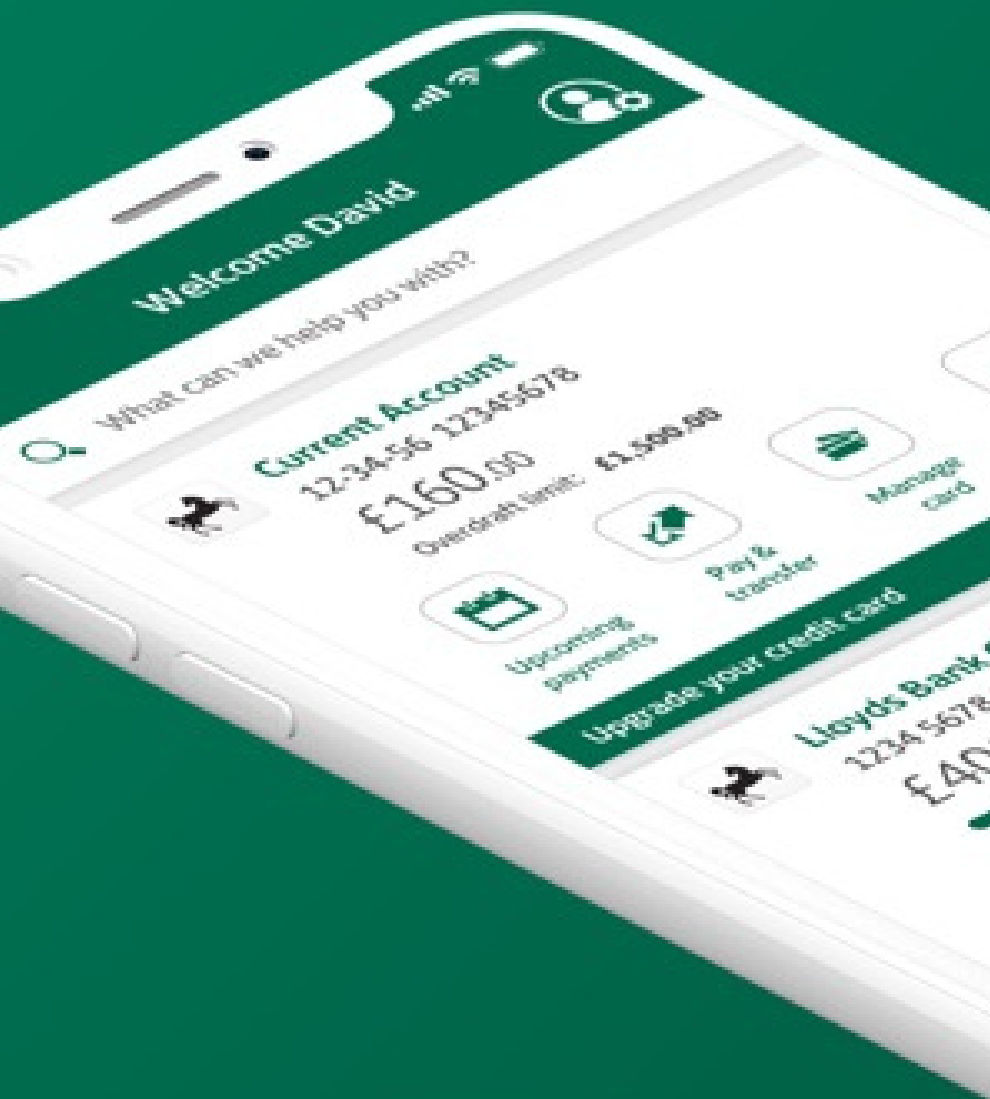


Everyday banking

Made easy, kept safe



Loan Prediction

Lloyd bank group Step Up Career Challenge

Presented By

Nguyet (Lana), Minh Le

Agenda



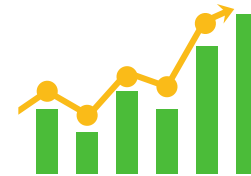
Data Summary

Briefly explain the dataset



Task A

Understand the different behaviours between customers who paid back their loan and customers who default



Task B

Using historical data, create a process that predicts the possibility of a new customer's ability to repay.

Data Summary

Field name	Missing value
emp_length	6%
emp_title	7%
avg_cur_bal	3%
mo_sin_old_il_acct	6%
mo_sin_old_rev_tl_op	3%
mo_sin_rcnt_rev_tl_op	3%
mo_sin_rcnt_tl	3%
mort_acc	2%
num_bc_tl	3%
num_il_tl	3%
num_op_rev_tl	3%
num_tl_90g_dpd_24m	3%
num_tl_op_past_12m	3%
percent_bc_gt_75	3%
total_bal_ex_mort	2%
inq_last_12m	49%
max_bal_bc	49%
mths_since_last_delinq	49%

