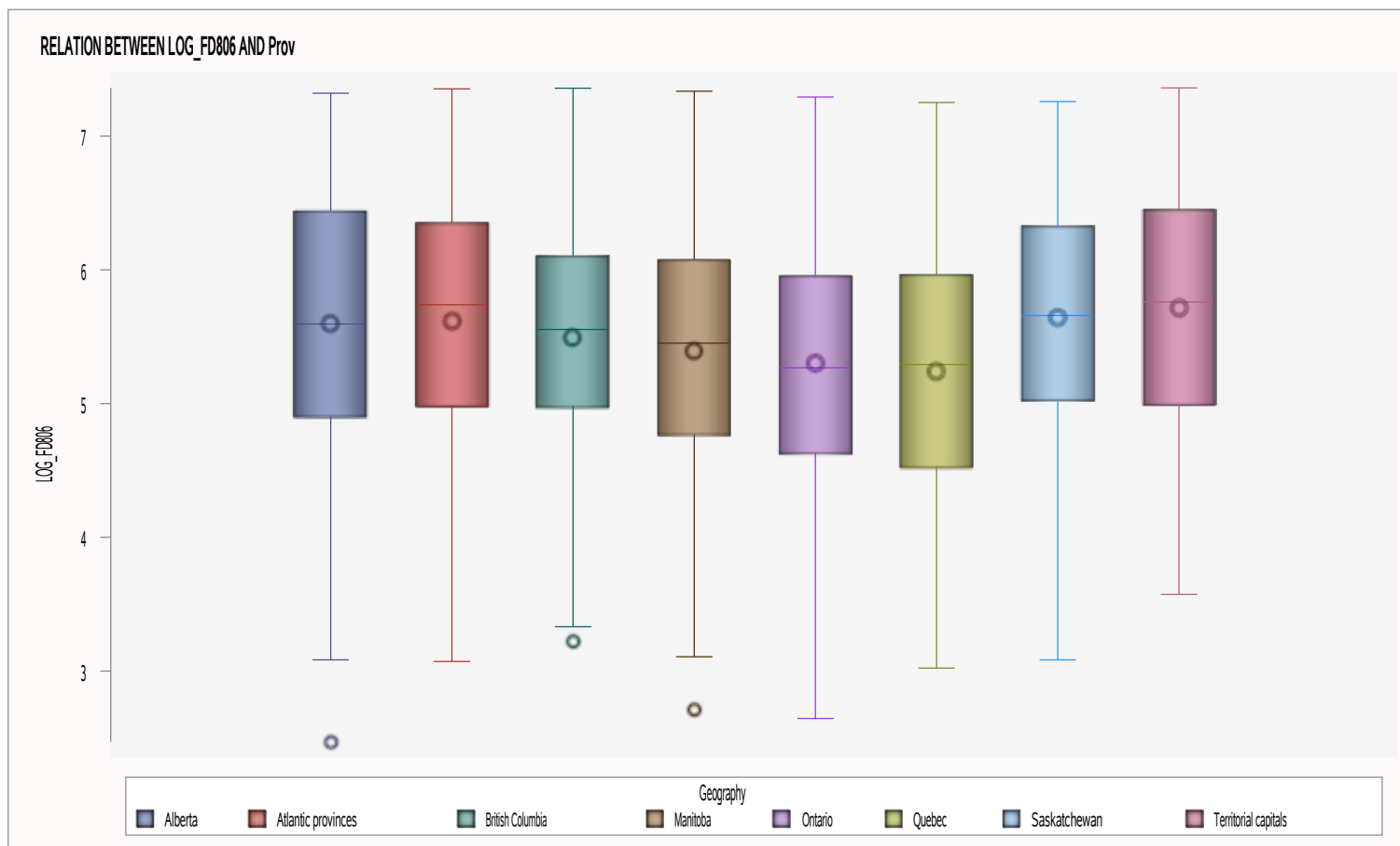


**BIVARIATE ANALYSIS OF Prov AND LOG_FD806 FOR ANA.MODEL2
RELATION BETWEEN LOG_FD806 AND Prov**

14:29 Sunday, November 21, 2021 1

The MEANS Procedure

Analysis Variable : LOG_FD806														
Geography	N Obs	N	N Miss	Minimum	Lower Quartile	Median	Mean	Upper Quartile	Maximum	Quartile Range	Coeff of Variation	Lower 95% CL for Mean	Upper 95% CL for Mean	Skewness
Alberta	958697	958697	0	2.47	4.90	5.60	5.60	6.44	7.32	1.54	17.11	5.60	5.60	-0.29
Atlantic provinces	629768	629768	0	3.07	4.97	5.74	5.62	6.35	7.35	1.38	16.90	5.61	5.62	-0.41
British Columbia	969907	969907	0	3.22	4.97	5.55	5.49	6.10	7.36	1.14	16.17	5.49	5.49	-0.29
Manitoba	299857	299857	0	2.71	4.76	5.45	5.39	6.07	7.34	1.32	17.20	5.39	5.40	-0.23
Ontario	2882375	2882375	0	2.64	4.62	5.27	5.30	5.95	7.29	1.33	17.62	5.30	5.30	0.02
Quebec	2104687	2104687	0	3.02	4.52	5.29	5.24	5.96	7.25	1.44	17.79	5.24	5.24	-0.20
Saskatchewan	271874	271874	0	3.08	5.01	5.66	5.64	6.33	7.26	1.31	15.62	5.64	5.64	-0.29
Territorial capitals	11711	11711	0	3.57	4.99	5.76	5.71	6.45	7.36	1.47	16.60	5.70	5.73	-0.28



One-way ANOVA Assumptions

In order to run a one-way ANOVA the following assumptions must be met:

1.The response of interest is continuous and normally distributed for each treatment group:

Normality test: PROC UNIVARIATE NORMAL and QQPlot for each group.

2.Treatment groups are independent of one another. Experimental units only receive one treatment,and they do not overlap.

3.There are no major outliers.

4.A check for unequal variances will help determine which version of a one-way ANOVA is most appropriate

(Levene's test, Null hypothesis: variances are equal between groups):

A .If variances are equal, then the assumptions of a standard one-way ANOVA are met.

B. If variances are unequal, then a Welch's one-way ANOVA is appropriate.

