

# Lending Case Study

Case study to study a dataset of loan and identify patterns for defaulters of loan so that the organization can analyse loan applications to avoid defaulters



# Background

---

When a person applies for a loan, there are **granted the loan**, the applicant can have below three status:

1. **Fully paid:** Applicant has fully paid the loan (the principal and the interest rate)
2. **Current:** Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
3. **Charged-off:** Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has **defaulted** on the loan

# Business Objective

---

To identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.

# DataSet Given for Analysis

Loan Data Set – Sample shown below

Columns – 111 , Rows - 39717

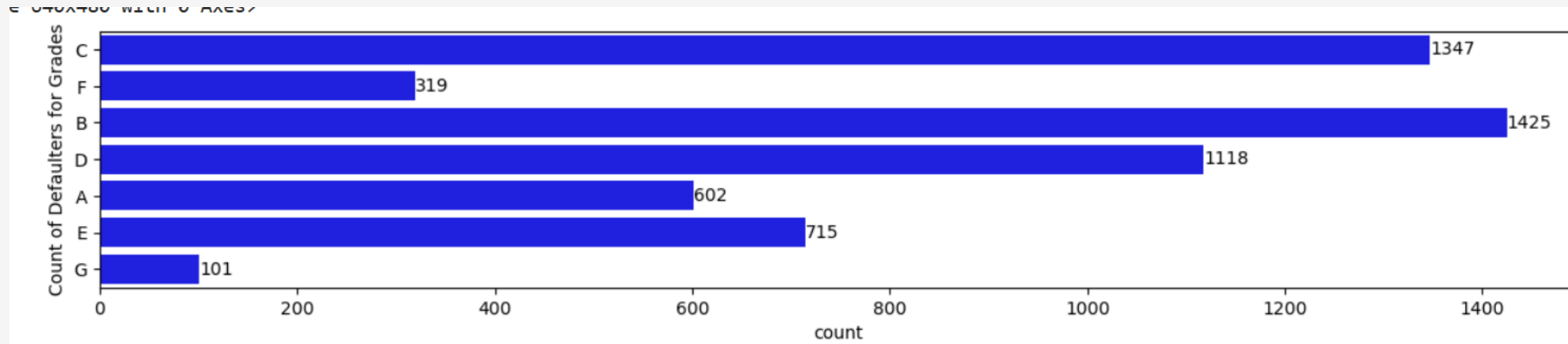
	1	2	3	4	5	6	7	8	9	10	11	
1	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	term	int_rate	installment	grade	sub_grade	emp_title	emp_
2	1077501	1296599	5000	5000	4975	36 months	10.65%	162.87	B	B2		10+ y
3	1077430	1314167	2500	2500	2500	60 months	15.27%	59.83	C	C4	Ryder	< 1 ye
4	1077175	1313524	2400	2400	2400	36 months	15.96%	84.33	C	C5		10+ y
5	1076863	1277178	10000	10000	10000	36 months	13.49%	339.31	C	C1	AIR RESOURCES BOARD	10+ y
6	1075358	1311748	3000	3000	3000	60 months	12.69%	67.79	B	B5	University Medical Group	1 yea
7	1075269	1311441	5000	5000	5000	36 months	7.90%	156.46	A	A4	Veolia Transportaton	3 yea
8	1069639	1304742	7000	7000	7000	60 months	15.96%	170.08	C	C5	Southern Star Photography	8 yea
9	1072053	1288686	3000	3000	3000	36 months	18.64%	109.43	E	E1	MKC Accounting	9 yea
10	1071795	1306957	5600	5600	5600	60 months	21.28%	152.39	F	F2		4 yea
11	1071570	1306721	5375	5375	5350	60 months	12.69%	121.45	B	B5	Starbucks	< 1 ye
12	1070078	1305201	6500	6500	6500	60 months	14.65%	153.45	C	C3	Southwest Rural metro	5 yea
13	1069908	1305008	12000	12000	12000	36 months	12.69%	402.54	B	B5	UCLA	10+ y
14	1064687	1298717	9000	9000	9000	36 months	13.49%	305.38	C	C1	Va. Dept of Conservation/Recreation	< 1 ye
15	1069866	1304956	3000	3000	3000	36 months	9.91%	96.68	B	B1	Target	3 yea
16	1069057	1303503	10000	10000	10000	36 months	10.65%	325.74	B	B2	SFMTA	3 yea
17	1069759	1304871	1000	1000	1000	36 months	16.29%	35.31	D	D1	Internal revenue Service	< 1 ye
18	1065775	1299699	10000	10000	10000	36 months	15.27%	347.98	C	C4	Chin's Restaurant	4 yea
19	1069971	1304884	3600	3600	3600	36 months	6.03%	109.57	A	A1	Duracell	10+ y
20	1062474	1294539	6000	6000	6000	36 months	11.71%	198.46	B	B3	Connection Inspection	1 yea
21	1069742	1304855	9200	9200	9200	36 months	6.03%	280.01	A	A1	Network Interpreting Service	6 yea
22	1069740	1284848	20250	20250	19142.16	108 60 months	15.27%	484.63	C	C4	Archdiocese of Galveston Houston	3 yea
23	1039153	1269083	21000	21000	21000	36 months	12.42%	701.73	B	B4	Osram Sylvania	10+ y
24	1069710	1304821	10000	10000	10000	36 months	11.71%	330.76	B	B3	Value Air	10+ y
25	1069700	1304810	10000	10000	10000	36 months	11.71%	330.76	B	B3	Wells Fargo Bank	5 yea
26	1069559	1304634	6000	6000	6000	36 months	11.71%	198.46	B	B3	bmg-educational	1 yea

