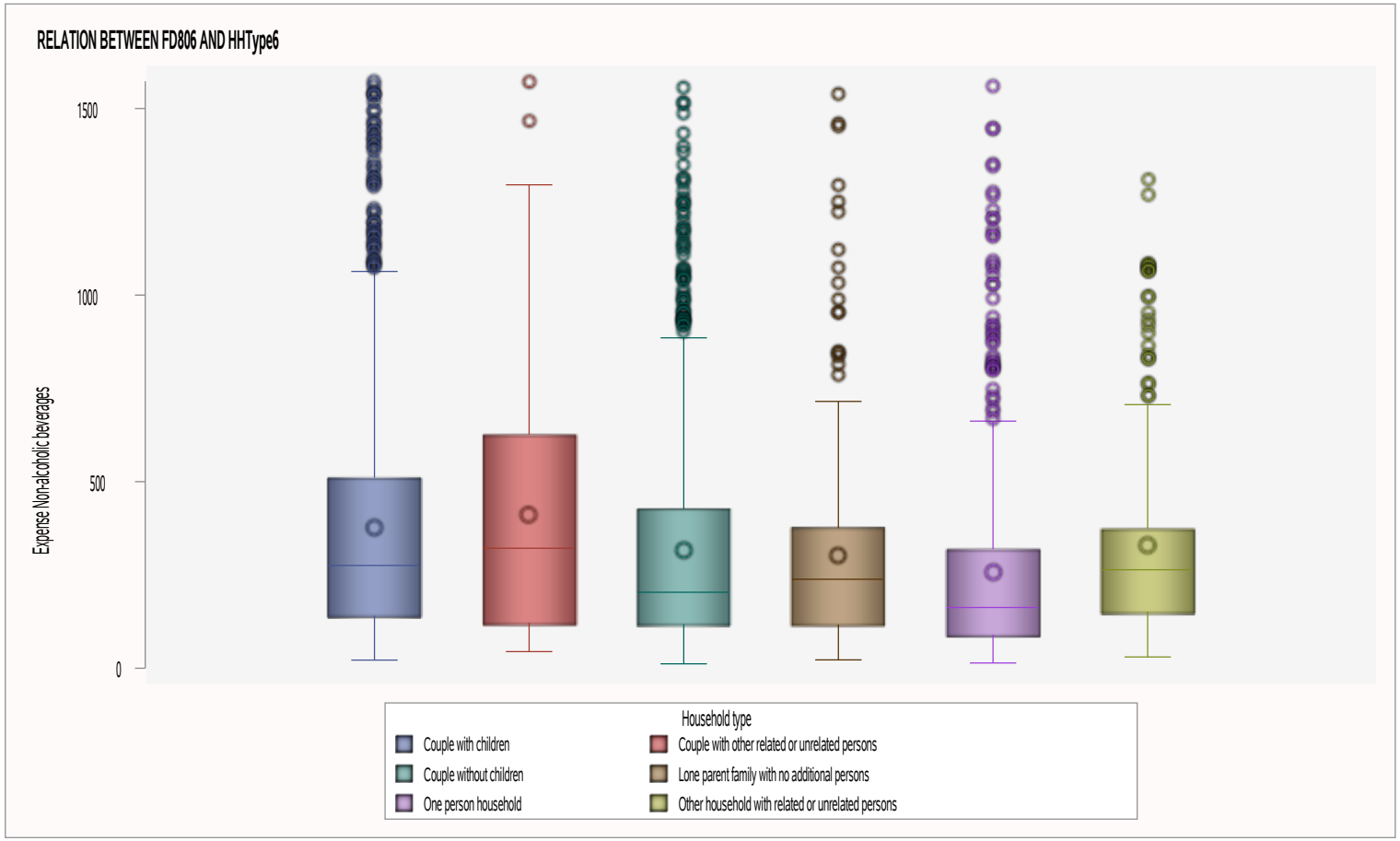
BIVARIATE ANALYSIS OF HHType6 AND FD806 FOR ANA.MODEL1
RELATION BETWEEN FD806 AND HHType6

