GRAMENER CASE STUDY

1-July-2018

Ravi Divecha
Anupam Rai
Hari B
Keshava A

Business Objective

- Company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study

APPROACH FOLLOWED

- Data is analyzed and cleaned in order to remove unwanted columns
  - Review of columns with NA values
  - Review of columns with zero values
  - Correlation between columns
  - Columns have the same value
  - Columns not having any significance with regard to this case study
- Create derived columns to aid in further analysis
- Perform bivariant analysis on derived columns to study possible trends.
- Identify driving factors i.e. the variables which are strong indicators for loan borrowers to default.

UNIVARIANT ANALYSIS

- Highest number of loan borrowed were for debt consolidation purpose.
- Inspection of income indicates need for outliers treatment for better analysis of data.
- Majority of loan borrowers are with 10 open credit lines

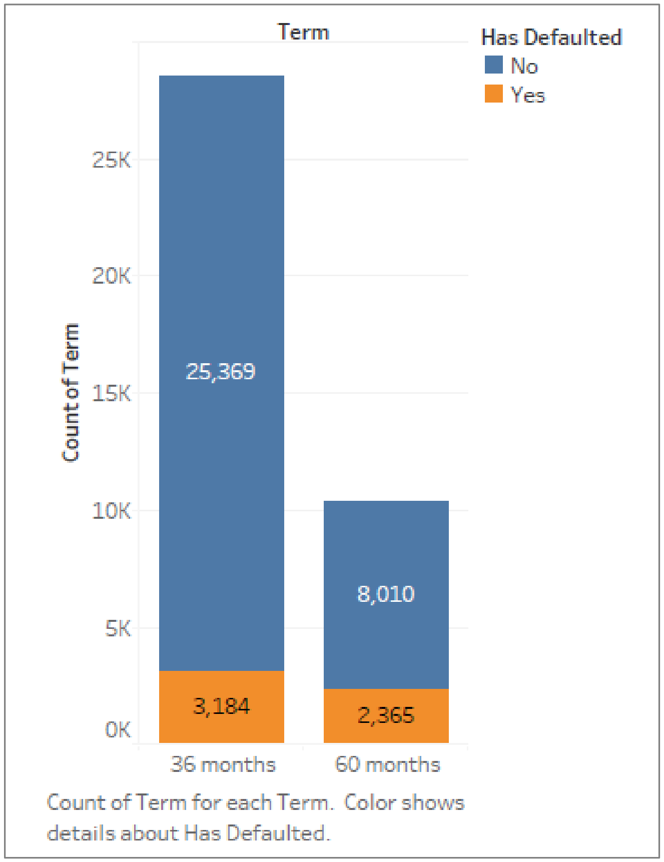
BIVARIANT ANALYSIS

- Analysis of loans defaulted vs amount of loan shows no significant trend between loan defaulter and loan non-defaulters.
- Analysis between borrower open credit lines and loan defaulters does not show any specific trend to identify possible driving factors for loan defaulters.
- Higher percentage of loan borrowers with longer payment term (6 years) have defaulted.
- Trend is observed that even though the verification status is verified or source verified, there are still high number of applicants who default on their loan
- Higher percentage of customers have defaulted as a trend when moved from Grade B to Grade G.
- High and very high interest rate show more number of loan defaulters
- Analysis of dti_category and loan defaulters show that higher percentage of customers are defaulting in the category of 20-30%
- Loan borrowers with high revolv_category are more likely to default on their loan.
- Analysis of inquiry in last 6 months with has defaulted indicates that as the enquires go up the number of defaulters also go up.
- Analysis of state with loan borrowers only shows that state of California has highest number of loan borrowers.

