Appendix 19

A19.1 t-estimator of μ

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

	A	B	C	D
1	**t-Estimate: Mean**			
2				
3				*Plastic bags*
4	Mean			1495
5	Standard Deviation			404
6	LCL			1451
7	UCL			1539

Estimate of the total number of plastic bags

LCL = 112 million(1451) = 162.512 billion

UCL = 112 million(1539) = 172.368 billion

A19.2 Spearman rank correlation coefficient test

H_0: $\rho_S = 0$

H_1: $\rho_S > 0$

$$z = r_S \sqrt{n-1}$$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*Satisfaction and Time*			
4	Spearman Rank Correlation			0.541
5	z Stat			8.984
6	P(Z<=z) one tail			0
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0
9	z Critical two tail			1.96

z = 8.98; p-value = 0. There is enough evidence to infer that those who do more research are more satisfied with their choice.

A19.3 Sign test

H_0: The location of the two populations are the same

H_1: The location of population 1 is to the right of the location of population2

$$z = \frac{x - .5n}{.5\sqrt{n}}$$

	A	B	C	D	E
1	**Sign Test**				
2					
3	Difference			*Cork rate - Metal rate*	
4					
5	Positive Differences			75	
6	Negative Differences			27	
7	Zero Differences			28	
8	z Stat			4.75	
9	P(Z<=z) one-tail			0	
10	z Critical one-tail			1.6449	
11	P(Z<=z) two-tail			0	
12	z Critical two-tail			1.96	

$z = 4.75$; p-value = 0. There is sufficient evidence to indicate that wine bottled with a screw cap is perceived to be inferior.

A19.4 Histograms (not shown) are approximately bell shaped

Unequal-variances t-test of $\mu_1 - \mu_2$

$$H_0: (\mu_1 - \mu_2) = 0$$

$$H_1: (\mu_1 - \mu_2) < 0$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Unequal Variances		
2			
3		*British*	*American*
4	Mean	7137	9304
5	Variance	38051	110151
6	Observations	28	33
7	Hypothesized Mean Difference	0	
8	df	53	
9	t Stat	-31.61	
10	P(T<=t) one-tail	2.26E-36	
11	t Critical one-tail	1.6741	
12	P(T<=t) two-tail	4.51E-36	
13	t Critical two-tail	2.0057	

$t = -31.61$; p-value = 0. There is overwhelming evidence to conclude that the total distance of American golf courses is greater than that of British courses.

A19.5 Spearman rank correlation coefficient test

H_0: $\rho_S = 0$

H_1: $\rho_S > 0$

$$z = r_S\sqrt{n-1}$$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*Parking and Visits*			
4	Spearman Rank Correlation			0.3194
5	z Stat			4.47
6	P(Z<=z) one tail			0
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0
9	z Critical two tail			1.96

z = 4.47; p-value = 0. There is enough evidence to infer that the ratings of parking and the number of visits is positively related. Thus, there is enough evidence to conclude that the problem of parking is one reason for the decline in downtown shopping.

A19.6 t-test of ρ or β_1

$H_0: \rho = 0$

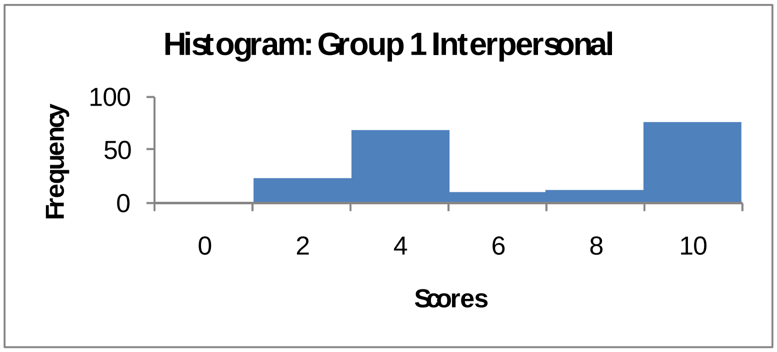
$H_1: \rho > 0$

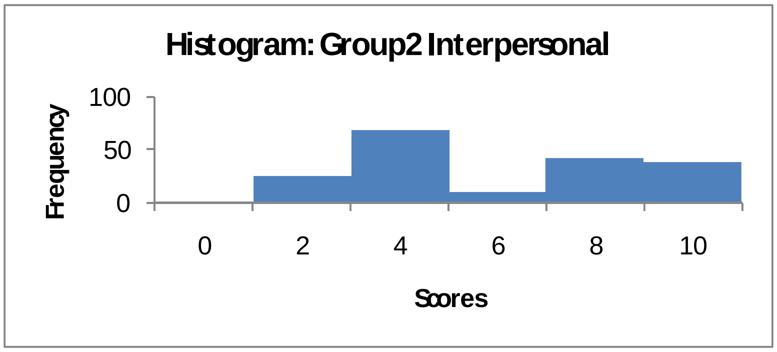
$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

	A	B
1	**Correlation**	
2		
3	*Visits and Income*	
4	Pearson Coefficient of Correlation	0.1747
5	t Stat	2.48
6	df	195
7	P(T<=t) one tail	0.0071
8	t Critical one tail	1.6527
9	P(T<=t) two tail	0.0142
10	t Critical two tail	1.9722

t = 2.48; p-value = .0071. There is enough evidence to infer that more affluent people shop more downtown than poorer people.

