

Factors Contributing to Customer Satisfaction When Utilizing the Airbnb
Platform in Crete
An Airbnb Reviews Sentiment and Regression Analysis

GEORGIOS PIROUNAKIS

A THESIS

submitted in partial fulfillment of the requirements for the degree
MASTER OF SCIENCE

DEPARTMENT OF ELECTRICAL
AND COMPUTER ENGINEERING

SCHOOL OF ENGINEERING
HELLENIC MEDITERRANEAN UNIVERSITY



2024

Approved by:
Major Professor
Dr. Nikolaos Papadakis

Copyright

GEORGIOS PIROUNAKIS

2024

The content provided in this dissertation is registered under the copyrights of the Hellenic Mediterranean University – H.M.U as well as of its author and can be used only for research or educational purposes. Any reuse of the content material beyond of the academic context it is prohibited without the legal consent of the university and of its author.

Abstract

The sharing economy represents a contemporary socio-economic model wherein individuals lease out their personal possessions, such as private cars or rooms in their homes, for short durations. This research endeavors to examine the factors influencing customer satisfaction in the utilization of proprietary rental websites within the sharing economy. Extensive datasets from Airbnb, comprising online reviews and listings in Crete, were subjected to analysis employing sentiment analysis, ordinal logistic regression, and visualization techniques. The results uncover that the sentiment of Airbnb guest reviews in Crete is significantly influenced by various factors, including hosts cleanliness, host communication, property type, host's response rate and the host's superhost and identity verified status. Furthermore, positive expressions of satisfaction among Airbnb guests in Crete primarily revolved around the property's type and host interaction, followed by the amenities surrounding the property, cleanliness, and room quality. Conversely, negative online reviews were predominantly associated with also property's type/location followed by challenges related to insufficient host interaction, location, and room quality. The findings underscore the importance for Airbnb hosts to provide transparent and seamless information and room quality services, emphasizing effective communication channels with guests to enhance their experience and elevate satisfaction levels. Future studies should explore the generalizability of these findings to other urban contexts.

Content

I.	INTRODUCTION	6
II.	LITERATURE REVIEW	7
A.	Sharing Economy	7
B.	Customers' Satisfaction	8
C.	Airbnb	8
D.	Sentiment Analysis	9
E.	Regression Analysis.....	11
III.	METHODOOGY	11
A.	Data Collection and Preparation	11
B.	Data Analysis	14
IV.	FINDINGS AND DISCUSSION.....	17
A.	Sentiment Analysis and Topic Clustering.....	17
B.	Factors Influencing Comments Polarity	21
V.	CONCLUSION AND FUTURE WORK.....	27
	REFERENCES	28

List of Figures

1.	Fig. 1 Detecting language using the =DETECTLANGUAGE function in Google Spreadsheet ...	12
2.	Fig. 2 Results of DETECTLANGUAGE function in Google Spreadsheet (the reviews.csv was divided into 4 sheets for not crushing Google Spreadsheet.....)	13
3.	Fig. 3 “host_since” to months in total using Excel function “datedif”	14
4.	Fig. 4 Sentiment Analysis using Azure in Microsoft excel environment.	15
5.	Fig. 5 Merging the two datasets(listings+reviews) by matching their id number using XLOOKUP function in Excel environment.	16
6.	Fig. 6 Merging the two datasets(listings+reviews) by matching their id number using XLOOKUP function in Excel environment.	17
7.	Fig. 7 Ordinal Regression Analysis in SPSS environment. In Factors we insert Nominal or Ordinal variables while on Covariates we add the scales variables.	17
8.	Fig. 8 Polarity Likelihood of Reviews (221 negative, 4910 neutral, 44305 positive).	19
9.	Fig. 9. Word Cloud for Airbnb Negative Reviews	21
10.	Fig. 10. Word Cloud for Airbnb Positive Reviews	21
11.	Fig. 11 Polarity Distribution based on review scores in cleanliness	24
12.	Fig. 14 Superhost Proportion in Positive and Negative Comments.	26
13.	Fig. 13 Property Type: Room in boutique hotel proportion in Positive and Negative comments .	27
14.	Fig. 12 Property Type: Room and breakfast in bed proportion in Positive and Negative comments.....	27

List of Tables

I.	Table I	20
II.	Table 2	22

I. INTRODUCTION

The advancement of media technology has played a pivotal role in fostering the success of an emerging business model known as the sharing economy [1]. This progressive approach to conducting business allows individuals to share their resources through decentralized or peer-to-peer (P2P) platforms in exchange for monetary transactions [2]. The sharing economy has not only created novel business prospects, particularly for those lacking essential resources but has also provided property owners with opportunities to monetize and leverage their underutilized assets on a global scale, enabling them to serve customers from diverse regions and cultures [3]. Operating around the clock, the sharing economy delivers necessary services in varying qualities and price ranges to meet individual needs and accommodate diverse purchasing power levels. Consequently, consumption and production sustainability have experienced accelerated growth [4]. The principles of the sharing economy have found application in various domains, including shared mobility, hospitality, on-demand staffing, and media streaming. Examples of sharing economy applications include Lime, JustPark, and Uber. In the realm of hospitality, numerous companies have emerged, such as Airbnb, Booking, HomeAway, Roomorama, and HomeExchange. These organizations function as intermediaries, facilitating communication between guests and hosts [5].

Examining Airbnb as an alternative to traditional hotels, numerous prior studies have investigated this platform from diverse angles. In [6], the author explored the substitution and complementary effects of Airbnb supply on hotel sales performance. Meanwhile, [7] delved into the extent to which Airbnb acts as a substitute for hotels, and [8] scrutinized the price disparities between hotel properties and adjacent Airbnb listings. Furthermore, researchers have delved into the social aspects of guest experiences in Airbnb accommodations. For example, [9] delved into the social interactions of Airbnb guests during their stays, considering various types of contacts such as guest-to-host, guest-to-community, and guest-to-guest communications. Similarly, [10] focused on the reciprocal nature of social interactions in peer-to-peer (P2P) accommodations. Various studies have also explored factors influencing guests' booking intentions, including investigating the impact of gender congruity between guests and hosts on bookings [11], as well as assessing the influence of property and host descriptions on booking decisions and post-stay reviews [12]. Additionally, researchers have recognized the significance of studying online advertising strategies [13], price determination [14], trust-building mechanisms [15], and customers' psychological behavior [16]. Despite the diverse fields of Airbnb research, only a limited number of studies have examined guests' reviews regarding their Airbnb stays. [17] delved into the attributes shaping Airbnb users' experiences using text mining and sentiment analysis.

Similarly, [18] explored latent dimensions in textual reviews through Latent Aspect Rating Analysis (LARA). The analysis of substantial data such as customer reviews stands out as a critical success factor for businesses operating in the sharing economy [19]. It provides business owners with insights into organizational strengths and weaknesses, customer habits, facilitates the prediction of future trends, improves decision-making, detects fraud, and ultimately enhances business revenues.

Considering the prior, research that explores the factors influencing customer satisfaction by concurrently analyzing guests' online reviews alongside listing features is feasible. Thus, to gain a deeper understanding of customers' needs and achieve satisfaction, it is crucial to identify the elements that contribute to customer contentment. Utilizing various data mining techniques to analyze user comments and reviews aids in uncovering the factors that significantly impact user satisfaction. Therefore, the central research question addressed in this study is: What factors contribute to customer satisfaction when utilizing the Airbnb platform in Crete? The dataset comprising Airbnb online reviews in Crete will be analyzed concurrently with the details from the listing's dataset using techniques such as sentiment analysis, text analysis, regression, and visualization. The structure of our paper will involve a review of relevant literature, an elucidation of the applied methodology, presentation of results and discussions, and finally, the conclusion of the paper.

II. LITERATURE REVIEW

A. Sharing Economy

As indicated by reference [20], the sharing economy concept represents a significant transformation rather than a novel occurrence per se. This is attributed to the fact that the sharing economy constitutes an online community encompassing economic facets like buying, selling, and renting, all of which preexisted. However, the novelty lies in its distinctive feature of enabling users to engage with various resources, spanning both economic and social dimensions, including humans, merchandise, services, and properties [21]. A more detailed elucidation is provided by [22], defining the sharing economy as a peer-to-peer (P2P) environment that allows an individual to simultaneously act as both a client and a provider. The sharing economy fosters social involvement and collaborative consumption, thereby contributing to environmental preservation, reduction of resource wastage, and heightened community awareness [4]. It encompasses five primary domains: product-service systems, redistribution markets, collaborative lifestyles, access-based consumption, and commercial sharing systems.

As previously mentioned, sharing economy enterprises have experienced rapid growth over the past decade. The author in [5] provides justification, linking this surge to the cost-effectiveness of sharing economy businesses for consumers. This is particularly evident in the realm of accommodation, where

travelers find it more economical to opt for Airbnb stays rather than renting hotel rooms. Furthermore, [19] has illustrated that the growth of the sharing economy is significantly influenced by science and technology revolutions, particularly the evolution of electronic social networking (e-SN) platforms. These platforms have played a substantial role in advancing both societal and economic development.

While business sectors vary, online hospitality stands out as one of the most prominent fields within the sharing economy. The term "hospitality," as defined by [22], encompasses two interconnected meanings. From the viewpoint of guests, hospitality involves the provision of high-quality services. Conversely, from the perspective of hosts, it revolves around offering services and accommodations with a keen emphasis on profitability. In the context of hospitality within the sharing economy, it entails enabling individuals to leverage their properties for hosting visitors in exchange for a modest financial gain, facilitated through online sharing economy platforms like the Airbnb website.

B. Customers' Satisfaction

The term "satisfaction" is defined as "a psychological statement reflecting the fulfillment of a need or desire and the pleasure derived from such fulfillment" [19]. Customers' satisfaction can be construed as their judgments, opinions, feelings, impressions, or emotional responses to their overall experiences with a product or service [23]. It has been highlighted by [22] that comprehending customers' satisfaction and its influencing factors is crucial for ensuring customer commitment and cultivating loyalty. Enhancing customer satisfaction not only leads to increased customer retention but also fosters loyalty, directly impacting an organization's revenues and profits [24]. However, it is imperative to note that elevating customer satisfaction involves associated costs. Therefore, organizations must exercise caution to avoid reaching a level of satisfaction that incurs substantial expenses without proportional returns. Striking a balance, organizations should aim to achieve the highest level of customer satisfaction at a minimum cost. This necessitates a deep understanding of customers and the factors that contribute to their satisfaction [24].

