Practice Exam for the Relationship Between Variables

1. Which distribution is used to make the accept/reject decision for a hypothesis test for the slope coefficient (β_1) within linear regression:

A. chi-squared distribution
B. T-distribution
C. Binomial distribution
D. Normal distribution
E. None of the above

2. Which statistical tool should be used to determine if a linear relationship exists between two variables?

A. ANOVA
B. Chi-Squared Test
C. T-Test
D. Linear Regression
E. SPC
F. None of the above

3. Identify all of the statements below regarding linear regression that are false:

A. The dependent variable is also often called an Input.
B. The simple linear regression model is fully defined by only 2 coefficients.
C. The Slope coefficient defines the change in Y over a specific change in X.
D. The Y-intercept coefficient is called $\beta_{1.}$

4. Identify all of the statements below regarding linear regression that are true:

A. The independent variable is often called an explanatory variable.
B. The Y Variable is often called the Response Variable.
C. The Y intercept coefficient is the value of X when Y is zero.
D. Simple Linear Regression is Robust enough to apply to curvilinear data.

5. **Fill in the Blanks: The linear regression method models the relationship between ______A_______ mathematically with _________B_________.**
    A. A – Variables
    B. A –Coefficients
    C. A – Statistics
    D. A – Parameters
    E. B – Statistics
    F. B – Coefficients
    G. B – Equations
    H. B – Parameters

6. **Identify all of the statements below regarding linear regression that are false:**

A. If our R^2 value was 0.10, then 90% of the variation in Y can be explained by the variation in X.
B. Correlation means that a change in one variables also implies a change in another variable.
C. The coefficient of determination ranges from -1 to +1, with 0 meaning no measurable correlation.
D. The Pearson correlation coefficient reflects the strength of the linear relationship between two variables.

7. **Identify all of the statements below regarding linear regression that are true:**

A. The null hypothesis used within linear regression assumes that the β_1 coefficient is zero.
B. A correlation coefficient greater than 0.90 is sufficient to conclude that a change in the independent variable causes a change in the dependent variable.
C. The sign of the correlation coefficient will always match the sign of the Y-intercept coefficient.
D. The coefficient of determination, R^2 reflects the proportion of the total variability in the Y variable that can be explained by the regression line.

8. **Identify all of the statements below regarding linear regression that are true:**

A. Don't make projections for Y that are outside of the range of observed Y values.
B. The strength of a linear relationship is reflected in the Coefficient of Determination
C. A correlation coefficient of negative 1 implies that as X increases, Y will also increase.
D. If two variables had a positive correlation coefficient. Then as X decreased, Y would be expected to decrease.

9. **You're creating a linear regression model for your data and you've calculated the following values.**

SSE = 12.24, S_{yy} = 4.21 S_{xy} = 6.28

What is the slope of your model?

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 1.49
B. -1.49
C. 1.95
D. 1.28
E. -1.28
F. Not Enough Information Available

10. **You're creating a linear regression model for your data and you've calculated the following values.**

S_{yy} = 102, S_{xy} = 168, S_{xx} = 142

What is the slope coefficient for your regression model?

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 0.61
B. 0.72
C. 0.85
D. 1.18
E. 1.39
F. 1.65

11. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 8.4, S_{xy} = 21.2, S_{xx} = 13.6

What is the slope coefficient for your regression model?

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 0.64
B. 1.61
C. 1.55
D. 0.40
E. 2.52
F. 0.62

12. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	**15**
2	**30**
3	**45**
4	**60**

Calculate S_{xx} for your linear regression model:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 5
B. 10
C. 30
D. 25
E. 75
F. 150

13. **You're creating a linear regression model and you've measured the following pairs of data:**

X	Y
1	**15**
2	**30**
3	**45**
4	**60**

Calculate S_{xy} for your linear regression model:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 5
B. 10
C. 30
D. 25
E. 75
F. 150

14. **You're creating a linear regression model and you've measured the following pairs of data:**

X	Y
1	**15**
2	**30**
3	**45**
4	**60**

Calculate S_{yy} for your linear regression model:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 75
B. 150
C. 450
D. 1125
E. 6750

15. You're creating a linear regression model for your data and you've calculated the following values.

$$S_{yy} = 1125, S_{xy} = 75, S_{xx} = 5, \beta_0 = -12$$

What is the predicted value of Y when X = 10.