# Summary of Observations from the Dataset

---

## 1 Univariate - Grade

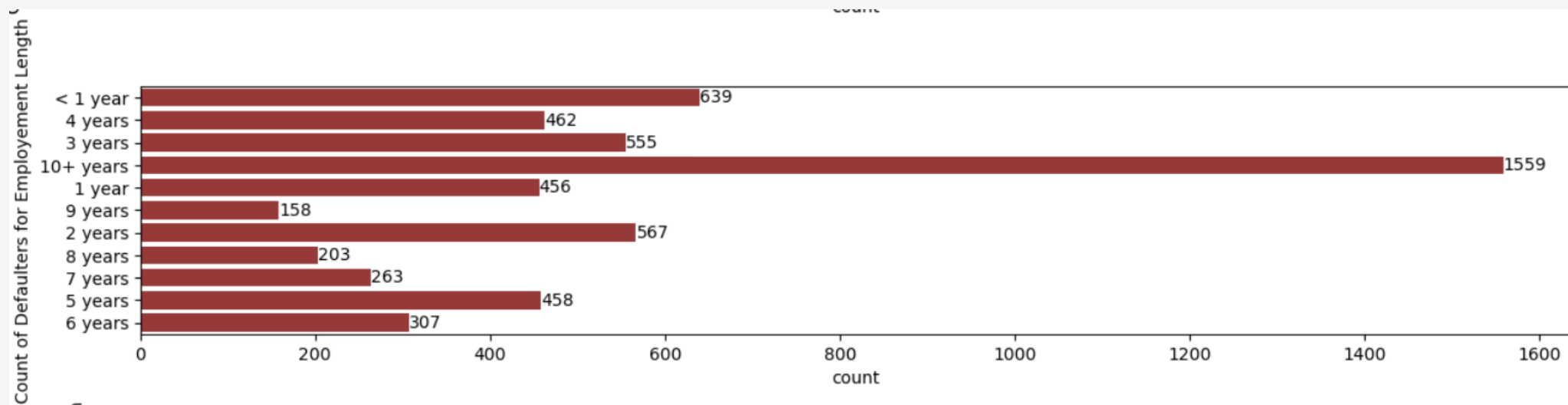
Most defaulters have grade of "B"  
followed by "A" and "C"



# Summary of Observations from the Dataset

## 2 Univariate - Employment Length

Most defaulters have Employment length for more than 10 years



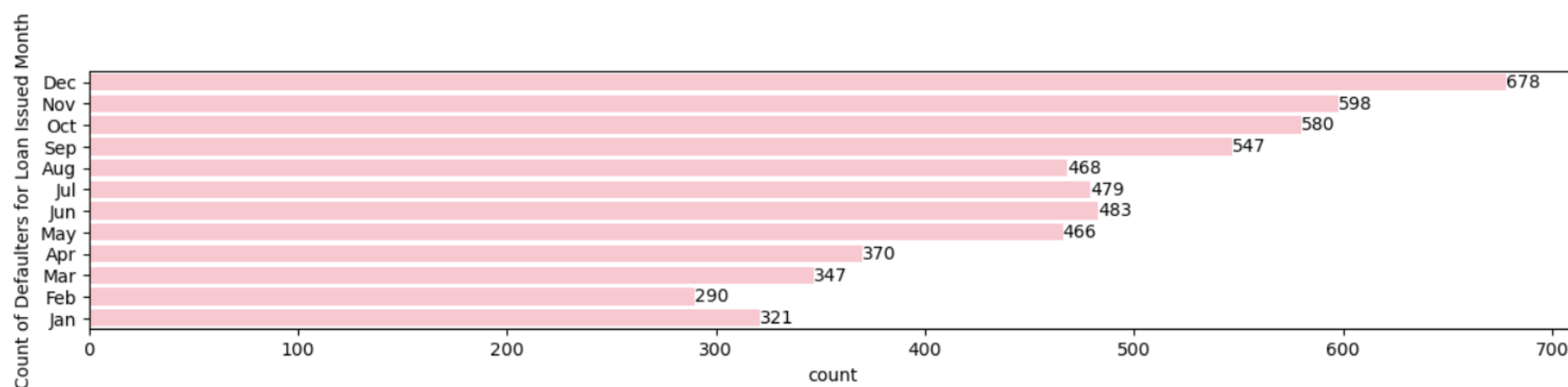
# Summary of Observations from the Dataset

---

## 3 Univariate - Issue Month

(Last quater)

Most defaulters have taken loan in months of December" followed by "November" and "October"

