Exercice 1.

Exploring Data

a)

Quantiles		
100.0%	maximum	4
99.5%		4
97.5%		3,85
90.0%		2,4
75.0%	quartile	2
50.0%	median	2
25.0%	quartile	0
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	1,4222222
Std Dev	1,0332845
Std Err Mean	0,154033
Upper 95% Mean	1,7326552
Lower 95% Mean	1,1117892
N	45

Quantiles		
100.0%	maximum	80
99.5%		80
97.5%		79,7
90.0%		67,8
75.0%	quartile	23
50.0%	median	17
25.0%	quartile	11,5
10.0%		4
2.5%		3
0.5%		3
0.0%	minimum	3

Summary Statistics	
Mean	22,133333
Std Dev	20,304052
Std Err Mean	3,0267494
Upper 95% Mean	28,233346
Lower 95% Mean	16,033321
N	45

Quantiles		
100.0%	maximum	412
99.5%		412
97.5%		410,65
90.0%		255,2
75.0%	quartile	154
50.0%	median	98
25.0%	quartile	61,5
10.0%		32,6
2.5%		16,6
0.5%		16
0.0%	minimum	16

Summary Statistics	
Mean	122,02222
Std Dev	91,611951
Std Err Mean	13,656703
Upper 95% Mean	149,5455
Lower 95% Mean	94,498945
N	45

e)

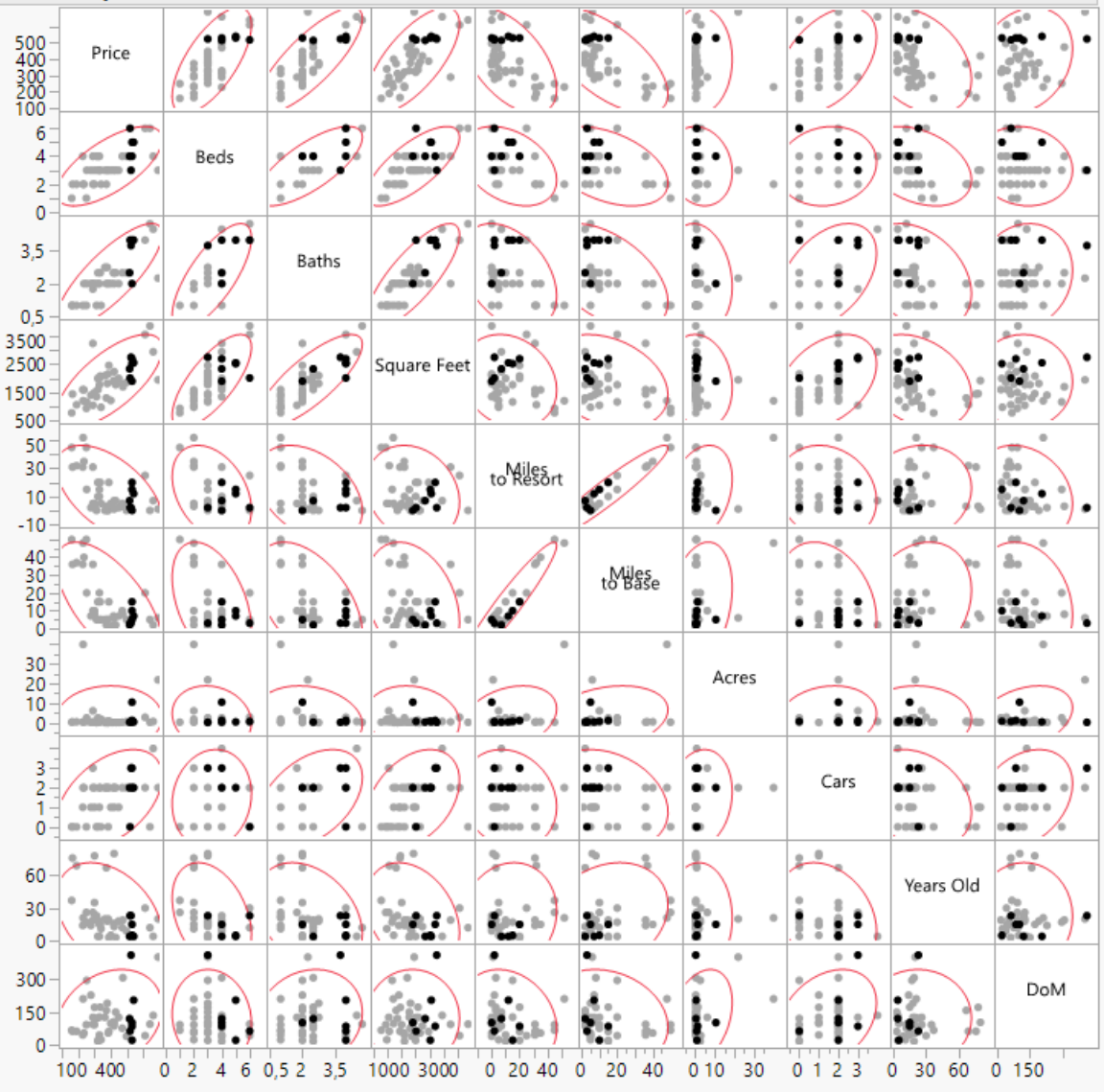
Multivariate

Correlations

	Price	Beds	Baths	Square Feet	Miles to Resort	Miles to Base	Acres	Cars	Years Old	DoM
Beds	0,6753	1,0000	0,7332	0,7282	-0,3509	-0,4241	-0,1473	0,1045	-0,3403	-0,0971
Baths	0,8001	0,7332	1,0000	0,7901	-0,3745	-0,4880	-0,1930	0,4357	-0,3267	0,1205
Square Feet	0,6970	0,7282	0,7901	1,0000	-0,1895	-0,2972	-0,1456	0,3728	-0,3037	-0,0110
Miles to Resort	-0,5391	-0,3509	-0,3745	-0,1895	1,0000	0,9480	0,2958	-0,1584	0,1082	-0,2219
Miles to Base	-0,6332	-0,4241	-0,4880	-0,2972	0,9480	1,0000	0,2634	-0,2612	0,2211	-0,1841
Acres	0,0251	-0,1473	-0,1930	-0,1456	0,2958	0,2634	1,0000	0,1474	-0,0295	0,3288
Cars	0,4523	0,1045	0,4357	0,3728	-0,1584	-0,2612	0,1474	1,0000	-0,2714	0,3137
Years Old	-0,3551	-0,3403	-0,3267	-0,3037	0,1082	0,2211	-0,0295	-0,2714	1,0000	0,0077
DoM	0,2298	-0,0971	0,1205	-0,0110	-0,2219	-0,1841	0,3288	0,3137	0,0077	1,0000

The correlations are estimated by Row-wise method.

Scatterplot Matrix



Build the Model

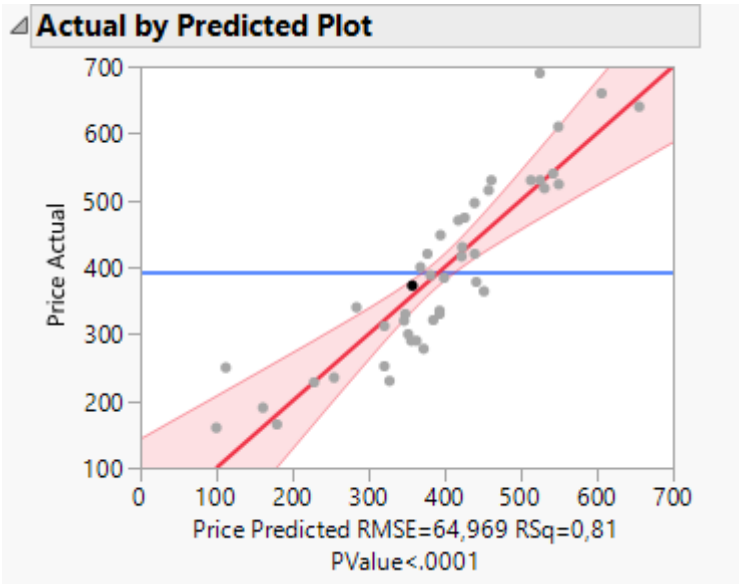
b)

Effect Summary

Source	LogWorth	PValue
Acres	1,880	0,01318
Baths	1,846	0,01424
Square Feet	0,982	0,10425
Miles to Base	0,319	0,48023
Miles to Resort	0,285	0,51915
Beds	0,265	0,54336
Cars	0,209	0,61769
Years Old	0,207	0,62093
DoM	0,092	0,80958

Remove Add Edit ☐ FDR

c)



d)

Summary of Fit

RSquare	0,809043
RSquare Adj	0,759939
Root Mean Square Error	64,96897
Mean of Response	391,1911
Observations (or Sum Wgts)	45

e)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	9	625914,08	69546,0	16,4763
Error	35	147733,84	4221,0	Prob > F
C. Total	44	773647,92		<,0001*

g)

Effect Summary

Source	LogWorth	PValue
Square Feet	2,067	0,00856
Acres	1,179	0,06629
Beds	1,062	0,08675
Miles to Base	0,667	0,21549
Cars	0,619	0,24055
DoM	0,398	0,39962
Years Old	0,156	0,69775
Miles to Resort	0,056	0,87910

Remove Add Edit Undo ☐ FDR

h)

