Assignment - 1

Session 5 – Data

Management using R

**5. Problem Statement**

> #

> states = rownames(USArrests)

> library(stringr)

> str\_count(states,"a")

[1] 3 2 1 2 2 1 0 2 1 1 2 1 0 2 1 2 0 2 1 2 2 1 1 0 0 2 2 2 1 0 0 0 2 2 0 2 0 2 1 2 2 0

[43] 1 1 0 1 1 1 0 0

> str\_count(tolower(states),"a")

[1] 4 3 2 3 2 1 0 2 1 1 2 1 0 2 1 2 0 2 1 2 2 1 1 0 0 2 2 2 1 0 0 0 2 2 0 2 0 2 1 2 2 0

[43] 1 1 0 1 1 1 0 0

> vowels = c("a","e","i","o","u")

> num\_vowels = vector(mode = "integer",length = 5)

> num\_vowels[j] = sum(num)

> for (j in seq\_along(vowels)) {

+ num = str\_count(tolower(states),vowels[j])

+ num\_vowels[j] = sum(num)

+ }

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

> names(num\_vowels) = vowels

> num\_vowels

a e i o u

61 28 44 36 8

> sort(num\_vowels,decreasing = TRUE)

a i o e u

61 44 36 28 8

1. Visualize the vowels distribution.
2. barplot(num\_vowels, main = "Number of vowels in USA States names",

+ border = "red", ylim = c(0, 80))