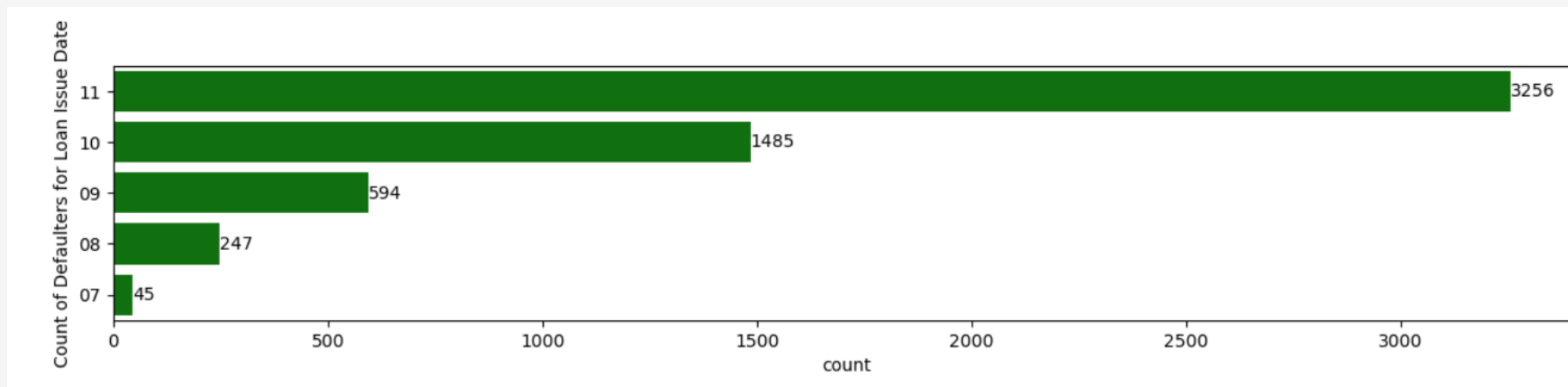


# Summary of Observations from the Dataset

---

## 4 Univariate - Issue Date

Most defaulters have taken loan in  
dates 11th followed by 10th



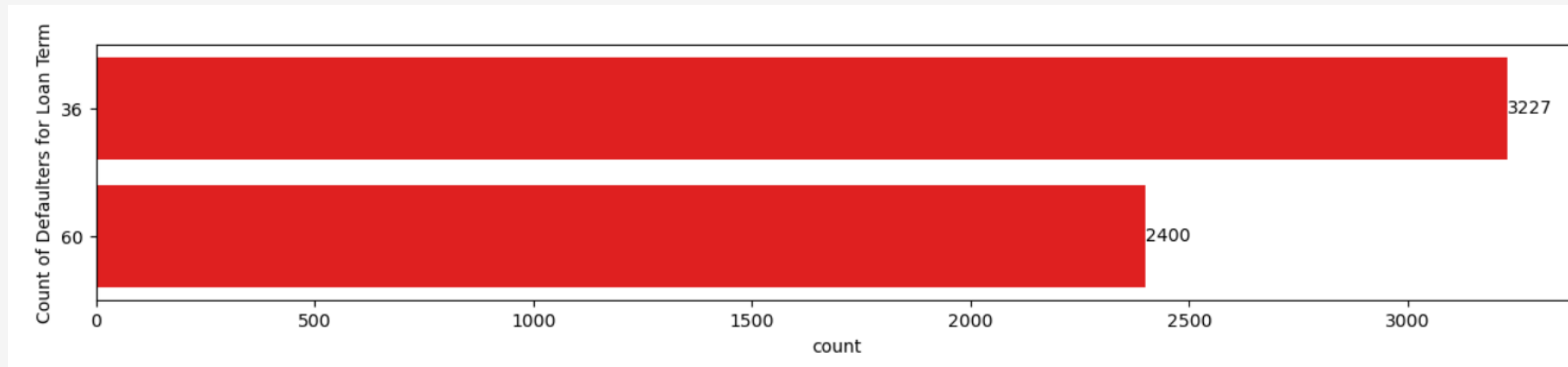


# Summary of Observations from the Dataset

---

## 5 Univariate - Term

Most defaulters have loan term of  
36 months

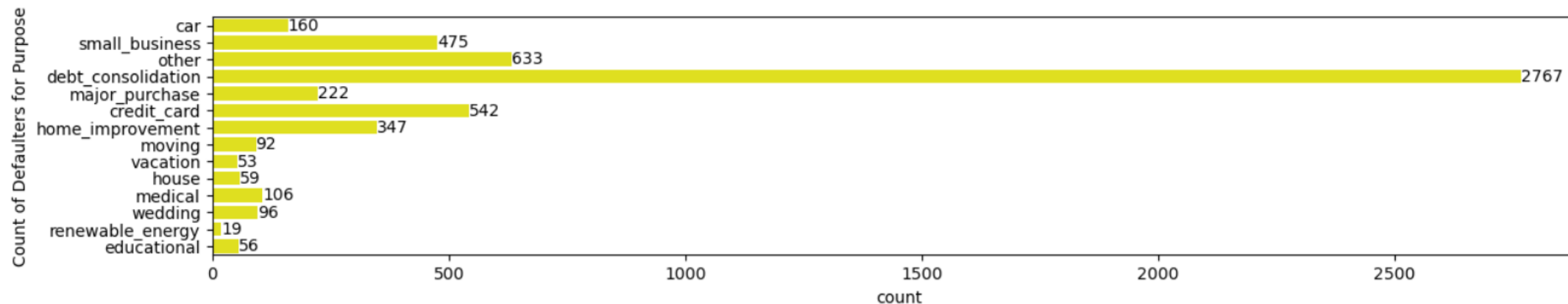


# Summary of Observations from the Dataset

---

## 6 Univariate - Purpose

Most defaulters take loan for the purpose of "Debt Consolidation"

