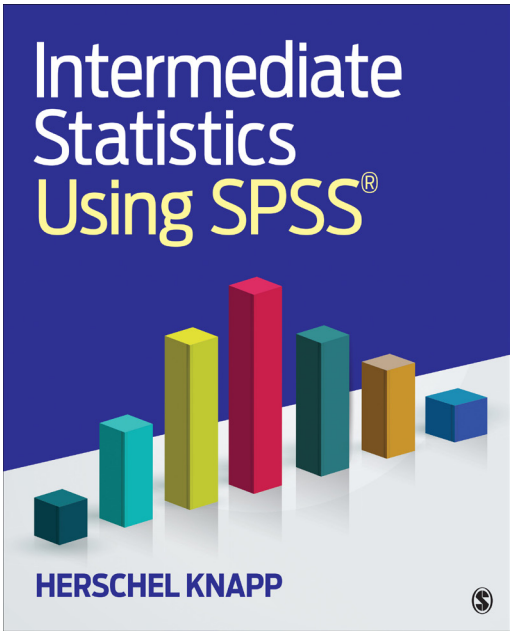


C H A P T E R 5

ANOVA and Kruskal-Wallis Test

Solutions to Odd-Numbered Exercises

EXERCISE 5.1, DATA SET A

(a)

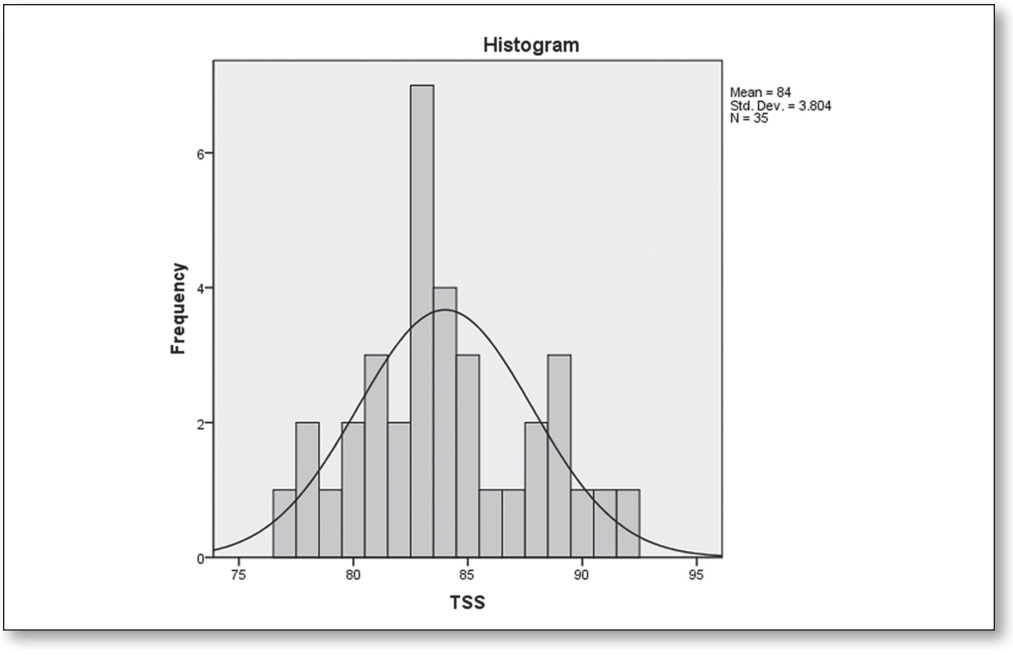
H_0: The number of students in a tutoring group has no effect on student satisfaction.

H_1: The number of students in a tutoring group has an effect on student satisfaction.

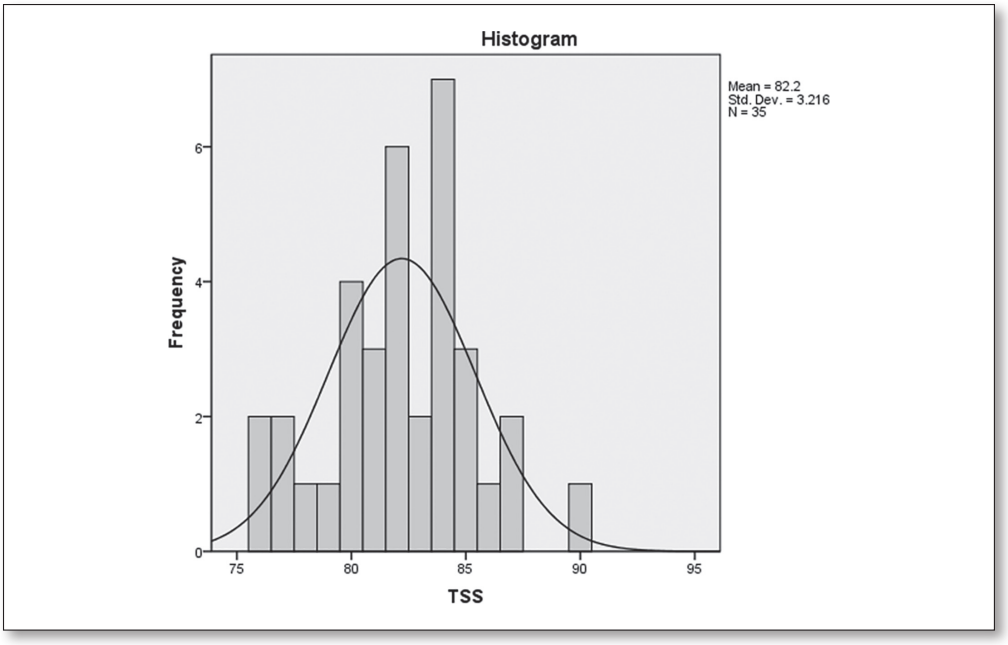
(b)

Histograms with normal curve plots show a normal distribution of *TSS* for all three groups as shown in the three figures below, hence, the pretest criterion of *normality* is satisfied.

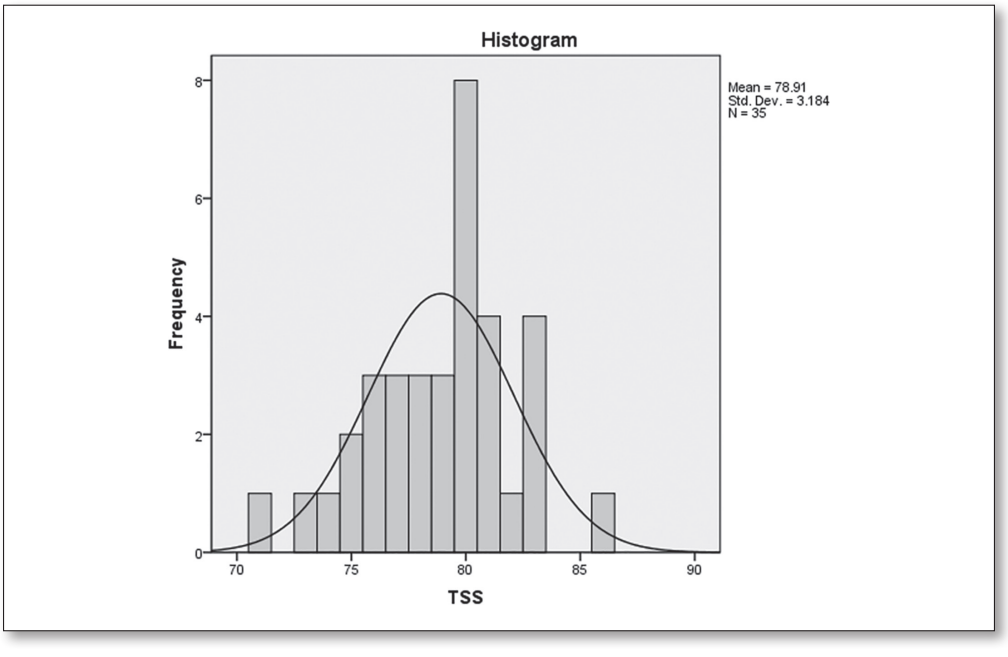
Normal distribution for *TSS* in Group 1 (One-to-one tutoring)



Normal distribution for *TSS* in Group 2 (Two-to-one tutoring)



Normal distribution for *TSS* in Group 3 (Five-to-one tutoring)



Test of Homogeneity of Variances

TSS

Levene Statistic	df1	df2	Sig.
.593	2	102	.555

The homogeneity of variance score shows a significance (*p*) of .555; since this is greater than the α level of .05, this suggests that there is no statistically significant difference among the variances of the three groups, hence, this pretest criterion passes.

The *n* for each group, as shown in the *Descriptives* table below is 35 for each group; since the *n*s are greater than 30, this criterion passes also.

(c)

The ANOVA revealed the following:

Descriptives

TSS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
One-to-one	35	84.00	3.804	.643	82.69	85.31	77	92
Two-to-one	35	82.20	3.216	.544	81.10	83.30	76	90
Five-to-one	35	78.91	3.184	.538	77.82	80.01	71	86
Total	105	81.70	3.988	.389	80.93	82.48	71	92

ANOVA

TSS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	465.505	2	232.752	19.978	.000
Within Groups	1188.343	102	11.650		
Total	1653.848	104			

Multiple Comparisons

TSS
Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
One-to-one	Two-to-one	1.800	.816	.075	-.14	3.74
	Five-to-one	5.086*	.816	.000	3.15	7.03
Two-to-one	One-to-one	-1.800	.816	.075	-3.74	.14
	Five-to-one	3.286*	.816	.000	1.35	5.23
Five-to-one	One-to-one	-5.086*	.816	.000	-7.03	-3.15
	Two-to-one	-3.286*	.816	.000	-5.23	-1.35

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

The Tukey post hoc test was used since the *n*s for each group were the same (35 each).

NOTE: Since the ANOVA test renders results involving multiple comparisons, it may be helpful to organize the findings as shown in the table below. SPSS does not generate this table directly, but you can construct it manually. You can copy the *group names* and *means* from the *Descriptives* table, and the *p* values from the *Sig.* column in the *Multiple Comparisons* table.

Groups (Tutor Satisfaction Survey)	*p*
One-to-one (M = 84.00) : Two-to-one (M = 82.20)	.075
One-to-one (M = 84.00) : Five-to-one (M = 78.91)	.000*
Two-to-one (M = 82.20) : Five-to-one (M = 78.91)	.000*

*Statistically significant difference ($\alpha = .05$).

These findings reveal that the highest student satisfaction was in the 1:1 tutoring group, however the 2:1 group was a close second, with no statistically significant difference between these two groups. Both the 1:1 and 2:1 tutor groups statistically significantly outperformed the (lowest) 5:1 group. Based on these findings, we reject H_0, and we do not reject H_1.

(d)

This study analyzed the effects that tutor group size had on student satisfaction scores. We recruited 105 students and randomly assigned them to one of three tutoring conditions: 1:1, 2:1, or 5:1. At the end of the term, each student completed the Tutor Satisfaction Survey (0 = very unsatisfied . . . 100 = very satisfied). We detected no statistically significant difference between the students who received 1:1 tutoring (M = 84.0, SD = 3.8) and those who received 2:1 tutoring (M = 82.2, SD = 3.2) ($p = .075$, $\alpha = .05$). The students in both of these groups had a statistically significantly higher Tutor Satisfaction score than those in the 5:1 group (M = 78.9, SD = 78.9), ($p < .001$, $\alpha = .05$). Based on these findings, we reject H_0 and accept H_1.

EXERCISE 5.1, DATA SET B

(a)

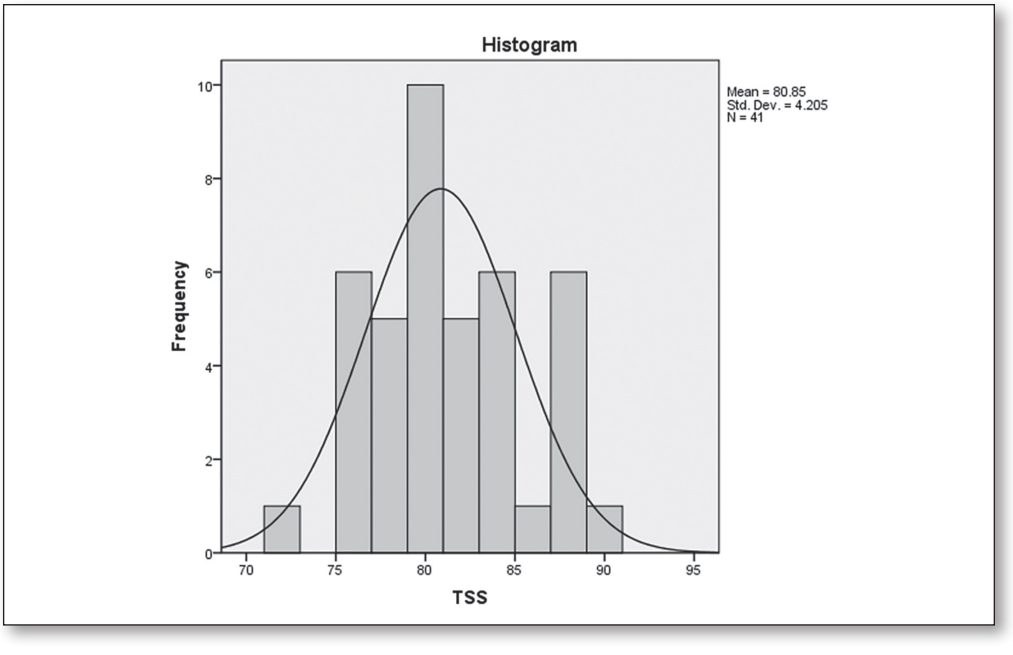
H_0: The number of students in a tutoring group has no effect on student satisfaction.

