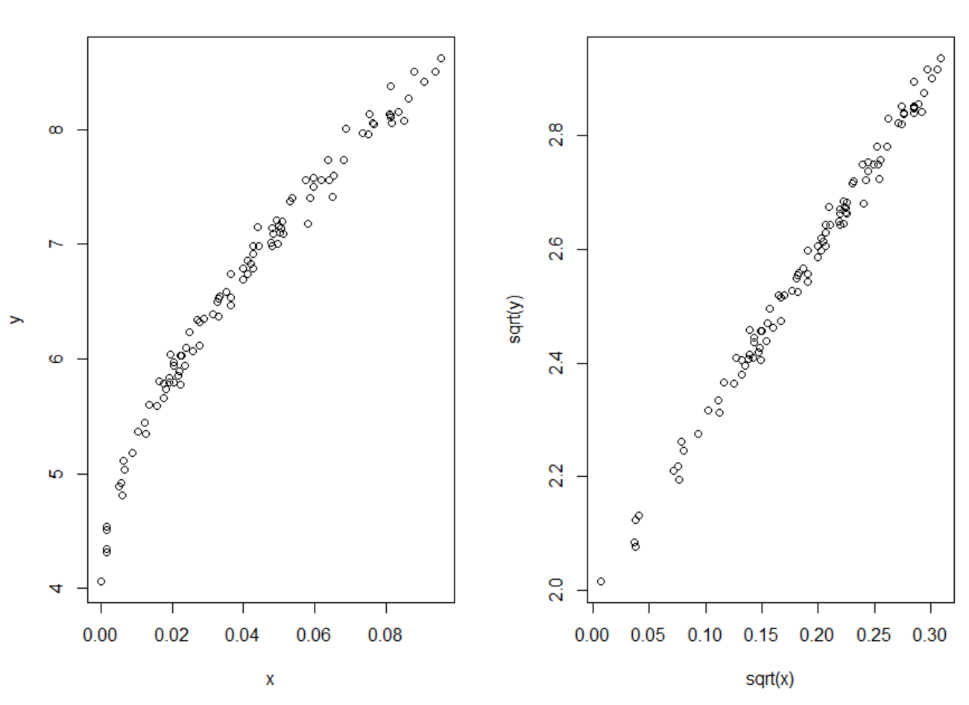
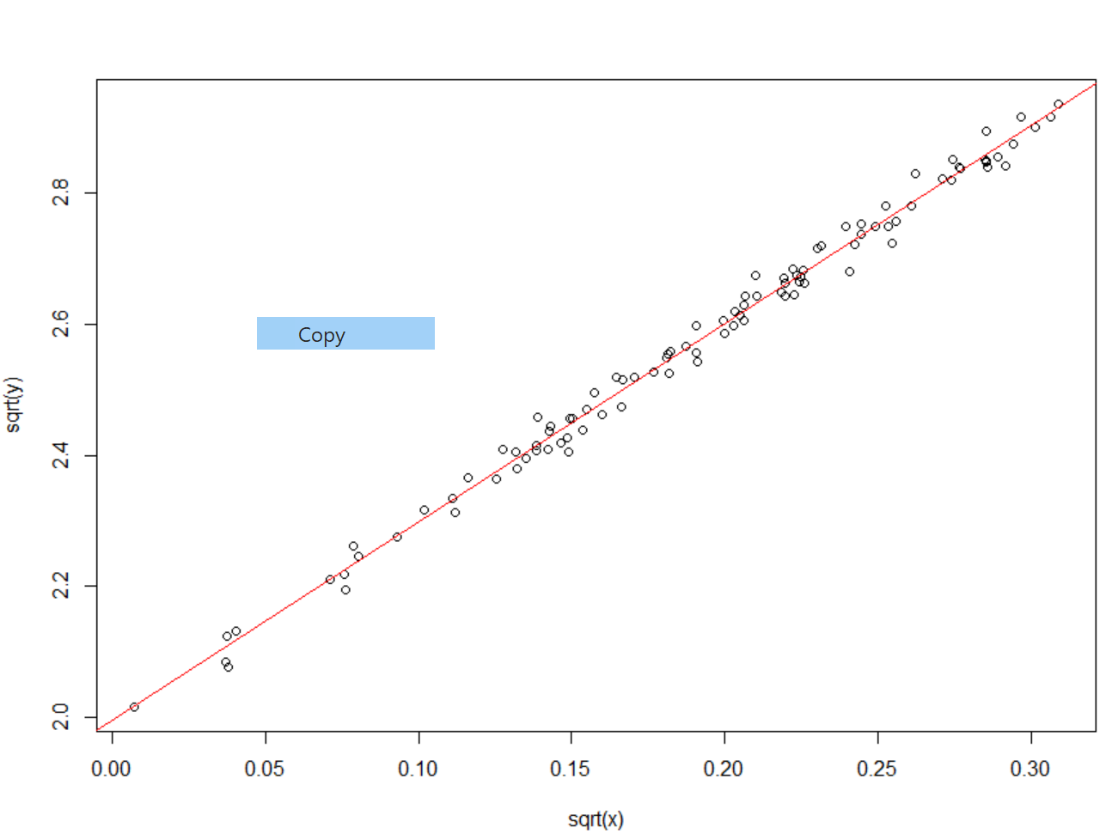
STAT 390 HW 5