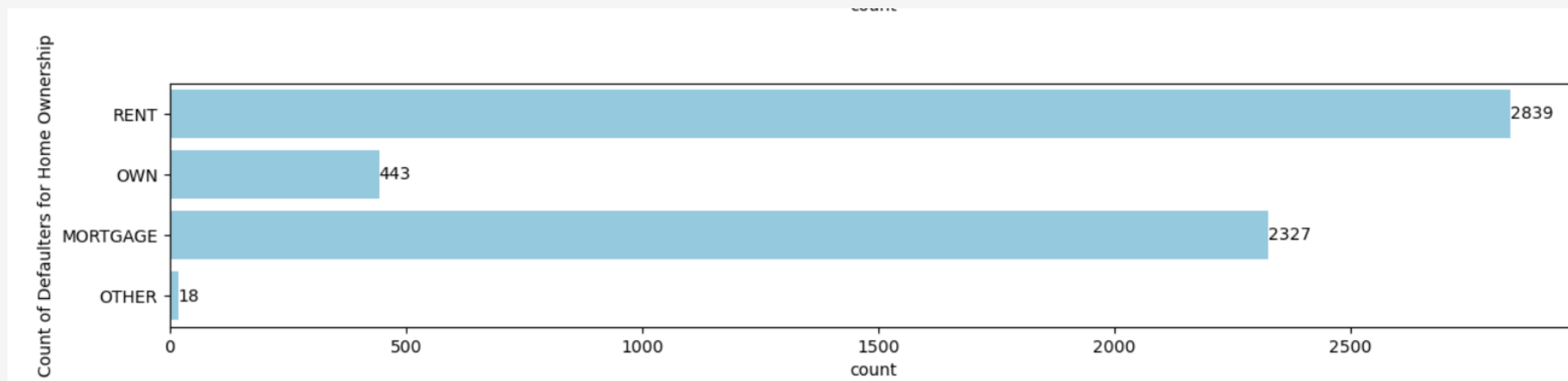


# Summary of Observations from the Dataset

---

## 7 Univariate - Home Ownership

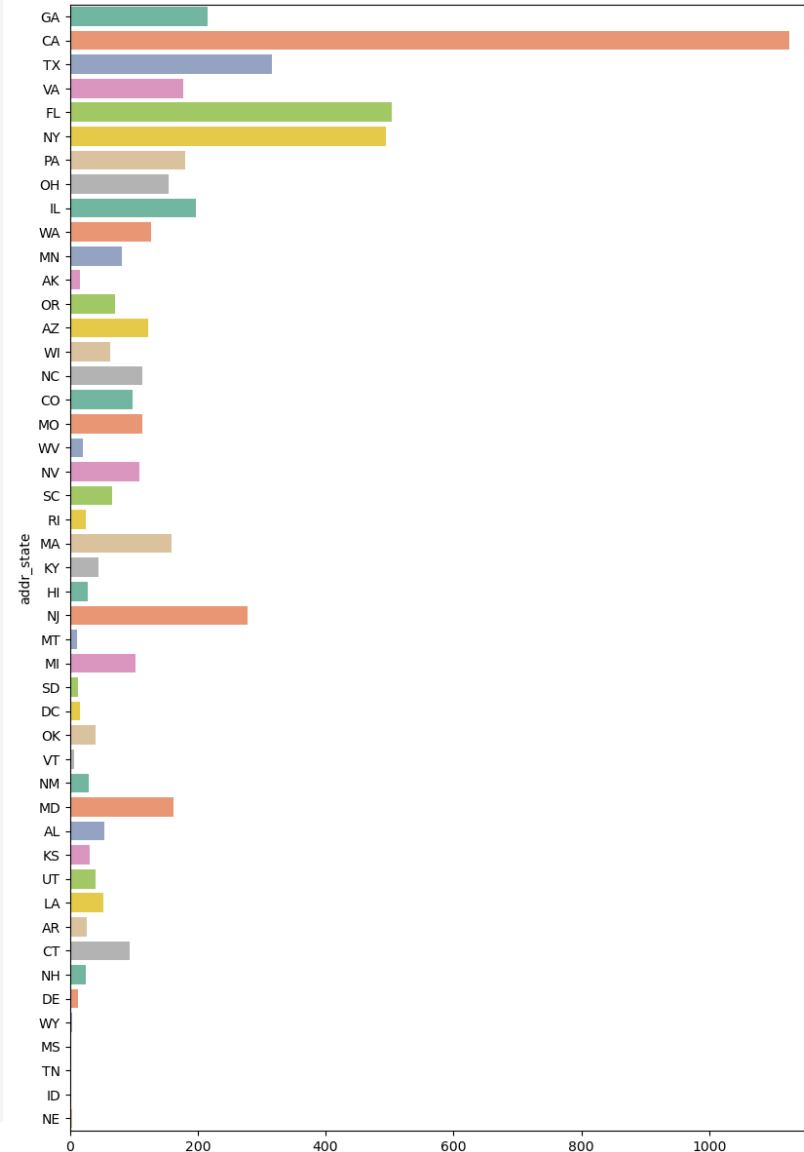
Most defaulters has home ownership status of "Rent" and "Mortgage"



# Summary of Observations from the Dataset

## 8 Univariate - Address State

Most defaulters are from  
state CA

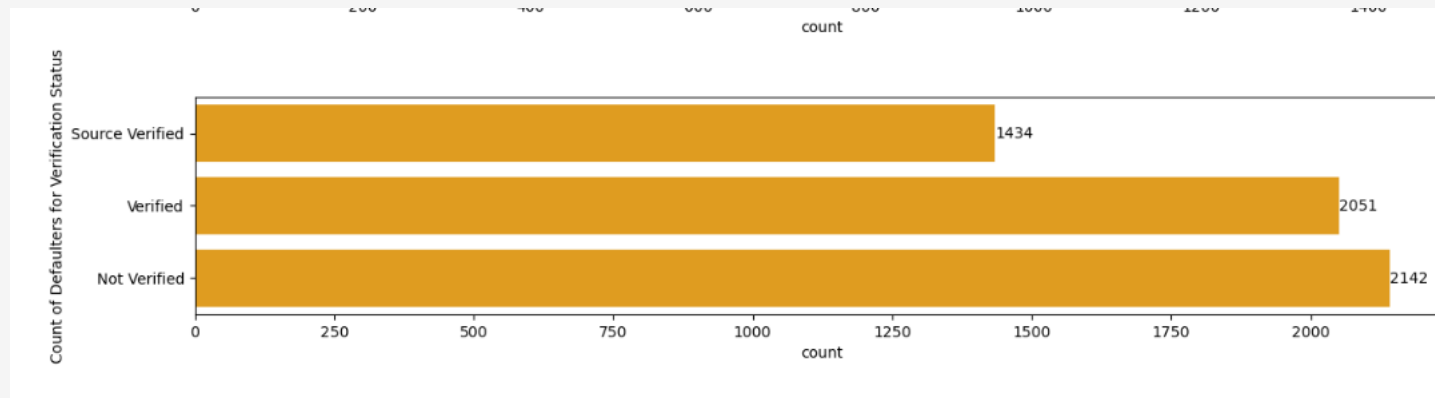


# Summary of Observations from the Dataset

---

## 9 Univariate - Verification Status

Verification status has no significant difference on the Defaulters. The number of defaulters with "Source Verified" are less compared to the other verification status

