LENDING CLUB ASSIGNMENT

SUBMISSION

Submitted by:

Anish Palan

Sudhasrinivas Pallam

Lending Club Case Study Analysis

The Lending club specialises in lending loans to urban customers. They want to understand the risk of giving a loan to an applicant based on the factors which contribute to a loan default. This information is utilised for portfolio and risk assessment.

Business understanding

- 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.

- Company can either accept a loan in which case it can be either fully paid, current or charged off, or they can reject a loan
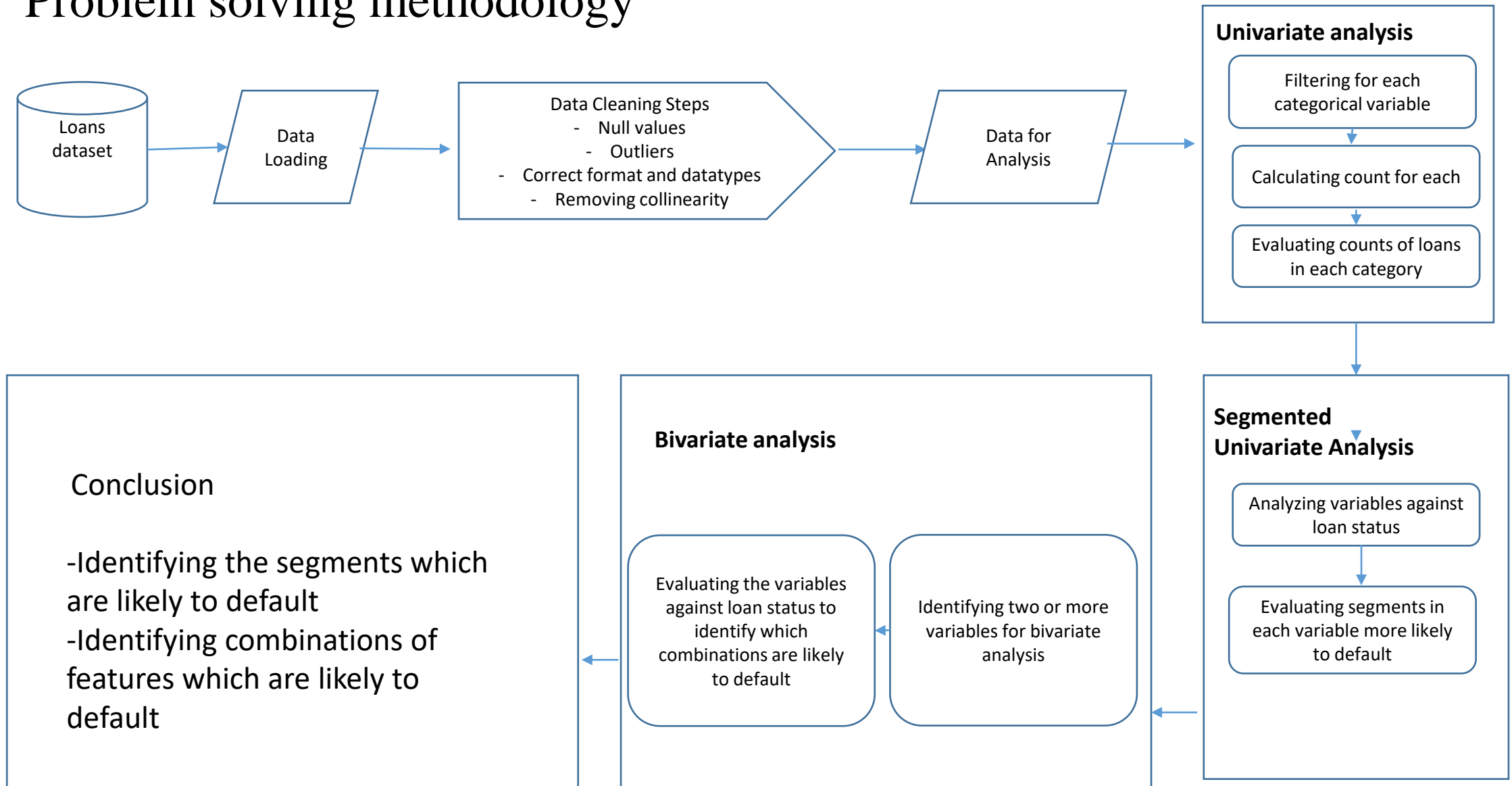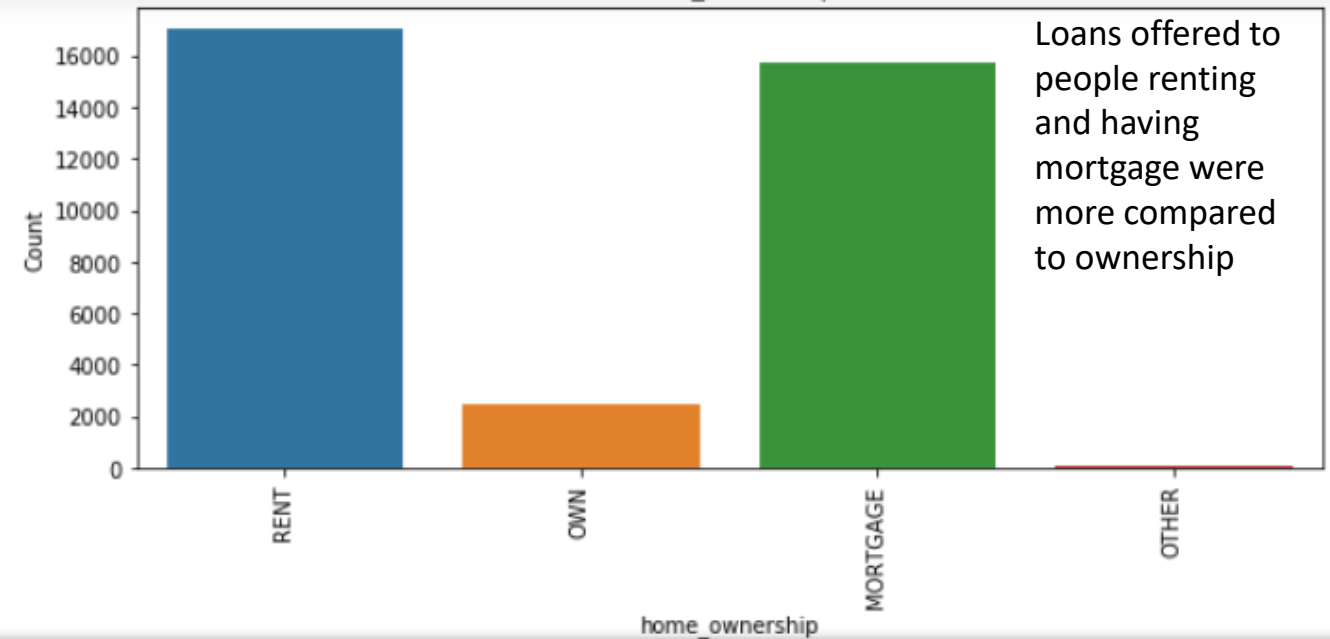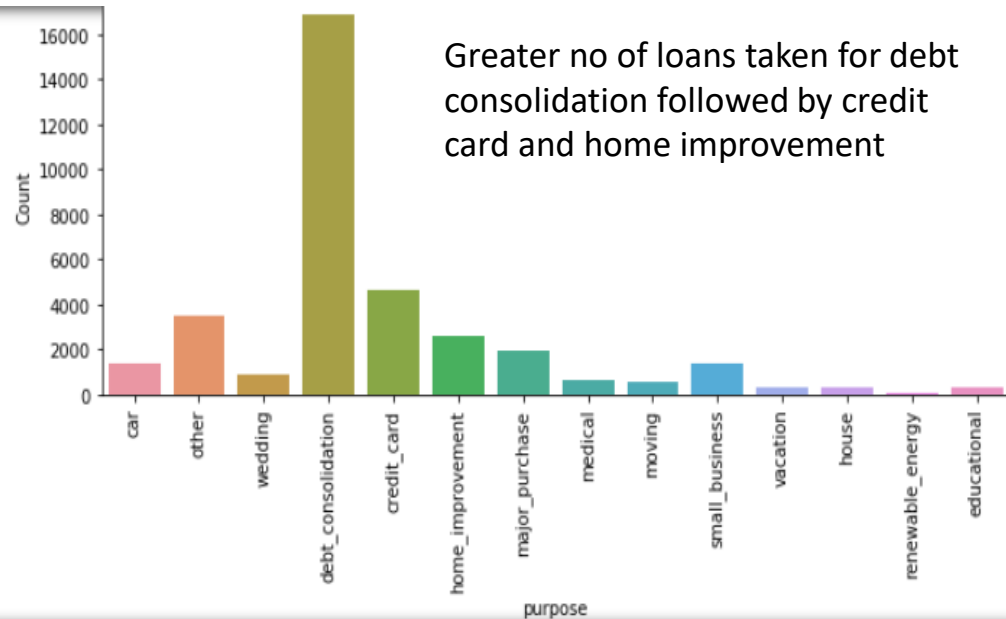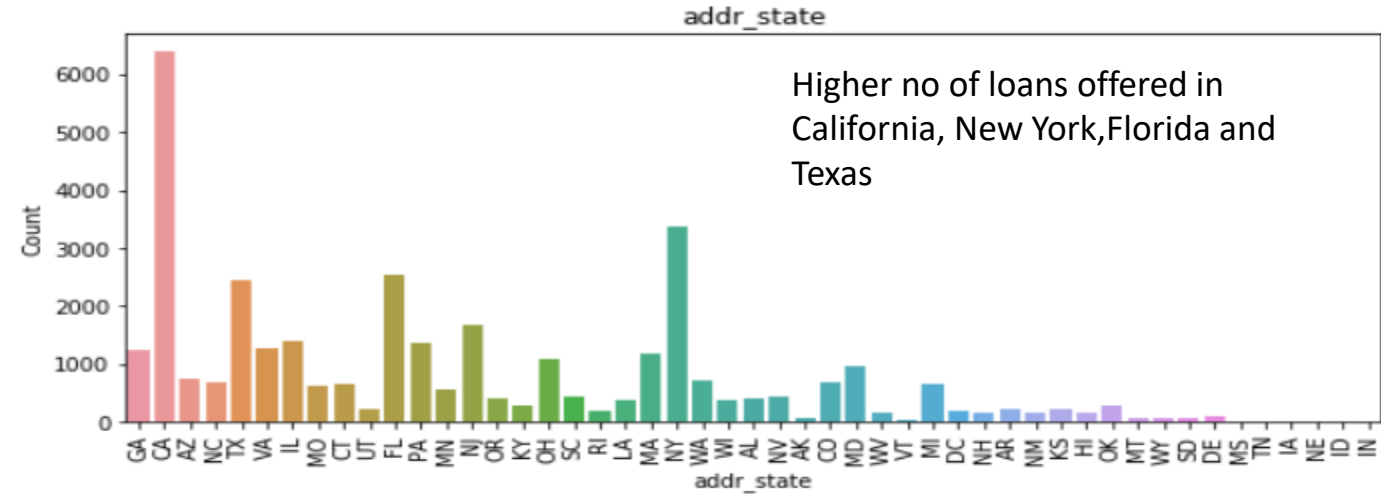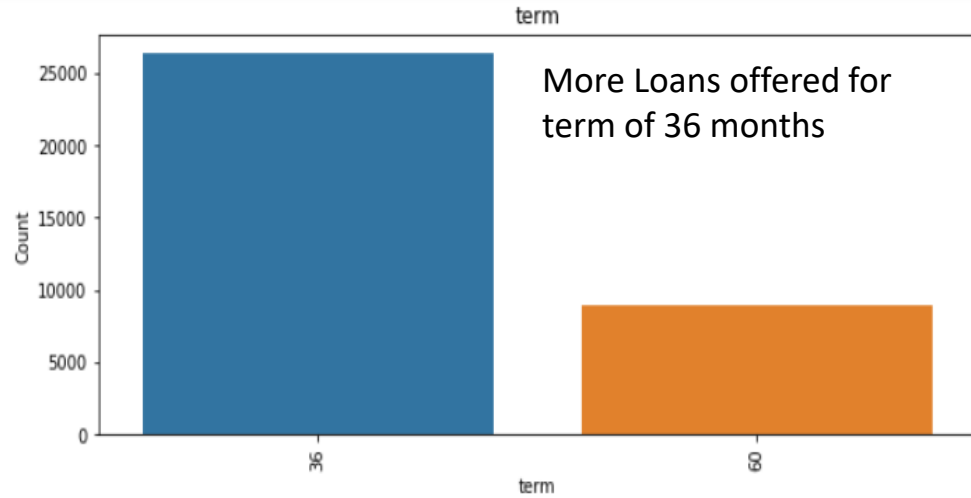
Business objective

