



# How hotel stars affecting customers' sentiment in Astana

DAlper Kürşat UYSAL<sup>1</sup>, <sup>D</sup>Egemen Güneş TÜKENMEZ<sup>2</sup>, <sup>D</sup>Nurzhan ABDIRAZAKOV<sup>3</sup>, <sup>D</sup>Murat Alper BAŞARAN<sup>4\*</sup>, <sup>D</sup>Kemal KANTARCI<sup>2,5</sup>

<sup>1</sup>Alanya Alaaddin Keykubat University, Rafet Kayis Faculty of Engineering, Industrial Engineering Department, Alanya-Antalya/Türkiye.

<sup>2</sup>Alanya Alaaddin Keykubat University, Tourism Faculty, Tourism Management Department Alanya-Antalya/Türkiye. <sup>3</sup>International University of Tourism and Hospitality, Turkestan, Kazakhstan.

<sup>4</sup>Alanya Alaaddin Keykubat University, Rafet Kayis Faculty of Engineering, Computer Engineering Department, Alanya-

Antalya/Türkiye.

<sup>5</sup>Hoca Ahmet Yesevi International Turkish-Kazakh University, Faculty of Economics, Administrative Sciences and Law, Department of Management and Tourism, Turkistan / Kazakhstan.

Corresponding author: Alper Kürşat UYSAL (Email: murat.basaran@alanya.edu.tr)

## Abstract

Hotel stars are key indicators for assessing locations, facilities, and other attributes. Although the World Tourism Organization has long advocated for such systems, no consensus has been reached on standardizing them. Italy and Spain use their own systems. Kazakhstan has no system. Since visitors rate hotels, Web 2.0 implementations have started using their rating systems. Hotels in countries without a standard system would be negatively affected and would be vulnerable to manipulation by Web 2.0. Topic modeling and sentiment analysis are implemented to explore how hotel stars converge or diverge regarding hotel attributes mentioned by visitors. TripAdvisor's 5,894 online reviews of Astana's hotels in Kazakhstan between 2006 and 2023 are used. The study pinpoints the same attributes as remaining issues, even though hotels differ in stars. Customers' sentiments about derived hotel attributes mostly converge for hotels rated 2, 3, and 4 stars; however, they diverge in 5-star hotels. Specifically, even though hotels with 2, 3, or 4 stars differ, they share similar issues. Thus, official rating systems should be adopted in Kazakhstan soon. Kazakhstan's hotels can directly adopt an official rating system to enhance attribute qualities.

Keywords: Kazakhstan, Sentiment analysis, Starred hotels, Topic modeling.

**Funding:** This study was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan through (Grant Number AP19679687 between 2023-2025).

History: Received: 12 March 2025 / Revised: 14 April 2025 / Accepted: 16 April 2025 / Published: 2 May 2025

**Copyright:** @ 2025 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

**Competing Interests:** The authors declare that they have no competing interests.

Publisher: Innovative Research Publishing

DOI: 10.53894/ijirss.v8i3.6666

Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

**Transparency:** The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

## **1. Introduction**

Kazakhstan is a Central Asian country with a long communist Soviet heritage whose tourism perspective was shaped by a communist economic doctrine in which the tourism industry was not aimed at developing as a driving factor of the economy, playing an income-generating component that can also boost economic development, employment, and diversification of its economy. Thus, its financial and other economic contributing factors were long overlooked. Since the collapse of the Soviet Union at the beginning of 1990, Kazakhstan has been a sovereign state with increasing oil, gas, and rare mineral exports that bring huge potential to reshape and diversify its economy. One of the steps toward this aim has been devising a strong and bulky service sector to expand, advance, and diversify its economy as several developing and developed countries do. One of the service industries that contribute to several components of the economy is the tourism industry, which brings job opportunities, local investments, and foreign investments, synchronization with the global economy by adapting global rules and regulations, and helping create a competitive and educated workforce in the field. Thus, better development of the tourism industry in Kazakhstan is a must-have operation that has been constantly surveilled and upgraded.

Even though Kazakhstan is not a hot spot for generic tourism concepts, it has huge advantages in attracting a variety of tourists whose interests are nature, ecology, authenticity, nostalgia, gastronomy, and genealogy tourism since Kazakhstan has a huge territory containing several mountainous areas, lakes, forests, and wildlife reservations, unique ancient towns such as Taraz (dated back 36 B.C.), Shymkent (dated back 4<sup>th</sup> century), Turkistan (dated back between 2<sup>nd</sup> and 3<sup>rd</sup> century) and a root with the ethnic origin of Turkish-speaking people from China to the Balkan peninsula and ex-soviet peoples. Also, with its 2 populous neighbors, China and India, Kazakhstan could use its attractions to market products to these 2 countries and develop new products to focus on these markets, for example, China is one of the top countries regarding outbound tourism globally. The recent cooperation accords with China and the announcement of the Kazakhstan year in China in 2024 anticipate an extra 500.000 Chinese tourists to visit Kazakhstan in 2024. The number of Indian tourists utilizing accommodation facilities in Kazakhstan is reported to have more than tripled by the end of 2023, rising from 19,000 in 2022 to 54,000 in 2023, 178.57%. This significant rise is attributed to the availability of daily direct flights, a 14-day visa-free regime introduced for Indian citizens in 2022, and the growing interest in Kazakhstan as a tourist destination [1]. Besides, Kazakhstan had a large German population at the beginning of 1990, reaching almost 1 million left progressively and a small fraction of the German population have lived there, which could trigger nostalgia tourism.

To develop a more lucrative tourism industry, the lodging industry of Kazakhstan needs first-hand assessments to speed up reaching a competitive hotel industry for inbound and outbound tourists, even though several challenges accompany it. The rapid changes occurring in information and mobile technology turn each customer into an evaluator and put much more stress on the whole hotel industry to manage effectively. These first-hand assessments, called customer reviews, are valuable to accelerate the self-development of the hotel industry if used and analyzed properly. For this purpose, hotels in Astana, the second largest city with unique architecture in Kazakhstan, are investigated based on online reviews of customers (5894 online reviews) covering 2006 through 2023. The first objective is to determine how 2, 3, 4, and 5-star hotels impact attributes and their related sentiments when Kazakhstan has no standard hotel rating system, but hotels are rated by Web 2.0 implementations. The second objective is how hotel categories in Kazakhstan diverge from those well-studied hotel attributes in the literature, given that hotel categories are not a fixed benchmark in the current big data era. Namely, how convergences and divergences occur in hotel stars regarding hotel attributes. All data is collected from TripAdvisor. Note that even though the same and different hotel categories converge in some hotel attributes, the same hotel categories diverge in the different hotel attributes. Given that hotel categories are not a piece of fixed information that can be used by customers since Web 2.0 implementations, for example, TripAdvisor assigns their stars to hotels based on customers' streamlined reviews. This analysis provides unique relations between hotel stars and corresponding hotel attributes with persistent issues or varying degrees of issues for hotel attributes.