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	176,73111	55,12519	3,21	0,0028*	.
Beds	27,266996	15,48496	1,76	0,0868	2,9086042
Square Feet	0,0729189	0,026219	2,78	0,0086*	2,889324
Miles to Resort	-0,421878	2,753966	-0,15	0,8791	13,324826
Miles to Base	-3,470846	2,752906	-1,26	0,2155	13,713111
Acres	3,5179615	1,857516	1,89	0,0663	1,4282824
Cars	15,61035	13,08159	1,19	0,2406	1,6460007
Years Old	-0,230635	0,589127	-0,39	0,6977	1,289
DoM	0,1186562	0,139201	0,85	0,3996	1,4650622

j)

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	177,27617	54,27945	3,27	0,0024*	.
Beds	27,464191	15,22637	1,80	0,0794	2,8885039
Square Feet	0,0721643	0,02541	2,84	0,0073*	2,7873441
Miles to Base	-3,867308	0,925871	-4,18	0,0002*	1,5931999
Acres	3,456705	1,789866	1,93	0,0611	1,3620917
Cars	15,260838	12,70997	1,20	0,2375	1,5959308
Years Old	-0,210345	0,566419	-0,37	0,7125	1,2238454
DoM	0,1245076	0,132079	0,94	0,3520	1,3547481

k)

Stepwise Fit for Price

Stepwise Regression Control

Stopping Rule: Minimum BIC

Direction: Forward

Enter All | Make Model | Remove All | Run Model

Go | Stop | Step

SSE	DFE	RMSE	RSquare	RSquare Adj	Cp	p	AICc	BIC
153342,58	40	61,915785	0,8018	0,7820	1,9193847	5	507,9345	516,564

Current Estimates

Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio"	"Prob>F"
☑	☑	Intercept	197,150171	1	0	0,000	1
☐	☐	Beds	0	1	1148,056	0,294	0,59063
☐	☑	Baths	59,2080973	1	46387,91	12,100	0,00123
☐	☑	Square Feet	0,05112326	1	19597,28	5,112	0,02927
☐	☑	Miles to Base	-3,7985104	1	90774,52	23,679	1,81e-5
☐	☑	Acres	5,00610917	1	46854,78	12,222	0,00117
☐	☐	Cars	0	1	311,3432	0,079	0,77968
☐	☐	Years Old	0	1	969,519	0,248	0,62118
☐	☐	DoM	0	1	408,8354	0,104	0,7485

Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p	AICc	BIC	
1	Baths	Entered	0,0000	495319	0,6402	26,012	2	527,135	531,97	○
2	Miles to Base	Entered	0,0015	59830,74	0,7176	13,607	3	518,659	524,885	○
3	Acres	Entered	0,0021	45558,33	0,7765	4,6377	4	510,675	518,169	○
4	Square Feet	Entered	0,0293	19597,28	0,8018	1,9194	5	507,935	516,564	○
5	Beds	Entered	0,5906	1148,056	0,8033	3,643	6	510,413	520,032	○
6	Cars	Entered	0,5668	1325,569	0,8050	5,3238	7	512,992	523,445	○
7	DoM	Entered	0,6605	795,2096	0,8060	7,1324	8	515,897	527,014	○
8	Years Old	Entered	0,7181	549,7775	0,8067	9	9	519,06	530,656	○
9	Best	Specific	.	.	0,8018	1,9194	5	507,935	516,564	◉

l)

Stepwise Fit for Price

Stepwise Regression Control

Stopping Rule: Minimum AICc

Direction: Forward

Enter All | Make Model | Remove All | Run Model

Go | Stop | Step

SSE	DFE	RMSE	RSquare	RSquare Adj	Cp	p	AICc	BIC
153342,58	40	61,915785	0,8018	0,7820	1,9193847	5	507,9345	516,564

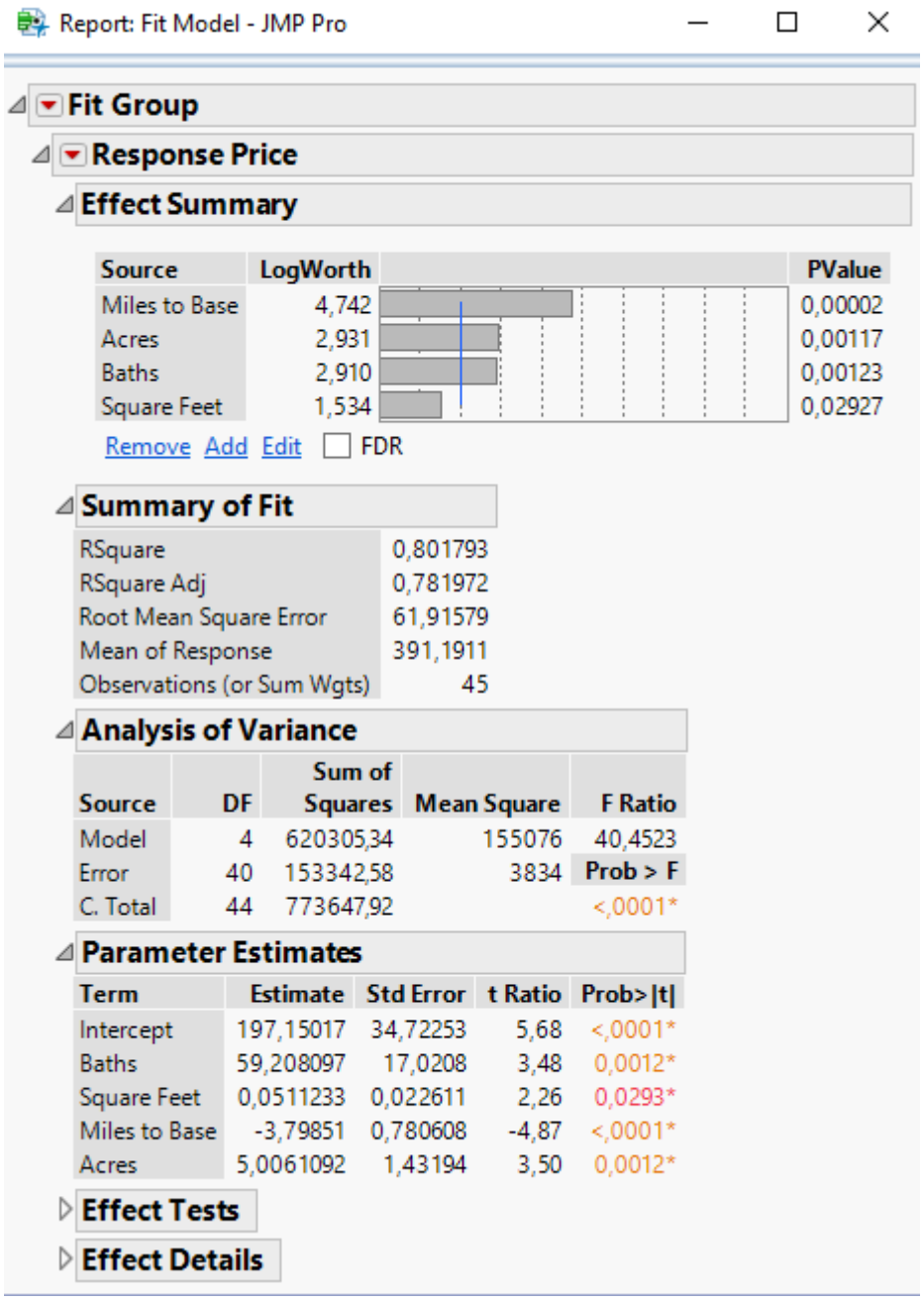
Current Estimates

Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio"	"Prob> F"
☑	☑	Intercept	197,150171	1	0	0,000	1
☐	☐	Beds	0	1	1148,056	0,294	0,59063
☐	☑	Baths	59,2080973	1	46387,91	12,100	0,00123
☐	☑	Square Feet	0,05112326	1	19597,28	5,112	0,02927
☐	☑	Miles to Base	-3,7985104	1	90774,52	23,679	1,81e-5
☐	☑	Acres	5,00610917	1	46854,78	12,222	0,00117
☐	☐	Cars	0	1	311,3432	0,079	0,77968
☐	☐	Years Old	0	1	969,519	0,248	0,62118
☐	☐	DoM	0	1	408,8354	0,104	0,7485

Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p	AICc	BIC	
1	Baths	Entered	0,0000	495319	0,6402	26,012	2	527,135	531,97	○
2	Miles to Base	Entered	0,0015	59830,74	0,7176	13,607	3	518,659	524,885	○
3	Acres	Entered	0,0021	45558,33	0,7765	4,6377	4	510,675	518,169	○
4	Square Feet	Entered	0,0293	19597,28	0,8018	1,9194	5	507,935	516,564	○
5	Beds	Entered	0,5906	1148,056	0,8033	3,643	6	510,413	520,032	○
6	Cars	Entered	0,5668	1325,569	0,8050	5,3238	7	512,992	523,445	○
7	DoM	Entered	0,6605	795,2096	0,8060	7,1324	8	515,897	527,014	○
8	Years Old	Entered	0,7181	549,7775	0,8067	9	9	519,06	530,656	○
9	Best	Specific	.	.	0,8018	1,9194	5	507,935	516,564	○
10	All	Removed	.	.	-0,000	143,27	1	570,84	574,167	○
11	Baths	Entered	0,0000	495319	0,6402	26,012	2	527,135	531,97	○
12	Miles to Base	Entered	0,0015	59830,74	0,7176	13,607	3	518,659	524,885	○
13	Acres	Entered	0,0021	45558,33	0,7765	4,6377	4	510,675	518,169	○
14	Square Feet	Entered	0,0293	19597,28	0,8018	1,9194	5	507,935	516,564	○
15	Beds	Entered	0,5906	1148,056	0,8033	3,643	6	510,413	520,032	○
16	Cars	Entered	0,5668	1325,569	0,8050	5,3238	7	512,992	523,445	○
17	DoM	Entered	0,6605	795,2096	0,8060	7,1324	8	515,897	527,014	○
18	Years Old	Entered	0,7181	549,7775	0,8067	9	9	519,06	530,656	○
19	Best	Specific	.	.	0,8018	1,9194	5	507,935	516,564	◉

m)