Securing customer satisfaction poses a significant challenge for numerous organizations. The initial phase in enhancing customer satisfaction involves assessing their current satisfaction levels and identifying the factors contributing to these levels. Effectively capturing and measuring customer experience serves as an excellent strategy to accomplish this. A valuable source of insight lies in analyzing rating scores and comments provided by customers following the completion of a service, offering a comprehensive understanding of their satisfaction [25]. In this study, an examination of reviews from Airbnb guests will be conducted to gain deeper insights into customer satisfaction and the influencing factors shaping their satisfaction levels.

C. Airbnb

Established in 2008, Airbnb, Inc. operates as an online marketplace, excelling in providing accommodations, homestays, and tourism experiences. It stands out as one of the foremost and lucrative platforms within the sharing economy [26]. Functioning as an intermediary, the organization facilitates connections between property owners and those seeking hospitality services. Specifically, it serves as a link between individuals wishing to rent out their homes (hosts) and those seeking lodging (guests) in a given region, earning a modest commission from each reservation [27]. Hosts engage with Airbnb as a means of generating income from their property, despite facing the risk of potential harm from guests. Conversely, guests take advantage of more affordable accommodation options compared to alternative choices, albeit

with the risk of the property not meeting their expectations in terms of quality. In essence, Airbnb offers services that aim to create a mutually beneficial outcome for both guests and hosts [28].

Various research studies have delved into Airbnb, exploring diverse aspects such as user experience, multimedia, administration, tourism, and architecture [29]. These investigations span a wide array of topics, including review bias, hospitality exchanges, price and neighborhood prediction, neighborhood ranking, socio-economic characteristics, listing recommendation systems, rentals' distribution, users' preferences and expectations, image mining, demand mining, grading schema, matching schema, consumer segmentation, trust evaluation, innovation adoption, and adoption evaluation. Additionally, some studies have specifically focused on analyzing dimensions of customer satisfaction [30], evaluating customer feedback through text mining [3], and examining public opinions using content analysis [21]. It's worth noting that prior research has predominantly concentrated on analyzing user reviews in isolation, without integrating them with listings data. There is a recognized need for a more comprehensive analysis to understand the factors influencing customer satisfaction. This analysis should be presented in a manner accessible to non-technical users, requiring minimal effort on their part.

D. Sentiment Analysis

According to [42] sentiment analysis, also known as opinion mining, is a field of study that examines people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions toward entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. This field encompasses a broad range of tasks and is referred to by various names such as sentiment analysis, opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, and review mining. Despite these variations, they are collectively considered under the umbrella of sentiment analysis or opinion mining. While the industry commonly uses the term "sentiment analysis," in academia, both "sentiment analysis" and "opinion mining" are frequently employed, essentially representing the same area of study.

The terms "sentiment analysis" and "opinion mining" can be used interchangeably. To simplify discussions, the term "opinion" is employed to encompass opinion, sentiment, evaluation, appraisal, attitude, and emotion, although these concepts are not equivalent, and distinctions are made when necessary. The focus of sentiment analysis and opinion mining is primarily on opinions expressing or implying positive or negative sentiments.

While linguistics and natural language processing (NLP) have a long history, research on people's opinions and sentiments gained momentum after the year 2000. The field has become highly active due to its wide range of applications across various domains, the growth of the industry, and the availability of a substantial volume of opinionated data on the Web, particularly in social media. Sentiment analysis plays a central role in social media research, impacting not only NLP but also management sciences, political science, economics, and social sciences, given their dependence on people's opinions. Although sentiment analysis research officially started in the early 2000s, there were earlier works on the interpretation of metaphors, sentiment adjectives, subjectivity, viewpoints, and affects [40][41].

Essential to nearly all human activities, opinions play a pivotal role in shaping our behaviors. Whenever decisions need to be made, seeking the opinions of others becomes a natural inclination. In the real world, businesses and organizations are consistently eager to gauge consumer or public sentiments regarding their products and services. On an individual level, consumers seek the opinions of current

product users before making a purchase, and voters often want to understand public perspectives on political candidates before casting their ballots.

In the past, individuals relied on friends and family for opinions, while organizations conducted surveys, opinion polls, and focus groups to gather public and consumer sentiments. The acquisition of public and consumer opinions has long been a significant business venture for marketing, public relations, and political campaign companies. However, with the rapid expansion of social media (including reviews, forum discussions, blogs, micro-blogs, Twitter, comments, and posts on social network sites) on the web, both individuals and organizations are increasingly turning to the content available in these media for decision-making.

Today, seeking opinions on consumer products no longer necessitates relying solely on friends and family, as there is a plethora of user reviews and discussions in public forums online. Similarly, organizations may find that traditional methods such as surveys and focus groups are no longer imperative, given the wealth of publicly available information on public opinions. Nevertheless, the challenge lies in locating and monitoring opinion sites on the web and distilling the valuable information they contain.

Opinions play a central role in virtually all human activities, acting as significant influencers of our behaviors. When faced with decisions, seeking the opinions of others becomes a natural inclination. In the real world, businesses and organizations consistently seek consumer or public opinions regarding their products and services. Similarly, individual consumers seek the opinions of existing product users before making purchases and desire insights into others' opinions about political candidates before casting votes in elections.

Traditionally, individuals turned to friends and family for opinions, while organizations conducted surveys, opinion polls, and focus groups to gather public or consumer sentiments. The acquisition of public and consumer opinions has been a substantial business endeavor for marketing, public relations, and political campaign companies. With the explosive growth of social media on the web, encompassing reviews, forum discussions, blogs, micro-blogs, Twitter, comments, and postings on social network sites, both individuals and organizations increasingly rely on the content within these media for decision-making.

In contemporary times, seeking opinions on consumer products is no longer confined to asking friends and family, given the abundance of user reviews and discussions on public forums online. Similarly, organizations may find that traditional methods like surveys and focus groups are no longer indispensable, thanks to the wealth of publicly available information on public opinions. However, the challenge lies in locating and monitoring opinion sites on the web and distilling the information they contain, a formidable task due to the proliferation of diverse sites. Each site typically houses a substantial volume of opinion text, often embedded in lengthy blogs and forum postings, making it challenging for the average human reader to identify relevant sites and extract and summarize opinions. Hence, the need for automated sentiment analysis systems.

In recent years, opinionated postings in social media have not only influenced businesses but also swayed public sentiments and emotions, profoundly impacting our social and political systems. Such postings have even mobilized masses for political changes, as witnessed in some Arab countries in 2011. Consequently, there is a growing necessity to collect and study opinions on the web. Opinionated documents are not limited to external data on the web; many organizations possess internal data, such as customer feedback from emails, call centers, or survey results. The applications of sentiment analysis have flourished across various domains, including consumer products, services, healthcare, financial services, social events, and political elections.

E. Regression Analysis

In accordance to [43] Regression analysis serves as a statistical tool for investigating relationships between variables. Typically, the objective is to determine the causal effect of one variable on another, such as the impact of a price increase on demand or the influence of changes in the money supply on the inflation rate. In order to address these inquiries, the investigator compiles data on the relevant variables and utilizes regression to estimate the quantitative effect of causal variables on the influenced variable. Additionally, the investigator commonly assesses the "statistical significance" of the estimated relationships, gauging the confidence level that the true relationship closely aligns with the estimated one.

Regression techniques have been integral to the field of economic statistics (econometrics) for an extended period. However, they have increasingly gained importance in legal contexts for lawyers and legal policymakers. Regression has been presented as evidence of liability under Title VII of the Civil Rights Act of 1964, as proof of racial bias in death penalty litigation, as a basis for assessing damages in contract actions, as evidence of violations under the Voting Rights Act, and as support for damage claims in antitrust litigation, among other applications.

The regression technique is employed to evaluate the strength of the relationship between a dependent variable and one or more independent variables. It aids in predicting the value of a dependent variable based on the independent variable(s).

While linear regression analysis requires the outcome or criterion variable to be measured as a continuous variable, there are instances when researchers seek to predict an outcome that is ordinal, such as ranging from "Strongly Disagree" to "Strongly Agree."

In such scenarios, scholars can utilize Ordinal Logistic Regression to assess the impact of one or more predictor variables on the outcomes. Ordinal Logistic Regression analysis is a method used to discern the cause-and-effect relationship between independent variable(s) and a dependent variable when the latter is ordinal in nature.

III. METHODOLOGY

A. Data Collection and Preparation

Over the years, a wealth of literature has affirmed the significance and utility of examining online review comments for researchers, business proprietors, and other consumers alike. Given that users' comments convey their impressions and sentiments developed during the use of a service or product, they serve as valuable resources for gauging and scrutinizing user satisfaction levels. Additionally, these comments provide insights into the key factors influencing the observed levels of satisfaction [31].

On the Airbnb platform, guests share reviews detailing their experiences and opinions regarding the properties they have stayed in. These reviews serve as valuable data that can be analyzed to discern the factors contributing to customer satisfaction [32]. For the present study, online review comments from Airbnb guests about their accommodations in Crete were collected, along with comprehensive listing data sourced from the Inside Airbnb website [33]. Given the extensive availability of Airbnb data as public datasets on the Inside Airbnb website, acquiring both datasets were a straightforward process. The listing dataset comprises 74 attributes describing hosts, property features, and scored reviews, while the review dataset consists of six attributes and over 290 thousand review records from 2010 till 09/2023. Before entering the analysis phase, the datasets underwent preprocessing or cleaning, a step known to enhance the reliability of results [34][38]. For the listing's dataset, this involved eliminating attributes with errors, blank values, or irrelevant information. Subsequently, new columns were created or derived from existing attributes, such as host experience in months obtained from the host's start date Fig. 3. On the other hand, the review dataset underwent checks for missing data and non-English content, with Google Online Spreadsheet (Fig.1, Fig. 2), employed to detect the language of comments and remove non-English data, constituting ~39% of the records. Unnecessary columns, including comment ID, reviewer ID, and reviewer name, were also removed. Furthermore, a filter was applied to consider only comments from the year 2022, reducing the volume of reviews to 59245 (49434 after merging with the listings dataset and removing rows with error values) for easier analysis on home devices with limited capabilities.

=DETECTLANGUAGE(F2)	
comments	
Very nice place to stay for a family og group of friends. a lot of outdoor space and good vibe in general.	en
My sister and I stayed at Elea Villa in June 2023, we arrived by taxi at night, Michail had advised of check-in	en
Lovely pool and surroundings. Great location as itâ€™s only a bus ride away to the harbour and walking dis	en
From arrival to departure the stay was exactly as we imagined it. Michail and his team made sure of that. T	en
Very quiet and calm place with a lovely swimming pool with plenty of seating to sunbathe in. The hosts	en
This in a beautiful and interesting historic house at a great location. Lena was very responsive and accomm	en
Sehr stilvolles kleines HÄuschen mit allem was man braucht. Zentral gelegen, zu Fuß gleich in der Altstadt	de
A friendly place! Overall was well. Bathroom needs some refreshment. Place was clean. With 20% lesser pr	en
We had the best of times staying in the beautiful little house located on the property of the smaller hotel. Sc	en
een prachtig ruim huisje met een prachtig zwembad. we hebben het zeer naar ons zin gehad.	nl
Stefanos is a wonderful host, and the house is a lovely place to stay in Crete.	en
Wir verbrachten als Familie (4 Personen) sehr schÄne Tage in der Unterkunft. Die Wohnung war wie auf d	de
Nous avons apprÄciÄ notre sÄjour chez Stefanos, hÄte rÄactif et chaleureux. Super piscine, m	fr

Fig. 1 Detecting language using the =DETECTLANGUAGE function in Google Spreadsheet.

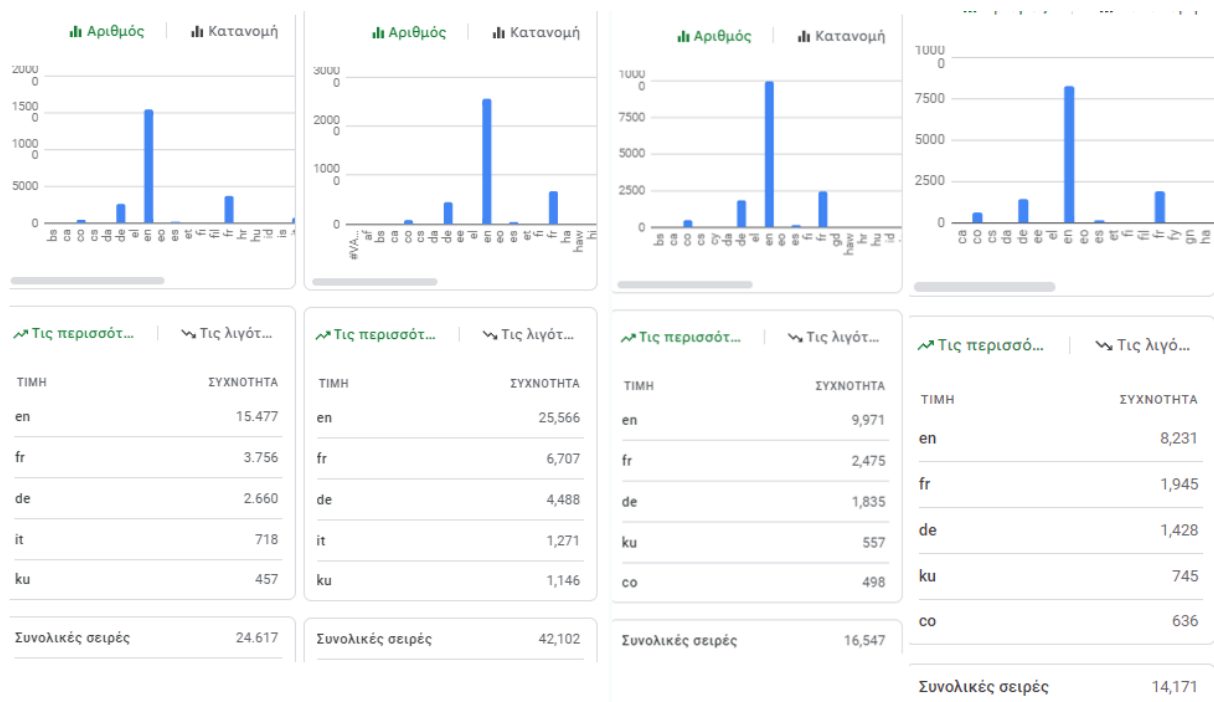


Fig. 2 Results of DETECTLANGUAGE function in Google Spreadsheet (the reviews.csv was divided into 4 sheets for not crushing Google Spreadsheet).

	A	B	C	D
1	host_since			
2	12/31/2021	=datedif(A2,today(),"m")		
3	4/18/2016			
4	6/9/2011			
5	8/31/2017			
6	1/12/2017			
7	4/25/2018			
8	8/7/2016			
9	10/2/2014			
10	9/13/2019			
11	8/1/2018			
12	5/14/2022			
13	9/29/2013			
14	5/4/2019			
15	10/10/2017			
16	9/29/2013			
17	3/10/2021			
18	6/24/2021			
19	1/10/2011			

Fig. 3 “host_since” to months in total using Excel function “datedif”

B. Data Analysis

1) Sentiment Analysis: The initial phase of the analysis involved the implementation of sentiment analysis. This sentiment analysis will be carried out utilizing the Text Analytics add-in by Microsoft® Azure [39] in the Excel environment Fig. 4. The output of Microsoft® Azure sentiment analysis comprises two columns encompassing global sentiment analysis results, detailing the polarity (Positive, Negative, or Neutral), along with a confidence score. The second sheet pertains to the categorization of comments based on topics, encompassing both topic categories and types.

B	C	D	E	F	G
positive	0.8311				
positive	0.9951				
positive	0.9967				
negative	0.0005				
positive	0.7883				
positive	0.7975				
positive	0.8189				
positive	0.9771				
positive	0.9462				
positive	0.9079				
positive	0.8549				
positive	0.6812				
positive	0.9662				
positive	0.807				
negative	0.004				
positive	0.9698				
positive	0.8906				
positive	0.956				
neutral	0.5198				
positive	0.9766				
positive	0.7185				
positive	0.7078				
positive	0.8625				
positive	0.9984				
positive	0.9521				
positive	0.7899				
neutral	0.5661				
positive	0.8881				
positive	0.806				
positive	0.9066				
positive	0.9789				
positive	0.8689				
positive	0.9969				
positive	0.6649				
positive	0.8901				
neutral	0.5815				
positive	0.7572				
positive	0.954				
negative	0.019				
positive	0.849				
positive	0.9168				
positive	0.8219				
positive	0.8261				

Fig. 4 Sentiment Analysis using Azure in Microsoft excel environment.

2)Regression analysis: The second phase of the analysis involved the identification of factors influencing user satisfaction. To initiate this process, the two datasets will be combined (Fig. 5, Fig. 6) and the research question will be addressed using the Ordinal Logistic Regression (OLR) test as suggesting in [35],[38] using the SPSS(a statistical software suite developed by IBM for data management, advanced analytics, multivariate analysis, business intelligence, and criminal investigation.)(Fig. 7) in order to identify which of the independent variables (IV) has statistically significant impact on our dependent variable (DV), which, in this context, is polarity.

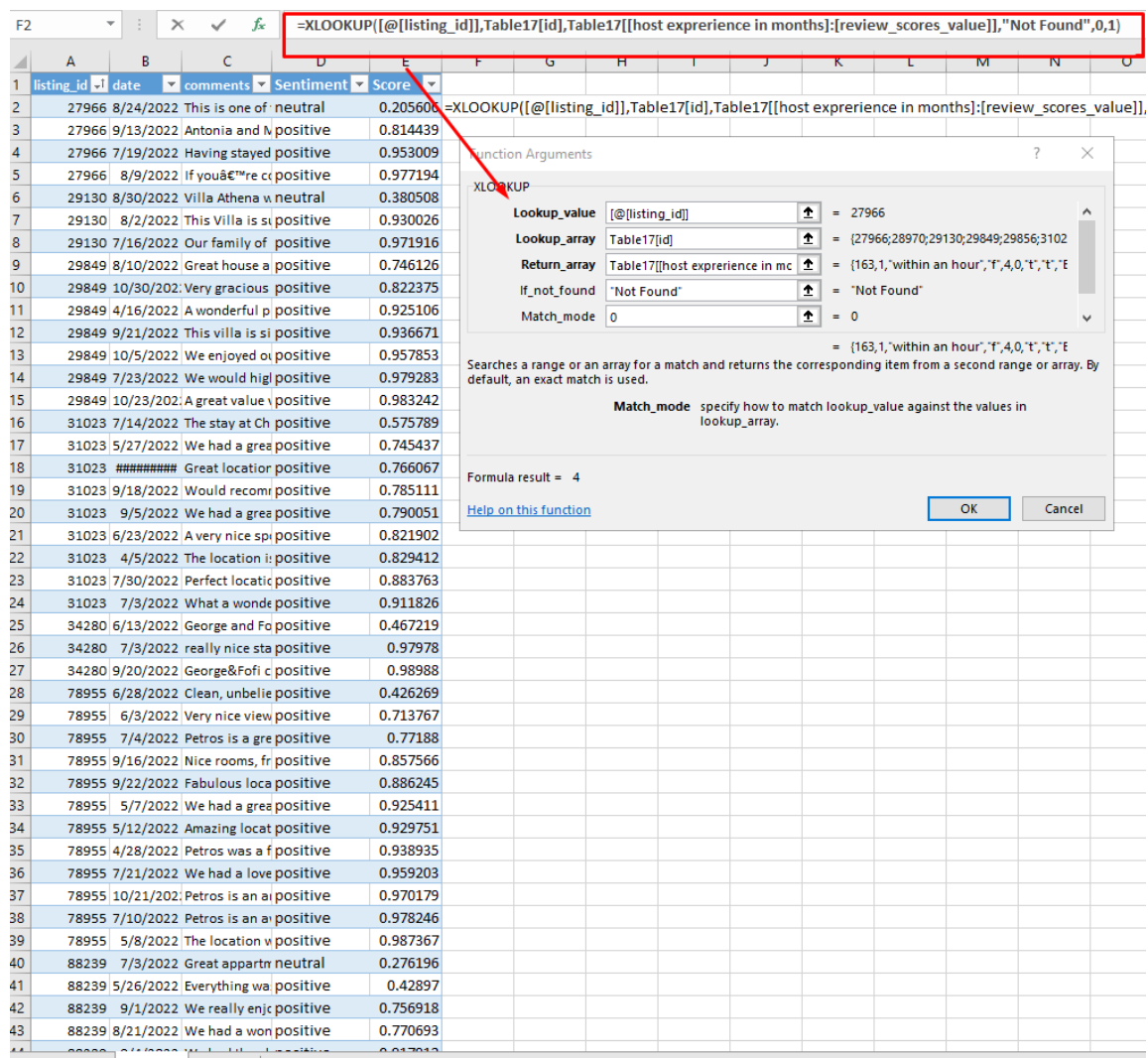


Fig. 5 Merging the two datasets(listings+reviews) by matching their id number using XLOOKUP function in Excel environment.

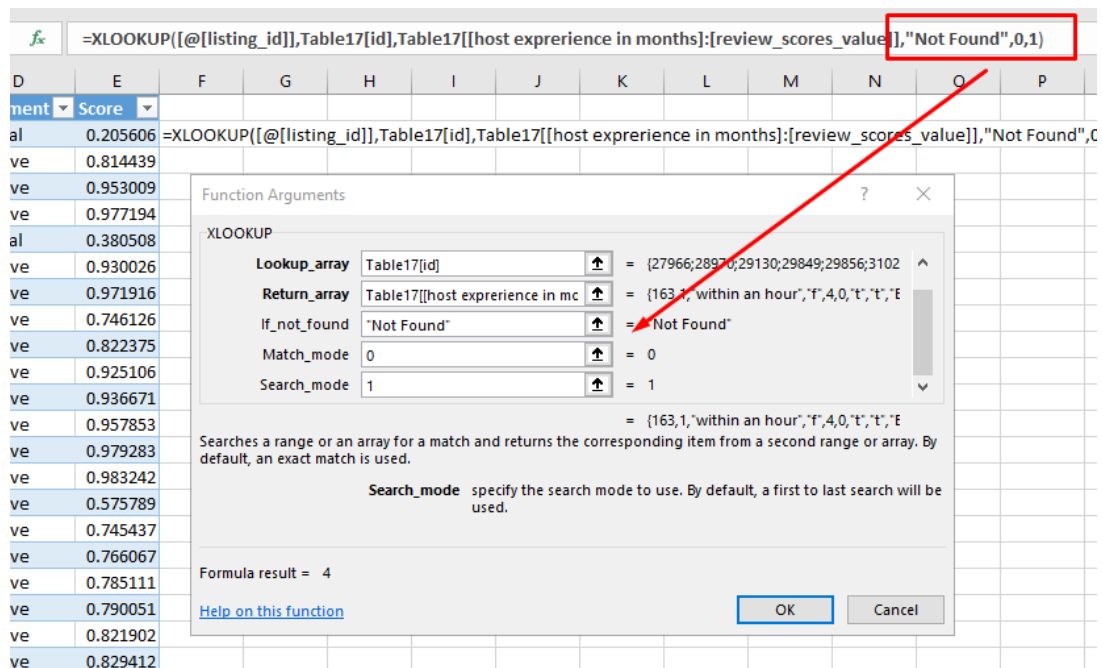


Fig. 6 Merging the two datasets(listings+reviews) by matching their id number using XLOOKUP function in Excel environment.

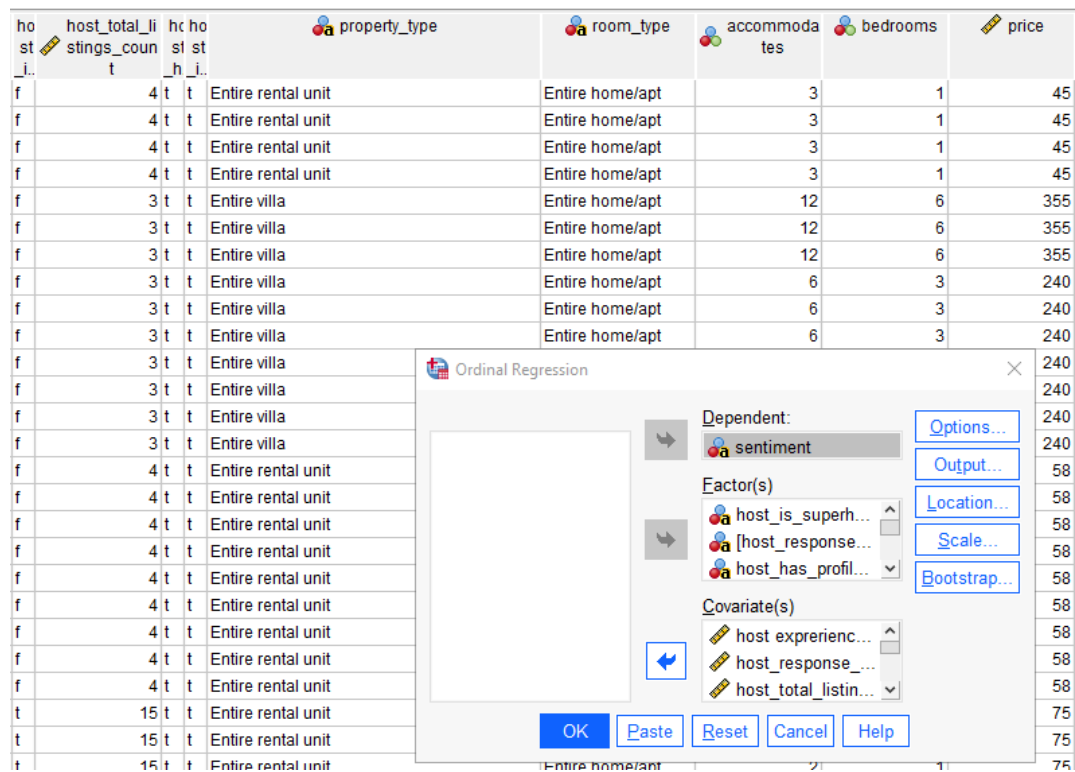


Fig. 7 Ordinal Regression Analysis in SPSS environment. In Factors we insert Nominal or Ordinal variables while on Covariates we add the scales variables.

The rationale behind choosing the OLR test for this study lies in the nature of the dependent variable (DV), which is an ordinal categorical variable represented on a three-point scale (Negative=0, Neutral=1, and Positive=2), while the independent variables (IVs) encompass a combination of continuous and categorical variables. To ensure the absence of highly correlated IVs, a multicollinearity test was conducted also using the SPSS. The results revealed that only two IVs, namely "accommodates" and "beds," exhibited a correlation exceeding 0.8. Consequently, the variable "beds" was excluded from the test. The multicollinearity test is crucial for avoiding issues in determining the variable contributing to the interpretation of the DV [35].

Following the elimination of highly correlated variables, the final set of variables comprised 17 attributes. These included host-related factors such as host experience in months, host response time, host response rate, super-host status, total listings count, profile picture, and identity verification. Additionally, property type, room specifications, accommodation details (bedrooms, number of accommodates), bookability, and various review scores encompassing cleanliness, check-in, communication, location were included in the final set.

Utilizing the outcomes of the OLR test, Microsoft Power BI software was employed to visualize the factors influencing polarity. Power BI is used to find insights within an organization's data. Power BI can help connect disparate data sets, transform and clean the data into a data model and create charts or graphs to provide visuals of the data. Additionally, Power BI was utilized to perform a text mining analysis, identifying the most frequently occurring words within comments. The visualization of words aided in extracting further insights regarding recurring terms across various polarity levels.

IV. FINDINGS AND DISCUSSION

A. Sentiment Analysis and Topic Clustering

The findings from the sentiment analysis point to a predominantly positive experience among Airbnb users in Crete, as illustrated in Figure 8. The outcome of topic clustering underscores the overwhelmingly positive sentiments users expressed, particularly regarding aspects of their experience: "apartment", "place", "host", "clean", "house", "beautiful" etc. as detailed in Table I. To provide context, these words associating with word "great" which most likely is combined with the words mentioned in most cases and which is the most frequent word in the positive reviews (17939 words count) can provide us a clear view what pleases the guests most. The importance of location (beach:8582 words count, close:5496 words count, walk:5290 words count, restaurants:5744 words count) as a crucial indicator for customer satisfaction is evident, reflecting users' concerns about the proximity of various facilities to their residence.

As a secondary major topic within positive sentiments, the term "host" represents a number of 10414 (Table I.) times seen in the positive reviews, indicating a high level of user satisfaction with the way hosts treated them. In contrast, consistently receiving a negative sentiment was the topic of "place"/ "apartment" which together were seen a number of 444 times in the negative reviews/comments (Table I). These two key factors represent the majority of the negative. This outcome suggests an opportunity for

hosts to enhance their properties either by a small renovation or improving decoration/electrical devices etc. Word clouds in Figure 9 and 10 visually support these findings.

While Airbnb comments generally lean towards positivity, a thorough examination of areas with negative sentiments empowers Airbnb hosts to rectify issues or manage future expectations, such as enhancing their properties. Furthermore, the notable prevalence of topics related to the city's environment (including location, facilities, restaurants, cleanliness) within positive sentiments, in contrast to their minimal presence in negative sentiments, suggests that Crete's overall environment significantly contributes to shaping a positive experience for Airbnb guests.

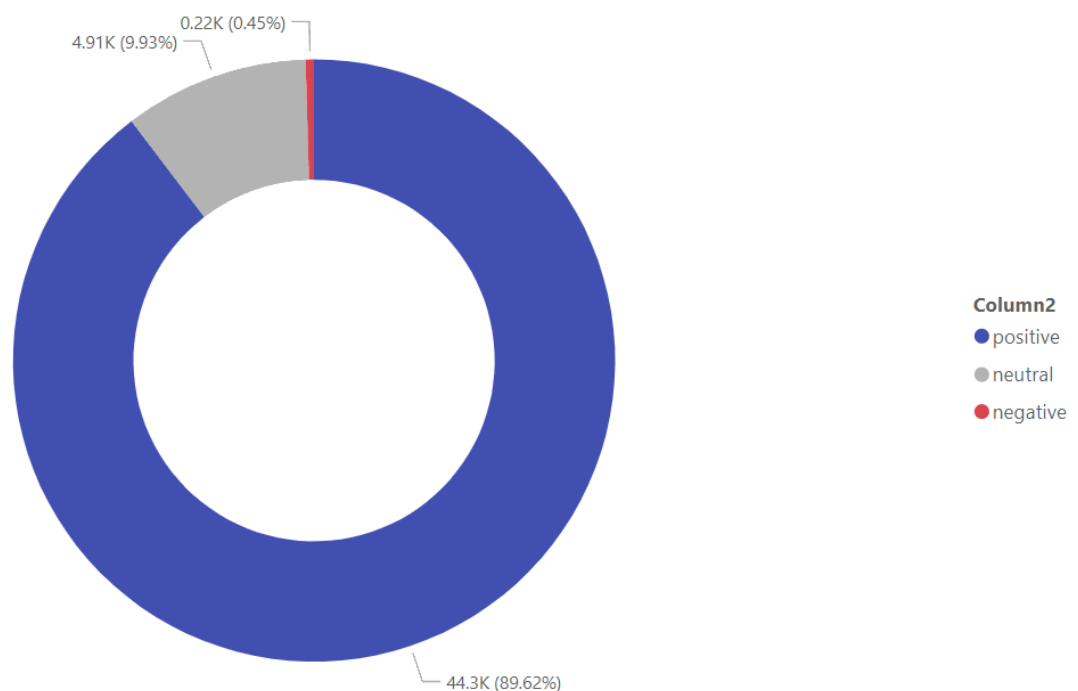


Fig. 8 Polarity Likelihood of Reviews (221 negative, 4910 neutral, 44305 positive).

TABLE I. WORDS' COUNT OF SENTIMENT ANALYSIS FOR AIRBNB USERS' ONLINE REVIEWS

Negative		Positive	
Words	Count	Words	Count
place	225	great	17939
apartment	219	place	16641
stay	196	apartment	13609
host	181	clean	11187
room	169	host	10414
house	163	house	9772
water	136	beautiful	8663
shower	128	beach	8582
location	124	good	7359
night	114	villa	5856
clean	113	restaurants	5744
time	109	time	5698
villa	106	close	5496
kitchen	98	Great	5347
car	98	pool	5233
bed	94	walk	5290
bathroom	89	quiet	4192
door	89	friendly	4159



Fig. 9. Word Cloud for Airbnb Negative Reviews



Fig. 10. Word Cloud for Airbnb Positive Reviews

B. Factors Influencing Comments Polarity

Following the implementation of an OLR test, the researchers identified that only eight out of the fourteen factors significantly influenced customers' satisfaction, with a P value (Sig.) of < 0.05 (refer to Table II). These influential factors encompass, property type ,the absence of the host being a super-host, host's response rate, having his identity not verified as well as review scores for, cleanliness.

TABLE II. ORDINAL REGRESSION RESULTS: PARAMETER ESTIMATES

Parameter Estimates							
	Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval Lower Bound	Upper Bound
[sentiment = negative]	-2.980	.982	9.212	1	.002	-4.905	-1.056
[sentiment = neutral]	.279	.980	.081	1	.776	-1.642	2.199
review_scores_cleanliness	.309	.087	12.518	1	<.001	.138	.480
review_scores_checkin	.002	.129	.000	1	.990	-.252	.255
review_scores_communication	-.188	.126	2.220	1	.136	-.436	.059
review_scores_location	-.082	.076	1.155	1	.282	-.232	.068
hostexperienceinmonths	.001	.001	1.492	1	.222	.000	.002
host_response_rate	1.122	.443	6.418	1	.011	.254	1.991
price	-5.072E-5	.000	.112	1	.738	.000	.000
host_total_listings_count	-4.248E-5	.000	.022	1	.881	-.001	.001
accommodates	-.004	.014	.092	1	.761	-.031	.023
bedrooms	-.007	.029	.064	1	.800	-.063	.049
[host_response_time=a few days or more]	-.064	.592	.012	1	.915	-1.223	1.096
[host_response_time=within a day]	.159	.136	1.357	1	.244	-.108	.426
[host_response_time=within a few hours]	.079	.067	1.408	1	.235	-.052	.210
[host_response_time=within an hour]	0 ^a	.	.	0	.	.	.
[host_is_superhost=f]	-.168	.032	27.158	1	<.001	-.232	-.105
[host_is_superhost=t]	0 ^a	.	.	0	.	.	.
[host_has_profile_pic=f]	.394	.257	2.346	1	.126	-.110	.899
[host_has_profile_pic=t]	0 ^a	.	.	0	.	.	.
[host_identity_verified=f]	.141	.041	11.838	1	<.001	.061	.221
[host_identity_verified=t]	0 ^a	.	.	0	.	.	.
[property_type=Barn]	.928	1.118	.689	1	.407	-1.263	3.118
[property_type=Bus]	1.076	1.114	.932	1	.334	-1.109	3.260
[property_type=Camper/RV]	1.075	1.028	1.095	1	.295	-.939	3.089
[property_type=Campsite]	17.528	3836.079	.000	1	.996	-7501.049	7536.105
[property_type=Casa particular]	17.475	2613.535	.000	1	.995	-5104.958	5139.909
[property_type=Castle]	.803	1.364	.346	1	.556	-1.871	3.476
[property_type=Cave]	.957	.935	1.047	1	.306	-.876	2.790
[property_type=Cycladic home]	1.470	.848	3.003	1	.083	-.193	3.133
[property_type=Dome]	.973	.983	.980	1	.322	-.953	2.900
[property_type=Earthen home]	1.036	.856	1.465	1	.226	-.642	2.713
[property_type=Entire bed and breakfast]	17.347	.000	.	1	.	17.347	17.347
[property_type=Entire home/apt]	.868	.851	1.040	1	.308	-.800	2.536

[property_type=Entire bed and breakfast]	17.347	.000	.	1	.	17.347	17.347
[property_type=Entire bungalow]	.868	.851	1.040	1	.308	-.800	2.536
[property_type=Entire cabin]	1.520	1.023	2.210	1	.137	-.484	3.525
[property_type=Entire chalet]	.531	1.047	.257	1	.612	-1.522	2.583
[property_type=Entire condo]	1.048	.829	1.597	1	.206	-.577	2.673
[property_type=Entire cottage]	1.148	.834	1.894	1	.169	-.487	2.784
[property_type=Entire guest suite]	1.102	.856	1.657	1	.198	-.575	2.779
[property_type=Entire guesthouse]	1.225	.868	1.994	1	.158	-.475	2.926
[property_type=Entire home]	1.076	.828	1.688	1	.194	-.547	2.700
[property_type=Entire home/apt]	17.336	3592.972	.000	1	.996	-7024.759	7059.432
[property_type=Entire loft]	1.022	.845	1.463	1	.227	-.634	2.679
[property_type=Entire place]	.841	.918	.838	1	.360	-.959	2.640
[property_type=Entire rental unit]	1.056	.828	1.627	1	.202	-.567	2.678
[property_type=Entire serviced apartment]	1.165	.835	1.949	1	.163	-.471	2.801
[property_type=Entire townhouse]	1.234	.838	2.168	1	.141	-.409	2.878
[property_type=Entire vacation home]	2.424	1.014	5.717	1	.017	.437	4.411
[property_type=Entire villa]	1.123	.829	1.838	1	.175	-.501	2.748
[property_type=Farm stay]	1.322	.871	2.303	1	.129	-.385	3.030
[property_type=Hut]	1.117	1.348	.686	1	.407	-1.526	3.760
[property_type=Island]	17.716	4163.248	.000	1	.997	-8142.100	8177.532
[property_type=Private room]	1.306	1.012	1.665	1	.197	-.678	3.290
[property_type=Private room in bed and breakfast]	.975	.885	1.213	1	.271	-.760	2.709
[property_type=Private room in cabin]	17.626	.000	.	1	.	17.626	17.626
[property_type=Private room in camper/rv]	1.028	1.383	.553	1	.457	-1.682	3.738
[property_type=Private room in condo]	.714	.948	.566	1	.452	-1.145	2.572
[property_type=Private room in cottage]	.259	1.044	.062	1	.804	-1.787	2.306
[property_type=Private room in cycladic house]	17.360	.000	.	1	.	17.360	17.360
[property_type=Private room in earthen home]	17.558	.000	.	1	.	17.558	17.558
[property_type=Private	1.326	1.360	.950	1	.330	-1.340	3.992

[property_type=Private room in earthen home]	17.558	.000	.	1	.	17.558	17.558
[property_type=Private room in farm stay]	1.326	1.360	.950	1	.330	-1.340	3.992
[property_type=Private room in guest suite]	1.235	.957	1.667	1	.197	-.640	3.110
[property_type=Private room in guesthouse]	.001	.932	.000	1	.999	-1.825	1.827
[property_type=Private room in home]	1.459	.891	2.680	1	.102	-.288	3.205
[property_type=Private room in hostel]	1.441	1.354	1.132	1	.287	-1.213	4.096
[property_type=Private room in loft]	1.055	1.017	1.077	1	.299	-.938	3.048
[property_type=Private room in rental unit]	1.143	.880	1.687	1	.194	-.582	2.868
[property_type=Private room in resort]	-.234	1.506	.024	1	.876	-3.187	2.718
[property_type=Private room in serviced apartment]	2.693	1.125	5.732	1	.017	.488	4.898
[property_type=Private room in tent]	.878	1.198	.537	1	.464	-1.469	3.225
[property_type=Private room in tiny home]	17.505	3061.321	.000	1	.995	-5982.574	6017.583
[property_type=Private room in townhouse]	.259	.946	.075	1	.785	-1.596	2.114
[property_type=Private room in villa]	.795	.994	.640	1	.424	-1.153	2.744
[property_type=Room in aparthotel]	1.388	.842	2.716	1	.099	-.263	3.039
[property_type=Room in bed and breakfast]	2.773	1.339	4.290	1	.038	.149	5.396
[property_type=Room in boutique hotel]	1.731	.874	3.925	1	.048	.019	3.443
[property_type=Room in hostel]	17.720	1740.033	.000	1	.992	-3392.682	3428.121
[property_type=Room in hotel]	1.303	.872	2.232	1	.135	-.406	3.012
[property_type=Room in nature lodge]	18.020	3840.306	.000	1	.996	-7508.841	7544.880
[property_type=Room in rental unit]	17.757	7185.944	.000	1	.998	-14066.434	14101.948
[property_type=Room in serviced apartment]	1.398	.903	2.397	1	.122	-.372	3.167
[property_type=Shared room]	-.688	1.433	.231	1	.631	-3.497	2.121
[property_type=Shared room in guest suite]	-2.553	2.265	1.270	1	.260	-6.992	1.886
[property_type=Shared room in hostel]	1.202	.956	1.581	1	.209	-.672	3.076
[property_type=Tent]	17.229	4142.093	.000	1	.997	-8101.123	8135.582

[property_type=Tent]	17.229	4142.093	.000	1	.997	-8101.123	8135.582
[property_type=Tiny home]	1.007	.872	1.332	1	.248	-.703	2.716
[property_type=Tower]	-.042	1.422	.001	1	.976	-2.830	2.745
[property_type=Treehouse]	0 ^a	.	.	0	.	.	.
[room_type=Entire home/apt]	.060	.265	.051	1	.822	-.460	.579
[room_type=Hotel room]	-.326	.191	2.911	1	.088	-.700	.048
[room_type=Private room]	0 ^a	.	.	0	.	.	.
[room_type=Shared room]	0 ^a	.	.	0	.	.	.
[instant_bookable=f]	-.029	.041	.507	1	.476	-1.109	.051

Cleanliness: The first factor according to our Ordinal Regression Analysis that is statistically significant with most of the positive comments to go with a 5-star review for cleanliness. Cleanliness of course is critical if you want to have a flourish business at the accommodation sector and at [44] concludes that has statistically significant effects on Airbnb customers trust and behavioral intentions. Seeing the chart at Figure 11 we can come to a conclusion that Airbnb hosts in Crete have it as their first priority with positives comments skyrocketing along with 5-star score for cleanliness.

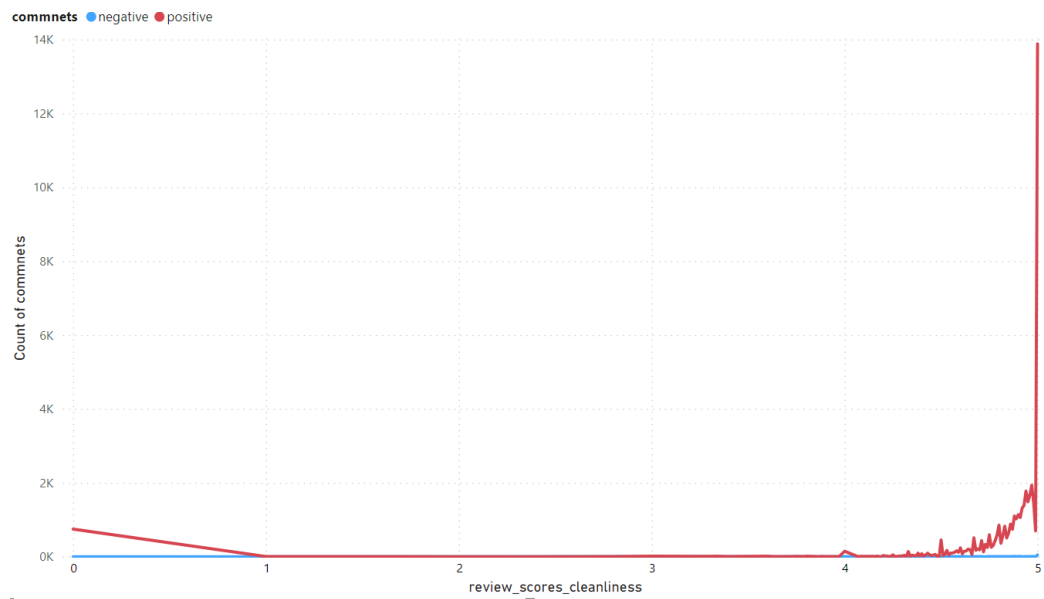
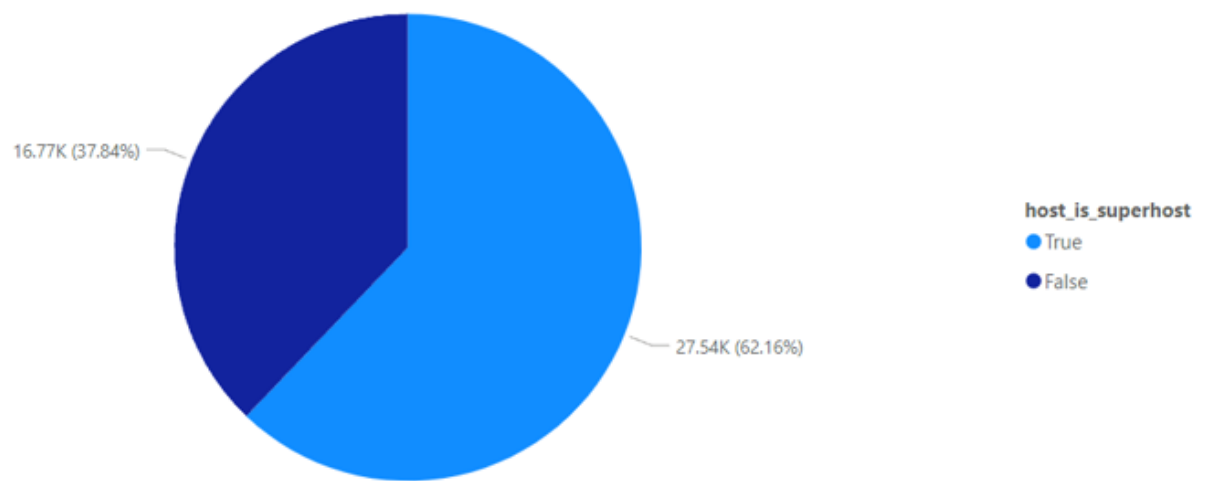


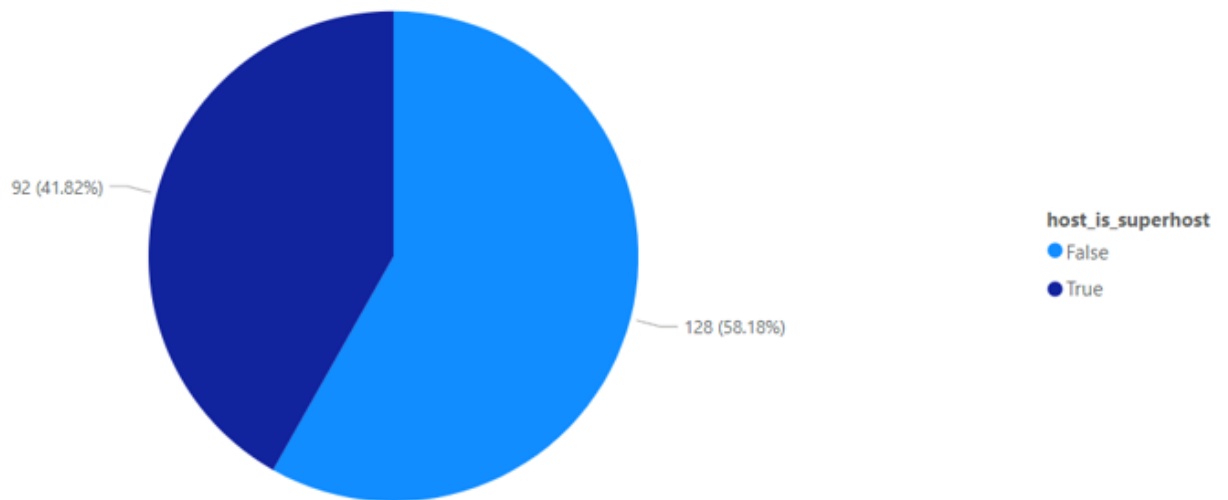
Fig. 11 Polarity Distribution based on review scores in cleanliness.

Super-host: Figure 14 depicts a 60% of super-hosts within positive comments, indicating that if a host holds super-host status does significantly influence the positivity of sentiments. Conversely, negative sentiments are notably affected when the host is not a super-host. Figure 10 illustrates that 58% of dissatisfied customers were hosted by non-super-hosts. This finding aligns with [38], who asserted that super host services have a more pronounced impact on satisfaction. Individuals tend to have certain expectations regarding service quality, and meeting these expectations may not notably enhance their satisfaction, while falling short of these expectations can significantly diminish their satisfaction.

As outlined in [10], what sets peer-to-peer (P2P) accommodation apart from traditional hotels is the creation of a sense of place through closer interaction with a resident (host). In this regard, both online and face-to-face interactions with the host can significantly influence customer satisfaction due to the atmosphere the host creates. Having his identity verified and responding to customer will make the customer feel safe during his stay and about his reservation, contributing to their satisfaction as our OLR test agrees and pinpoints the statistically significant of that. This observation aligns with our contribution, highlighting that minimal host interactions can lead to customer dissatisfaction.



Superhost Proportion in Positive Comments



Superhost Proportion in Negative Comments

Fig. 14 Superhost Proportion in Positive and Negative Comments.

Property Type: Another influential factor is the type of property, room in boutique hotel, entire vacation home, private room in serviced apartment, room in bed and breakfast according to the OLR test (Table II). Figure 12 and 13 illustrates the percentage of negative(red) and positive comments(blue) among the breakfast in bed room and room in boutique hotel. The percentage of positive comments is overwhelming in both cases and for entire vacation home and private room in serviced apartment there are only positive comments for these two categories. This is absolutely an indicator of customer's satisfaction when it comes to these kinds of properties making them a good choice for whoever wants to start an Airbnb accommodations business in Crete or for guest willing to visit Crete.

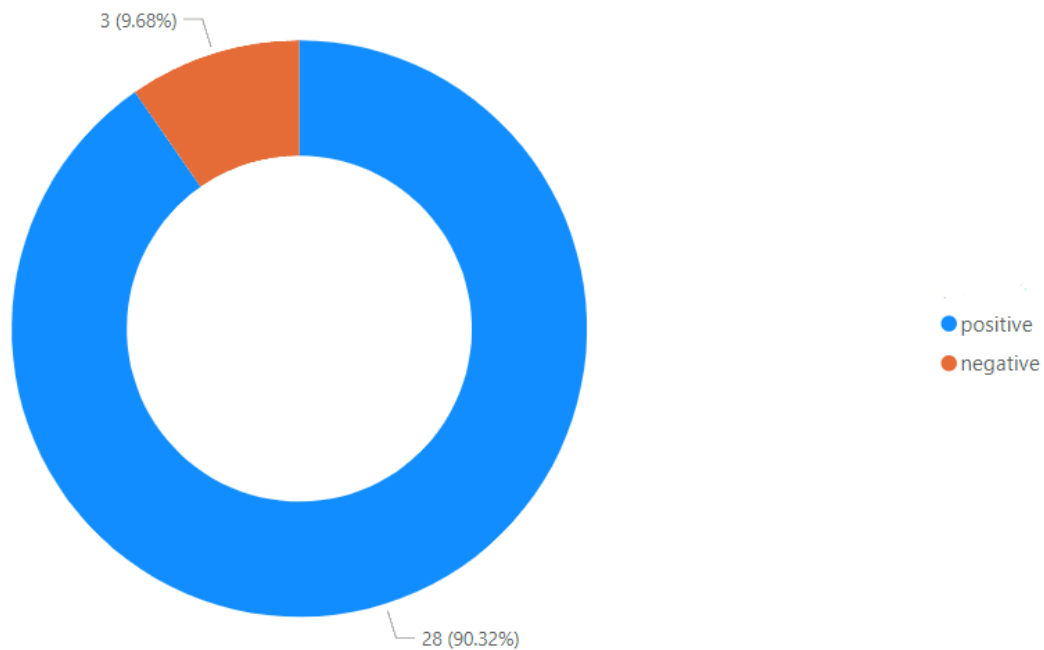


Fig. 12 Property Type: Room and breakfast in bed proportion in Positive and Negative comments.

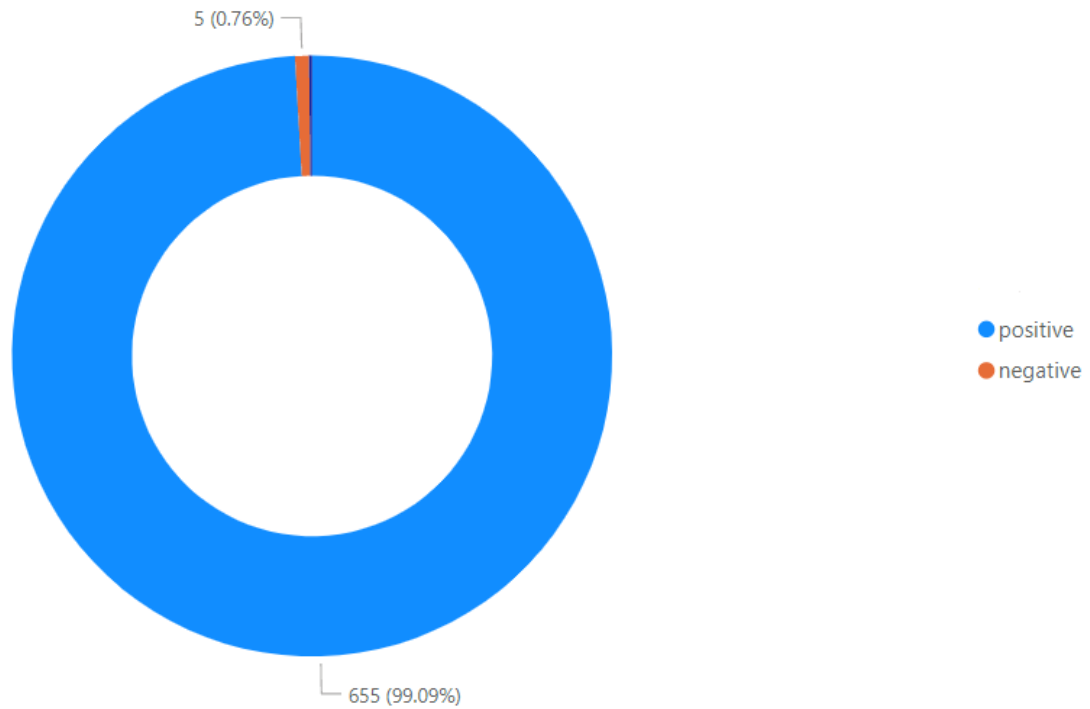


Fig. 13 Property Type: Room in boutique hotel proportion in Positive and Negative comments.

V. CONCLUSION AND FUTURE WORK

Airbnb serves as a prominent embodiment of the sharing economy at large, with a primary focus on hospitality. Despite continued scrutiny and attention within the industry, prior literature on Airbnb has predominantly scrutinized customer reviews in isolation. This study marks a pioneering effort to scrutinize Airbnb's listing features as determinants influencing review polarity. The findings reveal that cleanliness, the host's super-host status, identity verification, response rate and property type significantly impact customer satisfaction. Additionally, in the topic clustering analysis, positive comments prominently feature the word "host", "clean", "place/apartment" while negative comments are characterized by also "host", and "place/apartment". However, the study affirms that Crete itself plays a role in influencing customer satisfaction on Airbnb, potentially leading to positive feedback as comments tend to be overwhelmingly positive. To mitigate potential biased sampling effects, the study presents factors as a percentage of each factor's polarity within the same polarity total in visualization. However, limitations exist, particularly concerning the use of a single sample (Crete), which may impact the generalizability of the results due to non-random sampling.

This study opens avenues for future research. Comparative analyses across different countries could help discern common factors from cultural influences, shedding light on whether customer satisfaction is influenced by a country's overall environment. Additionally, exploring text attributes such as room

descriptions and host "about" sections in relation to polarity could provide valuable insights. Finally, the study introduces a novel method for identifying factors influencing customer satisfaction through the combination of sentiment analysis and regression tests, validated by visualization. This method holds potential applicability across various studies in hospitality and beyond.

REFERENCES

- [1] Georgiadis L, Iosifidis G, Tassioulas L. On the Efficiency of Sharing Economy Networks. *IEEE Trans Netw Sci Eng* 2019;4697:1–1. <https://doi.org/10.1109/tnse.2019.2904431>.
- [2] Ju Y, Back KJ, Choi Y, Lee JS. Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *Int J Hosp Manag* 2019; 77:342–52. <https://doi.org/10.1016/j.ijhm.2018.07.014>.
- [3] Sigala M, Rahimi R, Thelwall M. Big data and innovation in tourism, travel, and hospitality: Managerial approaches, techniques, and applications. *Big Data Innov Tour Travel Hosp Manag Approaches, Tech Appl* 2019:1–223. <https://doi.org/10.1007/978-981-13-6339-9>.
- [4] Cui L, Hou Y, Gao M, Yang Y. Exploring the influencing factors of sharing economy sustainability based on a two-mode social network analysis. 2019 16th Int Conf Serv Syst Serv Manag ICSSSM 2019 2019:1–5. <https://doi.org/10.1109/ICSSSM.2019.8887647>.
- [5] Martin-Fuentes E, Fernandez C, Mateu C, Marine-Roig E. Modelling a grading scheme for peer-to-peer accommodation: Stars for Airbnb. *Int J Hosp Manag* 2018;69:75–83. <https://doi.org/10.1016/j.ijhm.2017.10.016>.
- [6] Blal I, Singal M, Templin J. Airbnb's effect on hotel sales growth. *Int J Hosp Manag* 2018;73:85–92. <https://doi.org/10.1016/j.ijhm.2018.02.006>.
- [7] Guttentag DA, Smith SLJ. Assessing Airbnb as a disruptive innovation relative to hotelsSubstitution and comparative performance expectations. *Int J Hosp Manag* 2017;64:1–10. <https://doi.org/10.1016/j.ijhm.2017.02.003>.
- [8] Xie KL, Kwok L. The effects of Airbnb's price positioning on hotel performance. *Int J Hosp Manag* 2017;67:174–84. <https://doi.org/10.1016/j.ijhm.2017.08.011>.
- [9] Lin PMC, Fan DXF, Zhang HQ, Lau C. Spend less and experience more: Understanding tourists' social contact in the Airbnb context. *Int J Hosp Manag* 2019;83:65–73. <https://doi.org/10.1016/j.ijhm.2019.04.007>.
- [10] Moon H, Miao L, Hanks L, Line ND. Peer-to-peer interactions: Perspectives of Airbnb guests and hosts. *Int J Hosp Manag* 2019;77:405–14. <https://doi.org/10.1016/j.ijhm.2018.08.004>.

- [11] Su N, Mattila AS. Does gender bias exist? The impact of gender congruity on consumer's Airbnb booking intention and the mediating role of trust. *Int J Hosp Manag* 2019. <https://doi.org/10.1016/j.ijhm.2019.102405>.
- [12] Liang S, Schuckert M, Law R, Chen CC. The importance of marketergenerated content to peer-to-peer property rental platforms: Evidence from Airbnb. *Int J Hosp Manag* 2020;84:102329. <https://doi.org/10.1016/j.ijhm.2019.102329>.
- [13] Liu SQ, Mattila AS. Airbnb: Online targeted advertising, sense of power, and consumer decisions. *Int J Hosp Manag* 2017;60:33–41. <https://doi.org/10.1016/j.ijhm.2016.09.012>.
- [14] Falk M, Larpin B, Scaglione M. The role of specific attributes in determining prices of Airbnb listings in rural and urban locations. *Int J Hosp Manag* 2019; 83:132–40. <https://doi.org/10.1016/j.ijhm.2019.04.023>.
- [15] Yang SB, Lee K, Lee H, Koo C. In Airbnb we trust: Understanding consumers' trust-attachment building mechanisms in the sharing economy. *Int J Hosp Manag* 2019;83:198–209. <https://doi.org/10.1016/j.ijhm.2018.10.016>.
- [16] Wang C (Renee), Jeong M. What makes you choose Airbnb again? An examination of users' perceptions toward the website and their stay. *Int J Hosp Manag* 2018; 74:162–70. <https://doi.org/10.1016/j.ijhm.2018.04.006>.
- [17] Cheng M, Jin X. What do Airbnb users care about? An analysis of online review comments. *Int J Hosp Manag* 2019;76:58–70. <https://doi.org/10.1016/j.ijhm.2018.04.004>.
- [18] Luo Y, Tang R (Liang). Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *Int J Hosp Manag* 2019;80:144–54. <https://doi.org/10.1016/j.ijhm.2019.02.008>.
- [19] Zhang M, Luo X, Cheng X, Fu S. An empirical business study on service providers' satisfaction in sharing economy. *Proc - 2016 IEEE 1st Int Conf Data Sci Cyberspace, DSC 2016* 2017;514–9. <https://doi.org/10.1109/DSC.2016.115>.
- [20] Hawlitschek F, Teubner T, Gimpel H. Understanding the sharing economy - Drivers and impediments for participation in peer-to-peer rental. *Proc Annu Hawaii Int Conf Syst Sci* 2016;2016-March:4782–91. <https://doi.org/10.1109/HICSS.2016.593>.
- [21] Alamsyah A, Rochmah WY, Nugroho DDA. Understanding Public Opinion towards New Sharing Economy Business Model Using Content Analysis. *Proc 2019 Int Conf Inf Manag Technol ICIMTech* 2019 2019;1:300–4. <https://doi.org/10.1109/ICIMTech.2019.8843779>.
- [22] Sari R, Lisanti Y, Luhukay D. The Evaluation of Impact Sharing Economy Critical Success Factor Implementation (case study: Online hospitality). *Proc 2019 Int Conf Inf Manag Technol ICIMTech* 2019 2019;1:455–60. <https://doi.org/10.1109/ICIMTech.2019.8843746>.
- [23] Ch C, Gupta D. Factors influencing customer satisfaction with usage of shopping apps in India. *RTEICT 2017 - 2nd IEEE Int Conf Recent Trends Electron Inf Commun Technol Proc* 2017;2018-Janua:1483–6. <https://doi.org/10.1109/RTEICT.2017.8256844>.

- [24] Kumar V, Reinartz W. Applications of CRM in B2B and B2C Scenarios Part I. 2018. https://doi.org/10.1007/978-3-662-55381-7_16.
- [25] Zhao Y, Xu X, Wang M. Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *Int J Hosp Manag* 2019;76:111–21. <https://doi.org/10.1016/j.ijhm.2018.03.017>.
- [26] Lutz C, Newlands G. Consumer segmentation within the sharing economy: The case of Airbnb. *J Bus Res* 2018;88:187–96. <https://doi.org/10.1016/j.jbusres.2018.03.019>.
- [27] Edelman BG, Luca M. Digital Discrimination: The Case of Airbnb.com. *SSRN Electron J* 2014. <https://doi.org/10.2139/ssrn.2377353>.
- [28] Guo L, Li J, Wu J, Chang W, Wu J. A Novel Airbnb Matching Scheme in Shared Economy Using Confidence and Prediction Uncertainty Analysis. *IEEE Access* 2018;6:10320–31. <https://doi.org/10.1109/ACCESS.2018.2801810>.
- providers' satisfaction in sharing economy. *Proc - 2016 IEEE 1st Int Conf Data Sci Cyberspace, DSC 2016* 2017;514–9. <https://doi.org/10.1109/DSC.2016.115>.
- [20] Hawlitschek F, Teubner T, Gimpel H. Understanding the sharing economy - Drivers and impediments for participation in peer-to-peer rental. *Proc Annu Hawaii Int Conf Syst Sci* 2016;2016-March:4782–91. <https://doi.org/10.1109/HICSS.2016.593>.
- [21] Alamsyah A, Rochmah WY, Nugroho DDA. Understanding Public Opinion towards New Sharing Economy Business Model Using Content Analysis. *Proc 2019 Int Conf Inf Manag Technol ICIMTech* 2019 2019;1:300–4. <https://doi.org/10.1109/ICIMTech.2019.8843779>.
- [22] Sari R, Lisanti Y, Luhukay D. The Evaluation of Impact Sharing Economy Critical Success Factor Implementation (case study: Online hospitality). *Proc 2019 Int Conf Inf Manag Technol ICIMTech* 2019 2019;1:455–60. <https://doi.org/10.1109/ICIMTech.2019.8843746>.
- [23] Ch C, Gupta D. Factors influencing customer satisfaction with usage of shopping apps in India. *RTEICT 2017 - 2nd IEEE Int Conf Recent Trends Electron Inf Commun Technol Proc* 2017;2018-Janua:1483–6. <https://doi.org/10.1109/RTEICT.2017.8256844>.
- [24] Kumar V, Reinartz W. Applications of CRM in B2B and B2C Scenarios Part I. 2018. https://doi.org/10.1007/978-3-662-55381-7_16.
- [25] Zhao Y, Xu X, Wang M. Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *Int J Hosp Manag* 2019;76:111–21. <https://doi.org/10.1016/j.ijhm.2018.03.017>.
- [26] Lutz C, Newlands G. Consumer segmentation within the sharing economy: The case of Airbnb. *J Bus Res* 2018;88:187–96. <https://doi.org/10.1016/j.jbusres.2018.03.019>.
- [27] Edelman BG, Luca M. Digital Discrimination: The Case of Airbnb.com. *SSRN Electron J* 2014. <https://doi.org/10.2139/ssrn.2377353>.
- [28] Guo L, Li J, Wu J, Chang W, Wu J. A Novel Airbnb Matching Scheme in Shared Economy Using Confidence and Prediction Uncertainty Analysis. *IEEE Access* 2018;6:10320–31. <https://doi.org/10.1109/ACCESS.2018.2801810>.

- [29] Nguyen LS, Ruiz-Correa S, Mast MS, Gatica-Perez D. Check Out This Place: Inferring Ambiance from Airbnb Photos. *IEEE Trans Multimed* 2018;20:1499–511. <https://doi.org/10.1109/TMM.2017.2769444>.
- [30] Situmorang KM, Hidayanto AN, Wicaksono AF, Yuliawati A. Analysis on customer satisfaction dimensions in peer-to-peer accommodation using latent dirichlet allocation: A case study of airbnb. *Int Conf Electr Eng Comput Sci Informatics* 2018;2018-Octob:542–7. <https://doi.org/10.1109/EECSI.2018.8752912>.
- [31] Park E. Motivations for customer revisit behavior in online review comments: Analyzing the role of user experience using big data approaches. *J Retail Consum Serv* 2019;51:14–8. <https://doi.org/10.1016/j.jretconser.2019.05.019>.
- [32] Cuquet M, Fensel A, Bigagli L. A European research roadmap for optimizing societal impact of big data on environment and energy efficiency. *GloTS 2017 - Glob Internet Things Summit, Proc 2017*. <https://doi.org/10.1109/GIOTS.2017.8016274>. [33] Inside Airbnb n.d. <http://insideairbnb.com/>.
- [34] Li J, Lowe D, Wayment L, Huang Q. Text mining datasets of β hydroxybutyrate (BHB) supplement products' consumer online reviews. *Data Br* 2020;30. <https://doi.org/10.1016/j.dib.2020.105385>.
- [35] Osborne JW. Simple Linear Models With Polytomous Categorical Dependent Variables: Multinomial and Ordinal Logistic Regression. *Regres Linear Model Best Pract Mod Methods* 2020:133–56. <https://doi.org/10.4135/9781071802724.n6>.
- [36] Chattopadhyay M, Mitra SK. Do airbnb host listing attributes influence room pricing homogenously? *Int J Hosp Manag* 2019;81:54–64. <https://doi.org/10.1016/j.ijhm.2019.03.008>.
- [37] The Economist Intelligence Unit. Safe Cities Index 2019 2019.
- [38] Heyam Abdullah Bin Madhi1 , Muna M. Alhammad What Drives Airbnb Customers' Satisfaction in Amsterdam? A Sentiment Analysis
- [39] Rohan Singh, Pankaj Sharma Sentiment Analysis using Microsoft Azure Machine Learning and Python, *International Journal of Engineering Research & Technology (IJERT)* ,Vol. 10 Issue 11, November-2021
- [40] Vasileios Hatzivassiloglou and Kathleen R. McKeown. 1997. Predicting the Semantic Orientation of Adjectives
- [41] Janyce M. Wiebe, Rebecca F. Bruce, and Thomas P. O'Hara. 1999. Development and Use of a Gold-Standard Data Set for Subjectivity Classifications
- [42] Bing Liu. *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, May 2012
- [43] Alan O. Sykes, "An Introduction to Regression Analysis" (Coase-Sandor Institute for Law & Economics Working Paper No. 20, 1993).
- [44] Godovykh, M., Back, R.M., Bufquin, D., Baker, C. and Park, J.-Y. (2023), "Peer-to-peer accommodation amid COVID-19: the effects of Airbnb cleanliness information on guests' trust and behavioral intentions", *International Journal of Contemporary Hospitality Management*, Vol. 35 No. 4, pp. 1219-1237. <https://doi.org/10.1108/IJCHM-12-2021-1508>