LENDING CLUB CASE STUDY

SUBMISSION

Group Members: Aditya Bhave, Sivaprasad

Objective

The Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. The company wants to understand the driving factors (or driver variables) behind loan default

- **Approach:**
  1. Lending club has provided history of existing loans
  2. The loans were lended between year 2007 to 2011
  3. We analyzed the data using Exploratory Data Analysis (EDA) method

- **Goal:**
  1. This presentation explains the driving factors behind loan defaults

- **Assumptions:**
  1. We have considered the loan_status as provided in the data.
  2. We have ignored the loans which were ongoing.
  3. Some columns were cleaned based on analysis and relevance to our goal.

Data understanding

Following are the summary of the analysis:-

- Data was provided in 111 columns.
- The number of observations were 39717 (i.e. number of loans in data)
- Dataset was a mixture of Categorical and Continuous columns (variables)
- Most of the columns had no values, or same values for all records. Such columns were deleted.
- Loan Status column indicated the status of the loan. Defaulter was identified where Loan Status was Charged Off

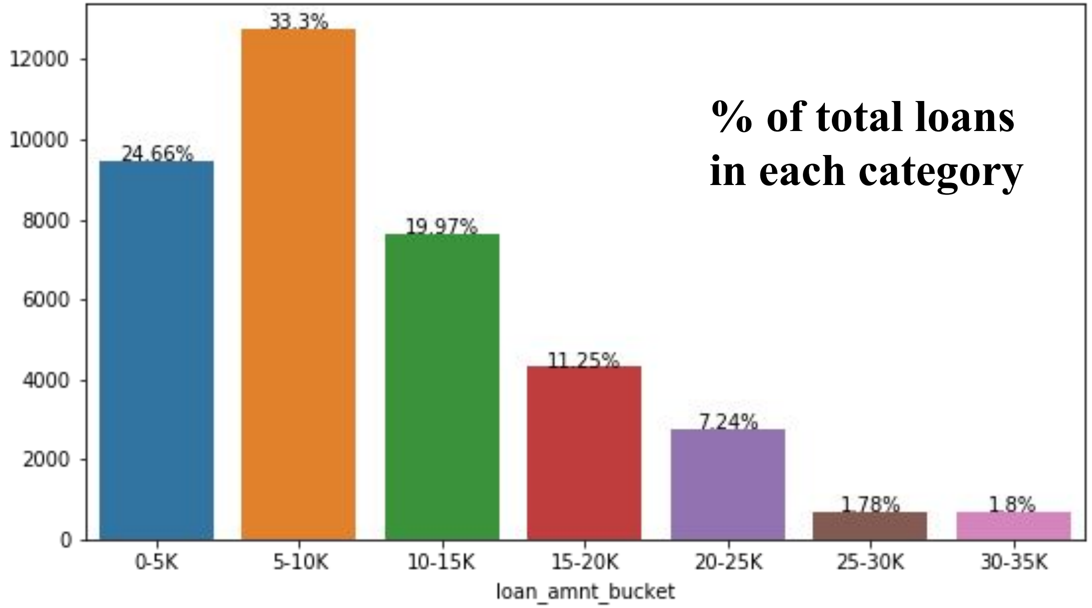
Data Cleaning and Manipulation

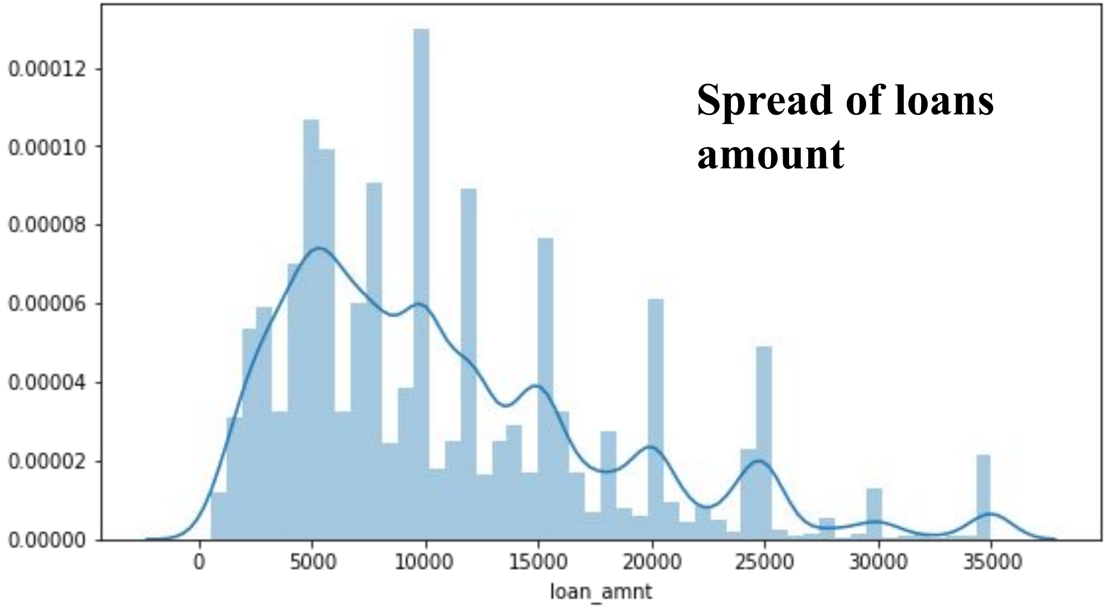
- Columns with only null values were removed from the dataset before analysis.
- Columns with more than 60% null values were removed as it was not feasible to impute the data.
- Columns with same value for the all the rows were removed as those couldn’t be used to derive any significant trend.
- Columns with redundant data or similar trend were removed as it might not give any significant information.
- Data types of some of the columns were changed to suit the analysis.

In the following slides, we will analyze each important variable separately

Loan Amount

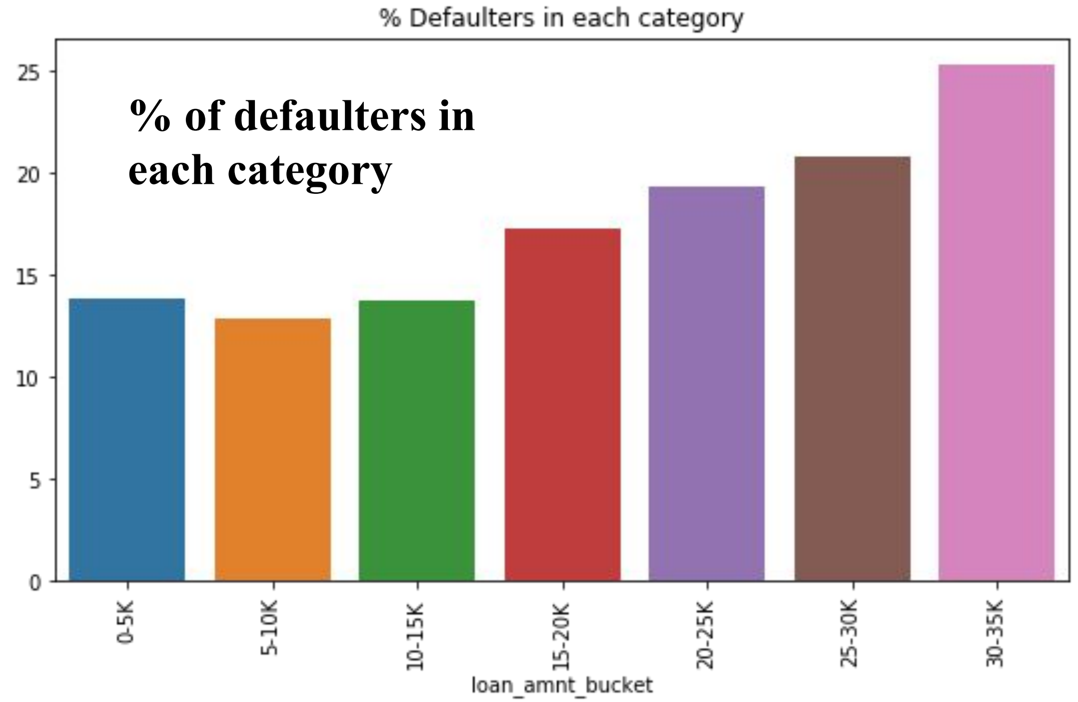
Definition : The listed amount of the loan applied for by the borrower. If at some point in time, the credit department reduces the loan amount, then it will be reflected in this value.





- Almost 75% of the loans were below 15,000
- Most of the loans (32%) fall in the range 5-10K.
- Loans of rounded figures are preferred such as 5000, 10000, 12000, 15000, 20000, 25000, 30000 and 35000.

Loan Amount (Contd.)

