**Forage Online Data Science Internship**

**Sentiment Analysis on British Airways Reviews: Understanding customer’s through Customer Feedback**

*Author: Ameti’obong*



[This Photo](http://commons.wikimedia.org/wiki/file:boeing_777-236er_-_british_airways_(g-viid).jpg) by Unknown Author is licensed under [CC BY-SA](https://creativecommons.org/licenses/by-sa/3.0/)

British Airways (BA), the flag carrier airline of the United Kingdom, serves millions of passengers each year. Over 10 years, they have flown at least [36 million people](https://www.statista.com/statistics/309349/british-airways-uk-passenger-numbers/) and this number only reduced in 2020 due to the onset of the Covid-19 pandemic.

BA receives a myriad of feedback or reviews from its customers, ranging from praise to complaints. Understanding this feedback is crucial for any business aiming to improve its services and meet customer expectations.

This article delves into a sentiment analysis conducted on a dataset of BA customer reviews, offering insights into the sentiments and major themes within the feedback.

**Project Outline**

**Dataset Overview**

The dataset comprises 1,000 reviews from BA customers. This is only a subsection of thousands of reviews, and this was done strictly for learning purposes. Each review encapsulates the experiences and feelings of passengers regarding various aspects of BA's services, from booking to inflight amenities.

**Methodology**

1. **Data collection:** Forage provided a website (Skytrax) for data to be collected. Web scraping was done on this website to collect 1,000 reviews. Forage provided a set of code that was used for web scraping and the Python library used is BeautifulSoup which helps in web scraping.

#connect to airline website to scrape data on reviews

base\_url = "https://www.airlinequality.com/airline-reviews/british-airways"

pages = 10

page\_size = 100

reviews = []

# for i in range(1, pages + 1):

for i in range(1, pages + 1):

print(f"Scraping page {i}")

# Create URL to collect links from paginated data

url = f"{base\_url}/page/{i}/?sortby=post\_date%3ADesc&pagesize={page\_size}"

# Collect HTML data from this page

response = requests.get(url)

# Parse content

content = response.content

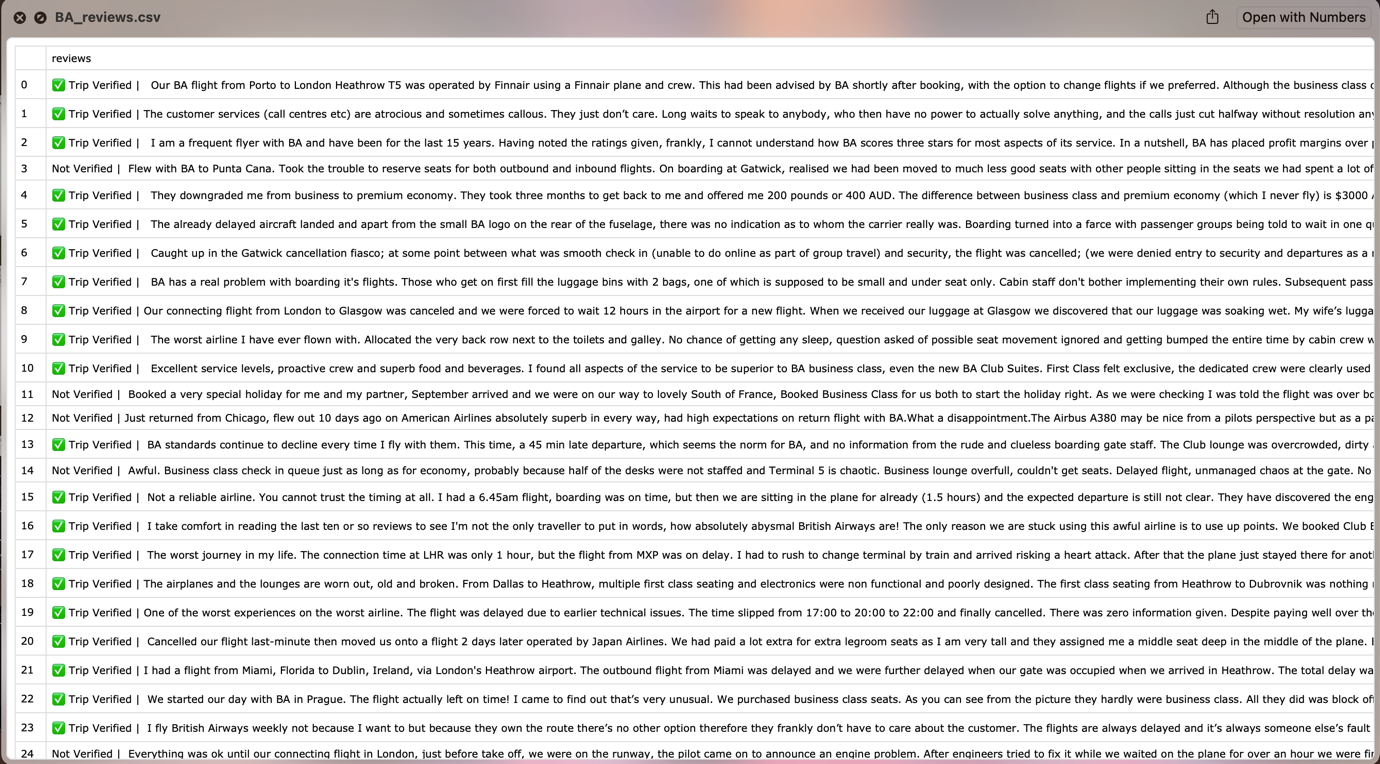
parsed\_content = BeautifulSoup(content, 'html.parser')

for para in parsed\_content.find\_all("div", {"class": "text\_content"}):

reviews.append(para.get\_text())

print(f" ---> {len(reviews)} total reviews")

