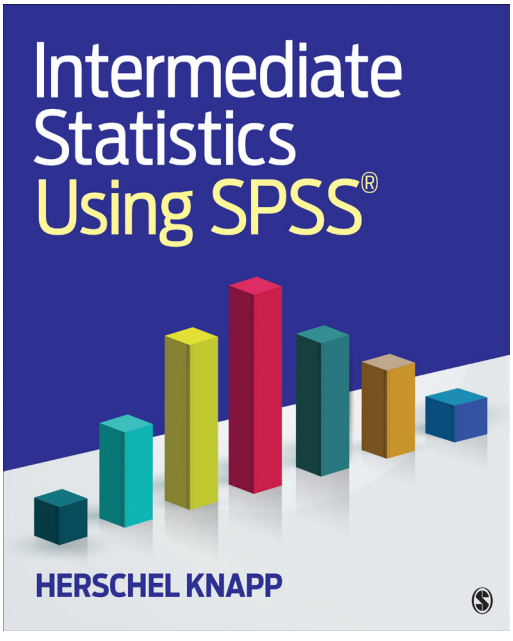


CHAPTER 11

Correlation and Regression — Pearson and Spearman

Solutions to Odd-Numbered Exercises

EXERCISE 11.1, DATA SET A

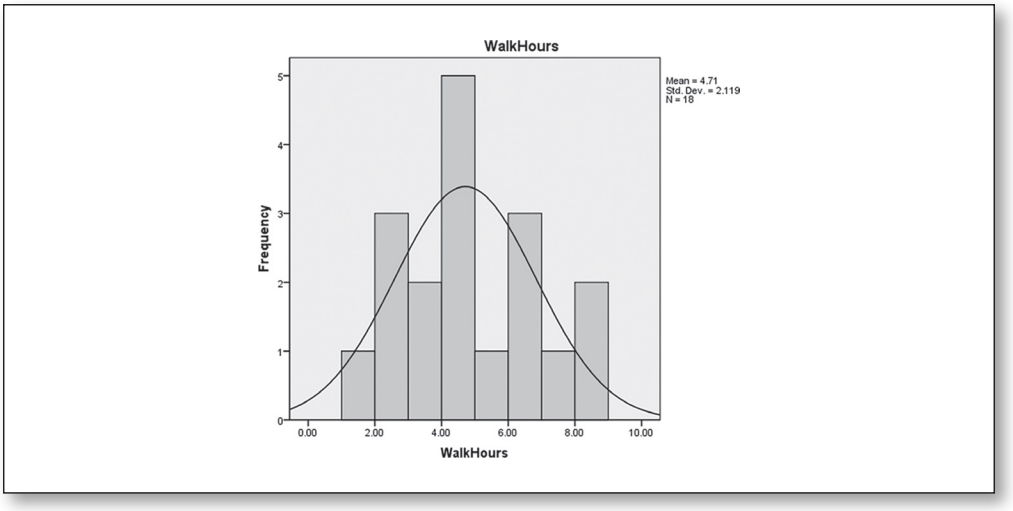
(a)

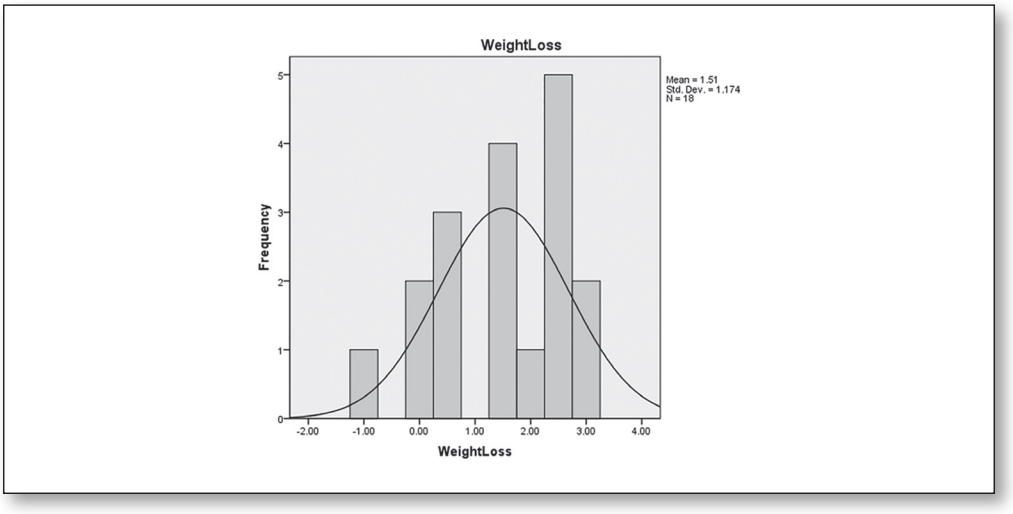
H_0: There is no significant correlation between hours of walking and weight loss

H_1: There is a significant correlation between hours of walking and weight loss.

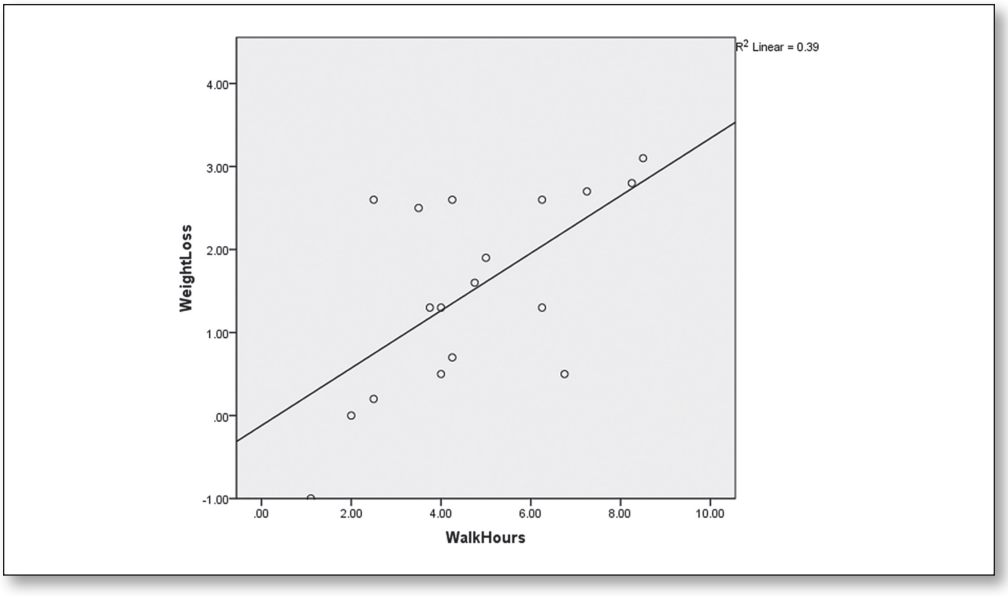
(b)

Histograms with normal curve plots show a normal distribution for *WalkHours* and *WeightLoss* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The scatterplot below shows that the straight regression line reasonably fits the field of points, hence the criterion of linearity is satisfied. The scatterplot also shows that the points are primarily clustered toward the center of the regression line with substantially fewer points toward the ends, hence the criterion of homoscedasticity is satisfied.



NOTE: Scatterplot with regression line for *WalkHours* and *WeightLoss* is linear and homoscedastic (pronounced: *hoe-moe-skuh-daz-tick*).

(c)

The correlations table below shows a positive correlation (r = .624) between *WalkHours* and *WeightLoss*; p = .006, which is less than the specified .05 α level, indicating that this is a statistically significant correlation.

