

# Can Twitter Predict NFL Game Outcomes? A Social Media Sentiment Analysis

Amogh Ghadge, DS 4002

May 2, 2025

Github Repository: <https://github.com/amoghghadge/DS4002-CS3>

## Topic

This case study investigates the relationship between NFL fans' sentiment on Twitter and actual game outcomes. You'll analyze whether the collective emotional tone of fans before games correlates with—or even predicts—which teams win or lose.

## Context

Social media has become a digital pulse of public opinion, with millions of sports fans expressing their thoughts, expectations, and emotions about teams and games in real-time. These vast datasets of fan sentiment potentially contain hidden patterns and predictive signals that traditional sports analytics might miss. Major sports organizations and betting companies are increasingly turning to social sentiment as a complementary data source alongside traditional performance metrics.

## Motivation

Understanding the relationship between fan sentiment and game outcomes could revolutionize sports prediction models. If collective fan intuition contains predictive power, this could provide:

- A competitive edge for sports betting algorithms
- New insights for team marketing and engagement strategies
- A deeper understanding of the psychology behind fan communities
- Practical applications of natural language processing in a high-interest domain

## Your Mission

As a junior data scientist at SportsTech Analytics, you've been assigned to investigate this fascinating question using our proprietary dataset of NFL-related tweets from the 2024 season. Each tweet has been pre-labeled with sentiment scores using a machine learning model.

Your task is to design and execute a data science investigation that determines whether pre-game fan sentiment on Twitter has any statistically significant relationship with actual game outcomes. This will require preprocessing social media data, implementing proper statistical analysis, and creating compelling visualizations that tell the story in your data.

The final deliverable should be a complete GitHub repository containing your code, analysis, and a one-page executive summary answering the fundamental question: Can we predict NFL game outcomes by analyzing fan sentiment on social media?

This investigation will test your skills in natural language processing, statistical analysis, data visualization, and critical thinking about real-world data applications. Your findings could influence our company's future investment in social media prediction tools.