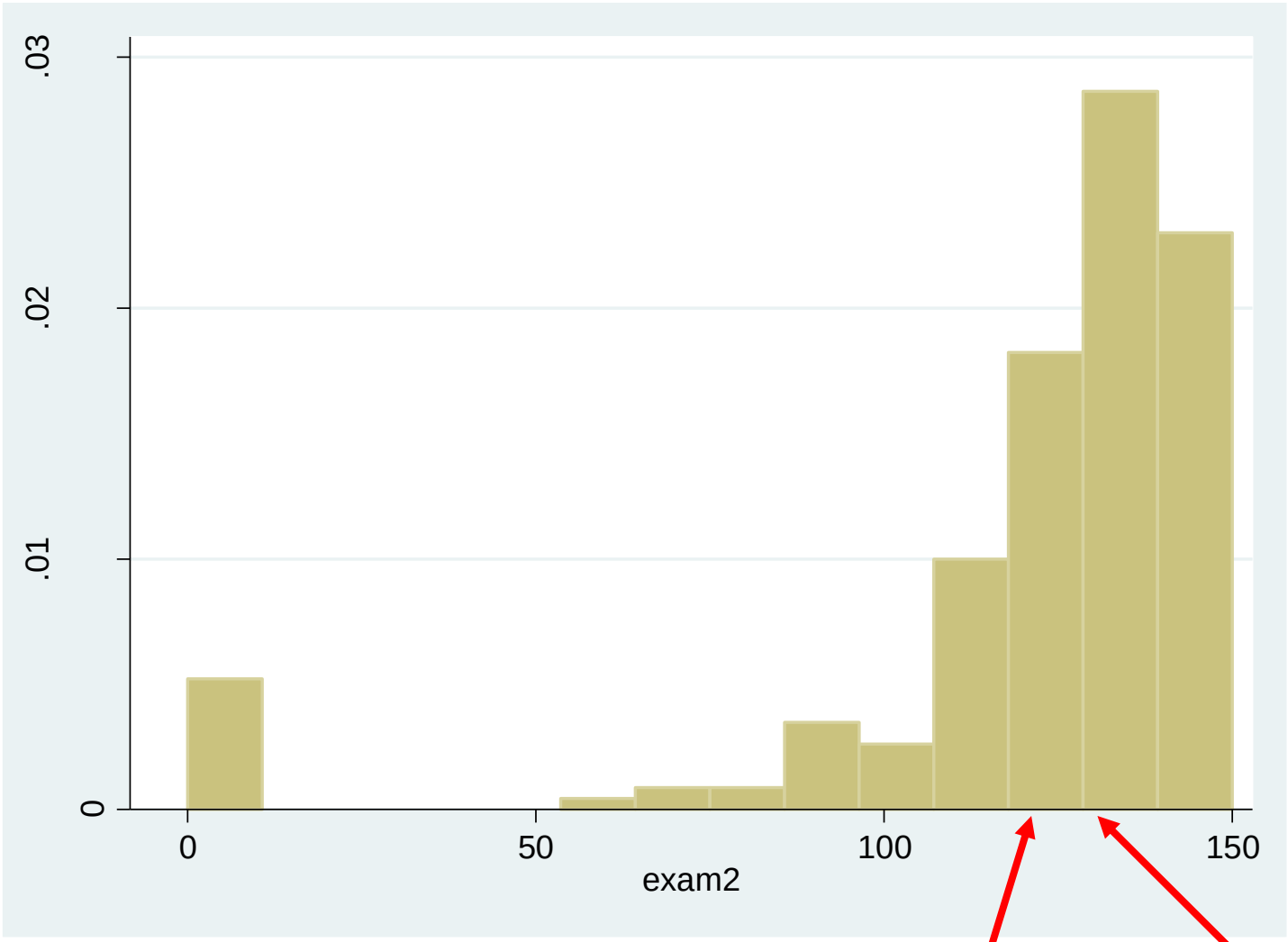
