**Assignment #3**

Complete the following tasks:

* Using the dataset of your choice, do the following
* Import the data using code

library(readr)

Deaths\_and\_Death\_Rates\_by\_Cause2 <- read\_csv("Desktop/Deaths-and-Death-Rates-by-Cause2.csv",

col\_types = cols(`2012` = col\_number(),

`2013` = col\_number(), `2014` = col\_number(),

`2015` = col\_number(), `2016` = col\_number()))

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

Death<-Deaths\_and\_Death\_Rates\_by\_Cause2[c(1:84), c(8,9,10)]

Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 84 obs. of 3 variables:

$ 2014: num 26956958 NA 87750 47774 78069 ...

$ 2015: num 27469114 57580 88255 NA 80527 ...

$ 2016: num 27862596 57734 NA NA 82733 ...

* Choose three variables of interest and complete the following tasks:

o Describe the central tendency of the variables

2014:Mean(866480), Median (141196), Mode (Mode does not exist)

2015: Mean(822645), Median (131705), Mode(Mode does not exist)

2016:Mean(872946), Median(145900), Mode(Mode does not exist)

o Describe the variables through variation

2014: Variance (1.268054e+13), Range(26937022), Sd(3560975)

2015: Variance (1.215033e+13), Range(27435421), Sd(3485732)

2016: Variance (1.310337e+13), Range (27828565), Sd(3619857)

o Discuss normality. Are your variables approximating normality? What are ways that you can show that they are or are not?

2014: 



2015: 



2016: 



**What to turn in**

* Publish your homework to GitHub under your user account, the appropriate format (check syllabus for instructions). This means turn in a separate document from your code that answers the questions above.
* Upload the code you used to GitHub.
* Provide me the links of these two files via e-mail no later than Thursday, October 4th at 6:00pm.