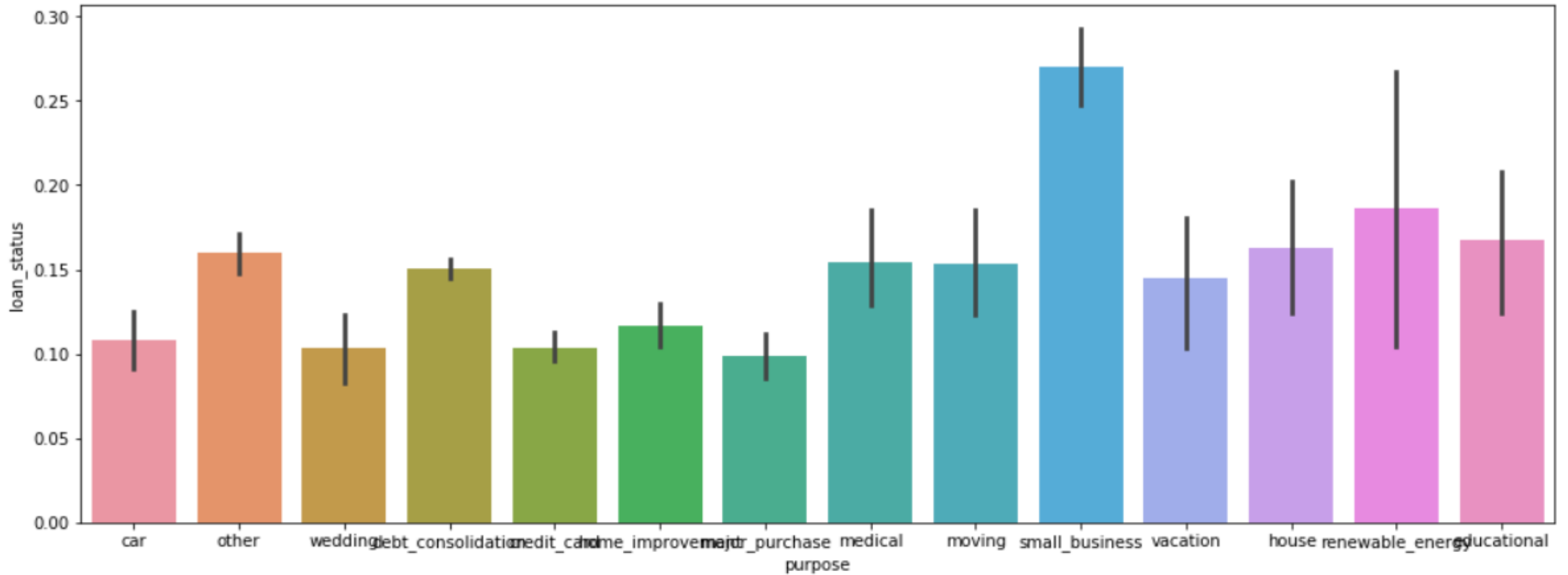# Conti..

Plotted graph between **loan_status** and **term**.

OBSERVATIONS –

- Among Charged Off loans, percentage of term 60 months rises to 45%.
- The higher term loans have a higher chance of default.

Conti..

