The Lifecycle of a Digital Food Trend

Kadeeja Zumreen, Sriya Kondury, Sayalee Chivate

30 November 2025

1 Abstract

"The Lifecycle of a Digital Food Trend" project analyzes how digital food trends rise, peak, and decline across online platforms from 2015 to 2025. For the purpose of this project, four major trends were targeted, namely, Baked Feta Pasta, Matcha, Dubai Chocolate, and Air Fryer. This analysis used data from Youtube, Reddit, and Google Trends to study which behavioral, geographic, and engagement factors affect the lifecycle of online food trends and whether early digital activity can predict a trend's popularity and longevity.[7, 5, 8]. The overarching research goal is to understand how the sentiment and themes of online discussions on social media reflect and predict the lifecycle of modern food trends.

Live data from YouTube Data API, Reddit's PRAW API, and the Google PyTrends Interface was collected, cleaned and then processed in Python to analyze the four trends. Analytical techniques like time series visualization, geographic choropleth mapping, sentiment analysis (VADER), topic modeling (LDA), linear regression, and Random Forest modeling were used throughout the project to address the sub-questions and the broader overarching question. To understand cross platform engagement patterns, statistical comparison was applied and then later on it was used to evaluate the relationships between search interest, content creation, and community discussion.

The results revealed distinct diffusion patterns and showed how each trend had a different lifecycle. Feta Pasta experienced short lived viral spikes, Matcha experienced steady growth and cultural adoption, Air Fryer became a practical household item and Dubai Chocolate became a product of an influencer marketing. The cross platform analysis showed that each trend appeared first in Google search activity and that fed into public curiosity. Then it was observed in videos and shorts on YouTube. And lastly it is seen in online discussions in Reddit which tend to last even when the trend loses its virality.[7, 5, 8]. Analyzing the geographic diffusion of the four trends showed how widespread their virality was and how they differed from each other. For instance, Matcha and Air Fryer were popular globally whereas Dubai Chocolate remained concentrated regionally. Using a combination of sentiment analysis and topic modeling, the four trends were further classified into long term trends and short lived viral trends based on the diverse and positive discussions on the online platforms. The machine learning results show that initial popularity was the strongest predictor of visibility on YouTube. Overall, these results illustrate how digital food trends change over time and how they are influenced by several factors like cultural context, early engagement signals, platform dynamics, etc. The results provide evidence that these factors are crucial in understanding and predicting digital trend's lifecycles.

2 Introduction

Food trends have rapidly spread online with the rise in social media [2]. Social media platforms like Tiktok, Instagram, Reddit, and YouTube revolutionized how recipes and ingredients gain worldwide popularity. These strongly affect customer behavior, inventories, and create new conversations around food [3, 6]. Current studies regarding food related conversations are fairly limited and are not enough to fully study and understand the lifecycle of a food trend. A food trend needs to be studied from its emergence to understand what caused it, to why it became popular, and then what caused the trend to disappear. Studies show that discussions about food on social media are vital in understanding real world behaviors and people's perceptions of food. This automatically links online discussion to public health outcomes, geographic identity, and dietary patterns. [4, 1] While there are studies that track discussions regarding food trends, there are not many studies which examine how a trend spreads throughout the world or even how early signs can be critical in predicting a trend's trajectory.

This project aims to use previous studies where online discussions were analyzed as a stepping stone to predict the lifecycle of food trends across three main platforms, namely, YouTube [5], Reddit [8], and Google Trends [7]. All trends behave differently on these platforms and that gives a unique perspective of people's opinions of a trend. For instance, YouTube captures content creation and influencer marketing, Reddit supplies community level discussion and sentiment, and Google Trends supplies search history, showing public curiosity regarding a trend. By using all these sources, this project aims to study the peak, decline, longevity of a food trend and also examine factors like the social, geographical and behavioral mechanisms that drive a food trend.

This paper builds off of two studies which are Diffusion of Innovation Theory [9] and Simmel's [10] Fashion Theory. Rogers' study models adoption as an S-curve, showing that most trends follow a predictable pattern. Some people try something new, in Rogers' case he calls them innovators and then a few more people join in, the early adopters and they are followed by the last group to join in, the laggards. When it comes to social media and viral trends, influencers and creators act as early adopters who make the trend viral. Simmel's theory gives a complimentary approach to Rogers' theory as it talks about how a trend moves in a cycle. Some people try it, it becomes unique. When the number of people trying it increases, the trend loses its uniqueness and becomes normal and then it drops. Then influencers start a new trend and this cycle of rise and fall continues. When combined, these two studies help in understanding that viral trends on social media are patterns which can be measured and studied over time to predict their longevity.

Using these two studies as a starting point, this project targets three primary research questions:

1. **What factors drive the rise and decline of food trends across digital platforms?**
2. **Can early digital activity predict a trend's peak popularity and longevity?**
3. **How do food trends originate and spread geographically and culturally across markets?**

This project addresses these questions by conducting exploratory data analysis, inferential statistics, sentiment analysis, and machine learning. It specifically uses cross platform modeling, regional mapping, and topic analysis to understand the spread of these specific food trends. This project structure examines how food trends move through Google Trends, YouTube, and Reddit and how the patterns differ between short and long term trends. The lifecycle of digital food trends offers valuable insights into modern cultural diffusion, customer behavior, and how social media shapes global food conversations. This research combines sociological theory with statistical analysis to contribute to relevant studies and the fields of marketing, public health, and predictive analytics.

3 Methods

3.1 Data Sources

This project targets data from three major digital platforms, YouTube [5], Reddit [8], and Google Trends [7]. The objective was to use three different platforms to analyze the differences in online engagement with food trends between the decade of 2015 to 2025. The four trends we chose for this project behave differently throughout the three platforms, with Youtube representing content creation, Reddit showing sentiment across online communities and Google Trends capturing public curiosity through search activity.

YouTube: For the Youtube platform, YouTube Data API v3 with python scripts was used to collect data [5]. This API gave metadata and provided engagement metrics for trending videos for our four trends which were Baked Feta Pasta, Matcha, Dubai Chocolate, and Airfryer. For each trend, the script came up with almost 500 most recent videos, including attributes such as title, description, publication date, view count, comment count, video duration, and channel details. In total, over 2000 videos were extracted into the dataset for all four of the trends to analyze the engagement behavior and visibility dynamics.

Reddit: For the Reddit platform, the Python Reddit API Wrapper (PRAW) was used to collect the data.[8] This data mainly consisted of community driven discussion and engagement regarding the four trends. To capture the data needed, a global keyword search of the word "food" was conducted across all of Reddit to capture any and all viral discussions and subreddits. This resulted in data retrieval from over thirty relevant communities such as r/food, r/recipes, and r/veganrecipes, etc. Similar to YouTube, for each post the python script came up with attributes such as title, subreddit, upvote count, comment count, author, URL, timestamp, and associated search word. This dataset was particularly important for the project as it is the one where we can perform sentiment analysis and analyze the tone, engagement style, evolution of food trends and use it to compliment the other two data sources by highlighting the importance of discussion rather than just content based engagement.

Google Trends: For the third data source, PyTrends was used to get the data for Google Trends. PyTrends is an unofficial Python API that captures data from Google's search analytics platforms. [7]. For the four trends, Matcha, Baked Feta Cheese Pasta, Dubai Chocolate and Air Fryer, weekly search interest was gathered for the targeted decade of 2015-2025. To address the third sub-question about the geographical spread of the trends, four main regions were taken into consideration, namely, United States (US), India (IN), United Arab Emirates (AE), and South Korea (KR). The data was collected on both the global level and specifically for all four of the countries to have a country-wise comparison of each trend. This approach provides both global and regional perspectives on how food trends evolve and align across markets.

3.2 Data Cleaning and Feature Engineering

Youtube: The data was collected, cleaned, and processed in Python for consistency. The timestamps were converted to Coordinated Universal Time (UTC) and the video durations were standardized into seconds for cross comparison. To capture engagement behavior and dynamics, derived variables were created. The engagement rate was calculated as:

$$\text{Engagement Rate} = \frac{\text{Likes} + \text{Comments}}{\text{Views}}$$

This variable checked the fairness in user interaction with videos, as there is a difference in visibility per video. The video age was also calculated by counting the days between collection and posted date to help analyze engagement patterns over time. A title length variable was created to examine influencers of visibility. Videos were also placed into short or long groups based on video duration. Finally, trend labels were created for each video to allow for trend comparisons within YouTube.

Reddit: Reddit posts were collected using the Python Reddit API Wrapper (PRAW) and placed together in a dataframe. First, duplicate and deleted posts were removed to clean the raw data. For consistent analysis, posts without titles or authors were removed as well. Time markers were changed to standard format and length of post, Trend ID, and amount of interaction received(upvotes, comments, etc.). For preparation for modeling steps, all letters were lowercased, website links were removed, and punctuation marks were taken out. Finally, a trend label was added to each post (ex: 'matcha', 'feta-pasta') to easily compare different food trends. The final dataset was then used to track the main topics of discussion and sentiment over time.

Google Trends: The data were collected using the PyTrends API and then cleaned and processed for consistency across regions and timeframes [7]. The weekly normalized Search Interest Values (0-100) were then verified and partial results marked as (isPartial = True) were discarded. The dataset was composed of the keywords and geographical regions, and the values were aligned by week, and the column names were made consistent for easier comprehension (ex: matcha_US, matcha_GB). We then built automated pauses and retries into the retrieval process to follow Google's rate limits.

The dataset was then validated for the consistent date years 2015 to 2025 and exported to a .csv file for easy reproduction. Python libraries such as pandas, numpy, and pycountry were used during the data cleaning process. The final dataset is a great basis for inter-regional and longitudinal research on the spread of food trends.

3.3 Exploratory Data Analysis

The exploratory data analysis of this project concentrated on global, behavioral and spatial patterns across the three data sources. To analyze the three datasets, Python libraries, such as pandas, numpy, matplotlib, seaborn and plotly. We performed descriptive statistics to analyze engagement, frequency, and search metrics. To explore the structure and relationships within and across platforms, we used various visualization tools to picture the data.

To examine patterns of activity over time in the YouTube data, we generated daily, weekly and monthly aggregations of video uploads, views, and engagement rates[5]. Similarly to capture geographic diffusion and spread of the trend using Google Trends data, search indexes were normalized and analyzed. Then this data was used to make choropleth maps and study the regions where the trends originated[7]. Lastly, to understand discussions regarding trends, reddit data was explored through distributions of various posts, their comments and engagement levels across subreddits[8]. We performed preliminary sentiment analysis and word cloud visualization on Reddit text data to capture the tone and most common words in user discussions. This was crucial in providing an initial perspective on public sentiment towards each trend.

The main idea behind the EDA was to validate data integrity, identify key relationships among the variables, guide the selection of features and parameters for subsequent regression, machine learning and topic modeling. The EDA was also beneficial for understanding the geographical patterns of each trend which in turn answered the third research question on how food trends spread across various regions and the context for their point of origin.

3.4 Statistical and Machine Learning Methods

Regression Analysis (Google Trends): We used simple linear regression for each region with OLS in Python to study the data except for global averages. T-tests were used to show mean differences between regions if one region had significantly higher search interest than another. This analysis showed whether the time or region had an effect on the search interest in Matcha.

Youtube: We trained a random forest regression model to predict the amount of views a video would get, using features like count, comment count, trend ID, video duration, and video age. The feature importance analysis showed that initial popularity, especially likes and comments, was the main predictor of video popularity. The model's accuracy was then tested by comparing its predictions against the view counts, which proved a close match and small bias.

Topic Modeling (Reddit): VADER analysis was used on the Reddit posts to measure the sentiment of the food trend discussions. For the topic analysis of the Reddit posts, the titles were standardized using an NLP pipeline. A document term matrix was created with CountVectorizer (1,000 max features; min df = 5; max df = 0.85) and LDA with $K \in \{3, 5, 7, 10\}$ were created, and the model was then evaluated. K= 5 was the most interpretable for this study, and was used to assign the most probable topic to the Reddit threads and enable the comparison of topics and their trends.

3.5 Analytical Framework and Tools

The data preprocessing and exploratory analysis were done in Python using Pandas, Numpy, Scikit-Learn, Gensim, and Plotly. The geographic and statistical visualizations were done on Jupyter Notebook for easy reproducibility. The full process from data collection to data visualization and predictive modeling is documented and available in the public GitHub repository mentioned in the appendix.

4 Results

4.1 Descriptive and Comparative Analysis of Trend Behavior

4.1.1 Geographic Diffusion

To answer our third sub-question and explore how food trends diffuse geographically across online markets and communities, we analyzed Google Trends data for India, the United States, the United Arab Emirates, and South Korea. Figures 2-5 show maps of search interest for the four selected trends. The patterns in the geographical maps reveal both regional and cultural differences in the global diffusion of digital food trends.

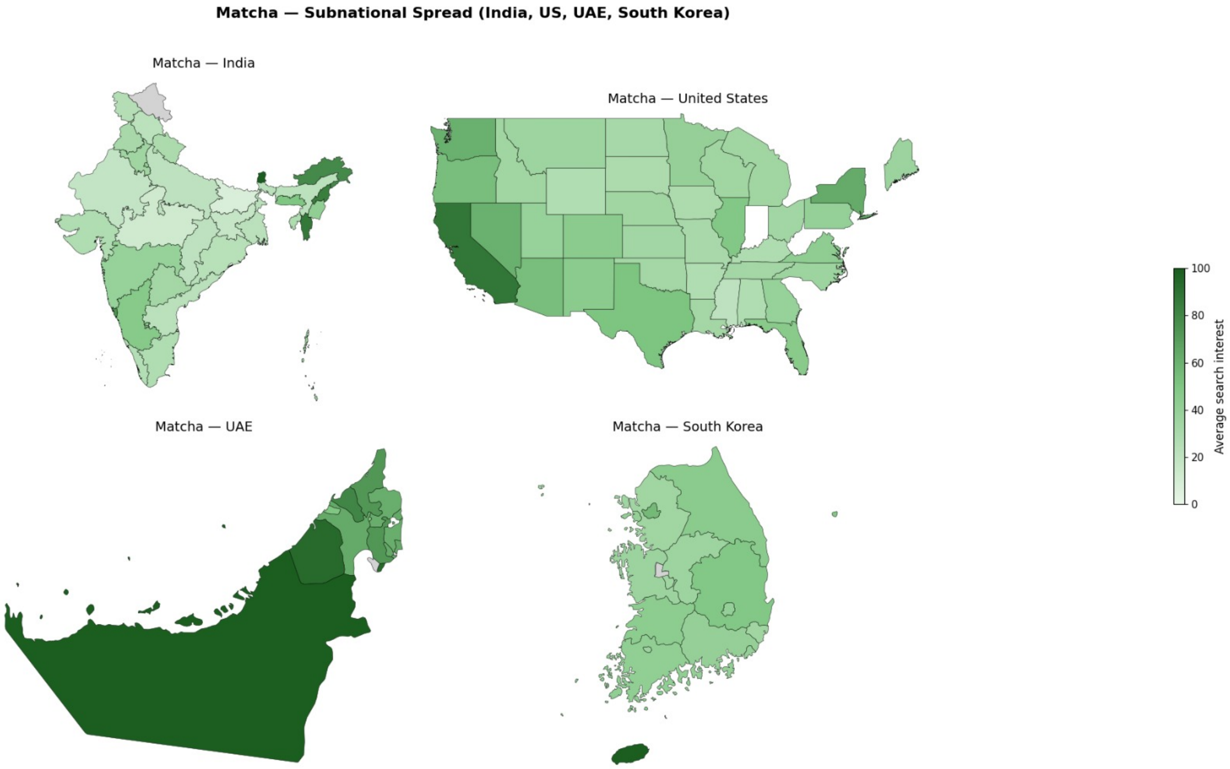


Figure 1: Regional Mapping for Matcha in India, US, UAE and South Korea

The Matcha trend shows that it stems from wellness culture and global cafe aesthetics. In the United States, search activity is higher along influencer driven states like California where the cafe culture is booming and influencers promote health oriented drinks like matcha. In South Korea and the UAE, the search interest is spread out showing that matcha is incorporated in regular beverages and desserts. In India however, the search interest is highest in the tea producing states. This pattern shows that matcha had a sustained growth and is still culturally relevant and integrated in various food items.

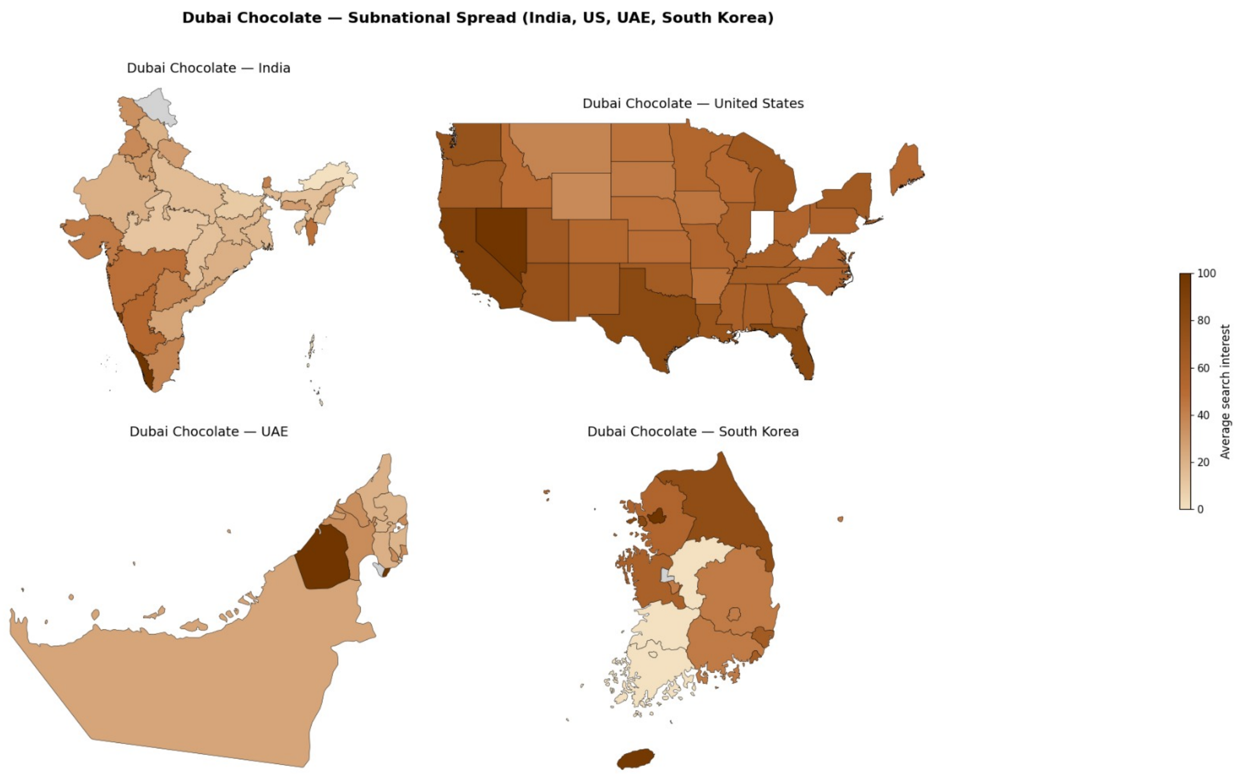


Figure 2: Regional Mapping for Dubai Chocolate in India, US, UAE and South Korea

Dubai Chocolate is vastly different from Matcha as it presents the perfect example of an influencer driven trend. Places like UAE, particularly Dubai, show the highest search interest which aligns with the fact that Dubai is the point of origin for Dubai Chocolate. Moreover, the trend's popularity in Dubai confirms the trend's emergence as a luxury sweet made popular by local influencers. The search interest in countries like India and South Korea is pretty limited aside from specific regions in the countries, showing that even though it is mostly popular in UAE, the influencer culture has driven it out of the country through social media as compared to the traditional methods of local adoption. In countries like the United States, the interest is fragmented and not particularly centered around one place, rather it is pretty spread out, showing the trend's virality throughout the country.

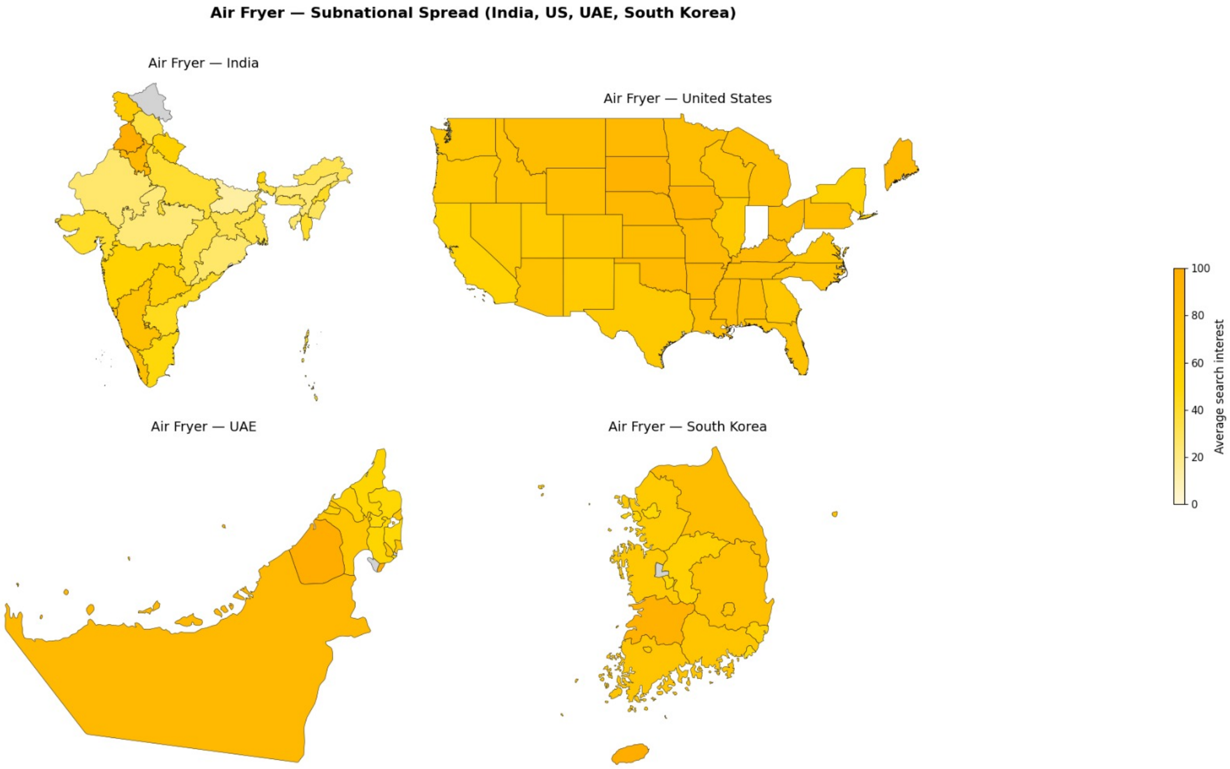


Figure 3: Regional Mapping for Airfryer in India, US, UAE and South Korea

The Air Fryer trend is different from all the other trends because instead of categorizing it as a trend, it is more of a household item. It had a steady, consistent growth before it was permanently adopted as a kitchen appliance. Across all the countries we have chosen, the interest is wide spread and it is consistent throughout all the regions, showing that it is adopted by everyone in their everyday lives. Countries like the United States show near universal engagement showing that the Air Fryer is not just a trend but a very useful item. Similarly countries like India and UAE show consistent growth of the trend throughout all the regions. In countries like South Korea however, the map shows that the search interest is nearly uniform, indicating that the Air Fryer is so well known that it is used in everyday cooking. This even distribution of this trend shows a sustained, utility driven adoption process rather than just ignoring it because it is a trend.

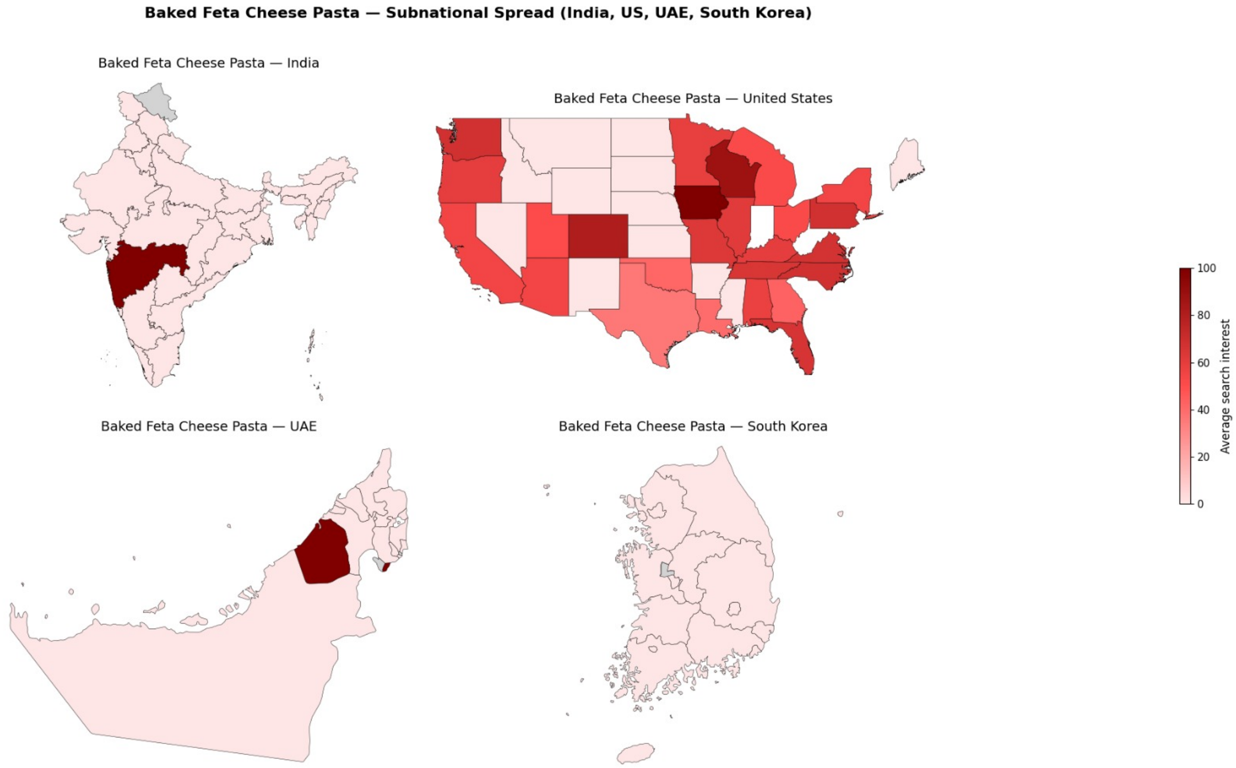


Figure 4: Regional Mapping for Feta Pasta in India, US, UAE and South Korea

Baked Feta Cheese Pasta is a lesser known trend and is the perfect example of a viral-burst trend. It became popular during the pandemic, showing its advantage as an easy dish to make at home but it didn't last for long. It disappeared as soon as it became viral, originating in the United States. It peaked in early 2021 with the search interest being strongest in the central US which aligns with its growing popularity on TikTok and other social media platforms during the pandemic period. Countries like UAE and South Korea show small, clustered, limited search interests in some places, which aligns with the fact that Baked Feta Cheese Pasta was a brief trend. Countries like India have only one state where the search interest is highest, showing its brief exposure through social media. This trend fits the pattern of a typical viral lifestyle with a rapid rise and then a sharp decline as soon as it gained some traction.

Together, these spatial patterns reveal three distinct diffusion models across digital food trends: (1) sustained cultural adoption, as in the case of Matcha; (2) localized influencer driven diffusion, represented by Dubai Chocolate; and (3) transient viral burst diffusion, represented by Baked Feta Cheese Pasta. The Air Fryer trend occupies a fourth category of technological adoption that has consistent and practical engagement. These models underscore the multifaceted ways digital food trends move across platforms, cultures, and geographies.

4.1.2 Sentiment and Discussion Analysis

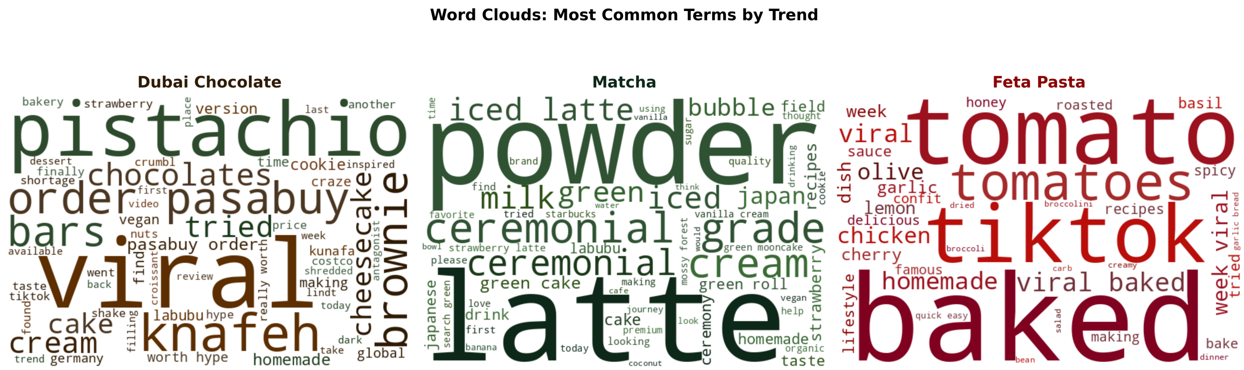


Figure 5: Word Clouds: Most Common Terms by Food Trends

This figure depicts the most frequently used terms for each trend. The word clouds uncover unique patterns that represent how users discuss these foods online. For feta pasta, these posts have an importance on recipe creation and virality with words like “tomato”, “tiktok” and “baked”. This reflects that feta pasta was a short term viral recipe. Matcha, on the other hand, has discussions that focus around preparation with terms like “ceremonial”, “powder”, and “latte”. This differs from feta pasta as matcha’s terms have more of an emphasis on everyday consumption rather than a short term viral trend. Meanwhile, Dubai Chocolate posts are more purchasing and spending forward than the other two, as they contain terms like “pistaschio”, “order”, and “pasabuy”.

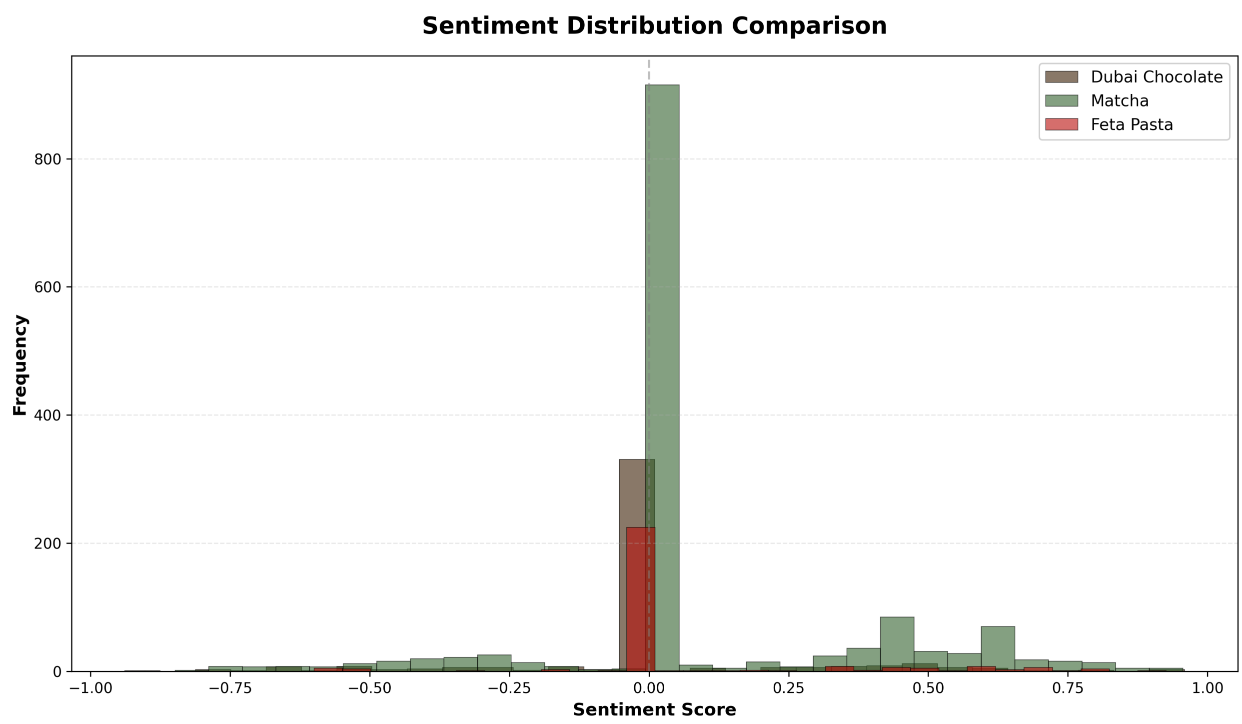


Figure 6: Sentiment Distribution Comparison

This figure shows the distribution of sentiment across Reddit discussions for each food trend. The majority of the posts cluster near zero for all 3 trends, suggesting that most discussions are more neutral compared to strongly emotional. Matcha however, does have a wider spread toward positive sentiment. This may be due to Matcha’s associations with health and everyday consumption. On the other hand, Feta Pasta has slightly more negative posts, which may be due to a decline in enthusiasm after the trend faded. Dubai Chocolate has a narrow and modest positive sentiment implying that it has favorable but less engagement.

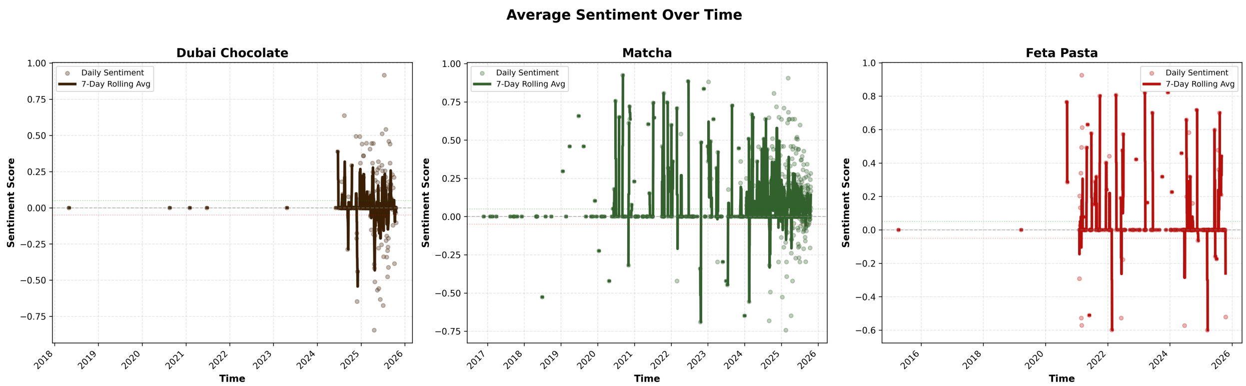


Figure 7: Average Sentiment Over Time

This plot depicts the sentiment for each trend with daily sentiment scores and a 7 day rolling average. Matcha has a stable and slightly positive sentiment over time. This may be because matcha has been adapted into everyday routines and therefore has sustained its popularity. Feta pasta has changed drastically over time. It had a spike in 2021 when it was viral before dropping as the interest declined. The sentiment for Dubai Chocolate is still increasing in 2025, which shows the growing popularity of this continuing trend. The patterns seen in this plot show how long term trends are able to sustain positive sentiment, while short term viral trends have temporary peaks in enthusiasm before inevitably decaying.

4.1.3 Trend Dynamics

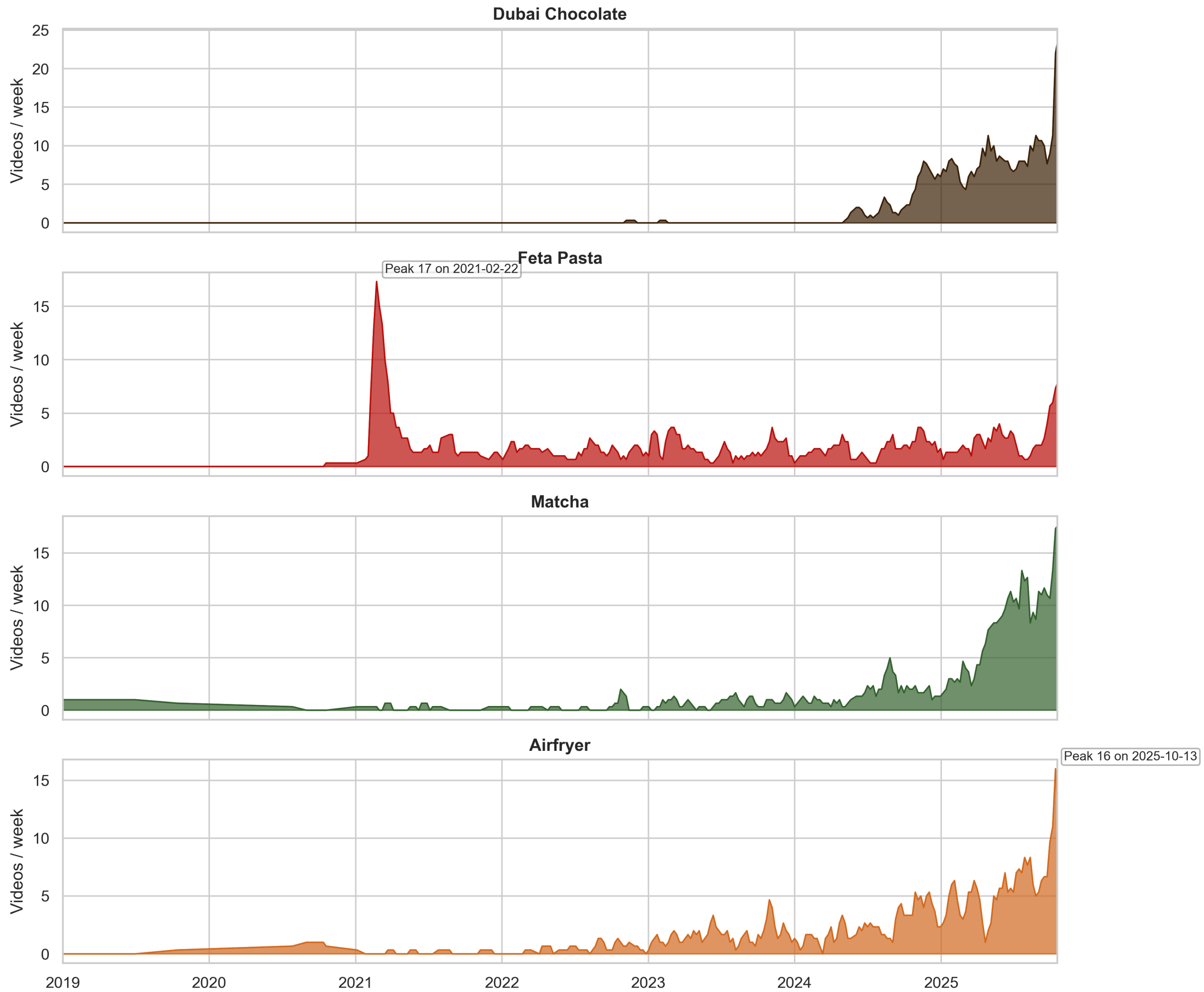


Figure 8: Weekly Youtube video trends for Dubai Chocolate, Feta Pasta, Matcha, and Airfryer.

Figure 8 shows the number of weekly Youtube videos uploaded for each trend between 2019 and 2025. The plot shows the unique lifecycle patterns of the four food trends: Air Fryer, Feta Pasta, and Dubai Chocolate. In early 2021, the feta pasta saw a sharp and brief spike that coincided with its viral social media spread during the COVID-19 lockdown. The short lived nature of novelty recipes is highlighted by the sharp rise and equally steep decline. After its peak, video production quickly diminished, indicating the trend did not sustain creator interest over time. In contrast, matcha and airfryer show steadier growth beginning in 2022, with periodic fluctuations that indicate sustained audience interest over time. Beginning in 2024, Dubai Chocolate shows a late but quick surge, emerging as the most popular upload category by 2025. The steep upward path suggests that the trend is at an early stage of diffusion and still spreading at a faster rate.

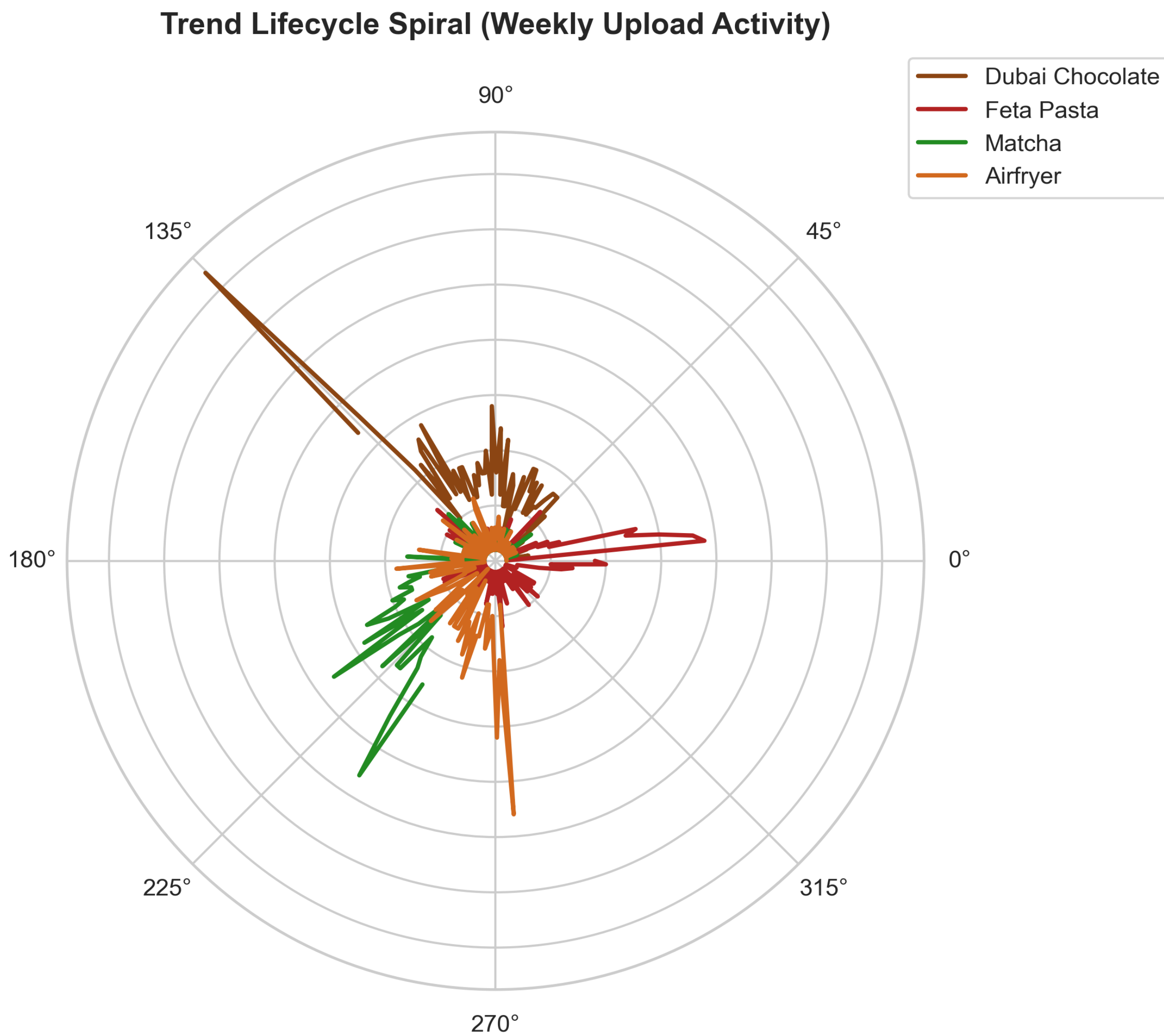


Figure 9: Lifecycle of Trends by Weekly Upload Activity

Figure 9 shows a spiral visualization, which maps weekly activity outward from each trend's point of emergence, providing a comparison of the upload paths taken by the different food trends. The compact, tightly clustered spiral for Feta Pasta reflects its condensed lifecycle of rapid ascent, sharp peak, and sudden decline. The short duration of the trend's prominence is visually reinforced by the concentration of activity close to the inner rings. On the other hand, the spirals for Matcha and Air Fryer show long-term relevance as they grow steadily outward with a consistent density. The broader and more evenly distributed rings show that these trends experienced longer periods of content creation, with Matcha in particular showing continuous growth aligned with its cultural and health-oriented associations. With rapid growth in the last few months, Dubai chocolate shows an outer ring explosion. Its sudden entry into the world of food trends and its quickly growing popularity are visually depicted by this pattern.

4.1.4 Cross-Platform Analysis

Weekly activity on Youtube (content creation and views), Google Trends (search curiosity), and Reddit (community discussion) was compared in order to understand how digital food trends vary across these three platforms. Figure 10, which consists of a multi-panel visualization shows how each platform reacts over time to the same four food trends, revealing different emergence, amplification, and saturation courses throughout the various digital media.

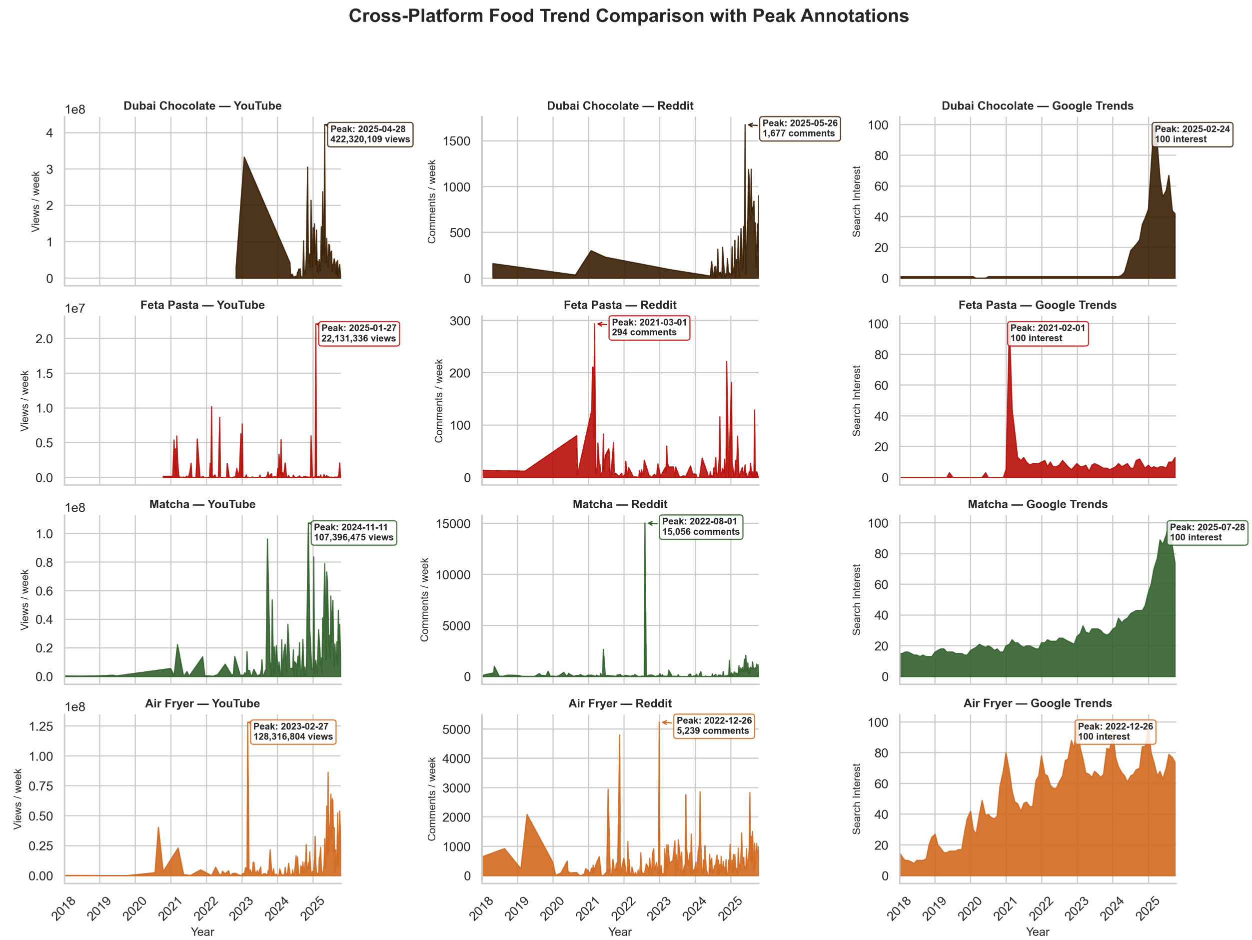


Figure 10: Weekly cross-platform activity for four food trends across YouTube, Reddit, and Google Trends from 2018 to 2025

Each trend exhibits a unique lifecycle pattern across the three digital platforms, as seen in Figure 10, depicting how public interest changes across various digital engagement platforms. Early in 2025, Dubai Chocolate saw a dramatic increase in Google search activity (February 2025), which was followed by a peak in YouTube video views (April 2025), and Reddit comments (May 2025). This pattern implies that the trend spread from early consumer interest to content production, and eventually to community level discussions.

In contrast, Feta pasta exhibits a highly condensed viral cycle that is concentrated in early 2021 on Reddit [8] and Google Trends [7]. In February 2021, there was an immediate global spike in Google Trends, which was immediately followed by an increase in Reddit discussion (March, 2021). Around this time, we can see a peak in YouTube views as well, even though the highest peak is observed to be in early 2025.

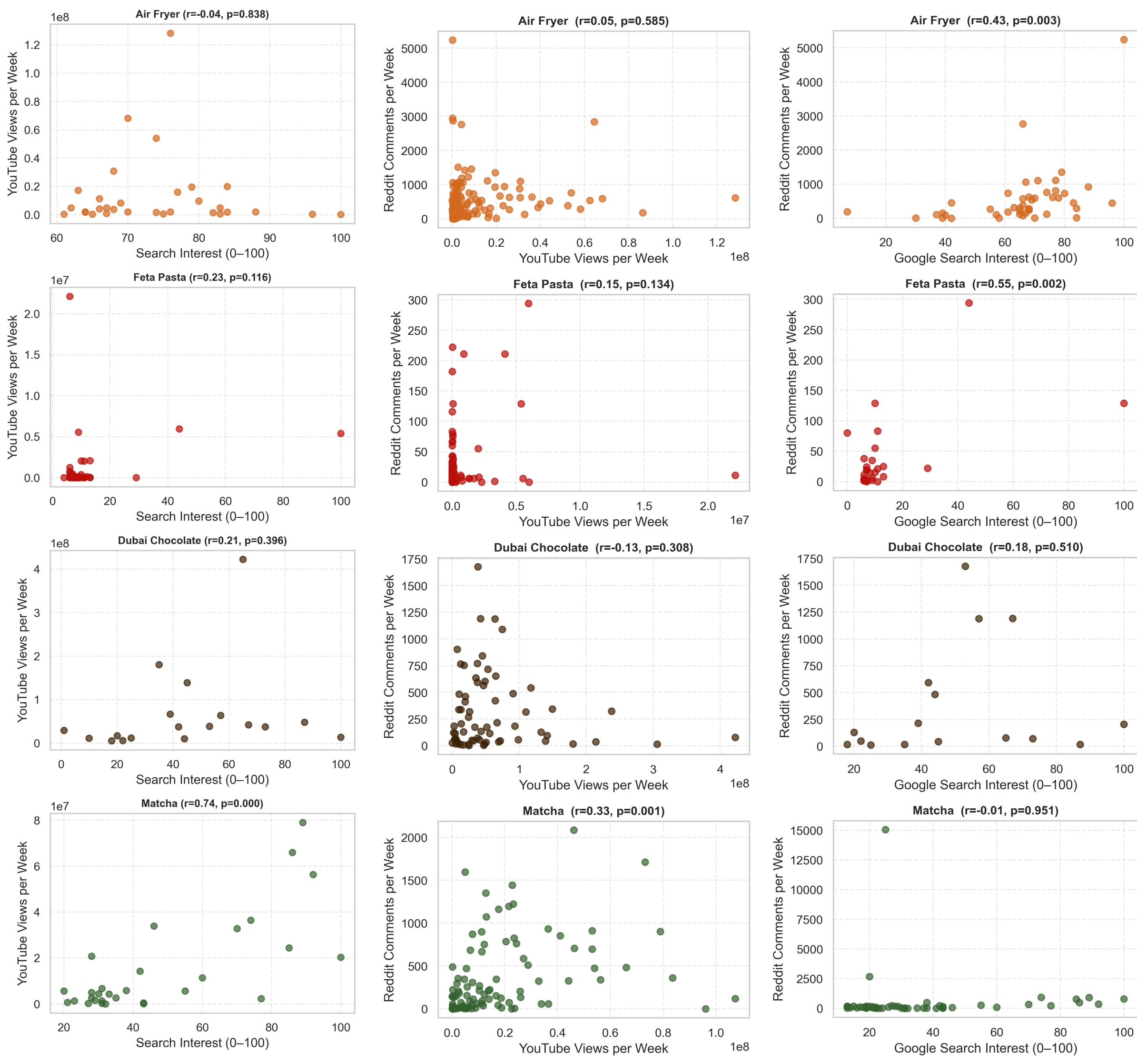
The growth of matcha is steady and long-lasting, with visible peaks across all three platforms. Between late 2024 and mid 2025, both Youtube [5] and Google Trends [7] show a strong increase in activity with major peaks on November 11 2024 (107 million views on YouTube) and July 28 2025 (Google Interest = 100). This suggests that the matcha trend experienced another global popularity wave during this time period, driven by continuous public interest and media exposure.

However, Reddit data [8] shows a spike for August 2022 that is unusually high at 15,056 comments, an outlier. This single large peak overshadows the rest of the Reddit timeline and makes it difficult to see smaller increases that likely occurred during 2024-2025. Even though those later fluctuations are less visible on the same scale, the overall pattern still suggests that Reddit activity followed a similar upward trend to YouTube [5] and Google Trends [7] toward 2025.

Finally, Airfryer shows overlapping peaks across platforms; Google Trends [7] show maximum engagement in December 2022 along with Reddit [8], while YouTube [5] shows its peak in February 2023. The consistency over the years reflects a stable utility driven adoption rather than novelty driven virality.

Overall, these patterns show that Google Trends [7] typically captures the earliest rise in public curiosity, followed by YouTube activity [5] as creators amplify interest through content, and finally Reddit [8] discussions as communities engage and share experiences. Feta Pasta represents a short-term viral trend with a sharp, synchronized peak across all platforms, while Matcha and Air Fryer maintain steady engagement over several years, reflecting lasting cultural adoption. Dubai Chocolate, a more recent trend, shows a strong upward trajectory across all platforms but is too recent to determine its longevity. This pattern of search to creation to discussion, indicates distinct phases of discovery, amplification, and sustained participation of digital food trends within online platforms.

To quantify these relationships, weekly activity levels across platforms were compared using scatterplots and Pearson correlation coefficients (Figure 11). This is an analytical way to show how changes in search interest correspond to changes in video engagement or community discussion, and how these relationships vary by trend.



(a) YouTube and Google Trends (b) Youtube and Reddit (c) Google Trends and Reddit

Figure 11: Pairwise correlation plots showing relationships between weekly activity across all 3 platforms for all 4 trends

The results show that Matcha has the highest positive correlation between Youtube [5] and Google Trends [7] with r = 0.74, $p < 0.001$, indicating that public search interest and video engagement increased together during its peak in 2024–2025. Feta Pasta also shows moderate positive correlations, with simultaneous viral spikes across platforms. Air Fryer shows weaker or insignificant correlations in most combinations, reflecting its steady, non-viral pattern of engagement. Dubai Chocolate correlations are mild and not statistically significant, likely because its rise is recent and still evolving. These comparisons emphasize that search interest and video activity tend to move together more closely than community discussions on Reddit [8], which often happens independently of initial trend discovery.

4.2 Statistical and Machine Learning Results

4.2.1 Regression Analysis

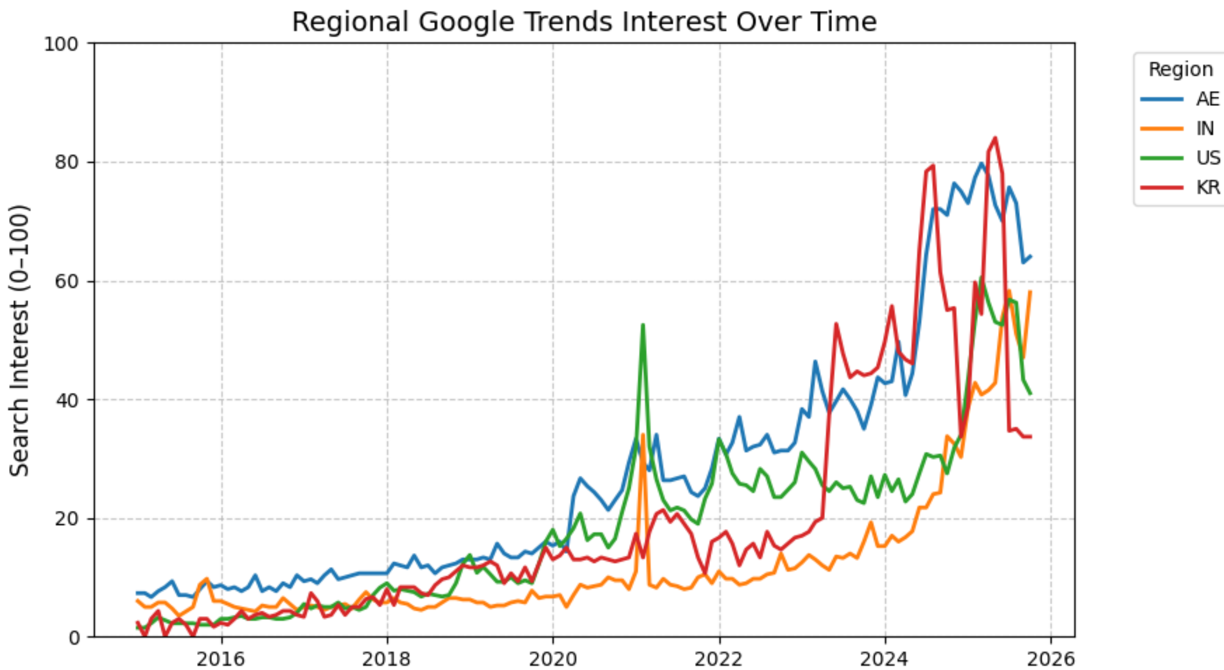


Figure 12: Regression Analysis of Google Trends

Matcha: Inferential statistics were used to analyze if region or time had a significant effect on search interest in Matcha. A simple linear regression (interest $\sim$ time) approach was conducted for each region with OLS in Python except for global averages. T-tests were used to demonstrate mean interest differences between regions for significantly higher interest in one region when compared to another.

Results from regression analysis indicate that Matcha interest has consistently trended upward in all regions, with the strongest growth in the UAE and South Korea, along with steady increases in the U.S. and India. Significant differences in means were also confirmed between regions based on T-test results ($p < 0.05$). Thus, there are geographic differences in Matcha interest.

Dubai Chocolate: In testing the impact of regional factors on search interest, a similar regression-based approach was adopted for Dubai Chocolate. Linear models tested the association of time with interest, and t-tests tested whether the means of interest were significantly different between all countries.

The UAE reported the most growth and aggregate search interest in Dubai Chocolate, reflective of its development and prestige in Middle Eastern markets. Other regions showed smaller and less stable growth rates. The statistically significant t-test results suggest that recognition and cultural significance regionally impact interest in search.

Airfryer: To measure temporal trends in interest across regions, for the Air Fryer trend, regression models were used. T-tests were also conducted to evaluate mean differences in regional engagement levels for Air Fryers to determine whether the engagement levels across regions were statistically different.

After 2020, interest in Air Fryers sharply increased across all regions, illustrating global adoption during the pandemic. The U.S. and UAE displayed the strongest trends, whereas interest in India and South Korea increased more gradually. Results were statistically significant (p ¡ 0.05), demonstrating that searches varied significantly by region.

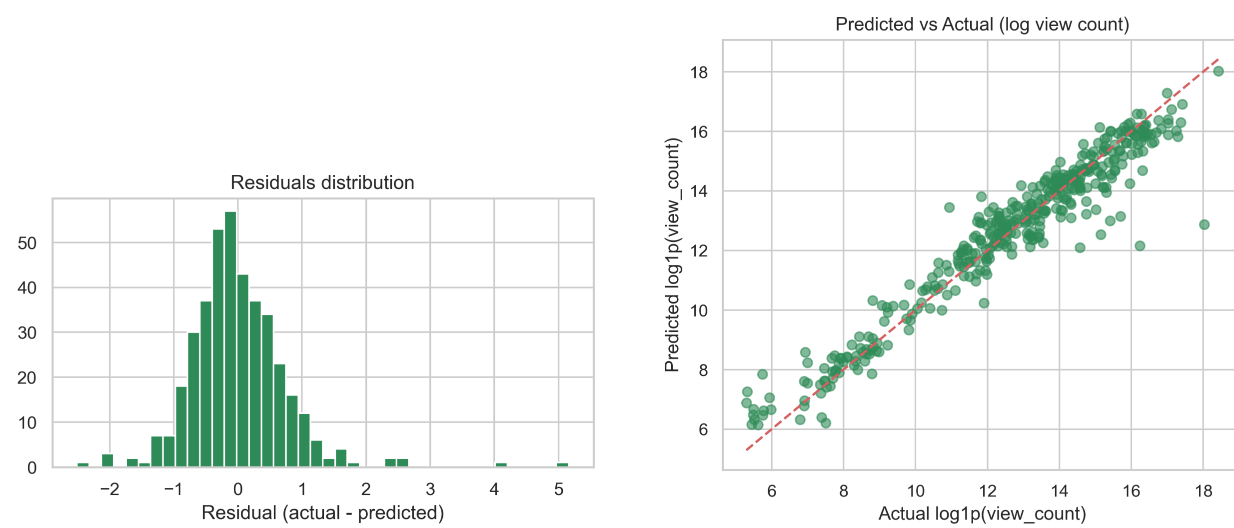
Feta Pasta: To examine regional trends in subscriber interest for Baked Feta Cheese Pasta, a linear regression and t-tests were performed. Time was used as the independent variable to capture trends in growth, and mean state values were compared to assess significance.

Interest in Baked Feta Cheese Pasta spiked sharply around 2021 across all regions, led by the U.S. and UAE. The trend was short-lived but highly pronounced, suggesting a viral effect rather than sustained growth. Regional mean differences were statistically significant, confirming varying degrees of trend adoption globally.

4.2.2 Random Forest Modeling

A Random Forest regression model was trained on the cleaned dataset of approximately 2,000 videos across the four food trends to identify which video level characteristics best predict content visibility on YouTube. The target variable view count was log transformed to correct for the extreme right skew common with online engagement metrics. The data was preprocessed and categorical variables encoded before splitting the dataset into an 80/20 training-testing split.

A baseline Random Forest Regressor was trained with 300 trees, with strong predictive performance during 5-fold cross-validation (mean $R^2 = 0.946 \pm 0.006$), indicating a stable and well generalizing model. However, Hyperparameter tuning using RandomizedSearchCV identified a slightly different configuration, but the tunings did not improve performance significantly, which suggests that the baseline model already provided near optimal predictive accuracy. On the holdout test set, the final model had an R^2 of 0.94, with an RMSE of 0.696, which reflects agreement between the predicted and observed values. As seen in Figure 13, points are closely clustered along the diagonal on the predicted versus actual plot, and the residuals are symmetrically distributed, confirming that the model is not under or over predicting.



(a) Predicted vs. Actual Log-Transformed View Counts

(b) Distribution of Model Residuals

Figure 13: Model Evaluation Diagnostics for the Random Forest Regression Model.

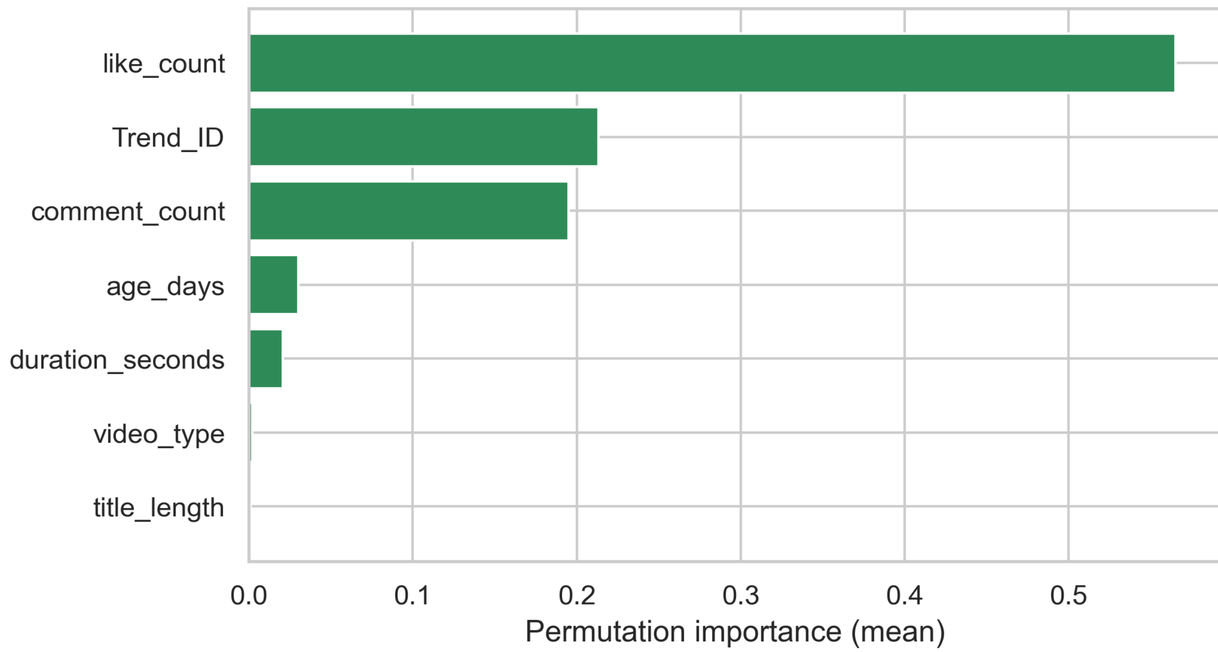


Figure 14: Permutation feature importance from the Random Forest regression model trained

Permutation feature importance (Figure 14) shows that like count was the best predictor of visibility, followed by Trend ID and comment count. These findings suggest that early interactive engagement particularly likes plays a central role in determining how far a video travels within the algorithmic ecosystem of the platform. Secondary predictors such as video age, duration, and title length contributed minimally, indicating that structural video characteristics matter less than early viewer response when determining visibility. This aligns with known dynamics of recommendations on YouTube, where high initial engagement signals algorithmic relevance and leads to greater exposure. Overall, the Random Forest model shows that behavioral signals drive YouTube visibility, not structural video attributes.

4.2.3 LDA (Topic Modeling)

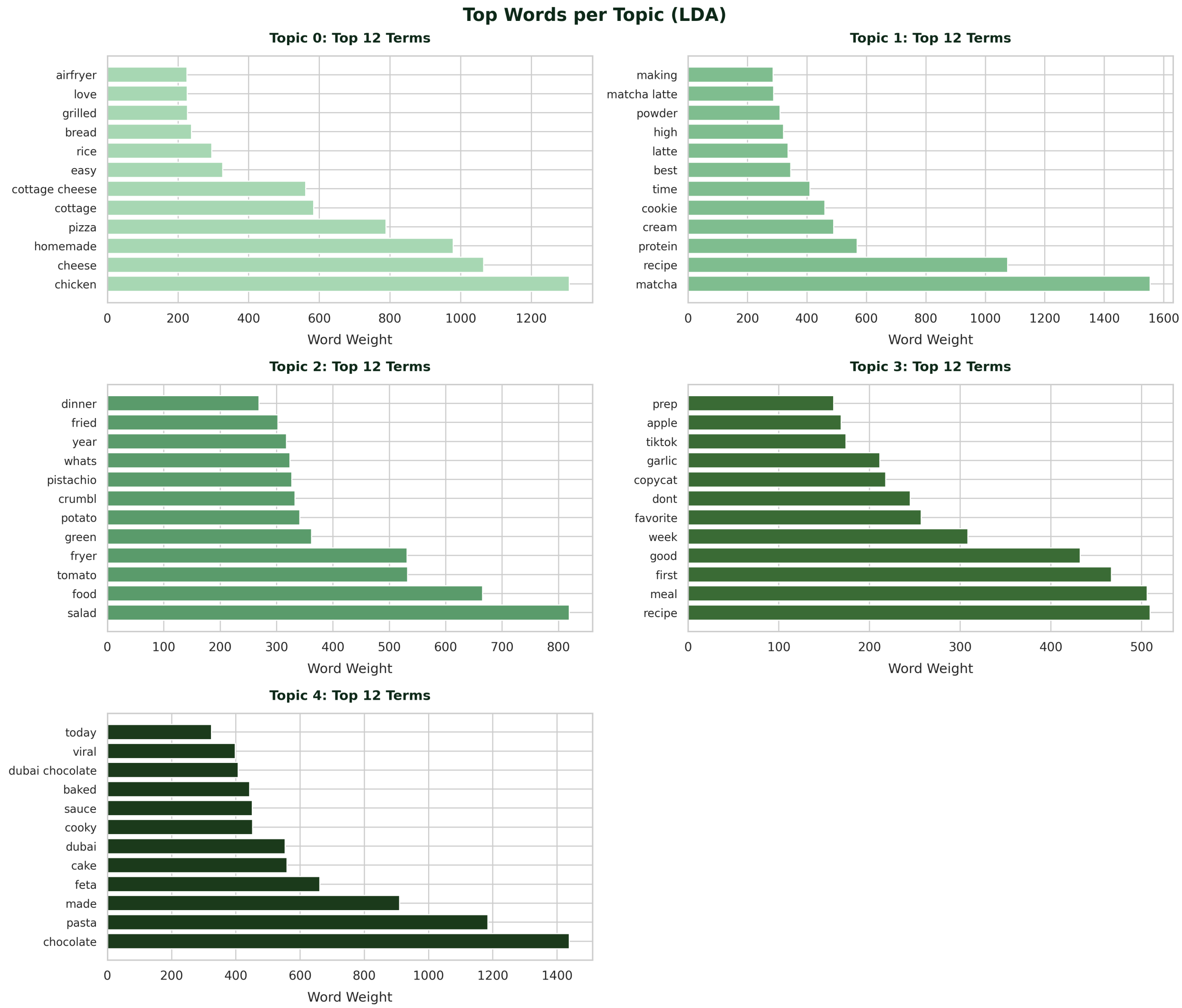


Figure 15: Top Words per Topic (LDA)

LDA uncovered the 5 distinct topics within the food trend discussions [8].

Topic 0: Describes cooking and preparation of high protein meals. Has words like chicken, cottage cheese, homemade, and airfryer.

Topic 1 References Matcha and drink preparation. Has words like matcha, latte, recipe, and protein. This may show some experimentation with ingredients.

Topic 2 Represents Air Fryer Recipes. Has words like fried, fryer, potato, pistachio, and salad. May be describing unconventional ways to use the airfryer.

Topic 3 Describes step by step recipes. Has words like recipe, meal, first, week, and copycat. Maybe depicting how users try recipes and show their results.

Topic 4 Represents viral food items including Dubai Chocolate and baked feta pasta. Has words like feta, pasta, chocolate, viral, and cake.

Overall, these topics show how Reddit users converse in food related conversations. Short term viral trends like Dubai Chocolate and Feta Pasta had more novelty related discussions, while long term trends like Matcha and Air Fryer had much more diverse conversation topics.

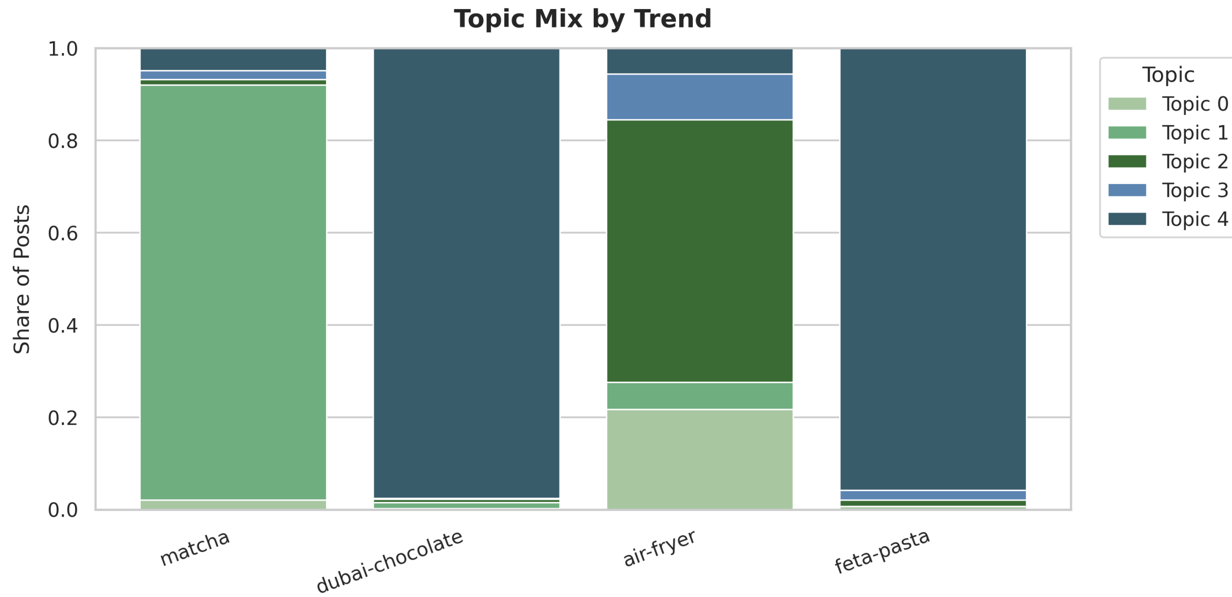


Figure 16: Topic Distribution by Trend

This stacked bar chart shows how the LDA topics are distributed across the food trends of focus.

Matcha discussions are almost entirely dominated by Topic 1, which has words like matcha, latte, and powder. This shows Matcha posts are mostly about recipes and drink preparation.

Dubai Chocolate and Feta Pasta are almost completely composed of Topic 4, which was focused on short term viral trends.

Air Fryer has the most diverse distribution as it contains discussions around cooking routines (Topic 0), experimentation (Topic 2), and step by step recipes (Topic 3). This shows Air Fryer as a sustained long term trend.

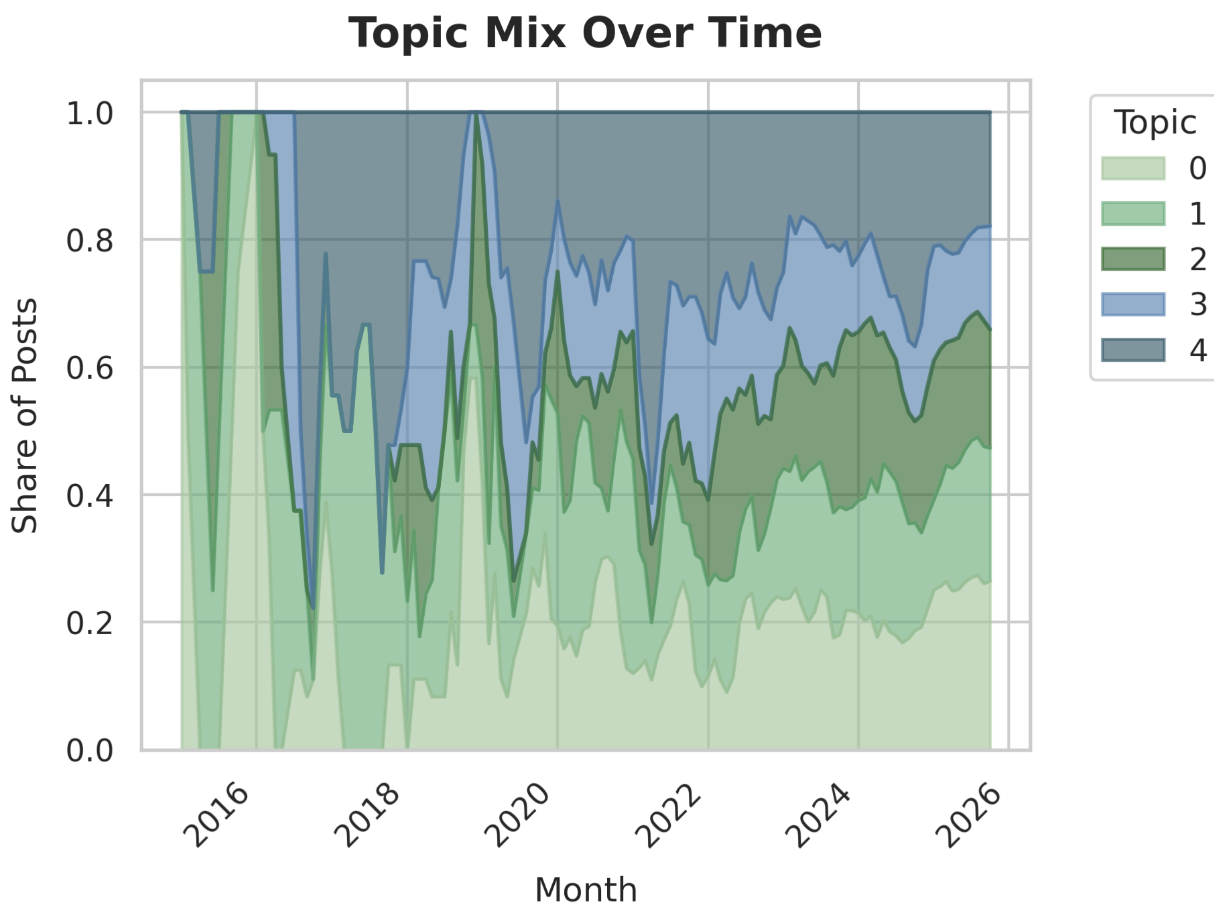


Figure 17: Evolution of Discussion Themes Over Time

This plot shows how conversation topics on Reddit [8] have changed from 2015-2025. In the earlier years, the discussion was broken across multiple topics with limited trend identity. Around 2020-2021, Topic 4 (viral food discussion) becomes dominant with the increase in TikTok inspired recipes from the pandemic. After this increase, the distribution is more balanced and Topics 1 and 2 grow at a steady rate over time. This depicts that even though viral trends can temporarily take over the conversation, long term trends persist as they become a part of everyday behavior.

5 Discussion

5.1 Overview

The focus of this research was to investigate the rise, peak, and fall of digital food trends using Google Trends, YouTube, and Reddit for the time period of 2015-2025[7, 5, 8]. Using a combination of various approaches like time series analysis, regression modelling, Random Forest, and topic modelling, we identified the relationships among curiosity, visibility, and community engagement as they relate to each of the platforms and the various types of trends. As a whole, the data indicate that food trends can be categorised into predictable lifecycle patterns that are statistically different between those long-term trends that have a lifestyle component (i.e., Matcha and Air Fryers) and those short-term viral trends (i.e., Baked Feta Pasta and Dubai Chocolate).

5.2 Answers to Research Questions

Our first research question was *What contributes to the increase and decrease of food trends on digital platforms?* Based on our analysis of searches, it seems that spikes of interest in food will normally start in

Google Trends and then gradually move on to YouTube where we observed increased video viewership. After these spikes of activity, we found that there tends to be a delay before there is any real growth in conversation about those food items posted on Reddit and that Reddit conversations tend to be more gradual over time when compared to YouTube views. Finally, we did not find that all food trends behave the same way. There are food trends that had very consistent long term engagement patterns conversation themes, while there are food trends that have had very large peaks of activity but then experience very rapid drops off of interest.

The Random Forest finding associated with answering the second question, *Can we anticipate a trend's greatest height and length of time through the early digital activity*", indicates that the amount of early engagement (likes and comments) has a strong correlation to the number of people who see a video through YouTube. Therefore, based on the way that people reacted to a trend (like, before it even happens), this shows you how far that trend is going to go. However, the number of days or weeks (or even months) a trend is trending seems to have more correlation with how well it fits into our culture than it does to how much it is viewed/engaged by audiences.

To address our third research question, *How and why do food trends originate, and spread geographically and culturally through the marketplace?*, Google Trends's geographic maps illustrate diverse geographical patterns in the way these trends are diffusing through time and regions[7]. For example, Matcha and Air Fryers are experiencing widespread geographical adoption, while Feta Pasta and Chocolate Dubai are being spread within smaller, more localized areas. Therefore, the results of our analysis indicate that food trends have been defined and developed online through both the interaction of digital platforms and the underlying social and cultural meanings associated with these products. While describing the manner in which food trends develop and escalate within the digital marketplace, the results of our study shed light on general questions surrounding how culture develops, survives, and eventually disappears in the digital landscape, as well.

Through a variety of methods, these findings illustrate how trends related to foods digitally manifest and progress across numerous platforms. To fully understand these trends, all aspects must be examined according to a broader conceptual framework, along with their methodological and social ramifications. The following sections explore these ramifications more thoroughly.

5.3 Theoretical Implications

This project further validates existing diffusion agencies in that it demonstrates that digital platforms compress the existing adoption cycle. Consistent with Rogers' (1983) *Diffusion of Innovation Theory*, the lifecycle of online food trends follows an accelerated S-curve trajectory and, as such, the roles of innovators, early adopters and majority users occur almost simultaneously[9]. Results show that search-based curiosity (Google Trends) [7] correlates with early adoption; content amplification (YouTube) [5] corresponds to uptake by early majority users; and sustained discussions (Reddit) [8] characterizes late majority users or decline of the trend.

Bourdieu's (1984) cultural capital concept offers insight into the reasons behind the differing lengths of time trends exist in social media [11]. Just because an idea/invention has been viral does not mean it will not continue as a trend after the wow factor of that trend has died down. For example, air fryers and matcha are good examples of continuing trends because they represent changing attitudes toward health. However, feta cheese pasta and dubai chocolates won't remain relevant once the fad factor wears off. The social context and everyday utility of these ideas/inventions may determine their continuing popularity beyond their initial virility.

5.4 Scientific and Methodological Implications

The methodology adopted in this research demonstrates how an integration of multiple datasets can lead to a more holistic view of the evolution of an online phenomenon from beginning to end. By triangulating three different types of datasets (i.e., search volume trends from Google [7], visual engagement on YouTube[5], and online discussion forums like Reddit [8]), the research found that patterns of curiosity, visibility, and conversation could all be captured within a consistent framework that can be replicated. Moreover, the correlation between the three platforms indicated a temporal relationship where the actions performed on the three platforms interacted with one another over time. The regression and Random Forest models revealed that early indicators of engagement predicted later visibility on the platforms.

By providing evidence of the relationship between quantitative modelling and the qualitative interpretation of data as a way to explain social-cultural processes, this research also provides a basis for future study into the evolution of digital lifecycles across a range of subjects including: fashion, sustainability, and consumer behaviour. In addition, by providing open-source tools, detailed documentation of processes, and the ability to replicate results, this research is aligned with the Standards of Practice in Computational Social Science.

5.5 Ethical and Societal Implications

While the data we analyzed is accessible to the general public, it also brings into question how representative the data is of global trends and how algorithms used for this type of research are biased culturally. The internet is not equally accessible to every region; and while many people who speak English have an advantage over non-English speakers, this means there will always be populations represented at lower rates than others, distorting the representation of "global" trends.

Understanding how these feedback loops work is important, because what happens digitally (virality) can affect what happens offline (consumption, supply chain and public health behavior). For example, viral recipes may create a surge of demand for a specific ingredient or appliance that is either a waste product or has an unsustainable production cycle. When the algorithm-driven popularity of something influences the way society thinks and behaves, it creates opportunities for better designed and marketed products and policies when the designers, marketers and policymakers recognize the social implications of their actions. Ethical research, therefore, requires a balance between providing insights through analytics and cultural sensitivity and recognizing that participation in online spaces is unequal.

5.6 Limitations

The multi-platform nature of this study provides a broad perspective on digital food trends, however, several limitations exist. First, this study is reliant upon data from YouTube, Reddit, and Google Trends and each platform provides only a partial picture of the digital space [7, 5, 8]. Each platform uses its own proprietary ranking and recommendation algorithms and the patterns identified are influenced by both the level of user interest and the degree to which the algorithms being utilized have shaped user experience.

Second, the focus of this study was on four food trends and a small number of geographic locations. We selected these case studies to provide examples of various ways in which trends diffuse, but they do not encompass all the trends that exist in the digital space, especially those that emerge within specific communities or regions that are not well represented on these platforms.

Third, our use of VADER and LDA to analyze sentiment and topics respectively oversimplify language and do not allow us to capture the nuances of context, humor and visual components that contribute to much of the online discourse around food trends. A more in-depth qualitative or multimodal approach would likely provide a more comprehensive understanding of how users perceive and interact with these trends.

Finally, the timing of activity across platforms does not align perfectly. Each site reports data at different intervals and with different levels of detail, which prevents us from determining the exact relationships between search interest, content creation, and online discussion. As a result, we interpret our findings as general patterns that evolve together rather than precise causal sequences.

6 Conclusion

This study shows how digital food trends are born, spread, and evolve across Google Trends [7], YouTube [5], and Reddit [8]. By exploring search patterns, creator participation, and community engagement, we learn how different cross-platform trends operate over time with Matcha and Air Fryer emerging as long lasting trends and Baked Feta Pasta and Dubai Chocolate as fads. Thus, established trends evolve from increasingly favored attention and platform benefits and cultural significance assigned to them by their respective online users over time.

This method more accurately assesses the nature of cross-platform relevance since Google's Trends is where the initial search is found, YouTube is where the general participatory creation exists and the Reddit posting provides the context of cultural adoption and detriment [7, 5, 8]. Should researchers attempt to analyze each platform in a vacuum, they risk losing context to how attentions transition into different facets of the digital space. These results show a natural progression across all 3 focus areas, providing a more accurate picture of how trends and fads evolve over time.

While this study has limitations, it adds information to the implications of digital food trends that anticipated behavioral patterns are driven by forces from an unknown source. Thus, this study not only supports a hypothesized model for future studies when applied to digital food trends, but also provides cultural insights to highlight why these trends occur, fade or flourish.

Ultimately, this is an applicable framework through which to assess how digital trends emerge and evolve as composites of online behavior, platform-driven algorithms and sociocultural meaning which fuels interest or leads to decline.

7 Code Appendix

All code used in this project is available in the public GitHub repository below:

GitHub Repository:
https://github.com/kzumreen/FoodTrendsPrediction
The repository is organized into the following main components:

- `EDA_Plots/` — Visualizations for YouTube trend dynamics and cross-platform comparisons plots.
- **Jupyter Notebooks:**
  - `GooglePytrends.ipynb` — Retrieval, cleaning, and exploratory analysis of Google Trends data.
  - `pytrends_eda_global.ipynb` — Country-level and global time-series visualizations.
  - `pytrends_eda_notebook_countries_with_Air_Fryer.ipynb` — Regional diffusion analysis across four countries.
  - `Trend_Country_Mapjpynb.ipynb` — Construction of choropleth geographic maps.
  - `Reddit_Access.ipynb` — Reddit API (PRAW) data collection.
  - `RedditInferentialStats.ipynb` — Basic regression and statistical analysis on Reddit data.
  - `RedditLDA.ipynb` — LDA topic modeling workflow (preprocessing, dictionary creation, topic extraction).
  - `youtubedata.ipynb` — YouTube API scraping, data cleaning, feature engineering, and dataset export.
  - `Cross_analysis.ipynb` — Cross-platform comparisons of search interest, content creation, and community discussion.
  - `Inferential_Statistics.ipynb` — Regression and statistical testing for Google Trends interest by region.
- **Python Scripts:**
  - `youtubedata.py` — Automated YouTube API retrieval for all four food trends.
- **Datasets:**
  - `youtube_data.csv` — Cleaned dataset of YouTube video-level metadata.
  - `reddit_food_trends_data.csv` — Cleaned Reddit dataset with posts, metadata, and sentiment fields.
  - `py_trends.csv` — Weekly Google Trends dataset (global + regional).

For complete reproducibility, all scripts, notebooks, and intermediate data files are available in the GitHub repository.

References

[1] Sofiane Abbar, Yelena Mejova, and Ingmar Weber. You tweet what you eat: Studying food consumption through twitter. *arXiv preprint arXiv:1412.4361*, 2015.

[2] Sean B. Cash, Saleem Alhabash, Gabriela Fretes, and Mengyan Ma. Food, beverages and social media: trends and tools for economic research. In *A Modern Guide to Food Economics*. Edward Elgar Publishing, 2022.

[3] Larissa S. Drescher, Carola Grebitus, and Jutta Roosen. Exploring food consumption trends on twitter with social media analytics: The example of #veganuary. *Journal of Consumer Affairs*, 2023.

[4] Daniel Fried, Mihai Surdeanu, Stephen Kobourov, Melanie Hingle, and Dane Bell. Analyzing the language of food on social media. *arXiv preprint arXiv:1409.2195*, 2014.

[5] Google. Youtube data api v3, 2023.

[6] Devashish Khulbe and Manu Pathak. Modeling food popularity dependencies using social media data. *arXiv preprint arXiv:1906.12331*, 2019.

[7] Patat, Damien. pytrends: Unofficial api for google trends, 2023. Accessed: [Date of your access].

[8] Reddit, Inc. and PRAW Authors. Praw: The python reddit api wrapper, 2023.

[9] Everett M. Rogers. *Diffusion of Innovations*. The Free Press, New York, 3rd edition, 1983.

[10] Georg Simmel. Fashion. *The American Journal of Sociology*, 62(6), 1957.

[11] David Swartz. *Culture and Power: The Sociology of Pierre Bourdieu*. University of Chicago Press, Chicago and London, 1997.