```
library(devtools)
library("jiebaR")
library(tm)
library(tmcn)
library('wordcloud2')

data <- read.csv("women_clothes.csv")

#將資料分成兩組
Recommend <- data[data$Recommended.IND == 1,]
Unrecommend <- data[data$Recommended.IND == 0,]

#擷取心得欄
Recommend_words <- Recommend$Review.Text
Unrecommend_words <- Unrecommend$Review.Text


x <- VectorSource(Recommend_words)
x <- VCorpus(x)

myStopWords <- c(stopwords()) #remove some words
x <- tm_map(x, removeWords, myStopWords)
head(myStopWords)

tdm <- TermDocumentMatrix(x, control =list(wordLengths = c(2, Inf)))

m1 <- as.matrix(tdm) #轉Matrix
v <- sort(rowSums(m1), decreasing = TRUE)
d <- data.frame(word = names(v), freq = v) #count freq
new_d <- d[d$freq > 500,]
head(new_d)

wordcloud2(new_d,size=0.5)

extract_d <- d[d$freq > 2000,]
extract_d %>%
  filter(freq > 6) %>%
  mutate(word = reorder(word, freq)) %>%
  ggplot(aes(word,freq))+
  theme(text=element_text(family="微軟正黑體", size=14))+
  geom_col() +
  xlab(NULL) +
  coord_flip()



y <- VectorSource(Unrecommend_words)
y <- VCorpus(y)

myStopWords <- c(stopwords()) #remove some words
y <- tm_map(y, removeWords, myStopWords)
head(myStopWords)

tdm2 <- TermDocumentMatrix(y, control =list(wordLengths = c(2, Inf)))

m2 <- as.matrix(tdm2) #轉Matrix
v2 <- sort(rowSums(m2), decreasing = TRUE)
```

```
d2 <- data.frame(word = names(v2), freq = v2) #count freq
new_d2 <- d2[d2$freq > 200,]
head(new_d2)

wordcloud2(new_d2,size=0.5)

extract_d2 <- d2[d2$freq > 500,]
extract_d2 %>%
  filter(freq > 6) %>%
  mutate(word = reorder(word, freq)) %>%
  ggplot(aes(word,freq))+
  theme(text=element_text(family="微軟正黑體", size=14))+
  geom_col() +
  xlab(NULL) +
  coord_flip()
```