**INTERNAL PRACTICAL**

**AIM:**

To implement basic functions and commands in R Programming. To build WordCloud, a text mining method using R for easy to understand and better visualization than a data table.

**PRACTICAL:**

library("tm")

library("SnowballC")

library("wordcloud")

library("RColorBrewer")

**# library("Rcpp")**

**# load file**

txt = readLines(file.choose())

**# corpus funnction**

docs = Corpus(VectorSource(txt))

toSpace <- content\_transformer(function (x, pattern ) gsub(pattern, " ", x))

docs <- tm\_map(docs, toSpace, "/")

docs <- tm\_map(docs, toSpace, "@")

docs <- tm\_map(docs, toSpace, "\\|")

docs <- tm\_map(docs, toSpace, "#")

**# Convert the text to lower case**

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

**# Remove numbers**

docs = tm\_map(docs, removeNumbers)

**# Remove white spaces**

docs = tm\_map(docs, stripWhitespace)

**# Remove english common stopwords**

docs <- tm\_map(docs, removeWords, stopwords("english"))

**# Remove your own stop word, specify your stopwords as a character vector**

docs <- tm\_map(docs, removeWords, c("blablal", "blabla2"))

**# Remove punctuations**

docs <- tm\_map(docs, removePunctuation)

dtm <- TermDocumentMatrix(docs)

m <- as.matrix(dtm)

v <- sort(rowSums(m), decreasing=TRUE)

d <- data.frame(word = names(v), freq=v)

head(d, 10)

wordcloud(words = d$word,

freq = d$freq,

min.freq = 1,

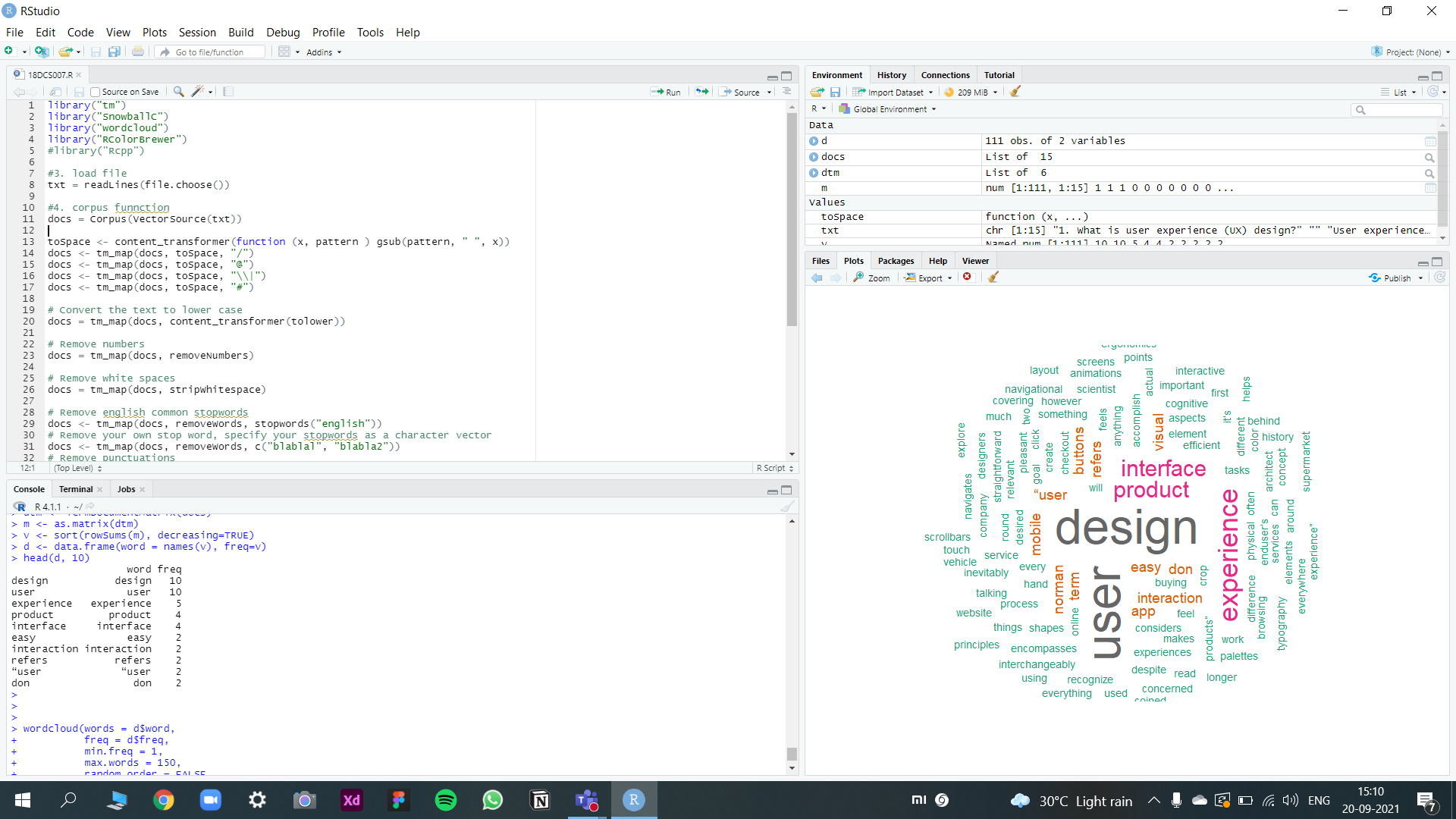
max.words = 150,

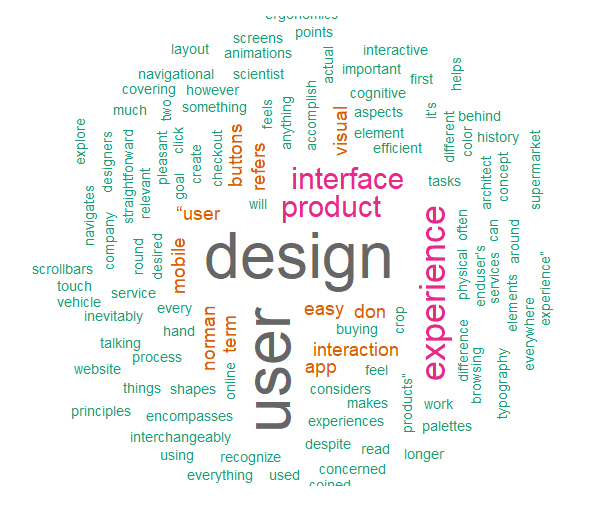
random.order = FALSE,

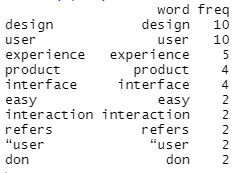
rot.per = 0.35,

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

**OUTPUT:**







**CONCLUSION:**

In this practical, we learnt about R and implemented Word Cloud using R.