Starbucks Food Review Analysis

FINAL REPORT

Loshana

TABLE OF CONTENTS

1.Introduction

Starbucks is a globally leading coffeehouse chain in the world. It has a rich history since 1971 when it opened its first store in Seattle's Pike Place Market. It Inspired by the seafaring tradition and Starbucks took its name from the famous novel "Moby-Dick."

The company's journey took a critical point when Howard Schultz drawn to Starbucks' coffee quality, joined the team in 1982. Schultz's visit to Milan in 1983 lighted his motive to bring the coffee culture of Italy to Starbucks, leading to the development of the brand into a coffeehouse by 1987.

Since then, Starbucks has expanded its impression across the globe, from the United States to Japan, Europe, and China. With a commitment to quality, it has become a beloved fixture in so many neighborhoods also welcoming millions of customers weekly.

Beyond profit, Starbucks is dedicated to making a positive impact on the communities. This commitment draws to uplifting people. Starbucks aims to be "people positive," prioritizing the well-being of partners, coffee farmers, and customers. Also, the company is working towards becoming "resource positive," striving to give back more to the planet than it takes.

Starbucks isn't just about serving coffee. It's about cherishing connections, supporting its partners, and making a change in this world.

Justification

In our case study, we lack information about how customers feel about the food offerings at Starbucks. Also, we are unsure about the reasons behind the reduced business during a certain season. By analyzing reviews, we can understand customer satisfaction and areas for improvement. Analyzing Starbucks review dataset gives important value for multiple reasons.

Insight into Customer Sentiment: Analyzing reviews provides valuable insights into how customers discover Starbucks' food offerings. By categorizing reviews as positive, negative, or neutral, can check overall sentiment and identify common themes, preference, or issues.

Identification of Areas for Improvement: Through sentiment analysis, can find specific areas where the food offerings are falling short or not meeting customer expectations. This helps Starbucks to prioritize improvements and enhancements to its menu options, services, quality, and presentation.

Customization of Offerings: By understanding customer desires, needs and preferences through analysis, Starbucks can get its food offerings to meet the preferences of its customer. This categorization can help attract new customers.

Enhanced Decision Making: Insights that get from reviews enable Starbucks to make decisions regarding menu development, marketing strategies, good service, and operational improvements. This data-driven approach arrange that resources are allocated to initiatives that will have the greatest impact on customer satisfaction and business success.

Competitive Advantage: By continuously analyzing customer feedback on food offerings and services, Starbucks can develop the highly competitive coffee and foodservice industry. And addressing issues and adapting to changing consumer preferences gives Starbucks a strengthens its brand reputation.

A food review case study for Starbucks food reviews provides valuable insights into customer sentiment, areas for improvement, and opportunities for customization. By identifying consumer feedback and reviews can enhance its food offerings, drive customer satisfaction, and achieve business growth.

2.Methodology

a. Process:

- The process involved collecting data from a dataset containing reviews of Starbucks food. Kaggle offers datasets with various domains. So, my dataset was acquired from the Kaggle platform and was imported into Python for further analysis. It begins by importing necessary libraries (pandas, NumPy, seaborn etc.)

- It has some basic exploratory data analysis (EDA) tasks like checking the basic information about the dataset (e.g., data types, null values). Then examining the dimensions (shape) of the dataset, checking null values, and removing them from the dataset. Also cleaned the 'Date' column and segregated them into 'month', 'year', and 'date' columns. Then renamed columns and reordered them. After that removed unsupported characters from reviews and converted them into sentence case. The text data was cleaned by converting it to lowercase, removing punctuation, eliminating stop words, and executing lemmatization and stemming procedures to enhance its quality for the analyses.

- Next, the sentiment analysis and categorical analysis were using the Vader sentiment analyzer and visualized compound scores by ratings. Also visualized positive, negative, or neutral scores by ratings. Those categorized ratings visualized using a pie chart and total rating count per month using a line plot too.

- Finally analyzed the frequency of ratings and visualized it using a bar plot. Additionally explored the most frequent locations mentioned in the reviews and visualized them using a horizontal bar plot.