The rest of the article is outlined as follows: Section 2 presents the literature between hotel attributes and hotel stars in which hotel stars can play a key factor to account for several investigated attributes given that hotel stars are considered fixed benchmarks. Also, hotel stars are used as mediators to establish a relationship between hotel attributes and some key attributes such as price performance and overall satisfaction. However, in the big data era, hotel stars are not fixed benchmarks for understanding these relations. Therefore, limited knowledge of how hotel stars impact hotel attributes and how they converge and diverge regarding hotel attributes is presented given that hotel stars are not a fixed piece of information. Section 3 briefly explains topic modeling and sentiment analysis with solid references. The results are presented in Section 4. The research is concluded in Section 5.

## 2. Literature Review

Hotel stars and their related services are key components when tourists assess the quality in different forms such as customers' overall satisfaction or price-performance relation. The standards and pre-set quality characteristics of hotel stars and their attributes have long been under scrutiny. One of the paid-attention attributes of hotels by customers is their star categories. Also, Web 2.0 sites, for example, TripAdvisor rate hotels by assigning their stars and sharing them with their customers. This is a unique implementation and assumes that star rating is not a fixed indicator, and is open to further assessments of customers' streamlined reviews. Hence, note that hotel stars, either not available in some countries, for example, Kazakhstan, or assigned by some regularity body, for instance, in Turkey have faced challenges to be used as an indicator for customers when decisions are made. Therefore, considering fixed hotel categories when trying to account for the effects of hotel attributes on some quality characteristics could be misleading if divergences exist between preset hotel

categories and assessed hotel categories since as a significant indicator, hotel stars are assumed to help travelers assess several aspects of hotels in advance.

Qi et al. [2] examined hotel guest perception of Indoor Environmental Quality (IEQ) as an important indicator for hotel management by using 160.000 online reviews collected from seventy 5-star hotels. They found that the 3 most underlined factors are air conditioning, noise, and humidity. Hung [3] studied the website quality of hotels to reveal the difference between online website quality and offline star ratings for hotels in Taiwan. It is a fact that the hotel star category is a comprehensive subject of investigation and is also a critical factor for customers when the hotel industry tries to relocate its services based on online reviews. Rhee and Yang [4] studied 6 hotel features, namely, value, location, sleep quality, rooms, cleanliness, and service among 4 well-known brands of chain hotels situated in the U.S.A. The issue occurs between the general rating expressed by previous guests and guests' actual experiences of hotel services due to the level of the hotel segment (hotel star classification). Zhang and Niu [5] investigated online reviews to forecast demand for hotels. They derived signals from the customers' reviews to reach complicated interrelated features. The data used contains 1, 2, 3, 4, and 5-star hotels with 15,002, 5852, 10,179, 5697, and 9892 reviews, respectively. Ding et al. [6] examined the differential roles of negative and positive reviews on the growth of branded and independent hotels. They found that positive reviews had more positive impacts on the growth of independent hotels than branded ones. Finally, higher growth rates provide extra positive reviews. On the other hand, negative reviews negatively impact the growth of independent hotels.

In general, however, star ratings have some limitations in several cases. Qiu et al. [7] claimed that they are biased. Zhang et al. [8] suggested that star ratings and sentiments do not match. For example, a high star score accompanies negative sentiment or vice versa. Sentiment analysis effectively detects customers' opinions and can be utilized as a complement or an alternative to product star ratings, Al-Natour and Turetken [9]. Jeong and Mindy Jeon [10] claimed that TripAdvisor ratings of performance attributes such as rooms, value, cleanliness, and service substantially differed when hotel characteristics were a concern, for instance, hotel stars. In a study conducted by Levy, et al. [11], one-star hotels categorized into 3 different groups in the Washington D.C. area are associated with customers' complaints using 10 popular review websites. The review complaints varied remarkedly by reviewers' purposes for travel and nationality. So, it shows that hotels even in the same star category end up with a vast varying set of complaints, implying divergences even in the same hotel category. However, the parking problem is persistent as a converged problem for all. Guillet and Law [12] claimed that prior hospitality research rarely investigated the differences between hotel stars. Hotel service quality can be highly subjective. Fernández and Bedia [13] claimed that a higher star rating is not necessarily a good hotel quality index it needs to refer to meeting customers' expectations when services are delivered. It is not obvious that a 4- to 5-star hotel is more luxurious and expensive than a 1- to 2-star one in the same location. Li et al. [14] claimed that the associations between hotel features and customers' satisfaction (positive sentiment) could be asymmetric by using the three-factor theory and mentioned that limited research systematically categorizes hotel attributes per the three-factor theory. Furthermore, limited knowledge is available on whether hotel star rating moderates the asymmetric effects of hotel attributes on customers' positive sentiments regarding features such as nationality, Manes and Tchetchik [15], and Albayrak and Caber [16]. Cser and Ohuchi [17] claimed that star rating encapsulates the ranking of hotels using quality dimensions like price, facilities, and level of service. Lee and Blum [18] found that hotels' categories between 2-2.5 are associated with location, followed by cleanliness, room, and service whereas 4-4.5 and 5-star hotels are associated with a combination of location, cleanliness, room, and service. Mid-class hotels such as 3-3.5 are associated with location, room, and service. Also, the differences exist between domestic and international visitors. Soifer, et al. [19] found that 1–2.5-, 3–3.5-, and 4–4.5-star hotels generate more positive sentiments when free Wi-Fi, pool, complimentary breakfast, free Wi-Fi, affiliation, free breakfast, and free Wi-Fi, respectively, are given to customers. However, for 5-star hotels, these do not have any impact on positive reviews. Ba et al. [20] suggested that the aggregated outcomes indicate the robust performance of 5-star hotels in 11 topics. The systematic review conducted by Mehraliyev et al. [21] in sentiment analysis suggested that five main topics have appeared in tourism and hospitality research, dominantly focused on market intelligence, which mostly copes with market and customers and their related areas, such as customer satisfaction, reviewer behavior, and engagement. To conduct sentiment analysis, market intelligence mostly focuses on deriving insights and patterns from customers. This systematic review just mentioned 1 study how how hotel star affects hotel attributes. They claimed that the hotel star category affects customers' sleep quality. Also, another perspective was discussed by Mao et al. [22], Kim et al. [23], Vagena and Papakonstantinidis [24], and Vagena and Papakonstantinidis [25] by claiming that a star rating system, as a "living organism", is obliged to adapt to the quickly changing conditions of the hotel market, considering the swift alterations in online marketing and distribution.

