How Does Customer Feedback Vary Across Leading Airline Review Platforms?

Adeim Suvanbekova

Christian González-Martel

Teresa Aguiar-Quintana

2024-07-25

Abstract

This study investigates score heterogeneity in online consumer reviews (OCRs) for airlines across TripAdvisor and Skytrax. The research addresses the variability in review scores, focusing on two airlines: Ryanair and Lufthansa. Previous studies have largely ignored cross-platform score variability, and this study fills that gap by applying a Bayesian hierarchical model. This model is suitable for its ability to incorporate prior knowledge and handle complex data structures. The dataset spans from January 1, 2023, to December 31, 2023. Results indicate no significant differences in review scores between platforms or airlines, suggesting consistent customer satisfaction. However, a significant interaction effect between TripAdvisor and Ryanair indicates lower scores for this combination. These findings highlight the importance of platform-airline interactions in customer feedback analysis. The study’s implications underscore the need for tailored strategies in managing OCRs and validate the use of Bayesian methodologies in hospitality research. Future research should expand the dataset for broader validation.

## 1 Introduction

As technology continues to advance industries globally and digitalization transforms operations, the hospitality sector has undergone significant changes. This evolution , driven by Web 2.0 technologies, has given rise to online review platforms featuring user-generated content (UGC), providing valuable insights for prospective customers and fundamentally reshaping customer engagement with businesses. Thus, online consumer reviews (OCRs) have begun to significantly influence consumer decision-making processes (Mendes-Filho et al. 2018), shaping brand image (Lam, Ismail, and Lee 2020) , and enhancing electronic Word-of-Mouth (eWOM) (Muda and Hamzah 2021; Nusairat et al. 2021). Furthermore, within the academic field, scholars are increasingly recognizing the advantages of applying UGC due to its accessibility and capacity for in-depth analysis. Consequently, there has been a notable growth in scholarly research incorporating UGC.

Given the unique characteristics of airline services, both in-flight and on-ground, effective customer communication and feedback are indicators of the operational efficiency and overall customer experience. Existing literature on airline services heavily relies on surveys as the primary data source to explore perceptions of service quality and customer satisfaction. The recent increasing integration of online reviews into scholarly research is attributed to accessibility, diverse perspectives, and rich data content, thereby overcoming the sampling limitations associated with traditional surveys. The incorporation of UGC within the airline industry has been widely utilized for various research aims, such as assessing passengers’ emotions following flight delays (Song, Guo, and Zhuang 2020), investigating the impact of airline alliance strategy on service quality (Migdadi 2022), and predicting airline recommendations. (Jain et al. 2022; Siering, Deokar, and Janze 2018). However, despite the growing literature utilizing custom-shared experiences, there remains a gap in understanding the variability in online customer ratings within the airline industry. To address this gap, this study proposes a Bayesian approach for measuring and managing score heterogeneity across major online review platforms.

**(Methodogy/Need to check)** Recently, (Martel–Escobar, González-Martel, and Vázquez-Polo 2023) noted a lack of research on the application of Bayesian techniques in publications within the tourism and hospitality sectors, particularly in studies focused on eWOM. While Bayesian approaches have found extensive adoption across various fields such as finance, healthcare, and marketing over the past decades, their application in hospitality research remains limited. Despite numerous advantages, including the incorporation of prior data, flexibility, and the ability to handle small sample sizes, the integration of Bayesian methods in hospitality studies remains constrained. According to Bianchi and Heo (2021) Bayesian statistics provide a more nuanced analysis compared to classical methods, which is particularly beneficial for the hospitality sector. Through the analysis of data obtained from primary OCR platforms for airlines, namely TripAdvisor and Skytrax, the study presents a novel Bayesian approach to understanding score heterogeneity among the principal review portals for airlines.

The airline industry has experienced notable transformations following the introduction of Low-Cost Carriers (LCCs), which have impacted passengers’ perceptions and shaped distinct expectations according to airline’s business model. The evaluation of services provided by LCCs differ compared to Full-Service Carriers (FSCs) due to disparities in airfare pricing and customer anticipations. This distinction is underscored by the customer satisfaction coefficient, revealing the heightened importance of different attributes (Byun and Lee 2016). Thus, to conduct a comprehensive analysis, both LCCs and FSCs will be examined.

Given research presents the methodology and findings of Bayesian analysis, based on data collected from leading online review platforms for both LCCs and FSCs. Through in-depth approach, the study not only contributes to the existing literature on UGC and airline service quality but also brings attention to rating heterogeneity, offering a comprehensive understanding of score variations. These insights are valuable for theoretical scholars utilizing UGC for their analysis, as the majority of studies tend to concentrate on a single UGC platform, while also holding practical implications for stakeholders seeking to enhance service quality, foster eWOM, and manage brand image effectively within the highly competitive industry.

## 2 Literature Review

### 2.1 The importance of online consumer reviews in tourism.

Online reviews have captured considerable attention of both academia and industry due to their significant impact on consumer behavior and purchasing decisions. Thus, UGC has become increasingly essential in the tourism sector, particularly considering the intangible nature of hospitality and the growing reliance on reviews for travel planning. Notably, such content is considered more credible than official sites and tourist destination pages (Bilos, Budimir, and Hrustek 2022), influences travelers’ behavior more than marketing-generated-content (Tsiakali 2018), and positively influences tourists’ intentions to book a hotel and visit a destination (Ukpabi and Karjaluoto 2018). Research by Cox et al. (2009) in the accommodation sector reveals that the majority of consumers prioritize the views of other travelers over the hotel’s self-description available online. Furthermore, UGC has more credibility and trustworthiness compared to professionally generated content, and this trustworthiness remains consistent regardless of whether participants encounter positive or negative information (Cheong and Morrison 2008).

According to Muritala, Sánchez-Rebull, and Hernández-Lara (2020), the research on online reviews and electronic eWOM in tourism and hospitality is reaching a mature stage, characterized by a rapidly growing volume of studies built on a solid base of seminal research and the application of the latest analytical techniques, including artificial neural networks and deep learning. Academics have extensively studied UGC in relation to travel planning behaviour (Bilos, Budimir, and Hrustek 2022; Cox et al. 2009; Mendes-Filho et al. 2018; Tsiakali 2018), corporate reputation management (Baka 2016; Nicoli and Papadopoulou 2017), motivations for content contribution (Bronner and Hoog 2011; Daugherty, Eastin, and Bright 2008; Kaur and Singh 2021; Yoo and Gretzel 2009), credibility perception (Cheong and Morrison 2008; Filieri et al. 2021; Lo and Yao 2019), helpfulness determinants (Filieri et al. 2021; Hu and Yang 2020; Huang et al. 2015; Kwon et al. 2021; Lo and Yao 2019), predicting recommendations (Jain et al. 2022; Siering, Deokar, and Janze 2018), and service recovery (Etemad-Sajadi and Bohrer 2019; Xu, Liu, and Gursoy 2019).

The increasing adoption of OCRs in scholarly research is attributed to numerous advantages. Firstly, it provides direct access to extensive datasets with organically generated and diverse data. In certain instances, OCRs enable researchers to gather socio-demographic details of content contributors, supporting cross-cultural studies and analyses of age or gender disparities (Dike et al. 2024; Punel, Al Hajj Hassan, and Ermagun 2019; Stamolampros et al. 2019; Sudhakar and Gunasekar 2020). **(DO I NEED TO PUT ARTICLES?)**. Additionally, given the unique nature of airline services, reviews often contain valuable information on cabin class, airline type, and routes for evaluating moderating effect and comparative analysis, along with additional ratings on certain service attributes for extended examination (Lucini et al. 2020; Sezgen, Mason, and Mayer 2019; Tahanisaz and shokuhyar 2020). Furthermore, reviews allow researchers to track changes in customer perceptions, preferences, and behaviors in response to significant external factors such as the COVID-19 pandemic (Hassan and Salem 2022; Nilashi et al. 2022; Sulu, Arasli, and Saydam 2022) and economic crises. However, study by W. Lu and Stepchenkova (2015) underscores that the anonymity of UGC creators, sample representativeness and the generalizability of results poses challenges to research validity, thereby questioning the reliability of findings.

Acknowledging the importance of online reviews, scholars have begun investigating their credibility and willingness to use. Research conducted by Lo and Yao (2019) revealed higher perceived credibility scores among reviews that are negative , written by experts and with a consistent rating. Consumers seek the opinions of others to minimize risks and gather prepurchase information, making these opinions more valuable than advertising. According to Bilos, Budimir, and Hrustek (2022) respondents rely on UGC more than their friends’ recommendations and promotional content. For analyzing the intention to use online consumer reviews, scholars rely on the Technology Acceptance Model (TAM) (Davis 1989) as a predominant framework in hospitality literature, which identifies perceived usefulness and ease of use as key factors. In addition, Changchit, Klaus, and Lonkani (2020) identified perceived credibility and importance of online reviews, along with computer self-efficacy, as additional factors influencing users’ intentions. According to Mendes-Filho et al. (2018), there is a heightened intention to use UGC in travel planning if the content provides perceived empowerment and usefulness. The study by Assaker (2020) found that gender and age significantly moderate travellers’ intention to use UGC, with perceived usefulness being more influential for male travellers and perceived ease of use being more impactful for female and older travellers.

Online travel review authors are primarily motivated by a desire to help service providers, concern for other consumers, and the need for enjoyment and positive self-enhancement (Yoo and Gretzel 2009). Kaur and Singh (2021) analysing motivations behind review submission, defined altruism, reciprocity, egocentric behavior, knowledge self-efficacy, and collectivism as a key antecedents. Identifying profile of vacationers contributing to review sites, Bronner and Hoog (2011) reveal them to typically be under 55 years, from high and lower-middle-income groups, and often couples with or without children. The authors highlighted the importance of distinguishing between self-directed and other-directed motivations, which influence the type of review site and the manner of expression. Self-directed motivations favor marketer-generated sites, commenting on fewer aspects, predominantly negative reviews, and contributing to less accessible platforms. Vacationers with other-directed motivations prefer consumer-generated sites, engaging in broader commenting, posting mainly positive reviews, using both text and ratings, and contributing to sites accessible to other vacationers. According to Wilson, Murphy, and Fierro (2012), national attitudes and motivations significantly influence the type of UGC posted. In their study, authors found that British and Swiss users tend to share photos and videos on social media to relive vacation experiences, while Spanish users prefer using text and comments to make recommendations to other travelers.

With the widespread availability of online reviews and the consequent information overload, scholars have begun to study reviews in relation to their perceived helpfulness. However, despite extensive research into the determinants of review helpfulness within tourism and hospitality, findings remain mixed. A recent study by Hu and Yang (2020) identified review length as the most influential factor, followed by reviewer expertise and age, while the impact of review valence and readability was found to be insignificant. Conversely, Filieri, Hofacker, and Alguezaui (2018) contended that long reviews are not necessarily perceived as helpful. Instead, consumers favor fact-based, objective, and logically structured reviews especially in the context of high involvement purchase decisions. According to Kwok and Xie (2016), the helpfulness of online hotel reviews is negatively affected by review length and rating, whereas positively impacted by manager response and reviewer experience. Huang et al. (2015) determined the relationship between word count and helpfulness remained significant for reviews containing 144 words or fewer. However, once the word count exceeded, the relationship loose statistical significance. Analysing extremely negative ratings, Filieri, Raguseo, and Vitari (2019), found that these ratings are often considered helpful and diagnostic, as they provide valuable insights to justify the extreme evaluation, thereby significantly contributing to consumer decision-making.

### 2.2 The research on OCR applied to airlines industry

Despite the growing adoption of reviews in scholarly research, studies specifically addressing online reviews related to airlines remain limited. Analyzing service quality within tourism and hospitality industry using UGC, W. Lu and Stepchenkova (2015) discovered that the majority of studies focused on the lodging setting, with relatively fewer investigations conducted in the airline, online travel agent, and restaurant settings. Similarly, an analysis of online customer evaluations in tourism and hospitality by Schuckert, Liu, and Law (2015) revealed that over half of the reviewed articles concentrated on hotels. Existing research of airline service through online customer-shared experiences are primarily focused on perceptions of service quality (Brochado, Duarte, and Mengyuan 2023; Korfiatis et al. 2019; Rasool and Pathania 2021), customer satisfaction (Ban and Kim 2019; Chatterjee et al. 2023; Dike et al. 2024; Lucini et al. 2020), and behavioural intentions (Siering, Deokar, and Janze 2018; Wang et al. 2023).