Correlations

		WalkHours	WeightLoss
WalkHours	Pearson Correlation	1	.624**
	Sig. (2-tailed)		.006
	N	18	18
WeightLoss	Pearson Correlation	.624**	1
	Sig. (2-tailed)	.006	
	N	18	18

**. Correlation is significant at the 0.01 level (2-tailed).

(d)

In a study examining the effects of brisk walking on weight loss, 18 adults were recruited and instructed to walk as many days of the week they could for as long as they could. The researcher gathered each participant's daily walking time, as well as their weight at the beginning and end of the week and found a positive, statistically significant correlation between the amount of walking time and the amount of weight lost ($r = .624$, $p = .006$, $\alpha = .05$). In other words, the more the participants walked, the more weight they lost. These findings suggest that moderate daily exercise may be effective in facilitating weight loss. As such, we reject H_0, and do not reject H_1.

NOTE: You might also inspect the descriptive statistics for *WalkHrs* and W*tLoss*, and consider including the supplement in your findings: "Participants averaged 4.1 hours of walking per week and had an average weight loss of 1.5 pounds."

EXERCISE 11.1, DATA SET B

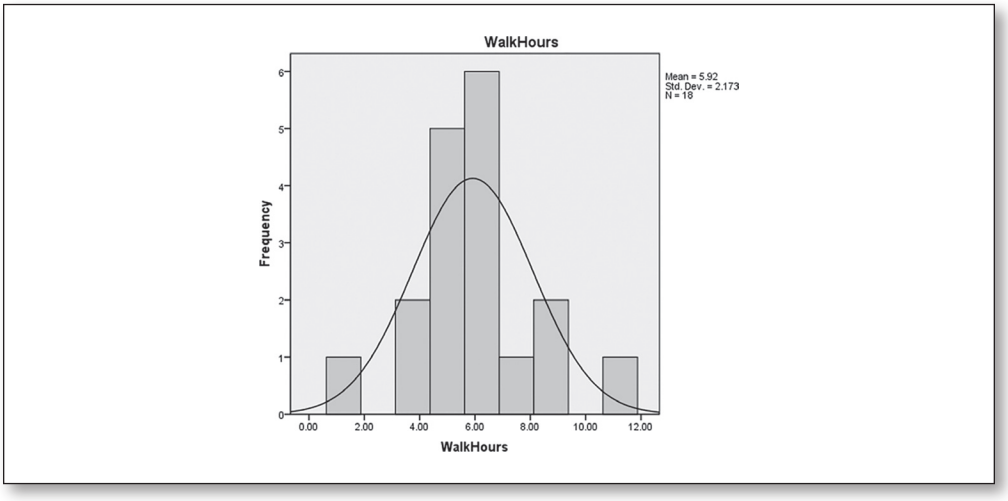
(1)

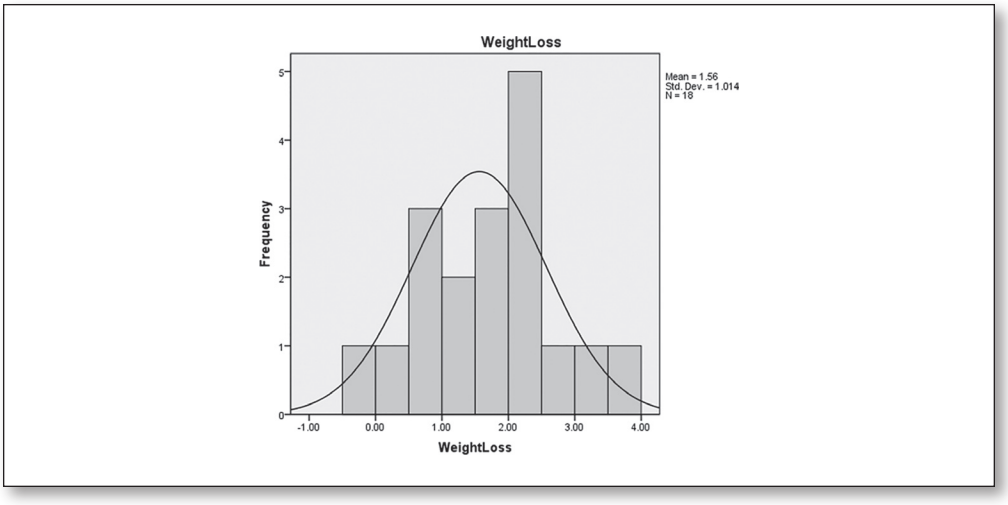
H_0: There is no significant correlation between hours of walking and weight loss

H_1: There is a significant correlation between hours of walking and weight loss.

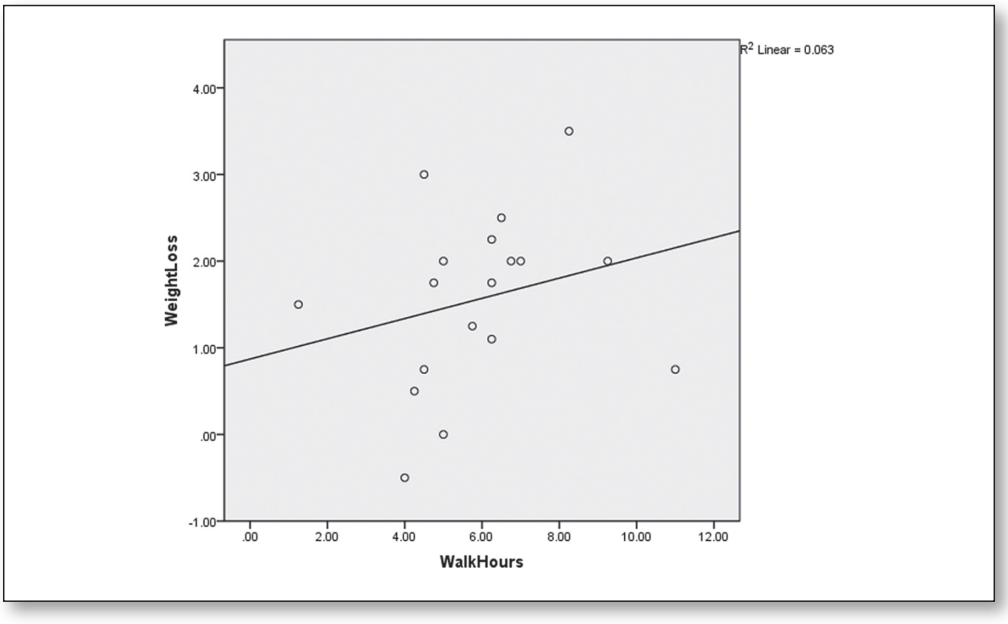
(b)

Histograms with normal curve plots show a normal distribution for *WalkHours* and *WeightLoss* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The scatterplot below shows that the straight regression line reasonably fits the field of points, hence the criterion of linearity is satisfied. The scatterplot also shows that the points are primarily clustered toward the center of the regression line with substantially fewer points toward the ends, hence the criterion of homoscedasticity is satisfied.



(c)

The correlations table below shows a positive correlation ($r = .250$) between *WalkHrs* and *WeightLoss*; however, since $p = .317$, which is greater than the specified .05 α level, this indicates that this is not a statistically significant correlation.

Correlations

		WalkHours	WeightLoss
WalkHours	Pearson Correlation	1	.250
	Sig. (2-tailed)		.317
	N	18	18
WeightLoss	Pearson Correlation	.250	1
	Sig. (2-tailed)	.317	
	N	18	18

(d)

In a study examining the effects of brisk walking on weight loss, 18 adults were recruited and instructed to walk as many days of the week they could for as long as they could. The researcher gathered each participant's daily walking time, as well as their weight at the beginning and end of the week and found a positive, correlation between the amount of walking time and the amount of weight lost ($r = .250$). Although the participants experienced some weight loss in this study, the correlation between walking and weight loss was not statistically significant ($p = .317$, $\alpha = .05$). As such, we do not reject H_0, and reject H_1.

EXERCISE 11.3, DATA SET A

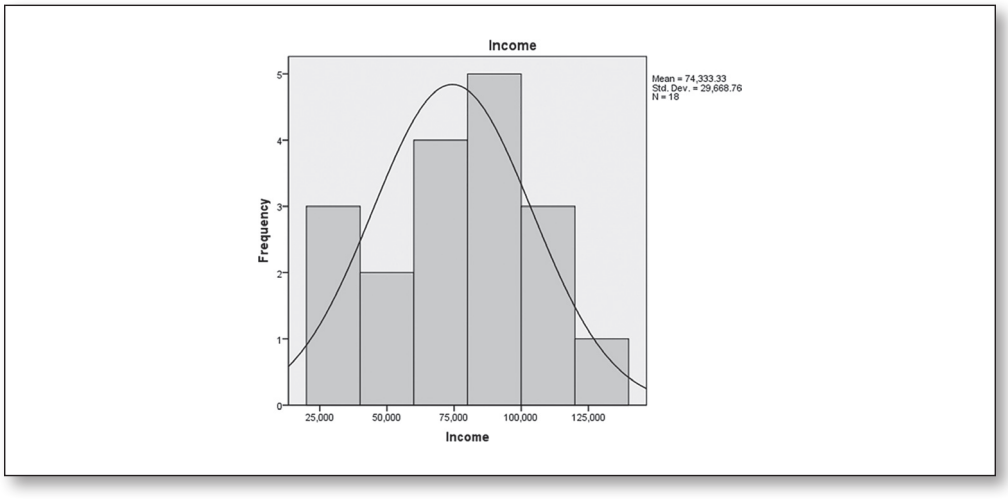
(a)

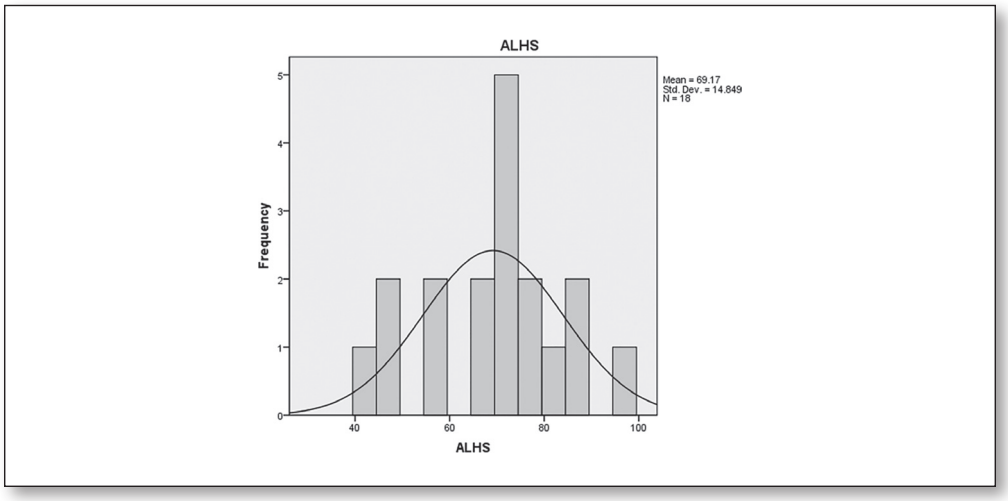
H_0: There is no significant correlation between income and happiness.

H_1: There is a significant correlation between income and happiness.

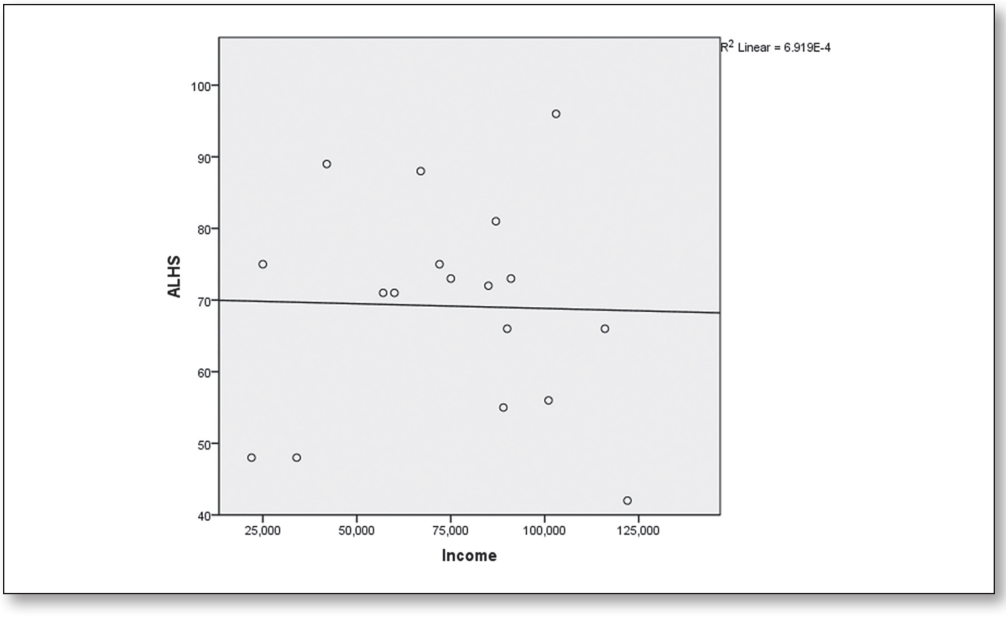
(b)

Histograms with normal curve plots show a normal distribution for *Income* and *ALHS* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The relatively broad spread of these points in relation to the regression line do not satisfy the criteria of linearity or homoscedasticity. As such, the Spearman statistic is a better alternative than Pearson.



(c)

The correlations table below shows a negative correlation ($\rho = -.109$) between Income and ALHS; however, $p = .668$, which is greater than the specified .05 α level, indicating that this is not a statistically significant correlation, as illustrated in the above scatterplot. Note the flatness of the regression line, along with the distance of the points from the regression line.

Correlations

			Income	ALHS
Spearman's rho	Income	Correlation Coefficient	1.000	-.109
		Sig. (2-tailed)	.	.668
		N	18	18
	ALHS	Correlation Coefficient	-.109	1.000
		Sig. (2-tailed)	.668	.
		N	18	18

(d)

In a study exploring the relationship between money and happiness, 18 adults were recruited and given a self-administered survey, which asked their annual income. The survey also included the Acme Life Happiness Scale (ALHS), a 10-question survey that renders a score from 0 to 100, with 0 being very unhappy, and 100 being very happy. Our findings revealed a slight negative correlation ($\rho = -.109$), meaning that earning more money was slightly associated with less happiness, however, this (negative) correlation was not found to be statistically significant ($p = .668$, $\alpha = .05$). As such, we did not reject H_0, and we rejected H_1.

