# iemisctext: Example Analysis using the tm package

### Irucka Embry

#### 2024-05-31

#### Contents

Examples	1
EcoC <sup>2</sup> S Links	6
Copyright and License	6

## Examples

```
install.load::load_package("iemisctext", "tm", "data.table", "ggwordcloud")
# load needed packages using the load_package function from the install.load
# package (it is assumed that you have already installed these packages)
# appeal_womanhood_world
data(appeal_womanhood_world)
# create a document term matrix of appeal_womanhood_world
aww <- DocumentTermMatrix(appeal_womanhood_world, control = list(removePunctuation = TRUE,</pre>
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc")))
aww
## <<DocumentTermMatrix (documents: 1, terms: 170)>>
## Non-/sparse entries: 0/170
## Sparsity
                     : 100%
## Maximal term length: 13
## Weighting
                      : SMART ntc (SMART)
# get the word frequency table
aww_tf <- as.data.table(termFreq(appeal_womanhood_world$content[[1]]$content, control = list(removePunc</pre>
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)
# create a word cloud
ggwordcloud(aww_tf$V1, aww_tf$V2)
```

# women

```
sons questions
# anarchy
data(anarchy)
# create a document term matrix of anarchy
ancy <- DocumentTermMatrix(anarchy, control = list(removePunctuation = TRUE, stopwords = TRUE,</pre>
    weighting = function(x) weightSMART(x, spec = "ntc")))
ancy
## <<DocumentTermMatrix (documents: 1, terms: 785)>>
## Non-/sparse entries: 0/785
## Sparsity
## Maximal term length: 14
                       : SMART ntc (SMART)
## Weighting
# get the word frequency table
ancy_tf <- as.data.table(termFreq(anarchy$content[[1]]$content, control = list(removePunctuation = TRUE</pre>
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)
# create a word cloud
ggwordcloud(ancy_tf$V1, ancy_tf$V2)
                                  stand
                            horses
                             see
                                  earth
    food
                                                    weak heard
                                  WIII till
       make
                                                                    following
                                              pale
                                                    passed
                                         lay
                                 rest rode
                                                                tis
       english
                                          find
                                             multitude
                                                                     laws
                                    le gold around
                     knew
                          great
well away
                      name
                            every
          death
                                    anarchy arms
must dew
                   mighty
          sea
                          nogu
                                                            thee
                                        england thy
      children
                    came
                              trampling
                                      feel
                                           toil
                                              god
          tyrants
                      home looked
                                      even
# war_prayer
data(war_prayer)
# create a document term matrix of war_prayer
wp <- DocumentTermMatrix(war_prayer, control = list(removePunctuation = TRUE, stopwords = TRUE,
    weighting = function(x) weightSMART(x, spec = "ntc")))
wp
## <<DocumentTermMatrix (documents: 1, terms: 523)>>
```

```
## Non-/sparse entries: 0/523
## Sparsity
                      : 100%
## Maximal term length: 13
                      : SMART ntc (SMART)
## Weighting
# get the word frequency table
wp_tf <- as.data.table(termFreq(war_prayer$content[[1]]$content, control = list(removePunctuation = TRU
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)
# create a word cloud
ggwordcloud(wp_tf$V1, wp_tf$V2)
                   spirit stather victory
                         uttered aid also pale
part
                        young
                                time heard
                                       hearts
# war racket
data(war_racket)
# create a document term matrix of war_racket
wr <- DocumentTermMatrix(war_racket, control = list(removePunctuation = TRUE, stopwords = TRUE,
    weighting = function(x) weightSMART(x, spec = "ntc")))
wr
## <<DocumentTermMatrix (documents: 1, terms: 1475)>>
## Non-/sparse entries: 0/1475
## Sparsity
                      : 100%
## Maximal term length: 15
## Weighting
                     : SMART ntc (SMART)
# get the word frequency table
wr_tf <- as.data.table(termFreq(war_racket$content[[1]]$content, control = list(removePunctuation = TRU</pre>
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)
# create a word cloud
```

