```
1  libname cecl "/home/dtiwari40/sasuser.v94/Deepak";
2  proc import datafile="/home/dtiwari40/sasuser.v94/Wintrust/loan1.csv"
3            dbms=csv out=cecl.loans replace;
4  run;
5
6
7
8  proc contents data=cecl.loans;
9  run;
10
11 data loan_data;
12 set  CECL.LOANS ;
13 keep loan_amnt  term int_rate  installment grade sub_grade emp_title emp_length
14      home_ownership  annual_inc verification_status issue_d  loan_status  title
15      dti delinq_2yrs  earliest_cr_line  mths_since_last_delinq out_prncp
16      total_pymnt  total_rec_int total_rec_late_fee  total_rec_prncp  recoveries  last_pymnt_d last_pymnt_amnt
17      next_pymnt_d  mths_since_last_major_derog tot_coll_amt tot_cur_bal avg_cur_bal
18      tax_liens  chargeoff_within_12_mths;
19 run;
20
21
22
23 proc sql;
24 select
25   sum(case when loan_amnt is null then 1 else 0 end) as LA,
26   sum(case when term is null then 1 else 0 end) as TM,
27   sum(case when int_rate is null then 1 else 0 end) as IR,
28   sum(case when installment is null then 1 else 0 end) as IL,
29   sum(case when grade is null then 1 else 0 end) as GD,
30   sum(case when sub_grade is null then 1 else 0 end) as SG,
31   sum(case when emp_title is null then 1 else 0 end) as ET,
32   sum(case when emp_length is null then 1 else 0 end) as EL,
33   sum(case when home_ownership is null then 1 else 0 end) as HO,
34   sum(case when annual_inc is null then 1 else 0 end) as AI,
35   sum(case when verification_status is null then 1 else 0 end) as VS,
36   sum(case when issue_d is null then 1 else 0 end) as ID,
37   sum(case when loan_status is null then 1 else 0 end) as LS,
38   sum(case when title is null then 1 else 0 end) as TL,
39   sum(case when dti is null then 1 else 0 end) as DTI,
40   sum(case when delinq_2yrs is null then 1 else 0 end) as DL,
41   sum(case when earliest_cr_line is null then 1 else 0 end) as ECL,
42   sum(case when mths_since_last_delinq is null then 1 else 0 end) as MLD,
43
```

```
  sum(case when out_prncp is null then 1 else 0 end) as OP,
  sum(case when total_pymnt is null then 1 else 0 end) as TP,
  sum(case when total_rec_int is null then 1 else 0 end) as TI,
  sum(case when total_rec_late_fee is null then 1 else 0 end) as RLF,
  sum(case when total_rec_prncp is null then 1 else 0 end) as TP,
  sum(case when recoveries is null then 1 else 0 end) as RE,
  sum(case when last_pymnt_d is null then 1 else 0 end) as LPD,
  sum(case when last_pymnt_amnt is null then 1 else 0 end) as LPA,
  sum(case when next_pymnt_d is null then 1 else 0 end) as NPD,
  sum(case when mths_since_last_major_derog is null then 1 else 0 end) as MLMD,
  sum(case when tot_coll_amt is null then 1 else 0 end) as TCA,
  sum(case when tot_cur_bal is null then 1 else 0 end) as TCB,
  sum(case when avg_cur_bal is null then 1 else 0 end) as ACB,
  sum(case when tax_liens is null then 1 else 0 end) as TAX,
  sum(case when chargeoff_within_12_mths is null then 1 else 0 end) as CHRG
from loan_data;
run;
quit;

/* Since this has many missing values so, it's better to drop this column */
data loan_data;
    set loan_data;
    drop mths_since_last_major_derog;
run;

/* Drooping missing values as we have enough data to train and fit the model */
data loan_data1;
    set loan_data;
    if nmiss(of _numeric_) + cmiss(of _character_) > 0 then delete;
run;


/* Looking loans by number of times chargeoff with 12 months */
proc sql;
    select chargeoff_within_12_mths as chargeoff,count(chargeoff_within_12_mths) as count
    from loan_data1
    group by chargeoff_within_12_mths;
quit;

/* Looking loans by loan status */
proc sql;
    create table loan_sts as
    select loan_status as ls,count(loan_status) as count
```

```
88      from loan_data1
89      group by loan_status;
90  quit;
91
92  /* Looking loans by month of the Year */
93  proc sql;
94      create table year_loans as
95      select distinct(month(issue_d)) as month ,count(loan_status) as count
96      from loan_data1
97      group by month(issue_d);
98  quit;
99
100 /* Making plot for years and no. of loans */
101 proc sgplot data=year_loans;
102     vbar month / response=count;
103     Title "Plot of number of loans vs month";
104     run;
105
106 /* Seeing for what purpose maximum loans are taken */
107 proc sql;
108     create table loan_purpose as
109     select title as purpose ,count(loan_status) as count
110     from loan_data1
111     group by title;
112 quit;
113
114 /* seeing loans by home ownership */
115 proc sql;
116     create table home_ownership as
117     select home_ownership as home_ownership ,count(loan_status) as count
118     from loan_data1
119     group by home_ownership;
120 quit;
121
122 /* seeing number of loans by employment length */
123
124 data loan_data1;
125     set loan_data1;
126     if emp_length="< 1 year" then job_length="<1";
127     else if emp_length in ("1 year","2 years","3 years") then job_length="1-3";
128     else if emp_length in ("4 years","5 years") then job_length="3-5";
129     else if emp_length in ("6 years","7 years","8 years","9 years") then job_length="5-9";
130     else job_length=">10";
131
```

```
132 run;
133
134 proc sql;
135     create table emplmnt2 as
136     select job_length as job_length,count(loan_status) as count
137     from loan_data1
138     group by job_length;
139 quit;
140
141 /* plotting number of loans vs job tenure */
142 proc sgplot data=emplmnt2;
143     vbar job_length / response=count;
144     Title "Plot of number of loans vs job tenure";
145 run;
146
147 /* Getting number of loans by loan risk grade */
148 proc sql;
149     create table grade_tbl as
150     select grade as grade,count(loan_status) as count
151     from loan_data1
152     group by grade;
153 quit;
154
155 /* plotting number of loans vs risk grade */
156 proc sgplot data=grade_tbl;
157     vbar grade / response=count;
158     Title "Plot of number of loans vs risk grade";
159 run;
160
161 /* Making new column for bad loans */
162 data badloans;
163     set loan_data1;
164     if loan_status in ("Default","In Grac","Late (1","Late (3") then bad_loan_status=1;
165     else bad_loan_status=0;
166 run;
167
168
169 proc sql;
170     create table bad_debt as
171     select *
172     from badloans(drop= emp_length verification_status loan_status sub_grade emp_title earliest_cr_line tax_liens)
173     where bad_loan_status=1;
174     run;
175 quit;
```

```
176
177 /* Getting number of bad loans per year */
178 proc sql;
179     create table year_bad as
180     select distinct(Year(issue_d)) as year,count(loan_amnt) as bad_loans
181     from bad_debt
182     group by Year(issue_d);
183     run;
184     quit;
185 proc sgplot;
186     vbar year / response=bad_loans;
187     title "Plot of bad loans vs year";
188     run;
189 /* Now let's see which grade are having maximum bad loans */
190 proc sql;
191     create table grade_bad as
192     select grade as risk_grade,count(loan_amnt) as bad_loan_grade
193     from bad_debt
194     group by risk_grade;
195     run;
196     quit;
197 proc sgplot;
198     vbar risk_grade / response=bad_loan_grade;
199     title "Plot of bad loans vs risk_grade";
200     run;
201
202
203 /* Let's see what employment_length are having maximum bad loans */
204
205 proc sql;
206     create table tenure_bad as
207     select job_length as job_tenure,count(loan_amnt) as bad_loan_tenure
208     from bad_debt
209     group by job_tenure;
210     run;
211     quit;
212 proc sgplot;
213     vbar job_tenure / response=bad_loan_tenure;
214     title "Plot of bad loans vs job_tenure";
215     run;
216 /* categorizing bad loans by loan purpose */
217 proc sql;
218     create table purpose_bad as
219     select title as purpose,count(loan_amnt) as bad_loan_purpose
```

```
220     from bad_debt
221     group by purpose;
222     run;
223     quit;
224 proc sgplot;
225     vbar purpose / response=bad_loan_purpose;
226     title "Plot of bad loans vs job_purpose";
227     run;
228 
229 /* Getting default rate by loan purpose */
230 proc sql;
231     create table default_count_purpose as
232     select a.purpose as loan_purpose, a.count as total_loans,b.bad_loan_purpose as default
233     from loan_purpose as a
234     inner join purpose_bad as b
235     on a.purpose=b.purpose;
236     run;
237     quit;
238 
239 data default_rate_purpose;
240     set default_count_purpose;
241     default_rate_percentage=(default/total_loans)*100;
242     run;
243 
244 proc sgplot;
245     vbar loan_purpose / response=default_rate_percentage;
246     title "Plot of default rate percentage vs loan purpose";
247     run;
248 
249 /* Now segregate the loans and calculate loss given default */
250 
251 data LGD;
252 set bad_debt;
253 loss_given_default=(tot_cur_bal-tot_coll_amt-recoveries-total_rec_late_fee);
254 run;
255 
256 proc sql;
257     create table loss_default as
258     select distinct(year(issue_d)) as year,sum(loan_amnt) as loan_amnt,sum(loss_given_default) as lgd
259     from LGD
260     group by year(issue_d);
261     run;
262     quit;
263 
```

```
264
265 /* Modelling begins here */
266 /* dropping columns and creating dummies for grades */
267
268 data modelling_data;
269     set badloans;
270     drop term last_pymnt_d next_pymnt_d sub_grade emp_title verification_status issue_d title
271     earliest_cr_line job_length emp_length home_ownership loan_status ;
272     if grade="A" then grade=0;
273     else if grade="B" then grade=1;
274     else if grade="C" then grade=2;
275     else if grade="D" then grade=3;
276     else if grade="E" then grade=4;
277     else if grade="F" then grade=5;
278     else if grade="G" then grade=6;
279     grade1=input(grade,1.);
280     drop grade;
281     run;
282
283
284 proc sql;
285 select grade1,count(grade1)
286  from modelling_data
287  group by grade1;
288  run;
289  quit;
290
291 proc contents data=modelling_data;
292 run;
293
294 /* Standarizing data */
295
296 PROC STANDARD data=modelling_data mean=0 std=1 out=scaled_data;
297 var annual_inc avg_cur_bal chargeoff_within_12_mths
298      delinq_2yrs  dti  grade1  installment  int_rate  last_pymnt_amnt  loan_amnt  mths_since_last_delinq
299       out_prncp  recoveries  tax_liens tot_coll_amt  tot_cur_bal  total_pymnt  total_rec_int
300       total_rec_late_fee  total_rec_prncp;
301 run;
302
303 /* Using logistic regression on loan level to calculate PD model. Gives out probability of default
304 for each loan, bad loan status 0 means good loans and 1 means loan which is going to default
305  */
306
307 proc logistic data=scaled_data outest=betas covout;
```

```
    Model bad_loan_status(event='1')=  annual_inc avg_cur_bal chargeoff_within_12_mths
     delinq_2yrs  dti  grade1  installment  int_rate  last_pymnt_amnt  loan_amnt  mths_since_last_delinq
      out_prncp  recoveries  tax_liens tot_coll_amt  tot_cur_bal  total_pymnt  total_rec_int
      total_rec_late_fee  total_rec_prncp
                / selection=stepwise
                  slentry=0.3
                  slstay=0.35
                  details
                  lackfit;
   output out=pred p=phat lower=lcl upper=ucl
          predprob=(individual crossvalidate);
   ods output Association=Association;
run;
proc print data=betas;
   title2 'Parameter Estimates and Covariance Matrix';
run;
```

The CONTENTS Procedure

Data Set Name	CECL.LOANS	**Observations**	1500000
Member Type	DATA	**Variables**	146
Engine	V9	**Indexes**	0
Created	11/20/2019 16:14:43	**Observation Length**	1008
Last Modified	11/20/2019 16:14:43	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	11629
First Data Page	1
Max Obs per Page	129
Obs in First Data Page	108
Number of Data Set Repairs	0
Filename	/home/dtiwari40/sasuser.v94/Deepak/loans.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	13497657731
Access Permission	rw-r--r--
Owner Name	dtiwari40
File Size	1GB
File Size (bytes)	1524367360

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	**Variable**	**Type**	**Len**	**Format**	**Informat**
1	VAR1	Num	8	BEST12.	BEST32.
58	acc_now_delinq	Num	8	BEST12.	BEST32.
76	acc_open_past_24mths	Num	8	BEST12.	BEST32.
25	addr_state	Char	2	$2.	$2.
71	all_util	Num	8	BEST12.	BEST32.
15	annual_inc	Num	8	BEST12.	BEST32.
55	annual_inc_joint	Num	8	BEST12.	BEST32.
54	application_type	Char	10	$10.	$10.
77	avg_cur_bal	Num	8	BEST12.	BEST32.
78	bc_open_to_buy	Num	8	BEST12.	BEST32.
79	bc_util	Num	8	BEST12.	BEST32.
80	chargeoff_within_12_mths	Num	8	BEST12.	BEST32.
46	collection_recovery_fee	Num	8	BEST12.	BEST32.
51	collections_12_mths_ex_med	Num	8	BEST12.	BEST32.
140	debt_settlement_flag	Char	1	$1.	$1.
141	debt_settlement_flag_date	Char	1	$1.	$1.
128	deferral_term	Char	1	$1.	$1.
27	delinq_2yrs	Num	8	BEST12.	BEST32.
81	delinq_amnt	Num	8	BEST12.	BEST32.
21	desc	Char	1	$1.	$1.
139	disbursement_method	Char	9	$9.	$9.
26	dti	Num	8	BEST12.	BEST32.
56	dti_joint	Num	8	BEST12.	BEST32.
28	earliest_cr_line	Num	8	MONYY7.	MONYY7.
13	emp_length	Char	9	$9.	$9.
12	emp_title	Char	27	$27.	$27.
5	funded_amnt	Num	8	BEST12.	BEST32.
6	funded_amnt_inv	Num	8	BEST12.	BEST32.
10	grade	Char	1	$1.	$1.
129	hardship_amount	Char	1	$1.	$1.
134	hardship_dpd	Char	1	$1.	$1.
131	hardship_end_date	Char	1	$1.	$1.
124	hardship_flag	Char	1	$1.	$1.
138	hardship_last_payment_amount	Char	1	$1.	$1.
133	hardship_length	Char	1	$1.	$1.
135	hardship_loan_status	Char	1	$1.	$1.

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
137	hardship_payoff_balance_amount	Char	1	$1.	$1.
126	hardship_reason	Char	1	$1.	$1.
130	hardship_start_date	Char	1	$1.	$1.
127	hardship_status	Char	1	$1.	$1.
125	hardship_type	Char	1	$1.	$1.
14	home_ownership	Char	8	$8.	$8.
2	id	Char	1	$1.	$1.
67	il_util	Num	8	BEST12.	BEST32.
37	initial_list_status	Char	1	$1.	$1.
73	inq_fi	Num	8	BEST12.	BEST32.
75	inq_last_12m	Num	8	BEST12.	BEST32.
29	inq_last_6mths	Num	8	BEST12.	BEST32.
9	installment	Num	8	BEST12.	BEST32.
8	int_rate	Num	8	BEST12.	BEST32.
17	issue_d	Num	8	MONYY7.	MONYY7.
50	last_credit_pull_d	Num	8	MONYY7.	MONYY7.
48	last_pymnt_amnt	Num	8	BEST12.	BEST32.
47	last_pymnt_d	Num	8	MONYY7.	MONYY7.
4	loan_amnt	Num	8	BEST12.	BEST32.
18	loan_status	Char	7	$7.	$7.
70	max_bal_bc	Num	8	BEST12.	BEST32.
3	member_id	Char	1	$1.	$1.
82	mo_sin_old_il_acct	Num	8	BEST12.	BEST32.
83	mo_sin_old_rev_tl_op	Num	8	BEST12.	BEST32.
84	mo_sin_rcnt_rev_tl_op	Num	8	BEST12.	BEST32.
85	mo_sin_rcnt_tl	Num	8	BEST12.	BEST32.
86	mort_acc	Num	8	BEST12.	BEST32.
30	mths_since_last_delinq	Num	8	BEST12.	BEST32.
52	mths_since_last_major_derog	Num	8	BEST12.	BEST32.
31	mths_since_last_record	Num	8	BEST12.	BEST32.
65	mths_since_rcnt_il	Num	8	BEST12.	BEST32.
87	mths_since_recent_bc	Num	8	BEST12.	BEST32.
88	mths_since_recent_bc_dlq	Num	8	BEST12.	BEST32.
89	mths_since_recent_inq	Num	8	BEST12.	BEST32.
90	mths_since_recent_revol_delinq	Num	8	BEST12.	BEST32.
49	next_pymnt_d	Num	8	MONYY7.	MONYY7.

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	**Variable**	**Type**	**Len**	**Format**	**Informat**
91	num_accts_ever_120_pd	Num	8	BEST12.	BEST32.
92	num_actv_bc_tl	Num	8	BEST12.	BEST32.
93	num_actv_rev_tl	Num	8	BEST12.	BEST32.
94	num_bc_sats	Num	8	BEST12.	BEST32.
95	num_bc_tl	Num	8	BEST12.	BEST32.
96	num_il_tl	Num	8	BEST12.	BEST32.
97	num_op_rev_tl	Num	8	BEST12.	BEST32.
98	num_rev_accts	Num	8	BEST12.	BEST32.
99	num_rev_tl_bal_gt_0	Num	8	BEST12.	BEST32.
100	num_sats	Num	8	BEST12.	BEST32.
101	num_tl_120dpd_2m	Num	8	BEST12.	BEST32.
102	num_tl_30dpd	Num	8	BEST12.	BEST32.
103	num_tl_90g_dpd_24m	Num	8	BEST12.	BEST32.
104	num_tl_op_past_12m	Num	8	BEST12.	BEST32.
32	open_acc	Num	8	BEST12.	BEST32.
61	open_acc_6m	Num	8	BEST12.	BEST32.
62	open_act_il	Num	8	BEST12.	BEST32.
63	open_il_12m	Num	8	BEST12.	BEST32.
64	open_il_24m	Num	8	BEST12.	BEST32.
68	open_rv_12m	Num	8	BEST12.	BEST32.
69	open_rv_24m	Num	8	BEST12.	BEST32.
136	orig_projected_additional_accrue	Char	1	$1.	$1.
38	out_prncp	Num	8	BEST12.	BEST32.
39	out_prncp_inv	Num	8	BEST12.	BEST32.
132	payment_plan_start_date	Char	1	$1.	$1.
105	pct_tl_nvr_dlq	Num	8	BEST12.	BEST32.
106	percent_bc_gt_75	Num	8	BEST12.	BEST32.
53	policy_code	Num	8	BEST12.	BEST32.
33	pub_rec	Num	8	BEST12.	BEST32.
107	pub_rec_bankruptcies	Num	8	BEST12.	BEST32.
22	purpose	Char	18	$18.	$18.
19	pymnt_plan	Char	1	$1.	$1.
45	recoveries	Num	8	BEST12.	BEST32.
34	revol_bal	Num	8	BEST12.	BEST32.
113	revol_bal_joint	Num	8	BEST12.	BEST32.
35	revol_util	Num	8	BEST12.	BEST32.

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	**Variable**	**Type**	**Len**	**Format**	**Informat**
121	sec_app_chargeoff_within_12_mths	Num	8	BEST12.	BEST32.
122	sec_app_collections_12_mths_ex_m	Num	8	BEST12.	BEST32.
114	sec_app_earliest_cr_line	Num	8	MONYY7.	MONYY7.
115	sec_app_inq_last_6mths	Num	8	BEST12.	BEST32.
116	sec_app_mort_acc	Num	8	BEST12.	BEST32.
123	sec_app_mths_since_last_major_de	Char	1	$1.	$1.
120	sec_app_num_rev_accts	Num	8	BEST12.	BEST32.
117	sec_app_open_acc	Num	8	BEST12.	BEST32.
119	sec_app_open_act_il	Num	8	BEST12.	BEST32.
118	sec_app_revol_util	Num	8	BEST12.	BEST32.
144	settlement_amount	Char	1	$1.	$1.
143	settlement_date	Char	1	$1.	$1.
145	settlement_percentage	Char	1	$1.	$1.
142	settlement_status	Char	1	$1.	$1.
146	settlement_term	Char	1	$1.	$1.
11	sub_grade	Char	2	$2.	$2.
108	tax_liens	Num	8	BEST12.	BEST32.
7	term	Char	9	$9.	$9.
23	title	Char	23	$23.	$23.
59	tot_coll_amt	Num	8	BEST12.	BEST32.
60	tot_cur_bal	Num	8	BEST12.	BEST32.
109	tot_hi_cred_lim	Num	8	BEST12.	BEST32.
36	total_acc	Num	8	BEST12.	BEST32.
110	total_bal_ex_mort	Num	8	BEST12.	BEST32.
66	total_bal_il	Num	8	BEST12.	BEST32.
111	total_bc_limit	Num	8	BEST12.	BEST32.
74	total_cu_tl	Num	8	BEST12.	BEST32.
112	total_il_high_credit_limit	Num	8	BEST12.	BEST32.
40	total_pymnt	Num	8	BEST12.	BEST32.
41	total_pymnt_inv	Num	8	BEST12.	BEST32.
43	total_rec_int	Num	8	BEST12.	BEST32.
44	total_rec_late_fee	Num	8	BEST12.	BEST32.
42	total_rec_prncp	Num	8	BEST12.	BEST32.
72	total_rev_hi_lim	Num	8	BEST12.	BEST32.
20	url	Char	1	$1.	$1.
16	verification_status	Char	15	$15.	$15.

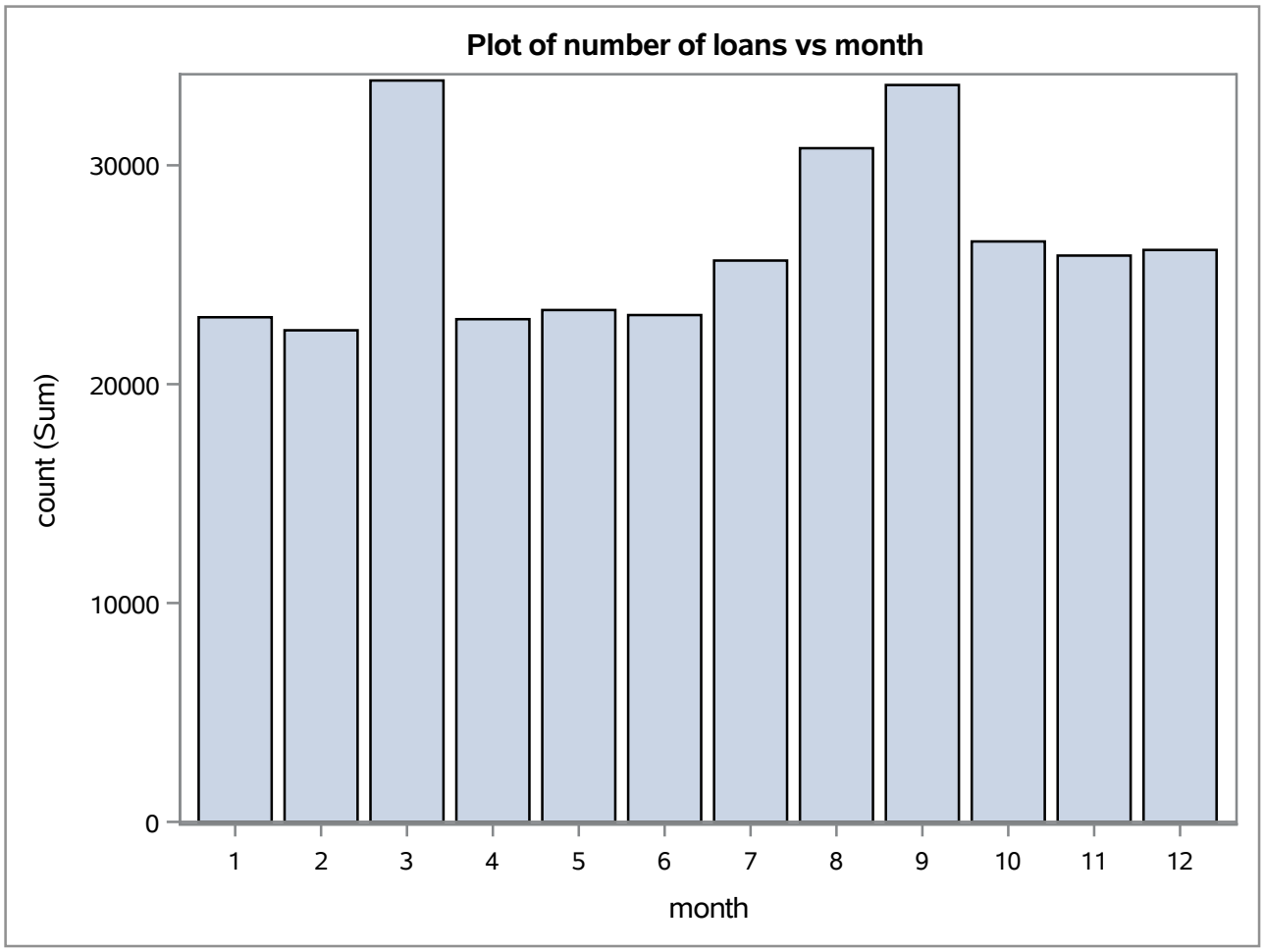
The CONTENTS Procedure

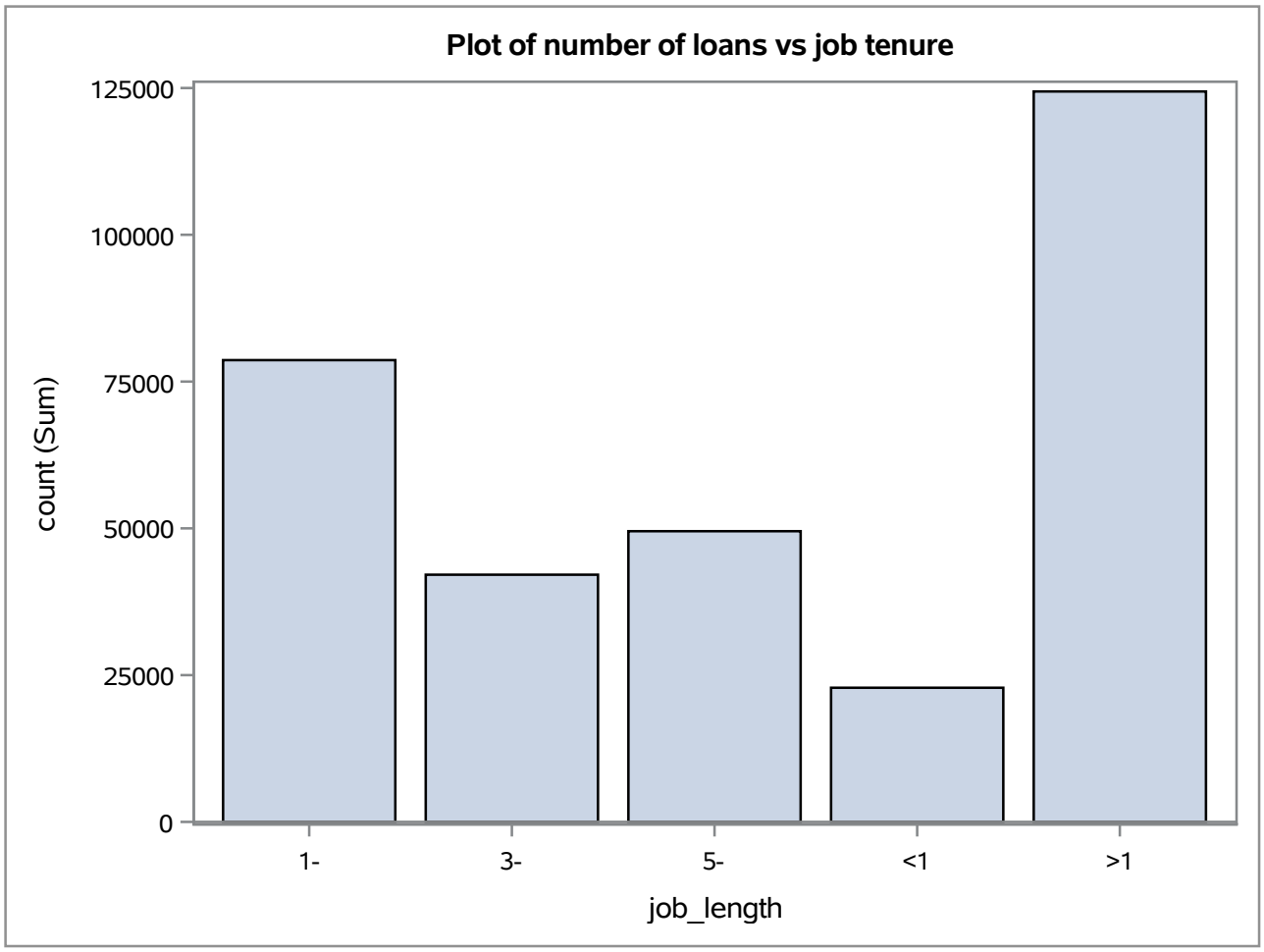
Alphabetic List of Variables and Attributes					
#	**Variable**	**Type**	**Len**	**Format**	**Informat**
57	verification_status_joint	Char	12	$12.	$12.
24	zip_code	Char	5	$5.	$5.

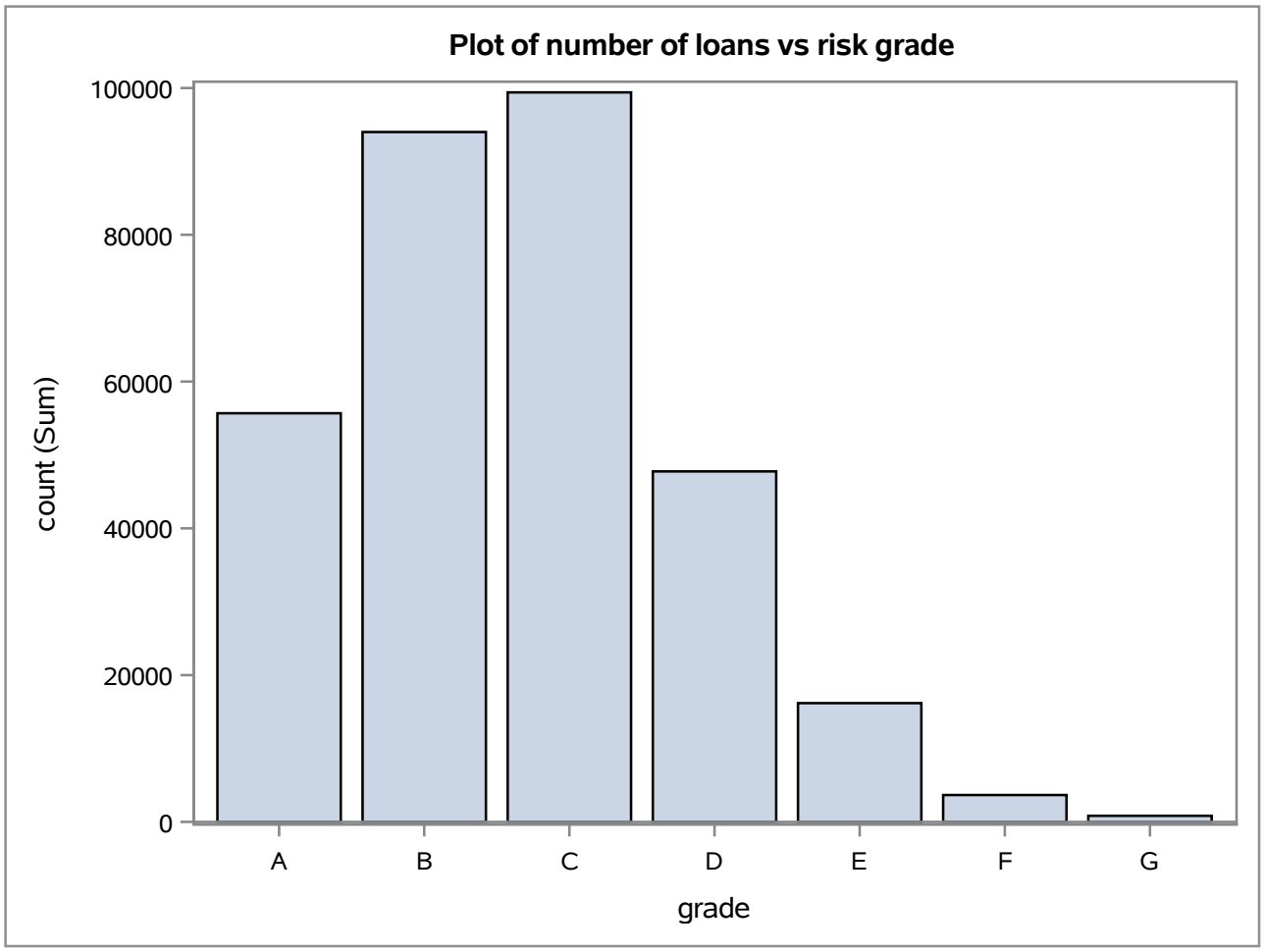
LA	TM	IR	IL	GD	SG	ET	EL	HO	AI	VS	ID	LS	TL	DTI	DL	ECL	MLD	OP	TP	TI	RLF	TP	RE
0	0	0	0	0	0	118056	104867	0	0	0	0	0	23305	1332	0	0	758292	0	0	0	0	0	0

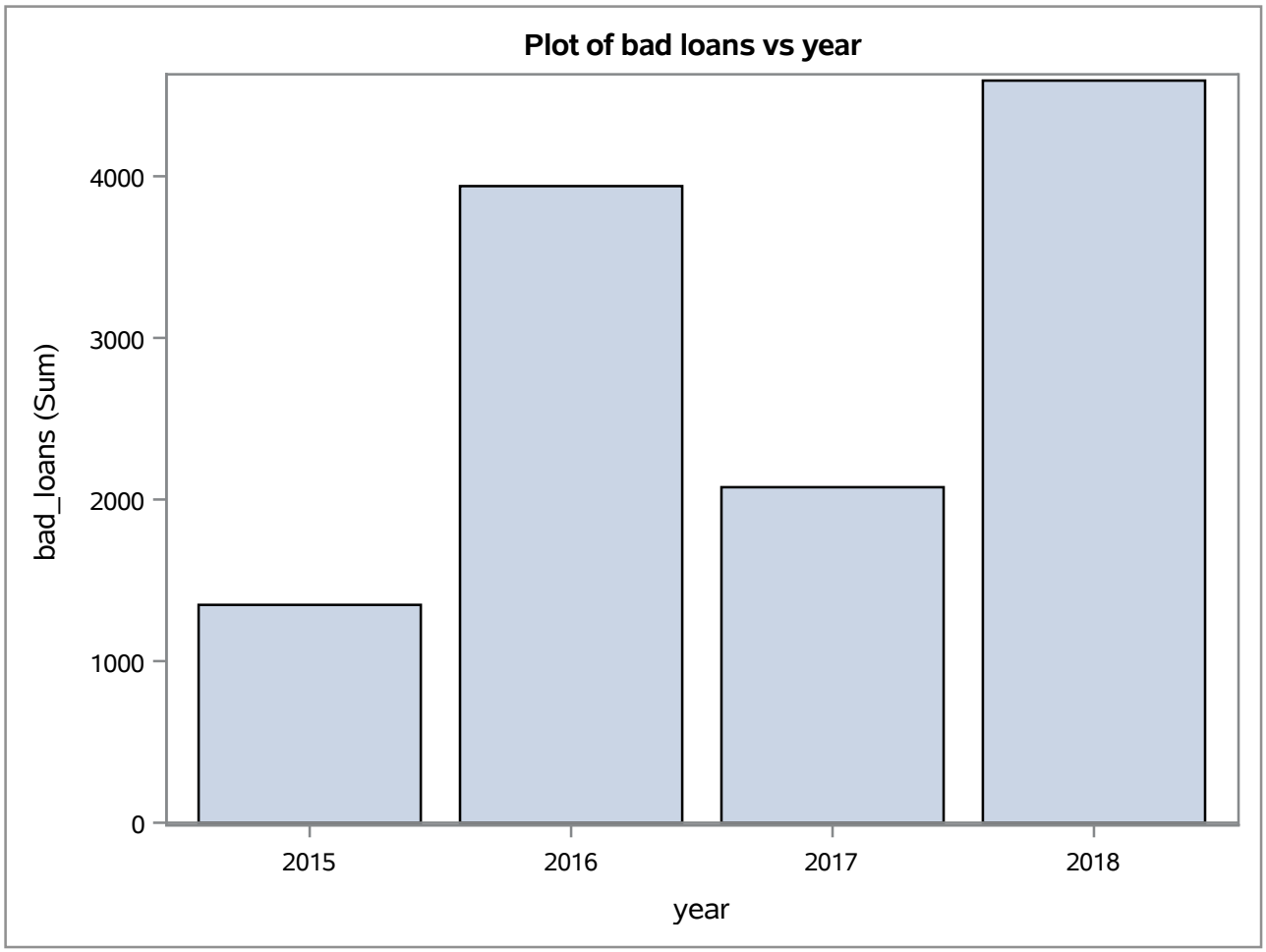
LPD	LPA	NPD	MLMD	TCA	TCB	ACB	TAX	CHRG
1643	0	756514	1095931	0	0	42	0	0

