**Analyzing The College Football Playoff Selections Based On Twitter Sentiment**

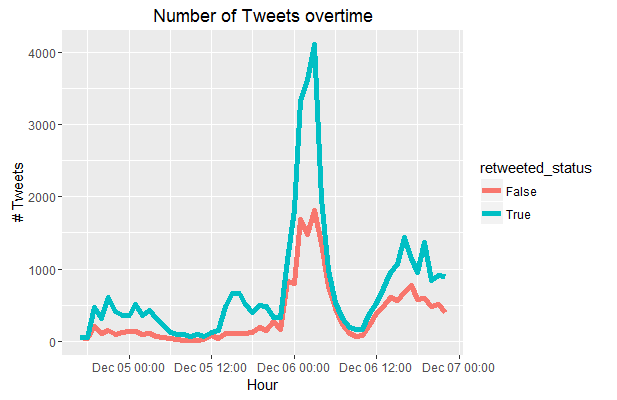
|  |  |  |
| --- | --- | --- |
|  | Table 1 |  |
| Rank | **Team** | **Handle/Hashtag** |
| 1 | Alabama | @AlabamaFTBL |
| 2 | Clemson | @ClemsonFB |
| #ALLIN |
| 3 | Ohio State | @OhioStateFB |
| #GoBucks |
| 4 | Washington | @UW\_Football |
| #PurpleReign |
| 5 | Penn State | @PennStateFball |
| #PSUnrivaled |
| 6 | Michigan | @UMichFootball |
| #GoBlue |
| 7 | Oklahoma | @OU\_Football |
| 8 | Wisconsin | @BadgerFootball |
| #OnWisconsin |

On December 4 the College Football Playoff Committee met for the last time during the 2016 football season. On that day they announced that the four teams that would be vying for chance at the title would be; Alabama, Clemson, Ohio State, and Washington. These four teams were ranked 1-4 respectively. But the question is… *Did the Committee get it right?* In this analysis I will try to either confirm or deny the committees picks based on the sentiment of tweets related to these teams in the following couple of days after the announcement. This analysis could be helpful to the Committee as well as sports enthusiasts for years to come to either increase or decrease the validity of the system that is now in its 3rd year.

Beyond the top four teams I also gathered tweets related to the teams ranked 4-8. To gather the data, I used the twitter REST API a couple of days after the announcement and searched for tweets that contained either the teams twitter handle or the hash tag that was identified within their descriptions (if one was identified), like #ALLIN for Clemson. You can see a list of all the terms and teams that were searched for in table 1. The API search returned about 5000 tweets for each of the search terms. Once all of the tweets were in a tabular form the value for sentiment was derived using NLTK’s VADER algorithm, which is a program that reads in the text and gives its best judgment if the tweet was positive, neutral, or negative.

|  |  |  |
| --- | --- | --- |
|  | **Table 2** | |
| **Team** | **Avg. Sentiment** | |
| **With Retweets** | **Without Retweets** |
| **Alabama** | 0.50 | 0.18 |
| **Clemson** | 0.11 | 0.06 |
| **Ohio State** | 0.20 | 0.12 |
| **Washington** | 0.12 | 0.04 |
| **Penn State** | 0.24 | 0.16 |
| **Michigan** | 0.39 | 0.21 |
| **Oklahoma** | 0.15 | 0.20 |
| **Wisconsin** | 0.34 | 0.17 |

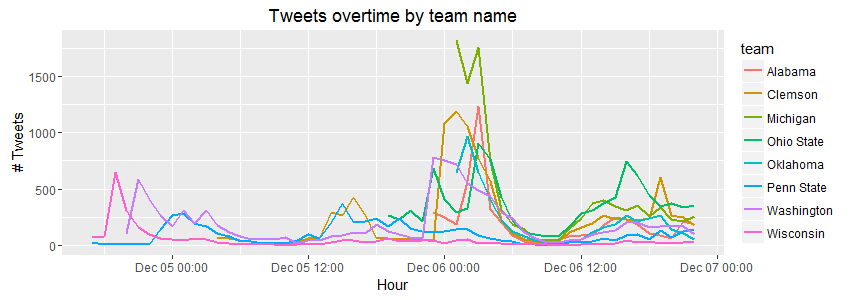
|  |
| --- |
| **Figure 1** |

 You can see in Figure 1 that there is a large amount of retweeted statuses (represented with the light blue line) over the course of time that the data was collected. Because of this there is a possibility that the overall sentiment of a given fan base could adversely or favorably affected. This could be a problem to some but typically a person will retweet a tweet because they agree with what was already said. Because of this some of the charts bellow will exclude retweets from there data. In Table 2 we can see the overall average sentiment of the tweets related to each team. It is obvious to see that far and beyond in the data that includes retweets that Alabama is in the rightful place of a #1 seed and the fans agree. They also rank highly in sentiment in the data not including retweets. The two teams that show up as interesting are Clemson and Washington both of these teams (which both made the playoffs) returned a sentiment of the two lowest as compared to all eight teams sampled, in both categories. This could mean that the fans at a whole tend to disagree that these two teams should be in the playoffs. What then also seems to be of interests is the teams that just missed the playoffs, for these teams one would expect their sentiment to be more on the low side as fans share there disgust with the

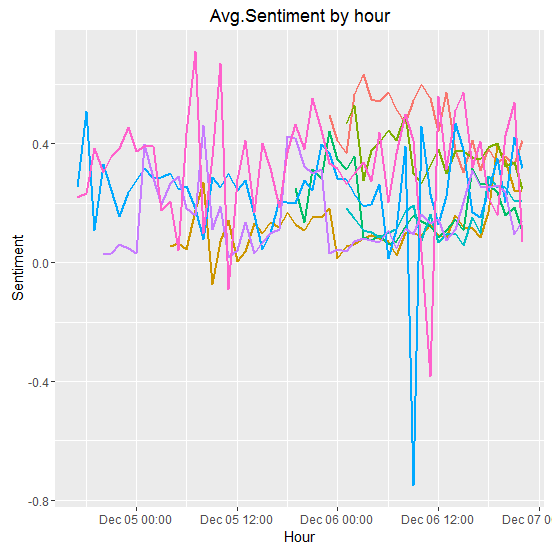
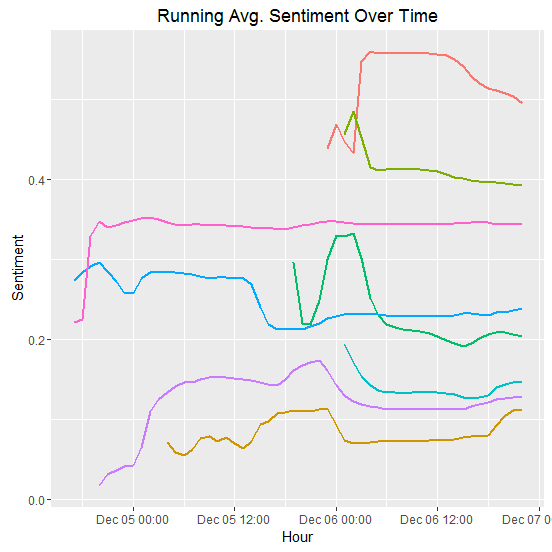
Committees choices. But surprisingly enough this is not true as in the data without retweets show Michigan, Oklahoma, and Wisconsin being 1st, 2nd, and 4th respectively in positive sentiment out of the eight teams. This could mean that these teams are happy with where they were placed within the final rankings. When we start to look at teams 5-8 in the rankings we see that of these 4, Penn State and Oklahoma could be the two teams that are the most bitter about their ranking, as they both have considerable low scores in overall sentiment when including retweets.

So if Deciding the 4 teams that would play for the ultimate goal of a National Championship were up to the people on twitter, would they choose the same teams as the committee? Well that still may be unknown. But now the Committee can see how the fan base reacted to their selection and maybe they will now take another look at Clemson and Washington.

|  |
| --- |
| **Figure 2** |



|  |  |
| --- | --- |
| **Figure 3** | **Figure 4** |

|  |  |
| --- | --- |
| **Figure 5** | **Figure 6** |

