Question 3

```
library(tm)
```

```
## Loading required package: NLP
```

```
library(wordcloud)
```

```
## Loading required package: RColorBrewer
```

```
text <- readLines(file.choose())
docs <- Corpus(VectorSource(text))
docs <- tm_map(docs, tolower)
```

```
## Warning in tm_map.SimpleCorpus(docs, tolower): transformation drops documents
```

```
docs <- tm_map(docs, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(docs, removeNumbers): transformation drops
## documents
```

```
docs <- tm_map(docs, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(docs, removePunctuation): transformation drops
## documents
```

```
docs <- tm_map(docs, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(docs, removeWords, stopwords("english")):
## transformation drops documents
```

```
docs <- tm_map(docs, stripWhitespace)
```

```
## Warning in tm_map.SimpleCorpus(docs, stripWhitespace): transformation drops
## documents
```

```
toSpace <- content_transformer(function(x, pattern) gsub(pattern, " ", x))
docs <- tm_map (docs, toSpace, "/")
```

```
## Warning in tm_map.SimpleCorpus(docs, toSpace, "/"): transformation drops
## documents
```

```
docs <- tm_map (docs, toSpace, "?")
```

```
## Warning in tm_map.SimpleCorpus(docs, toSpace, "?"): transformation drops
## documents
```

```
docs <- tm_map (docs, toSpace, ",")
```

```
## Warning in tm_map.SimpleCorpus(docs, toSpace, ","): transformation drops
## documents
```

```
docs <- tm_map (docs, toSpace, ".")
```

```
## Warning in tm_map.SimpleCorpus(docs, toSpace, "."): transformation drops
## documents
```

```
docs <- tm_map (docs, toSpace, "@")
```

```
## Warning in tm_map.SimpleCorpus(docs, toSpace, "@"): transformation drops
## documents
```

```
docs <- Corpus(VectorSource(text))
docs <- tm_map(docs, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(docs, removeWords, stopwords("english")):
## transformation drops documents
```

```
dtm <- TermDocumentMatrix(docs)
m   <-  as.matrix(dtm)
v <- sort(rowSums(m), decreasing = TRUE)
final <- data.frame(word = names(v) , freq = v)
head(final, 10)
```

```
##              word freq
## will         will  100
## articles articles   64
## the           the   40
## project   project   39
## website   website   38
## research research   37
## user         user   36
## database database   36
## article   article   36
## phase       phase   36
```

```
wordcloud(words = final$word, freq = final$freq,min.freq = 10,max.words = 200 ,colors = brewer.pal(8, "
```


information
created
postman
figure
article
system
website
within
download
development
convert
the
articles
pdf
keywords
relational
store
requests
user's
phase
these
unit
tasks
get/post
testing
this
user
use
able
procedure
can
create
data
'texts'
used
article's
users
one
text
tables
using
interactive
computer
team
project
research
datatype
will
web
database.
database