# PRINTABLE VERSION

## Quiz 7

You scored 100 out of 100

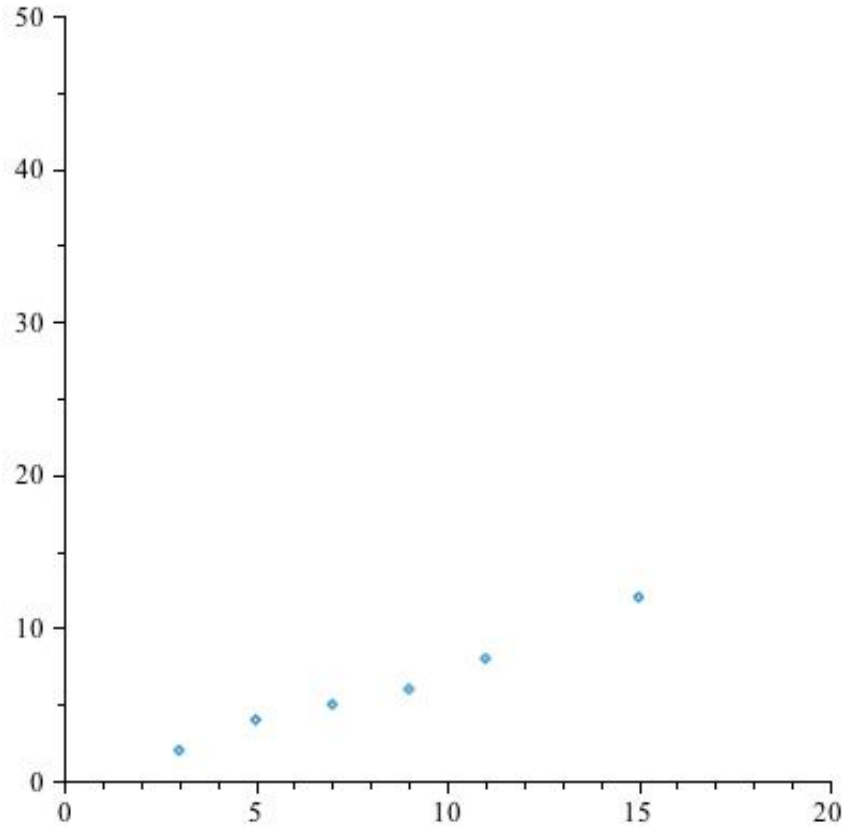
**Question 1**

Your answer is CORRECT.

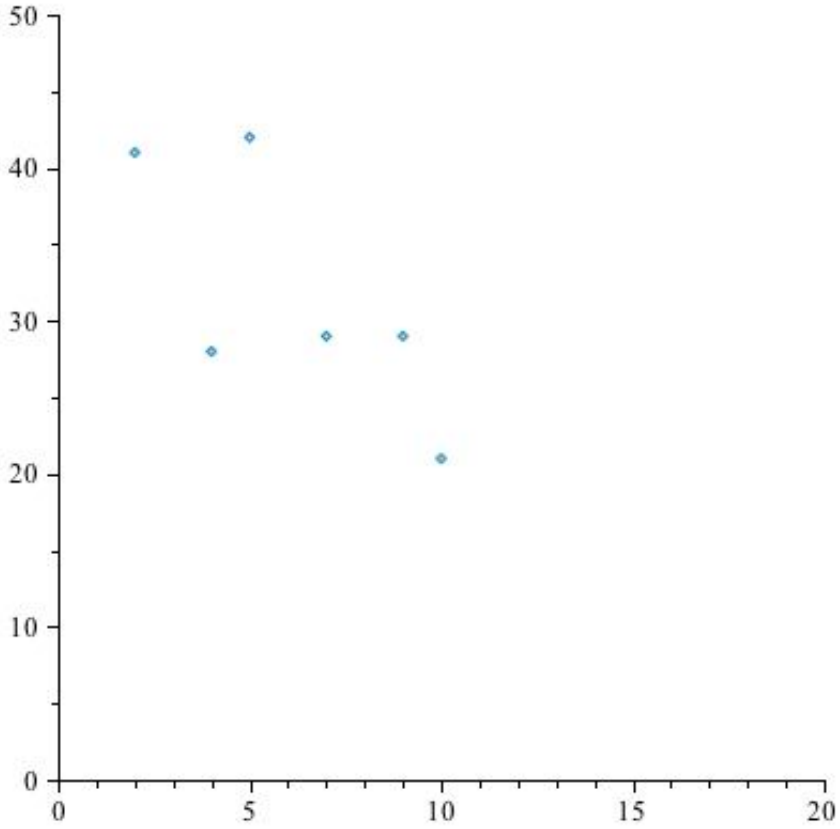
Which of the following would be the scatterplot for the given data?

| $x$ | 4 | 7 | 8 | 12 | 14 | 18 |
|---|---|---|---|---|---|---|
| $y$ | 21 | 29 | 29 | 42 | 28 | 41 |

