The SAS System

The CONTENTS Procedure

Data Set Name	WORK.LOANDATA	**Observations**	42538
Member Type	DATA	**Variables**	60
Engine	V9	**Indexes**	0
Created	12/05/2017 15:02:01	**Observation Length**	1080
Last Modified	12/05/2017 15:02:01	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information	
Data Set Page Size	90112
Number of Data Set Pages	513
First Data Page	1
Max Obs per Page	83
Obs in First Data Page	76
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	C:\Users\Saloni\AppData\Local\Temp\SAS Temporary Files_TD4224_DESKTOP-IQ5UG81_\loandata.sas7bdat
Release Created	9.0401M2
Host Created	W32_8HOME

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat
53	acc_now_delinq	Num	8	BEST12.	BEST32.
21	addr_state	Char	2	$2.	$2.
12	annual_inc	Num	8	BEST12.	BEST32.
52	annual_inc_joint	Char	1	$1.	$1.
51	application_type	Char	10	$10.	$10.
60	categorizedIncome	Num	8		
44	collection_recovery_fee	Num	8	BEST12.	BEST32.
23	delinq_2yrs	Num	8	BEST12.	BEST32.
54	delinq_amnt	Num	8	BEST12.	BEST32.
17	desc	Char	588	$588.	$588.
22	dti	Num	8	BEST12.	BEST32.
24	earliest_cr_line	Char	6	$6.	$6.
10	emp_length	Char	9	$9.	$9.
9	emp_title	Char	35	$35.	$35.
57	fico_avg	Num	8		

26	fico_range_high	Num	8	BEST12.	BEST32.
25	fico_range_low	Num	8	BEST12.	BEST32.
2	funded_amnt	Num	8	BEST12.	BEST32.
3	funded_amnt_inv	Num	8	BEST12.	BEST32.
7	grade	Char	1	$1.	$1.
11	home_ownership	Char	8	$8.	$8.
35	initial_list_status	Char	1	$1.	$1.
27	inq_last_6mths	Num	8	BEST12.	BEST32.
6	installment	Num	8	BEST12.	BEST32.
5	int_rate	Char	6	$6.	$6.
14	issue_d	Char	6	$6.	$6.
48	last_credit_pull_d	Char	6	$6.	$6.
58	last_fico_avg	Num	8		
49	last_fico_range_high	Num	8	BEST12.	BEST32.
50	last_fico_range_low	Num	8	BEST12.	BEST32.
46	last_pymnt_amnt	Num	8	BEST12.	BEST32.
45	last_pymnt_d	Char	6	$6.	$6.
1	loan_amnt	Num	8	BEST12.	BEST32.
59	loan_paid	Num	8		
15	loan_status	Char	11	$11.	$11.
28	mths_since_last_delinq	Num	8	BEST12.	BEST32.
29	mths_since_last_record	Char	1	$1.	$1.
47	next_pymnt_d	Char	6	$6.	$6.
30	open_acc	Num	8	BEST12.	BEST32.
36	out_prncp	Num	8	BEST12.	BEST32.
37	out_prncp_inv	Num	8	BEST12.	BEST32.
31	pub_rec	Num	8	BEST12.	BEST32.
55	pub_rec_bankruptcies	Num	8	BEST12.	BEST32.
18	purpose	Char	18	$18.	$18.
16	pymnt_plan	Char	1	$1.	$1.
43	recoveries	Num	8	BEST12.	BEST32.
32	revol_bal	Num	8	BEST12.	BEST32.
33	revol_util	Char	6	$6.	$6.
8	sub_grade	Char	2	$2.	$2.
56	tax_liens	Num	8	BEST12.	BEST32.
4	term	Char	9	$9.	$9.
19	title	Char	37	$37.	$37.
34	total_acc	Num	8	BEST12.	BEST32.
38	total_pymnt	Num	8	BEST12.	BEST32.
39	total_pymnt_inv	Num	8	BEST12.	BEST32.
41	total_rec_int	Num	8	BEST12.	BEST32.
42	total_rec_late_fee	Num	8	BEST12.	BEST32.
40	total_rec_prncp	Num	8	BEST12.	BEST32.
13					

	verification_status	Char	15	$15.	$15.
20	zip_code	Char	5	$5.	$5.

The SAS System

The FREQ Procedure

categorizedIncome	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	5088	11.96	5088	11.96
2	11878	27.92	16966	39.88
3	19433	45.68	36399	85.57
4	6117	14.38	42516	99.95
5	22	0.05	42538	100.00

The SAS System

The FREQ Procedure

loan_paid	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	8952	21.04	8952	21.04
1	33586	78.96	42538	100.00

The SAS System

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
loan_amnt	42535	11089.72	7410.94	500.0000000	35000.00
funded_amnt	42535	10821.59	7146.91	500.0000000	35000.00
funded_amnt_inv	42535	10139.83	7131.69	0	35000.00
installment	42535	322.6230627	208.9272165	15.6700000	1305.19
annual_inc	42531	69136.56	64096.35	1896.00	6000000.00
dti	42535	13.3730431	6.7263149	0	29.9900000
delinq_2yrs	42506	0.1524491	0.5124065	0	13.0000000
fico_range_low	42535	713.0525450	36.1884385	610.0000000	825.0000000
fico_range_high	42535	717.0525450	36.1884385	614.0000000	829.0000000
inq_last_6mths	42506	1.0814238	1.5274548	0	33.0000000
mths_since_last_delinq	15609	35.0176180	22.4184269	0	120.0000000
open_acc	42506	9.3439514	4.4962739	1.0000000	47.0000000
pub_rec	42506	0.0581565	0.2457131	0	5.0000000
revol_bal	42535	14297.86	22018.44	0	1207359.00
total_acc	42506	22.1244060	11.5928113	1.0000000	90.0000000
out_prncp	42535	11.8495618	125.3765634	0	3555.85
out_prncp_inv	42535	11.8158909	125.0618595	0	3553.30
total_pymnt	42535	12005.64	9067.24	0	58563.68
total_pymnt_inv	42535	11298.59	9010.32	0	58563.68
total_rec_prncp	42535	9664.27	7087.47	0	35000.02
total_rec_int	42535	2239.55	2583.72	0	23823.75
total_rec_late_fee	42535	1.5153209	7.8277999	0	209.0000000
recoveries	42535	100.3073533	718.0787158	0	29623.35
collection_recovery_fee	42535	13.9567374	159.6218609	0	7002.19
last_pymnt_amnt	42535	2612.72	4385.20	0	36115.20
last_fico_range_high	42535	689.3053015	80.2028819	0	850.0000000
last_fico_range_low	42535	676.3827436	119.0115547	0	845.0000000
acc_now_delinq	42506	0.000094104	0.0097004	0	1.0000000
delinq_amnt	42506	0.1430386	29.3595794	0	6053.00
pub_rec_bankruptcies	41170	0.0452271	0.2087374	0	2.0000000
tax_liens	42430	0.000023568	0.0048547	0	1.0000000
fico_avg	42535	715.0525450	36.1884385	612.0000000	827.0000000
last_fico_avg	42535	682.8440226	95.9865335	0	847.5000000
loan_paid	42538	0.7895529	0.4076310	0	1.0000000
categorizedIncome	42538	2.6263811	0.8740549	1.0000000	5.0000000

Logistic Regression for Credit Risk Analysis

The LOGISTIC Procedure

Model Information	
Data Set	WORK.LOANDATA
Response Variable	loan_paid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	42538
Number of Observations Used	42506

Response Profile		
Ordered Value	loan_paid	Total Frequency
1	1	33586
2	0	8920

Probability modeled is loan_paid=1.

Note: 32 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information					
Class	Value	Design Variables			
categorizedIncome	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	43677.891	28205.316
SC	43686.548	28291.890
-2 Log L	43675.891	28185.316

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15490.5746	9	<.0001
Score	14982.6229	9	<.0001

Wald	7415.5945	9	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
categorizedIncome	4	43.4539	<.0001
fico_avg	1	1024.3950	<.0001
loan_amnt	1	3765.2344	<.0001
total_pymnt	1	3716.7876	<.0001
delinq_2yrs	1	2.2429	0.1342
last_fico_avg	1	2129.9624	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-17.9456	0.6913	673.8685	<.0001
categorizedIncome	1	1	1.5105	0.5808	6.7641	0.0093
categorizedIncome	2	1	1.6664	0.5801	8.2520	0.0041
categorizedIncome	3	1	1.7877	0.5800	9.5006	0.0021
categorizedIncome	4	1	1.7402	0.5814	8.9580	0.0028
fico_avg		1	0.0172	0.000536	1024.3950	<.0001
loan_amnt		1	-0.00039	6.402E-6	3765.2344	<.0001
total_pymnt		1	0.000342	5.61E-6	3716.7876	<.0001
delinq_2yrs		1	0.0410	0.0274	2.2429	0.1342
last_fico_avg		1	0.00878	0.000190	2129.9624	<.0001

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
categorizedIncome 1 vs 5	4.529	1.451	14.138
categorizedIncome 2 vs 5	5.293	1.698	16.501
categorizedIncome 3 vs 5	5.976	1.917	18.624
categorizedIncome 4 vs 5	5.698	1.823	17.809
fico_avg	1.017	1.016	1.018
loan_amnt	1.000	1.000	1.000
total_pymnt	1.000	1.000	1.000
delinq_2yrs	1.042	0.987	1.099
last_fico_avg	1.009	1.008	1.009

Association of Predicted Probabilities and Observed Responses

Percent Concordant	86.2	Somers' D	0.726
Percent Discordant	13.6	Gamma	0.727
Percent Tied	0.2	Tau-a	0.241
Pairs	299587120	c	0.863

Logistic Regression for Credit Risk Analysis without delinq_2yrs

The LOGISTIC Procedure

Model Information	
Data Set	WORK.LOANDATA
Response Variable	loan_paid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	42538
Number of Observations Used	42535

Response Profile		
Ordered Value	loan_paid	Total Frequency
1	1	33586
2	0	8949

Probability modeled is loan_paid=1.