Report: Fit Model - JMP Pro
Fit Group
Response Price
Effect Summary
Source LogWorth PValue
Miles to Base 4,742 0,00002
Acres 2,931 0,00117
Baths 2,910 0,00123
Square Feet 1,534 0,02927
Remove Add Edit FDR
Summary of Fit
RSquare 0,801793
RSquare Adj 0,781972
Root Mean Square Error 61,91579
Mean of Response 391,1911
Observations (or Sum Wgts) 45
Analysis of Variance
Source DF Sum of Squares Mean Square F Ratio
Model 4 620305,34 155076 40,4523
Error 40 153342,58 3834 Prob > F
C. Total 44 773647,92 <,0001*
Parameter Estimates
Term Estimate Std Error t Ratio Prob>|t|
Intercept 197,15017 34,72253 5,68 <,0001*
Baths 59,208097 17,0208 3,48 0,0012*
Square Feet 0,0511233 0,022611 2,26 0,0293*
Miles to Base -3,79851 0,780608 -4,87 <,0001*
Acres 5,0061092 1,43194 3,50 0,0012*
Effect Tests
Effect Details

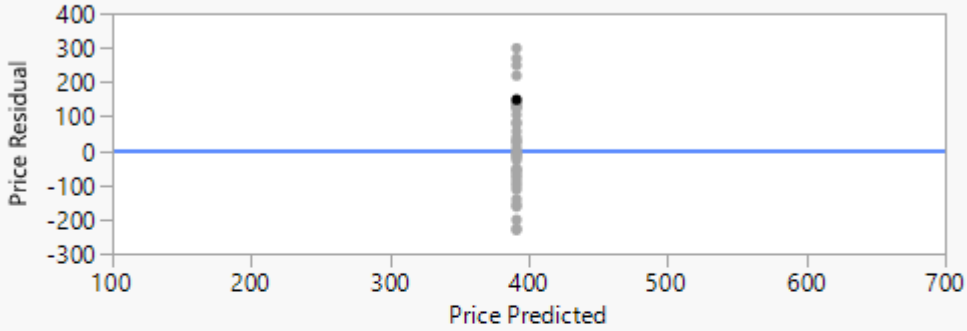
Examining Residuals

a)

Fit Group

Response Price

Residual by Predicted Plot



Summary of Fit

RSquare	0
RSquare Adj	0
Root Mean Square Error	132,6006
Mean of Response	391,1911
Observations (or Sum Wgts)	45

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	0	0,00	0,0	.
Error	44	773647,92	17582,9	Prob > F
C. Total	44	773647,92		.

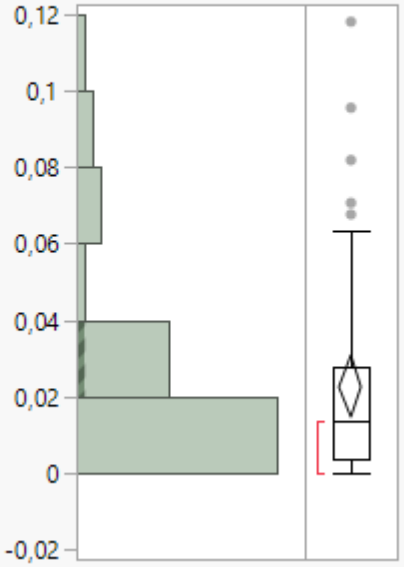
Parameter Estimates

Effect Details

b)

Distributions

Cook's D Influence Price



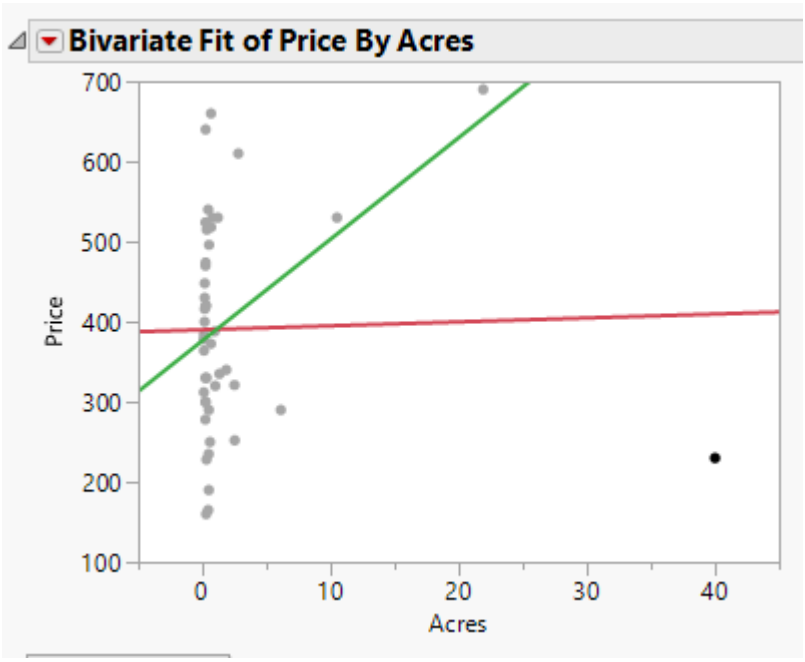
Quantiles

100.0%	maximum	0,1180330154
99.5%		0,1180330154
97.5%		0,1146563737
90.0%		0,0688436618
75.0%	quartile	0,0278132698
50.0%	median	0,0135363385
25.0%	quartile	0,003819107
10.0%		0,0003661547
2.5%		0,0000216966
0.5%		0,0000134617
0.0%	minimum	0,0000134617

Summary Statistics

Mean	0,0227273
Std Dev	0,0275851
Std Err Mean	0,0041121
Upper 95% Mean	0,0310148
Lower 95% Mean	0,0144398
N	45

d)



i)

Fit Group

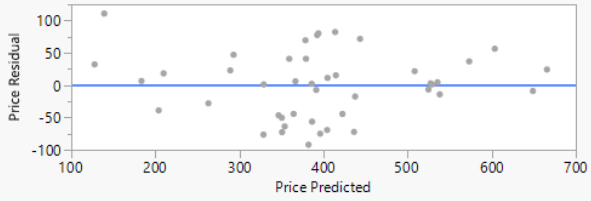
Response Price

Effect Summary

Source	LogWorth	PValue
Acres	5,687	0,00000
Baths	4,041	0,00009
Miles to Base	4,030	0,00009
Square Feet	1,720	0,01905

Remove Add Edit FDR

Residual by Predicted Plot



Summary of Fit

RSquare	0,854004
RSquare Adj	0,83903
Root Mean Square Error	52,88352
Mean of Response	394,8545
Observations (or Sum Wgts)	44

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	638004,83	159501	57,0326
Error	39	109070,00	2797	Prob > F
C. Total	43	747074,83		<,0001*

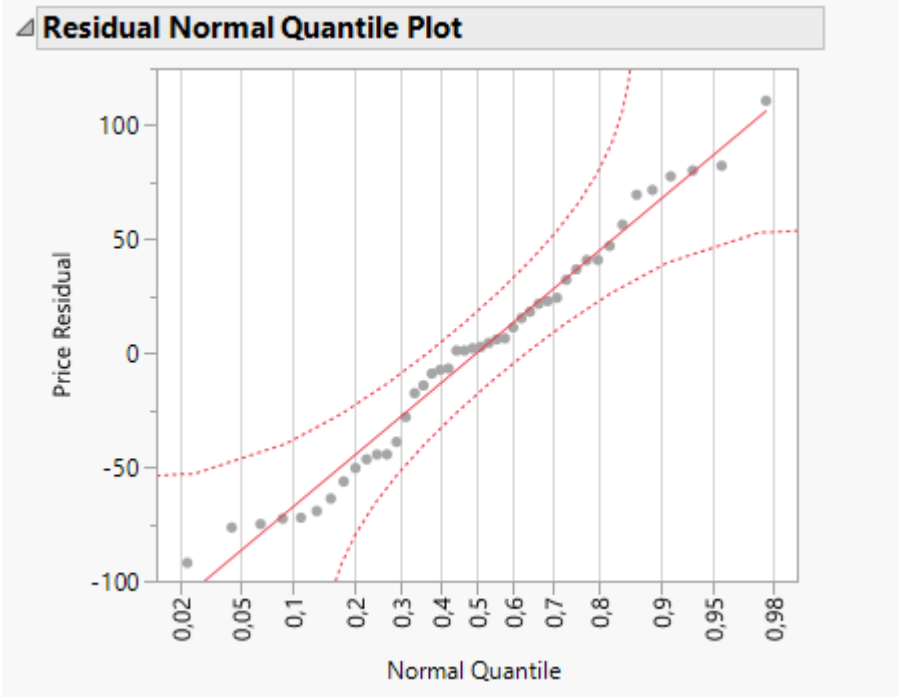
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	175,46632	30,1538	5,82	<,0001*
Baths	63,61719	14,57999	4,36	<,0001*
Square Feet	0,0472992	0,019337	2,45	0,0191*
Miles to Base	-3,025002	0,694498	-4,36	<,0001*
Acres	12,462947	2,237932	5,57	<,0001*

