

Case Study Lending Club

Revan Datta Parimi

Topics

- Important Factors for accepting an application
- Cleaning the Data Set and Statistics
- Analysis of Fully Paid Applicants
- Analysis for Loan Amount, Grade, Annual Income, Purpose and Interest Rate
- Distribution Plots
- Contents in Jupyter Plots

Important Factors for accepting an application

- loan_status – Loan Status
- loan_amnt – Loan Amount
- term – Term for Instalments
- int_rate – Interest Rate
- grade – Grade of Applicant
- sub_grade – Sub Grade of Applicant
- annual_inc – Annual income
- purpose – Purpose of Loan
- dti – ratio of monthly debt to total debt
- emp_length – Employee Experience
- last_pymnt_d – Last Payment Date
- home_ownership – Home Ownership
- verification_status – Verification Status

Cleaning the Data Set

- First we have dropped down the rows with missing values.
- Dropped the rows with irrelevant category like
 - Home Ownership is 'NONE'
 - Loan Status is 'Current'
- Converting the string to numeric values like
 - “36 months” to 36
 - “10.65%” to 10.65
- Removed the Outliers (Extreme max and min values)
- Now the rows reduced from 39,717 to 34,293 (13.6% rows are cleaned)

Data Set Statistics

	loan_amnt	term	int_rate	annual_inc	dti	emp_length
count	37475.000000	37475.000000	37475.000000	3.747500e+04	37475.000000	37475.000000
mean	11124.050033	41.918185	11.959722	6.944315e+04	13.286631	5.168859
std	7352.633429	10.344776	3.682582	6.472065e+04	6.659910	3.888199
min	500.000000	36.000000	5.420000	4.000000e+03	0.000000	0.000000
25%	5500.000000	36.000000	8.940000	4.100000e+04	8.160000	2.000000
50%	10000.000000	36.000000	11.830000	6.000000e+04	13.390000	4.000000
75%	15000.000000	36.000000	14.420000	8.300000e+04	18.550000	9.000000
max	35000.000000	60.000000	24.400000	6.000000e+06	29.990000	11.000000

Analysis of Fully Paid Applicants

- Out of all the categories, we need to check the applicants with loan status
 - Fully Paid – Low Risk Applicants
 - Current – Cannot be considered as Fully Paid or Charged Off, so removed in data cleaning
 - Charged Off – High Risk Applicants as they have a history of not paying on time.
- The Complete Analysis depends on the comparison of applicants category details based on ‘Fully Paid’ and “Charged Off” loan status

Loan Amount

- Loan 0 - 5000 = 86.78%
- Loan 5000 - 10000 = 87.4%
- Loan 10000 - 15000 = 86.38%
- Loan 15000 - 20000 = 82.38%
- Loan 20000 - 25000 = 80.44%
- Loan 25000 - 30000 = 78.83%

There is a proportional decrease with the increase in Loan

Assumption

- 80% being the margin of risk
- Loan Beyond 25k is risk

Grade

- Grade A - 94.13%
- Grade B - 88.07%
- Grade C - 83.13%
- Grade D - 78.3%
- Grade E - 73.71%
- Grade F - 68.46%
- Grade G - 67.01%

- We can see there is an proportional decrease from Grade A Applicants to Grade G Applicants

Assumption

- 70% as the safer risk Margin
- Grade A to E are good to process loans

Annual Income

- Annual Income 0 - 20000 = 81.8%
- Annual Income 20000 - 40000 = 82.73%
- Annual Income 40000 - 60000 = 85.04%
- Annual Income 60000 - 80000 = 87.35%
- Annual Income 80000 - 100000 = 89.28%
- Annual Income 100000-120000 = 90.58%
- Annual Income 100000+ = 90.25%

We can see Gradual increase of Fully Paid customers

Assumption

- 85% being the same margin
- Annual income more than 40,000 is preferred

Purpose

- | | | | |
|----------------------|--------|-------------------------|----------------------|
| • Car | 89.1% | • Medical | 84.44% |
| • Credit Card | 89.74% | • Moving | 84.44% |
| • Debt Consolidation | 85.22% | • Other | 84.38% |
| • Educational | 83.88% | • Renewable Energy | 82.14% |
| • Home Improvement | 87.87% | • Small Business | 73.53% (Risk) |
| • House | 84.24% | • Vacation | 85.5% |
| • Major Purchase | 89.85% | • Wedding | 90.57% |

Interest Rate

- Interest Rate 5 - 10 % = 93.48%
- Interest Rate 10 – 15% = 85.44%
- Interest Rate 15 – 20 % = 75.54%
- Interest Rate 20 – 25 % = 61.14%

We can the Proportion to Low Interest Rate

Assumption

- 75% as margin of safer risk
- Interest rate beyond 20 is applicable for higher Annual Income and Grade

