A text mining analysis of customer evaluations in terms of gastronomy tourism*

Gastronomi turizmi açısından müşteri değerlendirmelerinin metin madenciliği ile analizi

Ahmet BÜYÜKEKE**1
Tufan ÖZSOY2

ABSTRACT: Nutritional alternatives, which were limited to regional diversity in the past, have increased extraordinarily over time. Besides being the basic element to sustain life, it has come to the fore as hedonic consumption. It is already known that discovering the local cuisine and the pleasure of eating are very important for tourists. Social media platforms have become the most effective tool for tourists in making decisions. They significantly influence the decisions of tourists on where to go, where to stay, what to eat and drink. The primary aim of this research is to analyze and make sense of TripAdvisor reviews of restaurants serving Kaş and Belek, which have different accommodation alternatives. For this purpose, topic modelling, sentiment, and name-entity recognition analyses were carried out with 10,829 customer comments from 147 businesses. Reviews are clustered under the most appropriate 3 distinct subjects (Experience, Food, and Atmosphere). The satisfaction level in the comments is 89.52% for Kaş and 95.64% for Belek. In total, 800 and 445 different food names were discovered in Kaş and Belek reviews, respectively. Most liked foods: Meat dishes such as steak, burger, and stroganoff with cream, pepper, tomato, garlic, and spicy sauces.
Keywords: Gastronomy, Food tourism, Customer reviews, Text mining

ÖZ: Geçmişte bölgesel çeşitlilikle sınırlı olan beslenme alternatifleri zaman içinde olağanüstü artış göstermiştir. Yaşamı sürdürmenin temel unsuru olmasının yanı sıra hedonik tüketim olarak da ön plana çıkmıştır. Turistler için yöresel mutfağı keşfetmenin ve yemek yeme zevkinin çok önemli olduğu bilinmektedir. Sosyal medya platformları karar verme sürecinde en etkili araçlardan biridir. Turistlerin nereye gidecekleri, nerede kalacakları, ne yiyip içecekleri konusundaki kararlarında sosyal medyanın etkisi büyüktür. Bu araştırmanın temel amacı, birbirinden farklı konaklama alternatiflerine sahip Kaş ve Belek'te hizmet veren restoranların TripAdvisor yorumlarını analiz etmek ve anlamlandırmaktır. Bu amaç doğrultusunda 147 işletmeden 10,829 müşteri yorumu ile konu modelleme analizi, duygu analizi ve isim-varlık tanıma analizi gerçekleştirilmiştir. Yorumlar, 3 farklı konu başlığı (Deneyim, Yemek ve Atmosfer) altında gruplandırılmıştır. Yorumlardaki memnuniyet oranı Kaş için %89,52, Belek için %95,64'tür. Kaş ve Belek yorumlarından sırasıyla 800 ve 445 farklı yemek ismi keşfedilmiştir. En beğenilen yemekler: biber, domates, sarımsak ve baharat gibi soslar ile sunulan biftek, hamburger ve straganof gibi et yemekleridir.
Anahtar Kelimeler: Gastronomi, Yemek turizmi, Müşteri yorumları, Metin madenciliği

* This study come out of the development of the conference paper presented at UTK21 (21st National Tourism Congress).
** Sorumlu Yazar / Corresponding Author
1 PhD, Adana Alparslan Türkeş Science and Technology University/Faculty of Business/Management Information Systems/ Information Technologies, abuyukeke@atu.edu.tr, https://orcid.org/0000-0002-6103-7646
2 Assoc. Prof., Adana Alparslan Türkeş Science and Technology University/Faculty of Business//Business/Marketing, tozsoy@atu.edu.tr, https://orcid.org/0000-0002-0138-4179

GENİŞLETİLMİŞ ÖZET
Giriş
Beslenme, insanın hayatta kalabilmesi için en temel fizyolojik ihtiyacıdır. Fakat turizm sektöründe beslenme faaliyetleri fizyolojik bir ihtiyacın ötesinde konumlandırılmaktadır. Turistler yeni lezzetler, yeni kültürler, yeni hikayeler yaşamak istemektedirler. Türkiye, Yunanistan, Lübnan, Ürdün, Hindistan ve Çin gibi birçok ülke bu eğilimin farkındadır ve bu nedenle turizm pazarlama faaliyetlerinde mutfak ve mutfak geleneklerine odaklanmaktadır (Avraham ve Ketter, 2016).

Online kullanıcı görüşleri diğer tüketicilerin satın alma davranışlarında önemli bir belirleyici olurken, turizm işletmeleri için de önemli bir tanıtım aracı olarak görülmektedir. Bu sayede işletmeler reklamlarını yapabilir, halkla ilişkiler faaliyetlerini daha geniş kitlelere iletebilir, şikâyet ve önerileri dikkate alarak iş stratejilerini yenileyebilirler. Bu araştırmada TripAdvisor çevrimiçi seyahat platformu üzerinden toplanan restoran yorumları metin madenciliği yöntemleri ile analiz edilerek, işletmeler için stratejik sayılabilecek bilgiler elde edilmeye çalışılmıştır.

Literatür taraması
Yiyecek hizmetleri, destinasyonun tedarik sisteminin önemli bir bileşenidir ve gastronomi turistlerinin ilgisini çekmek için önemli bir faaliyet olarak kabul edilebilir. Boyne ve diğerleri (2002), turistlerin seyahat ederken bütçelerinin neredeyse %40'ını harcadığını bildirmiştir. Diğer bir çalışmaya göre, yiyecek ve içecek harcamaları temel olarak toplam turizm harcamasının üçte birini temsil etmektedir (Mak Vd., 2012). Piyasanın büyüklüğü göz önüne alındığında, yiyecek ve içeceklerin bir destinasyon için etkili bir tanıtım ve konumlandırma aracı olabileceği açıktır (Hjalager ve Richards, 2003). Mutfak deneyimleri, kültürel turizmin güçlü bir unsurudur.

Turistlerin karar alma süreçlerini doğru anlamak, onları turizm türüne, destinasyona ve kültürel yapıya göre değerlendirmek başarılı turizm stratejilerinin geliştirilmesi açısından hayati önem taşımaktadır. Gretzel ve diğerleri (2006) gıda turizminin yaygınlaşmasını ve popülerleşmesini teknolojik gelişmelere ve internete bağlamaktadır. Gastronomi söz konusu olduğunda, turistlerin bilmedikleri bir destinasyonu değerlendirmeleri için en etkili araçlardan biri kullanıcı yorumlarıdır. Ticari reklamlardan daha güvenilir olarak kabul edilen tüketici yorumları, böylece ürün ve işletmelerin geleceğinin teyit edilmesinde önemli bir unsur haline gelmiştir.

Metodoloji
Bu çalışmada turistik destinasyonlarda sunulan yeme-içme hizmetlerine ilişkin tüketici yorumları incelenerek turistlerin duygu ve düşüncelerinin anlaşılması ve yeme-içme deneyimlerinin temel örüntülerinin belirlenmesi amaçlanmaktadır. Bu amaçla, Kaş ve Belek turizm bölgelerinde faaliyet gösteren restoran işletmelerine yönelik sosyal medya yorumları otomatik olarak toplanarak duygu analizi, konu modelleme analizi ve isim-varlık tanıma analizi gerçekleştirilmiştir.

Veriler, Python programlama dili ile geliştirilen crawler yardımı ile TripAdvisor platformundan toplanmıştır. Belek'ten 34 restorana ait 3,659, Kaş'tan ise 113 restorana ait 7,170 yorum elde edilmiştir. Toplam yorum sayısı 10,829'dur.

Analiz öncesinde yorumların ön işleme tabi tutularak cümleye anlam katmayan fazlalıklardan arındırılması gerekmektedir. Bu sebeple sırası ile; harf dışında bütün karakterlerin silinmesi, bütün karakterlerin küçük harfe çevrilmesi, lemmatize işlemi ile kelimelerin temel formuna dönüştürülmesi ve son olarak da stop-words listesinde bulunan kelimelerin filtrelenmesi gerçekleştirilmiştir. Stop-words kelimeler; cümlede çok sık geçen fakat çok fazla anlamı olmayan kelimelerin listesini ifade etmektedir.

Konu modelleme analizi sonrasında Belek yorumları Deneyim, İçecek ve Kahvaltı ve Yemek başlıkları altında gruplandırılmıştır. Deneyim konusu yorumlarda en çok konuşulan konu başlığıdır. Kaş yorumları Deneyim, Yemek ve Atmosfer başlıkları altında gruplandırılmıştır. Kaş yorumlarında da en fazla Deneyim konusu konuşulmuştur.

Yorumlara duygu puanı atayabilmek için sözlük-tabanlı duygu analizi çalışması yapılmış ve her bir yorumun duygu puanı hesaplanmıştır. Hesaplanan bu puanlar isim-varlık tespiti analizi ile keşfedilen yiyecek ve içeceklerin duygu durumu puanlarının hesaplanmasında kullanılmıştır. Yemek isimlerinin keşfedilmesinde Amerika Tarım bakanlığına ait güncel ve yaklaşık 1,2 milyon yemek isminin kayıtlı olduğu bir liste kullanılmıştır. Bu liste ile tüm yorumların taratılması sonucunda Kaş yorumlarından 800, Belek yorumlarından 445 farklı yiyecek-içecek ismi yakalanarak toplamda 1,245 yiyecek-içecek ismi yakalanmıştır. Bütün yiyecek-içecek isimlerinin duygu puanı hesaplanmış Kaş ve Belek bölgeleri için ayrı ayrı en yüksek puana sahip 20'şer isim aşağıda tablo 4'te sunulmuştur.

Bulgular ve tartışma

Kaş ve Belek'te faaliyet gösteren restoranlara ait yorumlardan sırasıyla 800 ve 455 yemek ismi tespit edilmiştir. Her bir yemek için duygu puanı hesaplanmıştır. Aşağıda bulanan Tablo 5'ten her iki bölge için de sosların en yüksek puanları aldığı görülmektedir. Belek bölgesindeki soslar ve ana yemekler birlikte incelendiğinde et ürünleri ve soslardan memnuniyet düzeylerinin oldukça yüksek olduğu görülmektedir. Kaş bölgesinde öne çıkan soslar (pesto, domates sosu gibi) bu bölgede İtalyan lezzetlerinin çok tutulduğu ve müşteri memnuniyetinin yüksek olduğu anlamına gelebilir. Ayrıca Tablo 5'teki "Ekmek, Pizza, Makarna" grubuna bakıldığında İtalyan sebzeli pizza ve mantı makarnanın popüler yiyecekler arasında olduğu görülmektedir. Deniz ürünlerine bakılırsa Kaş bölgesinin Belek bölgesine göre daha başarılı olduğu söylenebilir.

Sonuç ve öneriler

Çalışmanın başında gastronomi açısından yerel lezzetlerin daha çok ön plana çıkacağı düşünülse de sonuç dünya mutfağının yerel lezzetlerden daha baskın olduğunu göstermektedir. Bu durumun nedeni, turistlerin taleplerini dikkate alan işletmelerin yöresel lezzetleri daha az sunması olarak düşünülebilir. Aynı zamanda yiyecek-içecek isimlerini belirlemek için yapılan analizlerde Türkçe isim listesinin olmaması da bu duruma sebep olarak gösterilebilir.

Bu çalışmada yöresel lezzetleri yorumlardan tespit edebilmek için ihtiyaç duyulan Türkçe yemek isimleri (İngilizce yazılı) listesinin olmayışı çalışmanın eksik kısmı olmuştur. Gelecekte yapılacak bu tür çalışmalar için dünya dillerine veya en azından Avrupa dillerine uygun yerel lezzetler listesinin olması akademik çalışmalara ve sektöre önemli katkılar sağlayacaktır.

Introduction

Nutrition is the most basic physiological need of humans in order to survive. When it comes to tourism, tourists (especially for a destination they are visiting for the first time) may be concerned about how they can meet their nutritional needs. Today, however, the situation is discussed in a different context. Now, tourists position their nutritional activities beyond a physiological need on their trips and want to experience new tastes, new cultures, and new stories. Many countries such as Turkey, Greece, Lebanon, Jordan, India, and China, are aware of this tendency, and therefore, focus on cuisine and culinary traditions in their tourism marketing activities (Avraham & Ketter, 2016). Thanks to the developments in information and communication technologies, the volume of user-generated content in the digital environment is increasing day by day. The traces and impact of word-of-mouth (WOM) marketing on social media platforms such as Facebook and Instagram, and travel and accommodation platforms such as Tripadvisor, Booking, Expedia and Yelp attract the attention of academics.

Food services are an important component of the destination's supply system and can be considered as an important activity in attracting gastronomy tourists to a destination. Boyne et al. (2002) reported that tourists spend almost 40% of their budget on food while traveling. According to another study, food and beverage expenditures represent essentially one-third of total tourism spending (Mak et al., 2012). Considering the size of the market, it is clear that food and beverages can be an effective promotional and positioning tool for a destination (Hjalager & Richards, 2003).

Online user opinions become an important determinant in the purchasing behaviour of other consumers, and it is also seen as an important promotional tool for tourism businesses. In this way, businesses can make their advertisements, convey their public relations activities to wider audiences, and revise their business strategies by taking into account the complaints and suggestions. In summary, they have an interactive medium that allows for multi-faceted communication. However, it cannot be said that every tourism business can benefit equally from this existing potential. Existing information and communication technologies offer opportunities to analyze the attitudes and behaviours of tourists better and in detail. Big data analysis is a fundamental instrument that can be used in this context. Especially in an environment where 2.5 exabytes (1 billion GB) of data are produced per day, sentiment analysis emerges as an efficient tool in making use of such a large and unstructured data set. Sentiment analysis also called opinion mining, can produce meaningful information from texts. Treating online customer opinions as just a score eliminates the possibility of understanding the implicit comments they contain and turning them into information. It will be much more goal-oriented to make inferences about how consumers feel on the basis of individuals, rather than considering them as a mass.

The next section reflects the literature review of gastronomic tourism. In the third chapter, the collection of data and the application of the relevant analysis methods are discussed. Later, the findings were shared, and the conclusions and recommendations were made in the last section.

Literature review

The analysis carried out by collecting customer comments from online travel platforms in the field of tourism has increased significantly in recent years. These analyses to understand tourists' thoughts and experiences are invaluable to the academic community and industry. On the other hand, the large amount of data that is automatically collected through the crawler also makes these analyses valuable. Text mining and various machine learning methods are used in analysing large volumes of data gathered.

In Turkey, there are some studies using tourist comments in the field of gastronomy and based on content analysis (such as: Şahin et al., 2018; Akyol, 2019; Arslan, 2020; Kendir, 2020; İbiş, 2021). Studies in this field tend to increase over the years. The study conducted by (Eren & Çelik, 2017) on "determining the most prominent food and beverages in the formation of a destination's online gastronomy image" is a good example. Content analysis was also used as a methodology for this study. In studies where the number of comments is much higher, text mining methods are used that allow the automatic interpretation of texts. It is believed that increasing the number of studies that consider customer comments on social media platforms and use large datasets will enrich the literature.

Sharing economy and new trends in tourism

The change in economy and innovative business models with a tendency towards digitalization are widely accepted on a global basis. As of March 2021, there are more than 5.1 billion internet users in the world, which shows that 65.6% of the entire world population is in the digital environment (Miniwatths Marketing Group, 2021). In a day, 500 million tweets and 293 billion emails are sent, 65 billion messages are transmitted on WhatsApp, 5 billion searches are made on search engines, and 4 petabytes (= 4 million Gb) of data are created only on Facebook. By 2025, it is estimated that 463 exabytes of data per day (212,765,957 DVD-equivalent data volume) will be generated globally (Desjardins, 2019).

While sharing between people dates back to the first days of human history, the current debate on the sharing economy and common consumption goes back to the early 2000s. Using the internet, users began to exchange goods and services with platforms such as Couchsurfing and Freecycle. Over time, these non-profit initiatives have pioneered large business enterprises that connect consumers and providers around the world (Cheng, 2016). Some claim that the sharing economy will reduce travel costs, encourage more people to participate in touristic activities, increase tourism revenues, provide a better allocation of resources, and distribute tourism revenues better, as well as have direct benefits for tourism. On the other hand, some argue that the sharing economy is inadequately regulated and taxed, poorly organized, predatory, and exploitative (Quattrone et al., 2016).

Developing information technologies have started to create new structures that support internet user participation in recent years. The internet, which started as a one-way information platform, has become a process that allows multi-directional interaction that enabling users to create a huge amount of content. Consumer evaluation and rating systems allow consumers to freely express their experiences, guide others, and at the same time benefit from the information provided by others.

Food tourism as a special type of tourism

The relationship between tourism and food dates back to the past. Wolf (2002) states that spices, wine, fruits, and other food products were valued as currency in the past. On the other hand, Mennell et al. (1992) argue that taverns and inns were used to serve basic food, and this developed the concept of commercial hospitality service and formed the basis of the culture of eating out. Over time, this dining out experience has caused the increase in the number of restaurants in resort areas and their diversity in terms of the products and services they offer (Beardsworth & Keil, 2002). Culinary experiences are an important element of cultural tourism. Drinks and food have become an integral part of destinations. Previously, alternatives to nutrition were perceived as a basic need for tourists, but now this phenomenon has changed and has become an important motivation for visits. Now, tourists are curious about the culinary diversity of the destination as an important part of their cultural experience, and those with culinary richness are prioritized among other alternatives in their destination preference.

The Global Food Tourism Report of 2013 published by the United Nations World Tourism Organization, the food tourism market is growing more and more every year, and food/beverage has moved to the centre of tourism experiences. In fact, a destination's cuisine determines the overall experience of touristic activities (UNWTO, 2014). Jiménez-Beltrán et al. (2016) underline that local cuisine can be the main motivation for tourists to choose a destination again, while Çetin and Bilgihan (2016) argue that most of the tourists are in search of local food experiences. According to Quan and Wang (2004), meals and beverages can be considered as the main or secondary motivation source for tourists' destination preferences. Symons (1999) sees local food as an essential component of a destination's attractiveness and overall tourist experience. Research by Reynolds (1993) found that food and drink can help tourists to discover the social and economic lifestyle of a destination and even a country. According to Hall and Sharples (2003), as the level of consumer interest in food tourism increases, it is expected that the main or sole motivation of travel will be a food-related activity. Researchers state that there is a high level of interest for gourmet-gastronomy tourism, a moderate interest for culinary tourism, and a low or non-existent level of interest for city-country tourism.

Digital reputation and e-wom

A good understanding of consumers' experiences is vital for the food and beverage industries in order to build brand image and ensure customer loyalty. It is necessary to follow the post-purchase comments and behaviours of previous customers in order to understand many variables, from which local flavours will be offered to what should be the basic characteristics of an atmosphere that will increase the service quality. The taste of the food, the speed and shape of the presentation, the temperature, as well as the physical environment where the presentation takes place (location, decoration, scenery, music, the distance between tables, ventilation, parking lot, children's playgrounds, etc.) and cost can be decisive in customer satisfaction. Undoubtedly, academically addressing a process affected by so many variables contains various difficulties. The fact that the target audience is composed of several societies with different cultural structures also complicates the situation.

Understanding the decision-making processes of tourists correctly, interpreting them according to tourism type, destination, and cultural structure is of vital importance for developing successful tourism strategies. This is where user-developed content comes into play. Because tourists take into account user-developed content while shaping transportation, accommodation, nutrition, entertainment, and travel plans. Zhang et al. (2010), user-generated content affects the consumer more than editor opinions. Sparks et al. (2013) similarly state that tourists find their previous consumer experiences more useful and reliable.

Gretzel, et al. (2006) attribute the widespread and popularization of food tourism to technological developments and the internet. The ease of accessing information about different cuisines has become one of the main motivations in planning touristic activities. When it comes to gastronomy, one of the most effective tools for tourists to evaluate a destination that they do not know about is WOM communication. WOM is the state of tourists who have visited the destination before, sharing their experiences (satisfaction, unhappiness) and thus informing others about their level of satisfaction. The concept known as WOM marketing refers to a process in which information transfer is ensured among consumers, previous experiences have the power to influence subsequent ones, and goods and services are more easily adopted by potential buyers. A negative difference between the expected-perceived consumer value and the experienced consumer value, that occur when a customer experiences below his expectation, will cause a dissatisfaction. The dissatisfaction can also be observed in different digital channels as a negative WOM. Similarly, an expected or higher level of experience designates a positive WOM.

With the development and spread of web-based information technologies, it has evolved into a structure called E-WOM. Since WOM communication takes place one-on-one between individuals in the traditional sense, it is not possible to monitor this process and its effects. E-WOM, on the other hand, has a traceable structure. In addition, E-WOM can be used in the decision-making process by consumers from many different segments, depending on its widespread access potential. Consumer reviews, which are considered more reliable than commercial advertisements, have thus become an important element in confirming the future of products and businesses. The fact that E-WOM has become such an important tool has been effective in finding its place in many different environments (e-mail, WhatsApp, websites, blogs, virtual communities, newsgroups, chat rooms, product review sites, online forums, etc.). Many variables such as the user's competencies in information technology use, education level, and participation level regarding the product can be effective in the preference of the relevant environment.

Methodology

The impact of the digital economy is felt more and more in daily life. The transformation from the traditional economy to the digital one has also caused significant changes in consumer behaviour. The restaurant choice of consumers is no longer a random behaviour, it has become the result of an evaluation that takes into account the experiences of others. While professional reviews such as Michelin Guide and Gault & Millau in the past only appealed to a high-income group, today, thanks to user-generated content, everyone has a structure that can both enter data and make decisions by evaluating the existing content.

Consumer rating systems on online sites mean nothing more than an effort to quantify existing qualitative inputs. Due to the unstructured format of online reviews, these ratings are widely used in scoring and ranking businesses and services. However, since text inputs, consumers are going to express themselves easily, contain many valuable meanings, it is vital to analyze them.

Social media platforms and especially internet sites where online consumer opinions are shared are considered as the main source of sentiment analysis studies, given the volume of consumer-generated content they have (Zhang et al., 2010). However, Kreeger et al. (2018) stated that the number of marketing-based studies using sentiment analysis is very limited. When it comes to the marketing of tourism, it is seen that academic studies on sentiment analysis are quite limited. This indicates a gap in the literature for understanding the change in the economy and consumption patterns.

In this study, it is aimed to understand the feelings and thoughts of tourists and to determine the basic patterns of food experiences by examining the consumer evaluations about the food services offered in tourist destinations. For this purpose, the social media comments on the restaurant businesses operating in Kaş and Belek tourism regions were automatically collected and analysed with sentiment analysis, topic modelling, and named entity recognition.

Kaş is a tourism region that generally has hostels, apartments, or boutique-style accommodation facilities. The tourists staying in this region visit historical and touristic areas, participate in boat tours, or engage in activities in the city centre (such as shopping, eating, and drinking) (Büyükeke et al., 2020), rather than spending time in the accommodation facility. Belek stands out with its golf tourism, eco-tourism, and high-quality 5- and 7-star hotels (Erkuş-Öztürk, 2009). The tourism establishments in Belek and Kas are not as diverse and extensive as in other parts of Antalya. Thanks to the analyses carried out with the comments gathered from the Kaş destination, it will be possible to obtain customer experiences and opinions about small companies. On the other hand, with comments belong to Belek, important information can be obtained about a destination with 5 and 7-star businesses.

The data were collected from the TripAdvisor web site (tripadvisor.com) via the crawler developed with the Python programming language. TripAdvisor is the most preferred community-based online comment platform by academicians as a data source in their studies in the field of tourism (Xiang et al., 2017). The crawler has identified 56 restaurants serving in Belek and collected a total of 3,698 comments. In order to increase the validity of the study, the 22 restaurants which have less than 5 comments were excluded from the analysis. Eventually, a total of 3.659 comments were preserved from the remaining 34 businesses. The average number of reviews per business is approximately 108. In the first stage, 7.278 comments belonging to 165 businesses were collected for Kaş, and when the research constraint was applied, 113 businesses with 7.170 comments have remained. A total of 10.829 consumer reviews were introduced to the analysis. The average number of reviews for each business is approximately 73.

Figure 1: Number of comments by year (Belek)
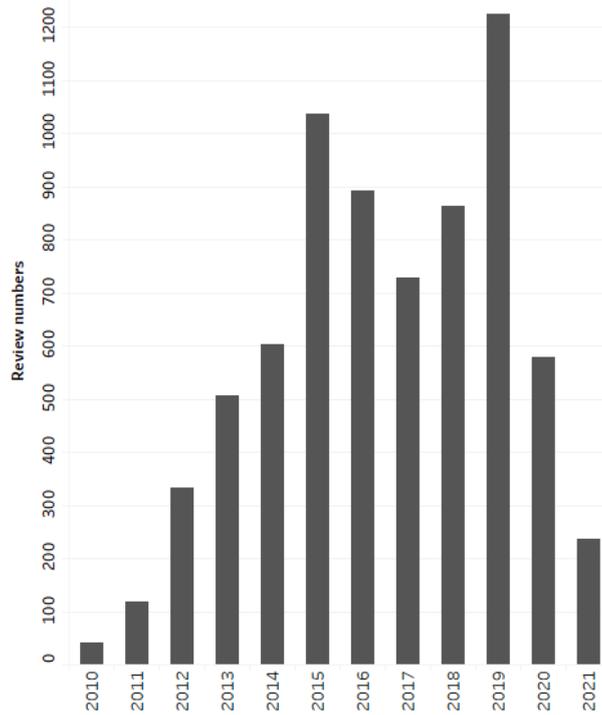
Source: (Created by the author)

Customer reviews contain the comment text, date, and rating. Figure 1 illustrates the distribution of comments by year for the Belek area. While an increase has been observed in the number of comments in recent years, the decrease in 2020 draws attention. It can be thought that the reason for this is the COVID-19 epidemic. The year 2021 continues and the number of comments supposed to increase until the end of the year.

Figure 2 shows the number of reviews of Kaş region restaurants by year. It can be noticed that it has a characteristic similar to the graph from the Belek region.

Figure 2: Number of comments by year (Kas)
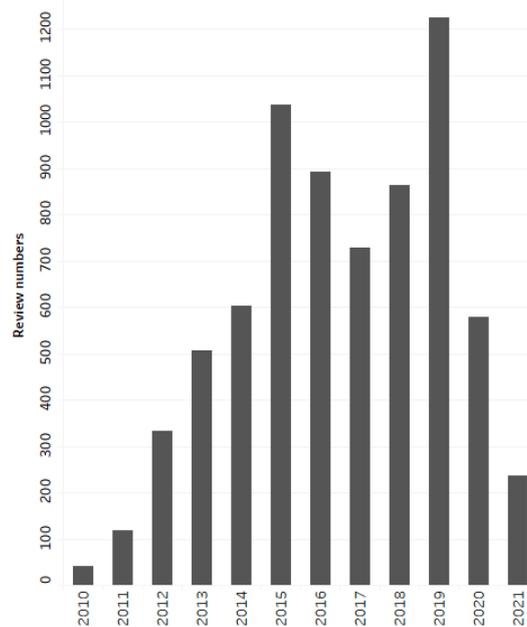
Source: (Created by the author)

Comments according to rating values are shown in Table 1. The average rating of Belek comments is 4.78, and the average rating of Kaş comments is 4.48. It can be seen that the comments mostly have high rating values. As the rating values decrease, the average word count of the comments increases. This shows that the complaints in the reviews are expressed in longer sentences.

Table 1: Number of comments by rating

Rating	Belek			Kaş		
	Review Numbers	%	Average Words	Review Numbers	%	Average Words
1	55	0.02	110.34	279	0.04	109.83
2	33	0.01	85.06	226	0.03	97.14
3	83	0.02	67.19	417	0.06	79.62
4	311	0.08	53.74	1.132	0.16	64.87
5	3.177	0.87	51.71	5.116	0.71	58.06

In this study, topic analysis was used to group the comments by certain titles, and sentiment analysis methods were used to determine the emotional intensity of the comments. In addition, with the help of regular expressions, the names of the food and beverage mentioned in the comments were discovered.

Pre-processing
Some preprocessing is required for the comments in unstructured text format to be analysed by the computer. Computer processors can operate on binary number system. For this reason, a dictionary is created in which each unique word, special character, or punctuation mark in the comments is expressed with a unique numerical value. This dictionary is called bag of words- BOW. However, the large size of the dictionary complicates the analysis. Therefore, special characters and numeric values that do not have much meaning in a sentence are removed. In this way, the size of the matrix in which the comments are expressed numerically is reduced.
Pre-processing steps:
- All characters except letters have been cleared.
- All letters converted to lowercase because Hello and hello are separate words for computers.
- Wordnet Lemmatizer belonging to Nltk[3], the Natural Language Processing (NLP) library, has been used to restore words to their basic form. E.g., *walking* to *walk*.
- Words that do not make sense in a sentence and are repeated frequently (a, an, the, which…) were removed from the comments using the Nltk stop words list.

Analyses
The main analysis methods used to obtain information from textual data are text classification, text summarization, text categorization, topic analysis, and sentiment analysis (Zhai & Massung, 2016). In this study, subject analysis was used to group the comments under certain headings, and sentiment analysis methods were used to determine the emotional intensity of the comments. In addition, with the help of regular expressions, the names of the food and beverage (foods) mentioned in the comments were discovered.

Topic modelling analysis
Subject analysis refers to grouping a document under certain headings as a result of statistical operations. Statistical operations are the process of capturing repetitive, often co-existing word patterns. LDA (Latent Dirichlet Allocation) method is mostly used in subject analysis studies (Blei et al., 2003).

---

[3] https://www.nltk.org/

In this method:
- • Documentation is a mix of topics, and
- • Topics are a mix of words.

This expression means the following for a corpus composed of interpretations; each comment contains all topics with a probability value ranging between 0 and 1, with the total probability equal to 1. The topic with the highest probability is called the dominant topic for that comment. On the other hand, each word has a certain weight value in determining each topic title. The higher this value, the greater the influence of the word in determining which topic to group a comment under. In this research, the Python Gensim library was used to implement the LDA model.

In order to find out the number of appropriate topics the Belek and Kaş comments can be grouped under, the numbers 3 to 7 were tested with a loop, and there was a limit of at least 20 comments under each topic. As a result, it was observed that the comments were grouped under the 3 most appropriate topics. In addition, the consistency score (Coherence Score) was the highest for the number 3. Table 2 shows the titles for the Belek region and the words with the highest weights in determining the subjects. The subject titles were determined according to the high-frequency words.

Table 2: Topics and words (Belek)

	Food	Drinks&Breakfast	Experience
1	steak	bar	great
2	meal	drink	service
3	year	english	visit
4	order	breakfast	staff
5	cook	beer	recommend
6	starter	price	friendly
7	restaurant	owner	place
8	dish	lovely	amazing
9	prawn	fish	time
10	eat	fresh	excellent
11	main	watch	nice
12	turkish	look	come
13	chicken	watch	back
14	garlic	full	enjoy
15	reasonable	local	definitely

Figure 3 shows the distribution of comments according to the dominant topics. The numerical values at the end of the graphic bar line are the average rating of the comments. It is discerned that the average rating of the comments about Experience is higher than the others.

Figure *3*: Topic distribution of comments (Belek)
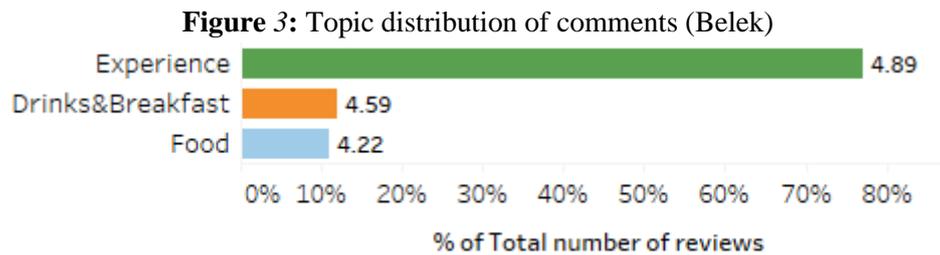


Source: (Created by the author)

Subject titles determined according to high-frequency words for eyebrow comments are given in table 3 below. When Figure 4 is examined, it is seen that the average rating of the comments on the title Food is lower. The comments of the words wait and bad, which express negativity among the words, can be examined separately.

Table 3: Topics and words (Kaş)

	Food	Atmosphere	Experience
1	Fish	Table	food
2	Order	View	restaurant
3	Get	Find	place
4	Dish	Sea	service
5	Wait	Wine	great
6	Waiter	Drink	staff
7	Serve	Little	enjoy
8	Bad	harbour	friendly
9	Main	Menu	Nice
10	Take	Beach	Visit
11	Salad	Well	recommend
12	Starter	Steak	lovely
13	Say	Coffee	Time
14	chicken	Hotel	price
15	meze	Terrace	Eat

Figure 4: Topic distribution of comments (Kaş)
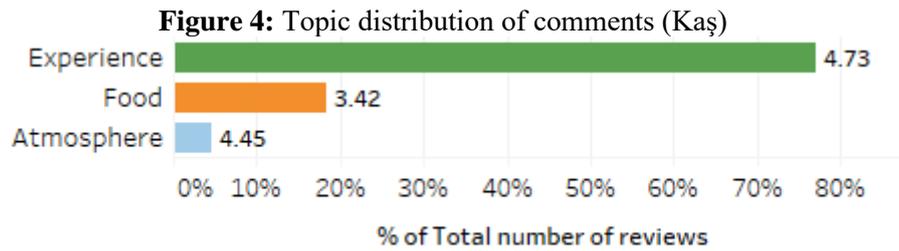
Source: (Created by the author)

When Figure 4 is examined, it is seen that the average rating of the comments on the title Food is lower.

Sentiment analysis

Sentiment analysis is a subfield of the text classification. It aims to determine the polarity and point of view in texts such as movie reviews, product reviews, or tweets. There are two methods in sentiment analysis, dictionary-based and machine learning methods. In machine learning methods, a predictive model is created using labelled data. Within the scope of this study, the dictionary-based method was used instead of machine learning because the average rating values of the comments were as high as 4.78 and 4.48 for Kaş and Belek regions, respectively. Expressions indicating the appreciation or dissatisfaction of the customers are frequently used in the comments of restaurant businesses. Thanks to these expressions, the positive or negative status of a comment can be determined. For this purpose, the SentiWordNet 3.0 dictionary, which was developed from the WordNet[4] dictionary, was used.

SentiWordNet is one of the most used dictionaries in dictionary-based sentiment analysis studies. In the dictionary, each synonym (word) has three types of points: positive, negative, and neutral. A word's score varies between 0 and 1. E.g., When the word amazing is examined: the positive score is 0.875, the negative score is 0.125. The neutral score can be calculated by subtracting the negative score from the positive score.

---

[4] https://wordnet.princeton.edu/

In order to find the emotional polarity of the comments, first of all, adjectives, adverbs, nouns, and verbs that make sense in a sentence were determined by the part of speech (POS) method. Python Spacy[5] library is used for this process. Then, using the SentiWordNet dictionary, the average of the difference between the positive and negative scores of each word in the comments was calculated, and the emotion scores of the comments were created. These scores were used to calculate the positive or negative scores of the food and beverage names mentioned in customer reviews. The formula for calculating the emotion score of a comment is shown in formula 1. This formula was applied to each comment one by one as a cycle, and the emotion scores of all comments were calculated.

$$E = \left(\frac{1}{n}\right) * \sum_{k=0}^{n}(posV(w_k) - negV(w_k)) \tag{1}$$

E: emotion score of a review
n: word count of a review
k: a sequentially increasing numeric value
w: k-th word in a review
posV: the positive value of the word (w_k) of a review found in the SentiWordNet dictionary
negV: the negative value of the word (w_k) of a review found in the SentiWordNet dictionary

Detection of food names
A current U.S. Department of Agriculture list[6] as of April 2021 was used to discover food names from the reviews. Food names with a maximum of 3 words (approximately 40,000) were selected from the nearly 1.2 million names on the list, and all comments regarding Kaş and Belek were scanned with the help of regular expressions. In this scan, a new list was created for both Kaş and Belek with finding food names. There are 800 food names in the list of Kaş and 445 in the list of Belek. Since it was not possible to mention all food names in the study, those with less than 10 frequencies were excluded from the names in the lists. In the last situation, there are 161 different food names in the Kaş list and 80 different food names in the Belek list. Table 4 lists the 20 most common food names found in Kaş and Belek reviews.

Table 4: Food names and frequency values

Belek (top 20)				Kaş (top 20)			
steak	653	bacon	90	steak	619	beer	227
beer	194	dessert	84	pizza	609	casserole	211
chicken	190	pizza	81	salad	535	dessert	211
garlic	184	pasta	81	bread	500	water	204
bread	148	cream	77	chicken	448	mushroom	200
sauce	128	meat	76	meat	439	pasta	183
curry	121	ice cream	71	seafood	354	coffee	175
salad	120	beef	69	sauce	297	shrimp	172
cocktail	113	dip	63	calamari	253	burger	154
ice	93	burger	60	cheese	228	cocktail	152

The emotional score for each food was calculated based on the scores obtained as a result of the sentiment analysis. For example, the name shrimp is repeated 172 times in all of the comments. The total emotion score of these comments is approximately 387 (It was explained in the Sentiment analysis section.). Consequently, the average emotional score (2.25) for shrimp can be easily found with process

---

[5] https://spacy.io/

[6] (Current Food Names List) https://fdc.nal.usda.gov/download-datasets.html

387/172. The high score of a food indicates that the emotional score of the reviews made about that product is high. By looking at the score of a food, it can be understood whether it is liked by customers. For example, if shrimp had a negative score, it should be considered that it is often mentioned in negative reviews.

All food names have positive scores when entire reviews are considered. This shows that, in general, every product is liked by customers. Indeed, the average rating values indicate this situation, and it can be understood that the ratings given by the customers for the comments are also consistent. However, if only negative comments are taken as a basis, it can be revealed which food received negative comments from some customers. Table 5 is obtained by grouping similar foods. In the list, the first five foods from each group are presented in order of points.

Tablo 5: Similar food groups with emotional score

	Belek		Kaş	
	Food names	**Average emotional score**	**Food names**	**Average emotional score**
Main Course	pepper steak	4.19	ratatouille	4.05
	heart	4.08	heart	3.69
	stroganoff	3.65	risotto	3.45
	burger	3.43	stroganoff	2.97
	meat	3.10	moussaka	2.95
Sea	shrimp	2.73	stuffed calamari	3.25
	calamari	2.69	squid	3.06
	salmon	2.64	salmon	2.89
	sushi	2.55	calamari	2.61
	seafood	2.36	seafood	2.55
Pastry	noodle	3.37	loaf	3.54
	bruschetta	3.25	treat	3.09
	naan	3.18	veggie	3.06
	bread	3.00	ravioli	2.83
	pasta	2.96	flatbread	4.05
Beverage	soft drink	3.38	soft drink	3.45
	cocktail	3.20	juice	3.17
	coffee	2.91	soda	3.07
	tea	2.71	sangria	2.95
	water	2.70	cocktail	2.87
Sauces	creamy	4.54	tomato sauce	4.75
	garlic sauce	4.16	pesto	3.87
	pepper sauce	3.73	creamy	3.82
	dip	3.59	date	3.78
	spicy	3.26	apple	3.37
Sweet, Snack	falafel	3.35	baklava	3.72
	ice cream	3.20	cheesecake	3.70
	slice	3.04	slice	3.54
	pudding	3.01	ice cream	3.18
	chocolate	2.92	crisp	3.14

Findings and discussion

As a result of the topic analysis, the comments of both regions were gathered under the most appropriate 3 different groups. Subject headings according to high-frequency words are Experience, Drink&Breakfast, and Food for Belek; It was determined by the authors as Experience, Food, and Atmosphere for Kaş. In approximately 77% of Belek and Kaş comments, the subject of Experience was found to be the dominant topic, and the average rating of these comments is quite high, such as 4.89 and 4.73, respectively. These comments usually contain words that include advice, likes, entertainment, and repetition of behaviour. This shows that the satisfaction level of tourists is high. 12% of Belek reviews have Drink&Breakfast with a rating average of 4.59%, and 11% have Food with an average rating of

4.22. 18% of Kaş comments have Food with a rating average of 3.42, and Atmosphere with an average of 4.45 in 5% of them. Considering the Food title for the Kaş region, the average of the rating values remained below 4.42 with 3.42.

From the comments of the restaurants operating in Kaş and Belek, 800 and 455 food names were ascertained, respectively. The average emotion score was calculated for each food. It is seen from Table 5 that the sauces for both regions received the highest scores. When the sauces and main dishes in the Belek region are examined together, it is seen that the satisfaction levels for meat products and sauces are very high. The sauces that stand out in the Kaş region (such as pesto, tomato sauce) show that Italian flavours are very popular with high customer satisfaction in this region. In addition, looking at the "Bread, Pizza, Pasta" group in Table 5, it is seen that Italian veggie pizza and ravioli pasta are among the popular foods. When the seafood is examined, the scores of the Kaş region are slightly higher than those of the Belek region. Although it is not enough to examine the customer comments alone, it can be interpreted that the small businesses in the Kaş region are more successful in this area and at least the customer satisfaction is higher. In the "Dessert, Snacks" group, falafel ranks first in Belek and baklava in Kaş.

Conclusion and recommendation
In this study, it is aimed to investigate the place of gastronomy in tourism and the positive and negative aspects of its effect on tourism. For this purpose, the tourist reviews of restaurants in Kaş and Belek destinations were collected automatically from the TripAdvisor platform and analysed with text mining methods. Customer reviews are a reality where they share their own experiences, wishes, and desires, and opinions, and thoughts impartially. The research results are completely based on this phenomenon. The inclusion of both Kaş, which has small businesses, and Belek, which is one of the big businesses, in the research can lead to a broad perspective on the sector.

In the research, the topic analysis of the comments was carried out and it was ensured that the comments could be summarized under certain titles. In addition, the flavours of the regions and customer satisfaction were revealed by using the sentiment analysis and the information filtered from the food names together.
Meat dishes such as steak, stroganoff, beef, burger for the Belek; It is thought that customer appreciation increases when used with sauces such as creamy, garlic sauce, pepper sauce, and spicy. Seafood reached a higher emotion score in Kaş than Belek. As well, in Kaş, Italian and French cuisine dishes such as ratatouille, risotto, heart came to the forefront with sauces such as tomato sauce, pesto, apple, and creamy. While Kaş is in search of more local, unpopular, and quality flavours, Belek focuses on more traditional and popular, easy-to-prepare flavours. This may affect the all-inclusive system. The fact that accommodation in Belek is mostly provided by big hotels based on the all-inclusive system may cause this type of evaluation by the tourists who do not leave the hotel and have to choose from the standard options at the open buffet. In addition, the fact that this study takes into account the current comments of the years coinciding with the pandemic period may also be affective in the findings. Because in the relevant period, tourists can be expected to choose accommodation facilities that are more sheltered, more precautionary, and standard security services are offered. In addition, the possibility that activities aimed at recognizing the local culture were more limited during the pandemic period should not be ignored.

Although it was thought at the beginning of the study that local flavours would come to the fore more in terms of gastronomic terms, the result shows that world cuisine is more dominant than local flavours. The reason for this situation can be considered as the fact that businesses that take into account the demands of tourists offer local tastes less. At the same time, the lack of a Turkish name list in the analyses carried out to determine the names of the foods can be shown as a reason for this situation.

In this study, the lack of a list of Turkish food names written in foreign languages, which is needed in order to identify local flavours from the comments, was the missing part. Having a list of local flavours suitable for world languages or at least European languages for such studies to be carried out in the future will make significant contributions to the studies and the sector.

References

Akyol, İ. Ö. (2019). Elektronik ağızdan ağıza iletişim, Destinasyona yönelik Tutum, Destinasyon Ve Gastronomi imajının Turistlerin Ziyaret Niyetine Etkisi: Türkiye örneği (Doctoral dissertation, Marmara Universitesi (Turkey).

Arslan, E. (2020). Çevrimiçi Gastronomik Turist Deneyimlerinin İçerik Analiziyle İncelenmesi. *AHBVÜ Turizm Fakültesi Dergisi*, 23 (2), 442-460

Avraham, E., & Ketter, E. (2016). *Tourism marketing for developing countries: Battling stereotypes and crises in Asia, Africa and the Middle East*. Springer.

Beardsworth, A., & Keil, T. (2002). *Sociology on the menu: An invitation to the study of food and society*. Routledge.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *the Journal of machine Learning research*, *3*, 993-1022.

Boyne, S., Williams, F., & Hall, D. (2003). On the trail of regional success: Tourism, food production and the Isle of Arran Taste Trail. *In Tourism and gastronomy* (pp. 105-128). Routledge.

Büyükeke, A., Sökmen, A., & Gencer, C. (2020). Metin madenciliği ve duygu analizi yöntemleri ile sosyal medya verilerinden rekabetçi avantaj elde etme: Turizm sektöründe bir araştırma. *Journal of Tourism and Gastronomy Studies*, *8*(1), 322-335.

Cetin, G., & Bilgihan, A. (2016). Components of cultural tourists' experiences in destinations. *Current Issues in Tourism*, *19*(2), 137-154.

Cheng, M. (2016). Sharing economy: A review and agenda for future research. *International Journal of Hospitality Management*, 57, 60-70.

Desjardins, J. (2019, April). How much data is generated each day. *In World economic forum* (Vol. 17).

Eren R. & Çelik M. (2017). Çevrimiçi Gastronomi İmajı: Türkiye Restoranlarının Tripadvisor Yorumlarının İçerik Analizi, *Turizm Akademik Dergisi*, 4 (2), 121-138.

Erkuş-Öztürk, H. (2009). The role of cluster types and firm size in designing the level of network relations: The experience of the Antalya tourism region. *Tourism management*, *30*(4), 589-597.

Gretzel, U., Fesenmaier, D. R., & O'leary, J. T. (2006). The transformation of consumer behaviour. *Tourism business frontiers: Consumers, products and industry*, *9*, 18.

Hall, C. M. & Sharples, E. (2003). The consumption of experiences or the experience of consumption? An introduction to the tourism of taste. In: C. M. Hall, E. Sharples, R. Mitchell *et al*. (eds), *Food Tourism Around the World: Development,Management and Markets*. Oxford: Butterworth-Heinemann, pp. 1–24.

Hjalager, A. M., & Richards, G. (Eds.). (2003). *Tourism and gastronomy*. Routledge.
Ignatov, E., & Smith, S. (2006). Segmenting Canadian culinary tourists. *Current issues in tourism*, *9*(3), 235-255.

İbiş, Salim. (2021). İstanbul'daki asırlık restoranlar ve bu restoranlara yönelik çevrimiçi yorumların incelenmesi. *Yönetim, Ekonomi ve Pazarlama Araştırmaları Dergisi*. 5(2) 127-139. ss.

Miniwatths Marketing Group. (2021, March 31). World internet Users Statistics and 2021 World Population Stats. *Internet World Stats Usage and Population Statistic.* https://www.internetworldstats.com/stats.htm

Jiménez Beltrán, J., López-Guzmán, T., & Santa-Cruz, F. G. (2016). Gastronomy and tourism: Profile and motivation of international tourism in the city of Córdoba, Spain. *Journal of Culinary Science & Technology*, *14*(4), 347-362

Kendir, H. (2020). Gastronomik Lezzetlere Yönelik Çevrimiçi Turist Deneyimlerinin İncelenmesi: Edirne Tava Ciğeri Örneği. *Safran Kültür ve Turizm Araştırmaları Dergisi*, 3 (3), 408-423 . Retrieved from https://dergipark.org.tr/tr/pub/saktad/issue/59328/845512

Kreeger, J. C., Parsa, H. G., Smith, S. J., & Kubickova, M. (2018). Calendar effect and the role of seasonality in consumer comment behavior: A longitudinal study in the restaurant industry. *Journal of foodservice business research*, 21(3), 342-357.

Mak, A. H., Lumbers, M., Eves, A., & Chang, R. C. (2012). Factors influencing tourist food consumption. *International Journal of Hospitality Management*, *31*(3), 928-936.

Mennell, S., Murcott, A., & Van Otterloo, A. H. (1992). The sociology of food: eating, diet, and culture.
Quan, S., & Wang, N. (2004). Towards a structural model of the tourist experience: An illustration from food experiences in tourism. *Tourism management*, *25*(3), 297-305.

Quattrone, G., Proserpio, D., Quercia, D., Capra, L., & Musolesi, M. (2016, April). Who benefits from the" sharing" economy of Airbnb? In *Proceedings of the 25th international conference on world wide web* (pp. 1385-1394).

Reynolds, P. C. (1993). Food and tourism: towards an understanding of sustainable culture. *Journal of Sustainable Tourism*, *1*(1), 48-54.

Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, *39*, 1-9.

Symons, M. (1999). Gastronomic authenticity and sense of place. In CAUTHE 1999: Delighting the Senses; Proceedings from the Ninth Australian Tourism and Hospitality Research Conference (p. 401). Bureau of Tourism Research.

Şahin, S., Girgin, G. K., Kazoğlu, İ. H., & Nur, G. (2018). Gastronomi Turistlerinin E-Şikâyet Davranışları Üzerine Bir Araştırma: Alaçatı. *Journal of Tourism and Gastronomy Studies,* 68, 82.

World Tourism Organization (2013), UNWTO Annual Report 2013, UNWTO, Madrid

Wolf, E. (2002). Culinary tourism: A tasty economic proposition. Portland. *International Culinary Tourism Taskforce*.

Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, *58*, 51-65.

Zhai, C., & Massung, S. (2016). *Text data management and analysis: a practical introduction to information retrieval and text mining*. Morgan & Claypool.

Zhang, Z., Ye, Q., Law, R., & Li, Y. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, *29*(4), 694-700.

Ethical approval
This study is among the studies that do not require ethics committee approval due to not containing any personal data of humans and is not an experimental study of animals.
Contribution rate of researchers
1st author contributed 60%, 2nd author contributed 40% to the study.
Conflict of interest
There is no potential conflict of interest in this study.