Effect Tests

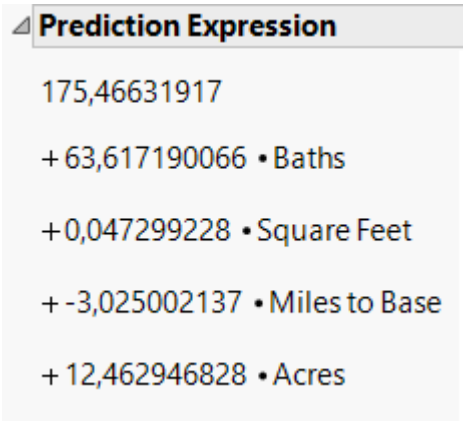
Effect Details

k)

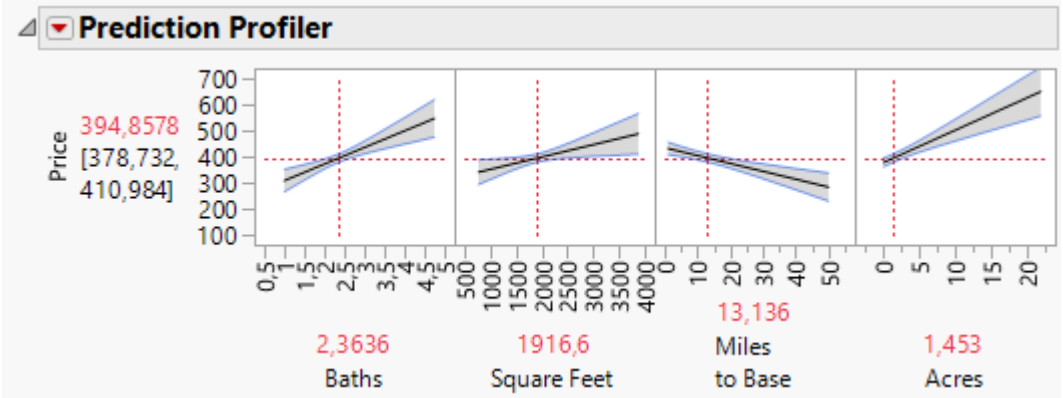


Predicted response

a)



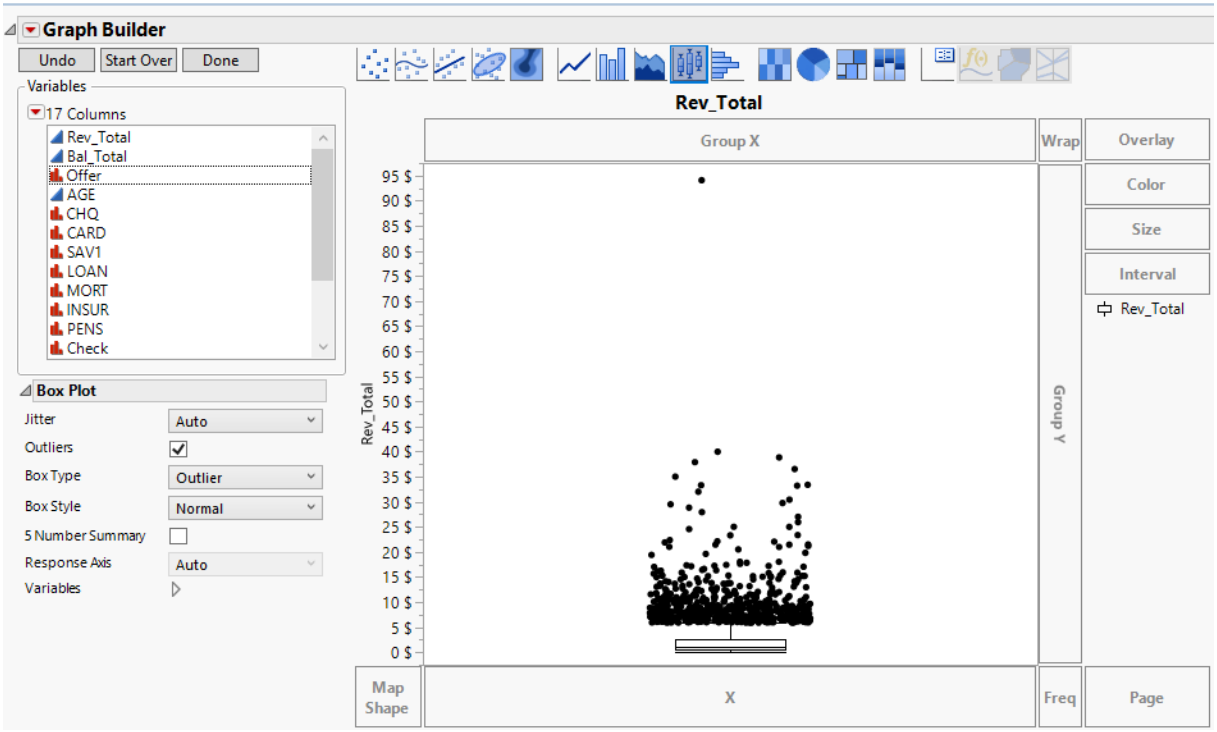
b)



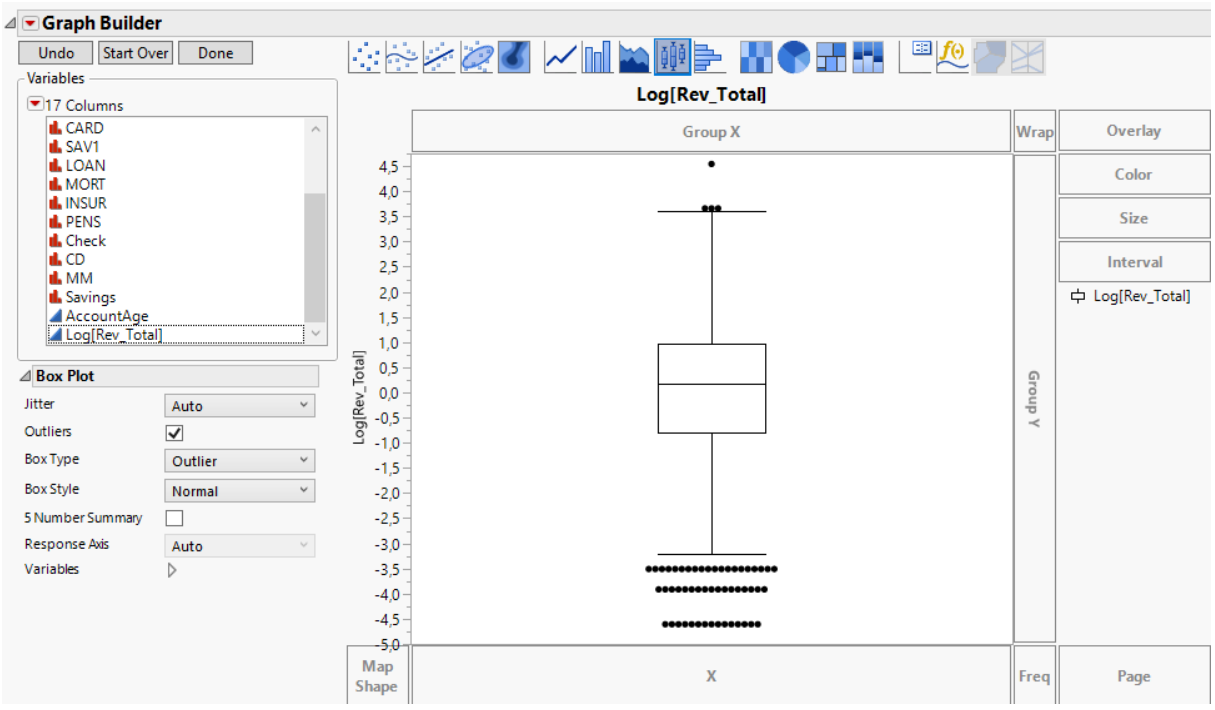
Exercice 2

Prepare for Modeling

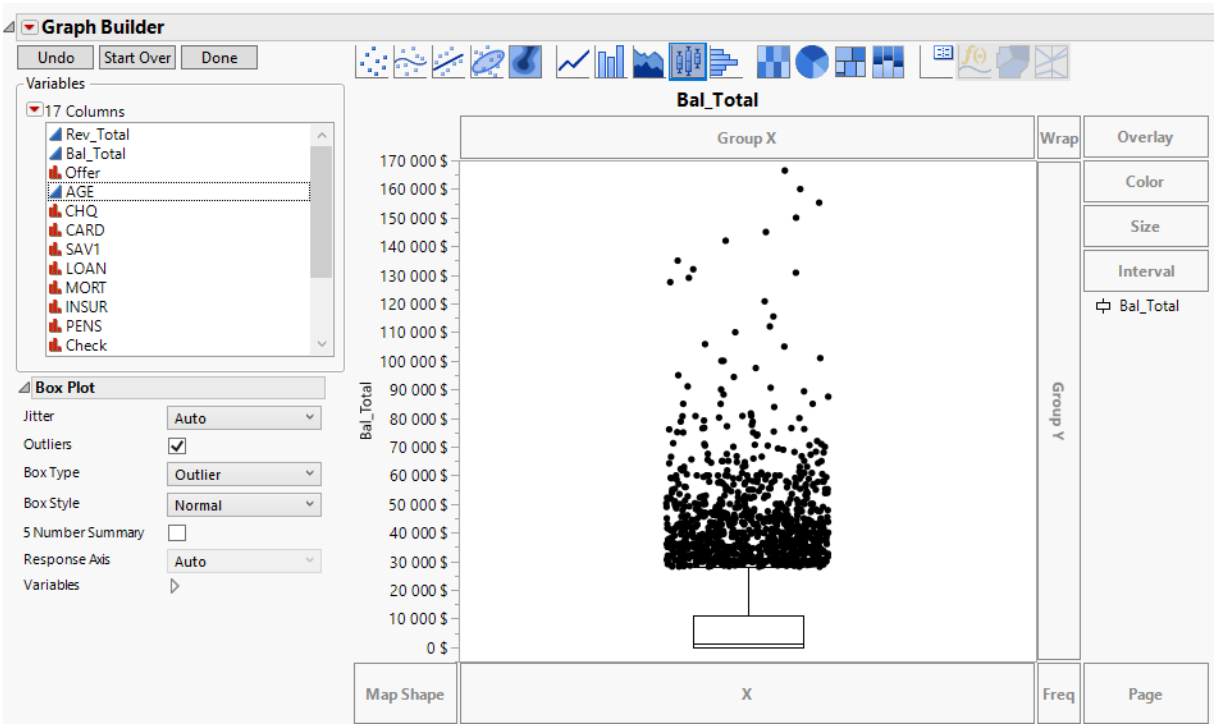
b)

