**T.C.**

**ANTALYA BILIM UNIVERSITY**

**INSTITUTE OF POSTGRADUATE EDUCATION**

**INSTITUTE OF POSTGRADUATE EDUCATION**

**MASTER'S THESIS**

# SENTIMENT ANALYSIS: BRAND PERCEPTION AND CUSTOMER SATISFACTION OF AIRLINE INDUSTRY

**Noor Ul Huda Nasir SIDDIQUI**

**JANUARY 2025**

**ANTALYA**

**T.C.**

**ANTALYA BILIM UNIVERSITY**

**INSTITUTE OF POSTGRADUATE EDUCATION**

**INSTITUTE OF POSTGRADUATE EDUCATION**

**MASTER'S THESIS**

# SENTIMENT ANALYSIS: BRAND PERCEPTION AND CUSTOMER SATISFACTION OF AIRLINE INDUSTRY

**Noor Ul Huda Nasir SIDDIQUI**

**JANUARY 2025**

**ANTALYA**

**T.C.**

**ANTALYA BILIM UNIVERSITY**

**INSTITUTE OF POSTGRADUATE EDUCATION**

# SENTIMENT ANALYSIS: BRAND PERCEPTION AND CUSTOMER SATISFACTION OF AIRLINE INDUSTRY

**Noor Ul Huda Nasir SIDDIQUI**

This thesis was accepted by the jury (with unanimous vote / majority vote) on the date …./…../202… in Business Administration of Business Administration Department.

Assoc. Prof. Dr. A. Mohammad ABUBAKAR (Supervisor) [signature]

Assoc. Prof. Dr. Ahmet TURKMENT [signature]

Assist. Prof. Dr. Aslihan DURSUN [signature]

Director of The Institute

Prof. Dr. Ibrahim Sani MERT [signature]

Thesis Submission Date: …./…../202….

## DECLARATION

Msc Thesis of this study named “Sentiment Analysis: Brand Perception and Customer Satisfaction of Airline Industry”, which I presented, I declare that scientific moral principles were followed in the preparation of this study, in case of benefiting from the works of others, reference is made in accordance with scientific norms, no falsification has been made in the data used, and that any part of this study is not presented as another academic study.

... / ... / 20..

Noor Ul Huda Nasir SIDDIQUI

**CONTENTS**

## ABSTRACT

### SENTIMENT ANALYSIS: BRAND PERCEPTION AND CUSTOMER SATISFACTION OF AIRLINE INDUSTRY

**Noor Ul Huda Nasir SIDDIQUI**

**MSc Thesis in Business Administration**

**Supervisor: Assoc. Prof. Dr. A. Mohammad ABUBAKAR**

**January 2025; 73 pages**

This research investigates customer satisfaction and brand loyalty in the airline industry through an in-depth analysis of customer reviews and overall ratings. By employing sentiment analysis and natural language processing techniques, the study identifies critical factors influencing satisfaction and dissatisfaction across various airlines. Key variables such as overall ratings, review length, traveller demographics, and travel class are examined to uncover patterns in customer sentiment. The findings reveal that positive sentiment is strongly associated with higher ratings, while lower ratings highlight dissatisfaction drivers, including delays and seating comfort. Longer reviews often provide more nuanced insights into customer experiences, reflecting stronger emotional responses. Additionally, the analysis shows that passengers in business and premium economy classes express significantly higher levels of satisfaction compared to those in economy class, particularly on long-haul routes. These results offer practical recommendations for airlines to enhance customer service, improve satisfaction, and strengthen brand loyalty. By pinpointing areas needing improvement and capitalizing on strengths, this study advances the understanding of customer experience management in the aviation sector. It also emphasizes the importance of sentiment analysis as a strategic tool for addressing customer needs, optimizing service quality, and sustaining a competitive edge in the airline industry.



**KEYWORDS:** Brand Loyalty, Customer Satisfaction, NLP, Perception, Sentiment Analysis, Value for Money.

**COMMITTEE:** Assoc. Prof. Dr. A. Mohammad ABUBAKAR (Supervisor)

Assoc. Prof. Dr. Ahmet TURKMEN

Assist. Prof. Dr. Aslihan DURSUN

# ÖZET

## DUYGU ANALİZİ: HAVAYOLU SEKTÖRÜNÜN MARKA ALGISI VE MÜŞTERİ MEMNUNİYETİ

**Noor ul Huda Nasir SIDDIQUI**

**Yüksek Lisans Tezi, İşletme Anabilim Dalı**

**Danışman: Doç. Dr. A. Mohammed ABUBAKAR**

**Ocak 2025; 73 sayfa**

Bu araştırma, müşteri yorumlarının ve genel değerlendirmelerin derinlemesine analizi yoluyla havayolu sektöründe müşteri memnuniyetini ve marka sadakatini araştırmaktadır. Duygu analizi ve doğal dil işleme tekniklerini kullanan çalışma, çeşitli havayollarında memnuniyet ve memnuniyetsizliği etkileyen kritik faktörleri tanımlamaktadır. Genel değerlendirmeler, değerlendirme uzunluğu, yolcu demografisi ve seyahat sınıfı gibi temel değişkenler, müşteri duyarlılığındaki kalıpları ortaya çıkarmak için incelenmiştir. Bulgular, olumlu duyguların daha yüksek derecelendirmelerle güçlü bir şekilde ilişkili olduğunu ortaya koyarken, düşük derecelendirmeler gecikmeler ve oturma konforu gibi memnuniyetsizlik faktörlerini vurgulamaktadır. Daha uzun değerlendirmeler genellikle müşteri deneyimleri hakkında daha incelikli bilgiler sunmakta ve daha güçlü duygusal tepkileri yansıtmaktadır. Ayrıca analiz, business ve premium ekonomi sınıflarındaki yolcuların, özellikle uzun mesafeli rotalarda, ekonomi sınıfındakilere kıyasla önemli ölçüde daha yüksek memnuniyet düzeyleri ifade ettiklerini göstermektedir Bu sonuçlar, havayolu şirketlerine müşteri hizmetlerini geliştirmek, memnuniyeti artırmak ve marka sadakatini güçlendirmek için pratik öneriler sunmaktadır. Bu çalışma, iyileştirilmesi gereken alanları tespit ederek ve güçlü yönlerden yararlanarak, havacılık sektöründe müşteri deneyimi yönetiminin anlaşılmasını ilerletmektedir. Ayrıca, müşteri ihtiyaçlarının karşılanması, hizmet kalitesinin optimize edilmesi ve havayolu endüstrisinde rekabet avantajının sürdürülmesi için stratejik bir araç olarak duygu analizinin önemini vurgulamaktadır.



**ANAHTAR KELİMELER:** Marka Sadakati, Müşteri Memnuniyeti, NLP, Algısı, Duygu Analizi, Paranın Değeri

**JÜRİ:** Doç. Dr. A. Mohammad ABUBAKAR (Supervisor)

Doç. Dr. Ahmet TURKMEN

Dr. Ögr Üyesi Aslihan DURSUN

## SYMBOLS AND ABBREVIATIONS

**Symbols**

$R^2$ : R-squared (Coefficient of Determination)

**Abbreviations**

AUC-ROC : Area Under the Receiver Operating Characteristic Curve

SMOTE : Synthetic Minority Over-sampling Technique

PCA : Principal Component Analysis

RMSE : Root Mean Squared Error

MAE : Mean Absolute Error

EVS : Explained Variance Score

JIT : Just-In-Time

NLP : Natural Language Processing

TF-IDF : Term Frequency-Inverse Document Frequency

EDA : Exploratory Data Analysis

GBDT : Gradient Boosting Decision Tree

CRM : Customer Relationship Management

CS : Customer Satisfaction

BUS : Business

ECO : Economy

INT : International

## LIST OF TABLES

## LIST OF FIGURES

## PREFACE

When I embarked on this thesis journey, I felt as though I was standing at the edge of an unknown path, unsure of what lay ahead. But with time, persistence, and the encouragement of those around me, the shadows of doubt gradually gave way to clarity and determination. I began to discover the topic of my interest and field which led me to choose this topic.

This research is deeply rooted in my interest in understanding the dynamics of sentiment analysis, customer satisfaction and brand loyalty, especially within the ever-evolving and competitive airline industry.

I owe my deepest gratitude to my Professor, Dr. A. Mohammed AbuBakar, whose insightful guidance, constructive critique, and unwavering support provided me with the direction and clarity needed to complete this work when I needed it most. The thoughtful advice often reminded me that progress is not about perfection but about consistent effort, and for that, I am profoundly thankful.

This research, focused on **sentiment analysis, brand perception, and customer satisfaction in the airline industry**, demanded not just intellectual rigor but also a deep curiosity to uncover hidden patterns and insights. Exploring this intersection of data and human sentiment has been both challenging and deeply rewarding. It has given me a newfound appreciation for the power of analytics to address real-world problems and improve experiences.

This thesis reflects not only my academic pursuits but also my deep-seated curiosity about the intricacies of human sentiments and how they influence decision-making and brand perceptions. The process of diving into natural language processing and sentiment analysis felt like deciphering a unique language—one that reveals the unspoken truths behind customer satisfaction and dissatisfaction.

The journey was not without its moments of frustration and doubt. Yet, each obstacle became an opportunity to learn and grow, and I am immensely grateful to everyone who supported me along the way. To my family, whose unwavering belief in me was a source of strength, and to my friends, who lent a listening ear and offered words of encouragement during difficult times—thank you for walking beside me.

Now, as I write these words, I feel a profound sense of fulfillment. Completing this thesis has been more than an academic exercise; it has been a journey of personal growth, resilience, and discovery. I hope this work not only contributes meaningfully to the academic community but also inspires others to embrace challenges, knowing that perseverance always leads to growth. This work is dedicated to everyone who believes in the power of data-driven strategies to enhance the human experience.

# 1. INTRODUCTION

## 1.1. Airline Industry Scope

The global airline industry serves as a cornerstone of the global economy, reflecting its vast scale and economic significance. In 2023, airlines carried approximately **4.4 billion passengers** globally, with 1.8 billion on international routes and 2.6 billion on domestic flights, showcasing a remarkable recovery nearing pre-pandemic levels (ATAG). The industry directly employed **11.6 million people**, including nearly 4 million in airlines, air navigation, and airports, 1.7 million in aircraft manufacturing, and 5.9 million in other on-airport roles. Additionally, the sector indirectly supported a total of **86.5 million jobs globally** (ATAG). From a financial perspective, global airline revenues were approximately **$918 billion** in 2023, with passenger air traffic contributing **$522 billion**, reflecting a 19% year-over-year increase (Statista; Statista). These statistics underscore the airline industry's integral role in global economic recovery and connectivity. With millions of passengers flying annually, customer satisfaction and brand loyalty have become paramount for airlines to maintain competitiveness and profitability. However, despite advancements in service quality, many airlines struggle to meet rising customer expectations due to evolving travel demands, economic fluctuations, and the immediacy of social media feedback.

Stakeholders, including airlines, policymakers, and researchers, have a vested interest in understanding what drives satisfaction or dissatisfaction among customers. Insights derived from sentiment analysis empower these stakeholders to improve service delivery, enhance customer experiences, and foster loyalty. This study addresses the growing need to decode customer sentiments as airlines navigate the challenges of maintaining relevance and trust in an increasingly competitive market. Understanding customer sentiment provides airlines with valuable information to improve service quality, address customer concerns, and build brand loyalty, which is essential for customer retention and competitive positioning (Kim and Lee, 2020). The brands have customers with different opinions, and knowing how each individual feels about the brand, their products, and their services has become critical for the brands to develop structured insights (Keller, 2020).

This study focuses to assess customer satisfaction and brand loyalty by examining customer reviews from multiple airlines, utilizing sentiment analysis techniques to identify patterns and key drivers of satisfaction and dissatisfaction. By analysing variables like overall rating, review length, and traveller demographics, this research seeks to uncover actionable insights for improving airline services and enhancing the overall customer experience. This approach aligns with the objectives of assessing sentiment trends across different airlines, evaluating satisfaction by travel class, and identifying factors influencing recommendation likelihood, thereby contributing to existing literature on customer experience management in the airline industry.

## 1.2. Problem Statement

Existing research establishes that customer satisfaction and loyalty are influenced by factors such as seating comfort (Ban and Kim 2019), in-flight entertainment and cleanliness are important for economy-plus class passengers (Salah et al, 2024), punctuality (Jiang X, Zhang Y 2022), and service quality e.g. the most satisfied passengers are traveling from business class for business purposes. As they enjoy amenities from the business class, they tend to be more satisfied (Salah et al, 2024). It employs various machine learning algorithms to predict customer satisfaction levels. Secondary dataset has been used to evaluate these factors. However, what remains underexplored is the utilization of unstructured textual data, such as customer reviews, to uncover deeper insights into customer sentiment.

Without leveraging sentiment insights, airlines risk failing to address customer concerns, enhance brand perception, and preempt potential crises. This gap highlights the importance of understanding nuanced sentiment patterns that can only be extracted through advanced techniques like Natural Language Processing (NLP). By addressing this gap, this study contributes to the field of customer experience management, empowering airlines to leverage data-driven decision-making for better outcomes.

## 1.3. Aims and Objectives

This study aims to uncover actionable insights into customer satisfaction and brand loyalty within the airline industry through the application of sentiment analysis. By analysing online reviews, the research seeks to identify the primary factors that drive customer satisfaction, such as seating comfort, in-flight entertainment, and service quality. Additionally, it investigates the relationship between satisfaction levels and loyalty indicators, including repeat booking intentions and the likelihood of recommendation. The study also explores variations in customer sentiment across different passenger classes, such as economy and business, and route types, including domestic and international flights, to provide a deeper understanding of distinct customer segments. Furthermore, it highlights the practical utility of sentiment analysis as a tool for enhancing real-time customer engagement, managing crises, and improving overall brand strategy. These insights aim to equip airlines with the data-driven tools necessary to refine their service delivery and foster long-term customer loyalty.

## 1.4. Research Question

- RQ1: How do specific service factors, such as seating comfort, in-flight entertainment, and staff responsiveness, influence customer satisfaction levels as expressed through sentiment in airline reviews?
- RQ2: What is the relationship between customer satisfaction and brand loyalty indicators (e.g., repeat booking intentions, recommendation likelihood) as reflected in sentiment across multiple airlines?

- RQ3: How does customer sentiment differ across flight characteristics, such as class (economy vs. business) and route type (domestic vs. international), and what do these differences reveal about satisfaction and loyalty within distinct customer segments?

### 1.5. Dissertation Structure

This dissertation comprises five chapters. Chapter 1 introduces the study, outlining its background, aims, research questions, and methodology. Chapter 2 reviews literature on sentiment analysis and brand perception to establish a theoretical foundation. Chapter 3 details the research methodology, including data collection, sentiment analysis, and statistical evaluation. Chapter 4 presents key findings with visual trends and comparative models. Finally, Chapter 5 summarizes the findings, highlight limitations, provides recommendations, and suggests directions for future research.

# 2. LITERATURE REVIEW

## 2.1. Theoretical Framework

This study underlined the Theory of Planned Behaviour (TPB). This psychological model makes predictions regarding behaviours based on attitudes, subjective norms, perceived behavioural control, and their influential role on intention and action (Ajzen, 2020). According to this research context, TPB accounts for how the social environment and past experiences influencing consumer perceptions of a specific brand impact their behaviour towards that brand. Sentimental analysis on the pages of an organization on social networks gives results on user preferences and perceived behavioral control that can aid the brand in predicting consumer behavior and appeal.

Brand Equity Theory points out the importance of consumer perception for creating a brand's value. Coelho et al. (2018) note that brand equity is accumulated through positive brand associations and experiences. This theory relates to the goals of sentiment analysis because the consumer sentiment that brands monitor can be used to immediately quantify and influence brand value. Likewise, Social Media Engagement Theory as postulated by Luttrell and Wallace (2024) argues that engagement on social media makes brand emotions real. To enhance the strategies in multinational markets, Airline industry can know the areas that CA visitors find interesting and engaging by observing consumer activity on social network sites.

Therefore, a strong theoretical foundation of consumer sentiment analysis and behavior monitoring derived from the integration of the TPB, together with Brand Equity Theory and Social Media Engagement Theory comes into play when the mentioned theoretical models are integrated. These behavioural predictors serve as valuable learning in its aim to enhance the overall brand perception and, thus, customer loyalty.

## 2.2. Brand Perception and Sentiment Analysis: An Overview

Brand perception refers to how customers view and interpret a brand from their experiences, engagements, and the communications that come to them through the different channels (Azemi et al.,2022). Honorable brand perception is an important factor for organizations (Khan et al.,2019). Product Quality, service to the customer, price, and cultural relevance are basic issues as per customers. They enable brands and their consumers to interact directly with each other instantaneously (Dzilska et al.,2024). This may lead to real-time responses, reviews, and discussions of the brand. Public sentiment through social media mainly influences how a brand perceives its audience (Kotras,2020).

The procedure for evaluating public opinions and emotions toward a brand is known as sentiment analysis, which is usually run through automated textual content analysis (Xu et al.,2022). This would mean processing the consumer comments, reviews, and posts to understand which one would prevail, whether the total tendency of sentiments is positive or negative, or if it is neutral (Nawaz and Kashif, 2021). It can provide actionable insights regarding customer satisfaction, preferences, and areas for improvement. It allows brands to customize their marketing mix, improve customer experience, and even offset negative reviews by engaging proactively (Rane,2023).

Sentiment analysis employs natural language processing algorithms and techniques coupled with machine learning algorithms to identify specific words, tones, and emotional cues in text data (Aftab et al., 2023; Du Chau,2024). Sentiment analysis is a very important tool in telling how public perception will develop over digital platforms and with insights that can shape effective strategies in brand management. Brand identity is tied to portraying traditional craftsmanship, meaning cultural alignment becomes a very important responsibility to maintain relevance and appeal (Khawaja, 2023).

### 2.3 The Role of Social Media

Social media has changed the face of brand-audience engagement. It has been the beginning of public perception, through the largest social network, Facebook, and the fast-paced Twitter (Ismadiyarov, 2024). The platforms facilitate two-way communication between the brand and its customers, and can obtain insights into the perceptions that it receives and respond promptly to praise and criticism (Bozkurt et al., 2024). Social media networking allows for fast dissemination of brand-related content, which has an influence on positioning and perception across different demographic segments. Posts, comments, and reviews shared by consumers contribute to the total brand narrative which could either be positive or negative (Koiou, 2021). Consumer-generated content is rich enough for detailed analysis that would allow it to understand its perceived brand. Through tracking conversations, mentions, and hashtags from its brand, and can obtain a general view of the public's sentiment for its brands and the main trends in how consumers are giving feedback (Bell et al., 2024).

It enables a brand to relate intimately with its customers while running focused campaigns, promotions, and interactions and can project a customer-centric image by responding to the wants and preferences of its audience in real time (Al-Ansari, 2019). The ability to respond and interact with consumers in an appropriate and timely manner reinforces positive perceptions of the brand. The most tangible benefits of social media for brands are that the platform can amplify positive brand experiences through testimonials from customers, partnerships with influencers, and other user-generated content, thus dramatically expanding a brand's reputation (Hosain and Mamun, 2023). However, it has a negative aspect that bad reviews reach as far and as fast as good ones, causing a threat to brand perception. A solitary sour review or complaint may mature eventually into a full-fledged public relations crisis that can tarnish the brand image (Jiang et al., 2021). Social media is the most important platform for creating an impression of the brand, and it has both opportunities and challenges. Due to the immediacy, reach, and interactivity of this platform, the tool is highly influential in deciding how consumers perceive the brand (Payton, 2021).

## 2.4 Sentiment Analysis

Social Media forums are excellent analysis grounds for sentiment analysis because they give insight into consumer opinions, preferences, and emotions towards a brand (Bruce, 2020). The dynamic nature of the fashion industry necessitates constant monitoring of consumer sentiments on these platforms to maintain a positive brand image, identify areas for improvement, and respond swiftly to emerging trends and issues (Tejpal, 2020). These platforms are utilized for feedback regarding personal experiences with products or services and public reactions to interactions with stakeholders, making them valuable sources for sentiment analysis (Zunic et al., 2020). They offer interactive engagement through posts, comments, likes, and shares, which brands can leverage to track conversations on brand-related topics and understand customer feedback tones. Positive feedback may indicate satisfaction with product quality, service, or available deals, while negative feedback may reflect dissatisfaction with aspects like quality, pricing, or delivery times. On Facebook, customers can provide detailed reviews, offering deeper insights into potential issues affecting brand perception. Users often express opinions promptly in response to new products, promotions, or controversies (Zunic et al., 2020).

The brevity of tweets, coupled with the use of hashtags, facilitates easier trend monitoring. Twitter's dynamic environment allows for real-time tracking of public sentiment, enabling brands to be more responsive to shifts in consumer opinions. Influencers and public figures can significantly amplify positive sentiments or exacerbate negative ones. Through sentiment analysis, brands can determine the emotions—positive, neutral, or negative—expressed by consumers, streamlining the process using natural language processing algorithms that extract key themes and direct attention to critical areas (Eginli and Taş, 2023). Brands may detect an increase in negative sentiments toward a particular product line and implement corrective measures accordingly.

However, sentiment analysis on these platforms is not without challenges. Sarcasm, humor, and cultural differences can complicate the interpretation of consumer sentiments, potentially skewing results. A comprehensive survey by Ghosh and Veale (2023) discusses the challenges of sarcasm detection in sentiment analysis, highlighting the complexities involved. Despite these challenges, sentiment analysis on platforms like Facebook and Twitter offers substantial opportunities for understanding public opinion and proactively managing brand reputation.

Sentiment analysis, especially on Facebook and Twitter/X, is extremely crucial in identifying the possible emergence of crises, which it can curb before they spin out of control. This would allow to react on time to emerging issues before it develops into a more significant threat to its brand image (Hobbs and O'Keefe, 2024). Sentiment analysis tools allow brands to wade through vast amounts of social data and sense even minute changes in consumer emotions (Sümer and Parilti, 2023). Brand may flag spiky negative sentiments or recurring

negative themes to determine the root causes of such dissatisfaction. The sooner a crisis can be detected, the easier it will be for the brand to resolve problems directly through customer interaction and come up with quick fixes before the situation turns into a full-scale crisis. The brand might then respond to the complaint through public forums or provide compensation that might help regain consumer loyalty (Yuan et al.,2020). If consumers feel that the quality of the product is deteriorating or they are not satisfied with the launch of a product, negative word-of-mouth on social media multiplies overnight and becomes a serious public relations disaster. With such an attitude, sentiment analysis helps observe the negative sentiments beforehand and correct them immediately (Cooper et al., 2019).

### 2.5 Comparative Studies: Sentiment Analysis across Platforms

Comparative studies across social media platforms highlight the varying dynamics of sentiment expression on Facebook and Twitter/X. Facebook's design encourages conformity due to its real-name policy and closer social networks, while Twitter's relative anonymity fosters more open expression of opinions and sentiments (Swaminathan et al., 2020; Park et al., 2023). Research demonstrates that Facebook users are more likely to self-disclose within their network, while Twitter/X enables broader discussions with diverse audiences, making it more suitable for real-time sentiment tracking (Nasuto and Rowe, 2024; Kapoor et al., 2022). These platform-specific features significantly influence the results of sentiment analysis, aiding brands in determining the best marketing channels for engaging with their audience and adapting to platform dynamics.

Facebook creates a scenario that incites longer interactions where the users can enter long comments, post long comments and hold discussions in detail. These more detailed forms of interaction are richer data to the social media, enabling brands to know about specific features of offers that customers can like or dislike (Jena, 2019). For instance, it would be comments containing long reviews about more features or issues about customer services or delivery of the product. Analyzing this detailed feedback enables brands to identify areas for improvement and tailor their offerings to better meet customer needs.

Comparative sentiment analysis enables to understand differences in consumers' behaviour and emotions on each platform (Singh et al.,2020). Sentiment on Facebook might allow us to dive deeper into which brand loyalty will strengthen the company in the long term and which customers might appreciate the experience of using the product or service (Maree and Van Heerden 2021). In contrast, the reaction of the customer towards the time-sensitive event could be better captured through a Twitter/X (Polanský and Sylvester-Hvid, 2022). This would transform the way they should create or develop their content and engagement strategies on these platforms, towards developing a coherent, as well as positive, brand image.

Comparing the sentiment analysis across platforms, brands will have a complete view of how it is perceived as a brand in the different digital areas and hence provide better targeted and effective brand management strategies (Xu et al., 2020).

### 2.6 Gaps in Existing Literature

A lot of work has been done in the field of the concept of Sentiment Analysis and Brand Perception but still, there is a huge gap in the existing knowledge, particularly when sentiment analysis regarding emerging markets.

**Table 1:** Gaps in the Existing Literature

| Authors | Year | Study Focus | Concept/ Method | Key Findings | Identified Gaps |
|---|---|---|---|---|---|
| Srinivas S, Ramachandiran S. | 2024 | Analysis of online customer reviews to extract airline-specific insights | Unsupervised text analytics, topic modelling, sentiment analysis. | Identified 11 service quality aspects; provided airline-specific sentiment summaries. | Lacks focus on class-specific (e.g., Business vs. Economy) satisfaction levels and route-based sentiment differences. |
| Xie H, Li Y, Pu Y, Zhang C, Huang J. | 2022 | Evaluation of airline service quality using online reviews. | Text mining, sentiment analysis, multi-criteria decision-making models (TOPSIS, VIKOR, AISM). | Developed a ranking framework for airline service quality. | Does not explore the impact of value for money on customer satisfaction and loyalty. |
| Wu S, Gao Y. | 2022 | Sentiment analysis of airline passengers' tweets to measure customer satisfaction. | Machine learning classifiers, lexical analysis, time series methods. | Effective in detecting abrupt changes in passenger sentiment. | Limited examination of factors like overall rating, class, routes, and recommendations on sentiment. |

**Table 1**: Continues

| | | | | | |
|---|---|---|---|---|---|
| Hasib MM, Rahman MS, Islam MS. | 2021 | Review of sentiment analysis and topic modeling in the airline industry. | Systematic literature review, machine learning, deep learning techniques. | Summarized key tasks and datasets in airline sentiment analysis. | Does not address recent trends or specific factors influencing customer satisfaction and loyalty. |
| Lucini FR, Tonetto LM, Fogliatto FS, Anzanello MJ. | 2020 | Exploration of airline customer satisfaction dimensions through online reviews. | Text mining, latent semantic analysis. | Identified key dimensions influencing customer satisfaction | Lacks analysis of class-specific satisfaction and the role of value for money. |
| Kumar S, Zymbler M | 2019 | Analyzing customer satisfaction from airline tweets using ML techniques. | Text analysis using Random Forest, machine learning models. | Sentiment analysis showed correlation between satisfaction and ratings. | Lacks emphasis on real-time analysis for route-based factors. |
| Namukasa J | 2013 | Influence of airline service quality on satisfaction and loyalty. | Survey-based analysis with statistical evaluation. | Fair value perceptions foster satisfaction and repeat patronage. | Limited research on modern airline contexts and data sources. |
| Yas H, et al. | 2021 | Effects of service quality and satisfaction on passenger loyalty. | Systematic review with empirical validation. | Loyalty influenced by fare affordability, onboard amenities | Lacks focus on cross-class and multi-route sentiment variations. |

In addition, the airline industry is understudied together with the analysis they have incorporated into their operations as another gap. Prior researchers have discussed sentiment analysis with a context to technology and consumer goods sectors. However, the findings of this issue suggest that the fields related to fashion and trend analysis must be more researched because of their exceptional focus on appearances, cultural trends, and self-identifications (Palalic et al., 2021).

However, there should be more research done that compares frequency in terms of positive or negative sentiment across several social media platforms. Even though numerous studies focus on specific platforms, there are very few that discuss changes in brand perception across those channels. Thus, interacting with a brand, which is highly visible on many platforms, customer emotions and interactions can be different depending on the environment of the platform (Kauffmann et al., 2019). Therefore, the problems having regard to the organisational properties of SA tools such as the inability to address sarcasm, irony, and culture can be considered as another research agenda. Further studies are required to refine these tools, especially in non-western environments, to help brands get reliable information that captures the sentiment of consumers appropriately (Wankhade et al., 2022).

### 2.7 Positive Sentiments in Customer Reviews and Their Correlation with Customer Satisfaction and Brand Loyalty

Sentiment analysis has gained recognition in measuring customer satisfaction and brand loyalty across industries, including the airline sector. Positive sentiments expressed in customer reviews are reliable indicators of higher satisfaction and loyalty levels. With the proliferation of online platforms, understanding the sentiments behind customer opinions has become crucial for organizations to make informed decisions. By analyzing sentiments expressed in customer reviews, businesses can enhance customer satisfaction and loyalty (Tan KL et al. 2023).

Furthermore, sentiment analysis has been applied to measure service encounter satisfaction with customer service chatbots, indicating its broader applicability in understanding customer interactions. (Feine J et al 2019).

Similarly, Islam MT el al. (2021) compares various machine learning algorithms combined with NLP techniques to perform sentiment analysis on U.S. airline Twitter data. The study highlights the effectiveness of sentiment analysis in understanding customer opinions, which can be correlated with satisfaction and loyalty metrics.

However, while existing studies establish a general relationship, gaps persist in analyzing quantifiable links between positive sentiment scores and customer loyalty metrics specific to airline customer experiences. Identifying this gap by empirically examining the relationship between positive customer sentiments, satisfaction levels, and loyalty indicators such as repeat recommendations and brand preference.

*H1: Positive Sentiments in customer reviews is significantly correlated with higher levels of customer satisfaction and brand loyalty.*

## 2.8 Comparative Satisfaction Levels Between Business Class and Economy Class Passengers

Passenger satisfaction is significantly influenced by service differentiation across travel classes, with Business Class passengers often reporting higher satisfaction due to superior amenities and personalized services. Research by Lucini et al. (2020) highlights that Business Class customers prioritize seat comfort, attentive staff, and in-flight amenities as critical drivers of satisfaction. Similarly, Xie et al. (2022) found that features such as spacious seating, high-quality meals, and exclusive lounges play a significant role in shaping positive sentiments among Business Class passengers.

Despite such findings, limited research compares satisfaction levels between Business and Economy Class using real-time sentiment analysis. Studies show that Economy Class passengers frequently highlight dissatisfaction linked to seat discomfort, delays, and lack of amenities. This study bridges the gap by conducting a comparative analysis of sentiment differences across classes, providing actionable insights to airlines on improving Economy Class experiences while maintaining Business Class satisfaction.

*H2: Business Class pasengers exhibit higher satisfaction levels compared to Economy Class Passengers.*

## 2.9 Sentiment Differences Between Domestic and International Routes

Route type (domestic vs. international) is a critical factor influencing service expectations and satisfaction. Research by Park et al. (2023) revealed that domestic flights tend to receive higher positive sentiments due to shorter flight durations, punctuality, and consistent service standards. On the other hand, according to Jiang, Zhang Y. (2022). international routes often face challenges such as service delays, higher expectations for amenities, and inconsistent onboard experiences, leading to mixed customer perceptions. Wu and Gao (2022) further demonstrated that sentiment analysis can identify route-specific service gaps, emphasizing the need for airlines to improve offerings on long-haul international flights. Sentiment analysis offers actionable insights into how service quality differences between domestic and international routes impact passenger satisfaction, helping airlines tailor their services.

Despite these insights, there remains a lack of studies that directly compare sentiment scores for domestic and international routes. This research bridges the gap by exploring how service provisions influence passenger sentiments across route types, contributing to a more targeted approach in airline service management.

*H3: Domestic routes receive higher positive sentiment due to enhanced service provisions compared to international routes.*

## 2.10 Value for Money as a Determinant of Customer Satisfaction and Loyalty

Value for money is a critical determinant of customer satisfaction and brand loyalty in the airline industry. Research by Namukasa (2013) underscores that passengers' perception of receiving fair value for their expenditures is fundamental in fostering satisfaction and repeat patronage. Additionally, Yas et al. (2021) found that factors like fare affordability, onboard amenities, and service efficiency collectively shape value perceptions.

However, existing literature often fails to provide a multi-class, multi-route perspective on value for money in the airline sector. By analyzing how perceived value impacts customer satisfaction across travel classes (Business vs. Economy) and route types (Domestic vs. International), the research addresses this critical gap, offering airlines actionable insights to enhance perceived value and drive brand loyalty.

*H4: Value for money is a critical determinant of customer satisfaction and loyalty across all classes and routes.*

# 3. METHOD

## 3.1 Overview of the Research Methodology Adopted

The study will be based on sentiment analysis that involves natural language processing, NLP, in analysing text data from dataset (Garg et al., 2020). From data collection and analysis of publicly available posts, comments and texts will be categorised into different sentiments as positive, negative, or neutral. Python will be used with the help of some other tools in doing the data scraping, cleaning, and labelling.

Sentiment analysis will provide an objective measurement of the attitude of the customers toward the brand (Hartmann et al., 2023). Further, using machine-learning algorithms, time-series pattern extraction for trends will be done in sentiment data for detecting trends (Mohan et al., 2022). Hence, airline can upgrade its marketing strategy and responses towards the customers. It will conclude with a large, interpretive report that will also visualize the trends using charts and graphs for easier data-driven decision-making for the brand.

In this regard, sentiment analysis is conducted based on the customer evaluations from Kaggle (Obiedat et al., 2022). Using this analysis technique, the study focuses on systematically categorizing and quantifying customer sentiment referring to the reviews of various airlines. With the application of quantitative analysis techniques, this study examines how aspects such as comfort, consumer service, and travel class significantly contribute to the level of satisfaction and loyalty. Furthermore, the methodology tends to leverage NLP, i.e. natural language processing tools for processing and interpreting the review text, while ensuring that a data-driven approach is used to understand the customer perceptions (Dash, 2022). Hence, the chapter shared details regarding data sources; the research approach used the steps of data preparation, the sentiment analysis model, the data analysis techniques, and the ethical considerations for using publicly accessible Kaggle data.

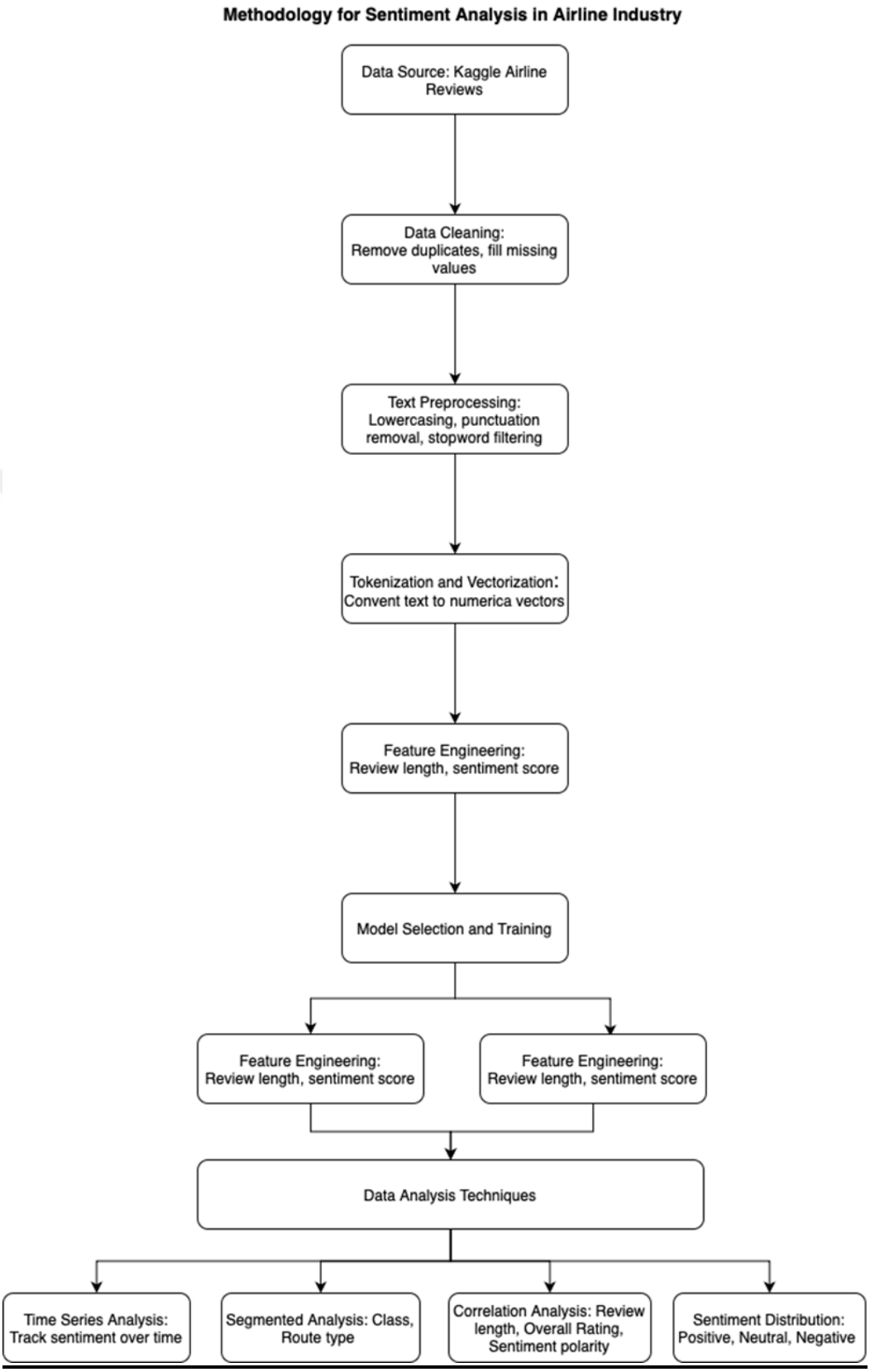



**Figure 1:** Proposed Methodology Diagram

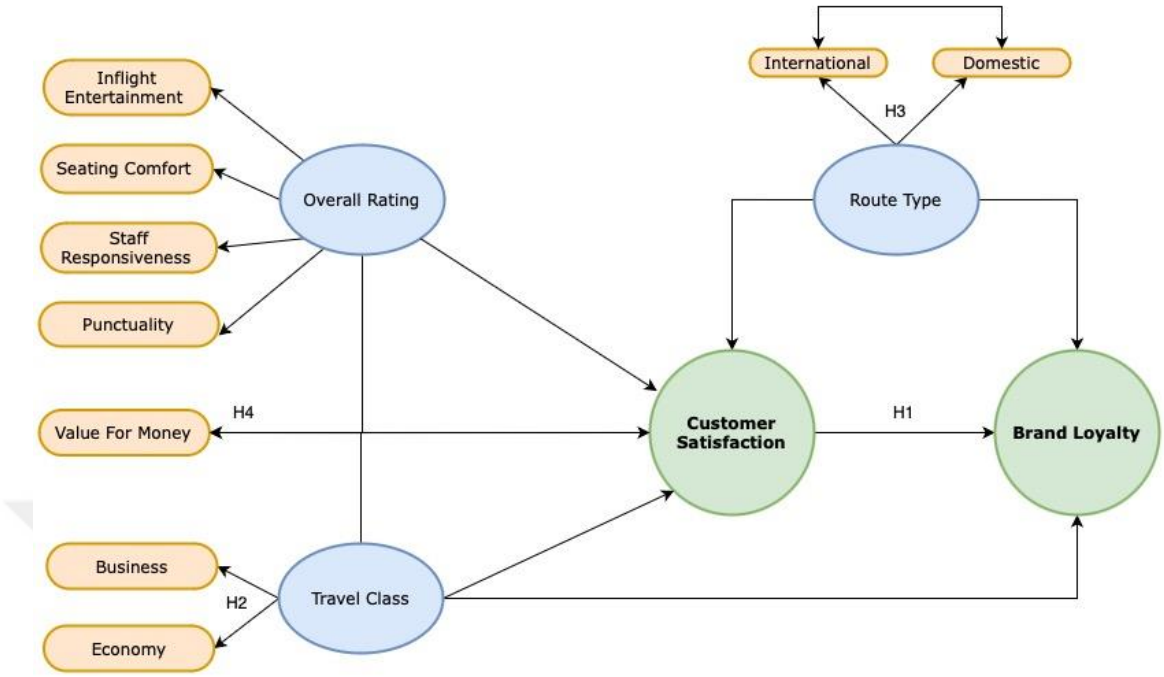



**Figure 2:** Conceptual Framework based on Hypothesis

### 3.2 Data Source and Description

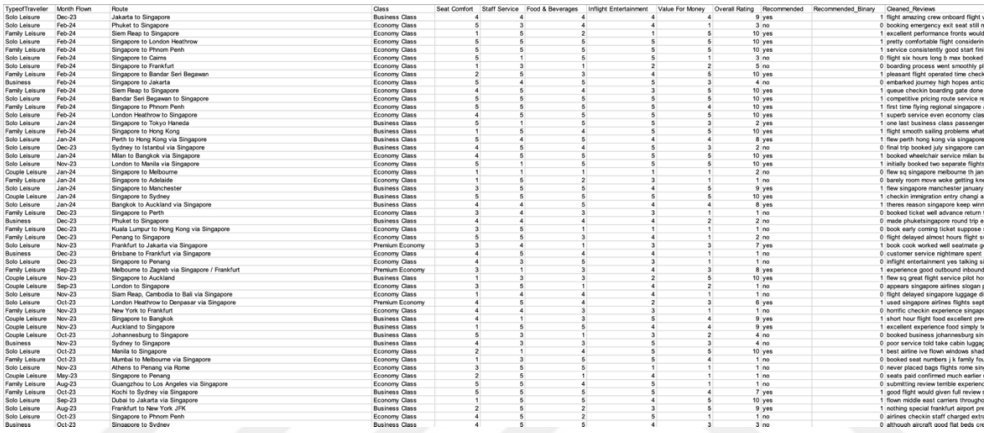
One of the primary data sources used in this study is a Kaggle dataset named "Airline Reviews" (Patel et al., 2023). It provides a comprehensive collection of consumer reviews on many airlines. Moreover, the dataset is suitable for examining customer satisfaction with loyalty indicators as it involves structured and unstructured data, both showing the consumer experiences in their own words. The core data fields involve,

