# Understanding Credit Risk

A group project by:
Vidhi Thakkar
Vighneshwar Mishra
Asha Katla
Anjala A R

Credit Risk problem

- 5300+, ie 14% loans have been defaulted, which has contributed to a loss considerable for the organization

- The organization is in need for a better metrics to identify the defaulters

- The data available is huge, and needs analytical understanding to understand the trend of defaulters
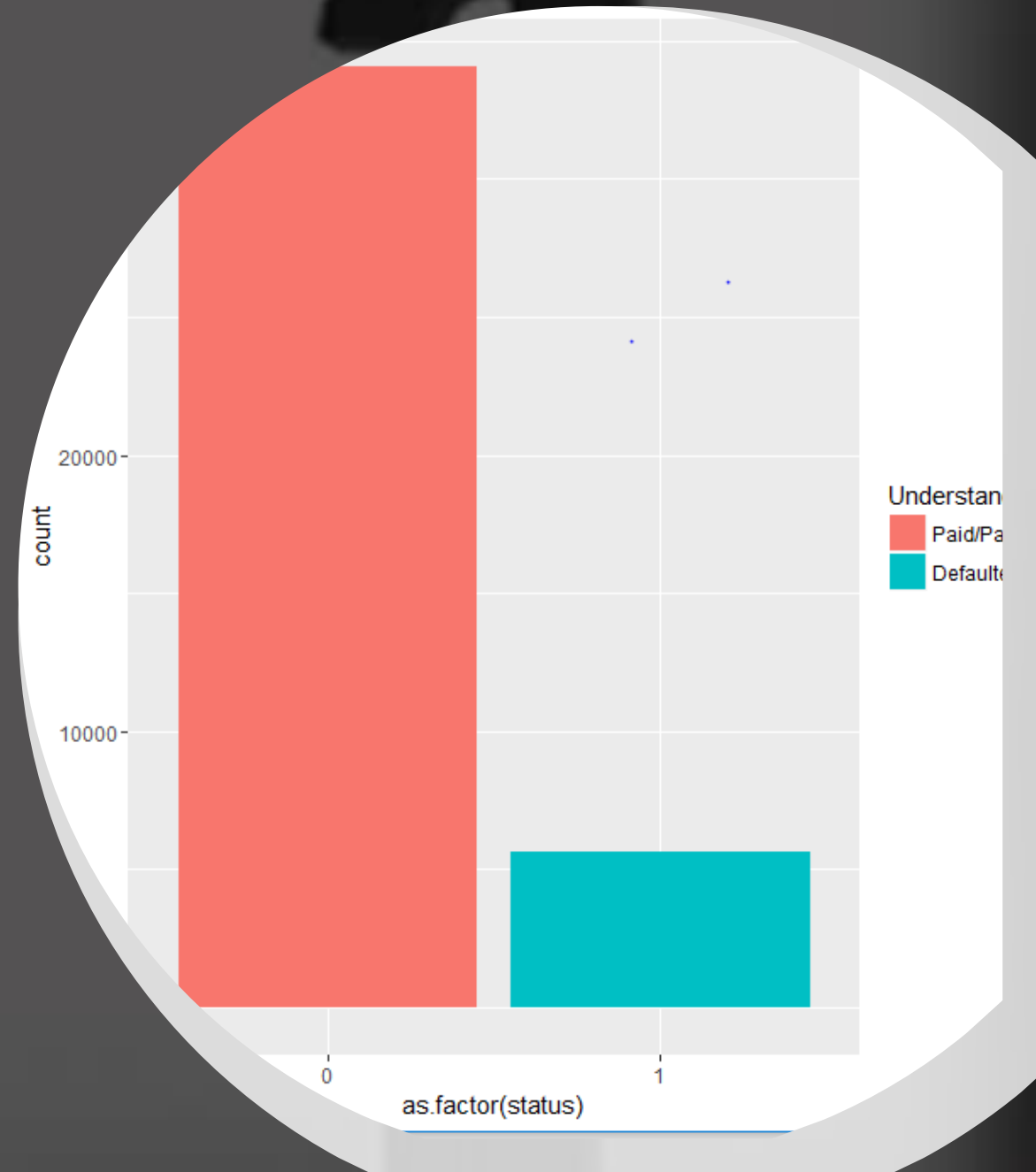
Summary: Flow of execution

