



네이버를 이용한 실시간 검색 Wordcloud



프로그램의 목적

실시간 검색어를 어떻게 하면 정리할 수 있을까 고민하던 중

특정 시간의 실시간 검색어들을 종합하여 각 실시간 검색어들의 빈도수를 체크하고

하루 단위로 각 실시간 검색어가 어떤 빈도로 쓰였는지 워드클라우드를 통해 시각화 하여 보여준다



원리

네이버라는 검색사이트의 주소를 빌려옴(크롤링)
그과정에서 매우 단순히 몇개의 숫자를
변경하는것만으로도 의미있는 결과를 얻을수 있음.
숫자를 변경하는 과정에서 **for** 문과 변수를 활용하였음.



사용법

코드에 년도와 달, 날을 의미하는 변수를
변경하여 자신이 원하는 날짜를 설정하여
코드실행.

이미지 예시





주의점

변수를 입력하는 과정에서

년도는 올해와 작년만이 가능하며(특정 기간만 가능)- 다른 기간은 주소값을 불러올수 없음.

달과 일의 경우 한자리수의 숫자를 표현하는 과정에서 0을 자동적으로 넣어주지 못하고
수동적으로 0과 한자리수의 숫자를 입력해야 할뿐만 아니라 문자형식으로 입력해야 한다.

숫자형식으로 예를 들어 09를 입력했을 경우 9로 인식하므로 제대로된 사이트 값을 표시못함.



문제점

월드 클라우드 실행시 모든 단어가 다 뜨지 않는 경우가 자주 발생하며 이미지를 확대해야 볼수 있는 경우가 많으므로 이미지를 확대해서 봐야함.(size=1)

size=0.6으로 변경한 후에야 이미지를 확대한 경우에 한해 단어들을 볼수 있게됨.

이 경우 작은 단어들은 더 안보이게 됨.



코드

```
library(rvest)
library(dplyr)
library(httr)
library(wordcloud2)
```

```
y<-8
m<-'01'
d<-'19'
url<-c()
for (i in 0:23){
  ifelse(i<10,ur<-paste('https://datalab.naver.com/keyword/realtimeList.naver?datetime=201',y,'-',m,'-',d,'T0',i,':00:00',sep=''),
    ur<-paste('https://datalab.naver.com/keyword/realtimeList.naver?datetime=201',y,'-',m,'-',d,'T',i,':00:00',sep=''))
  url<-c(url,ur)}
```




```
number<-c()
num<-c()
for (i in 1:24){
  u<-content(GET(url[i]))
  nu<-html_nodes(u,css=".section_lst_area .keyword_rank .rank_list li")%>%
  %>%html_nodes('a')%>%html_nodes('.title')%>%html_text()
  num<-nu[81:100]
  number<-c(number,num)}
cn<-table(number)
wordcloud2(cn,size=0.6,shape='rectangle')
```