Contents

Results

Analysis Visualizations and Graphs

Analysis figures and graphs derived from SPSS are provided for each analysis method. Some of this analysis was done in SAS but this paper is limited to SPSS output.

Table 8: Anova results including Robust test of equality of means and Anova effect sizes. The Welch F-statistic and Adjusted omega-squared are used in this analysis because homogeneity of variance violation occurs.

ANOVA

Expenses

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	361549380.965	2	180774690.482	9.747	<.001
Within Groups	19102764749.627	1030	18546373.543		
Total	19464314130.591	1032			

Robust Tests of Equality of Means

Expenses

	Statistic[a]	df1	df2	Sig.
Welch	11.655	2	436.298	<.001
Brown-Forsythe	10.756	2	749.506	<.001

a. Asymptotically F distributed.

ANOVA Effect Sizes[a]

			95% Confidence Interval	
		Point Estimate	Lower	Upper
Expenses	Eta-squared	.019	.005	.037
	Epsilon-squared	.017	.003	.035
	Omega-squared Fixed-effect	.017	.003	.035
	Omega-squared Random-effect	.008	.002	.018

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

$$Adjusted\ \omega^2 = \frac{(df_b)(F-1)}{(df_b)(F-1)+N_{total}} = \frac{2(11.655-1)}{2(11.655-1)+1033} = 0.0202$$

Table 9: Games-Howell Post-hoc analysis used because homoscedasticity is violated.

Games-Howell Post-hoc Multiple Comparisons

Dependent Variable: Expenses

Games-Howell

					95% Confidence Interval	
(I) Famsize_num	(J) Famsize_num	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
small family size	normal family size	-306.10085	292.96456	.549	-993.9455	381.7438
	large family size	-1773.27379*	382.52129	<.001	-2673.7469	-872.8007
normal family size	small family size	306.10085	292.96456	.549	-381.7438	993.9455
	large family size	-1467.17294*	352.08662	<.001	-2297.1158	-637.2301
large family size	small family size	1773.27379*	382.52129	<.001	872.8007	2673.7469
	normal family size	1467.17294*	352.08662	<.001	637.2301	2297.1158

*. The mean difference is significant at the 0.05 level.

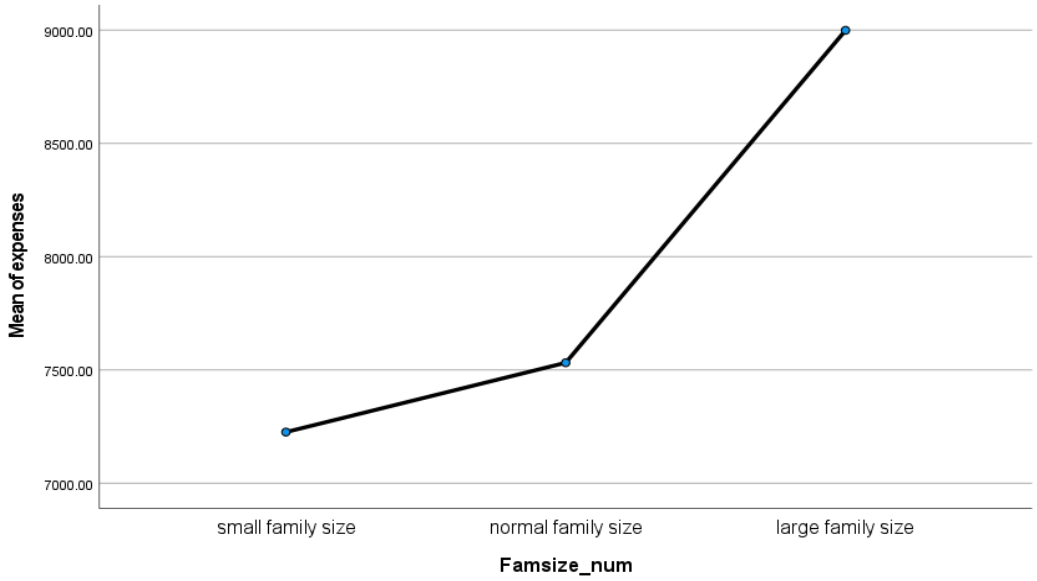

Figure 1: Mean Plots for 3-levels of family sized analyzed using Anova.

