Research Article

Exploring Global Twitter Sentiments on Millet Confections

Journal of Tourism, Hospitality & Culinary Arts (JTHCA)
2023, Vol. 15 (2) pp 60-89

Reprints and permission:
UiTM Press
Submit date: 29th August 2023
Accept date: 24th October 2023
Publish date: 30th December 2023

Parminder Singh Dhillon*
Department of Tourism, Hospitality & Hotel Management, Punjabi University, Patiala, Punjab, India
parminderhm@pbi.ac.in

Proposed citation:
Dhillon, P.S. (2023). Exploring Global Twitter Sentiments on Millet Confection. *Journal of Tourism, Hospitality & Culinary Arts, 15*(2), 60-89

Abstract
To assess the market potential of millet-based confections, this study aimed to elucidate the attitudes, opinions, and ongoing discussions of social media users from the top ten economies of the world. Text analysis and machine learning approaches were applied to classify the sentiments of the tweets as positive, neutral, or negative. The Naive Bayes classifier was applied to improve the precision of sentiment analysis. The process revealed the top themes presented as the top 50 phrases by their frequency in tweet data. Furthermore, the subjectivity distributions and polarity in the results provide intricate emotional perspectives. Additionally, a bar chart was used to display the distribution of positive, neutral, and negative tweets while word popularity was visualized through word clouds and word pair clouds. The insights from this research are valuable for businesses working with millets, marketers, and legislators globally.

Keywords:
Sentiment analysis, Naïve Bayes classifier, millet-based confectionery, Twitter, word cloud, text analysis, millets.

1 Introduction

Food is a vital element to survive. While examining dietary habits, the primary emphasis is on understanding what individuals consume (Wirfält et al., 2013). There are several factors that help in identifying the dietary habits of individuals; some important factors are natural components, price, well-being, mood, ethics, availability, sensory qualities, taste, and convenience (Lima et al., 2021). Consumer's food habits are mostly shaped by their own circumstances and feelings. People who work long hours tend to consume more packaged food because they see it as being more convenient (Okrent & Kumcu, 2016). The consumption of foods with heavy fat, sugar, and salt increases the chance of noncommunicable diseases (Mediratta & Mathur, 2023). However, nowadays, people are becoming health conscious and adopting good food habits. According to Kraft and Goodell (1993), health-conscious consumers (HCCs) are people who live "wellness-oriented" lives and are worried about their surroundings, stress, and fitness. As a result, this consumer group's food consumption behavior is influenced by several motivational factors.

Social media channels are among the best places to look for peer-generated knowledge, as they offer people a place to talk to one another on a variety of issues (Massagli et al., 2010). The usage of social media has tremendously increased among teenagers and adults in the age group 20 to 45 (Kaplan & Haenlein, 2010). Blogs, chat rooms, discussion boards, forums, micro-blogging services, and social networking websites are just a few examples of the various online tools collectively known as "social media" which allow users to communicate with one another while creating their own profiles (Mangold & Faulds, 2009). With over 330 million active users worldwide, Twitter has established itself as a prominent platform for sharing daily activities and expressing opinions (Karimiziarani et al., 2022). It is now one of the most popular platforms in the social media landscape, catching the attention of marketing and consumer science researchers (Mangold & Faulds, 2009; Kontopoulos et al., 2013; Worch, 2014; Carr et al., 2015).

Food and drinks consumption plays a significant role in human existence (Köster, 2009). Twitter presents a unique opportunity for studying consumer food trends due to its abundant user-generated content (Caiazza & Bigliardi, 2020). Academics and companies seeking a deeper understanding of eating habits and preferences can extract valuable insights from the abundant food-related information available online (Tao et al., 2020). Social media platforms including Twitter facilitate the collection of user-specific data, such as geographic location, enhancing the ability to gather regional insights into food preferences and behaviours (Croitoru et al., 2013). Leveraging Twitter data allows researchers to gain better understanding of customer behaviours, assisting companies in adapting their strategies to meet evolving consumer expectations (Krafft et al., 2021). Social media platforms like Twitter are increasingly popular for discussing eating preferences and routines (Citrin et al., 2003). Researchers and companies can access crucial information on consumer behaviour related to food choices by exploring this social media platform (Vidal et al., 2016).

According to Garcia-Closas et al. (1997), 'confectionery' refers to sugary and sticky meals. It's divided into two main subcategories: bakers' confectionery and sugar confectionery, both emphasizing sweet, carbohydrate-rich foods (Caggia et al., 2020). Sugar confectionery includes toffees, chocolates, candied nuts, bubble gum, etc., while baker's confectionery comprises cakes, pastries, and baked goods. A recent statistical analysis by Statista (2022) revealed that preserved pastries and cakes were the most consumed confectionery category, accounting for over 62.57 billion kilograms in the global market in 2021. Chocolate candy was the least popular category, but Statista projects increased consumption levels across all categories in the coming years.

It's important to recognize that many confections contain wheat and sugar, which can lead to elevated blood sugar levels (Konnar et al., 2022). Sweets often include high-fructose corn syrup, a common ingredient linked to obesity (Forshee et al., 2007). Therefore, it's crucial to monitor and control confectionery consumption (Goryńska-Goldmann et al., 2021) since excessive intake can result in allergies, type-2 diabetes, tooth decay, and other health issues (Tandel, 2011). The growing awareness of health has prompted individuals to be more mindful of their diets and the impact of their food choices on overall well-being (Elbel et al., 2009). While candies and similar treats may be tempting, it's essential to consider the potential adverse effects of excessive sugar and processed carbohydrate consumption (McCrickerd et al., 2016).

This study primarily aims to address the following research questions: How are millet-based confections perceived and discussed in the top 10 global economies? How do tweets contribute to attitudes related to millet-based confectionery? Researchers can gain insights into the factors influencing people's decisions on consuming confections and their attitudes toward health by analyzing user-generated content on Twitter (Saura et al., 2020). For researchers and organizations seeking a better understanding of customer behavior and food-related attitudes, Twitter sentiment analysis proves to be a valuable tool (Singh et al., 2018). It enables the discovery of patterns, preferences, and emotions related to various types of sweets by examining tweets about confectionery and meal options (Dondokova et al., 2019). Sentiment analysis can reveal favorable or negative sentiments associated with specific brands or varieties of confections, assisting companies in making informed decisions about product development and marketing strategies (Humphreys & Wang, 2018). Additionally, it sheds light on aspects like flavor, cost, health benefits, and social norms that influence people's food preferences (Nesmith, 2020). This data can help firms tailor their messages and marketing strategies to target audiences (Smith, 2011). Academics can also use sentiment analysis to understand how consumers perceive healthy eating and identify the barriers preventing them from making healthier choices (Wankhade et al., 2022). Overall, sentiment analysis on Twitter provides valuable insights for organizations and researchers to develop strategies promoting healthy eating and enhancing customer well-being (Wankhade et al., 2022).

2 Literature Review

2.1 Health properties of Millet based confectionary

The term 'millet,' a French word meaning 'a handful of millet contains thousands of grains' (Taylor & Emmambux, 2008), refers to one of the earliest grains used for household purposes and among the oldest cereals known to humanity (Michaelraj & Shanmugam, 2013). Millet is a highly nutritious grain that is rich in fiber, promoting digestive health and regulating bowel movements. It is also a source of protein with antibacterial properties, offering health benefits such as aiding in the management of conditions like obesity, diabetes, and high cholesterol levels (Veena, 2003). Millets are rich in micronutrients, including niacin, Vitamin B6, and folic acid (Pathak, 2013; Habiyaremye et al., 2017). Anitha et al. (2021) suggests that millet may reduce the risk of type 2 diabetes and help manage blood glucose levels. Fatma et al. (2016) examined the health benefits of gluten-free confections and cakes made with rice flour enriched with germinated millet flour. The study compared samples with varying proportions of millet flour, concluding that using germinated millet flour (GMF) enhanced the nutritional content of the baked goods, including fat, protein, zinc, iron, calcium, phenolic compounds, and flavonoids.

Due to the rise in health-conscious eating habits, businesses have responded to fulfill customer demands by offering wholesome and sustainable solutions (Olayanju, 2019). The demand for millet-based products is increasing due to their nutritional and environmental advantages (Shah et al., 2021). Policymakers and international organizations have also shown interest, sparking the attention of entrepreneurs in the millet-based goods market (Chera, 2017). Food entrepreneurs have embraced millet as a key ingredient in response to changing customer preferences for healthier options (Kline et al., 2014; Ganguly, 2018). Businesses are adapting to the dynamic market demand by forming partnerships, collaborating with suppliers, and streamlining distribution channels despite facing various challenges (Talwar et al., 2021). The dynamic market and changing consumer demand underscore the potential for millet-based confections in culinary choices.

2.2 Social media and food product perception

Social media platforms serve as a convenient means for individuals to express their attitudes, opinions, and experiences. The growth in Twitter's active user base and the wealth of user-generated content it hosts provide valuable opportunities for scientific investigations of public opinion (Pater et al., 2016). To gain a deeper understanding of consumer sentiments and perspectives across various sectors of society and the economy, it is crucial to analyze the ideas and information present in user-generated content disseminated through social media (Mostafa, 2013; Yu et al., 2013). Examining public opinions shared openly in social media is particularly essential for comprehending consumer dietary choices (Hemmerling et al., 2015; Danner & Menapace, 2020).

The Twitter micro-blogging app has emerged as a valuable source of information for studying human behavior through data mining methods (Hilbert, 2016). Researchers have effectively utilized this platform for various studies, including investigations into consumer consumption, restrictions, and behavior (Paschen et al., 2020). Notably, Twitter users not only provide the content of their published messages but also share supplementary data, such as their location (Stefanidis et al., 2013). This location data conveys contextual information about people's views and preferences, in addition to their geographic information (Stefanidis et al., 2013).

Social media has evolved into a significant communication medium, with people sharing insights into their daily activities, observations, and lifestyles online. Consequently, social media networks are increasingly harnessed to gather insights into customer sentiments (Dodds et al., 2011) and beliefs (Dwivedi et al., 2021). Moreover, customers are more inclined to express their opinions on products through social media compared to traditional questionnaire surveys (Rathore et al., 2016). Social media platforms serve as arenas for information exchange, communication, and even conflict (Nomisma, 2019). By analyzing tweets, businesses can gain valuable insights into how customers perceive specific topics or products.

2.3 Sentiment analysis

A well-established technique known as sentiment analysis (Pang & Lee, 2008) is employed to determine the polarity of opinions expressed in a specific user's review. The polarity of a review is typically categorized as positive, neutral, or negative based on the mood conveyed within the text (Hu & Liu, 2004; Liu, 2012). Sentiment analysis, which processes natural language, is used to monitor public opinions regarding a product or issue (Wankhade, 2022). It assesses beliefs, attitudes, judgments, feelings, and sentiments conveyed through written language (Liu, 2012). Sentiment analysis serves as a valuable tool for enhancing the quality of products and services in commercial decision-making and facilitating information retrieval for individuals (Bueno et al., 2002). It is a crucial technique for examining the relationship between food and emotions, a topic that has long been studied to understand human behavior. The surge in social media usage and the importance of sentiment analysis are intertwined. Both individuals and organizations utilize information from various sources when making decisions (Liu, 2012).

2.4 Early Works on Sentiment Analysis in the Food Domain

Many researchers have used Twitter data to study various topics in the food and beverage domain which includes sentiment analysis, consumer preference, and other food-related aspects like food quality or health disorders arising from food habits. These qualitative studies on Twitter data give insights into the sentiments of people surrounding food, diet, and culinary experiences. Some of the prominent research works in the food domain have been presented in Table 1 with their objectives and techniques.

Twitter data has been extensively employed by numerous researchers to investigate a wide range of topics within the food and beverage domain, encompassing sentiment analysis, consumer preferences, and various aspects related to food, such as food quality and health issues stemming from dietary habits. Qualitative studies using Twitter data offer valuable insights into the sentiments of individuals concerning food, dietary choices, and culinary experiences. An overview of notable research endeavors in the food domain is provided in Table 1, with their objectives and the techniques employed to achieve those objectives outlined.

Table 1: Previous works on sentiment analysis in food domain

Study	Aim	Types of Techniques
Zhou et al. (2020)	To identify eating disorder-related topics and explore factors associated with eating disorders using a topic modeling approach.	The classification was performed using CNN, LSTM, SVM, and Naïve Bayes, and the analysis was conducted using the CorEx topic modeling method.
Tibra et al. (2017)	To study the popularity of healthy foods and to predict geo-located food-based sentiment analytics.	Popularity is examined through data mining and feature-based sentiment analysis, achieved using a machine learning algorithm.
Oduru et al. (2022)	An attempt was made to predict the healthiness level of food images.	An image classifier, RESNET152, was applied to achieve the research objectives in the field of Deep Learning.

Wang et al. (2019)	To investigate dropout behaviors in individuals and estimate the causal effects of personal emotions and social networks on these behaviors.	The tweets were collected using the snowball sampling method. Emotions were quantified employing an automated sentiment analysis tool, and network centralities were assessed based on users' following networks. Linear and survival regression instrumental variables models were applied to estimate the effects of emotions and network centrality on dropout behaviors.
Sukunesan et al. (2021)	To investigate the engagement of Pro-Eating Disorder (Pro-ED) communities on Twitter.	Posting styles across various user metrics are compared using Mann-Whitney U tests. The factors of Pro-ED Twitter users are subsequently examined through the application of linear models.
Samoggia et al. (2020)	To study the content and sentiment of Twitter messages related to health attributes by examining consumer perceptions and emotions.	The objective was achieved through the application of frequency analysis, keyword-in-context analysis, and sentiment analysis based on the Lexicon approach.
Vydiswaran et al. (2020)	To examine the validity of Twitter as a source of information for neighborhood-level analysis of dietary choices and attitudes	Keyword analysis is performed to identify differences in food healthiness between the most and least affluent tracts.
Mostafa (2017)	To investigate the sentiments towards halal food.	Lexicon based sentiment-analysis was applied.
Vatambeti et al. (2023)	Analyzing customer sentiments towards different app-based meal delivery brands.	A sentiment analysis was applied, employing a framework that combines Convolutional Neural Network (CNN) and Bi-directional Long Short Term Memory (Bi-LSTM) models to achieve the research objectives.
Pindado and Barrena (2021)	To study the social representations of various regions worldwide regarding emerging food trends.	Regions with consumers sharing content about food trends are identified using a density-based clustering algorithm. The attitude is addressed through sentiment analysis, and sub-regional differences are explored through the use of grid maps.
Eskandari et al. (2022)	To investigate conversations about food poverty on Twitter during the COVID-19 pandemic and the subsequent national lockdown.	Tweets containing the terms 'food' and 'poverty' were gathered within specific date ranges using NodeXL Pro software. Content analysis was conducted by applying social network analysis tools to achieve the research objectives.

Sproagis and Rikters (2020)	To introduce the Latvian Twitter Eater Corpus, a compilation of tweets pertaining to food, beverages, eating, and drinking.	Corpus-based sentiment classification and analysis is employed to achieve research objectives.
Nova et al. (2022)	To understand individual-level influences over discrete online eating disorder (ED).	Objectives are achieved through the application of a combination of network analysis and content analysis, utilizing computational topic modeling and qualitative thematic analysis.
Vydiswaran et al. (2018)	To examine the level of "healthiness" in tweets related to food and analyze the sentiments expressed in those tweets originating from the metropolitan Detroit area.	Classification techniques were applied to achieve the objectives, including: (a) Support Vector Machine (SVM), (b) Random Forest, (c) Multinomial Naïve Bayes, and (d) Logistic Regression with the Stochastic Gradient Descent method.
Hamshaw et al. (2018)	To investigate how the intention of food hypersensitivity customers is influenced by specific characteristics of tweets.	Hypersensitivity was assessed using the Elaboration Likelihood Model and the Uses and Gratifications Theory.
Recuero-Virto and Valilla-Arróspide (2022)	To understand consumers' perceptions of food-related technologies.	To comprehend consumers' perceptions of food-related technologies, sentiment analysis was conducted utilizing a machine learning approach, and precision was assessed using Krippendorff's alpha value (KAV).
Singh and Glińska-Neweś (2022)	To understand consumer attitudes towards sustainable food consumption.	To comprehend consumer attitudes toward sustainable food consumption, tweets were gathered using the streaming Application Programming Interface (API), and topic modeling was accomplished through the application of the Latent Dirichlet Allocation (LDA) algorithm. Subsequently, sentiment analysis was conducted employing the Syuzhet package, utilizing Lexicon-based techniques.
Singh et al. (2018)	To identify supply chain management issues in food industries focusing on customer feedback and flow/quality issues of food products.	To address supply chain management issues in the food industry, particularly those related to customer feedback and the flow/quality of food products, Twitter data was analyzed using text analysis techniques. This involved applying a support vector machine (SVM) and hierarchical clustering with multiscale bootstrap resampling to identify clusters of words associated with supply chain issues.

Cooper et al. (2022)	To understand consumer perception of the value proposition of vegan food, particularly in relation to ethical, personal health, and environmental factors.	The predominant themes of conversation surrounding vegan food were identified through the application of Leximancer analysis.
Nguyen et al. (2017)	To build a national food environment database using geo-tagged Twitter and Yelp data and to test associations between state food environment indicators and health outcomes.	Tweets were classified into happy and not happy categories using the Machine Learning approach and the Maximum Entropy text classifier in MALLET.
Vidal et al. (2016)	To analyze consumers' spontaneous expressions of food-related emotions.	Consumers' spontaneous food-related emotional expressions were analyzed. Numbers of emojis in the content were manually calculated, and the Chi-square test was applied to identify differences in the percentage of tweets with emojis.
Benítez-Andrades et al. (2022)	To categorize tweets related to eating disorders, machine learning models were applied.	To categorize tweets related to eating disorders, machine learning models, including Classification with ML and BERT model, were applied.
Pater et al. (2016)	To systematically analyze online eating disorder (ED) discussions from user-generated content.	Cross-platform content analysis was employed to systematically analyze online eating disorder (ED) discussions from user-generated content, aiming to create a comprehensive corpus of ED-related posts.

2.5 Research Gap

While sentiment analysis and machine learning approaches have individually contributed to our understanding of public sentiment, there remains a need for research on how to effectively integrate these methods for a comprehensive examination of specific food products, such as millet-based confections. Current research often focuses solely on either sentiment analysis or machine learning, missing the potential synergy of combining both approaches. This study aims to bridge these gaps by developing a comprehensive framework for sentiment categorization in discussions about millet-based confections on Twitter. The research focuses on the top 10 economies, which are globally influential and economically significant. Its intent is to provide valuable insights for candy producers, marketers, and policymakers looking to tailor their strategies to the preferences and sentiments of consumers worldwide.

3 Methodology

The study aimed to investigate how millet-based confections were perceived and discussed across the top 10 global economies. Text analysis and machine learning techniques were employed to gain a deeper understanding of public opinions and perceptions of millet-based confectionary products. The objective was to create a classifier that could, based on the language in the tweets, classify them as conveying positive, negative, or neutral sentiments.

Information about millet-based confections was extracted from Twitter chats using several approaches. The initial listing of the top 50 words shed light on recurring themes and critical elements in this gastronomy sector, representing one of the key findings. Sentiment classification accuracy was enhanced through the utilization of a Naive Bayes classifier to simplify sentiment analysis. Additionally, the extraction of informative features enabled the identification of keywords that supported sentiment categorization. Visual representations of subjectivity distributions and polarity added depth to the understanding of emotional nuances. A bar chart was employed to provide insights into the distribution of positive, neutral, and negative tweets within the sentiment analysis. Word clouds were used to create visual representations of frequently associated phrases for all word combinations, as well as for neutral, positive, and negative sentiments. An exhaustive list of the top 20 terms and word pairs, customized to specific attitude categories and the entire dataset, further illuminated the essential linguistic factors influencing discussions about millet-based confections. Figure 1 illustrates the step-by-step research methodology, starting from data extraction and culminating in the presentation of results.

To gain insightful knowledge on the acceptability of and feelings towards millet-based confections, the power of data pre-processing methods, sentiment analysis, and machine learning classifiers was harnessed in this study. The data may prove useful to manufacturers of millet-based products, marketers, and policymakers seeking to understand consumer attitudes and preferences in the international market.

3.1 Why 10 Top Economies

This study aimed to gain comprehensive insights into public attitudes and discussions related to millet-based confections. To achieve this, data was collected from the top 10 economies, as they play a substantial role in the global economy and have a significant customer base. These economies often influence global trends and market dynamics, making their social media discussions highly relevant. Focusing on these nations allowed us to gather a wide range of perspectives and emotions, establishing a robust foundation for understanding the global adoption of millet-based confections. Additionally, the knowledge gained from these leading economies can be invaluable to confectionery producers and marketers, helping them tailor their products and business strategies to cater to the diverse tastes and needs of customer segments worldwide.

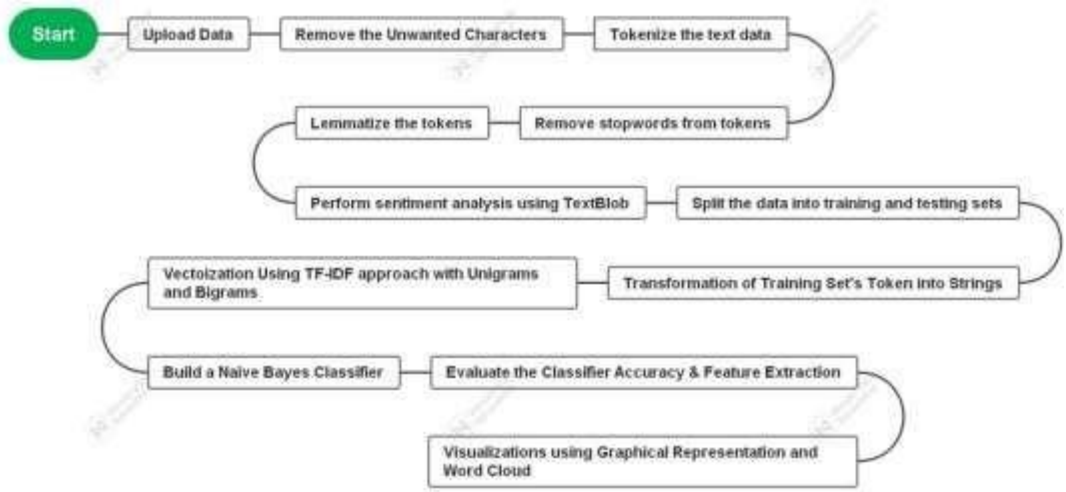

Figure 1: Highlights of Methodology

3.2 Data Collection

Data from the top 10 economies in the world were gathered for this study by utilizing data scraping methods from various social media sites, including Twitter. To locate tweets related to millet-based confectioneries, a keyword-based search approach was employed. A total of 1699 tweets were compiled, containing words such as 'millet confectionary,' 'millet sweets,' 'millet chocolate,' 'millet-based confectionary,' 'millet desserts,' 'millet candy,' 'millet snack foods,' and 'millet bakery.' This study aimed to advance our understanding of the sentiments and discussions surrounding millet-based confectioneries.

3.3 Pre-processing

Data Loading: The dataset of 1699 tweets on millet-based sweets was loaded into the Python environment using the Pandas package. The dataset had been stored in a CSV file.

Text Cleaning: Regular expressions were used to clean the text data from the tweets, eliminating unnecessary letters and symbols. This phase ensured that the analysis was not hampered by removing any special characters, emojis, or other non-alphanumeric characters from the data.

Tokenization: The word_tokenize function of the NLTK library was employed to tokenize the cleaned text into individual words. Tokenization reduced the text to a list of words, enabling subsequent word-level processing and analysis.

Stopword Removal: The NLTK stopwords list for the English language was used to eliminate stopwords, sometimes referred to as common and uninformative words, from the list of tokens. These stopwords, which had no bearing on the analysis, included words like "the," "a," "an," "and," etc.

Lemmatization: To normalize the words and reduce them to their root or base form, the remaining tokens were lemmatized. Lemmatization was performed using the WordNetLemmatizer function of the NLTK library, which assisted in reducing word variants and maintaining analytical consistency.

3.4 Sentiment Analysis

The TextBlob package was employed to conduct sentiment analysis on the pre-processed text data. The polarity and subjectivity of each tweet were determined using this technique, revealing whether the conveyed sentiments were positive, negative, or neutral. Histograms of polarity and subjectivity were utilized to visualize the sentiment distribution, providing an overview of the general sentiment patterns in the dataset. The polarity score indicated the sentiment of the tweet, ranging from -1 (most negative) to 1 (most positive), while the subjectivity score indicated the level of subjectivity or objectivity in the tweet, ranging from 0 (most objective) to 1 (most subjective).

3.5 Naïve Bayes Classification

A Naive Bayes classifier was trained using the Scikit-learn toolkit to gain further insight into the sentiments expressed in the tweets. Eighty percent of the data was employed for training, while the remaining twenty percent was designated for testing. To facilitate vectorization, the tokens from the training set were converted into strings. The text data was vectorized using the 'term frequency-inverse document frequency (TF-IDF)' approach with unigrams and bigrams. The Naive Bayes classifier was generated using the training data, and its accuracy was assessed using the test data. The training dataset constituted 80% of the total data, and the remaining 20% comprised the testing data.

3.6 Classifier Accuracy Evaluation and Feature Extraction

The accuracy of the Naive Bayes classifier was assessed by training it on the training dataset and testing its performance on the test dataset. The accuracy was used to measure the precision of sentiment measurement, determined by comparing the classifier's predictions to the actual sentiment labels in the test dataset. It is acknowledged that the Naive Bayes classifier was used, recognizing its effectiveness could vary with different datasets and might not capture more subtle sentiment nuances similar to those encountered in the current study. The most useful characteristics of the classifier were further investigated by examining the log probabilities corresponding to each feature. These descriptive elements revealed the words and word combinations most closely associated with both positive and negative emotions. Polarity and subjectivity histograms were constructed to depict the overall sentiments conveyed in the dataset, providing a deeper understanding of the sentiment distribution. Additionally, a bar chart was created to display the distribution of tweets in the positive, neutral, and negative sentiment categories. This offered a comprehensive overview of the general sentiments expressed by social media users regarding millet-based confectioneries."

3.7 Word Clouds

The most commonly used terms across the entire dataset and within each sentiment category were visually presented using word clouds. These word clouds provided readers with a quick and straightforward understanding of the terms that frequently appeared in discussions about millet-based sweets. The frequency of a word in the text data for a particular sentiment was represented by its size and boldness in the word cloud. The analysis of the word clouds could reveal recurring themes or emotions associated with specific categories. They could help in identifying the key words and expressions that were frequently employed and contributed to the overall emotion conveyed in the text data.

3.8 Software and Libraries Use

Python programming was used to carry out the pre-processing, analysis, and visualization of the data. In this study, the following libraries were used:

- Pandas: For data alteration, transformation, restructuring, and analysis
- NLTK (Natural Language Toolkit): For text cleaning, tokenization, stopwords removal, lemmatization, and hashtag extraction
- TextBlob: For sentiment analysis
- Scikit-learn: For building the Naive Bayes classifier and TF-IDF vectorization
- WordCloud: For generating word clouds
- Matplotlib: For data visualization

4 Results

4.1 Top Words

When the tweets on millet-based confections were analyzed, a comprehensive examination of the top 50 terms by frequency was conducted, providing insightful information on recurring themes and ingredients in this gastronomic domain (Refer: Table 2). Phrases such as 'gum' and 'Gum' were found to be the most frequently used keywords in the context of millet-based confections, with 1060 and 970 occurrences, respectively, followed by 'millet' (970) and 'Millet' (480). This suggests that 'gum/Gum' and 'millet/Millet' (regardless of capitalization) were the two frequently mentioned components in the tweets, underscoring their significance in the realm of confections made with millet. Terms such as 'flour' and 'Flour' were also notable, with 610 and 180 occurrences, respectively, followed by 'rice' (500). This highlights the prevalence of discussions about the use of rice in confections and the mention of various types of flour, possibly including millet flour. A wide array of ingredients, including 'sorghum,' 'xanthan,' 'bean,' 'starch,' 'oil,' 'tapioca,' 'maize,' 'sesame,' 'banana,' 'corn,' 'nut,' 'seed,' 'sugar,' 'potato,' 'kernel,' 'cassava,' 'cocoa,' 'soybean,' 'groundnut,' and 'oat,' were commonly referenced in discussions about millet-based confections. This underscores the diversity of ingredients utilized in these confections. The appearance of names like 'Lisa,' 'PM,' 'Dealer,' 'Cool,' 'Ralf,' and 'Dub' suggests references to individuals or entities associated with the topic. The use of the terms 'Ft' (feature) and 'DJ' (disc jockey) indicated discussions about upcoming projects or events. In the context of millet-based confections, terms such as 'Arabic' and 'Arabic' implied a cultural background or influence, potentially indicating regional distinctions or preferences. Words like 'use,' 'crop,' and 'yeast' indicated discussions about the utilization, cultivation, and processing of materials in millet-based confections. Words such as 'Lovin,' 'Trippin,' and 'also' may imply positive sentiments or favorable associations with millet-based confections. These findings reflect a diverse range of debates and discussions surrounding millet-based confections, encompassing topics related to ingredients, preparation, cultural considerations, and potential positive associations. These results shed light on the gastronomic elements and linguistic nuances that characterize discussions about millet-based confections. A more in-depth analysis of the contextual content around these phrases can unveil subtle patterns and meanings, contributing to a deeper understanding of this culinary discourse and enhancing the research's granularity.

Table 2: Top 50 Words By Frequency:

Words	Frequencies	Words	Frequencies
gum	1060	Dj	150
millet	970	tapioca	150
flour	610	maize	150
Gum	520	sesame	140
rice	500	like	130
Millet	480	banana	130
sorghum	390	corn	120
amp	280	nut	110
xanthan	250	seed	110
bean	240	Sorghum	110
Lisa	230	glutenfree	110
PM	220	sugar	100
starch	210	potato	100
Dealer	200	use	100
Cool	200	Ft	100
Ralf	200	kernel	100
oil	200	cassava	100
Trippin	180	cocoa	100
Dub	180	soybean	100
Flour	180	groundnut	100
Lovin	170	oat	90
Arabic	170	Maize	90
Arabic	170	also	90
crop	170	DJ	90
palm	160	yeast	90

4.2 Accuracy

A fairly high accuracy of 99.12% was achieved by the Naive Bayes classifier when it was used for the sentiment analysis of tweets on millet-based confections. This demonstrated a remarkable level of accuracy in classifying tweet sentiment as positive, negative, or neutral. The classifier's ability to distinguish and categorize emotions accurately highlighted its strength and effectiveness in capturing the subtle emotional content of the tweets, underlining its potential as a valuable tool for sentiment analysis within the domain of millet-based confections.

4.3 Informative Features

The top 50 informative features of the Naive Bayes classifier were analyzed to obtain valuable insights into the essential words and phrases crucial for the classification of sentiment in tweets about millet-based confections (Refer: Table 3). The log probability scores associated with these features indicated their relative importance in distinguishing sentiments. Phrases such as 'toilet asking,' 'bought,' and 'wheat' displayed higher negative log probabilities, suggesting an association with unfavorable attitudes. Conversely, terms like 'natural,' 'starter,' and 'starch' exhibited lower negative log probabilities, indicating a connection to positive sentiments. It is worth noting that the presence of specific components like 'cocoa,' 'palm,' and 'sesame' appeared to influence people's moods. Furthermore, words like 'maize,' 'sorghum,' and 'millet,' referring to gum, flour, and other grains, demonstrated greater significance.

Table 3: Top Informative Features

Informative Features	Log Probability	Informative Features	Log Probability
toilet asking	-7.323305751	sesame	-7.138362728
bought	-7.321121404	cocoa	-7.128964579
wheat	-7.305203393	palm	-7.119721812
starch millet	-7.300121553	seeds	-7.114686387
chewing	-7.292851606	guar	-7.113772642
chewing gum	-7.292851606	guar gum	-7.113772642
give	-7.284459987	xanthan	-7.107196973
natural	-7.278145605	xanthan gum	-7.101440926
starter	-7.272065421	rice millet	-7.064211933
starch	-7.268730125	etc	-7.050354476
rye	-7.249491886	buy	-7.040910231
acid	-7.248818505	flour organic	-7.027850613
groundnut	-7.237788601	cheese	-6.993323355
produce	-7.216089717	maize	-6.956821874
yam cassava	-7.20178603	friend	-6.951369366
base	-7.200182836	rice	-6.943836559
sugar	-7.181008459	mac	-6.925543644
seed	-7.169428106	bean	-6.889907086
gum stale	-7.143813462	arabic	-6.859088897
pk	-7.143813462	gum arabic	-6.840175039
pk sour	-7.143813462	flour	-6.626695159
sour	-7.143813462	organic	-6.624611483
sour chewing	-7.143813462	sorghum	-6.555542524
stale	-7.143813462	millet	-6.452474821
stale millet	-7.143813462	gum	-6.413377903

The study of the polarity and subjectivity scores revealed important details about the attitudes that people had toward millet-based confections. It was clear from looking at the polarity score distribution that the bulk of tweets fell between -0.15 and 0.6 (Refer: Figure 2). In particular, the neutral point, which is represented by a polarity score of around 0, was found to have the highest frequency of tweets (as shown in Figure 2: Polarity Distribution). The frequency then reached noticeable maxima at polarity scores of 0.35 and 0.15. This distribution indicates a neutral mood landscape, with a sizeable fraction of tweets expressing a favorable opinion of millet-based confections.

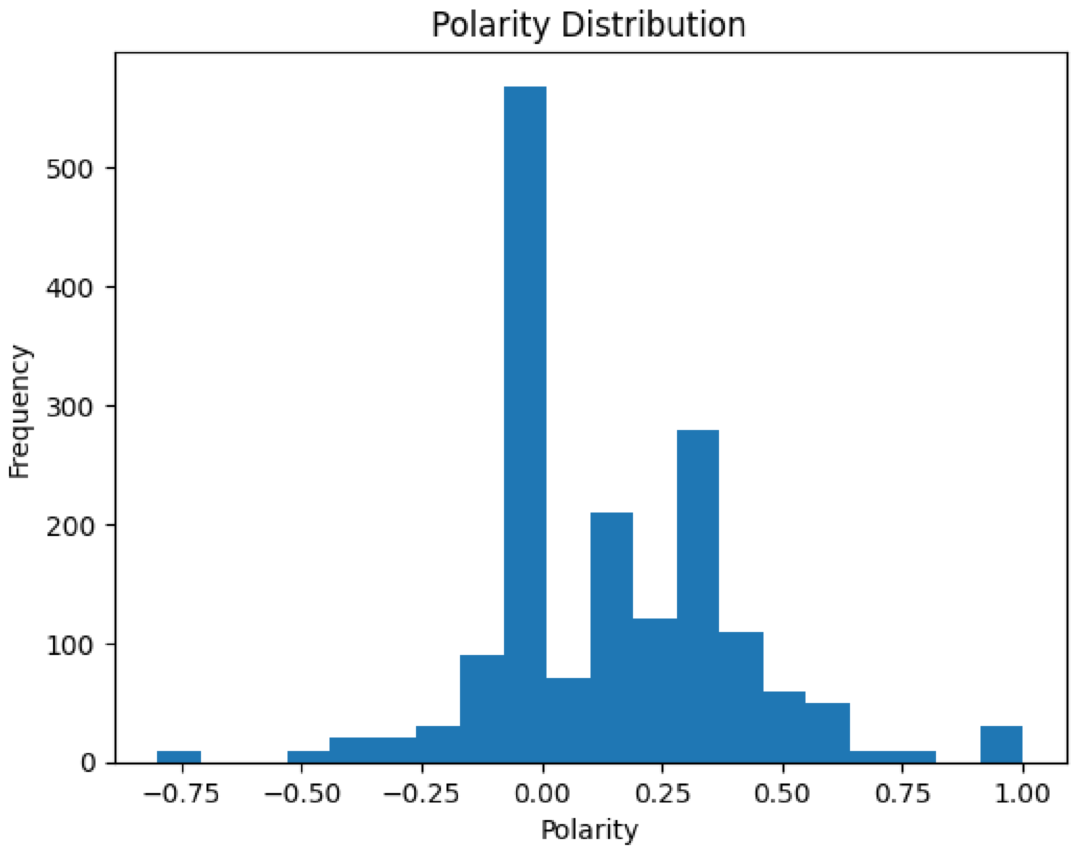

Figure 2: Polarity Distribution

The study of subjective scores provided further context for the sentiment analysis of people regarding millet-based confections. The subjectivity distribution analysis found significant trends in the subjectivity distribution of tweets. Notably, the subjectivity score range between 0 and 0.05 showed the highest frequency, indicating that tweets about millet-based confectionary tended to use objective and factual language more frequently (Refer: Figure 3). Next, subjectivity ratings in the range of

0.25 to 0.6, then 0.65 to 0.7, were connected to the next highest frequencies. These patterns offer a detailed view of the many levels of subjectivity employed in tweets on this culinary topic, as seen in Figure 3: Subjectivity Distribution. The research shows that millet-based confectionary discourse is frequently balanced between objective and subjective expressions among consumers, demonstrating a diverse interest in the topic.

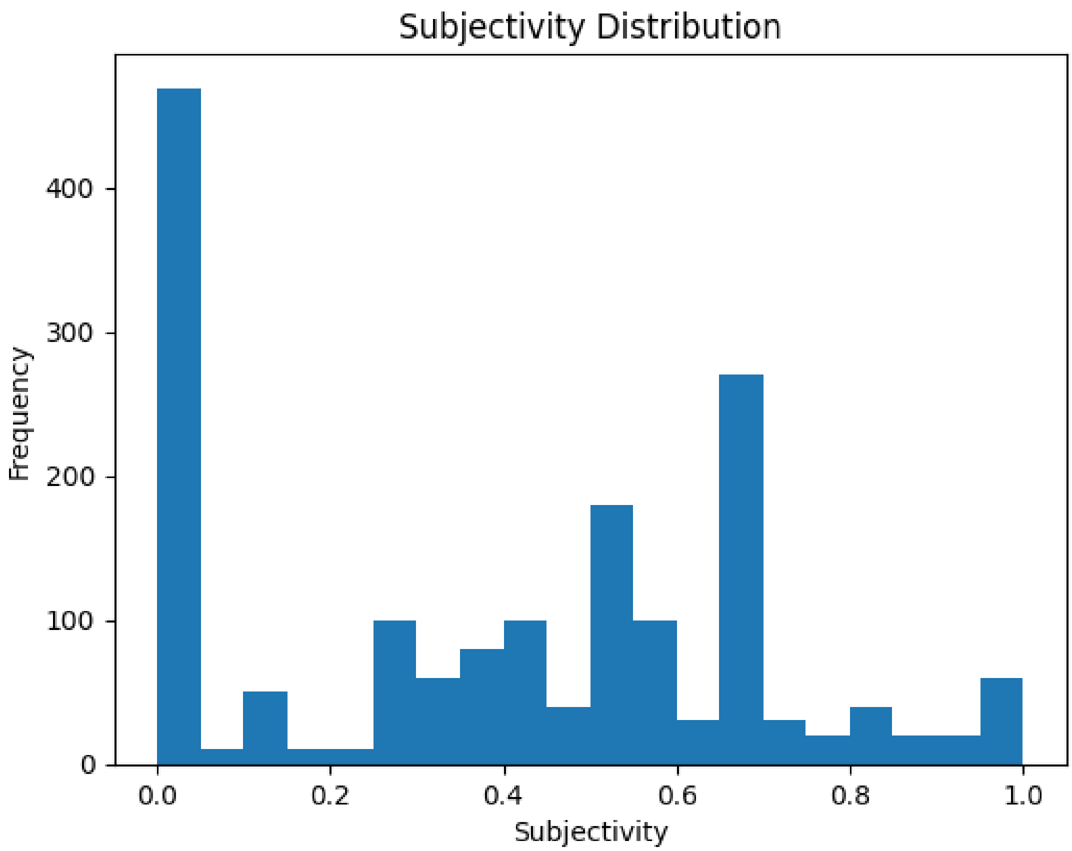

Figure 3: Subjective Distribution

The sentiment analysis of the dataset offered valuable insights into the prevailing user attitudes regarding millet-based confections. Sentiment classification revealed a predominantly positive sentiment, with 950 tweets expressing favorable and pleasant comments about confectionery items made from millet (See: Figure 4). This indicates strong support and appreciation for these items within the discussions. Simultaneously, a substantial portion of tweets—519 in total—was categorized as neutral, suggesting that consumers adopted a fair-minded and objective perspective, reflecting either the presentation of factual information or a lack of strong sentiment. A smaller but still significant fraction of tweets—230 in number—expressed negative sentiment, likely conveying unfavorable opinions or dissatisfaction with specific aspects of millet-based confections. The sentiment analysis, as displayed, underscores the dominance of positive sentiment, reflecting users' overall preferences for these confections. This finding could influence the industry, as a positive sentiment trend may indicate an opportunity for the promotion and further development of millet-based confectionery products.

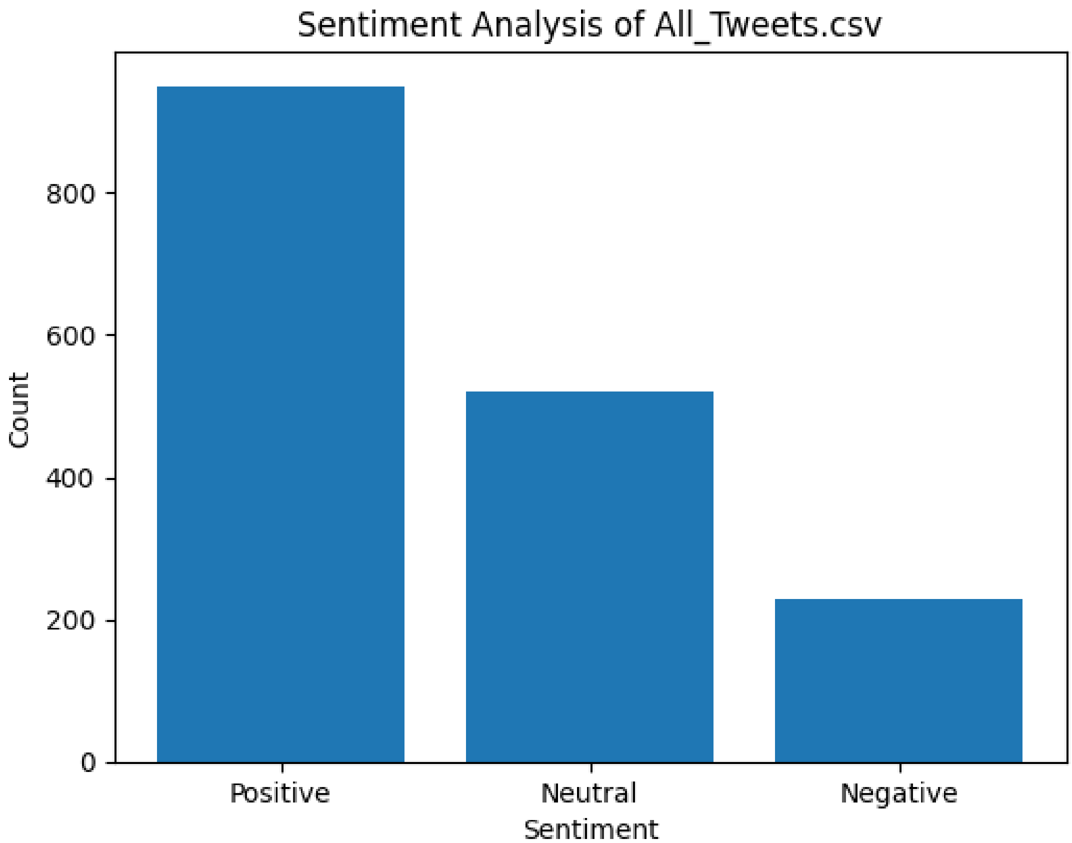

Figure 4: Sentiment Analysis

4.4 Most Frequent Positive Words

The Twitter sentiment analysis of millet-based confections was conducted, offering an abundance of fascinating insights into the positive word sentiments. The term 'gum' (570) and 'Gum' (340), with a total count of 910, were found to be the most frequently occurring positive words, followed by 'millet' (580) and 'Millet' (280), totaling 860. Furthermore, the term 'flour' appeared 380 times, demonstrating its diverse application in confectionery. The variety of elements under examination was further enriched by 'rice,' appearing 340 times. Notably, the research revealed the significance of several names, with 'Lisa' and 'Sorghum' each receiving 220 occurrences, potentially indicating a positive association. Terms like 'Dealer,' 'Cool,' 'Ralf,' and 'Trippin,' each occurring 200 times and conveying elements of delight and enthusiasm, contributed to the overall positive sentiment. Words such as 'Dub,' 'Bean,' 'PM,' and 'Lovin' added to the diverse and vibrant sentiment landscape, with frequencies ranging from 180 to 170. In conclusion, this investigation shed light on the complex nature of discussions surrounding millet-based confections and provided a comprehensive understanding of the predominantly favorable attitudes within this culinary domain.

4.5 Most Frequent Neutral Words

A range of neutral phrases were discovered during our investigation, providing significant contextual insight into the discussion of millet-based confections. Notably, the term 'gum,' in both capitalized and lowercase forms, was observed 440 times, underscoring its flexibility and importance. Similarly, 'millet' emerged as a prominent phrase with a total of 380 occurrences, indicating its widespread usage. The word 'flour,' with 220 instances in various capitalization forms, demonstrated a nuanced usage pattern. Furthermore, the data unveiled the significance of terms such as 'rice,' neutrally mentioned 110 times, and 'sorghum,' which appeared 160 times in both capitalized and lowercase variants. Specific words like 'starch,' 'amp,' 'Maize,' and 'cup' were each encountered 80 times. The research also identified specific references, such as 'Arabic,' which surfaced 90 times, suggesting potential cultural contexts. Words like 'tapioca,' 'cassava,' 'xanthan,' and 'cotton' were found between 70 and 50 times, subtly underscoring their presence. 'Crop' and 'tonne,' each appearing 50 times, provided subtle hints within the conversation about millet-based confections. This investigation has shed light on the intricacies of neutral sentiments and their interplay within the domain of food.

4.6 Most Frequent Negative Word

The investigation of language surrounding millet-based confections revealed numerous unfavorable terms. The unfavourableness was attributed to the approach of the tweets and demanded cautious attention. Prominent among these were the words "gum," "millet," and "flour," with frequencies of 230, 210, and 100, respectively, casting a shadow over elements associated with these compounds. Notably, capitalization differences did not diminish the impact. The word "flour," with a count of 100, reflected issues surrounding this ingredient. The word "seed" appeared 80 times, regardless of capitalization, and stood out as one of the top 20 negative words. Additionally, "Organic," "rice," "bean," and "Arabic" played significant roles, with each phrase appearing 50 times, highlighting the various dimensions associated with negative sentiment. The study also revealed characteristics linked to the words "xanthan" and "maize," each appearing 40 times among the top twenty negative words. The research also underscored the relevance of the words "etc," "sugar," "natural," "amp," and "sesame," each appearing 30 times among the top negative words, calling for attention to complex issues. These findings regarding unfavorable attitudes toward millet-based confections, shedding light on intricate interactions, emphasize the need for in-depth investigation in these areas.

4.7 Word Pair Cloud

The cluster of terms that appeared as the most common pairs in the discussion of millet-based confections was presented in the form of a word pair cloud, and their frequencies were extracted. The words 'Gum Arabic,' 'gum Arabic,' and 'gum Arabic,' were found to be the most significant word pair, with 90, 70, and 140 occurrences, regardless of capitalization. 'Xanthan gum' emerged as the next dominant word pair with a frequency of 240, emphasizing its crucial role in millet-based confections. The names 'Lisa Millet' and 'Millet Cool' both appeared 190 times, underscoring specific individuals and entities. 'Rice flour' and 'Ralf Gum' prominently appeared 180 times, evoking a symphony of textures and flavors. 'Cool Lovin' and 'Lovin Ralf,' resonating with 170 occurrences, delved into a range of experiences for emotional investigations. 'Trippin Dub' and 'Gum Trippin,' intertwined 170 times, continued to create a vivid mosaic of emotions. 'Dj Dealer' and 'Dealer Ft' intersected 130 times, contributing a rhythmic cadence to the confection narrative. 'Ft. Lisa' and 'millet flour' each appeared 100 times, enriching the conversation with diverse flavor and compositional aspects. Additionally, 'palm kernel' and 'brown rice' made 80 and 70 appearances, introducing various textures to the dialogue. The terms 'rice millet' and 'sorghum millet,' both captured 70 instances, underwent further relationship analysis, revealing a tableau of culinary fusions and influences. These related word pairs offered a mosaic of ideas that enriched the discussion surrounding millet-based confections by weaving together a rich tapestry of tastes, personalities, and culinary creations. Figure 5 displayed the word pair cloud of the tweets.

Figure 5: Word Pair Cloud

4.8 Positive Word Pairs

A symphony of positive expressions was unfolded in the world of millet-based confections through frequent combinations that eloquently elicited a range of emotions. The linked phrases 'Lisa Millet' and 'Millet Cool,' each resonating with 190 occurrences and evoking visions of delightful experiences, reached their peak. In complementing this, 'Ralf Gum' and 'Cool Lovin' unfolded with equal significance at 180 and 170 instances, respectively, capturing joyful and adoring moments. The dynamic interplay of 'Loving Ralf,' 'Trippin Dub,' and 'Gum Trippin,' which reverberated at 170 and 160 frequencies, depicted stories of companionship and adventure within these culinary realms, echoing similar emotions. The phrase 'xanthan gum' was used 140 times, emphasizing the crucial role of this component in confections. When it came to fusing audio experiences with the culinary journey, 'DJ Dealer' or 'DJ Dealer' took the stage with 130 and 70 occurrences, respectively. The appearance of 'rice flour' 120 times reflected its significance in flavor and texture. Additionally, 'Dealer Ft' and 'Ft Lisa' together revealed 100 instances, representing the convergence of elements in these creations. The term 'gum Arabic' appeared 80 times, while the phrases 'palm kernel' and 'millet flour' intersected 70 times each, orchestrating the composition of ingredients in these confections. An expertly crafted story of flavors that evoked moments of pleasure was unveiled by the ethereal coupling of 'cashew nut,' 'nut cassava,' 'cassava cocoa,' and 'cocoa bean,' each appearing 60 times. Figure 6 presented the word pair cloud illustrating the positive sentiments expressed by people in the tweets.

Figure 6: Positive Word Cloud

4.9 Neutral Word Pairs

A tapestry of neutral statements was formed within the realm of millet-based confections, created through the recurrence of word pairs that intricately embellished the sensory experience. While 'rice flour' was delicately converged with them, resonating peacefully, the delicate combination of 'Gum Arabic' and 'xanthan gum,' each echoing at 60 instances, played a role in providing texture and cohesion. The words 'rice flour' and 'flour cup' appeared 60 and 50 times, respectively, symbolizing the fundamental components of the confections. 'Tapioca flour' and 'potato starch' interlaced 30 times each among the neutral expressions, highlighting the complexity and nuance of flavor profiles. This contributed to a sense of balance and wholesomeness by embracing the simplicity of 'sorghum millet' (regardless of capitalization) and 'millet rice,' which appeared 50 and 30 times, respectively. The echoes of 'guar gum' and the harmony of 'Sorghum Maize,' both with a frequency of 30, further revealed the intricately woven textures and harmonies within this sonic tapestry. Intricate tales of flavors and craftsmanship were woven by 'Wheat FlourOatsGum,' manifesting delicate infusions by blending ingredients in 20 different instances, and by 'FlourOatsGum Arabic SugarMilletJuniour,' 'Mammoth pecan,' and 'ArabicSugarMilletJunior Mammoth,' each resonating 20 times. The fusion of the words 'starch millet,' 'Arabic Sogom,' and 'sogom millet' appeared 20 times, reflecting the diversity of the confections. The term 'market omdurman,' appearing 20 times, indicated the popularity of these millet-based confections in local markets. Figure 7 presents the word pair cloud depicting the neutral sentiments expressed by individuals in the tweets.

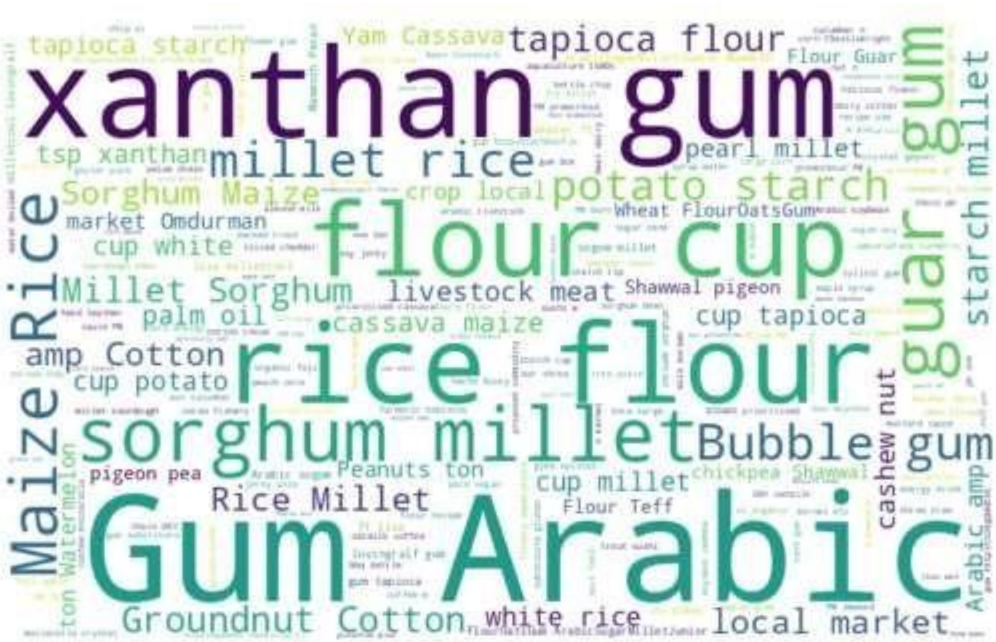

Figure 7: Neutral Word Cloud

4.10 Negative Word Pairs

In a separate tapestry of unfavorable mood, a subtle limit of gastronomic discovery was hinted at by incongruous word combinations. The words "xanthan gum" and "gum Arabic" were each found 40 times among the top twenty negative words, casting a shadow over the sensory experience and prompting contemplation on texture and structure. The contradictory interaction of "Flour Organic," resonating 30 times, signified discontent and dissatisfaction, while the rhythmic union of "rice millet," interlacing at 30 frequencies, indicated a potential amalgamation of grains. The words "guar gum" and "sesame seed," both appearing 20 times, combined to reveal complexity with a hint of bitterness. As "psyllium husk" beckoned with its own weight of significance, echoes of "xanthum gum" and "gum psyllium," both occurring 20 times, triggered contemplation regarding the amalgamation of materials. Amidst the layers, "millet flour" appeared 20 times, adding depth to the narratives. Meanwhile, "indifferent Krusteaz" and the identity of the "Krusteaz brand," each recurring 10 times, injected comments on indifference and brand connections. "Sorghum brown" and "brown rice," both echoing at 10 occurrences, provided a sensation of earthiness and subdued contentment. The inclusion of words like "brand composed" and "composed sorghum," with 10 appearances each, reflected the complexities between millet types and their underlying characteristics. "Gum etc" and the enigmatic "etc like," each resonating at 10 frequencies, echoed a world of uncertainties. The terms "brown rice," "millet quinoa," and "quinoa xanthan," each appearing ten times, shed light on various textures and represented interactions within the world of millet compositions. Figure 8 displayed the word pair cloud of negative sentiments realized from the tweets by people.

Figure 8: Negative Word Cloud

5 Discussion

The examination of tweets centered on millet-based confections involves the analysis of top-frequency phrases, yielding valuable insights into the recurring themes and influential ingredients within this culinary context. Notably, terms like 'gum' and 'millet' are frequently encountered, indicating their significance in confection-related discussions. The frequent occurrence of words such as 'flour' and 'rice' underscores the diversity of components and their varied applications. Beyond these components, the presence of phrases like 'Arabic' suggests cultural influences, while names like 'Lisa,' 'PM,' 'Dealer,' and others add a personal dimension to the discourse. The sentiment analysis reveals a predominantly favorable attitude toward millet-based confections, signifying a strong preference among users. The high accuracy of the classifier highlights its ability to capture subtle emotional nuances. Particular words, such as 'cocoa,' 'palm,' and 'sesame,' play a pivotal role in influencing sentiments, shedding light on the critical terms contributing to sentiment categorization. In terms of subjective scores, a well-balanced mix of objective and subjective statements is evident, while the polarity distribution indicates a primarily neutral stance.

A complex tapestry of concerns and perceptions is unveiled through further examination of neutral and negative sentiments. Terms such as 'gum,' 'millet,' and 'flour,' which accentuate their multifaceted applications, is imbued with neutral emotions. Conversely, negative sentiments accentuate the challenges associated with terms like 'Organic,' 'rice,' 'bean,' and others. These intricate sentiments offer a comprehensive insight into the diverse perspectives and emotions that are encompassed within the realm of millet-based confections.

The complex linkages and feelings surrounding millet-based confections are revealed through word pair analysis of the discourse. Notably, certain combinations, such as 'xanthan gum,' underscore the significance of specific components in these confections and showcase the meticulous balance of flavors and textures. Additionally, the frequent mentions of names like 'Lisa Millet' and 'Millet Cool' impart a personal touch and allude to heartwarming interactions and shared culinary adventures. Further investigation uncovers a intricate interplay of feelings in pairs like 'Cool Lovin,' 'Lovin Ralf,' and 'Trippin Dub,' suggesting a spectrum of pleasure, affection, and adventure evoked by these confections. Furthermore, the recurrent use of 'Gum Arabic' in various capitalizations not only draws attention to its widespread utilization but also hints at potential cultural implications associated with these treats. The exploration of neutral and unfavorable pairings unveils the intricate interplay of tastes and ingredients, highlighting the exquisite balance that characterizes the realm of millet-based confections. Overall, word pair analysis contributes to a deeper understanding of this culinary domain by enhancing comprehension of the intricate tapestry of flavors, emotions, and cultural nuances that permeate discussions surrounding millet-based confections.

5.1 Implications

5.1.1 Administrative Implications

The conclusions derived from the sentiment analysis of tweets related to millet-based confections bear substantial administrative implications and provide value to a range of stakeholders. Firstly, the favorable perception of millet-based confections, as revealed by these findings, can guide public health authorities and policymakers in formulating dietary recommendations that align with consumer preferences. Secondly, this study can be utilized by food regulatory agencies to monitor and evaluate the level of interest in specific ingredients, such as 'gum' and 'millet,' to ensure their safe and ethical use. Additionally, marketing and advertising firms can draw upon the prevalent positive sentiments and cultural associations identified in the tweets to develop effective promotional strategies for millet-based confections.

5.1.2 Managerial Implications

The management implications resulting from this research provide valuable insights for executives in the food industry and confection producers. Firstly, the emphasis on specific ingredients such as 'flour' and 'rice' suggests the potential for product innovation and diversification, encouraging the exploration of new recipes and variations. Secondly, the popularity of certain names, such as 'Lisa' and 'Sorghum,' suggests the possibility of celebrity endorsements or collaborations to enhance brand recognition. Furthermore, when managers are informed about the emotional associations linked with particular ingredients, they are better equipped to tailor their messaging and packaging to resonate with the emotions of their target consumers.

5.1.3 Research Implications

The results of this study offer opportunities for further investigation and understanding from a research perspective. Firstly, the subtle interactions between emotions and ingredients open the door to in-depth sensory inquiries that unveil the emotional and gustatory factors influencing consumer choices. The identified phrases and their correlations with sentiments also form the basis for linguistic and cultural studies, contributing to a deeper understanding of how language influences food perception. The proficient application of sentiment analysis in a culinary context also encourages the exploration of analogous methods in other gastronomic investigations, showcasing the potential for interdisciplinary discoveries.

5.2 Limitations

While this study offers valuable insights, it is not without limitations. Firstly, the accuracy of sentiment categorization can be influenced by the potential inability of sentiment analysis to recognize highly context-dependent emotions, idiomatic phrases, or subtle feelings. Secondly, the study is constrained by the available dataset, which may not fully represent the complete spectrum of feelings and attitudes toward millet-based confections. Furthermore, the study restricts its analysis to tweets written in English, possibly excluding opinions expressed in other languages. Thirdly, there is an absence of approach to explore the in-depth reason behind the sentiments. While we have conducted sentiment analysis, it doesn't capture the underlying factors influencing these sentiments fourth, this study lacks in the incorporation of an accurate and representative training dataset for the machine learning model, which is essential for its effectiveness. Fifth, due to the frequent use of brand name or trending topics related to our keyword “millets”, the sentiment analysis, word cloud and word-pair cloud discloses the instances where the terms associated with personal names and unrelated subjects crops up. In our current study, these associations may not be readily discernible. Finally, the research does not take into account time hence this study fails to define the evolution of the sentiments over time and the potential influence from the outside sources over the time. Recognizing these limitations opens the door to future studies that can address these issues and provide a more comprehensive understanding of attitudes related to millet-based confections.

5.3 Future Scope

In the realm of future research possibilities, several avenues open up for expanding our understanding of millet-based confections and the sentiments surrounding them. Firstly, the inclusion of sentiment analysis from multiple geographic locations will create opportunities for cross-cultural and regional assessments. Understanding how millet-based confections are perceived across various cultures and nations can help unveil their universal appeal and cultural nuances. Secondly, the utilization of more advanced natural language processing methods, such as deep learning models, will provide a deeper understanding of the subtle emotional nuances within conversations. These models can be employed to uncover nuances, sarcasm, and context-specific emotions that may be overlooked by current approaches. Thirdly, the incorporation of multimedia content analysis, such as photos and videos posted on social media, will provide a comprehensive understanding of the verbal and visual elements that influence sentiment. These elements can be leveraged to create a more holistic understanding of consumer sentiments toward millet-based confections. Fourth, the adoption of more advanced techniques such as Part-of-Speech tagging, Named Entity Recognition, sentiment lexicon analysis, and Latent Dirichlet Allocation (LDA) can also be applied for the identification of sentiment-related phrases, key themes, and topics within the tweets. The LDA approach may not provide the reasons of the sentiments directly but can contribute to the study by figuring out the reason of the sentiments from the themes those are associated with the sentiments. Understanding these topics can potentially shed light on the reasons behind sentiments expressed in the data. The themes can also clarify the emergence of personal names in the word cloud or word pair cloud. Moreover, by organizing the data by date and time and by applying trend analysis can shed lights on the evolution of sentiments over the time and can give more contextual meaning of the sentiments.

6 About the author

Dr. Parminder Singh Dhillon is the founder faculty member of the department of Tourism, Hospitality & Hotel management He is an alumni of Food Craft Institute Chandigarh, 1992 batch and is armed with more than 29 years of teaching and industry exposure in the glamorous field of hospitality. He has worked with Taj group and ITC Welcomgroup of hotels. He won first prize from CITCO, Chandigarh for standing first in the certificate course in Cookery. He did his PhD. in Gastronomic Tourism, Masters in Tourism management and M.B.A. in Hospitality management. Dr. Parminder has got to his credit 12 research paper publications and has presented more than 8 research papers in various International and National conferences. He is also the student counselor and project supervisor for Bachelors and Masters programmes in tourism for the Indira Gandhi National Open University, New Delhi (IGNOU). He has authored eight books on Food production/cooking for hospitality industry published by reputed International and National publishers.

7 References

Anitha, S., Kane-Potaka, J., Tsusaka, T. W., Botha, R., Rajendran, A., Givens, D. I., ... & Bhandari, R. K. (2021). A systematic review and meta-analysis of the potential of millets for managing and reducing the risk of developing diabetes mellitus. *Frontiers in nutrition,* 386.

Benítez-Andrades, J. A., Alija-Perez, J. M., Vidal, M. E., Pastor-Vargas, R., & García- Ordas, M. T. (2022). Traditional Machine Learning Models and Bidirectional Encoder Representations From Transformer (BERT)-Based Automatic Classificationof Tweets About Eating Disorders: Algorithm Development and Validation Study. *JMIR Medical Informatics*, *10*(2), 1–13. https://doi.org/10.2196/34492

Bueno, I., Carrasco, R. A., Ureña, R., & Herrera-Viedma, E. (2022). A business context aware decision-making approach for selecting the most appropriate sentiment analysis technique in e-marketing situations. *Information Sciences*, *589*, 300–320. https://doi.org/10.1016/j.ins.2021.12.080

Caggia, C., Palmeri, R., Russo, N., Timpone, R., Randazzo, C. L., Todaro, A., & Barbagallo, S. (2020). Employ of citrus by-product as fat replacer ingredient for bakery confectionery products. *Frontiers in nutrition, 7*, 46.

Caiazza, R., & Bigliardi, B. (2020). Web marketing in agri-food industry: Challenges and opportunities. *Trends in Food Science & Technology, 103*, 12-19.

Carr, J., Decreton, L., Qin, W., Rojas, B., Rossochacki, T., & wen Yang, Y. (2015). Social media in product development. *Food Quality and Preference, 40*, 354–364.

Chera, M. (2017). Transforming millets: Strategies and struggles in changing taste in Madurai *Food, Culture & Society, 20*(2), 303–324.https://doi.org/10.1080/15528014.2017.1305830.

Citrin, A. V., Stem Jr, D. E., Spangenberg, E. R., & Clark, M. J. (2003). Consumer need fortactile input: An internet retailing challenge. *Journal of Business research*, *56*(11), 915-922.

Cooper, K., Dedehayir, O., Riverola, C., Harrington, S., & Alpert, E. (2022). Exploring Consumer Perceptions of the Value Proposition Embedded in Vegan Food Products Using Text Analytics. *Sustainability (Switzerland)*, *14*(4), 1–16. https://doi.org/10.3390/su14042075

Croitoru, A., Crooks, A., Radzikowski, J., & Stefanidis, A. (2013). Geosocial gauge: a system prototype for knowledge discovery from social media. *International Journal of Geographical Information Science, 27*(12), 2483-2508.

Danner, H., & Menapace, L. (2020). Using online comments to explore consumer beliefs regarding organic food in German-speaking countries and the United States. *Food Quality and Preference*, *83*, 103912.

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PloS one*, *6*(12), e26752.

Dondokova, A., Aich, S., Kim, H. C., & Huh, G. H. (2019). A text mining approach to study individuals' food choices and eating behavior using Twitter feeds. In *FrontierComputing: Theory, Technologies and Applications (FC 2018) 7* (pp. 520-527).SpringerSingapore.

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., Jain, V., Karjaluoto, H., Kefi, H., Krishen, A. S., Kumar, V., Rahman, M. M., Raman, R., Rauschnabel, P. A., Rowley, J., Salo, J., Tran, G. A., & Wang, Y. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, *59*(July 2020), 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168

Elbel, B., Kersh, R., Brescoll, V. L., & Dixon, L. B. (2009). Calorie Labeling And Food Choices: A First Look At The Effects On Low-Income People In New York City: Calorie information on menus appears to increase awareness of calorie content, but not necessarily the number of calories people purchase. *Health affairs, 28*(Suppl1), w1110-w1121.

Eskandari, F., Lake, A. A., & Butler, M. (2022). COVID-19 pandemic and food poverty conversations: Social network analysis of Twitter data. *Nutrition Bulletin*, *47*(1), 93–105. https://doi.org/10.1111/nbu.12547

Forshee, R. A., Storey, M. L., Allison, D. B., Glinsmann, W. H., Hein, G. L., Lineback, D. R., ... & White, J. S. (2007). A critical examination of the evidence relating highfructose corn syrup and weight gain. *Critical reviews in food science and nutrition, 47*(6), 561-582.

Ganguly, T. (2018, January 4). Food trends 2018: The ones that will really matter to Indians. *National Restaurant Association of India*. https://nrai.org/food-trends- 2018-the-ones-that-will-really-matter-to-indians/.

GarcIa-Closas, R., GarcIa-Closas, M., & Serra-Majem, L. (1997). A cross-sectional study of dental caries, intake of confectionery and foods rich in starch and sugars, and salivary counts of Streptococcus mutans in children in Spain. *The American journalof clinical nutrition, 66*(5), 1257-1263.

Goryńska-Goldmann, E., Gazdecki, M., Rejman, K., Łaba, S., Kobus-Cisowska, J., & Szczepański, K. (2021). Magnitude, causes and scope for reducing food losses in the baking and confectionery industry —a multi-method approach. *Agriculture (Switzerland)*, *11*(10), 1–20. https://doi.org/10.3390/agriculture11100936

Habiyaremye, C., Matanguihan, J. B., D'Alpoim Guedes, J., Ganjyal, G. M., Whiteman, M. R., Kidwell, K. K., & Murphy, K. M. (2017). Proso millet (Panicum miliaceum L.) and its potential for cultivation in the Pacific Northwest, US: a review. *Frontiers in plant science*, 1961. Doi: 10.3389/fpls.2016.01961.

Hamshaw, R. J. T., Barnett, J., & Lucas, J. S. (2018). Tweeting and Eating: The Effect of Links and Likes on Food-Hypersensitive Consumers' Perceptions of Tweets. *Frontiers in Public Health*, *6*(April), 1–12.https://doi.org/10.3389/fpubh.2018.00118

Hemmerling, S., Hamm, U., & Spiller, A. (2015). Consumption behaviour regarding organic food from a marketing perspective—a literature review. *Organic Agriculture*, *5*, 277-313.

Hilbert, M. (2016). Big data for development: A review of promises and challenges. *Development Policy Review*, *34*(1), 135-174.

Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews.In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).

Humphreys, A., & Wang, R. J. H. (2018). Automated text analysis for consumer research. *Journal of Consumer Research, 44*(6), 1274-1306.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, *53*(1), 59-68.

Karimiziarani, M., Jafarzadegan, K., Abbaszadeh, P., Shao, W., & Moradkhani, H. (2022).Hazard risk awareness and disaster management: Extracting the information content of twitter data. *Sustainable Cities and Society, 77*, 103577.

Kaur, S., Kumar, K., Singh, L., Sharanagat, V. S., Nema, P. K., Mishra, V., & Bhusan, B. (2022). Gluten-free grains: Importance, processing and its effect on quality of gluten-free products. *Critical Reviews in Food Science and Nutrition*, *0*(0), 1–28. https://doi.org/10.1080/10408398.2022.2119933

Kline, C., Shah, N., & Rubright, H. (2014). Applying the positive theory of social entrepreneurship to understand food entrepreneurs and their operations. *TourismPlanning & Development, 11*(3), 330–342.https://doi.org/10.1080/21568316.2014.890126

Konar, N., Gunes, R., Palabiyik, I., & Toker, O. S. (2022). Health conscious consumers and sugar confectionery: Present aspects and projections. *Trends in Food Science and Technology*, *123*(February), 57–68. https://doi.org/10.1016/j.tifs.2022.02.001

Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology based sentiment analysis of twitter posts. *Expert Systems and Applications, 40*, 4065–4074.

Köster, E. P. (2009). Diversity in the determinants of food choice: A psychological perspective. *Food Quality and Preference, 20*, 70–82.

Kraft, F. B., & Goodell, P. W. (1993). Identifying the health conscious consumer. *Marketing Health Services, 13*(3), 18.

Krafft, M., Kumar, V., Harmeling, C., Singh, S., Zhu, T., Chen, J., Duncan, T., Fortin, W., & Rosa, E. (2021). Insight is power: Understanding the terms of the consumer-firm data exchange. *Journal of Retailing*, *97*(1), 133–149. https://doi.org/10.1016/j.jretai.2020.11.001

Lima, J. P. ., Costa, S. A., Brandão, T. R. S., & Rocha. (2021). Food Consumption Determinants and Barriers for Healthy Eating at the Workplace—A University Setting. *Foods*, *10*(4), 695.

Liu, B. (2012). Sentiment analysis: A fascinating problem. In *Sentiment Analysis and Opinion Mining* (pp. 1-8). Cham: Springer International Publishing.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business horizons*, *52*(4), 357-365.

Massagli M.-Frost J. Brownstein C. Okun S. Vaughan T. Bradley R. Wicks, P. and Heywood, J. (2010). Sharing health data for better outcomes on PatientsLikeMe. *Journal of medical Internet research 12*(2), e19. https://doi.org/10.2196/jmir.1549

McCrickerd, K., & Forde, C. G. (2016). Sensory influences on food intake control:moving beyond palatability. *Obesity Reviews, 17*(1), 18-29.

Mediratta, S., & Mathur, P. (2023). Determinants of food choices among adults (20-40years old) residing in Delhi, India. *Current Developments in Nutrition*, 100029.

Michaelraj, P. S. J., & Shanmugam, A. (2013). A study on millets based cultivation and consumption in India. *International Journal of Marketing, Financial Services &Management Research*, *2*(4), 49-58. www.indianresearchjournals.com

Mostafa, M.M. (2013). More than words: social networks' text mining for consumer brand sentiments. *Expert Systems with Applications, 40*(10), 4241-4251. doi:10.1016/j.eswa.2013.01.019.

Mostafa, M. M. (2017). Mining and mapping halal food consumers : A geo- located Twitter opinion polarity analysis Mining and mapping halal food consumers : A geo-located. *Journal of Food Products Marketing*, *24*(7), 858–879. https://doi.org/10.1080/10454446.2017.1418695

Nesmith, J. D. N. (2020). Voices "Herd": A Social and Sentiment Analysis of Consumers Perceptions of Fair Oaks Farms.

Nguyen, Q. C., Meng, H., Li, D., Kath, S., McCullough, M., Paul, D., Kanokvimankul, P., Nguyen, T. X., & Li, F. (2017). Social media indicators of the food environment andstate health outcomes. *Public Health*, *148*(801), 120–128. https://doi.org/10.1016/j.puhe.2017.03.013

Nomisma (2019). Osservatorio Caff_e, available at:
http://www.datalytics.it/osservatoriocaff%C3%A8.html.

Nova, F. F., Coupe, A., Mynatt, E. D., Guha, S., & Pater, J. A. (2022). Cultivating the Community. *Proceedings of the ACM on Human-Computer Interaction*, *6*(GROUP), 1–33. https://doi.org/10.1145/3492826

Oduru, T., Jordan, A., & Park, A. (2022). Healthy vs. Unhealthy Food Images: Image Classification of Twitter Images. *International Journal of Environmental Research and Public Health*, *19*(2), 923. https://doi.org/10.3390/ijerph19020923

Okrent, A. M., & Kumcu, A. (2016). *US households' demand for convenience foods* (No.1477-2017-3961).

Olayanju, J. (2019). Top trends driving change in the food industry. *Forbes*. https://www.forbes.com/sites/juliabolayanju/2019/02/16/top-trends-driving-change-in-the-food-industry/?sh=52112e446063.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, *2*(1–2), 1-135.

Paschen, J., Wilson, M., & Robson, K. (2020). # BuyNothingDay: investigating consumerrestraint using hybrid content analysis of Twitter data. *European Journal of Marketing*, *54*(2), 327-350.

Pater, J. A., Haimson, O. L., Andalibi, N., & Mynatt, E. D. (2016). Hunger Hurts but starving works:" characterizing the presentation of eating disorders online. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, *27*, 1185–1200. https://doi.org/10.1145/2818048.2820030

Pathak, H. C. (2013). Role of millets in nutritional security of India. *New Delhi: National Academy of Agricultural Sciences*, 1-16.

Pindado, E., & Barrena, R. (2021). Using Twitter to explore consumers' sentiments and their social representations towards new food trends. *British Food Journal*, *123*(3),1060–1082. https://doi.org/10.1108/BFJ-03-2020-0192

Rathore, A.K., Ilavarasan, P.V. & Dwivedi, Y.K. (2016). Social media content and productcocreation: an emerging paradigm. *Journal of Enterprise Information Management*, *29*(1), 7-18, doi: 10.1108/JEIM-06-2015-0047.

Recuero-Virto, N., & Valilla-Arróspide, C. (2022). Forecasting the next revolution: food technology's impact on consumers' acceptance and satisfaction. *British Food Journal*, *124*(12), 4339–4353. https://doi.org/10.1108/BFJ-07-2021-0803

Samoggia, A., Riedel, B., & Ruggeri, A. (2020). Social media exploration for understanding food product attributes perception: the case of coffee and health with Twitter data. *British Food Journal*, *122*(12), 3815–3835. https://doi.org/10.1108/BFJ-03-2019-0172

Saura, J. R., Reyes-Menendez, A., & Thomas, S. B. (2020). Gaining a deeper understanding of nutrition using social networks and user-generated content.*Internet Interventions, 20*, 100312.

Shah, P., Dhir, A., Joshi, R., & Tripathy, N. (2021). Drivers and barriers in the consumption of alternative staples. A systematic literature review and futureresearch agenda. *British Food Journal, 123*(11), 3726–3759. https://doi.org/10.1108/bfj-12-2020-1098.

Singh, A., & Glińska-Neweś, A. (2022). Modeling the public attitude towards organic foods: a big data and text mining approach. *Journal of Big Data*, *9*(1), 1–21.https://doi.org/10.1186/s40537-021-00551-6

Singh, A., Shukla, N., & Mishra, N. (2018). Social media data analytics to improve supplychain management in food industries. *Transportation Research Part E: Logistics and Transportation Review*, *114*, 398–415.https://doi.org/10.1016/j.tre.2017.05.008

Smith, K. T. (2011). Digital marketing strategies that Millennials find appealing, motivating, or just annoying. *Journal of Strategic marketing, 19*(6), 489-499.

Sproagis, U., & Rikters, M. (2020). What can we learn from almost a decade of food tweets. *Frontiers in Artificial Intelligence and Applications*, *328*(April), 191–198. https://doi.org/10.3233/faia200622

Statista (2022). Worldwide confectionery consumption by category. Statista. https://www.statista.com/forecasts/1310409/worldwide-confectionery- consumption-by-category

Stefanidis, A., Crooks, A., & Radzikowski, J. (2013). Harvesting ambient geospatialinformation from social media feeds. *GeoJournal*, *78*, 319-338.

Talwar, M., Talwar, S., Kaur, P., Tripathy, N., & Dhir, A. (2021). Has financial attitude impacted the trading activity of investors during the COVID-19 pandemic? *Journal of Retailing and Consumer Services, 58*, Article 102341.https://doi.org/10.1016/j.jretconser.2020.102341

Tandel, K. R. (2011). Sugar substitutes: Health controversy over perceived benefits.*Journal of Pharmacology and Pharmacotherapeutics, 2*(4), 236-243.

Tao, D., Yang, P., & Feng, H. (2020). Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive reviews in food science and food safety, 19*(2), 875-894.

Taylor, J. R., & Emmambux, M. N. (2008). Gluten-free foods and beverages from millets.In *Gluten-free cereal products and beverages*. *Academic Press,* 119-148. https://www.researchgate.net/279430250

Tibra, M., Saxena, A., Caytiles, R. D., & Iyengar, N. C. S. N. (2017). Predicting Geo- located Food Based Sentiment Analytics using Twitter for Healthy Food Consumption across India. *Advanced Science and Technology Letters*, *143*(February), 235–239. https://doi.org/10.14257/astl.2017.143.47

Vatambeti, R., Mantena, S. V., Kiran, K. V. D., Manohar, M., & Manjunath, C. (2023). Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique. *Cluster Computing*, *7*, 1–17. https://doi.org/10.1007/s10586-023-03970-7

Veena, B. (2003). Nutritional, functional and utilization studies on barnyard millet. M. Sc. Thesis. Submitted to University of Agricultural Sciences, Dharwad, Karnataka, India.

Vidal, L., Ares, G., & Jaeger, S. R. (2016). Use of emoticon and emoji in tweets for food- related emotional expression. *Food Quality and Preference*, *49*, 119–128. https://doi.org/10.1016/j.foodqual.2015.12.002

Vydiswaran, V. G. V., Romero, D. M., Zhao, X., Yu, D., Gomez-Lopez, I., Lu, J. X., Iott, B., Baylin, A., Clarke, P., Berrocal, V., Goodspeed, R., & Veinot, T. (2018). “Bacon baconbacon”: Food-related tweets and sentiment in metro detroit. *12th International AAAI Conference on Web and Social Media, ICWSM 2018*, *Icwsm*, 692–695. https://doi.org/10.1609/icwsm.v12i1.15060

Vydiswaran, V. G. V., Romero, D. M., Zhao, X., Yu, D., Gomez-Lopez, I., Lu, J. X., Iott, B. E., Baylin, A., Jansen, E. C., Clarke, P., Berrocal, V. J., Goodspeed, R., & Veinot, T. C.(2020). Uncovering the relationship between food-related discussion on Twitter and neighborhood characteristics. *Journal of the American Medical Informatics Association*, *27*(2), 254–264. https://doi.org/10.1093/jamia/ocz181

Wang, T., Mentzakis, E., Brede, M., & Ianni, A. (2019). Estimating determinants of attrition in eating disorder communities on twitter: An instrumental variables approach. *Journal of Medical Internet Research*, *21*(5), 1–16. https://doi.org/10.2196/10942

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

Wansink, B., Painter, J. E., & Lee, Y. K. (2006). The office candy dish: proximity'sinfluence on estimated and actual consumption. *International journal of obesity, 30*(5), 871-875.

Wirfält, E., Drake, I., & Wallström, P. (2013). What do review papers conclude about food and dietary patterns? *Food & nutrition research, 57*(1), 20523.

Worch, T. (2014). What should you know about analysing social media data using twitteR: The experience of a practitioner. In 6th European Conference on Sensory and Consumer Research, 7–10 September 2014. Copenhagen, Denmark.

Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firmequity value: A sentiment analysis approach. *Decision support systems*, *55*(4), 919-926.

Zhou, S., Zhao, Y., Bian, J., Haynos, A. F., & Zhang, R. (2020). Exploring eating disorder topics on twitter: Machine learning approach. *JMIR Medical Informatics*, *8*(10), e18273. https://doi.org/10.2196/18273