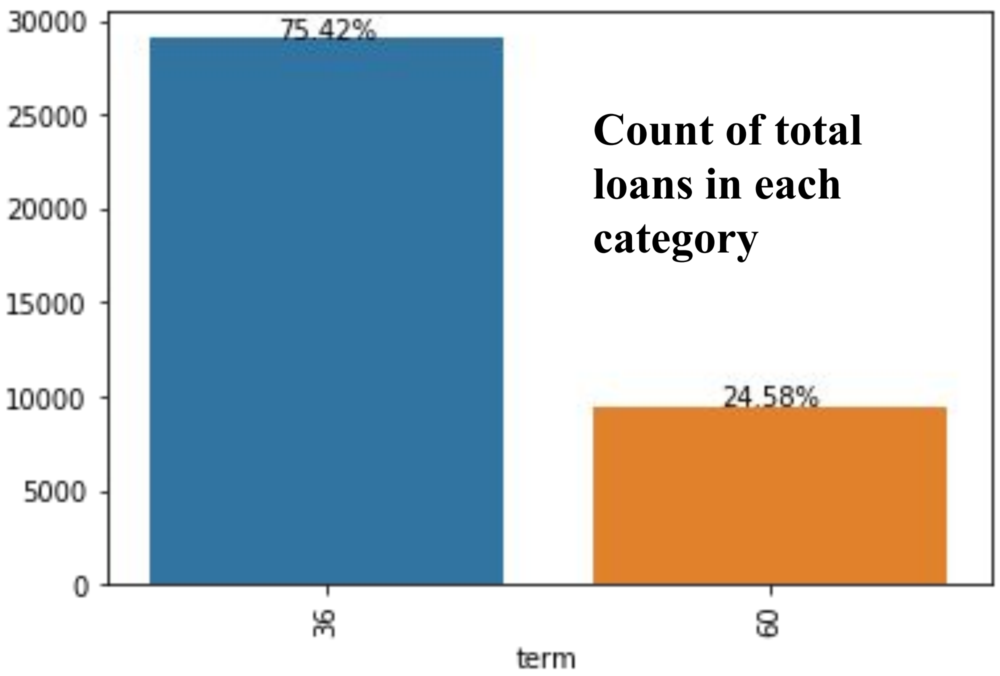


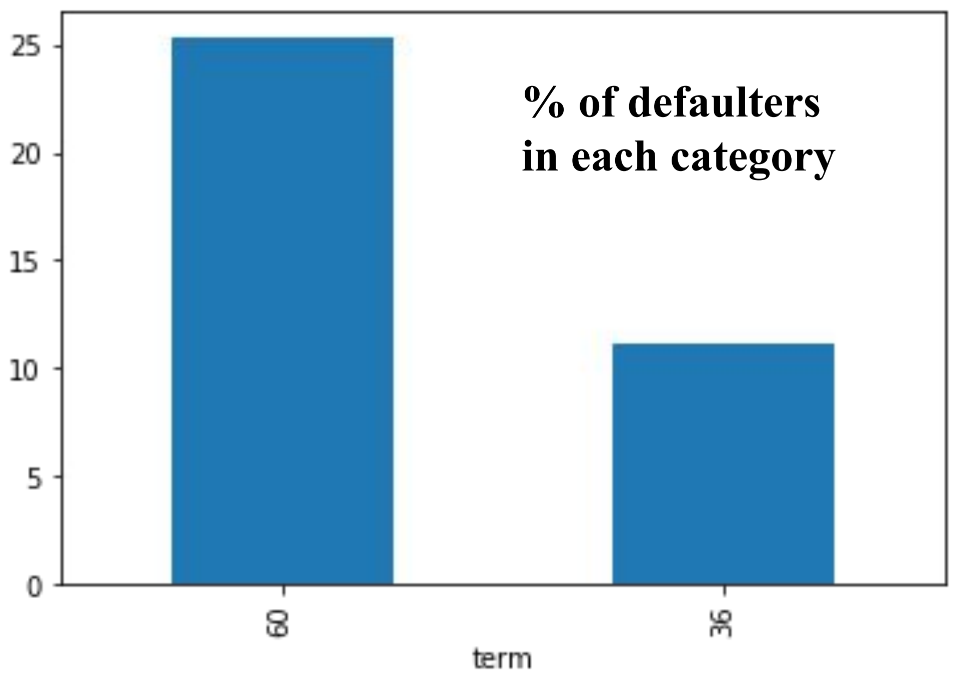
Loan Amount is a driving factor behind loan default

1. Defaulters Percentage graph shows that as the loan amount increases, the chance of defaulting increases
2. Risk of loans increases above 20% as loan amount increases above 20k
3. Smaller loans are comparatively safer loans which are also 50% of the business (as seen in previous slide)

Term

Definition : The number of payments on the loan. Values are in months and can be either 36 or 60.



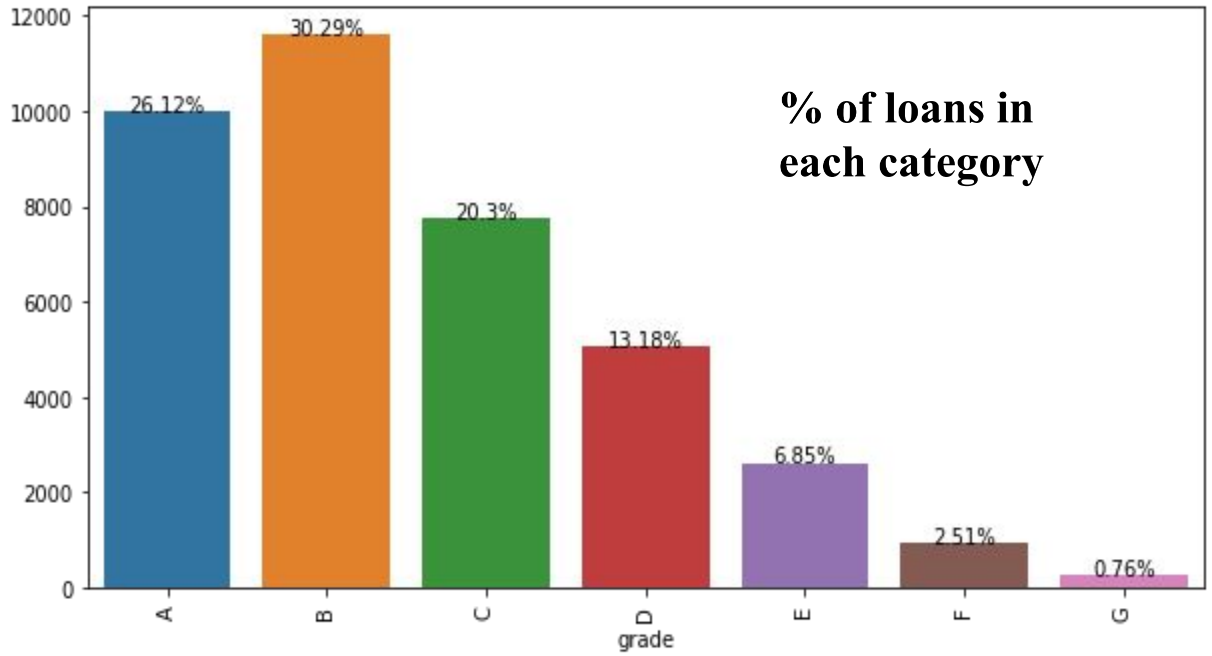


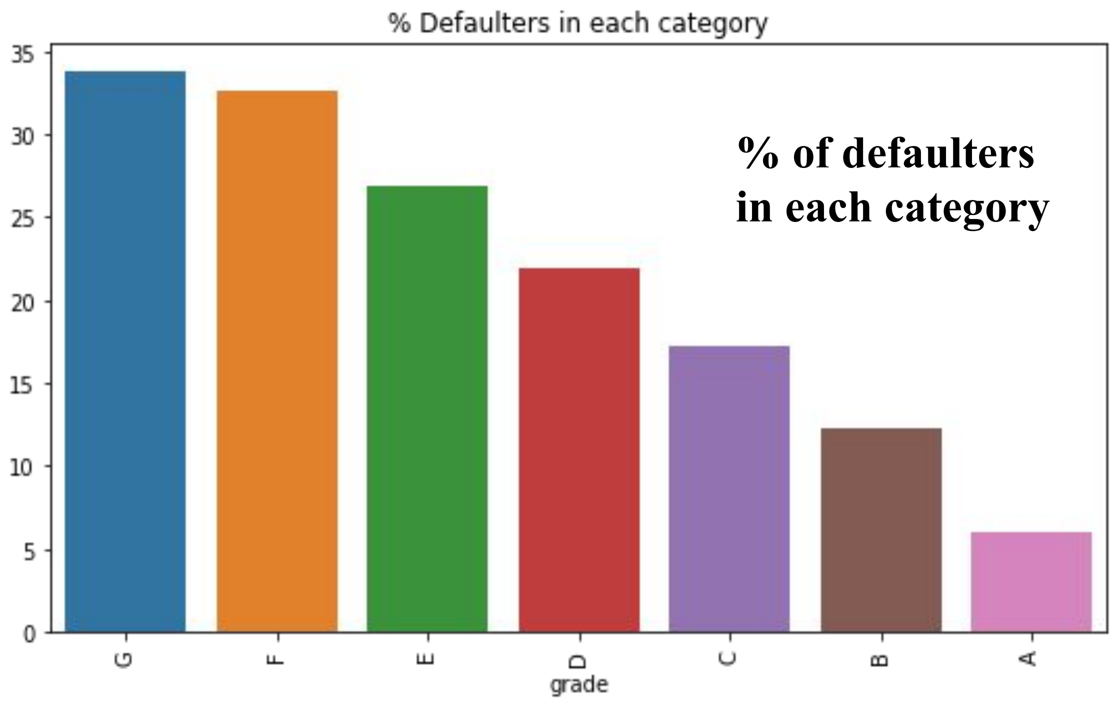
Term is a driving factor behind loan default

1. Long term loans are riskier
2. Almost 25% of total loans were long term, 60 months
3. Out of which 25% of long term loans were Charged Off, so it's evident that long term loans are riskier

