Di Zhen (1717719)

dizhen

2020/2/29

Question 1

```
Y1 <- c(8.04, 6.95, 7.58, 8.81, 8.33, 9.96, 7.24, 4.26, 10.84, 4.82, 5.68)
X1 <- c(10, 8, 13, 9, 11, 14, 6, 4, 12, 7, 5)
Y2 <- c(9.14, 8.14, 8.74, 8.77, 9.26, 8.10, 6.13, 3.10, 9.13, 7.26, 4.74)
X2 <- c(10, 8, 13, 9, 11, 14, 6, 4, 12, 7, 5)
Y3 <- c(7.46, 6.77, 12.74, 7.11, 7.81, 8.84, 6.08, 5.39, 8.15, 6.42, 5.73)
X3 <- c(10, 8, 13, 9, 11, 14, 6, 4, 12, 7, 5)
Y4 <- c(6.58, 5.76, 7.71, 8.84, 8.47, 7.04, 5.25, 12.50, 5.56, 7.91, 6.89)
X4 <- c(8, 8, 8, 8, 8, 8, 8, 19, 8, 8, 8)
```

1.1

```
cor(Y1,X1)
```

```
## [1] 0.8164205
```

```
cor(Y2,X2)
```

```
## [1] 0.8162365
```

```
cor(Y3,X3)
```

```
## [1] 0.8162867
```

```
cor(Y4,X4)
```

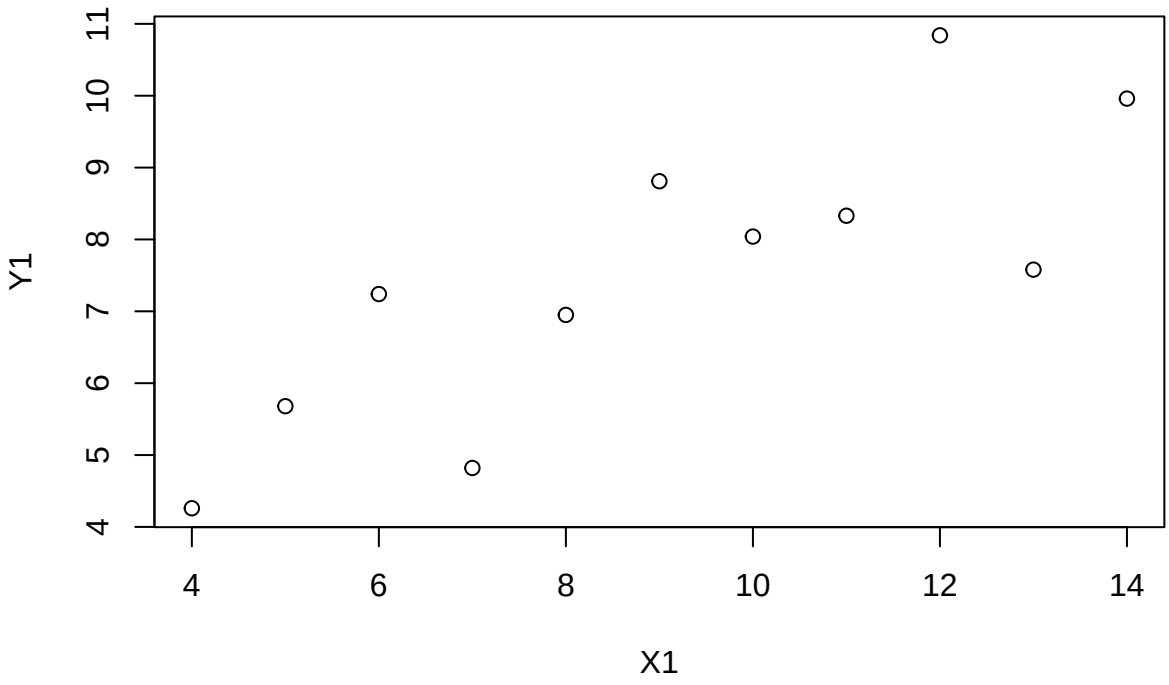
```
## [1] 0.8165214
```

The correlation coefficient of X1 and Y1 is around 0.816. The correlation coefficient of X2 and Y2 is around 0.816. The correlation coefficient of X3 and Y3 is around 0.816. The correlation coefficient of X4 and Y4 is around 0.817.