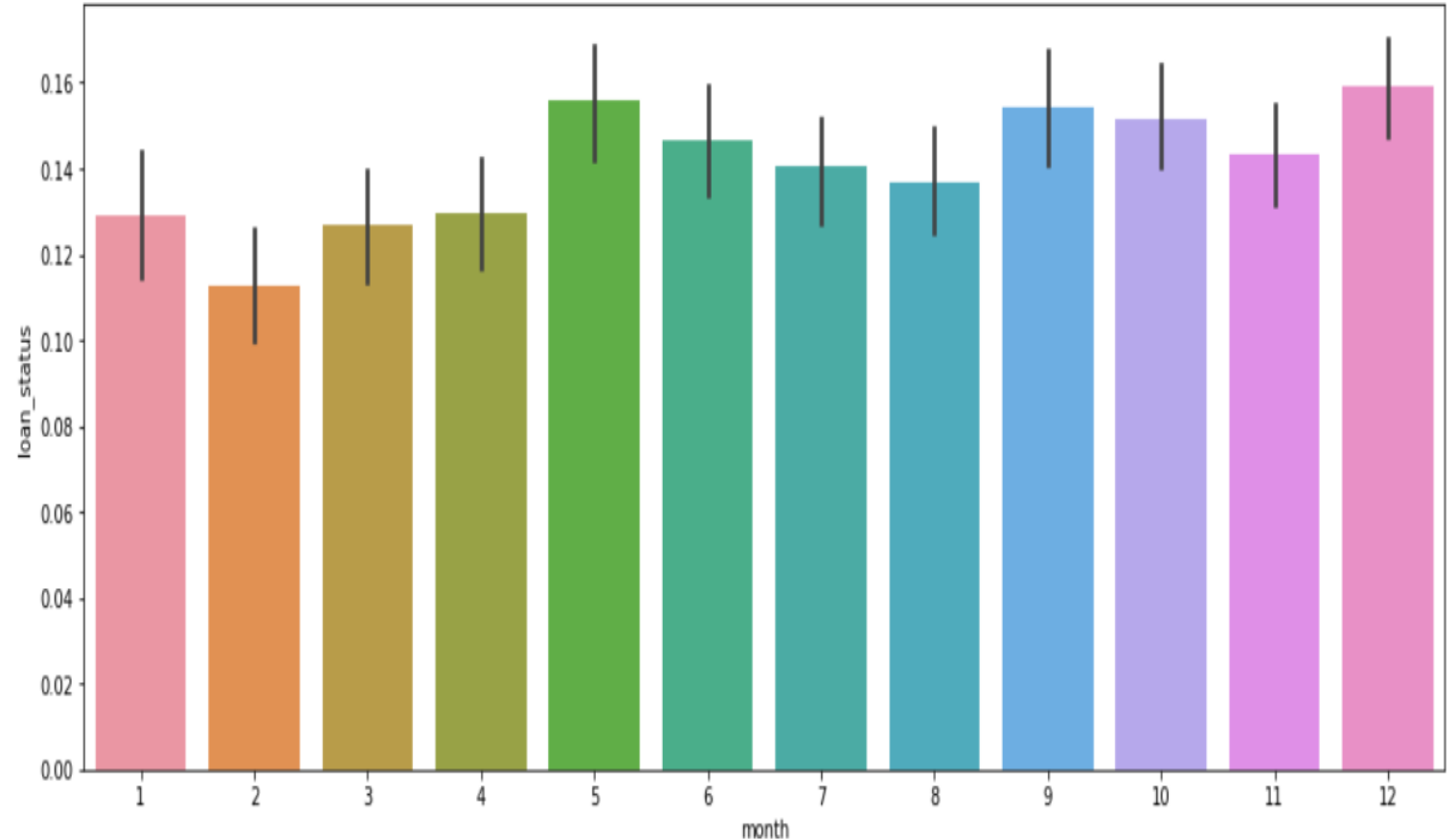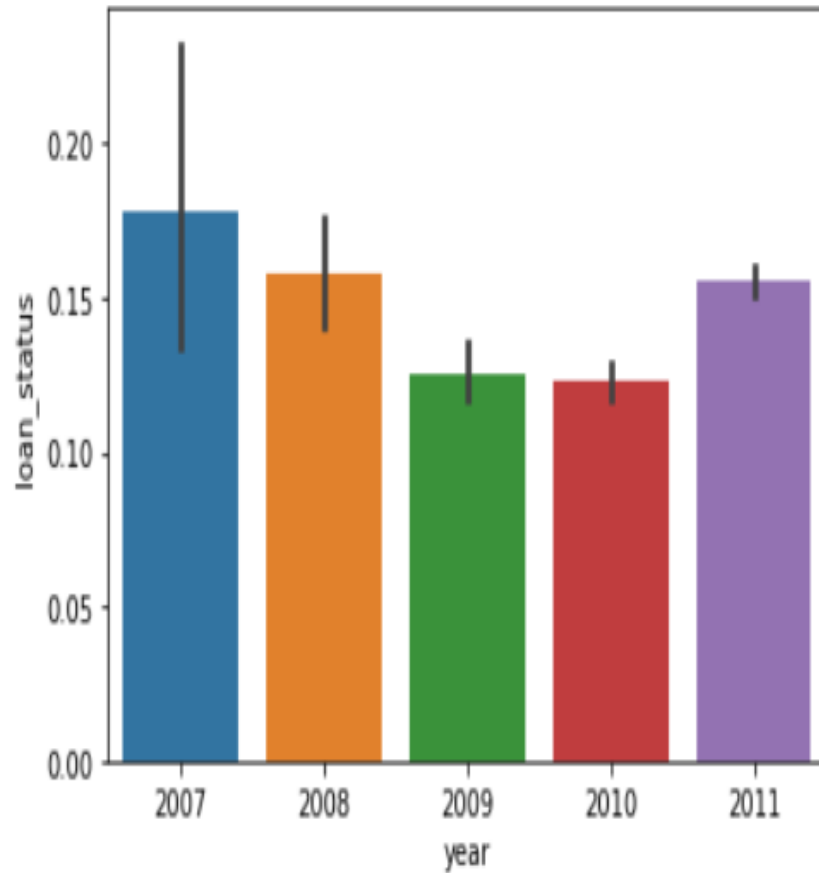Plotted graph between **loan_status** and **purpose**.

OBSERVATIONS –
- 26% of loans for small business are Charged Off. Making them the most risky purpose.
- 17% of the loans for renewable_enrgy are charged Off, but the number is too less to be of significance.

Conti..

Plotted graph between **loan_status – year** and **loan_status – month** separately .

OBSERVATIONS –
- Over the years, LC as an organization has been giving more loans. From Aug,2007 to 2011 the loans issued have risen significantly.
- Within a year, the number of loan issued rises over the month from jan to Dec. Dec is the month of the year were the maximum number of loans are being issued.
- Year of loan has no significant effect on likelihood of default, but we can see that loan taken in certain months are likely to be in defaulter's list.
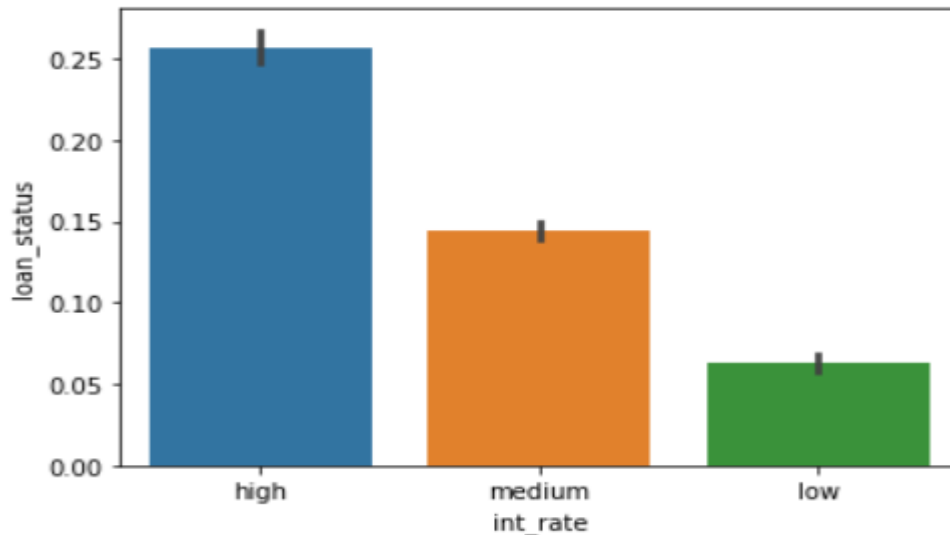
Conti..

Plotted graph between **loan_status** and **interest rate**.
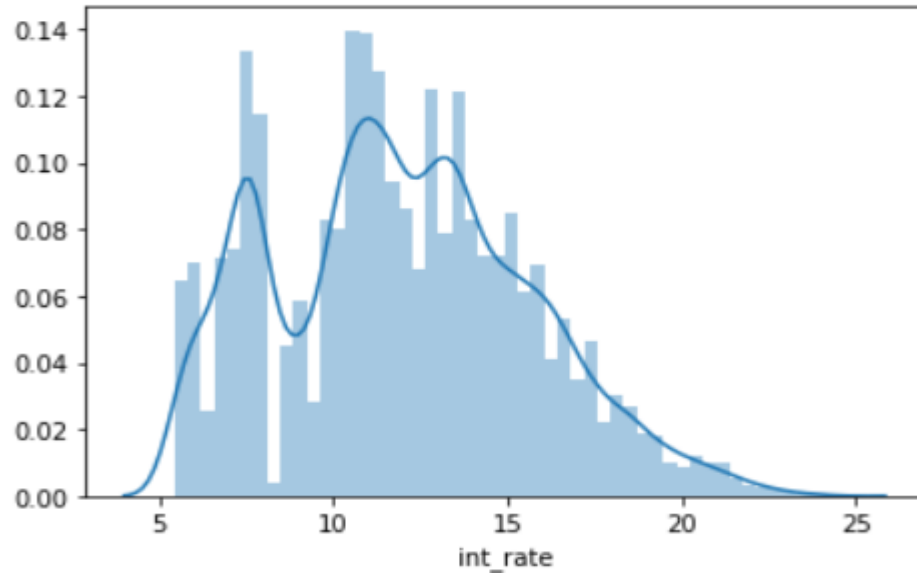
We can categorize
- Rate < = 10 as low
- Rate between 10 and 15 as medium.
- Rest of the amounts as high.

OBSERVATIONS –

- Overall, the intrest rate varies from 5.42% to 24.4% with average intrest rate of 11.8%.
- The intrest rate for Charged Off loans appear to be higher than for Fully paid. This is naturally expected. As, the risk increases the rate of interest imposed on the loan also increases.

Conti..

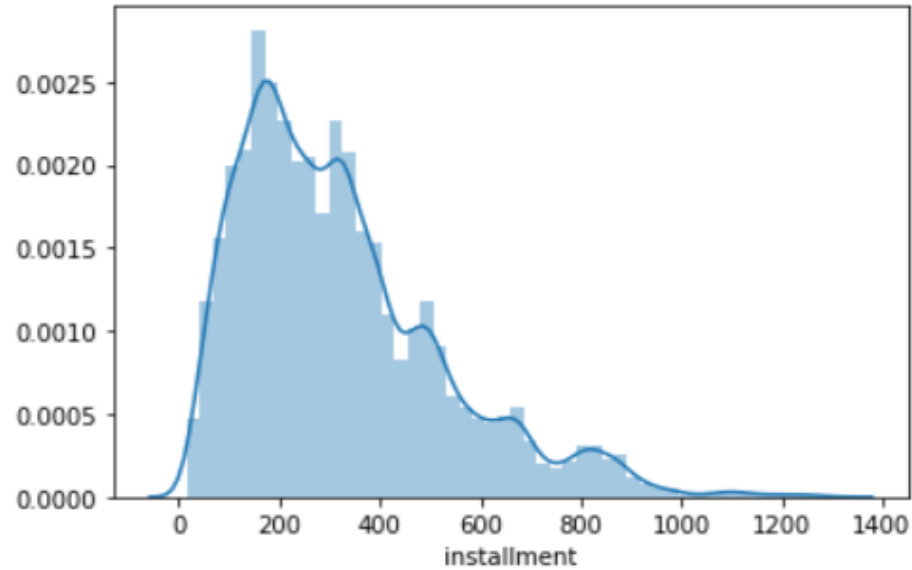Plotted graph between **loan_status** and **dti**.
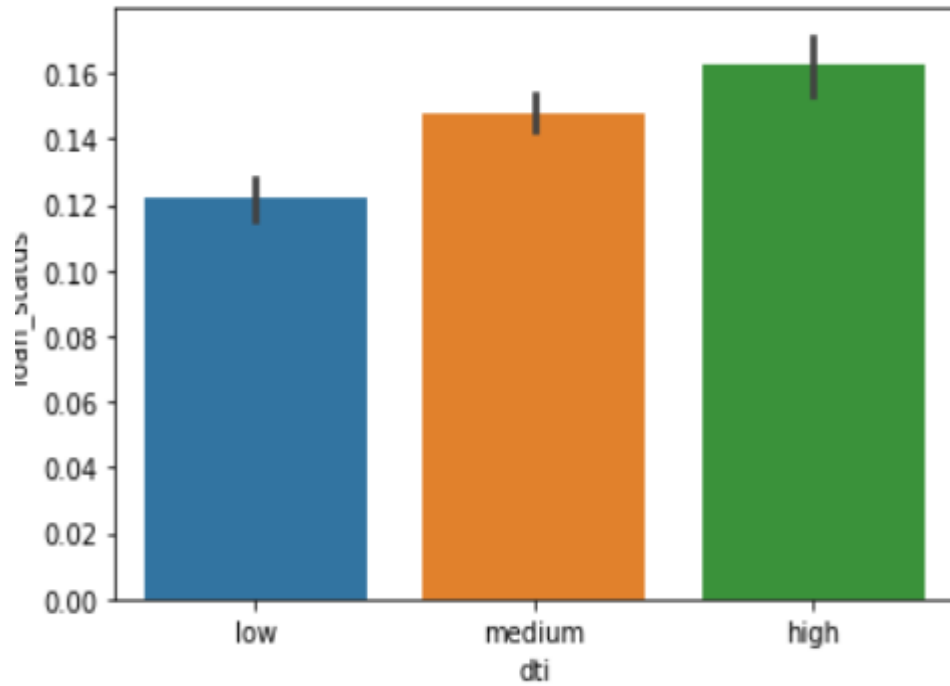
OBSERVATIONS –