EXERCISE 11.3, DATA SET B

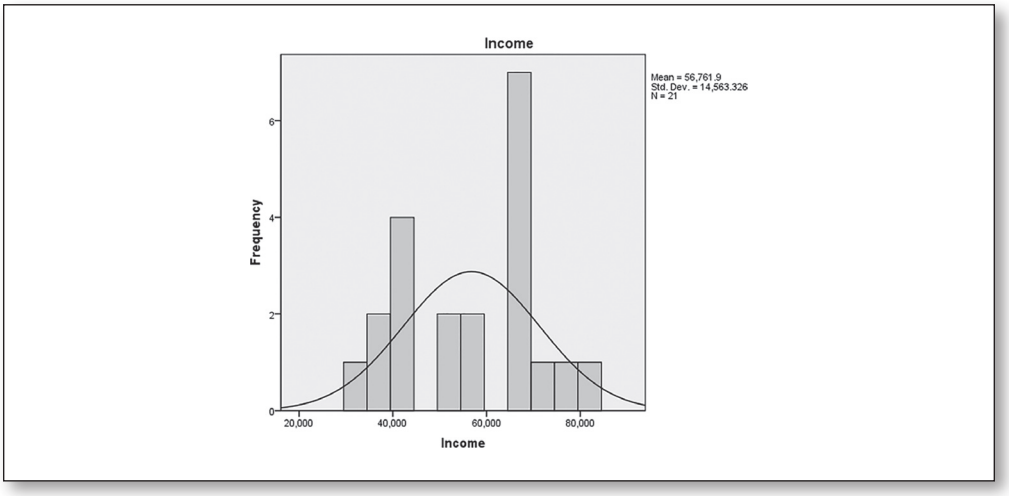
(a)

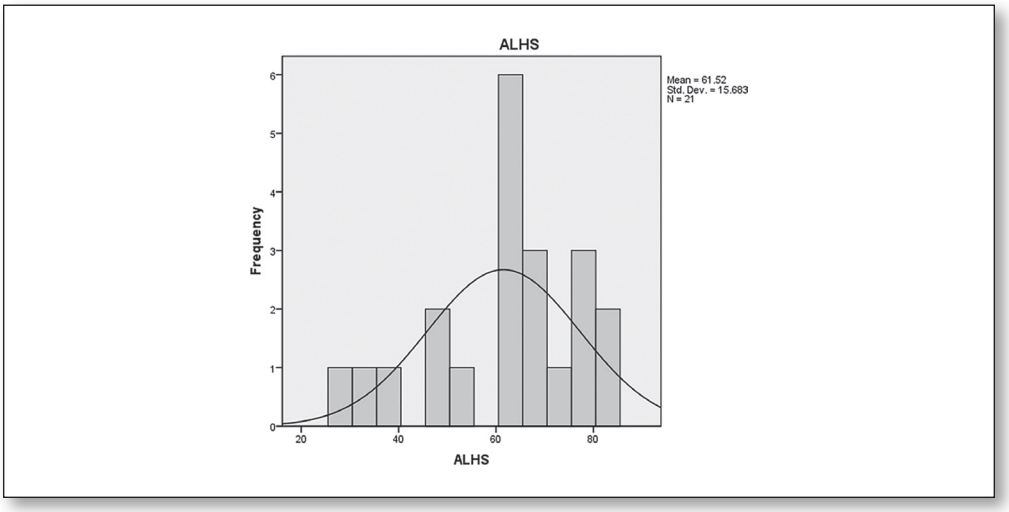
H_0: There is no significant correlation between income and happiness.

H_1: There is a significant correlation between income and happiness.

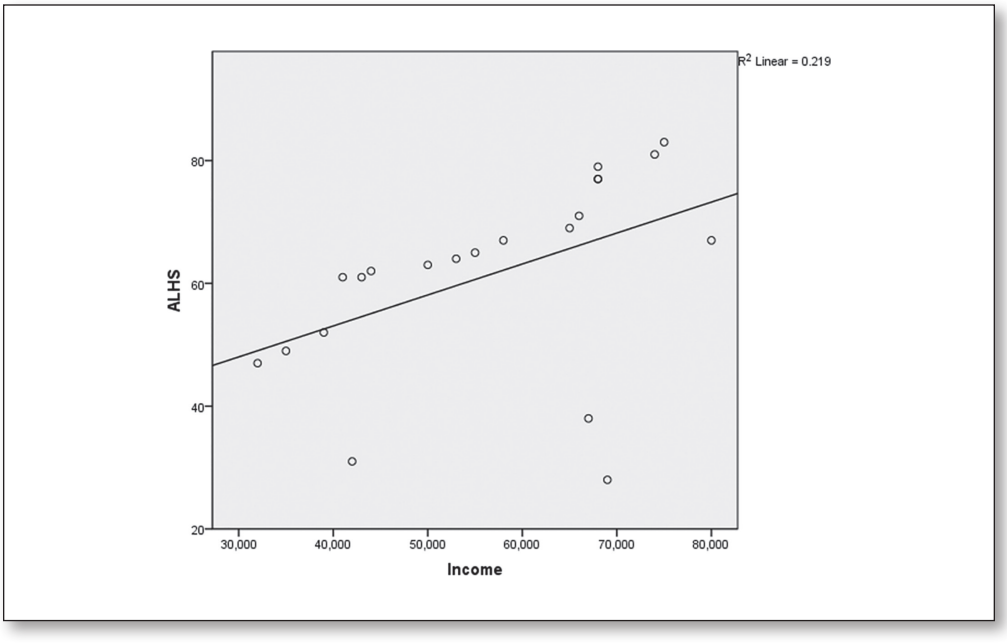
(b)

Histograms with normal curve plots show a normal distribution for *Income* and *ALHS* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The scatterplot below shows that the straight regression line reasonably fits the field of points, hence the criterion of linearity is satisfied. The scatterplot also shows that the points are primarily clustered toward the center of the regression line with substantially fewer points toward the ends, hence the criterion of homoscedasticity is satisfied.



(c)

The correlations table below shows a positive correlation ($r = .468$) between *Income* and *ALHS*. Since $p = .032$, is less than the specified .05 α level, this indicates that this is a statistically significant correlation.

Correlations

			Height	ASCI
Spearman's rho	Height	Correlation Coefficient	1.000	-.001
		Sig. (2-tailed)	.	.994
		N	41	41
	ASCI	Correlation Coefficient	-.001	1.000
		Sig. (2-tailed)	.994	.
		N	41	41

(d)

In a study exploring the relationship between money and happiness, 21 adults were recruited and given a self-administered survey that asked their annual income; the survey also included the Acme Life Happiness Scale (ALHS), a 10-question survey that renders a score from 0 to 100 (0 = very unhappy, 100 = very happy). Our findings revealed a positive correlation ($r = .468$), meaning that earning more money was associated with higher levels of happiness; this (positive) correlation was found to be statistically significant ($p = .032$, $\alpha = .05$). As such, we reject H_0 but do not reject H_1.

EXERCISE 11.5, DATA SET A

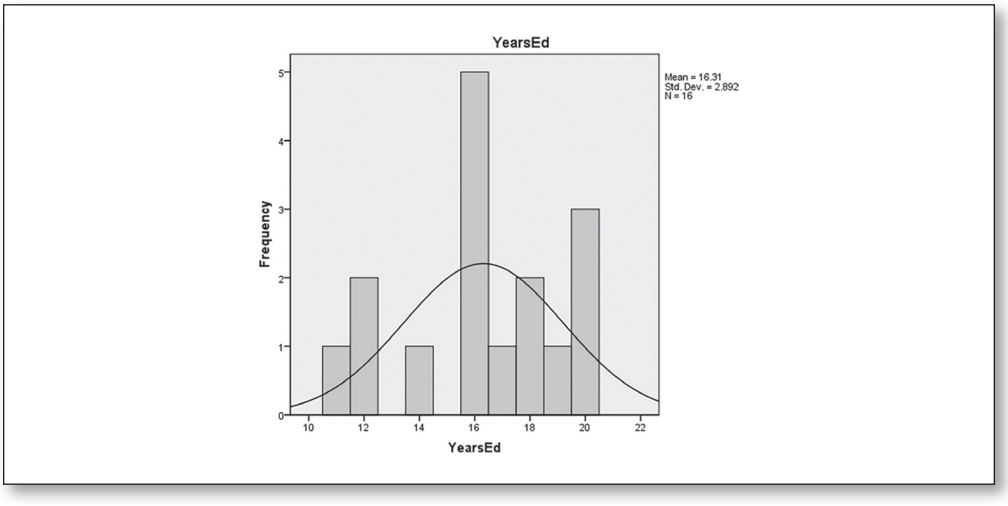
(a)

