CST3990 Undergraduate Individual Project
2024-2025

Analysing the Impact of Airline Service Factors on Passenger Well-being Using NLP

Student Name: Chuar Kia Yi

Student ID: M01022086

Supervisor: Dr Olugbenga Oluwagbemi

Deadline for Submission: 13th April 2025

Declaration

I, Chuar Kia Yi, hereby certify that all of the work presented in this report and all other related materials are my own. The information derived from the literature has been appropriately cited in the text, and a list of references has been supplied. No section of this dissertation has been presented previously for another degree or diploma at this or another institution.

Signature Date: 10 April 2025

Abstract

This study investigates the relationship between airline service factors and passenger well-being by using various Natural Language Processing (NLP) techniques. The aim is to bridge the gap between airline service performance and its impact on passenger health by focusing on key factors like seat comfort, food and drink, staff and service, and flight experience. The study utilises British Airways passenger reviews and applies Latent Dirichlet Allocation (LDA) for topic modelling, Aspect-Based Sentiment Analysis (ABSA) using VADER, and emotion detection with DistilBERT. The study identifies how different airline service factors affect passengers' dietary, emotional, mental, and physical body well-being. The findings indicate that uncomfortable seats and flight delays both have a negative impact on physical body and mental health. The emotional analysis findings show that anger and joy are the most common emotions associated with various service aspects. This study not only demonstrates the value of sentiment and emotion analysis in evaluating passenger health but also provides informed decisions for enhancing airline services. The findings indicate that airlines may improve both customer satisfaction and well-being by addressing key service-related issues. Lastly, this study provides a foundation for future investigations into the broader influence of service quality on health in other industries.

Acknowledgements

I would like to express my gratitude to my dissertation supervisor, Dr. Olugbenga Oluwagbemi, who provides me with helpful guidance and constructive feedback when I am performing my research. I am also grateful for his mentorship which has helped me develop a deeper understanding of this subject. I would like to express my appreciation to my family, friends and colleagues for their unconditional support and compassion throughout my dissertation writing.

Table of Contents

List of Figures

List of Tables

1 Background of Study

1.1 Introduction

Nowadays, millions of people travel throughout the world by airplane and it has become the primary source of transportation. As passengers spend more time in the air, their overall well-being and travel experience have become important topics to pay attention. Most airlines are evaluated not only on the safety and operational efficiency but also on the quality of services across different areas. For illustration, there are some crucial aspects like service and staff, flight experience, food and drink, and seat comfort that help in shaping passenger satisfaction. However, the impact of these service aspects on passenger well-being is often ignored by the airline industry. This research will examine how these service aspects affect four kinds of well-being, which involve physical health, mental health, emotional health and dietary health by using advanced text analysis techniques.

The rise of Natural Language Processing (NLP) enables the extraction of meaningful insights from passenger reviews to better understand how these factors affect human well-being. NLP is a field of artificial intelligence that enables computers to understand and interpret human language by using machine learning. It also allows researchers to discover different trends from the unstructured text reviews to better understand customer preferences. This project utilises Latent Dirichlet Allocation (LDA) which applies unsupervised learning on the large dataset to identify hidden topics in text. In addition, Aspect-Based Sentiment Analysis (ABSA) has been used to determine the positive and negative sentiments towards different airline service factors. The results are mapped to different passenger well-being dimensions to evaluate how they affect passenger health and overall well-being.

This research is intended for airline industry stakeholders and researchers who are seeking to understand the relationship between airline service factors and passenger well-being. By applying different NLP techniques, this research aims to generate meaningful insights into how airline service factors influence passenger well-being, while also supporting airlines in enhancing passenger travel experience and addressing the existing problems.

1.2 Problem Statement

The purpose of this study is to explore how airline service factors impact passenger well-being. Specifically, this study will address the following research questions:

1. What are the key airline service factors frequently mentioned by passengers in the reviews?
2. How can these identified airline service factors be mapped to different dimensions of passenger well-being?
3. How do these airline service factors impact passenger well-being based on sentiment analysis and root cause analysis?

Airlines often struggle to extract meaningful insights as the huge amount of online passengers' reviews is in an unstructured and messy form. Traditional sentiment analysis approaches categorise consumer comments as positive, negative, or neutral without identifying the root causes or specific services. This makes it more difficult for airlines to find areas for improvement. Furthermore, these strategies may not be efficient in dealing with the subtleties and context-specific terminology found in customer evaluations, which may lead to misunderstanding. Gupta et al. (2024) found that standard sentiment analysis struggles with comments containing both positive and negative parts, as it tends to classify only one side instead of providing a balanced view. British Airways believes that understanding consumer opinion in depth is critical to boosting passenger pleasure. Seat comfort, flight experience, food & drink, and staff service may all have a negative impact on consumer loyalty and an airline's image. However, many previous studies have just examined general consumer attitudes without delving into individual service areas that generate frustration. Furthermore, the sheer volume of unstructured consumer feedback from various sources makes it impossible to organise and assess properly, resulting in inefficient data management and costly storage.

1.3 Aim and Objectives

The aim of this research is to analyse how airline service factors impact the passenger well-being by using Natural Language Processing technique. To achieve aim, the objectives which have been set for the work are as follows:

1. To identify and extract the key airline service factors from passenger reviews using topic modelling.

2. To map the extracted airline service factors to different passenger well-being dimensions.

3. To evaluate how the identified service aspects affect different dimensions of passenger well-being.

1.4 Report Structure

Chapter 1

Chapter 1 will be the introduction and background of the study. It shows the general overview of the topic. It mentions about the problems and aims of the study and covers about the industry background of this topic, which will provide a better foundation for another researcher. Additionally, it explains the key concepts of the topic, like text preprocessing, sentiment analysis, and topic modelling which are very crucial for understanding how the study analyses passenger reviews. It likewise discusses the current issues in the industry related to passenger satisfaction and health by setting the stage for the research focus.

Chapter 2

Chapter 2 is literature review that examines the existing studies related to the analysis of airline passenger reviews by using different methods like preprocessing techniques, sentiment analysis, and topic modelling. It also discusses how previous studies implement the natural language processing (NLP) techniques to gain meaningful insights from customer feedback. In addition, this chapter explores the research on health and well-being in air travel and finally shows the gap presented in existing studies.

Chapter 3

This chapter describes the methodologies and tools utilised to conduct the natural language processing (NLP) analysis. It discusses the data preprocessing methods, topic modelling with Latent Dirichlet Allocation (LDA), Aspect-Based Sentiment Analysis with VADER, and emotion categorisation using DistilBERT. Then, it demonstrates the mapping technique between the airline service factors and health categories, followed by the root cause analysis using bigram and Word2vec. Lastly, Data visualization approaches were discussed to show how the results are interpreted.

Chapter 4

This chapter covers the research analysis and results. It demonstrates the topics found by LDA and Word Cloud, as well as sentiment and emotion distribution across different airline service factors. The chapter also shows the visualisations that reveal the crucial trends and insights, such as how negative sentiment and emotional response vary across different service factors. In addition, it presents findings from the root cause analysis that was conducted through bigram analysis and semantic expansion techniques.

Chapter 5

Chapter 5 provides an evaluation and discussion of the research findings. It links the results to the research objectives and existing literature. This chapter delves further into the root causes of health impacts and emotional distribution across various airline service factors. It also outlines the implications for both the airline industry and passenger health. Lastly, the limitations of the study have been specified.

Chapter 6

The final chapter involves the conclusion and recommendations. It covers the primary findings of the study and demonstrates how airline service factors are connected to different dimensions of well-being. This chapter also makes recommendations for future research and technique enhancement suggestions. Practical recommendations are also proposed to airlines and policymakers to improve not only their service quality but also support passenger health.

2 Literature Review

2.1 Industry Background

In the aviation industry, customer satisfaction is crucial for determining an airline's reputation, customer loyalty, and competitive edge. A positive passenger experience leads to loyalty and word-of-mouth recommendations, while dissatisfaction can damage an airline's reputation and market share. As Lestari and Murjito (2020) point out, growing competition has drastically squeezed profit margins, leading to significant financial losses for some airlines. Consequently, airlines like British Airways have increasingly turned to passenger feedback as a means of evaluating service quality, improving operational efficiency, and identifying areas for improvement. Customer satisfaction has become the key performance indicator (KPI) in all airlines, most people consider this as an essential driver of long-term success in the industry.

Historically, airlines gathered passenger feedback through post-flight surveys and feedback forms, focusing on customer satisfaction and operational shortcomings. However, these methods often suffer from low response rates and biases, limiting their effectiveness in capturing the full range of passenger sentiment. Moreover, the structured nature of these surveys, such as rating scales and multiple-choice questions, restricts the ability to uncover unexpected issues. Although the feedback is meaningful, it did not fully reflect the complex sentiments expressed by passengers, especially in terms of dissatisfaction with different service sectors.

The rise of online review platforms and social media, such as Skytrax, TripAdvisor, and Expedia, has revolutionized passenger feedback. Passengers now share their experiences freely, providing valuable insights across various services (Ban and Kim, 2019). These online platforms have democratized the feedback process by allowing users to voice their opinions publicly and directly to the airline industry. The transition from structured to unstructured feedback creates a rich but complicated source of qualitative data. Despite the abundance of reviews, the sheer volume and unstructured nature of the data present significant challenges for manual analysis, highlighting the need for advanced computational methods to process and analyse this information effectively. As a result, this research aims to use sentiment analysis and natural language processing to find important insights from the vast amount of airline passenger reviews.

2.2 Foundation of Natural Language Processing (NLP)

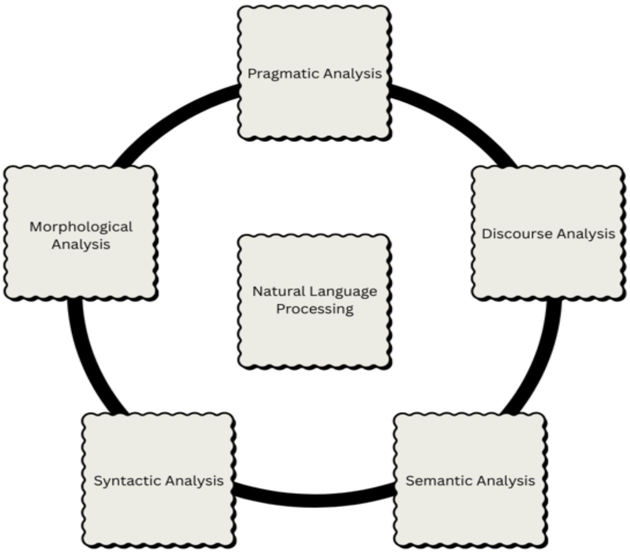

Figure 1: Overview of NLP Workflow

Natural Language Processing (NLP) is a set of rules where symbols are combined and used for transmitting information. Since some people may not master the machine language, NLP offers people who don't have enough time to learn and apply new languages (Khurana et al., 2022). With the rising need for machines to interact with people, NLP has become an essential tool for humans to utilise across various industries and applications.

In the past few years, the implementation of deep learning techniques in NLP has improved the research performance. There are a lot of examples that collaborate with NLP such as Internet of Things (IoT) devices and virtual assistants like Cortana, Siri and Google Assistant, which can normally be found in homes and offices. These systems heavily rely on NLP to handle voice commands and provide a real-time response accurately.

In addition, NLP can also be utilised for sentiment analysis, which plays a crucial role in the business intelligence sector. By implementing sentiment analysis, it can identify the customers' opinions automatically, enabling organizations to make informed decisions based on the results of NLP. For the airline business sector, sentiment analysis can help to analyse passenger feedback and make the right decision for improvement. However, the vast volume of customer reviews on social media makes it impossible for the company to manually process this data because it would be time-consuming.

According to Johri et al. (2021), there are five components for NLP:

1. Morphological Analysis: It is used to examine the nature of words by different morphemes like prefixes and suffixes, which is essential for further syntactic and semantic analysis.
2. Syntactic Analysis: It has a main step which is Parsing that determines the syntactic structure of the sentence. It builds a parse tree where the sentence acts as the root, intermediate nodes will represent nouns or verbs, and leaves will represent the individual words.
3. Semantic Analysis: It deduces the meaning of words and sentences by semantic validation and word sense disambiguation to determine the correct sense of the words.
4. Discourse Analysis: It determines the body structure of the text to maintain coherence and context.
5. Pragmatic Analysis: It involves three tasks which are reference resolution, discourse analysis, and dialog interpretation that interprets information in a specific text and takes into account external knowledge.

These components will work together to ensure that machines are able to understand the complexity of human language. This is because challenges like ambiguous grammar and different regional slang make NLP very complex to understand but fascinate people to study it. A previous study by Jim et al. (2024) has pointed out that despite significant advances in NLP, challenges like context understanding, bias and domain adaptation still exist. This research has highlighted the potential for further improvement, especially in the context of context understanding in phrases and words.

2.3 Sentiment Analysis in the Airline Industry

In today's world, the airline industry relies completely on customer feedback to improve service quality, operational performance, and brand loyalty. In contrast to conventional information gathering techniques, the introduction of internet review platforms such as Skytrax, TripAdvisor, and Google Reviews allows passengers to express their thoughts in real time. However, the massive and unstructured information makes it hard to be analysed manually. Hence, Natural Language Processing (NLP) techniques have emerged as valuable tools in airline research for better decision making by categorising passengers’ concerns as positive, neutral, and negative sentiments (Srinivas and Ramachandiran, 2023).

