JÖNKÖPING UNIVERSITY
Jönköping International
Business School

Twitter Sentiment Analysis on the Cryptocurrency Market

MASTER DEGREE PROJECT

THESIS WITHIN: Business Administration

NUMBER OF CREDITS: 15 ECTS

PROGRAMME OF STUDY: International Financial Analysis

AUTHORS: Frank van Engelen & Levente Kulcsár

JÖNKÖPING May 2023

Abstract

This thesis investigates the effect of sentiment on cryptocurrencies, and it aims to contribute to the literature on cryptocurrencies by extending the methods that are used in analyzing the impact of investor emotions on price development. Using extensive datasets and daily time-series-based variables, the study compares sentiment scores with financial metrics. By analyzing this relationship, the paper finds valuable information for companies, investors, and researchers interested in understanding the effect of sentiment on crypto. The relationship between sentiment and cryptocurrency prices is investigated by analyzing over 5 million tweets and price data of Bitcoin, Ethereum, and Cardano over a four-year period. Using the VADER sentiment tool, the study measures Twitter sentiment and tests the relationship between sentiment and returns using three different tools: Pearson Correlation, Cointegration, and Granger Causality. The study finds an overall positive sentiment for all the cryptocurrencies and establishes cointegration between all the cryptocurrencies and their sentiment, indicating a long-term relationship between the variables. However, the study only finds a Granger-Causing relationship of returns to sentiment for Bitcoin. The authors conclude that there is a definite long-term relation between sentiment and cryptocurrency price development, which is consistent with findings from previous literature. However, the direction of this relationship between the two variables remains inconclusive.

Acknowledgements

We would like to express our gratitude to our coordinator, S. Papadopoulos, for his support throughout this master's thesis. His guidance and expertise have been instrumental in ensuring the successful completion of this project. We would also like to thank him for providing us with the data access required for our analysis.

Furthermore, we would like to acknowledge the contributions of our thesis colleagues who provided us with feedback during the various stages of this research. Their constructive feedback and insightful comments helped us to refine our work and improve the overall quality of this paper.

Finally, we would like to express our gratitude to the faculty and staff of Jönköping University who have supported us throughout our academic journey. Especially Pär Sjölander, with helping us in different stages of the statistical analysis. We are deeply grateful for the knowledge, skills, and experiences we have gained during the process of this paper.

Table of Contents

1. Background

Crypto is one of the most interesting upcoming financial products, which has made its entrance into the financial world in the past years. With the current rapid technological developments, crypto and FinTech can take a new place in the world's financial system. Some call cryptocurrencies, and in particular Bitcoin (BTC), the new money. Others think that crypto is just another bubble in the financial world that can burst at any moment. Some even claim that the crypto bubble has already burst in the recent fall of 2022 (Mark & de Vynck, 2022). However, the technology behind blockchain and crypto, including smart contracts, seems very promising for the financial future. With the recent rise in popularity, there has been an increasing interest in understanding how investor sentiment impacts this market.

Among assets managers, private investors, and businesses there is a lot of discussion about whether crypto is an attractive form of investment or not. Whether it is functioning as a store of value, an investment that can return profits, a resource for transactions, or a substitute for other types of investments (Beg, 2022).

On the other hand, crypto is known for its high risks. One of those risks is the price volatility (Gianti, 2022). The cryptocurrency world is often compared to the Wild West of the financial markets. The prices are based on 'cowboy behavior' since the market is known to have a lot of beginners- and retail investors who are, on average, more sensitive to making emotional investment decisions. This increases price volatility. Furthermore, crypto is known to be vulnerable to various forms of fraud. The above-mentioned factors are all affecting investors-, asset managers- and business decisions.

However, the technology behind crypto is very promising and can be a disruptive innovation for the financial industry, even with the above-mentioned risks. Since a lot of the risks within cryptocurrencies are tied to emotions, the authors believe that researching the role of sentiment on crypto is relevant, because of the potential cryptocurrencies have in the modern-day financial system. Therefore, this paper is going to explore how sentiment and emotions are related to well-known financial metrics within the crypto world.

1.1 Problem Statement

As cryptocurrencies become more prominent in the fast-developing technological- and financial world, it is relevant to analyze their risks and rewards, especially in combination with sentiment and emotions. Sentiment analysis is a technique used to understand the emotion and opinions of people about a given subject from textual data (MonkeyLearn, 2018). An elaboration of the exact definition is in part 1.6. An analysis of the effect of sentiment on the price developments of crypto is relevant because it can, for example, impact investment & strategy decisions. A complementary reason to research the subject is that crypto price developments have reached their unpredictable limits (Parekh et al., 2021). Therefore, it has become increasingly harder to forecast or predict crypto prices, which makes decisions about, for example investments or strategies on crypto-based subjects harder to make. Using sentiment analysis to predict price development can therefore help make more informed decisions.

In existing literature authors perform a sentiment analysis on crypto and its' price development, including a few papers with a prediction of future movements of the prices. However, there was no found literature where the authors analyze multiple cryptocurrencies in combination with a transformation of the sentiment into a daily time-series process in the long-term. In this paper, a daily number regarding the sentiment is computed. This number can ultimately compare the sentiment to the price development of different cryptocurrencies.

In several research articles sentiment is found as a non-time series-based variable, but rather as a 'whole' number. In this paper, the sentiment score will be a time-series-based variable. We believe that analyzing the results this way should lead to more trustworthy conclusions since the price developments of any assets are time-based as well. Therefore, our problem statement is:

"The lack of involving multiple cryptocurrencies and a time-based sentiment variable may lead to inaccurate conclusions about price and/or volume predictions in previous literature."

1.2 Purpose

The main goal of this paper is to conduct sentiment analysis in the context of cryptocurrencies, to explore how the market is affected by the sentiment of investors. Sentiment analysis can help investors, companies, or researchers to understand how the market is reacting to news, events, and trends. By analyzing the sentiment of crypto with a time-based sentiment variable, we hope to add to the literature, so market participants can make more informed decisions regarding crypto investments.

Sentiment analysis is relevant in several ways. Firstly, by tracking the sentiment of social media posts and news articles, one can get a sense of the prevailing market opinion. The analysis can identify critical real-time issues. This information can also help to make informed decisions in less time about whether to buy, sell, or hold crypto investments (MonkeyLearn, 2018). Secondly, by using sentiment analysis tools, one can monitor the sentiment of individual cryptocurrencies or tokens. This information can help to identify trends and patterns that may signal the need for a change in strategy.

Finally, sentiment analysis can be used to forecast future movements to a degree. There is an association between sentiment and market movements. Analysis of the sentiment on the market or on an individual currency can provide valuable insights and one can make predictions about the future behavior or movement of the market (Yilmaz, 2023). For example, if sentiment is negative, it may be an indication that the market will experience a downturn in the near future. Conversely, if sentiment is positive, it may be an indication that the market will experience an upswing.

The further aim of the thesis is to contribute to the literature with the use of extensive datasets, the use of daily time-series-based variables, and the implementation of the analysis. Combining well-known frameworks with sentiment analysis will fulfill this goal. The research gap that is to be filled is to compare sentiment scores with price development. Doing so, while introducing a daily time-series-based variable for the sentiment in combination with the analysis of multiple cryptocurrencies. Researching the topics with this method, the paper hopes to give companies, investors, and researchers helpful insights into the effect of sentiment on cryptocurrency markets. Therefore, our research question will be:

“Does sentiment have an impact on the development of crypto prices & returns?”

In section 1.3 we will further explain why the research gap is relevant to study and what practical implications are particularly important.

1.3 Perspective

As for practical implications, our results could help companies, individuals & researchers who are making investment/resource allocation decisions. The thesis can also provide insights for researching the price movements of crypto and what affects those changes. First, it discusses what is relevant for investors and traders. By analyzing the sentiment of investors, traders can make informed decisions about whether to change their crypto portfolio.

For companies, there is an extra practical implication, next to the investment side, namely the strategic side. Sentiment analysis can be used as one of many information tools to forecast future movements to a degree. This will help companies that, for example, provide crypto transactions. They can make better predictions about the upcoming volume. The more accurate these predictions are, the better the company can divide its resources and manage costs.

For researchers, this thesis will provide information on how to analyze sentiment and compare that to the prices of assets, for example, stocks or bonds. Using this, they can form their own conclusions about other asset classes.

1.4 Definitions

1.4.1 Crypto & Cryptocurrencies

Cryptocurrencies are a form of digital or virtual currency that is secured through cryptography and operates independently of governments or financial institutions. These decentralized systems allow anyone to send and receive payments via peer-to-peer networks, all built on the blockchain. Considered by many as the future of money, Bitcoin is the first and most widely recognized cryptocurrency. Other cryptocurrencies have since been developed, each with their own distinct features. Cryptocurrencies are often used in decentralized finance (DeFi) and are popular as both a means of exchange and as a store of value (Kaspersky, 2022; Allen et al., 2022). Hassan et al. (2020) clarify that cryptocurrency, enabled by blockchain technology within the Internet of Things (IoT) paradigm, aims to reduce payment and transaction fees by eliminating intermediaries such as financial institutions. Cryptocurrency encompasses financial and non-monetary applications and includes coins and tokens. There is a difference between coins and tokens and their functionality. Coins are digital assets that have their own blockchain, whereas tokens rely on other blockchains (BOTS, 2022). The term altcoins refers to cryptocurrencies other than Bitcoin, like Ethereum (ETH) or Cardano (ADA).