Consumer feedback platforms have undergone a significant expansion, gaining increasing popularity with the advent of Web 2.0. Online reviews assessing airline services within the scholarly literature primarily rely on data sourced from Skytrax and TripAdvisor platforms, with some studies also incorporating UGC sourced from social media, particularly Twitter. (DO I NEED TO ADD TWITTER ARTICLES). A study by L. Lu et al. (2023) underscores the importance of incorporating content from multiple review channels to gain a more accurate understanding of service quality, recognizing the potential disparities across platforms. The study on US airlines revealed that Skytrax reviewers prioritize boarding and baggage experiences, whereas Twitter users emphasize customer service interactions. Bilos, Budimir, and Hrustek (2022) providing a comparative analysis with Airbnb, found that TripAdvisor reviews are perceived as more valuable for making purchasing decisions. Through conducting a comparative analysis between results sourced from Twitter and the Department of Transportation’s official service quality report for US airlines, (Tian et al., 2020) **SEARCH REFERENCE** established the reliability of service quality metrics derived from social media.

Cross-cultural investigations in the field have uncovered differences in the evaluation of provided services, influenced by the cultural backgrounds of reviewers. Punel, Al Hajj Hassan, and Ermagun (2019) revealed the influence of passengers’ country of residence to travel experience, perception, and the evaluation of airline services. Notably, North American passengers prioritize cost over in-flight services and complain more about their national airlines, while East and Southeast Asian passengers are satisfied with Asian airlines and value in-flight services. Middle Easterners are critical of non-local airlines, while Oceania travelers expect high service standards and South Americans have lower expectations for in-flight amenities with an emphasis on price. According to Dike and Davis (2023) **SEARCH REFERENCE** passengers from individualistic cultural backgrounds tend to perceive service quality as lower compared to customers from collectivist cultural backgrounds. Stamolampros et al. (2019) found that airline passengers’ satisfaction with service quality varies depending on their cultural values, impacting both overall satisfaction and specific aspects such as staff interactions, food quality, seat comfort, and cleanliness. Additionally, the study discovered that cultural characteristics significantly shape both numerical ratings and textual content in reviews.

The existing literature examining the airline industry underscores distinct disparities between Low-Cost Carriers (LCCs) and Full-Service Carriers (FSCs), emphasizing the significant impact of airline category on passenger expectations and perceptions of service quality. Consequently, each type of carrier should strategically prioritize specific features to enhance overall service delivery. According to Lim and Lee (2020), FSCs should focus on enhancing tangible features, while LCCs should concentrate on reliability. Furthermore, variations in recommendation drivers exist across different airline business models. (Siering et al., 2018) revealed that LCCs should focus on improving augmented service aspects like amenities to enhance recommendations, while FSNCs should prioritize core service aspects such as in-flight amenities and customer service quality.

#### 2.2.1 OCR studies based on data collected in Tripadvisor.

Tripadvisor, the world’s largest travel guidance platform, features over 1 billion reviews and is accessible in 43 markets and 22 languages (Tripadvisor, 2024) **SEARCH REFERENCE**. In addition to its general community guidelines, the platform has established specific posting guidelines for reviews, which are displayed on an organization’s listing page and are associated with a bubble rating ranging from 1 to 5. Beyond overall ratings, contributors evaluate additional service attributes, including legroom, seat comfort, in-flight entertainment, onboard experience, customer service, value for money, cleanliness, check-in and boarding processes, as well as food and beverage quality.

The platform has been extensively utilized as a data source by various investigators worldwide, contributing to diverse research findings. In particular, online reviews serve as an efficient tool for examining customer satisfaction. Sudhakar and Gunasekar (2020) analyzed Indian airlines using TripAdvisor feedbacks to reveal varying factors influencing satisfaction across nationalities and airline types, with Indian passengers prioritizing value for money and foreign passengers prioritizing service attributes. Sezgen, Mason, and Mayer (2019) examined 45 international airlines to define the key drivers of passenger satisfaction and dissatisfaction across full-service and low-cost carriers, as well as economy and premium cabins. Despite variations in satisfaction drivers, fundamental factors such as friendly staff, value, and service quality remain universally important. In the US airline market, Park, Lee, and Nicolau (2020) underscores the need for sustainable strategies beyond price competition, highlighting the asymmetric impact of service quality attributes on customer satisfaction and their critical importance in airline operations. According to Korfiatis et al. (2019), integrating unstructured data from online reviews into service quality assessment enhances understanding and predictability of customer satisfaction. Examining Tripadvisor reviews, the study revealed temporal variations in dimensions, thereby emphasizing the importance of monitoring and adapting to changing customer preferences over time.

Additionally, Tripadvisor reviews have been used to identify key themes by Brochado et al. (2019) and Sulu, Arasli, and Saydam (2022) during the pandemic. The study by Brochado et al. (2019) was aimed to identify the main themes in traveler reviews and their relationship with value-for-money ratings by examining airlines affiliated with the largest alliances, SkyTeam and Star Alliance, across France, the United States, Indonesia, Canada, Taiwan, and Germany. Whereas, Sulu’s (2022) research focused on online airline reviews during the COVID-19 outbreak, aiming to define principal themes and identify disparities among passengers of various nationalities. The analysis revealed that European passengers exhibited the most polarized opinions, ranging from highly critical to highly satisfied, while American and Australian passengers predominantly expressed positive reviews.

#### 2.2.2 OCR studies based on data collected in Skytrax.

Skytrax is the international air transport rating organisation with Air Travel review forum (www.airlinequality.com) that has become a leading review site for airline, airport and associated air travel traveller reviews **SEARCH REFERENCE** (Skytrax, 2024). Reviews submitted to the platform undergo a multi-stage checking process to ensure compliance with the Editorial Policy and requires reviewers to submit feedback that follows a structured evaluation framework. The platform enables passengers to share detailed opinions on their experiences across four categories: airline, seat, airport, and lounge. Passengers can rate airlines on a scale from 1 to 10 and provide textual comment, along with additional details such as country of residence, date of flight, route, cabin class, aircraft type, type of travel, and likelihood of recommending the airline. Moreover, contributors assess supplementary service attributes including value for money, ground service, seat comfort, cabin staff service, food & beverages, in-flight entertainment, and cabin Wi-Fi & connectivity.