A19.7 Histograms of Groups 1 and 2: Interpersonal





The histograms indicate that the scores are extremely nonnormal.

Interpersonal skills: Wilcoxon rank sum test

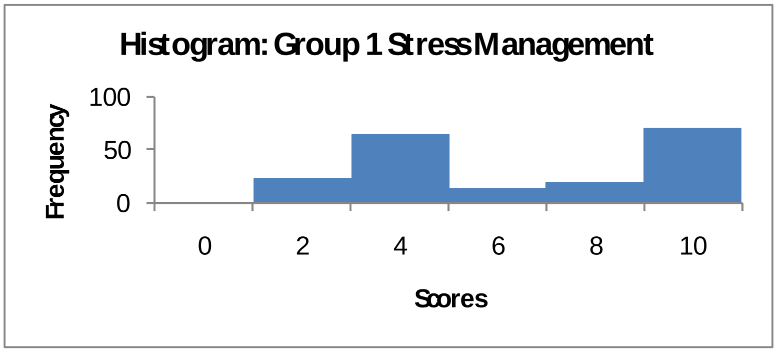
H_0: The two population locations are the same

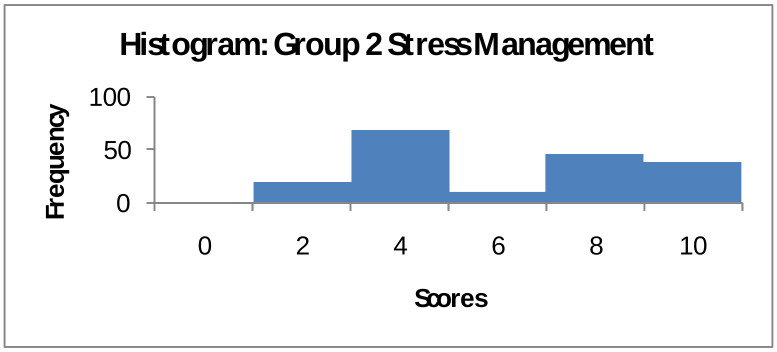
H_1: The location of population 1 is to the right of the location of population 2

	A	B	C	D
1	**Wilcoxon Rank Sum Test**			
2				
3			Rank Sum	Observations
4	*Gp 1 Interpersonal*		36392	187
5	*Gp 2 Interpersonal*		30769	179
6	z Stat		2.05	
7	P(Z<=z) one-tail		0.0200	
8	z Critical one-tail		1.6449	
9	P(Z<=z) two-tail		0.0400	
10	z Critical two-tail		1.96	

z = 2.05; p-value = .0200. There is enough evidence to conclude that students whose university GPA is greater than 3.0 score higher on interpersonal skills than students whose university GPA is less than 2.

Histograms of Groups 1 and 2: Stress Management





The histograms indicate that the scores are extremely nonnormal.

Stress management: Wilcoxon rank sum test

H_0: The two population locations are the same

H_1: The location of population 1 is to the right of the location of population 2

	A	B	C	D
1	**Wilcoxon Rank Sum Test**			
2				
3			Rank Sum	Observations
4	*Gp 1 Stress Mgmt*		36131.5	187
5	*Gp 2 Stress Mgmt*		31029.5	179
6	z Stat		1.80	
7	P(Z<=z) one-tail		0.0363	
8	z Critical one-tail		1.6449	
9	P(Z<=z) two-tail		0.0726	
10	z Critical two-tail		1.96	

z = 1.80; p-value = .0363. There is enough evidence to conclude that students whose university GPA is greater than 3.0 score higher on stress management than students whose university GPA is less than 2.

A19.8 All histograms (not shown) are somewhat bell shaped.

One-way analysis of variance

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1: At least two means differ

Miami-Dade

	A	B	C	D	E	F	G
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	6,409,467,776	2	3,204,733,888	231.37	1.1E-57	3.03
13	Within Groups	3,476,645,492	251	13,851,177			
14							
15	Total	9,886,113,268	253				

F = 231.37; p-value – 0. There is overwhelming evidence to infer that there are differences between the three groups of Americans residing in Miami-Dade.

Florida

	A	B	C	D	E	F	G
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	4,160,159,751	2	2,080,079,875	99.21	7.62E-34	3.03
13	Within Groups	6,331,923,528	302	20,966,634			
14							
15	Total	10,492,083,279	304				

F = 99.21; p-value – 0. There is overwhelming evidence to infer that there are differences between the three groups of Americans residing in the state of Florida.

United States

	A	B	C	D	E	F	G
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	5,742,495,149	2	2,871,247,574	110.48	9.27E-41	3.01
13	Within Groups	13,618,559,386	524	25,989,617			
14							
15	Total	19,361,054,535	526				

F = 110.48; p-value – 0. There is overwhelming evidence to infer that there are differences between the three groups of Americans residing in the United States.

