**CODE FOR COMPARISON WORD CLOUD**

library(qdap)

library(dplyr)

library(tm)

library(wordcloud)

library(plotrix)

library(dendextend)

library(ggplot2)

library(ggthemes)

library(RWeka)

library(reshape2)

library(quanteda)

recom\_n<-subset(comm, subset=comm$no)

recom\_n<-comm$no

View(recom\_n)

recom\_y<-comm$si

corpus\_review\_y=Corpus(VectorSource(recom\_y))

corpus\_review\_n=Corpus(VectorSource(recom\_n))

corpus\_review\_n=tm\_map(corpus\_review\_n, tolower)

corpus\_review\_y=tm\_map(corpus\_review\_y, removePunctuation)

corpus\_review\_n=tm\_map(corpus\_review\_n, removePunctuation)

corpus\_review\_y=tm\_map(corpus\_review\_y, removeWords, stopwords("it"))

corpus\_review\_n=tm\_map(corpus\_review\_n, removeWords, stopwords("it"))

term\_count\_y <- freq\_terms(corpus\_review\_y, 20)

plot(term\_count\_y)

term\_count\_n <- freq\_terms(corpus\_review\_n, 20)

plot(term\_count\_n)

all\_yes <- paste(recom\_y, collapse = "")

all\_no <- paste(recom\_n, collapse = "")

all\_combine <- c(all\_yes, all\_no)## Creating corpus for combination

corpus\_review\_all=Corpus(VectorSource(all\_combine))

corpus\_review\_all=tm\_map(corpus\_review\_all, tolower)#Remove punctuation

corpus\_review\_all=tm\_map(corpus\_review\_all, removePunctuation)#Remove stopwords

corpus\_review\_all=tm\_map(corpus\_review\_all, removeWords, stopwords("english"))

corpus\_review\_all=tm\_map(corpus\_review\_all, removeWords,c("also", "get","like", "company", "made", "can", "im", "dress","just","i"))#Stem document

corpus\_review\_all=tm\_map(corpus\_review\_all, stemDocument)

review\_tdm\_all <- TermDocumentMatrix(corpus\_review\_all)

all\_m=as.matrix(review\_tdm\_all)

colnames(all\_m)=c("Yes","No")#Sum rows and frequency data frame

review\_term\_freq\_all <- rowSums(all\_m)

review\_word\_freq\_all <- data.frame(term=names(review\_term\_freq\_all), num = review\_term\_freq\_all)#Make commonality cloud

commonality.cloud(all\_m,

colors = "steelblue1",

max.words = 50)

# Create comparison cloud

comparison.cloud(all\_m,

colors = c("green", "red"),

max.words = 50)