Sports Fans' Behaviors on Twitter: A Big Data Analysis of Sentiments 2018 World Cup Final Match

The purpose of the present study is to determine the words that came to the forefront in the sharing of the fans on 2018 World Cup Final Match, the most frequently expressions that were used together with these words, and the emotional tendencies of the fans. For this purpose, 56.877 tweets that were written in English and that were shared on Twitter on 2018 World Cup Final Match were received through the software that is called "The R-Project", and were then analyzed. According to the analysis results, it was concluded that a total of twenty statements that were used with the highest frequency in the sharing of the fans were positive; and it was also determined that the positive emotional trend was dominant compared to the negative trend no matter what the result of the final match was. As a result, it may be claimed that the perceptions and reactions of the fans about the World Cup Final Match are different from the competitions among the clubs at national level, and that mostly the positive emotions came to the forefront.

Introduction

The year 2018 witnessed one of the festivals that are held every four years, the World Cup Football Olympics, one of the biggest football festivals of humanity. The World Cup that was hosted by Russia in 2018 was carried out between 14 June and 15 July, 2018. At the end of the football festival that lasted thirty days, France became the owner of the World Cup by winning against Croatia by 4-2 in the final match. Sports -especially football- is a social phenomenon with which people can reflect their reactions in a peaceful manner. Because this institution has eliminated the differences like religion, belief and ideological differences (Fişek, 1998, p. 66). As a matter of fact, football competitions provide the fans with the opportunity for acting in an emotional manner and for releasing their emotions to reveal their positive and negative reactions (Raney, 2006, p. 315). Again, these competitions may be considered as a tool for establishing good relations among people and for bringing them together (Krilenko & Stepchenkova, 2017). Today, this unity and communication has reached global dimensions with the means of the social media like football stadiums and the Twitter.

Thanks to the social media platforms -especially the Twitter-, the relation and communication among people and societies have become stronger, and have had a global status. The Twitter is one of the most popular means of communication with its 297 million users among social media platforms (Yu & Wang, 2015; Zeng & Gerritsen, 2014). One of the fields in which the means of

social media communication is widely used is football. It is possible to claim that football fans use the Twitter, which is one of the most popular social media tools, in a very intensive manner. Fans share their reactions and emotions on competitions via the Twitter in a real-time manner (Wang, 2015; Ji & Raney, 2014). Actually, after the matches are played on green fields, football fans criticize this match on social media, and share their feelings and thoughts through the Twitter. In addition to being a very popular means of communication, the Twitter has also become a gigantic data source for academic studies, and opened the way for new studies in many fields like social sciences, politics and sport sciences (Kreiss, 2016; Himelboim et al., 2016; Ceron et al., 2014).

Today, one of the studies that are conducted by authors by employing the Twitter data is sentiment analysis studies (Yu & Wang, 2015; Yu, Duan & Cao, 2013; Liu, 2012, p. 42; Krilenko & Stepchenkova, 2017). Because, the Twitter has become an important data source by providing a broad sentiment analysis platform for researchers (Pandey et al., 2017). Sentiment analysis may be defined as a method enabling researchers to determine the opinions, thoughts, and attitudes of people on products and services by obtaining their public opinions (Liu, 2012, p. 43; Baudad et al., 2017). In addition, it may be defined as the reporting and making sense of the large data volume on electronic platforms (Kaynar et al., 2016). In Sentiment Analysis, this making sense of process may be carried out by employing the analysis methods that are dictionary-based (Öztürk & Ayvaz, 2018) or machine learning-based (Mackey et al., 2018). It is observed in the literature that the data are analyzed with sentiment analyses studies conducted in the fields like education, politics and tourism (Pang & Lee, 2008; Workewych et al., 2017; Aldayel & Azmi, 2017; Kim et. al., 2016; Mostafa, 2018; Khoo & Johnkhan, 2018; Deiner et al., 2017); and it was reported in previous studies that sentiment analysis studies may be conducted for football fans in sports sciences as well (Lucas et al., 2017; Yu & Wang, 2015).

Although supporters use the Twitter commonly in football in the field of sports sciences, sentiment analysis is a field that has not been investigated at an adequate manner yet. Right at this point, only three studies in which the words "Twitter" "World Cup", and "sentiment" were used together were detected in the advanced search carried out in the Title, Abstract and Keywords sections of the Science Direct Database to examine the literature (Lucas et al., 2017; Souza et al., 2016; Yu & Wang, 2015). It was seen that these three studies were related with the 2014 World Cup. In the present study, the sentiment analysis of the fans in the 2018 World Cup final match was performed.

It is expected that this study, which was conducted on sentiment analysis for the 2018 World Cup Final Match, will contribute to the relevant literature with the use of more than one sentiment glossary and the sharing of programming codes for authors, who will conduct emotion analysis studies on social media. For this reason, the purpose of this research is to determine the expressions and emotional tendencies that came to the forefront in the tweets

about the 2018 World Cup final match. For this purpose, the answers to the following study questions were sought in the present study. In the tweets that were written about the 2018 World Cup Final Match;

1) what are the statements that come to the forefront, and what are the words which are associated with these statements?
2) how is the emotional tendency that come to the forefront (Positive, Negative, Neutral)?