**HW 14H**

hw-lect14-1

#a



#b

#c

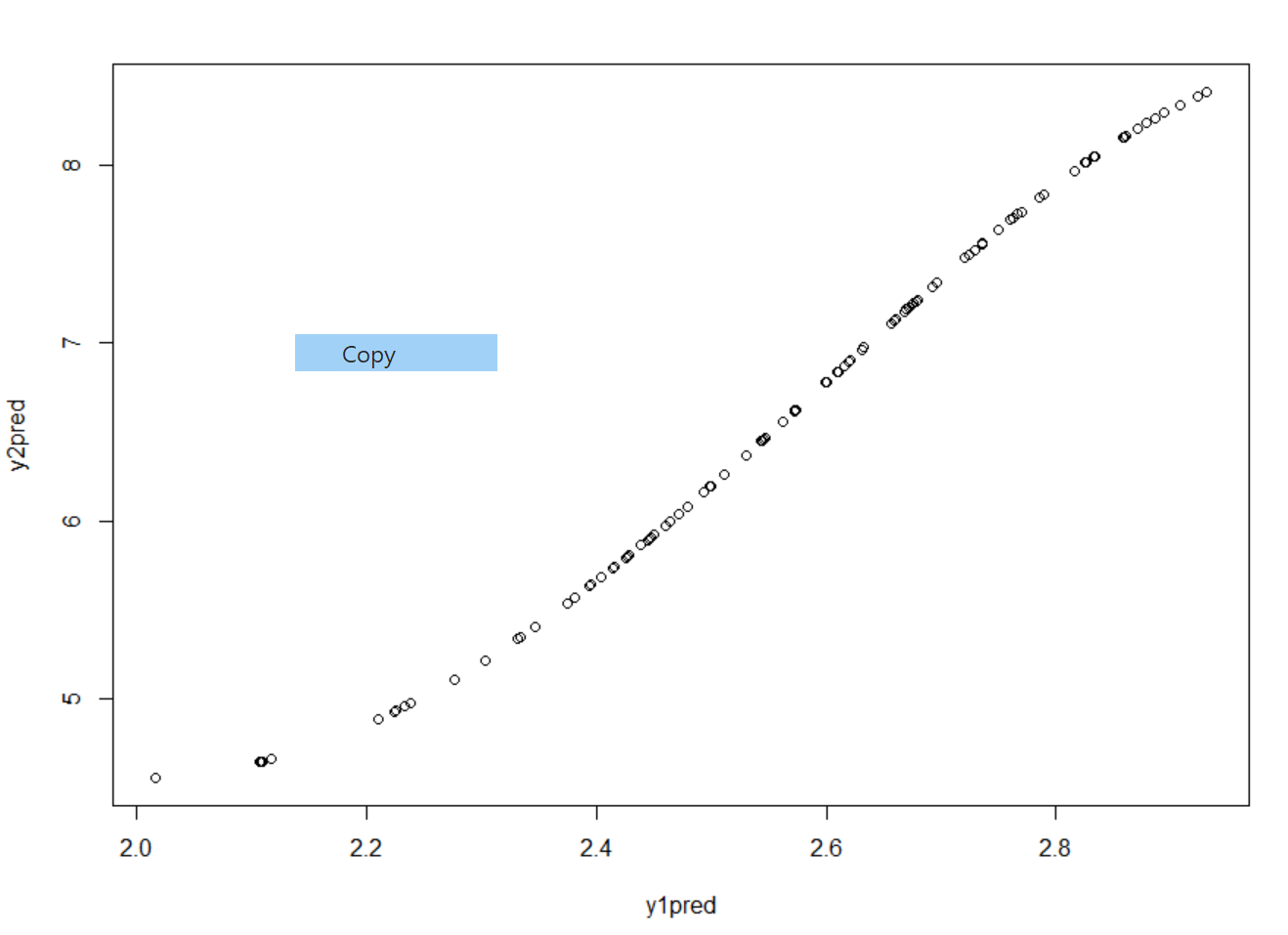
The percentage of the variability in the transformed y explained by the transformed x is 99.22%. The typical error in the prediction of the transformed y is 0.01911

#d

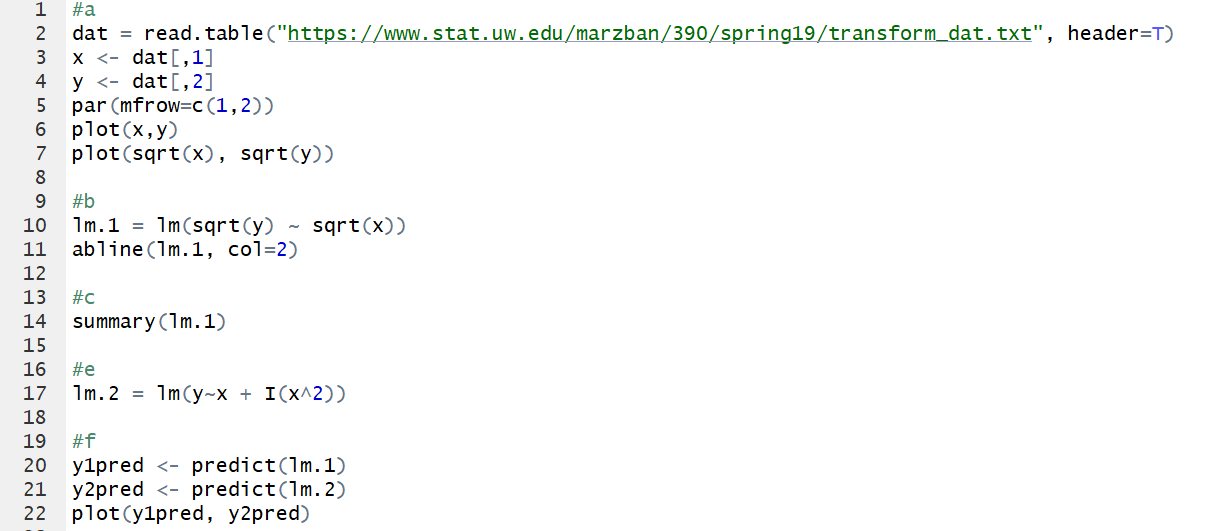
(On next paper page)