A19.9 t-estimator of μ

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Cars

	A	B	C	D
1	**t-Estimate: Mean**			
2				
3				*Cars*
4	Mean			12.20
5	Standard Deviation			2.94
6	LCL			11.86
7	UCL			12.55

Estimate of the total number of miles driven by cars:

LCL = 136 million(11.86 thousand) = 1,613 billion miles

UCL = 136 million(12.55 thousand) = 1,707 billion miles

Buses

	A	B	C	D
1	**t-Estimate: Mean**			
2				
3				*Buses*
4	Mean			8.50
5	Standard Deviation			2.05
6	LCL			8.11
7	UCL			8.88

Estimate of the total number of miles driven by buses:

LCL = 776 thousand (8.11 thousand) = 6,293 million miles

UCL = 776 thousand (8.88thousand) = 6,891 million miles

Vans, pickups, SUV's

	A	B	C	D
1	**t-Estimate: Mean**			
2				
3				*Vans, pickups, SUVs*
4	Mean			11.50
5	Standard Deviation			2.83
6	LCL			11.09
7	UCL			11.91

Estimate of the total number of miles driven by vans, pickups, SUV's

LCL = 87 million (11.09 thousand) = 965 billion miles

UCL = 87 million (11.91 thousand) = 1,036 billion miles

A19.10 All three histograms (not shown) are bell shaped.

One-way analysis of variance

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1: At least two means differ

	A	B	C	D	E	F	G
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	1116.46	2	558.23	73.45	3.90E-29	3.01
13	Within Groups	4369.92	575	7.60			
14							
15	Total	5486.38	577				

F = 73.45; p-value = 0. There is overwhelming evidence to infer that there are differences in distance driven between cars, buses, and vans, pickups, and SUV's.

A19.11 t-estimator of μ

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

	A	B	C	D
1	**t-Estimate: Mean**			
2				
3				*Debt*
4	Mean			72.15
5	Standard Deviation			37.54
6	LCL			69.11
7	UCL			75.18

Estimate of total debt:

LCL = 10 million(69.11 thousand) = $691 billion

UCL = 10 million(75.18 thousand) = $751.8 billion

A19.12 Both histograms (not shown) are positively skewed but not sufficiently so to violate the normality requirement of the t-test of $\mu_1 - \mu_2$.

Equal-variances t-test of $\mu_1 - \mu_2$

H_0: $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 - \mu_2 < 0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Equal Variances		
2			
3		*5 years ago*	*This year*
4	Mean	7.14	7.81
5	Variance	12.87	18.64
6	Observations	84	91
7	Pooled Variance	15.87	
8	Hypothesized Mean Difference	0	
9	df	173	
10	t Stat	-1.11	
11	P(T<=t) one-tail	0.1338	
12	t Critical one-tail	1.6537	
13	P(T<=t) two-tail	0.2677	
14	t Critical two-tail	1.9738	

$t = -1.11$; p-value = .1338. There is not enough evidence to allow us to infer that investors' portfolios are becoming more diverse.

A19.13 Both histograms (not shown) are bell shaped.

Equal-variances t-test of $\mu_1 - \mu_2$

H_0: $\mu_1 - \mu_2 = 5$

H_1: $\mu_1 - \mu_2 > 5$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Equal Variances		
2			
3		*MBA*	*BBA*
4	Mean	96.53	89.08
5	Variance	156.70	106.44
6	Observations	79	121
7	Pooled Variance	126.24	
8	Hypothesized Mean Difference	5	
9	df	198	
10	t Stat	1.51	
11	P(T<=t) one-tail	0.0665	
12	t Critical one-tail	1.6526	
13	P(T<=t) two-tail	0.1330	
14	t Critical two-tail	1.9720	

t = 1.51; p-value = .0665. There is not enough evidence to infer that an MBA is worthwhile.

A19.14 Chi-squared test of a contingency table

H_0 : The two variables are independent

H_1 : The two variables are dependent

$$\chi^2 = \sum_{i=1}^{15} \frac{(f_i - e_i)^2}{e_i}$$

	A	B	C	D	E	F	G	H
1	**Contingency Table**							
2								
3		*Group*						
4	*Side Effect*		1	2	3	4	5	TOTAL
5		1	11	10	10	30	106	167
6		2	126	61	54	101	440	782
7		3	111	55	41	78	473	758
8		TOTAL	248	126	105	209	1019	1707
9								
10								
11		chi-squared Stat			20.6415			
12		df			8			
13		p-value			0.0082			
14		chi-squared Critical			15.5073			

$\chi^2 = 20.64$; p-value = .0082. There is not enough evidence to conclude that there are differences in side effects between the three groups.

A19.15 Multiple regression

	A	B	C	D	E	F
1	SUMMARY OUTPUT					
2						
3	*Regression Statistics*					
4	Multiple R	0.8189				
5	R Square	0.6705				
6	Adjusted R Square	0.6493				
7	Standard Error	2.97				
8	Observations	100				
9						
10	ANOVA					
11		*df*	*SS*	*MS*	*F*	*Significance F*
12	Regression	6	1667.21	277.87	31.54	2.01E-20
13	Residual	93	819.23	8.81		
14	Total	99	2486.44			
15						
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	
17	Intercept	1.005	1.387	0.725	0.4704	
18	Dexterity	-0.059	0.150	-0.393	0.6953	
19	Detail	-0.177	0.159	-1.108	0.2708	
20	Teamwork	0.182	0.157	1.161	0.2485	
21	Math	0.065	0.162	0.404	0.6874	
22	ProbSolve	1.002	0.160	6.268	1.13E-08	
23	Tech	1.782	0.159	11.239	5.10E-19	

Only problems-solving skills and technical knowledge are linearly related to quality.

