# Lending Club Case Study

Data Driven Analysis

Business Problem

- A Financial Cooperation wants to reduce the risk of credit loss by taking strategic decisions to approve loan for prospective customers based on historic data available.

- As a Data Analyst , We are here going to Identify patterns which indicate if a person is likely to default or not.

- Based on the analysis , FC can make strategy and policies to :

- Deny/Approve the loan.

- Reducing or purposing higher amount of loan.

- Lending at higher interest rate to risky applicants.
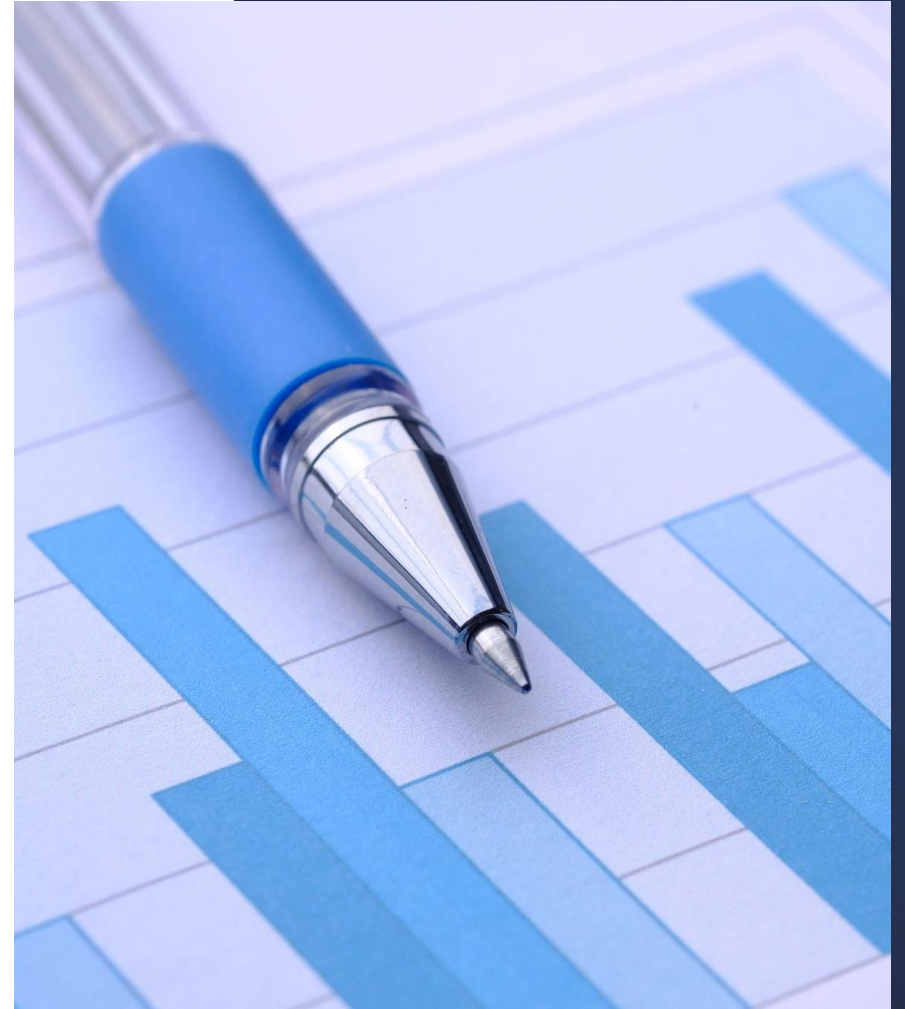
Major factors considered while lending a loan to a Customer are separated in two groups:

S. No.	Loan Attributes	Consumer Attributes
1	Loan Funded Amount.	Annual Income
2	Term of the offered loan.	Employment Length
3	Interest Rate	Purpose of the Loan
4	Loan Grade	Home ownership

Assumptions:

Customers whose Employment Title is not mentioned but their employment length is greater than 0 are considered as "Self Employed".

Employment time for customers with 10+ is considered as 11 Years and below 1 is considered as 0.

For the Analysis Consideration was made only for Charged off and Fully Paid Customers.

Outliers above 75% have been discarded from Analysis Data.

The Columns having more than 50 % null values have been dropped from analysis data.

Our analysis approach consists of below steps:

1. Understand the data through data dictionary. And identify the consumer attributes and loan attributes which would determine the recommendations after analysis.

2. Identify data quality issues and handle the data in such a way that we don't lose the original information and don't impute misleading information.

3. Fill the Missing Values , remove insignificant /duplicate information ,standardise the values inside columns, and derive few more columns based on existing columns to simply the analysis.