- Filtering the variables from 115 to 23 has been done after careful consideration of the effect the variable has with the credit risk

- All variables which has no effect on the credit tendency has been removed. More details on the important variables can be seen in the upcoming pages.

- Derived metrics has been created which will help the analysis

- We then studied how credit risk varied with each of the selected variable

- We have also included measures can take to mitigate the risk and pointers to identify good loans

Variable	Type	Meaning
Loan Amount	Continuous	Loan amount applied by a customer
Instalment	Continuous	Monthly instalment to the bank by the customer
Grade	Categorical	Credit grade the customer belongs to
Sub-grade	Categorical	Credit sub-grade the customer belongs to
Employee Experience level	Categorical	Years of experience in the current organization
Home ownership status	Categorical	Indicator if a customer owns a house or not
Annual Income	Continuous	Annual Income of the customer
Issue date	Categorical	Loan issue date
Purpose	Categorical	Loan purpose as given by the employee
No of delinquencies in 2 years	Categorical	
Earliest Credit line	Categorical	Earliest reported credit issue date of the customer

# Continuous & Categorical Variables used
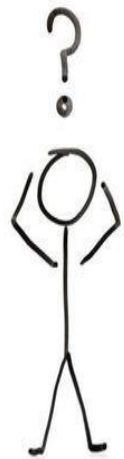
| Variable	Type	Meaning
No of inquiries in last 6 months	Categorical	No of inquiries for a loan in last 6 months by the customer
No of open accounts	Categorical	No of open accounts by the customer
Public records	Categorical	
Status of the loan	Categorical	Status of the loan
Interest Rate	Continuous	The interest at which loan has been issued by the bank
Term of the loan	Categorical	Term of the loan
Income	Continuous	Monthly income of the customer
Credit Line in Days	Continuous	No of the days since the first credit date
Annual Instalment	Continuous(Derived variable)	Annual Instalment made to the bank by the customer
Income to Instalment ratio	Continuous(Derived variable)	Ratio of annual income to annual instalment by the customer
Loan to income ratio	Continuous(Derived variable)	Applied loan to annual income ratio

Continuous & Categorical Variables used
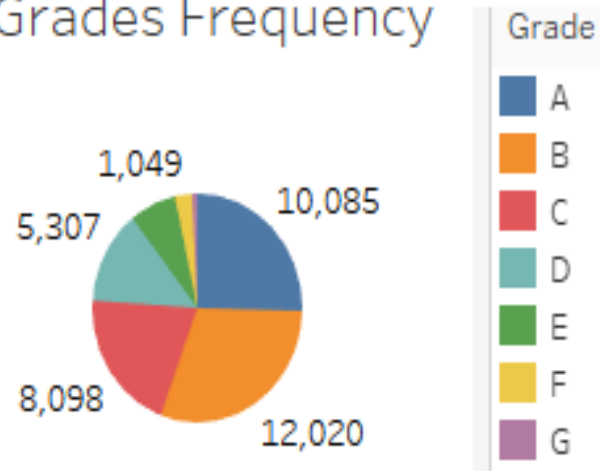
Summary: Data at hand

