

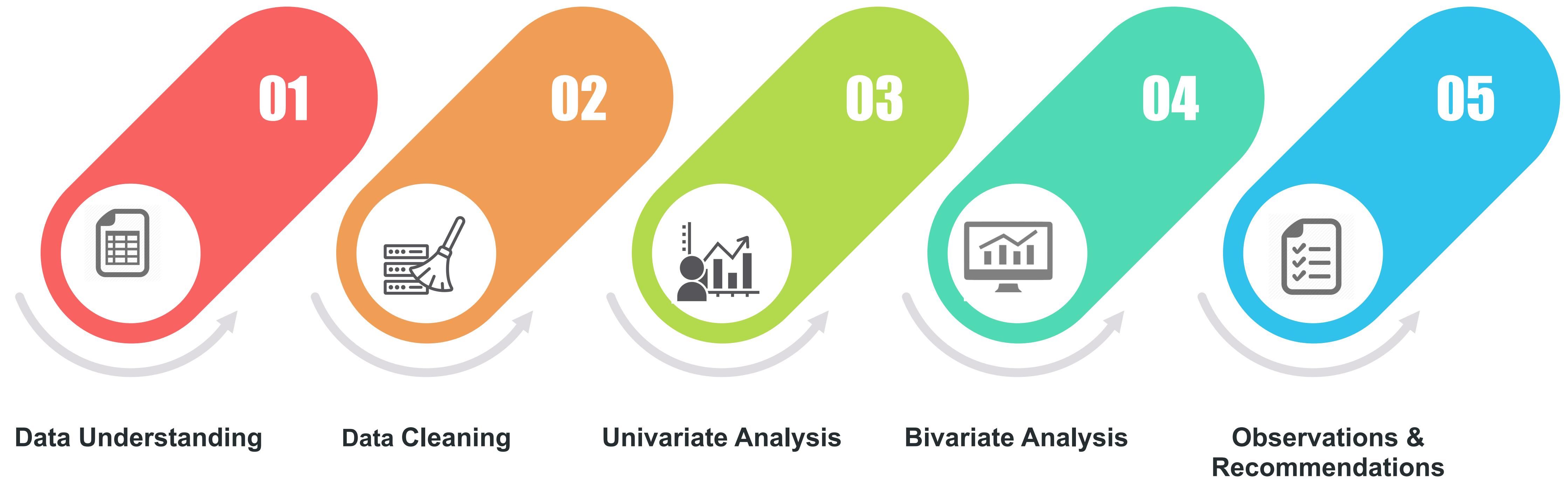
Lending Club Case Study

Naveen Kumar Banala
Tegbir Singh

Case Study overview

- Context
 - Finance company lending various types of loans to urban customers. The data provided contains the information about the loan applicants and their status. Identifying various factors influencing the customer to default the loan using EDA.
- Problem Statement
 - While lending the loans to customers, there are two types of risks associated with the bank's decision.
 1. Approving the loans to the defaulters will lead to financial loss.
 2. Not approving loans to those who are likely to repay the loan, will also result in financial loss.
- Objectives
 - Identify the patterns that indicate if the person is likely to default which will help the financial company to take necessary actions like denying the loan, reducing the loan amount, lending at higher interest rate.

Analysis Approach

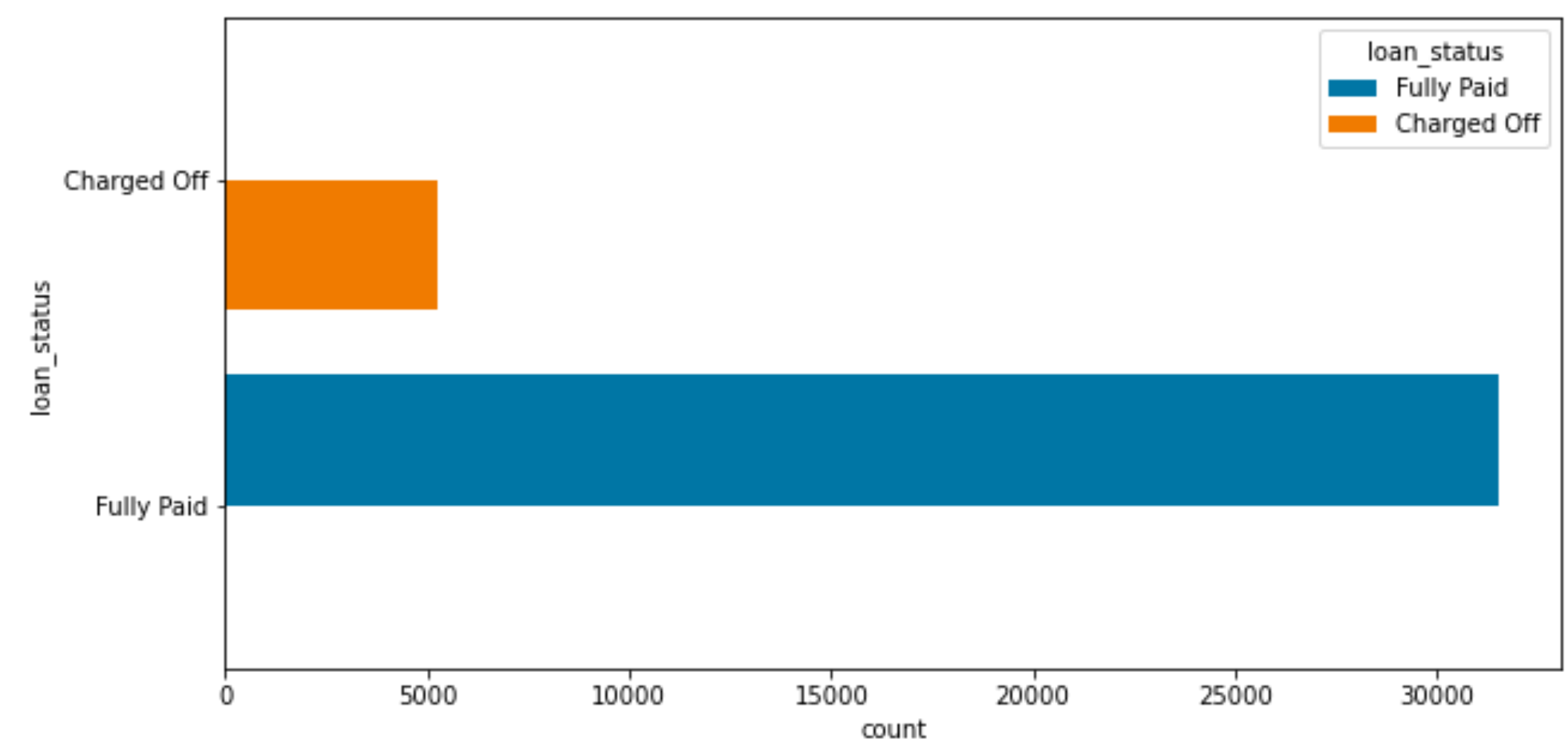


Data Cleaning

- Dropped null columns(54 columns), Single value columns (8 columns) and unique value columns(7 columns)
- Dropped missing values which has less than 2% missing values (4 columns).
- Dropped rows for loan status as “current” as this is not useful for the analysis and converted the loan status to numeric for analysis.
- Removing the columns which are not relevant to the analysis for example the values which are calculated after issuing the loan.
- Removed the outliers which are identified using box plot.

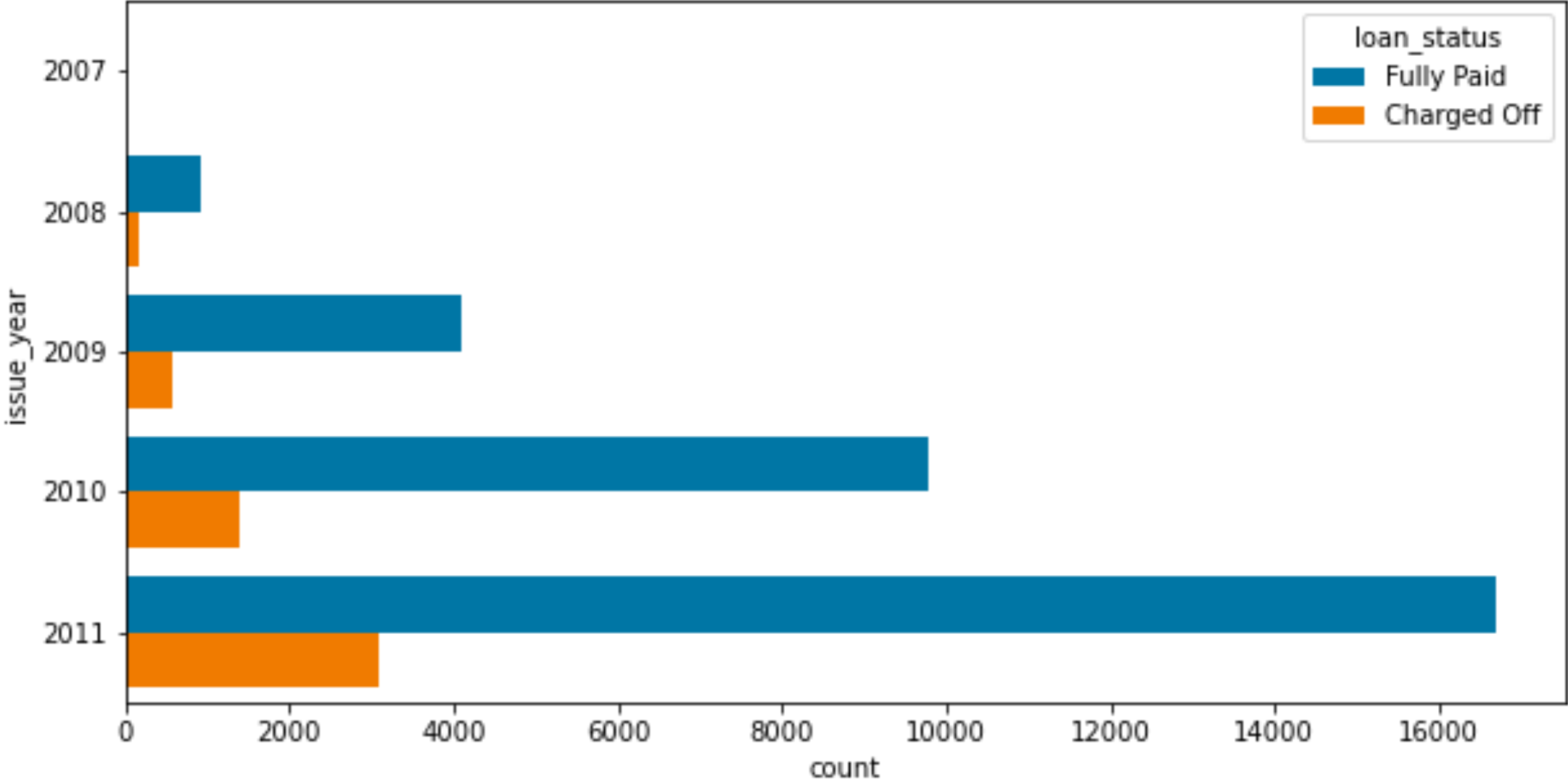
Loan Status

Loan Status



FullyPaid	31005	85.8%
Charged Off	5131	14.2%

Issue Year with respect to Loan Status

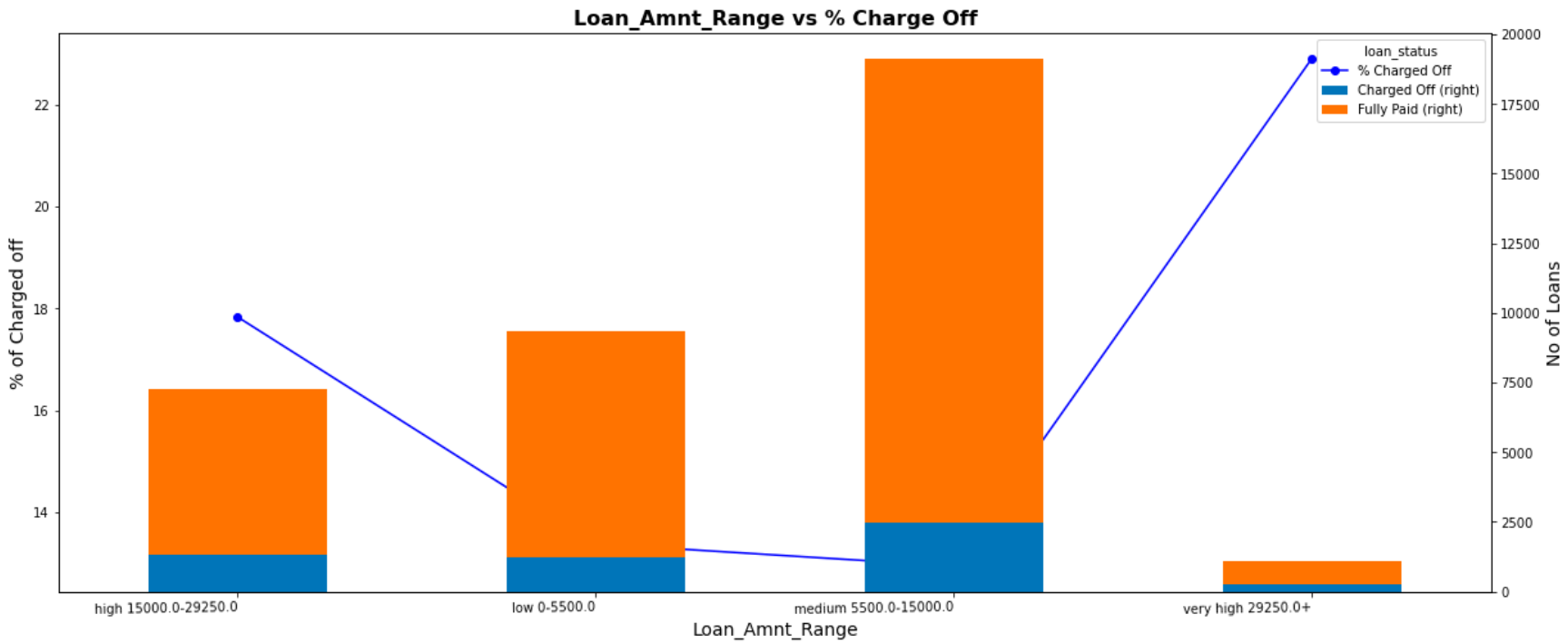
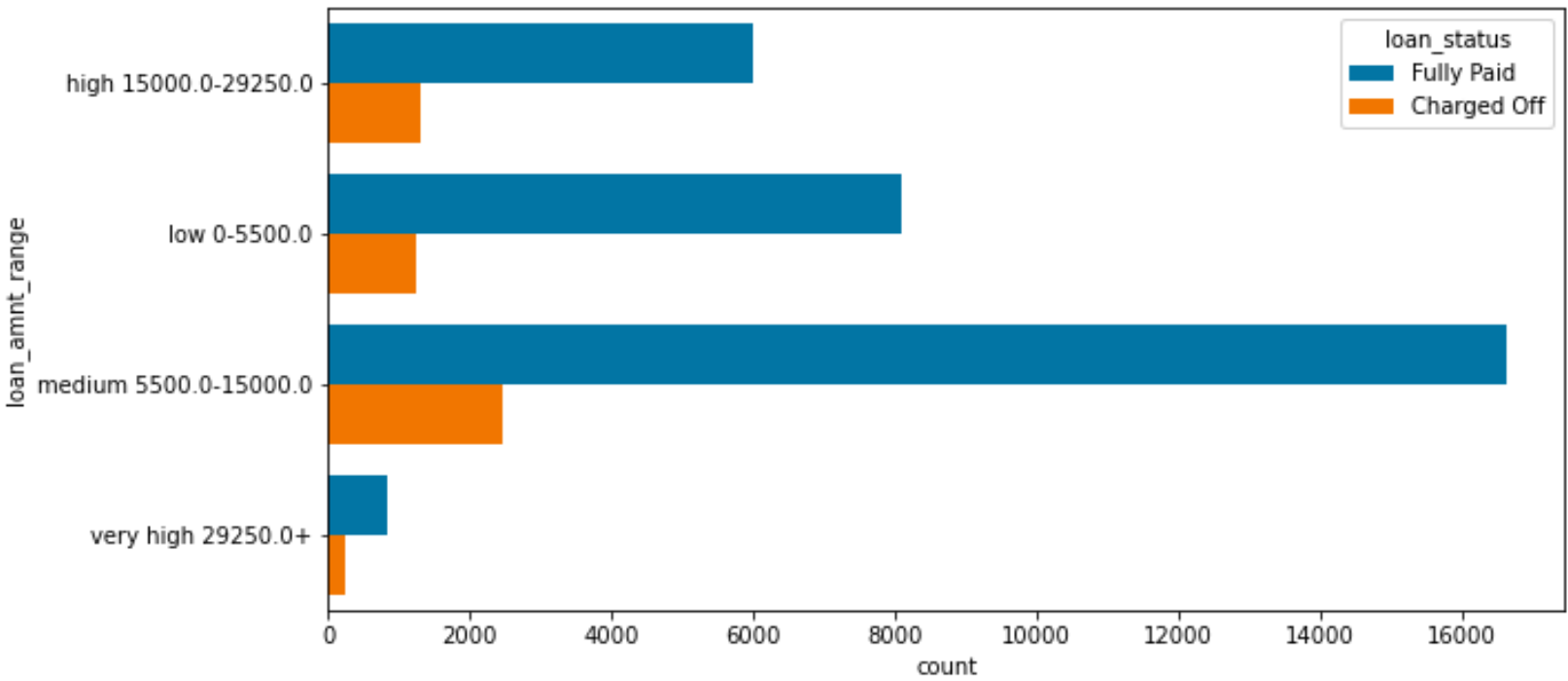


Insights:

- Total defaulted are around 14.2% from the total applicants after removing the current on going loans.
- Issuing of the loan are increasing year by year and also the defaulters.

