# Lending Club Case Study

#### **Group Members:**

- Moumita
- Jashwanth

# Introduction

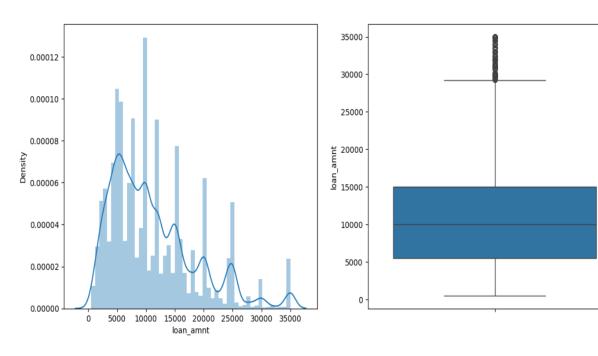
- The consumer finance industry faces risks when deciding to approve or reject loan applications based on applicant profiles.
- Two major risks are associated with loan decisions:
  - 1. Loss of business if a reliable borrower is denied a loan.
  - 2. **Credit loss** if a borrower defaults after approval.
- The dataset contains past loan applicant data, focusing on those who defaulted, which is critical for identifying high-risk borrowers.
- The analysis aims to identify factors that indicate whether a borrower is likely to default, helping reduce credit losses.
- By identifying risky applicants, the company can:
  - 1. Reject or minimise high-risk loans.
  - 2. Offer higher interest rates for riskier borrowers

# Problem Solving Methodology:

- •Understanding the features and data:
- •Explored the dataset to comprehend the features and their significance.
- •Cleaning the data:
- •Handled missing values, and outliers, and ensured data consistency for accurate analysis.
- •Univariate analysis:
- Analyzed the frequency distribution of each feature to understand individual characteristics.
- •Segmented univariate analysis:
- •Performed targeted univariate analysis to explore feature distributions within specific categories.
- •Bivariate analysis:
- •Examined the relationships and correlations between two features to identify patterns.
- •Drawing inferences:
- Derived insights based on the analysis to identify key factors influencing loan default.

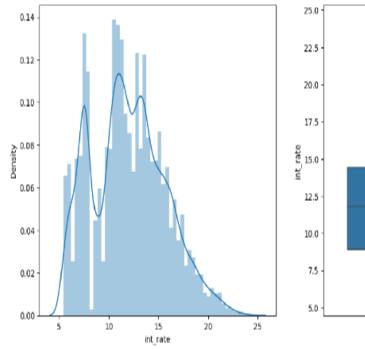
- As we see, the distribution of loan amount follows the Gaussian curve which is right-skewed.
- Majority of the customers are having a loan amount of around 10000 and the rest of the people are also having more than 10000

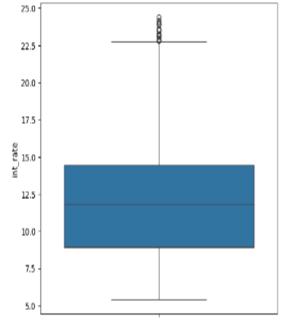




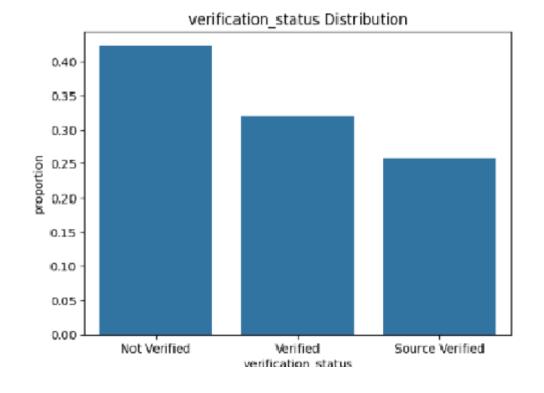
- Based on the analysis of the two plots, we can conclude that the majority of interest rates fall within the range of 9% to 14.5%, with a median rate of approximately 12.5%.
- A small number of borrowers, however, took loans with significantly higher interest rates, reaching up to 22.5%.

