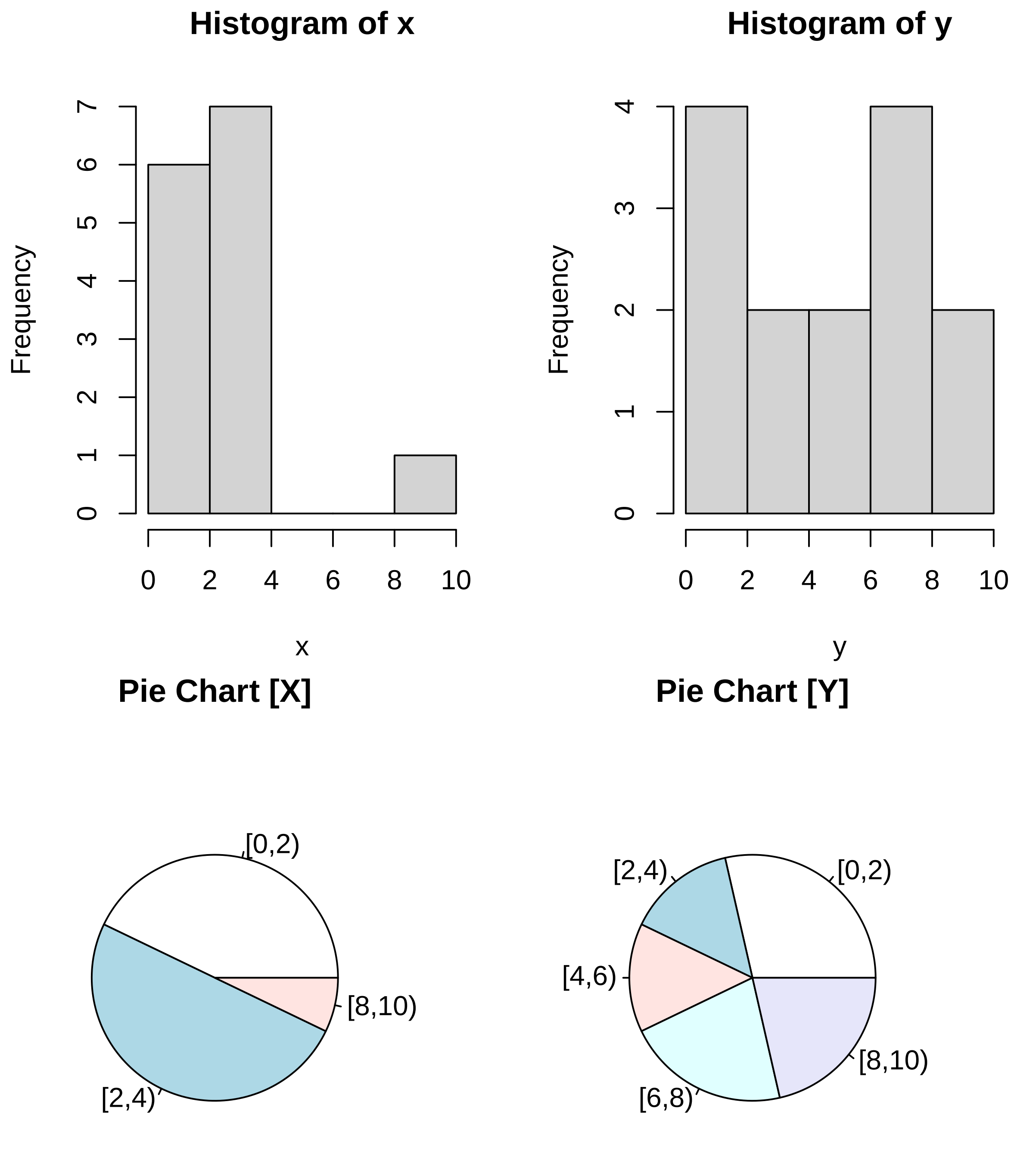
Homework 01

Ming Lin

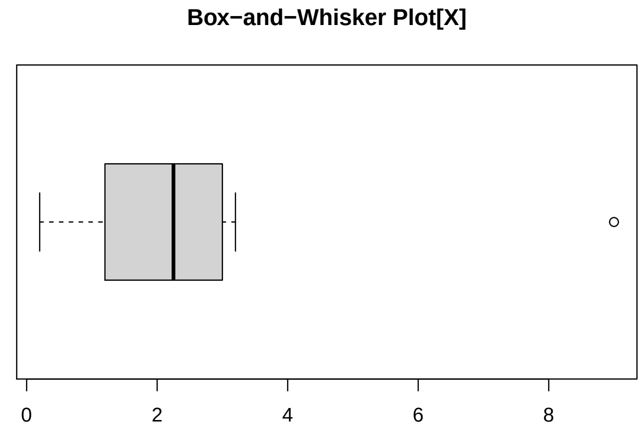
I pledge my honor that I have abided by the Stevens Honor System

Problem 1(i)



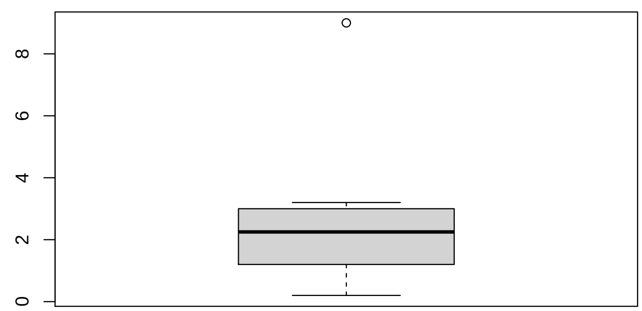
It seems that for the values within x pie chart ~50% of its values are greater than 2 and less than 4. Whereas y pie chart all values are distributed evenly.

Problem 1(ii)

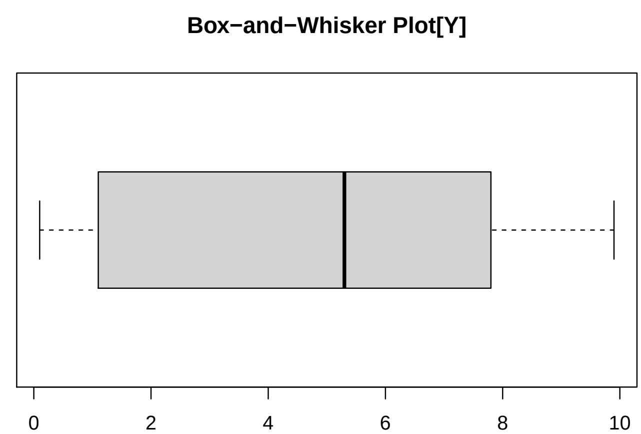