#e

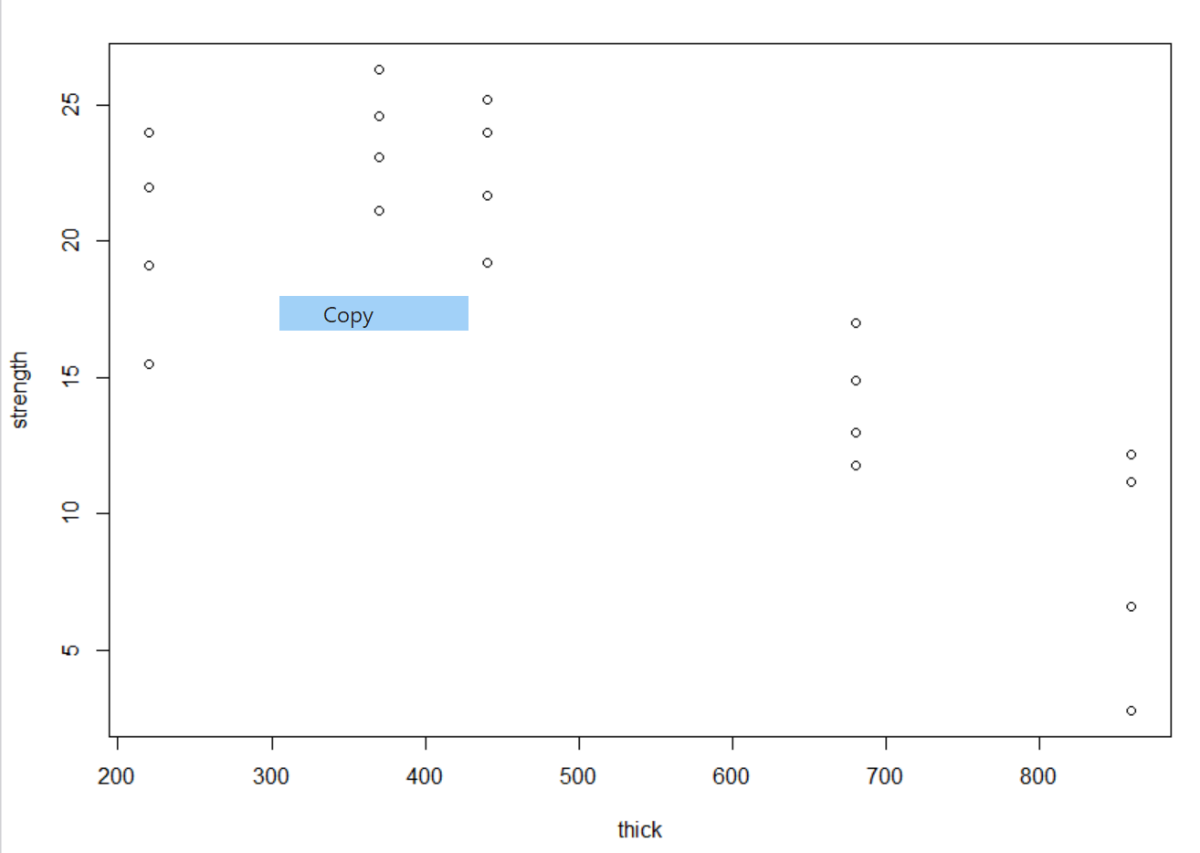
Look at code used below

#f

Code used for this section is below



3.33



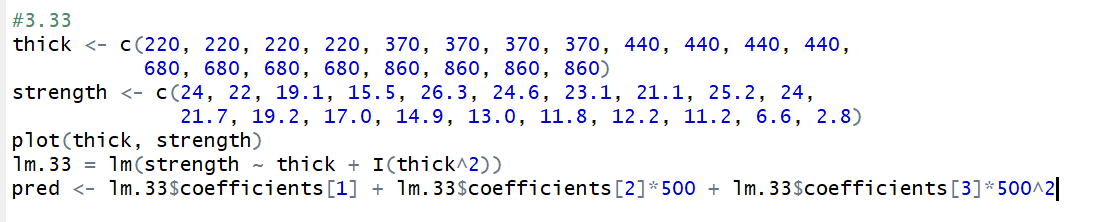
#a

This plot is not monotonic because y increases then decreases. As such, no transformation could create a linear relationship.