- Review Text: Overall content regarding every review, in which the consumer details their opinions and experiences. Moreover, the review text is essential for sentiment analysis because it lets the researcher directly capture the emotive nuances of the customers' feelings (Bharadwaj, 2023).
- Rating: A numeric rating is linked with every review on a scale of 1 to 5 or even 10 (Higgins, 2022). The rating works as a measurable indicator regarding customer satisfaction and lets for the correlation having sentiment polarity determined from the review text.
- Sentiment Polarity: Definite sentiment labels (constructive, neutral, negative) depending upon the review content analysis. Such sentiment groups are critical to determining the general tone of consumer feedback and linking it to brand perception (EDA, 2022).
- Review Length: Overall, words in every review indicate the intensity level regarding customer feedback (Samir et al., 2023). Lengthier reviews tend to

suggest significant opinions, which can be positive or negative, and provide extensive insights into aspects affecting the level of satisfaction.

- Additional Variables: Other fields involve the travel class (either economy or business) along with the type of route (domestic or international) that allows segmentation in the analysis (Park and Storey, 2023). Such variables support the investigation of satisfaction across the customer segments and address the aim of the study to understand the class- and route-led differences in sentiment.

The combination of datasets regarding structured variables including the free-text content allows an inclusive sentiment analysis that includes many aspects of customer satisfaction and loyalty.



| TypeofTraveller | Month Flown | Route | Class | Seat Comfort | Staff Service | Food & Beverages | Inflight Entertainment | Value For Money | Overall Rating | Recommended | Recommended_Binary | Cleaned_Reviews |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Solo Leisure | Dec-23 | Jakarta to Singapore | Business Class | 4 | 4 | 4 | 4 | 4 | 9 | yes | 1 | flight amazing crew onboard flight |
| Solo Leisure | Feb-24 | Phuket to Singapore | Economy Class | 5 | 3 | 4 | 4 | 1 | 3 | no | 0 | booking emergency exit seat still n |
| Family Leisure | Feb-24 | Siem Reap to Singapore | Economy Class | 1 | 5 | 2 | 1 | 5 | 10 | yes | 1 | excellent performance fronts would |
| Solo Leisure | Feb-24 | Singapore to London Heathrow | Economy Class | 5 | 5 | 5 | 5 | 5 | 10 | yes | 1 | pretty comfortable flight considerin |
| Family Leisure | Feb-24 | Singapore to Phnom Penh | Economy Class | 5 | 5 | 5 | 5 | 5 | 10 | yes | 1 | service consistently good start flei |
| Solo Leisure | Feb-24 | Singapore to Cairns | Economy Class | 5 | 1 | 5 | 5 | 1 | 3 | no | 0 | flight six hours long b max booked |
| Solo Leisure | Feb-24 | Singapore to Frankfurt | Economy Class | 1 | 3 | 1 | 2 | 2 | 5 | no | 0 | boarding process went smoothly pl |
| Family Leisure | Feb-24 | Singapore to Bandar Seri Begawan | Economy Class | 2 | 5 | 3 | 4 | 5 | 10 | yes | 1 | pleasant flight operated time check |
| Business | Feb-24 | Singapore to Jakarta | Economy Class | 5 | 4 | 5 | 5 | 3 | 4 | no | 0 | embarked journey high hopes antic |
| Family Leisure | Feb-24 | Siem Reap to Singapore | Economy Class | 4 | 5 | 4 | 3 | 5 | 10 | yes | 1 | queue checkin boarding gate done |
| Solo Leisure | Feb-24 | Bandar Seri Begawan to Singapore | Economy Class | 5 | 5 | 5 | 5 | 5 | 10 | yes | 1 | competitive pricing route service re |
| Family Leisure | Feb-24 | Singapore to Phnom Penh | Economy Class | 5 | 5 | 5 | 5 | 4 | 10 | yes | 1 | first time flying regional singapore |
| Solo Leisure | Feb-24 | London Heathrow to Singapore | Economy Class | 4 | 5 | 5 | 5 | 5 | 10 | yes | 1 | superb service even economy clas |
| Solo Leisure | Jan-24 | Singapore to Tokyo Haneda | Business Class | 5 | 1 | 5 | 5 | 3 | 2 | yes | 1 | one last business class passenger |
| Family Leisure | Feb-24 | Singapore to Hong Kong | Business Class | 1 | 5 | 4 | 5 | 5 | 10 | yes | 1 | flight smooth sailing problems what |
| Solo Leisure | Jan-24 | Perth to Hong Kong via Singapore | Business Class | 5 | 4 | 5 | 4 | 4 | 8 | yes | 1 | flew perth hong kong via singapore |
| Solo Leisure | Dec-23 | Sydney to Istanbul via Singapore | Business Class | 4 | 5 | 4 | 5 | 3 | 2 | no | 0 | final trip booked july singapore can |
| Solo Leisure | Jan-24 | Milan to Bangkok via Singapore | Economy Class | 4 | 5 | 5 | 5 | 5 | 10 | yes | 1 | booked wheelchair service milan ba |
| Solo Leisure | Nov-23 | London to Manila via Singapore | Economy Class | 5 | 1 | 5 | 5 | 5 | 10 | yes | 1 | initially booked two seperate flights |
| Couple Leisure | Jan-24 | Singapore to Melbourne | Economy Class | 1 | 1 | 1 | 1 | 1 | 2 | no | 0 | flew sq singapore melbourne th jan |
| Family Leisure | Jan-24 | Singapore to Adelaide | Economy Class | 1 | 5 | 2 | 3 | 1 | 1 | no | 0 | barely room move woke getting kne |
| Solo Leisure | Jan-24 | Singapore to Manchester | Business Class | 3 | 5 | 5 | 4 | 5 | 9 | yes | 1 | flew singapore manchester januar |
| Couple Leisure | Jan-24 | Singapore to Sydney | Business Class | 5 | 5 | 5 | 5 | 5 | 10 | yes | 1 | checkin immigration entry changi a |
| Solo Leisure | Jan-24 | Bangkok to Auckland via Singapore | Business Class | 4 | 4 | 5 | 4 | 4 | 8 | yes | 1 | theres reason singapore keep winn |
| Family Leisure | Dec-23 | Singapore to Perth | Economy Class | 3 | 4 | 3 | 3 | 1 | 1 | no | 0 | booked ticket well advance return |
| Business | Dec-23 | Phuket to Singapore | Business Class | 4 | 4 | 4 | 4 | 2 | 2 | no | 0 | made phuketsingapore round trip e |
| Family Leisure | Dec-23 | Kuala Lumpur to Hong Kong via Singapore | Economy Class | 3 | 5 | 1 | 1 | 1 | 1 | no | 0 | book early coming ticket suppose |
| Family Leisure | Dec-23 | Penang to Singapore | Economy Class | 5 | 5 | 3 | 4 | 1 | 2 | no | 0 | flight delayed almost hours flight w |
| Solo Leisure | Nov-23 | Frankfurt to Jakarta via Singapore | Premium Economy | 3 | 4 | 1 | 3 | 3 | 7 | yes | 1 | book cock worked well seatmate g |
| Business | Dec-23 | Brisbane to Frankfurt via Singapore | Economy Class | 4 | 5 | 4 | 4 | 1 | 1 | no | 0 | customer service nightmare spent |
| Solo Leisure | Dec-23 | Singapore to Penang | Economy Class | 4 | 3 | 5 | 3 | 1 | 1 | no | 0 | inflight entertainment yes talking si |
| Family Leisure | Sep-23 | Melbourne to Zagreb via Singapore / Frankfurt | Premium Economy | 3 | 1 | 3 | 4 | 3 | 8 | yes | 1 | experience good outbound inbound |
| Couple Leisure | Nov-23 | Singapore to Auckland | Business Class | 1 | 3 | 3 | 2 | 5 | 10 | yes | 1 | flew sq great flight service pilot ho |
| Couple Leisure | Sep-23 | London to Singapore | Economy Class | 3 | 5 | 1 | 4 | 2 | 1 | no | 0 | appears singapore airlines slogan |
| Solo Leisure | Nov-23 | Siam Reap, Cambodia to Bali via Singapore | Economy Class | 1 | 4 | 4 | 4 | 1 | 1 | no | 0 | flight delayed singapore luggage di |
| Solo Leisure | Oct-23 | London Heathrow to Denpasar via Singapore | Premium Economy | 4 | 5 | 4 | 2 | 3 | 9 | yes | 1 | used singapore airlines flights sept |
| Family Leisure | Nov-23 | New York to Frankfurt | Economy Class | 4 | 4 | 3 | 3 | 1 | 1 | no | 0 | horrific checkin experience singap |
| Couple Leisure | Nov-23 | Singapore to Bangkok | Business Class | 4 | 1 | 3 | 5 | 4 | 9 | yes | 1 | short hour flight food excellent pre |
| Couple Leisure | Nov-23 | Auckland to Singapore | Business Class | 1 | 5 | 5 | 4 | 4 | 9 | yes | 1 | excellent experience food simply lu |
| Couple Leisure | Oct-23 | Johannesburg to Singapore | Business Class | 5 | 3 | 1 | 3 | 2 | 4 | no | 0 | booked business johannesburg sin |
| Business | Nov-23 | Sydney to Singapore | Business Class | 4 | 3 | 3 | 5 | 3 | 4 | no | 0 | poor service told take cabin luggag |
| Solo Leisure | Oct-23 | Manila to Singapore | Economy Class | 2 | 1 | 4 | 5 | 5 | 10 | yes | 1 | best airline ive flown windows shad |
| Family Leisure | Oct-23 | Mumbai to Melbourne via Singapore | Economy Class | 1 | 3 | 5 | 5 | 4 | 1 | no | 0 | booked seat numbers j k family flo |
| Solo Leisure | Nov-23 | Athens to Penang via Rome | Economy Class | 3 | 5 | 5 | 1 | 1 | 1 | no | 0 | never placed bags flights rome sin |
| Couple Leisure | May-23 | Singapore to Penang | Economy Class | 2 | 5 | 1 | 4 | 1 | 1 | no | 0 | seats paid confirmed much earlier |
| Family Leisure | Aug-23 | Guangzhou to Los Angeles via Singapore | Economy Class | 5 | 5 | 4 | 5 | 1 | 1 | no | 0 | submitting review terrible experienc |
| Family Leisure | Oct-23 | Kochi to Sydney via Singapore | Business Class | 5 | 5 | 5 | 5 | 4 | 7 | yes | 1 | good flight would given full review |
| Solo Leisure | Sep-23 | Dubai to Jakarta via Singapore | Economy Class | 1 | 5 | 5 | 4 | 5 | 10 | yes | 1 | flown middle east carriers throughc |
| Solo Leisure | Aug-23 | Frankfurt to New York JFK | Business Class | 2 | 5 | 2 | 3 | 5 | 9 | yes | 1 | nothing special frankfurt airport pre |
| Solo Leisure | Oct-23 | Singapore to Phnom Penh | Economy Class | 4 | 5 | 2 | 5 | 1 | 1 | no | 0 | airlines checkin staff charged extr |
| Business | Oct-23 | Singapore to Sydney | Business Class | 4 | 5 | 5 | 4 | 3 | 3 | no | 0 | although aircraft good flat beds cre |

**Figure 3:** Airline Dataset

### 3.3 Data Preparation and Preprocessing

It is essential to have an effective data preparation to extract dependable sentiment insights from review text. Moreover, the data preparation with preprocessing involved the given steps:

- Data Cleaning: It involved the removal of duplicates and unrelated entries, as well as reviews with the missing values (Wang et al., 2020). Making sure that data quality is up to the mark, the research reduces noise and improves the level of accuracy regarding the sentiment analysis. Furthermore, reviews with imperfect information, like missing ratings or text, must be included to maintain constancy.

- Text Preprocessing: Techniques of NLP were used to clean the text data. Moreover, the review text was transformed to lowercase, whereas the inappropriate characters, like punctuation, were deleted. Stopwords, which are common words with little sentiment value (for instance, "and" or "the"), were even filtered to focus on words that have sentiment (Chai, 2020).

- Tokenisation and Vectorisation: Cleaned text had been tokenized—broken into specific words or even tokens. They were then vectorised to convert the text into numeric data. TF-IDF (Term Frequency-Inverse Document Frequency) weighed the terms and considered their importance in and across reviews (Al-Obaydy et

al., 2022). These helps identify significant words in sentiment classification. Such transformation regarding text data into statistical vectors lets machine learning algorithms for the interpretation and analysis of the text systematically.

- Feature Engineering: Further features were made to enhance the sentiment analysis. This involved reviewing the length and scores of sentiments (positive, neutral, negative) with the rating categories (Al-Natour and Turetken, 2020). Moreover, review length was an indicator of consumer engagement, having lengthier reviews that even reflect intense opinions. Such features led the study to analyze how the review length and other factors link with satisfaction and sentiment.

The blend of such preprocessing steps makes the data for vigorous sentiment analysis and enables precise categorization regarding the customer feedback depending upon the emotional tone.

### 3.4 Sentiment Analysis Model

For the classification of customer sentiment in reviews, the research used a machine learning-led sentiment analysis model (Akritidis and Bozanis, 2023). Moreover, considering the textual content, the model classifies all the reviews into positive, neutral, or negative classes of sentiment. A sentiment analysis model was selected after preliminary testing of various machine learning algorithms, including Logistic Regression and Naive Bayes, with Logistic Regression chosen for its high accuracy, precision, and recall scores (Itoo et al., 2021). The dataset was split into training and testing sets, ensuring the model could generalize well to new data and maintain consistent results (Uçar et al., 2020). Evaluation metrics such as accuracy, precision, recall, and F1-score assessed the model's performance, confirming its ability to classify sentiment effectively (Yacouby and Axman, 2020). This model systematically analyzes consumer feedback, offering valuable insights into customer satisfaction trends across airlines.

### 3.5 Data Analysis Techniques

The research uses many quantitative data analysis techniques for the exploration of patterns regarding customer satisfaction as well as brand loyalty:

- Sentiment Distribution Analysis: The technique assesses the quantity of positive, neutral, and negative sentiments in different airlines (Obiedat et al., 2022). With the analysis of the sentiment distribution, the research identifies significant satisfaction trends and reveals how airlines tend to compare referring to customer sentiment, satisfaction and loyalty.

- Correlation Analysis: To explore the link between variables, it is essential to use correlation analysis, which evaluates the strength and direction of relations between review length, ratings, and sentiment polarity (Birjali et al., 2021). For instance, a positive correlation between ratings and sentiment polarity will signify that greater ratings align with the positive sentiment. The analysis shows core drivers of satisfaction factors and their effect on loyalty indicators, such as recommendations.

- Segmented Sentiment Analysis: Sentiment trends are assessed through the class (economy vs. business) along with the route type (domestic vs. international) (Birjali et al., 2021). This helps in the examination of differences regarding satisfaction across the customer segments. Moreover, the segmented approach gives insights into the class- and route-based variations in the sentiment and supports the study's objective of understanding how such factors impact satisfaction and loyalty.

  *H2: Business Class passengers exhibit higher satisfaction levels compared to Economy Class Passengers.*

  *H3: Domestic routes receive higher positive sentiment due to enhanced service provisions compared to international routes.*

- Time-Series Analysis: Sentiment trends are assessed with time to detect the changes regarding customer satisfaction with events such as promotional campaigns or service changes (Wankhade et al., 2022). Moreover, the time-series analysis lets the researcher note seasonal or event-driven transformations in the sentiment, which provides insights regarding particular factors affecting satisfaction with time.

- The analysis of Value for Money as a determinant of customer satisfaction and loyalty was conducted using a sentiment analysis approach, leveraging customer reviews to quantify perceptions. Machine learning models, such as Logistic Regression, were employed to identify patterns in the dataset and classify sentiments into positive, neutral, or negative categories. The model evaluated key indicators related to perceived value for money, including ticket pricing, service quality, and benefits offered across different travel classes and routes. This data-driven approach provided a robust understanding of how customers value airline services and their willingness to recommend or remain loyal to a specific brand. Such insights are essential for identifying areas of improvement and optimizing pricing strategies to enhance customer satisfaction.

  *H4: Value for money is a critical determinant of customer satisfaction and loyalty across all classes and routes.*

Such analysis techniques let the study quantify the factor of satisfaction with the loyalty trends systematically and offer actionable insights to improve the customer experience along with the brand perception.

### 3.6 Ethical Considerations

The study follows some ethical standards regarding data handling while using publicly accessible data only from Kaggle (Mohammad, 2022). The ethical considerations in this study include privacy and anonymisation. Much heed is given to compliance considering the platform guidelines. Apart from this, the study included transparent and reliable reporting. The study's ethical approach ensures that the results are accurately reported. Furthermore, integrity is maintained in sentiment analysis.

# 4. FINDINGS

## 4.1 Data Preparation and Preprocessing Results

Data preparation and pre-processing are essential steps to make clean, trusted data for sentiment analysis. Missing values, date formatting, deduplication, and finalising critical variables, including Overall Rating, Recommended status and Review Length, were achieved as essential steps. The impact of missing data on analysis accuracy was determined, and missing values were either filled out or excluded based on the treatment's effect on analysis accuracy. To keep things consistent across time-based analyses, we standardised the date formatting so we could identify trends and seasonality in customer satisfaction. To prevent the results from being skewed, we removed duplicate entries. The variables Overall Rating, Recommended and Review Length were finalised to analyse levels of customer satisfaction and likelihood of customer recommendation.

Natural language processing (NLP) techniques were used for text preprocessing in sentiment analysis such that the review text was prepared for analysis. The words were then tokenised, and individual or so-called tokens were each reviewed to test word-level patterns. To get rid of common, non-informative words (e.g., "and" "the", "is") that do not add to the sentiment, we have implemented stop word removal. Additionally, punctuation and unnecessary symbols were eliminated to reduce review text ambiguity and lack of consistency. We standardised the text to better standardise customer expressions without interfering with noise or harmful data, while the sentiment model could still capture the sentiment. They turn out to be extremely important for the accuracy of the sentiment analysis because a good model can know its path and avoid words such as "comfortable," "poor," and "friendly," which correspond to emotional words in customer reviews. We restricted some variability by establishing a featured format to the review text where the model was better able to give consistently strong identification of positive and negative sentiment patterns. It has become such a preprocessing foundation that the sentiment analysis results come from genuine customer opinions and insights on brand loyalty and pleasure.

```
                                                  Reviews  \
0    Flight was amazing. The crew onboard this fl...
1    Booking an emergency exit seat still meant h...
2    Excellent performance on all fronts. I would...
3   Pretty comfortable flight considering I was f...
4   The service was consistently good from start ...

                                        Cleaned_Reviews
0  flight amazing crew onboard flight welcoming g...
1  booking emergency exit seat still meant huge d...
2  excellent performance fronts would definitely ...
3  pretty comfortable flight considering flying e...
4  service consistently good start finish cabin c...
```

**Figure 4:** Data Pre-processing

### 4.2 Descriptive Analysis and Data Insights

Descriptive analysis of the airline review dataset reveals the demographics of the travellers, review distribution and overall satisfaction pattern. Analysis of the timeline of reviews shows an increase in review volume after 2014, with the peak year being 2019. In this trend, customers and feedback may have an increasing fascination because they increasingly rely on digital platforms for sharing these travel experiences. Seasonal fluctuations and external events, such as the COVID-19 pandemic, impact travel demand and cause periodic drops in review volume.

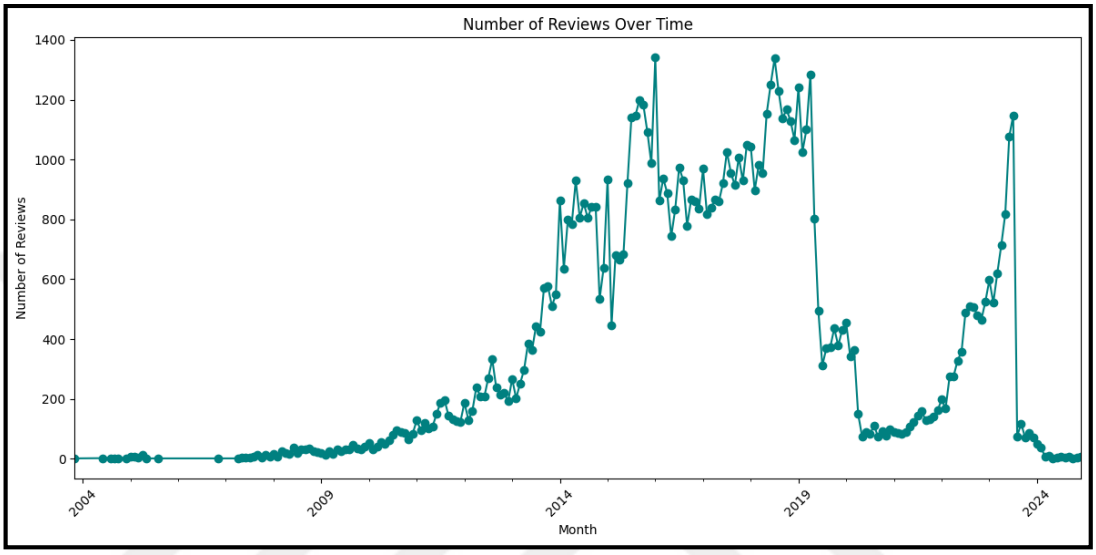



**Figure 5:** Number of Reviews over Time

Another essential variable is class distribution, which has an overwhelming proportion of Economy Class reviews, Business, Premium Economy, and First Class. This indicates that the economy class is the majority customer base and, consequently, needs to be more varied concerning service quality. Because of better service offerings, such as business and premium classes, receive less feedback but reflect more satisfaction.

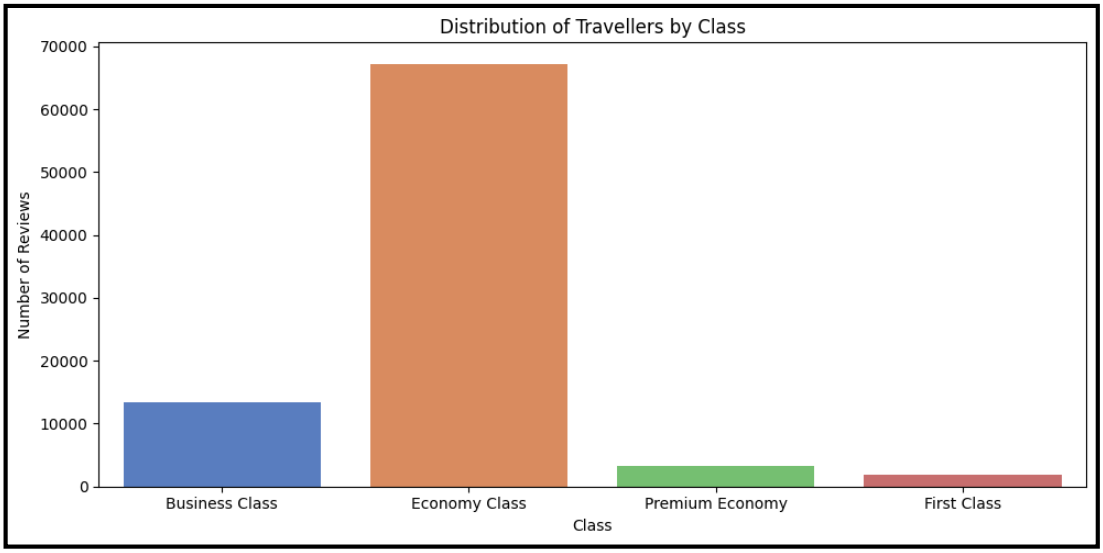



**Figure 6:** Distribution of Travellers by Class

The rating distribution has a skewed distribution with a large concentration on the bottom end of ratings around the 2-3 range and the spectrum on the top end of ratings of 8-10. This distribution points out the areas for improvement with ratings less than 5, where customers would continually be dissatisfied with delays, comfort, or customer service. The recommendation rates are presented as a pie chart and show that 54.9% of customers recommend their airline. In comparison, 45.1% do not, meaning that the airlines have room to improve customer experiences to increase the recommendation rates.

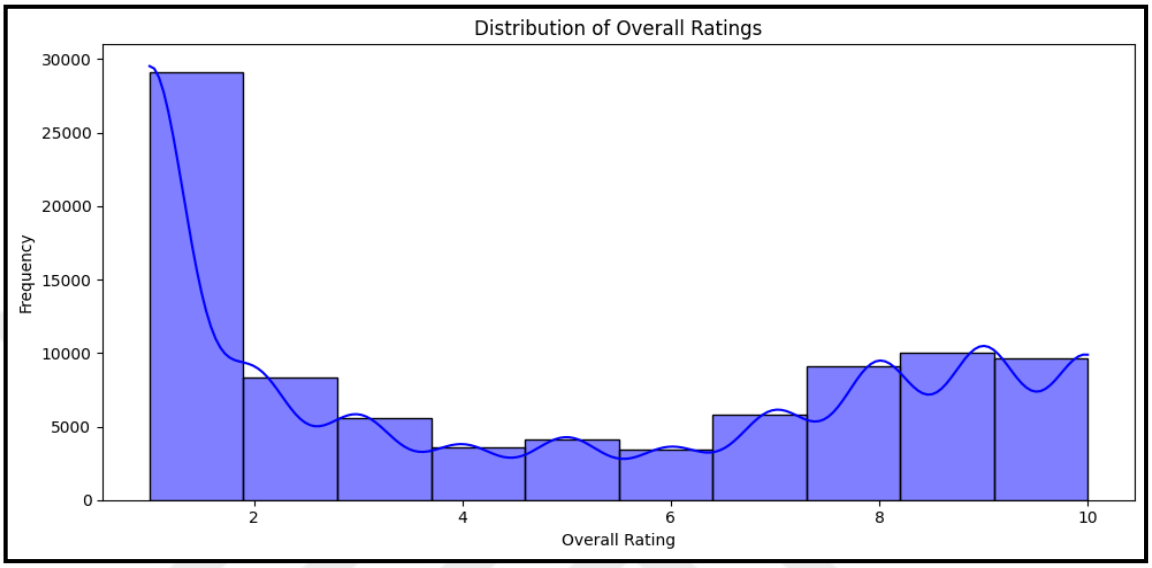



**Figure 7:** Distribution of Overall Ratings

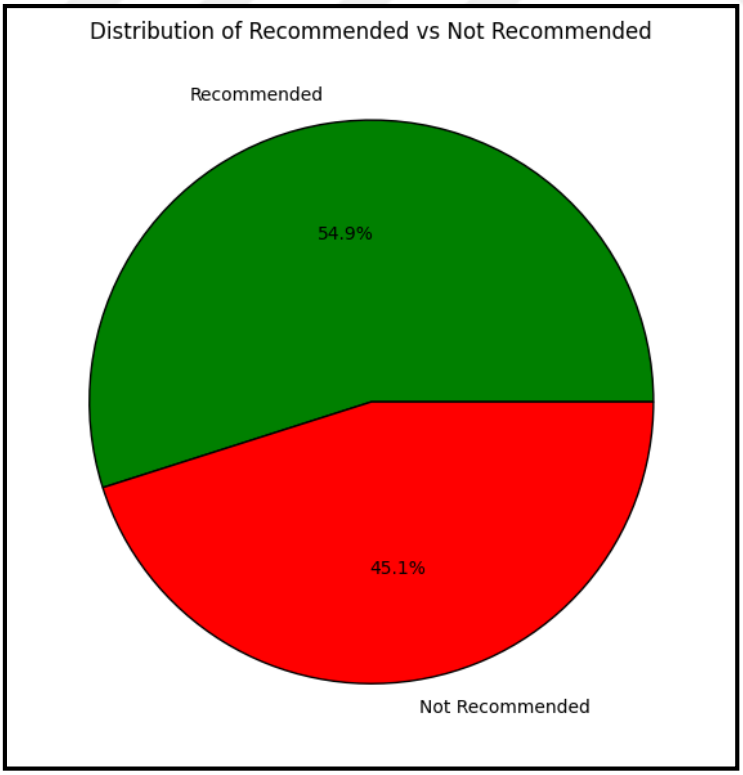



**Figure 8:** Distribution of Recommended and Non-Recommended

A correlation matrix of Rating categories (Overall Rating, Seat Comfort, Staff Service, Food & Beverages, Inflight Entertainment, and Value for Money) indicates significant relationships. The t-value represents the strength of the relationship between the predictor and the dependent variable. In the Table 2, all the t-values are very high (e.g., Value for Money = 93.79, Staff Service = 25.12, Inflight Entertainment = 20.83), which indicates that these predictors have a strong positive influence on overall customer satisfaction. Higher t-values suggest that the regression coefficients (estimates) are significantly different from zero, meaning the predictors are reliable indicators of customer satisfaction. The p-values for all predictors are < 0.001, which indicates that the results are highly statistically significant. Therefore, predictors like Value for Money, Staff Service, and Inflight Entertainment significantly contribute to overall satisfaction levels. Additionally, the estimate values indicate the magnitude and direction of the relationship. For instance, Value for Money(Estimate = 0.7689) has the highest positive effect, meaning an increase in perceived value significantly boosts customer satisfaction.

**Table 2:** Regression Model

Model Coefficients - Overall Rating

| Predictor | Estimate | SE | t | p |
|---|---|---|---|---|
| Intercept [a] | −0.1320 | 0.01998 | −6.61 | < .001 |
| Value For Money | 0.7689 | 0.00820 | 93.79 | < .001 |
| Inflight Entertainment | 0.1233 | 0.00592 | 20.83 | < .001 |
| Food & Beverages | 0.0643 | 0.00685 | 9.39 | < .001 |
| Seat Comfort | 0.1056 | 0.00689 | 15.33 | < .001 |
| Staff Service | 0.1634 | 0.00651 | 25.12 | < .001 |
| Recommended_Binary: | | | | |
| 1 – 0 | 3.0147 | 0.02462 | 122.46 | < .001 |

[a] Represents reference level

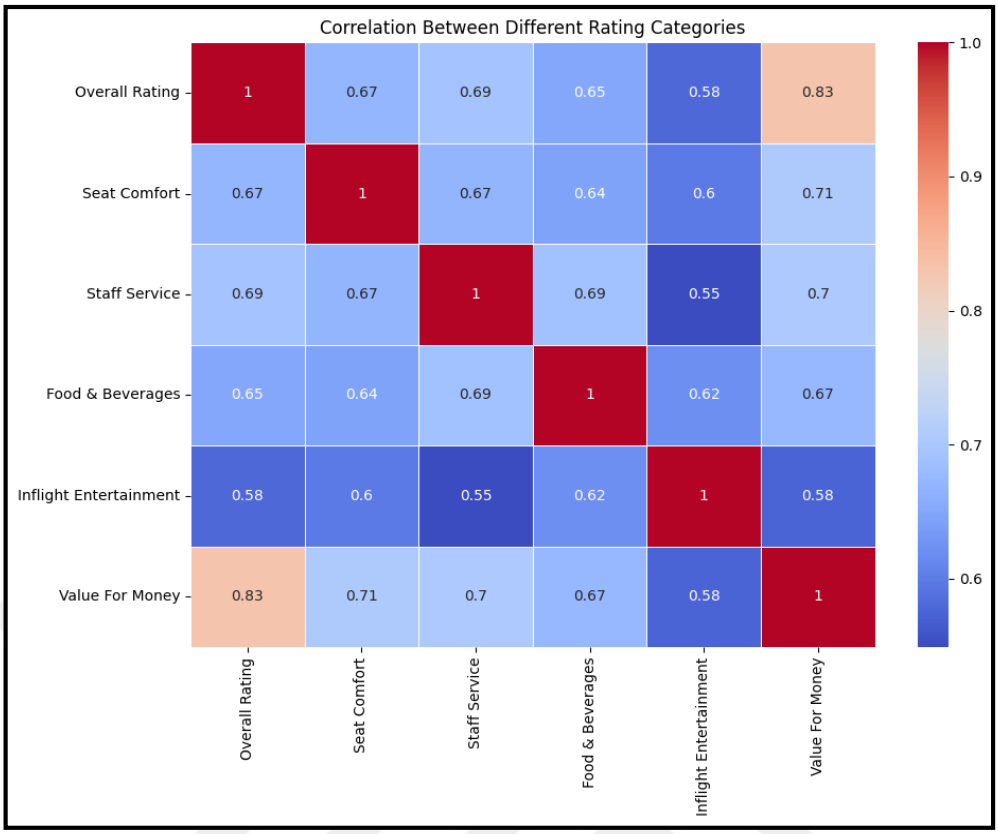



**Figure 9:** Correlation Matrix

A word cloud of standard review terms shows forms of words that appear frequently, such as 'seat,' 'food,' 'plane,' and 'staff.' Seating comfort, staff interactions and timeliness are critical to customers' sentiment, according to favourable terms such as "friendly" and "good" juxtaposed with negative ones such as "delayed" and "poor." From a holistic view of the review patterns, airlines can view their customer satisfaction and give actionable insights for when to enhance their service and improve customer loyalty.

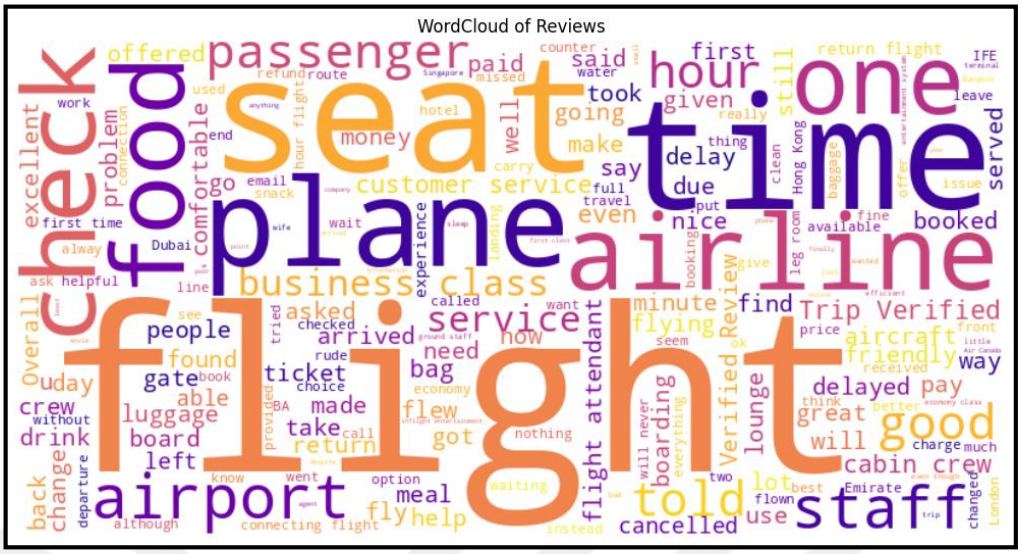



**Figure 10:** Word Cloud of Reviews

### 4.3. Sentiment Analysis Results

#### **4.3.1** Model Performance Evaluation

Two models, Logistic Regression and Multinomial Naive Bayes, were trained and tested using metrics like Accuracy, Precision, Recall, and F1-score to evaluate the quality of sentiment classification.

1. Logistic Regression: The tuned Logistic Regression model was accurate to 91%. For both classes (0 = Not Recommended; 1 = Recommended), the precision, recall, and F1 scores were about 0.90, nearly 0.90, implying the goodness of fit in predicting customer recommendations. A confusion matrix of Logistic Regression showed lower misclassification rates here than the confusion matrix of Logistic Regression from the previous example, with 764 false positives (those incorrectly labelled as Recommended) and 795 false negatives (incorrectly labelled as Not Recommended). The outcome of this strong performance shows that Logistic Regression was able to identify patterns in customer sentiment.

```
Logistic Regression Results:
Accuracy: 0.91
Precision: 0.90
Recall: 0.90
F1 Score: 0.90

Classification Report:
              precision    recall  f1-score   support

           0       0.92      0.92      0.92      9682
           1       0.90      0.90      0.90      8048

    accuracy                           0.91     17730
   macro avg       0.91      0.91      0.91     17730
weighted avg       0.91      0.91      0.91     17730
```

**Figure 11:** Logistic Regression Model Result

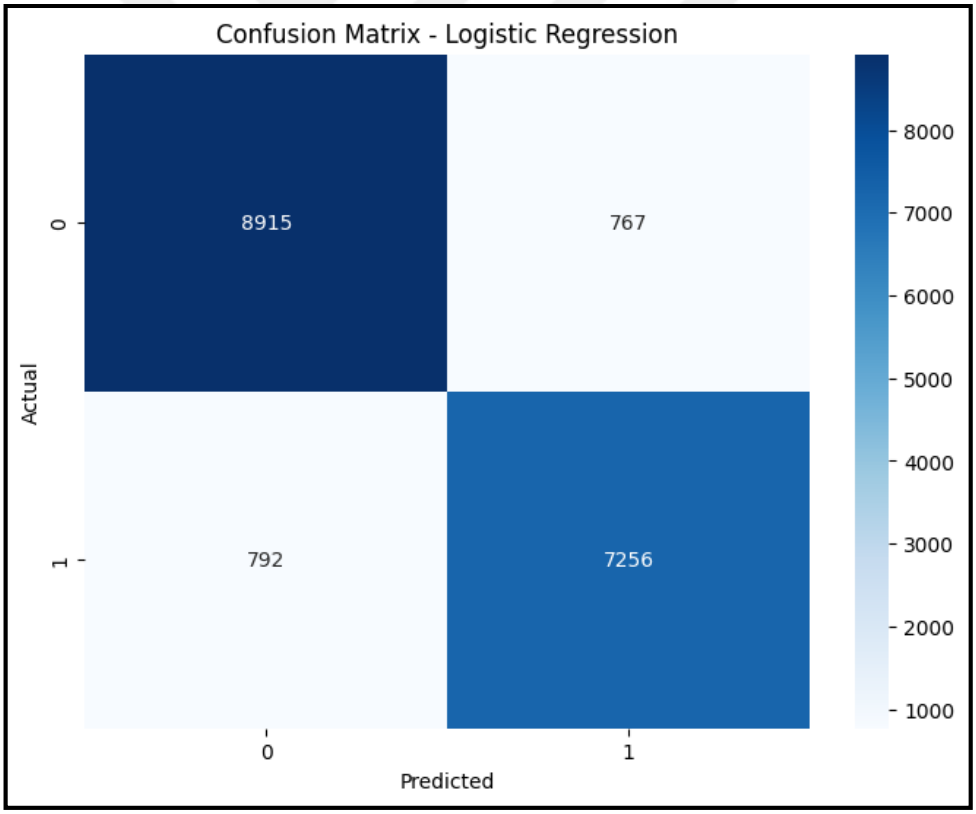



**Figure 12:** Logistic Regression Confusion Matrix

2. Multinomial Naive Bayes: The Naive Bayes model reached a slightly lower accuracy of 86%. For the Not Recommended class, its precision was 0.89, and for the Recommended class, 0.83. The recall scores were 0.85 and 0.88, respectively,

meaning the F1 score of the Recommended class is 0.85. Naive Bayes was confused with a higher misclassification rate than Logistic Regression, with 1,474 false positives and 970 false negatives. Finally, these results show that Naive Bayes performed relatively well but was less effective than Logistic Regression for this dataset.

```
Multinomial Naive Bayes Results:
Accuracy: 0.86
Precision: 0.83
Recall: 0.88
F1 Score: 0.85

Classification Report:
              precision    recall  f1-score   support

           0       0.89      0.85      0.87      9682
           1       0.83      0.88      0.85      8048

    accuracy                           0.86     17730
   macro avg       0.86      0.86      0.86     17730
weighted avg       0.86      0.86      0.86     17730
```

**Figure 13:** Multinomial Naive Bayes Model Results

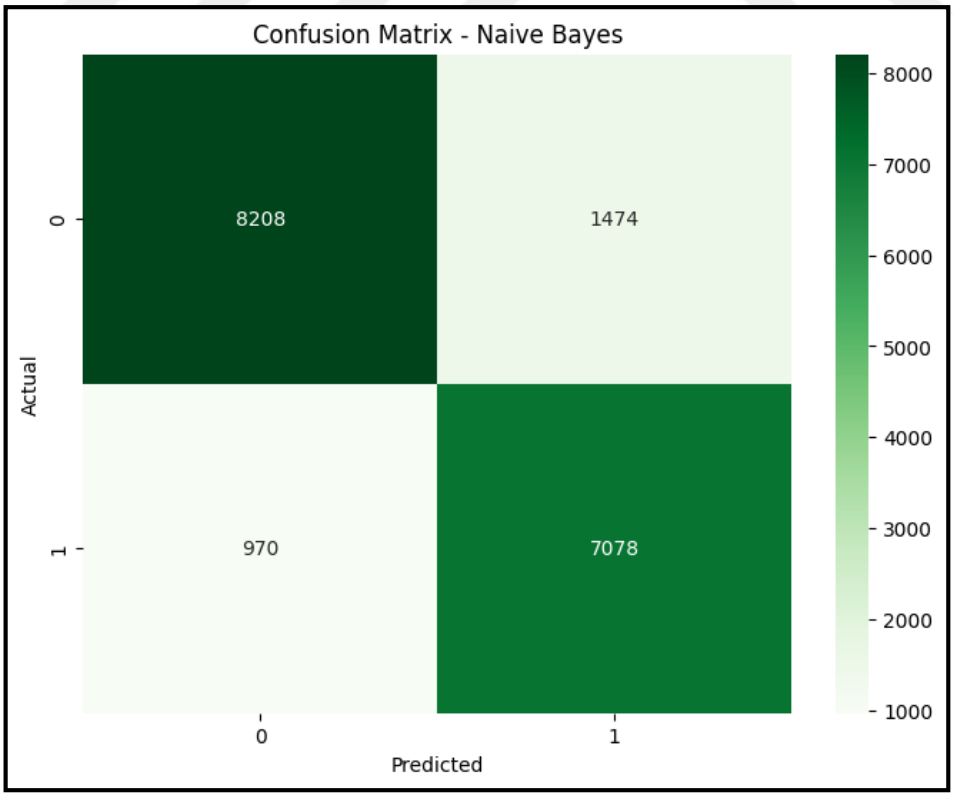



**Figure 14**: Confusion Matrix of Multinomial Naive Bayes

### 4.3.2 Sentiment Distribution Across Airlines

Customer satisfaction was further divided into groups by airline with apparent differences. The analysis sorted through reviews into positive, neutral, and negative sentiments, providing cues for how each airline is rated regarding customer views. Positive sentiment was highest for some airlines that were known for good customer service, comfortable seating, and for a price, too. These airlines responded to customer expectations and received positive feedback and higher recommendation rates. However, most airlines with large negative sentiments saw their consumers complaining of delayed flights, poor customer care, and low food quality. The lower recommendation rates suggest which service areas need improving in these airlines. Airlines that focus on providing premium experiences generally have better positive sentiment. In contrast, budget airlines – often severely constrained by operating limitations and cost-cutting measures – tend to have more negative sentiment. This compares the effect of business model and service level on customer satisfaction. Budget airline customers may be disappointed if their comfort and quality expectations are unmet, while premium airline customers appreciate the extra value of comfort and quality.

### 4.3.3 Hypothesis Results:

H1: Positive sentiment in customer reviews and satisfaction/loyalty

The results show a strong positive correlation and significant regression between customer reviews and customer satisfaction/brand loyalty.

Correlation Coefficient (r): 0.83 (Strong Positive Correlation)

Regression Results:

R-squared: 0.69 (69% of variation explained by Overall Rating)

Significant positive coefficient: $\beta = 0.119$ ($p < 0.001$)

**Table 3:** Strong positive correlation and significant regression result

| Measure | Value |
|---|---|
| Correlation (r) | 0.83 |
| R-squared | 0.69 |
| Coefficient (β) | 0.119 |
| p-value | <0.001 |

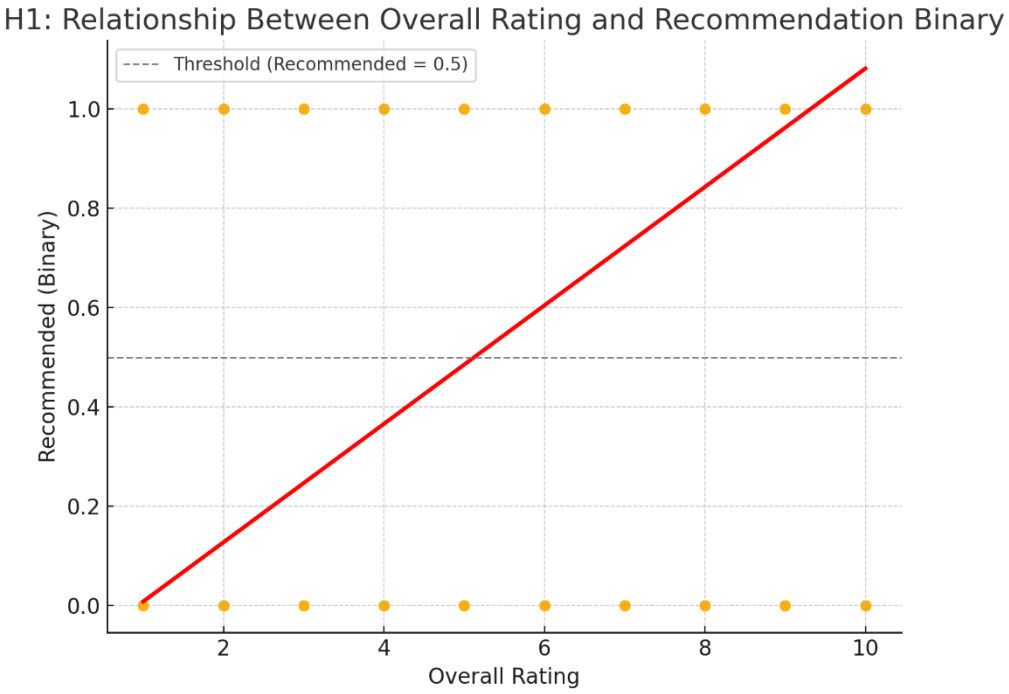



**Figure 15:** Relationship b/w Overall Rating & Recommendation Binary

### 4.3.3 Sentiment by Flight Class and Route

Sentiment analysis results were also segmented by flight class (economy vs business) and route type (domestic vs international) to show differing satisfaction levels across categories.

H2: Business-class vs. Economy-class satisfaction levels

The result shows a significant difference in the satisfaction levels of classes.

T-test results: Significant difference ($t = 60.48$, $p < 0.001$)

**Table 4:** Significant difference in satisfaction levels

| Class | Mean Overall Rating | T-test (t) | p-value |
|---|---|---|---|
| Business Class | 6.29 | 60.48 | <0.001 |
| Economy Class | 4.35 | 60.48 | <0.001 |

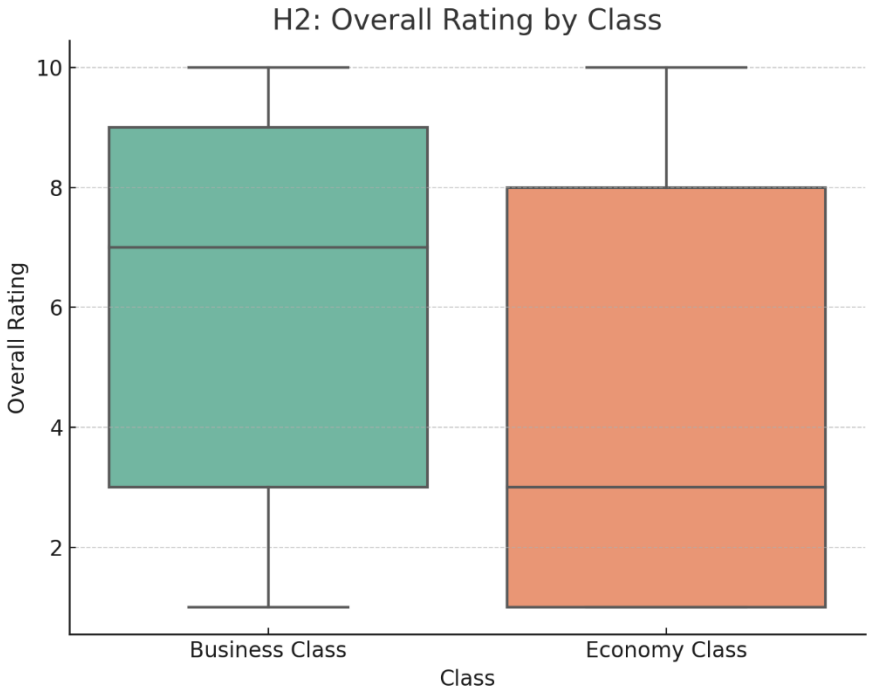



**Figure 16:** Box Plot - Overall Rating by Class

Sentiment by Flight Class:

- Economy Class: There was a more significant proportion of negative sentiment among economy class passengers. These ranged from discomfort, cramped seating, and inadequate service, pointing to what could be improved for economy passengers. Because economy travellers want value, failure to meet basic amenities and service expectations can cost sentiment. Increasing comfort and service quality in economy seats might help avoid negative feedback.

- Business Class: On the other hand, Business Class travellers were likelier to complain. Customers interviewed as business travellers emphasised the benefits associated with spacious seating, attentive service and good quality food, indicating that when they pay for the comfort of Business class, they get what they expect. This sentiment trend indicates that giving attention to high-quality business transport services for business travellers is very positive for clients' contentment and loyalty; they favour the additional relaxation and services.

Sentiment by Route Type:

- Domestic Routes: domestic flights yielded distributions of the sentiments to be more balanced, with a mixture of upbeat, near neutral, and negative reviews. Punctuality and efficient service were generally the reasons for positive feedback on both domestic routes, and this feedback was commonly negative to delays and inconsistent service quality. Airlines need to pay closer attention to short-haul

flights and can do it by addressing issues of convenience and timeliness to the delight of domestic travellers.

- International Routes: The sentiment around those same flights was more polarised, pointing to comfort, food quantity, and flight entertainment. In some cases, there was a positive sentiment on international routes, where amenities such as comfortable seating and high-quality meals were most influential on long journeys. A running theme in negative feedback was how uncomfortable economy seats were, the limited offerings (allegedly) available as meals, and the generally poor quality of in-flight entertainment, which was a call for improvement in long-haul travel.

H3: Domestic routes receive higher positive sentiment due to enhanced service provisions compared to international routes.

There is a statistically significant difference in the overall ratings between Domestic and International routes ($F=1083$, $p<0.001$).

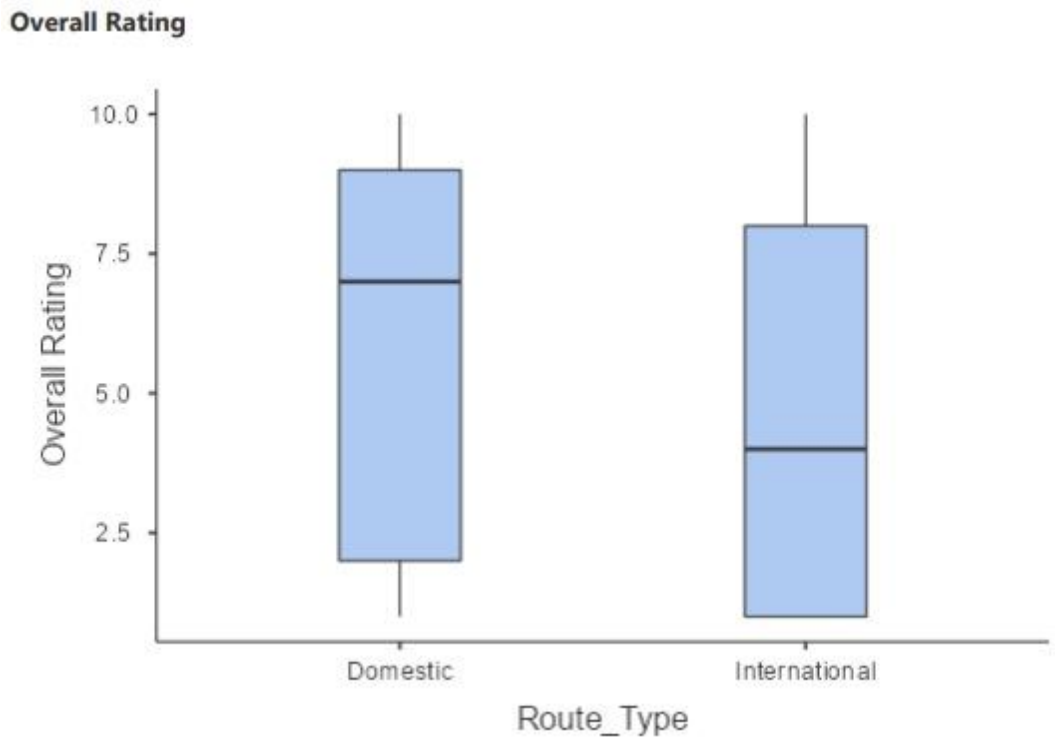



**Figure 17:** Overall Rating by Route Type

The mean overall rating for Domestic routes is higher (5.83) compared to International routes (4.81). The median for Domestic routes (7.00) is also higher than that for International routes (4.00).

**Table 5:** Domestic vs International Routes

Group Descriptives

| | Group | N | Mean | Median | SD | SE |
|---|---|---|---|---|---|---|
| Overall Rating | Domestic | 16694 | 5.830 | 7.00 | 3.318 | 0.02568 |
| | International | 39246 | 4.808 | 4.00 | 3.466 | 0.01750 |
| Recommended_Binary | Domestic | 16694 | 0.588 | 1.00 | 0.492 | 0.00381 |
| | International | 39246 | 0.459 | 0.00 | 0.498 | 0.00252 |

This hypothesis reflects the observed statistical difference in overall ratings. The results suggest that customers perceive domestic routes more favourably, possibly due to factors like consistent service quality, affordability, or reliability compared to international routes.

H4: Value for money is a critical determinant of customer satisfaction and loyalty across all classes and routes.

Findings: Strong positive correlation and significant regression results.

**Table 6:** Regression b/w Value For Money and Customer Satisfaction

| **Measure** | **Value** |
|---|---|
| Correlation (r) | 0.841 |
| R-squared | 0.786 |
| Coefficient (β) | 0.804 |
| p-value | <0.001 |


**Figure 18:** Relationship b/w Value for Money and Overall Rating

### 4.3.4 Insights and Implications

Logistic Regression with Naive Bayes sentiment analysis gives insights into the change in customer satisfaction trends between airlines, flight classes and route types. However, this research found key patterns in how class and route alter the passenger's sentiment. For example, economy passengers generally expect airlines to deliver more value, and international passengers demand comfort and amenities, areas which airlines could address. Results of the analysis suggest that it is easier for airlines with positive sentiment to keep customers and gather brand loyalty. On the contrary, airlines with negative sentiments have difficulties that could harm their reputation and customer retention. Class and route type differences in sentiment provide actionable insights. This could include doing better in the basics of economy class to counteract feelings of lousy service from budget travellers or excelling with the premium side of international flights to please long-haul customers.

## 4.4 Prediction Probabilities and Key Word Analysis

The prediction probabilities are 42% "Not Recommended" and 58% "Recommended", indicating a somewhat positive bent. Words such as "poor", "disappointing", "tiny", and "standard" add despair to the "Not Recommended" rating because customers don't like comfort and food quality. However, favourable terms such as sound, overall, well, and comfortable are likely to encourage recommendation – suggesting customer satisfaction with aspects of seating, for example, and service quality.

Sample reviews highlight both positive and negative feedback of language with actionable insights for airlines to qualify target areas to improve and reinforce strengths.

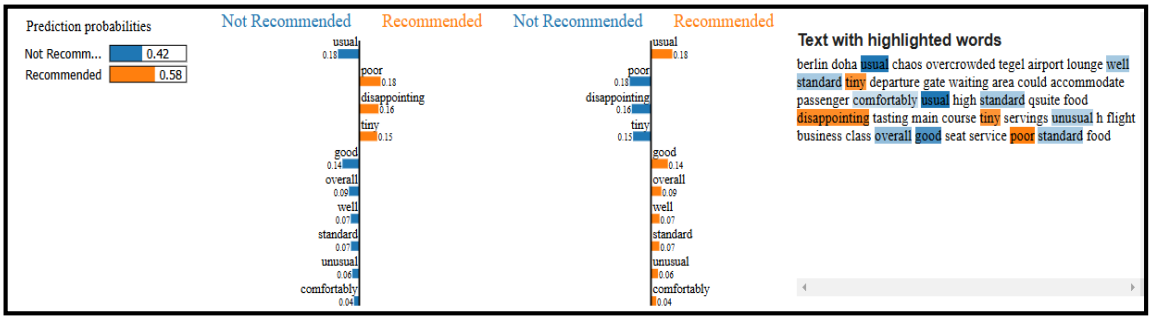



**Figure 19:** Model Interpretation

## 4.5 Sentiment Analysis Finding Discussion

This section discusses the key findings of sentiment analysis while placing them within the airline industry's customer satisfaction framework and brand loyalty. Examining customer reviews in different airlines, the research study aims to reveal how sentiment trends show travellers' level of satisfaction with airline services, notifying significant brand loyalty (Mehraliyev et al., 2020). The analysis is aligned with the goals of the study of not just identifying substantial factors that contribute to positive as well as negative sentiments but even assessing how such factors have an impact on customer perceptions based on the brand value.