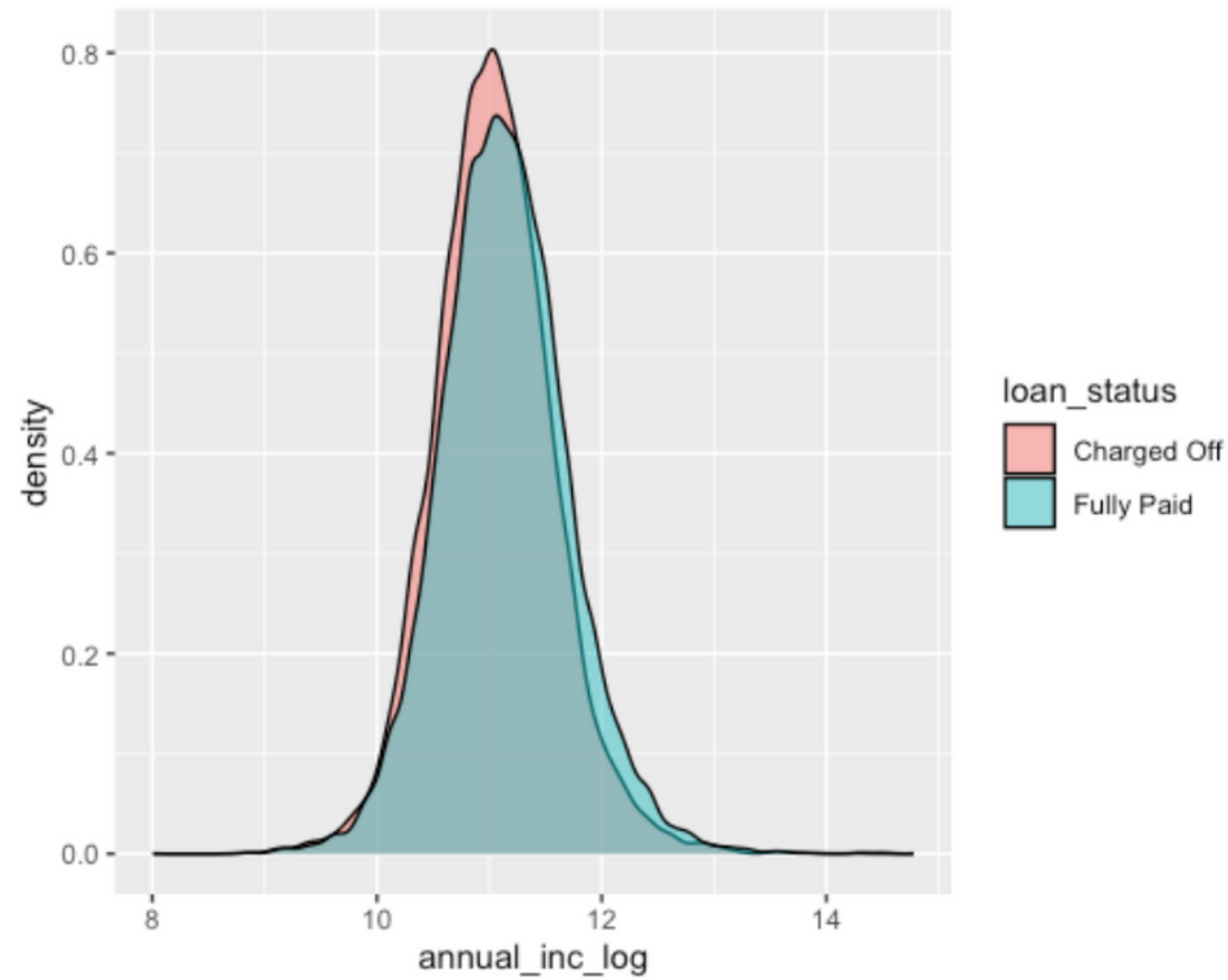
Data Evaluation

Field Name	Description
annual_inc	The annual income provided by the borrower during registration.
emp_length	Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years. Data captured as categorical.
home_ownership	The home ownership status provided by the borrower during registration. Our values are: RENT, OWN, MORTGAGE, OTHER
installment	The monthly payment owed by the borrower if the loan originates.
loan_amnt	The listed amount of the loan applied for by the borrower.