#b

The strength when the thickness is 500 is 21.2.

Code used can be seen below



**HW 15**

hw-lect15-1

#a

The coefficients are as follows: Intercept = -140.23, x1 = -16.48, x2 = 12.83, x3 = 0.0956, x4 = -0.2434, x5 = 0.499

#b

You cannot interpret the regression coefficients. Since x5 = x1\*x2, this variable depends on two other variables. As such, to interpret y’s relationship with say x1, x2 through x5 must be held constant. This is impossible because x5 = x1\*x2.

#c

The R^2 value is 0.75613. As such, 75.61% of the variability in y is explained by the model for regression that we have generated.

#d

The residual standard error is 7.023. This means that, on average, the observed values fall a distance of 7.023 from the regression line.

#f

(see code below)

#g

The R^2 value is 0.447034. This means that 44.7% of the variability in y is explained by the regression model using x1 and x2.

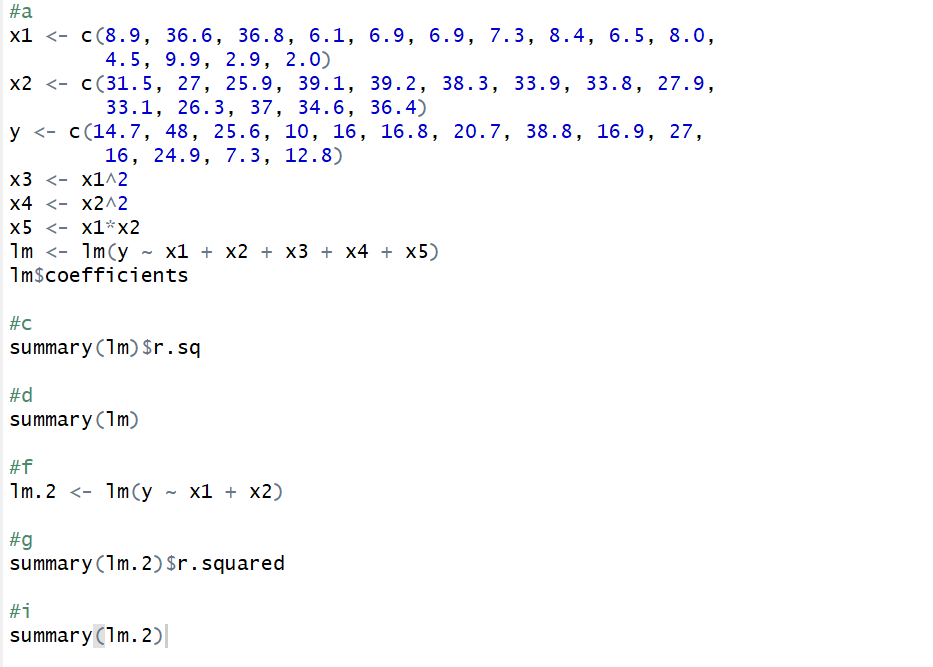
#h

Yes. Since the R^2 value was higher in our initial model we know that one of the higher order terms help construct a better regression model.

#i

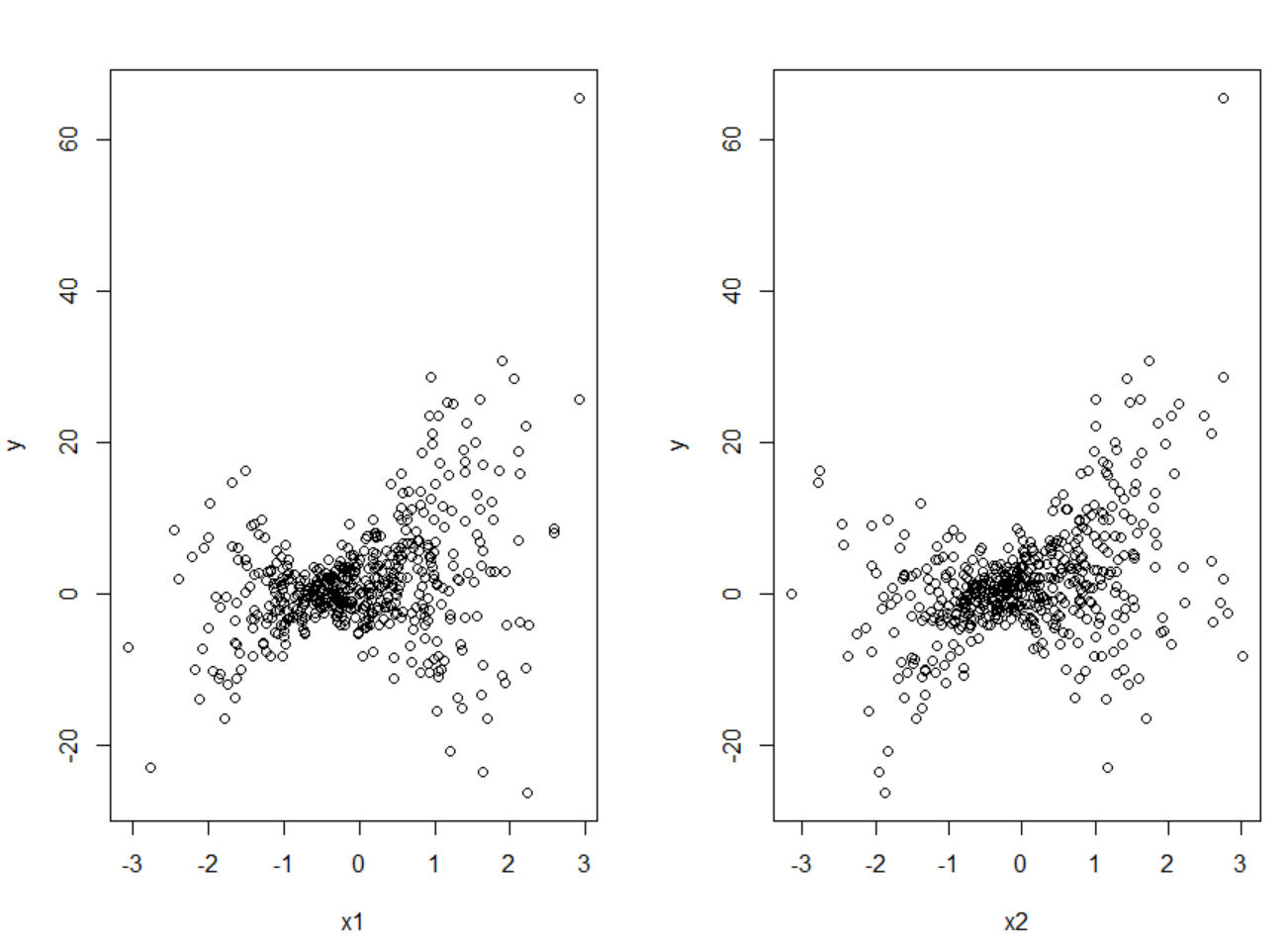
The residual standard error is lower in our initial model. This also suggests that the initial model provided a more accurate representation of the data.