Kruskal-Wallis analysis visualizations and Graphs

Ranks			
	Famsize_num	**N**	**Mean Rank**
Health expenses	small family size	458	478.71
	normal family size	424	521.47
	large family size	151	620.60
	Total	1033	

Test Statistics[a,b]	
	Health expenses
Kruskal-Wallis H	25.849
df	2
Asymp. Sig.	<.001
a. Kruskal Wallis Test	
b. Grouping Variable: Famsize_num	

Pairwise Comparisons of Famsize_num					
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
small family size-normal family size	-42.760	20.107	-2.127	.033	.100
small family size-large family size	-141.895	27.997	-5.068	<.001	.000
normal family size-large family size	-99.136	28.274	-3.506	<.001	.001
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is .017.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

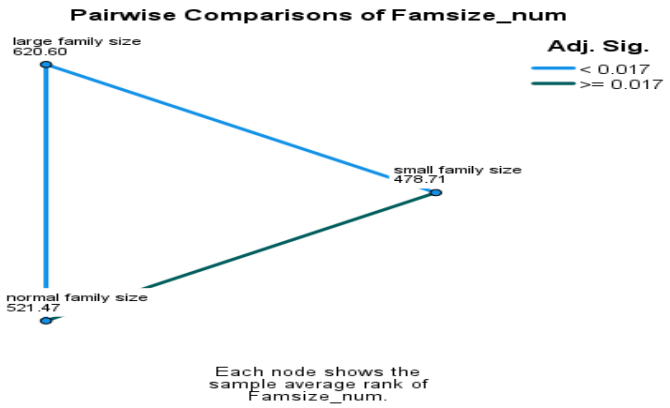

Figure 2: Visual of Dunn's pairwise comparison for family size levels. A significant difference between LFS vs SFS and LFS vs NFS, indicate by the blue lines.

Chi-square analysis visualizations and Graphs

Goodness of fit test for patient proportions distribution

Table 10: Chi-square goodness of fit test. (Chi-sq stat (E, U) = 175.37,361.756, p<.001, p<.0001); REJECT null

Groups based on family size and health expense-Equal(E)			
Categories	**Observed N**	**Expected N**	**Residual**
SFS_AA	210	172.2	37.8
NFS_AA	190	172.2	17.8
LFS_AA	85	172.2	-87.2
SFS_BA	248	172.2	75.8
NFS_BA	234	172.2	61.8
LFS_BA	66	172.2	-106.2
Total	1033		
Groups based on family size and health expense-Unequal(U) (SAS and Excel)			
Categories	**Observed N**	**Expected N**	**Residual**
SFS_AA	210	154.95 (15%)	55.05
NFS_AA	190	206.6 (20%)	-16.60
LFS_AA	85	309.9 (30%)	-224.90
SFS_BA	248	206.6 (20%)	41.40
NFS_BA	234	103.3 (10%)	130.70
LFS_BA	66	51.65 (5%)	14.35
Total	1033		

Chi-square Statistics		
	Family size+Spending amount group (Equal)	**Family size+Spending amount group (Unequal)**
Chi-Square	175.370[a]	361.756[b]
df	5	5
Asymp. Sig.	<.001	<.0001

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 172.2.

b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 51.65. **Obtained from SAS and merged with the SPSS table.**

Long version of project has snippet of SAS output. SAS code will be provided.

