*Ioana Ispir*

*Tiberiu Stoica*

*Adriana Nicoleta Stroe*

*Coca Cola:*

How did Ronaldo’s comments affect the customers perception of the brand? Sentiment Analysis

Table of Contents

[Introduction 3](#_Toc90469612)

[Methods 4](#_Toc90469613)

[Results 5](#_Toc90469614)

[Discussion 7](#_Toc90469615)

[References 9](#_Toc90469616)

# Introduction

Social Media refers to internet-based applications that have as a main characteristic the facilitation of online communication, networking and collaboration (Russo *et al.*, 2008; Carr and Hayes, 2015). As such, social media platforms, such as Facebook, Instagram and Twitter (to name a few), allow creating, sharing and exchanging user generated content (Kaplan and Haenlein, 2010; Carr and Hayes, 2015).

As social media platforms have grown rapidly in the past years and with millions of people using these platforms to gather information, share opinions and express their emotions, the platforms became a powerful tool for businesses

Another important aspect of social media platforms consists of gathering and analyzing the evolution of users’ sentiment regarding different topics and companies

Due to the power of social networks, some researchers believe that analyzing the data on the social media platforms, especially the ones on Twitter, can help us understand financial turnovers (Bollen, Mao and Pepe, 2011; Cazzoli *et al.*, 2016). However, even though a high number of quality datasets exists, there are limited research papers showing a clear correlation between stock prices and the sentiment of Twitter users (Ranco *et al.*, 2015).

As such, this paper is meant to analyze the Twitter users’ sentiment regarding a company before and after an important event that potentially influenced the opinion of consumers. In addition, the effects on the stock prices were also analyzed.

This year, during the 2020 UEFA’s European Championship held in the summer of 2021, prior to one of the pre-match press conferences, Cristiano Ronaldo promptly removed the Coca Cola bottles (one of the main sponsors of the tournament) from the shot and encouraged his viewers to drink water (Husband, 2021). This event highlighted the potential risks faced in the social media era by companies when associating with sports stars with an impressive social media following (Sweney, 2021).

# Methods

In order to be able to make the analysis, Anaconda-Navigator was installed, while the application Spyder (Python 3.9) was used to write the code that sits at the basis of this research. Afterwards, the data needed to be collected. Two data sources were identified, as follows:

* Twitter – directed online social network (Cazzoli *et al.*, 2016);
* Yahoo Finance – online financial platform from where financial data can be viewed and downloaded (Cazzoli *et al.*, 2016).

To be able to access the Tweets, an account for Twitter Developer was created. With the credentials given, the Twitter API was accessed, and the Tweets were fetched. Due to the nature of the analysis, the Tweets containing “#CocaCola” were collected in two separate variables:

* TweetsBefore: fetching the Tweets containing “#CocaCola” between 4th of June and 14th of June (the day of the pre-match press conference), around the time of the conference;
* TweetsAfter: fetching the Tweets containing “#CocaCola” between 14th of June (the day of the pre-match press conference), around the time of the conference and 24th of June.

For the next step, each individual Tweet stored in both variables mentioned above were processed in order to remove any parts that do not add any value to the sentiment analysis, as follows:

* Converting all characters to lower-case;
* Dropping the duplicate values;
* Removing any links;
* Removing user names;
* Removing stopwords;
* Removing punctuation.

Visualization tools were also created for each variable, including:

* Wordclouds – containing the most frequent words used in the fetched Tweets;
* Bar charts – showing the frequency of the words used in the fetched Tweets;

For the sentiment analysis, the sentiment needs to be defined first. As such, the following three categories were identified:

* Positive: values between 0 and 1;
* Neutral: values equal to 0;
* Negative: values between -1 and 0.

A function called “sentiment” was developed to analyze each Tweet and assign it into one of the three categories mentioned above. The attribute “polarity” was used to return a value between [-1;1], which indicated in which category the Tweet would be assigned. To make it more visual, two pie charts were created, showing the percentage of Tweets in each category.

In order to compare the sentiment of Twitter users before and after the pre-match press conference, an independent t-test was performed, which compared the means of the Twitter users sentiment. In addition, one sample T-test was performed on the average sentiment to test if the mean of a small group is similar to the mean of the sample.

Lastly, we looked at Coca-Cola’s financial data on Yahoo! Finance to see if the event had any effect on the company’s stock prices. Firstly, the data was accessed, fetched from the website and presented in a chart. Lastly, an independent t-test was performed to compare the means of the stock price from a set period before the event and after the event.

# Results

Firstly, we gathered the most frequently used words in the fetched tweets before the incident that contained the word “CocaCola”. There were no surprises there, as the results were mostly general words.

However, by using the same tag, the most frequently used words in tweets after the incident changed drastically. There were fewer mainstream opinions and more directly linked to the event. We can conclude from this that the event was a major one, that attracted a lot of attention and generated a lot of discussions and debates.