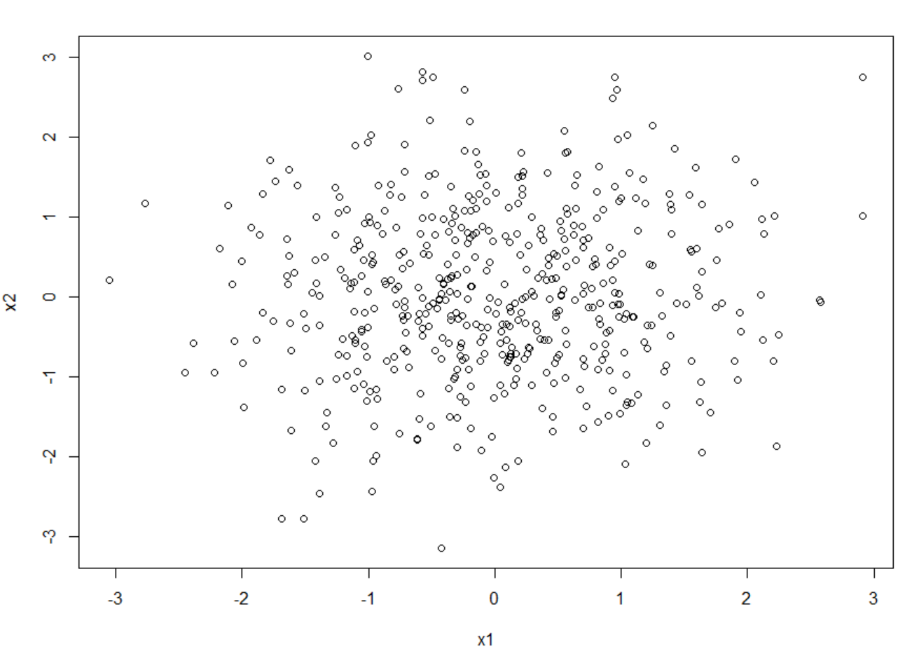
Code used can be seen below



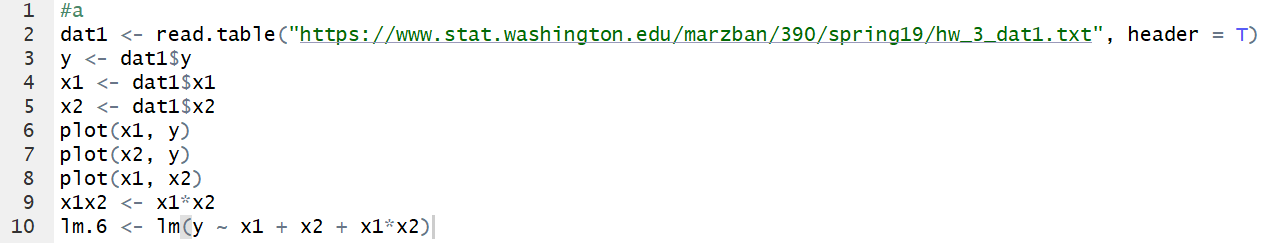
hw-lect15-2

#a

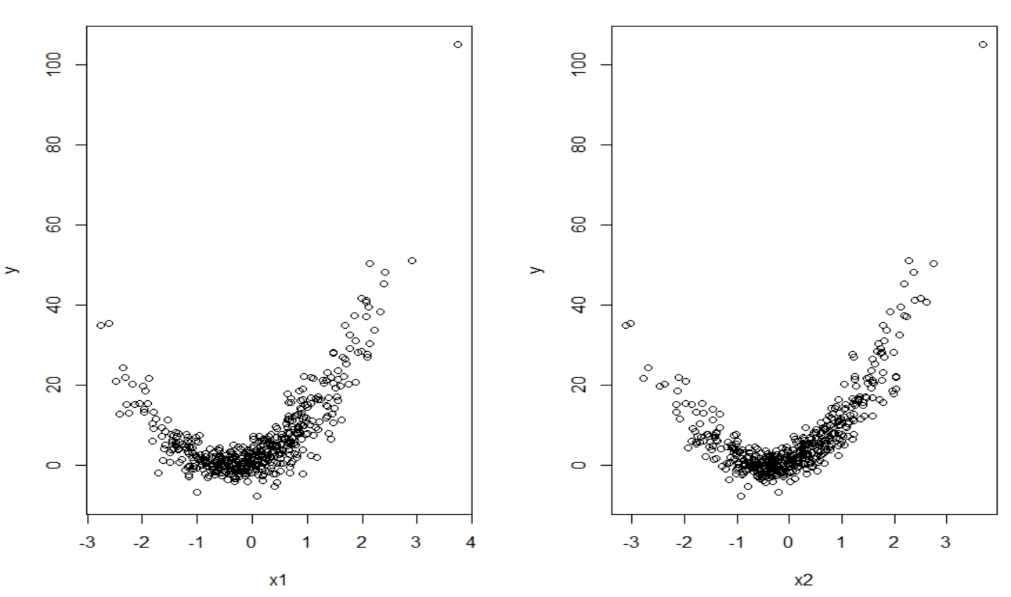


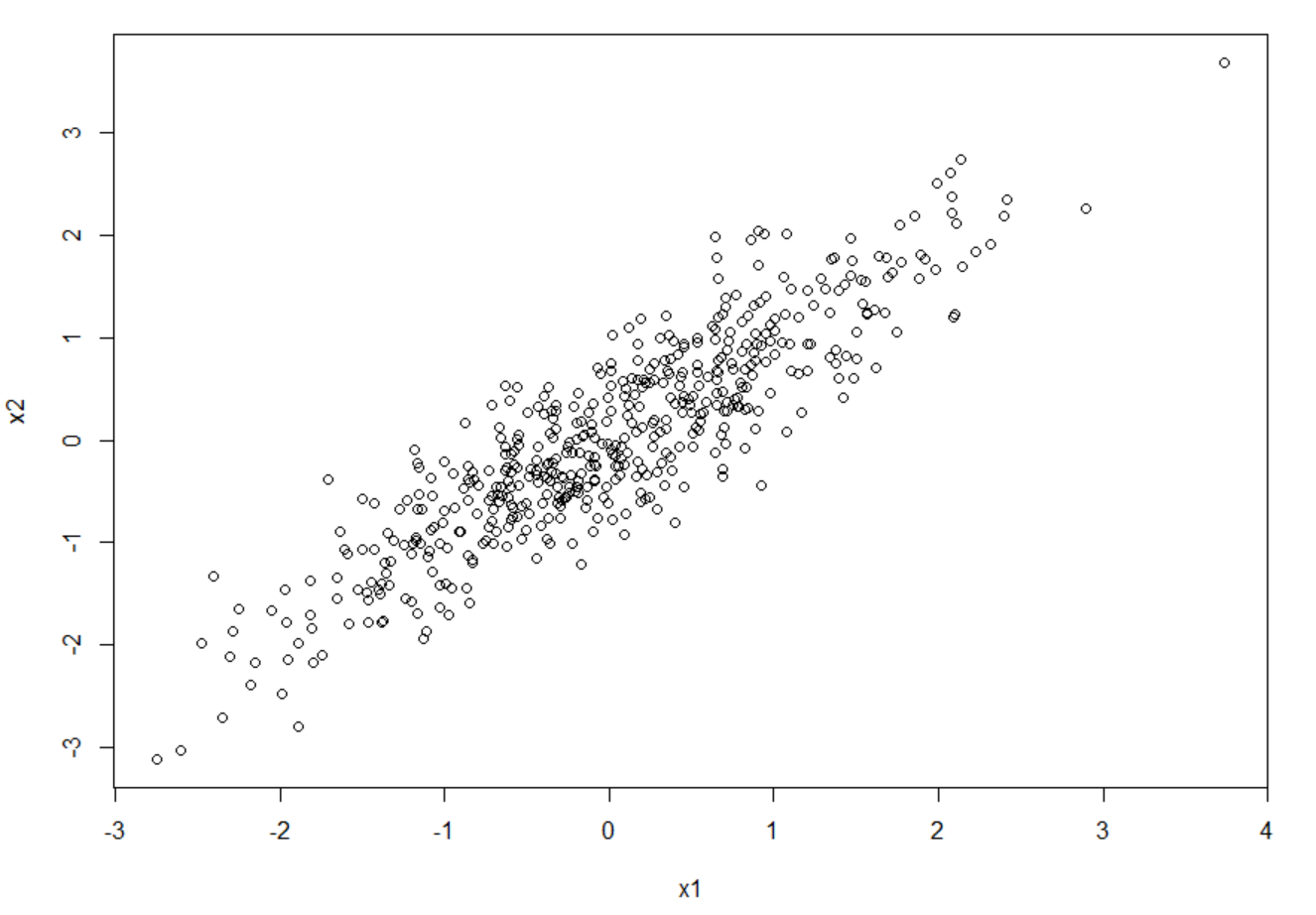


From the above graphs there does not seem to be any signs of nonlinearity. As such, we can use both x1 and x2 in the model. Additionally, there seems to be a saddle in the first two graphs, indicating the need for an interaction