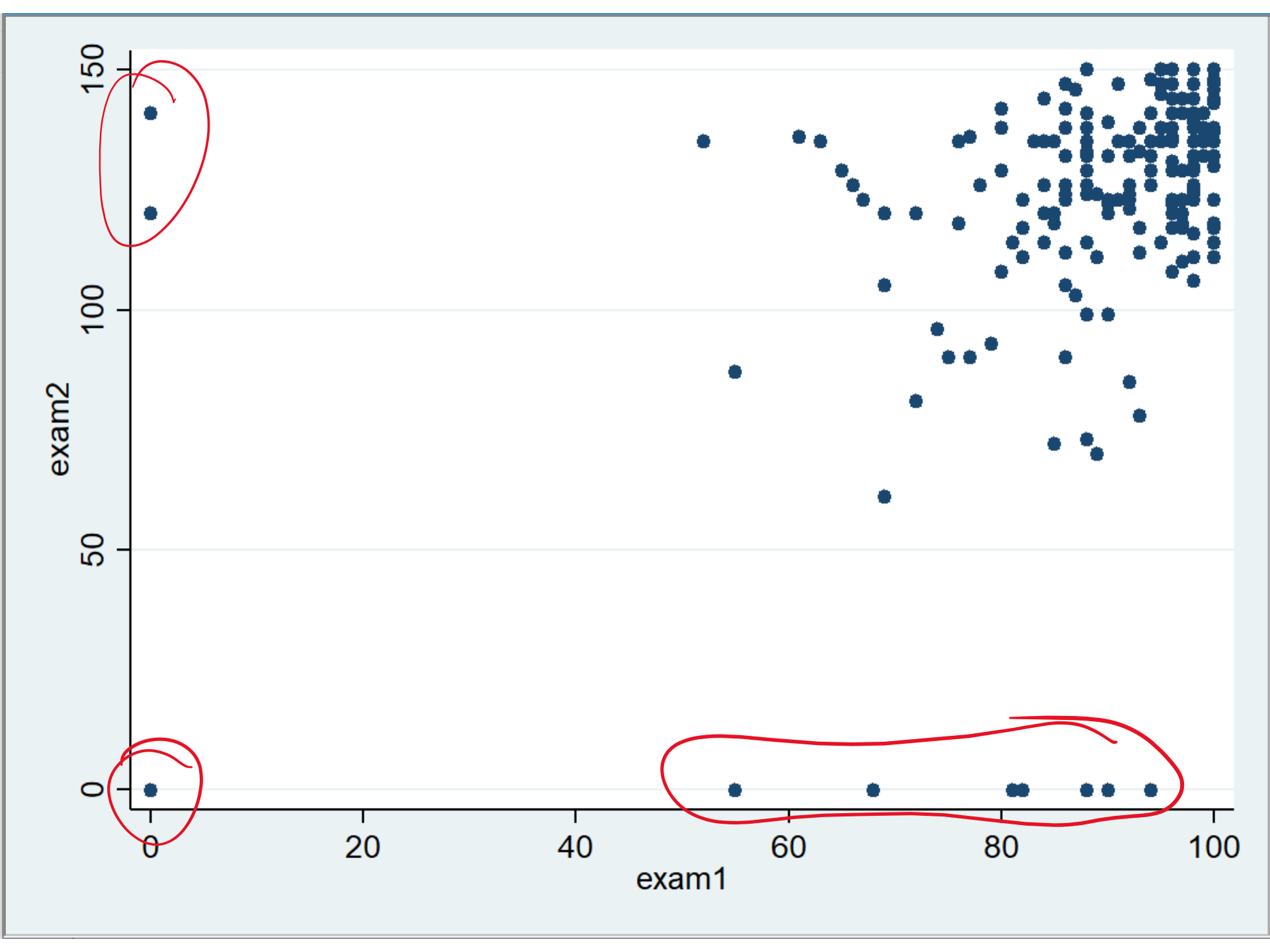
SOC 3811/5811:
BASIC SOCIAL STATISTICS