Understanding the sentiment trends and their implications for consumer experience is essential as it gives valuable insight into core areas that the airlines can focus on for improvement (Shahid et al., 2024). The discussion considers how significant variations across aspects such as class and route contribute to satisfaction while comparing these insights to the existing literature regarding brand loyalty and sentiment analysis. The chapter even concludes with the implications regarding industry practices and outlines recommendations that leverage such findings for tactical customer relationship management with digital engagement.

### 4.5.1 Key Findings Interpretation

#### 4.5.1.1 Sentiment Trends and Customer Satisfaction

The sentiment analysis proved that distinctive satisfaction trends exist across airlines. Moreover, positive sentiment is usually linked with aspects like attentive consumer service, flight punctuality, and the apparent value of ticket prices (Park, 2020). Airlines with more excellent ratings usually demonstrate positive sentiments in the reviews, which suggests a direct relationship between higher satisfaction levels and favourable ratings.

Contrariwise, recurring problems that negatively affected the sentiment involved delays in flight, poor entertainment options, or discomforting seating. For instance, passengers often expressed dissatisfaction with restricted legroom, mainly on long-haul flights. The finding links with the preceding research, which shows that seat comfort is one of the critical drivers of satisfying customers in the airline industry (Farzadnia and

Vanani, 2022). The existence of higher ratings with negative feedback in specific areas, though, signifies the complexity of customer satisfaction; a favourable rating may not indicate consistently positive sentiments in the service factors.

The observations underline the need to adopt a holistic approach to evaluating customer satisfaction, incorporating quantitative ratings with qualitative feedback. In addition, the association between sentiment trends and the ratings validates that specific service factors have a greater weight in the overall satisfaction perspective of the customers (Eachempati et al., 2022). Airlines may anticipate and deal with customer expectations by considering such crucial factors.

#### 4.5.1.2 Insights on Brand Loyalty

Brand loyalty was strongly influenced by the higher ratings, positive sentiments, and recurrent recommendations in customer reviews (Ahmad and Guzmán, 2021). Positive feedback even included the customers overtly suggesting the airline or mentioning their intentions for booking their future flights. Reviews that praised the airlines for constancy in service quality, appropriate time management, and staff approachability were frequently linked with loyal customers, displaying positive sentiment and a higher level of commitment to the brand.

Furthermore, airlines with mixed reviews on particular features can maintain customer loyalty even when they constantly excel in significant areas, like customer service. For example, airlines that get positive feedback on customer service replies regardless of the negative remarks in the other regions proved as evidence regarding sustained loyalty (Rane et al., 2023). Airlines strengthen brand loyalty while paying attention to the service attributes that significantly affect travellers' experiences. However, the other factors, such as amenities, may need more convincing feedback.

#### 4.5.1.3 Role of Class and Route in Satisfaction

The analysis even discussed disparities in sentiment, considering the flight class and the route. The passengers of business class constantly articulated higher levels of satisfaction than those in economy class, which cite better seating comfort, improved food quality, and customised services (Zhou et al., 2020). Such a trend indicates that premium class amenities significantly contribute to positive customer sentiment and strengthen brand loyalty amongst business class travellers.

On the other hand, economy class travellers are critical, specifically on the longer flights where service or comfort limitations are evident. Such a pattern of feedback underlines the significance of addressing the service factors that matter much to economy passengers, like seat comfort, options of meals, and in-flight entertainment availability. The study even found positive sentiment on the international routes instead of the domestic ones, mainly because of the increased in-flight services with more aircraft used for lengthier routes. Such variations by class and route pay heed to maintaining higher standards at service levels, including routes, since all the segments contribute uniquely to the brand perception of airlines, including customer satisfaction (Lau, 2022).

## 4.6 Comparison with Existing Literature

### 4.6.1 Arrangement with Sentiment Analysis Studies

The findings match the previous research on sentiment analysis and customer satisfaction, which highlights the significance of sentiment being a tool to monitor the level of satisfaction and manage brand perception. Research has demonstrated that real-time sentiment analysis lets firms promptly heed the key concerns, preventing negative sentiments from poor brand loyalty (Park, 2020). This study confirms such insights since the sentiment model discovered nuanced interpretations regarding consumer satisfaction beyond numerical ratings here.

The use of sentiment analysis in the study even aligns with the research emphasising the value of real-time consumer insights for the airlines. In the competitive airline industry, it is essential to understand customer sentiment on a gritty level letting the brands address the level of dissatisfaction before it intensifies and reinforces brand loyalty (Farzadnia and Vanani, 2022). The research findings signify the applicability of sentiment analysis in gaining actionable insights into consumer experiences and satisfaction levels.

### 4.6.2 Comparison to Brand Perception and Loyalty Research

This study is aligned with the theory of Brand Equity, which pays heed to the significance of reliable, positive experiences to foster brand loyalty. Moreover, the relationship between higher ratings and significant sentiments, with the probability of recommendation, tends to support the theory, which underscores how airlines cultivate brand loyalty through dependable and high-quality service. In addition, the Theory of Planned Behavior (TPB) is pertinent here, regarding TPB, favourable attitudes regarding a brand influence intention regarding purchasing or recommendation (Ahmad and Guzmán, 2021). Significant sentiments recognized in the research support such behavioural intentions, predominantly when supported by the strong recommendations in the reviews. The insights attained from the analysis suggest that the factor of brand loyalty in the airline industry pivots not just on the direct experiences but on customers' belief that they will get the same positive experiences in future interactions. It strengthens the theoretical opinion that sentiment-driven perceptions significantly shape customers' intentions to remain loyal to the brand.

### 4.6.3 Significance of core variables

Variables like ratings and review length gave crucial insights concerning understanding customer satisfaction in this research. Moreover, ratings were generally correlated with positive sentiment, whereas review length conveyed complicated perceptions of particular service factors. The relevance of such variables is aligned with the present literature, which signifies that the ratings give rapid, calculable satisfaction indicators. In contrast, longer reviews offer an explicit picture of the customer experience (Zhou et al., 2020). The extended reviews even showed sentiment having a critical view, suggesting that the customers taking time to write detailed feedback might be motivated by stronger positive or negative perceptions. It is linked with the studies that pay heed to the complexity of information in the longer reviews regarding the sentiment analysis and

highlight that comprehensive customer feedback provides airlines with a detailed understanding of consumer satisfaction levels.

### 4.7 Implication for Airlines and the Industry: Customer Experience Strategy

The findings from the study prove that airlines may improve customer satisfaction while prioritizing the core factors found to be the primary drivers of positive sentiment. For example, convenience, comfort, and attentive consumer service constantly affected favorable reviews, suggesting that airlines must improve such aspects to meet customer expectations (Hasanov et al., 2020). Moreover, investments in seating comfort, variety of meals, and well-trained staff can significantly improve satisfaction levels, creating a positive brand perception. Incorporation of sentiment analysis into consumer experience approaches can help airlines monitor levels of satisfaction in real time, which enables them to respond to the trends quickly. While understanding the areas contributing to the positive or negative sentiment, airlines may adjust their services to address the critical feedback areas and enhance customer satisfaction

# 5. ARGUMENT

This chapter delves into the findings of sentiment analysis, interpreting them within the broader context of the airline industry's customer satisfaction and brand loyalty frameworks. By examining customer reviews and feedback, the study sheds light on critical factors that shape sentiment trends, satisfaction, and loyalty. These findings are analyzed in relation to existing research, highlighting their implications for theory and practice in the airline industry.

The results reveal distinctive sentiment trends across airlines, where positive sentiments are closely associated with attentive customer service, punctuality, and value for money. Conversely, recurring issues like delays, inadequate entertainment options, and seating discomfort consistently drive negative sentiment. For instance, restricted legroom, particularly on long-haul flights, emerged as a frequent source of dissatisfaction. These findings align with prior research that identifies seat comfort as a pivotal driver of customer satisfaction in the airline sector (Farzadnia and Vanani, 2022). However, the coexistence of high ratings alongside negative feedback in specific service areas underscores the complexity of customer satisfaction, suggesting that a favorable overall rating does not necessarily indicate uniform satisfaction across all service aspects. This complexity highlights the importance of integrating qualitative feedback with quantitative metrics to form a holistic evaluation of customer satisfaction.

The relationship between sentiment trends and ratings further validates that certain service factors carry greater weight in shaping overall customer satisfaction. Airlines can anticipate and address customer expectations by prioritizing these critical factors. For instance, attributes like flight punctuality and staff responsiveness not only enhance satisfaction but also contribute to brand loyalty. Positive sentiment in customer reviews often translates into recommendations and repeat bookings, underscoring the direct link between satisfaction and loyalty. Additionally, airlines that excel in key areas, such as customer service, can maintain loyalty even when they receive mixed feedback in other domains. This finding demonstrates the significance of consistently meeting customer expectations in critical service areas to sustain brand loyalty (Rane et al., 2023).

The analysis also reveals significant disparities in sentiment based on flight class and route. Business-class passengers frequently report higher satisfaction levels due to superior seating comfort, enhanced meal options, and personalized services (Zhou et al., 2020). On the other hand, economy-class passengers often express dissatisfaction, particularly on long-haul flights, where limitations in comfort and amenities are more pronounced. Furthermore, international routes tend to garner more positive sentiment than domestic routes, attributed to the enhanced in-flight services offered on longer flights. These variations emphasize the need for airlines to address specific service factors that matter most to their diverse customer base, ensuring consistent standards across all routes and classes.

The findings of this study align with existing literature on sentiment analysis and its role in monitoring customer satisfaction. Sentiment analysis has been widely recognized as a powerful tool for capturing real-time consumer insights, enabling firms to address dissatisfaction promptly and enhance brand loyalty (Park, 2020). This study reaffirms the applicability of sentiment analysis in uncovering nuanced customer

experiences that go beyond numerical ratings. Moreover, the research supports the theoretical foundations of Brand Equity and the Theory of Planned Behavior (TPB). Positive sentiments, high ratings, and recommendations are consistent with the principles of Brand Equity, which emphasize the value of reliable, high-quality experiences in fostering loyalty. Similarly, TPB underscores how favorable attitudes toward a brand influence intention to recommend or repurchase, as evidenced by the strong correlations observed in this study (Ahmad and Guzmán, 2021).

In summary, the findings provide compelling evidence of the interplay between sentiment, satisfaction, and loyalty in the airline industry. By addressing the factors that drive positive sentiment and mitigating sources of dissatisfaction, airlines can strategically enhance their customer experience and brand perception. The integration of sentiment analysis into industry practices offers a robust framework for understanding customer needs and fostering loyalty in an increasingly competitive market.

# 6. CONCLUSIONS AND SUGGESTIONS

## 6.1 Summary of Findings

### 6.1.1 Overview of Customer Satisfaction Insights

The sentiment analysis regarding airline reviews signified many significant trends referring to customer satisfaction, with sentiment polarity linked to specific service attributes. Moreover, positive sentiments were frequently linked with excellent customer service, speedy flight schedules, and first-class amenities. Airlines that constantly provide attentive and friendly service have seen an uptick in positive sentiment, which indicates that customers tend to value human interaction as a factor of responsiveness. In fact. Comfort, mainly in business class seating, even evolved as a critical determinant of satisfaction, with passengers frequently citing the need for seat space and legroom.

Contrariwise, negative sentiment was even related to delays, restricted options of in-flight entertainment, and discomfort in the economy seating, particularly on long-haul flights. Even the passengers were dissatisfied with apparent inadequacies in such areas, which signaled that airlines might benefit while paying heed to bring improvements that address such pain points. The sentiment analysis discovered that overdue flights significantly impact customer satisfaction since the travelers showed frustration with the missed connections and inconvenience. In addition, feedback regarding in-flight amenities, for instance, inadequate meal options or poor entertainment on lengthy flights, has resulted in the dissatisfaction of economy class passengers, highlighting the need to prioritize customer comfort with a factor of satisfaction in airlines even in reduced fare categories.

This study showed that passengers' ratings are only sometimes fully revealing regarding particular service strengths or even weaknesses. As high ratings are usually linked with positive sentiment, specific negative factors were yet visible in the reviews of significantly rated airlines. All the findings suggested that sentiment analysis helped provide granular insights regarding customer satisfaction instead of the ratings alone and offered airlines a clear understanding of the driving forces of satisfaction and dissatisfaction.

### 6.1.2 Key Contributions to Brand Loyalty Understanding

The analysis provided significant insights regarding factors contributing to brand loyalty amongst airline passengers. Moreover, positive sentiment and high ratings were highly linked with loyalty aspects, such as recommendations and repeat bookings. Consumers having positive experiences explicitly endorsed the airline in the reviews while recommending it to others or expressing the intention of flying with the airline again. The correlation between positive experiences and loyalty links with theories such as Brand Equity Theory, which postulates that positive brand associations strengthen consumer retention and loyalty.

The study even demonstrated that specific service factors, like the quality of consumer service, had a, significantly stronger effect on loyalty. Airlines that provided consistent and responsive consumer support got positive mentions from consumers, and

this interaction strengthened brand loyalty. Moreover, the travellers valued prompt responses to complaints and inquiries, which helped build trust with reassurance, making them likely loyal to the airline. Hence, the relationship between consumer service and loyalty strengthens the significance of high-quality service as an introductory element of brand perception in the airline industry.

Overall, the sentiment analysis found many factors that influence customer satisfaction and loyalty, including comfort, consistency, and quality of customer service. The insights attained by the analysis give the airlines a strong framework to understand how the sentiment trends affect brand loyalty and offer guidance for specific improvements for fostering robust customer relationships.

### 6.2 Practical Suggestions for Airlines

To strengthen customer satisfaction and loyalty, airlines must consider a multi-layered approach that pays heed to three basic areas: enhancement in the customer experience, optimisation of digital engagement, and establishment of a real-time crisis management approach. Based on sentiment analysis insights, all the recommendations provide actionable phases for the airlines to address customers' needs and enhance loyalty, with the protection of the brand reputation.

#### 6.2.1 Improvement in the Customer Experience

Customer experience is a central sentiment trend in the airline industry, and factors like seating comfort, quality, service, and in-flight amenities even influence customer satisfaction (Li et al., 2024). The given steps specify how airlines can address such areas for the enhancement of the passenger experience while reducing negative feedback,

#### 6.2.2 Seating Comfort and Cabin Quality

Economy passengers, particularly on long-haul flights, tend to cite discomfort as a significant concern. Airlines may invest in ergonomic seating with modifiable headrests, extra legroom, and supportive cushioning (Al-Murrakshi, 2021). Moreover, provided the effect of comfort on sentiment, these enhancements can lead to prominent increases in customer satisfaction.

#### 6.2.3 Optimisation of Digital Engagement

Social media is a vital channel for customer interaction, and sentiment analysis can give insights into how passengers observe the airline's brand. Thus, airlines must use such insights to engage effectively on digital platforms, promote positive experiences, and proactively address undesirable sentiments.

#### 6.2.4 Ensuring a Real-Time Crisis Management Strategy

In the competitive airline industry, it is important to have effective crisis management to maintain customer trust and loyalty. The real-time sentiment monitoring approach lets the airlines detect any early warning signs regarding dissatisfaction and enables quick responses to mitigate the negative sentiment before it affects the brand perception (Numbi, 2024). Thus, it is necessary to implement real-time sentiment analysis

tools that monitor online reviews, overall social media mentions, and customer feedback platforms. This lets airlines stay informed of the possible crises as they explain. While having alerts for points in the negative sentiment, airlines can rapidly identify problems that need instant attention, like late flights, cancellations, or even challenges regarding customer service. With the development of a strong real-time crisis management system, the airlines can minimize the effect of negative sentiment on the brand, ensure customer trust retention, and foster a brand reputation for reliability even in difficult situations.

### 6.3 Limitations of the Analysis

The research had some limitations inherent in sentiment analysis and data interpretation. Sentiment analysis tools might need help with complicated language features, like irony, sarcasm, or mixed sentiments, that lead to misclassification. Furthermore, reviews might have various languages or cultural expressions, and the sentiment models must be revised for interpretation. These limitations are relevant to the international industry, such as airlines, where consumers may change between languages, such as English or regional dialects, affecting the correctness of sentiment classification.

As the analysis gives valuable insights, it also has some more limitations. One prominent challenge is the sentiment model's ability to accurately understand complicated language structures, like sarcasm or nuanced cultural orientations. These subtleties lead to sentiment classification inaccuracies and impact sentiment trends' interpretations. Cultural nuances, particularly in the reviews involving mixed language, have additional issues, such as the model might need to be more accurate in the overall context, mainly in multilingual settings. In addition, using data from various airlines provides general insights. However, it introduces the variability per differences in service standards, which may only partially represent the airlines.

Sentiment trends, core variables, and the effect of flight class with route were found in the Brand Equity Theory and the Theory of Planned Behavior. Moreover, the insights of this study gave airlines a roadmap to enhance the customer experience along with the management of digital engagement by sentiment monitoring. Furthermore, the chapter even addressed significant limitations, stressing the areas of improving the sentiment analysis along with the dataset comprehensiveness.

Sentiment analysis also revealed essential trends regarding customer satisfaction within the airline industry. Positive sentiment increased with higher ratings, and lower ratings had a negative relationship with rating comfort, service quality, and delay. The most positive sentiment, however, was found for premium services, such as business class and long-haul international routes, while economy class and domestic routes had more mixed feedback. Over time, we found that more positive reviews were longer and included deeper details about satisfaction or issues. The implications of these findings, particularly for improving brand perception and customer loyalty through focused service improvements, are significant, and the areas to precisely target are outlined.

## 6.4 Future Research Directions

### 6.4.1 Enhancement in Sentiment Analysis Models

An area regarding future research refers to enhancing the sentiment analysis models for capturing complicated emotional nuances. The traditional sentiment models even struggle to accurately interpret sarcasm, causticness, and mixed sentiments, which are common in customer feedback. For example, a passenger may praise the airline ironically while saying, "The amazing five-hour delay in history!" (Venugopalan and Gupta, 2022). Present models may misinterpret this sarcasm as a positive sentiment, which skews the analysis results.

Future research may explore the advanced natural language processing (NLP) ways, like deep learning models having contextualized embeddings (for instance, BERT, Roberta), that better know the sentiment cues while considering the case of words in the sentences (Venugopalan and Gupta, 2022). In addition, hybrid sentiment models combined with rule-led approaches and machine learning help distinguish nuanced sentiments in diversified customer reviews. Hence, improving the sentiment analysis models may yield better insights into customer perceptions, enabling airlines to respond effectively to the feedback..

### 6.4.2 Industry-based Sentiment Studies

One more promising direction is conducting industry-based sentiment studies that pay heed to the longitudinal trends in individual airlines. Examination of the sentiment patterns with time for an airline, as opposed to the cross-sectional studies of many airlines, tends to reveal significant changes in policies, quality of service, or branding, which affects customer satisfaction, including loyalty (Xiang et al., 2022). The longitudinal approach lets researchers track their shifts in customer sentiment, which provides airlines with a comprehensive view of the enduring impacts of strategic decisions. Moreover, sector-based studies in the airline industry, like examination of the sentiment based on the budget airlines against premium carriers, will provide targeted insights. Such insights can help guide the airlines to refine their competitive positioning with service enhancements as per customer expectations that are unique to all the market segments.

### 6.4.3 Cross-Platform Analysis

Lastly, cross-platform analysis regarding sentiment across social media channels such as Facebook, Instagram, and Twitter can give a wider view of customer engagement (Mutinda et al., 2023). All the platforms have different user bases and interaction styles, which can influence how consumers express their satisfaction or dissatisfaction. For instance, the character limit of Twitter even leads towards brief, precise reviews, but Facebook gives more detailed feedback. Hence, understanding such platform-led differences is essential for airlines to seek and tailor their engagement approaches.

## 6.5 Conclusion

This research underlines the value of sentiment analysis in understanding and improving customer satisfaction with brand loyalty in the airline industry. By comprehensive analysis based on the customer reviews, core factors that affect the positive and negative sentiments were found, which included customer service quality, level of comfort, and digital engagement. Thus, the findings provide actionable insights for the airlines, demonstrating how strategic enhancements can improve service quality with real-time engagement. All this can foster overall customer loyalty while strengthening the brand perception. Thus, the practical applications regarding sentiment analysis, mainly for the proactive customer relationship management, crisis management, and digital engagement, pay heed to the tool's importance for competitive advantage in the service-oriented industry.

