

## M12\_Exercise\_NLP\_Abhishek\_Jain

### R code

```
library(readtext)
# Reading the words from data
stopwords <- readLines("stopwords-big") # Words to ignore
loveWords <- readLines("roget-love") # Words about love
fearWords <- readLines("roget-fear") # Words about fear
evilText <- readLines("C-evilgenius.txt") # Text of Evil Genius by Wilkie Collins
# Counting the number of love words and fear words
numLoveWords <- 0
numFearWords <- 0
# Count the number of unique words about love longer than 5 characters
loveWords <- paste(loveWords, collapse=" ")
loveWords <- strsplit(tolower(loveWords), "[^a-z]") [[1]]
loveWords <- loveWords[nchar(loveWords) > 5]
loveWords <- unique(loveWords)
# Count the number of unique words about fear longer than 5 characters
fearWords <- paste(fearWords, collapse=" ")
fearWords <- strsplit(tolower(fearWords), "[^a-z]") [[1]]
fearWords <- fearWords[nchar(fearWords) > 5 ]
fearWords <- unique(fearWords)
# Count the number of unique words in Evil Genius longer than 5 characters and
# not in stopwords
evilText <- paste(evilText, collapse = " ")
evilText <- strsplit(tolower(evilText), "[^a-z]") [[1]]
evilText <- evilText[nchar(evilText) > 5 ]
evilText <- evilText[!is.element(evilText, stopwords)]
evilTextDF <- as.data.frame(table(evilText))

for (word in loveWords) {
  currentWord <- evilTextDF[evilTextDF$evilText == word, 2]
  if (length(currentWord)>0) {
    numLoveWords <- numLoveWords + currentWord
  }
}
for (word in fearWords) {
  currentWord <- evilTextDF[evilTextDF$evilText == word, 2]
  if (length(currentWord) > 0) {
    numFearWords <- numFearWords + currentWord
  }
}

cat(numLoveWords, "&", numFearWords)

#code for the barplot
sentimentAnalysis <- c(numLoveWords, numFearWords)
```

```
sentimentAnalysisPlot <- barplot(sentimentAnalysis, names.arg=c("Love", "Fear"),
                                main="Sentiment Analysis of 'Evil Genius' by Wilkie
Colkins",
                                ylab="Number of Words related", ylim = c(0, 500),
                                col = rainbow(2))
text(x = sentimentAnalysisPlot, y = sentimentAnalysis + 20,
     labels = sentimentAnalysis, col = "red")
```

## Screenshot