```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.200   1.275   2.250   2.407   2.900   9.000
```

```
## Warning in if (plot) {: the condition has length > 1 and only the first element
## will be used
```

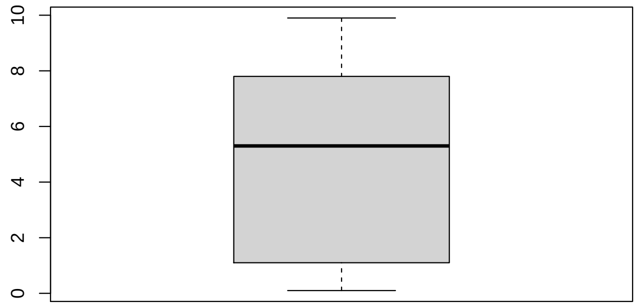


Variance[X]:4.5684066 Outlier[X]:9



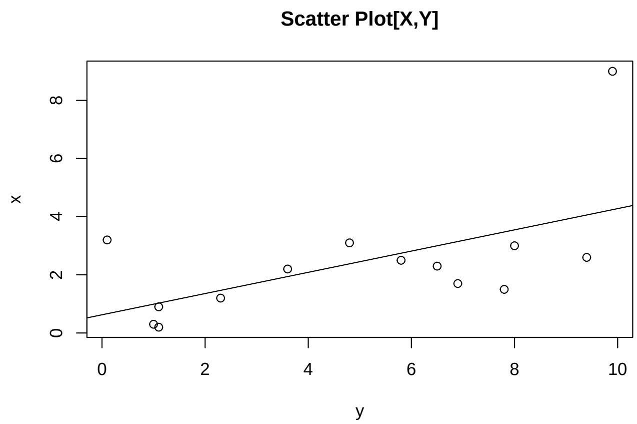

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.100   1.400   5.300   4.879   7.575   9.900
```

