CHAPTER 7

RELATIONS BETWEEN CORONAVIRUS (COVID-19) PANDEMIC DATA AND SOCIAL MEDIA INTERACTIONS: BIG DATA MINING IMPLEMENTATION – TURKEY 2020 CASE STUDY

İlayda DİLEK[1]
Mustafa COŞKUN[2]

AIM

The main research questions of this study are to what extent the effects of the global pandemic caused by the COVID-19 virus on society reflect on social media and whether people's emotions and mood analysis can be made with social media big data. Although various studies addressing these issues were revealed with literature review, any hybrid qualitative and quantitative study could not be covered in the related literature, specifically as Turkey case study. In this context, the main purpose of this study is to determine the reactions that reflect from society in COVID-19 pandemic process and to analyze the relationship between pandemic data, features of their accounts and sentiments of social media users in Turkey in these reactions' perspective.

INTRODUCTION

The COVID-19 global pandemic, which started in Wuhan, China in December 2019 and spread all over the world in a short time, had an impact in many areas such as the psychological and physical health, economic situation, and security of societies. Many events such as social distance, curfews (lackdowns), transportation restrictions, distance education of schools, reduction in job opportunities, etc. have demonstrated their effects in Turkey as all over the world. On the other hand, the COVID-19 global pandemic process has provided a great opportunity for scientists to study how societies respond to crisis situations. In this period which reduced face-to-face communication, people satisfied most of their needs of communication such as socializing and obtaining information through social media. Social media environments, where feelings and thoughts

[1] Researcher, İzmir Bornova College Science High School, ilaydadilek35@gmail.com
[2] PhD. İzmir Provincial Directorate for National Education, mustafacoskun@hotmail.com

can be freely announced, have inevitably come to contain the most valuable social analysis data in this period when people, who are a social entity, are confined to home. Thus, social media has turned into a big data mining source that can be used for effective and meaningful social analysis. Cinelli, Quattrociocchi, and Galeazzi (2020) and Ferrara (2020) showed in their studies how important social media is as a news source, especially during the pandemic period, moreover, there are bots that aim to create various perceptions in the society as well as accurate news. In addition, Ahmad and Murad HR (2020) and Li, Chaudhary and Zhang (2020) discussed the relationship between social media and people's sentiments in their studies and emphasized that with accurate analysis, sentiments can be captured with social media data. In limited number of studies of COVID-19 pandemic process in Turkey based on social media data it has been focused on their effects on education (Akbulut, Sahin ve Esen, 2020) or has been analyzed how affected their psychological state (Karaşar ve Canlı, 2020). When the limited number of COVID-19 studies in the literature are reviewed, to the best of authors' knowledge; a country-based analysis of the pandemic-social media relationship has not been done yet. Thus, a Turkey Case study was questioned. Therefore, how information chaos in social media impact on various aspects such as frequency of using social media, moods, feeling changes of users in Turkey during COVID-19 pandemic process was studied by analyzing data collected from twitter.com, the most popular online short message sharing platform. As a result of analyzing the limited number of studies obtained through the literature review, the following hypotheses were formed to be tested in the study.

H1. There is a scientifically significant relationship between the daily number of tests, illnesses, deaths, seriously illnesses, recovering patients and the daily number of tweets.

H2. The number of retweets of medical subject follower social media accounts increases on days that are crucial in terms of the number of COVID-19 cases.

H3. There is a scientifically significant relationship between number of COVID-19 cases and rates of feeling change of active medical subject follower social media accounts.

H4. There are significant similarities between the words used in social media and their usage densities in the important days of the year 2020 in terms of the number of cases, when COVID-19 occurred.

METHODS

In the study, relational, causal comparison and time-series relationship analyses were used as research methods. In this context, a hybrid result was tried to be obtained by using both qualitative and quantitative methods.

Firstly, social media accounts were determined, then their tweets in the relevant time period were obtained, and then these tweets were subjected to sentiment analysis with dictionaries and algorithms obtained by reviewing the literature. In the last stage, the hypotheses were tested by analyzing the relationships between pandemic and social media time-based data. It is possible to divide the big data mining process, which will be discussed in detail, into 3 main stages.

1. Determining and filtering the accounts
2. Tweeting and sentiment analysis process
3. Analysis of the relationships between pandemic data and Twitter users' account and tweet properties

1. Determining and filtering the accounts:

As the first step of the study, data of 209887 accounts of the followers of @ saglikbakanligi account on the twitter.com site that include user id, user name, profile picture, date of establishment, location, number of likes, whether protected account, number of followers, whether approved account, biography, number of tweets posted, data on the number of accounts they follow mutually, their display names and the number of lists they have created were collected using the User API (Advanced Programming Interface) offered in the Twitter Developer Netwrok in PHP coding language. @saglikbakanligi account had 2 million followers on the data collection phase of this study. However, due to the time limitation, only the information of 209887 accounts could be reached, randomly. It is predicted that real and meaningful accounts can be reached by making the necessary filtering on these first 209887 randomly accessed accounts.

The reason for choosing the accounts that follow the @saglikbakanligi as the sample is that the users should be active medical events followers. Among the accounts providing news in the field of health, @saglikbakanligi has been chosen as the account that can be described as the most unbiased since it belongs to the T.C. Ministry of Health (the formal account). This account is followed by people who have both positive and negative thoughts about the process. Therefore, it will be an accurate resource for a society analysis. In accordance with the KVKK (Personal Data Protection Law), only "unprotected, open to public accounts" were examined and user information was anonymized in all phases of the study. Any personal data was not collected and included in the data mining sets. After reaching the account information, the data from the MYSQL database were transferred to the Microsoft Excel program and passed through the following filtering stages.

