

KABIR CHATURVEDI

Semester: Spring 2024

Sentiment Analysis of Fan Discussions During the 2022 FIFA World Cup: Uncovering Global Emotional Trends (1346 words)

Introduction:

The 2022 FIFA World Cup, a stage for 32 nations, was not merely a football tournament; it was an emotionally charged global event that invoked a spectrum of sentiments from frustration to euphoria. This edition of the world cup was particularly riveting, climaxing in a final that pitted Lionel Messi against Kylian Mbappe, living up to and exceeding the hype it generated. Additionally, the tournament was marked by unexpected turns, most notably Morocco's remarkable journey, which saw them overcoming football giants such as Cristiano Ronaldo's Portugal and a formidable Spanish squad. This led to an increased volume of narratives and interactions on platforms like Reddit, where fans passionately discussed and debated these underdog triumphs. Building on the foundations laid by Wunderlich & Memmert (2021) and Gruettner et al. (2020), this paper investigates the sentiment trends within Reddit communities, providing a window into the collective emotional response to one of the world's most uniting sports events. This intrigue laid the groundwork for my research question.

Research Question:

How did these pivotal events of the 2022 FIFA World Cup influence the sentiment of global fans on social media?

Method:

Data:

Data collection for the study was meticulously conducted through manual curation from the 'r/worldcup' and 'r/soccer' subreddits, focusing on posts with highest user interaction in the form of upvotes from November to December 2022. A total of 80 posts and comments were chosen, 40 from each subreddit, beginning with those receiving the highest upvotes to capture community resonance. To ensure a broad perspective, comments were also selected using the 'controversial' filter in the reddit comment section. Targeted searches using keywords such as "excited," "WC 2022," "World Cup 2022," "Frustrated," and "2022". The keywords were not case-sensitive, as the data preprocessing stage involved normalizing the text to lower case to maintain consistency. The chosen dataset comprised a diverse emotional spectrum, capturing fan reactions before, during, and after games, embodying the wide array of sentiments across the tournament's timeline. This strategy facilitated a comprehensive dataset, capturing a wide spectrum of public sentiment and providing a holistic overview of the global emotional response throughout the World Cup narrative.

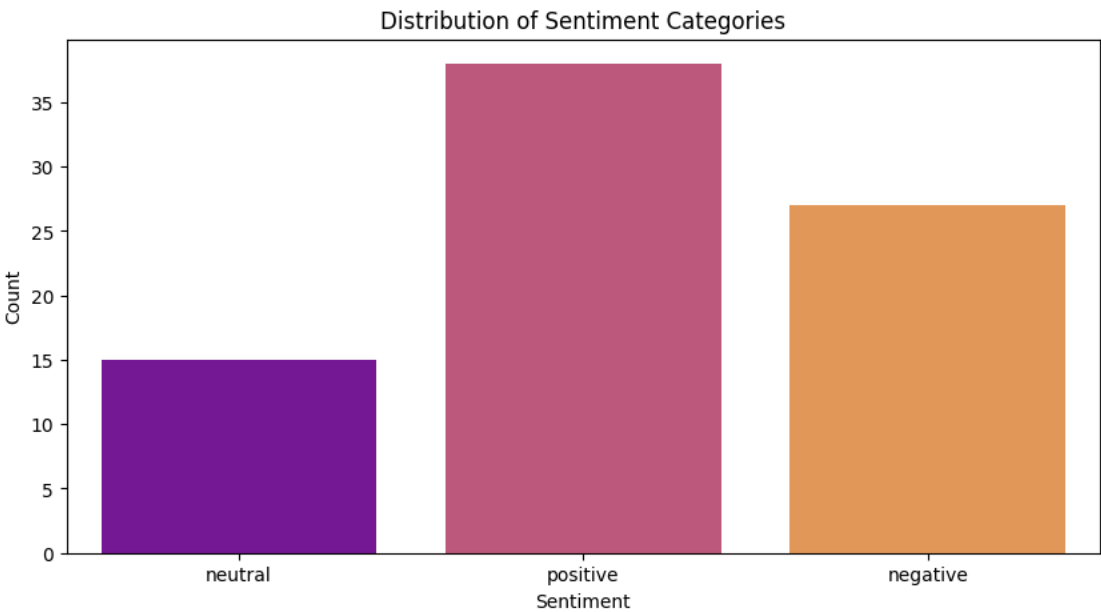
The dataset's temporal scope extended from the start of the 2022 FIFA World Cup on November 20th to the end of December 2022, approximately ten days after the tournament. This extended timeframe was purposefully chosen to capture not only the immediate reactions during the matches, but also the sentiments that emerged as fans processed and accepted the game's outcomes. By including data up to the