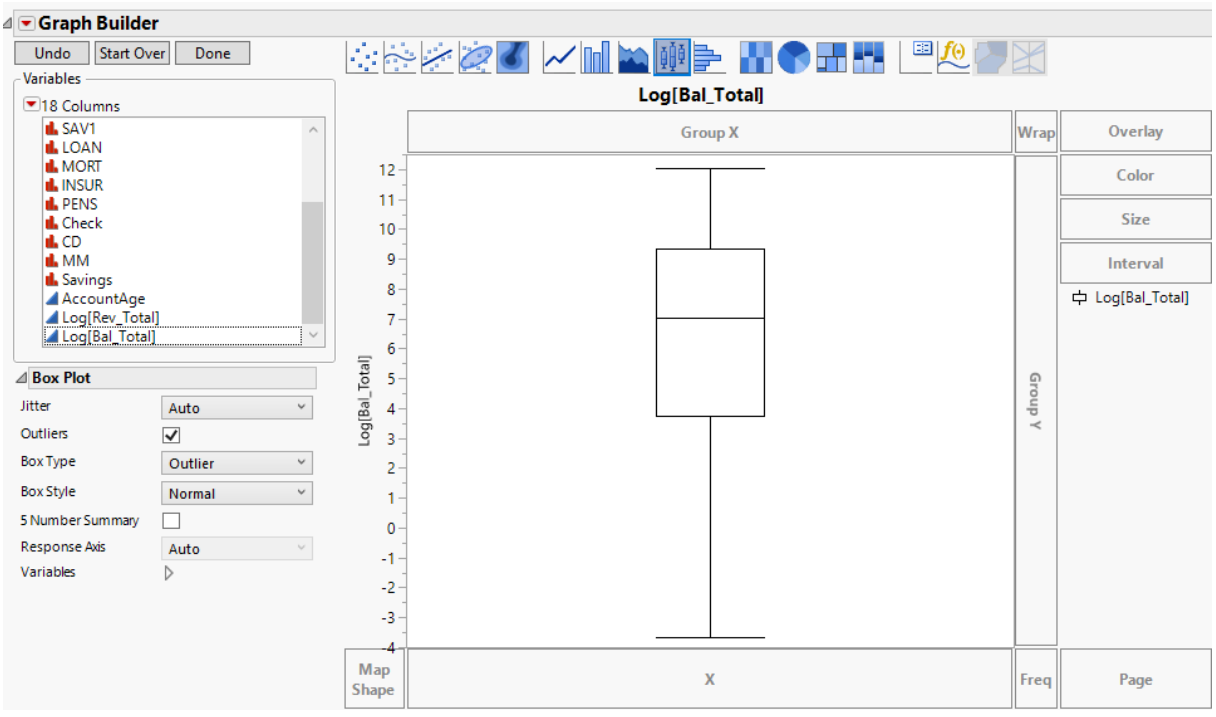


d)

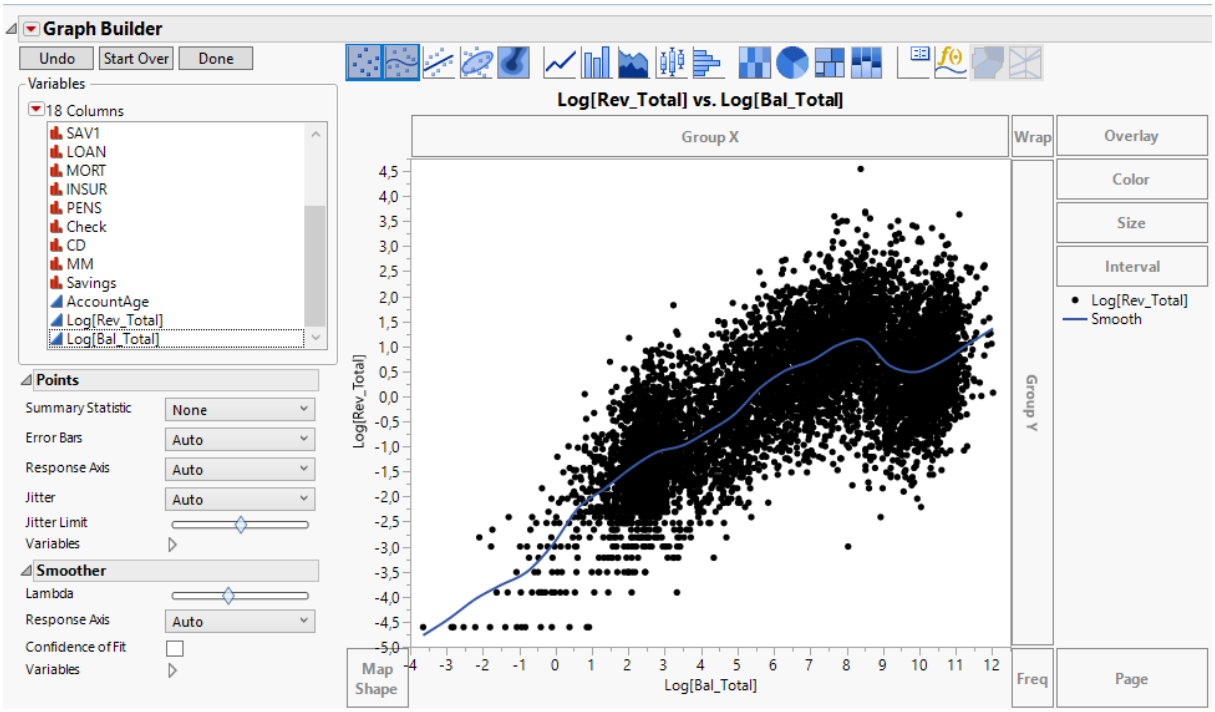


e)

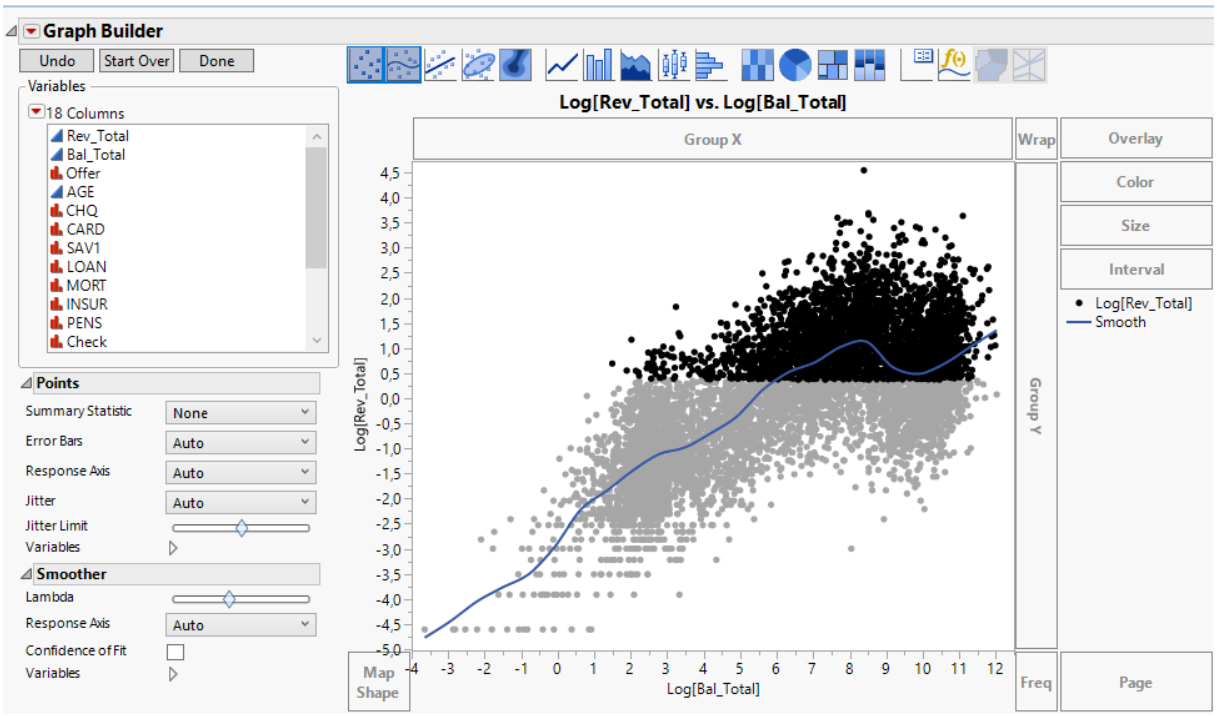

Graph Builder
Undo
Start Over
Done
Variables
17 Columns
Rev_Total
Bal_Total
Offer
AGE
CHQ
CARD
SAV1
LOAN
MORT
INSUR
PENS
Check
Box Plot
Jitter
Auto
Outliers
Box Type
Outlier
Box Style
Normal
5 Number Summary
Response Axis
Auto
Variables
Bal_Total
Group X
Wrap
Overlay
Color
Size
Interval
Bal_Total
170 000 $
160 000 $
150 000 $
140 000 $
130 000 $
120 000 $
110 000 $
100 000 $
90 000 $
80 000 $
70 000 $
60 000 $
50 000 $
40 000 $
30 000 $
20 000 $
10 000 $
0 $
Group Y
Map Shape
X
Freq
Page