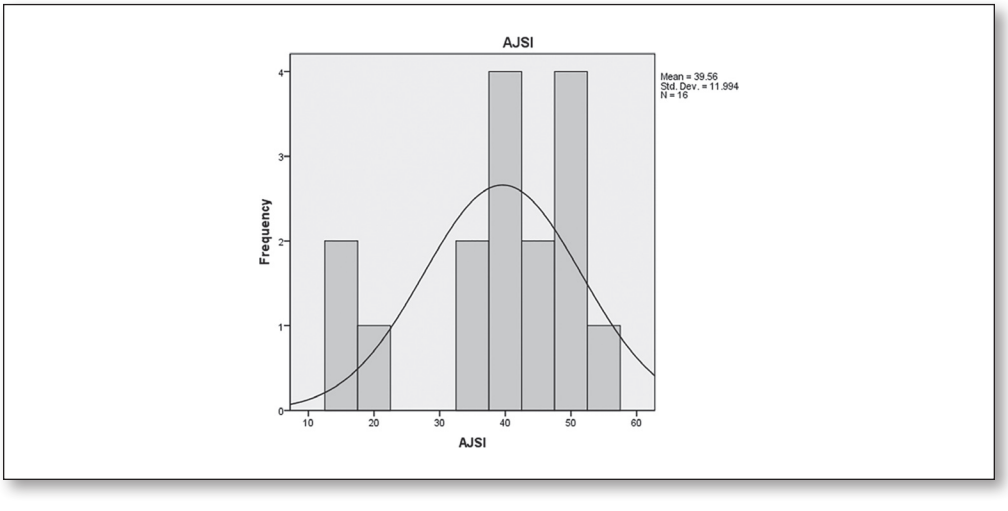
H_0: There is no significant correlation between years of education and job satisfaction.

H_1: There is a significant correlation between years of education and job satisfaction.

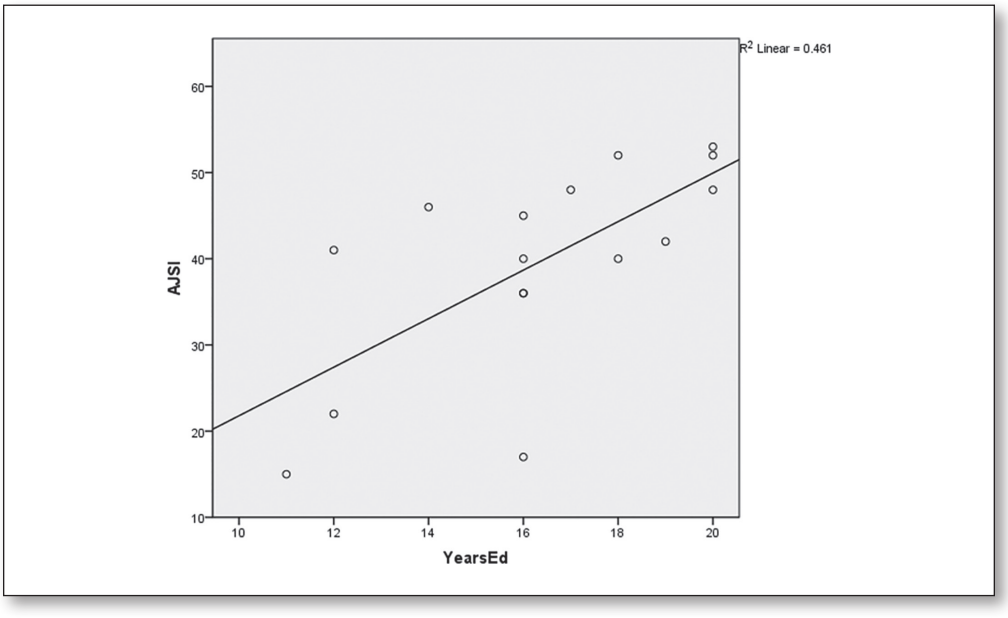
(b)

Histograms with normal curve plots show a normal distribution for *YearsEd* and *AJSI* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The scatterplot below shows that the straight regression line reasonably fits the field of points; hence, the criterion of linearity is satisfied. The scatterplot also shows that the points are primarily clustered toward the center of the regression line with substantially fewer points toward the ends; hence, the criterion of homoscedasticity is satisfied.



(c)

The correlations table below shows a positive correlation ($r = .679$) between *YearsEd* and *AJSI*; $p = .004$, which is less than the specified .05 α level, indicating that this is a statistically significant correlation, as illustrated in the scatterplot above; note the positive slope of the regression line and that most of the points are clustered fairly close to the regression line.

Correlations

		YearsEd	AJSI
YearsEd	Pearson Correlation	1	.679**
	Sig. (2-tailed)		.004
	N	16	16
AJSI	Pearson Correlation	.679**	1
	Sig. (2-tailed)	.004	
	N	16	16

**. Correlation is significant at the 0.01 level (2-tailed).

(d)

In order to assess the possible correlation between years of education and job satisfaction, 16 adults were recruited and given a self-administered survey which asked how many years of education they had; the survey also included the Acme Job Satisfaction Inventory (AJSI), a six question survey that renders a score from 0 to 60, with 0 being very dissatisfied with one's job, and 60 being very satisfied. The participants' education level ranged from 11th grade through master's degree. Our findings revealed a relatively strong positive correlation ($r = .679$), suggesting that those with more education tended to get more satisfaction from their jobs; this correlation was found to be statistically significant ($p = .004$, $\alpha = .05$). As such, we reject H_0, and we do not reject H_1.

EXERCISE 11.5, DATA SET B

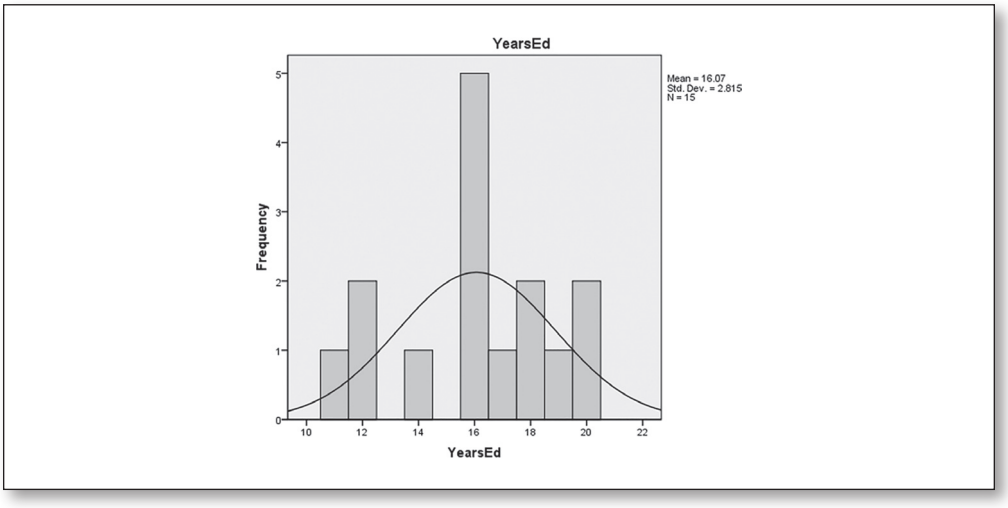
(a)

