The Most Effective Brand of Super Bowl LVIII on Various Perspectives

2024 Game Day Analytics Challenge

Undergrade 118

Jimin Go | TackKyoon Woo | Taehyun Noh

21 February 2024

TABLE OF CONTENTS

Abstract

This paper aims to analyze which brand had the most effective Super Bowl advertisement and to provide valuable insights into large-scale events like the Super Bowl by utilizing Twitter data. Companies consider advertising their product or services during the Super Bowl as an unparalleled opportunity. And, these advertisements often spark heated reactions on social media platforms such as twitter. Therefore, companies are willing to pay a significant amount of money, $7 million for a 30-second spot, to advertise during the Super Bowl. To evaluate how worthwhile the investments in these companies' Super Bowl advertisements were, we employed various techniques. Based on analysis of tweet data, we applied multi-criteria decision-making techniques, and additionally conducted social network analysis. This study proposes novel techniques for analyzing advertising effectiveness for companies and is expected to aid in understanding the characteristics of Twitter users.

Keywords

Time Based Analysis, Cost Analysis, Sentiment Analysis, AHP, TOPSIS,

Social Network Analysis

1. Introduction

1.1 Why are Super Bowl ads so popular?

The Super Bowl is one of the most-watched sporting events globally, with average 123.7 million of viewers tuning in [1]. Also, it provides an opportunity for many people to enjoy sports, advertisements, music and more, including halftime show. The level of interest in Super Bowl is incredibly high, as evidenced by the fact that our of the top 30 most watched TV broadcasts in US history, 22 of them are Super Bowls. It means the Super Bowl presents an unparalleled chance for companies seeking to advertise their products or services. Although Super Bowl ads cost a pretty penny, their impact on brands' visibility can be massive, as they are a favorite part of the game for almost 30% of viewers [1]. For example, Apple sold $3.5 million worth of computers just after it ran and $100 million within 100 days after the Super Bowl.

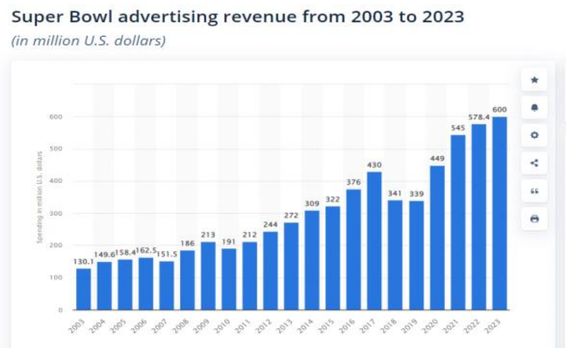

Fig1. Super Bowl advertising revenue from 2003 to 2023

According to a 2022 study by Kantar [2], an average Super Bowl ad is over 20 more effective than a regular TV ad. So, most companies have been willing to invest millions of dollars for mere seconds of screen time to display their advertisements during the event. To make this investment worthwhile, each company created a high-quality and innovative advertisements through their own marketing tactics. This trend has continued into this year as well. Given the opportunity to reach such a vast audience, companies spent a 30 second Super Bowl ad spot cost $7 million, and that is not counting in how much was spent on developing it. While much more expensive to work with, the star power of beloved actors, musicians, and athletes is invaluable for creating connections between the audience and the brand, some companies featured popular celebrities in their advertisements. Also, some other companies incorporated elements in their advertisements aimed at encouraging viewer engagement and social sharing, leveraging the fact that people share their Super Bowl viewing experience with others. Besides those strategies, various strategies were employed to leave a lasting impression on people.

1.2 Purpose of the study

It is not unsurprising that Super Bowl ads are often known for their quality and approaches. The question remains, however, which strategies are truly effective? Which brand's strategy has left the deepest impression on current or potential customers?

To clearly answer above questions, this paper provides a thorough and extensive analysis of the most recent Super Bowl commercial advertisements that was held on February 11, 2024. By analyzing Twitter data for all advertisements aired during the Super Bowl game, particularly, the paper determines which brand leveraged the Super Bowl game to deliver the most effective advertising. In the paper, the methodology used to analyze data pertaining to Super Bowl advertising is outlined in detail. Summarizing all these analyzations, we answered to the question of "what makes an ad successful?". The strategies that a brand wishing to increase awareness should adopt differ from those of a brand that already has sufficient awareness. Based on this evaluation result, the paper also provides practical strategies that tailored to the purpose of companies advertising.

This comprehensive analysis undertaken in the paper can provide insights into effective advertising within the context of large-scale events like the Super Bowl game. Additionally, they can better understand the psychology of viewers watching the Super Bowl and tailor their strategies accordingly. Furthermore, it is expected to serve as valuable material for companies and marketers seeking a better understanding of the somewhat unique advertising environment of the Super Bowl, thereby assisting companies in making informed decisions when planning future advertisements.

1.3 Paper structure

The remainder of this paper is structured as follows. The following, Section 2 presents how we handle given data. Section 3 describes the various methodologies used to analyze data. Based on these analytical findings, Section 4 presents evaluation results conducted using the AHP(Analytic Hierarchy Process) and TOPSIS(Technique for Order Preference by Similarity to Ideal Solution) model as criteria. Also discusses companies strategies synthesized all these analyses. Section 5 examines the given data from a network perspective. Section 6 closes with a conclusion.

2. Data Preprocessing

2.1. Dataset overview

During analysis, we used two datasets provided to us: count_three_days.csv and 'Final keywords_2024.csv'. The 'count_three_days.csv' dataset consists of three columns: 'date', 'brandname', and 'count', representing the tweet counts for each brand for the two days prior to the Super Bowl and the day of the Super Bowl. And the 'Final keywords_2024.csv' dataset consists of 34 columns and contains tweets written on the day of Super Bowl along with their contents.

2.2 Handling missing values

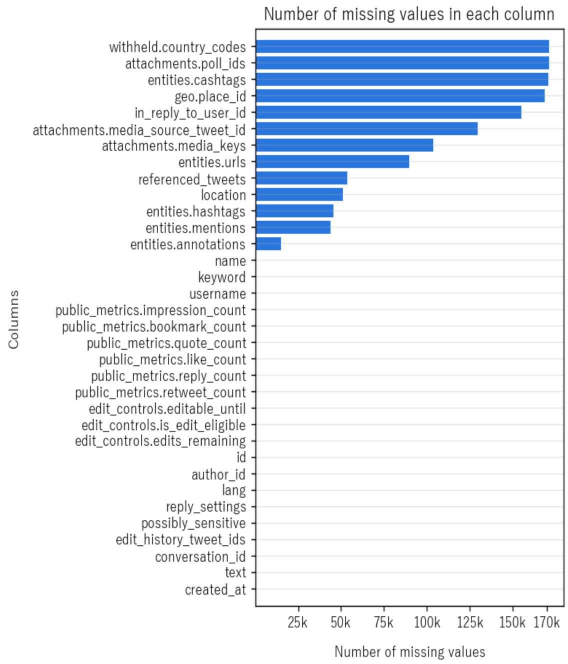

Fig 2. Number of missing values in each column

The first step we conducted was handling missing values. Upon reviewing the entire dataset, there were some columns with numerous missing values. Fig 2 shows the number of missing values in each column. From the figure, it can be observed that the top 7 features have more than 50% missing values out of total 171,454 entries. These columns were removed from the dataset due to the insufficient quantity of data, which could potentially lead to biased results.

