WORD CLOUD

#packages for text mining and word cloud

library(tm)

library(wordcloud)

# getting the directory

#getwd()

# setting the directory

#setwd("F:\\rahul\\DS-Class\\DS\_class\\Logistic")

#reading the csv file

wrd\_cld = read.csv("comments.csv")

#Step 1 = to corpus

crps = Corpus(VectorSource(wrd\_cld$Comments))

class(crps)

#shows data at second sqbkts data

crps[[1]][1]

#show complete data

strwrap(crps)

#show 1st pos data

strwrap(crps)[12]

#Step 2 = Lower case

crps = tm\_map(crps, content\_transformer(tolower))

#Step 3 = Remove nos

crps = tm\_map(crps,removeNumbers)

#Step 4 = Remove stop words

crps = tm\_map(crps,removeWords,stopwords("english"))

#Step 5 = Remove Punctuation

crps = tm\_map(crps,removePunctuation)

#Step 6 = Stemming

crps = tm\_map(crps,stemDocument)

#Step 7 = Remove Selected Stopwords

crps = tm\_map(crps,removeWords,c("get","tools","game"))

#Removing unnecessary white spaces

crps = tm\_map(crps,stripWhitespace)

#Step 8 = TDM

tdm = TermDocumentMatrix(crps)

#Step 9 = To matrix and also data frame

mat = as.matrix(tdm)

v = sort(rowSums(mat), decreasing = TRUE) #getting d ttl occurance

df = data.frame(word = names(v),Frequency = v) #writing it into datafarame

#Step 10 = Word Cloud

wordcloud(df$word,df$Frequency)

#control structure

wordcloud(df$word,df$Frequency,

random.order = FALSE,

rot.per = 0.3,

scale = c(4,0.5),

max.words = 100,

colors = brewer.pal(8,"Dark2"))

#give title

title(main = "WORD CLOUD")