Insights from sentiment analysis to leverage local tourism business in restaurants

Ting Yu (Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal)

Paulo Rita (NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal)

Sérgio Moro (Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal)

Cristina Oliveira (Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal)

This is the accepted author *manuscript of the following article published by Elsevier:* Yu, T., Rita, P., Moro, S., & Oliveira, C. (2021). Insights from sentiment analysis to leverage local tourism business in restaurants. International Journal of Culture, Tourism, and Hospitality Research. [Advanced online publication on 23 November 2021]. https://doi.org/10.1108/IJCTHR-02-2021-0037

Funding Information:

The work by S. Moro and C.Oliveira was partially funded by national funds through FCT - Fundação para a Ciência e Tecnologia, I.P., under the project FCT UIDB/04466/2020. The work by P. Rita was partially funded by national funds through FCT - Fundação para a Ciência e Tecnologia, I.P., under the project FCT UIDB/04152/2020 - Centro de Investigação em Gestão de Informação (MagIC).

Insights from sentiment analysis to leverage local tourism business in restaurants

Abstract

Purpose

Social media has become a main venue for users to express their opinions and feelings, generating a vast number of available and valuable data to be scrutinized by researchers and marketers. This paper extends previous studies analyzing social media reviews through text mining and sentiment analysis in order to provide useful recommendations for management in the restaurant industry.

Design/methodology/approach

The Lexalytics, a text mining artificial intelligence tool, is applied to analyze the text of the online reviews of the restaurants in a touristic Dutch village extracted from the most frequently used social media platforms focusing on the four restaurant quality factors, namely food and beverage, service, atmosphere and value.

Findings

Findings of this research are presented by the identified key themes with comparisons of the customers' review sentiment between a selected restaurant, Zwaantje, vis-à-vis its benchmark restaurants set by a specific approach under the abovementioned quality dimensions, in which the F&B and service are the most commented by customers. Results demonstrate that text mining can generate insights from different aspects and that the proposed approach is valuable to restaurant management.

Originality

The paper provides a relatively big scale in numbers and resources of social media reviews to further explore the most important service dimensions in the restaurant industry in a specific tourist area. It also offers a useful framework to apply the text mining business intelligence tool by comparison of peers for local small business restaurant practitioners to improve their management skills beyond manually reading social media reviews.

Keywords: Restaurant Business; Social Media; Online Reviews; Text Mining; Sentiment Classification; Lexalytics; Giethoorn.

1 Introduction

Social media collects 7.6 billion people in the world and 53 percent of them are active social media users (Shaw, 2018) who have been cultivated and encouraged to share their purchase experiences through online reviews, which have been taken as references for product and service purchases. Applying analytics get new possibilities from the data sources generated by the massive growth of User-Generated Content (UGC) (Calheiros *et al*., 2017; Liu *et al*.,2017; Oliveira *et al.,* 2019), which makes it more important to explore how to apply UGC to serve for the development of businesses.

To better understand UGC and transfer the available online text reviews into valuable information, text mining and sentiment analysis through business intelligence tools have been developed rapidly in the last few years (Moro *et al*., 2019). Lexalytics, a natural language processing tool, was applied to analyze customers' reviews from the primary social media platforms Google Maps Restaurants Review, Facebook and TripAdvisor, with a focus on a sample of all the restaurants in Giethoorn, a Dutch tourist village. While analysing online reviews has become quite common to both scholars and practitioners, most of the research within hospitality has been conducted in hotels (D'Acunto et al., 2020: Osman et al., 2019). In comparison, restaurants have received less attention. Furthermore, most of that attention is turned to large restaurants, chains, or online reservations (e.g., Li et al., 2020). With our study, we intend to address such research gap by focusing on a small restaurant that aims to survive and even to excel in face of competition in a locally known tourism destination. Restaurant Zwaantje was chosen to compare with the benchmark restaurants in the area, to explore how a family running restaurant could take advantage from the available social media reviews and text mining tools generating sentiment analysis to support managerial decisions.

Restaurant managers should understand and meet consumers' needs, wants, and demands to succeed in this competitive industry (Gregoire *et al*., 1995). Text mining and sentiment analysis are one of the most frequently approaches applied to identify the meaningful information behind the loads of UGC in social media (Guerreiro and Rita, 2020). This study applied the text mining and sentiment analysis approach using Lexalytics to provide the relevant themes from UGC under the dimensions of food and beverage (F&B), service, atmosphere and value in a specific tourism area with 28 restaurants meeting research requirements in order to extend previous research in the restaurant industry.

The present study also aims to analyse the reviews through a scientific method to understand what are customers' experiences and expectations, and to assist restaurant managers to improve performance in different service dimensions. TripAdvisor, Google Maps Restaurants Review and Facebook are the most used social media platforms by the tourists in Giethoorn (Ting, 2020), where UGC is categorized into four service dimensions, namely F&B, Service, Atmosphere and Value to be analyzed in this study to answer the following research questions:

RQ1: Among the F&B, Service, Atmosphere and Value dimensions, which are the ones more commented by Giethoorn restaurants' consumers?

RQ2: Compared with the benchmark restaurants in Giethoorn under the abovementioned four service dimensions, what are the key themes from UGC that the manager of a restaurant (in this case, Zwaantje) should pay attention to?

This study performed a text mining analysis of UGC from the most used social media platforms through Lexalytics, to discover the sentiment of customer reviews, setting up the benchmark restaurants and allocating the key themes that may affect customers' satisfaction. By comparing the targeted restaurant Zwaantje and the benchmark restaurants in terms of sentiment themes, several recommendations were generated to address restaurant management. Furthermore, this research also adds deepness to the discussion about restaurant quality factors perceived by customers that may lead to their behavioural intentions. Findings of this study can help restaurant operators to extract valuable information from UGC and then converted it to competitive intelligence and actionable decisions. It also provides a useful framework for future research on the analysis of UGC in the restaurant sector.

2 Literature Review

Nowadays, there is a large quantity of data available for the general public which has been accelerated by an enhancing popularity of social media and the prevalence of UGC (Nave *et al.*, 2018). UGC refers to content created by consumers regarding a brand or experience with the aim of supporting other consumers in the process of consumption (Smith *et al.*, 2012). As a consequence of the increasing social media usage as a marketing communication channel, consumers' voices have become more important. Electronic word-of-mouth (eWOM) makes an essential element of the tourism sector as it reflects consumers' attitudes, intentions and experiences (Doosti *et al.*, 2016). Among consumers characterised by little expertise in a product category (Gilly *et al.*, 1998), who perceive decision making as high risk (Bansal and Voyer, 2000), or who are deeply involved in the purchasing decision (Beatty and Smith, 1987) and much more focused on listening others' opinions for product or service advice, makes eWOM an important tool for marketers to deliver product and service information to consumers (Levy and Gvili, 2015).

Over the years, reviewing via social media has become one of the most important reference sources before customers make their purchasing decisions (Cheung and Thadani, 2012). In line with a survey conducted by Channel Advisor, 90 percent of online consumers take reviews into consideration, and 83 percent of them sense the influence of online reviews on the actual decisions when purchasing (Nittala, 2015). As highlighted by Ryu *et al.*, (2012), consumer behaviour is influenced by attributes such as environment, food, services and value. The overall image of a restaurant affects the perceived value that a customer shares about it, which in return determines the satisfaction of customers, illustrating a predictor of consumer behaviour (Ryu *et al.*, 2012). The increase in the rating of online reviews supported it has a positive influence on the sales and the price of hotels (Lu *et al.*, 2013, Moro *et al.*, 2018; Öğüt and Onur Taş, 2012).

Studies have identified meal experience, food and beverages, service, atmosphere and price as well as value as decisive for consumers when selecting a restaurant (Hansen *et al.*, 2005; Jensen and Hansen, 2007). One of the first instruments introduced for dining experience refers to SEVQUAL with 22 items across five service dimensions: tangibles, reliability, responsiveness, assurance and empathy (Parasuraman *et al.*, 1988), while DINESERV shares the same dimension structure but across 29 items (Stevens *et al.*, 1995). Raajpoot (2002) developed another instrument TANGSERV to cover ambience/social factors, layout/design factors and product/service factors. Later, Antun *et al.* (2010) proposed DinEX including five attributes: food, service, atmosphere, food healthfulness and social. Bufquin *et al.* (2017) applied the instrument for research and found that food, service, atmosphere and value were the most important attributes. Pekar and Ou (2008) conducted a study in sentiment analysis on the pre-defined dimensions such as food, room, service, facilities and price by extracting 268 reviews from social media platform epinions.com, proposing a method to recognize the relationships between subjective expressions and references to features of hotel products. Moreover, previous studies showed that the most frequently mentioned word in both positive and negative reviews was "food" (Bilgihan *et al.*, 2017), which is also shared by Namkung and Jang (2007), indicating that food quality has been perceived as a vital component in customers' satisfaction. According to Bilgihan *et al.* (2017), the service quality or value of money may taint the customer's experience and generate negative comments. Our research applies F&B, service, atmosphere and value aligned with some factors from DinEX.

Table 1 summarizes a list of studies that have addressed cases of small restaurant businesses by analyzing online reviews written in social media in search for knowledge that leverages restauranteurs' decision making. As it can be observed, specific case studies on this domain are rather scarce and are usually focused on helping managers to understand their own clients through online reviews (e.g., Fernandes et al., 2021) instead of taking a broader approach by using automated techniques such as sentiment analysis to compare themselves with the competition. Such gap justifies the contribution of the present study.

3 Materials and Methods

3.1 Case Description

Giethoorn is a Dutch tourist village with 2,624 inhabitants and 2,032 hectares, of which 152 hectars is lake (Oozo.nl, 2018). Fairy-tale houses have been built alongside the lake and all transportation is done by boat. Therefore, it gained the reputation as the "North Venice". It reached recently 1.5 million visitors annually (Valkeman, 2018). There are 29 restaurants in the village providing customers with many choices of international food and a lovely environment. The restaurant Zwaantje has been run by the family since 1992 with restaurant and boat rental business, located in the entrance of Giethoorn alongside the river, providing Dutch cuisine with an average pricing strategy.

3.2 Data Collection

Twenty-nine restaurants of Giethoorn were selected based on TripAdvisor, Google Maps Restaurants Review and Facebook under F&B, Service, Atmosphere and Value. A total of 11,140 reviews were initially identified: 6,245 reviews from Google, 2,745 reviews from Facebook and 2,150 reviews from TripAdvisor by the end of August of 2018. Since we needed to assure that sufficient online reviews existed for the analysis, we selected only restaurants containing at least 20 reviews for each selected data source platform. As a result, one restaurant was excluded, leaving a total of 28 restaurants which were analysed. Then, for each 28 restaurants and for each platform, we selected the 100 most recent reviews. However, if a pair <restaurant,platform> contained less than 100 (but equal or more than 20, according to previous selection criterion), we selected all reviews. Such approach enables to assure balancing between the selected restaurants, by choosing restaurants that have at least been reviewed 20 times while avoiding an over representativeness by popular restaurants that have been often reviewed. In total, out of 11,140 available reviews, 4,832 were chosen for analysis in this research.

Lexalytics was applied as research supporting tool since it is considered a good tool to perform sentiment analysis (Lak & Turetken, 2014).

3.3 Proposed Approach

Figure 1 illustrates the proposed way for extracting useful knowledge from the unstructured text reviews from TripAdvisor, Google Maps Restaurants Review and Facebook, under the four quality factors: F&B, Service, Atmosphere and Value. Macroscopically, there were three parts in the approach: Input, Procedure and Output.

In the Input part, the reviews were collected from the selected social media sites. Different languages were translated into English by Google Translate which has already been found a useful tool when using bag-of-words text models by showing a considerable overlap in the set of features generated from human-translated and machine-translated texts and portraying high similarity with only small differences across languages (De Vries et al., 2018). Moreover, Hampshire and Salvia (2010) showed that Google Translate is the top-tier when compared to other online machine translating services. In addition, Schwarz et al. (2017) used Google Translate in the multilingual Swiss context to estimate intra-party preferences by comparing speeches to votes. The Procedure was executed manually and by Lexalytics, in which the categorization of the original reviews into the four service dimensions, was done manually as shown in Table 2. Unlike some previous studies (Zahoor et al., 2020), we did not train a classifier to perform sentiment classification. Instead, we adopted Lexalytics, which is based on an already trained classifier model. This is a similar approach to other studies in hospitality and tourism (e.g., Piccinelli et al., 2021). We conducted aspect-based analysis by identifying four categories, namely food & beverage, service, atmosphere, and value. Then, for each sentence related to each aspect, we directly used Lexalytics to compute sentiment classification. Then, Lexalytics was used to break the categorized unstructured reviews into corpus and came up with sentiment themes and clouds in the end. One should stress that the benchmark restaurants setting up was executed by the classification of sentiment words. The Output part was based on the comparison between the benchmarks and restaurant Zwaantje on key sentiment themes.

3.4 Text Mining and Sentiment Analysis

Text mining is the process of transforming unstructured text data into meaningful and actionable information by utilizing different artificial intelligence (AI) technologies to automatically process data and generate valuable insights (Hung and Zhang, 2012). Social media platforms have vast UGC, which are generally presented in digital attributes and open unstructured text features that are freely expressed by users. The digital attributes include rating valence, emotional polarity, number of online user reviews and score variance, while the unstructured text contains a wealth of information such as the users' interests, preferences, intentions, perceptions, emotions and attitudes towards products or services. Sentiment analysis is the determining process if the text is positive, neutral or negative (Lexalytics, 2019), which is widely used on UGC in more recent studies (Colace *et al.*, 2015).

The sentiment analysis approach must guarantee a fast and relevant accurate analysis of UGC. As a consequence, Lexalytics, where words are annotated with respect to their sentiment score, was applied. Lexalytics is an on-premises and multi-lingual text analysis engine to provide sentiment analysis from unstructured text and come up with the readable information via Natural Language Processing (NLP) and text mining (Lexalytics, 2019). Lexalytics analysis procedure includes: (1) Breaking each text document down into its component parts (sentences, phrases, tokens and parts of speech); (2) Identifying each sentiment-bearing phrase and component; (3) Assigning a sentiment score to each phrase and component (-1 to +1); (4) Combining scores for multi-layered sentiment analysis, if necessary. The whole process of Lexalytics text analytics technology and the NLP feature stack shows the layers of processing each text document goes through to be transformed into structured data (Lexalytics, 2019).

3.5 Text Mining for Sentiment Classification

To ensure that all the reviews were analyzed in a more accurate way, each review text was divided into F&B, Service, Atmosphere and Value and the unrelated content deleted manually. Table 3 illustrates that in 4,832 reviews, there were 3,049 reviews under F&B, 2,615 reviews under Service, 1,674 reviews under Atmosphere and 596 reviews under Value.

The sentiment words and themes selected for text mining combine NLP and machine learning techniques to assign weighted sentiment scores within a sentence or phrase. As Table 4 illustrates, Lexalytics first highlights the text, then picks the word out, makes scores and comes up with the sentiment valuation. NPL aims at answering four questions: Who is talking? What are they talking about? How do they feel? Why do they feel that way? This last question is a question of context. Context analysis in NLP involves breaking down sentences to extract the n-grams, noun phrases, themes, and facets present within. The foundation of context determination is the noun. N-grams are combinations of one or more words that represent entities, phrases, concepts, and themes that appear in text. Generally speaking, the lower the value of "n", the more general the phrase or entity (Lexalytics, 2019). For example, "wonderful meals over there, the cook is very fantastic", the words extracted via N-gram is fantastic, the word intensifier is "very". The theme in the text is "cook". Even though the "wonderful" and "meals" are not extracted, this review text is already categorized into F&B by the author, the word "fantastic" extracted has expressed the sentiment under F&B, which decreases the uncomplexity of the detection of the NLP.

3.6 Sentiment Words Classification and Method for Setting the Benchmark Restaurants

Lexalytics recognized the words from the targeted documents first and came up with the sentiment classification in terms of negative, neutral and positive sentiment and grand total from the whole targeted reviews defined by the four categories and specific restaurant from the three selected social media sites as shown in Table 5.

Benchmark restaurants were set by the percentage of the sum of the neutral and positive sentiment out of the Grand Total{(Neutral + Positive)/Grand Total}. To make the research to get maximum useful information, only the restaurants equal to and more than ten Grand Total and the first three restaurants with most Grand Total were selected as the benchmarks. Table 6 illustrates how the benchmark restaurants were set under F&B in TripAdvisor, by ranking the restaurants from the largest percentage to the smallest. The first ten restaurants with the largest percentage were considered, among which the first three with the largest Grand Total were chosen to be the benchmark restaurants. Hence, Lindenhof with 100 Grand Total, Grachthof with 75 Grand Total and Achterhuus with 72 Grand Total were the benchmarks under F&B in TripAdvisor.

4 Results

Thirty-five restaurants were set as the benchmark restaurants, among which 9 restaurants were chosen under F&B, Service and Atmosphere respectively, while 8 restaurants met the requirement under Value as shown in Table 7.

The themes are divided by key terms with similar terms and words that help Lexalytics to identify and compute sentiment scores, such as adjectives. The frequencies of the theme, the average sentiment scores and the sentiment classification, in which negative are sentiment scores below -0.5, neutral are sentiment scores from -0.5 to 0.5 and positive are sentiment sores above 0.5 are also presented.

The key terms of Zwaantje and Benchmarks under F&B are shown in Figure 2, in which food got the most frequencies, Zwaantje got neutral sentiments overall, in which except pancakes got positive sentiment with score 0.52, restaurant, chocolate milk, cup, menu and cafe obtained neutral sentiment. All the key terms from benchmark restaurants, including wines, quality, appetizers and delicious, generated positive sentiments.

The key terms of Zwaantje and Benchmarks under Service are presented in Figure 3, in which staff and service got the most frequencies, Zwaantje got negative to neutral sentiment, while benchmark restaurants got positive sentiment, which gives a very clear view on the existing gap between Zwaantje and Benchmarks.

The key terms of Zwaantje and Benchmarks under Atmosphere are shown in Figure 4. Key terms in Zwaantje such as place with the most frequencies, restaurant, décor, location got positive sentiment, while ambiance got neutral and toilets got negative. The key terms in the benchmark restaurants were place, location, terrace, atmosphere, view and restaurant, which all got positive sentiment.

The key terms of Zwaantje and Benchmarks under Value are presented in Figure 5. This dimension is less mentioned in the reviews from Zwaantje. However, from the benchmark restaurants, price got most frequencies and price quality ratio followed, both obtained neutral sentiments, which raises the question that in this specific touristic place the Value is not a main dimension generating the sentiments from customers.

5 Discussion

RQ1: Among the F&B, Service, Atmosphere and Value dimensions, which are more commented by Giethoorn restaurants' consumers?

Bougoure and Neu (2010) found that tangibles such as physical facilities, equipment and appearance of personnel were perceived to be important restaurant attributes in fast food restaurants. Qin and Prybutok (2008) found that a sympathetic and reassuring service with quick operation is preferred for students. In turn, Bufquin *et al*. (2015) found that food, service and atmosphere were the most important attributes in American casual-dining restaurant. The most important factors were also various by geography, culture and dining intention, among others. The study by Tripathi and Dave (2017) explored the underlying key dimensions of service quality in Indian restaurant industry, culture orientation was extracted along with ambient settings, empathy, privacy and entertainment, reliability and responsiveness by their high significances related to the service quality dimensions. The research on the success factors on the customer perception on the fast food restaurant industry in Greece, conducted by Mamalis (2009), had presented the factors such as adaption to locality, food quality, service, facilities, place to be and sales incentive programme. Polyorat and Sophonsiri (2010) conducted the research of the chain restaurants in Thailand and found that the tangibles and empathy under SERVQUAL were most relevant to locals, which was influenced by culture of the individualism-collectivism and masculinity-femininity. In this research, according to the percentages of F&B, Service, Atmosphere and Value out of all the reviews collected from TripAdvisor, Google and Facebook, the F&B (38%) and Service (33%) were the categories that collected more comments while Atmosphere (21%) was mentioned less and Value (8%) the least. This study explored further in a specific tourism area, what were the most important attributes for customers or tourists, reinforcing the existing literature which found that food and service were the most important factors.

RQ2: Compared with the benchmark restaurants in Giethoorn under the above mentioned service dimensions, what are the key themes that the manager of restaurant Zwaantje should pay attention to?

Figure 2 illustrates the key terms with frequencies and sentiment scores for both Zwaantje and Benchmarks. Food, wine, quality and suite were extracted from benchmarks with positive sentiment, which referred to food taste and quality, matching drinks and food alignment. Moreover, authentic and Dutch style were mentioned as terms, pancake and coffee were extracted more than once with neutral sentiment, which can provide customers' perceptions that Zwaantje is a Dutch authentic restaurant with tasty pancake and can make improvements when establishing comparisons according to the key phrases generating positive sentiments of the benchmark restaurants.

The main service key terms extracted from Zwaantje were service, staff, operation and attention, in which staff got negative sentiment, while for benchmark restaurants were service and staff with positive sentiments (Figure 3). The negative sentiment modifiers of Zwaantje's staff were unfriendly, arrogant, disrespectful, rude and discriminating. However, the benchmark restaurants only got friendly and nice. The operation speed and smoothness were also the factors influencing service experience.

The negative themes from Zwaantje under atmosphere was dirty toilets. Place, restaurant, décor and location got positive sentiment while ambience got neutral, and according to customers' perceptions the following should be improved: cosy, authentic, good location and environment, great décor, nostalgic design, somehow historic, home style, wooden house and furniture (Figure 4). Generally, customers were satisfied with the atmosphere of Zwaantje, and the cleanness shall be stressed by the restaurant manager. The benchmarks' terms can be summarized into four aspects: place, location, terrace, atmosphere and view.

As showed in Figure 5, customer experience on value could be strengthened by improving the quality of the food and beverage, adjusting the portion reasonably and proceed with reasonable pricing aligning with the key themes found, namely price, price quality ratio, quantity and quality.

6 Conclusions

The goal of this research, as shown in Figure 6, was to provide a scalable sentiment analysis process in a specific unit in order to explore the area of restaurant service management. This endeavour was performed through the collection and categorization of unstructured data from several social media sites aiming to understand customer perceptions and feelings. A text mining approach was used to find hidden information and also filling the void in a tourist destination (Giethoorn, The Netherlands) restaurant industry by applying a scientific approach to provide suggestions on restaurant management.

This research makes contributions to extend current limited empirical studies available on applying business intelligence tools to analyse social media reviews via a text mining approach in the restaurant industry. More specifically, this study provides research outcomes based on a relatively big volume of data in the context of a small but intensive tourist area in the Netherlands. The value of our contribution is especially relevant for small business restaurant managers who need to deal with competition but who usually do not have the required skills to analyse data using automated tools and instead rely on manually reading a sample of reviews. With our research, we show that existing tools can be easily used to extract valuable insights. This research also offers a more comprehensive understanding of the key terms that customers emphasize to dine at a restaurant by empirically identifying the factors through analysing social media reviews by comparison with peers. Additionally, the research model offered a framework for future researcher to conduct an empirical research on the similar theme in the restaurant industry.

This research has extended the previous studies by analysing social media reviews under F&B, Service, Atmosphere and Value, and found that the qualities of F&B and Service were mentioned more than the other factors in the studied tourist village in the Netherlands.

Management of Zwaantje restaurant could improve decision making based on the extracted themes to make customers achieve more satisfied experiences and leave positive reviews in social media in order to generate more sales. This study also found that staff behavior, operation and attention were the key themes under Service. Zwaantje performing quite well under the Atmosphere dimension and the value of money was an overall experience that customers had and also depended on other dimensions. Social media serves as the communication and interaction platform between restaurant management and its customers, which is getting an increasingly valuable information. Applying the text mining NPL tool to analyse social media reviews about restaurants, with comparison among competitors, enables to transform raw data into valuable knowledge and support management decision to improve their competitiveness.

As this study selected all the restaurants meeting the set research standard in Giethoorn without distinguishing the restaurant formality spectrum, price, food type and so on, the results of this research may not be applied conclusively to compare the performances of all the peers. Since the text mining technology is still a work in progress, the word dictionary in the text mining programs still have room to be more adapted to the researchers' context. Existing intensifier and negators in the text document make it difficult to analyze the real sentiment. In addition, the variance of language of the online reviews suggests that more studies in this area are in need as English was not the main language. While we used Google translator, it is still a rather limited tool in properly translating between many languages to English.

The use of different platforms as data sources means that metadata information (e.g., date of visit, nationality of visitor), which is important to better characterize the travelers and their visits to the restaurants, is missing in many cases or is inconsistent. For example, we can know the date when the review was written, but not when the traveler visited the restaurant. Even in an information rich platform such as TripAdvisor, the visit date is not a mandatory field (therefore is missing in many cases). While using many sources has the advantage of providing the perspectives of visitors using different platforms, it ultimately leads to some inconsistency and prevents the dataset from being better characterized, which could otherwise strengthen the validity of the results. Another important limitation that should be stated is that we did not consider the quantitative ratings, which offer "excellent cues by providing a quick tone of a review" (Lak and Turetken, 2014; p. 796). However, Facebook still does not provide an individual quantitative rating, which means that opinions made on this important platform would be underrepresented, in comparison to TripAdvisor and Google, where a user can rate his/her review with a quantitative score. Such limitation prevented us from including ratings in our study.

Future research might also extend the focus on the similar types of restaurants, concentrating for instance on particular cuisine(s), price and atmosphere. Furthermore, future studies can conduct a comparative analysis among different cuisines or among different consumers from different regions. Future research might as well deepen the study to involve other restaurant service dimensions excluded in this research such as experience, perceptions and attitudes.

References

Antun, J.M., Frash Jr, R.E., Costen, W. and Runyan, R.C. (2010), "Accurately assessing expectations most important to restaurant patrons: The creation of the DinEX Scale", *Journal of Foodservice Business Research*, Vol.13 No. 4, pp.360-379. DOI: 10.1080/15378020.2010.524539

Bansal H.S. and Voyer PA (2000), "Word-of-mouth processes within a services purchase decision context", *Journal of Service Research*, Vol. 3 No. 2, pp.166-177. DOI: 10.1177/109467050032005

Beatty, S.E. and Smith, S.M. (1987), "External search effort: an investigation across several product categories", *Journal of Consumer Research*, Vol. 14 No.1, pp.83-95.

Bilgihan, A., Seo, S. and Choi, J. (2017), "Identifying restaurant satisfiers and dissatisfiers: suggestions from online reviews", *Journal of Hospitality Marketing & Management*, Vol. 27 No. 5, pp.601-625. DOI: 10.1080/19368623.2018.1396275

Bougoure, U.S. and Neu, M.K. (2010), "Service quality in the Malaysian fast food industry: An examination using DINESERV", *Services Marketing Quarterly*, Vol. 31 No. 2, pp.194-212. DOI: 10.1080/15332961003604360

Bufquin, D., DiPietro, R. and Partlow, C. (2017), "The influence of the DinEX service quality dimensions on casual-dining restaurant customers' satisfaction and behavioral intentions", *Journal of Foodservice Business Research*, Vol. 20 No. 5, pp.542-556. DOI: 10.1080/15378020.2016.1222744

Bufquin, D., Partlow, C.G. and DiPietro, R.B. (2015), "Measuring restaurant patrons' perceptions and expectations: An importance-performance analysis using the DinEX model", *Journal of Foodservice Business Research*, Vol. 18, pp.226-243. DOI: 10.1080/15378020.2015.1051430

Calheiros, A.C., Moro, S. and Rita, P. (2017), "Sentiment classification of consumer-generated online reviews using topic modeling", *Journal of Hospitality Marketing & Management*, Vol. 26 No. 7, pp.675-693. DOI: 10.1080/19368623.2017.1310075

Cheung, C. and Thadani, D. (2012), "The impact of electronic word-of-mouth communication: A literature analysis and integrative model", *Decision Support Systems*, Vol. 54 No.1, pp.461-470. DOI: 10.1016/j.dss.2012.06.008

Colace, F., Casaburi, L., De Santo, M. and Greco, L. (2015), "Sentiment detection in social networks and in collaborative learning environments", *Computers in Human Behavior*, Vol.51, pp.1061-1067. DOI: 10.1016/j.chb.2014.11.090

D'Acunto, D., Tuan, A., Dalli, D., Viglia, G. and Okumus, F. (2020), "Do consumers care about CSR in their online reviews? An empirical analysis", *International Journal of Hospitality Management*, Vol. 85, pp.102-342. DOI: 10.1016/j.ijhm.2019.102342

De Vries, E., Schoonvelde, M. and Schumacher, G. (2018), "No longer lost in translation: evidence that Google Translate works for comparative bag-of-words text applications", *Political Analysis*, Vol. 26 No. 4, pp.417-430. DOI: 10.1017/pan.2018.26

Doosti, S., Jalilvand, M., Asadi, A., Khazaei Pool, J. and Mehrani Adl, P. (2016), "Analyzing the influence of electronic word of mouth on visit intention: the mediating role of tourists' attitude and city image", *International Journal of Tourism Cities*, Vol. 2 No. 2, pp.137-148. DOI: 10.1108/IJTC-12-2015-0031

Fernandes, E., Moro, S., Cortez, P., Batista, F. and Ribeiro, R. (2021), "A data-driven approach to measure restaurant performance by combining online reviews with historical sales data", *International Journal of Hospitality Management*, Vol. 94, pp.102-830. DOI: 10.1016/j.ijhm.2020.102830

Guerreiro, J. and Rita, P. (2020), "How to predict explicit recommendations in online reviews using text mining and sentiment analysis", *Journal of Hospitality and Tourism Management*, Vol. 43, pp.269-272. DOI: 10.1016/j.jhtm.2019.07.001

Gilly, M.C., Graham, J.L., Wolfinbarger, M.F. et al. (1998), "A dyadic study of interpersonal information search", *Journal of the Academy of Marketing Science*, Vol. 26, pp.83-100. DOI: 10.1177/0092070398262001

Gregoire, M.B., Shanklin, C.W. and Greathouse, K.R. (1995), "Factors influencing restaurant selection by travelers who stop at visitor information centers", *Journal of Travel & Tourism Marketing*, Vol. 4 No.2, pp.41-49. DOI: 10.1300/J073v04n02_03

Hampshire, S. and Salvia, C. P. (2010), "Translation and the Internet: evaluating the quality of free online machine translators", *Quaderns: Revista de Traducció*, Vol.17, pp.197–209.

Hansen, K.V., Jensen, Ø. and Gustafsson, I.B. (2005), "The meal experiences of á la carte restaurant customers", *Scandinavian Journal of Hospitality and Tourism*, Vol.5 No.2, pp.135-151. DOI: 10.1080/15022250510014417

Hlee, S., Lee, J., Yang, S. and Koo, C. (2019), "The moderating effect of restaurant type on hedonic versus utilitarian review evaluations", *International Journal of Hospitality Management*, Vol.77, pp.195-206. DOI: 10.1016/j.ijhm.2018.06.030

Hung, J. and Zhang, K. (2011), "Examining mobile learning trends 2003–2008: a categorical meta-trend analysis using text mining techniques", *Journal of Computing in Higher Education*, Vol.24 No.1, pp.1-17. DOI: 10.1007/s12528-011-9044-9

Jensen, Ø. and Hansen, K. (2007), "Consumer values among restaurant customers", *International Journal of Hospitality Management*, Vol.26 No.3, pp.603-622. DOI: 10.1016/j.ijhm.2006.05.004

Lak, P. and Turetken, O. (2014), "Star ratings versus sentiment analysis - a comparison of explicit and implicit measures of opinions", in *47th Hawaii International Conference on System Sciences*, January, pp. 796-805. IEEE. DOI: 10.1109/HICSS.2014.106

Levy, S. and Gvili, Y. (2015), "How credible is e-word of mouth across digital-marketing channels? the roles of social capital, information richness, and interactivity." *Journal of Advertising Research*, Vol.55 No.1, pp.95-109. DOI: 10.2501/JAR-55-1-095-109

Lexalytics (2019), "Sentiment analysis explained", Lexalytics, available at: https://www.lexalytics.com/technology/sentiment-analysis (accessed 20 June 2019)

Li, H., Xie, K.L. and Zhang, Z. (2020), "The effects of consumer experience and disconfirmation on the timing of online review: field evidence from the restaurant business", *International Journal of Hospitality Management*, Vol.84, pp.102-344. DOI: 10.1016/j.ijhm.2019.102344

Liu, X., Burns, A.C. and Hou, Y. (2017), "An investigation of brand-related user-generated content on Twitter", *Journal of Advertising*, Vol.46 No.2 , pp.236-247. DOI: 10.1080/00913367.2017.1297273

Lu, X., Ba, S., Huang, L. and Feng, Y. (2013), "Promotional marketing or word-of-mouth? Evidence from online restaurant reviews", *Information Systems Research*, Vol.24 No.3, pp. 596-612. DOI: 10.1287/isre.1120.0454

Mamalis, S. (2009), "Critical success factors of the food service industry", *Journal of International Food & Agribusiness Marketing*, Vol.21 No.2-3, pp.191-206. DOI: 10.1080/08974430802589709

Moro, S., Rita, P., Esmerado, J. and Oliveira, C. (2019), "Unfolding the drivers for sentiments generated by Airbnb Experiences", *International Journal of Culture, Tourism and Hospitality Research*, Vol.13 No.4, pp.430-442. DOI: 10.1108/IJCTHR-06-2018-0085

Moro, S., Rita, P. and Oliveira, C. (2018), "Factors influencing hotels' online prices", *Journal of Hospitality Marketing & Management*, Vol.27 No.4, pp.443-464. DOI: 10.1080/19368623.2018.1395379

Namkung, Y. and Jang, S. (2007), "Does food quality really matter in restaurants? Its impact on customer satisfaction and behavioral intentions", *Journal of Hospitality & Tourism Research*, Vol.31 No.3, pp.387-409. DOI: 10.1177/1096348007299924

Nave, M., Rita, P. anda Guerreiro, J. (2018), "A decision support system framework to track consumer sentiments in social media", *Journal of Hospitality Marketing & Management*, Vol.27 No.6, pp.693-710. DOI: 10.1080/19368623.2018.1435327

Nittala, R. (2015), "Factors influencing online shopping behavior of urban consumers in India", *International Journal of Online Marketing*, Vol.5 No.1, pp.38-50. DOI: 10.4018/IJOM.2015010103

Öğüt, H. and Onur Taş, B. (2012), "The influence of internet customer reviews on the online sales and prices in hotel industry", *The Service Industries Journal*, Vol.32 No.2, pp.197-214. DOI: 10.1080/02642069.2010.529436

Oliveira, C., Brochado, A., Moro, S. and Rita, P. (2019), "Consumer perception of tourist experience through online reviews", *Worldwide Hospitality and Tourism Themes*, Vol.11 No.6, pp.696-717. DOI: 10.1108/WHATT-09-2019-0052

Oozo.nl. (2018), "Wetenswaardigheden, cijfers en statistieken over Giethoorn-Zuid - Oozo.nl.", available at: https://www.oozo.nl/cijfers/steenwijkerland/giethoorn-zuid/giethoorn-zuid(Accessed 3 October 2020)

Osman, H., D'Acunto, D. and Johns, N. (2019), "Home and away: Why do consumers shy away from reporting negative experiences in the peer-to-peer realms?", *Psychology & Marketing*, Vol.36 No.12, pp.1162-1175. DOI: 10.1002/mar.21264

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol.64, pp.12-40.

Pekar, V. and Ou, S. (2008), "Discovery of subjective evaluations of product features in hotel reviews", *Journal of Vacation Marketing*, Vol.14 No.2, pp.145-155. DOI: 10.1177/1356766707087522

Patodkar, V. and I.R, S. (2016), "Twitter as a corpus for sentiment analysis and opinion mining", *IJARCCE*, Vol.5 No.12, pp.320-322.

Piccinelli, S., Moro, S., and Rita, P. (2021), "Air-travelers' concerns emerging from online comments during the COVID-19 outbreak", *Tourism Management*, Vol.85, 104313. DOI: 10.1016/j.tourman.2021.104313

Polyorat, K. and Sophonsiri, S. (2010), "The influence of service quality dimensions on customer satisfaction and customer loyalty in the chain restaurant context: A Thai case", *Journal of Global Business and Technology*, Vol.6 No.2, pp.64-76.

Qin, G. and Prybutok, V.R. (2008), "Determinants of customer-perceived service quality in fast-food restaurants and their relationship to customer satisfaction and behavioral intentions", *Quality Management Journal*, Vol.15 No.2, pp.35-50. DOI: 10.1080/10686967.2008.11918065

Raajpoot, N.A. (2002), "TANGSERV: A multiple item scale for measuring tangible quality in foodservice industry", *Journal of Foodservice Business Research*, Vol.5 No.2, pp.109-127. DOI: 10.1300/J369v05n02_08

Ryu, K., Lee, H.R. and Kim, W.G. (2012), "The influence of the quality of the physical environment, food, and service on restaurant image, customer perceived value, customer satisfaction, and behavioral intentions", *International Journal of Contemporary Hospitality Management*, Vol.24 No.2, pp.200-223. DOI: 10.1108/09596111211206141

Shaw, A., 2018, "Do people really look at Facebook reviews? Here's why it matters", *Forbes*, available at: http://Do People Really Look At Facebook Reviews? Here's Why It Matters (Accessed 3 October 2020)

Schwarz, D., Traber, D. and Benoit, K. (2017), "Estimating Intra-Party Preferences: Comparing Speeches to Votes", *Political Science Research and Methods*, Vol.5 No.2, pp.379-396. DOI: 10.1017/psrm.2015.77

Smith, A., Fischer, E. and Yongjian, C. (2012), "How does brand-related user-generated content differ across YouTube, Facebook, and Twitter?", *Journal of Interactive Marketing*, Vol.26 No.2, pp.102-113. DOI: 10.1016/j.intmar.2012.01.002

Stevens, P., Knutson, B. and Patton, M. (1995), "DINESERV: a tool for measuring service quality in restaurants", *The Cornell Hotel and Restaurant Administration Quarterly*, Vol.36 No.2, pp.56-60. DOI: 10.1016/0010-8804(95)93844-K

Tripathi, G. and Dave, K. (2017), "Exploration of service quality factors in the restaurant industry: a study of selected restaurants in the New Delhi region", In *Hospitality Marketing and Consumer Behavior* (pp. 189-207). Apple Academic Press.

Valkeman, A. (2018). "'Niet alleen Amsterdam kampt met groei toerisme' | NU - Het laatste nieuws het eerst op NU.nl.", available at: https://www.nu.nl/reizen/4342792/niet-alleen-amsterdam-kampt-met-groei-toerisme.html (Accessed 3 October 2020)

Xia Liu, Burns, Alvin C. and Hou, Yingjian (2017). "An Investigation of brand-related user-generated content on Twitter", *Journal of Advertising*, Vol.46 No.2, 236-247. DOI: 10.1080/00913367.2017.1297273

Ting, Y. (2020), “Sentiment analysis in restaurants on social media reviews: the case of Giethoorn restaurants”, Master’s dissertation, Iscte - Instituto Universitário de Lisboa, available at https://www.iscte-iul.pt/tese/11038 (accessed 24 August 2021).

Yu, C. and Zhang, X. (2020). “The embedded feelings in local gastronomy: a sentiment analysis of online reviews”, *Journal of Hospitality and Tourism Technology*, Vol.11 No.3, pp.461-478. DOI: 10.1108/JHTT-02-2019-0028

Zahoor, K., Bawany, N.Z., and Hamid, S. (2020), “Sentiment Analysis and Classification of Restaurant Reviews using Machine Learning”, In *2020 21st International Arab Conference on Information Technology* (ACIT) (pp. 1-6). IEEE.

Tables

Table 1 Summary of studies that have addressed online reviews within small restaurant businesses.

Reference	Source	Nr. of Reviews	Approach	Findings
(Hlee et al., 2019)	Yelp	2629 of 3 casual and 3 luxury restaurants	Descriptive statistics and hierarchical regression	The effect of content richness and source credibility on utilitarian evaluations are greater for a casual restaurant than for a luxury restaurant
(Yu & Zhang, 2020)	TripAdvisor	814 reviews from 78 traditional Macau restaurants	Sentiment analysis	Positive feelings towards the food and the service do not show a linear relationship
(Fernandes et al., 2021)	TripAdvisor	1220 reviews for 6 restaurants in Algarve, Portugal	Descriptive statistics, text mining and sentiment analysis	The use of internal restaurant data combined with online reviews leverages decision making

Table 2 Categorization of Reviews

Items	Description
Text Review (Zwaantje)	Goede locatie maar bediening kan beter Het terras ligt aan de gracht en daardoor mooi en leuk uitzicht. Bediening is te krap voor drukke dagen waardoor het rommelig overkomt. Drankjes zijn vlot geserveerd maar etenswaren kwamen in etappes. Kwaliteit van gerechten redelijk maar niet meer dan dat.Op terras afrekenen lukt niet dus dan maar binnen - dit duurde echter lang. Ruimte voor verbetering zeker gezien locatie.
Translation	"Good location but control is better" May 21, 2018 Jeroen S The terrace is located on the canal and thus beautiful and nice views. Operation is too tight for busy days making it happen messy. Drinks are served quickly but food came in ages. Quality of food reasonable but no more than that. On desk Checkout not so successful but then inside - however, this was long. Room for improvement especially considering location.
Social Media	TripAdvisor
F&B	Quality of food reasonable but no more than that.
Service	Operation is too tight for busy days making it happen messy. Drinks are served quickly but food came in ages. On desk Checkout not so successful but then inside - however, this was long.
Atmosphere	The terrace is located on the canal and thus beautiful and nice views. Room for improvement especially considering location.
Value	N/A

Table 3 Number of Reviews under Service Dimensions

Service Dimensions	TripAdvisor	Google	Facebook	Total
F&B	1,205	1,134	710	**3,049**
Service	1,041	956	618	**2,615**
Atmosphere	717	675	282	**1,674**
Value	314	230	52	**596**
Total				**11,140**

Table 4 Sample of the Word Sentiment Analysis Report from Restaurant Zwantje

ID	Highlighted Text	Word	Word Sentiment	Word Sentiment +/-	Word Intensifier	WordNegator
24	wonderful meals over there, the cook is very **fantastic**.	fantastic	1.068	positive	very	
25	**Good** food and extensive menu.	good	0.5	neutral		
29	**terrible** food	terrible	-0.75	negative		

Table 5 The Sentiment Word from the Restaurant Achterhuus in TripAdvisor under F&B

Word	Negative	Neutral	Positive	Grand Total
good		17	6	23
delicious		10	3	13
great	2		6	8
very good		7		7
excellent			6	6
tasteful		2	1	3
the best			3	3
very tasty			3	3
pleasant		2	1	3
nice		1	2	3
Grand Total	**2**	**39**	**31**	**72**

Note: the numbers represent the times that the phrases appear in the reviews

Table 6 The Benchmark Restaurants under F&B in TripAdvisor

Name	Ranking	F&B Negative	F&B Neutral + Positive	F&B Grand Total	F&B (Positive+Neutral)/Total
Eetkamer	1	0	45	45	100.00%
Otterskooi	2	0	45	45	100.00%
Geythorn	3	0	28	28	100.00%
Piccola	4	0	55	55	100.00%
Het Wapen	5	0	16	16	100.00%
Smidse	6	0	17	17	100.00%
Lindenhof	7	1	99	100	99.02%
Grachthof	8	1	74	75	98.67%
Sloothaak	9	1	36	37	97.30%
Achterhuus	10	2	70	72	97.22%

The restaurants in red are the benchmark restaurants

Table 7 The Benchmark Restaurants

	Service Dimensions			
Social Media	F&B	Service	Atmosphere	Value

Tripadvisor	Lindenhof Grachthof Achterhuus	Fanfare Fratelli Eetkamer	Fanfare Grachthof Fratelli	Smit Eetkamer Rietstulp
Google	Hollands Venetie Achterhuus Vishandel Gerrits	Otterskooi Fanfare Jonge Hotel	Achterhuus Fanfare Smit's Paviljoen	Achternhuus Rietstulp Geythorn
Facebook	Lindenhof Grachthof Otterskooi	Jonge Hotel Lindenhof Fratelli	Lindenhof Fanfare Grachthof	Fanfare Fratelli

Figures

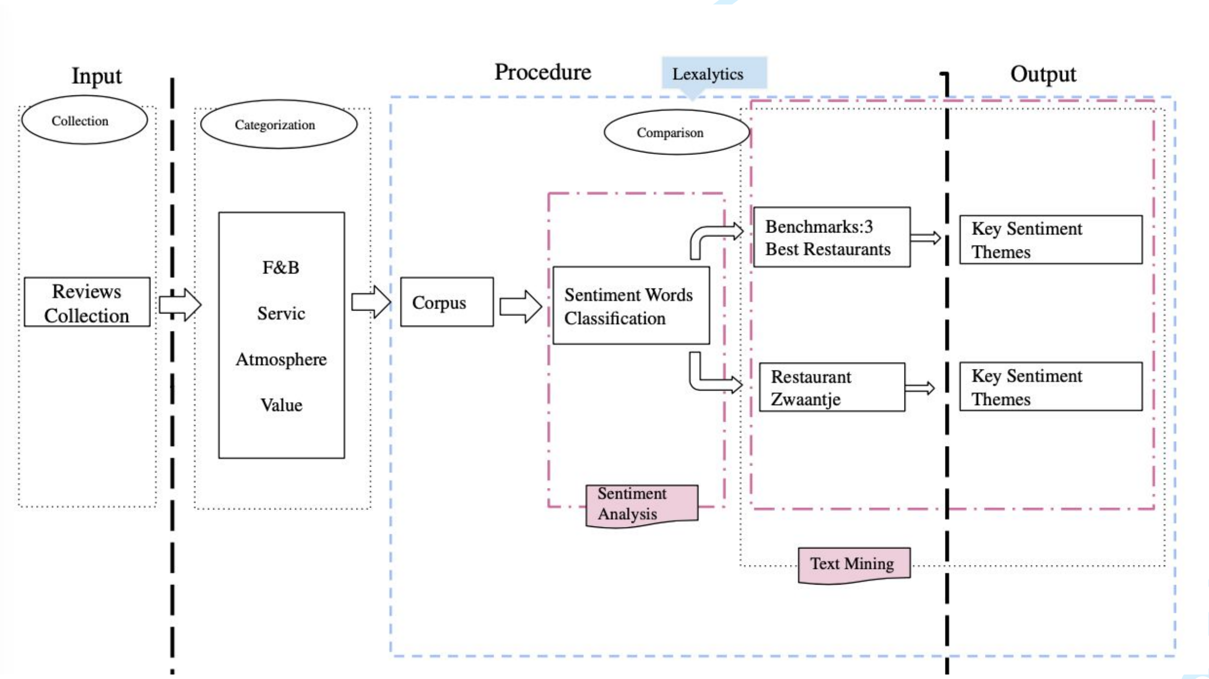

Figure 1 - Proposed Approach

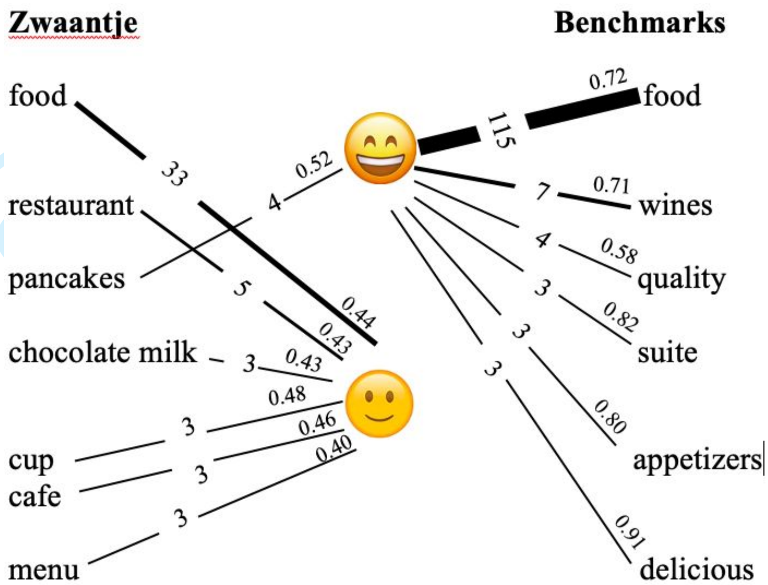

Figure 2 – Key Themes Extraction between Zwaantje and Benchmarks under F&B

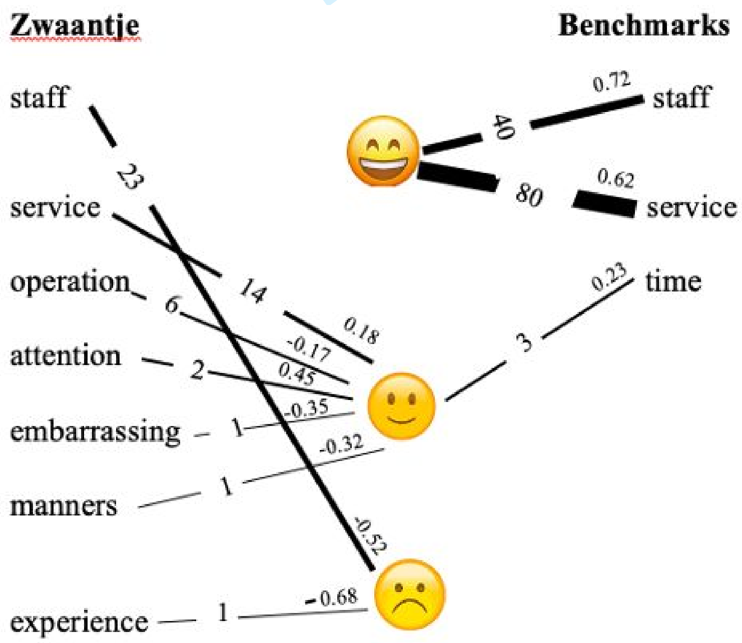

Figure 3 - Key Themes Extraction between Zwaantje and Benchmarks under Service

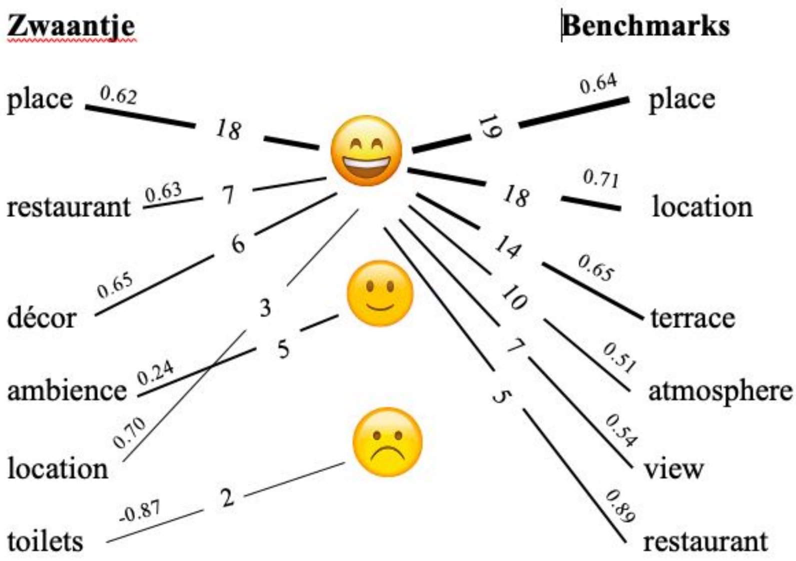

Figure 4 - Key Themes Extraction between Zwaantje and Benchmarks under Atmosphere

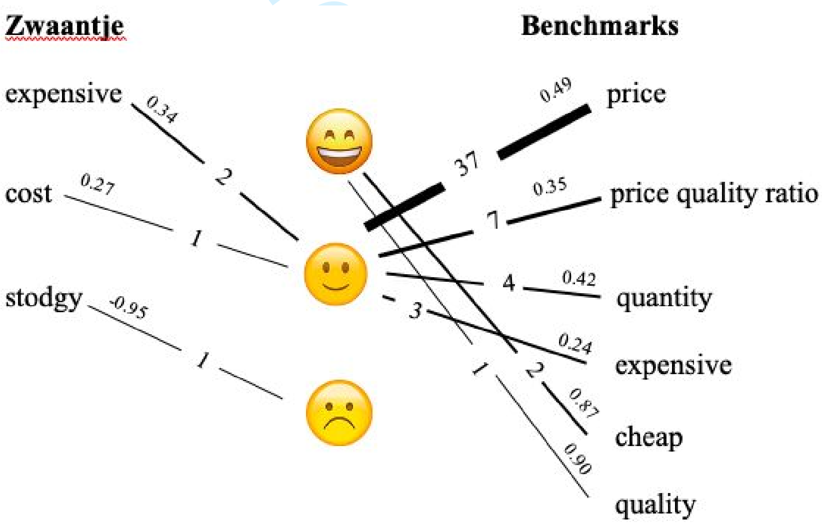

Figure 5 - Key Themes Extraction between Zwaantje and Benchmarks under Value

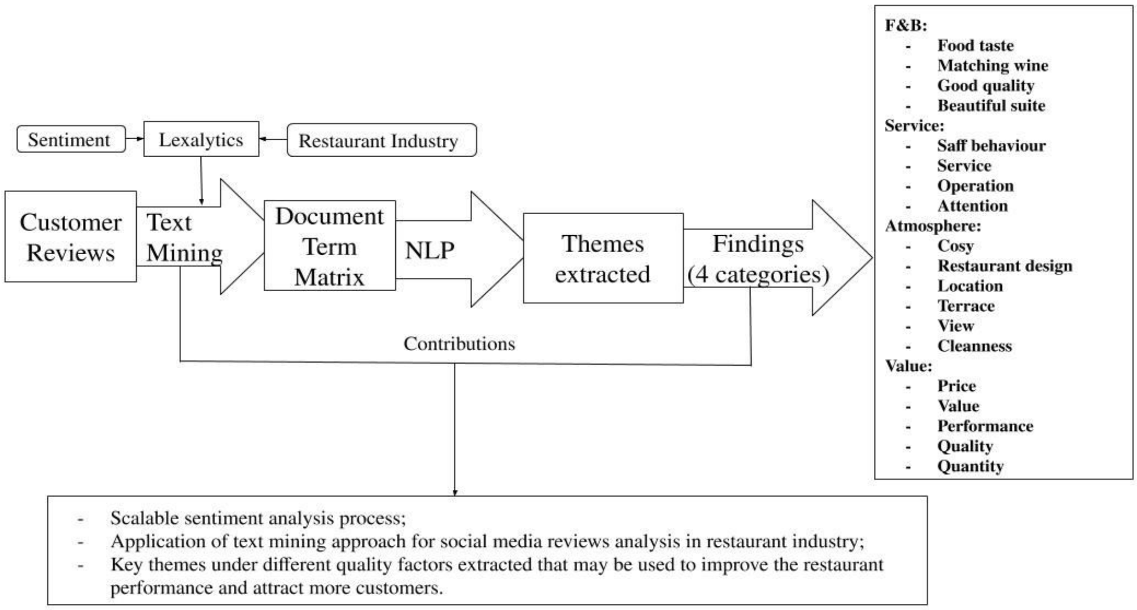

Figure 6 - Graphical Summary of Research Methods and Findings