```
> cname <- file.path("~", "Desktop", "texts")
> cname
[1] "~/Desktop/texts"

> dir(cname)
[1] "sample.txt"


> install.packages("tm")
> library(tm)
> docs <- VCorpus(DirSource(cname))
> docs <- tm_map(docs,removePunctuation)
> docs <- tm_map(docs, removeNumbers)
> docs <- tm_map(docs, tolower)
> docs <- tm_map(docs, removeWords, stopwords("english"))
> docs <- tm_map(docs, removeWords, c("syllogism", "tautology"))
> docs <- tm_map(docs, stripWhitespace)
> docs <- tm_map(docs, PlainTextDocument)
> dtm <- DocumentTermMatrix(docs)
> tdm <- TermDocumentMatrix(docs)
> freq <- colSums(as.matrix(dtm))

> length(freq)
[1] 29
> ord <- order(freq)
> m <- as.matrix(dtm)
> dim(m)
[1]  1 29
> write.csv(m, file="DocumentTermMatrix.csv")
> dtms <- removeSparseTerms(dtm, 0.1)
> head(table(freq), 20)
freq
 1  2  4
25  3  1
> tail(table(freq), 20)
freq
 1  2  4
25  3  1
> freq <- colSums(as.matrix(dtms))
> freq
      also      cabin    current       dear       dept  direction     issues
         1          2          1          1          1          1          1
   machine maintenance        one        per  permanent permanently  permission
         4          1          2          1          1          1          1
      pict    problem   properly       pune     rajesh        rao  repairing
         1          1          1          1          1          1          1
       sir       smps      staff    subject       told   transfer transferred
         1          1          2          1          1          1          1
   working
         1

> install.packages("ggplot2")
```

```
> library(ggplot2)

Attaching package: ggplot2

The following object is masked from package:NLP:

    annotate

> wf <- data.frame(word=names(freq), freq=freq)
> p <- ggplot(subset(wf, freq>50), aes(word, freq))
> p <- p + geom_bar(stat="identity")
> p <- p + theme(axis.text.x=element_text(angle=45, hjust=1))
> p
> findAssocs(dtm, c("country" , "american"), corlimit=0.85)
$country
numeric(0)

$american
numeric(0)

> findAssocs(dtms, "think", corlimit=0.70)
$think
numeric(0)


> install.packages("wordcloud")
> library(wordcloud)
Loading required package: RColorBrewer
> dtms <- removeSparseTerms(dtm, 0.15) # Prepare the data (max 15% empty space)
> freq <- colSums(as.matrix(dtm))
> dark2 <- brewer.pal(6, "Dark2")
> wordcloud(names(freq), freq, max.words=100, rot.per=0.2, colors=dark2)
```