2\_norm\_stand.r

abhia

2019-11-13

# IMPORT BY USING THE SUITABLE ENCODING  
fifa = read.csv( "1\_cleaned\_data.csv")  
  
#SELECTS NUMERICAL COLUMNS ONLY  
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

numfifa = select\_if(fifa, is.numeric)  
  
#NORMALIZE FUNTION USING MIN-MAX VALUES  
normalize <- function(x){  
 return ((x-min(x))/(max(x)-min(x)))  
}  
  
#NORMALIZING THE DATA USING USER-DEFINED NORMALIZED FUNCTIONS  
#UNITIZATION WITH ZERO MINIMUM ((x-min)/range)  
mmfifa <- as.data.frame(lapply(numfifa, normalize))  
write.csv(mmfifa, "2\_minmax.csv", row.names = FALSE)  
  
#NORMALIZING THE DATA THROUGH NATURAL LOGARITHMS  
logfifa = as.data.frame(lapply(numfifa, log))  
write.csv(logfifa, "2\_log.csv", row.names = FALSE)  
  
#STANDARDIZATION ((x-mean)/sd)  
library("clusterSim")

## Loading required package: cluster

## Loading required package: MASS

##   
## Attaching package: 'MASS'

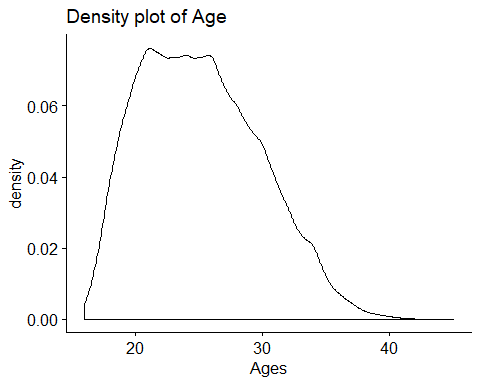
## The following object is masked from 'package:dplyr':  
##   
## select

sfifa = data.Normalization(numfifa, type="n1", normalization = "column")  
write.csv(sfifa, "2\_standard.csv", row.names = FALSE)  
  
#NORMALIZATION ((x-mean)/sqrt(sum((x-mean)^2)))  
nfifa = data.Normalization(numfifa, type="n12", normalization = "column")  
write.csv(nfifa, "2\_norm.csv", row.names = FALSE)  
  
  
# #NORMALITY TEST  
  
# This library is already loaded: library("dplyr")  
library("ggpubr")

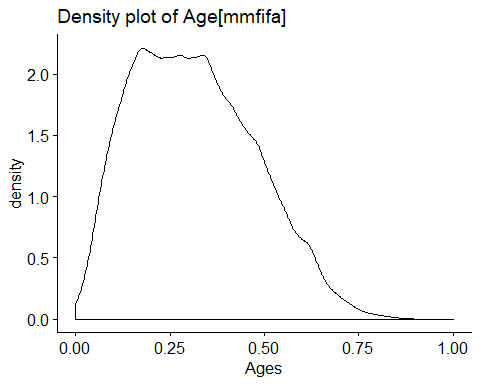
## Loading required package: ggplot2

## Loading required package: magrittr

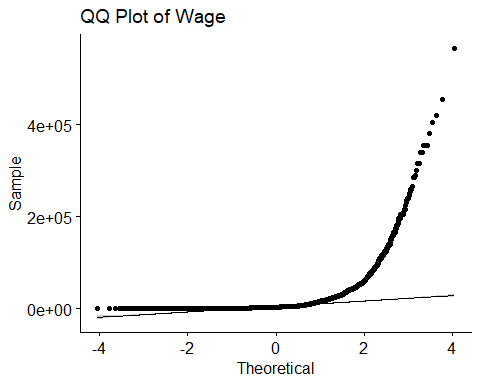
#DENSITY PLOT OF AGE  
print(ggdensity(fifa$Age, main="Density plot of Age", xlab = "Ages"))



#DENSITY PLOT OF AGE[mmfifa]  
print(ggdensity(mmfifa$Age, main="Density plot of Age[mmfifa]", xlab = "Ages"))



#QQ PLOT OF WAGE  
print(ggqqplot(fifa$Wage, main="QQ Plot of Wage"))



#QQ PLOT OF WAGE[sfifa]  
print(ggqqplot(sfifa$Wage, main="QQ Plot of Wage[sfifa]"))

