

2 문제 1

인터넷에서 적당한 한글 텍스트 자료 하나를 내려받아서 워드클라우드를 만드시오.

```
library(rJava)
## Warning: package 'rJava' was built under R version 3.4.4

library(KoNLP)
## Warning: package 'KoNLP' was built under R version 3.4.4
## Checking user defined dictionary!

library(dplyr)
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(wordcloud)
## Warning: package 'wordcloud' was built under R version 3.4.4
## Loading required package: RColorBrewer

setwd('C:/2018-2/wo/')
text<- readLines('nambuk.txt')
head(text)
## [1] "남과 북은 남북 관계의 전면적이며 획기적인 개선과 발전을 이룩함으로써
##      끊어진 민족의 혈맥을 잇고 공동번영과 자주통일의 미래를 앞당겨 나갈 것이다.
## "
## [2] "남북관계를 개선하고 발전시키는 것은 온 겨레의 한결같은 소망이며 더 이상
##      미룰 수 없는 시대의 절박한 요구이다.
## "
## [3] "남과 북은 우리 민족의 운명은 우리 스스로 결정한다는 민족 자주의 원칙을
##      확인하였으며 이미 채택된 남북 선언들과 모든 합의들을 철저히 이행함으로써 관계
##      개선과 발전의 전환적 국면을 열어나가기로 하였다. "
## [4] "남과 북은 고위급 회담을 비롯한 각 분야의 대화와 협상을 빠른 시일 안에
##      개최하여 정상회담에서 합의된 문제들을 실천하기 위한 적극적인 대책을 세워나가기로
##      하였다. "
## [5] "남과 북은 당국 간 협의를 긴밀히 하고 민간교류와 협력을 원만히 보장하기
##      위하여 쌍방 당국자가 상주하는 남북공동연락사무소를 개성지역에 설치하기로 하였다.
## "
## [6] "남과 북은 민족적 화해와 단합의 분위기를 고조시켜 나가기 위하여
##      각계각층의 다방면적인 협력과 교류 왕래와 접촉을 활성화하기로 하였다. "
```

```
useNIADic()
```

```
## Backup was just finished!
## 983012 words dictionary was built.
```

```
nouns <- sapply(text, extractNoun, USE.NAMES = F)
nouns = unlist(nouns)
```

```
wordcount = table(nouns2)
wc = sort(wordcount, decreasing = T)
wc[1:80]
## nouns2
```

##	하기	민족	비롯	한반도	적극	발전	협력	
##	13	7	7	7	6	5		5
##	개최	공동	남북	노력	대책	문제	합의	
##	4	4	4	4	4	4		4
##	회담	개선	구축	긴장	당면	비핵화	실질적	
##	4	3	3	3	3	3		3
##	자주	추진	평화	평화체제	해결	협의를	각계각층	
##	3	3	3	3	3	3		2
##	계기	고조시	공고	공동번영	관계	교류	남북관계	
##	2	2	2	2	2	2		2
##	단합	당국	보장하	분위기	사회	선언	신뢰	
##	2	2	2	2	2	2		2
##	실천	실현	쌍방	올해	왕래	우리	의의	
##	2	2	2	2	2	2		2
##	이룩	이상	전면	전환	접촉	정전협정	중대	
##	2	2	2	2	2	2		2
##	중지하	충돌	통일	항구적	해소	화해	확인	
##	2	2	2	2	2	2		2
##	활성화	10. 1	차적으로	2018 4	선언에서	4	자회담	6.15
##	2	1	1	1	1	1		1
##	65	8.15	가지	개성	거래	결정	경의선	
##	1	1	1	1	1	1		1
##	고위급	공간	공동행사					
##	1	1	1					

```
nouns2 <- gsub(pattern = "하기", replacement = "", nouns2)
nouns2 <- gsub(pattern = "비롯", replacement = "", nouns2)
nouns2 <- gsub(pattern = "고조시", replacement = "", nouns2)
nouns2 <- gsub(pattern = "가지", replacement = "", nouns2)
nouns2 <- gsub(pattern = "^65", replacement = "", nouns2)
nouns2 <- gsub(pattern = "경우", replacement = "", nouns2)
nouns2 <- gsub(pattern = "언제", replacement = "", nouns2)
nouns2 <- gsub(pattern = "이번", replacement = "", nouns2)
nouns2 <- gsub(pattern = "그제", replacement = "", nouns2)
nouns2 <- gsub(pattern = "번째", replacement = "", nouns2)
nouns2 <- gsub(pattern = "만큼", replacement = "", nouns2)
nouns2 <- gsub(pattern = "10.", replacement = "", nouns2)
nouns2 = nouns2[nchar(nouns2) >= 2]
wordcount = table(nouns2) # 단어 빈도 계산
```

```
random.order=F, family="malgun")
```

