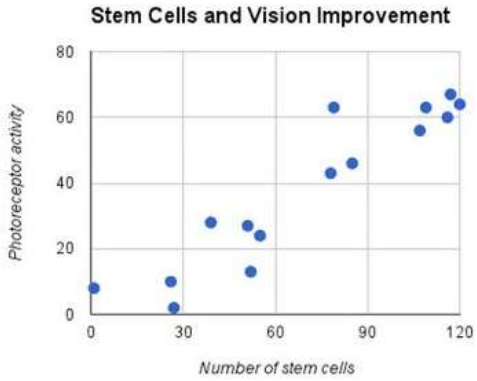
Dimensionless Technologies Private Limited
Visit us at: www.dimensionless.in
- info@dimensionless.in
- 9923170071, 8108094992

Problem Set – Lesson 15

Q1.

You can restore impaired vision using stem cells. Scientists found a strong correlation between the number of stem cells added to the eyes of mice and their vision improvement (all mice started at a photoreceptor activity* of 0). Note: This example has been simplified for the purposes of this lesson.

* Photoreceptor activity is a measure of how well eyes perceive light. The higher the photoreceptor activity, the better one's vision.



This data is found in the Google spreadsheet. Calculate the slope of the linear regression line $y = bx + a$, where $b = r\left(\frac{S_y}{S_x}\right)$.

$b =$

https://docs.google.com/spreadsheets/d/1b-0_jxZQz6BKsQX1gzFDQdatlJh4QRvgKbO8_if--ps/edit?usp=sharing

Q2.

Calculate r^2. $r^2 =$

What does this value mean? (Check all that apply.)

- ○ A proportion of r^2 of the variation in photoreceptor activity is related to the number of stem cells.
- ○ A proportion of r^2 of mice that received the stem cells had better photoreceptor activity.
- ○ Mice that received stem cells had a percent increase equal to r^2 in photoreceptor activity.
- ○ A proportion of r^2 of the variation in photoreceptor activity is caused by the number of stem cells.

Q3.

Calculate the y-intercept a, given that the linear regression line goes through the point $(\bar{x}, \bar{y})$ and has the slope you calculated in the last exercise.

$a =$

Q4.

Now that you know the slope and y-intercept, what is the equation of the regression line?

$y = \square x + \square$

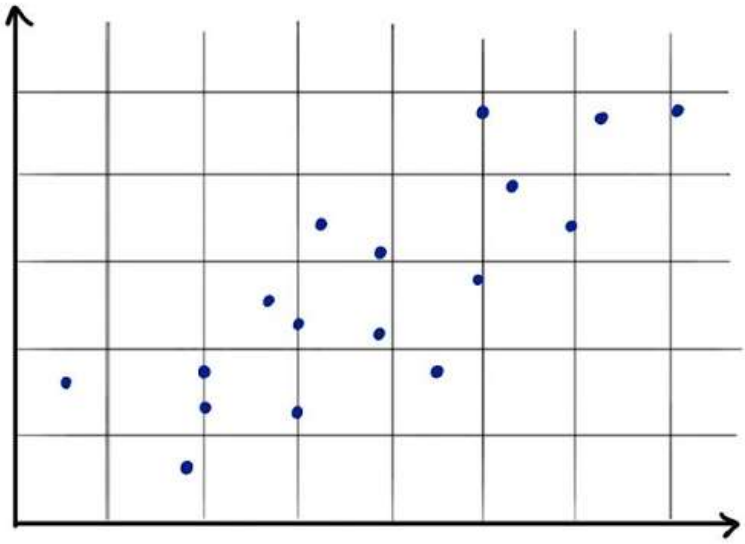
Q5.

Photoreceptor activity of 70 is considered normal. How many stem cells would we expect are needed for a visually impaired mouse (meaning photoreceptor activity is 0) to have normal vision?

Q6.

According to our line of best fit, the mice need at least _____ stem cells added to their eyes to have a positive value of photoreceptor activity.

Q7.



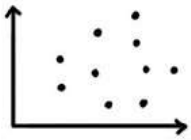
Estimate the slope of the regression line.

$b \approx$

Q8.

Indicate the slope of the regression lines for the following data.

Zero slope ☐ Positive slope ☐ Negative slope ☐



☐ ☐ ☐

☐ ☐ ☐

Q9.

The confidence interval for the slope of the regression line is (−0.684, 1.733). What can we conclude?

Check all that apply.

- ☐ We should reject the null; there is evidence of a linear relationship between x and y.
- ☐ The correlation coefficient r is not significant.
- ☐ The slope of the regression line for the sample data is definitely between −0.684 and 1.733.
- ☐ The slope of the regression line for the population is definitely between −0.684 and 1.733.
- ☐ The true slope for the population (β_1) could be negative.

Q10.

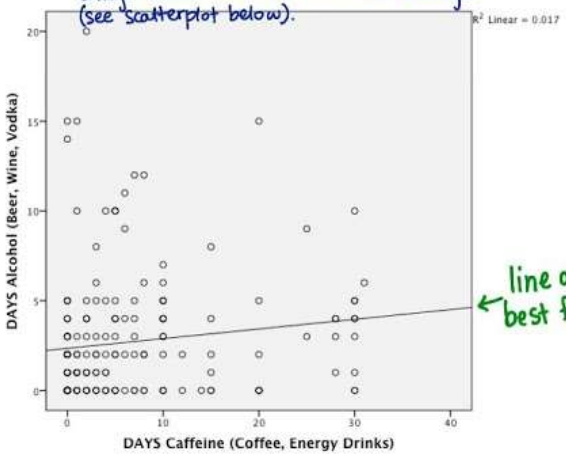
Match each symbol with the appropriate definition.

HINT: regression equation for sample: $\hat{y}_i = bx_i + a$
regression equation for population: $\hat{y}_i = \beta_1 x_i + \beta_0$

A	Observed values of y	____ a
B	Observed values of x	____ x_i
C	Expected values of y	____ $\hat{\beta}_1$
D	Slope of regression line based on sample data	____ $\hat{y}_i$ ____ b
E	Y-intercept of regression line based on sample data	____ $\hat{\beta}_0$
F	True slope	____ y_i
G	True y-intercept	

Q11.

Laraway and Snycerski (2013) examined the relationship between the number of days in the last month that US college students (N=186) drank caffeinated beverages (such as coffee, tea, energy drinks) and the number of days in the last month these same college students drank alcoholic beverages (see scatterplot below).



line of best fit

The linear regression equation for this relationship is

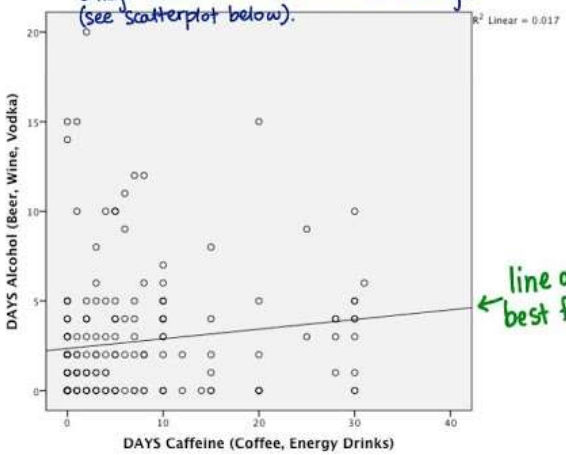
$$\hat{Y} = 2.35 + 0.05x.$$

How many alcoholic beverages would we predict were consumed by the following hypothetical students? (Round to two decimal places and write them down since you'll need them for the next question:

Luis (10 caffeinated beverages):
Johanna (5 caffeinated beverages):
Ravi (0 caffeinated beverages):
Patrick (20 caffeinated beverages):

Q12.

Laraway and Snycerski (2013) examined the relationship between the number of days in the last month that US college students (N=186) drank caffeinated beverages (such as coffee, tea, energy drinks) and the number of days in the last month these same college students drank alcoholic beverages (see scatterplot below).



line of best fit

Remember that the standard error of the estimate is like a standard deviation around the regression line. The standard error of the estimate in this relationship is 3.50. Use this to create a confidence interval for the number of alcoholic beverages consumed (expected value ± standard error of the estimate) for each hypothetical student in the last example.

Luis (10 caffeinated beverages): (____, ____)
Johanna (5 caffeinated beverages): (____, ____)
Ravi (0 caffeinated beverages): (____, ____)
Patrick (20 caffeinated beverages): (____, ____)

Q13.

In this example, $r^2 \approx 0.2$.
What does this mean? (Select all that apply.)

- ○ 20% of the variation in alcohol consumption is caused by caffeine consumption.
- ○ 20% of the variation in alcohol consumption is related to caffeine consumption.
- ○ 20% of all students drank both alcohol and caffeinated beverages last month.
- ○ Every 1-unit increase in caffeine consumption results in a 2% increase in alcohol consumption.

