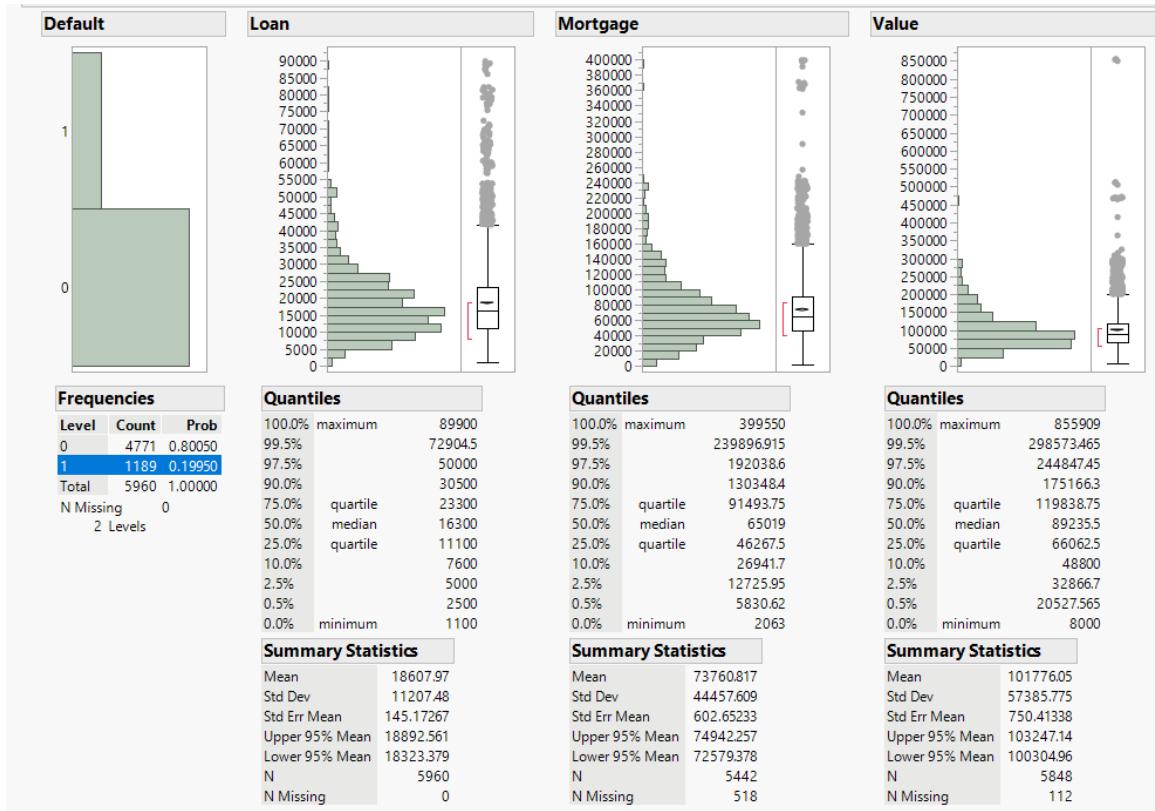


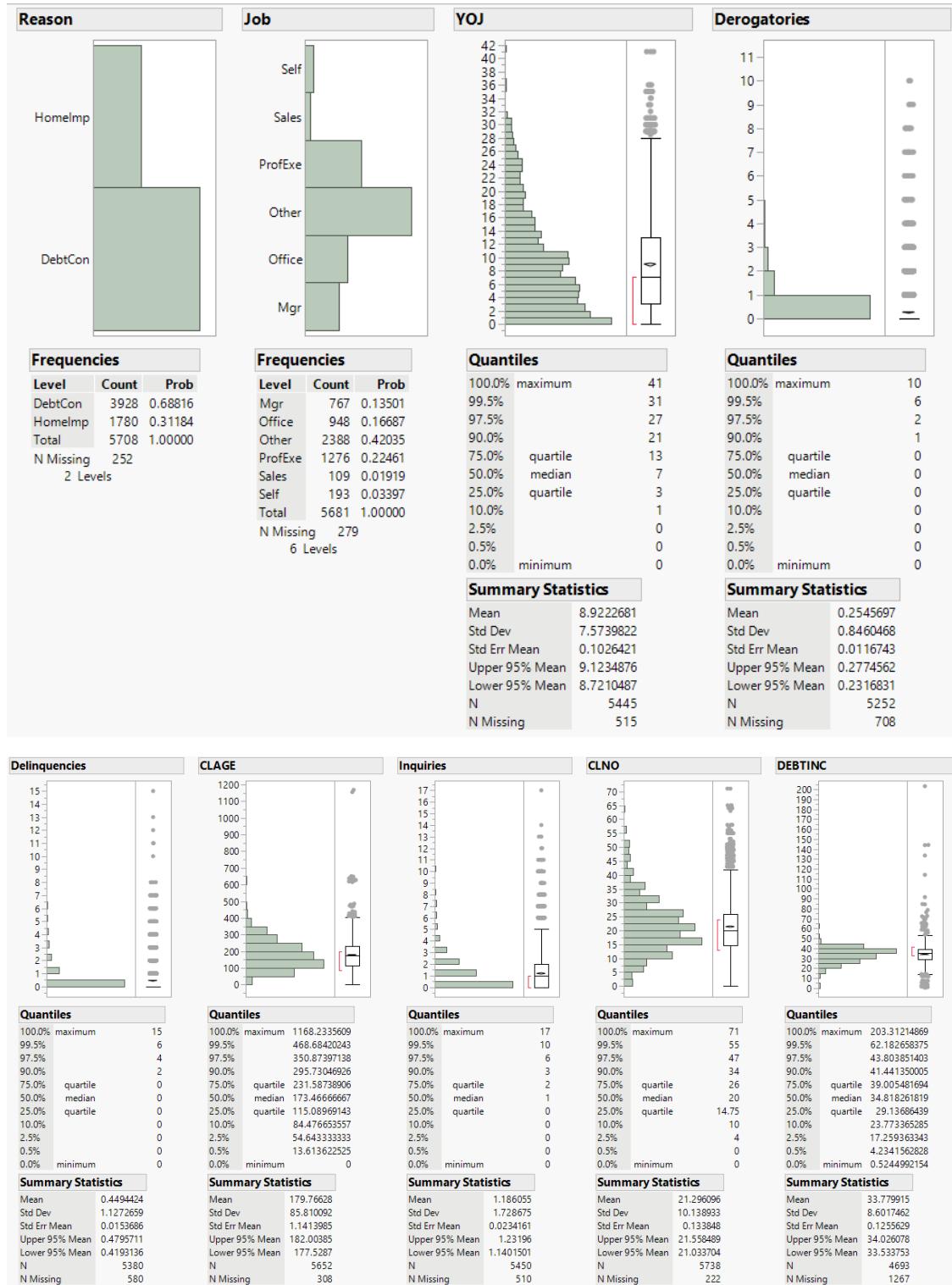
Final

Before I go into the data below, I wanted to state that I am using loan as the dependent variable. Also, note that Logistic regression and Decision Trees were not done as the dependent variable is not categorical. While I do understand that I could have used Decision Tree, it did not make that much sense to do.

Interpreting the Data

The first thing that needs to be done is looking over the data and running a stats summary.

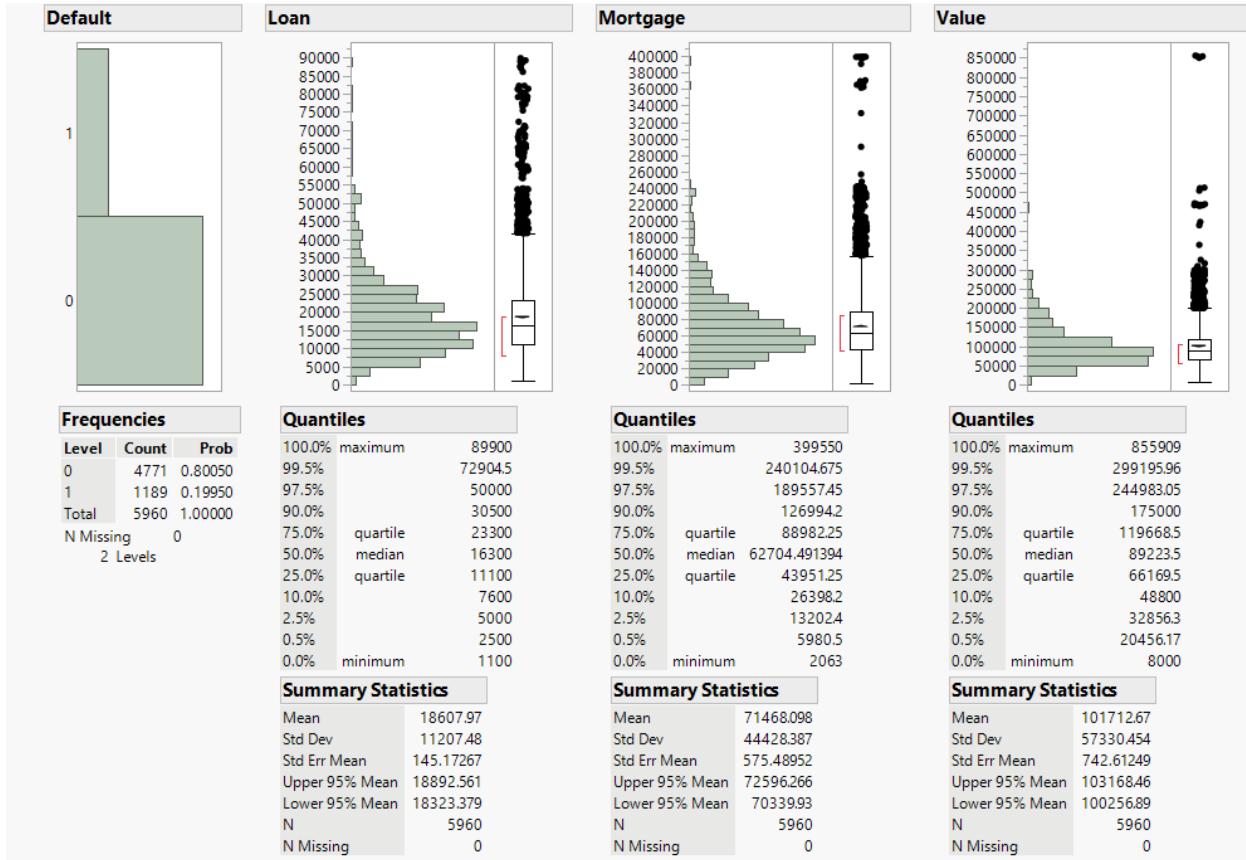


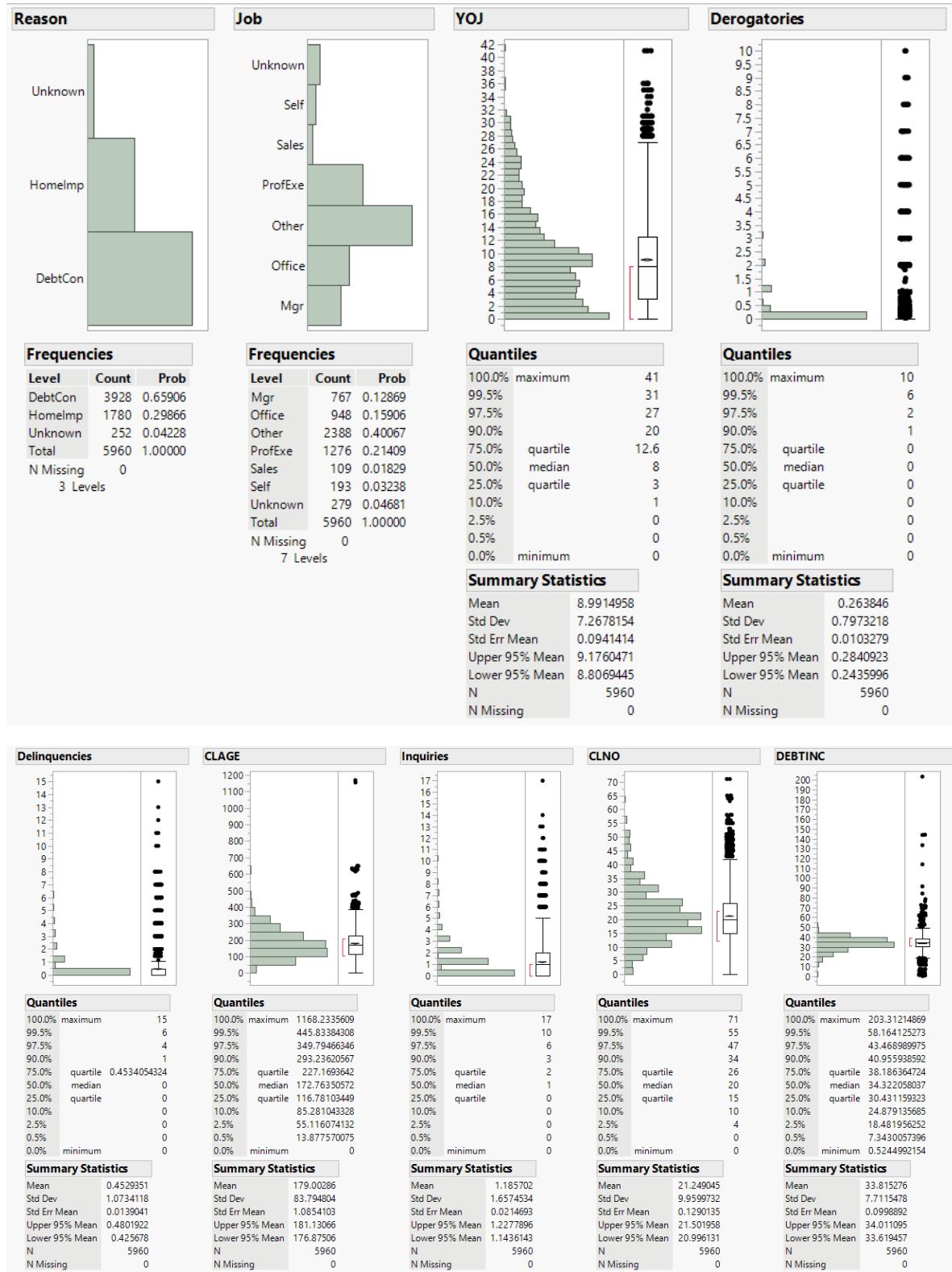


Based on the initial statistics report, it should be noted that there is missing data that needs to be handled before anything can be done with the data.

For the missing continuous variables, the data was imputed.

For the categorical data, the data was broken out into indicator columns. Initially the data was imputed, but it seemed to favor one variable over the other. Because of the skewed imputed data, the initial indicator columns were deleted. The missing categorical data was recoded to “unknown”. Below is the updated data.



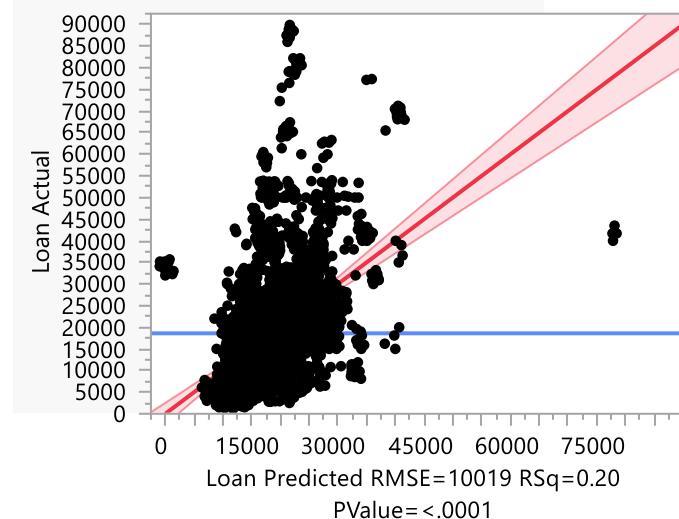


After filling the missing data with values, the next part is to look at the new stats summary. By looking at the distribution, the first thing that jumps out is that all the continuous variables are skewed. Because of the skewed data, the best thing to do is to take the log of the dependent

variable. What about the other variables? Well without seeming to be log happy, but since the data is skewed it would be better to take the log of the variables.