Graph Builder
Undo
Start Over
Done
Variables
18 Columns
SAV1
LOAN
MORT
INSUR
PENS
Check
CD
MM
Savings
AccountAge
Log[Rev_Total]
Log[Bal_Total]
Box Plot
Jitter
Auto
Outliers
Box Type
Outlier
Box Style
Normal
5 Number Summary
Response Axis
Auto
Variables
Log[Bal_Total]
Group X
Wrap
Overlay
Color
Size
Interval
Log[Bal_Total]
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3
-4
Group Y
Map Shape
X
Freq
Page

f)


Graph Builder
Undo
Start Over
Done
Variables
18 Columns
SAV1
LOAN
MORT
INSUR
PENS
Check
CD
MM
Savings
AccountAge
Log[Rev_Total]
Log[Bal_Total]
Points
Summary Statistic
None
Error Bars
Auto
Response Axis
Auto
Jitter
Auto
Jitter Limit
Variables
Smoother
Lambda
Response Axis
Auto
Confidence of Fit
Variables
Log[Rev_Total] vs. Log[Bal_Total]
Group X
Wrap
Overlay
Color
Size
Interval
Log[Rev_Total]
Smooth
Group Y
Log[Rev_Total]
Log[Bal_Total]
Map Shape
Freq
Page

h)


Graph Builder
Undo
Start Over
Done
Variables
18 Columns
Rev_Total
Bal_Total
Offer
AGE
CHQ
CARD
SAV1
LOAN
MORT
INSUR
PENS
Check
Points
Summary Statistic
None
Error Bars
Auto
Response Axis
Auto
Jitter
Auto
Jitter Limit
Variables
Smoother
Lambda
Response Axis
Auto
Confidence of Fit
Variables
Log[Rev_Total] vs. Log[Bal_Total]
Group X
Wrap
Overlay
Color
Size
Interval
Log[Rev_Total]
Smooth
Group Y
Log[Rev_Total]
Log[Bal_Total]
Map Shape
Freq
Page

Build the Model

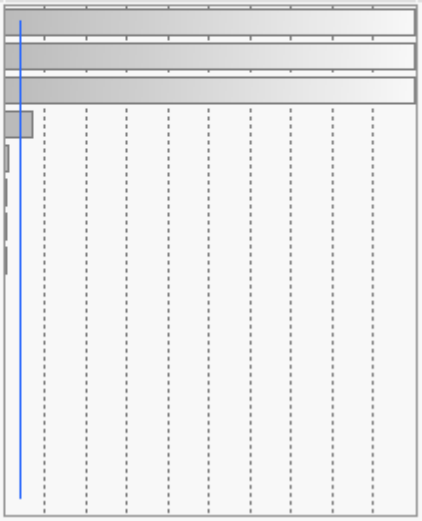
a)

Singularity Details

Term	Details
LOAN[0]	=CD[0]
INSUR[0]	=MM[0]=Savings[0]

Effect Summary

Source	LogWorth	PValue
Log[Bal_Total]	1184,077	0,00000
CARD	158,527	0,00000
Check	121,653	0,00000
Offer	3,508	0,00031
AGE	0,601	0,25064
MORT	0,440	0,36279
AccountAge	0,429	0,37271
SAV1	0,401	0,39731
CHQ	0,217	0,60672
PENS	0,011	0,97514
Savings	.	.
MM	.	.
CD	.	.
INSUR	.	.
LOAN	.	.



Remove Add Edit ☐ FDR

Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	7399	4857,3230	0,65648	0,6077
Pure Error	8	8,6419	1,08023	Prob > F
Total Error	7407	4865,9648		0,8935
				Max RSq
				0,9993

Summary of Fit

RSquare	0,598864
RSquare Adj	0,598214
Root Mean Square Error	0,810519
Mean of Response	0,059558
Observations (or Sum Wgts)	7420

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	12	7264,490	605,374	921,5041
Error	7407	4865,965	0,657	Prob > F
C. Total	7419	12130,455		<,0001*

Parameter Estimates

Term		Estimate	Std Error	t Ratio	Prob>|t|
Intercept		-2,515773	0,04768	-52,76	<,0001*
Offer[0]		-0,069321	0,019212	-3,61	0,0003*
AGE		-0,000526	0,000458	-1,15	0,2506
CHQ[0]		-0,00545	0,010587	-0,51	0,6067
CARD[0]		-0,790471	0,028682	-27,56	<,0001*
SAV1[0]		0,0107845	0,01274	0,85	0,3973
LOAN[0]	Biased	0,0596287	0,018157	3,28	0,0010*
MORT[0]		0,0154212	0,016944	0,91	0,3628
INSUR[0]	Biased	0,0359458	0,013843	2,60	0,0094*
PENS[0]		-0,000298	0,009562	-0,03	0,9751
Check[0]		0,6886466	0,028726	23,97	<,0001*
CD[0]	Zeroed	0	0	.	.
MM[0]	Zeroed	0	0	.	.
Savings[0]	Zeroed	0	0	.	.
AccountAge		-0,002321	0,002604	-0,89	0,3727
Log[Bal_Total]		0,4421557	0,004931	89,67	<,0001*

Effect Tests

Effect Details

b)

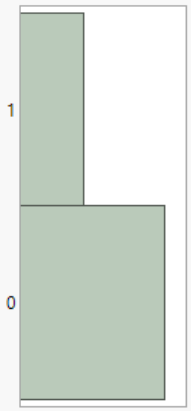
Singularity Details

Term	Details
LOAN[0]	=CD[0]
INSUR[0]	=MM[0]=Savings[0]

c)

Distributions

INSUR

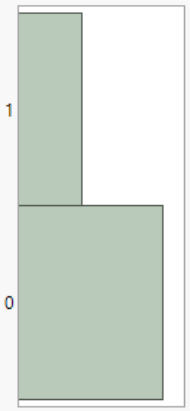


Frequencies

Level	Count	Prob
0	5170	0,69677
1	2250	0,30323
Total	7420	1,00000

N Missing 0
2 Levels

MM

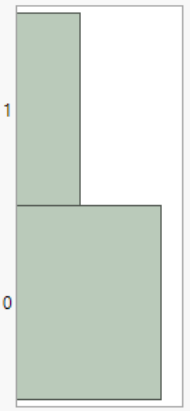


Frequencies

Level	Count	Prob
0	5170	0,69677
1	2250	0,30323
Total	7420	1,00000

N Missing 0
2 Levels

Savings



Frequencies

Level	Count	Prob
0	5170	0,69677
1	2250	0,30323
Total	7420	1,00000

N Missing 0
2 Levels

e)

Effect Summary

Source	LogWorth	PValue
Log[Bal_Total]	1184,077	0,00000
CARD	158,527	0,00000
Check	121,653	0,00000
Offer	3,508	0,00031
LOAN	2,988	0,00103
INSUR	2,025	0,00943
AGE	0,601	0,25064
MORT	0,440	0,36279
AccountAge	0,429	0,37271
SAV1	0,401	0,39731
CHQ	0,217	0,60672
PENS	0,011	0,97514

Remove Add Edit Undo ☐ FDR

g)

Stepwise Regression Control

Stopping Rule: Minimum AICc

Direction: Forward

Rules: Combine

Enter All | Make Model | Remove All | Run Model

Go | Stop | Step

SSE	DFE	RMSE	RSquare	RSquare Adj	Cp	p	AICc	BIC
4868,873	7413	0,8104332	0,5986	0,5983	5,4268869	7	17946,9	18002,17

h)

Current Estimates

Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio"	"Prob>F"
☑	☑	Intercept	-2,538458	1	0	0,000	1
☐	☑	Offer{0-1}	-0,0699765	1	8,731206	13,294	0,00027
☐	☐	AGE	0	1	1,263258	1,924	0,1655
☐	☐	CHQ{0-1}	0	1	0,00549	0,008	0,92716
☐	☑	CARD{0-1}	-0,7963998	1	733,8014	1117,234	3e-228
☐	☐	SAV1{0-1}	0	1	0,662701	1,009	0,31518
☐	☑	LOAN{0-1}	0,05720648	1	7,111245	10,827	0,001
☐	☐	MORT{0-1}	0	1	0,629	0,958	0,32781
☐	☑	INSUR{1-0}	-0,0400496	1	5,899812	8,983	0,00273
☐	☐	PENS{0-1}	0	1	0,001181	0,002	0,96618
☐	☑	Check{0-1}	0,70491156	1	622,6609	948,019	5e-196
☐	☐	AccountAge	0	1	0,758714	1,155	0,2825
☐	☑	Log[Bal_Total]	0,44240234	1	5303,953	8075,421	0