Bar chart

Description automatically generated with low confidenceA picture containing table

Description automatically generated

Word Cloud before the event: Word Cloud after the event:

Text

Description automatically generatedText

Description automatically generated

For the next part, a sentiment analysis was performed, which showed some interesting results. The sentiment was defined as follows: positive, negative and neutral. The results for the tweets before the event were 55.9% neutral, 17.4% negative and 26.7% positive. However, after the event, things changed in a surprising manner. The neutral tweets increased to 62.7%, the negative decreased to 9% and the positive ones increased to 28.3%. Even though it was widely regarded as an event that damaged the brand’s reputation, our results tell a different story. In reality, the people were not that invested (the overwhelming majority being neutral) or, even if they were, it did not seem as something with a negative connotation and, if anything, the event spiked up Coca-Cola’s reputation.

Chart, pie chart

Description automatically generatedChart, pie chart

Description automatically generated

Several t-tests were also performed. Firstly, an independent (unpaired) t-test in relation to the mean of the sentiment scores before and after the event was executed. As the p-value was not significant (p = 0.188 > 0.05), we can conclude that that the sentiment of the two variables is not significantly different. In addition, a one-sample t-test was also performed to determine if a number of tweets randomly chosen would score an average sentiment comparable with that of the whole group of reference. As the p-value was not significant (p = 0.433 > 0.05), we can conclude that the sentiment of the sample is not significantly different than 0.06, which is roughly equal to the average sentiment of the group of reference.

The findings related to the stock market were interesting as well. The brand’s share price was not completely stable (although the fluctuations in price were smaller than $1) and, as expected, on the day of the event the price dropped. What is interesting is that, the next day, it gained back some momentum (before dropping further). At the end of the period we analysed, however, the price grew back to more than its original stock price (from 01.06.2021), though lower than the maximum value in this period. An independent t-test in relation to the mean of the stock price before and after the event was also performed. As the p-value was not significant (p = 1.259 > 0.05), we can conclude that the share prices between the two periods mentioned before are not significantly different.

Chart, line chart

Description automatically generated

(Bedre, 2019; Dallanoce, 2021; GeeksforGeeks, 2021; Kite, 2021; Kuzminykh, 2021; NLTK Project, 2021; Pandas Development Team, 2021; Roesslein, 2021; The SciPy Community, 2021a, 2021b; Twitter, 2021; Twitter Developer Platform, 2021; Yahoo, 2021)

# Discussion

Whether Ronaldo knew what the implications would be for moving the Coca-Cola bottles or not, is difficult to assess. The event created a massive wave of controversy, that many correlated with the decrease of stock prices for Coca-Cola, which resulted in a loss of $4 billion of the brand’s value (IESE Business School, 2021; Yahoo, 2021). The press roared with opinions regarding the event, some praising the power that influencers have, others criticizing it. However, the premise that only one event could move the stock prices might be a bit shaky.

Stock prices move with respect to many indicators and that is a complex process. To assess that only one event, that many have described as a media stunt, is farfetched. While the number 4 billion might seem as extremely high, we must understand the context: Coca-Cola is one of the biggest corporations in the world; according to Yahoo! Finance, it has 4.3 billion shares, which, on Friday before the event, were priced at $56.16, thus conferring the company the value of $242 billion (Yahoo, 2021).

While it is true that, after the event, the value dropped by $4 billion, we must consider all the factors. In that morning, before the press conference, the company opened lower than the $56.16 value, and by 9:40 a.m., 3 minutes before Ronaldo moved the bottles, it had gotten to $55.26 (thus creating the $4 billion drop) (Yahoo, 2021).

There are several reasons for this occurrence, and Ronaldo might not be one of them. Firstly, at the time, the US stock market was not trading as much (for example, Ford Motor Company went down for more than $2 billion that day) (IESE Business School, 2021). In addition, Coca-Cola became ex-dividend (it has paid its dividends recently) so, naturally, the prices decreased marginally (Ibid). Furthermore, at the end of the trading day, the price went up by $0.30, adding $1.3 billion in value for the company (Yahoo, 2021).

In conclusion, the evidence that we gathered do not support the hypothesis that Ronaldo moving the Coca-Cola bottles during the press conference damaged the brand’s reputation or its stock prices. The general perception of the brand, if anything, improved and the stock prices were probably to be decreased no matter the superstar’s reaction. It was a natural occurrence, as the stocks are a volatile market, and the fact that, at the end of the period that we have analyzed, the price went higher than the original value, is an argument that supports this hypothesis.

# References

Banu, S. (2019) ‘Analysis of User Perception using Twitter Data’, *International Journal for Research in Applied Science and Engineering Technology*, 7, pp. 2429–2434. doi:10.22214/ijraset.2019.5402.

Bedre, R. (2019) *Perform three types of t-test in Python*, *Data science blog*. Available at: https://www.reneshbedre.com/blog/ttest.html (Accessed: 14 December 2021).

Bollen, J., Mao, H. and Pepe, A. (2011) ‘Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena’, *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), pp. 450–453.

Carr, C.T. and Hayes, R.A. (2015) ‘Social Media: Defining, Developing, and Divining’, *Atlantic Journal of Communication*, 23(1), pp. 46–65. doi:10.1080/15456870.2015.972282.

Cazzoli, L. *et al.* (2016) ‘A Large Scale Study to Understand the Relation between Twitter and Financial Market’, in *2016 Third European Network Intelligence Conference (ENIC)*. *2016 Third European Network Intelligence Conference (ENIC)*, pp. 98–105. doi:10.1109/ENIC.2016.022.

Dallanoce, F. (2021) *ANOVA, T-test and other statistical tests with Python*, *Medium*. Available at: https://towardsdatascience.com/anova-t-test-and-other-statistical-tests-with-python-e7a36a2fdc0c (Accessed: 14 December 2021).

GeeksforGeeks (2021) *Graph Plotting in Python | Set 1 - GeeksforGeeks*, *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/graph-plotting-in-python-set-1/ (Accessed: 14 December 2021).

Giachanou, A. and Crestani, F. (2016) ‘Like It or Not: A Survey of Twitter Sentiment Analysis Methods’, *ACM Computing Surveys*, 49(2), p. 28:1-28:41. doi:10.1145/2938640.

Go, A., Huang, L. and Bhayani, R. (2009) ‘

Husband, B. (2021) ‘Cristiano Ronaldo, Coca Cola and the truth behind the long running saga’, *Mirror*, 15 June. Available at: https://www.mirror.co.uk/sport/football/news/cristiano-ronaldo-coca-cola-truth-24324624 (Accessed: 13 December 2021).

IESE Business School (2021) ‘A Post-Truth World: Why Ronaldo Did Not Move Coca-Cola Share Price’, *Forbes*. Available at: https://www.forbes.com/sites/iese/2021/06/19/a-post-truth-world-why-ronaldo-did-not-move-coca-cola-share-price/ (Accessed: 15 December 2021).

Kaplan, A.M. and Haenlein, M. (2010) ‘Users of the world, unite! The challenges and opportunities of Social Media’, *Business Horizons*, 53(1), pp. 59–68. doi:10.1016/j.bushor.2009.09.003.

Kite (2021) *Code Faster with Line-of-Code Completions, Cloudless Processing*, *Kite*. Available at: https://www.kite.com (Accessed: 14 December 2021).

Kuzminykh, N. (2021) *Sentiment Analysis in Python With TextBlob*, *StackAbuse*. Available at: https://stackabuse.com/sentiment-analysis-in-python-with-textblob/ (Accessed: 14 December 2021).

NLTK Project (2021) *NLTK :: Example usage of NLTK modules*. Available at: https://www.nltk.org/howto.html (Accessed: 14 December 2021).

Pandas Development Team (2021) *pandas documentation — pandas 1.3.5 documentation*. Available at: https://pandas.pydata.org/docs/index.html (Accessed: 14 December 2021).

Ranco, G. *et al.* (2015) ‘The Effects of Twitter Sentiment on Stock Price Returns’, *PLOS ONE*, 10(9), p. e0138441. doi:10.1371/journal.pone.0138441.

Roesslein, J. (2021) *Twitter API v1.1 documentation*. Available at: https://docs.tweepy.org/en/latest/api.html#tweepy.API.search\_full\_archive (Accessed: 14 December 2021).

Russo, A. *et al.* (2008) ‘Participatory Communication with Social Media’, *Curator: The Museum Journal*, 51, pp. 21–31. doi:10.1111/j.2151-6952.2008.tb00292.x.

Sweney, M. (2021) ‘Coca-Cola’s Ronaldo fiasco highlights risk to brands in social media age’, *The Guardian*, 18 June. Available at: https://www.theguardian.com/media/2021/jun/18/coca-colas-ronaldo-fiasco-highlights-risk-to-brands-in-social-media-age (Accessed: 13 December 2021).

The SciPy Community (2021a) *scipy.stats.ttest\_1samp — SciPy v1.7.1 Manual*. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest\_1samp.html (Accessed: 14 December 2021).

The SciPy Community (2021b) *scipy.stats.ttest\_ind — SciPy v1.7.1 Manual*. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest\_ind.html (Accessed: 14 December 2021).

Twitter (2021) *Twitter. It’s what’s happening*, *Twitter*. Available at: https://twitter.com/ (Accessed: 15 December 2021).

Twitter Developer Platform (2021) *Premium search APIs*. Available at: https://developer.twitter.com/en/docs/twitter-api/premium/search-api/api-reference/premium-search (Accessed: 14 December 2021).

Yahoo (2021) *The Coca-Cola Company (KO) Stock Price, News, Quote & History - Yahoo Finance*, *Yahoo Finance*. Available at: https://finance.yahoo.com/quote/KO/ (Accessed: 15 December 2021).