HomeWork4

Katherine Penney

8/5/2021

```
library(moments)
```

Step 1

#Questions 1-2 printVecInfo function as a vector as input and put in functions

```
printVecInfo <- function(m) {
  meanM <- mean(m)
  medianM <- median(m)
  minM <- min(m)
  maxM <- max(m)
  sdM <- sd(m)
  quantile <- quantile(m, probs= c(.05,.95))
  skewM <- skewness(m)
  cat('mean:', meanM, '\n')
  cat('median:', medianM, '\n')
  cat('min:', minM, '\n')
  cat('max:', maxM, '\n')
  cat('std:', sdM, '\n')
  cat('quantile', quantile, '\n')
  cat('Skewness:', skewM, '\n\n')
}
```

Step 1

#Question 3

```
m <- c(1,2,3,4,5,6,7,8,9,10,50)
printVecInfo(m)
```

```
## mean: 9.545455
## median: 6
## min: 1
## max: 50
## std: 13.72125
## quantile 1.5 30
## Skewness: 2.620396
```

Step 2

#Question 4 create variable 'jar' with 50 red and 50 read marbles

```
startjar <- c("redmar","bluemar")
jar <- rep(startjar,50)
```

#Question 5 #Confirm that there are 50 red marbles

```
length(which(jar == 'redmar'))
```

```
## [1] 50
```

#Confirm that there are 50 blue marbles

```
length(which(jar == "bluemar"))
```

```
## [1] 50
```

#Total amount in jar

```
length(jar)
```

```
## [1] 100
```

#Question 6 #Sample 10 marbles from jar

```
thesample <- sample(jar, 10, replace = T)
```

#How many were red? (6)

```
redone <- length(which(thesample == "redmar"))
redone
```

```
## [1] 7
```