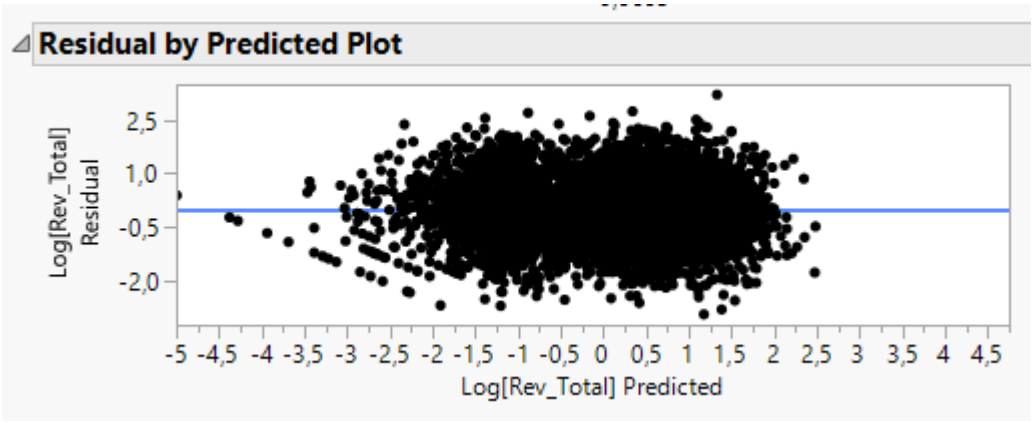
i)

Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p	AICc	BIC	
1	Log[Bal_Total]	Entered	0,0000	5260,963	0,4337	3040,8	2	20491,1	20511,8	○
2	CARD{0-1}	Entered	0,0000	1074,107	0,5222	1407,8	3	19231,4	19259,1	○
3	Check{0-1}	Entered	0,0000	905,034	0,5969	32,121	4	17973,5	18008,1	○
4	Offer{0-1}	Entered	0,0001	10,21159	0,5977	18,577	5	17960	18001,5	○
5	LOAN{0-1}	Entered	0,0043	5,366693	0,5981	12,408	6	17953,9	18002,2	○
6	INSUR{1-0}	Entered	0,0027	5,899812	0,5986	5,4269	7	17946,9	18002,2	○
7	AGE	Entered	0,1655	1,263258	0,5987	5,5039	8	17947	18009,2	○
8	MORT{0-1}	Entered	0,3403	0,59734	0,5988	6,5947	9	17948,1	18017,2	○
9	AccountAge	Entered	0,3844	0,496928	0,5988	7,8382	10	17949,3	18025,3	○
10	SAV1{0-1}	Entered	0,4506	0,373804	0,5988	9,2692	11	17950,8	18033,7	○
11	CHQ{0-1}	Entered	0,6045	0,176237	0,5989	11,001	12	17952,5	18042,3	○
12	PENS{0-1}	Entered	0,9751	0,000638	0,5989	13	13	17954,5	18051,2	○
13	Best	Specific	.	.	0,5986	5,4269	7	17946,9	18002,2	◉

Examining Residuals

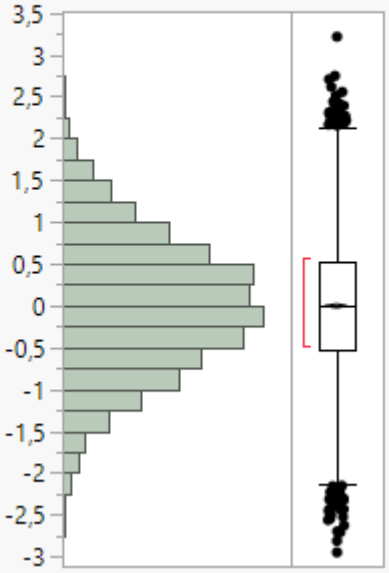
a)



c)

Distributions

Residual Log[Rev_Total]



Quantiles

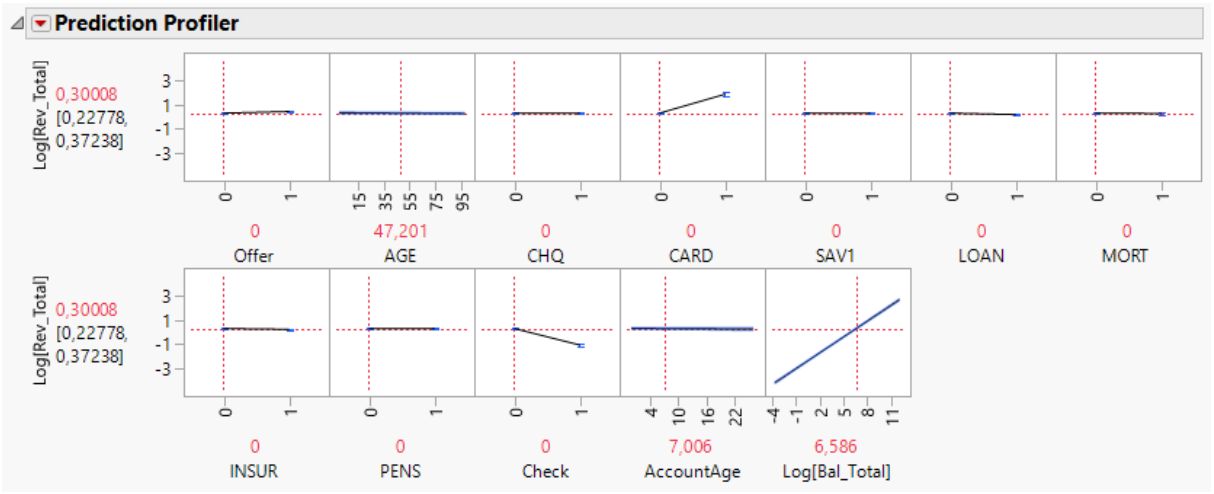
100.0%	maximum	3,2136027531
99.5%		2,0529756348
97.5%		1,5879487032
90.0%		1,0448271584
75.0%	quartile	0,5340866984
50.0%	median	-0,003960858
25.0%	quartile	-0,538273325
10.0%		-1,038397356
2.5%		-1,603561729
0.5%		-2,101350258
0.0%	minimum	-2,948163877

Summary Statistics

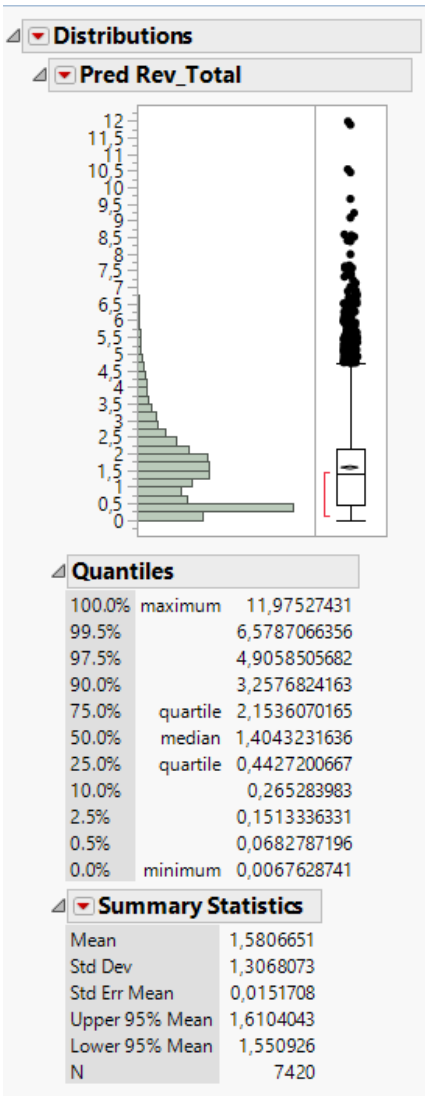
Mean	8,44e-16
Std Dev	0,8098634
Std Err Mean	0,0094018
Upper 95% Mean	0,0184301
Lower 95% Mean	-0,01843
N	7420

Predicted response

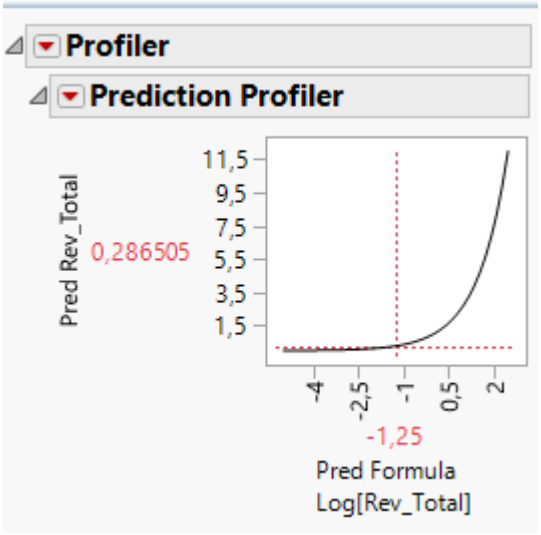
a)



e)



f)


Profiler
Prediction Profiler
Pred Rev_Total
0,286505
11,5
9,5
7,5
5,5
3,5
1,5
-4
-2,5
-1
0,5
2
-1,25
Pred Formula
Log[Rev_Total]

Exercise 3

a)

Fit Group

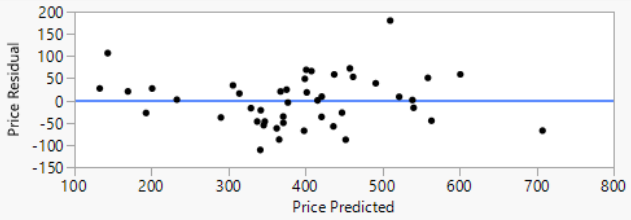
Response Price

Effect Summary

Source	LogWorth	PValue
Acres	1,934	0,01163
Baths	1,537	0,02902
Square Feet	1,421	0,03792
Miles to Base	0,820	0,15140
Miles to Base*Miles to Resort	0,464	0,34386
Beds*Baths	0,281	0,52355
Miles to Resort	0,266	0,54170 ^
Beds	0,221	0,60051 ^

