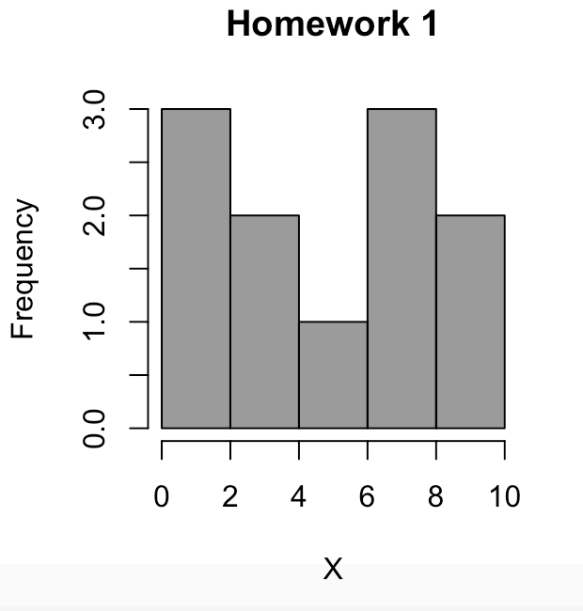
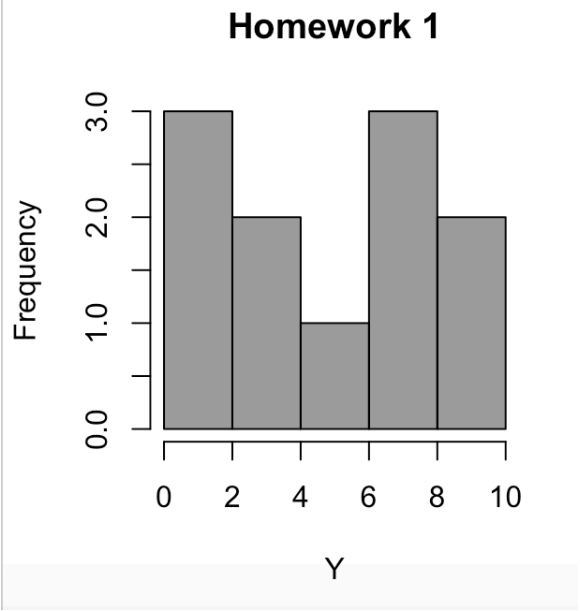
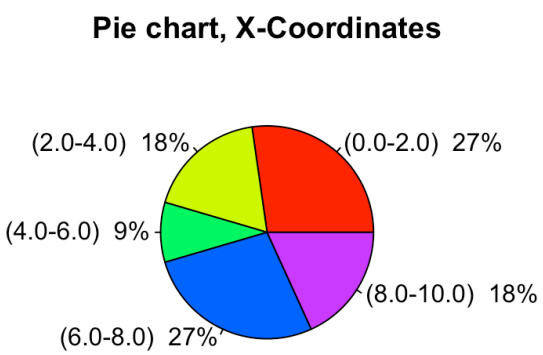
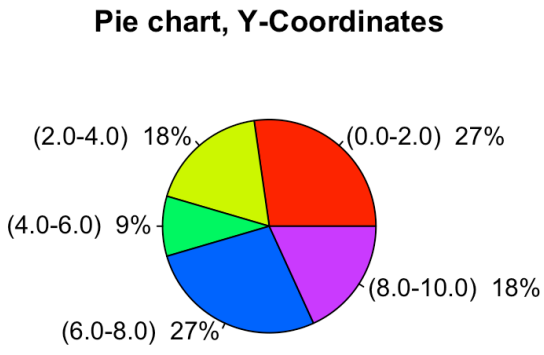
Homework 1
Kaitlynn Prescott
I pledge my honor that I have abided by the Stevens Honor System.

1. Plot histogram, pie chart, and describe the distribution of x and y.



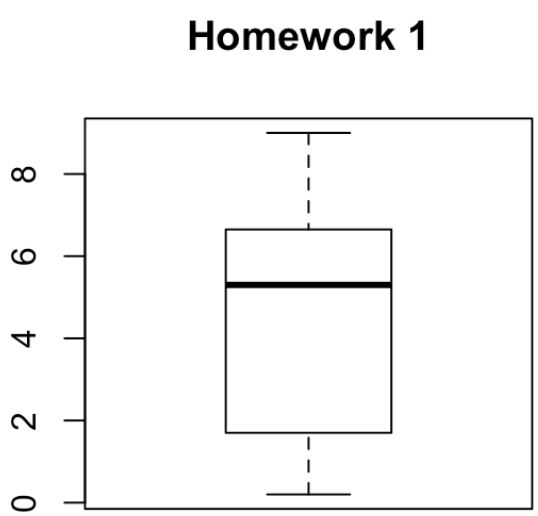



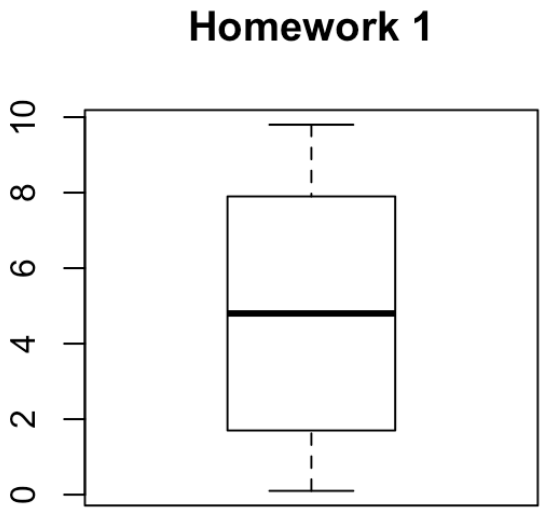




Based on the information, there is an equal amount of x coordinates and y coordinates in the 0-2 range and the 6-8 range. These are also where the bulk of the x coordinates lie. The range with the fewest x and y coordinates is the 4-6 range.

2. For x and y, build the box-plot, compute their five-number summaries and variances. Are there any outliers of x and y respectively?




X

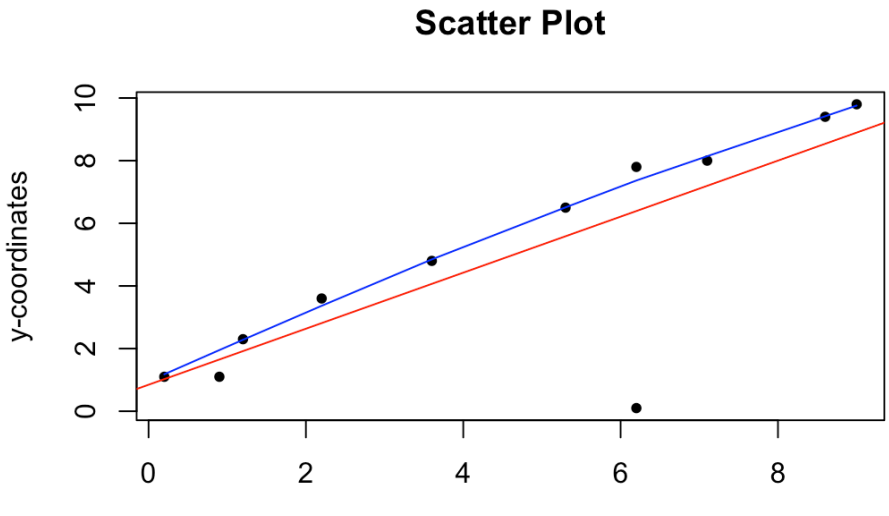



Y

Five Number Summaries:

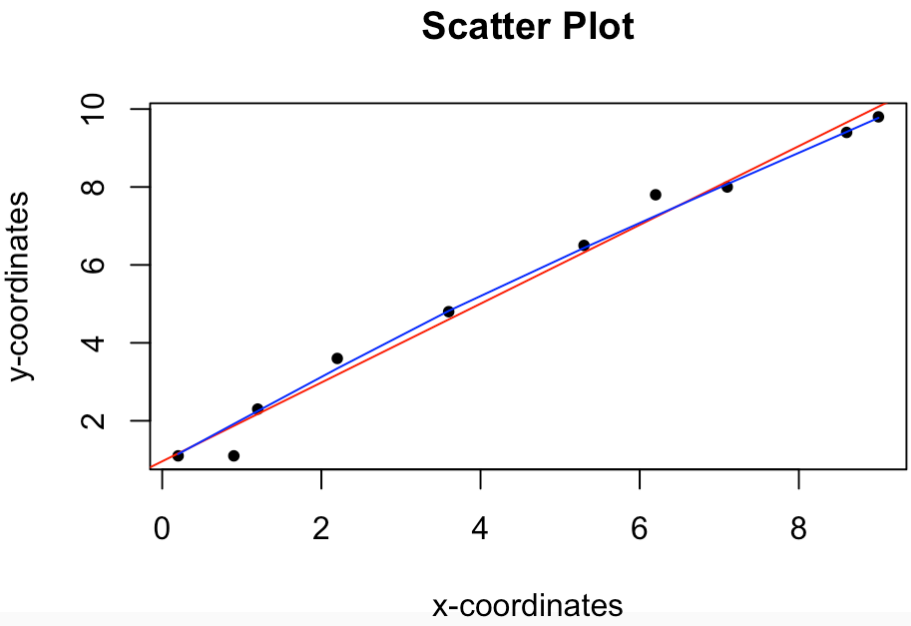
| | | |
|---|---|---|
| X: | Minimum: 0.2 | Y: Minimum: 0.1 |
| | Q1: 1.2 | Q1: 1.1 |
| | Q2: 5.3 | Q2: 4.8 |
| | Q3: 7.1 | Q3: 8.0 |
| | Maximum: 9.0 | Maximum: 9.8 |
| | Outliers: None | Outliers: None |
| | Variance: 9.86 | Variance: 12.52 |

3. Obtain the scatter plot of (x, y) and evaluate their correlation coefficient. Describe the linear association between x and y.




The correlation coefficient is 0.7942.
The scatter plot has a positive linear relationship, with one outlier, (6.2, 0.2).

4. Are there any outliers of (x,y)? If yes, remove them and compute the correlation coefficient again. What difference do you observed?




The correlation coefficient is now .972. This scatter plot has a much stronger linear relationship.