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 63
B. 75
C. 138
D. 150
E. 738

16. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	18
6	27
9	48
14	46
21	78
24	84
29	77

Calculate S_{xx} for your linear regression model:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 104
B. 627
C. 1527
D. 2172
E. 4130

17. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	18
6	27
9	48
14	46
21	78
24	84
29	77

Calculate S_{xy} for your linear regression model:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 104
B. 627
C. 1527
D. 2172
E. 4130

18. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	18
6	27
9	48
14	46
21	78
24	84
29	77

Calculate S_{yy} for your linear regression model:

Hint: CQE Academy Equation List page 9 - You'll be able to use this on the actual exam.

A. 104
B. 627
C. 1527
D. 2172
E. 4130

19. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 4130, S_{xy} = 1527, S_{xx} = 626.86, β_0 = 17.81

What is the predicted value of Y when X = 23.

Hint: CQE Academy Equation List page 9 - You'll be able to use this on the actual exam.

A. 48
B. 54
C. 56
D. 66
E. 74
F. Not Enough Information Provided

20. You're creating a linear regression model based on the following 5 sample data points. what is the predicted value of Y when X = 55:

X	Y
65	32
60	71
57	122
50	135
45	191

Hint: CQE Academy Equation List page 9 - You'll be able to use this on the actual exam.

A. 129
B. 120
C. 117
D. 113
E. 128

21. You've created a linear regression model based on the following data.

$$Y(x) = 522.94 - 7.45(x)$$

Calculate the SSE associated with this model

X	Y
65	32
60	71
57	122
50	135
45	191

Hint: CQE Academy Equation List page 9 - You'll be able to use this on the actual exam.

A. 551
B. 881
C. 110
D. 1202
E. 522
F. 293

22. You've created a linear regression model based on the following data.

$$Y(x) = 522.94 - 7.45(x)$$

You want to perform a hypothesis test to confirm that a linear relationship truly exists. Below are the key sample statistics associated with your linear regression model:

SSE = 17,287, n = 5, S_{yy} =14,934, S_{xx} = 253.20,

What is the conclusion of your hypothesis test.

Hint: CQE Academy Equation List page 9 - You'll be able to use this on the actual exam.

A. The t-statistic is greater than the t-critical, thus the null hypothesis cannot be rejected and a linear relationship cannot be confirmed.
B. The t-statistic is less than the t-critical, thus the null hypothesis must be rejected and a linear relationship can be confirmed.
C. The t-statistic is greater than the t-critical, thus the null hypothesis must be rejected and a linear relationship can be confirmed.
D. The t-statistic is less than the t-critical, thus the null hypothesis cannot be rejected and a linear relationship cannot be confirmed.
E. Not enough information has been provided to make a conclusion.

23. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 4130, S_{xy} = 1527, S_{xx} = 626.86

What is correlation coefficient for this data set:

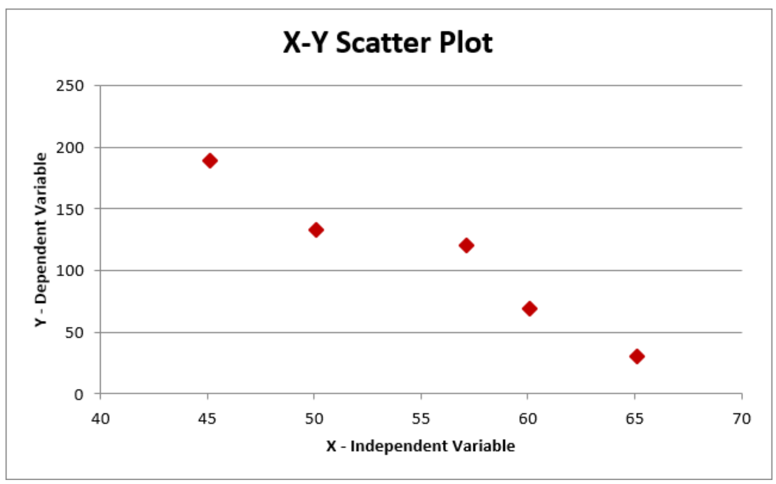
Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. -1
B. -0.4
C. 0
D. 0.25
E. 0.72
F. 0.95
G. 1
H. Not Enough Information Provided