purpose	A category provided by the borrower for the loan request.
term	The number of payments on the loan. Values are in months and can be either 36 or 60.
int_rate	Interest Rate on the loan
avg_cur_bal	Average current balance of all current credit lending products / accounts
mort_acc	Number of mortgage accounts.
num_tl_op_past_12m	the number of accounts opened by the borrower in the past 12 months

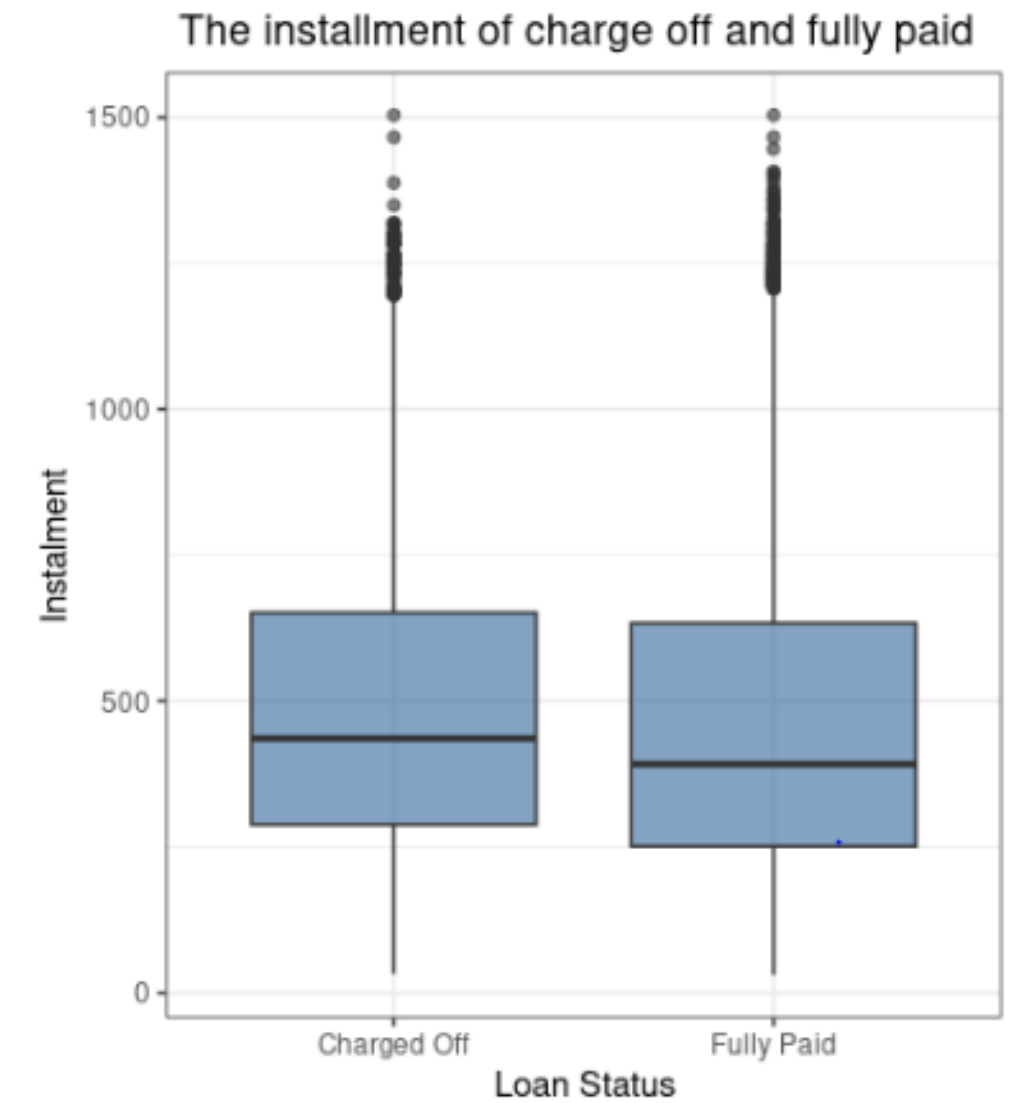
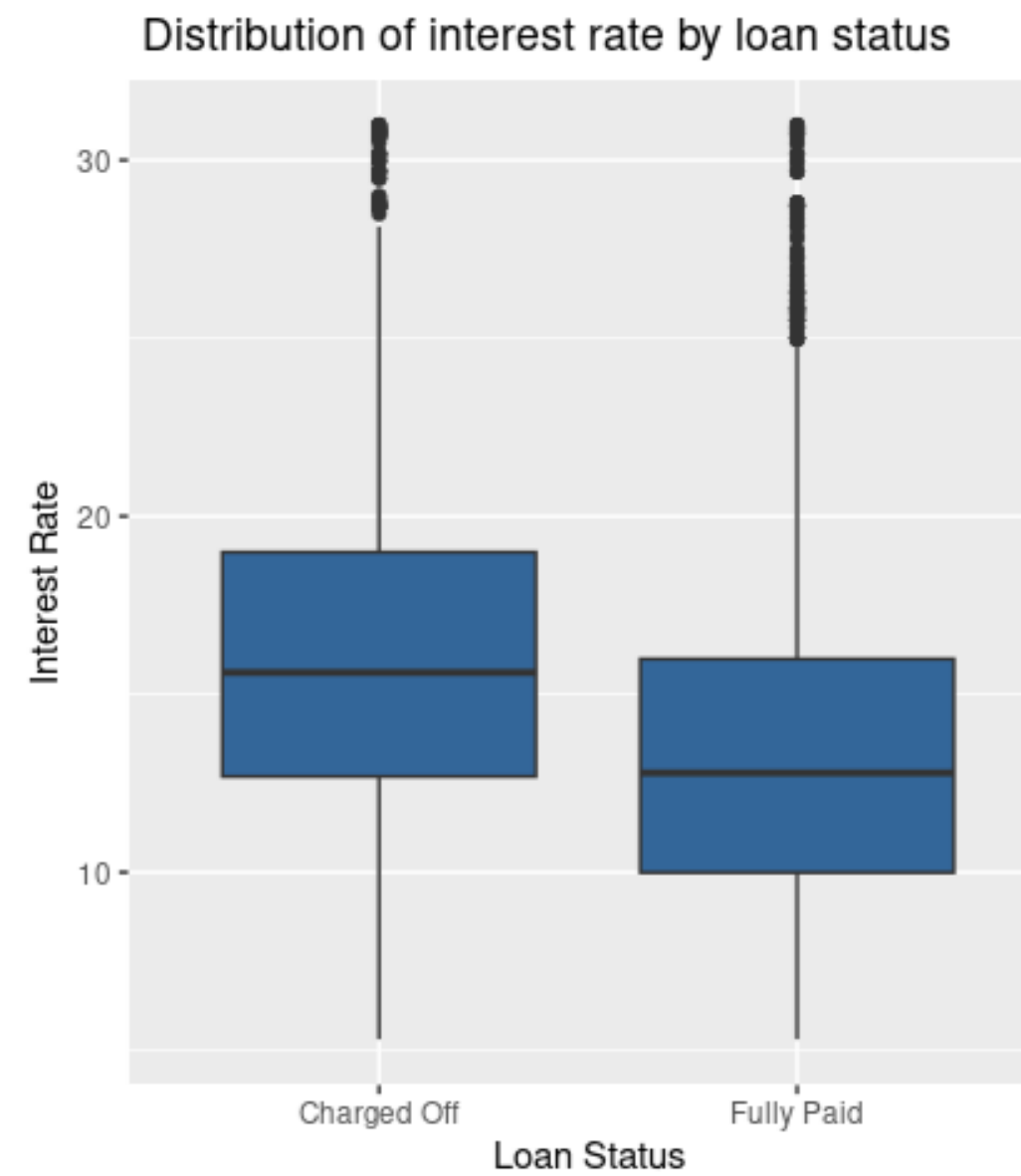
Data Dictionary

