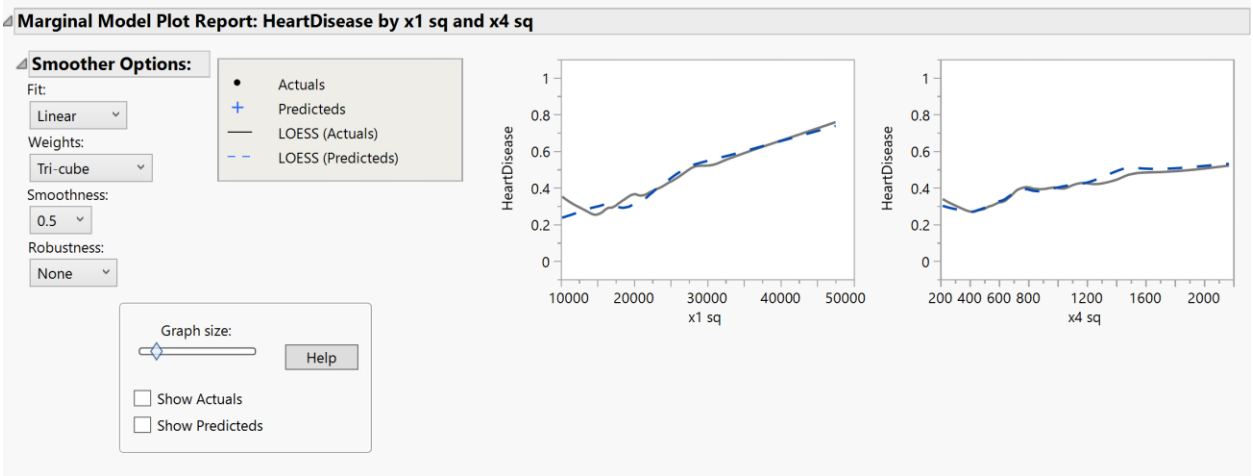
Exercise 4

a)
Based on the marginal plots and the summary output, we can tell that:
X2 and x5 are valid predictors – reasonable marginal plots and significant p-values.
For the exact opposite reasons, x1 and x4 cannot be included in the model – along the tails they are not aligned.
Hence, this model cannot be considered valid

b)
From Fig.8.18, the kernel density estimate, we see that both x1 and x4 are skewed towards the right.
Hence, I'd recommend trying a log transformation – log(x1) and log(x4)

c)
Yes, the model is valid from the marginal plot as all predictor are properly specified.



d)
The logit transformation of this variable(x3) has a linear relationship with the Y. The coefficient here represents the change in the log-odds for one unit change in this variable x3 or whenever x3=1 here (when the patient has family history, the log-odds of the patient having the same is increased by 0.941)

Exercise 5

a)
Nominal Logistic Fit for TARGET_B
Effect Summary

Source	LogWorth	PValue
totalmonths	3.442	0.00036
NUMCHLD	1.626	0.02367
Icmed	1.152	0.07040

Source	LogWorth	PValue
Icavg	1.087	0.08194
LASTGIFT	0.708	0.19608
IC15	0.631	0.23398
INCOME	0.576	0.26528
gender	0.381	0.41558
homeowner	0.364	0.43261
HV	0.242	0.57235
RAMNTALL	0.238	0.57874
WEALTH	0.231	0.58766
TIMELAG	0.125	0.74954
AVGGIFT	0.111	0.77450
MAXRAMNT	0.054	0.88342
NUMPROM	0.025	0.94333

Converged in Gradient, 3 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	21.0895	16	42.17893	0.0004*
Full	1060.2201			
Reduced	1081.3096			

RSquare (U)	0.0195
AICc	2154.84
BIC	2245.43
Observations (or Sum Wgts)	1560

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	1543	1060.2201	2120.44
Saturated	1559	0.0000	**Prob>ChiSq**
Fitted	16	1060.2201	<.0001*

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.7945455	0.6286885	8.15	0.0043*
homeowner[0]	0.05063279	0.0645332	0.62	0.4327
gender[0]	0.04328971	0.0531747	0.66	0.4156