4. Identify and Remove the outliers from major attributes to provide effective Analysis results.

5. Do Univariate Analysis to understand the columns/ attribute deeper.

6. Differentiate the customers based on the Loan status in the dataset.

7. Do Segmented Univariate Analysis to understand how the target behaviour varies in groups.

8. Do Bivariate Analysis to get knowledge on how one attribute influence the others.

9. Multivariate Analysis was done to find out the correlation among various non categorical attributes.

Analysis Approach

DATA UNDERSTANDING

Understanding the Borrowers:

- We see that majority of loan Borrowers don't own a house but have rented a house.
- Annual income for half of the borrowers (middle)lies between 34k to 60k.
- Loan fund amount Vs annual Income ratio majorly lies between 0.1 to 0.3.

Loan Status Information of Borrowers:

- 85% of Loan borrowers pays their loan fully without any loss to the company.
- Due to 15 % of Loan borrowers do not pay the loan properly, company incur the loss.
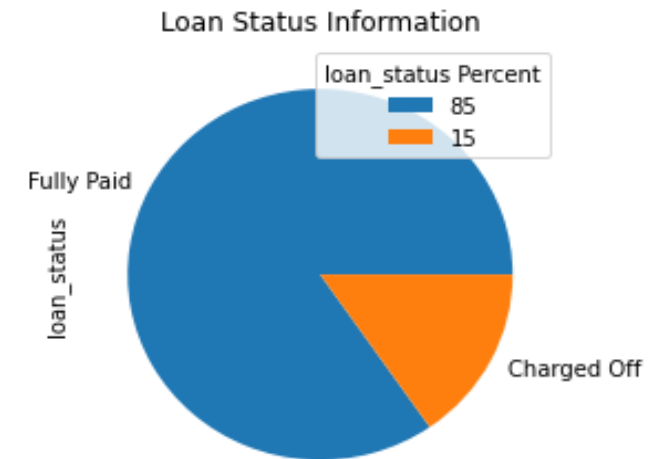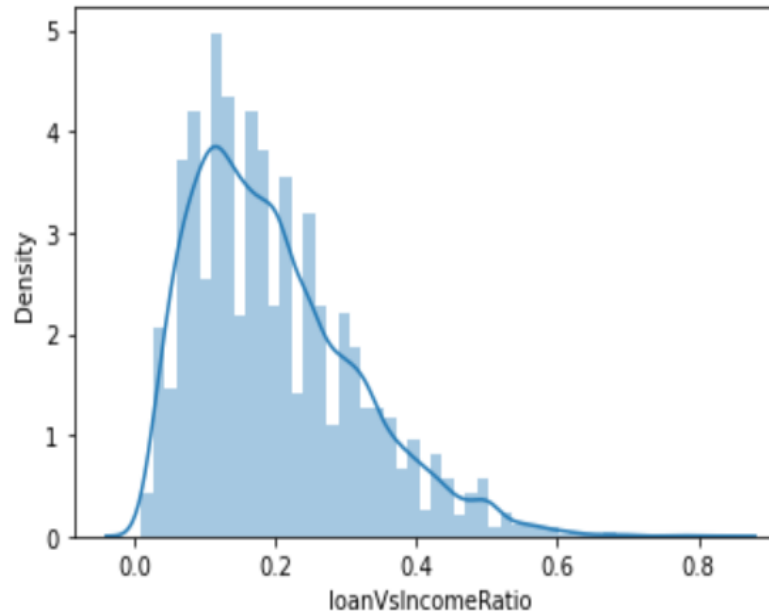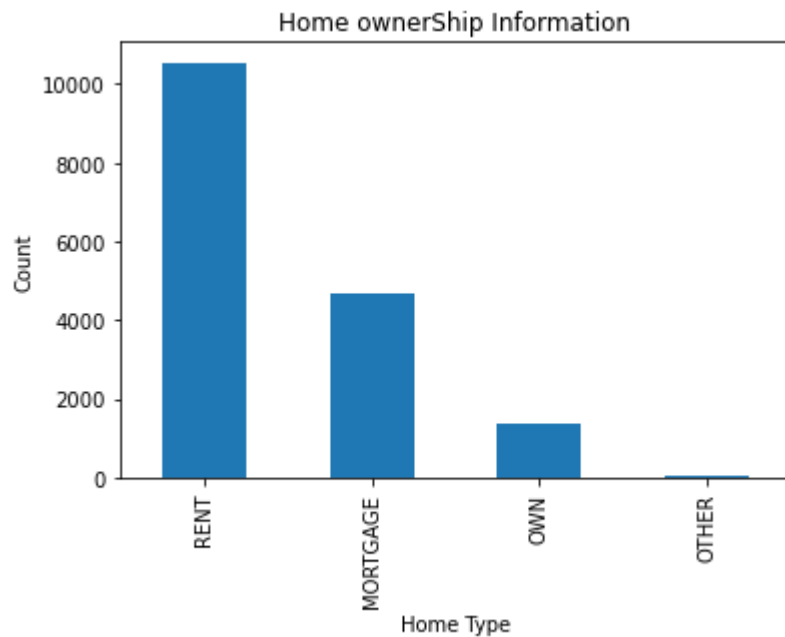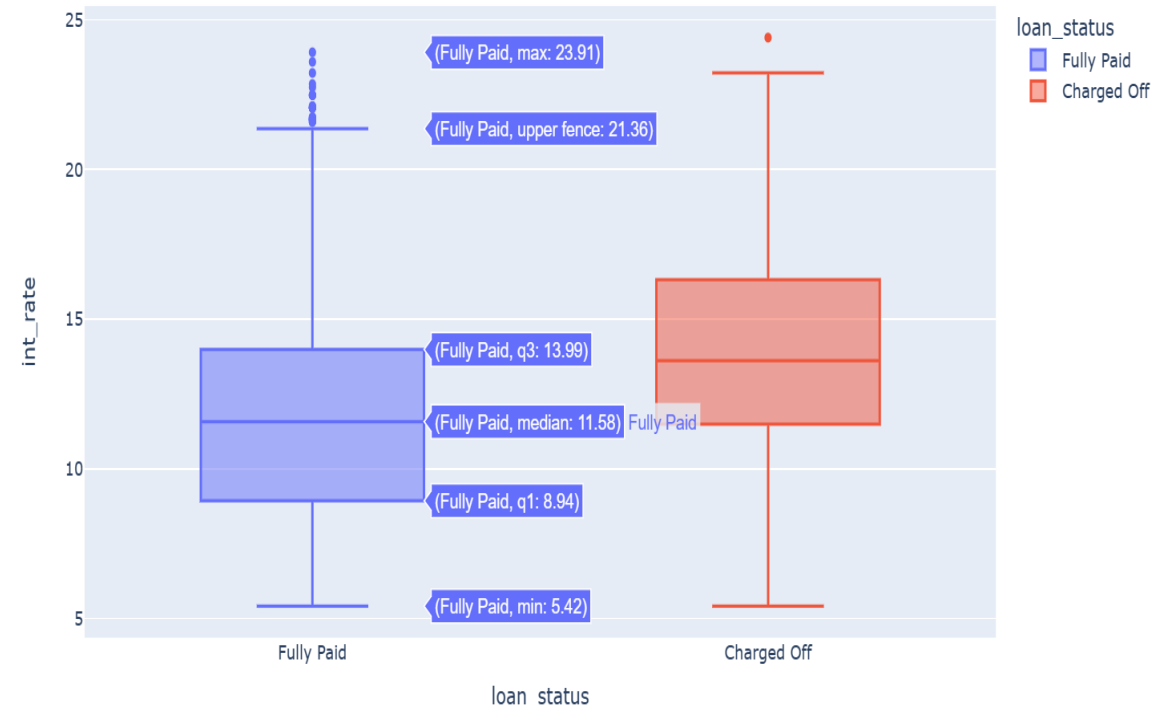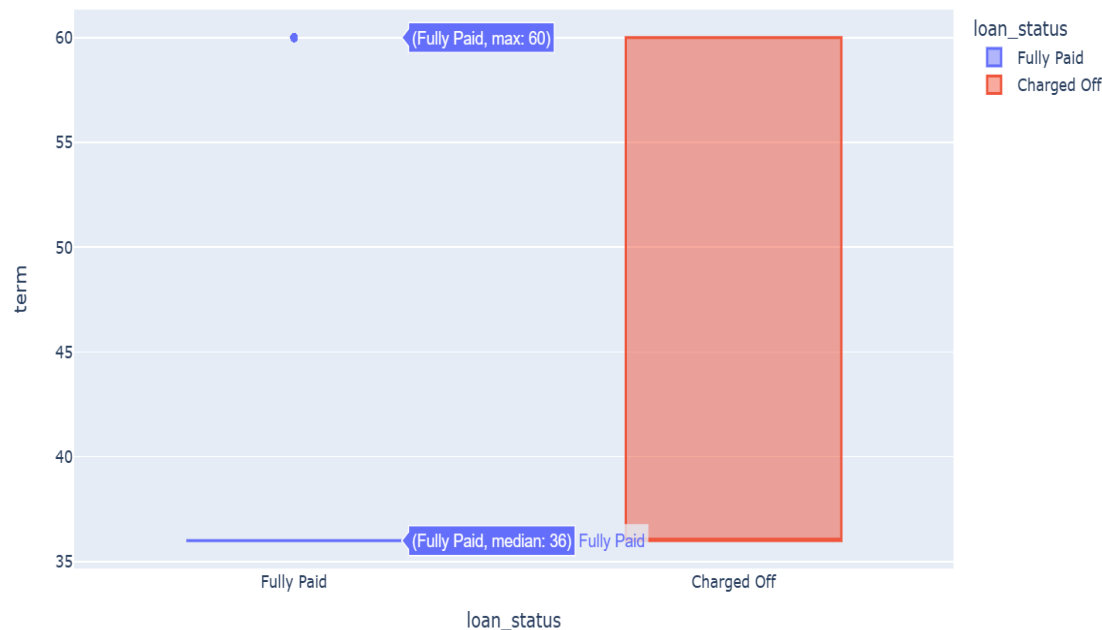
Factors Influencing Borrowers