Normal Distribution?
Null hypothesis: sample has a normal distribution
CLT :
a.If it looks normal and each group have more than 30 observations
b.If moderately skewed, each group must have more than 100 observations
*rule of thumb: If skewness is between -1 and -0.5 or between 0.5 and 1, the distribution is moderately skewed.
*if the sample size is over 2000, the Kolmogorov test should be used. If the sample size is less than 2000, the Shapiro test is better.

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Alberta

Moments			
N	958697	Sum Weights	958697
Mean	5.59782167	Sum Observations	5366614.84
Std Deviation	0.95784837	Variance	0.91747349
Skewness	-0.2933758	Kurtosis	-0.4036799
Uncorrected SS	30920931	Corrected SS	879578.166
Coeff Variation	17.1110911	Std Error Mean	0.00097826

Basic Statistical Measures			
Location		Variability	
Mean	5.597822	Std Deviation	0.95785
Median	5.595083	Variance	0.91747
Mode	6.439478	Range	4.85031
		Interquartile Range	1.54433

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Alberta

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	5722.199	Pr > t	<.0001
Sign	M	479348.5	Pr >= M	<.0001
Signed Rank	S	2.298E11	Pr >= S	<.0001

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.065737	Pr > D	<0.0100
Cramer-von Mises	W-Sq	548.8601	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	3884.935	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.32095
99%	7.30953
95%	7.08369
90%	6.84050
75% Q3	6.43948
50% Median	5.59508
25% Q1	4.89515
10%	4.35773
5%	4.01241
1%	3.39786
0% Min	2.47064

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Alberta

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
2.47064	3940	86	7.18200	4355	49
3.08374	3250	43	7.27414	2361	20
3.39786	5372	151	7.30361	5912	110
3.43108	3237	35	7.30953	8411	40
3.52812	3248	32	7.32095	5649	80