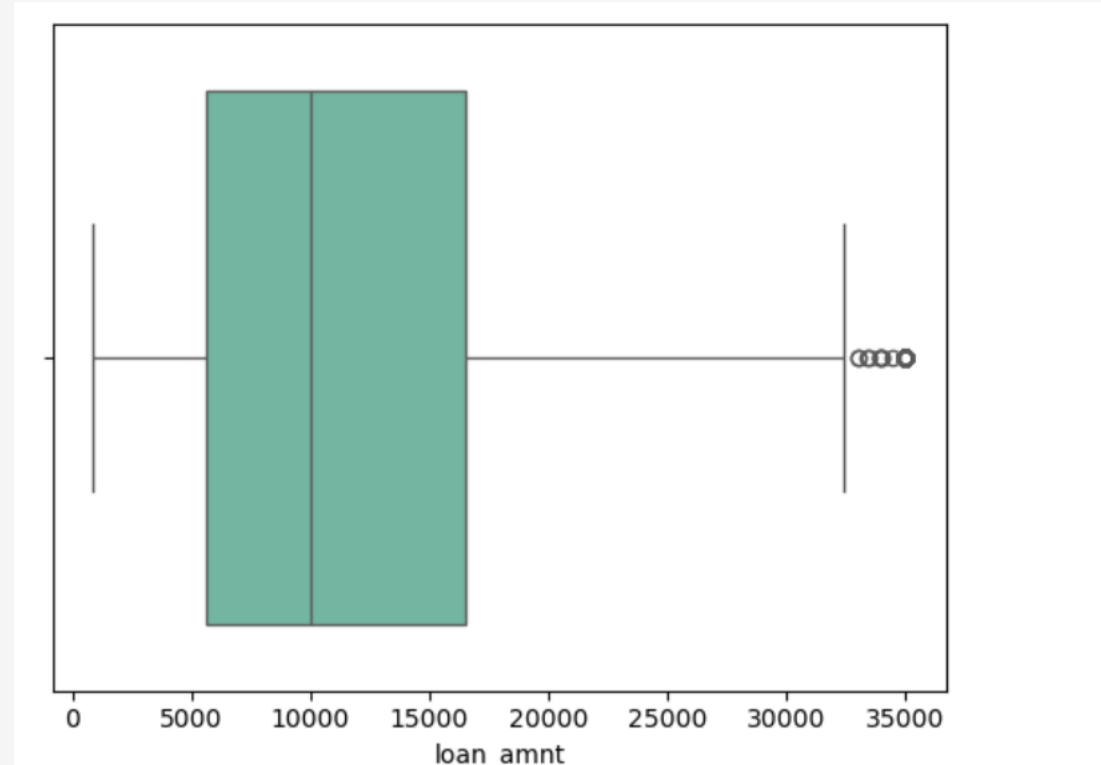


# Summary of Observations from the Dataset

---

## 10 Univariate - Loan Amount

**Most defaulters have loan amount  
between 5000-15000**

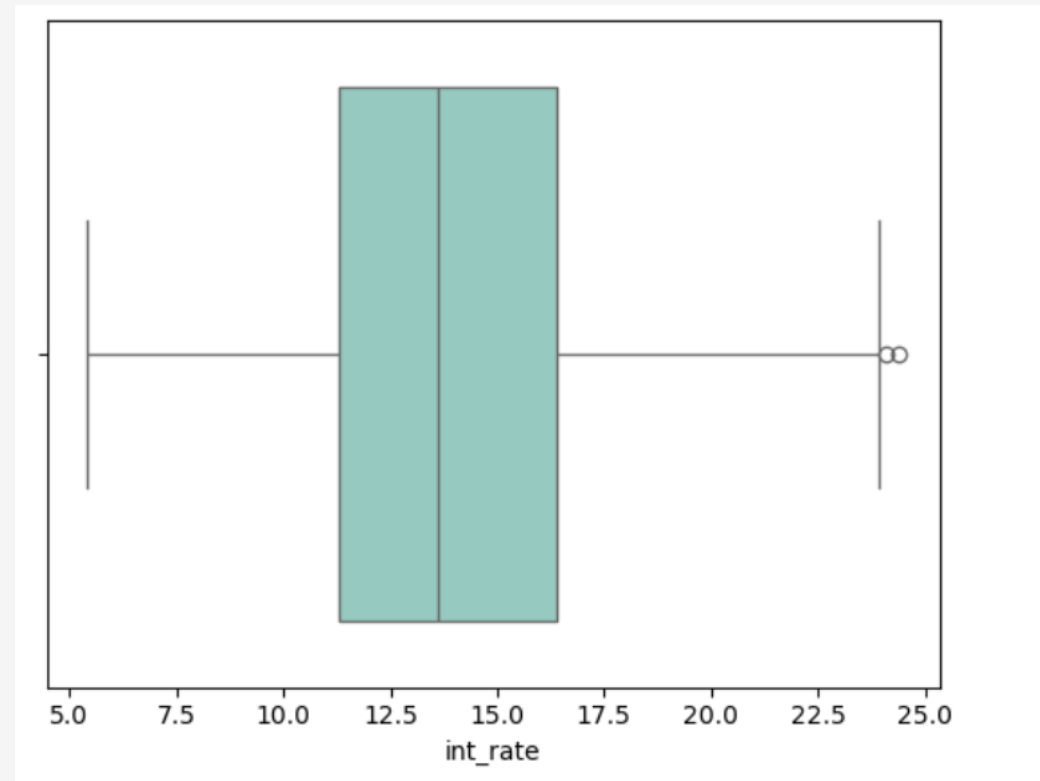


# Summary of Observations from the Dataset

---

