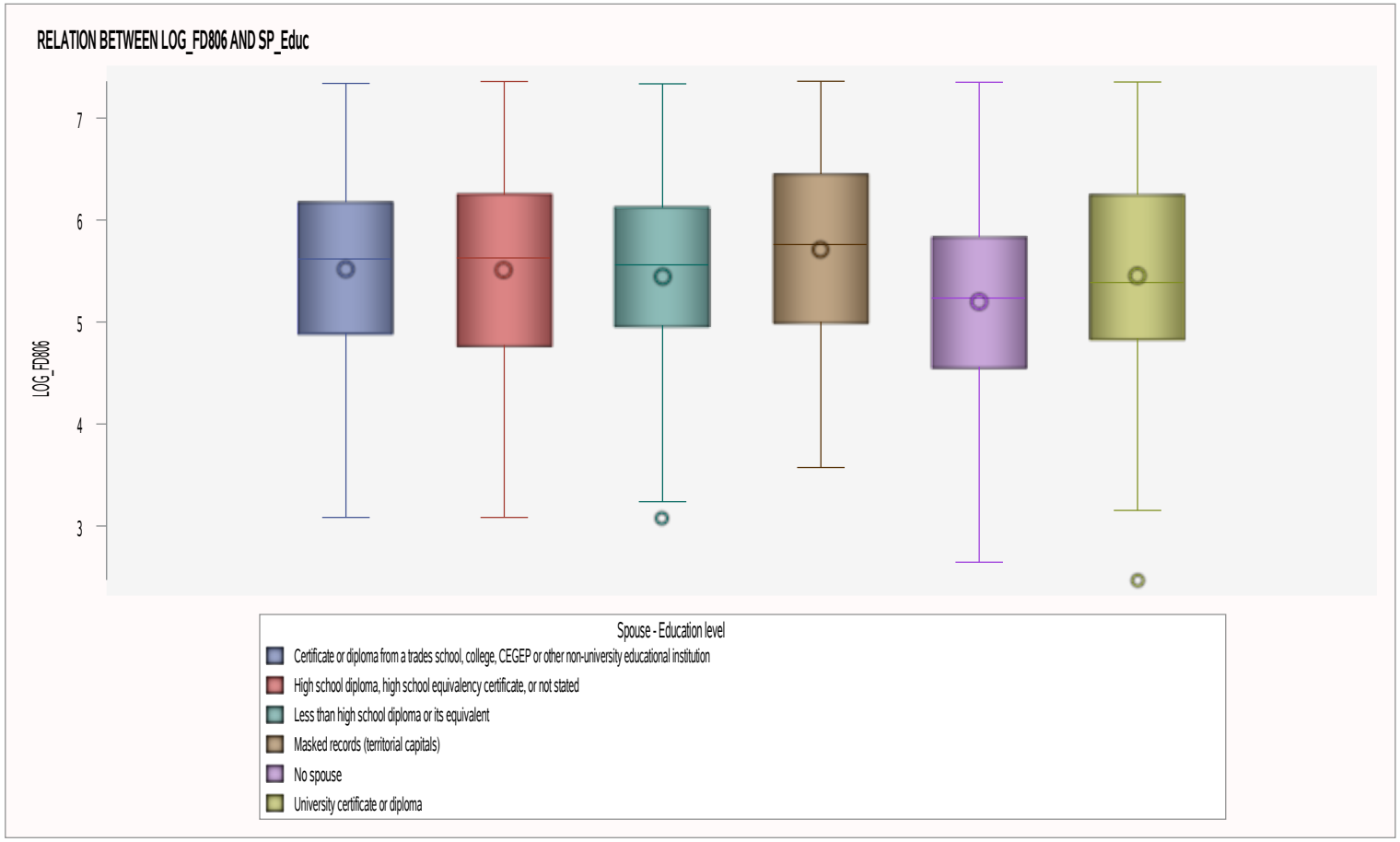
BIVARIATE ANALYSIS OF SP_Educ AND LOG_FD806 FOR ANA.MODEL2
RELATION BETWEEN LOG_FD806 AND SP_Educ

The MEANS Procedure

Analysis Variable : LOG_FD806														
Spouse - Education level	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
Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution	1636840	1636840	0	3.08	4.88	5.62	5.51	6.17	7.34	1.30	16.57	5.51	5.51	-0.29
High school diploma, high school equivalency certificate, or not stated	1358512	1358512	0	3.08	4.76	5.63	5.50	6.25	7.36	1.49	17.09	5.50	5.51	-0.16
Less than high school diploma or its equivalent	546094	546094	0	3.07	4.95	5.56	5.44	6.12	7.34	1.17	17.48	5.44	5.45	-0.46
Masked records (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
No spouse	2903242	2903242	0	2.64	4.54	5.23	5.20	5.83	7.35	1.29	18.07	5.20	5.20	-0.06
University certificate or diploma	1672477	1672477	0	2.47	4.82	5.39	5.45	6.25	7.35	1.42	16.87	5.45	5.45	-0.07



