

Lending Club Case Study

This project is to find the riskiest borrowers of a banking firm. This data can help a bank identify the at risk borrowers and scrutinize them.

Total Number of borrowers the bank has is 39717

There were some unnecessary columns as well as columns with NULL values which were cleaned up first.

After clean up, these are the list of columns

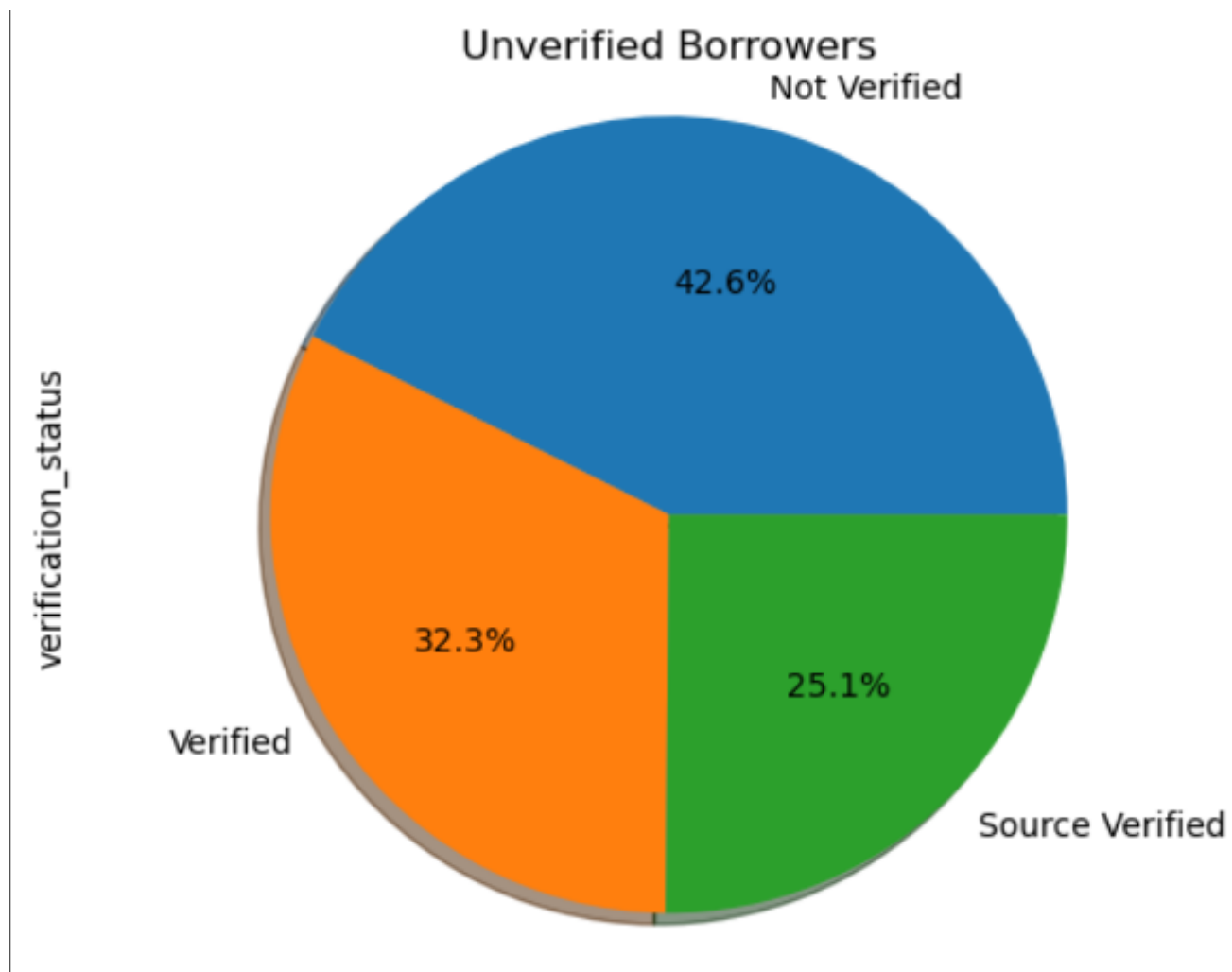
```
df2.isnull().sum()
```

[210]		
...	id	0
	member_id	0
	loan_amnt	0