First of all, accounts opened after January 1, 2020 were eliminated. The reason for this filtering is that the data between January 1, 2020 and December 31, 2020 will be used in our study. As a result of this elimination, 68659 accounts remained. Later, the accounts whose tweets were protected were eliminated. Because access to the tweets of these accounts for analysis is subject to the special permission of the persons in accordance with the KVKK and Twitter contract. As a result of this elimination, 52587 accounts remained. The accounts that remained after these filtering could contribute to our work, but there were still bots and fake accounts among them. Two criteria were determined to eliminate these accounts. The first criterion was the elimination of all accounts whose biography was between 18.34479216 and 85.97610054, because the average number of characters in the biographies of all users was calculated as 18.34479216 and the standard deviation as 33.81565419. In order to obtain reliable dataset, accounts containing length of biography two standard deviation away from mean were eliminated. As a result of this elimination, 11682 accounts remained. As the second elimination criterion, the number of daily tweets posted by the users was reviewed. Those whose average number of tweets per day was not between 0.49 and 11.16 were eliminated. Again, the reason was being two standard deviation away from mean similarly. With this elimination, 3024 accounts remained. Finally, 2891 of the remaining 3024 accounts, whose tweets can be accessed, were selected.

Tweeting And Sentiment Analysis Process:

The tweets of the determined 2891 open accounts were pulled by the local application written in PHP with the Get Status API from Twitter Developer APIs. In terms of confidentiality and protection of data, the tweets were kept on a local computer and certainly were not put on the server or cloud environment (although they are anonymous and open accounts). The tweets in the time period were subjected to sentiment analysis with the sentiment analysis algorithm presented in the studies of Coşkun and Özturan (2018) and the Turkish Emotion dictionary prepared by Vural (2013). 2642347 tweets were pulled from 2891 users.

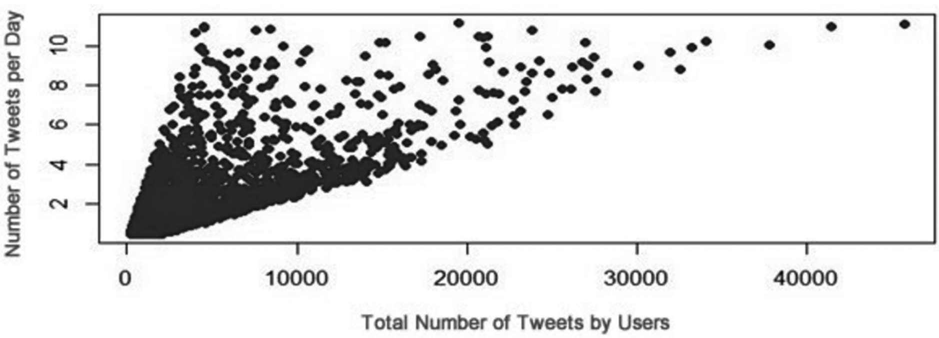
Analysis Of The Relationships Between Pandemic Data And Twitter Users' Account And Tweet Properties:

Firstly, the frequencies and the relationships were reviewed. Possible relationships of the data between numerical data such as the number of pandemic cases and daily tweets were reviewed through Pearson Product Moment Correlation Coefficients. Two tailed significance tests were performed. These statistical analyzes were made with IBM SPSS version 25 and R 4.0.3. The positive and negative peaks in the graphic that emerged as a result of the sentiment analysis and the events that took place nowadays were reviewed and it was revealed the types of social events caused emotion changes in social media. Analysis results are presented in the findings section.

RESULTS

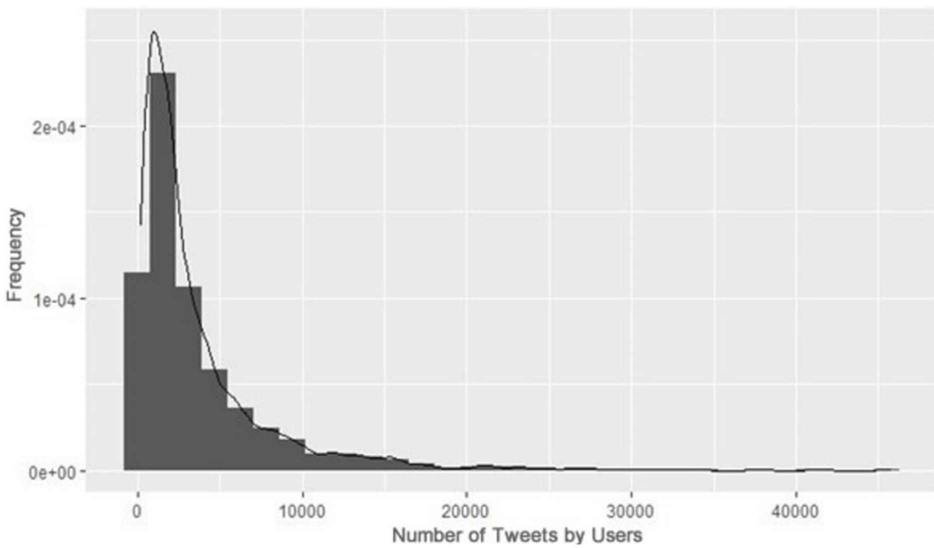
1. Descriptive Statistics:

Firstly, the total number of tweets of users selected for analysis and whose tweets were collected (from the date they created their accounts to today) and the number of tweets sent daily (during the pandemic period) were reviewed. This relationship is presented in Graphic 1.



Graphic 1: *Relationship Between Total Number of Tweets by Users – Number of Tweets per Day*

As it can be seen in Graphic 1, the number of daily tweets posted by users with a huge number of tweets in total is highly proportionate; fewer number of daily tweets posted by users with a relatively small number of tweets in total have a low proportionate again. This consistency is accepted as one of the proofs that the users in the data set are real and correct accounts (not bots). Afterwards, the distribution of the number of tweets was analyzed (Graphic-2).