For those unfamiliar with the concept of cryptocurrency, it is a type of digital money that utilizes encryption methods to ensure the secure creation and transfer of funds. Unlike traditional currencies, cryptocurrencies are decentralized and not regulated by central authorities like banks or governments. This means that transactions are not subject to the typical regulations of traditional financial systems. The most popular cryptocurrency is Bitcoin, but there are several others available such as Ethereum, Litecoin, and Ripple. Cryptocurrencies are created through a process known as mining. Mining consists of using powerful computers to solve complex mathematical equations. When solving these equations, new units of the currency are made and added to the blockchain (Frankenfield, 2023; Ashford, 2020; Reutzel, 2021).

The transactions are recorded on a decentralized public ledger known as the blockchain. This secure and tamper-proof database is an ideal method for recording financial transactions. In addition to being used to purchase goods and services, cryptocurrencies have other purposes, such as facilitating (smart) contracts. Investors must use cryptocurrency exchanges, which are online platforms that allow trading cryptocurrencies and other assets, such as Dollars or Euros (Begatta, 2021; Iansiti & Lakhani, 2018; Westra, 2018; Genç, 2022).

The value of cryptocurrencies is known for its rapid and drastic fluctuation, which makes them a speculative investment. Various factors, including global economic conditions, technological developments, and investor sentiment influence their value. (Yellin et al., 2018).

Cryptocurrencies come with risks such as hacking, theft, fraud, and price fluctuations. However, they also offer increased privacy and security because transactions are pseudonymous, meaning personal information is not required for transactions to occur. This is a valuable benefit for those who prefer privacy and want to retain control over their financial transactions (Ashford, 2020; Securities and Exchange Commission, 2017).

1.4.2 Sentiment Analysis

Sentiment analysis, which is also known as opinion mining, is a process that uses natural language processing (NLP) and machine learning (ML) techniques to extract subjective information from textual data. This process classifies text as positive, negative, or neutral sentiment. Sentiment analysis is a powerful tool that can help make more informed decisions by analyzing public opinion and recognizing trends and it can be used in predicting the stock market. By analyzing social media and news articles, one can obtain public opinion insights and adjust their strategies accordingly. In the finance industry, sentiment analysis is becoming increasingly important as more data becomes available, but it can be applied in various areas beyond finance, including customer service, product reviews, and political campaigns. (Yilmaz, 2023).

When breaking down the process of sentiment analysis, you mainly have three steps: text preprocessing and classification. The first step, text preprocessing, involves cleaning and normalizing the text data. The second step, feature extraction, involves identifying relevant features in the text, such as keywords and phrases. Finally, classification is the step where the text is categorized based on the sentiment it expresses (Monkeylearn, 2018). Generally, these three steps are used in sentiment analysis, but this paper only uses preprocessing and categorizations. The methodology of our analysis is further discussed in the Methodology part.

2. Frame of reference

2.1 Previous Literature & Hypotheses

A study conducted by Caferra (2020) analyzed 720 daily observations of 13 different cryptocurrencies between 1-1-2018 and 1-1-2020. The study focused on two key points: financial herding and the possible convergence of opinion. While the analysis of financial herding was not important for this study, the convergence of opinion & price development is relevant in understanding the relationship between sentiment and price in the cryptocurrency market. The study of Pano & Kashef (2020) uses Twitter for data scraping for sentiment analysis as well. The paper develops different text preprocessing strategies for correlating the sentiment scores of Twitter text with Bitcoin prices during the COVID-19 pandemic. It tries to identify which tweet preprocessing process results in the best correlating sentiment score with BTC closing prices. It uses VADER (a sentiment analysis tool) to analyze a tweet in 3 forms:

negative, positive, or neutral. After the preprocessing, the sentiment results are compared to the BTC prices. The paper tries to contribute to the already existing BTC price forecasting literature. The preprocessing of the data is in the following way: The data is gathered with the Twitter API and Python is used to perform the analysis. The tweets were manually chosen by a set of keywords. A total of 4.169.709 tweets were collected in this research. A strength of the paper compared to other papers is the big amount of data used in the study (with over 4 million tweets). During the preprocessing, the use of different Python functions cleaned the data and tokenized words in a way VADER could better capture different aspects of sentiment. To further compare the results of the sentiment analysis with the bitcoin prices, a Pearson correlation method between the scores of Twitter data and Bitcoin price development has been used. In the study of Kraaijeveld & De Smedt (2020), 24 million tweets were collected for the top nine cryptocurrencies as of 2018 over a period of two months between 04-06-2018, and 04-08-2018. Here as well, the sentiment analysis tool VADER was utilized with a mean score per interval, both on a daily and hourly basis. Toda & Yamamoto's Granger Causality test was employed to evaluate the relationship between sentiment and cryptocurrencies, tested on an hourly and daily interval.

There is one common major limitation of these papers:

- They consider short periods for analyzing the sentiment & cryptocurrency prices. Most papers use data that covers only a few months, weeks or even days. The reasoning behind this approach can be that they can analyze a higher volume of data during the given timeframe. The existing literature has some gaps. For example, the same type of research is not conducted in the long run.

Other limitations, for the papers of Caferra (2020) and Pano & Kashef (2020) are that:

- They were searching for the best data processing strategy that correlates the most with the Bitcoin price. This feels like 'matching' the strategy and the Bitcoin prices with each other. This is the wrong way of conducting research, because such approach can lead to biases, if you for example cherry pick data or methodologies.
- It only analyzes the Bitcoin prices and no other cryptocurrencies.

Shen et al. (2019) do not use sentiment analysis to research whether tweets have an influence on Bitcoin returns and trading volume, but research whether the volume of tweets influences these. The paper aimed to investigate whether Twitter data could be used to predict Bitcoin prices. The authors collected tweets related to Bitcoin from January 2014 to August 2018 and applied sentiment analysis to classify the tweets as positive, negative, or neutral. They then used a Granger Causality test to examine whether changes in Twitter sentiment could predict changes in Bitcoin prices, after controlling for other factors such as volatility, trading volume, and news sentiment.

The study by Caferra (2020) found that optimistic news had a significant impact on reducing returns dispersion, which highlighted a convergence of price expectations. This result suggests that sentiment has a strong effect on the price of cryptocurrencies, and eventually, price development will converge to market expectations. This is also covered in one of the most known theories in the financial market, namely the Efficient Market Hypothesis (EMH), where market prices will eventually move to price levels that reflect all available information (Fama, 1970).

According to Caferra (2020), the media's attention played a crucial role in this effect. It acted as an informative signal for the convergence of prices and price expectations. When positive news regarding cryptocurrencies was widely disseminated by the media, it led to an increase in prices. It concluded that it is essential for investors and traders to stay informed about the sentiment and media attention surrounding the cryptocurrency market to make informed investment decisions. Pano & Kashef (2020) found the same results but in the short-term. The sentiment has an impact on the price development of cryptocurrency prices. To be more precise, for the short-term, the VADER model can be used in actual BTC price forecasting. Selecting the optimum preprocessing strategy (for the data) would prompt machine learning prediction models to achieve better accuracy as compared to the actual prices. In the long-term, the results are inconclusive because of unpredictable correlation polarity. Moreover, the paper speculates that a substantial spike in correlation (between sentiment and Bitcoin prices may be stronger when analyzing it on a shorter time scale - minutes instead of days). Shen et al. (2019) found that even the volume of tweets already correlates with the price development of Bitcoin, so independent of any sentiment. The authors found that Twitter sentiment was a statistically significant predictor of Bitcoin returns in the short-term, up to two days. However, the predictive power of Twitter sentiment decreased as the prediction horizon increased, indicating that other factors may have a stronger influence on Bitcoin prices in the long-term.

Kraaijeveld & De Smedt (2020) include an analysis of the eight other biggest cryptocurrencies next to Bitcoin, from 04-06-2018. Different conclusions were found for each coin at different intervals. Here only the relevant conclusions for this paper are being discussed for Bitcoin, Ethereum, and Cardano. According to the study's authors, Twitter users tend to express more positivity or negativity towards Bitcoin when its price returns are high on a daily basis. However, during hourly intervals, there is no indication of predictive power in either direction. This suggests that Twitter is merely reacting to Bitcoin's price returns, rather than causing any Granger-Causal effect on them. On the other hand, the analysis of Ethereum on a daily basis revealed that Twitter sentiment has no predictive power on price returns. However, it was observed that price returns have an impact on the volume of messages on Twitter. The lagged analysis of the VAR model showed that price returns Granger-Cause message volume for selected lags during intraday and in some cases on daily intervals. This indicates that Twitter is simply reacting to Ethereum's price returns, rather than causing any Granger-Causal effect on them. Regarding Cardano, the study found that Twitter sentiment has a short-term predictive effect on daily trading volume (lag 1). No causal relations were found intraday (Kraaijeveld & De Smedt, 2020).

