Zi Xuan Li

Professor Vahdani

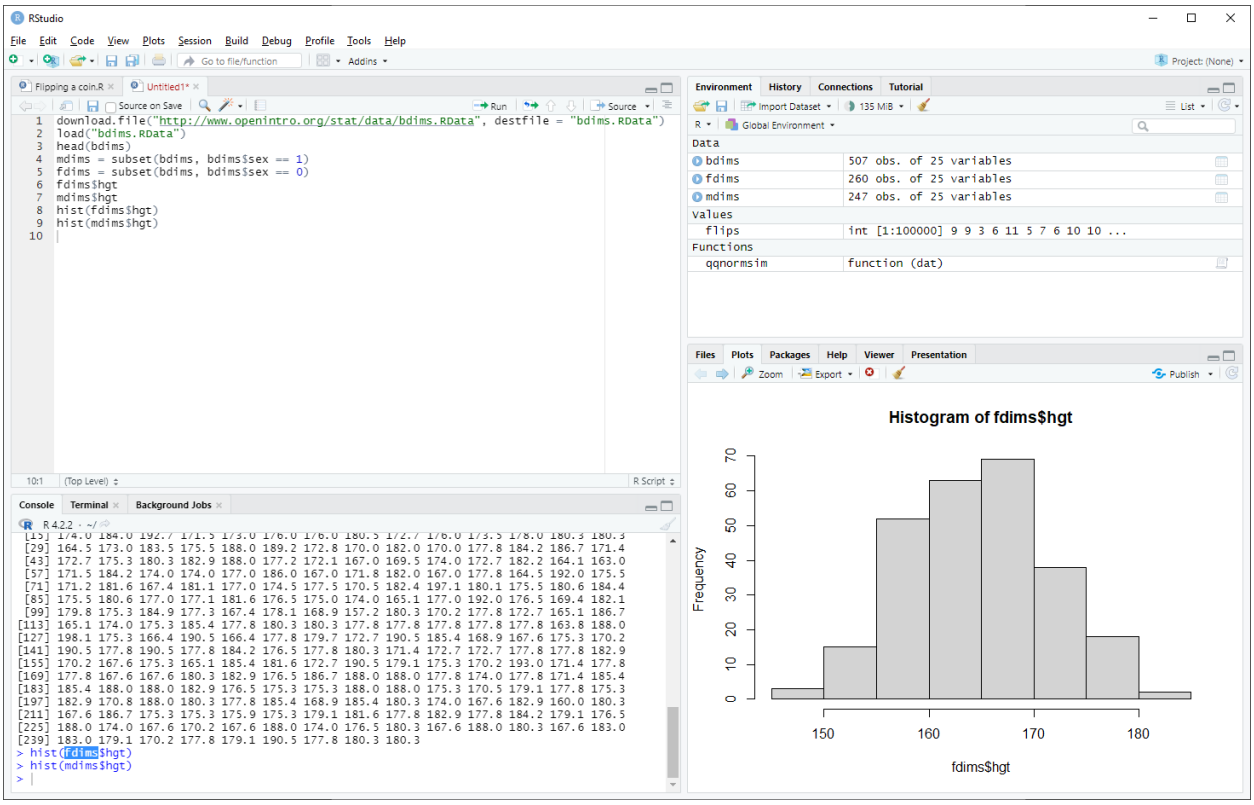
CSC 21700

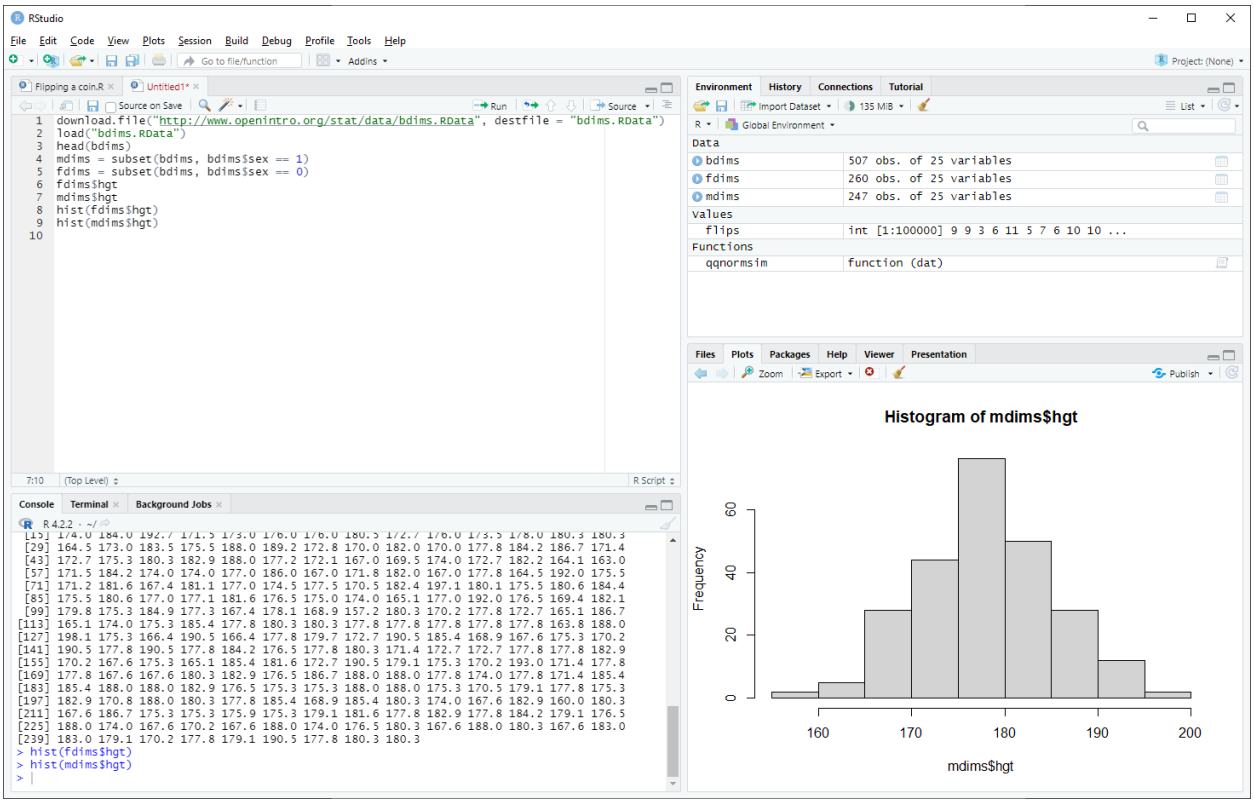
December 18th, 2022

R Assignment #2

Exercise 1:

Plot the histograms. How would you compare the various aspects of the two distributions?