A19.16 a t-test of μ

$$H_0: \mu = 0$$

$$H_1: \mu > 0$$

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

	A	B	C	D
1	**t-Test: Mean**			
2				
3				*Decrease*
4	Mean			24.73
5	Standard Deviation			17.92
6	Hypothesized Mean			0
7	df			222
8	t Stat			20.61
9	P(T<=t) one-tail			0
10	t Critical one-tail			1.6517
11	P(T<=t) two-tail			0
12	t Critical two-tail			1.9707

t = 20.61; p-value = 0. There is overwhelming evidence to infer that there is a decrease in metabolism when children watch television.

b Both histograms (not shown) are roughly bell shaped.

Unequal-variances t-test of $\mu_1 - \mu_2$

$$H_0: (\mu_1 - \mu_2) = 0$$

$$H_1: (\mu_1 - \mu_2) > 0$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Unequal Variances		
2			
3		*Obese*	*Nonobese*
4	Mean	30.86	23.35
5	Variance	112.85	358.58
6	Observations	41	182
7	Hypothesized Mean Difference	0	
8	df	106	
9	t Stat	3.46	
10	P(T<=t) one-tail	0.0004	
11	t Critical one-tail	1.6594	
12	P(T<=t) two-tail	0.0008	
13	t Critical two-tail	1.9826	

t = 3.46; p-value = .0004. There is enough evidence to conclude that the decrease in metabolism is greater among obese children.

A19.17 a z-test of p (success = mathematics major wins)

$$H_0: p = .5$$

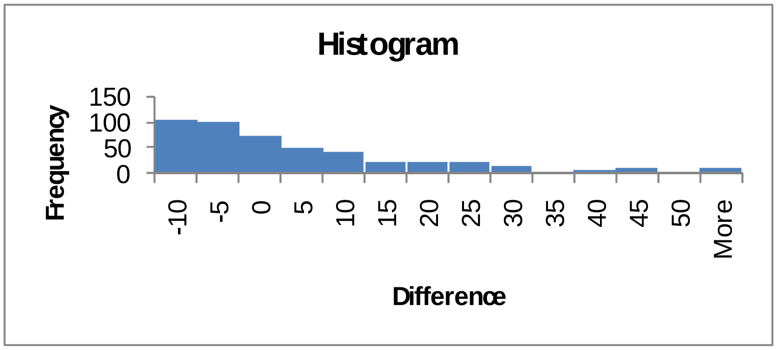
$$H_1: p > .5$$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

	A	B	C	D
1	**z-Test: Proportion**			
2				
3				*Math win*
4	Sample Proportion			0.558
5	Observations			500
6	Hypothesized Proportion			0.5
7	z Stat			2.59
8	P(Z<=z) one-tail			0.0047
9	z Critical one-tail			1.6449
10	P(Z<=z) two-tail			0.0094
11	z Critical two-tail			1.9600

z = 2.59; p-value = .0047. There is enough evidence to infer that mathematics majors win more frequently than English majors.

b Histogram of differences



The matched pairs differences are extremely nonnormal.

Wilcoxon signed rank sum test

H_0 : The two population locations are the same.

H_1 : The location of population 1 is to the right of the location of population 2.

	A	B	C	D
1	**Wilcoxon Signed Rank Sum Test**			
2				
3	Difference		*English - Math*	
4				
5	T+		67513.5	
6	T-		57736.5	
7	Observations (for test)		500	
8	z Stat		1.512	
9	P(Z<=z) one-tail		0.0652	
10	z Critical one-tail		1.6449	
11	P(Z<=z) two-tail		0.1304	
12	z Critical two-tail		1.96	

z = 1.512, p-value = .0652. There is not enough evidence to infer that English majors outscore mathematics majors.

c Part (a) ignores the magnitude of the paired differences.

A19.18a Histograms (not shown) are bell shaped.

Equal-variances t-test of $\mu_1 - \mu_2$

$$H_0 : (\mu_1 - \mu_2) = 0$$

$$H_1 : (\mu_1 - \mu_2) < 0$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Equal Variances		
2			
3		*4 or more*	*Less*
4	Mean	6.00	7.40
5	Variance	5.62	8.61
6	Observations	100	100
7	Pooled Variance	7.11	
8	Hypothesized Mean Difference	0	
9	df	198	
10	t Stat	-3.71	
11	P(T<=t) one-tail	0.0001	
12	t Critical one-tail	1.6526	
13	P(T<=t) two-tail	0.0003	
14	t Critical two-tail	1.9720	

t = –3.71; p-value = .0001. There is enough evidence to infer that children who wash their hands four or more times per day have less sick days due to cold and flu.

b Histograms (not shown) are bell shaped.