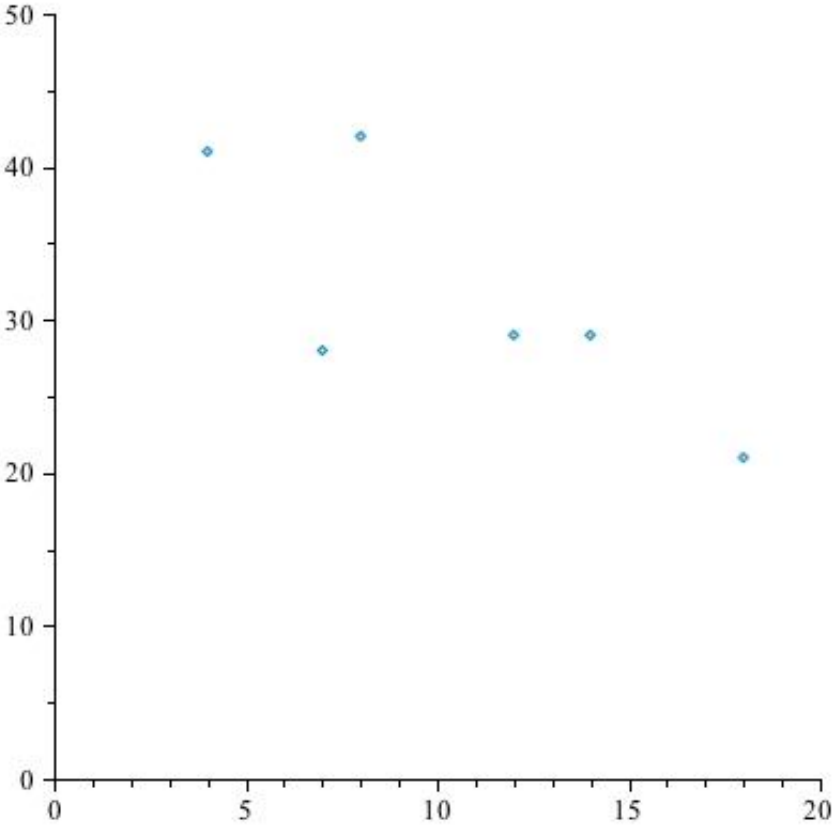
**a)**

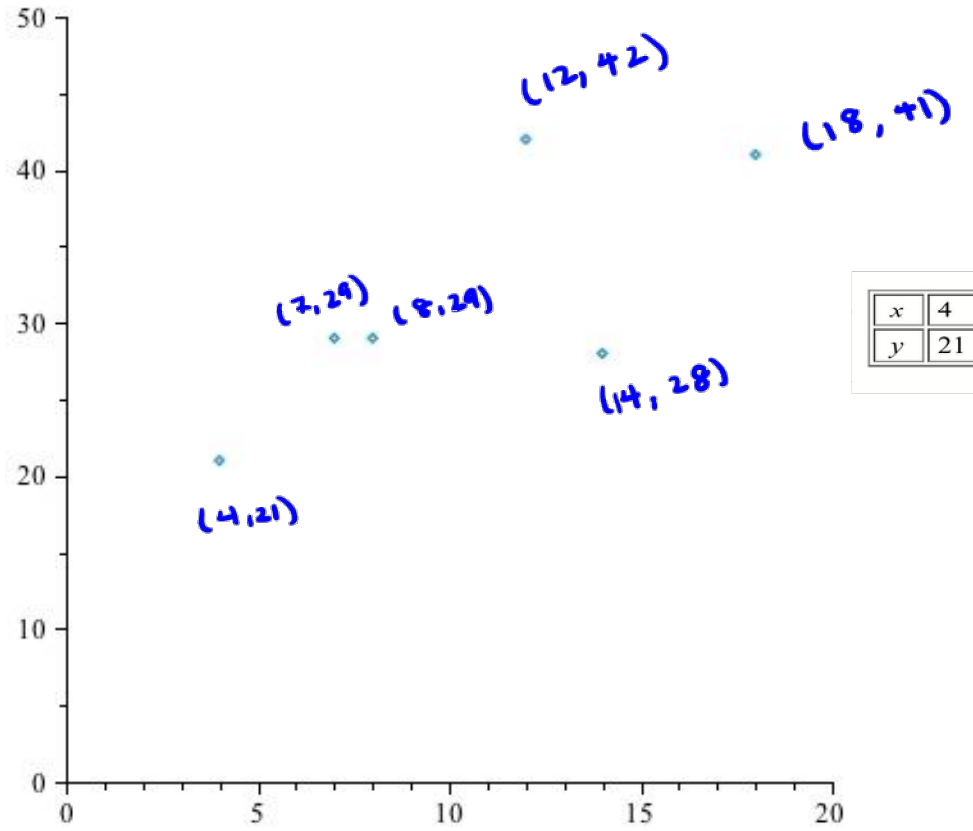


b)




c)






| $x$ | 4 | 7 | 8 | 12 | 14 | 18 |
|---|---|---|---|---|---|---|
| $y$ | 21 | 29 | 29 | 42 | 28 | 41 |

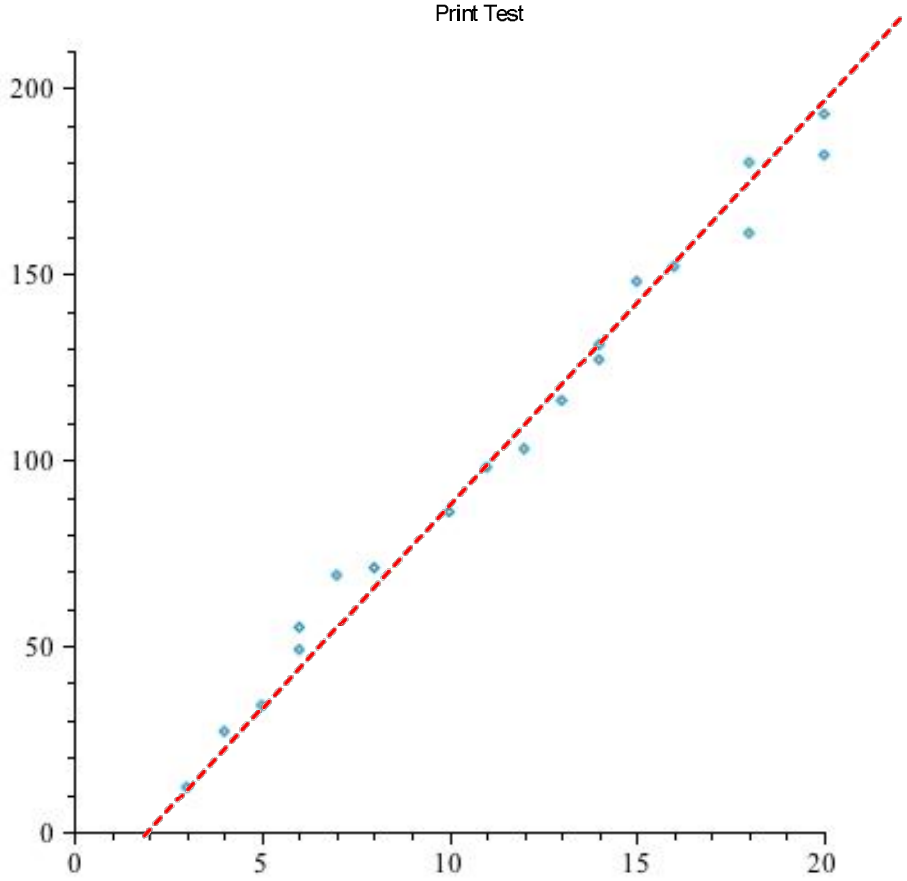
**d)**

**e)** None of the above

**Question 2**

Your answer is CORRECT.

What can be said about the relationship between the explantory variable and the response variable in the following scatterplot?




**a)** There is a weak positive linear association.

**b)** The explanatory variable causes the responses.

**c)** There is a strong positive linear association.

**d)** There is a strong negative linear association.

**e)** None of the above

**Question 3**

Your answer is CORRECT.

Determine the correlation coefficient for the data shown in this table:

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ | 2 | 7 | 8 | 11 | 17 | 16 |
| $y$ | 23 | 28 | 29 | 40 | 28 | 42 |

```
x=c(2,7,8,11,17,16)
> y=c(23,28,29,40,28,42)
> cor(x,y)
[1] 0.6046549
```

**a)** 0.6047

**b)** 0.3656

**c)** 0.3023

**d)** -0.3656

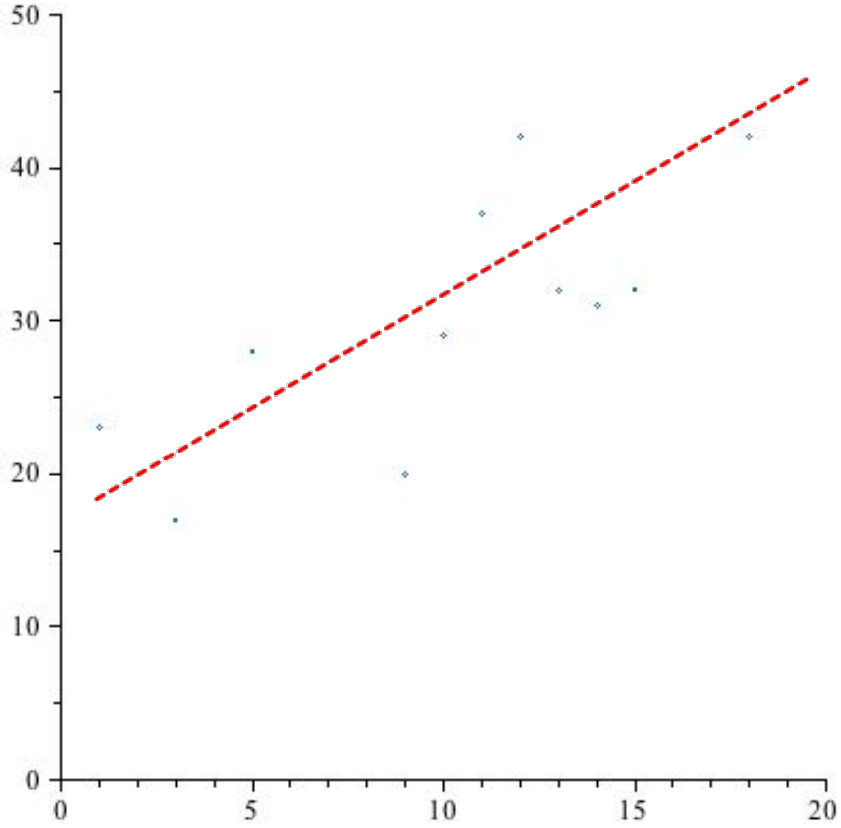
**e)** -0.6047

**f)** None of the above

**Question 4**

Your answer is CORRECT.

Choose the best correlation coefficient for the data shown in this scatterplot:




**a)** -0.7483

**b)** 0.7483

**c)** 1.4513

**d)** -0.3742

**e)** 0.3742

**Question 5**

Your answer is CORRECT.

Which of the following is a true statement? see notes for section 5.3 & 5.4

**a)** The variable that is being predicted in regression analysis is the independent variable.

**b)** If there is no correlation between the independent and dependent variables, then the value of the correlation coefficient must be -1.

**c)** The correlation coefficient $r$ is always between -1 and +1.

**d)** The coefficient of determination can assume negative values.

**e)** A negative correlation indicates that as values of $x$ decrease, values of $y$ will decrease.

**f)** None of the above

**Question 6**

Your answer is CORRECT.

