LENDING CLUB CASE STUDY

ABSTRACT

- **Background:** This project centers around a consumer finance company specializing in urban lending, striving to balance business growth and minimize financial losses due to loan defaults.
- **Business Problem:** The core challenge is to identify loan applicants likely to default, thereby enabling the company to make more informed lending decisions. The primary goal is to reduce credit loss and optimize lending strategies.
- **Dataset:** The project utilizes historical data on past loan applicants, including their repayment outcomes. This dataset provides valuable insights into applicant profiles and their loan repayment behavior, serving as the foundation for exploratory data analysis (EDA) and risk analytics.

APPROACH

Data Cleaning & Understanding → **Univariate Analysis** → **Bivariate Analysis** → **Observations** → **Recommendations**

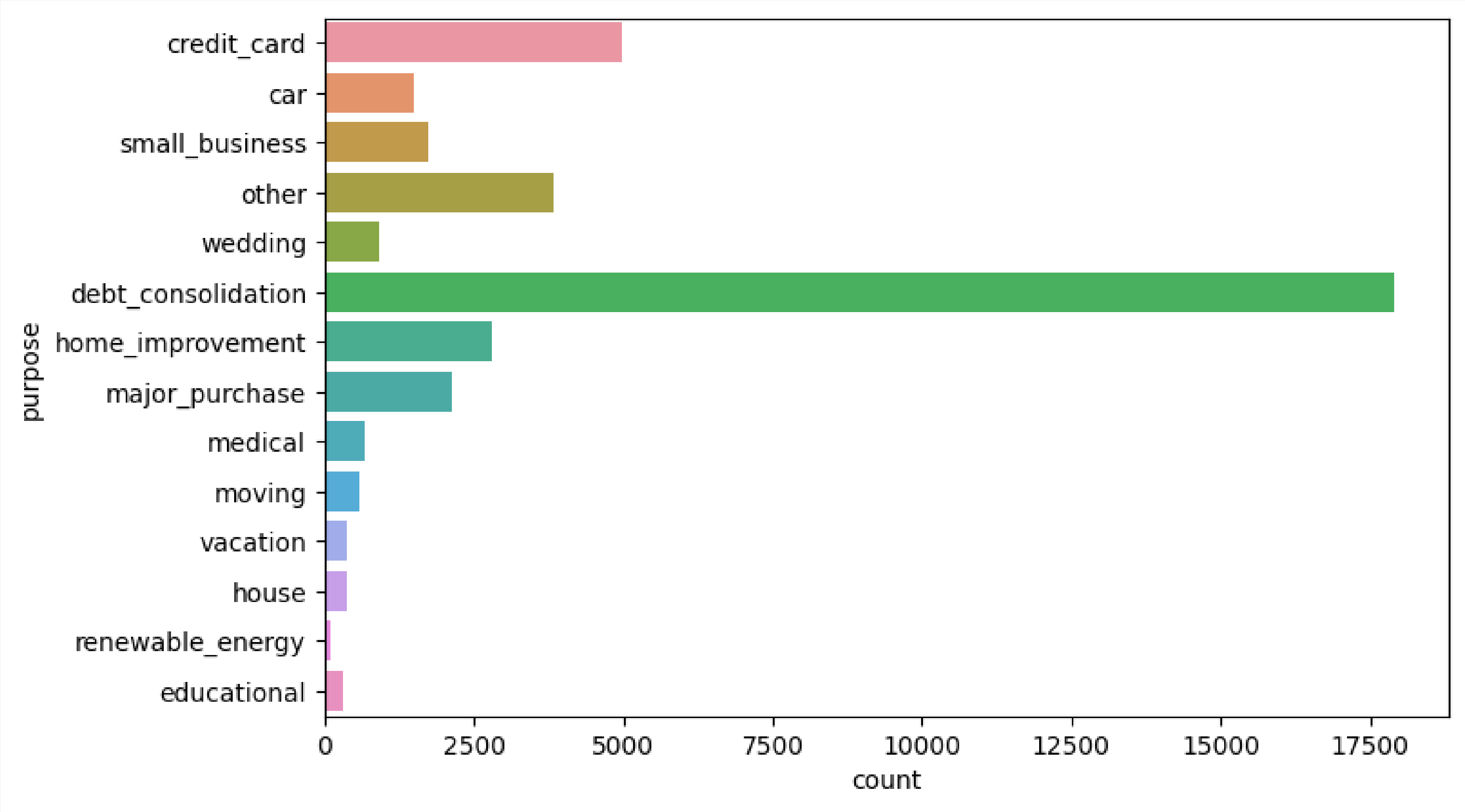
Load the data and perform necessary cleaning, remove outliers, checking null values, dropping irrelevant columns, etc.

Analyze relevant columns, plotting bar plot, box plots, etc.