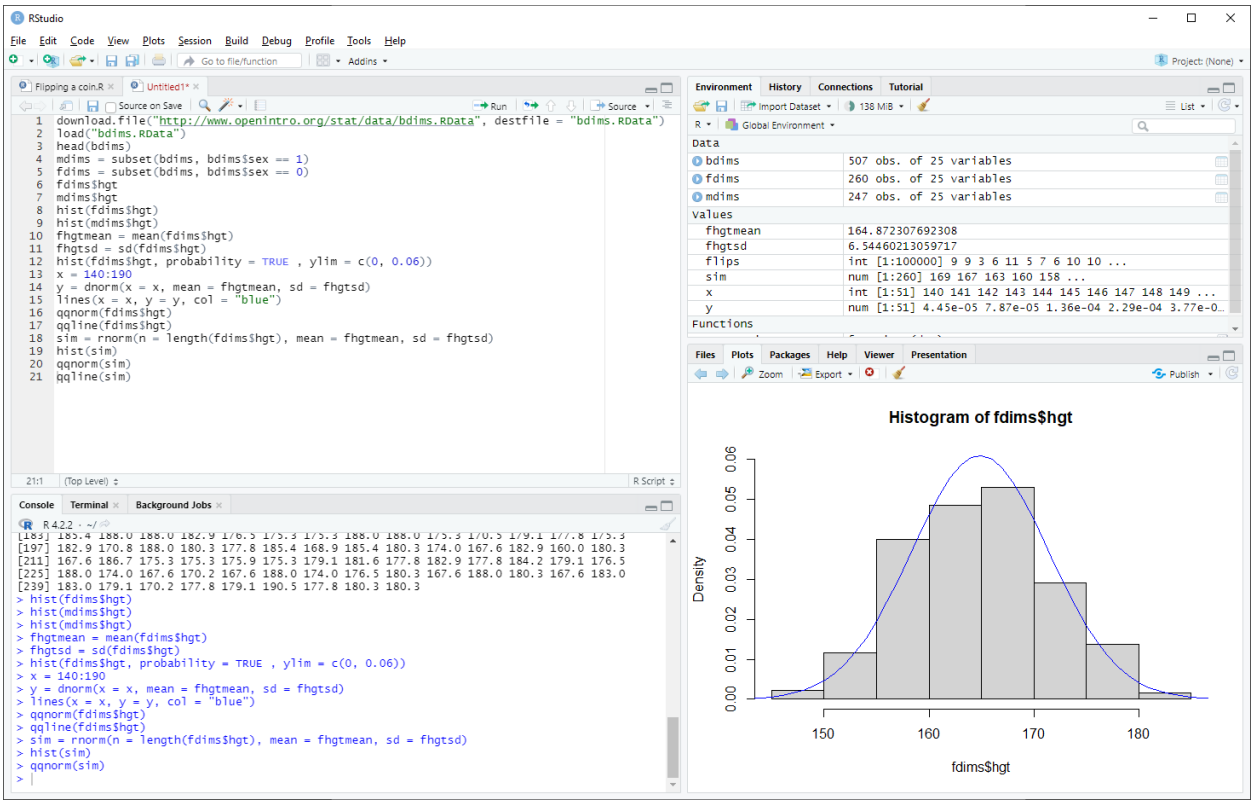




Comparing the histograms from men's heights to women's heights, I would say the histogram of men's heights seems more normal because it looks more like a