John & Stantic (2022) have tried to forecast crypto price fluctuations with Granger Causality analysis. The authors use a time series of public sentiment, expressed as the quantification of a large-scale collection of daily tweets (5 million + in total between 5 and 31 May 2021), to predict the price of the cryptocurrency Dogecoin. They focus specifically on Dogecoin and analyze the textual content of each tweet mentioning Dogecoin using a modified version of the lexicon-based sentiment polarity analysis method, known as VADER. The generated Twitter sentiment time series is then compared to a time series of the closing prices of Dogecoin for each day. John & Stantic (2022) found a unidirectional Granger-Causal relation for the Twitter sentiment on Dogecoin prices, in lags ranging from 2 to 4 days, but they did not find a bivariate relationship between the variables. The Pearson correlation coefficient confirms the correlation between Twitter and Dogecoin. Although the results of the study seem compelling, we have to be wary and critical of using this study. The time span of the data set is relatively small (just over 21 days), and the authors only analyze one token, namely Dogecoin. This token is known to be very responsive to sentiment (Dablander, 2021). Furthermore, no tests have been conducted for stationarity or autocorrelation in the variables, which can lead to spurious regressions.

To conclude, all the papers find a relation between crypto prices and sentiment. Crypto prices correlate with sentiment, or the volume of tweets correlates with crypto prices. There is a (strong) connection, however, the results are inconclusive regarding the direction of the relationship.

The hypotheses that are tested in this paper will give direction on how to answer the research question and will give the opportunity to compare the results found in the literature. To research whether the crypto market is affected by sentiment, the following hypotheses will be tested:

Hypothesis 1: “Sentiment impacts the development of crypto prices.”

Hypothesis 2: “Crypto prices impact sentiment.”

Hypothesis 3: “The relationship between crypto price development and sentiment is bidirectional.”

These hypotheses will give us insight on how crypto price development is affected by sentiment via different channels. By testing for both correlation and Granger Causality, this paper is confident it can test hypotheses 1, 2, and 3 and ultimately answer the research question.

Testing these hypotheses and making conclusions can be important for investors & businesses to make more informed decisions when investing in crypto.

There is almost non existing literature that contradicts the statement that sentiment does not correlate with or cause price development of crypto in any direction. The reviewed papers either find a connection between the number of tweets or the sentiment of tweets (or other textual data) and crypto price developments or find a connection between tweets and volume. Noticeable is that most of the found connections are based on short-term periods. The predictions for price development are only significant in the short-term (Caffera, 2020; Pano & Kashef, 2020; Karalevicius et al., 2018; Shen et al., 2019). Kraaijeveld & De Smedt (2020) add to these conclusions that sentiment has more of a ‘causal’ power than an ‘effect’ power on cryptocurrency price developments on a daily basis. However, when the hourly intervals are analyzed, the opposite is observed, indicating that sentiment is more reactive than ‘causal’ in this shorter time frame.

Combining the different methodologies and proposed gaps in the above-mentioned studies will form the basis for our analysis. How the research gaps proposed by the other papers are treated will be discussed in the next part.

3. Methodology

3.1 Data

The methodology used in the paper of Hassen et al. (2020) and Chen et al. (2019) made us think about how to analyze sentiments from Twitter. Hassan et al., (2020) conducted a study on 15,000 tweets related to cryptocurrency, and sentiment analysis was performed using machine learning to create an emotion score. The emotions tested included anger, anticipation, disgust, fear, joy, sadness, surprise, and trust, with negative and positive sentiment specifications. The results of the study did not draw any conclusions on prices or price development, the data processing part of the study is particularly relevant for us. The first step was to collect data by scraping Twitter, followed by pre-processing, and cleaning the tweets. They used a sentiment lexicon analysis to evaluate eight emotions and two sentiments. Similarly, in the present study, the pre-processing was aimed at cleaning the sampled tweets and converting them into text without hashtags, or tags (mentions). While using this technique the noise in the data is minimized even further, which ultimately drives more reliable results. A further description of the preprocessing process can be found on page 15.

Chen et al. (2019) use data from Stocktwits (a platform similar to Twitter, but it is dedicated to financial conversations) and Reddit for their analysis to conduct daily sentiment indicators. In their paper, Chen et al. (2019) create a crypto-specific lexicon for their sentiment analysis to examine the connection between individual investor confidence and crypto returns. They argue that the domain-specific language (that usually contains slang and emojis as well) differs from everyday financial jargon and it is less formal, so analyzing the specific context can help to enhance accuracy. An example of that jargon can be “to the moon”, which means that prices will go up (sometimes with extreme highs). However, they find that using this crypto-specific lexicon has no significant effect on the results of the sentiment analysis. Therefore, this paper does not need to implement a similar technique. The steps taken in these studies for text mining were similar to other papers referenced, which confirms that the process of scraping data that we had in mind from Twitter is the right way to proceed.

This paper will use the Twitter API for the same reason as Pano & Kashef (2020) did. To ensure compliance with Twitter's regulations, we as authors do not use any web scrapers that might violate the restrictions imposed for the Twitter API. Such restrictions were implemented to safeguard Twitter users. Any violation of the privacy of Twitter users is of course not desirable.

Rouhani and Abedin (2019) used Twitter data for their paper as well. They collected 5,376,281 tweets related to several coins. Thanks to duplicate values, the final dataset contained only 101,161 unique posts. This paper addresses the issue of filtering data already at the data downloading. Setting some data standards for downloading the data can be important for the prevention of data pollution, already in an early stage. After the data collection, Rouhani, and Abedin (2019) pre-processed the data to prepare it for sentiment analysis. This process included noise cleaning, tokenization (splitting text into words), and other text conversion steps to prepare the data for the sentiment models.

The above-mentioned papers provided the core concepts to scrape our data and eventually perform the sentiment analysis. This paper will run a similar data collection and data preprocessing procedure, but deviate from them in the following:

In this paper, 3 coins are considered, namely: Bitcoin, Ethereum, and Cardano. These cryptocurrencies cover around 60% of the whole market capitalization of cryptocurrencies at the time of the analysis, the 27th of February 2023. For the analysis, the Adjusted Closing prices of Bitcoin, Ethereum, and Cardano (against USD) are downloaded from the period of 01-01-2019 until 31-12-2022. The data is obtained from https://finance.yahoo.com/.

Originally, the Binance coin (BNB) and Ripple were both considered as well. The most important reason to include these coins is that they are in the top 5 biggest cryptocurrencies at that point in time. Logical reasoning is that the biggest cryptocurrencies are affected the most by a change in sentiment. However, during the data scraping the analysis found that these coins were not feasible to the data standards. The tweets generated from Ripple are often not crypto related since the word ‘ripple’ has several meanings in English and slang. For this reason, Ripple is excluded from the analysis.

When scraping data for BNB, it was found that the Binance coin is involved in discussions related to the Binance site. Users can have bad or good experiences with the platform, which will impact the sentiment analysis. These impacts are not related to prices or price development. Since the aim of the paper is not to research the sentiment on the Binance platform, BNB is excluded from this research paper.

Moreover, the analysis in this paper will use a bigger dataset with unique values. During the download process the data is already filtered to keep the dataset clean, applying the following criteria in the scraping:

- Only original tweets are downloaded - No retweets in the dataset. This criterion ensures that the same tweet is not duplicated by being downloaded as an original and a retweeted data point as well.
- The language is English. The packages for the sentiment analysis are optimized for English. It also ensures data homogeneity.
- Tweets that contain the word “giveaway” are excluded, limiting the noise by applying this criterion. Crypto-related contests or other activities with crypto prizes are not in focus, which is why this paper tries to filter out this type of content.
- During the process of downloading data for one coin, the other coins are excluded. For example, when data for Bitcoin is downloaded, Ethereum and Cardano are excluded. The goal of this process is to limit duplications and sentiment scores for individual coins can be calculated.

For example, the text query to download the tweets for Cardano was constructed as follows: *"cardano -is:retweet -giveaway -ethereum -bitcoin lang:en"*.

With the above-mentioned query, the data is downloaded from 01-01-2019 to 31-12-2022. To reach even distribution during the period 1000-1500 tweets every day for each coin were downloaded. After the download process, this paper has 5,170,534 unique data points that can be analyzed.

3.2 VADER Sentiment Analysis

In the research of Pano & Kashef (2020), when analyzing the optimal preprocessing strategy to develop the most accurate machine learning model, the authors mention the use of ‘Valence Aware Dictionary and sEntiment Reasoner (VADER) model. VADER is a sentiment analysis tool that uses a combination of a lexicon and rules to analyze text. It is designed to handle various forms of language commonly used in social media, such as slang, abbreviations, emoticons, and emojis. Unlike machine learning algorithms, VADER does not require any training and is known for its speed and high performance (Hutto & Gilbert, 2014).

For the sentiment analysis, we use the VADER tool mentioned by Pano & Kashef (2020), Kraaijeveld & De Smedt (2020), and John & Stantic (2022). This tool is able to transform qualitative text data into quantitative values without the need for training. The VADER tool will compute daily sentiment scores.