Analyze multiple columns together to find relation, plot correlation plot, analyze numeric columns, etc.

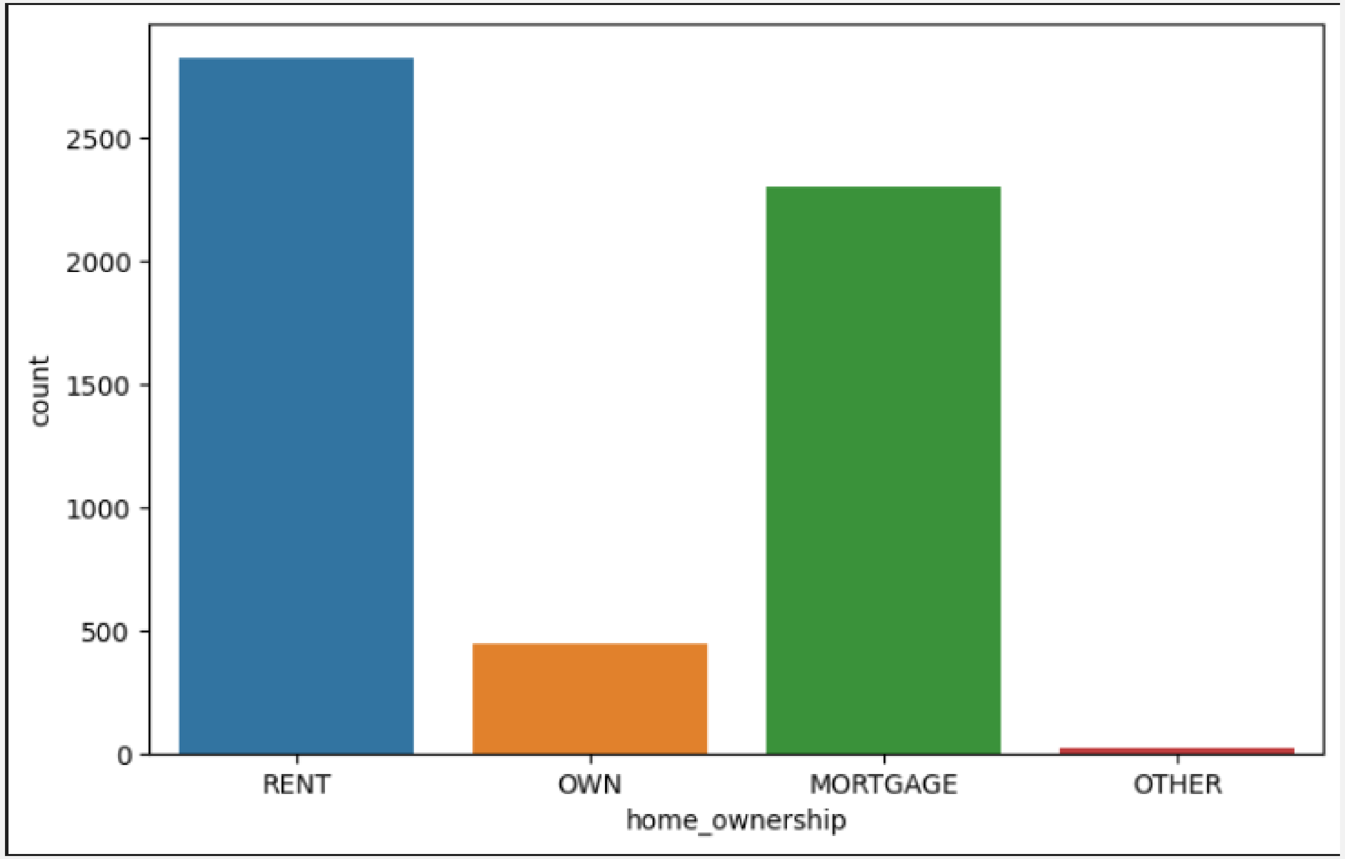
Conclusions and recommendations based on the analysis to reduce the risk of approving loans to the customers who are likely to default

ANALYSIS



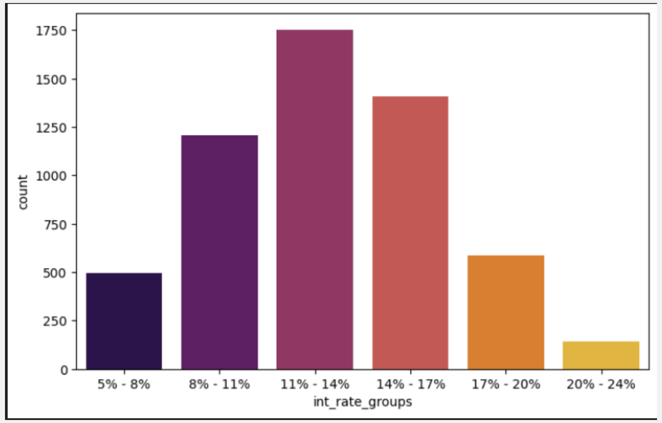
- Most of the loans were taken to clear debt.
- Credit Card and Other reasons also contribute to a good portion of loans.

ANALYSIS

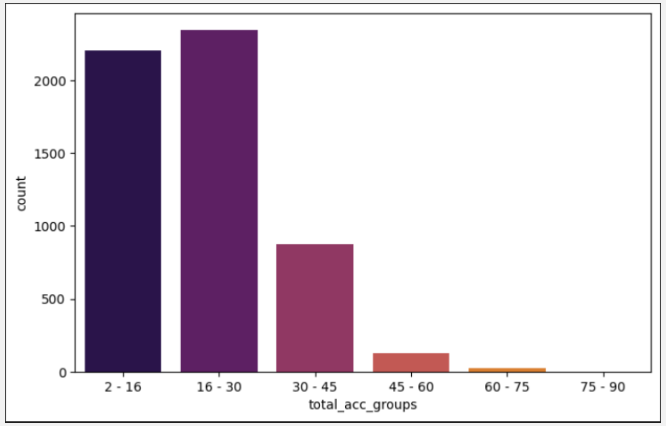


- The above visual show that only a minority of applicants own a house. Majority of them either rent or are on mortgage.
- Most applicants take loan for debt consolidation, followed by credit card dues.

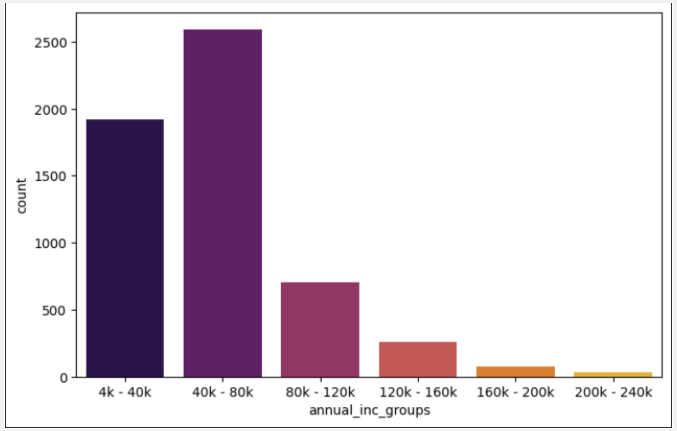
ANALYSIS



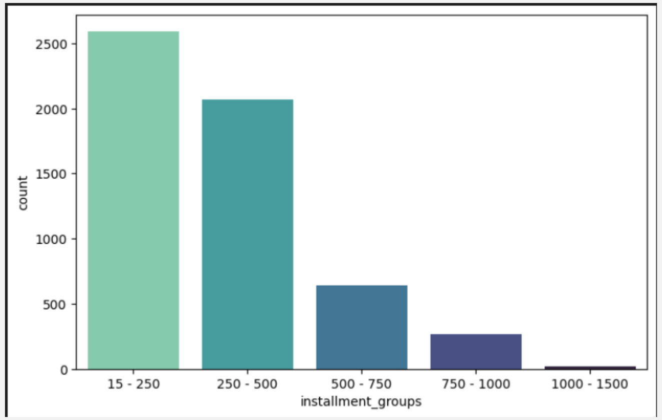
1750
1500
1250
1000
750
500
250
0
count
5% - 8%
8% - 11%
11% - 14%
14% - 17%
17% - 20%
20% - 24%
int_rate_groups



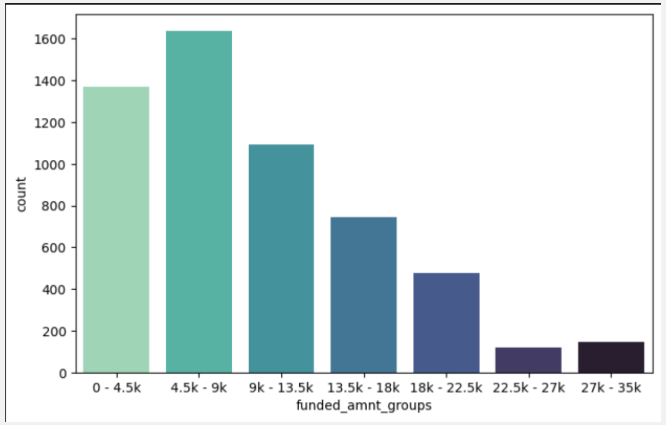
2000
1500
1000
500
0
count
2 - 16
16 - 30
30 - 45
45 - 60
60 - 75
75 - 90
total_acc_groups



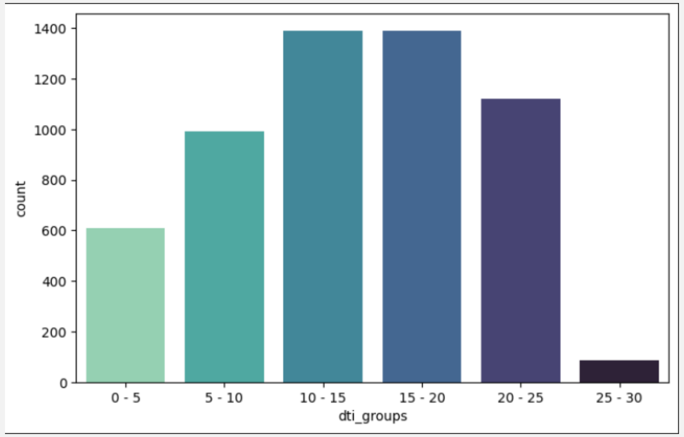
2500
2000
1500
1000
500
0
count
4k - 40k
40k - 80k
80k - 120k
120k - 160k
160k - 200k
200k - 240k
annual_inc_groups



2500
2000
1500
1000
500
0
count
15 - 250
250 - 500
500 - 750
750 - 1000
1000 - 1500
installment_groups

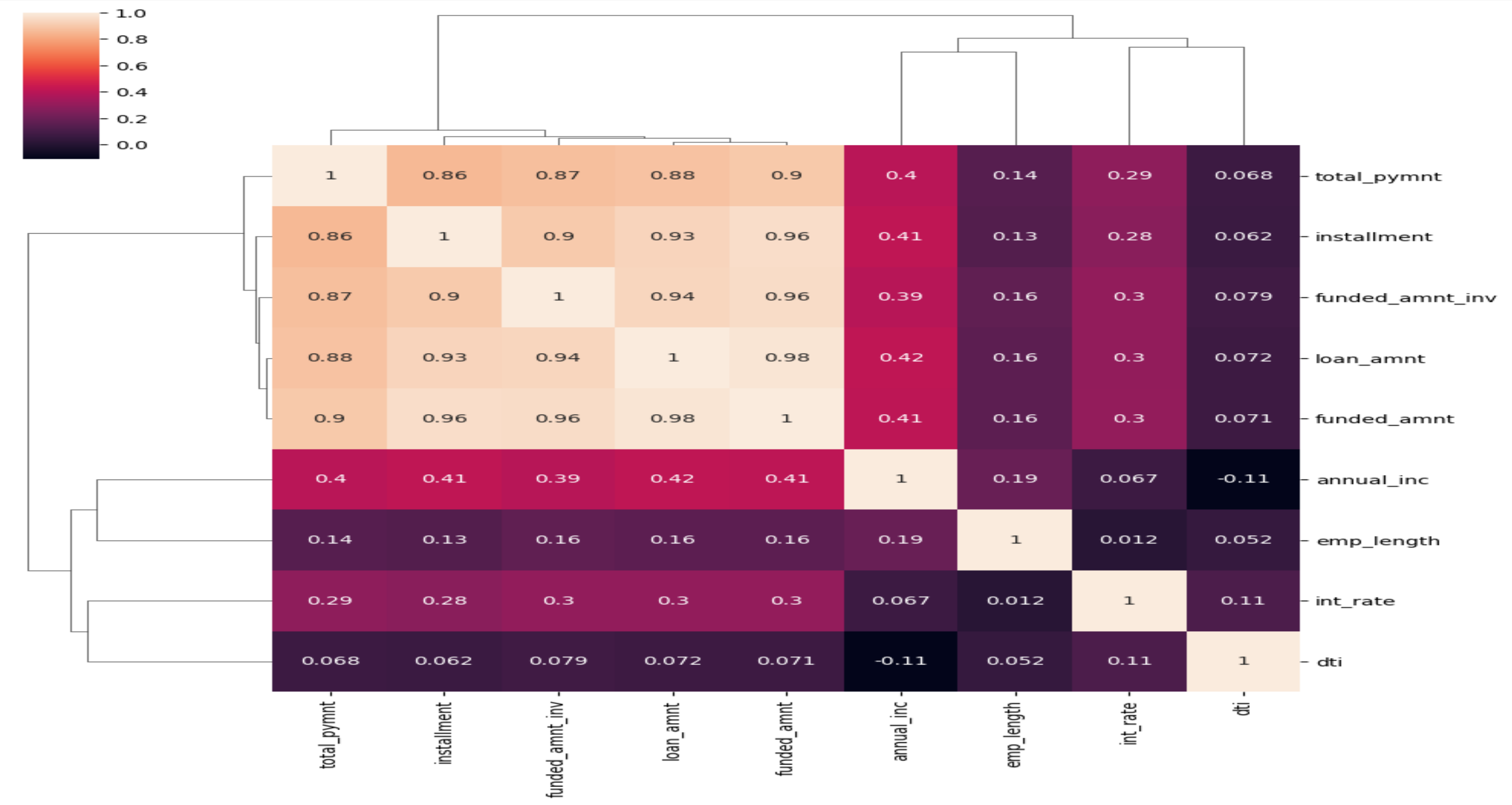


1600
1400
1200
1000
800
600
400
200
0
count
0 - 4.5k
4.5k - 9k
9k - 13.5k
13.5k - 18k
18k - 22.5k
22.5k - 27k
27k - 35k
funded_amnt_groups