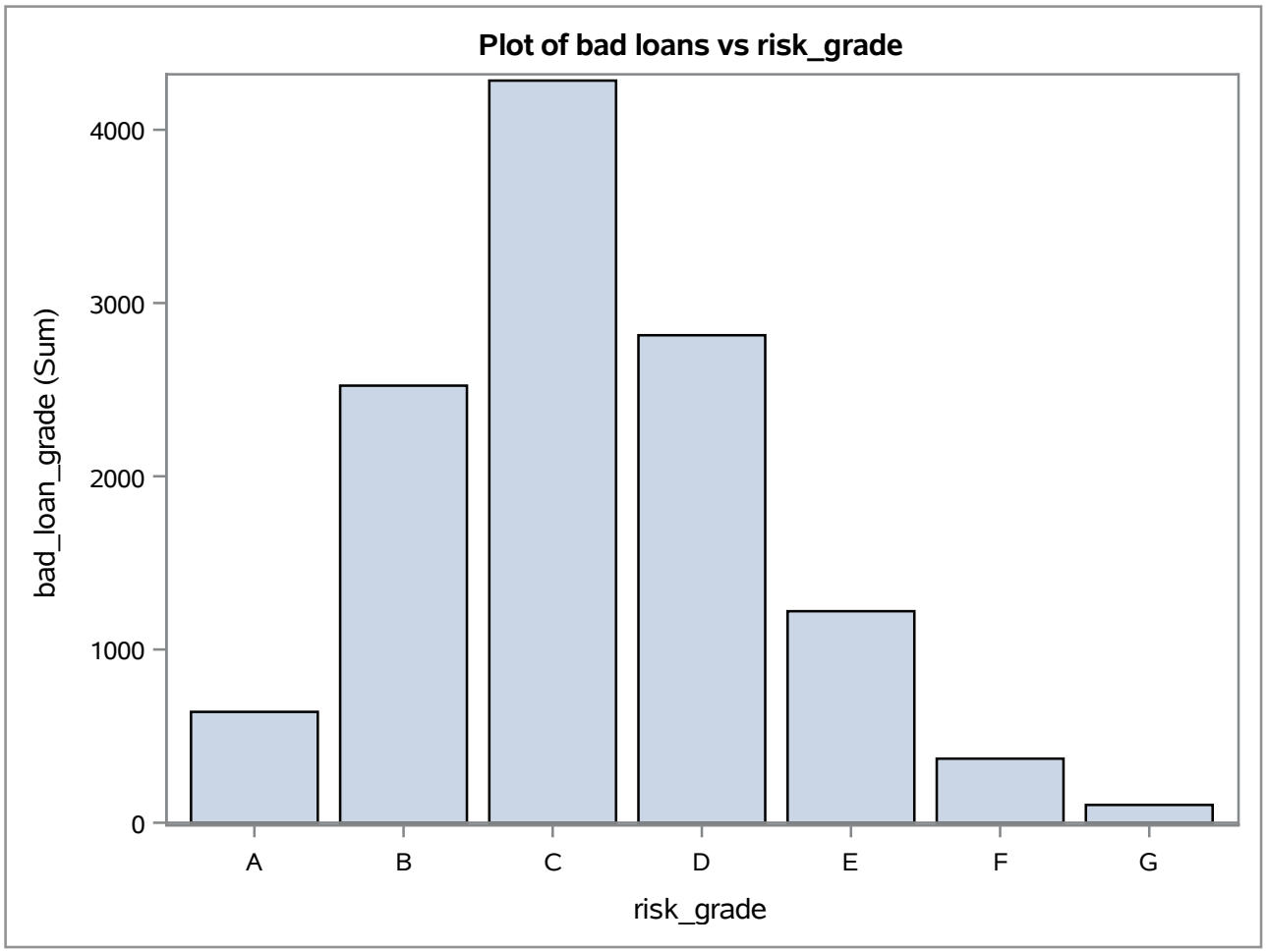
chargeoff	count
0	313247
1	3983
2	249
3	33
4	12
5	3
6	3
7	1
9	1

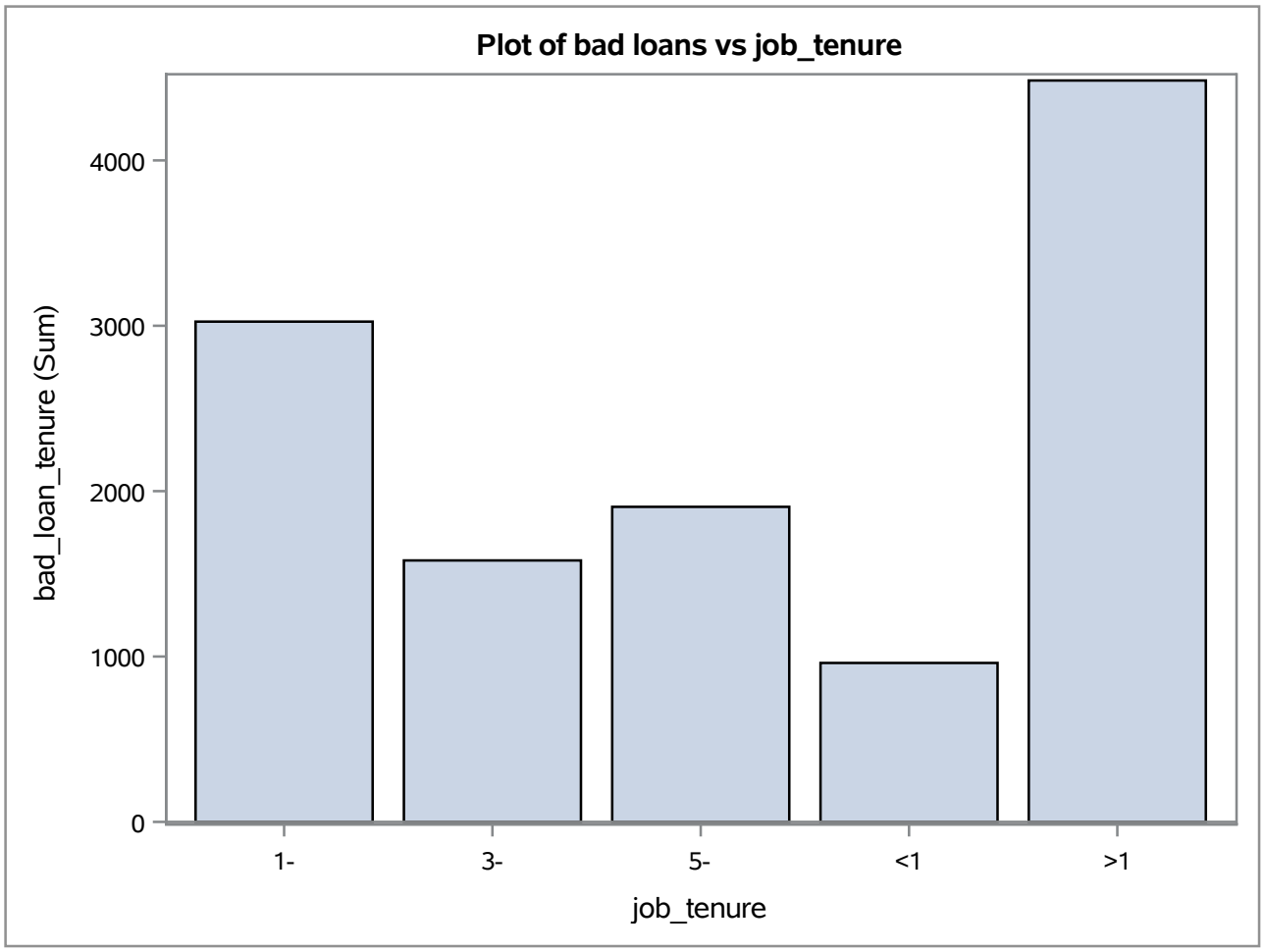

Plot of number of loans vs month
count (Sum)
30000
20000
10000
0
1
2
3
4
5
6
7
8
9
10
11
12
month

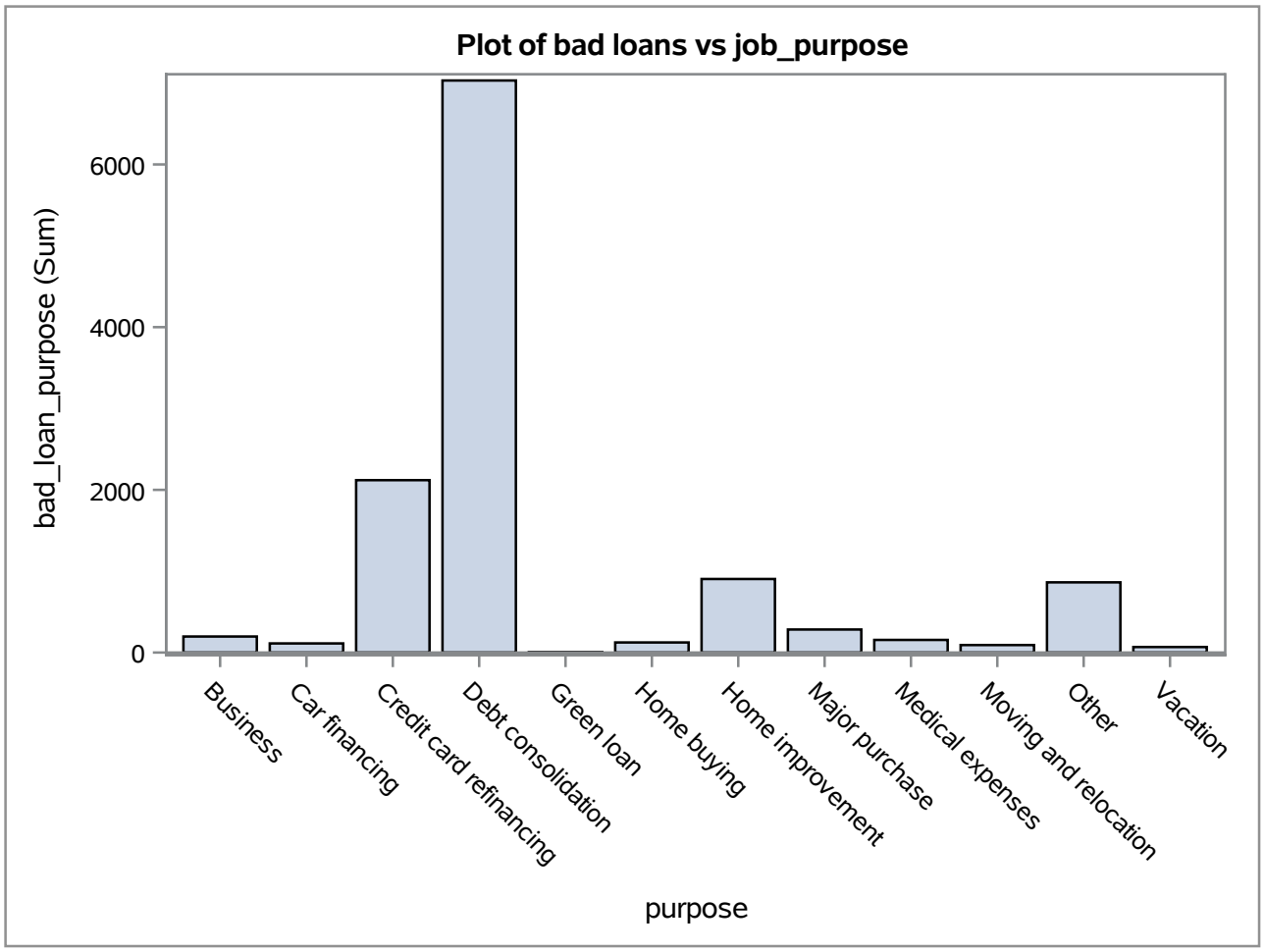

Plot of number of loans vs job tenure
125000
100000
75000
50000
25000
0
count (Sum)
1-
3-
5-
<1
>1
job_length