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
NUMCHLD	0.33869019	0.1538529	4.85	0.0277*
INCOME	-0.0400956	0.0360149	1.24	0.2656
WEALTH	-0.013642	0.0251622	0.29	0.5877
HV	-0.0000492	8.7159e-5	0.32	0.5724
Icmed	-0.0023858	0.0013237	3.25	0.0715
Icavg	0.00248868	0.0014341	3.01	0.0827
IC15	-0.0073824	0.0062146	1.41	0.2349
NUMPROM	0.00026327	0.0037067	0.01	0.9434
RAMNTALL	-0.0004237	0.0007744	0.30	0.5843
MAXRAMNT	-0.0017079	0.0116696	0.02	0.8836
LASTGIFT	0.01677852	0.0130627	1.65	0.1990
totalmonths	0.04908645	0.0139036	12.46	0.0004*
TIMELAG	-0.0029872	0.0093686	0.10	0.7498
AVGGIFT	0.00509292	0.0177836	0.08	0.7746

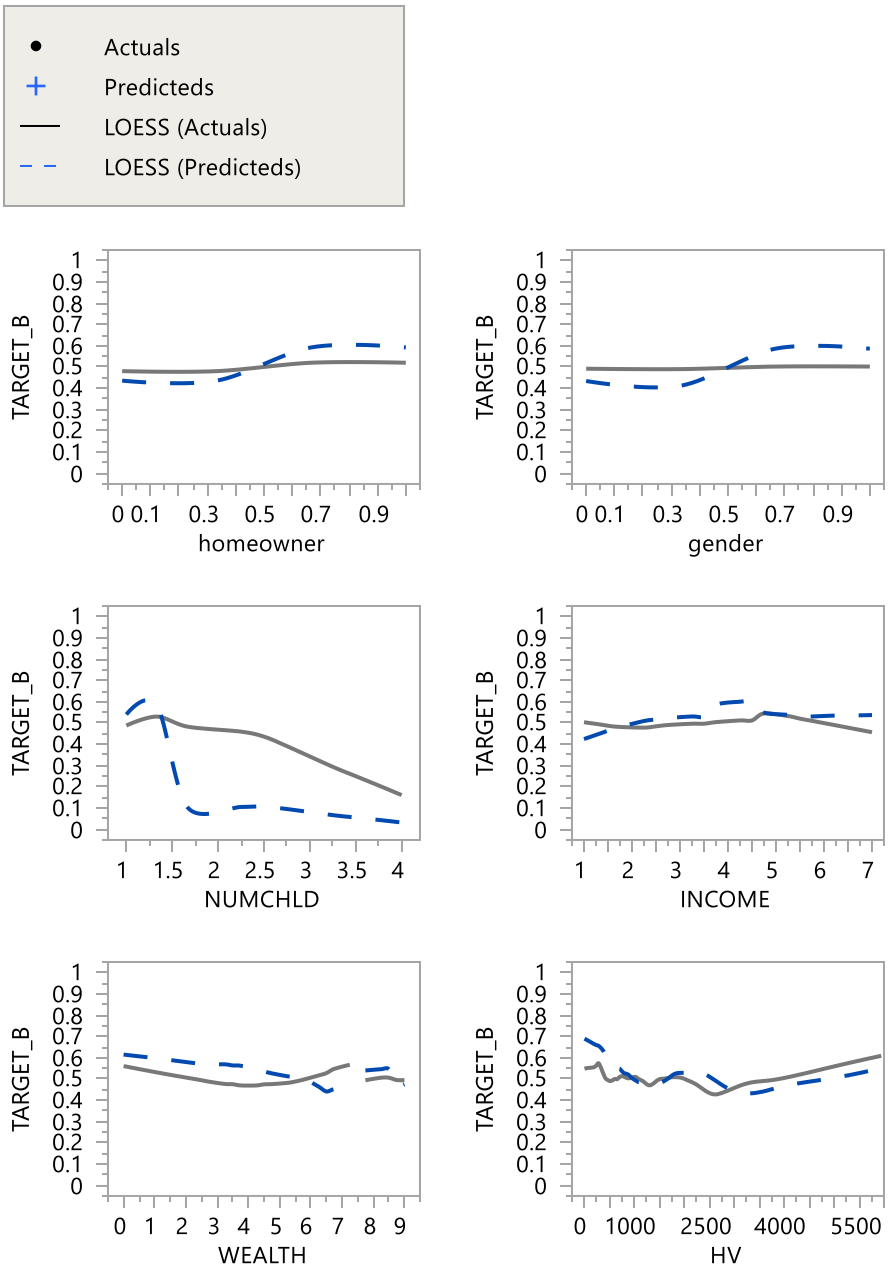
For log odds of 0/1

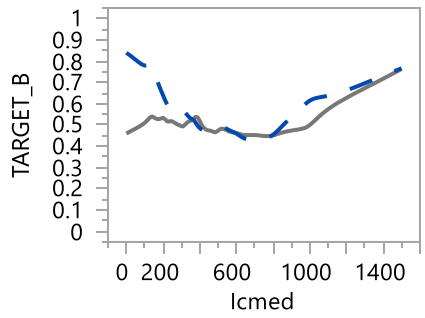
Effect Likelihood Ratio Tests

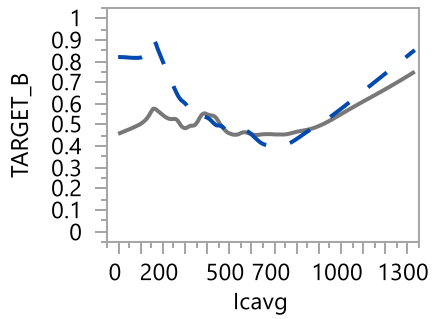
Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
homeowner	1	1	0.61579382	0.4326
gender	1	1	0.66278386	0.4156
NUMCHLD	1	1	5.11846201	0.0237*
INCOME	1	1	1.24097833	0.2653
WEALTH	1	1	0.29400731	0.5877