1400
1200
1000
800
600
400
200
0
count
0 - 5
5 - 10
10 - 15
15 - 20
20 - 25
25 - 30
dti_groups

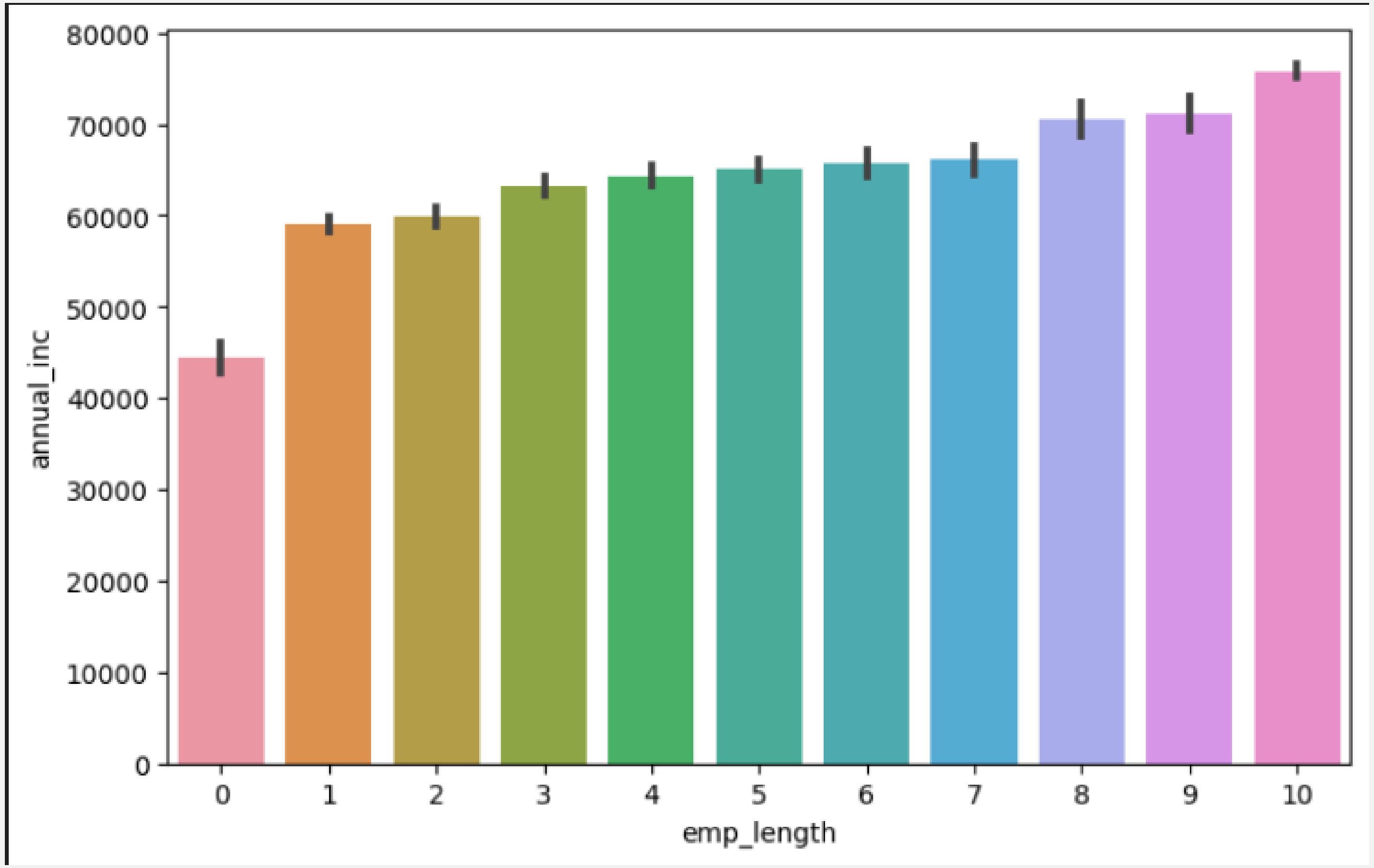
ANALYSIS



	total_pymnt	installment	funded_amnt_inv	loan_amnt	funded_amnt	annual_inc	emp_length	int_rate	dti
total_pymnt	1	0.86	0.87	0.88	0.9	0.4	0.14	0.29	0.068
installment	0.86	1	0.9	0.93	0.96	0.41	0.13	0.28	0.062
funded_amnt_inv	0.87	0.9	1	0.94	0.96	0.39	0.16	0.3	0.079
loan_amnt	0.88	0.93	0.94	1	0.98	0.42	0.16	0.3	0.072
funded_amnt	0.9	0.96	0.96	0.98	1	0.41	0.16	0.3	0.071
annual_inc	0.4	0.41	0.39	0.42	0.41	1	0.19	0.067	-0.11
emp_length	0.14	0.13	0.16	0.16	0.16	0.19	1	0.012	0.052
int_rate	0.29	0.28	0.3	0.3	0.3	0.067	0.012	1	0.11
dti	0.068	0.062	0.079	0.072	0.071	-0.11	0.052	0.11	1

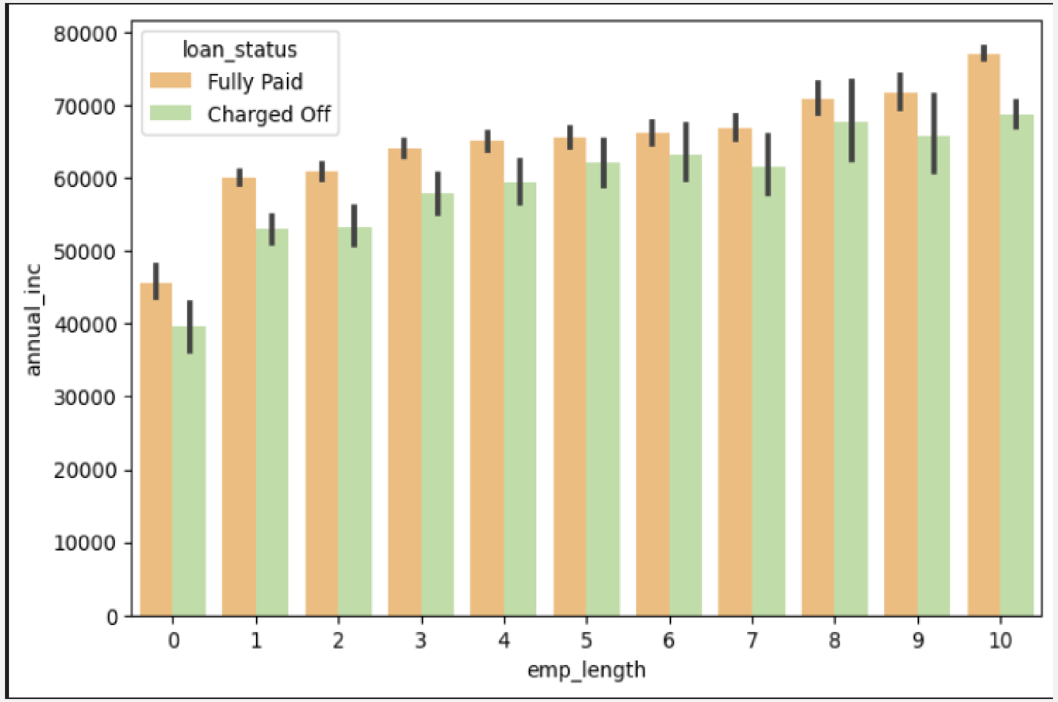
- As expected, there's a high correlation among total_pymnt, installment, funded_amnt_inv, loan_amnt, funded_amnt.
- There's a negative correlation between annual_inc and dti. DTI is the ratio of total monthly debt payments and monthly income.
- A negative correlation here makes sense because the more you earn, the less likely you'll be in debt, and vice-versa.