The MEANS Procedure

Analysis Variable : FD806 Expense Non-alcoholic beverages														
Household type	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
Couple with children	2457237	2457237	0	21.84	135.72	275.47	377.09	510.64	1573.00	374.92	85.91	376.69	377.50	1.46
Couple with other related or unrelated persons	557180	557180	0	44.55	116.87	321.54	409.09	626.08	1569.37	509.21	81.56	408.21	409.96	1.28
Couple without children	2206478	2206478	0	11.83	112.84	203.58	317.62	426.14	1556.41	313.30	91.54	317.23	318.00	1.66
Lone parent family with no additional persons	513225	513225	0	22.36	112.28	238.48	302.25	375.52	1536.76	263.24	88.73	301.51	302.98	1.81
One person household	1964999	1964999	0	14.08	85.45	162.76	257.47	319.02	1557.89	233.57	106.69	257.08	257.85	2.11
Other household with related or unrelated persons	429757	429757	0	30.16	146.64	263.90	327.27	373.84	1308.58	227.20	80.88	326.48	328.06	1.69



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: FD806 (Expense Non-alcoholic beverages)

Freq: WeightD

Household type=Couple with children

Moments			
N	2457237	**Sum Weights**	2457237
Mean	377.092158	**Sum Observations**	926604804
Std Deviation	323.959134	**Variance**	104949.521
Skewness	1.45931195	**Kurtosis**	1.83638136
Uncorrected SS	6.07301E11	**Corrected SS**	2.57886E11
Coeff Variation	85.9098041	**Std Error Mean**	0.20666489

Basic Statistical Measures			
Location		**Variability**	
Mean	377.0922	**Std Deviation**	323.95913
Median	275.4700	**Variance**	104950
Mode	293.0500	**Range**	1551
		Interquartile Range	374.92000

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

Freq: WeightD

Household type=Couple with children

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1824.655	**Pr > |t|**	<.0001
Sign	**M**	1228619	**Pr >= |M|**	<.0001
Signed Rank	**S**	1.51E12	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.159684	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	18814.1	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	110124.9	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1573.00
99%	1469.40
95%	1145.04
90%	785.84
75% Q3	510.64
50% Median	275.47
25% Q1	135.72
10%	70.72
5%	51.74
1%	28.08
0% Min	21.84

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

Freq: WeightD

Household type=Couple with children

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
21.84	3250	653	1540.64	1345	483
23.40	132	22	1540.76	776	396
24.96	186	161	1546.22	669	680
26.00	1093	172	1561.82	491	456
26.26	1144	636	1573.00	43	306

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

Freq: WeightD

Household type=Couple with other related or unrelated persons

Moments			
N	557180	Sum Weights	557180
Mean	409.085512	Sum Observations	227934265
Std Deviation	333.63924	Variance	111315.142
Skewness	1.28183958	Kurtosis	1.74708531
Uncorrected SS	1.55267E11	Corrected SS	6.20225E10
Coeff Variation	81.5573347	Std Error Mean	0.44697102

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

Freq: WeightD

Household type=Couple with other related or unrelated persons

Basic Statistical Measures			
Location		**Variability**	
Mean	409.0855	**Std Deviation**	333.63924
Median	321.5400	**Variance**	111315
Mode	657.3000	**Range**	1525
		Interquartile Range	509.21000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	915.2395	**Pr > |t|**	<.0001
Sign	**M**	278590	**Pr >= |M|**	<.0001
Signed Rank	**S**	7.761E10	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.137284	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	2203.392	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	16222.43	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1569.37
99%	1464.24
95%	1022.41
90%	775.63

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

Freq: WeightD

Household type=Couple with other related or unrelated persons

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	626.08
50% Median	321.54
25% Q1	116.87
10%	59.10
5%	52.00
1%	44.55
0% Min	44.55