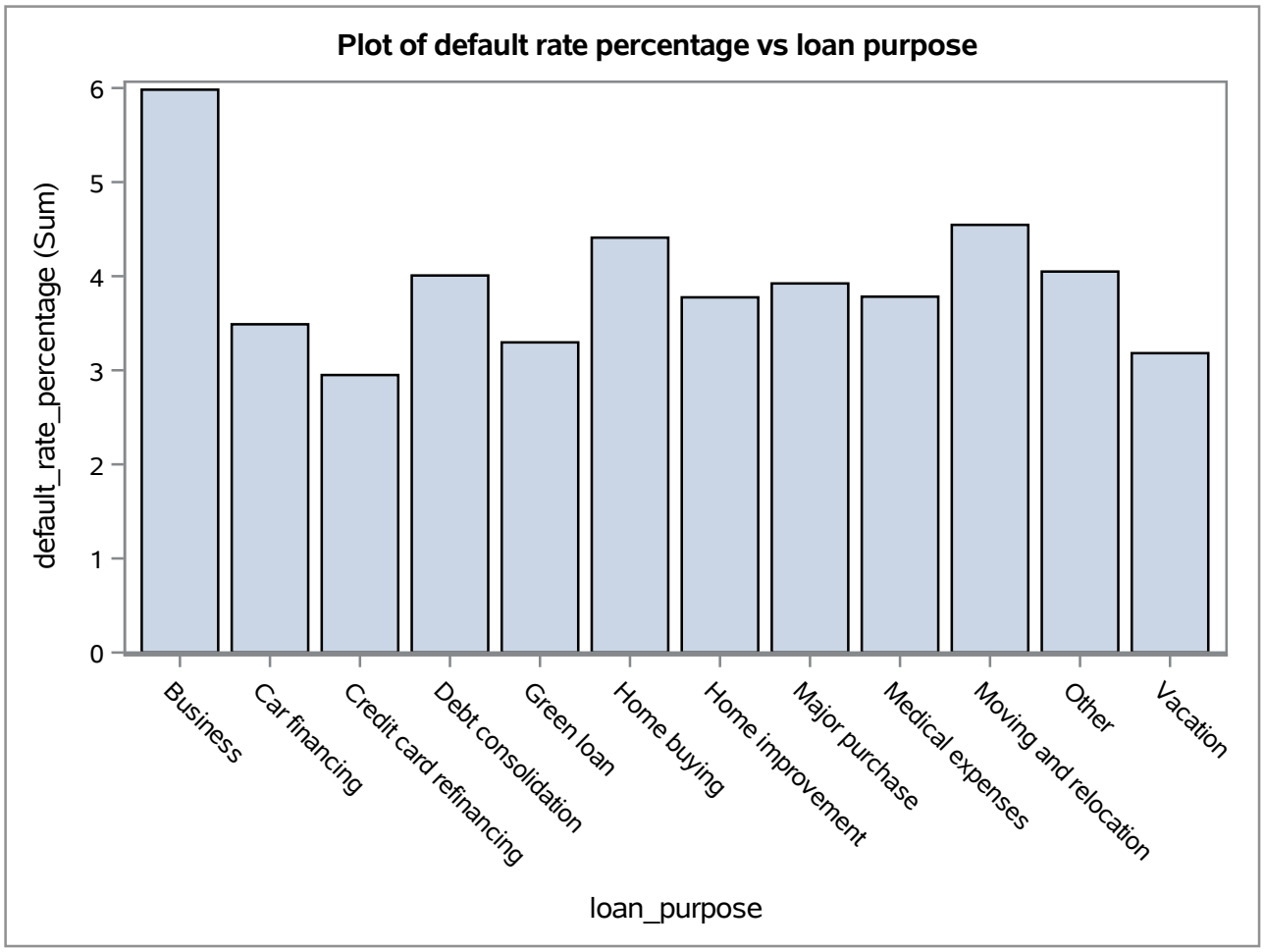

Plot of number of loans vs risk grade
100000
80000
60000
40000
20000
0
count (Sum)
A
B
C
D
E
F
G
grade


Plot of bad loans vs year
bad_loans (Sum)
4000
3000
2000
1000
0
2015
2016
2017
2018
year


Plot of bad loans vs risk_grade
bad_loan_grade (Sum)
4000
3000
2000
1000
0
A
B
C
D
E
F
G
risk_grade


Plot of bad loans vs job_tenure
bad_loan_tenure (Sum)
4000
3000
2000
1000
0
1-
3-
5-
<1
>1
job_tenure


Plot of bad loans vs job_purpose
bad_loan_purpose (Sum)
6000
4000
2000
0
Business
Car financing
Credit card refinancing
Debt consolidation
Green loan
Home buying
Home improvement
Major purchase
Medical expenses
Moving and relocation
Other
Vacation
purpose


Plot of default rate percentage vs loan purpose
default_rate_percentage (Sum)
0
1
2
3
4
5
6
Business
Car financing
Credit card refinancing
Debt consolidation
Green loan
Home buying
Home improvement
Major purchase
Medical expenses
Moving and relocation
Other
Vacation
loan_purpose

Plot of default rate percentage vs loan purpose

grade1	
0	55685
1	94005
2	99406
3	47758
4	16182
5	3664
6	832

Plot of default rate percentage vs loan purpose

The CONTENTS Procedure

Data Set Name	WORK.MODELLING_DATA	**Observations**	317532
Member Type	DATA	**Variables**	21
Engine	V9	**Indexes**	0
Created	11/20/2019 16:15:22	**Observation Length**	168
Last Modified	11/20/2019 16:15:22	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	408
First Data Page	1
Max Obs per Page	779
Obs in First Data Page	751
Number of Data Set Repairs	0
Filename	/saswork/SAS_work0D9F0001C2E0_odaws02-prod-us.oda.sas.com/SAS_work71500001C2E0_odaws02-prod-us.oda.sas.com/modelling_data.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	1074815024
Access Permission	rw-r--r--
Owner Name	dtiwari40
File Size	51MB
File Size (bytes)	53608448

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat
4	annual_inc	Num	8	BEST12.	BEST32.
17	avg_cur_bal	Num	8	BEST12.	BEST32.
20	bad_loan_status	Num	8		
18	chargeoff_within_12_mths	Num	8	BEST12.	BEST32.
6	delinq_2yrs	Num	8	BEST12.	BEST32.
5	dti	Num	8	BEST12.	BEST32.

Plot of default rate percentage vs loan purpose

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
21	grade1	Num	8		
3	installment	Num	8	BEST12.	BEST32.
2	int_rate	Num	8	BEST12.	BEST32.
14	last_pymnt_amnt	Num	8	BEST12.	BEST32.
1	loan_amnt	Num	8	BEST12.	BEST32.
7	mths_since_last_delinq	Num	8	BEST12.	BEST32.
8	out_prncp	Num	8	BEST12.	BEST32.
13	recoveries	Num	8	BEST12.	BEST32.
19	tax_liens	Num	8	BEST12.	BEST32.
15	tot_coll_amt	Num	8	BEST12.	BEST32.
16	tot_cur_bal	Num	8	BEST12.	BEST32.
9	total_pymnt	Num	8	BEST12.	BEST32.
11	total_rec_int	Num	8	BEST12.	BEST32.
12	total_rec_late_fee	Num	8	BEST12.	BEST32.
10	total_rec_prncp	Num	8	BEST12.	BEST32.

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

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

Number of Observations Read	317532
Number of Observations Used	317532

Response Profile		
Ordered Value	**bad_loan_status**	**Total Frequency**
1	0	305577
2	1	11955

Probability modeled is bad_loan_status=1.

Stepwise Selection Procedure

Step 0. Intercept entered:

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

-2 Log L	=	101865.29

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-3.2411	0.00932	120852.265	<.0001

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
17350.3952	18	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	0.6477	0.4210
avg_cur_bal	1	14.9928	0.0001
chargeoff_within_12_	1	11.5596	0.0007
delinq_2yrs	1	74.1419	<.0001
dti	1	23.5456	<.0001
grade1	1	3231.9517	<.0001
installment	1	511.9189	<.0001
int_rate	1	2851.3095	<.0001
last_pymnt_amnt	1	128.2643	<.0001
loan_amnt	1	321.2653	<.0001
mths_since_last_deli	1	90.0384	<.0001
out_prncp	1	58.2623	<.0001
recoveries	0	.	.
tax_liens	1	60.5273	<.0001
tot_coll_amt	1	0.4893	0.4842
tot_cur_bal	1	13.9567	0.0002
total_pymnt	1	920.1058	<.0001
total_rec_int	1	2008.3352	<.0001
total_rec_late_fee	1	14712.7605	<.0001
total_rec_prncp	1	362.0701	<.0001

Step 1. Effect total_rec_late_fee entered:

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	96611.695
SC	101877.96	96633.031
-2 Log L	101865.29	96607.695

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	5257.5925	1	<.0001
Score	14712.7605	1	<.0001
Wald	3902.8197	1	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-3.3153	0.00972	116287.472	<.0001
total_rec_late_fee	1	0.4182	0.00669	3902.8197	<.0001

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
total_rec_late_fee	1.519	1.499	1.539

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	27.5	**Somers' D**	0.254
Percent Discordant	2.1	**Gamma**	0.858
Percent Tied	70.4	**Tau-a**	0.018
Pairs	3653173035	**c**	0.627

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
2738.6589	17	<.0001

Analysis of Effects Eligible for Removal			
Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
total_rec_late_fee	1	3902.8197	<.0001

Note: No effects for the model in Step 1 are removed.

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	13.4055	0.0003
avg_cur_bal	1	66.6548	<.0001
chargeoff_within_12_	1	11.6292	0.0006
delinq_2yrs	1	28.8424	<.0001
dti	1	26.0555	<.0001
grade1	1	2556.8625	<.0001
installment	1	142.4592	<.0001
int_rate	1	2340.0075	<.0001
last_pymnt_amnt	1	16.7172	<.0001
loan_amnt	1	72.0478	<.0001
mths_since_last_deli	1	34.1141	<.0001
out_prncp	1	47.0690	<.0001
recoveries	0	.	.
tax_liens	1	24.8243	<.0001
tot_coll_amt	1	1.4007	0.2366
tot_cur_bal	1	69.7308	<.0001
total_pymnt	1	176.5710	<.0001
total_rec_int	1	770.5418	<.0001
total_rec_prncp	1	16.8607	<.0001

Step 2. Effect grade1 entered:

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94210.007
SC	101877.96	94242.012
-2 Log L	101865.29	94204.007

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	7661.2797	2	<.0001
Score	17123.7993	2	<.0001
Wald	6306.6756	2	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-3.4140	0.0105	104776.007	<.0001
grade1	1	0.4435	0.00887	2500.5506	<.0001
total_rec_late_fee	1	0.3887	0.00660	3466.1982	<.0001

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
grade1	1.558	1.531	1.585
total_rec_late_fee	1.475	1.456	1.494

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	63.8	**Somers' D**	0.434
Percent Discordant	20.3	**Gamma**	0.516
Percent Tied	15.9	**Tau-a**	0.031
Pairs	3653173035	**c**	0.717

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
155.2492	16	<.0001

Analysis of Effects Eligible for Removal			
Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
grade1	1	2500.5506	<.0001
total_rec_late_fee	1	3466.1982	<.0001

Note: No effects for the model in Step 2 are removed.