The decline of salmon fisheries along the Columbia River in Oregon has caused great concern among commercial and recreational fishermen. The paper 'Feeding of Predaceous Fishes on Out-Migrating Juvenile Salmonids in John Day Reservoir, Columbia River' (Trans. Amer. Fisheries Soc. (1991: 405-420)) gave the accompanying data on y = maximum size of salmonids consumed by a northern squaw fish (the most abundant salmonid predator) and x = squawfish length, both in mm. Use the following statistics to give the equation of the least squares regression line.

$$\bar{x} = 525.180, \quad \bar{y} = 372.258, \quad s_x = 13.431, \quad s_y = 12.100, \quad r = 0.9768$$

$b = r\left(\frac{s_y}{s_x}\right) = 0.9768\left(\frac{12.1}{13.431}\right)$

$b = 0.88$

$a = \bar{y} - b\bar{x} = 372.258 - (0.88)(525.2)$

$a = -89.92$

$\hat{y} = a + bx$

$\hat{y} = -89.92 + 0.88x$

**a)** $\hat{y} = 0.880x + 89.900$

**b)** $\hat{y} = -89.900x + 0.880$

**c)** $\hat{y} = 1.084x - 89.900$

**d)** $\hat{y} = 1.084x + 89.900$

**e)** $\hat{y} = 0.880x - 89.900$

**f)** None of the above

**Question 7**

Your answer is CORRECT.

Suppose that you are given the following results. Find the correlation coefficient of the data.

$s_x = 1.791,\ \ s_y = 13.400,\ \ b = -6.480$

$b = r\left(\frac{s_y}{s_x}\right)$

**a)** 0.155

**b)** -0.433

$\frac{(1.791)(-6.480)}{13.400} = r$

**c)** -0.866

$r = -0.866$

**d)** 0.866

**e)** -0.155

**f)** None of the above

**Question 8**

Your answer is CORRECT.

Suppose you find that the correlation coefficient for a set of data is -0.824. What is the coefficient of determination and what does it mean?

$r^2 * 100\% = (-0.824)^2 * 100\%$

**a)** -0.824; This means that we are 82.4% accurate with our prediction of the LSRL equation.

**b)** 0.824; This means that 82.4% of the variation of $y$ is explained by the LSRL of $y$ on $x$.

**c)** 0.679; This means that we are 67.9% accurate with our prediction of the LSRL equation.

**d)** 0.679; This means that 67.9% of the variation of $y$ is explained by the LSRL of $y$ on $x$.

$= 67.9\%$

**e)** None of the above

**Question 9**

Your answer is CORRECT.

Suppose that the LSRL for the appraised value (in thousands of dollars) and number of rooms for houses in East Meadow, New York is $\hat{y} = 19.718x + 74.80$. Predict the price of a 14 room house (in thousands of dollars).
tbl

$\hat{y} = 19.718(14) + 74.80$

$\hat{y} = 350.852$

**a)** 1052.556

**b)** 454.852

**c)** 355.852

**d)** 464.852

**e)** 350.852

**f)** None of the above

**Question 10**

Your answer is CORRECT.

Select the equation of the least squares line for the data: (51.00, 1.30), (48.75, 3.25), (52.50, .65), (46.50, 6.50), (45.00, 5.85), (41.25, 8.45), (43.50, 6.50).

```
 x=c(51,48.75,52.5,46.5,45,41.
25,43.5)
>
y=c(1.30,3.25,.65,6.50,5.85,8.4
5,6.50)
> lm(y~x)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)          x
   37.6429      -0.7032
```

**a)** $\hat{y} = 41.407 - 0.77316x$

**b)** $\hat{y} = 37.643 - 0.77316x$

**c)** $\hat{y} = 37.643 - 0.70287x$

**d)** $\hat{y} = 0.70287x - 37.643$

**e)** $\hat{y} = -37.643 - 0.70287x$

**f)** None of the above

# PRINTABLE VERSION

## Quiz 8

You scored 100 out of 100

**Question 1**

Your answer is CORRECT.

Which of the following would be the LSRL for the given data?

| $x$ | 1 | 4 | 8 | 12 | 15 | 19 |
|---|---|---|---|---|---|---|
| $y$ | 43 | 37 | 26 | 20 | 24 | 11 |

```
x=c(1,4,8,12,15,19)
> y=c(43,37,26,20,24,11)
> lm(y~x)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)          x
     42.815      -1.625
```

**a)** $\hat{y} = 42.82 + 1.625\,x$

**b)** $\hat{y} = 42.82 - 1.625\,x$

**c)** $\hat{y} = -1.625 + 42.82\,x$

**d)** $\hat{y} = -1.625 - 42.82\,x$

**e)** None of the above

**Question 2**

Your answer is CORRECT.

Determine the correlation coefficient for the data shown in this table:

| $x$ | 3 | 6 | 11 | 13 | 16 | 18 |
|---|---|---|---|---|---|---|
| $y$ | 22 | 28 | 27 | 40 | 27 | 43 |

```
x=c(3,6,11,13,16,18)
> y=c(22,28,27,40,27,43)
> cor(x,y)
[1] 0.7018695
```

**a)** -0.7019

**b)** 0.7019

**c)** 0.3509

**d)** -0.4926

**e)** 0.4926

**f)** None of the above

**Question 3**

Your answer is CORRECT.

Suppose you have the following data:

Residual = $y - \hat{y}$

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $y$ | 24 | 29 | 28 | 42 | 33 | 41 |

and the LSRL is $\hat{y} = 21.73 + 3.171\,x$. Find the residual value for $x = 1$.

$y(1) - \hat{y}(1) = 24 - 24.091 = -0.901$

**a)** 24

**b)** -0.901

$\hat{y}(1) = 21.73 + 3.171 = 24.901$

**c)** 24.901

**d)** 0.901

**e)** None of the above

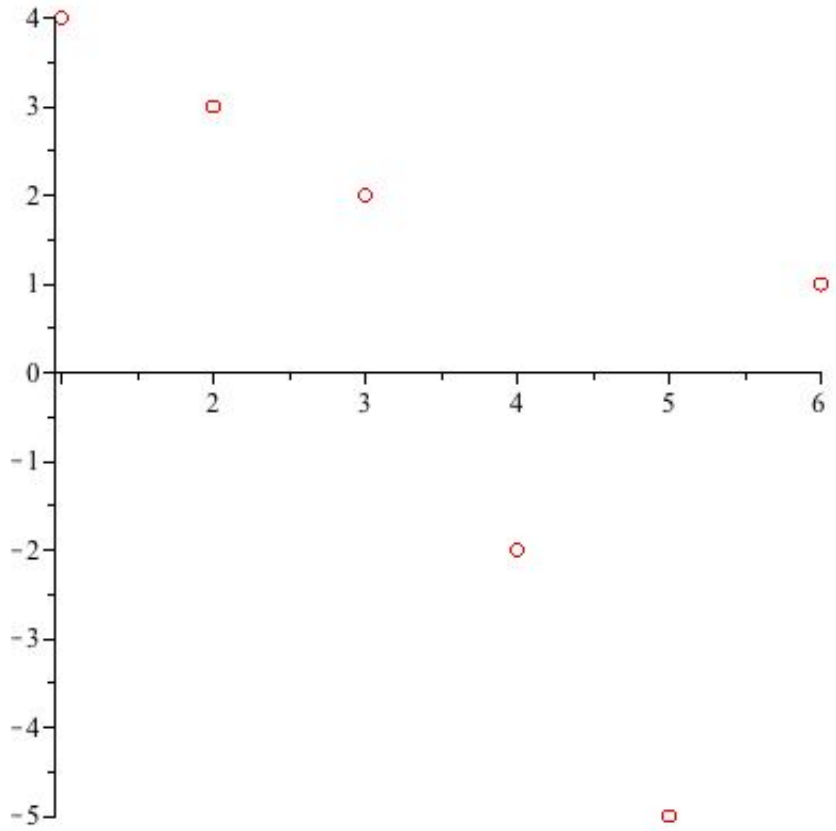
**Question 4**

Your answer is CORRECT.

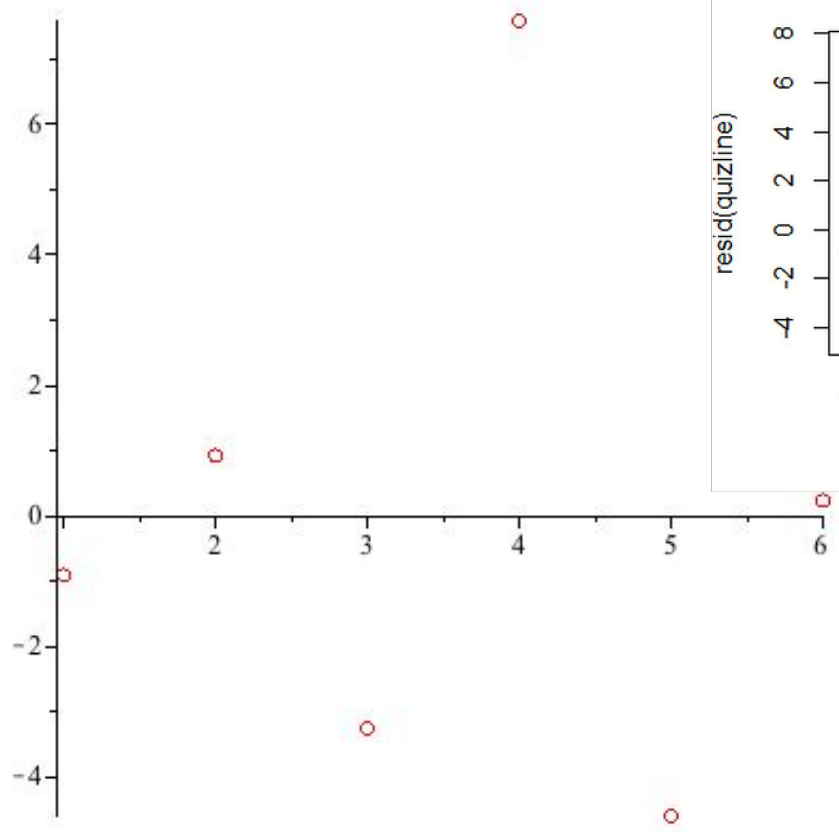
Which of the following is the residual plot for the data in the given table?

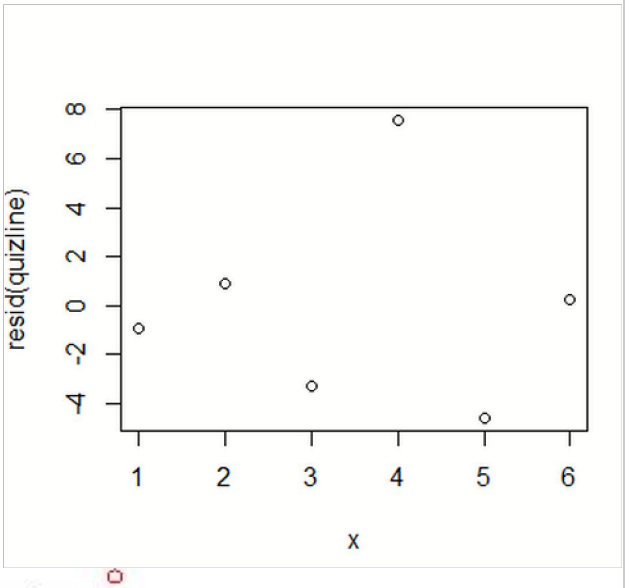
| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $y$ | 14 | 19 | 18 | 32 | 23 | 31 |

```
x=c(1,2,3,4,5,6)
> y=c(14,19,18,32,23,31)
> quizline=lm(y~x)
> plot(x,resid(quizline))
```

