

Homework 1  
Data Science, PCS – UConn, 2022  
Due Date: July 12, 2022 (by 8pm, via HuskyCT)

1. What is wrong with the following R program?

```
d <- 3
y <- D*3
cat(mean(y),"\n")
```

2. Create a variable called Sevens which contains the first 100 integers which are divisible by 7 without remainder.
3. Use the examples shown in class to generate 10000 values from a binomial( $N,p$ ) distribution, and plot a histogram and a boxplot. Do it for  $N=100$  and  $p=0.1, 0.2, \dots, 0.9$ . What happens when you change  $N$  to be 10?
4. We also saw an example of the syntax of a “for” loop. Write a for loop which runs for 100 iterations, and in each one a random sample of size  $n=30$  is drawn from a standard normal distribution (using the rnorm( $n$ ) function). In each iteration print to the screen the mean of the sample, the median, the difference between the mean and the median, and the standard deviation of the sample (with the sd() function). Also, use the examples shown in class to collect all the means from all the iterations and store these 100 values in a variable called allMeans. What is the mean of the vector allMEans? What is its standard deviation?
5. Run the following program which generates the first 15 elements famous Fibonacci sequence:

```
# Fibonacci
n <- 15
fib <- rep(0,n)
fib[1] <- 1
fib[2] <- 1
for (i in 3:n) {
  fib[i] <- fib[i-1] + fib[i-2]
}
cat(fib,"\n")
```

Add R code which will keep the ratio between pairs of consecutive elements in a vector called FibRatio. Note that this vector will have 14 elements. Print the values of FibRatio using the cat() function, and show them as a graph, using the plot() function. What do you observe about the sequence of ratios?