Regression Model with the Original Data

Regression Model- all original data



Effect Summary

Source	LogWorth	PValue
Value	107.926	0.00000
Mortgage	27.732	0.00000
YOJ	17.019	0.00000
Job_Self	12.747	0.00000
Reason_DebtCon	8.199	0.00000
DEBTINC	5.560	0.00000
Job_Sales	5.087	0.00001
Delinquencies	2.942	0.00114
Inquiries	2.666	0.00216
CLAGE	2.542	0.00287
Job_ProfExe	1.497	0.03181
Derogatories	1.262	0.05473
CLNO	0.667	0.21505
Reason_HomelImp	0.343	0.45440
Job_Mgr	0.274	0.53187
Job_Other	0.206	0.62252
Job_Office	0.190	0.64586

Summary of Fit

RSquare	0.203198
RSquare Adj	0.200919
Root Mean Square Error	10018.52
Mean of Response	18607.97
Observations (or Sum Wgts)	5960

Analysis of Variance

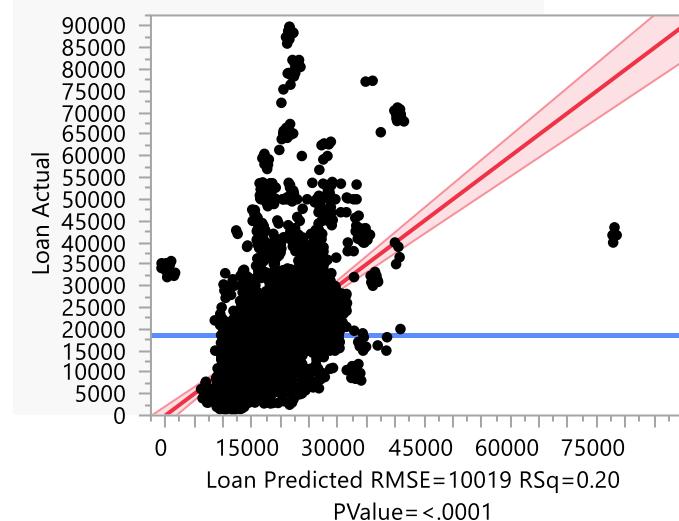
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	17	1.5209e+11	8.9467e+9	89.1361
Error	5942	5.964e+11	100370689	Prob > F
C. Total	5959	7.485e+11		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	5293.283	940.4478	5.63	<.0001*	.
Mortgage	-0.073365	0.006596	-11.12	<.0001*	5.0984555
Value	0.1124255	0.004972	22.61	<.0001*	4.8246104
Reason_DebtCon	4116.2007	707.6832	5.82	<.0001*	6.6822042
Reason_HomeImp	-543.173	726.0237	-0.75	0.4544	6.5560873
Job_Mgr	-472.1466	755.2014	-0.63	0.5319	3.7973974
Job_Office	-337.5826	734.6055	-0.46	0.6459	4.2862213
Job_Other	-338.8783	688.3373	-0.49	0.6225	6.7560859
Job_ProfExe	-1558.726	725.8822	-2.15	0.0318*	5.2643739
Job_Sales	-5258.896	1177.971	-4.46	<.0001*	1.4793559
Job_Self	7350.9257	995.949	7.38	<.0001*	1.8455606
YOJ	161.66064	18.78441	8.61	<.0001*	1.1065459
Derogatories	329.84644	171.6714	1.92	0.0547	1.1123202
Delinquencies	-411.5015	126.4475	-3.25	0.0011*	1.0937537
CLAGE	5.0248203	1.684874	2.98	0.0029*	1.1834103
Inquiries	253.36028	82.55812	3.07	0.0022*	1.1116515
CLNO	-18.67432	15.06069	-1.24	0.2150	1.3358963
DEBTINC	84.49091	18.00477	4.69	<.0001*	1.1445241

This regression model shows an R Squared Adjusted of .20, which makes this a bad model. A majority of the p-values look to significant and the VIF doesn't look terrible. The graph looks terrible, it looks like the data is pooled in one area.

Regression Model- Stepwise



Effect Summary

Source	LogWorth	PValue
Value	111.141	0.00000
Reason_DebtCon	56.552	0.00000
Mortgage	31.480	0.00000
Job_Self	22.785	0.00000
YOJ	16.468	0.00000
Job_Sales	6.218	0.00000
DEBTINC	5.012	0.00001
Job_ProfExe	3.899	0.00013
Inquiries	3.134	0.00073
Delinquencies	2.790	0.00162
CLAGE	2.130	0.00741

Summary of Fit

RSquare	0.202332
RSquare Adj	0.200857
Root Mean Square Error	10018.91
Mean of Response	18607.97
Observations (or Sum Wgts)	5960

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	11	1.5144e+11	1.377e+10	137.1579
Error	5948	5.9705e+11	100378466	Prob > F
C. Total	5959	7.485e+11		<.0001*

Parameter Estimates

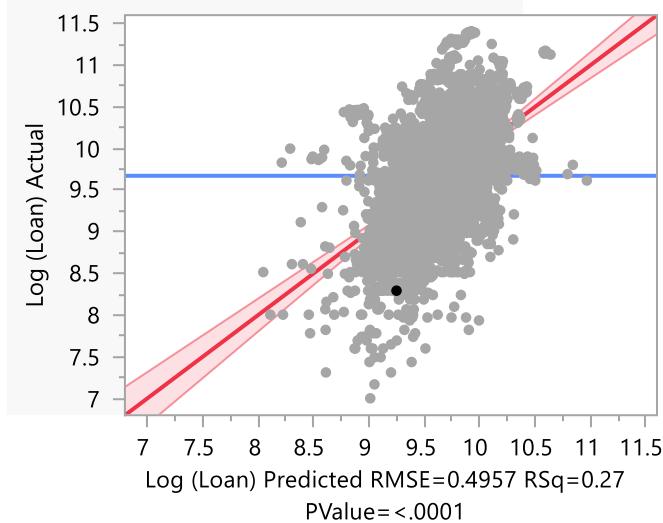
Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	4653.2814	699.6209	6.65	<.0001*	.
Mortgage	-0.075829	0.006381	-11.88	<.0001*	4.7709778
Value	0.1133354	0.004936	22.96	<.0001*	4.7534778
Reason_DebtCon	4521.9325	280.4521	16.12	<.0001*	1.0493632
Job_ProfExe	-1291.767	336.709	-3.84	0.0001*	1.1326353
Job_Sales	-4887.166	978.4387	-4.99	<.0001*	1.0205564
Job_Self	7598.1348	757.1431	10.04	<.0001*	1.0665381
YOJ	158.12603	18.69651	8.46	<.0001*	1.0961289
Delinquencies	-383.5878	121.6448	-3.15	0.0016*	1.0121689
CLAGE	4.3704676	1.631617	2.68	0.0074*	1.109695
Inquiries	273.30384	80.89759	3.38	0.0007*	1.0673002
DEBTINC	77.82403	17.58008	4.43	<.0001*	1.0910825

After running the stepwise and comparing it to the first regression, not much has changed in terms of the R squared adjustment and RMSE. The only thing that has changed is that the variables have become more significant and the VIF is under 5.

Regression Model with the Log of the Original Data

The only variables that it didn't make sense to take the log were Derogatories, Delinquencies, and inquiries because of the rating system (like 1-6, 1-10, etc). While the imputed data is not a whole number it is better than removing the data. The reason why CLNO is not included in the exception is that while the numbers are whole there seems to a more random set.

Regression Model- Log of variables



Effect Summary

Source	LogWorth	PValue
Log(Value)	165.666	0.00000
log (Mortgage)	50.643	0.00000
Job_Self	17.461	0.00000
Default	15.760	0.00000
Log(YOJ)	9.912	0.00000
Reason_DebtCon	6.658	0.00000
Job_Other	5.194	0.00001
Reason_HomeImp	4.404	0.00004
Inquiries	4.136	0.00007
Log(DEBTINC)	3.945	0.00011
Log(CLAGE)	3.872	0.00013
Derogatories	3.607	0.00025
Job_Mgr	3.586	0.00026
Job_Office	3.294	0.00051
Job_ProfExe	2.665	0.00216
Job_Sales	1.728	0.01869
Delinquencies	1.506	0.03121
Log(CLNO)	0.419	0.38111

Summary of Fit

RSquare	0.274201
RSquare Adj	0.271817

Root Mean Square Error	0.495687
Mean of Response	9.672645
Observations (or Sum Wgts)	5499

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	18	508.6846	28.2603	115.0168
Error	5480	1346.4659	0.2457	Prob > F
C. Total	5498	1855.1505		<.0001*

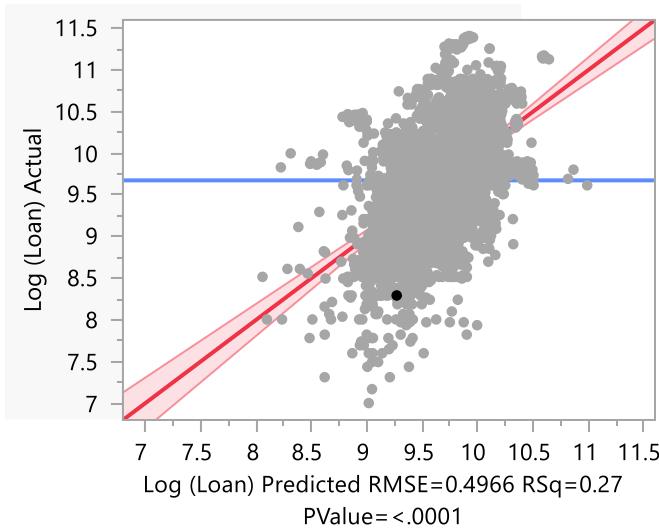
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	4.3274003	0.185778	23.29	<.0001*	.
Default	-0.15754	0.019061	-8.26	<.0001*	1.3074009
log (Mortgage)	-0.27498	0.018048	-15.24	<.0001*	2.9985605
Log(Value)	0.6533199	0.022948	28.47	<.0001*	2.9925315
Reason_DebtCon	0.1868694	0.036016	5.19	<.0001*	6.5939663
Reason_Homelmp	-0.152393	0.037043	-4.11	<.0001*	6.5244872
Job_Mgr	0.1410608	0.03859	3.66	0.0003*	3.814745
Job_Office	0.1311535	0.037705	3.48	0.0005*	4.2016974
Job_Other	0.1598126	0.035379	4.52	<.0001*	6.6856593
Job_ProfExe	0.1138072	0.037093	3.07	0.0022*	5.2752959
Job_Sales	-0.142987	0.060783	-2.35	0.0187*	1.4763116
Job_Self	0.4388432	0.050293	8.73	<.0001*	1.8787529
Log(YOJ)	0.0483312	0.007495	6.45	<.0001*	1.0859917
Derogatories	0.0332656	0.00907	3.67	0.0002*	1.1487626
Delinquencies	-0.014568	0.00676	-2.15	0.0312*	1.2124765
Log(CLAGE)	0.0548549	0.014355	3.82	0.0001*	1.2584794
Inquiries	0.0174958	0.004408	3.97	<.0001*	1.1360206
Log(CLNO)	-0.013787	0.01574	-0.88	0.3811	1.3914217
Log(DEBTINC)	0.1084565	0.028078	3.86	0.0001*	1.1339487

This model is slightly better than the original regression model. The R Squared Adjustment went from .20 to .27. All the variables are significant and the VIFs are pretty good.

Regression Model with the Log of the Original Data-Stepwise

It is worth noting that nothing changed significantly by doing this regression as you can see below, which is why I am going to add it and move on to the next step.



Effect Summary

Source	LogWorth	PValue
Log(Value)	166.253	0.00000
log (Mortgage)	53.393	0.00000
Job_Self	17.337	0.00000
Default	16.113	0.00000
Log(YOJ)	10.032	0.00000
Reason_DebtCon	6.449	0.00000
Inquiries	6.412	0.00000
Job_Other	5.518	0.00000
Reason_Homelmp	4.724	0.00002
Log(CLAGE)	3.486	0.00033
Job_Mgr	3.421	0.00038
Job_Office	3.317	0.00048
Log(DEBTINC)	3.124	0.00075
Derogatories	2.896	0.00127
Job_ProfExe	2.530	0.00295
Delinquencies	1.818	0.01522
Job_Sales	1.770	0.01699

Summary of Fit

RSquare	0.274501
RSquare Adj	0.272269
Root Mean Square Error	0.496569
Mean of Response	9.675073
Observations (or Sum Wgts)	5544

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
				Prob > F
Model	17	515.5573	30.3269	122.9896
Error	5526	1362.6067	0.2466	

Source	DF	Sum of Squares	Mean Square	F Ratio
C. Total	5543	1878.1640		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	4.4265748	0.181703	24.36	<.0001*	.
Default	-0.157152	0.018793	-8.36	<.0001*	1.288594
log (Mortgage)	-0.279762	0.017864	-15.66	<.0001*	2.9472717
Log(Value)	0.6536142	0.022921	28.52	<.0001*	2.99965
Reason_DebtCon	0.1833079	0.03596	5.10	<.0001*	6.5947642
Reason_HomeImp	-0.158313	0.036977	-4.28	<.0001*	6.5165845
Job_Mgr	0.1368754	0.038488	3.56	0.0004*	3.8035969
Job_Office	0.1313121	0.037597	3.49	0.0005*	4.1809277
Job_Other	0.1648001	0.035265	4.67	<.0001*	6.6882233
Job_ProfExe	0.1098151	0.036927	2.97	0.0030*	5.2247302
Job_Sales	-0.14492	0.060697	-2.39	0.0170*	1.4671187
Job_Self	0.4367746	0.050246	8.69	<.0001*	1.8691499
Log(YOJ)	0.0483823	0.007454	6.49	<.0001*	1.081918
Derogatories	0.0292027	0.009058	3.22	0.0013*	1.1428485
Delinquencies	-0.016254	0.006695	-2.43	0.0152*	1.1866718
Log(CLAGE)	0.0491551	0.013671	3.60	0.0003*	1.1454878
Inquiries	0.0218481	0.0043	5.08	<.0001*	1.1173493
Log(DEBTINC)	0.0917386	0.027206	3.37	0.0008*	1.1206036

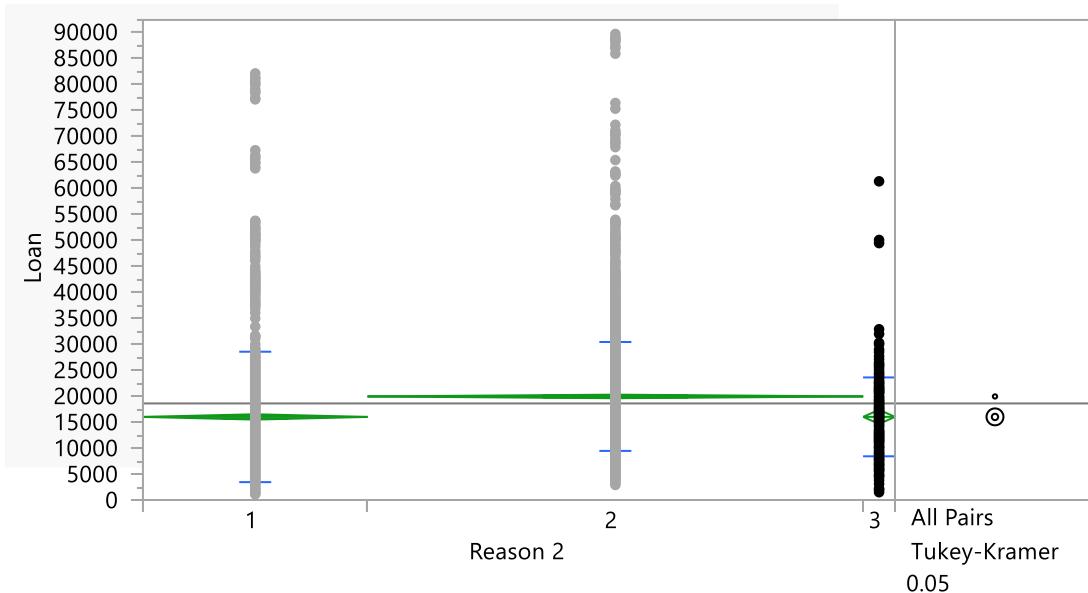
ANOVA

To run the ANOVA, Loan will be the dependent variable ad it will be compared to Reason and Job. Below will be 2 one way ANOVA's (one for reason and one for job), then a Two way ANOVA. To define the Data, please see the following:

- Reason:
 - HomeImp=1
 - DebtCon=2
 - Unknown=3
- Job:
 - Mgr=1
 - Office=2
 - Other=3
 - ProfExe=4
 - Sales=5
 - Self=6
 - Unknown=7

ANOVA with Loan and the categorical data Reason and Job

ANOVA One-Way Loan by Reason



Oneway Anova Summary of Fit

Rsquare	0.027844
Adj Rsquare	0.027518
Root Mean Square Error	11052.2
Mean of Response	18607.97
Observations (or Sum Wgts)	5960

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Reason 2	2	2.0841e+10	1.042e+10	85.3102	<.0001*
Error	5957	7.2765e+11	122151136		
C. Total	5959	7.485e+11			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	1780	16006.6	261.96	15493	16520
2	3928	19953.0	176.34	19607	20299
3	252	16017.9	696.22	14653	17383

Std Error uses a pooled estimate of error variance

Means Comparisons Comparisons for all pairs using Tukey-Kramer HSD Confidence Quantile

q*	Alpha
2.34429	0.05

HSD Threshold Matrix

Abs(Dif)-HSD

	2	3	1
2	-584.6	2251.4	3206.0
3	2251.4	-2308.2	-1732.6
1	3206.0	-1732.6	-868.5

Positive values show pairs of means that are significantly different.

