#### ROTINA R PARA CÁLCULO DAS FRONTEIRAS DE EFICIÊNCIA - CUSTOS OPERACIONAIS EFICIENTES

#### ROTINA BÁSICA PARA REALIZAÇÃO DAS ESTIMAÇÕES POR ANÁLISE ENVOLTORIA DE DADOS

# Limpar os dados do R

rm(list=ls())

## O script utilizar os seguintes pacotes

if (!require(openxlsx)) {

install.packages('openxlsx')

library(openxlsx)

}

if (!require(Benchmarking)) {

install.packages('Benchmarking')

library(Benchmarking)

}

if (!require(svDialogs)) {

install.packages('svDialogs')

library(svDialogs)

}

if (!require(plotly)) {

install.packages('plotly')

library(plotly)

}

if (!require(dplyr)) {

install.packages('dplyr')

library(dplyr)

}

dlgMessage('Selecione o arquivo "COE\_2021"')

base\_dados <- read.xlsx(file.choose(),sheet = "Snis", startRow = 2)

View(base\_dados)

## SelecÌ§aÌƒo de prestadores que ofertam os dois tipos de serviço (abastecimento de água e esgotamento sanitário)

base\_corte1 <- subset(base\_dados, serv == 'Água e Esgoto')

View(base\_corte1)

# ExclusÃ£o de autarquias

base\_corte2 <- subset(base\_corte1, natjur != 'Autarquia')

View(base\_corte2)

# Escala mÃ­nima de serviÃ§o (número de economias de Água) de pelo menos 10% da COPASA

base\_corte2B <- subset(base\_corte2, prest == 'COPASA')

econ\_agua\_med\_copasa <- mean(base\_corte2B$econativas\_a)

econativas\_a\_med <- with(base\_corte2,tapply(econativas\_a,prest,mean,na.rm=TRUE))

econativas\_a\_med <- data.frame(row.names(econativas\_a\_med),econativas\_a\_med)

names(econativas\_a\_med)[1] <-'prest'

base\_corte2 <- merge(base\_corte2,econativas\_a\_med,by='prest')

base\_corte3 <- base\_corte2[base\_corte2$econativas\_a\_med>=econ\_agua\_med\_copasa\*0.1,]

View(base\_corte3)

# Nível Máximo de Hidrometração = 70%

hidrom\_med <- aggregate(hidrom~prest,data=base\_corte3,FUN=mean)

names(hidrom\_med)[2]<-'hidrom\_med'

base\_corte3<- merge(base\_corte3,hidrom\_med,by='prest')

base\_corte4<-subset(base\_corte3,base\_corte3$hidrom\_med >= 70)

View(base\_corte4)

## base\_final

base\_final<-base\_corte4

View(base\_final)

## Construção das perdas (1000 m³/ano)

base\_final$perdas = with(base\_final, volprod\_a + volimp\_a - volcons\_a)

## A modelagem DEA utiliza as médias para os anos de 2016 a 2019 para produtos e insumos

base\_medias <- aggregate(cbind(opex\_defl,econativas\_a,econativas\_e,esg\_tratado, perdas)~ prest, data=base\_final, FUN=mean)

summary(base\_medias)

## Escolhendo insumos e produtos

nx = c("opex\_defl", "perdas" )

ny = c("econativas\_a", "econativas\_e", "esg\_tratado")

X=as.matrix(base\_medias%>% select(opex\_defl, perdas))

rownames(X)<- base\_medias$prest

Y=as.matrix(base\_medias%>% select(econativas\_a, econativas\_e, esg\_tratado))

rownames(Y)<- base\_medias$prest

X\_16 = as.matrix(base\_final%>% filter(ano == 2016) %>% select(opex\_defl, perdas) )

rownames(X\_16)<- base\_medias$prest

Y\_16 = as.matrix(base\_final%>% filter(ano == 2016) %>% select(econativas\_a, econativas\_e, esg\_tratado))

rownames(Y\_16)<- base\_medias$prest

X\_17 = as.matrix(base\_final%>% filter(ano == 2017) %>% select(opex\_defl, perdas) )

rownames(X\_17)<- base\_medias$prest

Y\_17 = as.matrix(base\_final%>% filter(ano == 2017) %>% select(econativas\_a, econativas\_e, esg\_tratado))

rownames(Y\_17)<- base\_medias$prest

X\_18 = as.matrix(base\_final%>% filter(ano == 2018) %>% select(opex\_defl, perdas) )

rownames(X\_18)<- base\_medias$prest

Y\_18 = as.matrix(base\_final%>% filter(ano == 2018) %>% select(econativas\_a, econativas\_e, esg\_tratado))

rownames(Y\_18)<- base\_medias$prest

X\_19 = as.matrix(base\_final%>% filter(ano == 2019) %>% select(opex\_defl, perdas) )

rownames(X\_19)<- base\_medias$prest

Y\_19 = as.matrix(base\_final%>% filter(ano == 2019) %>% select(econativas\_a, econativas\_e, esg\_tratado))

rownames(Y\_19)<- base\_medias$prest

gm\_mean = function(x, na.rm=TRUE){

exp(sum(log(x[x > 0]), na.rm=na.rm) / length(x))

}

## CALCULAR O DEA - COM RESTRIÇÕES

dea.arsae <- function(X , Y, XREF = NULL, YREF = NULL) {

if (is.null(XREF)) {XREF = X} else {XREF = XREF}

if (is.null(YREF)) {YREF = Y} else {YREF = YREF}

orgKr <- dim(X)

m = dim(X)[2] # Insumos

n = dim(Y)[2] # Produtos

K = dim(X)[1] # DMUs

okr <- orgKr[1]

Kd <- 0

# Saida

u <- matrix(NA, K, m) # Peso Insumo

v <- matrix(NA, K, n) # Peso Produto

v\_p <- matrix(NA, K, n) # Peso composto do produto

objval <- rep(NA, K) # Vetor da FO

gamma <- matrix(NA, K, 1) # Peso escala

# Implementacao

lps <- make.lp(1 + K , m + n + 1)

rownames(lps)[1]=paste(rownames(lps)[1],c("DMU"))

rownames(lps)[1:K+1]=paste(rownames(lps)[2:K+1],rownames(X))

colnames(lps)=c(nx,ny,"scale")

for (i in 1:m) set.column(lps,i, c(0, -XREF[, i])) # Custos

for (h in 1:n) set.column(lps, m + h, c(0, YREF[, h])) # Produtos

set.column(lps, m + n + 1, c(0, rep(1, K))) # Escala

{ lp.control(lps, sense = "max");set.rhs(lps, 1, 1)}

# Restricoes de pesos

res.p<-matrix(0,8,5)

# | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# | Max | Min | Max | Min | Max | Min | Max | Min |

res.p[,1]<-c( - 8128 , 1209, - 695 , 305 , - 1275 , 150 , - 9882090 , 4186076 ) # Custos Operacionais

res.p[,2]<-c( 1, -1, 0, 0, 0, 0, 0, 0) # Perdas

res.p[,3]<-c( 0, 0, 1, -1, 0, 0, 0, 0) # econativas\_a\_med

res.p[,4]<-c( 0, 0, 0, 0, 1, -1, 0, 0) # econativas\_e\_med

res.p[,5]<-c( 0, 0, 0, 0, 0, 0, 1, -1) # esg\_tratado\_med

for(k in 1:dim(res.p)[1])

{

add.constraint(lps, c(res.p[k,],0), "<=", 0)

}

teste=rep(NA,K)

# Laco para todas as empresas

for (k in 1:K)

{

set.row(lps, 1, c(X[k, ]), 1:m)

objrow <- c(rep(0, m), Y[k, ], 1)

set.objfn(lps, objrow)

status <- solve(lps)

teste[k]=status

objval[k] <- get.objective(lps)

pesos <- get.variables(lps)

u[k, ] <- pesos[1:m] # Peso insumo

v[k, ] <- pesos[(m + 1):(m + n)] # Peso produto

gamma[k, ] <- pesos[(m + n + 1)] # Peso escala

}

# Teste

if (sum(teste)>0)

{

paste("########## ATENCAO! OCORREU ALGUM ERRO NA OTIMIZACAO DA DMU DE NUMERO:", which(teste>0),"##########")

}

# Relatorio de saida

for (z in 1:ncol(YREF))

{

v\_p[,z]<- v[,z]\*YREF[,z] # Pesos ponderados pela variavel

}

colnames(v\_p)=paste("Y.",ny,sep="")

colnames(v)=paste("Vs\_Y.",ny,sep="")

colnames(X)=nx

colnames(Y)=ny

colnames(u)=nx

p.sombra.u1<-matrix(0,nrow=nrow(YREF),ncol=ncol(YREF))

p.sombra.u2<-matrix(0,nrow=nrow(YREF),ncol=ncol(XREF))

for (z in 1:nrow(YREF)){

p.sombra.u1[z,]<- v[z,]/u[z,1]

}

for (z in 1:nrow(XREF)){

p.sombra.u2[z,]<- u[z,]/u[z,1]

