

# Analysis of British Airways Reviews – Case Study (The Forage, Data Science)

## Introduction

This case study focuses on analyzing customer reviews of **British Airways**, conducted as part of a data science project on **The Forage** platform. The objective was to perform **web scraping** from the **Skytrax** review website, clean and analyze the text data using various **Natural Language Processing (NLP)** methods, and interpret the results.

The key goals of this analysis were:

- **Extract reviews related to British Airways,**
- **Analyze the sentiment of the reviews (positive/negative),**
- **Visualize the most common words using a WordCloud,**
- **Eliminate common words without emotional significance (stopwords) for more accurate analysis.**

The outcome is a comprehensive study of customer sentiments regarding British Airways and the most frequently expressed opinions.

---

## Analysis Process

### 1. Web Scraping Reviews Using BeautifulSoup

The first step was to collect review data from **Skytrax** using the **BeautifulSoup** library, which allows extracting text content from HTML pages.

The reviews were scraped from the following page:

 [Skytrax – British Airways Reviews](#)

This process involved:

- **Identifying the HTML structure of the website,**
  - **Extracting text reviews and their metadata (e.g., date, rating, verification status),**
  - **Saving the data into a CSV file for further analysis.**
-

## 2. Data Cleaning

Once the data was collected, it had to be **cleaned**. This process involved removing unnecessary elements from the text, such as:

- "✅ Trip Verified |"
- "❌"
- "✅ Verified Review |"
- Other special characters and excessive spaces

These elements had no relevance for sentiment analysis and only added noise to the data. After cleaning, the data was ready for further analysis.

---

## 3. Sentiment Analysis of Reviews

To determine whether reviews were **positive, neutral, or negative**, we used the **TextBlob** library.

Each review was converted into a **sentiment score**, where:

- **Values > 0** indicate a **positive review** (the higher, the more positive),
- **Value = 0** means a **neutral review**,
- **Values < 0** indicate a **negative review** (the lower, the more negative).

Based on this analysis, we found that:



**Percentage of positive reviews: 69.25%**

**Percentage of negative reviews: 30.24%**

A **sentiment distribution graph** was created to visualize the split between **positive and negative reviews**.

## Key Advantages of Sentiment Analysis

1. **Brand Reputation Monitoring** – Tracks public perception on social media, reviews, and news, helping companies detect and respond to crises early.
2. **Market Research & Competitor Analysis** – Identifies customer sentiment towards competitors, industry trends, and product positioning opportunities.
3. **Customer Experience Improvement** – Helps businesses understand pain points, improve service quality, and personalize marketing strategies.
4. **Social Media & Political Insights** – Analyzes public opinion on policies, elections, and social issues, detecting misinformation and trends.
5. **Automated Large-Scale Data Processing** – Saves time and resources by quickly categorizing text as positive, negative, or neutral.
6. **Financial Market Predictions** – Tracks investor sentiment and stock market trends based on news and social media analysis.
7. **Workplace & Employee Sentiment** – Analyzes internal feedback to improve company culture and prevent high turnover.

---

## 4. WordCloud – Visualization of the Most Common Words

Another part of the analysis was creating a **WordCloud** – a graphical representation of **the most frequently used words in reviews**.

To ensure that results were not skewed by **common words without emotional significance** (e.g., *the, in, and, to*), we applied a **stopwords filter**.

Stopwords included common **functional words, conjunctions, pronouns, numbers, and general terms** that have no emotional impact.

After removing these words, the WordCloud revealed frequently discussed topics such as **"British Airways," "business," "service," "return," "seat," "economy,"** helping identify key themes in customer feedback.

### Advantages of WordCloud: Why Use It?

- **Quick and Intuitive** – Highlights frequently occurring words in a text.
- **Easy to Interpret** – Visually appealing and simple to understand.
- **Great for Exploratory Data Analysis (EDA)** – Helps analyze text before deeper processing.
- **Aids Text Preprocessing** – Identifies stopwords, misspellings, and redundant words.
- **Supports Sentiment and Topic Analysis** – Shows key themes in reviews, social media, etc.

---

## Conclusion

The analysis of British Airways reviews provided valuable insights into how customers perceive the airline. Key findings include:

- **Most reviews were positive (69.25%), but a significant portion (30.24%) were negative.**
- **The most frequently discussed topics included business class, economy class, and customer service.**
- **The WordCloud method highlighted key words in the reviews, while stopwords helped remove less relevant expressions.**
- **Sentiment analysis confirmed that British Airways generally receives positive feedback, but there is still room for improvement.**

This study offers a **deeper understanding of customer perception** of British Airways and can be valuable for further improving the customer experience. 🚀