

The Triumphs and Tragedies of Fandom: Emotional Arcs in NFL Tweets

Elisabeth Kollrack

Undergraduate Honors Thesis

Department of Mathematics and Statistics
College of Engineering & Mathematical Sciences
The University of Vermont

Advised by Christopher M. Danforth and Peter Sheridan Dodds

Contents

1	Abstract	2
2	Introduction	2
2.1	Background and Motivation	2
2.2	Research Questions	6
3	Methods	7
3.1	Data Collection	7
3.2	Spatial Analysis	8
3.3	Sentiment Analysis	11
4	Results	12
4.1	Geographic Fandom Structure	12
4.2	In-Game Patterns	16
4.3	Sentiment Variations	24
5	Conclusion	30
6	Acknowledgments	32
A	Data Processing	38
B	Additional Figures	40

1 Abstract

This research examines how fan sentiment toward National Football League (NFL) teams is expressed on Twitter and how it varies with geography, game outcomes, and team performance. Geotagged tweets from the 2011–2014 NFL seasons were used to estimate a geographic fandom radius for each team based on normalized tweet activity across metropolitan areas, identifying the regions where engagement with a team exceeds background levels of discussion. Tweet sentiment was measured using the Hedonometer, which assigns happiness scores to words based on human evaluations. Sentiment was analyzed both within each team’s fandom radius and across all tweets mentioning the team, revealing only small differences between local and overall sentiment patterns. Additional analysis comparing tweets associated with winning and losing teams showed that winning teams tend to use more positive language and fewer negative words. Sentiment for both groups is similar before games, drops at kickoff, and rebounds slightly by halftime. After halftime, the trajectories diverge: sentiment for winning teams increases toward the end of the game, while sentiment for losing teams remains low. Finally, a comparison between average team sentiment and win percentage revealed a weak positive relationship, suggesting that while team success contributes to more positive fan sentiment, other factors also influence how fans discuss the NFL on social media. Our work contributes to a growing body of computational social science research that quantifies the relationship between human emotion and collective attention.

2 Introduction

2.1 Background and Motivation

The internet has allowed fans to engage with professional sports without being physically present at games [1]. Through digital platforms such as Twitter, fans can interact with one another and share moments of joy, frustration, and record-breaking achievements [2]. Twitter

provides time-stamped and geographically tagged text, enabling fans to react to games in real time and making it a valuable source for studying fan engagement [3]. More recently, the proliferation of sports gambling and prediction markets has further driven online engagement. Twitter allows sports bettors to exchange information, react to live events, and influence one another’s wagering behavior [4].

Fandom often extends beyond a team’s physical location, and this research measures the geographic extent of sports fandom through an empirical estimate of what we refer to as the “fandom radius”. Rather than relying on traditional measures such as ticket sales or surveys, we use geotagged tweets to identify where fans are located. Some fandoms may exhibit more positive or negative sentiment, and this study analyzes how sentiment differs across fandoms as well as between winning and losing teams during games.

Storywrangler is a language processing tool used to analyze Twitter data that can track sociolinguistic and cultural trends [5]. The system analyzes billions of tweets and allows one to explore word popularity, cultural shifts, and social amplification. Tweets are broken up into n-grams (sequences of words) and analyzed by frequency, rank, and popularity as measured by retweets. Our research uses a similar approach, as it searches for Twitter hashtags (team names) and tracks their frequencies through time. Storywrangler has been used previously to analyze social movements ([#BlackLivesMatter](#), [#MeToo](#)), politics, COVID-19, and popular culture, providing useful insights into human communication dynamics[6, 7, 8, 9, 10].

This methodology can be used to analyze attention dynamics across different types of events. Arnold et al. [11] examined tweets involving the names of hurricanes and compared public attention on social media to hurricane severity. The key findings were that stronger hurricanes generated more engagement regardless of death toll, and that storms affecting the continental United States received substantially more attention than those impacting regions such as Puerto Rico. Their results suggest that Twitter engagement can influence disaster response and the allocation of funding and aid.

Twitter is also used for real time event detection. Sakaki et al. [12] created a Support

Vector Machine to track tweets based on keywords and context as an early indicator of events. Filters were also used to track the location and severity of events. Event tracking could be helpful in the present research to determine which team the author of a tweet likely supports. When implemented in Japan, earthquakes of seismic intensity 3 or higher were detected by the algorithm. In some cases, this system detected earthquakes faster than the Japan Meteorological Agency, taking advantage of social media users as a proxy for seismometer sensors.

Twitter is a key platform for social media engagement in the context of sports. Einsle et al. examined how user interactions such as likes, comments, and shares boost fan engagement [13]. By enabling retweets and replies, Twitter allows fans to interact with one another in real time and quickly share updates about games. Social media allows fans to remain connected and avoid missing major events if they are not watching a game live. Fear of Missing Out (FoMO) leads to Continuous Partial Attention (CPA), where fans interact on social media while watching games [14]. Twitter is especially useful for aggregating discussions about sporting events because hashtags help organize conversations and increase the visibility of trending topics.

Smith researched differences between fans who tweet during games and those who tweet after games [15]. The study explores how Twitter use affects fans' enjoyment of events and the role of Twitter as a second screen. The authors found that fans who used Twitter while watching games appeared to enjoy the experience more than those who only read tweets. They also found that fans who interacted with others online were more interested in the game and reported higher enjoyment. These results suggest that Twitter interaction provides fans with a sense of community, which enhances enjoyment.

Geographic identity also plays a role in how fans maintain loyalty to teams outside their local market. Reifurth et al. surveyed 920 nonlocal U.S. sports fans and found that physical distance from a team did not weaken team identification [16]. Fans who identified with the city of their favorite team showed stronger loyalty than those who identified with a

broader region. These results suggest that fandom is shaped not just by proximity, but by psychological connections to place. For this research, geotagged tweet distributions may reflect these geographic identity patterns, helping explain why certain teams attract fan activity far from their home markets.

Chierichetti et al. [17] used real time event data to examine how large populations responded to major events such as the Super Bowl [18], award shows, and the World Cup. These events were detected using tweet volume and recent activity. The authors found that during games, original tweets increase while replies and retweets decrease. After games, the opposite pattern emerges, with replies and retweets becoming more common. The researchers used a logistic regression classifier to detect events and a probabilistic model to differentiate between general tweeting and event based tweeting. For our research, it is important to ensure that tweets referring to the game are tracked rather than all tweets posted during the game.

Corney [19] conducted similar research and used word frequency to identify popular topics. The study focused on FA Cup finals and examined significant moments during matches. Their system used clustering to detect important moments, including those occurring during overlapping events. The authors were able to distinguish between different teams' fans and how they reacted differently to the same event. Their methods successfully detected most major events, such as goals and saves, during World Cup finals.

Twitter data has also been used to measure societal happiness over time. Dodds et al. [20] created the Hedonometer, a tool that assigns happiness scores based on the words used in tweets. The Hedonometer found that Saturday is the "happiest" day of the week and Tuesday the "saddest." Holidays tend to be associated with higher happiness scores, while natural disasters and celebrity deaths correspond to decreases. This metric analyzes how words such as "love" and "food" are associated with higher happiness, while words such as "hate" and "funeral" are associated with lower happiness. The Hedonometer can be used as a real time tracker of population level mood, as well as for event detection and analysis

of public responses. This instrument is especially helpful for examining how fandoms differ based on team success.

Reagan et al. used the Hedonometer to analyze emotional arcs in a corpus of fiction from Project Gutenberg [21]. To construct these arcs, the authors divided each story into 10,000 word segments and assigned a happiness score to each segment based on the words it contained. These scores formed a time series representing the emotional trajectory of the story. They applied three independent methods for this analysis: Singular Value Decomposition (SVD), hierarchical clustering, and a neural network. Notably, all three approaches converged on the same set of results, strengthening their findings. The analysis revealed six fundamental emotional arc patterns that characterize the majority of stories. One of these, the “Man in a Hole” arc, consists of a decline in sentiment followed by a recovery, and is particularly relevant to our research, as discussed later in this paper.

Together, these studies demonstrate that Twitter data can be used to detect events, measure attention, and infer collective behavior in real time. Building on this literature, the following section outlines the data and methods used to identify and analyze geographic patterns of sports fandom using geotagged tweets.

2.2 Research Questions

- How does geographic fandom size vary by team?
- Is there a difference in sentiment of tweets based on fandom?
- How do tweet volume and sentiment change in real time during a game, and how are these patterns affected by the game outcome?

3 Methods

3.1 Data Collection

Twitter data used in this project were collected through the Vermont Advanced Computing Center (VACC). Twitter’s Decahose API provides access to a random 10% sample of all public tweets from 2008 through 2023. Each tweet includes metadata such as the tweet text, timestamp, and geographic coordinates when available.

The dataset for this study consisted of tweets authored during the 2011 through 2014 National Football League (NFL) regular seasons. The 2011 season began on September 8, 2011, and the 2014 regular season concluded on December 28, 2014. Only tweets authored during the regular season were included in the dataset. We chose this time period based on the fact that GPS tagged messages, a functionality that became available in 2010, became less common following design changes in the Twitter app in 2015.

Tweets were collected surrounding each NFL game played during the study period (2011-2014). For every game involving a given team, a seven day window was constructed spanning three days before the game through three days after the game. Tweets were retrieved if they occurred within this window and contained team related hashtags.

Tweet timestamps were recorded in Coordinated Universal Time (UTC), while official game kickoff times were listed in Eastern Standard Time (EST). To ensure consistency, all game times were converted to UTC using Python’s `pytz` package prior to constructing the time windows.

Hashtags were defined for each team using a predefined set of team identifiers, including team mascots (e.g., `#Bills`) and full team names (e.g., `#BuffaloBills`; see Table A1 in Appendix A.1). In addition, matchup specific hashtags were generated for each game using both team abbreviations (e.g., `#NEvsBUF` and `#BUFvsNE`). Hashtag matching was performed in a case insensitive manner.

Tweets were retrieved separately for each hashtag and then combined into a single dataset.

Duplicate tweets resulting from multiple hashtag matches were removed using unique tweet identifiers. For each tweet, metadata were recorded including the associated team, their opponent, game identifier, season, home/away designation, and game timestamp.

Metropolitan population data were obtained at the Core Based Statistical Area (CBSA) level using annual population estimates from 2011 through 2014. The dataset was restricted to Metropolitan Statistical Areas. Season specific population estimates were used for analyses conducted at the yearly level.

For analyses aggregating across all seasons, the average population across the four years was calculated and used as the population measure. Population estimates were merged with a CBSA shapefile using CBSA identifiers, allowing population data to be spatially joined with the metropolitan regions used in the geographic analysis.

3.2 Spatial Analysis

To quantify the geographic extent of each NFL team’s fanbase, we define and compute a *fandom radius* based on the spatial distribution of geolocated Twitter activity. Each tweet is associated with a latitude and longitude coordinate, and distances are measured relative to a team specific city center.

For consistency across teams, the city center is defined using the primary city, often referenced in the team’s name, rather than the physical location of the stadium. For example, the New England Patriots play their home games in Foxborough, Massachusetts, but the geographic center of Boston is used as the team’s reference point. This choice reflects the broader metropolitan identity of each franchise rather than stadium specific geography.

Tweets are spatially joined to U.S. Core Based Statistical Areas (CBSAs), representing major metropolitan regions, using `geopandas`. Distances between tweet locations and team city centers are computed using a projected coordinate system (EPSG:5070) to ensure accurate measurements in kilometers. For each metropolitan area, we calculated the mean distance of its associated tweets from the city center as well as the total number of tweets

originating within that region.