Univariate Analysis

Loan Amount with respect to Loan Status

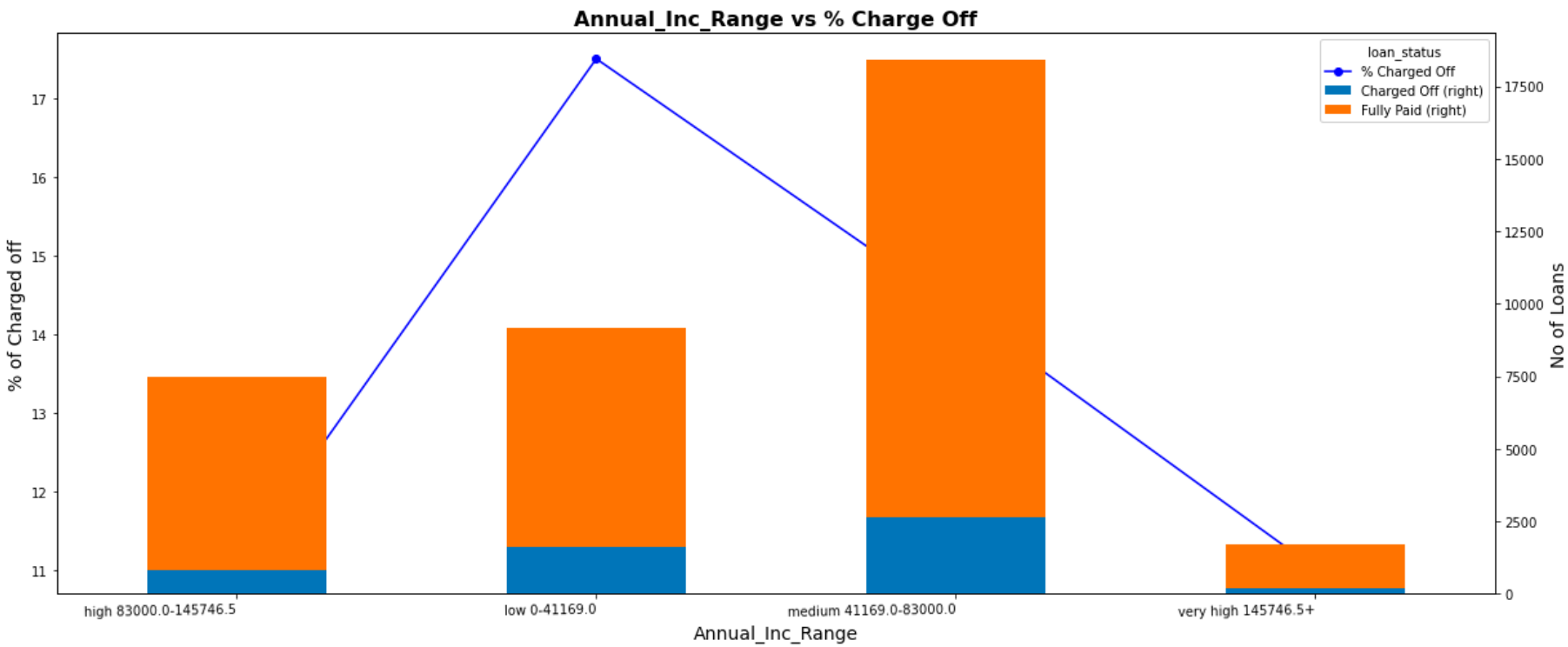
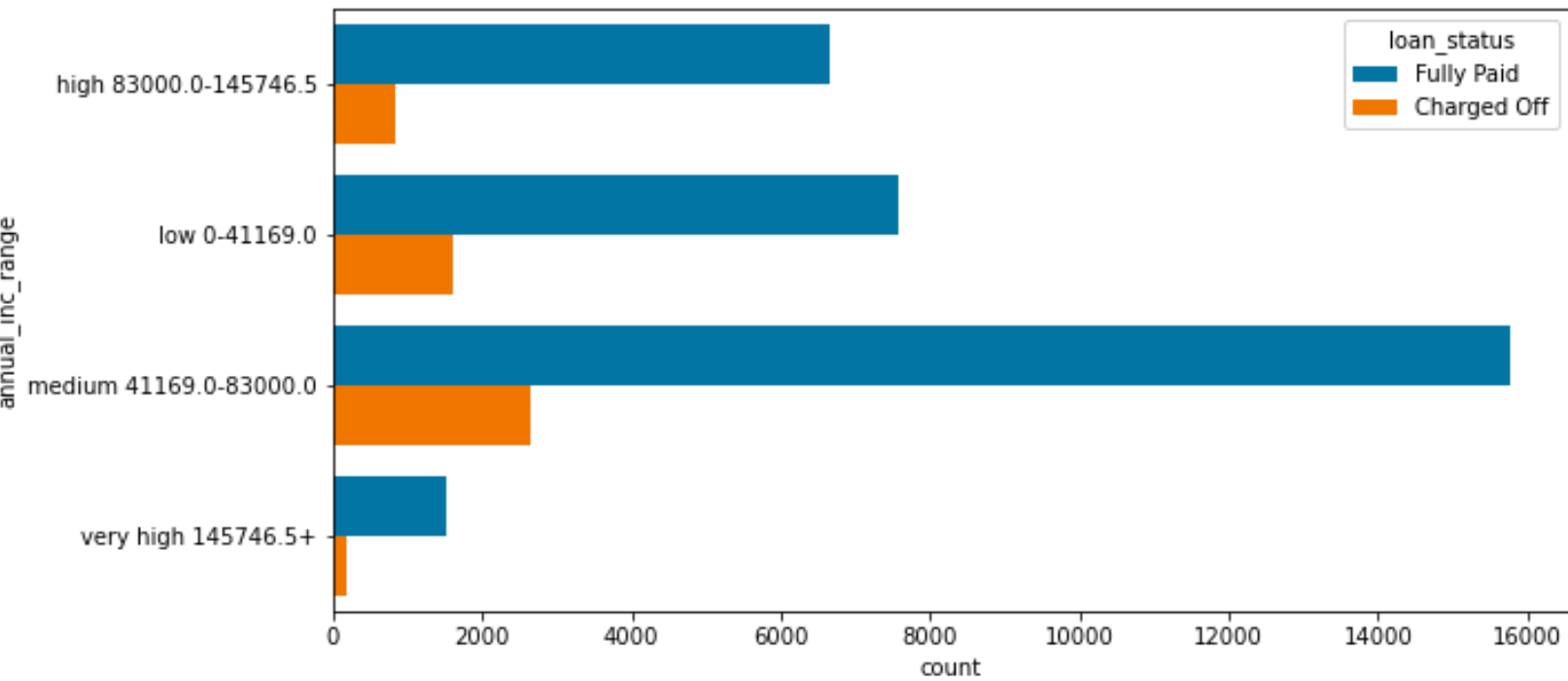


loan_status	Charged Off	Fully Paid	All	% Charged Off
very high 29250.0+	247	831	1078	22.913
high 15000.0-29250.0	1297	5977	7274	17.831
low 0-5500.0	1250	8093	9343	13.379
medium 5500.0-15000.0	2471	16624	19095	12.941

Insights:

- There is a positive correlation between the rate of defaulters and loan amount
- For loan amounts greater than 30000, the rate of defaulters is very high

Annual Income with respect to Loan Status



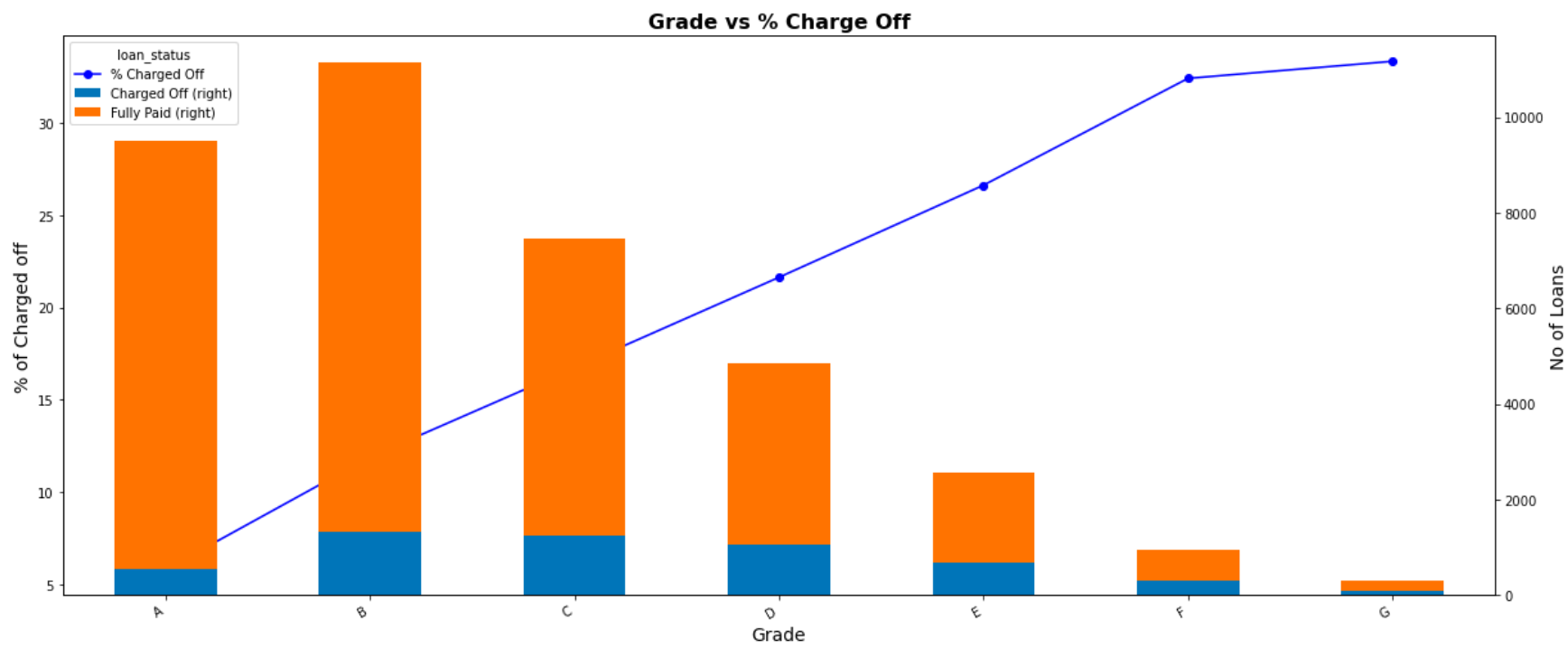
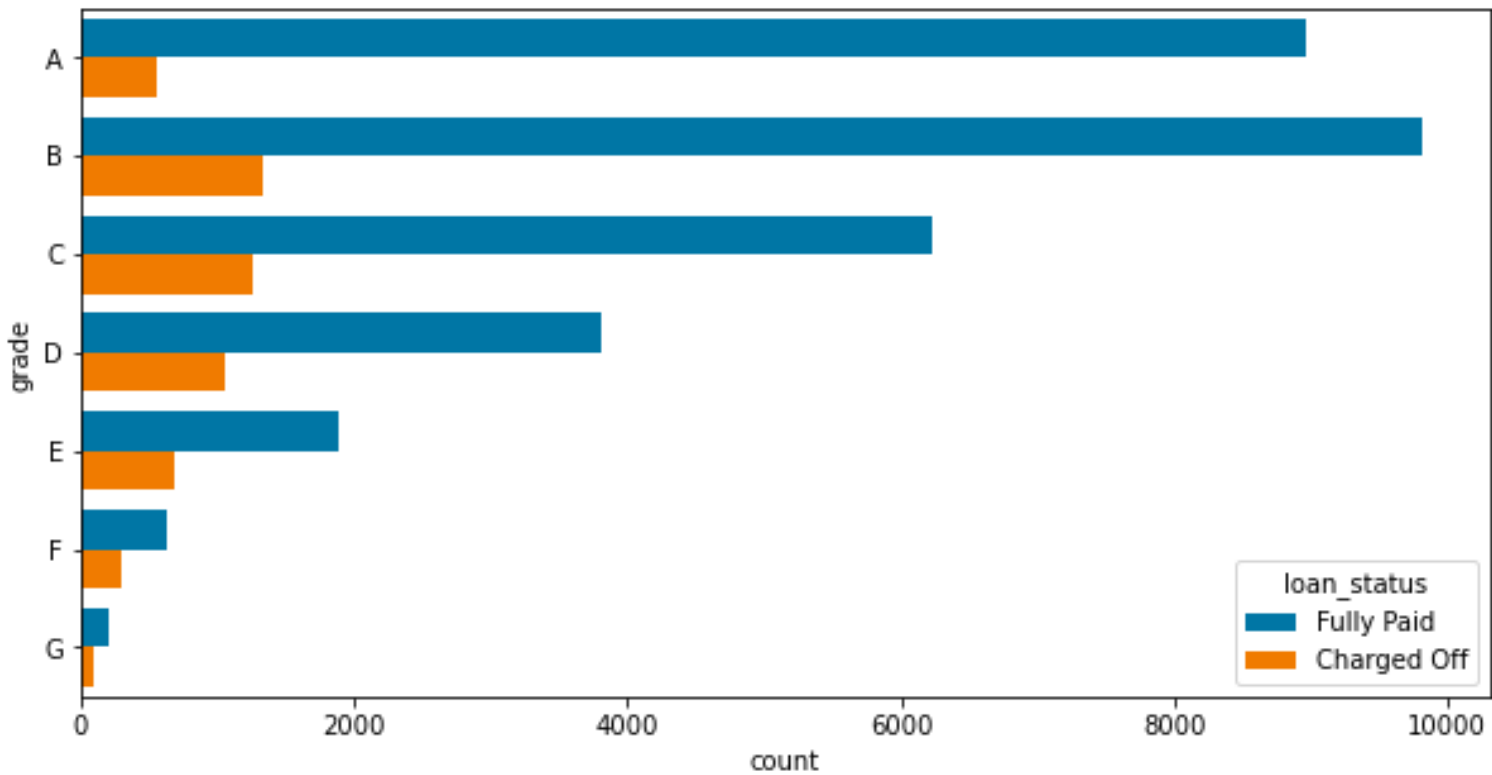
loan_status	Charged Off	Fully Paid	All	% Charged Off
low 0-41169.0	1610	7588	9198	17.504
medium 41169.0-83000.0	2640	15775	18415	14.336
high 83000.0-145746.5	827	6645	7472	11.068
very high 145746.5+	188	1517	1705	11.026

Insights

- Defaulters rate reduces with increase in annual income

Univariate Analysis

Grade with respect to Loan Status

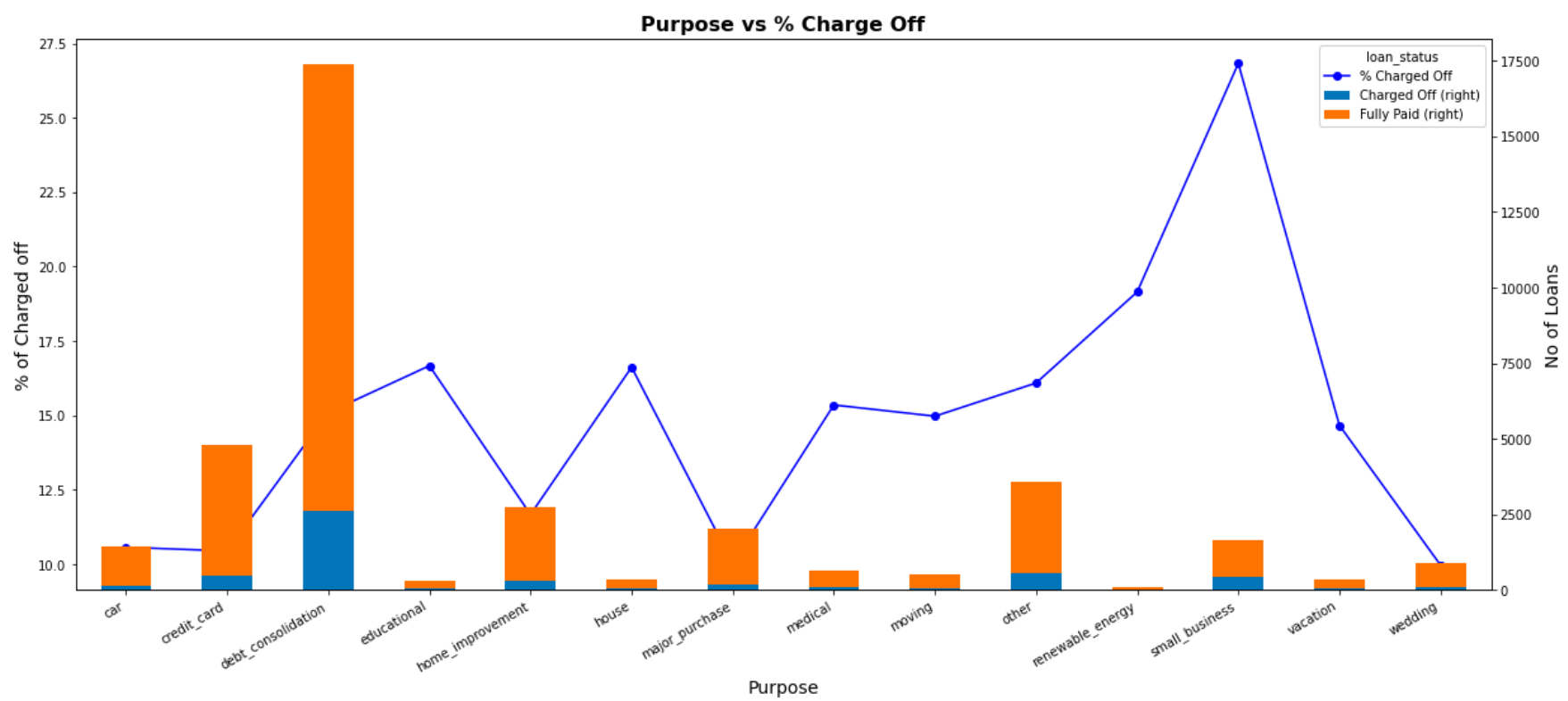
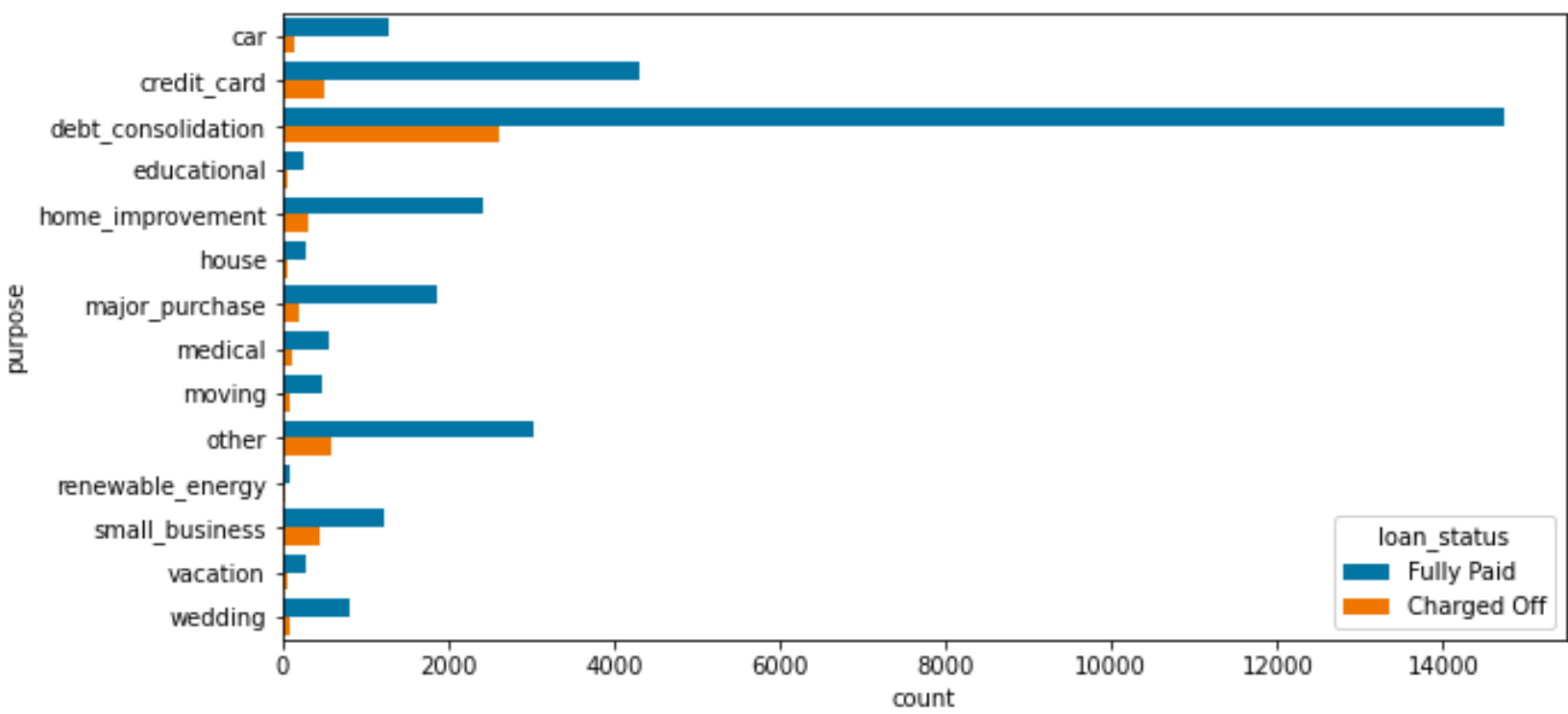


loan_status	Charged Off	Fully Paid	All	% Charged Off
G	98	196	294	33.3
F	304	634	938	32.409
E	685	1888	2573	26.623
D	1049	3800	4849	21.633
C	1251	6222	7473	16.740
B	1326	9820	11146	11.897
A	552	8965	9517	5.800

Insights

- Grade A, B and C are good for the LC as the default percentage is low
- Defaulter percentage increases as the grade goes from A to G