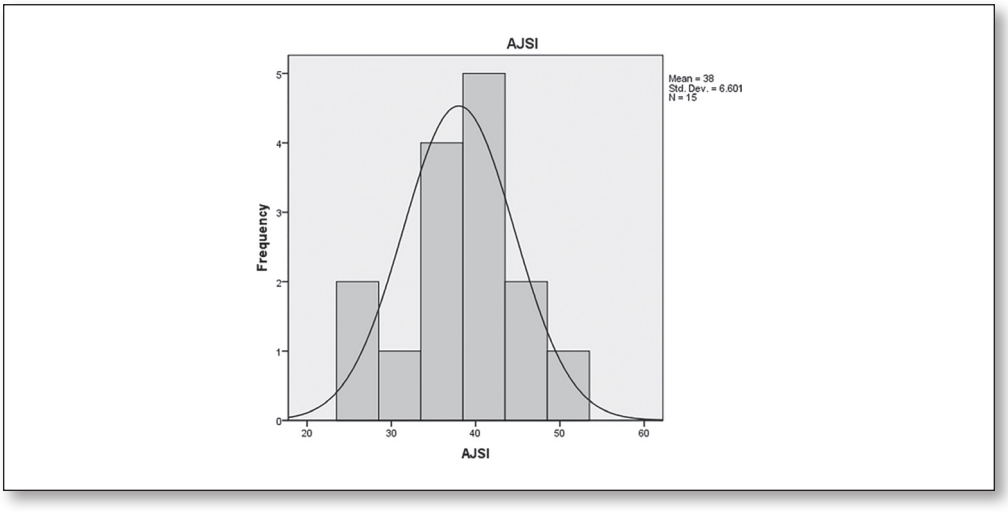
H_0: There is no significant correlation between years of education and job satisfaction.

H_1: There is a significant correlation between years of education and job satisfaction.

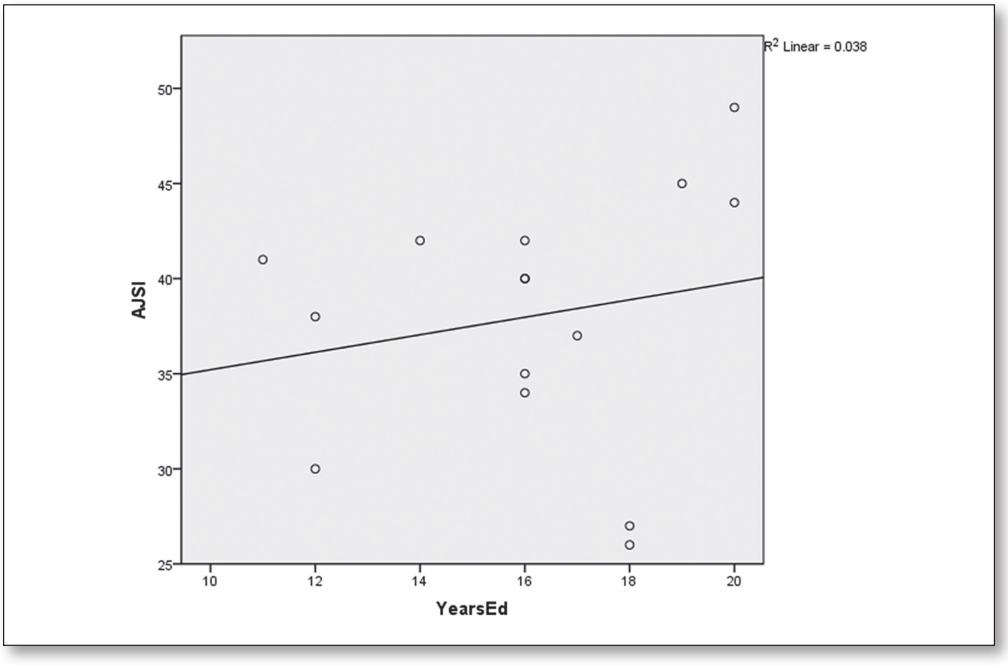
(b)

Histograms with normal curve plots show a normal distribution for *YearsEd* and *AJSI* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The scatterplot below shows that the straight regression line reasonably fits the field of points; hence, the criterion of linearity is satisfied. The scatterplot also shows that the points are primarily clustered toward the center of the regression line with substantially fewer points toward the ends; hence, the criterion of homoscedasticity is satisfied.



(c)

The correlations table below shows a relatively weak positive correlation (r = .196) between *YearsEd* and *AJSI*, however, p = .484, which is greater than the specified .05 α level, indicating that this is not a statistically significant correlation, as illustrated in the scatterplot above; note the slight positive slope of the regression line and that most of the points are clustered fairly far from the regression line.

Correlations

		YearsEd	AJSI
YearsEd	Pearson Correlation	1	.196
	Sig. (2-tailed)		.484
	N	15	15
AJSI	Pearson Correlation	.196	1
	Sig. (2-tailed)	.484	
	N	15	15

(d)

In order to assess the possible correlation between years of education and job satisfaction, 15 adults were recruited and given a self-administered survey which asked how many years of education they had; the survey also included the Acme Job Satisfaction Inventory (AJSI), a six question survey that renders a score from 0 to 60, with 0 being very dissatisfied with one's job, and 60 being very satisfied. The participant's education level ranged from 11th grade through master's degree. Our findings revealed a relatively weak positive correlation ($r = .196$), suggesting that those with more education tended to get somewhat more satisfaction from their jobs, however this correlation was not found to be statistically significant ($p = .484$, $\alpha = .05$). As such, we do not reject H_0, and we reject H_1.

EXERCISE 11.7, DATA SET A

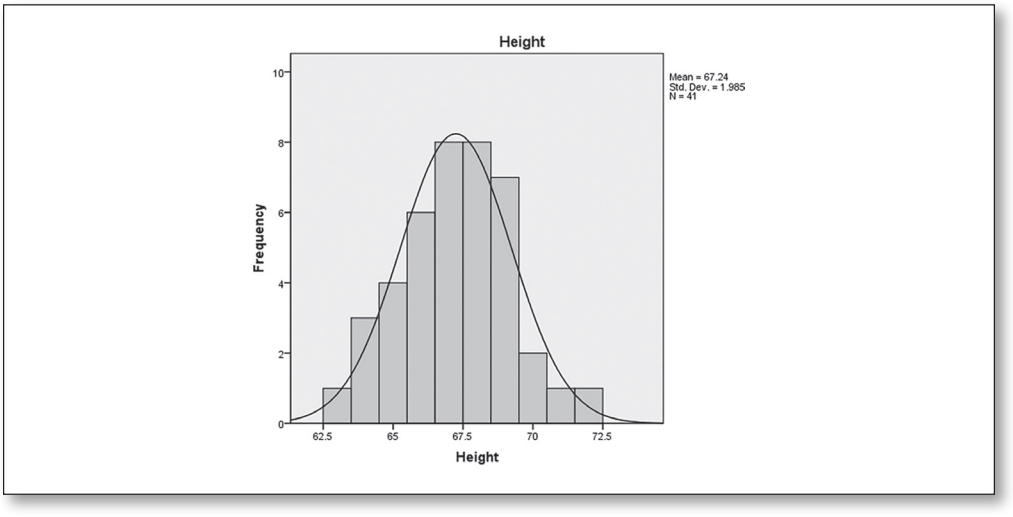
(a)