#### Interest Rate Distribution

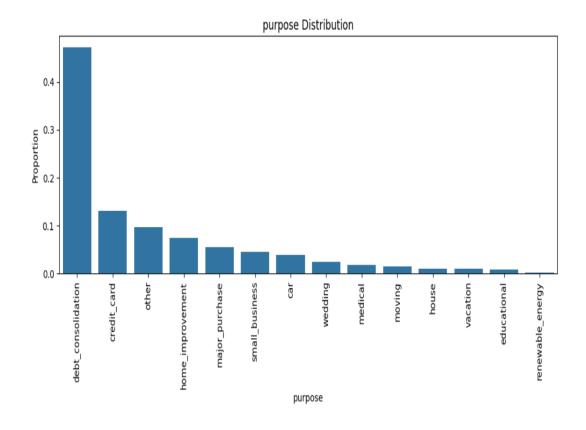




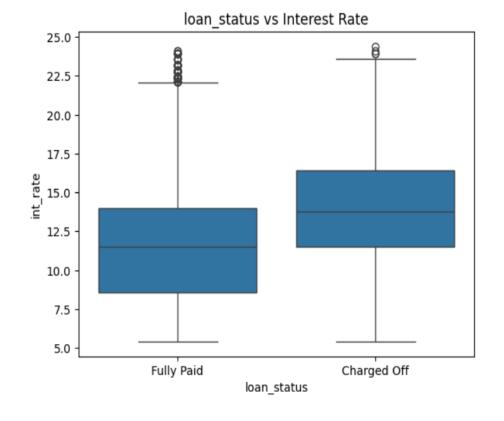
- As we can observe, the majority of the customer got their loan sanctioned from Lending Club with proper background verification.
- More than 40% of the customers were grated the loans without verification, which can cause the more defaulters and credit loss to the company



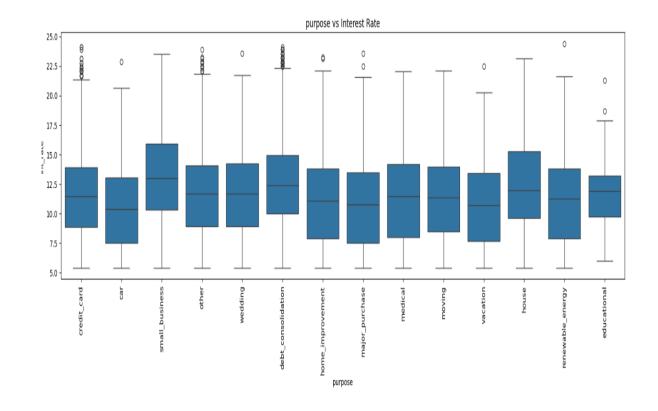
 Customers who are having the loan purpose as debt consolidation have the higher take followed by the credit card holders and home improvements



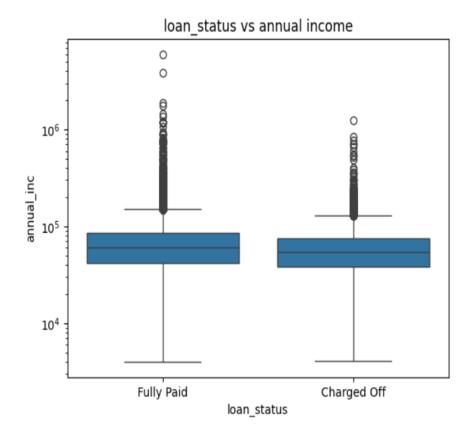
- We can infer that the customers who have high interest rates have been prone to default on the loan with most of the interest rates being around 14%
- While the customers who are having lesser interest rates are able to clear the loans.