Purpose with respect to Loan Status



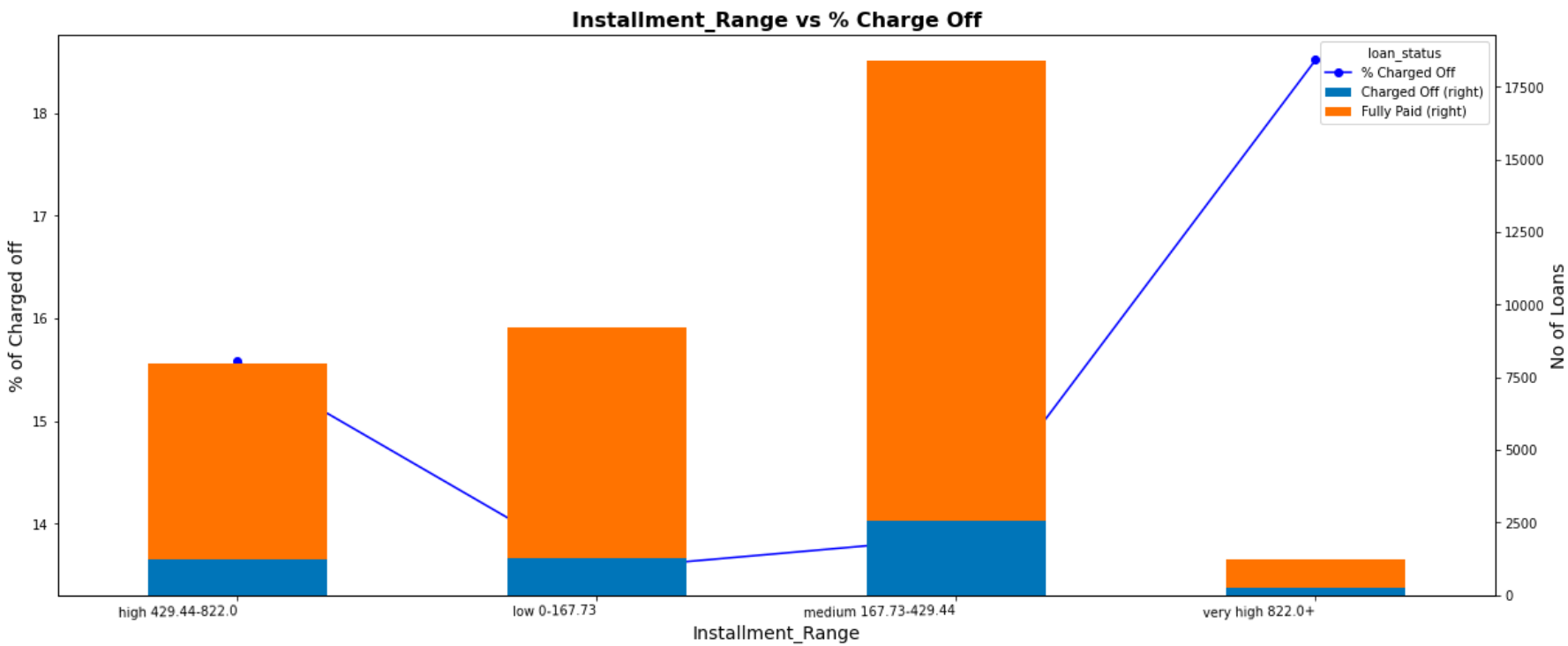
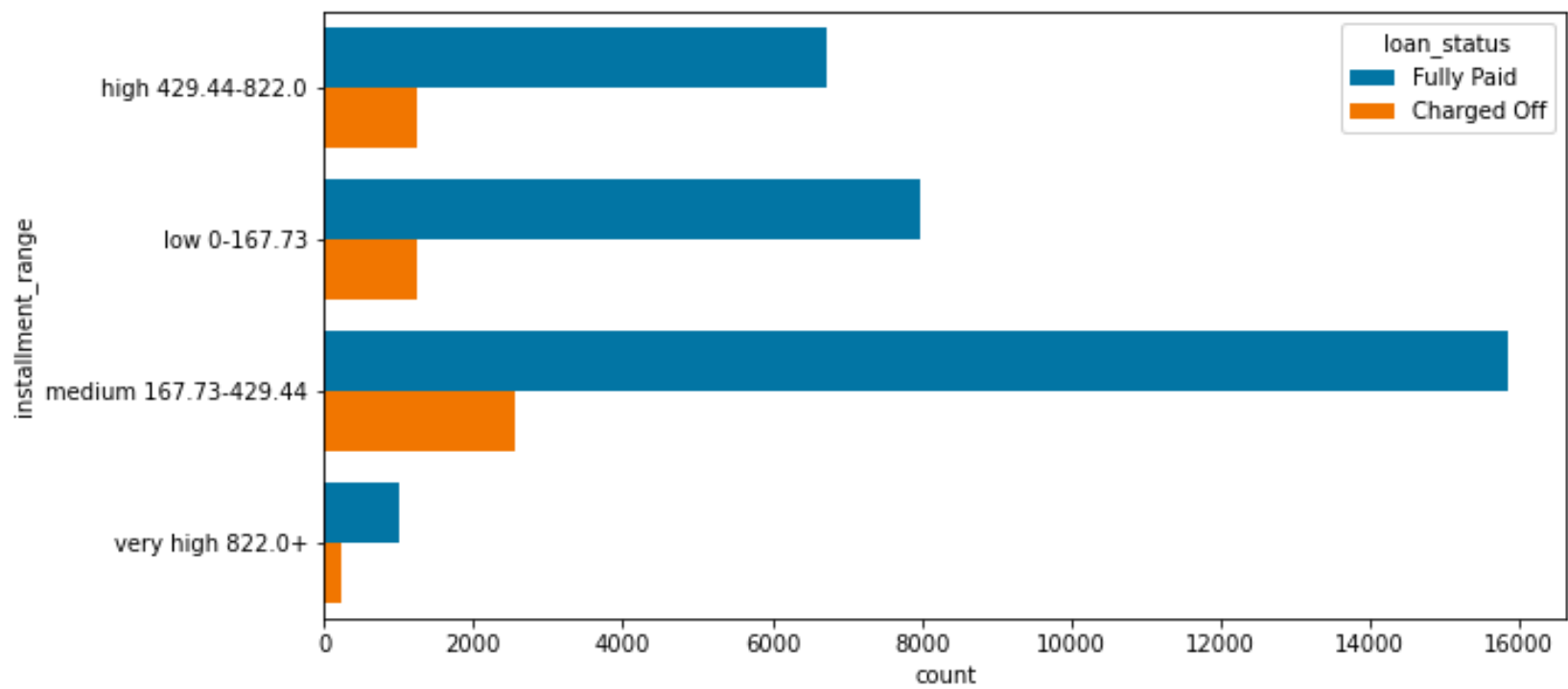
loan_status	Charged Off	Fully Paid	All	% Charged Off
small_business	444	1211	1655	26.828
educational	49	245	294	16.667
house	58	291	349	16.619
other	578	3015	3593	16.087
medical	99	546	645	15.349
debt_consolidation	2622	14748	17370	15.095
moving	81	460	541	14.972
vacation	50	291	341	14.663
home_improvement	319	2415	2734	11.668
credit_card	501	4297	4798	10.442

Insights

- Credit card and wedding related loans are safer investments for LC
- small businesses are a risky investment since a higher %age of loans are defaulted for this category

Univariate Analysis

Installment Range with respect to Loan Status

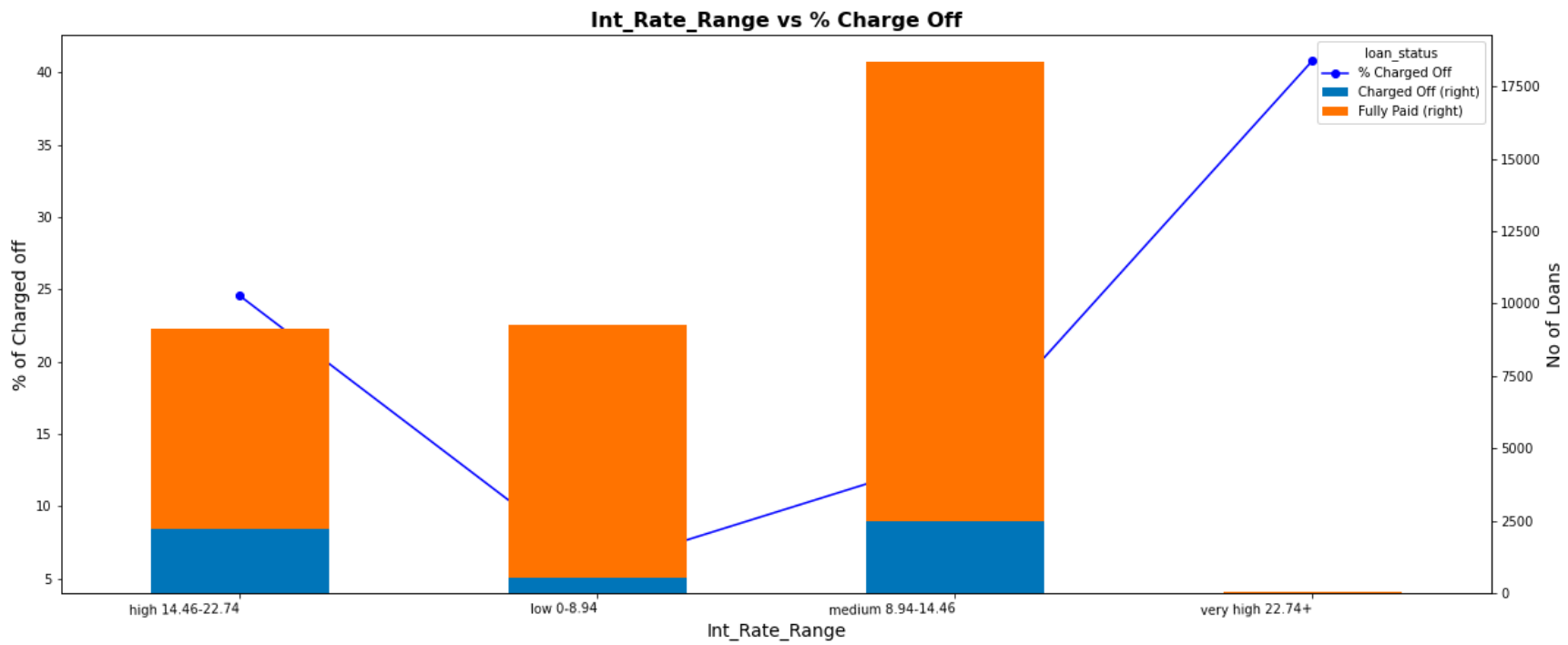
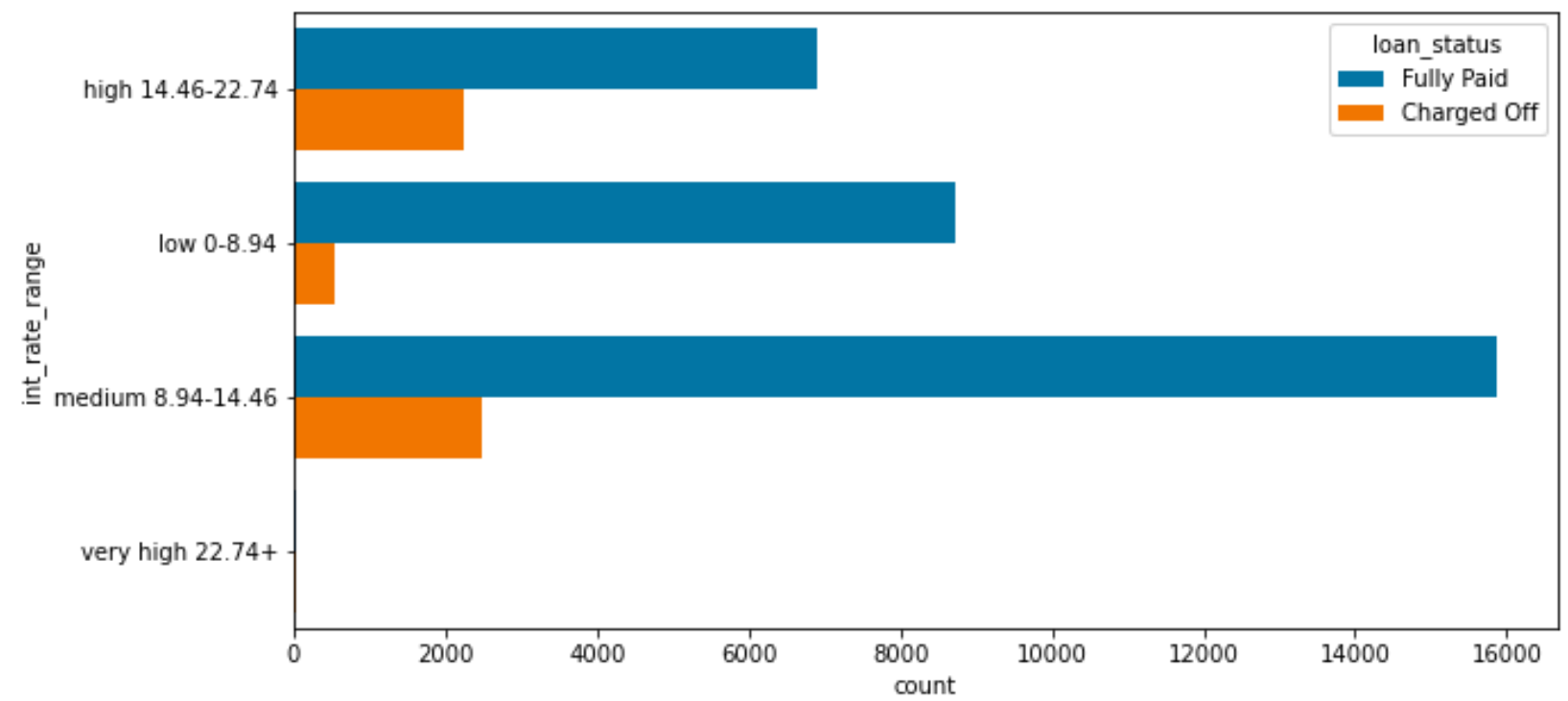


loan_status	Charged Off	Fully Paid	All	% Charged Off
very high 822.0+	227	999	1226	18.515
high 429.44-822.0	1242	6730	7972	15.580
medium 167.73-429.44	2547	15829	18376	13.860
low 0-167.73	1249	7967	9216	13.553

Insights:

- Very high installment amount has a higher chance of default(822+)

Interest Rate with respect to Loan Status



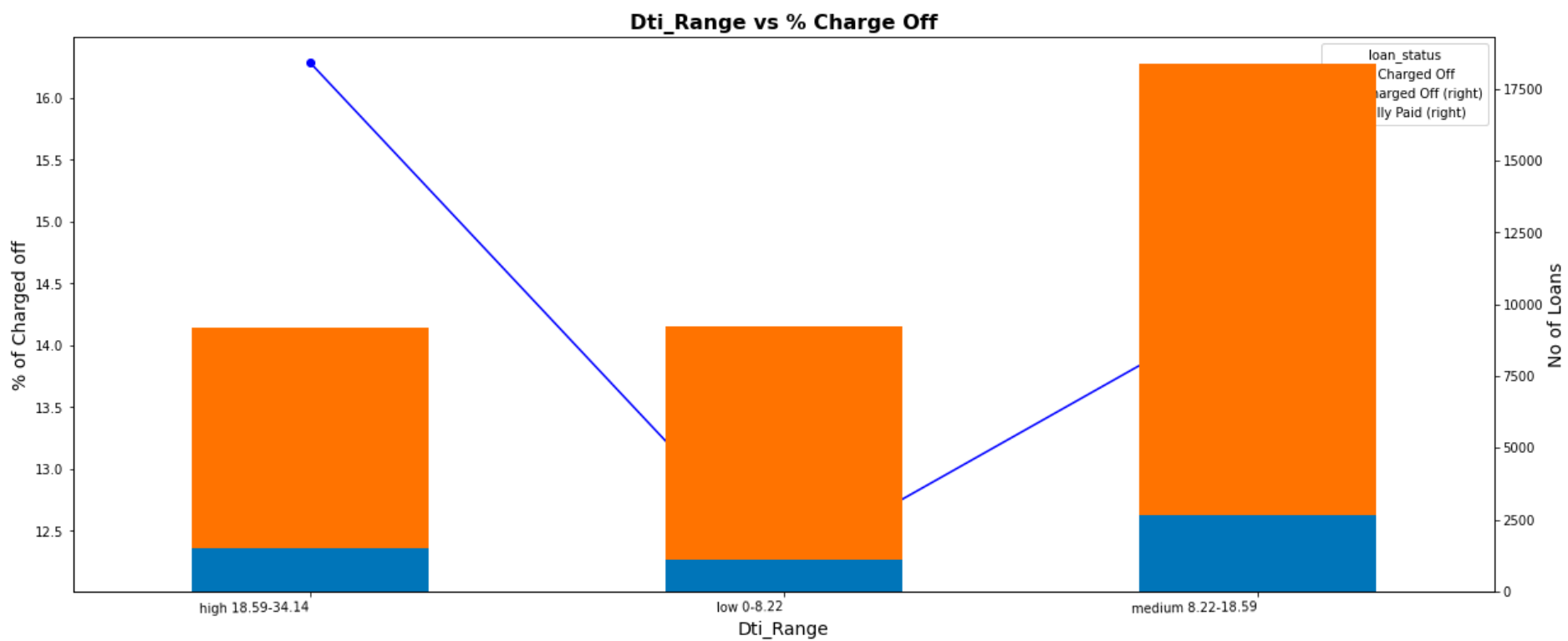
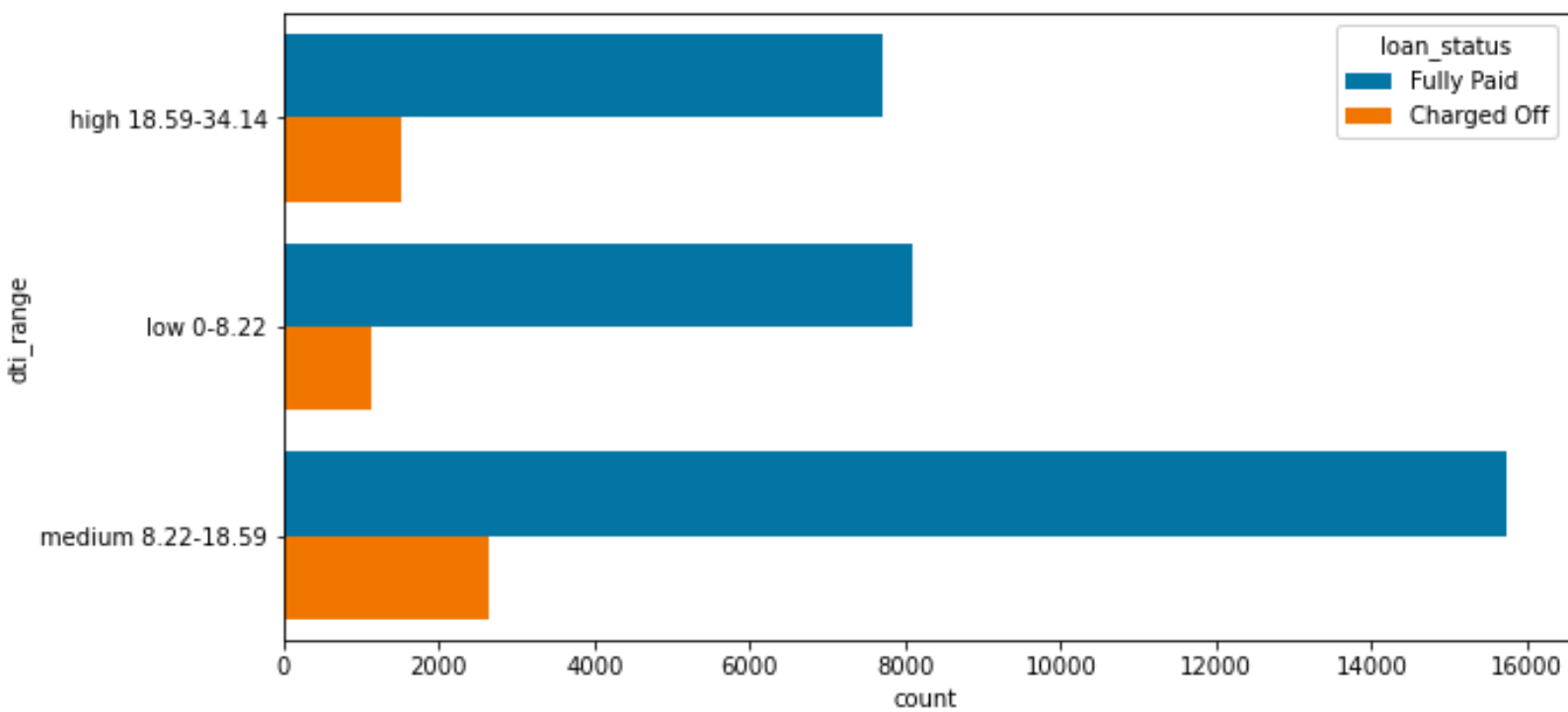
loan_status	Charged Off	Fully Paid	All	% Charged Off
very high 22.74+	20	29	49	40.816
high 14.46-22.74	2241	6890	9131	24.543
medium 8.94-14.46	2471	15882	18353	13.464
low 0-8.94	533	8724	9257	5.758