Status Frequency

Status
- ■ Defaulter
- ■ Paid/Paying

5,627

34,090
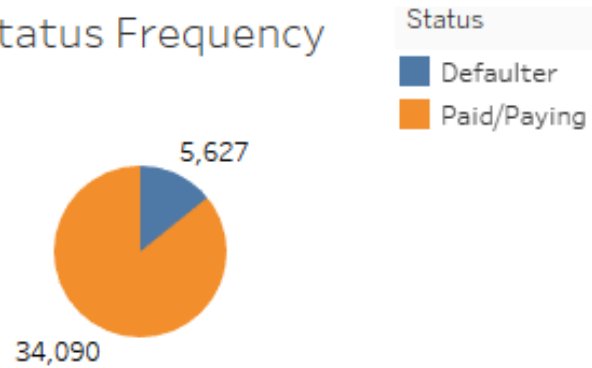
The data contains details about 39717 borrowers, with 14.16 % of the loans have lapsed

Home Ownership Frequency

Home Ownership
- ■ MORTGAGE
- ■ NONE
- ■ OTHER
- ■ OWN
- ■ RENT

17,659

18,899

3,058
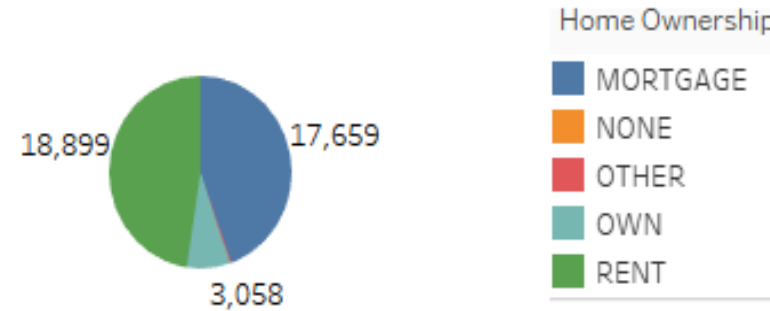
We can see that most borrowers do not own a loan.

Grades Frequency

Grade
- ■ A
- ■ B
- ■ C
- ■ D
- ■ E
- ■ F
- ■ G

1,049

5,307

10,085

8,098

12,020

Borrowers are of 7 grades and have 5 sub grades with in them. The most of the borrowers are from B grade and followed by A grade.

Employee Experience Frequency

Emp Experience
- ■ High Experienced
- ■ Junior/Not available
- ■ mid-Level
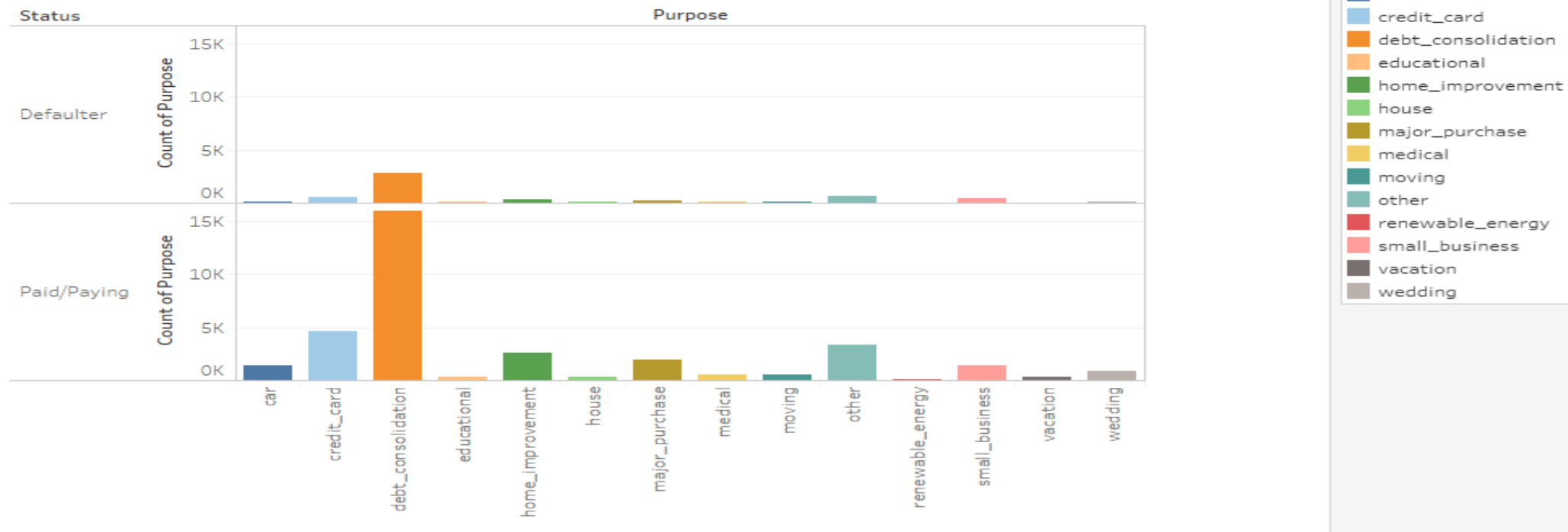
13,042