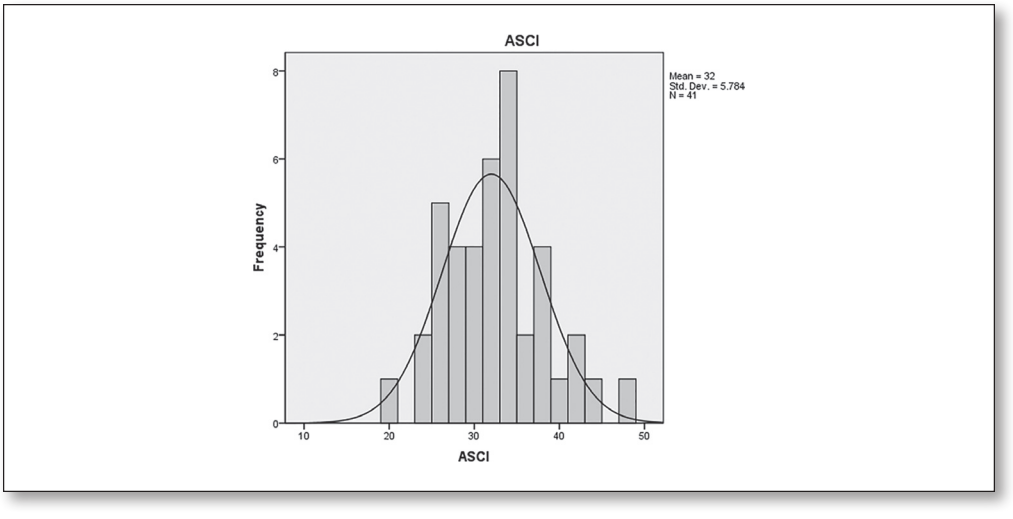
H_0: There is no significant correlation between height and self-confidence.

H_1: There is a significant correlation between height and self-confidence.

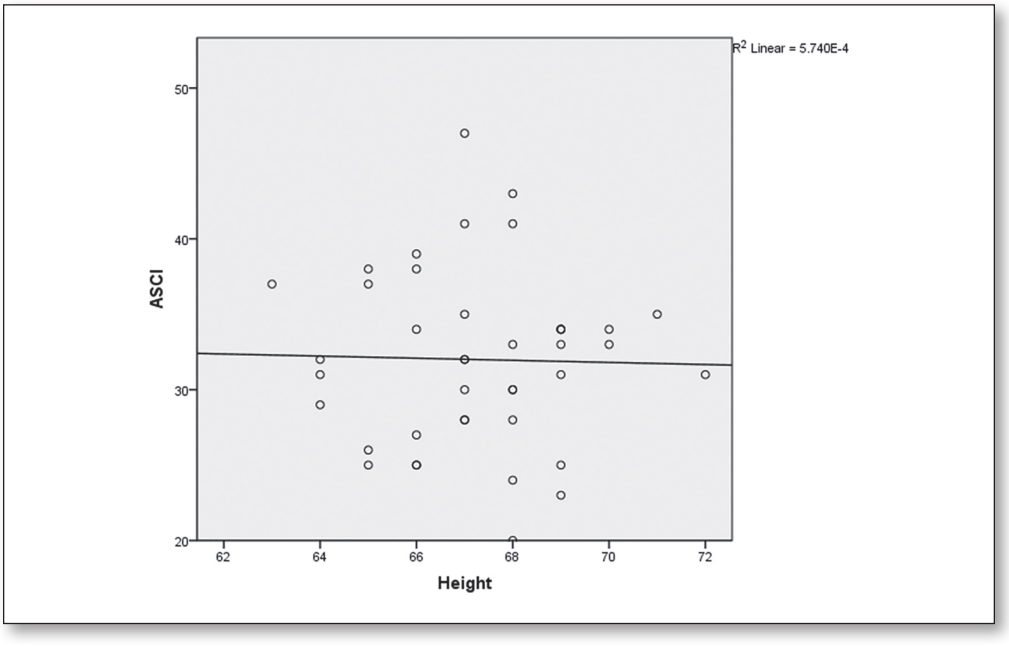
(b)

Histograms with normal curve plots show a normal distribution for *Height* and *ASCI* as shown in the two figures below; hence, the pretest criterion of *normality* is satisfied.





The relatively broad spread of these points in relation to the regression line do not satisfy the criteria of linearity or homoscedasticity. As such, the Spearman statistic is a better alternative than Pearson.



(c)

The correlations table below shows a relatively weak negative correlation ($\rho = -.001$) between Height and ASCI; $p = .994$, which is greater than the specified .05 α level, indicating that this is not a statistically significant correlation. As illustrated in the above scatterplot, the regression line is relatively flat and most of the points are clustered fairly far from the regression line.

Correlations

			Height	ASCI
Spearman's rho	Height	Correlation Coefficient	1.000	-.001
		Sig. (2-tailed)	.	.994
		N	41	41
	ASCI	Correlation Coefficient	-.001	1.000
		Sig. (2-tailed)	.994	.
		N	41	41

(d)

To evaluate the possible correlation between height and self-confidence, 41 adults were recruited and given a self-administered survey which asked them to indicate their height; the survey also included the Acme Self-Confidence Index (ASCI), a 10-question survey that renders a score from 0 to 50 (0 = very low self-confidence, 50 = very high self-confidence). Our findings suggest that although there was a negative correlation between the variables ($\rho = -.001$), suggesting that taller people were slightly less self-confident than shorter people. Further analysis revealed that this correlation is not statistically significant ($p = .994$, $\alpha = .05$). As such, we do not reject H_0, and we reject H_1.

EXERCISE 11.7, DATA SET B

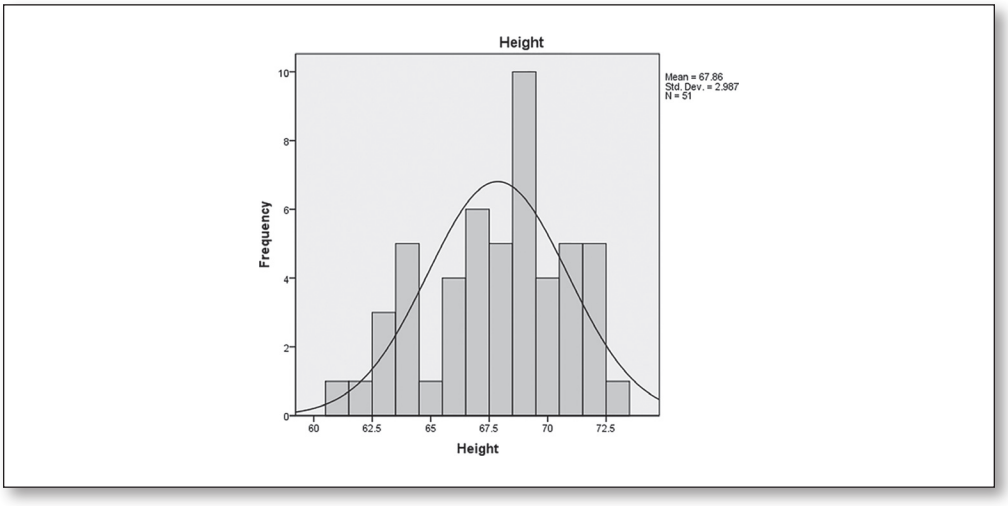
(a)