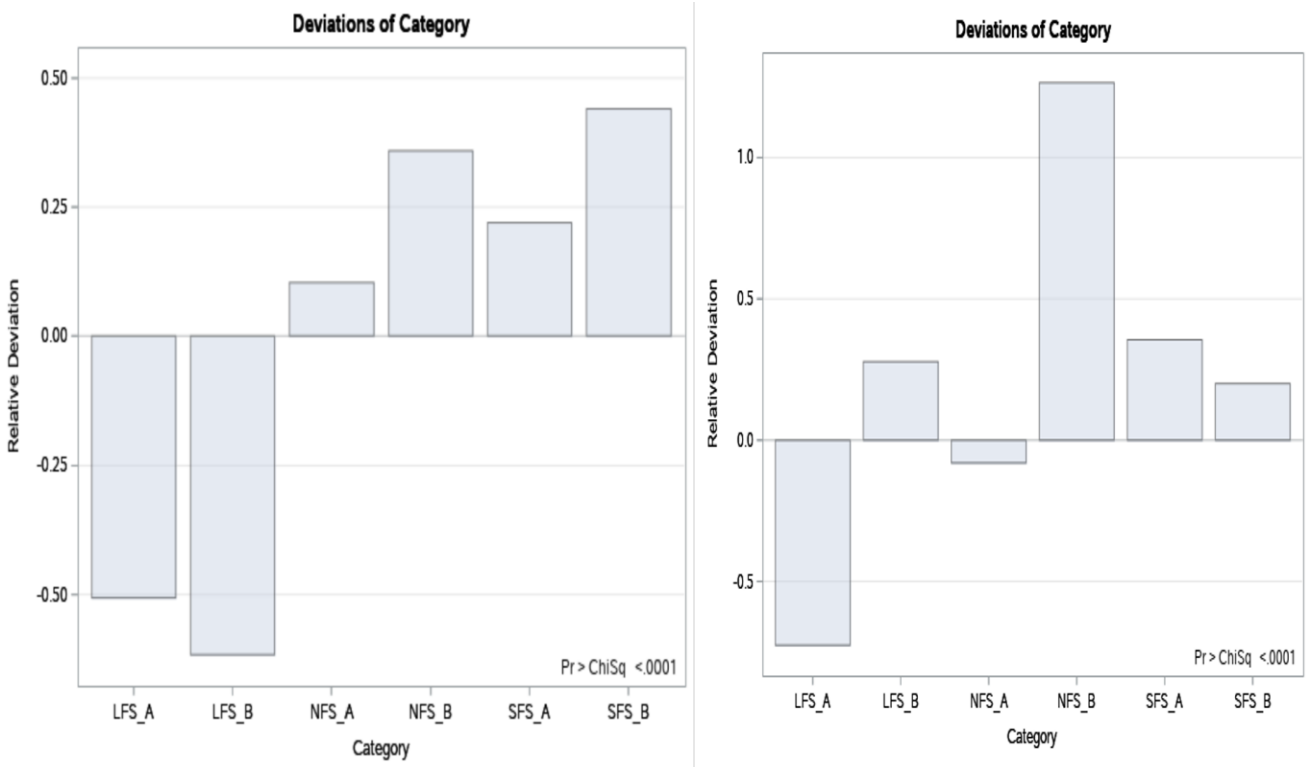

Figure 3: Chi-square goodness of fit test (left = equal proportions hypothesis, right = unequal proportions hypothesis). Unequal proportions hypothesis follows the rule LFSA>NFSA>SFSA & LFSB<NFSB<SFSB. The SAS code provided for this analysis has to be followed exactly as provided to avoid errors.

Goodness of fit test insuree proportions by expense type per level

Test whether within each family size level, health expenses (expense type) will generate unequal proportions of insurees. Does the distribution for the proportions of insurees per health expense type follow an equal proportions hypothesis? Is there anything unique about spending above vs below average that will drive an unequal proportion of patients between the expense type groups, for each level?