There are some fresh publications have highlighted the importance of sentiment analysis in evaluating airline service quality. To illustrate this, Patel et al. (2023) analysed airline passenger evaluations using several categorisation models, including Naïve Bayes, Support Vector Machines (SVM), and Decision Trees, and they discovered some dissatisfaction patterns with luggage, flight delays, and in-flight service. This suggests that sentiment analysis not only captures surface-level satisfaction, but also can explore deeper for service-related concerns. Moreover, Srinivas and Ramachandiran (2023) have used an unsupervised text analytics method to identify different airline service issues by demonstrating how sentiment analysis helps airlines better understand passengers' difficulties across different service factors.

Furthermore, sentiment analysis has been used to assess the success of airline brands as they are more emphasizing on customer satisfaction and service quality. According to Shrestha et al. (2024), they have conducted sentiment analysis on British Airways passengers' evaluations but mostly focusing on the numerical rating scores. The research did not evaluate the review contents, and this has limited their capability to understand deeper passengers' opinions. This highlights a gap in existing studies as passengers' contextual reviews could provide more important information compared to numerical evaluations.

Although traditional machine learning played as a significant role in airline sentiment analysis, recent deep learning approaches have increased the accuracy of sentiment analysis to a new height. According to Patel et al.'s (2023) study, transformer-based models like BERT and RoBERTa outperform traditional models as they better capture the contextual value of passengers' feedback. However, most current research primarily categorises sentiments rather than gathering insights from different aspects. Moreover, another study only utilised topic modelling to analyse airline ratings but did not perform ABSA for discovering sentiment across different aspects (Srinivas and Ramachandiran, 2023).

2.4 Review of Preprocessing and Extraction Methods in Airline Review Analysis

Regarding analysing airline passenger reviews, text preprocessing is a crucial step to enhance the quality of textual data by removing noise and transforming raw data into a standardized format. It is necessary to conduct preprocessing for accurate sentiment analysis as passengers' reviews often include informal language, acronyms, and misspellings. There are some common preprocessing techniques like text normalization, stopword removal, tokenization, stemming, and lemmatization, which can be used to improve natural language processing. They can handle different languages, dialects, writing styles and can be trained for specific areas to enhance accuracy and reduce processing time (Abubakar Ahmad Aliero et al., 2023).

Lemmatization and stemming are very important methods for simplifying words to their most basic forms. Stemming helps improve the textual consistency by reducing words to root form, while lemmatization only ensures the output word becomes a dictionary-based form (Khyani et al., 2021). Stemming is faster and more efficient as it utilises fixed rules and string manipulation, whereas lemmatization is slower as it employs dictionary lookups and more complex algorithms. For illustration, lemmatization is more useful while conducting online reviews analysis to preserve the meaning of domain-specific terms (e.g. “delay”, “delays”, “delayed”), which might be crucial themes related to service aspects.

Additionally, feature extraction is another crucial step in sentiment analysis as it converts text into numerical representations that the machine can understand. Traditional techniques like Term Frequency-Inverse Document Frequency (TF-IDF) identify the value of words by comparing their frequency of appearance in documents to the entire text collection. However, TF-IDF is very limited for deep sentiment analysis as it does not capture semantic relationships between words (Zhan, 2025). For example, phrases like “The VIP lounge is very comfortable, but there is a limited number of food choices” may confuse the TF-IDF as it treats each word independently and cannot capture the meaning of “but”, which may lead to misinterpretation. According to Semary et al. (2024), they stated that word embeddings could address this issue by mapping words into high-dimensional frameworks, and the popular examples are Word2Vec and GloVe. TF-IDF is more suitable for a smaller database as it focuses on word frequency, while word embeddings are suitable for a larger scale as they capture contextual meaning.

In recent research, deep learning-based approaches like BERT and RoBERTa have been widely used for feature extraction. They offered contextualised word representation that significantly increased the sentiment classification accuracy (Semary et al., 2024). A huge amount of labelled data and processing resources are required for these models to achieve the optimal performance.

2.5 Enhancing Sentiment Classification Through Aspect-Based Sentiment Analysis

Aspect-Based Sentiment Analysis (ABSA) is a sentiment analysis technique that identifies and categorises opinions regarding different aspects of the writing content. In contrast to traditional sentiment analysis that only provides results like positive, neutral and negative, ABSA offers more precise information by linking emotions to specific aspects or qualities (Georgios Kontonatsios et al., 2023). This method can be beneficial when analysing passengers' reviews since a single review might provide different opinions on many aspects of a product or service.

In recent studies, many researchers have explored different ways to enhance ABSA. For example, Chang et al. (2022) stated that there are three essential steps to enhance it, which include extracting aspects, calculating sentiment and aggregating sentiment scores. Their findings showed that deep learning models like BERT-based can improve ABSA accuracy by better recognising more context. Another research conducted by Gu et al. (2024) evaluated SVM and Naïve Bayes to advanced deep learning models like LSTM networks and Transformers. They found out that deep learning models outperform other models as they can detect more complex speech patterns.

In addition, Graph-based algorithms are also a new way to improve sentiment categorisation accuracy. Gu et al. (2024) introduced a syntax-aware ABSA model by using Graph Convolutional Networks (GCNs) to improve the accuracy of sentiment predictions. Their findings suggested that including grammatical dependencies in aspect-based sentiment categorisation enhances performance, especially for detecting implicit sentiment.

In Nazir et al. (2020) research paper, they have identified several issues faced by ABSA such as inconsistencies in aspect extraction, handling of sarcasm, and the need for huge datasets with annotations. Furthermore, their research highlighted how domain-specific limits ABSA's generalisability and requires adjustments for many applications. Manual annotation of aspect-based sentiment data is still extremely demanding on resources and this required more automated and unsupervised learning systems to be developed (Kumar et al., 2023).

2.6 Latent Dirichlet Allocation (LDA) in Sentiment Analysis

Latent Dirichlet Allocation (LDA) is a topic modelling and unsupervised statistical framework that is used to identify hidden topics in a huge text dataset. LDA helps discover potential topics by examining the word distribution inside the text data, which is completely different compared to traditional sentiment analysis that only categorises text (Rana et al., 2016). Therefore, LDA is an important tool for extracting unstructured customer evaluations and categorising them to a variety of themes.

According to Çallı and Çallı (2022), they utilised LDA to analyse the airline passengers' complaints during the COVID-19 pandemic by categorising them into different categories such as flight ticket refund, baggage delivery, and seat convenience. Their research has demonstrated how LDA assists the airline in identifying major service concerns and making service changes by clustering similar complaints in a category. Moreover, Farkhod et al. (2021) proposed the Topic/Document/Sentence (TDS) model that combines with LDA topic modelling to improve the sentiment classification accuracy. They have mentioned that LDA can perform better when combined with sentiment classification models for topic identification.

Furthermore, LDA was also carried out by the researcher to evaluate product performance and customer feedback. Farrikh Alzami et al. (2023) have investigated social media products reviews by using LDA-based topic modelling and demonstrated how LDA identify significant topics from large amounts of unstructured reviews data. They also mentioned that preprocessing techniques are very important in enhancing the LDA topic extraction accuracy, especially the stopword removal and stemming.

Nevertheless, the major problem faced by LDA is its inability to capture the sentiment polarity automatically and this has emphasised the demand for additional sentiment classification models for further analysis (Rana et al., 2016). The quality of LDA-derived topics can be affected by parameter adjustments such as the number of topics and word distributions inside the textual data.

2.7 Exploring Health and Well-Being in the Context of Air Travel

Most research in the airline industry has only focused on its safety, service efficiency and customer satisfaction, but recent studies have started to explore how air travel affects passengers' health and well-being. Although this field of study is still in its early stages, there are still some research shows that various airline factors can have a significant influence on passengers' physical, mental, and emotional health.

According to Nicholson et al study. (2003), they have investigated many medical issues related to air travel and highlighted the key concerns about cabin conditions and long flight durations. They found out that the low humidity, reduced oxygen tension, and limited mobility in the airplane can affect human health conditions, especially for those passengers who already have existing respiratory problems. They also discussed the danger of deep vein thrombosis (DVT) and neurological symptoms possibly linked to prolonged inactivity and exposure to toxins like tri-ortho-cresyl phosphate (TOCP), which is a neurotoxic motor oil that is easily found in cabin air.

In addition, Clapp et al. (2021) investigated the effect of physical seating comfort on psychological results. They found out that seat discomfort may affect humans to achieve psychological flow, which is linked with deep attention and well-being. Although the research focused on office environments, the findings can be highly relatable to airline settings. They mentioned that poor seating conditions may lead to fatigue, discomfort and affect both physical and mental well-being.

Furthermore, environmental noise will also influence health aspects. In Benz et al. (2022) study, they pointed out that continuous exposure to aircraft noise was related to an increased risk of cardiovascular and psychological problems among humans who live near airports. The findings highlighted how environmental stresses associated with aviation can affect both physical and mental health. Although this study focuses on ground-level exposure, it still provided important insights that can be applied to flight experience as well.

As a result, these studies show that air quality, seat ergonomics, and environmental stressors will affect not only the comfort issues but also human health risks. However, there is a significant gap in linking the online passengers' feedback to the health and well-being. This study will demonstrate how passenger reviews can be analysed through NLP and discover possible effects on different well-being. It also investigates the root cause of complaints to identify patterns that may not yet be reflected in existing studies.

2.8 Data Visualization for Sentiment Analysis

Data visualization is crucial for sentiment analysis as it displays the results in a clearer and more interactive format that humans can understand easily. These visual display enables companies to identify trends and make brilliant decisions through bar charts, plots, and heatmaps (Ameni Dhaoui Boumaiza, 2016). These techniques can highlight the key insights effectively in a straightforward and visually engaging way.

There is a popular tool used for visualising sentiment analysis which is R programming. It employs the Python codes with packages like ggplot2 to draft sentiment patterns in creative ways (Shara, 2021). The dashboard in R programming allows users to analyse real-time data and track sentiment changes dynamically. In research conducted by Shara (2021), she visualised the sentiment polarity by using different histograms and time-series charts, which make it easier to evaluate customers' sentiment changes over time.

In the aviation industry, sentiment visualisation has been widely used to determine their service quality and customer satisfaction. In Samah et al. (2022) research, they have investigated multilingual sentiment analysis of Malaysian airline businesses by using machine-learning based sentiment methods. They have created a website which is used to demonstrate the visualised airline sentiment trends and emphasize the airline performance into various categories like customer service reports, ticket price, and in-flight experiences. This study revealed how sentiment dashboards assist airlines in tracking passenger major complaints and adjusting services according to customers' feedback.

2.9 Challenges and Limitations in Existing Studies

It is noted that sentiment analysis is widely used for various research in the current dataset, however, it is yet to be able to fulfil all the particular extent. For example, many studies focus on general sentiment, categorizing feedback into limited classifications such as positive, negative, or neutral. This approach overlooks the details that could provide deeper insights.

According to research, numerical ratings may not accurately reflect passenger opinion since they rarely covered critical service areas such as check-in, in-flight experience, and customer assistance (Chang et al., 2022). This emphasises the need for more study in Aspect-Based Sentiment Analysis (ABSA) to acquire a more in-depth knowledge of passenger experiences.

According to research by Chang et al. (2022), sentiment analysis has a significant gap in cross-cultural and multilingual analysis. Most sentiment analysis methods tend to focus primarily on English-language evaluations, which limits the ability to generalize the findings to other passenger groups. This creates a barrier for understanding sentiment from a diverse, global audience. To address this, future research should prioritize the development of multilingual sentiment analysis algorithms that can accurately process and interpret evaluations from a broader range of consumers, allowing for more inclusive and globally relevant insights.

Another important area for future research is combining Aspect-Based Sentiment Analysis (ABSA) with data visualization techniques. While sentiment analysis can reveal insights into customer attitudes, the value of these insights depends on how easily they can be interpreted. Samah et al. (2022) highlighted that visualizing sentiment patterns can enhance decision-making in the aviation industry. However, most recent studies have not explored interactive visualization methods for ABSA results, which has limited the accessibility and practical use of these insights for business stakeholders.

Moreover, current ABSA models still struggle with understanding more complex elements of language, such as sarcasm, denial, and subtle emotional cues (Raghunathan and Saravanakumar, 2023). These challenges lead to less accurate sentiment predictions, which makes it clear that there’s a need for improvements in deep learning models. To truly reflect the complexities of real-world conversations, these models will need to be fine-tuned to better capture these nuances. Another issue holding back sentiment analysis in the airline industry is the shortage of publicly available ABSA datasets. Gupta et al. (2024) point out that having well-annotated datasets that cover a wide range of aspects of airline services is crucial. Without these, it’s difficult to train and evaluate models that can accurately capture customer sentiment and improve the decision-making process for businesses.

3 Methodology

3.1 Overview of Methodology

This chapter introduces the methodology that has been utilised to help achieve the research objectives successfully. The methodology of this paper is to analyse the British Airways airline review dataset through a series of NLP techniques, which include text preprocessing, bigram extraction, topic modelling, aspect-based sentiment analysis and emotion detection. This chapter is mainly to introduce the methodology used in this research to achieve the first and second objectives: (1) To identify and extract the key airline service factors from passenger reviews using topic modelling, and (2) To map the extracted airline service factors to different passenger well-being dimensions.

3.2 Dataset Description and Text Preprocessing

The dataset used in this research is the British Airways Review Dataset (2012-2023) which was obtained from Kaggle. It contains a total of 1324 passenger reviews related to British Airways flights. Each review includes multiple attributes like cabin class, route, ratings and open-ended feedback. For this research, only two columns were analysed which are the date of the review and review content.

To streamline the analysis, data cleaning and text preprocessing have been carried out to remove the unnecessary data and retain only the textual content. A total of 17 irrelevant columns have been dropped from the original dataset as they did not contribute any meaningful content to the text-based analysis. The goal of preprocessing was to ensure the final data was clean and consistent, which would be suitable for further natural language processing tasks.