24. You're creating a linear regression model for your data and your scatter plot looks something like this:

Based on this initial data, estimate the Pearson Correlation Coefficient

A. -1
B. -0.95
C. -0.60
D. 0
E. 0.95
F. 1
G. Not Enough Information Provided



25. You're creating a linear regression model for your data and you've calculated the following values.

$S_{yy} = 31$, $S_{xy} = 12$, $S_{xx} = 76$

What is correlation coefficient for this data set:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. -1
B. -0.4
C. 0
D. 0.25
E. 0.72
F. 0.95
G. 1
H. Not Enough Information Provided

26. You're creating a linear regression model for your data and you've calculated the following values.

$S_{yy} = 31$, $S_{xy} = 12$, $S_{xx} = 76$

Based on these results, what percentage of variation in Y, can be explained by the variation in X.

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 100%
B. 81%
C. 76%
D. 43%
E. 28%
F. 13%
G. 6%
H. 1%
I. Not Enough Information Provided

27. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 20,478, S_{xy} = -8,719, S_{xx} = 13,241

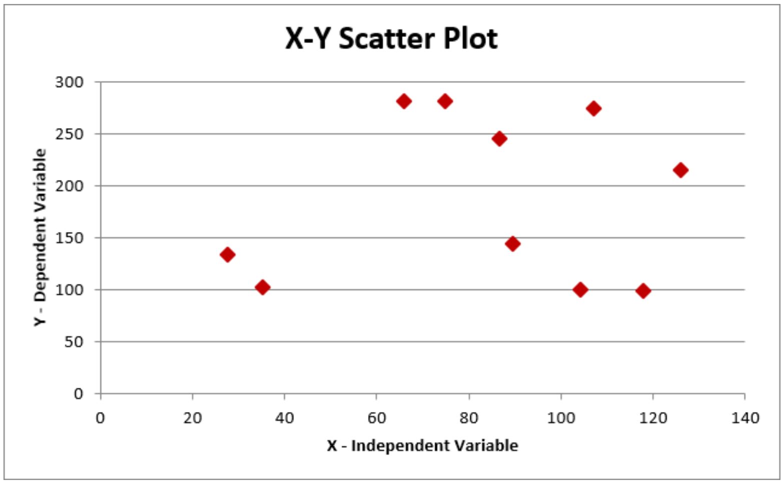
What is correlation coefficient for this data set:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. -0.72
B. -0.53
C. -0.21
D. 0
E. 0.25
F. 0.72
G. Not Enough Information Provided

28. You're creating a linear regression model for your data and your scatter plot looks something like this:

Based on this initial data, estimate the Pearson Correlation Coefficient



A. -1
B. -0.95
C. -0.60
D. 0
E. 0.95
F. 1
G. Not Enough Information Provided

29. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 181, S_{xy} = -58, S_{xx} = 112

What is correlation coefficient for this data set:

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. -0.72
B. -0.4
C. 0.16
D. 0.40
E. 0.72
F. Not Enough Information Provided

30. You're creating a linear regression model for your data and you've calculated the following values.

31. S_{yy} = 181, S_{xy} = -58, S_{xx} = 112

Based on these results, what percentage of variation in Y, can be explained by the variation in X.

Hint: CQE Academy Equation List - You'll be able to use this on the actual exam.