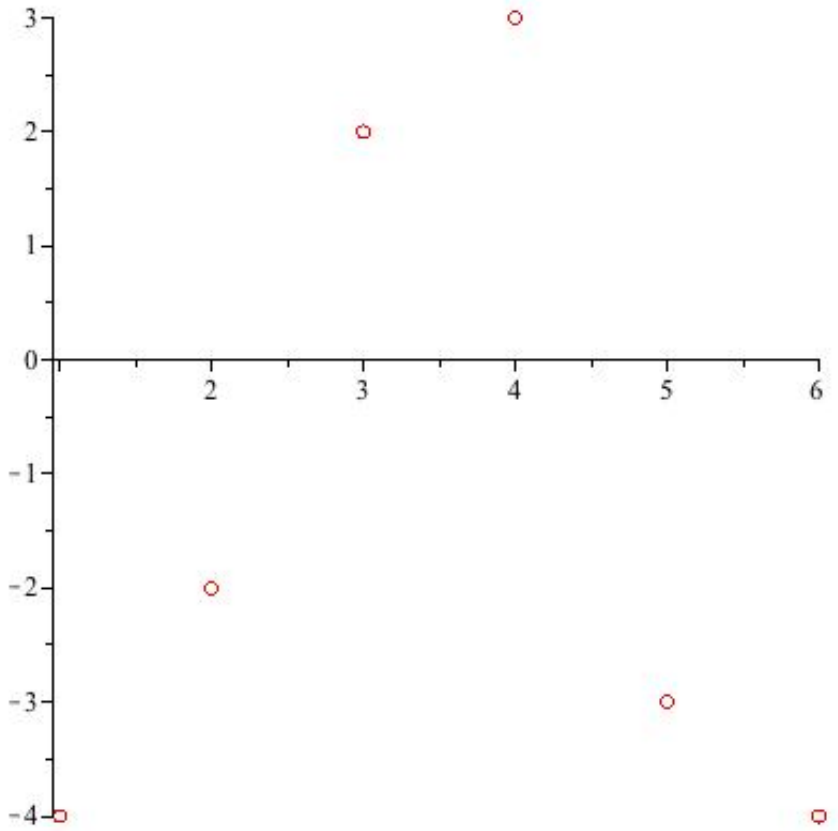

4
3
2
1
0
-1
-2
-3
-4
-5
2
3
4
5
6

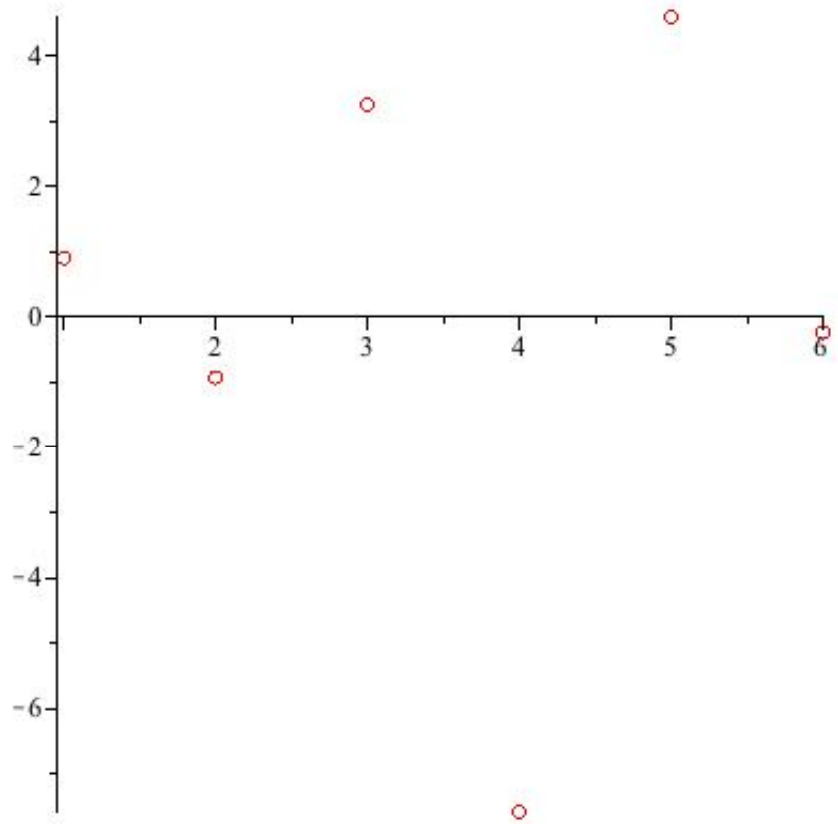
**a)**


6
4
2
0
-2
-4
2
3
4
5
6


resid(quizline)
8
6
4
2
0
-2
-4
1
2
3
4
5
6
x

**b)**

**c)** 

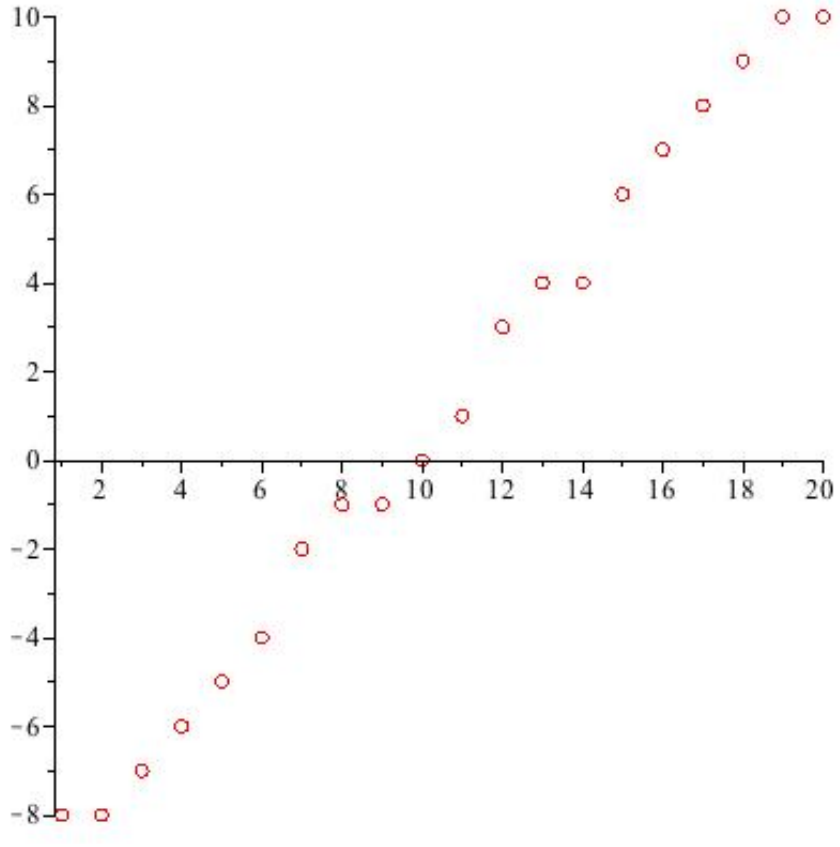

