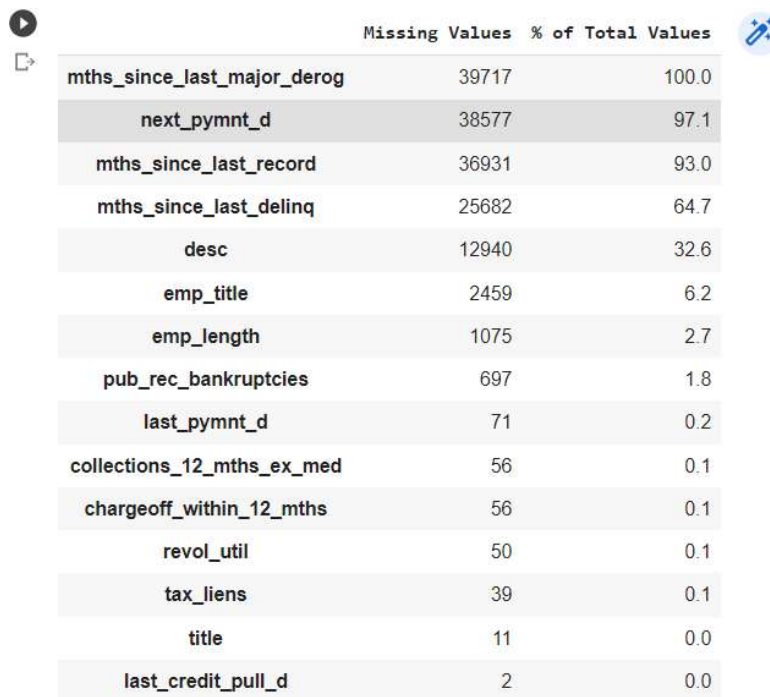Data understanding:

- Data talks about the borrowers info such as loan amount, repayment etc
- NA columns are removed
- Identified string & no are mixed in column "term" and that is removed

Data understanding:

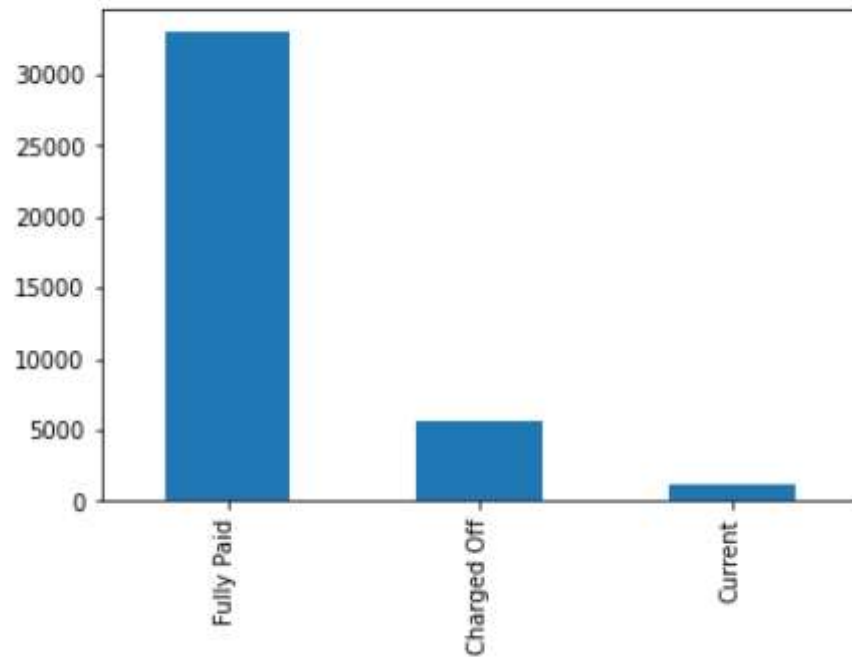- Identified percentage of missing values and those columns were removed

	Missing Values	% of Total Values
mths_since_last_major_derog	39717	100.0
next_pymnt_d	38577	97.1
mths_since_last_record	36931	93.0
mths_since_last_delinq	25682	64.7
desc	12940	32.6
emp_title	2459	6.2
emp_length	1075	2.7
pub_rec_bankruptcies	697	1.8
last_pymnt_d	71	0.2
collections_12_mths_ex_med	56	0.1
chargeoff_within_12_mths	56	0.1
revol_util	50	0.1
tax_liens	39	0.1
title	11	0.0
last_credit_pull_d	2	0.0