Q14.

Which of the following definitions best describes the slope?

- It is the change in X given a decrease in the Y-intercept
- It is the change in Y given a one-unit change in X
- It is the change in the Y-intercept given a one-unit change in X
- It is the average deviation of Y scores around the regression line
- It is the value of Y when X = 0
- It is the value of X when Y = 0
- It is the predicted value for Y given a value of X

Q15.

Which of the following definitions best describes the y-intercept?

- It is the change in X given a decrease in the Y-intercept
- It is the change in Y given a one-unit change in X
- It is the change in the Y-intercept given a one-unit change in X
- It is the average deviation of Y scores around the regression line
- It is the value of Y when X = 0
- It is the value of X when Y = 0
- It is the predicted value for Y given a value of X

Q16.

Which of the following definitions best describes y-hat ($\hat{y}$)?

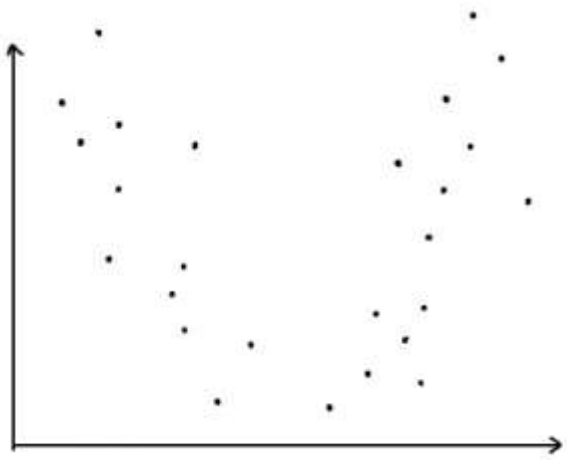
- It is the change in X given a decrease in the Y-intercept
- It is the change in Y given a one-unit change in X
- It is the change in the Y-intercept given a one-unit change in X
- It is the average deviation of Y scores around the regression line
- It is the value of Y when X = 0
- It is the value of X when Y = 0
- It is the predicted value for Y given a value of X

Q17.

Which of the following definitions best describes the standard error of the estimate?

- It is the change in X given a decrease in the Y-intercept
- It is the change in Y given a one-unit change in X
- It is the change in the Y-intercept given a one-unit change in X
- It is the average deviation of Y scores around the regression line
- It is the value of Y when X = 0
- It is the value of X when Y = 0
- It is the predicted value for Y given a value of X

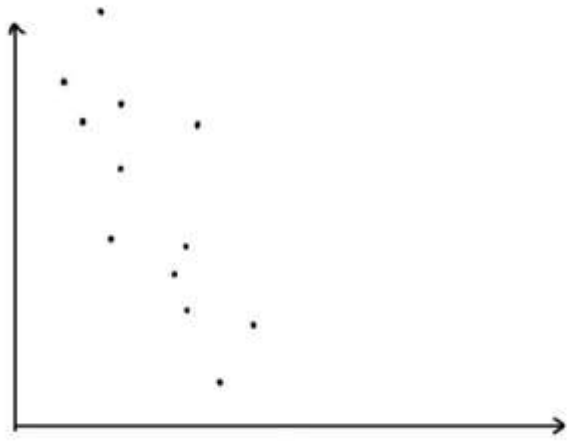
Q18.



How would you describe this relationship? (Select all that apply.)

- Linear
- Nonlinear
- Curvilinear

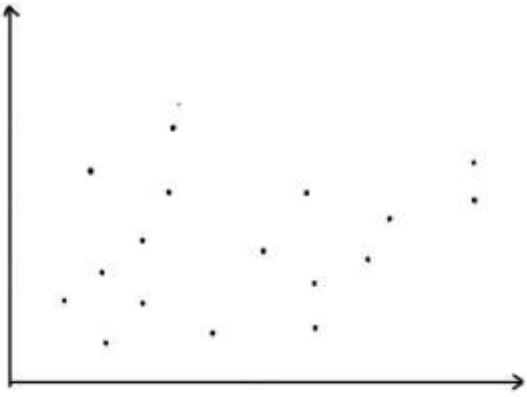
Q19.



How would you describe this relationship? (Select all that apply.)

- Linear
- Nonlinear
- Curvilinear
- Negative
- Positive
- Strong
- Weak

Q20.



How would you describe this relationship? (Select all that apply.)

- Linear
- Nonlinear
- Curvilinear
- Negative
- Positive
- Strong
- Weak