H_1: The number of students in a tutoring group has an effect on student satisfaction.

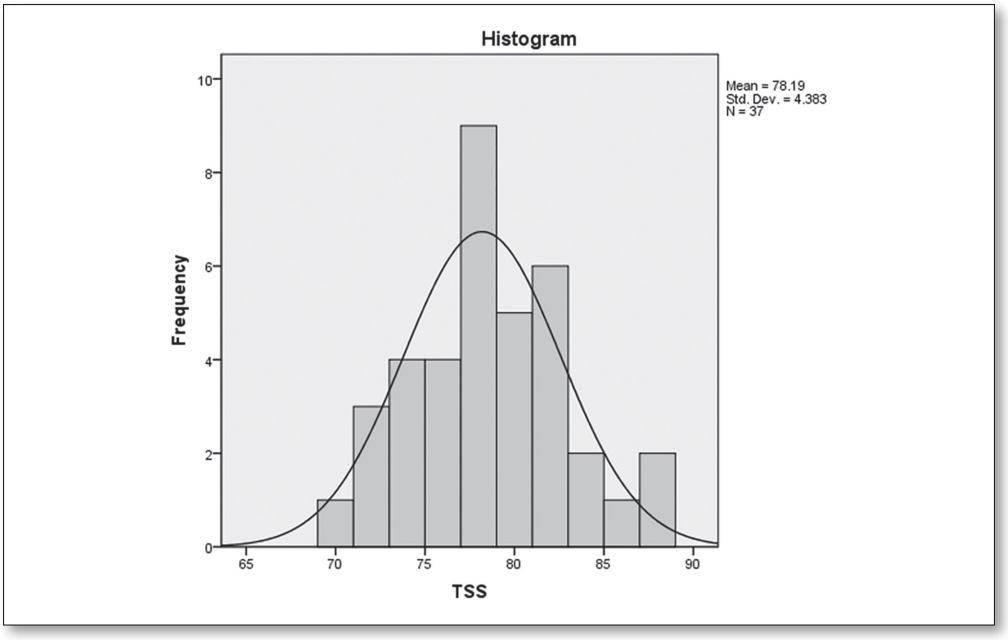
(b)

Histograms with normal curve plots show a normal distribution of *TSS* for all three groups as shown in the three figures below, hence, the pretest criterion of *normality* is satisfied.

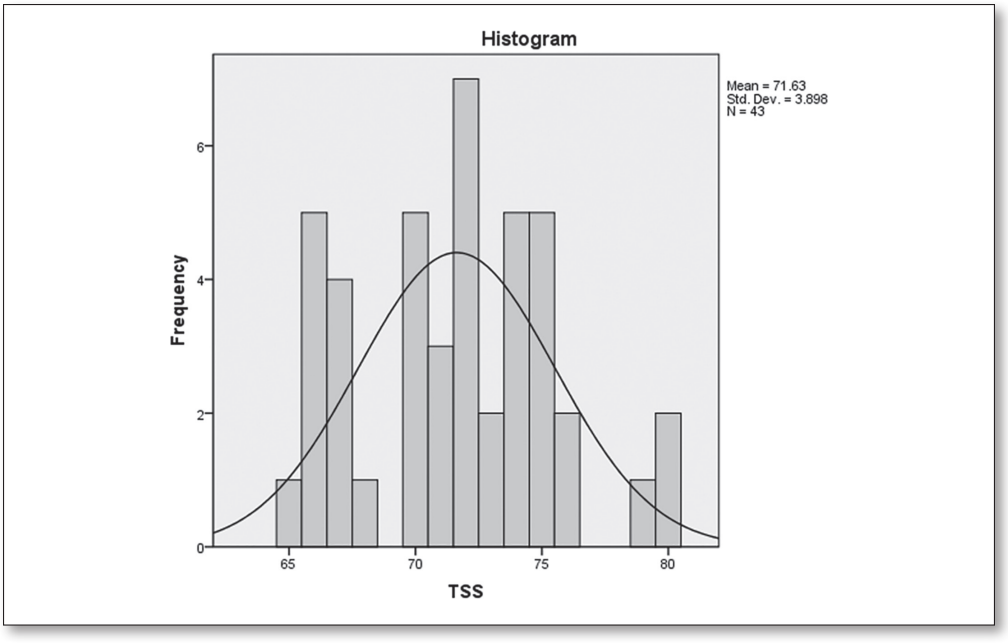
Normal distribution for *TSS* in Group 1 (One-to-one tutoring)



Normal distribution for *TSS* in Group 2 (Two-to-one tutoring)



Normal distribution for *TSS* in Group 3 (Five-to-one tutoring)



Test of Homogeneity of Variances

TSS

Levene Statistic	df1	df2	Sig.
.238	2	118	.789

The homogeneity of variance score shows a significance (*p*) of .789; since this is greater than the α level of .05, this suggests that there is no statistically significant difference among the variances of the three groups, hence, this pretest criterion passes.

The *n* for each group, as shown in the *Descriptives* table below are 41, 37, and 43; since the *n*s are greater than 30, this criterion passes also.

(c)

The ANOVA revealed the following:

Descriptives

TSS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
One-to-one	41	80.85	4.205	.657	79.53	82.18	72	90
Two-to-one	37	78.19	4.383	.721	76.73	79.65	70	87
Five-to-one	43	71.63	3.898	.594	70.43	72.83	65	80
Total	121	76.76	5.724	.520	75.73	77.79	65	90

ANOVA

TSS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1895.205	2	947.603	54.897	.000
Within Groups	2036.844	118	17.261		
Total	3932.050	120			

Multiple Comparisons

TSS

Sidak

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
One-to-one	Two-to-one	2.664*	.942	.016	.38	4.95
	Five-to-one	9.226*	.907	.000	7.03	11.42
Two-to-one	One-to-one	-2.664*	.942	.016	-4.95	-.38
	Five-to-one	6.561*	.932	.000	4.30	8.82
Five-to-one	One-to-one	-9.226*	.907	.000	-11.42	-7.03
	Two-to-one	-6.561*	.932	.000	-8.82	-4.30

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

The Sidak post hoc test was used since the *n*s for each group were not the same (*n*s = 41, 37, and 43).

NOTE: Since the ANOVA test renders results involving multiple comparisons, it may be helpful to organize the findings as shown in the table below. SPSS does not generate this table directly, but you can construct it manually. You can copy the *group names* and *means* from the *Descriptives* table, and the *p* values from the *Sig.* column in the *Multiple Comparisons* table.

Groups (Tutor Satisfaction Survey)	*p*
One-to-one (M = 80.85) : Two-to-one (M = 78.19)	.016*
One-to-one (M = 80.85) : Five-to-one (M = 71.63)	.000*
Two-to-one (M = 78.19) : Five-to-one (M = 71.63)	.000*

*Statistically significant difference ($\alpha = .05$).

These findings reveal that the highest student satisfaction was highest in the 1:1 tutoring group, followed by the 2:1 group, and finally, the 5:1 group. All group comparisons produced statistically significant differences. Based on these findings, we reject H_0, and we do not reject H_1.

(d)

This study analyzed the effects that tutor group size had on student satisfaction scores. We recruited 121 students and randomly assigned to one of three tutoring conditions: 1:1, 2:1, or 5:1. At the end of the term, each student completed the Tutor Satisfaction Survey (TSS) (0 = very unsatisfied . . . 100 = very satisfied). We detected that smaller tutor groups statistically significantly outperformed each of the larger groups when it comes to the TSS. Specifically, using an α level of .05, 1:1 tutoring (M = 80.9, SD = 4.2) outperformed 2:1 tutoring (M = 78.19, SD = 4.4) ($p = .016$), 2:1 tutoring (M = 78.19, SD = 4.4) outperformed 5:1 tutoring (M = 71.63, SD = 3.9) ($p < .001$). Also, 1:1 tutoring outperformed 5:1 tutoring ($p < .001$). Based on these findings, we reject H_0 and accept H_1.

EXERCISE 5.3, DATA SET A

(a)

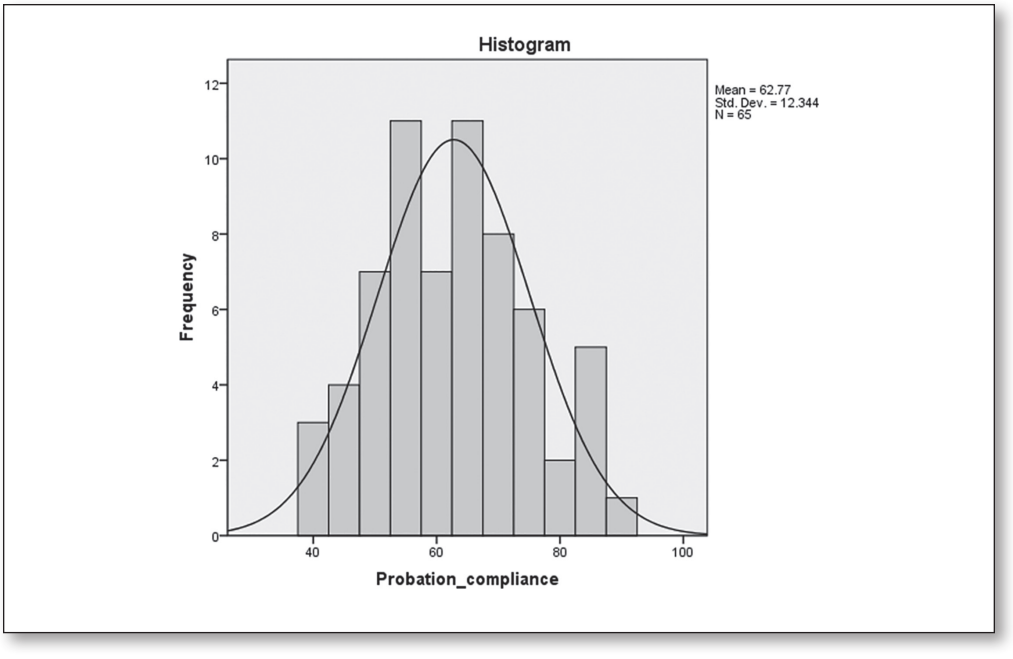
H_0: Having a mentor will have no effect on probation compliance.

H_1: Having a mentor will have an effect on probation compliance.

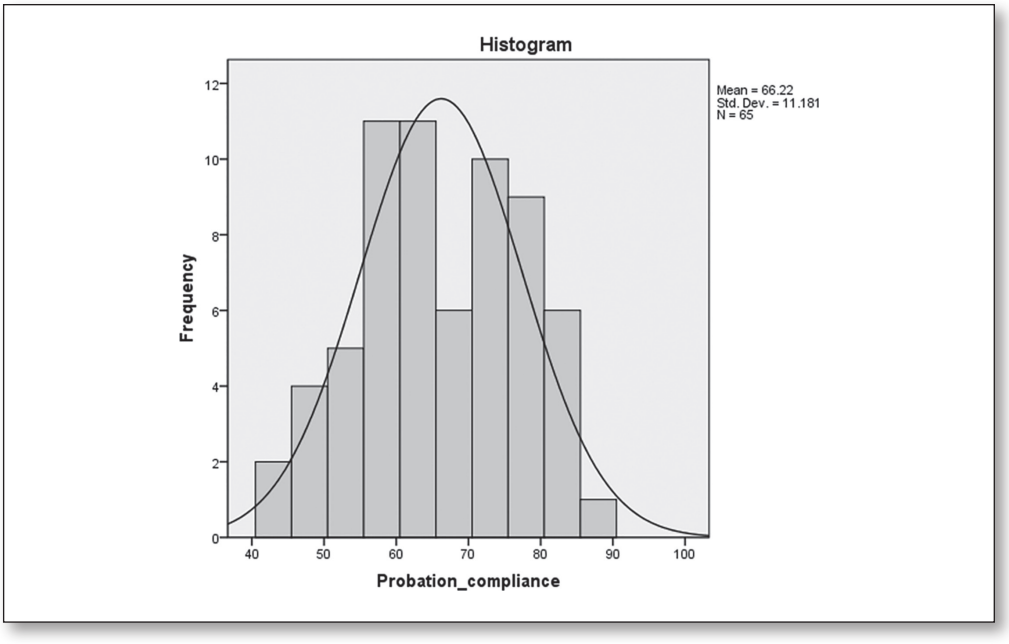
(b)

The Histograms with normal curve plots show a normal distribution of the *Probation_compliance* variable for all three groups as shown in the three figures below, hence, the pretest criterion of *normality* is satisfied.