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
44.55	12589	725	1145.38	1539	767
52.00	21675	739	1191.01	60	709
52.26	36	745	1295.58	1523	759
55.44	16745	769	1464.24	12702	747
59.10	20541	765	1569.37	4907	734

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

Freq: WeightD

Household type=Couple without children

Moments			
N	2206478	**Sum Weights**	2206478
Mean	317.617514	**Sum Observations**	700816056
Std Deviation	290.741589	**Variance**	84530.6718
Skewness	1.65772878	**Kurtosis**	2.70480882
Uncorrected SS	4.09106E11	**Corrected SS**	1.86515E11
Coeff Variation	91.5382738	**Std Error Mean**	0.19572998

Basic Statistical Measures			
Location		**Variability**	
Mean	317.6175	**Std Deviation**	290.74159
Median	203.5800	**Variance**	84531
Mode	190.3400	**Range**	1545
		Interquartile Range	313.30000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1622.733	**Pr > |t|**	<.0001
Sign	**M**	1103239	**Pr >= |M|**	<.0001
Signed Rank	**S**	1.217E12	**Pr >= |S|**	<.0001

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

Freq: WeightD

Household type=Couple without children

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.160741	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	21068.53	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	120508.2	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1556.41
99%	1309.62
95%	934.96
90%	751.92
75% Q3	426.14
50% Median	203.58
25% Q1	112.84
10%	59.54
5%	49.14
1%	29.64
0% Min	11.83

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

Freq: WeightD

Household type=Couple without children

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
11.83	3940	1110	1430.56	1379	1108
21.58	347	952	1485.65	5912	1186
21.84	1372	1370	1511.64	5649	1341
25.22	463	1187	1513.44	2134	1094
25.48	987	983	1556.41	1360	889

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

Freq: WeightD

Household type=Lone parent family with no additional persons

Moments			
N	513225	**Sum Weights**	513225
Mean	302.248729	**Sum Observations**	155121604
Std Deviation	268.197191	**Variance**	71929.7332
Skewness	1.80743776	**Kurtosis**	3.45271654
Uncorrected SS	8.38014E10	**Corrected SS**	3.69161E10
Coeff Variation	88.733935	**Std Error Mean**	0.37436938

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

Freq: WeightD

Household type=Lone parent family with no additional persons

Basic Statistical Measures			
Location		**Variability**	
Mean	302.2487	**Std Deviation**	268.19719
Median	238.4800	**Variance**	71930
Mode	54.0800	**Range**	1514
		Interquartile Range	263.24000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	807.3543	**Pr > |t|**	<.0001
Sign	**M**	256612.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	6.585E10	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.186172	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	4891.038	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	28248.84	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1536.76
99%	1217.58
95%	849.05
90%	694.20

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

Freq: WeightD

Household type=Lone parent family with no additional persons

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	375.52
50% Median	238.48
25% Q1	112.28
10%	54.08
5%	44.86
1%	29.90
0% Min	22.36

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
22.36	381	1693	1250.86	990	1666
24.96	1553	1652	1292.20	1554	1690
29.90	5372	1658	1449.41	505	1684
35.92	14896	1655	1456.00	442	1687
40.67	629	1558	1536.76	836	1713

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

Freq: WeightD

Household type=One person household

Moments			
N	1964999	**Sum Weights**	1964999
Mean	257.465002	**Sum Observations**	505918471
Std Deviation	274.692386	**Variance**	75455.907
Skewness	2.11090445	**Kurtosis**	4.28838971
Uncorrected SS	2.78527E11	**Corrected SS**	1.48271E11
Coeff Variation	106.691156	**Std Error Mean**	0.19595911

