# COMP3162 Project 1

### $620156895\ 620156898$

2025-03-29

#### Part I a

```
library(wordcloud)
## Loading required package: RColorBrewer
library(RColorBrewer)
library(wordcloud2)
library(tm)
## Loading required package: NLP
library(syuzhet)
library(SnowballC)
travel_review.df <- read.csv("traveler_reviews.csv", stringsAsFactors = FALSE)</pre>
thailand_review <- travel_review.df [travel_review.df $Country=="Thailand", ]
malaysia_review <- travel_review.df[travel_review.df$Country=="Malaysia", ]
japan_review <- travel_review.df[travel_review.df$Country=="Japan", ]</pre>
Thailand
thailand_corpus <- Corpus(VectorSource(thailand_review$ReviewText))</pre>
thailand_corpus <- tm_map(thailand_corpus, content_transformer(tolower))</pre>
## Warning in tm_map.SimpleCorpus(thailand_corpus, content_transformer(tolower)):
## transformation drops documents
thailand_corpus <- tm_map(thailand_corpus, removePunctuation)</pre>
## Warning in tm_map.SimpleCorpus(thailand_corpus, removePunctuation):
## transformation drops documents
thailand_corpus <- tm_map(thailand_corpus, removeWords, c(stopwords("en"), "bampb", "etc", "amp"))
## Warning in tm_map.SimpleCorpus(thailand_corpus, removeWords, c(stopwords("en"),
## : transformation drops documents
```

```
honeymoon
                                                                                                                                 beautiful buses although average opened
                                                                                                  check
                                                                                piazza young conditioning provided
                                                                                                                                                                                           /ay
                                                                                             outside early easy
                                                                                                                                                                                                             sleep buffet
                                                                centre available leaning ভূল parking
                                  service every co
                                                                                                                                                 bathroom problem
                                                                                                                    like
                                           distance
                  facilities friendly day Just airportwill gave window
                                                                                                                                                                                                                                 central
                                                                                                                                                                                                                                                 placeflight
                                                                                                                                                                                                                                                wifi
                                                     window
foundhelptul
                                                                                                                                                                                                                           also really plenty
                        looked can
                                                                                                                                                                                                                      alk nights
                                     free bar
                                gotnext DISA
                 city dispersion of the control of th
                                                                                                                                                                                                                                               noise appropriate made punts, restauran
                                                                                      Ē
             within
beds
drinks wop u.w
                                                                                                                                                                                                                                                                            overall
                                                ē
                                                                                       aiı
                                                morning little
                                                                                                                                                                            station minute
                                                                                                                                                                            large quite me so de street get get
                                                      huge comfortable first
                                           reception restaurants twobed
                                                                                                                         excellent
                                                                 loved main
                                                                                                                                                                                                                         fresh
                                                               shopping places welcome people people
                                                                                                          places directly
                                                                                                                                                                                    probably
                                                                                                                 recommend bathrooms
```

#### Malaysia

```
malaysia_corpus <- Corpus(VectorSource(malaysia_review$ReviewText))
malaysia_corpus <- tm_map(malaysia_corpus, content_transformer(tolower))

## Warning in tm_map.SimpleCorpus(malaysia_corpus, content_transformer(tolower)):
## transformation drops documents

malaysia_corpus <- tm_map(malaysia_corpus, removePunctuation)

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

malaysia_corpus <- tm_map(malaysia_corpus, removeWords, c(stopwords("en"), "bampb", "'d", "etc", "amp"

## Warning in tm_map.SimpleCorpus(malaysia_corpus, removeWords, c(stopwords("en"),
## : transformation drops documents

wordcloud(malaysia_corpus, max.words = 150, random.order = FALSE, colors = brewer.pal(8, "Dark2"))</pre>
```

```
tookeuro longer
                    mins alberto
                    want across osleep located shopping
              however minutes choice around service coffee. Service coffee.
            price
              value great glocation small trip cheap
               side
      near visit right to day
                                                      comfortable
                                    bit clean
provided even One O
        free can
                                                     night nextcar
tuscany close
                                                                leaning
shuttle dont
  turn friendly
   restaurants
                                                               early<sub>make</sub>
 quitelong light
walking large onice available shortwifi
                                                              placewalls
         e left airport staff et also floor placewalls although stayeduse modern really Walkarea find pay sure
        ertwill
ଞ୍ଚ left airpor
ଝୁ stayeduse
            river little will minute helpfullovelyfirst take
           highly double shower didnt needed desk
                    water apartment bookedmay see
                homemade
                            restaurant stop dinner something windows
                           something
```

#### Japan

```
japan_corpus <- Corpus(VectorSource(japan_review$ReviewText))
japan_corpus <- tm_map(japan_corpus, content_transformer(tolower))

## Warning in tm_map.SimpleCorpus(japan_corpus, content_transformer(tolower)):

## transformation drops documents

japan_corpus <- tm_map(japan_corpus, removePunctuation)

## Warning in tm_map.SimpleCorpus(japan_corpus, removePunctuation): transformation

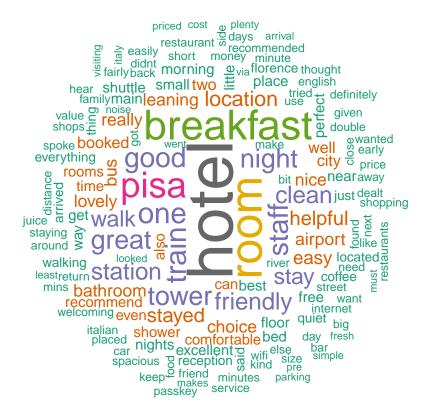
## drops documents

japan_corpus <- tm_map(japan_corpus, removeWords, c(stopwords("en"), "bampb", "etc", "amp"))

## Warning in tm_map.SimpleCorpus(japan_corpus, removeWords, c(stopwords("en"), :

## transformation drops documents

wordcloud(japan_corpus, max.words = 150, random.order = FALSE, colors = brewer.pal(8, "Dark2"))</pre>
```



Similarities: The words "hotel", "room" and "pisa" were the most common words in all three countries word cloud. The word "tried" was common in Malaysia's and Japan's word cloud, the word "fresh" was common in Thailand's and Japan's word cloud and the word "one" was common in Malaysia's and Thailand's word cloud.

Differences: The word "croissants" was only present in Thailand's word cloud, the word euro was only present in Malaysia's word cloud and the word "spacious" was only present in Japan's word cloud.

#### Part I b

Thailand Dominant Sentiment

```
thailand_sentiments <- get_nrc_sentiment(thailand_review$ReviewText)
dominant_thailand_sentiment <- names(which.max(colSums(thailand_sentiments)))
print(paste("The dominant sentiment for Thailand is:", dominant_thailand_sentiment))
## [1] "The dominant sentiment for Thailand is: positive"

Malaysia Dominant Sentiment
```

```
malaysia_sentiments <- get_nrc_sentiment(malaysia_review$ReviewText)
dominant_malaysia_sentiment <- names(which.max(colSums(malaysia_sentiments)))
print(paste("The dominant sentiment for Malaysia is:", dominant_malaysia_sentiment))</pre>
```

## [1] "The dominant sentiment for Malaysia is: positive"

## Japan Dominant Sentiment

```
japan_sentiments <- get_nrc_sentiment(japan_review$ReviewText)
dominant_japan_sentiment <- names(which.max(colSums(japan_sentiments)))
print(paste("The dominant sentiment for Japan is:", dominant_japan_sentiment))</pre>
```

## [1] "The dominant sentiment for Japan is: positive"