Graphic 2: *Distribution of Number of Tweets by Users*

Graphic 2 demonstrates that number of tweets of almost all users are between 0-10000. It has been concluded that there is no need for a filtering on this issue due to the lack of a high proportion of users against the standard distribution. Afterwards, social media account features of the users such as tweets, likes and followers were reviewed. (Table 1).

Table 1: Descriptive Statistics of Users

	Number of Daily Tweets	Number of Likes	Number of Followers	Verified	Length of Biography	Number of Tweets	Number of Lists
Number of Values	3023	3023	3023	10	3023	3023	3023
Number of Zero Value	0	6	2	0	0	0	0
Number of Null Value	0	0	0	3013	0	0	0
Minimum	0,485309866592615	0	0	1	18	191	3
Maximum	11,1454741464459	255872	151873	1	85	45690	32511
Range	10,6601642798533	255872	151873	0	67	45499	32508
Total	5923,05476479482	27430016	3243154	10	134876	11249840	2625886
Median	1,21974263356952	3853	281	1	42	2105	457
Mean	1,95933336579385	9073,77307310619	1072,82633145882	1	44,6166060205094	3721,41581210718	868,635792259345
Standard Error of the Mean	0,0348798911128608	311,577176627673	88,779313236108	0	0,346781873618834	81,4497559481602	28,3897277725992
Confidence Interval	0,0683907218778695	610,924729130763	174,073975758491	0	0,67995231390523	159,702551480073	55,66513868985
Variance	3,67780238862811	293473858,736707	23826580,0045745	0	363,53892997279	20054771,6751569	2436467,39179592
Standard Deviation	1,91775972651115	17131,0787382671	4881,24779176129	0	19,0666968815469	4478,25542763663	1560,91876527766
Coefficient of a Variance	0,978781742806762	1,88797742683713	4,54989558759602	0	0,42734529992672	1,20337410645356	1,79697725926958

In the row specified as "Number of Zero Value" in Table 1, 0 counted values in the data are shown. For example, there are 2 users with 0 followers. "Number of Null Value" refers to those that have no value. While determining the approved accounts, 1 is assigned if it is approved, and left blank if it is not approved. As it can be seen from the table, there are only 10 approved accounts.

After the analysis of the accounts, the statistics of the tweets were also obtained (Table 2).

Table 2: Descriptive Statistics of Tweets

	Number of Tweets	Number of Replies	Number of Retweets	Number of Mentions	Mean of Tweet Length	Mean of Likes	Retweeted Tweets	Sentiment Average
Number of Values	296	296	296	296	296	296	296	296
Number of Zero Value	0	0	0	0	0	0	0	2
Number of N/A Value	0	0	0	0	0	0	0	0
Minimum	3621	1105	1549	3587	106,821661759058	1,67587711325214	144,823295759528	-0,0292755234389476
Maximum	13102	3402	7140	21681	129,25531167691	8,9929955290611	4814,76402401159	0,0253240518482957
Range	9431	2297	5591	18094	22,4336499178518	7,31711841580897	4669,94072825206	0,0545995752872433
Total	1951168	552266	896677	2202912	34946,2542350612	1074,67940984184	242778,058036411	-2,02169877782756
Median	6209	1781	2797	6722	117,74652970192	3,36504289774718	677,815500214028	-0,00590244961328948
Mean	6591,78378378378	1865,76351351351	3029,31418918919	7442,27027027027	118,061669713044	3,63067368189809	820,196142014901	-0,00683006343860662
Standard Error of the Mean	100,280748603933	26,2248945115747	54,7160313636328	164,381495241296	0,203943766538372	0,0700090317822562	35,9624596596518	0,000385316956009768
Confidence Interval	197,35633544229	51,6115919577608	107,683235219172	323,509048017557	0,401369105841839	0,137780442934058	70,7754913154057	0,000758318455740798
Variance	2976643,64800733	203572,547274393	886177,850103069	7998297,68941823	12,3115457333179	1,45077430120233	382816,357412537	4,39468703502356e-05
Standard Deviation	1725,29523502713	451,190145364893	941,37019822335	2828,1261798969	3,50878123189775	1,20448092604338	618,721550790448	0,00662924357300556
Coefficient of a Variance	0,261734196936415	0,241826009618568	0,310753569762703	0,380008529278284	0,0297199018142471	0,331751358445876	0,754358035957705	-0,970597657341522

Table 2 contains the statistics of the tweets shared by the users whose statistics are given in Table 1 between January 1, 2020 – December 31, 2020.

"Number of Zero Value" shows how many values are equal to 0 in the dataset. For example, there are 2 days with an average sentiment state of 0. "Number of N / A Value" shows how many values are empty or unusable in the dataset.

Finally, descriptive statistics of the data of COVID-19 pandemic in Turkey are listed. (Table 3).

Table 3: Descriptive Statistics of COVID-19 in Turkey

	Total Number of Tests	Total Number of Patients	Total Number of Deaths	Number of Severely Ill Patients	Number of Recovered Patients	Daily Number of Patients	Daily Number of Tests	Daily Number of Deaths	Daily Number of Recovered Patients
Number of Values	280	296	290	156	280	280	280	272	272
Number of Zero Value	0	0	0	0	0	0	0	0	0
Number of N/A Value	16	0	6	140	16	16	16	24	24
Minimum	47823	1	1	542	42	786	7533	256	256
Maximum	24504567	2208652	20881	5988	2100650	7381	208873	35511	35511
Range	24456744	2208651	20880	5446	2100608	6595	201340	35255	35255
Total	2257119179	107966109	2063552	385140	91100904	640209	24464277	996901	996901
Median	5556976	236688,5	5836,5	1553	228518,5	1637	67053	1401,5	1401,5
Mean	8061139,925	364750,368243243	7115,69655172414	2468,84615384615	325360,371428571	2286,46071428571	87372,4178571429	3665,07720588235	3665,07720588235
Standard Error of the Mean	413646,623921902	28254,3202877839	276,067399444182	144,361326716227	26327,6616810256	93,907264992818	3319,81450422679	414,810296515036	414,810296515036
Confidence Interval	814264,672476469	55605,5792366818	543,35763221904	285,169507017278	51826,0843338085	184,856744743975	6535,06523113128	816,660390302869	816,660390302869
Variance	47908988254956,4	236298758017,705	22101830,6204033	3251070,05359802	194080815485,353	2469200,83716078	3085927135,89287	46802382,3298106	46802382,3298106
Standard Deviation	6921631,90692458	486105,706629437	4701,25840817151	1803,07239277795	440546,042412541	1571,36909641267	55551,1218238918	6841,2266685011	6841,2266685011
Coefficient of a Variance	0,858641826258162	1,33270792561686	0,660688433521296	0,730329992401103	1,35402489392983	0,687249549749454	0,635797007640557	1,86659824178359	1,86659824178359