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information					
Class	Value	Design Variables			
categorizedIncome	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	43768.375	28325.539
SC	43777.033	28403.462
-2 Log L	43766.375	28307.539

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15458.8355	8	<.0001
Score	14950.3404	8	<.0001

Wald	7421.1319	8	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
categorizedIncome	4	60.5332	<.0001
fico_avg	1	1030.3561	<.0001
loan_amnt	1	3755.1808	<.0001
total_pymnt	1	3716.9871	<.0001
last_fico_avg	1	2124.4543	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-18.7370	0.6668	789.5672	<.0001
categorizedIncome	1	1	2.5543	0.5496	21.6023	<.0001
categorizedIncome	2	1	2.7129	0.5489	24.4268	<.0001
categorizedIncome	3	1	2.8339	0.5488	26.6660	<.0001
categorizedIncome	4	1	2.7820	0.5503	25.5544	<.0001
fico_avg		1	0.0168	0.000524	1030.3561	<.0001
loan_amnt		1	-0.00039	6.388E-6	3755.1808	<.0001
total_pymnt		1	0.000342	5.604E-6	3716.9871	<.0001
last_fico_avg		1	0.00874	0.000190	2124.4543	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
categorizedIncome 1 vs 5	12.862	4.381	37.767
categorizedIncome 2 vs 5	15.072	5.140	44.198
categorizedIncome 3 vs 5	17.012	5.803	49.877
categorizedIncome 4 vs 5	16.152	5.492	47.497
fico_avg	1.017	1.016	1.018
loan_amnt	1.000	1.000	1.000
total_pymnt	1.000	1.000	1.000
last_fico_avg	1.009	1.008	1.009

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	86.1	Somers' D	0.724
Percent Discordant	13.7	Gamma	0.726
Percent Tied	0.2	Tau-a	0.241
Pairs	300561114	c	0.862

Multiple Regression for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

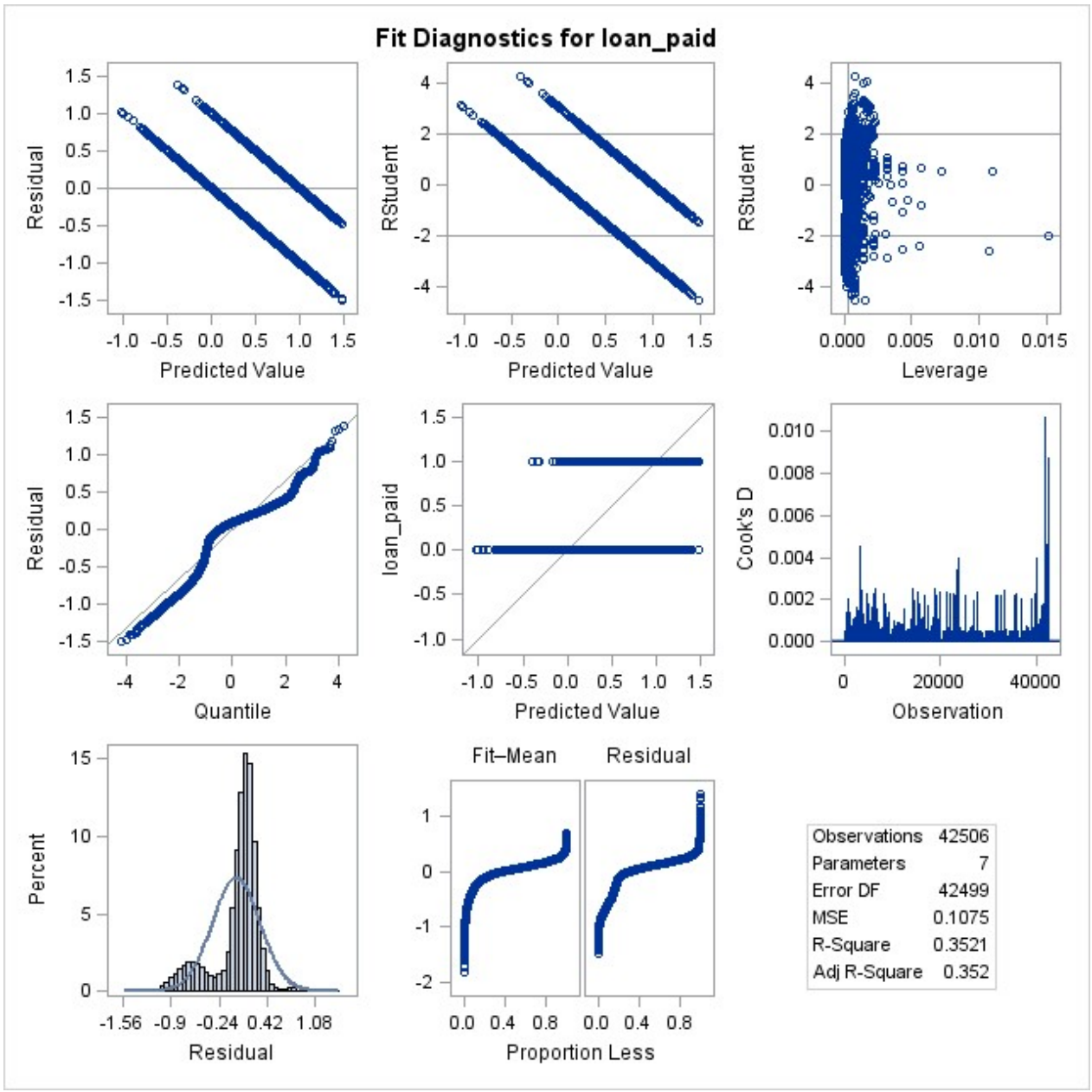
Root MSE	0.32780	R-Square	0.3521
Dependent Mean	0.79015	Adj R-Sq	0.3520
Coeff Var	41.48622		

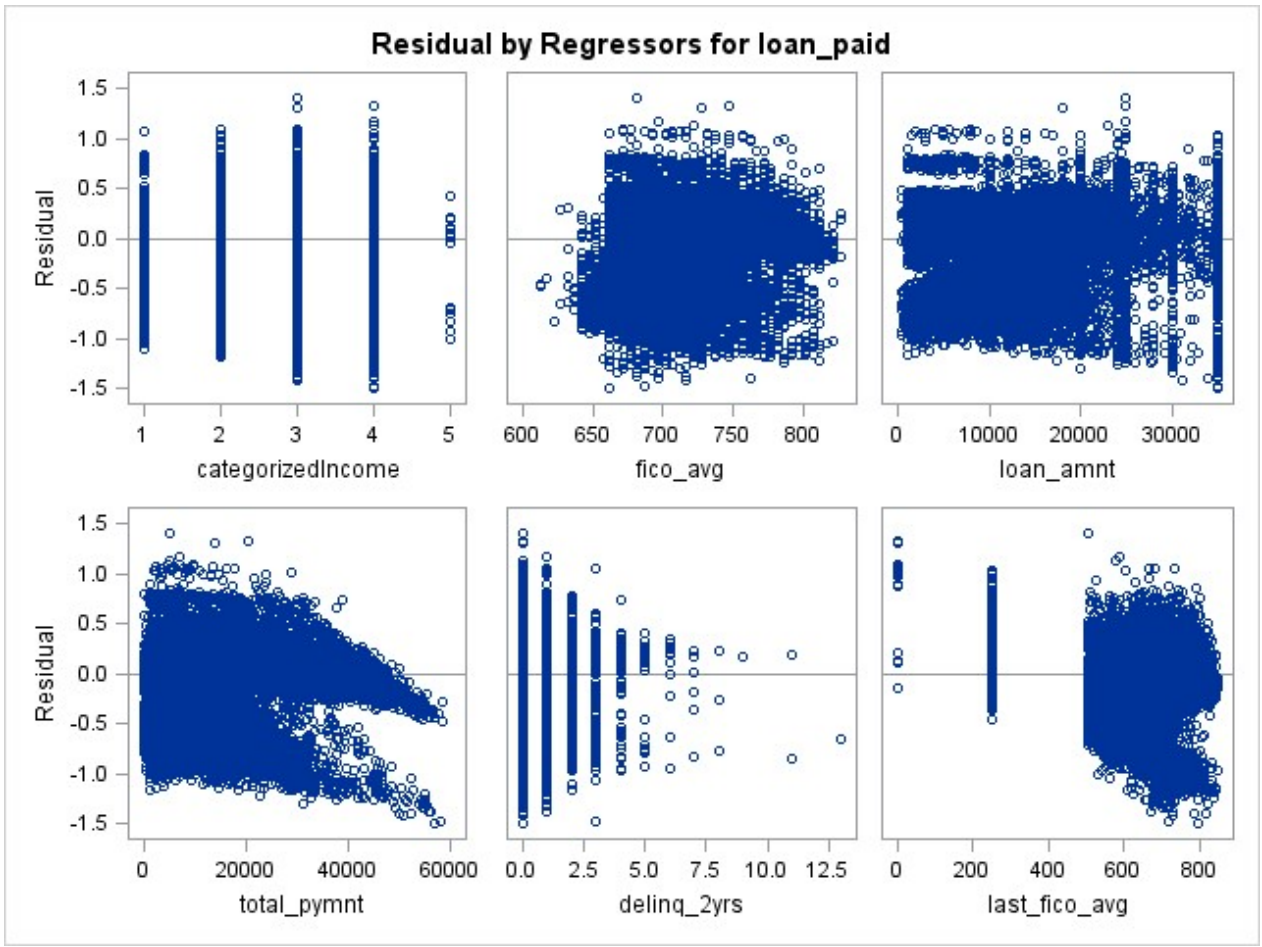
Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-1.13925	0.03256	-34.99	<.0001
categorizedIncome	1	0.01325	0.00202	6.57	<.0001
fico_avg	1	0.00151	0.00004847	31.12	<.0001
loan_amnt	1	-0.00004634	4.802088E-7	-96.49	<.0001
total_pymnt	1	0.00003839	3.894729E-7	98.57	<.0001
delinq_2yrs	1	0.00352	0.00318	1.11	0.2687
last_fico_avg	1	0.00127	0.00001841	69.09	<.0001

Multiple Regression for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid




Residual by Regressors for loan_paid
Residual
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1
2
3
4
5
categorizedIncome
600
650
700
750
800
fico_avg
0
10000
20000
30000
loan_amnt
0
20000
40000
60000
total_pymnt
0.0
2.5
5.0
7.5
10.0
12.5
delinq_2yrs
0
200
400
600
800
last_fico_avg