Metropolitan areas vary substantially in population size, and so raw tweet counts are normalized by population to allow accurate comparison across regions. Tweet activity is expressed as tweets per 100,000 residents using season specific metropolitan population estimates. Normalized tweet activity for each metropolitan area, $A(M_i)$, is defined as follows:

$$A(M_i) = \frac{\text{Tweets}(M_i)}{\text{Population}(M_i)} \times 100,000 \quad (1)$$

Metropolitan areas are then ordered by increasing distance from the team’s city center, and their populations are cumulatively summed. The metropolitan areas are represented by M_1, M_2, \dots, M_n and are sorted by increasing distance d_i from the team’s city center. We define a core region as the set of metropolitan areas whose combined population accounts for at least 70% of the total population represented in the data. The baseline distance, d_{baseline} , is defined as the distance of the metropolitan area at which this 70% population threshold is first reached. Formally, this distance is given by:

$$d_{\text{baseline}} = d_k \quad \text{where} \quad k = \min \left\{ m : \sum_{i=1}^m \text{Population}(M_i) \geq 0.70 \sum_{i=1}^n \text{Population}(M_i) \right\} \quad (2)$$

Here, k represents the index of the metropolitan area at which the cumulative population first reaches at least 70% of the total population. In other words, M_k is the outermost metropolitan area included in the core region when aggregating outward from the team’s city center. d_{baseline} is intended to capture a team’s primary regional market.

Baseline tweet activity, A_{baseline} , is calculated as the average normalized tweet activity across all metropolitan areas located at distances greater than or equal to the baseline

distance. Baseline activity is defined as:

$$A_{\text{baseline}} = \frac{1}{N} \sum_{i: d_i \geq d_{\text{baseline}}} A(M_i) \quad (3)$$

where $N = |\{i : d_i \geq d_{\text{baseline}}\}|$ is the number of metropolitan areas at or beyond the baseline distance. Metropolitan areas with $d_i \geq d_{\text{baseline}}$ are treated as the baseline (background) region used to estimate typical non-local engagement. The *fandom radius*, R_{fandom} , is then defined as the distance from the city center at which normalized tweet activity falls below this baseline level.

$$R_{\text{fandom}} = \min\{d_i : A(M_i) \leq A_{\text{baseline}}\} \quad (4)$$

This distance represents the spatial boundary beyond which engagement with the team is indistinguishable from background interest. Table 1 provides examples of the intermediate quantities for several teams, including the values of $A(M_1)$, k , d_{baseline} , N , A_{baseline} , and the resulting fandom radius R_{fandom} . These examples demonstrate how variation in baseline distance and activity levels leads to differences in the geographic reach of team fandoms.

Table 1: Example calculations of fandom radius components for selected NFL teams. For each team, the table reports normalized tweet activity at the city center $A(M_1)$, the index k at which the 70% cumulative population threshold is reached, the corresponding baseline distance d_{baseline} , the number of metropolitan areas N used to compute baseline activity, the baseline activity level A_{baseline} , and the resulting fandom radius R_{fandom} .

Team	$A(M_1)$ (city center)	k	d_{baseline} (km)	N	A_{baseline}	R_{fandom} (km)
Seattle Seahawks	56	250	3738	55	0.92	782
Dallas Cowboys	29	266	1975	86	3.03	369
New England Patriots	29	241	2578	83	1.6	232
Baltimore Ravens	45	213	2011	73	0.94	121

This methodology is applied separately for each season from 2011 to 2014, as well as to the aggregated data across all four seasons, producing both season specific and overall fandom radius estimates for each NFL team.

3.3 Sentiment Analysis

After defining fandom regions, the next step was to analyze the sentiment of tweets associated with each fandom. Sentiment was measured using the Hedonometer [20]. The Hedonometer assigns happiness scores to individual words based on human ratings of happiness. Scores range from 1 to 9, where 1 represents very negative sentiment and 9 represents very positive sentiment.

Tweets were grouped by fandom and assigned an overall happiness score based on the words they contained. To prevent certain terms from biasing sentiment, words such as team names and player names (e.g., “Cowboys,” “Redskins,” “Eagles,” “Cousins”) were removed using pattern matching with regular expressions (See Table A2 in Appendix A.1). After this cleaning step, tweets containing game-specific hashtags (e.g., #NEvsBUF) were included in the datasets for both teams involved in the game. This means certain tweets were included twice, once for each team, but this double counting occurs uniformly across all games and teams, ensuring that the relative comparisons of team and fandom sentiment remain valid. Most NFL related tweets in the dataset fell within a range of approximately 5.8 to 6.2, indicating slightly positive sentiment on average, comparable to the range observed for all English tweets on <http://www.hedonometer.org>.

Sentiment was analyzed using two approaches. The first approach considered all tweets mentioning a given team, regardless of geographic location. For each team, tweet text containing aforementioned hashtags was aggregated and an average happiness score was calculated.

The second approach analyzed sentiment at the fandom level. In this case, only tweets originating from within the geographic radius defining a team’s fandom were included. The average happiness score of these tweets was then calculated to represent the sentiment of that local fan base.

4 Results

4.1 Geographic Fandom Structure

Calculating fandom radii provides insight into the geographic “reach” of each NFL team’s fan base. Larger radii indicate a more geographically dispersed fan base, while smaller radii indicate a more concentrated fan base centered around the team’s home region.

Figure 1 shows the distribution of fandom radii across all NFL teams. Radii vary substantially across the league, ranging from 121 kilometers for the Baltimore Ravens to 782 kilometers for the Seattle Seahawks. Teams such as the Seattle Seahawks and Minnesota Vikings have the largest radii, suggesting fan bases that extend beyond their home metropolitan areas. In contrast, teams such as the Baltimore Ravens and Washington Redskins exhibit smaller radii, indicating more geographically concentrated fan bases. We note that the basic geographic constraints associated with national borders and oceans complicate interpretation of these rankings.

The patterns we identify appear to reflect the geographic distribution of NFL teams. Franchises located in regions with fewer nearby teams tend to have larger fandom radii, as they serve as the primary team for fans across multiple states. Conversely, teams located in densely populated regions with many nearby franchises tend to have smaller fandom radii due to competition for local fan bases.

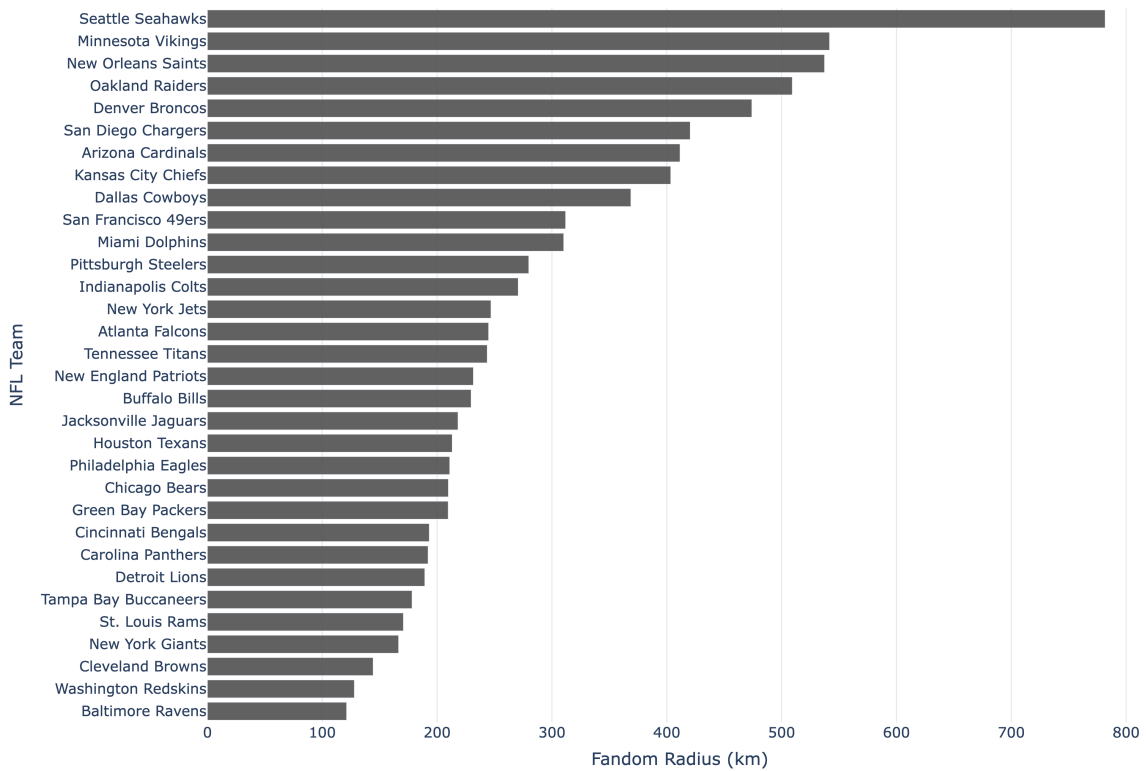


Figure 1: Estimated fandom radius for each NFL team calculated by using geotagged Twitter activity. The radius reflects the geographic distance required to capture the majority of fan tweets associated with a team, with larger values indicating more geographically dispersed fan bases.

The spatial distribution of these fandom regions is illustrated in Figure 2. The map also highlights significant overlap between several fandom regions, particularly in the Midwest and Northeast. These overlaps suggest that certain geographic areas are divided among multiple fan bases rather than being dominated by a single team. For example, teams located in close proximity, such as Oakland and San Francisco, show overlapping fandom regions, indicating shared or competing fan engagement within the same geographic area.

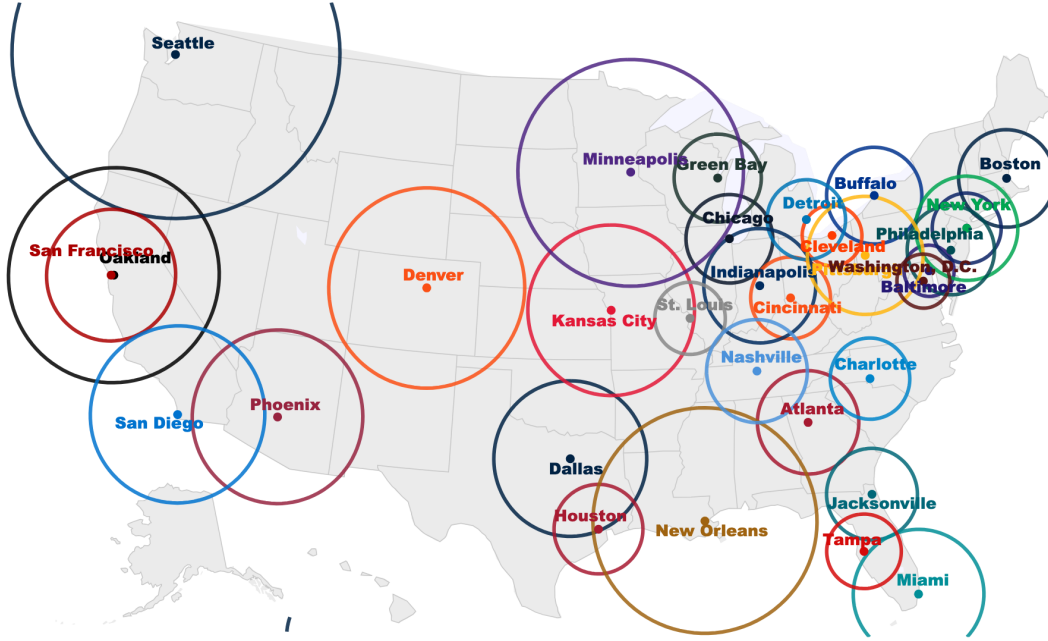


Figure 2: Fandom radius map of NFL teams across the United States. Each circle represents the approximate geographic reach of a team’s fanbase. Franchises located in regions with fewer nearby teams tend to have larger fandom radii, as they serve as the primary team for fans across multiple states. Conversely, teams located in densely populated regions with many nearby franchises tend to have smaller fandom radii due to competition for local fan bases.