Insights

- Defaulters rate is higher for high interest rates (14.4% +)

Univariate Analysis

Dti with respect to Loan Status

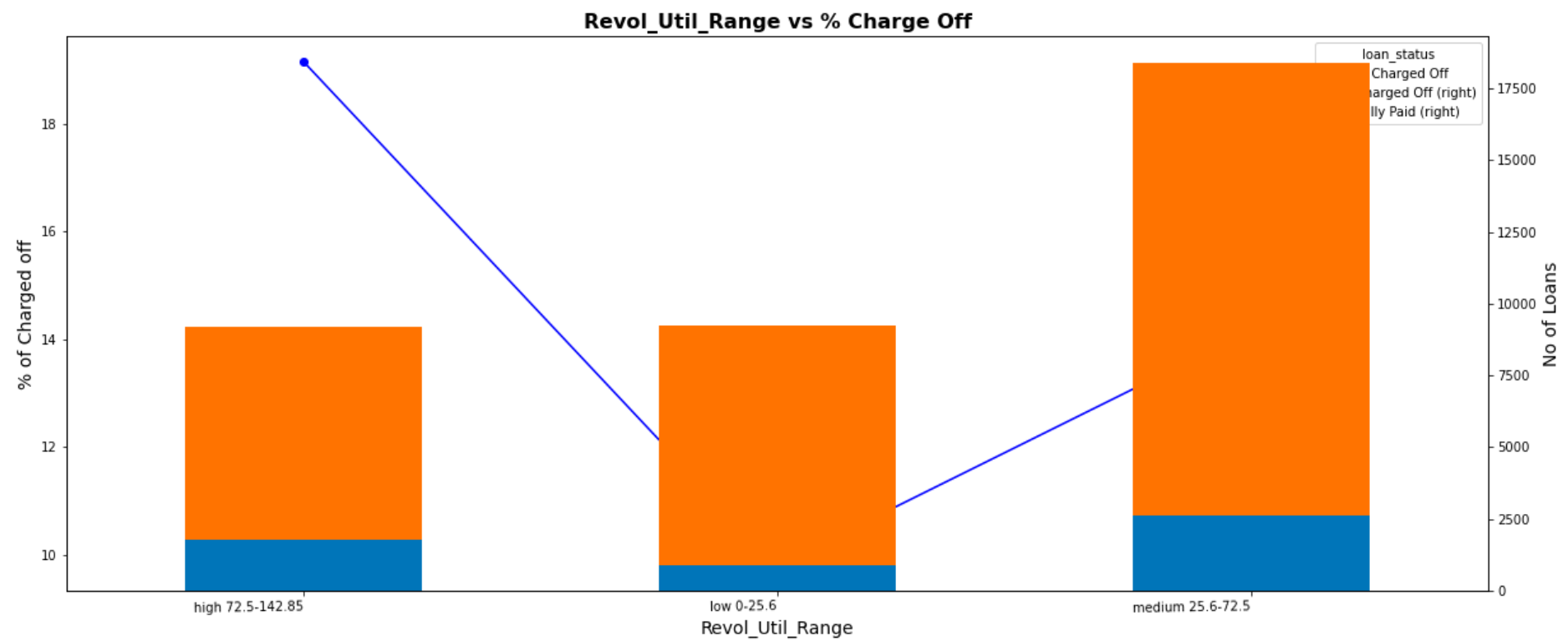
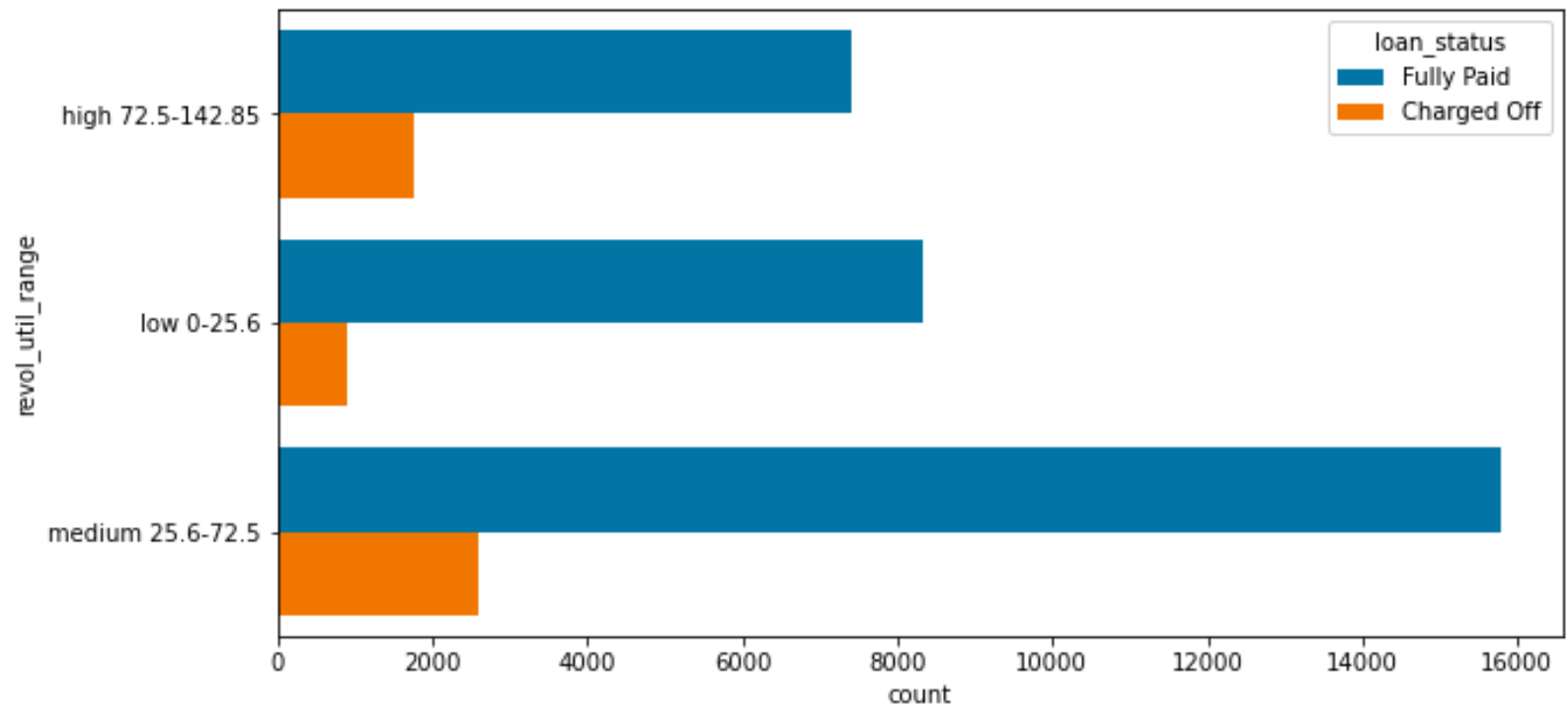


loan_status	Charged Off	Fully Paid	All	% Charged Off
high 18.59-34.14	1497	7693	9190	16.289
medium 8.22-18.59	2643	15746	18389	14.373
low 0-8.22	1125	8086	9211	12.214

Insights

- There seems to be a positive correlation between the dti and defaulters rate

Revolving Utilisation Rate with respect to Loan Status



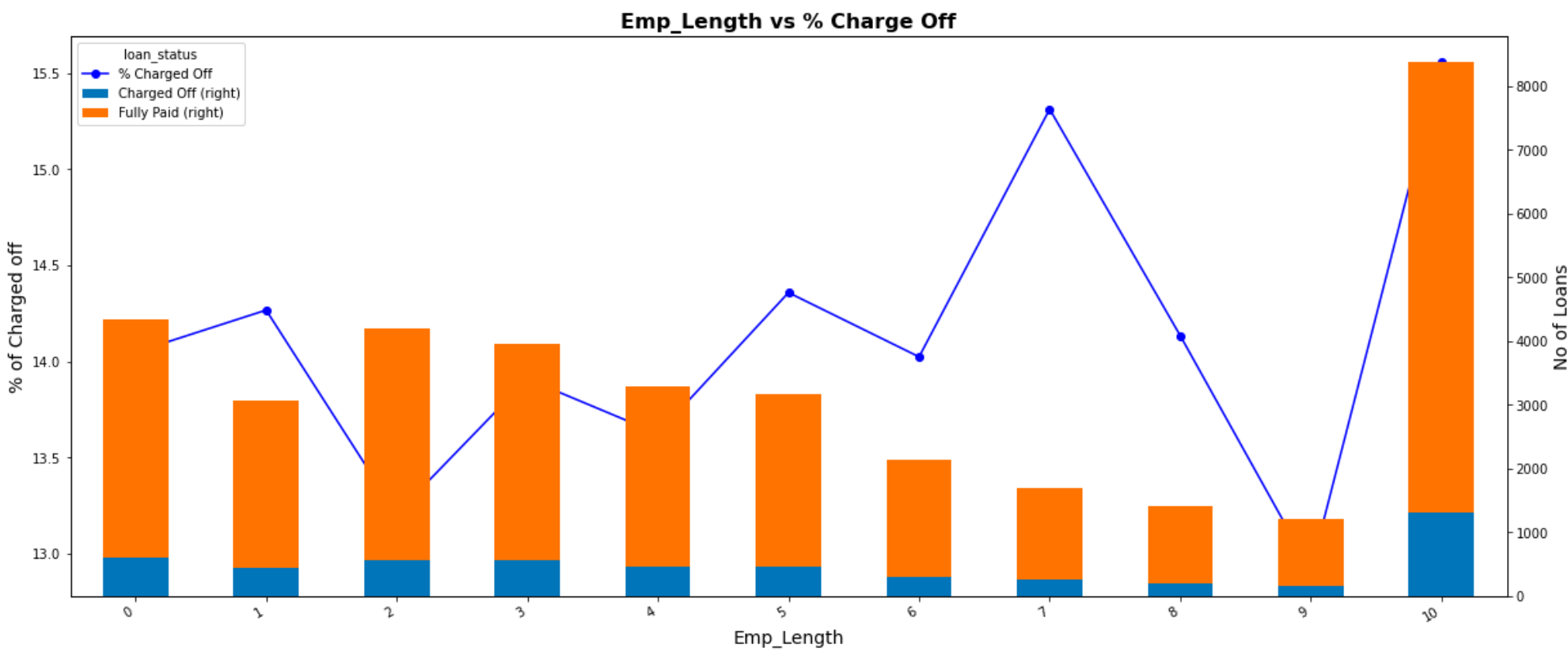
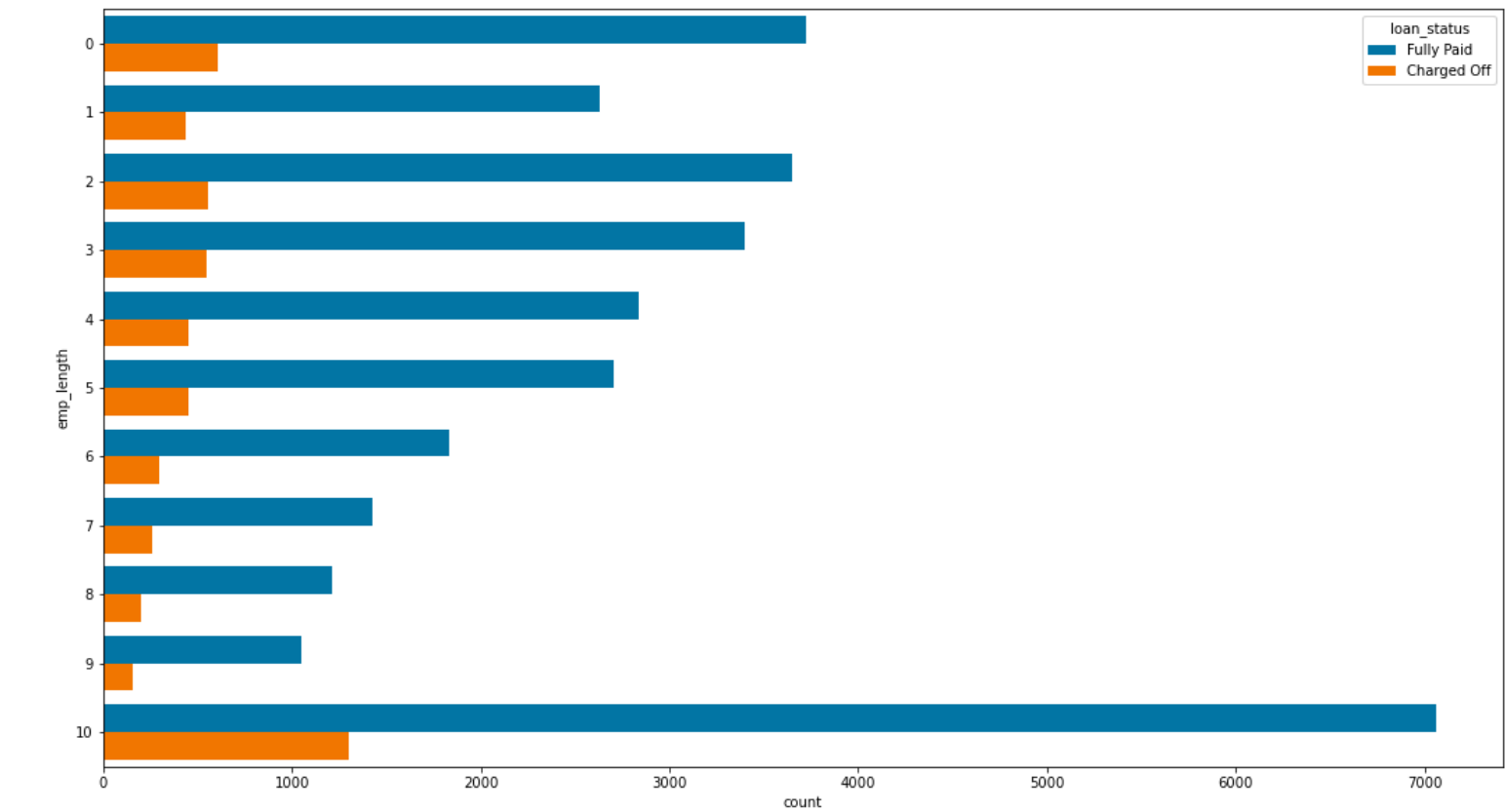
loan_status	Charged Off	Fully Paid	All	% Charged Off
high 72.5-142.85	1758	7422	9180	19.150
medium 25.6-72.5	2604	15787	18391	14.159
low 0-25.6	903	8316	9219	9.795

Insights

- As the revolving utilization increases, we see an increase in default

Univariate Analysis

Employee Experience with respect to Loan Status

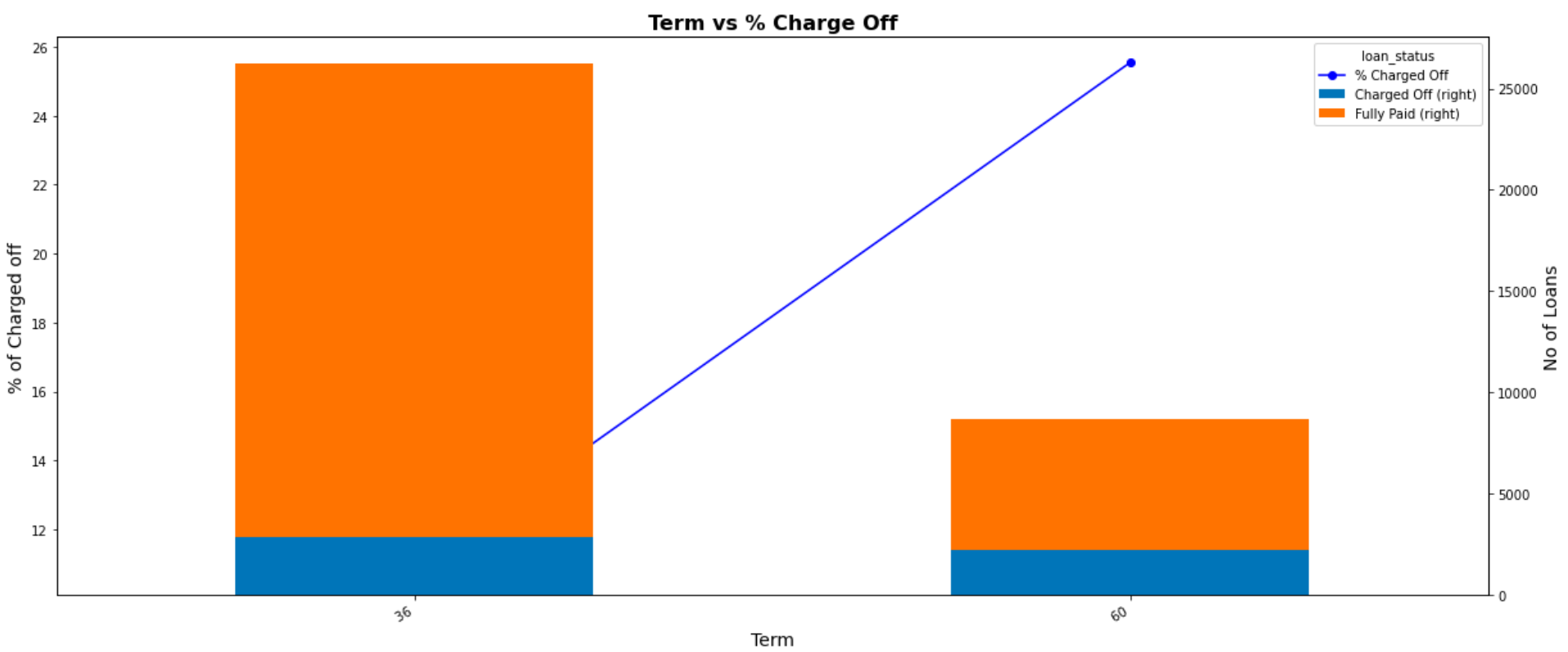
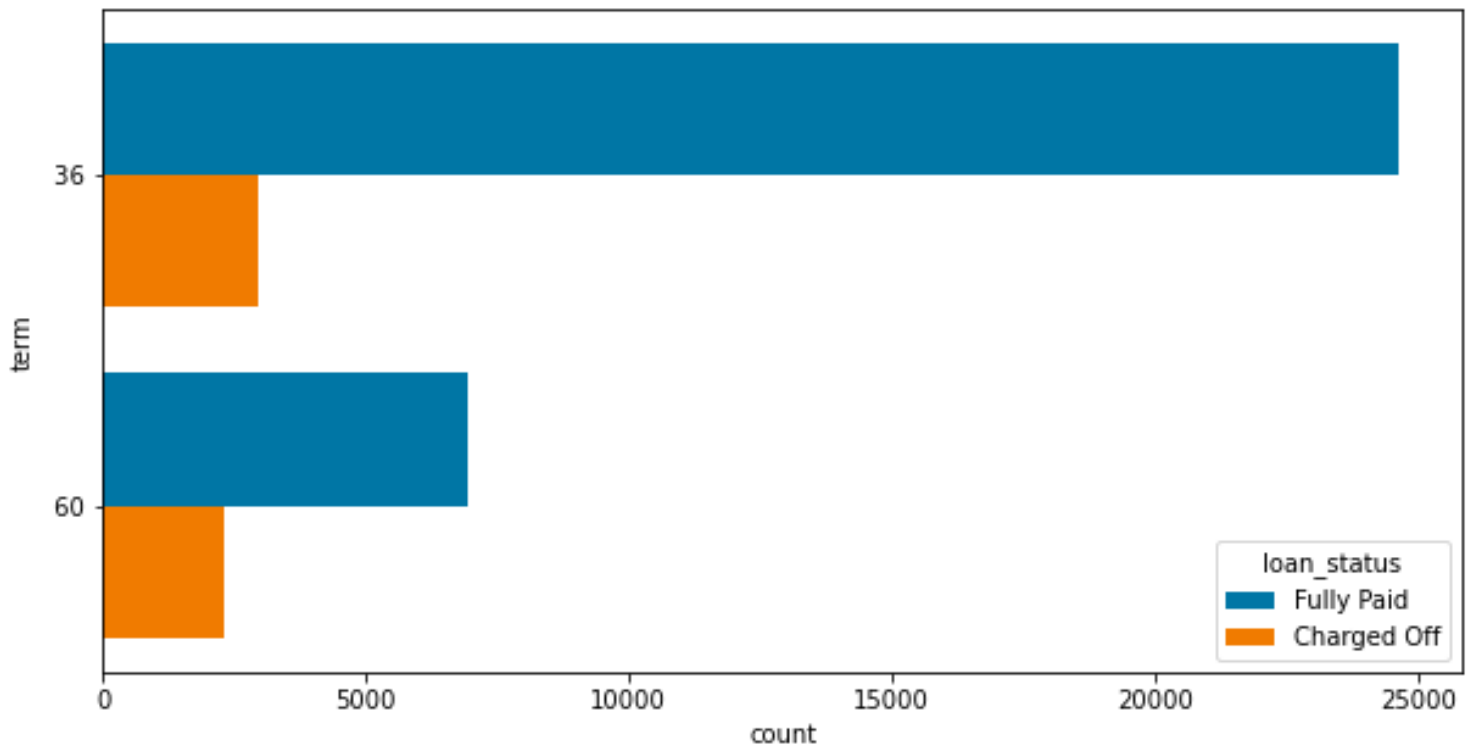


