Executive PG Program in ML and AI

ML C35 Lending Club Case Study

Submitted by :- Rahul R

Group Facilitator :- Rahul R

Objective of the Case Study:-

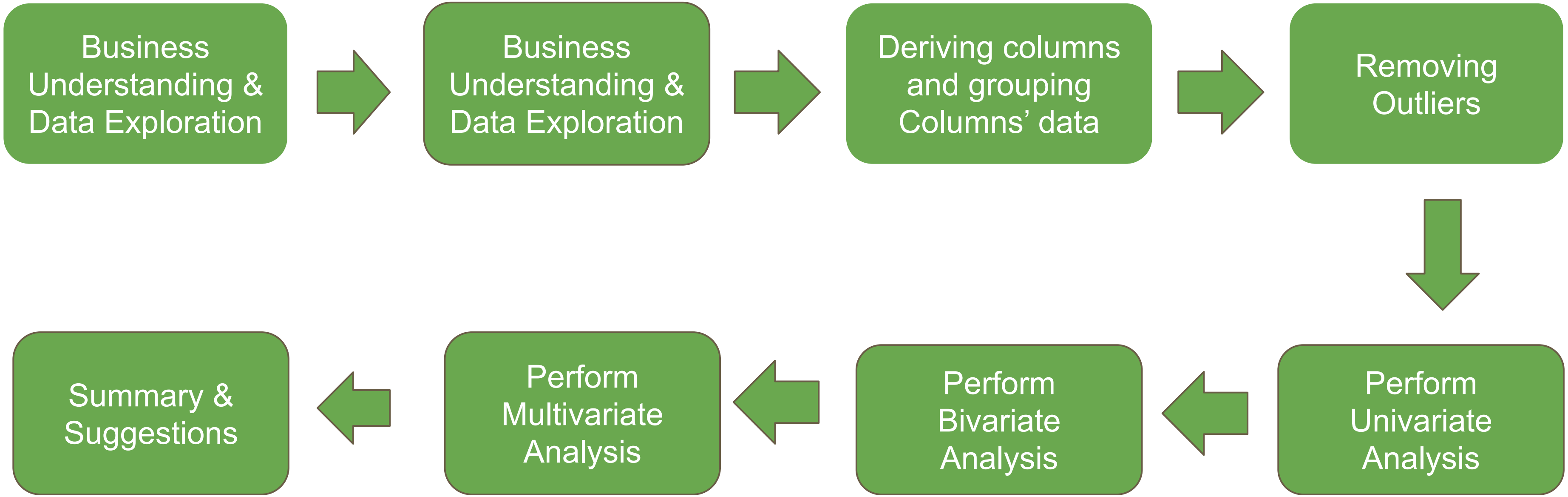
- Lending Club is a marketplace for personal loans It matches the borrowers who are seeking a loan with investors looking to lend money and make a return

- There may be chances of risk occurrence while funding loan to the borrowers These risks could lead to financial loss to the Lending Club

- The main objective of this Case Study is to understand and analyse the driving factors or variables leading to Default risk and provide some recommendations for its prevention or betterment

Approach :-

Below is the approach followed for Case Study:



Data Understanding and Assumptions :-

- In the loan status, there are three options

  A. Fully Paid
  B. Current
  C. Charged Off

- For efficient analysis, current loan status is ignored as we couldn’t draw conclusion whether it would fall under full paid or charged off

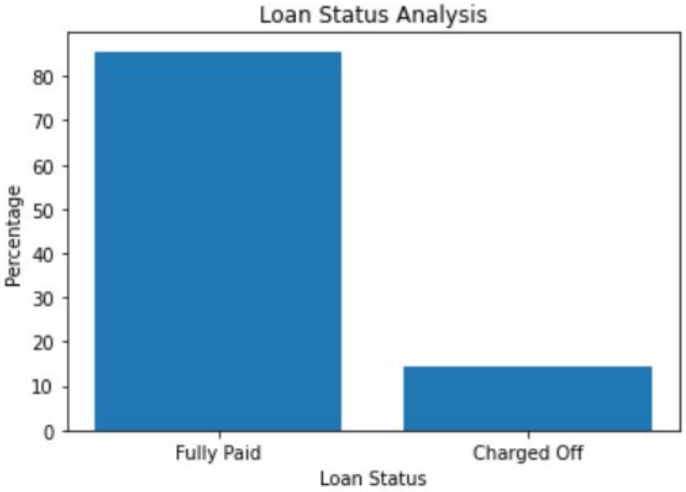
- The customers with Charged off loan status are considered to be the Bad Customers and a high amount of risk could be expected from them

Data Cleaning and Formatting:-

- Removal of columns with null values mostly 90 of null values
- Dropping of rows with loan status as “Current”
- Removal of unnecessary and extra columns
- Removal of Outliers from annual income for better analysis
- Imputing the null tabs with suitable value
- Formatting the data in column (Ex: Removal of ‘%’ sign and reducing 10+ or <5 years of experience to numbers)
- Deriving the separate columns for separating and grouping data
- Categorizing the columns and creating bins for easy analysis

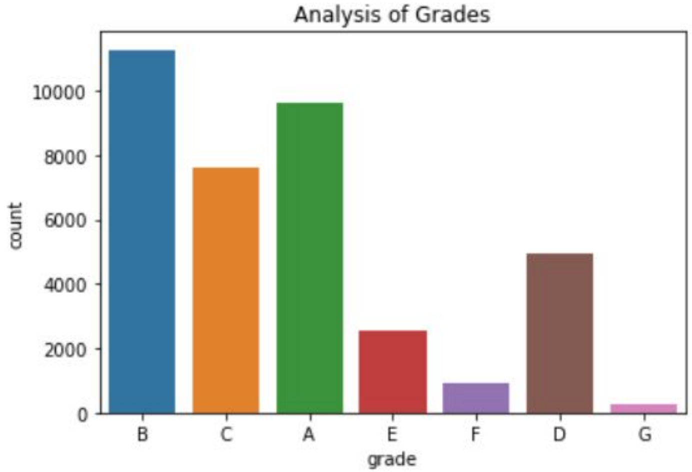
Univariate Analysis:-

Loan Status Analysis



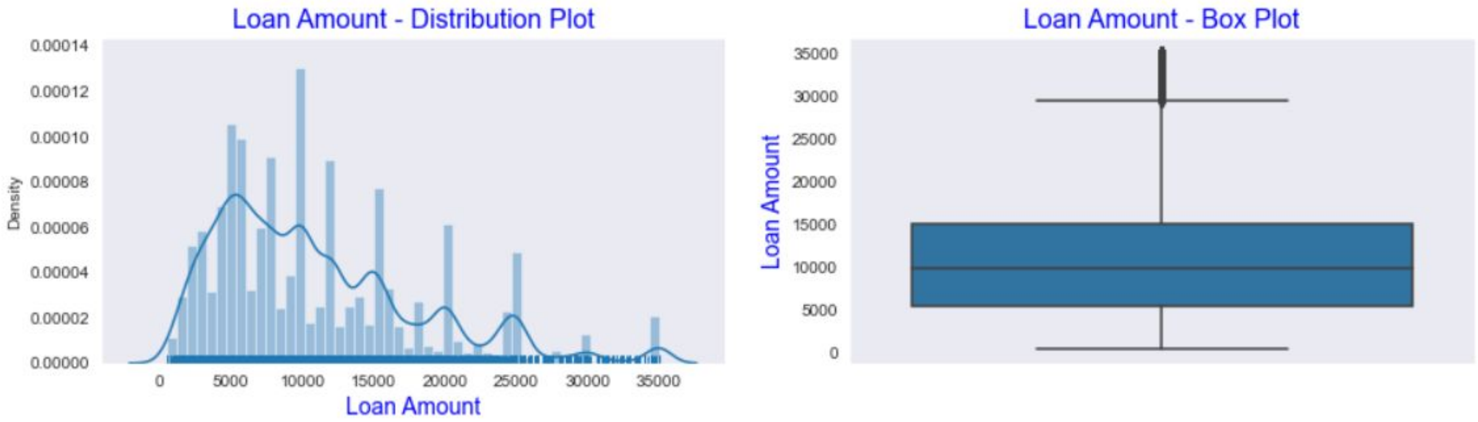
It is observed that the dataset shows 86% of Fully paid and 14% of Charged Off customers.