#What was the percentage of red marbles? 60%

```
redone/10
```

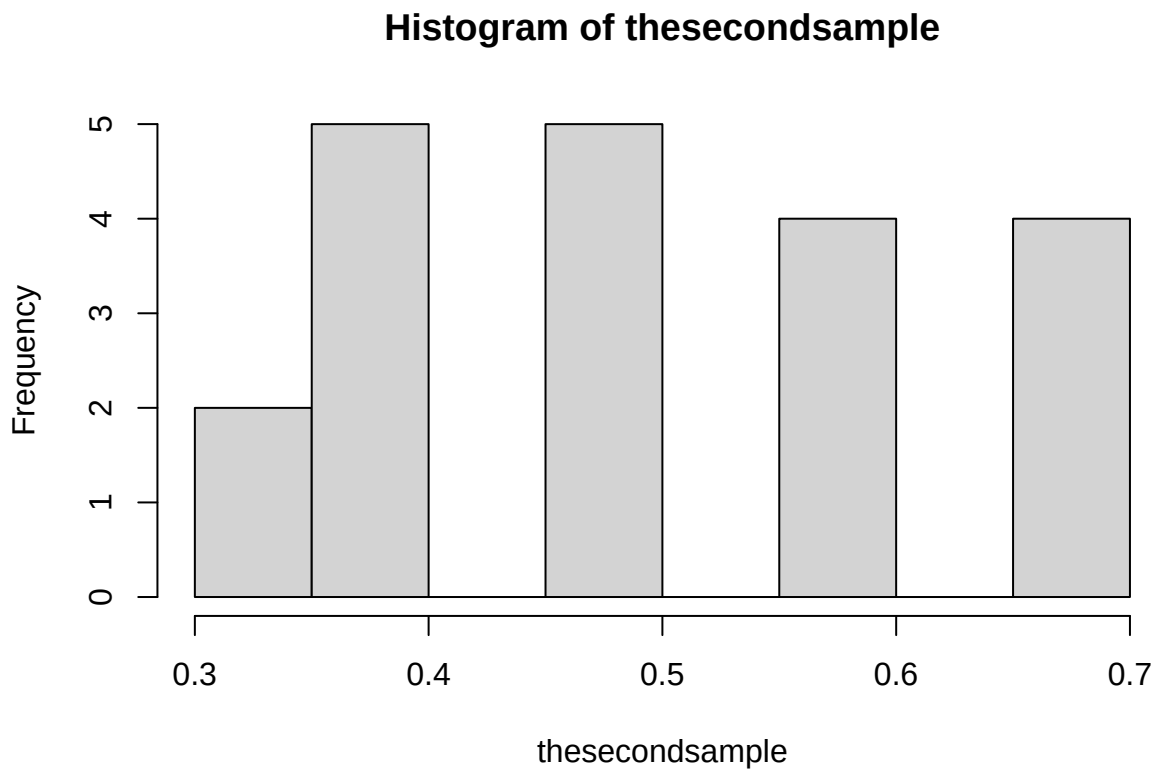
```
## [1] 0.7
```

#Replicate command 20 times #Use printVecInfo

```
thesecondsample <- replicate(20, length(which((sample(jar, 10, replace = T)) == 'redmar'))/10)
```

#Question 7 #Histogram

```
hist(thesecondsample)
```



#Use printVecInfo

```
printVecInfo(thesecondsample)
```

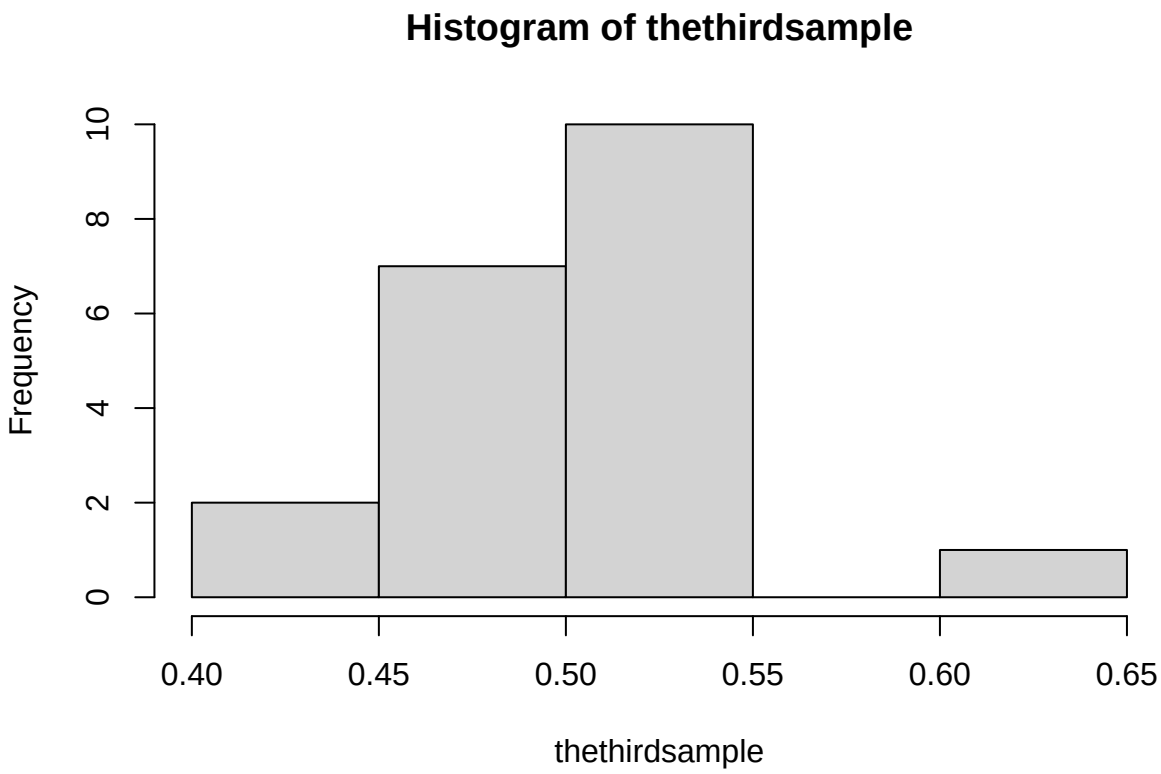
```
## mean: 0.515
## median: 0.5
## min: 0.3
## max: 0.7
## std: 0.1308877
## quantile 0.3 0.7
## Skewness: 0.006863306
```