}

colnames(p.sombra.u1)<-paste("v.",colnames(YREF),"/u1",sep="")

colnames(p.sombra.u2)<-paste("u.2",colnames(XREF),"/u1",sep="")

re = list(Eficiencia = objval, pesos.u = u, pesos.v = v, preco.sombra.v = p.sombra.u1, preco.sombra.u = p.sombra.u2 )

return(re)

}

## Cálculo COE (DEA com restrições)

e <- dea.arsae(X, Y)

COE <- data.frame(base\_medias$prest, e$Eficiencia, X, Y, e$preco.sombra.v, e$preco.sombra.u)

format(COE,decimal.mark=",")

## Cálculo do FP entre 2016 e 2017

e00 <- dea.arsae(X\_16, Y\_16)

e00 <- e00$Eficiencia

e10 <- dea.arsae(X\_17, Y\_17, XREF=X\_16,YREF=Y\_16)

e10 <- e10$Eficiencia

e11 <- dea.arsae(X\_17, Y\_17)

e11 <- e11$Eficiencia

e01 <- dea.arsae(X\_16, Y\_16, XREF=X\_17,YREF=Y\_17)

e01 <- e01$Eficiencia

tc\_16a17 <- sqrt(e10/e11 \* e00/e01)

desloc\_17 <- gm\_mean (tc\_16a17)

tc\_16a17 <- data.frame(base\_medias$prest,tc\_16a17)

colnames(tc\_16a17)<- c("Prestador","Mudança Tecnológica")

format(tc\_16a17,decimal.mark=",")

## Cálculo do FP entre 2017 e 2018

e00 <- dea.arsae(X\_17, Y\_17)

e00 <- e00$Eficiencia

e10 <- dea.arsae(X\_18, Y\_18, XREF=X\_17,YREF=Y\_17)

e10 <- e10$Eficiencia

e11 <- dea.arsae(X\_18, Y\_18)

e11 <- e11$Eficiencia

e01 <- dea.arsae(X\_17, Y\_17, XREF=X\_18,YREF=Y\_18)

e01 <- e01$Eficiencia

tc\_17a18 <- sqrt(e10/e11 \* e00/e01)

desloc\_18 <- gm\_mean (tc\_17a18)

tc\_17a18 <- data.frame(base\_medias$prest,tc\_17a18)

colnames(tc\_17a18)<- c("Prestador","Mudança Tecnológica")

format(tc\_17a18,decimal.mark=",")

## Cálculo do FP entre 2018 e 2019

e00 <- dea.arsae(X\_18, Y\_18)

e00 <- e00$Eficiencia

e10 <- dea.arsae(X\_19, Y\_19, XREF=X\_18,YREF=Y\_18)

e10 <- e10$Eficiencia

e11 <- dea.arsae(X\_19, Y\_19)

e11 <- e11$Eficiencia

e01 <- dea.arsae(X\_18, Y\_18, XREF=X\_19,YREF=Y\_19)

e01 <- e01$Eficiencia

tc\_18a19 <- sqrt(e10/e11 \* e00/e01)

desloc\_19 <-gm\_mean (tc\_18a19)

tc\_18a19 <-data.frame(base\_medias$prest,tc\_18a19)

colnames(tc\_18a19)<- c("Prestador","Mudança Tecnológica")

format(tc\_18a19,decimal.mark=",")

deslocs <- cbind(desloc\_17,desloc\_18,desloc\_19)

FP <- gm\_mean(deslocs) - 1

Tab\_NT <- cbind(deslocs,FP)

colnames(Tab\_NT) <- c("Deslocamento 16/17","Deslocamento 17/18","Deslocamento 18/19","Fator de Produtividade")

# Planilha com os dados de saída

wb <- createWorkbook()

addWorksheet(wb, "COE")

addWorksheet(wb, "Bienio\_16\_17")

addWorksheet(wb, "Bienio\_17\_18")

addWorksheet(wb, "Bienio\_18\_19")

addWorksheet(wb, "FP")

writeData(wb, "COE", COE, startCol = 2, startRow = 3, rowNames = TRUE)

writeData(wb, "Bienio\_16\_17", tc\_16a17, startCol = 2, startRow = 3, rowNames = TRUE)

writeData(wb, "Bienio\_17\_18", tc\_17a18, startCol = 2, startRow = 3, rowNames = TRUE)

writeData(wb, "Bienio\_18\_19", tc\_18a19, startCol = 2, startRow = 3, rowNames = TRUE)

writeData(wb, "FP", Tab\_NT, startCol = 2, startRow = 3, rowNames = TRUE)

saveWorkbook(wb, "Fator\_Produtividade.xlsx", overwrite = TRUE)