bell-shaped curve than the histogram of the women's heights.

Exercise 2:

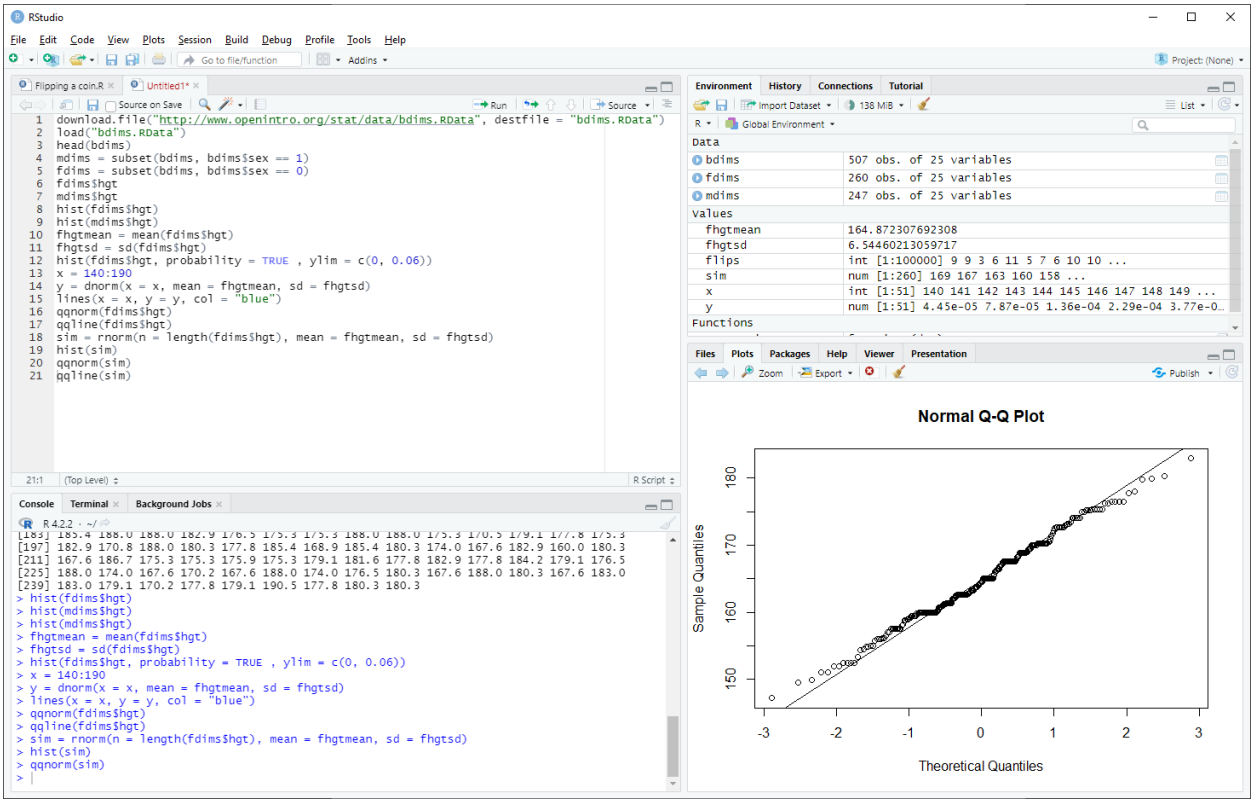
Based on this plot, does it appear that the data follow a nearly normal distribution?