A. 100%
B. 81%
C. 76%
D. 41%
E. 28%
F. 16%
G. 6%
H. 1%
I. Not Enough Information Provided

Problem Set Solution:

1. **Which distribution is used to make the accept/reject decision for a hypothesis test for the slope coefficient (β_1) within linear regression:**

A. chi-squared distribution
B. T-distribution
C. Binomial distribution
D. Normal distribution
E. None of the above

2. **Which statistical tool should be used to determine if a linear relationship exists between two variables?**

A. ANOVA
B. Chi-Squared Test
C. T-Test
D. Linear Regression
E. SPC
F. None of the above

3. **Identify all of the statements below regarding linear regression that are false:**

A. The ~~dependent~~ independent variable is also often called an Input (False)
B. The simple linear regression model is fully defined by only 2 coefficients (True)
C. The Slope coefficient defines the change in Y over a specific change in X. (True)
D. The Y-intercept coefficient is called ~~β_1~~ β_0 (False)

4. **Identify all of the statements below regarding linear regression that are true:**

A. The independent variable is often called an explanatory variable (True)
B. The Y Variable is often called the Response Variable (True)
C. The Y intercept coefficient is the value of X when Y is zero. (False)
D. Simple Linear Regression is Robust enough to apply to curvilinear data (False)

5. Fill in the Blanks: The linear regression method models the relationship between ______A_______ mathematically with _________B_________.

- **A. A – Variables**
- B. A –Coefficients
- C. A – Statistics
- D. A – Parameters
- E. B – Statistics
- **F. B – Coefficients**
- G. B – Equations
- H. B – Parameters

6. Identify all of the statements below regarding linear regression that are false:

A. If our R^2 value was 0.10, then ~~90%~~ 10% of the variation in Y can be explained by the variation in X.

B. Correlation means that a change in one variables also implies a change in another variable.

C. The correlation coefficient ~~coefficient of determination~~ ranges from -1 to +1, with 0 meaning no measurable correlation.

D. The Pearson correlation coefficient reflects the strength of the linear relationship between two variables.

7. Identify all of the statements below regarding linear regression that are true:

A. The null hypothesis used within linear regression assumes that the β_1 coefficient is zero (True)

B. A correlation coefficient greater than 0.90 is sufficient to conclude that a change in the independent variable ~~causes~~ correlates with a change in the dependent variable. (False)

C. The sign of the correlation coefficient will always match the sign of the ~~Y-intercept~~ slope coefficient. (False)

D. The coefficient of determination, R^2 reflects the proportion of the total variability in the Y variable that can be explained by the regression line (True)

8. Identify all of the statements below regarding linear regression that are true:

A. Don't make projections for Y that are outside of the range of observed Y values.

B. The strength of a linear relationship is reflected in the ~~Coefficient of Determination~~ correlation coefficient

C. A correlation coefficient of negative 1 implies that as X increases, Y will decrease ~~also increase~~.

D. If two variables had a positive correlation coefficient. Then as X decreased, Y would be expected to decrease.

9. **You're creating a linear regression model for your data and you've calculated the following values.**

SSE = 12.24, S_{yy} = 4.21 S_{xy} = 6.28

What is the slope of your model?

A. 1.49
B. -1.49
C. 1.95
D. 1.28
E. -1.28
F. Not Enough Information Available

The slope of the linear regression model is β_1 which we can calculate using the sum of squares equation below. We know SSE, S_{yy} and S_{xy}, and if we plug in those values we can solve for the slope β_1.

$$SSE = S_{yy} - \beta_1 S_{xy}$$

$$\mathrm{SSE} = 12.24, \mathrm{S_{yy}} = 4.21 \quad \mathrm{S_{xy}} = 6.28$$

$$12.24 = 4.21 - \beta_1 * 6.28$$

$$\beta_1 = \frac{12.24 - 4.21}{-6.28} = -1.278$$

10. **You're creating a linear regression model for your data and you've calculated the following values.**

S_{yy} = 102, S_{xy} = 168, S_{xx} = 142

What is the slope coefficient for your regression model?

A. 0.61
B. 0.72
C. 0.85
D. 1.18
E. 1.39
F. 1.65

The slope of the linear regression model, $\beta_1 = S_{xy} / S_{xx} = 168 / 142 = 1.18$

11. **You're creating a linear regression model for your data and you've calculated the following values.**

$$S_{yy} = 8.4, S_{xy} = 21.2, S_{xx} = 13.6$$

What is the slope coefficient for your regression model?