loan_status	Charged Off	Fully Paid	All	% Charged Off
10	1302	7065	8367	15.561
7	258	1427	1685	15.312
5	453	2702	3155	14.358
1	438	2632	3070	14.267
8	199	1209	1408	14.134
0	608	3720	4328	14.048
6	299	1833	2132	14.024
3	549	3398	3947	13.909
4	448	2840	3288	13.625
2	555	3647	4202	13.208
9	156	1052	1208	12.914

Insights

- No exact pattern and correlation

Term with respect to Loan Status



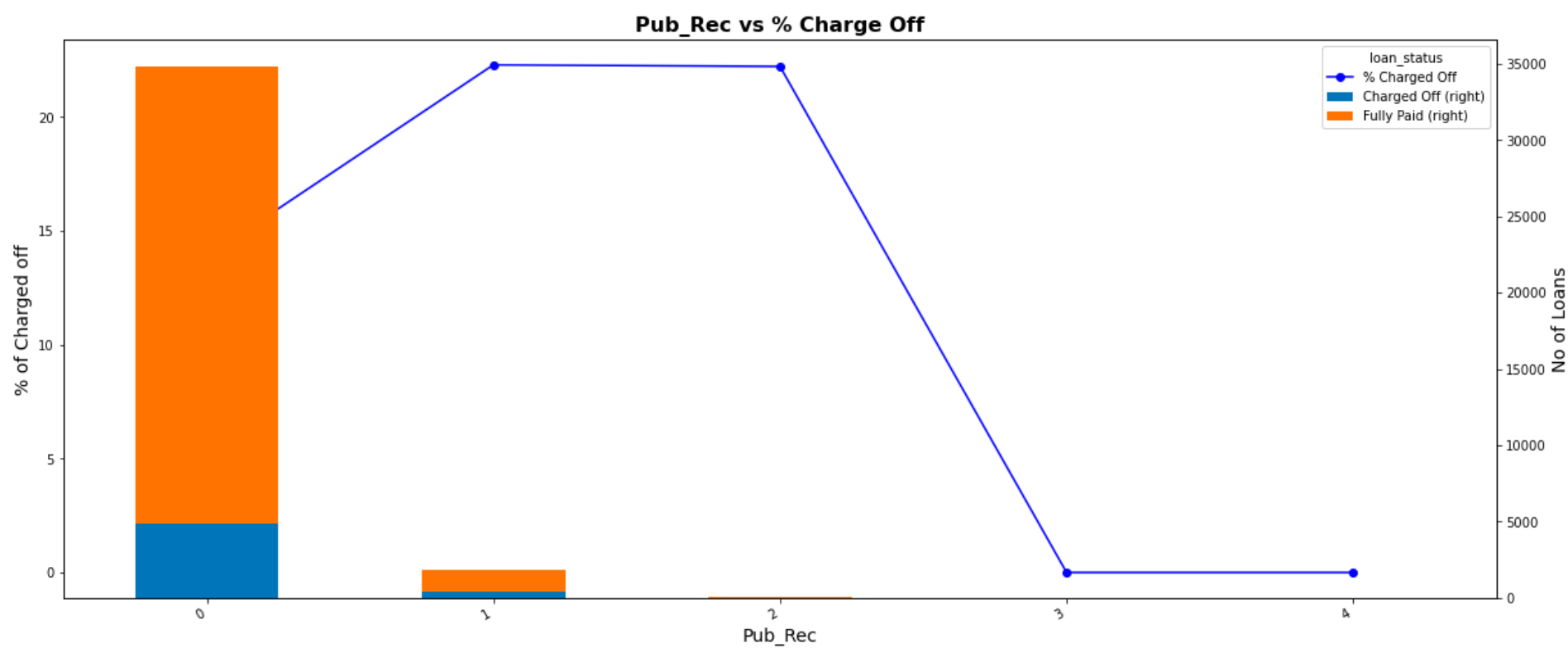
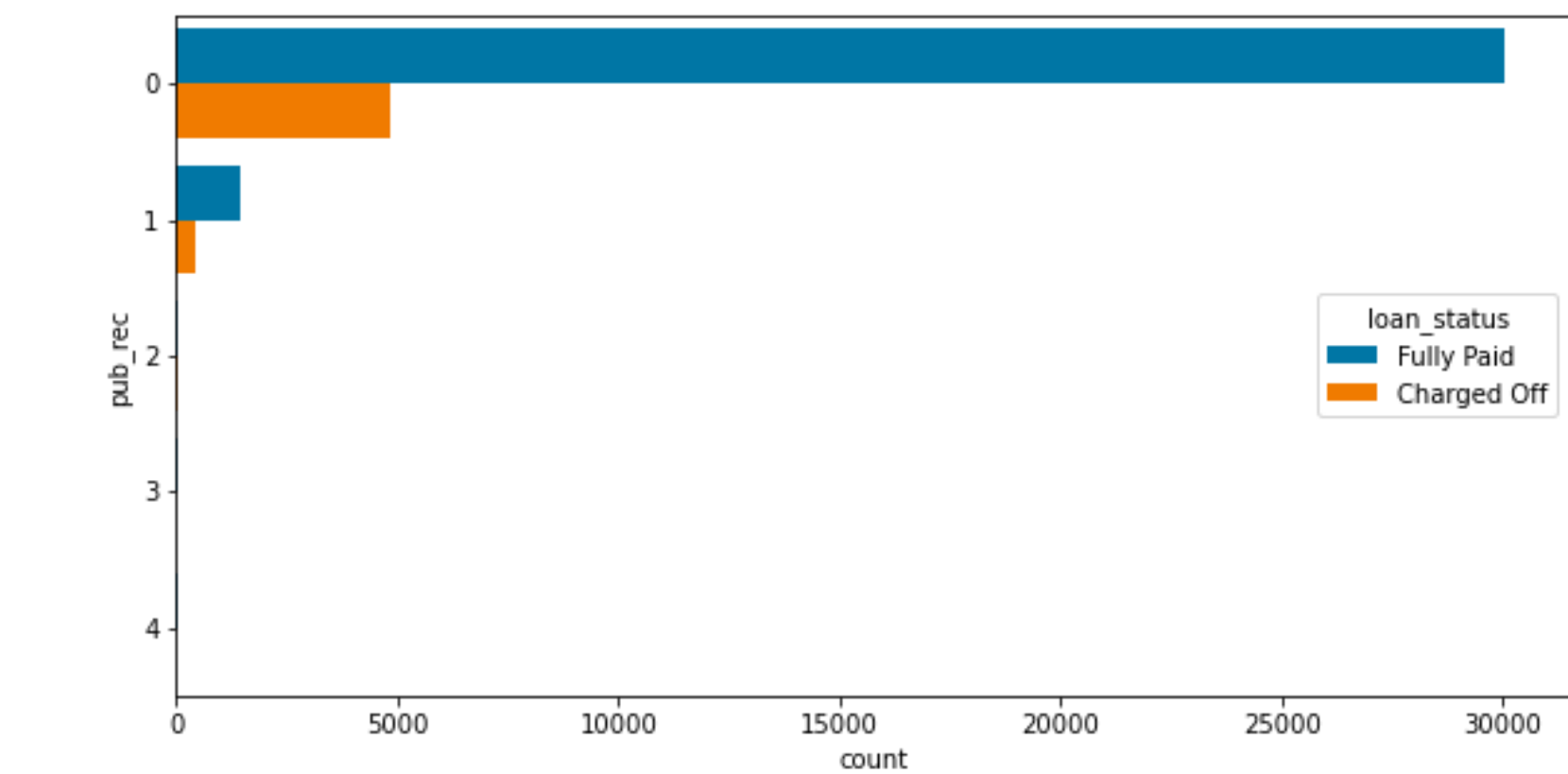
loan_status	Charged Off	Fully Paid	All	% Charged Off
60	2226	6486	8712	25.551
36	2841	23396	26237	10.828

Insights

- There are more defaults for 60 months duration than shorter term
- Majority of the loans are taken for a period of 36 months (74.85%)

Univariate Analysis

Public Derogatory with respect to Loan Status

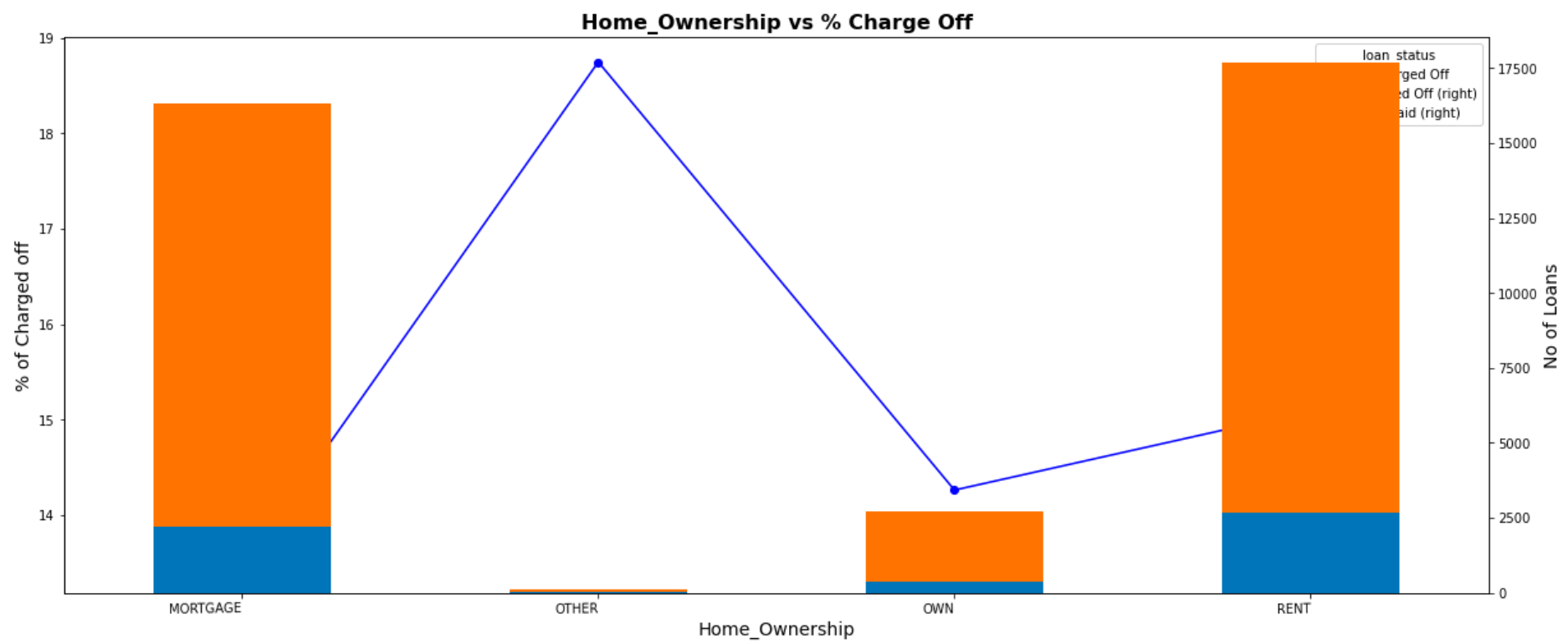
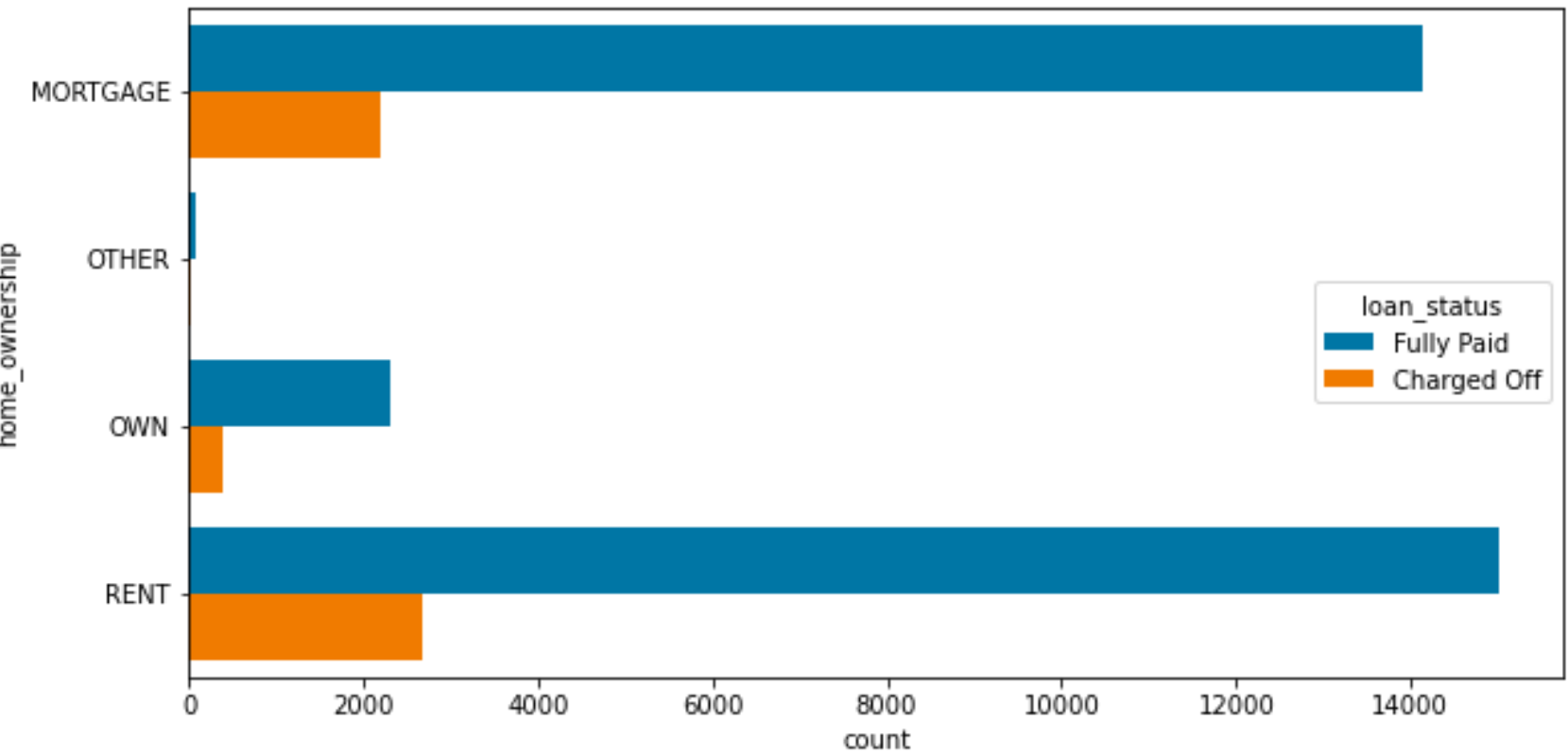


loan_status	Charged Off	Fully Paid	All	% Charged Off
1	418	1457	1875	22.293
2	10	35	45	22.222
0	4837	30024	34861	13.875
3	0	7	7	0.000
4	0	2	2	0.000

Insights

- Looks like defaulters rate increase with increase in public Derogatory with 1 & 2.