Unequal-variances t-test of $\mu_1 - \mu_2$

$$H_0 : (\mu_1 - \mu_2) = 0$$

$$H_1 : (\mu_1 - \mu_2) < 0$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Unequal Variances		
2			
3		*Four or more*	*Less*
4	Mean	1.76	3.17
5	Variance	1.48	3.19
6	Observations	100	100
7	Hypothesized Mean Difference	0	
8	df	174	
9	t Stat	-6.52	
10	P(T<=t) one-tail	0.0000	
11	t Critical one-tail	1.6537	
12	P(T<=t) two-tail	0.0000	
13	t Critical two-tail	1.9737	

t = –6.52; p-value = 0. There is enough evidence to infer that children who wash their hands four or more times per day have less sick days due to stomach illness.

A19.19 Spearman rank correlation coefficient test

H_0: $\rho_S = 0$

H_1: $\rho_S > 0$

$z = r_S \sqrt{n-1}$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*EI and Satisfaction*			
4	Spearman Rank Correlation			0.1887
5	z Stat			2.55
6	P(Z<=z) one tail			0.0054
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0.0108
9	z Critical two tail			1.96

z = 2.55; p-value = .0054. There is evidence to conclude that workers who use EI more often are more satisfied with their employment situation.

A19.20 a Histograms (not shown) of sick days are bell shaped.

Unequal-variances t-test of $\mu_1 - \mu_2$

$H_0: (\mu_1 - \mu_2) = 0$

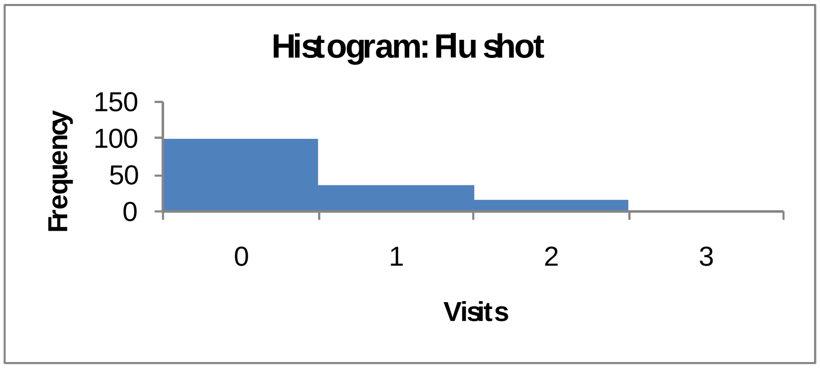
$H_1: (\mu_1 - \mu_2) < 0$

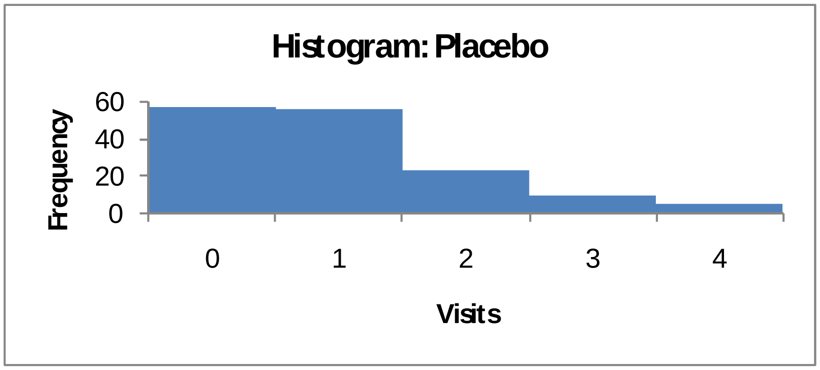
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

	A	B	C
1	t-Test: Two-Sample Assuming Unequal Variances		
2			
3		*Days flu shot*	*Days placebo*
4	Mean	2.82	3.22
5	Variance	1.25	2.07
6	Observations	150	150
7	Hypothesized Mean Difference	0	
8	df	281	
9	t Stat	-2.69	
10	P(T<=t) one-tail	0.0038	
11	t Critical one-tail	1.6503	
12	P(T<=t) two-tail	0.0076	
13	t Critical two-tail	1.9684	

t = –2.69; p-value = .0038. There is overwhelming evidence to indicate that the number of sick days is less for those who take the flu shots.

b Histograms of number of visits





The number of visits is extremely nonnormal.

Wilcoxon rank sum test

H_0: The two population locations are the same

H_1: The location of population 1 is to the left of the location of population 2

	A	B	C	D
1	**Wilcoxon Rank Sum Test**			
2				
3			Rank Sum	Observations
4	*Visits flu shot*		19152	150
5	*Visits placebo*		25998	150
6	z Stat		-4.56	
7	P(Z<=z) one-tail		0	
8	z Critical one-tail		1.6449	
9	P(Z<=z) two-tail		0	
10	z Critical two-tail		1.9600	

$z = -4.56$; p-value = 0. There is overwhelming evidence to indicate that those who take the flu shots visit their doctors less frequently

A19.21 Spearman rank correlation coefficient test

a $H_0 : \rho_S = 0$

$H_1 : \rho_S \neq 0$

$z = r_S \sqrt{n-1}$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*Satisfied and Severity*			
4	Spearman Rank Correlation			-0.2604
5	z Stat			-3.90
6	P(Z<=z) one tail			0
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0
9	z Critical two tail			1.96

$z = -3.90$; p-value = .0001. There is overwhelming evidence to conclude that satisfaction level is affected by severity of illness.

b $H_0 : \rho_S = 0$

$H_1 : \rho_S < 0$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*Satisfied and Days*			
4	Spearman Rank Correlation			-0.3846
5	z Stat			-5.76
6	P(Z<=z) one tail			0
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0
9	z Critical two tail			1.96

$z = -5.76$; p-value = 0. There is overwhelming evidence to conclude that satisfaction level is higher for patients who stay for shorter periods of time.