#Question 8 #100 times sampled from jar

```
thethirdsample <- replicate(20, length(which((sample(jar, 100, replace = T)) == 'redmar'))/100)
```

#Histogram

```
hist(thethirdsample)
```



#Use printVecInfo

```
printVecInfo(tthethirdsample <- replicate(20, length(which((sample(jar, 100, replace = T)) == 'redmar'))
```

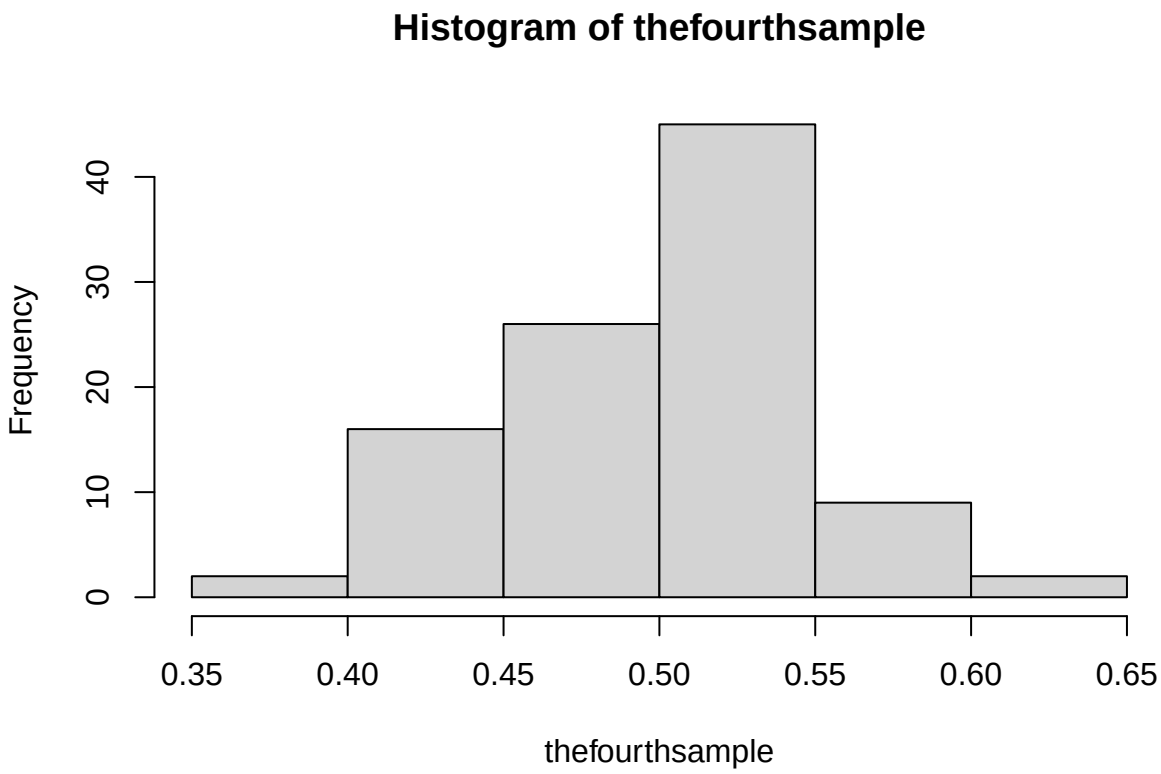
```
## mean: 0.5015
## median: 0.505
## min: 0.41
## max: 0.59
## std: 0.0482619
## quantile 0.4195 0.59
## Skewness: 0.03256155
```

#Question 9 #Replicating the sampling 100 times

```
thefourthsample <- replicate(100, length(which((sample(jar, 100, replace = T)) == 'redmar')))/100)
```