- High dti results into high default rate.
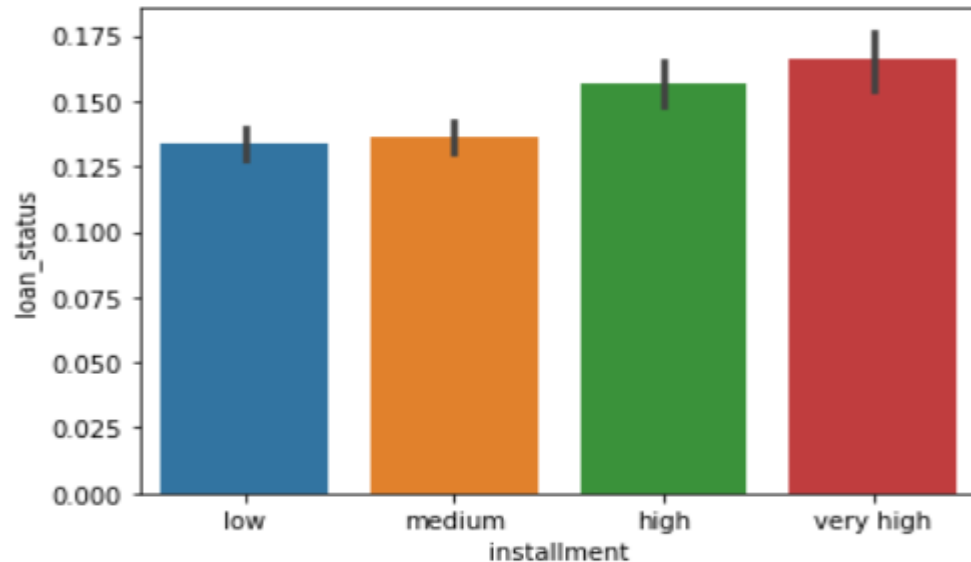- If there is High DTI then the chances of Charge Off is high.

Conti..

Plotted graph between **loan_status** and **installment** .

We can categorize
- Installment < =200 as low
- Installment between 200 and 400 as medium.
- Installment between 400 and 600 as high.
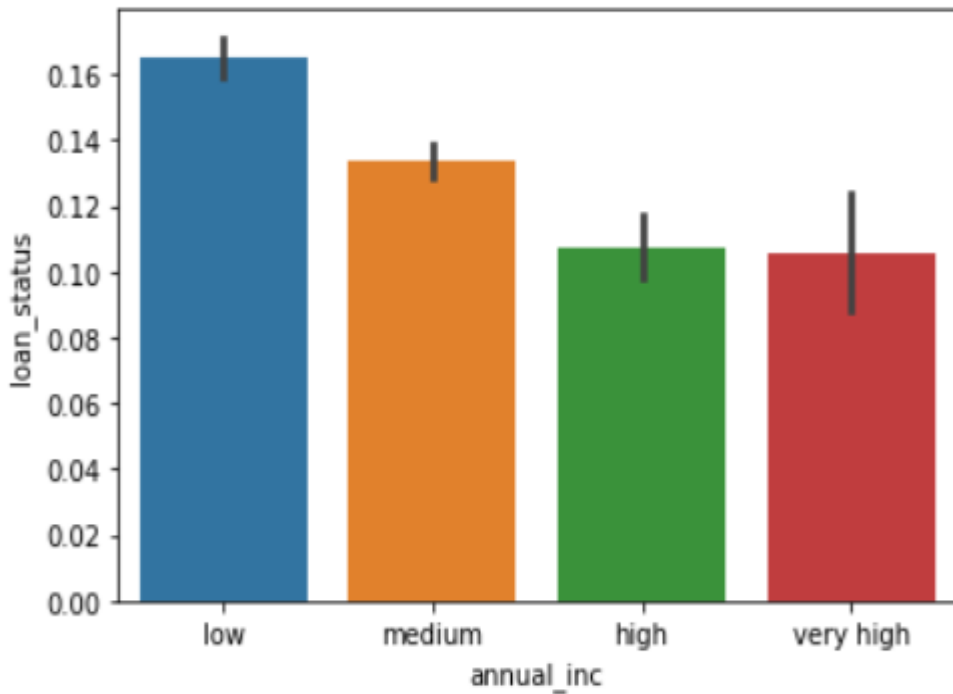- Rest of the amounts as very high.

OBSERVATIONS –

- The higher the installment amount, the higher the default rate.
- Installment amount varies largely between 170.08 to 420 with median of 280 loans Charged Off have high installment on average

Conti..