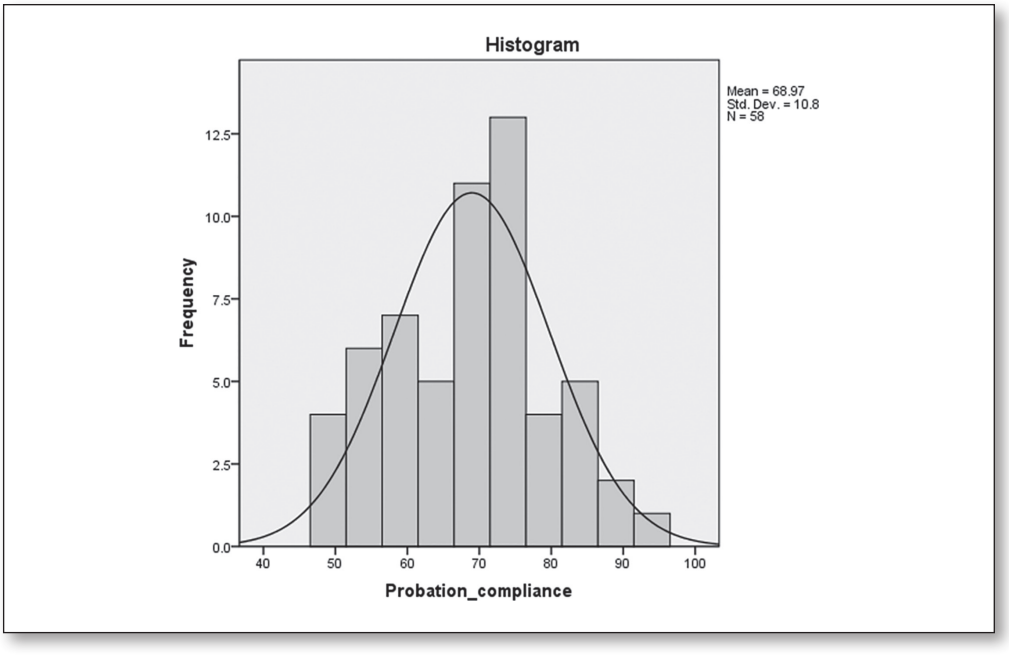
Normal distribution for *Probation_compliance* in Group 1 (No mentor)



Normal distribution for *Probation_compliance* in Group 2 (Peer mentor)



Normal distribution for *Probation_compliance* in Group 3 (Adult mentor)



Test of Homogeneity of Variances

Probation_compliance

Levene Statistic	df1	df2	Sig.
.685	2	185	.505

The homogeneity of variance score for *mood* shows a significance (*p*) of .505; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances among the three groups, hence, this pretest criterion passes.

The *n* for the groups are 65, 65, and 58 (see *Descriptives* table below), which satisfies the 30 per group minimum criterion.

(c)

The ANOVA test revealed the following:

Descriptives

Probation_compliance

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
No mentor	65	62.77	12.344	1.531	59.71	65.83	40	89
Peer mentor	65	66.22	11.181	1.387	63.44	68.99	43	87
Adult mentor	58	68.97	10.800	1.418	66.13	71.81	49	93
Total	188	65.87	11.698	.853	64.19	67.56	40	93

ANOVA

Probation_compliance

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1188.482	2	594.241	4.505	.012
Within Groups	24400.454	185	131.894		
Total	25588.936	187			

Multiple Comparisons

Probation_compliance
Sidak

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
No mentor	Peer mentor	-3.446	2.015	.243	-8.30	1.41
	Adult mentor	-6.196*	2.074	.010	-11.19	-1.20
Peer mentor	No mentor	3.446	2.015	.243	-1.41	8.30
	Adult mentor	-2.750	2.074	.462	-7.75	2.25
Adult mentor	No mentor	6.196*	2.074	.010	1.20	11.19
	Peer mentor	2.750	2.074	.462	-2.25	7.75

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

Groups (Probation Compliance)	*p*
No mentor (M = 62.77) : Peer mentor (M = 66.22)	.243
No mentor (M = 62.77) : Adult mentor (M = 68.97)	.010*
Peer mentor (M = 66.22) : Adult mentor (M = 68.97)	.462

*Statistically significant difference ($\alpha = .05$).

We discovered that youth offenders had significantly higher probation compliance when paired with an adult mentor (M = 68.97, SD = 10.80) compared to having no mentor (M = 62.77, SD = 12.34) (p = .010, α = .05). There was no statistically significant difference between those who had an adult mentor compared to those who had a peer mentor (M = 66.22, SD = 11.81) (p = .462, = .05). Finally, we detected no statistically significant difference between those who had a peer mentor and those who had no mentor (p = .243, α = .05).

Since those in the Adult mentor group had a statistically significantly higher probation compliance than those who had no mentor, we would reject H_0. By that same reasoning, we would not reject H_1.

(d)

A judge appointed us to evaluate the effectiveness of a new mentorship program for juvenile offenders with priors. The 188 juveniles were randomly assigned to one of three groups: no mentor, a trained peer mentor who is 3 to 5 years older than the offender, or a trained adult mentor. Those paired with an adult mentor had the highest average level of probation compliance (M = 68.97, SD = 10.80), statistically significantly outperforming those who had no mentor (M = 62.77, SD = 12.34) (p = .010, α = .05). No statistically significant differences in probation compliance were detected between those in no mentor group (M = 62.77, SD = 12.34) and the peer mentor group (M = 66.22, SD = 11.18) (p = .243, α = .05), or the peer mentor (M = 66.22, SD = 11.18) and the adult mentor (M = 68.97, SD = 10.80) (p = .462, α = .05). Based on these findings, we reject H_0 and accept H_1. These results suggest that adult mentors are the optimal choice for enhancing probation compliance for this population.

EXERCISE 5.3, DATA SET B

(a)

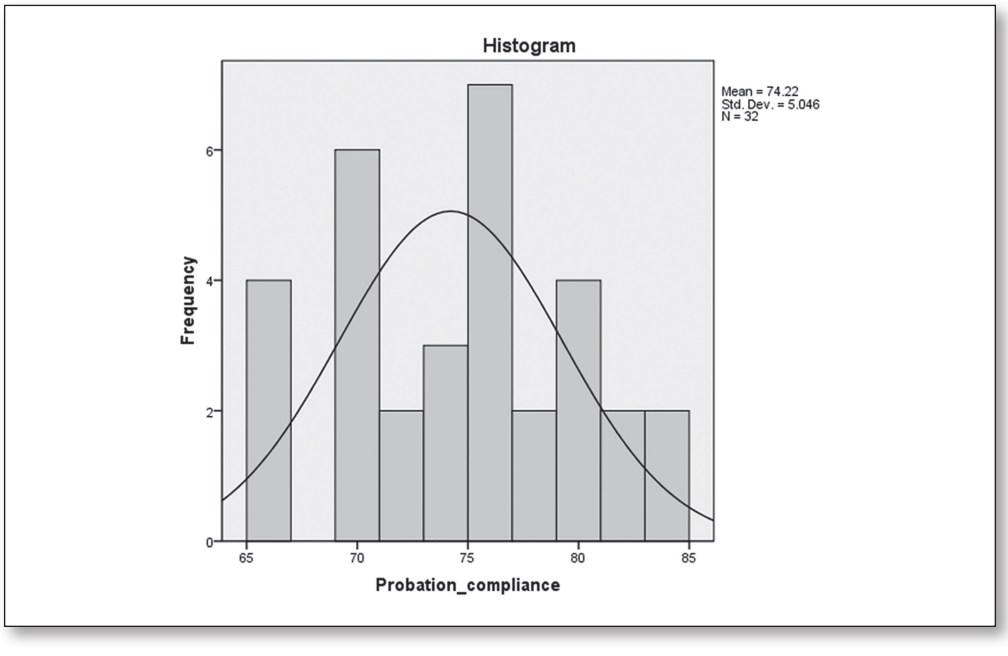
H_0: Having a mentor will have no effect on probation compliance.

H_1: Having a mentor will have an effect on probation compliance.

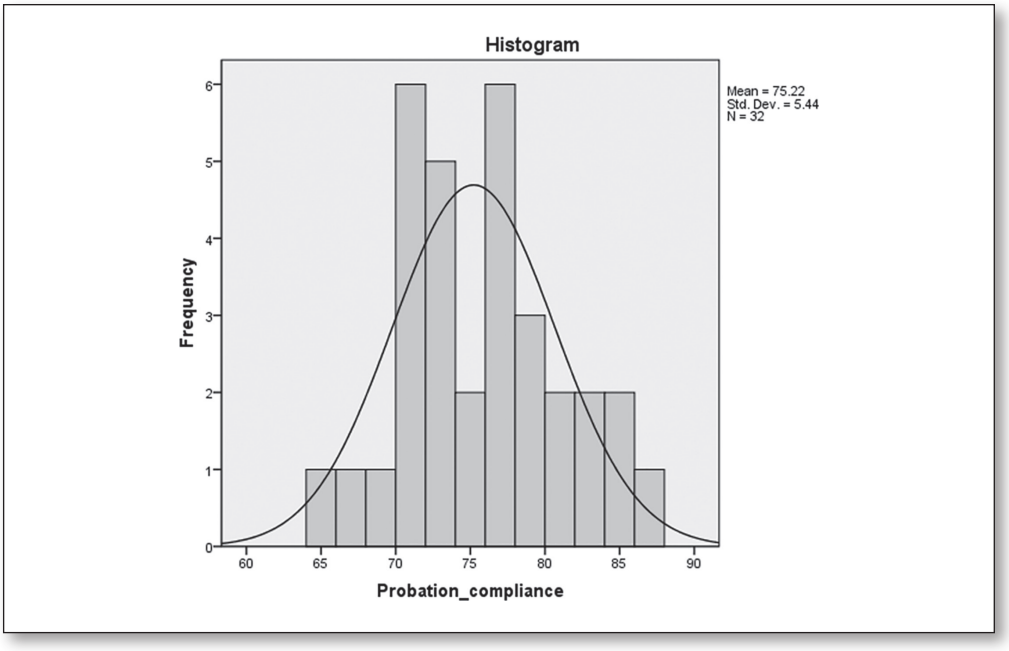
(b)

The Histograms with normal curve plots show a normal distribution of the *Probation_compliance* variable for all three groups as shown in the three figures below, hence, the pretest criterion of *normality* is satisfied.

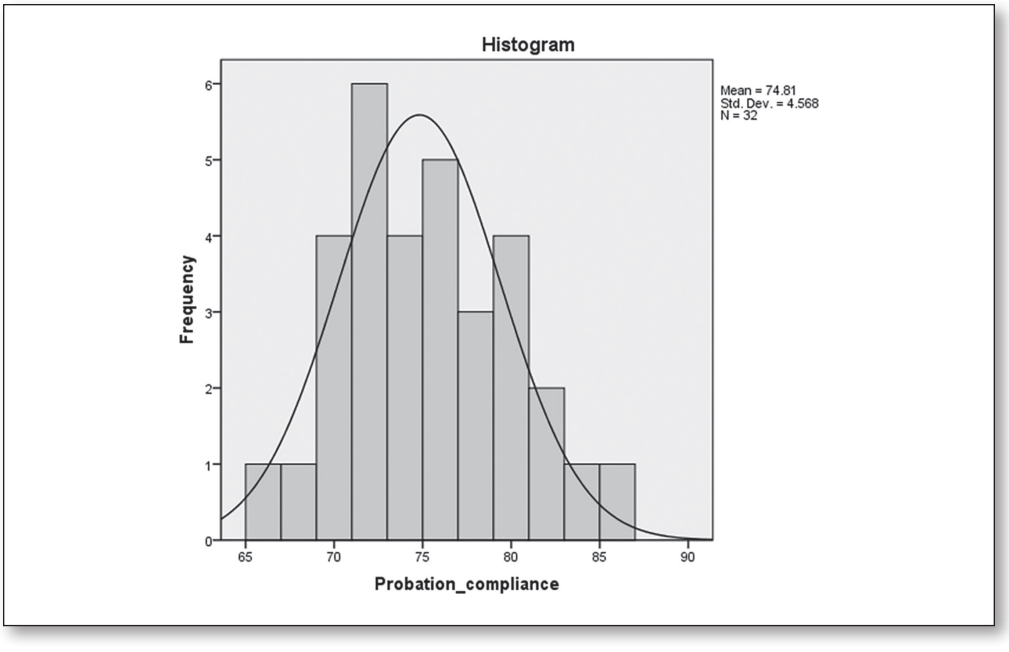
Normal distribution for *Probation_compliance* in Group 1 (No mentor)



Normal distribution for *Probation_compliance* in Group 2 (Peer mentor)



Normal distribution for *Probation_compliance* in Group 3 (Adult mentor)



Test of Homogeneity of Variances

Probation_compliance

Levene Statistic	df1	df2	Sig.
.475	2	93	.623

The homogeneity of variance score for *Probation_compliance* shows a significance (*p*) of .623; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances among the three groups, hence, this pretest criterion passes.

The *n* for the groups are 32 each (see *Descriptives* table below), which satisfies the 30 per group minimum criterion.

(c)

The ANOVA test revealed the following:

Descriptives

Probation_compliance

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
No mentor	32	74.22	5.046	.892	72.40	76.04	66	84
Peer mentor	32	75.22	5.440	.962	73.26	77.18	65	86
Adult mentor	32	74.81	4.568	.808	73.17	76.46	66	85
Total	96	74.75	4.995	.510	73.74	75.76	65	86

ANOVA

Probation_compliance

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16.188	2	8.094	.320	.727
Within Groups	2353.813	93	25.310		
Total	2370.000	95			

Multiple Comparisons

Probation_compliance
Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
No mentor	Peer mentor	-1.000	1.258	.707	-4.00	2.00
	Adult mentor	-.594	1.258	.885	-3.59	2.40
Peer mentor	No mentor	1.000	1.258	.707	-2.00	4.00
	Adult mentor	.406	1.258	.944	-2.59	3.40
Adult mentor	No mentor	.594	1.258	.885	-2.40	3.59
	Peer mentor	-.406	1.258	.944	-3.40	2.59

Groups (Probation Compliance)	*p*
No mentor (M = 74.22) : Peer mentor (M = 75.22)	.707
No mentor (M = 74.22) : Adult mentor (M = 74.81)	.885
Peer mentor (M = 75.22) : Adult mentor (M = 74.81)	.944

We discovered that youth offenders had higher probation compliance when paired with a peer mentor (M = 75.2, SD = 5.4), the next highest group was those who were paired with an adult (M = 74.8, SD = 4.5), followed by those who had no mentor (M = 74.2, SD = 5.0). No statistically significant differences were detected among any pairs of these groups; p values ranged from .707 to .944 ($\alpha = .05$), suggesting that mentorship has an insignificant impact when it comes to probation compliance within this population. As such, we would accept H_0 and reject H_1.