Grade

Definition : LC assigned loan grade



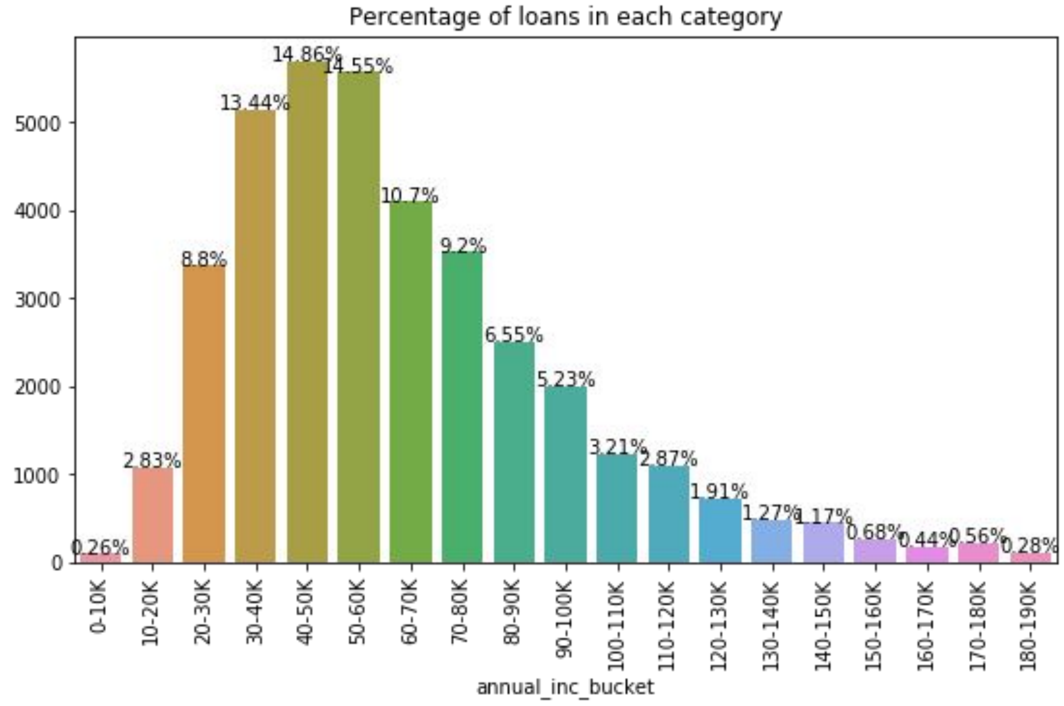


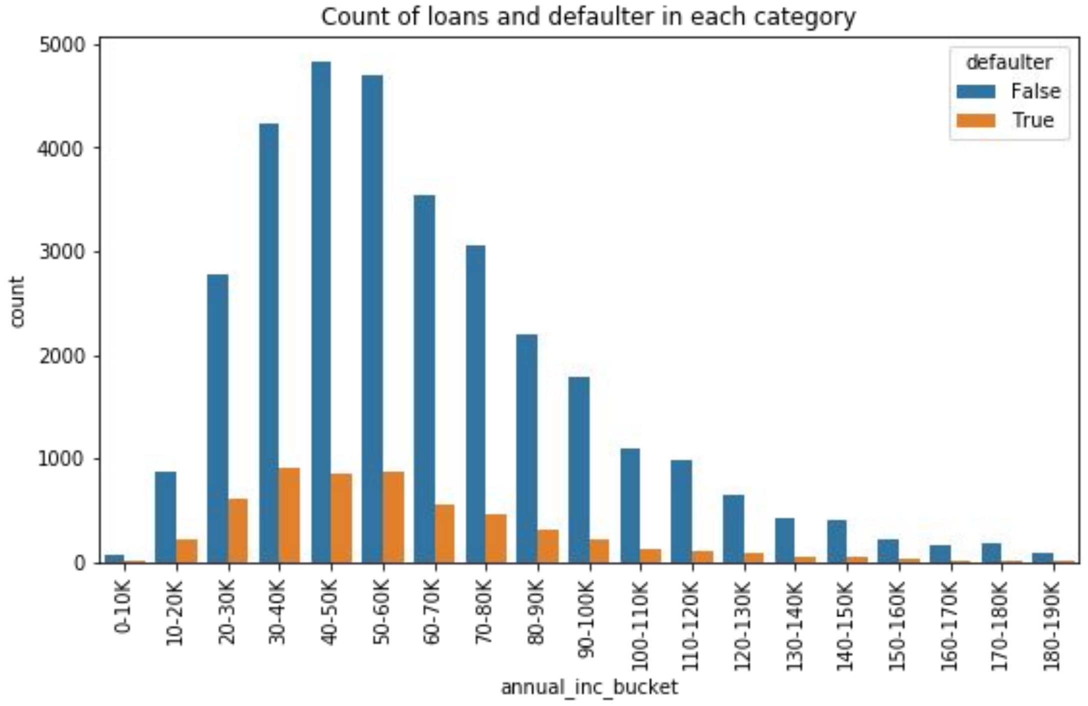
Grade is a driving factor behind loan default

1. Top 3 popular loan grades are B, A and C. Almost 30% of the loans are grade B
2. Lower grade loans such as G, F, E, D are riskier
3. D Grade loans carries a risk of 23% default and it increases as loan degrades with 34% risk for G grade

Annual Income

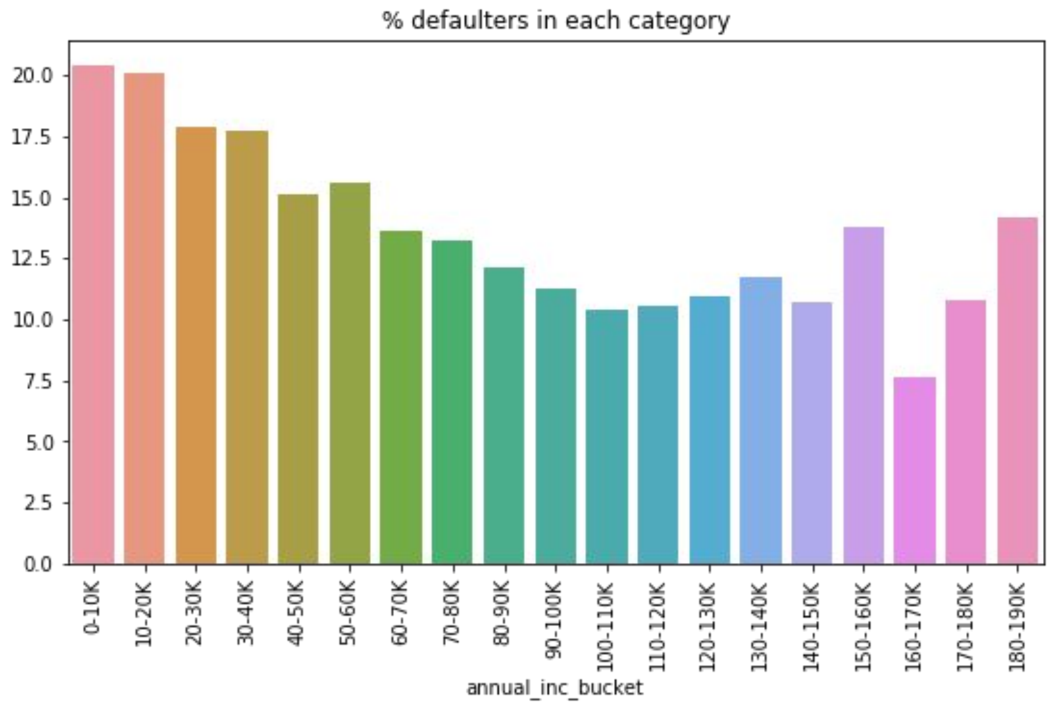
The self-reported annual income provided by the borrower during registration

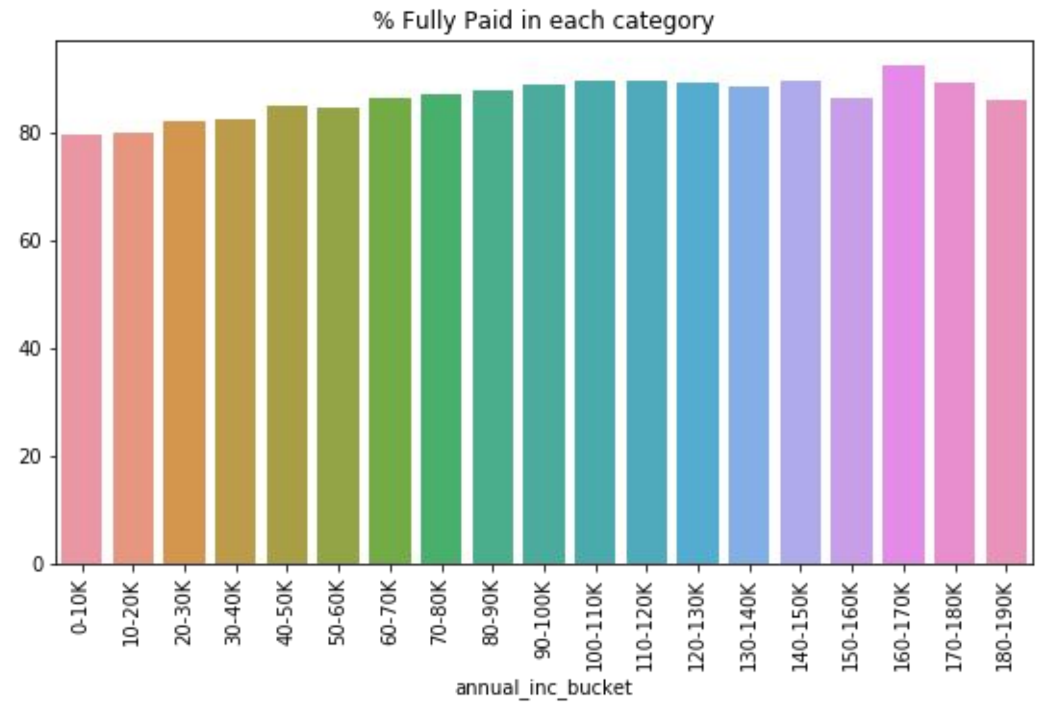




- The number of loans issued to people with lower (0-10k) and higher (150k+) income were less.
- 75% of the borrowers had an annual income equal to or lower than \$10k- \$80k.

Annual Income (Contd..)



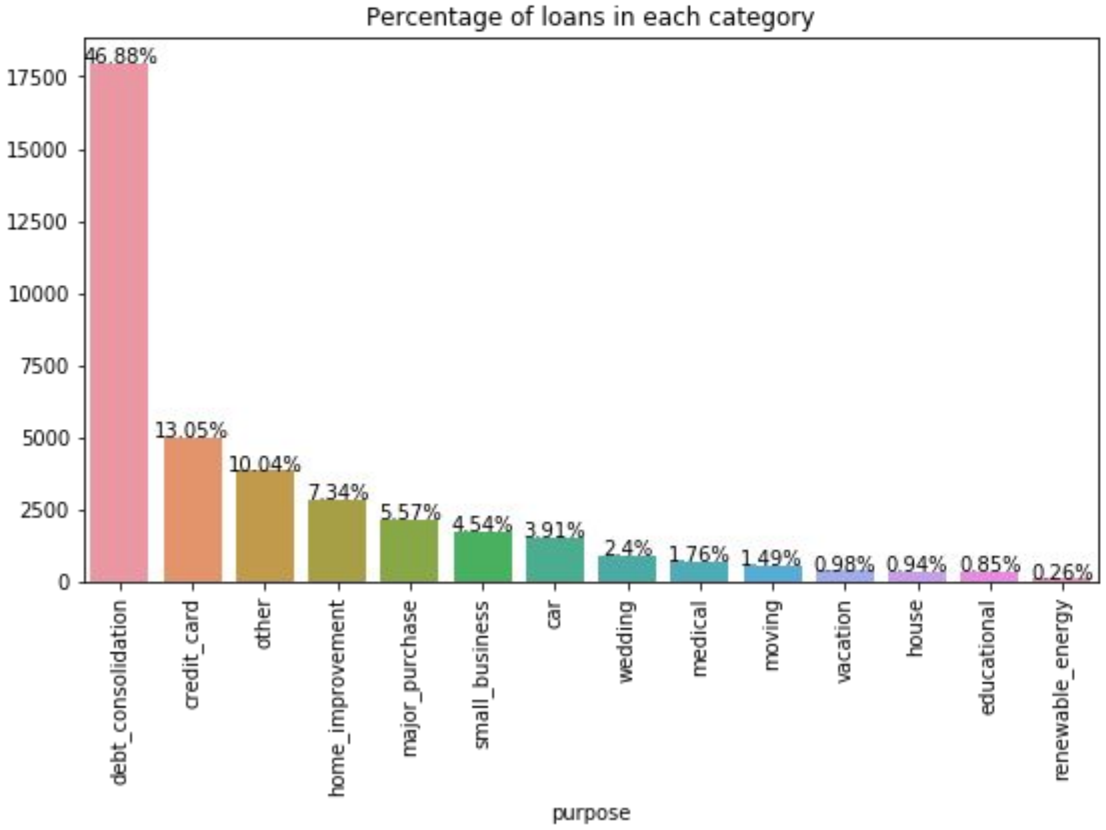


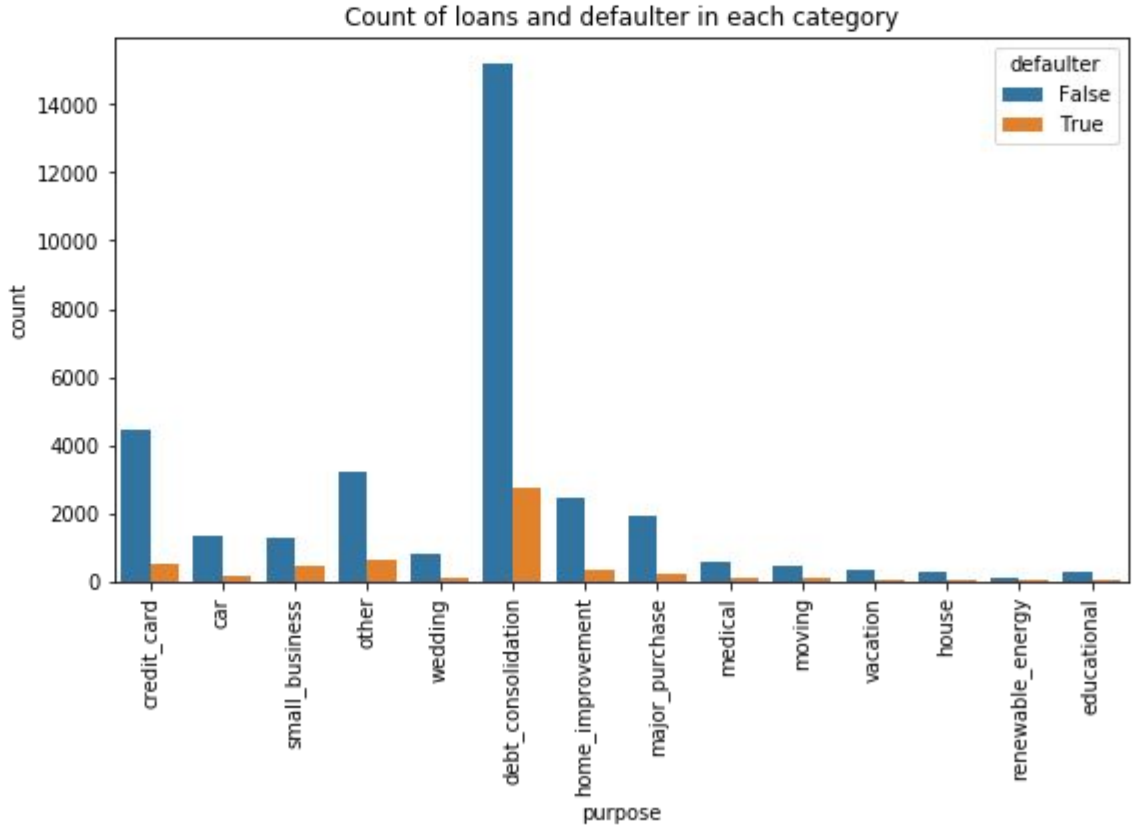
Annual Income is a driving factor behind loan default

1. People with lower income are more likely to default
2. Almost 19% of loans were defaulted in the 10-40k. In these categories, the number of loans issued was significantly higher in 30-40K category.
3. As the Annual income increases, chances of defaulting decreases

Purpose

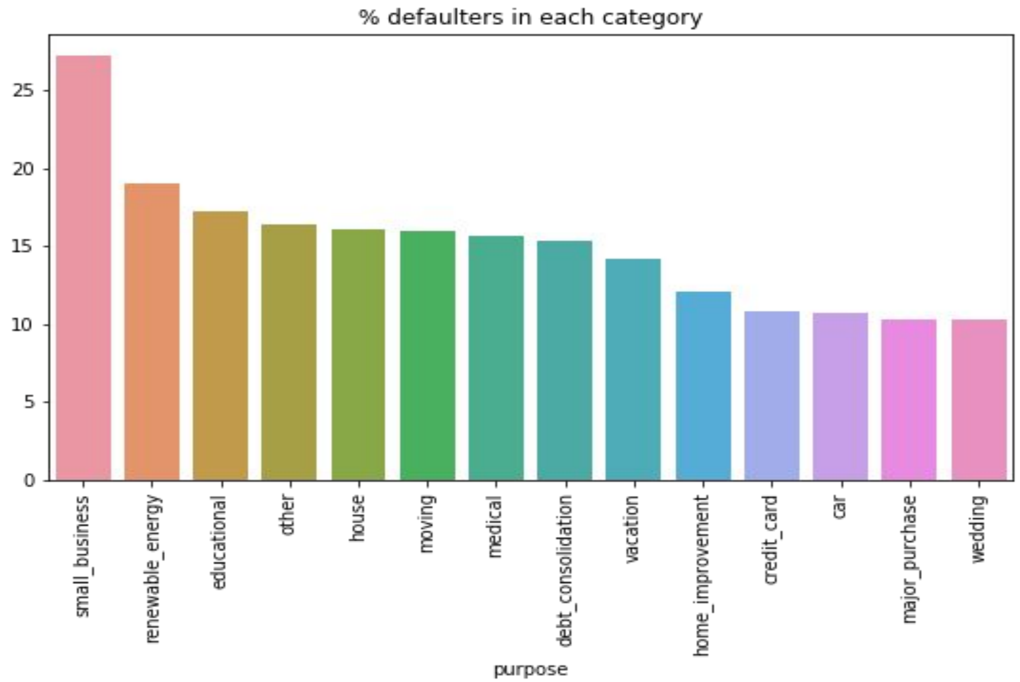
A category provided by the borrower for the loan request

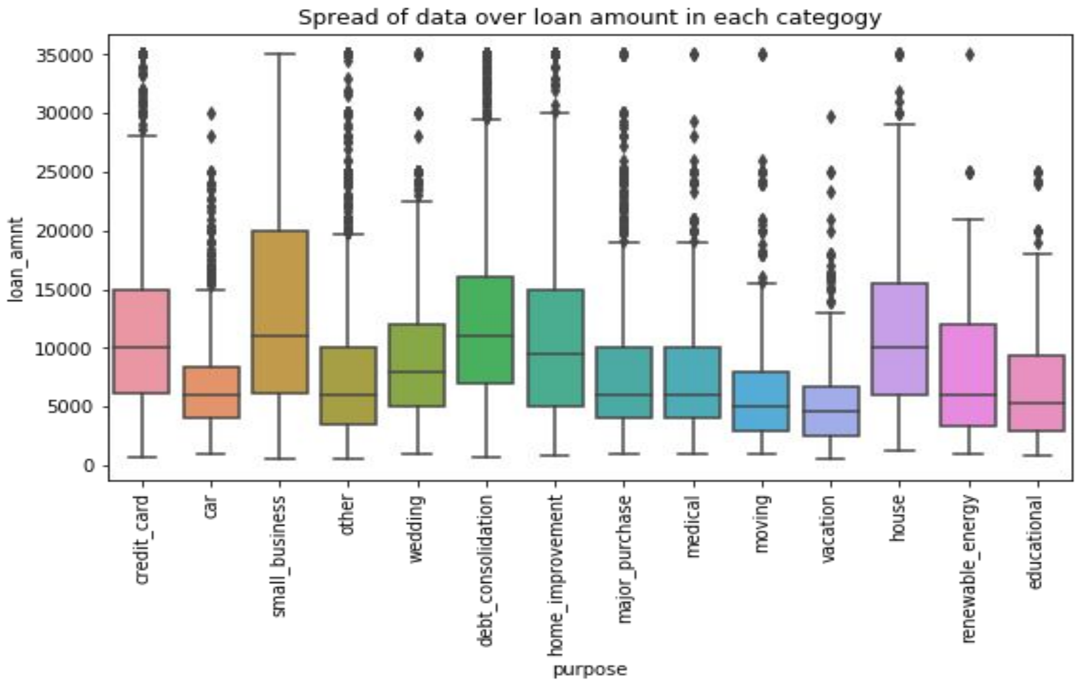




- Top 3 purposes are Debt Consolidation (46.8%), Credit Card (13.03%), Other (10.02%)
- Some categories like house, renewable energy, educational are very low in numbers
- Majority of the business happens in Debt Consolidation and credit card

Purpose (Contd..)



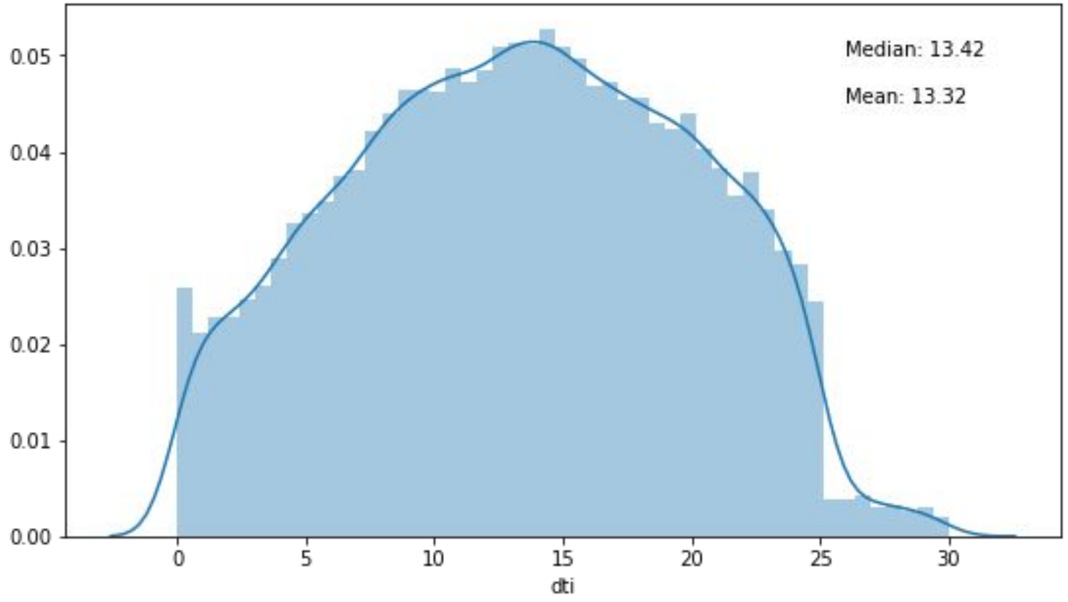


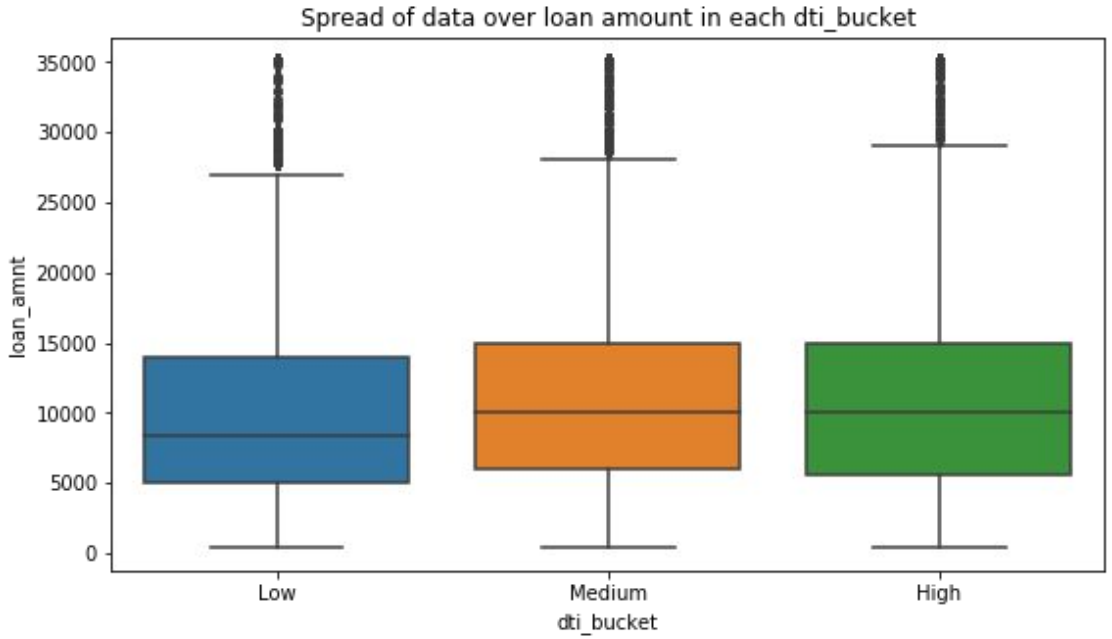
When Purpose is Small Business, chances of default are more

1. Loans where purpose is Debt Consolidation and Credit Card are 70% of total with 13% chance of default
2. Loans where purpose is Other are 10% with 17% chance of default
3. **Loans where purpose is Small Business are 4.55% with 27% chance of default**
   a. Also, box plot shows that higher amount of loans are more in this category

Debt to Income ratio (DTI)

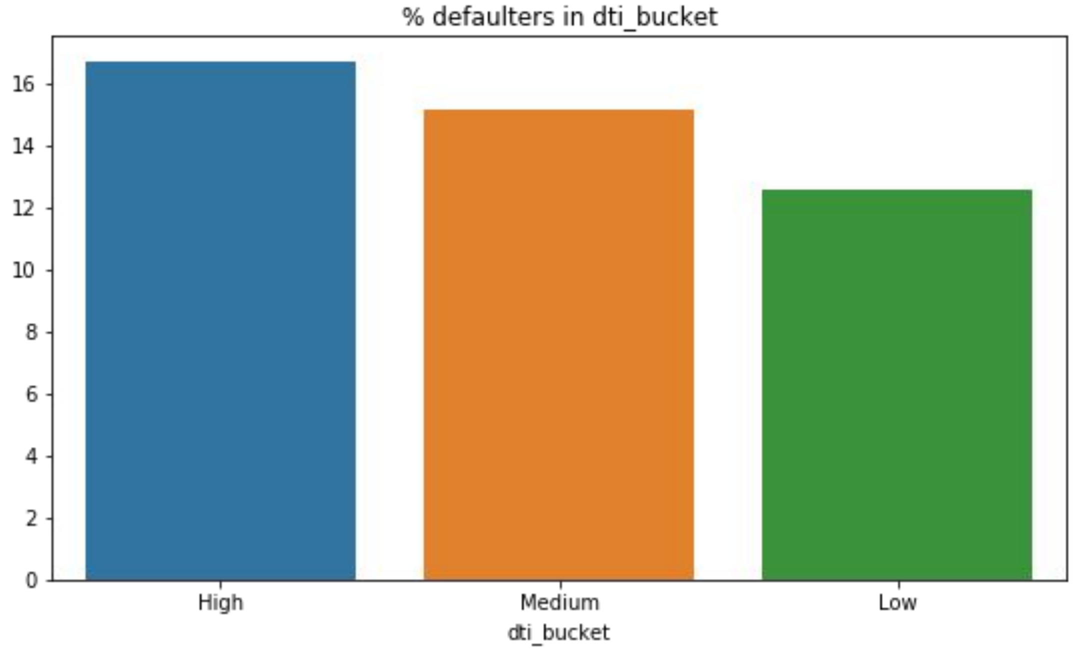
A ratio calculated using the borrower's total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower's self-reported monthly income

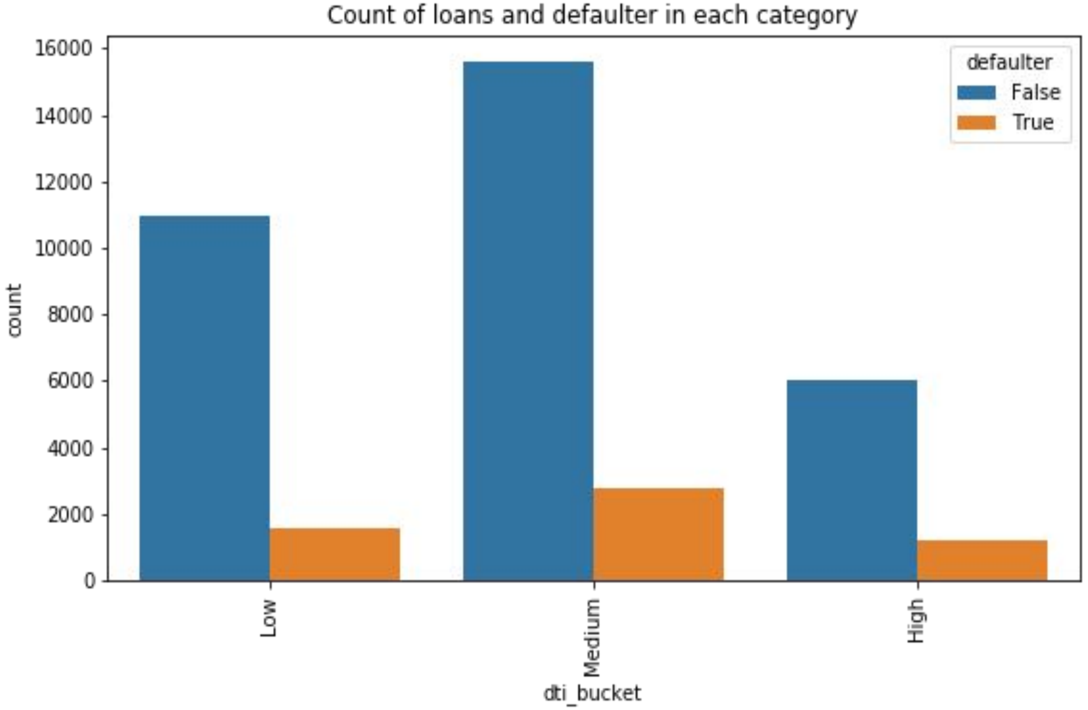




- Most of the loans have debt to total income ratio of 13.3 (Mean)
  - Which means that 13.3 % of individual's monthly salary goes in paying debt
- There is a small spike at 0 which indicates no or very minor current debt
- There is a drop at 25 which indicates that lending club rarely offers loan with dti > 25

Debt to Income ratio (DTI) (contd..)



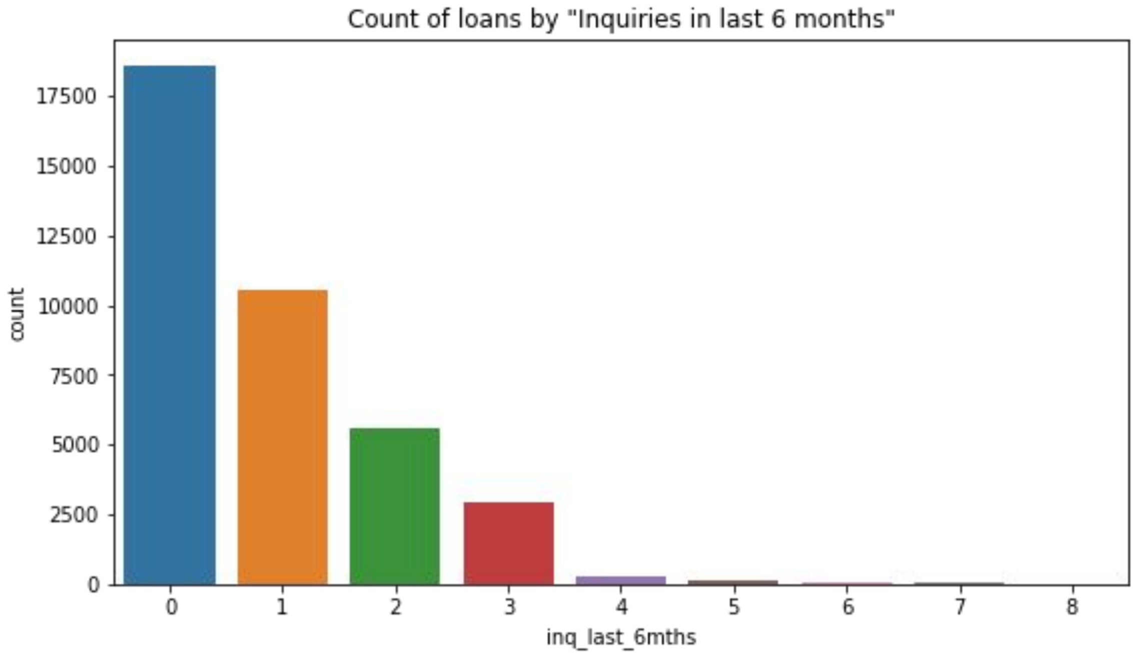


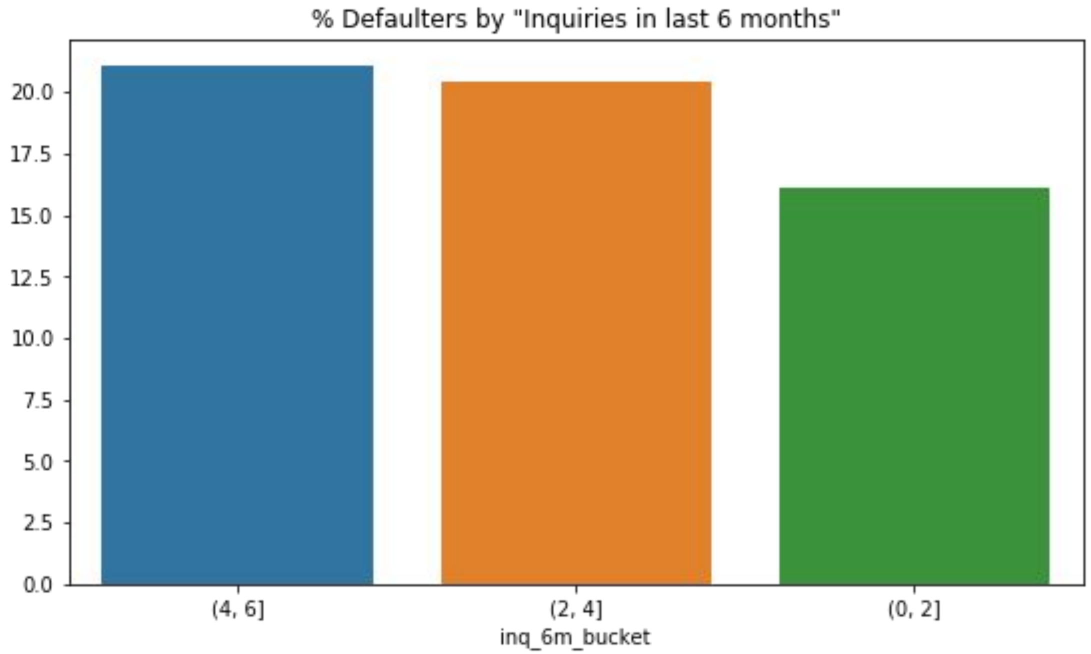
DTI is driving factor behind loan default

- We see that with dti chances of defaulting increases gradually
- In the High bucket (20-30) its 17%
- Count of loans are more Medium bucket than in High bucket

Inquiries in last 6 months

The number of inquiries in past 6 months (excluding auto and mortgage inquiries)



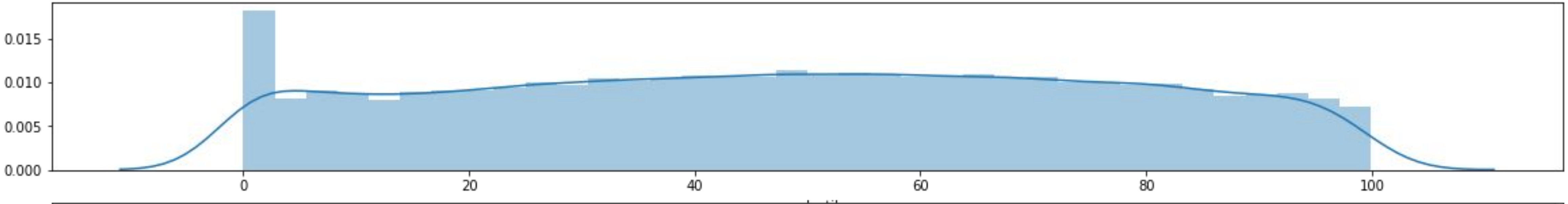


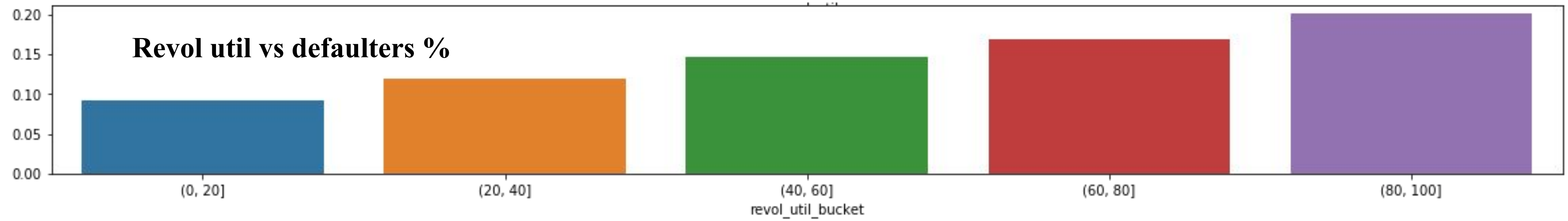
Inquiries in last 6 months is one of the driving factors behind loan default

- % of defaulters gradually decreases with decrease in # of inquiries
- If number of inquiries in last 6 months is between 4 to 6 there is 20% chance of defaulting
- Also we observe that most of the people had 0 enquiries and number loans gradually decreases as inquiries increase

Revolving utilization

Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit.

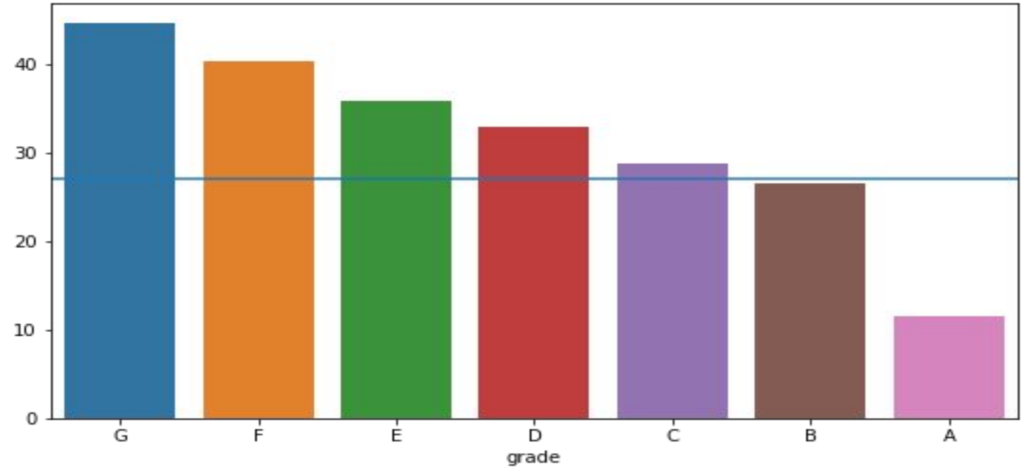




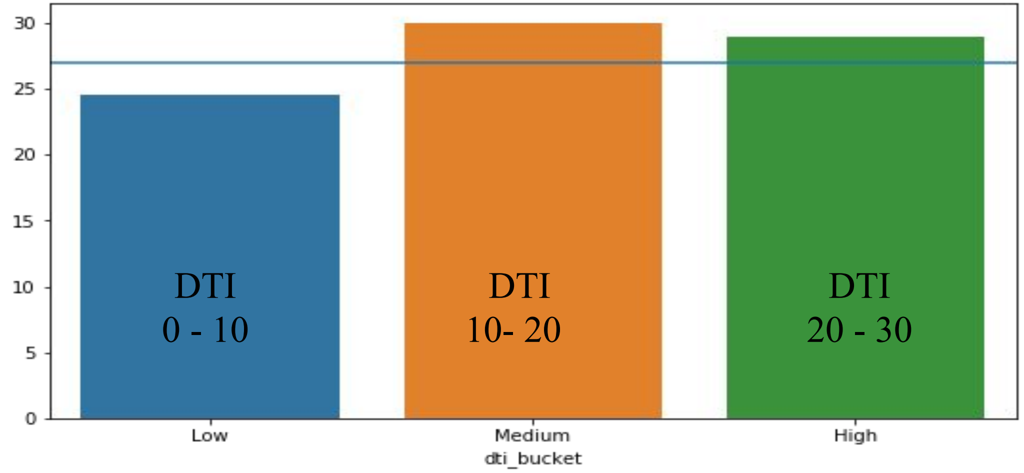
Revolving utilization is one of the driving factors behind loan default

- The graph shows that the number of defaulters increases with increase in revol_util
- So, as the revol_util increases, the chances of default increases
- Plot shows that at 0 we have a peak
  - This tells us that most of the people who got loan are not utilizing their credit line at all
- There are some people who are utilizing upto 80 - 100% of their credit line. Such people had almost 20% chance of default

Small business loans (Compared with others)



Small Business loans vs Grade



Small Business loans vs DTI

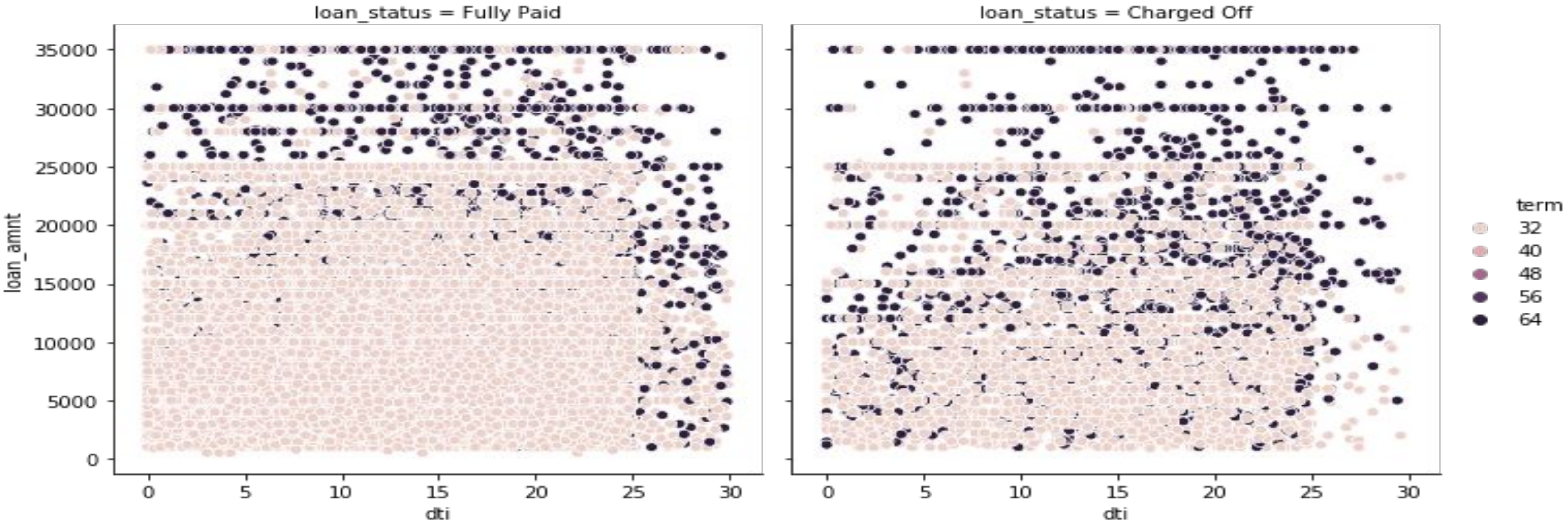
Above graph shows that for Small Business loans, as the loan grade is low, the chances of defaulters are more. We have seen that Small Business and Grade both are factors behind loan default. The horizontal lint is 27% which is defaulters % in small business

Avoid loans for Small Businesses with loans of B or lower grades

Above graph shows that for Small Business loans, as the DTI is high, the chances of defaulters are more. We have seen that Small Business and DTI both are factors behind loan default. The horizontal lint is 27% which is defaulters % in small business

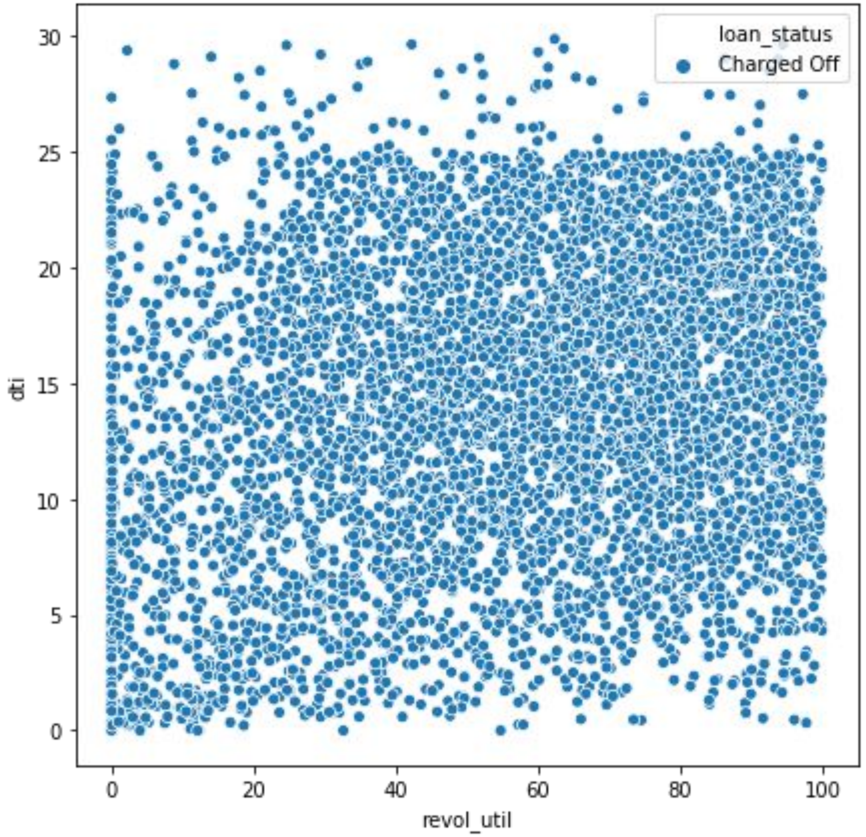
Avoid loans for Small Businesses with high DTI (10 or more)

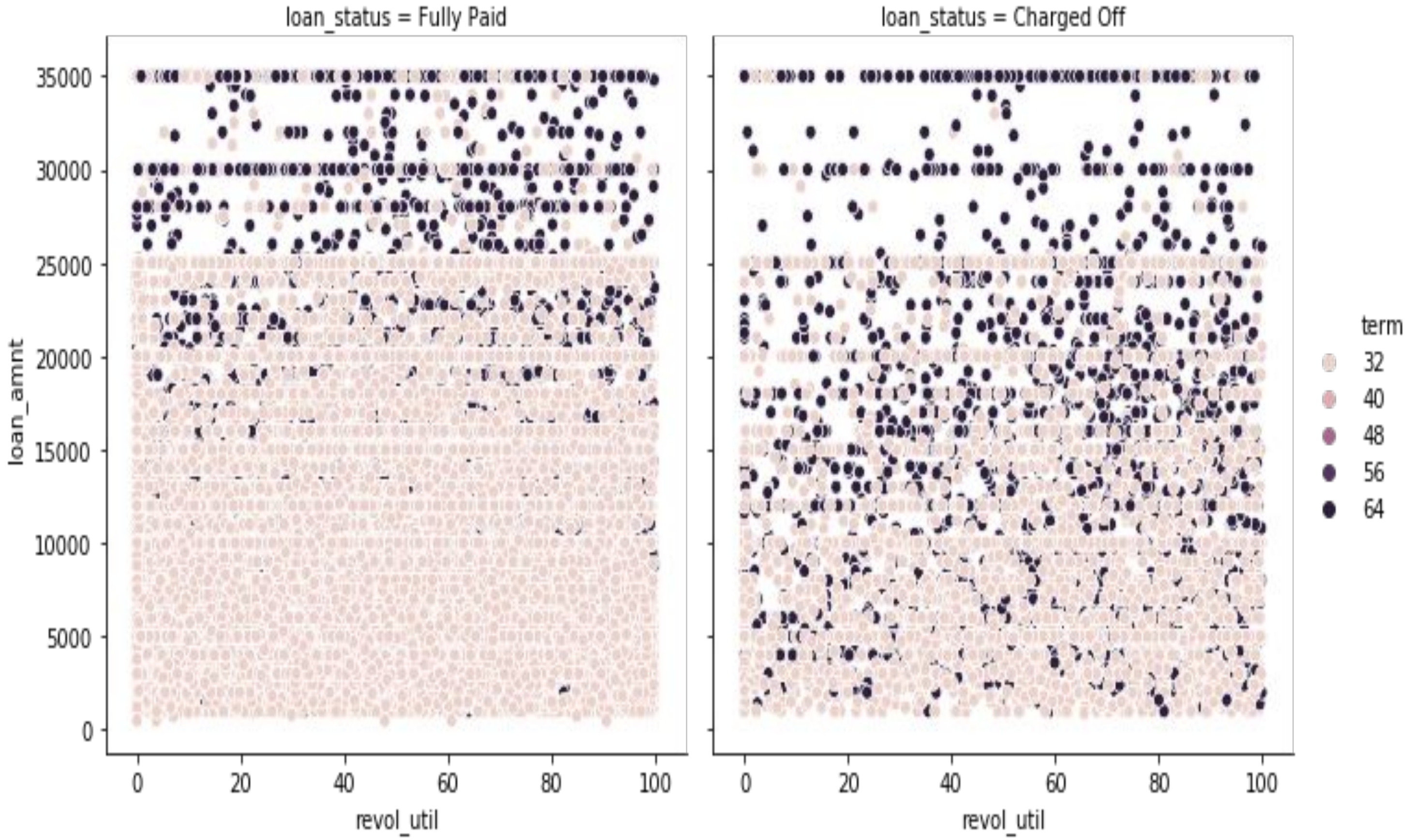
Debt to Income (Compared with others)



- Above graph shows that Fully paid loans are mostly in low term and low loan amount bracket. When DTI crosses 25 and Loan amount crosses 25000, long term loans starts
- Graph on right side shows that if DTI and Loan amount are low but if loan is long term, it is risky loan
- DTI, Loan Amount and term as all increases, the chances of default increases
- **Avoid Long Term Loans for customers with high loan amount requirement and high DTI**

Revolving utilization compared

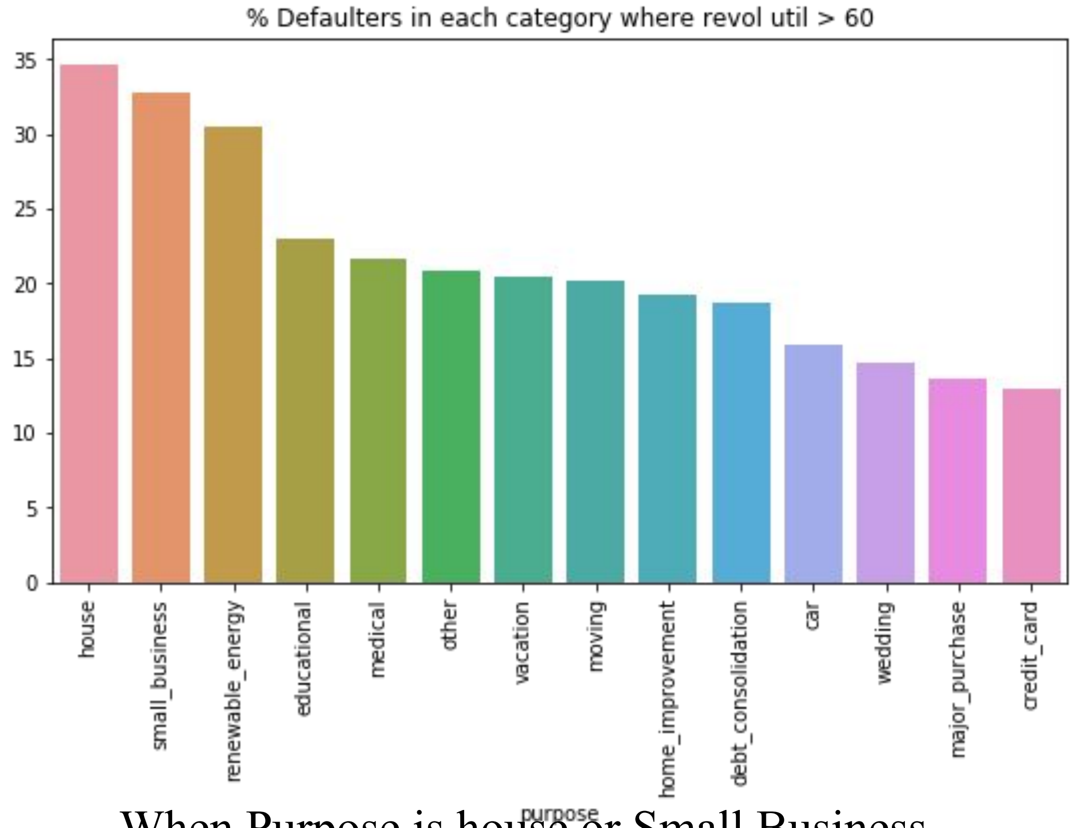


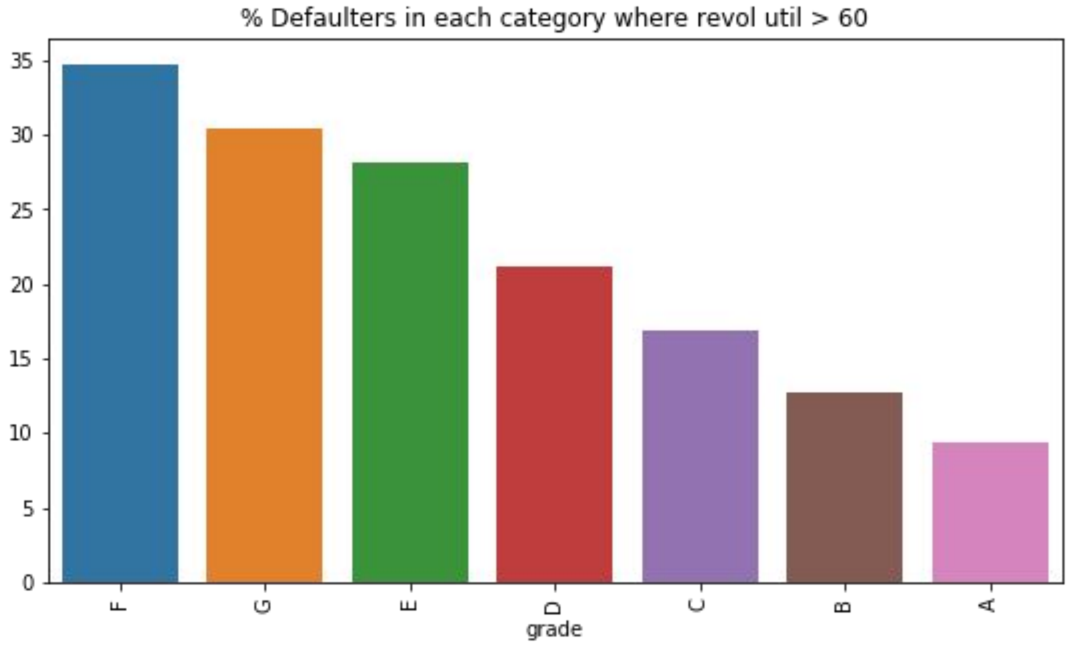


Above graph shows that Defaulters increase as Revolving utilization and DTI increase

Above graph shows again that short term loans with smaller loan amounts results in successfully paid loans. If loan Amount is small, term is high then it is a risky loan. Also we see a bit more saturation on lower right corner, which shows that High revolve util is bad even for smaller loans

Revolving utilization compared (contd.)





When Purpose is house or Small Business, with revolving utilization more than 60%, the chances of default is close to 35%

Avoid loans with high Revolving utilization

When grade is low with revolving utilization more than 60%, the chances of default is close to 35%

Avoid loans with high Revolving utilization and low grade

Conclusion

Chances of Default increases with following variables

- Loan Amount ⬆
- Term ⬆
- DTI ⬆
- # of inquiries ⬆
- Revol utill ⬆
- Grade = G, F, E
- Purpose = Small Business

Recommendations

- Avoid high loan amounts (Loans > 25k carries 20% risk of default)
- Avoid high term term (Long term loans are more risky)
- Avoid loans with G, F, E Grades
- Avoid loans for purpose Small Business
- Avoid lending to people with high Debt to Income ratio
- Avoid lending to people having high # of inquiries in last 6 months
- Avoid lending to people having high revolving utilization
- Avoid giving relatively higher loans to people with lower income.