These codes demonstrate a comprehensive analysis of Starbucks reviews, including data preprocessing, text cleaning, sentiment analysis, categorical analysis and various visualizations to gain insights into customer feedback and sentiments.

b. Assumptions

- The dataset used for analysis is representative of overall customer sentiments towards Starbucks foods and offerings. sentiment scores generated by the VADER analyzer accurately reflect the sentiment of the reviews and ratings.

c. Limitations

- The analysis depends on ratings and reviews. It does not take into other factors such as location, demographics, or purchase history.

d. Data Collection Instruments

- The primary instrument used for data collection was the Python programming language, along with libraries such as Pandas, NumPy, Matplotlib, and Seaborn for data manipulation, analysis, and visualization.

f. Data Collection Period

The data was collected over a period specified in the dataset, ranging from 2000 to 2023.

3.The Problem

a. Identify the problem.

Starbucks faces challenges like lack of customer feedback on food offerings and reduced business in certain periods.

- **Lack of Customer Feedback on Food Offerings:**

  Detailed information about how customers feel about Starbucks' food offerings and services is a significant problem. Without this feedback, Starbucks face challenges in insights into customer preferences, satisfaction levels, and areas for improvement.

- **Decreased business in certain seasons**:

  The reasons behind reduced business during certain seasons create a challenge for Starbucks. Without clarity on the factors contributing to this decline, it may struggle to implement solutions to mitigate its impact.

b. Explain why the problem is important.

- Understanding customer preference is important for maintaining and enhancing customer satisfaction and loyalty.
- Without customer feedback, Starbucks may miss opportunities to address customer concerns and improve their food offerings.
- Lack of insights into customer perceptions can delay decision-making processes related to menu development, offerings, services, marketing strategies, and operational improvements.
- Identifying the root causes of seasonal decresement in business is crucial for developing targeted strategies to address and overcome them.
- Without understanding the reasons behind fluctuations in the business Starbucks may continue to experience revenue losses during certain periods.
- Lack of insights into seasonal trends can hinder Starbucks' power to optimize resource allocation and maximize profitability throughout the year.
- Customer satisfaction is important for Starbucks' success and maintaining loyalty.

- Identifying areas for improvement in food offerings can enhance customer experience and sales.

c. How was the problem identified?

Lack of Information on Customer Sentiment Regarding Food Offerings:

This problem was identified by the acknowledgment of the lack of information about customer feelings towards Starbucks' food offerings.

Uncertainty about Reasons for Reduced Business During Certain Seasons:

It mentions the reasons behind reduced business during certain seasons, indicating a need for further investigation and analysis.

Limited Insight into Customer Preferences and Expectations:

It is figured the justification, as it maintains the importance of analyzing reviews to identify areas for improvement and customization based on customer needs and preferences.

Challenges in Decision Making:

The need for enhanced decision-making is highlighted. where it mentions the importance of insights gained from reviews in guiding menu development, marketing strategies, and food offerings.

Risk of Falling Behind Competitors in the Industry:

The importance of maintaining a competitive advantage is emphasizing the need for Starbucks to continuously analyze customer feedback to stay ahead of competitors.

D. Assessment of the Effectiveness of the Problem Identification Process:

The process for identifying the problems was highly effective. By recognizing the customer feedback on food offerings and the seasonal business. I was able to identify critical areas that require attention. Through a thorough analysis of the situation. And set the stage for implementing data-driven solutions to address these challenges and drive business success.

4.Steps Taken to Address the Problems:

Enhancing Customer Feedback Mechanisms for Food Offerings: Starbucks should implement systematic review analysis of customer feedback from various sources like social media, utilize sentiment analysis techniques to categorize customer opinions towards food offerings and conduct targeted customer satisfaction surveys focusing on food preferences help to gather direct feedback.

Understanding Reasons for Reduced Business During Certain Seasons: should analyze historical sales data to identify seasonal patterns in business performance, conduct market research to understand external factors impacting customer behavior and utilize predictive modeling techniques to anticipate factors contributing to reduced business during specific seasons.