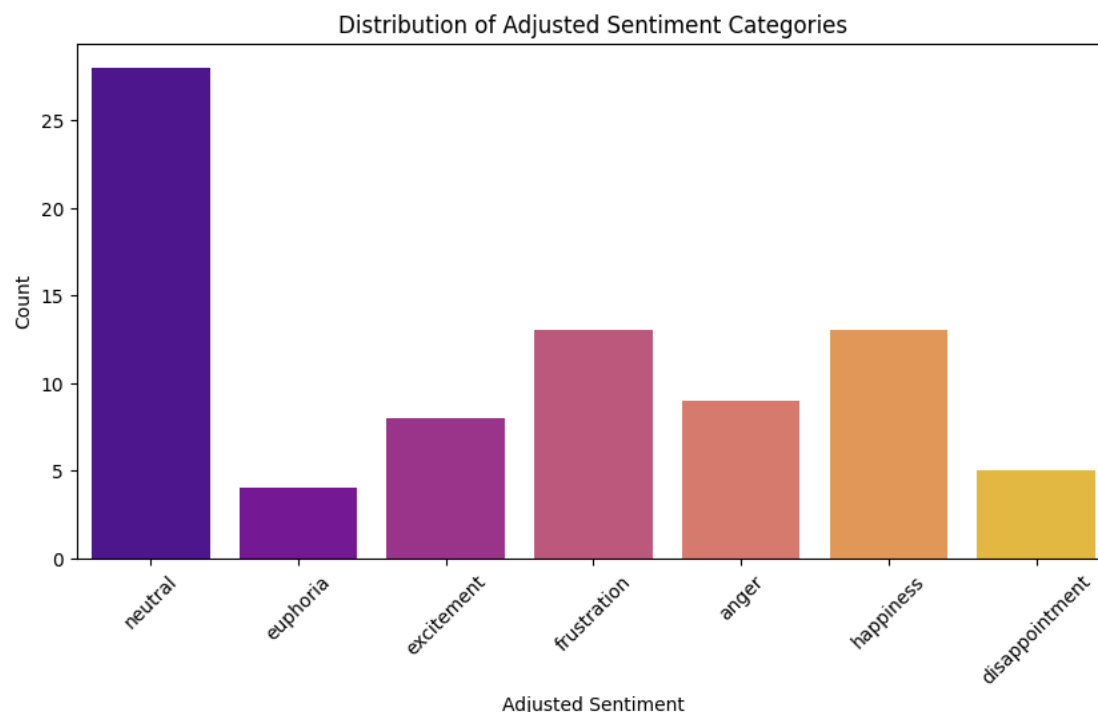
end of December, the study hopes to capture the transition from in-the-moment reactions to more contemplative post-event discussions, providing a comprehensive picture of the emotional arc experienced by fans during and after the World Cup.

Analysis:

This project harnessed the ‘TextBlob’ library's capabilities, which was selected for its ease in conducting sentiment analysis, providing immediate sentiment categorization suitable for processing the diverse emotional content from Reddit posts.

