#4/29(금)\_HW2

#Crawling 문제. 대한성서공회의 창세기를 제외한 다른 권을 선택한 후

이를 크롤링하여 막대그래프와 구름모양으로 각데이터의 패턴 상위20개를 시각화하여 나타내시오

#Crawling for Bible

getwd()

library(rvest)

library(stringr)

library(dplyr)

library(ggplot2)

library(wordcloud)

library(wordcloud2)

library(KoNLP)

cnt=c()

url="https://www.bskorea.or.kr/bible/korbibReadpage.php?version=GAE&book=rev&chap="

for (i in 1:22) {

base\_url=paste0(url,i)

t\_css="#tdBible1 span"

hdoc=read\_html(base\_url,encoding = "UTF-8")

n\_css=html\_nodes(hdoc,t\_css)

cnt\_part=html\_text(n\_css)

cnt\_part=gsub("\\d+","",cnt\_part)

cnt\_part#확인용출력

cnt\_part=str\_trim(cnt\_part,side = "both")

cnt=c(cnt,cnt\_part)

}

txt=sapply(cnt, extractNoun,USE.NAMES = F)

txt=unlist(txt)#백터화 >>분석가능

class(txt)

count=Filter(function(x){nchar(x)>=2},txt)

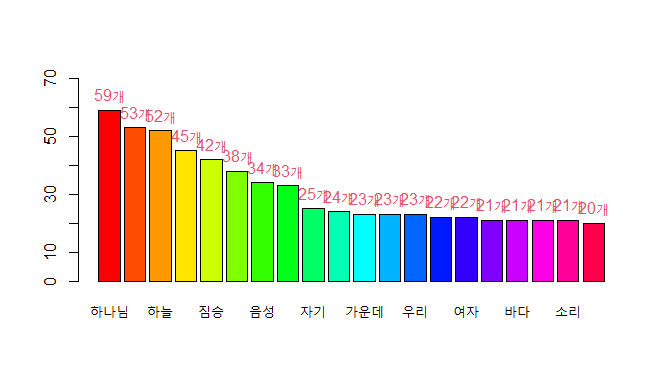
word=table(count)

tt=head(sort(word,decreasing = T),20)

kk=barplot(tt,ylim=c(0,300),col=rainbow(20))

text(kk,tt,paste0(tt,"개"),pos=3,las=2,col=2,cex=2)

plot.new()



library(RColorBrewer)

display.brewer.all()

palate=brewer.pal(12,"Set3")

wordcloud(names(word),

freq = word,

min.freq = 2,

scale=c(5,0.2),

random.order = F,

random.color = T,

colors = palate

)

wordcloud2(data=word,

size=0.4,

shape="cloud")

텍스트이(가) 표시된 사진

자동 생성된 설명