Grade Analysis



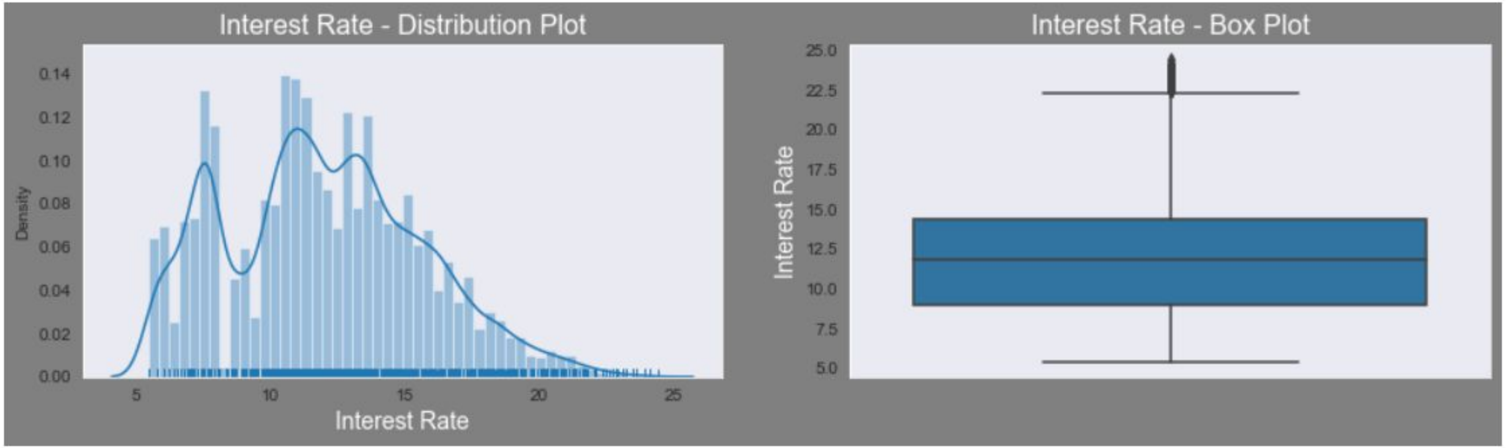
Most of the loans have grade of A and B. Therefore stating most of the loans are high graded loans.

Loan Amount Distribution



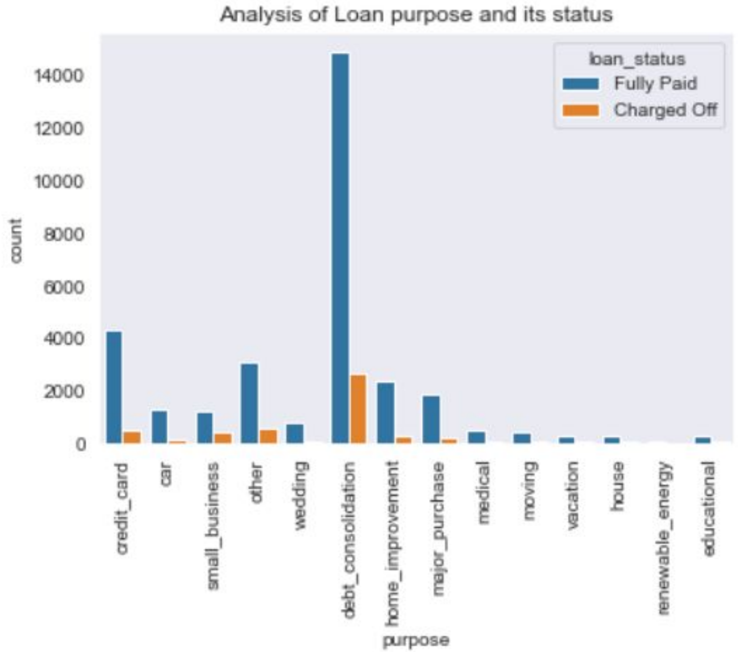
Plots show that most of the Loan amounts are in range of 5000-15000.

Analysis of Rate of Interest



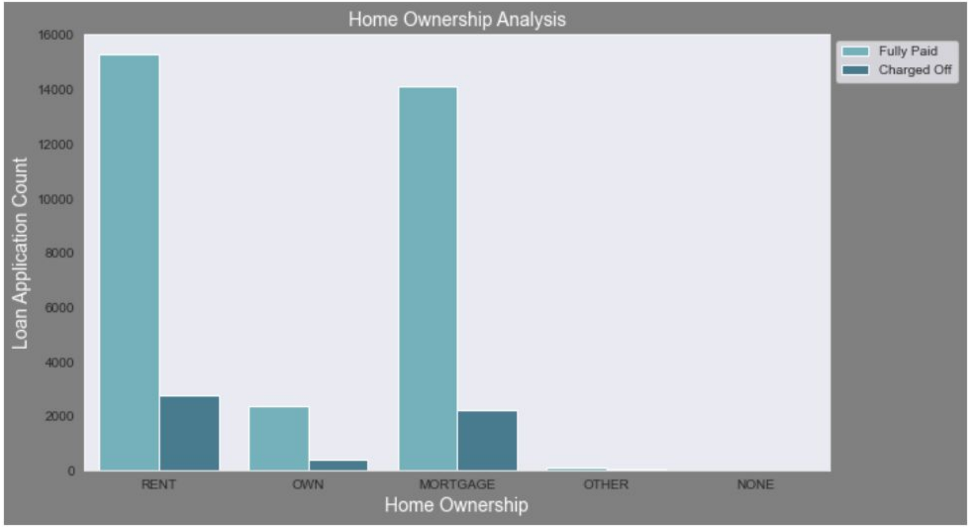
Segmented Univariate Analysis:-

Analysis of Loan Purpose with its status



- Majority of loan has been given for the debt consolidation and credit card purpose and has been fully paid.
- Charged off percentage is also high for debt consolidation.