VADER is an open-source, lexicon, and rule-based social media sentiment analysis tool. It has several characteristics that make the tool especially useful for social media sentiment analysis (Hutto & Gilbert, 2014):

- It recognizes conventional use of punctuation to signal increased sentiment intensity (e.g., "Good!!!")
- use of contractions as negations (e.g., "wasn't very good")
- understanding many sentiment-laden slang words
- understanding many sentiment-laden emoticons such as :) and :D
- understanding sentiment-laden initialisms and acronyms (for example: 'lol')

VADER computes three scores. “Positive”, “Negative” and “Neutral” scores are proportions to indicate the percentage of text that fall under each category. The fourth output value from the VADER analysis is called “Compound”. This score is a mix of the 3 other outputs. It is a normalized value between -1 and +1 where -1 is the most extreme negative and +1 is the most extreme positive. This is a broadly used “normalized, weighted composite score” to measure sentiment. The mean of the compound score will be computed for every day, which will eventually give the daily sentiment scores.

Text can be translated to the negative - neutral - positive spectrum by using the following thresholds:

1. positive sentiment: compound score >= 0.05
2. neutral sentiment: (compound score > -0.05) and (compound score < 0.05)
3. negative sentiment: compound score <= -0.05

To give a better perspective, examples for the application of VADER is mentioned from Hutto & Gilbert (2014):

Text	positive	neutral	negative	compound
The book was good.	0.492	0.508	0	0.440
At least it isn't a horrible book.	0.363	0.637	0	0.431
Make sure you :) or :D today!	0.706	0.294	0	0.863
VADER is not smart, handsome, nor funny.	0	0.354	0.646	-0.742

Table 1: Examples for VADER sentiment scoring system.

While Table 2 contains examples from the analysis of this paper:

Text	positive	neutral	negative	compound
You should definitely own Bitcoin 🙏	0.310	0.690	0.000	0.402
New post: "Bitcoin no longer worth the mining cost, JPMorgan says"	0.000	0.700	0.300	-0.434
Reality is very bad for Ethereum	0.000	0.569	0.431	-0.585
Great project with huge potential heading to the moon 🚀🚀🚀🥳 #Ethereum #ETH #HarpyFinance #Airdrop	0.375	0.625	0.000	0.840
I'm buying $ADA, Cardano's ADA token is seeing some modest gains. ✌️	0.167	0.833	0.000	0.340

Table 2: Examples of tweets with their VADER sentiment scores.

3.3 Sentiment Score and Currency Price Comparison

First, this paper will use a simple Pearson Correlation analysis for the sentiment scores and the cryptocurrency prices, adopted from Pano & Kashef (2020) and John & Stantic (2022). A Pearson Correlation test will be performed for each individual cryptocurrency (for both the Adjusted Closing Prices and logarithmic returns) and its mean sentiment scores. The Pearson

Correlation coefficient measures the ratio of the shared variance of two variables in a linear relation (Cramer, 1997). The Pearson Correlation coefficient will be used as a supplemental analysis to give more reliability to the results found in the Granger Causality test and the Johansen cointegration test, which will be further explained below.

Shen et al. (2019) examined the link between investor attention and Bitcoin returns, trading volume, and realized volatility. They use the number of tweets that mention Bitcoin between 04-09-2014 and 31-08-2018. They studied the dynamics between the variables by applying a bivariate vector autoregressive (VAR) model.

They determine the lag length with the Schwarz-Bayesian information criterion method. After the VAR model, linear and nonlinear Granger Causality tests were conducted, where the VAR model is expressed as:

$$x_t = \beta_0 + \sum_{i=1}^{n} \beta_{1i} x_{t-1} + \sum_{i=1}^{m} \beta_{2i} y_{t-1} + \varepsilon_{1t}$$

$$y_t = \delta_0 + \sum_{i=1}^{n} \delta_{1i} y_{t-1} + \sum_{i=1}^{m} \delta_{2i} x_{t-1} + \varepsilon_{2t}$$

where:

x = Sentiment daily mean
y = Logarithmic daily returns
β = Constant of the sentiment variable VAR formula
δ = Constant of the returns variable VAR formula
ε = Residual

The log returns (y) for the Adjusted Closing Prices are specified as follows:

$$y_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log\,(P_{t-1})$$

In our analysis, we follow the process implemented by Shen et al. (2019), Kraaijeveld & De Smedt (2020) and John & Stantic (2022). Instead of the number of tweets, we will use our sentiment score in the regression and causality models. The Granger Causality test is a perfect statistical tool for us to examine whether one variable can cause another. We are using Granger Causality tests to explore whether sentiment can predict prices, whether sentiment and prices affect each other, and whether prices can predict sentiment.

Granger Causality will be tested bidirectionally, so the paper will examine if sentiment Granger caused the returns, and vice versa. For Granger-Causality, the Adjusted Closing Prices will be transferred into logarithmic returns. This way, every possible relation between the time-series variables will be tested. In the paper, the connection between Bitcoin and its sentiment, the relationship between Ethereum and its sentiment, and the relationship between Cardano and its sentiment will be tested separately by testing the hypothesis mentioned above, for each coin.

It's worth noticing that Granger Causality does not establish an actual causal relationship between two variables, but it can detect a statistically significant pattern in lagged values of the variables. So variable Y can have a predictive power for variable X or in any other given direction (Kraaijeveld and De Smedt, 2020). Furthermore, Engle and Granger (1987) mention that the Granger Causality test is invalid if the time series of the variables are non-stationary and co-integrated. This will cause a spurious regression. For this reason, this paper will follow the same trajectory as Kraaijeveld and De Smedt (2020) and continue with a Toda and Yamamoto (1995) ("T&Y") Granger Causality test, since the level data of the Adjusted Closing Prices of cryptocurrency is not stationary (p-values respectively for Bitcoin-, Ethereum- and Cardano Adjusted Closing prices; 0.535, 0.587, 0.481, as presented in Appendix Table 1). However, when the Adjusted Closing Prices are transferred into logarithmic returns the data becomes stationary. Therefore the Adjusted Closing prices are integrated order 1, I(1). The data of the compounded number for sentiment is stationary (p-values, respectively for Bitcoin-, Ethereum- and Cardano sentiment data; 0.023, 0.005 and 0.002, as presented in Appendix Table 1). Hence, the sentiment data is integrated order 0, I(0).

Toda and Yamamoto (1995) proposed a series of steps that are ideal for analyzing series with varying orders of integration. Using the T&Y Granger Causality test, this paper does not have to perform differencing, which can introduce biases. It can be used with any kind of variable, irrespective of whether it is in state I(0), I(1), I(2), co-integrated or not. To examine the maximum order of integration and to determine whether there is stationarity in the variables, an Augmented Dickey-Fuller test is conducted (ADF-test), as mentioned above with results. To test for autocorrelation in the residuals of the VAR models specified above, the Portmanteau autocorrelation test will be used (Hatemi-J, 2004). The VAR models in the T&Y Granger Causality test are specified above as well.

The appropriate maximum number of lags is chosen with the Akaike Information Criterion (AIC), which will be further discussed in section '4. Results. AIC is a statistical measure that can be used to determine the optimal number of lags in a time series regression model. The AIC balances the goodness of fit with the complexity of the model, penalizing models with a higher number of lags. Therefore, using the AIC information criterion can help to select a parsimonious model that fits the data well while avoiding overfitting (Bevans, 2020).

4. Results

4.1 Sentiment Analysis

For the sentiment analysis we found an overall mean of 0.168 for Bitcoin, 0.184 for Ethereum and 0.268 for Cardano. The overall mean of the sentiment was 0.207, which is positive. An interesting finding, since this reveals that the overall sentiment over the analyzed period was positive (Appendix Table 2).

To get a general idea about the connection between our analyzed values, a simple Pearson correlation test was conducted. The results are presented in Table 3. While Bitcoin and Cardano show a moderate correlation between sentiment and Adjusted Closing Prices with 0.554 and 0.446 coefficients respectively, the 0.255 value for Ethereum shows only a weak connection. Interestingly the results are not transferred to the logarithmic returns. All the 3 coins show no meaningful correlation with 0.066 for Bitcoin and 0.008 for Ethereum, while Cardano even turned negative in this aspect with a coefficient of - 0.012. An explanation for the difference can be that the correlation between sentiment and Adjusted Closing Prices measures a linear relationship. It indicates the degree to which changes in sentiment are associated with changes in the price levels of the assets. For the correlation between sentiment and returns, returns represent the change in value of an asset over time and are often expressed as a percentage of the initial investment. The percentage change for returns is different from the absolute value change for the Adjusted Closing Prices.

Bitcoin:

	Sentiment	Log Return	Price
Sentiment	1.000	0.067	0.554
Log Return	0.067	1.000	0.003
Price	0.554	0.003	1.000

Ethereum:

	Sentiment	Log Return	Price
Sentiment	1.000	0.008	0.255
Log Return	0.008	1.000	0.006
Price	0.255	0.006	1.000

Cardano:

	Sentiment	Log Return	Price
Sentiment	1.000	-0.012	0.446
Log Return	-0.012	1.000	0.012
Price	0.446	0.012	1.000

Table 3: Correlation matrixes for the considered coins and sentiment

When performing the ADF-tests for the sentiment data, the paper finds that the sentiment data is stationary for every cryptocurrency. However, as mentioned in the methodology, the Adjusted Closing Prices of cryptocurrencies are not. Because the levels of data are different and the data is not stationary for every variable, this paper will continue discussing the results of the T&Y Granger Causality procedure.