**d)** 


**e)** None of the above

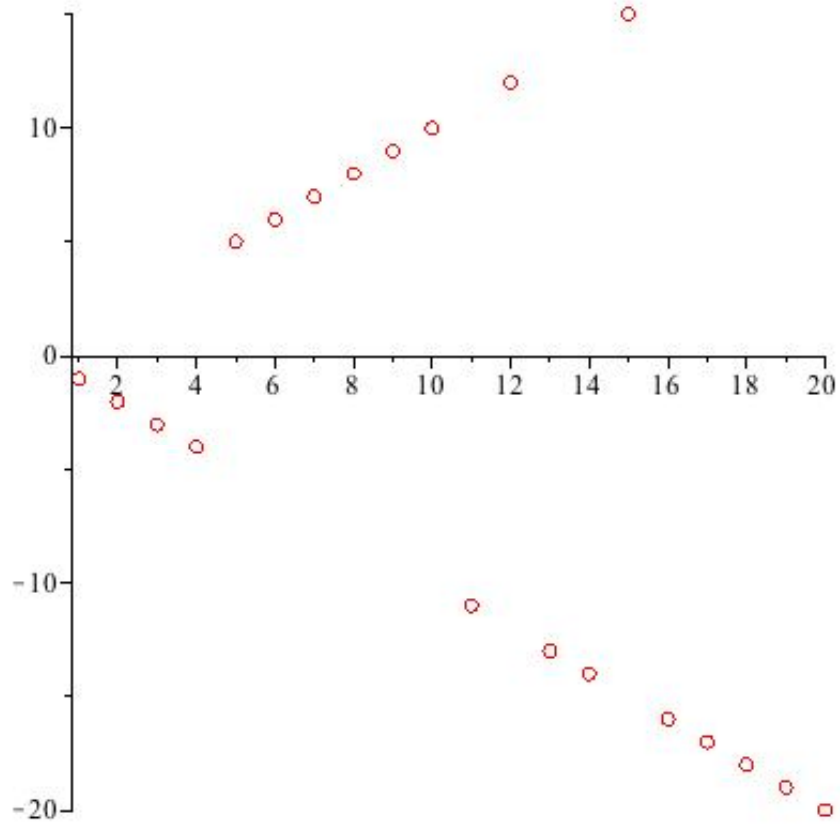
## Question 5

Your answer is CORRECT. → no pattern

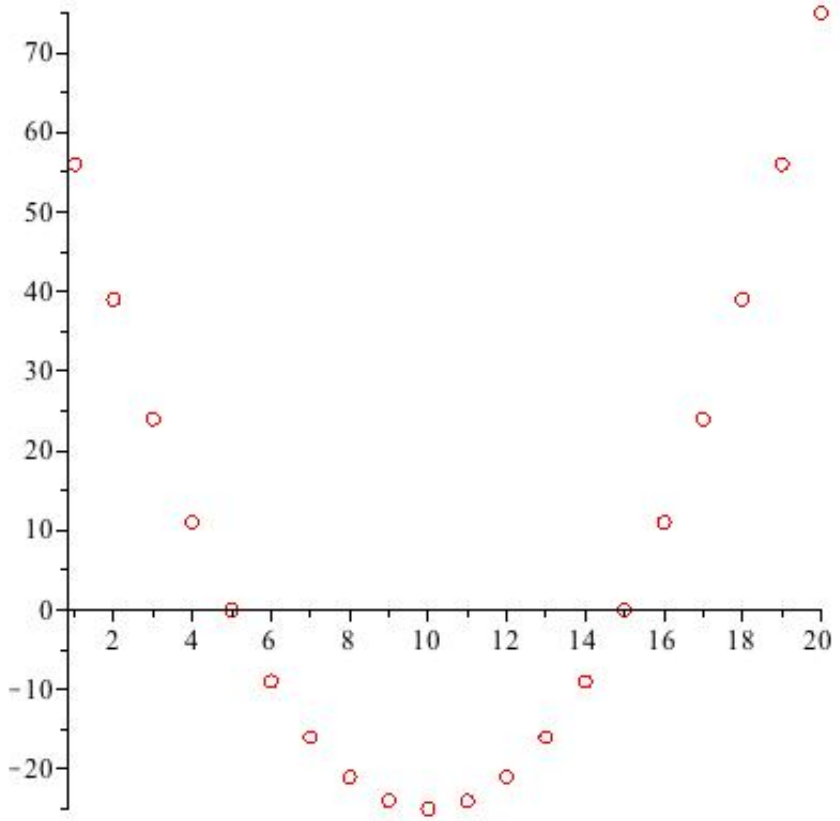
Which of the following residual plots would indicate a good LSRL model?