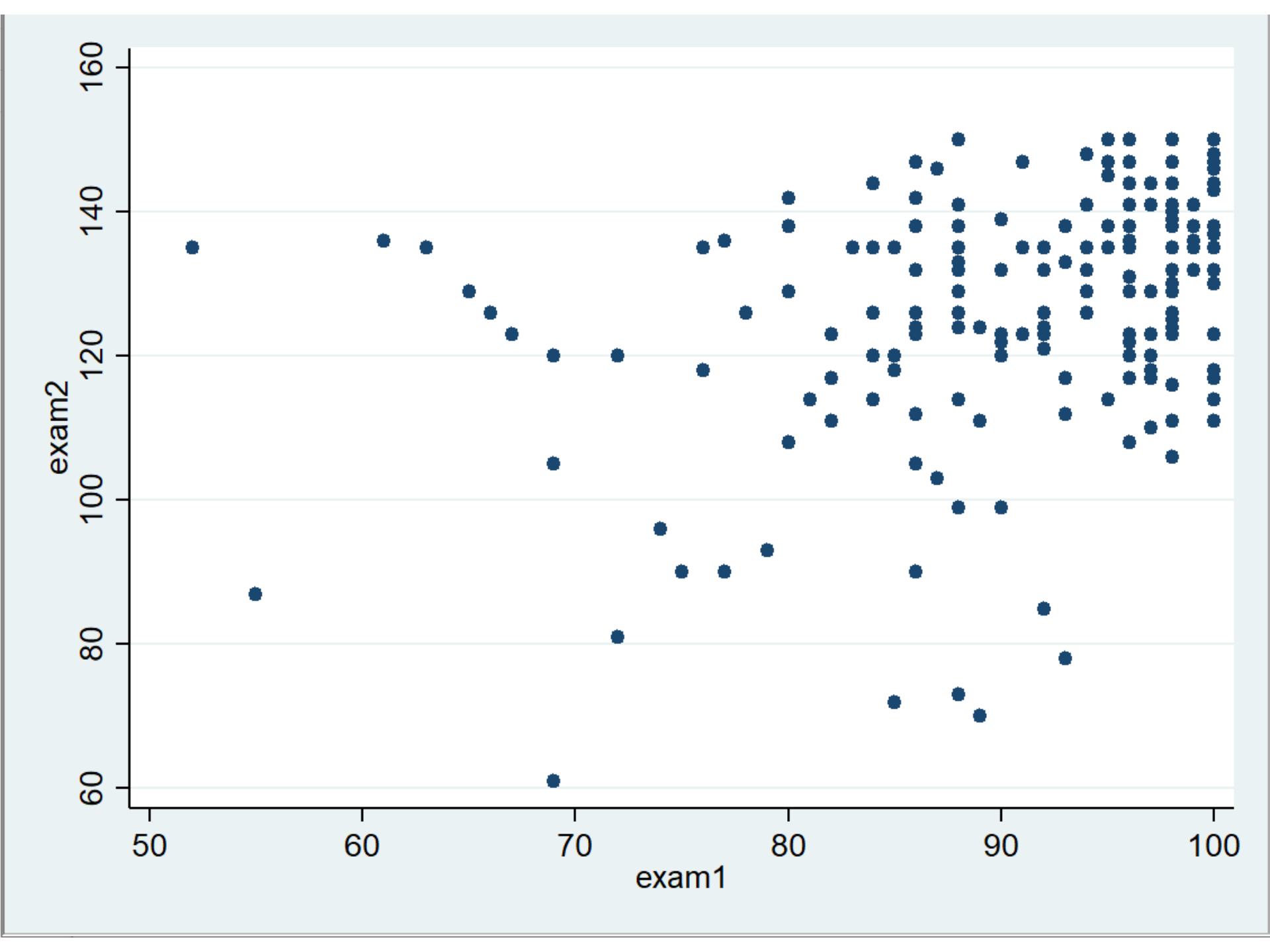
Associations Between Continuous Variables