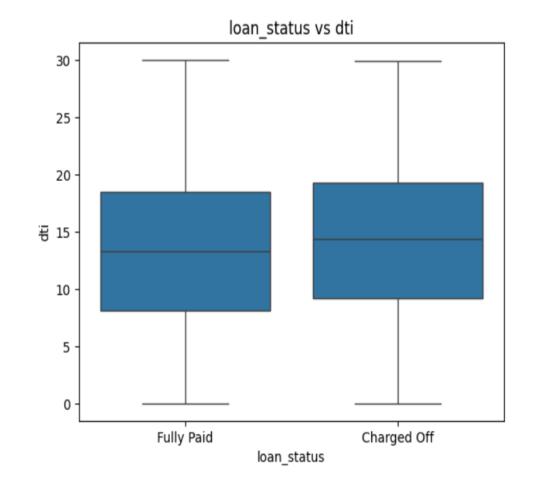
- Customers for the loan purposes like small business, debt consolidation and house loans have higher interest rates than others.
- Also, a good volume of customers have taken the higher interest loans who are having debt consolidation and credit card payments etc



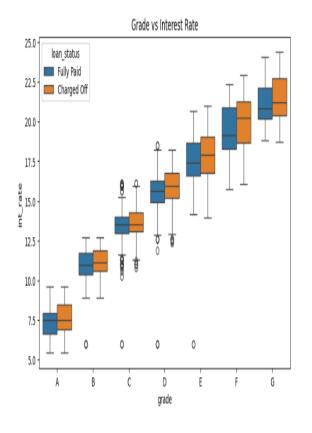
- Customers having a higher annual income can successfully repay the loan.
- While customers who have less annual income are prone to default on the loan.

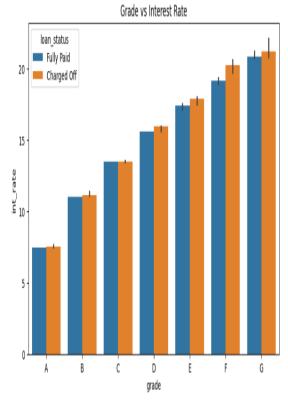


- Although it's not too significant and evident from dti feature, but there is a slight indication that customers having more dti have more chances to default on the loan.
- Customers having relative lesser dti are able to close the loan payment successfully.

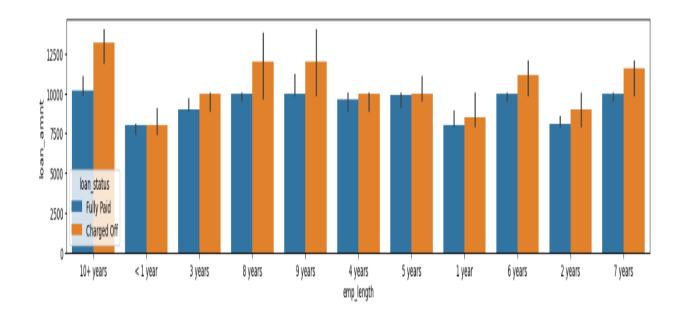


- As the grade of the customer increases, the interest also linearly increases with in turn leads to a higher rise in the defaulter's chances.
- Customer having A, B and C grades can clear the loans with fewer loan default rate

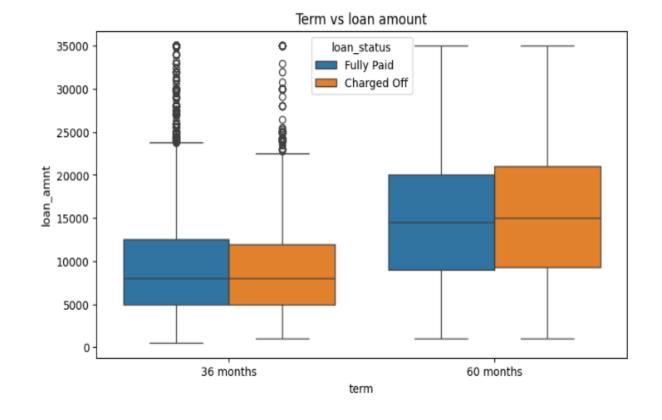




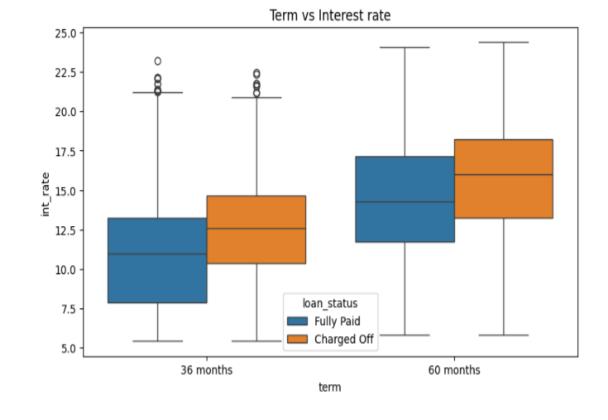
- Customers with high employment length have defaulted the loan more than less employment lengths.
- Shorter loan tenures make is yielding the good loan recovery and good credit maintenance for the company.