After removing the unnecessary columns, a full text preparation pipeline was established. Figure 2 illustrates the sequential steps taken during the preprocessing steps, which include various techniques like normalization, lemmatization, tokenization and the list goes on.

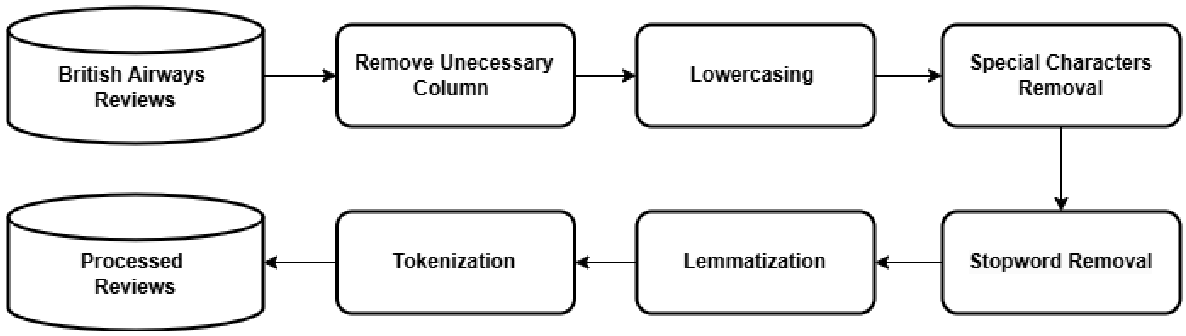

Figure 2: Preprocessing pipeline

The individual steps in the preprocessing pipeline are shown below:

- Lowercasing: Convert text or letters to their lowercase equivalents to maintain consistency, which can be helpful in data analysis.
- Special Characters Removal: Removes the punctuation and non-alphabetic symbols that will only produce noise in the data.
- Stopword Removal: Eliminates common stop words like “the”, “and”, and “we” by using a predefined stopwords list in the corpus. Customized domain stopwords have also been applied in this process as those words are typically non-informative for the topic modelling and sentiment analysis.
- Lemmatization: By importing “WordNetLemmatizer” in Python, it will loop through the WordNet lexical database to find the base form of each word, and this will enable more accurate results for topic modelling. For example, “delayed” and “delays” will be converted to “delay”, which is the original form.
- Tokenization: By using the built-in NLTK’s word_tokenize function, it splits the passenger reviews into individual words that are referred to as tokens. This will enable the content to be ready for vectorization and model input.

3.3 Bigram Detection

In this section, Bigram detection has been applied to identify the word pairs that indicate the specific passenger experiences by using the Pointwise Mutual Information (PMI) approach. This step will identify meaningful two-word phrases that occur together more frequently, and this will help to enhance the semantic representation of the text. The Phrases model from the Gensim library has been utilised to train the tokenized review corpus with a minimum occurrence threshold of 5 and a PMI threshold of 10. This will ensure that only important bigrams will be retained and other tokens with remain as individual words.

After the model was trained, it would loop through the tokenized reviews and replace the sequences of frequently occurring terms with their combined bigram form. To illustrate this, separate tokens like “flight” and “delay” were merged as a single token which is “flight_delay”. These changes will improve the interpretability of subsequent analysis as they are in meaningful word pairs. Figure 3 demonstrates the bigram detection technique in code form.

```
def build_phrases(tokenized_reviews):
    bigram = Phrases(tokenized_reviews, min_count=5, threshold=10)
    bigram_model = Phraser(bigram)
    tokenized_reviews = [bigram_model[review] for review in tokenized_reviews]
    return tokenized_reviews
```

Figure 3: Bigram Detection

3.4 Applying LDA and Word Cloud for Topic Modelling

After employing data preprocessing and bigram build phases, topic modelling was conducted using Latent Dirichlet Allocation (LDA) and Word Cloud visualisation to provide a quick overview of each topic. The process started with transforming the tokenized reviews into a Bag-of-Words (BoW) representation by using the corpus tool. To improve the accuracy of the model, a filter function has been applied to ensure that no rare words and common words which appeared more than 70% in the corpus.

Next step, the LDA model was trained using the Bow corpus to identify 5 distinct topics. The model has been configured with parameters like “passes = 30” which will loop through the dataset 30 times during training, while the “alpha = auto” and “eta = auto” are used to find how topics spread across reviews and how words are grouped within topics. The configuration will allow the model to better fit the current structure of the review data and provide better topic coherence. When the training is completed, the model will demonstrate the top 10 keywords for each topic.

In this research, two kinds of visualization were generated to display the LDA results. The first one was generated by using pyLDAvis which enables us to view how all topics are distributed and identify the most relevant terms for each topic. In addition, another interactive visualization technique has also been developed which is known as Word Cloud. In this process, it is going to extract the top 50 terms from each topic and generate the Word Cloud images.

3.5 Aspect-Based Sentiment Analysis Using VADER

After topic modelling, aspect-based sentiment analysis (ABSA) was conducted to identify the sentiments across different airline service factors. This approach will provide a more detailed sentiment categorisation since it can capture how passengers feel about different sectors of their flying experience.

First, the ABSA approach started by defining the domain-specific keyword dictionary under the four key factors: service_and_staff, flight_experience, food_and_drink, and seat_comfort. Each factor has been assigned a comprehensive list of terms which related to that specific aspect, such as “service”, “crew”, and “rude” were assigned under the service_and_staff. These keywords were chosen to guarantee that they align and cover various dimensions of the airline service factors. By utilising the rule-based approach, any predefined keywords found in the passenger review will be marked as that specified factor. This will ensure that one single review will have multi-aspect sentiment evaluation.

Next, the sentiment analysis tool VADER (Valence Aware Dictionary and Sentiment Reasoner) will be employed to calculate the sentiment polarity of every single passenger review. The reason I chose VADER is because of its efficiency and accuracy when dealing with informal text. The sentiment polarity score will be divided into three categories: positive (compound score $\geq$ 0.05), negative (compound score $\leq$ -0.05), and neutral (score in between). The sentiment label was applied to each factor for the passenger reviews.

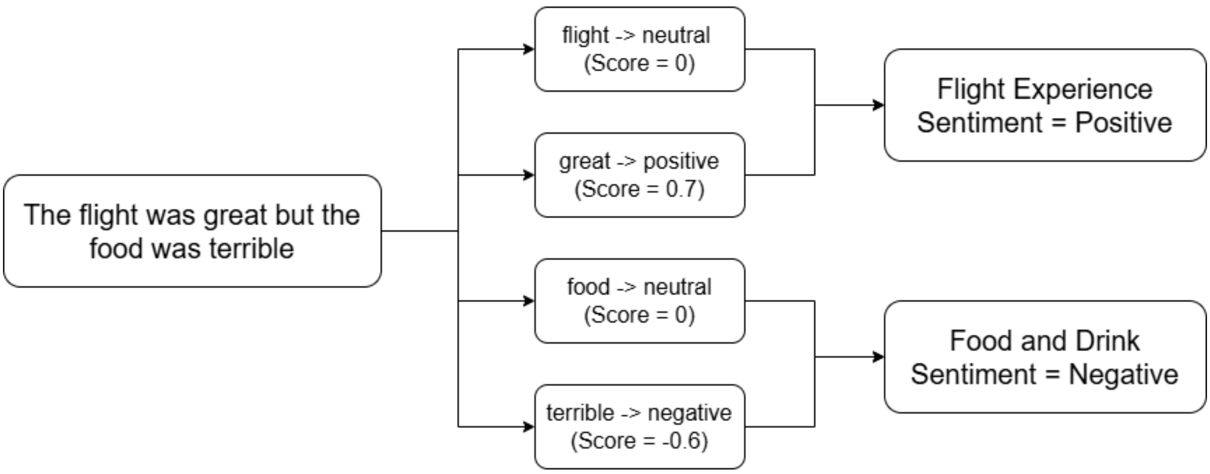

Figure 4:Example of Aspect-Based Sentiment Analysis

According to Figure 4, the sentence "The flight was great but the food was terrible" was calculated separately for different aspects. Since great possess a positive term in the sentiment lexicon, the result will be examined as positive sentiment. In contrast, the negative term 'terrible' contributes to the negative sentiment.

This process is used to produce a structured summary for each review that includes the relevant airline service factors associated with sentiment polarity. These results will be the foundation for the subsequent process, which includes a more detailed investigation of their relationship to passenger health aspects.

3.6 Emotional Analysis Using DistilBERT

The following step utilised emotional analysis to acquire a deeper understanding of the passenger's emotional states about different aspects of well-being. Unlike sentiment analysis, which only identifies the passenger sentiment as positive, negative and neutral, emotional analysis offers more detailed insights by categorising textual information into various emotions such as joy, anger, love, and the list goes on. By implementing this additional layer of analysis, it will enhance our understanding of how different aspects of the flight experience may affect passengers' health and well-being.

To analyse emotional analysis, a pre-trained transformer model DistilBERT was utilised via the "bhadresh-savani/distilbert-base-uncased-emotion" in the Hugging Face library. This model has been customised to perform emotion classification and recognise a variety of emotions expressed in natural language. The model was built using the Hugging Face pipeline with the configuration of truncation enabled at a maximum length of 512 tokens to ensure compatibility with the model's limitations.

```
emotion_classifier = pipeline(
    'sentiment-analysis',
    model='bhadresh-savani/distilbert-base-uncased-emotion',
    truncation=True,
    max_length=512
)

def detect_emotion(text):
    result = emotion_classifier(text)
    return result[0]['label']
```

Figure 5: Emotion Classifier function

According to Figure 5, it demonstrates how the emotion classifier works by passing each passenger review and labelling the predicted emotion. The output label will be used to indicate the emotion identified in the review text. The results were then combined with the aspect-based sentiment findings to investigate how different service factors affect passenger sentiment and trigger emotional responses among them.

By analysing the passengers' emotions in the entire workflow, the project aims to offer a deeper understanding of how passengers emotionally respond to different airline service aspects, rather than just determining their flight experience as positive or negative.

3.7 Mapping Airline Service Factors to Health Categories

In this phase, the sentiment and emotional results determined through the Aspect-based sentiment analysis and emotional analysis were mapped to particular health categories. This stage will discuss more about how different factors of airline service might affect passengers' health and well-being. The four health categories are identified as emotional health, mental health, dietary health, and physical body health. The mapping process was conducted based on the assumption that different aspects of the passenger experience might have a direct impact on their health and well-being. The following section will provide a detailed explanation of how each airline service factor was mapped to the corresponding health categories.

1. **Service and Staff → Emotional Health**

   The first airline service factor "service and staff" was mapped to the emotional health. To achieve this, keywords about customer service were extracted from the passenger reviews like "rude", "helpful", "polite", and "attentive" were identified as reflecting the passenger's emotional experience with the airline staff. To illustrate this, a review that mentions a "rude" crew member might trigger a negative emotional sentiment that affects a passenger's emotional health. In contrast, a review shows that helpful staff would contribute to positive emotional outcomes. Sentiments for this category were then classified as positive, neutral, and negative based on the tone and emotional response to the service.

2. **Flight Experience → Mental Health**

   The second airline service factor "flight experience" was mapped the mental health. The keywords for this category are related to flight experience, such as "delays", "boarding", "exhaustion", and "jetlag". These keywords were then linked with the mental health consequences. For example, passenger who describe their experience with delayed boarding may show signs of stress and frustration, which would negatively affect their mental health. For positive mental health outcomes, reviews like smooth flight experiences and timely arrivals are considered as indicators of a calm and stress-free journey. By defining these keywords, the airline service factor can be clearly linked to the mental health category.

3. **Food and Drink → Dietary Health**

   The following airline service factor is "food and drink", which will be linked to dietary health. The food-related experiences keywords in the reviews were analysed, such as "taste", "hygienic", "food poisoning" and "nutrition". These factors are strongly related to passengers' dietary health during the flight. For instance, reviews that describe "delicious" or "tasty" meals were linked to positive dietary health, while "food poisoning" and "lack of nutrition" were mapped to negative dietary health. Therefore, food-related experiences play a crucial role in shaping dietary health perceptions as factors like the variety of meals, food nutrition, and portion size would significantly affect the overall satisfaction and dissatisfaction of passengers.

4. **Seat Comfort → Physical Body Health**

   The last airline service factor "seat comfort" was mapped to the physical body health. This aspect focused on keywords related to physical discomfort such as "legroom", "seat recline", "comfort", and "cramped". These features are important indicators of passengers' physical body health during the flight. For illustration, reviews that mention "seat cramped" or "lack of legroom" often show passengers' discomfort, which can lead to physical body issues like back pain, leg cramps, and fatigue. On the other hand, positive comments regarding the comfortable seats and sufficient space were linked to positive physical body health. By linking these physical aspects to sentiment, the mapping process revealed how seat comfort affects passengers' physical body health.

After analysing the comments in the airline service factors, their corresponding sentiment polarity and emotional results were assigned to one of four health categories. This structured mapping will enable a better understanding of how various service factors, such as flight experience, seat comfort or service and staff affect mental, physical body and emotional health.

3.8 Root Cause Analysis (RCA)

Root Cause Analysis (RCA) is one of the most important steps in finding and comprehending the reasons for health-related complaints made by passengers in the airline reviews. This section will outline the approach for conducting RCA for these four health categories: mental health, physical health, emotional health, and dietary health. By using a combination of predefined bigrams and a Word2Vec model, the root causes can be finally identified with greater clarity and depth.

3.8.1 Health Aspect Identification

Before proceeding to the root cause analysis, predefined bigrams are necessary to group the common airline passenger concerns under their relevant health aspects. These bigrams are generated in the data preprocessing steps, which include the frequent two-word phrases and will be used to highlight the problems like dissatisfaction and negative sentiments in different aspects.