- Lending loans to 'risky' applicants is the largest source of financial loss (called credit loss)

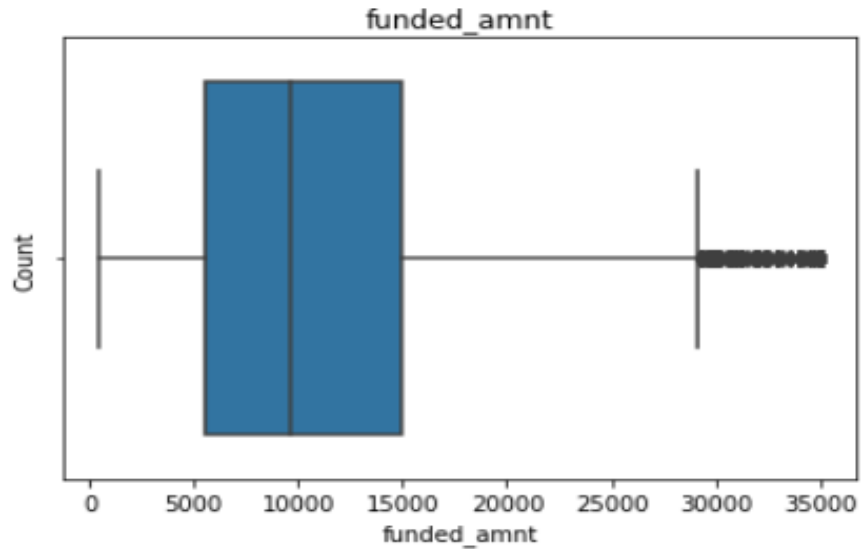- Identification of the risky loan applicants using the driver variables behind the loan default.