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Atlantic provinces

Moments			
N	629768	Sum Weights	629768
Mean	5.61588116	Sum Observations	3536702.25
Std Deviation	0.94909276	Variance	0.90077707
Skewness	-0.4111892	Kurtosis	-0.3415584
Uncorrected SS	20428979.2	Corrected SS	567279.674
Coeff Variation	16.9001575	Std Error Mean	0.00119596

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Atlantic provinces

Basic Statistical Measures			
Location		Variability	
Mean	5.615881	Std Deviation	0.94909
Median	5.738828	Variance	0.90078
Mode	4.333099	Range	4.28184
		Interquartile Range	1.37771

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	4695.691	Pr > t	<.0001
Sign	M	314884	Pr >= M	<.0001
Signed Rank	S	9.915E10	Pr >= S	<.0001

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.052851	Pr > D	<0.0100
Cramer-von Mises	W-Sq	365.0699	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	2531.387	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.35361
99%	7.32214
95%	7.03277
90%	6.73863

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Atlantic provinces

Quantiles (Definition 5)	
Level	Quantile
75% Q3	6.35141
50% Median	5.73883
25% Q1	4.97369
10%	4.32665
5%	3.89182
1%	3.25810
0% Min	3.07177

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.07177	840	797	7.34003	776	400
3.07177	347	659	7.34357	669	308
3.15274	132	361	7.35014	1360	487
3.21165	209	863	7.35109	759	881
3.21727	186	213	7.35361	491	375

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=British Columbia

Moments			
N	969907	Sum Weights	969907
Mean	5.49258536	Sum Observations	5327296.99
Std Deviation	0.88839624	Variance	0.78924789
Skewness	-0.2858383	Kurtosis	-0.3704013
Uncorrected SS	30026129.7	Corrected SS	765496.26
Coeff Variation	16.174464	Std Error Mean	0.00090207

Basic Statistical Measures			
Location		Variability	
Mean	5.492585	Std Deviation	0.88840
Median	5.553541	Variance	0.78925
Mode	6.246301	Range	4.14116
		Interquartile Range	1.13596

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	6088.848	Pr > t 	<.0001
Sign	M	484953.5	Pr >= M 	<.0001
Signed Rank	S	2.352E11	Pr >= S 	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=British Columbia

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.051237	Pr > D	<0.0100
Cramer-von Mises	W-Sq	491.1901	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	3618.11	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.35843
99%	7.20510
95%	7.04319
90%	6.60932
75% Q3	6.10216
50% Median	5.55354
25% Q1	4.96620
10%	4.14804
5%	3.87867
1%	3.36246
0% Min	3.21727

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=British Columbia

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.21727	1553	1060	7.10462	9539	1053
3.33220	6907	1089	7.19090	3791	946
3.36246	3183	932	7.20510	5467	1076
3.44042	4303	1064	7.21127	1683	1094
3.63627	3779	1007	7.35843	4907	987

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Manitoba

Moments			
N	299857	Sum Weights	299857
Mean	5.39213173	Sum Observations	1616868.45
Std Deviation	0.92721768	Variance	0.85973263
Skewness	-0.2306575	Kurtosis	-0.4166667
Uncorrected SS	8976163.64	Corrected SS	257795.988
Coeff Variation	17.1957535	Std Error Mean	0.00169326

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Manitoba

Basic Statistical Measures			
Location		Variability	
Mean	5.392132	Std Deviation	0.92722
Median	5.451725	Variance	0.85973
Mode	4.481872	Range	4.62243
		Interquartile Range	1.31947

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	3184.461	Pr > t 	<.0001
Sign	M	149928.5	Pr >= M 	<.0001
Signed Rank	S	2.248E10	Pr >= S 	<.0001

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.037604	Pr > D	<0.0100
Cramer-von Mises	W-Sq	95.38586	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	635.7575	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.33580
99%	7.16667
95%	6.82778
90%	6.63637

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Manitoba

Quantiles (Definition 5)	
Level	Quantile
75% Q3	6.07496
50% Median	5.45172
25% Q1	4.75548
10%	4.16200
5%	3.81110
1%	3.25810
0% Min	2.71337

