

## snippet.shell

```
#
install.packages("rJava")
install.packages("memoise")
install.packages("KoNLP")

#
library(KoNLP)
library(dplyr)

useNIADic()

#
setwd('C:/Users/Playdata/weekend_R/ R ')
txt <- readLines('./data/hiphop.txt')
head(txt)

install.packages('stringr')
library(stringr)

txt <- str_replace_all(txt, '\\W', ' ')

extractNoun(' ')

R.Version()
R.version
R.version.string
version

# 가
nouns <- extractNoun(txt)
# list
wordcount <- table(unlist(nouns))

#
df_word <- as.data.frame(wordcount, stringasFactors = F)
head(df_word)

#
df_word <- rename(df_word, word = Var1, freq = Freq)
head(df_word)

# ?nchar

#
df_word <- filter(df_word, nchar(word) >= 2)
#as.character(df$coll)
df_word <- filter(df_word, nchar(as.character(df_word$word)) >= 2)
head(df_word)

top_20 <- df_word %>%
```

```
    arrange(desc(freq)) %>%
    head(20)
top_20

#
install.packages('wordcloud')

#
library(wordcloud)
library(RColorBrewer)

# Dark2          8
pal <- brewer.pal(8, 'Dark2')
set.seed(1234)

wordcloud(words = df_word$word,
          freq = df_word$freq,
          min.freq = 2 ,
          max.words = 200,
          random.order = F,
          rot.per = .1,
          scale = c(4, 0.3),
          colors = pal          )

pal <- brewer.pal(9, 'Blues')[5:9]
set.seed(1234)

wordcloud(words = df_word$word,
          freq = df_word$freq,
          min.freq = 2 ,
          max.words = 200,
          random.order = F,
          rot.per = .1,
          scale = c(4, 0.3),
          colors = pal)

# 10-2

#
twitter <- read.csv('./data/twitter.csv',
                  header = T,
                  stringsAsFactors = F,
                  fileEncoding = 'UTF-8')

head(twitter)

#
twitter <- rename(twitter,
                 no =      ,
                 id =      ,
                 date =    ,
                 tw =      )
```

```
head(twitter)

#
twitter$tw <- str_replace_all(twitter$tw, '\\W', ' ')
head(twitter$tw)

#
#
nouns <- extractNoun(twitter$tw)
#           list
wordcount <- table(unlist(nouns))
#
df_word <- as.data.frame(wordcount, stringsAsFactors = F)
#
df_word <- rename(df_word,
                  word = Var1,
                  freq = Freq)
#
df_word <- filter(df_word, nchar(df_word$word) >= 2)
View(df_word)

top20 <- df_word %>%
  arrange(desc(freq)) %>%
  head(20)

top20

library(ggplot2)

#
order <- arrange(top20, freq)$word

ggplot(data=top20, aes(x=word, y=freq)) +
  ylim(0, 2500) +
  geom_col() +
  coord_flip() +
  scale_x_discrete(limit = order) +          #
  geom_text(aes(label=freq), hjust=-0.3)     #

pal <- brewer.pal(8, 'Dark2')
set.seed(1234)

wordcloud(words = df_word$word,
          freq=df_word$freq,
          min.freq = 10,
          max.words = 200,
          random.order = F,
          rot.per = .1,
          scale = c(6, 0.2),
          colors = pal)

pal <- brewer.pal(9, 'Blues')[5:9]
```

```
set.seed(1234)

wordcloud(words = df_word$word,
          freq=df_word$freq,
          min.freq = 10,
          max.words = 200,
          random.order = F,
          rot.per = .1,
          scale = c(6, 0.2),
          colors = pal)
```

---

## Plugin Backlinks:

From:  
<https://moro.kr/> - **Various Ways**

Permanent link:  
<https://moro.kr/open/%ED%85%8D%EC%8A%A4%ED%8A%B8-%EB%A7%88%EC%9D%B4%EB%8B%9D>

Last update: **2020/06/02 09:25**