Various scholars have used Skytrax reviews to define key factors influencing customer experience and satisfaction within the airline sector. Ban and Kim (2019) identified six key clusters, including seat comfort, staff service, entertainment, ground service, value for money, and airline brand, with all factors except entertainment significantly impacting customer satisfaction and recommendation. Shadiyar et al. (2020) **SEARCH REFERENCE** explored the utilization of big data analytics and online customer reviews to enhance customer satisfaction in airlines, identifying factors such as seat comfort, staff behavior, food and beverage quality, ground service, and particularly value for money as main. The study by Lucini et al. (2020) was aimed to measure customer satisfaction by analyzing Skytrax reviews to provide airlines with guidelines to improve competitiveness. The study revealed that the type of cabin significantly impacts customer satisfaction dimensions, prioritizing customer service for first-class passengers, comfort for premium economy passengers, and efficient baggage handling and minimal waiting times for economy class travellers, whereas the type of passenger had minimal impact on satisfaction dimensions. Dike et al. (2024) analyzed Skytrax feedbacks to explore how customer culture, travel purpose, and seat type shape customer expectations and satisfaction. The study highlighted that customer service quality, delays, and baggage handling significantly impact perceived service performance, with empathy and reliability having the strongest influence on perceived satisfaction, while cultural backgrounds affect levels of customer satisfaction. Chatterjee et al. (2023) conducted a ten-year analysis of Skytrax reviews, identifying core attributes such as value for money, seating comfort, and cabin staff behaviour as more strongly related to customer satisfaction and recommendation than augmented aspects like food and beverages and in-flight entertainment. Authors found that core service attributes such as responsiveness and reliability are crucial in shaping customer satisfaction, whereas tangibles and empathy significantly influence customers’ likelihood of recommending an airline.

Also, Skytrax forum have been extensively used to analyze service quality perceptions worldwide. Brochado’s (2023) analysis of four leading Chinese airlines highlighted value for money as consistently the most influential service attribute, with varying responses from different traveler types to service attributes. Kwon et al. (2021) used the platform to understand customer needs and preferences in the Asian airline market. Lim and Lee (2020) used Skytrax feedbacks to compare service quality perceptions between full-service carriers and low-cost carriers , finding tangibles as the most significant dimension for FSCs and reliability for LCCs. In terms of varied expectations of service quality, Punel, Al Hajj Hassan, and Ermagun (2019) revealed that different regions exhibit distinct priorities and expectations concerning in-flight services and overall satisfaction. Passengers in first or business class prioritize seat comfort, food, and beverages, as well as in-flight entertainment. In contrast, passengers in economy class place a higher emphasis on value for money.

**The rest of this paper is organised as follows**

## 3 Data & Methods

### 3.1 Bayesian Approach

This study adopts a Bayesian approach to analyze the score heterogeneity of online consumer reviews (OCRs) across major airline review platforms, specifically TripAdvisor and Skytrax.

Bayesian methods have been increasingly recognized in various fields for their robustness in handling complex data structures and integrating prior knowledge. The researcher starts with an initial belief about a parameter () and its associated probability distribution (), named *prior probability distribution*. Then data are collected . These data follow a distribution that depends on the parameter , known as *likelihood function*. The researcher’s belief of are updated applying Bayes’ theorem, combinig both into a *posterior probability function*

In the context of online consumer reviews (OCRs), Bayesian approaches are particularly advantageous because they allow us to account for the variability in review scores across different platforms and airlines. Bianchi and Heo (2021) highlight the advantages of Bayesian statistics in hospitality research, particularly for analyzing subjective guest reviews. Martel–Escobar, González-Martel, and Vázquez-Polo (2023) highlighted the importance of managing score heterogeneity between online consumer review websites, offering an alternative to frequentist methods. Gómez-Déniz, Martel-Escobar, and Vázquez-Polo (2024) use a robust Bayesian methodology to manage uncertainty with prior distributions, effectively transforming measures of interest into actionable decision-making intervals.

### 3.2 Data description

The data for this study was collected from two primary OCR platforms: TripAdvisor and Skytrax. These platforms were chosen due to their popularity and extensive use in academic research on airline services. **[Citation needed]** The dataset comprises reviews from two distinct airlines: Ryanair, representing the Low-Cost Carriers (LCCs), and Lufthansa, representing the Full-Service Carriers (FSCs). This differentiation allows for a comparative analysis of customer perceptions and expectations across different business models within the airline industry. The reviews span the period from January 1, 2023, to December 31, 2023, covering a full year of reviews.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Descriptive Statistics of Airline Review Scores by Site for the period from 2023-01-01 to 2023-12-31   | Descriptive Statistics of Airline Review Scores by Site | | | | | | --- | --- | --- | --- | --- | | For the period from 2023-01-01 to 2023-12-31 | | | | | | Site | Airline | Score | | N | | Mean | SD | | SKYTRAX | LUFTHANSA | 3.008734 | 2.795660 | 229 | | SKYTRAX | RYANAIR | 3.520661 | 3.356832 | 121 | | TRIPADVISOR | LUFTHANSA | 2.885443 | 2.026448 | 1999 | | TRIPADVISOR | RYANAIR | 2.833981 | 2.088866 | 3861 | |

Source: [Exploratory Analysis](https://chrglez.github.io/hierarchical_bayesian_airlines/notebooks\exploratory-analysis-preview.html#cell-tbl-descriptive)

[Table 1](#tbl-descriptive) presents various descriptive statistics about the distributions of ratings for Skytrax and TripAdvisor for the two airlines studied. These descriptive statistics provide a detailed overview of the data, highlighting central tendencies, variability, and other relevant aspects of the rating distributions.

