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

The MEANS Procedure

Analysis Variable : FD806 Expense Non-alcoholic beverages														
Reference person - 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	2688986	2688986	0	15.08	116.22	249.62	351.14	470.34	1546.22	354.12	91.09	350.76	351.53	1.52
High school diploma, high school equivalency certificate, or not stated	1668821	1668821	0	20.54	87.10	187.72	302.08	404.82	1569.37	317.72	98.92	301.62	302.53	1.82
Less than high school diploma or its equivalent	1043308	1043308	0	20.80	129.38	223.60	305.61	409.24	1534.26	279.86	86.78	305.10	306.12	1.56
Masked records (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
University certificate or diploma	2716050	2716050	0	11.83	116.48	230.62	325.70	402.82	1561.82	286.34	91.14	325.35	326.06	1.71

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

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

Moments			
N	2688986	**Sum Weights**	2688986
Mean	351.143647	**Sum Observations**	944220350
Std Deviation	319.853519	**Variance**	102306.274
Skewness	1.51944997	**Kurtosis**	1.93067102
Uncorrected SS	6.06657E11	**Corrected SS**	2.751E11
Coeff Variation	91.0890805	**Std Error Mean**	0.1950549

Basic Statistical Measures			
Location		**Variability**	
Mean	351.1436	**Std Deviation**	319.85352
Median	249.6200	**Variance**	102306
Mode	72.0200	**Range**	1531
		Interquartile Range	354.12000

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

Freq: WeightD

Reference person - 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**	1800.23	**Pr > |t|**	<.0001
Sign	**M**	1344493	**Pr >= |M|**	<.0001
Signed Rank	**S**	1.808E12	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.162213	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	23208.65	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	137700.4	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1546.22
99%	1384.86
95%	1133.16
90%	777.14
75% Q3	470.34
50% Median	249.62
25% Q1	116.22
10%	66.74
5%	47.73
1%	27.30
0% Min	15.08

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

Freq: WeightD

Reference person - 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
15.08	1520	493	1511.64	5649	388
21.84	1372	371	1513.44	2134	422
21.84	3250	141	1536.76	836	688
24.82	209	458	1540.64	1345	570
25.70	1166	236	1546.22	669	186

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

Freq: WeightD

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

Moments			
N	1668821	Sum Weights	1668821
Mean	302.077883	Sum Observations	504113915
Std Deviation	298.813978	Variance	89289.7936
Skewness	1.81826406	Kurtosis	3.41411976
Uncorrected SS	3.0129E11	Corrected SS	1.49009E11
Coeff Variation	98.9195154	Std Error Mean	0.23131086

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

Freq: WeightD

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

Basic Statistical Measures			
Location		**Variability**	
Mean	302.0779	**Std Deviation**	298.81398
Median	187.7200	**Variance**	89290
Mode	107.6400	**Range**	1549
		Interquartile Range	317.72000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1305.939	**Pr > |t|**	<.0001
Sign	**M**	834410.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	6.962E11	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.17929	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	18475.93	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	104559.4	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1569.37
99%	1469.40
95%	917.02
90%	729.72

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

Freq: WeightD

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

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	404.82
50% Median	187.72
25% Q1	87.10
10%	52.00
5%	48.01
1%	31.13
0% Min	20.54

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
20.54	4062	1159	1340.82	1455	774
24.96	186	1130	1346.28	5467	885
25.22	463	1069	1346.69	1060	807
25.74	1311	981	1469.40	12196	796
28.00	6907	1140	1569.37	4907	1186

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

Freq: WeightD

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

Moments			
N	1043308	**Sum Weights**	1043308
Mean	305.607525	**Sum Observations**	318842776
Std Deviation	265.196401	**Variance**	70329.1309
Skewness	1.5588862	**Kurtosis**	2.18099046
Uncorrected SS	1.70816E11	**Corrected SS**	7.33749E10
Coeff Variation	86.7767901	**Std Error Mean**	0.25963388

Basic Statistical Measures			
Location		**Variability**	
Mean	305.6075	**Std Deviation**	265.19640
Median	223.6000	**Variance**	70329
Mode	487.1400	**Range**	1513
		Interquartile Range	279.86000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1177.071	**Pr > |t|**	<.0001
Sign	**M**	521654	**Pr >= |M|**	<.0001
Signed Rank	**S**	2.721E11	**Pr >= |S|**	<.0001

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

Freq: WeightD

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

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.164155	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	8900.523	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	52408.25	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1534.26
99%	1178.84
95%	928.19
90%	644.54
75% Q3	409.24
50% Median	223.60
25% Q1	129.38
10%	57.46
5%	39.00
1%	26.00
0% Min	20.80

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

Freq: WeightD

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

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
20.80	8060	1251	1178.84	5611	1225
21.58	840	1405	1199.64	1941	1241
22.36	381	1243	1222.26	2000	1226
25.48	987	1326	1362.61	680	1319
26.00	1093	1227	1534.26	1566	1216

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

Freq: WeightD

Reference person - Education level=Masked records (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	**Corrected SS**	1576619520
Coeff Variation	82.0272407	**Std Error Mean**	3.39068942

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

Freq: WeightD

Reference person - Education level=Masked records (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	**M**	5855.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	34289808	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**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

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

Freq: WeightD

Reference person - Education level=Masked records (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					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
35.62	18	1475	1447.85	46	1632
43.96	115	1541	1462.25	45	1560
47.35	157	1584	1494.74	58	1598
50.18	58	1613	1522.61	59	1470
51.48	33	1585	1573.00	43	1513

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

Freq: WeightD

Reference person - Education level=University certificate or diploma

Moments			
N	2716050	**Sum Weights**	2716050
Mean	325.703246	**Sum Observations**	884626301
Std Deviation	296.846008	**Variance**	88117.5527
Skewness	1.70910262	**Kurtosis**	2.87801966
Uncorrected SS	5.27457E11	**Corrected SS**	2.39332E11
Coeff Variation	91.1400215	**Std Error Mean**	0.18012016

Basic Statistical Measures			
Location		**Variability**	
Mean	325.7032	**Std Deviation**	296.84601
Median	230.6200	**Variance**	88118
Mode	657.3000	**Range**	1550
		Interquartile Range	286.34000

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1808.255	**Pr > |t|**	<.0001
Sign	**M**	1358025	**Pr >= |M|**	<.0001
Signed Rank	**S**	1.844E12	**Pr >= |S|**	<.0001

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

Freq: WeightD

Reference person - Education level=University certificate or diploma

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.165836	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	26160.55	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	149862.3	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	1561.82
99%	1430.56
95%	991.74
90%	753.81
75% Q3	402.82
50% Median	230.62
25% Q1	116.48
10%	59.10
5%	51.22
1%	21.58
0% Min	11.83

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

Freq: WeightD

Reference person - Education level=University certificate or diploma

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
11.83	3940	2139	1494.48	8411	1637
14.08	15313	2265	1540.76	776	1822
21.28	7665	2125	1556.41	1360	2240
21.58	347	2151	1557.89	759	2169
23.40	132	1655	1561.82	491	2023

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**
RP_Educ	5	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) 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: FD806 Expense Non-alcoholic beverages

Frequency: WeightD

The GLM Procedure

Dependent Variable: FD806 Expense Non-alcoholic beverages

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	3255416238.8	813854059.7	8959.63	<.0001
Error	8.13E6	738391713603	90835.703212		
Corrected Total	8.13E6	741647129841			

R-Square	Coeff Var	Root MSE	FD806 Mean
0.004389	92.20625	301.3896	326.8646

Source	DF	Type I SS	Mean Square	F Value	Pr > F
RP_Educ	4	3255416239	813854060	8959.63	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
RP_Educ	4	3255416239	813854060	8959.63	<.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
RP_Educ	4	1.9315E9	4.8288E8	12191.4	<.0001
Error	8.13E6	3.22E11	39608.3		

Welch's ANOVA for FD806			
Source	DF	F Value	Pr > F
RP_Educ	4.0000	8669.90	<.0001
Error	92346.8		

The GLM Procedure

Level of RP_Educ	N	FD806	
		Mean	Std Dev
Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution	2688986	351.143647	319.853519
High school diploma, high school equivalency certificate, or not stated	1668821	302.077883	298.813978
Less than high school diploma or its equivalent	1043308	305.607525	265.196401
Masked records (territorial capitals)	11711	447.328840	366.931505
University certificate or diploma	2716050	325.703246	296.846008