Additional analysis examined how fandom density scales with metropolitan area size. Figure 3 shows a comparison between fandom intensity and spatial reach. The horizontal axis, ρ_{fandom} , is defined as $\log_{10}(\text{fandom size}/\text{area})$, where fandom size is measured as tweets per 100,000 residents. The vertical axis, ρ_{team} , is defined as $\log_{10}(1/\text{area})$. Each point represents a metropolitan statistical area with at least one tweet, aggregated across all teams and seasons from 2011–2014.

To quantify the relationship between these variables, we estimate α in

$$\rho_{\text{team}} \propto \rho_{\text{fandom}}^{\alpha},$$

using Reduced Major Axis (RMA) regression. The RMA slope is given by $\alpha = \text{sign}(r) \sigma_y / \sigma_x$, where r is the Pearson correlation coefficient and σ_x, σ_y are the standard deviations of the respective variables.

This scaling relationship captures how fandom intensity varies with the spatial reach of metropolitan areas. We find $\alpha = 0.53$ and $R^2 = 0.64$, indicating a moderate positive correlation. An alpha value less than 1 suggests that fandom is relatively spread out, increasing more slowly than would be expected under proportional scaling with spatial density. The observed value implies that fandom is less spatially concentrated and extends beyond densely populated metropolitan areas.

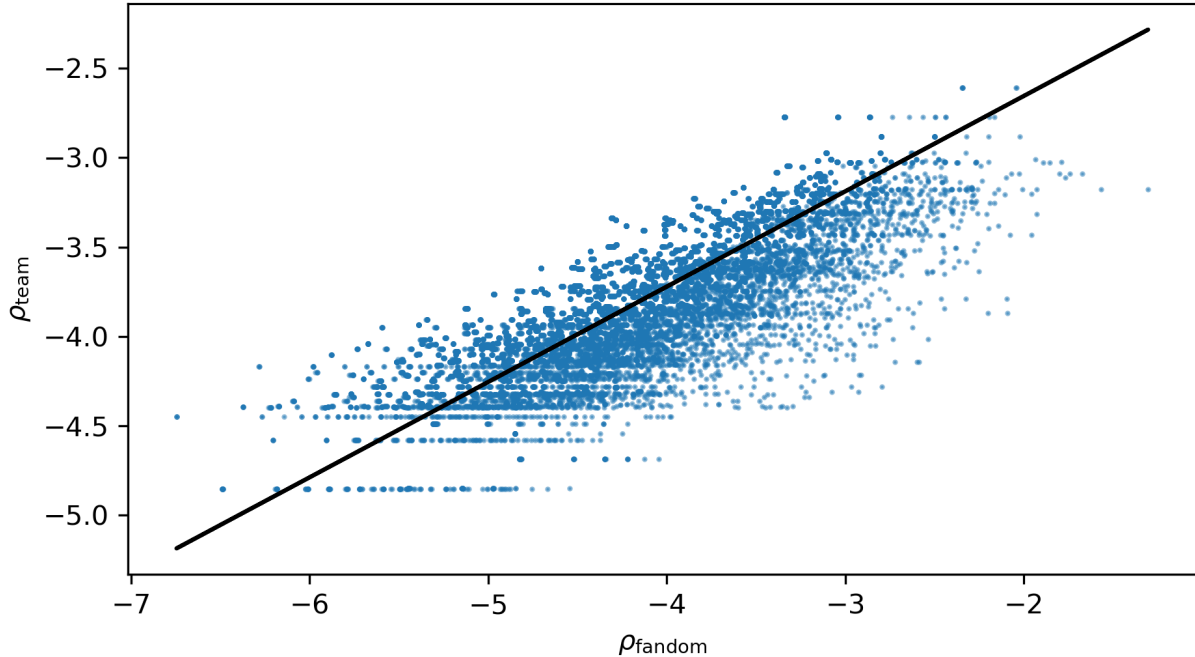


Figure 3: Log-log relationship between fandom density and metropolitan area size. Four seasons of data for 32 teams were represented (~ 1500 tweets per team per season). 19,174 data points included, with each point representing a team–season–metropolitan area combination that tweeted about the team. The horizontal axis, ρ_{fandom} , is $\log_{10}(\text{tweets per } 100\text{k residents/area})$, and the vertical axis, ρ_{team} , is $\log_{10}(1/\text{area})$. The black line shows the Reduced Major Axis (RMA) fit with slope $\alpha = 0.53$. The scaling relationship ($\alpha < 1$) indicates that fandom intensity increases more slowly than spatial density, consistent with a geographically dispersed fan base.

4.2 In-Game Patterns

To understand fan engagement in real time, we analyzed tweet activity for all NFL teams from three hours before kickoff to four hours after kickoff on game day. Figure 4 shows the total number of tweets mentioning each team in 30 minute intervals relative to kickoff. Across all teams, three key moments stand out: kickoff, halftime, and game end. Tweet activity

generally spikes at kickoff, dips during halftime (typically around 100 minutes after kickoff), and reaches a larger peak at the conclusion of the game (roughly three hours after kickoff [22]).

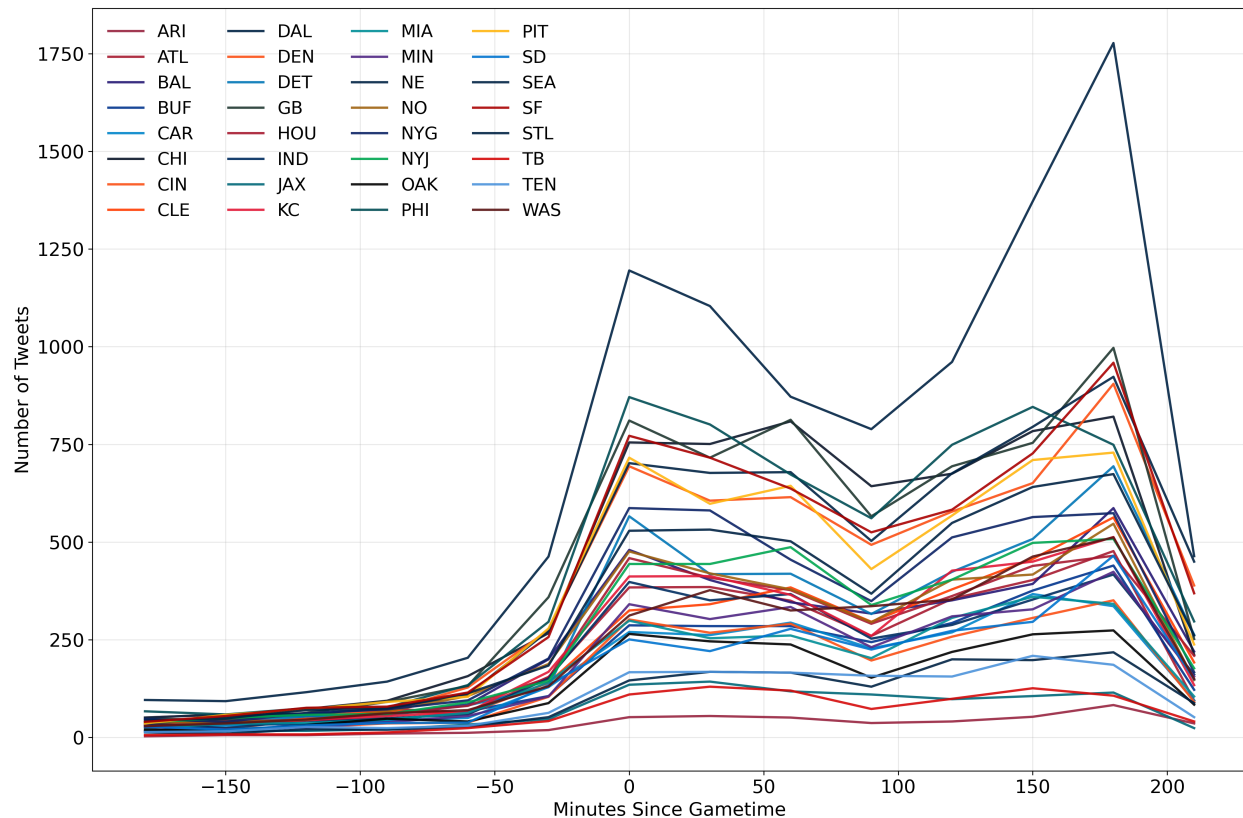


Figure 4: Tweet activity around game time for all NFL teams. Each line represents the total number of tweets mentioning a given team, aggregated in 30 minute intervals relative to kickoff (negative values indicate tweets before the game, positive values indicate tweets during and after the game). Tweets were counted from 3 hours before kickoff until 4 hours after.

Tweet activity gradually decreases after kickoff, likely reflecting fans shifting their focus from posting pregame messages to watching the game itself. The lowest point occurs during halftime, when game play pauses and online discussion temporarily decreases. Tweet volume picks up again after halftime, and steadily grows until the conclusion of the game, where fans likely share celebrations or frustrations about the final result.

The analysis also revealed that some teams generate significantly more in game conversation than others. The Dallas Cowboys, Green Bay Packers, and Philadelphia Eagles

consistently had the highest tweet volumes, while the Jacksonville Jaguars, Tampa Bay Buccaneers, and Arizona Cardinals had the fewest. Teams with larger fan engagement displayed the characteristic pattern of pregame, halftime, and post-game peaks, whereas teams with smaller in game audiences exhibited flatter, less distinct patterns.

A similar pattern appears when tweet activity is grouped by game competitiveness. Figure 5 shows the average tweet activity for different types of game outcomes. The goal of this analysis was to determine whether fans are more engaged during close games compared to games where the outcome is relatively certain.

For this analysis, games were grouped into four buckets: very close games, where the score difference was between 1 and 3 points (within a field goal of tying or winning), close games, with a 4 to 7 point margin (within a touchdown), comfortable wins with margins between 8 and 14 points, and blowouts, where a team led by at least 15 points.

Close games tend to spike around 150 minutes after kickoff (around the fourth quarter), which is consistent with the idea that fans become more attentive when the outcome is uncertain. Interestingly, blowout games show slightly higher average tweet volume during portions of the game. One possible explanation is that these games may include multiple scoring events or highlight worthy plays that prompt online discussion even when the overall outcome is less competitive.

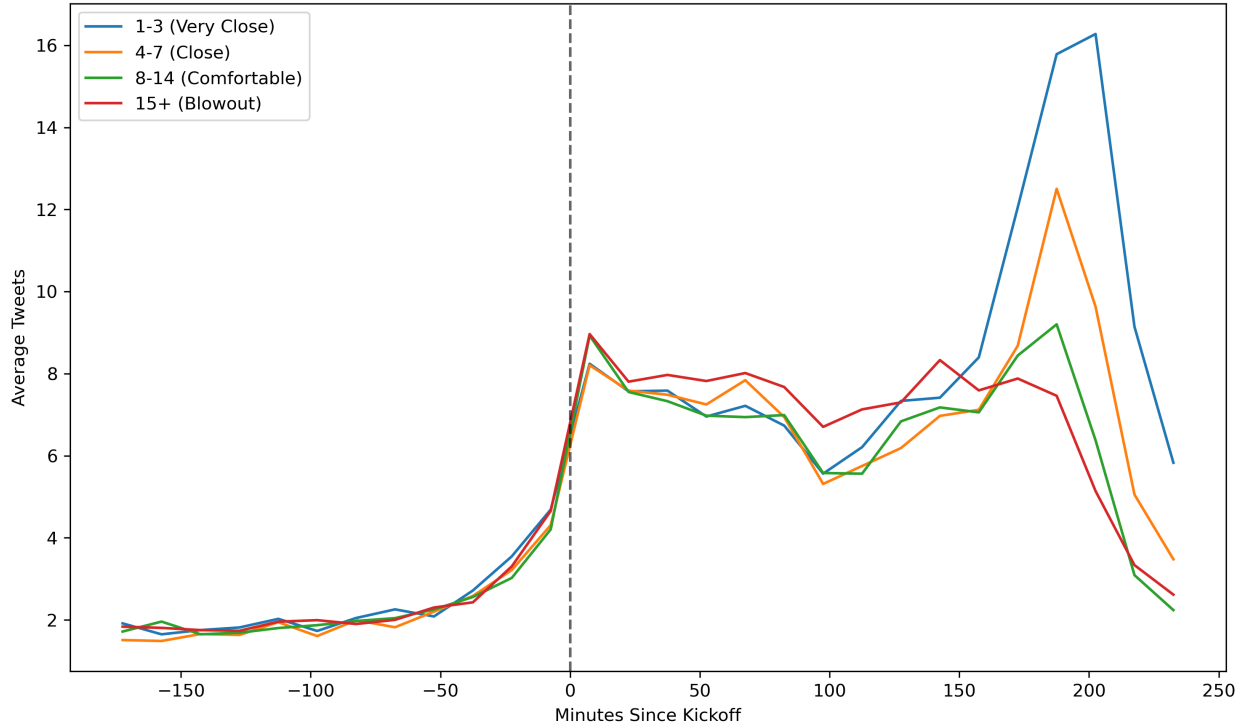


Figure 5: Average tweet volume per 15 minutes during games by point differential. Tweet counts are aggregated in 15 minute intervals from 3 hours before kickoff to 4 hours after kickoff. Lines are colored by game outcome: 1-3 points (Very Close), 4-7 points (Close), 8-14 points (Comfortable), and 15+ points (Blowout).