A. 0.64
B. 1.61
C. 1.55
D. 0.40
E. 2.52
F. 0.62

The slope of the linear regression model, $\beta_1 = S_{xy} / S_{xx} = 21.2 / 13.6 = 1.55$

12. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	**15**
2	**30**
3	**45**
4	**60**

Calculate S_{xx} for your linear regression model:

A. 5
B. 10
C. 30
D. 25
E. 75
F. 150

Below is the equation to calculate S_{xx}.

$$S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

Let's go back to our data to calculate the sum of X^2, and then we can square the sum of X.

	X	**Y**	**X²**
	1	15	1
	2	30	4
	3	45	9
	4	60	16
Total	**10**	**150**	**30**

$$S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n} = 30 - \frac{10^2}{4} = 30 - 25 = 5$$

13. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	**15**
2	**30**
3	**45**
4	**60**

Calculate S_{xy} for your linear regression model:

A. 5
B. 10
C. 30
D. 25
E. 75
F. 150

Below is the equation to calculate S_{xy}.

$$S_{xy} = \sum_{i}^{n}(X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n}$$

Let's go back to our data to calculate these values.

	X	Y	X*Y
	1	15	15
	2	30	60
	3	45	135
	4	60	240
Total	**10**	**150**	**450**

$$S_{xy} = \sum_{i}^{n}(X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n} = 450 - \frac{10 * 150}{4} = 75$$

14. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	**15**
2	**30**
3	**45**
4	**60**

Calculate S_{yy} for your linear regression model:

A. 75
B. 150
C. 450
D. 1125
E. 6750

Below is the equation to calculate S_{yy}.

$$S_{yy} = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

Let's go back to our data to calculate these values.

	X	Y	Y^2
	1	15	225
	2	30	900
	3	45	2025
	4	60	3600
Total	**10**	**150**	**6750**

$$S_{yy} = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = 6750 - \frac{150^2}{4} = 1125$$

15. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 1125, S_{xy} = 75, S_{xx} = 5, β_0 = -12

What is the predicted value of Y when X = 10.

A. 63
B. 75
C. 138
D. 150
E. 738

Below is the regression equation to solve for Y.

$$Y(x) = \beta_0 + \beta_1 * x$$

To solve for Y, we need to calculate the slope coefficient:

$$\beta_1 = \frac{S_{xy}}{S_{xx}} = \frac{75}{5} = 15$$

Now we can solve for Y(10).

$$Y(10) = -12 + 15 * 10 = 138$$

16. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	18
6	27
9	48
14	46
21	78
24	84
29	77

Calculate S_{xx} for your linear regression model:

A. 104
B. 627
C. 1527
D. 2172
E. 4130

Below is the equation to calculate S_{xx}.

$$S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

Let's go back to our data to calculate the sum of X^2, and then we can square the sum of X.

	X	Y	x^2
	1	18	1
	6	27	36
	9	48	81
	14	46	196
	21	78	441
	24	84	576
	29	77	841
Total	**104**	**378**	**2172**

$$S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n} = 2172 - \frac{104^2}{7} = 626.86$$

17. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	18
6	27
9	48
14	46
21	78
24	84
29	77

Calculate S_{xy} for your linear regression model:

A. 104
B. 627
C. 1527
D. 2172
E. 4130

Below is the equation to calculate S_{xy}.

$$S_{xy} = \sum_{i}^{n}(X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n}$$

Let's go back to our data to calculate these values.

	X	Y	x*y
	1	18	18
	6	27	162
	9	48	432
	14	46	644
	21	78	1638
	24	84	2016
	29	77	2233
Total	**104**	**378**	**7143**

$$S_{xy} = \sum_{i}^{n}(X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n} = 7143 - \frac{104 * 378}{7} = 1527$$

18. You're creating a linear regression model and you've measured the following pairs of data:

X	Y
1	18
6	27
9	48
14	46
21	78
24	84
29	77

Calculate S_{yy} for your linear regression model:

A. 104
B. 627
C. 1527
D. 2172
E. 4130

Below is the equation to calculate S_{xy}.

$$S_{yy} = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

Let's go back to our data to calculate these values.