Yes, with the line of the normal distribution marked on the histogram of women's heights the distribution does look nearly normal with the exception at 165 (in the center of the curve).

Exercise 3:

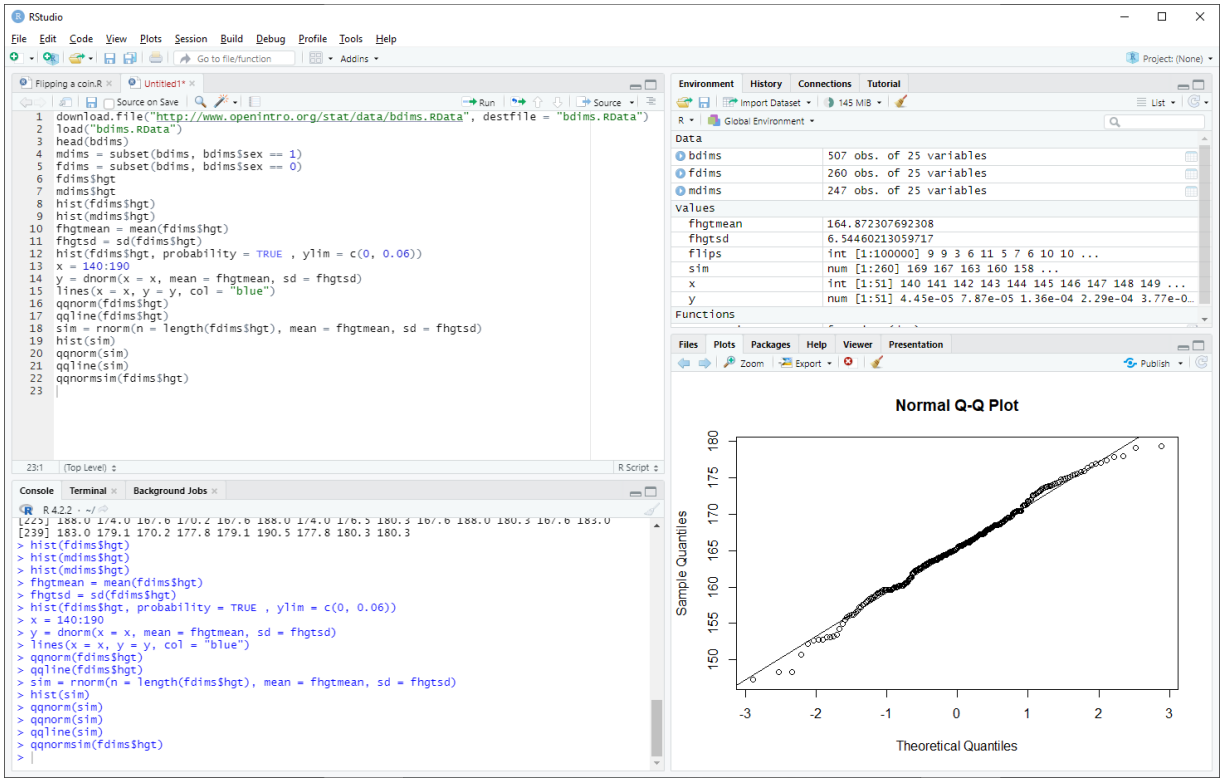
Make a normal probability plot of sim. Do all of the points fall on the line? How does this plot compare to the probability plot for the real data?