13,389

13,286
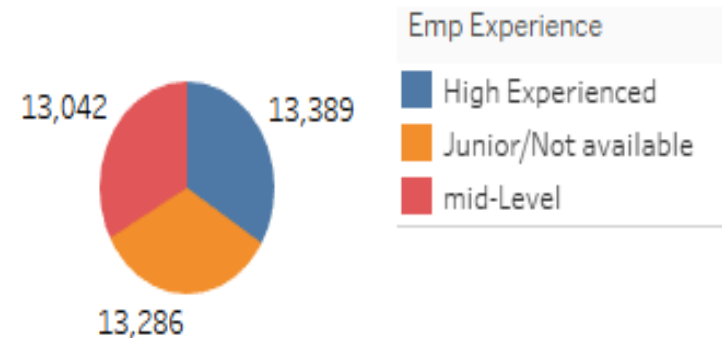
Borrowers are classified into 3 based on the years of exp in the current organization

Understanding Borrowers Across categories

- Loans taken for starting small business, having contributed to the most number of defaults are the riskiest category to sanction loans and loans taken for education and moving are the safest loans.
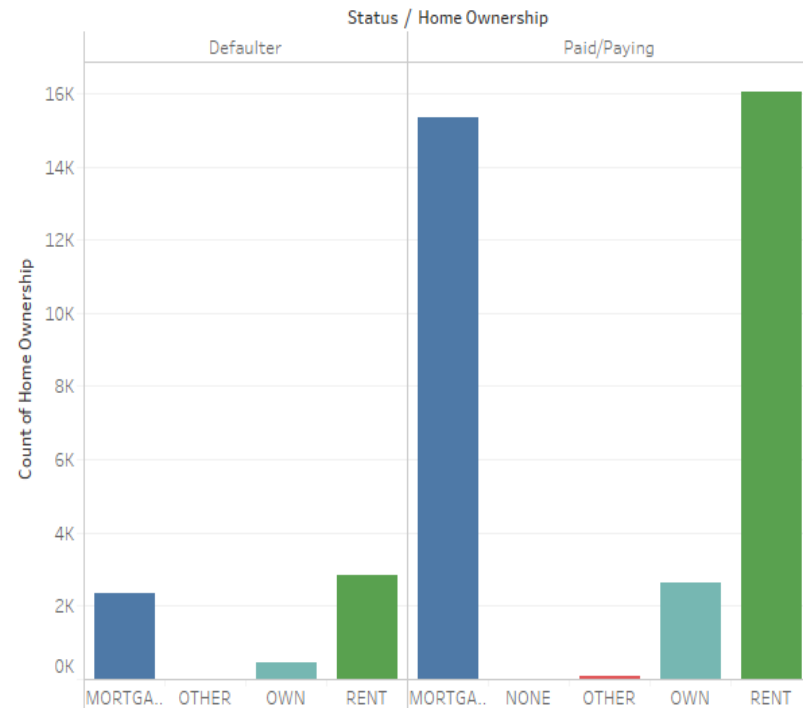
Understanding Borrowers/Risk trend across categories

- borrowers can be bucketed into two types based on whether they own a home or not.

