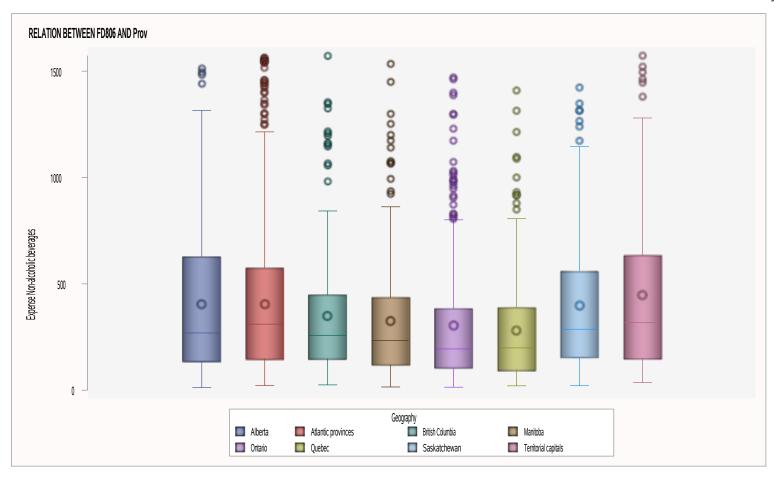
BIVARIATE ANALYSIS OF Prov AND FD806 FOR ANA.MODEL1 **RELATION BETWEEN FD806 AND Prov**

The MEANS Procedure

	Analysis Variable: FD806 Expense Non-alcoholic beverages													
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	11.83	133.64	269.10	404.36	626.08	1511.64	492.44	87.30	403.65	405.06	1.28
Atlantic provinces	629768	629768	0	21.58	144.56	310.70	404.59	573.30	1561.82	428.74	84.86	403.75	405.44	1.36
British Columbia	969907	969907	0	24.96	143.48	258.15	346.77	446.82	1569.37	303.34	86.11	346.17	347.36	1.68
Manitoba	299857	299857	0	15.08	116.22	233.16	324.21	434.83	1534.26	318.61	89.20	323.17	325.24	1.59
Ontario	2882375	2882375	0	14.08	101.66	193.91	305.37	382.58	1469.40	280.92	98.27	305.03	305.72	1.84
Quebec	2104687	2104687	0	20.54	91.95	198.90	279.25	388.94	1410.32	296.99	89.11	278.91	279.58	1.59
Saskatchewan	271874	271874	0	21.84	150.46	286.58	396.78	558.48	1420.62	408.02	79.99	395.59	397.97	1.15
Territorial capitals	11711	11711	0	35.62	146.21	317.31	447.33	633.10	1573.00	486.89	82.03	440.68	453.98	1.02



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.lf 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: FD806 (Expense Non-alcoholic beverages)

Freq: WeightD

Geography=Alberta

Moments					
N	958697	Sum Weights	958697		
Mean	404.357996	Sum Observations	387656798		
Std Deviation	353.006082	Variance	124613.294		
Skewness	1.28102241	Kurtosis	1.01480062		
Uncorrected SS	2.76218E11	Corrected SS	1.19466E11		
Coeff Variation	87.3003836	Std Error Mean	0.36053008		

Basic Statistical Measures					
Location		Variability			
Mean	404.3580	Std Deviation	353.00608		
Median	269.1000	Variance	124613		
Mode	626.0800	Range	1500		
		Interquartile Range	492.44000		

Freq: WeightD

Geography=Alberta

Tests for Location: Mu0=0					
Test	St	atistic	p Val	ue	
Student's t	t	1121.565	Pr > t	<.0001	
Sign	М	479348.5	Pr >= M	<.0001	
Signed Rank	s	2.298E11	Pr >= S	<.0001	

Tests for Normality					
Test	Statistic p Value			ue	
Kolmogorov-Smirnov	D	0.165296	Pr > D	<0.0100	
Cramer-von Mises	W-Sq	7815.48	Pr > W-Sq	<0.0050	
Anderson-Darling	A-Sq	45258.44	Pr > A-Sq	<0.0050	

Quantiles (Definition 5)		
Level	Quantile	
100% Max	1511.64	
99%	1494.48	
95%	1192.36	
90%	934.96	
75% Q3	626.08	
50% Median	269.10	
25% Q1	133.64	
10%	78.08	
5%	55.28	
1%	29.90	
0% Min	11.83	

