##in group perform kmenas cluster with 3, 4 groups and 5 groups. Discuss what you found

k3 <- kmeans(df, centers = 3, nstart = 25)

k4 <- kmeans(df, centers = 4, nstart = 25)

k5 <- kmeans(df, centers = 5, nstart = 25)

p2 <- fviz\_cluster(k3, geom = "point", data = df) + ggtitle("k = 3")

p3 <- fviz\_cluster(k4, geom = "point", data = df) + ggtitle("k = 4")

p4 <- fviz\_cluster(k5, geom = "point", data = df) + ggtitle("k = 5")

library(gridExtra)

grid.arrange(p1, p2, p3, p4, nrow = 2)

#In group perform the distance. How many clusters will you retain?

###distance

distance\_x <- get\_dist(df\_com)

fviz\_dist(distance\_x, gradient = list(low = "#00AFBB", mid = "white", high = "#FC4E07"))

fviz\_nbclust(df\_com, kmeans, method = "wss")

gap\_stat <- clusGap(df\_com, FUN = kmeans, nstart = 25,

K.max = 10, B = 50)

fviz\_gap\_stat(gap\_stat)

###two clusters

k2\_com <- kmeans(df\_com, centers = 2, nstart = 50)

fviz\_cluster(k2\_com, data = df\_com)