

Assignments for session on "DATA MANAGEMENT USING R"

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

There are 177 vowels

```
USA_States<-rownames(USArrests) # Get the row names, i.e the USA states
USA_States

changed_string1 <-gsub("[^aeiouAEIOU]","c",USA_States)#Substitute Consonants with 'c'
changed_string2<-gsub("[^c]","v",changed_string1)# Replace the remaining letters with 'v'
changed_string2

library(stringr)
str_count(changed_string2,"v") # Counts the occurrences of "v" in a string
sum(str_count(changed_string2,"v")) # Sums up all the occurrences of "v" i.e vowels
# 177 vowel
```

2. Visualize the vowels distribution.

```
USA_States <- rownames(USArrests) # Get the names of states
```

```
USA_States <- paste(USA_States, collapse = "") # Collapse the separate names into a string
```

```
USA_States
```

```
USA_States <- tolower(strsplit(USA_States, "")[[1]]) # Convert each letter into lower case and splitting it
```

```
USA_States
```

```
USA_States <- USA_States[USA_States %in% letters]
```

```
table(USA_States) # Get the occurrences of each letter
```

```
spread_df <- as.data.frame(table(USA_States)) # Convert the table into a data frame
```

```
names(spread_df)
```

```
colnames(spread_df) <- c("letters", "Count") # Give the desired column names
```

```
names(spread_df)
```

```
library(dplyr)
```

```
filter(spread_df, letters %in% c("a","e","i","o","u")) # Finding number of vowels in df
```

```
#There are 61 a, 28 e , 44 i , 36 o and 8 u in the names of USA States
```

```
# type 1
```

```
vowel_count <- filter(spread_df, letters %in% c("a","e","i","o","u"))
```

```
vowel_count
```

```
barplot(vowel_count$Count, axes = TRUE, axisnames = TRUE, xlab = "Vowels", ylab = "Count",  
col = "red")
```