Connecting Letters Report

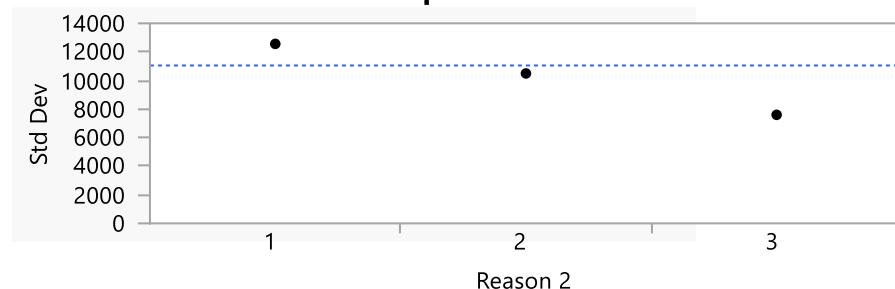
Level	Mean	
2	A	19952.953
3	B	16017.857
1	B	16006.629

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value	
2	1	3946.324	315.7875	3206.03	4686.621	<.0001*	
2	3	3935.096	718.2091	2251.41	5618.785	<.0001*	
3	1	11.228	743.8756	-1732.63	1755.087	0.9999	

Tests that the Variances are Equal



Level	Count	Std Dev	MeanAbsDif to Mean	MeanAbsDif to Median
1	1780	12563.57	8458.910	7890.337
2	3928	10492.83	7580.012	7426.935
3	252	7598.08	5608.305	5570.238

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	13.3766	2	5957	<.0001*
Brown-Forsythe	8.5544	2	5957	0.0002*
Levene	17.5736	2	5957	<.0001*
Bartlett	69.9447	2	.	<.0001*

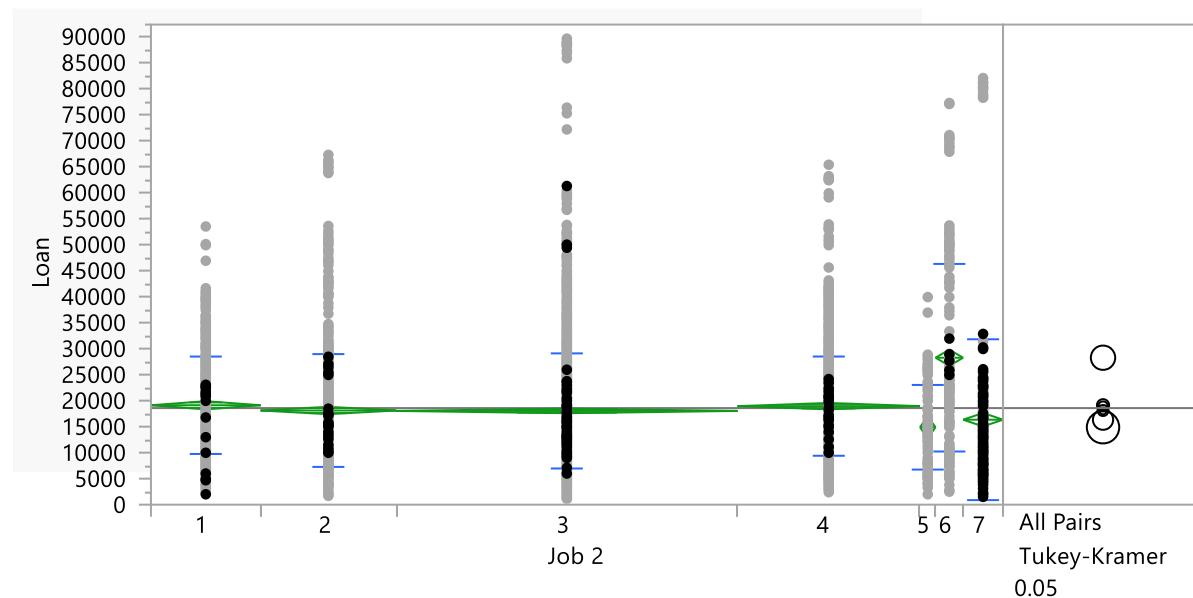
Welch's Test

Welch Anova testing Means Equal, allowing Std Devs Not Equal

F Ratio	DFNum	DFDen	Prob > F
84.4634	2	720.52	<.0001*

The P-value is small, which means that there is a difference in the means. Based on the Levene test the p-value is small and suggests the variance is not equal. The Welch test is also small, which means that there is a significance difference in the population means among HomeImp, DebtCon, and unknown. Based on the Connecting Letters Report, HomeImp and Unknown are similar in Means and DebtCon is different.

ANOVA One-Way Loan by Job



Oneway Anova Summary of Fit

Rsquare	0.029919
Adj Rsquare	0.028941
Root Mean Square Error	11044.11
Mean of Response	18607.97
Observations (or Sum Wgts)	5960

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Job 2	6	2.2394e+10	3.7324e+9	30.6002	<.0001*
Error	5953	7.261e+11	121972369		
C. Total	5959	7.485e+11			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	767	19155.3	398.8	18374	19937
2	948	18142.6	358.7	17439	18846
3	2388	18061.7	226.0	17619	18505
4	1276	18983.5	309.2	18377	19590
5	109	14913.8	1057.8	12840	16987
6	193	28314.5	795.0	26756	29873
7	279	16371.7	661.2	15076	17668

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons for all pairs using Tukey-Kramer HSD

Confidence Quantile

q*	Alpha
2.94936	0.05

HSD Threshold Matrix

Abs(Dif)-HSD

	6	1	4	2	3	7	5
6	-3315.8	6536.1	6815.3	7599.6	7815.3	8893.2	9498.0
1	6536.1	-1663.3	-1316.4	-569.3	-258.3	506.3	907.3
4	6815.3	-1316.4	-1289.6	-555.8	-207.7	459.0	819.2
2	7599.6	-569.3	-555.8	-1496.1	-1169.5	-447.6	-65.6
3	7815.3	-258.3	-207.7	-1169.5	-942.7	-370.9	-42.4
7	8893.2	506.3	459.0	-447.6	-370.9	-2757.9	-2221.3
5	9498.0	907.3	819.2	-65.6	-42.4	-2221.3	-4412.3

Positive values show pairs of means that are significantly different.

Connecting Letters Report

Level	Mean	
6	A	28314.508
1	B	19155.280
4	B	18983.464
2	B C	18142.616
3	B C	18061.683
7	C	16371.685
5	C	14913.761

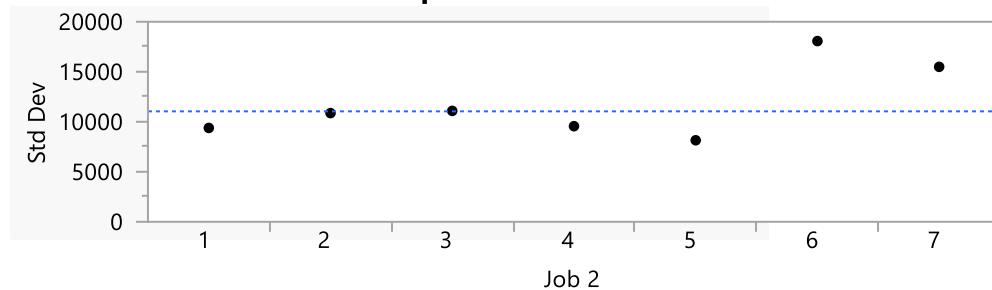
Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value	Plot
6	5	13400.75	1323.251	9498.00	17303.49	<.0001*	
6	7	11942.82	1034.001	8893.18	14992.46	<.0001*	
6	3	10252.82	826.473	7815.26	12690.39	<.0001*	
6	2	10171.89	872.149	7599.61	12744.17	<.0001*	

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value			
6	4	9331.04	852.978	6815.31	11846.78	<.0001*			
6	1	9159.23	889.385	6536.11	11782.35	<.0001*			
1	5	4241.52	1130.503	907.26	7575.78	0.0034*			
4	5	4069.70	1102.090	819.24	7320.16	0.0042*			
2	5	3228.85	1116.994	-65.56	6523.27	0.0591			
3	5	3147.92	1081.707	-42.42	6338.27	0.0559			
1	7	2783.60	772.141	506.27	5060.92	0.0058*			
4	7	2611.78	729.909	459.02	4764.54	0.0064*			
2	7	1770.93	752.223	-447.65	3989.51	0.2186			
3	7	1690.00	698.752	-370.87	3750.87	0.1908			
7	5	1457.92	1247.473	-2221.32	5137.17	0.9059	<	≤	>
1	3	1093.60	458.369	-258.30	2445.49	0.2046	<	≤	>
1	2	1012.66	536.366	-569.27	2594.60	0.4884	<	≤	>
4	3	921.78	382.971	-207.74	2051.30	0.1956	<	≤	>
4	2	840.85	473.553	-555.83	2237.53	0.5647	<	≤	>
1	4	171.82	504.594	-1316.41	1660.05	0.9999	<	≤	>
2	3	80.93	423.958	-1169.47	1331.34	1.0000	<	≤	>

Tests that the Variances are Equal



Level	Count	Std Dev	MeanAbsDif to Mean	MeanAbsDif to Median
1	767	9385.69	7572.79	7533.77
2	948	10868.21	7545.64	7339.45
3	2388	11090.79	7901.88	7629.52
4	1276	9559.92	7128.57	6993.03
5	109	8152.48	6714.81	6709.17
6	193	18069.48	14565.66	13650.26
7	279	15496.54	8747.40	8208.96

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	17.7770	6	5953	<.0001*
Brown-Forsythe	19.3850	6	5953	<.0001*
Levene	28.2628	6	5953	<.0001*
Bartlett	52.5623	6	.	<.0001*

Welch's Test

Welch Anova testing Means Equal, allowing Std Devs Not Equal

F Ratio	DFNum	DFDen	Prob > F
15.7708	6	819.96	<.0001*

The P-value is small, which means that there is a difference in the means. Based on the Levene test the p-value is small and suggests the variance is not equal. The Welch test is also small, which means that there is a significance difference in the population means among Mgr, Office, Other, ProfExe, Sales, Self, and Unknown. Based on the Connecting Letters Report; people who are listed as Mgr and ProfExe have similar means; the people who are listed as Unknown and in Sales are similar in means; people who are listed as office and other fall between the means for Mgr and ProfExe and Unknown and in Sales; and those that are listed as Self is completely different from the other means.

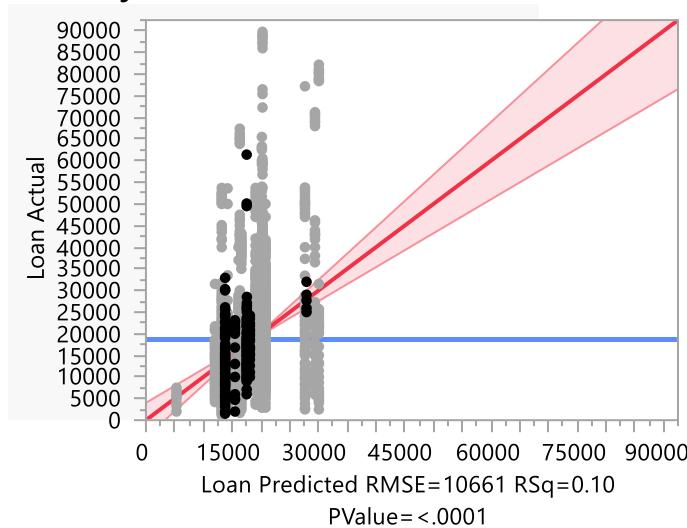
ANOVA Two-Way Loan by Reason and Job

Response Loan

Singularity Details

Term	Details
Intercept	=Reason 2[1] + Reason 2[2] + Job 2[1] + Job 2[2] + Job 2[3] + Job 2[4] - 6*Job 2[5] + Job 2[6] - Reason 2[1]*Job 2[1] - Reason 2[1]*Job 2[2] - Reason 2[1]*Job 2[3] - Reason 2[1]*Job 2[4] + 6*Reason 2[1]*Job 2[5] - Reason 2[1]*Job 2[6] - Reason 2[2]*Job 2[1] - Reason 2[2]*Job 2[2] - Reason 2[2]*Job 2[3] - Reason 2[2]*Job 2[4] + 6*Reason 2[2]*Job 2[5] - Reason 2[2]*Job 2[6]

Actual by Predicted Plot



Effect Summary

Source	LogWorth	PValue
Reason 2*Job 2	40.453	0.00000
Job 2	7.963	0.00000 ^

Source	LogWorth	PValue
Reason 2	3.150	0.00071 ^

Summary of Fit

RSquare	0.098099
RSquare Adj	0.095214
Root Mean Square Error	10660.58
Mean of Response	18607.97
Observations (or Sum Wgts)	5960

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	19	7.3426e+10	3.8646e+9	34.0046
Error	5940	6.7507e+11	113648032	Prob > F
C. Total	5959	7.485e+11		<.0001*

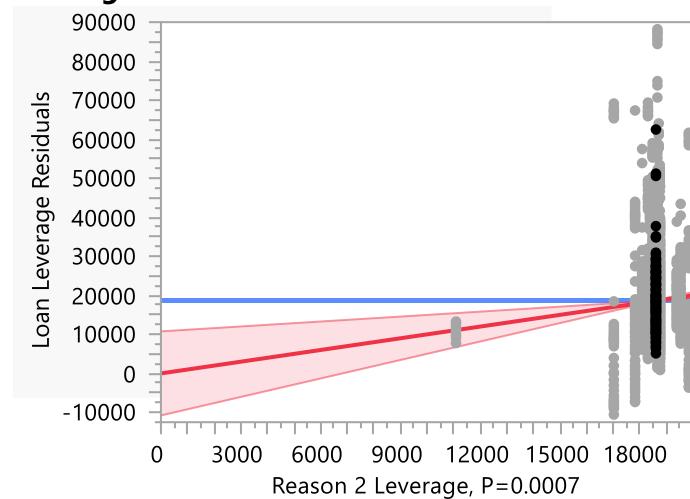
Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Reason 2	2	1	1304698545	11.4802	0.0007* LostDFs
Job 2	6	5	5202668491	9.1558	<.0001* LostDFs
Reason 2*Job 2	12	11	2.5546e+10	20.4344	<.0001* LostDFs

Effect Details

Reason 2

Leverage Plot



Least Squares Means Table

Level	Least Sq		Std Error	Mean
	Mean			
1	17614.691		533.04659	16006.6
2	19681.155		296.36278	19953.0
3	0.000	NonEstimable	.	16017.9

LSMeans Differences Tukey HSD

$\alpha=0.050$ Q=2.34429

LSMean[i] By LSMean[j]

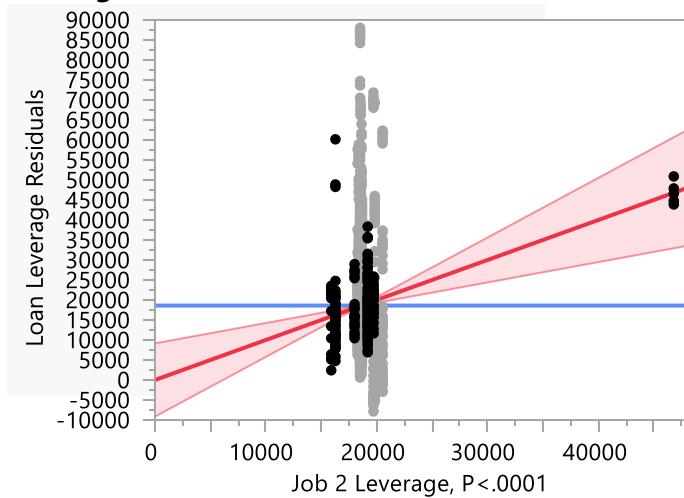
Mean[i]-Mean[j]	1	2	3	
Std Err Dif				
Lower CL Dif				
Upper CL Dif				
1	0 0 0 0	-2066.5 609.893 -3496.2 -636.7	17614.7 5393.43 4970.93 30258.5	
2	2066.46 609.893 636.698 3496.23	0 0 0 0	19681.2 5398.62 7025.22 32337.1	
3	-17615 5393.43 -30258 -4970.9	-19681 5398.62 -32337 -7025.2	0 0 0 0	

Level	Least Sq	
	Mean	
2	A	19681.155
1	B	17614.691
3	C	0.000

Levels not connected by same letter are significantly different.