Note that the point differential used in this analysis reflects the final score margin, rather than the margin at each moment in the game. As a result, some games classified as blowouts may have been competitive earlier before one team pulled away late.

The next step of this analysis was to compare tweet volume to sentiment. Figure 6 shows both average tweet volume and average sentiment, separated by winning and losing teams. In the top panel, tweet activity follows a consistent pattern for both groups, with spikes around kickoff and at the conclusion of the game and lower activity during the middle of the game (as shown in Figures 4 and 5). In the bottom panel, sentiment begins relatively high before kickoff and declines once the game starts. As the game progresses, sentiment for tweets referencing the eventual winning team increases, while sentiment for the losing team decreases at first, and then remains constant. This demonstrates different fan reactions as the outcome becomes clearer.

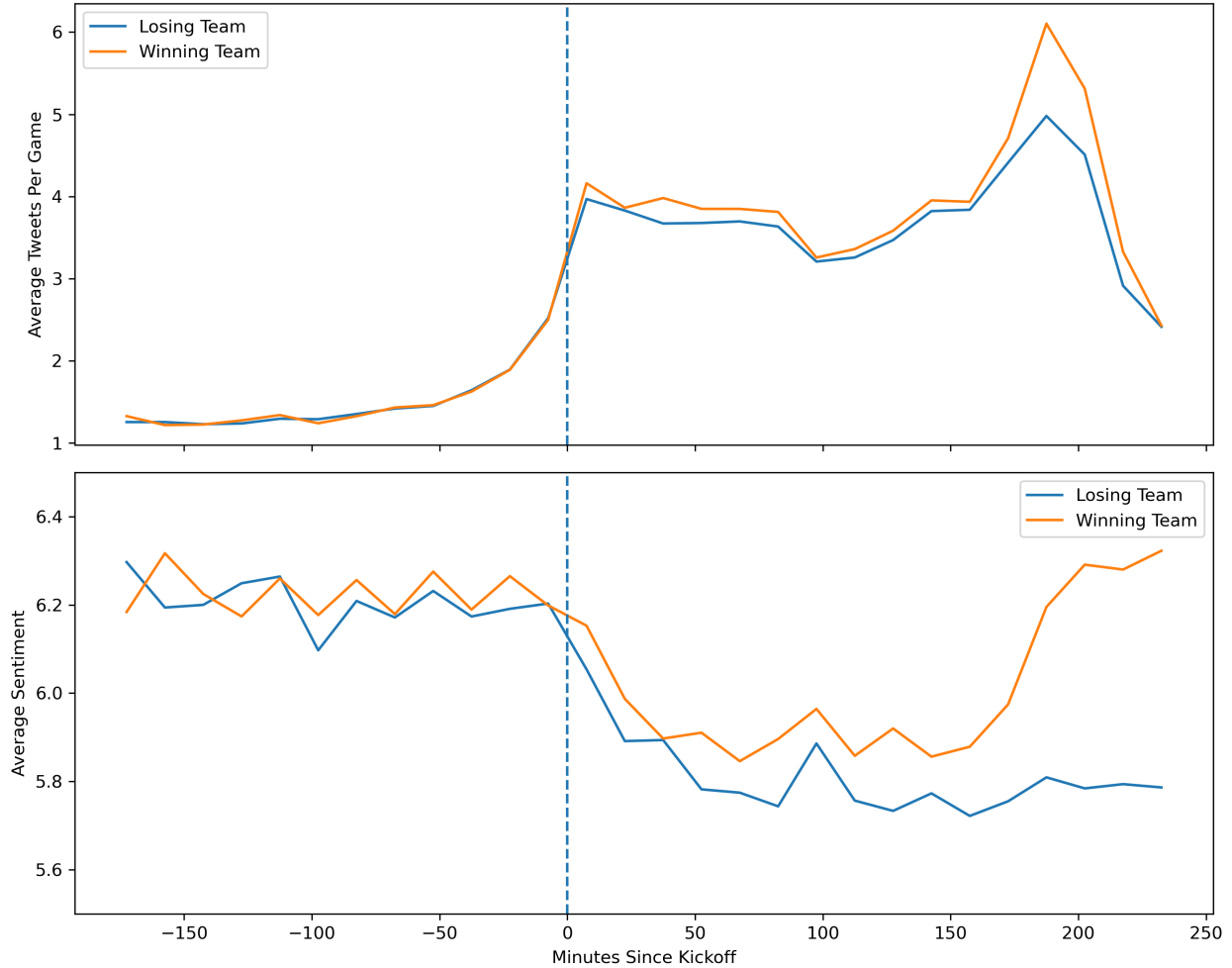


Figure 6: Tweet activity and sentiment during games for winning and losing teams. The top panel shows the average number of tweets (aggregated in 15 minute intervals) relative to kickoff, while the bottom panel shows the average tweet sentiment over the same time period. Tweets are grouped by whether they reference the eventual winning or losing team. At the beginning of the game, both winning and losing teams had an average of 4 tweets per hour (approximately 40 tweets per hour after accounting for the 10% sample).

Tweet volume is similar for both winning and losing teams until near the end of the game, when tweets mentioning the winning team increase significantly compared to those mentioning the losing team. This difference may reflect fans of the losing team being less likely to post online following a disappointing result, while fans of the winning team are more likely to celebrate and share reactions after the game ends.

Sentiment for tweets referencing winning and losing teams is very similar before kickoff, and both measures decline noticeably once the game begins. This pattern likely reflects

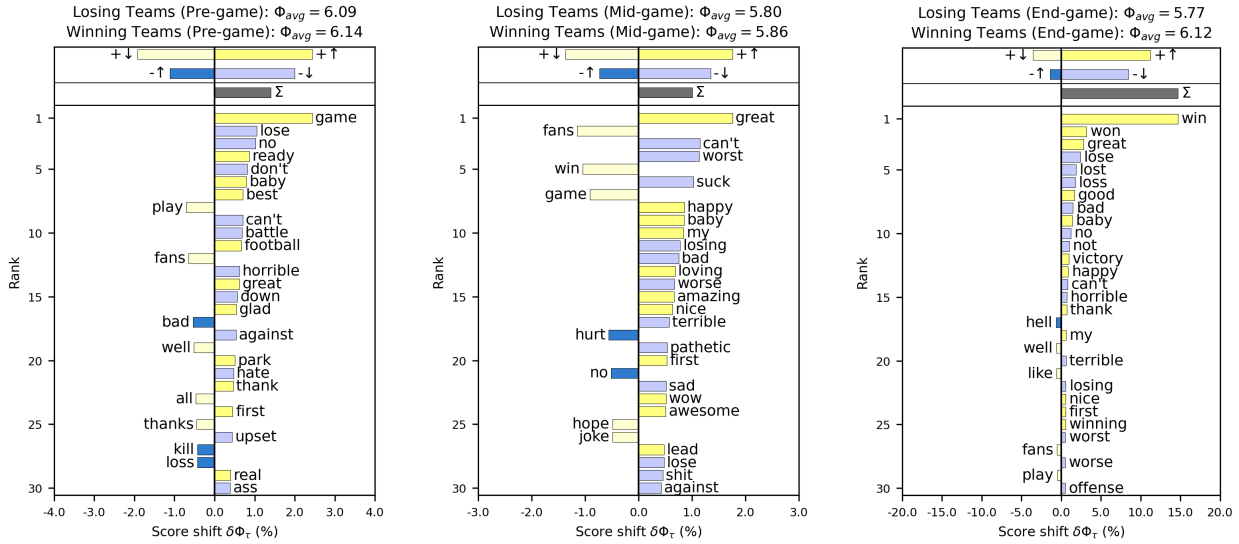
pregame optimism from both fan bases, followed by more critical or emotional reactions once the game begins. Throughout the game, sentiment for the eventual winning team remains slightly higher than that of the losing team.

Around the 100 minute mark, corresponding roughly to halftime, sentiment increases slightly for both groups. This increase may occur because fans temporarily shift discussion toward other NFL related topics or general game commentary during the break in game play. After halftime, sentiment begins to diverge more clearly. Sentiment associated with the winning team gradually increases, while sentiment for the losing team remains relatively stable. This pattern suggests that fans of the winning team increasingly express positive reactions as the outcome becomes more certain, whereas fans of the losing team continue to express more neutral or negative emotions.

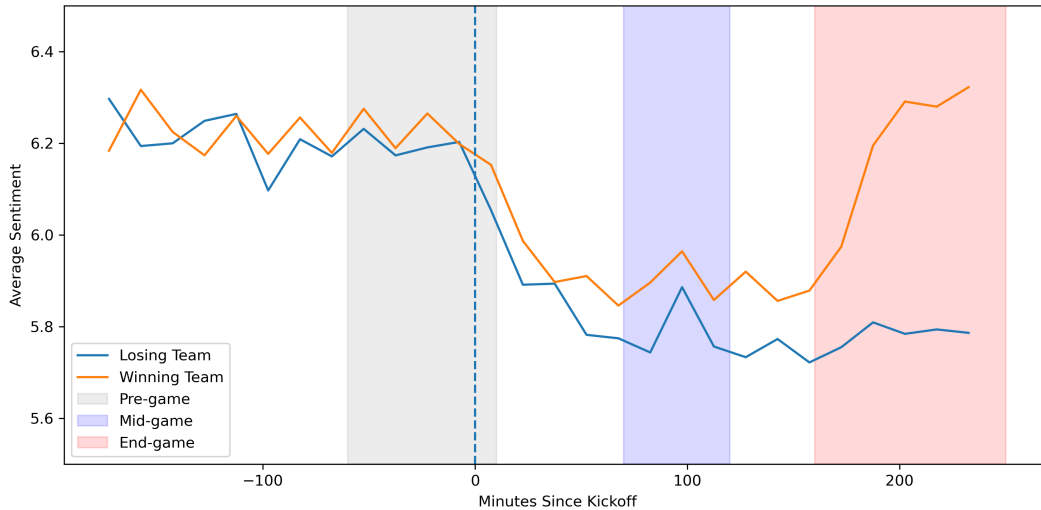
These results support a hypothesis popularized by Kurt Vonnegut that the stories we love to tell exhibit a small handful of emotional arcs [23]. Fans of both teams begin the game with optimism, posting tweets whose happiness is on par with the best individual days measured by the Hedonometer. As the game begins, fans of both winning and losing teams experience plummeting sentiment. Consistent with Vonnegut’s “Man in a Hole” emotional arc, fans of winning teams subsequently experience an emotional rise beginning near the start of the fourth quarter and ending in triumph at a sentiment slightly higher than the optimistic pre-game excitement [23]. The losing team on the other hand experiences Vonnegut’s “Tragedy” emotional arc, exemplified by *Romeo & Juliet*, and finishing with a sentiment comparable to the saddest days measured by the Hedonometer. Future work might explore the relationship between the shapes of individual game scores and fan experience [24].