On further Analysis on 15% of defaulters. We concluded some factors plays the major role in influencing the borrowers.

Term period of the Loan offered.

• The chance of loan getting defaulted is less for 36 month periods than 60 months as shown below.

Interest Rate of the Loan offered.

• It has been observed that, if the interest rate is below 10, there are minimum chances of borrowers being defaulted.

• If the interest rate is above 14%, the borrowers need to pay high amount gradually it defaults to a defaulter category.

Factors influencing Borrowers

- We also observed that Grade of a borrower helps in Understanding the defaulter's category.

- Borrower with Grade A are less likely to be Defaulters than people in Grade B and so on moving towards Grade G

0.068 Grade A

0.12 Grade B

0.18 Grade C

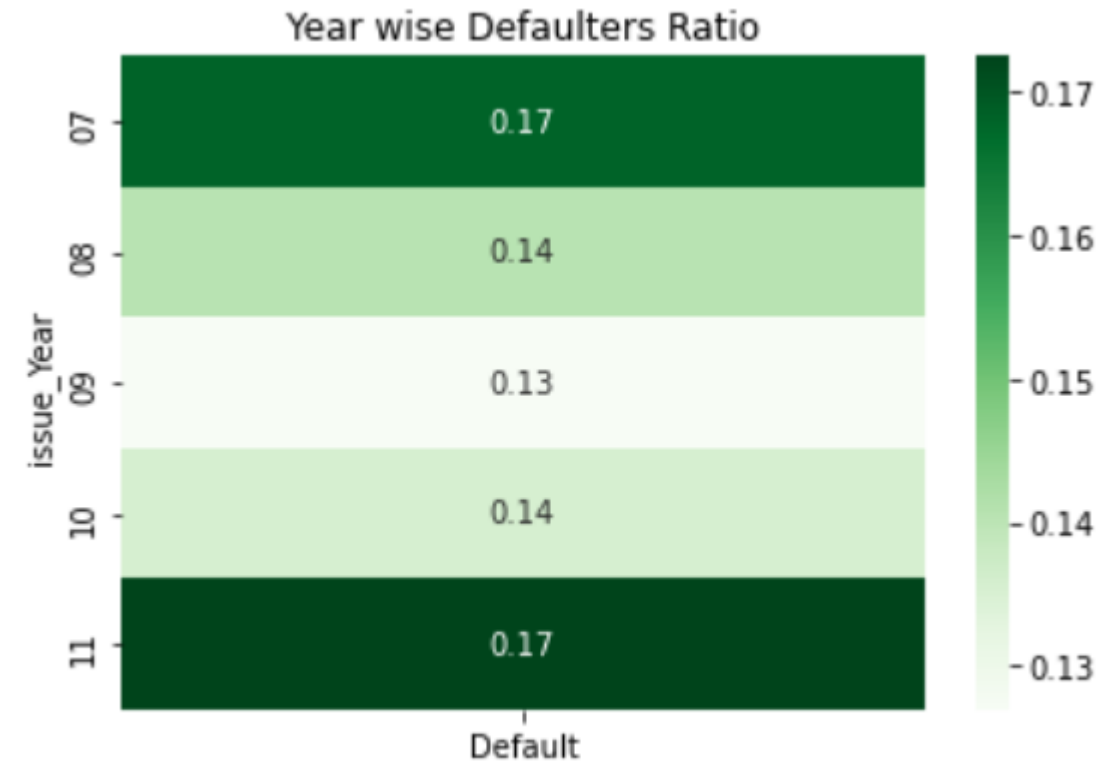0.23 Grade D
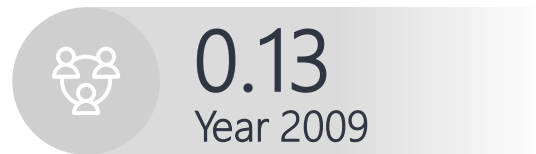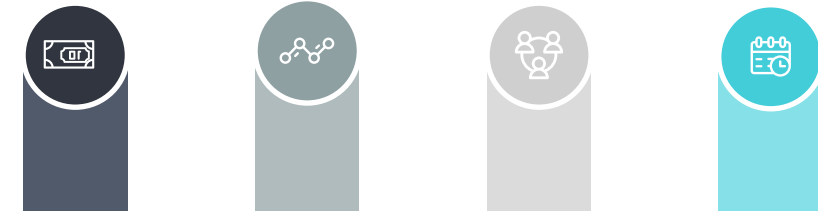
Grade wise Defaulters Ratio
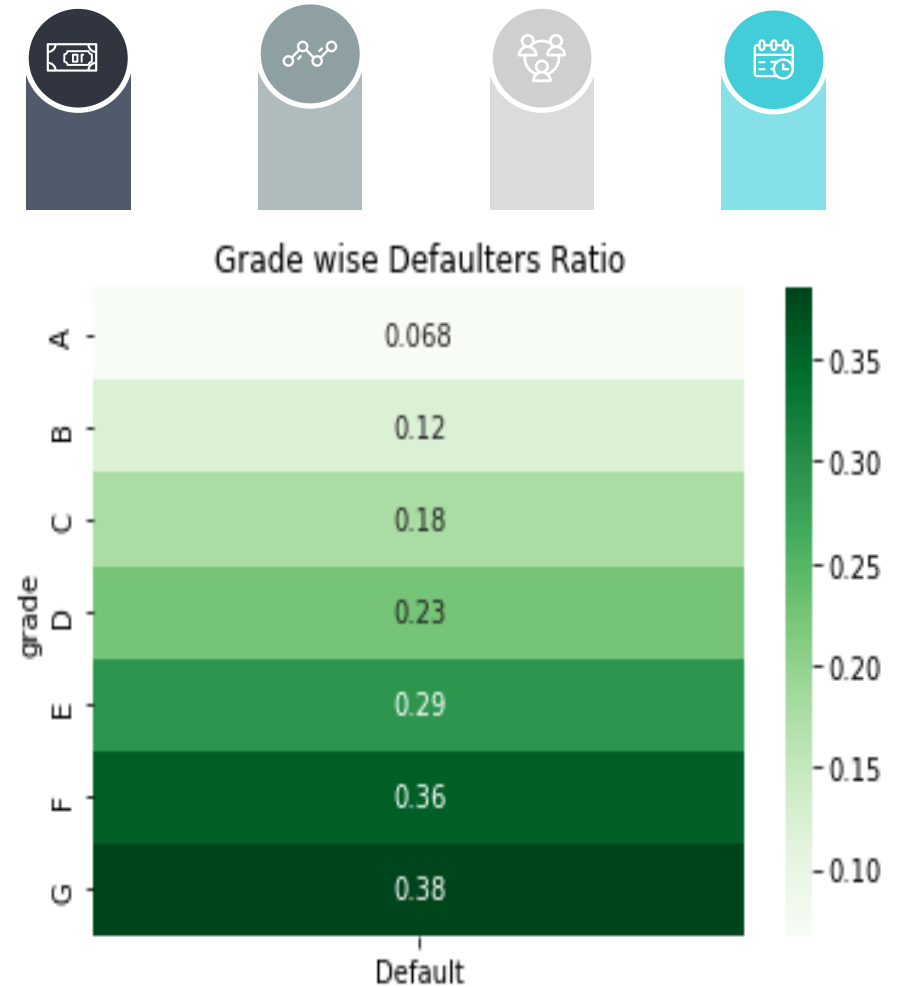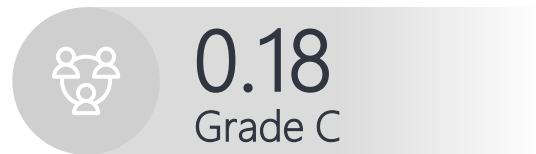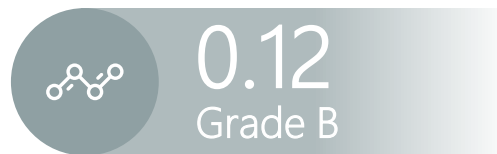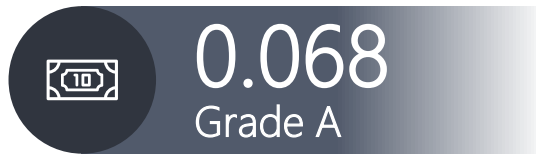
External Factors influencing Borrowers

From the below Heatmap for 'Years' we can infer that the ratio of defaulters was quite high in Year 2007 and 2011.

- This information can be used to identify the other factors during these years which led to increased number of defaulter.

0.17 Year 2007

0.17 Year 2011

0.13 Year 2009

0.14 Year 2008-10

Year wise Defaulters Ratio

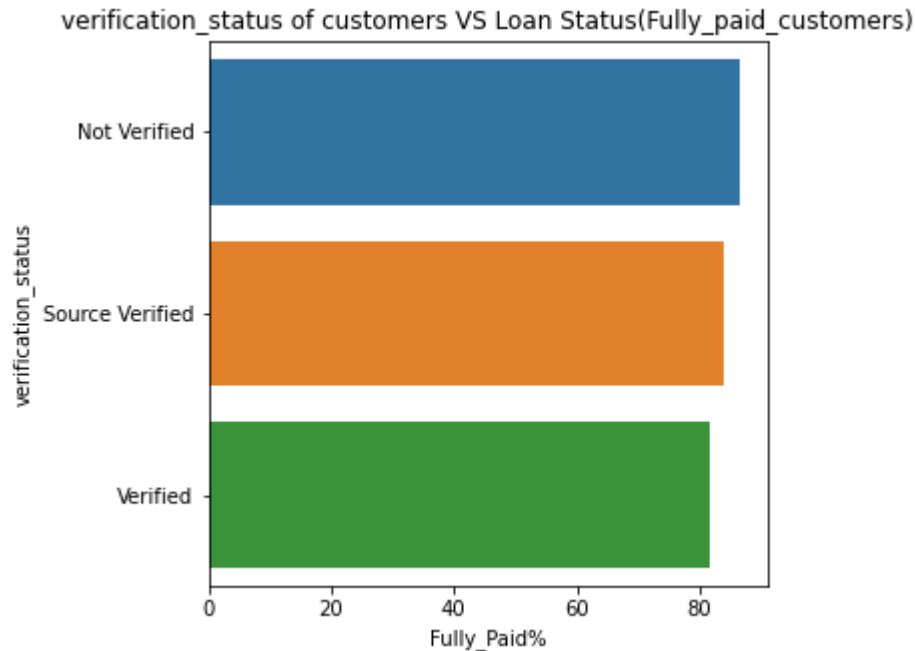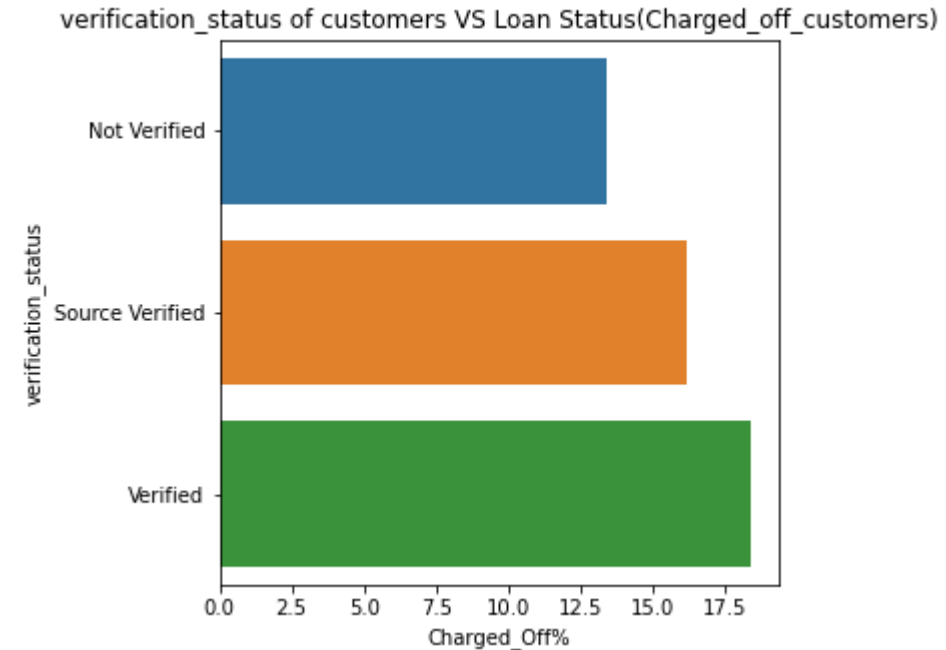| issue_Year	Default
07	0.17
08	0.14
09	0.13
10	0.14
11	0.17

Comparing two factors for the Analysis

Verification status for
fully paid Borrowers

Verification status for charged
off Borrowers

Normally Loan approver assumes that probability of verified customers in defaulter's category will be low and approves the loan. But during our analysis we interpreted that even there is a high possibility of verified customers to be in defaulter's category.