Job 2

Leverage Plot



Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
1	16823.440	834.2426	19155.3
2	17639.464	728.0154	18142.6
3	16984.913	459.5739	18061.7
4	18272.246	756.4628	18983.5
5	0.000 NonEstimable	.	14913.8
6	28319.535	1675.7974	28314.5
7	18613.724	670.3399	16371.7

LSMeans Differences Tukey HSD

$\alpha=0.050$ Q=2.94936

LSMean[i] By LSMean[j]

Mean[i]-Mean[j]	1	2	3	4	5	6	7
Std Err Dif							
Lower CL Dif							
Upper CL Dif							
1	0	-816.02	-161.47	-1448.8	16823.4	-11496	-1790.3
	0	1107.23	952.454	1126.14	11220.1	1871.97	1070.19
	0	-4081.7	-2970.6	-4770.2	-16269	-17017	-4946.7
	0	2449.61	2647.66	1872.6	49915.7	-5975	1366.11
2	816.024	0	654.551	-632.78	17639.5	-10680	-974.26
	1107.23	0	860.938	1049.88	11198.4	1827.1	989.627
	-2449.6	0	-1884.7	-3729.3	-15389	-16069	-3893
	4081.66	0	3193.77	2463.69	50667.6	-5291.3	1944.51
3	161.473	-654.55	0	-1287.3	16984.9	-11335	-1628.8
	952.454	860.938	0	885.124	11159	1737.67	812.751
	-2647.7	-3193.8	0	-3897.9	-15927	-16460	-4025.9
	2970.61	1884.67	0	1323.22	49896.7	-6209.6	768.285
4	1448.81	632.782	1287.33	0	18272.2	-10047	-341.48

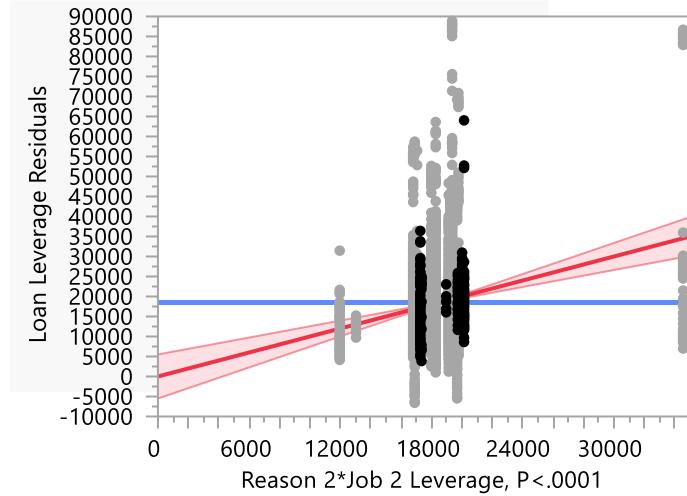
	1126.14	1049.88	885.124	0	11205.5	1838.62	1010.74
	-1872.6	-2463.7	-1323.2	0	-14777	-15470	-3322.5
	4770.21	3729.25	3897.88	0	51321.3	-4624.5	2639.55
5	-16823	-17639	-16985	-18272	0	-28320	-18614
	11220.1	11198.4	11159	11205.5	0	9963.54	11163.6
	-49916	-50668	-49897	-51321	0	-57706	-51539
	16268.8	15388.6	15926.9	14776.8	0	1066.54	14311.7
6	11496.1	10680.1	11334.6	10047.3	28319.5	0	9705.81
	1871.97	1827.1	1737.67	1838.62	9963.54	0	1804.9
	5974.99	5291.28	6209.6	4624.53	-1066.5	0	4382.52
	17017.2	16068.9	16459.6	15470.1	57705.6	0	15029.1
7	1790.28	974.26	1628.81	341.479	18613.7	-9705.8	0
	1070.19	989.627	812.751	1010.74	11163.6	1804.9	0
	-1366.1	-1944.5	-768.29	-2639.6	-14312	-15029	0
	4946.68	3893.03	4025.91	3322.51	51539.1	-4382.5	0

Level	Least Sq Mean	
6	A	28319.535
7	B	18613.724
4	B	18272.246
2	B	17639.464
3	B	16984.913
1	B	16823.440
5	A B	0.000

Levels not connected by same letter are significantly different.

Reason 2*Job 2

Leverage Plot



LSMeans Differences Tukey HSD

$\alpha=0.050$ Q=3.5707

LSMean[i] By LSMean[j]

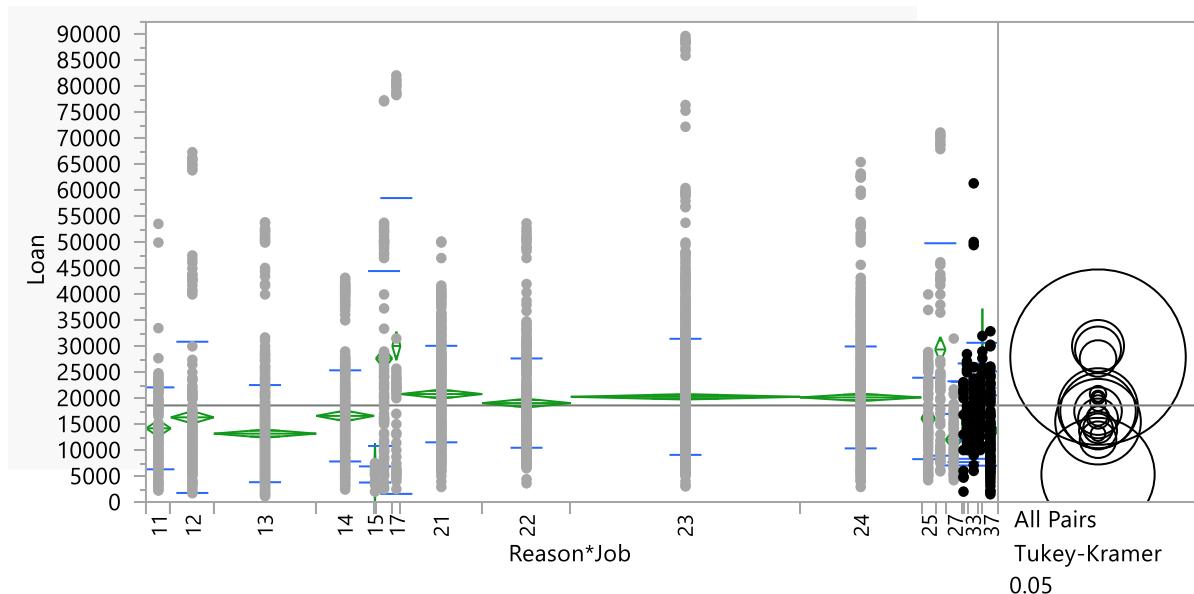
Mean[i]-Mean[j]	1,1	1,2	1,3	1,4	1,5	1,6	1,7	2,1	2,2	2,3	2,4	2,5
Std Err Dif												
Lower CL Dif												
Upper CL Dif												
1,1	0	-2119.1	1019.41	-2399.2	8886.49	-13455	-15872	-6604.5	-4861.5	-6071.3	-5957	-1907.2
	0	1015.24	901.042	966.314	3181.79	1281.17	1626.95	922.949	914.578	850.883	887.314	1350.84
	0	-5744.2	-2197.9	-5849.7	-2474.7	-18029	-21681	-9900	-8127.1	-9109.5	-9125.3	-6730.7
	0	1506	4236.75	1051.17	20247.7	-8880.1	-10062	-3308.9	-1595.8	-3033	-2788.7	2916.22
1,2	2119.13	0	3138.53	-280.12	11005.6	-11336	-13753	-4485.3	-2742.3	-3952.1	-3837.9	211.892
	1015.24	0	732.321	811.283	3138.19	1168.68	1539.93	759.114	748.914	669.643	715.364	1244.67
	-1506	0	523.638	-3177	-199.9	-15509	-19251	-7195.9	-5416.5	-6343.2	-6392.2	-4232.4
	5744.25	0	5753.43	2616.73	22211.1	-7162.6	-8254.1	-1774.8	-68.194	-1561	-1283.5	4656.22
1,3	-1019.4	-3138.5	0	-3418.7	7867.09	-14474	-16891	-7623.9	-5880.9	-7090.7	-6976.4	-2926.6
	901.042	732.321	0	662.826	3103.13	1070.97	1467.16	597.839	584.833	479.144	541.206	1153.41
	-4236.8	-5753.4	0	-5785.4	-3213.2	-18298	-22130	-9758.6	-7969.1	-8801.5	-8908.9	-7045.1
	2197.94	-523.64	0	-1051.9	18947.4	-10650	-11652	-5489.2	-3792.6	-5379.8	-5043.9	1191.84
1,4	2399.25	280.121	3418.66	0	11285.7	-11055	-13473	-4205.2	-2462.2	-3672	-3557.8	492.012
	966.314	811.283	662.826	0	3122.7	1126.44	1508.12	692.313	681.114	592.845	644.042	1205.09
	-1051.2	-2616.7	1051.9	0	135.513	-15078	-18858	-6677.3	-4894.3	-5788.9	-5857.4	-3811
	5849.66	3176.97	5785.41	0	22436	-7033.3	-8087.5	-1733.2	-30.167	-1555.1	-1258.1	4795.02
1,5	-8886.5	-11006	-7867.1	-11286	0	-22341	-24758	-15491	-13748	-14958	-14843	-10794
	3181.79	3138.19	3103.13	3122.7	0	3234.02	3385.93	3109.56	3107.08	3088.94	3099.17	3262.25
	-20248	-22211	-18947	-22436	0	-33889	-36848	-26594	-24842	-25987	-25910	-22442
	2474.73	199.903	3213.24	-135.51	0	-10794	-12668	-4387.7	-2653.5	-3928.1	-3777.3	854.788
1,6	13454.7	11335.6	14474.1	11055.5	22341.2	0	-2417.1	6850.26	8593.27	7383.48	7497.74	11547.5
	1281.17	1168.68	1070.97	1126.44	3234.02	0	1726.87	1089.46	1082.38	1029.12	1059.44	1469.65
	8880.07	7162.61	10650	7033.33	10793.5	0	-8583.2	2960.12	4728.42	3708.79	3714.78	6299.83
	18029.4	15508.6	18298.2	15077.6	33888.9	0	3749.01	10740.4	12458.1	11058.2	11280.7	16795.2
1,7	15871.8	13752.7	16891.2	13472.6	24758.3	2417.1	0	9267.37	11010.4	9800.58	9914.84	13964.6
	1626.95	1539.93	1467.16	1508.12	3385.93	1726.87	0	1480.71	1475.51	1436.9	1458.77	1779.17
	10062.5	8254.08	11652.5	8087.54	12668.2	-3749	0	3980.19	5741.78	4669.85	4706.02	7611.72
	21681.2	19251.3	22130	18857.6	36848.4	8583.22	0	14554.5	16279	14931.3	15123.7	20317.5
2,1	6604.47	4485.35	7623.88	4205.23	15491	-6850.3	-9267.4	0	1743.01	533.216	647.471	4697.24
	922.949	759.114	597.839	692.313	3109.56	1089.46	1480.71	0	618.052	519.171	576.942	1170.6
	3308.9	1774.78	5489.18	1733.19	4387.68	-10740	-14555	0	-463.87	-1320.6	-1412.6	517.366
	9900.04	7195.91	9758.58	6677.27	26594.3	-2960.1	-3980.2	0	3949.89	2387.02	2707.56	8877.11
2,2	4861.46	2742.34	5880.87	2462.22	13748	-8593.3	-11010	-1743	0	-1209.8	-1095.5	2954.23
	914.578	748.914	584.833	681.114	3107.08	1082.38	1475.51	618.052	0	504.139	563.454	1164.02
	1595.78	68.1942	3792.61	30.1667	2653.5	-12458	-16279	-3949.9	0	-3009.9	-3107.5	-1202.1
	8127.14	5416.48	7969.13	4894.27	24842.4	-4728.4	-5741.8	463.867	0	590.333	916.383	7110.57
2,3	6071.26	3952.13	7090.67	3672.01	14957.8	-7383.5	-9800.6	-533.22	1209.79	0	114.255	4164.02
	850.883	669.643	479.144	592.845	3088.94	1029.12	1436.9	519.171	504.139	0	452.802	1114.67
	3033.01	1561.04	5379.79	1555.14	3928.1	-11058	-14931	-2387	-590.33	0	-1502.6	183.887
	9109.5	6343.22	8801.54	5788.88	25987.4	-3708.8	-4669.9	1320.59	3009.92	0	1731.07	8144.16
2,4	5957	3837.88	6976.41	3557.76	14843.5	-7497.7	-9914.8	-647.47	1095.54	-114.25	0	4049.77
	887.314	715.364	541.206	644.042	3099.17	1059.44	1458.77	576.942	563.454	452.802	0	1142.72
	2788.67	1283.53	5043.93	1258.08	3777.31	-11281	-15124	-2707.6	-916.38	-1731.1	0	-30.534
	9125.33	6392.22	8908.89	5857.43	25909.7	-3714.8	-4706	1412.62	3107.46	1502.56	0	8130.07
2,5	1907.23	-211.89	2926.64	-492.01	10793.7	-11548	-13965	-4697.2	-2954.2	-4164	-4049.8	0
	1350.84	1244.67	1153.41	1205.09	3262.25	1469.65	1779.17	1170.6	1164.02	1114.67	1142.72	0
	-2916.2	-4656.2	-1191.8	-4795	-854.79	-16795	-20317	-8877.1	-7110.6	-8144.2	-8130.1	0
	6730.69	4232.44	7045.12	3811	22442.2	-6299.8	-7611.7	-517.37	1202.12	-183.89	30.5335	0
2,6	15194.2	13075.1	16213.6	12795	24080.7	1739.48	-677.63	8589.74	10332.8	9122.96	9237.21	13287