TRIVARIANT ANALYSIS

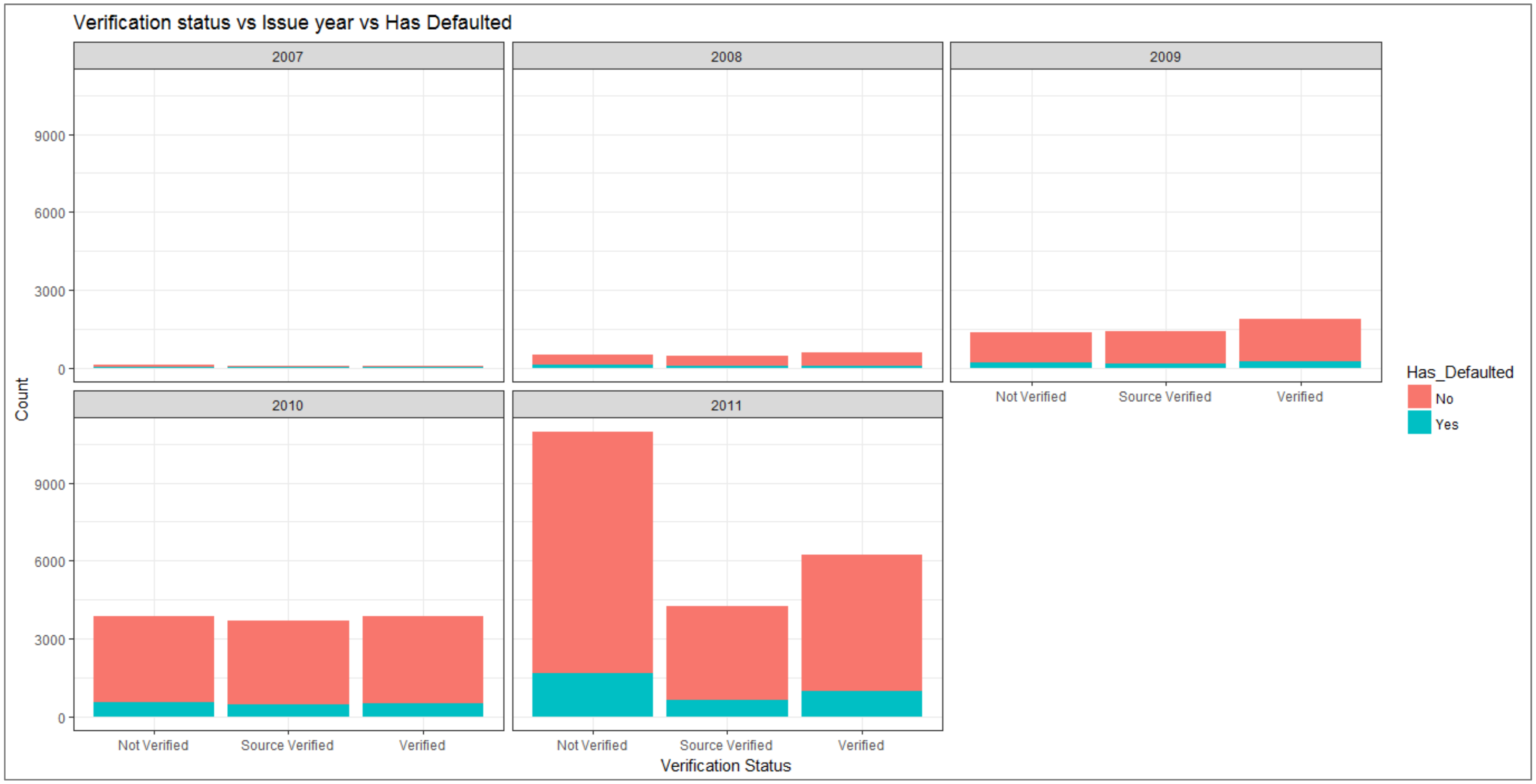
- Applicants default more with high interest rates for 60 months tenure.
- Analysis shows that applicants default more with grades B, C and D for 60 months tenure.
- Analysis shows that medium interest rate applicants default more for mortgage loans.
- Analysis for issue year vs has defaulted vs term shows that there are no 36 months loans taken in 2007, 2008 and 2009
- Analysis of verification status vs loan defaulters vs issue year indicates that in 2011 there is high number of loans and there are more defaulters even though their source was verified and income was verified.

Loan Term vs Defaulted Customers



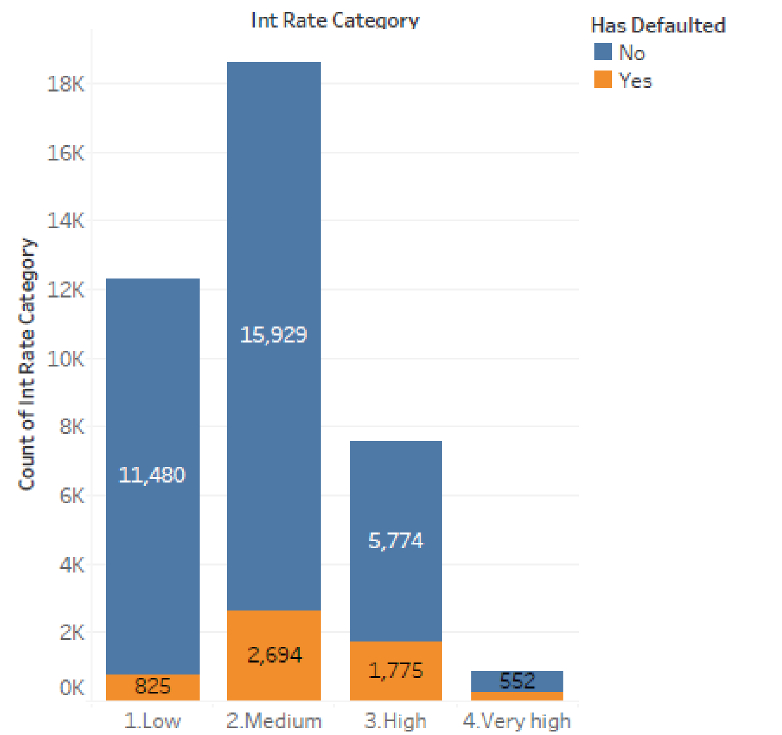
- Higher percentage of loan borrowers with longer payment term (6 years) show the possibility of defaulting on their loan.

Source Verification trend over the years



- Trend shows that even though there are loans offered with source verified and verified, there are still loan defaulters.
- There seems to be a need to calibrate source verification and verification procedures due to increase in loans offered over the years.

Interest Rate vs Has Defaulted

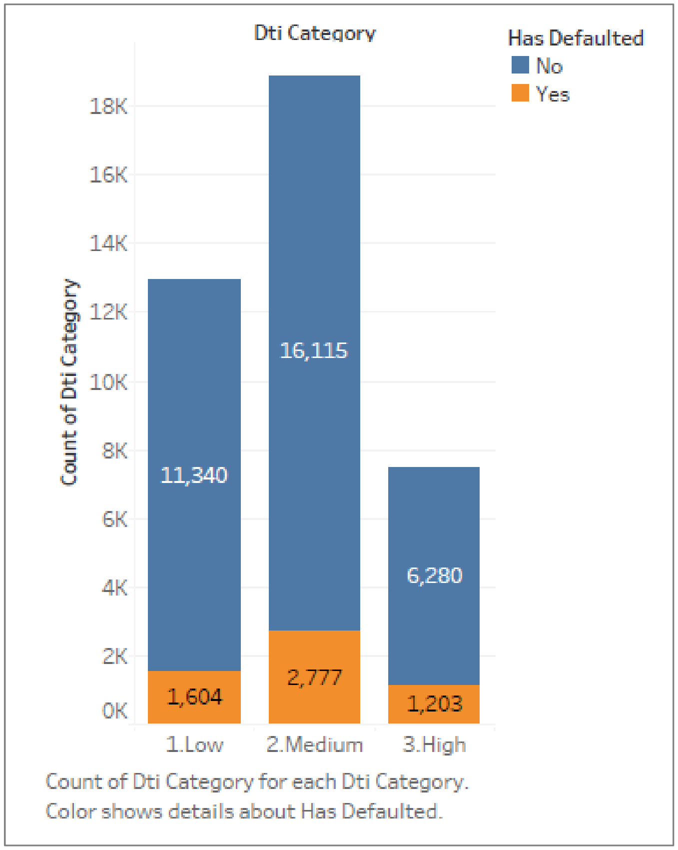


Count of Int Rate Category for each Int Rate Category. Color shows details about Has Defaulted.

Count of Int Rate Category for each Int Rate Category broken down by Term. Color shows details about Has Defaulted. The data is filtered on Revol Util Category, which keeps 1.Low, 2.Medium, 3.High and 4.Very high.

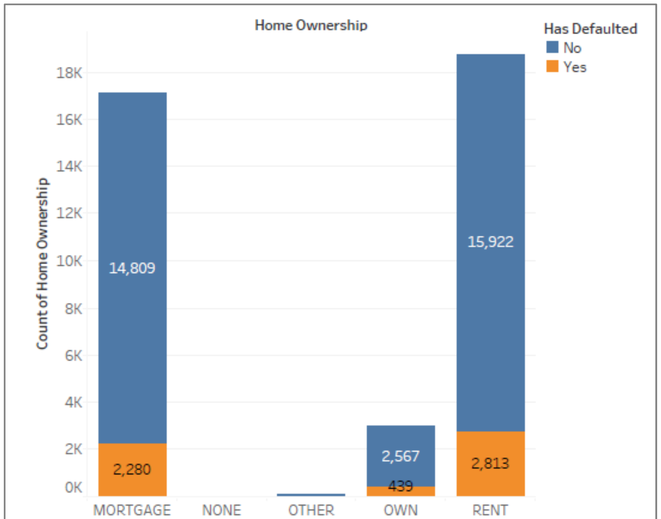
- Trend shows that loan borrowers with high and very high interest rates are likely to default.
- Loans of tenure 5 years with high and very high interest rate have an even higher probability to default

DTI vs Defaulted Customers

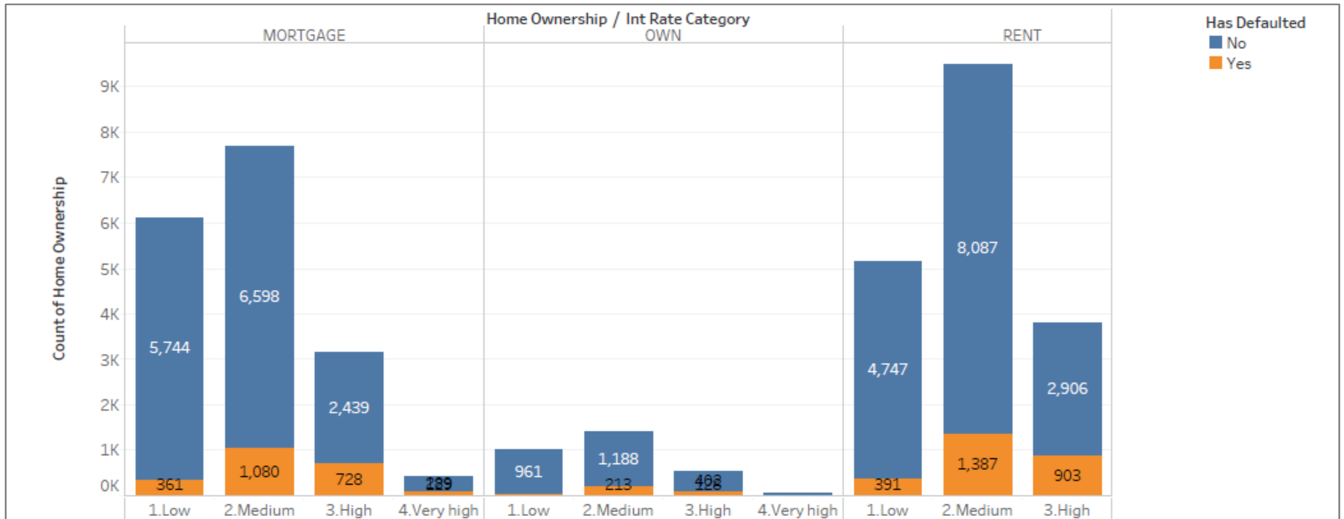


- Higher percentage of loan borrowers with higher debt to income ratio have more probability to default on their loan.

Home_ownership vs Defaulted Customers



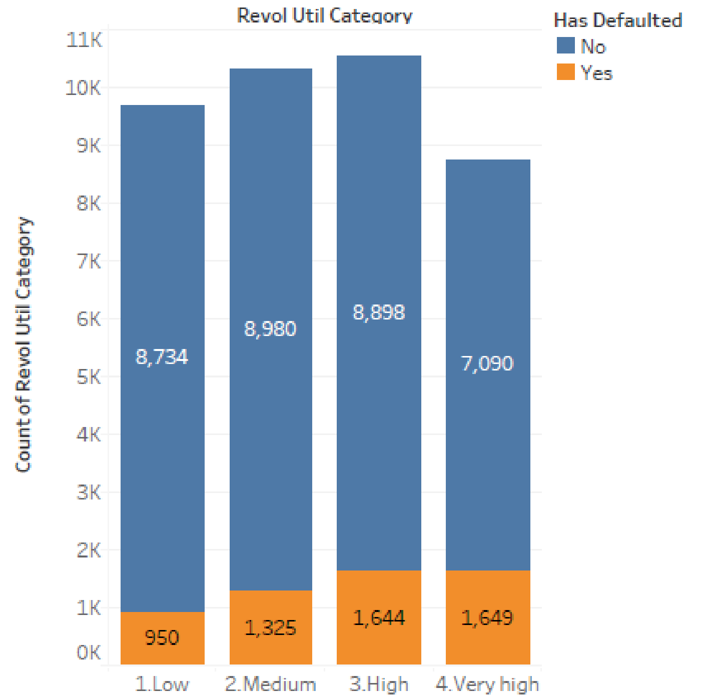
Count of Home Ownership for each Home Ownership. Color shows details about Has Defaulted.



Count of Home Ownership for each Int Rate Category broken down by Home Ownership. Color shows details about Has Defaulted. The data is filtered on Revol Util Category, which keeps 1.Low, 2.Medium, 3.High and 4.Very high. The view is filtered on Home Ownership, which excludes NONE and OTHER.

- Maximum number of loans borrowed are by applicants who currently are in a Mortgage or rented place.
- Loan borrows who are already paying mortgage and rent find it difficult to pay high and very high interest rate.

