

SESSION 5: Data Management Using R

Assignment 1

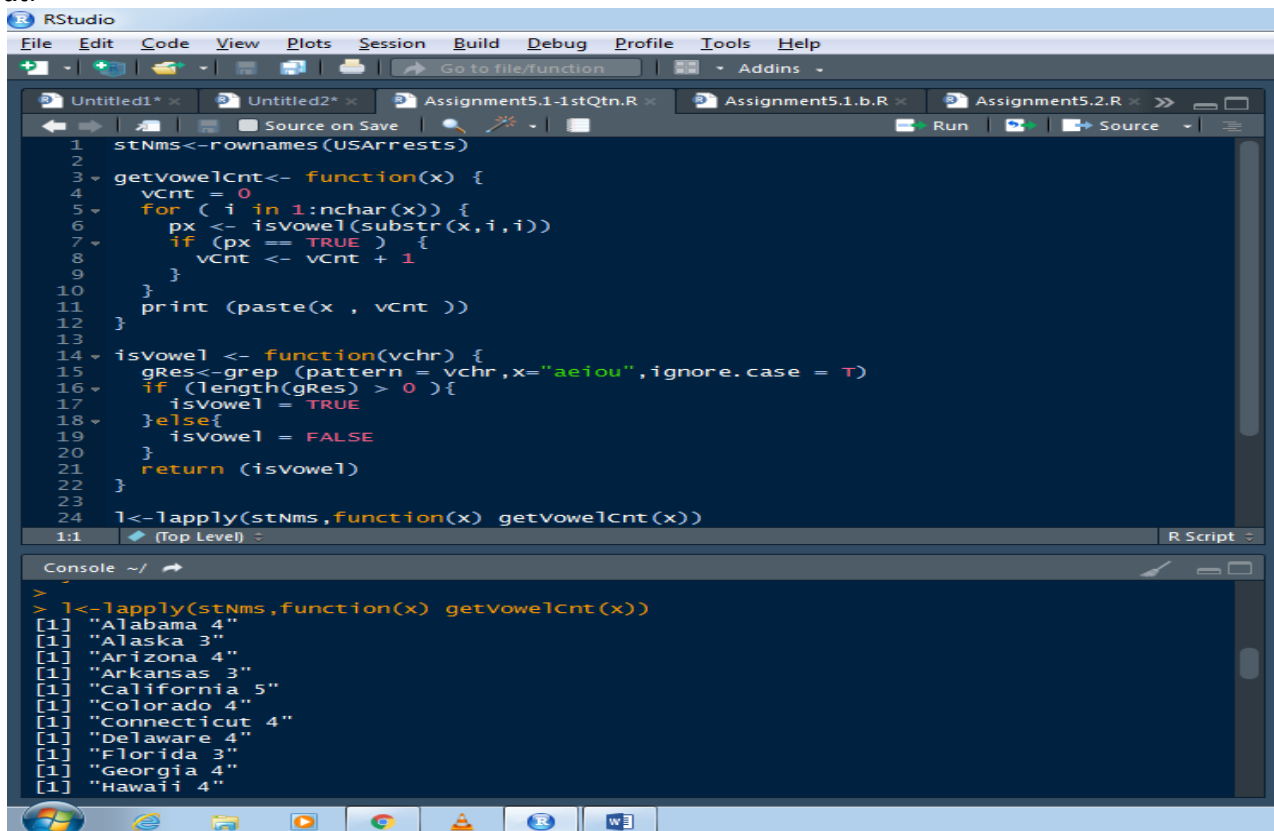
1. How many vowels are there in the names of USA States?

Answer:

```
stNms<-rownames(USArrests)
getVowelCnt<- function(x) {
  vCnt = 0
  for ( i in 1:nchar(x)) {
    px <- isVowel(substr(x,i,i))
    if (px == TRUE ) {
      vCnt <- vCnt + 1
    }
  }
  print (paste(x , vCnt ))
}
isVowel <- function(vchr) {
  gRes<-grep (pattern = vchr,x="aeiou",ignore.case = T)
  if (length(gRes) > 0 ){
    isVowel = TRUE
  }else{
    isVowel = FALSE
  }
  return (isVowel)
}
```

```
l<-lapply(stNms,function(x) getVowelCnt(x))
```

Output:



The screenshot shows the RStudio interface. The source editor contains the R code defined in the previous block. The console shows the output of the `lapply` function, which is a list of character strings representing the state names followed by the count of vowels. The output is as follows:

```
> l<-lapply(stNms,function(x) getVowelCnt(x))
[1] "Alabama 4"
[1] "Alaska 3"
[1] "Arizona 4"
[1] "Arkansas 3"
[1] "California 5"
[1] "Colorado 4"
[1] "Connecticut 4"
[1] "Delaware 4"
[1] "Florida 3"
[1] "Georgia 4"
[1] "Hawaii 4"
```

2. Visualize the vowels distribution.

Answer:

```
vowelLtr<-c("a","e","i","o","u")
vowelCnt<-c(0,0,0,0,0)
df <- data.frame(vowelCnt)

getVowelCnt<- function(x) {
  for ( i in 1:nchar(x)) {
    v <- substr(x,i,i)
    for ( j in 1:length(vowelLtr)) {
      if (toupper(vowelLtr[j]) == toupper(v)) {
        vowelCnt[j] <- vowelCnt[j] + 1
        break
      }
    }
  }
  return (vowelCnt)
}

mainFunc <- function(){
  stNms<-rownames(USArrests)
  df<- cbind(df,lapply(stNms,function(x) vowelCnt<- getVowelCnt(x)))
  print(df)
  vowelCnt<-apply(df,1,sum)
  per <- round(vowelCnt/sum(vowelCnt)*100)
  lbls<-paste(vowelLtr, per, "%", sep = " ")
  pie(vowelCnt, labels=lbls, main="Vowel Distribution")
}

mainFunc()
```

Output:

