

Psych 315, Winter 2021, Homework 3 Answer Key

Due Friday, January 22 by 5pm.

Name _____ ID _____

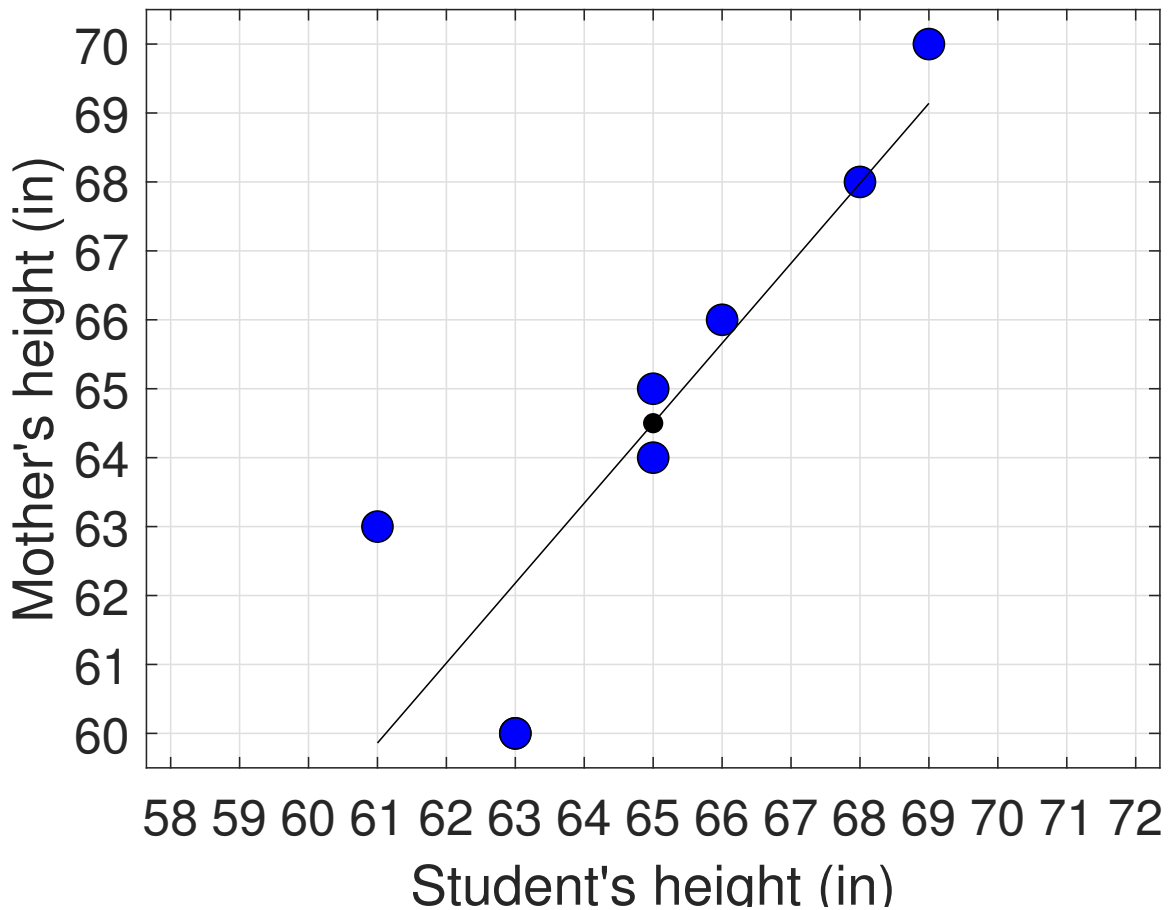
Section [AA] (Natalie), [AB] (Natalie), [AC] (Ryan), [AD] (Ryan), [AE] (Kelly), [AE] (Kelly)

The following is a table containing the heights of the 8 Female students in this class that chose Yellow for their favorite color, along with their mother's heights (in inches).

We'll define the variable X to be for the student's height and the variable Y for the corresponding mother's height.

X = Student	Y = Mother
65	65
63	60
63	60
68	68
66	66
65	64
61	63
69	70

1) Make a scatterplot of these scores on the axes below



Evaluate the following expressions:

2) $\sum X =$

$$65 + 63 + 63 + 68 + 66 + 65 + 61 + 69 = 520$$

3) $\bar{X} = \frac{\sum X}{n}$

$$\frac{520}{8} = 65$$

4) $\sum Y =$

$$65 + 60 + 60 + 68 + 66 + 64 + 63 + 70 = 516$$

5) $\bar{Y} = \frac{\sum Y}{n}$

$$\frac{516}{8} = 64.5$$

6) $\sum XY =$

$$4225 + 3780 + 3780 + 4624 + 4356 + 4160 + 3843 + 4830 = 33598$$

7) $\sum X^2 =$

$$4225 + 3969 + 3969 + 4624 + 4356 + 4225 + 3721 + 4761 = 33850$$

8) $\sum Y^2 =$

$$4225 + 3600 + 3600 + 4624 + 4356 + 4096 + 3969 + 4900 = 33370$$

9) $(\sum X)^2 =$

$$520^2 = 270400$$

10) $(\sum Y)^2 =$

$$516^2 = 266256$$

11) SS_X

$$SS_X = \sum X^2 - \frac{(\sum X)^2}{n} =$$
$$33850 - \frac{270400}{8} = 50$$

12) SS_Y

$$SS_Y = \sum Y^2 - \frac{(\sum Y)^2}{n} =$$
$$33370 - \frac{266256}{8} = 88$$

13) The standard deviation of x:

$$s_X = \sqrt{\frac{SS_X}{n}} =$$
$$\sqrt{\frac{50}{8}} = 2.5$$

14) The standard deviation of y:

$$s_Y = \sqrt{\frac{SS_Y}{n}} =$$
$$\sqrt{\frac{88}{8}} = 3.32$$

15) The correlation:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{SS_X SS_Y}} =$$
$$r = \frac{33598 - \frac{(520)(516)}{8}}{\sqrt{(50)(88)}} = 0.87$$

16) The slope of the regression line:

$$m = r\left(\frac{s_Y}{s_X}\right)$$

$$b = 0.87\left(\frac{3.32}{2.5}\right) = 1.16$$

17) Plot the point (\bar{X}, \bar{Y}) on the scatter plot.

18) Use the point-slope formula:

$$Y - \bar{Y} = m(X - \bar{X})$$

to find the equation of the line and write in the format of the slope-intercept:

$$Y = mX + b$$

$$(Y-64.5) = (1.16)(X-65)$$

$$Y = (1.16)(X-65) + 64.5$$

$$Y = (1.16)X - (1.16)(65) + 64.5$$

$$Y = (1.16)X - 75.4 + 64.5$$

$$Y = (1.16)X + -10.9$$

19) Draw the regression line on the scatter plot. Hint: You know it goes through (\bar{X}, \bar{Y}) , so pick another point on the line, plot it, and use a ruler to draw the line.

20) Use this regression line to predict the height of the mother of a student who is 68.7 inches tall (round to 1/10th of an inch).

$$Y = mX+b = (1.16)X + -10.9 = (1.16)(68.7) + -10.9 = 68.8 \text{ inches.}$$

21) Now use R to generate a scatterplot and correlation for the data we've been working with. You'll need to find the Female students in this class that chose Yellow for their favorite color and find their heights and their mother's heights (in inches).

Here's a hint of how to find those heights and put them into variables to be plotted:

```
# load in the survey data
survey <- read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# find the heights of the Female students that prefer Yellow:
height.student <- survey$height[survey$gender == "Female" & survey$color == "Yellow"]
height.mother <- survey$mheight[survey$gender == "Female" & survey$color == "Yellow"]
```

More hints can be found in the R files associated with the tutorial on correlation:

ScatterplotsAndCorrelations.R

Answer:

```
# Homework 2 problem 21
#
# Load in the data
survey <- read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# Get the heights for the students and their mothers for
# Students that chose "Yellow" as their favorite color:
height.student <- survey$height[survey$gender == "Female" & survey$color == "Yellow"]
height.mother <- survey$mheight[survey$gender == "Female" & survey$color == "Yellow"]
# Scatterplot:
plot(height.student,height.mother,
      xlab = "Student's Height",
      ylab = "Mother's Height",
      pch = 19,
      col = "blue",
      asp = 1,
      cex = 2)
# Correlation:
cor(height.student,height.mother,use = "complete.obs")
[1] 0.8743829
```