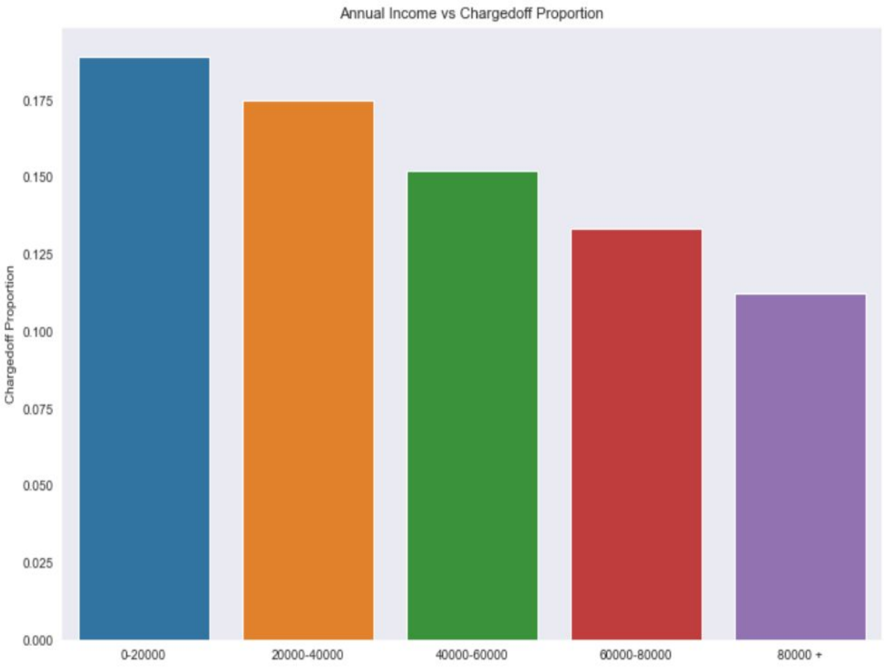
Analysis of Loan Status based on Home Ownership



- The plot shows that most of them living in rented home or mort gazed their home
- Applicant numbers are high from these categories so charged off is high too

Bivariate Analysis:-

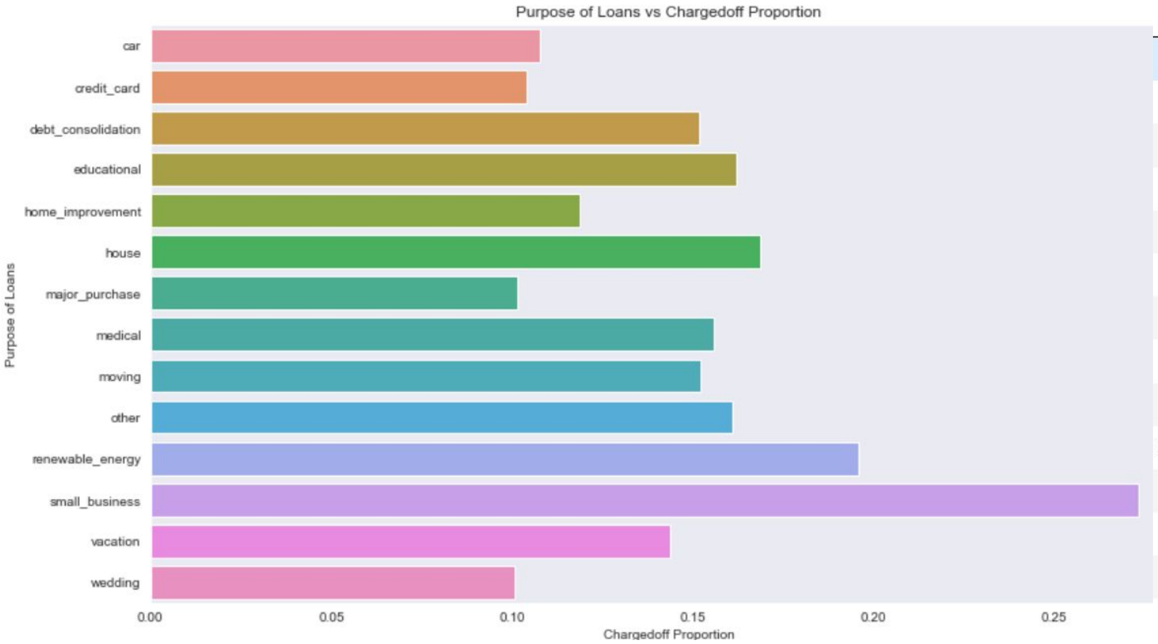
Analysis of Annual Income vs Charged Proportion:



annual_inc_category	Charged Off	Fully Paid	Total	Chargedoff_Proportion
0-20000	191	818	1009	0.189296
20000-40000	1421	6709	8130	0.174785
40000-60000	1675	9326	11001	0.152259
60000-80000	1002	6503	7505	0.133511
80000 +	1068	8452	9520	0.112185

- Income category of 80000 has less chances of charged off
- Income category of 0 20000 has high chances of charged off
- With an increase in annual income, charged off proportion got decreased

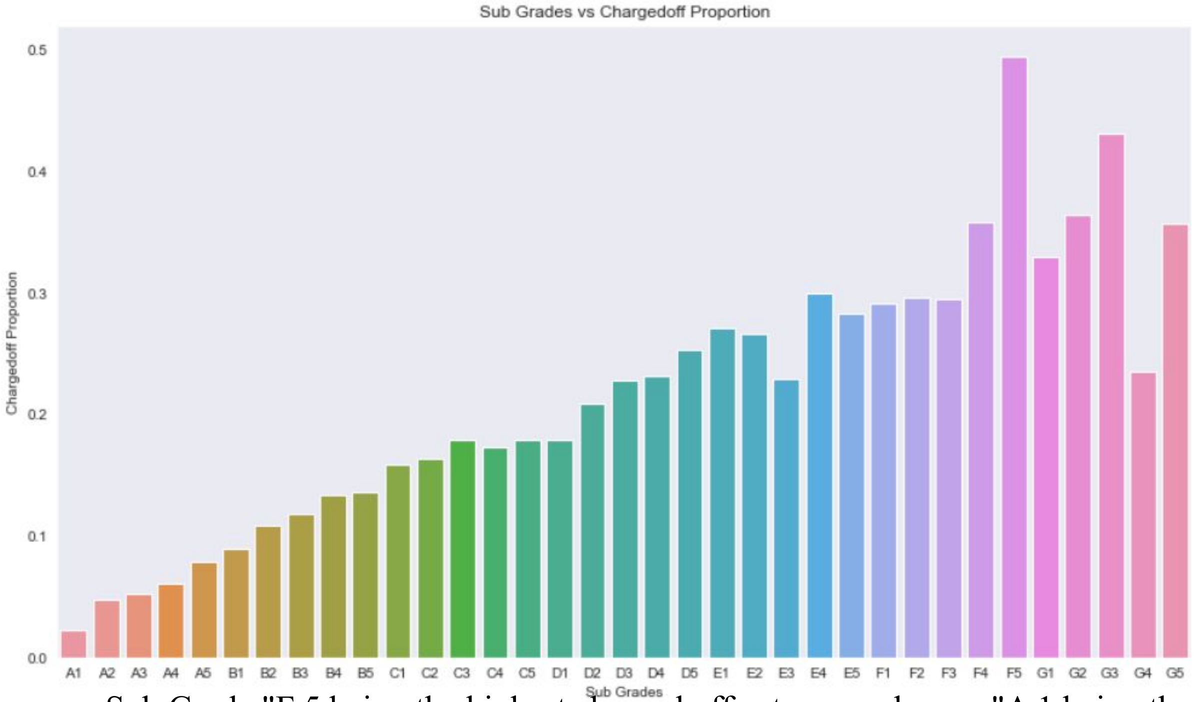
Purpose of loans vs Charged off Proportion



purpose	Charged Off	Fully Paid	Total	Chargedoff_Proportion
small_business	459	1221	1680	0.273214
renewable_energy	18	74	92	0.195652
house	58	286	344	0.168605
educational	51	264	315	0.161905
other	592	3088	3680	0.160870
medical	101	548	649	0.155624
moving	83	463	546	0.152015
debt_consolidation	2664	14875	17539	0.151890
vacation	50	298	348	0.143678
home_improvement	321	2386	2707	0.118581
car	155	1286	1441	0.107564
credit_card	505	4355	4860	0.103909
major_purchase	209	1850	2059	0.101506
wedding	91	814	905	0.100552