2.3 Feature selection and extraction

We adopted only a subset of columns that we deemed important from the entire columns of dataset for our analysis: 'created_at', 'text', 'username', and 'entities.hashtags'. To conduct analysis over time, we utilized the 'created_at' column, which indicates the time when the tweets were posted. Additionally, to analyze whether the tweets contain contents related to specific brands, we used the 'text' column containing the tweet content and the 'entities.hashtags' column containing hashtags. Finally, to analyze the network between users and brands, adopted 'username' column, which contains the usernames of those who posted the tweets. We also considered using the 'lang' column indicating the language used in the tweet to determine the global influence of the brands. However, since the vast majority of the values were in English and other languages were very few in number, we decided not to utilize this column. Furthermore, upon inspecting the 'keyword' column, it was noted that there was a significant mismatch between the keyword and both text and hashtags of the tweets.

RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @porktendencia: ??Bob Esponja??:	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @porktendencia: ??Bob Esponja??:	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @CultureCrave: Full version of SpongeBob performing ??Sweet Victory?? at the #SuperBowl??????	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
Apparently this is the Flashdance #SuperBowl.	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @CultureCrave: Full version of SpongeBob performing ??Sweet Victory?? at the #SuperBowl??????	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy
RT @DiscussingFilm: SpongeBob performs the full version of ??Sweet Victory?? at the #SuperBowl https://t.co/EZqpPuT6LC	NERDS Gummy

Fig 3. Example of irrelevance between Keyword and contents

In Fig 3, it can be observed that the keyword is 'NERDS Gummy', however, there is no relevant content found within the text. While the words 'sweet' and 'Gummy' may appear to match, in context, 'Sweet' is used to describe happiness of victory, making the two words unrelated. Similar cases were observed with keywords such as 'Progressive', 'Uber Eats', and 'Bud Light'. Consequently, it was decided to exclude these brands from the 'keyword' column and directly search for tweets related to these brands. Additionally, we examined the columns corresponding to 'public_metrics'. For data corresponding to these columns, we observed that when a tweet is retweeted, it inherits information from the original tweet being retweeted. Therefore, we concluded that utilizing this data for analysis might lead to excessive aggregation of data of biased result. So, we ultimately decided to use the four columns mentioned earlier: 'created_at', 'text', 'username', and 'entities.hashtags' for our analysis.

2.4 Data transformation

We confirmed that the 'created_at' column was represented in Coordinated Universal Time (UTC), and to avoid confusion, we converted it to Daylight Saving Time (DST). Additionally, to facilitate analysis, we converted the time represented from seconds to minute. And the remaining columns were deemed suitable for analysis in their current form. To conduct brand-level analysis, collected brand names using the 'count_three_days.csv' data. The 'brandname' column in the dataset mostly consisted of brand names, but it also included entries that were not brand names. For example, movies such as 'Wicked', 'Monkey man', and 'If' were included, which are from 'Universal'. And since 'Universal' itself was also listed as a separated brand, we consolidated these entries under the brand name 'Universal'. We performed the same process for other data with similar characteristics and ultimately obtained a total 96 brand names.

3. Data analysis

3.1 Tweet analysis over time

To understand how people's reactions to brands evolve during the Super Bowl game, we analyzed tweets written over time for each brand. We expected that if the advertisement was effective, there would be an increase in the number of tweets related to the brand of that advertisement during the time when the advertisement was aired. Before conducting the analysis over time, we first extracted the top 10 brands that accounted for a significant proportion of the total tweets, believing that these brands would demonstrate notable changes over time. Each tweet associated with a brand was considered relevant if the brand's name appeared in the 'text', 'entities.hashtags', or 'keyword' column of the tweet. Tweets meeting these criteria were counted as relevant to the respective brand.

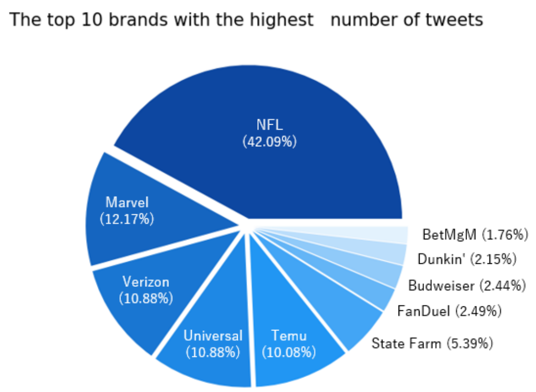

Fig 4. The top 10 brands with the highest number of tweets

Fig 4 represents the top 10 brands holding the highest number of tweets. This graph shows that NFL stands out, occupying a disproportionately high proportion of the total tweets and the rest of the brands are distributed relatively evenly in proportion. Next, we analyzed how the number of tweets changes over time for each of the top 10 acquired brands, in relation to the total tweets. The method used to count tweets remained consistent with the approach used to extract the top 10 tweets, which involved checking for the presence of the brand name in the 'text', 'entities.hashtags', or 'keyword' columns. However, the difference lies in segmenting these tweets into minute intervals, reflecting the progression of time. The result of this analysis is represented in the Fig 5. It can be observed that tweets related NFL were actively posted from the start of the game at 16:40 until 17:00. At this point, tweets related to NFL dominated the majority of the overall tweets, while tweets related to other brands, including State Farm, were also actively posted. This indicates that the majority of Twitter users mentioned NFL and other brands simultaneously. Thus, this can be interpreted as the Super Bowl and the brands being interconnected through the link of Super Bowl advertisements. Furthermore, tweet related to State Farm were particularly abundant in the early stages of the game. Tweets about State Farm are not prominently visible after this point; however, the attention gained during this time period has made it the sixth most tweeted brand among all brands. The reason State Farm received significant attention during this time is believed to be because State Farm's advertisement was aired during the