Freq: WeightD

Geography=Alberta

Extreme Observations						
Lowest			Highest			
Value	Freq	Obs	Value	Freq	Obs	
11.83	3940	103	1315.54	4355	6	
21.84	3250	66	1442.51	2361	14	
29.90	5372	152	1485.65	5912	108	
30.91	3237	59	1494.48	8411	3	
34.06	3248	19	1511.64	5649	123	

The UNIVARIATE Procedure Variable: FD806 (Expense Non-alcoholic beverages)

Freq: WeightD

Geography=Atlantic provinces

Moments					
N	629768	Sum Weights	629768		
Mean	404.594424	Sum Observations	254800621		
Std Deviation	343.356582	Variance	117893.743		
Skewness	1.35825064	Kurtosis	1.46298435		
Uncorrected SS	1.77336E11	Corrected SS	7.42456E10		
Coeff Variation	84.8643882	Std Error Mean	0.43266831		

Freq: WeightD

Geography=Atlantic provinces

Basic Statistical Measures					
Location		Variability			
Mean	404.5944	Std Deviation	343.35658		
Median	310.7000	Variance	117894		
Mode	76.1800	Range	1540		
		Interquartile Range	428.74000		

Tests for Location: Mu0=0					
Test	St	atistic	p Val	lue	
Student's t	t	935.1145	Pr > t	<.0001	
Sign	М	314884	Pr >= M	<.0001	
Signed Rank	s	9.915E10	Pr >= S	<.0001	

Tests for Normality				
Test	Statistic p Value			ue
Kolmogorov-Smirnov	D	0.133816	Pr > D	<0.0100
Cramer-von Mises	W-Sq	4094.321	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	25014.54	Pr > A-Sq	<0.0050

Quantiles (Definition 5)				
Level	Quantile			
100% Max	1561.82			
99%	1513.44			
95%	1133.16			
90%	844.40			

Freq: WeightD

Geography=Atlantic provinces

Quantiles (Definition 5)					
Level	Quantile				
75% Q3	573.30				
50% Median	310.70				
25% Q1	144.56				
10%	75.69				
5%	49.00				
1%	26.00				
0% Min	21.58				

Extreme Observations						
ı	owest		Hi	ighest		
Value	Freq	Obs	Value	Freq	Obs	
21.58	840	866	1540.76	776	314	
21.58	347	477	1546.22	669	397	
23.40	132	211	1556.41	1360	458	
24.82	209	844	1557.89	759	825	
24.96	186	246	1561.82	491	335	

Freq: WeightD

Geography=British Columbia

Moments							
N	969907	Sum Weights	969907				
Mean	346.767865	Sum Observations	336332580				
Std Deviation	298.615022	Variance	89170.9313				
Skewness	1.68376037	Kurtosis	2.83724693				
Uncorrected SS	SS 2.03117E11 Corrected SS		8.64874E10				
Coeff Variation	86.1138103	Std Error Mean	0.30321215				

Basic Statistical Measures							
Location Variability							
Mean	346.7679	Std Deviation	298.61502				
Median	258.1500	Variance	89171				
Mode	516.1000	Range	1544				
		Interquartile Range	303.34000				

Tests for Location: Mu0=0							
Test	Statistic p Value						
Student's t	t	1143.648	Pr > t	<.0001			
Sign	М	484953.5	Pr >= M	<.0001			
Signed Rank	s	2.352E11	Pr >= S	<.0001			

Freq: WeightD

Geography=British Columbia

Tests for Normality							
Test Statistic p Value							
Kolmogorov-Smirnov	D	0.141356	Pr > D	<0.0100			
Cramer-von Mises	W-Sq	7488.665	Pr > W-Sq	<0.0050			
Anderson-Darling	A-Sq	46314.23	Pr > A-Sq	<0.0050			

Quantiles (Definition 5)				
Level	Quantile			
100% Max	1569.37			
99%	1346.28			
95%	1145.04			
90%	741.98			
75% Q3	446.82			
50% Median	258.15			
25% Q1	143.48			
10%	63.31			
5%	48.36			
1%	28.86			
0% Min	24.96			

Freq: WeightD

Geography=British Columbia

	Extreme Observations							
ı	Lowest		н	ighest				
Value Freq Obs			Value	Freq	Obs			
24.96	1553	1057	1217.58	9539	1056			
28.00	6907	1095	1327.30	3791	936			
28.86	3183	951	1346.28	5467	1104			
31.20	4303	1074	1354.61	1683	1094			
37.95	3779	1046	1569.37	4907	990			