Limitations

This study was limited with the 56.877 tweets that were written in English and that were shared about the 2018 World Cup Final Match between 15-23 July 2018. The reason why only English tweets were considered was that the RSentiment and Bin Liu Sentiment Glossaries that are used for emotion analysis were in English. In addition, as required by the Data Sharing Policy of the Twitter, a maximum of 10.000 tweet data per day could be downloaded from the system. The data that were downloaded from the Twitter consisted of all the tweets that contained the phrase "worldcupfinal" in their hashtags.

Materials and Methods

In the World Cup Final Matches, it may be observed that although one team lost, the supporters of both teams had positive feelings. In this study, in which the directions and levels of these sentiments were examined, a sentiment analysis was carried out on the tweets that were shared through the Twitter, which is one of the social media tools. For sentiment analysis; the tweets that were in English were scanned with #worldcupfinal hashtag on the days after the final match, and 56.877 tweets were received. These data were withdrawn with the R-Project Software, which is open-source and free, and the sentiment analysis was made again with the same software. The details on the method employed in the present study are given in the following parts under titles.

Sampling

In this study, the sampling consisted of 56.877 tweets that were in English and that were written in the Twitter with #worldcupfinal hashtag between July 15-23, 2018.

It was observed that the tweets that were written specifically for the Final Match of 2018 World Cup were shared with hashtags such as #Worldcupfinal, #Worldcupfinals, # Worldcupfinal18 and # Worldcupfinal2018. For this reason, all the tweets that included the #worldcupfinal hashtag were withdrawn from the Twitter to reach the sharing that was made with all these hashtags.

The other point that was considered when downloading the data that would constitute the sampling was that the expressions that were used in the tweets were written in English. As the RSentiment and Bing Liu Glossaries, which

would be employed during the sentiment analysis in the study had English words, a filter was applied to ensure that the tweets were in English. Right at this point, the filtering was carried out with the lang = "en" parameter, which is provided by the Twitter, during the downloading of the data; and in this way, the tweets that were written in the English language about the World Cup 2018 final match constituted the sampling of the present study.

When the number of sampling was determined, the situation in which the number of daily tweets about the match being decreased was taken into consideration. Right at this point, when a great decline was observed in the number of the tweets that were written in English with the #worldcupfinal hashtag, the data collection process for the sampling was ended on the 9th day following the match. In this way, the 56.877 tweets that were shared between 15-23 July, 2018 constituted the sampling of the present study. The graphic showing the numbers of the tweets by days is presented in Figure 1 below.

Figure 1. *The Distribution of the Tweets that were Written about 2018 World Cup Final Match according to Days*
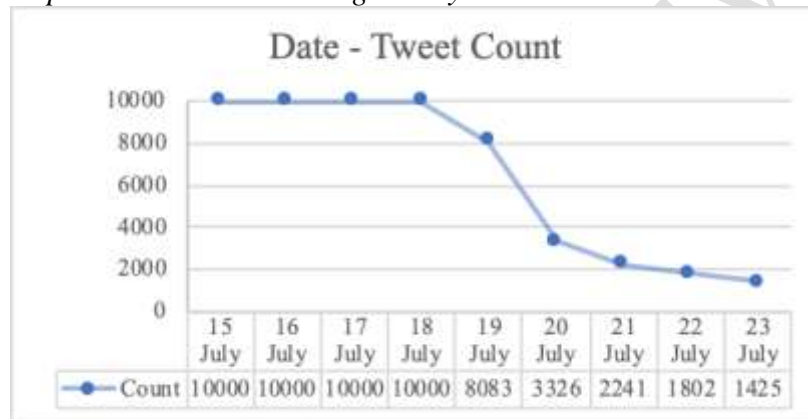
When Figure 1 is examined, it is seen that over 10.000 tweets were shared during the first 4 days; and as of the 5th day, the sharing decreased in time. In addition, since the Twitter only allows the downloading of maximum 10.000 tweet data daily, the tweets that were downloaded were limited with 10.000 for the first 4 days.