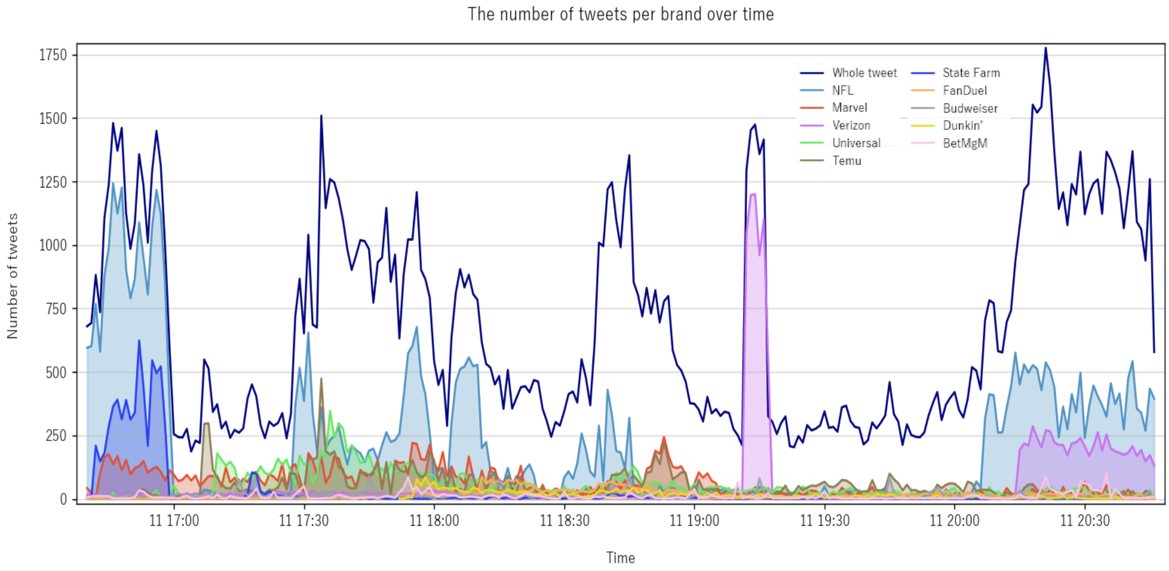

Fig 5. The number of tweets over time for the top 10 brands

first quarter. It is judged that State Farm's cleverly crafted advertisement was captivating enough to grab viewer's attention. Another notable brand is Verizon. The number of tweets related to Verizon is observed to sharply increase between 17:00 and 17:30. This characteristic is similar to that of State Farm, but Verizon shows an even sharper increase in tweets within a shorter timeframe. Verizon is the only brand that generated such a large number of tweets within a short period of time, approximately 10 minutes, among all the brands. This time corresponds to the third quarter of the game, during which Verizon's advertisement was aired. Therefore, it is deemed that Verizon's advertisement sparked a burst of interest among viewers. Furthermore, at this point, there are also a significant increase in the total number of tweets, and it can be attributed to Verizon. Tweets related to Verizon were also actively posted from the latter part of game through overtime. However, we believe the time when advertising impact was most keenly felt was between 17:00 and 17:30, as this period had the highest density of posted tweets. Unlike State Farm and Verizon, Marvel did not generate a large number of tweets within a short period. However, tweets related to Marvel were consistently posted throughout the game, making Marvel the second most tweeted brand among all brands. This fact highlights the importance, from a marketing perspective, of instilling an impression that can endure as long as obtaining explosive interest through advertising. Then, is capturing people's attention the only important factor in marketing?

3.2 Cost analysis per tweet

In advertising, it is important not only how well captured people's attention but also how much investment has been made. In the previous section, we assessed how many tweets were generated per brand to gauge the level of interest captured by each. Based on this data, we analyzed how much investment was made per brand to generate that number of tweets. Given that the average cost of a Super Bowl ad was $7 million for 30 seconds, we derived that approximately $0.23 million was

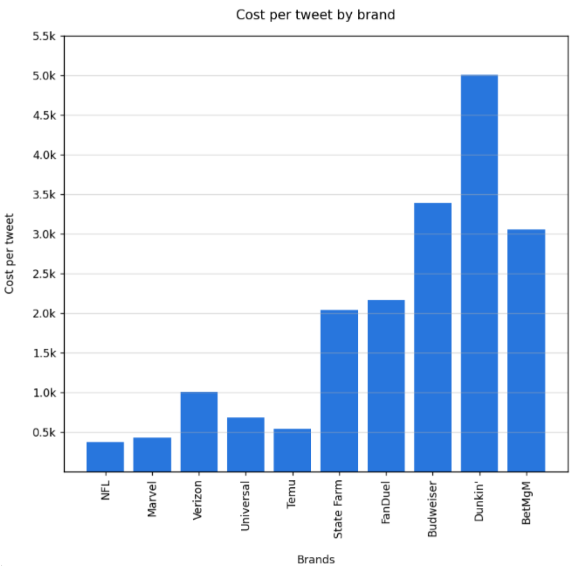

Fig 6. Cost per tweet by brand

spent per second by dividing $7 million by 30 seconds. Furthermore, when calculating the advertising duration for each brand, one of two methods was selected. First one is involved summing the durations of multiple ads aired by each brand, while the second one entailed calculating the average duration of advertisement for each brand. Due to the observation that a single ad was exposed multiple times during the Super Bowl, we opted for the second method and calculated the advertising duration for each brand accordingly. Using this calculation, we then multiplied the duration of each brand's ad to ascertain the advertising expenditure for each brand. And we divided the advertising cost for each brand the number of tweets per brand to calculate the amount of money spent per tweet for each brand. In Fig 6, displays the analysis results of the cost per tweet for the top 10 brands with the highest number of tweets. When comparing the results from Fig 4, it is evident that despite generating a larger number of tweets, certain brands in Fig 5 have incurred higher costs compared to brands that generated fewer tweets. Verizon and Universal exemplify this case. Verizon aired a single 61-second ad, while Universal aired two ads-one lasting 60 seconds and the other 30 seconds-resulting in an average ad duration of 45 seconds. With a total difference of 15 seconds in the duration of their advertisement, and given that Super Bowl advertisements require a higher cost per second, Universal's cost per tweet was consequently lower than it of Verizon's, as the number of tweets per brand did not vary significantly. In situations like this, it can be challenging to determine which brand's advertisement was more effective - the one that generated slightly more tweets through advertising but with higher investment costs, or the one that generated slightly fewer tweets but with lower investment costs. Therefore, we also want to analyze other factors that can help us asses the effectiveness of the brand's advertisement.

3.3 Sentiment analysis

The fact that there are many tweets on Twitter shows interest in the advertisement, but the number of tweets and cost per tweet are not absolute. A lot of negative tweets may mean that the advertisement was that bad. It is also very important to analyze the positive and negative expressions in the tweets. So, in this study, an emotional analysis was conducted to evaluate whether the response of Tweets to the advertisement was positive or negative.

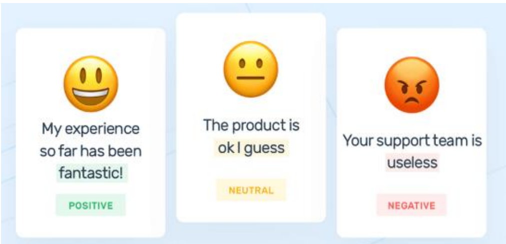

Fig 7. Example of emotional analysis [3]

Emotional analysis is the process of understanding and analyzing the emotional content of the text data to identify the positive, negative, and neutral emotions contained in each sentiment or document. VADER was used for emotional analysis. VADER is an open-source library for emotional analysis in text data and was developed in the Python language. VADER calculates the emotional score by considering the positivity, negativity, and neutrality of each word. In particular, VADER can recognize bullish expressions and is suitable for text analysis in the form of sentences such as tweets. It has a scale between -1 and 1, and the closer to -1, the more negative, and the closer to 1, the more positive it is.

In this study, it was considered strong negative if the emotional score was -1 to -0.7, negative if -0.7 to -0.3, neutral if -0.3 to 0.3, positive if 0.3 to 0.7, and strong positive if 0.7 to 1. In addition, VADER was used except for certain meaningless words to increase the efficiency of emotional analysis. The following are excluded words.

@
RT
rt
https
Emoticons
_(underbar)

Fig 8. The excluded words

By integrating the sentiment analysis of company-specific tweets, the company's number of tweets and the response of the tweets were examined, and the results of the sentiment analysis for a specific company in the early and late stages of the game were compared to evaluate how the advertising effect in the middle of the game caused the company's response.

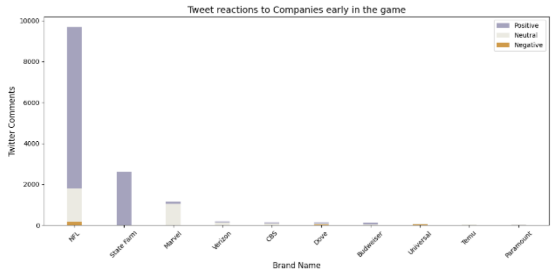

Fig 9. Tweet reactions to companies before in the game

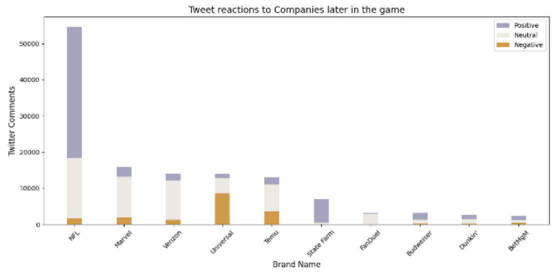

Fig 10. Tweet reactions to companies later in the game

At the beginning of the game, the total number of tweets was 14,637, and at the end of the game, the number of tweets was 169,460, more than 10 times the difference in the number of tweets. This can be seen as evidence that the effectiveness of the advertisement is certain. There have also been changes for the top 10 brands in the tweets section. In the case of the NFL, it ran an overwhelming first place. However, it is very difficult to guarantee the effectiveness of the advertisement because Super Bowl games are run by the NFL, making it difficult to tell whether people have seen the advertisement or just commented on it because it meant the game. FanDuel was ranked 35th early in the game and then 7th late in the game. This indicates that the advertisement was effective in any direction. In terms of emotion, State Farm

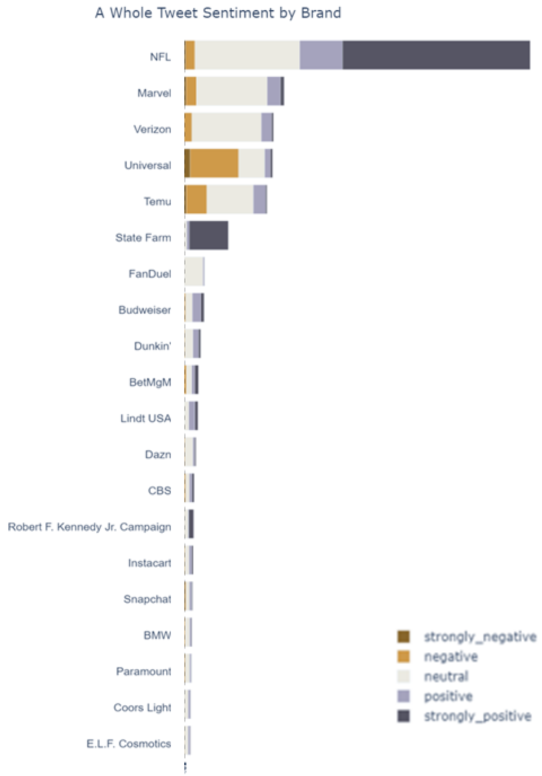

Fig 11. A whole tweet sentiment by brand

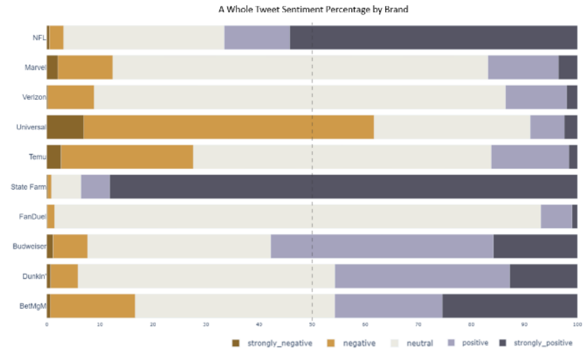

Fig 12. A whole tweet sentiment percentage by brand (top10)

maintained its 10th place in the number of positive tweets before and after the game, maintaining more than 90% of the number of positive tweets. This could mean that State Farm's usual brand image was originally good, and advertisements were also effective. On the contrary, in the case of Universal, the number of tweets in the second half of the game remained in fourth place, making the advertisement effect look good, but negative tweets accounted for more than half, which was higher than any other company in the top 10. It was also confirmed that the number of positive tweets was not even large. This is an example of the proper use of emotional analysis, and it can be seen that the advertisement was not good or the perception of the company was very bad.

Through emotional analysis of tweets, we once again confirmed that the total amount of tweets cannot necessarily determine whether the advertisement is well made or has an effect. Our team was convinced that it was appropriate to use the number of tweets, the ratio of negative tweets, and the ratio of positive tweets as the criteria for determining which company made the most effective advertisement later in the study.

4. Evaluating Results

4.1 MCDM

As a result of analyzing the tweet about the current Super Bowl commercial, it was confirmed that there is not only one measure in evaluating advertisements, but various criteria and perspectives exist. That is why our team decided that we should use a multi-criteria decision model. Multi-criteria decision-making (MCDM) refers to the process of making an optimal decision by considering several criteria or factors. This can better reflect the complexity of the real world than by judging on a single basis. The goal is to derive the best choice by considering various factors in a given situation. Among the many techniques, it was decided that it would be appropriate to use a mixture of AHP and TOPSIS techniques. Weights will be determined through the AHP model, and the TOPSIS model will evaluate how close the alternatives are to the given criteria to find the most ideal alternative.

4.2 AHP

The AHP (Analytic Hierarchy Process) is a method of evaluating the relative importance between each criterion or alternative by decomposing complex decision-making problems into hierarchical structures. In this study, AHP used to calculate weights for each criterion is largely composed of four steps.