These descriptive statistics indicate the mean scores, standard deviations, and the number of reviews for each airline on both platforms, providing a clear picture of the data distribution and the volume of customer feedback available for analysis.

|  |
| --- |
| Figure 1: Number of reviews by rating according to website (‘Skytrax’ and ‘Tripadvisor’) and airlines (Ryanair and Lufthansa) |

Source: [Exploratory Analysis](https://chrglez.github.io/hierarchical_bayesian_airlines/notebooks\exploratory-analysis-preview.html#cell-fig-histogram)

Our analysis employs a standard 1–5 scale, with the Skytrax sample data linearly transformed to match this conventional scale.

The histograms in [Figure 1](#fig-histogram) illustrate the frequency of different review scores, showing a clear left skew (negative skewness) in the distribution of scores for both airlines on both platforms. This left skewness indicates that a majority of the reviews are clustered towards the lower end of the rating scale. Specifically, a significant number of reviews have scores of 1 or 2, while higher scores are less frequent. This pattern suggests a tendency for customers to leave lower ratings, which could be indicative of more frequent negative experiences or higher expectations not being met.

### 3.3 Bayesian Hierarchical Model

To analyze score heterogeneity, we applied a Bayesian hierarchical model. Hierarchical models, also known as multilevel models, are useful when data are structured in groups or clusters. In this study, reviews are grouped by the site and airline. This hierarchical structure allows us to account for the variability at both the site and airline levels.

Bayesian hierarchical models are powerful tools for analyzing data with nested or grouped structures. These models have two main levels:

* **Fixed Effects**: These are the overall effects that apply to the entire population. In our study, the fixed effects include the site and airline.
* **Random Effects**: These account for the variability within groups. In our study, the random effects capture the variability within each site and airline.

In our analysis, following Bianchi and Heo (2021), we model the review scores using a Poisson distribution.

This choice is appropriate because the scores are count data, which can naturally be modeled using a Poisson distribution. The Poisson distribution is widely used for modeling the number of events occurring within a fixed interval of time or space when these events happen with a known constant mean rate and independently of the time since the last event. This fits well with the nature of review scores, which can be considered as counts of discrete ratings.

The general form of a Bayesian hierarchical model can be expressed as:

Where:

* is the expected score for review on site for airline .
* is the overall intercept.
* is the effect of site .
* is the effect of airline .
* is the interaction effect between site and airline .
* is the random error term.

The results of the Bayesian hierarchical model applied to analyze the heterogeneity of review scores are presented in **?@tbl-results**

::: {#tbl-results .cell tbl-alt=‘Summary of Bayesian Hierarchical Model Results including Fixed and Random Effects.’ tbl-cap=’ Summary of Bayesian Hierarchical Model Results including Fixed (population-level) and Random Effects (group-level).’}

#| label: tbl-results  
#| tbl-alt: "Summary of Bayesian Hierarchical Model Results including Fixed and Random Effects."  
#| tbl-cap: " Summary of Bayesian Hierarchical Model Results including Fixed (population-level) and Random Effects (group-level)."  
  
fit |> tbl\_regression(tidy\_fun = broom.mixed::tidy,  
 intercept=TRUE,  
 estimate\_fun = ~ style\_sigfig(.x, digits = 3))

Characteristic

Beta

95% CI

1

(Intercept)

1.05

-3.67, 6.17

site

    SKYTRAX

—

—

    TRIPADVISOR

-0.034

-4.65, 4.39

airline

    LUFTHANSA

—

—

    RYANAIR

0.144

-5.67, 5.87

site \* airline

    TRIPADVISOR \* RYANAIR

-0.178

-0.299, -0.050

airline.sd\_\_(Intercept)

1.71

0.095, 4.84

site.sd\_\_(Intercept)

1.68

0.074, 5.02

1

CI = Credible Interval

Source: [Hierarchical Bayesian Model](https://chrglez.github.io/hierarchical_bayesian_airlines/notebooks\hierarchical-bayesian-model-preview.html#c77b80fb-5926-4250-a51e-5e1fb0d7bd62)

Plot the results

:::

The overall intercept of 1.05 indicates the baseline level of review scores when other factors are held constant. However, the wide 95% confidence interval ranging from -3.67 to 6.17 suggests high uncertainty around this estimate, indicating that the baseline level can vary significantly.

The effect of the site being TripAdvisor compared to Skytrax is estimated at -0.03. This small effect size, coupled with a 95% confidence interval between -4.65 and 4.39, indicates that there is no significant difference in review scores between TripAdvisor and Skytrax. The wide confidence interval reflects substantial uncertainty and variability in this comparison.

The effect of Ryanair compared to Lufthansa is estimated at 0.14. Similar to the site effect, the 95% confidence interval of -5.67 to 5.87 shows a wide range, indicating that there is no significant difference in review scores between these two airlines. This suggests that the review scores for Ryanair and Lufthansa are statistically similar.

The interaction between the site TripAdvisor and the airline Ryanair is estimated at -0.18, with a 95% confidence interval from -0.30 to -0.05. This negative interaction effect indicates that the combination of TripAdvisor and Ryanair tends to have lower review scores compared to other combinations. The narrow confidence interval suggests that this interaction effect is statistically significant, implying a specific negative effect when these two factors are combined.

The standard deviation of the intercept for the airlines is 1.71, with a 95% confidence interval between 0.10 and 4.84. This indicates a considerable amount of variability in the baseline review scores across different airlines. The wide confidence interval suggests that while there is some variability, the exact amount is uncertain.

The standard deviation of the intercept for the sites is 1.68, with a 95% confidence interval between 0.07 and 5.02. Similar to the airline intercept variability, this shows significant variability in the baseline review scores across different sites. The wide confidence interval reflects uncertainty in the extent of this variability.

## 4 Conclusion

This study explored the score heterogeneity of online consumer reviews (OCRs) across two major airline review platforms—TripAdvisor and Skytrax—using a Bayesian hierarchical model. Our findings provide valuable insights into how review scores vary across different platforms and airlines, contributing to the broader understanding of customer feedback in the airline industry.

The analysis revealed no significant differences in review scores between TripAdvisor and Skytrax or between Ryanair and Lufthansa. This suggests that, on average, customer satisfaction as reflected in the review scores is consistent across these platforms and airlines. However, the interaction effect between TripAdvisor and Ryanair indicates that specific platform-airline combinations can influence customer perceptions, highlighting the nuanced nature of OCRs.

