Research Article

Academic Platform Journal of Engineering and Smart Systems (APJESS) 11(3), 174-183, 2023

Received: 21-Feb-2022 Accepted: 18-Sep-2023

homepage: https://dergipark.org.tr/tr/pub/apjess

https://doi.org/10.21541/apjess.1076756

An Explorative Analysis of Tweets Sentiments for Investment Decision in Stock Markets

[1]Emine ATEŞ, *[2]Aysun GÜRAN

[1] Department of Computer Engineering, Dogus University, İstanbul, Türkiye, emine.ates.34@gmail.com

[2] Department of Computer Engineering, Dogus University, İstanbul, Türkiye, adogrusoz@dogus.edu.tr

Abstract

Nowadays lots of researches report that positive or negative social media posts have significant effects on stock market returns. In particular, as well as long-term effects, it may be important to identify the short-term effects of public opinions on the stock market. Unfortunately, the number of studies conducted in this context for Turkey is very limited. Therefore, this study investigates the relationship between Turkish tweets and the trend of Bist30 index returns for short-term analysis. Correlation analysis is a statistical method used to measure the strength and direction of relationship between two variables. The existence of a short-term correlation between sentiments of financial tweets and the stock market returns may be an indicator that social media can be used as a resource to predict the sudden changes of stock markets returns. In this study, the correlation analysis steps are conducted on the polarity scores and changes in Bist30 index returns with all the necessary statistical tests. The contribution of our study is that analyzes are made on a daily, weekly and monthly basis. The experimental results show that there are significant positive correlations between the sentiment polarity values and changes in Bist30 index returns for even short-term analysis. As a result, this study points out a useful pathway for the future researches to show that social media posts may convey useful information for financial markets even for short term analysis.

Keywords: Correlation analysis, sentiment analysis, Turkish tweets, Bist30

1. INTRODUCTION

In our age where a customer-oriented approach prevails, financial enterprises have started to give more importance to the use of social media. Not only financial enterprises but also individual investors try to reach positive or negative general opinions about companies by following the news shared on social media. They update their asset portfolios based on this information and/or may turn to new investment channels. Liu et al. [1] state that people invested in stocks that are familiar to them. One of the reasons that feeds the familiarity of stocks is the news sources shared about financial markets [2]. Social media enables developments to spread much faster than many sources and allows more ideas to be collected at the same time [3]. Sprenger et al. [4] state that individual investors who follow Twitter are more profitable.

Nowadays, when the use of social media is so widespread, the number of studies examining whether positive or negative comments shared about financial businesses create a dynamism in financial markets is increasing. Mao et al. [5] investigate whether the daily number of tweets can be used to predict the S&P 500 stock market index over a period of 4 months. Sprenger et al. [6] report a relationship between stock-related tweets sentiments and stock market returns on a daily basis. Valle-Cruze et al. [7] analyzes the behavior of stock markets based on tweets during the COVID-19 pandemic and reports that the markets reacted 0 to 10 days after the tweets were shared. Karabulut [8] states that Gross National Happiness index of Facebook is efficient to estimate US daily returns. The number of studies conducted in this context for Turkey is very limited [2, 9, 10, 11]. These studies did not evaluate the analyzed period by dividing it into shorter periods. Based on such studies, the main motivation in this study is to investigate whether there is a correlation between the sentiments of tweets and the trend of the Bist30 index returns on a daily, weekly and monthly basis. Thus, it will be possible to investigate whether there are significant short-term correlation results between social media posts and the trend of stock markets. Additionally, it can be seen whether social media platforms provide useful information for short term analysis. Therefore, our aim is to investigate the relationship between the Turkish tweets related to Borsa Istanbul 30 (Bist30) companies and the trend of the Bist30 index returns on a daily, weekly and monthly basis.

In order to determine the existence of the relationship between Bist30 index returns and tweets, the tweets are labeled as positive, negative or neutral after passing through sentiment analysis stages. For this process, dictionary-based or machine learning-based approaches can be used. In dictionary-based approaches, textual data is labeled using domain-independent sentiment dictionaries such as Afinn [12], OpinionMiner [13], and Sent WordNet [14]. Although these dictionaries speed up the labeling process, they may be insufficient because they cannot carry the variation of emotions in special areas or different cultures. Loughran and McDonalds [15] state that the results of the analysis are affected negatively, if general-purpose sentiment dictionaries are used to classify financial text data. Unfortunately, there is no open-source Turkish emotion dictionary created specifically for the finance field. For all these reasons, to create the training dataset, the opinions of financial investors are taken as a basis in the labeling phase of social media posts. During the period between 07.05.2018 and 30.04.2019, when our study was started, 57933 tweets shared on Twitter regarding the stocks listed in the Bist30 index are labeled as positive, negative and neutral by financial investors manually. By using these labeled tweets as a training dataset, a test dataset consisting of two hundreds thousand tweets from the same time period are assigned to three classes as positive, negative and neutral by different machine learning algorithms (Support Vector Machines (SVM), Logistic Regression (LR), Naïve Bayes (NB), Decision Trees (DT), K-Nearest Neighborhood (KNN) classifiers and Majority Voting (MO) ensemble learner). Actually, we conducted these stages in our previous study [11]. This study differs from our previous study [11] in that the analyzes are carried out daily, weekly and monthly. We conducted the experiments for daily, weekly and monthly time periods in the time interval (07.05.2018 - 30.04.2019) in which tweets are examined. Hence, we can see whether social media platforms provide useful information for short term analysis. By using the number of tweets under the positive and negative categories, we calculate sentiment polarity scores of the tweets for daily, weekly and monthly basis. After that, the correlation analysis steps are applied on the polarity scores and the trend of Bist30 index returns for daily, weekly and monthly basis analysis with all the necessary statistical tests. The experimental results show that there are significant correlations between the sentiment polarity values and the changes of Bist30 index returns. We think this study points out a useful pathway for the future researches to show that social media posts may also convey useful information for short-term analysis.