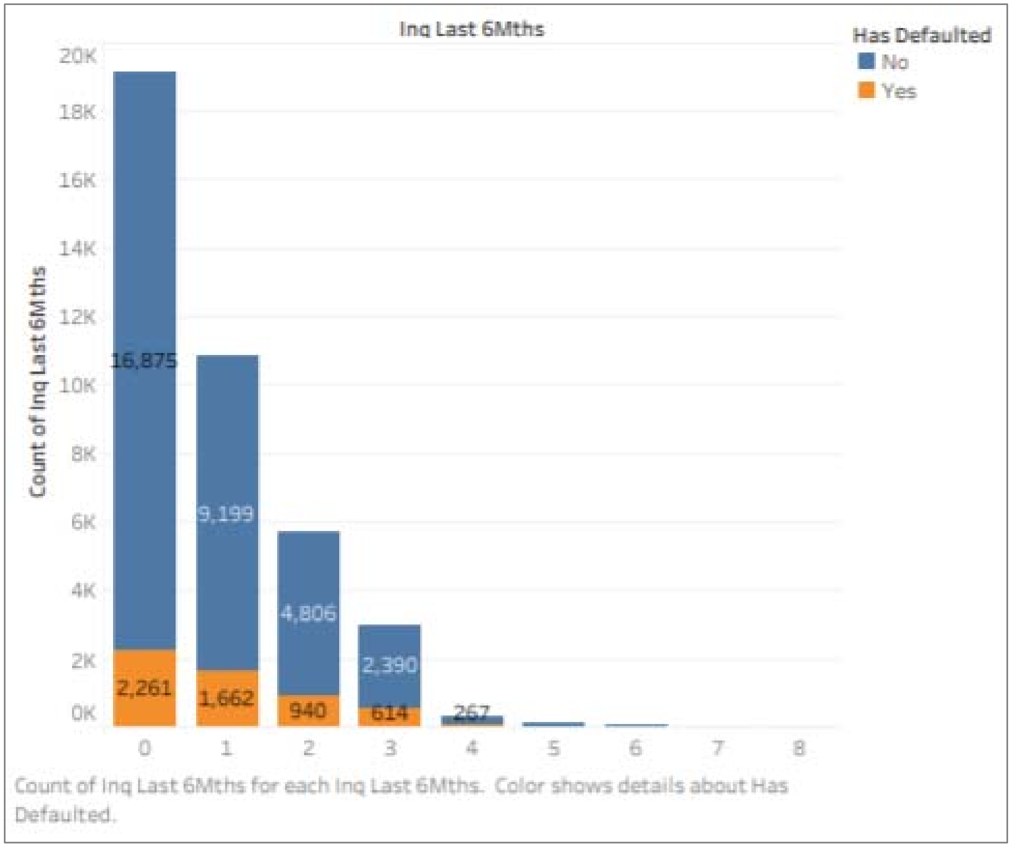
Revolv util category vs has defaulted



Count of Revol Util Category for each Revol Util Category. Color shows details about Has Defaulted. The view is filtered on Revol Util Category, which keeps 1.Low, 2.Medium, 3.High and 4.Very high.

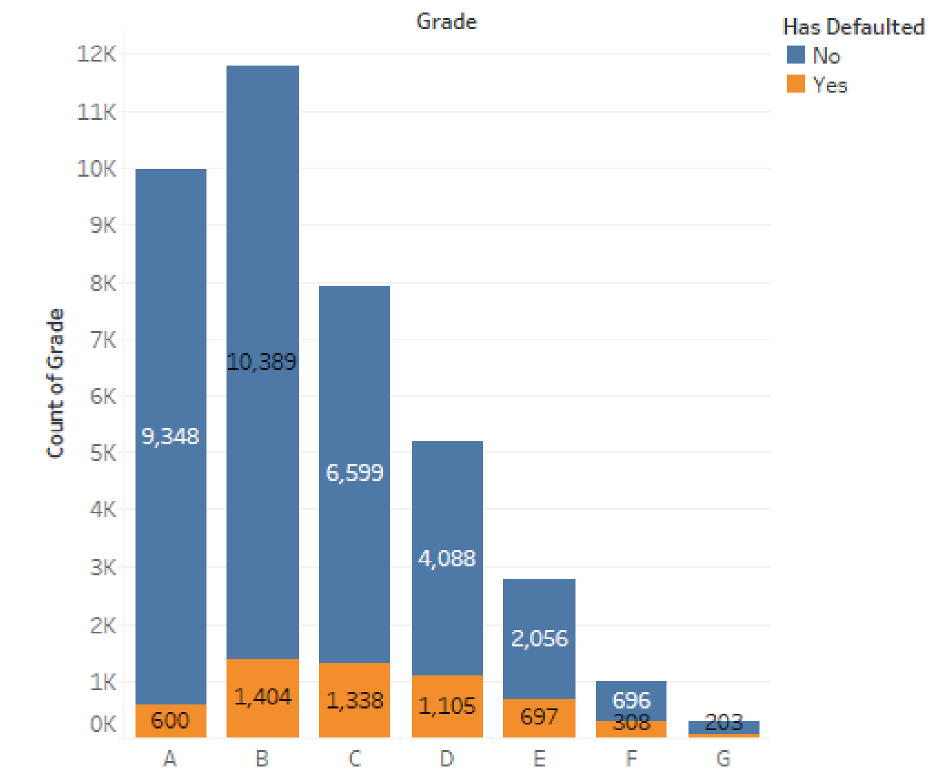
- Loan borrowers who have a higher revolving utilization show probability to default on their loans