Home Ownership with respect to Loan Status



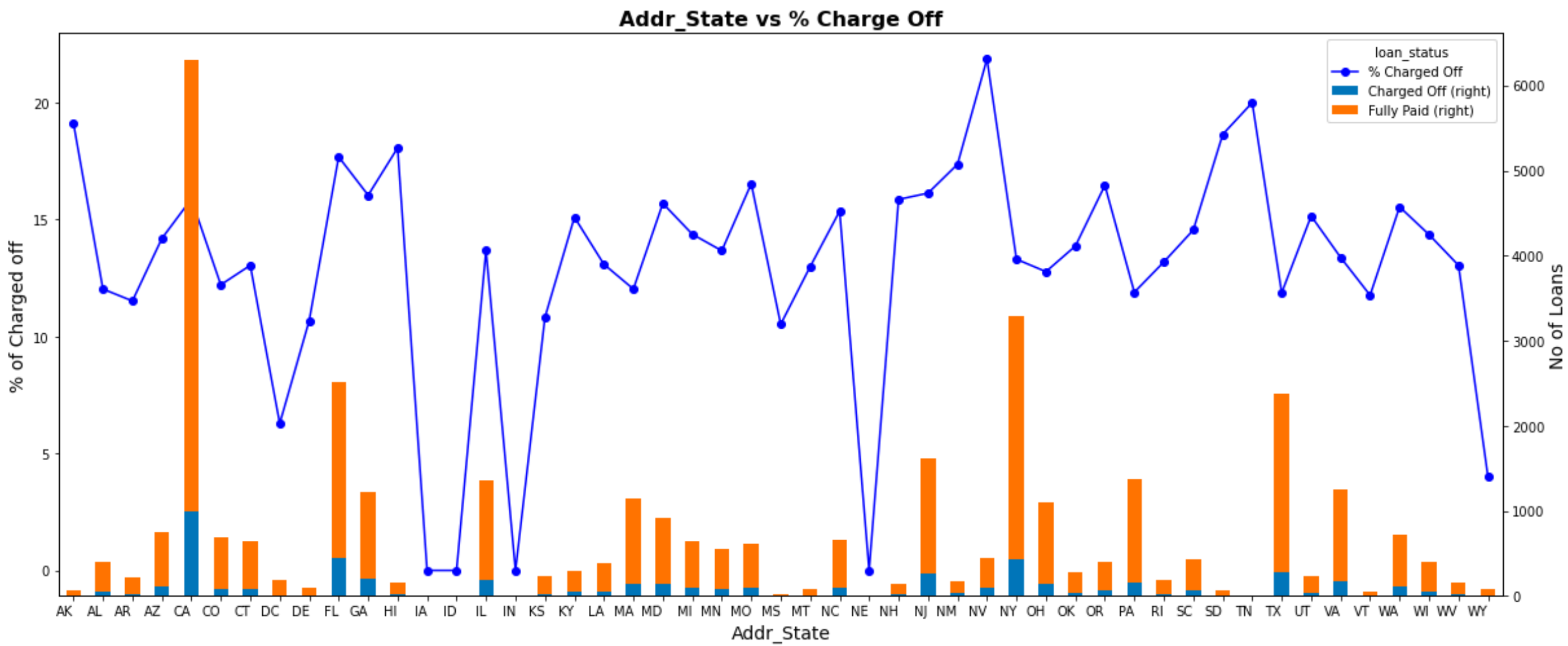
loan_status	Charged Off	Fully Paid	All	% Charged Off
OTHER	18	78	96	18.750
RENT	2666	14997	17663	15.094
OWN	385	2315	2700	14.259
MORTGAGE	2196	14135	16331	13.447

Insights

- No exact pattern and correlation

Univariate Analysis

State with respect to Loan Status



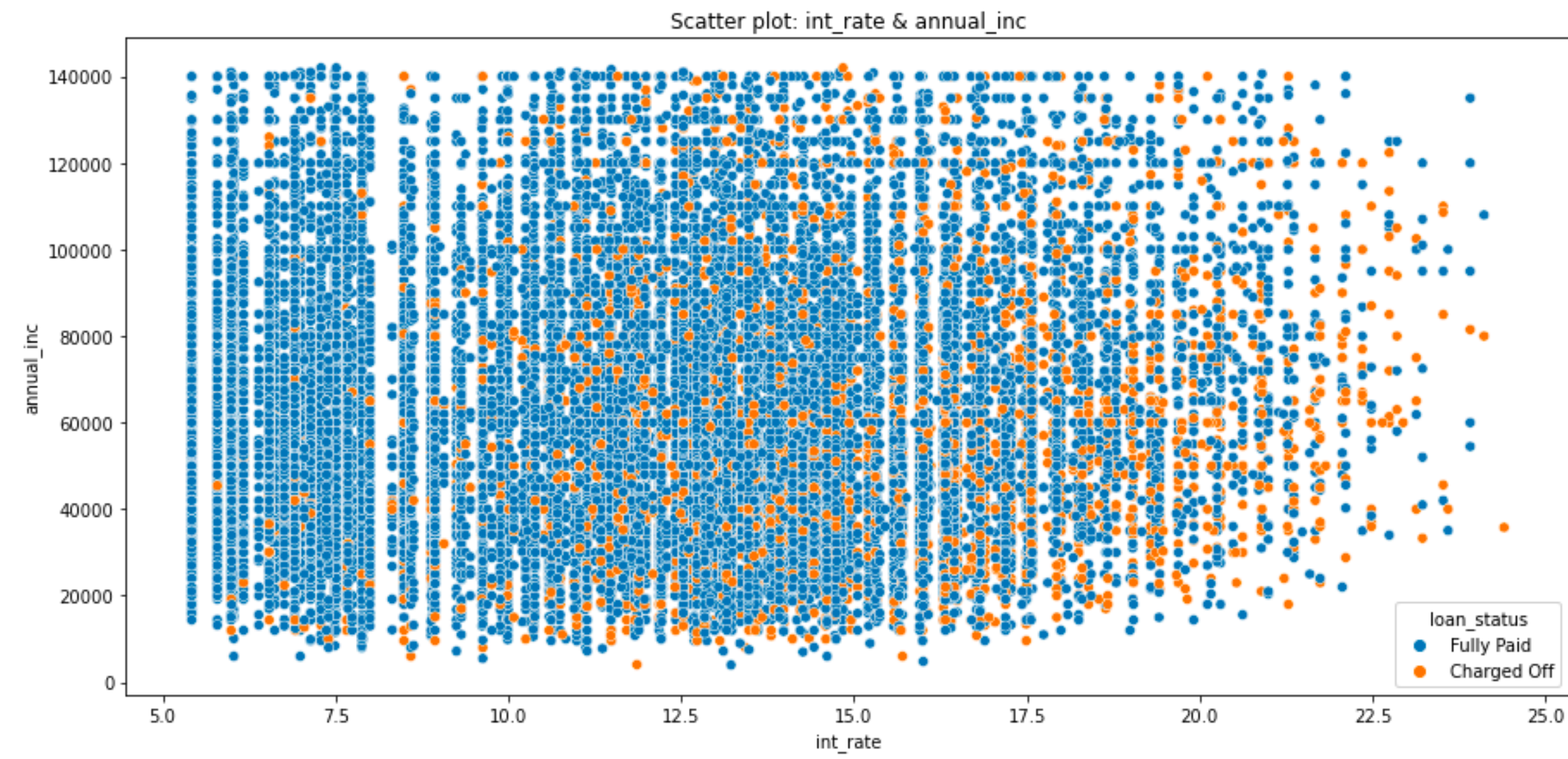
Insights

- Highest number of applicants are from CA state and the highest number of defaulters.
- FL state has more number of defaulter with relative percentage of 17.68 where the financial loss will be more.
- NV state has highest seat of defaulters with relative percentage of 21.86 but the applicants are comparatively less than other states.
- CA, FL, NJ, GA, MD, WA, MO are states with more than 15% relative defaulters where the loss will be more comparatively.

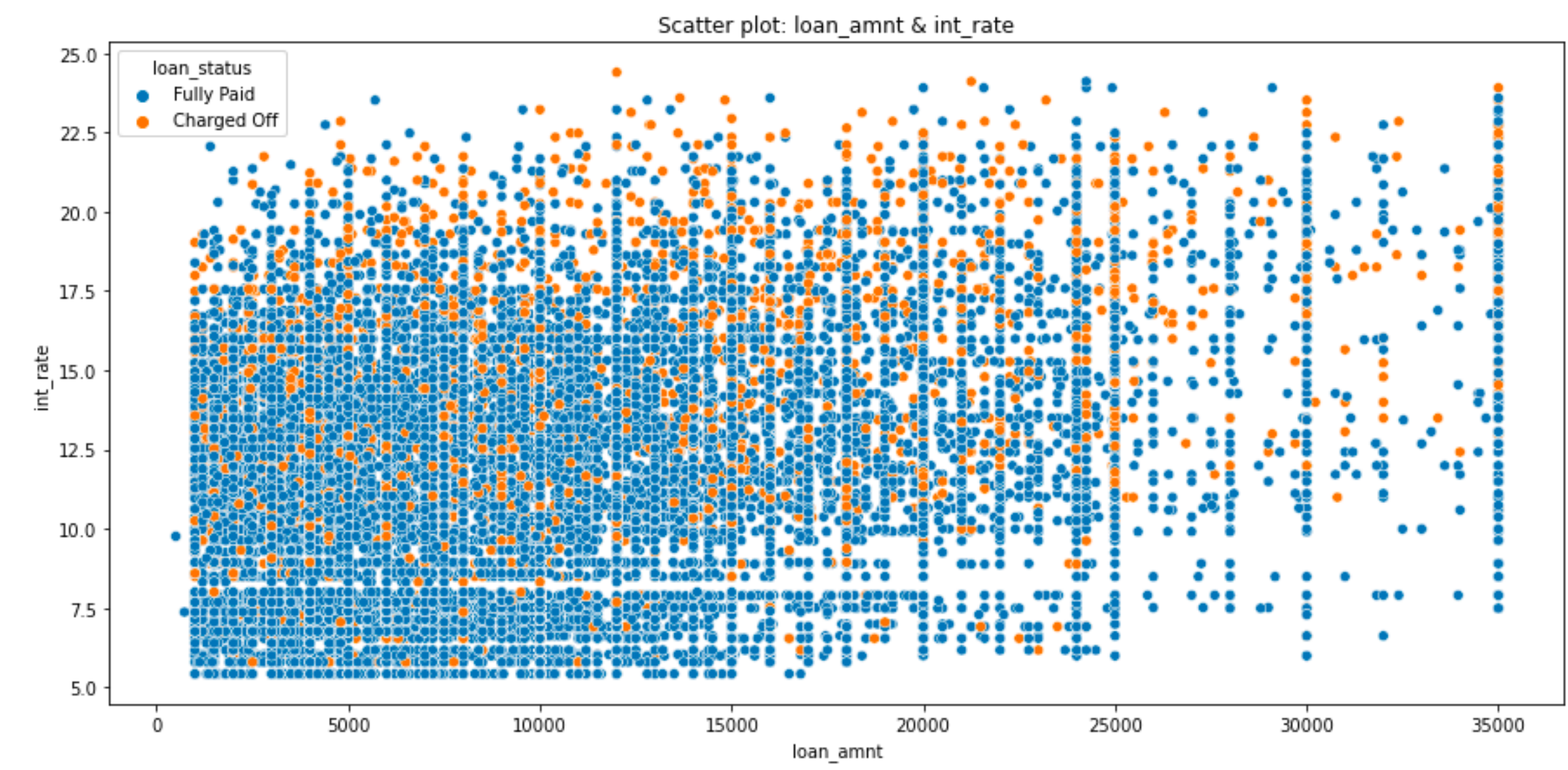
loan_status	Charged Off	Fully Paid	All	% Charged Off
CA	1000	5297	6297	15.881
FL	445	2072	2517	17.680
NY	438	2854	3292	13.305
TX	282	2094	2376	11.869
NJ	261	1357	1618	16.131
GA	197	1031	1228	16.042
IL	186	1173	1359	13.687
VA	167	1081	1248	13.381
PA	164	1215	1379	11.893
MD	145	779	924	15.693
OH	140	956	1096	12.774
MA	138	1009	1147	12.031
WA	113	614	727	15.543
AZ	107	647	754	14.191
MO	102	515	617	16.532
NC	102	562	664	15.361
NV	96	343	439	21.868
MI	92	548	640	14.375
CO	85	611	696	12.213
CT	84	560	644	13.043
MN	76	480	556	13.669
OR	67	340	407	16.462
SC	62	364	426	14.554
WI	57	340	397	14.358
LA	50	332	382	13.089
AL	48	351	399	12.030
KY	45	253	298	15.101
OK	38	236	274	13.869
UT	35	196	231	15.152
NM	29	138	167	17.365
HI	28	127	155	18.065
AR	25	192	217	11.521
KS	25	206	231	10.823
RI	24	158	182	13.187
NH	23	122	145	15.862
WV	21	140	161	13.043
AK	13	55	68	19.118
DC	12	179	191	6.283
SD	11	48	59	18.644
DE	11	92	103	10.680
MT	10	67	77	12.987
VT	6	45	51	11.765
WY	3	72	75	4.000
TN	2	8	10	20.000
MS	2	17	19	10.526
addr_state				

Bivariate Analysis

Interest Rate vs Annual income



Loan Amount vs Interest Rate



Insights

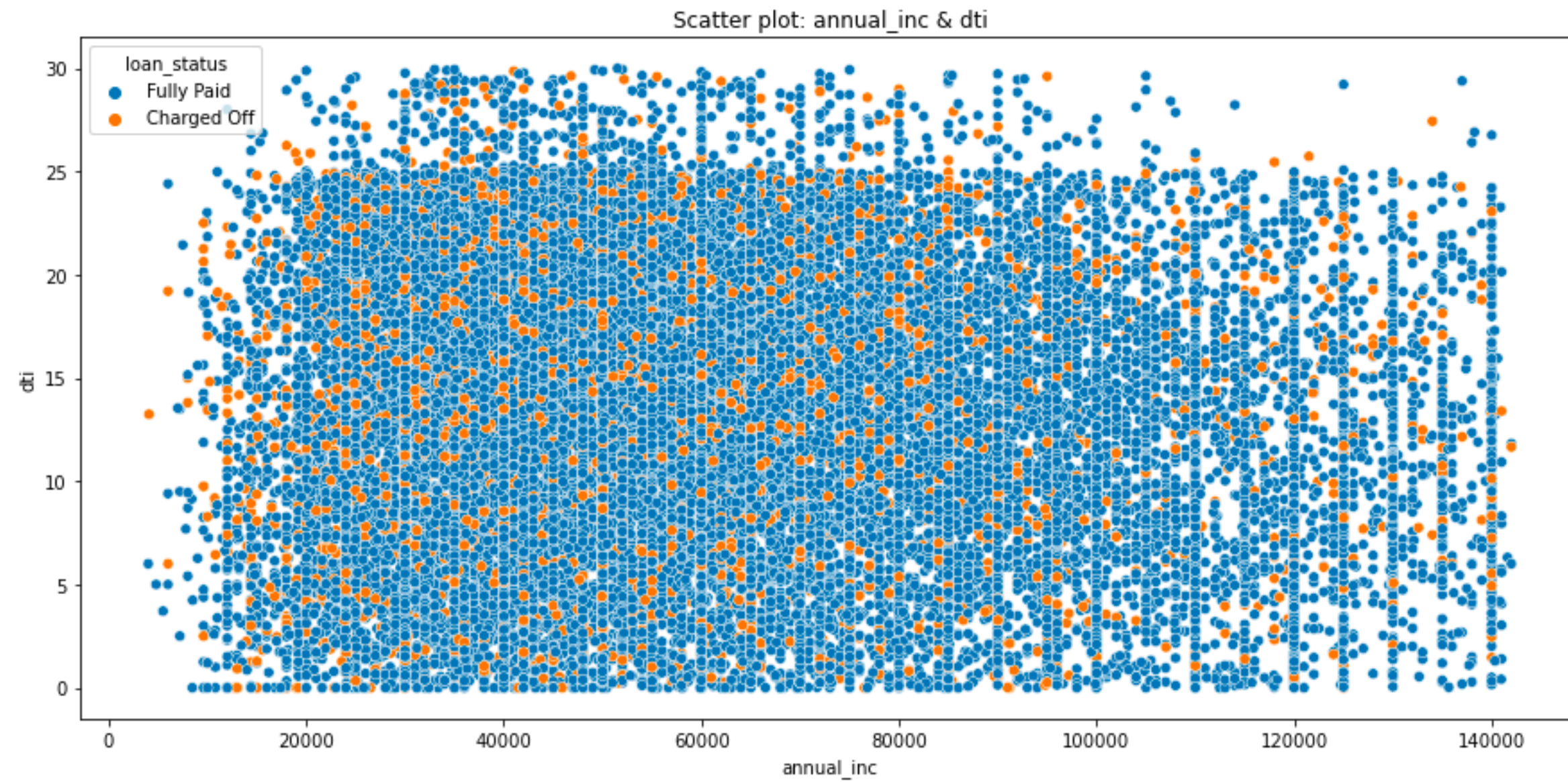
- Higher the interest rate higher the defaulters irrespective of the annual income.

Insights

- higher the interest rate higher charged off ratio
- Another interesting fact is higher interest rates for less loan amount results in more number of defaulters.
- Reducing the loan amount will also help company to reduce the financial loss from defaulters.

Bivariate Analysis

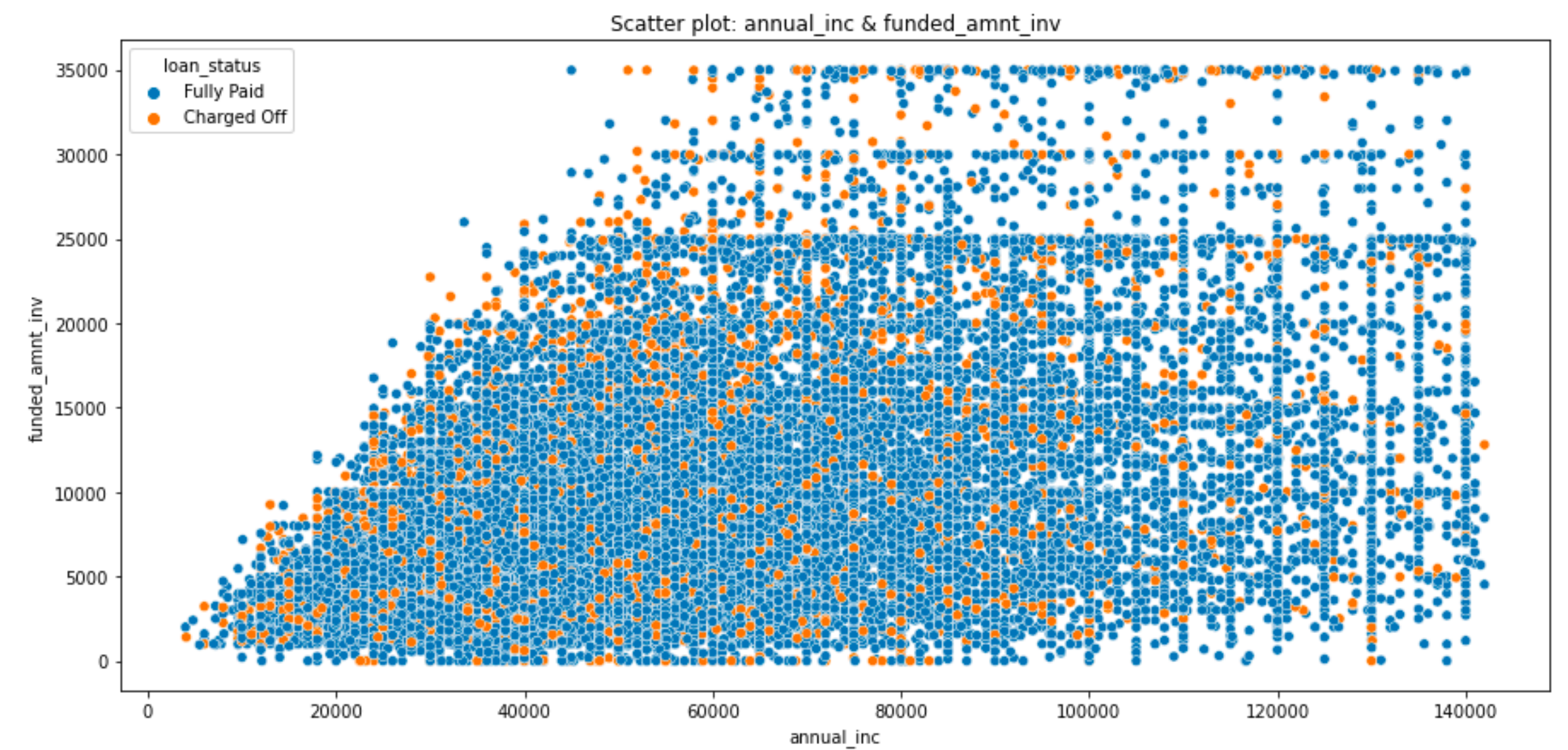
Annual income vs Dti



Insights

- Higher Dti ratio have more defaulters across all annual income ranges.
- There is a positive correlation between dpi and defaulters.