We present our study in this paper with six more sections, which include related studies, dataset and methods, results, and conclusion.

2. RELATED WORKS

Researches that analyze whether there is a relationship between the sentiment values of tweets and stock market indexes' returns make use of methods such as sentiment analysis, correlation analysis and Granger causality analysis.

Sentiment analysis aims to evaluate the feelings and thoughts of communities on specific or general subjects. The studies most dominantly use machine learning approaches to extract positive, negative or neutral opinions of communities. Kaukaras et al. [16] apply sentiments analysis on financial tweets with seven different machine learning algorithms. De Oliveira Carosia et al. [17] utilize artificial neural network architecture to perform sentiment analysis in financial news in Brazilian Portuguese and state that investment strategies based on sentiment analysis can bring profitability for investors. Utilizing financial related reviews on Eastmoney.com, Qui et al. [18] proposes a modified sentiment index depend on the number of positive, negative and neutral reviews that are extracted by Baidu AI Cloud sentiment analysis model. Yue et al. [19] presents a survey paper about sentiment analysis in social media. Hamraoui and Boubaker [20] investigate the correlations between tweets and the Tunisian financial market considering over a 12-month period and show that tweets can be utilized for price volatility. By using dictionary-based software, Bollen et al. [21], divided approximately 9.8 million tweets shared in a 10-month period into six sentiment categories (Calm, Alert, Sure, Vital, Kind, and Happy). They stated that there was a Granger causality between the tweets and the Dow Jones Index (DJIA) return values. Similarly, Mittal and Goel [22] examined the relationship between DJIA values and 476 million Twitter posts divided into 4 different sentiment categories with a dictionary-based approach and stated that there was a Granger causality between DIJA index returns and public moods captured on tweets. The paper reported that two public moods 'calm' and 'happy' could be used to predict future DJIA movements.

Eliaçık and Erdoğan [9] labeled 2408 Twitter posts, covering a 6-month period, with positive and negative labels, and proposed an approach that demonstrates the credibility of the person who shared the post in order to calculate the sentiment polarity values of tweets. As a result, they found a significant correlation between the proposed sentiment polarity value and the weekly value changes of the Bist100 index. In addition, they reported that the Pearson coefficient increased when the extraordinary events occurred during the collection period of the data set.

Yıldırım and Yüksel [10] classified 500 Twitter posts collected about a company in the telecommunications sector and traded on Borsa Istanbul with positive and negative tags using machine learning methods. They performed the daily correlation analysis between the sentiment polarity values obtained from tweets classified in their study and the firm's stock values and reported that there was a moderate and negative relationship between the two samples.

Ateş and Güran [11] applied the correlation and Granger causality analysis methods for Turkish financial tweets. They utilize different text representations (tokens, n-grams, Doc2Vec combined vectors etc.) and apply six different machine learning algorithms to assign the Turkish financial-related tweets into positive, negative and neutral classes. Based on the sentiment analysis of Turkish financial-related tweets, the results of correlations and Granger causality analyses steps for two different time periods indicated a causality from the stock returns to tweet sentiments for a long term and significant correlations for both short and long terms.

Ranco et al. [23] labeled 100000 stock related tweets with positive, negative and neutral labels manually. By using this training dataset, they classified 1.5 million tweets with support vector machine algorithm. Based on the calculated sentiment polarity values they stated a low Pearson correlation and Granger causality results between sentiments and stocks.

After labeling 3216 posts shared on Twitter as positive, negative and neutral, Pagolu et al. [24] classified a test data set consisting of 250 thousand tweets with N-gram and Word2Vec methods. After obtaining the DIJA index changes for a period of one year, they designed a system to transform the relationship between tweets and index value changes into a classification problem by using the number of tweets under positive and negative categories. As a result, they determined that the emotional changes in tweets were effective on the index value changes.

Deng et al. [25], with a dictionary-based software, labeled approximately 18 million tweets collected from the StockTwit platform in a 4-year period as positive and negative. Afterwards, they analyzed whether there was Granger causality between the daily and hourly value changes of the DJIA index and the sentiment polarity values obtained from tweets. As a result, they stated that there is no Granger causality when the daily changes between the two-time series are taken as basis, but negative tweets in hourly analysis are effective in the causality relationship.

Zhao [26] conducted the correlation and Granger causality analysis between the Singapore Straits Times Index (STI) values, which consist of the first 30 stocks with the highest market value, and the happiness indexes obtained from Twitter posts. This happiness index is calculated by determining the word consisting of 10 thousand emotion indicators such as love, happiness, smile and fun in approximately 50 million tweets shared by Twitter users by natural language processing methods. In the end, they stated that there is a positive correlation between the Twitter happiness index and the STI Index return values covering the 4-year period, which is also supported by the Granger causality analysis.

3. DATASET AND SENTIMENT ANALYSIS PROCESS

In this study, two hundred thousand tweets posted between 07.05.2018 - 30.04.2019 are used as a corpus. By labelling 57933 tweets manually, the study applied sentiment analysis steps on two hundred thousand test tweets and labeled those tweets as positive, negative and neutral by using LR, SVM, NB, DT, KNN and MV classifiers. Before classifiers are applied, words starting with '@', '$' characters, expressions containing URL addresses and punctuation marks in tweets are cleaned. Tokens such as '%+', '+%', '-%', '%-', ':)', ':(' are not removed because they are assumed to reflect positive/negative moods. In all experiments where the performance values of different classifiers are compared, 80% of the training dataset is used for training and 20% for testing. In addition, 10-fold cross-correction method is applied. Python programming language is used during the experiments. The linear kernel function is chosen for the SVM algorithm. In the KNN method, the Euclidean distance is taken as a basis, taking into account the 3 nearest neighbors. The Gini index is used for DT classifier. In MV ensemble learning, majority voting strategy is used. Classification results of 57933 tweets vectorized with the use of 50 thousand features with high tf-idf values are as indicated in Table 1. At this stage, apart from word-based 1-grams, the effects of 2-grams and 3-grams are also examined.