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

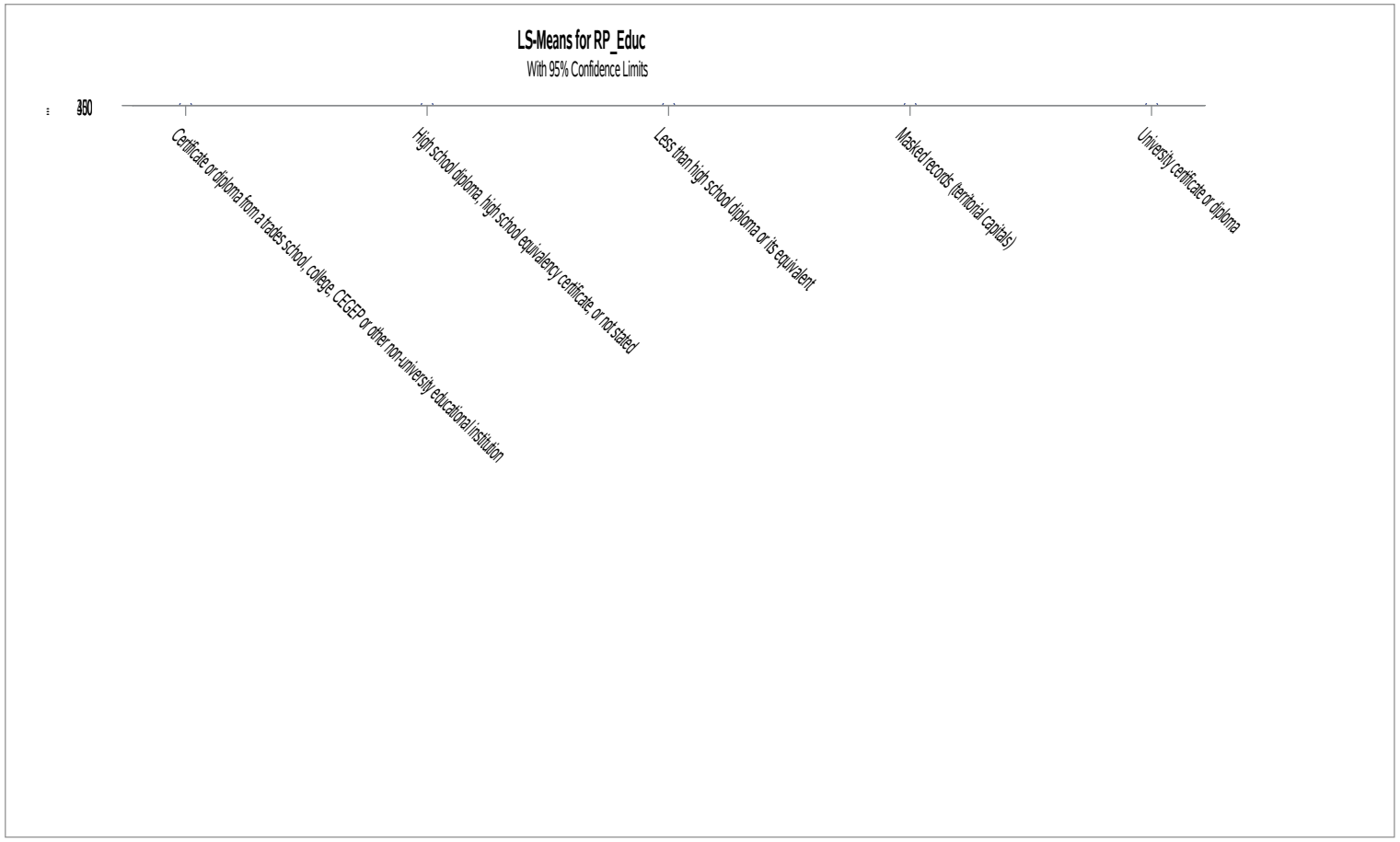
RP_Educ	FD806 LSMEAN	LSMEAN Number
Certificate or diploma from a trades school, college, CEGEP or other non-university educational institution	351.143647	1
High school diploma, high school equivalency certificate, or not stated	302.077883	2
Less than high school diploma or its equivalent	305.607525	3
Masked records (territorial capitals)	447.328840	4
University certificate or diploma	325.703246	5

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

Dependent Variable: FD806

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