Inquiry vs Defaulted Customers

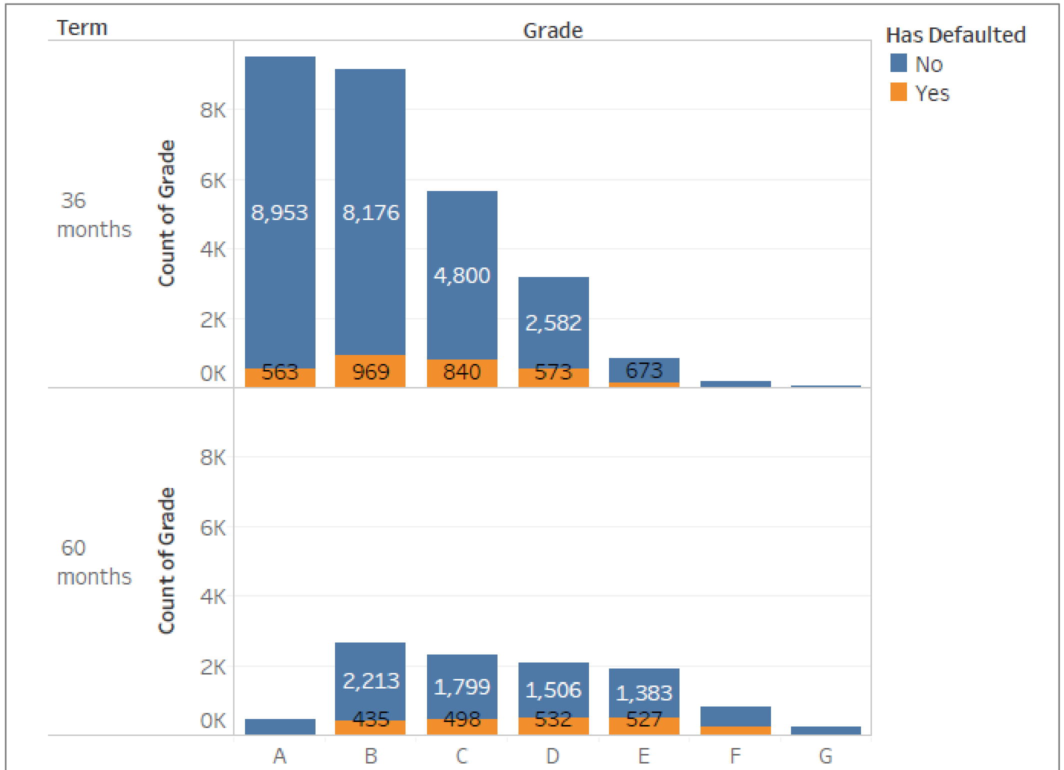


- Loan borrowers who have had higher number of inquiries regarding their monthly payments in the last 6 months show a likelihood of defaulting on their loan.