Table 1. Classification results of 57933 tweets

	1-Gram	2-Gram	3-Gram
LR	%63,41 (+/- 0,28)	**%63,60 (+/- 0,29)**	%63,47 (+/- 0,24)
SVM	%62,53 (+/- 0,48)	%62,32 (+/- 0,54)	%62,32 (+/- 0,47)
NB	%59,55 (+/- 0,32)	%61,58 (+/- 0,43)	%61,40 (+/- 0,56)
KA	%58,01 (+/- 0,30)	%57,60 (+/- 0,36)	%57,47 (+/- 0,41)
KNN	%56,16 (+/- 0,45)	%55,43 (+/- 0,45)	%55,34 (+/- 0,49)
MV	% 63,03 (+/- 0,24)	% 63,39 (+/- 0,22)	% 63,29 (+/- 0,20)

As it can be seen from Table 1, the best performance value belongs to the LR algorithm with 63.60% accuracy value. The best performance result has been obtained with 2-grams. Many experiments are carried out with the majority voting method, which is the ensemble learning method, and different combinations and individual classifiers are used. The best performance results are obtained by combining LR, SVM and NB algorithms (63.39%). However, this result is still lower than LR classifier. After this step, two hundred thousand tweets are classified with LR algorithm based on 2-grams and the number of positive and negative tweets are detected. Figure 1-3 show the daily, weekly and monthly distribution of the classified tweets that are posted between 07.05.2018 - 30.04.2019 for short-term analysis.

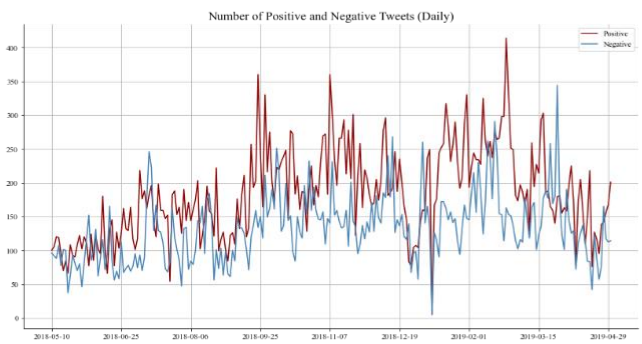

Figure 1. Distribution of the number of tweets per day

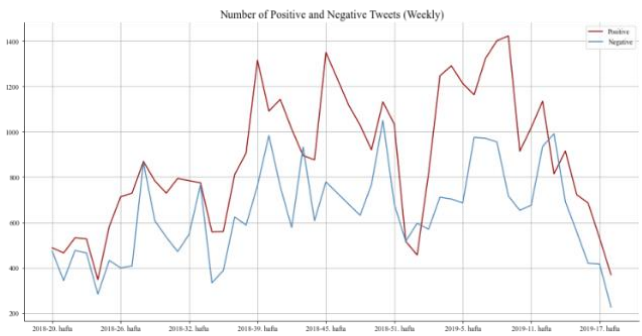

Figure 2. Distribution of the number of tweets per week

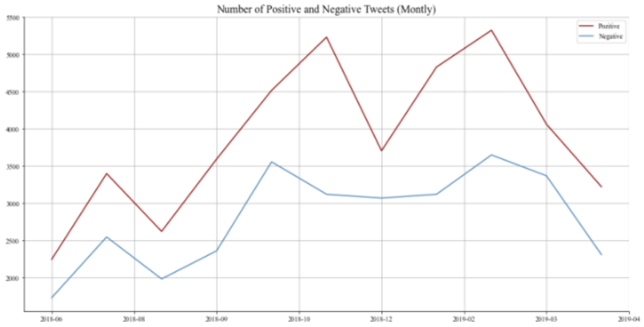

Figure 3. Distribution of the number of tweets per month

Detecting the number of positive and negative tweets per day, per week and per month, sentiment polarity values are calculated for each time period. We use the following equations for calculating the sentiment polarity values on a daily, weekly and monthly basis.

$$Sd = \frac{p_{td}}{p_{td}+n_{td}} - \frac{p_{td-1}}{p_{td-1}+n_{td-1}} \quad (1)$$

$$Sw = \frac{p_{tw}}{p_{tw}+n_{tw}} - \frac{p_{tw-1}}{p_{tw-1}+n_{tw-1}} \quad (2)$$

$$Sm = \frac{p_{tm}}{p_{tm}+n_{tm}} - \frac{p_{tm-1}}{p_{tm-1}+n_{tm-1}} \quad (3)$$

where p_{ti} and $, n_{ti}$ indicate the number of positive and negative tweets (i= d, w, m). After this step, we extract the daily, weekly and monthly movements of the closing prices of Bist30 with the following equations (Eq 4-5-6):

$$R_d = \frac{P_d - P_{d-1}}{P_{d-1}} \quad (4)$$

$$R_w = \frac{P_w - P_{w-1}}{P_{w-1}} \quad (5)$$

$$R_m = \frac{P_m - P_{m-1}}{P_{m-1}} \quad (6)$$

where P_d , P_w and P_m are the closing prices of the stock at day d, week w and month m respectively.

Then correlation analysis steps are performed on the sentiment polarity values and the changes of Bist30 index returns.

4. CORRELATION ANALYSIS

Pearson Correlation is a statistical analysis used to depict the strength and direction of relationship between two variables. In correlation analysis, the degree of correlation is shown with the correlation coefficient (r) that lies between -1 and +1. Positive one indicates a perfect positive linear relationship between the variables; whereas negative one points out a perfect negative association between variables and a value of zero coefficient indicates no correlation. The absolute value of r indicates the strength of the existing relationship. Although the strength may vary by discipline the following general guidelines can be used to interpret the correlation coefficient [27].