Basic Statistical Measures			
Location		**Variability**	
Mean	257.4650	**Std Deviation**	274.69239
Median	162.7600	**Variance**	75456
Mode	386.9000	**Range**	1544
		Interquartile Range	233.57000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1313.871	**Pr > |t|**	<.0001
Sign	**M**	982499.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	9.653E11	**Pr >= |S|**	<.0001

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

Freq: WeightD

Household type=One person household

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.199652	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	29015.43	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	165709.1	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1557.89
99%	1230.11
95%	915.33
90%	599.04
75% Q3	319.02
50% Median	162.76
25% Q1	85.45
10%	50.70
5%	34.39
1%	20.54
0% Min	14.08

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

Freq: WeightD

Household type=One person household

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
14.08	15313	2018	1346.28	5467	2127
15.08	1520	2154	1354.61	1683	2058
20.54	4062	1917	1444.43	1311	1818
20.80	8060	1722	1447.85	46	1943
21.28	7665	1892	1557.89	759	1975

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

Freq: WeightD

Household type=Other household with related or unrelated persons

Moments			
N	429757	**Sum Weights**	429757
Mean	327.270551	**Sum Observations**	140646810
Std Deviation	264.683575	**Variance**	70057.3949
Skewness	1.6888428	**Kurtosis**	2.62426534
Uncorrected SS	7.61371E10	**Corrected SS**	3.01076E10
Coeff Variation	80.8760747	**Std Error Mean**	0.40375277

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

Freq: WeightD

Household type=Other household with related or unrelated persons

Basic Statistical Measures			
Location		**Variability**	
Mean	327.2706	**Std Deviation**	264.68358
Median	263.9000	**Variance**	70057
Mode	141.5000	**Range**	1278
		Interquartile Range	227.20000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	810.5717	**Pr > |t|**	<.0001
Sign	**M**	214878.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	4.617E10	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.189182	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	4031.699	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	23929.75	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1308.58
99%	1266.72
95%	1062.90
90%	706.42

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

Freq: WeightD

Household type=Other household with related or unrelated persons

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	373.84
50% Median	263.90
25% Q1	146.64
10%	77.48
5%	53.30
1%	30.16
0% Min	30.16

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
30.16	12248	2308	1077.70	2800	2216
35.97	838	2230	1083.16	644	2241
45.60	6607	2322	1086.02	10294	2260
52.00	1422	2313	1266.72	1518	2317
53.30	8041	2224	1308.58	3082	2325

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**
HHType6	6	Couple with children Couple with other related or unrelated persons Couple without children Lone parent family with no additional persons One person household Other household with related or unrelated persons

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	5	19929588843	3985917768.6	44894.3	<.0001
Error	8.13E6	721717540999	88784.485543		
Corrected Total	8.13E6	741647129841			

R-Square	Coeff Var	Root MSE	FD806 Mean
0.026872	91.15922	297.9673	326.8646

Source	DF	Type I SS	Mean Square	F Value	Pr > F
HHType6	5	19929588843	3985917769	44894.3	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
HHType6	5	19929588843	3985917769	44894.3	<.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**
HHType6	5	5.0149E9	1.003E9	25667.0	<.0001
Error	8.13E6	3.177E11	39076.8		

Welch's ANOVA for FD806			
Source	**DF**	**F Value**	**Pr > F**
HHType6	5.0000	44119.9	<.0001
Error	1933052		

The GLM Procedure

Level of HHType6	**N**	**FD806**	
		Mean	**Std Dev**
Couple with children	2457237	377.092158	323.959134
Couple with other related or unrelated persons	557180	409.085512	333.639240
Couple without children	2206478	317.617514	290.741589
Lone parent family with no additional persons	513225	302.248729	268.197191
One person household	1964999	257.465002	274.692386
Other household with related or unrelated persons	429757	327.270551	264.683575