	1486.6	1390.82	1309.79	1355.52	3320.77	1595.33	1884.32	1324.96	1319.14	1275.8	1300.38	1651.8
	9886.03	8108.88	11536.8	7954.82	12223.3	-3956.9	-7405.9	3858.73	5622.51	4567.45	4593.93	7388.89
	20502.4	18041.3	20890.5	17635.1	35938.2	7435.9	6050.69	13320.8	15043	13678.5	13880.5	19185.1
2,7	-2191.3	-4310.5	-1171.9	-4590.6	6695.14	-15646	-18063	-8795.8	-7052.8	-8262.6	-8148.4	-4098.6
	1281.17	1168.68	1070.97	1126.44	3234.02	1405.88	1726.87	1089.46	1082.38	1029.12	1059.44	1469.65
	-6766	-8483.5	-4996	-8612.8	-4852.6	-20666	-24229	-12686	-10918	-11937	-11931	-9346.3
	2383.31	-137.48	2652.16	-568.44	18242.9	-10626	-11897	-4905.7	-3188	-4587.9	-4365.4	1149.09
3,1	1281.36	-837.76	2300.77	-1117.9	10167.9	-12173	-14590	-5323.1	-3580.1	-4789.9	-4675.6	-625.87
	2462.71	2406.11	2360.2	2385.88	3857.78	2529.83	2721.33	2368.65	2365.4	2341.51	2354.99	2565.82
	-7512.2	-9429.3	-6126.8	-9637.1	-3607.1	-21207	-24308	-13781	-12026	-13151	-13085	-9787.6
	10075	7753.73	10728.3	7401.37	23942.8	-3140.1	-4873.4	3134.61	4866.02	3570.92	3733.32	8535.9
3,2	3353.32	1234.19	4372.73	954.074	12239.8	-10101	-12519	-3251.2	-1508.1	-2717.9	-2603.7	1446.09
	2205.07	2141.67	2089.96	2118.91	3698.63	2279.79	2490.58	2099.49	2095.83	2068.83	2084.07	2319.66
	-4520.3	-6413.1	-3089.9	-6611.9	-966.86	-18242	-21412	-10748	-8991.7	-10105	-10045	-6836.7
	11227	8881.45	11835.3	8520.07	25446.5	-1961	-3625.4	4245.5	5975.42	4669.21	4837.91	9728.88
3,3	3318.41	1199.28	4337.82	919.161	12204.9	-10136	-12553	-3286.1	-1543.1	-2752.8	-2638.6	1411.17
	1524.61	1431.38	1352.78	1397.11	3337.96	1630.81	1914.45	1367.47	1361.84	1319.9	1343.68	1686.1
	-2125.5	-3911.8	-492.56	-4069.5	286.068	-15959	-19389	-8168.9	-6405.8	-7465.8	-7436.5	-4609.4
	8762.34	6310.32	9168.19	5907.8	24123.7	-4313.2	-5717.5	1596.76	3319.65	1960.13	2159.27	7431.71
3,4	3876.01	1756.88	4895.41	1476.76	12762.5	-9578.7	-11996	-2728.5	-985.46	-2195.3	-2081	1968.77
	2321.31	2261.17	2212.25	2239.63	3769.09	2392.4	2594.06	2221.27	2217.8	2192.3	2206.7	2430.42
	-4412.7	-6317.1	-3003.9	-6520.3	-695.76	-18121	-21258	-10660	-8904.5	-10023	-9960.4	-6709.5
	12164.7	9830.84	12794.7	9473.8	26220.8	-1036.2	-2733.2	5203	6933.63	5632.79	5798.45	10647.1
3,5	-14195	-16314	-13175	-16594	-5308.3	-27650	-30067	-20799	-19056	-20266	-20152	-16102
	33481.3	33468.9	33459.1	33464.6	33874	33289.1	33541.3	33443.1	33443.8	33448.8	33446	33399.4
	-133746	-135821	-132648	-136086	-126262	-146515	-149832	-140214	-138474	-139702	-139577	-135361
	105357	103193	106297	102898	115645	91215.8	89699.1	98615.8	100361	99169.5	99273.6	103157
3,6	13725.2	11606	14744.6	11325.9	22611.7	270.435	-2146.7	7120.7	8863.71	7653.92	7768.17	11817.9
	4835.57	4806.99	4784.18	4796.9	5674.53	4870.1	4972.27	4788.35	4786.74	4774.98	4781.61	4888.89
	-3541.2	-5558.3	-2338.3	-5802.3	2349.65	-17119	-19901	-9977	-8228.3	-9396.1	-9305.5	-5638.8
	30991.5	28770.4	31827.4	28454.2	42873.7	17660.1	15607.8	24218.4	25955.7	24703.9	24841.8	29274.7
3,7	-423.8	-2542.9	595.609	-2823	8462.69	-13879	-16296	-7028.3	-5285.3	-6495.1	-6380.8	-2331
	1309.69	1199.87	1104.92	1158.77	3245.43	1431.91	1748.13	1122.86	1115.99	1064.42	1093.76	1494.58
	-5100.3	-6827.3	-3349.7	-6960.7	-3125.7	-18991	-22538	-11038	-9270.1	-10296	-10286	-7667.7
	4252.7	1741.46	4540.96	1314.56	20051.1	-8765.6	-10054	-3018.9	-1300.4	-2694.3	-2475.3	3005.65

Level	Least Sq Mean					
1,7	A					
2,6	A					
3,6	A B C D E					
1,6	A					
2,1	B					
2,3	B					
2,4	B E					
2,2	B E					
3,4	B C D E F G					
3,2	B C D E F G					
3,3	B C D E					
1,4	C					
1,2	C F					
2,5	C D E F G					

Level							Least Sq Mean
3,1	B	C	D	E	F	G	15476.190
1,1		C	D		F	G	14194.828
3,7		C	D		F	G	13771.028
1,3			D		G		13175.419
2,7			D		G		12003.478
1,5				F	G		5308.333
3,5	A	B	C	D	E	F	0.000

Levels not connected by same letter are significantly different.



Oneway Anova Summary of Fit

Rsquare	0.098099
Adj Rsquare	0.095214
Root Mean Square Error	10660.58
Mean of Response	18607.97
Observations (or Sum Wgts)	5960

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Reason*Job	19	7.3426e+10	3.8646e+9	34.0046	<.0001*
Error	5940	6.7507e+11	113648032		
C. Total	5959	7.485e+11			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
11	174	14194.8	808.2	12611	15779

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
12	301	16314.0	614.5	15109	17519
13	716	13175.4	398.4	12394	13956
14	405	16594.1	529.7	15556	17633
15	12	5308.3	3077.4	-725	11341
16	115	27649.6	994.1	25701	29598
17	57	30066.7	1412.0	27299	32835
21	572	20799.3	445.7	19925	21673
22	620	19056.3	428.1	18217	19896
23	1604	20266.1	266.2	19744	20788
24	847	20151.8	366.3	19434	20870
25	97	16102.1	1082.4	13980	18224
26	73	29389.0	1247.7	26943	31835
27	115	12003.5	994.1	10055	13952
31	21	15476.2	2326.3	10916	20037
32	27	17548.1	2051.6	13526	21570
33	68	17513.2	1292.8	14979	20048
34	24	18070.8	2176.1	13805	22337
36	5	27920.0	4767.6	18574	37266
37	107	13771.0	1030.6	11751	15791

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons for all pairs using Tukey-Kramer HSD

Confidence Quantile

q*	Alpha
3.54542	0.05

HSD Threshold Matrix

Abs(Dif)-HSD

	17	26	36	16	21	23	24	22	34	32
17	-7080	-6003	-15482	-3705	4018	4706	4743	5779	2799	3688
26	-6003	-6256	-16003	-3917	3892	4600	4627	5656	2425	3327
36	-15482	-16003	-23904	-16996	-9856	-9275	-9185	-8107	-8731	-8030
16	-3705	-3917	-16996	-4984	2988	3735	3742	4756	1097	2019
21	4018	3892	-9856	2988	-2235	-1307	-1398	-448	-5147	-4192
23	4706	4600	-9275	3735	-1307	-1335	-1491	-578	-5577	-4617
24	4743	4627	-9185	3742	-1398	-1491	-1837	-902	-5743	-4785
22	5779	5656	-8107	4756	-448	-578	-902	-2147	-6878	-5922
34	2799	2425	-8731	1097	-5147	-5577	-5743	-6878	-10911	-10081
32	3688	3327	-8030	2019	-4192	-4617	-4785	-5922	-10081	-10287
33	5766	5506	-7107	4354	-1562	-1927	-2125	-3285	-8416	-8563
14	8126	7989	-5681	7062	1751	1570	1274	47	-6464	-6558
12	8293	8144	-5437	7192	1794	1578	1302	87	-6260	-6359
25	7657	7431	-5515	6337	547	212	-2	-1173	-6648	-6778
31	4942	4554	-6364	3204	-3075	-3512	-3674	-4806	-8699	-8925
11	10104	9924	-3419	8912	3332	3055	2811	1619	-4354	-4465
37	10098	9880	-3144	8802	3047	2721	2503	1329	-4237	-4363
13	11690	11570	-2217	10677	5504	5392	5058	3807	-2948	-3037
27	11941	11729	-1350	10662	4933	4614	4392	3215	-2415	-2538
15	12754	12307	2493	10875	4466	4006	3856	2732	-601	-873

Positive values show pairs of means that are significantly different.

Connecting Letters Report

Level		Mean
17	A	30066.667
26	A	29389.041
36	A B C D E	27920.000
16	A	27649.565
21	B	20799.301
23	B	20266.085
24	B E	20151.830
22	B E	19056.290
34	B C D E F G	18070.833
32	B C D E F G	17548.148
33	B C D E	17513.235
14	C	16594.074
12	C F	16313.953
25	C D E F G	16102.062
31	B C D E F G	15476.190
11	C D F G	14194.828
37	C D F G	13771.028
13	D G	13175.419
27	D G	12003.478
15	F G	5308.333

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Diff	Lower CL	Upper CL	p-Value	Diagram
17	15	24758.33	3385.926	12753.8	36762.88	<.0001*	
26	15	24080.71	3320.767	12307.2	35854.23	<.0001*	
36	15	22611.67	5674.529	2493.1	42730.28	0.0104*	
16	15	22341.23	3234.024	10875.2	33807.22	<.0001*	
17	27	18063.19	1726.867	11940.7	24185.66	<.0001*	
26	27	17385.56	1595.326	11729.5	23041.67	<.0001*	
17	13	16891.25	1467.158	11689.6	22092.94	<.0001*	
17	37	16295.64	1748.129	10097.8	22493.50	<.0001*	
26	13	16213.62	1309.790	11569.9	20857.38	<.0001*	
36	27	15916.52	4870.098	-1350.0	33183.08	0.1161	
17	11	15871.84	1626.953	10103.6	21640.08	<.0001*	
16	27	15646.09	1405.876	10661.7	20630.51	<.0001*	
26	37	15618.01	1618.318	9880.4	21355.64	<.0001*	
21	15	15490.97	3109.559	4466.3	26515.67	0.0001*	
26	11	15194.21	1486.597	9923.6	20464.83	<.0001*	
23	15	14957.75	3088.935	4006.2	25909.34	0.0002*	
24	15	14843.50	3099.169	3855.6	25831.36	0.0003*	
36	13	14744.58	4784.175	-2217.4	31706.51	0.1902	
17	31	14590.48	2721.330	4942.2	24238.75	<.0001*	
16	13	14474.15	1070.967	10677.1	18271.18	<.0001*	
36	37	14148.97	4877.677	-3144.5	31442.41	0.2894	

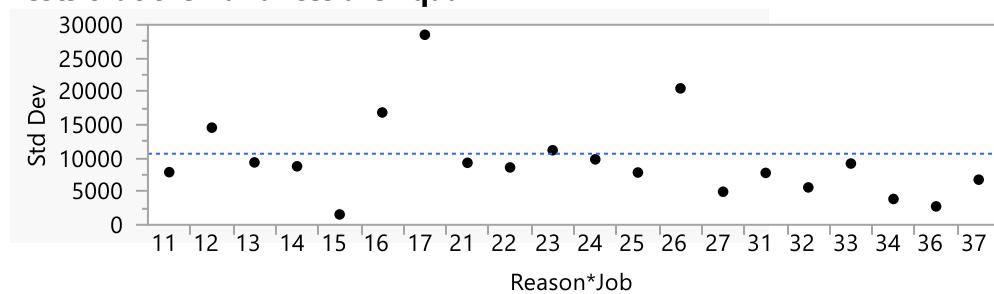
Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value				
17	25	13964.60	1779.172	7656.7	20272.53	<.0001*				
26	31	13912.85	2639.817	4553.6	23272.12	<.0001*				
16	37	13878.54	1431.913	8801.8	18955.28	<.0001*				
17	12	13752.71	1539.933	8293.0	19212.43	<.0001*				
22	15	13747.96	3107.084	2732.0	24763.89	0.0017*				
36	11	13725.17	4835.572	-3419.0	30869.33	0.3297				
17	14	13472.59	1508.124	8125.7	18819.53	<.0001*				
16	11	13454.74	1281.169	8912.5	17997.02	<.0001*				
26	25	13286.98	1651.803	7430.6	19143.32	<.0001*				
26	12	13075.09	1390.824	8144.0	18006.15	<.0001*				
26	14	12794.97	1355.520	7989.1	17600.86	<.0001*				
34	15	12762.50	3769.085	-600.5	26125.51	0.0824				
17	33	12553.43	1914.450	5765.9	19340.97	<.0001*				
17	32	12518.52	2490.585	3688.3	21348.70	<.0001*				
36	31	12443.81	5304.848	-6364.1	31251.75	0.7020				
32	15	12239.81	3698.629	-873.4	25353.02	0.1033				
33	15	12204.90	3337.958	370.4	24039.38	0.0345*				
16	31	12173.37	2529.833	3204.0	21142.71	0.0003*				
17	34	11995.83	2594.062	2798.8	21192.89	0.0007*				
26	33	11875.81	1796.696	5505.8	18245.86	<.0001*				
26	32	11840.89	2401.251	3327.4	20354.35	0.0002*				
36	25	11817.94	4888.889	-5515.2	29151.12	0.6483				
36	12	11606.05	4806.992	-5436.8	28648.87	0.6505				
16	25	11547.50	1469.651	6337.0	16758.04	<.0001*				
16	12	11335.61	1168.680	7192.1	15479.08	<.0001*				
36	14	11325.93	4796.897	-5681.1	28332.96	0.6907				
26	34	11318.21	2508.417	2424.8	20211.61	0.0011*				
14	15	11285.74	3122.704	214.4	22357.05	0.0398*				
16	14	11055.49	1126.435	7061.8	15049.18	<.0001*				
17	22	11010.38	1475.510	5779.1	16241.68	<.0001*				
12	15	11005.62	3138.190	-120.6	22131.84	0.0566				
25	15	10793.73	3262.254	-772.3	22359.80	0.1035				
36	33	10406.76	4939.727	-7106.7	27920.19	0.8538				
36	32	10371.85	5190.259	-8029.8	28773.52	0.9037				
26	22	10332.75	1319.138	5655.8	15009.66	<.0001*				
31	15	10167.86	3857.782	-3509.6	23845.33	0.4773				
16	33	10136.33	1630.809	4354.4	15918.24	<.0001*				
16	32	10101.42	2279.787	2018.6	18184.23	0.0016*				
17	24	9914.84	1458.767	4742.9	15086.79	<.0001*				
36	34	9849.17	5240.700	-8731.3	28429.67	0.9439				
17	23	9800.58	1436.899	4706.2	14895.00	<.0001*				
16	34	9578.73	2392.400	1096.7	18060.81	0.0096*				
17	21	9267.37	1480.713	4017.6	14517.12	<.0001*				
26	24	9237.21	1300.384	4626.8	13847.63	<.0001*				
26	23	9122.96	1275.804	4599.7	13646.22	<.0001*				
11	15	8886.49	3181.795	-2394.3	20167.31	0.3607				
36	22	8863.71	4786.743	-8107.3	25834.75	0.9511				