Multiple Regression(Forward Selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Forward Selection: Step 1

Variable last_fico_avg Entered: R-Square = 0.1927 and C(p) = 10452.70

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1357.90452	1357.90452	10143.1	<.0001
Error	42504	5690.20916	0.13387		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48230	0.01276	191.30899	1429.02	<.0001
last_fico_avg	0.00186	0.00001850	1357.90452	10143.1	<.0001

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable total_pymnt Entered: R-Square = 0.2047 and C(p) = 9664.043

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1442.86370	721.43185	5470.41	<.0001
Error	42503	5605.24998	0.13188		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48131	0.01266	190.52655	1444.71	<.0001
total_pymnt	0.00000502	1.9782E-7	84.95918	644.22	<.0001
last_fico_avg	0.00177	0.00001870	1186.24899	8994.99	<.0001

Bounds on condition number: 1.0371, 4.1483

Forward Selection: Step 3

Variable loan_amnt Entered: R-Square = 0.3362 and C(p) = 1041.560

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2369.60136	789.86712	7175.56	<.0001
Error	42502	4678.51232	0.11008		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.18701	0.01201	26.71054	242.65	<.0001
loan_amnt	-0.00004318	4.706425E-7	926.73766	8418.96	<.0001
total_pymnt	0.00003676	3.903271E-7	976.58189	8871.77	<.0001
last_fico_avg	0.00149	0.00001737	805.48109	7317.40	<.0001

Bounds on condition number: 4.8374, 31.82

Forward Selection: Step 4

Variable fico_avg Entered: R-Square = 0.3514 and C(p) = 48.3177

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2476.54437	619.13609	5755.99	<.0001
Error	42501	4571.56931	0.10756		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.10928	0.03157	132.84175	1235.00	<.0001
fico_avg	0.00150	0.00004748	106.94301	994.23	<.0001
loan_amnt	-0.00004576	4.723744E-7	1009.53920	9385.49	<.0001
total_pymnt	0.00003845	3.895487E-7	1048.19632	9744.88	<.0001
last_fico_avg	0.00128	0.00001836	523.90925	4870.68	<.0001

Bounds on condition number: 4.9307, 48.658

Forward Selection: Step 5

Variable categorizedIncome Entered: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001
last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

Forward Selection: Step 6

Variable delinq_2yrs Entered: R-Square = 0.3521 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13925	0.03256	131.55878	1224.32	<.0001
categorizedIncome	0.01325	0.00202	4.63725	43.16	<.0001
fico_avg	0.00151	0.00004847	104.09608	968.75	<.0001
loan_amnt	-0.00004634	4.802088E-7	1000.52274	9311.15	<.0001
total_pymnt	0.00003839	3.894729E-7	1044.02119	9715.96	<.0001
delinq_2yrs	0.00352	0.00318	0.13146	1.22	0.2687
last_fico_avg	0.00127	0.00001841	512.91934	4773.37	<.0001

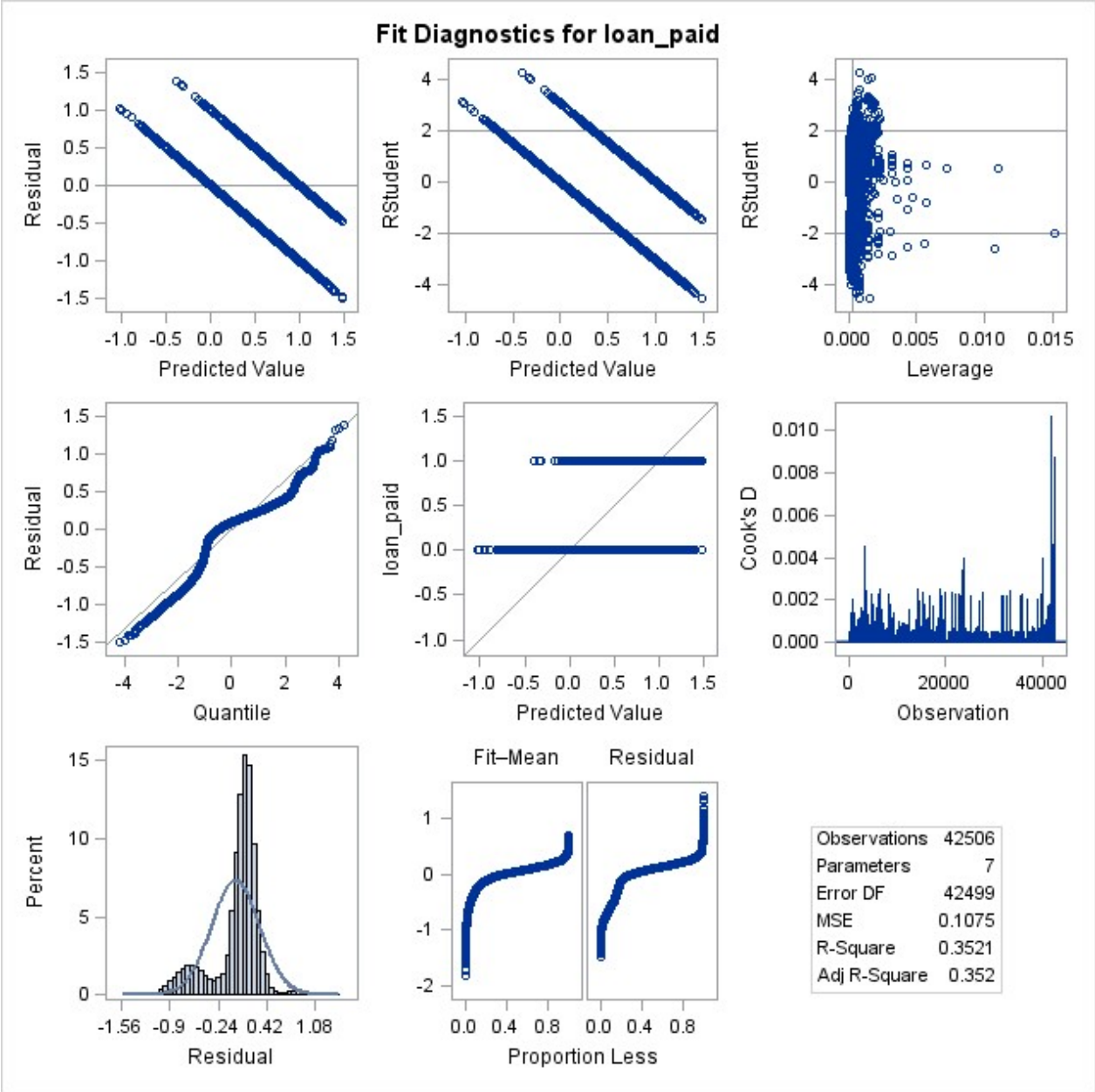
Bounds on condition number: 5.0097, 88.047

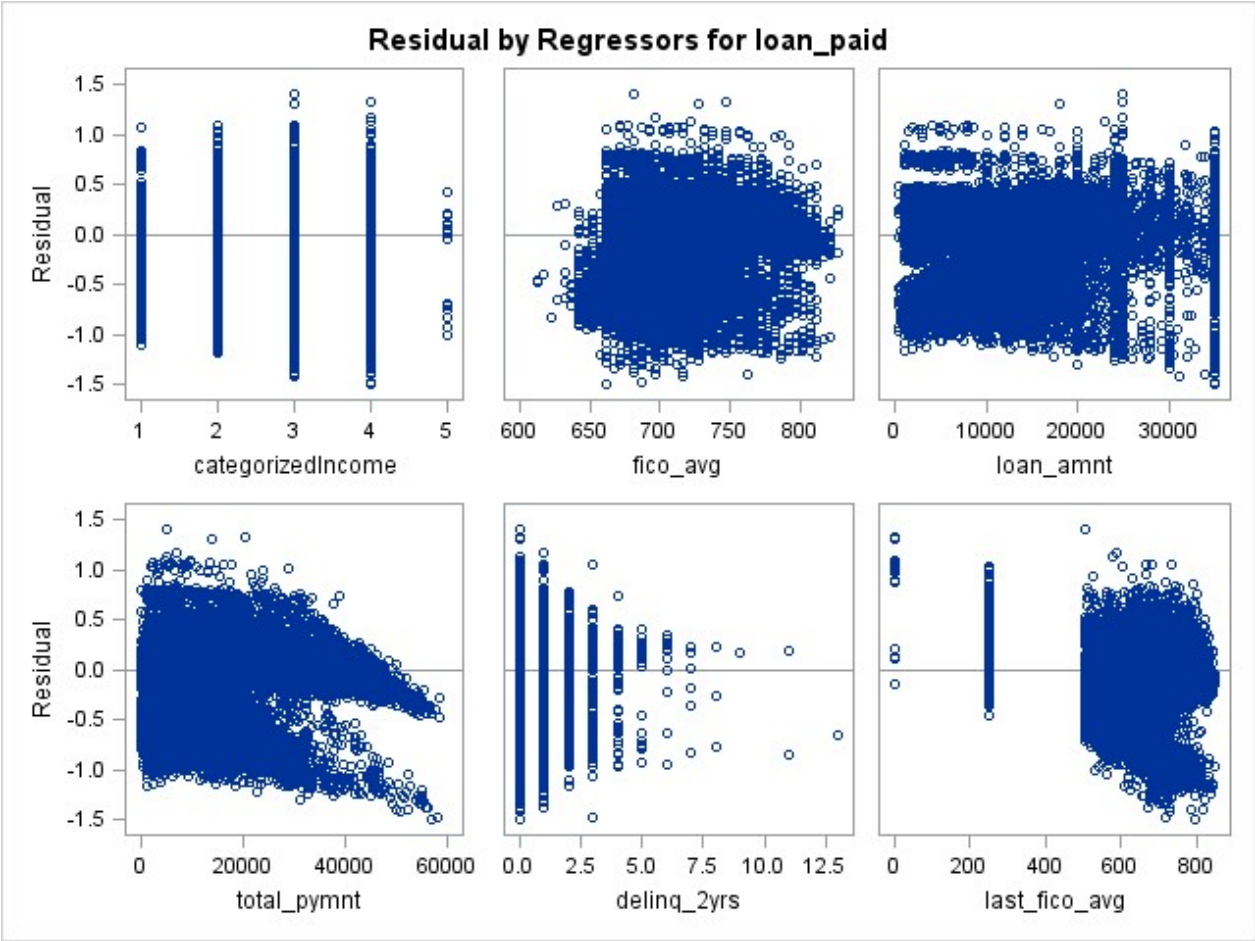
All variables have been entered into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	last_fico_avg	1	0.1927	0.1927	10452.7	10143.1	<.0001
2	total_pymnt	2	0.0121	0.2047	9664.04	644.22	<.0001
3	loan_amnt	3	0.1315	0.3362	1041.56	8418.96	<.0001
4	fico_avg	4	0.0152	0.3514	48.3177	994.23	<.0001
5	categorizedIncome	5	0.0007	0.3520	6.2234	44.09	<.0001
6	delinq_2yrs	6	0.0000	0.3521	7.0000	1.22	0.2687