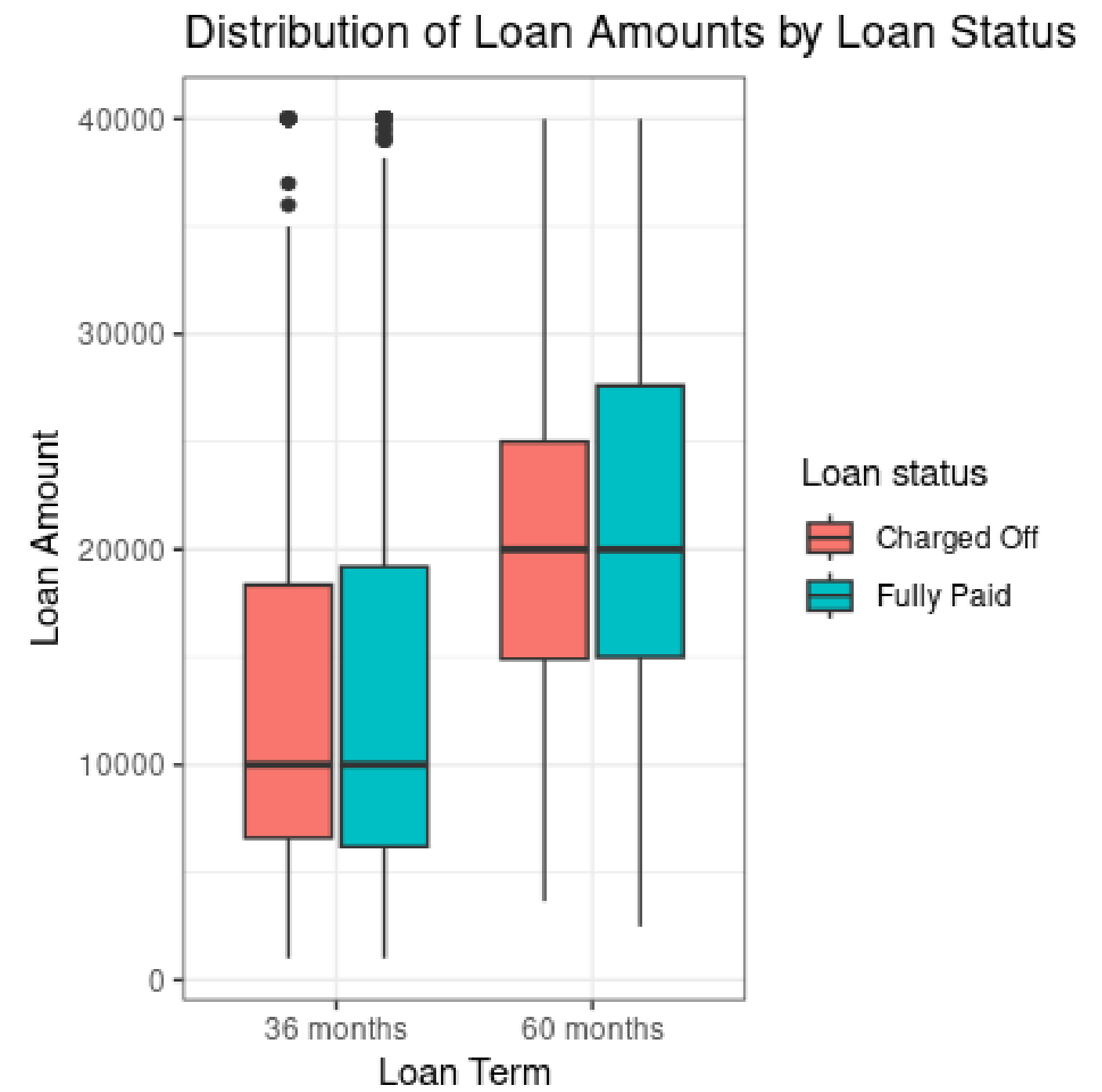
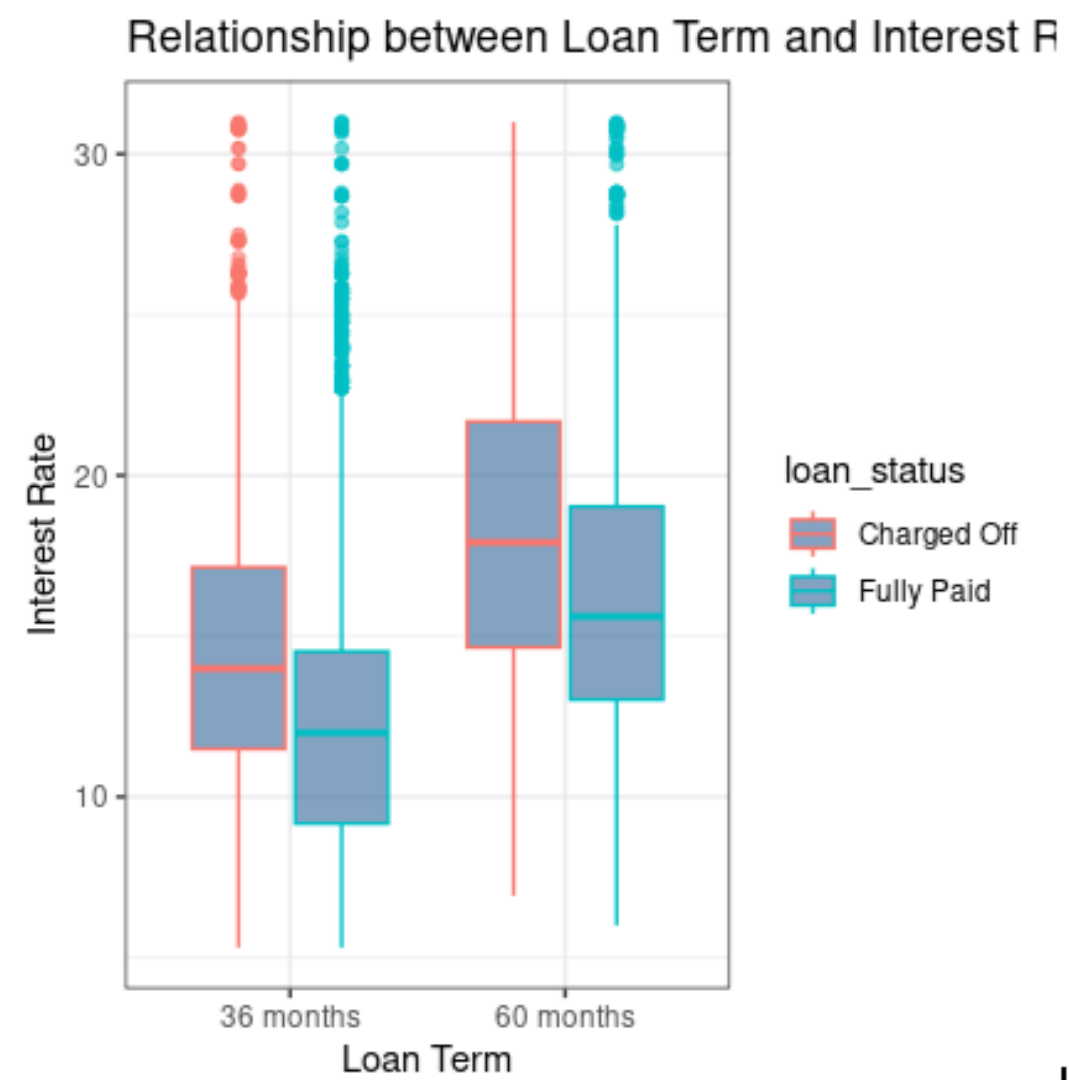


