

N.M. Abdirazakov^{*1}, Z.T. Kushebayev², F.A. Numanova³, P.Z. Orynbet⁴, G.E. Myltykbayeva⁵

¹*International University of Tourism and Hospitality, Kazakhstan;*

²*University of International Business, Kazakhstan;*

^{3,5}*JSC "National Center of Science and Technology Evaluation", Kazakhstan;*

⁴*Turan University, Kazakhstan*

¹*nurzhan.abdirazakov@gmail.com, ²kushebayev@gmail.com, ³firuza_numanova@mail.ru,*

⁴perizat.orynbet@mail.ru, ⁵mgauhar2809@gmail.com

¹*https://orcid.org/0000-0002-6733-2873, ²https://orcid.org/0000-0002-5243-2222,*

⁴https://orcid.org/0000-0001-5442-7913

Understanding the relationships between hotel variables in Almaty, Kazakhstan: an investigation using booking.com data

Abstract

Object: The purpose of this article is to study relationships between hotel variables such as star ratings, prices, customer ratings, distance from the city center, and review numbers for hotels in Almaty, Kazakhstan.

Methods: Data for this study was collected quantitatively from Booking.com using the web scraping tool Parsehub, a popular tool in contemporary research for web data extraction. Following data collection and the removal of missing data, 69 cases out of 96 were retained. To address the research questions, the study uses descriptive statistics to analyze data by star rating categories, performs One Way ANOVA test to determine differences between variables, and conducts correlation analysis. The analysis tools used include Jamovi and SPSS 23.

Findings: The findings indicate significant differences in customer ratings based on star ratings, a positive correlation between star ratings and prices, a weak, negative relationship between customer ratings and distance from the city center, and a weak positive correlation between review numbers and customer ratings.

Conclusions: These insights contribute to a better understanding of the hotel industry in Almaty. Results demonstrate the importance of star ratings, customer reviews, price, and location in understanding the dynamics of the hotel market.

Keywords: Kazakhstan, Tourism, Booking.com, OTAs, Hotels, Destination, Hospitality.

Introduction

Tourism can be an important sector for economy of Kazakhstan, with the potential to contribute significantly to the country's growth and development. However, as a number of performance indicators show, tourism sector of Kazakhstan is currently just beginning to grow. The World Bank (2022) reports that although 8.5 million tourists in Kazakhstan spent \$2.9 billion in 2019, average spending per person was only \$343, which is lower than expected compared to countries with well-developed tourism economies. Additionally, average occupancy was low at 25%, although the number of properties increased by 220% between 2010 and 2019. As of June 2022, the average hotel occupancy rate in the Republic of Kazakhstan was only 22.8% (Bureau of National Statistics, 2022).

As the largest city and former capital of Kazakhstan, Almaty is a primary destination for tourists visiting the country. In 2019, 1.3 million arrivals were recorded in Almaty, of which 898,000 were domestic tourists and 435,000 were inbound tourists (Pchelyanskaya, 2020). Most of these tourists come from neighboring countries such as Russia, China, Uzbekistan, and Kyrgyzstan (ibid, 2020). However, despite the high number of tourists, Almaty's tourist accommodation facilities remain largely unclassified and provide minimal services, with a total of 335 facilities and 20,669 beds available as of the first half of 2020 (Kazakh Tourism, 2022). The occupancy rate in Almaty's accommodation places was only 29.2%, indicating the need for further development and improvement in the city's tourism industry. Despite this fact, it is important to note that there is intense competition among hotels in this city. This can be attributed to the fact that Almaty is the largest and most prominent business center in Kazakhstan, attracting both domestic and international business travelers, as well as tourists. As such, hotels in Almaty are constantly striving to improve their services and meet the needs of their guests to remain competitive in the market. It has become crucial for them to understand customers' needs and preferences to improve their services and stand out from their competitors.

* Corresponding author. E-mail address: *nurzhan.abdirazakov@gmail.com*

Booking.com, one of the world's largest online travel agencies, provides a platform for customers to review and rate hotels they have stayed in. These reviews and ratings can be used as valuable sources of information for hotel managers to assess their performance and make necessary improvements. In recent years, there has been a surge in scholarly interest towards the analysis of hotels and tourism through online travel agencies (OTAs) such as booking.com (Agušaj et al., 2017; Díaz & Rodríguez, 2018; Mariani & Borghi, 2018; Martin-Fuentes et al., 2018; Moro et al., 2018; Figini et al., 2020; Borges-Tiago et al., 2021; Ruytenbeek et al., 2021; Rita et al., 2022). This trend can be attributed to the growing importance of OTAs in the tourism industry, which has resulted in increased research efforts aimed at understanding the impacts of these platforms on various aspects of the sector, including consumer behavior, marketing strategies, and distribution channels. Scholars have recognized the potential of OTA data for conducting large-scale empirical studies that can inform industry stakeholders and policymakers on key issues related to tourism and hospitality. As such, research in this area is likely to continue to receive significant attention in the future.

The references provided above examine various aspects of the impact of online travel agencies (OTAs) and Booking.com in the tourism industry. Some studies investigate the impact of the rating systems on customer sentiment and tone of voice, and others compare the differences in branding co-creation between TripAdvisor and Booking.com. There are also studies that analyze the reliability and validity of online reputation databases for lodging, the influence of online ratings on hotel pricing power, and the factors that affect hotels' online prices. Overall, the studies demonstrate the significance of OTAs and Booking.com in the tourism industry and provide insights into the different aspects of their influence on travelers' decision-making processes. It should be noted that the studies mentioned were conducted in various locations worldwide, and it's probable that no research has been carried out on this specific subject in Almaty or even in Kazakhstan. Thus, conducting a pioneering exploratory or descriptive research on the impact of Booking.com in the tourism industry in Almaty or Kazakhstan would be beneficial. Such study will fill the existing gap in literature and provide significant insights for local practitioners in the hospitality and tourism industry to effectively utilize the data from Booking.com in the tourism market.

Thus, the purpose of this study is to explore and comprehend the links between factors for hotels in Almaty, Kazakhstan, including star ratings, prices, customer ratings, and distance from the city center. Through the examination of these variables, this study will add to our present understanding of Kazakhstan's local tourist business by offering insights into the tastes and behaviors of hotel guests in Almaty. This research has important applications for hotel managers and owners who want to become more competitive in the market. They will be better able to comprehend the preferences of their clientele and make defensible choices about cost, location, and level of service according to the study's findings.

Literature review

Tourism is one of the most significant sectors of the world's economy, having a significant impact on employment and socio-economic development (Holjevac, 2003). The competitiveness of a destination directly affects the income generated by visitor numbers and spending, as well as related industries such as hotels and retail (Tsai et al., 2009). In other words, destination competitiveness and its impact on the tourism industry and related businesses such as hotels and retail are crucial factors for the economic and social development of a region or country. The hotel industry, in particular, plays a vital role in the socio-economic and cultural development of a destination (Go et al., 1994). The performance of hotels is affected by social and economic factors, and at the same time, the effectiveness of the hotel industry is significantly influenced by the overall tourism development of a destination (Zhang et al., 2020). In addition, economic status and tourism development have a strong impact on the overall performance of hotels, where hotels can enhance their occupancy rate and sales income when tourism and economic development of a certain country shows positive performance (Chen, 2010). On the contrary, a country's competitiveness is influenced by the performance of businesses, including hotel businesses (Barros, 2005; Tsai et al., 2009). Therefore, the effectiveness of the hotel industry and its contribution to the overall tourism product of a destination has a direct impact on a country's competitiveness. Overall, the literature suggests that the tourism industry, particularly the hotel industry, plays a vital role in the socio-economic and cultural development of a destination. Furthermore, the competitiveness of a destination and the effectiveness of the hotel industry are interlinked, where a positive performance in the tourism and economic development of a country can enhance the performance of the hotel industry and, in turn, contribute to the country's competitiveness.

According to Cibere, Başaran, and Kantarci (2020), business owners and managers now see a greater need for assessing business performance as a result of globalization and growing market competitiveness in

recent decades. Consultants, academics, and industry professionals use a range of operational indicators, such as revenue per available room, occupancy, total revenue per available room, average room rate, and gross operating profit per available room, to evaluate how effectively hotels run their operations (Sainaghi et al., 2013). In the last thirty years, researchers have used a variety of methods to gauge hotel performance. These strategies have included non-financial measures, individual financial strategies, and holistic strategies. Yet, there can be certain issues with using conventional financial accounts to gauge hotel performance (Atkinson, & Brown, 2001; Zigan, & Zeglat, 2010). The majority of competitive and financial metrics for hotel performance are narrow (Harrington and Akerhurst, 1996). Furthermore, when utilized as the only means of assessing performance, some indicators — like customer satisfaction and complaint levels, average daily rate and revenue per available room, and hotel occupancy — may have drawbacks (Banker et al., 2005; Tuominen, 2011; Cibere et al., 2020). According to Kantarci et al. (2017), customer-generated material offers information that might impact potential customers' decision-making process in addition to hotels.

Online review platforms have been divided into several categories by Chang et al. (2019): online travel agencies that rely on transactions, like Booking.com and Expedia, and community-based websites like Yelp, LonelyPlanet, and Tripadvisor. The combination of customer reviews generates electronic word-of-mouth in transaction-based online travel agencies (Gligorijevic, 2016; Xiang et al., 2017). These platforms are all made up of user-generated content that offers trustworthy reviews based on real experiences, assisting customers in making educated decisions. For instance, Tripadvisor integrates a variety of user data and information tools to offer resources and business models (Yoo et al., 2016).

By serving as intermediate between travelers and hotels, Booking.com and other online travel agencies (OTAs) have a big impact on the hotel and tourism sectors. These platforms make it simple for customers to search for and reserve hotels, flights, and other travel-related services, which facilitates a more accessible and convenient booking process. Booking.com, one of the largest travel marketplaces that serves both big and small business owners, enables properties worldwide to reach a worldwide audience and expand their operations (Booking.com, 2023). With over 28 million accommodation listings— including over 6.6 million unique listings of homes, apartments, and other unique places to stay— the platform is available in 43 languages. Booking.com strives to make travel comfortable for everyone, regardless of destination or activity, with round-the-clock customer service (ibid, 2023). Booking.com determines a property's overall rating using a rating system that is based on reviews from previous visitors. A visitor is requested to post a review of their stay after their visit is over. A numerical rating on a scale of 1 to 10 is included in the review for various categories, including staff, location, amenities, cleanliness, comfort, and value for money. Also, visitors can post remarks regarding their experience. The average of all the numerical ratings submitted by visitors who have stayed at the property and written reviews is used to determine the overall rating of the establishment. In order to make sure that the overall rating accurately represents the condition of the property right now, Booking.com also considers the frequency and recentness of the reviews. The property provides Booking.com with the star rating, which is typically assigned by an official hotel rating agency or another independent third party (Booking.com, 2023). The latitude and longitude of the property and the center of the city or town are usually used to calculate the distance from the center of a city or town that is displayed on Booking.com. Although the service providers set the prices that are displayed on Booking.com, the platform retains the ability to use its own resources to finance rewards and other benefits (Booking.com, 2023). In the field of tourism research, a lot of attention has been paid to studying Booking.com and online travel agencies (OTAs).

The location of a hotel, in addition to its star rating and membership in a larger hotel chain, has long been recognized in academic discourse as a critical factor in determining room rates. Moreover, certain studies have revealed a significant association between hotel room rates and the adoption of environmentally sustainable practices by the hotel's staff.

In a study conducted by Fleischer (2012), a comprehensive dataset comprising 2,819 hotel rooms during the peak season and 2,406 rooms during the off-peak season was meticulously gathered from Booking.com. The findings of this investigation revealed a striking result: hotel room rates exhibited a notable increase of approximately 10% when they boasted a scenic view, as opposed to their counterparts lacking this particular specification. Research indicates that tour prices differ among various tour operators due to market segmentation. A study conducted by Espinet, Saez, Coenders, and Fluvia (2003) revealed significant price variations between four-star hotels and others. They also noted that factors like location, hotel size, proximity to the beach, and the availability of parking had a significant impact on prices. It was highlighted that attributes like the type of tour operator, location, breakfast, restaurant, TV, room type, and distance from

the beach played a significant role in determining tour prices (Thrane, 2005). Haroutunian, Mitsis, and Pashardes (2005) found that price variation was often more influenced by the destination country and the tour operator rather than the specific attributes of the holiday package. Furthermore, Israeli's (2002) hedonic analysis of hotel prices in Israel discovered that star ratings consistently had a positive impact on prices, in contrast to the inconsistent impact of corporate affiliation. The research of Abrate et al. (2011) affirms the results commonly observed in the predominant hedonic price literature. The models used in this study demonstrated that both the star rating of a property and its appeal based on location contribute to explaining variations in prices. They reveal a robust connection with the pricing premium principle, where a higher price is attributed to quality signals based on reputation, irrespective of the season. These findings align with the underlying assumptions, indicating that the hotel industry's pricing strategy aims to assess a premium in correlation with the identified factors driving value.

Moreover, Hu and Chen (2016) in their findings confirmed the interaction effect between hotel star class and review rating. They state that one unexplored assumption in existing research is the idea that predictor variables do not interact with each other. However, it's possible that some predictors might have an interactive effect (ibid, 2016). For instance, two important variables, Hotel star class (HSC) and review rating (RR), are known to be linked to review helpfulness (O'Mahony & Smyth, 2010; Zhu et al., 2014; Hu & Chen, 2016, 930). HSC represents a hotel's quality with a star rating, while RR reflects how reviewers perceive hotel quality. Common intuition might suggest a positive correlation between these two variables, but that's not always the case. When a review rating doesn't align with the hotel star class (for example, a high-class hotel with a low review rating or a low-class hotel with a high review rating), it tends to attract attention, increasing the likelihood of the review being read (ibid, 2016). Consequently, the chances of these reviews receiving a helpful vote may not be the same as reviews that match the expected pattern (high RR for high HSC or low RR for low HSC). The interaction effect between RR and HSC on review helpfulness appears highly plausible, but existing literature hasn't addressed this question (ibid, 2016). Agušaj, Bazdan & Lujak (2017) stated that hotels with higher star ranking are usually higher priced. A hotel's favorable location enhances its overall worth, enabling hotel managers to command higher prices compared to similar hotels situated in less advantageous areas (Alegre et al., 2013, 421). Yang, Wong, and Wang (2012) illustrate that the hotel's location plays a pivotal role in its competitive strategy. Factors like accessibility, the clustering of amenities, and proximity to public services can contribute to increased profitability since they influence tourists' choices when selecting accommodations. Location is considered a crucial factor in the hotel industry, affecting pricing, market share, and a hotel's overall performance (Mellinas, Nicolau, & Park, 2019). Additionally, it plays a significant role in shaping the guest experience (Mellinas, Nicolau, & Park, 2019, 421).

Although researchers and practitioners often use OTA analysis to measure hotel performance, there is a severe lack of empirical studies in Kazakhstan's hotels and the wider Central Asian region. Some attempts have been made and published on scientific platforms like eLibrary.ru in the Russian language, including papers by Bayandinova and Djoldybayev (2015), Kanbayeva (2021), Duysembayev and Kayrbekova (2020), Krasnova (2018), Nurpeisova et al. (2018), Abenova et al. (2021) and Babaevskaya (2019). However, all of these papers are based on secondary data and do not employ empirical research methods to study hotel performance in Kazakhstan using Booking.com. Only one paper by Yergesh et al. (2017) was found to analyze hotel reviews of Kazakhstan on TripAdvisor, but it focused more on the linguistic features of Kazakh language than examining hotel performance based on OTA analysis. Therefore, there is a need for more empirical studies on OTA-based hotel performance in Kazakhstan and the wider Central Asian region.

Methods

To address the research questions, the study will use quantitative data collected from Booking.com. The researchers utilized a web scraping tool, Parsehub, to extract data from the website based on a defined structure. Parsehub is commonly utilized in contemporary research studies for data collection. It is commonly used in both academic and non-academic research projects to extract data from web pages, eliminating the need for complex code writing. The researchers chose neutral dates for one individual's check-in and check-out in order to illustrate the list of accommodations in Almaty. 846 lodging options with a variety of characteristics, including price, breakfast availability, and lodging name, were found through this search. As the study's primary focus was hotels, the researchers narrowed down the results to show only lodging options, yielding 96 hotels in Almaty. The author used the scraping tool to gather information about these 96 hotels on five different variables: price, star rating, average customer rating, review count, and distance from the city center. 69 cases were kept in total after the data was transferred to Jamovi and any missing information

was eliminated. As stated in the literature review, official organizations or government agencies usually bestow star ratings (Booking.com, 2023). It is possible that some hotels will not meet the requirements for a rating or will not choose to participate in the rating process. But, it's crucial to remember that a hotel's lack of stars does not always imply that it is of poor quality; in fact, there might be other aspects of the hotel that should be taken into account when examining the data.

The study will use descriptive statistics, such as the average price, average review numbers, and average customers' average overall rating, to analyze the data by categories of star rating. The researchers will also conduct One Way ANOVA test to determine significant differences between star rating categories and distance from the city center categories. In addition, the study will perform a correlation analysis. To conduct the analysis, the researchers will use Jamovi and SPSS 23 version. The tools will enable the study to provide a detailed analysis of the data extracted from Booking.com to address the research questions.

Research questions: 1) Is there a significant difference in “customers' overall ratings” between hotels with different “star ratings?” 2) What is the relationship between the “star rating” and the “price” for hotels in Almaty? 3) Do hotels with higher customer overall ratings in Almaty have a closer distance from the city center compared to lower-rated hotels? 4) Do hotels with closer “distance from the city center” have a higher “prices” in hotels of Almaty? 5) What is the relationship between the “review number” and the “customers' overall average rating” for hotels in Almaty?

Results

The Table 1 provides summary descriptive statistics for four variables: “price”, “reviews_number”, “rating”, and “distance_from_centre”, categorized by different star ratings (3, 4, and 5).

Table 1. Descriptive Statistics for Hotel Data in Almaty derived from Booking.com

Descriptives					
	star	price	reviews_number	rating	distance_from_centre
N	3	27	27	27	27
	4	34	34	34	34
	5	8	,	8	8
Missing	3	0	0	0	0
	4	0	0	0	0
	5	0	0	0	0
Mean	3	25836	675	7.69	4.03
	4	44475	1138	8.33	2.92
	5	107337	847	8.46	4.30
Std. error mean	3	2130	137	0.196	0.584
	4	2568	207	0.155	0.532
	5	17704	268	0.164	1.38
Median	3	22000	424	7.90	3.10
	4	40900	708	8.60	1.70
	5	116067	666	8.50	2.50
Standard deviation	3	11067	714	1.02	3.04
	4	14974	1208	0.903	3.10
	5	50074	758	0.463	3.89
Minimum	3	10000	14	4.90	0.600
	4	15000	20	5.80	0.400
	5	49712	181	7.70	1.00
Maximum	3	50000	3204	9.30	10.4
	4	90000	5571	9.70	14.2
	5	196000	2569	9.20	12.4

Note – compiled by the authors using the jamovi software

According to Table 1, on average, the 5-star hotels have the highest prices (KZT107,337), followed by 4-star hotels (KZT44,475), and then 3-star hotels (KZT25,836). The 4-star hotels have the highest average number of reviews (1138), followed by 3-star hotels (675), and then 5-star hotels (847) (Table 1). The range of review numbers is substantial, with 4-star hotels having the highest variability (from 20 to 5571 reviews). On average, the 5-star hotels have the highest customer ratings (8.46), followed by 4-star hotels (8.33), and then 3-star hotels (7.69). On average, the 4-star hotels are the closest to the city center (2.92), followed by

5-star hotels (4.30), and then 3-star hotels (7.69). In summary, these summary statistics provide insights into the central tendencies, variability, and distributions of the “price”, “reviews_number”, “rating”, and “distance_from_centre” variables across different star rating categories for hotels in Almaty.

In order to address the research question “Is there a significant difference in “customers' overall ratings” between hotels with different “star ratings?””, One-Way ANOVA analysis was conducted in Jamovi to compare the means of three groups (Star ratings 3, 4, and 5) based on a dependent variable (rating). One-Way ANOVA assesses whether there are statistically significant differences between the means of the three groups (Kenton, 2023). A larger F-value indicates greater differences between group means (Kim, 2014). The p-value is the probability of observing the data if there were no real differences between the groups (Dahiru, 2008). In our case, $p = 0.015$ (Welch's) and $p = 0.016$ (Fisher's). Both p-values are less than 0.05, indicating that there is a statistically significant difference between at least one pair of group means. Shapiro-Wilk Normality Test checks whether the data in each group follows a normal distribution. The p-value for the “rating” variable is less than 0.001, indicating a violation of the assumption of normality. In other words, the “rating” variable does not follow a normal distribution, as indicated by the Shapiro-Wilk test ($p < .001$). However, the assumption of homogeneity of variances is not significantly violated, as suggested by the Levene's test ($p = 0.158$). Researchers should consider the non-normal distribution of the “rating” variable when performing statistical analyses and choose appropriate tests or transformations as needed. Homogeneity of Variances Test of Levene checks whether the variances in the different groups are roughly equal (Statistics solutions). The p-value for “rating” is 0.158, which suggests that the assumption of equal variances is not significantly violated. This is a positive result because it means that the groups have roughly equal variances (Charlesworth Author Services, 2022).

When there is a significant result in ANOVA, post hoc tests is typically conducted to determine which specific groups differ from each other. In this case, we have conducted a Tukey post-hoc test for the “rating” variable between the three groups (Star ratings 3, 4, and 5). For the comparison between Star ratings 3 and 4, the mean difference is -0.641, and the p-value is 0.023. This difference is statistically significant ($p < 0.05$). For the comparison between Star ratings 3 and 5, the mean difference is -0.774, and the p-value is 0.098. This difference is not statistically significant ($p > 0.05$). For the comparison between Star ratings 4 and 5, the p-value is 0.927, indicating that there is no statistically significant difference between these two groups.

In the Table 2, conducting comprehensive correlation analysis, we delve into the relationships between key variables relevant to the hotel industry in Almaty, Kazakhstan. The analysis tried to investigate various aspects, ranging from the proximity of hotels to the city center, the number of customer reviews, the overall customer ratings, hotel prices, and their star ratings. Through Pearson's r and Spearman's rho correlation analyses, we attempted to explore the strength and significance of the associations among these factors.

Table 2. Correlation Matrix for Hotel Data in Almaty

Correlation Matrix						
		distance_from_centre	reviews_number	rating	price	star
distance_from_centre	Pearson's r	—				
	df	—				
	p-value	—				
	Spearman's rho	—				
	df	—				
	p-value	—				
reviews_number	N	—				
	Pearson's r	-0.348**	—			
	df	67	—			
	p-value	0.003	—			
	Spearman's rho	-0.381**	—			
	df	67	—			
rating	p-value	0.001	—			
	N	69	—			
	Pearson's r	-0.250*	0.259*	—		
	df	67	67	—		
	p-value	0.038	0.031	—		

Correlation Matrix						
		distance_from_centre	reviews_number	rating	price	star
	Spearman's rho	-0.345**	0.232	—		
	df	67	67	—		
	p-value	0.004	0.055	—		
	N	69	69	—		
price	Pearson's r	-0.074	-0.006	0.327**	—	
	df	67	67	67	—	
	p-value	0.545	0.961	0.006	—	
	Spearman's rho	-0.210	0.216	0.471***	—	
	df	67	67	67	—	
	p-value	0.083	0.074	< .001	—	
	N	69	69	69	—	
Star	Pearson's r	-0.053	0.135	0.321**	0.701***	—
	df	67	67	67	67	—
	p-value	0.665	0.270	0.007	< .001	—
	Spearman's rho	-0.127	0.174	0.348**	0.724***	—
	df	67	67	67	67	—
	p-value	0.297	0.152	0.003	< .001	—
	N	69	69	69	69	—

Note. * p < .05, ** p < .01, *** p < .001.
Note – compiled by the authors using the jamovi software

Research Question 2: What is the relationship between the “star” and the “price” for hotels in Almaty? To answer the research question based on the correlation matrix provided, we can analyze the Pearson's correlation coefficients (Pearson's r) and Spearman's rank correlation coefficients (Spearman's rho) between “star” and “price” (Table 2). Pearson's r for “star” and “price” is 0.701, which is statistically significant at the $p < 0.001$ level, denoted as ***. This indicates a strong positive correlation between the “star” rating of hotels and their prices. In other words, as the star rating increases, the price tends to increase as well. Spearman's rho for “star” and “price” is 0.724, which is also statistically significant at the $p < 0.001$ level (***). This reinforces the strong positive correlation between the “star” rating and the price of hotels in Almaty. In summary, the results suggest that there is a strong and positive correlation between the “star” rating and the price of hotels in Almaty.

Research Question 3: Do hotels with higher customer overall ratings in Almaty have a closer distance from the city center compared to lower-rated hotels? The Pearson's r correlation between “rating” and “distance_from_centre” is -0.250 ($p = 0.038$), suggesting a weak negative correlation (Table 2). The Spearman's rho correlation between these variables is -0.345 ($p = 0.004$), also indicating a weak negative correlation. Based on these correlations, it appears that there is a statistically significant, albeit weak, negative relationship between customer overall ratings and the distance from the city center. In the Almaty hotels' case the interaction between location (displayed as “distance from centre” in our case) and customer review ratings is not to be found.

Research Question 4: Do hotels with closer “distance from the city center” have a higher “price” in hotels of Almaty? The Pearson's r correlation between “distance_from_centre” and “price” is 0.327 ($p = 0.006$), indicating a weak positive correlation (Table 2). The Spearman's rho correlation between these variables is 0.348 ($p < 0.001$), also indicating a weak positive correlation. Based on these correlations, there is a weak positive relationship between the distance from the city center and hotel prices.

Research Question 5: What is the relationship between the “review number” and the “customers' overall average rating” for hotels in Almaty? The Pearson's r correlation between “reviews_number” and “rating” is 0.259 ($p = 0.031$), indicating a weak positive correlation (Table 2). The Spearman's rho correlation between these variables is 0.232 ($p = 0.055$), also suggesting a weak positive correlation. These correlations show that there is a statistically significant, albeit weak, positive relationship between the number of reviews and customer ratings. Hotels with more reviews tend to have slightly higher customer ratings.

Discussions

As it is seen in Table 1, on average, the 5-star hotels have the highest prices. It matches with Israeli's (2002) hedonic analysis of hotel prices, where it was found that star ratings consistently had a positive impact on prices, in contrast to the inconsistent impact of corporate affiliation. The research of Abrate et

al. (2011) affirms the results commonly observed in the predominant hedonic price literature. The models used in this study demonstrated that both the star rating of a property and its appeal based on location contribute to explaining variations in prices. They reveal a robust connection with the pricing premium principle, where a higher price is attributed to quality signals based on reputation, irrespective of the season. On average, the 5-star hotels have the highest customer ratings, followed by 4-star and 3-star hotels. This result is similar to Hu and Chen (2016) findings, because, they found a correlation effect between review rating and hotel star class.

A Tukey post-hoc test was conducted for the “rating” variable between the three groups (Star ratings 3, 4, and 5). For the comparison between Star ratings 3 and 4, the mean difference is -0.641, and the p-value is 0.023. This difference is statistically significant ($p < 0.05$). This also conforms with Hu and Chen (2016) findings, where they confirmed the interaction effect between hotel star class and review rating. In summary, the results suggest that there is a strong and positive correlation between the “star” rating and the price of hotels in Almaty. Higher-rated hotels tend to have higher prices. As it was clearly reported by Schamel (2012) in literature review, important factors that influence hotel room pricing are the hotel's star rating, popularity rating based on customer reviews, advance reservation time, and specific hotel amenities such as express check-out, room service, or Internet access. As HSC represents a hotel's quality with a star rating (Hu and Chen, 2016), hotels with higher star ranking are usually higher priced (Agušaj et al., 2017). The Pearson's r correlation between “rating” and “distance_from_centre” is -0.250 ($p = 0.038$), suggesting a weak negative correlation (Table 2). The Spearman's ρ correlation between these variables is -0.345 ($p = 0.004$), also indicating a weak negative correlation. Based on these correlations, it appears that there is a statistically significant, albeit weak, negative relationship between customer overall ratings and the distance from the city center. In the Almaty hotels' case the interaction between location (displayed as “distance from centre” in our case) and customer review ratings is not to be found.

The Pearson's r correlation between “distance_from_centre” and “price” is 0.327 ($p = 0.006$), indicating a weak positive correlation (Table 2). The Spearman's ρ correlation between these variables is 0.348 ($p < 0.001$), also indicating a weak positive correlation. Based on these correlations, there is a weak positive relationship between the distance from the city center and hotel prices. Espinet, Saez, Coenders, and Fluvia (2003) noted that factors like location, hotel size, proximity to the beach, and the availability of parking had a significant impact on prices. It was highlighted that attributes like the type of tour operator, location, breakfast, restaurant, TV, room type, and distance from the beach played a significant role in determining tour prices (Thrane, 2005). In fact, Almaty city is surrounded by Alatau mountains and in most cases, hotels located at mountains have good views and they are located far from city centre. This ambiguous result was also demonstrated in Sinclair, Clewer, and Pack's (1990) *нет в лит-ре* paper, where they say that prices for British package holidays tend to rise when hotels are situated away from the central area of Malaga. On the contrary, some other authors identified a negative correlation between motel prices and the distance from the center of Ballina, Australia. However, hotel types must be taken into consideration here and should be analysed separately.

Conclusions

In summary, this research study used quantitative data scraped from Booking.com to examine the relationships between variables including star ratings, prices, customer ratings, distance from the city center, and review numbers for hotels in Almaty, Kazakhstan. In order to respond to the following research questions, these variables were examined: 1) Is there a significant difference in “customers' overall ratings” between hotels with different “star ratings?”. According to the study, hotels with varying star ratings significantly differ in terms of overall customer satisfaction. Customer ratings for hotels are positively correlated with their star ratings, indicating a positive relationship between the two variables. 2) What is the relationship between the “star rating” and the “price” for hotels in Almaty? The results of the analysis showed a direct relationship between hotel costs and star ratings. The general belief that hotels with more stars provide better amenities and services is consistent with the tendency for higher-rated hotels to command higher prices. 3) Do hotels with higher customer overall ratings in Almaty have a closer distance from the city center compared to lower-rated hotels? The findings showed a weak, negative correlation between the distance from the city center and the overall ratings of customers. 4) Do hotels with closer “distance from the city center” have higher “prices” in hotels in Almaty? The research showed that there is a weak positive relationship between the distance from the city center and hotel prices. 5) What is the relationship between the “review number” and the “customer's overall average rating” for hotels in Almaty? The study discovered a weak but positive relation-

ship between a hotel's average customer rating and the quantity of reviews it gets. Customers tend to give hotels with more reviews a little higher rating.

It is important to recognize some limitations even in light of the insightful knowledge this research has provided. First off, the study only used information gathered from Booking.com, which might not be representative of all lodging options in Almaty. The hotels that were listed on Booking.com were the only accommodations included in the sample. Furthermore, web scraping technologies might not be able to guarantee perfect accuracy, and data availability might have changed since the data was first gathered. Moreover, statistical analyses and correlation rather than causation are the basis for the study's conclusions. Although correlations between variables have been found, more research may be necessary to determine the causes of these correlations. Furthermore, other variables like client demographics and travel intentions that might affect hotel rankings, costs, and locations were not taken into account in the study.

In summary, this study offers insightful information about Almaty, Kazakhstan's hotel sector. It illustrates how crucial it is to comprehend the dynamics of the hotel market by taking into account star ratings, customer reviews, pricing, and location. Nonetheless, additional investigation and an all-encompassing methodology for gathering data are imperative to completely grasp the intricacies of this sector and the diverse elements impacting hotels in Almaty.

References

- Abenova, E. A., Ustenova, O. Zh., Akhtymbaeva, B. I., & Smykova, M. R. (2021). Marketing management of the competitiveness of hotels in Kazakhstan. *Statistics, account and audit statistics, account and audit*, 122. doi: 10.51579/1563-2415.2021-1.24.
- Abrate, G., Capriello, A., & Fraquelli, G. (2011). When quality signals talk: Evidence from the Turin hotel industry. *Tourism Management*, 32(4), 912-921.
- Agušaj, B., Bazdan, V., & Lujak, Đ. (2017). The relationship between online rating, hotel star category and room pricing power. *Ekonomiska misao i praksa*, (1), 189-204.
- Alegre, J., Cladera, M., & Sard, M. (2013). Tourist areas: Examining the effects of location attributes on tour-operator package holiday prices. *Tourism Management*, 38, 131-141.
- Atkinson, H., & Brown, J. B. (2001). Rethinking performance measures: assessing progress in UK hotels. *International Journal of Contemporary Hospitality Management*, 13(3), 128-136.
- Babaevskaya, E. A. (2019). Analysis of the development of the hotel business in the republic of Kazakhstan. *In Trends in the development of tourism and hospitality in Russia*, 399-403
- Banerjee, S., & Chua, A. Y. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management*, 53, 125-131.
- Banker, R. D., Potter, G., & Srinivasan, D. (2005). Association of nonfinancial performance measures with the financial performance of a lodging chain. *Cornell Hotel and Restaurant Administration Quarterly*, 46(4), 394-412.
- Barros, C. P. (2005). Measuring efficiency in the hotel sector. *Annals of tourism research*, 32(2), 456-477.
- Bayandinova, A. M., & Djoldybaev, A. O. (2015). Problems of development of hotel business in the Republic of Kazakhstan. *In Innovative development of socio-economic systems: conditions, results and opportunities* (pp. 6-6).
- Borges-Tiago, M. T., Arruda, C., Tiago, F., & Rita, P. (2021). Differences between TripAdvisor and Booking.com in branding co-creation. *Journal of Business Research*, 123, 380-388.
- Bureau of National Statistics (2022). Statistics of tourism. Adapted from <https://stat.gov.kz/en/industries/business-statistics/stat-tourism/>.
- Chang, Y. C., Ku, C. H., & Chen, C. H. (2019). Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management*, 48, 263-279.
- Charlesworth Author Services (2022). Significance and use of Post-hoc Analysis studies. [Cwauthors]. Retrieved from <https://www.cwauthors.com/article/significance-and-use-of-post-hoc-analysis-studies#:~:text=Simply%20put%2C%20a%20post%2Dhoc,also%20called%20multiple%20comparison%20tests>.
- Chen, M. H. (2010). The economy, tourism growth and corporate performance in the Taiwanese hotel industry. *Tourism Management*, 31(5), 665-675.
- Cibere, G. U., Bařaran, M. A., & Kantarci, K. (2020). Evaluation of hotel performance attributes through consumer generated reviews: The case of Bratislava. *Advances in Hospitality and Tourism Research (AHTR)*, 8(1), 48-75.
- Clewer, A., Pack, A., & Sinclair, M. T. (1990). Forecasting models for tourism demand in city dominated and coastal areas. *In Papers of the Regional Science Association*, 69(1), 31-42. Berlin/Heidelberg: Springer-Verlag.
- Dahiru, T. (2008). P-value, a true test of statistical significance? A cautionary note. *Annals of Ibadan postgraduate medicine*, 6(1), 21-26.
- de la Cruz del Río-Rama, M., Maldonado-Erazo, C. P., Álvarez-García, J., & Ramírez-Coronel, R. L. (2019). Social Media and Hotels in Iran: Online Reputation of Hotels on Booking.com. *In Experiencing Persian Heritage: Perspectives and Challenges* (pp. 189-205). Emerald Publishing Limited.

- Díaz, M. R., & Rodríguez, T. F. E. (2018). Determining the reliability and validity of online reputation databases for lodging: Booking.com, TripAdvisor, and HolidayCheck. *Journal of Vacation Marketing*, 24(3), 261-274.
- Duysembayev, A. A., & Kayrbekova, D. N. (2020). Analysis of the development of hotel services in the regions of Kazakhstan.
- Espinet, J. M., Saez, M., Coenders, G., & Fluvià, M. (2003). Effect on prices of the attributes of holiday hotels: a hedonic prices approach. *Tourism Economics*, 9(2), 165-177.
- Figini, P., Vici, L., & Viglia, G. (2020). A comparison of hotel ratings between verified and non-verified online review platforms. *International Journal of Culture, Tourism and Hospitality Research*, 14(2), 157-171.
- Fleischer, A. (2012). A room with a view—A valuation of the Mediterranean Sea view. *Tourism Management*, 33(3), 598-602.
- Gligorićević, B. (2016). Review platforms in destinations and hospitality. In *Open tourism* (pp. 215-228). Springer, Berlin, Heidelberg.
- Go, F., Pine, R., & Yu, R. (1994). Hong Kong: Sustaining competitive advantage in Asia's hotel industry. *Cornell Hotel and Restaurant Administration Quarterly*, 35(5), 50-61.
- Harrington, D., & Akehurst, G. (1996). Service quality and business performance in the UK hotel industry. *International Journal of Hospitality Management*, 15(3), 283-298.
- Holjevac, I. A. (2003). A vision of tourism and the hotel industry in the 21st century. *International Journal of Hospitality Management*, 22(2), 129-134.
- Hu, Y. H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management*, 36(6), 929-944.
- Israeli, A. A. (2002). Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. *International Journal of Hospitality Management*, 21(4), 405-424.
- Kanbayeva, A. K. (2021). Franchising as a factor of increasing the economic stability of the hotel. *Global Science and Innovations: Central Asia* (see in books), 2(7), 133-135.
- Kantarci, K., Basaran, M. A., & Ozyurt, P. M. (2017). Determining hotel performance through consumer generated travel 2.0 reviews: A case of Kyrgyzstan. *Manas Sosyal Arastirmalar Dergisi*, 6(5), 131-145.
- Kazakh Tourism. (2020). Sights of Almaty. Accessed in 25.11.2022 from <https://tourstat.kz/almaty>
- Kenton, W. (2023). Analysis of Variance (ANOVA) Explanation, Formula, and Applications. [Investopedia]. Retrieved from <https://www.investopedia.com/terms/a/anova.asp#:~:text=The%20one%2Dway%20ANOVA%20is,variable%20affecting%20a%20dependent%20variable.>
- Kim, H. Y. (2014). Analysis of variance (ANOVA) comparing means of more than two groups. *Restorative dentistry & endodontics*, 39(1), 74-77.
- Krasnova, O. N. (2018). Formation of competitive advantages in the hotel business of kazakhstan. *Bulletin of the Faculty of Management*. St. Petersburg State University of Economics, (3-1), 58-61.
- Mariani, M. M., & Borghi, M. (2018). Effects of the Booking.com rating system: Bringing hotel class into the picture. *Tourism Management*, 66, 47-52.
- Martin-Fuentes, E., Mateu, C., & Fernandez, C. (2018). Does verifying uses influence rankings? Analyzing Booking.com and TripAdvisor. *Tourism Analysis*, 23(1), 1-15.
- Mellinas, J. P., Nicolau, J. L., & Park, S. (2019). Inconsistent behavior in online consumer reviews: The effects of hotel attribute ratings on location. *Tourism Management*, 71, 421-427.
- Moro, S., Rita, P., & Oliveira, C. (2018). Factors influencing hotels' online prices. *Journal of Hospitality Marketing & Management*, 27(4), 443-464.
- Nurpeisova, L. S., Zamanbekova, A. B., & Akylova, G. S. (2018). Development of the hotel business in the Republic of Kazakhstan. In *The Europe and the Turkic World: Science, Engineering and Technology* (pp. 223-230).
- O'Mahony, M. P., & Smyth, B. (2010). A classification-based review recommender. *Knowledge-Based Systems*, 23(4), 323-329.
- Pchelyanskaya, I. (2020). Krisha. V bolee chem 90 % mest Almaty zakrytie turistov ne imeet kategorii i uslugi. Retrieved from <https://krisha.kz/content/news/2020/v-almaty-bolee-90-mest-razmeshcheniya-turistov-ne-imeyut-kategorii-i-servisa>
- Rita, P., Ramos, R., Borges-Tiago, M. T., & Rodrigues, D. (2022). Impact of the rating system on sentiment and tone of voice: A Booking.com and TripAdvisor comparison study. *International Journal of Hospitality Management*, 104, 103245.
- Ruytenbeek, N., Verschraegen, M., & Decock, S. (2021). Exploring the impact of platforms' affordances on the expression of negativity in online hotel reviews. *Journal of Pragmatics*, 186, 289-307.
- Sainaghi, R., Phillips, P., & Corti, V. (2013). Measuring hotel performance: Using a balanced scorecard perspectives' approach. *International Journal of Hospitality Management*, 34, 150-159.
- Schamel, G. (2012). Weekend vs. midweek stays: Modelling hotel room rates in a small market. *International Journal of Hospitality Management*, 31(4), 1113-1118.
- Statistics Solutions. (2023). The Assumption of Homogeneity of Variance. [Statistics Solutions]. ugi. Retrieved from <https://www.statisticssolutions.com/about-us/>

- Thrane, C. (2005). Hedonic price models and sun-and-beach package tours: The Norwegian case. *Journal of Travel Research*, 43(3), 302-308.
- Tsai, H., Song, H., & Wong, K. K. (2009). Tourism and hotel competitiveness research. *Journal of travel & tourism marketing*, 26(5-6), 522-546.
- Tuominen, P. (2011). The influence of TripAdvisor consumer-generated travel reviews on hotel performance.
- 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.
- Yang, Y., Wong, K. K., & Wang, T. (2012). How do hotels choose their location? Evidence from hotels in Beijing. *International journal of hospitality management*, 31(3), 675-685.
- Yergesh, B., Bekmanova, G., & Sharipbay, A. (2017). Sentiment analysis on the hotel reviews in the Kazakh language. *In 2017 International Conference on Computer Science and Engineering (UBMK)*. (pp. 790-794). IEEE.
- Yoo, K. H., Sigala, M., & Gretzel, U. (2016). Exploring TripAdvisor. In *Open tourism* (pp. 239-255). Springer, Berlin, Heidelberg.
- Zhang, D., Tu, J., Zhou, L., & Yu, Z. (2020). Higher tourism specialization, better hotel industry efficiency? *International Journal of Hospitality Management*, 87, 102509.
- Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, 15(4), 267-280.
- Zigan, K., & Zeglat, D. (2010). Intangible resources in performance measurement systems of the hotel industry. *Facilities*, 28(13/14), 597-610.

Н.М. Әбдіразақов, Ж.Т. Көшебаев, Ф.А. Нуманова, П.Ж. Орынбет, Г.Е. Мылтыкбева

Қазақстанның Алматы қаласындағы қонақ үйге қатысты айнымалылар арасындағы байланыстарды түсіну: Booking.com деректерін пайдалану арқылы зерттеу

Аңдатпа:

Мақсаты: Зерттеу жұмысы Қазақстанның, оның ішінде Алматы қаласындағы қонақ үйлерді жұлдызды рейтингісі, бағалары, клиенттердің рейтингтері, қала орталығынан қашықтығы және қонақ үйлер туралы пікірлер саны арқылы қонақ үй айнымалылары арасындағы өзара байланысты зерттеуге бағытталған.

Әдісі: Бұл зерттеуге арналған деректер заманауи зерттеулердегі танымал веб-деректерден ақпарат алатын құрал Parsehub веб-бетінің парсинг құралын қолдану арқылы Booking.com сайтынан сандық әдісімен жиналды.

Деректерді жинау және жетіспейтін деректерді жоюдан кейін 96 мәліметтің 69-ы сақталды. Зерттеу сұрақтарына жауап іздеу үшін зерттеуде жұлдызды рейтинг санаттары бойынша деректерді талдауға сипаттамалы статистика пайдаланылды, айнымалылар арасындағы айырмашылықтарды анықтауға One Way ANOVA тестісі орындалды және корреляциялық талдау жүргізілді. Пайдаланылған талдау бағдарламалары ретінде Jamovi және SPSS 23 нұсқасы қолданылды.

Қорытынды: Нәтижелер жұлдызды рейтингтерге негізделген тұтынушы рейтингтеріндегі айтарлықтай айырмашылықтарды, жұлдызды рейтингтер мен бағалар арасындағы оң корреляцияны, тұтынушы рейтингтері мен қала орталығынан қашықтығы арасында теріс байланысты және кері байланыс саны мен тұтынушы рейтингтері арасындағы әлсіз оң корреляцияны көрсетеді.

Тұжырымдама: Бұл идеялар Алматыдағы қонақ үй бизнесін жақсы түсінуге ықпал етеді. Нәтижелер қонақ үй нарығының динамикасын түсіну үшін жұлдызды рейтингтердің, тұтынушылардың пікірлерінің, бағалардың және орналасудың маңыздылығын көрсетеді.

Кілт сөздер: Қазақстан, туризм, Booking.com, OTAs, қонақ үйлер, дестинация, қонақжайлық.

Н.М. Абдиразақов, Ж.Т. Кушебаев, Ф.А. Нуманова, П.Ж. Орынбет, Г.Е. Мылтыкбаева

Понимание взаимосвязей между переменными, связанными с гостиницами в Алматы (Казахстан): исследование с использованием данных Booking.com

Аннотация:

Цель: Настоящая исследовательская работа направлена на изучение взаимосвязей между переменными отеля, такими как звездный рейтинг, цены, рейтинги клиентов, расстояние от центра города и количество отзывов об отелях в Алматы (Казахстан).

Методы: Данные для этого исследования были собраны количественным методом с *Booking.com* с применением инструмента парсинга веб-страниц *Parsehub*, популярного в современных исследованиях инструмента для извлечения веб-данных. После сбора данных и удаления недостающих данных 69 случаев из 96 были сохранены. Для решения вопросов исследования в статье применена описательная статистика для анализа данных по категориям звездного рейтинга, выполнен тест *One Way ANOVA* для определения различий

между переменными и проведен корреляционный анализ. Используемые инструменты анализа включают *Jamovi* и *SPSS 23*.

Результаты: Результаты указывают на значительные различия в рейтингах клиентов, основанных на звездных рейтингах, положительную корреляцию между звездными рейтингами и ценами, слабую и отрицательную связь между рейтингами клиентов и расстоянием от центра города, а также слабую положительную корреляцию между количеством отзывов и рейтингами клиентов.

Выводы: Эти идеи способствуют лучшему пониманию гостиничного бизнеса в Алматы. Результаты демонстрируют важность звездных рейтингов, отзывов клиентов, цен и местоположения для понимания динамики гостиничного рынка.

Ключевые слова: Казахстан, туризм, *Booking.com*, *OTAs*, гостиницы, дестинация, гостеприимство, рейтинг, отзыв, корреляция.