	X	Y	y^2
	1	18	324
	6	27	729
	9	48	2304
	14	46	2116
	21	78	6084
	24	84	7056
	29	77	5929
Total	**104**	**378**	**24542**

$$S_{yy} = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = 24,542 - \frac{378^2}{7} = 4130$$

19. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 4130, S_{xy} = 1527, S_{xx} = 626.86, β_0 = 17.81

What is the predicted value of Y when X = 23.

A. 48
B. 54
C. 56
D. 66
E. 74
F. Not Enough Information Provided

Below is the regression equation to solve for Y.

$$\boldsymbol{Y(x) = \beta_0 + \beta_1 * x}$$

To solve for Y, we need to calculate the slope coefficient:

$$\boldsymbol{\beta_1 = \frac{S_{xy}}{S_{xx}} = \frac{1527}{626.86} = 2.44}$$

Now we can solve for Y(10).

$$\boldsymbol{Y(23) = 17.81 + 2.44 * 23 = 73.84}$$

20. You're creating a linear regression model based on the following 5 sample data points. what is the predicted value of Y when X = 55:

X	Y
65	32
60	71
57	122
50	135
45	191

A. 129
B. 120
C. 117
D. 113
E. 128

	X	Y	x^2	x*y
	65	32	4225	2080
	60	71	3600	4260
	57	122	3249	6954
	50	135	2500	6750
	45	191	2025	8595
	277	551	15599	28639
Total	**277**	**551**	**15,599**	**28,639**

Below is the regression equation to solve for Y. To solve for Y, we need to calculate the slope coefficient (**β_1**) and the Y-intercept (**β_0**).

$$Y(x) = \beta_0 + \beta_1 * x$$

$$\beta_1 = \frac{S_{xy}}{S_{xx}} \quad \& \quad \beta_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$S_{xy} = \sum_i^n (X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n} \qquad S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

$$S_{xy} = \sum_i^n (X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n} = 28,639 - \frac{277 * 551}{5} = -1,886.40$$

$$S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n} = 15,599 - \frac{277^2}{5} = 253.20$$

$$\beta_1 = \frac{S_{xy}}{S_{xx}} = \frac{-1,886.40}{253.20} = -7.45$$

$$\beta_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 110.20 - (-7.40 * 55.40) = 522.94$$

$$Y(55) = \beta_0 + \beta_1 * x = 522.94 - 7.45(55) = 113$$

21. You've created a linear regression model based on the following data.

$$Y(x) = 522.94 - 7.45(x)$$

Calculate the SSE associated with this model

X	Y
65	32
60	71
57	122
50	135
45	191

A. 551
B. 881
C. 110
D. 1202
E. 522
F. 293

Error is calculated as the difference between the predicted and actual Y values, squared, then summed.

$$SSE = Sum\ of\ Squares\ of\ the\ Error = \sum \varepsilon rror^2 = \sum (y - \hat{y})^2$$

	X	Y	Y-hat	error (ε) (Y_i - Y_{hat})	ε^2
	65	32	38.68	6.68	44.59
	60	71	75.93	4.93	24.29
	57	122	98.28	-23.72	562.66
	50	135	150.43	15.43	238.12
	45	191	187.68	-3.32	11.01
Total					**880.67**

To calculated the predicted value of Y, we must plug in each X value (65, 60, 57, 50, 45) into our regression model, which is where we get the following predicted values of Y (38.68, 75.93, 98.28, 150.43, 187.68).

We can then take the difference between the predicted and the actual to calculate the error.

Then we square those values to get ε^2. Now finally we can sum up those squared values to get a total sum of square error of 880.67 (881).

22. You've created a linear regression model based on the following data.

$$Y(x) = 522.94 - 7.45(x)$$

You want to perform a hypothesis test to confirm that a linear relationship truly exists. Below are the key sample statistics associated with your linear regression model:

SSE = 17,287, n = 5, S_{yy} =14,934, S_{xx} = 253.20

What is the conclusion of your hypothesis test.