**Code provided by forage for web scraping.**



**Screenshot of Dataset**

import os

# Define the directory path where you want to save the CSV file

directory\_path = '/Users/yummy/Desktop/Data'

# Ensure the directory exists, create it if it doesn't

if not os.path.exists(directory\_path):

os.makedirs(directory\_path)

# Specify the full file path, including the directory and file name

file\_path = os.path.join(directory\_path, 'BA\_reviews.csv')

# save the DataFrame to the CSV file

df.to\_csv(file\_path)

# Verify that the file has been saved

if os.path.exists(file\_path):

print(f"File saved successfully at: {file\_path}")

else:

print(f"File was not saved at: {file\_path}")

**Code for mapping dataset/file to a directory on the local device**

1. **Data Cleaning**: Before conducting the analysis, the dataset was first cleansed by removing the irrelevant column which is the Trip verified column. This column wasn’t needed for the analysis, so it was dropped. Symbols and numbers were also removed to make it easy for analysis. I also checked for missing values and there was none.
2. **Exploratory Data Analysis (EDA):** Wordcloud was used to generate the image below. This indicates the most frequent words present throughout the reviews. This didn’t provide lots of insight as seen below, but further analysis gave more insights.

****

**A graph with blue lines and a blue line to show distribution of words.
**

1. **Sentiment Analysis**: Leveraging VADER for Sentiment Analysis

Sentiment analysis is a crucial tool in understanding customer feedback, and it has several methodologies. One of the robust methods, especially for analyzing social media and short texts, is VADER (Valence Aware Dictionary and sEntiment Reasoner).

**Why VADER?:**

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool. It is a pre-trained model within the NLTK package. NLTK itself is a comprehensive Python library for natural language processing and text analytics. It's particularly adept at gauging the sentiment in English text, especially in environments like social media where short, nuanced messages are prevalent.

It is a part of the Natural Language Toolkit (NLTK), which is a leading platform for building Python programs to work with human language data. Within the NLTK package, VADER sentiment analysis can be accessed as **SentimentIntensityAnalyzer**.

**Key Features of VADER**:

* **Lexicon-based**: VADER has a predefined list of words, emojis, and emoticons with assigned polarity scores that it uses for sentiment analysis.
* **Handles Polarity and Intensity**: VADER provides compound scores that describe the combined polarity (positive, neutral, negative) and intensity of the sentiment.
* **Context Awareness**: It understands the context in which words are used. For instance, VADER knows that "not good" has a different sentiment than just "good".
* **Recognizes Slang and Emojis**: Modern communication often involves slang and emojis. VADER is equipped to handle these, making it particularly useful for analyzing social media content.

VADER, as a part of the NLTK library in Python, offers a powerful, context-aware toolset for sentiment analysis, especially for short and nuanced texts commonly found in social media.

VADER, with its unique approach to sentiment analysis, offered invaluable insights into the British Airways reviews dataset. By capturing sentiment intensity and context, it provided a more comprehensive view of customer feedback, enabling a clearer path for potential service improvements.

It illuminated specific areas of concern and praise for British Airways. While many reviews emphasized excellent service, others indicated room for improvement, especially in areas like customer support and boarding procedures. The intensity scores provided by VADER highlighted the strength of these sentiments, offering a deeper layer of insight.

**Results:**

**A bar chart with yellow and green squares

Description automatically generated**

**Sentiment Distribution**:

* **Negative Reviews**: Out of 1000 reviews, A significant portion of the reviews were negative (57% ~ >500), highlighting areas where BA might need improvements.
* **Positive Reviews**: A slightly smaller number of reviews were positive (41% ~ <500), indicating satisfied customers and areas where BA excels.
* **Neutral Reviews**: A minimal portion of feedback was neutral (2% ~ <50).

**Findings**:

1. **Booking & Customer Service**: Many reviews discussed issues related to booking and interactions with customer service.
2. **Boarding & Staff**: Another theme revolved around boarding procedures and staff interactions.
3. **Seating & Premium Experiences**: Reviews frequently touched on seating comfort and experiences in premium classes.
4. **Luggage & In-flight Experience**: Some passengers voiced concerns about luggage handling and overall in-flight experience.

**Conclusion & Recommendations**

The sentiment analysis provided a structured view into the vast array of feedback BA receives. While many customers expressed satisfaction, there's a clear indication of areas that need attention.

**Recommendations**:

1. **Improve Customer Service**: Given the prevalence of negative feedback related to customer service, BA should consider enhanced training for its customer service representatives.
2. **Review Boarding Procedures**: Streamlining the boarding process can alleviate some of the concerns raised by passengers.
3. **Enhance In-flight Experience**: Feedback about seating and in-flight amenities suggests room for improvement in these areas.

In today's digital age, customer feedback is invaluable. By leveraging sentiment analysis and other advanced forms of analysis, businesses like British Airways can gain deep insights into customer perceptions, enabling them to refine their services and better cater to their clientele.

*Do you have any experience with British Airways? Share your thoughts in the comments below!*

*Disclaimer: This analysis is based on a sample dataset and might not represent the complete range of customer feedback received by British Airways. The analysis was conducted as part of a free online internship program with* [*Forage*](https://www.theforage.com/)*. (*[*Forage Online Internship Program*](https://www.theforage.com/))Top of FormBottom of Form