(d)

A judge appointed us to evaluate the effectiveness of a new mentorship program for juvenile offenders with priors. The 96 juveniles were randomly assigned to one of three groups: no mentor, a trained peer mentor who is 3 to 5 years older than the offender, or a trained adult mentor. The highest probation compliance was found among those paired with a peer mentor (M = 75.2, SD = 5.4), followed by the adult mentor group (M = 74.8, SD = 4.5), and those with no mentor had the lowest level of probation compliance (M = 74.2, SD = 5.0). No statistically significant differences were detected among any pairs of these groups; p values ranged from .707 to .944 ($\alpha = .05$), suggesting that mentorship has an insignificant impact when it comes to probation compliance within this population. As such, we would accept H_0 and reject H_1. Our intent is to continue our research and revise the mentor training protocol, and reassess our mentor recruitment protocol.

EXERCISE 5.5, DATA SET A

(a)

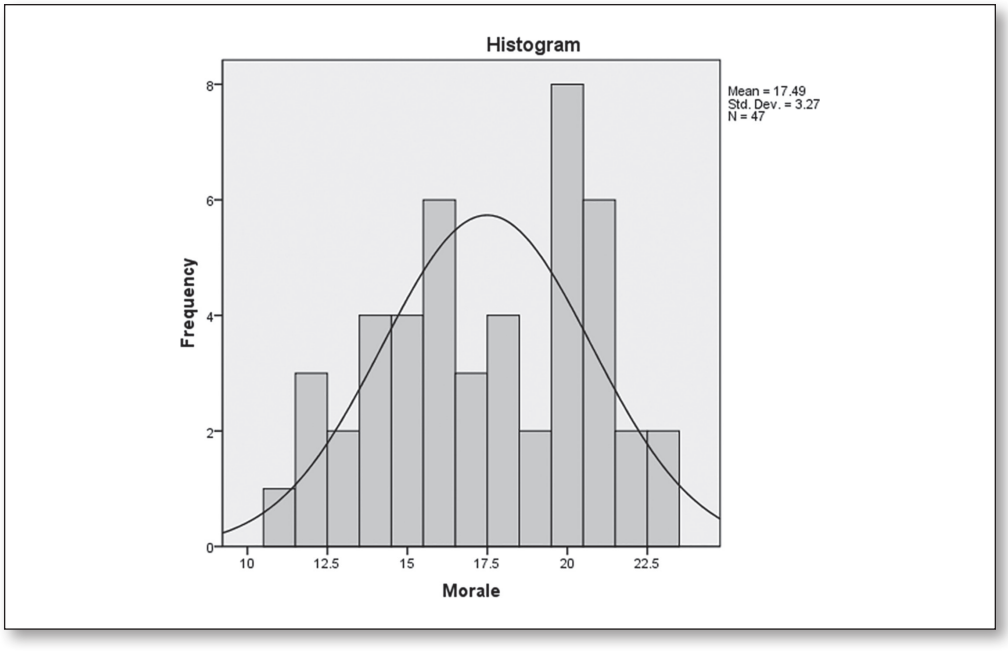
H_0: Increasing paid time off will not affect employee morale.

H_1: Increasing paid time off will affect employee morale.

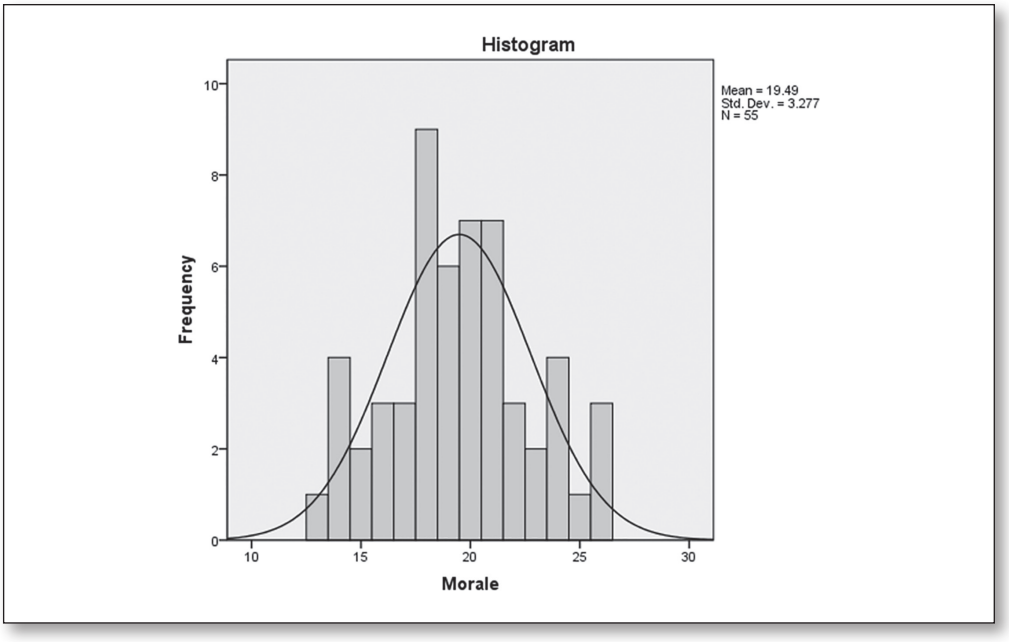
(b)

Histograms with normal curve plots show a normal distribution of *Morale* for all groups per the three figures below, hence, the pretest criterion of *normality* is satisfied.

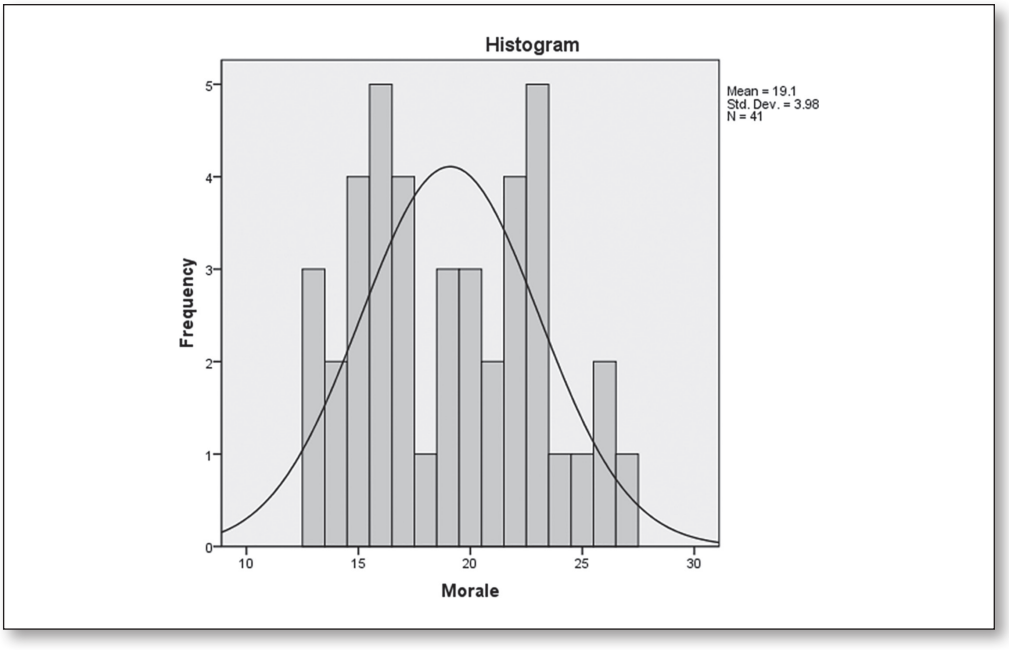
Site 1—2 Weeks PTO (Paid Time Off)



Site 2—2 Weeks PTO (Paid Time Off) + 4th Friday of the month off



Site 3—3 Weeks PTO (Paid Time Off)



Test of Homogeneity of Variances

Morale

Levene Statistic	df1	df2	Sig.
2.364	2	140	.098

The homogeneity of variance score shows a significance (*p*) of .098; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances of the groups; hence, this pretest criterion passes. The *n* for each group, as shown in the *Descriptives* table below, is over 30 for each group, which satisfies that criterion.

(c)

The ANOVA test revealed the following:

Descriptives

Morale

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
2 Weeks PTO	47	17.49	3.270	.477	16.53	18.45	11	23
2 Weeks PTO + 4th Fridays off	55	19.49	3.277	.442	18.61	20.38	13	26
3 Weeks PTO	41	19.10	3.980	.622	17.84	20.35	13	27
Total	143	18.72	3.575	.299	18.13	19.31	11	27

ANOVA

Morale

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	109.711	2	54.856	4.504	.013
Within Groups	1705.100	140	12.179		
Total	1814.811	142			

Multiple Comparisons

Morale

Sidak

(I) Site	(J) Site	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2 Weeks PTO	2 Weeks PTO + 4th Fridays off	-2.002*	.693	.013	-3.68	-.33
	3 Weeks PTO	-1.608	.746	.095	-3.41	.19
2 Weeks PTO + 4th Fridays off	2 Weeks PTO	2.002*	.693	.013	.33	3.68
	3 Weeks PTO	.393	.720	.929	-1.35	2.13
3 Weeks PTO	2 Weeks PTO	1.608	.746	.095	-.19	3.41
	2 Weeks PTO + 4th Fridays off	-.393	.720	.929	-2.13	1.35

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

Sites (Morale)	*p*
2 weeks PTO (M = 17.49) : 2 Weeks PTO + 4th Fridays off (M = 19.49)	.013*
2 weeks PTO (M = 17.49) : 3 weeks PTO (M = 19.10)	.095
2 Weeks PTO + 4th Fridays off (M = 19.49) : 3 weeks PTO (M = 19.10)	.929

*Statistically significant difference ($\alpha = .05$).

We discovered that employees working at a site which provides 2 weeks of PTO per year with the last Friday of the month off (with pay) had a statistically significantly higher morale (M = 19.49) compared to those at a different site that had 2 weeks of PTO per year (M = 17.49) (p = .013). We detected no statistically significant differences in morale between the site that received 2 weeks of PTO (M = 17.49) and the site that received 3 weeks of PTO (M = 19.10) (p = .095) or the site that had 2 weeks of PTO + the last Friday of the month off (M = 19.49) and the site that had 3 weeks of PTO off per year (M = 19.10) (p = .929). Since a difference was detected between one pair of sites, I would reject H_0 and accept H_1.

(d)

In order to assess methods of improving employee morale, we conducted a three-site study: The 47 employees at site 1 received the usual 2 weeks of PTO (Paid Time Off) per year scored a mean of 17.49 (SD = 3.27) on the Acme Morale Scale, wherein 1 = extremely low morale, and 25 = extremely high morale. The 55 employees at site 2 received the same 2 weeks of PTO per year plus the last Friday of each month off (with pay); their mean morale score was 19.49 (SD = 3.28). The 41 employees at site 3, who have 3 weeks of PTO per year had a mean morale score of 19.10 (SD = 3.98). Pairwise comparisons showed that employees who received 2 weeks of PTO plus every 4th Friday off had statistically significantly higher morale compared to those who received (only) 2 weeks of PTO (p = .013, $\alpha = .05$). No statistically significant differences were detected between the site that received 2 weeks of PTO and the site that received 3 weeks of PTO (p = .095). Additionally, no statistically significant differences were detected between the site that received 2 weeks of PTO plus the 4th Friday off and the site that received 3 weeks of PTO (p = .929). Based on these findings, we would reject H_0 in favor of H_1. Since we detected no statistically significant difference in morale between sites 2 and 3, we are considering changing site 2 to 3 weeks of PTO per year.

EXERCISE 5.5, DATA SET B

(a)

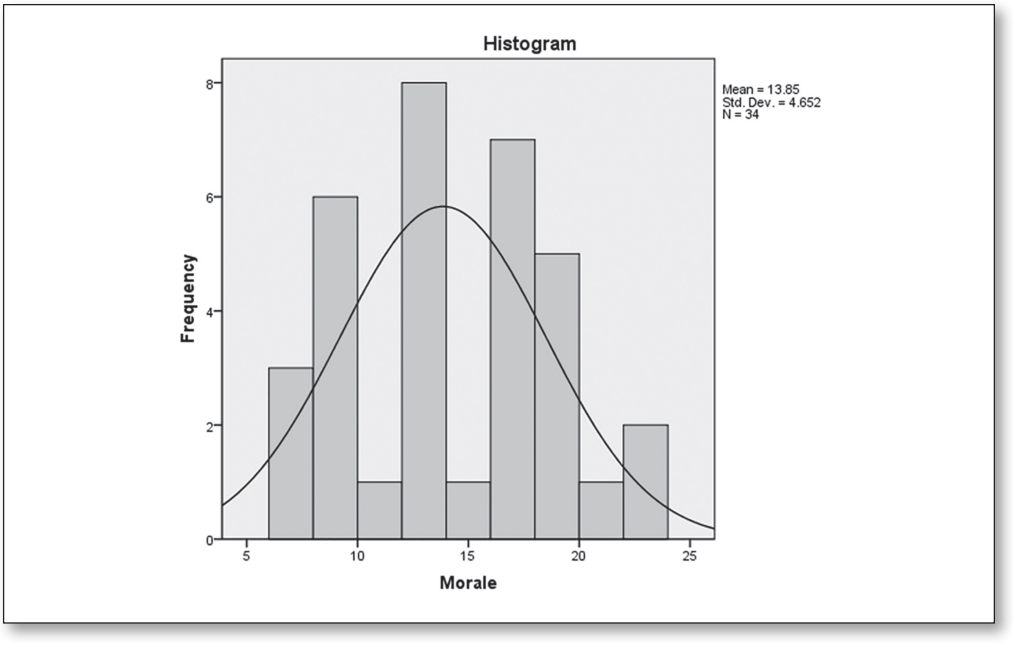
H_0: Increasing paid time off will not affect employee morale.