To illustrate this, Mental Health will be discussed in this section as an example. The bigrams were divided into subcategories to indicate distinct types of psychological distress. These include:

- Flight Delay Problem: Includes bigram phrases like "delayed_flight" and "missed_connection"

- Uncomfortable Flight: Includes bigram phrases like "cabin_temperature" and "long_flight"
- Negative Experience: Involves bigram phrases like "bad_experience" and "disappointing_experience"
- Noise Disturbance: Consists of bigram phrases like "loud_passengers" and "crying_baby"

Each group focusses on a different dimension of how service issues may negatively impact a passenger's mental state. Figure 6 illustrates the complete list of predefined bigrams for all health aspects.

```
predefined_health_root_causes = {
    'Mental Health': {
        'Flight Delay Problems': ['delayed_flight', 'flight_delay', 'flight_delayed_hour', 'late_departure', 'missed_connection'],
        'Uncomfortable Flight': ['uncomfortable_flight', 'long_flight', 'long_time', 'cabin_temperature', 'poor_air'],
        'Negative Experience': ['poor_experience', 'bad_experience', 'disappointing_experience'],
        'Noise Disturbance' : ['loud_passengers', 'crying_baby', 'engine_noise', 'noisy_cabin'],
    },
    'Physical Body Health': {
        'Uncomfortable Seat': ['uncomfortable_seat', 'seat_uncomfortable', 'hard_seat', 'seat_painful'],
        'Cramped Legroom': ['seat_cramped', 'cramped_seat', 'tight_space', 'no_legroom'],
        'Inadequate Sleeping Conditions': ['seat_recline', 'broken_recline', 'no_pillow','sleep_disrupted'],
        'Limited Movement': ['narrow_seat', 'aisle_blocked', 'restricted_movement', 'seatbelt_sign', 'long_sitting'],
    },
    'Emotional Health': {
        'Poor Service': ['poor_service', 'service_poor', 'service_slow', 'unresponsive_crew', 'ignored_request'],
        'Rude Staff': ['staff_rude', 'rude_attendant', 'impolite_crew', 'hostile_behavior'],
        'Lack of Empathy': ['no_apology', 'no_explanation', 'staff_uncaring', 'robotic_response'],
        'Boarding Chaos': ['boarding_mess', 'boarding_confusion', 'unclear_instructions', 'late_boarding'],
    },
    'Dietary Health': {
        'Bad Food Quality': ['food_poor', 'poor_food', 'bad_meal', 'cold_food', 'tasteless_meal'],
        'Limited Options': ['limited_food', 'no_vegetarian', 'no_meal_choice', 'meal_unavailable'],
        'Meal Timing Issues': ['late_meal', 'missed_meal', 'no_snack', 'meal_delayed'],
        'Unhygienic Food': ['dirty_utensils', 'spoiled_food', 'unclean_tray', 'unhygienic_conditions']
    }
}
```

Figure 6: Predefined Health Root Causes Across Health Aspects

3.8.2 Word2Vec Model

In this section, the Word2Vec model was trained using the tokenized review data. Word2Vec is a common approach for generating vector representations of words or phrases based on their context. It will create a "vector space" to classify similar words and place them together. For example, "rude_staff" was used to find its nearest vector, and one of the most similar results was "unprofessional_staff" based on their vector representations. Figure 7 shows the parameters used to train the Word2Vec model.

```
model = Word2Vec(sentences=tokenized_reviews, vector_size=100, window=5, min_count=2, workers=4)
```

Figure 7: Word2Vec Model Configuration Parameters

The Word2Vec model was configured with several parameters. First, the vector size will be set to 100 which is represented as a 100-dimensional vector. The window has been set to 5 which specifies the maximum distance between the present and predicted words in a phrase. To ensure the significant patterns, a minimum count of 2 has been configured to ensure that only bigrams that appear at least twice can be included. Lastly, the workers parameter was adjusted for the model to speed up the processing by 4 threads. This trained model will examine the predefined bigrams and identify other bigrams in the dataset that appear in similar contexts, allowing it to find semantically related phrases.

3.8.3 Bigram Expansion

This section will discuss the bigram expansion. After training the Word2Vec model, it used the discovered cosine similarity to locate semantically similar bigrams to the predefined root causes. Cosine similarity is used to calculate two vectors and determine how similar they are in terms of meaning. The function works as follows:

1. A vector representation is created for each predefined bigram
2. Cosine similarity is determined between the vector of the predefined bigram and the vector of all other bigrams in the dataset.
3. The bigram with the highest similarity to the predefined bigram is selected and displayed.

Therefore, this will be the result of the expanded list of bigrams correlated to each health aspect. This results in an expanded list of bigrams related to each health aspect. For instance, if “delayed_flight” is recognised as a root cause under the Mental Health category, the model will broaden it to include more related words like “late_departure”, and “flight_went”.

```
def find_similar_bigrams(seed_bigram):
    seed_vec = get_bigram_vector(seed_bigram)
    if seed_vec is None:
        return []

    similarities = []
    for bigram in bigram_df['Bigram'].unique():
        vec = get_bigram_vector(bigram)
        if vec is not None:
            sim = cosine_similarity([seed_vec], [vec])[0][0]
            similarities.append((bigram, sim))

    sorted_bigrams = sorted(similarities, key=lambda x: -x[1])
    return [bigram for bigram, sim in sorted_bigrams[:topn]]
```

Figure 8: Bigram Expansion Function Implementation

Figure 8 illustrates the implementation of the bigram expansion function. There are two functions to be mentioned which are “get_bigram_vector” and “cosine_similarity”. The first function is to retrieve the vector representation of a bigram and if both words are present in the Word2Vec model, their vectors are averaged. The second function is to compare the vector of predefined bigrams with other bigrams to find the most similar ones.

3.8.4 Expanded Root Causes and Analysis

After the expansion process, a more comprehensive list of bigrams related to each health aspect will be shown. The expanded list includes both the old bigrams and new bigrams which were identified through the semantic similarity. This enabled a more in-depth knowledge of passenger complaints regarding their health and well-being during the flight journey. These results will be saved in a dictionary for further analysis.

By integrating semantically similar bigrams, the research covered a greater range of health-related topics by offering a more comprehensive view of the factors affecting passenger health. This expanded list of bigrams ensures that the analysis reflects the complexities of passenger feedback, exposing new issues that may have an impact on their overall health and well-being. The analysed results will serve as a foundation for future visualization of these root causes that affect passenger health.

3.9 Data Visualization of Health and Sentiment Insights

This section will demonstrate how the visualization techniques are applied to provide a better understanding of the relationship between airline service factors, passenger sentiment, emotions, and the health impacts. There are a total of four visualizations that were created using the processed results.

1. Bar Chart of Health-Related Aspects Distribution

A grouped bar chart was created to show the distribution of good, neutral, and bad impacts across the four identified health categories: emotional health, mental health, dietary health, and physical body health. This chart will assist in visualizing how frequently each health aspect was mentioned in different impacts.

2. Emotion Distribution by Airline Service Factors

A grouped bar chart was generated to depict the distribution of detected emotions like joy, anger, fear, sadness and others across different airline service factors. The chart will provide a visual representation of the frequency of emotion within these factors that allows a better understanding of how different elements of flight experience contribute to passengers' emotional responses.

3. Root Cause Visualization per Health Category

For each health category, a separate bar chart was plotted to display the top root causes based on the bigram frequency analysis. These root causes were determined based on predefined bigrams, and their frequencies were calculated by aggregating occurrences in the dataset. These plotted graphs will provide a clearer understanding of the existing problems that passengers reflect during their flight experience.

4. Yearly Trends of Health-Related Sentiment

Line graphs were plotted to show how the sentiment trends changed over time for each health category. The review publishing date was extracted and grouped by years, which allows for temporal analysis of passenger feedback on health and well-being.

All interactive visualisations in this project were implemented using Python's Seaborn and Matplotlib tools. Every plotted graph has its annotations to enhance clarity. These visual tools are an essential component for the analysis pipeline because they provide both qualitative and quantitative insights into how airline service factors affect passengers' health and well-being.

4 Analysis and results

4.1 LDA Topic Modelling and Word Cloud Interpretation

In this section, the results of Latent Dirichlet Allocation (LDA) and Word Cloud visualization have been demonstrated to explore the main topic discussed in airline passenger reviews. Although the LDA model has been applied to discover hidden topics from the data, but the automatically generated data were not always clearly interpretable. Therefore, manual interpretation was conducted to make the topic more meaningful and clearer based on the representative keywords from each topic. Table 1 summarises the four refined topics through a combination of LDA output and keyword grouping.

Topic	**Top Keywords (LDA & Word Cloud)**
Service and Staff	staff, service, crew, customer service
Flight Experience	check, time, business class, passenger, long haul, experience, delay
Food and Drink	meal, food, drink, snack, breakfast, water
Seat Comfort	seat, leg, legroom, space, cramped, comfort

Table 1: Results of Topic Modelling

Based on the table above, we can notice that the first discovered topic is ‘Service and Staff’. This topic focused on perceptions of airline staff and the quality of service provided during flights. The most frequent keywords associated with this topic are staff, service, crew and customer service. The second topic is ‘Flight Experience’ which reflects passengers’ perception of different elements like flight delay problems and service quality that affect the overall experience. The most frequent keywords like check, time, business class, passenger, long haul, experience, and delay point out the problems during the journey. The third topic is ‘Food and Drink’ which captures the feedback regarding the meal and beverages provided during the flight. The most frequent keywords in this topic are meal, food, drink, snack, breakfast, and water as passengers are very concerned about the overall dining experience. Lastly, the topic ‘Seat Comfort’ has been determined which reflects that passengers are highly concerned with the amount of personal space during flight. Based on the analysis, the most frequent words are seat, leg, legroom, space, cramped and comfort.

Besides, word clouds were generated to offer a visual representation of the most common terms found in passenger reviews. Even though word clouds are useful for identifying repeating phrases, some overlap between topics and generic terms limits their ability to clearly define the topics. As a result, they were employed to support the final topic interpretations. Figures 9, 10, 11, and 12 show the word clouds which are labelled according to their LDA topic number.

Figure 9: Word Cloud for LDA Topic 1

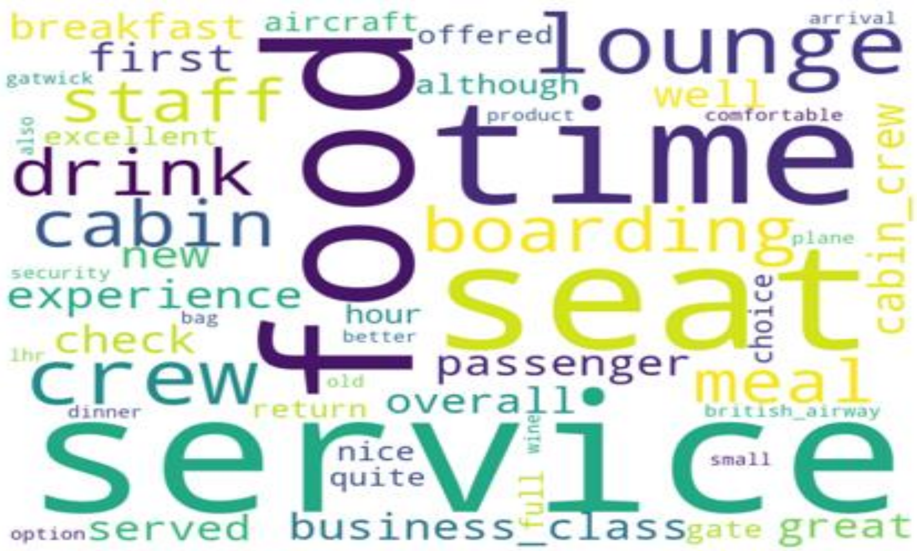

Figure 10: Word Cloud for LDA Topic 2

Figure 11: Word Cloud for LDA Topic 3

Figure 12: Word Cloud for LDA Topic 4

4.2 Sentiment Analysis Results on Health-Related Aspects

This section illustrates the findings of sentiment analysis results on health-related aspects identified in airline passengers' reviews. The analysis focused on four main health aspects: mental health, physical body health, dietary health and emotional health. The results were classified into three categories which are good impact, neutral, and bad impact.

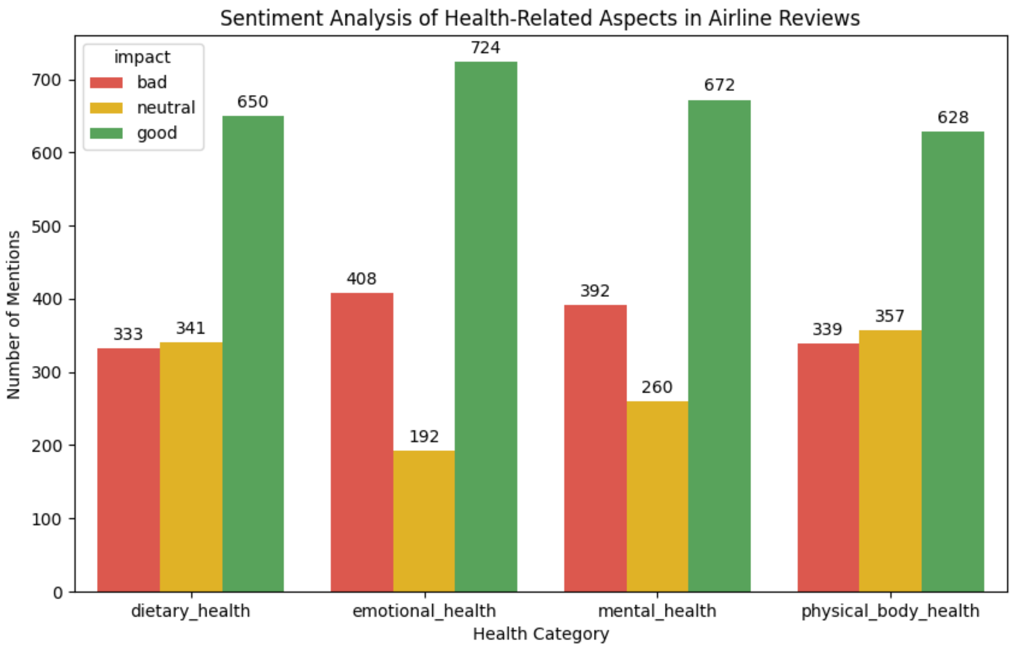

Figure 13: Result of Sentiment Analysis on Health-Related Aspects

According to Figure 13, the sentiment distribution for all the health-related aspects is displayed. In summary, all health aspects indicate a larger number of good impacts compared to the bad impacts. This result showed that British Airways passengers had a very good experience with their health and well-being while travelling.

Dietary Health

Dietary health has the lowest volume of bad impacts which is 333, with 341 neutrals and 650 good impacts. This has shown that British Airways is performing relatively well in the food and beverages sector, while bad impact was pointed out by cold meals or a lack of healthy food options.

Emotional Health

Emotional health received the most good impact reviews at 724, but it also has the most bad impact reviews at 408. Some passengers might face friendly and helpful staff, while a significant proportion reported that they have faced rude or inattentive treatment that directly impacts their emotional well-being.

Mental Health

For the mental health results, it received the second-highest good impact for 672 reviews, 260 neutral impacts, and 392 negative impacts. This has shown that passengers were relaxed and satisfied with the airline's smooth process, but the bad impacts represented problems like aircraft delays or uncomfortable environments.

Physical Body Health

For physical body health, 628 reviews were categorized as good impacts, 357 as neutral, 339 as bad impacts. Good impacts represented seats are very comfortable and have adequate space, while the negative impacts illustrated the passengers felt discomfort due to cramped seating and lack of legroom.

4.3 Emotion Distribution Across Airline Service Factors

This section displays the emotion distribution across different airline service factors, which are flight experience, food and drink, seat comfort, as well as service and staff. It was analysed by categorising passenger emotions into seven specific emotional states: anger, fear, joy, love, neutral, sadness, and surprise. The results demonstrate that joy was the most common emotion across all airline service factors. According to Figure 14, there are 511 occurrences of joy emotion for the flight experience and 557 occurrences for the service and staff aspect. A total of 1324 reviews were analysed to assess the emotion responses associated with these service factors.

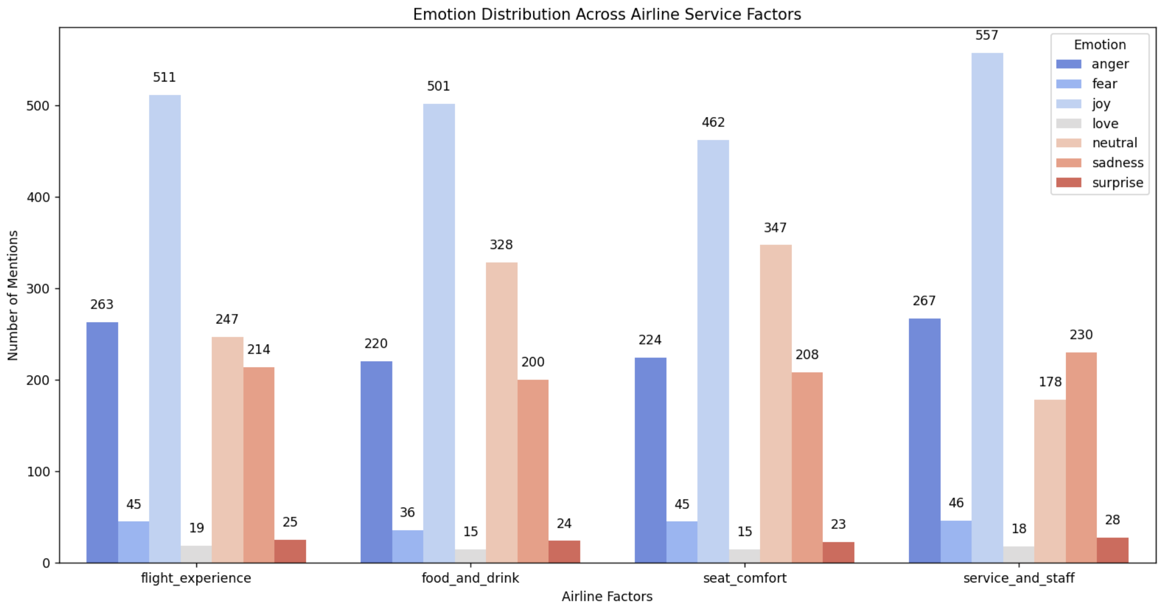

Figure 14: Result of Emotion Distribution Across Airline Service Factors

However, anger was notably high for flight experience (263 occurrences), food and drink (220 occurrences), seat comfort (224 occurrences), with service and staff (267 occurrences) also mentioning that passengers felt anger during the flight. A total of 214 passengers felt sadness for the flight experience, 200 passengers for the food and drink aspect, 208 passengers for seat comfort, 230 passengers for service and staff. The highest number of neutral emotions was observed in seat comfort (347 occurrences) and food and drink (328 occurrences). Emotions like fear, love, and surprise were comparatively low across all service factors.

In summary, the data highlighted that joy was the most popular emotion across all airline service factors. Neutral feelings were also common because passengers did not provide any feedback on these service factors in the reviews. In contrast, negative emotions like anger and sadness were notably prevalent in all airline service factors, which also shows that passengers experienced dissatisfaction and frustration.

4.4 Root Cause Analysis Results

This section shows the findings of the root cause analysis by identifying the top root causes affecting different aspects of passenger health and well-being based on the reviews. The analysis focuses on four key health categories: dietary health, emotional health, mental health, and physical body health. The results were obtained through a combination of rule-based analysis and bigram analysis. The rule-based analysis will focus on identifying predefined terms associated with each health aspect, while the bigram analysis will evaluate the frequency of occurring word pairs that are related to the health concerns. By identifying these root causes, British Airways can obtain a better understanding of what the major elements affecting passengers' health throughout their flights, therefore make enhancements to them.

4.4.1 Root Causes Affecting Dietary Health

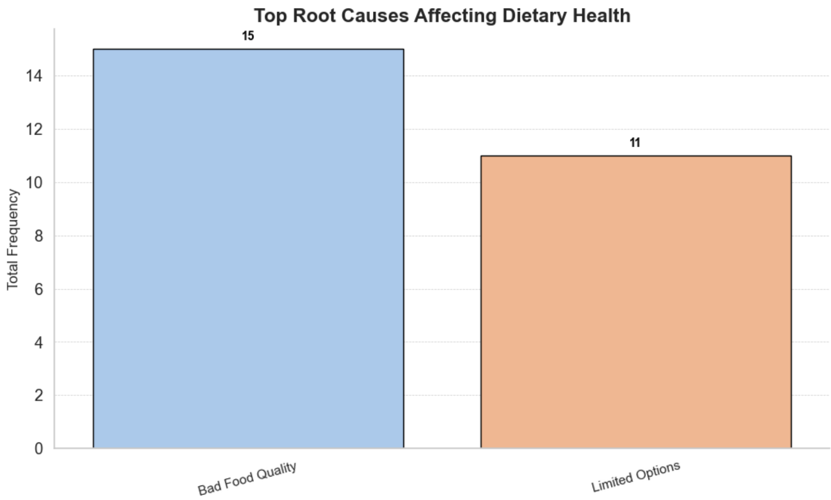

Figure 15: Top Root Causes Affecting Dietary Health

Figure 15 demonstrates the top root causes that affect dietary health. It was found that 15 passengers reported bad food quality, and 11 passengers mentioned limited meal options. These issues show the passenger's dissatisfaction with food-related aspects during the flight.

4.4.2 Root Causes Affecting Emotional Health

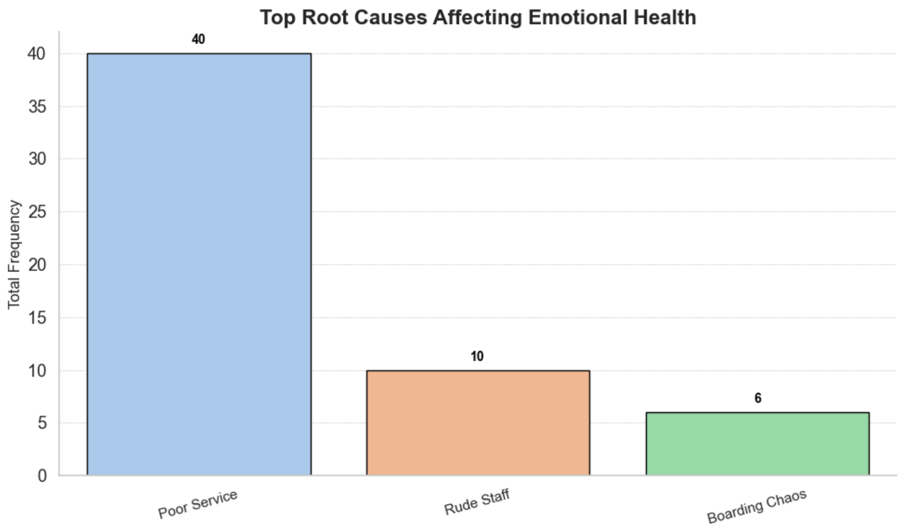

Figure 16: Top Root Causes Affecting Emotional Health

Based on Figure 16, 40 passengers have reported the poor service of the airline, which is the top root cause affecting emotional health. Other factors like rude staff are mentioned by 10 passengers and boarding chaos is mentioned by 6 passengers. These results emphasize that negative interactions during flight contribute to emotional health issues, hence making passengers feel stress and frustration.

4.4.3 Root Causes Affecting Mental Health

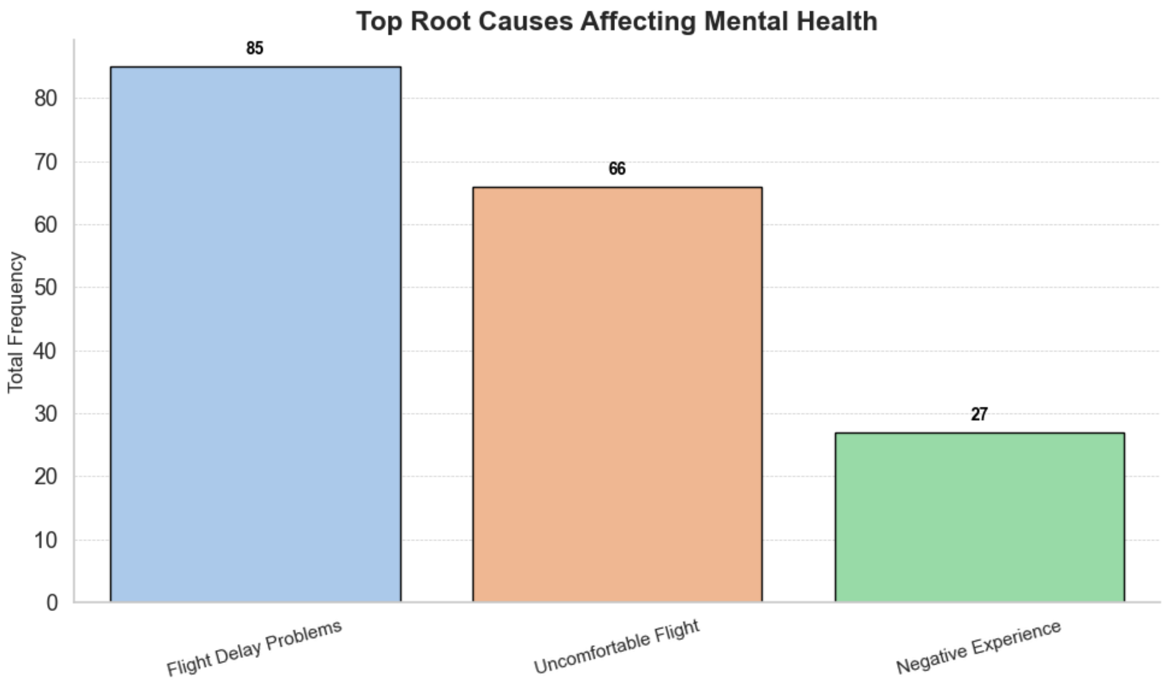

Figure 17: Top Root Causes Affecting Mental Health

According to Figure 17, the top root cause that affects mental health is flight delay problems. 85 passengers have reported flight delays, 66 passengers mentioned uncomfortable flights, and negative experience was noted by 27 passengers. These factors highlight the influence of bad flight conditions on passenger mental health, showing that delays and discomfort are the major sources of passenger stress.