term. There does not seem to be much collinearity because there is no linear relationship between x1 and x2. This information gives us the following model: y = a +B1x1 + B2x2 + B2x1x2. This model gives us an R^2 value of 0.9356 and a residual error of 2.017.

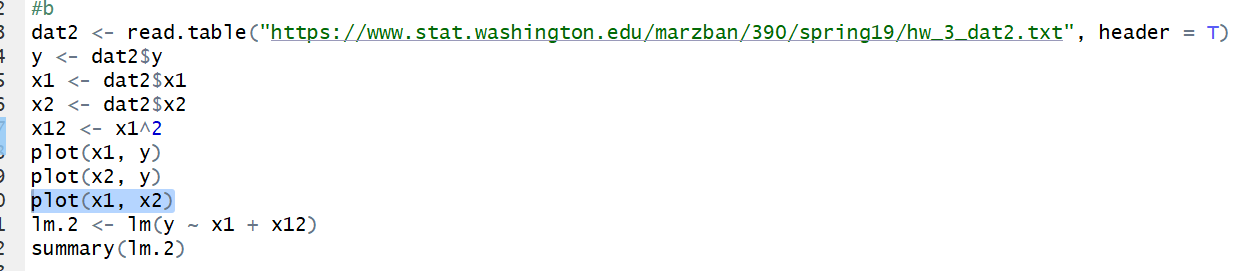


#b





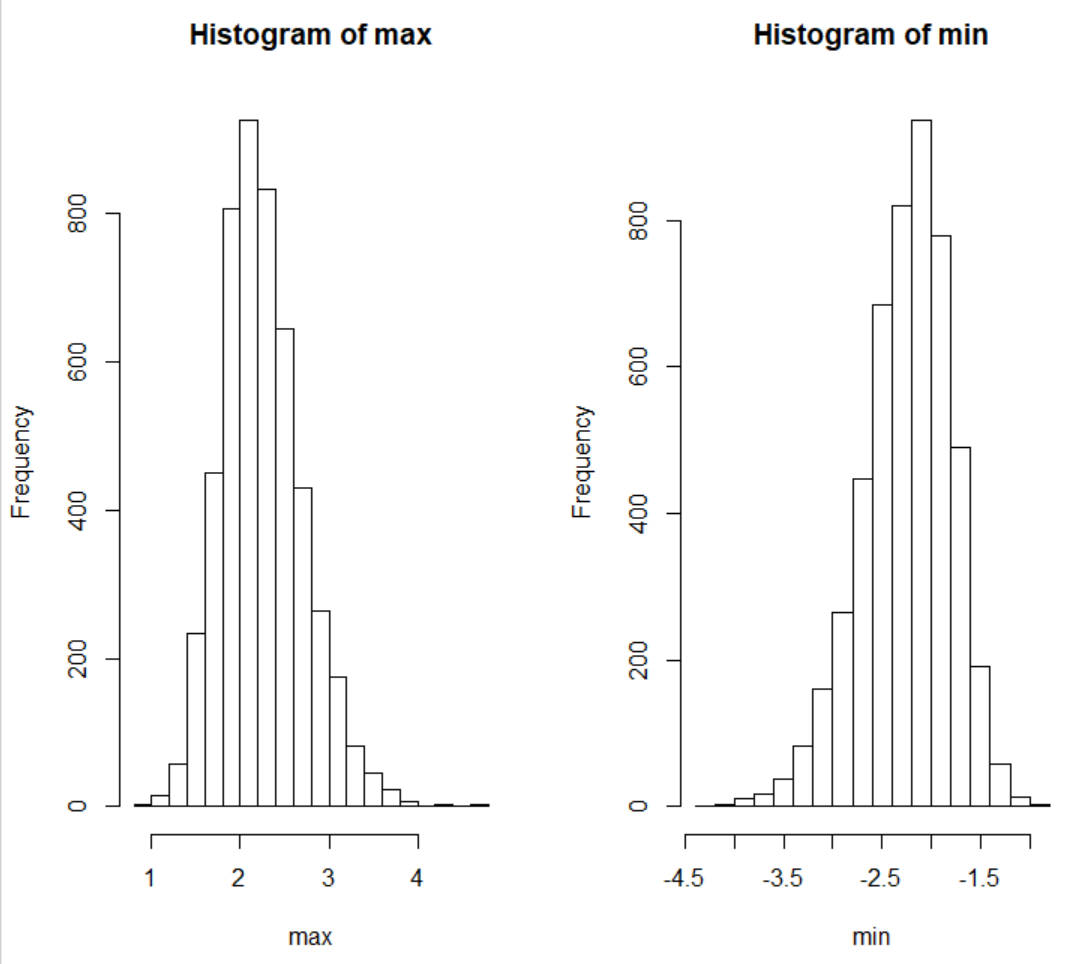
In this part there were no signs of a saddle, accordingly there was no interaction term included. Additionally, there seemed to be a linear relationship between x1 and x2, meaning that there was collinearity. As such, one of these variables needed to be held constant, so x2 was dropped from the model. Since there’s no x2 term, nonlinearity is not relevant in this case. The first two graphs seem parabolic necessitating a x1^2 term. This led to the following model: y = a + B1x1 + B2x1^2. This gives an R^2 value of .8639 and a residual error of 3.839.

