Assignment Week 7 Option 1

Jihun Bae

Install "friends" package (install.packages("friends")). Use "friends" package and write codes for following tasks:

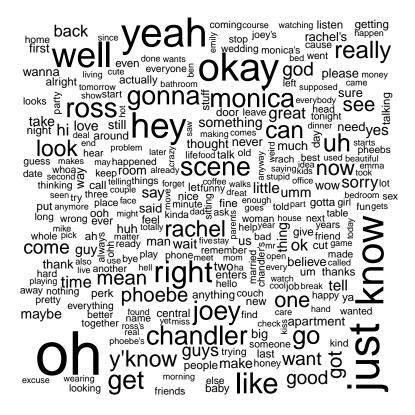
1. Tokenize the text and remove stop-words.

```
## Joining, by = "word"
## # A tibble: 355,914 x 6
      speaker
                    season episode scene utterance word
##
      <chr>
                     <int>
                             <int> <int>
                                              <int> <chr>
   1 Monica Geller
                       1
                                       1
                                                  1 nothing
   2 Monica Geller
                                 1
                                                  1 tell
                         1
                                       1
   3 Monica Geller
                         1
                                 1
                                       1
                                                  1 just
## 4 Monica Geller
                         1
                                        1
                                                  1 guy
## 5 Monica Geller
                         1
                                 1
                                       1
                                                  1 work
## 6 Joey Tribbiani
                         1
                                 1
                                       1
                                                  2 c'mon
   7 Joey Tribbiani
                         1
                                 1
                                       1
                                                  2 going
## 8 Joey Tribbiani
                         1
                                 1
                                      1
                                                  2 guy
## 9 Joey Tribbiani
                                 1
                         1
                                       1
                                                  2 gotta
## 10 Joey Tribbiani
                                       1
                                                 2 something
```

2. Calculate word frequency.

... with 355,904 more rows

```
## # A tibble: 17,203 x 2
      word
##
      <chr> <int>
   1 oh
             8181
##
   2 just
             5735
##
   3 yeah
             5051
   4 okay
             4926
             4873
  5 know
             4229
##
  6 well
   7 hey
             4219
  8 right
           3569
## 9 like
             3238
## 10 ross
             3196
## # ... with 17,193 more rows
```



- 3. Make a wordcloud.
- 4. Count the frequency of those positive and negative words. Show the most frequent positive and negative words, respectively, using head() function.

```
## Joining, by = "word"
##
   # A tibble: 10 x 2
##
      word
                 n
##
       <chr> <int>
##
    1 well
              4229
##
    2 right
              3569
##
    3 like
              3238
##
      good
              1756
##
    5
      great
              1353
    6 love
##
              1040
##
               756
    7 thank
##
    8
               750
##
    9 fine
               648
   10 work
               620
## Joining, by = "word"
##
   # A tibble: 10 x 2
##
      word
                    n
##
       <chr>
               <int>
##
                1460
    1 sorry
##
    2 bad
                 414
                 325
##
    3 wrong
```

```
4 funny
                 314
##
                 298
##
    5 hard
                 276
##
    6 weird
    7 stupid
                 258
##
##
    8 crazy
                 255
##
    9 excuse
                 236
## 10 problem
```

5. Visualize the frequency of positive and negative words by each character.

Joining, by = "word"

jative-Positive Ratio in all seasons by using bing