- Small Business borrowers have high chances of getting charged off
- Renewable Energy applicants are in second place of getting charged off
- Applicants who borrow loan for wedding has least chances of getting charged off

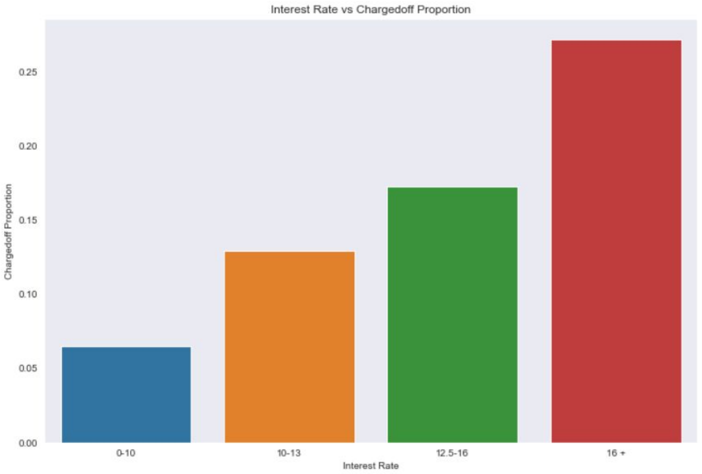
Analysing Sub grades against Charged Proportion



- Sub Grade "F 5 being the highest charged off category whereas "A 1 being the least
- Sub Grades of " has very less chances of charged off
- Sub Grades of " and " have very high chances of charged off
- Sub Grades of charged off is increasing with sub grades moving from sub grades of " towards sub grades of "

sub_grade	Charged Off	Fully Paid	Total	Chargedoff_Proportion
A1	24	1021	1045	0.022967
A2	68	1364	1432	0.047486
A3	90	1640	1730	0.052023
A4	168	2605	2773	0.060584
A5	207	2406	2613	0.079219
B1	153	1558	1711	0.089421
B2	208	1714	1922	0.108221
B3	321	2393	2714	0.118276
B4	317	2059	2376	0.133418
B5	346	2198	2544	0.136006
C1	317	1675	1992	0.159137
C2	307	1575	1882	0.163124
C4	202	967	1169	0.172797
C3	258	1181	1439	0.179291
D1	163	745	908	0.179515
C5	199	909	1108	0.179603
D2	259	979	1238	0.209208
D3	246	831	1077	0.228412
E3	113	380	493	0.229209
D4	207	687	894	0.231544
G4	12	39	51	0.235294
D5	204	601	805	0.253416
E2	159	437	596	0.266779
E1	190	511	701	0.271041
E5	103	261	364	0.282967
F1	86	209	295	0.291525
F3	49	117	166	0.295181
F2	67	159	226	0.296460
E4	124	290	414	0.299517
G1	30	61	91	0.329670
G5	10	18	28	0.357143
F4	52	93	145	0.358621
G2	27	47	74	0.364865
G3	19	25	44	0.431818
F5	52	53	105	0.495238

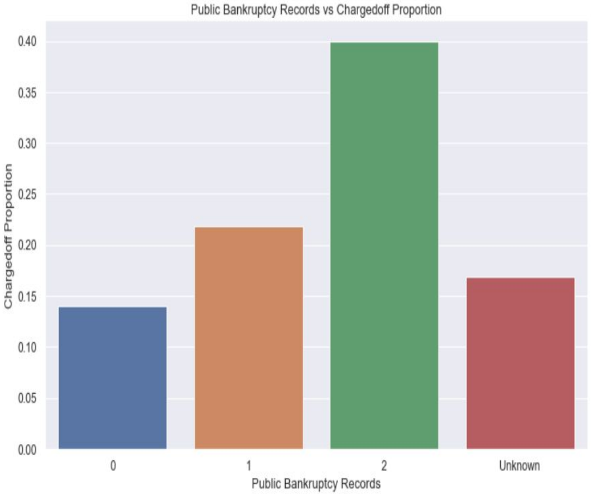
Interest rate vs Charged off proportion



int_rate_category	Charged Off	Fully Paid	Total	Chargedoff_Proportion
0-10	758	10988	11746	0.064533
10-13	1177	7918	9095	0.129412
12.5-16	1923	9204	11127	0.172823
16 +	1214	3251	4465	0.271892

- Interest rate less than 10% has very less chances of charged off Interest rates are starting from minimum 5%.
- Interest rate more than 16% has good chances of charged off as compared to other category interest rates
- Charged off proportion is increasing with higher interest rates

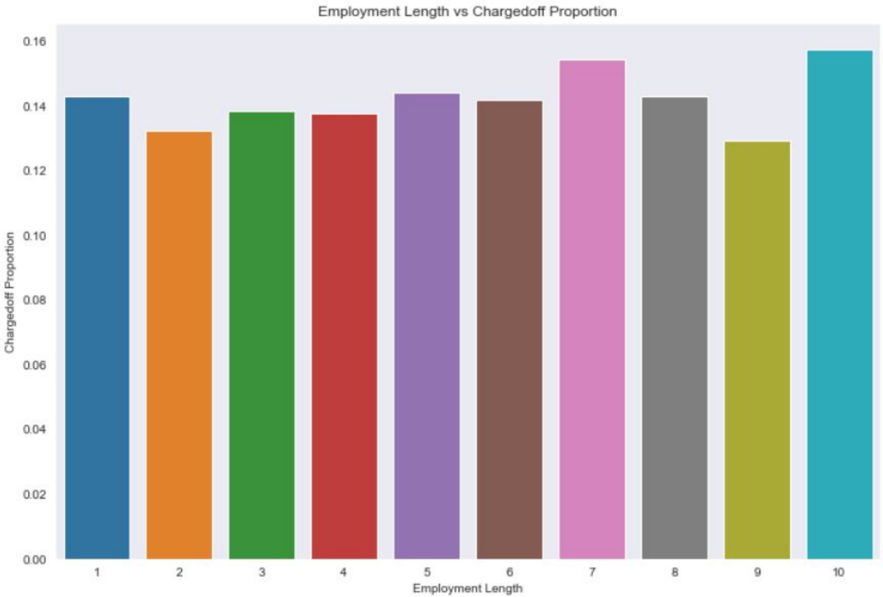
Analyzing Public Record Bankruptcies against Charged off Proportion



pub_rec_bankruptcies	Charged Off	Fully Paid	Total	Chargedoff_Proportion
0	4905	30041	34946	0.140359
Unknown	116	572	688	0.168605
1	334	1192	1526	0.218873
2	2	3	5	0.400000