4.4.4 Root Causes Affecting Physical Body Health

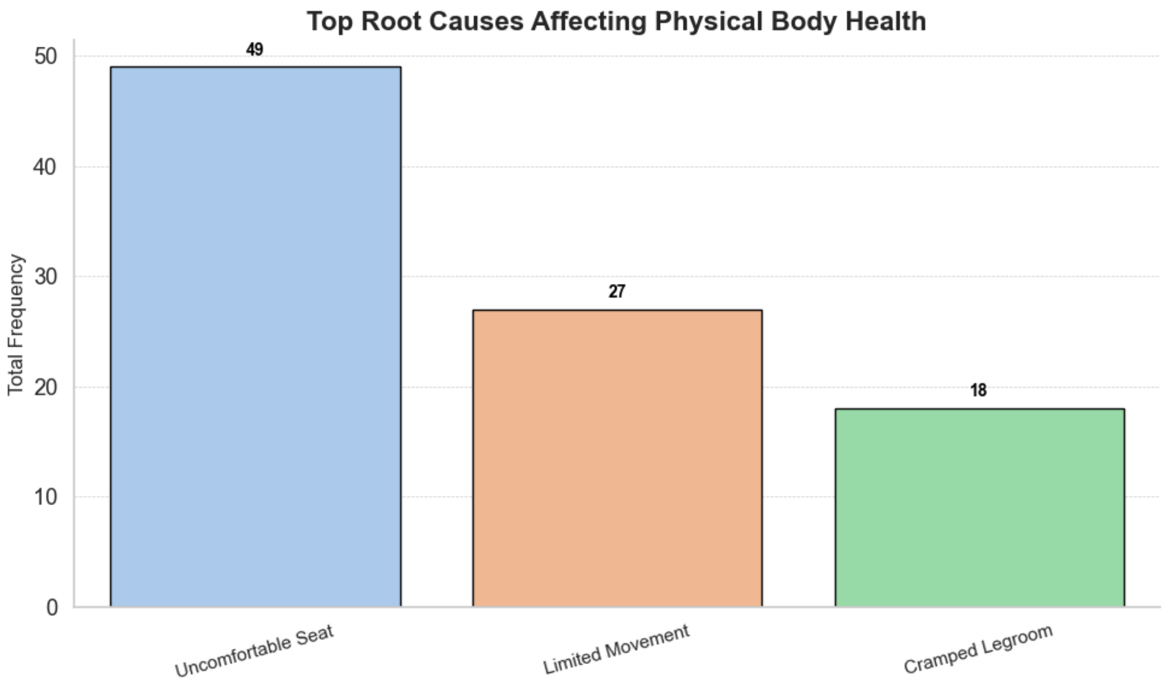

Figure 18: Top Root Causes Affecting Physical Body Health

In Figure 18, physical body health issues were mostly affected by the uncomfortable seat as 49 passengers reported on this problem. In addition, 27 passengers mentioned that the airline space is narrow which causes limited movement, while cramped legroom was reported by 18 passengers. These root causes can harm the passengers' physical body health by causing them fatigue and physical strain during the flight.

4.5 Yearly Trends in Passenger Sentiment Across Health Aspects

In this section, the results of yearly trends in passenger sentiment across different health aspects have been demonstrated. The sentiment can be divided into three categories which are good health impact, neutral health impact, and bad health impact. It is used to determine how passengers feel about various health-related factors during their flight over time. The graphs are examined in four major health categories: dietary health, emotional health, mental health, and physical body health.

4.5.1 Dietary Health Trends

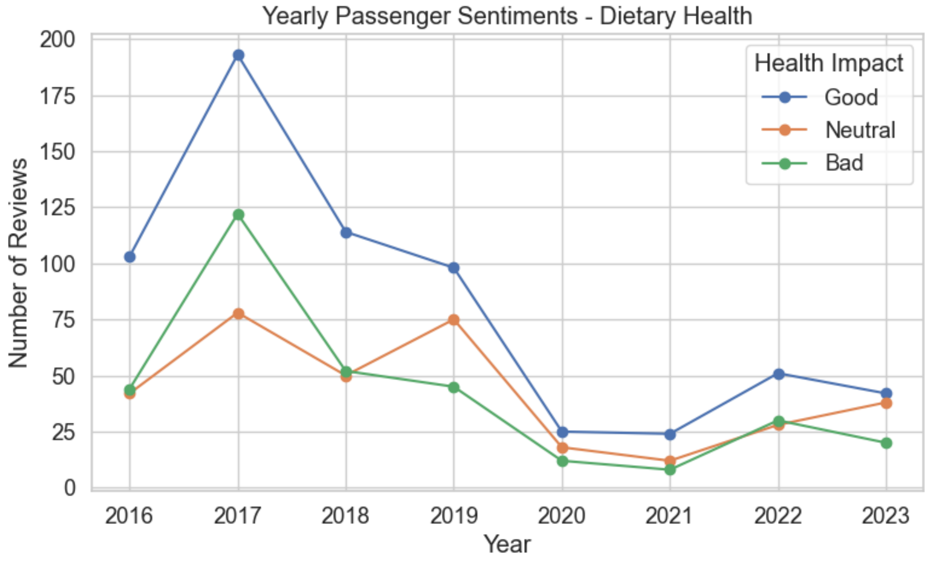

Figure 19: Yearly Trends in Passenger Sentiments Toward Dietary Health

Figure 19 demonstrates the sentiment trend for dietary health from 2016 to 2023. In 2017, the good health impact reviews were at their peak, with over 175 reviews. After that, the number of reviews started to drop until 2020. Although there was a minor rebound after 2021, the rating remained lower compared to the previous peak, which shows that passengers' initial optimism about food faded over time.

Next, the reviews for neutral impacts were on a more stable path. It peaked in 2019 before dropping in 2020 and 2021, then rebounded in 2022 and 2023. This reflects that the pandemic has disrupted the standard service expectations. Moving on, the reviews of bad health impacts peaked in 2017, which indicates passengers were dissatisfied with the food quality and variety. After 2017, the sentiments dropped continuously until 2021 and had a minor increase in 2022.

4.5.2 Emotional Health Trends

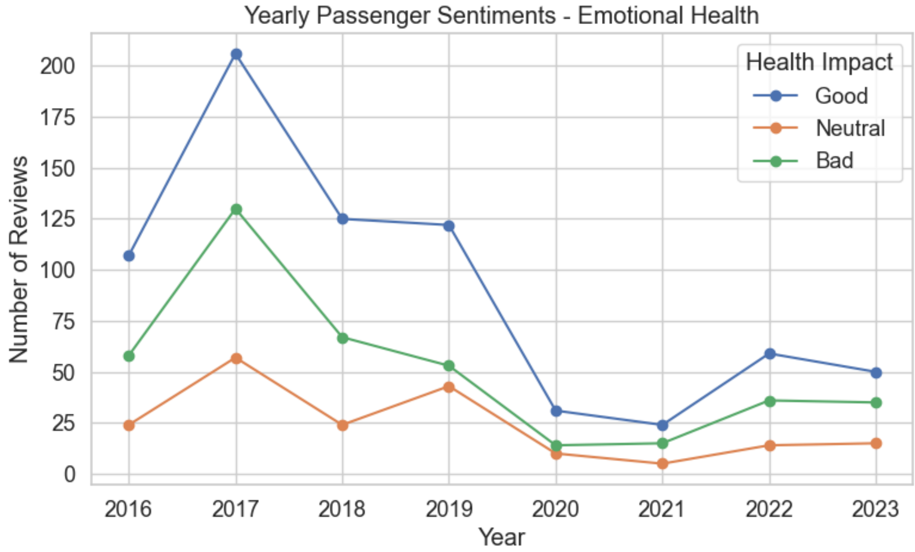

Figure 20: Yearly Trends in Passenger Sentiments Toward Emotional Health

Figure 20 illustrates the emotional health sentiments trends from 2016 to 2023. The reviews for good health impacts reached a peak in 2017 with over 200 mentions. However, the reviews started to drop in 2018, and it was mostly caused by the COVID-19 pandemic. A tiny rebound occurred in 2022 and 2023 as the epidemic is getting better.

In 2017, the reviews of bad health impacts peaked and declined continuously throughout the year. The lowest point for the bad health impacts reviews was in 2020, with a minor rebound in the following years. As mentioned previously, the fall could reflect the reduced volume of flights during the pandemic season.

The review indicating neutral impacts remained at a lower point during the 8 years. It spiked in 2017 and 2019 before declining to the lowest point in 2021. In summary, these health impact patterns showed that passengers have a strong emotional reaction in the early years, and this decreased year by year due to operational constraints and reduced service customisation.

4.5.3 Mental Health Trends

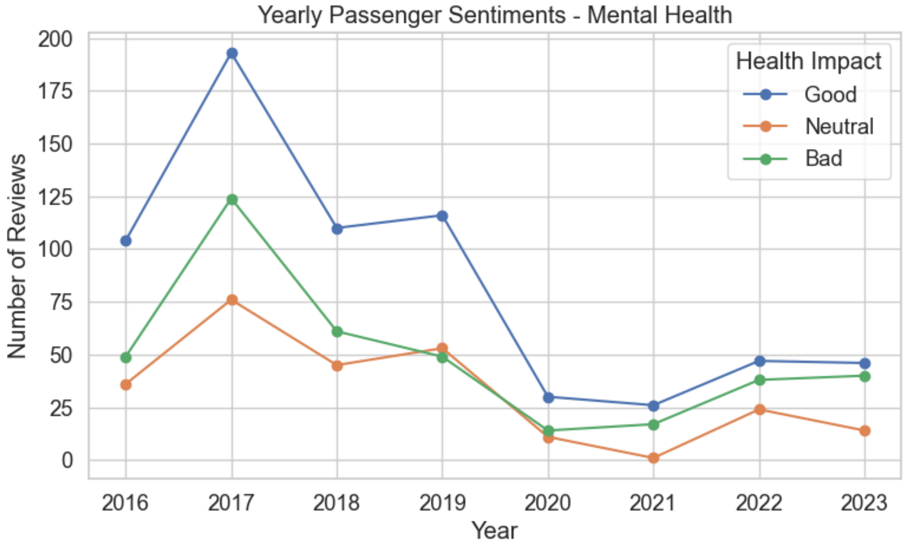

Figure 21: Yearly Trends in Passenger Sentiments Toward Mental Health

Figure 21 demonstrates the yearly passenger sentiments related to mental health, which mainly focus on issues like flight delay or other negative travel experiences. The number of reviews indicating good health impacts peaked in 2017 and dropped steadily until 2021. There was some rebound after the pandemic in 2022 and 2023, although the good health impacts are still low compared to previous years.

The bad health impacts rose from 2016 to 2017, which indicates that passengers are facing problems with mental stress when travelling. By 2020, the number of reviews had declined dramatically, mostly due to the fewer flights in the pandemic season. However, a consistent level of health-adverse maintained in the years afterwards.

For the neutral health impacts, there were no major changes like the good health and bad health impacts. It reached a small peak in 2017 and dropped continuously until 2021. It can be observed that the impacts started to grow in 2022, and this has highlighted that the pandemic season might be over.

4.5.4 Physical Body Health Trends

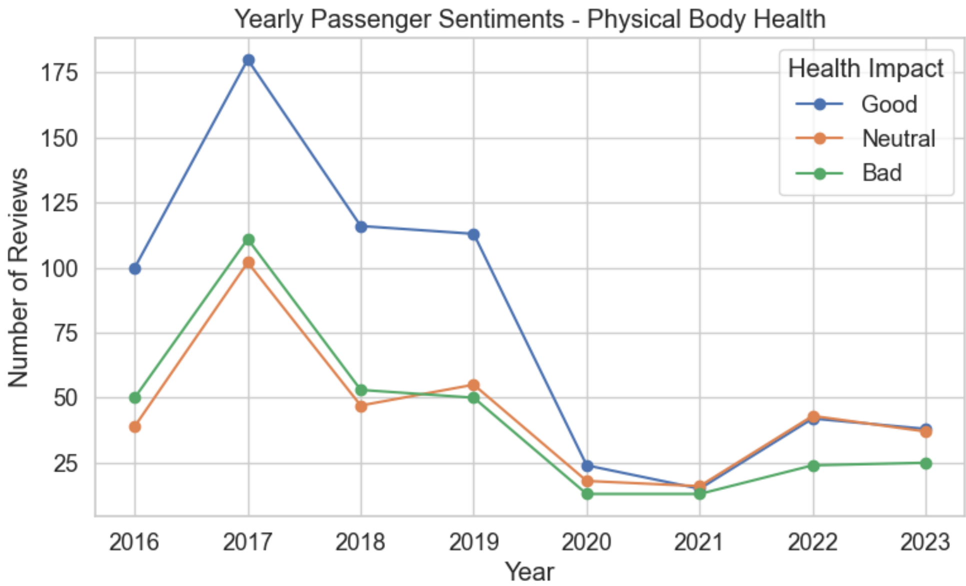

Figure 22: Yearly Trends in Passenger Sentiments Toward Physical Body Health

Figure 22 depicts the yearly trends in passenger sentiments toward physical body health from 2016 to 2023. As shown in the figure, the number of reviews indicating good health impact peaked in 2017 but then dropped continuously until 2021. There is also a minor rebound for the years 2022 and 2023. This suggests that although physical comfort was previously valued, it has yet to fully recover as a source of good health impacts.

The bad health impacts demonstrated a similar pattern, which peaked in 2017 and fell continuously over the next few years. However, the minor increase began in 2022, which shows that discomfort complaints about the seat or space have become more common.

Based on the figure above, the neutral impact remained constant over time. Some passengers might think that physical comfort issues are a permanent aspect of the travel experience, neither worsening nor improving much.

5 Evaluation and Discussion

5.1 Brief Summary of Chapter

This chapter will evaluate and discuss the sentiment and emotion results gathered from the airline passengers' reviews, with particular emphasis on the impact of passenger health and well-being. The goal is to interpret the findings using the NLP techniques and explore what passengers are concerned about in the flight experience based on their emotional, mental, dietary, and physical body health.

The discussion starts with the examination of the emotional distribution in different airline service factors, for example, flight experience, food and drink, service and staff, as well as seat comfort. These emotional patterns illustrate how passengers respond on a deeper level to various service factors. Next, this chapter will delve into the most important aspects of the evaluation, which is the root cause analysis discussion. This section will explore the fundamental causes of poor passenger reviews and tie them to different kinds of health issues. This study will be supported by a visualization graph called a fishbone diagram.

