**IST 687 PREP EXERCISE 04**

**Student Name: Bhavish Kumar**

**Prep Exercise No: 04**

**Date Due: 18th September 2019**

1. **Create a Function**
2. Create a new function ‘printVecInfo’ and have it take one numeric vector as its input argument. Here’s a shell of the function:

***ANSWER***

printVecInfo <- function(numVector)  
{  
 # Use print( ) inside of a function when you want some output  
 # to appear on the console.  
 print(mean(numVector) )  
}  
 test <- c(1,2,3,4,5)

printVecInfo(test)

### Mean = 3

1. Make the function print all of the following information for the vector supplied in the argument:
   1. Mean (hey great that is already in the sample above!)
   2. Median
   3. Min and Max
   4. Standard Deviation
   5. 0.05 and .95 quantiles *Hint: use the quantile( ) function.*

***ANSWER***

printVecInfo <- function(numVector)

{

# Use print( ) inside of a function when you want some output

# to appear on the console.

print(mean(numVector)) ### Printing the mean of the vector

print(median(numVector)) ### Printing the median of the vector

print(min(numVector)) ### Printing Min value of the vector

print(max(numVector)) ### Printing Max value of the vector

print(sd(numVector)) ### Printing std deviation of the vector

print(quantile(numVector, probs = c(0.05,0.95))) ### Printing 5th and 95th percentile values

}

printVecInfo(test) ### test <- c(1,2,3,4,5)

### test <- c(1,2,3,4,5)

### Mean = 3

### Median = 3

### Min = 1

### Max = 5

### Standard Deviation = 1.581139

### 5% = 1.2; 95% = 4.8

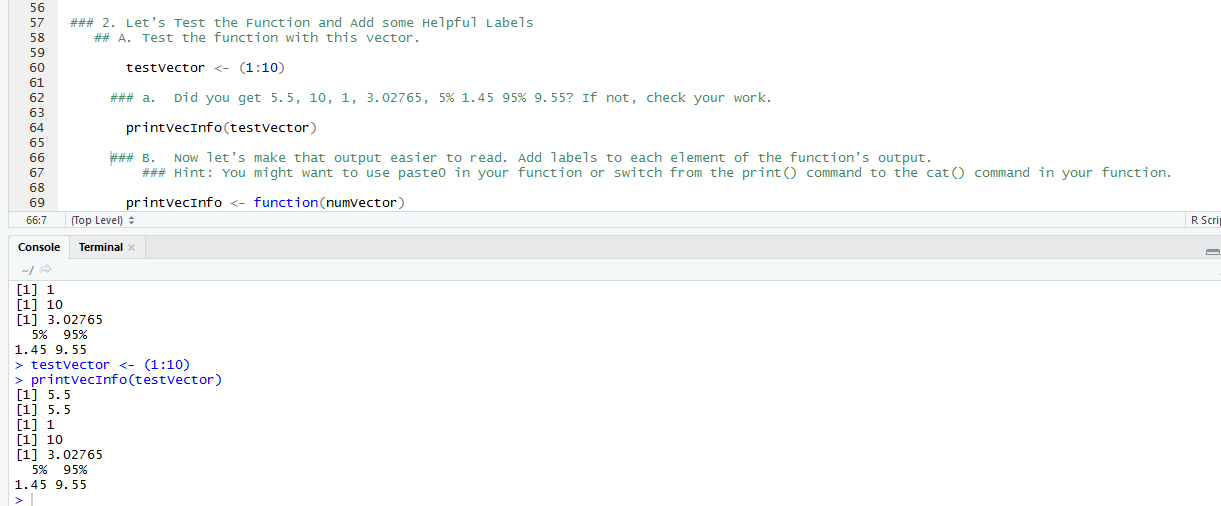
1. **Let’s Test the Function and Add some Helpful Labels**
2. Test the function with this vector. testVector <- (1:10).
   1. Did you get 5.5, 10, 1, 3.02765, 5% 1.45 95% 9.55? If not, check your work.