Data Collection Tool

In this study, the tweet data was collected from an open source statistical computing software which is named as 'R Project'. There are lots of studies in which the R Project was used for statistical computing [27,28] and especially analyzing the sentiments of Twitter users [16,29,30]. R Project provides ease of use with many ready-made libraries for researchers and one of these libraries is 'twitteR'. In this study, all the tweet data was collected with thank to R Project software, twitteR library and R programming codes written by researchers. Used R Project libraries and library codes are below.

```
1   library(twitteR)              # Gets the tweets on twitter
2   library(tm)                   # Creates term document matrix
3   library(wordcloud)            # Creates word cloud
4   library(RSentiment)           # Analysis sentiment
```

Data Collection Process

In the present study, the data collection process was initiated following the 2018 World Cup Final Match. According to its free use policy, the Twitter allows the withdrawal of the tweets that are made maximum one week ago. For this reason, with the help of the R programming codes that were written by the researchers, the shares that were made one day ago were downloaded. To enable the researchers, who would like to conduct studies on data analysis in social media, to withdraw data from the Twitter, the R-Programming codes that were employed in the present study are presented below.

```
16   cekilecekKelime="worldcupfinal"        # Define hashtag
17   cekilecekTweetSayisi=10000             # Define number of max tweets
18   setwd("C:/worldcup2018")               # Define working folder in computer
19                                          # Define twitter account below
20   setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)
21     # More info: https://developer.twitter.com/en/docs/tweets/search/overview/standard
22
23                                          # The following code was repeated for all
24   days
25   buTarihten="2018-07-15"                # Choose starting date (yyyy-mm-dd) *
26   buTariheKadar="2018-07-16"             # Choose ending date (yyyy-mm-dd) *
27
28                                          # R Project gets the defined tweets
29   tweets = searchTwitter (cekilecekKelime, n = cekilecekTweetSayisi, lang = "en", since =
30   buTarihten, until = buTariheKadar, locale = NULL, geocode = NULL, sinceID = NULL, maxID
31   = NULL, resultType = NULL, retryOnRateLimit = 120)
32
33   tweet.arif <-sapply(tweets, function(x) x$getText())
34   write.csv2(tweet.arif, file = "15.csv")    # Exporting data to a CSV file *
35                                          # The above code was repeated for all
36   days
```

* Date in this line was changed for each day.

With the help of the programming codes given above, the tweets that had the "worldcupfinal" hashtag and that were written in the English language (maximum 10.000) were withdrawn; and these data were transferred to a CSV format file. When the shares on the subject matter decreased in number, all the data in the CSV files where the 9-day data were kept were collected in a single CSV file; and then the process of clearing the data was initiated for the analyses.

The Clearing of the Data

Two different clearing procedures were carried out for the tweet data that would be analyzed in line with the study questions in the present study. The clearing process that was carried out for the first study question was performed for the purpose of reaching more accurate findings on the frequency of using of the relevant words. In this clearing process, the punctuation marks, links,

1 specific characters, and numbers were deleted, and more filtering was carried
2 out. The R-Programming codes and their functions that were used for the
3 purpose of clearing the data for the first study question are given below.
4

```
5   text <- readLines(file.choose())                # Choose the CSV file that includes the
6   tweets
7   stopWords <- readLines(file.choose())           # Choose the CSV file that includes the
8   StopWords
9
10  text = gsub('RT ','', text)                      # Removes "RT " word (retweet)
11  text = tolower(text)                             # Converts text to lowercase
12  text = gsub('\'s ',' ', text)                    # Removes "'s " character
13  text = gsub('\'m ',' ', text)                    # Removes "'m " character
14
15  for(a in 1:NROW(stopWords))                      # Removes defined StopWords
16          { text = gsub(stopWords[a],' ', text) }
17
18  sil<-c("#","@","http","<")                       # Removes all the words
19    for(x in 1:NROW(sil)){                         #   which starts with
20          BaslayaniSil=sil[x]                      #    the defined character(s)
21          silinecekUzunluk=nchar(BaslayaniSil)
22          for(a in 1:length(text)){
23          harfSayisi=nchar(text[a])
24      for(b in 1:harfSayisi){
25              if(substr(text[a],b,b+silinecekUzunluk-1)==toString(BaslayaniSil))
26              {
27                      oncesi=substr(text[a],1,b-1)
28                      sonKelimemi="evet"
29                          for(c in b:harfSayisi){
30                              if(substr(text[a],c,c)==" "){
31                              sonrasi=substr(text[a],c+1,harfSayisi)
32                                  gecici <- c(oncesi,sonrasi)
33                                  text[a]=toString(gecici)
34                                  remove(gecici)
35                                  sonKelimemi="hayir"
36                              break
37                              }
38                          }
39                      if(sonKelimemi=="evet"){ text[a]=oncesi}
40
41          }
42  } } }
43  text = gsub('[[:punct:]]', '', text)             # Removes punctuation
44  text = gsub('[[:cntrl:]]', '', text)             # Removes control characters
45  text = gsub('\\d+', '', text)                    # Removes numbers
46  text = gsub('\\s+', ' ', text)                   # Converts multiple spaces to one space
```

47
48   When the clearing process was carried out with the help of the codes that
49 were given above, a total of 174 words like "I", "me" "my" "is", which are
50 defined in the R-Project Software and given in the StopWords library, were
51 employed.
52   In the CSV file, where the words of the StopWords were transferred, one
53 character space () was added to the front and back of each word as a solution
54 for the deletion problem that would stem from the presence of these
55 expressions in other words (For example, the "me" characters in "come" were
56 not deleted).

6

In the clearing process for the second study question, control characters, which prevented the RSentement Glossary from functioning, the punctuation characters, and some words (never, no, not, nothing, nobody, none) were deleted. Since these words were not among the words that meant sentiments in the RSentement and Bing Liu Glossaries that were employed in the study, their deletion did not constitute a problem for the study. When the RSentement Glossary is executed without this clearing process, an error message appears. The error message that was received when the clearing process was not performed is shared below to help researchers for future studies.

```
# Researchers who encounter this error message can continue their operations if they perform the
specified cleaning process.
Error: .onLoad failed in loadNamespace() for 'rJava', details:
  call: dirname(this$RuntimeLib)
```

The Analysis of the Data

A total of 56.877 tweet data that were collected in 9 days were analyzed to determine the emotions, the most frequently used words, and the other words with which these words were used, which came to the forefront in the shares in the present study. Firstly, the Term Document Matrix was formed based on the data that were cleared for the first study question; then frequency calculation, finding related words, forming a word cloud, and drawing the graphics were carried out.

For the second study question, each word that was detected in each tweet was analyzed in sentimental terms by employing the RSentement and Bing Liu Glossaries; and the direction of the tweets was determined according to these two glossaries (very positive, positive, neutral, negative, very negative). Since there were no specific glossaries for sports sciences, two different glossaries were employed in the present study to increase the reliability of the findings on sentiment analysis. The R-Programming codes and their explanations that were employed for the analyses of the study questions in the present study are given below for the benefit of researchers who would like to conduct studies on data analysis in social media.

```
# 1. Analysis for the first research question
# 1.1. Creation of term document matrix
docs <- Corpus(VectorSource(text))
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)

# 1.2. The most used 20 words
head(d, 20)

# 1.3. Identifying the words associated with 20 words one by one (example word: croatia)
findAssocs(dtm, terms = "croatia", corlimit = 0.3)

# 1.4. Creation of Word Cloud
set.seed(1234)
```

```
1    wordcloud(words = d$word, freq = d$freq, min.freq = 1,
2         max.words=200, random.order=FALSE, rot.per=0.35,
3         colors=brewer.pal(8, "Dark2"))
4
5    # 1.5. Graphics of the most frequently used words
6    barplot(d[1:20,]$freq, las = 2, names.arg = d[1:20,]$word,
7         col ="lightblue", main ="Most frequent words",
8         ylab = "Word frequencies")
9
10   # 2. Analysis for the second research question
11   # 2.1. Preparation
12   text <- readLines(file.choose)        # Choose the CSV file that includes the tweets
13
14   text = gsub('[[:punct:]]', '', text)      # Removes punctuation characters
15   text = gsub('[[:cntrl:]]', '', text)      # Removes control characters
16
17   # 2.2. Sentiment analysis using RSentiment's English lexicon
18   calculate_total_presence_sentiment(text)
19
20   # 2.3. Sentiment analysis using Bing Liu's English lexicon
21   # Lexicon was downloaded from https://github.com/jeffreybreen/twitter-sentiment-analysis-
22   tutorial-201107/tree/master/data/opinion-lexicon-English
23
24   # 2.3.1. Identify positive and negative words to the R Project
25   pos_words = scan("words-positive.txt", what="character", comment.char=";")
26   neg_words = scan("words-negative.txt", what="character", comment.char=";")
27
28   # 2.3.2. Calculate the sentiments of tweets
29   calculate_custom_total_presence_sentiment(text, pos_words, neg_words, check = 0)
```

Results

In this section, the findings that were obtained as a result of the analyses conducted on the data in the scope of the study are given. In this context, the statements in the tweets about the 2018 World Cup final match that came to the forefront in the scope of the first study question, and the words that were related to these statements are given. Then, a Word cloud was formed for the 200 words that were used most commonly in these tweets. The statements and the words related to these statements in the tweets that were shared about the 2018 World Cup final match are given in Table 1 below.

Table 1. *The most used words in tweets and the most frequently expressions used with these words*

Word	Frequency	Used words together (%)				
croatia	11071	attended (65)	grabarkitarovic (62)	kolinda (62)	hats (61)	every (60)
france	10920	congrats (62)	wowwhat (62)	game (60)		
kids	10601	reenacted (82)	match (67)			
match	10235	reenacted (83)	kids (67)			
game	9945	wowwhat (97)	congrats (94)	good (87)	france (60)	
reenacted	7261	match (83)	kids (82)			
world	6308	cup (85)				
good	5574	wowwhat (89)	game (87)	congrats (86)		
cup	5143	world (85)				
congrats	4841	wowwhat (96)	game (94)	good (86)	france (62)	
work	4719	hard (97)	sides (97)			
president	4561	kolinda (74)	grabarkitarovic (71)	attended (66)		
hard	4537	sides (100)	work (97)			
sides	4510	hard (100)	work (97)			
wowwhat	4497	game (97)	congrats (96)	good (89)	france (62)	
kylian	4176	boots (84)	continuar (84)	dust (84)	equaling (84)	igualar (84)
		keeps (84)	records (84)	may (83)	again (82)	like (71)
like	4078	boots (85)	continuar (85)	dust (85)	equalling (85)	igualar (85)
		keeps (85)	records (85)	may (84)	again (83)	kylian (71)
amp	3835	recreated (86)	perfect (85)	beautiful (84)		
again	3158	boots (97)	continuar (97)	dust (97)	equalling (97)	igualar (97)
		keeps (97)	records (97)	like (83)	kylian (82)	
beautiful	3126	recreated (97)	perfect (96)	amp (84)		

When Table 1 is examined, it is seen that the statements about the 2018 World Cup Final Match are listed as follows: "Croatia" "France" "Kids" "Match" "Game" "Reenacted" "World" "Good" "Cup" "Congrats" "Work" "President" "Hard" "Sides" "Wowwhat" "Kylian" "Like" "Amp" "Again" "Beautiful". Following the determination of the 20 statements that came to the forefront in Twitter shares, it was determined with which words they were most commonly used that written together over than 60% frequency. In this respect,

1   it was seen that the word "Croatia", which came to the forefront, was used
2   together with the words "attended", "kolinda" and "grabarkitarovic". Another
3   word, which came to the forefront was "France" was used with the words
4   "congrats", "game" and "wowwhat".
5   A Word Cloud, which consisted of 200 words that were most commonly
6   used in the tweets that were written about the 2018 World Cup final match to
7   consider the words from a wider viewpoint. This Word Cloud is given in
8   Figure 2 below.
9
10 **Figure 2.** *The Word Cloud that was Formed for the 200 Words that were Most*
11 *Commonly Used.*

12
13   As it may be understood from the Word Cloud, which was prepared for the
14 200 most commonly used words, it is seen that the words that represent the
15 unifying power of football like "equaling", "perfect", "muslims", "bringing",
16 "celebrate", "woman", "won", "brothers", "alive", "winner" were used in the
17 shares.
18   In the scope of the second study question, a Sentiment Analysis was
19 conducted in order to determine the sentimental tendencies of the twitter shares
20 about the final match of the fans. Two different sentiment analysis glossaries
21 (i.e. RSentement and Bing Liu) were employed for the purpose of determining
22 the sentimental distribution of the tweets. The graphic that shows the

sentimental distribution of the tweets according to the RSentimens English Lexicon and Bing Liu Sentiment Lexicon is given below in Figure 3.

Figure 3. *The Sentimental Distribution of the Tweets*

RSentment's English Lexicon

Bing Liu's English Lexicon
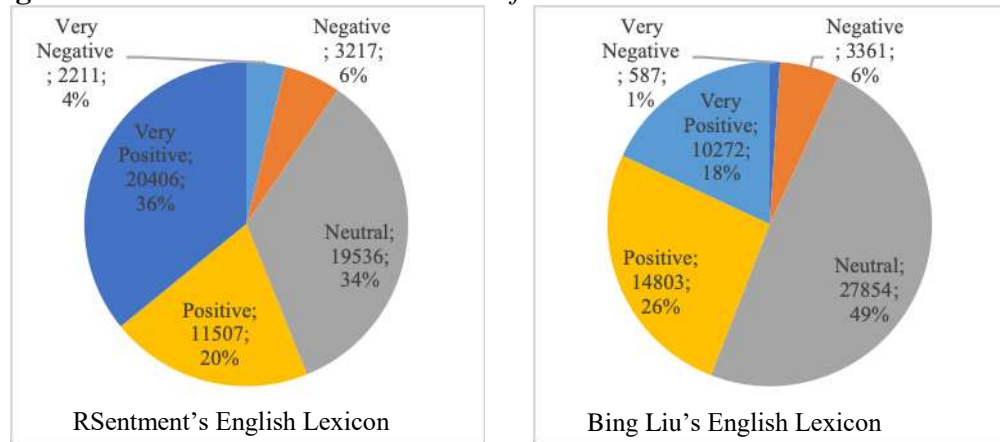
When Figure 3 is examined it is seen that the word counts in the twitter shares according to the results of the analysis made with the RSentement Lexicon were; very negative 2211 (%4), negative 3217 (%6), neutral 19536 (%34), Positive 11507 (%20) and very positive 20406 (%36)'dır. the word counts in the twitter shares according to the results of the analysis made with the Bing Liu Lexicon were; very negative 587 (%1), negative 3361 (%6), neutral 27854 (%49), positive 14803 (26%) and very positive 10272 (%18). As a result of the Sentiment Analysis that was carried out with RSentiment Lexicon, it was determined that the positive tweet rate was 56%, and the negative tweet rate was 10%. As a result of the Sentiment Analysis that was carried out with Bing Liu Sentiment Lexicon, the rate of positive tweets was 44%, and the negative tweet rate was 7%. For this reason, it is understood that the rate of the positive tweets is higher at a significant level than the rate of the negative tweets according to both glossaries. However, according to the analysis results obtained with the RSentiment Lexicon, the rate of the Neutral tweets is 34%; and 49% according to the analysis results obtained with the Bing Liu Sentiment Lexicon.

Discussion

Since fans can express themselves in a comfortable manner in great sports organizations like the World Cup, the sentimental reactions can be measured (Lucas et al., 2017). Based on this viewpoint, the purpose of the present study was to determine the statements of the fans in their tweets about the 2018 world cup final match, the words related to these statements, and the emotional tendencies in the twitter shares.

As the first step, the statements about the world cup final match in the twitter shares were determined in the scope of the first study question. In this respect, the statements like "Croatia", "France", "kids", "match" and "game" came to the forefront in the tweets about the final match. Unlike the violent and hatred statements, which surpassed the football in the fields from time to time, the fact that the expressions in the twitter shares about the World Cup final match are positive attracts attention and are exciting. When the statements used were examined with a closer look, it is possible to claim that twitter shares have positive statements like "Congrats", "Beautiful" and "Good". Right at this point, it is understood that the competition in the world cup is different from the one in club football; and that the nations might as well support their teams in a gentle manner. It may also be claimed that fans who come from different languages, religions, races or nationalities may unite around the beauties of football in terms of fair play. Because sports -especially football- is a universal activity field that increases the differences between cultures and improves international relations and contributes to the world peace (Üstünel & Alkurt, 2015).

Following the determination of the statements that came to the forefront in the Twitter shares, the words that were used most frequently were determined. In this respect, it was determined that the word "Croatia", which came to the forefront, was used together with the words "attended", "kolinda" and "grabarkitarovic". It was also understood that another word "France", which was used most frequently, was used together with the words "congrats", "game" and "wowwhat". It may be claimed that using the word "France" especially with the word "congrats" was due to the fact that it completed the tournament as the champion and had the world cup. It was also determined that the word "Croatia" was used in an intense manner together with the name of the head of Croatia (Kolinda Grabar-Kitarovic). This can be explained in two ways. Firstly, the president of Croatia supported the Croatian national team throughout the world cup, and was in the tribunes; and secondly, the relation between football and politics showed itself in this world cup specifically with Croatia. Ince (2016), who examined the football-politics relation, reported that governments adopted being within sports especially in football as a policy, and underlined that there is a close relation between politicians and sports. Goig (2017), who analyzed the relation between football and politics on a club basis with the fans of Barcelona, Atletico Madrid and Real Madrid in Spain, reported that the ideologies that were represented by the clubs overlapped with the political viewpoints of the fans. For this reason, the word "Croatia" being used together with the name of a head of state may be accepted as the indicator of the fact that football and politics are in very close contact.

For the purpose of considering the words that came to the forefront in the scope of the study, a Word Cloud that had 200 words was prepared. According to this Word Cloud that was prepared, it may be claimed that the words "equaling", "perfect", "muslims", "bringing", "celebrate" and "woman", which represented the peaceful and unifying power of football, were used. As a matter

of fact, Afacan et al. (2017) pointed out to the fact that football, which was a popular sports branch, had a unifying and integrative structure. However, unlike this positive finding, it is observed that the studies reported in the literature examining the relation between fans and football, were generally shaped around violence and hooliganism (Newson et al., 2018; Stott et al., 2008; Raspaud & Bastos, 2013; Sekulic et al., 2015; Spaaij, 2008; Wakefield & Wann, 2006). The fact that the findings of the present study is different from the findings reported in the literature may be explained with the difference in the viewpoints of the fans on the world cup organization.

In the scope of the second study question, the sentimental distributions of the tweets that were shared by the fans on the final match were determined. For this purpose, the RSentiment English Lexicon and the Bing Liu English Lexicon were employed in Sentiment Analysis. The results of the Sentiment Analysis that was carried out with both dictionaries showed that the rate of the positive tweets about to the final match was much higher than the negative tweets in the twitter shares. However, in the study that was conducted on the world cup for a Sentiment Analysis by Yu & Wang (2015) on the 2014 World Cup, it was concluded that the sentimental tendencies of the fans were negative in their twitter shares when their team lost the game. However, it was concluded in another Sentiment Analysis study on the 2014 Olympic Games in Sochi that the emotional tendencies of the supporters, who followed the matches were positive in their twitter shares (Krilenko & Stepchenkova, 2017). For this reason, it may be said that the findings of the study both overlapped and did not overlap with the literature. This situation may be interpreted as that the aggressive attitude among the fans in club-level competitions does not exist in national teams' level like the 2018 world cup.

Again, according to the results of the analysis that were made with the RSentiment Lexicon, it was seen that the rate of the neutral tweets was 34%; and 49% according to the analysis made with the Bing Liu Sentiment Lexicon. According to the results of the analysis that were made with both glossaries, it may be claimed that fans' being neutral in terms of sentiments was caused by the fact that final match was followed by the supporters of countries other than France or Croatia; and that they remained neutral about this competition.