Figure 7 provides insight into the sentiment differences between tweets about winning and losing teams throughout a game. Pre-game tweets show similar sentiment, with the winning team at 6.14 and the losing team at 6.09, reflecting general excitement; common positive words include “game”, “ready”, “play”, and “fans.” During halftime, sentiment drops slightly to 5.86 for the winning team and 5.80 for the losing team. By the end of the

game, the difference is more pronounced, with the winning team at 6.12 and the losing team at 5.77. These differences are driven by the greater use of positive words such as “win”, “great”, and “victory” in tweets about winning teams, as well as the more frequent use of negative words such as “lose”, “bad”, and “horrible” for losing teams. The four color summary bars at the top of the graph indicate that the dominant contributors to the overall difference are an increase in positive language, followed by a decrease in negative language in winning team tweets. Words going against this trend, namely relatively happy words used less often in winning team tweets and relatively sad words used more often in winning team tweets, are far less common.



(a) Word shifts from 60 minutes before the game until 10 minutes after kickoff. (b) Word shifts from 70 minutes after kickoff until 120 minutes after kickoff. (c) Word shifts from 160 minutes after kickoff until 250 minutes after kickoff.



(d) Tweet sentiment during games for winning and losing teams (aggregated in 15-minute intervals) relative to kickoff. The colors correspond to the word shift graphs above (Figures 7a-7c).

Figure 7: Comparison of word shifts and tweet sentiment for the eventual winning and losing teams during the game. Panels (a)-(c) show word shifts at three different time windows relative to kickoff, where dark yellow bars represent positive words used more frequently by the winning team and less frequently by the losing team, light blue bars represent negative words used less frequently by the winning team and more frequently by the losing team, light yellow bars represent positive words used less frequently by the winning team, and dark blue bars represent negative words used more frequently by the winning team. Panel (d) shows in-game sentiment trajectory for winning and losing teams, with colored sections matching the corresponding word shift panels (a-c).

4.3 Sentiment Variations

In addition to comparing the sentiment of tweets from winning and losing teams, sentiment was analyzed for tweets posted within a team’s fandom radius compared to all tweets posted. Table 2 summarizes tweet activity and sentiment for each NFL team. Specifically, the table includes the total number of tweets associated with each team, the number and percentage originating within the fandom radius, and the average sentiment both within the radius and across all tweets. Comparing these values helps reveal whether fans located near a team’s geographic fan base express sentiment differently than the broader online audience.

Table 2: Geographic fandom radius and sentiment comparison for NFL teams. Each row reports the estimated fandom radius (km), total tweet volume, the number and percentage of tweets originating within that radius, and average Hedonometer [20] sentiment scores inside the radius versus overall.

Team	Radius (km)	Total Mentions	Mentions (Inside)	% Inside Radius	Happiness (Inside)	Happiness (Total)
Arizona Cardinals	411.50	2078	203	9.77	5.95	5.97
Atlanta Falcons	244.78	4047	1338	33.06	5.95	5.95
Baltimore Ravens	121.17	4948	1653	33.41	5.97	5.93
Buffalo Bills	229.57	3972	1586	39.93	5.96	5.94
Carolina Panthers	192.10	3445	1095	31.79	6.09	6.02
Chicago Bears	209.78	8052	3184	39.54	5.81	5.86
Cincinnati Bengals	193.15	2884	1016	35.23	5.99	5.98
Cleveland Browns	144.26	4691	1735	36.99	5.95	5.90
Dallas Cowboys	368.69	12565	2876	22.89	6.03	5.97
Denver Broncos	474.04	7088	1391	19.62	6.08	6.04
Detroit Lions	189.24	5207	1454	27.92	5.90	5.95
Green Bay Packers	209.61	8028	1633	20.34	5.99	5.96
Houston Texans	213.10	4617	2241	48.54	5.94	5.96
Indianapolis Colts	270.58	3854	1292	33.52	6.10	6.04
Jacksonville Jaguars	218.16	1495	529	35.38	6.06	5.98
Kansas City Chiefs	403.40	4476	1978	44.19	5.97	5.94
Miami Dolphins	310.20	3352	851	25.39	6.10	6.01
Minnesota Vikings	541.69	3720	1395	37.50	5.94	5.98
New England Patriots	231.55	7800	2223	28.50	5.98	5.96
New Orleans Saints	537.33	5027	1333	26.52	6.11	6.03

Continued on next page

Table 2: Geographic fandom radius and sentiment comparison for NFL teams (continued).

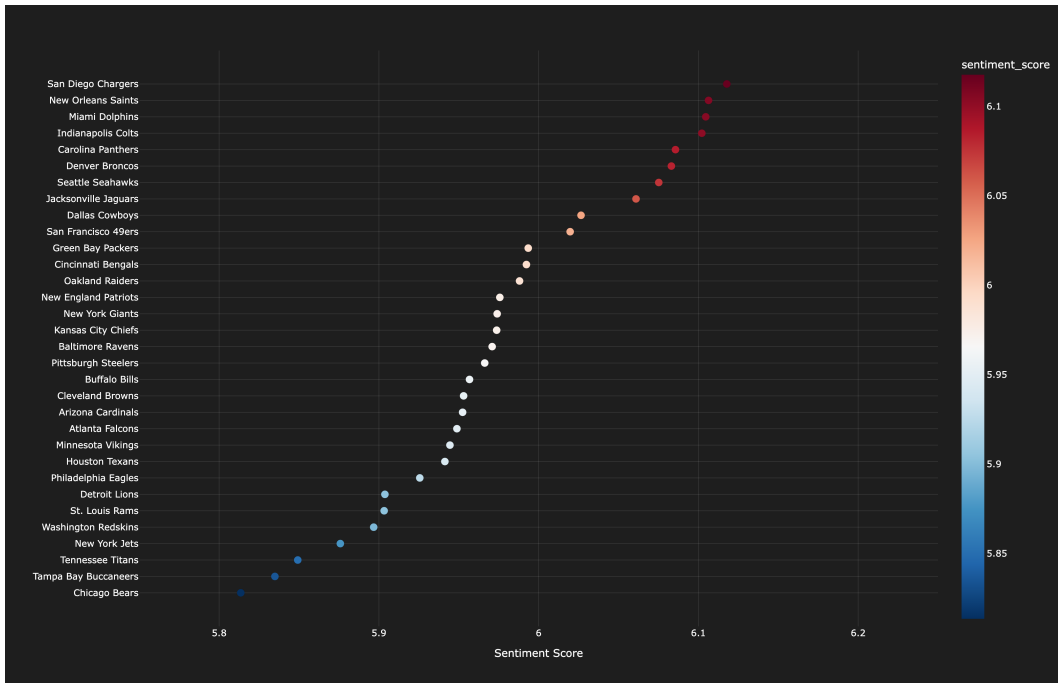
Team	Radius (km)	Total Mentions	Mentions (Inside)	% Inside Radius	Happiness (Inside)	Happiness (Total)
New York Giants	166.41	6347	1993	31.40	5.97	5.99
New York Jets	246.87	5418	2412	44.52	5.88	5.88
Oakland Raiders	509.26	2984	789	26.44	5.99	5.96
Philadelphia Eagles	210.93	8359	3741	44.75	5.93	5.95
Pittsburgh Steelers	279.76	7947	3070	38.63	5.97	5.96
Seattle Seahawks	781.69	6160	2509	40.73	6.08	6.05
San Francisco 49ers	311.85	8015	2696	33.64	6.02	6.00
San Diego Chargers	420.35	3208	1191	37.13	6.12	6.03
St. Louis Rams	170.58	2067	343	16.59	5.90	5.95
Tampa Bay Buccaneers	178.15	1127	229	20.32	5.83	5.93
Tennessee Titans	243.61	1995	639	32.03	5.85	5.90
Washington Redskins	127.89	4361	1426	32.70	5.90	5.89

Table 2 shows considerable variation across the league in the percentage of tweets posted within a team’s fandom radius. For some teams, such as the Houston Texans, Kansas City Chiefs, and Philadelphia Eagles, over 40% of tweets originate within this radius, suggesting a strong geographic concentration of fan activity. In contrast, teams such as the Arizona Cardinals and St. Louis Rams have fewer than 20% of tweets within their radius, indicating that discussion of these teams is more geographically dispersed.

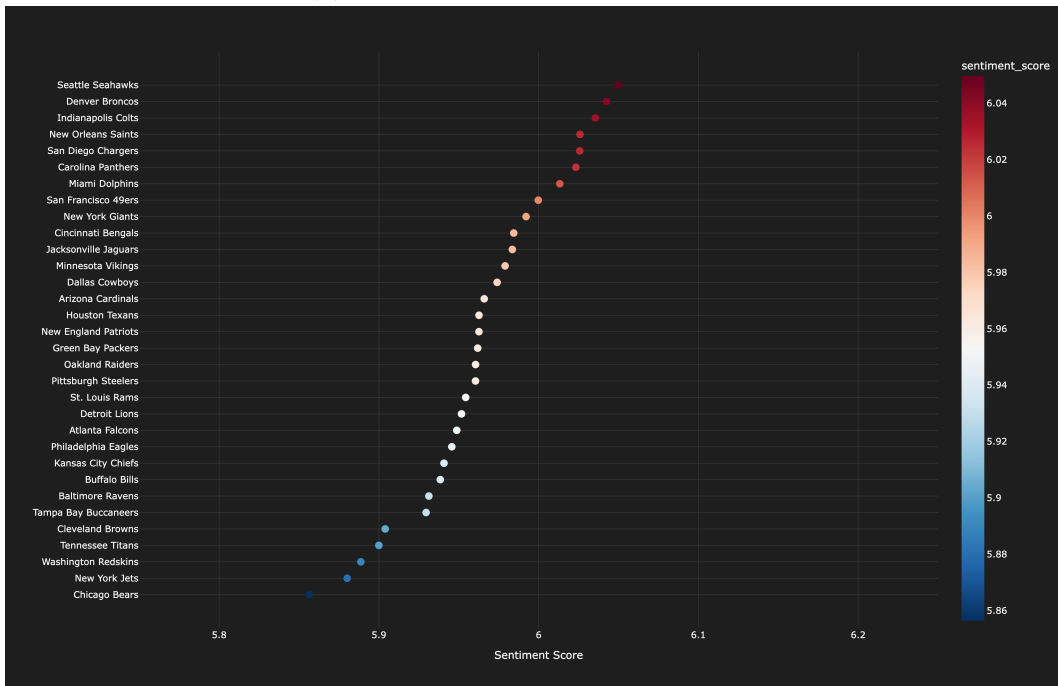
This variation is expected given how the fandom radius is defined (Section 2.2). The radius represents the geographic boundary within which fan engagement exceeds background levels, rather than the full extent of the fan base, and is intentionally conservative. It is calculated as the distance at which normalized tweet activity falls to baseline, capturing only the densest region of above background engagement. Consequently, a low percentage of tweets within the radius does not indicate low overall fan activity, but rather reflects the broader spatial distribution of a team’s fans. In other words, teams with a low percentage of tweets inside the radius may have many distant fans whose activity contributes to baseline levels, rather than extending the radius itself.

Despite these differences in geographic concentration, the average sentiment of tweets posted inside a fandom region compared to the overall sentiment remains relatively similar. This suggests that location is not a major factor influencing sentiment variation in tweets. Fans tend to express similar tones when composing tweets regardless of where they are located. The differences between sentiment inside the fandom radius and overall sentiment are generally small across teams, indicating that the emotional tone of online discussions is relatively consistent across geographic regions.

