Load the iris dataset

data("iris")

names(iris)

new\_data<-subset(iris, select = c(-Species))

new\_data

cl<-kmeans(new\_data, 3)

cl

data<-new\_data

wss<-sapply(1:15, function(k){kmeans(data, k)$tot.withinss})

wss

plot(1:15, wss, type='b', pch=19, frame = FALSE, xlab = "Number of clusters K", ylab ="Total within-sum of squares")

install.packages("cluster")

library(cluster)

clusplot(new\_data, cl$cluster, color = TRUE, shade = TRUE, labels = 2, lines = 0)

cl$cluster

cl$centers

clusters<-hclust(dist(iris[, 3:4]))

plot(clusters)

clusterCut<-cutree(clusters, 3)

table(clusterCut, iris$Species)

install.packages("ggplot2")

library(ggplot2)

ggplot (iris, aes (Petal.Length, Petal.Width, color = Species)) +

geom\_point (alpha = 0.4, size = 3.5) + geom\_point (col = clusterCut) +

scale\_color\_manual (values = c("black", "red", "green"))

install.packages("cluster")

install.packages("factoextra")

# Load the required libraries

library(cluster)

library(factoextra)

library(datasets)

# Load the Iris dataset

data("iris")

iris\_data <- iris[, 1:4] # Selecting only the numeric columns

# Determine the optimal number of clusters using the elbow method

set.seed(123)

wss <- numeric(10)

for (i in 1:10) {

kmeans\_model <- kmeans(iris\_data, centers = i, nstart = 10)

wss[i] <- kmeans\_model$tot.withinss

}

# Plot the elbow method graph

plot(1:10, wss, type = "b", xlab = "Number of Clusters",

ylab = "Within-cluster Sum of Squares")

# Determine the optimal number of clusters using silhouette coefficient

silhouette <- numeric(10)

for (i in 2:10) {

kmeans\_model <- kmeans(iris\_data, centers = i, nstart = 10)

silhouette[i] <- silhouette\_avg(kmeans\_model$cluster, dist(iris\_data))

}

# Plot the silhouette coefficient graph

plot(2:10, silhouette[2:10], type = "b", xlab = "Number of Clusters",

ylab = "Average Silhouette Width")

# Perform K-means clustering with the selected number of clusters

optimal\_clusters <- 3 # You can choose the optimal number from the plots