Grade vs has defaulted



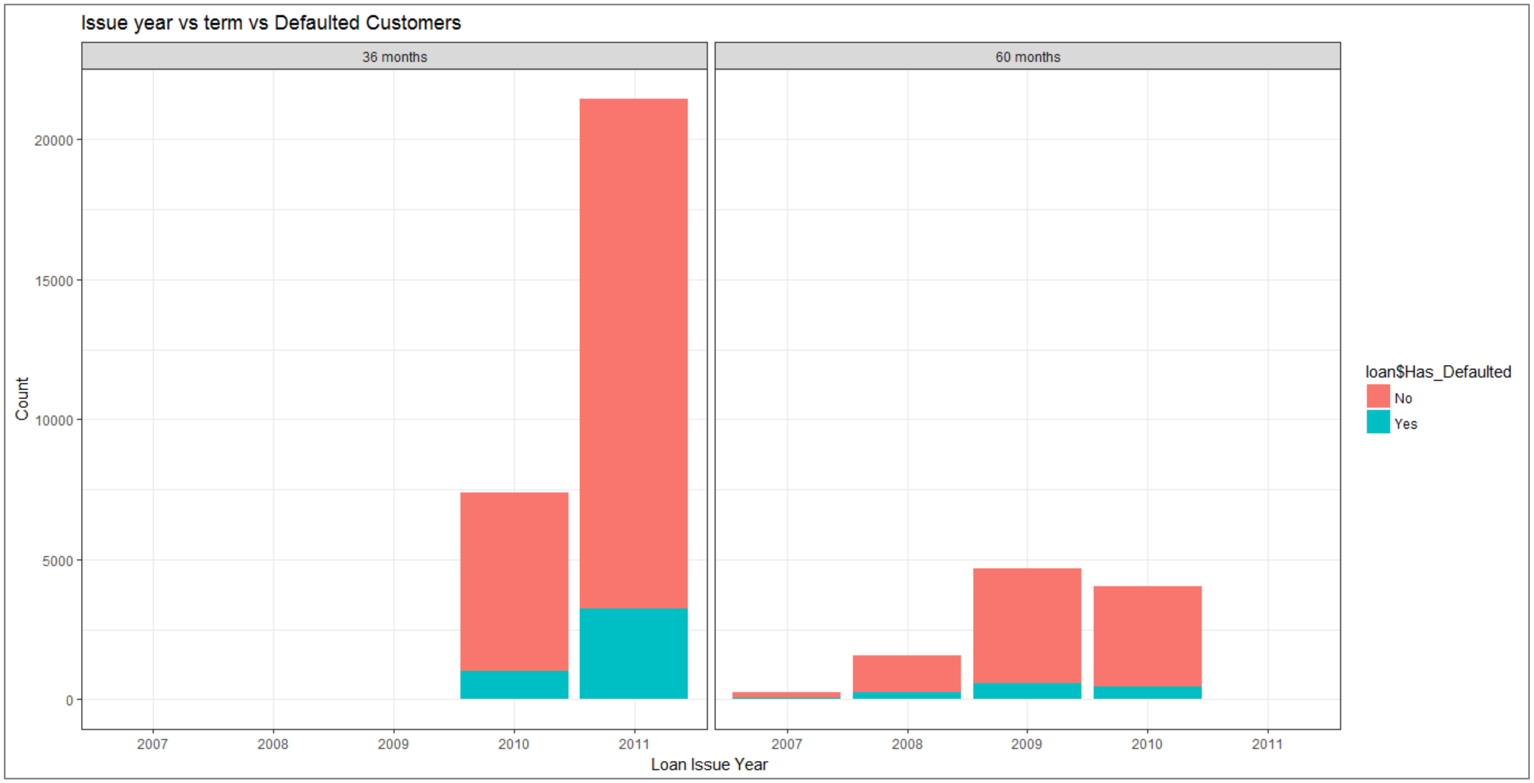
Count of Grade for each Grade. Color shows details about Has Defaulted.



Count of Grade for each Grade broken down by Term. Color shows details about Has Defaulted.

- Trend shows that loan borrowers with grades B, C, D and above have high likely hood to default.
- Loans with tenure of 5 years have a higher probability to default for above mentioned grades.

Issue year vs term vs Defaulted Customers



- Loans of tenure 36 months were not prevalent in the years 2007-2009 and as soon as it was in introduced, there was a spike in loan offerings in the year 2010 and 2011.
- Loans of tenure 60 months were not prevalent in the year 2011 indicating loan borrowers to take loans of shorter term in order to pay less amount of interest.

CONCLUSIONS

The factors that affect borrowers to default on their loan are mentioned below. Monitoring these factors can reduce the amount of credit loss for the bank.

1. Loans offered with tenure of 5 years are more likely to default.
2. Loans accepted at higher interest rates are more likely to default.
3. Loans offered with grades B, C, D and above are more likely to default.
4. Loan borrowers with high debt to income ratio are more likely to default.
5. Loan borrowers with high revolving utilization rate are more likely to default.
6. Loans borrowers verification status and source verification status should be calibrated since there are loan defaulters even though verification status is good.
7. Loan borrowers who have more number of inquiries in 6 months show a tendency to default on their home lone.