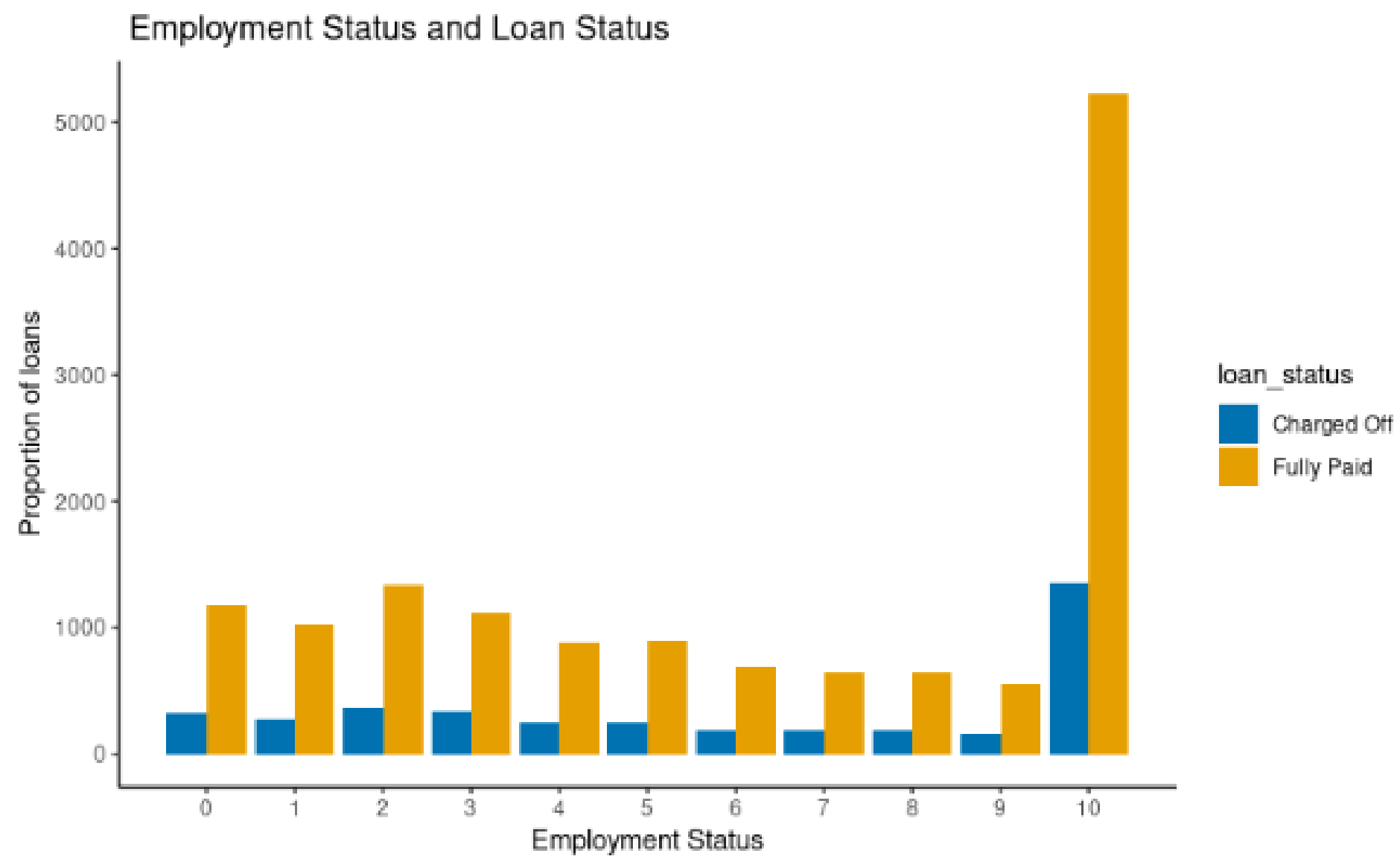
Task A



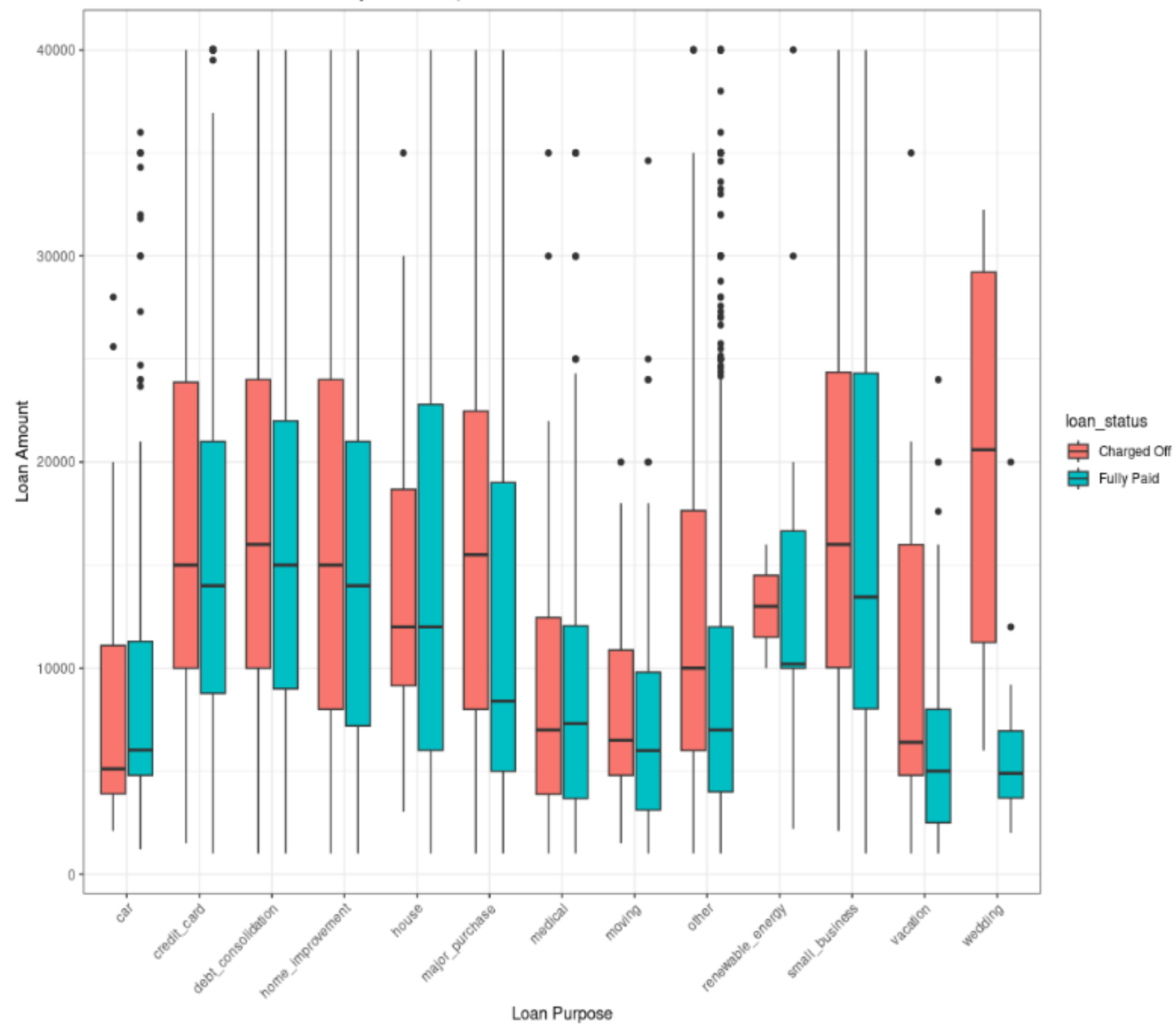
Distribution of annual income by loan status



