library(KoNLP)

library(wordcloud)

library(RColorBrewer)

install.packages('devtools')

library(devtools)

library(wordcloud2)

devtools::install\_github("lchiffon/wordcloud2")

#단어사전 불러옴

useSejongDic()

useNIADic()

#텍스트파일 읽어 옴

setwd("c:/r") #본인 파일 위치로 수정

getwd()

data<-readLines("c:/r/인천 아띠아모-세제.txt",encoding = "UTF-8")

#1차 데이터 정제, 공백, 숫자,영어, 한글 외 삭제

data1 <- gsub("[^ 가-힣]","", data)

data1

head(data1)

#단어사전이용해서 단어추출

nouns <-extractNoun(data1)

nouns

data6<-unlist(nouns)

data6 #단어로 분리됨

#80개 단어들 빈도수 많은 것 부터 보기

wordcount<-table(data6)

head(sort(wordcount, decreasing=T),80)

#위에서 필요없는 단어들 버리기

#불필요한 단어 수작업으로 삭제 후 삭제 됐는지 다시 확인 필요

data6 <- gsub('세제','',data6)

data6 <- gsub('니예','',data6)

data6 <- gsub('해서','',data6)

data6 <- gsub('하시','',data6)

wordcount<-table(data6)

head(sort(wordcount, decreasing=T),80)

#하고 싶은 사람만 하세요!!(한글자지만 의미 있을 수 있음 ex)옷 등등)

#data6 <- data6[nchar(data6)>=2] # 2글자 이상 단어만 추출

#data6

#상위 50개 단어 워드 클라우드 돌리기

wordcount<-head(sort(wordcount, decreasing=T),50)

palete <- brewer.pal(5,"Set1") #글꼴 색, 글꼴 색 확인 display.brewer.all()

windowsFonts(word\_font=windowsFont("맑은 고딕")) #글꼴

wordcloud2(data=wordcount, size=2, col="random-dark", rotateRatio= 0)