Goal of Data Analysis

- Univariate analysis

- Segmented Univariate analysis

- Bivariate analysis

Problem solving methodology

Loans dataset → Data Loading → Data Cleaning Steps
- Null values
- Outliers
- Correct format and datatypes
- Removing collinearity
→ Data for Analysis →

Univariate analysis
- Filtering for each categorical variable
- Calculating count for each
- Evaluating counts of loans in each category

Segmented Univariate Analysis
- Analyzing variables against loan status
- Evaluating segments in each variable more likely to default

Bivariate analysis
- Evaluating the variables against loan status to identify which combinations are likely to default
- Identifying two or more variables for bivariate analysis

Conclusion

-Identifying the segments which are likely to default
-Identifying combinations of features which are likely to default

Univariate Analysis
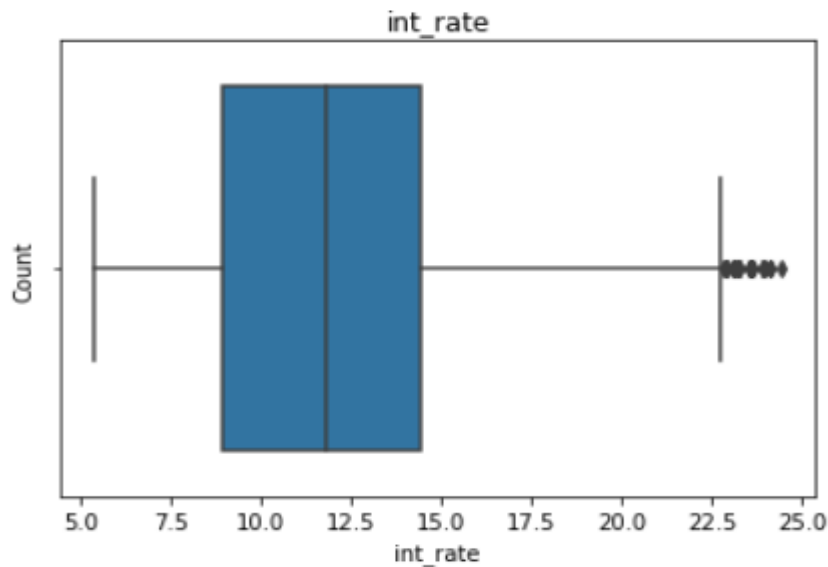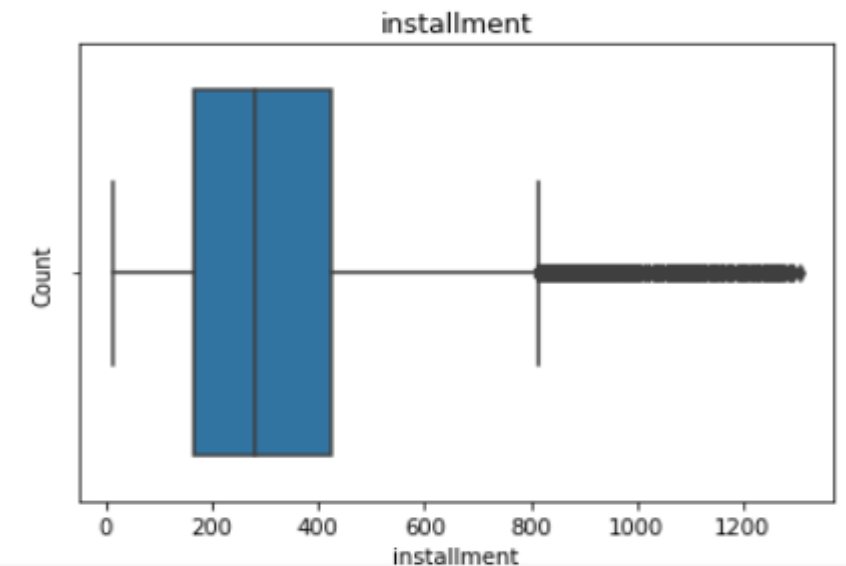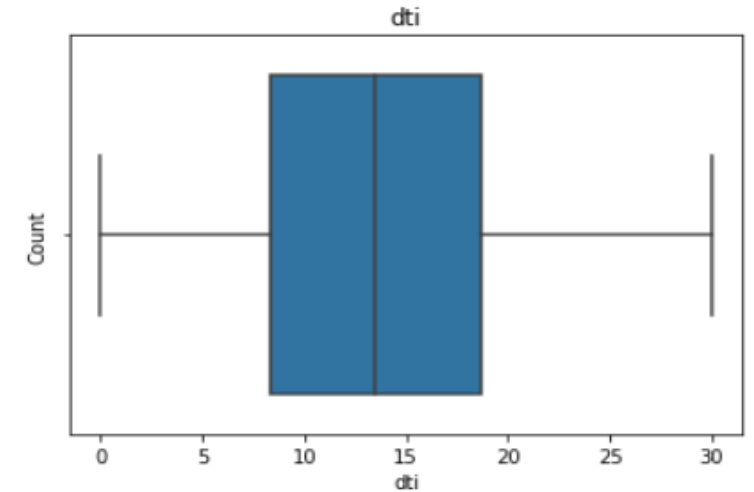
funded_amnt