A. The t-statistic is greater than the t-critical, thus the null hypothesis cannot be rejected and a linear relationship cannot be confirmed.
B. The t-statistic is less than the t-critical, thus the null hypothesis must be rejected and a linear relationship can be confirmed.
C. The t-statistic is greater than the t-critical, thus the null hypothesis must be rejected and a linear relationship can be confirmed.
D. The t-statistic is less than the t-critical, thus the null hypothesis cannot be rejected and a linear relationship cannot be confirmed.
E. Not enough information has been provided to make a conclusion.

The **null hypothesis** associated with this test is that the **β_1 coefficient, for the slope of the line, is zero**. The **alternative hypothesis** is that your slope is not equal to zero, thus proving a linear relationship between the two variables.

H_0: $\beta_1 = 0$ & H_A: $\beta_1 \neq 0$

This hypothesis is conducted using the **t-distribution**, and the following calculation for the **t-statistic**.

$$t = \frac{\hat{\beta}_1 - \beta_1}{S_e / \sqrt{S_{xx}}}$$

To complete this equation, we must calculate the standard deviation of the error (S_e).

$$\textit{Sample Standard Deviation of the Error}: S_e = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{17,287}{5-2}} = 75.91$$

$$\textit{Degrees of Freedom}, d.f. = n - 2 = 5 - 2 = 3$$

$$t = \frac{\hat{\beta}_1 - \beta_1}{S_e / \sqrt{S_{xx}}} = \frac{-7.45 - 0}{75.91 / \sqrt{253.20}} = -1.56$$

Using the standard alpha risk of 5% (0.05), and our sample size of 3, we can look up the critical t-value from the NIST table.

$$t - critical = t_{95,3} = +/- 2.353$$

Since our t-statistic (-1.56) is less than our critical t-value (-2.353) we must fail to reject the null hypothesis that β_1 = 0. Thus, we cannot conclude that a linear relationship exists for these two variables.

23. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 4130, S_{xy} = 1527, S_{xx} = 626.86

What is correlation coefficient for this data set:

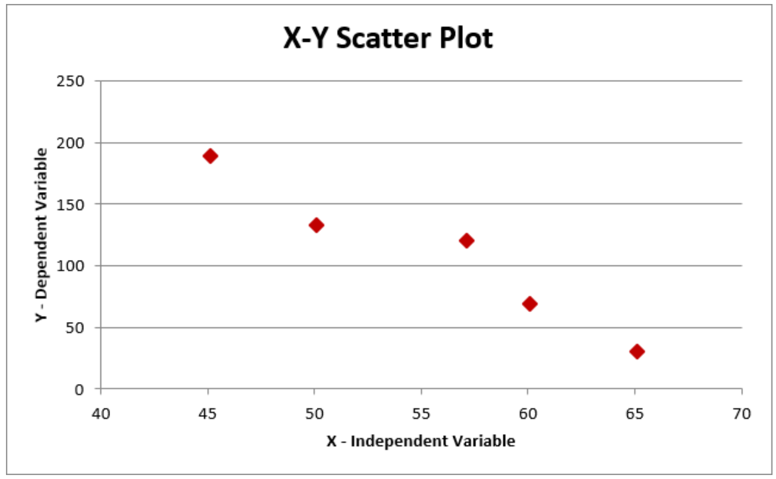
A. -1
B. -0.4
C. 0
D. 0.25
E. 0.72
F. 0.95
G. 1
H. Not Enough Information Provided

Below is the equation to solve for the Pearson Correlation Coefficient.

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}} = \frac{1527}{\sqrt{626.86} * \sqrt{4130}} = 0.95$$

24. You're creating a linear regression model for your data and your scatter plot looks something like this:

Based on this initial data, estimate the Pearson Correlation Coefficient



A. -1
B. -0.95
C. -0.60
D. 0
E. 0.95
F. 1
G. Not Enough Information Provided

25. You're creating a linear regression model for your data and you've calculated the following values.

$S_{yy} = 31$, $S_{xy} = 12$, $S_{xx} = 76$

What is correlation coefficient for this data set:

A. -1
B. -0.4
C. 0
D. 0.25
E. 0.72
F. 0.95
G. 1
H. Not Enough Information Provided

Below is the equation to solve for the Pearson Correlation Coefficient.