H_1: Increasing paid time off will affect employee morale.

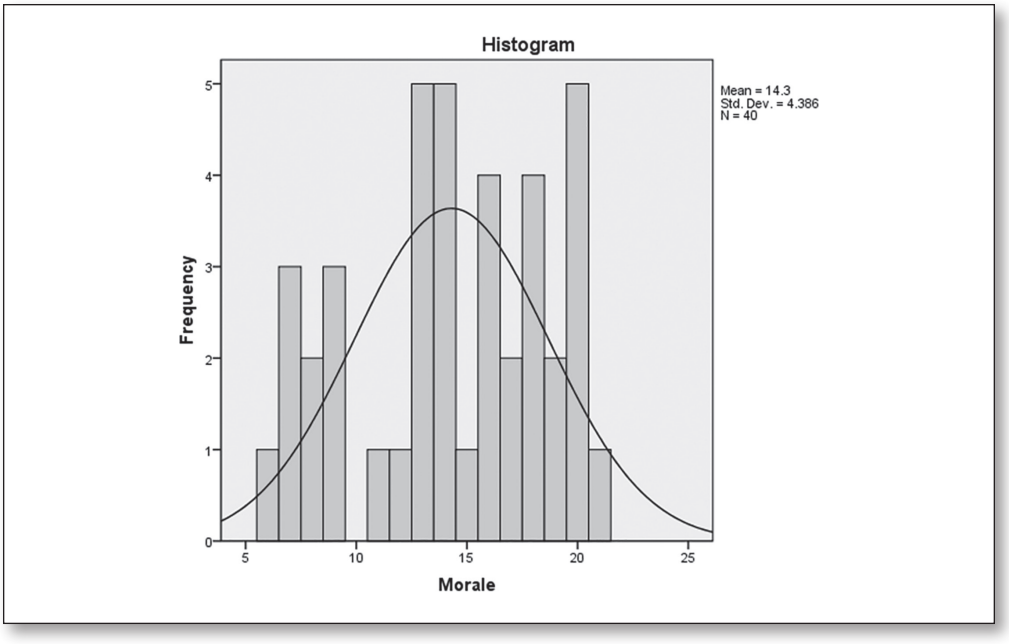
(b)

Histograms with normal curve plots show a normal distribution of *Morale* for all groups per the three figures below, hence, the pretest criterion of *normality* is satisfied.

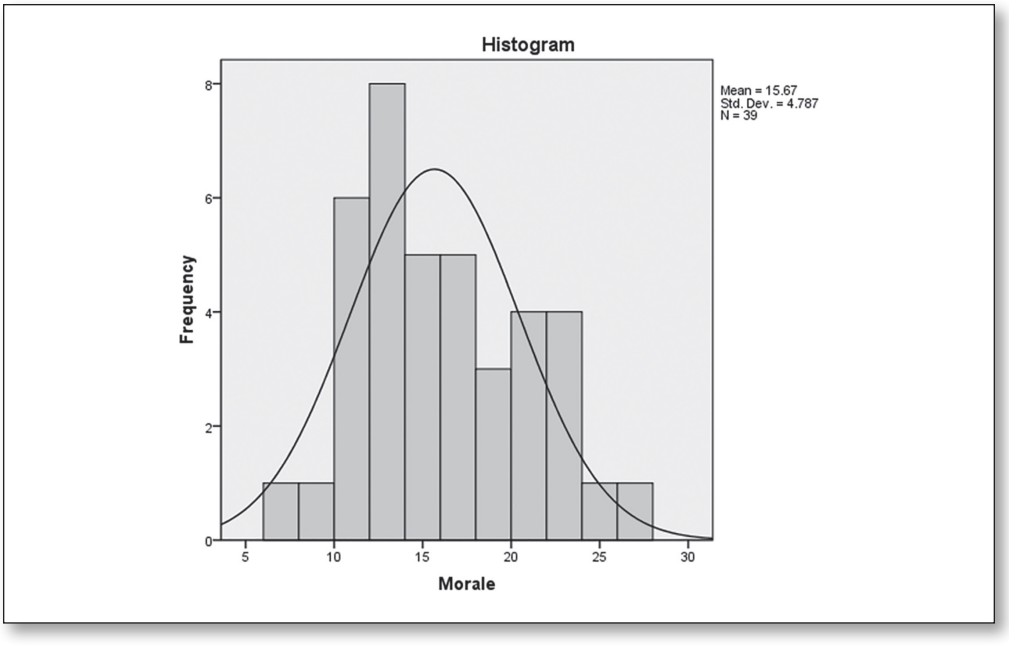
Site 1—2 Weeks PTO (Paid Time Off)



Site 2—2 Weeks PTO (Paid Time Off) + 4th Friday of the month off



Site 3—3 Weeks PTO (Paid Time Off)



Test of Homogeneity of Variances

Morale

Levene Statistic	df1	df2	Sig.
.259	2	110	.772

The homogeneity of variance score shows a significance (*p*) of .772; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances of the groups; hence, this pretest criterion passes. The *n* for each group, as shown in the *Descriptives* table below, is over 30 for each group, which satisfies that criterion.

(c)

The ANOVA test revealed the following:

Descriptives

Morale

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
2 Weeks PTO	34	13.85	4.652	.798	12.23	15.48	7	23
2 Weeks PTO + 4th Fridays off	40	14.30	4.386	.694	12.90	15.70	6	21
3 Weeks PTO	39	15.67	4.787	.766	14.12	17.22	7	26
Total	113	14.64	4.631	.436	13.77	15.50	6	26

ANOVA

Morale

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	66.793	2	33.396	1.573	.212
Within Groups	2335.331	110	21.230		
Total	2402.124	112			

Multiple Comparisons

Morale

Sidak

(I) Site	(J) Site	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2 Weeks PTO	2 Weeks PTO + 4th Fridays off	-.447	1.075	.967	-3.05	2.16
	3 Weeks PTO	-1.814	1.081	.262	-4.44	.81
2 Weeks PTO + 4th Fridays off	2 Weeks PTO	.447	1.075	.967	-2.16	3.05
	3 Weeks PTO	-1.367	1.037	.469	-3.88	1.15
3 Weeks PTO	2 Weeks PTO	1.814	1.081	.262	-.81	4.44
	2 Weeks PTO + 4th Fridays off	1.367	1.037	.469	-1.15	3.88

Sites (Morale)	*p*
2 weeks PTO (M = 13.85) : 2 Weeks PTO + 4th Fridays off (M = 14.30)	.967
2 weeks PTO (M = 13.85) : 3 weeks PTO (M = 15.67)	.262
2 Weeks PTO + 4th Fridays off (M = 14.30) : 3 weeks PTO (M = 15.67)	.469

*Statistically significant difference ($\alpha = .05$).

The highest mean employee morale was found at the site that gave 3 weeks of PTO (M = 15.67), followed by the site that gave 2 weeks of PTO plus the last Friday of each month off (M = 14.30), and the lowest mean morale was found at the site that received 2 weeks of PTO (M = 13.85). All of the pairwise comparisons produced p values between .262 and .967 ($\alpha = .05$); since no statistically significant differences were detected among these sites, we accept H_0 and reject H_1.

(d)

In order to assess the effect that paid time off (PTO) has on morale, we conducted a three-site study using the Acme Morale Scale (1 = extremely low morale . . . 25 = extremely high morale). The highest mean employee morale was found at the site that gave 3 weeks of PTO (M = 15.67, SD = 4.79), followed by the site that gave 2 weeks of PTO plus the last Friday of each month off (M = 14.30, SD = 4.39), and the lowest mean morale was found at the site that received 2 weeks of PTO (M = 13.85, SD = 4.65). All of the pairwise comparisons produced p values between .262 and .967 ($\alpha = .05$): 2 weeks PTO : 2 weeks PTO with every 4th Friday off (p = .967), 2 weeks PTO : 3 weeks PTO (p = .292), and 2 weeks PTO with every 4th Friday off : 3 weeks PTO (p = .469). Since we detected no statistically significant differences among these sites, we accept H_0 and reject H_1.

EXERCISE 5.7, DATA SET A

(a)

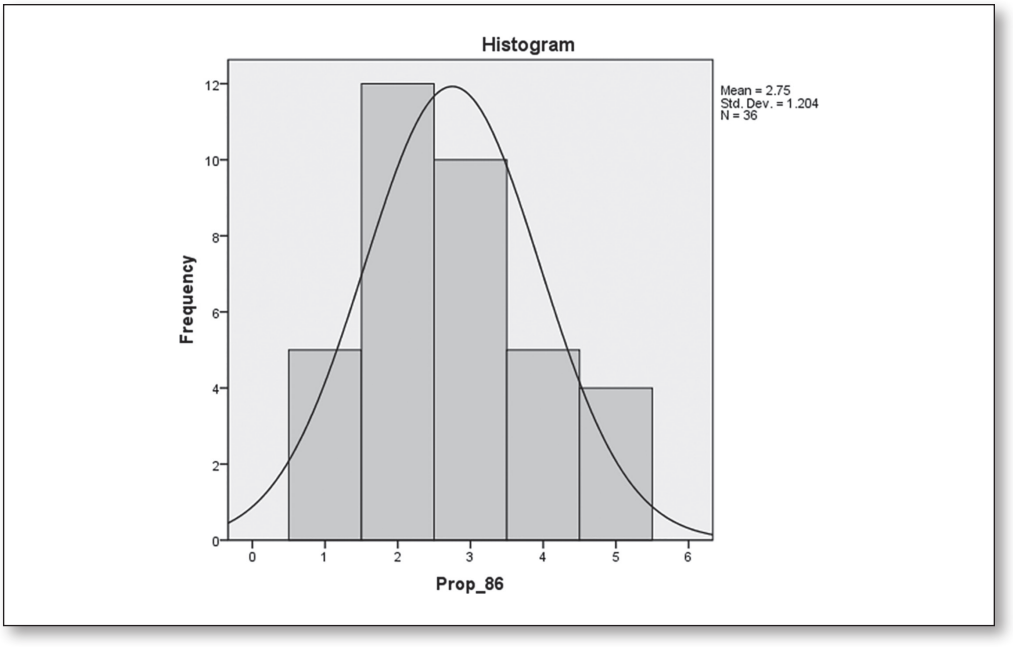
H_0: Advertising media will not influence voter choice.

H_1: Advertising media will influence voter choice.

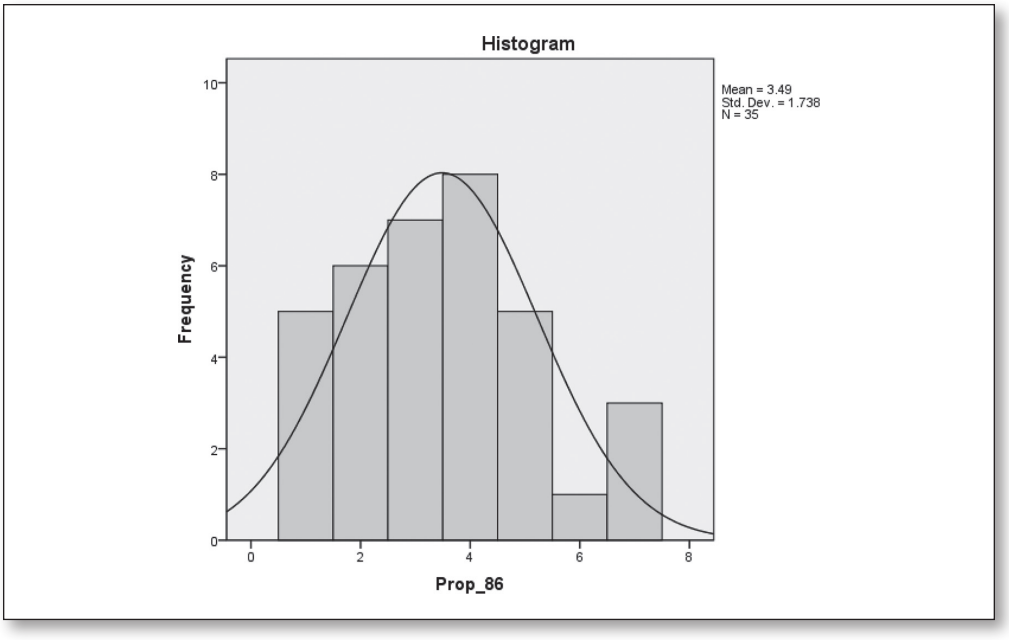
(b)

Histograms with normal curve plots show a normal distribution of *Prop_86* for the groups as shown in the three figures below, hence, the pretest criterion of *normality* is satisfied.