Comparing two factors for the Analyses
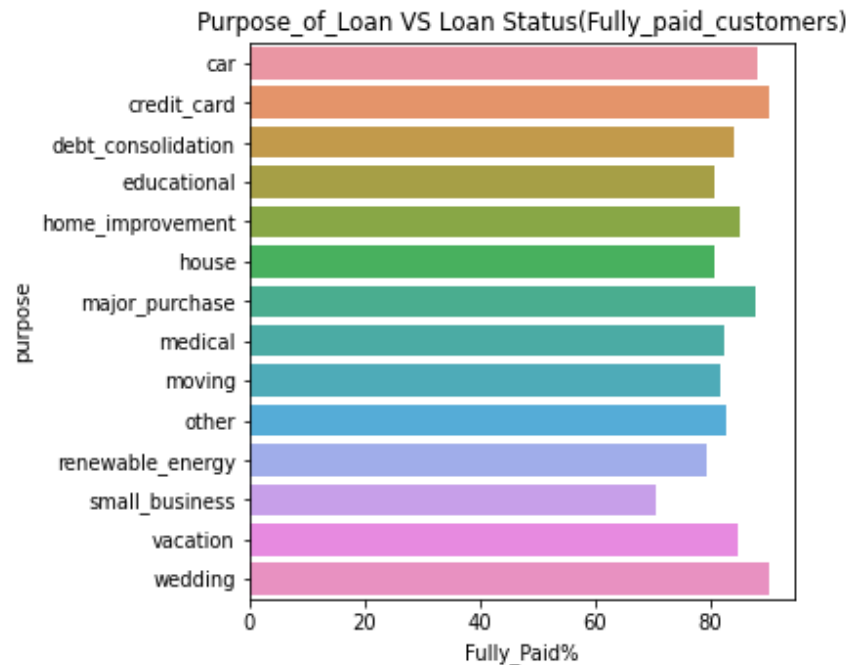
Understanding the Borrowers:

- Loan fund amount Vs annual Income ratio lies between 0.1 to 0.3
- For high value of the Loan Vs income Ratio , there will be slightly higher tendency of Borrower Defaulting the Loan
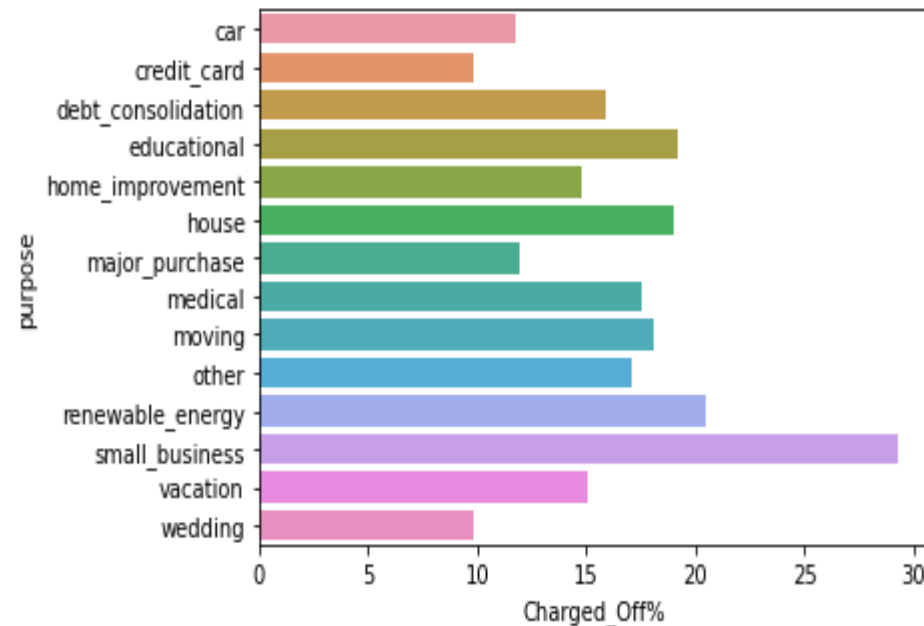
Comparing two factors for the Analysis

