

# Project\_DS

Naufal Nur Fahriza/Kevin Bimo S

11/26/2021

untuk memanggil library

```
library(twitteR) #menyediakan akses ke API twitter  
library(wordcloud)
```

```
## Warning: package 'wordcloud' was built under R version 4.1.2
```

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

```
library(corpus) #menganalisis teks
```

```
## Warning: package 'corpus' was built under R version 4.1.2
```

```
library(tm) #membersihkan data
```

```
## Warning: package 'tm' was built under R version 4.1.2
```

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

```
library(RCurl)
```

import data dari twitter API

```
customer_key="xIDCpqOZ2foGIDpxADlPXs5r7"  
customer_secret="ONzVwCMf0ePCtcctWcPoKcs5e8xbh7yUvB15pARzk7RJSGiEew"  
access_token="1518224958-EL54T97jUrK7gKILrJLbHW5XHDYYVpnuCNpf88W"  
access_secret="h4AGywi0bequEGYBBmqVGp1Qdm2WC2H3X06quHMgXe73H"  
setup_twitter_oauth(customer_key,customer_secret,access_token,access_secret)
```

```
## [1] "Using direct authentication"
```

```
data_vaksin <- searchTwitter("astrazeneca", n= 1000, lang="id")
```

menyiapkan data

```
vaksin.df <- twListToDF(data_vaksin)
View(vaksin.df)
```

membersihkan data

```
ulasan<-vaksin.df$text
ulasan1<-Corpus(VectorSource(ulasan))
removeURL<-function(x) gsub("http[^[:space:]]*", "", x)
textclean<-tm_map(ulasan1, removeURL)
```

```
## Warning in tm_map.SimpleCorpus(ulasan1, removeURL): transformation drops
## documents
```

```
removeNL<-function(y) gsub("\n", "", y)
textclean<-tm_map(ulasan1, removeNL)
```

```
## Warning in tm_map.SimpleCorpus(ulasan1, removeNL): transformation drops
## documents
```

```
replacecomma<-function(y) gsub(",", "", y)
textclean<-tm_map(ulasan1, replacecomma)
```

```
## Warning in tm_map.SimpleCorpus(ulasan1, replacecomma): transformation drops
## documents
```

```
removetitik2<-function(y) gsub(":", "", y)
textclean<-tm_map(ulasan1, removetitik2)
```

```
## Warning in tm_map.SimpleCorpus(ulasan1, removetitik2): transformation drops
## documents
```

```
removetitikcoma<-function(y) gsub(";", "", y)
textclean<-tm_map(ulasan1, removetitikcoma)
```

```
## Warning in tm_map.SimpleCorpus(ulasan1, removetitikcoma): transformation drops
## documents
```

```
removesymbol<-function(y) gsub("@", "", y)
textclean<-tm_map(ulasan1, removesymbol)
```

```
## Warning in tm_map.SimpleCorpus(ulasan1, removesymbol): transformation drops
## documents
```

```
removetitik3<-function(y) gsub("...", "", y)
textclean<-tm_map(ulasan1, removetitik3)
```

```
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd14
```

```

## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd14
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de25
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de25
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd23
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd23
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd7a
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd7a
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd23
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd23
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de11
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de11
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de02
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de02
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de4f
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de4f
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de43
## Warning in gsub("...", "", y): unpaired surrogate Unicode point de43
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd72
## Warning in gsub("...", "", y): unpaired surrogate Unicode point dd72

## Warning in tm_map.SimpleCorpus(ulasan1, removetitik3): transformation drops
## documents

removehashtag<-function(y) gsub("#", "", y)
textclean<-tm_map(ulasan1, removehashtag)

## Warning in tm_map.SimpleCorpus(ulasan1, removehashtag): transformation drops
## documents

remove.all<-function(xy) gsub("[^[:alpha:][:space:]]*", "", xy)
textclean<-tm_map(ulasan1, remove.all)

## Warning in tm_map.SimpleCorpus(ulasan1, remove.all): transformation drops
## documents

```

```
dataframe<-data.frame(text=unlist(sapply(textclean, '[')), stringsAsFactors=F)
View(dataframe)
write.csv(dataframe, file="vaksin.csv")
```