The UNIVARIATE Procedure Variable: FD806 (Expense Non-alcoholic beverages)

Freq: WeightD

Geography=Manitoba

Moments							
N	299857	299857					
Mean	324.209684	Sum Observations	97216543.3				
Std Deviation	289.205287	Variance	83639.6978				
Skewness	1.58681127	Kurtosis	2.46460341				
Uncorrected SS	5.65984E10	Corrected SS	2.50799E10				
Coeff Variation	89.203161	Std Error Mean	0.52814009				

Freq: WeightD

Geography=Manitoba

Basic Statistical Measures							
Loc	Location Variability						
Mean	324.2097	Std Deviation	289.20529				
Median	233.1600	Variance	83640				
Mode	88.4000	Range	1519				
		Interquartile Range	318.61000				

Tests for Location: Mu0=0							
Test	Statistic p Value						
Student's t	t	613.8706	Pr > t	<.0001			
Sign	М	149928.5	Pr >= M	<.0001			
Signed Rank	s	2.248E10	Pr >= S	<.0001			

Tests for Normality							
Test Statistic p Value							
Kolmogorov-Smirnov	D	0.160308	Pr > D	<0.0100			
Cramer-von Mises	W-Sq	2594.704	Pr > W-Sq	<0.0050			
Anderson-Darling	A-Sq	14984.59	Pr > A-Sq	<0.0050			

Quantiles (Definition 5)			
Level	Quantile		
100% Max	1534.26		
99%	1295.52		
95%	923.14		
90%	762.32		

Freq: WeightD

Geography=Manitoba

Quantiles (Definition 5)				
Level	Quantile			
75% Q3	434.83			
50% Median	233.16			
25% Q1	116.22			
10%	64.20			
5%	45.20			
1%	26.00			
0% Min	15.08			

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
15.08	1520	1317	1199.64	1941	1130
22.36	381	1272	1250.08	1147	1222
26.00	2580	1253	1295.52	1520	1141
26.26	1551	1237	1449.41	505	1270
29.55	846	1231	1534.26	1566	1186

Freq: WeightD

Geography=Ontario

Moments					
N	2882375	Sum Weights	2882375		
Mean	305.372478	Sum Observations	880197996		
Std Deviation	300.076977	Variance	90046.192		
Skewness	1.84284509	Kurtosis	3.20371507		
Uncorrected SS	5.28335E11	Corrected SS	2.59547E11		
Coeff Variation	98.2658879	Std Error Mean	0.17674919		

Basic Statistical Measures				
Location Variability				
Mean	305.3725	Std Deviation	300.07698	
Median	193.9100	Variance	90046	
Mode	107.6400	Range	1455	
		Interquartile Range	280.92000	

Tests for Location: Mu0=0					
Test	Statistic p Value				
Student's t	t	1727.716	Pr > t	<.0001	
Sign	М	1441188	Pr >= M	<.0001	
Signed Rank	s	2.077E12	Pr >= S	<.0001	

Freq: WeightD

Geography=Ontario

Tests for Normality				
Test	Statistic p Value			
Kolmogorov-Smirnov	D	0.1755	Pr > D	<0.0100
Cramer-von Mises	W-Sq	34395.01	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	196675.8	Pr > A-Sq	<0.0050

Quantiles (Definition 5)			
Level	Quantile		
100% Max	1469.40		
99%	1403.07		
95%	991.74		
90%	723.14		
75% Q3	382.58		
50% Median	193.91		
25% Q1	101.66		
10%	55.44		
5%	49.53		
1%	27.30		
0% Min	14.08		

Freq: WeightD

Geography=Ontario

Extreme Observations						
	Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs	
14.08	15313	1591	1300.78	18757	1394	
27.30	15881	1597	1384.86	5785	1338	
29.12	1772	1455	1403.07	6213	1386	
35.92	14896	1545	1464.24	12702	1429	
36.08	6267	1599	1469.40	12196	1353	

The UNIVARIATE Procedure Variable: FD806 (Expense Non-alcoholic beverages)

Freq: WeightD

Geography=Quebec

Moments					
N	2104687	Sum Weights	2104687		
Mean	279.245849	Sum Observations	587725108		
Std Deviation	248.826064	Variance	61914.4104		
Skewness	1.59416598	Kurtosis	2.70037194		
Uncorrected SS	2.9443E11	Corrected SS	1.3031E11		
Coeff Variation	89.1064506	Std Error Mean	0.171515		