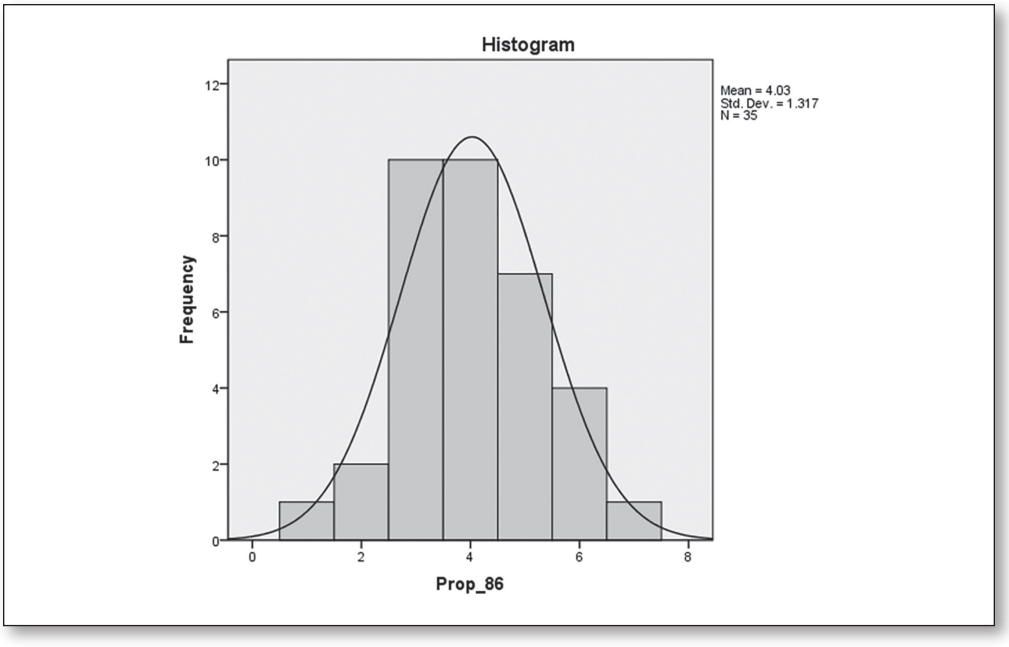
Normal distribution for *Prop_86* in Group 1 (Control group)



Normal distribution for *Prop_86* in Group 2 (Print group)



Normal distribution for *Prop_86* in Group 3 (Video group)



Test of Homogeneity of Variances

Prop_86

Levene Statistic	df1	df2	Sig.
2.921	2	103	.058

The homogeneity of variance score for *mood* shows a significance (*p*) of .058; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances of the three groups, hence, this pretest criterion passes.

The *n* for the three groups are 36, 35, and 35, which satisfies the 30 per group minimum criterion (see *Descriptives* table below).

(c)

The ANOVA test revealed the following:

Descriptives

Prop_86

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control	36	2.75	1.204	.201	2.34	3.16	1	5
Print	35	3.49	1.738	.294	2.89	4.08	1	7
Video	35	4.03	1.317	.223	3.58	4.48	1	7
Total	106	3.42	1.517	.147	3.12	3.71	1	7

ANOVA

Prop_86

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	29.272	2	14.636	7.095	.001
Within Groups	212.464	103	2.063		
Total	241.736	105			

Multiple Comparisons

Prop_86
Sidak

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control	Print	-.736	.341	.096	-1.56	.09
	Video	-1.279*	.341	.001	-2.11	-.45
Print	Control	.736	.341	.096	-.09	1.56
	Video	-.543	.343	.311	-1.38	.29
Video	Control	1.279*	.341	.001	.45	2.11
	Print	.543	.343	.311	-.29	1.38

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

Groups (Prop 86)	*p*
Control (M = 2.75) : Print (M = 3.49)	.096
Control (M = 2.75) : Video (M = 4.03)	.001*
Print (M = 3.49) : Video (M = 4.03)	.311

*Statistically significant difference ($\alpha = .05$).

The mean voter likelihood level for those in the Video (advertisement) group was 4.03 which is significantly higher (p = .001) than the mean for the control group (2.75). The other comparisons among the groups were statistically insignificant. Since there was one group that statistically significantly outperformed another, I would reject H_0 and accept H_1.

(d)

In order to determine the most persuasive form of advertisement to encourage people to vote yes on Proposition 86, we convened focus groups consisting of registered voters. We recruited 106 participants and randomly assigned them to one of three media groups control (no media), print advertisement, and video advertisement. Prior to dismissing the participants, each was asked to indicate the likelihood that they would vote "yes" on proposition 86 using a 1 to 7 scale (1 = will definitely vote no . . . 7 = will definitely vote yes). The video (M = 4.03, SD = 1.32) significantly outperformed the control group (M = 2.75, SD = 1.20) (p = .001, $\alpha = .05$). Although the video group produced a higher mean than the print group (M = 3.94, SD = 1.74), the difference between the groups is insignificant (p = .311, $\alpha = .05$). Additionally, there was no statistically significant difference detected between the control group and the print group (p = .096, $\alpha = .05$). Based on these findings, we reject H_0 and accept H_1. Considering that there is no statistically significant difference between the video and the print version of the ad, if video is not accessible or affordable, the print version could be used.

EXERCISE 5.7, DATA SET B

(a)

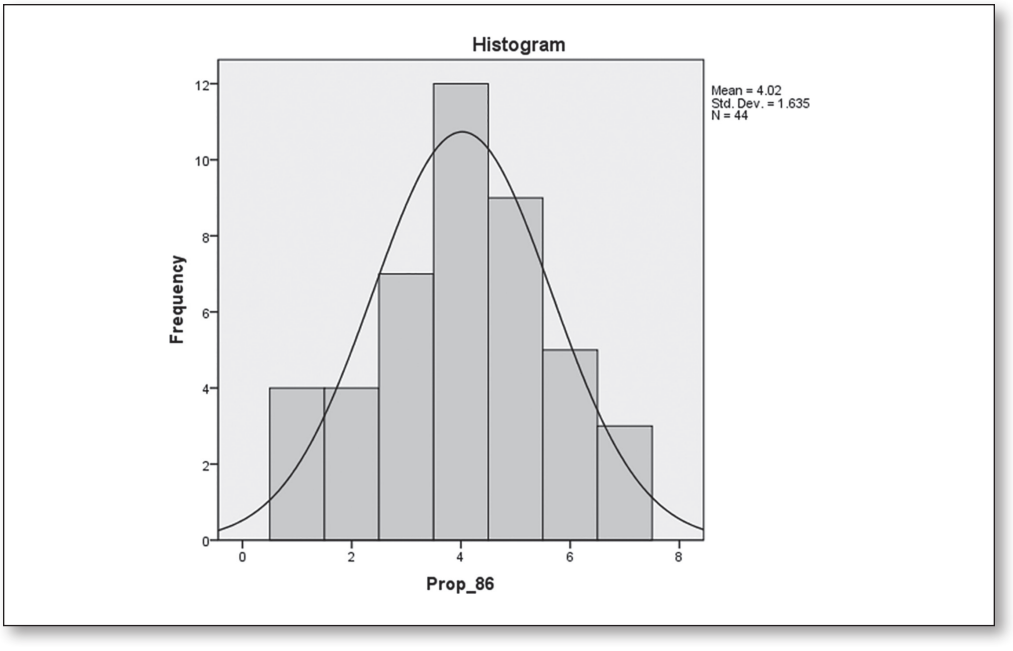
H_0: Advertising media will not influence voter choice.

H_1: Advertising media will influence voter choice.

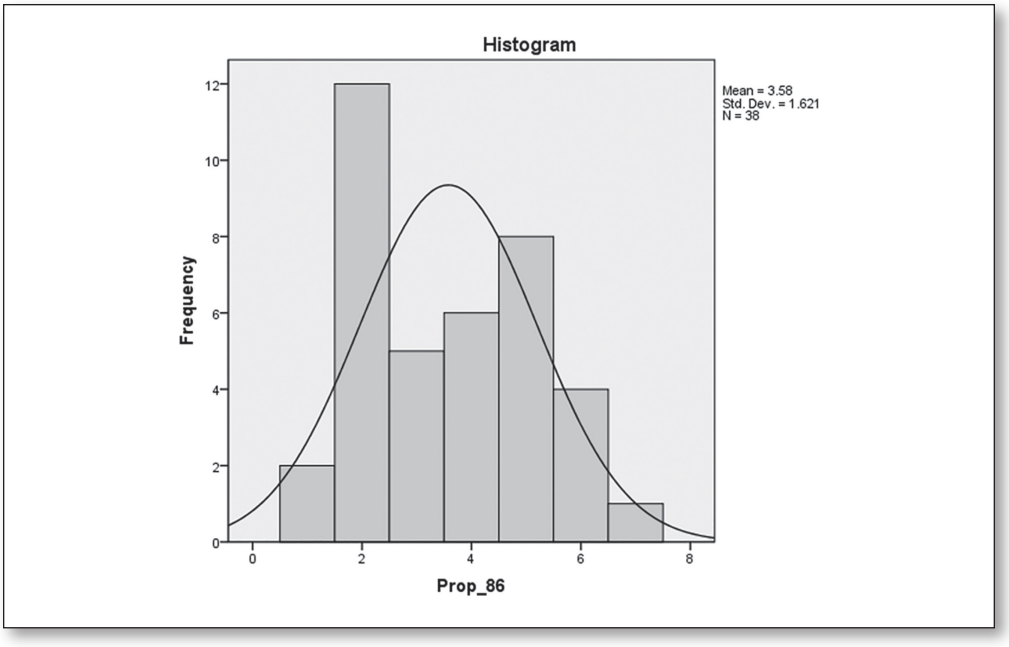
(b)

Histograms with normal curve plots show a normal distribution of *Prop_86* for the groups as shown in the three figures below, hence, the pretest criterion of *normality* is satisfied.

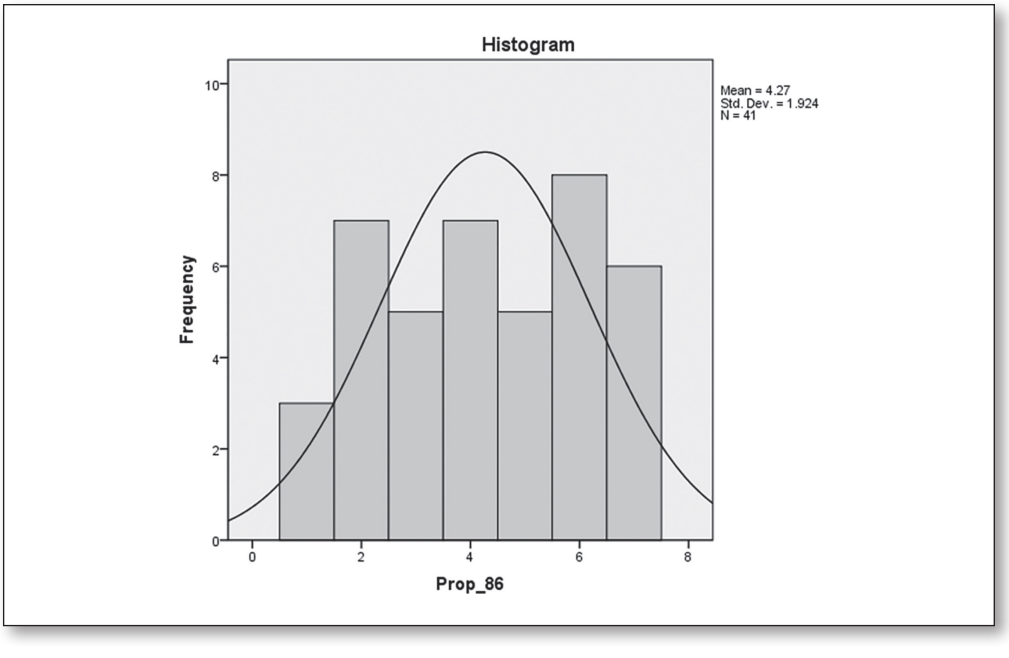
Normal distribution for *Prop_86* in Group 1 (Control group)



Normal distribution for *Prop_86* in Group 2 (Print group)



Normal distribution for *Prop_86* in Group 3 (Video group)



Test of Homogeneity of Variances

Prop_86

Levene Statistic	df1	df2	Sig.
1.992	2	120	.141

The homogeneity of variance score for *mood* shows a significance (*p*) of .141; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances of the three groups, hence, this pretest criterion passes.

The *n* for the three groups are 44, 38, and 41, which satisfies the 30 per group minimum criterion (see *Descriptives* table below).

(c)

The ANOVA test revealed the following:

Descriptives

Prop_86

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control	44	4.02	1.635	.247	3.53	4.52	1	7
Print	38	3.58	1.621	.263	3.05	4.11	1	7
Video	41	4.27	1.924	.300	3.66	4.88	1	7
Total	123	3.97	1.741	.157	3.66	4.28	1	7

ANOVA

Prop_86

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9.581	2	4.790	1.596	.207
Within Groups	360.289	120	3.002		
Total	369.870	122			

Multiple Comparisons

Prop_86
Sidak

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control	Print	.444	.384	.578	-.49	1.37
	Video	-.246	.376	.886	-1.16	.67
Print	Control	-.444	.384	.578	-1.37	.49
	Video	-.689	.390	.221	-1.63	.26
Video	Control	.246	.376	.886	-.67	1.16
	Print	.689	.390	.221	-.26	1.63

Groups (Prop 86)	*p*
Control (M = 4.02) : Print (M = 3.58)	.578
Control (M = 4.02) : Video (M = 4.27)	.886
Print (M = 3.58) : Video (M = 4.27)	.221

These findings indicate that there are no statistically significant differences in voter likelihood among the three groups; the means range from 3.58 to 4.27 and the p values range from .221 to .886. Since none of the pairwise comparisons produced a p value less than .05, I would accept H_0 and reject H_1.