Annual income vs Funded Amount

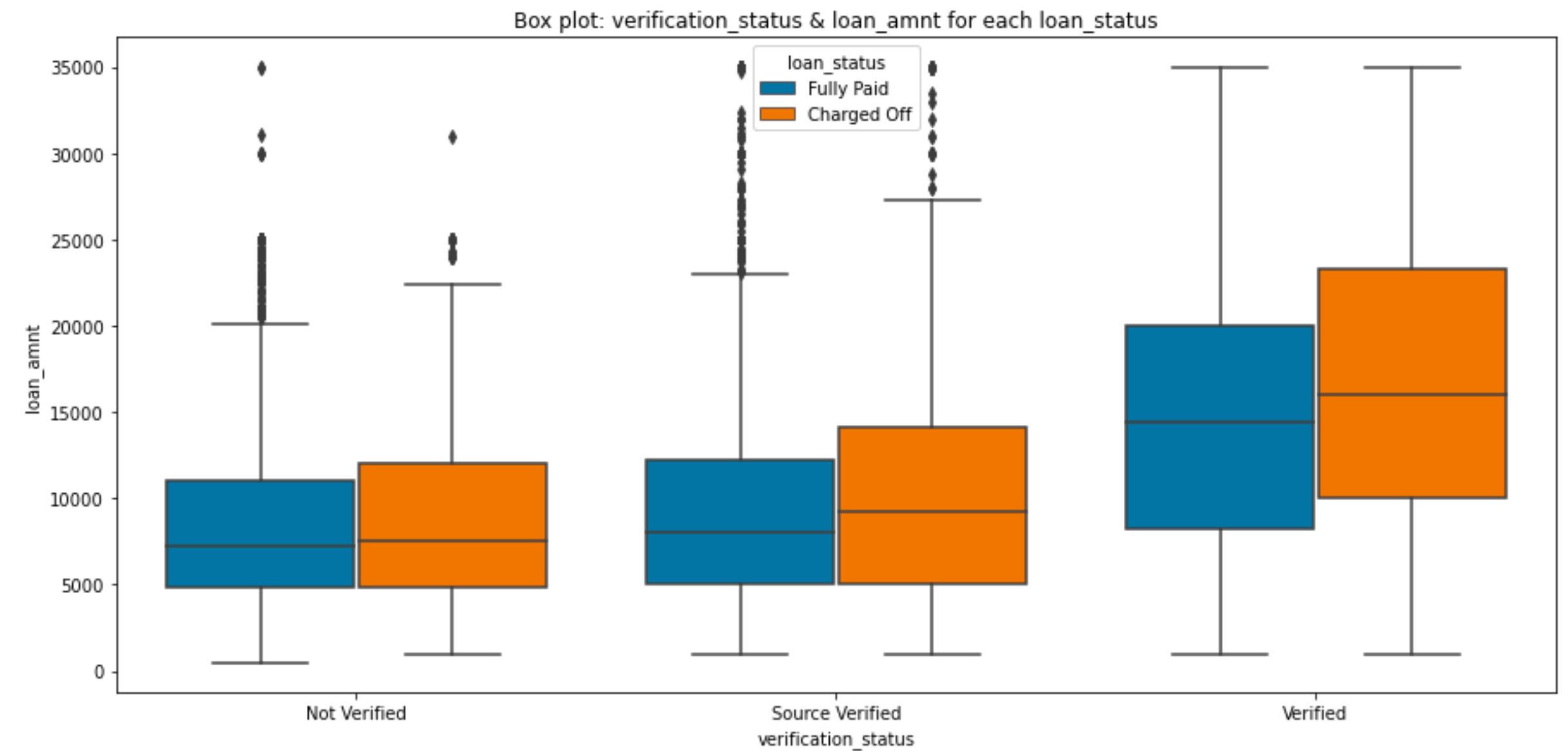
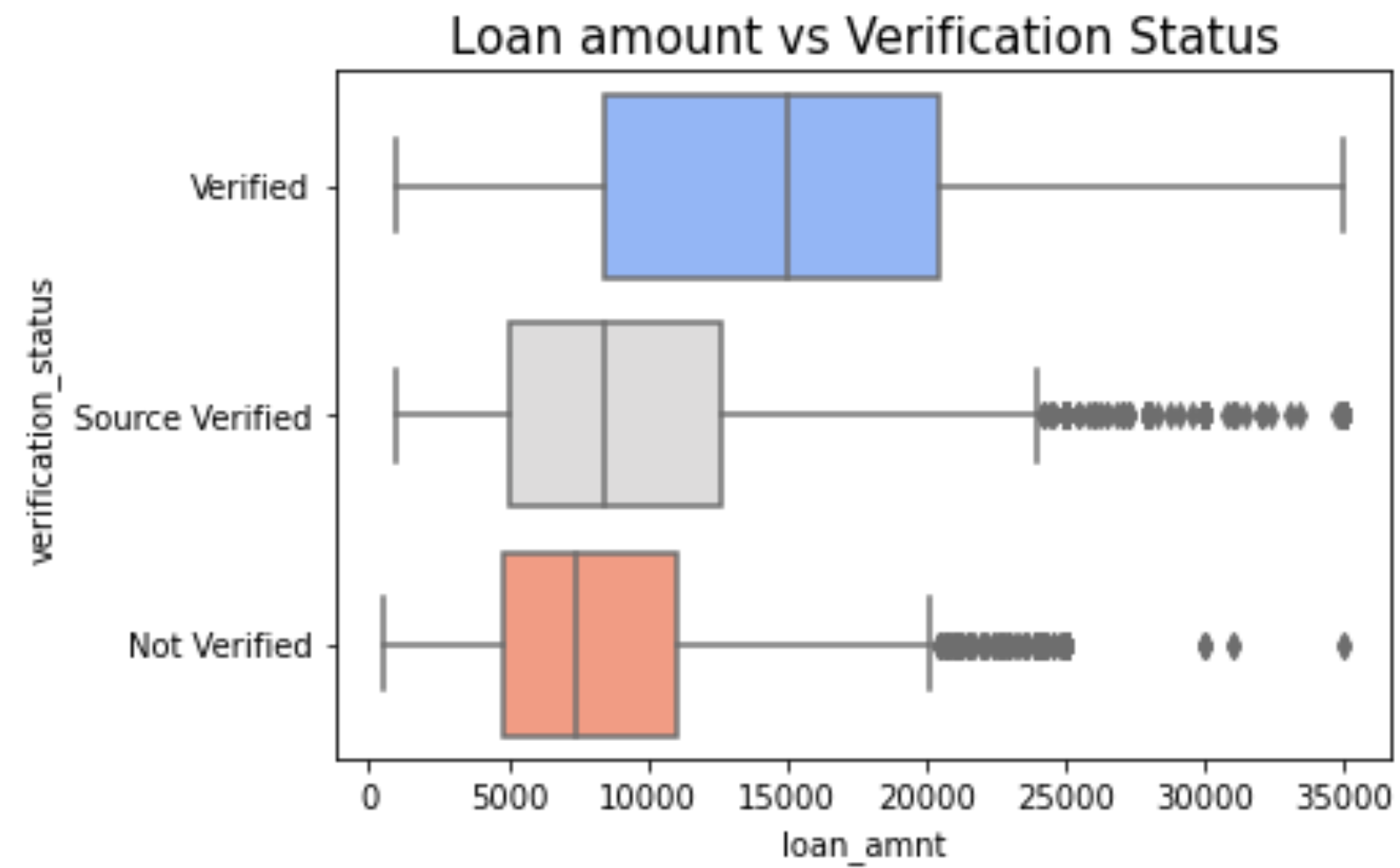


Insights

- higher the annual income higher the loan amount.
- Defaulter rate is more where funded amount is higher than 25000 and the annual income less than 120000.
- Majority of the loan are funded less than 25k which is good for the company to reduce the number of defaulters.

Bivariate Analysis

Loan amount vs Verification Status

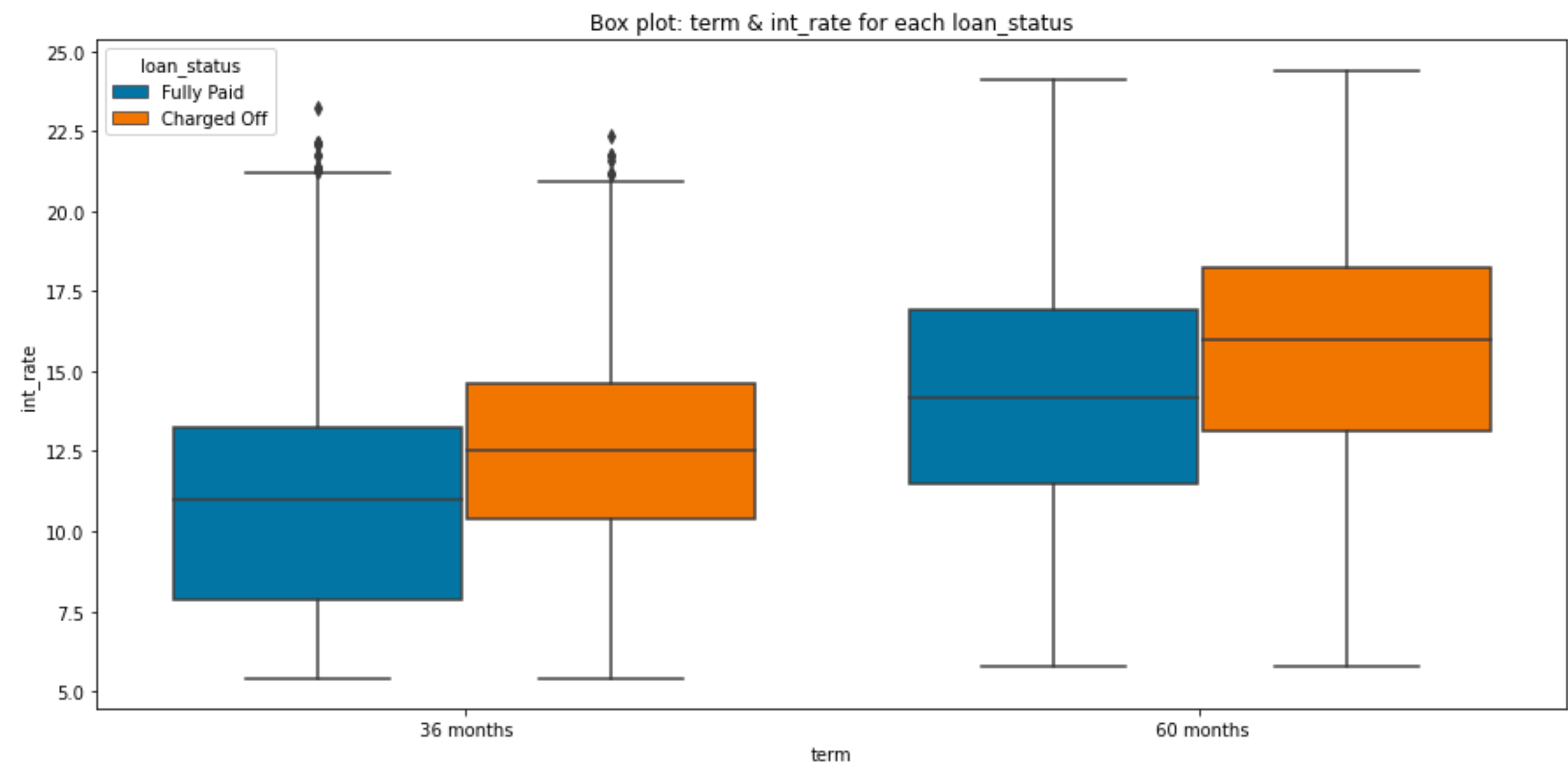


Insights

- Higher rate of defaulter are with account status as verified. Approving high amount loans to verified applicants is risky.

Bivariate Analysis

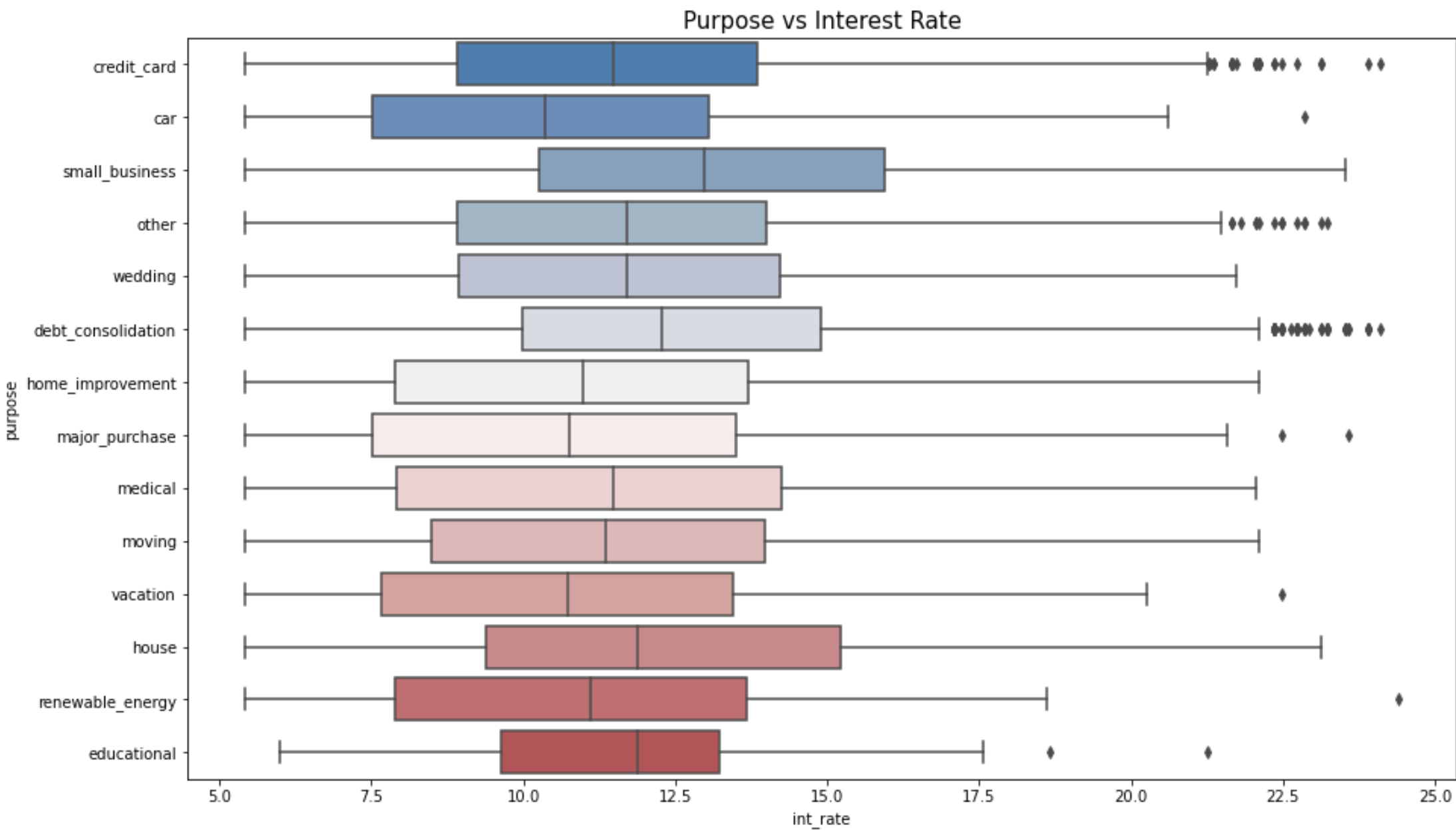
Term vs Interest Rate



Insights

- higher the annual income higher the loan amount.
- Defaulter rate is more where funded amount is higher than 25000 and the annual income less than 120000.
- Majority of the loan are funded less than 25k which is good for the company to reduce the number of defaulters.

Interest Rate vs Purpose

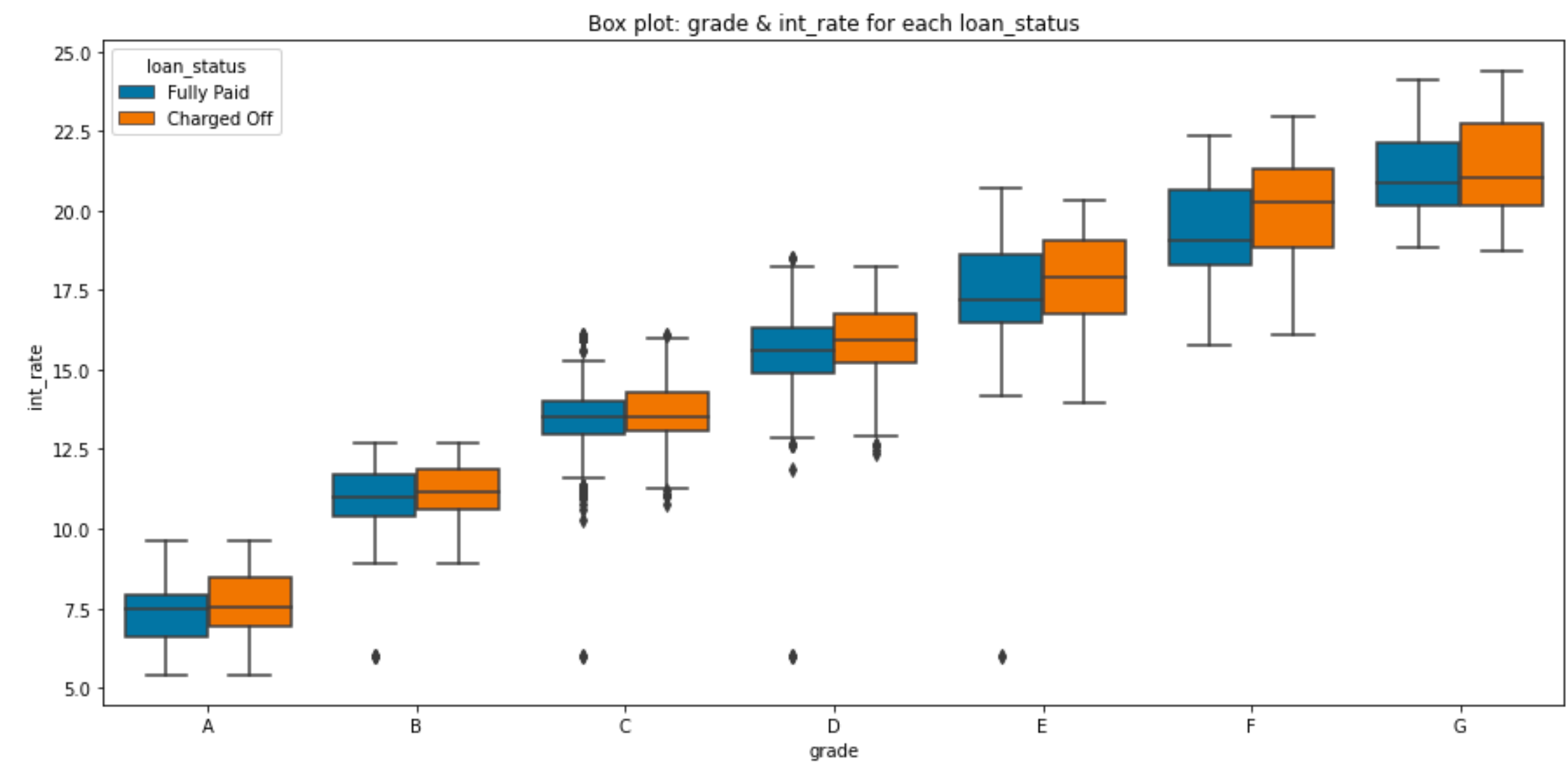


Insights

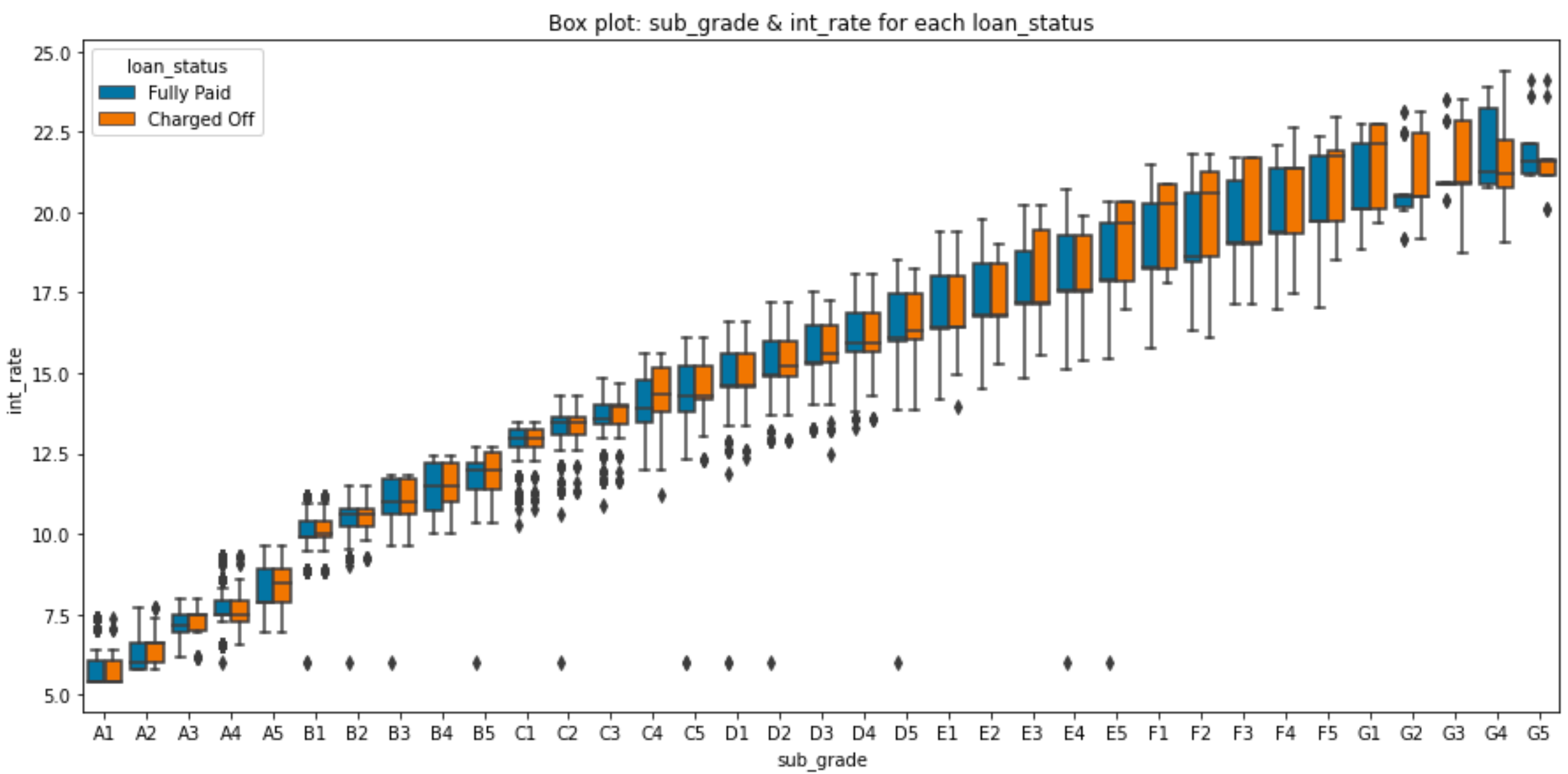
- Higher interest rate has been offered to the small business and then to debt consolidation.