Multiple Regression(Forward Selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid




Residual by Regressors for loan_paid
Residual
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1
2
3
4
5
categorizedIncome
600
650
700
750
800
fico_avg
0
10000
20000
30000
loan_amnt
0
20000
40000
60000
total_pymnt
0.0
2.5
5.0
7.5
10.0
12.5
delinq_2yrs
0
200
400
600
800
last_fico_avg

Multiple Regression(Backward Elimination) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Backward Elimination: Step 0

All Variables Entered: R-Square = 0.3521 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13925	0.03256	131.55878	1224.32	<.0001
categorizedIncome	0.01325	0.00202	4.63725	43.16	<.0001
fico_avg	0.00151	0.00004847	104.09608	968.75	<.0001
loan_amnt	-0.00004634	4.802088E-7	1000.52274	9311.15	<.0001
total_pymnt	0.00003839	3.894729E-7	1044.02119	9715.96	<.0001
delinq_2yrs	0.00352	0.00318	0.13146	1.22	0.2687
last_fico_avg	0.00127	0.00001841	512.91934	4773.37	<.0001

Bounds on condition number: 5.0097, 88.047

Backward Elimination: Step 1

Variable delinq_2yrs Removed: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001
last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

All variables left in the model are significant at the 0.1000 level.

Summary of Backward Elimination							
Step	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	delinq_2yrs	5	0.0000	0.3520	6.2234	1.22	0.2687

Multiple Regression(Stepwise Selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Stepwise Selection: Step 1

Variable last_fico_avg Entered: R-Square = 0.1927 and C(p) = 10452.70

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1357.90452	1357.90452	10143.1	<.0001
Error	42504	5690.20916	0.13387		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48230	0.01276	191.30899	1429.02	<.0001
last_fico_avg	0.00186	0.00001850	1357.90452	10143.1	<.0001

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable total_pymnt Entered: R-Square = 0.2047 and C(p) = 9664.043

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1442.86370	721.43185	5470.41	<.0001
Error	42503	5605.24998	0.13188		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48131	0.01266	190.52655	1444.71	<.0001
total_pymnt	0.00000502	1.9782E-7	84.95918	644.22	<.0001
last_fico_avg	0.00177	0.00001870	1186.24899	8994.99	<.0001

Bounds on condition number: 1.0371, 4.1483

Stepwise Selection: Step 3

Variable loan_amnt Entered: R-Square = 0.3362 and C(p) = 1041.560

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2369.60136	789.86712	7175.56	<.0001
Error	42502	4678.51232	0.11008		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.18701	0.01201	26.71054	242.65	<.0001
loan_amnt	-0.00004318	4.706425E-7	926.73766	8418.96	<.0001
total_pymnt	0.00003676	3.903271E-7	976.58189	8871.77	<.0001
last_fico_avg	0.00149	0.00001737	805.48109	7317.40	<.0001

Bounds on condition number: 4.8374, 31.82

Stepwise Selection: Step 4

Variable fico_avg Entered: R-Square = 0.3514 and C(p) = 48.3177

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2476.54437	619.13609	5755.99	<.0001
Error	42501	4571.56931	0.10756		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.10928	0.03157	132.84175	1235.00	<.0001
fico_avg	0.00150	0.00004748	106.94301	994.23	<.0001
loan_amnt	-0.00004576	4.723744E-7	1009.53920	9385.49	<.0001
total_pymnt	0.00003845	3.895487E-7	1048.19632	9744.88	<.0001
last_fico_avg	0.00128	0.00001836	523.90925	4870.68	<.0001

Bounds on condition number: 4.9307, 48.658

Stepwise Selection: Step 5

Variable categorizedIncome Entered: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001
last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	last_fico_avg		1	0.1927	0.1927	10452.7	10143.1	<.0001
2	total_pymnt		2	0.0121	0.2047	9664.04	644.22	<.0001
3	loan_amnt		3	0.1315	0.3362	1041.56	8418.96	<.0001
4	fico_avg		4	0.0152	0.3514	48.3177	994.23	<.0001
5	categorizedIncome		5	0.0007	0.3520	6.2234	44.09	<.0001

Multiple Regression(Best subset selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Maximum R-Square Improvement: Step 1

Variable last_fico_avg Entered: R-Square = 0.1927 and C(p) = 10452.70

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1357.90452	1357.90452	10143.1	<.0001
Error	42504	5690.20916	0.13387		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48230	0.01276	191.30899	1429.02	<.0001
last_fico_avg	0.00186	0.00001850	1357.90452	10143.1	<.0001

Bounds on condition number: 1, 1

The above model is the best 1-variable model found.

Maximum R-Square Improvement: Step 2

Variable total_pymnt Entered: R-Square = 0.2047 and C(p) = 9664.043

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1442.86370	721.43185	5470.41	<.0001
Error	42503	5605.24998	0.13188		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48131	0.01266	190.52655	1444.71	<.0001

total_pymnt	0.00000502	1.9782E-7	84.95918	644.22	<.0001
last_fico_avg	0.00177	0.00001870	1186.24899	8994.99	<.0001

Bounds on condition number: 1.0371, 4.1483

Maximum R-Square Improvement: Step 3

Variable last_fico_avg Removed: R-Square = 0.2219 and C(p) = 8535.594
Variable loan_amnt Entered

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1564.12027	782.06014	6061.26	<.0001
Error	42503	5483.99340	0.12903		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.80960	0.00314	8595.48251	66618.2	<.0001
loan_amnt	-0.00005045	5.011724E-7	1307.50557	10133.7	<.0001
total_pymnt	0.00004498	4.095906E-7	1556.19567	12061.1	<.0001

Bounds on condition number: 4.5443, 18.177

The above model is the best 2-variable model found.

Maximum R-Square Improvement: Step 4

Variable last_fico_avg Entered: R-Square = 0.3362 and C(p) = 1041.560

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2369.60136	789.86712	7175.56	<.0001
Error	42502	4678.51232	0.11008		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.18701	0.01201	26.71054	242.65	<.0001
loan_amnt	-0.00004318	4.706425E-7	926.73766	8418.96	<.0001
total_pymnt	0.00003676	3.903271E-7	976.58189	8871.77	<.0001
last_fico_avg	0.00149	0.00001737	805.48109	7317.40	<.0001

Bounds on condition number: 4.8374, 31.82

The above model is the best 3-variable model found.

Maximum R-Square Improvement: Step 5

Variable fico_avg Entered: R-Square = 0.3514 and C(p) = 48.3177

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2476.54437	619.13609	5755.99	<.0001
Error	42501	4571.56931	0.10756		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.10928	0.03157	132.84175	1235.00	<.0001
fico_avg	0.00150	0.00004748	106.94301	994.23	<.0001
loan_amnt	-0.00004576	4.723744E-7	1009.53920	9385.49	<.0001
total_pymnt	0.00003845	3.895487E-7	1048.19632	9744.88	<.0001
last_fico_avg	0.00128	0.00001836	523.90925	4870.68	<.0001

Bounds on condition number: 4.9307, 48.658

The above model is the best 4-variable model found.

Maximum R-Square Improvement: Step 6

Variable categorizedIncome Entered: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001

last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

The above model is the best 5-variable model found.

Maximum R-Square Improvement: Step 7

Variable delinq_2yrs Entered: R-Square = 0.3521 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13925	0.03256	131.55878	1224.32	<.0001
categorizedIncome	0.01325	0.00202	4.63725	43.16	<.0001
fico_avg	0.00151	0.00004847	104.09608	968.75	<.0001
loan_amnt	-0.00004634	4.802088E-7	1000.52274	9311.15	<.0001
total_pymnt	0.00003839	3.894729E-7	1044.02119	9715.96	<.0001
delinq_2yrs	0.00352	0.00318	0.13146	1.22	0.2687
last_fico_avg	0.00127	0.00001841	512.91934	4773.37	<.0001

Bounds on condition number: 5.0097, 88.047

The above model is the best 6-variable model found.

No further improvement in R-Square is possible.

Variance Inflation

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Root MSE	0.32780	R-Square	0.3521
Dependent Mean	0.79015	Adj R-Sq	0.3520
Coeff Var	41.48622		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	-1.13925	0.03256	-34.99	<.0001	0
categorizedIncome	1	0.01325	0.00202	6.57	<.0001	1.22825
fico_avg	1	0.00151	0.00004847	31.12	<.0001	1.21654
loan_amnt	1	-0.00004634	4.802088E-7	-96.49	<.0001	5.00969
total_pymnt	1	0.00003839	3.894729E-7	98.57	<.0001	4.93378
delinq_2yrs	1	0.00352	0.00318	1.11	0.2687	1.05340
last_fico_avg	1	0.00127	0.00001841	69.09	<.0001	1.23290

Variance Inflation

The CONTENTS Procedure

Data Set Name	WORK.LOANDATA	**Observations**	42538
Member Type	DATA	**Variables**	60
Engine	V9	**Indexes**	0
Created	12/05/2017 15:06:11	**Observation Length**	1080
Last Modified	12/05/2017 15:06:11	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information	
Data Set Page Size	90112
Number of Data Set Pages	513
First Data Page	1
Max Obs per Page	83
Obs in First Data Page	76
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	C:\Users\Saloni\AppData\Local\Temp\SAS Temporary Files_TD4224_DESKTOP-IQ5UG81_\loandata.sas7bdat
Release Created	9.0401M2
Host Created	W32_8HOME

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat
53	acc_now_delinq	Num	8	BEST12.	BEST32.
21	addr_state	Char	2	$2.	$2.
12	annual_inc	Num	8	BEST12.	BEST32.
52	annual_inc_joint	Char	1	$1.	$1.
51	application_type	Char	10	$10.	$10.
60	categorizedIncome	Num	8		
44	collection_recovery_fee	Num	8	BEST12.	BEST32.
23	delinq_2yrs	Num	8	BEST12.	BEST32.
54	delinq_amnt	Num	8	BEST12.	BEST32.
17	desc	Char	588	$588.	$588.
22	dti	Num	8	BEST12.	BEST32.
24	earliest_cr_line	Char	6	$6.	$6.
10	emp_length	Char	9	$9.	$9.
9	emp_title	Char	35	$35.	$35.
57	fico_avg	Num	8		

26	fico_range_high	Num	8	BEST12.	BEST32.
25	fico_range_low	Num	8	BEST12.	BEST32.
2	funded_amnt	Num	8	BEST12.	BEST32.
3	funded_amnt_inv	Num	8	BEST12.	BEST32.
7	grade	Char	1	$1.	$1.
11	home_ownership	Char	8	$8.	$8.
35	initial_list_status	Char	1	$1.	$1.
27	inq_last_6mths	Num	8	BEST12.	BEST32.
6	installment	Num	8	BEST12.	BEST32.
5	int_rate	Char	6	$6.	$6.
14	issue_d	Char	6	$6.	$6.
48	last_credit_pull_d	Char	6	$6.	$6.
58	last_fico_avg	Num	8		
49	last_fico_range_high	Num	8	BEST12.	BEST32.
50	last_fico_range_low	Num	8	BEST12.	BEST32.
46	last_pymnt_amnt	Num	8	BEST12.	BEST32.
45	last_pymnt_d	Char	6	$6.	$6.
1	loan_amnt	Num	8	BEST12.	BEST32.
59	loan_paid	Num	8		
15	loan_status	Char	11	$11.	$11.
28	mths_since_last_delinq	Num	8	BEST12.	BEST32.
29	mths_since_last_record	Char	1	$1.	$1.
47	next_pymnt_d	Char	6	$6.	$6.
30	open_acc	Num	8	BEST12.	BEST32.
36	out_prncp	Num	8	BEST12.	BEST32.
37	out_prncp_inv	Num	8	BEST12.	BEST32.
31	pub_rec	Num	8	BEST12.	BEST32.
55	pub_rec_bankruptcies	Num	8	BEST12.	BEST32.
18	purpose	Char	18	$18.	$18.
16	pymnt_plan	Char	1	$1.	$1.
43	recoveries	Num	8	BEST12.	BEST32.
32	revol_bal	Num	8	BEST12.	BEST32.
33	revol_util	Char	6	$6.	$6.
8	sub_grade	Char	2	$2.	$2.
56	tax_liens	Num	8	BEST12.	BEST32.
4	term	Char	9	$9.	$9.
19	title	Char	37	$37.	$37.
34	total_acc	Num	8	BEST12.	BEST32.
38	total_pymnt	Num	8	BEST12.	BEST32.
39	total_pymnt_inv	Num	8	BEST12.	BEST32.
41	total_rec_int	Num	8	BEST12.	BEST32.
42	total_rec_late_fee	Num	8	BEST12.	BEST32.
40	total_rec_prncp	Num	8	BEST12.	BEST32.
13					

	verification_status	Char	15	$15.	$15.
20	zip_code	Char	5	$5.	$5.

Variance Inflation

The FREQ Procedure

categorizedIncome	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	5088	11.96	5088	11.96
2	11878	27.92	16966	39.88
3	19433	45.68	36399	85.57
4	6117	14.38	42516	99.95
5	22	0.05	42538	100.00

Variance Inflation

The FREQ Procedure

loan_paid	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	8952	21.04	8952	21.04
1	33586	78.96	42538	100.00

Variance Inflation

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
loan_amnt	42535	11089.72	7410.94	500.0000000	35000.00
funded_amnt	42535	10821.59	7146.91	500.0000000	35000.00
funded_amnt_inv	42535	10139.83	7131.69	0	35000.00
installment	42535	322.6230627	208.9272165	15.6700000	1305.19
annual_inc	42531	69136.56	64096.35	1896.00	6000000.00
dti	42535	13.3730431	6.7263149	0	29.9900000
delinq_2yrs	42506	0.1524491	0.5124065	0	13.0000000
fico_range_low	42535	713.0525450	36.1884385	610.0000000	825.0000000
fico_range_high	42535	717.0525450	36.1884385	614.0000000	829.0000000
inq_last_6mths	42506	1.0814238	1.5274548	0	33.0000000
mths_since_last_delinq	15609	35.0176180	22.4184269	0	120.0000000
open_acc	42506	9.3439514	4.4962739	1.0000000	47.0000000
pub_rec	42506	0.0581565	0.2457131	0	5.0000000
revol_bal	42535	14297.86	22018.44	0	1207359.00
total_acc	42506	22.1244060	11.5928113	1.0000000	90.0000000
out_prncp	42535	11.8495618	125.3765634	0	3555.85
out_prncp_inv	42535	11.8158909	125.0618595	0	3553.30
total_pymnt	42535	12005.64	9067.24	0	58563.68
total_pymnt_inv	42535	11298.59	9010.32	0	58563.68
total_rec_prncp	42535	9664.27	7087.47	0	35000.02
total_rec_int	42535	2239.55	2583.72	0	23823.75
total_rec_late_fee	42535	1.5153209	7.8277999	0	209.0000000
recoveries	42535	100.3073533	718.0787158	0	29623.35
collection_recovery_fee	42535	13.9567374	159.6218609	0	7002.19
last_pymnt_amnt	42535	2612.72	4385.20	0	36115.20
last_fico_range_high	42535	689.3053015	80.2028819	0	850.0000000
last_fico_range_low	42535	676.3827436	119.0115547	0	845.0000000
acc_now_delinq	42506	0.000094104	0.0097004	0	1.0000000
delinq_amnt	42506	0.1430386	29.3595794	0	6053.00
pub_rec_bankruptcies	41170	0.0452271	0.2087374	0	2.0000000
tax_liens	42430	0.000023568	0.0048547	0	1.0000000
fico_avg	42535	715.0525450	36.1884385	612.0000000	827.0000000
last_fico_avg	42535	682.8440226	95.9865335	0	847.5000000
loan_paid	42538	0.7895529	0.4076310	0	1.0000000
categorizedIncome	42538	2.6263811	0.8740549	1.0000000	5.0000000

Variance Inflation

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Root MSE	0.32780	R-Square	0.3521
Dependent Mean	0.79015	Adj R-Sq	0.3520
Coeff Var	41.48622		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	-1.13925	0.03256	-34.99	<.0001	0
categorizedIncome	1	0.01325	0.00202	6.57	<.0001	1.22825
fico_avg	1	0.00151	0.00004847	31.12	<.0001	1.21654
loan_amnt	1	-0.00004634	4.802088E-7	-96.49	<.0001	5.00969
total_pymnt	1	0.00003839	3.894729E-7	98.57	<.0001	4.93378
delinq_2yrs	1	0.00352	0.00318	1.11	0.2687	1.05340
last_fico_avg	1	0.00127	0.00001841	69.09	<.0001	1.23290

Logistic Regression for Credit Risk Analysis

The LOGISTIC Procedure

Model Information	
Data Set	WORK.LOANDATA
Response Variable	loan_paid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	42538
Number of Observations Used	42506

Response Profile		
Ordered Value	loan_paid	Total Frequency
1	1	33586
2	0	8920

Probability modeled is loan_paid=1.

Note: 32 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information					
Class	Value	Design Variables			
categorizedIncome	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	43677.891	28205.316
SC	43686.548	28291.890
-2 Log L	43675.891	28185.316

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15490.5746	9	<.0001
Score	14982.6229	9	<.0001

Wald	7415.5945	9	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
categorizedIncome	4	43.4539	<.0001
fico_avg	1	1024.3950	<.0001
loan_amnt	1	3765.2344	<.0001
total_pymnt	1	3716.7876	<.0001
delinq_2yrs	1	2.2429	0.1342
last_fico_avg	1	2129.9624	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-17.9456	0.6913	673.8685	<.0001
categorizedIncome	1	1	1.5105	0.5808	6.7641	0.0093
categorizedIncome	2	1	1.6664	0.5801	8.2520	0.0041
categorizedIncome	3	1	1.7877	0.5800	9.5006	0.0021
categorizedIncome	4	1	1.7402	0.5814	8.9580	0.0028
fico_avg		1	0.0172	0.000536	1024.3950	<.0001
loan_amnt		1	-0.00039	6.402E-6	3765.2344	<.0001
total_pymnt		1	0.000342	5.61E-6	3716.7876	<.0001
delinq_2yrs		1	0.0410	0.0274	2.2429	0.1342
last_fico_avg		1	0.00878	0.000190	2129.9624	<.0001

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
categorizedIncome 1 vs 5	4.529	1.451	14.138
categorizedIncome 2 vs 5	5.293	1.698	16.501
categorizedIncome 3 vs 5	5.976	1.917	18.624
categorizedIncome 4 vs 5	5.698	1.823	17.809
fico_avg	1.017	1.016	1.018
loan_amnt	1.000	1.000	1.000
total_pymnt	1.000	1.000	1.000
delinq_2yrs	1.042	0.987	1.099
last_fico_avg	1.009	1.008	1.009

Association of Predicted Probabilities and Observed Responses

Percent Concordant	86.2	Somers' D	0.726
Percent Discordant	13.6	Gamma	0.727
Percent Tied	0.2	Tau-a	0.241
Pairs	299587120	c	0.863

Logistic Regression for Credit Risk Analysis without delinq_2yrs

The LOGISTIC Procedure

Model Information	
Data Set	WORK.LOANDATA
Response Variable	loan_paid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	42538
Number of Observations Used	42535

Response Profile		
Ordered Value	loan_paid	Total Frequency
1	1	33586
2	0	8949

Probability modeled is loan_paid=1.