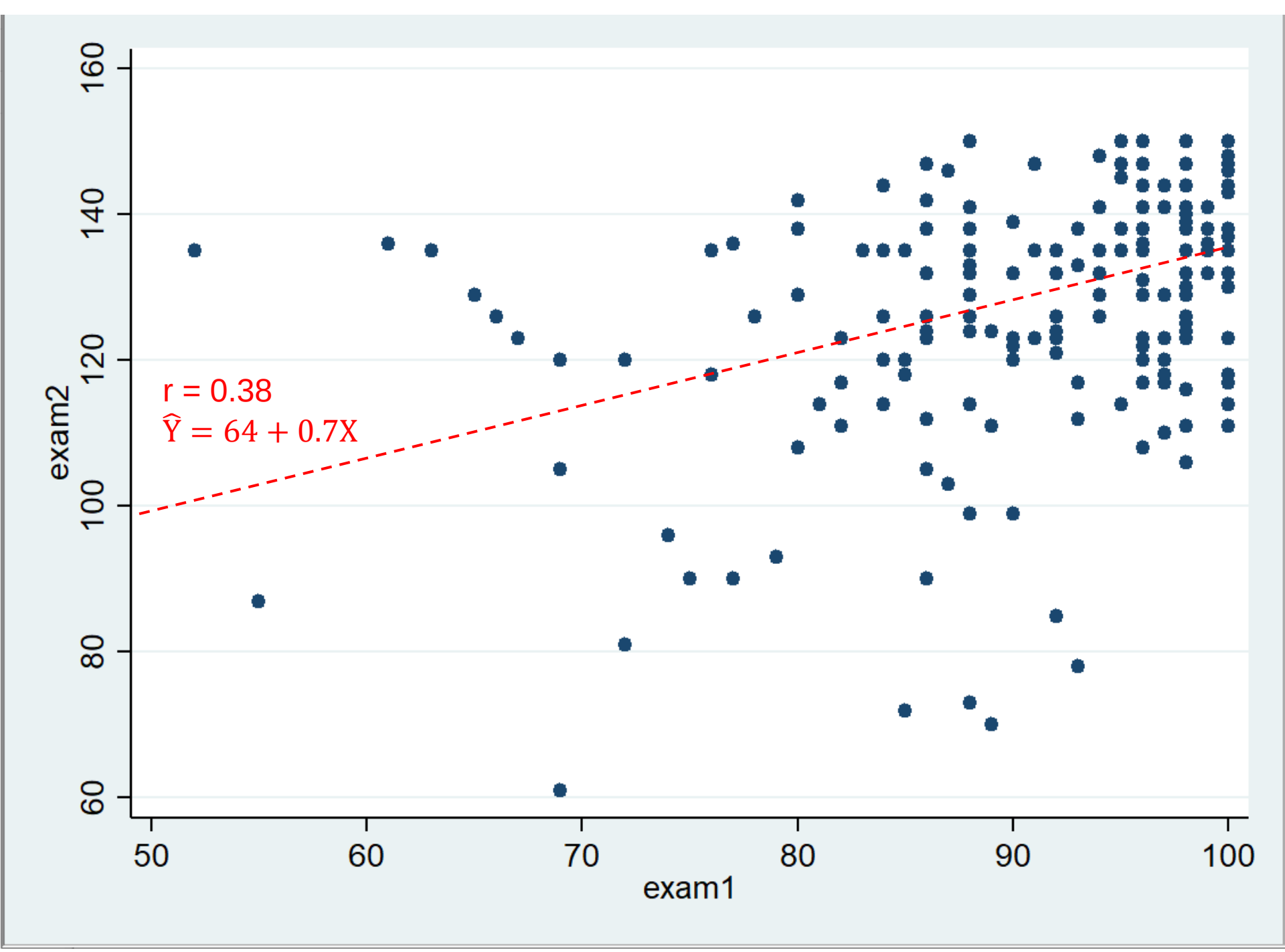


Mean = 121
(81%)
B-

Median = 132
(88%)
B+


150
100
50
0
exam2
0
20
40
60
80
100
exam1


160
140
120
100
80
60
exam2
50
60
70
80
90
100
exam1


r = 0.38
$\hat{Y} = 64 + 0.7X$
exam2
160
140
120
100
80
60
50
60
70
80
90
100
exam1

Recommended Formulas

The easiest way (and order in which) to compute the statistics we are covering today:

Compute the mean and standard deviation of each variable

Compute r_{YX} as $r_{YX} = \left(\frac{1}{N-1}\right)\sum_{i=1}^{N}\left(\frac{X_i - \overline{X}}{s_x}\right)\left(\frac{Y_i - \overline{Y}}{s_y}\right)$