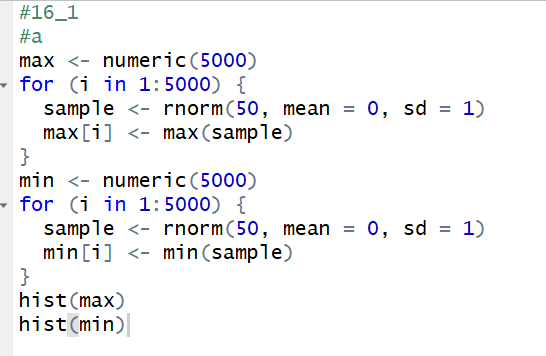


**HW 16**

hw-lect16-1

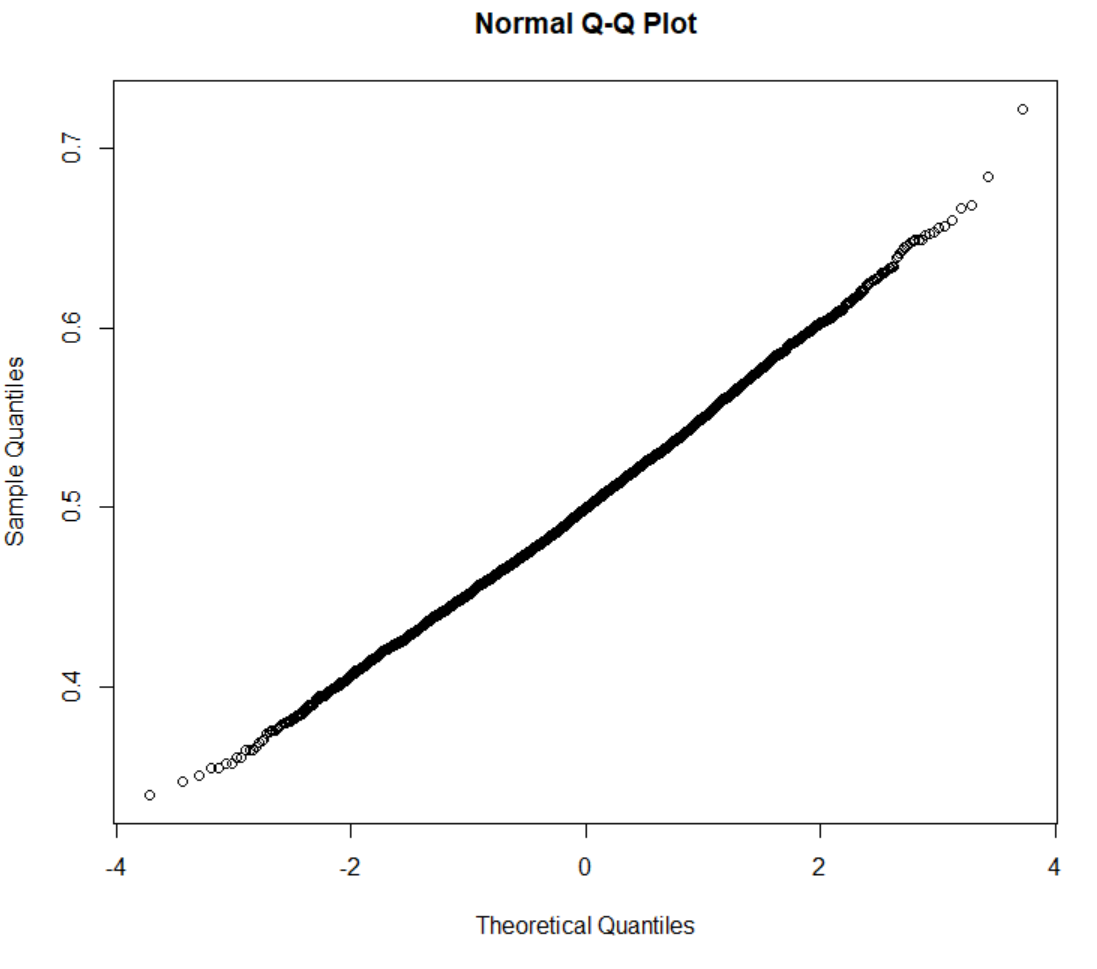
#a

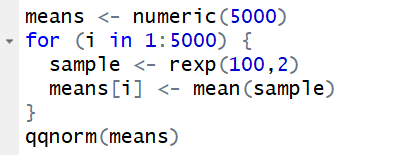




hw-lect16-2

#a





#b

Using the qq plot to approximate values, it looks like the mean is around 0.5 and that the standard deviation is around 0.02. We can justify these approximations with the following calculations: