



# İstanbul'daki Airbnb Misafir Deneyimlerini Ne Şekillendiriyor: Konu Modellemesi Yoluyla İlgörüler

## Exploring What Shapes Guest Experiences in Istanbul's Airbnb: Insights from Topic Modeling

Yavuz Selim Balcioğlu<sup>1</sup>

### öz

Bu çalışma, özellikle Airbnb'ye odaklanarak, İstanbul'un hızla gelişen konaklama paylaşım ekonomisinde müşteri deneyimini ve memnuniyetini etkileyen çeşitli faktörleri kapsamlı bir şekilde incelemek için endüktif bir yaklaşım benimsemektedir. Araştırma içerisinde, bu alandaki müşteri tercihleri ve beklentilerinin inceliklerini anlamaya yönelik önemli bir çaba ile birlikte, İstanbul'dan toplanan 508.746 Airbnb yorumu üzerinden geniş bir veri setinin kapsamlı bir analizi gerçekleştirilmiştir. Analiz edilen bilgilerin açıklığı ve ilgililiğini sağlamak üzere metinsel verilerin kapsamlı bir ön işlemeden başlanmıştır. Bunun ardından, çalışma, kullanıcı tarafından üretilen içerikten 32 ayrı konuyu tespit etmek ve çıkarmak için sofistike bir istatistiksel model olan Latent Dirichlet Allocation (LDA) yöntemini kullanmaktadır. Bu konular, incelemelere gömülü olarak, konuk deneyimi hakkında zengin bilgiler sağlamaktadır. Bu sayede ortaya çıkarılan konular, analiz için yapılandırılmış bir çerçeve ile birkaç ana boyuta sistematik olarak kategorize edilmiştir. Bu boyutlar, konuklar tarafından yapılan detaylı değerlendirmeleri, konaklama yerlerinin merkezi kentsel alanlardan daha çevresel yerlere kadar uzanan konum özelliklerini ve Airbnb listelerinin hem somut hem de soyut yönlerini içermektedir. Ek olarak, çalışma, konuk memnuniyetini şekillendirmede kritik faktörler olan ev sahiplerinin yönetim uygulamalarını ve genel hizmet kalitesini incelemektedir. Bu boyutların her biri, paylaşımlı konaklama sektöründeki müşteri deneyiminin karmaşık yönlerini anlamak ve değerlendirmek için birer araç olarak işlev görmektedir. Bu konular arasındaki karmaşık ilişkileri daha derinlemesine keşfetmek için çalışma, çeşitli konular arasındaki karmaşık etkileşimi ve ince bağlantıları ortaya çıkaran istatistiksel bir teknik olan hiyerarşik Ward Kümeleme yöntemini kullanılmıştır. Bu yaklaşım, yorumlar arası konaklama bağlamında müşteri deneyiminin çok yönlü doğasını aydınlatmakta hayati bir rol oynamaktadır. Bu analiz sonuçları, İstanbul'un Airbnb sektöründeki misafir deneyimlerini şekillendiren belirleyiciler hakkında kapsamlı ve katmanlı bir anlayış sağlamayı hedeflemektedir. Müşteri memnuniyetinin etkin faktörlerine ilişkin ayrıntılı ve çok yönlü bir bakış açısı sunarak, İstanbul'un konaklama paylaşım ekonomisinin dinamik ve çeşitli manzarasında misafir deneyimlerini ve memnuniyetini etkileyen ana faktörlerin anlaşılmasını geliştiren ve bu alandaki bilgi birikimine önemli bir katkıda bulunmaktadır.

**Anahtar Kelimeler:** Airbnb Müşteri Deneyimi, Latent Dirichlet Allocation (LDA), Konaklama Ekonomisi, Hiyerarşik Ward Kümeleme, İstanbul Konaklama Sektörü

### ABSTRACT

This study adopts an inductive approach to comprehensively examine the various factors influencing customer experience and satisfaction within Istanbul's rapidly evolving accommodation-sharing economy, with a particular focus on Airbnb. The research involves an extensive analysis of a large dataset comprising 508,746 Airbnb reviews collected from Istanbul, marking a significant endeavor in understanding the nuances of customer preferences and expectations in this domain. The process begins with a thorough preprocessing of the textual data to ensure clarity and relevance in the information analyzed. Following this, the study employs Latent Dirichlet Allocation (LDA), a sophisticated

<sup>1</sup> Corresponding Author | Yetkili Yazar: Gebze Technical University, Business Administration Faculty, Management Information System, [ysbalcioglu@gtu.edu.tr](mailto:ysbalcioglu@gtu.edu.tr), ORCID: 0000-0001-7138-2972



statistical model, to identify and extract 32 distinct topics from the user-generated content. These topics, embedded within the reviews, provide a rich source of insights into the guest experience. The extracted topics are systematically categorized into several key dimensions, offering a structured framework for analysis. These dimensions include detailed assessments made by guests, locational attributes of the accommodations ranging from central urban areas to more peripheral locations, and both the tangible and intangible aspects of the Airbnb listings. Additionally, the study examines the management practices of the hosts and the overall quality of service, factors that are crucial in shaping guest satisfaction. Each of these dimensions serves as a lens through which the intricate aspects of customer experience in the shared accommodation sector can be understood and evaluated. To explore deeper into the intricate relationships among these topics, the study employs hierarchical Ward Clustering. This statistical technique is instrumental in revealing the complex interplay and subtle connections between the various topics, playing a vital role in elucidating the multifaceted nature of customer experience in the peer-to-peer accommodation context. The analysis aims to provide a comprehensive and layered understanding of the determinants that shape guest experiences in Istanbul's Airbnb sector. By offering a detailed, multi-faceted perspective on the drivers of customer satisfaction, this study significantly contributes to the body of knowledge in the field, enhancing the understanding of key factors that influence guest experiences and satisfaction in the dynamic and diverse landscape of Istanbul's accommodation-sharing economy.

**Keywords:** Airbnb Customer Experience, Latent Dirichlet Allocation (LDA), Accommodation-Sharing Economy, Hierarchical Ward Clustering, Istanbul Hospitality Sector

## INTRODUCTION:

The elusiveness of determining antecedents for guest satisfaction and dissatisfaction in the hospitality industry, particularly within its sharing economy segment, presents a significant challenge for industry stakeholders (Brunetti et al., 2020). This complexity is further amplified in the peer-to-peer (P2P) accommodation sector, such as Airbnb, where traditional models of guest satisfaction may not be applicable due to its relatively recent emergence and the decentralized nature of its offerings (Chen & Bellavitis, 2020). Although user-generated content (UGC) provides a rich source of information, there is a noticeable paucity of research specifically addressing its significance in this context. Much of the existing literature continues to apply conventional models of customer experience to the P2P paradigm, despite these models being potentially inappropriate for such a decentralized environment (Massoud et al., 2009). Recent developments highlight the evolving landscape of human resources management (HRM) in the post-pandemic tourism contexts, with a focus on algorithmic management as seen in Airbnb. This approach to performance monitoring and rewards raises important questions about the quantification of human performance and the broader implications of algorithm-driven management in the sharing economy (Turnsek & Ladkin, 2024).

As highlighted by Canina et al. (Canina et al., 2005), there is a disproportionate academic focus on higher-end hotel accommodations, underscoring the need for more research in diverse accommodation segments, especially in the burgeoning field of P2P accommodation. The characteristics influencing guest experiences in P2P settings should be recognized as distinct from those in traditional accommodation sectors. The sharing economy in accommodation demonstrates several notable differences from traditional models, such as hotels and motels. For instance, guests in shared economies encounter a wider variety of accommodation types, ranging from couches in apartments to mansions (Cromley, 2004). Interactions between customers and hosts in shared accommodations are markedly different from those between guests and employees in traditional settings (Guest, 2002). Additionally, the shared accommodation market exhibits a greater variance in service and experience quality, coupled with a lower degree of standardization in available units (Supply et al., 2006). This lack of uniformity extends to brand recognition, which is significantly less pronounced in shared accommodations compared to traditional ones. Consequently, customers often rely more heavily on user-generated reviews to assess property quality, as these reviews often represent the most reliable source of information in the absence of standard brand indicators.

The imperative role of user-generated reviews and the extensive knowledge they offer cannot be overstated (O'Hern & Kahle, 2013), yet academic research is still grappling with the challenge of fully harnessing the extensive range of user feedback available. Firstly, the nature of online user-generated content (UGC) diverges from traditional hotel reviews and survey methods commonly used in existing

hospitality research. UGC typically offers vastly larger sample sizes and the data are predominantly unstructured (Zhang et al., 2016). This lack of predefined constructs allows for a broader range of topics to be explored, although it simultaneously presents challenges in extracting meaningful insights. Secondly, advancements in information technology, computing, and big data analytics have facilitated the analysis of UGC (Liu, 2020), offering new avenues for gaining insights. However, given the recent development of text analytics tools such as sentiment analysis and topic modeling, research leveraging UGC within the hospitality industry's sharing economy is likely still nascent.

This study, therefore, seeks to employ advanced text analytics methodologies to analyze UGC, aiming to bridge these research gaps. Among these methodologies, sentiment analysis, while commonly used in hospitality research (Wilson et al., 2012), has limitations in extracting insights from large datasets of aggregated reviews. Sentiment analysis tends to overlook the individualized nature of guest experiences, as aggregating sentiments may not accurately represent the diverse and subjective preferences of individuals. In this research, the primary focus is directed towards the thematic content of discussions rather than individual sentiment expression. To achieve this, the study employs Topic Modeling, specifically Latent Dirichlet Allocation (LDA), a method originally conceptualized by Blei, Ng, and Jordan and further discussed in Jelodar et al., 2019. LDA is particularly adept at identifying latent topics within user-generated content (UGC), owing to its stochastic admixture model. This model's flexibility allows for words to be associated with multiple topics, and documents to contain a blend of various topics, thereby enabling a comprehensive semantic analysis. Such an approach is instrumental in parsing through extensive volumes of text-based user reviews, facilitating the extraction of significant, meaningful insights. By leveraging LDA, the study efficiently distills the essence of the large datasets of UGC, thereby unearthing underlying thematic structures that offer a deeper understanding of user perspectives and experiences.

The primary objective of this research is to extract actionable insights from Airbnb customer user-generated content (UGC) in Istanbul. The aim is to provide valuable information to both academic researchers and industry practitioners by identifying the principal topics of interest, assessing their relative significance, and exploring the interrelationships among these topics. Istanbul is chosen for this study due to its extensive data availability, the diverse range of guests and hosts it attracts, and the rich variety of its neighborhoods. This research is guided by three key objectives: (1) To extract latent topics that mirror customer interests as evidenced in UGC discussions within the Istanbul Airbnb context. (2) To validate and organize these identified topics into meaningful categories, thereby facilitating a clearer understanding of guest preferences and concerns. (3) To conduct an analysis of the similarities and differences among these topics, which will aid in comprehending the complex web of factors influencing guest experiences in Istanbul's Airbnb market.

The investigation initiates with an exhaustive examination of the scholarly literature pertaining to customer experience in conventional lodging domains, including hotels and motels. This groundwork sets the stage for a focused exploration of the peer-to-peer (P2P) accommodation sector, with an emphasis on its unique attributes and relevance in the contemporary accommodation landscape. Central to this study is a critical evaluation of existing research that probes into the determinants of customer experience within the P2P accommodation sphere, with a particular lens on the city of Istanbul. Methodologically, the study is anchored in the deployment of Latent Dirichlet Allocation (LDA), a sophisticated analytical tool used to uncover latent thematic elements from textual data. This approach is instrumental in achieving the study's preliminary objective of discerning the underlying thematic currents that shape customer experiences in the P2P accommodation context. For the second objective, the extracted topics are rigorously validated through multiple methods. Primarily, this involves an in-depth analysis of the documents, which exhibit the highest representation of each topic. In addition, to ensure the relevance and specificity to Istanbul, validation is further reinforced by

geographically mapping the locations of accommodation units and correlating them with the thematic content of reviews pertinent to those locations.

Following the initial analysis, the study categorizes the identified topics into four clearly defined types for better clarity and adherence to academic norms. These types are 'evaluation', 'location', 'unit characteristics', and 'management'. The categorization is informed by conceptual frameworks prevalent in prior academic literature. In its conclusive phase, the research applies Ward hierarchical clustering, a robust statistical technique, to systematically map out the interrelationships between the extracted topics. This phase of the study is aimed at uncovering the degrees of association among these topics, thereby shedding light on the intricate web of factors that shape guest experiences within Istanbul's peer-to-peer accommodation sector. This approach allows for a more layered and detailed understanding of the dynamics at play in influencing guest perceptions and experiences in this specific market.

### 1. Literature Review:

The burgeoning growth of Indian tourism has garnered substantial attention from practitioners, researchers, and governmental bodies, underscoring its significance in the service sector. A pivotal aspect of this interest lies in understanding the service experiences at the point of encounter, which profoundly influences the consumption of tourism services. However, the challenge arises in effectively measuring these service experiences during the encounter itself. Addressing this, the study leverages the immediacy and richness of social media, particularly focusing on tourists' tweets, which are often laden with sentiments and emotional content. Analyzing 791,804 tweets, the research employs topic modeling, topic association mining, and sentiment analysis to unearth key factors impacting customer service experience in the Indian tourism context. These factors include accessibility, accommodation, assurance, cultural attraction, Jugaadu (innovative) service flexibility, cleanliness, hospitality, price, restaurant options, and security. The study extends its analysis across five distinct geographical zones of India - North, South, East, West, and Northeast - revealing that the significance and influence of these factors vary markedly across these zones (Kar et al., 2021). This variation highlights the location-specific nature of factors influencing customer service experiences in Indian tourism. The findings offer profound theoretical and practical implications, providing nuanced insights into the region-specific drivers of service experience in the dynamic Indian tourism landscape.

This research explores into the role of social media (Musanovic & Dorcic, 2023), particularly Facebook, as a pivotal platform for interaction and information exchange between hotel brands and customers. Focusing specifically on Croatian five-star hotel brands, the study aims to understand their social media activities and to extract latent topics from their Facebook posts using the Latent Dirichlet Allocation (LDA) model. The use of descriptive statistical analysis and probabilistic generative topic modeling enables the identification of prevalent themes and strategies employed by these hotel brands in their social media marketing efforts. The findings indicate that Croatian five-star hotel brands primarily leverage Facebook for objectives such as relationship building, publicity and messaging strategies, brand development, market research, and sales promotion. The paper not only discusses the practical application of LDA in analyzing social media content but also provides insightful recommendations and challenges related to the implementation of marketing strategies on social media platforms, particularly in the context of the Croatian hospitality industry. These insights contribute to a deeper understanding of the dynamic interplay between social media marketing and the hospitality sector, underlining the importance of tailored, strategic content in engaging customers and building brand value.

The increasing relevance of user-generated content (UGC) in the domain of online knowledge sharing, particularly within the travel industry, forms the cornerstone of this study. Travel websites commonly feature UGC, encompassing diverse aspects like attractions, accommodations, and dining experiences, offering invaluable insights for both industry professionals and consumers. Prior research predominantly concentrates on identifying useful attractions, hotels, or reviews for customers through UGC analysis. This paper (Hwang et al., 2014), however, posits the crucial role of reviews, particularly those significant from a hotel management perspective, in ensuring the success of hotels in the competitive travel industry. An in-depth examination of hotel reviews from Taiwan on Tripadvisor.com, conducted by two hotel managers, reveals that noteworthy reviews can be distinguished by specific attributes such as content features, sentiments, and writing quality. The study introduces three methods for representing content features and, through empirical analysis using Tripadvisor.com data, discovers that the Latent Dirichlet Allocation (LDA) method demonstrates comparable, if not superior, performance to the TF-IDF method, particularly in terms of higher recall with fewer features. The findings underscore the importance of all three features types in identifying significant hotel reviews, with content features largely influencing precision, while sentiments and review qualities affect recall. This underscores the multifaceted nature of UGC in the travel industry and highlights the necessity for hotel management to pay attention to diverse aspects of customer reviews for strategic and operational insights.

The tourism literature has extensively focused on understanding the determinants of guest satisfaction and consumer behavior. Departing from the traditional deductive methodologies prevalent in existing literature, this study adopts an inductive approach to analyze a large corpus of unstructured text data, comprising 104,161 online reviews from Korean accommodation customers. Employing Latent Dirichlet Allocation (LDA) — a generative, Bayesian, hierarchical statistical model — the study efficiently extracts and validates key topics of interest from these reviews (Sutherland et al., 2020). While the findings resonate with prior literature regarding the significance of factors like location and service quality, this research goes a step further by offering a more nuanced distinction within these dimensions. Moreover, the study contrasts the characteristics of metropolitan versus rural accommodations and different types of lodging to unveil variations in the topics deemed important. Notably, it uncovers a greater emphasis on competitive and unique aspects among varying accommodation types. These insights not only provide actionable strategies for managers in enhancing customer satisfaction but also equip researchers with a more refined framework for assessing customer satisfaction in the hospitality sector. This approach demonstrates the value of leveraging large-scale unstructured data to gain a deeper, more granular understanding of guest preferences and expectations in the evolving landscape of tourism.

This study pioneers in dissecting Airbnb user satisfaction and dissatisfaction by meticulously analyzing a corpus of 59,766 reviews from 27,980 listings across 12 cities (Ding et al., 2021), employing both Latent Dirichlet Allocation (LDA) and supervised LDA (sLDA). Diverging from previous Airbnb-focused research that predominantly used LDA, this study distinctively segregates positive and negative reviews to unearth the heterogeneity in satisfaction and dissatisfaction attributes. A notable finding is the emergence of 'guest conflicts' as a topic, steering future research towards understanding guest interactions in shared accommodation spaces. The study's topic distribution analysis illuminates the variance in importance attributed to identical service attributes across different Airbnb property types. Furthermore, the correlation analysis underscores the connection between a home-like experience, host assistance, and guests' intent to revisit. Through sLDA analysis, the study identifies key attributes that significantly predict user satisfaction and dissatisfaction, offering crucial strategic insights for enhancing guest experiences. Methodologically, this research exemplifies the transformation of social media data into actionable knowledge on customer satisfaction, presenting vital implications for

Airbnb practitioners. The study's nuanced approach to analyzing online reviews extends beyond conventional methodologies, providing a richer, more detailed understanding of user experiences in the sharing economy.

The role of peer-to-peer (P2P) platforms in the economy is increasingly significant, particularly as they facilitate direct interactions between buyers and sellers, often comprising individuals without extensive marketing expertise. This study explores into the field of P2P accommodation platforms, with a specific focus on Airbnb (Subroyen et al., 2023), to address the previously underexplored area of both host (marketer-generated content, MGC) and guest (user-generated content, UGC) data. Prior research has underscored the importance of MGC in influencing factors such as pricing, demand, and customer experience. UGC, typically in the form of customer reviews, has also been highlighted for its impact. Bridging the gap in existing literature, this study employs Latent Dirichlet Allocation (LDA) for topic modeling to analyze data from both hosts and guests on Airbnb. The resultant topics are ranked and scrutinized using data visualization tools like Tableau, considering various dimensions such as geographical location, review scores, and the volume of reviews. This analysis reveals notable discrepancies between the topics emphasized by hosts and those mentioned by guests, even among top-rated properties. The study makes a significant contribution by demonstrating the utility of topic modeling and analysis in extracting practical insights from the extensive data accumulated on an accommodation booking platform like Airbnb. It highlights the critical need for hosts to align their marketing content more closely with guest expectations and preferences, thereby enhancing the efficacy of their marketing strategies in the competitive P2P accommodation sector.

The utility of customer reviews in informing business decisions is undeniable, yet the challenge lies in distilling key information from vast quantities of text. This study (Hong et al., 2018), addresses this challenge by proposing a novel multi-topic sentiment analysis approach utilizing Latent Dirichlet Allocation (LDA) for user-generated content (UGC). Employing a methodical design, the study collected 104,039 hotel reviews from seven top tourist destinations via TripAdvisor. Using LDA, 30 relevant hotel-related topics were extracted from these reviews, with a focus on six major dimensions: value, cleanliness, rooms, service, location, and sleep quality. The R programming language was instrumental in conducting the data analysis. The study's primary contribution lies in its development of a lexicon-based sentiment analysis approach that targets sentences embedded with keywords related to these six dimensions. The efficacy of this model was ascertained by comparing its sentiment analysis results with the actual attribute ratings provided on the platform, showcasing superior accuracy and recall. This model offers a significant advantage in sentiment analysis, particularly for reviews lacking detailed attribute ratings, thereby providing a more nuanced understanding of customer sentiments across various topics. The findings from this study not only enhance methodological approaches in sentiment analysis but also offer valuable insights for businesses in the hospitality sector seeking to understand and address customer needs and preferences more effectively.

## 2. Materials And Methods

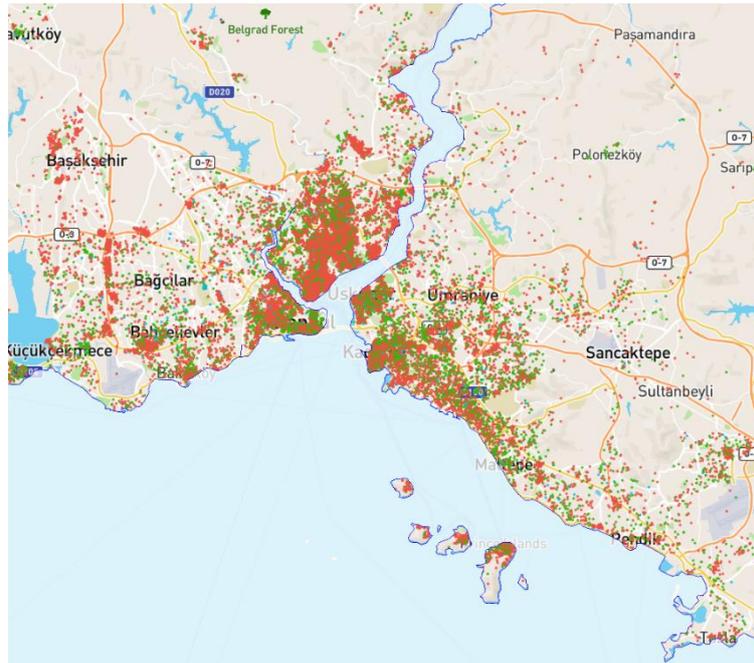
Employing a single-city sample in this research, specifically Istanbul, allows for a focused and detailed exploration of the relevant topics, albeit at the potential expense of the broader applicability of the findings. This methodological choice enhances the efficiency of the algorithmic process by streamlining the scope of location-specific keywords and effectively reducing the impact of confounding variables, such as disparities in subsample sizes across different cities or the nuanced contextual variations inherent in diverse neighborhoods and landmarks. The selection of Istanbul as the focal city is grounded in a range of persuasive reasons, aligning with the study's objectives to explore into the specificities of the peer-to-peer accommodation market within a defined urban context.

Firstly, Istanbul's cultural and historical diversity offers rich insights into a variety of attractions, guest and host backgrounds, and differing motivations for participating in the sharing economy. Secondly, the extensive sample size available within Istanbul allows for a robust analysis and clear identification of converging topics. Thirdly, the city's distinctive neighborhood characteristics, ranging from historical districts to modern urban areas, provide a wealth of diverse traits in the reviews.

Istanbul is a city that spans two continents, Europe and Asia, divided by the Bosphorus Strait. This unique geographical feature adds to the diversity of its neighborhoods. Notable areas include the historic Sultanahmet district, the cosmopolitan Beyoğlu, and the Asian side's Kadıköy and Üsküdar. Each district has its own unique charm and attracts different types of tourists, contributing to a varied peer-to-peer accommodation market. The accommodation options in Istanbul range from traditional houses in historical neighborhoods to modern apartments in bustling city centers. Additionally, Istanbul's rich history and cultural heritage, combined with its modern metropolitan dynamics, make it an ideal setting to explore a wide spectrum of guest experiences in the peer-to-peer accommodation sector.

In this study, Latent Dirichlet Allocation (LDA) was employed as the primary method for topic modeling, aimed at uncovering the latent thematic structure within the large corpus of Airbnb reviews. LDA, a generative probabilistic model, identifies topics based on the co-occurrence patterns of words within the documents. This approach assumes that documents are mixtures of topics and that topics themselves are mixtures of words. Each document in the dataset is viewed as a combination of various topics, where each word in the document is attributed to a particular topic, based on its distributional properties. The LDA process begins with the random assignment of each word in the documents to a topic, followed by iterative adjustments based on the assumption that words more likely to occur together belong to the same topic. Through this iterative process, the algorithm converges, yielding a distribution of topics across documents and a distribution of words across topics. This method allows for the extraction of meaningful patterns and themes from unstructured text data, providing insights into the prevalent topics and their relative weights within the dataset. The LDA analysis was facilitated by the topicmodels package in R, with parameter settings optimized for the specific characteristics of the dataset.

Figure (1): Airbnb listing locations in Istanbul



### 3. Data Collection and Sample

The dataset underpinning this study was procured from Inside Airbnb, a notable independent and open-source entity that compiles publicly accessible data from Airbnb across various global locales, also offering research tools. Utilizing open-source technologies, primarily Python, Inside Airbnb extracts data directly from the Airbnb platform. This study's focus was narrowed to the user-generated reviews and geographical location data of Airbnb listings situated in Istanbul. The initial dataset, amassed from Inside Airbnb, encompassed 508,746 reviews across 47,312 Istanbul listings. Rigorous data processing and analysis were undertaken by the research team. The dataset was meticulously screened to exclude reviews comprising four or fewer words and those not in English. Post-screening, the refined dataset for subsequent modeling and analytical endeavors included 347,258 reviews linked to 41,364 properties, averaging about 8 reviews per listing. This process ensured a robust and relevant sample for the study, facilitating a thorough examination of guest experiences within Istanbul's Airbnb market.

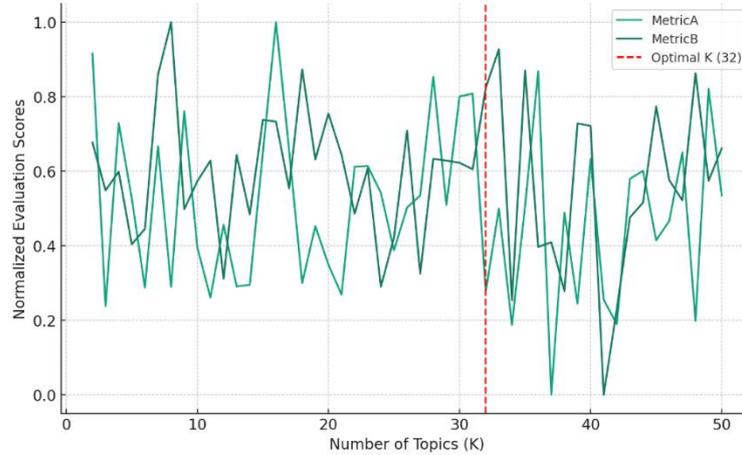
#### 3.1. Optimal Number of Topics

In determining the optimal number of topics ( $K$ ) for our study, we utilized a comprehensive approach to evaluate the performance of various  $K$  values in the Latent Dirichlet Allocation (LDA) topic models. This process involved normalizing the results of each method on a scale from 0 to 1 and employing a grid search to explore all possible numbers of topics ranging from 2 to 50. The selection criteria were based on methodologies proposed by (Naldi et al., 2011), which aim to minimize their respective criteria, and on approaches by Griffiths & Steyvers (2004), which focus on maximizing the log-likelihood of the word-topic probability in the documents.

Therefore, we defined  $K = 32$  as the total number of topics for our study. This decision was further substantiated by executing 2000 iterations of the Gibbs sampling procedure, which provided a robust probabilistic distribution for topic assignment. The comprehensive evaluation of the extracted

keywords for each topic corroborated our decision, as the keywords across these 32 topics accurately encapsulated the rich diversity of content in the Airbnb posts.

Figure (2): Appropriate K value for the optimal number of topics



The dataset is notable for its scope and granularity, providing a comprehensive basis for extracting 32 distinct topics of interest through Latent Dirichlet Allocation (LDA). This rich dataset, particularly focused on Istanbul's unique accommodation market, presents an opportunity to explore deeply into the nuances of guest experiences and preferences within the peer-to-peer accommodation sector.

In this research, the entirety of data handling, preprocessing, modeling, analysis, and visualization tasks were methodically performed using the R programming language. The preprocessing of textual data entailed the exclusion of brief reviews consisting of four words or less and the elimination of non-English reviews, identified via Google's Compact Language Detector 3, integrated into the R environment. Further refinement included the removal of a standard list of stopwords, punctuation, numerals, and a tailored list of non-conventional stopwords unique to our dataset, such as programming symbols and emoticons (e.g., 'xD', '\t', '\n'). For the purpose of streamlining the Latent Dirichlet Allocation (LDA) process, stemming was employed, utilizing the hunspell stemming dictionary accessible through the hunspell package in R. This particular stemming tool was selected for its efficacy in excluding most proper nouns, thereby reducing potential biases in the analysis.

The optimal number of topics for the LDA was ascertained using the ldatuning and topicmodels packages in R, with the LDA model relying on the Gibbs Sampling algorithm, set to complete 2000 iterations. Notably, the LDA process required approximately 48 hours of computational time on a system equipped with a 3.6 GHz Intel i9 10-Core processor and 64 GB of RAM. For geographical mapping of the data, the ggmap package was utilized, drawing on Google Map data. In terms of statistical analysis, the textmineR package was employed to compute squared distances, aligning with the classical Ward1 algorithm and the Lance-Williams update formula for Ward Clustering. Throughout the study, base R coding was utilized to ensure practicality and efficiency in data processing and analysis.

#### 4. Results

A methodological approach that is well suited for discerning the underlying thematic structures within large sets of unstructured text data. Initially, the comprehensive corpus of Airbnb reviews specific to Istanbul was preprocessed to refine the data quality, involving the removal of excessively short reviews, non-English content, and other irrelevant elements. Subsequent to this preprocessing, the LDA algorithm was applied. This involved the probabilistic modeling of words and documents, where

the algorithm iteratively assigns words to topics and topics to documents, based on their distributional patterns. Through this iterative process, the LDA was able to detect and categorize words into coherent topics, each representing a different aspect of the guest experience as reflected in the review texts. The algorithm's effectiveness in discerning nuanced themes was evident in its capacity to identify distinct and interpretable topics, ranging from specific aspects of the accommodations to broader experiential and environmental factors. This methodical extraction of topics using LDA provided a foundational basis for the subsequent analysis, facilitating a deeper understanding of guest preferences and perceptions in the context of Istanbul's Airbnb market.

Table 1 presents a structured categorization of key themes identified from Airbnb reviews in Istanbul, each defined by a set of relevant keywords. The first two topics, "Historical Proximity" and "Bosphorus Views," focus on the geographical and scenic attributes of Airbnb listings, highlighting the importance of cultural landmarks and picturesque views in guest preferences. Topics three and four, "Host Hospitality" and "Language Proficiency," shift the focus to the interpersonal aspects of the Airbnb experience, emphasizing the significance of host-guest interactions and the ability to communicate effectively in multiple languages. The next two topics, "Traditional Decor" and "Modern Amenities," represent the physical characteristics of the accommodations, contrasting traditional Turkish aesthetics with modern conveniences and facilities. Finally, topics seven and eight, "Culinary Experience" and "Nightlife Proximity," illustrate the influence of local environment on the guest experience, encompassing the rich culinary landscape of Istanbul and the vibrancy of its nightlife. Together, these topics provide a comprehensive overview of the diverse factors that shape guest experiences in Istanbul's Airbnb market.

Table (1): Latent topics regarding location

Topic	Relevant Keywords
1. Historical Proximity	historical, proximity, Sultanahmet, Hagia Sophia, Topkapi Palace, old city, cultural, landmark, walking distance, convenient, accessible, tourist spots
2. Bosphorus Views	Bosphorus, view, scenic, waterfront, panorama, seaside, balcony, picturesque, serene, waterfront dining, ferry, strait
3. Host Hospitality	host, welcoming, helpful, friendly, accommodating, informative, local tips, guidance, support, communication, responsive, gracious
4. Language Proficiency	English speaking, multilingual, communication, understanding, language barrier, fluent, interpret, guide, translate, communicate
5. Traditional Decor	Ottoman, traditional, authentic, decor, Turkish design, artistic, cultural, vintage, aesthetic, antique, style, unique
6. Modern Amenities	modern, amenities, updated, technology, Wi-Fi, air conditioning, contemporary, luxury, comfortable, state-of-the-art, equipped, facilities
7. Culinary Experience	cuisine, food, Turkish, restaurants, culinary, dining, local dishes, kebab, seafood, street food, gourmet, flavors
8. Nightlife Proximity	nightlife, entertainment, bars, clubs, vibrant, Taksim, Beyoğlu, night out, live music, social scene, party, lively

Table 2, titled 'Latent topics regarding transportation,' presents a nuanced categorization of themes pertinent to guest experiences in Istanbul's Airbnb accommodations, specifically focusing on

transportation and related aspects. Topic 9, 'Urban Connectivity,' highlights the significance of convenient public transport options and easy access to the city center, reflecting the importance guests place on mobility and infrastructure. 'Family Friendliness,' as captured in Topic 10, emphasizes the need for safe, child-friendly environments, indicating a demand for accommodations catering to family needs. Topic 11, 'Business Travel Amenities,' addresses the requirements of business travelers, underscoring the necessity of amenities like high-speed Wi-Fi and workspaces. In Topic 12, 'Shopping Proximity' comes into focus, demonstrating guests' interest in being near shopping centers, markets, and the famous Grand Bazaar. The 'Historical Character' of Istanbul is encapsulated in Topic 13, highlighting the allure of its rich architectural heritage and cultural landmarks. Topic 14, 'Arts and Culture Scene,' underscores the appeal of local arts, museums, and cultural events, reflecting guests' desire for cultural engagement. Lastly, Topic 15, 'Serene Retreats,' reveals a preference for peaceful and tranquil accommodation settings, illustrating the guests' pursuit of relaxation and privacy. Collectively, these topics provide a comprehensive understanding of the diverse preferences and considerations of guests when selecting Airbnb accommodations in Istanbul, encompassing both practical necessities and experiential desires.

Table (2): Latent topics regarding transportation

Topic	Relevant Keywords
9. Urban Connectivity	public transport, metro, bus, tram, connectivity, easy access, city center, convenience, mobility, transit, commute, infrastructure
10. Family Friendliness	family-friendly, children, safe, playgrounds, parks, spacious, kid-friendly attractions, amenities for children, safety, comfort, family accommodation
11. Business Travel Amenities	Wi-Fi, workspace, quiet, convenient location, business district, desk, Internet speed, meeting facilities, accessibility, corporate
12. Shopping Proximity	shopping, markets, boutiques, souvenirs, Grand Bazaar, malls, retail, fashion, local crafts, shopping districts, convenience
13. Historical Character	architecture, heritage, history, historic, Ottoman architecture, Byzantine, landmark buildings, cultural heritage, preservation, authentic
14. Arts and Culture Scene	galleries, museums, exhibitions, cultural events, art scene, performances, concerts, local artists, cultural immersion, artistic
15. Serene Retreats	quiet, peaceful, retreat, garden, terrace, relaxing, serene, tranquil, private, getaway, nature, calm

Table 3, titled 'Latent topics regarding eco-conscious,' delineates a diverse array of themes that are central to the eco-conscious aspect of guest experiences in Istanbul's Airbnb market. Topic 16, 'Waterfront Properties,' captures the allure of accommodations with scenic waterfront and seafront views, emphasizing the desirability of properties near the Bosphorus shore and coastal areas. 'Local Neighborhood Experience,' under Topic 17, sheds light on guests' preference for authentic local living experiences in residential neighborhoods, highlighting the appeal of immersion in daily community life. Topic 18, 'Pet-Friendly Accommodations,' reflects the growing demand for accommodations that cater to guests with pets, encompassing amenities and facilities for animal care. The focus on 'Eco-Friendly Initiatives' in Topic 19 underscores the increasing importance of sustainability and

environmentally conscious choices in guest accommodations. 'Health and Wellness Facilities,' outlined in Topic 20, represents the guest interest in accommodations offering wellness and fitness amenities, indicative of a trend towards health-oriented travel experiences. Topic 21, 'Romantic Getaways,' identifies the specific needs of couples seeking intimate and romantic settings, emphasizing accommodations suited for honeymoons or romantic retreats. 'Historical Tours and Activities,' as presented in Topic 22, highlights the interest in cultural and educational experiences through guided tours and local sightseeing. Finally, Topic 23, 'Outdoor Adventures,' caters to the adventurous spirit of guests, focusing on accommodations that offer or are near outdoor activities and natural exploration opportunities. Together, these topics provide a comprehensive insight into the multifaceted eco-conscious preferences of guests choosing Airbnb accommodations in Istanbul, spanning from environmental considerations to experiential desires.

Table (3): Latent topics regarding eco-conscious

Topic	Relevant Keywords
16. Waterfront Properties	waterfront, seafront, coastal, marina, seaside, Bosphorus shore, beach access, sea views, coastal charm, maritime, pier
17. Local Neighborhood Experience	local life, neighborhood, residential area, community, authentic experience, daily life, local markets, off the beaten path, local vibe
18. Pet-Friendly Accommodations	pets allowed, pet-friendly, dog walking, animal care, veterinary access, pet amenities, garden space, animal lovers
19. Eco-Friendly Initiatives	sustainability, eco-friendly, green, environmentally conscious, recycling, energy-efficient, organic, nature-friendly, carbon footprint
20. Health and Wellness Facilities	spa, wellness, fitness center, gym, health, relaxation, yoga, meditation, rejuvenation, healthy living, massage, retreat
21. Romantic Getaways	romantic, couples, private, intimate, honeymoon, anniversary, cozy, secluded, love, picturesque, charming
22. Historical Tours and Activities	guided tours, historical tours, cultural activities, sightseeing, heritage walks, local guides, educational, expeditions, excursion
23. Outdoor Adventures	hiking, trekking, outdoor, adventure, nature trails, parks, biking, exploration, camping, scenic, natural beauty

Table 4, titled 'Latent topics regarding accommodation,' encapsulates a comprehensive range of themes central to the Airbnb guest experience in Istanbul, each defined by a distinct set of relevant keywords. Topic 24, 'Architectural Marvels,' focuses on the appeal of accommodations with unique architectural designs and proximity to iconic buildings and landmarks, highlighting guests' appreciation for aesthetic and structural beauty. 'Local Culinary Delights,' as presented in Topic 25, emphasizes the importance of local gastronomy, catering to food enthusiasts seeking authentic Turkish culinary experiences. Topic 26, 'Festival and Event Proximity,' identifies the preference for accommodations near cultural and seasonal events, reflecting the desire for immersive local experiences. 'Educational Opportunities,' outlined in Topic 27, showcases the interest in accommodations that offer or are near opportunities for cultural and skill-based learning. Topic 28, 'Group Accommodations,' addresses the needs of larger groups or families, focusing on spacious and group-friendly properties. The 'Luxury and Opulence' of Topic 29 explores into the high-end segment of the market, highlighting the demand for

lavish and exclusive amenities. 'Budget-Friendly Options,' in Topic 30, caters to cost-conscious travelers, emphasizing the importance of economical and value-oriented accommodations. Topic 31, 'Digital Nomad Suitability,' reflects the growing trend of remote work and the demand for accommodations equipped with facilities conducive to productivity. Lastly, 'Accessible Accommodations' in Topic 32 highlights the need for inclusivity in lodging options, focusing on properties that are disability-friendly and offer ease of access. Together, these topics provide a rich and diverse understanding of the varied preferences and requirements of guests choosing Airbnb accommodations in Istanbul, covering everything from architectural interests and culinary delights to practical needs and specific lifestyle accommodations.

Table (4): Latent topics regarding accommodation

Topic	Relevant Keywords
24. Architectural Marvels	unique architecture, design, iconic buildings, aesthetic, structural beauty, innovative design, architectural tours, landmarks
25. Local Culinary Delights	local cuisine, foodie experiences, Turkish delicacies, street food, traditional dishes, culinary tours, food markets, gourmet experiences
26. Festival and Event Proximity	festivals, events, cultural celebrations, local festivities, event proximity, concert venues, seasonal events, fairs, exhibitions
27. Educational Opportunities	educational, workshops, cultural learning, language classes, cooking classes, art workshops, history lessons, learning experiences
28. Group Accommodations	large groups, family reunions, group-friendly, spacious, multiple rooms, common areas, group activities, capacity
29. Luxury and Opulence	luxury, opulent, high-end, exclusive, premium, luxury amenities, upscale, prestigious, lavish, elite
30. Budget-Friendly Options	budget, affordable, economical, value, cheap, low-cost, savings, deals, budget travel, cost-effective
31. Digital Nomad Suitability	digital nomad, remote work, long-term stay, Wi-Fi, workspaces, coworking, internet speed, productivity spaces, quiet areas
32. Accessible Accommodations	accessible, disability friendly, wheelchair access, ease of access, inclusive design, mobility friendly, adapted facilities, accessibility features

The unique geographical setting of Istanbul, straddling two continents, coupled with its storied past and vibrant local neighborhoods, significantly enriches the Airbnb experience, distinguishing it from other global destinations. Furthermore, the findings underscore the pivotal role of hosts in shaping guest experiences through their hospitality, responsiveness, and provision of local insights. These aspects, along with the accommodation's proximity to key attractions and the availability of modern amenities, emerge as critical determinants of guest satisfaction. This study's insights highlight the nuanced dynamics of the peer-to-peer accommodation market in Istanbul, suggesting that both tangible and intangible elements of the Airbnb experience contribute to its appeal among global travelers. The implications of these findings are manifold, offering valuable guidance for hosts seeking to enhance their offerings and for policymakers aiming to foster a sustainable tourism ecosystem that leverages the unique cultural and historical assets of Istanbul.

## 5. Discussion and Conclusions

The implementation of Latent Dirichlet Allocation (LDA) on a substantial dataset comprising Airbnb reviews from Istanbul has yielded detailed insights into the myriad factors influencing guest experiences within this distinctive setting. The analysis has successfully distilled 32 key topics, methodically organized across Tables 1 to 4. These topics encompass a broad spectrum of elements pivotal to guest perception, spanning areas such as their overall assessment of the stay, the precise location of the accommodation, intrinsic features of the lodging, and the hosts' management styles. This rich tapestry of findings offers both parallels and contrasts with conventional lodging experiences. Commonalities are observed in areas like perceived value, proximity to transportation and local amenities, aesthetic and hygienic aspects of the rooms, as well as host hospitality. Nevertheless, the study also uncovers stark divergences, most notably in the decentralized nature of Airbnb properties. This decentralization manifests in heightened focus on aspects such as the safety and security of the venues, the challenges posed by navigating to these unique accommodations, and a pronounced emphasis on the distinctive characteristics of each location, attributable to the heterogeneous nature of Airbnb offerings.

The study highlights a pronounced preference among guests for accommodations that not only provide modern conveniences but also offer a genuine taste of Istanbul's rich cultural tapestry. This dual expectation challenges hosts to curate spaces that reflect the local character while ensuring the global standards of comfort and convenience that travelers expect. Moreover, the significance of the local neighborhood experience in shaping guest satisfaction points to the importance of Airbnb's role in broader urban tourism dynamics. It prompts a reflection on how the platform can contribute to sustainable tourism practices that benefit both visitors and local communities without compromising the authenticity that makes destinations like Istanbul so appealing. This aspect underscores the need for a collaborative approach among hosts, local authorities, and Airbnb to promote responsible tourism practices that harness the benefits of the sharing economy while preserving the cultural integrity and sustainability of the host cities.

In further enriching our discussion, it becomes imperative to contextualize our findings within the broader spectrum of research on peer-to-peer accommodation experiences. For instance, our analysis identified "Host Hospitality" and "Local Neighborhood Experience" as pivotal to guest satisfaction, a theme resonant with the findings of Lee et al. (2023), who emphasized the significance of host-guest interactions and localized experiences in shaping positive Airbnb stays. Conversely, while our study highlights the critical role of "Digital Nomad Suitability" and "Eco-Friendly Initiatives" in Istanbul, similar studies in other urban contexts, such as those by Gunter (2018) in San Francisco, did not prioritize these aspects as heavily. This discrepancy underscores the influence of cultural and geographical specificity on guest preferences, suggesting that the factors influencing satisfaction in the Airbnb market can significantly vary across different locales. Moreover, the emphasis our participants placed on "Architectural Marvels" and "Culinary Experiences" aligns with recent trends highlighted by Orellana (2023), who noted an increasing guest interest in accommodations offering unique aesthetic and gastronomic experiences. However, our study extends beyond these findings by illustrating how such interests are intricately linked with the historical and cultural fabric of Istanbul, thereby enriching the narrative on the uniqueness of guest experiences in diverse markets. This comparative analysis not only situates our work within the existing body of literature but also highlights the nuanced understanding our study contributes, particularly regarding the dynamic interplay between guest experiences and the specific characteristics of Istanbul's Airbnb offerings.

Additionally, our study reveals a more nuanced understanding of guest-host interactions in the Airbnb context, with a variety of topics dedicated to this dynamic, reflecting the direct and often personal nature of these interactions. Unlike traditional accommodations where guest interactions with employees might be more transactional, Airbnb experiences are often marked by a deeper level of personal engagement with hosts. Another distinctive aspect uncovered is the emphasis on the accuracy and authenticity of listing descriptions, a concern less prevalent in the more standardized traditional accommodation sector.

Our study's exclusive focus on Istanbul allows for the extraction of context-specific topics that are particularly relevant to this city's unique characteristics. The large number of topics used in our analysis enabled a more detailed exploration of these themes, allowing for a richer and more precise understanding of what drives customer satisfaction in the peer-to-peer accommodation sector in Istanbul. For instance, the distinction between general recommendations and more specific advice about location, units, and management practices offers deeper insights into guest preferences.

The intricate relationships between these topics, as analyzed through Ward hierarchical clustering, reveal broader thematic clusters that encompass emotive and social experiences, navigation information, neighborhood characteristics, unit attributes, host hospitality, convenience, comfort, and communication and listing management. These findings provide valuable insights for both industry practitioners and researchers, offering a more nuanced framework for conceptualizing and measuring the dimensions of customer experience in the context of Istanbul's Airbnb market. The detailed breakdown of topics not only enhances the understanding of customer satisfaction drivers but also offers actionable insights for improving guest experiences in this dynamic and evolving sector.

### 5.1. Practical and Managerial Implications

The insights derived from the 32 latent topics identified in this study provide valuable guidance for both researchers and industry practitioners within the shared accommodation sector, particularly in a culturally and historically rich city like Istanbul. The diverse range of topics, from architectural marvels to local culinary experiences and eco-friendly initiatives, underscores the varied interests and preferences of guests in the peer-to-peer accommodation market. These findings emphasize the importance of tailoring the guest experience to individual preferences, which is a cornerstone of effective micro-segmentation in marketing strategies. Accommodation providers can leverage these insights to refine their promotional activities, ensuring that they highlight attributes most likely to resonate with their target segments.

From a theoretical standpoint, this research underscores the distinct nature of guest experiences in peer-to-peer accommodations compared to traditional lodging options. Unlike traditional hotels, Airbnb guests in Istanbul show a pronounced interest in engaging with the local community, exploring neighborhood-specific attractions, and experiencing the unique cultural and historical aspects of the city. This nuanced understanding aids in identifying key areas for customer satisfaction research and informs how hosts and managers can provide relevant and appealing information to their guests.

The utilization of statistical distances for the hierarchical clustering of topics in this study also offers methodological implications for future research. Traditional analyses in hospitality and tourism often overlook the complex relationships between various guest experience dimensions. By employing Ward clustering, this study reveals higher-level conceptualizations of the semantic environment, enhancing the understanding of the broader customer experience landscape.

For Airbnb hosts and other players in the shared accommodation market in Istanbul, these insights are instrumental in optimizing listings to appeal to specific guest preferences, thereby enhancing guest

satisfaction and bolstering online reputation. The detailed breakdown of topics offers hosts a roadmap to align their offerings with the diverse interests of their guests, whether it be highlighting proximity to historical sites, offering unique culinary experiences, or ensuring eco-friendly practices.

Moreover, traditional accommodation providers can leverage these findings to strategize more effectively. By understanding the unique dimensions of customer satisfaction in the peer-to-peer market, they can adapt their services to meet evolving guest expectations. This could involve providing more detailed information on local attractions, emphasizing unique aspects of their locations, or even rethinking their service offerings to include elements that have been successful in the peer-to-peer market, such as personalized experiences or greater engagement with local culture.

The findings not only reinforce the value of integrating modern amenities with the unique cultural and historical elements of a location but also highlight the critical role of interpersonal interactions between hosts and guests in enriching travel experiences. This research extends our understanding of the accommodation-sharing economy beyond mere transactions, framing it as a conduit for cultural exchange and deeper engagement with the local fabric of destinations. Consequently, it calls for a more nuanced approach to managing and promoting peer-to-peer accommodations, one that considers the delicate balance between offering authentic experiences and ensuring guest comfort. As the sharing economy continues to evolve, this study underscores the need for ongoing dialogue among stakeholders to navigate the challenges and opportunities presented by this dynamic sector, ensuring it contributes positively to both the tourism industry and local communities.

## 5.2. Limitations and Future Research

This study, while providing comprehensive insights into guest preferences and experiences within Istanbul's Airbnb market, is subject to certain limitations that must be acknowledged. The focus on Istanbul as the sole research context, while allowing for a detailed exploration of the city's unique Airbnb market, limits the generalizability of the findings. The cultural, historical, and geographical specifics of Istanbul mean that the results may not be directly applicable to other cities or accommodation markets with different characteristics.

Additionally, the exclusive reliance on data from Airbnb, despite its prominence in the peer-to-peer accommodation sector, introduces another limitation. The platform's specific user base, policies, and market dynamics might influence the nature and range of the reviews, potentially leading to a skewed representation of customer experiences and preferences. The unique features of Airbnb could therefore shape the findings in a manner that might not be entirely representative of the broader peer-to-peer accommodation market.

Future research should aim to expand the scope of analysis to include multiple cities and countries, encompassing a variety of cultural and geographical contexts. This would allow for a more holistic understanding of customer experiences in the global peer-to-peer accommodation market and would enable the validation and generalization of the findings from this study. Additionally, examining data from other peer-to-peer accommodation platforms could provide a more rounded view of the market, accounting for platform-specific biases and idiosyncrasies.

Further research could also explore the dynamic interplay between the various factors influencing guest satisfaction. Advanced analytical techniques could be employed to examine the complex relationships and interdependencies among the different dimensions of customer experience identified in this study. Such investigations would deepen our understanding of the customer experience landscape in peer-to-peer accommodations and could yield valuable insights for both academic research and industry practice.

**Etik Standart ile Uyumluluk**

**Çıkar Çatışması:** [TR] Yazar / yazarlar, kendileri ve / veya diğler üçüncü kiři ve kurumlarla çıkar çatışmasının olmadığını beyan eder.

[EN] There is no conflict of interest between the authors or any third party individuals or institutions.

**Etik Kurul İzni:** Bu makalede etik kurul iznine gerek yoktur, buna ilişkin ıslak imzalı etik kurul kararı gerekmediğine ilişkin onam formu sistem üzerindeki makale süreci dosyalarına eklenmiştir

**Finansal Destek:** ..... Destek varsa nereden olduđu ve kim olduđu belirtilecek.

**Teşekkür:** Teşekkürümüz yoktur.

**REFERENCE:**

- Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., Pedrini, G., & Orzes, G. (2020). Digital transformation challenges: strategies emerging from a multi-stakeholder approach. *The TQM Journal*, 32(4), 697-724.
- Chen, Y., & Bellavitis, C. (2020). Blockchain disruption and decentralized finance: The rise of decentralized business models. *Journal of Business Venturing Insights*, 13, e00151.
- Cromley, E. (2004). Domestic Space Transformed, 1850–2000. *Architectures: Modernism and After*, 163-201.
- Ding, K., Choo, W. C., Ng, K. Y., Ng, S. I., & Song, P. (2021). Exploring sources of satisfaction and dissatisfaction in Airbnb accommodation using unsupervised and supervised topic modeling. *Frontiers in psychology*, 12, 659481.
- Enz, C. A., Canina, L., & Liu, Z. (2008). Competitive dynamics and pricing behavior in US hotels: the role of co-location. *Scandinavian Journal of Hospitality and Tourism*, 8(3), 230-250.
- Guest, D. (2002). Human resource management, corporate performance and employee wellbeing: Building the worker into HRM. *The journal of industrial relations*, 44(3), 335-358.
- Gunter, U. (2018). What makes an Airbnb host a superhost? Empirical evidence from San Francisco and the Bay Area. *Tourism Management*, 66, 26-37.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl\_1), 5228-5235.
- Hong, T. H., Niu, H., Ren, G., & Park, J. Y. (2018). Multi-Topic Sentiment Analysis using LDA for Online Review. *The Journal of Information Systems*, 27(1), 89-110.
- Hwang, S. Y., Lai, C. Y., Jiang, J. J., & Chang, S. (2014). The identification of noteworthy hotel reviews for hotel management. *Pacific Asia Journal of the Association for Information Systems*, 6(4), 1.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78, 15169-15211.

- Kar, A. K., Kumar, S., & Ilavarasan, P. V. (2021). Modelling the service experience encounters using user-generated content: A text mining approach. *Global Journal of Flexible Systems Management*, 22(4), 267-288.
- Lee, C. K. H., Tse, Y. K., Zhang, M., & Wang, Y. (2023). What have hosts overlooked for improving stay experience in accommodation-sharing? Empirical evidence from Airbnb customer reviews. *International Journal of Contemporary Hospitality Management*, 35(2), 765-784.
- Liu, X. (2020). Analyzing the impact of user-generated content on B2B Firms' stock performance: Big data analysis with machine learning methods. *Industrial marketing management*, 86, 30-39.
- Massoud, M. A., Tarhini, A., & Nasr, J. A. (2009). Decentralized approaches to wastewater treatment and management: applicability in developing countries. *Journal of environmental management*, 90(1), 652-659.
- Mušanović, J., & Dorčić, J. (2023). Topic modelling of Croatian five-star hotel brands posts on Facebook using Latent Dirichlet Allocation. *ECONOMICS AND BUSINESS OF THE POST COVID-19 WORLD*, 161.
- Naldi, M. C., Campello, R. J., Hruschka, E. R., & Carvalho, A. C. P. L. F. (2011). Efficiency issues of evolutionary k-means. *Applied Soft Computing*, 11(2), 1938-1952.
- Prentice, C., & Pawlicz, A. (2024). Addressing data quality in Airbnb research. *International Journal of Contemporary Hospitality Management*, 36(3), 812-832.
- Subroyen, J., Turpin, M., de Waal, A., & Van Belle, J. P. (2023). Topic Analysis and Visualisation of Peer-to-Peer Platform Data: An Airbnb Case Study. In *Computational Intelligence: Select Proceedings of InCITe 2022* (pp. 157-166). Singapore: Springer Nature Singapore.
- Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rüscher, S., Willery, E., ... & van Soelingen, D. (2006). Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *Journal of clinical microbiology*, 44(12), 4498-4510.
- Sutherland, I., Sim, Y., Lee, S. K., Byun, J., & Kiatkawsin, K. (2020). Topic modeling of online accommodation reviews via latent dirichlet allocation. *Sustainability*, 12(5), 1821.
- Turnsek, M., & Ladkin, A. (2024). The Algorithmic Management: Reflecting on the Practices of Airbnb. In *Human Relations Management in Tourism* (pp. 57-81). IGI Global.
- O'Hern, M. S., & Kahle, L. R. (2013). The empowered customer: User-generated content and the future of marketing. *Global Economics and Management Review*, 18(1), 22-30.
- Orellana, K. (2023). Short stop, Long Memories in Helsinki: Crafting guided tours for cruise ship travelers with a short layover.
- Wilson, A., Murphy, H., & Fierro, J. C. (2012). Hospitality and travel: The nature and implications of user-generated content. *Cornell hospitality quarterly*, 53(3), 220-228.
- Zhang, X., Yu, Y., Li, H., & Lin, Z. (2016). Sentimental interplay between structured and unstructured user-generated contents: an empirical study on online hotel reviews. *Online Information Review*, 40(1), 119-145.