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
2.71337	1520	1314	7.08978	1941	1131
3.10727	381	1270	7.13096	1147	1256
3.25810	2580	1228	7.16667	1520	1148
3.26805	1551	1243	7.27891	505	1267
3.38608	846	1252	7.33580	1566	1130

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Ontario

Moments			
N	2882375	Sum Weights	2882375
Mean	5.30040293	Sum Observations	15277748.9
Std Deviation	0.93412537	Variance	0.87259021
Skewness	0.02419881	Kurtosis	-0.5652113
Uncorrected SS	83493356.2	Corrected SS	2515131.33
Coeff Variation	17.6236672	Std Error Mean	0.00055021

Basic Statistical Measures			
Location		Variability	
Mean	5.300403	Std Deviation	0.93413
Median	5.267394	Variance	0.87259
Mode	4.678792	Range	4.64785
		Interquartile Range	1.32530

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	9633.386	Pr > t 	<.0001
Sign	M	1441188	Pr >= M 	<.0001
Signed Rank	S	2.077E12	Pr >= S 	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Ontario

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.039775	Pr > D	<0.0100
Cramer-von Mises	W-Sq	847.0641	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	7154.464	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.29261
99%	7.24642
95%	6.89946
90%	6.58360
75% Q3	5.94694
50% Median	5.26739
25% Q1	4.62163
10%	4.01530
5%	3.90258
1%	3.30689
0% Min	2.64476

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Ontario

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
2.64476	15313	1596	7.17072	18757	1363
3.30689	15881	1579	7.23335	5785	1355
3.37143	1772	1464	7.24642	6213	1341
3.58129	14896	1552	7.28909	12702	1432
3.58574	6267	1572	7.29261	12196	1415

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Quebec

Moments			
N	2104687	Sum Weights	2104687
Mean	5.23877013	Sum Observations	11025971.4
Std Deviation	0.93216345	Variance	0.86892871
Skewness	-0.1950948	Kurtosis	-0.6476333
Uncorrected SS	59591351.7	Corrected SS	1828822.08
Coeff Variation	17.7935552	Std Error Mean	0.00064254

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Quebec

Basic Statistical Measures			
Location		Variability	
Mean	5.238770	Std Deviation	0.93216
Median	5.292802	Variance	0.86893
Mode	5.754000	Range	4.22920
		Interquartile Range	1.44218

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	8153.255	Pr > t	<.0001
Sign	M	1052344	Pr >= M	<.0001
Signed Rank	S	1.107E12	Pr >= S	<.0001

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.041222	Pr > D	<0.0100
Cramer-von Mises	W-Sq	875.923	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	6349.62	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.25157
99%	7.09923
95%	6.65562
90%	6.42652

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Quebec

Quantiles (Definition 5)	
Level	Quantile
75% Q3	5.96343
50% Median	5.29280
25% Q1	4.52124
10%	3.96119
5%	3.53777
1%	3.27790
0% Min	3.02237

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.02237	4062	1878	6.99027	10294	1908
3.03495	8060	1899	6.99866	5455	1907
3.05777	7665	1906	7.09923	7410	1893
3.27790	16121	1864	7.17749	11846	1787
3.30689	7416	1635	7.25157	2957	1686

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Saskatchewan

Moments			
N	271874	Sum Weights	271874
Mean	5.63878298	Sum Observations	1533038.48
Std Deviation	0.88073409	Variance	0.77569254
Skewness	-0.2887928	Kurtosis	-0.5745512
Uncorrected SS	8855361.17	Corrected SS	210889.857
Coeff Variation	15.619223	Std Error Mean	0.00168912

Basic Statistical Measures			
Location		Variability	
Mean	5.638783	Std Deviation	0.88073
Median	5.658018	Variance	0.77569
Mode	3.814631	Range	4.17511
		Interquartile Range	1.31152

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	3338.293	Pr > t 	<.0001
Sign	M	135937	Pr >= M 	<.0001
Signed Rank	S	1.848E10	Pr >= S 	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Saskatchewan