Remove Add Edit Undo FDR ('^' denotes effects with containing effects above them)

Residual by Predicted Plot



Summary of Fit

RSquare	0,815712
RSquare Adj	0,77476
Root Mean Square Error	62,93155
Mean of Response	391,1911
Observations (or Sum Wgts)	45

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	8	631074,24	78884,3	19,9184
Error	36	142573,68	3960,4	Prob > F
C. Total	44	773647,92		<,0001*

Parameter Estimates

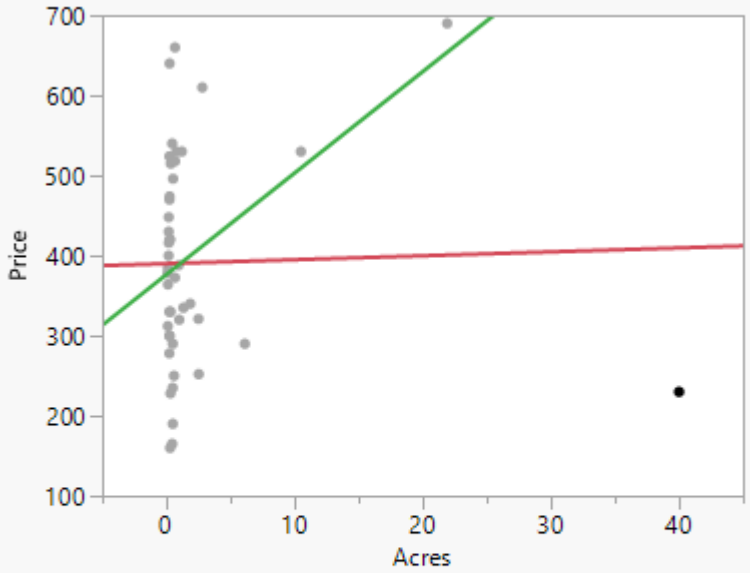
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	192,96208	47,84171	4,03	0,0003*
Baths	47,844064	21,0391	2,27	0,0290*
Square Feet	0,0557743	0,025881	2,16	0,0379*
Miles to Base	-3,845811	2,62382	-1,47	0,1514
Acres	4,4090811	1,658362	2,66	0,0116*
(Beds-3,28889)*(Baths-2,33333)	5,7057536	8,857551	0,64	0,5235
Beds	7,4444377	14,0906	0,53	0,6005
(Miles to Base-13,9111)*(Miles to Resort-12,4)	0,0675319	0,070404	0,96	0,3439
Miles to Resort	-1,42622	2,31491	-0,62	0,5417

Effect Tests

Effect Details

b)

Bivariate Fit of Price By Acres



Linear Fit
Linear Fit

Linear Fit

Price = 390,05851 + 0,4904919*Acres

Summary of Fit

RSquare	0,000629
RSquare Adj	-0,02261
Root Mean Square Error	134,0914
Mean of Response	391,1911
Observations (or Sum Wgts)	45

Lack Of Fit

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	486,40	486,4	0,0271
Error	43	773161,51	17980,5	Prob > F
C. Total	44	773647,92		0,8701

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	390,05851	21,14205	18,45	<,0001*
Acres	0,4904919	2,982185	0,16	0,8701

Linear Fit

Price = 376,45798 + 12,665452*Acres

Summary of Fit

RSquare	0,122152
RSquare Adj	0,101251
Root Mean Square Error	124,9588
Mean of Response	394,8545
Observations (or Sum Wgts)	44

Lack Of Fit

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	91256,84	91256,8	5,8443
Error	42	655817,99	15614,7	Prob > F
C. Total	43	747074,83		0,0200*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	376,45798	20,31719	18,53	<,0001*
Acres	12,665452	5,23908	2,42	0,0200*

c) 1)

Wybrano poszczególne interakcje i efekt kwadratowy aby polepszyć dopasowanie modelu i uzupełnić go o brakujące ważne zmienne

c) 2)

Efekt kwadratowy polepsza model, przez co jest bardziej dopasowany do danych wyjściowych i ułatwia interpretacje

c) 3)

Poprawia nasz model który pokazuje wartości prognozowane od rzeczywistych. Przykładem może być gdy przewidywana wycena domu wynosi 200 tys. a model pomylił się o około 30 tyś. Natomiast poprzedni model nie jest interpretowalny ponieważ dla wszystkich domów przewidział tę samą cene, różniącą się tylko błedem w obliczeniach.

c) 4)

Po dodaniu acres * cars, zmniejszyła się cena pomyłki dla niektórych przypadków.

Fit Group

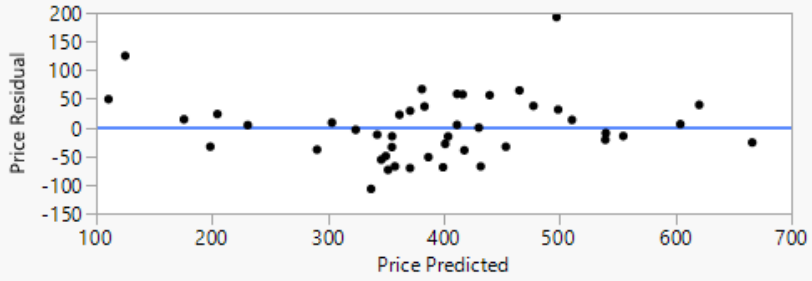
Response Price

Effect Summary

Source	LogWorth	PValue
Miles to Base	4,553	0,00003
Beds*Baths	1,314	0,04856
Square Feet	1,233	0,05850
Acres	1,019	0,09583
Cars	0,544	0,28553
Beds	0,501	0,31558 ^
Acres*Cars	0,476	0,33406
Baths	0,102	0,78983 ^

Remove Add Edit Undo FDR ('^' denotes effects with containing effects above them)

Residual by Predicted Plot



Summary of Fit

RSquare	0,825232
RSquare Adj	0,786394
Root Mean Square Error	61,28468
Mean of Response	391,1911
Observations (or Sum Wgts)	45

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	8	638438,70	79804,8	21,2484
Error	36	135209,22	3755,8	Prob > F
C. Total	44	773647,92		<,0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	243,67382	51,71507	4,71	<,0001*
Baths	7,8038406	29,06264	0,27	0,7898
Square Feet	0,0495306	0,025347	1,95	0,0585
Miles to Base	-4,845225	1,010285	-4,80	<,0001*
Acres	8,7155841	5,096098	1,71	0,0958
(Beds-3,28889)*(Baths-2,33333)	23,478754	11,49928	2,04	0,0486*
Beds	14,613278	14,35823	1,02	0,3156
(Acres-2,30911)*(Cars-1,42222)	-8,86243	9,051554	-0,98	0,3341
Cars	16,828315	15,52285	1,08	0,2855

Effect Tests

Effect Details

Exercise 4

a)

Fit Group

Response Log[Rev_Total]

Effect Summary

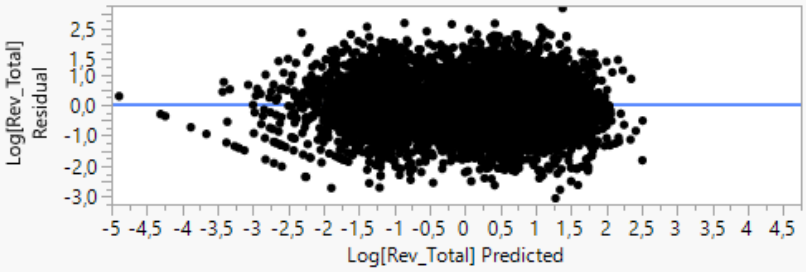
Source	LogWorth	PValue
Log[Bal_Total]	1368,799	0,00000
Check	275,558	0,00000
CARD	235,456	0,00000
Offer	4,083	0,00008

Remove Add Edit ☐ FDR

Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	5992	3932,2562	0,656251	0,9852
Pure Error	1423	947,8833	0,666116	Prob > F
Total Error	7415	4880,1395		0,6431
				Max RSq
				0,9219

Residual by Predicted Plot



Summary of Fit

RSquare	0,597695
RSquare Adj	0,597478
Root Mean Square Error	0,811261
Mean of Response	0,059558
Observations (or Sum Wgts)	7420

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	7250,315	1812,58	2754,075
Error	7415	4880,140	0,66	Prob > F
C. Total	7419	12130,455		<,0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-2,45907	0,025795	-95,33	<,0001*
Offer[0]	-0,075493	0,019166	-3,94	<,0001*
CARD[0]	-0,765025	0,022471	-34,05	<,0001*
Check[0]	0,6434943	0,017353	37,08	<,0001*
Log[Bal_Total]	0,4342763	0,004361	99,57	<,0001*

Effect Tests

Effect Details

b)

Nowy model posiada mniejszy RSquare przez co mniej dokładnie wyjaśnia nasz model. Adjusted R Square jest większy, natomiast RMSE jest mniejszy. P-wartość różni się minimalnie, ale nie przekracza 0.05, przez to dyskwalifikuje to hipoteze zerową i przyjmuje alternatywną.

c)

Pierwszy model jest lepszy, ponieważ posiada większy ARS oraz mniejszy RMSE. ARS zwiększa się wtedy, gdy, poprzez dostosowanie liczby terminów w modelu, nowy termin poprawia dopasowanie modelu bardziej niż przedtem.