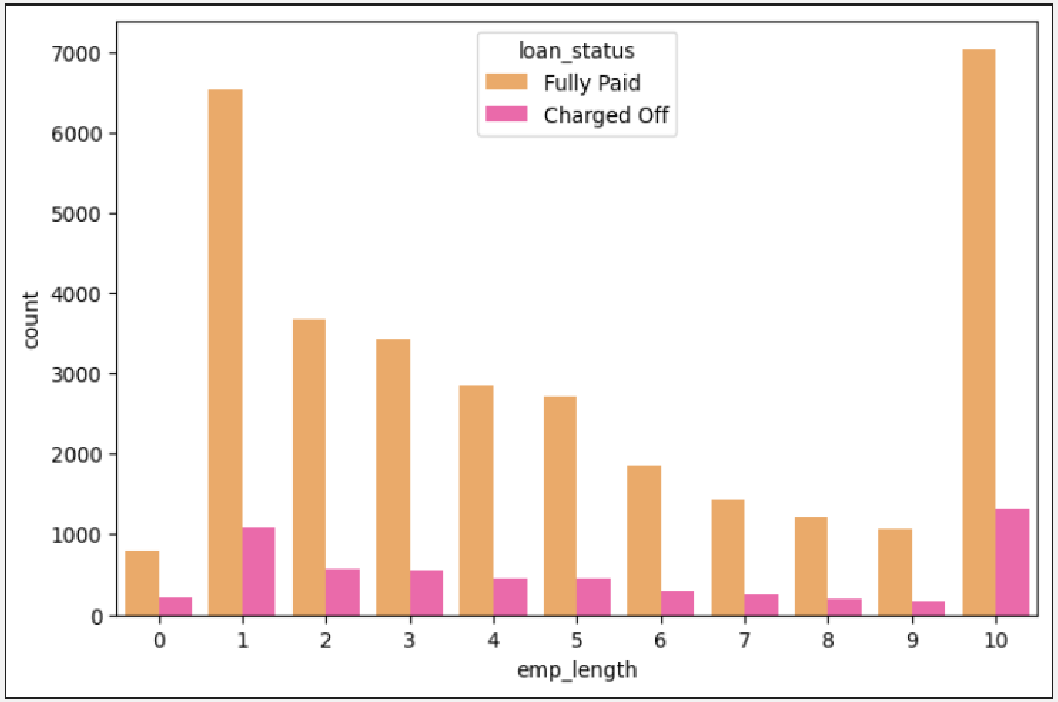
ANALYSIS



- The above visual shows that as an applicant grows in professional experience, his annual income also increases.
- It also shows there's a sudden increase in annual income from fresher to an year of experience.

