

Scrapping terhadap review Hotel Astoria di Lombok melalui website tiket.com :

Didapatkan review pengunjung sebagai berikut :



Dari visualisasi diatas dapat disimpulkan bahwa terlihat ada beberapa kata yang memiliki ukuran yang lebih menonjol (lebih besar) dalam hal ini terkait karena kata-kata tersebutlah yang sering muncul pada review para pengunjung hotel. Adapun beberapa kata yang memiliki ukuran yang lebih besar yaitu ada orang terbantu, nyaman, makanan, kamar, bagus, enak, dan lain-lain. Dari visualisasi diatas saya tidak menemukan komentar negative dari hotel tersebut. Sehingga wajar dikatakan bahwa hotel ini merupakan salah satu hotel terbaik yang ada di Pulau Lombok, oleh sebab itu banyak orang-orang bahkan turis-turis baik domestic maupun non domestic memilih hotel tersebut untuk disinggahi. Dari review tersebutlah para pihak hotel Astoria

dapat semakin terus meningkatkan kualitas dan fasilitas yang sudah baik menjadi semakin baik dan tetap konstan dalam segala aspek hal yang sudah sekiranya berkesan dihati pelanggan.

Adapun syntax dalam pembuatan visualisasi tersebut sebagai berikut :

```
library(xml2)
library(rvest)

Astoria = read_html("https://www.tiket.com/hotel/indonesia/lombok-astoria-
formerly-golden-tulip-mataram-lombok-108001534490332766?checkin=2021-10-
29&checkout=2021-10-30&room=1&adult=1")
Astoria

review<-html_nodes(Astoria,".review-content")
reviewtext<-html_text(review)
reviewtext

reviewtextbaru<-gsub("\n","",reviewtext)
reviewtextbaru

reviewtextbaru <- gsub("[[:punct:]]", " ", reviewtextbaru)
reviewtextbaru <-gsub("[[:space:]]", " ", reviewtextbaru)
write.csv(reviewtextbaru,"/Users/macbook/Documents/SEMESTER 5/BUSINESS
INTELLIGENCE DAN MACHINE LEARNING/TUGAS/reviewtextbaru.csv")

library(tm)
library(SnowballC)
library(wordcloud)
library(RColorBrewer)
library(stringr)

dokumen<-readLines("/Users/macbook/Documents/SEMESTER 5/BUSINESS INTELLIGENCE
DAN MACHINE LEARNING/TUGAS/reviewtextbaru.csv")
dokumen

dokumen <- VCorpus(VectorSource(dokumen))
dokumen
```

```

str(dokumen)

dokumen[[1]]

dokumen[[2]]$content

dokumenDTM<-DocumentTermMatrix(dokumen,control=list(tolower=TRUE,
                                                    removeNumbers=TRUE,
                                                    stopwords=TRUE,
                                                    removePunctuation=TRUE,
                                                    stemming=TRUE))

dokumenDTM
inspect(dokumenDTM)
str(dokumenDTM)
dokumenDTM$dimnames$Terms

dokumenfreq<-findFreqTerms(dokumenDTM,3)
dokumenfreq
dokumenfreq<-findFreqTerms(dokumenDTM,3)
dokumenfreq

dokkudtm <- TermDocumentMatrix(dokumen)
em <- as.matrix(dokkudtm)
em
ve <- sort(rowSums(em),decreasing=TRUE)
de <- data.frame(word = names(ve),freq=ve)
head(de, 15)

stopw = readLines("/Users/macbook/Documents/SEMESTER 5/BUSINESS INTELLIGENCE
DAN MACHINE LEARNING/TUGAS/id.stopwords.txt")
de = de[!is.element(de$word, stopw),]

wordcloud(words = de$word, freq = de$freq, min.freq = 1,
          max.words=70, random.order=FALSE, rot.per=0.20,
          colors=brewer.pal(8, "Dark2"))

vee <-as.list(findAssocs(dokkudtm, terms =c("fasilitas", "hotel", "sangat"),
corlimit = c(0.15,0.15,0.15,0.15,0.15,0.15)))

```