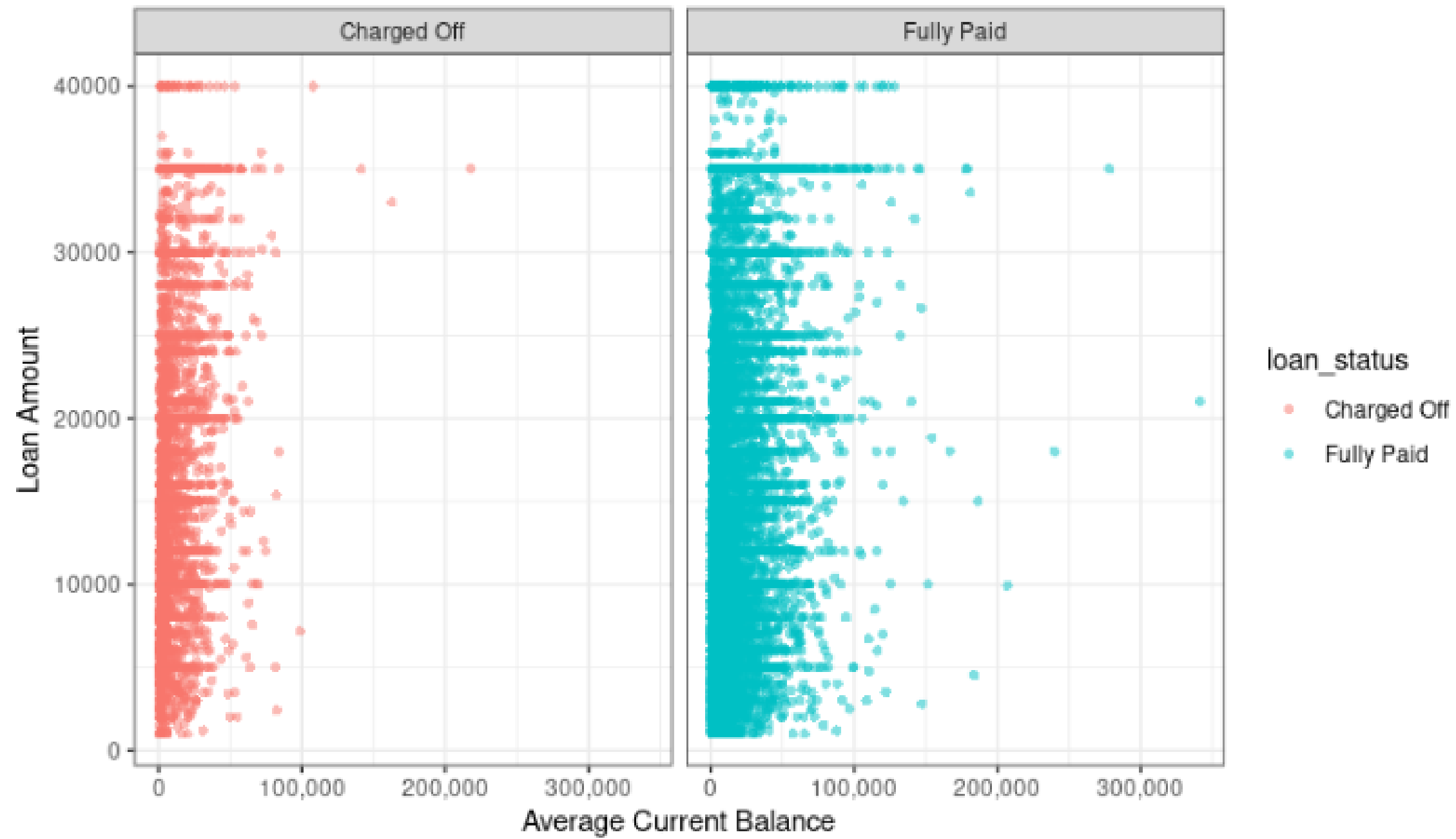




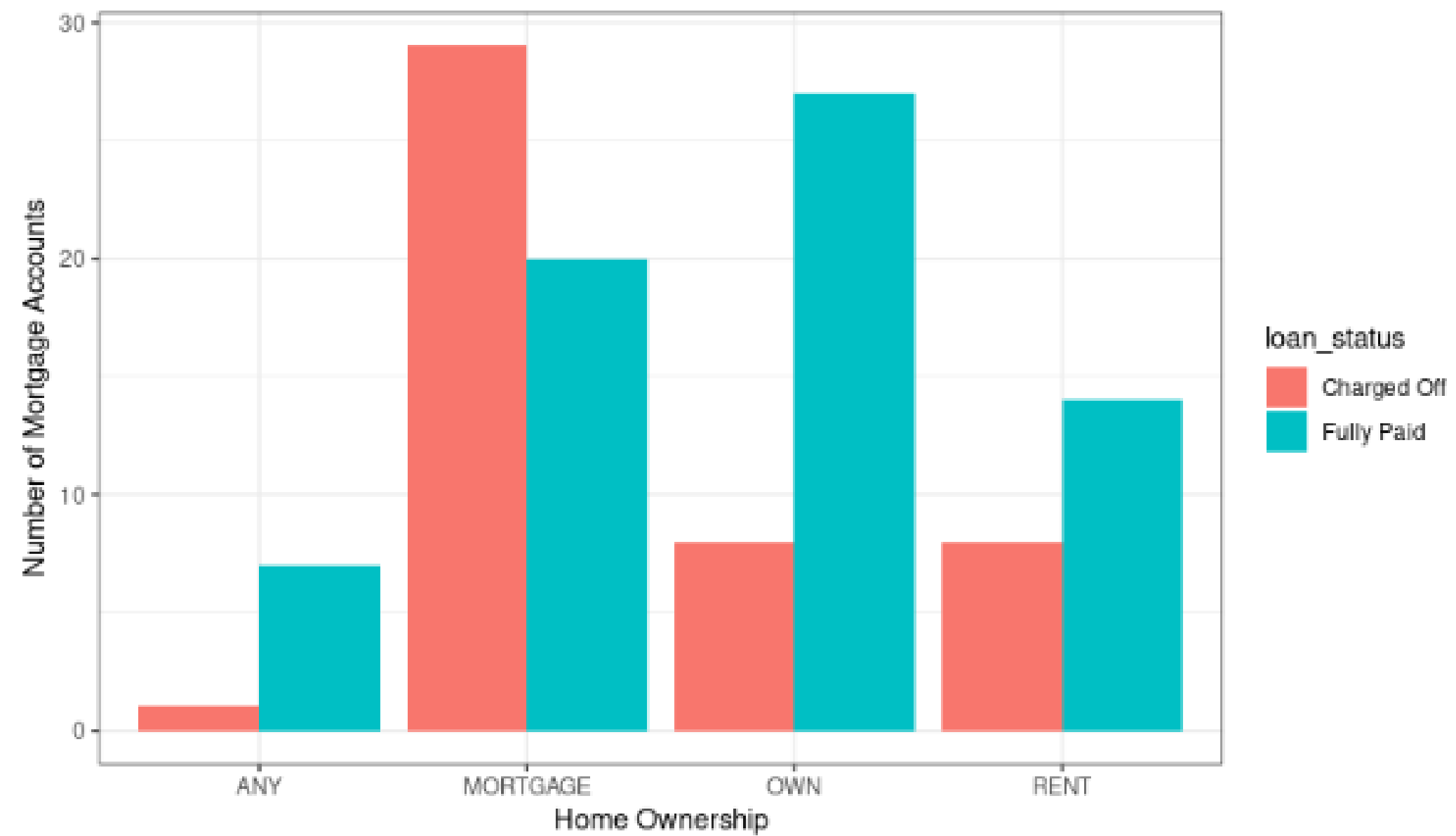
Distribution of Loan Amounts by Loan Purpose