$0.1 < \lvert r \rvert < 0.3$	small / weak correlation
$0.3 < \lvert r \rvert < 0.5$	medium / moderate correlation
$0.5 < \lvert r \rvert$	large / strong correlation

Many statistical methods make assumptions about normality. In general, graphical and numerical methods can be used to assess the normality of the data [28, 29]. Skewness, kurtosis, Shapiro–Wilk test, Kolmogorov–Smirnov test, histogram and normal Q–Q plot are the most popular methods to evaluate the normal distribution.

Skewness refers to the deviation from symmetry and kurtosis reflects the degree of the sharpedness / peakedness of a distribution [18]. In order to give "excess" kurtosis SPSS subtracts 3 from the kurtosis. By using skewness and excess kurtosis, the *z*-test can be used for testing the normality. In order to calculate the z-scores, we need to divide the skew values or excess kurtosis by their standard errors. If sample size is less than 50 and the *z*-values fall between +1.96 and -1.96 for either skewness or kurtosis, with an alpha level 0.05, then the data is said to be normally distributed [30]. If sample size is between $50 \leq n < 300$ and *z*-values fall between -3.29 and +3.29 for $\alpha = 0.05$, the data is assumed to be normally distributed [31]. For sample size greater than 300, the histograms and the absolute values of skewness and kurtosis should be examined for normality of the data (the absolute skewness value that is less than or equal to 2 or the absolute kurtosis (excess) that is less than or equal to 4 could be the limit values for normality) [32].

Pearson and Spearman statistical tests are the most commonly used tests to assess whether the data is normally distributed. If sample size is less than fifty, the Shapiro–Wilk test is preferred; otherwise, Kolmogorov–Smirnov test is used [31].

Histograms are used to show the distribution of a dataset. For the normally distributed dataset, the histogram graphs are nearly bell-shaped and symmetric about the mean [32,33].

In statistics, a normal Q–Q plot shows whether a dataset comes from Gaussian distribution. If the dataset is normally distributed, data points are seen as a straight line on the graph.

In our study, the suitability of sentiment polarity measurements and the stock value changes to normal distribution were assessed with the tests specified (Skewness& Kurtosis z-values, Kolmogorov-Smirnov or Shaphiro-Wilk tests, the Histograms and Normal Q-Q plots). We use IBM SPSS [34] software to perform the required steps.

5. NORMALITY TEST AND CORRELATION ANALYSIS RESULTS

The aim of this section is to perform the correlation analysis between the sentiment polarity values (S_d, S_w, S_m) and daily, weekly and monthly movements of the closing prices of Bist30 (R_d, R_w, R_m). Before starting the correlation analysis, it is tested whether the sentiment polarity values (S_d, S_w, S_m) and (R_d, R_w, R_m) values fit the normal distribution. For normality test, we analyze different numerical and visual methods: skewness and kurtosis z-values, the Kolmogorov-Smirnov/Shapiro-Wilk tests p-values, histogram and normal Q-Q plots. These numerical and visual outputs will indicate whether our analyzed variables are approximately normally distributed.

5.1. Normality Test Results

In this section, in order to test the normality of R_i and S_i values (i=d, w, m), first of all we look at the skewness and kurtosis z-values, then the p-values of Kolmogorov-Smirnov/Shapiro-Wilk tests and finally histogram and normal Q-Q plots.

5.1.1. Daily Analysis Results

Using SPSS to test for normality of daily R_d and S_d values, first of all, we focus on skewness and excess kurtosis and divide the skew values and excess kurtosis by their standard errors to get the z-scores. According to Table 2, skewness and excess kurtosis z-scores for S_d are (|-0.040/0.156|=0.25641, |-0.238/0.312|=0.76282) respectively. Since we have 242 trading dates and the absolute values of these numbers should not exceed 3.29 with an alpha level 0.05, it can be said that S_d values are approximately normally distributed. Similarly, the skewness and kurtosis z-scores for R_d are (|-0.398/0.156|= 2.55128, |0.966/0.312|= 3.096154) respectively. Again, since the absolute values of these numbers are not exceed 3.29, we can accept the null hypothesis and say that R_d values are approximately normally distributed.

Table 2. Skewness and Kurtosis values for daily analysis

		Statistic	Std. Error
S_d	Skewness	-.040	.156
	Kurtosis	-.238	.312
R_d	Skewness	-.398	.156
	Kurtosis	.966	.312

Table 2 presents the results of Kolmogorov-Smirnov and the Shapiro-Wilk tests at a significance level of 0.05. As stated before the number of trading days considered between 07.05.2018 and 30.04.2019 was 242. Because of this, we focus on the Kolmogorov-Smirnov test as our numerical means of assessing normality. Since p values are greater than 0.05, Table 3 shows that daily S_d and R_d values are normally distributed.

Table 3. Tests of normality on a daily basis

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
S_d	.041	242	.200*	.996	242	.719
R_d	.047	242	.200*	.987	242	.028

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

As a next step, we look at the graphical figures (Figure 4 and Figure 5) for daily S_d and R_d values. The histogram diagrams show us that the S_d and R_d are approximately normally distributed.