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.055935	Pr > D	<0.0100
Cramer-von Mises	W-Sq	175.6252	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1242.856	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.25885
99%	7.18062
95%	6.97271
90%	6.76542
75% Q3	6.32522
50% Median	5.65802
25% Q1	5.01370
10%	4.50347
5%	4.15685
1%	3.58269
0% Min	3.08374

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Saskatchewan

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.08374	1372	2065	7.14419	1518	2141
3.34850	523	2010	7.18054	888	1925
3.58269	838	2152	7.18062	1691	2027
3.70549	629	2082	7.20541	1060	1976
3.72014	497	1953	7.25885	1377	1981

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Territorial capitals

Moments			
N	11711	Sum Weights	11711
Mean	5.71392282	Sum Observations	66915.7502
Std Deviation	0.94859878	Variance	0.89983964
Skewness	-0.2810241	Kurtosis	-0.9449775
Uncorrected SS	392888.554	Corrected SS	10537.1222
Coeff Variation	16.6015329	Std Error Mean	0.00876568

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Territorial capitals

Basic Statistical Measures			
Location		Variability	
Mean	5.713923	Std Deviation	0.94860
Median	5.759879	Variance	0.89984
Mode	4.574092	Range	3.78783
		Interquartile Range	1.46558

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	651.8517	Pr > t 	<.0001
Sign	M	5855.5	Pr >= M 	<.0001
Signed Rank	S	34289808	Pr >= S 	<.0001

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.082104	Pr > D	<0.0100
Cramer-von Mises	W-Sq	16.42393	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	115.1312	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Level	Quantile
100% Max	7.36074
99%	7.30971
95%	7.08256
90%	6.93674

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Geography=Territorial capitals

Quantiles (Definition 5)	
Level	Quantile
75% Q3	6.45063
50% Median	5.75988
25% Q1	4.98504
10%	4.25135
5%	4.13772
1%	3.78328
0% Min	3.57291

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.57291	18	2172	7.27783	46	2311
3.78328	115	2270	7.28773	45	2168
3.85757	157	2298	7.30971	58	2171
3.91562	58	2310	7.32818	59	2180
3.94119	33	2299	7.36074	43	2163

Null hypothesis: equal variances

a.If variances are equal, then a pooled t-test is appropriate

b.If variances are unequal, then a Satterthwaite (also known as Welch's) test is appropriate

The GLM Procedure

Class Level Information		
Class	Levels	Values
Prov	8	Alberta Atlantic provinces British Columbia Manitoba Ontario Quebec Saskatchewan Territorial capitals

Number of Observations Read	2327
Number of Observations Used	2327
Sum of Frequencies Read	8128876
Sum of Frequencies Used	8128876

The GLM Procedure

Dependent Variable: LOG_FD806

Frequency: WeightD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	172580.537	24654.362	28485.7	<.0001
Error	8.13E6	7035530.477	0.865		
Corrected Total	8.13E6	7208111.014			

R-Square	Coeff Var	Root MSE	LOG_FD806 Mean
0.023943	17.28520	0.930322	5.382190

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Prov	7	172580.5368	24654.3624	28485.7	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Prov	7	172580.5368	24654.3624	28485.7	<.0001

The GLM Procedure

Levene's Test for Homogeneity of LOG_FD806 Variance ANOVA of Absolute Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Prov	7	4166.6	595.2	2107.67	<.0001
Error	8.13E6	2295703	0.2824		

Welch's ANOVA for LOG_FD806			
Source	DF	F Value	Pr > F
Prov	7.0000	28314.2	<.0001
Error	220971		

The GLM Procedure

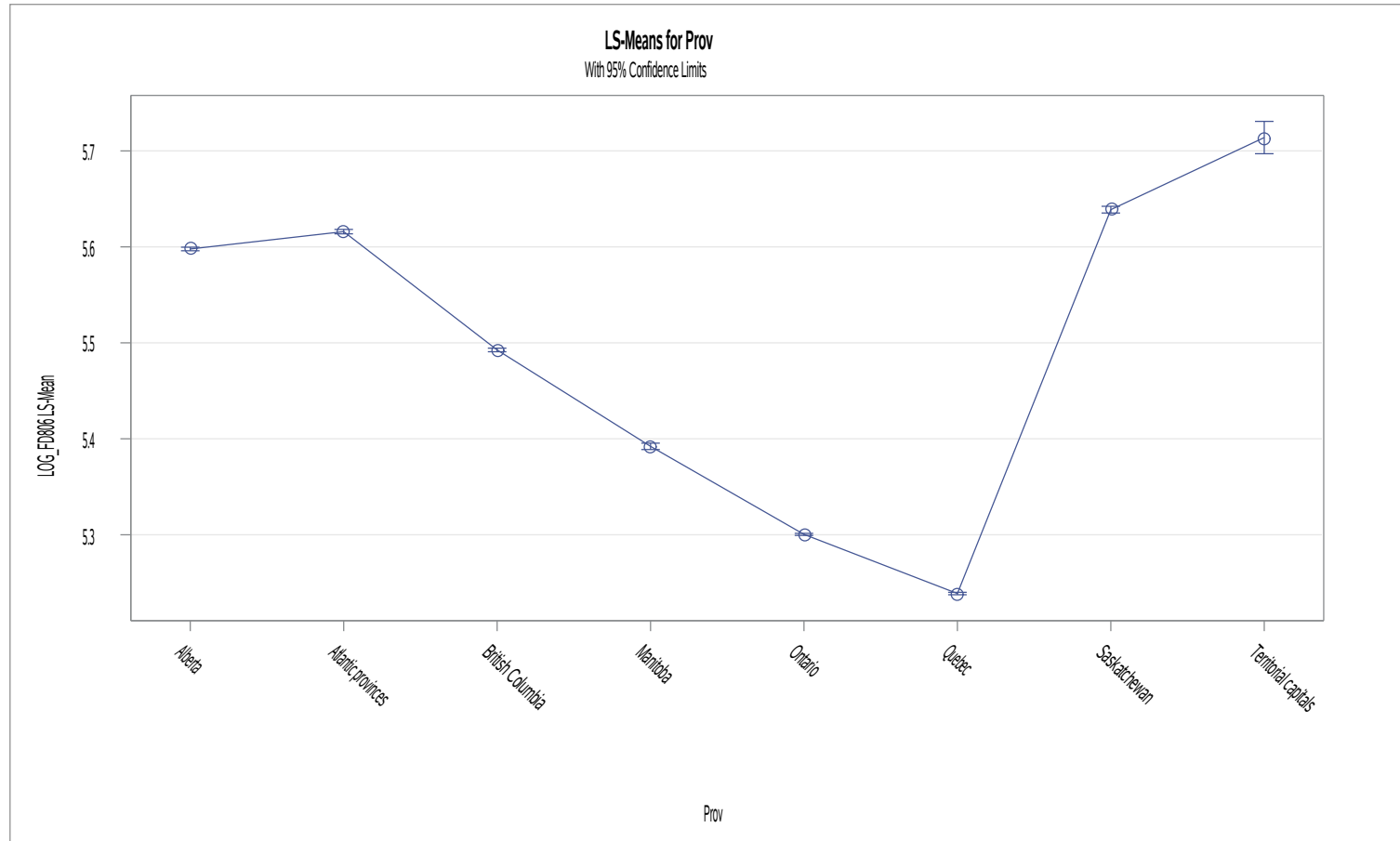
Level of Prov	N	LOG_FD806	
		Mean	Std Dev
Alberta	958697	5.59782167	0.95784837
Atlantic provinces	629768	5.61588116	0.94909276
British Columbia	969907	5.49258536	0.88839624
Manitoba	299857	5.39213173	0.92721768
Ontario	2882375	5.30040293	0.93412537
Quebec	2104687	5.23877013	0.93216345
Saskatchewan	271874	5.63878298	0.88073409
Territorial capitals	11711	5.71392282	0.94859878

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

Prov	LOG_FD806 LSMEAN	LSMEAN Number
Alberta	5.59782167	1
Atlantic provinces	5.61588116	2
British Columbia	5.49258536	3
Manitoba	5.39213173	4
Ontario	5.30040293	5
Quebec	5.23877013	6
Saskatchewan	5.63878298	7
Territorial capitals	5.71392282	8

[illegible]

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer



The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