Freq: WeightD

Geography=Quebec

	Basic Statistical Measures				
Loc	Location Variability				
Mean	279.2458	Std Deviation	248.82606		
Median	198.9000	Variance	61914		
Mode	315.4500	Range	1390		
		Interquartile Range	296.99000		

Tests for Location: Mu0=0					
Test	Statistic p Value				
Student's t	t 1628.113		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 0.15269		Pr > D	<0.0100	
Cramer-von Mises	W-Sq	16885.38	Pr > W-Sq	<0.0050	
Anderson-Darling	A-Sq	97685.35	Pr > A-Sq	<0.0050	

Quantiles (Definition 5)				
Level Quanti				
100% Max	1410.32			
99%	1211.03			
95%	777.14			
90%	618.02			

Freq: WeightD

Geography=Quebec

Quantiles (Definition 5)				
Level Quantil				
75% Q3	388.94			
50% Median	198.90			
25% Q1	91.95			
10%	52.52			
5% 34.39				
1%	26.52			
0% Min	20.54			

Extreme Observations					
	Lowest		F	lighest	
Value	Freq	Obs	Value	Freq	Obs
20.54	4062	1865	1086.02	10294	1909
20.80	8060	1835	1095.16	5455	1875
21.28	7665	1858	1211.03	7410	1873
26.52	16121	1868	1309.62	11846	1705
27.30	7416	1627	1410.32	2957	1687

Freq: WeightD

Geography=Saskatchewan

Moments					
N	271874	Sum Weights	271874		
Mean	396.778273	Sum Observations	107873696		
Std Deviation	317.363383	Variance	100719.517		
Skewness	1.14716933	Kurtosis	0.63577485		
Uncorrected SS	7.01849E10	Corrected SS	2.73829E10		
Coeff Variation	79.9850709	Std Error Mean	0.6086575		

Basic Statistical Measures					
Location Variability					
Mean	396.7783	Std Deviation	317.36338		
Median	286.5800	Variance	100720		
Mode	45.3600	Range	1399		
		Interquartile Range	408.02000		

Tests for Location: Mu0=0						
Test	Statistic p Value					
Student's t	t 651.8909		Pr > t	<.0001		
Sign	M 135937		Pr >= M	<.0001		
Signed Rank	s	1.848E10	Pr >= S	<.0001		

Freq: WeightD

Geography=Saskatchewan

Tests for Normality					
Test	Statistic p Value				
Kolmogorov-Smirnov	D 0.142037		Pr > D	<0.0100	
Cramer-von Mises	W-Sq 1757.807		Pr > W-Sq	<0.0050	
Anderson-Darling	A-Sq	10351.25	Pr > A-Sq	<0.0050	

Quantiles (Definition 5)				
Level	Quantile			
100% Max	1420.62			
99%	1313.72			
95%	1067.11			
90%	867.33			
75% Q3	558.48			
50% Median	286.58			
25% Q1	150.46			
10%	90.33			
5%	63.87			
1%	35.97			
0% Min	21.84			

Freq: WeightD

Geography=Saskatchewan

Extreme Observations					
ı	Lowest		н	ighest	
Value	Freq	Obs	Value	Freq	Obs
21.84	1372	2064	1266.72	1518	2153
28.46	523	2027	1313.62	888	1965
35.97	838	2140	1313.72	1691	2058
40.67	629	2083	1346.69	1060	1940
41.27	497	1944	1420.62	1377	1961

The UNIVARIATE Procedure Variable: FD806 (Expense Non-alcoholic beverages)

Freq: WeightD

Geography=Territorial capitals

Moments					
N	11711	Sum Weights	11711		
Mean	447.32884	Sum Observations	5238668.05		
Std Deviation	366.931505	Variance	134638.729		
Skewness	1.0227851	Kurtosis	0.2342045		
Uncorrected SS	3920026824	20026824 Corrected SS			
Coeff Variation	82.0272407	Std Error Mean	3.39068942		

Freq: WeightD

Geography=Territorial capitals

Basic Statistical Measures					
Location Variability					
Mean	447.3288	Std Deviation	366.93150		
Median	317.3100	Variance	134639		
Mode	96.9400	Range	1537		
		Interquartile Range	486.89000		

Tests for Location: Mu0=0						
Test	Statistic p Value					
Student's t	t	131.9286	Pr > t	<.0001		
Sign	М	5855.5	Pr >= M	<.0001		
Signed Rank	S	34289808	Pr >= S	<.0001		

