Text Mining

Environment

```
source("code.R")
```

```
## -- Attaching packages ---------------------------------------------------------------------------
## v ggplot2 3.3.0     v purrr   0.3.4
## v tibble  3.0.1     v dplyr   0.8.5
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
## -- Conflicts ------------------------------------------------------------------------------------ ti
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Filtering stop words

Top 20 unique words in the text, less stop words

```
## # A tibble: 20 x 3
##    word            n percent
##    <chr>       <int>   <dbl>
##  1 workers        39       2
##  2 singapore      35       2
##  3 people         31       2
##  4 brother        27       1
##  5 day            25       1
##  6 government     24       1
##  7 leadership     22       1
##  8 team           21       1
##  9 it's           19       1
## 10 unions         19       1
## 11 lee            18       1
## 12 exceptional    16       1
## 13 swee           14       1
## 14 country        13       1
## 15 don't          13       1
## 16 leaders        12       1
## 17 minister       12       1
## 18 ntuc           12       1
## 19 jobs           11       1
## 20 job            10       0
```

Sentiments

we use NRC to get our sentiments. A description of NRC, taken from NRC website: The NRC Emotion Lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation,

trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive). The annotations were manually done by crowdsourcing.

For more info about National Research Council Canada (NRC) Emotion Lexicon, please visit: http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm All the sentiments in the text

```
## # A tibble: 10 x 3
##    sentiment        n percent
##    <chr>        <int>   <dbl>
##  1 positive       294      27
##  2 trust          209      19
##  3 anticipation   119      11
##  4 negative       119      11
##  5 joy            107      10
##  6 fear            87       8
##  7 anger           59       5
##  8 sadness         51       5
##  9 surprise        37       3
## 10 disgust         25       2
```