A19.22 Chi-squared test of a contingency table

H_0 : The two variables are independent

H_1 : The two variables are dependent

$$\chi^2 = \sum_{i=1}^{6} \frac{(f_i - e_i)^2}{e_i}$$

	A	B	C	D	E
1	**Contingency Table**				
2					
3		*City*			
4	*Outcome*		1	2	TOTAL
5		1	503	201	704
6		2	536	215	751
7		3	308	83	391
8		TOTAL	1347	499	1846
9					
10					
11		chi-squared Stat			8.47
12		df			2
13		p-value			0.0145
14		chi-squared Critical			5.9915

$\chi^2 = 8.47$; p-value = .0145. There is enough evidence to conclude that there are differences in the death rate between the three cities. Recommendation: Spend less time at accident scene and get patient to the hospital as soon as possible.

A19.23 a z-estimator of p

$$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$$

	A	B
1	**z-Estimate: Proportion**	
2		*Golfer?*
3	Sample Proportion	0.1586
4	Observations	1116
5	LCL	0.1372
6	UCL	0.1800

Estimate of the total number of golfers:

LCL = 207.7 million (.1372) = 28.5 million

UCL = 207.7 million (.1800) = 37.4 million

b. z-estimator of p

$$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$$

There were 53 people in the highest income category who did not play golf and 68 who did.

	A	B	C	D	E
1	**z-Estimate of a Proportion**				
2					
3	**Sample proportion**	0.562	**Confidence Interval Estimate**		
4	**Sample size**	121	**0.562**	±	**0.088**
5	**Confidence level**	0.95	**Lower confidence limit**		**0.474**
6			**Upper confidence limit**		**0.650**

Estimate of the number of golfers who earn at least $75,000:

LCL = 17.1 million (.474) = 8.1 million

UCL = 17.1 million (.650) = 11.1 million

c Chi-squared test of a contingency table

H_0 : The two variables are independent

H_1 : The two variables are dependent

$$\chi^2 = \sum_{i=1}^{12} \frac{(f_i - e_i)^2}{e_i}$$

	A	B	C	D	E
1	**Contingency Table**				
2					
3		*Income*			
4	*Golfer?*		1	2	TOTAL
5		1	208	6	214
6		2	169	7	176
7		3	148	12	160
8		4	172	30	202
9		5	189	54	243
10		6	53	68	121
11		TOTAL	939	177	1116
12					
13					
14		chi-squared Stat			209.4
15		df			5
16		p-value			0
17		chi-squared Critical			11.0705

$\chi^2 = 209.4$; p-value = 0. There is enough evidence to conclude that income is a determinant in who plays golf.

A19.24 Chi-squared test of a contingency table

H_0 : The two variables are independent

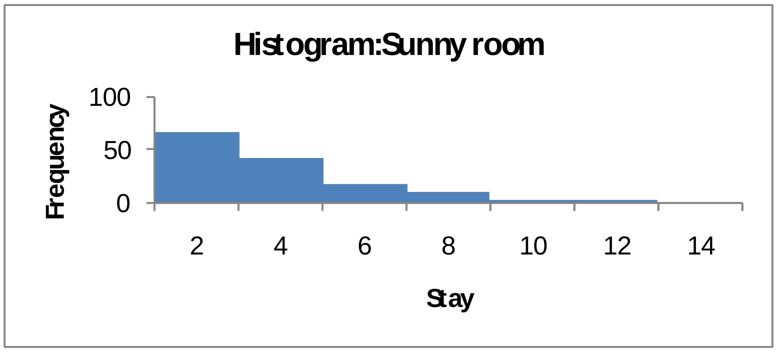
H_1 : The two variables are dependent

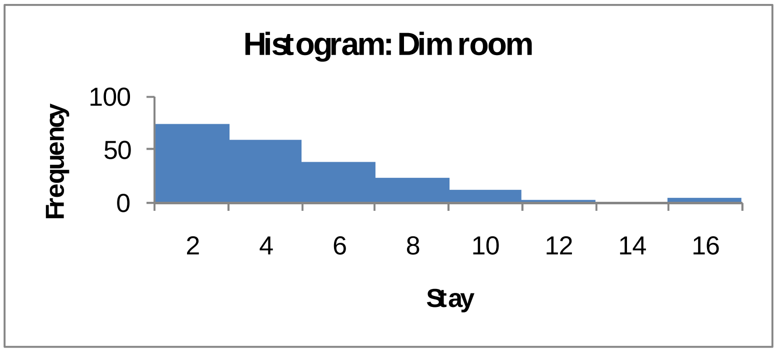
$$\chi^2 = \sum_{i=1}^{6} \frac{(f_i - e_i)^2}{e_i}$$

	A	B	C	D	E
1	**Contingency Table**				
2					
3		*Income*			
4	*University?*		1	2	TOTAL
5		1	101	29	130
6		2	261	87	348
7		3	40	25	65
8		TOTAL	402	141	543
9					
10					
11		chi-squared Stat			6.35
12		df			2
13		p-value			0.0417
14		chi-squared Critical			5.9915

$\chi^2 = 6.35$; p-value = .0417. There is enough evidence to conclude that family income affects whether children attend university.