Tests for Normality							
Test	Sta	atistic	p Value				
Kolmogorov-Smirnov	D	0.155138	Pr > D	<0.0100			
Cramer-von Mises	W-Sq	69.88919	Pr > W-Sq	<0.0050			
Anderson-Darling	A-Sq	422.3186	Pr > A-Sq	<0.0050			

Quantiles (Definition 5)				
Level	Quantile			
100% Max	1573.00			
99%	1494.74			
95%	1191.01			
90%	1029.41			

Freq: WeightD

Geography=Territorial capitals

Quantiles (Definition 5)					
Level	Quantile				
75% Q3	633.10				
50% Median	317.31				
25% Q1	146.21				
10%	70.20				
5%	62.66				
1%	43.96				
0% Min	35.62				

Extreme Observations							
ı	Lowest		н	ighest			
Value	Freq	Obs	Value	Freq	Obs		
35.62	18	2183	1447.85	46	2293		
43.96	115	2265	1462.25	45	2171		
47.35	157	2299	1494.74	58	2215		
50.18	58	2278	1522.61	59	2169		
51.48	33	2302	1573.00	43	2179		

Null hypothesis: equal variances

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

b.lf 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: FD806 Expense Non-alcoholic beverages

Frequency: WeightD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	17551256049	2507322292.7	28147.8	<.0001
Error	8.13E6	724095873793	89077.085984		
Corrected Total	8.13E6	741647129841			

R-Square	Coeff Var	Root MSE	FD806 Mean
0.023665	91.30931	298.4578	326.8646

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Prov	7	17551256049	2507322293	28147.8	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Prov	7	17551256049	2507322293	28147.8	<.0001

The GLM Procedure

Levene's Test for Homogeneity of FD806 Variance ANOVA of Absolute Deviations from Group Means							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Prov	7	7.3095E9	1.0442E9	27277.3	<.0001		
Error	8.13E6	3.112E11	38281.4				

Welch's ANOVA for FD806				
Source	DF	F Value	Pr > F	
Prov	7.0000	25845.7	<.0001	
Error	220418			

The GLM Procedure

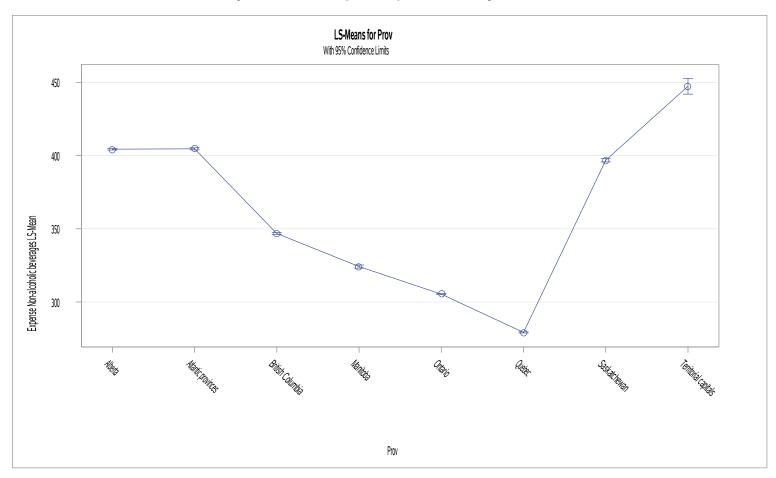
		FD806		
Level of Prov	N	Mean	Std Dev	
Alberta	958697	404.357996	353.006082	
Atlantic provinces	629768	404.594424	343.356582	
British Columbia	969907	346.767865	298.615022	
Manitoba	299857	324.209684	289.205287	
Ontario	2882375	305.372478	300.076977	
Quebec	2104687	279.245849	248.826064	
Saskatchewan	271874	396.778273	317.363383	
Territorial capitals	11711	447.328840	366.931505	

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

Prov	FD806 LSMEAN	LSMEAN Number
Alberta	404.357996	1
Atlantic provinces	404.594424	2
British Columbia	346.767865	3
Manitoba	324.209684	4
Ontario	305.372478	5
Quebec	279.245849	6
Saskatchewan	396.778273	7
Territorial capitals	447.328840	8

Least Squares Means for effect Prov Pr > t for H0: LSMean(i)=LSMean(j)									
	Dependent Variable: FD806								
i/j	1	2	3	4	5	6	7	8	
1		0.9997	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
2	0.9997		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
3	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	
4	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	
6	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	
7	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	
8	<.0001	<.0001	<.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