The analysis refined the understanding of sentiment by introducing 'sentiment' and 'adjusted_sentiment' categories. These captured a spectrum from basic sentiments—positive, negative, and neutral, included in the 'sentiment' column—to more nuanced emotions such as anger, disappointment, euphoria, excitement, frustration, happiness, and neutral, encompassed within the 'adjusted_sentiment' column. This dual categorization facilitates a comprehensive exploration of both general and specific emotional responses, enabling a detailed examination of fan reactions throughout the tournament. For example, 'euphoria' was classified as the utmost positive sentiment, marked by a significantly high positive polarity, for expressions of intense admiration, exemplified by "Greatest player of all time, what a goal and assist." Critiques of performance, such as "Lautaro was so poor the entire tournament" identified 'frustration', highlighting disappointment in player performance. Conversely, 'anger' represented the extreme negative sentiment, characterized by a notably negative polarity score, for vehement criticisms like "You should shut the hell up after that ball in the dugout. Miserable cheaters should have red twice.", reflecting intense negative sentiment. This detailed sentiment classification enabled an in-depth exploration of fan emotions, offering insights into the varying fan experiences and reactions throughout the tournament's key moments.





Results:

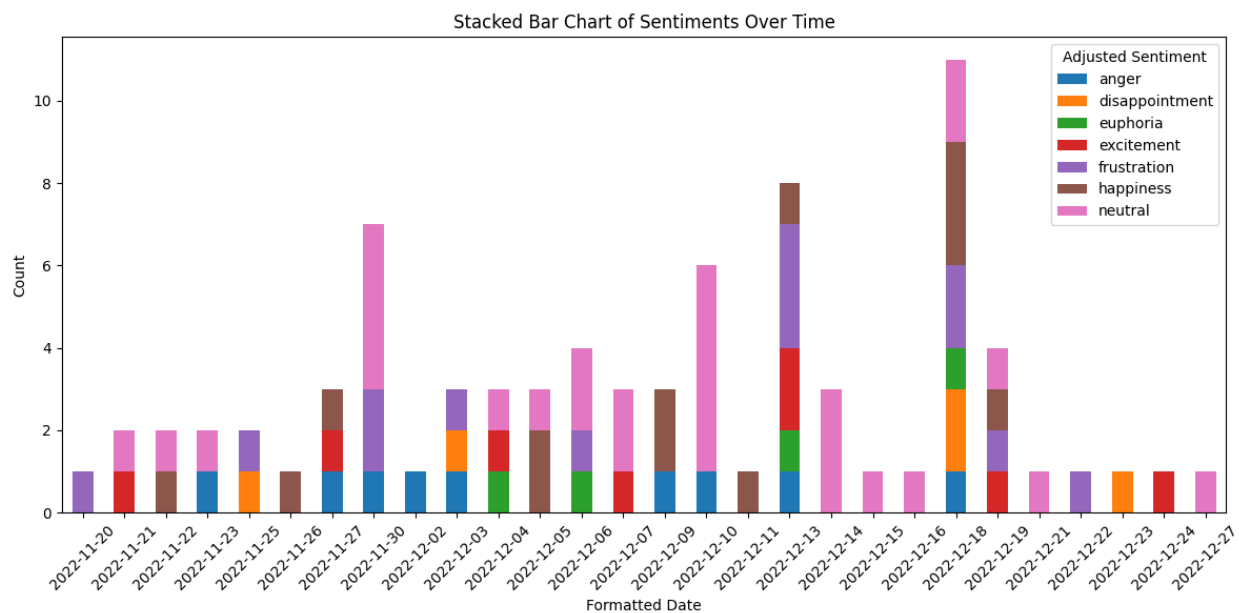
The sentiment analysis conducted around the 2022 FIFA World Cup delves deeply into the emotional responses shared on Reddit's r/worldcup and r/soccer communities. Examining post-game, pre-game, and in-game discussions revealed a clear pattern: the post-game period had the highest sentiment intensity, with over 61% of emotional expressions recorded. This suggested that the outcomes of matches served as powerful emotional triggers for fans.

The analysis also distinguished the emotional landscape of the discussions, revealing 'neutral' sentiments as the most common, implying that a significant number of discussions were balanced or lacked strong sentiment. Both 'frustration' and 'happiness' were equally prevalent, reflecting the contradictory nature of fan reactions to their teams' wins and losses. 'Anger' and 'excitement' trailed, highlighting the depth of passion fans held and underscoring the depth of passion fans held for the tournament's narrative.