The significant variability in baseline scores across platforms and airlines underscores the importance of considering both fixed and random effects in understanding score heterogeneity. This finding aligns with previous research (Martel–Escobar, González-Martel, and Vázquez-Polo 2023; Bianchi and Heo 2021), emphasizing the need for robust statistical methods to manage the diverse sources of OCR data.

Despite the strengths of our study, including the application of advanced Bayesian hierarchical models, there are limitations that need to be addressed. The analysis was constrained by the specific dataset, which covered only one year of reviews for two airlines. Future research should consider a broader dataset, including more airlines and a longer timeframe, to validate and extend our findings.

Furthermore, the Bayesian hierarchical approach provides a framework for incorporating prior knowledge and handling complex data structures, making it a powerful tool for analyzing OCRs. However, further work is needed to refine these models and explore their application in other contexts within the tourism and hospitality sectors (Gómez-Déniz, Martel-Escobar, and Vázquez-Polo 2024).

In summary, our study contributes to the understanding of score heterogeneity in OCRs and offers practical implications for airline industry stakeholders. By highlighting the importance of platform-airline interactions and the variability in customer feedback, we provide insights that can help improve service quality and manage brand reputation effectively. Future research should continue to explore these dynamics, leveraging the strengths of Bayesian methodologies to enhance our understanding of customer satisfaction in the digital age.

The findings of this study underscore the necessity for tailored strategies in managing OCRs and provide a foundation for future research in this area. By continuing to investigate the complex interactions between review platforms, airlines, and customer perceptions, we can develop more effective approaches to improving customer satisfaction and fostering positive electronic word-of-mouth (eWOM) in the highly competitive airline industry.

## References

Assaker, G. 2020. “Age and Gender Differences in Online Travel Reviews and User-Generated-Content (UGC) Adoption: Extending the Technology Acceptance Model (TAM) with Credibility Theory.” *Journal of Hospitality Marketing and Management* 29 (4). <https://doi.org/10.1080/19368623.2019.1653807>.

Baka, V. 2016. “The Becoming of User-Generated Reviews: Looking at the Past to Understand the Future of Managing Reputation in the Travel Sector.” *Tourism Management* 53. <https://doi.org/10.1016/j.tourman.2015.09.004>.

Ban, H. J., and H. S. Kim. 2019. “Understanding Customer Experience and Satisfaction Through Airline Passengers’ Online Review.” *Sustainability (Switzerland* 11 (15). <https://doi.org/10.3390/su11154066>.

Bianchi, G., and C. Y. Heo. 2021. “A Bayesian Statistics Approach to Hospitality Research.” *Current Issues in Tourism* 24 (22). <https://doi.org/10.1080/13683500.2021.1896486>.

Bilos, A., B. Budimir, and A. Hrustek. 2022. “The Role of User-Generated Content in Tourists’ Travel Planning Behavior: Evidence from Croatia.” *Journal of Tourism and Development* 39. <https://doi.org/10.34624/rtd.v39i0.27277>.

Brochado, A., M. Duarte, and Z. Mengyuan. 2023. “Passengers’ Perceptions of Chinese Airlines’ Service Quality: A Mixed Methods Analysis of User-Generated Content.” *Journal of China Tourism Research* 19 (3). <https://doi.org/10.1080/19388160.2022.2122647>.

Brochado, A., P. Rita, C. Oliveira, and F. Oliveira. 2019. “Airline Passengers’ Perceptions of Service Quality: Themes in Online Reviews.” *International Journal of Contemporary Hospitality Management* 31 (2). <https://doi.org/10.1108/IJCHM-09-2017-0572>.

Bronner, F., and R. Hoog. 2011. “Vacationers and eWOM: Who Posts, and Why, Where, and What?” *Journal of Travel Research* 50 (1). <https://doi.org/10.1177/0047287509355324>.

Byun, H. J., and B. C. Lee. 2016. “Classifying Service Quality Attributes of Low-Cost Carriers and Full-Service Carriers Based on an Analytical Kano Model.” *Global Business and Finance Review* 21 (2). <https://doi.org/10.17549/gbfr.2016.21.2.34>.

Changchit, C., T. Klaus, and R. Lonkani. 2020. “Online Reviews: What Drives Consumers to Use Them.” *Journal of Computer Information Systems*. <https://doi.org/10.1080/08874417.2020.1779149>.

Chatterjee, S., A. Ghatak, R. Nikte, S. Gupta, and A. Kumar. 2023. “Measuring SERVQUAL Dimensions and Their Importance for Customer-Satisfaction Using Online Reviews: A Text Mining Approach.” *Journal of Enterprise Information Management* 36 (1). <https://doi.org/10.1108/JEIM-06-2021-0252>.

Cheong, H. J., and M. A. Morrison. 2008. “Consumers’ Reliance on Product Information and Recommendations Found in UGC.” *Journal of Interactive Advertising* 8 (2). <https://doi.org/10.1080/15252019.2008.10722141>.

Cox, C., S. Burgess, C. Sellitto, and J. Buultjens. 2009. “The Role of User-Generated Content in Tourists’ Travel Planning Behavior.” *Journal of Hospitality and Leisure Marketing* 18 (8). <https://doi.org/10.1080/19368620903235753>.

Daugherty, T., M. S. Eastin, and L. Bright. 2008. “Exploring Consumer Motivations for Creating User-Generated Content.” *Journal of Interactive Advertising* 8 (2). <https://doi.org/10.1080/15252019.2008.10722139>.

Davis, F. D. 1989. “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology.” *MIS Quarterly: Management Information Systems* 13 (3). <https://doi.org/10.2307/249008>.

Dike, S. E., Z. Davis, A. Abrahams, A. Anjomshoae, and P. Ractham. 2024. “Evaluation of Passengers’ Expectations and Satisfaction in the Airline Industry: An Empirical Performance Analysis of Online Reviews.” *Benchmarking* 31 (2). <https://doi.org/10.1108/BIJ-09-2021-0563>.

Etemad-Sajadi, R., and L. Bohrer. 2019. “The Impact of Service Recovery Output/Process on Customer Satisfaction and Loyalty: The Case of the Airline Industry.” *Tourism and Hospitality Research* 19 (2). <https://doi.org/10.1177/1467358417743080>.