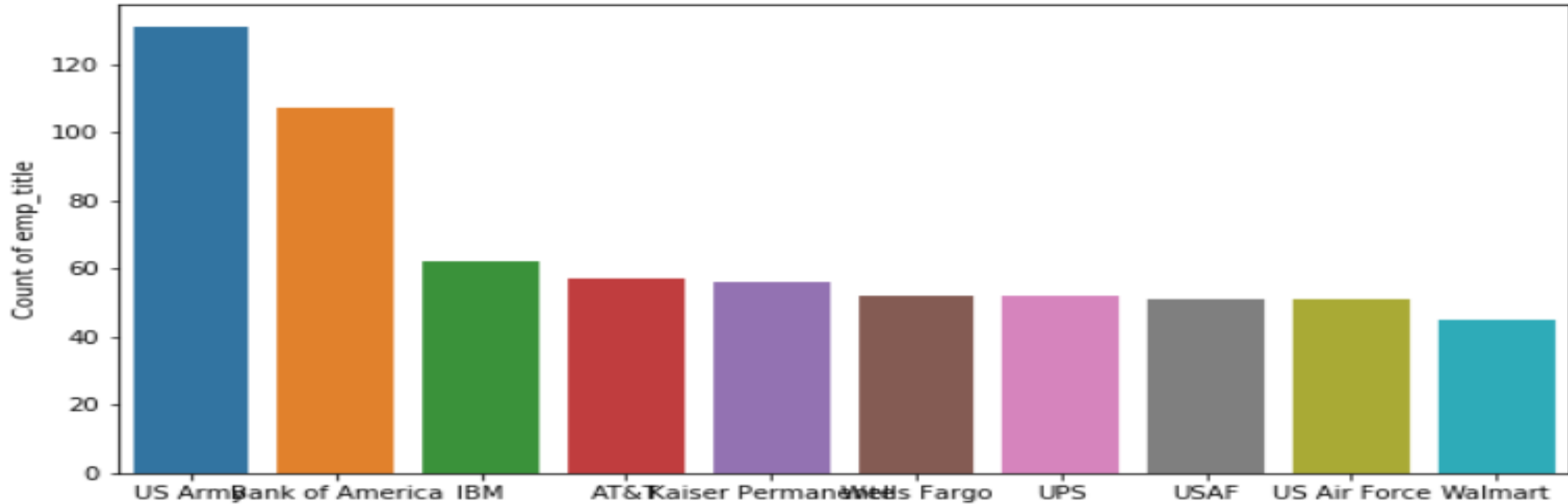Plotted graph between **loan_status** and **annual income**.

OBSERVATIONS –

- The lower the income higher will be the chance of getting Charged Off as the default rate is more.

Conti..

Plotted graph between **loan_status** and **employee_title**.