- This is important to understand to understand if the borrower can pay the monthly income over the rent to be paid

- Also, owning a home is a good collateral

- But from the data, we can see that 91% of the defaulters do not own a home

Defaulters as per Home Ownership

Home Ownership
- MORTGAGE
- NONE
- OTHER
- OWN
- RENT

Understanding Borrowers/Risk trend across categories

- borrowers may also be categorized based on the number of years of experience. We have bucketed the employees into 3 based on their years of experience –
  - a. Junior/Not working (NA, 1-2 yrs of exp)
  - b. Mid level employee (3-6 yrs of exp)
  - c. High Experienced (7 – 10+ yrs of exp)
- Years of experience of the borrower is crucial to understand the career growth of the individual, if the income will be stagnant over a period etc.
- From the graph, we can infer that the riskiest loans are for senior level employees, followed by Junior/not working borrowers


Defaulters as per Employee Experience
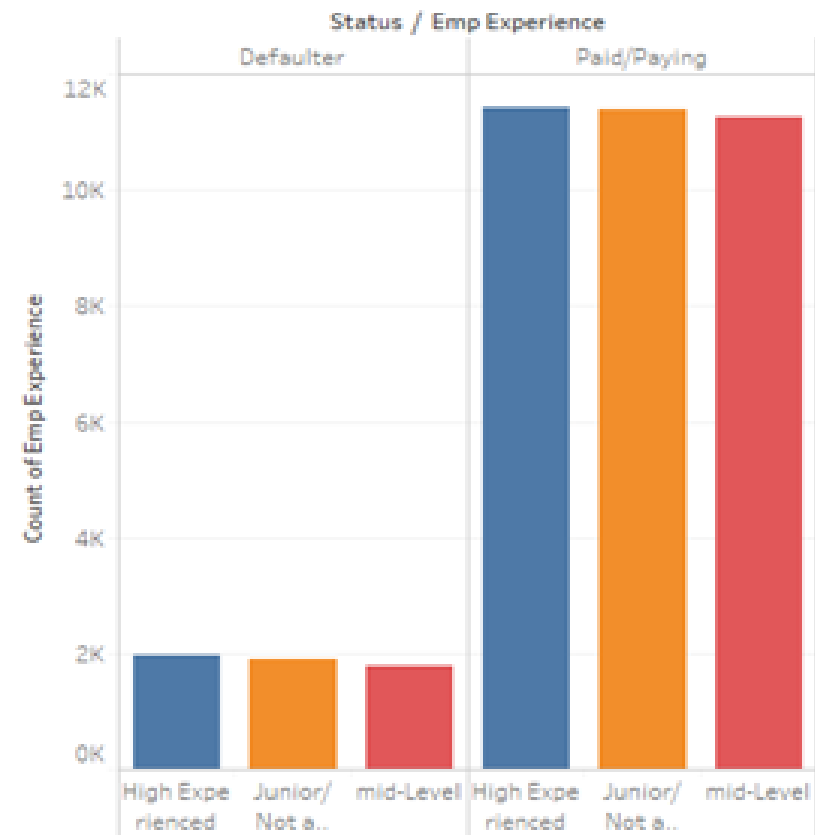
Understanding Borrowers/Risk trend across categories

- We have also classified the borrowers on their income level. There are 5 categories
  - ➤ a. Federal Poverty Level (0- 24999 USD)
  - ➤ Low Income (25000 – 34999 USD)
  - ➤ Middle Income (35000 – 49999 USD)
  - ➤ High Income (50000 – 74999 USD)
  - ➤ Super Rich (75000 USD and above)

  - It is important to know which category of borrowers contributes to the most defaulters, so that necessary precaution may be added for such borrowers in future

Understanding Borrowers/Risk trend across categories

- Counter intuitively, the graph suggests that borrowers below federal poverty are less riskier than borrowers than all the other category.