int_rate

dti

installment

Funded Amount was more towards value of 10000, and varying from 5000 to 15000
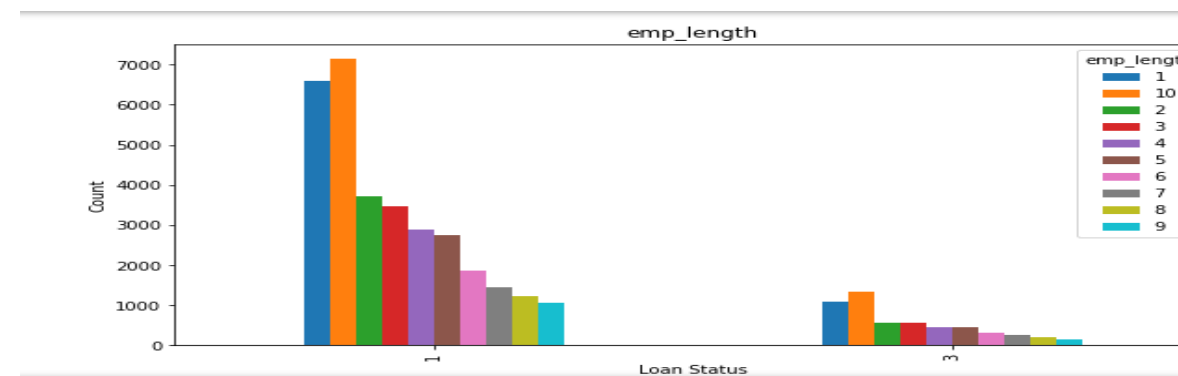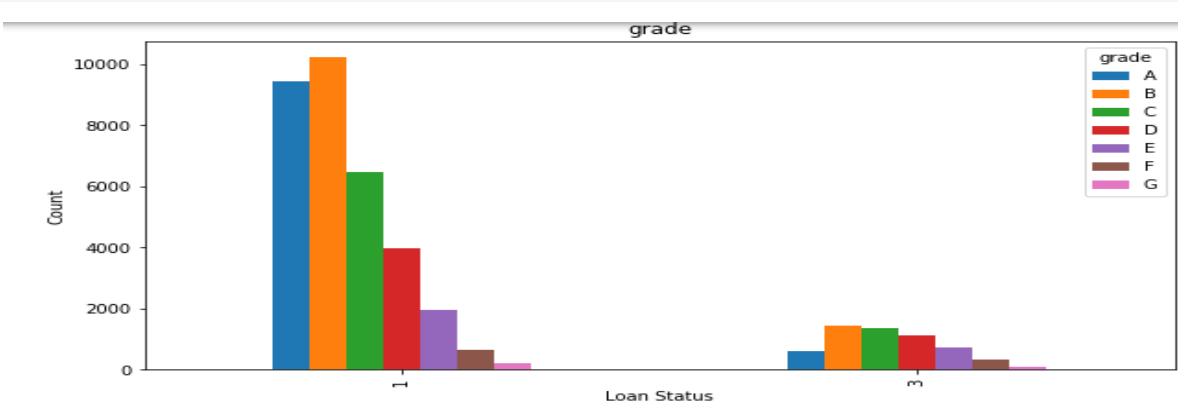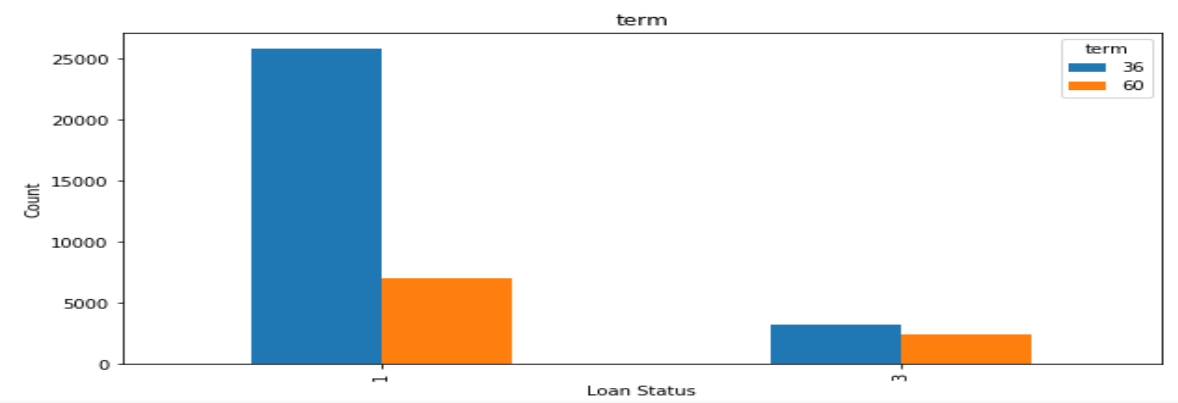
Interest rate varied from 8% to 15% with more at around 11%

Debt to Income ratio varied from 8 to 18 being more on 14

Installments varied between 200 to 400+ with more number being at around 250-270
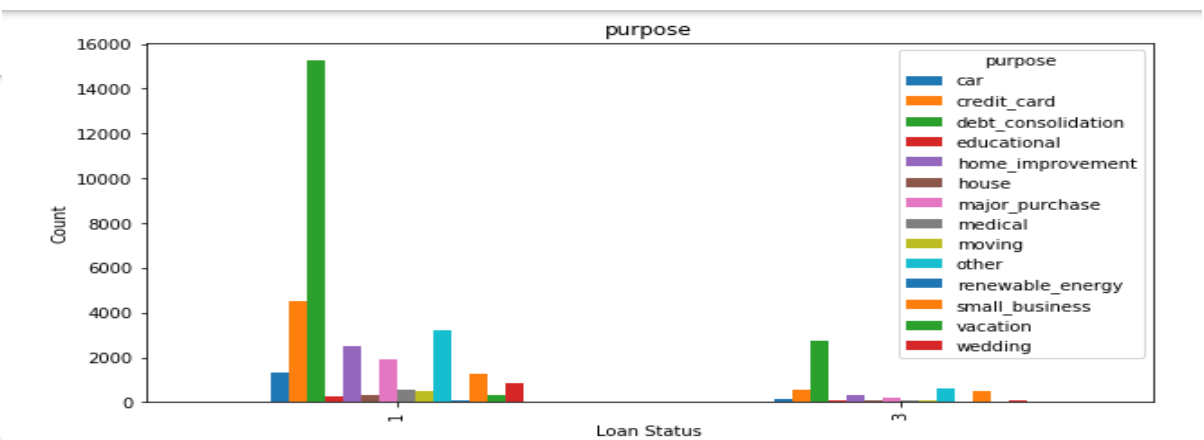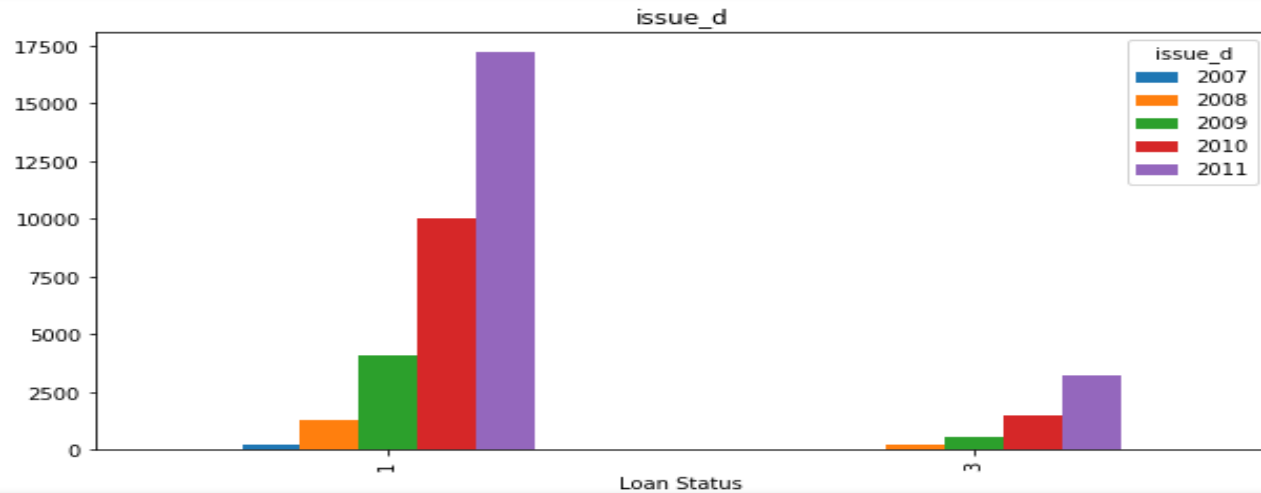
Segmented univariate Analysis

Loans given for 36 months tend to default more compared to 60 months

Grades B, C and D were the top defaulters

People with employment length of >=10 years tend to default more followed by those with <=1 year
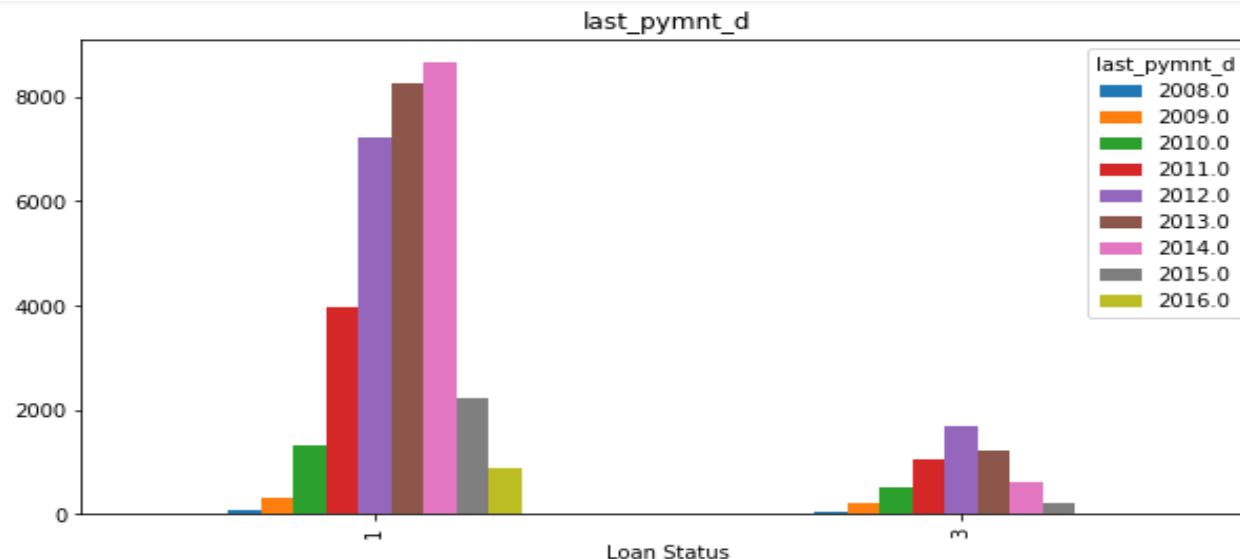
Debt consolidation turned out to be top purpose for default followed by credit card, and small businesses

More number of default loans observed on loans issues in the year 2011.

It shows a trend that as years increased, the default loans also increased over the years

Loans having last payment date of 2012 tend to default more followed by those in 2013 and 2011

It also shows that these years as the last payment was received tend to show more defaulters compared to other years
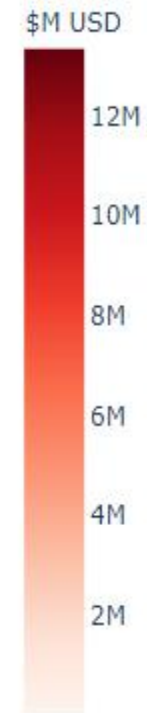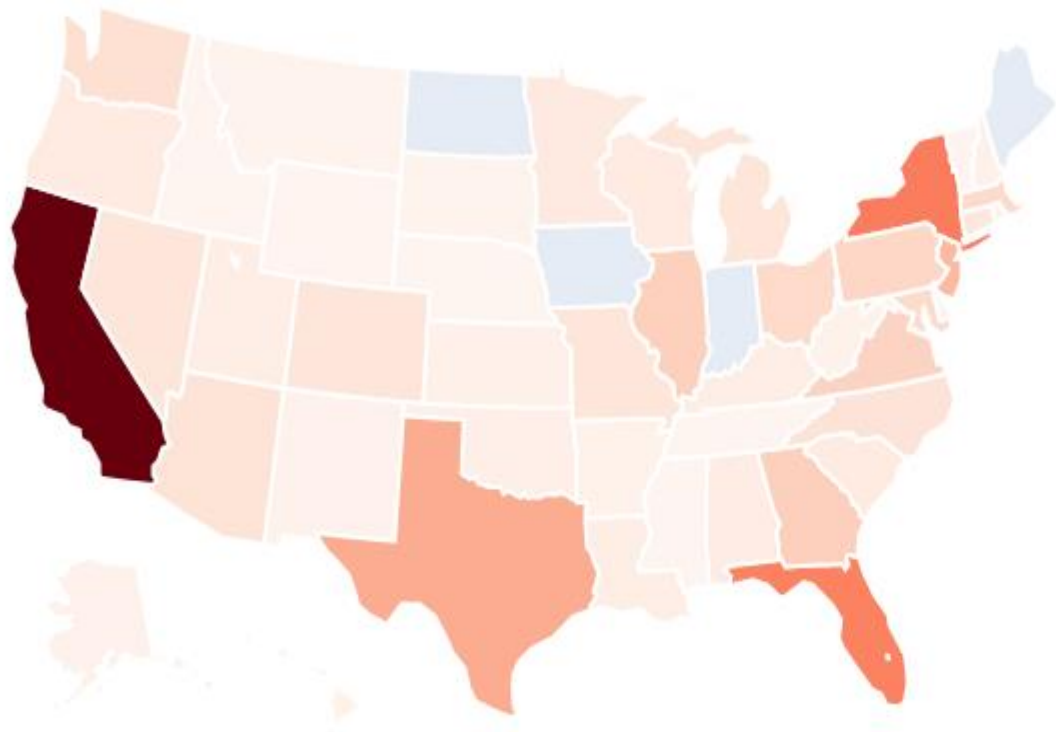
CA, NY, FL, TX were the top state with defaulters in the respective order

NY has lowest DTI while TX has the highest CA has highest avg annual income followed by TX

Interest rates are more in TX and NY

State	Interest Rate	Average Annual inc	Average DTI
CA	13.77	$65781	13.51
NY	13.85	$65383	13.27
FL	13.52	$60895	13.65
TX	14.3	$65545	14.09

Term Analysis

1. With 36 month term with exp <4 years and 60 month term with >10 years exp more likely to default
2. 36 month term is likely to default with Other home ownership type followed by Renting type
3. 36 month term is more likely to default when annual income is Low (<40000)
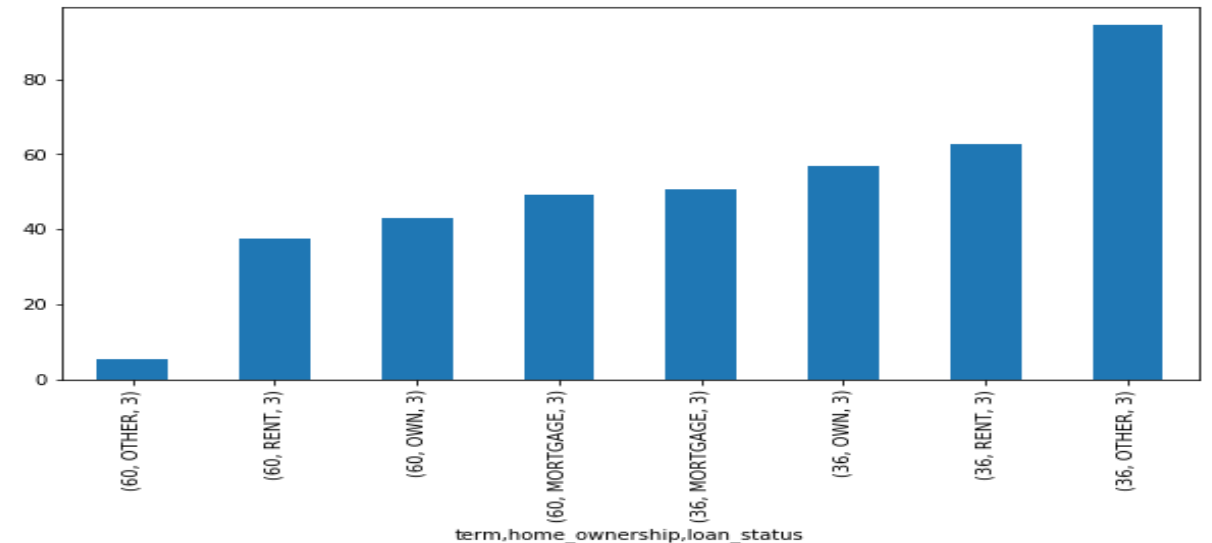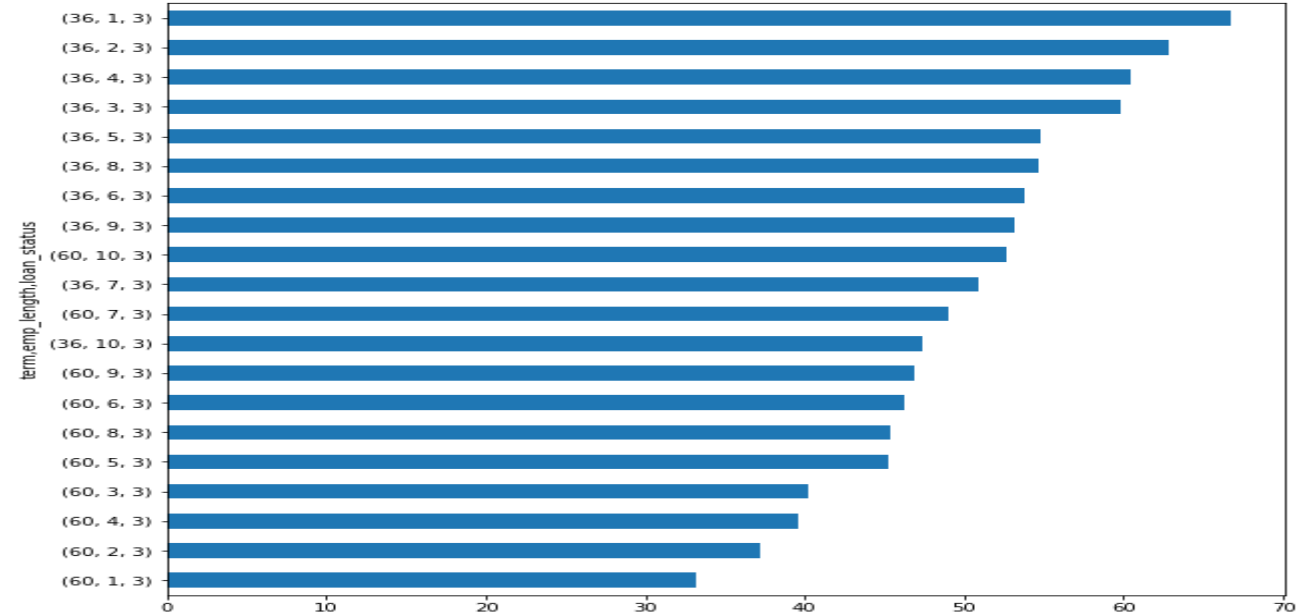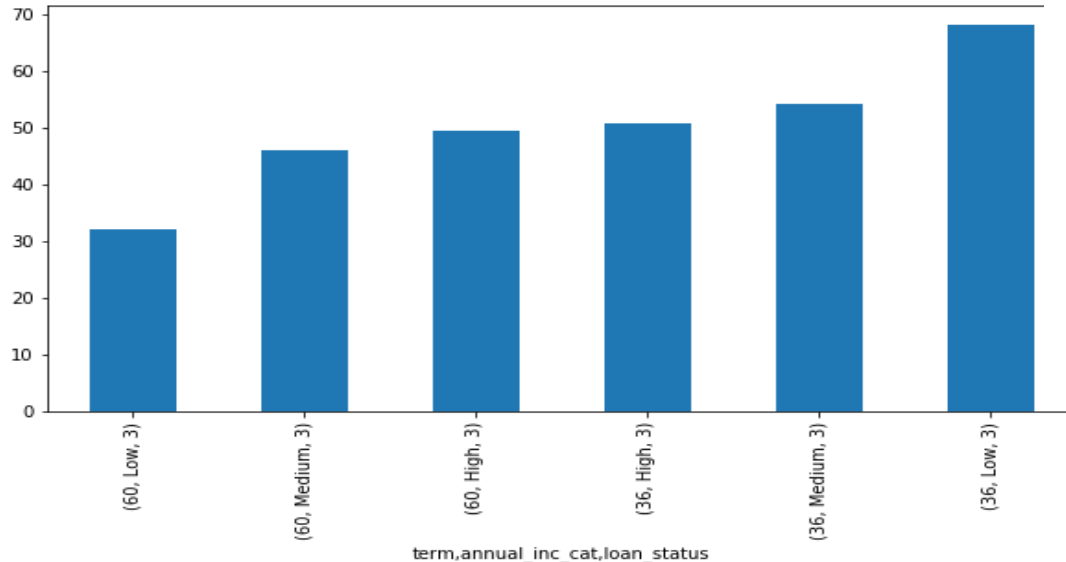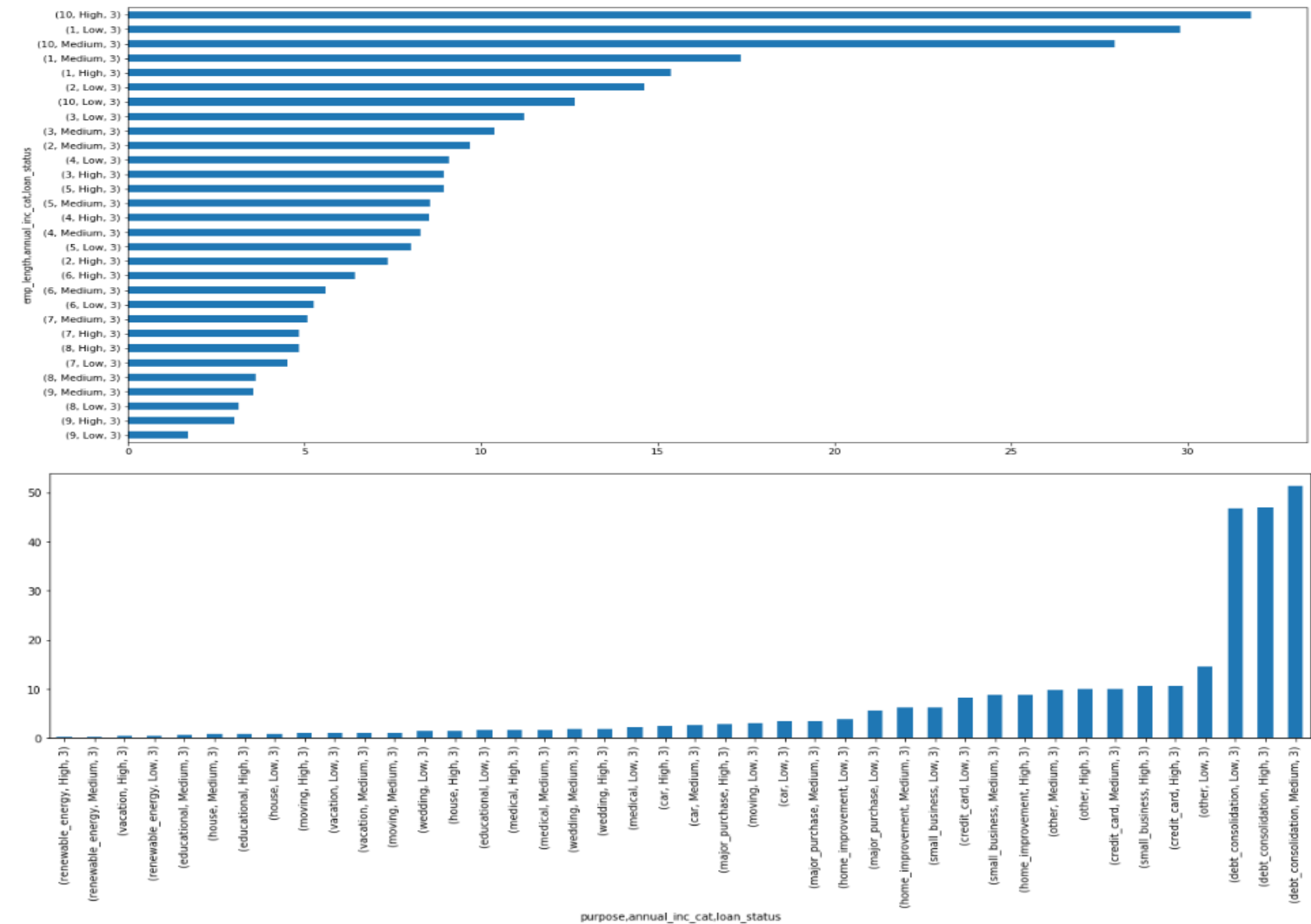
Annual Income analysis