OBSERVATIONS –

- Max of the loan takers are from the US Army, then Bank of America followed by IBM and AT&T.
- Kaiser Permanents, wells Fargo, UPS, USAF, US Air Force and Walmart are also among the top 10 emp_title loans.
- The numbers per employer are quite low, and this variable is unlikely to have an impact on default rates.
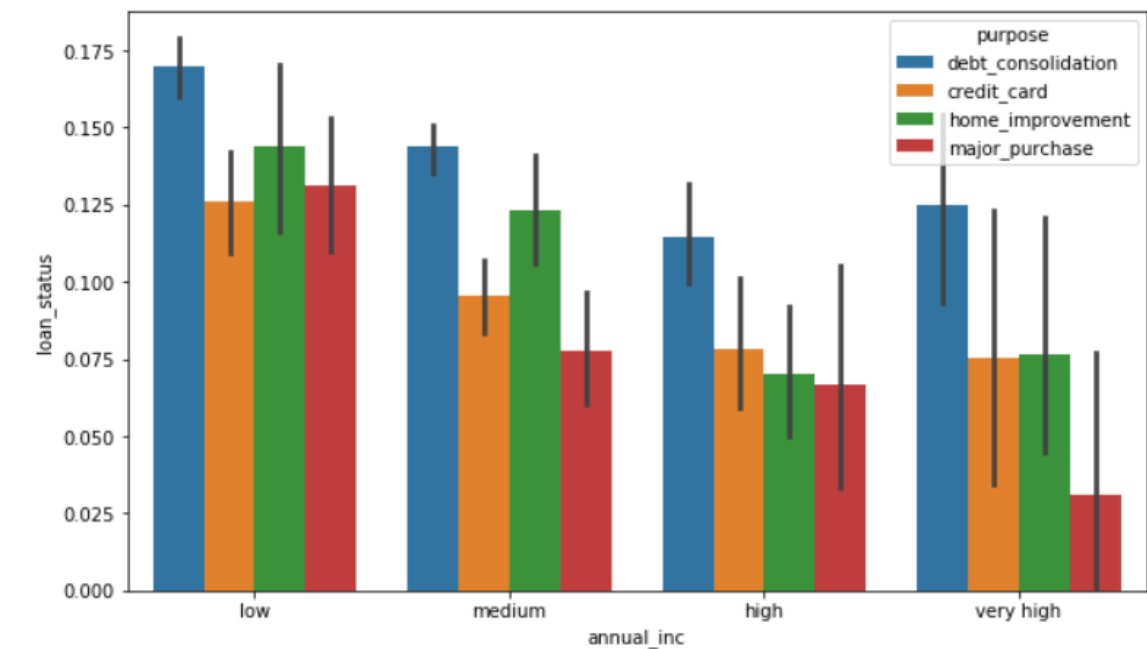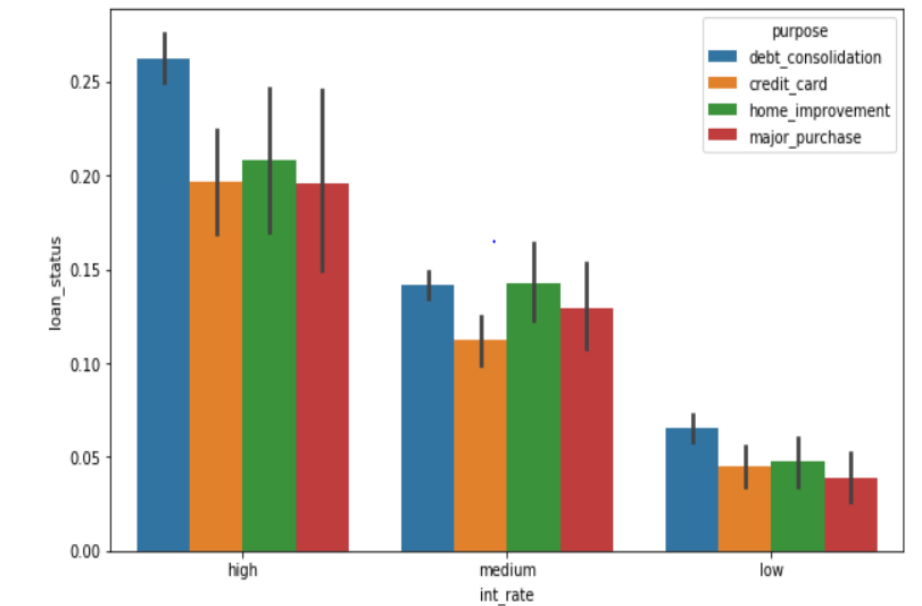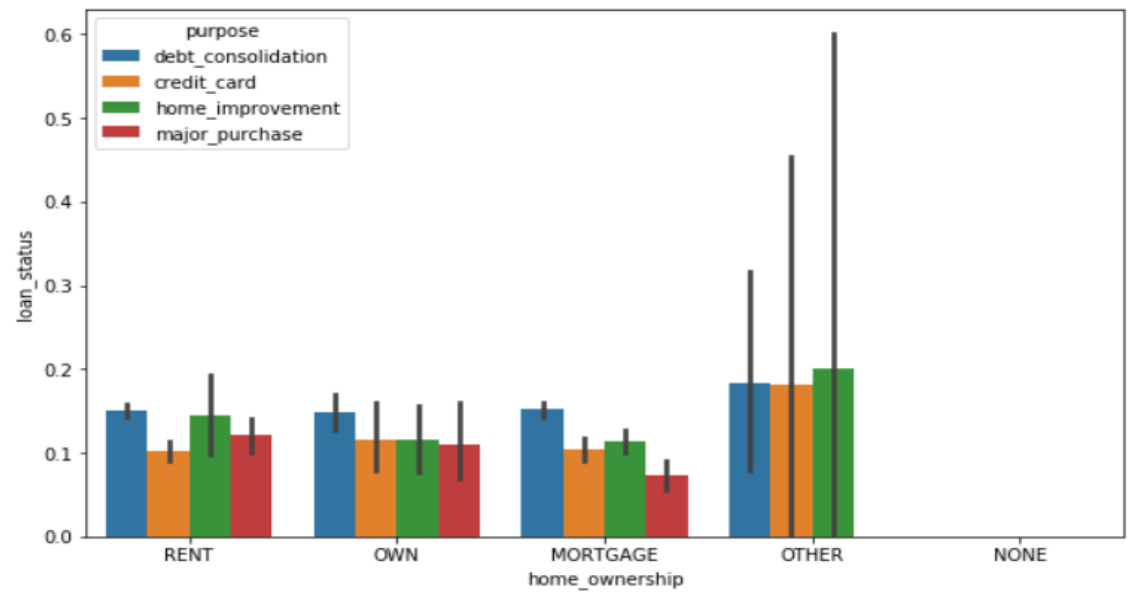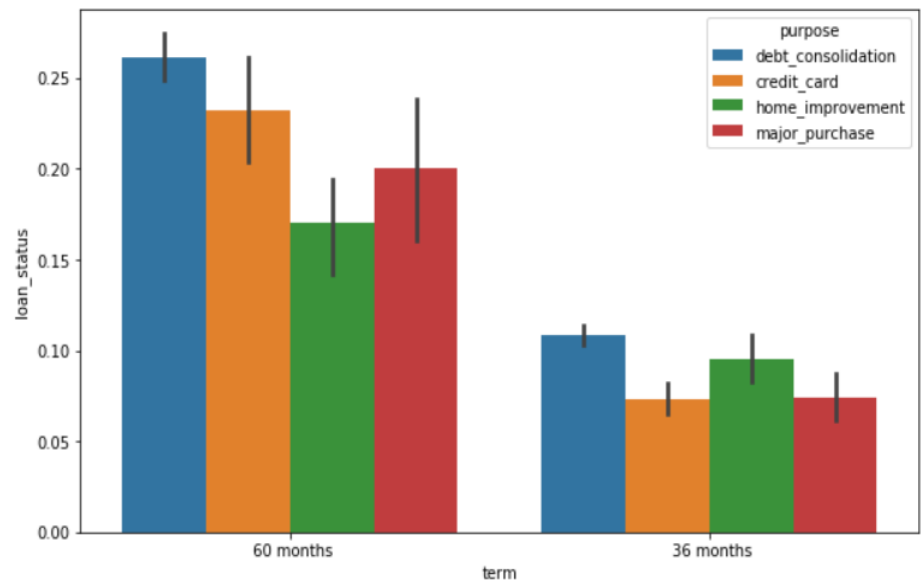
SEGMENTED ANALYSIS

Plotted graph of **top four purpose** of taking loan.
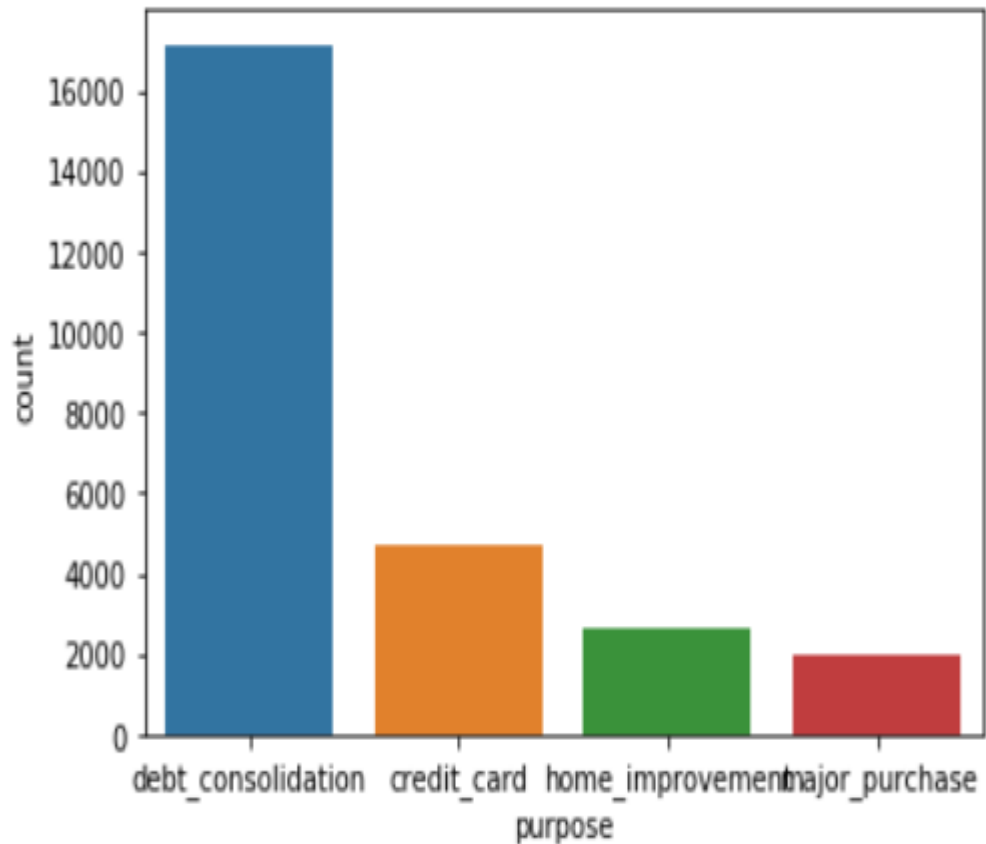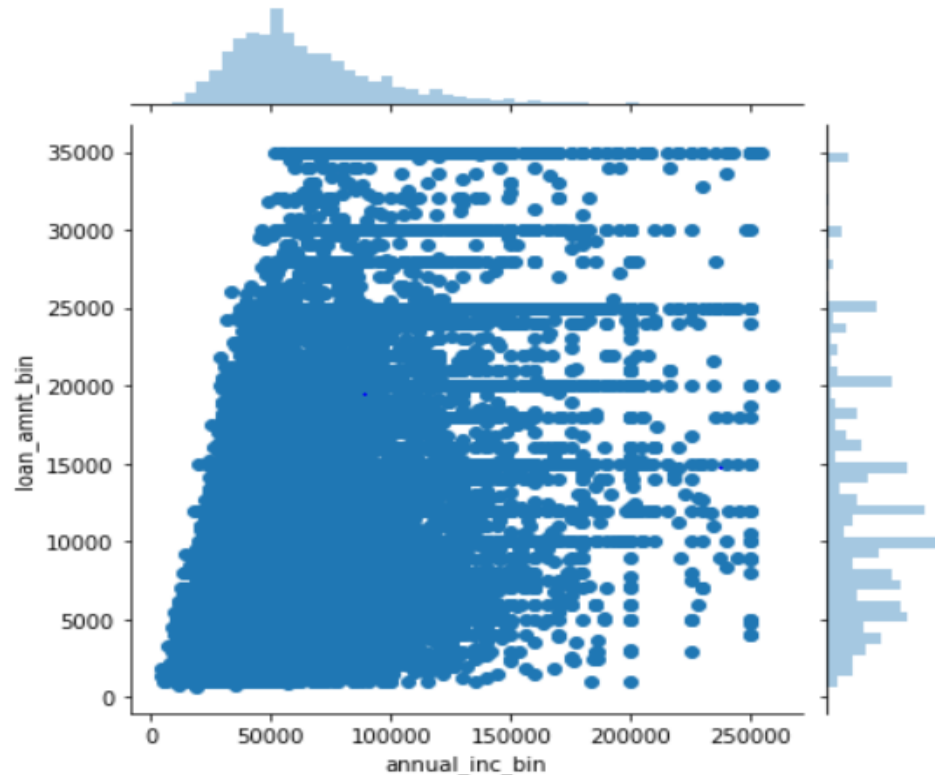

OBSERVATIONS –

- We can see that most of the loans are taken for debt consolidation, credit card, home improvement and major purchases.