Defaulters as per categorised Income Levels
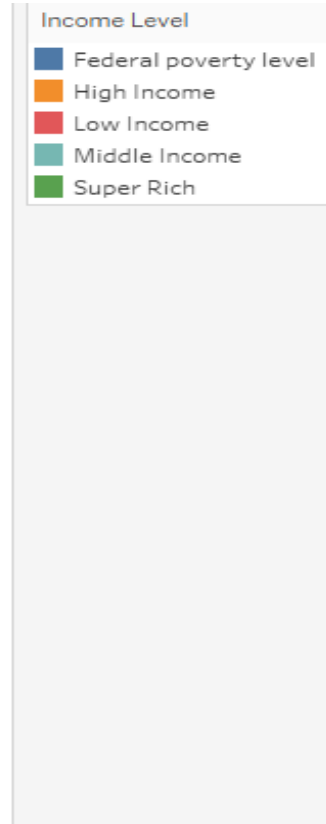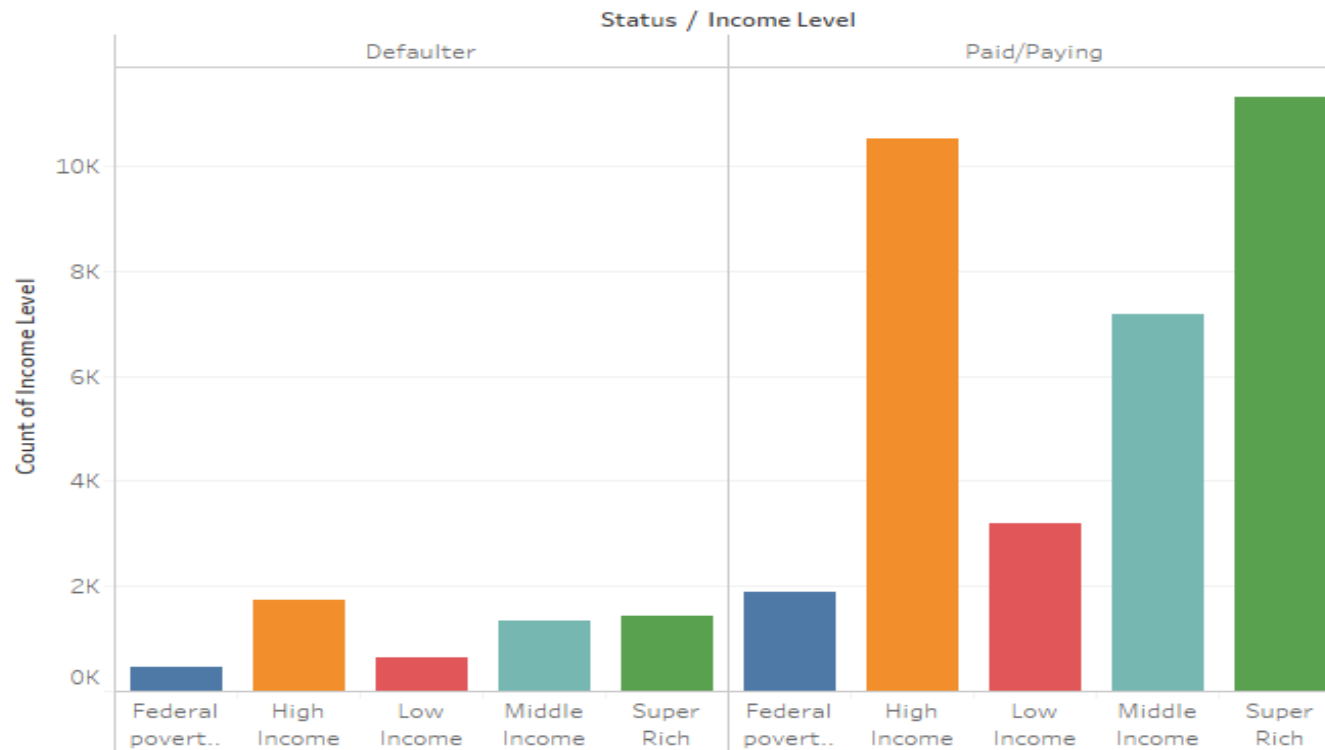
Understanding Borrowers/Risk trend across categories