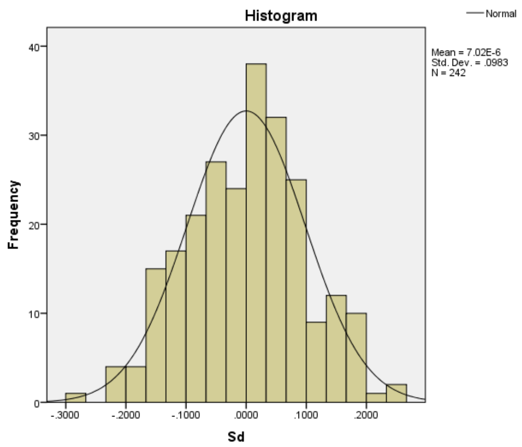

Figure 4. Histogram diagram of daily S_d

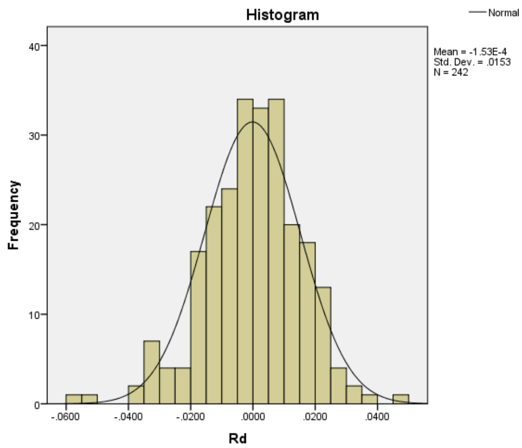

Figure 5. Histogram diagram of daily R_d

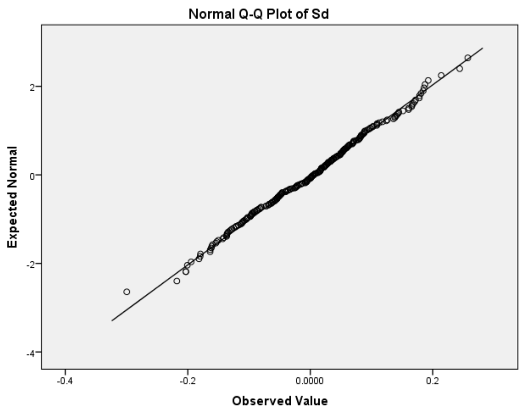

Figure 6. Q-Q plot for diagram of daily S_d

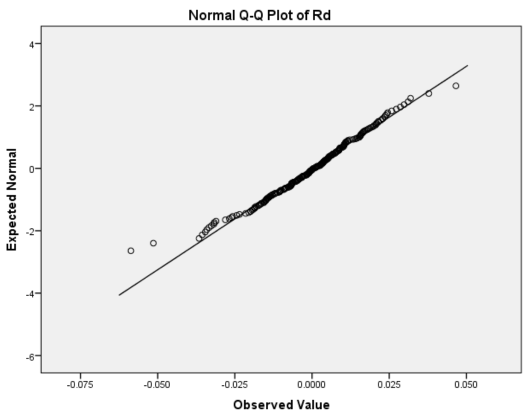

Figure 7. Q-Q plot for diagram of daily R_d

Then we look at the normal Q-Q diagrams and since the dots are along the lines, these figures indicate that S_d and R_d are approximately normally distributed.

All the results indicate that S_d and R_d are approximately normally distributed.

5.1.2. Weekly Analysis Results

Table 4 presents normality test results for weekly analysis. It is seen that skewness and excess kurtosis z-scores for weekly S_w are (|-0.522/0.337|=1.5489614, |0.315/0.662|=0.475831) respectively. Since we have 50 trading weeks and the absolute values of these numbers should not exceed 1.96 with an alpha level 0.05, it can be said that S_w values come from a normal distribution.

Table 4 also indicates that the skewness and excess kurtosis z-scores for weekly R_w are (|-0.371/0.337|=1.10089, |0.541/0.662|=0.817221) respectively. Having 50 trading weeks and since the absolute values of these numbers are not exceed 3.29 with an alpha level 0.05, we can accept the null hypothesis and say that R_w values are also approximately normally distributed.

Table 4. Skewness and kurtosis values for weekly analysis

		Statistic	Std. Error
S_w	Skewness	-.522	.337
	Kurtosis	.315	.662
R_w	Skewness	-.371	.337
	Kurtosis	.541	.662

There are several methods for normality test such as Kolmogorov-Smirnov normality test and Shapiro-Wilk's test. Since it is more convenient to use Shapiro-Wilk's test for small sample sizes, we focus on Shapiro-Wilk test statistics. The null hypothesis is determined as follows: "H0: The data is normally distributed". It will be rejected if the corresponding p-value is below 0.05. In SPSS output, the p-value is labeled as "Sig." Since the Sig. values (0.485 and 0.837) are greater than 0.05 for weekly S_w and R_w respectively, we keep the null hypothesis. That's why it can be said that S_w and R_w are approximately normally distributed.

Table 5. Tests of normality on a weekly basis

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
S_w	.114	50	.124	.978	50	.485
R_w	.078	50	.200*	.987	50	.837

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The graphical figures for S_w and R_w values can be examined with Figure 8-11. The histogram diagrams (Figure 8 and Figure 9) appear to support that the S_w and R_w are approximately normally distributed. As it is seen the diagrams are quiet symmetrical. There are big peaks and small tails on the sides.

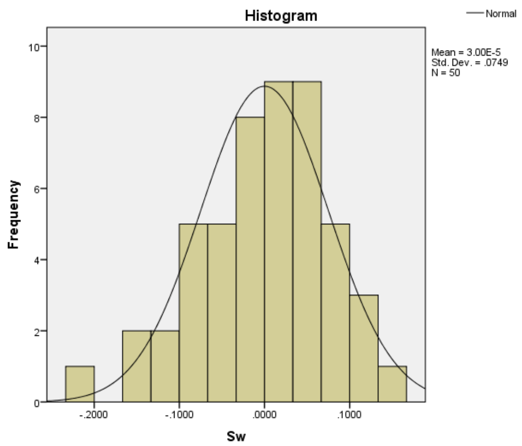

Figure 8. Histogram diagram of weekly S_w

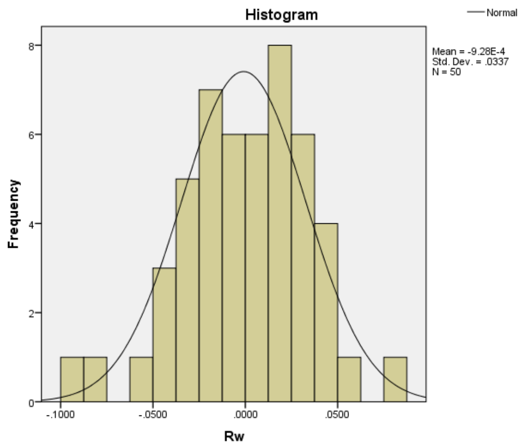

Figure 9. Histogram diagram of weekly R_w

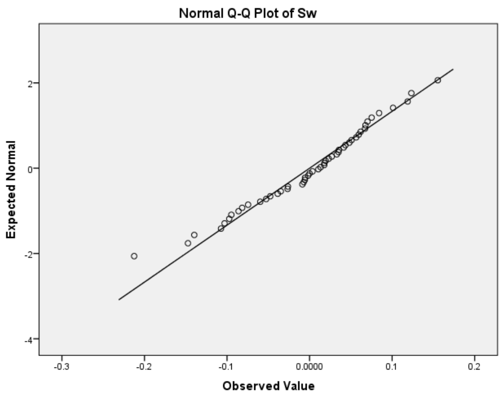

Figure 10. Q-Q plot for diagram of weekly S_w

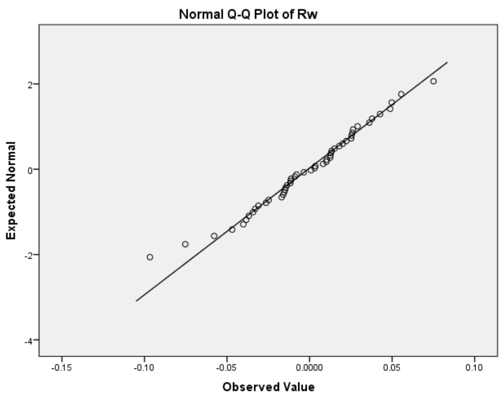

Figure 11. Q-Q plot for diagram of weekly R_w

Then we look at the normal Q-Q diagrams and since the dots are along the lines, Figure 10 and Figure 11 indicate that weekly S_w and R_w are approximately normally distributed.

All the results indicate that S_w and R_w are approximately normally distributed.

5.1.3. Monthly Basis Results

Table 6 shows normality test results for monthly analysis. Skewness and excess kurtosis z-scores for monthly S_m are (|-0.425/0.661||=0.642965, |-0.839/1.279|=0.6559812) respectively. Since we have 11 trading months and the absolute values of these numbers should not exceed 1.96 with an alpha level 0.05, it can be said that monthly S_m values are approximately normally distributed. The skewness and excess kurtosis z-scores for monthly R_m are (|-0.371/0.337|=1.10089, |0.541/0.662|=0.817221) respectively. Since we have 11 trading months and the absolute values of these numbers are not exceed 3.29 with an alpha level 0.05, we can accept the null hypothesis and say that monthly R_m values are also approximately come from a normal distribution.

Table 6. Skewness and kurtosis values for monthly analysis

		Statistic	Std. Error
S_m	Skewness	-.425	.661
	Kurtosis	-.839	1.279
R_m	Skewness	1.171	.661
	Kurtosis	2.005	1.279

It is more appropriate to use Shapiro-Wilk's test for small sample sizes, because of this reason we focus on Shapiro-Wilk test statistics for monthly analysis. "H0: The data is normally distributed" is the null hypothesis for this test of normality and it should be rejected if the Sig. Value (p-value) is less than 0.05. Ass it is seen from Table 7, the Sig. values (0.600 and 0.279) are greater than 0.05 for S_m and R_m respectively. This means we should keep the null hypothesis and assume that S_m and R_m are approximately normally distributed.

Table 7. Tests of normality on a monthly basis

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
S_m	.186	11	.200*	.947	11	.600
R_m	.172	11	.200*	.915	11	.279

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Figure 12 and 13 depict the histogram diagrams of S_m and R_m.

Based on the smallest-sized sample histogram diagrams don't point out a symmetrical bell shape.

Figure 14 and 15 show Q-Q diagrams and since the dots are along the lines, they point out that monthly S_m and R_m are approximately normally distributed.