ggwordcloud(wr\_tf\$V1, wr\_tf\$V2)

```
veterans
                                                                                    huge days 200
              wartime
                                                                                                     nettina
                       mosquito
                                                trained
                                                                    suddenly
company
                                                            every
                                                                               bonds yearly
                                       One miles spent
                                                                                                   japanese
                         promptly
                                                                    somebody
              period
                                                              dont
                                                                                         shipbuilders sailors
                                        total
                                                                             sold
                      philadel
                            along
                                                  shoulder
                                                                  nation
                                                                            arms
                                                                                              nothing conscript
                                      plebiscite
                                                others
                                                          say
                          changed
                                                                                        lets
                                                                  families
 apparently
                                   democracy
                                                                             knows
                                                       life
                                                                                               racket
                                             perhaps
                         mentally
                 fought
                                                                300 end
                                                                                       enemy
                                                                          cost
                mean let
                                                 broken
                                                                                 naval
                                                                                        shoes
                                          stand
                          wav
                                                                  territory
                                                                                               sum
                                                                                                    united
                                                                           arent
                                                                                  pacific
                  uncle 19141918
                                                collected
                                                                                       prize
                                      whether
                                                                                                country
 dollars
                                                             without committee
                                                                                 worth
                                                                                               forces trade
                                        defense
                              instance
                                                                   shunted
               course purposes
                                                            bargain
                                                                            matter fleet
                                          left
                                               losses
                                                                                               maneuvers 100
                                                       tremendous
                                                                    paying
                                five
             shouldnt
                away
                      germany
                                        lot
                                                                     earnings bills
                                                                                               back
                                            hundreds
called
                                                                                                    men die
      buv
                                                      soldier
                                wage
                                         told
                         least
eighteen
                 upon
                                              per
                                                                    medals power
                                                    times
                                                                                             get
                              month
                                                                                  hell
                                                          else
        world
                                                                                     came
                        able
                  hospital
                                                                cent
                                             outside
                                                                               capital government
                                yes
                                        put
                          lives
                                                     president
        airplane
                                                                 shot
                                                                          god
                see
                                              showed
                                 time
                                         army
                                                                                           rather something
                                                       congress
                                                                                always
                                                                  normal taken
           chapter
                                               children
                               powder
                                   19101914
                       wounded
                                                         coast
                                                                                          money loss
                                                                    munitions
                side
                                   average
                                              business
                        uniform
    japan
                                                        italy
                                                                        just risk
                                                                three g
                    trenches
                                                                                                    medal
              navy
                                                  fight
wrenches
                                                                                    blood
                                                                                            jumped
                                                        greater
                               smash
                                                                             must
                                                                ever
                                                                                                    built
                                                  bill
            even billions
    industry
                                liberty
                                                           averaged
                                                                                           take years
                                                                                    battle
                                                                      speculators
                                        two
        <sub>destroyed</sub> bought
                                                    china
                                                             never
                                                                                         young wars
                                                                        general
                            international
                                          sam
                                                     believe
                                                           steel
                                                                                         minds killed
                     group
      american
                                                                                                         use
                                                                  thousands
                                           probably
                                                                             shall
                                 make
                                                                                         safe
                                                                                                      states
     bad
                                                          couldnt
                                                                          hungry
                   made
                                           russia
                                                   workers
                                                                  much
                                great
                                                                                              4000000
                                                another ship
                                                                           voting
                                                                40000000
                                                                                                       6000000
    jugoslavia civil
                                                                                        good
                                        still
# connect_dots
data(connect_dots)
# create a document term matrix of connect_dots
cds <- DocumentTermMatrix(connect_dots, control = list(removePunctuation = TRUE,</pre>
     stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc")))
cds
## <<DocumentTermMatrix (documents: 1, terms: 337)>>
## Non-/sparse entries: 0/337
## Sparsity
## Maximal term length: 19
## Weighting
                            : SMART ntc (SMART)
# get the word frequency table
cds_tf <- as.data.table(termFreq(connect_dots$content[[1]]$content, control = list(removePunctuation =</pre>
     stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)
# create a word cloud
ggwordcloud(cds_tf$V1, cds_tf$V2)
```

went

germans

# petroleum

```
industrial
                       created earth
 automobile
          hemp
                   ignorance
  people
                            corporations
                  states water companies 5
                     marihuana engine
automobiles
          ^{\text{tax}} fear _{\text{line}}
                               diesel
# us them
data(us_them)
# create a document term matrix of us them
ut <- DocumentTermMatrix(us_them, control = list(removePunctuation = TRUE, stopwords = TRUE,
    weighting = function(x) weightSMART(x, spec = "ntc")))
ut
## <<DocumentTermMatrix (documents: 1, terms: 247)>>
## Non-/sparse entries: 0/247
## Sparsity
                       : 100%
## Maximal term length: 71
## Weighting
                       : SMART ntc (SMART)
# get the word frequency table
ut_tf <- as.data.table(termFreq(us_them$content[[1]])$content, control = list(removePunctuation = TRUE,
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)
# create a word cloud
ggwordcloud(ut_tf$V1, ut_tf$V2)
                                  freedom
                           world
                    well people
                                    states
                        constitution
# climate_strange
data(climate strange)
# create a document term matrix of climate_strange
cs <- DocumentTermMatrix(climate_strange, control = list(removePunctuation = TRUE,</pre>
    stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc")))
CS
## <<DocumentTermMatrix (documents: 1, terms: 525)>>
## Non-/sparse entries: 0/525
## Sparsity
                      : 100%
## Maximal term length: 24
```

```
## Weighting : SMART ntc (SMART)

# get the word frequency table

cs_tf <- as.data.table(termFreq(climate_strange$content[[1]]$content, control = list(removePunctuation stopwords = TRUE, weighting = function(x) weightSMART(x, spec = "ntc"))), keep.rownames = TRUE)

# create a word cloud
ggwordcloud(cs_tf$V1, cs_tf$V2)

Can \( \varphi\) away \( \varphi\) yeighting food odn to consciousness dont ove act energy way global others

| Can \( \varphi\) away \( \varphi\) global others

| Can \( \varphi\) away \( \varphi\) global others
```

### EcoC<sup>2</sup>S Links

```
EcoC<sup>2</sup>S Home - https://www.ecoccs.com/
About EcoC<sup>2</sup>S - https://www.ecoccs.com/about-ecoc<sup>2</sup>s.html
Services - https://www.ecoccs.com/services.html
1 Stop Shop - https://www.ecoccs.com/other-biz.html
Products - https://www.questionuniverse.com/products.html
Media - https://www.ecoccs.com/media.html
Resources - https://www.ecoccs.com/resources.html
R Trainings and Resources provided by EcoC<sup>2</sup>S (Irucka Embry, E.I.T.) - https://www.ecoccs.com/rtraining.html
```

# Copyright and License

All R code written by Irucka Embry is distributed under the GPL-3 (or later) license, see the GNU General Public License {GPL} page.

All written content originally created by Irucka Embry is copyrighted under the Creative Commons Attribution-ShareAlike 4.0 International license. All other written content retains the copyright of the original author(s).

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International license.