In addition, this chapter also makes comparisons with previous studies, either supporting or questioning what they mentioned about airline services and their impact on passenger health. In the following section, the implications of the findings have been discussed in two subsections, which are implications for airlines and implications for health and well-being. Finally, the chapter shows the limitations of the study.

5.2 Emotion Distribution Discussion

This section investigates the distribution of emotions across various airline service factors, including flight experience, food and drink, seat comfort, as well as service and staff. By studying these emotional responses, we can get important insights into passenger experiences and determine which airline service factors contribute to positive and negative passenger health and well-being.

According to Figure 14, joy emotion is the most common feeling across all airline service factors. This shows that passengers felt highly satisfied with these services, indicating the services exceed passenger expectations and promote good emotional states. Furthermore, the number of mentions for neutral feelings across all airline service factors supports the idea that passengers may not feel strongly enough about particular aspects to express their feelings. The neutral feeling emotions represented by customers have reasonable expectations that do not elicit significant emotional responses.

However, the distribution of negative emotions like anger and sadness shows the significant impact of airline service factors on passengers' emotional health. Anger emotion was particularly high in airline service factors like flight experience (263 instances), food and drink (220 instances), service and staff (267 instances), and seat comfort (224 instances). These findings indicate that negative experiences such as flight delay, poor service, uncomfortable seats, or bad food quality have a major influence on passenger anger emotions. The high amount of sadness also shows that passengers faced bad treatment in flight, which led to their disappointment and frustration.

Furthermore, emotions like fear, love and surprise were less mentioned by passengers, showing that passengers' emotions are typically influenced by negative or positive experiences rather than rare emotions. These findings are consistently related to a previous study by Nguyen Ngoc Hien et al. (2024), the emotional states like anger, sadness and joy can elicit various behavioural responses. These emotional states are also more commonly observed in passenger reviews than surprise or love.

In summary, the emotional response reflects the complexities of the airline service experience. Although joy emotions reflect well airline service quality, negative emotions like anger and sadness should be considered by British Airways to make informed decisions and enhance customer experience. A deeper understanding of these emotional reactions may help airlines a lot to modify their services to better match the passengers' demands, therefore improving customer loyalty and ensuring long-term success.

5.3 Root Cause Analysis Discussion

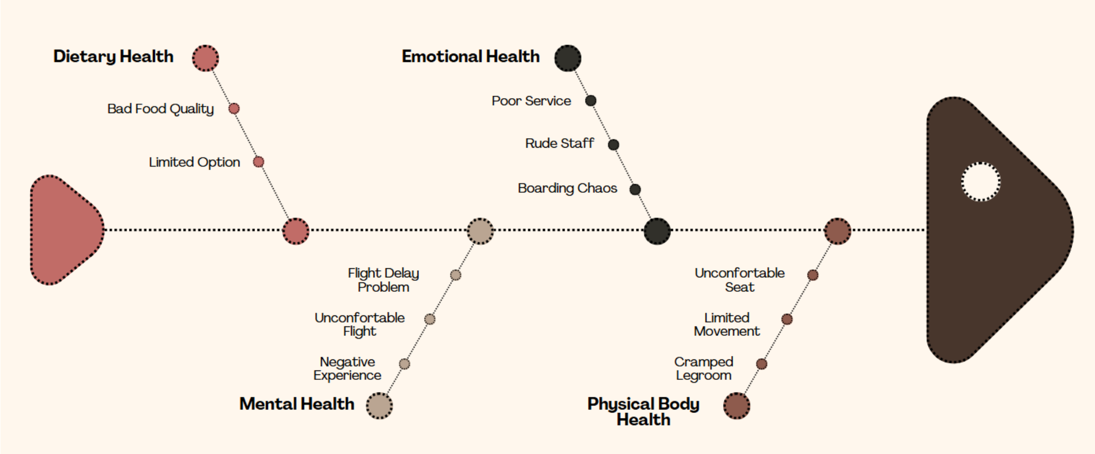

Figure 23: Root Cause Analysis of Health-Related Aspects Using Fishbone Diagram

This section will discuss the root cause analysis of health-related aspects using a fishbone diagram. Based on Figure 23, it can be observed that the four main health aspects have been analysed, which are dietary health, emotional health, mental health and physical body health. Understanding these root causes will enable us to investigate how they affect passengers' well-being and provide insights for improvement to reduce these health effects.

Dietary Health

Two major root causes that affect dietary health are bad food quality and limited food options. Poor food quality like cold meals or food that does not match passengers' nutritional criteria can lead to body discomfort, particularly for people with special dietary requirements. This can also result in health issues like digestive problems or a lack of nutrients.

For some passengers with food allergies or passengers who prefer a vegan diet, they might not be able to find suitable meals because of the limited options. Since they can only serve the unsuitable food or go without, this will result in low energy levels and indigestion, especially for passengers who taking long-haul flights.

Emotional Health

Poor service, rude staff, and boarding chaos are the root causes that affect passenger emotional health. The interactions between passengers and staff are very crucial in an airline's operations. When passengers meet unhelpful or impolite staff, they may feel frustrated and angry. These negative emotional responses not only have an immediate impact on the flying experience but can also have long-term psychological consequences, such as stress and anxiety.

In addition, boarding chaos is also determined as a root cause that impacts negative emotional health. When passengers already feel anxious about their flight, a disorganised boarding procedure can cause worry and feelings of helplessness. These emotional reactions might cause a drop in overall passenger satisfaction and affect their emotional well-being throughout the journey.

Mental Health

The main root causes of mental health issues are flight delay problems, uncomfortable flight and negative experience. First, long waiting periods and unexpected delays are major stressors for passengers because these problems might disrupt their travel planning. For individuals who already have anxiety disorders, the problems of flight delay may result in serious mental health outcomes.

The next root cause is the uncomfortable flight. The poor cabin environment and lack of in-flight entertainment are the problems that lead to mental fatigue and increased stress. Passengers on long-haul flights who cannot rest properly or get relax may affect their mood to being anger and affect their emotional stability.

Physical Body Health

The top root causes that affect physical body health include uncomfortable seat, limited movement, and cramped legroom. The long durations of sitting on the uncomfortable seat can cause physical discomfort like back pain, neck stiffness, as well as muscle cramps. This problem often occurs in long-haul flights because passengers are forced to remain seated for lengthy periods without being able to move freely.

Moreover, limited movement space inside the aircraft also increases the risk of physical body health issues like poor circulation and deep vein thrombosis (DVT). When passengers want to stretch or walk around the flight, the limited space will hinder them from relaxing. The cramped legroom will cause discomfort for taller people and people with larger body types. This physical strain not only affects the passenger flight experience but also leads to back pain or joint problems.

By conducting the root cause analysis, we can assess the airline service factors that impact passengers' dietary, mental, emotional and physical body health. This detailed analysis is critical for understanding how the root cause, like food quality, seat comfort, flight delays and service aspects, affects the passenger's health and well-being. This finding also contributes to the third objective, which is to evaluate how the identified service aspects affect different dimensions of passenger well-being. By addressing these core problems, British Airways can improve passenger satisfaction and minimise the health-related complaints.

5.4 Comparison with Previous Studies

The previous studies only investigated either airline service analysis or the health effects of air travel, and none of them combined these fields. In recent research, Kontonatsios et al. (2023) and Chang et al. (2022) only focused on using Aspect-Based Sentiment Analysis (ABSA) to discover customer satisfaction across service aspects. Their research still gave helpful insights, although they did not apply their results to health-related issues. In this study, the aspect sentiment results are linked with different health impacts like such as physical, mental, dietary, and emotional health.

Furthermore, the previous studies on health in aviation by Nicholson et al. (2003) and Benz et al. (2022) did not use the NLP techniques and relied on survey-based and observation methodologies. Their studies emphasized the health hazards caused by factors like limited mobility and environmental noise. They did not use passenger reviews to get closer to the health problems. Therefore, this study proposed an innovative framework for analysing health-related topics from online passenger reviews by integrating ABSA and emotion detection. Moreover, the study from Çallı and Çallı (2022) only used LDA topic modelling to identify the main complaints from airline passenger reviews. So, this study utilised root cause analysis and health impact mapping to improve the interpretability of passengers' issues in another way.

In summary, this study makes a new contribution by bridging the gap between NLP analysis and health-related impact evaluation. It demonstrates how passenger reviews may be systematically analysed to reveal their service performance and impact on passengers’ health, which is an area the current academic study hasn’t explored yet.

5.5 Implications of Findings

This study offers valuable insights that may be used for both the airline sector and passenger health and well-being. By integrating Aspect-Based Sentiment Analysis with emotion detection and health-related evaluation, the study bridges the gap between airline service performance and impacts on passenger health and well-being.

Implications for Airlines

For airline aspects, understanding how airline service factors affect passengers' health is crucial for improving overall service quality and customer satisfaction. This study emphasises service factors like seat comfort, service and staff, flight experience, as well as food and drink that have a direct impact on passenger well-being. Airlines may utilise these findings to enhance their service not only for comfort, but also to reduce health hazards. For illustration, airlines can make enhancements for seat ergonomics and resolve cabin environment issues to increase customer satisfaction and reduce health-related complaints.

Furthermore, airlines may track passengers' emotions and sentiments in real time to handle the problems before they worsen. Airlines can also adjust their services to better match passengers' needs according to the results from the different NLP analyses, especially in terms of emotional and psychological aspects.

Implications for Passenger Health

The integration of ABSA with health-related issues creates a unique framework for measuring how various airline service factors influence passengers' health. By mapping sentiment results to various health categories, airlines and healthcare professionals may acquire a better understanding of how these service factors lead to health concerns like anxiety or physical pain. This study creates a way for future research on the relationship between travel experience and health. It also enables greater investigation into how air travel environments can be optimised for passenger well-being.

Moreover, this study can help policymakers in the aviation industry and healthcare professionals by giving them scientific proof of the health impacts of air travel. It can also improve the standards that prioritize passenger health like recommended seating arrangements, fight environmental conditions, and noise control.

5.6 **Limitations of the Study**

This study provided valuable insights into the relationship between airline service factors and passenger well-being. However, there are several limitations to consider.

The first limitation is that the size of the dataset is too small for natural language processing. The dataset used for analysis only contained 1324 reviews, which may limit the depth of root cause analysis. A larger dataset would be more suitable for this study for offering more comprehensive insights into airline service factors that affect passenger health.

The second limitation was found when conducting topic modelling. It is shown that the coherence score of the topic modelling results was relatively low, and this suggests that the model did not capture the topics well. This limitation may influence the accuracy of the identified service aspects and their relationship to passenger health. Future research might improve the model by conducting deeper data cleaning or using another model to increase the topic coherence.

The last limitation lies in the sentiment analysis tool in the study, which includes VADER for Aspect-Based Sentiment Analysis (ABSA) and DistilBERT for emotion detection. Although these tools are very effective, they are not able to fully capture the complexity of human sentiment when sarcasm or mixed emotions occur. This limitation could affect the accuracy of sentiment and emotional analysis results. Therefore, future research could use more advanced techniques to overcome these limitations.

6 Conclusion and Recommendation

To summarise, this study investigated the relationship between airline service factors and passenger well-being using natural language processing techniques such as Latent Dirichlet Allocation (LDA) topic modelling, Aspect-Based Sentiment Analysis (ABSA), and DistilBERT emotion detection. This study aimed to narrow the gap between airline service performance and its influence on passenger health by focusing on factors like seat comfort, food and drink, service and crew, and flight experience. This study analysed the British Airways passenger reviews to provide significant insights into how these service factors directly affect passengers' physical body, mental, dietary and emotional well-being.

The findings demonstrates that uncomfortable seats, poor service and flight delay problems had a negative impact on passengers' dietary and mental well-being. Furthermore, the emotional analysis revealed strong sentiments of joy and anger related to many service aspects, providing a deeper insight into passenger experiences.

The study faced some limitations, including a small dataset size, challenges with topic modelling coherence, and constraints of sentiment analysis tools used in handling complex contextual reviews. However, it has provided a solid foundation for future research into how airline service factors affect passenger health and well-being. For recommendations, future studies can utilise larger or real-time datasets to enhance the depth of root cause analysis and emotional insights.

Finally, this research framework could be applied to other industries like healthcare or public transportation, where the customer experience and health are interconnected. Airlines and governments might also benefit from using real-time sentiment and emotion tracking systems, which would allow them to respond proactively to passenger requirements and develop services that promote both satisfaction well-being.

7 Ethics Approval Form

Do not amend before use

Middlesex University London

Research Ethics Screening Form for Students

Only for students on taught programmes – e.g., BSc, MSc, MA, LLM etc

NOT for PostGraduate Researchers – e.g., MRes/MPhil/PhD degrees

Middlesex University is concerned with protecting the rights, health, safety, dignity, and privacy of its research participants. It is also concerned with protecting the health, safety, rights, and academic freedom of its students and with safeguarding its own reputation for conducting high quality, ethical research.

This Research Ethics Screening Form will enable students to self-assess and determine whether the research requires ethical review and approval via the Middlesex Online Research Ethics (MORE) form before commencing the study. Supervisors must approve this form after consultation with students.