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value					
21	27	8795.82	1089.463	4933.2	12658.43	<.0001*					
16	22	8593.27	1082.380	4755.8	12430.77	<.0001*					
26	21	8589.74	1324.956	3892.2	13287.27	<.0001*					
37	15	8462.69	3245.428	-3043.7	19969.11	0.4992					
23	27	8262.61	1029.124	4613.9	11911.29	<.0001*					
24	27	8148.35	1059.444	4392.2	11904.53	<.0001*					
13	15	7867.09	3103.127	-3134.8	18868.99	0.5561					
36	24	7768.17	4781.609	-9184.7	24721.00	0.9874					
36	23	7653.92	4774.983	-9275.4	24583.25	0.9892					
21	13	7623.88	597.839	5504.3	9743.48	<.0001*					
16	24	7497.74	1059.444	3741.6	11253.91	<.0001*					
16	23	7383.48	1029.124	3734.8	11032.16	<.0001*					
36	21	7120.70	4788.350	-9856.0	24097.43	0.9956					
23	13	7090.67	479.144	5391.9	8789.44	<.0001*					
22	27	7052.81	1082.380	3215.3	10890.31	<.0001*					
21	37	7028.27	1122.861	3047.3	11009.29	<.0001*					
24	13	6976.41	541.206	5057.6	8895.21	<.0001*					
16	21	6850.26	1089.463	2987.7	10712.87	<.0001*					
27	15	6695.14	3234.024	-4770.8	18161.13	0.8721					
21	11	6604.47	922.949	3332.2	9876.72	<.0001*					
23	37	6495.06	1064.417	2721.2	10268.87	<.0001*					
24	37	6380.80	1093.759	2503.0	10258.64	<.0001*					
23	11	6071.26	850.883	3054.5	9088.00	<.0001*					
34	27	6067.36	2392.400	-2414.7	14549.43	0.5554					
24	11	5957.00	887.314	2811.1	9102.91	<.0001*					
22	13	5880.87	584.833	3807.4	7954.35	<.0001*					
32	27	5544.67	2279.787	-2538.1	13627.48	0.6369					
33	27	5509.76	1630.809	-272.2	11291.67	0.0843					
21	31	5323.11	2368.649	-3074.8	13720.97	0.7706					
22	37	5285.26	1115.990	1328.6	9241.92	0.0004*					
34	13	4895.41	2212.252	-2948.0	12738.79	0.7928					
22	11	4861.46	914.578	1618.9	8104.03	<.0001*					
23	31	4789.89	2341.509	-3511.7	13091.54	0.8835					
21	25	4697.24	1170.604	546.9	8847.53	0.0093*					
24	31	4675.64	2354.992	-3673.8	13025.09	0.9088					
14	27	4590.60	1126.435	596.9	8584.29	0.0073*					
21	12	4485.35	759.114	1794.0	7176.73	<.0001*					
32	13	4372.73	2089.955	-3037.0	11782.51	0.8612					
33	13	4337.82	1352.783	-458.4	9134.00	0.1376					
12	27	4310.48	1168.680	167.0	8453.94	0.0309*					
34	37	4299.81	2407.793	-4236.8	12836.45	0.9657					
21	14	4205.23	692.313	1750.7	6659.77	<.0001*					
23	25	4164.02	1114.667	212.1	8115.99	0.0262*					
25	27	4098.58	1469.651	-1112.0	9309.12	0.3636					
24	25	4049.77	1142.719	-1.7	8101.19	0.0502					
23	12	3952.13	669.643	1578.0	6326.30	<.0001*					
34	11	3876.01	2321.311	-4354.0	12106.04	0.9830					

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value					
24	12	3837.88	715.364	1301.6	6374.14	<.0001*					
32	37	3777.12	2295.935	-4362.9	11917.18	0.9855					
33	37	3742.21	1653.308	-2119.5	9603.88	0.7599					
23	14	3672.01	592.845	1570.1	5773.90	<.0001*					
22	31	3580.10	2365.399	-4806.2	11966.44	0.9945					
24	14	3557.76	644.042	1274.4	5841.16	<.0001*					
31	27	3472.71	2529.833	-5496.6	12442.04	0.9984					
14	13	3418.66	662.826	1068.7	5768.66	<.0001*					
32	11	3353.32	2205.071	-4464.6	11171.23	0.9942					
33	11	3318.41	1524.613	-2087.0	8723.81	0.8149					
21	33	3286.07	1367.472	-1562.2	8134.33	0.6592					
21	32	3251.15	2099.493	-4192.4	10694.75	0.9928					
12	13	3138.53	732.321	542.1	5734.92	0.0030*					
22	25	2954.23	1164.016	-1172.7	7081.16	0.5540					
25	13	2926.64	1153.410	-1162.7	7015.97	0.5544					
14	37	2823.05	1158.768	-1285.3	6931.37	0.6337					
23	33	2752.85	1319.904	-1926.8	7432.47	0.8645					
22	12	2742.34	748.914	87.1	5397.55	0.0339*					
21	34	2728.47	2221.265	-5146.9	10603.80	0.9997					
23	32	2717.94	2068.826	-4616.9	10052.80	0.9991					
24	33	2638.59	1343.678	-2125.3	7402.50	0.9170					
24	32	2603.68	2084.074	-4785.2	9992.61	0.9996					
34	31	2594.64	3185.458	-8699.2	13888.44	1.0000					
12	37	2542.93	1199.875	-1711.1	6796.99	0.8472					
22	14	2462.22	681.114	47.4	4877.05	0.0397*					
17	16	2417.10	1726.867	-3705.4	8539.58	0.9979					
14	11	2399.25	966.314	-1026.7	5825.24	0.5974					
25	37	2331.03	1494.577	-2967.9	7629.94	0.9921					
31	13	2300.77	2360.199	-6067.1	10668.68	1.0000					
23	34	2195.25	2192.302	-5577.4	9967.89	1.0000					
11	27	2191.35	1281.169	-2350.9	6733.64	0.9780					
17	36	2146.67	4972.266	-15482.1	19775.46	1.0000					
12	11	2119.13	1015.243	-1480.3	5718.59	0.8637					
24	34	2081.00	2206.697	-5742.7	9904.67	1.0000					
32	31	2071.96	3101.773	-8925.1	13069.06	1.0000					
33	31	2037.04	2661.410	-7398.8	11472.87	1.0000					
34	25	1968.77	2430.425	-6648.1	10585.66	1.0000					
25	11	1907.23	1350.844	-2882.1	6696.55	0.9977					
37	27	1767.55	1431.913	-3309.2	6844.29	0.9996					
34	12	1756.88	2261.173	-6259.9	9773.70	1.0000					
21	22	1743.01	618.052	-448.2	3934.27	0.3419					
26	16	1739.48	1595.326	-3916.6	7395.58	0.9999					
31	37	1705.16	2544.394	-7315.8	10726.12	1.0000					
22	33	1543.06	1361.836	-3285.2	6371.34	0.9999					
22	32	1508.14	2095.827	-5922.5	8938.74	1.0000					
34	14	1476.76	2239.631	-6463.7	9417.20	1.0000					
26	36	1469.04	4928.126	-16003.3	18941.34	1.0000					

Level	- Level	Difference	Std Err Diff	Lower CL	Upper CL	p-Value						
32	25	1446.09	2319.658	-6778.1	9670.26	1.0000	<	>	<	>	<	>
33	25	1411.17	1686.097	-4566.8	7389.10	1.0000	<	>	<	>	<	>
31	11	1281.36	2462.714	-7450.0	10012.73	1.0000	<	>	<	>	<	>
32	12	1234.19	2141.671	-6358.9	8827.33	1.0000	<	>	<	>	<	>
23	22	1209.79	504.139	-577.6	2997.18	0.6617	<	>	<	>	<	>
33	12	1199.28	1431.385	-3875.6	6274.15	1.0000	<	>	<	>	<	>
13	27	1171.94	1070.967	-2625.1	4968.97	0.9999	<	>	<	>	<	>
14	31	1117.88	2385.880	-7341.1	9576.84	1.0000	<	>	<	>	<	>
24	22	1095.54	563.454	-902.1	3093.22	0.9238	<	>	<	>	<	>
11	13	1019.41	901.042	-2175.2	4213.98	0.9999	<	>	<	>	<	>
22	34	985.46	2217.800	-6877.6	8848.50	1.0000	<	>	<	>	<	>
32	14	954.07	2118.915	-6558.4	8466.53	1.0000	<	>	<	>	<	>
33	14	919.16	1397.107	-4034.2	5872.50	1.0000	<	>	<	>	<	>
12	31	837.76	2406.113	-7692.9	9368.45	1.0000	<	>	<	>	<	>
17	26	677.63	1884.316	-6003.1	7358.33	1.0000	<	>	<	>	<	>
21	24	647.47	576.942	-1398.0	2692.98	0.9999	<	>	<	>	<	>
25	31	625.87	2565.822	-8471.1	9722.80	1.0000	<	>	<	>	<	>
37	13	595.61	1104.924	-3321.8	4513.03	1.0000	<	>	<	>	<	>
34	33	557.60	2531.132	-8416.3	9531.54	1.0000	<	>	<	>	<	>
21	23	533.22	519.171	-1307.5	2373.90	1.0000	<	>	<	>	<	>
34	32	522.69	2990.739	-10080.8	11126.12	1.0000	<	>	<	>	<	>
14	25	492.01	1205.090	-3780.5	4764.57	1.0000	<	>	<	>	<	>
11	37	423.80	1309.687	-4219.6	5067.20	1.0000	<	>	<	>	<	>
14	12	280.12	811.283	-2596.2	3156.46	1.0000	<	>	<	>	<	>
36	16	270.43	4870.098	-16996.1	17537.00	1.0000	<	>	<	>	<	>
12	25	211.89	1244.668	-4201.0	4624.77	1.0000	<	>	<	>	<	>
23	24	114.25	452.802	-1491.1	1719.63	1.0000	<	>	<	>	<	>
32	33	34.91	2424.970	-8562.6	8632.46	1.0000	<	>	<	>	<	>

Tests that the Variances are Equal



Level	Count	Std Dev	MeanAbsDif to Mean	MeanAbsDif to Median
11	174	7895.55	6313.83	6261.49
12	301	14570.20	9910.87	8824.92
13	716	9357.47	6191.82	5821.23
14	405	8779.71	6562.94	6400.74
15	12	1553.56	1140.28	1108.33

Level	Count	Std Dev	MeanAbsDif	MeanAbsDif
			to Mean	to Median
16	115	16847.92	13682.04	13068.70
17	57	28500.53	22960.23	19601.75
21	572	9298.47	7720.23	7662.59
22	620	8603.56	6165.07	6089.84
23	1604	11182.17	7837.78	7661.35
24	847	9818.08	7278.24	7162.57
25	97	7844.38	6420.89	6345.36
26	73	20467.99	16840.08	15205.48
27	115	4961.29	4233.26	4206.09
31	21	7784.21	6991.38	6219.05
32	27	5608.06	4301.23	4040.74
33	68	9188.65	5668.12	5539.71
34	24	3872.81	3223.26	3204.17
36	5	2759.89	2064.00	2020.00
37	107	6769.42	5446.31	5429.91

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	23.3341	19	5940	<.0001*
Brown-Forsythe	20.2920	19	5940	<.0001*
Levene	33.2820	19	5940	<.0001*
Bartlett	44.5083	19	.	<.0001*

Warning: Small sample sizes. Use Caution.

Welch's Test

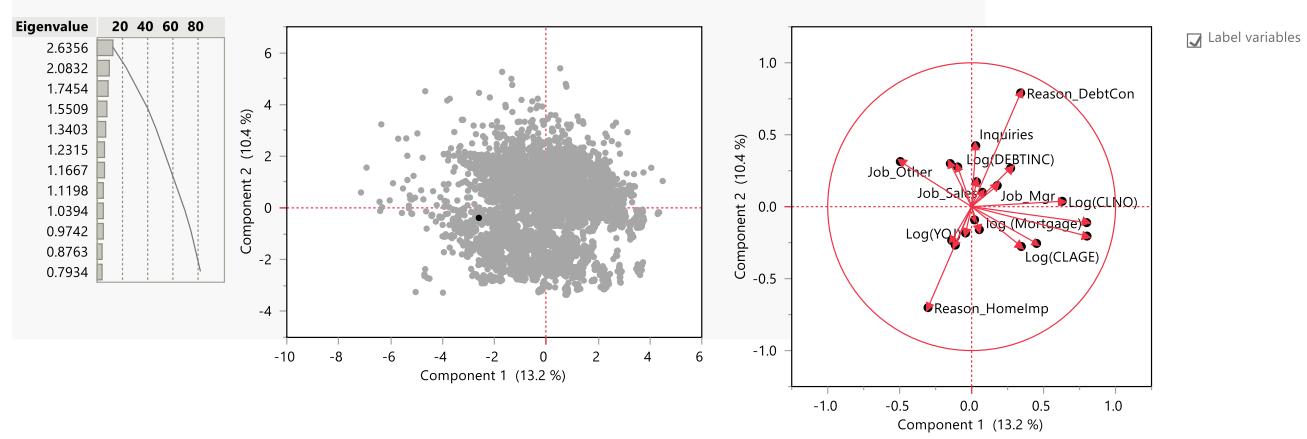
Welch Anova testing Means Equal, allowing Std Devs Not Equal

F Ratio	DFNum	DFDen	Prob > F
74.5043	19	241.68	<.0001*

The P-value is small, which means that there is a difference in the means. Based on the Levene test the p-value is small and suggests the variance is not equal. The Welch test is also small, which means that there is a significance difference in the population means among. Based on the connecting letters it becomes obvious that the means are different there are so many different combinations

Factor Analysis and Principal Components

Principal Components: on Correlations Summary Plots



Select component Component 1 Component 2

Warning: the Correlation matrix is not positive definite.

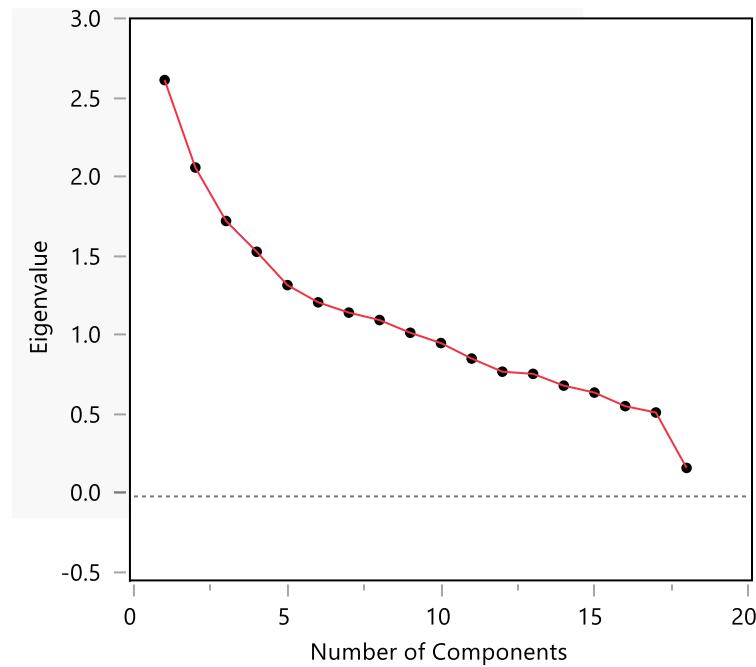
Eigenvalues

Number	Eigenvalue	Percent	Cum Percent	ChiSquare	DF	Prob > ChiSq
1	2.6356	13.178	13.178	4503.52	190.054	<.0001*
2	2.0832	10.416	23.594	91.427	174.987	1.0000
3	1.7454	8.727	32.321	.	159.537	.
4	1.5509	7.755	40.076	.	144.315	.
5	1.3403	6.702	46.778	.	129.408	.
6	1.2315	6.158	52.935	.	114.977	.
7	1.1667	5.834	58.769	.	101.274	.
8	1.1198	5.599	64.368	.	88.364	.
9	1.0394	5.197	69.565	.	76.188	.
10	0.9742	4.871	74.436	.	64.850	.
11	0.8763	4.381	78.817	.	54.417	.
12	0.7934	3.967	82.784	.	44.844	.
13	0.7797	3.899	86.683	.	36.117	.
14	0.7053	3.526	90.209	.	28.409	.
15	0.6614	3.307	93.516	.	21.606	.
16	0.5757	2.878	96.395	.	15.468	.
17	0.5356	2.678	99.073	.	10.381	.
18	0.1854	0.927	100.000	.	6.455	.