Data Analysis:
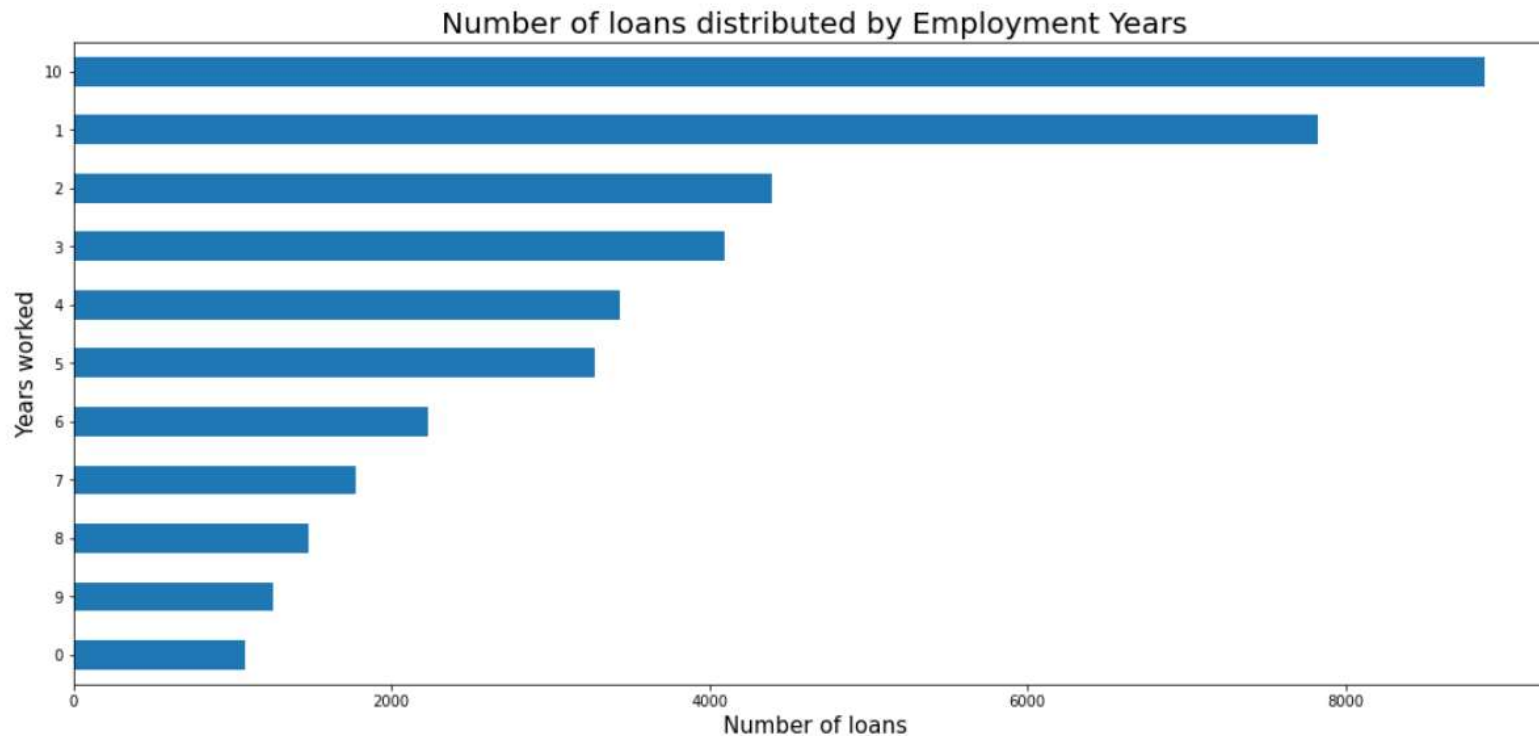Fully paid members are more than the other category