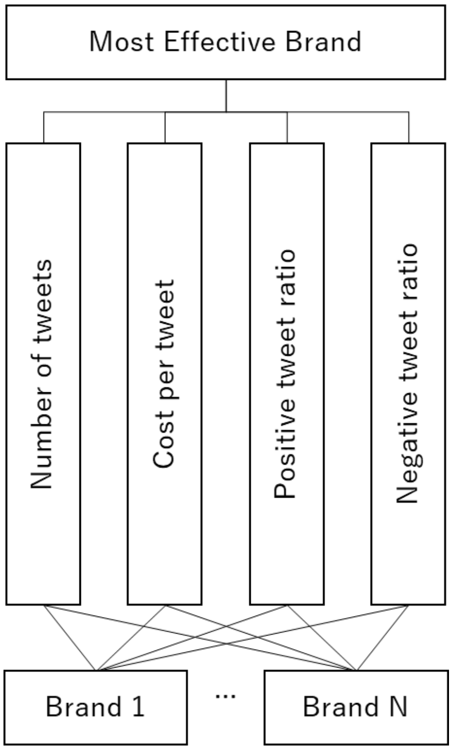

Fig 13. AHP Model architecture

a. Define the problem and create a hierarchy: First, clearly define the purpose of the decision and identify all the criteria and alternatives involved. Then list these criteria and alternatives in a hierarchical structure. Typically, the goal is at the top and the criteria are below it, followed by alternatives. In this study, as a result of analyzing the previous data, the number of Tweets (C1), advertising costs per comment (C2), the ratio of positive tweets(C3), and ratio of negative tweets (C4) were determined as criteria for evaluating advertisements.

b. Relative Importance Assessment: The step of evaluating the relative importance between each criterion and alternative. Decision makers evaluate the relative importance between each factor through a pairwise comparison. For example, compare the relative importance between two criteria on a scale of 1 to 5, and assign relative weights based on the results of the comparison. The pairwise comparison of the criteria selected in this study is as follows.

	C1	C2	C3	C4
C1	1	3	2	4
C2	0.33	1	3	4
C3	0.5	0.33	1	2
C4	0.25	0.25	0.5	1

Table 1. Pair Comparison Table

c. Consistency Review: When evaluating relative importance, decision makers must check the consistency of each pair of comparisons. This is done by performing the same comparisons in different ways to ensure consistency. If there is a lack of consistency, it is necessary to review the comparisons again to improve consistency. The formula for evaluating the consistency of the pairwise comparison is as follows.

$$CR = RI / CI$$

- CR: Consistency Review

- RI: Random Index These values are predefined according to the dimension of the matrix and are used in AHP for consistency assessment. In this study, the number of criteria is 4, so n=4, where the RI value is 0.9.
- CI: = $\frac{\lambda_{\max}-n}{n-1}$ Here, $\lambda_{\max}$ represents the maximum eigenvalue.

If the CR value is less than 0.1, it is judged to be consistent. The steps to obtain $\lambda_{\max}$ in this study are as follows:

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 0.33 & 1 & 3 & 4 \\ 0.5 & 0.33 & 1 & 2 \\ 0.25 & 0.25 & 0.5 & 1 \end{bmatrix}$$

- Form a matrix model of the pairwise comparison table.

$$\begin{bmatrix} 0.481 & 0.655 & 0.307 & 0.364 \\ 0.159 & 0.218 & 0.461 & 0.364 \\ 0.240 & 0.072 & 0.154 & 0.182 \\ 0.120 & 0.055 & 0.077 & 0.091 \end{bmatrix}$$

- Each element is normalized by dividing it by the sum of the columns to which the element belongs.

W_1= (0.481+0.655+0.307+0.364)/4=0.452

W_2= (0.159+0.218+0.461+0.364)/4=0.301

W_3= (0.240+0.072+0.154+0.182)/4=0.162

W_4= (0.120+0.055+0.077+0.091)/4=0.086

- The values in each column are summed and divided by 4 to obtain a normalized criterion weight.

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 0.33 & 1 & 3 & 4 \\ 0.5 & 0.33 & 1 & 2 \\ 0.25 & 0.25 & 0.5 & 1 \end{bmatrix} \begin{bmatrix} 0.452 \\ 0.401 \\ 0.162 \\ 0.086 \end{bmatrix} = \begin{bmatrix} 2.020 \\ 1.278 \\ 0.658 \\ 0.355 \end{bmatrix}$$

- Find the matrix multiplied by the previously obtained pairwise comparison table matrix and the weight matrix for each reference

$$\lambda_{\max} = \frac{\left[\frac{2.020}{0.452}+\frac{1.278}{0.401}+\frac{0.658}{0.162}+\frac{0.355}{0.086}\right]}{4} = 4.232$$

- Find the average of the sum of the values obtained by dividing the newly obtained matrix by the weight matrix.

Finally, we get the CR value:

$$CR = \frac{4.232-4}{4-1} = 0.077$$

Since the CR value is less than 0.1, the pairwise comparison performed earlier can be judged to be consistent. Therefore, since the weights for the four criteria are stable, we will use weights to draw conclusions using the TOPSIS technique.

4.3 TOPSIS

TOPSIS (The Technique for Order Preference by Simplicity to Ideal Solution) is a method of making the most appropriate choice between a given alternative and is determined by comparing the similarity between an ideal solution and a real solution. The purpose is to select an alternative that meets the given criteria and is close to the ideal solution.

	C1	C2	C3	C4
AllState	20.00	1213333.33	0.10	0.30
Aws	151.00	47902.87	0.03	0.05
⋮	⋮	⋮	⋮	⋮
BetMgM	2292.00	3054.10	0.17	0.46
BMW	1275.00	10980.39	0.15	0.29

Table 2. Example of a numerical table by each company's criteria

a. Standardization: To match the scale of each element, standardize the values to compare the performance of alternatives for each criterion. In this study, we used vector normalization techniques to standardize data. Here's how to standardize vectors: When the vector

$$\mathrm{v} = [v_1 + v_2 + \cdots + v_n]$$

is given, the length of the vector is obtained using a corresponding equation.

$$\|\mathrm{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

By dividing each element of vector v by the length of the corresponding vector, a normalized vector may be obtained.

$$\mathrm{v}_{normalized} = \frac{\mathrm{v}}{\|\mathrm{v}\|} = \left[\frac{v_1}{\|\mathrm{v}\|} + \frac{v_2}{\|\mathrm{v}\|} + \cdots + \frac{v_n}{\|\mathrm{v}\|}\right]$$

b. Weighted Normalization: Normalize each criterion by weighting it according to its importance. In this process, the weight is the value obtained through the pairwise comparison of the AHP technique. Therefore, with the same standardization formula, you can standardize the tables by the companies mentioned above and multiply the weights by criteria to get a new table. This step allows for a fair comparison of alternatives when each criterion has a different importance.

c. Calculation of Ideal Solution and Non-ideal Solution: The ideal solution is the maximum value for each criterion, the maximum value for the criteria to be maximized and the minimum value for the criteria to be minimized. Non-ideal solutions are the minimum values for each criterion, the minimum values for the criteria to be maximized and the maximum values for the criteria to be minimized. In this study, the ideal solution is the number of Twitter and the ratio of positive tweets, and the negative solution is the ratio of cost per Twitter and negative tweets. Find the maximum and minimum values in each column. The maximum value is A+ and the minimum value is A-

Brand	C1	C2	C3	C4
AllState	0.000	0.021	0.009	0.008
Aws	0.001	0.001	0.003	0.001
⋮	⋮	⋮	⋮	⋮
BetMgM	0.017	0.000	0.014	0.012
BMW	0.009	0.000	0.013	0.008
A+	0.395	0.000	0.087	0.000
A-	0.000	0.298	0.000	0.024