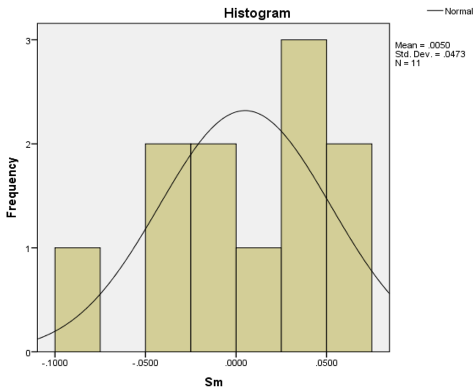

Figure 12. Histogram diagram of monthly S_m

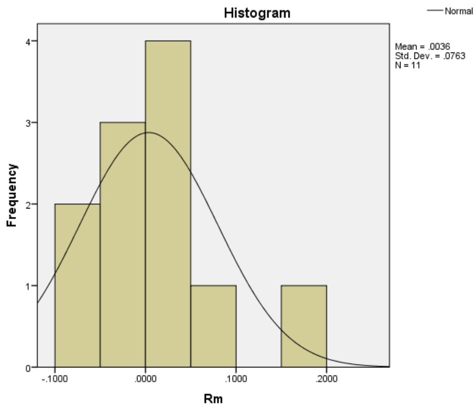

Figure 13. Histogram diagram of monthly R_m

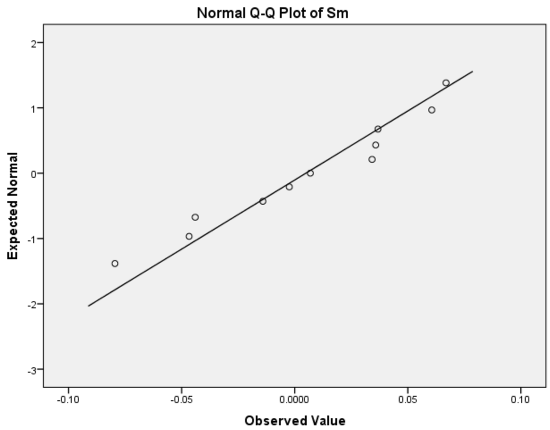

Figure 14. Q-Q plot for diagram of monthly S_m

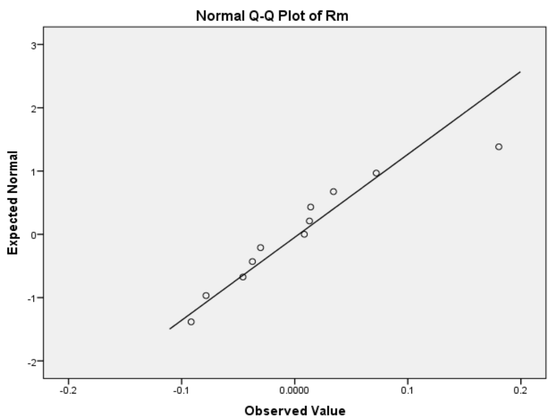

Figure 15. Q-Q plot for diagram of monthly R_m

Although the histogram diagram the belongs to the smallest-sized sample doesn't reflect a symmetrical bell shape, we could conclude that based on the other examined methods S_m and R_m satisfy the normality assumptions.

5.2. Correlation Results

We conduct various visual and statistical normality tests on our variables to investigate the assumption of normality. The results show that daily, weekly and monthly variables seem to support the normality assumptions. This section gives Pearson correlation results for daily, weekly and monthly analysis.

Table 8. Daily correlations between S_d and R_d

		S_d	R_d
S_d	Pearson Correlation	1	.509**
	Sig. (2-tailed)		.000
	N	242	242
R_d	Pearson Correlation	.509**	1
	Sig. (2-tailed)	.000	
	N	242	242

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

Table 8 depicts the correlation results among S_d and R_d values. According to Table 8, it is seen that S_d and R_d have a statistically significant linear relationship (r=.509, $p < .001$). When the direction of the relationship is analyzed, it is seen that S_d and R_d tend to increase together. Since $| r | > .5$ we can say that the association is approximately strong.

Table 9. Weekly correlations between S_w and R_w

		S_w	R_w
S_w	Pearson Correlation	1	.602**
	Sig. (2-tailed)		.000
	N	50	50
R_w	Pearson Correlation	.602**	1
	Sig. (2-tailed)	.000	
	N	50	50

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

Table 9 indicates the weekly correlation results between S_w and R_w values. As it is seen the linear relationship between S_w and R_w is also statistically significant (r=.602, $p < .001$). There is a positive association and the magnitude of the association is quite strong ($| r | > .5$).

Table 10. Monthly correlations between S_m and R_m

		S_m	R_m
S_m	Pearson Correlation	1	.707*
	Sig. (2-tailed)		.015
	N	11	11
R_m	Pearson Correlation	.707*	1
	Sig. (2-tailed)	.015	
	N	11	11

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

Table 10 presents the monthly correlation results among S_m and R_m values. S_m and R_m have a significant linear relationship. Since ($| r |= 0.707 > 0.5$) it can be said that the association is strong. Due to the positive correlation coefficient, it can be said that S_m and R_m moves together in the same direction.

6. CONCLUSION

Nowadays instead of making the analyzes suggested by classical theories, people can make their individual financial investment decisions by taking certain shortcuts. Positive and negative comments shared on social media have become one of those shortcuts that are taken into consideration when making investment decisions. With the effect of those comments, investors can revise their portfolios. In literature there are lots of researches that investigate the relations between social media posts and stock markets. Twitter is one of the most important social media platforms. Researches indicate that there is a correlation between tweets sentiments and the changes in stock prices for both long-term and short-term analysis. Unfortunately, most of the studies are related with analysis of social media platforms in English language. There are very limited number of studies related with Turkish language, especially for short-term analysis. Hence, with this motivation, the main aim of this study is to investigate the correlation between sentiments of tweets related to the Bist30 index and the index movement on daily, weekly and monthly basis.

We conduct sentiment analysis stages and correlation analysis steps to investigate whether there is a relation between tweet sentiments and returns of Bist30 index returns for short-term analysis. The analyzes are carried out for the time interval of 07.05.2018 - 30.04.2019. Before starting the correlation analysis, the normal distribution suitability of all variables are tested by using different graphical and numerical methods. After that the results of the correlation analysis between the sentiments of tweets and the change of index returns are presented. The presented results indicate that there are significant positive correlations between sentiments of tweets and the change of index returns. With these results, it has been seen that social media posts are likely to affect financial markets. The findings obtained as a result of our study are in parallel with many studies in the literature [6, 7, 8, 9, 10, 11, 22, 24, 26]. This study, which is based on short-term analysis, also supports that social media resources can be effective to predict the trend of stock markets.

For future studies, Granger causality analysis steps for short-term analysis will be studied and also we will investigate whether tweets can be an indicator that can be used to predict the returns of stock market indices. We think the present study points out a useful direction of future research to improve the understanding of the effects of social media on financial markets.

Author contributions: Emine Ateş: Data collection, Software, Literature review; Aysun Güran: Literature review, Investigation, Software, Writing-Review & Editing, Analysis of results.
Conflict of Interest: No conflict of interest was declared by the authors.
Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES

[1] Liu, L. X., Ann E. S., Yong Z., The long-run role of the media: Evidence from initial public offerings, Management Science, 60 (8), 1945-1964, 2014.

[2] Atan, S., Çinar, Y. Borsa Istanbul'da finansal haberler ile piyasa degeri iliskisinin metin madenciligi ve duygu (sentiment) analizi ile incelenmesi, Ankara Üniversitesi Sbf Dergisi,74 (1),1-34, 2019.

[3] Teti, E., Dallocchio, M., Aniasi, A. The relationship between twitter and stock prices. Evidence from the US technology industry, Technological Forecasting and Social Change, 149 (119747), 1-9, 2019.