SFS level goodness of fit

Table 11: Chi-sq goodness of fit (gof) test for insuree proportions in SFS by expense type. The observed distribution DOES NOT differ significantly from the equal proportions' distribution (Chi-sq = 3.153, p = 0.076); ACCEPT the null (alpha = 0.017)

Expense type code (SFS)			
SFS	**Observed N**	**Expected N**	**Residual**
BA	248	229.0	19.0
AA	210	229.0	-19.0
Total	458		

Chi-sq Test Statistics (SFS)	
	Expense type code (SFS)
Chi-Square	3.153[a]
df	1
Asymp. Sig.	.076
a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 229.0.	

NFS level goodness of fit

Table 12: Chi-sq goodness of fit (gof) test for insuree proportions in NFS by expense type. The observed distribution DOES NOT differ significantly from the equal proportions' distribution (Chi-sq = 4.566, p = 0.033); ACCEPT the null

Expense type code (NFS)			
NFS	**Observed N**	**Expected N**	**Residual**
BA	234	212.0	22.0
AA	190	212.0	-22.0
Total	424		

Chi-sq Test Statistics (NFS)	
	Expense type code (NFS)
Chi-Square	4.566[a]
df	1
Asymp. Sig.	.033
a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 212.0.	

LFS level goodness of fit

Table 13: Chi-sq goodness of fit (gof) test for insuree proportions in LFS by expense type. The observed distribution DOES NOT differ significantly from the equal proportions' distribution (Chi-sq = 2.391, p = .122); ACCEPT the null

Expense type code (LFS)			
LFS	**Observed N**	**Expected N**	**Residual**
BA	66	75.5	-9.5
AA	85	75.5	9.5
Total	151		

Chi-sq Test Statistics (LFS)	
	Expense type code (LFS)
Chi-Square	2.391[a]
df	1
Asymp. Sig.	.122
a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 75.5.	

Chi-square test of association between family size and expense type

Table 1: Summary of Chi-square test of association. With respect to the Bonferroni corrected alpha-level, statistically significant association is observed between family size and expense type, in the NFS vs LFS pairwise comparison (alpha = 0.017). Refer to details of these comparisons in the appendix.

Pairwise	Family sizes	Family sizes	Pearson χ^2	p-value	Phi-stat	p-value	Decision
1-SFS v NFS	SFS	NFS	.096	.757	.010	.757	Independent
2-SFS v LFS	SFS	LFS	4.956	.026	-.090	.026	Independent
3- NFS v LFS	NFS	LFS	5.881	.015	-.101	.015	Association

