John Rene Lorelli

DataScience Foundations

Week 12

NLP

1. The entire R code used when creating this bar plot.

> library(readtext)

Warning message:

package ‘readtext’ was built under R version 3.6.3

> stopword <- readLines("stopwords-big")

> wlove<-readLines("roget-love")

> wfear<-readLines("roget-fear")

> wlove<-paste(wlove, collapse = " ")

> wlove<-strsplit(tolower(wlove),"[^a-z]") [(1)]

> wlove<-wlove[nchar(wlove)>3]

> wlove<-unique(wlove)

> wfear<-paste(wfear, collapse = " ")

> wfear<-strsplit(tolower(wfear),"[^a-z]") [(1)]

> wfear<-wfear[nchar(wfear)>3]

> wfear<-unique(wfear)

> swd<-data.frame(character(), integer(), integer(), stringsAsFactors = FALSE)

> aus<-readLines("C-evilgenius.txt")

> aus<-paste(aus, collapse =" " )

> aus<-strsplit(tolower(aus), "[^a-z]") [(1)]

> aus<-aus[nchar(aus)>5]

> aus<-aus[!is.element(aus, stopword)]

> leng<-length(aus)

> auswds<-as.data.frame(table(aus))

> nn<-0

> for(aw in wlove){ cn<-auswds[auswds$aus==aw, 2]; if(length(cn)>0){nn<nn+cn;}}

Warning messages:

1: In `==.default`(auswds$aus, aw) :

longer object length is not a multiple of shorter object length

2: In is.na(e1) | is.na(e2) :

longer object length is not a multiple of shorter object length

> fn<-0

> for(aw in wfear){ cn<-auswds[auswds$aus==aw, 2]; if(length(cn)>0){fn<-fn+cn;}}

Warning messages:

1: In `==.default`(auswds$aus, aw) :

longer object length is not a multiple of shorter object length

2: In is.na(e1) | is.na(e2) :

longer object length is not a multiple of shorter object length

> for(aw in wlove){ cn<-auswds[auswds$aus==aw, 2]; if(length(cn)>0){nn<-nn+cn;}}

Warning messages:

1: In `==.default`(auswds$aus, aw) :

longer object length is not a multiple of shorter object length

2: In is.na(e1) | is.na(e2) :

longer object length is not a multiple of shorter object length

> nn

[1] 27286

> fn

[1] 27286

> nn

[1] 27286

> fn

[1] 27286

> sen<-c(nn,fn)

>barplot(sen, names.arg = c("Love", "Fear"), main=" Sentiment Analysis of Evil Genius")

1. Screenshot of the bar plot created.