Filieri, R., F. Acikgoz, V. Ndou, and Y. Dwivedi. 2021. “Is TripAdvisor Still Relevant? The Influence of Review Credibility, Review Usefulness, and Ease of Use on Consumers’ Continuance Intention.” *International Journal of Contemporary Hospitality Management* 33 (1). <https://doi.org/10.1108/IJCHM-05-2020-0402>.

Filieri, R., C. F. Hofacker, and S. Alguezaui. 2018. “What Makes Information in Online Consumer Reviews Diagnostic over Time? The Role of Review Relevancy, Factuality, Currency, Source Credibility and Ranking Score.” *Computers in Human Behavior* 80. <https://doi.org/10.1016/j.chb.2017.10.039>.

Filieri, R., E. Raguseo, and C. Vitari. 2019. “What Moderates the Influence of Extremely Negative Ratings? The Role of Review and Reviewer Characteristics.” *International Journal of Hospitality Management* 77. <https://doi.org/10.1016/j.ijhm.2018.07.013>.

Gómez-Déniz, E., M. Martel-Escobar, and F. J. Vázquez-Polo. 2024. “A Bayesian Model for Online Customer Reviews Data in Tourism Research: A Robust Analysis.” *Cogent Business & Management* 11 (1). <https://doi.org/10.1080/23311975.2024.2363592>.

Hassan, T. H., and A. E. Salem. 2022. “Impact of Service Quality of Low-Cost Carriers on Airline Image and Consumers’ Satisfaction and Loyalty During the Covid-19 Outbreak.” *International Journal of Environmental Research and Public Health* 19 (1). <https://doi.org/10.3390/ijerph19010083>.

Hu, X., and Y. Yang. 2020. “What Makes Online Reviews Helpful in Tourism and Hospitality? A Bare-Bones Meta-Analysis.” *Journal of Hospitality Marketing and Management*. <https://doi.org/10.1080/19368623.2020.1780178>.

Huang, A. H., K. Chen, D. C. Yen, and T. P. Tran. 2015. “A Study of Factors That Contribute to Online Review Helpfulness.” *Computers in Human Behavior* 48. <https://doi.org/10.1016/j.chb.2015.01.010>.

Jain, P. K., A. Patel, S. Kumari, and R. Pamula. 2022. “Predicting Airline Customers’ Recommendations Using Qualitative and Quantitative Contents of Online Reviews.” *Multimedia Tools and Applications* 81 (5). <https://doi.org/10.1007/s11042-022-11972-7>.

Kaur, K., and T. Singh. 2021. “What Motivates Consumers to Write Online Reviews? Qualitative Research in the Indian Cultural Context.” *Journal of Global Marketing* 34 (3). <https://doi.org/10.1080/08911762.2021.1882022>.

Korfiatis, N., P. Stamolampros, P. Kourouthanassis, and V. Sagiadinos. 2019. “Measuring Service Quality from Unstructured Data: A Topic Modeling Application on Airline Passengers’ Online Reviews.” *Expert Systems with Applications* 116. <https://doi.org/10.1016/j.eswa.2018.09.037>.

Kwok, L., and K. L. Xie. 2016. “Factors Contributing to the Helpfulness of Online Hotel Reviews: Does Manager Response Play a Role?” *International Journal of Contemporary Hospitality Management* 28. <https://doi.org/10.1108/IJCHM-03-2015-0107>.

Kwon, H. J., H. J. Ban, J. K. Jun, and H. S. Kim. 2021. “Topic Modeling and Sentiment Analysis of Online Review for Airlines.” *Information Switzerland* 12 (2). <https://doi.org/10.3390/info12020078>.

Lam, J. M. S., H. Ismail, and S. Lee. 2020. “From Desktop to Destination: User-Generated Content Platforms, Co-Created Online Experiences, Destination Image and Satisfaction.” *Journal of Destination Marketing and Management* 18. <https://doi.org/10.1016/j.jdmm.2020.100490>.

Lim, J., and H. C. Lee. 2020. “Comparisons of Service Quality Perceptions Between Full Service Carriers and Low Cost Carriers in Airline Travel.” *Current Issues in Tourism* 23 (10). <https://doi.org/10.1080/13683500.2019.1604638>.

Lo, A. S., and S. S. Yao. 2019. “What Makes Hotel Online Reviews Credible?: An Investigation of the Roles of Reviewer Expertise, Review Rating Consistency and Review Valence.” *International Journal of Contemporary Hospitality Management* 31 (1). <https://doi.org/10.1108/IJCHM-10-2017-0671>.

Lu, L., P. Xu, Y. Y. Wang, and Y. Wang. 2023. “Measuring Service Quality with Text Analytics: Considering Both Importance and Performance of Consumer Opinions on Social and Non-Social Online Platforms.” *Journal of Business Research* 169. <https://doi.org/10.1016/j.jbusres.2023.114298>.

Lu, W., and S. Stepchenkova. 2015. “User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software.” *Journal of Hospitality Marketing and Management* 24 (2). <https://doi.org/10.1080/19368623.2014.907758>.

Lucini, F. R., L. M. Tonetto, F. S. Fogliatto, and M. J. Anzanello. 2020. “Text Mining Approach to Explore Dimensions of Airline Customer Satisfaction Using Online Customer Reviews.” *Journal of Air Transport Management* 83. <https://doi.org/10.1016/j.jairtraman.2019.101760>.

Martel–Escobar, M., C. González-Martel, and F. J. Vázquez-Polo. 2023. “Managing Score Heterogeneity Between Online Consumer Review Websites.” *Cogent Social Sciences* 9 (2). <https://doi.org/10.1080/23311886.2023.2267261>.

Mendes-Filho, L., A. M. Mills, F. B. Tan, and S. Milne. 2018. “Empowering the Traveler: An Examination of the Impact of User-Generated Content on Travel Planning.” *Journal of Travel and Tourism Marketing* 35 (4). <https://doi.org/10.1080/10548408.2017.1358237>.

Migdadi, Y. K. A. A. 2022. “The Impact of Airline Alliance Strategy on the Perceived Service Quality: A Global Survey.” *Journal of Quality Assurance in Hospitality and Tourism* 23 (2). <https://doi.org/10.1080/1528008X.2021.1884929>.