	funded_amnt	0
	funded_amnt_inv	0
	term	0
	int_rate	0
	installment	0
	grade	0
	sub_grade	0
	emp_title	0
	emp_length	1075
	home_ownership	0
	annual_inc	0
	verification_status	0
	issue_d	0
	loan_status	0
	pymnt_plan	0
	url	0
	desc	12940
	purpose	0
	title	11
	zip_code	0
	addr_state	0
	dti	0
	...	
	chargeoff_within_12_mths	56
	delinq_amnt	0
	pub_rec_bankruptcies	697
	tax_liens	39

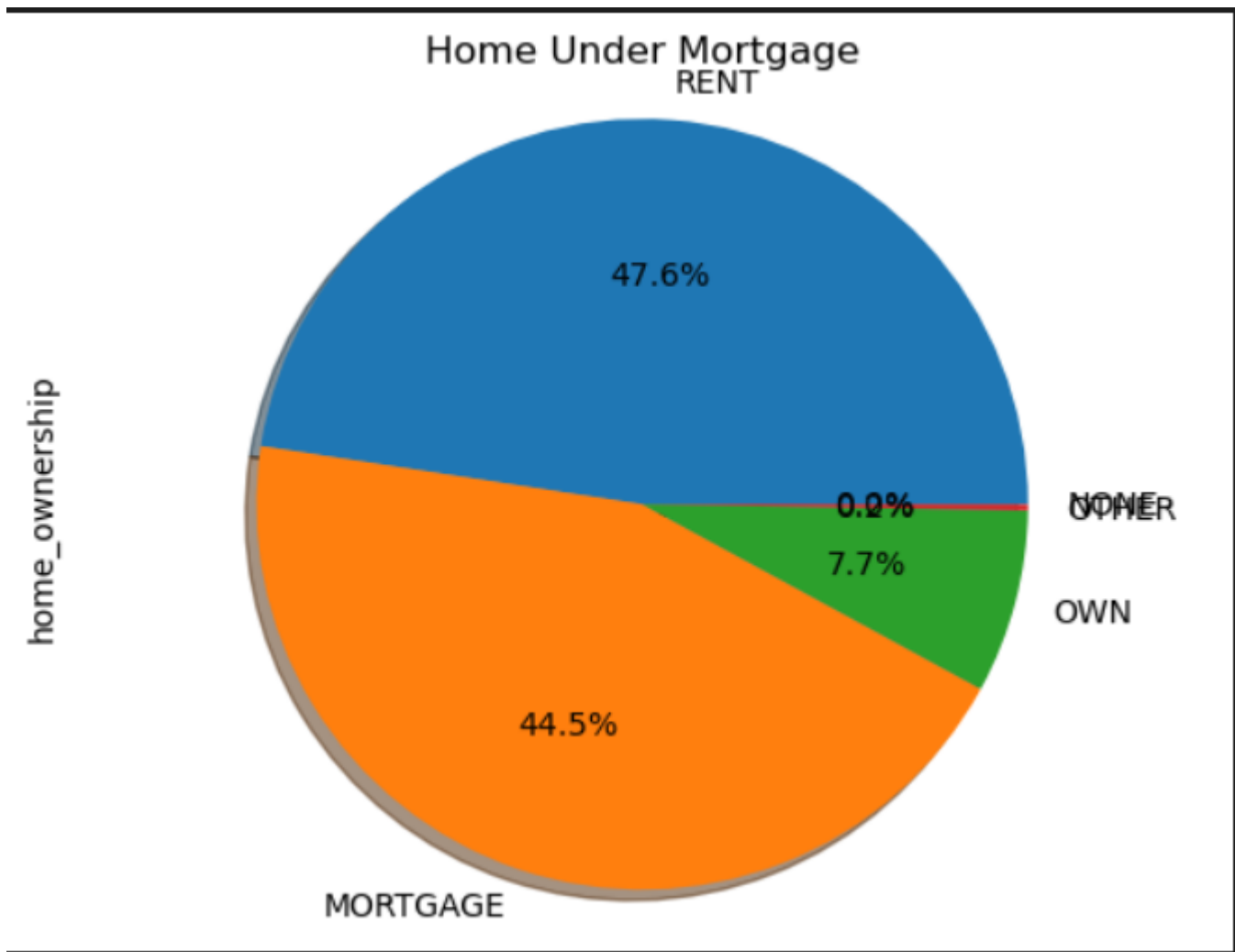
To identify Risky borrowers, below factors can be applied

1. If Source of borrower is not verified, then its risky
2. If home of borrower is already in mortgage, then its a risk factor
3. Ratio of salary vs loan disbursed should be less. So higher the percentage of loan amount compared to salary greater the risk
4. Borrowers tagged as "Charged Off" are risks

So the Percentage of Unverified Borrowers was found to be as Below:



Similarly, percentage of borrowers with their home already in mortgage was as below:

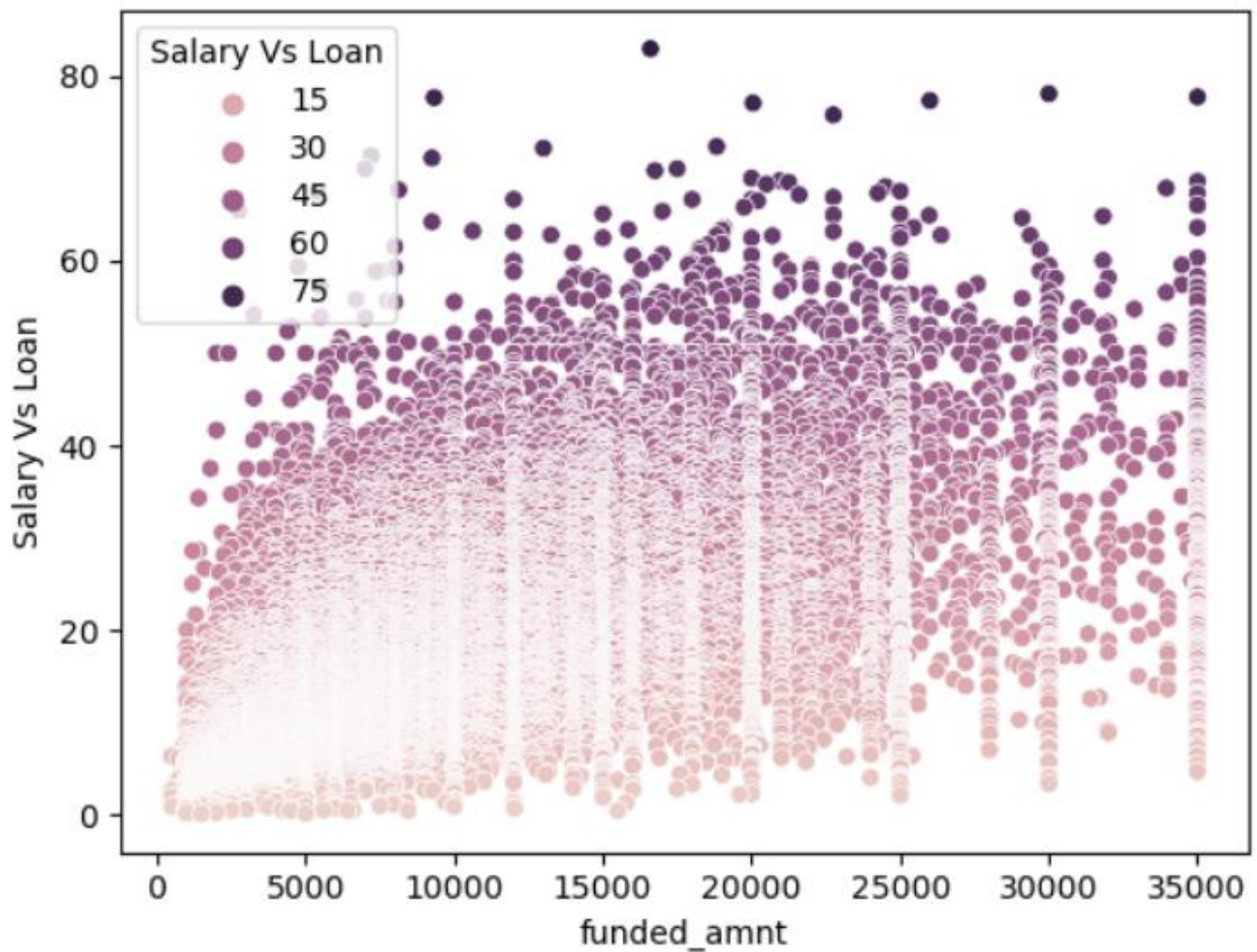


When we calculated the borrowers whose loan paying capacity based on their salary is at high risk, we found the top 10 at risk borrowers:

	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	term	int_rate	installment	grade	sub_grade	...	last_pymnt_d	last_pymnt_amnt	last
39200	242642	237046	16600	16600	3590.670000	36 months	9.76%	533.77	B	B2	...	11-Feb	542.62	
10965	822845	1031350	30000	30000	27212.914320	60 months	19.29%	783.02	E	E4	...	16-May	783.02	
11969	805663	1011728	35000	35000	34975.000000	60 months	12.99%	796.18	C	C1	...	12-Dec	28069.18	
39431	200886	200841	9325	9325	1868.290000	36 months	8.38%	293.85	A	A5	...	11-Jan	300.60	
19724	670629	857337	26000	26000	25975.000000	60 months	18.25%	663.78	F	F1	...	11-Mar	26398.06	
39230	249699	249616	20050	20050	3025.004566	36 months	9.76%	644.71	B	B2	...	9-Dec	644.71	
9333	838602	1048698	22750	22750	22725.000000	60 months	15.62%	548.66	D	D1	...	14-Jun	548.66	
14561	763793	964449	18825	18825	18825.000000	60 months	11.99%	418.66	B	B5	...	15-Sep	4016.99	
30701	497118	637218	13000	13000	12173.306770	36 months	7.88%	406.65	A	A5	...	12-Sep	2387.53	
7203	884937	1100520	7200	7200	7175.000000	36 months	6.03%	219.14	A	A1	...	14-Oct	221.23	

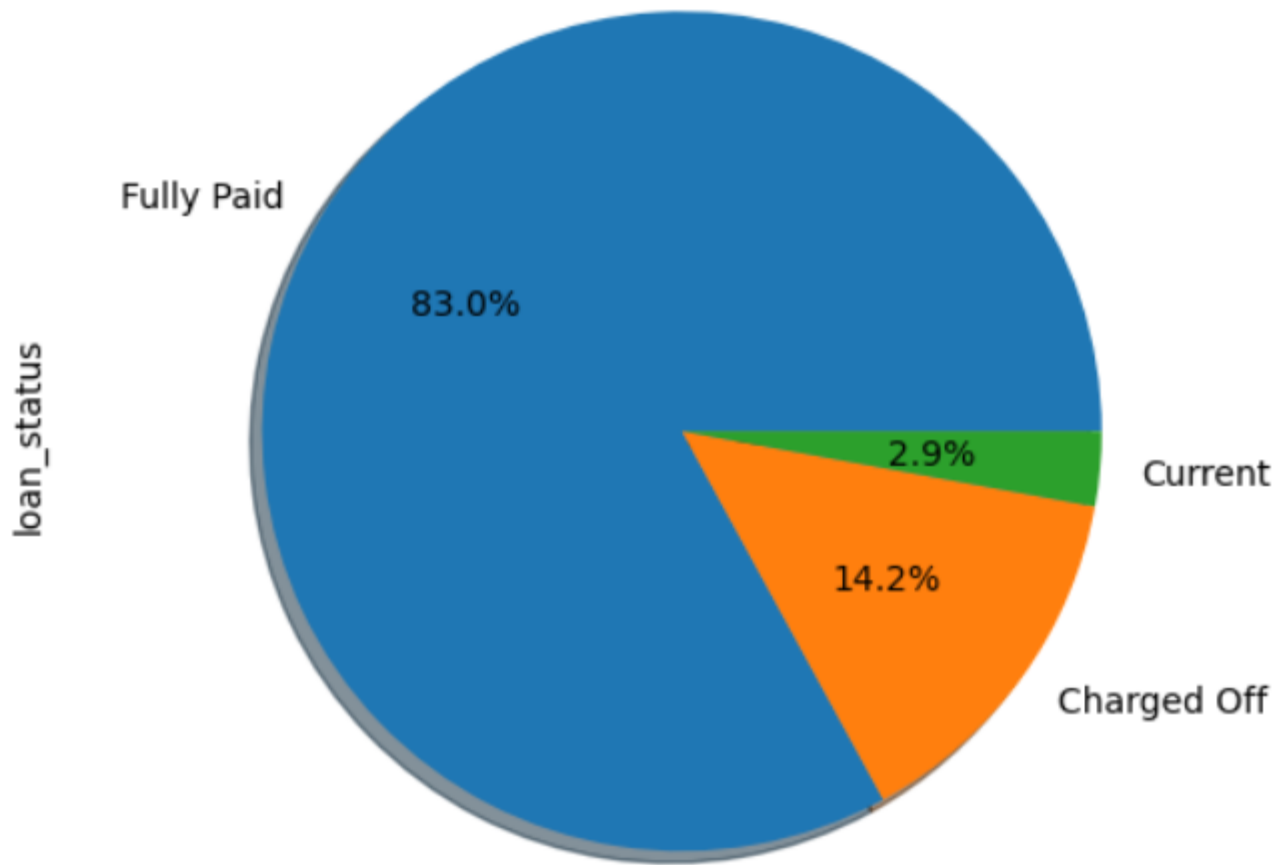
10 rows × 52 columns

If we plot a scatter plot based on funded amount and the loan paying capacity risk, we get as below. Most at risk are in the 10K-20K range.



Lastly, the percentage of borrowers who are charged off are depicted in below chart:

Charged Off Borrowers



If we combine the above 4 points, we get the most riskiest borrowers. Upon calculation we get 836 number of borrowers who should be under top scrutiny.

Below are the top 10 most riskiest borrowers:

	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	term	int_rate	installment	grade	sub_grade	...	last_pymnt_d	last_pymnt_amnt
16303	734645	931193	31000	31000	28412.413900	60 months	13.06%	706.30	C	C2	...	12-Aug	1447.91
32845	476966	604474	25000	25000	22125.000000	36 months	15.31%	870.39	D	D3	...	10-Apr	870.39
31867	487311	621181	25000	25000	22975.000000	36 months	10.25%	809.62	B	B2	...	11-Oct	43.65
38264	361356	369771	25000	25000	2711.178345	36 months	11.46%	823.93	B	B5	...	9-Mar	823.93
39102	256365	256357	25000	25000	750.000311	36 months	10.71%	815.04	B	B5	...	9-Jan	815.04
32320	483962	615973	25000	25000	24925.000000	36 months	16.45%	884.44	E	E1	...	10-Oct	884.44
30441	506881	653616	25000	25000	24842.450000	36 months	10.99%	818.38	B	B4	...	11-Feb	12818.38
32842	476988	604531	25000	25000	22395.410000	36 months	18.09%	904.88	F	F1	...	11-Aug	120.05
38956	292921	292891	25000	25000	9574.999856	36 months	13.55%	848.99	D	D4	...	8-Jul	848.99
39341	224381	223623	25000	25000	1750.005358	36 months	16.15%	880.78	F	F3	...	10-Jul	880.88

10 rows x 52 columns