- Loan grade is an indicator of the credit score of the individual. borrowers with higher credit score are expected to be less risky.

- The data completely agrees with this theory.

- The trend of defaulters increases from grade A(5.9%) to G (31%)

- But the number of defaulters are more in mid categories – B and C.

- Data shown in the graph are absolute numbers.

Defaulters across Grades
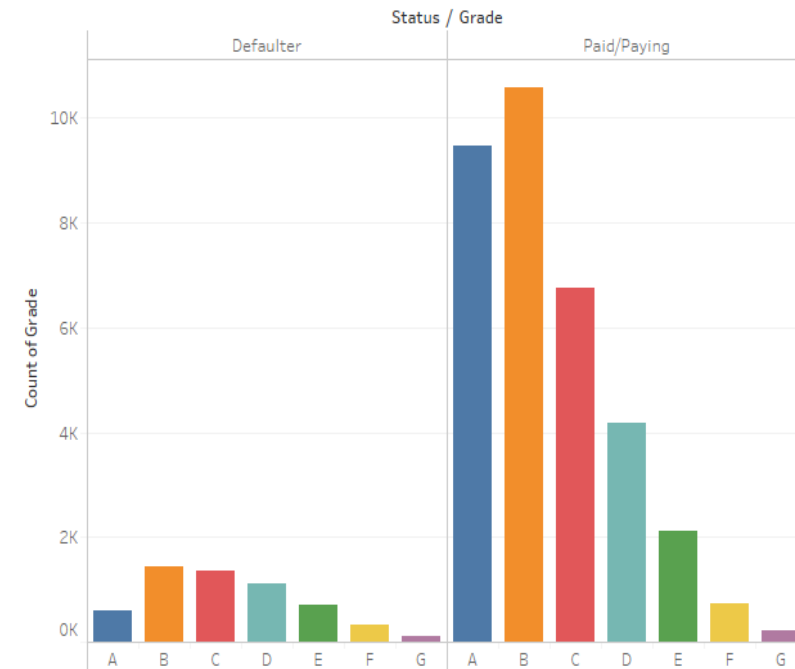
Grade
- A
- B
- C
- D
- E
- F
- G

Summary of Analysis – Part 1

- We have defaulters spread across multiple categories, and it is complex to predict the exact defaulters with this data

- However, we can see some clear indicators– Though grade is a good measure to pick defaulters, with the number of defaulters proportional to the merit of the grade.

- At the same time, it is interesting here to note that the most number of defaulters are from B and C grade. It is important to look at why these mid grades contribute to the most number of defaulters.

- Purpose of the loan is also a good qualitative measure to look at while giving a loan. The loan taken for debt clearing is a high risk loan.

Summary of Analysis – Part 2

- The data clearly shows that 90% of the defaulters do not own a house. Now this combined with other risk variables like a low grade or a high income may not be advisable.

- We have also found that the income to instalment ratio for defaulters are around 0.73. It will be recommended to check these for borrowers in future.

Suggestions

- To look deeper at additional risk indicators among grade B and C borrowers and across the income categories – middle, high and super rich before approving a loan to these categories

- Borrowers under poverty level and F and G grades have the smallest number of defaulters. It is advisable to give them more loan under controlled conditions