## 11 Univariate - Interest Rate

**Most defaulters have interest rate  
between 10%-15%**

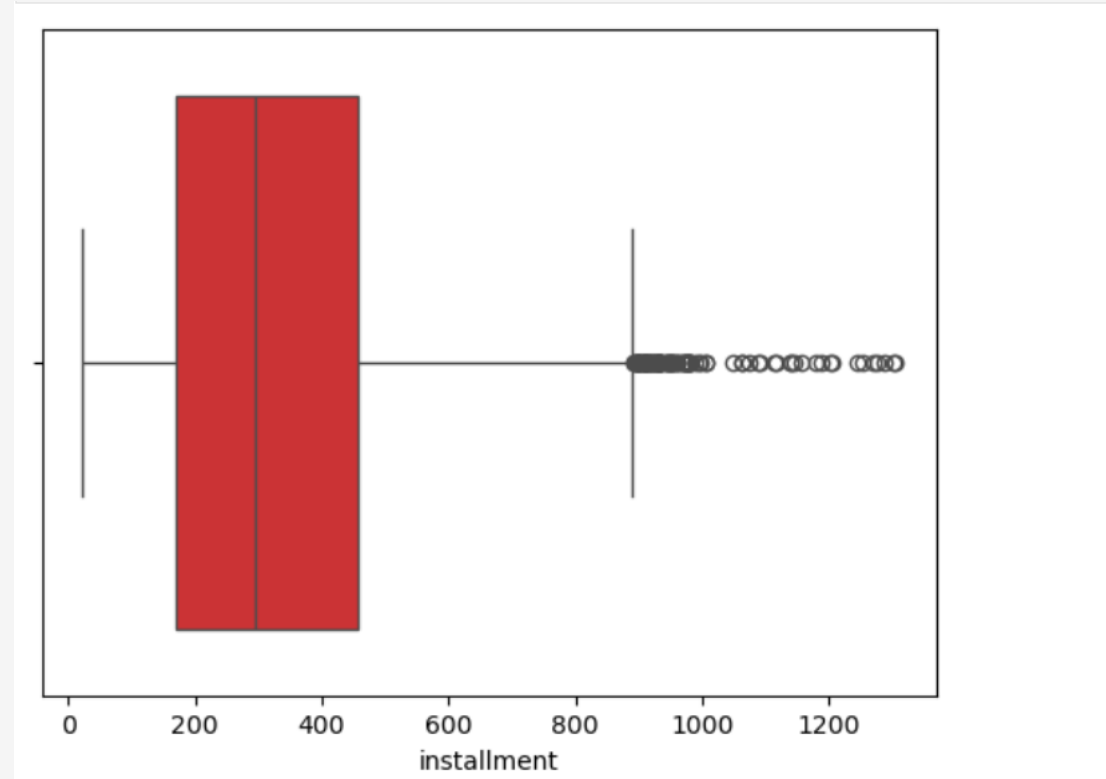


# Summary of Observations from the Dataset

---

## 12 Univariate - Installment

**Most defaulters have installment  
between 200-400**





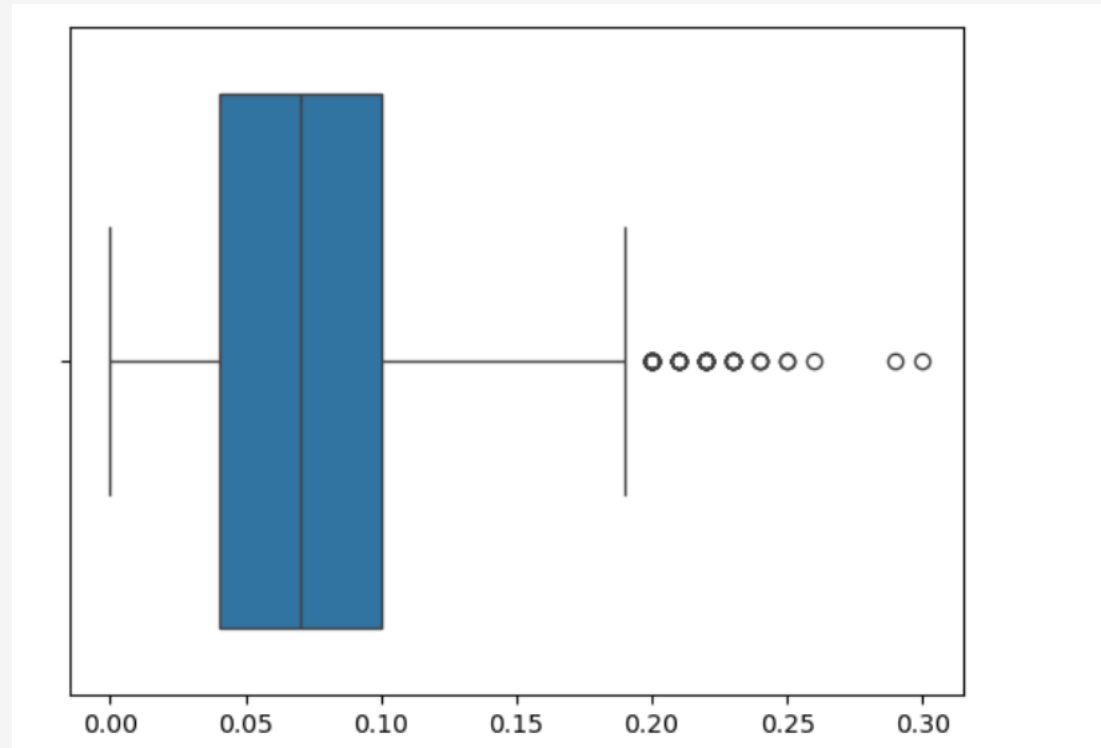
# Summary of Observations from the Dataset