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information					
Class	Value	Design Variables			
categorizedIncome	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	43768.375	28325.539
SC	43777.033	28403.462
-2 Log L	43766.375	28307.539

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15458.8355	8	<.0001
Score	14950.3404	8	<.0001

Wald	7421.1319	8	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
categorizedIncome	4	60.5332	<.0001
fico_avg	1	1030.3561	<.0001
loan_amnt	1	3755.1808	<.0001
total_pymnt	1	3716.9871	<.0001
last_fico_avg	1	2124.4543	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-18.7370	0.6668	789.5672	<.0001
categorizedIncome	1	1	2.5543	0.5496	21.6023	<.0001
categorizedIncome	2	1	2.7129	0.5489	24.4268	<.0001
categorizedIncome	3	1	2.8339	0.5488	26.6660	<.0001
categorizedIncome	4	1	2.7820	0.5503	25.5544	<.0001
fico_avg		1	0.0168	0.000524	1030.3561	<.0001
loan_amnt		1	-0.00039	6.388E-6	3755.1808	<.0001
total_pymnt		1	0.000342	5.604E-6	3716.9871	<.0001
last_fico_avg		1	0.00874	0.000190	2124.4543	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
categorizedIncome 1 vs 5	12.862	4.381	37.767
categorizedIncome 2 vs 5	15.072	5.140	44.198
categorizedIncome 3 vs 5	17.012	5.803	49.877
categorizedIncome 4 vs 5	16.152	5.492	47.497
fico_avg	1.017	1.016	1.018
loan_amnt	1.000	1.000	1.000
total_pymnt	1.000	1.000	1.000
last_fico_avg	1.009	1.008	1.009

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	86.1	Somers' D	0.724
Percent Discordant	13.7	Gamma	0.726
Percent Tied	0.2	Tau-a	0.241
Pairs	300561114	c	0.862

Multiple Regression for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

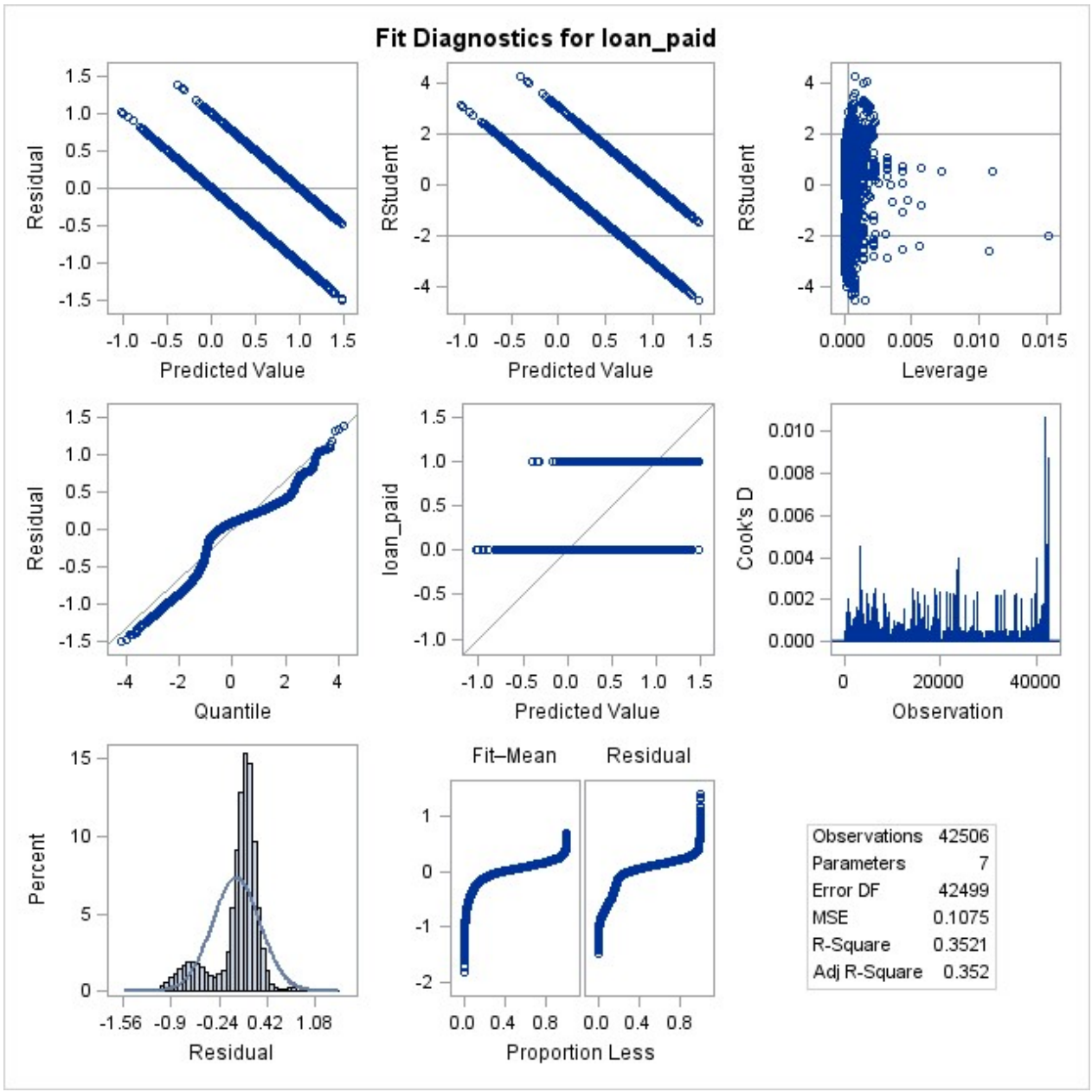
Root MSE	0.32780	R-Square	0.3521
Dependent Mean	0.79015	Adj R-Sq	0.3520
Coeff Var	41.48622		

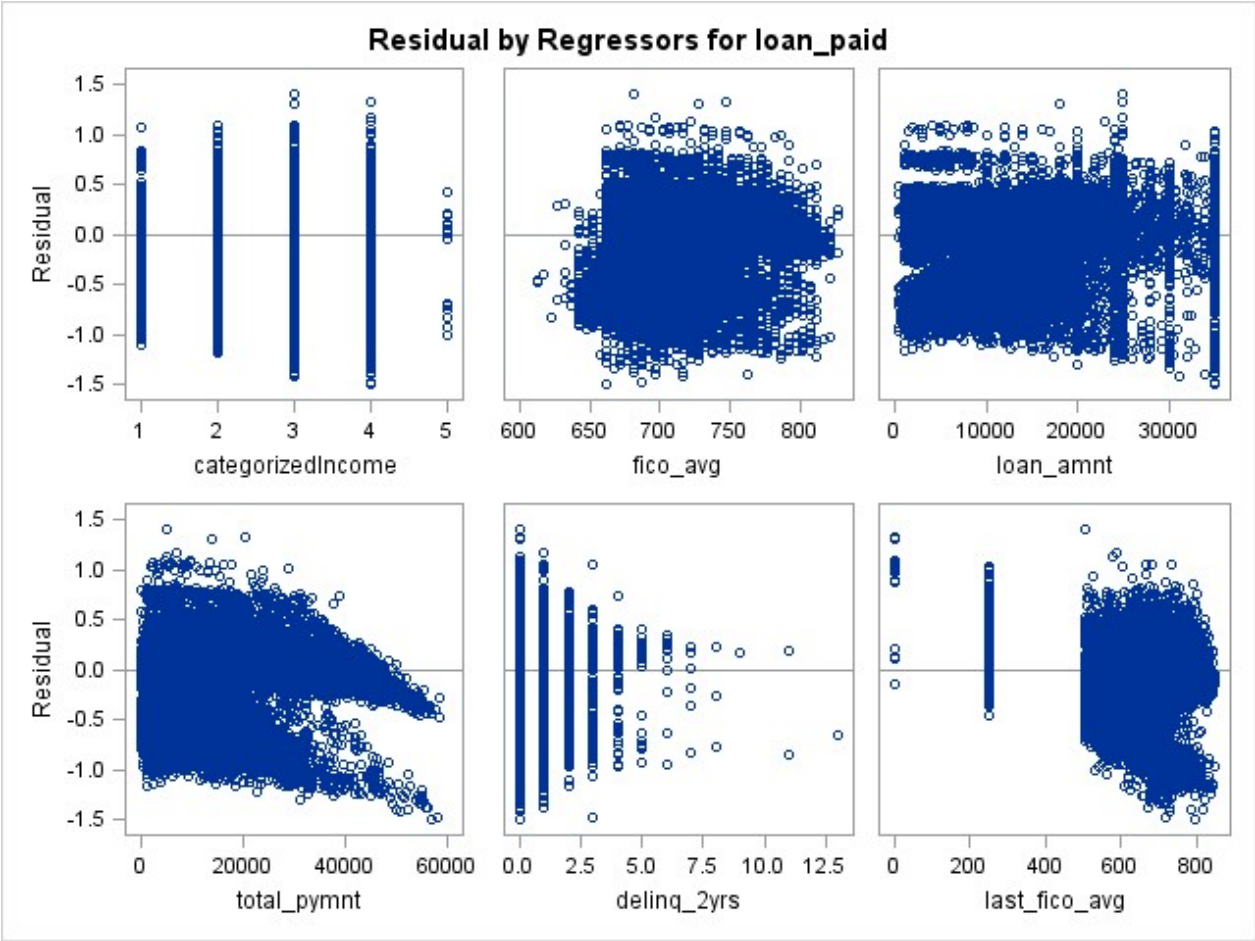
Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-1.13925	0.03256	-34.99	<.0001
categorizedIncome	1	0.01325	0.00202	6.57	<.0001
fico_avg	1	0.00151	0.00004847	31.12	<.0001
loan_amnt	1	-0.00004634	4.802088E-7	-96.49	<.0001
total_pymnt	1	0.00003839	3.894729E-7	98.57	<.0001
delinq_2yrs	1	0.00352	0.00318	1.11	0.2687
last_fico_avg	1	0.00127	0.00001841	69.09	<.0001

Multiple Regression for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid




Residual by Regressors for loan_paid
Residual
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1
2
3
4
5
categorizedIncome
600
650
700
750
800
fico_avg
0
10000
20000
30000
loan_amnt
0
20000
40000
60000
total_pymnt
0.0
2.5
5.0
7.5
10.0
12.5
delinq_2yrs
0
200
400
600
800
last_fico_avg

Multiple Regression(Forward Selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Forward Selection: Step 1

Variable last_fico_avg Entered: R-Square = 0.1927 and C(p) = 10452.70

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1357.90452	1357.90452	10143.1	<.0001
Error	42504	5690.20916	0.13387		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48230	0.01276	191.30899	1429.02	<.0001
last_fico_avg	0.00186	0.00001850	1357.90452	10143.1	<.0001

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable total_pymnt Entered: R-Square = 0.2047 and C(p) = 9664.043

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1442.86370	721.43185	5470.41	<.0001
Error	42503	5605.24998	0.13188		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48131	0.01266	190.52655	1444.71	<.0001
total_pymnt	0.00000502	1.9782E-7	84.95918	644.22	<.0001
last_fico_avg	0.00177	0.00001870	1186.24899	8994.99	<.0001

Bounds on condition number: 1.0371, 4.1483

Forward Selection: Step 3

Variable loan_amnt Entered: R-Square = 0.3362 and C(p) = 1041.560

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2369.60136	789.86712	7175.56	<.0001
Error	42502	4678.51232	0.11008		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.18701	0.01201	26.71054	242.65	<.0001
loan_amnt	-0.00004318	4.706425E-7	926.73766	8418.96	<.0001
total_pymnt	0.00003676	3.903271E-7	976.58189	8871.77	<.0001
last_fico_avg	0.00149	0.00001737	805.48109	7317.40	<.0001

Bounds on condition number: 4.8374, 31.82

Forward Selection: Step 4

Variable fico_avg Entered: R-Square = 0.3514 and C(p) = 48.3177

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2476.54437	619.13609	5755.99	<.0001
Error	42501	4571.56931	0.10756		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.10928	0.03157	132.84175	1235.00	<.0001
fico_avg	0.00150	0.00004748	106.94301	994.23	<.0001
loan_amnt	-0.00004576	4.723744E-7	1009.53920	9385.49	<.0001
total_pymnt	0.00003845	3.895487E-7	1048.19632	9744.88	<.0001
last_fico_avg	0.00128	0.00001836	523.90925	4870.68	<.0001

Bounds on condition number: 4.9307, 48.658

Forward Selection: Step 5

Variable categorizedIncome Entered: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001
last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

Forward Selection: Step 6

Variable delinq_2yrs Entered: R-Square = 0.3521 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13925	0.03256	131.55878	1224.32	<.0001
categorizedIncome	0.01325	0.00202	4.63725	43.16	<.0001
fico_avg	0.00151	0.00004847	104.09608	968.75	<.0001
loan_amnt	-0.00004634	4.802088E-7	1000.52274	9311.15	<.0001
total_pymnt	0.00003839	3.894729E-7	1044.02119	9715.96	<.0001
delinq_2yrs	0.00352	0.00318	0.13146	1.22	0.2687
last_fico_avg	0.00127	0.00001841	512.91934	4773.37	<.0001

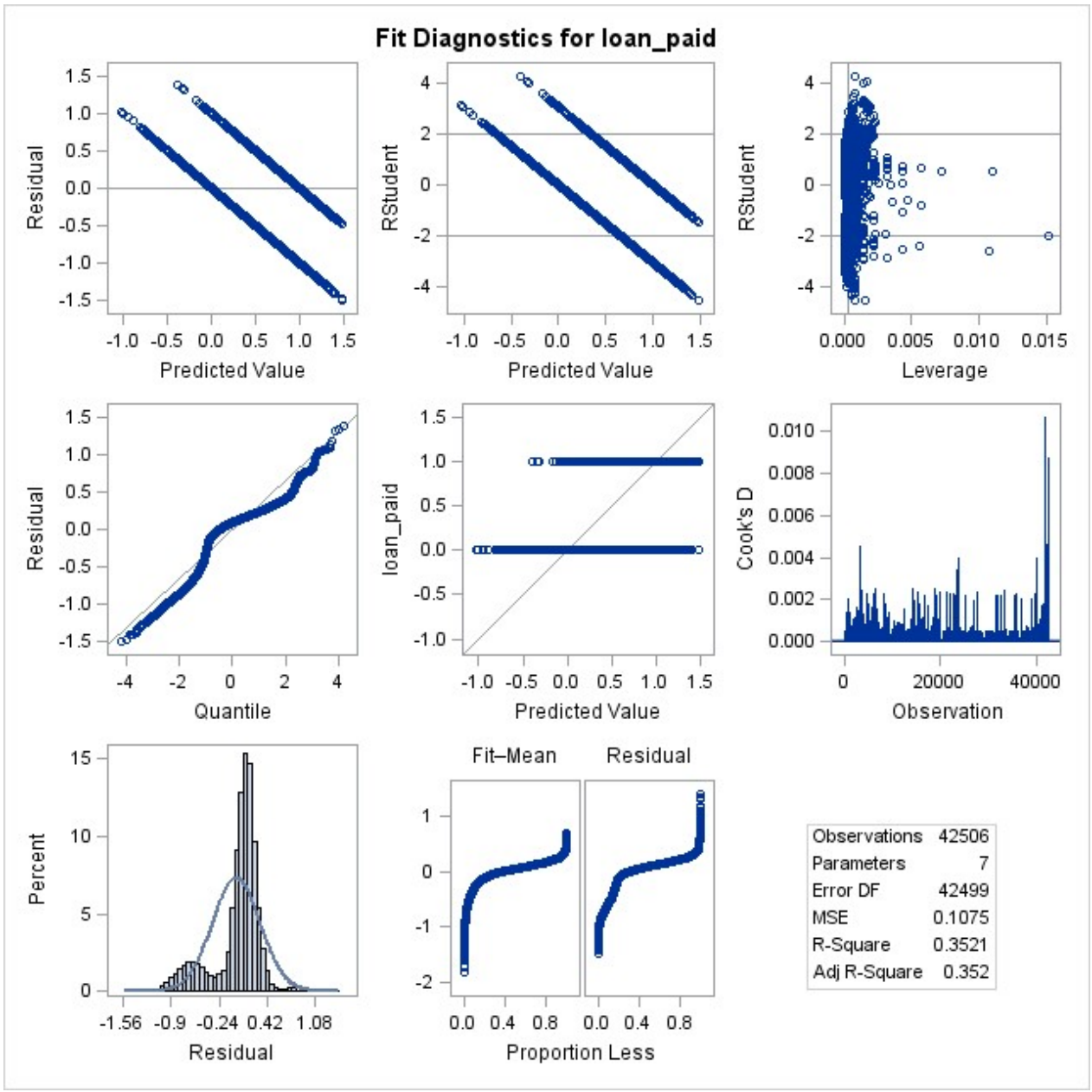
Bounds on condition number: 5.0097, 88.047

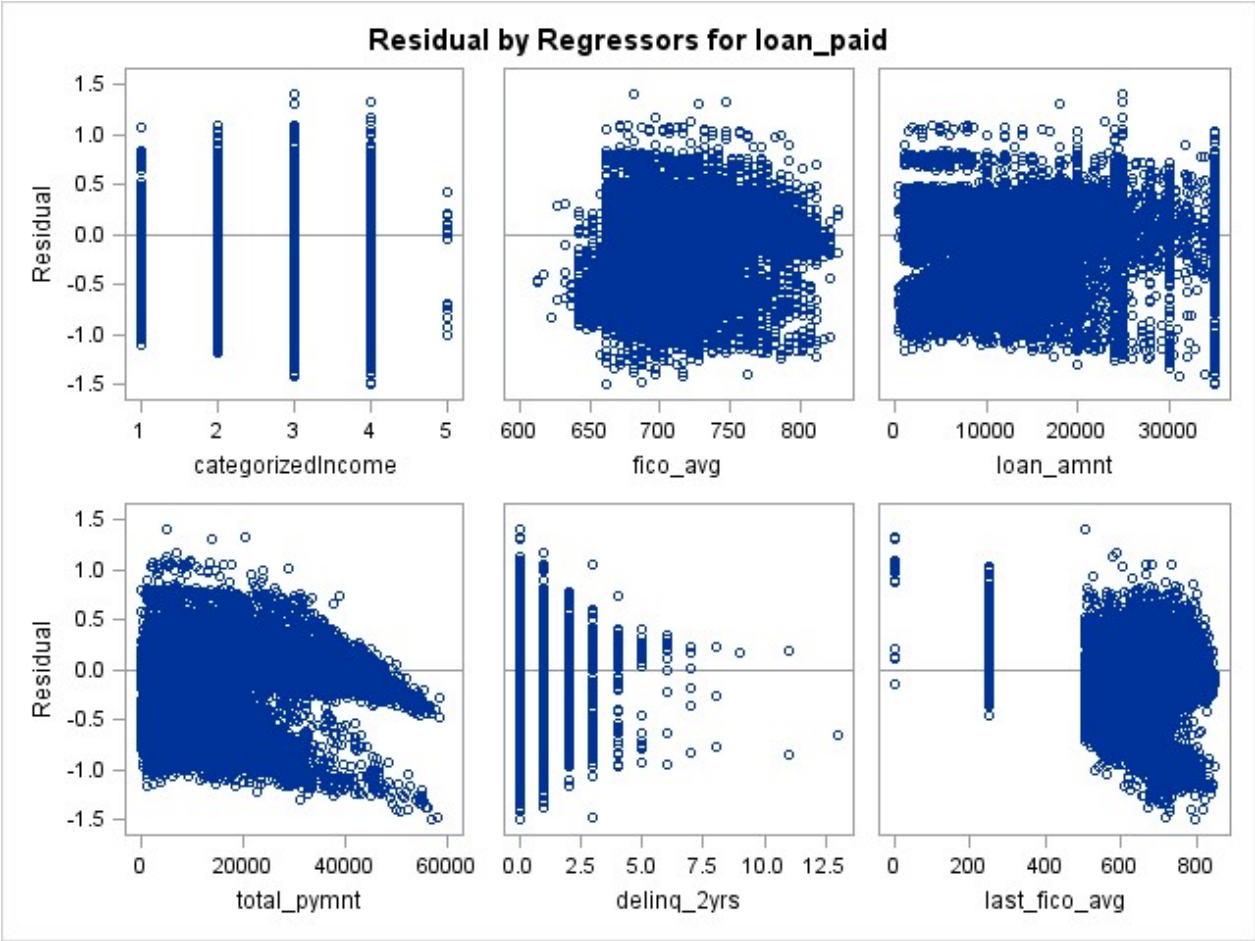
All variables have been entered into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	last_fico_avg	1	0.1927	0.1927	10452.7	10143.1	<.0001
2	total_pymnt	2	0.0121	0.2047	9664.04	644.22	<.0001
3	loan_amnt	3	0.1315	0.3362	1041.56	8418.96	<.0001
4	fico_avg	4	0.0152	0.3514	48.3177	994.23	<.0001
5	categorizedIncome	5	0.0007	0.3520	6.2234	44.09	<.0001
6	delinq_2yrs	6	0.0000	0.3521	7.0000	1.22	0.2687

Multiple Regression(Forward Selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid




Residual by Regressors for loan_paid
Residual
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1
2
3
4
5
categorizedIncome
600
650
700
750
800
fico_avg
0
10000
20000
30000
loan_amnt
0
20000
40000
60000
total_pymnt
0.0
2.5
5.0
7.5
10.0
12.5
delinq_2yrs
0
200
400
600
800
last_fico_avg

Multiple Regression(Backward Elimination) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Backward Elimination: Step 0

All Variables Entered: R-Square = 0.3521 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13925	0.03256	131.55878	1224.32	<.0001
categorizedIncome	0.01325	0.00202	4.63725	43.16	<.0001
fico_avg	0.00151	0.00004847	104.09608	968.75	<.0001
loan_amnt	-0.00004634	4.802088E-7	1000.52274	9311.15	<.0001
total_pymnt	0.00003839	3.894729E-7	1044.02119	9715.96	<.0001
delinq_2yrs	0.00352	0.00318	0.13146	1.22	0.2687
last_fico_avg	0.00127	0.00001841	512.91934	4773.37	<.0001

Bounds on condition number: 5.0097, 88.047

Backward Elimination: Step 1

Variable delinq_2yrs Removed: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001
last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

All variables left in the model are significant at the 0.1000 level.

Summary of Backward Elimination							
Step	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	delinq_2yrs	5	0.0000	0.3520	6.2234	1.22	0.2687

Multiple Regression(Stepwise Selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Stepwise Selection: Step 1

Variable last_fico_avg Entered: R-Square = 0.1927 and C(p) = 10452.70

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1357.90452	1357.90452	10143.1	<.0001
Error	42504	5690.20916	0.13387		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48230	0.01276	191.30899	1429.02	<.0001
last_fico_avg	0.00186	0.00001850	1357.90452	10143.1	<.0001

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable total_pymnt Entered: R-Square = 0.2047 and C(p) = 9664.043

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1442.86370	721.43185	5470.41	<.0001
Error	42503	5605.24998	0.13188		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48131	0.01266	190.52655	1444.71	<.0001
total_pymnt	0.00000502	1.9782E-7	84.95918	644.22	<.0001
last_fico_avg	0.00177	0.00001870	1186.24899	8994.99	<.0001

Bounds on condition number: 1.0371, 4.1483

Stepwise Selection: Step 3

Variable loan_amnt Entered: R-Square = 0.3362 and C(p) = 1041.560

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2369.60136	789.86712	7175.56	<.0001
Error	42502	4678.51232	0.11008		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.18701	0.01201	26.71054	242.65	<.0001
loan_amnt	-0.00004318	4.706425E-7	926.73766	8418.96	<.0001
total_pymnt	0.00003676	3.903271E-7	976.58189	8871.77	<.0001
last_fico_avg	0.00149	0.00001737	805.48109	7317.40	<.0001

Bounds on condition number: 4.8374, 31.82

Stepwise Selection: Step 4

Variable fico_avg Entered: R-Square = 0.3514 and C(p) = 48.3177

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2476.54437	619.13609	5755.99	<.0001
Error	42501	4571.56931	0.10756		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.10928	0.03157	132.84175	1235.00	<.0001
fico_avg	0.00150	0.00004748	106.94301	994.23	<.0001
loan_amnt	-0.00004576	4.723744E-7	1009.53920	9385.49	<.0001
total_pymnt	0.00003845	3.895487E-7	1048.19632	9744.88	<.0001
last_fico_avg	0.00128	0.00001836	523.90925	4870.68	<.0001

Bounds on condition number: 4.9307, 48.658

Stepwise Selection: Step 5

Variable categorizedIncome Entered: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001
last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	last_fico_avg		1	0.1927	0.1927	10452.7	10143.1	<.0001
2	total_pymnt		2	0.0121	0.2047	9664.04	644.22	<.0001
3	loan_amnt		3	0.1315	0.3362	1041.56	8418.96	<.0001
4	fico_avg		4	0.0152	0.3514	48.3177	994.23	<.0001
5	categorizedIncome		5	0.0007	0.3520	6.2234	44.09	<.0001

Multiple Regression(Best subset selection) for Credit Risk Analysis

The REG Procedure
Model: MODEL1
Dependent Variable: loan_paid

Number of Observations Read	42538
Number of Observations Used	42506
Number of Observations with Missing Values	32

Maximum R-Square Improvement: Step 1

Variable last_fico_avg Entered: R-Square = 0.1927 and C(p) = 10452.70

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1357.90452	1357.90452	10143.1	<.0001
Error	42504	5690.20916	0.13387		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48230	0.01276	191.30899	1429.02	<.0001
last_fico_avg	0.00186	0.00001850	1357.90452	10143.1	<.0001

Bounds on condition number: 1, 1

The above model is the best 1-variable model found.

Maximum R-Square Improvement: Step 2

Variable total_pymnt Entered: R-Square = 0.2047 and C(p) = 9664.043

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1442.86370	721.43185	5470.41	<.0001
Error	42503	5605.24998	0.13188		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.48131	0.01266	190.52655	1444.71	<.0001

total_pymnt	0.00000502	1.9782E-7	84.95918	644.22	<.0001
last_fico_avg	0.00177	0.00001870	1186.24899	8994.99	<.0001

Bounds on condition number: 1.0371, 4.1483

Maximum R-Square Improvement: Step 3

Variable last_fico_avg Removed: R-Square = 0.2219 and C(p) = 8535.594
Variable loan_amnt Entered

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1564.12027	782.06014	6061.26	<.0001
Error	42503	5483.99340	0.12903		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.80960	0.00314	8595.48251	66618.2	<.0001
loan_amnt	-0.00005045	5.011724E-7	1307.50557	10133.7	<.0001
total_pymnt	0.00004498	4.095906E-7	1556.19567	12061.1	<.0001

Bounds on condition number: 4.5443, 18.177

The above model is the best 2-variable model found.

Maximum R-Square Improvement: Step 4

Variable last_fico_avg Entered: R-Square = 0.3362 and C(p) = 1041.560

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2369.60136	789.86712	7175.56	<.0001
Error	42502	4678.51232	0.11008		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.18701	0.01201	26.71054	242.65	<.0001
loan_amnt	-0.00004318	4.706425E-7	926.73766	8418.96	<.0001
total_pymnt	0.00003676	3.903271E-7	976.58189	8871.77	<.0001
last_fico_avg	0.00149	0.00001737	805.48109	7317.40	<.0001

Bounds on condition number: 4.8374, 31.82

The above model is the best 3-variable model found.

Maximum R-Square Improvement: Step 5

Variable fico_avg Entered: R-Square = 0.3514 and C(p) = 48.3177

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2476.54437	619.13609	5755.99	<.0001
Error	42501	4571.56931	0.10756		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.10928	0.03157	132.84175	1235.00	<.0001
fico_avg	0.00150	0.00004748	106.94301	994.23	<.0001
loan_amnt	-0.00004576	4.723744E-7	1009.53920	9385.49	<.0001
total_pymnt	0.00003845	3.895487E-7	1048.19632	9744.88	<.0001
last_fico_avg	0.00128	0.00001836	523.90925	4870.68	<.0001

Bounds on condition number: 4.9307, 48.658

The above model is the best 4-variable model found.

Maximum R-Square Improvement: Step 6

Variable categorizedIncome Entered: R-Square = 0.3520 and C(p) = 6.2234

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2481.28249	496.25650	4618.28	<.0001
Error	42500	4566.83119	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13112	0.03172	136.63927	1271.60	<.0001
categorizedIncome	0.01338	0.00201	4.73812	44.09	<.0001
fico_avg	0.00150	0.00004745	107.04142	996.15	<.0001
loan_amnt	-0.00004634	4.80173E-7	1000.96217	9315.19	<.0001
total_pymnt	0.00003839	3.894739E-7	1044.00589	9715.76	<.0001

last_fico_avg	0.00127	0.00001841	512.90037	4773.17	<.0001

Bounds on condition number: 5.0089, 67.83

The above model is the best 5-variable model found.

Maximum R-Square Improvement: Step 7

Variable delinq_2yrs Entered: R-Square = 0.3521 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	2481.41395	413.56899	3848.79	<.0001
Error	42499	4566.69973	0.10745		
Corrected Total	42505	7048.11368			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1.13925	0.03256	131.55878	1224.32	<.0001
categorizedIncome	0.01325	0.00202	4.63725	43.16	<.0001
fico_avg	0.00151	0.00004847	104.09608	968.75	<.0001
loan_amnt	-0.00004634	4.802088E-7	1000.52274	9311.15	<.0001
total_pymnt	0.00003839	3.894729E-7	1044.02119	9715.96	<.0001
delinq_2yrs	0.00352	0.00318	0.13146	1.22	0.2687
last_fico_avg	0.00127	0.00001841	512.91934	4773.37	<.0001

Bounds on condition number: 5.0097, 88.047

The above model is the best 6-variable model found.

No further improvement in R-Square is possible.