Table 3. Example of a standardized and weighted numerical table by each company's criteria

d. Calculate similarity and distance: Calculate the similarity between ideal and non-ideal solutions. In our study, we use Euclidean distance. The shorter the distance from the ideal solution, the closer the alternative is to the ideal solution. Use the following formula:

$$Di^{+} = \sqrt{\sum\left(v_{ij} - v_j^{+}\right)^2}$$

$$Di^{-} = \sqrt{\sum\left(v_{ij} - v_j^{-}\right)^2}$$

e. Determine alternative's ranking: Rank the alternatives according to their distance from the ideal solution. The one with the shortest distance is considered the best option.

$$C = \frac{Di^{+}}{Di^{+} + Di^{-}}$$

Brand	D+	D-	C	Rank
AllState	0.403	0.278	0.408	76
Aws	0.403	0.298	0.425	54
⋮	⋮	⋮	⋮	⋮
BetMgM	0.386	0.299	0.436	10
BMW	0.393	0.298	0.431	14

Table 4. Example of a standardized and weighted numerical table by each company's criteria

4.4 Result

The results of setting the weights of the four criteria selected by our team using the AHP technique and ranking the alternatives (company) with the TOPSIS technique are as follows:

Brand	Rank
NFL	1
Marvel	2
Universal	3
Verizon	4
Temu	5
StateFarm	6
FanDuel	7
Budweiser	8
Dunkin'	9
BetMgM	10

Table 5. The top 10 spots for the best advertising companies

It is impossible not to completely exclude subjective evaluation in the process of pairwise comparisons, but our team is confident that the results of determining the weights of the criteria and ranking companies according to the weights will be much more objective than the results derived from just looking at the data.

These are the results of the calculation with four weights created through the four pairwise comparisons focused on each criterion with a CR value less than 0.1. As a result of the analysis, the NFL topped the four results. Despite the emphasis on importance on each of the four criteria, the reason for this result is that the number of tweets in the NFL is much higher than that of other companies, so it is expected that the weight does not change significantly. The results of the second to fourth places when calculated with weights emphasizing the ratio of negative tweets were very different from the other three results. In particular, Universal, which had a lot of negative tweets, went down to the bottom, and State Farm, which had a lot of positive tweets, ranked high.

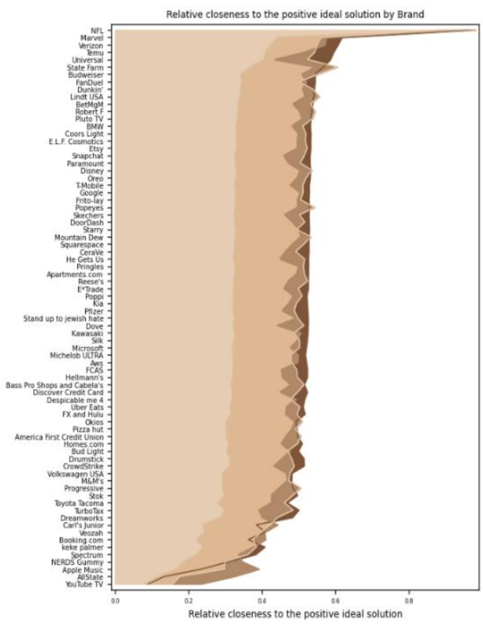

Fig 14. The result of TOPSIS with different weights on each criterion

5. **Social** Network Analysis

In the process of evaluating the impact of brand advertisements during the Super Bowl period, we have contemplated the potential for approaching it from a new perspective different from the methods we have previously discussed. Twitter, being a social media platform where users post and share short text messages, as well as explore topics using hashtags, could provide new insights if focus on its characteristics. In that point, utilizing Social Network Analysis (SNA) is highly effective. Social network analysis is the process of investigating social structures using networks and graph theory [4]. Through network visualization, one can not only discern relationships and interactions among users but also identify influential users based on centrality measures, shedding light on the extent of their influence. Moreover, SNA allows for gauging Twitter users' reactions to advertisements aired during the Super Bowl game period. We believe this will enable a deeper understanding of users and derive meaningful insights.

5.1 Experiment settings

A simple structure composed of nodes (vertices) and edges (links) generates network structures through various combinations. Nodes are usually representing entities in the network, and can hold self-properties, such as weight, size, position, and any other attribute. Also, it has network-based properties, such as Degree, number of neighbors, or Cluster, a connected component the node belongs to etc. Edges represent the connections between the nodes, and might hold properties as well, such as representing the strength of the connection, direction in case of asymmetric relation.

In the paper, we define nodes as Twitter users and brands, assuming a link between them if a Twitter user mentions a specific brand in their posts or hashtags. For instance, if user A wrote "Wow, @Drumstick with a @SuperBowl commercial.", then we considered that user A is linked between Drumstick and Super Bowl.

A node with many direct connections is considered influential, as nodes and edges are inherently intertwined, with direct connections holding more significance. Also, whether the edges are incoming or outgoing from the nodes is a critical issue. It is evident that a node with many connections is receiving or giving popularity, attention, or recognition from/to other nodes. So, in the paper, we calculated the size of nodes by reflecting these degree centrality measures. We set the maximum value to 1,000 and the minimum value to 1.5. Additionally, the color of nodes was randomly determined. If a user mentioned two brands, the color of the user node was randomly chosen from the colors of the two mentioned brands. Furthermore, in order to effectively represent situations where users mentioned brands, we depicted a directed graph.

5.2 Used tools for analysis

The process of creating nodes and links was carried out using Python, and to visualize social network, we utilized NodeXL. This program is a social network analysis and visualization tool based on Microsoft Excel, offering various network analysis techniques and visualization formats for large-scale data.

5.3 Experiment results

To compare the impact of Super Bowl game on brand advertisements, we divided the given data into three parts.

5.3.1 Initial game time

We defined the "initial game time" as the first 10 minutes following the game. During that period, 2,651 nodes were created, and 2,902 edges were established between nodes. Out of the total nodes, 57 represented brands, accounting for approximately 59.4% of the total 96 brands. It is interesting to note that within the first 10 minutes since the start of the Super Bowl game, most of the brands were mentioned.

Fig 15 illustrates the network during this time. In that figure we notice that while nodes are scattered, there are certain nodes around which edges converge. In particular, we can observe three nodes of interest. The navy blue node represents National Football League (NFL), the red node represents Marvel, and the orange node represents Verizon. The node representing the NFL, the organization hosting the Super Bowl, being the largest in size, aligns with our expectations. It will be necessary to observe how the other nodes evolve over time.