1.2

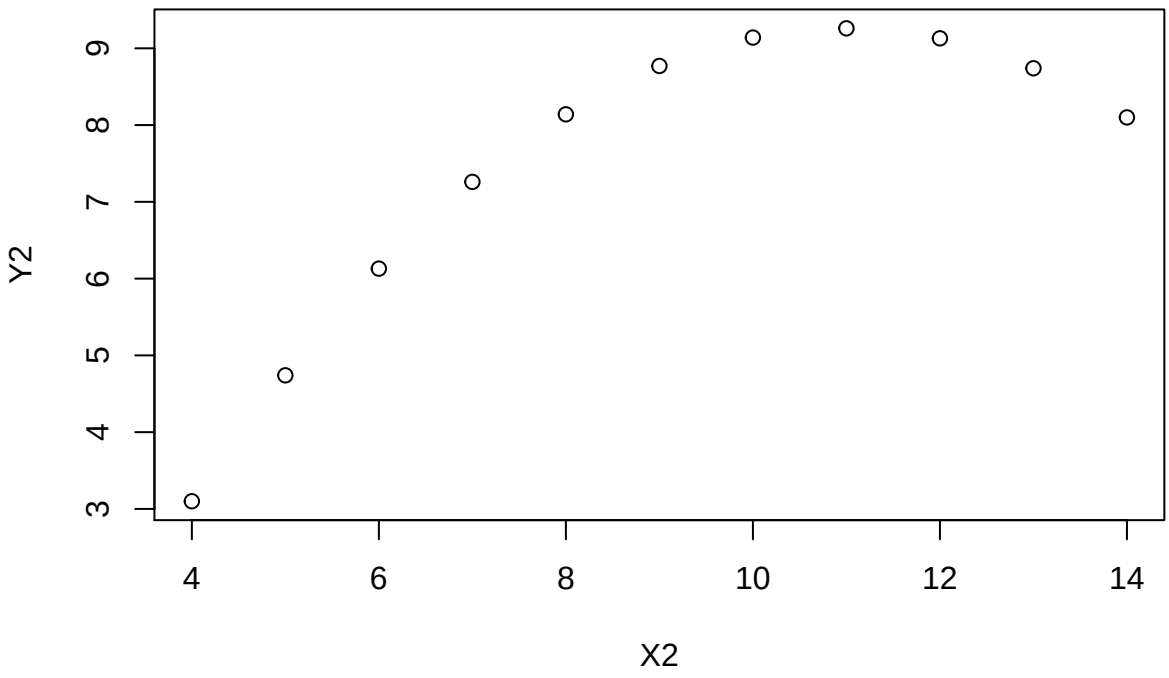
```
plot(X1, Y1,
     main = "Figure 1. X1 and Y1",
     ylab = "Y1",
     xlab = "X1",
     cex.main = 0.9)
```