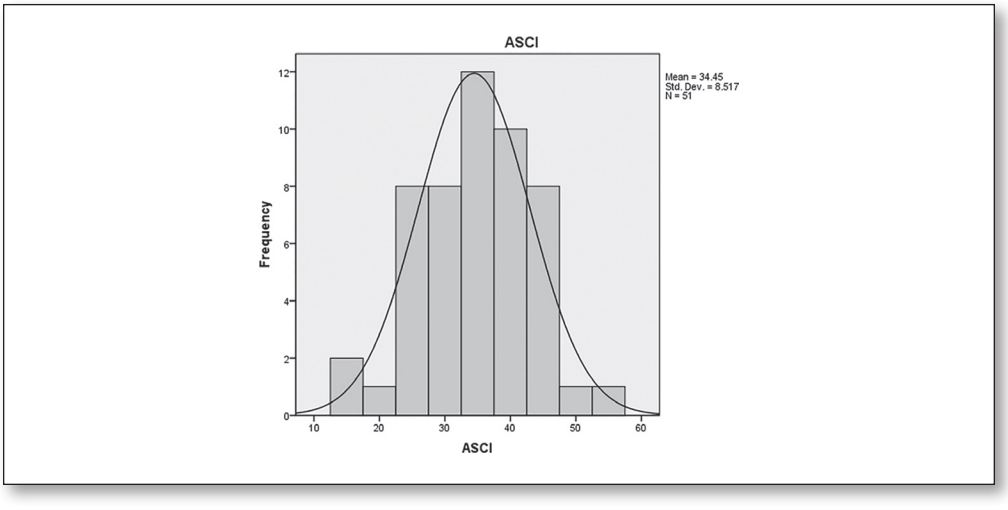
H_0: There is no significant correlation between height and self-confidence.

H_1: There is a significant correlation between height and self-confidence.

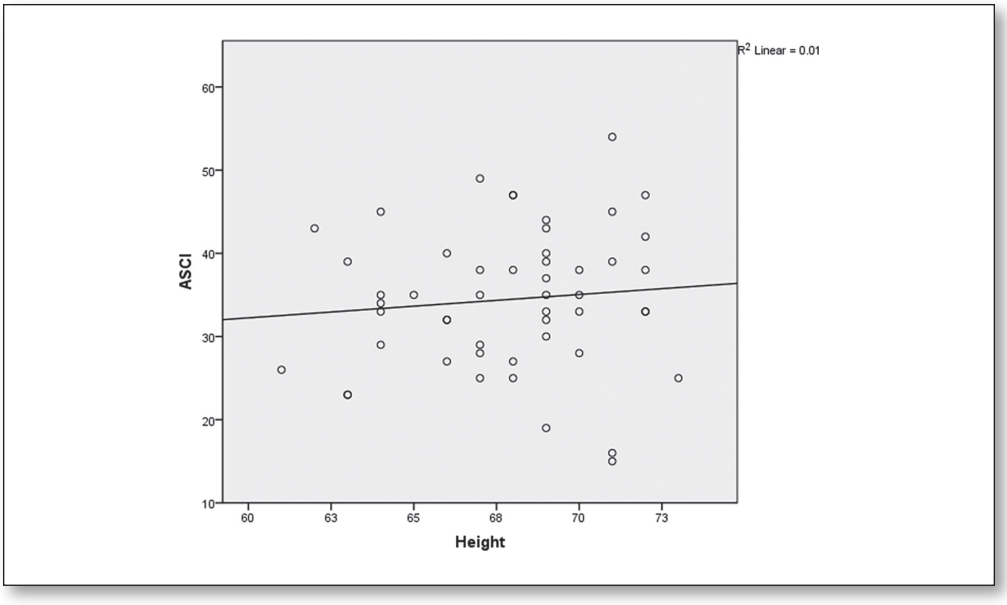
(b)

Histograms with normal curve plots show a normal distribution for *Height* and *ASCI* as shown in the two figures below, hence, the pretest criterion of *normality* is satisfied.





The relatively broad spread of these points in relation to the regression line do not satisfy the criteria of linearity or homoscedasticity. As such, the Spearman statistic is a better alternative than Pearson.



(c)

The correlations table below shows a relatively weak positive correlation ($\rho = .122$) between Height and ASCI, however, $p = .394$, which is greater than the specified .05 α level, indicating that this is not a statistically significant correlation. As illustrated in the above scatterplot, the regression line is relatively flat and most of the points are clustered fairly far from the regression line.

Correlations

			Height	ASCI
Spearman's rho	Height	Correlation Coefficient	1.000	.122
		Sig. (2-tailed)	.	.394
		N	51	51
	ASCI	Correlation Coefficient	.122	1.000
		Sig. (2-tailed)	.394	.
		N	51	51

(d)

To evaluate the possible correlation between height and self-confidence, 51 adults were recruited and given a self-administered survey which asked to indicate their height; the survey also included the Acme Self-Confidence Index (ASCI), a 10-question survey that renders a score from 0 to 50 (0 = very low self-confidence, 50 = very high self-confidence). Our findings suggest that although there was a slight positive correlation between the variables ($\rho = .122$), implying that taller people tend to have more self-confidence than shorter people. However, it was determined that this correlation is not statistically significant ($p = .394$, $\alpha = .05$). As such, we do not reject H_0, and we reject H_1.

EXERCISE 11.9, DATA SET A

(a)

H_0: There is no significant correlation in consumer car color preferences between the Pico and Sepulveda dealerships.

H_1: There is a significant correlation in consumer car color preferences between the Pico and Sepulveda dealerships.

(b)

We gathered car sales data from the Pico and Sepulveda sites, noting how many of each color (black, blue, red, silver, white, yellow) were sold at each site.

(c)

The Spearman's rho produced a positive correlation (.886) between *Pico* and *Sepulveda*; $p = .019$, which is less than the specified .05 α level, indicating that this is a statistically significant correlation.

Correlations

			Ariel	Dusty
Spearman's rho	Ariel	Correlation Coefficient	1.000	-.335
		Sig. (2-tailed)	.	.263
		N	13	13
	Dusty	Correlation Coefficient	-.335	1.000
		Sig. (2-tailed)	.263	.
		N	13	13

(d)

To evaluate the similarity of consumer car color preferences between the Pico dealership and the Sepulveda dealership, we reviewed the sales database and ranked the color from most sold to least sold for the specified time frame for each site. Spearman analysis revealed a statistically significant positive correlation (.886) suggesting that consumer car color preferences were similar between these two sites ($p = .019$, $\alpha = .05$). As such, we reject H_0, and we do not reject H_1.

EXERCISE 11.9, DATA SET B

(a)

H_0: There is no significant correlation in consumer car color preferences between the Pico and Sepulveda dealerships.

H_1: There is a significant correlation in consumer car color preferences between the Pico and Sepulveda dealerships.

(b)

We gathered car sales data from the Pico and Sepulveda sites, noting how many of each color (black, blue, red, silver, white, yellow) were sold at each site.

(c)

The Spearman's rho produced a positive correlation (.143) between *Pico* and *Sepulveda*; $p = .787$, which is greater than the specified .05 α level, indicating that this is not a statistically significant correlation.

Correlations

			Ariel	Dusty
Spearman's rho	Ariel	Correlation Coefficient	1.000	.747**
		Sig. (2-tailed)	.	.003
		N	13	13
	Dusty	Correlation Coefficient	.747**	1.000
		Sig. (2-tailed)	.003	.
		N	13	13

**. Correlation is significant at the 0.01 level (2-tailed).

(d)

To evaluate the similarity of consumer car color preferences between the Pico dealership and the Sepulveda dealership, we reviewed the sales database and ranked the color from most sold to least sold for the specified time frame for each site. Spearman analysis revealed a statistically insignificant positive correlation (.143) suggesting that consumer car color preferences were not similar between these two sites ($p = .787$, $\alpha = .05$). As such, we do not reject H_0, and reject H_1.