---

13

**Univariate - Percentage of Amount  
towards installements**

**Most defaulters have installments  
between 5%-10% of their monthly  
income**



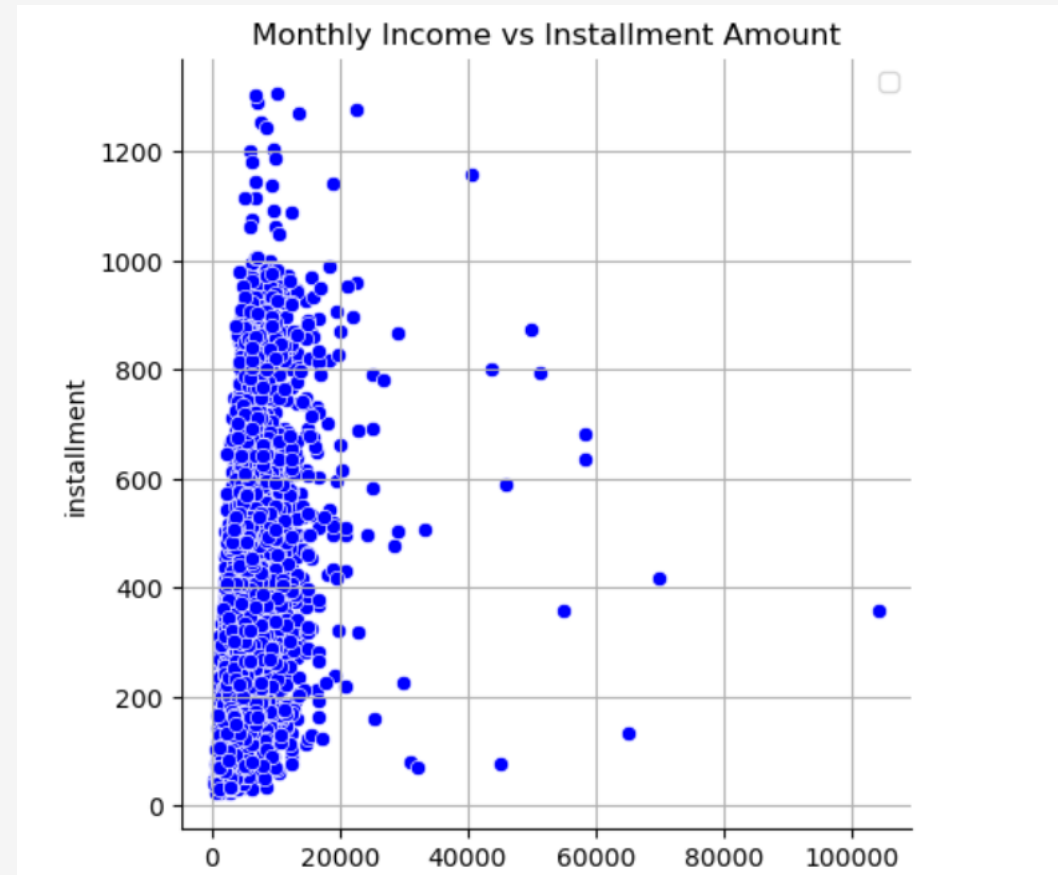
# Summary of Observations from the Dataset

---

14

**Bivariate - Monthly Income v/s Installment Amount**

Most defaulters have monthly income between 3000-5000 and installation amount between 200-320 and show similar pattern

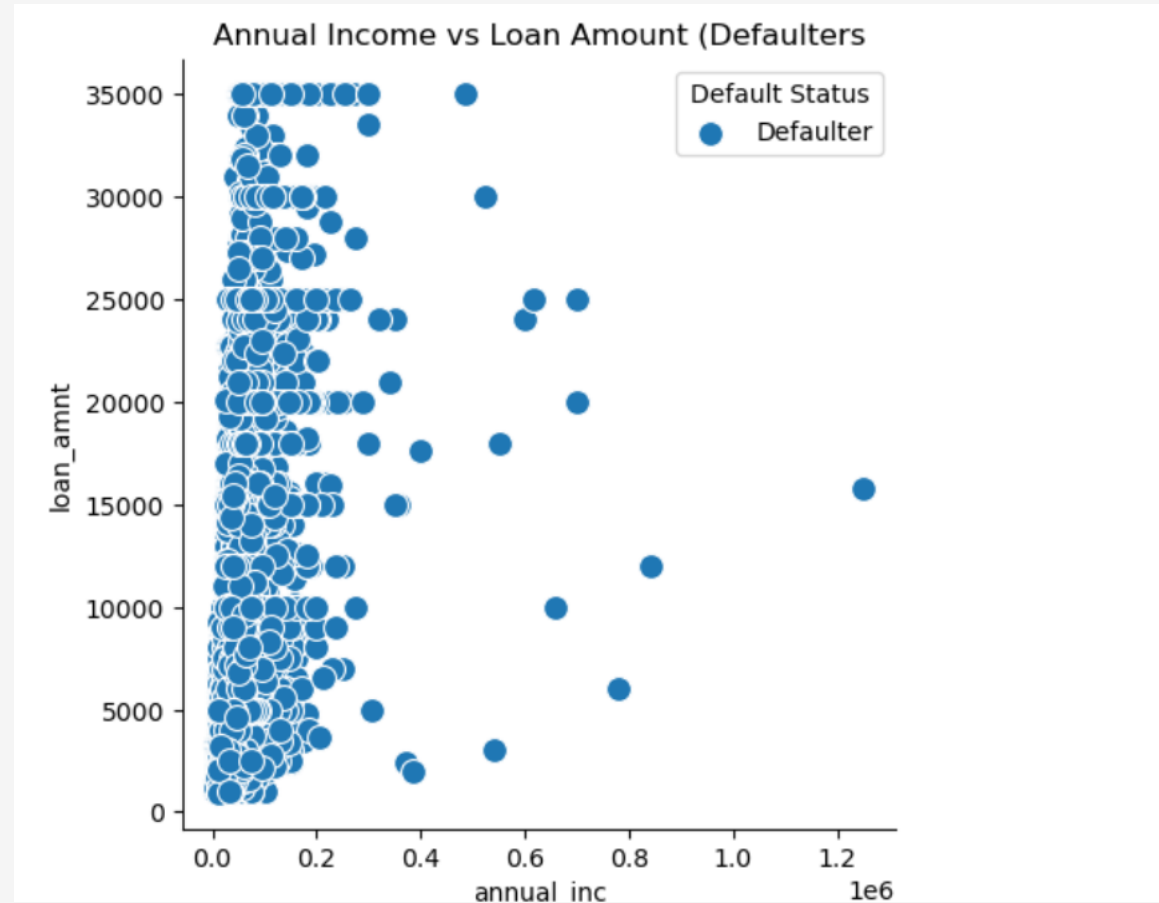


# Summary of Observations from the Dataset

15

## Bivariate - Annual Income v/s Loan Amount

Most defaulters have annual income between 30000-60000 and loan amount between 10000-25000 and show similar pattern