HV	1	1	0.31876709	0.5723
Icmed	1	1	3.27377074	0.0704
Icavg	1	1	3.02602505	0.0819
IC15	1	1	1.41649262	0.2340
NUMPROM	1	1	0.0050532	0.9433
RAMNTALL	1	1	0.30827307	0.5787
MAXRAMNT	1	1	0.02150186	0.8834
LASTGIFT	1	1	1.67132156	0.1961
totalmonths	1	1	12.7198486	0.0004*
TIMELAG	1	1	0.10192213	0.7495
AVGGIFT	1	1	0.08207763	0.7745

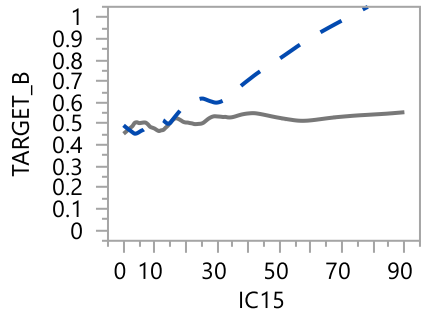
b)

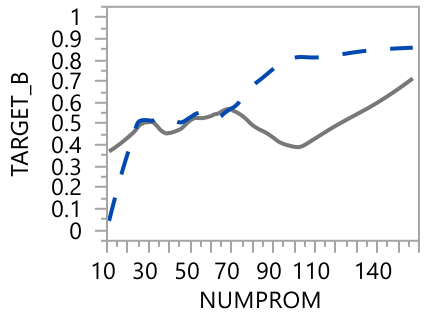
Marginal Model Plot Report: TARGET_B by homeowner, gender, NUMCHLD, INCOME, WEALTH, HV, Icmed, Icavg, IC15, NUMPROM, RAMNTALL, MAXRAMNT, LASTGIFT, totalmonths, TIMELAG and AVGGIFT

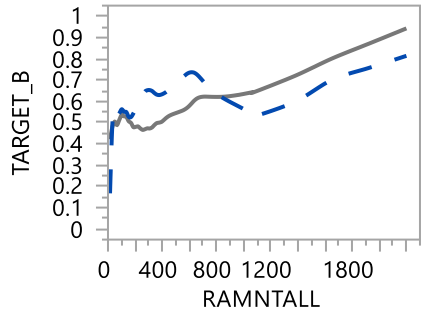

Actuals
Predicteds
LOESS (Actuals)
LOESS (Predicteds)
TARGET_B
homeowner
gender
NUMCHLD
INCOME
WEALTH
HV

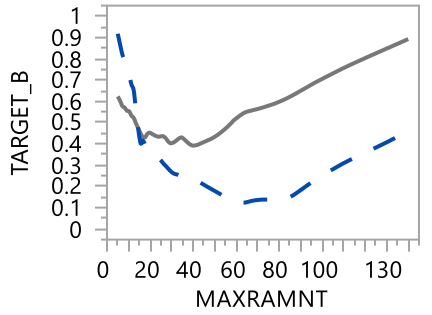

TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 200 600 1000 1400
lcmed

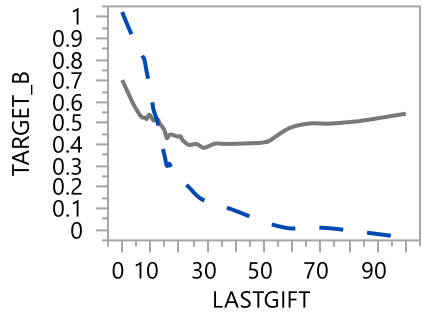

TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 200 500 700 1000 1300
lcavg