Eigenvectors

	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8	Prin9	Prin10	Prin11	Prin12	Prin13	Prin14	Prin15	Prin16	Prin17	Prin18
Default	-0.09113	0.20720	0.47483	0.13958	0.00152	0.00355	-0.19911	-0.01455	0.06233	0.06084	-0.24695	0.20332	-0.13640	-0.12976	0.12399	0.67909	-0.20489	0.00864
log (Mortgage)	0.49244	-0.07674	0.11231	-0.01044	-0.02100	-0.21745	0.12755	0.14901	-0.05318	-0.10664	-0.19121	0.14174	0.27591	-0.10689	-0.06557	0.00817	0.04225	0.69867
Log(Value)	0.49375	-0.14156	0.10232	-0.02125	0.03016	-0.17259	0.10329	0.13337	-0.02434	0.02184	-0.14427	0.15501	0.32231	-0.07918	-0.03036	0.07571	0.12929	-0.69827
Reason_DebtCon	0.20992	0.54757	-0.33215	0.01156	-0.03089	-0.00423	-0.09326	-0.05325	-0.01530	0.20000	-0.03907	0.00390	0.01491	0.04430	-0.09182	0.07263	0.04539	-0.00159
Reason_Homelmp	-0.18669	-0.48586	0.37439	-0.26074	0.00362	0.03581	0.04441	0.05190	0.03557	-0.15899	0.09050	-0.04403	0.03428	-0.06206	-0.17721	0.00393	-0.02390	0.00220
Reason_Unknown	-0.06987	-0.18486	-0.06905	0.56580	0.06452	-0.07146	0.11866	0.24353	-0.04478	-0.10957	-0.11380	0.09076	-0.11316	0.03676	0.61927	-0.18008	-0.05286	-0.00138
Job_Mgr	0.10973	0.10163	0.10983	0.12642	-0.14674	0.14628	0.57369	-0.59200	0.16047	-0.09864	-0.11902	-0.13825	-0.00071	0.00312	0.08942	0.01092	-0.00543	-0.01718
Job_Office	0.01271	-0.06352	-0.08665	-0.06898	-0.53911	0.53781	-0.12278	0.29104	-0.25291	-0.01432	0.03541	0.13706	0.15500	-0.08620	0.07359	0.05044	-0.03822	0.00330
Job_Other	-0.30518	0.21715	-0.02421	-0.23696	0.49158	-0.15250	0.13142	0.32319	-0.04328	-0.09039	-0.20173	-0.06895	0.16534	-0.08668	-0.03231	-0.03678	-0.05685	-0.01942
Job_ProfExe	0.27815	-0.17700	0.00485	-0.02578	0.14046	-0.21745	-0.58015	-0.36550	-0.14949	-0.06465	0.11084	0.04339	-0.21595	0.10722	0.14740	-0.07153	0.05164	0.01744
Job_Sales	0.04575	0.06933	-0.00017	0.01787	-0.10943	-0.01517	-0.16942	0.16062	0.87952	-0.10447	0.21262	0.16276	0.02626	-0.16246	0.03141	-0.12396	0.04938	0.00673
Job_Self	0.03294	-0.11054	0.20709	-0.05163	-0.16035	-0.25249	0.17812	0.18887	0.04386	0.79768	0.18837	-0.16347	-0.14002	0.10305	0.08461	-0.02222	-0.00456	0.06378
Job_Unknown	-0.08521	-0.16253	-0.15087	0.55054	-0.04399	-0.16654	0.08293	0.13400	-0.02067	-0.08572	0.08663	0.09205	-0.13006	0.15419	-0.63806	0.21739	0.08319	-0.02424
Log(YOJ)	-0.02576	-0.12668	-0.13147	-0.00589	0.41663	0.28408	0.16347	-0.19176	-0.01640	0.28907	0.27628	0.67691	0.12700	0.05073	0.01379	0.04279	-0.07786	0.07465
Derogatories	-0.05943	0.19127	0.34933	0.24211	0.07456	0.09439	-0.16058	-0.01799	-0.00147	-0.03236	0.27518	-0.16298	0.59065	0.53091	0.04233	-0.04272	0.00631	0.03051
Delinquencies	0.02066	0.11965	0.38784	0.23533	0.15042	0.33393	-0.15968	0.00054	-0.02899	0.21219	-0.35097	0.05812	-0.14225	-0.18384	-0.25055	-0.49135	0.29849	-0.00820
Log(CLAGE)	0.21267	-0.19167	-0.12727	0.02426	0.35947	0.40622	0.00590	0.13466	0.12478	0.03783	0.07137	-0.42008	-0.08381	-0.00814	0.11688	0.40334	0.44613	0.08846
Inquiries	0.01784	0.29357	0.23319	0.09895	0.00372	-0.13032	0.13124	0.04388	-0.28730	-0.18140	0.62360	0.00052	-0.08442	-0.48682	0.06119	-0.00789	0.24609	-0.01929
Log(CLNO)	0.38786	0.02591	0.07766	0.08969	0.22154	0.23632	0.02925	0.15684	0.00448	-0.05438	0.17591	-0.24267	-0.18721	-0.06900	-0.15348	-0.11827	-0.72937	-0.05545
Log(DEBTINC)	0.16693	0.18504	0.20982	-0.27499	-0.01605	0.06818	0.21723	0.24611	-0.01158	-0.25178	0.04726	0.27317	-0.46642	0.56066	0.02916	-0.03540	0.17478	-0.02582

Scree Plot



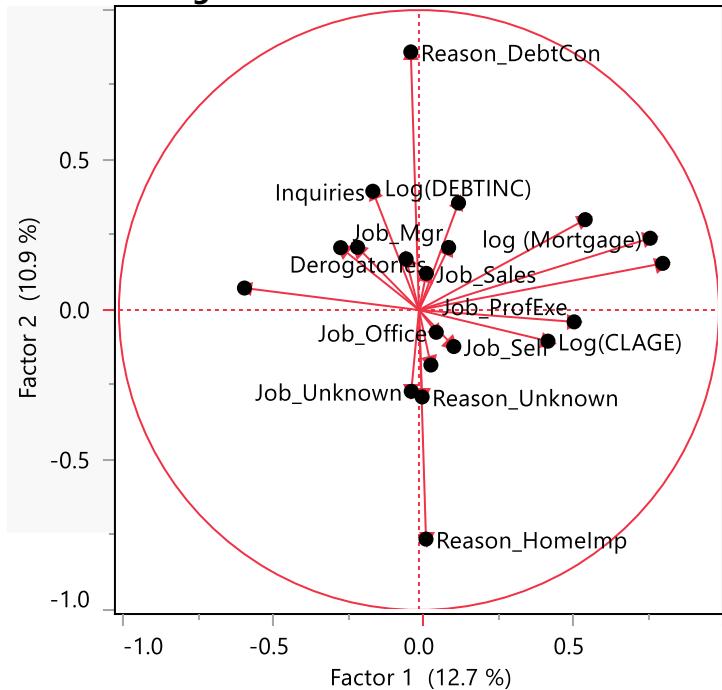
Factor Analysis: Principal Component / Varimax

Rotated Factor Loading

	Factor 1	Factor 2
Default	-0.260961	0.207893
log (Mortgage)	0.770791	0.239300
Log(Value)	0.812465	0.155514
Reason_DebtCon	-0.027185	0.860251
Reason_Homelmp	0.023494	-0.763587
Reason_Unknown	0.010644	-0.289726
Job_Mgr	0.098964	0.208469
Job_Office	0.057626	-0.074236
Job_Other	-0.581664	0.073295
Job_ProfExe	0.517311	-0.039473

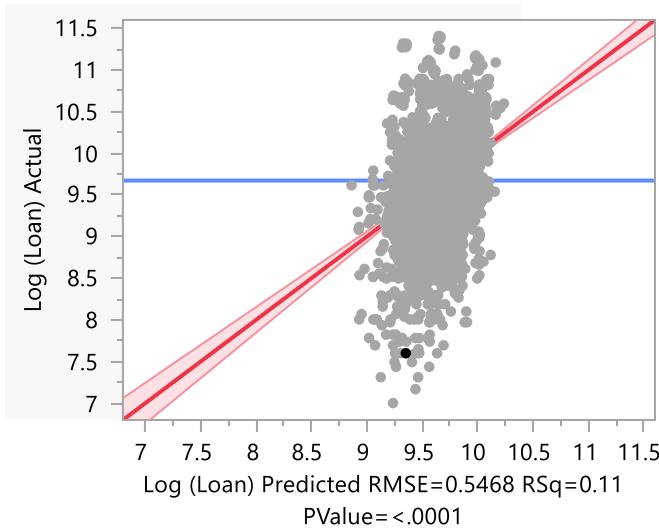
	Factor 1	Factor 2
Job_Sales	0.024736	0.122136
Job_Self	0.116190	-0.121723
Job_Unknown	-0.025587	-0.271131
Log(YOJ)	0.039798	-0.183291
Derogatories	-0.204610	0.208942
Delinquencies	-0.042986	0.170588
Log(CLAGE)	0.430078	-0.103799
Inquiries	-0.153764	0.395902
Log(CLNO)	0.554158	0.301321
Log(DEBTINC)	0.131897	0.356903

Factor Loading Plot



[x] Label variables

Regression Model with only FA/PCA



Effect Summary

Source	LogWorth	PValue
Prin1	133.826	0.00000
Prin2	14.181	0.00000

Summary of Fit

RSquare	0.114138
RSquare Adj	0.113815
Root Mean Square Error	0.546826
Mean of Response	9.672645
Observations (or Sum Wgts)	5499

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	2	211.7424	105.871	354.0618	
Error	5496	1643.4081	0.299		Prob > F
C. Total	5498	1855.1505			<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	9.6706925	0.007376	1311.1	<.0001*	.
Prin1	0.1166077	0.004593	25.39	<.0001*	1.0000392
Prin2	0.0402451	0.00515	7.81	<.0001*	1.0000392

Effect Tests

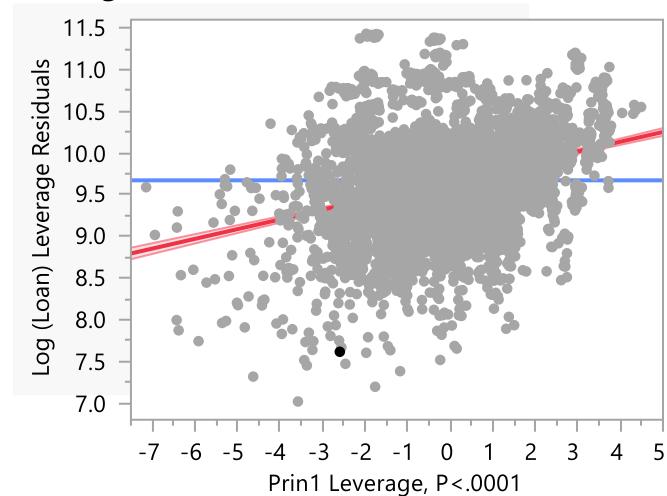
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Prin1	1	1	192.73432	644.5555	<.0001*

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Prin2	1	1	18.25734	61.0575	<.0001*

Effect Details

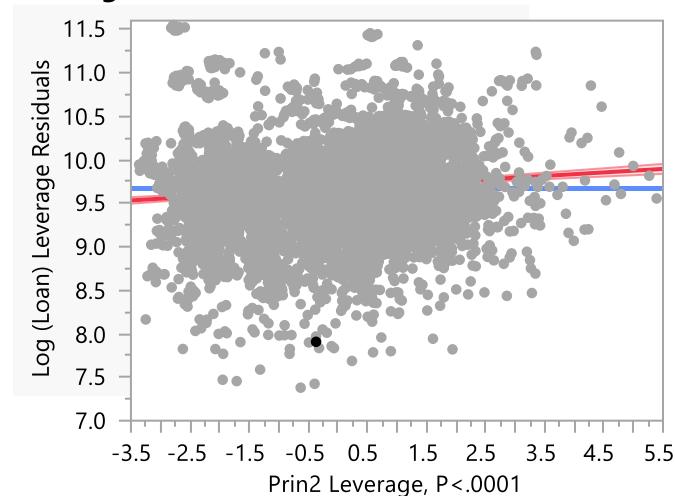
Prin1

Leverage Plot



Prin2

Leverage Plot



Compared to the regression model that was done with the logged variables; this is a substantially worse model. The R squared adjustment went from .27 to .11, it dropped a little more than double. The RMSE is greater at .55 more than the .49. While the variables are significant and the VIF is low.

Cluster Analysis

Method	NCluster	CCC	Best
K Means Cluster	3	-21.207	
K Means Cluster	4	-1.6881	
K Means Cluster	5	16.8703	Optimal CCC
K Means Cluster	6	-23.464	
K Means Cluster	7	47.1674	
K Means Cluster	8	91.1737	
K Means Cluster	9	96.0266	
K Means Cluster	10	138.618	

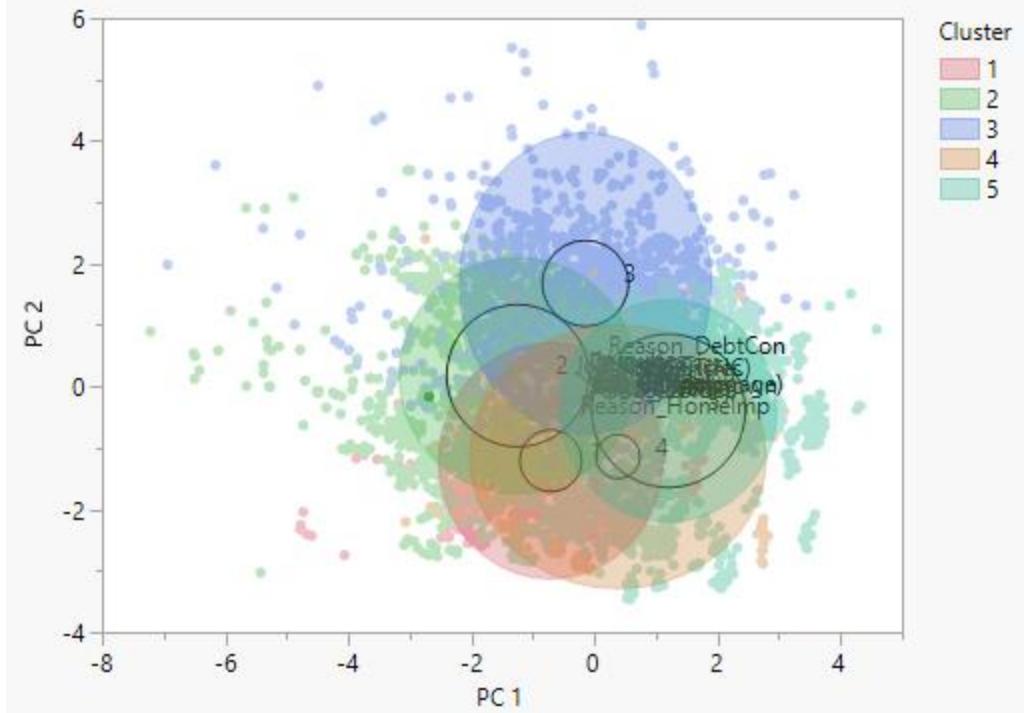
Based on the result, it seems like 5 would be the most optimal. The info for the 5 clusters is listed below.