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	2.2532	0.1333
avg_cur_bal	1	24.6821	<.0001
chargeoff_within_12_	1	6.5299	0.0106
delinq_2yrs	1	6.2721	0.0123
dti	1	3.2989	0.0693
installment	1	42.8513	<.0001
int_rate	1	1.6557	0.1982
last_pymnt_amnt	1	3.1745	0.0748
loan_amnt	1	19.8243	<.0001
mths_since_last_deli	1	9.3252	0.0023
out_prncp	1	5.1749	0.0229
recoveries	0	.	.
tax_liens	1	19.1062	<.0001
tot_coll_amt	1	0.3544	0.5516
tot_cur_bal	1	24.7306	<.0001
total_pymnt	1	11.5564	0.0007
total_rec_int	1	2.5289	0.1118
total_rec_prncp	1	16.4877	<.0001

Step 3. Effect installment entered:

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94169.859
SC	101877.96	94212.532
-2 Log L	101865.29	94161.859

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	7703.4286	3	<.0001
Score	17195.7808	3	<.0001
Wald	6383.9251	3	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-3.4153	0.0106	104727.111	<.0001
grade1	1	0.4357	0.00892	2383.9035	<.0001
installment	1	0.0621	0.00948	42.8303	<.0001
total_rec_late_fee	1	0.3819	0.00666	3291.9281	<.0001

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
grade1	1.546	1.519	1.573
installment	1.064	1.044	1.084
total_rec_late_fee	1.465	1.446	1.484

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.7	**Somers' D**	0.434
Percent Discordant	28.3	**Gamma**	0.434
Percent Tied	0.0	**Tau-a**	0.031
Pairs	3653173035	**c**	0.717

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
109.5050	15	<.0001

Analysis of Effects Eligible for Removal			
Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
grade1	1	2383.9035	<.0001
installment	1	42.8303	<.0001
total_rec_late_fee	1	3291.9281	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Note: No effects for the model in Step 3 are removed.

Analysis of Effects Eligible for Entry			
Effect	**DF**	**Score Chi-Square**	**Pr > ChiSq**
annual_inc	1	6.9981	0.0082
avg_cur_bal	1	37.1597	<.0001
chargeoff_within_12_	1	6.8104	0.0091
delinq_2yrs	1	6.7032	0.0096
dti	1	4.0224	0.0449
int_rate	1	2.4364	0.1185
last_pymnt_amnt	1	0.4869	0.4853
loan_amnt	1	20.6924	<.0001
mths_since_last_deli	1	9.5304	0.0020
out_prncp	1	11.7838	0.0006
recoveries	0	.	.
tax_liens	1	17.9118	<.0001
tot_coll_amt	1	0.4963	0.4811
tot_cur_bal	1	44.3603	<.0001
total_pymnt	1	0.0000	0.9964
total_rec_int	1	2.8083	0.0938
total_rec_prncp	1	0.8362	0.3605

Step 4. Effect tot_cur_bal entered:

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

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	101867.29	94125.227
SC	101877.96	94178.568
-2 Log L	101865.29	94115.227

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	7750.0603	4	<.0001
Score	17235.1863	4	<.0001
Wald	6409.2281	4	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-3.4174	0.0106	104475.889	<.0001
grade1	1	0.4294	0.00898	2288.0950	<.0001
installment	1	0.0782	0.00975	64.2999	<.0001
tot_cur_bal	1	-0.0715	0.0107	44.6780	<.0001
total_rec_late_fee	1	0.3843	0.00668	3310.8062	<.0001

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
grade1	1.536	1.510	1.564
installment	1.081	1.061	1.102
tot_cur_bal	0.931	0.912	0.951
total_rec_late_fee	1.469	1.450	1.488

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.8	**Somers' D**	0.436
Percent Discordant	28.2	**Gamma**	0.436
Percent Tied	0.0	**Tau-a**	0.032
Pairs	3653173035	**c**	0.718

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
64.7967	14	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
grade1	1	2288.0950	<.0001
installment	1	64.2999	<.0001
tot_cur_bal	1	44.6780	<.0001
total_rec_late_fee	1	3310.8062	<.0001

Note: No effects for the model in Step 4 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	1.7326	0.1881
avg_cur_bal	1	1.4481	0.2288
chargeoff_within_12_	1	6.8957	0.0086
delinq_2yrs	1	8.9654	0.0028
dti	1	3.2968	0.0694
int_rate	1	2.9677	0.0849
last_pymnt_amnt	1	0.4219	0.5160
loan_amnt	1	13.3689	0.0003
mths_since_last_deli	1	12.5955	0.0004
out_prncp	1	8.9101	0.0028
recoveries	0	.	.
tax_liens	1	17.7926	<.0001
tot_coll_amt	1	0.4076	0.5232
total_pymnt	1	0.0386	0.8442
total_rec_int	1	1.8626	0.1723
total_rec_prncp	1	1.0330	0.3094

Step 5. Effect tax_liens entered:

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

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94114.570
SC	101877.96	94178.580
-2 Log L	101865.29	94102.570

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7762.7175	5	<.0001
Score	17252.7642	5	<.0001
Wald	6426.4774	5	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4178	0.0106	104449.602	<.0001
grade1	1	0.4292	0.00898	2285.5185	<.0001
installment	1	0.0777	0.00976	63.4290	<.0001
tax_liens	1	0.0238	0.00609	15.2773	<.0001
tot_cur_bal	1	-0.0715	0.0107	44.5979	<.0001
total_rec_late_fee	1	0.3836	0.00668	3299.2288	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
grade1	1.536	1.509	1.563
installment	1.081	1.060	1.102
tax_liens	1.024	1.012	1.036
tot_cur_bal	0.931	0.912	0.951
total_rec_late_fee	1.468	1.448	1.487

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.8	Somers' D	0.437
Percent Discordant	28.2	Gamma	0.437
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.718

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
47.1067	13	<.0001

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
grade1	1	2285.5185	<.0001
installment	1	63.4290	<.0001
tax_liens	1	15.2773	<.0001
tot_cur_bal	1	44.5979	<.0001
total_rec_late_fee	1	3299.2288	<.0001

Note: No effects for the model in Step 5 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	1.9493	0.1627
avg_cur_bal	1	1.5179	0.2179
chargeoff_within_12_	1	6.9217	0.0085
delinq_2yrs	1	9.0604	0.0026
dti	1	3.0649	0.0800
int_rate	1	2.5528	0.1101
last_pymnt_amnt	1	0.4088	0.5226
loan_amnt	1	13.3345	0.0003
mths_since_last_deli	1	12.7482	0.0004
out_prncp	1	7.7031	0.0055
recoveries	0	.	.
tot_coll_amt	1	0.3673	0.5445
total_pymnt	1	0.0017	0.9675
total_rec_int	1	2.3686	0.1238
total_rec_prncp	1	0.6008	0.4383

Step 6. Effect loan_amnt entered:

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

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94103.139
SC	101877.96	94177.817
-2 Log L	101865.29	94089.139

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7776.1482	6	<.0001
Score	17286.5589	6	<.0001
Wald	6450.5970	6	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4177	0.0106	104508.787	<.0001
grade1	1	0.4269	0.00898	2259.3140	<.0001
installment	1	0.1692	0.0268	39.9396	<.0001
loan_amnt	1	-0.1015	0.0278	13.3297	0.0003
tax_liens	1	0.0238	0.00609	15.1886	<.0001
tot_cur_bal	1	-0.0660	0.0108	37.6110	<.0001
total_rec_late_fee	1	0.3833	0.00668	3296.0382	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
grade1	1.533	1.506	1.560
installment	1.184	1.124	1.248
loan_amnt	0.904	0.856	0.954
tax_liens	1.024	1.012	1.036
tot_cur_bal	0.936	0.917	0.956
total_rec_late_fee	1.467	1.448	1.486

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.0	**Somers' D**	0.439
Percent Discordant	28.0	**Gamma**	0.439
Percent Tied	0.0	**Tau-a**	0.032
Pairs	3653173035	**c**	0.720

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
33.7067	12	0.0008

Analysis of Effects Eligible for Removal			
Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
grade1	1	2259.3140	<.0001
installment	1	39.9396	<.0001
loan_amnt	1	13.3297	0.0003
tax_liens	1	15.1886	<.0001
tot_cur_bal	1	37.6110	<.0001
total_rec_late_fee	1	3296.0382	<.0001

Note: No effects for the model in Step 6 are removed.

Analysis of Effects Eligible for Entry			
Effect	**DF**	**Score Chi-Square**	**Pr > ChiSq**
annual_inc	1	1.6889	0.1937
avg_cur_bal	1	1.5106	0.2190
chargeoff_within_12_	1	6.8288	0.0090
delinq_2yrs	1	8.4647	0.0036
dti	1	3.3284	0.0681
int_rate	1	5.9751	0.0145
last_pymnt_amnt	1	0.4135	0.5202
mths_since_last_deli	1	11.8752	0.0006
out_prncp	1	1.1887	0.2756
recoveries	0	.	.
tot_coll_amt	1	0.3184	0.5726
total_pymnt	1	0.6458	0.4216

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
total_rec_int	1	0.0017	0.9672
total_rec_prncp	1	1.1893	0.2755

Step 7. Effect mths_since_last_deli entered:

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94093.204
SC	101877.96	94178.551
-2 Log L	101865.29	94077.204

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7788.0827	7	<.0001
Score	17291.5477	7	<.0001
Wald	6461.5452	7	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4186	0.0106	104364.410	<.0001
grade1	1	0.4253	0.00900	2232.5646	<.0001
installment	1	0.1669	0.0268	38.8364	<.0001
loan_amnt	1	-0.0982	0.0278	12.4600	0.0004
mths_since_last_deli	1	-0.0338	0.00979	11.8738	0.0006
tax_liens	1	0.0239	0.00609	15.4087	<.0001
tot_cur_bal	1	-0.0687	0.0108	40.3696	<.0001
total_rec_late_fee	1	0.3820	0.00668	3270.5546	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
grade1	1.530	1.503	1.557
installment	1.182	1.121	1.245
loan_amnt	0.907	0.858	0.957
mths_since_last_deli	0.967	0.948	0.986
tax_liens	1.024	1.012	1.036
tot_cur_bal	0.934	0.914	0.954
total_rec_late_fee	1.465	1.446	1.485

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.9	Somers' D	0.438
Percent Discordant	28.1	Gamma	0.438
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.719

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
21.6314	11	0.0274

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
grade1	1	2232.5646	<.0001
installment	1	38.8364	<.0001
loan_amnt	1	12.4600	0.0004
mths_since_last_deli	1	11.8738	0.0006
tax_liens	1	15.4087	<.0001
tot_cur_bal	1	40.3696	<.0001
total_rec_late_fee	1	3270.5546	<.0001

Note: No effects for the model in Step 7 are removed.