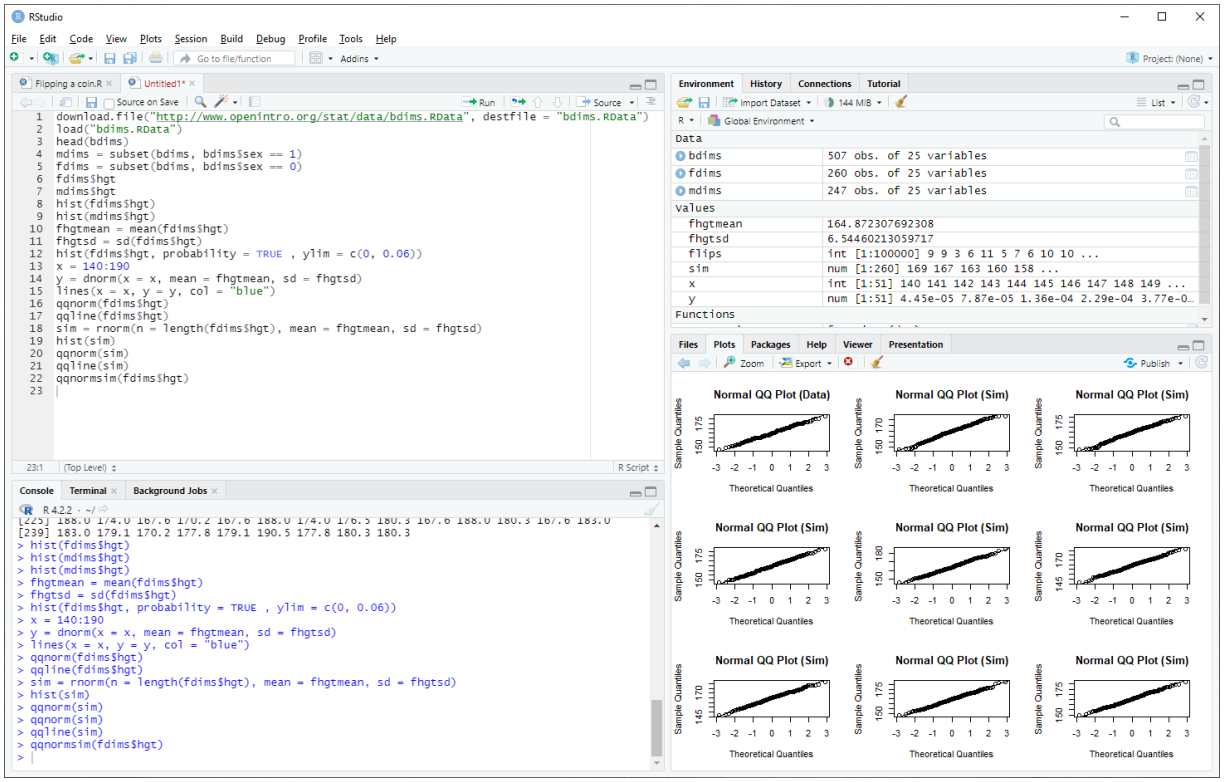


All of the points do not fall on the line, some of the points are over the line while others are under the line. The normal probability plot makes sense when compared to the probability plot for the real data.

Exercise 4:

Does the normal probability plot for fdims$hgt look similar to the plots created for the simulated data? That is, do plots provide evidence that the female heights are nearly normal?