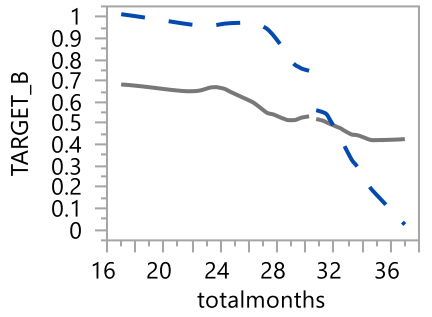

TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 30 50 70 90
IC15

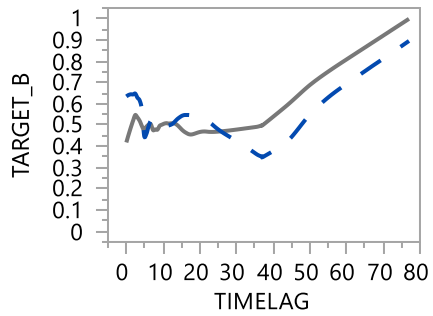

TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10 30 50 70 90 110 140
NUMPROM

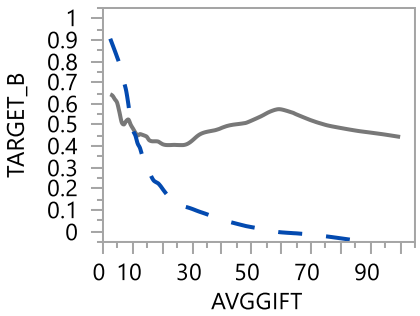

TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 400 800 1200 1800
RAMNTALL


TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 20 40 60 80 100 130
MAXRAMNT


TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 30 50 70 90
LASTGIFT


TARGET_B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
16 20 24 28 32 36
totalmonths





c)

which predictor variables may benefit from being transformed	Reasonable transformation
NUMCHILD	Square transformation
IC15	Log
Lcavg	Log
NUMPROM	Log
LASTGIFT	Log
Totalmonths	Log
AVGGIFT	Log
Wealth	Square
HV	Log
Lcmed	Log
Lcavg	log

Variables that do not require transformation – Home owner, gender, income

d)

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.0621077	0.1369726	0.21	0.6502
WEALTH*zipconvert_2[0]	0.00599654	0.0121727	0.24	0.6223
WEALTH*zipconvert_3[0]	-0.0071289	0.0119135	0.36	0.5496
WEALTH*zipconvert_4[0]	0.0149102	0.0119961	1.54	0.2139
WEALTH*zipconvert_5[0]	0.01173042	0.0121261	0.94	0.3334

For log odds of 0/1

Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
WEALTH*zipconvert_2	1	1	0.24274927	0.6222
WEALTH*zipconvert_3	1	1	0.35804821	0.5496
WEALTH*zipconvert_4	1	1	1.54628617	0.2137
WEALTH*zipconvert_5	1	1	0.936366	0.3332

Source	LogWorth	PValue
WEALTH*zipconvert_4	0.670	0.21368
WEALTH*zipconvert_5	0.477	0.33321
WEALTH*zipconvert_3	0.260	0.54959
WEALTH*zipconvert_2	0.206	0.62223

Remove Add Edit ☐ FDR

e)

Logistic Regression Model using part c) and d)

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	18.4719	18	36.94382	0.0053*
Full	993.5230			
Reduced	1011.9949			