A19.25a Histograms of stays





Length of stays is extremely nonnormal.

Wilcoxon rank sum test

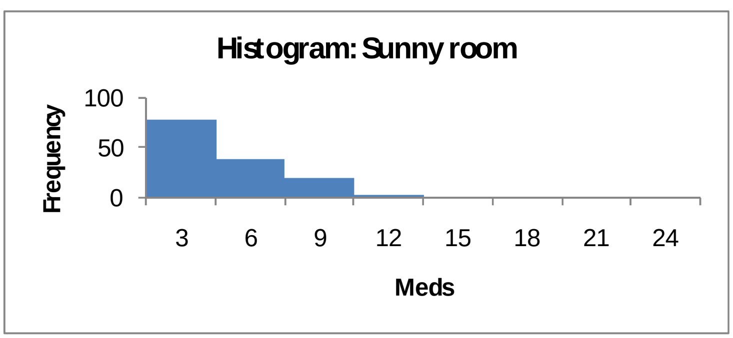
H_0: The two population locations are the same

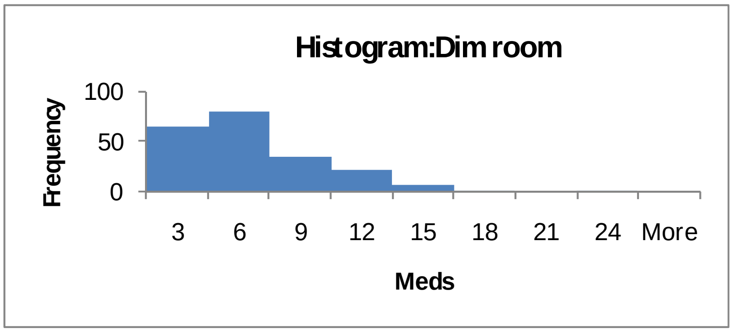
H_1: The location of population 1 is to the left of the location of population 2

	A	B	C	D	E
1	**Wilcoxon Rank Sum Test**				
2					
3			Rank Sum	Observations	
4	*Sunny stay*		22655.5	143	
5	*Dim stay*		39472.5	209	
6	z Stat		-2.76		
7	P(Z<=z) one-tail		0.0029		
8	z Critical one-tail		1.6449		
9	P(Z<=z) two-tail		0.0058		
10	z Critical two-tail		1.96		

z = -2.76; p-value = .0029. There is enough evidence to conclude that the length of stay is smaller in bright rooms than in dim ones.

Histograms of amounts of pain medication





b. Amount of pain medication is extremely nonnormal.

Wilcoxon rank sum test

H_0: The two population locations are the same

H_1: The location of population 1 is to the left of the location of population 2

	A	B	C	D	E
1	**Wilcoxon Rank Sum Test**				
2					
3			Rank Sum	Observations	
4	*Sunny med*		21197.5	143	
5	*Dim med*		40930.5	209	
6	z Stat		-4.31		
7	P(Z<=z) one-tail		0		
8	z Critical one-tail		1.6449		
9	P(Z<=z) two-tail		0		
10	z Critical two-tail		1.96		

z = −4.31; p-value = 0. There is enough evidence to conclude that the amount of pain medication is lower in bright rooms than in dim ones.

A19.26 Case 13.1: Wilcoxon rank sum test

H_0 : The two population locations are the same.

H_1 :The location of population 1 is to the right of the location of population 2.

	A	B	C	D
1	**Wilcoxon Rank Sum Test**			
2				
3			Rank Sum	Observations
4	*W Rate*		64365.5	101
5	*M Rate*		490565.5	952
6	z Stat		3.83	
7	P(Z<=z) one-tail		0.0001	
8	z Critical one-tail		1.6449	
9	P(Z<=z) two-tail		0.0002	
10	z Critical two-tail		1.9600	

z = 3.83, p-value = .0001. There is enough evidence to conclude that women pay higher rates of interest than men.

Case A16.1 Relationship between interest rates and sales: Spearman rank correlation coefficient test

$$H_0: \rho_S = 0$$

$$H_1: \rho_S \neq 0$$

$$z = r_S\sqrt{n-1}$$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*Rates and Sales*			
4	Spearman Rank Correlation			-0.2629
5	z Stat			-8.53
6	P(Z<=z) one tail			0
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0
9	z Critical two tail			1.9600

z = –8.53, p-value = 0. There is overwhelming evidence to infer that interest rates and sales are linearly related.