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}} = \frac{12}{\sqrt{76} * \sqrt{31}} = 0.25$$

26. You're creating a linear regression model for your data and you've calculated the following values.

$S_{yy} = 31$, $S_{xy} = 12$, $S_{xx} = 76$

Based on these results, what percentage of variation in Y, can be explained by the variation in X.

A. 100%
B. 81%
C. 76%
D. 43%
E. 28%
F. 13%
G. 6%
H. 1%
I. Not Enough Information Provided

This question is asking you to solve for the Coefficient of Determination, R^2.

This coefficient reflects the proportion of variation in Y that can be explained by the variation in X.

Below is the equation to solve for the Pearson Correlation Coefficient, which then can be squared to find the coefficient of determination.

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}} = \frac{12}{\sqrt{76} * \sqrt{31}} = 0.25$$

$$R^2 = 0.25^2 = 0.06 = 6\%$$

27. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 20,478, S_{xy} = -8,719, S_{xx} = 13,241

What is correlation coefficient for this data set:

A. -0.72
B. -0.53
C. -0.21
D. 0
E. 0.25
F. 0.72
G. Not Enough Information Provided

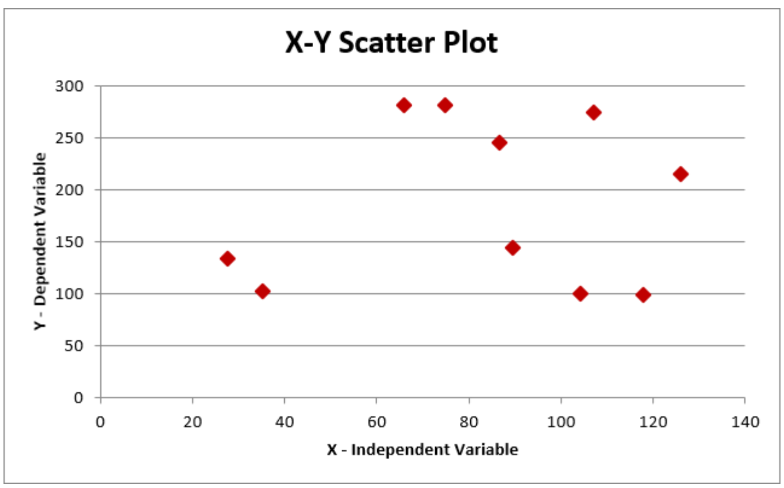
Below is the equation to solve for the Pearson Correlation Coefficient.

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}} = \frac{-8{,}719}{\sqrt{13{,}241} * \sqrt{20{,}478}} = -0.53$$

28. You're creating a linear regression model for your data and your scatter plot looks something like this:

Based on this initial data, estimate the Pearson Correlation Coefficient

A. -1
B. -0.95
C. -0.60
D. 0
E. 0.95
F. 1
G. Not Enough Information Provided



29. You're creating a linear regression model for your data and you've calculated the following values.

S_{yy} = 181, S_{xy} = -58, S_{xx} = 112

What is correlation coefficient for this data set:

A. -0.72
B. -0.4
C. 0.16
D. 0.40
E. 0.72
F. Not Enough Information Provided

Below is the equation to solve for the Pearson Correlation Coefficient.

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}} = \frac{-58}{\sqrt{112} * \sqrt{181}} = -0.41$$

30. You're creating a linear regression model for your data and you've calculated the following values.

31. S_{yy} = 181, S_{xy} = -58, S_{xx} = 112

Based on these results, what percentage of variation in Y, can be explained by the variation in X.

A. 100%
B. 81%
C. 76%
D. 41%
E. 28%
F. 16%
G. 6%
H. 1%
I. Not Enough Information Provided

This question is asking you to solve for the Coefficient of Determination, R^2. This coefficient reflects the proportion of variation in Y that can be explained by the variation in X.

Below is the equation to solve for the Pearson Correlation Coefficient, which then can be squared to find the coefficient of determination.

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}} = \frac{-58}{\sqrt{112} * \sqrt{181}} = -0.41$$

$$R^2 = -0.41^2 = 0.1659 = 16\%$$