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 Kolmgorov 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

Spouse - Education level=Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution

Moments			
N	1636840	**Sum Weights**	1636840
Mean	5.51243153	**Sum Observations**	9022968.43
Std Deviation	0.91359992	**Variance**	0.83466481
Skewness	-0.2903323	**Kurtosis**	-0.5377288
Uncorrected SS	51104707.6	**Corrected SS**	1366211.91
Coeff Variation	16.573447	**Std Error Mean**	0.00071409

Basic Statistical Measures			
Location		**Variability**	
Mean	5.512432	**Std Deviation**	0.91360
Median	5.618479	**Variance**	0.83466
Mode	3.931041	**Range**	4.25621
		Interquartile Range	1.29659

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	7719.52	**Pr > |t|**	<.0001
Sign	**M**	818420	**Pr >= |M|**	<.0001
Signed Rank	**S**	6.698E11	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.0592	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	1013.45	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	6911.437	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.33995
99%	7.30953
95%	6.90938
90%	6.65639
75% Q3	6.17476
50% Median	5.61848
25% Q1	4.87817
10%	4.10710
5%	3.93104
1%	3.31200
0% Min	3.08374

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.08374	3250	188	7.25923	1473	4
3.21727	186	439	7.30953	8411	61
3.22764	463	393	7.32095	5649	331
3.25810	2580	472	7.32214	2134	254
3.25810	975	366	7.33995	1345	364

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=High school diploma, high school equivalency certificate, or not stated

Moments			
N	1358512	Sum Weights	1358512
Mean	5.50446956	Sum Observations	7477887.96
Std Deviation	0.94074817	Variance	0.88500712
Skewness	-0.1577265	Kurtosis	-0.8783725
Uncorrected SS	42364098.6	Corrected SS	1202291.91
Coeff Variation	17.0906235	Std Error Mean	0.00080713

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=High school diploma, high school equivalency certificate, or not stated

Basic Statistical Measures			
Location		**Variability**	
Mean	5.504470	**Std Deviation**	0.94075
Median	5.628340	**Variance**	0.88501
Mode	5.680343	**Range**	4.27469
		Interquartile Range	1.48858

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	6819.834	**Pr > |t|**	<.0001
Sign	**M**	679256	**Pr >= |M|**	<.0001
Signed Rank	**S**	4.614E11	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.065	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	1311.182	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	8675.946	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.35843
99%	7.29261
95%	6.89278
90%	6.77203

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=High school diploma, high school equivalency certificate, or not stated

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	6.24630
50% Median	5.62834
25% Q1	4.75772
10%	4.09351
5%	3.95124
1%	3.52812
0% Min	3.08374

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
3.08374	1372	590	7.20541	1060	592
3.26805	1144	647	7.26582	1379	656
3.30689	7416	577	7.28220	1128	484
3.32215	969	811	7.29261	12196	501
3.38912	1988	816	7.35843	4907	533

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=Less than high school diploma or its equivalent

Moments			
N	546094	**Sum Weights**	546094
Mean	5.44420986	**Sum Observations**	2973050.34
Std Deviation	0.95165419	**Variance**	0.90564569
Skewness	-0.4554079	**Kurtosis**	-0.4824408
Uncorrected SS	16680476.7	**Corrected SS**	494566.772
Coeff Variation	17.4801158	**Std Error Mean**	0.00128779

Basic Statistical Measures			
Location		**Variability**	
Mean	5.444210	**Std Deviation**	0.95165
Median	5.560682	**Variance**	0.90565
Mode	5.560682	**Range**	4.26404
		Interquartile Range	1.16998

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	4227.556	**Pr > |t|**	<.0001
Sign	**M**	273047	**Pr >= |M|**	<.0001
Signed Rank	**S**	7.455E10	**Pr >= |S|**	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=Less than high school diploma or its equivalent

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.094912	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	1096.292	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	6833.732	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.33580
99%	7.30361
95%	6.80801
90%	6.59266
75% Q3	6.12030
50% Median	5.56068
25% Q1	4.95032
10%	4.00497
5%	3.63627
1%	3.38912
0% Min	3.07177

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=Less than high school diploma or its equivalent

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.07177	347	887	7.24213	495	889
3.23789	987	981	7.25885	1377	869
3.26805	1551	967	7.27338	1081	858
3.38912	7954	989	7.30361	5912	909
3.47010	606	884	7.33580	1566	850

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=Masked records (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

Spouse - Education level=Masked records (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

Spouse - Education level=Masked records (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	1025	7.27783	46	1177
3.78328	115	1144	7.28773	45	1067
3.85757	157	1150	7.30971	58	1091
3.91562	58	1176	7.32818	59	1006
3.94119	33	1151	7.36074	43	1057