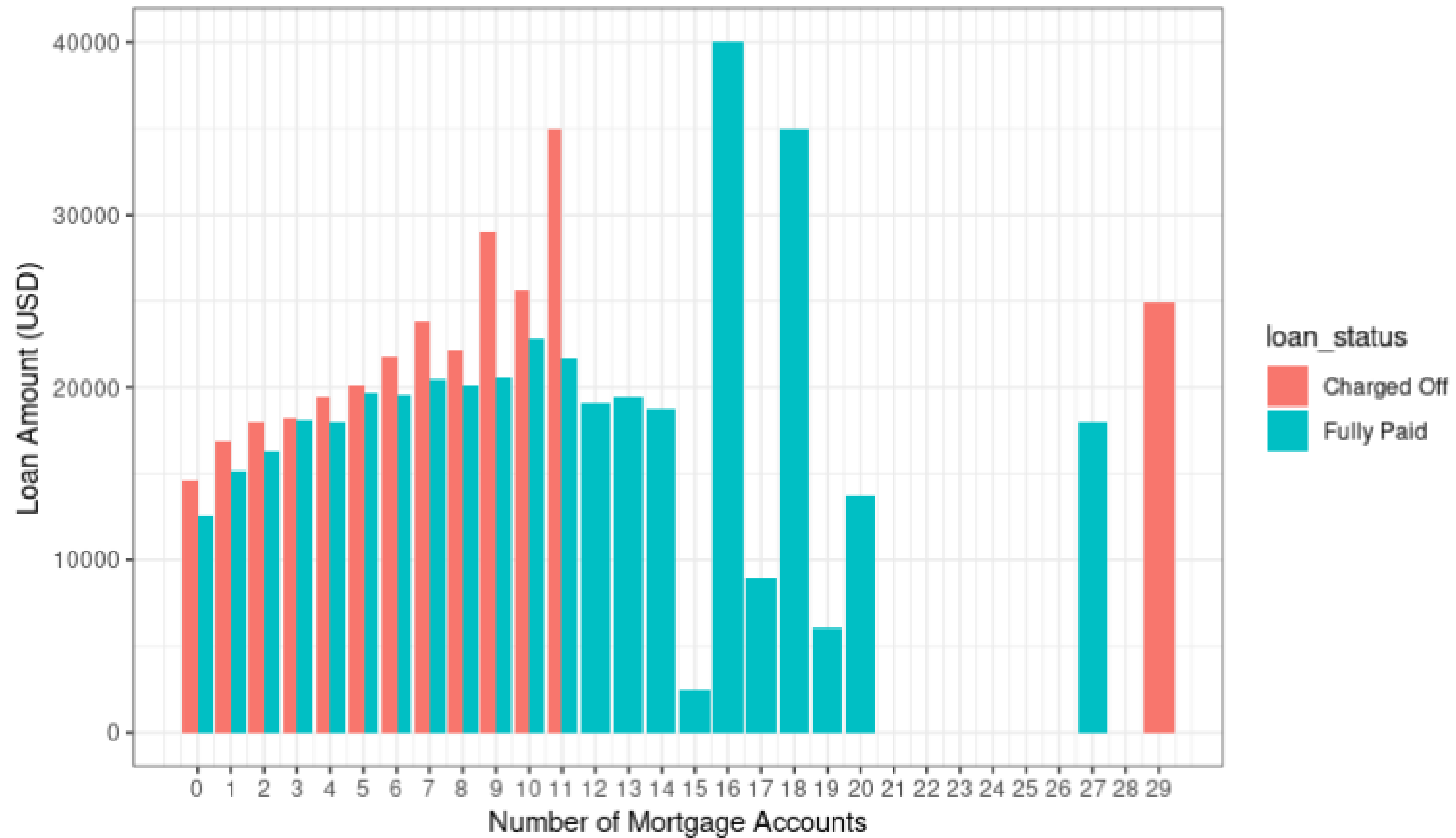
Relationship between Loan Status, Average Current Balance, and Loan Amount



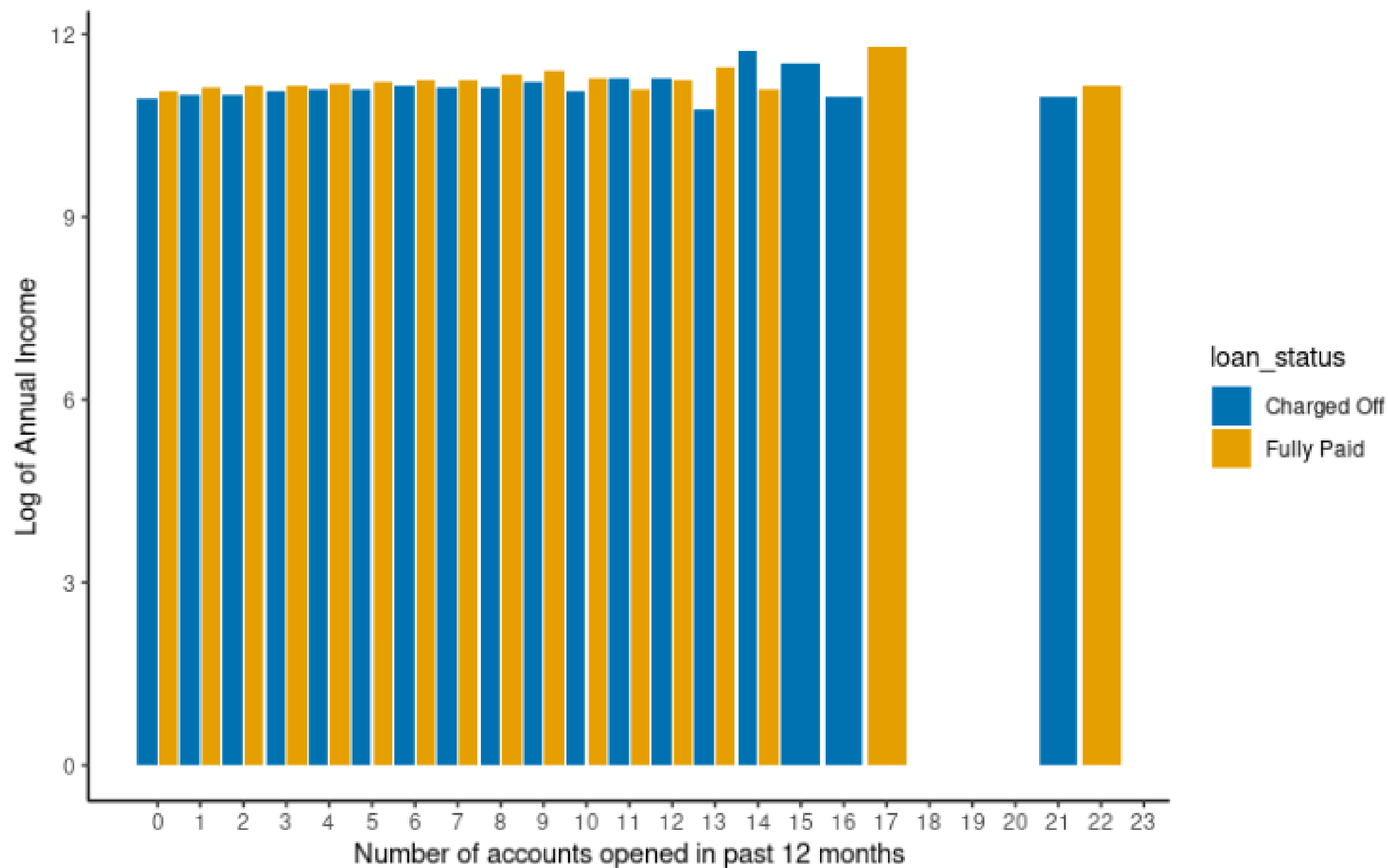
Relationship between Home Ownership, Number of Mortgage Accounts, and Loan Status



Average Loan Amount by Number of Mortgage Accounts and Loan Status



Relationship between Num_tl_op_past_12m, Annual_inc_log, and Loan Status



Task B

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.7061	0.3664	0.5535	0.7095	1.6617

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.116e+00	5.969e-01	-1.870	0.061414	.
annual_inc_log	3.713e-01	5.561e-02	6.677	2.45e-11	***
term60 months	-5.758e-01	5.222e-02	-11.026	< 2e-16	***
int_rate	-8.429e+00	5.293e-01	-15.925	< 2e-16	***
pub_rec_bankruptcies	-1.282e-01	5.515e-02	-2.324	0.020102	*
installment	-6.164e-04	1.014e-04	-6.081	1.19e-09	***
avg_cur_bal	7.796e-06	2.070e-06	3.767	0.000166	***
num_tl_op_past_12m	-7.514e-02	1.160e-02	-6.477	9.37e-11	***
mort_acc	8.304e-02	1.491e-02	5.570	2.55e-08	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Call:

roc.default(response = testData\$loan_status, predictor = predicted)

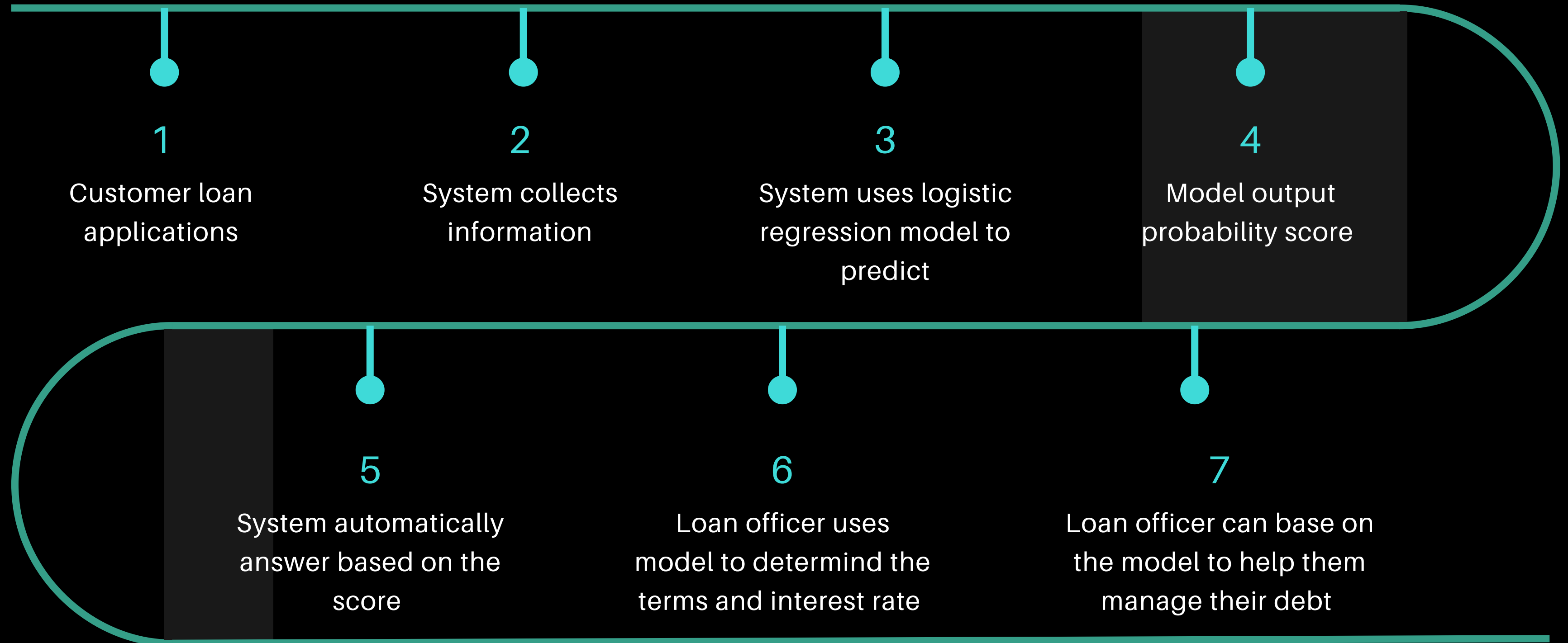
Data: predicted in 1169 controls (testData\$loan_status 0) < 4328 cases (testData\$loan_status 1).

Area under the curve: 0.6866

A logistic regression model result

Model evaluation result

Loan prediction process



Application information

Annual income (log)

Term

Loan Amount

Rate

Number mortgaged account

Number account open last 12 months

Employment status (Length)

House status (Home Ownership)

Purpose

Average current balance

Talks and Workshops

Contact Me

Loan Prediction

For inquiries,
contact me.

Email

lanalee297@gmail.com

Contact

+44 7548908922