4.2 Determine Lags with AIC

Choosing the appropriate number of lags to include in the model can be challenging. In this paper, the Akaike Information Criterion (AIC) is used for the VAR models. The results indicate that the Bitcoin VAR model is best fitted with 10 lags, the Ethereum VAR model is best fitted with 18 lags, and the Cardano VAR model is best fitted with 11 lags. The fitted VAR models can be found in the Appendix (Appendix Table 3, 4, and 5).

4.3 Autocorrelation

To test for autocorrelation in the residuals of the bivariate VAR models for Bitcoin, Ethereum and Cardano, this paper used the Portmanteau autocorrelation test. In the case where autocorrelation is found in the residuals, the procedure to remove the autocorrelation is to add lags to the VAR model. The VAR models tested are the VAR models with the number of lags specified by the AIC (as mentioned above).

In the case when performing the test for Bitcoin, the test indicates that up until 11th lag, there is no significant autocorrelation present in the residuals (p-value: 0.265). This suggests that a model with 10 lags, as determined by AIC, is the optimal fit for the data. Including more lags will not improve the accuracy of the model.

In the case of Ethereum, the Portmanteau autocorrelation test indicates that autocorrelation is present in the residuals up to 19 lags. However, based on the AIC, the optimal number of lags to use in the model is 18. Therefore, the Ethereum VAR model needs to include 20 lags. Here we cannot reject the null hypothesis of no autocorrelation, indicating that autocorrelation is no longer present from this point onwards. Therefore, it is recommended to use 20 lags in the Granger Causality test for optimal results when analyzing Ethereum data.

The results of the autocorrelation test for the Cardano VAR model suggest that there is no significant autocorrelation present in the residuals up until the 12th lag. Furthermore, the AIC suggested that the optimal number of lags to use in the model is 11. These findings imply that incorporating more than 11 lags may not significantly improve the accuracy of the model. Therefore, it is recommended to use 11 lags when modeling Cardano data, as this is the optimal choice based on both the Portmanteau autocorrelation test and the AIC.

The Portmanteau test results can be found in Appendix Table 6.

4.4 Plotting Adjusted Closing Prices and Sentiment

When plotting the Adjusted Closing Prices and sentiment scores for individual coins, a pattern is evident in some plots. However, based solely on the graphical representations, no definitive conclusions can be drawn regarding the correlation or directionality of the relationships between the variables. Nonetheless, it appears that correlation is more plausible. Notably, during periods when Bitcoin prices experienced an increase, the overall sentiment (mean) was more positive as compared to the periods preceding and following the Bitcoin price hike.

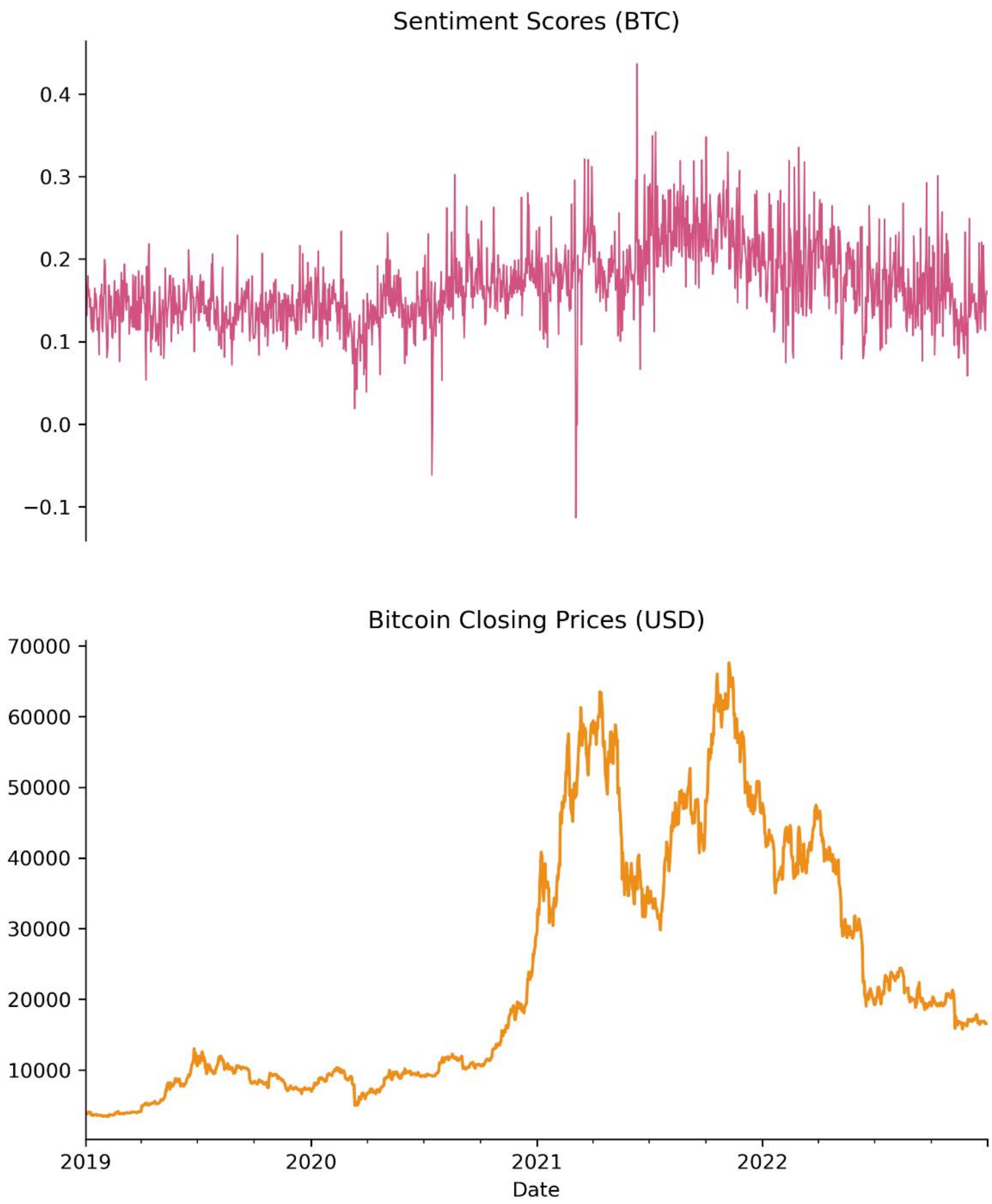

Figure 1: Bitcoin sentiment and Adjusted Closing Prices plotted together.

The same observation can be made for Cardano. At the beginning of the examined period, the sentiment followed an upward trend, while the price rose slightly and only increased significantly from the beginning of 2021. After a short period of high returns (high Adjusted Closing Prices), the price fell, but the sentiment score only followed the movement after the beginning of 2022, which indicates a correlation purely based on this plot observation.

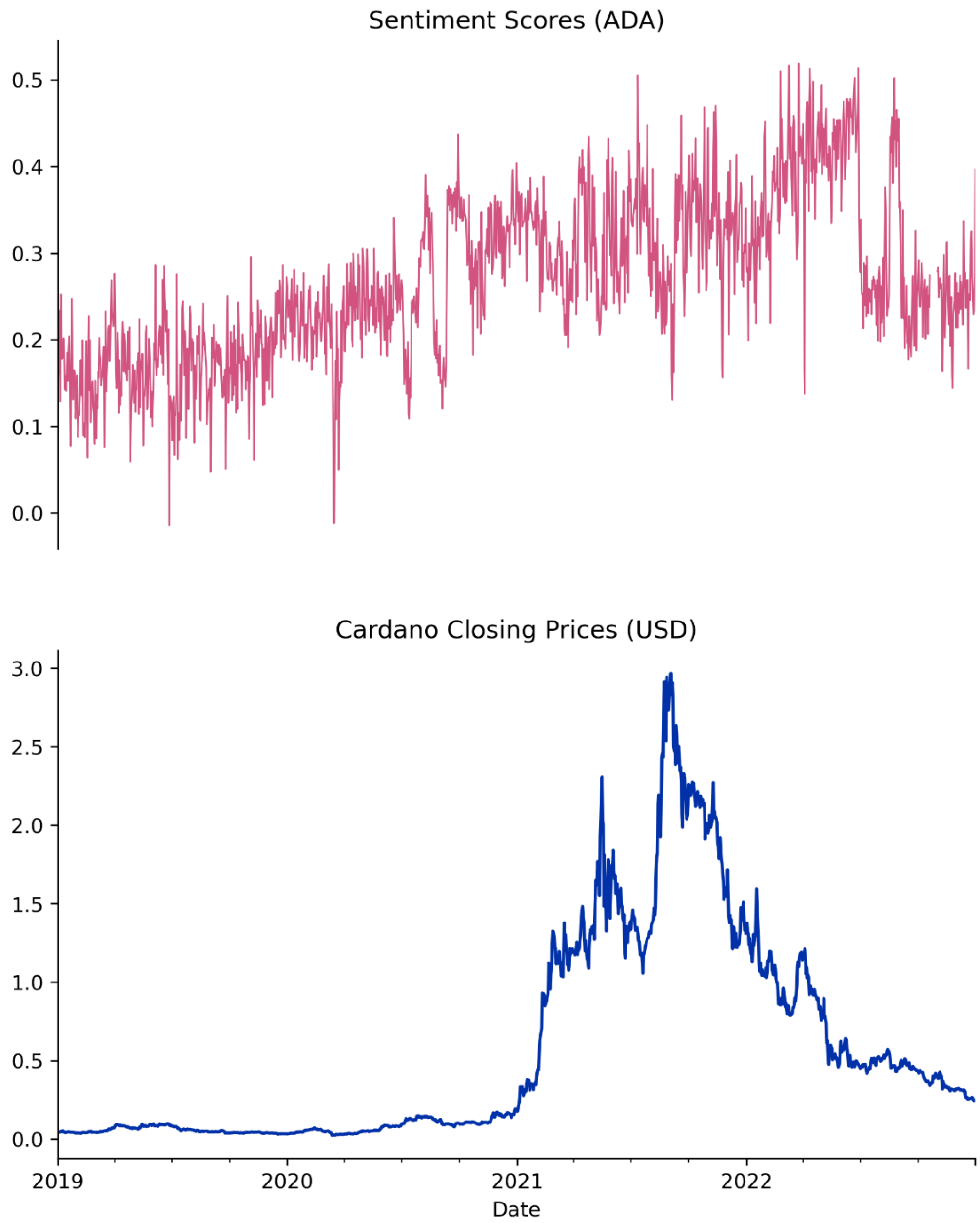

Figure 2: Cardano sentiment and Adjusted Closing Prices plotted together.

On the other hand, regarding Ethereum, the connection appears to be less evident when examining the graphical representations of the data. Nevertheless, a trend becomes apparent upon closer inspection of the price surge that occurred in May 2021, where the sentiment scores seem to align with the same pattern.

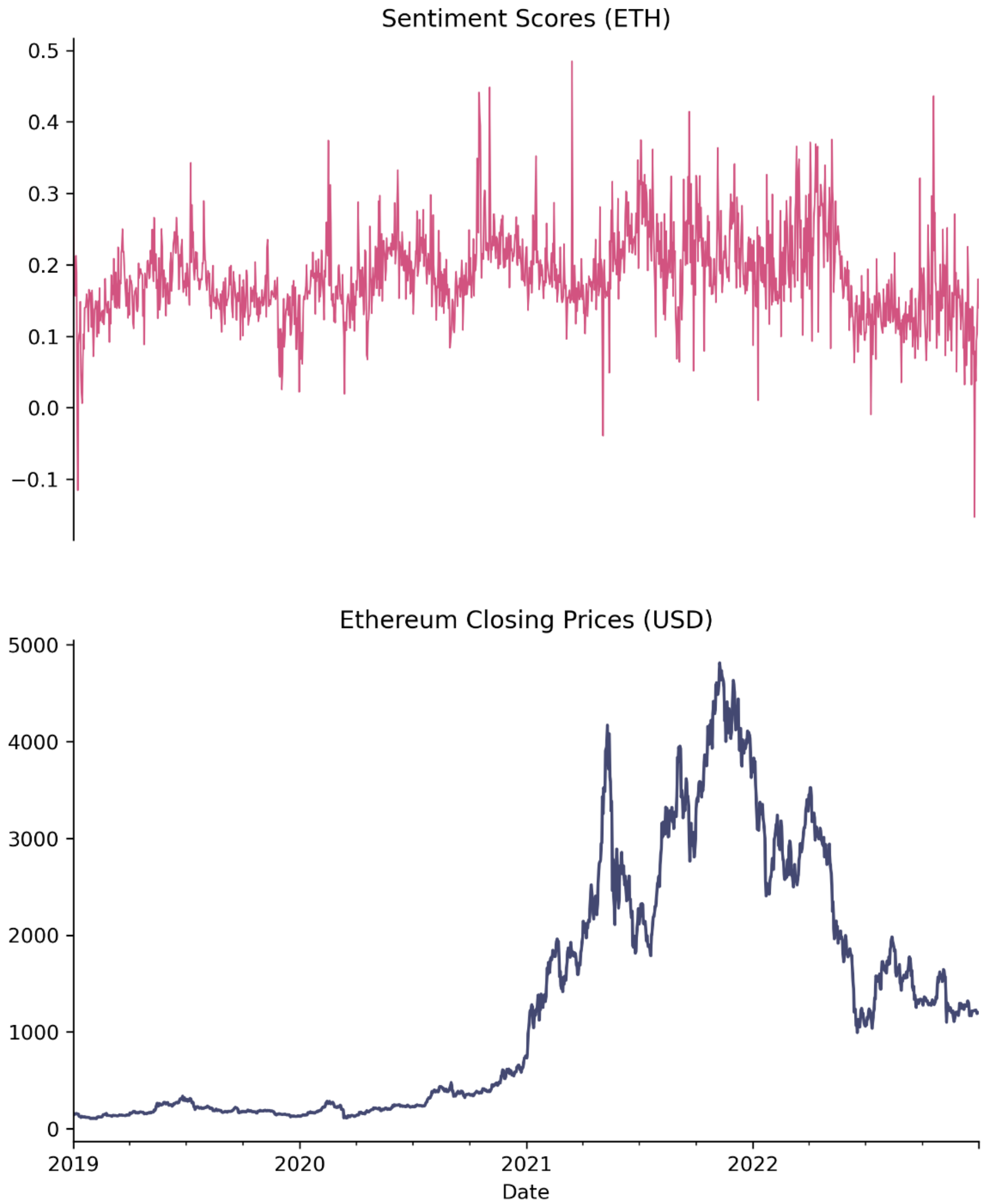

Figure 3: Ethereum sentiment and Adjusted Closing Prices plotted together.

4.5 Cointegration tests

In addition to the Toda and Yamamoto (T&Y) procedure, this study employs Johansen's Trace and Maximum Eigenvalue tests to validate the T&Y findings for series containing two-time series with the same order of integration. Both the log returns and the sentiment are I(0). Again, following the research technique from Kraaijeveld & De Smedt (2020). By comparing the results of the T&Y procedure with those of Johansen's tests, a reliable and robust approach is achieved, as co-integration between two-time series implies the existence of Granger Causality in the long-term. For the Johansen test, this paper compares the maximum eigenvalue statistics with the critical values at a 5% significance level. This paper tested for no cointegration equations (r=0) or at least one cointegration equation (r=1) since a bivariate VAR model is analyzed (Hjalmarsson & Österholm, 2007).

First, the test studies a cointegration analysis for Bitcoin and it reports that the variables under consideration exhibit cointegration. Notably, when the null hypothesis (H0) is tested for r = 0 at a 5% significance level, it is rejected, thereby confirming the presence of cointegration (test value: 120.07 > critical value: 14.26). We can also reject H0 for r = 1 (test value: 17.92 > critical value: 3.84). So, the generated results provide significant evidence of the existence of cointegration between Bitcoin returns and sentiment, as evidenced by the findings presented in Table 6 of the Appendix.

When testing cointegration for Ethereum we find the same results. For Ethereum, the variables sentiment and returns are cointegrated. When testing for r = 0 at a 5% significance level, H0 is rejected, which identifies cointegration (test value: 67.25 > critical value: 14.26). For r = 1 the null hypothesis is also rejected (test value: 13.15 > critical value: 3.84). Therefore, the outcome of the test indicates significant evidence that there is cointegration for Ethereum returns and sentiment (Table 7 in Appendix).

When testing cointegration for Cardano, we observe similar results. Specifically, that sentiment and returns variables for Cardano are cointegrated. The null hypothesis is rejected at a 5% significance level when testing for r = 0, which confirms the existence of cointegration (test value: 95.16 > critical value: 14.26). Moreover, we can as well reject H0 for r = 1 (test value: 12.15 > critical value: 3.84). Again, the results of the test provide significant evidence that there is a cointegration between Cardano prices and sentiment, as presented in Table 8 in the Appendix.

To conclude, there is significant evidence for a long-term relationship between the variables. This can eventually provide additional support when examining the Granger-Causal relationships in the following section.

4.6 Granger Causality Results

This paper aims to answer the research question if there exists a ‘causal’ relationship between sentiment and cryptocurrency prices. To achieve this, the study utilizes the Granger Causality test, which investigates the direction of the relationship between the variables and whether one variable affects the other over time. The full test results are disclosed in the appendix, where readers can find all the lagged p-values and their Granger-Causing effects.

Formulating the hypothesis, and the alternative hypothesis:

H0: No Granger Causality exists,

H1: Granger Causality exists.

The p-value determines whether the null hypothesis can be rejected at a 5% significance level, with a p-value greater than 0.05 indicating that no causality exists between the variables. The tables shown below present the last three lags of the Granger Causality tests because the last lag in the model contains the most important information based on the AIC.

The effect of sentiment to returns – Bitcoin

For Bitcoin, the p-values associated with each lag in the Granger Causality test of sentiment to returns exceed the threshold of 0.05 (Table 4). As a result, the null hypothesis (H0) that no Granger-Causing effect exists cannot be rejected, leading to the conclusion that there is no statistically significant evidence of a Granger-Causing relationship of sentiment to returns.

number of lags (no zero) 8		
ssr based F test	F=0.3213	p=0.9582
ssr based chi2 test	chi2=2.6010	p=0.9569
likelihood ratio test	chi2=2.5987	p=0.9570
parameter F test	F=0.3213	p=0.9582

number of lags (no zero) 9		
ssr based F test	F=0.6637	p=0.7424
ssr based chi2 test	chi2=6.0526	p=0.7346
likelihood ratio test	chi2=6.0400	p=0.7359
parameter F test	F=0.6637	p=0.7424

number of lags (no zero) 10		
ssr based F test	F=0.5862	p=0.8264
ssr based chi2 test	chi2=5.9482	p=0.8196
likelihood ratio test	chi2=5.9360	p=0.8206
parameter F test	F=0.5862	p=0.8264

Table 4: Granger-Causality test results, sentiment to returns – Bitcoin.

number of lags (no zero) 8		
ssr based F test	F=4.5464	p=0.0000
ssr based chi2 test	chi2=36.8019	p=0.0000
likelihood ratio test	chi2=36.3433	p=0.0000
parameter F test	F=4.5464	p=0.0000

number of lags (no zero) 9		
ssr based F test	F=4.0603	p=0.0000
ssr based chi2 test	chi2=37.0279	p=0.0000
likelihood ratio test	chi2=36.5633	p=0.0000
parameter F test	F=4.0603	p=0.0000

number of lags (no zero) 10		
ssr based F test	F=3.5768	p=0.0001
ssr based chi2 test	chi2=36.2934	p=0.0001
likelihood ratio test	chi2=35.8466	p=0.0001
parameter F test	F=3.5768	p=0.0001

Table 5: Granger-Causality test results, returns to sentiment – Bitcoin.

The effect of returns to sentiment - Bitcoin

The Granger Causality test conducted to examine the causal relationship between returns and sentiment revealed p-values below the significance level of 0.05 for each lag (Table 5). Thus, the null hypothesis (H0) of the absence of a Granger-Causing effect is rejected, and it can be concluded that there is a statistically significant Granger-Causing effect of returns on sentiment. Specifically, for Bitcoin, the returns significantly influencing sentiment, but there is no evidence to suggest that the relationship operates bidirectionally.

The effect of sentiment to returns - Ethereum

Regarding Ethereum, the p-values for each lag in the Granger Causality test from sentiment to returns are above the 0.05 threshold, as presented in Table 6. Consequently, we fail to reject the null hypothesis (H0) that there is no Granger-Causing effect, leading us to conclude that there is no statistically significant evidence of a Granger-Causing relationship of sentiment to returns for Ethereum.

number of lags (no zero) 18		
ssr based F test	F=0.6397	p=0.8704
ssr based chi2 test	chi2=11.8185	p=0.8565
likelihood ratio test	chi2=11.7703	p=0.8589
parameter F test	F=0.6397	p=0.8704

number of lags (no zero) 19		
ssr based F test	F=0.6268	p=0.8886
ssr based chi2 test	chi2=12.2406	p=0.8751
likelihood ratio test	chi2=12.1889	p=0.8774
parameter F test	F=0.6268	p=0.8886

number of lags (no zero) 20		
ssr based F test	F=0.5951	p=0.9185
ssr based chi2 test	chi2=12.2515	p=0.9071
likelihood ratio test	chi2=12.1997	p=0.9090
parameter F test	F=0.5951	p=0.9185

Table 6: Granger-Causality test results, sentiment to returns – Ethereum.

The effect of returns to sentiment - Ethereum

A similar result is found for the opposite direction: the test reports that the p-values for each lag in the Granger Causality test, conducted to assess the causal relationship between sentiment and returns, are higher than the significance level of 5%, as shown in Table 7. Therefore, we do not reject the null hypothesis (H0) that there is no Granger-Causing effect, which ultimately leads us to the conclusion that there is no statistically significant evidence to support the existence of a Granger-Causing relationship of returns to sentiment for Ethereum.

number of lags (no zero) 18		
ssr based F test	F=1.0350	p=0.4161
ssr based chi2 test	chi2=19.1199	p=0.3845
likelihood ratio test	chi2=18.9943	p=0.3922
parameter F test	F=1.0350	p=0.4161
number of lags (no zero) 19		
ssr based F test	F=1.0054	p=0.4512
ssr based chi2 test	chi2=19.6335	p=0.4169
likelihood ratio test	chi2=19.5010	p=0.4251
parameter F test	F=1.0054	p=0.4512
number of lags (no zero) 20		
ssr based F test	F=0.9526	p=0.5189
ssr based chi2 test	chi2=19.6110	p=0.4825
likelihood ratio test	chi2=19.4786	p=0.4909
parameter F test	F=0.9526	p=0.5189

Table 7: Granger-Causality test results, returns to sentiment – Ethereum.

The effect of sentiment to returns - Cardano

Concerning Cardano, Table 8 displays p-values for the first 5 lags in the Granger Causality test from sentiment to returns, all of which are above the 5% significance level. Moreover, all the included lags have a p-value > 5%. As a result, we cannot reject the null hypothesis that there is no Granger-Causing effect. This leads us to the conclusion that there is no statistically significant evidence of a Granger-Causing relationship of sentiment to returns for Cardano.

number of lags (no zero) 9		
ssr based F test	F=0.2614	p=0.9845
ssr based chi2 test	chi2=2.3839	p=0.9838
likelihood ratio test	chi2=2.3819	p=0.9839
parameter F test	F=0.2614	p=0.9845

number of lags (no zero) 10		
ssr based F test	F=0.3725	p=0.9587
ssr based chi2 test	chi2=3.7806	p=0.9567
likelihood ratio test	chi2=3.7757	p=0.9569
parameter F test	F=0.3725	p=0.9587

number of lags (no zero) 11		
ssr based F test	F=0.7355	p=0.7049
ssr based chi2 test	chi2=8.2225	p=0.6932
likelihood ratio test	chi2=8.1991	p=0.6954
parameter F test	F=0.7355	p=0.7049

Table 8: Granger-Causality test results, sentiment to returns – Cardano.

number of lags (no zero) 9		
ssr based F test	F=0.5775	p=0.8165
ssr based chi2 test	chi2=5.2671	p=0.8104
likelihood ratio test	chi2=5.2575	p=0.8113
parameter F test	F=0.5775	p=0.8165

number of lags (no zero) 10		
ssr based F test	F=0.5098	p=0.8842
ssr based chi2 test	chi2=5.1732	p=0.8793
likelihood ratio test	chi2=5.1639	p=0.8800
parameter F test	F=0.5098	p=0.8842

number of lags (no zero) 11		
ssr based F test	F=0.5526	p=0.8676
ssr based chi2 test	chi2=6.1774	p=0.8613
likelihood ratio test	chi2=6.1642	p=0.8622
parameter F test	F=0.5526	p=0.8676

Table 9: Granger-Causality test results, returns to sentiment – Cardano.

The effect of returns to sentiment - Cardano

Furthermore, again a similar result is found in the opposite direction for Cardano. The results of the test, presented in Table 9, indicate that the p-values for each lag exceed the standard 5% significance level. Accordingly, we fail to reject the null hypothesis that states the absence of any Granger-Causing effect. This results in a conclusion that there is no statistically significant evidence to support the existence of a Granger-Causing relationship from returns to sentiment for Ethereum.

5. Conclusion & Discussion

5.1 Conclusion

The cryptocurrency market is becoming more and more important in our financial system. The market is known for its big value gains and value losses in a short period of time and for the impact sentiment has on this market. As discussed earlier, the sentiment analysis performed on three major cryptocurrencies, Bitcoin, Ethereum, and Cardano, revealed an overall positive sentiment towards these assets. The overall mean of sentiment for all three cryptocurrencies was 0.218, suggesting a relatively optimistic outlook on these assets among Twitter users.

First, the most basic test to test the relationship between sentiment and cryptocurrencies assets has been performed, namely the Pearson correlation test. The results of the Pearson correlation analysis indicate a moderate correlation between sentiment and Adjusted Closing Prices for Bitcoin and Cardano, while the correlation was found to be weak for Ethereum. However, the correlation between sentiment and returns was found to be insignificant for all three cryptocurrencies. This is likely due to the fundamental differences in the computation of returns and Adjusted Closing Prices.

While the Pearson correlation test provided a quantitative analysis of the relationship between sentiment and Adjusted Closing Prices, visual inspection of the plots also revealed interesting patterns. Specifically, the plots for Bitcoin and Cardano showed a correlation between sentiment scores and prices, with sentiment being more positive during periods of increased prices. However, the plot for Ethereum did not provide a clear indication of this relationship. Nonetheless, a closer inspection of the price surge that occurred in May 2021 revealed a potential alignment between sentiment scores and price development, suggesting that market sentiment may play a role in driving price movements in the cryptocurrency market or the other way around. It is important to note that these visual observations do not provide any definitive conclusions, and further analysis was necessary to determine the strength and significance of the relationship between these variables.