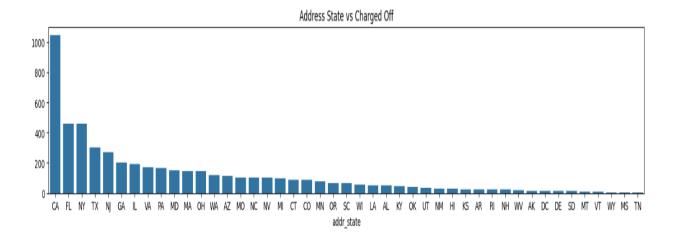
- Customers with short term were able to clear the loan amount effectively.
- The customers who are having higher term loan payments are subjected to default on the loan.



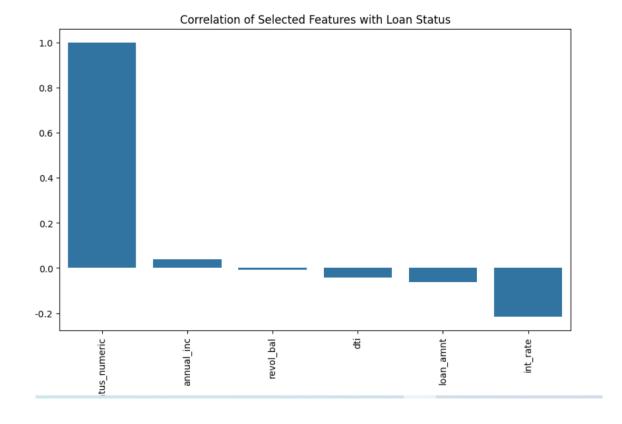
- Customers who opted to clear the loan for higher terms are subjected by the company to pay the higher interest rates.
- This has counter-fired as many of the customers have defaulted on the loan because of the higher interest rates and higher term.



• It's evident that customers from the CA, FL, and NY have a very high charged Off compared to customers from the other states.



 We can see from the Pearson correlation, that high interest rates, high loan amounts, and high DTI have a negative correlation with the loan status



