```
> # Install
> install.packages("tm")  # for text mining
Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
also installing the dependencies 'NLP', 'Rcpp', 'slam', 'xml2', 'BH'

trying URL 'http://package-proxy/src/contrib/NLP_0.2-0.tar.gz'
Content type 'application/x-tar' length 392231 bytes (383 KB)
==================================================
downloaded 383 KB

trying URL 'http://package-proxy/src/contrib/Rcpp_1.0.4.6.tar.gz'
Content type 'application/x-tar' length 3156751 bytes (3.0 MB)
==================================================
downloaded 3.0 MB

trying URL 'http://package-proxy/src/contrib/slam_0.1-47.tar.gz'
Content type 'application/x-tar' length 214787 bytes (209 KB)
==================================================
downloaded 209 KB

trying URL 'http://package-proxy/src/contrib/xml2_1.3.1.tar.gz'
Content type 'application/x-tar' length 509216 bytes (497 KB)
==================================================
downloaded 497 KB

trying URL 'http://package-proxy/src/contrib/BH_1.72.0-3.tar.gz'
Content type 'application/x-tar' length 11547022 bytes (11.0 MB)
==================================================
downloaded 11.0 MB

trying URL 'http://package-proxy/src/contrib/tm_0.7-7.tar.gz'
Content type 'application/x-tar' length 1288587 bytes (1.2 MB)
==================================================
downloaded 1.2 MB

* installing *binary* package 'NLP' ...
* DONE (NLP)
* installing *binary* package 'Rcpp' ...
* DONE (Rcpp)
* installing *binary* package 'slam' ...
* DONE (slam)
* installing *binary* package 'xml2' ...
* DONE (xml2)
* installing *binary* package 'BH' ...
* DONE (BH)
* installing *binary* package 'tm' ...
* DONE (tm)

The downloaded source packages are in
        '/tmp/RtmpR9rqON/downloaded_packages'
> install.packages("SnowballC") # for text stemming
```

```
Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
trying URL 'http://package-proxy/src/contrib/SnowballC_0.7.0.tar.gz'
Content type 'application/x-tar' length 606692 bytes (592 KB)
==================================================
downloaded 592 KB

* installing *binary* package 'SnowballC' ...
* DONE (SnowballC)

The downloaded source packages are in
        '/tmp/RtmpR9rqON/downloaded_packages'
> install.packages("wordcloud") # word-cloud generator
Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
also installing the dependency 'RColorBrewer'

trying URL 'http://package-proxy/src/contrib/RColorBrewer_1.1-2.tar.gz'
Content type 'application/x-tar' length 52941 bytes (51 KB)
==================================================
downloaded 51 KB

trying URL 'http://package-proxy/src/contrib/wordcloud_2.6.tar.gz'
Content type 'application/x-tar' length 366509 bytes (357 KB)
==================================================
downloaded 357 KB

* installing *binary* package 'RColorBrewer' ...
* DONE (RColorBrewer)
* installing *binary* package 'wordcloud' ...
* DONE (wordcloud)

The downloaded source packages are in
        '/tmp/RtmpR9rqON/downloaded_packages'
> install.packages("RColorBrewer") # color palettes
Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
trying URL 'http://package-proxy/src/contrib/RColorBrewer_1.1-2.tar.gz'
Content type 'application/x-tar' length 52941 bytes (51 KB)
==================================================
downloaded 51 KB

* installing *binary* package 'RColorBrewer' ...
* DONE (RColorBrewer)

The downloaded source packages are in
        '/tmp/RtmpR9rqON/downloaded_packages'
> # Load
> library("tm")
Loading required package: NLP
> library("SnowballC")
> library("wordcloud")
```

```
Loading required package: RColorBrewer
> library("RColorBrewer")
>
> filePath <- "http://www.sthda.com/sthda/RDoc/example-files/martin-luther-king-i
-have-a-dream-speech.txt"
> text <- readLines(filePath)
>
> docs <- Corpus(VectorSource(text))
> inspect(docs)
<<SimpleCorpus>>
Metadata:  corpus specific: 1, document level (indexed): 0
Content:  documents: 46

 [1]



 [2] And so even though we face the difficulties of today and tomorrow, I still h
ave a dream. It is a dream deeply rooted in the American dream.


 [3]




 [4] I have a dream that one day this nation will rise up and live out the true m
eaning of its creed:



 [5]




 [6] We hold these truths to be self-evident, that all men are created equal.



 [7]




 [8] I have a dream that one day on the red hills of Georgia, the sons of former
slaves and the sons of former slave owners will be able to sit down together at t
he table of brotherhood.
```

[9]

[10] I have a dream that one day even the state of Mississippi, a state sweltering with the heat of injustice, sweltering with the heat of oppression, will be transformed into an oasis of freedom and justice.

[11]

[12] I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character.

[13]

[14] I have a dream today!