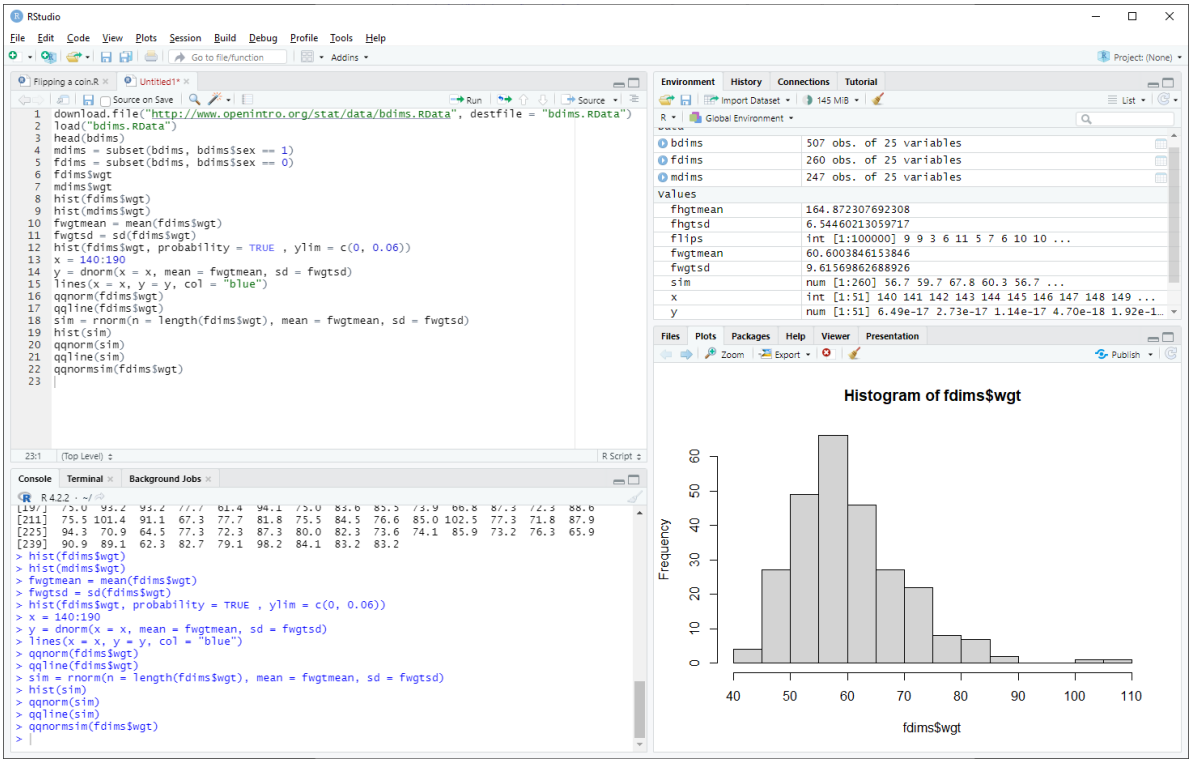


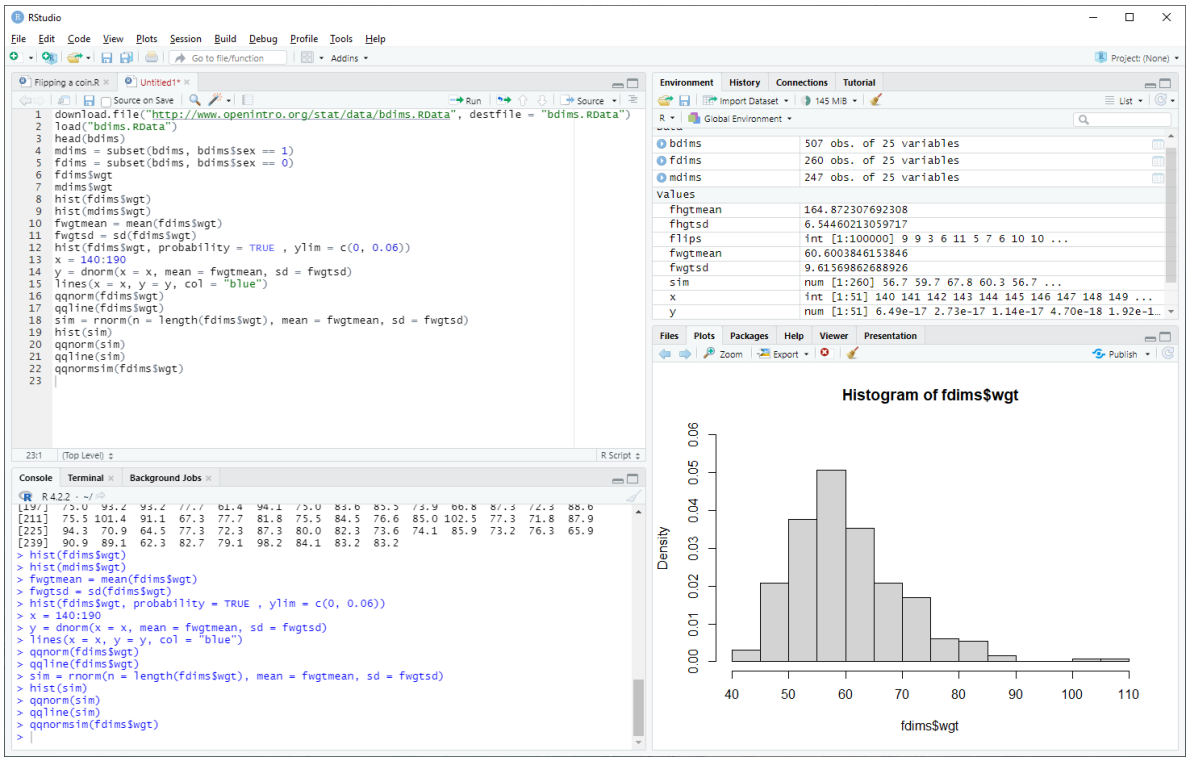


Yes, the simulated data plots provide evidence that the female heights are nearly normal.