Nunkoo, et al. [26] claimed that 2 substantial factors that play a role in generating high satisfaction among customers toward service quality in 1 and 2-star hotels are accommodation and infrastructure, on the other hand, safety, security, and room quality are significant factors in 3-star hotels, furthermore, 4 and 5-star hotels are mostly assessed based on accommodation, waiting time, and customer interaction. Finally, the authors suggest some specific guidelines for managerial intervention steps toward improving service quality and customer satisfaction based on star types. In a study investigating hotel booking intentions that use online reviews, three factors, brand image, star category, and price, are leading factors [27]. Guests staying in 3, 4, or 5-star hotels in Lisbon were examined to assess service quality and gauge specific dimensions to determine Importance-Performance-Analysis (IPA), Mohsin et al. [28]. Sayfuddin and Chen [29] studied how customers' reviews affected hotels' revenues based on hotel stars. They found that the signaling impact of a 1-star increase boosted 2.2–3.0% in hotel monthly revenues, whereas the reputational effect of a 1-star increase rose around 1.5–2.3% in hotel monthly revenues. More diversified outcomes [30-34] can be found in the literature. These suggest that, given that hotel categories are fixed, these associations are present in the literature. However, this perspective is a one-sided outlook.

In the era of big data generated by Web 2.0 implementations that reshape the hotel industry overall, the question of hotel star categories affects hotel attributes.

The first objective is to determine how 2, 3, 4, and 5-star hotels impact attributes and their related sentiments when Kazakhstan has no standard hotel rating system, but hotels are rated by Web 2.0 implementations. The second objective is to analyze how hotel categories in Kazakhstan diverge from those well-studied hotel attributes in the literature, given that hotel categories are not a fixed benchmark in the current big data era. Namely, how convergences and divergences occur in hotel stars regarding hotel attributes.

#### 3. Methodology

#### 3.1. Topic Modeling and Sentiment Analysis

In this article, hotels in Astana, Kazakhstan, are investigated based on online reviews of customers (5894 online reviews) covering a very large period between 2006 and 2023 to determine how 2, 3, 4, and 5-star hotels are assessed by the eye of the customers.

Topic modeling is a statistical methodology to classify text data into a pre-set quantity of topics and aggregate identical customers' assessments into the same group. Related verbal statements are placed into the same group as a statistical factor analysis does. Thus, reviews will be put under relevant topics to determine the primary concerns of customers. The methods implemented to derive the pre-set quantity of topics are called Non-negative Matrix Factorization (NMF) [35] and Latent Dirichlet Allocation (LDA) [36]. In the article, the NMF approach is implemented to extract topics. Alternatively, sentiment analysis detects reviews' polarity, depicting emotional structure by assigning it to positive, neutral, or negative. The VADER technique is implemented to find each sentiment's polarity [37].

#### 3.2. Topic Modeling: Non-negative Matrix Factorization (NMF)

The NMF methodology implements weighted TF-IDF (term frequency-inverse document frequency) to texts by partitioning the word matrix into 2 lower-ranked matrices [35]. Then, the NMF separates its input into a product of a terms-topics matrix and a topics-documents matrix [36]. Before conducting the NMF, pre-processing is run on the text data to obtain TF-IDF-weighted data composed of converting all letters to lowercase letters, lemmatization, and removing whitespace. Afterward, texts are converted into number forms [37].

### 3.3. Sentiment Analysis: VADER Approach

Sentiment analysis derives emotional tone from text data, which is characterized as positive, negative, or neutral. The VADER proposed for sentiment extraction is a lexicon and rule-based type approach [38].

## 4. Results

The data available on TripAdvisor covering 2006 through 2023 is collected for Astana hotels. All data analyses are conducted by using Orange Data Analysis software. Both topic modeling and sentiment analysis are conducted to find the pivotal factors in the eyes of customers. First, the analyses are conducted based on the whole dataset to depict the whole picture. Then, the same analyses are run by covering distinct hotel star categories to determine how they differ or converge regarding attributes. Table 1 depicts that 2 topics, Services and Room have the highest probabilities 0.32 and 0.24, respectively, followed by Food and Beverage and Transportation and Location. The smallest probability belongs to the Front Desk.

Table	1.		

Topics and the most related 20 words of all reviews.

Food and Beverage		Front Desk		Rooms		Service		Transportation and Location	
Marginal Topic Prob.	0.177291	Marginal Topic Prob	0.109733	Marginal Topic Prob.	0.239014	Marginal Topic Prob.	0.321902	Marginal Topic Prob.	0.151635
Breakfast	0.0634182	Ask	0.0245762	Room	0.162497	Hotel	0.198246	Great	0.0319776
Nice	0.0366647	Room	0.0221698	Floor	0.0204173	Staff	0.0276515	Location	0.0276178
Service	0.0342823	Reception	0.0184855	Bed	0.014924	Service	0.0230285	Excellent	0.0272982
Coffee	0.0300441	Night	0.0159715	Clean	0.0142498	Business	0.0175448	View	0.0229954
Buffet	0.02384	Pay	0.0159559	Sleep	0.0130924	Star	0.013203	Astana	0.0219273
Restaurant	0.0185411	Call	0.0158592	Hotel	0.0121896	Recommend	0.0124835	City	0.0197502
Bar	0.0168785	Check	0.0148406	Bathroom	0.0121636	Nice	0.0123167	Тор	0.0191366
Recommend	0.0146763	Bad	0.0142862	Water	0.0092684	Trip	0.0092527	Service	0.0175835
Include	0.0142872	Book	0.0119221	Large	0.0091384	Price	0.0087677	Hotel	0.0172322
Food	0.0131488	Guest	0.0109905	Window	0.0089807	Guest	0.0087196	Walk	0.0155565
Need	0.011085	Minute	0.0107888	Shower	0.0084252	Comfortable	0.0073271	Center	0.0146578
English	0.0104862	Wait	0.0106895	Service	0.0076519	Free	0.0072769	High	0.0143946
Choice	0.0100857	Tenge	0.0105888	Comfortable	0.0072255	Need	0.0067146	Park	0.0124103
Morning	0.0098003	Receptionist	0.0095977	View	0.00703	Standard	0.0066049	Distance	0.0118255
Option	0.0088713	Hour	0.0091341	Need	0.006981	Quite	0.0063714	Taxi	0.0097899
Great	0.0080938	Taxi	0.0077861	Door	0.0069206	Excellent	0.006091	Shop	0.0095326
Free	0.00763	Morning	0.0076353	Quite	0.0068108	High	0.0054954	Airport	0.0094467
Hotel	0.0075607	Arrive	0.0074851	Night	0.0066816	Center	0.0054333	Min	0.0089215
Holiday	0.0075083	Hotel	0.0073335	Tv	0.0064632	Friendly	0.005389	Speak	0.0077082
Center	0.007239	People	0.0073211	Wifi	0.0063842	Guest	0.0053022	Central	0.0074597

Figure 1 depicts the sentiments of the 5 topics. Customers evaluate Service with the largest positive sentiment, followed by Room. On the other hand, Food and Beverage and Transportation and Location are assessed by almost the same quantity of positive sentiments. The least positive sentiment belongs to the Front Desk. Even though Service and Room are assessed by positive sentiments, both have neutral and negative sentiments that account for one-third and one-fifth of all sentiments, respectively, which is relatively high and means that both Service and Room need further improvements in all hotels located in Astana for the study period. This also implies that these issues currently exist. Even though the positive sentiments are almost the same regarding Food and Beverage and Transportation and Location, the number of neutral and negative sentiments in Food and Beverage are quite fewer than those in transportation and location for the study period. It implies that hotels suffer from some degree of location and transportation issues. However, Food and Beverages with the least quantity of negative and neutral sentiments differ from other attributes. The Front Desk has an almost equal number of positive and the total number of both neutral and negative sentiments for the study period. It means that customers generally are not happy with the operations of the Front Desk. Since this is the picture of the whole data set, which hotel star category leads to these evaluations is a significant question to be answered. More information regarding how these attributes will change and which hotel star group contributes more either negative or positive sentiments will be presented subsequently.

To further analyze, hierarchical clustering and correspondence analysis are conducted to picture the findings. Figures 2 and 3 depict hierarchical clustering and correspondence results for the whole set of hotels in the study period.



Figure 1.

Topic-based sentiment distributions of all reviews.



Hierarchical Clustering Analysis of All Hotels.



Correspondence Analysis of All Hotels.

Table 2 presents the topics of 2 and 3-star hotels. We combined the online reviews of 2 and 3-star hotels since each group has quite a small quantity of reviews to run the analysis. The highest probability belongs to Location. The other attributes are ranked in order with quite small differences as follows: Staff, Food and Beverage, Room and Price, and Booking, whose probabilities are 0.184, 0.18, 0.164, and 0.162, respectively.

	Topic	distributions	with the	20 most	related	words for	the 2 a	nd 3-star Hotels.
--	-------	---------------	----------	---------	---------	-----------	---------	-------------------

Food and Beverage		Location		Price and Booking		Staff		Room	
Marginal Topic Prob.	0.180955	Marginal Topic Prob.	0.306892	Marginal Topic Prob.	0.1628	Marginal Topic Prob.	0.184695	Marginal Topic Prob.	0.164196
Breakfast	0.026749	Hotel	0.135401	Price	0.122389	Hotel	0.06626	Room	0.034862
Ask	0.020803	Astana	0.135401	Place	0.01421	Staff	0.061222	Water	0.027332
Service	0.017474	Excellent	0.025076	Reception	0.012134	Nice	0.024565	Staff	0.02156
Restaurant	0.016876	Locate	0.0196	Book	0.012026	Need	0.019402	Clean	0.01903
Food	0.015772	Walk	0.01638	Ask	0.011741	People	0.019279	Breakfast	0.01682
Pay	0.014651	Center	0.012893	Cheap	0.011671	Friendly	0.017889	Night	0.014875
Quite	0.013964	Recommend	0.011109	Quality	0.011087	Pleasant	0.013719	Hotel	0.012446
Bar	0.013615	Business	0.010487	Night	0.009102	English	0.012532	Shower	0.011529
Russian	0.013031	Location	0.010097	Open	0.008945	Seem	0.012257	Bathroom	0.01088
Eat	0.011903	Price	0.010063	Tenge	0.00886	General	0.011777	Location	0.010466
Choice	0.009959	Book	0.009736	General	0.008645	Reception	0.011237	Friendly	0.009965
Delicious	0.009728	Great	0.008651	High	0.008402	Open	0.010736	Cold	0.009777
Perfect	0.009688	Staff	0.008547	Option	0.007926	Arrive	0.010448	Need	0.009045
Plate	0.00965	Guest	0.007898	Cost	0.007566	Late	0.009132	Ac	0.008782
Value	0.009338	Ask	0.007735	Include	0.007547	Speak	0.009114	Comfortable	0.008681
Review	0.009078	City	0.007674	Lady	0.007485	Problem	0.008618	Bed	0.008322
Bring	0.008781	Place	0.007544	Online	0.006904	Enough	0.00839	Convenient	0.008189
Cafe	0.008419	Choose	0.006819	Extra	0.006887	Arrival	0.007663	Bad	0.008085
Order	0.008338	Include	0.006722	Ready	0.006877	Plenty	0.007586	Ask	0.007725
Tasty	0.008303	Trip	0.006625	Shift	0.006489	Lack	0.007428	Speak	0.007703

Figure 4 depicts the sentiments of 5 topics. Customers assess Food and Beverage with the highest number of positive sentiments, followed by Location. The rest of the topics have almost the same number of positive sentiments. Even though Food and Beverages have the highest number of positive sentiments, one-fifth of the sentiments are composed of neutral and negative sentiments, which means that hotels with 2 and 3-star should focus on Food and Beverages to improve the current conditions and solve the issues related to neutral and negative sentiments. For the rest of the topics, the same pattern is observed with a range of one-third to one-fifth of online reviews having neutral and negative sentiments. Hotels with 2 and 3 stars in Astana should focus on a comprehensive perspective to improve the current conditions and solve the issues stemming from neutral and negative sentiments. As a result, even though the number of positive sentiments is larger than both negative and neutral sentiments in Astana hotels, each topic needs further improvements.



Figure 4.

The sentiments of topics for 2 and 3-star hotels.

More statistically verifiable findings are also provided in Figures 5 and 6.



Figure 5.

Hierarchical Clustering Analysis of 2-3 Star Hotels.



**Figure 6.** Correspondence Analysis of 2-3 Star Hotels.

Table 3 depicts that Location has the largest probability, followed by Front Office and Welcoming and Staff Issues and Hotel Facilities. The other topics, Rooms and Breakfast and Price and Booking have the same probabilities around 0.15. It is observed that topics are slightly different than 2 and 3-star hotels and their importance of them is also changed. Some topics are merged into a new topic with additional attributes. It means that as the hotel star class increases the topic grows more related and interrelated in Astana hotels for the study period.

Table 3.	
Word Distributions of Topics	of 4-St

Word Distributions of Topics of 4-Star Hotels.										
<b>Rooms and Breakfast</b>		Staff Issues and Hotel		Location		Front Office	e and	Price and Transportation		
		Facilities				Welcoming				
Marginal Topic	0.155081	Marginal topic	0.195035	Marginal Topic	0.292148	Marginal Topic	0.204449	Marginal Topic	0.152914	
Probability		probability		Probability		Probability		Probability		
Breakfast	0.038964	Staff	0.061929	Astana	0.028359	Reception	0.138989	hotel	0.048617	
Room	0.032249	Restaurant	0.031081	location	0.027422	Ask	0.032679	Service	0.018531	
Bed	0.029377	Friendly	0.030453	City	0.01887	Service	0.015334	Nice	0.017605	
Night	0.024642	Nice	0.028971	center	0.01881	Call	0.01338	Price	0.016037	
Water	0.024021	Service	0.022619	clean	0.016637	Window	0.01214	Business	0.015348	
Clean	0.023915	Clean	0.020608	locate	0.01422	Wait	0.00895	Pay	0.014157	
Suitable	0.018209	Pool	0.019453	excellent	0.012669	Check	0.00792	Ask	0.013042	
Comfortable	0.015135	Helpful	0.019426	comfortable	0.012515	Tenge	0.007876	Sound	0.012669	
Bathroom	0.01315	English	0.019069	business	0.011044	Book	0.007291	Taxi	0.012329	
Shower	0.011321	Astana	0.018713	great	0.010304	People	0.007001	Walk	0.01226	
Floor	0.010309	Speak	0.017748	trip	0.00968	Receptionist	0.006827	Minute	0.012126	
Quite	0.008674	Bar	0.013987	View	0.008053	Change	0.006514	Night	0.011197	
Morning	0.008456	Recommend	0.013785	price	0.007974	Girl	0.006363	Check	0.010595	
Towel	0.00802	Place	0.012149	walk	0.007945	View	0.006353	Airport	0.009896	
Bad	0.007907	Spa	0.01052	helpful	0.007566	Bad	0.006212	Need	0.009646	
Place	0.006945	Thank	0.01033	nice	0.007438	Move	0.005991	Money	0.008135	
Need	0.006943	Offer	0.009945	park	0.007268	Return	0.005881	Street	0.008036	
Dinner	0.006753	Sauna	0.009505	close	0.006936	Arrive	0.005733	Problem	0.007517	
Buffet	0.00649	Buffet	0.00849	high	0.006722	Tired	0.005702	Call	0.007311	
Coffee	0.00631	Swimming	0.008258	taxi	0.006498	Check-in	0.005689	Wait	0.007148	

Figure 7 depicts that Staff Issues and Hotel Facilities have the largest quantity of positive sentiments, followed by Location. However, each attribute has quite a lot of neutral and negative sentiments. Also, except for Staff Issues and Hotel Facilities, the other attributes have a total number of neutral and negative sentiments close to the total number of positive sentiments, which implies that more improvements need to be conducted to satisfy customers. It is observed that as the attributes become more complicated, which means that more attributes are represented together under a common title, the total number of neutral and negative attributes increases and are almost equal to the total number of positive sentiments in general. The relatively worst attribute is Front Desk and Welcoming, since the total number of neutral and negative sentiments is larger than the total number of positive sentiments.



Figure 7.

Topic and Sentiment distribution for the 4-star hotels.

Both hierarchical clustering and correspondence analysis are presented to further verify the findings.



Figure 8.

Hierarchical Clustering Analysis of 4-Star Hotels.



Correspondence Analysis of 4-Star Hotels.

Table 4 summarizes that Quality Experience has the largest probability, followed by Service and Entertainment with 0.23 and 0.22. Location and Facilities and Room and Housekeeping have almost the same probabilities of 0.19. The least probability, which is 0.15, belongs to Food and Beverage. It is observed that as the hotel star category increases, the topics get more complicated and are represented by more words such as Service and Entertainment, Location and Facilities. Also, conceptual, for example, Quality Experience, emerges to define a set of terms or verbal statements used by customers. Also, topic probabilities converge, which implies that as the hotel star category increases, all aspects of the hotel industry are demanded by customers, which means that hotel attributes are a complete set of customers' expectations. This is only satisfied by 5-star hotels, partially in Astana for the study period.

Service and Entertainment		Location and Facilities		Rooms and Housekeeping		Food and Beverage		Quality Experience	
Marginal	0.223767	Marginal	0.191193	Marginal	0.191581	Marginal	0.150875	Marginal	0.242234
topic prob.		Topic Prob.		Topic Prob.		Topic Prob.		Topic Prob.	
Hotel	0.101805	Astana	0.052421	room	0.071635	Food	0.075853	hotel	0.100712
Astana	0.026831	City	0.051172	floor	0.021686	Quality	0.048799	star	0.08965
Service	0.020405	Business	0.046166	nice	0.021002	Bar	0.027671	staff	0.022635
Staff	0.018979	Pool	0.045564	View	0.019871	Buffet	0.027628	comfortable	0.015802
Floor	0.013873	Play	0.023516	night	0.014013	Facility	0.025728	excellent	0.013655
View	0.013099	Spa	0.023084	shower	0.012287	Wonderful	0.021417	business	0.013086
Restaurant	0.01234	Location	0.022195	bathroom	0.012171	Perfect	0.020294	lobby	0.01243
Food	0.011739	View	0.020487	clean	0.010988	Lunch	0.020214	price	0.011203
City	0.009072	Large	0.020374	reception	0.010862	Spend	0.018759	need	0.010351
Club	0.00862	Restaurant	0.020141	ask	0.010036	Include	0.016606	friendly	0.010201
Night	0.007735	Gym	0.0152	book	0.008864	Travel	0.014977	guest	0.009564
Lounge	0.007604	Center	0.01477	lobby	0.00768	Delicious	0.014857	spacious	0.008553
Guest	0.00731	Walk	0.013801	bed	0.007549	Recommend	0.014459	recommend	0.008485
High	0.007216	Swimming	0.012078	pay	0.007291	Eat	0.013296	trip	0.007883
Amazing	0.007169	Bar	0.011234	change	0.007114	Provide	0.012304	helpful	0.007139
Friendly	0.007097	Place	0.01111	quite	0.006939	Beautiful	0.01091	pleasant	0.00706
Level	0.007004	Lobby	0.010827	City	0.006856	Free	0.008761	choice	0.007057
Quality	0.006908	Facility	0.010711	large	0.006752	Atmosphere	0.008721	conference	0.006799
Help	0.00688	Sauna	0.010493	standard	0.0067	Coffee	0.007278	include	0.006728
Provide	0.006639	Central	0.00884	door	0.006535	Lobby	0.007271	convenient	0.006225

 Table 4.

 Word Distributions of Topics of 5-Star Hotels.

Figure 4 depicts that Service and Entertainment has the highest number of positive sentiments, followed closely by Room and Housekeeping and Quality Experience. Also, the negative and neutral sentiments are very few except for Quality Experience, which means that even though 5-star Hotels improves several hotel-related attributes to satisfy their customers, a more comprehensive form of attributes remains weaker when compared to singly and coupled defined attribute sets. This shows that customer expects more refined products, however, 5-star hotels failed partially to provide their customer in Astana for the study period. It means that more research and improvement should be conducted to pinpoint those specific issues. Nevertheless, Location and Facilities and food and Beverage have comparatively fewer positive sentiments when compared to other attributes. Location attribute is mostly articulated in all-star categories by customers.



Figure 10.

Topic and Sentiment Distribution of 5-Star Hotels.



Hierarchical Clustering Analysis of 5-Star Hotels.



Correspondence Analysis of 5-Star Hotels.

#### 5. Discussion and Conclusion

In the first half of this section, we will present how each hotel star category is associated with negative, neutral, and positive sentiments of customers regarding derived topics from 5,894 reviews from 2006 through 2023. Hotel star categories are a substantial indicator that has been investigated regarding several aspects of the hotel industry, namely, customer satisfaction, price-service performance, profit and price relations, online booking intentions, sustainability, waste management and disposal, green perspective, revisit intentions, and so on.

The research firstly focuses on how the hotels in Astana were assessed in the eyes of customers to pinpoint what main attributes are pertinent to 2, 3, 4, and 5-star hotels and what types of sentiments are observed in the long run.

When the sentiment of customers covering the whole hotel star category is under investigation, customers evaluate Service with the largest positive sentiment, followed by Room. On the other hand, Food and Beverage and Transportation and Location are assessed by almost the same quantity of positive sentiments. The least positive sentiment belongs to the Front Desk. Even though Service and Room are assessed by positive sentiments, both have neutral and negative sentiments that account for one-third and one-fifth of all sentiments in these topics, respectively, which is quite relatively high and means that both Service and Room need further improvements in all hotel star categories located in Astana. Besides, the number of positive sentiments is almost the same as the combination of neutral and negative sentiments regarding Food and Beverage, and Transportation and Location. The number of neutral and negative sentiments in Food and Beverage are quite fewer than those in Transportation and Location. It implies that hotels suffer from varying degrees of Location and Transportation issues. However, Food and Beverages with the least quantity of negative and neutral sentiments differ from other attributes. The Front Desk has an almost equal number of positive and the total number of both neutral and negative sentiments. It means that customers generally are not happy with the operation of the Front Desk.

In brief, hotel star categories lead to single or coupled issues such as Room, Service, Front Desk, Transportation and Location, and Food and Beverage in the aggregated data.

In the second half of the section, we examined how hotel star categories affect hotel attributes. For 2- and 3-star hotels, the 5 topics that customers assess are Room, Staff, Location, Food and Beverages, and Price and Booking. Customers evaluate Food and Beverages with the highest number of positive sentiments, followed by Location. The rest of the topics have almost the same number of positive sentiments. Even though Food and Beverages have the highest number of positive sentiments, one-fifth of the sentiments are composed of neutral and negative sentiments, which means that hotels with 2 and 3 stars should focus on Food and Beverages to improve the current conditions and address the issues related to neutral and negative sentiments. For the rest of the topics, the same pattern is observed, with a range of one-third to one-fifth of online reviews having neutral and negative sentiments. Both 2- and 3-star hotels in Astana satisfy the expectations of customers in Food and Beverages and Location. However, both 2- and 3-star hotels cannot cope with issues pertinent to Room, Staff, and Price and

Booking, which cause negative sentiments. Even though changing locations is almost impossible for 2- and 3-star hotels in the long run, the landscape of hotels can be improved, and better transportation options can be suggested to customers as an aid. Note that single issues or closely related coupled issues appear to be problematic. Thus, 2- and 3-star hotels cause negative or positive sentiments regarding single or very closely related coupled issues. The divergence mostly stems from Price and Booking, which is not seen in hotel categories in Astana nor observed as a globally persistent problem.

The 4-star hotels suggest that Staff and Hotel Facilities have the largest quantity of positive sentiments, followed by Location. However, each attribute has quite a lot of neutral and negative sentiments. Also, except for Staff and Hotel Facilities, the other attributes have a total number of neutral and negative sentiments close to the total number of positive sentiments. The relatively worst attribute is Front Desk and Welcoming since the total number of neutral and negative sentiments is larger than the total number of positive sentiments. Note that as the hotel star category increases, the number of hotel attributes attached to positive sentiments increases with the cost of an almost equal number of neutral and negative sentiment numbers. However, as the number of topics with positive sentiments increases, the cost of having an almost equal number of neutral and negative sentiments occurs, which means that the divergence between 2 and 3 stars, and 4-star hotels is not big enough to generate more positive sentiments even though some attributes are different. The divergence mostly occurs in Staff and Hotel Facilities when compared to 2 and 3-star hotels, which is not a coherent outcome from the global perspective. Generally, 4-star hotels are better equipped than 2 and 3-star hotels and the difference between them is quite straightforward. However, Location remains a persistent problem, which is not globally observed.

The 5-star hotels suggest that Service and Entertainment have the highest number of positive sentiments, followed closely by Room and Housekeeping and Quality Experience. Also, the negative and neutral sentiments are very few except for Quality Experience, which means that even though 5-start hotels improve several hotel-related attributes to satisfy their customers, a more comprehensive form of attributes remains weaker, for example, Quality Experience. This shows that customer expects more refined and integrated products, however, 5-star hotels failed partially to provide them to their customer, which diverges from the rest of the hotel market globally. Nevertheless, Location and Facilities and Food and Beverage have comparatively fewer positive sentiments when compared to other attributes, which also diverges from the rest of the hotel have a harmony of some attributes such as Location, Quality Assessment, Facilities, and Rooms, namely, customers' preferences cover a large set of attributes that need to be satisfied and paid attention to by managers of the hotels. Hence, the perspective of customers is almost totally different than the customers staying in 2 and 3-star hotels and 4-star hotels. Interestingly, Facilities are still weaker than what customers expect in 5-star hotels. It is a fact that the quality of Rooms is superior when compared to other star-type hotels. The divergence still exists for 5-star hotels in Astana by not providing a globally accepted service.

Therefore, managers of 2 3-star and 4-star hotels should pay attention to improving the quality of Rooms, Transportation, Infrastructure, and the Front Desk. Besides, 4-star hotels should differentiate themselves from 2 and 3-star hotels by improving globally underlined attributes.

It is a fact that as the hotel star category increases, the quality of the hotel attributes increases in Astana. The main divergences are Location, Room, and Infrastructure for 2, 3, and 4-star hotels. Transportation is a significant factor for customers for all 2, 3, and 4-star hotels. To solve this issue, since it is not directly related to hotels, professional help should be provided by the hotels to customers to show them how to directly and cheaply commute. The landscape around 2 and 3-star hotels should be improved. 2, 3, and 4-star hotels lead to similar problems with varying degrees, which is not a common situation globally.

Even though hotel star categories are different, some attributes persistently exist in all, for example, Room, Location, and Staff. Issues appear singly in 2 and 3-star categories, however, as the hotel star category increases, they are adjunct to other attributes that are related or interrelated, which implies managerial problems should be taken into consideration.

The limitation of the research is that the whole analysis is just dependent upon one source of the collected data, which is TripAdvisor. More comprehensive research can be conducted by adding more online reviews coming from other data sources.

Future research will focus on using more comprehensive data sources, expanding the period to cover more data if possible, and separating the sentiments of foreign and local tourists to uncover customer segments.

## References

- [1] Travel Trends Today, "Indian tourist arrivals to Kazakhstan triple to 540n in 2023," Retrieved: https://www.traveltrendstoday.in/indian-tourist-arrivals-to-kazakhstan-triple-to-540n-2023/, 2022.
- [2] M. Qi, X. Li, E. Zhu, and Y. Shi, "Evaluation of perceived indoor environmental quality of five-star hotels in China: An application of online review analysis," *Building and Environment*, vol. 111, pp. 1-9, 2017. https://doi.org/10.1016/j.buildenv.2016.09.027
- [3] C.-L. Hung, "Online positioning through website service quality: A case of star-rated hotels in Taiwan," *Journal of Hospitality and Tourism Management*, vol. 31, pp. 181-188, 2017. https://doi.org/10.1016/j.jhtm.2016.12.004.
- [4] H. T. Rhee and S.-B. Yang, "Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers' overall ratings," *Computers in Human Behavior*, vol. 50, pp. 576-587, 2015. https://doi.org/10.1016/j.chb.2015.02.069.
- [5] D. Zhang and B. Niu, "Leveraging online reviews for hotel demand forecasting: A deep learning approach," *Information Processing & Management*, vol. 61, no. 1, p. 103527, 2024. https://doi.org/10.1016/j.ipm.2023.103527
- [6] X. Ding, B. Gao, and S. Liu, "Understanding the interplay between online reviews and growth of independent and branded hotels," *Decision Support Systems*, vol. 152, p. 113649, 2022. https://doi.org/10.1016/j.dss.2021.113649
- [7] L. Qiu, J. Pang, and K. H. Lim, "Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence," *Decision Support Systems*, vol. 54, no. 1, pp. 631-643, 2012. https://doi.org/10.1016/j.dss.2012.08.020

- [8] Z. Zhang, Q. Ye, Z. Zhang, and Y. Li, "Sentiment classification of Internet restaurant reviews written in Cantonese," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7674-7682, 2011. https://doi.org/10.1016/j.eswa.2010.12.147
- [9] S. Al-Natour and O. Turetken, "A comparative assessment of sentiment analysis and star ratings for consumer reviews," *International Journal of Information Management*, vol. 54, p. 102132, 2020. https://doi.org/10.1016/j.ijinfomgt.2020.102132
- [10] M. Jeong and M. Mindy Jeon, "Customer reviews of hotel experiences through consumer generated media (CGM)," *Journal of Hospitality & Leisure Marketing*, vol. 17, no. 1-2, pp. 121-138, 2008.
- [11] S. E. Levy, W. Duan, and S. Boo, "An analysis of one-star online reviews and responses in the Washington, DC, lodging market," *Cornell Hospitality Quarterly*, vol. 54, no. 1, pp. 49-63, 2013. https://doi.org/10.1177/1938965512464513
- [12] D. B. Guillet and R. Law, "Analyzing hotel star ratings on third-party distribution websites," International Journal of Contemporary Hospitality Management, vol. 22, no. 6, pp. 797-813, 2010. https://doi.org/10.1108/09596111011063098
- [13] M. C. L. Fernández and A. M. S. Bedia, "Is the hotel classification system a good indicator of hotel quality?: An application in Spain," *Tourism Management*, vol. 25, no. 6, pp. 771-775, 2004.
- [14] H. Li, Y. Liu, C.-W. Tan, and F. Hu, "Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews," *International Journal of Contemporary Hospitality Management*, vol. 32, no. 5, pp. 1713-1735, 2020. https://doi.org/10.1108/IJCHM-06-2019-0581
- [15] E. Manes and A. Tchetchik, "The role of electronic word of mouth in reducing information asymmetry: An empirical investigation of online hotel booking," *Journal of Business Research*, vol. 85, pp. 185-196, 2018. https://doi.org/10.1016/j.jbusres.2017.12.019
- [16] T. Albayrak and M. Caber, "The symmetric and asymmetric influences of destination attributes on overall visitor satisfaction," *Current Issues in Tourism*, vol. 16, no. 2, pp. 149-166, 2013. https://doi.org/10.1080/13683500.2012.682978
- [17] K. Cser and A. Ohuchi, "World practices of hotel classification systems," *Asia Pacific Journal of Tourism Research*, vol. 13, no. 4, pp. 379-398, 2008.
- [18] H. Lee and S. C. Blum, "How hotel responses to online reviews differ by hotel rating: An exploratory study," *Worldwide Hospitality and Tourism Themes*, vol. 7, no. 3, pp. 242-250, 2015. https://doi.org/10.1108/WHATT-03-2015-0016
- [19] I. Soifer, E.-K. Choi, and E. Lee, "Do hotel attributes and amenities affect online user ratings differently across hotel star ratings?," *Journal of Quality Assurance in Hospitality & Tourism*, vol. 22, no. 5, pp. 539-560, 2021. https://doi.org/10.1080/1528008X.2020.1814935
- [20] H. Ba, V.-M. Huynh, and T. Tra, "Measuring hotel review sentiment: An aspect-based sentiment analysis approach. Seki H, Nguyen, C. H., Huynh, V-N, Inuiguchi M," in *Integrated Uncertainty in Knowledge in Modeling and Decision Making, 7th International Symposium, IUKM 2019, Nara, Japan, March 27-29, Proceedings, Springer, 2019*
- [21] F. Mehraliyev, I. C. C. Chan, and A. P. Kirilenko, "Sentiment analysis in hospitality and tourism: A thematic and methodological review," *International Journal of Contemporary Hospitality Management*, vol. 34, no. 1, pp. 46-77, 2022. https://doi.org/10.1108/IJCHM-02-2021-0132
- [22] Z. Mao, Y. Yang, and M. Wang, "Sleepless nights in hotels? Understanding factors that influence hotel sleep quality," *International Journal of Hospitality Management*, vol. 74, pp. 189-201, 2018. https://doi.org/10.1016/j.ijhm.2018.05.002
- B. Kim, S. Kim, B. King, and C. Y. Heo, "Luxurious or economical? An identification of tourists' preferred hotel attributes using [23] scaling," Journal of Vacation Marketing, vol. 25, no. 2, best-worst pp. 162-175. 2019 https://doi.org/10.1177/1356766718757789
- [24] A. Vagena and L. Papakonstantinidis, "Characteristics of official hotel classification systems," *International Journal of Research in Tourism and Hospitality*, vol. 6, no. 3, pp. 33-49, 2020. https://doi.org/10.20431/24550043.0603004
- [25] A. Vagena and L. Papakonstantinidis, "Creation of a model hotel classification system," *International Journal of Research in Tourism and Hospitality*, vol. 6, no. 3, pp. 50-61, 2020. https://doi.org/10.20431/24550043.0603005.
- [26] R. Nunkoo, V. Teeroovengadum, C. M. Ringle, and V. Sunnassee, "Service quality and customer satisfaction: The moderating effects of hotel star rating," *International Journal of Hospitality Management*, vol. 91, p. 102414, 2020. https://doi.org/10.1016/j.ijhm.2019.102414
- [27] O. A. El-Said, "Impact of online reviews on hotel booking intention: The moderating role of brand image, star category, and price," *Tourism Management Perspectives*, vol. 33, p. 100604, 2020. https://doi.org/10.1016/j.tmp.2019.100604
- [28] A. Mohsin, H. Rodrigues, and A. Brochado, "Shine bright like a star: Hotel performance and guests' expectations based on star ratings," *International Journal of Hospitality Management*, vol. 83, pp. 103-114, 2019. https://doi.org/10.1016/j.ijhm.2019.04.012
- [29] A. Sayfuddin and Y. Chen, "The signaling and reputational effects of customer ratings on hotel revenues: Evidence from TripAdvisor," *International Journal of Hospitality Management*, vol. 99, p. 103065, 2021. https://doi.org/10.1016/j.ijhm.2021.103065
- [30] V. Tiwari and A. Mishra, "The effect of a hotel's star-rating-based expectations of safety from the pandemic on during-stay experiences," *Journal of Retailing and Consumer Services*, vol. 71, p. 103223, 2023. https://doi.org/10.1016/j.jretconser.2022.103223
- [31] B. Walheer, L. Zhang, and Y. Luo, "Bidirectional technological spillover in the Chinese star-rated hotel sector: An empirical investigation," *Economic Modelling*, vol. 86, pp. 210-226, 2020. https://doi.org/10.1016/j.econmod.2019.06.013
- [32] T.-Y. Hsiao, C.-M. Chuang, and L. Huang, "The contents, determinants, and strategic procedure for implementing suitable green activities in star hotels," *International Journal of Hospitality Management*, vol. 69, pp. 1-13, 2018. https://doi.org/10.1016/j.ijhm.2017.10.005
- [33] E. T. Ampofo, "Mediation effects of job satisfaction and work engagement on the relationship between organisational embeddedness and affective commitment among frontline employees of star-rated hotels in Accra," *Journal of Hospitality and Tourism Management*, vol. 44, pp. 253-262, 2020. https://doi.org/10.1016/j.jhtm.2020.06.002
- [34] E. Martin-Fuentes, "Are guests of the same opinion as the hotel star-rate classification system?," *Journal of Hospitality and Tourism Management*, vol. 29, pp. 126-134, 2016. https://doi.org/10.1016/j.jhtm.2016.06.006
- [35] A. Obadimu, E. Mead, and N. Agarwal, "Identifying latent toxic features on YouTube using non-negative matrix factorization," presented at the The Ninth International Conference on Social Media Technologies, Communication, and Informatics, IEEE, Valencia, Spain, 2019.

- [36] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. Jan, pp. 993-1022, 2003.
- [37] Y. Chen, H. Zhang, R. Liu, Z. Ye, and J. Lin, "Experimental explorations on short text topic mining between LDA and NMF based Schemes," *Knowledge-Based Systems*, vol. 163, pp. 1-13, 2019. https://doi.org/10.1016/j.knosys.2018.08.011.
- [38] J. Wu and T. Yang, "Service attributes for sustainable rural tourism from online comments: Tourist satisfaction perspective," *Journal of Destination Marketing & Management*, vol. 30, p. 100822, 2023. https://doi.org/10.1016/j.jdmm.2023.100822