Conclusions

In this study, the statements of the fans that came to the forefront about the 2018 world cup final match, the words that were related with these statements, and the emotional tendencies of the twitter shares were determined. In this respect, when the 56.877 tweets that were shared about the World Cup final match were analyzed, it was determined that;

- all of the statements and the words that were used together with these statements were positive,

- according to the Word Cloud that was prepared in the context of the study, it was determined that the words that were used on twitter shares were positive in general,
- the rates of the positive tweets were much more than the rates of the negative tweets,
- the way of perception in which the competition at national teams' level in the world cup and the competition among clubs differentiated in a positive way.

In future studies, it may be recommended that a glossary may be formed for sentiment analysis in specific areas like sports sciences, and methods like machine learning may be employed to analyze tweets that are written in more than one language.

References

Afacan, E., Onağ, Z., Demiran, D. & Çobanoğlu, G. (2017). The reasons of violence in football and the ways to prevent them according to Professional football players views: The match between Tarsus İdman Yurdu Erkutspor and Yeni Salihlispor. *International Journal of Social Science Research* 6(2), 124-141.

Aldayel, H. & Azmi, A. M. (2016). Arabic tweets sentiment analysis – a hybrid scheme. *Journal of Information Science* 42(6), 782–797. https://doi.org/10.1177/0165551515610513

Baudad, N., Faizi, R., Thami, R. & Chibeb, R. (2017). Sentiment analysis in Arabic: A review of the literature. Ain Shams Engineering Journal 9(4), 2479-2490. https://doi.org/10.1016/j.asej.2017.04.007

Ceron, A., Curini, L., Iacus, S. M. & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society* 16(2), 340–358. https://doi.org/10.1177/1461444813480466

Deiner, M. S., Fathy, C., Kim, J., Niemeyer, K., Ramirez, D., Ackley, S. F. & Porco, T. C., (2017). Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Informatics Journal. https://doi.org/10.1177/1460458217740723

Fişek, K. (1998). Spor Yönetimi. [Sport Management]. Ankara: Bağırgan Yayınevi.

Goig, R. L. (2017). Football and politics in Spain: An empiricial analysis of the social base of the main football clubs. *Journal of Iberian And Latin American Literary And Cultural Studies* 1(4), 79-100.

Himelboim, I., Sweetser, K. D., Tinkham, S. F., Cameron, K., Danelo, M. & West, K. (2016). Valence-based homophily on Twitter: Network analysis of emotions and political talk in the 2012 presidential election. *New Media & Society* 18(7), 1382–1400. https://doi.org/10.1177/1461444814555096

İnce, M. (2016). Spor ile siyasetin ilişkisi üzerine bir analiz; Sporu siyasete alet etmek. [An Analysis on relation of sport and politics; play politics with something]. *Karabük Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 6(2), 456-464.

Ji, Q. & Raney, A. A. (2014). Morally judging entertainment: A case study of live tweeting during downton abbey. *Media Psychology* 18(2), 221-242. https://doi.org/10.1080/15213269.2014.956939

Kaynar, O., Yıldız, M., Görmez, Y. & Albayrak, A. (2016). *Sentiment Analysis with Machine Learning Techniques.* Paper presented International Artificial Intelligence and Data Processing Sysmposium, September 28-30, 2018, Malatya, Turkey.

Khoo, C. S. & Johnkhan, S. B. (2018). Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *Journal of Information Science* 44(4), 491–511. https://doi.org/10.1177/0165551517703514

Kim, E. H., Jeong, Y. K., Kim, Y., Kang, K. Y. & Song, M. (2016). Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. *Journal of Information Science* 42(6), 763–781. https://doi.org/10.1177/0165551515608733

Kızılkaya, Y. M. & Oğuzlar, A. (2017). Tweetlerin etkinliğinin sosyal ağlar ile irdelenmesi. [Effectiveness of tweets with social networks]. *International Journal of Social Inquiry* 10(2), 135-150.

Kızılkaya, Y. M. & Oğuzlar, A. (2018). Bazı denetimli öğrenme algoritmalarının R Programlama Dili ile kıyaslanması. [Comparison of some supervised learning algorithms R programming language]. *Karadeniz Uluslararası Bilimsel Dergi* 37(37), 90-98.

Kreiss, D. (2016). Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle. *New Media & Society* 18(8) 1473–1490. https://doi.org/10.1177/1461444814562445

Krilenko, A. P. & Stepchenkova, S. O. (2017). Sochi 2014 Olympics an twitter: Perspective of hosts and guests. *Tourism Management* 63(2017), 54-65. https://doi.org/10.1016/j.tourman.2017.06.007

Liu, B. (2012). *Sentiment analysis and opinion mining*. Williston: Morgan & Claypool Publishers.