[4] Sprenger T. O., Tumasjan A., Sandner P. G., Welpe, I. M., Tweets and trades: The information content of stock microblogs, European Financial Management, 20 (5), 926-957, 2014.

[5] Mao Y, Wei W, Wang B, Liu B, Correlating S&P 500 stocks with Twitter data. In: Proceedings of the 1st ACM international workshop on hot topics on interdisciplinary social networks research, 69–72, 2012.

[6] Sprenger T.O., Sandner P.G., Tumasjan A., Welpe I.M., News or noise? Using Twitter to identify and understand company-specific news flow, J Bus Finance Account 41(7–8), 791–830, 2014.

[7] Valle-Cruz, D., Fernandez-Cortez, V., López-Chau, A., & Sandoval-Almazán, R., Does twitter affect stock market decisions? financial sentiment analysis during pandemics: A comparative study of the h1n1 and the covid-19 periods. Cognitive computation, 14(1), 372-387, 2022.

[8] Karabulut, Y.,Can Facebook predict stock market activity? AFA 2013 San Diego Meetings Paper, Available at SSRN: https://ssrn.com/abstract=2017099, 2013.

[9] Eliaçik A. B., Erdogan N., Mikro Bloglardaki Finans Topluluklari için Kullanici Agirliklandirilmis Duygu Analizi Yöntemi, Ulusal Yazılım Mühendisliği Sempozyumu, İzmir, 782-793, 2015.

[10] Yildirim M., Yüksel C. A., Sosyal Medya ile Hisse Senedi Fiyatinin Günlük Hareket Yönü Arasindaki Iliskinin Incelenmesi: Duygu Analizi Uygulamasi. Uluslararasi Iktisadi ve Idari Incelemeler Dergisi, 22, 33-44, 2017.

[11] Ates, E., Guran, A. (2021). Pearson correlation and Granger causality analysis of Twitter sentiments and the daily changes in Bist30 index returns. Journal of the Faculty of Engineering and Architecture of Gazi University, 36(3), 1688-1701.

[12] Nielsen F., A new anew: evaluation of a word list for sentiment analysis in microblogs, In The ESQ2011 Workshop on Making Sense of Microposts, Heraklion, Crete, 93-98, 2011.

[13] T. Wilson, J. Wiebe, P. Hofimann, Recognizing contextual polarity in phrase-level sentiment analysis, in: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing—HLT, 347–354, 2005.

[14] Baccianella, S., Esuli, A., Sebastiani, F., Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining, In LREC'10, Valetta, 2200-2204, 2010.

[15] Loughran, T. I. M., McDonald, B., When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. Journal of Finance, 66(1), 35-65, 2011.

[16] Koukaras, P., Nousi, C., Tjortjis, C., Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning, In Telecom MDPI, 3(2), 358-378, 2022.

[17] De Oliveira Carosia, A. E., Coelho, G. P., da Silva, A. E. A., Investment strategies applied to the Brazilian stock market: a methodology based on sentiment analysis with deep learning. Expert Systems with Applications, 184, 115470, 2021.

[18] Qiu, Y., Song, Z., & Chen, Z. (2022). Short-term stock trends prediction based on sentiment analysis and machine learning. Soft Computing, 26(5), 2209-2224, 2022.

[19] Yue, L., Chen, W., Li, X., Zuo, W., Yin, M., A survey of sentiment analysis in social media. Knowledge and Information Systems, 60(2), 617-663, 2019.

[20] Hamraoui, I., Boubaker, A., Impact of Twitter sentiment on stock price returns. Social Network Analysis and Mining, 12(1), 1-15, 2022.

[21] Bollen J., Mao H., Zeng X., Twitter mood predicts the stock market, Journal of computational science, 2 (1), 1-8, 2011.

[22] Mittal A., Goel A., Stock prediction using twitter sentiment analysis, Working Paper Standford University, CS 229, 1-5, 2012.

[23] Ranco G., Aleksovski D., Caldarelli G., Grcar M., Mozetic I., The effects of Twitter sentiment on stock price returns, PloS one, 10 (9), 1-21, 2015.

[24] Pagolu V. S., Reddy K. N., Panda G., Majhi B, Sentiment analysis of Twitter data for predicting stock market movements, In 2016 international conference on signal processing, communication, power and embedded system (SCOPES), Paralakhemundi-Hindistan, 1345-1350, 3-5 Ekim, 2016.

[25] Deng S., Huang Z. J., Sinha A. P., Zhao H., The Interaction between Microblog Sentiment and Stock Return: An Empirical Examination, MIS quarterly, 42 (3), 895-918, 2018.

[26] Zhao R., Quantifying the correlation and prediction of daily happiness sentiment and stock return: The Case of Singapore, Physica A: Statistical Mechanics and its Applications, 533, 1-9, 2019.

[27] https://libguides.library.kent.edu/spss/pearsoncorr Access date: June 1, 2019.

[28] Bland M. 4th ed. Oxford: Oxford University Press; 2015. An Introduction to Medical Statistics.

[29] Campbell MJ, Machin D, Walters SJ. 4th ed. Chichester: John Wiley & Sons, Ltd; 2007. Medical Statistics: A text book for the health sciences.

[18] Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. Annals of cardiac anaesthesia, 22(1), 67.

[30] Ghasemi A, Zahediasl S. Normality tests for statistical analysis: A guide for non-statisticians. Int J Endocrinol Metab. 2012; 10:486–9.

[31] Kim HY. Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod. 2013; 38:52–4.

[32] Armitage P, Berry G. 2nd ed. London: Blackwell Scientific Publications; 1987. Statistical Methods in Medical Research.

[33] Barton B, Peat J. 2nd ed. Sydney: Wiley Blackwell, BMJ Books; 2014. Medical Statistics: A Guide to SPSS, Data Analysis and Clinical Appraisal.

[34] IBM SPPS software. https://www.ibm.com/tr-tr/analytics/spss-statistics-software. Access date: June 1, 2019.