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	**DF**	**Score Chi-Square**	**Pr > ChiSq**
annual_inc	1	1.8280	0.1764
avg_cur_bal	1	1.6174	0.2035
chargeoff_within_12_	1	5.0685	0.0244
delinq_2yrs	1	1.4566	0.2275
dti	1	2.9336	0.0868
int_rate	1	5.6901	0.0171
last_pymnt_amnt	1	0.4060	0.5240
out_prncp	1	0.7759	0.3784
recoveries	0	.	.
tot_coll_amt	1	0.4366	0.5088
total_pymnt	1	0.3674	0.5444
total_rec_int	1	0.0080	0.9288
total_rec_prncp	1	0.7765	0.3782

Step 8. Effect int_rate entered:

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

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	101867.29	94089.521
SC	101877.96	94185.536
-2 Log L	101865.29	94071.521

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	7793.7665	8	<.0001
Score	17292.5739	8	<.0001
Wald	6468.8948	8	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4191	0.0106	104303.514	<.0001
grade1	1	0.4917	0.0292	283.5724	<.0001
installment	1	0.1813	0.0274	43.6195	<.0001
int_rate	1	-0.0687	0.0288	5.6923	0.0170
loan_amnt	1	-0.1128	0.0285	15.6653	<.0001
mths_since_last_deli	1	-0.0333	0.00979	11.5927	0.0007
tax_liens	1	0.0234	0.00610	14.6792	0.0001
tot_cur_bal	1	-0.0685	0.0108	40.1381	<.0001
total_rec_late_fee	1	0.3810	0.00669	3244.5857	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
grade1	1.635	1.544	1.731
installment	1.199	1.136	1.265
int_rate	0.934	0.882	0.988
loan_amnt	0.893	0.845	0.945
mths_since_last_deli	0.967	0.949	0.986
tax_liens	1.024	1.011	1.036
tot_cur_bal	0.934	0.914	0.954
total_rec_late_fee	1.464	1.445	1.483

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.9	Somers' D	0.438
Percent Discordant	28.1	Gamma	0.438
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.719

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
15.9581	10	0.1008

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
grade1	1	283.5724	<.0001
installment	1	43.6195	<.0001
int_rate	1	5.6923	0.0170
loan_amnt	1	15.6653	<.0001
mths_since_last_deli	1	11.5927	0.0007
tax_liens	1	14.6792	0.0001
tot_cur_bal	1	40.1381	<.0001
total_rec_late_fee	1	3244.5857	<.0001

Note: No effects for the model in Step 8 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	1.8551	0.1732
avg_cur_bal	1	1.6097	0.2045
chargeoff_within_12_	1	5.0521	0.0246
delinq_2yrs	1	1.4100	0.2351
dti	1	2.7388	0.0979
last_pymnt_amnt	1	0.3935	0.5305
out_prncp	1	0.0275	0.8683
recoveries	0	.	.
tot_coll_amt	1	0.4126	0.5206
total_pymnt	1	0.2061	0.6499
total_rec_int	1	0.8239	0.3640
total_rec_prncp	1	0.0274	0.8685

Step 9. Effect chargeoff_within_12_ entered:

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

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94086.811
SC	101877.96	94193.494
-2 Log L	101865.29	94066.811

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7798.4760	9	<.0001
Score	17297.1634	9	<.0001
Wald	6472.9539	9	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4192	0.0106	104285.538	<.0001
chargeoff_within_12_	1	0.0188	0.00837	5.0560	0.0245
grade1	1	0.4914	0.0292	283.2636	<.0001
installment	1	0.1813	0.0274	43.6439	<.0001
int_rate	1	-0.0686	0.0288	5.6771	0.0172
loan_amnt	1	-0.1127	0.0285	15.6443	<.0001
mths_since_last_deli	1	-0.0312	0.00984	10.0659	0.0015
tax_liens	1	0.0234	0.00609	14.6927	0.0001
tot_cur_bal	1	-0.0683	0.0108	40.0100	<.0001
total_rec_late_fee	1	0.3810	0.00669	3245.6702	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
chargeoff_within_12_	1.019	1.002	1.036
grade1	1.635	1.544	1.731
installment	1.199	1.136	1.265
int_rate	0.934	0.882	0.988
loan_amnt	0.893	0.845	0.945
mths_since_last_deli	0.969	0.951	0.988
tax_liens	1.024	1.011	1.036

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
tot_cur_bal	0.934	0.914	0.954
total_rec_late_fee	1.464	1.445	1.483

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.9	Somers' D	0.439
Percent Discordant	28.1	Gamma	0.439
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.719

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
10.8897	9	0.2833

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
chargeoff_within_12_	1	5.0560	0.0245
grade1	1	283.2636	<.0001
installment	1	43.6439	<.0001
int_rate	1	5.6771	0.0172
loan_amnt	1	15.6443	<.0001
mths_since_last_deli	1	10.0659	0.0015
tax_liens	1	14.6927	0.0001
tot_cur_bal	1	40.0100	<.0001
total_rec_late_fee	1	3245.6702	<.0001

Note: No effects for the model in Step 9 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	1.8743	0.1710
avg_cur_bal	1	1.6172	0.2035
delinq_2yrs	1	1.0441	0.3069
dti	1	2.7146	0.0994
last_pymnt_amnt	1	0.3996	0.5273

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
out_prncp	1	0.0248	0.8748
recoveries	0	.	.
tot_coll_amt	1	0.4010	0.5266
total_pymnt	1	0.1993	0.6553
total_rec_int	1	0.8164	0.3662
total_rec_prncp	1	0.0247	0.8750

Step 10. Effect dti entered:

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94085.764
SC	101877.96	94203.115
-2 Log L	101865.29	94063.764

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7801.5233	10	<.0001
Score	17300.8660	10	<.0001
Wald	6478.6758	10	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4193	0.0106	104279.285	<.0001
chargeoff_within_12_	1	0.0188	0.00837	5.0236	0.0250
dti	1	-0.0183	0.0109	2.7913	0.0948
grade1	1	0.4926	0.0292	284.3593	<.0001
installment	1	0.1824	0.0275	44.1633	<.0001
int_rate	1	-0.0672	0.0288	5.4487	0.0196
loan_amnt	1	-0.1135	0.0285	15.8742	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
mths_since_last_deli	1	-0.0306	0.00985	9.6694	0.0019
tax_liens	1	0.0232	0.00609	14.4454	0.0001
tot_cur_bal	1	-0.0676	0.0108	39.2006	<.0001
total_rec_late_fee	1	0.3807	0.00669	3240.2014	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
chargeoff_within_12_	1.019	1.002	1.036
dti	0.982	0.961	1.003
grade1	1.636	1.545	1.733
installment	1.200	1.137	1.266
int_rate	0.935	0.884	0.989
loan_amnt	0.893	0.844	0.944
mths_since_last_deli	0.970	0.951	0.989
tax_liens	1.023	1.011	1.036
tot_cur_bal	0.935	0.915	0.955
total_rec_late_fee	1.463	1.444	1.483

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.9	Somers' D	0.438
Percent Discordant	28.1	Gamma	0.438
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.719

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
8.3796	8	0.3973

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
chargeoff_within_12_	1	5.0236	0.0250
dti	1	2.7913	0.0948
grade1	1	284.3593	<.0001
installment	1	44.1633	<.0001
int_rate	1	5.4487	0.0196
loan_amnt	1	15.8742	<.0001
mths_since_last_deli	1	9.6694	0.0019
tax_liens	1	14.4454	0.0001
tot_cur_bal	1	39.2006	<.0001
total_rec_late_fee	1	3240.2014	<.0001

Note: No effects for the model in Step 10 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	2.4799	0.1153
avg_cur_bal	1	2.5043	0.1135
delinq_2yrs	1	0.9980	0.3178
last_pymnt_amnt	1	0.4129	0.5205
out_prncp	1	0.0308	0.8606
recoveries	0	.	.
tot_coll_amt	1	0.3638	0.5464
total_pymnt	1	0.2144	0.6434
total_rec_int	1	0.8340	0.3611
total_rec_prncp	1	0.0307	0.8608

Step 11. Effect avg_cur_bal entered:

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

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94085.216
SC	101877.96	94213.236
-2 Log L	101865.29	94061.216

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7804.0716	11	<.0001
Score	17303.3431	11	<.0001
Wald	6480.0460	11	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4194	0.0106	104271.458	<.0001
avg_cur_bal	1	-0.0290	0.0184	2.5048	0.1135
chargeoff_within_12_	1	0.0188	0.00837	5.0294	0.0249
dti	1	-0.0214	0.0113	3.5899	0.0581
grade1	1	0.4925	0.0292	284.3175	<.0001
installment	1	0.1817	0.0275	43.8143	<.0001
int_rate	1	-0.0669	0.0288	5.4018	0.0201
loan_amnt	1	-0.1136	0.0285	15.8958	<.0001
mths_since_last_deli	1	-0.0307	0.00985	9.7245	0.0018
tax_liens	1	0.0232	0.00609	14.5052	0.0001
tot_cur_bal	1	-0.0435	0.0185	5.5121	0.0189
total_rec_late_fee	1	0.3810	0.00669	3241.2881	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
avg_cur_bal	0.971	0.937	1.007
chargeoff_within_12_	1.019	1.002	1.036
dti	0.979	0.957	1.001
grade1	1.636	1.545	1.733
installment	1.199	1.136	1.266
int_rate	0.935	0.884	0.990

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
loan_amnt	0.893	0.844	0.944
mths_since_last_deli	0.970	0.951	0.989
tax_liens	1.023	1.011	1.036
tot_cur_bal	0.957	0.923	0.993
total_rec_late_fee	1.464	1.445	1.483

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.9	Somers' D	0.439
Percent Discordant	28.1	Gamma	0.439
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.719

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
5.8766	7	0.5542

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
avg_cur_bal	1	2.5048	0.1135
chargeoff_within_12_	1	5.0294	0.0249
dti	1	3.5899	0.0581
grade1	1	284.3175	<.0001
installment	1	43.8143	<.0001
int_rate	1	5.4018	0.0201
loan_amnt	1	15.8958	<.0001
mths_since_last_deli	1	9.7245	0.0018
tax_liens	1	14.5052	0.0001
tot_cur_bal	1	5.5121	0.0189
total_rec_late_fee	1	3241.2881	<.0001

Note: No effects for the model in Step 11 are removed.

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
annual_inc	1	2.7568	0.0968
delinq_2yrs	1	0.9698	0.3247
last_pymnt_amnt	1	0.3996	0.5273
out_prncp	1	0.0381	0.8453
recoveries	0	.	.
tot_coll_amt	1	0.3462	0.5563
total_pymnt	1	0.2343	0.6283
total_rec_int	1	0.8685	0.3514
total_rec_prncp	1	0.0380	0.8455

Step 12. Effect annual_inc entered:

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94078.230
SC	101877.96	94216.919
-2 Log L	101865.29	94052.230

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7813.0566	12	<.0001
Score	17305.1805	12	<.0001
Wald	6483.9227	12	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4201	0.0106	104169.656	<.0001
annual_inc	1	-0.0774	0.0284	7.4157	0.0065
avg_cur_bal	1	-0.0357	0.0186	3.6761	0.0552
chargeoff_within_12_	1	0.0189	0.00837	5.1145	0.0237

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
dti	1	-0.0312	0.0122	6.5106	0.0107
grade1	1	0.4906	0.0292	282.1020	<.0001
installment	1	0.1861	0.0275	45.7771	<.0001
int_rate	1	-0.0670	0.0288	5.4077	0.0200
loan_amnt	1	-0.1094	0.0285	14.6757	0.0001
mths_since_last_deli	1	-0.0313	0.00985	10.1073	0.0015
tax_liens	1	0.0241	0.00610	15.5362	<.0001
tot_cur_bal	1	-0.0244	0.0199	1.5059	0.2198
total_rec_late_fee	1	0.3819	0.00670	3245.4237	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
annual_inc	0.926	0.875	0.979
avg_cur_bal	0.965	0.930	1.001
chargeoff_within_12_	1.019	1.003	1.036
dti	0.969	0.946	0.993
grade1	1.633	1.542	1.730
installment	1.205	1.141	1.271
int_rate	0.935	0.884	0.990
loan_amnt	0.896	0.848	0.948
mths_since_last_deli	0.969	0.951	0.988
tax_liens	1.024	1.012	1.037
tot_cur_bal	0.976	0.939	1.015
total_rec_late_fee	1.465	1.446	1.484

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.0	Somers' D	0.439
Percent Discordant	28.0	Gamma	0.439
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.720

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
3.3596	6	0.7625

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
annual_inc	1	7.4157	0.0065
avg_cur_bal	1	3.6761	0.0552
chargeoff_within_12_	1	5.1145	0.0237
dti	1	6.5106	0.0107
grade1	1	282.1020	<.0001
installment	1	45.7771	<.0001
int_rate	1	5.4077	0.0200
loan_amnt	1	14.6757	0.0001
mths_since_last_deli	1	10.1073	0.0015
tax_liens	1	15.5362	<.0001
tot_cur_bal	1	1.5059	0.2198
total_rec_late_fee	1	3245.4237	<.0001

Note: No effects for the model in Step 12 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
delinq_2yrs	1	1.0776	0.2992
last_pymnt_amnt	1	0.4149	0.5195
out_prncp	1	0.0207	0.8855
recoveries	0	.	.
tot_coll_amt	1	0.3528	0.5525
total_pymnt	1	0.2085	0.6479
total_rec_int	1	0.9170	0.3383
total_rec_prncp	1	0.0206	0.8857