Lucas, G. M., Gratch, J., Malandrakis, N., Szablowski, E., Fessler, E. & Nichols, J. (2017). GOAALLL! Using sentiment in the World cup to explore theories. *Image and Vision Computing* 65(2017), 58-65. https://doi.org/10.1016/j.imavis.2017.01.006

Mackey, T., Kalyanam, J., Klugman, J., Kuzmenko, E. & Gupta, R. (2018). Solution to detect, classify, and report ıllicit online marketing and sales of controlled substances via Twitter: using machine learning and web forensics to combat digital opioid access. *Journal of Medical Internet Research* 20(4), e10029.

Mostafa, M. M. (2018). Clustering halal food consumers: A Twitter sentiment analysis. *International Journal of Market Research*, https://doi.org/10.1177/1470785318771451

Newson, M., Bortolini, T., Buhrmester, M., Da Silva, S. R., Da Aquino, J. N. Q. & Whitehouse, H. (2018). Brazil's football worriors: Social bonding and inter-group violence. *Evolution and Human Behavior* 39(2018), 675-683. https://doi.org/10.1016/j.evolhumbehav.2018.06.010

Öztürk, N. & Ayvaz, S. (2017). Sentiment analysis on twitter: A text mining approach to the Syrian refugee crisis. *Telematics and Informatics* 35(1), 136-147.

Pandey, A. C., Rajpoot, D. S. & Saraswat, M. (2017). Twitter sentiment analysis using hybrid cuckoo search. *Information Processing and Management* 53(2017), 764-779. https://doi.org/10.1016/j.ipm.2017.02.004

Pang, B. & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2(1), 1-135.

Philander, K. & Zhong, Y. (2016). Twitter sentiment analysis: capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management* 55, 16-24.

Raney A.A. (2006). *Why we watch and enjoy mediated sports*. In: AA Raney & J Bryant (Eds), Handbook of Sports and Media. New York: Lawrence Erlbaum Associates, pp 313-329.

Raspaud, M. & Bastos, F. (2013). Torcedores de futebol: Violence and public policies in Brazil before the 2014 FIFA world cup. *Sport in Society, Culture Commerce Media Politics* 16(2), 192–204. https://doi.org/10.1080/17430437.2013.776251

Ripley, B. D. (2001). The R project in statistical computing. MSOR Connections. *The Newsletter of The LTSN Maths, Stats & OR Network* 1(1), 23-25.

Sekulic, M., Kühl, S., Connert, T., Krastl, G. & Filippi, A. (2015). Dental and jaw injuries sustained by hooligans. *Dental Traumatology* 31(6), 477–481. 10.1111/edt.12205.

Souza, A., Figueredo, M., Cacho, N., Araújo, D. & Prolo, C. A. (2016). Using big data and real-time analytics to support smart city initiatives. *IFAC-PapersOnLine* 49(30), 257-262.

Spaaij, R. (2008). Men like us, boys like them: Violence, masculinity, and collective ıdentity in football hooliganism. *Journal of Sport and Social Issues* 32(4), 369–392. https://doi.org/10.1177/0193723508324082

Stott, C., Adang, O., Livingstone, A. & Schreiber, M. (2008). Tackling football hooliganism: A quantitative study of public order policing and crowd psychology. *Psychology Public Policy and Law* 14(2), 115-141. http://dx.doi.org/10.1037/a0013419

Üstünel, R. & Alkurt, Z. (2015). Futbolda şiddet ve düzensizliğin önlenmesi için 6222 sayılı yasanın getirdiği yeni bir uygulama: Elektronik bilet ve yaşanan sorunlar. [A new application by law no.6222 for prevention of football violence and disorder: Electronic ticket and problems experienced]. *Science Journal of Turkish Military Academy* 25(2), 141-175.

Wakefield, K. L. & Wann, D. L. (2006). An examination of dysfunctional sport fans: Method of classification and relationships with problem behaviors. *Journal of Leisure Research* 38(2), 168-186. DOI:10.1080/00222216.2006.11950074

Wang, X. (2015). Using attitude functions self-efficacy and norms to predict attitudes and intentions to use mobile devices to access social media during sporting event attendance. *Mobile Media and Communication* 3, 75-90. https://doi.org/10.1177/2050157914548932

Workewych, A. M., Muzzi, M. C., Jing, R., Zhang, S., Topolovec-Vranic, J.& Cusimano, M. D. (2017). Twitter and traumatic brain injury: A content and sentiment analysis of tweets pertaining to sport-related brain injury. *SAGE Open Medicine* 5, 1-11. https://doi.org/10.1177/2050312117720057

Yu, Y. & Wang, X. (2015). World Cup 2014 in the twitter world: A big data analysis of sentiments in U.S. sports fans' tweets. *Computer in Human Behavior* 48(2015), 392-400. https://doi.org/10.1016/j.chb.2015.01.075

Yu, Y., Duan, W. & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support System* 55(4), 919-926. https://doi.org/10.1016/j.dss.2012.12.028

Zeng, B. & Gerritsen, R. (2014). What do we know about social median tourism? *A review. Tourism Management Perspective* 10, 27-36. https://doi.org/10.1016/j.tmp.2014.01.001