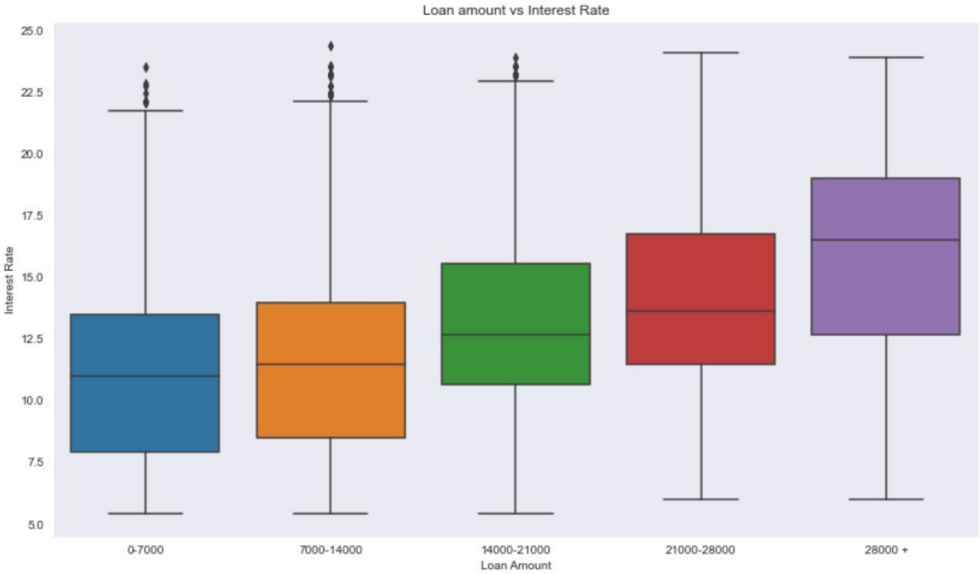
- Those who already have pub_rec_bankruptcies value 1, have charged off proportion higher than who have no pub_rec_bankruptcies.
- public_record_bankruptcies count 2 has even higher charged off proportion but those numbers are not significant to decide.
- Unknown is the column for which we don't have any information about borrower.

Employment Length vs Charged off Proportion



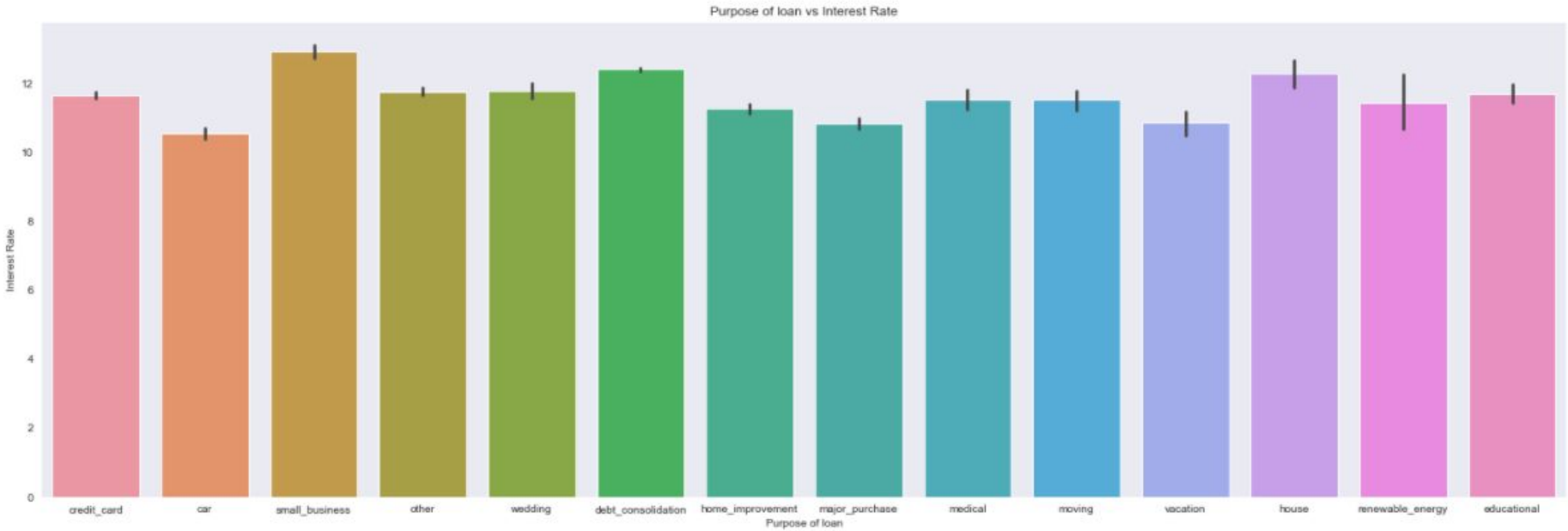
- Borrowers who are not working or have less than 1 year of work experience have high chances of getting charged off
- Probably borrowers with less or no experience they don't have source of income to repay loan
- Rest of the applicants have more or less same chances of getting charged off

Loan Amount vs Interest Rate



- The Interest Rate Increases with the loan amount
- Probably for more loan amount, the term would be more and the interest rate as well varies accordingly

Analyzing the Purpose of loan with Interest Rate

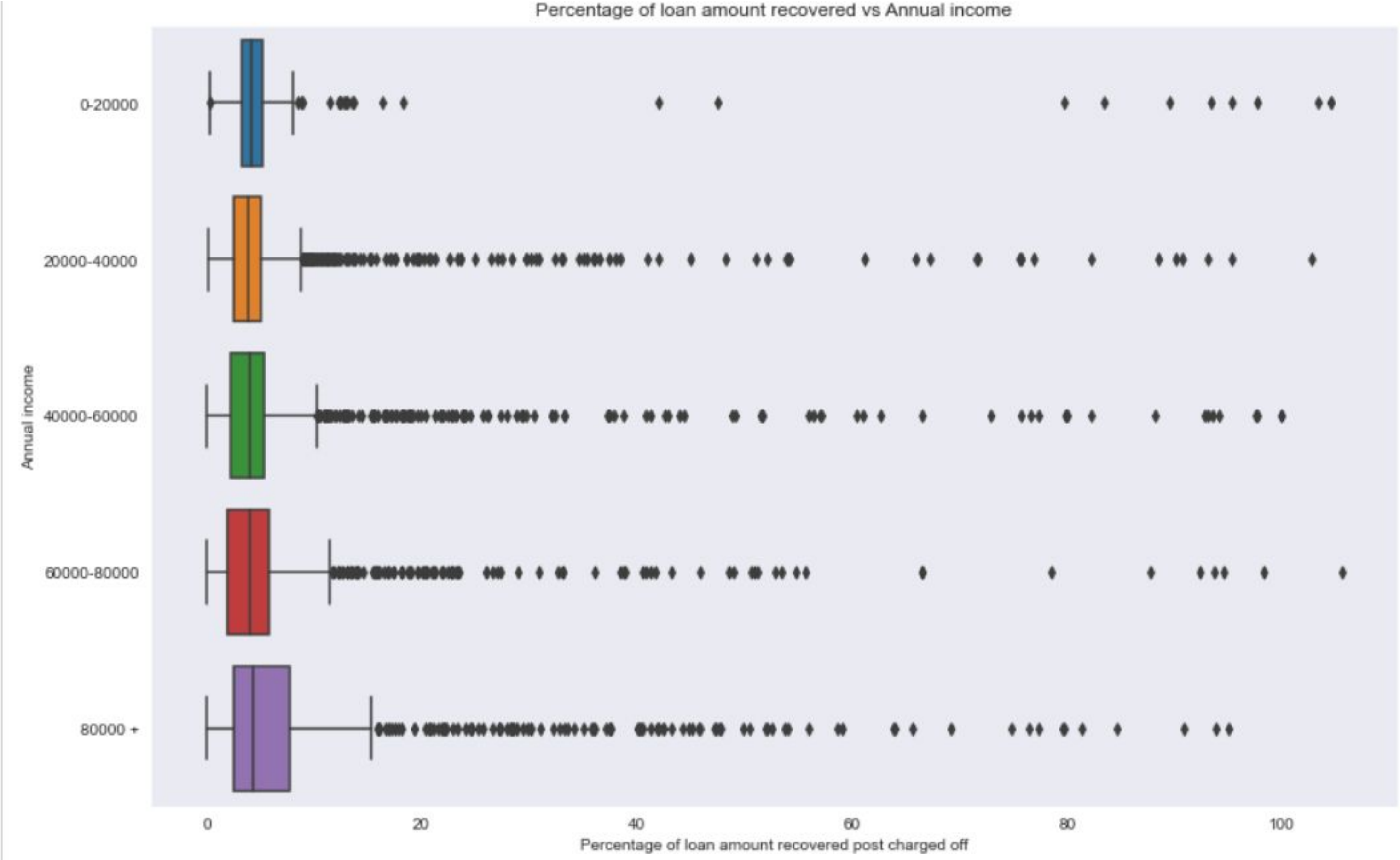


- Small Business applicants pay at a higher interest rate Probably, they take a long term to pay the loan amount The chances of being charged off is high with them

- Debt consolidation is 2 nd where borrowers had to pay more interest rate

- Car loan applicants pay with lesser interest rate comparatively

Analysis of Recovery with Annual Income

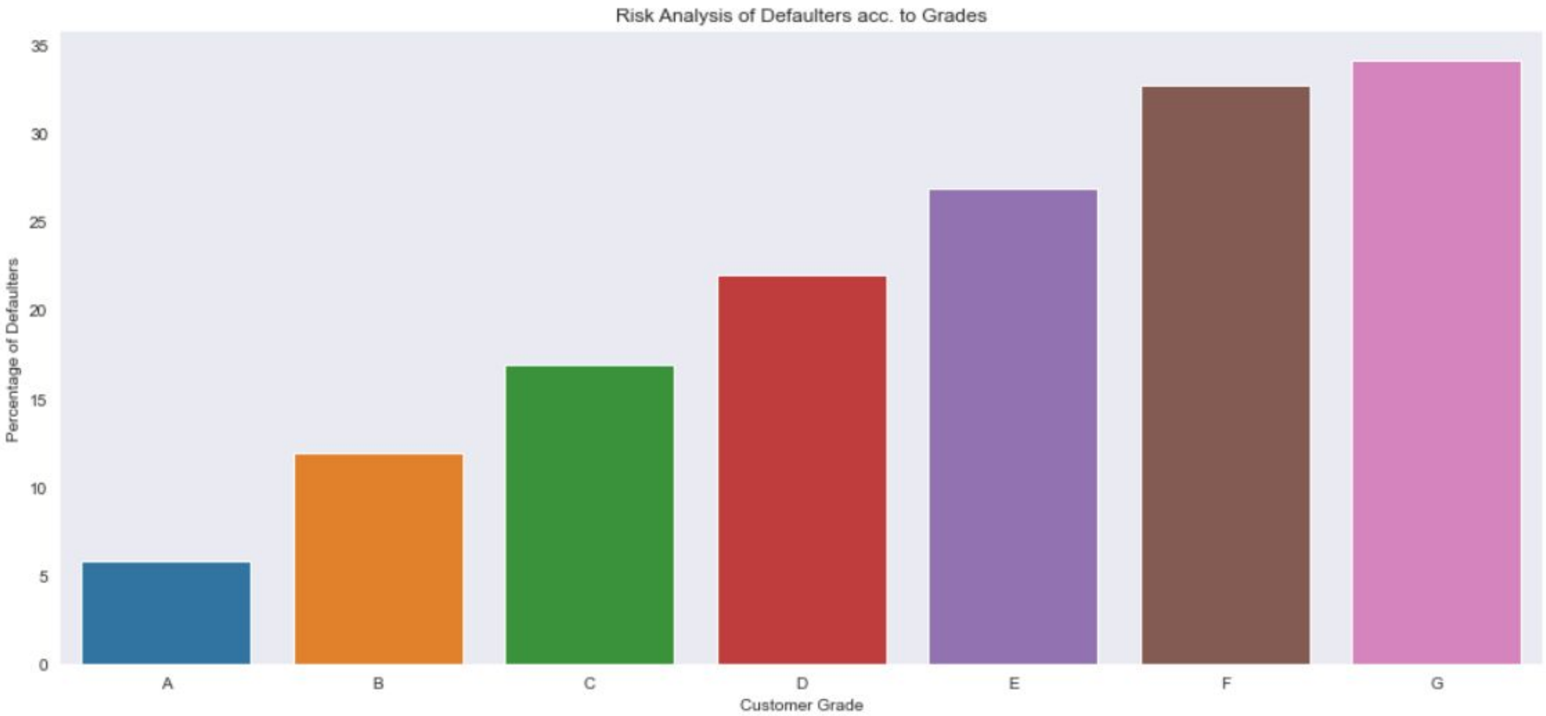


Higher percentage of loan amount is recovered when annual income is high.

Analysing Grade vs Interest rate

- From the above plot it has been visualize that based on grade interest rate is going higher
- “A” is having minimum interest rate, whereas “G” is having higher interest

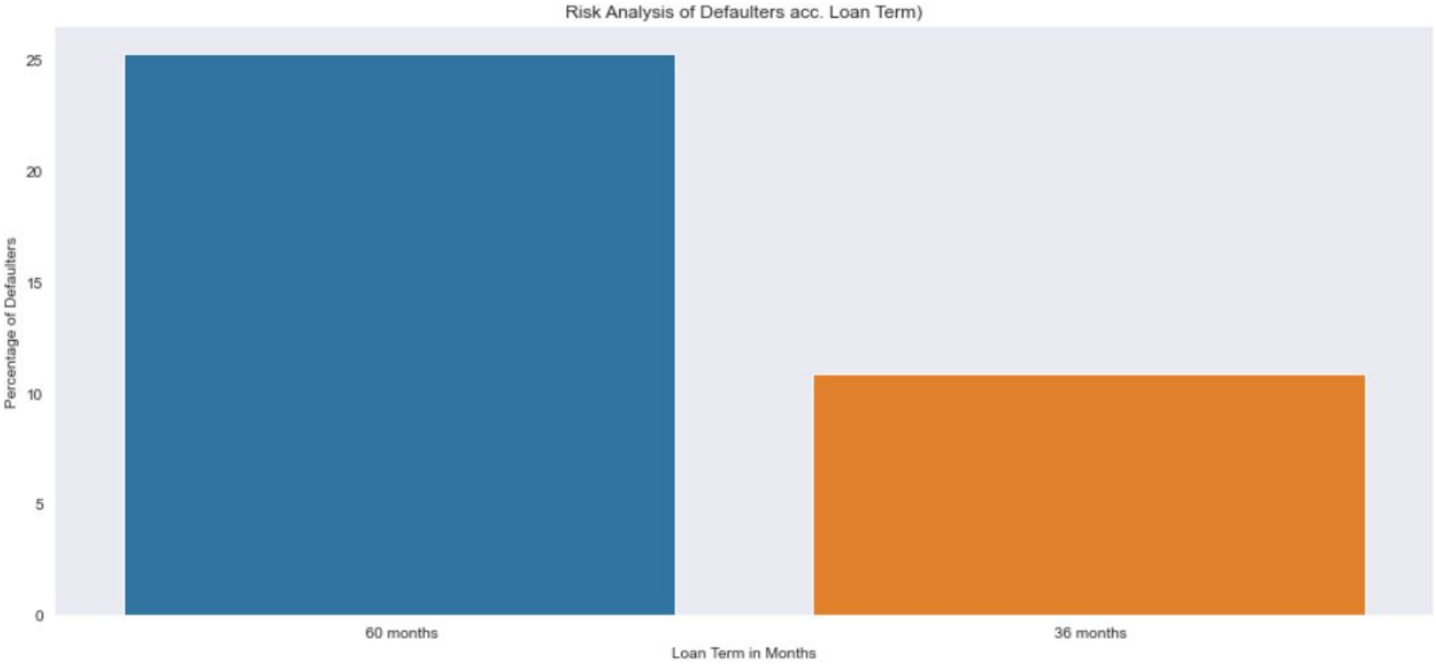
Percentage of Defaulters vs Customer Grade of loan



	defaulter_perc
A	5.81
B	11.94
C	16.90
D	21.92
E	26.83
F	32.66
G	34.03

- Grade G has highest percentage of Defaulters (~ 34%).
- Grade A has the least percentage of Defaulters (~ 6%).
- Grade G loan has highest amount of risk.

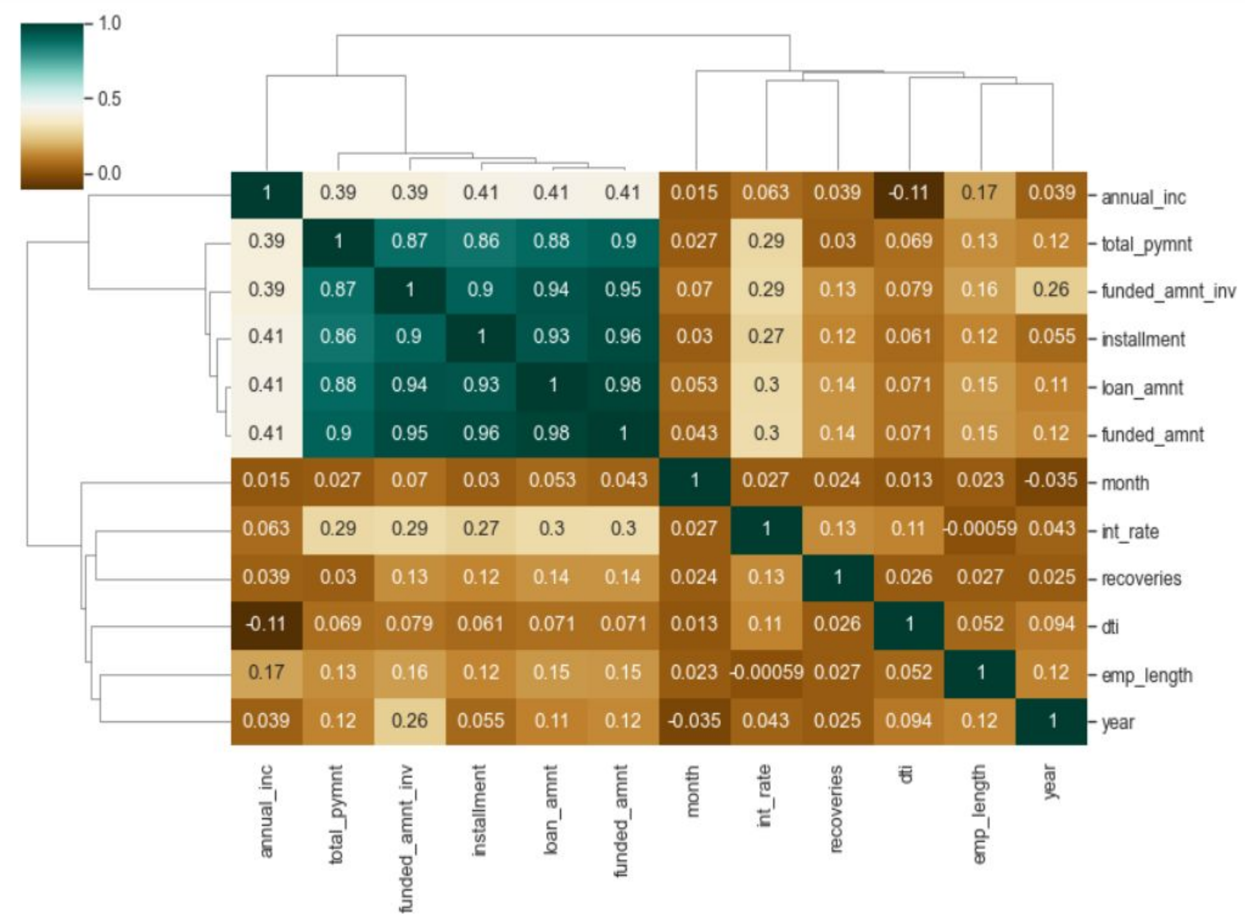
Risk Analysis of Defaulters according to Loan Term



	defaulter_perc
60 months	25.23
36 months	10.88

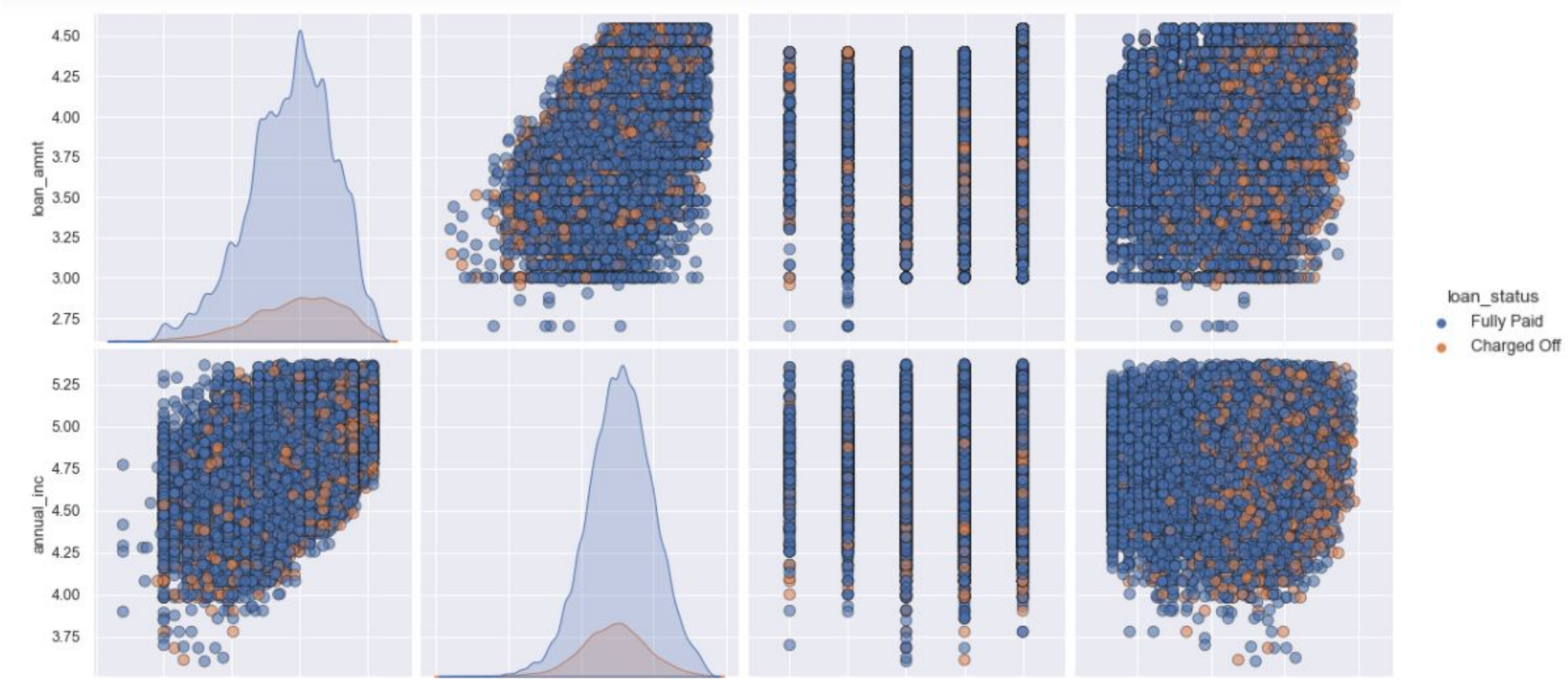
- Percentage of Defaulters is found to be highest in 60 months term (~ 25%) and for 30 months (~ 11%).