Figure 8 compares the average sentiment of tweets for each NFL team within their fandom radius (top) with overall sentiment across all tweets (bottom). Sentiment scores for tweets inside fandom regions range from approximately 5.81 to 6.12, while overall sentiment across all tweets spans a slightly narrower range of roughly 5.86 to 6.05. In both groupings, the Chicago Bears consistently exhibit the lowest sentiment, whereas the highest sentiment is observed for the San Diego Chargers within fandom regions and the Seattle Seahawks across all tweets. The Spearman correlation between fandom region sentiment and overall sentiment is $\rho = 0.85$ ($p < 0.001$), indicating a strong positive relationship. The close similarity in sentiment patterns between the two plots indicates that, for most teams, the geographic location of the fan has little influence on the overall emotional tone of tweets.



(a) Sentiment inside fandom radius.

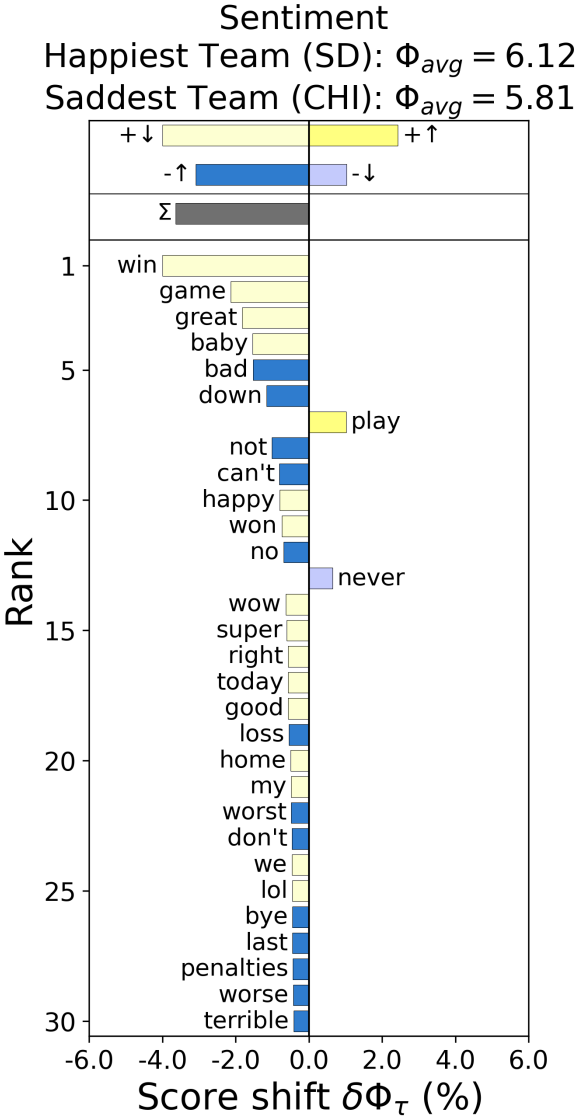


(b) Overall sentiment for all tweets.

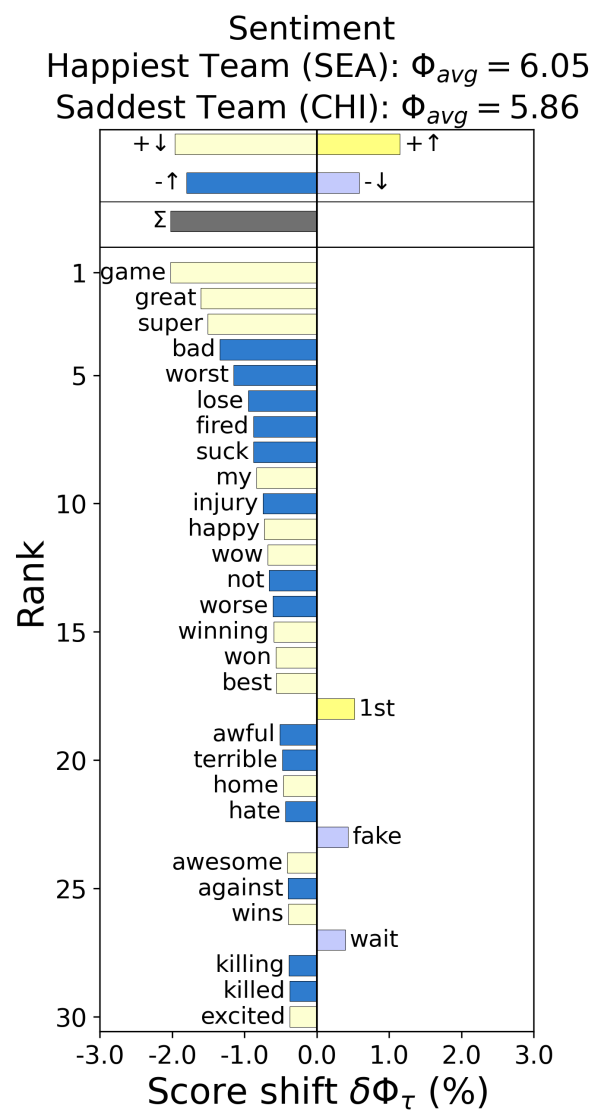
Figure 8: Comparison of sentiment scores for each NFL team. Each dot represents the average sentiment for a team, colored by sentiment score. Plot (a) shows tweets posted at a location inside the fandom radius and plot (b) shows all tweets mentioning the team, regardless of location.

Figure 9 shows two word-shift graphs comparing the language used by the happiest and saddest teams, both within their fandom radius and across all tweets[25]. The graph on the left compares tweets written inside the fandom radius for the San Diego Chargers and the Chicago Bears. Positive words such as “win,” “game,” and “great” appear less frequently in Bears tweets, while negative words such as “bad” and “down” appear more frequently, contributing to the lower sentiment score for Bears tweets. The interpretation of some words is less straightforward; for example, the word “down” may refer to a football down rather than negative sentiment depending on context. Interestingly, the word “play” appears more frequently in Bears tweets and less frequently in Chargers tweets, while the word “never” appears less frequently in Bears tweets and more frequently in Chargers tweets.

The graph on the right compares overall tweets for the Seattle Seahawks and the Chicago Bears. Bears tweets contain positive words such as “game,” “great,” and “super” less frequently than Seahawks tweets, which contributes to the lower overall sentiment score. Bears tweets also include more negative words such as “bad,” “worst,” “lose,” “fired,” and “suck,” further lowering their sentiment relative to Seahawks tweets. Overall, these word shifts provide insight into the differences in sentiment between fandoms by highlighting the relative frequency of positive and negative game related language.



(a) Word shift for Tweets written inside the fandom radius comparing the happiest versus saddest team.



(b) Word shift for all tweets posted comparing the happiest versus saddest team.

Figure 9: Comparison of word shifts for the happiest and saddest teams inside the fandom radius and across all tweets. Light yellow bars represent positive words used more frequently by the happiest team and less frequently by the saddest team. Dark blue bars represent negative words used more frequently by the saddest team and less frequently by the happiest team. Dark yellow bars represent positive words used more frequently by the saddest team, while light blue bars represent negative words used less frequently by the saddest team and more frequently by the happiest team.[25]

Another component of this analysis examined whether a team's record, measured by winning percentage, could help predict tweet sentiment. Figure 10 compares each team's

average sentiment score from 2011 to 2014 with their average win percentage over the same period. The scatter plot shows a weak positive correlation between winning percentage and sentiment, indicating that teams with higher win percentages tend to have slightly higher sentiment scores. A linear regression line illustrates this upward trend, and the Pearson correlation between win percentage and sentiment score is 0.33. This suggests that while winning more games is associated with more positive tweet sentiment, team performance alone does not fully explain sentiment differences across teams. Other factors, such as game context, geographic variation in fan bases, team culture, and fan expectations, may also influence how fans express sentiment on Twitter.

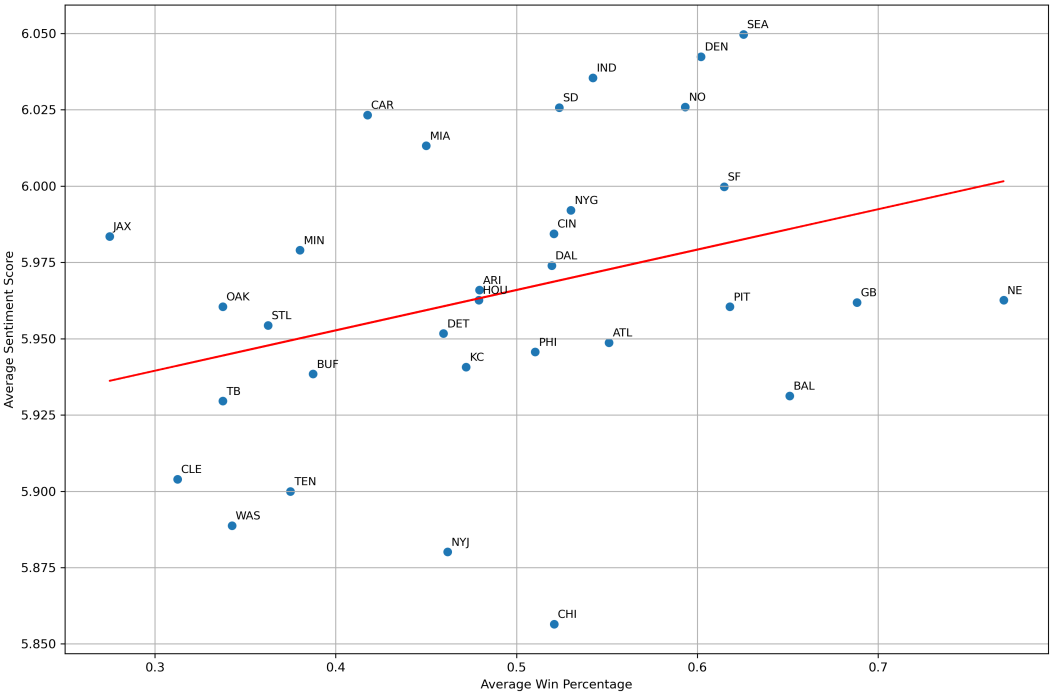


Figure 10: Relationship between average team win percentage (2011–2014) and average tweet sentiment. Each point represents an NFL team, with a regression line indicating the overall trend.

5 Conclusion

Using Twitter data to estimate fandom locations and analyze fan sentiment provides a new perspective on social media engagement among NFL fans. This analysis revealed

that fandom radii vary significantly across the league. Teams located in regions without nearby NFL teams tend to have larger fandom radii, while teams located in dense clusters of franchises often have smaller, overlapping radii. This research also provided insight into the sentiment of different fandoms and allowed for the quantification of relatively “happy” and “sad” fan bases. Sentiment varied across teams, and while the range of scores differed slightly depending on whether sentiment was calculated overall or only within the fandom radius, consistent groupings of happier, neutral, and less happy teams emerged.

Examining tweeting patterns during games also provided insight into how fans respond to game events. Tweet activity tended to spike shortly after kickoff and again toward the end of games, reflecting moments of heightened fan attention. Winning teams also generated slightly more tweets than losing teams, suggesting that fans of losing teams may be less inclined to share reactions online. Sentiment differences between winning and losing teams were especially visible near the end of games, which aligns with expectations that fans of winning teams are more likely to express positive reactions, often using celebratory language.

Social media platforms allow fans to react to games in real time regardless of their geographic location. The results of this research support the idea that fan emotion on Twitter is driven more strongly by in game events and final outcomes than by geographic location. Social media sentiment also has the potential to help measure fan engagement in real time. Geographic engagement patterns can help teams identify the difference between local and more widely distributed fan bases. Sentiment analysis can also provide insight into fan reactions to games, players, coaching decisions, and officiating decisions. These data could complement traditional measures of fan engagement such as television ratings and stadium attendance.

This study has several limitations. The analysis was limited to the seasons with the most available Twitter data (2011–2014) and only includes a 10% sample of tweets. Additionally, the analysis relied only on geolocated tweets, which represent a small subset of Twitter users. Many fans may not use Twitter to discuss sports or may not engage in real time

conversations during games due to factors such as age, internet access, or cultural preferences. The Hedonometer also evaluates words individually, which can miss contextual meaning such as sarcasm or football specific terminology. Furthermore, tweets were collected using specific game related hashtags, which excludes tweets that discuss games without hashtags, use alternative hashtags, or contain misspellings.