Relationship between interest rates and ages: Spearman rank correlation coefficient test

$$H_0: \rho_S = 0$$

$$H_1: \rho_S \neq 0$$

$$z = r_S\sqrt{n-2}$$

	A	B	C	D
1	**Spearman Rank Correlation**			
2				
3	*Rates and Age*			
4	Spearman Rank Correlation			-0.1853
5	z Stat			-6.01
6	P(Z<=z) one tail			0
7	z Critical one tail			1.6449
8	P(Z<=z) two tail			0
9	z Critical two tail			1.9600

z = –6.01, p-value = 0. There is overwhelming evidence to infer that interest rates and age of business are linearly related.

Difference between sales: Wilcoxon rank sum test

H_0 : The two population locations are the same.

H_1 : The location of population 1 is to the left of the location of population 2.

	A	B	C	D
1	**Wilcoxon Rank Sum Test**			
2				
3			Rank Sum	Observations
4	*W Sales*		12285	101
5	*M Sales*		542646	952
6	z Stat		-14.09	
7	P(Z<=z) one-tail		0	
8	z Critical one-tail		1.6449	
9	P(Z<=z) two-tail		0	
10	z Critical two-tail		1.9600	

z = –14.09, p-value = 0. There is sufficient evidence to conclude that businesses owned by women have lower sales than businesses owned by men.

Difference between ages: Wilcoxon rank sum test

H_0 : The two population locations are the same.

H_1 : The location of population 1 is to the left of the location of population 2.

	A	B	C	D
1	**Wilcoxon Rank Sum Test**			
2				
3			Rank Sum	Observations
4	*W Age*		35034.5	101
5	*M Age*		519896.5	952
6	z Stat		-6.26	
7	P(Z<=z) one-tail		0	
8	z Critical one-tail		1.6449	
9	P(Z<=z) two-tail		0	
10	z Critical two-tail		1.9600	

z = –6.26, p-value = 0. There is sufficient evidence to conclude that businesses owned by men are older than businesses owned by women.

Interest rates among the 3 types of businesses: Kruskal Wallis test

H_0 : The locations of all 3 populations are the same

H_1 : At least two population locations differ

$$H = \left[\frac{12}{n(n+1)}\sum\frac{T_j^2}{n_j}\right] - 3(n+1)$$

	A	B	C
1	**Kruskal-Wallis Test**		
2			
3	Group	Rank Sum	Observations
4	*Bus 1*	111554	193
5	*Bus 2*	42305.5	86
6	*Bus 3*	401071.5	774
7			
8	H Stat		7.22
9	df		2
10	p-value		0.0270
11	chi-squared Critical		5.9915

H = 7.22, p-value = .0270. There is enough evidence to conclude that there are differences in interest rates among the three types of business.

Case A19.1a Chi-squared goodness-of-fit test

H_0 : $p_1 = 1/3,\ p_2 = 1/3,\ p_3 = 1/3$

H_1 : At least one p_i is not equal to its specified value

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i}$$

	A	B	C
1	Type	Actual	Expected
2	1	43	50
3	2	45	50
4	3	62	50
5		p-value	0.1130

p-value = .1130. There is not enough evidence to infer that the proportions of each type of personality type are different.

b Kruskal Wallis test

H_0 : The locations of all 5 populations are the same

H_1 : At least two population locations differ

$$H = \left[\frac{12}{n(n+1)} \sum \frac{T_j^2}{n_j}\right] - 3(n+1)$$

	A	B	C
1	**Kruskal-Wallis Test**		
2			
3	Group	Rank Sum	Observations
4	*Type A*	2579	43
5	*Type B*	3773	45
6	*Type C*	4973	62
7			
8	H Stat		7.88
9	df		2
10	p-value		0.0195
11	chi-squared Critical		5.9915

H = 7.88, p-value = .0195. There is enough evidence to infer that satisfaction levels differ between the three personality types.

c One-way analysis of variance

$H_0: \mu_1 = \mu_2 = \mu_3$

$H_1:$ At least two means differ

	A	B	C	D	E	F	G
11	ANOVA						
12	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
13	Between Groups	70.20	2	35.10	3.51	0.0325	3.06
14	Within Groups	1471.38	147	10.01			
15							
16	Total	1541.58	149				

F = 3.51, p-value = .0325. There is enough evidence to infer that life insurance sales differ between the three personality types

Case A19.2 t-test of μ_D

$$H_0: \mu_D = 0$$

$$H_1: \mu_D > 0$$

$$t = \frac{\bar{x}_D - \mu_D}{s_D / \sqrt{n_D}}$$

	A	B	C
1	t-Test: Paired Two Sample for Means		
2			
3		*W Rate*	*M Rate*
4	Mean	1.44	1.40
5	Variance	0.320	0.376
6	Observations	53	53
7	Pearson Correlation	0.47	
8	Hypothesized Mean Difference	0	
9	df	52	
10	t Stat	0.48	
11	P(T<=t) one-tail	0.3158	
12	t Critical one-tail	1.6747	
13	P(T<=t) two-tail	0.6315	
14	t Critical two-tail	2.0066	

t = .48, p-value = .3158. There is not enough evidence to infer that female business owners pay higher rates of interest than male business owners.

Because this experiment controls for all other relevant factors besides gender we can be confident that this statistical conclusion is definitive.