(d)

In order to determine the most persuasive form of advertisement to encourage people to vote yes on Proposition 86, we convened focus groups consisting of registered voters. We recruited 123 participants and randomly assigned them to one of three media groups control (no media), print advertisement, and video advertisement. Prior to dismissing the participants, each was asked to indicate the likelihood that they would vote "yes" on proposition 86 using a 1 to 7 scale (1 = will definitely vote no . . . 7 = will definitely vote yes). Pairwise assessments of the control group (M = 4.02, SD = 1.64), print group (M = 3.58, SD = 1.62), and video group (M = 4.27, SD = 1.92) produced *p* values between ranging from .221 to .886, indicating that no statistically significant differences were detected among these groups ($\alpha = .05$). Based on these findings, we accept H_0 and reject H_1. These findings indicate that these print and video advertisements were essentially ineffective when it comes to prompting a change in voter opinion on Proposition 86. Further, it is notable that the print advertisement produced a lower score than the control group, which reviewed no advertisement(s).

EXERCISE 5.9, DATA SET A

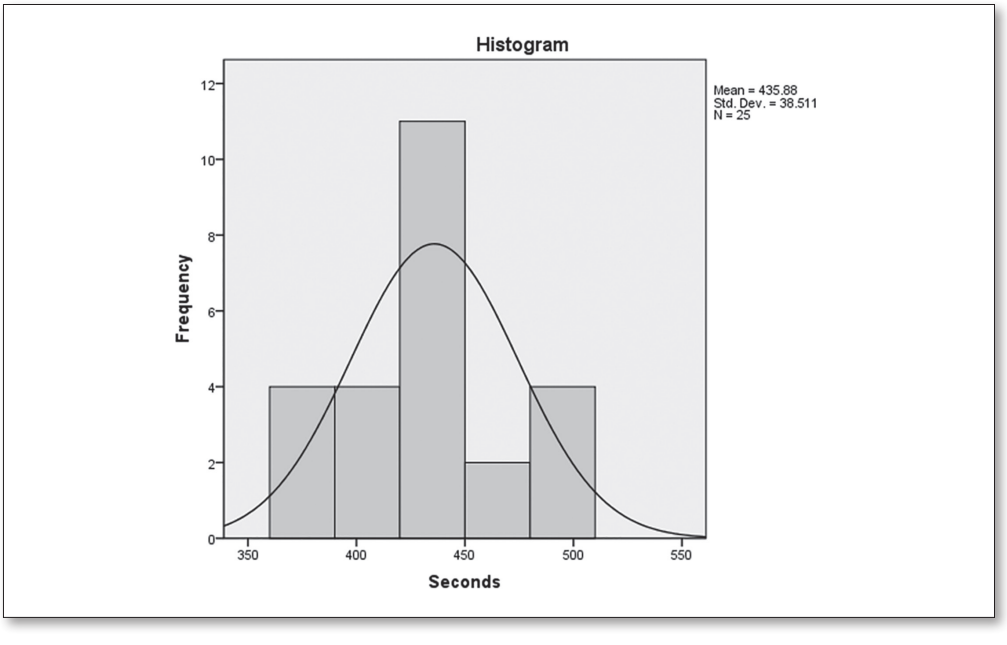
(a)

H_0: Lighting source has no effect on reading rate.

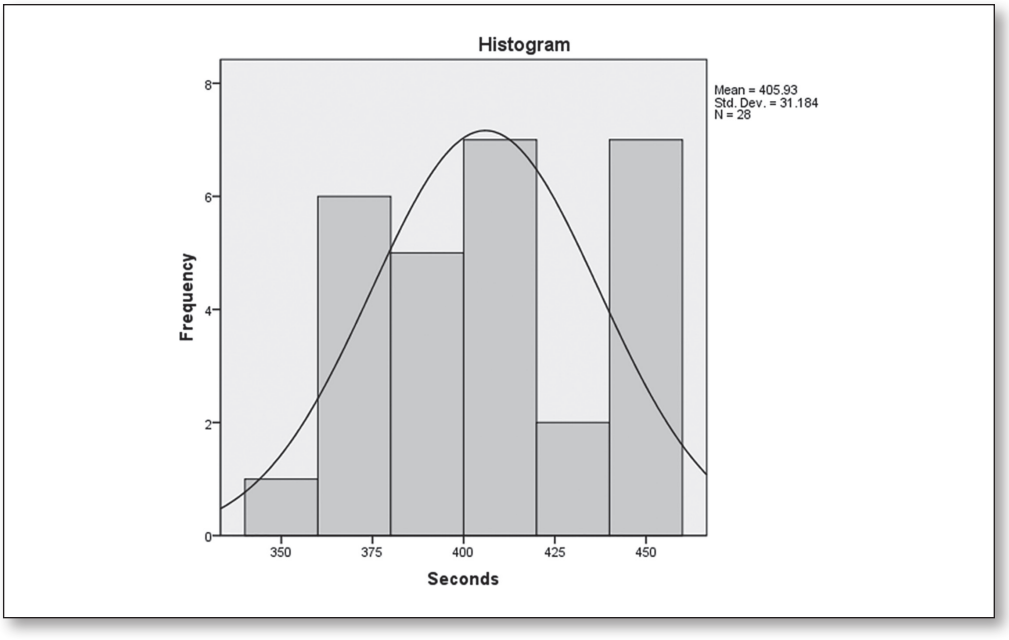
H_1: Lighting source has an effect on reading rate.

(b) Histograms with normal curve plots show a normal distribution of *Seconds* for all groups as shown in the four figures below, hence, the pretest criterion of *normality* is satisfied.

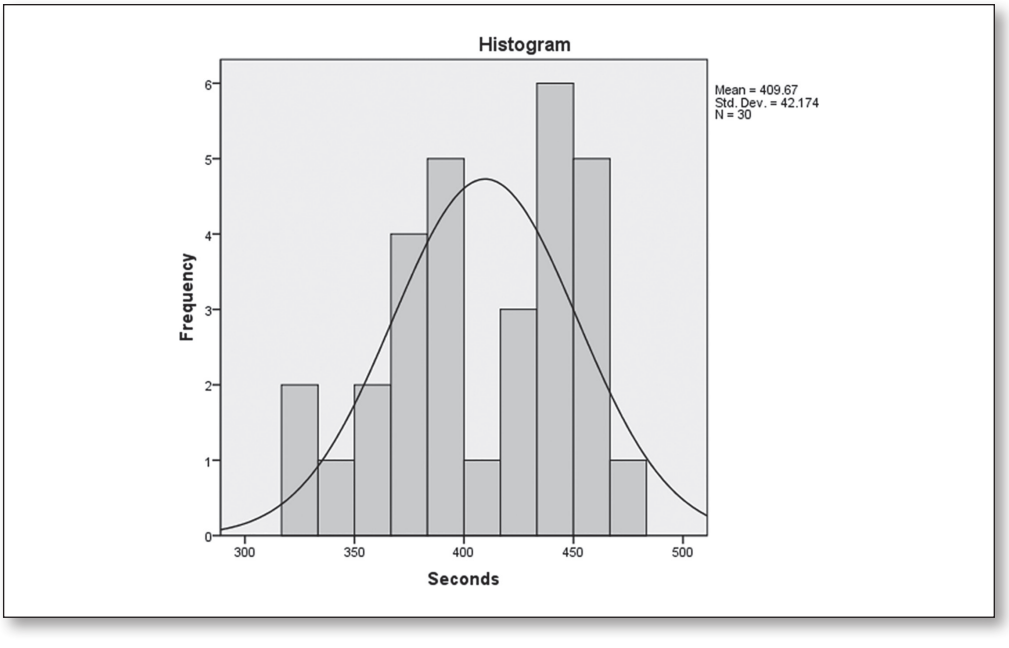
Normal distribution for *Seconds* in Group 1 (Room lighting)



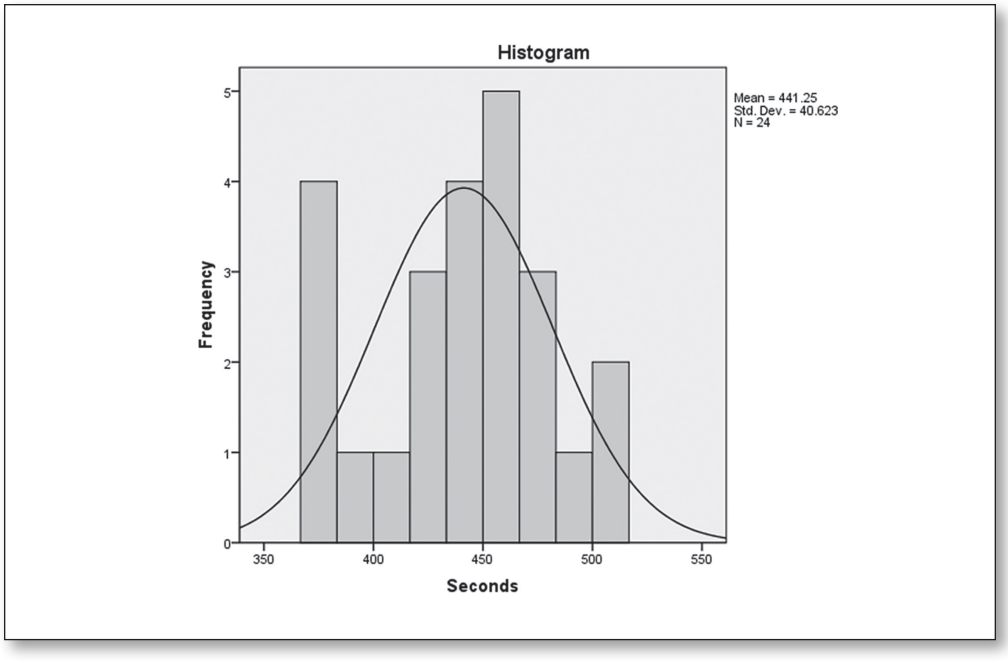
Normal distribution for *Seconds* in Group 2 (Acme reading lamp)



Normal distribution for *Seconds* in Group 3 (Generic reading lamp)



Normal distribution for *Seconds* in Group 4 (Flashlight)



Test of Homogeneity of Variances

Seconds

Levene Statistic	df1	df2	Sig.
1.499	3	103	.219

The homogeneity of variance score for *seconds* shows a significance (*p*) of .219; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances among the four groups, hence, this pretest criterion passes.

The Generic lamp group was the only group that had the minimal *n* of 30. The Room lighting, Acme lighting, and Flashlight groups had *n*s of 25, 28 and 24 respectively (see *Descriptives* table below). The findings of the ANOVA test would be more robust if the *n*s were slightly higher for these three groups.

(c) The ANOVA test revealed the following:

Descriptives

Seconds

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Room lighting	25	435.88	38.511	7.702	419.98	451.78	374	509
Acme lamp	28	405.93	31.184	5.893	393.84	418.02	357	455
Generic lamp	30	409.67	42.174	7.700	393.92	425.41	328	470
Flashlight	24	441.25	40.623	8.292	424.10	458.40	368	512
Total	107	421.90	40.851	3.949	414.07	429.73	328	512

ANOVA

Seconds

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	25504.205	3	8501.402	5.784	.001
Within Groups	151385.664	103	1469.764		
Total	176889.869	106			

Multiple Comparisons

Seconds
Sidak

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Room lighting	Acme lamp	29.951*	10.549	.032	1.65	58.25
	Generic lamp	26.213	10.382	.076	-1.64	54.06
	Flashlight	-5.370	10.956	.997	-34.76	24.02
Acme lamp	Room lighting	-29.951*	10.549	.032	-58.25	-1.65
	Generic lamp	-3.738	10.074	.999	-30.76	23.29
	Flashlight	-35.321*	10.665	.008	-63.93	-6.71
Generic lamp	Room lighting	-26.213	10.382	.076	-54.06	1.64
	Acme lamp	3.738	10.074	.999	-23.29	30.76
	Flashlight	-31.583*	10.499	.020	-59.75	-3.42
Flashlight	Room lighting	5.370	10.956	.997	-24.02	34.76
	Acme lamp	35.321*	10.665	.008	6.71	63.93
	Generic lamp	31.583*	10.499	.020	3.42	59.75

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

Groups (Reading Time in Seconds)	*p*
Room lighting (M = 436) : Acme Lamp (M = 406)	.032*
Room lighting (M = 436) : Generic lamp (M = 410)	.076
Room lighting (M = 436) : Flashlight (M = 441)	.997
Acme lamp (M = 406) : Generic lamp (M = 410)	.999
Acme lamp (M = 406) : Flashlight (M = 441)	.008*
Generic lamp (M = 410) : Flashlight (M = 441)	.020*

Means rounded to nearest second.

*Statistically significant difference ($\alpha = .05$).

Per the table above, since the mean reading time in the Acme reading lamp group is statistically significantly lower than scores of those who read using Room lighting and by Flashlight, we reject H_0. For the same reason, we would not reject H_1.

(d)

This study analyzed the effects that the Acme reading lamp had on reading speed. The 107 participants were randomly assigned to one of four groups; one group read a 1,000-word essay using regular room lighting, the second group read the same essay using the new Acme reading lamp, the third group read using a generic reading lamp, and the fourth group read using a flashlight. Results revealed that on the average, those who read using the Acme reading lamp read significantly faster (M = 406) than those who read using a flashlight (M = 441) (p = .008, α = .05), or regular room lighting (M = 436) (p = .032, α = .05)) using an α level of .05. Incidentally, those who used a generic reading lamp (M = 410) finished reading the essay significantly faster than those who read by flashlight (M = 441) (p = .020, α = .05). We also discovered that those who read using a Acme reading lamp (M = 406) completed the essay faster than those who used the generic reading lamp (M = 410), however this difference was not found to be statistically significant (p = .999, α = .05).