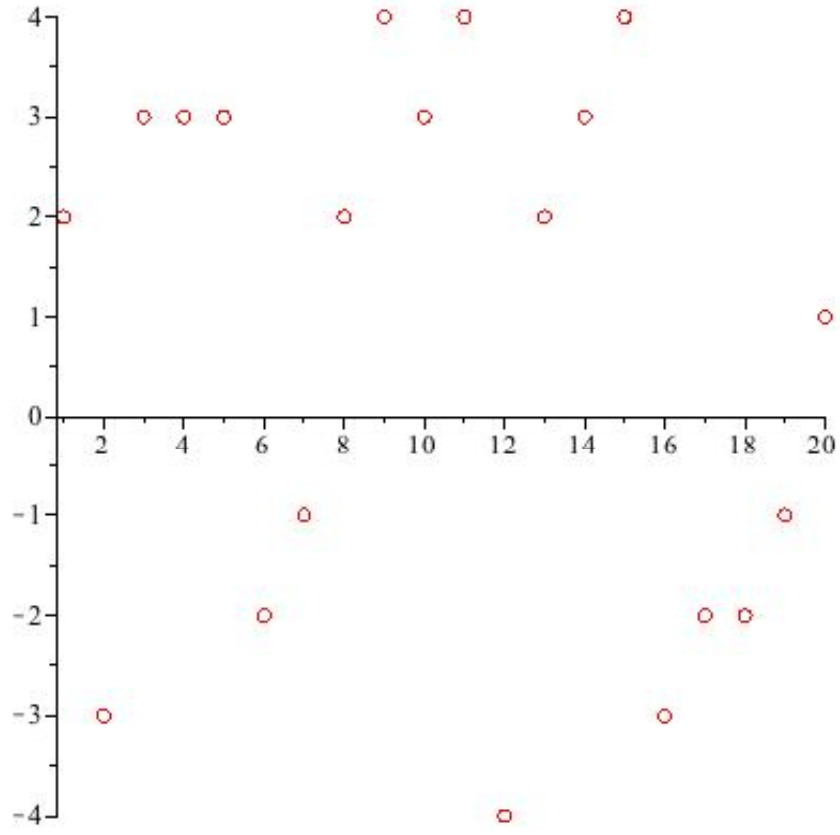

**a)**




**b)**




**c)**




no pattern

**d)**

**Question 6**

Your answer is CORRECT.

For children between the ages of 18 months and 29 months, there is approximately a linear relationship between *height* and *age*. The relationship can be represented by $\hat{y} = 61.05 + 0.63x$ where $y$ represents height (in centimeters) and $x$ represents age (in months). Joseph is 23.5 months old and is 82 centimeters tall. What is Joseph's residual?

Residual = $y - \hat{y}$

$y(23.5) - \hat{y}(23.5) = 82 - 75.855 = 6.145$

$\hat{y}(23.5) = 61.05 + 0.63(23.5)$

$= 75.855$

**a)** 75.855

**b)** 7.145

**c)** -30.710

**d)** 112.710

**e)** 6.145

**f)** None of the above

**Question 7**

Your answer is CORRECT.

If the LSRL relating the independent variable $x$ and the dependent variable $y$ for a given problem is $\hat{y} = 3x + 4$, then an increase of 1 unit in $x$ is associated with an increase of how many units in $y$?

$\hat{y} = 3x + 4$

↑ slope

for every x increase there is a 3 unit increase in y

**a)** 4

**b)** 1

**c)** 0

**d)** 3

**e)** 7

**Question 8**

Your answer is CORRECT.

If the correlation between body weight and annual income were high and positive, we could conclude that:

never say causation

body weight ↑

annual income ↑

never say causation

**a)** high incomes cause people to gain weight.

**b)** high incomes cause people to eat more food.

**c)** high-income people tend to spend a greater proportion of their income on food than low-income people.

**d)** high-income people tend to be heavier than low-income people.

never say causation

**e)** low incomes cause people to eat more food.

**Question 9**

Your answer is CORRECT.

The following two-way table describes the preferences in movies and fast food restaurants for a random sample of 100 people.

| | McDonalds | Taco Bell | Wendy's |
|---|---|---|---|
| Iron Man | 20 | 12 | 8 |
| Dispicable Me | 10 | 11 | 7 |
| Harry Potter | 5 | 15 | 12 |

What percent of people in the sample like the movie Iron Man?

$\frac{20+12+8}{100} = \frac{40}{100} = 40\%$

**a)** 28%

**b)** 27%

**c)** 35%

**d)** 38%

**e)** 40%

**Question 10**

Your answer is CORRECT.

The following two-way table describes the preferences in movies and fast food restaurants for a random sample of 100 people.

| | McDonald's | Taco Bell | Wendy's | |
|---|---|---|---|---|
| Iron Man | 20 | 13 | 7 | 40 |
| Dispicable Me | 10 | 11 | 7 | 28 |
| Harry Potter | 9 | 11 | 12 | 32 |
| | 39 | 35 | 26 | |

What percent of the Iron Man lovers also like Taco Bell?

$\frac{13}{40} = 32.5\% \approx 33\%$

**a)** 40%

**b)** 11%

**c)** 19%

**d)** 24%

**e)** 33%