Compute the slope b_{YX} as $b_{YX} = r_{YX}\frac{s_Y}{s_X}$

Compute the intercept a as $a = \overline{Y} - b\overline{X}$

Compute R^2_{YX} as $R^2_{YX} = r^2_{YX}$

Mean of X = 3.5

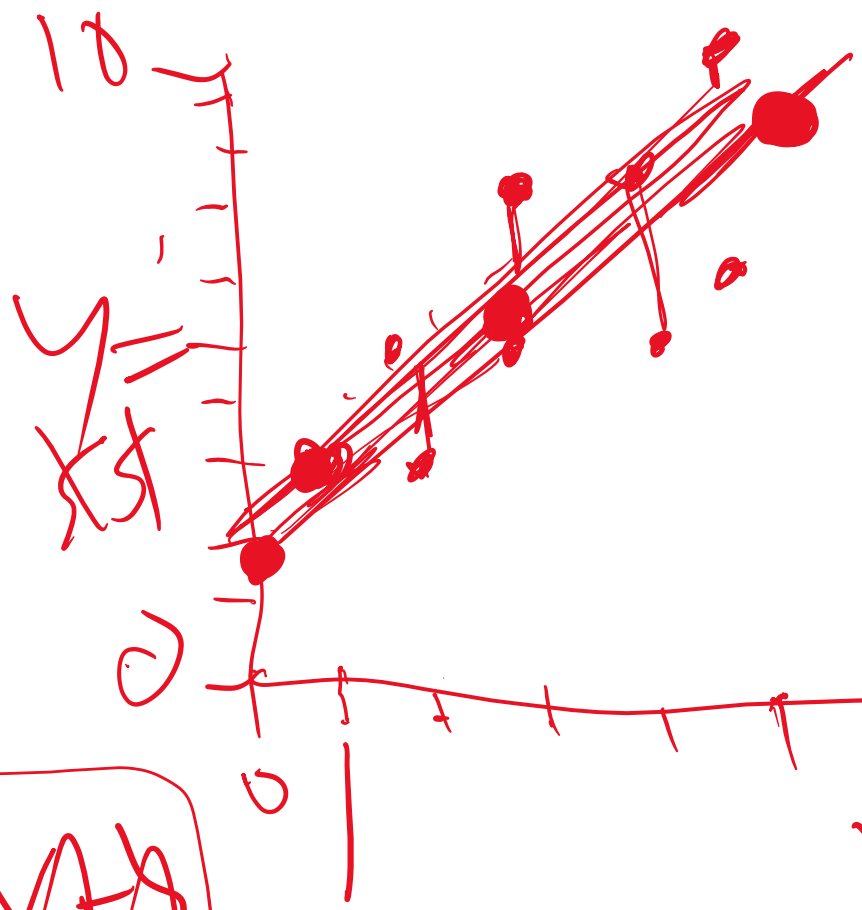
SD of X = 1.195

Mean of Y = 6.0

SD of Y = 2.070

$$r_{YX} = \left(\frac{1}{N-1}\right)\sum_{i=1}^{N}\left(\frac{X_i - \bar{X}}{s_x}\right)\left(\frac{Y_i - \bar{Y}}{s_y}\right)$$

X	Y
2	3
2	5
3	7
3	5
4	7
4	5
5	6
5	10



$\hat{y} = a + bx$

$\hat{y} = 1.8 + 1.2x$

$3 = 1.8 + (0.2)(1)$

Mean of X = 3.5 SD of X = 1.195

Mean of Y = 6.0 SD of Y = 2.070

$$r_{YX} = \left(\frac{1}{N-1}\right)\sum_{i=1}^{N}\left(\frac{X_i - \bar{X}}{s_x}\right)\left(\frac{Y_i - \bar{Y}}{s_y}\right)$$

X	Y
2	3
2	5
3	7
3	5
4	7
4	5
5	6
5	10

$\left(\frac{2-3.5}{1.195}\right)\left(\frac{3-6}{2.07}\right) =$

$\left(\frac{5-3.5}{1.195}\right)\left(\frac{10-6}{2.07}\right) =$

$r_{YX} = \left(\frac{1}{7}\right)(4.851) = .693$

$-1 \text{ to } 1$

$.693^2 = .48$

4.851

Mean of X = 3.5 SD of X = 1.195 r_{YX} = 0.693

Mean of Y = 6.0 SD of Y = 2.070

X	Y
2	3
2	5
3	7
3	5
4	7
4	5
5	6
5	10

$\hat{y} = 1.8 + 1.2(3)$

$\hat{y} = 5.4$

$\hat{y} = 1.8 + 1.2x$

$\hat{y} = 1.8 + (1.2)(0)$

$\hat{y} = 1.8$

$$b_{YX} = r_{YX}\frac{s_Y}{s_X} = .693\frac{2.070}{1.195} = 1.2$$

$$a = \bar{Y} - b\bar{X} = 6 - (1.2)(3.5) = 1.8$$

$$R^2_{YX} = r^2_{YX} = .693^2 = .48$$

y = what you saw

$\hat{y}$ = Predicted based on regression

Mean of X = 3.5

SD of X = 1.195

$r_{YX} = 0.693$

Mean of Y = 6.0

SD of Y = 2.070

X	Y
2	3
2	5
3	7
3	5
4	7
4	5
5	6
5	10

$$b_{YX} = r_{YX}\frac{s_Y}{s_X} = 0.693\,\frac{2.070}{1.195} = 1.2$$

$$a = \overline{Y} - b\overline{X} = 6.0 - 1.2(3.5) = 1.8$$

$$R^2_{YX} = r^2_{YX} = 0.693^2 = 0.48$$

Week 9 - Day 1

* Required

What is the chance that in the next year you personally will be infected with the coronavirus that causes the disease COVID-19? *

Your answer should be some number between 0 (absolutely no chance) and 100 (absolutely certain). Note that if you have already been infected, answer the question anyway ... about the risk that you will be infected again in the next year.

