

# Quantifying Fanbase Reactions To Controversial Blockbuster Trades Across Different Major Sports Markets

Parker Gibbons

IPHS 200 Programming Humanity (Fall 2022) Prof Elkins and Chun, Kenyon College

## Abstract

This project uses sentiment analysis and statistics to quantify 'impactful' players in Major League Baseball. In game performance was contrasted with social media sentiment over the course of the 2022 season during which such trades occurred. The focus was on five MLB teams and major markets in which they played. The operating hypothesis is that each market, independent of size and money, react differently to players based on their culture and competitiveness.

## Introduction

As Yankees outfielder Joey Gallo patiently waited on being traded at the 2022 Major League Baseball Trade Deadline, he sat down for an interview reflecting upon his experience in New York. He described New York as 'a tough place to not play well,' and that 'It's hard to deal with' (The Athletic). Gallo's tenure with the Yankees could easily be identified as 'disappointing' and 'lackluster' by fans and these opinions were widely reflected not only during games but also on social media.

As social media becomes more prominent in everyday life, it is becoming more common for individuals to use social media in order to express their opinions on a particular player and their performance. With these opinions so readily available online, it is becoming easier for athletes to access these opinions which have the potential to contribute to a player's mental state and consequently their performance. This issue could be amplified or mitigated if a player switches teams during the season as they are now exposed to a different market and fanbase.

The players chosen for this project are regarded as having the potential to have a large impact on their respective teams. Thus transactions involving them are often presented as significant news around Major League Baseball. The players used include Joey Gallo, Juan Soto, and Frankie Montas. Gallo is a twenty-nine year old outfielder who spent the first part of the 2022 season on the New York Yankees before being traded to the Los Angeles Dodgers on August 2nd. Soto is a twenty-four year old outfielder who spent the first half of 2022 on the team who originally signed him, the Washington Nationals, before being traded on August 2nd to the San Diego Padres, where he spent the back half of the 2022 season. Montas is a twenty-nine year old pitcher who spent the first half of the 2022 season with the Oakland Athletics before being traded to the New York Yankees on August 2nd. All three of these players were regarded to be significant to their teams and were traded to a market that contrasted their previous market.

This project will also use the method of sentiment analysis to generate an overview of each player's public perception. Sentiment analysis has become one of the most prominent topics in Silicon Valley and within the tech industry for its ability to analyze thousands of lines of text in a matter of minutes while also presenting a new understanding of the text being analyzed. Sentiment analysis combines natural language processing with machine learning techniques to determine the polarity of each text. It is a method that is becoming more pervasive throughout the industry and thus being able to understand it is becoming increasingly important.

## Methodology

In order to find the sentiment of individual opinions about each player, I used python library snsrape to collect publicly available tweets. I then went through multiple steps in preprocessing including some which are unique to tweets: emojis, emojicons, and slang. This allows us to use two different sentiment analysis models. The first is VADER (Valence Aware Dictionary for Sentiment Reasoning) which is sensitive to both polarity and intensity (strength) of emotion. The second model used is TextBlob which is a Lexicon-based sentiment analyzer that uses scores to calculate a sentence's polarity. This allows me to gain a general understanding on how each player was perceived on social media during the 2022 season and how their perception changed after the transaction to their second team. I will choose the model that is the most accurate in reflecting the perception of the players over time and compare the selected model to each player's performance.

Each player's performance will be evaluated using one statistic that is able to encapsulate the multiple different statistics into one number. For hitters, the statistic used for evaluation will be OPS adds on-base percentage and slugging percentage to get one number that unites the two. It's meant to combine how well a hitter can reach base, with how well he can hit for average and for power (MLB). For pitchers, the statistic used for evaluation will be FIP (fielding independent pitching) which is similar to the statistic ERA (earned-run average) but focuses solely on the events a pitcher has the most control over by eliminating the results on balls hit into the field of play in which the pitcher does not have control over. (MLB). While not completely representative of a player's performance, these statistics allow for a general overview of a player's performance over time. These statistics for each player were gathered from Baseball Reference.

Using snsrape, I was able to scrape 30,000 of the most recent tweets involving 'Frankie Montas', and each 100,000 of the most recent tweets involving 'Joey Gallo', and 'Juan Soto'. For Montas, the tweets went back as far as May, 2020 while the tweets for Soto went back to July 22, 2022. For Gallo, the tweets went back to August 17th, 2021. After cleaning the raw tweets, I was able to run both sentiment analysis models on the compilation of tweets.

Additionally, I was able use Google Sheets to create line graphs for each player's respective statistic over time and lined these graphs up with each player's sentiment analysis over the same period of time. For Montas and Gallo's case, I only used compared sentiment with statistics during the 2022 season in order to keep the scope of the project within the 2022 season.



Figure 1

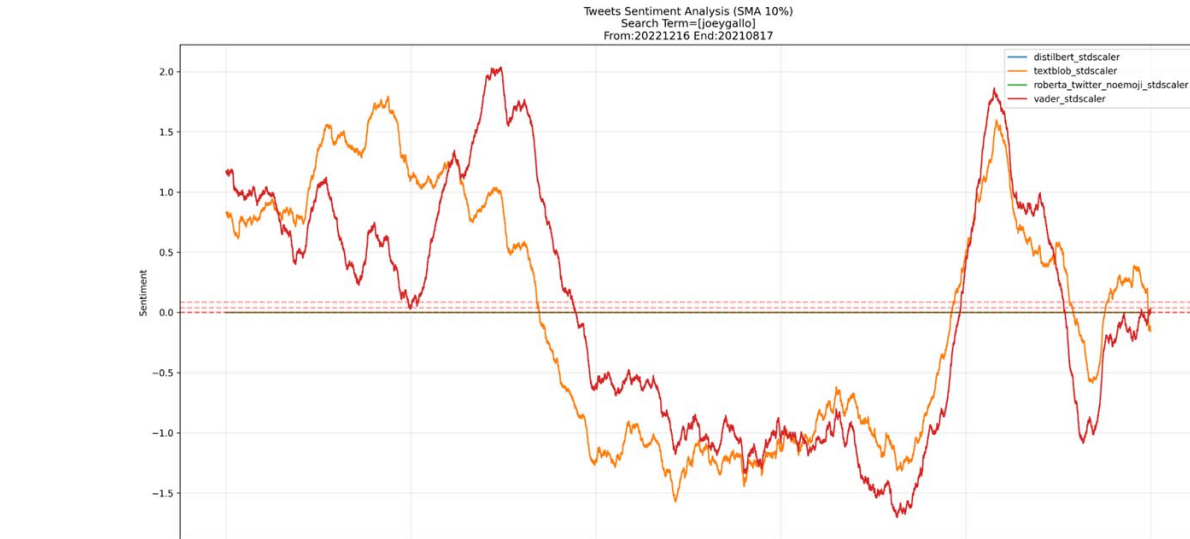


Figure 3

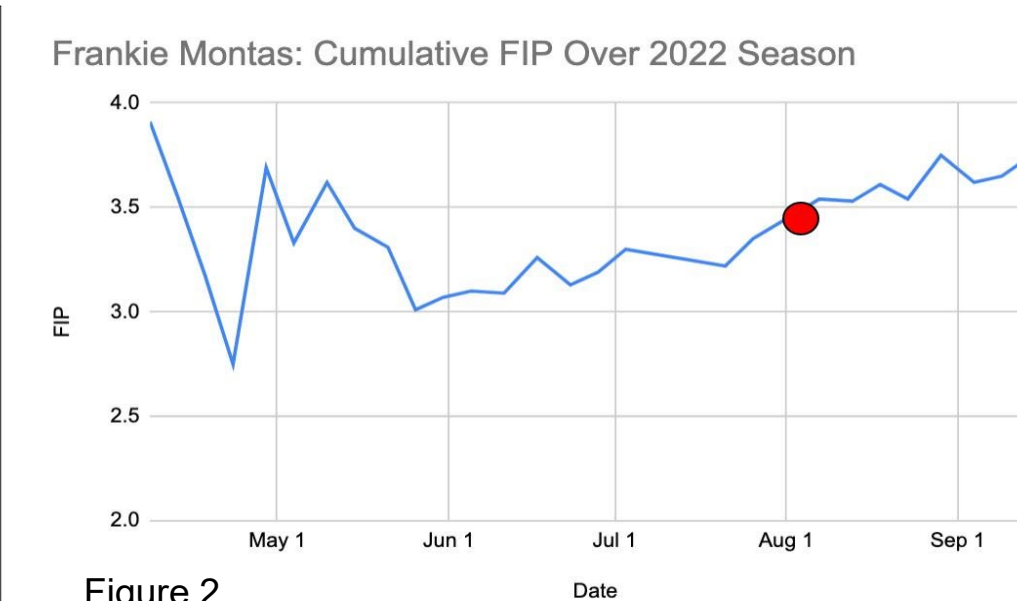


Figure 2



Figure 4

## Results

### Frankie Montas

When observing the sentiment of tweets regarding Frankie Montas (Figure 1), his time with the Oakland Athletics in 2022 (April to August) showed a fairly consistent sentiment ranging from -0.5 to 1.5 with a minimal standard deviation. However, once he was traded to New York in August, his sentiment began to rapidly decrease with an increasing standard deviation. This is correlated with his FIP over time. Despite their being a slight increase in FIP from June to August as seen in Figure 2, the rate of change increased subsequent to the trade to New York (for a pitcher, higher FIP is worse, lower is better). Thus there seems to be a general relationship between sentiment and FIP, as Montas' performance gets worse, so does the sentiment. However, it is clear from the sentiment analysis graph (Figure 1) that the sentiment after the trade to New York was much more reactive, showing a rate of change that was greater than it ever was during his time in Oakland. This shows New York as being a much more intemperate market than Oakland, as the sentiment after the trade exhibited much more drastic and frequent changes to a declining performance level that didn't differ much from the past.

### Joey Gallo

When observing the sentiment of tweets regarding Joey Gallo (Figure 3), his time with the Yankees in 2022 (April to August) shows a consistent negative sentiment ranging from -0.5 to -1.75 with a minimal standard deviation. Upon being traded to the Los Angeles Dodgers, the sentiment saw a large increase with a very high rate of change to a positive sentiment for a brief period of time before falling back to a negative sentiment followed by hovering around a sentiment of zero. This contrasts with Gallo's performance as shown in figure 4 because despite the large changes in sentiment, Gallo's performance remained consistently poor over the same period of time. However, it seems that the trade shifted the sentiment for a period of time, even to being positive at a point, before lowering again. This could have resulted from the slight increase in performance immediately following the trade which could have swayed individual's perceptions. However, it seems like a realization was made as the sentiment decreased shortly after the streak of improved performance.

### Juan Soto

When observing the sentiment of tweets regarding Juan Soto (Figure 5), his time with the Nationals in 2022 (April to August) shows a mostly consistent positive sentiment ranging from 0 to 1.5 with a minimal standard deviation. Upon being traded to the San Diego Padres, the sentiment saw a very large increase in sentiment with a high rate of change before peaking at nearly 2.5 before sharply declining to a negative sentiment consistently for the rest of the season. This contrasts from Soto's performance as shown in figure 6 because despite the large changes in sentiment, Soto's performance saw a slight negative change to no change at all compared to the time he spent on the Nationals. It seems that the trade caused a large positive sentiment which was quickly dissipated upon Soto's constant or slightly negative performance. This could suggest that individuals were excited over the trade and formulated expectations for Soto's performance based on his past success and these same individuals saw their expectations let down with Soto's subsequent performance. These individuals most likely expected Soto to perform at a higher level and were disappointed by his constant results in 2022.

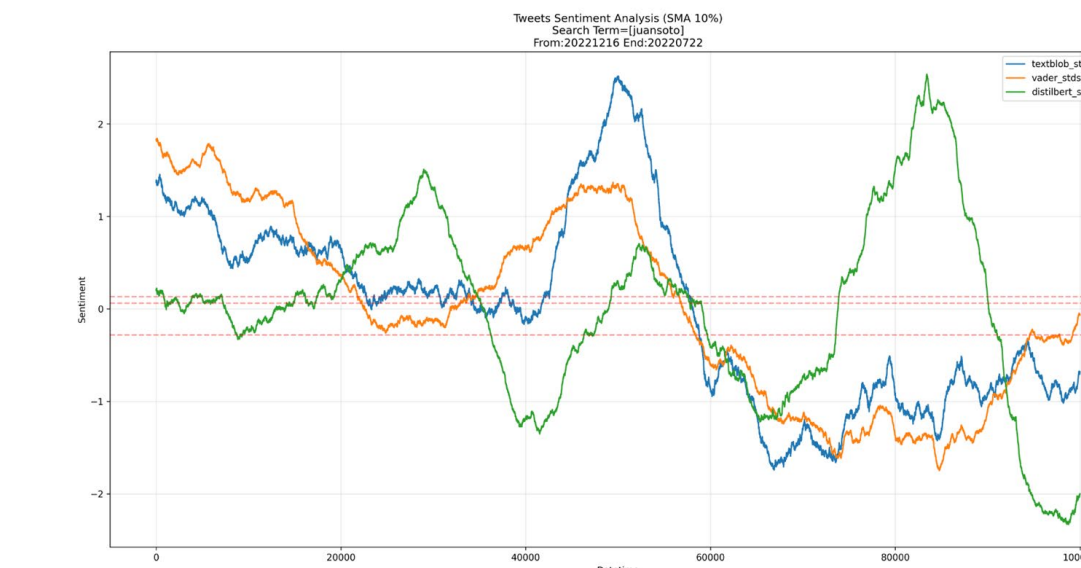


Figure 5

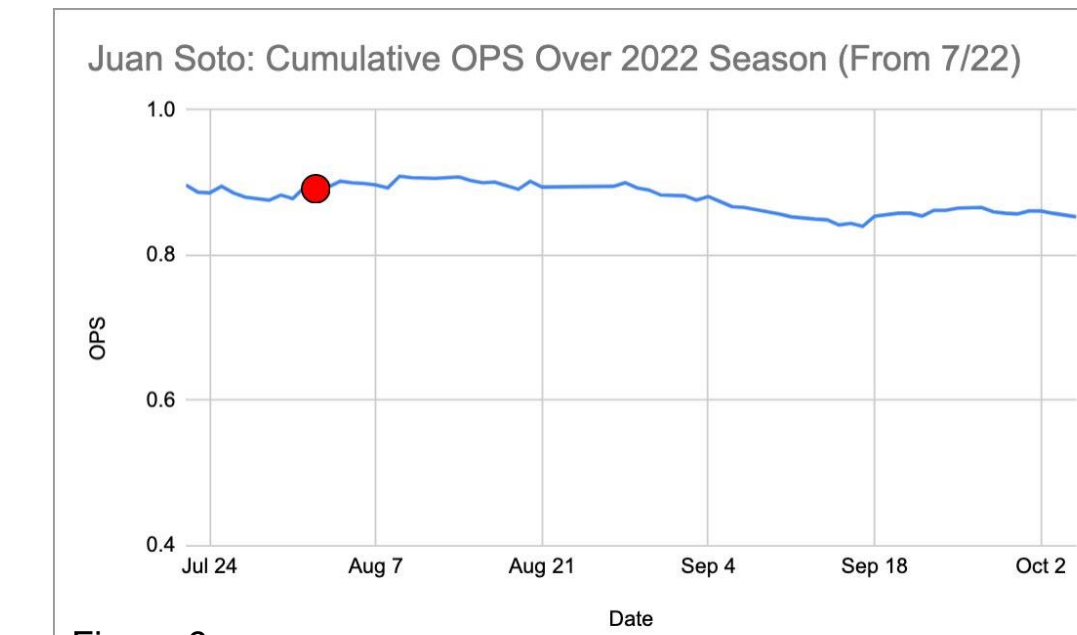


Figure 6

## Conclusion

I was also to use the sentiment analysis and statistics to correlate perception on social media to on-field performance. However, I was able to see the magnitude of this correlation change between different teams and their markets. I saw that in smaller markets like Oakland and Washington, performance whether good or bad, did not have large impacts on social media perception. However, bigger markets like New York, San Diego, and Los Angeles reacted more harshly to good and bad performance, as reflected with increased rates of change in terms of sentiment. This allowed me to deduce that different markets react differently to a player even if the performance trends are generally the same. Slight changes in performance could lead to large changes in sentiment in bigger markets while an even larger change in performance could cause a small or steady change in sentiment.

## Future and Ethics Statement

Prior to doing this analysis, my expectation for the work was that relationships between the sentiment of tweets and performance would ultimately be limited by the many external factors that impact a player's performance such as health, personal issues, and mental state. In addition, the sentiment analysis isn't entirely correct (70-90% accurate) due to contextual differences which could play a role in the analysis of a player's perception. In the future, I would extend this project by using a linear regression model to find the statistics that best evaluate a pitcher and hitter's performance and thus compare the model with the sentiment. I would also want to use more players in the sample to make more accurate conclusions. I would also want to use more sentiment analysis models, including Twitter RoBERTa, in order to obtain a more accurate sentiment analysis. One ethical drawback to this analysis could be the idea of analyzing a player's perception solely on Twitter which scope could be too minimal. It's obviously not possible to know what is going on in the life of all of these players, so the data will always be limited by not knowing all the factors that play into each player's performance and perception.

## References/Acknowledgements

- Adler, L. (2022, July 29). Joey Gallo on his time with yankees: 'I didn't live up to expectations'. The Athletic. Retrieved December 16, 2022, from <https://theathletic.com/3460126/2022/07/29/joey-gallo-yankees-expectations/>
- Earned run average (ERA): Glossary. MLB.com. (n.d.). Retrieved December 16, 2022, from <https://www.mlb.com/glossary/standard-stats/earned-run-average>
- Frankie Montas 2022 pitching game logs. Baseball. (n.d.). Retrieved December 16, 2022, from <https://www.baseball-reference.com/players/g/fcgr?id=montaf02&amp;t=p&year=2022>
- Joey Gallo 2022 batting game logs. Baseball. (n.d.). Retrieved December 16, 2022, from <https://www.baseball-reference.com/players/g/fcgr?id=gallajo01&am;p=t-b&year=2022>
- On-base plus slugging (OPS): Glossary. MLB.com. (n.d.). Retrieved December 16, 2022, from <https://www.mlb.com/glossary/standard-stats/on-base-plus-slugging>