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

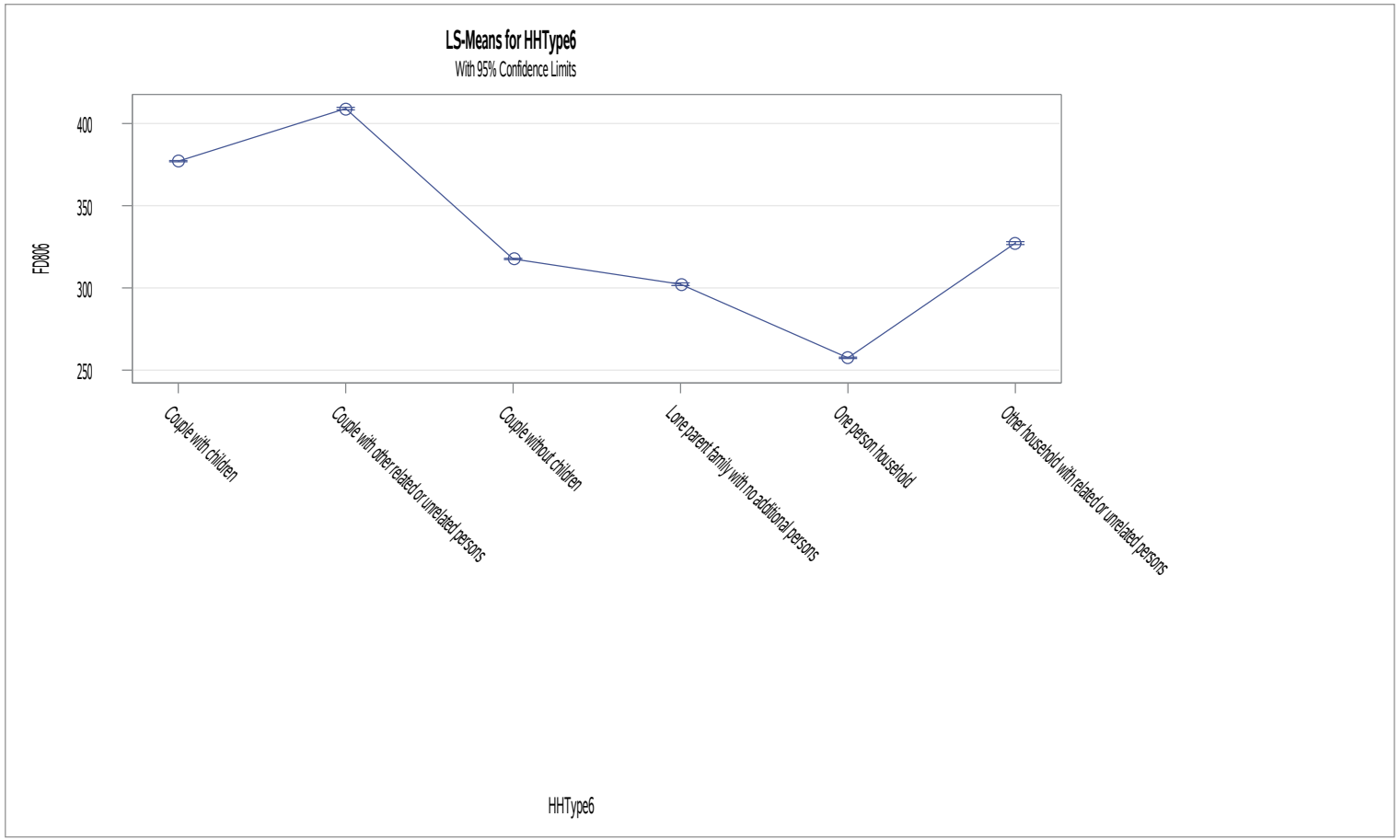
HHType6	FD806 LSMEAN	LSMEAN Number
Couple with children	377.092158	1
Couple with other related or unrelated persons	409.085512	2
Couple without children	317.617514	3
Lone parent family with no additional persons	302.248729	4
One person household	257.465002	5
Other household with related or unrelated persons	327.270551	6

Least Squares Means for effect HHType6
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: 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