***ANSWER***

testVector <- (1:10)

printVecInfo(testVector)

Yes, obtained the above values for Mean, Median, Min, Max, Standard Deviation and Quantiles as shown in the screenshot below:



1. Now let’s make that output easier to read. Add labels to each element of the function’s output. *Hint: You might want to use paste0 in your function or switch from the print() command to the cat() command in your function.*

***ANSWER***

printVecInfo <- function(numVector)

{

# Use print( ) inside of a function when you want some output

# to appear on the console.

cat("vector mean is: ",print(mean(numVector)))

cat("vector median is: ",print(median(numVector)))

cat("Min value of vector is: ",print(min(numVector)))

cat("Max value of vector is: ",print(max(numVector

cat("Standard Deviation of vector is: ",print(sd(numVector)))

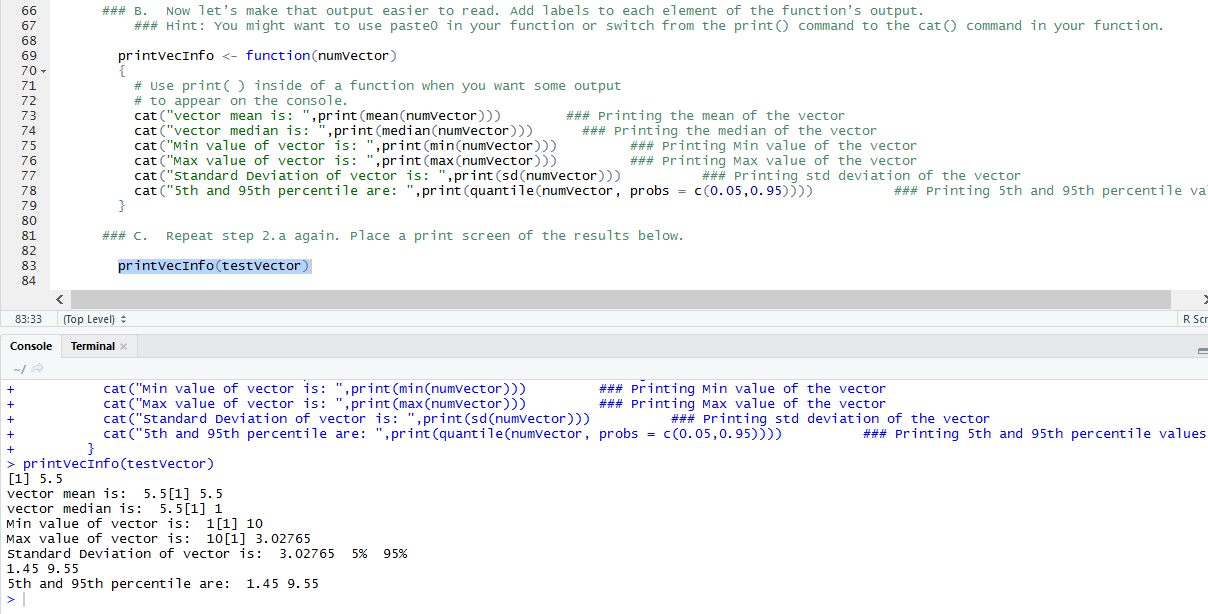
cat("5th and 95th percentile are: ",print(quantile(numVector, probs = c(0.05,0.95))))

}

1. Repeat step 2.a again. Place a print screen of the results below.

***ANSWER***

printVecInfo(testVector)



1. **Explore the dataframe!**
2. For this week’s homework we’ll practice sampling using a built-in data frame called airquality. Copy the dataframe into a new dataframe called myAQdata.

***ANSWER***

myAQdata <- airquality

1. Using the VIEW command, explore the data. You may also want to use the ?command. You may even have to do additional research to help you understand the elements of the dataset.

***ANSWER***

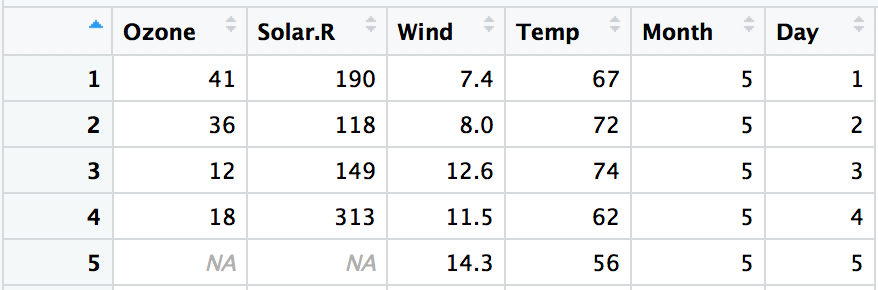
View(myAQdata) ### VIEW Command to explore the data

summary(myAQdata)

str(myAQdata)

### Also using the summary and str commands to further explore the data in more detail.

1. Please use your own words to describe the different elements of the dataset and their values based on the first 5 rows of the dataframe.

****

The dataset has 6 numeric columns which are Ozone, Solar, Wind, Temperature, Month and Day. The first 5 rows show us the Ozone, Solar, Wind and Temperature values for the first 5 days of the fifth month. We can observe that there may be some kind of liner relationship between ozone and wind.