EXERCISE 5.9, DATA SET B

(a)

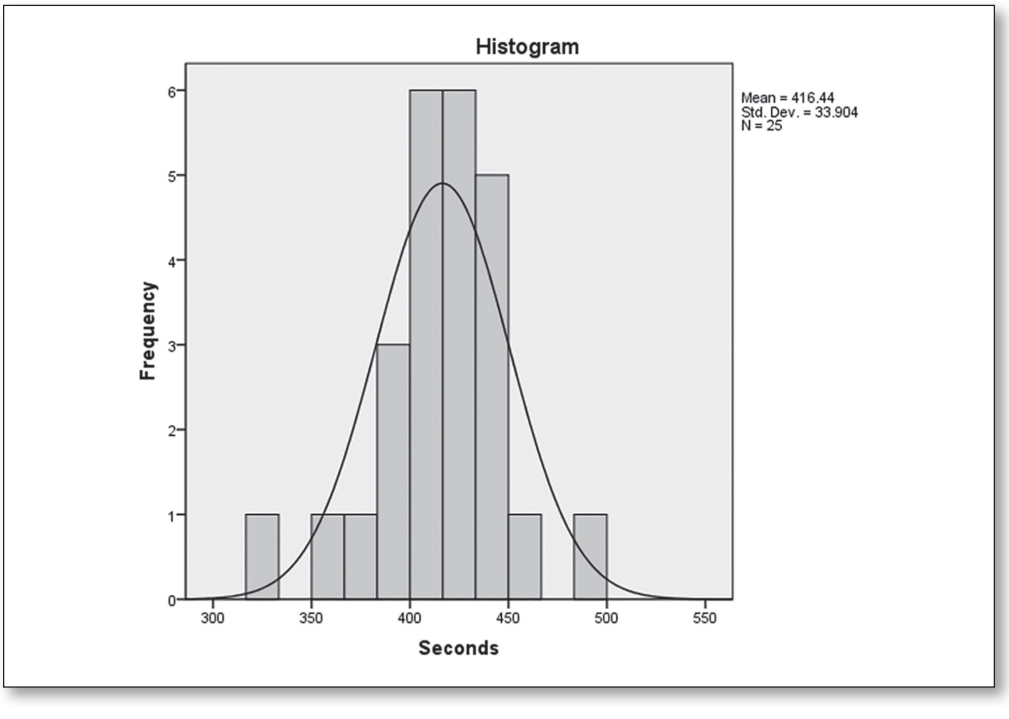
H_0: Lighting source has no effect on reading rate.

H_1: Lighting source has an effect on reading rate.

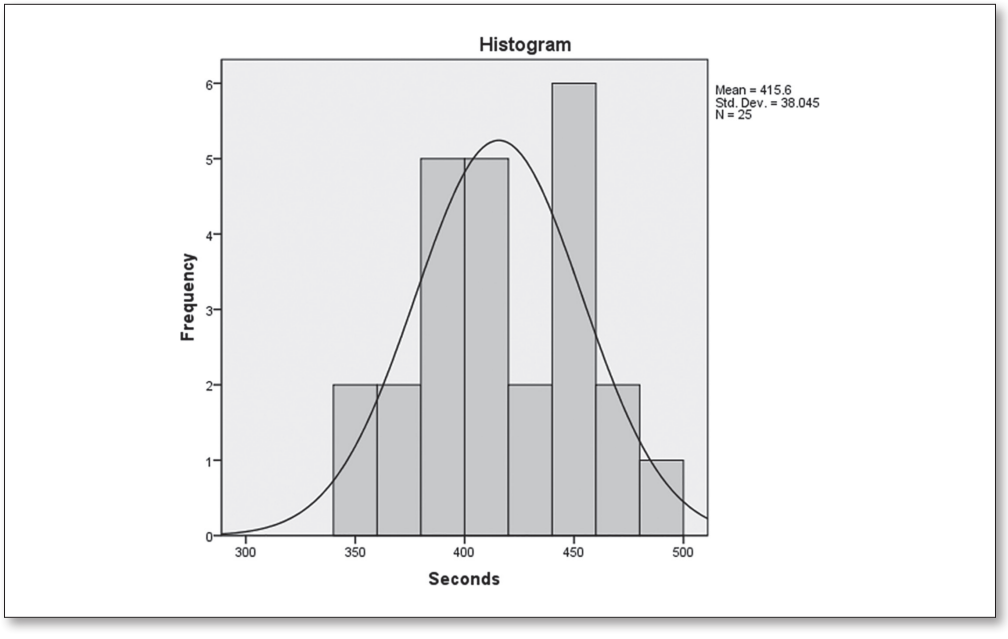
(b)

Histograms with normal curve plots show a normal distribution of *seconds* for all groups as shown in the four figures below, hence, the pretest criterion of *normality* is satisfied.

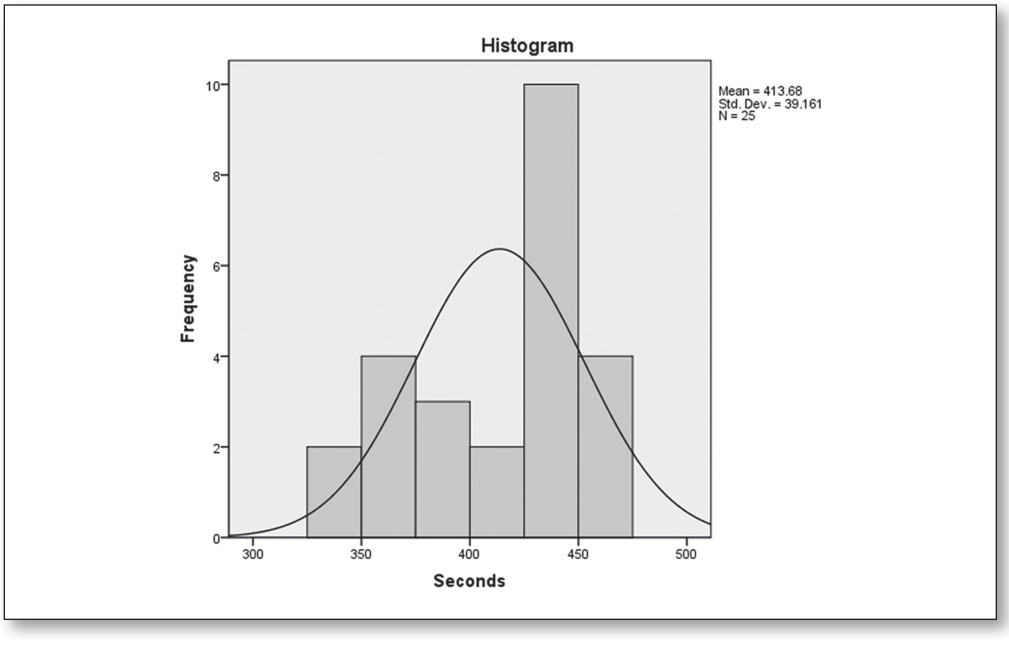
Normal distribution for *Seconds* in Group 1 (Room lighting)



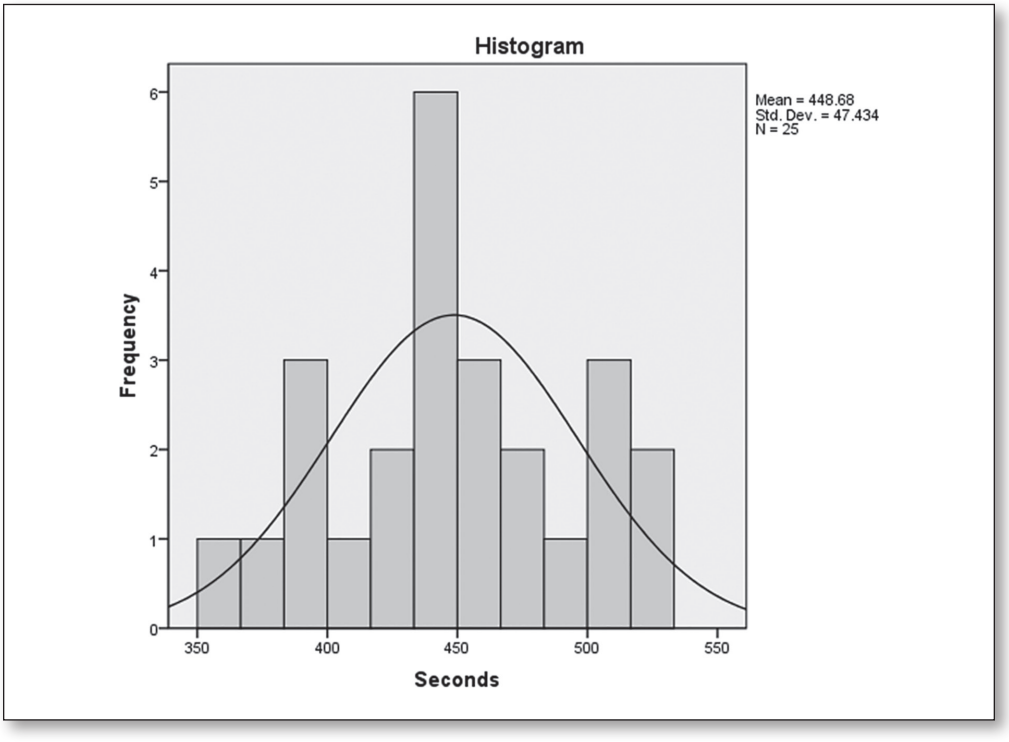
Normal distribution for *Seconds* in Group 2 (Acme reading lamp)



Normal distribution for *Seconds* in Group 3 (Generic reading lamp)



Normal distribution for *seconds* in Group 4 (Flashlight)



Test of Homogeneity of Variances

seconds

Levene Statistic	df1	df2	Sig.
1.163	3	96	.328

The homogeneity of variance score for *seconds* shows a significance (*p*) of .328; since this is greater than the α level of .05, this suggests that there is no statistically significant difference between the variances among the four groups, hence, this pretest criterion passes.

The groups each had an *n* of 25 (see *Descriptives* table below). The findings of the ANOVA test would be more robust if the *n*s were at least 30 per group.

(c)

The ANOVA test revealed the following:

Descriptives

seconds

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Room lighting	25	416.44	33.904	6.781	402.44	430.44	331	489
Acme lamp	25	415.60	38.045	7.609	399.90	431.30	344	481
Generic lamp	25	413.68	39.161	7.832	397.52	429.84	338	470
Flashlight	25	448.68	47.434	9.487	429.10	468.26	361	525
Total	100	423.60	41.947	4.195	415.28	431.92	331	525

ANOVA

seconds

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21066.960	3	7022.320	4.402	.006
Within Groups	153131.040	96	1595.115		
Total	174198.000	99			

Multiple Comparisons

seconds
Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Room lighting	Acme lamp	.840	11.296	1.000	-28.70	30.38
	Generic lamp	2.760	11.296	.995	-26.78	32.30
	Flashlight	-32.240*	11.296	.027	-61.78	-2.70
Acme lamp	Room lighting	-.840	11.296	1.000	-30.38	28.70
	Generic lamp	1.920	11.296	.998	-27.62	31.46
	Flashlight	-33.080*	11.296	.022	-62.62	-3.54
Generic lamp	Room lighting	-2.760	11.296	.995	-32.30	26.78
	Acme lamp	-1.920	11.296	.998	-31.46	27.62
	Flashlight	-35.000*	11.296	.013	-64.54	-5.46
Flashlight	Room lighting	32.240	11.296	.027	2.70	61.78
	Acme lamp	33.080*	11.296	.022	3.54	62.62
	Generic lamp	35.000*	11.296	.013	5.46	64.54

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

Groups (Reading Time in Seconds)	*p*
Room lighting (M = 416) : Acme Lamp (M = 416)	1.000
Room lighting (M = 416) : Generic lamp (M = 414)	.995
Room lighting (M = 416) : Flashlight (M = 449)	.027*
Acme lamp (M = 416) : Generic lamp (M = 414)	.998
Acme lamp (M = 416) : Flashlight (M = 449)	.022*
Generic lamp (M = 414) : Flashlight (M = 449)	.013*

Means rounded to nearest second.

*Statistically significant difference ($\alpha = .05$).

Per the table above, since the mean reading time in the Acme lamp group (M = 416) is statistically significantly lower than those who read using a Flashlight (M = 449) (p = .022, α = .05), based on the .05 α level, we reject H_0 and not reject H_1. Additionally, Room lighting (M = 416) statistically significantly outperformed the reading rate of the Flashlight (M = 449) (p = .027, α = .05), and finally, the reading rate for the generic lamp (M = 416) outperformed the Flashlight (p = .013, α = .05)).

(d)

This study analyzed the effects that the Acme reading lamp had on reading speed. The 100 participants were randomly assigned to one of four groups; one group read a 1,000-word essay using regular room lighting, the second group read the same essay using the new Acme reading lamp, the third group read using a generic reading lamp, and the fourth group read using a flashlight. There was no statistically significant difference in reading times among those who used the Acme reading lamp (M = 416), Room lighting (M = 416) or the Generic lamp (M = 414) using an α level of .05. All three of those groups read statistically significantly faster than the fourth group, who read using a flashlight (M = 449); p values ranged from .013 to .027.