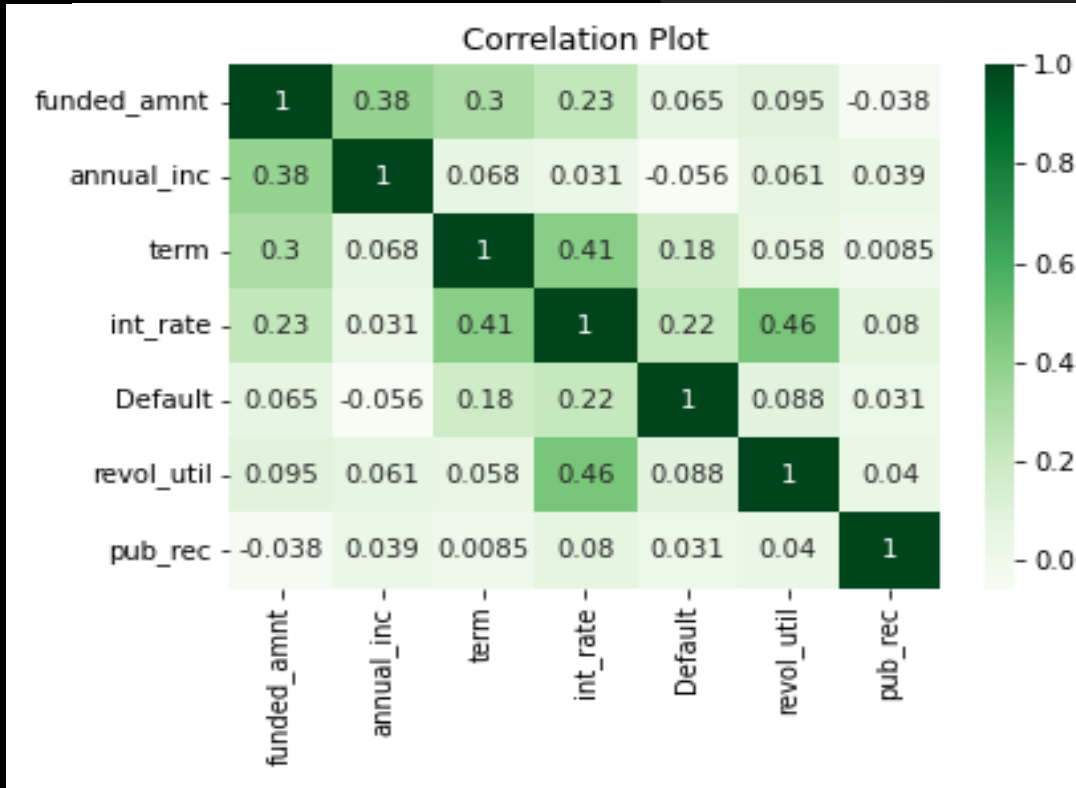Purpose of Loan for Fully Paid Customers

Purpose of Loan for charged off Customers

Observations:

•Customers opted loan for purpose of **car, credit card, major purchase and wedding** are likely to **pay the loan fully.**

•People who opted loan for **Small business** are not likely to pay the loan. Hence approving loan to small business can be done carefully.

4

Correlation Plot

Conclusive Analysis

- From the Heat map we observed the negative correlation between Tendency to default and Annual income. It means that borrowers with higher annual income are less likely to default.

- Positive correlation has been observed between term and tendency to default i. e. Keeping the term length short may be beneficial for the cooperation to reduce the tendency to default.

- Positive correlation between interest rate and tendency to default. i.e. The loan with lower interest rate are less likely to be defaulted.

- Positive correlation between Annual income , interest rate and funded amount shows that Financial cooperation funds high amount loan to those with higher annual income and higher loan amount needs high term plan to get repaid. And financial cooperation imposes higher interest rate on long term plan loans.

Recommendations

Loan Attributes
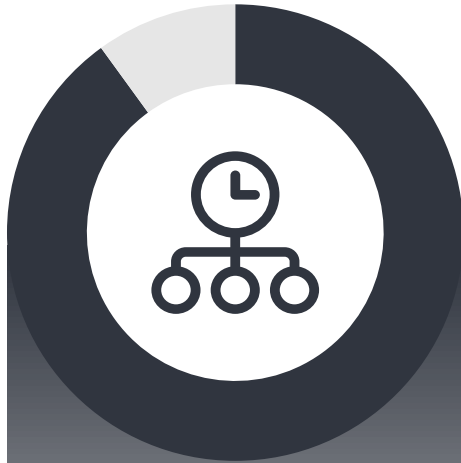
Financial corporation should focus on providing Loans with lower interest rate of higher Grade (A/B) for the short Term.

Keep the approved loan amount such that loan amount vs Income Ratio stays between 0.1 to 0.3.

Consumer Attributes

Customers with Annual Income in range between 35k to 60k and who have rented or Mortgage house are prospective customers.

Avoid approving the loan for Small business purpose.

External Factors

Identify the other economical factors causing increased loan defaults in year 2007 and 2011.Based on those, financial corporation can make strategy to avoid them.