Chi-square Bar Charts

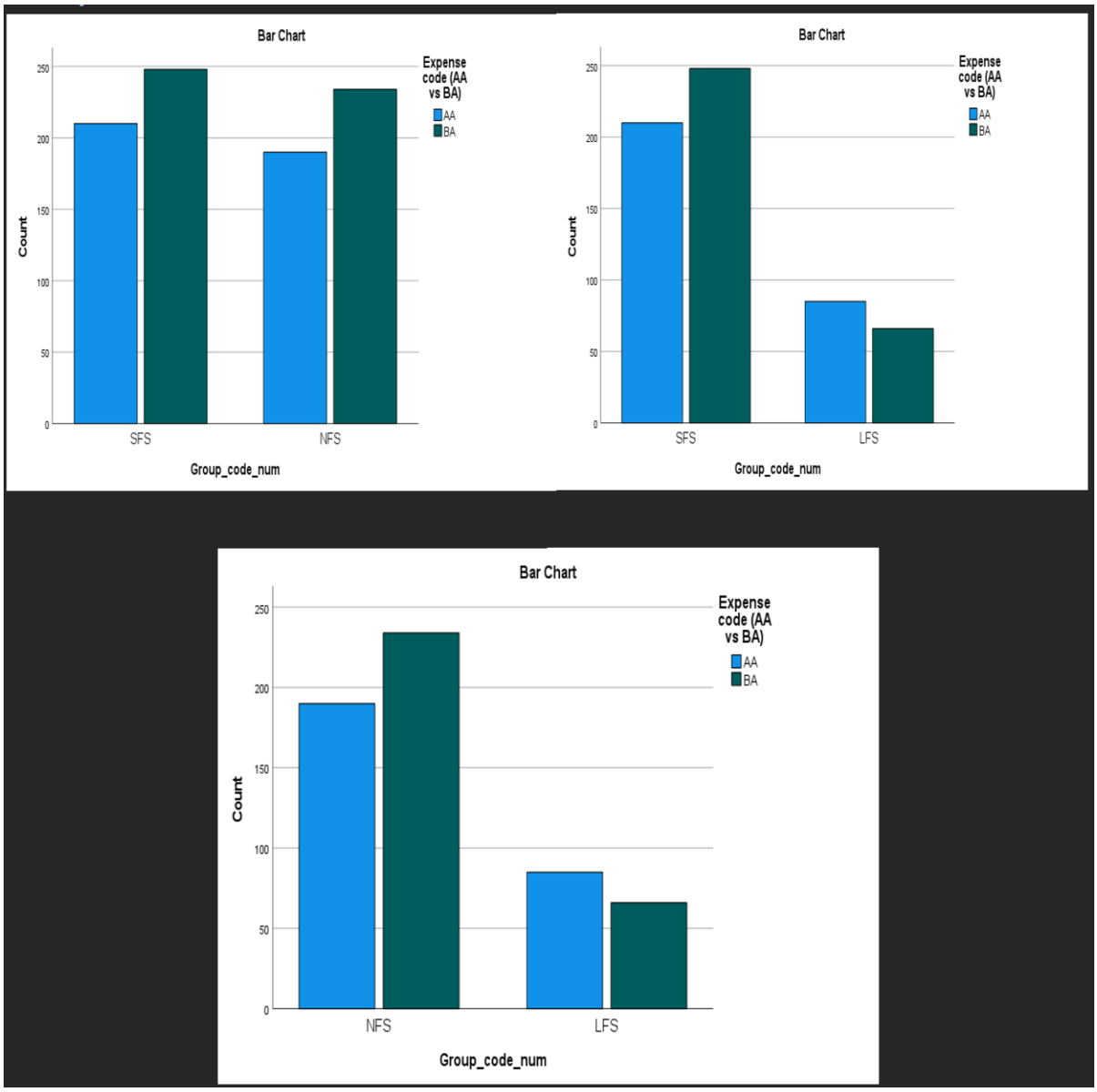

Figure 4: Bar charts of Chi-square pairwise tests of association (left = SFS vs NFS, right = SFS vs LFS, bottom = NFS vs LFS)

Chi-square test of association Detailed

SFS vs NFS pairwise Chi-square association test

Pairwise family size levels * Expense type (AA vs BA) Crosstabulation					
			Dichotomous Expense type (AA vs BA)		Total
			AA	BA	
Family size	SFS	Count	210	248	458
		Expected Count	207.7	250.3	458.0
		% within Group_code_num	45.9%	54.1%	100.0%
		% within Expense code (AA vs BA)	52.5%	51.5%	51.9%
		% of Total	23.8%	28.1%	51.9%
	NFS	Count	190	234	424
		Expected Count	192.3	231.7	424.0
		% within Group_code_num	44.8%	55.2%	100.0%
		% within Expense code (AA vs BA)	47.5%	48.5%	48.1%
		% of Total	21.5%	26.5%	48.1%
Total		Count	400	482	882
		Expected Count	400.0	482.0	882.0
		% within Group_code_num	45.4%	54.6%	100.0%
		% within Expense code (AA vs BA)	100.0%	100.0%	100.0%
		% of Total	45.4%	54.6%	100.0%

Chi-Square Association Tests					
Statistical test	**Value**	**df**	**Asymptotic Significance (2-sided)**	**Exact Sig. (2-sided)**	**Exact Sig. (1-sided)**
Pearson Chi-Square	.096[a]	1	.757		
Continuity Correction[b]	.059	1	.809		
Likelihood Ratio	.096	1	.757		
Fisher's Exact Test				.787	.404
N of Valid Cases	882				
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 192.29.					
b. Computed only for a 2x2 table					