Your answer

If you were to become infected in the next year with the coronavirus that causes the disease COVID-19, what is the chance that you would experience symptoms so severe that you would have to be hospitalized? *

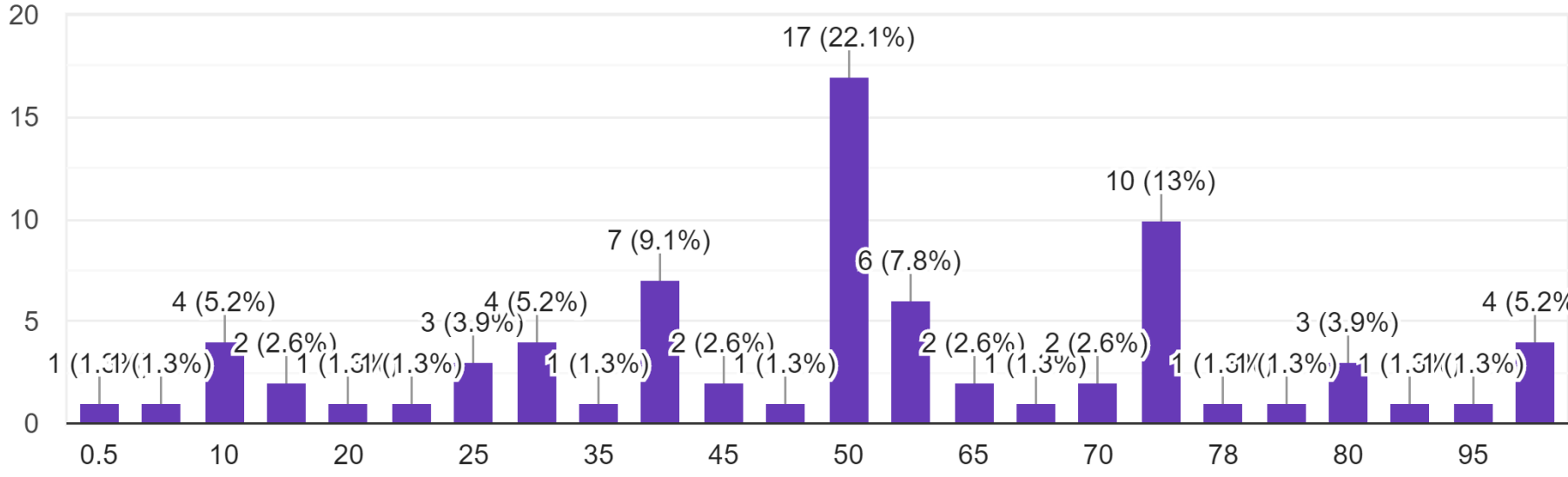
Your answer should be some number between 0 (absolutely no chance) and 100 (absolutely certain). Note that if you have already been infected, answer the question anyway ... about the risk of hospitalization if you were to be infected again in the next year.

Your answer

Submit

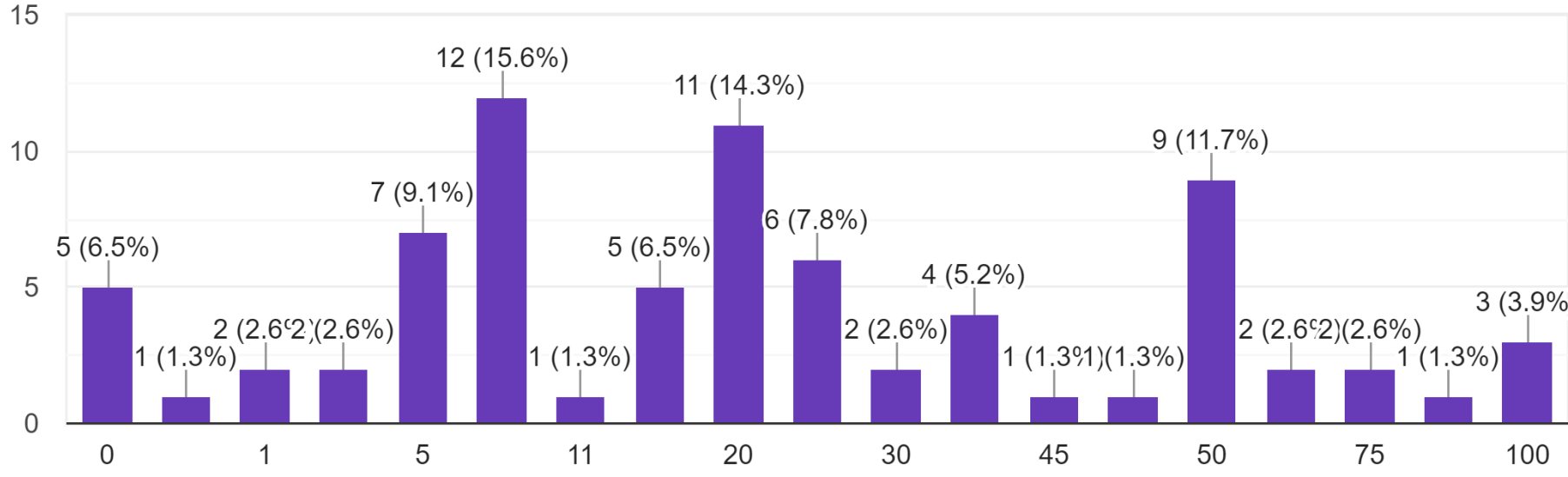
What is the chance that in the next year you personally will be infected with the coronavirus that causes the disease COVID-19?

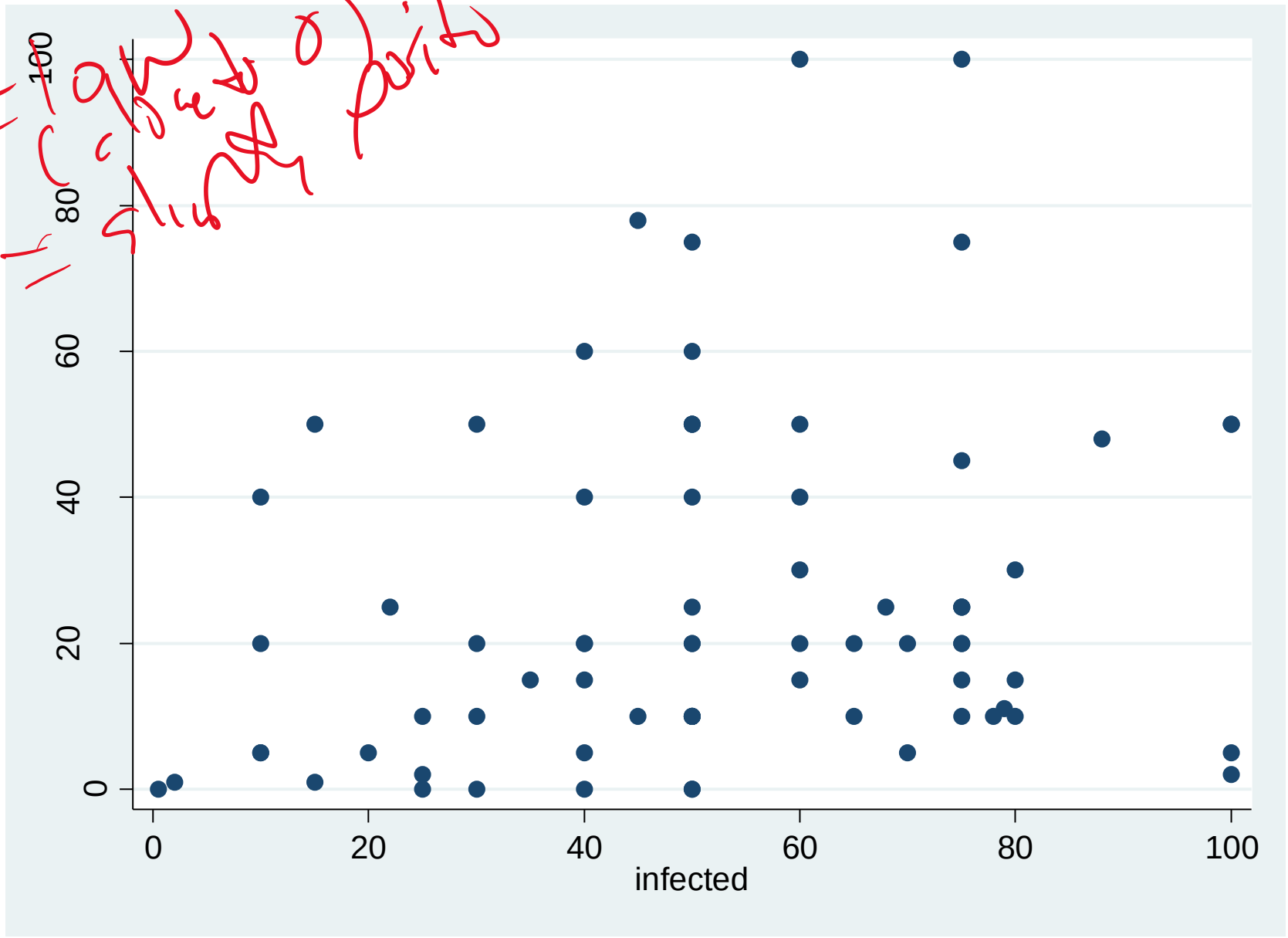
77 responses



If you were to become infected in the next year with the coronavirus that causes the disease COVID-19, what is the chance that you would exper...o severe that you would have to be hospitalized?