	Sentiment	Count	Percentage
0	positive	38	47.50
1	negative	27	33.75
2	neutral	15	18.75

Adjusted Sentiment Percentage		
0	neutral	35.00
1	frustration	16.25
2	happiness	16.25
3	anger	11.25
4	excitement	10.00
5	disappointment	6.25
6	euphoria	5.00

Visualizations like the stacked bar chart tracked sentiment fluctuations over time, correlating peaks to specific match days, implying that game results have a direct impact on public sentiment. Bar charts provided a quantitative view of the prevalence of each emotion, with 'neutral' outnumbering others, implying that the majority of the commentary was unbiased or informative. The significant presence of 'happiness' and 'frustration' demonstrated the common duality of sports fandom, while the less frequent 'anger' and 'excitement' captured moments of disagreement and celebration.



In synthesizing these insights, the study painted a dynamic portrait of the World Cup's emotional narrative, offering a comprehensive understanding of fan sentiment. Moreover, the examination of the 'match related' column showed that emotions peaked post-game, highlighting that the most vigorous fan interactions occurred after the match conclusions. The prevalence of 'neutral' sentiments, followed by

'frustration' and 'happiness', and the least frequent 'euphoria' and 'disappointment', mapped out the full emotional spectrum experienced by fans during the tournament's pivotal moments.

Conclusion

The sentiment analysis for the 2022 FIFA World Cup revealed an intriguing divergence in emotional responses among fans of various national teams. Brazilian fans demonstrated a dramatic shift from initial elation, fueled by their team's commanding victories, to profound disappointment after their team was eliminated. This sentiment contrasted sharply with that of Portuguese fans, who spoke with a persistent tone of dissatisfaction, influenced by the team's perceived underperformance and Ronaldo's controversial exclusion. The emotional journey of Argentine fans was particularly compelling, tracing a path from disappointment following an early upset to a crescendo of jubilation as Argentina won the title.

The temporal sentiment trends revealed a significant predominance of post-match discussions, implying that fans are most likely to express their feelings in the immediate aftermath of matches. This trend is consistent with the literature on sports psychology, which states that immediate post-game periods are frequently characterized by heightened emotional expressions as fans process match outcomes. Pre-match discussions, on the other hand, tended to be a mix of optimism and anxiety, whereas in-game reactions captured the volatility of sentiments as fans reacted in real time to what was happening on the pitch.

The in-depth examination of these trends provides a granular understanding of the sentiment landscape during the World Cup, highlighting the diverse and dynamic nature of fan emotions in response to their teams' results. The comprehensive analysis of post-game discussions, in particular, demonstrates how sports events can elicit strong emotional responses and drive engagement within digital fan communities.

Limitations

The study's methodology, while rigorous, is subject to limitations inherent in manual data collection. Despite the researcher's best efforts to be representative, the selection of data points may inadvertently reflect a selection bias, potentially skewing the dataset towards specific sentiments or discussions. Furthermore, the non-randomized nature of the data aggregation process may limit the findings' applicability to the entire population of Reddit users or all discourses occurring during the 2022 FIFA World Cup period. Such limitations are important for readers to consider when interpreting the study's findings, as they highlight the contextual boundaries within which the conclusions are drawn.

References:

- Wunderlich, F., & Memmert, D. (2021, December 29). *A big data analysis of Twitter data during Premier League matches: Do tweets contain information valuable for in-play forecasting of goals in football? - social network analysis and Mining*. SpringerLink.
- Gruettner, A., Vitisvorakarn, M., Wambsganss, T., Rietsche, R., & Back, A. (2020a). *The New Window to Athletes' Soul – What Social Media Tells Us About Athletes' Performances*. scholarspace.manoa.hawaii.edu.
- Shah, P. (2020, November 6). *My absolute go-to for sentiment analysis-textblob*. Medium.
<https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524>

Barai, M. K. (2023, November 8). *Sentiment analysis with textblob and vader*. Analytics Vidhya.
<https://www.analyticsvidhya.com/blog/2021/10/sentiment-analysis-with-textblob-and-vader/>