Table 3 was drawn using the General Coronavirus Table published by the T.C. Ministry of Health at https://covid19.saglik.gov.tr/TR-66935/genel-koronavirus-tablosu.html page. The data belong to between the dates of March 11th to December

31st, 2020 due to date of first case that emerged is March 11th, 2020 in Turkey. In the table, "Number of Zero Value" shows how many values are equal to 0 in the dataset. "Number of N / A Value" shows how many values are empty or unusable in the dataset.

1. Correlation analysess:

In this section, the possible relationships between pandemic data and social media data (H1 and H2 hypotheses) were analyzed through Pearson Product Moment Correlation Coefficients.

Table 4: Positive Relationships Between Daily Total Number of Patients and Social Media Data

		Daily Number of Total Patients	Number of Tweets	Number of Replies	Number of Retweets	Number of Mentions	Average Length of Tweets
Daily Number of Total Patients	Pearson Correlation	1	.734**	.676**	.729**	.802**	.246**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	296	296	296	296	296	296
Number of Tweets	Pearson Correlation	.734**	1	.918**	.960**	.953**	.167**
	Sig. (2-tailed)	.000		.000	.000	.000	.004
	N	296	296	296	296	296	296
Number of Replies	Pearson Correlation	.676**	.918**	1	.804**	.867**	-.026
	Sig. (2-tailed)	.000	.000		.000	.000	.651
	N	296	296	296	296	296	296
Number of Retweets	Pearson Correlation	.729**	.960**	.804**	1	.935**	.299**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	296	296	296	296	296	296
Number of Mentions	Pearson Correlation	.802**	.953**	.867**	.935**	1	.251**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	296	296	296	296	296	296
Average Length of Tweets	Pearson Correlation	.246**	.167**	-.026	.299**	.251**	1
	Sig. (2-tailed)	.000	.004	.651	.000	.000	
	N	296	296	296	296	296	296

**.Correlation is significant at the 0.01 level (2-tailed).

Table 4 and Table 5 shows that, five of the relationships between the daily total number of patients and social media interaction variables (Table 2) are positive; three of them are negatively correlated (H1). There is a positively high

(0.734, 0.729, 0.802) scientifically significant relationship between daily total number of patients and daily number of tweets, number of retweets, number of mentions. Also, there is a positively medium level (0.676) scientifically significant relationship between daily total number of patients and number of replies. And, there is a positive (0.246) scientifically significant relationship between between daily total number of patients and tweet lengths. Based on these data, it can be said that, on the days when the number of patients increased, information sharing (number of tweets, replies, retweets and mentions) and tweet length increased on social media (H2).

Table 5: Negative Relationships Between Daily Total Number of Patients and Social Media Data

		Daily Number of Total Patients	Average of Retweeted Tweets	Sentiment Average
Daily Number of Total Patients	Pearson Correlation	1	-.301“	-.136*
	Sig. (2-tailed)		.000	.019
	N	296	296	296
Average of Retweeted Tweets	Pearson Correlation	-.301**	1	-.003
	Sig. (2-tailed)	.000		.957
	N	296	296	296
Sentiment Average	Pearson Correlation	-.136*	-.003	1
	Sig. (2-tailed)	.019	.957	
	N	296	296	296

**.Correlation is significant at the 0.01 level (2-tailed).
*.Correlation is significant at the 0.05 level (2-tailed).

There is a moderate (-0.301) negatively scientifically significant relationship between the total number of patients and the average number of retweets of daily tweets. Thus, there is a negatively (-0.136) scientifically significant relationship between total number of patients and mood. It can be concluded that the increase in the total number of patients makes people unhappy and reduces social media interaction.

Table 6: Relationships Between Total Number of Deaths and Social Media Data

		Daily Number of Total Deaths	Number of Tweets	Number of Replies	Number of Retweets	Number of Mentions	Average Length of Tweets	Average of Retweeted Tweets
Daily Number of Total Deaths	Pearson Correlation	1	.761**	.733**	.695**	.778**	.039	-.456**
	Sig. (2-tailed)		.000	.000	.000	.000	.509	.000
	N	290	290	290	290	290	290	290
Number of Tweets	Pearson Correlation	.761**	1	.918**	.960**	.953**	.167**	-.222**
	Sig. (2-tailed)	.000		.000	.000	.000	.004	.000
	N	290	296	296	296	296	296	296
Number of Replies	Pearson Correlation	.733**	.918**	1	.804**	.867**	-.026	-.309**
	Sig. (2-tailed)	.000	.000		.000	.000	.651	.000
	N	290	296	296	296	296	296	296
Number of Retweets	Pearson Correlation	.695**	.960**	.804**	1	.935**	.299**	-.135*
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.020
	N	290	296	296	296	296	296	296
Number of Mentions	Pearson Correlation	.778**	.953**	.867**	.935**	1	.251**	-.268**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000
	N	290	296	296	296	296	296	296
Average Length of Tweets	Pearson Correlation	.039	.167**	-.026	.299**	.251**	1	.178**
	Sig. (2-tailed)	.509	.004	.651	.000	.000		.002
	N	290	296	296	296	296	296	296
Average of Retweeted Tweets	Pearson Correlation	-.456**	-.222**	-.309**	-.135*	-.268**	.178**	1
	Sig. (2-tailed)	.000	.000	.000	.020	.000	.002	
	N	290	296	296	296	296	296	296