77 responses




0
20
40
60
80
100
0
20
40
60
80
100
infected

n=78

FOR COVID-19 CLASS EXAMPLE

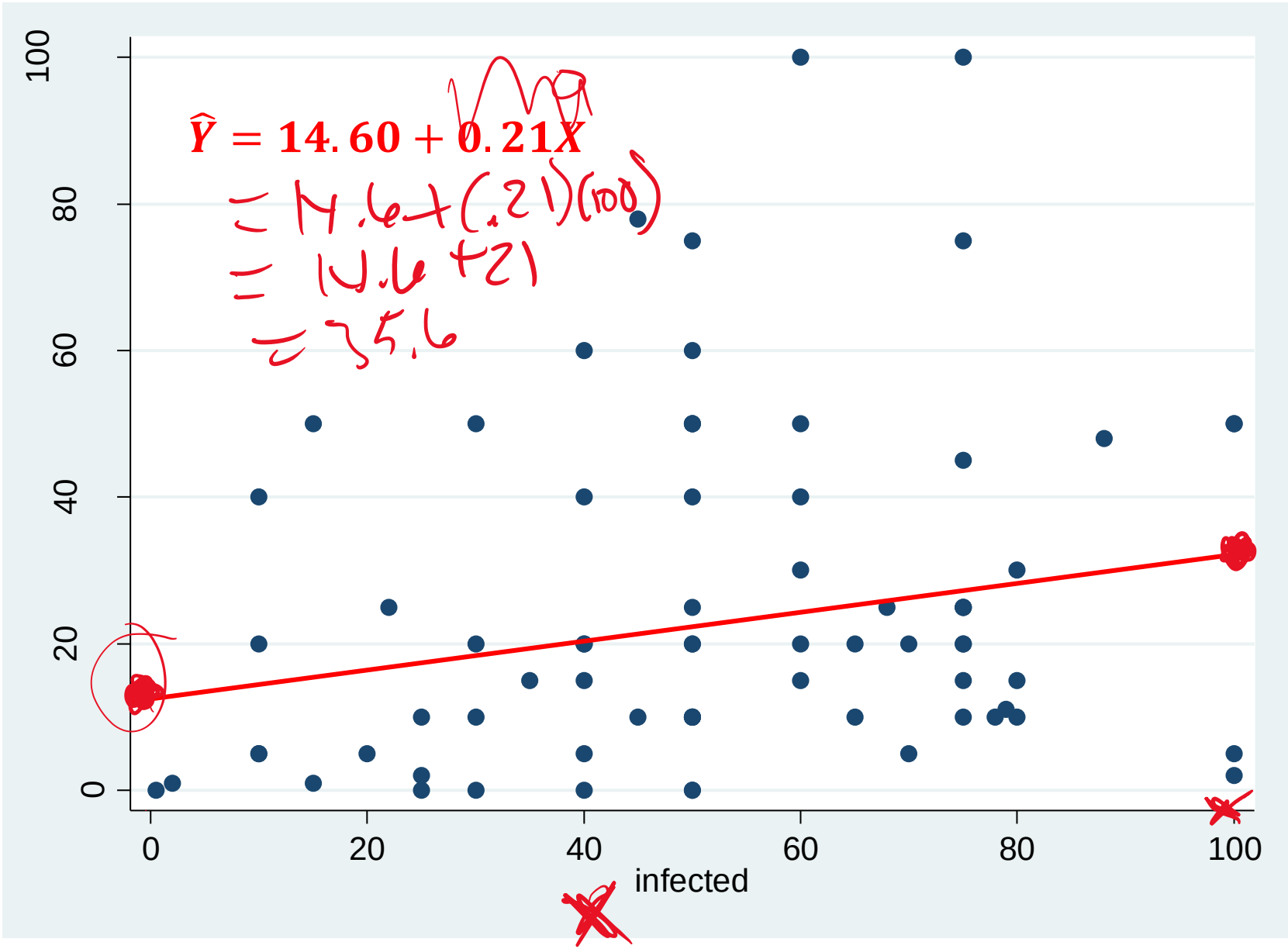
Mean of X = 52.0 SD of X = 24.4 r_{YX} = 0.21

Mean of Y = 25.3 SD of Y = 23.5

$$b_{YX} = r_{YX}\frac{s_Y}{s_X} = (.21)\frac{23.5}{24.4} = .21$$

$$a = \bar{Y} - b\bar{X} = 25.3 - (.21)(52) = 14.60$$

$$R^2_{YX} = r^2_{YX} = .21^2 = .046$$


$\hat{Y} = 14.60 + 0.21X$
$= 14.6 + (.21)(100)$
$= 14.6 + 21$
$= 35.6$
0
20
40
60
80
100
infected