```
## Warning in if (plot) {: the condition has length > 1 and only the first element
## will be used
```



Variance[Y]: 11.247967 Outlier[Y]: 0

Problem 1(iii)

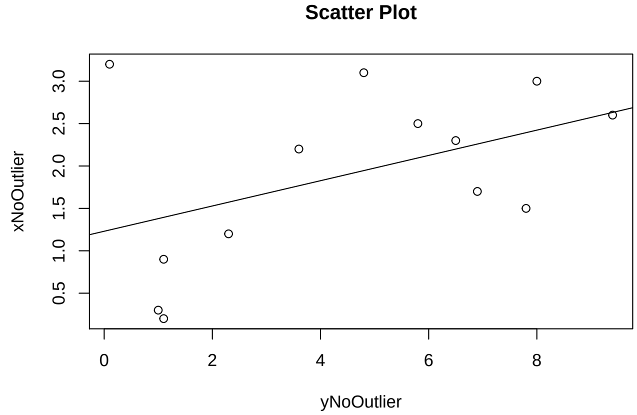


Correlation Coefficient:

```
## [1] 0.5730545
```

X and Y have a modern linear association because of their positive correlation coefficient

Problem 1(iv)

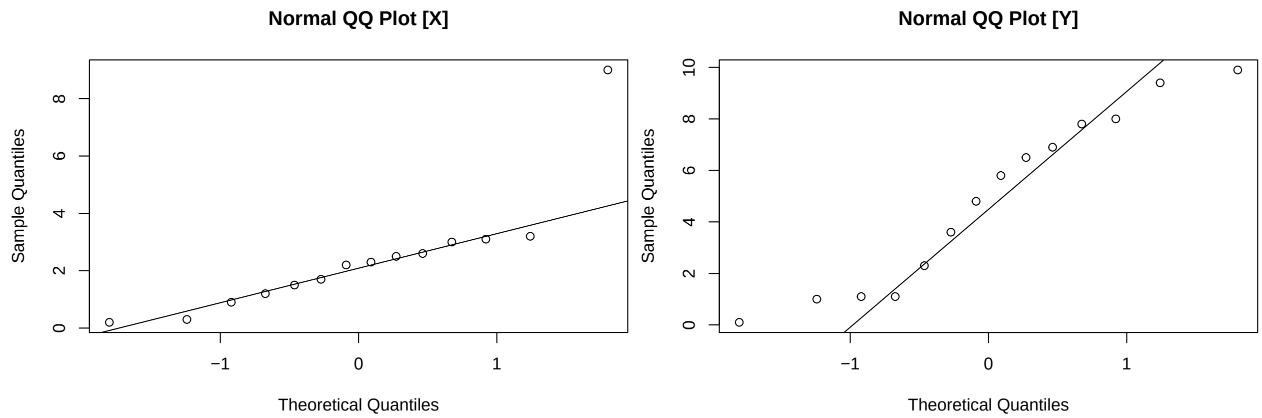


Removed (9.0, 9.8) and changed the correlaton coefficient to 0.4586256.

Problem 1(v)

The main difference observed between the numerical results in part 3 and 4 is that part 3 had a positive relation, while part 4 had a weaker relation. This change is a result of removing the outlier in part 4.

Problem 1(vi)



After observing the both QQ plots, it seems that x is more likely to be of normal distribution.

Problem 2

We can then write 9 as

$$\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(\bar{x} - a)^2$$

Equation 11 is clearly minimized when $a = \bar{x}$. Now consider part b of theorem 1. Expanc second expression in part b and simplify

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \bar{x})^2 &= \sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x}^2 \\ &= \sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ &= \sum_{i=1}^{n} x_i^2 - n\bar{x}^2\end{aligned}$$

Using the proof above this proves the first part of problem 2 and the second part becuase 1/n is a constant.