#Histogram

```
hist(thefourthsample)
```



Step 3

#Question 10 #Store airqulaity dataset

```
air <- airquality
summary(air)
```

```
##      Ozone           Solar.R          Wind             Temp
##  Min.   :  1.00   Min.   :  7.0   Min.   : 1.700   Min.   :56.00
##  1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400   1st Qu.:72.00
##  Median : 31.50   Median :205.0   Median : 9.700   Median :79.00
##  Mean   : 42.13   Mean   :185.9   Mean   : 9.958   Mean   :77.88
##  3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500   3rd Qu.:85.00
##  Max.   :168.00   Max.   :334.0   Max.   :20.700   Max.   :97.00
##  NA's   :37       NA's   :7
##      Month            Day
##  Min.   :5.000   Min.   : 1.0
##  1st Qu.:6.000   1st Qu.: 8.0
##  Median :7.000   Median :16.0
##  Mean   :6.993   Mean   :15.8
##  3rd Qu.:8.000   3rd Qu.:23.0
##  Max.   :9.000   Max.   :31.0
##
```

#Question 11 Clean dataset

```
cleandata <-na.omit(air)
cleandata
```

```
##     Ozone Solar.R Wind Temp Month Day
## 1      41     190  7.4   67     5   1
## 2      36     118  8.0   72     5   2
## 3      12     149 12.6   74     5   3
## 4      18     313 11.5   62     5   4
## 7      23     299  8.6   65     5   7
## 8      19      99 13.8   59     5   8
## 9       8      19 20.1   61     5   9
## 12     16     256  9.7   69     5  12
## 13     11     290  9.2   66     5  13
## 14     14     274 10.9   68     5  14
## 15     18      65 13.2   58     5  15
## 16     14     334 11.5   64     5  16
## 17     34     307 12.0   66     5  17
## 18      6      78 18.4   57     5  18
## 19     30     322 11.5   68     5  19
## 20     11      44  9.7   62     5  20
## 21      1       8  9.7   59     5  21
## 22     11     320 16.6   73     5  22
## 23      4      25  9.7   61     5  23
## 24     32      92 12.0   61     5  24
## 28     23      13 12.0   67     5  28
## 29     45     252 14.9   81     5  29
## 30    115     223  5.7   79     5  30
## 31     37     279  7.4   76     5  31
## 38     29     127  9.7   82     6   7
## 40     71     291 13.8   90     6   9
## 41     39     323 11.5   87     6  10
## 44     23     148  8.0   82     6  13
## 47     21     191 14.9   77     6  16
## 48     37     284 20.7   72     6  17
## 49     20      37  9.2   65     6  18
## 50     12     120 11.5   73     6  19
## 51     13     137 10.3   76     6  20
## 62    135     269  4.1   84     7   1
## 63     49     248  9.2   85     7   2
## 64     32     236  9.2   81     7   3
## 66     64     175  4.6   83     7   5
## 67     40     314 10.9   83     7   6
## 68     77     276  5.1   88     7   7
## 69     97     267  6.3   92     7   8
## 70     97     272  5.7   92     7   9
## 71     85     175  7.4   89     7  10
## 73     10     264 14.3   73     7  12
## 74     27     175 14.9   81     7  13
## 76      7      48 14.3   80     7  15
## 77     48     260  6.9   81     7  16
## 78     35     274 10.3   82     7  17
## 79     61     285  6.3   84     7  18
## 80     79     187  5.1   87     7  19
```

```
## 81     63     220 11.5   85     7  20
## 82     16       7  6.9   74     7  21
## 85     80     294  8.6   86     7  24
## 86    108     223  8.0   85     7  25
## 87     20      81  8.6   82     7  26
## 88     52      82 12.0   86     7  27
## 89     82     213  7.4   88     7  28
## 90     50     275  7.4   86     7  29
## 91     64     253  7.4   83     7  30
## 92     59     254  9.2   81     7  31
## 93     39      83  6.9   81     8   1
## 94      9      24 13.8   81     8   2
## 95     16      77  7.4   82     8   3
## 99    122     255  4.0   89     8   7
## 100    89     229 10.3   90     8   8
## 101   110     207  8.0   90     8   9
## 104    44     192 11.5   86     8  12
## 105    28     273 11.5   82     8  13
## 106    65     157  9.7   80     8  14
## 108    22      71 10.3   77     8  16
## 109    59      51  6.3   79     8  17
## 110    23     115  7.4   76     8  18
## 111    31     244 10.9   78     8  19
## 112    44     190 10.3   78     8  20
## 113    21     259 15.5   77     8  21
## 114     9      36 14.3   72     8  22
## 116    45     212  9.7   79     8  24
## 117   168     238  3.4   81     8  25
## 118    73     215  8.0   86     8  26
## 120    76     203  9.7   97     8  28
## 121   118     225  2.3   94     8  29
## 122    84     237  6.3   96     8  30
## 123    85     188  6.3   94     8  31
## 124    96     167  6.9   91     9   1
## 125    78     197  5.1   92     9   2
## 126    73     183  2.8   93     9   3
## 127    91     189  4.6   93     9   4
## 128    47      95  7.4   87     9   5
## 129    32      92 15.5   84     9   6
## 130    20     252 10.9   80     9   7
## 131    23     220 10.3   78     9   8
## 132    21     230 10.9   75     9   9
## 133    24     259  9.7   73     9  10
## 134    44     236 14.9   81     9  11
## 135    21     259 15.5   76     9  12
## 136    28     238  6.3   77     9  13
## 137     9      24 10.9   71     9  14
## 138    13     112 11.5   71     9  15
## 139    46     237  6.9   78     9  16
## 140    18     224 13.8   67     9  17
## 141    13      27 10.3   76     9  18
## 142    24     238 10.3   68     9  19
## 143    16     201  8.0   82     9  20
## 144    13     238 12.6   64     9  21
```