We will segment the loan applications across the **purpose** of the loan, like- the **term, interest rate, home ownership** and **income.**
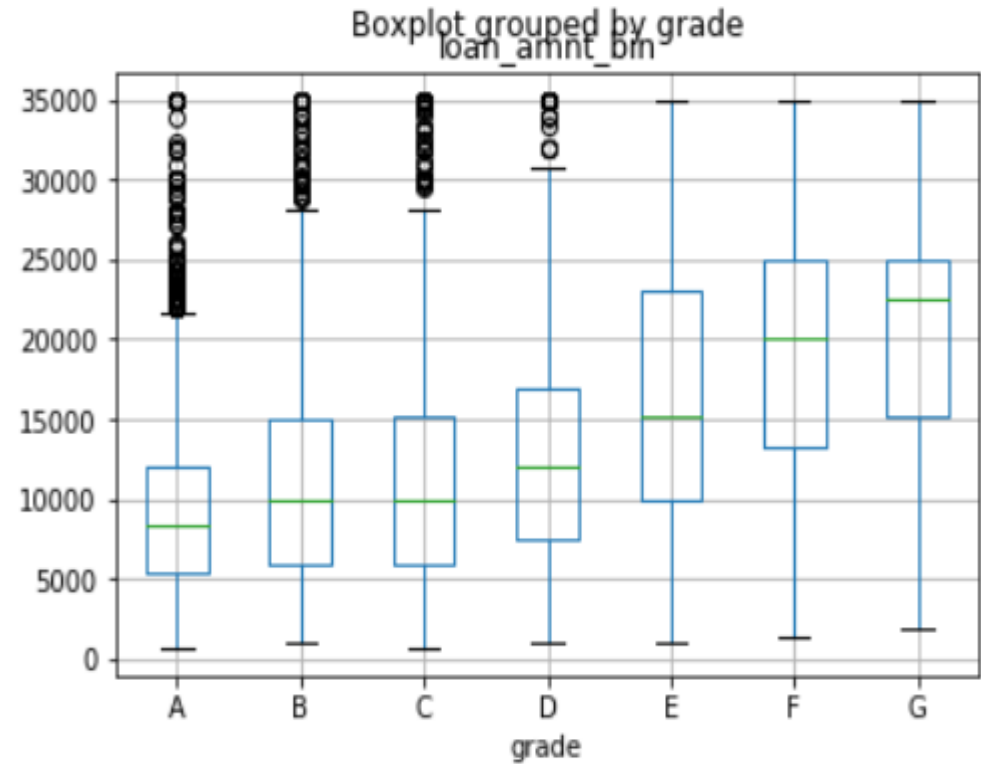
BIVARIATE ANALYSIS

Plotted graph between **loan amount** and **annual**

Plotted graph between **loan** and **grade**

OBSERVATIONS –
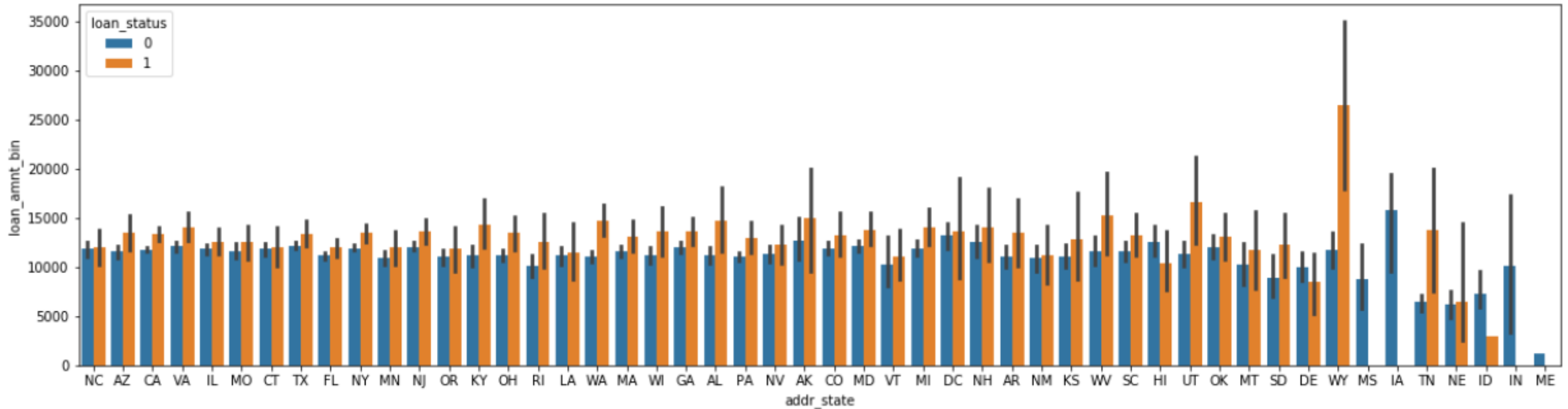- There are people with average income lower than 50000 taking loans of 25000 or higher. These would be risky loans.

OBSERVATIONS –
- Larger loans generally appear to be given to lower grade.
- With the median loan amount for a grade G loan being almost 10000 higher than that of a grade A, B, or C loan.

Conti..

Plotted graph between **loan_status** and **address** and **loan amount**

OBSERVATIONS –

- From above graph 0 indicates Charged Off and 1 indicates Fully Paid
- States WY: Wyoming has the highest average loan amount that was charged off.

CONCLUSION

After performing Univariate, Segment and Bivariate analysis we are now able to conclude the most affecting the defaulters (charged off) list. The top factors are –

- **Term of loan**
- **Interest Rate**
- **Debt to Income**
- **Purpose of loan**
- **Income of a person**
- **Instalment of loan**

In order to be saved from financial loss, we need to –
1. Keep the **term lower**.
2. **Avoid** giving loans for **small businesses**.
3. People having **more dti** are more likely to be charged off.
4. **Higher interest** rate can cause **high default rate**.
5. It is risky to grant higher loans for people with **less than 50000 annual income**.
6. People **living in WY** are more likely to be defaulter, therefore risky to grant them loan.
7. The chances **of US army and bank of America** being a defaulter is higher than other companies.
8. The chances of **higher loan amounts** with **higher instalments** getting charged off is high.

THANK YOU