[15]

[16] I have a dream that one day, down in Alabama, with its vicious racists, with its governor having his lips dripping with the words of interposition and nullification, one day right there in Alabama little black boys and black girls will be able to join hands with little white boys and white girls as sisters and brothers.
[17]

[18] I have a dream today!

[19]

[20] I have a dream that one day every valley shall be exalted, and every hill an
d mountain shall be made low, the rough places will be made plain, and the crooke
d places will be made straight; and the glory of the Lord shall be revealed and a
ll flesh shall see it together.

[21]

[22] This is our hope, and this is the faith that I go back to the South with.

[23]

[24] With this faith, we will be able to hew out of the mountain of despair a sto
ne of hope. With this faith, we will be able to transform the jangling discords o
f our nation into a beautiful symphony of brotherhood. With this faith, we will b
e able to work together, to pray together, to struggle together, to go to jail to
gether, to stand up for freedom together, knowing that we will be free one day.
[25]

[26] And this will be the day, this will be the day when all of God s children wi
ll be able to sing with new meaning:

[27]

[28] My country tis of thee, sweet land of liberty, of thee I sing.

[29] Land where my fathers died, land of the Pilgrim s pride,

[30] From every mountainside, let freedom ring!

[31] And if America is to be a great nation, this must become true.

[32] And so let freedom ring from the prodigious hilltops of New Hampshire.

[33] Let freedom ring from the mighty mountains of New York.

[34] Let freedom ring from the heightening Alleghenies of Pennsylvania.

[35] Let freedom ring from the snow-capped Rockies of Colorado.

[36] Let freedom ring from the curvaceous slopes of California.

[37]

[38] But not only that:

[39] Let freedom ring from Stone Mountain of Georgia.

[40] Let freedom ring from Lookout Mountain of Tennessee.

[41] Let freedom ring from every hill and molehill of Mississippi.

[42] From every mountainside, let freedom ring.

[43] And when this happens, when we allow freedom ring, when we let it ring from every village and every hamlet, from every state and every city, we will be able to speed up that day when all of God s children, black men and white men, Jews and Gentiles, Protestants and Catholics, will be able to join hands and sing in the words of the old Negro spiritual:
[44] Free at last! Free at last!

[45]

[46] Thank God Almighty, we are free at last!


```
>
> toSpace <- content_transformer(function (x , pattern ) gsub(pattern, " ", x))
> docs <- tm_map(docs, toSpace, "/")
Warning message:
In tm_map.SimpleCorpus(docs, toSpace, "/") : transformation drops documents
> docs <- tm_map(docs, toSpace, "@")
Warning message:
In tm_map.SimpleCorpus(docs, toSpace, "@") : transformation drops documents
> docs <- tm_map(docs, toSpace, "\\|")
Warning message:
In tm_map.SimpleCorpus(docs, toSpace, "\\|") :
  transformation drops documents
>
> # Convert the text to lower case
> docs <- tm_map(docs, content_transformer(tolower))
```

```
Warning message:
In tm_map.SimpleCorpus(docs, content_transformer(tolower)) :
  transformation drops documents
> # Remove numbers
> docs <- tm_map(docs, removeNumbers)
Warning message:
In tm_map.SimpleCorpus(docs, removeNumbers) :
  transformation drops documents
> # Remove english common stopwords
> docs <- tm_map(docs, removeWords, stopwords("english"))
Warning message:
In tm_map.SimpleCorpus(docs, removeWords, stopwords("english")) :
  transformation drops documents
> # Remove your own stop word
> # specify your stopwords as a character vector
> docs <- tm_map(docs, removeWords, c("blabla1", "blabla2"))
Warning message:
In tm_map.SimpleCorpus(docs, removeWords, c("blabla1", "blabla2")) :
  transformation drops documents
> # Remove punctuations
> docs <- tm_map(docs, removePunctuation)
Warning message:
In tm_map.SimpleCorpus(docs, removePunctuation) :
  transformation drops documents
> # Eliminate extra white spaces
> docs <- tm_map(docs, stripWhitespace)
Warning message:
In tm_map.SimpleCorpus(docs, stripWhitespace) :
  transformation drops documents
> # Text stemming
> # docs <- tm_map(docs, stemDocument)
>
>
> dtm <- TermDocumentMatrix(docs)
> m <- as.matrix(dtm)
> v <- sort(rowSums(m),decreasing=TRUE)
> d <- data.frame(word = names(v),freq=v)
> head(d, 10)
             word freq
will         will   17
freedom   freedom   13
ring         ring   12
dream       dream   11
day           day   11
let           let   11
every       every    9
one           one    8
able         able    8
together together    7
>
>
> set.seed(1234)
```

```
> wordcloud(words = d$word, freq = d$freq, min.freq = 1,
+           max.words=200, random.order=FALSE, rot.per=0.35,
+           colors=brewer.pal(8, "Dark2"))
```