Correlation Measures[c]			
		Value	**Approximate Significance**
Nominal by Nominal	Phi	.010	.757
	Cramer's V	.010	.757
N of Valid Cases		882	

SFS vs LFS pairwise Chi-square association test

Pairwise family size levels * Expense type (AA vs BA) Crosstabulation					
			Expense code (AA vs BA)		Total
			AA	BA	
Family size	SFS	Count	210	248	458
		Expected Count	221.9	236.1	458.0
		% within Group_code_num	45.9%	54.1%	100.0%
		% within Expense code (AA vs BA)	71.2%	79.0%	75.2%
		% of Total	34.5%	40.7%	75.2%
	LFS	Count	85	66	151
		Expected Count	73.1	77.9	151.0
		% within Group_code_num	56.3%	43.7%	100.0%
		% within Expense code (AA vs BA)	28.8%	21.0%	24.8%
		% of Total	14.0%	10.8%	24.8%
Total		Count	295	314	609
		Expected Count	295.0	314.0	609.0
		% within Group_code_num	48.4%	51.6%	100.0%
		% within Expense code (AA vs BA)	100.0%	100.0%	100.0%
		% of Total	48.4%	51.6%	100.0%

Chi-Square Association Tests					
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.956[a]	1	.026		
Continuity Correction[b]	4.546	1	.033		
Likelihood Ratio	4.961	1	.026		
Fisher's Exact Test				.031	.016
N of Valid Cases	609				
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 73.14.					
b. Computed only for a 2x2 table					

Correlation Measures[c]			
		Value	**Approximate Significance**
Nominal by Nominal	Phi	-.090	.026
	Cramer's V	.090	.026
N of Valid Cases		609	

NFS vs LFS pairwise Chi-square association test

Pairwise family size levels * Expense type (AA vs BA) Crosstabulation					
			Expense code (AA vs BA)		Total
			AA	BA	
Group_code_num	NFS	Count	190	234	424
		Expected Count	202.8	221.2	424.0
		% within Group_code_num	44.8%	55.2%	100.0%
		% within Expense code (AA vs BA)	69.1%	78.0%	73.7%
		% of Total	33.0%	40.7%	73.7%
	LFS	Count	85	66	151
		Expected Count	72.2	78.8	151.0
		% within Group_code_num	56.3%	43.7%	100.0%
		% within Expense code (AA vs BA)	30.9%	22.0%	26.3%
		% of Total	14.8%	11.5%	26.3%
Total		Count	275	300	575
		Expected Count	275.0	300.0	575.0
		% within Group_code_num	47.8%	52.2%	100.0%
		% within Expense code (AA vs BA)	100.0%	100.0%	100.0%
		% of Total	47.8%	52.2%	100.0%

Chi-Square Association Tests					
	Value	**df**	**Asymptotic Significance (2-sided)**	**Exact Sig. (2-sided)**	**Exact Sig. (1-sided)**
Pearson Chi-Square	5.881[a]	1	.015		
Continuity Correction[b]	5.430	1	.020		
Likelihood Ratio	5.884	1	.015		
Fisher's Exact Test				.018	.010
N of Valid Cases	575				
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 72.22.					
b. Computed only for a 2x2 table					

Directional Measures						
			Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Nominal by Nominal	Uncertainty Coefficient	Symmetric	.008	.007	1.216	.015[c]
		Group_code_num Dependent	.009	.007	1.216	.015[c]
		Expense code (AA vs BA) Dependent	.007	.006	1.216	.015[c]
a. Not assuming the null hypothesis.						
b. Using the asymptotic standard error assuming the null hypothesis.						
c. Likelihood ratio chi-square probability.						

Correlation Measures			
		Value	**Approximate Significance**
Nominal by Nominal	Phi	-.101	.015
	Cramer's V	.101	.015
N of Valid Cases		575	