> library(readxl)

> data=read\_excel("C:\\Users\\Dell\\Documents\\for knn1.xlsx",col\_names = TRUE)> head(data) #Top observations present in the data

# A tibble: 6 × 10

overall total score of individual (ou…¹ physical health scor…² `mental health`

<dbl> <dbl> <dbl>

1 49 60 60

2 60 60 55

3 45 60 50

4 61 20 50

5 78 100 60

6 69 73.3 55

> dim(data) #Check the dimensions of the data

[1] 460 10

> summary(data) #Summarise the data

physical health score(out of 100) mental health

Min. : 20.00 Min. :25.00

1st Qu.: 40.00 1st Qu.:50.00

Median : 53.33 Median :55.00

Mean : 55.43 Mean :56.14

3rd Qu.: 66.67 3rd Qu.:65.00

Max. :100.00 Max. :90.00

> install.packages("class")

> library(class)

> data1=data[,-1]

> dim(data1)

[1] 460 9

> View(data1)

> #Normalization

> normalize=function(x) {

+ return ((x - min(x)) / (max(x) - min(x))) }

> norm=as.data.frame(lapply(data1[,1:2], normalize))

> set.seed(123)

> data\_spl=sample(1:nrow(norm),size=nrow(norm)\*0.7,replace = FALSE)

> train2=data1[data\_spl,] # 70% training data

> test2=data1[-data\_spl,] # remaining 30% test data

> train\_labels=data1[data\_spl,]

> test\_labels=data1[-data\_spl,]

> knn\_pred=knn(train=train2[,-9], test=test2[,-9],train\_labels$y,k=2)

> confusionmatrix=table(knn\_pred,test\_labels$y)

> confusionmatrix

knn\_pred 0 1

0 117 5

1 2 14

> accuracy=sum(diag(confusionmatrix)/nrow(test\_labels))

> accuracy

[1] 0.9492754