Cluster Summary			
Cluster	Count	Step	Criterion
1	371	35	0
2	1952		
3	715		
4	188		
5	2273		

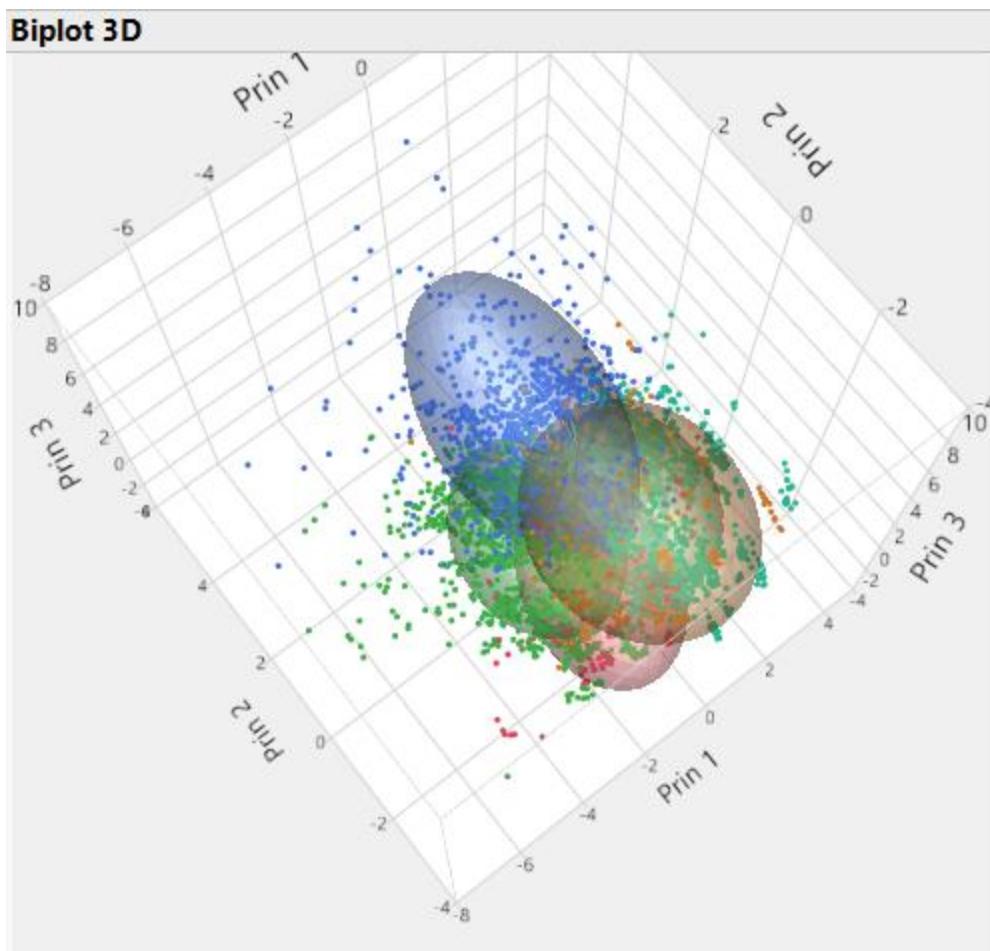
Cluster Means		log	Default	(Mortgage)	Log(Value)	Reason_DebtCon	Reason_HomeImp	Reason_Unknown	Job_Mgr	Job_Office	Job_Other	Job_ProfExe	Job_Sales	Job_Self	Job_Unknown	Log(YOJ)	Derogatories	Delinquencies	Log(CLAGE)	Inquiries	Log(CLNO)	Log(DEBTINC)
1	0.08086253	10.9149906	11.331562	0.39380054	0.12668464	0.57951482	0.0403127	0.07008386	0.16172507	0.05660377	0	0	0.67115903	2.05726426	0.19879441	0.37573187	5.14810623	0.87789186	2.88760549	3.38753863		
2	0.14241803	10.6209639	11.1085237	0.61160033	0.38831967	0	0.03432377	0.07479508	0.85706967	0.03125	0.00256348	0	0	0.20584509	0.13430517	0.24105831	5.03214417	0.85546761	2.75268036	3.46000953		
3	0.82657343	10.9226343	11.3347491	0.73706294	0.24335664	0.01958042	0.28951049	0.12867133	0.40979021	0.12867133	0.03636364	0.0013886	0.00559441	1.73278796	1.15678384	1.69487391	4.84296397	2.97340123	3.05271357	3.58002072		
4	0.29787234	11.2004048	11.7593011	0.38829787	0.58510638	0.02659574	0	0	0	0	0	1	0	1.75231128	0.2326209	0.49399016	4.92535699	1.30142316	2.94901891	3.5148848		
5	0.066687198	11.3418391	11.6745262	0.73823141	0.26176859	0	0.19181698	0.26264848	0.06071271	0.45446546	0.03035636	0	0	0.18883285	0.09810085	0.27116468	5.1916107	0.89319931	1.31314523	3.51389028		

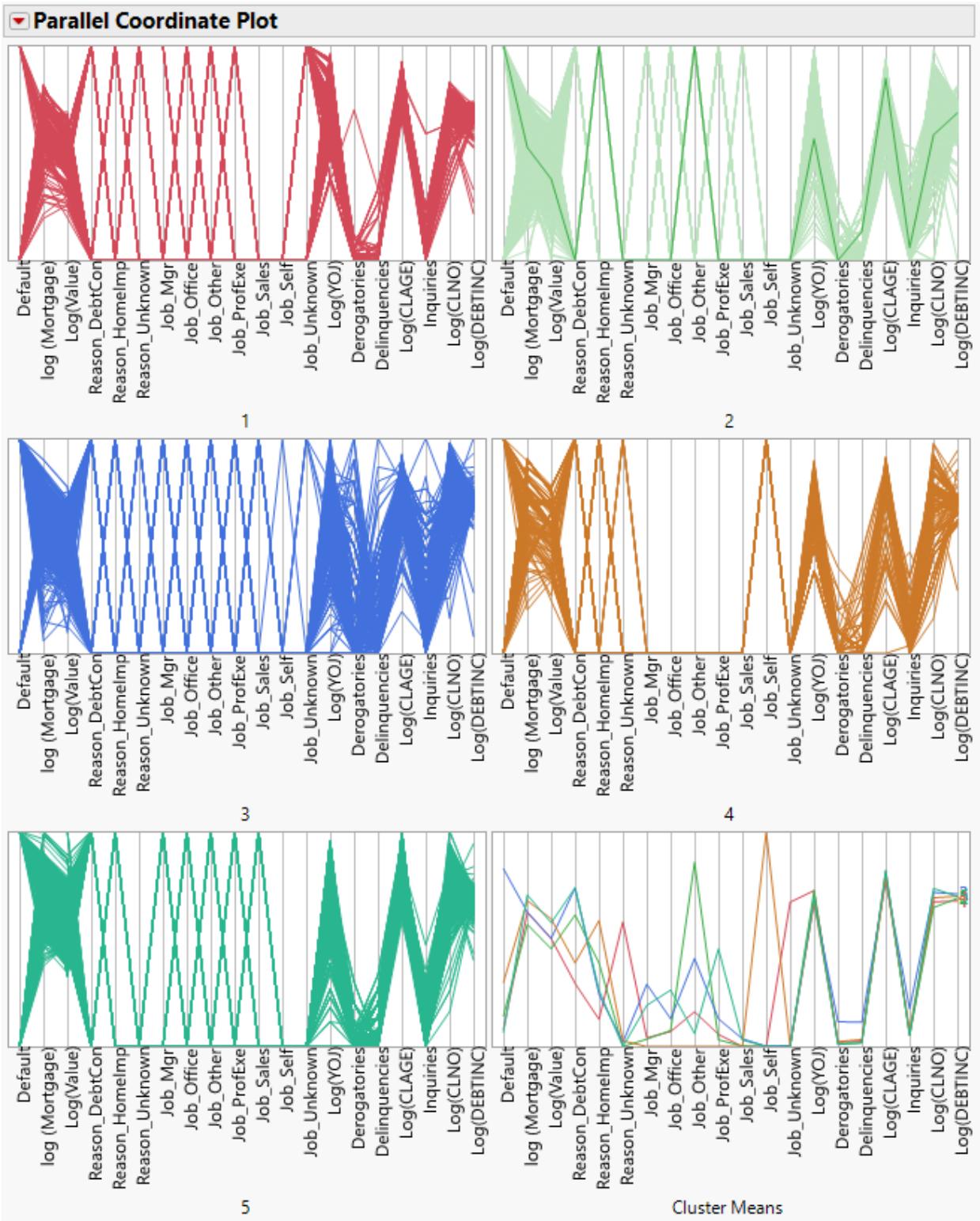
Cluster Standard Deviations		log	Default	(Mortgage)	Log(Value)	Reason_DebtCon	Reason_HomeImp	Reason_Unknown	Job_Mgr	Job_Office	Job_Other	Job_ProfExe	Job_Sales	Job_Self	Job_Unknown	Log(YOJ)	Derogatories	Delinquencies	Log(CLAGE)	Inquiries	Log(CLNO)	Log(DEBTINC)
1	0.27262389	0.57671108	0.44290223	0.45550168	0.33261936	0.4936369	0.19696847	0.25526232	0.36191981	0.23100394	0	0	0.46979207	0.78615185	0.41889042	0.75787049	0.3499724	1.03175184	0.3515488	0.32496505		
2	0.34947828	0.58002468	0.41516778	0.48736794	0	0.18205947	0.26306041	0.35000178	0.17399264	0.05054616	0	0	0.08536156	0.39089961	0.5507494	0.539941	1.11471197	0.54833086	0.2779981			
3	0.37861563	0.60942546	0.45461614	0.44022854	0.4291098	0.13855334	0.45355510	0.33483581	0.49179487	0.33483581	0.18719328	0.03737172	0.07458625	1.05482416	1.71940977	2.10954202	0.58286332	2.8140093	0.4858845	0.24389718		
4	0.45732308	0.78373684	0.56918683	0.48736294	0.49270367	0.16089878	0	0	0	0	0	0	0	0.71708844	0.57608256	1.22887546	0.98220667	1.36573254	0.68108093	0.32816426		
5	0.24980015	0.46832091	0.42606563	0.43959731	0.43959731	0	0.39372809	0.44007301	0.2388026	0.49792229	0.17156687	0	0	0.59079777	0.30177731	0.63091377	0.4103379	1.1171904	0.3817323	0.20290096		

Biplot

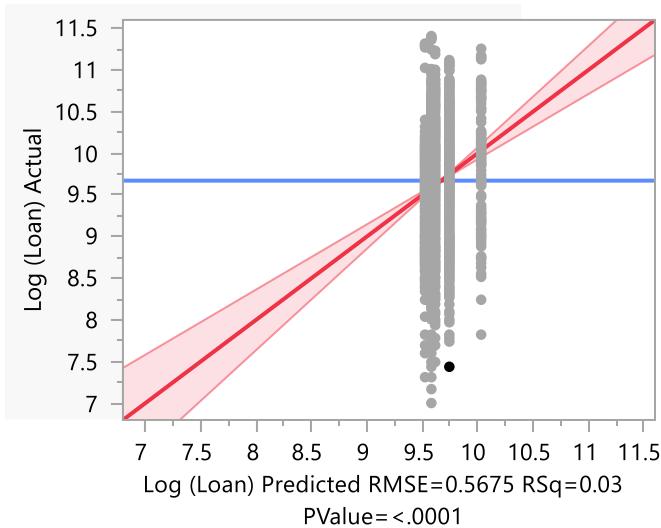


Biplot 3D





Regression with only Cluster Variables



Effect Summary

Source	LogWorth	PValue
Cluster_2	21.120	0.00000
Cluster_1	11.511	0.00000
Cluster_4	10.662	0.00000
Cluster_3	7.085	0.00000

Summary of Fit

RSquare	0.032568
RSquare Adj	0.031918
Root Mean Square Error	0.567467
Mean of Response	9.671748
Observations (or Sum Wgts)	5960

Analysis of Variance

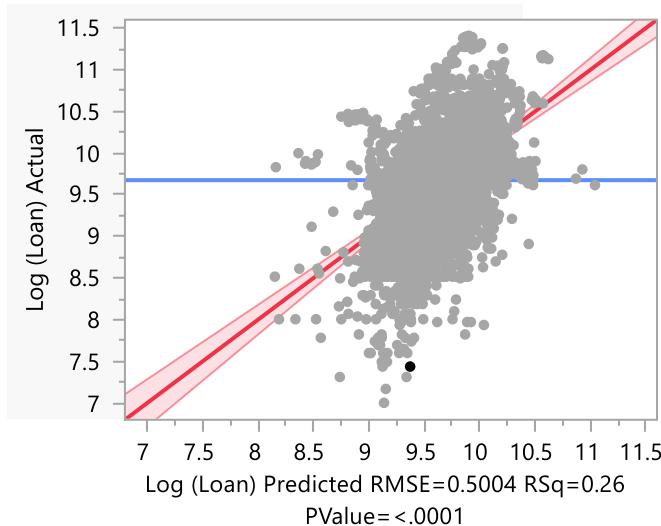
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	64.5552	16.1388	50.1176
Error	5955	1917.6233	0.3220	Prob > F
C. Total	5959	1982.1785		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	9.7448139	0.010853	897.91	<.0001*	.
Cluster_1	-0.219416	0.031397	-6.99	<.0001*	1.0650033
Cluster_2	-0.16215	0.016815	-9.64	<.0001*	1.1526175
Cluster_3	-0.127969	0.023836	-5.37	<.0001*	1.1101813
Cluster_4	0.2869491	0.042786	6.71	<.0001*	1.035051

This is the worst model so far, with an R square adjusted of .032 and RMSE is .57. The variables are all significant and the VIF is low. Overall it is a bad model.

Regression with Cluster Variables and Original-stepwise



Effect Summary

Source	LogWorth	PValue
Log(Value)	166.162	0.00000
Reason_DebtCon	113.196	0.00000
log (Mortgage)	52.726	0.00000
Log(YOJ)	12.022	0.00000
Job_Self	11.165	0.00000
Job_Sales	8.704	0.00000
Inquiries	6.412	0.00000
Log(CLAGE)	5.869	0.00000
Delinquencies	3.327	0.00047
Log(DEBTINC)	2.870	0.00135
Derogatories	2.311	0.00489
Cluster_3	2.272	0.00534
Job_ProfExe	1.792	0.01614

Summary of Fit

RSquare	0.26285
RSquare Adj	0.261117
Root Mean Square Error	0.50036
Mean of Response	9.675073
Observations (or Sum Wgts)	5544

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	13	493.6748	37.9750	151.6817
Error	5530	1384.4892	0.2504	Prob > F
C. Total	5543	1878.1640		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	4.2701249	0.178125	23.97	<.0001*	.
Cluster_3	-0.07576	0.027187	-2.79	0.0053*	1.8386299
log (Mortgage)	-0.279398	0.017958	-15.56	<.0001*	2.9334485
Log(Value)	0.6580654	0.023085	28.51	<.0001*	2.996669
Reason_DebtCon	0.3357256	0.014456	23.22	<.0001*	1.0496671
Job_ProfExe	-0.041687	0.017323	-2.41	0.0161*	1.1323921
Job_Sales	-0.307612	0.051185	-6.01	<.0001*	1.0275775
Job_Self	0.2649937	0.03854	6.88	<.0001*	1.0830459
Log(YOJ)	0.0533746	0.007461	7.15	<.0001*	1.0674594
Derogatories	0.026905	0.009556	2.82	0.0049*	1.2527947
Delinquencies	-0.024685	0.007055	-3.50	0.0005*	1.2978272
Log(CLAGE)	0.0663122	0.013708	4.84	<.0001*	1.1344227
Inquiries	0.0236007	0.004645	5.08	<.0001*	1.284066
Log(DEBTINC)	0.086835	0.027079	3.21	0.0014*	1.0933978

This is a way better than the model than the model that just the cluster data, but very close to the original model that was done using the log variables. This model is about .01 below the original at R squared adjusted .26 and RMSE at .50. All the variables are significant and the VIFs are a low amount.

Please note that the model after a stepwise was run, still needed further editing. The VIFs were key in my decision on which ones should be removed then whatever variable had a p-value that had a value more than .05.

Conclusion

Models	R Sq. Adj	RMSE
Original Data-stepwise	0.2	10018.91
Log of original data-stepwise	0.27	0.49
FA/PCA	0.11	0.55
Cluster	0.03	0.57
Cluster with Log of original data-stepwise	0.26	0.5

Based on the table of the models that were run, the best model is the one that was run with the stepwise log of the original data model. Following closely behind is the stepwise model of cluster with the log of the original data with some additional adjustments. It should be noted that the ANOVA that was done to compare Loan with Reason and Job, shows that the means are very different.

While the two models that were mentioned were the best model, it should be noted that they are not good models. Good models are closer to 1. Based on the graphs on all the regression models, it was obvious that the R squared adjusted would be closer to 0 as the variables were varied with no discernable pattern.