1. With High or Medium income >=10years exp is more likely to default
2. Low income with <=1 year exp with medium 1 year exp is likely to default
3. With any range of income debt consolidation purpose category emerges as the top defaulter
4. Credit card, home improvement and small business comes next usually when Annual income is High to Medium

Funding Amount Analysis

1. More funded amount was defaulted in small business followed by debt consolidation and house type of loans
2. Mortgage and other home ownership were more defaulted.
3. For Grades F, G and E there was more default in funded amount
4. Defaulted Loans for terms 60 month were the most funded ie most money lost there
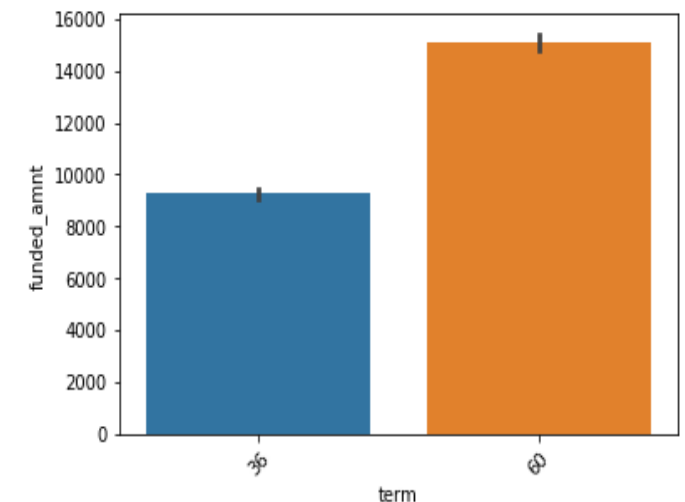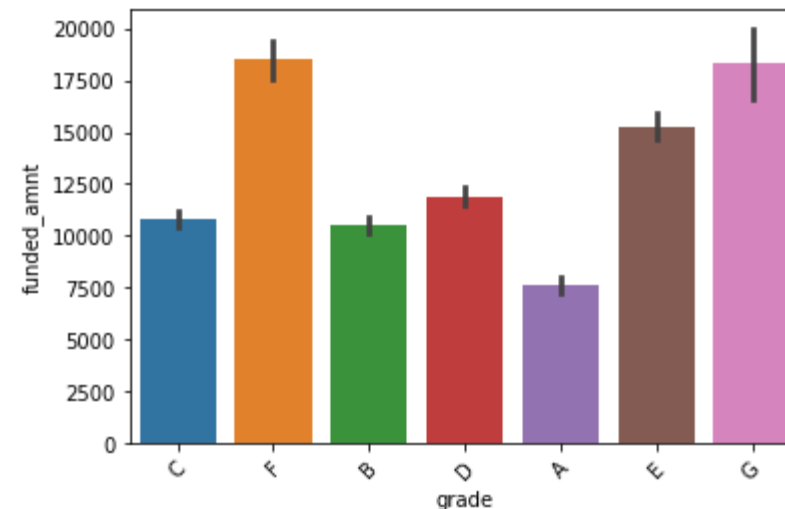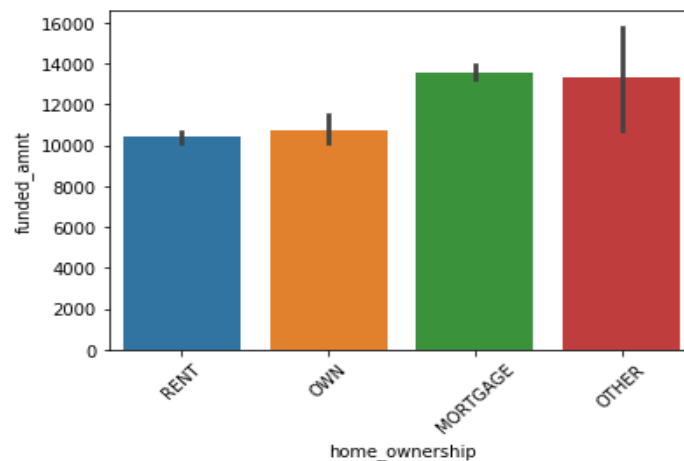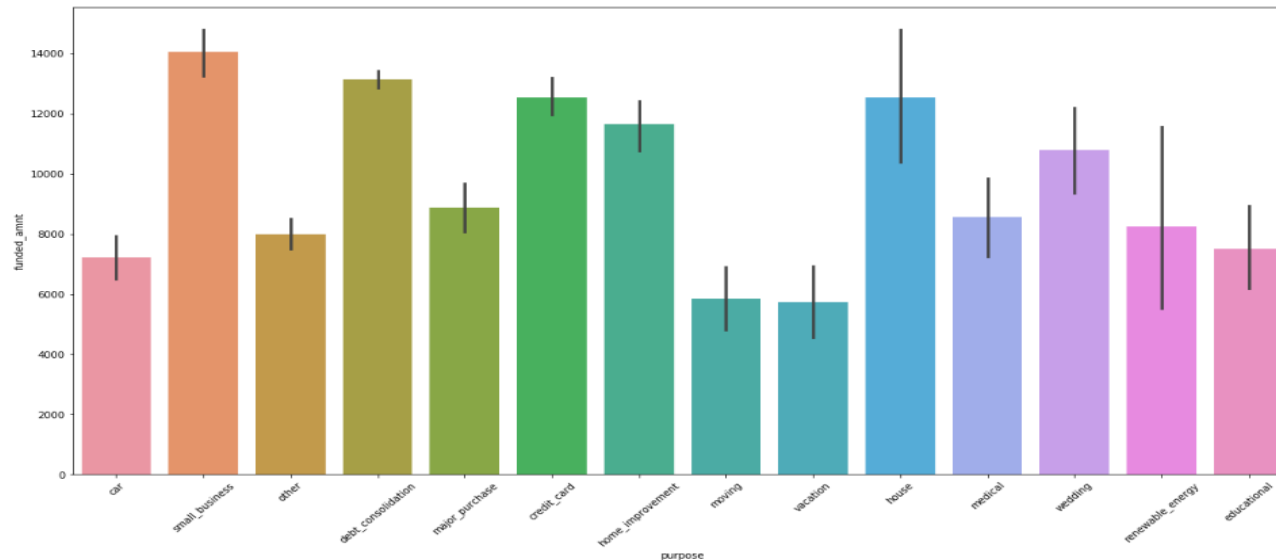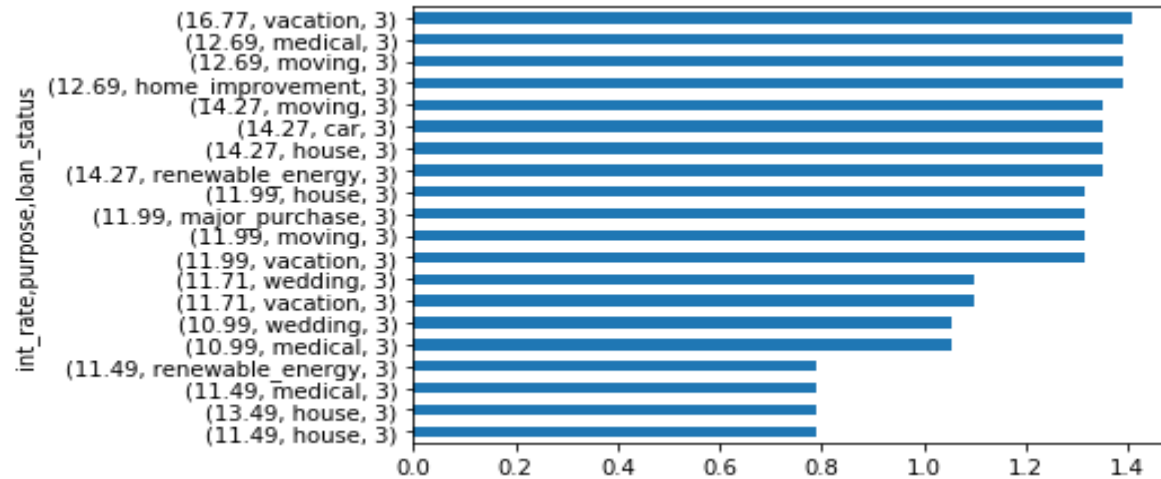
Interest rate vs Purpose Analysis

