

Session 5 – Data

Management using R

Assignment - 1



Contents

[Introduction](#page3) [2](#page3)



[Objective](#page3) [2](#page3)



[Prerequisites](#page3) [2](#page3)



[Associated Data Files](#page3) [2](#page3)



[Problem Statement](#page3) [2](#page3)



[Expected Output](#page3) [2](#page3)



|  |  |
| --- | --- |
| Copyrights© 2017, AcadGild. All Rights Reserved | 1 |
|  |  |



**Introduction**



This assignment will help you to understand the key concepts learnt in this session.

**Objective**



This assignment will test your skills on Data Management Functions in R.

**Prerequisites**



Not Applicable

**Associated Data Files**



Not Applicable

**Problem Statement**



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

**View(USArrests)**

**names(USArrests) #to get the column names**

**US\_States<-rownames(USArrests) # to get the row names, i.e the USA states**

**US\_States**

**b<-gsub("[^aeiouAEIOU]","C",US\_States)#is basically "if not any of 'aeiouAEIOU' then 'C'**

**#if aeiouAEIOU is NOT in [], its a complete pattern.**

**# using ^ means, exclude what comes after this**

**b<-gsub("[^C]","V",b)# is "if not 'C', then 'V**

**b**

**library(stringr)**

**str\_count(b,"V") #counts character in a string**

**sum(str\_count(b,"V")) # adds the count of "V" i.e vowels**

**# 177 Vowels**

1. Visualize the vowels distribution.

**vowel\_dist <- filter(distribution, letters %in% c("a","e","i","o","u"))**

**vowel\_dist**

**barplot(vowel\_dist$Freq, axes = TRUE, axisnames = TRUE, xlab = "Vowels", ylab = "frequency")**

**Expected Output**



Not Applicable

|  |  |
| --- | --- |
| Copyrights© 2017, AcadGild. All Rights Reserved | 2 |
|  |  |