Figure 1. X1 and Y1



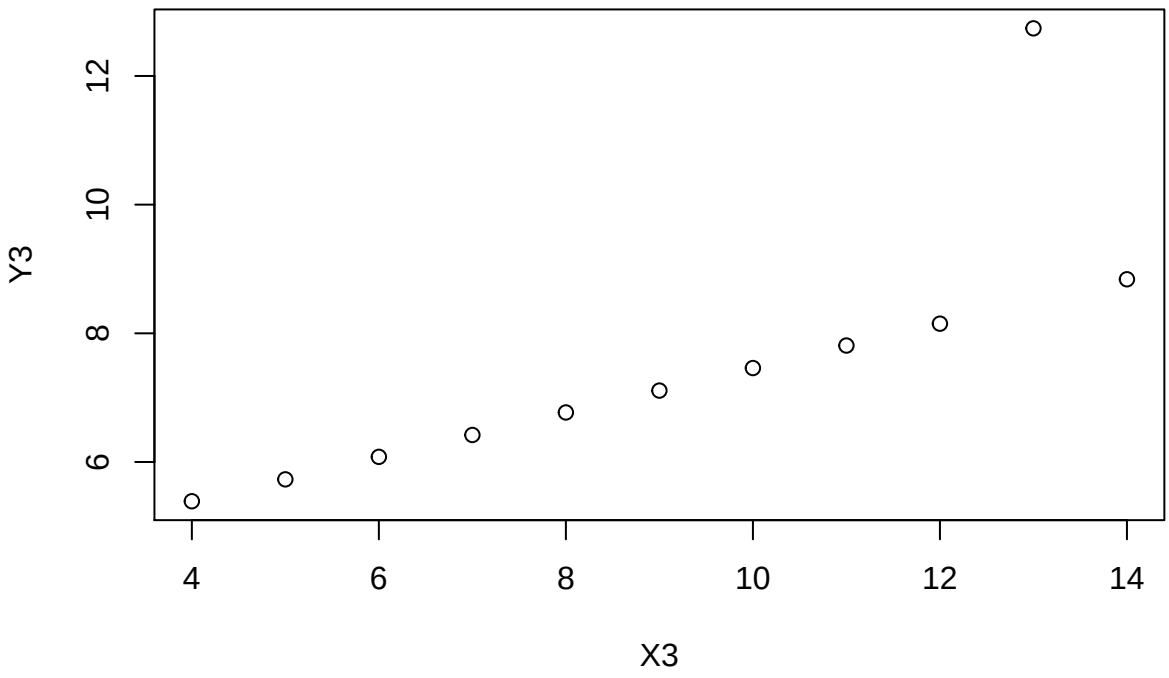

```
plot(X2, Y2,
     main = "Figure 2. X2 and Y2",
     ylab = "Y2",
     xlab = "X2",
     cex.main = 0.9)
```

Figure 2. X2 and Y2




```
plot(X3, Y3,
     main = "Figure 3. X3 and Y3",
     ylab = "Y3",
     xlab = "X3",
     cex.main = 0.9)
```

Figure 3. X3 and Y3




```
plot(X4, Y4,
     main = "Figure 4. X4 and Y4",
     ylab = "Y4",
     xlab = "X4",
     cex.main = 0.9)
```

Figure 4. X4 and Y4

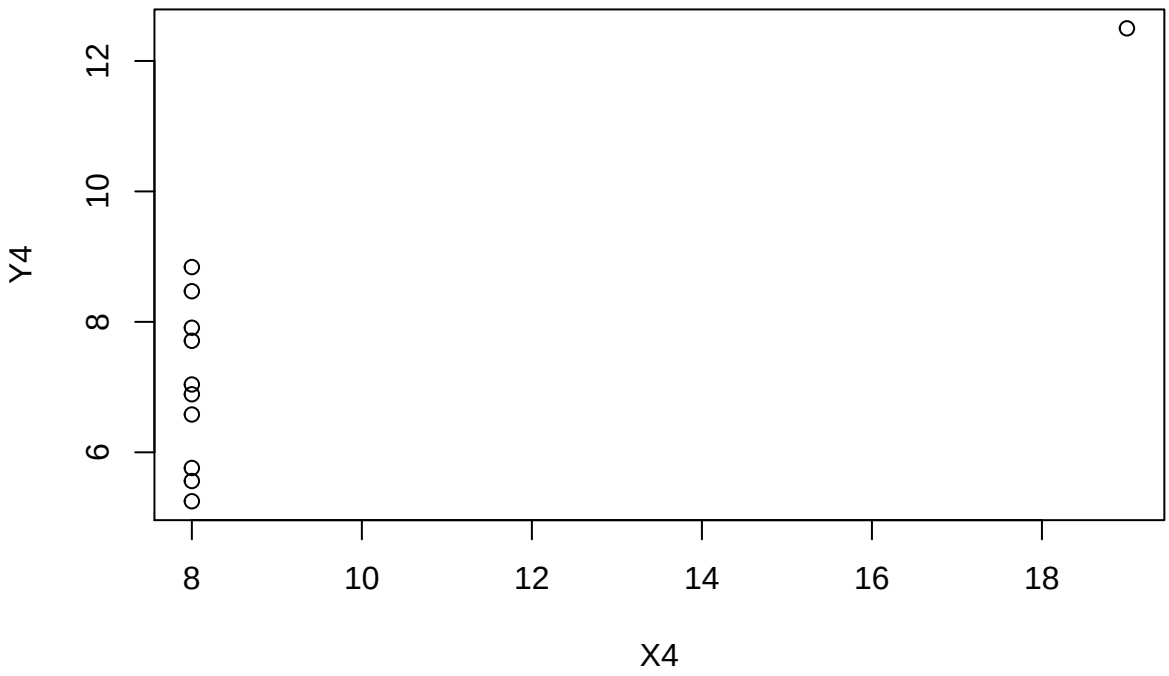


Figure 1-4 are the answer to the question 1.2.

1.3

The ρ of four pairs are similar(~0.816) meaning strong positive relationship. The scatter plots show that relationships between four pairs are different. $\rho = 0.816$ reflects true linear relation between first pair. There is a non-linear relationship in other plots but we cannot notice it from ρ. In third plot, the outlier decreases the ρ which is misleading. In fourth plot, there seems no relationship while an outlier increases the ρ which is misleading. So we should examine the plot of data before exploring the relationship. And the possibility of a non-linear relationship is not discounted when calculating ρ.

Question 2

```
H <- c(186, 180, 160, 186, 163, 172, 192, 170, 174, 191, 182, 178 ,181, 168 ,162 ,188, 168, 183, 188, 1
W <- c(175, 168, 154, 166, 162, 152, 179, 163, 172, 170, 170, 147, 165, 162, 154, 166, 167, 174, 173, 1
```

2.1

```
plot(H, W,
     main = "Figure 5. Heights of husbands and wives ",
     ylab = "Heights of husbands",
```

```
     xlab = "Heights of wives",
     cex.main = 0.9)
```

Figure 5. Heights of husbands and wives

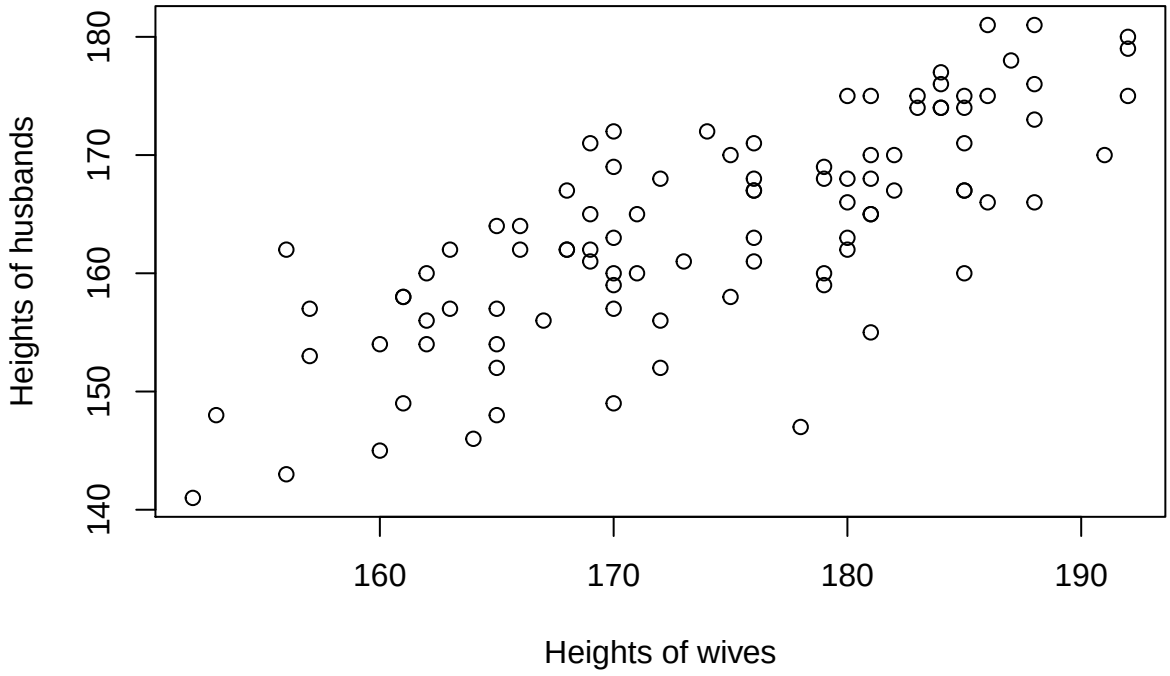


Figure 5 is the answer to the question 2.1.