Correlation Matrix

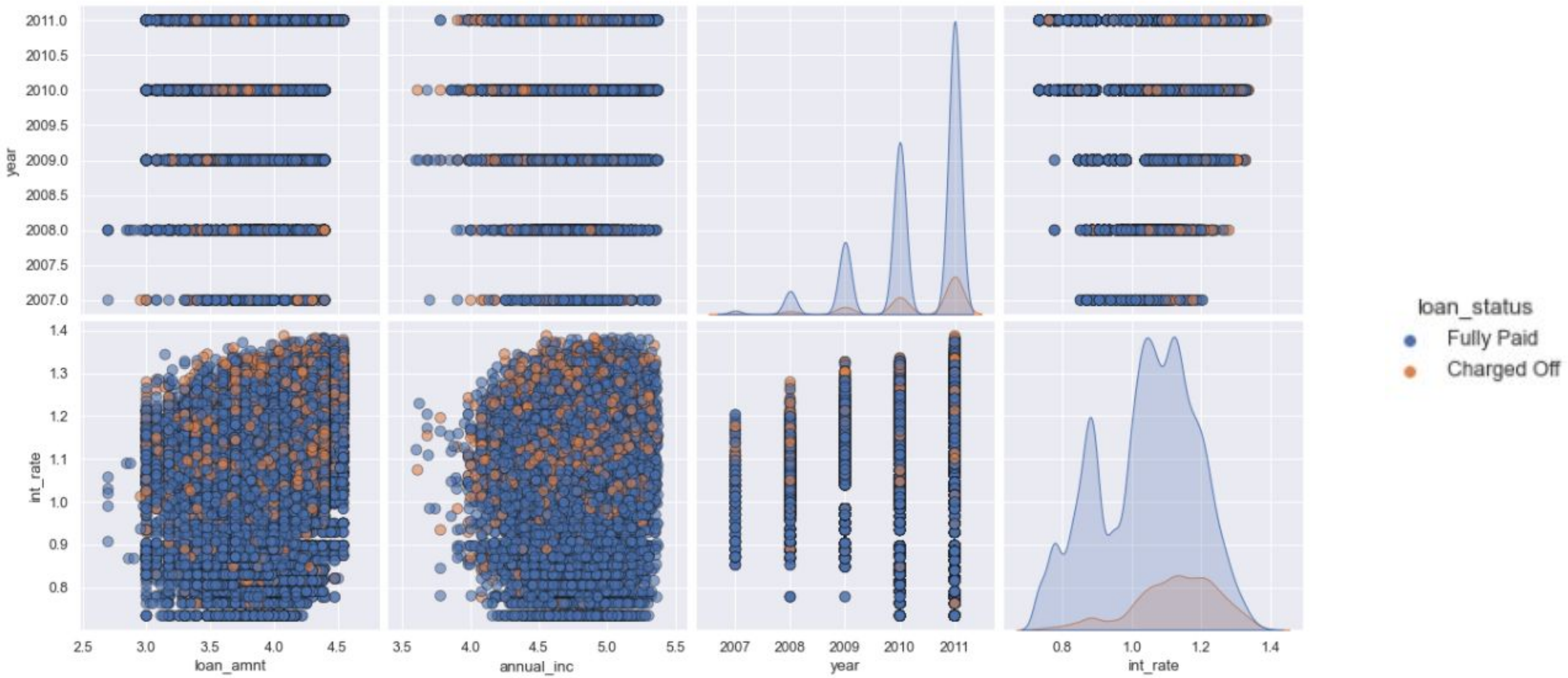


- The Loan amount, investor amount, funding amount are strongly correlated.
- Annual income with DTI(Debt to income ratio) is negatively correlated annual income is low DTI is high & vice versa.
- Debt income ratio is the percentage of a consumer's monthly gross income that goes toward paying debts.
- Positive correlation between annual income and employment years income increases with work experience.

Multivariate Analysis:



- Higher the interest rate higher loan amount.
- Higher the annual income higher the loan amount.



- Increase in number of charged off with increase in year
- Interest rate is increasing as the loan amount increases

Summary:-

- It is observed that the dataset shows 86 of Fully paid and 14 of Charged Off customers
- Customers have lend their loans mostly for debt consolidation(nearly 47 and for credit card(nearly 13
- Most of the borrowers are living in rented home or mort gazed their home
- Borrowers who are not working or have less than 1 year of work experience have high chances of getting charged off
- Small Business borrowers have high chances of getting charged off
- Borrowers most likely to opt for 36 months term and have less chances to be defaulters/charged off compared to 60 months
- Percentage of Defaulters is found to be highest in 60 months term 25 and for 30 months 11
- Interest rate less than 10 has very less chances of charged off
- Grade G has highest percentage of Defaulters 34 and grade A has the least percentage of Defaulters 6
- The Interest Rate Increases with the loan amount and higher percentage of loan amount is recovered when annual income is high

Recommendations:-

- With respect to the Employment length and charged off ratio, it is observed that the many applicants under 1 year or unemployed experience charged off. So it is recommended that the lending club don't provide them with higher loan amount. It also helps applicants with lower interest rate and no charged off problems.

- Lending club should carefully decide while giving loan for applicants with Public Bankruptcy Records.

- Loans for small business applicants should be reconsidered as they fall under charged off. Giving huge amount of loan with higher interest rate lead to their charged off conditions.

- Loan has to be provided taking into consideration of annual income. This results in easy recovery without financial loss.

- Percentage of Defaulters is found to be highest in 60 months term (~ 25%) and for 30 months (~ 11%). So, giving loan amount with shorter term should be beneficial as it would not cause loss.