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=No spouse

Moments			
N	2903242	**Sum Weights**	2903242
Mean	5.19785064	**Sum Observations**	15090618.3
Std Deviation	0.93931279	**Variance**	0.88230853
Skewness	-0.0588125	**Kurtosis**	-0.3863747
Uncorrected SS	81000334.2	**Corrected SS**	2561554.29
Coeff Variation	18.0711771	**Std Error Mean**	0.00055128

Basic Statistical Measures			
Location		**Variability**	
Mean	5.197851	**Std Deviation**	0.93931
Median	5.234951	**Variance**	0.88231
Mode	5.958166	**Range**	4.70633
		Interquartile Range	1.28658

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	9428.773	**Pr > |t|**	<.0001
Sign	**M**	1451621	**Pr >= |M|**	<.0001
Signed Rank	**S**	2.107E12	**Pr >= |S|**	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=No spouse

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.034146	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	363.4049	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	2804.009	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.35109
99%	7.11486
95%	6.81928
90%	6.40723
75% Q3	5.83115
50% Median	5.23495
25% Q1	4.54457
10%	3.95891
5%	3.66356
1%	3.05777
0% Min	2.64476

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=No spouse

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
2.64476	15313	1556	7.27547	1311	1266
2.71337	1520	1771	7.27891	505	1680
3.02237	4062	1537	7.28345	442	1191
3.03495	8060	1580	7.33743	836	1218
3.05777	7665	1594	7.35109	759	1403

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=University certificate or diploma

Moments			
N	1672477	**Sum Weights**	1672477
Mean	5.45282014	**Sum Observations**	9119716.26
Std Deviation	0.91964225	**Variance**	0.84574186
Skewness	-0.0672169	**Kurtosis**	-0.5518985
Uncorrected SS	51142655.5	**Corrected SS**	1414482.97
Coeff Variation	16.8654425	**Std Error Mean**	0.00071111

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=University certificate or diploma

Basic Statistical Measures			
Location		**Variability**	
Mean	5.452820	**Std Deviation**	0.91964
Median	5.387792	**Variance**	0.84574
Mode	6.488141	**Range**	4.88297
		Interquartile Range	1.42373

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	7668.004	**Pr > |t|**	<.0001
Sign	**M**	836238.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	6.993E11	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.057485	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	692.6115	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	4795.035	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.35361
99%	7.27414
95%	7.07052
90%	6.63154

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Spouse - Education level=University certificate or diploma

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	6.24731
50% Median	5.38779
25% Q1	4.82358
10%	4.23208
5%	3.95623
1%	3.41838
0% Min	2.47064

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
2.47064	3940	2267	7.28909	12702	2244
3.15274	132	2002	7.34003	776	2014
3.33506	5931	1949	7.34357	669	1999
3.34850	523	2268	7.35014	1360	2280
3.35341	1062	2013	7.35361	491	2153

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**
SP_Educ	6	Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution High school diploma, high school equivalency certificate, or not stated Less than high school diploma or its equivalent Masked records (territorial capitals) No spouse University certificate or diploma

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	5	158466.050	31693.210	36545.1	<.0001
Error	8.13E6	7049644.964	0.867		
Corrected Total	8.13E6	7208111.014			

R-Square	Coeff Var	Root MSE	LOG_FD806 Mean
0.021984	17.30253	0.931255	5.382190

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SP_Educ	5	158466.0500	31693.2100	36545.1	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SP_Educ	5	158466.0500	31693.2100	36545.1	<.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
SP_Educ	5	2139.7	427.9	1497.03	<.0001
Error	8.13E6	2323750	0.2859		

Welch's ANOVA for LOG_FD806			
Source	DF	F Value	Pr > F
SP_Educ	5.0000	36322.2	<.0001
Error	132308		

The GLM Procedure

Level of SP_Educ	N	LOG_FD806	
		Mean	Std Dev
Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution	1636840	5.51243153	0.91359992
High school diploma, high school equivalency certificate, or not stated	1358512	5.50446956	0.94074817
Less than high school diploma or its equivalent	546094	5.44420986	0.95165419
Masked records (territorial capitals)	11711	5.71392282	0.94859878
No spouse	2903242	5.19785064	0.93931279
University certificate or diploma	1672477	5.45282014	0.91964225

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

SP_Educ	LOG_FD806 LSMEAN	LSMEAN Number
Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution	5.51243153	1
High school diploma, high school equivalency certificate, or not stated	5.50446956	2
Less than high school diploma or its equivalent	5.44420986	3
Masked records (territorial capitals)	5.71392282	4
No spouse	5.19785064	5
University certificate or diploma	5.45282014	6

Least Squares Means for effect SP_Educ
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: LOG_FD806

i/j	1	2	3	4	5	6
1		<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001	

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

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