© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

Detection and Prediction of Customers Feedback on Hotel Businesses using Sentimental Machine Learning and Data Mining Algorithms

UWAKWE VIVIAN OLUCHI[1], ELEBERI EBELE[2] AND OGUOMA IKECHUKWU STANLEY[3]
[1,2]*Department of Computer Science, Imo State University Owerri*
[3]*Department of Computer Science, University of Agriculture and Environmental Science Umuagwo, Imo State, Nigeria*

Abstract*—Sentimental analysis has established enormous improvement in recent years that not even a private field has been left from its grips. The hotel business today requires a good analysis on the feedback gotten from customers so that great decision could be made. The motivation behind this research is to ensure that customers are provided with satisfactory goods and services and the hunger for accurate decision making and achieving rapid growth within the hotel business. The aim of the research is to adopt sentimental machine learning and data mining algorithm in identifying, predicting and analyzing Customers feedback on hotel businesses while the objectives includes: to analyze customers feedback so as to identify lapses in service delivery in hotel business, to provide accurate and reliable information for top decision making within the organization and hence make accurate future forecast. Data mining methodology called sentimental analysis was adopted while implementing the aspect-base sentiment modeling tool. The data was analyzed with R and SAS Enterprise Miner. The experiment was done using twenty (20) different hotel/restaurants sourced from kaggle repository. The result after the experiment shows that the use of aspect-based sentimental analysis is the best and more accurate method suitable for data prediction and hence has more percentage acceptance level of performance when it comes to opinion mining peoples opinion on some issues, therefore it could be adopted for future use by various organizations to understand customers view on their products and services so that decision can be made accurately.

Keywords*—Artificial Intelligence, Machine Learning, Sentiment Analysis, Aspect-Based Model and Data Mining, Customers Feedback.

I. INTRODUCTION

In business, opinion of customers decides the progress of the entire company. Once a particular product has been liked and highly demanded by customers it simply means that there is product satisfaction and thus increase in profit made in the company. Today, the birth of internet and its huge benefits to most business can never be over emphasized, as it has brought or rather provided various companies an open score card to capture, market, sale and analyze their products remotely. Looking at the rate, at which hotel business is springing up, one could begin to imagine how they will be able to satisfy their customers because of the nature of the business. Their business type deals with human taste of food and consumables and thus huge customers comments, reactions and reports are highly needed so that there can be an improvement in areas where necessary. Sentiment analysis as defined by (Deshmukh and Bardekar; 2020) deals with attitude, reaction to events, comment, responses both emotional and contextual to a document or report, or action of a speaker towards a particular information. The scholars further states that sentiment analysis could also refers to the usage of natural language processing, computational biometric or semantics, or textual analysis to either systematically quantify, extract, identify or verify the state of a given information. In this context, hotel business has been chosen as the dataset for this research; hence the research is going to look into customer's responses (Voice notes, written comments, video or animation of consumable products) in regards to social media handles, it will also look into areas like comments from the hotel website, and materials and reports. Deshmukh and Bardekar (2020) states that sentiment analysis has penetrated various domain area more importantly in the aspect of customers responses, marketing of products, customers services online and offline, customers services in medical field etc. The position of sentiment analysis in the area of natural language processing and large-scale data mining can never be underestimated. Handling and making decision in regards to business growth is not an easy task; it requires concrete facts and evidence before it could be achieved. Nevertheless, this research is

motivated base on ensuring that customers are provided with satisfactory goods and services and hunger for accurate decision making and achieving rapid growth. The aim of this research is to adopt sentimental machine learning and data mining algorithm in identifying and analyzing Customers feedback on hotel businesses while the objectives includes: to analyze customers feedback so as to identify lapses in service delivery in hotel business, to provide accurate and reliable information for top decision making within the organization and hence make accurate future forecast.

This paper is organized as follows: Introduction: presents an overview of the importance of sentimental analysis in an organization, definition of sentimental analysis, areas sentimental analysis could be applied, aim and objectives of the study. Literature Review: looks at generally the literature review on related works, types of sentimental analysis, Methodology: methodology adoption for the study, proposed system diagram, while Results: present results, conclusion and recommendation of the study.

II. LITERATURE REVIEW

As stated by (Deshmukh and Bardekar; 2020), that sentimental analysis data mining approach mostly used by researchers is bag-of-words approach data gotten from an online review or social media. This approach, instead of looking into the whole sentence or rather paragraph for the analysis, it will only consider individuals words and their count as the feature vectors. Based on the research conducted by Deshmukh and Bardekar, their work demonstrated that semantics constructed feature vector with ensemble classifier output was far better performed the traditional bag-of-words procedure with a single machine learning classifier of 3-5%. The researchers observed the ensemble method of the experiment outperforming the traditional bag-of-words approach with a very huge difference and thus among other methods used, ensemble approach showed that extremely randomized trees classification performance was far better than others. Kiritchenko et al. (2014) presented a review adopting lexicon and linguistic structures while (Castellucci et al; 2014) adopted the bag-of-words approach after the learning from an external data. On the same note, Hu and Liu (2004) implemented an association rule-based system used for aspect identification. Secondly Liu, B (2012), outlined four different procedures on how to extract aspects sentimental which includes: Target Relation, frequent phrases, opinion and supervised learning and topic models. Jakob and Gurevych (2010) in their research adopted the conditional random fields for the aspect sentimental term

TYPES OF SENTIMENT ANALYSIS

According to monkeylearn.com stated that sentiment analysis models deal mainly on polarity (i.e., positive, negative or neutral) also deals on the feelings and emotions i.e. (angry, happy or sad), in times of urgency deals on (urgent, not urgent) and finally, (interested or not interested). One could interpret customer's responses and queries base on the type of sentimental analysis needed.

FINE-GRAINED SENTIMENT ANALYSIS

This type of sentimental analysis deals with ranting. Some referred to it as 5-star-rating where one could have:

a. Very positive
b. Positive
c. Neutral
d. Negative
e. Very negative

Considering the business type one in involved in, ratings could start with, (Very Positive= 5 stars and Very Negative = 1 Star) depends on choice.

EMOTION DETECTION

Sentimental analysis to detect once emotions could be attributed with this type of sentimental analysis. Some of the facts it could detect includes: happiness, frustration, anger, sadness etc. Most researches for detection on sentiment analysis uses lexicons (List of words and the emotions which it conveys) or rather the complex machine learning algorithm approach as stated by (monkeylearn.com). Though using lexicon is good but its disadvantages is that people could express their emotions differently such as (Your customer voice is superb or your goods is on in town), hence which in turn means happiness towards the product or services delivered.

ASPECT-BASED SENTIMENT ANALYSIS

Aspect based sentimental analysis are usually applied when handling texts or say product reviews. Organizations would want to know which particular features end user or people are mentioning if in positive, neutral, or negative perspective. That is to

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

say, in hotel business, aspect sentimental analysis could analyze customers view or comments on a particular product, where the model will be able to pin point or determine that the sentence made by the customer is negative opinion or positive respect to the product or services rendered.

MULTILINGUAL SENTIMENT ANALYSIS

According to (Deshmukh and Bardekar; 2020), Multilingual sentiment analysis requires preprocessing and huge resources. The resources for adopting multilingual sentiment analysis are gotten online either from (sentiment lexicons, translated corpora or noise detection algorithms). An expert in program coding is highly needed before it could be used. Furthermore, detection of languages of text could also be achieved automatically with MonkeyLearn language classifier, after the training of a custom sentiment analysis model to classify any text to once choice has been achieved.

III. METHODOLOGY

The adopted methodology for the research is Aspect based sentimental analysis, that is to say it was structured using the aspect based features of (positive, and negative perspective). The hotel data set was gotten from the kaggle dataset repository and was saved as XLS (Microsoft Excel Format). From the excel file, 20 hotels where randomly selected and used for this research structured in the following categories:

ID, Hotel /Restaurant Name, Location, Review Date, and Review, Figure 1 shows the proposed system sentimental analysis on the hotel review.

EXPLANATION OF DATASET

This is a hotel review data set which contains the reviews for 537 different hotels.

And can be accessed through kaggle. Sentimental analysis can be supervised or unsupervised, According to medium.com. There is no dependent variable because this is an unsupervised learning approach in which the researcher tends to find the sentiment, opinion of the customers of the hotels based on the reviews in dataset provided. The dataset contains 53644 observations with 5 variables. The variables in the dataset are:

1. ID – a number that uniquely differentiate a review from another review (9-digit integer), that starts with "RN" meaning Review number
2. Hotel/Restaurant name – Name of the hotel
3. Location —Location of the hotel.
4. Review Date – description of the date when each review was conducted.
5. Review – The review of customers about the hotel/restaurant.

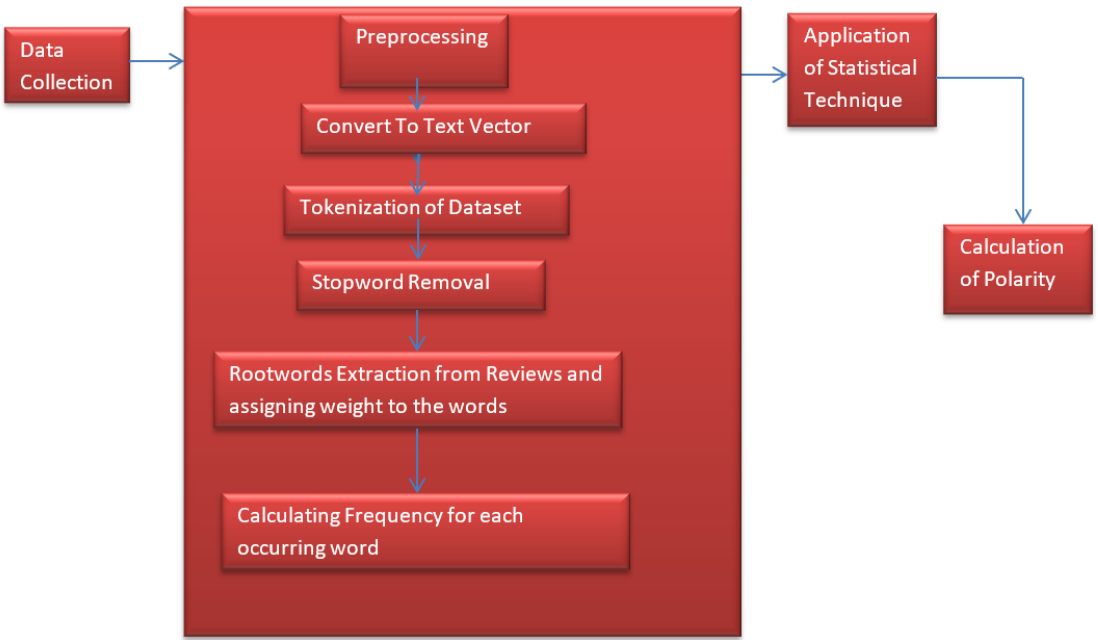


Figure 1: proposed system sentimental analysis on the hotel review

IV. RESULTS

EXPERIMENTS ON THE DATASET USING R

The main objective in this task is to find the thought, opinion, and idea of the customers of 20 randomly selected hotels based on how they

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

feel about the hotels product and services by grouping the review into negative and positive and to determine the percentage of the review with the negative and positive LEXICON.

Step1: Loading packages to be used (that is libraries)

##Load necessary packages
library(skimr) ##for data exploration
library(readxl) # read excel into r dataframe
library(tm)
library(SnowballC)
library(wordcloud) #for visualization
library(dplyr) ## for filteration
library(ggplot2) ###Create graphics and charts

Step 2: Loading My Dataset to R Dataframe

First was to create a variable name and assign it to the dataset while loading it to the R dataframe.

cn <- c("ID", "Hotel", "Location", "Date", "Review")
hotel_reviews <- read_excel("hotel_reviews.xlsx", col_names = cn)

Step 3: Exploring the data, at this stage, skimr and others was used and created report with the dataexplorer package. I viewed the data set, checked the names of the variable, checked the head (first six rows), the tail (the last six rows), the structure which shows the datatypes of the variables, the dimension of the dataset,

```
> head(hotel_reviews)
# A tibble: 6 x 5
  ID        Hotel               Location Date            Review
  <chr>     <chr>               <chr>    <chr>           <chr>
1 ID        Hotel/Restaurant n~ Location Review Date     "Review"
2 rn579778~ Thong Dee The Kath~ Kathu    Reviewed 1 wee~ "Just been for sunday roast lamb and beef truly excellent,11out of 10\r\r\~
3 rn576350~ Thong Dee The Kath~ Kathu    Reviewed 3 wee~ "Quietly set off the main road, nice atmosphere. Immaculate and friendly s~
4 rn574921~ Thong Dee The Kath~ Kathu    Reviewed 4 wee~ "I made a reservation for a birthday two days in advance assuming we would~
5 rn572905~ Thong Dee The Kath~ Kathu    Reviewed April~ "We visit here regularly and never fail to be impressed by the quality and~
6 rn572364~ Thong Dee The Kath~ Kathu    Reviewed April~ "Visited this wonderful place on my travels and i cannot complain about an~
> |
```

Figure 2: Showing the first six rows of the Data

```
> tail(hotel_reviews)
# A tibble: 6 x 5
  ID        Hotel   Location Date            Review
  <chr>     <chr>   <chr>    <chr>           <chr>
1 rn163047~ Bite in Patong   Reviewed June ~ "we came here for lunch with our family.\r\r\nRestaurant nice and cozy.\r\r\nFood was g~
2 rn162368~ Bite in Patong   Reviewed May 2~ "I love this small restaurant, for the great food, and the extreme friendliness of its ~
3 rn161843~ Bite in Patong   Reviewed May 2~ "We stopped at this restaurant after shopping at Jung Ceylon. It's close to main \"squa~
4 rn161734~ Bite in Patong   Reviewed May 2~ "Great times ,This is one of the best restaurant in Phuket town. The owners and staffs ~
5 rn161218~ Bite in Patong   Reviewed May 1~ "Bite in, it has become my favorite restaurant in Phuket. The food absolutely delicious~
6 rn161212~ Bite in Patong   Reviewed May 1~ "The restaurant put in a good location in Jungcylon department store,Patong Phuket. I w~
> |
```

Figure 3: showing the last six rows of the data

Step 4: Conversion of hotel column to factor to view the levels

Step 5: EXTRACTION/ filtration of 20 hotels at random and explore it with ggplot(bargraph) to get frequency of review per hotel

```
# A tibble: 20 x 2
   Hotel                                                `Number of reviews`
   <fct>                                                              <int>
 1 Sabai Sabai                                                          200
 2 Audy Restaurant                                                      100
 3 B-Hive Gallery, Bar and Restaurant                                   100
 4 DaVinci Restaurant                                                   100
 5 Dee-Lish Bar & Restaurant                                            100
 6 Istanbul Turkish Restaurant                                          100
 7 Kitchen Bistro Patong                                                100
 8 Kwality Indian Restaurant                                            100
 9 Legends Sports Bar & Grill                                           100
10 O-OH Farm Ta-Eiad                                                    100
11 Sabaijai Cafe                                                        100
12 Sawasdee Thai Cuisine                                                100
13 The Breakfast Hut                                                    100
14 The Coffee Club - The Turtle Village                                 100
15 Chomtalay restaurant at Mandarava Resort and Spa                      98
16 Bon Island Restaurant                                                 96
17 Jeffer                                                                96
18 Oasis Bar                                                             96
19 Dibuk restaurant                                                      95
20 Navrang Mahal                                                         93
> |
```

Figure 4: showing number of reviews per hotel (of the 20hotel)

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

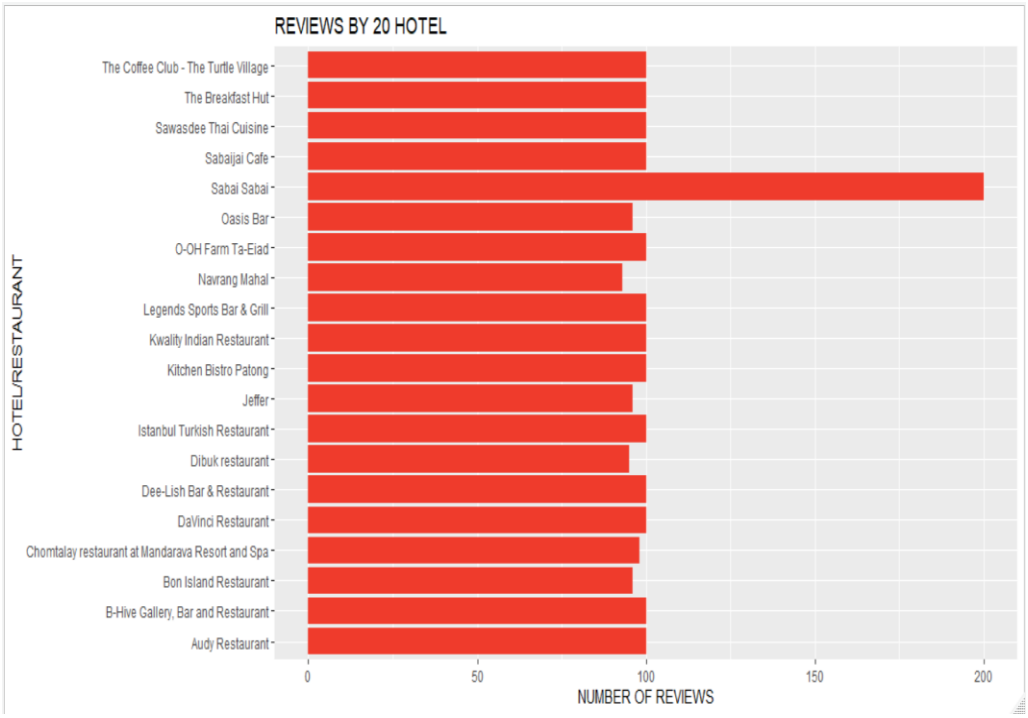


Figure 5: plot showing the reviews of the 20 number of the hotel

Step 6: creation of text vectors

	V1
1	second time to Phuket and couldn't go past visiting this rest...
2	Tonight my husband, Master 14 and Miss 11 ventured from ...
3	A very pleasant place to have diner, where great turkish foo...
4	Yum! Platter for two was more than enough for myself and ...
5	Really nice welcoming restaurant on the road side, worth th...
6	Nice ambiance, warm and friendly staff and very delicious di...
7	Super tasty food. Perfect service and perfect location. Very fr...
8	We have a holiday home at Kata and felt like something diff...
9	We found this restaurant via Trip Advisor and are glad we di...
10	We went here for breakfast and we were so satisfied after o...
11	When we want eat Turkish food , we just go to Istanbul rest...
12	I Come back again today, and the food Was more delicious ...
13	This is a must when coming to Phuket. Just up the hill from ...
14	knowing all our food is hand made, inc our cream, bread, b...

Showing 1 to 14 of 2,074 entries, 1 total columns

Figure 6: showing the text vectors

Step 7: Exploration of text vectors

```
> count_words(review_hotel_20)
  [1] 38 46 43 40 46 24 17 46 46 46 45 31 39 41 46 25 46 20 46 46 46 19 49 25 43 18 34 46 32 26 22 46 46 18 46 21 30 38 35 46
 [41] 17 20 35 48 45 45 16 36 18 46 46 45 46 45 18 46 46 46 46 39 35 46 27 39 46 23 41 29 46 19 46 46 46 22 30 32 45 46 44 46
 [81] 28 22 33 46 45 46 22 32 41 46 28 46 46 28 40 45 23 25 37 46 32 46 46 47 23 14 46 46 46 46 20 36 46 47 47 35 32 46 19 22
[121] 30 46 45 20 16 44 28 52 39 34 28 46 43 46 46 33 74 29 47 45 45 32 50 46 25 18 30 47 46 46 46 43 46 24 44 46 46 39 32 17
[161] 38 47 29 46 38 46 22 17 30 26 46 29 46 23 47 28 39 46 20 26 44 45 46 46 42 28 25 54 38 23 17 42 47 46 23 38 48 24 37 37
[201] 25 46 46 29 46 52 48 46 46 46 45 46 45 46 41 46 42 46 46 46 25 22 46 42 46 40 36 48 46 46 45 36 47 31 43 36 19 46 21 21
[241] 35 46 46 172 23 46 23 46 42 47 44 18 37 23 37 30 24 45 46 46 45 45 46 46 47 45 46 46 46 25 46 46 40 48 46 37 46 46 47 35
[281] 45 46 21 46 46 46 38 48 52 19 46 47 45 38 47 45 46 42 45 46 47 63 46 46 46 21 22 46 46 46 38 32 29 46 24 46 27 46 46 46
[321] 20 25 35 43 45 46 46 46 39 46 46 20 16 36 23 20 46 20 45 20 30 100 35 23 21 46 37 37 33 43 19 25 26 23 32 38 46 39 46 46
[361] 31 26 28 39 22 18 22 21 32 23 36 46 46 45 46 29 47 22 25 27 43 38 49 35 43 27 24 28 34 22 45 46 46 34 46 43 29 33 42 34
[401] 46 19 46 44 46 17 44 23 43 52 22 44 36 46 34 27 46 30 21 23 38 46 46 37 46 46 46 23 43 44 47 47 46 46 45 45 28 37 45
[441] 46 46 38 46 39 24 45 27 46 46 33 20 46 47 46 43 45 17 46 46 45 47 46 46 44 46 35 46 44 41 45 46 46 30 46 46 43 46 46 47
[481] 35 25 36 35 46 48 47 46 47 20 46 46 46 34 46 47 35 39 46 19 45 46 19 45 31 26 24 29 47 27 20 44 36 46 47 22 16 22 27 46
[521] 45 46 46 46 46 45 46 46 46 46 46 45 25 46 46 47 21 35 27 26 24 16 46 46 24 46 20 46 46 41 35 46 45 29 46 46 46 47 46 46
[561] 32 44 32 46 33 44 24 34 30 46 46 46 17 46 46 46 45 34 46 29 42 47 18 35 46 24 46 46 38 46 23 46 35 46 45 46 46 37 43 46
[601] 45 46 46 23 35 45 31 48 29 46 40 46 43 25 45 40 46 41 19 26 43 47 28 20 28 46 49 31 25 45 46 46 46 19 21 20 21 45 46 29
[641] 45 45 46 24 52 28 46 21 46 46 46 18 47 46 45 46 46 46 41 46 45 24 42 21 46 46 33 45 42 46 46 45 27 46 23 45 24 46 46 27
[681] 45 27 41 39 46 46 36 47 45 46 46 46 46 25 20 30 43 23 46 26 37 46 39 46 46 48 26 22 46 30 46 38 19 22 21 25 46 43 19 17
[721] 45 23 26 24 45 50 48 46 32 25 20 26 46 21 45 31 30 46 39 37 46 45 46 46 31 44 46 46 30 46 45 25 46 46 15 46 22 46 46 46
[761] 46 46 46 46 46 46 46 25 45 29 46 44 46 32 47 47 44 45 46 44 46 37 46 36 46 46 46 47 47 46 20 40 46 47 46 17 45 46 44 23
[801] 45 32 46 36 43 45 37 47 38 37 32 46 46 46 46 46 19 22 28 45 89 27 17 23 29 34 18 24 22 26 43 26 27 45 42 47 38 23 46
[841] 47 46 28 40 46 30 21 46 46 47 33 27 33 50 31 29 37 46 47 45 42 33 45 44 44 33 36 30 46 41 46 28 26 41 46 41 24 46 46 46
[881] 42 42 31 44 33 30 46 46 29 47 31 27 46 38 22 20 25 43 25 25 46 46 46 46 50 46 29 27 46 27 43 29 37 20 47 46 35 27 21 46
[921] 21 46 43 42 46 44 27 37 46 18 40 37 42 33 46 46 47 37 46 42 46 22 46 46 32 46 45 24 46 25 24 37 30 45 41 46 39 46 46 25
[961] 43 31 46 23 18 37 24 18 23 46 38 46 19 25 46 24 17 34 24 23 19 22 43 46 28 20 16 45 46 26 46 46 46 42 28 28 27 19 46 46
[ reached getOption("max.print") -- omitted 1074 entries ]
```

Figure 7: showing frequency of word in a row. E.g., 38, 46, 43, 40 and 46 words from 1st to 5th row

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

```
count_characters(review_hotel_20)
 [1] 218 277 248 224 272 141 111 274 228 228 239 167 211 231 262 132 262 134 252 234 241 118 274 137 237 106 185 252 186 130 132 281 238 114 228 132 171 184 207 219
[41] 113 107 188 266 244 217 117 190 118 234 259 249 270 274 107 273 244 221 244 219 203 246 145 230 241 115 250 151 280 100 231 237 280 121 162 162 242 244 235 268
[81] 144 139 209 266 264 253 123 185 236 244 172 223 238 157 220 259 119 165 214 250 163 262 265 238 119 109 240 265 262 265 113 238 248 263 227 197 174 247 119 106
[121] 143 255 262 116 105 230 167 275 234 197 153 238 221 243 254 167 319 161 242 282 265 177 255 269 129 122 192 270 267 275 251 251 265 117 234 252 233 219 187 112
[161] 188 282 171 265 203 257 139 101 170 143 249 156 264 135 274 146 202 272 131 135 229 240 258 257 231 156 146 297 225 120 113 229 254 258 149 241 268 157 227 202
[201] 128 246 238 168 250 278 280 239 238 234 261 228 240 287 208 256 197 261 244 230 123 117 241 218 250 223 188 263 238 249 230 204 249 162 257 211 101 277 124 105
[241] 164 236 252 734 125 236 128 221 216 264 244 118 216 121 181 152 152 227 255 263 270 276 234 258 299 240 217 246 229 125 272 216 230 279 243 175 248 245 248 183
[281] 250 246 109 239 262 244 210 252 265 102 243 233 229 212 229 226 247 226 250 233 229 313 265 287 258 116 120 276 244 272 233 198 167 252 142 232 144 226 250 249
[321] 116 144 184 232 249 244 246 243 214 259 261 119 101 210 147 120 271 118 226 118 160 467 192 138 125 294 185 216 185 221 106 131 148 152 187 243 280 222 247 260
[361] 205 161 151 239 144 123 122 120 193 138 202 234 272 240 227 142 272 127 126 137 241 195 242 193 237 164 126 153 188 122 237 240 266 177 288 228 157 173 243 177
[401] 255 126 251 215 223 101 244 120 232 309 141 256 201 243 185 154 248 162 117 131 191 264 249 206 237 262 260 253 138 249 232 255 248 244 245 234 277 128 189 234
[441] 214 243 222 242 217 122 223 170 271 236 205 146 232 252 233 241 252 110 244 279 232 269 249 244 222 252 195 258 208 222 258 243 248 160 245 241 260 236 256 258
[481] 186 139 187 204 238 242 268 264 225 117 249 258 238 182 240 260 180 221 252 116 222 223 118 263 154 165 137 154 268 142 117 214 190 243 252 117 104 119 150 229
[521] 260 276 250 248 259 254 239 240 242 240 262 265 155 243 241 242 110 179 141 136 126 108 237 258 129 238 125 251 244 211 196 281 253 148 237 259 228 267 240 242
[561] 181 218 169 259 180 238 139 180 167 246 255 232 105 239 231 256 245 215 237 159 254 284 103 188 254 126 271 241 226 233 126 254 193 247 228 240 240 210 214 252
[601] 255 221 240 126 213 251 162 240 146 238 211 243 251 129 247 192 267 238 101 166 245 254 159 111 120 264 268 183 140 232 253 236 257 137 133 105 105 237 235 173
[641] 242 242 249 142 256 163 269 125 252 253 249 105 256 287 267 228 231 245 241 258 260 132 245 135 228 233 174 276 245 264 242 259 165 248 124 246 148 258 239 160
[681] 228 156 221 266 272 236 185 262 248 224 268 249 258 156 110 160 238 119 244 165 236 252 259 259 253 266 144 126 251 172 255 252 119 117 110 163 253 224 112 104
[721] 232 133 142 145 257 291 255 245 197 141 131 155 260 123 249 163 158 269 228 229 271 235 261 264 164 259 287 272 191 262 264 148 239 270 110 260 127 253 242 236
[761] 246 264 253 231 241 257 262 141 231 168 227 251 257 191 275 250 261 228 253 277 280 210 255 203 251 256 257 228 266 248 130 240 261 277 233 141 259 291 270 109
[801] 245 178 264 205 208 229 221 241 211 196 198 228 179 255 235 235 229 105 134 151 240 400 161 109 134 159 188 108 126 142 129 231 145 146 236 225 242 221 124 250
[841] 265 237 133 205 246 160 118 226 263 252 185 164 192 288 174 148 196 286 251 250 218 216 234 234 239 190 202 151 233 197 264 146 134 232 234 224 121 244 271 255
[881] 208 219 182 234 163 147 258 260 147 284 146 142 258 194 148 114 144 219 139 139 253 252 247 237 281 245 162 160 291 150 237 152 209 128 264 261 203 144 118 245
[921] 124 285 229 211 270 275 159 202 277 104 230 197 220 175 251 283 266 223 269 251 208 110 260 251 218 254 222 135 264 141 153 208 144 244 207 242 212 244 257 144
[961] 249 174 234 136 103 237 158 114 138 218 212 286 109 140 271 156 108 177 138 137 106 148 264 250 166 136 100 245 241 142 269 256 266 246 164 156 153 109 244 262
[ reached getOption("max.print") -- omitted 1074 entries ]
```

Figure 8: showing frequency of characters in a row. E.g., 218,277,248,224 and 272 characters from 1st to 5th row

```
count_sentences(review_hotel_20)
  [1]  3  4  6  5  4  3  4  6  5  4  1  4  6  1  5  2  6  4  5  4  5  3  9  3  4  3  4  2  5  2  3  5  5  3  4  2  2  8  2  5  2  1  5  5  8  4  2  6  4  5  5  5  5
 [54]  9  5  5  8  4  4  4  4  3  4  3  6  3  5  1  4  2  6  7  6  2  3  3  6  7  3  4  4  2  5  6  6  2  4  5  4  5  3  4  5  3  3  4  1  4  5  7  2  4  8  4  2  3
[107]  4  5  7  5  2  9  4  8  3  4  3  6  5  2  2  7  4  2  2  5  5  8  2  2  4  5  6  2  1  2  2  6 12  9  3  2  4  2  2  3  7  5  6  6  3  5  2  4  8  3  5  5
[160]  4  6  5  6  7  3  6  1  2  1  4  5  4  4  4  6  3  2  6  5  3  2  3  5  4  4  7  3  7  4  1  2  5  5  4  2  6 10  1  3  3  4  4  4  3  7  6  6  5  6  5  5  5
[213]  6  5  6  7  2  5  4  4  1  5 14  4  3  2  3  5  7  4  5  4  4  2  3  4  3  5  2  4  3  4  3  3  4  4  5  4  2  7  3  3  5  2  4  2  3  5  3  9  4 10  5  4  2
[266]  3  4  4  5  1  9  3  8  6  5  3  7  6  6  4  3  4  2  3  4  8  2  4  6  3  5  3  8  5  6  5  6  6  6  4  4  7  5  5  2  2  8  5  8  5  7  2  4  3  5  3  4
[319]  4  6  3  3  2  7  6  6  4  3  1  2  4  1  2  3  3  2  7  2  4  2  4  7  5  3  3  8  4  6  3  3  3  2  4  2  1  3  5  4  8  5  3  2  1  7  3  5  1  2  6  1  2
[372]  3  4  5  5  2  4  1  2  2  6  2  3  2  4  5  4  3  2  5  5  5  6  2  4  6  5  2  4  3  5  7  3  7  2  5  3  4  3  7  4  6  7  4  3  5  2  5  2  1  3  3  6  5  4
[425]  7  5  5  4  4 10  4 10  4  5  2  4  3  1  6  4  2  5  5  3  6  2  3  4  8  4  5  4  2  6  4  2  3  4  6  5  3  5  6  4  2  2  7  2  3  3  2  4  6  3  3  2  4
[478]  5  5  4  3  2  3  4  5  4  4  4  4  3  4  7  4  2  7  2  4  3  5  1  4  4  3  7  2  2  2  2  8  4  2  8  2 12  4  4  9  1  2  4  7  4  5  6  6  2  5  4  4  5
[531]  7  7  3  5  6  5  2  2  1  2  2  3  5  5  1  7  2  5  7  3  3  7 10  4  4  3  5  4  5  4  3  7  3  6  4  8  1  4  5  5  5 10  2 12  6  5  3  6  6  2  4  5  1
[584]  5  5  1  5  4 15  6  1  5  3  3  4  5  4  6  4  5  5  5  4  2  2  6  2  4  3  4  3  5  5  2  3  4  5  2  2  2  3  2  2  4  1  5  6  6  3  5  6  4  4  2  2  1
[637]  2  4  7  3  2  5  8  3  4  3  3  2  6  4  3  1  5  4  3  8  3  7  3  5  4  3  5  6  4  7  3  8  7  5  6  3  3  5  3  5  4  9  3  2  9  1  3  6  6  4  3 10  5
[690]  4  4  6  7  2  2  3  7  3  5  5 15  7  6  5  8  6  2  3  6  6  6  5  3  3  3  1  7  5  3  3  6  3  3  1  5  3  6  5  4  2  5  9  6  1  6  3  3  3  4  4  4  2
[743]  5 10  1  5  6  6  1  5 16  2  3  6  2  4  4  5  6  5  6  5  5  5  5  7  3  4  1  4  2  5  4  5  4 10  6  5  3  4  5  5  4  3  6  5  7  4  6  2  4  6  5  3
[796]  3 10  1  5  1  5 11  5  3  3  4  4  4  4 12  7  5  4 12  1  3  3  4  7  1  2  1  5  6  4  1  4  3  2 10  3  7  1  6  4  4  5  5  6  2  4  3  4  1  4
[849]  7  5  4  2  2  5 11  1  5  4  4  4  4  5  4  5 10  4  1  2  4  5  5  2  2  1  5  4  1  3  4  7  2  4  3  2  3  4  8  5  4  5  2  1  4  2  2  3  2  3  2  3  8
[902]  6  5  6  2  5  3  2  7  4  4  2  2  3  5  3  3  5  1  3  5  4  4  4  5  4  3  4  1  6  3  5  3  4  5  4  3  5  3  4  2  5  4  2  5  3  3  5  2  1  2  4  6
[955]  4  5  6  4  4  3  3  3  5  4  3  4  4  2  4  4  3  5  2  1  7  4  2  3  3  2  1  4  3  5  2  4  1  7  4  1  5  3  4  4  2  2  3  1  5  9
[ reached getOption("max.print") -- omitted 1074 entries ]
```

Figure 9: showing frequency of sentence in a row. E.g., 218,277,248,224 and 272 sentences from 1st to 5th row

```
> ng1 <- ngram(review_hotel_20,n=1,)
> print(ng1)
An ngram object with 9862 1-grams
> get.phrasetable(ng1)
          ngrams freq          prop
1            and  3294 0.0427864445
2            the  3140 0.0407861067
3              a  1832 0.0237962253
4             to  1493 0.0193928845
5            was  1390 0.0180549963
6             is  1329 0.0172626547
7             of  1001 0.0130021952
8           food   993 0.0128982815
9             in   895 0.0116253393
10           for   886 0.0115084365
11           The   827 0.0107420733
12            we   715 0.0092872823
13             I   700 0.0090924442
14          very   663 0.0086118436
15            We   646 0.0083910271
16          with   603 0.0078324912
17          this   565 0.0073389014
18          were   557 0.0072349877
19          good   535 0.0069492252
20           but   489 0.0063517217
21            it   485 0.0062997649
22  restaurant   480 0.0062348189
23            on   469 0.0060919376
24           are   459 0.0059620455
25           had   456 0.0059230779
26            at   451 0.0058581319
27           you   446 0.0057931859
```

Figure 10: showing frequency of words that occurred in the text vector

```
> ng2 <- ngram(review_hotel_20, n = 2)
> print(ng2)
An ngram object with 40124 2-grams
> get.phrasetable(ng2)
              ngrams freq          prop
1            and the   359 0.0047922256
2             of the   232 0.0030969258
3           the food   213 0.0028432982
4           food was   201 0.0026831124
5           The food   192 0.0025629730
6            food is   159 0.0021224621
7           food and   152 0.0020290203
8             it was   147 0.0019622762
9             in the   146 0.0019489274
10            on the   139 0.0018554857
11            at the   139 0.0018554857
12             for a   138 0.0018421369
13            to the   137 0.0018287881
14          the best   136 0.0018154392
15             is a    126 0.0016819511
16  this restaurant   122 0.0016285558
17          is very   112 0.0014950676
18        this place   107 0.0014283235
19           had a    107 0.0014283235
20      friendly and   102 0.0013615794
21          was very    97 0.0012948353
22           we were    95 0.0012681377
23           a great    91 0.0012147424
24          from the    87 0.0011613472
25             was a    86 0.0011479983
26            we had    86 0.0011479983
```

Figure 11: showing frequency of two words that occurred in the text vector

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

```
> ng3 <- ngram(review_hotel_20, n = 3)
> print(ng3)
An ngram object with 60387 3-grams
> get.phrasetable(ng3)
                ngrams freq         prop
1          The food was   88 1.208144e-03
2          the food was   66 9.061080e-04
3           the food is   62 8.511924e-04
4          and the food   57 7.825478e-04
5           The food is   54 7.413611e-04
6            one of the   43 5.903431e-04
7            and it was   40 5.491564e-04
8           of the best   37 5.079696e-04
9       and the service   36 4.942407e-04
10        and the staff   31 4.255962e-04
11   very friendly and    30 4.118673e-04
12     The service was    28 3.844095e-04
13           We had a     28 3.844095e-04
14    times during our    28 3.844095e-04
15     The staff were     26 3.569516e-04
16        a few times     26 3.569516e-04
17       very good and    25 3.432227e-04
18            a lot of    25 3.432227e-04
19      the service is    24 3.294938e-04
20     the service was    24 3.294938e-04
21       the staff are    23 3.157649e-04
22            it was a    23 3.157649e-04
23         We ate here    23 3.157649e-04
24         a couple of    23 3.157649e-04
25        good and the    22 3.020360e-04
26      during our stay   22 3.020360e-04
```

Figure 12: showing frequency of three words that occurred in the text vector

Step 8: tokenization of text vectors
Step 9: convert text to lower case
Step 11: Removal of links from the reviews
Step 12: Removal of punctuation from the reviews
Step 13: Removal of digits from the reviews
Step 14: Removal of leading blank spaces at the beginning from the reviews
Step 15: Removal of blank spaces at the end from the reviews
Step 16: Remove redundant word, restaurant, and hotel from the reviews
Step 17: inspection of vector after cleaning
Step 18: Converting the text vectors to corpus
Step 19: Clean up corpus by removing stop words and Whitespace
Step 20: Stem the words to their root of all reviews present in the corpus
Step 21: Load the positive and negative lexicon data
Step 22: Inspect lexicons

```
> head(positive_lexicon)
          a.
1     abound
2    abounds
3  abundance
4   abundant
5 accessable
6 accessible
> tail(positive_lexicon)
             a.
2000      yay
2001 youthful
2002     zeal
2003   zenith
2004     zest
2005    zippy
>
```

```
> head(negative_lexicon)
    X2.faced
1    2-faces
2   abnormal
3    abolish
4 abominable
5 abominably
6  abominate
> tail(negative_lexicon)
      X2.faced
4777    zapped
4778      zaps
4779    zealot
4780   zealous
4781 zealously
4782    zombie
>
```

Figure 13: Checking the lexicon from head to tail

Step 23: Creating a function for sentimental analysis. A function that will generate wordcloud, the count of total positive and negative words in each review, Calculate overall percentage of positive and negative words of all the reviews, Create a dataframe with all the positive and negative reviews and print the dataframe to the console.

Step 24: Using the sentiment function created to calculate the percentage of positive reviews for the 20 hotel that was randomly selected.

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

```
  Review_Type Count
1     Postive  4309
2    Negitive   846
[1] "Percentage of Positive Reviews: 83.59 %"
>
```

Figure 14: showing positive and negative count with positive percentage review

Figure 15: showing wordcloud visualization of 20hotels

Step 25: Create a function that will do step 8 to step 24 and name it sent_all_proc (sentimental all procedure)
Step 26: FILTER/EXTRACTION of the 20 hotels separately from the dataset (step 5)
Step 27: Creation of the text vectors for each of them and apply the sent_all_proc. Repeat it for all the 20 hotel text vectors

```
sent_all_proc <- function(x){
  require(gsub)
  require(tokenizers)
  require(tm)
  require(wordcloud)
  z0 = tokenize_words(x)
  z1 = tolower(VectorSource(z0))
  z2 = gsub("http\\S+\\s*|[[:punct:]]|[[:digit:]]|^|$|hotel|restaurant", "",z1)
  z3 = Corpus(VectorSource(z2))
  z4 = tm_map(z3, removeWords, stopwords("english"))
  z5 = tm_map(z4, stripWhitespace)
  z6 = tm_map(z5, stemDocument)
  res = sentiment(z6)
  return(res)
}
```

Figure 16: Showing the Function Created

Step 28: For simplicity, the researcher will only show the wordcloud and percentage output for 5 hotels only

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442


```
  Review_Type Count
1     Postive    71
2    Negitive    45
[1] "Percentage of Positive Reviews: 61.21 %"
```

Sabai-Sabai hotel

Figure 17: wordcloud and percentage output for sabia-sabai hotel


```
  Review_Type Count
1     Postive    54
2    Negitive    24
[1] "Percentage of Positive Reviews: 69.23 %"
```

Audy Restaurant

Figure 18: wordcloud and percentage output for Audy restaurant

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

B-Hive Gallery, Bar and Restaurant

Figure 19: wordcloud and percentage output for B-hive gallery, bar and restaurant

DaVinci Restaurant

Figure 20: wordcloud and percentage output for DaVinci restaurant

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442


```
  Review_Type Count
1     Postive    50
2    Negitive    16
[1] "Percentage of Positive Reviews: 75.76 %"
```

Istanbul Turkish Restaurant

Figure 21: wordcloud and percentage output for Istanbul Turkish restaurant

V. ANALYSIS AND SUMMARY ON R

The summary focuses on how the experiment was done in RStudio. As stated earlier in this report, the main aim of this task is to find the thought, opinion, and idea of the customers of 20 randomly selected hotels based on how they feel about the hotels product and services by grouping the review into negative and positive and to determine the percentage of the review with the negative and positive LEXICON. After launching the RStudio, the loading of packages to be used (that is libraries) are done and then the dataset was loaded into the environment by first creating a variable name and assign it to the dataset while loading it to the R dataframe after which exploring the data is done, at this stage, skimr and others were used to create report with the dataexplorer package. After exploration, the dataset was viewed, and the names of the variable were checked including the head (first six rows), the tail (the last six rows), the structure which shows the datatypes of the variables in figure 2 and figure 3, and the dimension of the dataset. Then trying to check the levels of the hotels in the data set, there is need to Convert the hotel column to factor and then carryout the EXTRACTION/ FILTRATION of 20 hotels at random and explore it with ggplot(bargraph) to get frequency of review per hotel shown figure 4 while figure 5 shows a plot of the reviews of the 20 number of the hotel. Creating the text vectors shown in figure 6 while figure 7: shows the frequency of word in a row. E.g., 38, 46, 43, 40 and 46 words from 1st to 5th row, figure 8: showing frequency of characters in a row. E.g., 218,277,248,224 and 272 characters from 1st to 5th row, Figure 9: shows frequency of sentence in a row. E.g., 218,277,248,224 and 272 sentences from 1st to 5th row, figure 10: shows frequency of words that occurred in the text vector, figure 11: shows frequency of two words that occurred in the text vector and figure 12: shows frequency of three words that occurred in the text vector respectively. Tokenization of the text vector was done which led to conversion of the text to lower case and cleaning the data set by removal of some variables such as: removal of links from the reviews, removal of punctuation from the reviews, removal of digits from the reviews, removal of leading blank spaces at the beginning from the reviews, removal of blank spaces

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

at the end from the reviews, remove redundant word, restaurant, and hotel from the reviews and then inspection of the vector after cleaning. The Converting of the text vectors to corpus made it possible for the removal of the stopwords and whitespace and then stem the words to their root of all reviews present in the corpus. The positive and negative lexicon data are now loaded the inspected shown in figure 13. Now in other to create the wordcloud for the sentimental analysis, a function is called. The function generated the wordcloud and the count of total positive and negative words in each review are done and, calculation of overall percentage of positive and negative words of all the reviews and Creation of a dataframe with all the positive and negative reviews and print the dataframe to the console are achieved. Using the sentiment function created to calculate the percentage of positive reviews for the 20 hotel that was randomly selected shown in figure 14 while Figure 15: shows wordcloud visualization of 20hotels. Furthermore, because this research is concerned with only 20 hotels, step 8 to step 27 are repeated shown in Figure 16 the Function Created and then for simplicity, the researcher only show the wordcloud and percentage output for 5 hotels only that is figure 17 shows the wordcloud and percentage output for sabia-sabai hotel, figure 18 shows the wordcloud and percentage output for Audy restaurant, figure 19 shows the wordcloud and percentage output for B-hive gallery, bar and restaurant, figure 20 shows the wordcloud and percentage output for DaVinci restaurant and figure 21 shows the wordcloud and percentage output for Istanbul Turkish restaurant respectively.

EXPERIMENTS ON THE DATASET USING SAS ENTERPRISE MINER

The data set was first cleaned in R by extracting the 20 random hotels in which column ID, LOCATION, REVIEW DATE was removed. The objective of this experiment is to find test topic and cluster of word that makes up the sentiments and the opinion of customers of the randomly selected hotel.

Step 1: Launch the SAS enterprise miner

Step 2: Create a project

Step 3: Create a Diagram

Step 4: Drag and drop the file import node from the sample menu bar and import the dataset on the enterprise miner software

Step 5: exploring the dataset in SAS

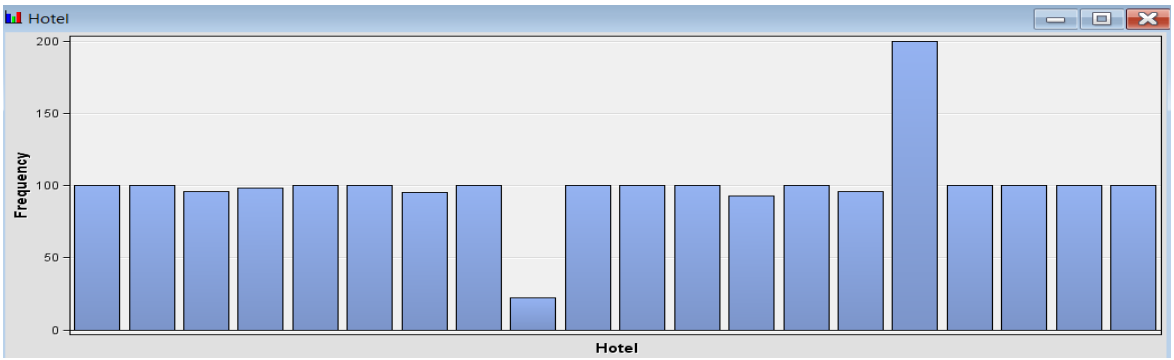


Figure 22: showing frequency pf reviews per hotel

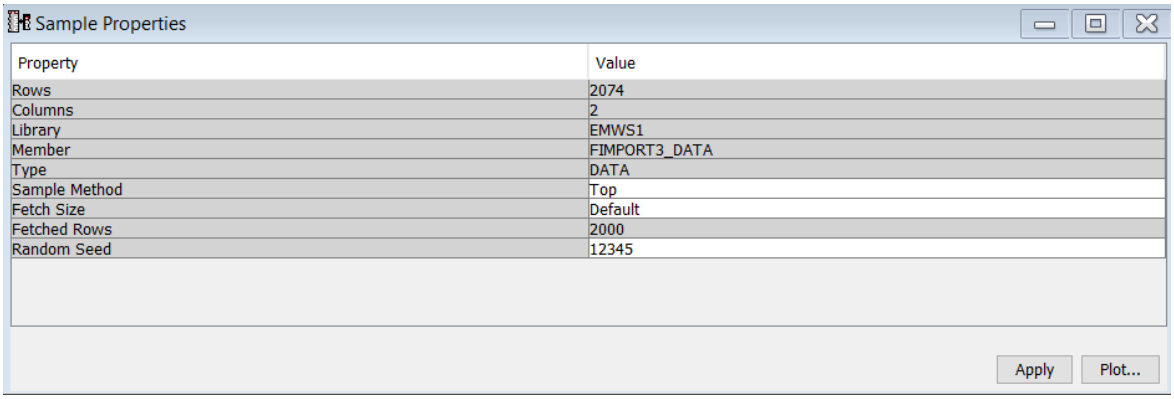
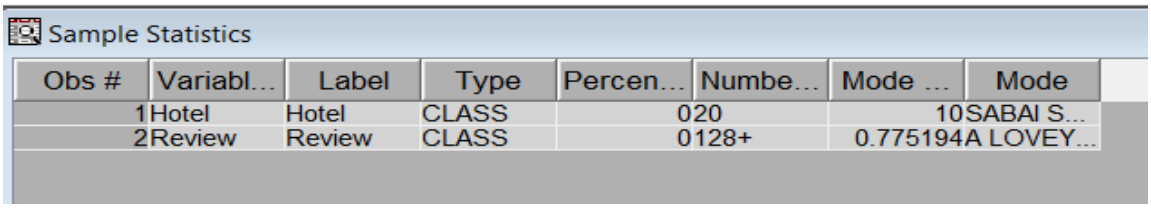


Figure 23: showing Sample report of the dataset

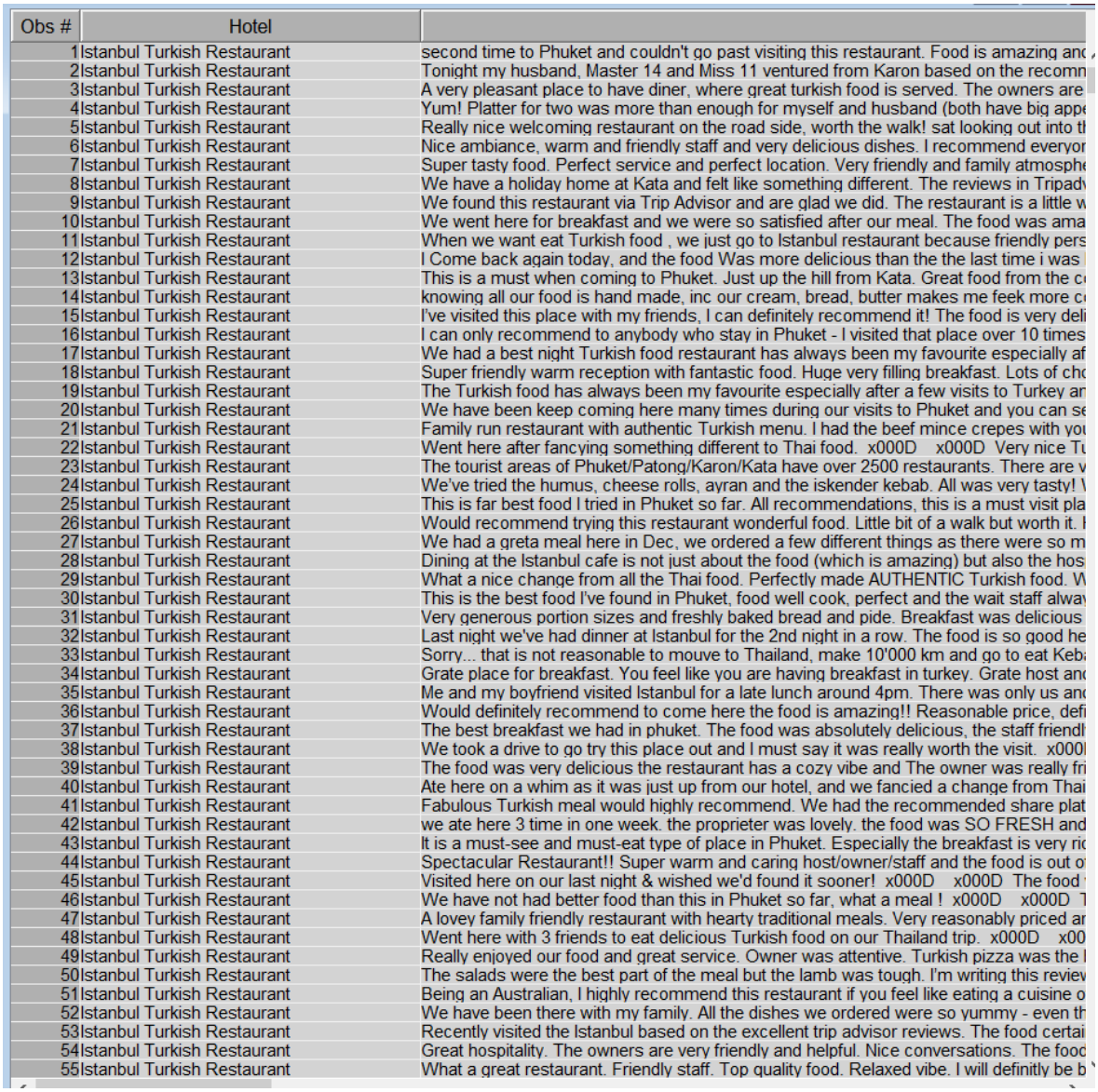
© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442



Sample Statistics

Obs #	Variabl...	Label	Type	Percen...	Numbe...	Mode ...	Mode
1	Hotel	Hotel	CLASS		020	10	SABAI S...
2	Review	Review	CLASS		0128+	0.775194	A LOVEY...

Figure 24: showing Sample statistics of the dataset



Obs #	Hotel	
1	Istanbul Turkish Restaurant	second time to Phuket and couldn't go past visiting this restaurant. Food is amazing and
2	Istanbul Turkish Restaurant	Tonight my husband, Master 14 and Miss 11 ventured from Karon based on the recomn
3	Istanbul Turkish Restaurant	A very pleasant place to have diner, where great turkish food is served. The owners are
4	Istanbul Turkish Restaurant	Yum! Platter for two was more than enough for myself and husband (both have big appe
5	Istanbul Turkish Restaurant	Really nice welcoming restaurant on the road side, worth the walk! sat looking out into th
6	Istanbul Turkish Restaurant	Nice ambiance, warm and friendly staff and very delicious dishes. I recommend everyor
7	Istanbul Turkish Restaurant	Super tasty food. Perfect service and perfect location. Very friendly and family atmosphe
8	Istanbul Turkish Restaurant	We have a holiday home at Kata and felt like something different. The reviews in Tripad
9	Istanbul Turkish Restaurant	We found this restaurant via Trip Advisor and are glad we did. The restaurant is a little w
10	Istanbul Turkish Restaurant	We went here for breakfast and we were so satisfied after our meal. The food was ama
11	Istanbul Turkish Restaurant	When we want eat Turkish food , we just go to Istanbul restaurant because friendly pers
12	Istanbul Turkish Restaurant	I Come back again today, and the food Was more delicious than the the last time i was
13	Istanbul Turkish Restaurant	This is a must when coming to Phuket. Just up the hill from Kata. Great food from the c
14	Istanbul Turkish Restaurant	knowing all our food is hand made, inc our cream, bread, butter makes me feek more c
15	Istanbul Turkish Restaurant	I've visited this place with my friends, I can definitely recommend it! The food is very deli
16	Istanbul Turkish Restaurant	I can only recommend to anybody who stay in Phuket - I visited that place over 10 times
17	Istanbul Turkish Restaurant	We had a best night Turkish food restaurant has always been my favourite especially af
18	Istanbul Turkish Restaurant	Super friendly warm reception with fantastic food. Huge very filling breakfast. Lots of cho
19	Istanbul Turkish Restaurant	The Turkish food has always been my favourite especially after a few visits to Turkey ar
20	Istanbul Turkish Restaurant	We have been keep coming here many times during our visits to Phuket and you can se
21	Istanbul Turkish Restaurant	Family run restaurant with authentic Turkish menu. I had the beef mince crepes with you
22	Istanbul Turkish Restaurant	Went here after fancying something different to Thai food. x000D x000D Very nice Tu
23	Istanbul Turkish Restaurant	The tourist areas of Phuket/Patong/Karon/Kata have over 2500 restaurants. There are v
24	Istanbul Turkish Restaurant	We've tried the humus, cheese rolls, ayran and the iskender kebab. All was very tasty! \
25	Istanbul Turkish Restaurant	This is far best food I tried in Phuket so far. All recommendations, this is a must visit pla
26	Istanbul Turkish Restaurant	Would recommend trying this restaurant wonderful food. Little bit of a walk but worth it. I
27	Istanbul Turkish Restaurant	We had a greta meal here in Dec, we ordered a few different things as there were so m
28	Istanbul Turkish Restaurant	Dining at the Istanbul cafe is not just about the food (which is amazing) but also the hos
29	Istanbul Turkish Restaurant	What a nice change from all the Thai food. Perfectly made AUTHENTIC Turkish food. W
30	Istanbul Turkish Restaurant	This is the best food I've found in Phuket, food well cook, perfect and the wait staff alwa
31	Istanbul Turkish Restaurant	Very generous portion sizes and freshly baked bread and pide. Breakfast was delicious
32	Istanbul Turkish Restaurant	Last night we've had dinner at Istanbul for the 2nd night in a row. The food is so good he
33	Istanbul Turkish Restaurant	Sorry... that is not reasonable to mouve to Thailand, make 10'000 km and go to eat Keb
34	Istanbul Turkish Restaurant	Grate place for breakfast. You feel like you are having breakfast in turkey. Grate host an
35	Istanbul Turkish Restaurant	Me and my boyfriend visited Istanbul for a late lunch around 4pm. There was only us an
36	Istanbul Turkish Restaurant	Would definitely recommend to come here the food is amazing!! Reasonable price, defi
37	Istanbul Turkish Restaurant	The best breakfast we had in phuket. The food was absolutely delicious, the staff friendl
38	Istanbul Turkish Restaurant	We took a drive to go try this place out and I must say it was really worth the visit. x000
39	Istanbul Turkish Restaurant	The food was very delicious the restaurant has a cozy vibe and The owner was really fri
40	Istanbul Turkish Restaurant	Ate here on a whim as it was just up from our hotel, and we fancied a change from Thai
41	Istanbul Turkish Restaurant	Fabulous Turkish meal would highly recommend. We had the recommended share plat
42	Istanbul Turkish Restaurant	we ate here 3 time in one week. the proprieter was lovely. the food was SO FRESH and
43	Istanbul Turkish Restaurant	It is a must-see and must-eat type of place in Phuket. Especially the breakfast is very ric
44	Istanbul Turkish Restaurant	Spectacular Restaurant!! Super warm and caring host/owner/staff and the food is out o
45	Istanbul Turkish Restaurant	Visited here on our last night & wished we'd found it sooner! x000D x000D The food
46	Istanbul Turkish Restaurant	We have not had better food than this in Phuket so far, what a meal ! x000D x000D T
47	Istanbul Turkish Restaurant	A lovey family friendly restaurant with hearty traditional meals. Very reasonably priced ar
48	Istanbul Turkish Restaurant	Went here with 3 friends to eat delicious Turkish food on our Thailand trip. x000D x00
49	Istanbul Turkish Restaurant	Really enjoyed our food and great service. Owner was attentive. Turkish pizza was the
50	Istanbul Turkish Restaurant	The salads were the best part of the meal but the lamb was tough. I'm writing this reviev
51	Istanbul Turkish Restaurant	Being an Australian, I highly recommend this restaurant if you feel like eating a cuisine o
52	Istanbul Turkish Restaurant	We have been there with my family. All the dishes we ordered were so yummy - even th
53	Istanbul Turkish Restaurant	Recently visited the Istanbul based on the excellent trip advisor reviews. The food certai
54	Istanbul Turkish Restaurant	Great hospitality. The owners are very friendly and helpful. Nice conversations. The fooc
55	Istanbul Turkish Restaurant	What a great restaurant. Friendly staff. Top quality food. Relaxed vibe. I will definitly be b

Figure 25: view of dataset showing the first 55 rows.

Step 6: Editing the variable to select the variables needed a specification of role (review= text, hotel = ID) and run

Step 7: Drag and dropped the text parsing node which parsed the text and showed some statistical visualization of the text component.

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

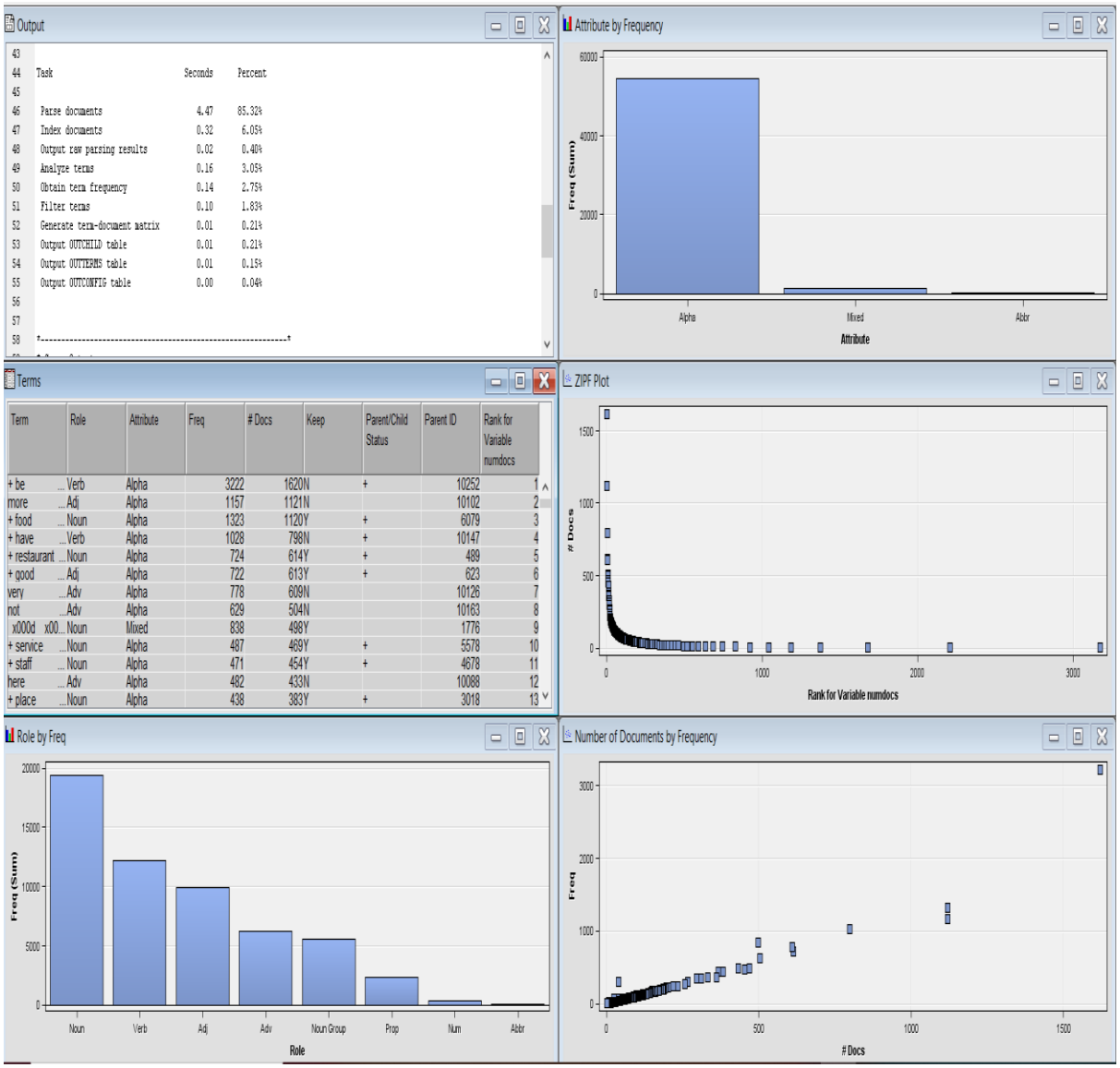
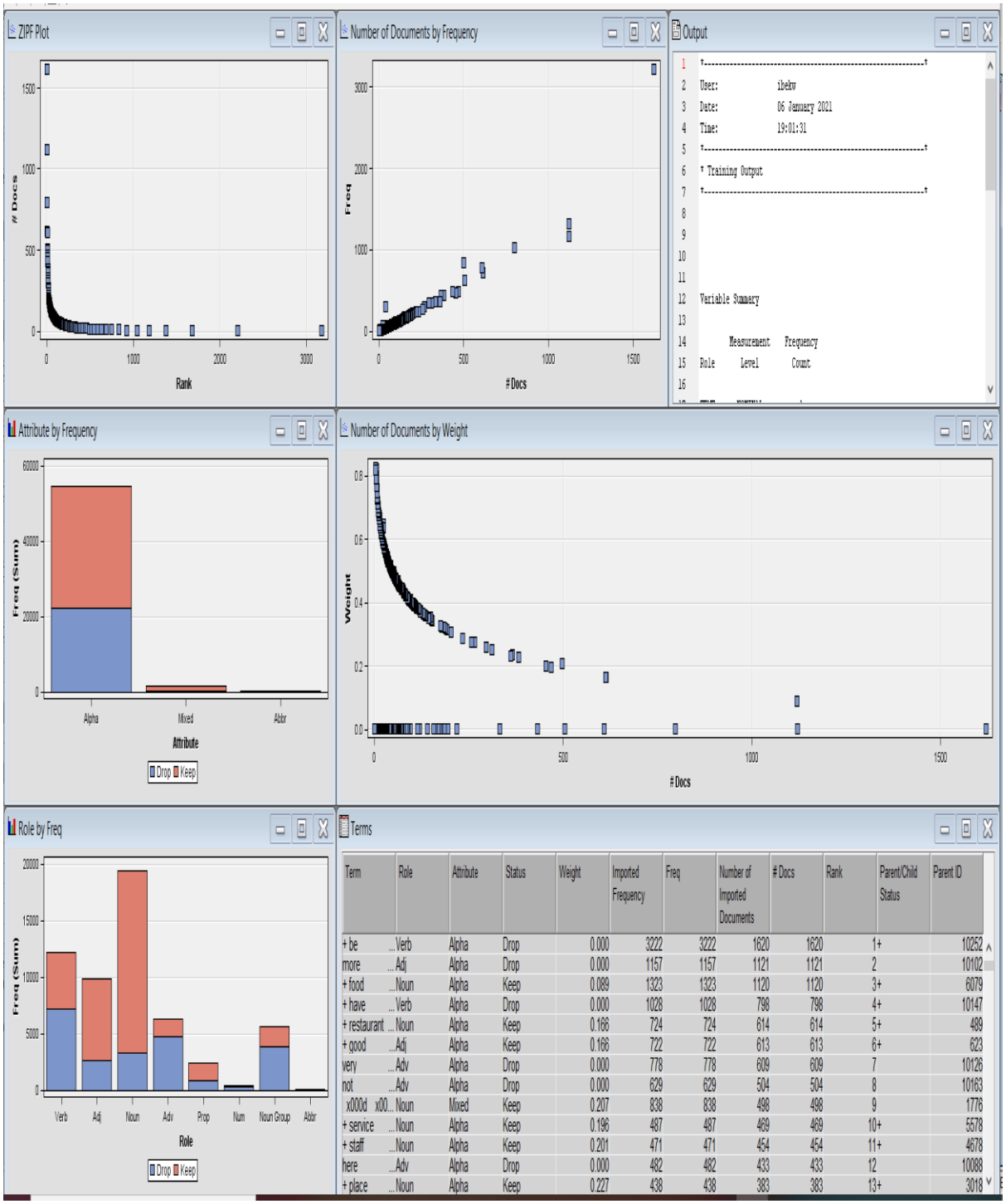


Figure 26: showing the output of text parsing node

Step 7: Drag and dropped the text filtered node which filtered the text and dropped some words and showed some statistical visualization of the text component.

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442



Term	Role	Attribute	Status	Weight	Imported Frequency	Freq	Number of Imported Documents	# Docs	Rank	Parent/Child Status	Parent ID
+ be	...Verb	Alpha	Drop	0.000	3222	3222	1620	1620	1+		10252
more	...Adj	Alpha	Drop	0.000	1157	1157	1121	1121	2		10102
+ food	...Noun	Alpha	Keep	0.089	1323	1323	1120	1120	3+		6079
+ have	...Verb	Alpha	Drop	0.000	1028	1028	798	798	4+		10147
+ restaurant	...Noun	Alpha	Keep	0.166	724	724	614	614	5+		489
+ good	...Adj	Alpha	Keep	0.166	722	722	613	613	6+		623
very	...Adv	Alpha	Drop	0.000	778	778	609	609	7		10126
not	...Adv	Alpha	Drop	0.000	629	629	504	504	8		10163
x000d x00...	...Noun	Alpha	Mixed	Keep	0.207	838	838	498	498	9		1776
+ service	...Noun	Alpha	Keep	0.196	487	487	469	469	10+		5578
+ staff	...Noun	Alpha	Keep	0.201	471	471	454	454	11+		4678
here	...Adv	Alpha	Drop	0.000	482	482	433	433	12		10088
+ place	...Noun	Alpha	Keep	0.227	438	438	383	383	13+		3018

Figure 27: showing the output of text filter node which consist of role by frequency plot, attribute by frequency plot, terms table (colored by dropped$blue and keep$red words)

Step 7: Drag and dropped the text topic node and edited the number of multi-term topic to FIVE and run. This generated statistical visualization of the text `topic component

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

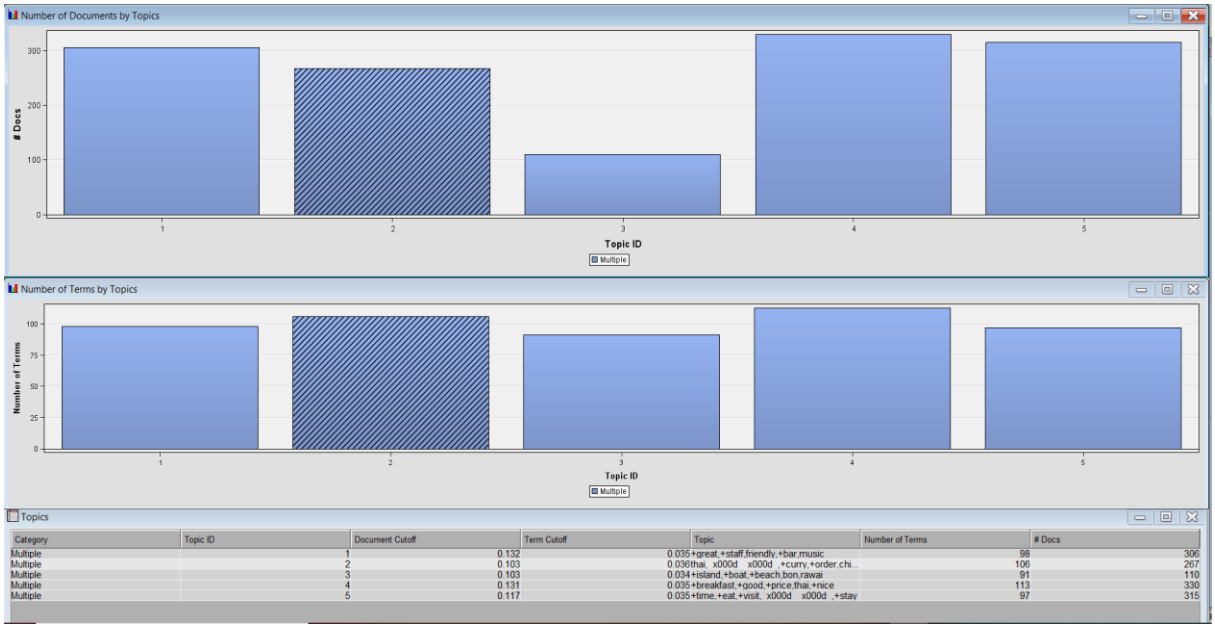


Figure 28: showing the table with five topics, bar plot of number of documents by table, bar plot of number of terms by topics.

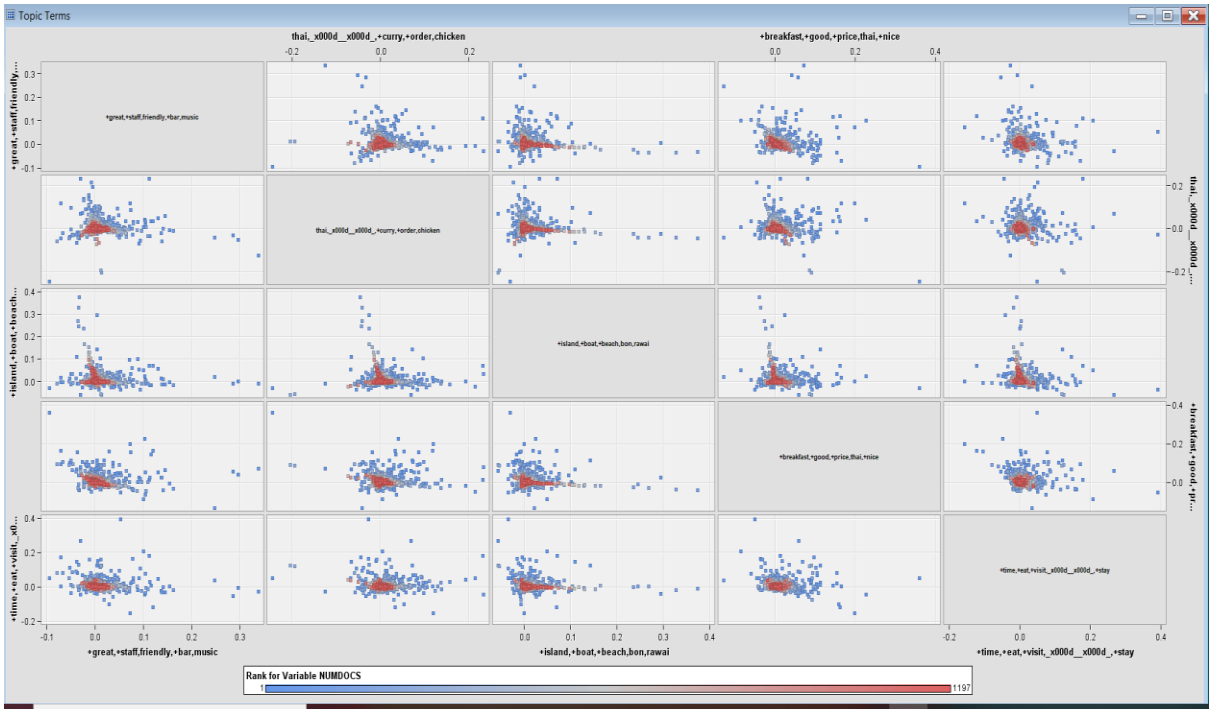
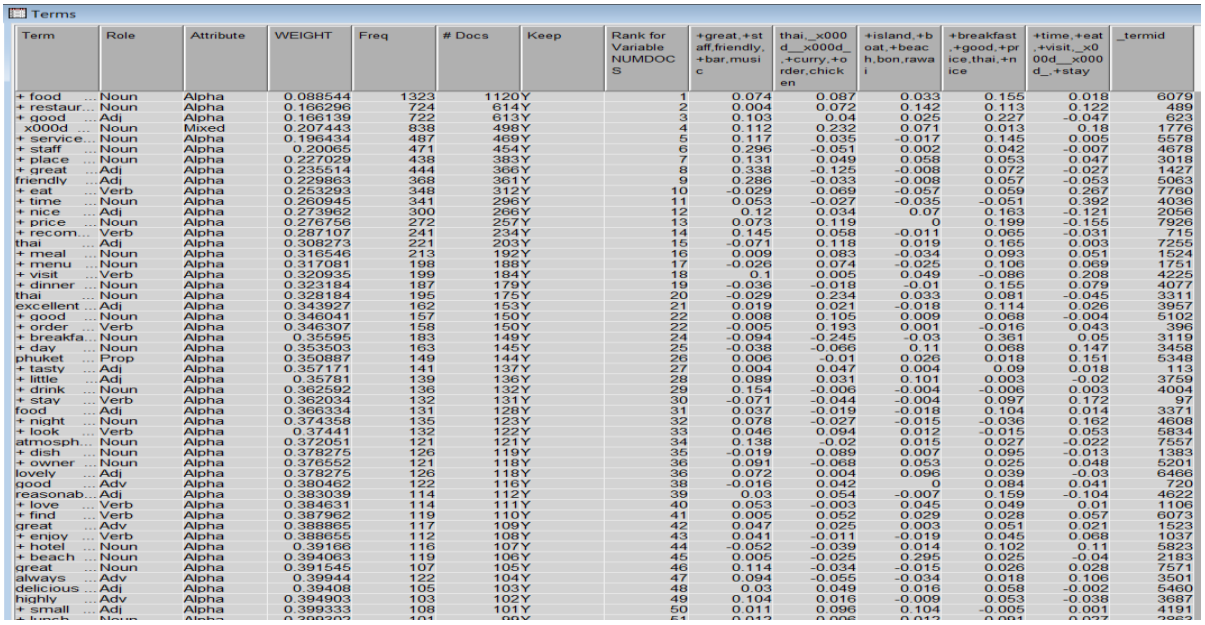


Figure 29: showing the topics term

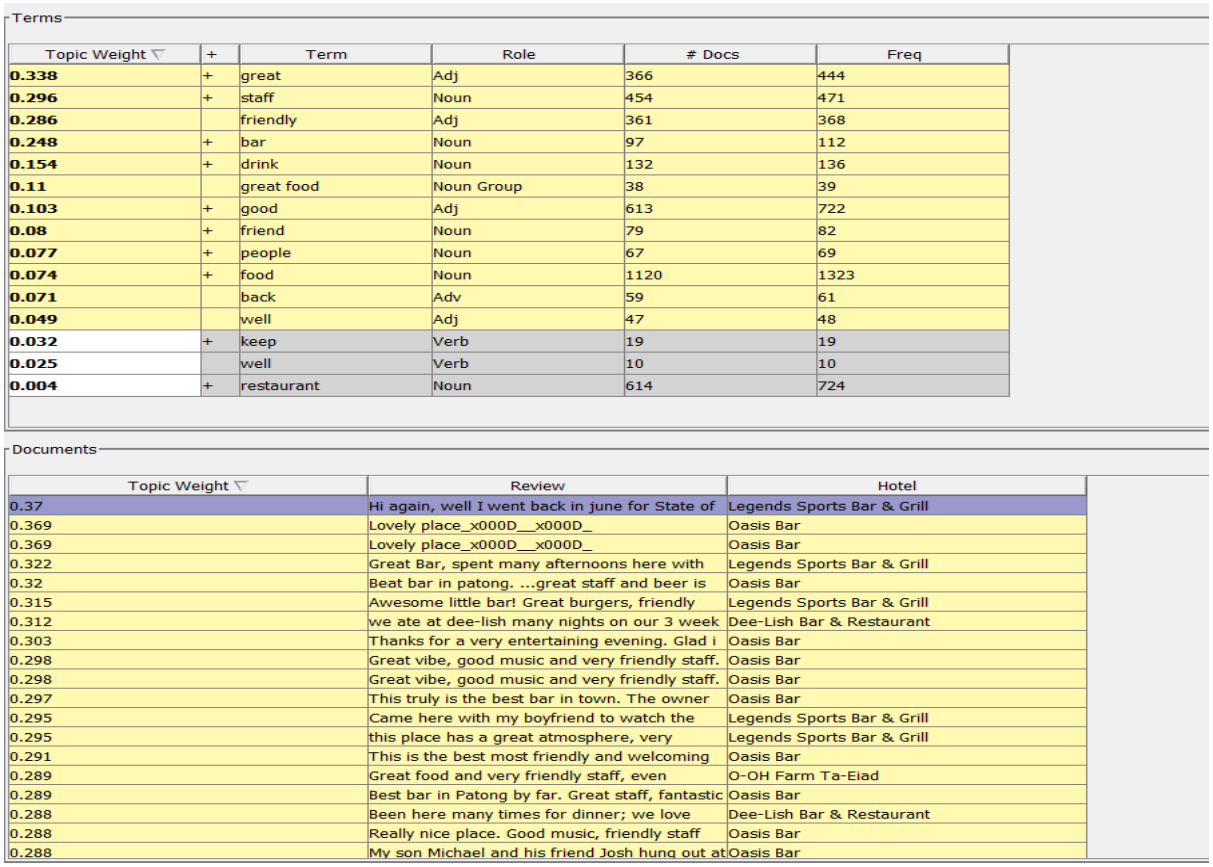
© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442



Terms

Term	Role	Attribute	WEIGHT	Freq	# Docs	Keep	Rank for Variable NUMDOCS	+great,+staff,friendly,+bar,music	thai,_x000d__x000d_,+curry,+order,chicken	+island,+boat,+beach,bon,rawai	+breakfast,+good,+price,thai,+nice	+time,+eat,+visit,_x000d_,+stay	_termid
+ food	Noun	Alpha	0.088544	1323	1120Y		1	0.074	0.087	0.033	0.155	0.018	6079
+ restaur...	Noun	Alpha	0.166296	724	614Y		2	0.004	0.072	0.142	0.113	0.122	489
+ good	Adj	Alpha	0.160139	722	613Y		3	0.103	0.04	0.025	0.227	-0.047	623
x000d	Noun	Mixed	0.207443	838	498Y		4	0.112	0.232	0.071	0.013	0.18	1776
+ service	Noun	Alpha	0.196434	487	469Y		5	0.117	0.035	-0.017	0.145	0.005	5578
+ staff	Noun	Alpha	0.20065	471	454Y		6	0.296	-0.051	0.002	0.042	-0.007	4678
+ place	Noun	Alpha	0.227029	438	383Y		7	0.131	0.049	0.058	0.053	0.047	3018
+ great	Adj	Alpha	0.235514	444	366Y		8	0.338	-0.125	-0.008	0.072	-0.027	1427
friendly	Adj	Alpha	0.229863	368	361Y		9	0.286	-0.033	-0.008	0.057	-0.053	5063
+ eat	Verb	Alpha	0.253293	348	312Y		10	-0.029	0.089	-0.057	0.059	0.267	7760
+ time	Noun	Alpha	0.260945	341	296Y		11	0.053	-0.027	-0.035	-0.051	0.392	4036
+ nice	Adj	Alpha	0.273962	300	266Y		12	0.12	0.034	0.07	0.163	-0.121	2056
+ price	Noun	Alpha	0.276756	272	257Y		13	0.073	0.119	0	0.199	-0.155	7926
+ recom...	Verb	Alpha	0.287107	241	234Y		14	0.145	0.058	-0.011	0.065	-0.031	715
thai	Adj	Alpha	0.308273	221	203Y		15	-0.071	0.118	0.019	0.165	0.003	7255
+ meal	Noun	Alpha	0.316546	213	192Y		16	0.009	0.083	-0.034	0.093	0.051	1524
+ menu	Noun	Alpha	0.317081	198	188Y		17	-0.026	0.074	-0.025	0.106	0.069	1751
+ visit	Verb	Alpha	0.320935	199	184Y		18	0.1	0.005	0.049	-0.086	0.208	4225
+ dinner	Noun	Alpha	0.323184	187	179Y		19	-0.036	-0.018	-0.01	0.155	0.079	4077
thai	Noun	Alpha	0.328184	195	175Y		20	-0.029	0.234	0.033	0.081	-0.045	3311
excellent	Adj	Alpha	0.343927	162	153Y		21	0.019	0.021	-0.018	0.114	0.026	3957
+ good	Noun	Alpha	0.346041	157	150Y		22	0.008	0.105	0.009	0.068	-0.004	5102
+ order	Verb	Alpha	0.346307	158	150Y		22	-0.005	0.193	0.001	-0.016	0.043	396
+ breakfa...	Noun	Alpha	0.35595	183	149Y		24	-0.094	-0.245	-0.03	0.361	0.05	3119
+ day	Noun	Alpha	0.353503	163	145Y		25	-0.038	-0.066	0.11	0.068	0.147	3458
phuket	Prop	Alpha	0.350887	149	144Y		26	0.006	-0.01	0.026	0.018	0.151	5348
+ tasty	Adj	Alpha	0.357171	141	137Y		27	0.004	0.047	0.004	0.09	0.018	113
+ little	Adj	Alpha	0.35781	139	136Y		28	0.089	0.031	0.101	0.003	-0.02	3759
+ drink	Noun	Alpha	0.362592	136	132Y		29	0.154	-0.006	-0.004	-0.006	0.003	4004
+ stay	Verb	Alpha	0.362034	132	131Y		30	-0.071	-0.044	-0.004	0.097	0.172	97
food	Adj	Alpha	0.366334	131	128Y		31	0.037	-0.019	-0.018	0.104	0.014	3371
+ night	Noun	Alpha	0.374358	135	123Y		32	0.078	-0.027	-0.015	-0.036	0.162	4608
+ look	Verb	Alpha	0.37441	132	122Y		33	0.046	0.094	0.012	-0.015	0.053	5834
atmosph...	Noun	Alpha	0.372051	121	121Y		34	0.138	-0.02	0.015	0.027	-0.022	7557
+ dish	Noun	Alpha	0.378275	126	119Y		35	-0.019	0.089	0.007	0.095	-0.013	1383
+ owner	Noun	Alpha	0.376552	121	118Y		36	0.091	-0.068	0.053	0.025	0.048	5201
lovely	Adj	Alpha	0.378275	126	118Y		36	0.072	0.004	0.096	0.039	-0.03	6466
good	Adv	Alpha	0.380462	122	116Y		38	-0.016	0.042	0	0.084	0.041	720
reasonab...	Adj	Alpha	0.383039	114	112Y		39	0.03	0.054	-0.007	0.159	-0.104	4622
+ love	Verb	Alpha	0.384631	114	111Y		40	0.053	-0.003	0.045	0.049	0.01	1106
+ find	Verb	Alpha	0.387962	119	110Y		41	0.005	0.052	0.029	0.028	0.057	6073
great	Adv	Alpha	0.388865	117	109Y		42	0.047	0.025	0.003	0.051	0.021	1523
+ enjoy	Verb	Alpha	0.388655	112	108Y		43	0.041	-0.011	-0.019	0.045	0.068	1037
+ hotel	Noun	Alpha	0.39166	116	107Y		44	-0.052	-0.039	0.014	0.102	0.11	5823
+ beach	Noun	Alpha	0.394063	119	106Y		45	0.005	-0.025	0.295	0.025	-0.04	2183
great	Noun	Alpha	0.391545	107	105Y		46	0.114	-0.034	-0.015	0.026	0.028	7571
always	Adv	Alpha	0.39944	122	104Y		47	0.094	-0.055	-0.034	0.018	0.106	3501
delicious	Adj	Alpha	0.39408	105	103Y		48	0.03	0.049	0.016	0.058	-0.002	5460
highly	Adv	Alpha	0.394903	103	102Y		49	0.104	0.016	-0.009	0.053	-0.038	3687
+ small	Adj	Alpha	0.399333	108	101Y		50	0.011	0.096	0.104	-0.005	0.001	4191

Figure 30: showing the topics term table.



Terms

Topic Weight	+	Term	Role	# Docs	Freq
0.338	+	great	Adj	366	444
0.296	+	staff	Noun	454	471
0.286		friendly	Adj	361	368
0.248	+	bar	Noun	97	112
0.154	+	drink	Noun	132	136
0.11		great food	Noun Group	38	39
0.103	+	good	Adj	613	722
0.08	+	friend	Noun	79	82
0.077	+	people	Noun	67	69
0.074	+	food	Noun	1120	1323
0.071		back	Adv	59	61
0.049		well	Adj	47	48
0.032	+	keep	Verb	19	19
0.025		well	Verb	10	10
0.004	+	restaurant	Noun	614	724

Documents

Topic Weight	Review	Hotel
0.37	Hi again, well I went back in june for State of	Legends Sports Bar & Grill
0.369	Lovely place_x000D__x000D_	Oasis Bar
0.369	Lovely place_x000D__x000D_	Oasis Bar
0.322	Great Bar, spent many afternoons here with	Legends Sports Bar & Grill
0.32	Beat bar in patong. ...great staff and beer is	Oasis Bar
0.315	Awesome little bar! Great burgers, friendly	Legends Sports Bar & Grill
0.312	we ate at dee-lish many nights on our 3 week	Dee-Lish Bar & Restaurant
0.303	Thanks for a very entertaining evening. Glad i	Oasis Bar
0.298	Great vibe, good music and very friendly staff.	Oasis Bar
0.298	Great vibe, good music and very friendly staff.	Oasis Bar
0.297	This truly is the best bar in town. The owner	Oasis Bar
0.295	Came here with my boyfriend to watch the	Legends Sports Bar & Grill
0.295	this place has a great atmosphere, very	Legends Sports Bar & Grill
0.291	This is the best most friendly and welcoming	Oasis Bar
0.289	Great food and very friendly staff, even	O-OH Farm Ta-Eiad
0.289	Best bar in Patong by far. Great staff, fantastic	Oasis Bar
0.288	Been here many times for dinner; we love	Dee-Lish Bar & Restaurant
0.288	Really nice place. Good music, friendly staff	Oasis Bar
0.288	My son Michael and his friend Josh hung out at	Oasis Bar

Figure 31: showing interactive topics viewer of review of different hotels with their topic weight.

Step 7: Drag and dropped the text cluster node to check the cluster of word that mostly occurred together, edit the number of clusters to 5(not to generate too many clusters), connect it to the text topic node and run. But 4 clusters were generated with its output.

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

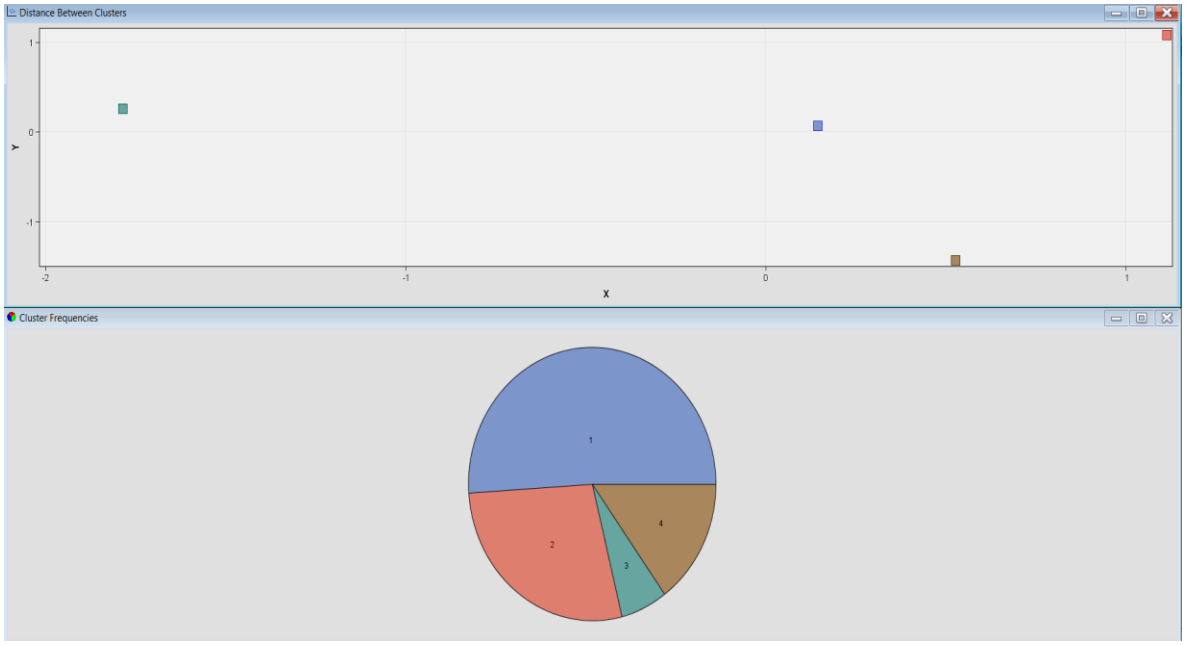


Figure 32: showing clusters and its distance.

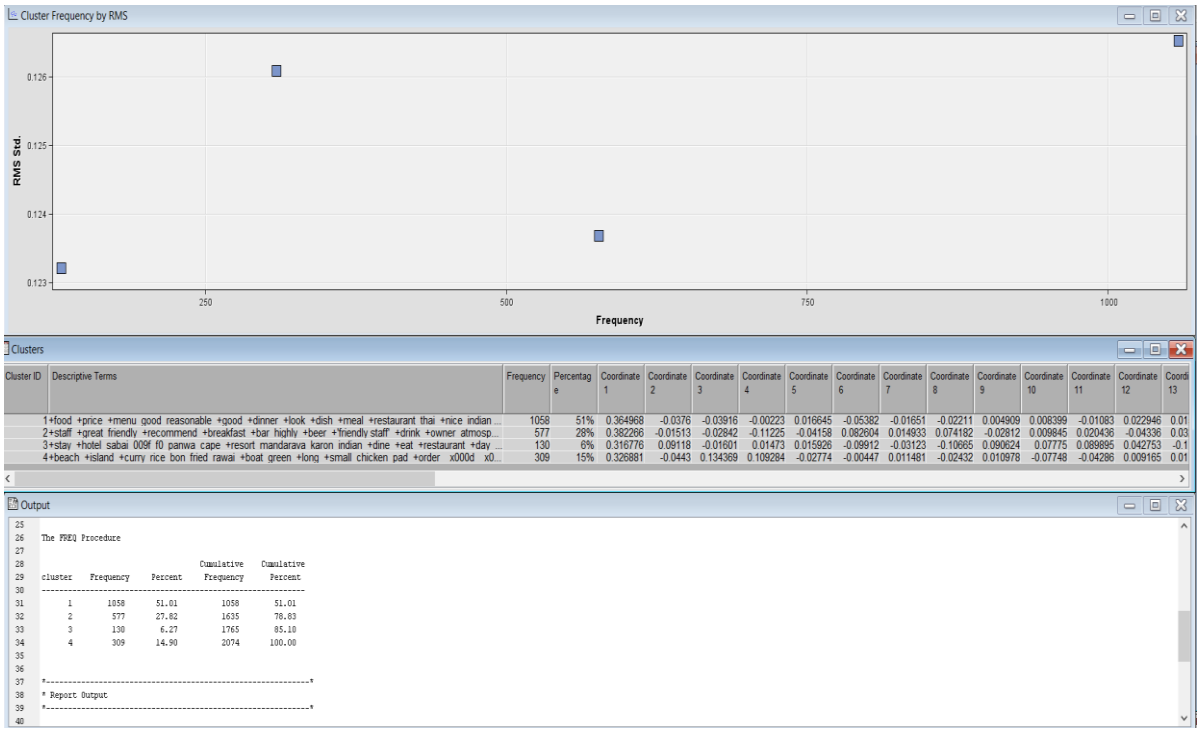


Figure 33: showing clusters frequency by RMS, cluster table, and the cluster output

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1714442

APPENDIX

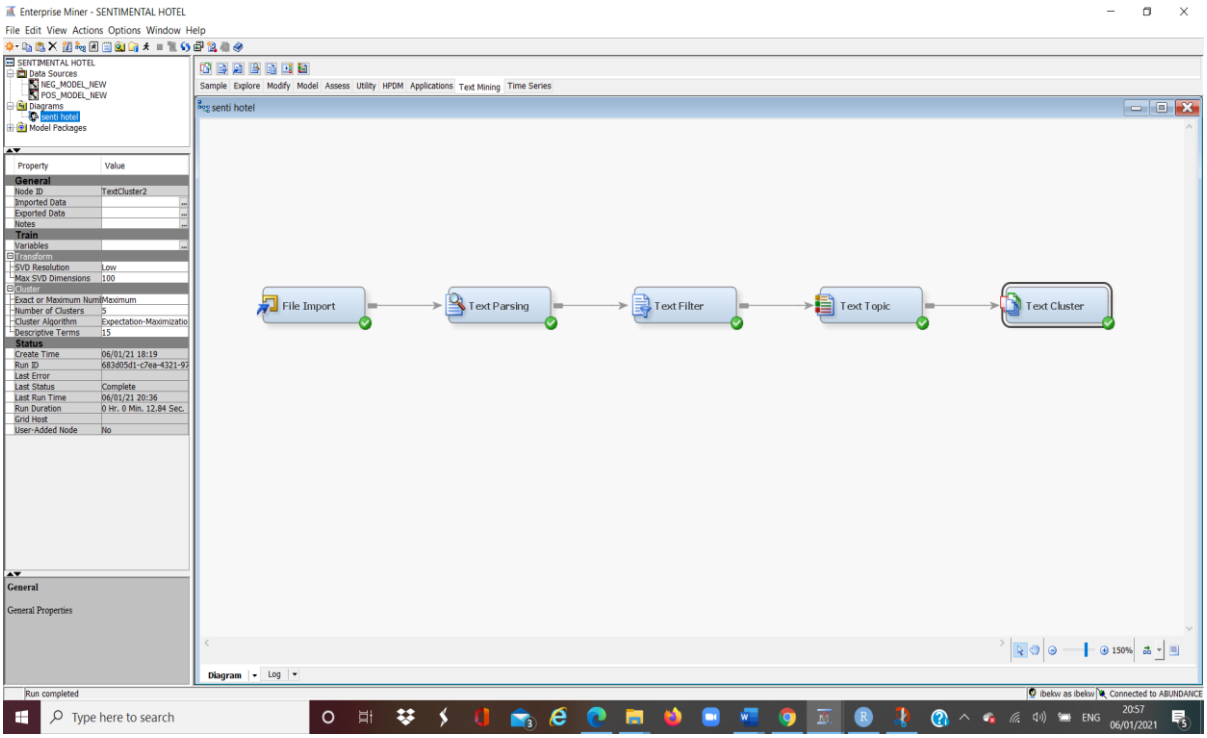


Figure 34: Screen shot of the Enterprise Miner on hotel_reviews.xlsx in SAS

VI. SUMMARY ON SAS ENTERPRISE MINER

This is summary on how the experiment was done in SaS enterprise miner software. The enterprise miner software was launched after which a project was created. Then the cleaned data set was imported by dragging and dropping from the import node from sample menu bar of the software then data exploration was done on the data set, on the frequency review of the hotel shown in figure 22 while figure 23 shows Sample report of the dataset and figure 24 shows Sample statistics of the dataset and figure 25 shows the view of dataset showing the first 55 rows. Editing the variable to select the variables, there is need to set the following specification of role like (review= text, hotel = ID) and then dragging and dropping the text parsing node which parsed the text and showed some statistical visualization of the text component the figure 26 shows the output of text parsing node while figure 27 shows the output of text filter node which consist of role by frequency plot, attribute by frequency plot, terms table (colored by dropped$blue and keep$red words). Performing the statistical visualization, text topic node was dragged and dropped and edited the number of multi-term topic to FIVE and runwhich produced the following figure 28 shows the bar plot of the result after run. Figure 30 shows the topics term table, figure 31 shows interactive topics viewer of review of different hotels with their topic weight then checking for the cluster word occurred mostly together after which, 5 clusters are generated with its output shown in figure 32 and figure 33.

VII. RESULT COMPARISM BETWEEN R AND SAS (sentimental) CRITICAL FINDINGS

R	SAS
Wordcluster in R was done earlier than in SAS	Wordcluster in Sas is done towards at the last of the experiment.
When the wordcloud and percentage output for 5 hotels were generated, it shows sabai-sabai hotel	wordcluster of 5 number of clusters produced a result: When frequency is 1050 when the cluster is 1, with cumulative precast of 51.01%.

with 71 positive, negative 45 with a percentage of positive review of 61.21%. Audy restaurant with review of positive – 54 negative – 24 percentage of positive 69.23% B-Hive Gallery Bar and Restaurant positive – 47 negative – 35 percentage of positive 57.32% DaVinci Restaurant positive – 54 negative – 25 percentage of positive 68.35% Istanbul Turkish Restaurant positive – 50 negative – 16 percentage of positive 75.76%	When frequency is 577 then cluster is 2 gave a cumulative precast of 78.83% When frequency is 130 then cluster is 3 produced a cumulative precast of 65.10% When frequency is 309 and cluster is 4 gave a cumulative precast of 100.00%.

REFERENCE

[1] Castellucci, G., et al.(2014) Unitor: Aspect based sentiment analysis with structured learning. in Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014). 2014.
[2] Hu, M. and B. Liu (2004) Mining opinion features in customer reviews. in AAAI.
[3] Kiritchenko, S., et al(2014) NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. in Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014).
[4] Renuka S. Deshmukh, Dr. Aashish A. Bardekar(2020) EVOLUTION OF SENTIMENTAL ANALYSIS, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072
[5] Liu, B (2012), Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 2012. 5(1): p. 1-167.
[6] Jakob, N. and I. Gurevych (2010) Extracting opinion targets in a single-and cross-domain setting with conditional random fields. in Proceedings of the 2010 conference on empirical methods in natural language processing. 2010. Association for Computational Linguistics.