#### Correlation between Columns

id -	1	0.12	0.14	0.2	0.046	0.082	0.0097	0.93	0.08	-0.0039	0.063	-0.038	0.019	-0.025	0.0024	0.088	0.032	0.11	0.17	0.14	-0.01	0.33	0.83	-0.018	0.064
loan_amnt -	0.12	1	0.98	0.95	0.3	0.93	0.27	0.13	0.06	-0.035	-0.18	0.011	0.17	-0.047	0.32	0.061	0.26	0.88	0.86	0.48	-0.032	-0.12	0.12	0.0087	-0.18
funded_amnt -	0.14	0.98		0.97	0.3	0.96	0.27	0.14	0.06	-0.035	-0.17	0.011	0.17	-0.048		0.065	0.25	0.9	0.88	0.48	-0.033	-0.035	0.13	0.0011	-0.17
funded_amnt_inv -	0.2	0.95	0.97		0.29	0.92	0.26	0.23	0.063	-0.039	-0.16	0.00016	0.16	-0.052	0.29	0.068	0.24	0.89	0.91	0.47	-0.038	0.14	0.22	0.0035	-0.16
int_rate -	0.046			0.29	1	0.28	0.048	0.033	0.11	0.16	0.11	0.14	0.0086	0.099	0.096		-0.046	0.29	0.28	0.17	0.083	0.0032	0.026	0.032	0.11
installment -	0.082	0.93	0.96	0.92	0.28	1	0.27	0.069	0.05	-0.023	-0.16	0.009	0.17	-0.043		0.091	0.23	0.86	0.83	0.41	-0.029	-0.037	0.061-0	0.00088	3-0.16
annual_inc -	0.0097	0.27	0.27	0.26	0.048	0.27	1	0.017	-0.12	0.021	-0.18	0.035	0.16	-0.014	0.28	0.016	0.24	0.26	0.25	0.14	-0.012	-0.03	0.015	0.0033	-0.18
issue_d -	0.93	0.13	0.14	0.23	0.033	0.069	0.017	1	0.081	0.00019	0.059	-0.051	0.015	-0.024	0.0036	0.075	0.042	0.12	0.2	0.14	-0.011	0.45	0.94	-0.024	0.06
dti -	0.08	0.06	0.06	0.063	0.11	0.05	-0.12	0.081	1	-0.035	-0.0370	0.00067	0.29	0.0054	0.23	0.28	0.23	0.058	0.06	0.0077	0.0064	0.014	0.081	0.0028	-0.037
delinq_2yrs -	-0.0039	-0.035	-0.035	-0.039	0.16	-0.023	0.021	0.00019	-0.035	1	-0.064	0.0084	0.011	0.011	-0.056	-0.044	0.067	-0.025	-0.03	-0.013	0.0058-	-0.0085	0.0043	0.011	-0.065
earliest_cr_line -	0.063	-0.18	-0.17	-0.16	0.11	-0.16	-0.18	0.059	-0.037	-0.064	1	-0.015	-0.21	-0.04	-0.24	0.036	-0.35	-0.16	-0.15	-0.066	-0.049	0.067	0.054	0.0029	1
inq_last_6mths -	-0.038	0.011	0.011	0.00016	5 0.14	0.009	0.035	-0.0510	.00067	0.0084	-0.015	1	0.094	0.023	-0.025	-0.07	0.12 -	0.0088	-0.016	0.028	0.014	-0.041	-0.059	0.0019	-0.014
open_acc -	0.019	0.17	0.17	0.16	0.0086	0.17	0.16	0.015	0.29	0.011	-0.21	0.094	1	0.0034	0.29	-0.088	0.69	0.16	0.15	0.08	0.0091	-0.04	0.015	0.0019	-0.21
pub_rec -	-0.025	-0.047	-0.048	-0.052	0.099	-0.043	-0.014	-0.024	0.0054	0.011	-0.04	0.023	0.0034	1	-0.059	0.057	-0.02	-0.05	-0.053	-0.034	0.84	-0.017	-0.016	0.0014	-0.04
revol_bal -	0.0024			0.29	0.096		0.28	-0.0036	0.23	-0.056	-0.24	-0.025	0.29	-0.059	1	0.3	0.31	0.29	0.28	0.13	-0.047	-0.071-	0.0085	0.0048	-0.24
revol_util -	0.088	0.061	0.065	0.068	0.46	0.091	0.016	0.075	0.28	-0.044	0.036	-0.07	-0.088	0.057	0.3	1	-0.069	0.073	0.075	-0.018	0.06	0.026	0.062-0	0.0006	10.036
total_acc -	0.032	0.26	0.25	0.24	-0.046	0.23	0.24	0.042	0.23	0.067	-0.35	0.12	0.69	-0.02		-0.069	1	0.22	0.22	0.17 -	0.0067	-0.035	0.043	0.0043	-0.35
total_pymnt -	0.11	0.88	0.9	0.89	0.29	0.86	0.26	0.12	0.058	-0.025	-0.16 -	0.0088	0.16	-0.05	0.29	0.073	0.22		0.98	0.51	-0.039	-0.0021	0.12 -0	0.00022	2-0.16
total_pymnt_inv -	0.17	0.86	0.88	0.91	0.28	0.83	0.25	0.2	0.06	-0.03	-0.15	-0.016	0.15	-0.053	0.28	0.075	0.22	0.98	1	0.5	-0.043	0.14	0.19	0.0025	-0.15
last_pymnt_amnt -	0.14	0.48	0.48	0.47	0.17	0.41	0.14	0.14	0.0077	-0.013	-0.066	0.028	0.08	-0.034	0.13	-0.018	0.17	0.51	0.5	1	-0.023	-0.0055	0.13	-0.012	-0.066
pub_rec_bankruptcies -	-0.01	-0.032	-0.033	-0.038	0.083	-0.029	-0.012	-0.011	0.0064	0.0058	-0.049	0.014	0.0091	0.84	-0.047	0.06	-0.0067	-0.039	-0.043	-0.023	1	-0.021-	0.0032	0.0038	-0.048
approved_loan_amnt_ratio -	0.33	-0.12	-0.035	0.14	0.0032	-0.037	-0.03	0.45	0.014	-0.0085	0.067	-0.041	-0.04	-0.017	-0.071	0.026	-0.035	0.0021	0.14	0.0055	-0.021	1	0.46	0.033	0.067
issue_d_year -	0.83	0.12	0.13	0.22	0.026	0.061	0.015	0.94	0.081	0.0043	0.054	-0.059	0.015	-0.016-	-0.0085	0.062	0.043	0.12	0.19	0.13 -	0.0032	0.46	1	0.01	0.055
issue_d_weekday -	-0.018	0.0087	0.0011	0.0035	0.032-	0.00088	0.0033	-0.024	0.0028	0.011	0.0029	0.0019	0.0019	0.0014-	0.0048	0.0006	LO.0043€	0.00022	<b>2</b> 0.0025	-0.012	0.0038	0.033	0.01	1	0.003
earliest_cr_line_year -	0.064	-0.18	-0.17	-0.16	0.11	-0.16	-0.18	0.06	-0.037	-0.065	1	-0.014	-0.21	-0.04	-0.24	0.036	-0.35	-0.16	-0.15	-0.066	-0.048	0.067	0.055	0.003	1
		loan_amnt -	funded_amnt -	funded_amnt_inv -	int_rate -	installment -	- amual inc	- p anssi	- Eff	delinq_2yrs -	earliest cr line -	ing_last_6mths -	- oben acc	- par qnd	- levol_bal -	- revo <u>l</u> util	- total acc	total_pymnt -	total_pymnt_inv -	last_pymnt_amnt -	pub_rec_bankruptcies -	approved_loan_amnt_ratio -	issue_d_year -	issue_d_weekday -	earliest_cr_line_year -