## REFERENCES

Aftab, F., Bazai, S.U., Marjan, S., Baloch, L., Aslam, S., Amphawan, A. & Neo, T.K., 2023. A Comprehensive Survey on Sentiment Analysis Techniques. International Journal of Technology, 14(6), p.12881298.

Agarwal, S., 2022. Deep learning-based sentiment analysis: Establishing customer dimension as the lifeblood of business management. Global Business Review, 23(1), pp.119-136.

Agnihotri, D., Kulshreshtha, K. & Tripathi, V., 2021, April. A study on firms' communication based on artificial intelligence and its influence on customers' complbehaviouravior in Social media environment. In IOP Conference Series: Materials Science and Engineering (Vol. 1116, No. 1, p. 012180). IOP Publishing.

Ahmad, F. & Guzmán, F., 2021. Negative online reviews, brand equity and emotional contagion. European Journal of Marketing, 55(11), pp.2825-2870.

Ajzen, I., 2020. The theory of planned behaviour: Frequently asked questions. Human behavior and emerging technologies, 2(4), pp.314-324.

Akritidis, L. & Bozanis, P., 2023. Low-dimensional text representations for sentiment analysis NLP tasks. SN Computer Science, 4(5), p.474.

Al-Ansari, J.A., 2019. From sentiment analysis to choreography of emotions: social media analysis for improved customer relationship management (CRM) in the Omani telecom sector.

Ali, S., 2024. Unveiling Brand Identity: A Semiotic Exploration of Logo Design in Pakistani Clothing Brands. Academy of Education and Social Sciences Review, 4(3), pp.329-340.

Al-Murrakshi, M., 2021. Seat comfort issues in Economy Class and their effect on long-haul passenger satisfaction and future re-flying intentions (Doctoral dissertation, Auckland University of Technology).

Al-Natour, S. & Turetken, O., 2020. A comparative assessment of sentiment analysis and star ratings for consumer reviews. International Journal of Information Management, 54, p.102132.

Al-Obaydy, W.I., Hashim, H.A., Najm, Y.A. & Jalal, A.A., 2022. Document classification using term frequency-inverse document frequency and K-means clustering. Indonesian Journal of Electrical Engineering and Computer Science, 27(3), pp.1517-1524.

Alvarado-Karste, D. & Guzmán, F., 2020. The effect of brand identity-cognitive style fit and social influence on consumer-based brand equity. Journal of Product & Brand Management, 29(7), pp.971-984.

ATAG - https://www.atag.org/facts-figures/ - 17/09/2023

Ausat, A.M.A., 2023. The role of social media in shaping public opinion and its influence on economic decisions. Technology and Society Perspectives (TACIT), 1(1), pp.35-44.

Azemi, Y., Ozuem, W., Wiid, R. & Hobson, A., 2022. Luxury fashion brand customers'

perceptions of mobile marketing: Evidence of multiple communications and marketing channels. Journal of Retailing and Consumer Services, 66, p.102944.

Ban HJ. & Kim HS. Understanding customer experience and satisfaction through airline passengers' online review. Sustainability. 2019;11(15):4066. [Page 3]

Bashir, T., Zhongfu, T., Sadiq, B., Niaz, U., Anjum, F. & Mahmood, H., 2022. An assessment of influential factors developing the intention to use social media sites: A technology acceptance model-based approach. Frontiers in Psychology, 13, p.983930.

Batista, J.M., Barros, L.S., Peixoto, F.V. & Botelho, D., 2022. Sarcastic or assertive: How should brands reply to consumers' uncivil comments on social media in the context of brand activism? Journal of interactive marketing, 57(1), pp.141-158.

Bell, C., Olukemi, A. & Broklyn, P., 2024. Social Media Sentiment Analysis for Brand Reputation Management.

Bellary, S., Bala, P.K. & Chakraborty, S., 2024. Exploring cognitive-behavioral drivers impacting consumer continuance intention of fitness apps using a hybrid approach of text mining, SEM, and ANN. Journal of Retailing and Consumer Services, 81, p.104045.

Bharadwaj, L., 2023. Sentiment analysis in online product reviews: mining customer opinions for sentiment classification. Int J Multidiscip Res, 5(5).

Birjali, M., Kasri, M. & Beni-Hssane, A., 2021. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowledge-Based Systems, 226, p.107134.

Bordovskaia, E., 2023. Public Communication in Crisis Response Management.

Bougie, R. & Sekaran, U., 2019. Research methods for business: A skill building approach. John Wiley & Sons.

Bozkurt, S., Gligor, D., Hollebeek, L.D. & Sumlin, C., 2024. Understanding the effects of firms' unresponsiveness on social media toward customer feedback on customers' engagement: the impact of ethnicity. Journal of Research in Interactive Marketing.

Bruce, K.L., 2020. Assessing Brand-Consumer Personality Congruence on Twitter: A Computational Textual Analysis Approach. The University of Alabama.

Chai, C.P., 2020. The importance of data cleaning: Three visualisation examples. Chance, 33(1), pp.4-9.

Clark, T., Foster, L., Bryman, A. and Sloan, L., 2021. Bryman's social research methods. Oxford University Press.

Coelho, P.S., Rita, P. & Santos, Z.R., 2018. On the relationship between consumer-brand identification, brand community, and brand loyalty. Journal of Retailing and Consumer Services, 43, pp.101-110.

Cooper, T., Stavros, C. & Dobele, A.R., 2019. Domains of influence: exploring negative sentiment in social media. Journal of Product & Brand Management, 28(5), pp.684-699.

Creswell, J.W. & Creswell, J.D., 2017. Research design: Qualitative, quantitative, and mixed methods approach. Sage publications.

Dash, B., 2022. Information Extraction from Unstructured Big Data: A Case Study of Deep Natural Language Processing in Fintech. University of the Cumberlands.

De Vadder, S.F., 2024. Using Twitter as a public communication strategy.

Du Chau, X.T., 2024. Exploring Customer Insights of Content Platforms with Big Data Analytics.

Duan, Y., Hemsley, J., Smith, A.O., Joh, U., Gray, L. & Khoury, C., 2024. Curating Virality: Exploring Curated Logics Within# BlackLivesMatter on Twitter/X. Social Media+ Society, 10(2), p.20563051241242799.

Dwivedi, Y.K., Ismagilova, E., Hughes, D.L., Carlson, J., Filieri, R., Jacobson, J., Jain, V., Karjaluoto, H., Kefi, H., Krishen, A.S. & Kumar, V., 2021. Setting the future of digital and social media marketing research: Perspectives and research propositions. International journal of information management, 59, p.102168.

Dzilská, Z., Kalamen, K., Pollák, F. & Ruschak, M., 2024. Effective Management of Marketing Communication in Social Media.

Eachempati, P., Srivastava, P.R., Kumar, A., de Prat, J.M. & Delen, D., 2022. Can customer sentiment impact firm value? An integrated text mining approach. Technological Forecasting and Social Change, 174, p.121265.

Ebo, D., 2024. Social Media Journalism and Its Impact on Social Change: The Case of Twitter (X) (Master's thesis, Western Michigan University).

EDA, C., 2022. Unveiling the Ambivalence from the Airline Reviews: The Airline Recommendation System using CNN.

Eginli AT. & Taş N. Sentiment Analysis in Social Media Marketing: Leveraging Natural Language Processing for Consumer Insights. In: Advances in Digital Marketing and eCommerce. Springer; 2023:567-580.

Eginli, A.T. & Taş, N.Ö., 2023. Emotions on social media: a sentiment analysis approach based on Twitter (x) data on the Russian-Ukraine war. International Journal of Social Inquiry, 16(2), pp.445-459.

Farzadnia, S. & Vanani, I.R., 2022. Identification of opinion trends using sentiment analysis of airlines passengers' reviews. Journal of Air Transport Management, 103, p.102232.

Feine J, Morana S, Gnewuch U. Measuring Service Encounter Satisfaction with Customer Service Chatbots using Sentiment Analysis. In: 14th Internationale Tagung Wirtschaftsinformatik (WI 2019). 2019:1115-1129.

Garay, L. & Morales, S., 2020. Decomposing and relating user engagement in festivals' virtual brand communities: An analysis of Sónar's Twitter and Facebook. Tourist Studies, 20(1), pp.96-119.

Garg, S., Panwar, D.S., Gupta, A. & Katarya, R., 2020, November. A literature review on sentiment analysis techniques involving social media platforms. In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) (pp. 254-259). IEEE.

Ghani, R., Hossain, E., & Karim, M. (2019). Customer satisfaction analysis in the airline industry using sentiment analysis.

Ghosh S. & Veale T. A Survey of Sentiment Analysis and Sarcasm Detection: Challenges, Techniques, and Trends. arXiv preprint arXiv:2301.12303. 2023.

Gundapu, S., 2022. Automatic Detection of Negativity in User-Generated Social Media Content: Building Models for Fake News, Sentiment Analysis, Propaganda Techniques, Offensive and Sarcastic Content Identification (Doctoral dissertation, International Institute of Information Technology Hyderabad).

Guo, X., Heinberg, M. & Zou, S., 2019. Enhancing consumer attitude toward culturally mixed symbolic products from foreign global brands in an emerging-market setting: the role of cultural respect. Journal of International Marketing, 27(3), pp.79-97.

Haris, M., Nasir, N. & Cheema, S., 2022. The impact of social media activities on emotional attachment with the mediating role of brand image and brand commitment of retail sector. Review of Education, Administration & Law, 5(2), pp.173-188.

Hartmann, J., Heitmann, M., Siebert, C. & Schamp, C., 2023. More than a feeling: Accuracy and application of sentiment analysis. International Journal of Research in Marketing, 40(1), pp.75-87.

Hasanov, J. (n.d.). The primary causes of unfavorable customer reviews in low-cost airline businesses: The case of Wizz Air.

Hasib, M. M., Rahman, M. S., & Islam, M. S. (2021). A systematic review on airline industries based on sentiment analysis and topic modeling. arXiv preprint. https://doi.org/10.48550/arXiv.2101.12487

Hasib MM, Rahman MS, Islam MS. A systematic review on airlines industries based on sentiment analysis and topic modeling. arXiv preprint arXiv:2101.12487. 2021.

Higgins, L. (2022). Classification of airline customer sentiment expressed in Twitter tweets using lexicons, decision tree, and naïve Bayes (Doctoral dissertation). National College of Ireland, Dublin. https://trap.ncirl.ie

Hobbs, M. J., & O'Keefe, S. (2024). Agonism in the arena: Analyzing cancel culture using a rhetorical model of deviance and reputational repair. Public Relations Review, 50(1), 102420. https://doi.org/10.1016/j.pubrev.2023.102420

Hosain, M.S. & Mamun, A.M.A., 2023. Facebook-based social media marketing and Facebook-based online purchases: evidence from the Facebook page admins of selected South Asian fashion retailers. Journal of Electronic Business & Digital Economics, 2(2), pp.191-212.

Islam MT, Khan T. A Comparative Study of Sentiment Analysis Using NLP and Different Machine Learning Techniques on US Airline Twitter Data. International Journal of Advanced Computer Science and Applications. 2021;12(9):634-641.

Itoo, F., Meenakshi & Singh, S., 2021. Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. International Journal of Information Technology, 13(4), pp.1503-1511.

Jain, N. & Kohli, P., 2024. Beyond Borders. DeepMisti Publication.

Jelodar, H., Wang, Y., Rabbani, M., Ahmadi, S.B.B., Boukela, L., Zhao, R. & Larik, R.S.A., 2021. A NLP framework based on meaningful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on youtube comments. Multimedia Tools and Applications, 80, pp.4155-4181.

Jena, R.K., 2019. Sentiment mining in a collaborative learning environment: capitalising on big data. Behaviour & Information Technology, 38(9), pp.986-1001.

Jiang H, & Zhang Y. Passenger satisfaction analysis of domestic vs. international airline routes using sentiment analytics. Journal of Transport & Tourism. 2022;31(5):245-260.

Jiang X, Zhang Y, Li Y, & Zhang B. Forecast and analysis of aircraft passenger satisfaction based on RF-RFE-LR model. Scientific Reports. 2022;12(1):11174. [Page 5]

Jiang, G., Liu, F., Liu, W., Liu, S., Chen, Y. & Xu, D., 2021. Effects of information quality on information adoption on social media review platforms: Moderating role of perceived risk. Data Science and Management, 1(1), pp.13-22.

Kapoor K, Gunta S, Rohit S. Self-Disclosure and Opinion Expression: A Cross-Platform Study on Facebook and Twitter. Journal of Behavioral Communication. 2022;19(3):121-136.

Kapoor, S. & Banerjee, S., 2021. On the relationship between brand scandal and consumer attitudes: A literature review and research agenda. International Journal of Consumer Studies, 45(5), pp.1047-1078.

Kashif, S. & Mubarik, S., 2020. An evolutionary historical perspective of Pakistan retail fashion industry. JISR management and social sciences & economics, 18(1), pp.53-75.

Kauffmann, E., Peral, J., Gil, D., Ferrández, A., Sellers, R. and Mora, H., 2019. Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining. Sustainability, 11(15), p.4235.

Kaur S, Singh J. Sentiment analysis of airline passenger reviews to compare satisfaction levels between travel classes. International Journal of Hospitality and Tourism Management. 2021;32(3):167-181.

Kaur S. Social Media Sentiment Analysis for Brand Monitoring. International Journal of Engineering Research & Technology (IJERT). 2024;13(10):1-5.

Keller, K.L., 2020. Consumer research insights on brands and branding: a JCR curation. Journal of ConResearchsearch, 46(5), pp.995-1001.

Khan, F., Sajid, A. & Kamran, S., 2019. Dimensions of Brand Equity affecting Consumption of Ḥalāl/Ḥarām Products: Moderating Role of Islamic Work Ethics. Journal of Islamic Business and Management, 9(2).

Khan, S., Khan, I.U., Sarwar, M. & Gul, S., 2021. Gender-Based Objectification In Commercial Advertisements: A Socio-Cultural Analysis. Webology (ISSN: 1735-188X), 18(6).

Khawaja, S.S., 2023. Refashioning Style from the Outside-In: A Pakistani Diasporic Response.

Kim, J., & Lee, W. (2020). Enhancing brand loyalty through customer experience management in airlines.

Klucarova, S. & He, X., 2022. Status consumption and charitable donations: The power of empowerment. Psychology & Marketing, 39(6), pp.1116-1128.

Koiou, M., 2021. Social Media use among brand evangelists: How Social Media Marketing affects the Greek consumers' perception of "generic" vs."popular" clothing brands.

Kotras, B., 2020. Opinions that matter: the hybridization of opinion and reputation measurement in social media listening software. Media, Culture & Society, 42(7-8), pp.1495-1511.

Kumar S, Zymbler M. A machine learning approach to analyze customer satisfaction from airline tweets. J Big Data.2019;6(1):1-15.

Kumaradeepan, V., Azam, F. and Tham, J., 2023. Factors influencing of social media on consumer perception and purchase intention towards brand loyalty: a conceptual paper. International Journal of Professional Business Review, 8(5), pp.e01571-e01571.

Kwon, H.J., Ban, H.J., Jun, J.K. & Kim, H.S., 2021. Topic modelling and sentiment analysis of online review for airlines. Information, 12(2), p.78.

Lancia, F., 2020. Social Sciences, Emotions and Algorithms.

Lau, A.C., 2022. The impact of airport service quality on passenger satisfaction at the Hong Kong International Airport. University of Wales Trinity Saint David (United Kingdom).

Li, Q., Jing, R. & Zhu, X., 2024. Determinants of travel satisfaction for commercial airlines: A data mining approach. Engineering Applications of Artificial Intelligence, 133, p.108597.

Lucini FR, Tonetto LM, Fogliatto FS, Anzanello MJ. Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. J Air Transp Manag. 2020;83:101760.

Luttrell, R. & Wallace, A.A., 2024. Social media and society: An introduction to the mass media landscape. Rowman & Littlefield.

Malik, S. & Hassan, K.A.K., 2019. An investigation of house designs in Lahore: Transformation of residential architecture from traditional to modern. Journal of Design and Built Environment, 19(1), pp.49-59.

Manimegalai, D., Sentiment analysis on the brand recall of twitter data with the tool of NCSU tweet sentiment visualization.

Manyanga, W., 2022. Mediating effects of customer satisfaction and word-of-mouth intention on the relationship between customer experience and loyalty: evidence from the banking sector in Zimbabwe (Doctoral dissertation, Chinhoyi University if Technology).

Manzoor, U., Baig, S.A., Malik, E. and Shahid, M.I., 2020. Consumer perceptions of brands in Pakistan's denim industry. Pakistan Journal of Multidisciplinary Research, 1(1), pp.1-13.

Maree, T. & Van Heerden, G., 2021. Beyond the "like": customer engagement of brand fans on Facebook. European Business Review, 33(2), pp.255-271.

McPhearson T. A qualitative analysis of sarcasm, irony, and related phenomena in social media interactions. Big Data & Society. 2018;5(2):1-12.

Mehraliyev, F., Kirilenko, A.P. & Choi, Y., 2020. From measurement scale to sentiment scale: Examining the effect of sensory experiences on online review rating behavior. Tourism Management, 79, p.104096.

Mohammad, S.M., 2022. Ethics sheet for automatic emotion recognition and sentiment analysis. Computational Linguistics, 48(2), pp.239-278.

Mohan, S., Solanki, A.K., Taluja, H.K. and Singh, A., 2022. Predicting the impact of the third wave of COVID-19 in India using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach. Computers in Biology and Medicine, 144, p.105354.

Muniesa, R.L. & Giménez, C.G., 2020. The importance of the loyalty of fashion brands through digital marketing. Journal of Spatial and Organizational Dynamics, 8(3), pp.230-243.

Mutinda, J., Mwangi, W. & Okeyo, G., 2023. Sentiment analysis of text reviews using lexicon-enhanced bert embedding (LeBERT) model with convolutional neural network. Applied Sciences, 13(3), p.1445.

Muzinda, M., 2021. Customer Service Strategies to Minimise Negative Customer Incidents in the Airline Industry (Doctoral dissertation, Walden University).

Namukasa J. The influence of airline service quality on passenger satisfaction and loyalty: The case of Uganda airline industry. The TQM Journal. 2013;25(5):520-532.

Nasuto, A. & Rowe, F., 2024. Understanding anti-immigration sentiment spreading on Twitter. PloS one, 19(9), p.e0307917.

Nawaz, N. & Kashif, S., 2021. Sentiment Analysis of Consumer Misbehavior via Apparel Tweets: Consumer Misbehavior. Sukkur IBA Journal of Management and Business, 8(1), pp.28-47.

Niazi, N.K., 2020. The Effects of a Visually Warm (Vs Cold) Retail Design on the Store Patronage Intentions (Doctoral dissertation, © Lahore School of Economics).

Numbi, L.K., 2024. The Impact of Service Quality, Logistic Management, And Safety Towards Customer Satisfaction to Use the National Airline Carriers in Africa: A Case Study of the Democratic Republic of Congo Aviation Sector (Doctoral dissertation, Bangkok University).

Obiedat, R., Qaddoura, R., Ala'M, A.Z., Al-Qaisi, L., Harfoushi, O., Alrefai, M.A. & Faris, H., 2022. Sentiment analysis of customers' reviews using a hybrid evolutionary SVM-based approach in an imbalanced data distribution. IEEE Access, 10, pp.22260-22273.

Palalic, R., Ramadani, V., Mariam Gilani, S., Gërguri-Rashiti, S. & Dana, L.P., 2021.

Social media and consumer buying behavior decision: what entrepreneurs should know?. Management Decision, 59(6), pp.1249-1270.

Park E, Jang Y, & Kim J. Comparative sentiment analysis of domestic and international airline routes. Journal of Travel Research. 2023;62(4):567-582.

Park, E.H. & Storey, V.C., 2023. Emotion ontology studies: A framework for expressing feelings digitally and its application to sentiment analysis. ACM Computing Surveys, 55(9), pp.1-38.