Gaining Deeper Insight into Customer Preferences: Starbucks should expand data collection efforts to capture a wider range of customer preferences. By employing data analytics tools, Starbucks can identify recurring themes and trends among its customer base, enabling the company to better understand and address their needs. Also collaborating with culinary experts to translate feedback into actionable menu improvements.

Leveraging Data-driven Insights for Decision Making: For the future success Starbucks should prioritize data-driven insights for decision-making. Implementing data visualization techniques for effective communication of insights. And conducting regular performance evaluations to monitor the impact of data-driven decisions.

Staying Ahead of Competitors in the Industry: To ensure future success, Starbucks must conduct thorough competitive analysis to benchmark its offerings. Continuously monitoring market trends and responding to online reviews also helps the business.

5.The Results and discussion

Sentiment analysis

VADER

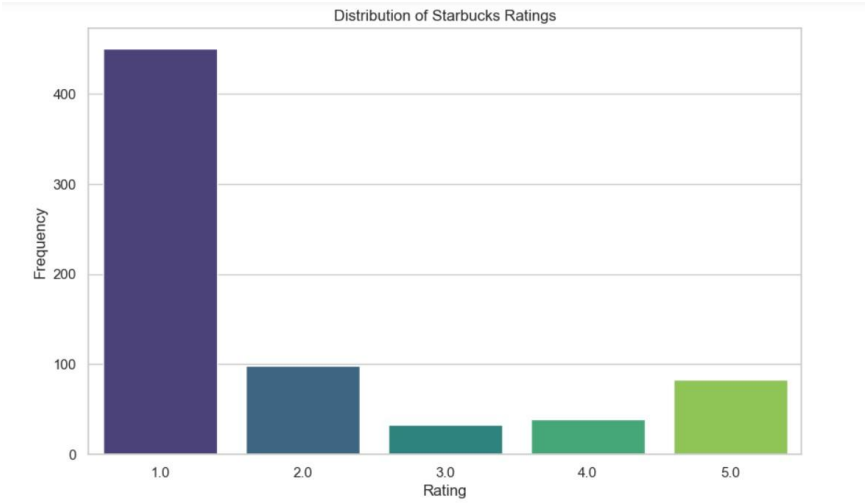
VADER (Valence Aware Dictionary and sentiment Reasoner) is a sentiment analysis library designed specifically for analyzing the sentiment of text. It provides a pre-trained model that can classify text into positive, negative, or neutral sentiment based on the words.

	neg	neu	pos	compound
Helen	0.000	0.797	0.203	0.8991
Courtney	0.101	0.757	0.142	0.7344
Daynelle	0.088	0.773	0.139	0.4215
Taylor	0.086	0.691	0.223	0.9232
Tenessa	0.000	0.669	0.331	0.9793

It's clear that our data is imbalanced in which the majority of review tends to be negative, this indicates that the company must review and improve its customer service by identifying and investigating problem areas. This helps to identify a customer is satisfied with the service or not!

Distribution of Starbucks Ratings

It's clear that our data is imbalanced in which the majorety of review tends to be negative , this indicates that the company must review and improve its customer service by identifing and investigating problem areas.