Factors Analysis

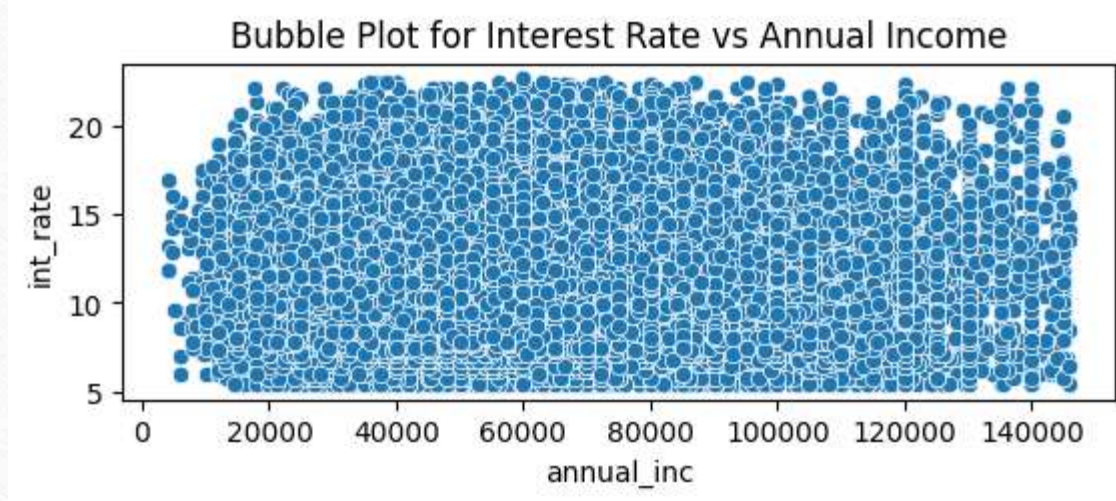
Effecting the Risk (Proportional)

- Loan Amount
- Grade
- Sub Grade
- Annual Income
- Purpose

Low or No Risk (Not Proportional)

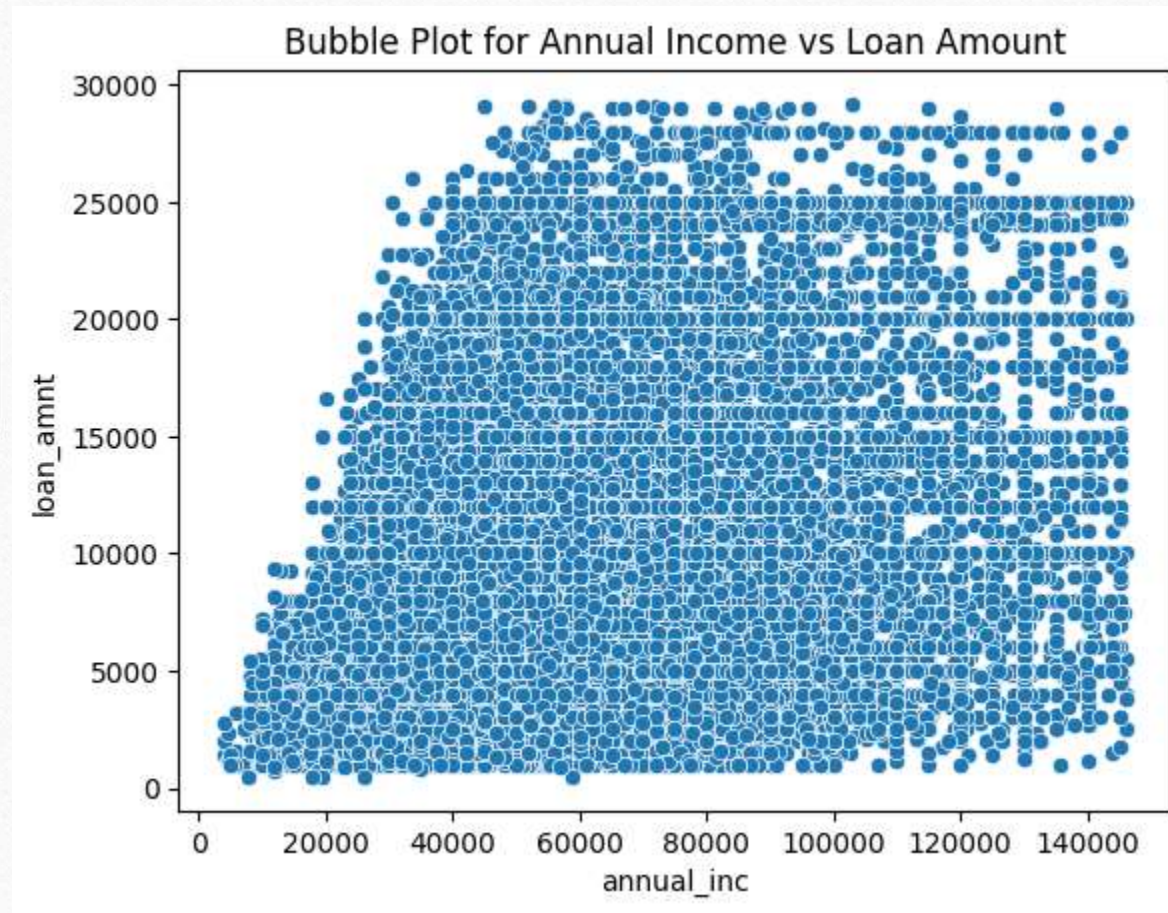
- Term
- Dti
- Employment Length
- Last Payment Date
- Home Ownership
- Verification Status

Interest Rate vs Annual Income



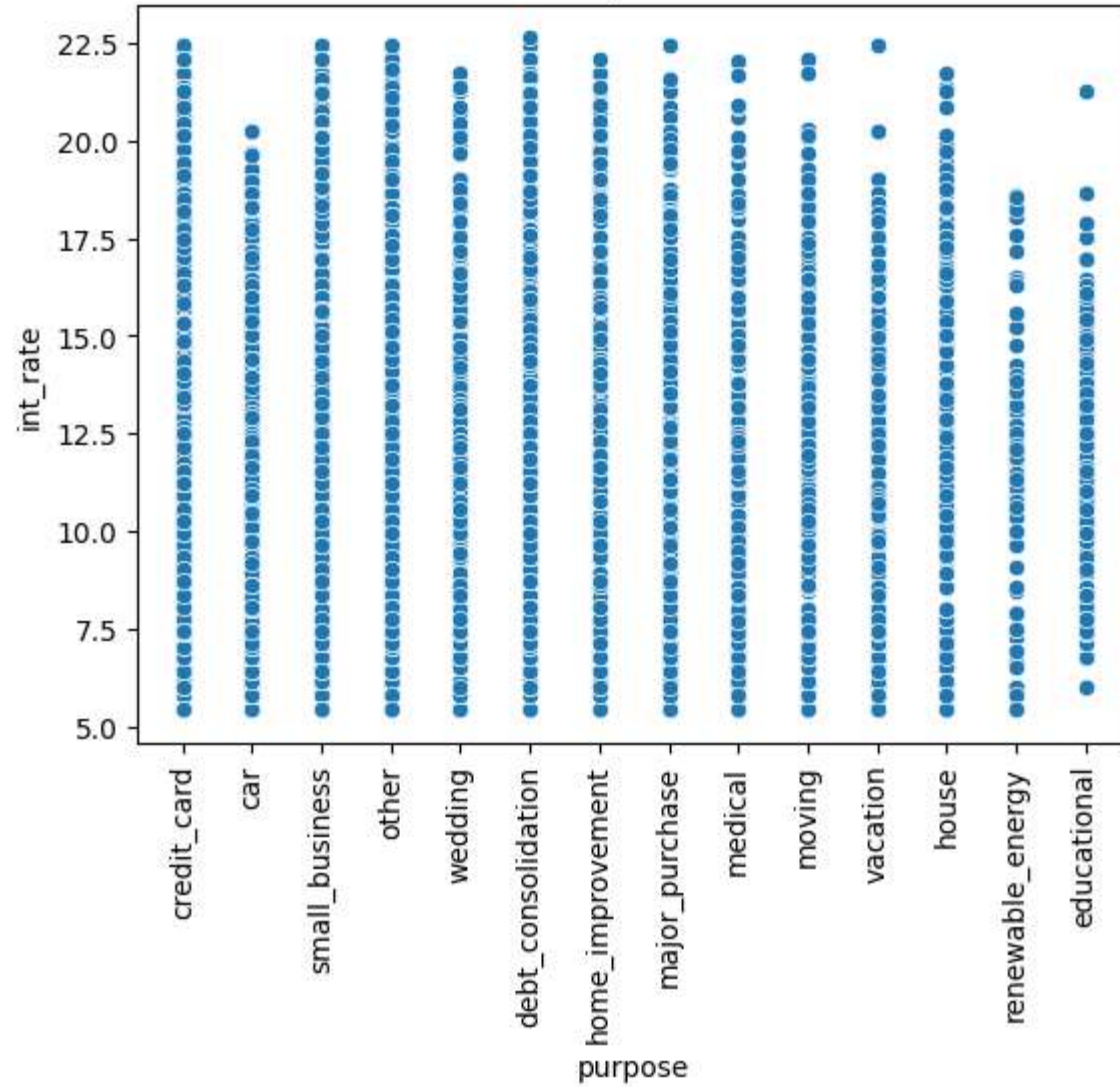
We can see many Applicants are comfortable between Annual Income of 40,000 to 1,00,000+ for higher interest rates of more than 20%

Annual Income vs Loan Amount



Applicants with Annual Income more than 60,000 are comfortable at Higher loan of 25,000+

Bubble Plot for Purpose vs Interest Rate



Purpose vs Interest Rate

Applicants pay higher interest rates for

- Credit Card
- Small Business
- Debt Consolidation
- Other
- Home Improvement

Contents in Jupyter Plots

- All the initial slides given with percentages for proportionate factors have Plots for
 - Univariate and Bivariate(Distribution Plots) Analysis using Histogram plots
 - Interesting relationship between Loan Amount, Interest rate, Annual Income and Purpose are designed using Bubble Plots.

Summary

- Applicant's Interest rate to be less than 15% for safe risk for low grades.
- Applicants with annual income from 40,000 to 1,00,000+ for higher interest rates of more than 20%
- Applicants are comfortable with higher Interest Rate for
 - Credit Card
 - Small Business
 - Debt Consolidation
 - Other
 - Home Improvement