Bivariate Analysis

Grade vs Interest Rate



Sub Grade vs Interest Rate



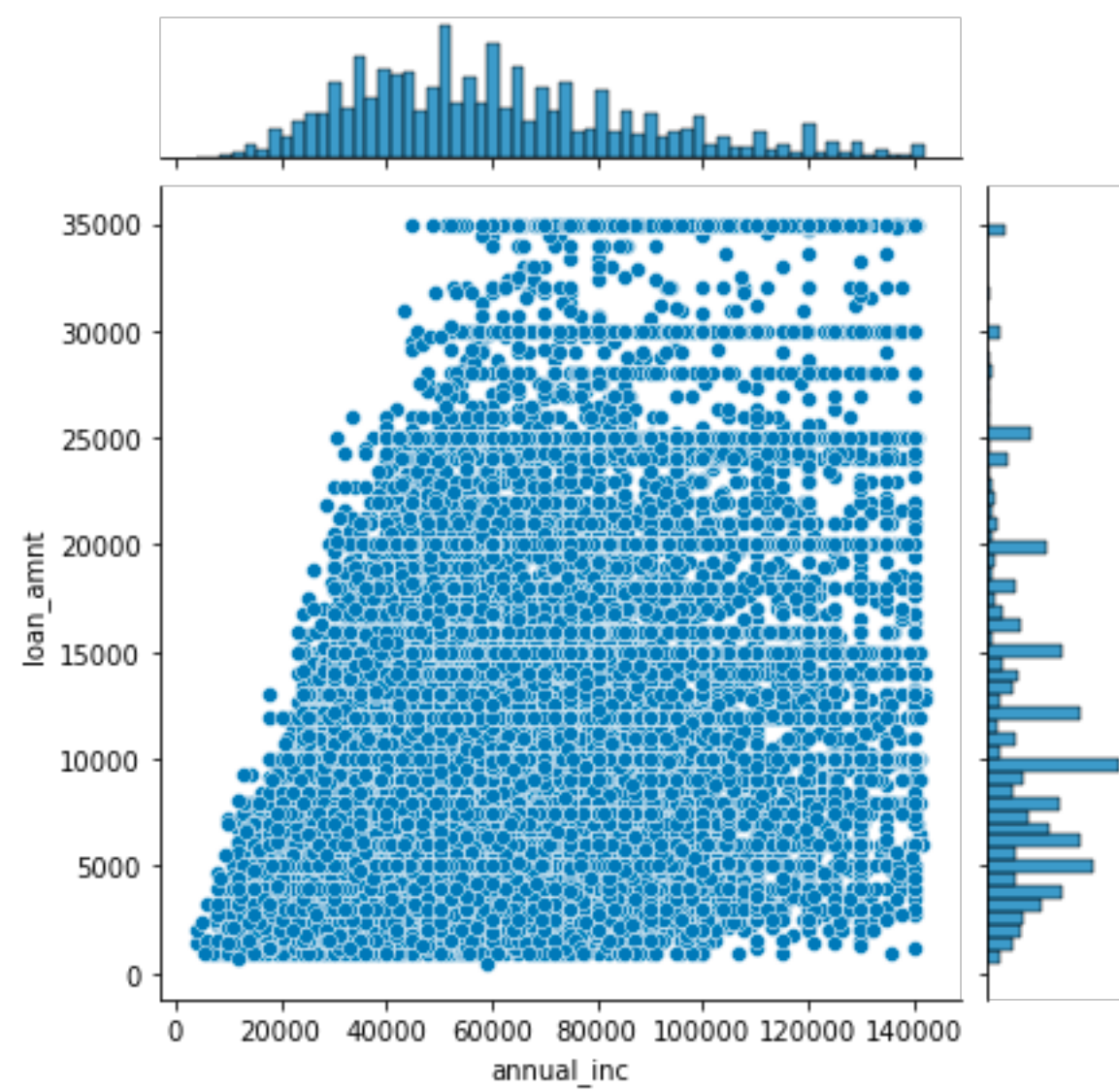
Insights

- Interest rate increased with the inverse of the grades which also have more defaulter at higher interest rates.

Insights

- Defaulter are more in the lower grades and the high loan amount will lead to more financial loss.

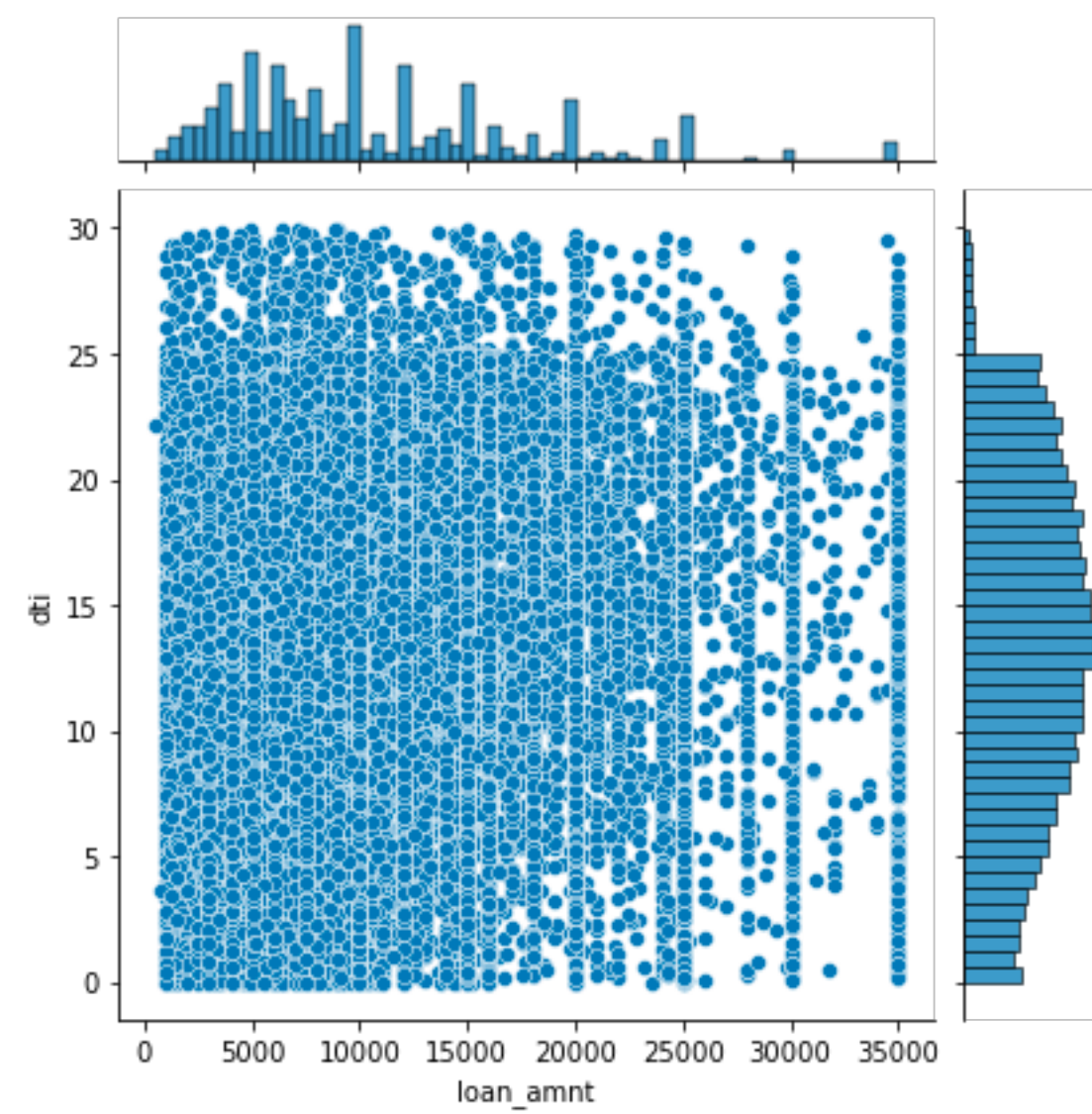
Annual income vs Loan Amount



Insights

- Higher loan amount has been issued to the mid level income applicant which is risky.

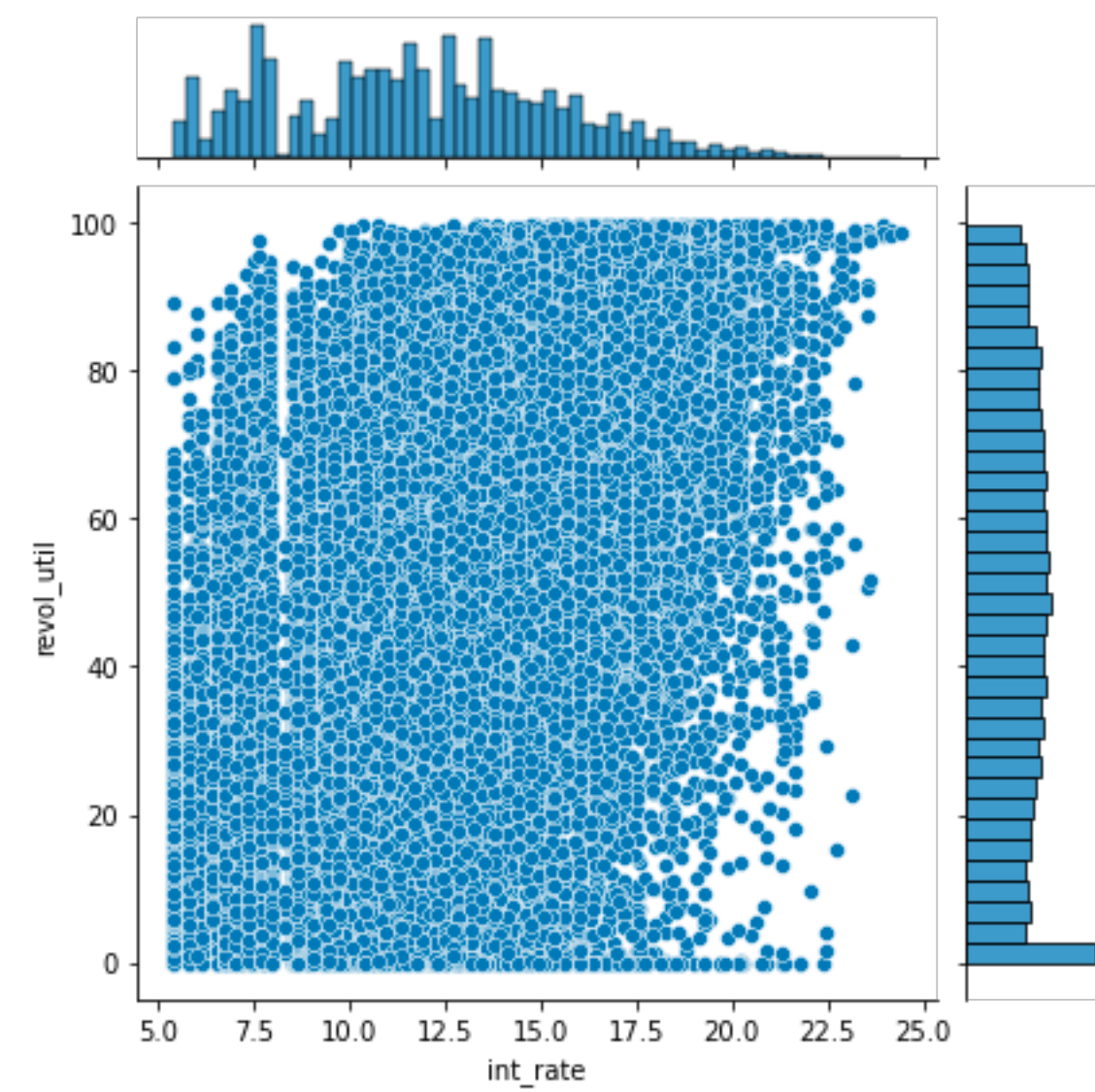
Loan Amount vs Dti



Insights

- Issuing high amount loan to applicants with less dpi will lead to financial loss

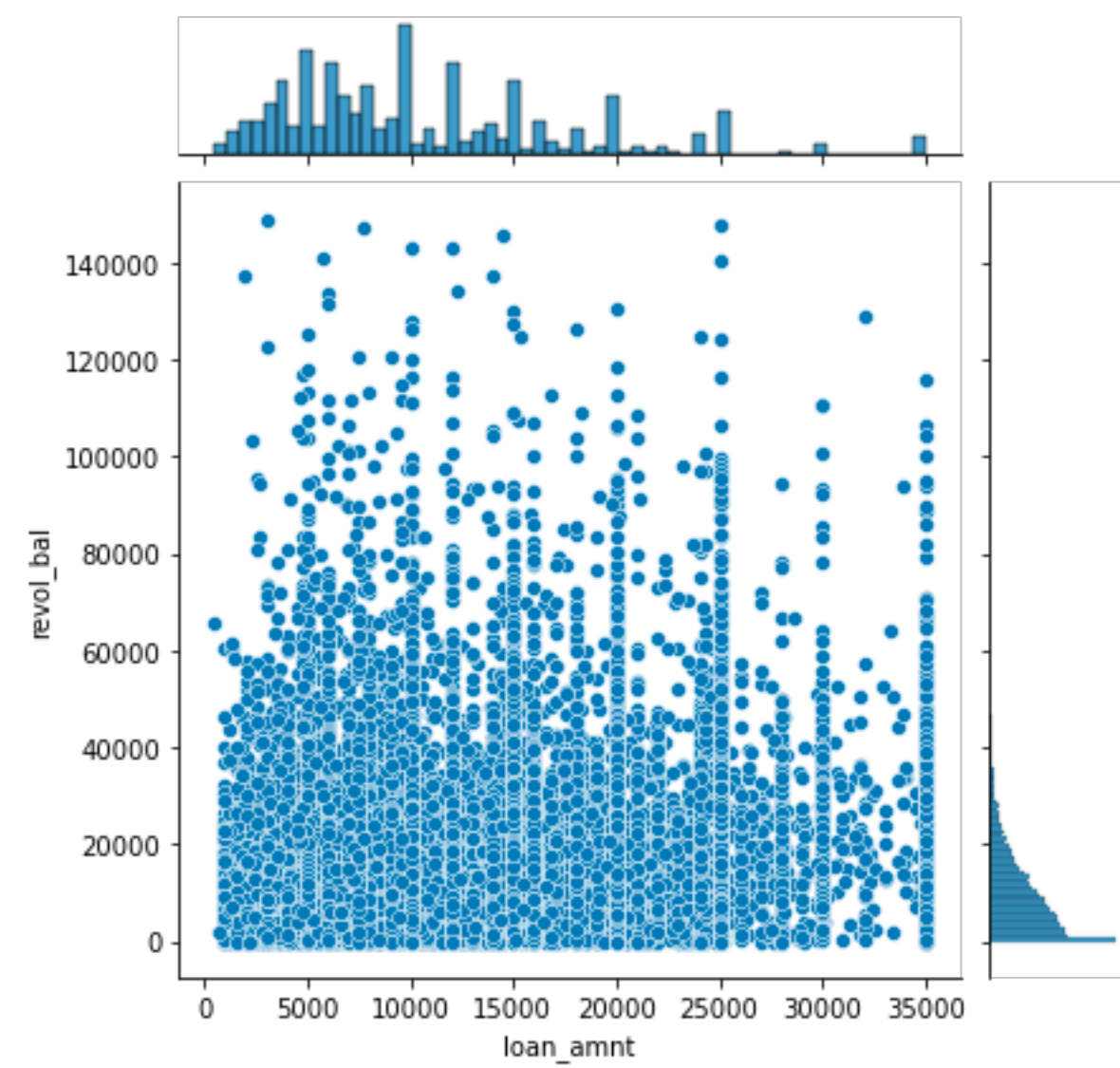
Revolving Utilization Rate vs Interest Rate



Insights

- Higher interest have to be provided to applicant with higher Revolving Utilisation Rate as the risk is more and the loan amount has to be reduced

Revolving Balance vs Loan Amount



Insights

- Approving high amount loans to applicants with high revolving balance is risky.

Observations

- The percentage of the defaulter is more for the high amount loans.
- Defaulters are more where funded amount is higher than 25K and the annual income less than 120K
- The percentage of defaulters increases as grade goes from A to G and same pattern is observed in sub grade as well.
- The percentage of defaulters is more for the applicants whose purpose is small business and number of defaulters are more for debt consolidation.
- More number of defaulters are observed with higher instalment amount.
- Higher interest rates have more defaulters and low interests are also offered to the high value loans.
- Applicants with high dti value have more tendency to default.
- Applicant with 10 years experience have high defaulting chances.
- Longer the term higher the defaulters.
- Applicants with more public derogatory have high chances of defaulting
- No particular pattern observed with address state but the state with more applicants have equal risk and few states like CA, FL, NJ, GA, MD, WA, MO have more defaulters.

Recommendations

Reducing the Financial loss from defaulters

- Stop approving the loan above 14% interest rate.
- Deny the loans for the applicants who's requirement is for small business.
- Reducing the loans for the applicants with high revolving utilisation rate.
- Reducing the loan amount for the long term loan.
- Stop issuing the loan to the applicants with high public derogatory background.
- Reducing the loan amount for the verified customer as there is high rate of defaulters in this group.

Increasing the profit by approving good loans

- Approving the loan to the applicants with annual income above 41K and the amount/income ratio above 30%.
- Providing the loan to the applicant whose dti rate is less.
- Approving the loans for the applicants whose requirement is for wedding, major purchase and credit card.
- Charging the higher interest rates to the applicant with revolving utilization rate is more as this accounts are risky.
- Approving the loans to the applicant with dti value less than 8.5.