`<matplotlib.axes._subplots.AxesSubplot at 0x7fd1e3ea1ad0>`

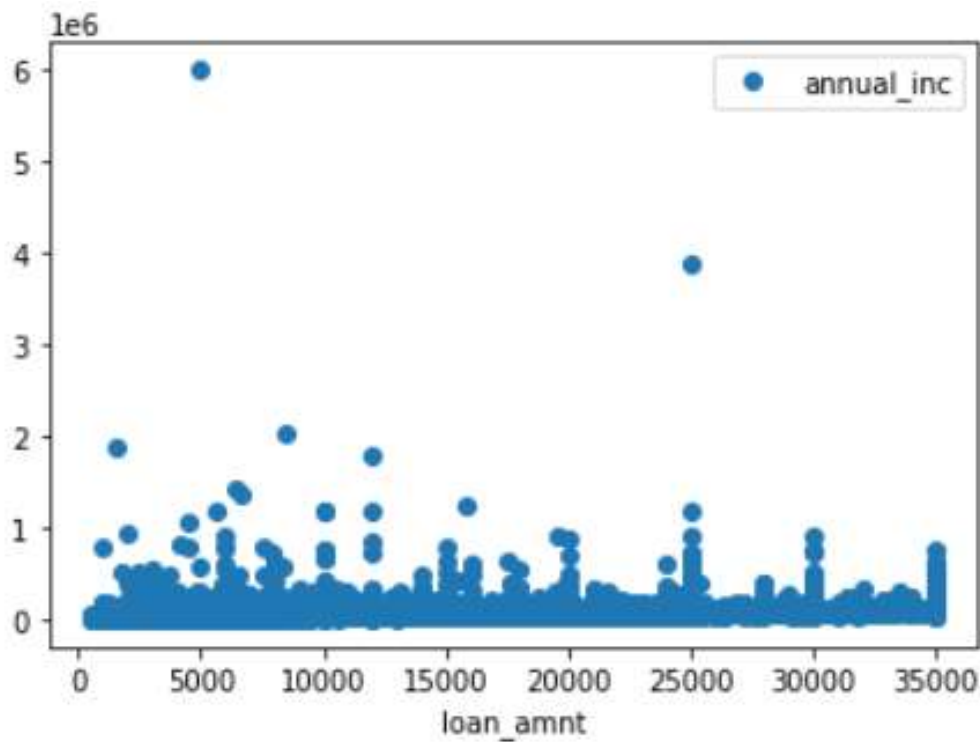Data Analysis:
The column looks fine. Also, it can be seen that people who have worked for 10 or more years are more likely to take a loan.

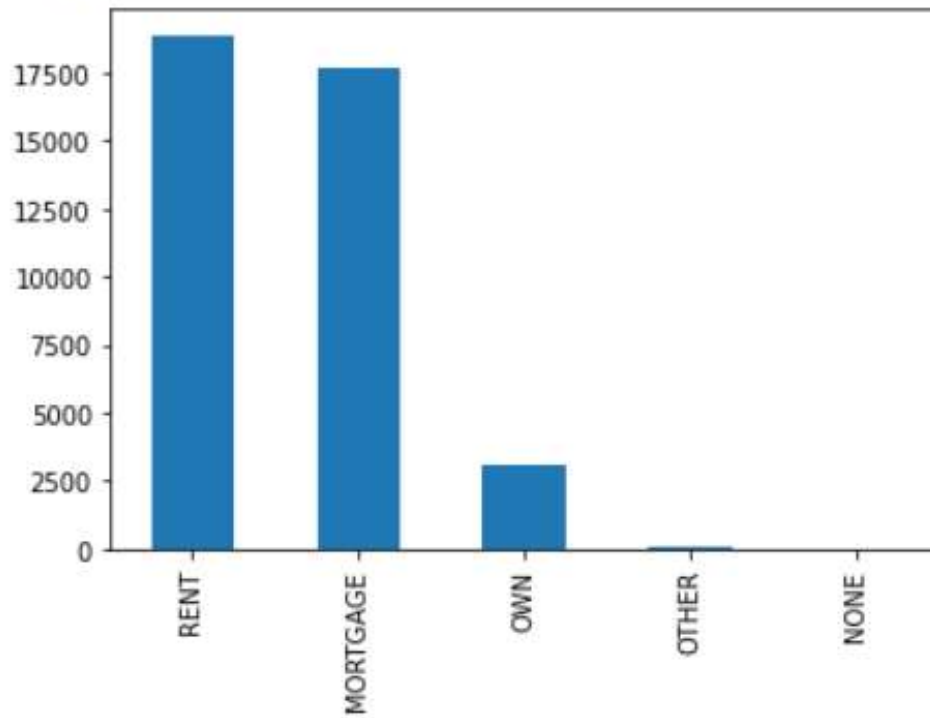Number of loans distributed by Employment Years

Data Analysis:
figure indicates that more than
90% borrowers taken loan amount upto 35000 who have got average income of
100000

Data Analysis:
Below graph indicates that most people are in RENT & MORTAGE homes getting loan than own house

Data Analysis:
Correlation

- Correlation between total_pymnt_inv and total_pymnt is: 0.97
→ Repayment structure follows the investment as per the given data

Recommendations:

- High interest payments leads to high losses, it is always better to consider low interest rates – So low risk at lower interest

- One of the driving factor for lender is , it can be seen that people who have worked for 10 or more years are more likely to take a loan.

- Repayment structure is good → Average income of 100000 with more than 10 years of employment . So using this analysis the fund limit can be deciding factor for finance company

- Data shows that the most (>90%) borrowers from Rented &Mortgage home than the Own home. Means more borrowers these categories approaching for loan.