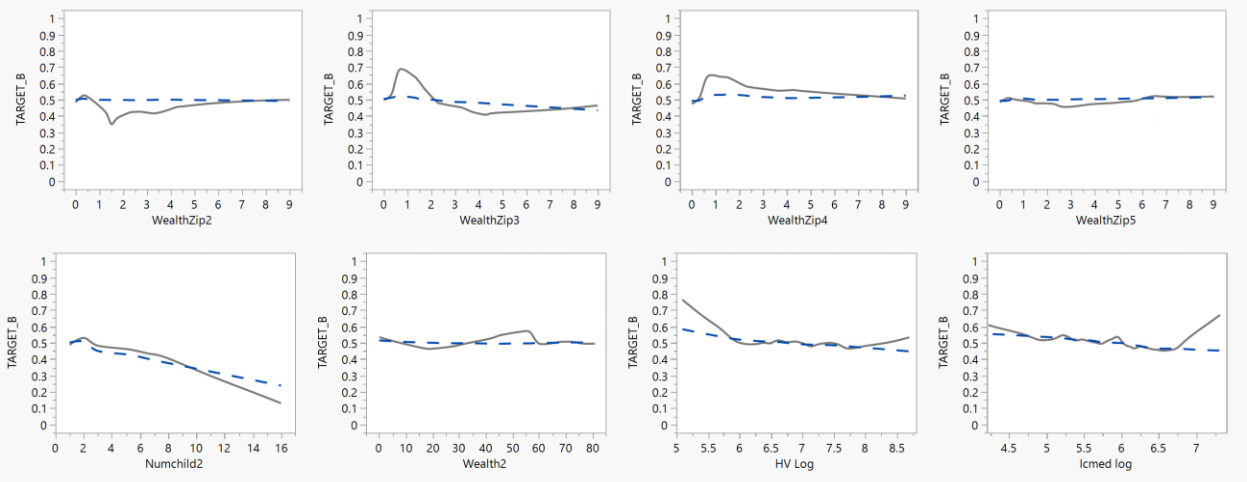
RSquare (U)	0.0183
AICc	2025.57
BIC	2125.48
Observations (or Sum Wgts)	1460

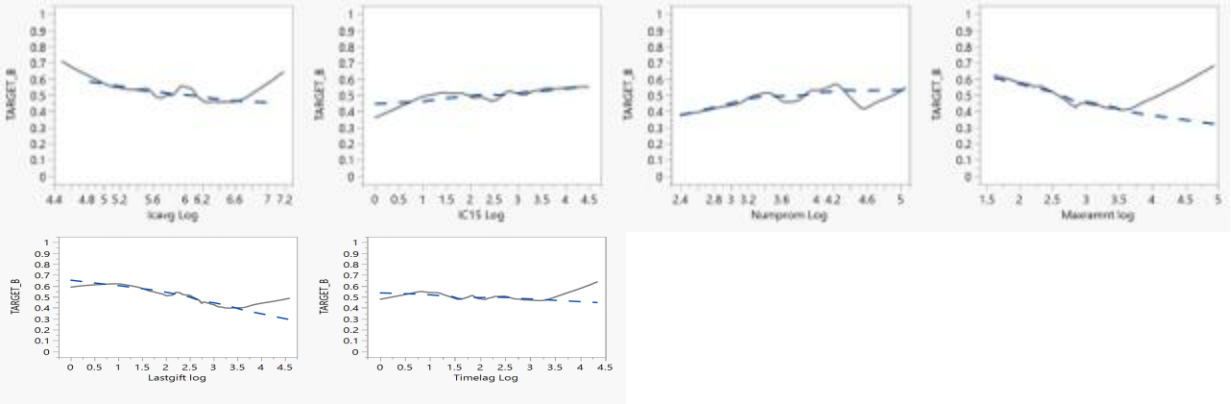
Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
homeowner	1	1	0.55478655	0.4564
gender	1	1	0.30145592	0.5830
INCOME	1	1	1.16792809	0.2798
WealthZip2	1	1	0.05720587	0.8110
WealthZip3	1	1	0.36131224	0.5478
WealthZip4	1	1	0.01599377	0.8994
WealthZip5	1	1	0.00070847	0.9788
Numchild2	1	1	3.96130417	0.0466*
Wealth2	1	1	0.29609731	0.5863
HV Log	1	1	0.67053459	0.4129
Icmed log	1	1	0.74515028	0.3880
Icavg Log	1	1	0.42937242	0.5123
IC15 Log	1	1	1.0858197	0.2974
Numprom Log	1	1	3.10684068	0.0780
Lastgift log	1	1	0.85917631	0.3540
Timelag Log	1	1	0.16776787	0.6821
AvgGift Log	1	1	0.02728313	0.8688
Maxramnt log	1	1	0.73974642	0.3897

f)




TARGET_B
Icavg Log
IC15 Log
Numprom Log
Maxramnt log
Lastgift log
Timelag Log

g)