Future research could expand this work by examining how fans react to different types of games. One potential direction is to compare pregame betting odds and point spreads with tweet sentiment to determine whether fan sentiment reflects expectations about which team is likely to win. Additional game level analysis could also examine how sentiment changes in response to specific scoring events during games to better understand which moments generate the strongest emotional reactions among fans. Social media data could further be used to identify which game features, such as close scores, impressive plays, or star player performance drive the highest levels of attention and engagement.

6 Acknowledgments

I would like to thank my advisor, Chris Danforth, whose guidance, encouragement, and support were invaluable throughout every stage of this project. I am also grateful to Peter Dodds for serving on my thesis committee. Thank you to Michael Arnold, as this work would not have been possible without his coding and data collection assistance. Finally, I owe a special thanks to my mom, who first got me into American football and unknowingly sparked the idea for this research.

References

- [1] Kiran Khanna. Social Media in Sports Drives Powerful Fan Engagement. <https://aidastory.com/social-media-in-sports-drives-powerful-fan-engagement/>, March 2025.
- [2] Katie Young. Why the Second Screen in Sports is Social. <https://wearesocial.com/us/blog/2017/09/second-screen-sports-social/>, September 2017.
- [3] Ed Green-Moore. The Role of Social Media Interaction in Sports on X. <https://socialtrend.co.uk/latest-news/sports-industry-social-media-engagement-on-x/>, February 2024.
- [4] David Purdum. The rise of Gambling Twitter: Social media and the popularity of sports betting. https://www.espn.com/chalk/story/_/id/34539260/the-rise-gambling-twitter-new-social-media-data-shows-exploding-popularity-sports-betting, September 2022.
- [5] Thayer Alshaabi, Jane L. Adams, Michael V. Arnold, Joshua R. Minot, David R. Dewhurst, Andrew J. Reagan, Christopher M. Danforth, and Peter Sheridan Dodds. Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using twitter. *Science Advances*, 7(29), July 2021. ISSN 2375-2548. doi: 10.1126/sciadv.abe6534. URL <http://dx.doi.org/10.1126/sciadv.abe6534>.
- [6] Henry H. Wu, Ryan J. Gallagher, Thayer Alshaabi, Jane L. Adams, Joshua R. Minot, Michael V. Arnold, Brooke Foucault Welles, Randall Harp, Peter Sheridan Dodds, and Christopher M. Danforth. Say their names: Resurgence in the collective attention toward black victims of fatal police violence following the death of george floyd. *PLOS ONE*, 18(1):e0279225, 2023. ISSN 1932-6203. doi: 10.1371/journal.pone.0279225. URL <https://dx.plos.org/10.1371/journal.pone.0279225>.

- [7] Morgan Weaving, Thayer Alshaabi, Michael V. Arnold, Khandis Blake, Christopher M. Danforth, Peter S. Dodds, Nick Haslam, and Cordelia Fine. Twitter misogyny associated with Hillary Clinton increased throughout the 2016 U.S. election campaign. *Scientific Reports*, 13(1):5266, March 2023. ISSN 2045-2322. doi: 10.1038/s41598-023-31620-w. URL <https://www.nature.com/articles/s41598-023-31620-w>.
- [8] Mikaela Irene Fudolig, Thayer Alshaabi, Michael V. Arnold, Christopher M. Danforth, and Peter Sheridan Dodds. Sentiment and structure in word co-occurrence networks on Twitter. *Applied Network Science*, 7(1):9, December 2022. ISSN 2364-8228. doi: 10.1007/s41109-022-00446-2. URL <https://appliednetsci.springeropen.com/articles/10.1007/s41109-022-00446-2>.
- [9] Thayer Alshaabi, Michael V. Arnold, Joshua R. Minot, Jane Lydia Adams, David Rushing Dewhurst, Andrew J. Reagan, Roby Muhamad, Christopher M. Danforth, and Peter Sheridan Dodds. How the world’s collective attention is being paid to a pandemic: COVID-19 related n-gram time series for 24 languages on Twitter. *PLOS ONE*, 16(1):e0244476, January 2021. ISSN 1932-6203. doi: 10.1371/journal.pone.0244476. URL <https://dx.plos.org/10.1371/journal.pone.0244476>.
- [10] Anne Marie Stupinski, Thayer Alshaabi, Michael V Arnold, Jane Lydia Adams, Joshua R Minot, Matthew Price, Peter Sheridan Dodds, and Christopher M Danforth. Quantifying Changes in the Language Used Around Mental Health on Twitter Over 10 Years: Observational Study. *JMIR Ment Health*, 9(3):e33685, March 2022. ISSN 2368-7959. doi: 10.2196/33685. URL <http://www.ncbi.nlm.nih.gov/pubmed/35353049>.
- [11] Michael V. Arnold, David Rushing Dewhurst, Thayer Alshaabi, Joshua R. Minot, Jane L. Adams, Christopher M. Danforth, and Peter Sheridan Dodds. Hurricanes and hashtags: Characterizing online collective attention for natural disasters. *PLOS ONE*, 16(5):1–20, 05 2021. doi: 10.1371/journal.pone.0251762. URL <https://doi.org/10.1371/journal.pone.0251762>.

- [12] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, page 851–860, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605587998. doi: 10.1145/1772690.1772777. URL <https://doi.org/10.1145/1772690.1772777>.
- [13] Carmen-Sarah Einsle, Gregorio Escalera-Izquierdo, and Jerónimo García-Fernández. Social media hook sports events: a systematic review of engagement. *Communication & Society*, 36(3):133–151, 2023. ISSN 2386-7876. doi: 10.15581/003.36.3.133-151. URL <https://revistas.unav.edu/index.php/communication-and-society/article/view/43790>.
- [14] Younghan Lee and Sangwon Na. Fear of missing out: An antecedent of online fan engagement of sport teams' social media. *Communication & Sport*, 12(5):908–937, 2024. ISSN 2167-4795, 2167-4809. doi: 10.1177/21674795231174565. URL <https://journals.sagepub.com/doi/10.1177/21674795231174565>.
- [15] Lauren Reichart Smith, Ann Pegoraro, and Sally Ann Cruikshank. Tweet, retweet, favorite: The impact of twitter use on enjoyment and sports viewing. *Journal of Broadcasting & Electronic Media*, 63(1):94–110, 2019. ISSN 0883-8151, 1550-6878. doi: 10.1080/08838151.2019.1568805. URL <https://www.tandfonline.com/doi/full/10.1080/08838151.2019.1568805>.
- [16] Katherine R. N. Reifurth, Matthew J. Bernthal, Khalid Ballouli, and Dorothy Collins. Nonlocal fandom: Effects of geographic distance, geographic identity, and local competition on team identification. *Sport Marketing Quarterly*, 28(4):195–208, 2019. ISSN 1061-6934, 1557-2528. doi: 10.32731/SMQ.284.122019.02. URL <https://journals.sagepub.com/doi/10.32731/SMQ.284.122019.02>.
- [17] Flavio Chierichetti, Jon Kleinberg, Ravi Kumar, Mohammad Mahdian, and Sandeep

- Pandey. Event detection via communication pattern analysis. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):51–60, 2014. doi: 10.1609/icwsm.v8i1.14536. URL <https://ojs.aaai.org/index.php/ICWSM/article/view/14536>.
- [18] Kelsey Linnell, Michael Arnold, Thayer Alshaabi, Thomas McAndrew, Jeanie Lim, Peter Sheridan Dodds, and Christopher M. Danforth. The sleep loss insult of Spring Daylight Savings in the US is observable in Twitter activity. *Journal of Big Data*, 8(1):121, September 2021. ISSN 2196-1115. doi: 10.1186/s40537-021-00503-0. URL <https://doi.org/10.1186/s40537-021-00503-0>.
- [19] David Corney, Craig Martin, and Ayse Göker. Spot the ball: Detecting sports events on twitter. In Maarten de Rijke, Tom Kenter, Arjen P. de Vries, ChengXiang Zhai, Franciska de Jong, Kira Radinsky, and Katja Hofmann, editors, *Advances in Information Retrieval*, volume 8416 of *Lecture Notes in Computer Science*, pages 449–454. Springer, 2014. doi: 10.1007/978-3-319-06028-6_40.
- [20] Peter Sheridan Dodds, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PLoS ONE*, 6(12):e26752, 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0026752. URL <https://dx.plos.org/10.1371/journal.pone.0026752>.
- [21] Andrew J. Reagan, Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1):31, November 2016. ISSN 2193-1127. doi: 10.1140/epjds/s13688-016-0093-1. URL <https://doi.org/10.1140/epjds/s13688-016-0093-1>.
- [22] Alex Kennedy. How long is a football game? Average length of NFL games, NCAA games, and the Super Bowl. [https://www.profootballnetwork.com/how-long-is-a-](https://www.profootballnetwork.com/how-long-is-a)

football-game-breaking-down-the-time-between-the-first-and-last-whistle/, September 2025.

- [23] David Comberg. Kurt Vonnegut on the Shapes of Stories, October 2010. URL <https://www.youtube.com/watch?v=oP3c1h8v2ZQ>.
- [24] Dilan Patrick Kiley, Andrew J. Reagan, Lewis Mitchell, Christopher M. Danforth, and Peter Sheridan Dodds. Game story space of professional sports: Australian rules football. *Physical Review E*, 93(5):052314, May 2016. doi: 10.1103/PhysRevE.93.052314. URL <https://link.aps.org/doi/10.1103/PhysRevE.93.052314>.
- [25] Ryan J. Gallagher, Morgan R. Frank, Lewis Mitchell, Aaron J. Schwartz, Andrew J. Reagan, Christopher M. Danforth, and Peter Sheridan Dodds. Generalized word shift graphs: a method for visualizing and explaining pairwise comparisons between texts. *EPJ Data Science*, 10(1):4, December 2021. ISSN 2193-1127. doi: 10.1140/epjds/s13688-021-00260-3. URL <https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-021-00260-3>.

A Data Processing

Table A1: NFL teams, their home cities (team names rather than stadium locations), and associated hashtags used for tweet querying.