5.3.2 Half game time

The second part refers to the period before the start of the half time show, until the beginning of the game, which we defined as the "Half game time". During this period, 14,559 nodes were created, and 17,550 edges were established between nodes. Out of the total 96 brands, 82 were mentioned, accounting for approximately 85.4%. Compared to the first part, the number of nodes increased by approximately 549%, and the number of links increased by 604.7%. This result intuitively demonstrates the impact of the Super Bowl game and helps us understand why brands invest significant amounts of money to advertise during the Super Bowl game time. Fig 16 represents the network during the half game time. As evident from the sharp increase in the number of nodes and links, we can observe a rise in the density of the network,

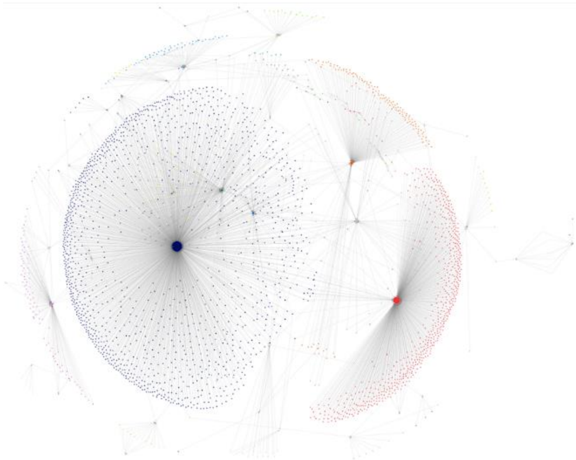

Fig 15. Initial game time Network

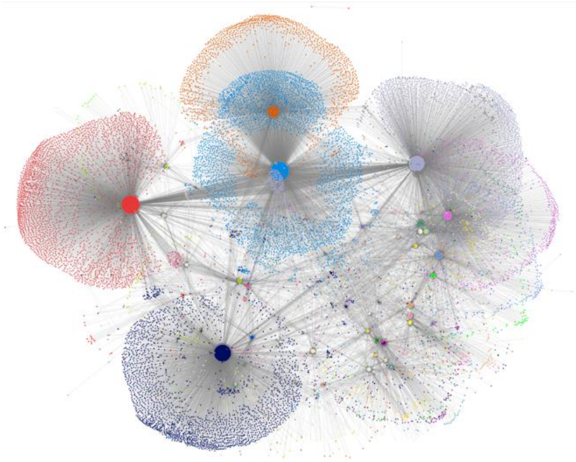

Fig 16. Half game time Network

Brand	Node Color	Initial game time		Half game time		Growth rate (%)
		Node size	Rank	Node size	Rank	
Temu		6.13	8	360.7	1	5,884.18
Marvel		78.58	2	351.1	2	446.81
NFL		169.9	1	301.1	3	177.22
Universal		6.37	7	248	4	3,893.25
Verizon		17.9	3	154.8	5	864.80
Budweiser		11.05	5	87.83	6	795.11
State Farm		1.5	-	50.22	7	3,348.00
Dove		13.7	4	38.81	8	283.28
Dunkin'		1.5	-	29.44	9	1,962.67
BMW		1.5	-	29.32	10	1,954.67

Table 6. Comparison of node sizes and growth rates between top 10 brands during half game time and the initial game time

compared to Fig 15. Additionally, this confirms that not only a closer proximity between nodes have increased but also there is a significant increase in the number of nodes where edges converge. The most interesting point is the change in the ranking of brand node sizes compared to the initial game time.

Table 6 displays the top 10 brands during the half game time and how they performed during the initial game time. We did not display nodes with a size of 1.5 during the initial game time, as they are not meaningful for ranking purpose. All brands experienced a significant increase in Twitter mentions. Particularly, Temu showed the most notable growth, surpassing the initial game time's top brand, NFL, to claim the first position. On the other hand, NFL, which exhibited the second lowest growth rate, slipped to third place. The incline ranking of Universal and the decline in ranking of Verizon can be understood in a similar context.

Additionally, it is noteworthy that State Farm, Dunkin, and BMW, which showed remarkable growth rates and ranked in the top 10 during the half game time, had a node size of 1.5 during the initial game time. We attributed these results to the timing of when brands advertised. Most of the companies that ranked in top 10 during the half game time had advertised before the half time show. Specifically, State Farm and BMW advertised in the first quarter, and Dunkin advertised in the second quarter.

However, it is worth contemplating why Budweiser, despite showing 1.78 times higher than Marvel, experienced a decrease in rank. This indicates that assessing the influence of brands cannot be solely based on growth rates. Budweiser exhibited a higher growth rate than Marvel; however, during the half game time, the node size of Budweiser was only slightly larger than Marvel's node size during the initial game time.

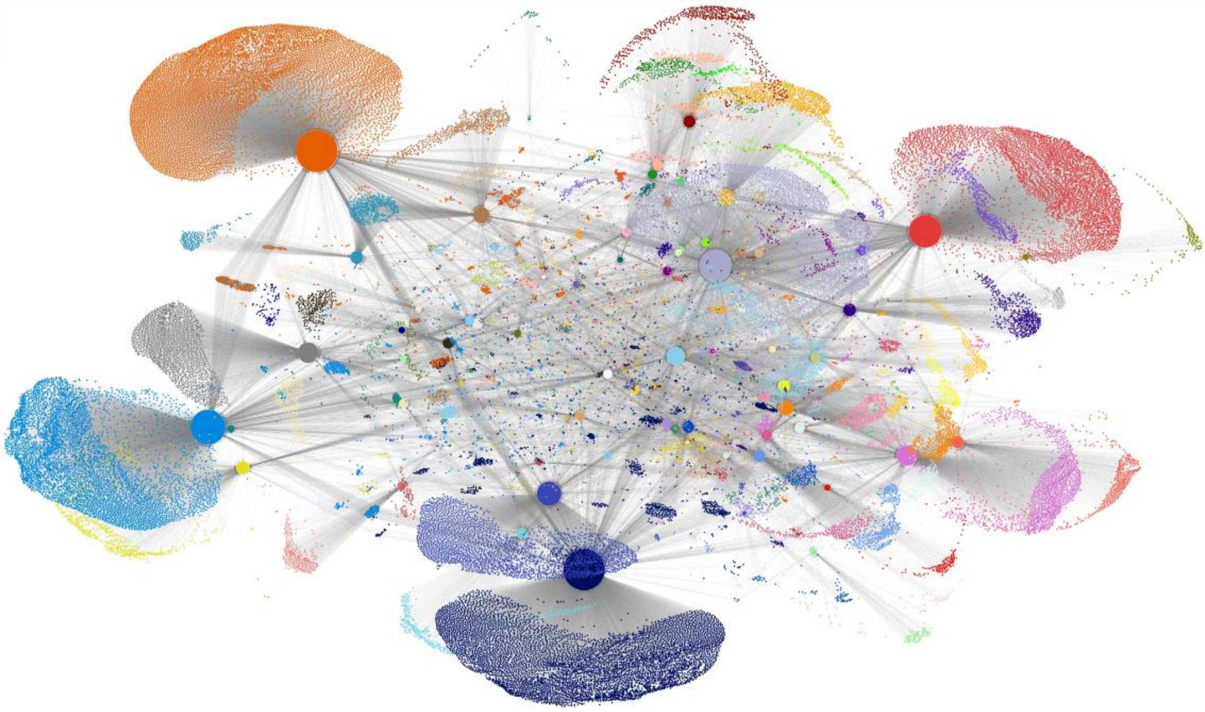

Fig 17. Whole game time network

5.3.3 Whole game time

We defined the period after all games ended as the final third part. Even just until the half time show, also known as the highlight of the Super Bowl game, we observed a significant increase in the number of nodes and edges. It was very intriguing to contemplate how the network might have changed immediately after the game ended. During that period, 51,024 nodes were created, and 66,160 edges were established between nodes. Also, all 96 companies that advertised were included in the overall node count. As expected, there were sharp increases in the number of nodes, which rose by approximately 350%, and edges, which increased by 378%, compared to half game time. This figure indicates that compared to the initial game time, there was approximately a 1,924.71% increase in the number of nodes and a 2,279.81% increase in the number of edges. The rapid increase in Twitter mentions within just a few hours was a remarkably surprising result for us. Fig 17 represents the whole network during the Super Bowl game time. Compared to previous networks, density of the network and closeness between the nodes is a notable point. Furthermore, as the volume of the network increased, we observed a greater number of nodes reaching a certain level of size, resulting in a more diverse network. Moreover, it creates the impression of interconnection between brands. A user tweeting about multiple brands causes edges to extend in multiple directions.

In the whole game time network, we focused on two points: First is the significant fluctuation in brand rankings based on node size, and the second is the intermediary nodes that connected each of the brands. Fig 18 represents the top 20 brands after the Super Bowl game, compared to the half game time. The node size of most companies exhibited a significant increase, particularly notable fact was that the top 5 groups remained unchanged from half game time to the whole game time.

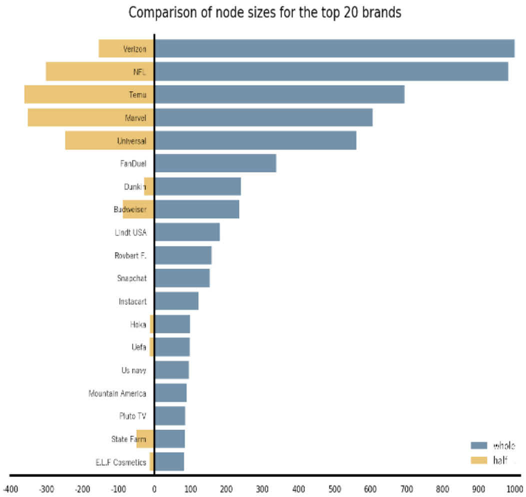

Fig 18. Comparison of node sizes

for the top 20 brands

Upon closer examination, it's worth noting the ascent of Verizon from 5th place to 1st place. We believe this can be explained using the same logic discussed in section 5.3.2 regarding the increase in node size during advertising time slots. In fact, Verizon advertised during the third quarter, which contributed to this shift in ranking. This signifies that their advertisements have resonated with people, promoting action. Additionally, we can consider the remarkable growth trajectory of Temu, a relatively new player in the United States. From Temu's perspective, large events like the Super Bowl present crucial opportunities to elevate brand visibility. Based on this, we can infer that they crafted advertisement aimed at garnering significant reactions from people, and their strategy is proved to be successful.

It is necessary to delve deeper into the strategies employed by these two groups as mentioned above. Because this is evident from the fact that they generated more tweets than even well-established entities like Marvel or Universal, both of which already have dedicated fan bases.

The second notable point of the whole network is about intermediary nodes. In fig, we can observe numerous edges that resemble connections between brands. This is one of the pieces of information that can be gleaned by representing Twitter data as a network. Intermediary nodes refer to instances where a user, who tweeted about multiple brands, occupies a central position among the nodes representing those brands. As the network volume increased, the presence of intermediary nodes became more evident. Companies need to focus on these intermediary nodes. For instance, if a user mentioned the brands they represent share common characteristics, they might become loyal users to one brand at some point.

5.4 Summary

In this section, we have represented tweets related to brands advertising during the Super Bowl game as a network structure and divided it into three-time parts to assess the impact of the Super Bowl game on advertising. Over time, we observed a significant increase in the number and size of nodes represented in the network, as well as the number of edges. Additionally, the ranking based on node sized contributed to change. Finally, Verizon took the top spot. But this result does not match with the TOPSIS results. Because when drawing the network, the data used does not account for instances where one person mentioned a specific brand multiple times. We believe that the analysis conducted from two perspectives is both valid. Therefore, it seems necessary to consider various perspectives comprehensively.

6. Conclusion

This paper discussed various methods for determining the most effective advertising brands during the Super Bowl game period by analyzing Twitter data to the advertisement aired by brands. First, we analyzed how tweets changed over time. This result left us questioning whether it is more effective to receive tweets continuously or intensively at specific points in time. This led to further analysis of how much money each brand spent to generate a single tweet. However, we were unable to find suitable insights to explain these analytical results. So, we determined that it would be beneficial to explore another factor and proceeded with sentiment analysis of the posts and hashtags written by users. The results were so remarkable that we deemed it crucial to consider these factors when assessing the influence of brands.

As we conducted analysis on the three factors, we realized the need for a comprehensive method to assess all these elements. Therefore, we adopted AHP and TOPSIS as one of the notable multi criteria decision making methods. We used the number of tweets, cost per tweet, the ratio of positive tweets, and the ratio of negative tweets as evaluation criteria, and calculated the weights using AHP. Subsequently, we conducted evaluations for each brand based on the calculated weights.

Moreover, conducting TOPSIS while varying the weights of each evaluation factor revealed an intriguing finding: NFL consistently dominated the first place in all experiments. Additionally, it is noteworthy that, except when assigning the highest weight to the Negative tweet ratio, the rankings for the second and the third places remained unchanged. We revisited our initial research questions considering these findings. For instance, we contemplated whether it is a valid assessment to conclude that the NFL, as the organizer of the Super Bowl, performed the best in advertising.

To address these concerns, we conducted social network analysis and obtained unexpected results. Through the changes in network size over time, we were able to intuitively observe the influence of Super Bowl. In TOPSIS, Verizon ranked third, but had the largest node size. However, this does not imply that network analysis is the only valid approach. Both perspectives are valid methods, and it is necessary to consider them comprehensively when devising strategies for each brand.

The AHP method we used has a limitation in that it subjectively assigns weights. Although in our experiment, we assessed the consistency of decision makers' preferences when the indicator was less than 0.1, various results could still emerge depending on the weights chosen. We hope that this analysis will help each brand formulate advertising strategies that align with their objectives.

References

[1] https://techreport.com/statistics/most-memorable-super-bowl-commer cials/#:~:text=And%20since%20many%20ads%20make,even%20more%20impact-ful%20and%20memorable.

[2] https://www.kantar.com/north-america/campaigns/super-bowl-2022/creative-effec-tiveness-download

[3] https://monkeylearn.com/sentiment-analysis/

[4] https://towardsdatascience.com/social-network-analysis-from-theory-to-applica-tions-with-python-d12e9a34c2c7