Muda, M., and M. I. Hamzah. 2021. “Should i Suggest This YouTube Clip? The Impact of UGC Source Credibility on eWOM and Purchase Intention.” *Journal of Research in Interactive Marketing* 15 (3). <https://doi.org/10.1108/JRIM-04-2020-0072>.

Muritala, B. A., M. V. Sánchez-Rebull, and A. B. Hernández-Lara. 2020. “A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality.” *Sustainability (Switzerland* 12 (ue 23)). <https://doi.org/10.3390/su12239977>.

Nicoli, N., and E. Papadopoulou. 2017. “TripAdvisor and Reputation: A Case Study of the Hotel Industry in Cyprus.” *EuroMed Journal of Business* 12 (3). <https://doi.org/10.1108/EMJB-11-2016-0031>.

Nilashi, M., R. A. Abumalloh, B. Minaei-Bidgoli, W. Abdu Zogaan, A. Alhargan, S. Mohd, S. N. F. Syed Azhar, S. Asadi, and S. Samad. 2022. “Revealing Travellers’ Satisfaction During COVID-19 Outbreak: Moderating Role of Service Quality.” *Journal of Retailing and Consumer Services* 64. <https://doi.org/10.1016/j.jretconser.2021.102783>.

Nusairat, N. M., M. A. Alroale, M.Al Qeed, J. A. Al-Gasawneh, Q. Hammouri, A. Ahmad, and H. Abdellatif. 2021. “USER-GENERATED CONTENT.” *Academy of Strategic Management Journal* 20 (4).

Park, S., J. S. Lee, and J. L. Nicolau. 2020. “Understanding the Dynamics of the Quality of Airline Service Attributes: Satisfiers and Dissatisfiers.” *Tourism Management* 81. <https://doi.org/10.1016/j.tourman.2020.104163>.

Punel, A., L. Al Hajj Hassan, and A. Ermagun. 2019. “Variations in Airline Passenger Expectation of Service Quality Across the Globe.” *Tourism Management* 75. <https://doi.org/10.1016/j.tourman.2019.06.004>.

Rasool, G., and A. Pathania. 2021. “Reading Between the Lines: Untwining Online User-Generated Content Using Sentiment Analysis.” *Journal of Research in Interactive Marketing* 15 (3). <https://doi.org/10.1108/JRIM-03-2020-0045>.

Schuckert, M., X. Liu, and R. Law. 2015. “Hospitality and Tourism Online Reviews: Recent Trends and Future Directions.” *Journal of Travel and Tourism Marketing* 32 (5). <https://doi.org/10.1080/10548408.2014.933154>.

Sezgen, E., K. J. Mason, and R. Mayer. 2019. “Voice of Airline Passenger: A Text Mining Approach to Understand Customer Satisfaction.” *Journal of Air Transport Management* 77. <https://doi.org/10.1016/j.jairtraman.2019.04.001>.

Siering, M., A. V. Deokar, and C. Janze. 2018. “Disentangling Consumer Recommendations: Explaining and Predicting Airline Recommendations Based on Online Reviews.” *Decision Support Systems* 107. <https://doi.org/10.1016/j.dss.2018.01.002>.

Song, C., J. Guo, and J. Zhuang. 2020. “Analyzing Passengers’ Emotions Following Flight Delays- a 2011–2019 Case Study on SKYTRAX Comments.” *Journal of Air Transport Management* 89. <https://doi.org/10.1016/j.jairtraman.2020.101903>.

Stamolampros, P., N. Korfiatis, P. Kourouthanassis, and E. Symitsi. 2019. “Flying to Quality: Cultural Influences on Online Reviews.” *Journal of Travel Research* 58 (ue 3)). <https://doi.org/10.1177/0047287518764345>.

Sudhakar, S., and S. Gunasekar. 2020. “Examining Online Ratings and Customer Satisfaction in Airlines.” *Anatolia* 31 (2). <https://doi.org/10.1080/13032917.2020.1747238>.

Sulu, D., H. Arasli, and M. B. Saydam. 2022. “Air‐travelers’ Perceptions of Service Quality During the COVID‐19 Pandemic: Evidence from TripAdvisor.com.” *Sustainability (Switzerland* 14 (1). <https://doi.org/10.3390/su14010435>.

Tahanisaz, S., and S. shokuhyar. 2020. “Evaluation of Passenger Satisfaction with Service Quality: A Consecutive Method Applied to the Airline Industry.” *Journal of Air Transport Management* 83. <https://doi.org/10.1016/j.jairtraman.2020.101764>.

Tsiakali, K. 2018. “User-Generated-Content Versus Marketing-Generated-Content: Personality and Content Influence on Traveler’s Behavior.” *Journal of Hospitality Marketing and Management* 27 (8). <https://doi.org/10.1080/19368623.2018.1477643>.

Ukpabi, D. C., and H. Karjaluoto. 2018. “What Drives Travelers’ Adoption of User-Generated Content? A Literature Review.” In *Tourism Management Perspectives*. Vol. 28. <https://doi.org/10.1016/j.tmp.2018.03.006>.

Wang, X., J. Zheng, L. Tang, and Y. Luo. 2023. “Recommend or Not? The Influence of Emotions on Passengers’ Intention of Airline Recommendation During COVID-19.” *Tourism Management* 95. <https://doi.org/10.1016/j.tourman.2022.104675>.

Wilson, A., H. Murphy, and J. C. Fierro. 2012. “Hospitality and Travel: The Nature and Implications of User-Generated Content.” *Cornell Hospitality Quarterly (Vol* 53 (ue 3)). <https://doi.org/10.1177/1938965512449317>.

Xu, X., W. Liu, and D. Gursoy. 2019. “The Impacts of Service Failure and Recovery Efforts on Airline Customers’ Emotions and Satisfaction.” *Journal of Travel Research* 58 (6). <https://doi.org/10.1177/0047287518789285>.

Yoo, K. H., and U. Gretzel. 2009. “What Motivates Consumers to Write Online Travel Reviews?” *Information Technology & Tourism* 10 (4). <https://doi.org/10.3727/109830508788403114>.