Team	Home City	Hashtags
Arizona Cardinals	Phoenix, AZ	#ArizonaCardinals, #Cardinals
Atlanta Falcons	Atlanta, GA	#AtlantaFalcons, #Falcons
Baltimore Ravens	Baltimore, MD	#BaltimoreRavens, #Ravens
Buffalo Bills	Buffalo, NY	#BuffaloBills, #Bills
Carolina Panthers	Charlotte, NC	#CarolinaPanthers, #Panthers
Chicago Bears	Chicago	#ChicagoBears, #Bears
Cincinnati Bengals	Cincinnati, OH	#CincinnatiBengals, #Bengals
Cleveland Browns	Cleveland, OH	#ClevelandBrowns, #Browns
Dallas Cowboys	Dallas, TX	#DallasCowboys, #Cowboys
Denver Broncos	Denver, CO	#DenverBroncos, #Broncos
Detroit Lions	Detroit, MI	#DetroitLions, #Lions
Green Bay Packers	Green Bay, WI	#GreenBayPackers, #Packers
Houston Texans	Houston, TX	#HoustonTexans, #Texans
Indianapolis Colts	Indianapolis, IN	#IndianapolisColts, #Colts
Jacksonville Jaguars	Jacksonville, FL	#JacksonvilleJaguars, #Jaguars
Kansas City Chiefs	Kansas City, MO	#KansasCityChiefs, #Chiefs
Miami Dolphins	Miami, FL	#MiamiDolphins, #Dolphins
Minnesota Vikings	Minneapolis, MN	#MinnesotaVikings, #Vikings
New England Patriots	Boston, MA	#NewEnglandPatriots, #Patriots
New Orleans Saints	New Orleans, LA	#NewOrleansSaints, #Saints
New York Giants	New York City, NY	#NewYorkGiants, #Giants
New York Jets	New York City, NY	#NewYorkJets, #Jets
Oakland Raiders	Oakland, CA	#OaklandRaiders, #Raiders
Philadelphia Eagles	Philadelphia, PA	#PhiladelphiaEagles, #Eagles
Pittsburgh Steelers	Pittsburgh, PA	#PittsburghSteelers, #Steelers
Seattle Seahawks	Seattle, WA	#SeattleSeahawks, #Seahawks
San Francisco 49ers	San Francisco, CA	#SanFrancisco49ers, #Niners
San Diego Chargers	San Diego, CA	#SanDiegoChargers, #Chargers
St. Louis Rams	St. Louis, MO	#StLouisRams, #Rams
Tampa Bay Buccaneers	Tampa, FL	#TampaBayBuccaneers, #Buccaneers
Tennessee Titans	Nashville, TN	#TennesseeTitans, #Titans
Washington Redskins	Washington, D.C.	#WashingtonRedskins, #Redskins

Table A2: Words Removed from Tweets to Prevent Bias in Sentiment Analysis

cowboys
redskins
cousins
washington
eagles
lynch
saints
bills
love
birds
fly
eagle
sea
miami
sun
beach
rivers
luck
golden
high
beast

B Additional Figures

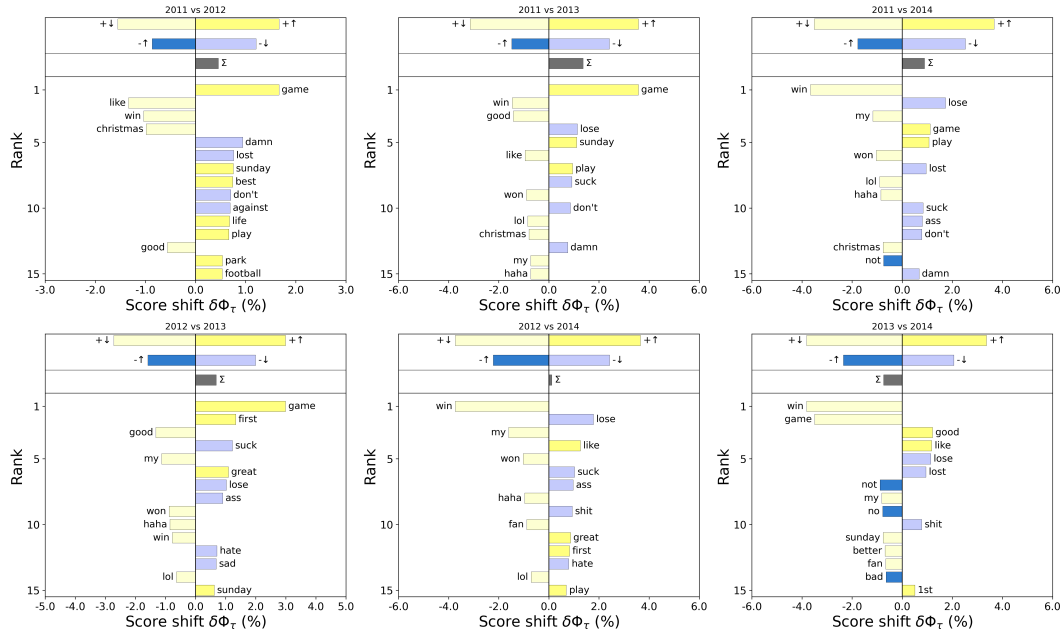


Figure B1: Word shift analysis for tweets over multiple years, illustrating changes in the most positive and negative words used by fans each season[25].

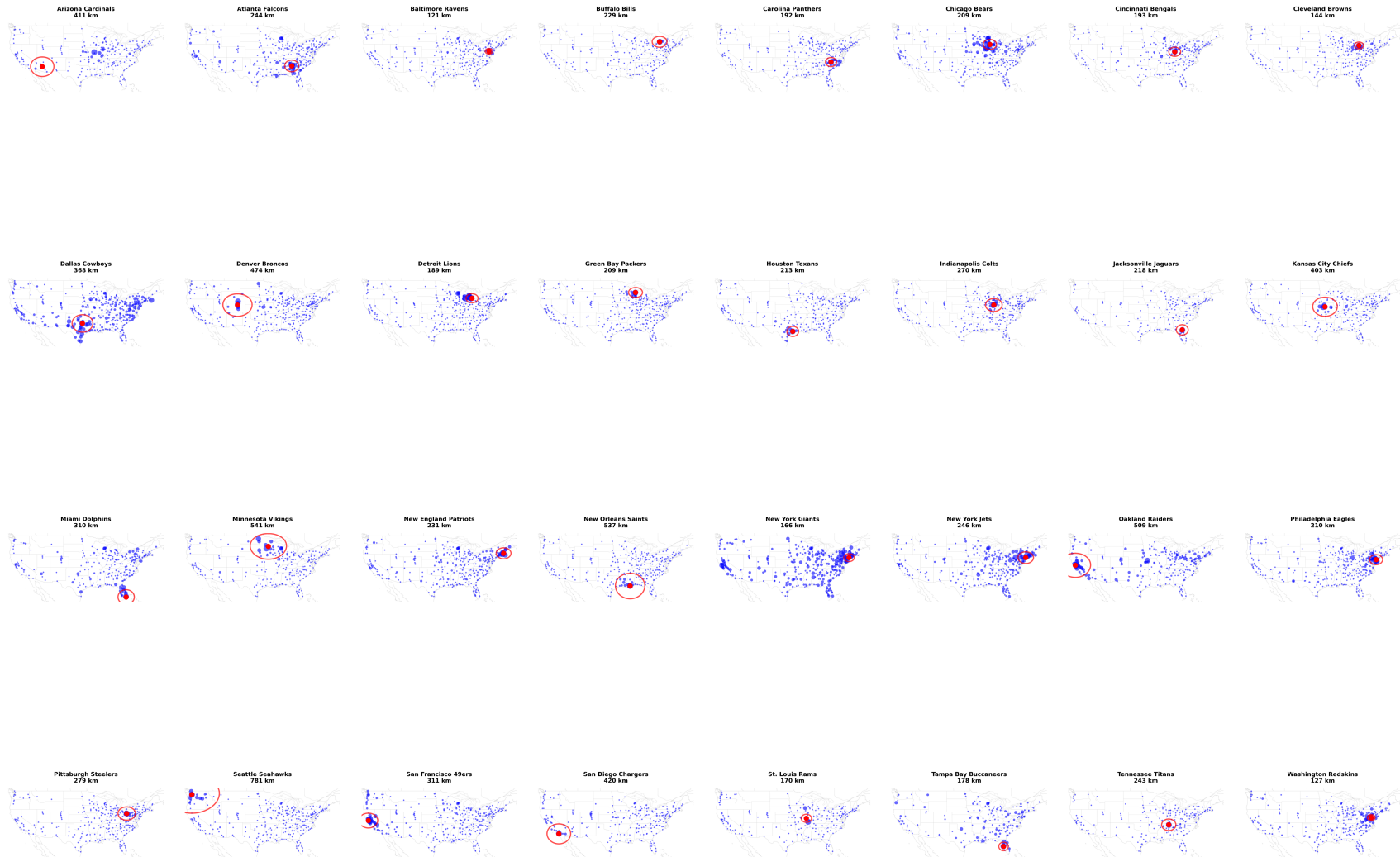


Figure B2: Estimated fandom radii and geographic tweet activity for all 30 NFL teams. Each team's home city is marked with a red dot and its fandom radius is shown as a red circle. Blue dots indicate tweet activity in metro areas, with larger dots corresponding to higher tweet volume per 100,000 residents.

Losing Teams: $\Phi_{avg} = 5.84$
 Winning Teams: $\Phi_{avg} = 6.00$

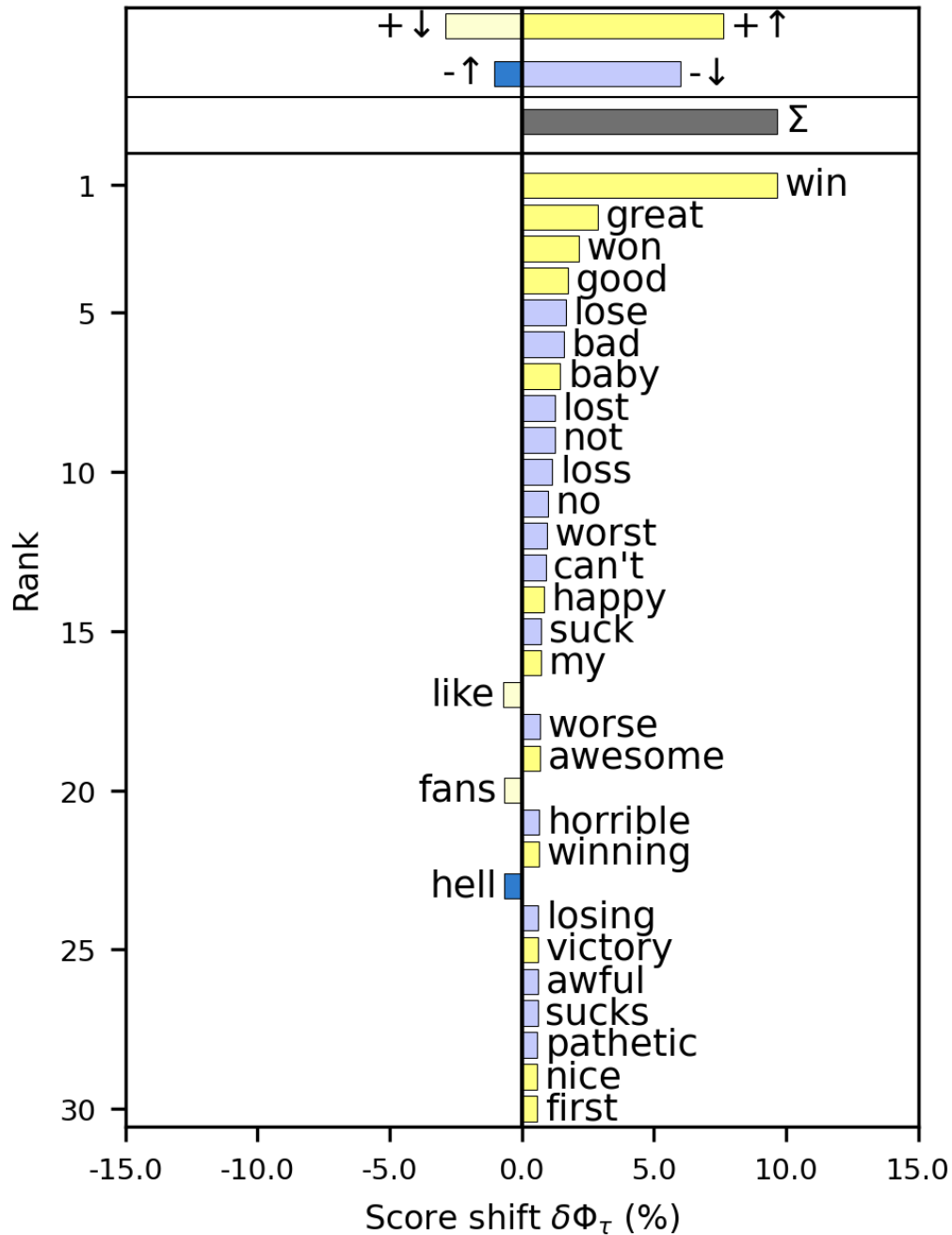


Figure B3: Word shifts for tweets posted from 3 hours before to 4 hours after an NFL game. Dark yellow bars represent positive words used more frequently by the winning team and less frequently by the losing team, light blue bars represent negative words used less frequently by the winning team and more frequently by the losing team, light yellow bars represent positive words used less frequently by the winning team, and dark blue bars represent negative words used more frequently by the winning team [25].