EDU Stats: Descriptive & Comparative: Assignment #3

**Complete the following tasks:**

**• Using the dataset of your choice, do the following**

**• Import the data using code**

**• Create a subset of your dataset with only the variables you want to use for this assignment**

[1] "TotUsingFrozenEggs" "Diag\_OvulatoryRate" "Diag\_MaleRate"

**Describe the central tendency of the variables**

(cdc2$TotUsingFrozenEggs)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.000 0.000 0.000 2.017 2.000 64.000

Mode:0

(cdc2$Diag\_OvulatoryRate)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

0.0000 0.1000 0.1500 0.1724 0.2100 0.8900 1

Mode:0.13

(cdc2$Diag\_MaleRate)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.000 0.240 0.320 0.337 0.425 1.000

Mode:0.33

**Describe the variables through variation**

(cdc2$TotUsingFrozenEggs)

[1] 28.56247

(cdc2$Diag\_OvulatoryRate)

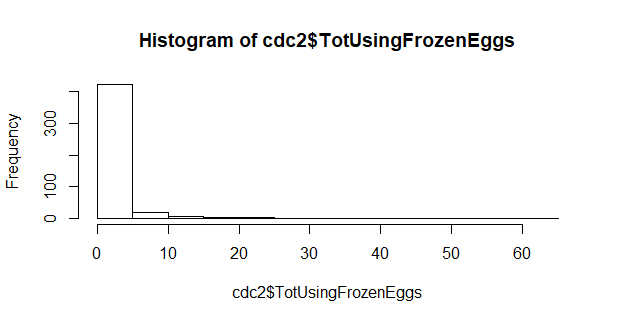
[1] NA

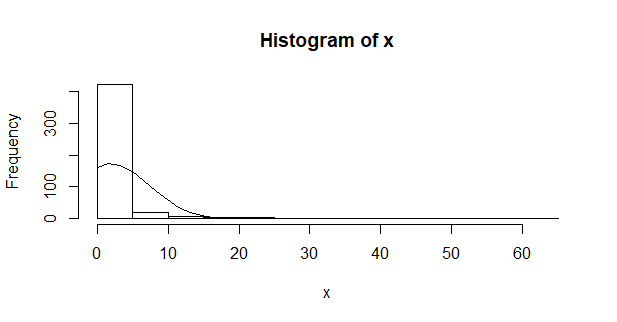
(cdc2$Diag\_MaleRate)

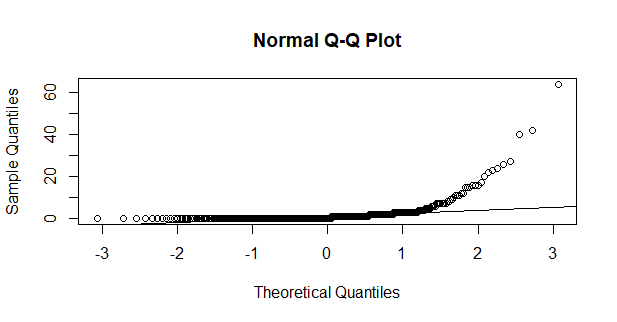
[1] 0.02217369

**Discuss normality. Are your variables approximating normality? What are ways**

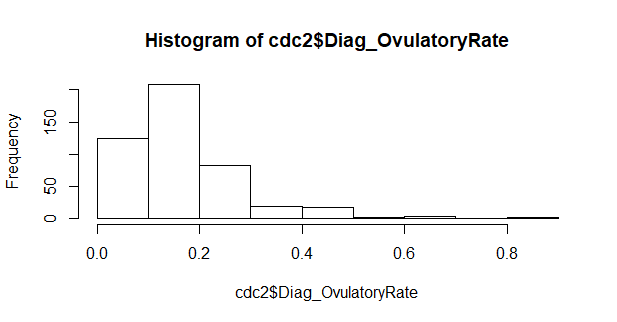
**that you can show that they are or are not?**





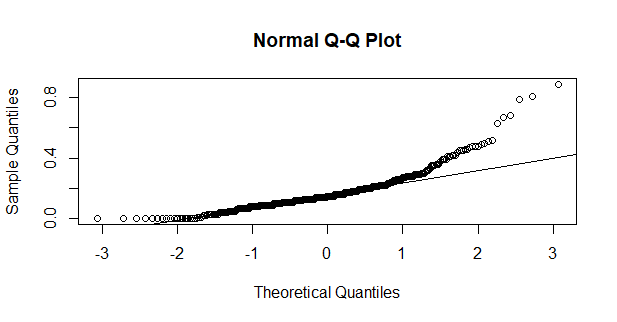


The histogram, distribution line, and q-q plot show this variable is skewed not approximating normality.

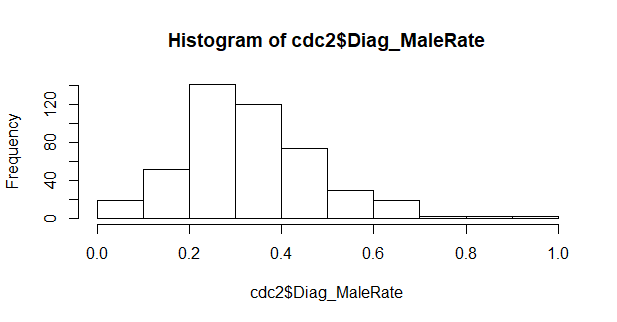


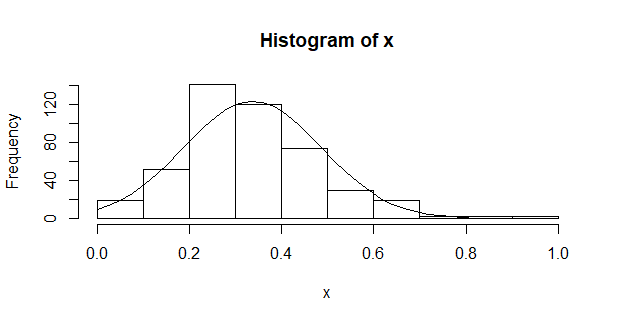
Error in seq.default(min(x), max(x), length = 40) : 'from' must be a finite number

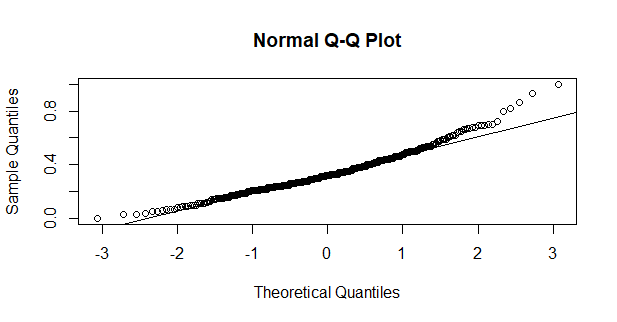
Distribution line code ended in error for this variable.



The histogram, distribution line, and q-q plot show this variable is skewed not approximating normality.







This variable approximates normality better than the other two variables. A few outliers do skew the data. However, not as far as the previous variables.