Step 13. Effect delinq_2yrs entered:

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

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	101867.29	94079.170
SC	101877.96	94228.527
-2 Log L	101865.29	94051.170

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7814.1171	13	<.0001
Score	17305.3939	13	<.0001
Wald	6485.0893	13	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.4202	0.0106	104157.194	<.0001
annual_inc	1	-0.0780	0.0284	7.5091	0.0061
avg_cur_bal	1	-0.0356	0.0186	3.6504	0.0561
chargeoff_within_12_	1	0.0182	0.00839	4.7079	0.0300
delinq_2yrs	1	0.0111	0.0107	1.0781	0.2991
dti	1	-0.0311	0.0122	6.4685	0.0110
grade1	1	0.4901	0.0292	281.4284	<.0001
installment	1	0.1860	0.0275	45.7234	<.0001
int_rate	1	-0.0668	0.0288	5.3720	0.0205
loan_amnt	1	-0.1090	0.0286	14.5740	0.0001
mths_since_last_deli	1	-0.0250	0.0116	4.6414	0.0312
tax_liens	1	0.0241	0.00610	15.5588	<.0001
tot_cur_bal	1	-0.0247	0.0199	1.5432	0.2141
total_rec_late_fee	1	0.3817	0.00670	3243.2022	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
annual_inc	0.925	0.875	0.978
avg_cur_bal	0.965	0.931	1.001
chargeoff_within_12_	1.018	1.002	1.035
delinq_2yrs	1.011	0.990	1.033

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
dti	0.969	0.946	0.993
grade1	1.632	1.542	1.729
installment	1.204	1.141	1.271
int_rate	0.935	0.884	0.990
loan_amnt	0.897	0.848	0.948
mths_since_last_deli	0.975	0.953	0.998
tax_liens	1.024	1.012	1.037
tot_cur_bal	0.976	0.938	1.014
total_rec_late_fee	1.465	1.446	1.484

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.0	Somers' D	0.440
Percent Discordant	28.0	Gamma	0.440
Percent Tied	0.0	Tau-a	0.032
Pairs	3653173035	c	0.720

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
2.2823	5	0.8089

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
annual_inc	1	7.5091	0.0061
avg_cur_bal	1	3.6504	0.0561
chargeoff_within_12_	1	4.7079	0.0300
delinq_2yrs	1	1.0781	0.2991
dti	1	6.4685	0.0110
grade1	1	281.4284	<.0001
installment	1	45.7234	<.0001
int_rate	1	5.3720	0.0205
loan_amnt	1	14.5740	0.0001
mths_since_last_deli	1	4.6414	0.0312
tax_liens	1	15.5588	<.0001

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Analysis of Effects Eligible for Removal			
Effect	DF	Wald Chi-Square	Pr > ChiSq
tot_cur_bal	1	1.5432	0.2141
total_rec_late_fee	1	3243.2022	<.0001

Note: No effects for the model in Step 13 are removed.

Analysis of Effects Eligible for Entry			
Effect	DF	Score Chi-Square	Pr > ChiSq
last_pymnt_amnt	1	0.4141	0.5199
out_prncp	1	0.0248	0.8749
recoveries	0	.	.
tot_coll_amt	1	0.3644	0.5461
total_pymnt	1	0.2195	0.6394
total_rec_int	1	0.9284	0.3353
total_rec_prncp	1	0.0247	0.8752

Note: No (additional) effects met the 0.3 significance level for entry into the model.

Summary of Stepwise Selection							
Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
1	total_rec_late_fee		1	1	14712.7605		<.0001
2	grade1		1	2	2556.8625		<.0001
3	installment		1	3	42.8513		<.0001
4	tot_cur_bal		1	4	44.3603		<.0001
5	tax_liens		1	5	17.7926		<.0001
6	loan_amnt		1	6	13.3345		0.0003
7	mths_since_last_deli		1	7	11.8752		0.0006
8	int_rate		1	8	5.6901		0.0171
9	chargeoff_within_12_		1	9	5.0521		0.0246
10	dti		1	10	2.7146		0.0994
11	avg_cur_bal		1	11	2.5043		0.1135
12	annual_inc		1	12	2.7568		0.0968
13	delinq_2yrs		1	13	1.0776		0.2992

Plot of default rate percentage vs loan purpose

The LOGISTIC Procedure

Partition for the Hosmer and Lemeshow Test					
		bad_loan_status = 1		bad_loan_status = 0	
Group	Total	Observed	Expected	Observed	Expected
1	31758	261	491.62	31497	31266.38
2	31754	364	592.21	31390	31161.79
3	31748	549	711.83	31199	31036.17
4	31752	732	775.25	31020	30976.75
5	31749	825	900.59	30924	30848.41
6	31748	982	1036.65	30766	30711.35
7	31750	1099	1114.44	30651	30635.56
8	31754	1384	1307.60	30370	30446.40
9	31755	1688	1587.66	30067	30167.34
10	31764	4071	3437.06	27693	28326.94

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
392.2494	8	<.0001

Plot of default rate percentage vs loan purpose
Parameter Estimates and Covariance Matrix

Obs	_LINK_	_TYPE_	_STATUS_	_NAME_	Intercept	annual_inc	avg_cur_bal	chargeoff_within_12_mths	delinq_2yrs
1	LOGIT	PARMS	0 Converged	bad_loan_status	-3.42016	-0.077954	-0.035554	0.018196	0.011141
2	LOGIT	COV	0 Converged	Intercept	0.00011	0.000011	0.000003	-0.000001	-0.000001
3	LOGIT	COV	0 Converged	annual_inc	0.00001	0.000809	0.000068	-0.000001	-0.000006
4	LOGIT	COV	0 Converged	avg_cur_bal	0.00000	0.000068	0.000346	-0.000000	0.000001
5	LOGIT	COV	0 Converged	chargeoff_within_12_mths	-0.00000	-0.000001	-0.000000	0.000070	-0.000008
6	LOGIT	COV	0 Converged	delinq_2yrs	-0.00000	-0.000006	0.000001	-0.000008	0.000115
7	LOGIT	COV	0 Converged	dti	0.00000	0.000101	0.000048	0.000000	0.000001
8	LOGIT	COV	0 Converged	grade1	-0.00004	0.000018	0.000002	-0.000001	-0.000005
9	LOGIT	COV	0 Converged	installment	-0.00001	-0.000044	0.000004	0.000000	-0.000001
10	LOGIT	COV	0 Converged	int_rate	0.00001	0.000001	-0.000003	0.000000	0.000002
11	LOGIT	COV	0 Converged	last_pymnt_amnt	.	.	.	.	.

Obs	dti	grade1	installment	int_rate	last_pymnt_amnt	loan_amnt	mths_since_last_delinq	out_prncp	recoveries	tax_liens
1	-0.031119	0.49011	0.18597	-0.066752	.	-0.10899	-0.024980	.	.	0.024068
2	0.000003	-0.00004	-0.00001	0.000008	.	0.00001	0.000003	.	.	-0.000001
3	0.000101	0.00002	-0.00004	0.000001	.	-0.00004	0.000003	.	.	-0.000009
4	0.000048	0.00000	0.00000	-0.000003	.	-0.00000	0.000002	.	.	-0.000001
5	0.000000	-0.00000	0.00000	0.000000	.	0.00000	0.000004	.	.	0.000000
6	0.000001	-0.00001	-0.00000	0.000002	.	0.00000	0.000066	.	.	0.000000
7	0.000150	-0.00001	-0.00001	-0.000010	.	0.00000	-0.000003	.	.	0.000000
8	-0.000006	0.00085	0.00014	-0.000799	.	-0.00016	0.000007	.	.	-0.000007
9	-0.000013	0.00014	0.00076	-0.000174	.	-0.00073	0.000007	.	.	-0.000002
10	-0.000010	-0.00080	-0.00017	0.000829	.	0.00018	-0.000003	.	.	0.000006
11	.	.	.	.	.	.	.	.	.	.

Obs	tot_coll_amt	tot_cur_bal	total_pymnt	total_rec_int	total_rec_late_fee	total_rec_prncp	_LNLIKE_	_ESTTYPE_
1	.	-0.024704	.	.	0.38174	.	-47025.58	MLE
2	.	0.000002	.	.	-0.00001	.	-47025.58	MLE
3	.	-0.000196	.	.	-0.00001	.	-47025.58	MLE
4	.	-0.000297	.	.	-0.00000	.	-47025.58	MLE
5	.	0.000001	.	.	0.00000	.	-47025.58	MLE
6	.	-0.000003	.	.	-0.00000	.	-47025.58	MLE
7	.	-0.000063	.	.	0.00000	.	-47025.58	MLE
8	.	0.000007	.	.	-0.00001	.	-47025.58	MLE
9	.	0.000018	.	.	-0.00001	.	-47025.58	MLE
10	.	0.000001	.	.	0.00001	.	-47025.58	MLE
11	.	.	.	.	.	.	-47025.58	MLE

Plot of default rate percentage vs loan purpose
Parameter Estimates and Covariance Matrix

Obs	_LINK_	_TYPE_	_STATUS_	_NAME_	Intercept	annual_inc	avg_cur_bal	chargeoff_within_12_mths	delinq_2yrs
12	LOGIT	COV	0 Converged	loan_amnt	0.00001	-0.000045	-0.000003	0.000000	0.000004
13	LOGIT	COV	0 Converged	mths_since_last_delinq	0.00000	0.000003	0.000002	0.000004	0.000066
14	LOGIT	COV	0 Converged	out_prncp	.	.	.	.	.
15	LOGIT	COV	0 Converged	recoveries	.	.	.	.	.
16	LOGIT	COV	0 Converged	tax_liens	-0.00000	-0.000009	-0.000001	0.000000	0.000000
17	LOGIT	COV	0 Converged	tot_coll_amt	.	.	.	.	.
18	LOGIT	COV	0 Converged	tot_cur_bal	0.00000	-0.000196	-0.000297	0.000001	-0.000003
19	LOGIT	COV	0 Converged	total_pymnt	.	.	.	.	.
20	LOGIT	COV	0 Converged	total_rec_int	.	.	.	.	.
21	LOGIT	COV	0 Converged	total_rec_late_fee	-0.00001	-0.000010	-0.000004	0.000001	-0.000001
22	LOGIT	COV	0 Converged	total_rec_prncp	.	.	.	.	.

Obs	dti	grade1	installment	int_rate	last_pymnt_amnt	loan_amnt	mths_since_last_delinq	out_prncp	recoveries	tax_liens
12	0.000001	-0.00016	-0.00073	0.000177	.	0.00082	-0.000009	.	.	0.000002
13	-0.000003	0.00001	0.00001	-0.000003	.	-0.00001	0.000134	.	.	-0.000000
14	.	.	.	.	.	.	.	.	.	.
15	.	.	.	.	.	.	.	.	.	.
16	0.000000	-0.00001	-0.00000	0.000006	.	0.00000	-0.000000	.	.	0.000037
17	.	.	.	.	.	.	.	.	.	.
18	-0.000063	0.00001	0.00002	0.000001	.	-0.00003	0.000004	.	.	0.000003
19	.	.	.	.	.	.	.	.	.	.
20	.	.	.	.	.	.	.	.	.	.
21	0.000001	-0.00001	-0.00001	0.000012	.	0.00000	0.000003	.	.	-0.000001
22	.	.	.	.	.	.	.	.	.	.

Obs	tot_coll_amt	tot_cur_bal	total_pymnt	total_rec_int	total_rec_late_fee	total_rec_prncp	_LNLIKE_	_ESTTYPE_
12	.	-0.000033	.	.	0.00000	.	-47025.58	MLE
13	.	0.000004	.	.	0.00000	.	-47025.58	MLE
14	.	.	.	.	.	.	-47025.58	MLE
15	.	.	.	.	.	.	-47025.58	MLE
16	.	0.000003	.	.	-0.00000	.	-47025.58	MLE
17	.	.	.	.	.	.	-47025.58	MLE
18	.	0.000395	.	.	0.00000	.	-47025.58	MLE
19	.	.	.	.	.	.	-47025.58	MLE
20	.	.	.	.	.	.	-47025.58	MLE
21	.	0.000001	.	.	0.00004	.	-47025.58	MLE
22	.	.	.	.	.	.	-47025.58	MLE