Court of Public Opinion: A Sentiment Analysis

Reddit's collective reaction to Freddie Kitchens, Kevin Stefanski, and other recent NFL head coaching hires

Clay Girdner

2/14/2020

Hand up, this time last year I thought Freddie Kitchens was a good hire for the Cleveland Browns. The strong offensive finish to 2018 combined with Freddie's seemingly close ties to Baker Mayfield had me convinced that Kitchens was the man for the job. I wasn't about to let the fact that Freddie lacked any head coaching experience come between me and my delusions of Browns grandeur. I was hooked on recency bias, and it was a hell of a drug.

Fast forward to the present day. Freddie's a goner, the Browns are once again riding the coaching carousel, and my hope-fueled dopamine high has long since faded. However, being so completely wrong left this armchair G.M.'s ego bruised and battered. I had to wonder if everyone bought into the hype or if I was among a gullible minority that fell for Freddie's good ole boy schtick? To answer this question, I set out to measure the consensus fan opinion of the Kitchens pick and then compare it to that of Kevin Stefanski and other recent NFL head coach selections.
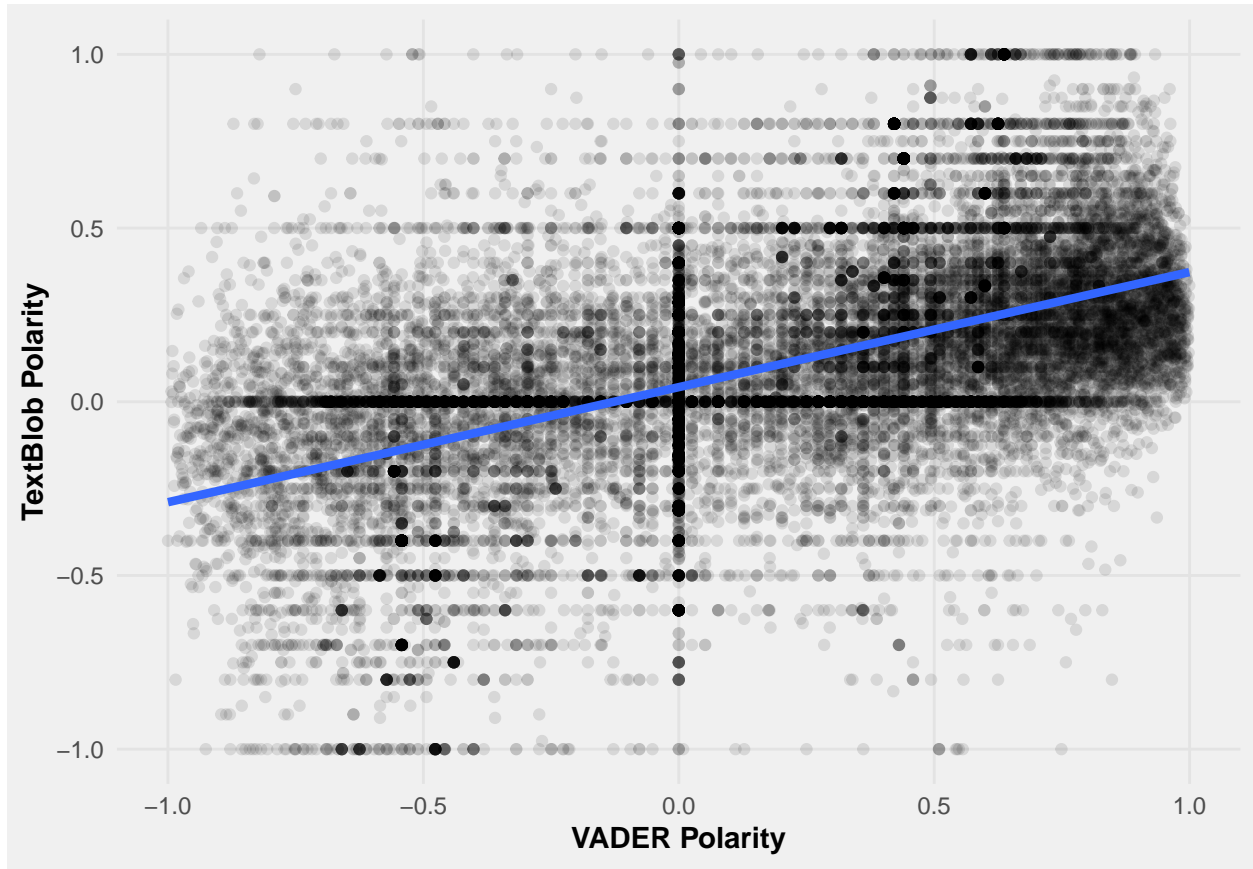
Data

I needed data to conduct this experiment, so I turned to Reddit threads – specifically posts on the NFL subreddit, r/nfl, and team-specific subreddits (i.e. r/browns for Kitchens and Stefanski, r/NYGiants for Joe Judge, etc.) linking to a Tweet or team release announcing a new hiring going back to 2018 (19 hires and 38 threads).

Python was used to gather the comments from each thread along with the PRAW library which allows users to tap into Reddit's free API. See reddit_hire_data_pull.py and comments_data.csv in the repository for the code and data - https://github.com/ClayGirdner/nfl_hires_reddit_sentiment.

In order to measure the overall sentiment for each thread, comments were fed through two different text analysis tools available in Python – VADER and TextBlob, both of which produce numeric scores indicating the degree of polarity for each comment (i.e. level of positivity/negativity). Two example comments and their resulting scores are provided below along with a scatter plot displaying the relationship between the two polarity scores for all 31,000+ comments in the data set (correlation = 0.51).

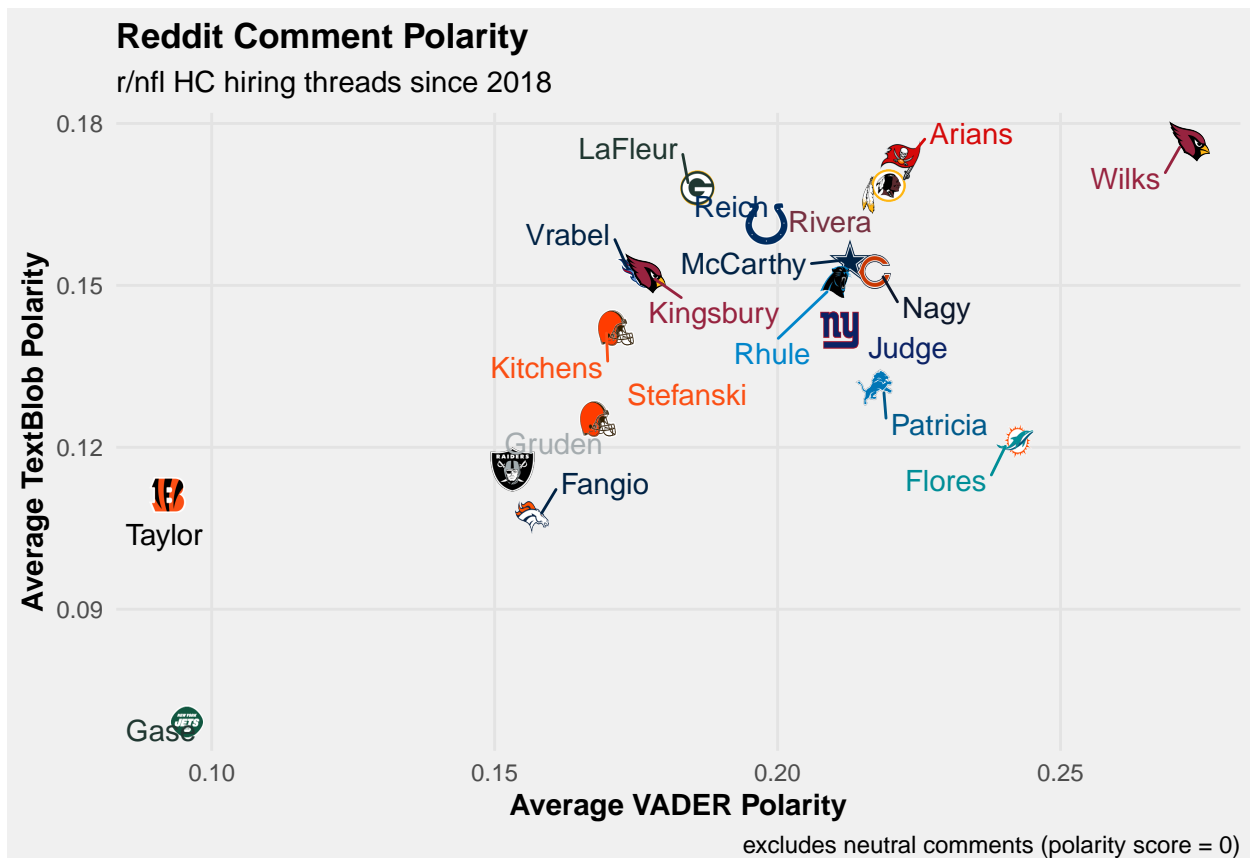(Note that all plots/calculations/analyses beyond this point will be done in R rather than Python.)

```
##                                  Comment  TextBlob   VADER
## 0     I really like this hire for our team!     0.25  0.4740
## 1  This guy stinks! Why would we hire him?    -0.75 -0.3164
```

We see that the polarity metrics have a tendency to concentrate around zero as both methods default to zero if the polarity cannot be detected. If we exclude these neutral observations (polarity = 0 in either method), the correlation improves to 0.54. Due to this quirk in the data, neutral comments will be excluded when calculating the polarity aggregates for each thread in the subsequent section. However, note that these neutral comments **are** included in the subjectivity section presented later as they provide context on the overall intensity of emotion within comment threads.
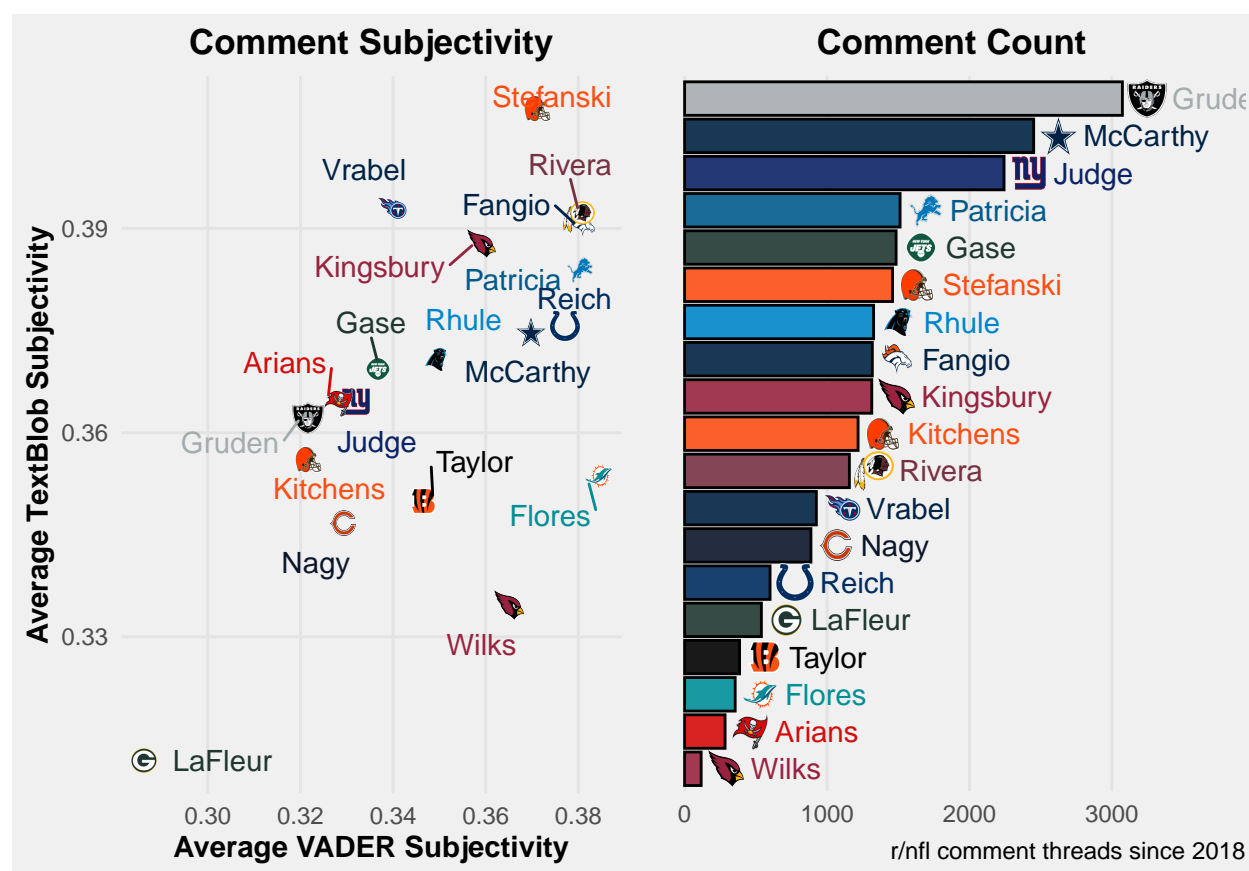
Polarity

With these polarity metrics in hand, scores were averaged for each hire thread on r/nfl since 2018. Results of these averaged polarity scores are presented below.

Reddit Comment Polarity

r/nfl HC hiring threads since 2018

Average TextBlob Polarity vs. Average VADER Polarity

excludes neutral comments (polarity score = 0)

In the plot we see that Kitchens scored near the middle of the pack in terms of VADER and TextBlob polarity while Stefanski just trailed in both variables. Surprisingly enough, Steve Wilks, the one-year coach of Arizona in 2018 scored the highest in both metrics, while Adam Gase scored last and next-to-last (unsurprisingly enough). However, if we look at comment counts and subjectivity metrics (intensity of emotion, in either direction), Wilks's high polarity scores begin to lose some of their luster and the Kitchens-Stefanski divide becomes more pronounced.
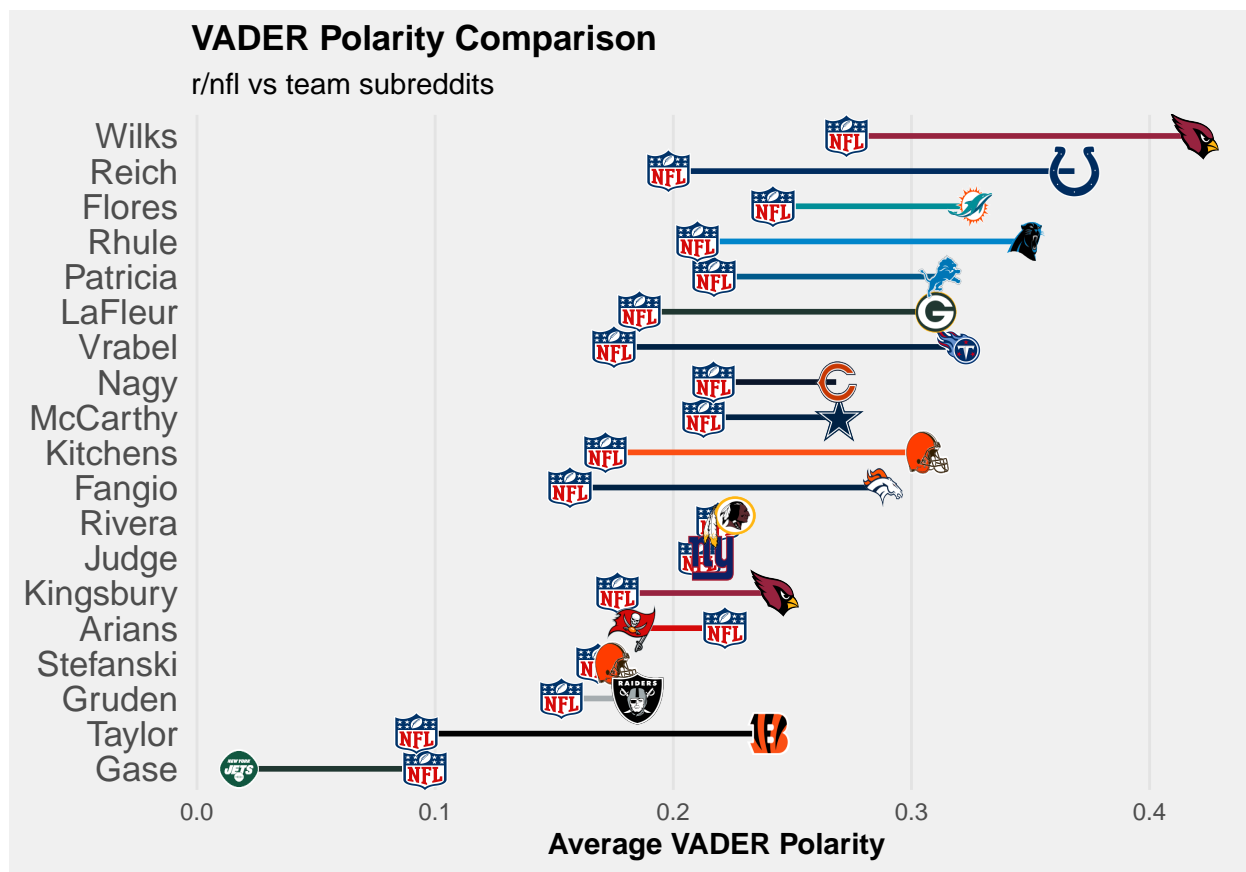
Subjectivity and Comment Counts

The Wilks thread was the smallest in terms of total comments (116) and one of the least subjective threads (at least according to TextBlob). I interpret this to mean that fans did not feel strongly about the hire and that his high polarity scores could be fluky due to the small sample size.

We also see that the Stefanski r/nfl hire thread was among the most opinionated discussions in the entire experiment and well ahead of the Kitchens thread which scored low in both subjectivity metrics. This suggests that fans are much more divided and emotional when it comes to Stefanski with only 23% of the Stefanski comments being neutral (according to VADER) compared to 34% of comments in the Kitchens thread.

Subreddit Comparisons

In addition to r/nfl, I also examined the hire threads within each team's respective subreddit to compare the sentiment of team-specific fans to the general NFL fanbase within r/nfl. Here I compare VADER polarity scores between the two subreddit groups.

VADER Polarity Comparison
r/nfl vs team subreddits

As intuition would suggest, this analysis shows that team subreddits tend to be more optimistic about "their guy" than the generalists commenting on r/nfl. This is especially true regarding the r/browns thread for Kitchens which had a VADER polarity score well above its r/nfl counterpart, proving that I wasn't the only Browns fan drinking the Kool-Aid. However, the same cannot be said for the Stefanski thread on r/browns with an average polarity nearly equal to r/nfl, suggesting that Browns fans are returning to their usual pessimistic demeanor. The last point to note from this plot is the r/nyjets reaction to the Adam Gase selection. Jets fans were not only more pessimistic than the community at large, but they generated the lowest VADER polarity score by a sizable margin. (Note that r/eagles was not represented in this analysis.)

Word Composition: Kitchens vs Stefanski

Lastly, I broke out the Kitchens and Stefanski discussions in order to compare the word composition between the two sets of comments. Here are wordclouds showing some of the most frequently used words from the Kitchens (orange) and Stefanski (brown) threads.

The word "baker" clearly sticks out when viewing the comment threads this way as it was heavily featured in the Kitchens discussion but barely mentioned in the Stefanski thread. In fact, Baker was mentioned nearly seven times more often when Freddie was hired! On the flip side, the word offense (or offensive) was mentioned more than twice as often in the Stefanski thread. I believe this is consistent with the narratives that Freddie was hired primarily due to his relationship with Mayfield and the other Browns players (word "players" 1.5x more often for Freddie) while Stefanski was hired for his ability to scheme up an offense and call plays.

Conclusion

In summary, this analysis shows that my affection for Kitchens wasn't completely outside the norm at the time of his hiring. However, it must also be noted that much of the Freddie love was concentrated within the Browns fanbase as Freddie's polarity scores were roughly average within the r/nfl thread but sky high in the r/browns comments, supporting the idea that team subreddits are often echo-chambers. Who knows, perhaps this just goes to show that the consensus on Reddit may not be the most reliable predictor of future success (shocker, I know). And if that's the case, it may be a good thing that the dummies on Reddit are talking negatively about my boy Stefanski! Oh boy, here I go getting my hopes up again...