2.2

```
cov(H,W)
```

```
## [1] 69.41294
```

The covariance of the height (cm) of husbands and wives is around 69.413

2.3

```
H.inch <- H/2.54
W.inch <- W/2.54
cov(H.inch, W.inch)
```

```
## [1] 10.75903
```

The covariance of the height(inch) of husbands and wives is around 10.759

2.4

When the unit of measurement changed from centimeter to inch, the covariance changed from 69.413 to 10.759. The positive value means that there is a positive relationship between height of husbands and wives. It does not tell how much about the strength of relationship because it is affected by unites of measurement.

2.5

```
cor(H, W)
```

```
## [1] 0.7633864
```

The correlation coefficient of the height (cm) of husbands and wives is around 0.763

2.6

```
cor(H.inch, W.inch)
```

```
## [1] 0.7633864
```

The correlation coefficient of the height (inch) of husbands and wives is around 0.763

2.7

The correlation coefficient of the height of husbands and wives is 0.763 which means positive relationship between them. The ρ did not change although the unit of measurement changed from centimeter to inch. That is because the ρ is covariance between the standardised height of husbands and standardised height of wives.

2.8

```
H.NEW <- W-5
```

The variable H.NEW is produced.

2.9

```
plot(H.NEW, W,
     main = "Figure 6. New heights of husbands and heights of wives ",
     ylab = "New heights of husbands",
     xlab = "Heights of wives",
     cex.main = 0.9)
```

Figure 6. New heights of husbands and heights of wives

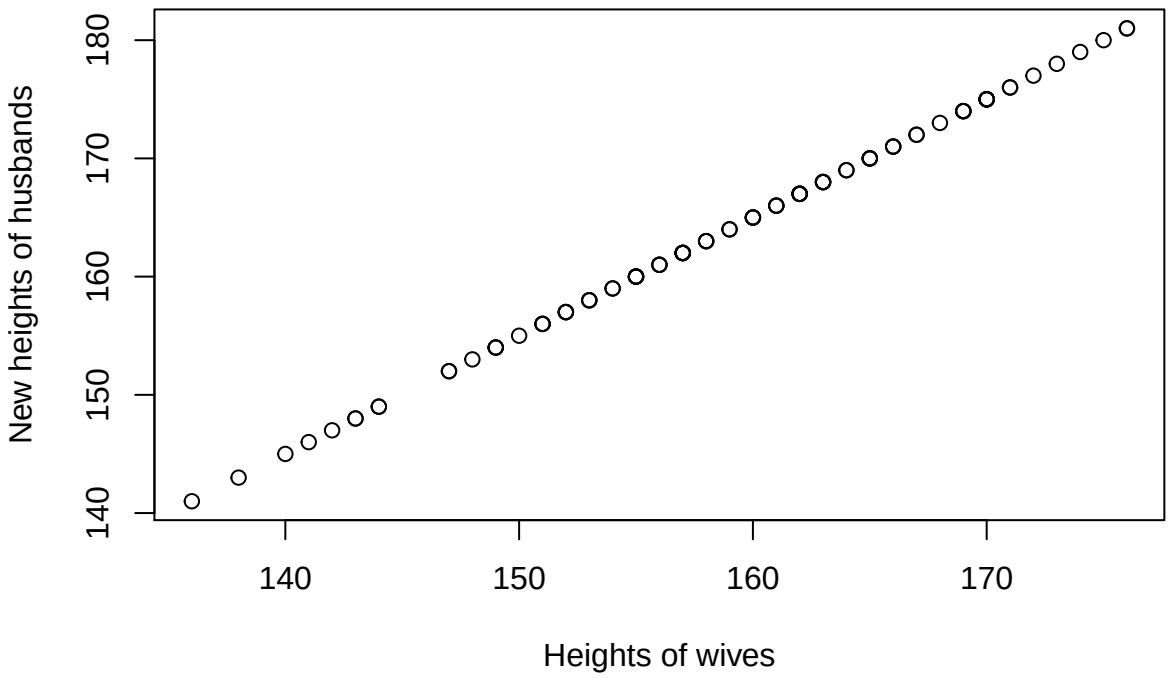


Figure 6 is the answer to the question 2.9

2.10

From the scatter plot, I estimate the correlation coefficient between W and H.NEW is 1.