Exercise 5:

Using the same technique, determine whether or not female weights appear to come from a normal distribution. If not, how would you describe the shape of this distribution? Note: You may use a histogram to help you decide.





The female weights do not come to a normal distribution, the shape of the distribution looks more like a positive skew than a symmetrical distribution. This can be seen in the histogram plotted with frequency as well as density.

Exercise 6:

Now let's consider some of the other variables in the body dimensions data set. Using the figures on the next page, match the histogram to its normal probability plot. All of the variables have been standardized (first subtract the mean, then divide by the standard deviation), so the units won't be of any help. If you are uncertain based on these figures, generate the plots in R to check.

(a) The histogram for female bi-iliac diameter (bii.di) belongs to normal probability plot letter _B_

(b) The histogram for female elbow diameter (elb.di) belongs to normal probability plot letter _C_

(c) The histogram for general age (age) belongs to normal probability plot letter _D_

(d) The histogram for female chest depth (che.de) belongs to normal probability plot letter _A_

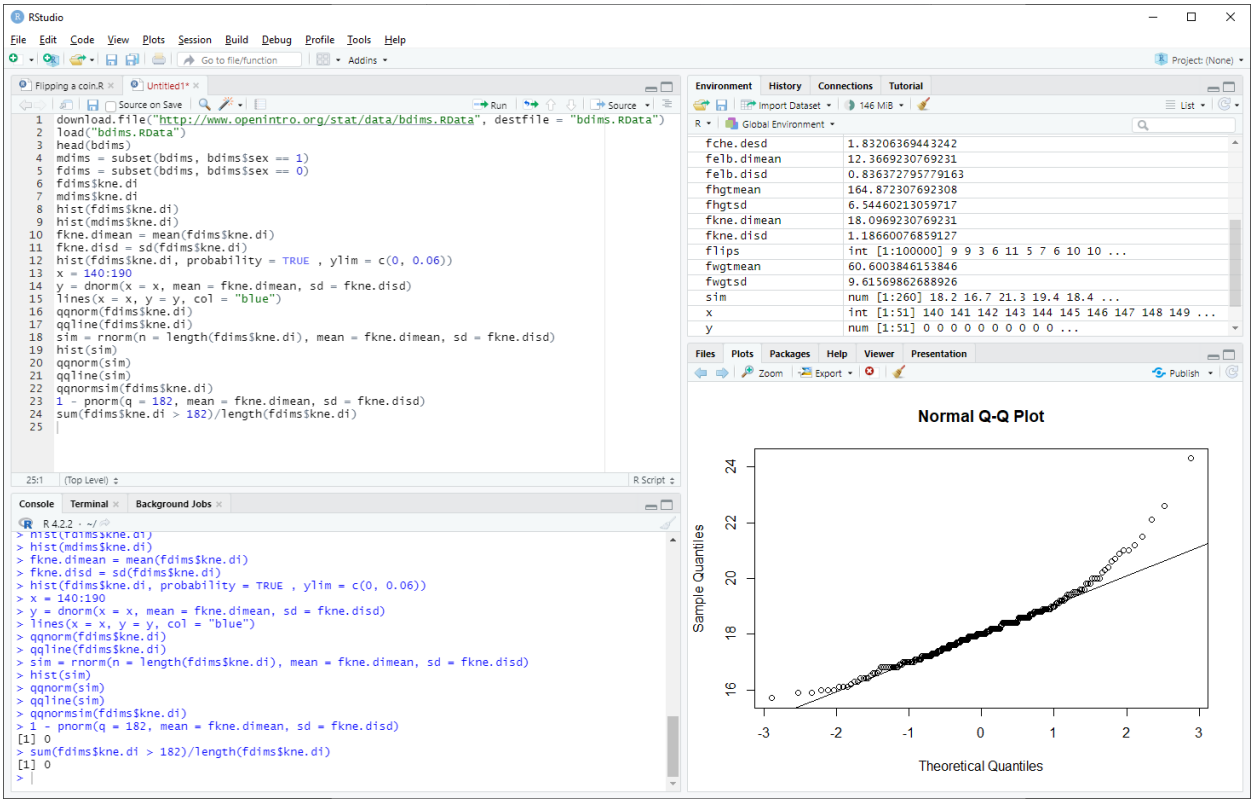
Exercise 7:

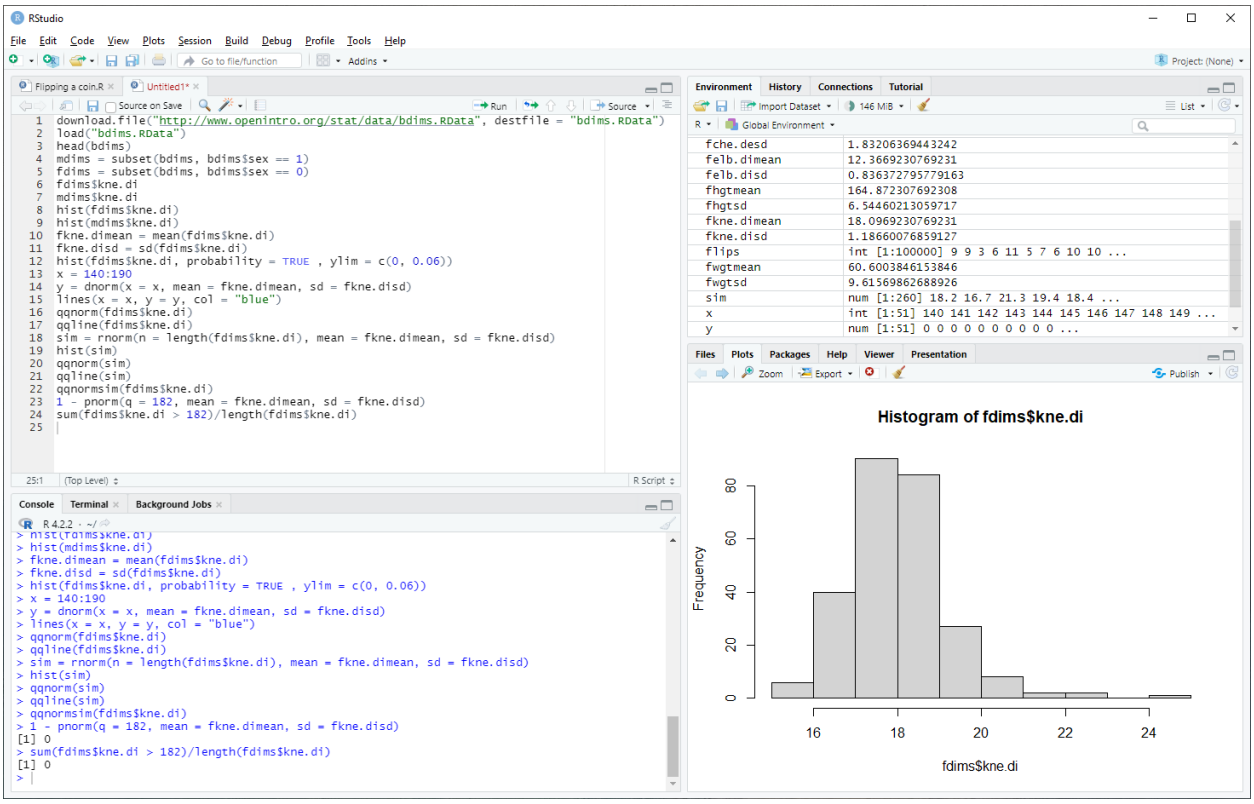
Note that normal probability plots C and D have a slight stepwise pattern. Why do you think this is the case?

Stepwise patterns are occur often in discrete data and the variables elbow diameter (Plot C) was probably not recorded as integers while chest depth (Plot A) probably were.

Exercise 8:

As you can see, normal probably plots can be used both to assess normality and visualize skewness. Make a normal probability plot for female knee diameter (kne.di). Based on this normal probability plot, is this variable left skewed, symmetric, or right skewed? Use a histogram to confirm your findings.





Based on the normal probability plot as well as the histogram, the variable is most likely right skewed.