

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       1 nothing
## 2 Monica Geller      1       1     1       1 tell
## 3 Monica Geller      1       1     1       1 just
## 4 Monica Geller      1       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       1     1       2 something
## # ... with 355,904 more rows
```

2. Calculate word frequency.

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



```
## 4 funny      314
## 5 hard       298
## 6 weird      276
## 7 stupid     258
## 8 crazy      255
## 9 excuse     236
## 10 problem   225
```

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

```
## Joining, by = "word"
```

Positive-Negative Ratio in all seasons by using bing

