K-Means clustering is unsupervised machine learning because there is not a target variable. Clustering can be used to create a target variable, or simply group data by certain characteristics.

Here’s a great and simple way to use R to find clusters, visualize and then tie back to the data source to implement a marketing strategy.

setwd

#import dataset

ABC <-read.table("AbcBank.csv",header=TRUE,

sep=",")

#choose variables to be clustered

# make sure to exclude ID fields or Dates

ABC\_num<- ABC[,2:5]

#scale the data! so they are all normalized

ABC\_scaled <-as.data.frame(scale(ABC\_num))

#kmeans function

k3<- kmeans(ABC\_scaled, centers=3, nstart=25)

#library with the visualization

library(factoextra)

fviz\_cluster(k3, data=ABC\_scaled,

ellipse.type="convex",

axes =c(1,2),

geom="point",

label="none",

ggtheme=theme\_classic())

#check out the centers

# remember these are normalized but

#higher values are higher values for the original data

k3$centers

#add the cluster to the original dataset!

ABC$Cluster<-as.numeric(k3$cluster)

Check out our awesome clusters: