Bowling.R

Suraj Durgesht

16EC06

Fri Mar 8 13:35:17 2019

odibowling2007 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2007odibowlingrating.csv")  
odibowling2008 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2008odibowlingrating.csv")  
odibowling2009 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2009odibowlingrating.csv")  
odibowling2010 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2010odibowlingrating.csv")  
odibowling2011 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2011odibowlingrating.csv")  
odibowling2012 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2012odibowlingrating.csv")  
odibowling2013 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2013odibowlingrating.csv")  
odibowling2014 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2014odibowlingrating.csv")  
odibowling2015 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2015odibowlingrating.csv")  
odibowling2016 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2016odibowlingrating.csv")  
  
testbowling2007 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2007testbowlingrating.csv")  
testbowling2008 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2008testbowlingrating.csv")  
testbowling2009 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2009testbowlingrating.csv")  
testbowling2010 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2010testbowlingrating.csv")  
testbowling2011 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2011testbowlingrating.csv")  
testbowling2012 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2012testbowlingrating.csv")  
testbowling2013 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2013testbowlingrating.csv")  
testbowling2014 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2014testbowlingrating.csv")  
testbowling2015 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2015testbowlingrating.csv")  
testbowling2016 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2016testbowlingrating.csv")  
  
twentybowling2007 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2007twenty20bowlingrating.csv")  
twentybowling2008 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2008twenty20bowlingrating.csv")  
twentybowling2009 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2009twenty20bowlingrating.csv")  
twentybowling2010 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2010twenty20bowlingrating.csv")  
twentybowling2011 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2011twenty20bowlingrating.csv")  
twentybowling2012 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2012twenty20bowlingrating.csv")  
twentybowling2013 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2013twenty20bowlingrating.csv")  
twentybowling2014 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2014twenty20bowlingrating.csv")  
twentybowling2015 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2015twenty20bowlingrating.csv")  
twentybowling2016 <- read.csv("D:\\Programming\\DA\\Lab 4\\Player Ratings\\2016twenty20bowlingrating.csv")  
  
  
  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

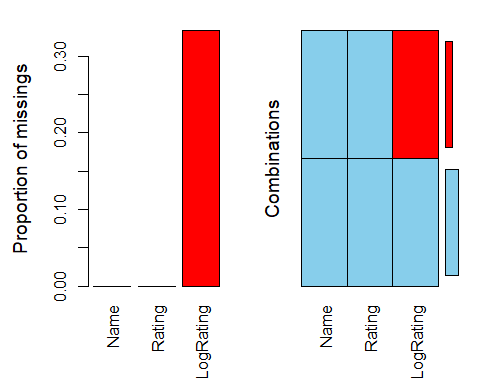
dataBowling <- bind\_rows(odibowling2007, odibowling2008, odibowling2009,odibowling2010,  
 odibowling2011, odibowling2012, odibowling2013, odibowling2014,  
 odibowling2015, odibowling2016,  
 testbowling2007, testbowling2008, testbowling2009, testbowling2010,  
 testbowling2011, testbowling2012, testbowling2013, testbowling2014,  
 testbowling2015, testbowling2016,  
 twentybowling2007, twentybowling2008, twentybowling2009,twentybowling2010,

summary(dataBowling)

## Name Rating LogRating   
## Length:3000 Min. : 0.0 Min. :0.8451   
## Class :character 1st Qu.:268.8 1st Qu.:2.4669   
## Mode :character Median :420.0 Median :2.6405   
## Mean :400.0 Mean :2.5258   
## 3rd Qu.:538.0 3rd Qu.:2.7437   
## Max. :916.0 Max. :2.9619   
## NA's :1000

library(VIM)

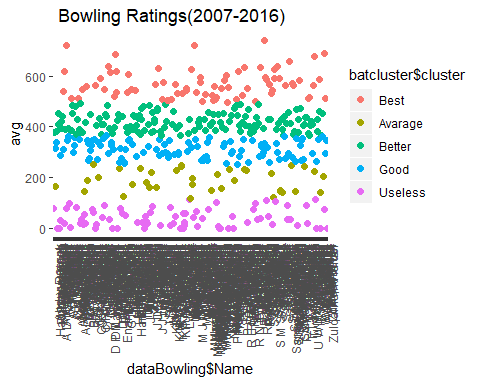
aggr(dataBowling)



dataBowling <- dataBowling %>%  
 group\_by(Name) %>%  
 summarise(avg = mean(Rating))  
  
set.seed(20)  
  
batcluster <- kmeans(dataBowling[, 2], 5)  
  
batcluster$cluster <- as.factor(batcluster$cluster)  
  
str(batcluster)

## List of 9  
## $ cluster : Factor w/ 5 levels "1","2","3","4",..: 5 3 2 4 3 4 5 3 3 5 ...  
## $ centers : num [1:5, 1] 570.2 189.2 417 310.5 39.8  
## ..- attr(\*, "dimnames")=List of 2  
## .. ..$ : chr [1:5] "1" "2" "3" "4" ...  
## .. ..$ : chr "avg"  
## $ totss : num 13640182  
## $ withinss : num [1:5] 249143 62750 170198 106112 91460  
## $ tot.withinss: num 679664  
## $ betweenss : num 1.3e+07  
## $ size : int [1:5] 78 37 149 103 78  
## $ iter : int 3  
## $ ifault : int 0  
## - attr(\*, "class")= chr "kmeans"

library(ggplot2)  
ggplot(dataBowling, aes(dataBowling$Name, avg, color = batcluster$cluster)) +  
 geom\_point(size = 2) +  
 scale\_color\_hue(labels = c("Best", "Avarage", "Better", "Good", "Useless")) +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 ggtitle(" Bowling Ratings(2007-2016)")



ggplot(dataBowling, aes(dataBowling$Name,avg,color = batcluster$cluster)) +  
 geom\_point(shape = 10,size = 5,stroke = 2) +  
 scale\_color\_hue(labels = c("Avarage", "Better", "Useless", "Good", "Best")) +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 ggtitle("ODI Bowling Ratings(2007-2016)")