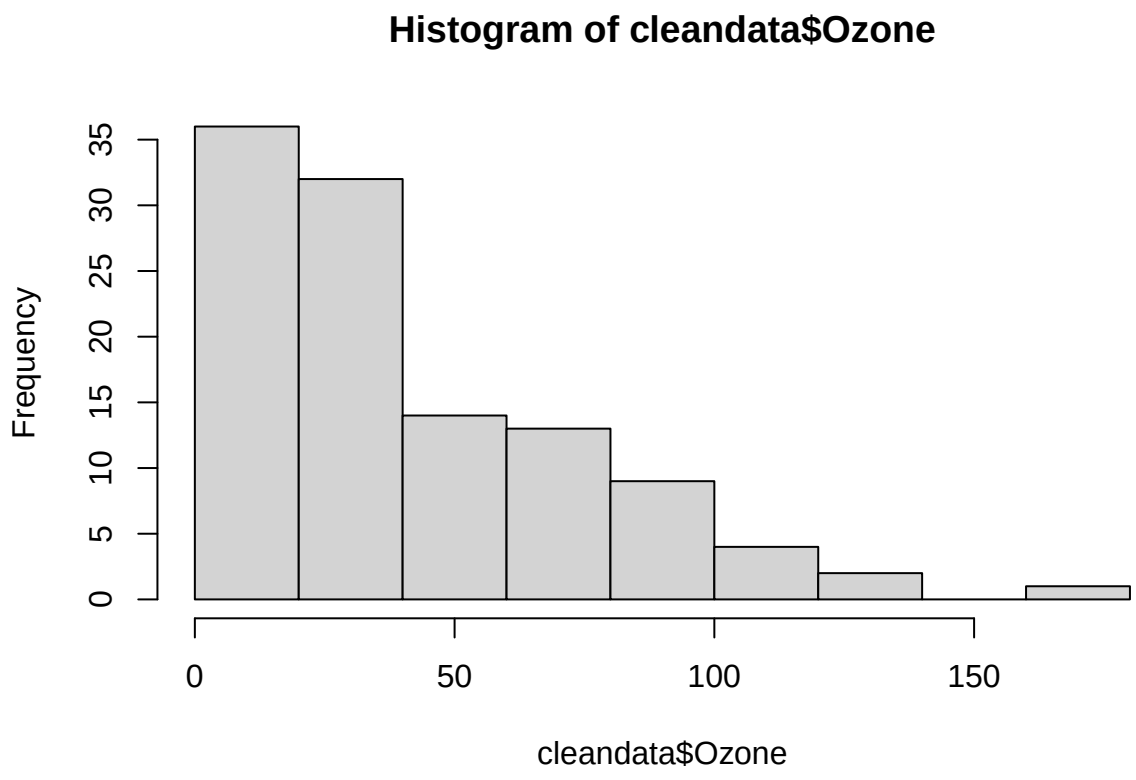
```
## 145    23      14  9.2   71     9  22
## 146    36     139 10.3   81     9  23
## 147     7      49 10.3   69     9  24
## 148    14      20 16.6   63     9  25
## 149    30     193  6.9   70     9  26
## 151    14     191 14.3   75     9  28
## 152    18     131  8.0   76     9  29
## 153    20     223 11.5   68     9  30
```

#Question 12 printVecInfo #Ozone

```
printVecInfo(cleandata$Ozone)
```

```
## mean: 42.0991
## median: 31
## min: 1
## max: 168
## std: 33.27597
## quantile 8.5 109
## Skewness: 1.248104
```

```
hist(cleandata$Ozone)
```

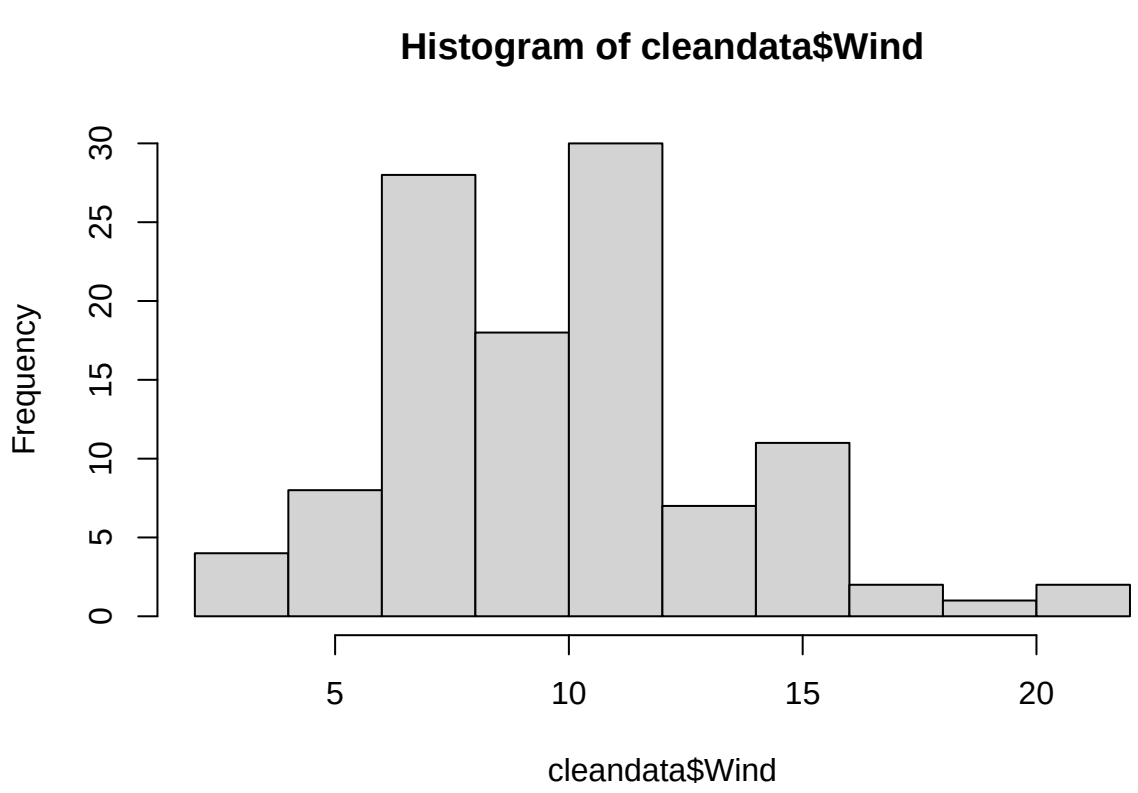


#Wind

```
printVecInfo(cleandata$Wind)
```

```
## mean: 9.93964
## median: 9.7
## min: 2.3
## max: 20.7
## std: 3.557713
## quantile 4.6 15.5
## Skewness: 0.4556414
```

```
hist(cleandata$Wind)
```



#Temp

```
printVecInfo(cleandata$Temp)
```

```
## mean: 77.79279
## median: 79
## min: 57
## max: 97
## std: 9.529969
## quantile 61 92.5
## Skewness: -0.2250959
```

```
hist(cleandata$Temp)
```