Student Name:	Chuar Kia Yi	Email:kc1287@live.mdx.ac.uk
Research project title:	Analysing the Impact of Airline Service Factors on Passenger Well-being Using NLP	
Programme of study/module:	Bs(C) Computer Science	
Supervisor Name:	Dr. Olugbenga Oluwagbemi	Email:O.Oluwagbemi@mdx.ac.uk

Please answer whether your research/study involves any of the following given below:		
1. HANIMALS or animal parts.	☐ Yes	☒ No
2. MCELL LINES (established and commercially available cells - biological research).	☐ Yes	☒ No
3. HCELL CULTURE (Primary: from animal/human cells- biological research).	☐ Yes	☒ No
4. HCLINICAL Audits or Assessments (e.g. in medical settings).	☐ Yes	☒ No
5. XCONFLICT of INTEREST or lack of IMPARTIALITY. If unsure see "Code of Practice for Research" (Sec 3.5) at: https://unihub.mdx.ac.uk/study/spotlights/types/research-at-middlesex/research-ethics	☐ Yes	☒ No
6. XDATA to be used that is not freely available (e.g. secondary data needing permission for access or use).	☐ Yes	☒ No
7. XDAMAGE (e.g., to precious artefacts or to the environment) or present a significant risk to society).	☐ Yes	☒ No
8. XEXTERNAL ORGANISATION – research carried out within an external organisation or your reseach is commissioned by a government (or government body).	☐ Yes	☒ No
9. MFIELDWORK (e.g biological research, ethnography studies).	☐ Yes	☒ No
10. HGENETICALLTY MODIFIED ORGANISMS (GMOs) (biological research).	☐ Yes	☒ No
11. HGENE THERAPY including DNA sequenced data (biological research).	☐ Yes	☒ No

Do not amend before use

Middlesex University London

12. [M]HUMAN PARTICIPANTS – ANONYMOUS Questionnaires (participants not identified or identifiable).	☐ Yes	☒ No
13. [X]HUMAN PARTICIPANTS – IDENTIFIABLE (participants are identified or can be identified): survey questionnaire/ INTERVIEWS / focus groups / experiments / observation studies/ evaluation studies.	☐ Yes	☒ No
14. [H]HUMAN TISSUE (e.g., human relevant material, e.g., blood, saliva, urine, breast milk, faecal material).	☐ Yes	☒ No
15. [H]ILLEGAL/HARMFUL activities research (e.g., development of technology intended to be used in an illegal/harmful context or to breach security systems, searching the internet for information on highly sensitive topics such as child and extreme pornography, terrorism, use of the DARK WEB, research harmful to national security).	☐ Yes	☒ No
16. [X]PERMISSION is required to access premises or research participants.	☐ Yes	☒ No
17. [X]PERSONAL DATA PROCESSING (Any activity with data that can directly or indirectly identify a living person). For example data gathered from interviews, databases, digital devices such as mobile phones, social media or internet platforms or apps with or without individuals'/owners' knowledge or consent, and/or could lead to individuals/owners being IDENTIFIED or SPECIAL CATEGORY DATA (GDPR[1]) or CRIMINAL OFFENCE DATA. [1]Special category data (GDPR- Art.9): "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation".	☐ Yes	☒ No
18. [X]PUBLIC WORKS DOCTORATES: Evidence of permission is required for use of works/artifacts (that are protected by Intellectual Property (IP) Rights, e.g. copyright, design right) in a doctoral critical commentary when the IP in the work/artifactis jointly prepared/produced or is owned by another body	☐ Yes	☒ No
19. [H]RISK OF PHYSICAL OR PSYCHOLOGICAL HARM (e.g., TRAVEL to dangerous places in your own country or in a foreign country (see https://www.gov.uk/foreigntravel-advice), research with NGOs/humanitarian groups in conflict/dangerous zones, development of technology/agent/chemical that may be harmful to others, any other foreseeable dangerous risks).	☐ Yes	☒ No
20. [X]SECURITY CLEARANCE – required for research.	☐ Yes	☒ No
21. [X]SENSITIVE TOPICS (e.g., anything deeply personal and distressing, taboo, intrusive, stigmatising, sexual in nature, potentially dangerous, etc).	☐ Yes	☒ No

M – Minimal Risk; X – More than Minimal Risk. H – High Risk

If you have answered 'Yes' to ANY of the items in the table, your application REQUIRES ethical review and approval using the MOREform **BEFORE commencing your research**. Please apply for ethical approval using the MOREform (https://moreform.mdx.ac.uk/). Consult your supervisor for guidance. Also see *Middlesex Online Research Ethics* (MyLearning area*) and www.tiny.cc/mdx-ethics*

If you have answered 'No' to ALL of the items in the table, your application is Low Risk and you may NOT require ethical review and approval using the MOREform before commencing your research. Your research supervisor will confirm this below.

Based on the details provided in the self-assesment form, I confirm that:	Insert Y or N
The study is Low Risk and *does not require* ethical review & approval using the MOREform	Y

Student Signature:............. .... Date:......28/02/2025.................

To be completed by the supervisor:

The study *requires* ethical review and approval using the MOREform.	

Superivsor Signature:......OLUWAGBEMI O.O....................... Date: February 28th, 2025

8 References

1. Abubakar Ahmad Aliero, Bashir, S., Hamzat Olanrewaju Aliyu and Bashar Umar Kangiwa (2023). *Systematic Review on Text Normalization Techniques and its Approach to Non-Standard Words*. [online] ResearchGate. Available at: https://www.researchgate.net/publication/374160463_Systematic_Review_on_Text_Normalization_Techniques_and_its_Approach_to_Non-Standard_Words.

2. Ameni Dhaoui Boumaiza (2016). A Survey on Sentiment Analysis and Visualization. *Qatar Foundation Annual Research Conference Proceedings Volume 2016 Issue 1*. doi:https://doi.org/10.5339/qfarc.2016.ictpp1203.

3. Ban, H.-J. and Kim, H.-S. (2019). Understanding Customer Experience and Satisfaction through Airline Passengers' Online Review. *Sustainability*, [online] 11(15), p.4066. doi:https://doi.org/10.3390/su11154066.

4. Benz, S., Kuhlmann, J., Jeram, S., Bartels, S., Ohlenforst, B. and Schreckenberg, D. (2022). Impact of Aircraft Noise on Health. *Aviation Noise Impact Management*, pp.173–195. doi:https://doi.org/10.1007/978-3-030-91194-2_7.

5. Çallı, L. and Çallı, F. (2022). Understanding Airline Passengers during Covid-19 Outbreak to Improve Service Quality: Topic Modeling Approach to Complaints with Latent Dirichlet Allocation Algorithm. *Transportation Research Record: Journal of the Transportation Research Board*, 2677(4), p.036119812211120. doi:https://doi.org/10.1177/03611981221112096.

6. Chang, Y.-C., Ku, C.-H. and Nguyen, D.-D.L. (2022). Predicting aspect-based sentiment using deep learning and information visualization: The impact of COVID-19 on the airline industry. *Information & Management*, 59(2), p.103587. doi:https://doi.org/10.1016/j.im.2021.103587.

7. Clapp, S.R., McCauley, P.R., Karwowski, W. and Hancock, P.A. (2021). The seat of happiness? The effect of seat comfort on the achievement of psychological flow during

transactional work. *Applied Ergonomics*, 96, p.103508. doi:https://doi.org/10.1016/j.apergo.2021.103508.

8. Farkhod, A., Abdusalomov, A., Makhmudov, F. and Cho, Y.I. (2021). LDA-Based Topic Modeling Sentiment Analysis Using Topic/Document/Sentence (TDS) Model. *Applied Sciences*, 11(23), p.11091. doi:https://doi.org/10.3390/app112311091.

9. Farrikh Alzami, Rama Aria Megantara, Dwi Puji Prabowo, Puri Sulistiyawati, Ricardus Anggi Pramunendar, Ika Novita Dewi and Ritzkal Ritzkal (2023). LDA Topic Analysis for Product Reviews in Social Media Platform. *Moneter Jurnal Keuangan dan Perbankan*, 11(2), pp.277–283. doi:https://doi.org/10.32832/moneter.v11i2.402.

10. Georgios Kontonatsios, Clive, J., Harrison, G.R., Metcalfe, T., Patrycja Sliwiak, Tahir, H. and Ghose, A. (2023). FABSA: An aspect-based sentiment analysis dataset of user reviews. *Neurocomputing*, 562, pp.126867–126867. doi:https://doi.org/10.1016/j.neucom.2023.126867.

11. Gu, Q., Wang, Z., Zhang, H., Sui, S. and Wang, R. (2024). Aspect-Level Sentiment Analysis Based on Syntax-Aware and Graph Convolutional Networks. *Applied sciences*, 14(2), pp.729–729. doi:https://doi.org/10.3390/app14020729.

12. Gupta, S., Ranjan, R. and Singh, S. (2024). *Comprehensive Study on Sentiment Analysis: From Rule based to modern LLM based system*. [online] Available at: https://arxiv.org/pdf/2409.09989.

13. Jim, J.R., Apon, M., Malakar, P., Kabir, M.M., Nur, K. and Mridha, M.F. (2024). Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, 6, pp.100059–100059. doi:https://doi.org/10.1016/j.nlp.2024.100059.

14. Johri, P., Khatri, S.K., Al-Taani, A.T., Sabharwal, M., Suvanov, S. and Kumar, A. (2021). Natural Language Processing: History, Evolution, Application, and Future Work. *Lecture Notes in Networks and Systems*, pp.365–375. doi:https://doi.org/10.1007/978-981-15-9712-1_31.

15. Khurana, D., Koli, A., Khatter, K. and Singh, S. (2022). Natural Language processing: State of the art, Current Trends and Challenges. *Multimedia Tools and Applications*, [online] 82(3), pp.3713–3744. doi:https://doi.org/10.1007/s11042-022-13428-4.

16. Khyani, D., B. S.Siddhartha, B.S., N. M.Niveditha, N.M., Divya, DivyaB.M. and Manu, Y.M. (2021). *(PDF) An Interpretation of Lemmatization and Stemming in Natural Language Processing*. [online] ResearchGate. Available at: https://www.researchgate.net/publication/348306833_An_Interpretation_of_Lemmatization_and_Stemming_in_Natural_Language_Processing.

17. Kumar, D., Gupta, A., Vishan Kumar Gupta and Gupta, A. (2023). Aspect-Based Sentiment Analysis using Machine Learning and Deep Learning Approaches. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(5s), pp.118–138. doi:https://doi.org/10.17762/ijritcc.v11i5s.6636.

18. Lestari, Y.D. and Murjito, E.A. (2020). Factor Determinants of Customer Satisfaction with Airline Services Using Big Data Approaches. *Jurnal Pendidikan Ekonomi Dan Bisnis (JPEB)*, 8(1), pp.34–42. doi:https://doi.org/10.21009/jpeb.008.1.4.

19. Nazir, A., Rao, Y., Wu, L. and Sun, L. (2020). Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey. *IEEE Transactions on Affective Computing*, pp.1–1. doi:https://doi.org/10.1109/taffc.2020.2970399.

20. Nguyen Ngoc Hien, Nguyen Thanh Long, Vo Tan Liem and Danh Xuan Luu (2024). Customer Responses to Airline Service Failure: Perspectives from Expectation Disconfirmation Theory. *SAGE open*, 14(2). doi:https://doi.org/10.1177/21582440241248334.

21. Nicholson, A.N., Cummin, A.R.C. and Giangrande, P.L.F. (2003). The airline passenger: current medical issues. *Travel Medicine and Infectious Disease*, 1(2), pp.94–102. doi:https://doi.org/10.1016/s1477-8939(03)00060-7.

22. PATEL, A., OZA, P. and AGRAWAL, S. (2023). Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model. *Procedia Computer Science*, 218, pp.2459–2467. doi:https://doi.org/10.1016/j.procs.2023.01.221.

23. Raghunathan, N. and Saravanakumar, K. (2023). Challenges and Issues in Sentiment Analysis: A Comprehensive Survey. *IEEE Access*, 11, pp.69626–69642. doi:https://doi.org/10.1109/access.2023.3293041.

24. Rana, T.A., Cheah, Y.-N. and Letchmunan, S. (2016). Topic Modeling in Sentiment Analysis: A Systematic Review. *Journal of ICT Research and Applications*, 10(1), pp.76–93. doi:https://doi.org/10.5614/itbj.ict.res.appl.2016.10.1.6.

25. Samah, K.A.F.A., Misdan, N.F.A., Jono, M.N.H.H. and Riza, L.S. (2022). The Best Malaysian Airline Companies Visualization through Bilingual Twitter Sentiment Analysis: A Machine Learning Classification. *JOIV : International Journal on Informatics Visualization*, [online] 6(1), pp.130–137. doi:https://doi.org/10.30630/joiv.6.1.879.

26. Semary, N.A., Ahmed, W., Amin, K., Paweł Pławiak and Hammad, M. (2024). Enhancing machine learning-based sentiment analysis through feature extraction techniques. *PLOS ONE*, [online] 19(2), pp.e0294968–e0294968. doi:https://doi.org/10.1371/journal.pone.0294968.

27. Shara, J. (2021). An Application Example of Sentiment Analysis Using R. *ICSNS-XVI-2021, ISBN: 978-9928-259-14-7*. [online] Available at: https://www.researchgate.net/publication/354921657_An_Application_Example_of_Sentiment_Analysis_Using_R.

28. Shrestha, D., Katkuri, S., Goel, U. and Lavanya, P. (2024). British Airways Reviews Analysis. *International Journal of Scientific Research & Engineering Trends*, [online] 10, pp.2395–566. Available at: https://ijsret.com/wp-content/uploads/2024/07/IJSRET_V10_issue4_274.pdf.

29. Srinivas, S. and Ramachandiran, S. (2023). Passenger intelligence as a competitive opportunity: unsupervised text analytics for discovering airline-specific insights from

online reviews. *Annals of Operations Research*. doi:https://doi.org/10.1007/s10479-022-05162-9.

30. Zhan, Z. (2025). Comparative Analysis of TF-IDF and Word2Vec in Sentiment Analysis: A Case of Food Reviews. *ITM Web of Conferences*, [online] 70, p.02013. doi:https://doi.org/10.1051/itmconf/20257002013.