**INPUT :**

#create data frame with village and household information

data <- data.frame(village =1:10, households=c(2,2,3,5,6,4,8,2,4,3))

#calculate the probability for each village based on the number of households

sum(data$households)

#set the random seed for of reproducibility

set.seed(123)

#select a sample of 6 village using the pps method ]

sampled\_village <- sample(data$village,size=6,prob=data$probability)

#display the sampled village

print(sampled\_village)

**OUTPUT :**

> sum(data$households)

[1] 39

> print(sampled\_village)

[1] 3 10 2 8 6 9