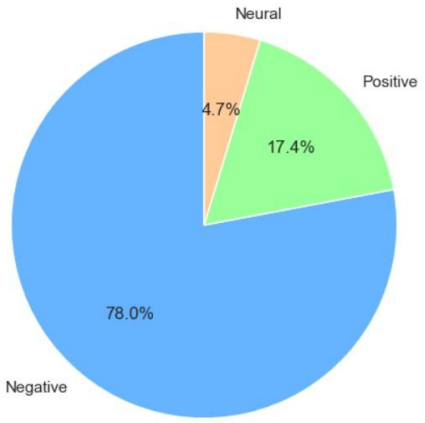
Categorical analysis

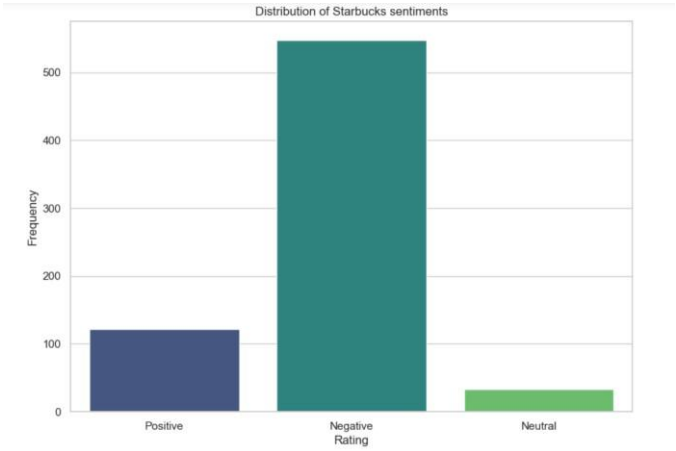
Positive, Neutral and Negative score by Rating

After dealing with our EDA, we can notice that our dataset Have mostly Negative Reviews which comes about 78% of the total dataset observations.

we splinted our data into 3 categories: Positive, Negative and Neutral in which Neutral:

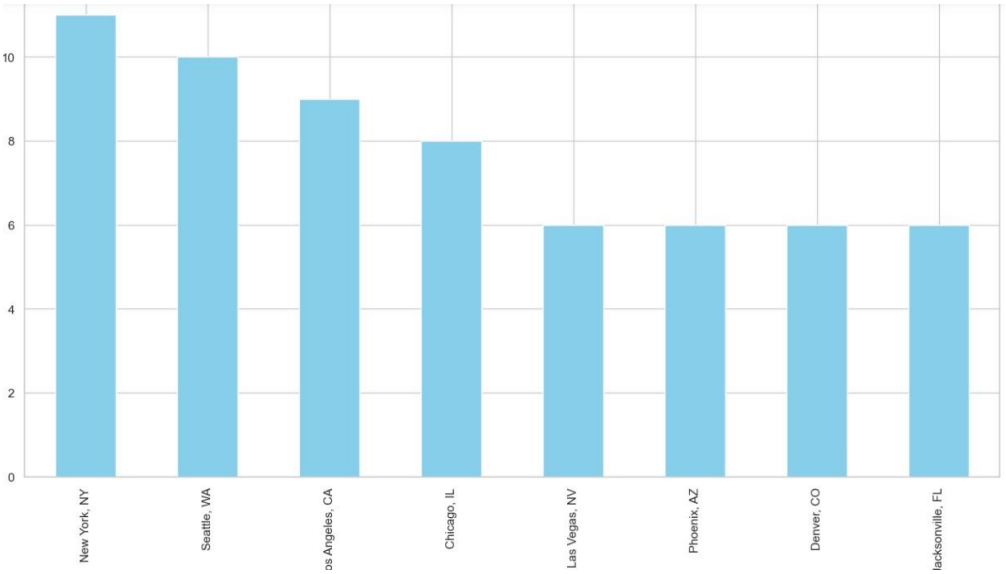
- 0 Positive: 1 Negative: 2





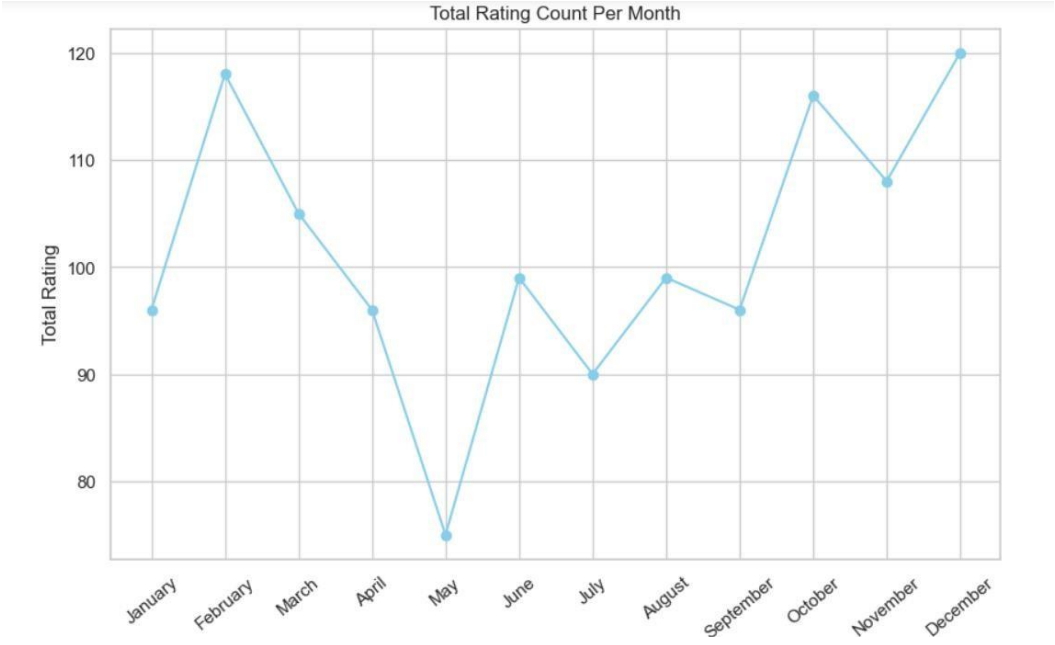
Most frequent locations in our dataset

We can notice that there is a peak in New York. But it's true that this chart gives us the variations of locations, but this doesn't give us a perspective about the relation between locations and Ratings.



Total Rating Count Per Month

The rating is higher in the early part of the year. Then there is a significant increase in the second half of the year and starting in July and peaking in December.



6.Conclusion

In conclusion, our analysis of Starbucks reviews revealed several key insights. Firstly, we observed that most reviews tend to be negative, indicating potential areas for improvement in customer service and overall satisfaction. This suggests the importance of addressing problem areas to enhance the customer experience.

Additionally, we found that sentiment analysis using VADER showed imbalanced data, with negative reviews dominating. This underscores the need for Starbucks to identify and investigate problem areas to improve customer satisfaction.

Furthermore, our analysis of ratings distribution showed fluctuations throughout the year, with higher ratings observed in the second half, particularly peaking in December. This suggests potential seasonal trends in customer satisfaction that Starbucks could leverage in its strategies.

Overall, by leveraging insights from sentiment analysis and ratings distribution, Starbucks can prioritize improvements in customer service, address problem areas, and capitalize on seasonal trends to enhance customer satisfaction and drive business success.

7.Recommendations

Enhance Customer Service: Prioritize addressing problem areas highlighted in negative reviews to improve overall satisfaction.

Investigate Problem Areas: Conduct thorough investigations into issues raised by customers to identify root causes and implement effective solutions.

Leverage Seasonal Trends: Capitalize on seasonal fluctuations in ratings by adjusting strategies and offerings accordingly.

Employee Training: Provide comprehensive training to employees to ensure consistent and high-quality customer service.

Continuous Monitoring: Implement regular monitoring and feedback mechanisms to track changes in customer sentiment and address issues promptly.

8.Appendices

The link to the dataset that have been used throughout the analysis –

https://www.kaggle.com/datasets/harshalhonde/starbucks-reviews-dataset

All the codes used to analyze the data are provided for reference. Below are the Python codes snippets used for data analysis and visualization:

```
# Importing the libraries
 import pandas as pd import
numpy as np from tqdm.notebook
import tqdm tqdm.pandas()

import seaborn as sns import
matplotlib.pyplot as plt import
plotly.express as px import
plotly.graph_objs as go import
plotly.figure_factory as ff
 import string import re from
collections import Counter
import nltk import transformers
# # Importation of data and EDA df =
pd.read_csv('D:/DATA CASE STUDY/reviews_data.csv')
df.head()  df.info()
```

```
#Dimension of the data.
df.shape
df.columns.tolist()
 import seaborn as sns from
matplotlib import pyplot as plt
sns.set()
plt.style.use('ggplot')

#Checking for null values df.isnull().sum()

# Let's exclude null values.
df = df.dropna()

# Check the unique values and frequency for 'Rating'
 df['Rating'].value_counts()

# dropping links df.drop('Image_Links',
axis=1, inplace=True) df.head()

# remove entries with no review df =
df[df['Review'] != 'No Review Text'] df =
df.reset_index(drop=True) df.head()

# Importing necessary module from
dateutil import parser

# expanding date
df['Date'] = df['Date'].str.replace('Reviewed', '').str.replace('[.,]', '')

# parsing dates into standard form df['Date']
= df['Date'].apply(parser.parse)

# segregating dates
df['month'] = df['Date'].dt.strftime('%b')
df['year'] = df['Date'].dt.year df['date']
= df['Date'].dt.day

# dropping old date column
df.drop('Date', axis=1, inplace=True)
 df.head()

# renaming columns df.columns =
['name','location','rating','review','month','year','date']
```

```
# rearanging columns df =
df[['name','location','date','month','year','review','rating']]
 df.head()

# function to remove unsupported characters from reviews
def clean_review(text):
    return ''.join(char for char in text if char.isalnum() or char.isspace())
# removing unupported characters df['review'] =
df['review'].apply(clean_review)

# making all statements sentence case df['review'] =
df['review'].apply(lambda x: x.capitalize())
# viewing data df.head()

# Basic statistical overview basic_stats
= {
    "Total Reviews": df.shape[0],
    "Unique Locations": df['location'].nunique(),
    "Date Range": (df['date'].min(), df['date'].max()),
    "Rating": {
    "Average Rating": df['rating'].mean(),
    "Min Rating": df['rating'].min(),
    "Max Rating": df['rating'].max()
    }
}
basic_stats

# # Sentiment analysis of the reviews
# Importing the Vader sentiment analyzer
from nltk.sentiment import SentimentIntensityAnalyzer sia
= SentimentIntensityAnalyzer()

# Applying the vader sia on all the dataset
 res = {} for i,row in tqdm(df.iterrows() , total =
len(df)):
    text = row['review']     myname =
row['name']     res[myname] =
sia.polarity_scores(text)

#Stocking the results in a dataframe vader_df
= pd.DataFrame(res).T vader_df.head()
 data = df.merge(vader_df.reset_index(names = 'name')  , how =
'left') data.head()
```

```
# # Rating def
pos_neg(x):
if x > 3:
        return "Positive"
elif x < 3:
        return "Negative"
else:
        return "Neutral"
 data["rating"] =
data["rating"].apply(pos_neg) data.head()

# Setting up the aesthetics for plots
sns.set(style="whitegrid")

# Plotting the distribution of ratings
plt.figure(figsize=(10, 6)) sns.countplot(x='rating',
data=data, palette="viridis") plt.title('Distribution
of Starbucks Ratings') plt.xlabel('Rating')
plt.ylabel('Frequency') plt.show()
df["rating"]=df["rating"].replace([1,2],2)
df["rating"]=df["rating"].replace(3,0)
df["rating"]=df["rating"].replace([4,5],1)
df = df.dropna(axis = 0, how ='any')
df["rating"]=df["rating"].astype(int)
 df["rating"].value_counts()
  plt.figure(figsize=(6, 6)) sns.set_style("whitegrid")
plt.pie(df["rating"].value_counts(),labels=["Negative","Positive","Neural"],
autopct='%1.1f%%',colors=['#66b3ff', '#99ff99', '#ffcc99'] , startangle=90);

# # Location
plt.figure(figsize=(16,8))
sns.set_style("whitegrid")
df["location"].value_counts().sort_values(ascending=
False).head(8).plot.bar(color='skyblue')
plt.title("Most frequent locations in our dataset");

# # Months df = df.groupby(['month'])['rating'].sum().reset_index()
month_order = ['January', 'February', 'March', 'April', 'May', 'June', 'July',
'August', 'September', 'October', 'November', 'December']
 plt.figure(figsize=(10, 6)) plt.plot(df['month'], df['rating'], marker='o',
color='skyblue', linestyle='-') plt.xlabel('month') plt.ylabel('Total Rating')
plt.title('Total Rating Count Per Month')

plt.xticks(df['month'], labels=month_order,rotation=40)
plt.grid(True) plt.show()
```

```
# # Categorical analysis
ef pos_neg(x):     if x
> 3:
        return "Positive"
elif x < 3:
        return "Negative"
else:
        return "Neutral"

data["rating"] = data["rating"].apply(pos_neg) data.head()

# Setting up the aesthetics for plots
sns.set(style="whitegrid")

# Plotting the distribution of ratings
plt.figure(figsize=(10, 7)) sns.countplot(x='rating',
data=data, palette="viridis") plt.title('Distribution
of Starbucks sentiments') plt.xlabel('Rating')
plt.ylabel('Frequency') plt.show()
```

9.Refferences

(Chuang, 2019)Chuang, H.-J. (2019). Starbucks in the World. *HOLISTICA – Journal of Business and Public Administration*, *10*(3), 99–110. https://doi.org/10.2478/hjbpa-20190031

KevinCoquitlam, Falls, H., CourtneyApopka, Twp, D., TaylorSeattle, JuanKissimmee, HadisehHouston, ChristinaCokeburg, Hills, S., & Beach, S. (n.d.). *Starbucks*. ConsumerAffairs. https://www.consumeraffairs.com/food/starbucks.html

Oraji, A. (2023b, December 1). *Unveiling the Starbucks experience: A data-driven journey*. Medium. https://medium.com/@ali.oraji/unveiling-the-starbucks-experience-a-datadriven-journey-3ce0a1b40d18

KevinCoquitlam, Falls, H., CourtneyApopka, Twp, D., TaylorSeattle, JuanKissimmee, HadisehHouston, ChristinaCokeburg, Hills, S., & Beach, S. (n.d.). *Starbucks*. ConsumerAffairs. https://www.consumeraffairs.com/food/starbucks.html

GfG. (2021, November 2). *Python: NLP analysis of Restaurant Reviews*. GeeksforGeeks. https://www.geeksforgeeks.org/python-nlp-analysis-of-restaurant-reviews/

(No date) *Starbucks launches winter menu with new iced hazelnut …* Available at: https://stories.starbucks.com/press/2024/starbucks-launches-winter-menu-withnew-iced-hazelnut-oatmilk-shaken-espresso-and-the-return-of-pistachio (Accessed: 22 March 2024).

Ali, M. (2023) *NLTK sentiment analysis tutorial: Text mining & analysis in python*, *DataCamp*. Available at: https://www.datacamp.com/tutorial/text-analyticsbeginners-nltk (Accessed: 22 March 2024).

Business studies (no date) *StudySmarter UK*. Available at: https://www.studysmarter.co.uk/explanations/business-studies/business-casestudies/starbucks-marketing-strategy/ (Accessed: 22 March 2024).

McKinnon, T. (2024) *The reasons behind Starbucks' great customer experience*, *Indigo9 Digital Inc.* Available at: https://www.indigo9digital.com/blog/starbuckscustomerexperience (Accessed: 22 March 2024).

Shrivastava, D. (2022) *Starbucks case study: Starbucks leads the Coffee Industry*, *StartupTalky*. Available at: https://startuptalky.com/starbucks-case-study/ (Accessed: 22 March 2024).

Starbucks delivering Customer Service (2016) *SlideShare*. Available at: https://www.slideshare.net/zhjermia/starbucks-delivering-customer-service58568775 (Accessed: 22 March 2024).

Starbucks: A customer loyalty story to learn from (no date) *NPS Benchmarks*. Available at: https://customergauge.com/benchmarks/blog/starbucks-a-story-of (Accessed: 22 March 2024).

Villegas, F. (2023) *The starbucks customer experience: Brewing success*, *QuestionPro*. Available at: https://www.questionpro.com/blog/starbucks-customer-experience/ (Accessed: 22 March 2024).