Kmeans -clustering

heartdata= read.csv("dataset/heart.csv")

#Exploring data with 0,1 values as they won't be of much help in making the clusters

heartdata\_c = heartdata[,-which(colnames(heartdata) %in% c("sex","fbs","exng","output"))]

# Standardize the dataset and apply K-means clustering

heartdata\_scale = scale(heartdata\_c)

km <- kmeans(heartdata\_scale, centers = 2, nstart = 10)

km

heartdata\_clustered <- as\_tibble(heartdata\_scale) %>% add\_column(cluster = factor(km$cluster))

ggplot(heartdata\_clustered,aes(x=heartdata\_clustered$age,y = heartdata\_clustered$thalachh, color = cluster)) + geom\_point()

#actual output

ggplot(heartdata, aes(x = heartdata$age, y = heartdata$thalachh, color = heartdata$output)) + geom\_point()