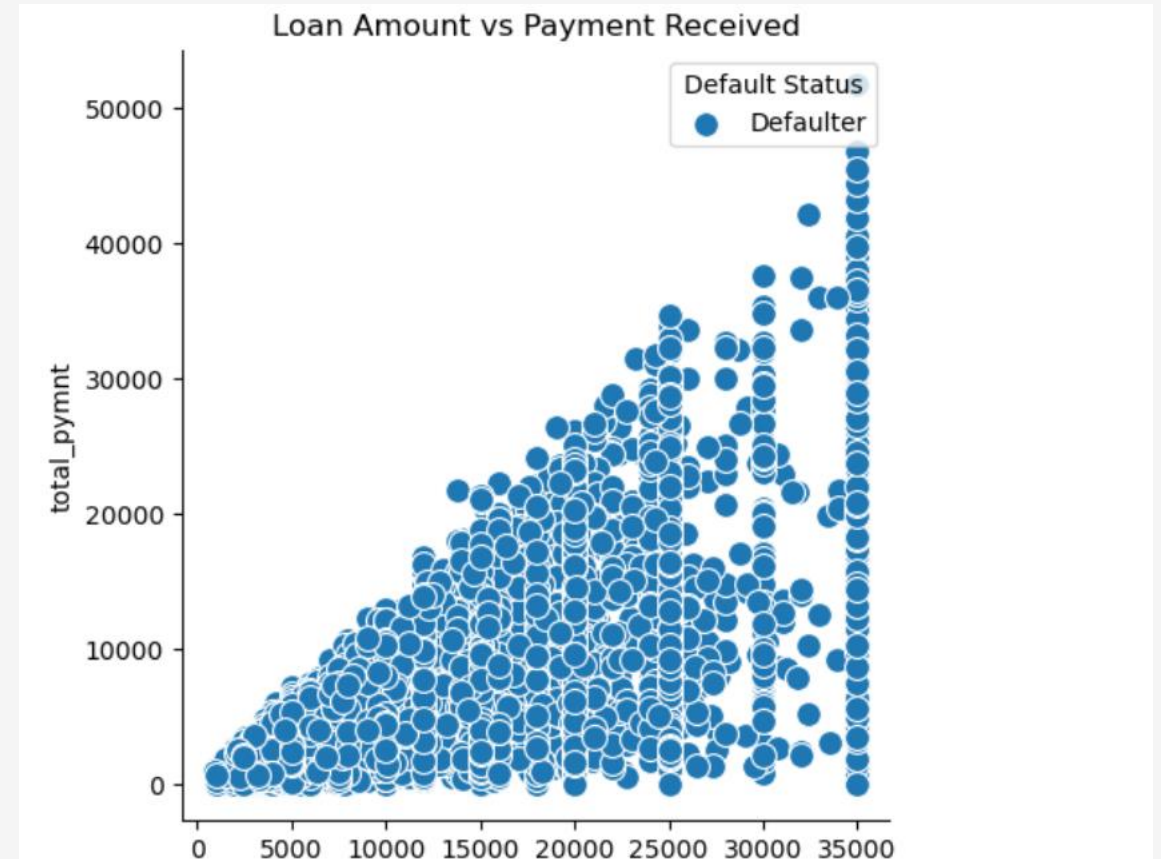
# Summary of Observations from the Dataset

---

16

**Bivariate –Loan Amnt v/s Payment Received**

**Payment received is proportional to the loan amount**



# Summary of Observations from the Dataset

## 16 Correlation -

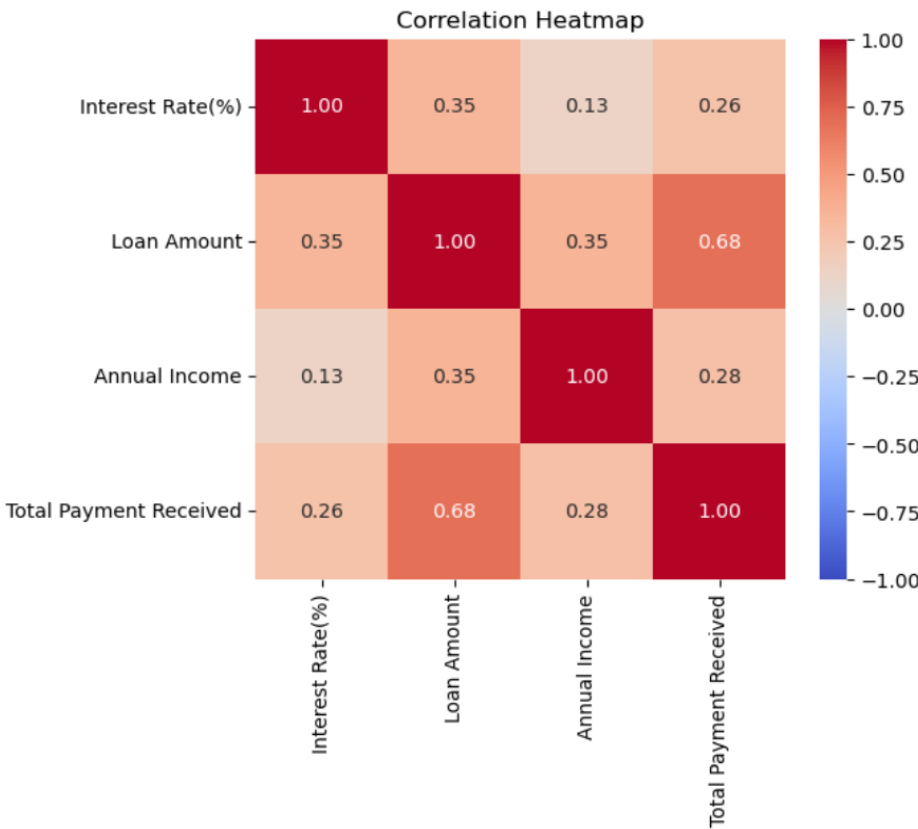
Interest Rate(%), Loan Amount, Annual Income, Total Payment Received

Weak

- Interest Rate(%) & Annual Income
- Interest Rate(%) & Total Payment Received
- Annual Income & Total Payment Received

Strong

- Loan Amount & Total Payment Received



	Interest Rate(%)	Loan Amount	Annual Income	Total Payment Received
Interest Rate(%)	1.000000	0.351216	0.132252	0.264457
Loan Amount	0.351216	1.000000	0.352359	0.680814
Annual Income	0.132252	0.352359	1.000000	0.279975
Total Payment Received	0.264457	0.680814	0.279975	1.000000

# Thank You

-Shreya (Batch C-71)