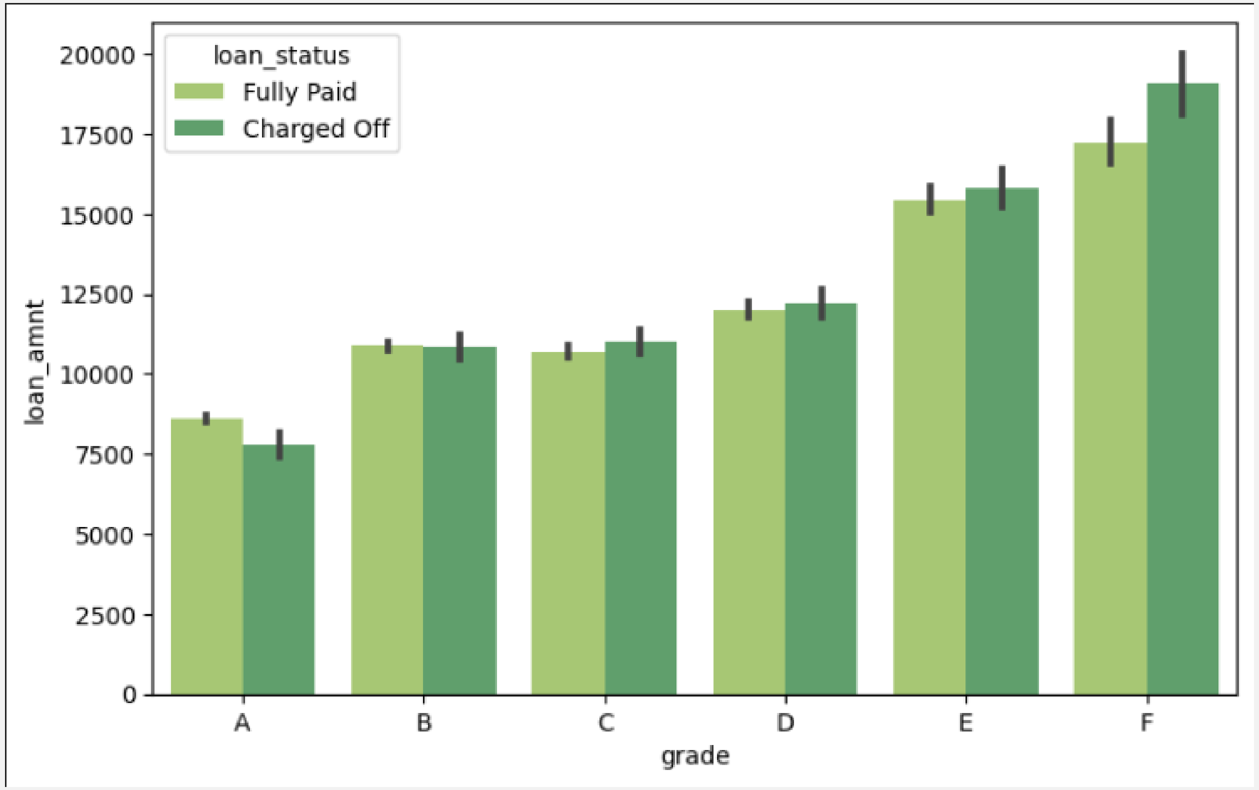
ANALYSIS





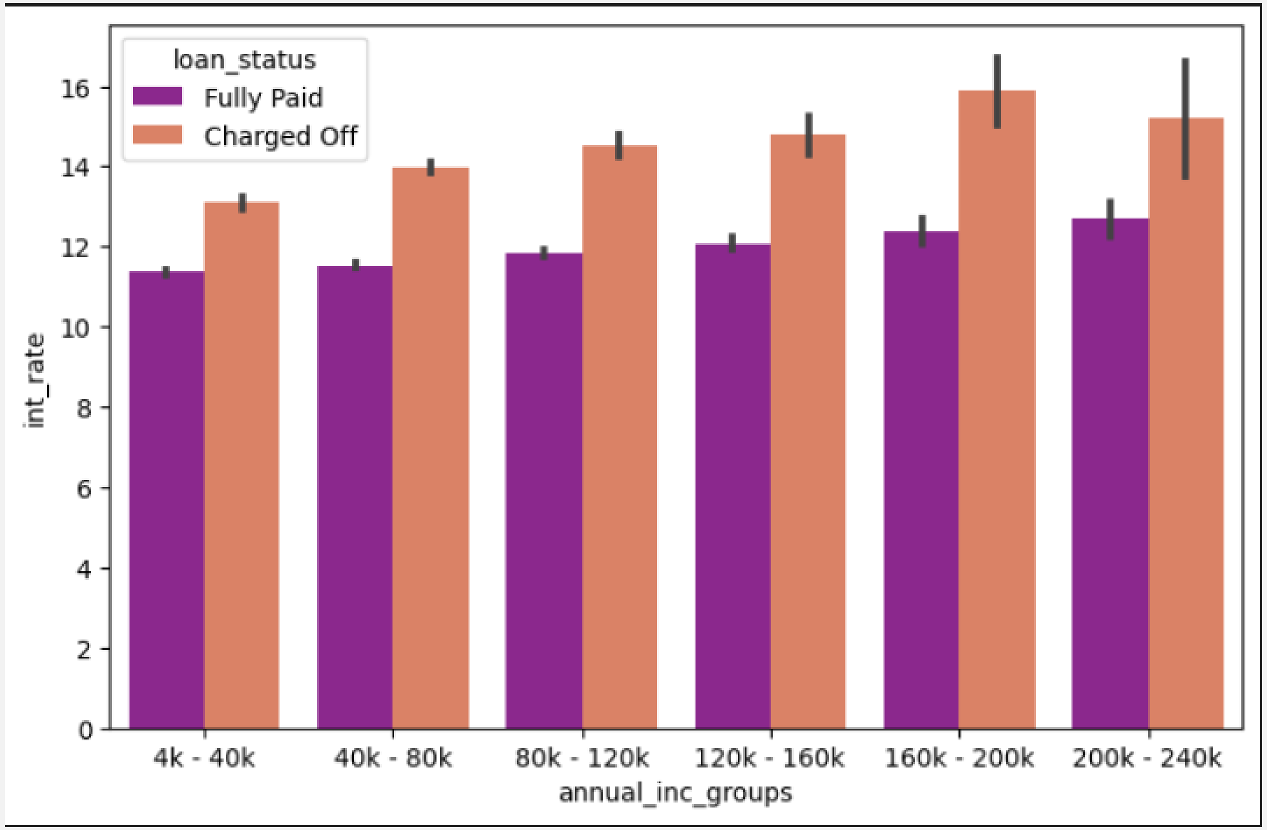
- Combining interpretation from both the plots above, most of the loans are issued to applicants having an year of experience or more than 10 years of experience.
- The default rate in both these years is high too. For applicants with an year of experience, this might be due to them taking loans with sudden increase in their annual income and later realizing they can't pay it off.

ANALYSIS



- Charged Off rate is high for grade E and F, with grade F being the highest.

ANALYSIS



- Charged Off ratio is the highest for applicants with annual_inc between 160k-200k and int_rate over 15%

CONCLUSION

Based on the Univariate Analysis, applicants may default when they have:

- house_ownership as "RENT"
- purpose as "debt_consolidation", that is, taking loan to clear other debts
- dti between 10-20
- annual_inc between 40k-80k
- int_rate between 11%-14%
- funded_amnt between 4.5k-9k

Based on the Bivariate Analysis, applicants may default when they have:

- emp_length more than 10 years and annual_inc above 60k
- int_rate between 20%-24% and annual_inc above 60k
- emp_length more than 10 years and loan_amnt greater than 14k
- verified loan and loan_amnt above 16k
- term as 60 months and loan_amnt more than between 14k-16k
- grade is F and loan_amnt greater than 15k