- 0.6

- 0.2

- 0.0

- -0.2

#### 1. Strong Positive Correlations:

Loan Amount, Funded Amount, Funded Amount Inv: These variables are highly correlated, as expected since they represent different aspects of the loan amount.

Total Payment, Total Payment Inv: Similarly, these variables are strongly correlated, indicating a direct relationship between the total amount paid and the total amount invested.

#### 2. Moderate Positive Correlations:

Loan Amount, Interest Rate: There's a moderate positive correlation, suggesting that larger loans might have higher interest rates.

*Installment, Loan Amount*: A larger loan amount typically corresponds to higher installments.

#### 3. Weak or No Correlations:

Public Records Bankruptcies, Interest Rate: There's a weak correlation or no correlation between these variables, indicating that bankruptcy history doesn't significantly influence interest rates.

#### 4. Specific Observations:

Interest Rate: While it has moderate positive correlations with loan amount and installment, it doesn't seem to be strongly correlated with other variables.

Delinquency: Delinquency (delinq\_2yrs) doesn't have strong correlations with most other variables, suggesting it might be a relatively independent factor.

*Public Records Bankruptcies:* As mentioned earlier, it doesn't seem to be strongly correlated with interest rates.

# **Conclusions:**

- The Lending Club should exercise heightened caution when offering loans to small businesses, as they have a higher likelihood of default.
- It is advisable not to issue loans with a repayment tenure exceeding 8 years, as shorter terms are more effective in ensuring loan recovery.
- Special attention should be given to applicants from California (CA), Florida (FL), and New York (NY), as these states exhibit higher default rates.
- The company must maintain tight control over interest rates, as higher rates tend to increase the risk of default, leading to significant credit losses.
- A robust customer verification process is critical. Unverified applicants have shown a significantly higher tendency to default on their loans.
- Borrowers in lower credit grades, particularly F and G, should be closely monitored, as they are more likely to default.