For most purposes, interest rate >=11% amount
for the most number of defaulters
Vacation tops the list of purpose with highest
interest rate for default loans

DTI vs Open Accounts Analysis

With High DTI and open accounts in range 31 to
34, and Medium with 27 open accounts they
are more likely to default
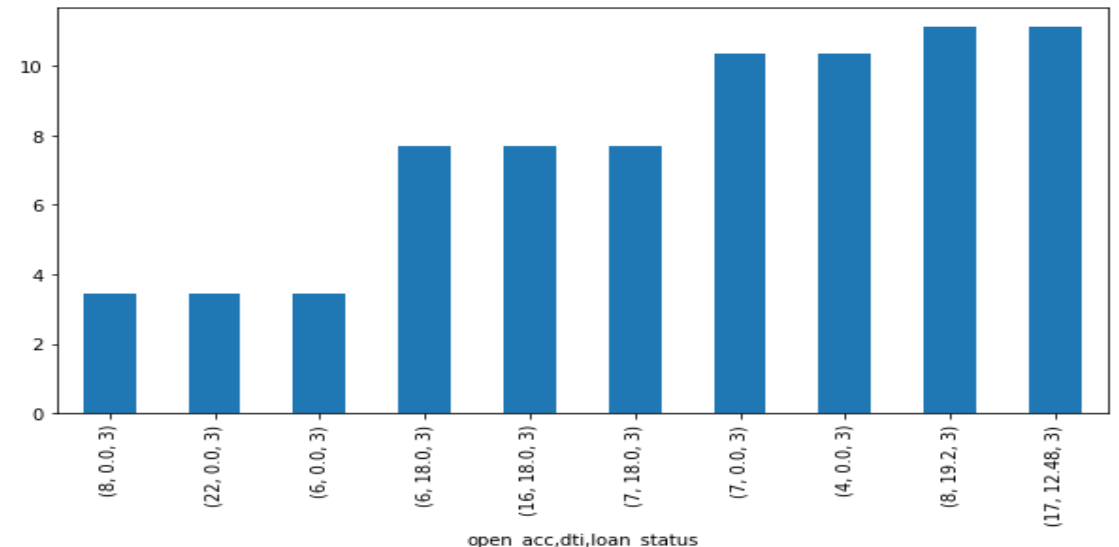
Based on the Analysis of the Lending club below are some observations of the loans which are being defaulted

1. With 36 term with exp <4 years and 60 term with >10 years exp more likely to default

2. Term 36 and Other is more likely to default

3. 36 and Low income more likely to default

4. With any range of income debt_consolidation comes as the highest defaulter

5. With High or Medium income >10years exp or Low income with <=1 year exp with medium 1 year exp is likely to default

6. With experience >7 years, and loan purpose like educational or wedding are likely to default

7. For medium DTI, with experience in range 6 to 8 years likely to default

8. For High DTI, with 7 years of experience is likely to default

9. Educational, vacation and wedding seems to default more across the DTI

10. With High DTI and open accounts in range 31 to 34, and Medium with 27 open accounts they are more likely to default

11. More funded amount was defaulted in small business followed by debt consolidation and house

12. Mortgage and other home ownership were more defaulted

13. For Grades F, G and E there was more default in funded amount

14. Loans offered for term 60 months caused more default to funded amount

15. For most purposes, interest rate >=11% amount for the most number of defaulters