**.Correlation is significant at the 0.01 level (2-tailed).
*.Correlation is significant at the 0.05 level (2-tailed).

When Table 6 is reviewed, five of the relationships between the total number of deaths and social media interaction variables (Table 2) are positive; one was found to be negatively correlated (H1). There is a positively high (0.761, 0.733, 0.778) scientifically significant relationship between total number of deaths and daily number of tweets, number of replies, the number of mentions, tweet lengths. There is a moderately positive (0.695) scientifically significant relationship between total number of deaths and the number of retweets. Based on this, it can be found that, on the days when the number of deaths increased, information sharing (tweet, reply, retweet and mention numbers) and tweet length increased (H2). There is also a moderate (-0.456) negatively scientifically significant relationship between the total number of deaths and the average number of retweets of tweets. It can be deduced from these relation that people tend not to spread tragic news such as the number of deaths.

Table 7: Relationships Between Daily Number of Patients and Social Media Data

		Number of Daily Patients	Number of Tweets	Number of Replies	Number of Retweets	Number of Mentions	Average of Retweeted Tweets
Number of Daily Patients	Pearson Correlation	1	.574**	.627**	.430**	.551**	-.316**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	280	280	280	280	280	280
Number of Tweets	Pearson Correlation	.574**	1	.918**	.960**	.953**	-.222**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	280	296	296	296	296	296
Number of Replies	Pearson Correlation	.627**	.918**	1	.804**	.867**	-.309**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	280	296	296	296	296	296
Number of Retweets	Pearson Correlation	.430**	.960**	.804**	1	.935**	-.135*
	Sig. (2-tailed)	.000	.000	.000		.000	.020
	N	280	296	296	296	296	296
Number of Mentions	Pearson Correlation	.551**	.953**	.867**	.935**	1	-.268**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	280	296	296	296	296	296
Average of Retweeted Tweets	Pearson Correlation	-.316**	-.222**	-.309**	-.135*	-.268**	1
	Sig. (2-tailed)	.000	.000	.000	.020	.000	
	N	280	296	296	296	296	296

**.Correlation is significant at the 0.01 level (2-tailed).

*.Correlation is significant at the 0.05 level (2-tailed).

There is a negatively (-0.316) and scientifically significant relationship between the number of daily patients and the average number of retweets of tweets. This relationship supports our interpretation about the total number of patients and demonstrates that the increase in the number of patients negatively affects the use of social media.

Table 8: Relationships Between Daily Number of Deaths and Social Media Data

		Daily Number of Deaths	Number of Tweets	Number of Replies	Number of Retweets	Number of Mentions	Average Length of Tweets	Average of Retweeted Tweets
Daily Number of Deaths	Pearson Correlation	1	.816**	.805**	.738**	.839**	.159**	-.354**
	Sig. (2-tailed)		.000	.000	.000	.000	.008	.000
	N	280	280	280	280	280	280	280
Number of Tweets	Pearson Correlation	.816**	1	.918**	.960**	.953**	.167**	-.222**
	Sig. (2-tailed)	.000		.000	.000	.000	.004	.000
	N	280	296	296	296	296	296	296
Number of Replies	Pearson Correlation	.805**	.918**	1	.804**	.867**	-.026	-.309**
	Sig. (2-tailed)	.000	.000		.000	.000	.651	.000
	N	280	296	296	296	296	296	296
Number of Retweets	Pearson Correlation	.738**	.960**	.804**	1	.935**	.299**	-.135*
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.020
	N	280	296	296	296	296	296	296
Number of Mentions	Pearson Correlation	.839**	.953**	.867**	.935**	1	.251**	-.268**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000
	N	280	296	296	296	296	296	296
Average Length of Tweets	Pearson Correlation	.159**	.167**	-.026	.299**	.251**	1	.178**
	Sig. (2-tailed)	.008	.004	.651	.000	.000		.002
	N	280	296	296	296	296	296	296
Average of Retweeted Tweets	Pearson Correlation	-.354**	-.222**	-.309**	-.135*	-.268**	.178**	1
	Sig. (2-tailed)	.000	.000	.000	.020	.000	.002	
	N	280	296	296	296	296	296	296

**.Correlation is significant at the 0.01 level (2-tailed).
*.Correlation is significant at the 0.05 level (2-tailed).

When Table 8 is reviewed, five of the relationships between the daily number of deaths and social media interaction variables (Table 2) are positive; one was found to be negatively correlated (H1).

There is a positively high (0.816, 0.805, 0.738, 0.839) scientifically significant relationship between daily number of deaths and daily number of tweets, number of replies, number of retweets, number of mentions.

There is a positively (but low) significant (0.159) relationship between daily numbers of deaths and the average tweet length (H2).

There is a negatively (-0.354) and scientifically significant relationship between the daily number of deaths and the average number of retweets of tweets. From these data, it can be interpreted that people are exposed to less interaction when sharing sad news.

Table 9: Relationships Between the Daily Number of Recovering Patients and Social Media Data

		Daily Number of Recovered Patients	Number of Tweets	Number of Replies	Number of Retweets	Number of Mentions	Average Length of Tweets	Average Number of Likes	Average of Retweeted Tweets
Daily Number of Recovered Patients	Pearson Correlation	1	.642**	.577**	.654**	.714**	.307**	-.166**	-.224**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.006	.000
	N	272	272	272	272	272	272	272	272
Number of Tweets	Pearson Correlation	.642**	1	.918**	.960**	.953**	.167**	.056	-.222**
	Sig. (2-tailed)	.000		.000	.000	.000	.004	.340	.000
	N	272	296	296	296	296	296	296	296
Number of Replies	Pearson Correlation	.577**	.918**	1	.804**	.867**	-.026	.108	-.309**
	Sig. (2-tailed)	.000	.000		.000	.000	.651	.065	.000
	N	272	296	296	296	296	296	296	296
Number of Retweets	Pearson Correlation	.654**	.960**	.804**	1	.935**	.299**	.022	-.135*
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.709	.020
	N	272	296	296	296	296	296	296	296
Number of Mentions	Pearson Correlation	.714**	.953**	.867**	.935**	1	.251**	.031	-.268**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.596	.000
	N	272	296	296	296	296	296	296	296
Average Length of Tweets	Pearson Correlation	.307**	.167**	-.026	.299**	.251**	1	-.160**	.178**
	Sig. (2-tailed)	.000	.004	.651	.000	.000		.006	.002
	N	272	296	296	296	296	296	296	296
Average Number of Likes	Pearson Correlation	-.166**	.056	.108	.022	.031	-.160**	1	-.304**
	Sig. (2-tailed)	.006	.340	.065	.709	.596	.006		.000
	N	272	296	296	296	296	296	296	296
Average of Retweeted Tweets	Pearson Correlation	-.224**	-.222**	-.309**	-.135*	-.268**	.178**	-.304**	1
	Sig. (2-tailed)	.000	.000	.000	.020	.000	.002	.000	
	N	272	296	296	296	296	296	296	296

**.Correlation is significant at the 0.01 level (2-tailed).
*.Correlation is significant at the 0.05 level (2-tailed).

When Table 9 is reviewed, five of the relationships between the daily number of recovering patients and social media interaction variables (Table 2) are positive; and two of them were found to be negatively correlated (H1).

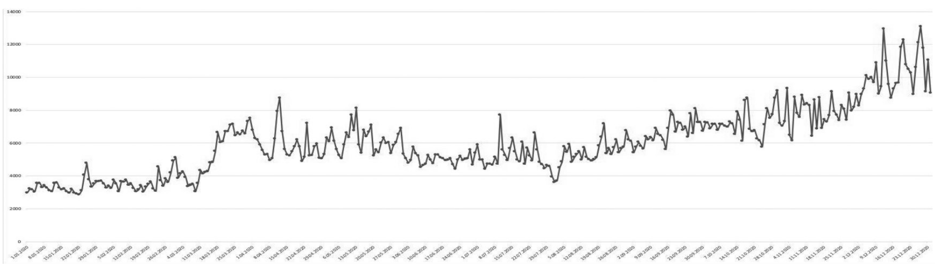
There is a positively medium level (0.642, 0.577, 0.654) scientifically significant relationship between daily number of recovering patients and daily number of tweets, number of replies, number of retweets, number of mentions, average tweet length. There is a positively high (0.714) scientifically significant relationship between the daily number of recovering patients and the number of mentions. There is a positively low (0.307) scientifically significant relationship between the daily number of recovering patients and the average number of tweets. (H2).

There is a negatively (-0.166, – 0.224) low scientifically significant relationship between the daily number of recovering patients and the average number of likes and tweets retweeted. Based on all these positive and negative scientifically significant relationships, it can be said that the pandemic has an effect on the rate of people's use of social media. In addition, it reveals that the level of social emotion and the reflection of these emotions can be measured by the data on social media during the pandemic process (H3).

Based on these results, it is thought that the emotion analysis algorithm (Coskun, Özturan 2018) that were used in this study and the analyzed big data are valid. Since data validity is provided, impact of important days and events on society in Turkey (comparision of magnitude of effect and intensities of effects of the events) has been reviewed by making sentiment analysis with data obtained.

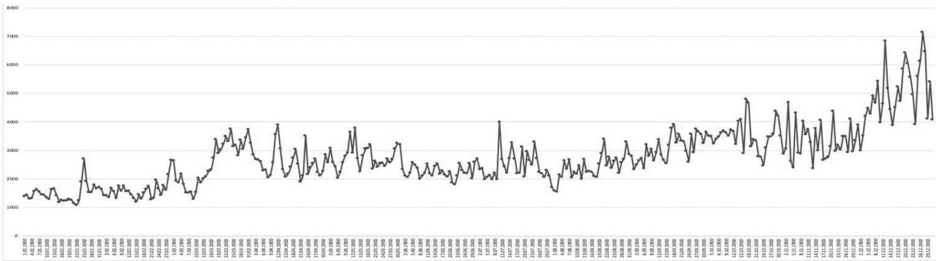
2. Validity Analysis of Relationships Between Social Media Data and Events That Are Lived:

In this section, in order to review whether the data used in the research support the results that are suitable for the purpose of the research, the relationships between the social media data and the events that took place within the selected time period (2020) were tried to be revealed (reliability check). It has been observed that important social events within the time period can be determined with the research data. The validity evidence obtained is presented below.



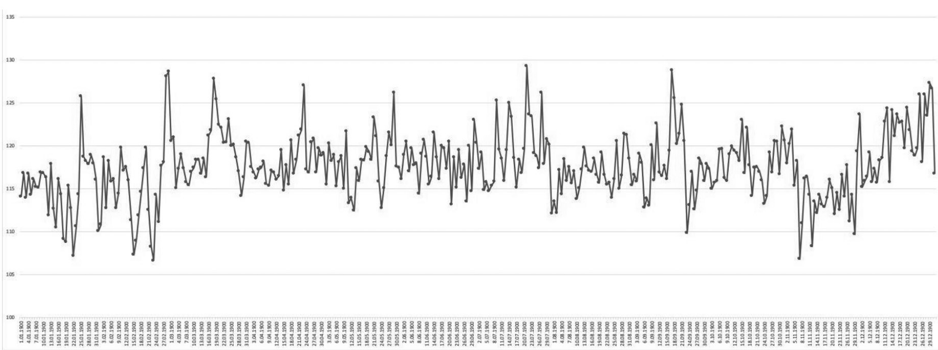
Graphic 3: *Change in Number of Tweets by Days*

Sudden increases occurred in the number of tweets shared on April 12 and December 12. On April 12, 2020, Minister of Interior Süleyman Soylu announced that he resigned by accepting his responsibility for the events that occurred after the curfew was announced to the public 2 hours before on April 10. Süleyman Soylu's resignation was not accepted by President Recep Tayyip Erdoğan. In the light of this information, it can be interpreted that both the first curfew and the resignation of the Minister of Interior greatly increased the number of tweets shared by people. On December 12, 2020, In the table published by the T.C. Ministry of Health, 20,191 patients recovered on December 12, while the total number of people recovered increased by more than 1 million in one day. Following the reactions on the social media, the section where "the number of patients recovered" was changed to "the number of people who recovered" in the table. We can see the reaction on social media by looking at the sudden increase in the number of tweets posted.



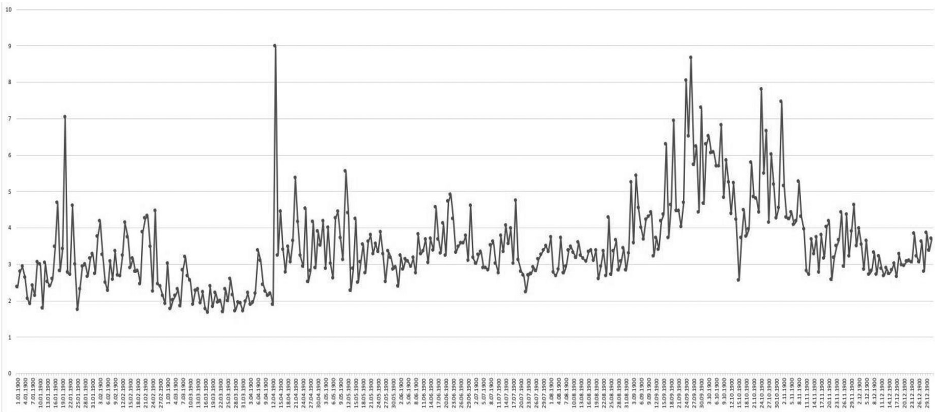
Graphic 4: *Change in the Number of Retweets by Days*

Additional measures were introduced to airports on January 24, 2020. (thermal cameras, etc.) Passengers who are coming from China began to be subjected to additional scans and it was reported that those showing signs of coronavirus would be quarantined. Based on the instant increase in the number of retweets on the same date, it was interpreted that people resorted to retweet to share these developments with their surroundings and spread the news. The increase in the number of retweets on 16 March 2020, when various restriction announcements were made, proves this interpretation. There is a sudden increase in the number of replies and tweets on April 12, when the Minister of Interior resigned. On December 12, when the change was made in the coronavirus table which is published by the Ministry of Health, there was a large increase in the number of retweets. This is another proof that users frequently use retweet to spread news.



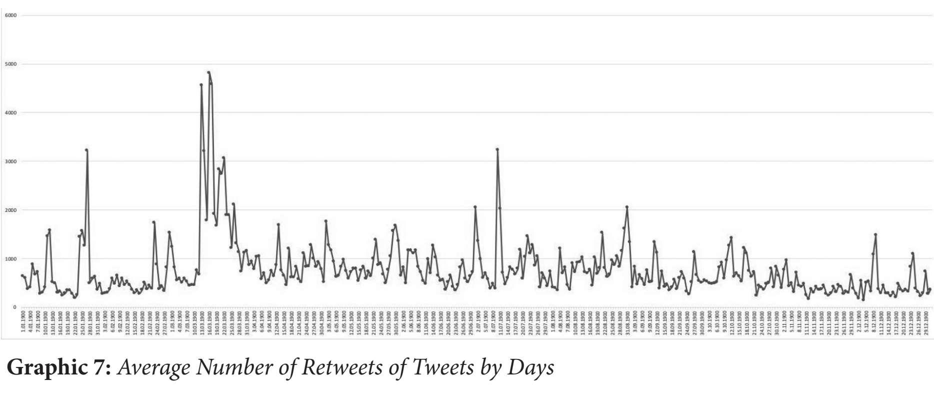
Graphic 5: *Change in Average Tweets Length by Days*

After 34 soldiers were martyred in Idlib on February 27, 2020, people expressed their feelings by sending shorter tweets. The length of the tweets posted on 18 March and 23 April 2020 due to the national holiday increased. Since the celebrations could not be held face to face increased the tendency towards social media. On March 19, 2020, ÖSYM postponed 9 exams, including TUS and MSU, which were planned to be held in the near future. It can be seen that people express their feelings and thoughts longer after this event.



Graphic 6: *Change in Average Number of Likes by Days*

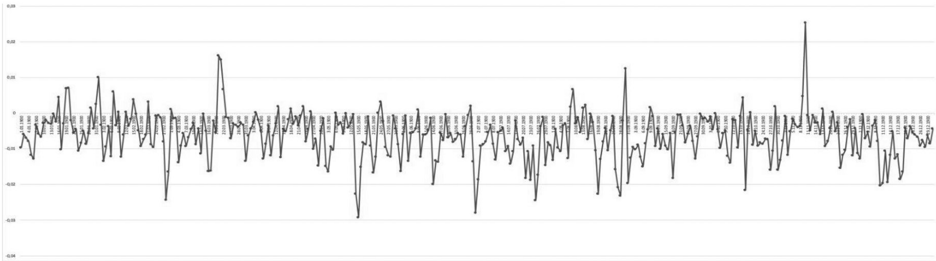
On April 13, 2020, a very clear increase was observed in the average number of likes. The reason for this situation may be that the resignation made by the Minister of Interior on April 12, 2020 was not accepted. On October 29, 2020, the average number of likes increased due to the national holiday.



Graphic 7: *Average Number of Retweets of Tweets by Days*

On March 16, 2020, various measures were taken, such as the temporary closure of places such as cinema, concert hall, wedding hall, cafe, coffee house, massage parlor, gymnasium, and interruption of prayers performed in congregation. In order to spread these news, people turned to retweets. On March 19, 2020, ÖSYM postponed 9 exams, including TUS and MSU, which were planned to be held in the near future. It has been observed that people try to spread this news to their surroundings through retweet. People have retweeted the developments regarding

Hagia Sophia as a mosque, which was on the agenda on 11 July 2020, and spread them around.



Graphic 8: *Change of Moods of People by Days*

May 15, 2020 has been observed as the day when is reacted most. It has been observed that there was not implementation of curfew for children between the ages of 15-20 between 11.00 and 15.00 on that day was not welcomed by the users. On 3 July 2020, the day that is reacted second most, an explosion occurred in a fireworks factory in Sakarya, 127 people were injured and 7 people died. The opening of the Hagia Sophia Mosque on July 24, 2020 and performing first prayer in congregation made this day the day when is reacted third most. The martyrdom of our 34 soldiers in Idlib on February 27, 2020 made February 27 the day is reacted fourth most. The reason why there was a sudden increase in emotional state in Graphic 8 on November 10 is that people talk about how much they love Atatürk in their tweets and that the love is described as a positive feeling (H4).

CONCLUSION AND DISCUSSION

In this study, the reactions which are reflected from society during COVID-19 pandemic on social media and based on these reactions, the relationships between pandemic data and moods of social media users with features of their social media accounts are analyzed in Turkey. In the study, the remaining 2891 Twitter users were included in the sample as a result of the filtering made to find real users from the users following the official Twitter account of the T.C. Ministry of Health. All tweets of these users for the year 2020 (2642347) and the features of these tweets are included in the analysis.

As a result of the correlation analysis made with this data set, it was concluded that there was a statistically significant relationship between pandemic data and several social media interaction variables. The number of tweets, retweets and mentions, which are among the social media variables, showed a high positive correlation with all pandemic data. It was also observed that the number of replies also showed a high positive correlation with all pandemic data except the daily numbers of patient.

With the increased sentiment intensity of the users, the length of the tweets also increased. This shows that people share their feelings in detail on social media. The reason for this may be that many people cannot find the opportunity to share their feelings with their surroundings and turn to social media, especially during the pandemic period.

Moreover, it was determined that the results which obtained in the study illuminated the research purpose what extent by making validity analysis. It has been concluded that the data set serves the purpose of the research by seeing that important social events that were lived can be determined with the research data.

Especially on the days when there are social events that concern the whole country, social media interactions have been observed to increase significantly compared to the days of individual events. It can be interpreted that this is caused from the need for unity and awareness in social events.

Among the research results, the finding that people use social media noticeably more, especially in times of crisis, draws attention. It can be said that the reasons for this are the needs both to obtain information and to be heard. Humans are social beings and no matter how restricted they are, they need to reveal this aspect in order to protect his psychology. They did this through social media during the pandemic period. During the pandemic period when activities like protests etc. were restricted, people used social media to react to various events and to create community. The need for people to be united with other people who share their thoughts and belong to a community is also considered to be a major factor here.

As a result, it is possible to review the dynamics of the society with the analysis of the social media big data obtained during the COVID-19 pandemic, which is an important social event; It has been seen that social events, tendencies, movements and their results can be predicted through these analyzes. It is thought that social media data mining methods can be used in developing policies for the benefit of humanity and in carrying out preventive social services.

SUGGESTIONS FOR FUTURE STUDIES

In the study, while filtering the users whose data will be taken, the profile photo of these users was determined as a "default" elimination criterion, but when applied, it was seen that too many real users who could benefit from the study were also eliminated. Therefore, this feature has been abandoned as a criterion. However, it is anticipated that there may be meaningful relationships between the account properties of users by defining account photos as "default". It is predicted that the examination of this situation in future studies will yield scientifically significant results.

In addition, the reason of the result of the negative mood of a society whose majority is Muslim on the day Hagia Sophia was used as a mosque is not known whether people prioritize their health before their religious beliefs or are influenced by other social dynamics. It is thought that it would be appropriate to make more detailed studies on this subject in future studies.

REFERENCES

Ahmad, A. R., & Murad, H. R. (2020). The impact of social media on panic during the COVID19 pandemic in Iraqi Kurdistan: online questionnaire study. *Journal of medical Internet research*, *22*(5), e19556.

Akbulut, M., Sahin, U., & Esen, A. C. (2020). More than a Virus: How COVID 19 Infected Education in Turkey? *Journal of Social Science Education*, *19*, 30-42.

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., ... & Scala, A. (2020). The covid-19 social media infodemic. Scientific Reports, 10(1), 1-10.

Coşkun, M., & Ozturan, M. (2018). # europehappinessmap: A framework for multi-lingual sentiment analysis via social media big data (a Twitter case study). *Information*, *9*(5), 102.

Ferrara, E. (2020). What types of covid-19 conspiracies are populated by twitter bots? *First Monday*.

Karaşar, B., & Canlı, D. (2020). Psychological resilience and depression during the Covid-19 pandemic in Turkey. Psychiatria Danubina, 32(2), 273-279.

Li, D., Chaudhary, H., & Zhang, Z. (2020). Modeling spatiotemporal pattern of depressive symptoms caused by COVID-19 using social media data mining. *International Journal of Environmental Research and Public Health*, *17*(14), 4988.

Vural, A. G., Cambazoglu, B. B., Senkul, P., & Tokgoz, Z. O. (2013). A framework for sentiment analysis in turkish: Application to polarity detection of movie reviews in turkish. In *Computer and Information Sciences III* (pp. 437-445). Spring