Park, J., 2020. Framework for sentiment-driven evaluation of customer satisfaction with cosmetics brands. IEEE Access, 8, pp.98526-98538.

Patel, A., Oza, P. & Agrawal, S., 2023. Sentiment analysis of customer feedback and reviews for airline services using language representation model. Procedia Computer Science, 218, pp.2459-2467.

Payton, P., 2021. Crisis Management Strategies for Sustaining Organizations During a Crisis (Doctoral dissertation, Walden University).

Polanský, M. and Sylvester-Hvid, L., 2022. A Comparative Study Concerning Twitter's Influence on Company Stock Market Performance.

Rane, N., 2023. Enhancing customer loyalty through Artificial Intelligence (AI), Internet of Things (IoT), and Big Data technologies: improving customer satisfaction, engagement, relationship, and experience. Internet of Things (IoT), and Big Data Technologies: Improving Customer Satisfaction, Engagement, Relationship, and Experience (October 13, 2023).

Rane, N.L., Achari, A. & Choudhary, S.P., 2023. Enhancing customer loyalty through quality of service: Effective strategies to improve customer satisfaction, experience, relationship, and engagement. International Research Journal of Modernization in Engineering Technology and Science, 5(5), pp.427-452.

Rialti, R., Zollo, L., Kim, K. & Yoon, S., 2024. Mega-influencers and brand dynamics: Shaping attitudes toward leading and challenger brands through electronic word of mouth. Psychology & Marketing.

Rust, R.T., Rand, W., Huang, M.H., Stephen, A.T., Brooks, G. & Chabuk, T., 2021. Real-time brand reputation tracking using social media. Journal of Marketing, 85(4), pp.21-43.

Salah-Ud-Din M, Lincy BT, & Al Ali H. Exploratory Data Analysis and Prediction of Passenger Satisfaction with Airline Services. IEEE. (Page 11-12)

Samir, H.A., Abd-Elmegid, L. & Marie, M., 2023. Sentiment analysis model for Airline customers' feedback using deep learning techniques. International Journal of Engineering Business Management, 15, p.18479790231206019.

Sánchez-Rada, J.F. & Iglesias, C.A., 2019. Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. Information Fusion, 52, pp.344-356.

Shahid, R., Mozumder, M.A.S., Sweet, M.M.R., Hasan, M., Alam, M., Rahman, M.A., Prabha, M., Arif, M., Ahmed, M.P. & Islam, M.R., 2024. Predicting Customer Loyalty in the Airline Industry: A Machine Learning Approach Integrating

Sentiment Analysis and User Experience. International Journal on Computational Engineering, 1(2), pp.50-54.

Shahid, S., Paul, J., Gilal, F.G. & Ansari, S., 2022. The role of sensory marketing and brand experience in building emotional attachment and brand loyalty in luxury retail stores. Psychology & Marketing, 39(7), pp.1398-1412.

Singh, N.K., Tomar, D.S. & Sangaiah, A.K., 2020. Sentiment analysis: a review and comparative analysis over social media. Journal of Ambient Intelligence and Humanized Computing, 11(1), pp.97-117.

Srinivas S, Ramachandiran S. Passenger intelligence as a competitive opportunity: unsupervised text analytics for discovering airline-specific insights from online reviews. Ann Oper Res. 2024;333:1045-1075.

StatCounter Global Stats. (n.d.). Social Media Stats Pakistan. [online] Available at: https://gs.statcounter.com/social-media-stats/all/pakistan.

Sümer, S.I. & Parilti, N. eds., 2023. Social Media Analytics in Predicting Consumer Behavior. CRC Press.

Swaminathan, V., Sorescu, A., Steenkamp, J.B.E., O'Guinn, T.C.G. and Schmitt, B., 2020. Branding in a hyperconnected world: Refocusing theories and rethinking boundaries. Journal of marketing, 84(2), pp.24-46.

Tan KL, Lee CP, & Lim KM. A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. Applied Sciences. 2023;13(7):4550.

Tejpal, M., 2020. Enhancing Customer Satisfaction through Fashion Retail. Journal of Textile and Apparel, Technology and Management, 11(3).

Tumuluru, P., Hussain, S.S., Kankanala, L., Shoaib, S., Madhu, D. & Kumar, C.N., 2023, October. Advancing Twitter Sentiment Analysis: An Ensemble Approach with Transformer-XL, RoBERTa, and XGBoost. In 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS) (pp. 944-950). IEEE.

Uçar, M.K., Nour, M., Sindi, H. & Polat, K., 2020. The effect of training and testing process on machine learning in biomedical datasets. Mathematical Problems in Engineering, 2020(1), p.2836236.

Venugopalan, M. & Gupta, D., 2022. An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis. Knowledge-based systems, 246, p.108668.

Wang, J., Wang, X., Yang, Y., Zhang, H. & Fang, B., 2020. A Review of Data Cleaning Methods for Web Information System. Computers, Materials & Continua, 62(3).

Wankhade, M., Rao, A.C.S. & Kulkarni, C., 2022. A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), pp.5731-5780.

Wu S, & Gao Y. Happy or grumpy? A machine learning approach to analyze the sentiment of airline passengers' tweets. arXiv preprint arXiv:2209.14363. 2022.

Wu S, & Gao Y. Route-specific sentiment analysis to uncover service gaps in airline offerings. Journal of Service Management. 2022;15(3):325-340.

Xiang, C., Zhang, J., Li, F., Fei, H. & Ji, D., 2022. A semantic and syntactic enhanced neural model for financial sentiment analysis. Information Processing & Management, 59(4), p.102943.

Xie H, Li Y, Pu Y, Zhang C, & Huang J. Evaluating airline service quality through comprehensive text-mining and TOPSIS-VIKOR-AISM analysis.

Xie H, Li Y, Pu Y, & Zhang C. Evaluating airline service quality through comprehensive sentiment analysis. Journal of Service Quality and Management. 2022;19(4):215-228.

Xu, Q.A., Chang, V. & Jayne, C., 2022. A systematic review of social media-based sentiment analysis: Emerging trends and challenges. Decision Analytics Journal, 3, p.100073.

Yacouby, R. & Axman, D., 2020, November. Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. In Proceedings of the first workshop on evaluation and comparison of NLP systems (pp. 79-91).

Yas H, et al. The effects of airline service quality and passenger satisfaction on passenger loyalty in the airline industry. ResearchGate. 2021.

Yousuf, O. & Siddiqui, D.A., 2019. The influence of brand equity on consumer responses: Evidence from a clothing brand of Pakistan. Available at SSRN 3397489.

Yuan, D., Lin, Z., Filieri, R., Liu, R. & Zheng, M., 2020. Managing the product-harm crisis in the digital era: The role of consumer online brand community engagement. Journal of business research, 115, pp.38-47.

Zhou, Y., Zhang, T., Mo, Y. & Huang, G., 2020. Willingness to pay for economy class seat selection: From a Chinese air consumer perspective. Research in Transportation Business & Management, 37, p.100486.

Zunic, A., Corcoran, P. & Spasic, I., 2020. Sentiment analysis in health and well-being: systematic review. JMIR medical informatics, 8(1), p.e16023.

# APPENDIX

Importing Libraries ⬇

```
[ ] import pandas as pd
    from dateutil import parser
    import matplotlib.pyplot as plt
    import seaborn as sns
    from wordcloud import WordCloud
    import numpy as np
    import re
    import os
    import nltk
    from nltk.corpus import stopwords
    from nltk.tokenize import word_tokenize
    from sklearn.feature_extraction.text import TfidfVectorizer
    from sklearn.model_selection import train_test_split
    from sklearn.linear_model import LogisticRegression
    from sklearn.naive_bayes import MultinomialNB
    from sklearn.model_selection import GridSearchCV
    from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, classification_report, confusion_matrix
    import joblib
    import lime
    from lime.lime_text import LimeTextExplainer
```

## Data Loading and Merging ⬇

```
[ ] df1 = pd.read_csv('airline_reviews_1.csv')
    df2 = pd.read_csv('airline_reviews_2.csv')
    df3 = pd.read_csv('airline_reviews_3.csv')

    # Dropping empty rows fromfinal_df3
    df3 = df3.drop(df3.index[df3.isna().all(axis=1)])

    # Renaming columns in df2 to match df1
    df2_renamed = df2.rename(columns={
        'Airline Name': 'Airline',
        'Overall_Rating': 'Overall Rating',
        'Review_Title': 'Title',
        'Review Date': 'Review Date',
        'Review': 'Reviews',
        'TypeOfTraveller': 'TypeofTraveller',
        'Seat Type': 'Class',
        'Date Flown': 'Month Flown',
        'Cabin Staff Service': 'Staff Service'
    })
```

```python
# Renaming columns in df3 to match df1
df3_renamed = df3.rename(columns={
    'airline': 'Airline',
    'overall': 'Overall Rating',
    'review_date': 'Review Date',
    'customer_review': 'Reviews',
    'traveller_type': 'TypeofTraveller',
    'cabin': 'Class',
    'route': 'Route',
    'date_flown': 'Month Flown',
    'seat_comfort': 'Seat Comfort',
    'cabin_service': 'Staff Service',
    'food_bev': 'Food & Beverages',
    'entertainment': 'Inflight Entertainment',
    'value_for_money': 'Value For Money',
    'recommended': 'Recommended'
})

# Dropping unique and unecessary columns
columns_to_drop = ['Aircraft', 'Ground Service', 'Wifi & Connectivity', 'Name', 'author', 'aircraft', 'ground_service']
df1 = df1.drop(columns=[col for col in columns_to_drop if col in df1.columns], errors='ignore')
df2_renamed = df2_renamed.drop(columns=[col for col in columns_to_drop if col in df2_renamed.columns], errors='ignore')
df3_renamed = df3_renamed.drop(columns=[col for col in columns_to_drop if col in df3_renamed.columns], errors='ignore')
```

```python
final_columns = [
    'Title', 'Review Date', 'Airline', 'Verified', 'Reviews', 'TypeofTraveller',
    'Month Flown', 'Route', 'Class', 'Seat Comfort', 'Staff Service',
    'Food & Beverages', 'Inflight Entertainment', 'Value For Money',
    'Overall Rating', 'Recommended'
]

for df in [df1, df2_renamed, df3_renamed]:
    for column in final_columns:
        if column not in df.columns:
            df[column] = None

df1 = df1[final_columns]
df2_renamed = df2_renamed[final_columns]
df3_renamed = df3_renamed[final_columns]

merged_df = pd.concat([df1, df2_renamed, df3_renamed], ignore_index=True)
merged_df.to_csv('final_airline_reviews.csv', index=False)
print("Datasets merged successfully!")
```

## Data Cleaning ⬇

```
# Formatting Date in dd/mm/yyyy format
final_df = pd.read_csv('final_airline_reviews.csv')

def format_date(date_str):
    try:
        return parser.parse(date_str).strftime('%d/%m/%Y')
    except Exception as e:
        print(f"Error parsing date: {date_str}, {e}")
        return None

final_df['Review Date'] = final_df['Review Date'].apply(format_date)
final_df['Review Date'] = pd.to_datetime(final_df['Review Date'], format='%d/%m/%Y', errors='coerce')
final_df = final_df.dropna(subset=['Review Date'])

final_df['Overall Rating'] = pd.to_numeric(final_df['Overall Rating'], errors='coerce')
final_df = final_df.dropna(subset=['Overall Rating'])

# Adding 'Recommended_Binary' based on the value of the 'Recommended' column
final_df['Recommended_Binary'] = final_df['Recommended'].apply(lambda x: 1 if str(x).strip().lower() == 'yes' else 0)

# Drop duplicate Reviews for the same Airline
final_df = final_df[~final_df.duplicated(subset=['Reviews', 'Airline'], keep='first')]

final_df.to_csv('final_airline_reviews.csv', index=False)

print("Data cleaned and file saved successfully!")
```

## Cleaning Reviews

```
current_directory = os.getcwd()  # Get the current working directory
nltk_data_path = os.path.join(current_directory, 'nltk_data')

# Set NLTK to use the specified directory
nltk.data.path.append(nltk_data_path)

# Download necessary NLTK resources to the current directory
nltk.download('punkt', download_dir=nltk_data_path)
nltk.download('stopwords', download_dir=nltk_data_path)

def clean_text(text):
    text = text.lower()
    text = re.sub(r'[^a-z\s]', '', text)

    words = word_tokenize(text)

    stop_words = set(stopwords.words('english'))
    words = [word for word in words if word not in stop_words]
    return ' '.join(words)

final_df['Cleaned_Reviews'] = final_df['Reviews'].apply(lambda x: clean_text(str(x)))
final_df.to_csv('final_airline_reviews.csv', index=False)
print(final_df[['Reviews', 'Cleaned_Reviews']].head())
```

## Data Visualisation ⬇

```
plt.figure(figsize=(12, 6))
final_df['Review Date'].groupby(final_df['Review Date'].dt.to_period('M')).count().plot(kind='line', marker='o', color='teal')
plt.title('Number of Reviews Over Time')
plt.xlabel('Month')
plt.ylabel('Number of Reviews')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 5))
sns.countplot(data=final_df, x='TypeofTraveller', )
plt.title('Distribution of Types of Traveller')
plt.xlabel('Type of Traveller')
plt.ylabel('Number of Reviews')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 5))
sns.histplot(final_df['Overall Rating'], bins=10, kde=True, color='blue')
plt.title('Distribution of Overall Ratings')
plt.xlabel('Overall Rating')
plt.ylabel('Frequency')
plt.tight_layout()
plt.show()
```

```
final_df['Word Count'] = final_df['Reviews'].astype(str).apply(lambda x: len(x.split()))
plt.figure(figsize=(10, 5))
sns.histplot(final_df['Word Count'], bins=30, color='purple')
plt.title('Word Density Distribution for Reviews')
plt.xlabel('Word Count')
plt.ylabel('Frequency')
plt.tight_layout()
plt.show()

all_reviews = ' '.join(final_df['Reviews'].dropna().astype(str).tolist())
wordcloud = WordCloud(width=800, height=400, background_color='white', max_words=200, colormap='plasma').generate(all_reviews)

plt.figure(figsize=(10, 6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('WordCloud of Reviews')
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 6))
sns.boxplot(data=final_df, x='TypeofTraveller', y='Overall Rating', palette='husl')
plt.title('Ratings by Type of Traveller')
plt.xlabel('Type of Traveller')
plt.ylabel('Overall Rating')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

```
plt.figure(figsize=(10, 5))
sns.countplot(data=final_df, x='Class', palette='muted')
plt.title('Distribution of Travellers by Class')
plt.xlabel('Class')
plt.ylabel('Number of Reviews')
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 8))
corr = final_df[['Overall Rating', 'Seat Comfort', 'Staff Service', 'Food & Beverages', 'Inflight Entertainment', 'Value For Money']].corr()
sns.heatmap(corr, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title('Correlation Between Different Rating Categories')
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 5))
sns.boxplot(data=final_df, x='Inflight Entertainment', y='Recommended_Binary', palette='Blues')
plt.title('Inflight Entertainment Rating vs. Recommendation Status')
plt.xlabel('Inflight Entertainment Rating')
plt.ylabel('Recommended (Binary)')
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 6))
sns.boxplot(data=final_df, x='Class', y='Value For Money', palette='cool')
plt.title('Value For Money Ratings by Class')
plt.xlabel('Class')
plt.ylabel('Value For Money Rating')
plt.tight_layout()
plt.show()
```

```
recommended_percent = final_df['Recommended_Binary'].value_counts(normalize=True) * 100
plt.figure(figsize=(8, 6))
recommended_percent.plot(kind='pie', autopct='%1.1f%%', labels=['Recommended', 'Not Recommended'], colors=['green', 'red'], wedgeprops={'edgecolor': 'black'})
plt.title('Distribution of Recommended vs Not Recommended')
plt.ylabel('')
plt.tight_layout()
plt.show()

distinct_reviews = final_df['Reviews'].nunique()
print(f"Total Distinct Reviews: {distinct_reviews}")

final_df['Review Length'] = final_df['Reviews'].astype(str).apply(len)
average_length = final_df['Review Length'].mean()
print(f"Average Review Length: {average_length:.2f} characters")

print(final_df[['Reviews', 'Recommended', 'Overall Rating']].describe(include='all'))
```

## Feature Extraction ⬇

```
final_df = pd.read_csv('final_airline_reviews.csv')

tfidf_vectorizer = TfidfVectorizer(max_features=5000)
X = tfidf_vectorizer.fit_transform(final_df['Cleaned_Reviews'])

y = final_df['Recommended_Binary']

print("Shape of TF-IDF Features:", X.shape)
print("Sample Target Values:", y.head())
```

## Data Split and Model Training ⬇

```
#Splitting the data and reserving 20% for testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Logistic Regression
logistic_model = LogisticRegression(max_iter=1000)
logistic_model.fit(X_train, y_train)

# Multinomial Naive Bayes
naive_bayes_model = MultinomialNB()
naive_bayes_model.fit(X_train, y_train)

# Model Evaluation

y_pred_logistic = logistic_model.predict(X_test)
print("Logistic Regression Results:")
print(f'Accuracy: {accuracy_score(y_test, y_pred_logistic):.2f}')
print(f'Precision: {precision_score(y_test, y_pred_logistic):.2f}')
print(f'Recall: {recall_score(y_test, y_pred_logistic):.2f}')
print(f'F1 Score: {f1_score(y_test, y_pred_logistic):.2f}')
print("\nClassification Report:")
print(classification_report(y_test, y_pred_logistic))

cm_logistic = confusion_matrix(y_test, y_pred_logistic)
plt.figure(figsize=(8, 6))
sns.heatmap(cm_logistic, annot=True, fmt='d', cmap='Blues')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix - Logistic Regression')
plt.show()
```

```
y_pred_nb = naive_bayes_model.predict(X_test)
print("Multinomial Naive Bayes Results:")
print(f'Accuracy: {accuracy_score(y_test, y_pred_nb):.2f}')
print(f'Precision: {precision_score(y_test, y_pred_nb):.2f}')
print(f'Recall: {recall_score(y_test, y_pred_nb):.2f}')
print(f'F1 Score: {f1_score(y_test, y_pred_nb):.2f}')
print("\nClassification Report:")
print(classification_report(y_test, y_pred_nb))

cm_nb = confusion_matrix(y_test, y_pred_nb)
plt.figure(figsize=(8, 6))
sns.heatmap(cm_nb, annot=True, fmt='d', cmap='Greens')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix - Naive Bayes')
plt.show()
```

## Hyper-Parameter Tuning ⬇

```
param_grid = {'C': [0.01, 0.1, 1, 10, 100], 'solver': ['lbfgs', 'liblinear']}
grid = GridSearchCV(LogisticRegression(max_iter=1000), param_grid, cv=5, scoring='accuracy')
grid.fit(X_train, y_train)

logistic_model_tuned = grid.best_estimator_

y_pred = logistic_model_tuned.predict(X_test)
print("Tuned Logistic Regression Results:")
print(f'Accuracy: {accuracy_score(y_test, y_pred):.2f}')
print(f'Precision: {precision_score(y_test, y_pred):.2f}')
print(f'Recall: {recall_score(y_test, y_pred):.2f}')
print(f'F1 Score: {f1_score(y_test, y_pred):.2f}')
print("\nClassification Report:")
print(classification_report(y_test, y_pred))

cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(8, 6))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix - Tuned Logistic Regression')
plt.show()
```

```
# Model Interpretation with LIME
explainer = LimeTextExplainer(class_names=['Not Recommended', 'Recommended'])
idx = 0
sample_review = final_df.iloc[X_test.indices[idx]]['Cleaned_Reviews']

def predict_proba_for_lime(texts):
    return logistic_model_tuned.predict_proba(tfidf_vectorizer.transform(texts))

explanation = explainer.explain_instance(sample_review, predict_proba_for_lime, num_features=10, labels=[0, 1])
explanation.show_in_notebook()


joblib.dump(logistic_model_tuned, 'sentiment_analysis_model.pkl')
print("Model saved successfully as 'sentiment_analysis_model.pkl'")
```