To further investigate the relationship between sentiment and cryptocurrency prices, this paper proceeded with the Toda & Yamamoto Granger Causality methodology. To account for the non-stationarity of the Adjusted Closing Prices in the cryptocurrency market, logarithmic returns in the VAR model are used. This allowed us to model the stationary data and capture the dynamics of the market more accurately. Using the Portmanteau autocorrelation test and AIC, the optimal number of lags for each cryptocurrency in the VAR model is determined. Specifically, the VAR model for Bitcoin was found to be perfectly fitted with 10 lags, the model for Ethereum with 20 lags, and the model for Cardano with 11 lags. These findings suggest that the inclusion of the appropriate number of lags in the VAR model is crucial in accurately modeling the relationship between sentiment and prices in the cryptocurrency market.

When further examining the relationship between the variables, the Johansen's Trace and Maximum Eigenvalue tests are used to validate the Toda and Yamamoto (T&Y) procedure findings for series containing at least one cointegrated variable. Both of the time-series are integrated in the same order I(0). The paper finds evidence of cointegration between Bitcoin returns and sentiment, Ethereum returns and sentiment, and Cardano returns and sentiment. The null hypothesis for r=0 and r=1 is rejected at a 5% significance level for all three cryptocurrencies, confirming the presence of cointegration in all three relations. To conclude for cointegration, there is significant evidence of a long-term relationship between the variables, and they are cointegrated (Appendix Table 7). To investigate the ‘causal-’ and more short-term relationship, the relation was further analyzed with Granger Causality.

Our analysis of the Granger-Causality between sentiment and returns in the cryptocurrency market revealed varying results for the coins. Specifically, for Bitcoin, a statistically significant Granger-Causing effect exists for returns on sentiment. This suggests that past returns of Bitcoin can be used to predict future sentiment in the market. However, we did not find any statistically significant evidence of a Granger-Causing relationship of sentiment to returns in Bitcoin. In contrast, for Ethereum, we did not find any statistically significant evidence of a Granger-Causing relationship between sentiment and returns. This suggests that past returns of Ethereum are not useful in predicting future sentiment in the market, and vice versa, in the short-term. Similarly, for Cardano, our analysis did not reveal any statistically significant evidence of a Granger-Causing relationship of sentiment to returns or from returns to sentiment. This implies that the past returns of Cardano do not have a significant impact on future sentiment, or vice versa.

The overall results show that while cointegration is present between sentiment and returns for Bitcoin, there is only one statistically significant Granger-Causing effect relation of returns on sentiment. This finding contradicts the cointegration results for the other cryptocurrencies and highlights the importance of considering short-term and long-term relationships and causal relationships between variables. It suggests that although sentiment and returns may have a long-term relationship, changes in one variable may not cause changes in the other variable in the short-term. This answers our H1, H2, and H3. For H1, where we have to reject that sentiment has an impact on cryptocurrency development in the short-term. H2 suggests that only Bitcoin returns have an impact on sentiment, but not the other cryptocurrencies. And for H3 we can only conclude that the relationship between crypto prices and sentiment is a long-term bidirectional relation, but no Granger-Causal effect in the short-term. So, to finally answer the research question; "Does sentiment have an impact on the development of crypto prices & returns?", sentiment does not have an impact on crypto prices in the short-term perspective, but there is a long-term relationship between sentiment and crypto prices.

5.2 Discussion

Our findings both agree and disagree with the previous literature. While previous studies have found mixed results, this study presents contradictory findings. The findings of previous studies, such as Pano & Kashef (2020) and Shen et al. (2019), found that sentiment has an impact on price development. However, this paper only supports the short-term relationship between Bitcoin returns and sentiment and not for other cryptocurrencies in the short-term.

The study also contradicts the findings of Kraaijeveld & De Smedt (2020) and John & Stantic (2022), who found that Twitter sentiment tends to reflect cryptocurrency price returns on a daily basis and in the short-term. However, this study found an opposite relation for Bitcoin. While Kraaijeveld & De Smedt (2020) found that price returns have an impact on the sentiment of Twitter for Ethereum, this study found no relation.

Overall, as mentioned in the literature review, all the studies (including this paper) show that there is a (positive) relationship between sentiment and cryptocurrency prices, but the direction of this relationship seems to be inconclusive and seems to be more prevalent in the long-term.

The results of the analysis highlight the importance of sentiment in the cryptocurrency market, and its potential impact on asset prices, especially for Bitcoin in the short-term and Ethereum and Cardano in the long run. A better understanding of market sentiment can provide valuable insights for investors and traders, enabling them to make informed decisions and mitigate risks. This sentiment analysis should be considered as an additional information source when making decisions about crypto investments. Furthermore, the sentiment analysis approach employed in this study can serve as a framework for future research, facilitating a deeper understanding of the dynamics of the cryptocurrency market and its relationship with market sentiment. However, these findings suggest that while sentiment can be a useful tool in predicting price movements, it may not be the sole determining factor for returns in the cryptocurrency market. It remains important to use multiple information sources when investing in these cryptocurrencies or cryptocurrencies in general.

6. Limitations & Future Research

The limitations of this paper mainly originate from the data, the data source, and the statistical methods used in the analysis.

Twitter is an open platform, with millions of users, where everyone can express their opinion on different topics. In a controversial subject, such as cryptocurrencies, heavy arguments can occur on the platform. Sometimes users apply clickbait titles and exaggerated wording to express their opinion, leading to extreme sentiment values. Besides these arguments we can find tweets that mention coin names but are not related to the market. Cryptocurrencies are usually involved in some form of gambling, giveaways and unfortunately, sometimes they can be the subject of scams. These factors result in heavily noisy data, where the cleaning would require human supervision for every datapoint, making it impossible at the scale the study is conducted.

Due to the limitations of computation power and data access, the number of daily tweets considered for the analysis covers only a small portion of the available data. From the vast number of daily tweets for each coin, a sample of 1000-1500 tweets per day is used.

Also, these tweets are clustered around a specific time during the day, so the paper did not achieve random daily sampling when one considers the tweeting times. Usually, the analyzed Twitter data is sampled in the evening. This limitation originates from the Twitter API properties.

Furthermore, Pano & Kashef (2020) have a good suggestion for future research. They propose to include a larger dataset with time spans at each minute, instead of each day. This way the correlation spikes can be monitored more precisely, so investors have access to more information, not just daily based but even minute-based data. The authors of this paper agree with this suggestion, since from the analysis it seems like spreading out the study for several years can significantly affect the precisions of the data and therefore the results.

Twitter is a leading social media site among crypto investors, but of course not the only one. Other similar webpages and forums (Reddit, Facebook, Stocktwits, Instagram) are excluded, however, they can have excessive influence on investors' views. This paper also ignored other non-social media resources that investors may consider relevant during their decision-making, such as financial news sites and papers or investment blogs.

From a methodological point of view, the T&Y procedure requires the addition of lags to the VAR models. This ultimately leads to a loss of one extra degree of freedom and some efficiency. Despite this drawback, T&Y remains one of the best procedures available to test the hypotheses in this paper.

The determination of price-driving factors is a challenging task as there are many variables that may influence prices or sentiment. In this paper, a bivariate Granger-Causality test was conducted. However, several other factors can affect the variables tested in the model. For example, if a lurking variable Z influences both time series X and Y, the model may overfit due to the presence of spurious relationships. Kraaijenveld & de Smedt (2020) also noted this issue, which is a disadvantage of using this statistical method.

To minimize the effects of these drawbacks, this paper carefully selected the appropriate statistical techniques and considered the limitations and potential biases inherent in the data.

Further research based on this paper should consider the above-mentioned limitations and these also open up opportunities for additional analysis in the future. In order to check the performance of the models, it is possible to perform an out-of-sample prediction technique to evaluate their performance. Nevertheless, we made the decision to exclude such standard check models due to the lack of substantial evidence supporting a 'causal' relation between sentiment and returns in either direction. Additionally, incorporating the out-of-sample prediction methods would have exceeded the scope of this research paper. Consequently, this aspect would be better suited for future research.

7. Appendix

Bitcoin:

Sentiment		**Price**		**Log Return:**	
Test Statistics:	-3.145	Test Statistics:	-1.497	Test Statistics:	-18.070
p-value:	0.023	p-value:	0.535	p-value:	2.605
critical value 5%:	-2.864	critical value 5%:	-2.864	critical value 5%:	-2.864
Series is stationary		**Series is not stationary**		**Series is stationary**	

Ethereum:

Sentiment		**Price**		**Log Return:**	
Test Statistics:	-3.617	Test Statistics:	1.390	Test Statistics:	-11.424
p-value:	0.005	p-value:	0.587	p-value:	6.764
critical value 5%:	-2.864	critical value 5%:	-2.864	critical value 5%:	-2.864
Series is stationary		**Series is not stationary**		**Series is stationary**	

Cardano:

Sentiment		**Price**		**Log Return:**	
Test Statistics:	-3.946	Test Statistics:	-1.606	Test Statistics:	-18.186
p-value:	0.002	p-value:	0.481	p-value:	2.426
critical value 5%:	-2.864	critical value 5%:	-2.864	critical value 5%:	-2.864
Series is stationary		**Series is not stationary**		**Series is stationary**	

Appendix Table 1: Augmented-Dicky-Fuller test for each coin and sentiment data.

Bitcoin:

	Price	Log Return	Sentiment
mean	23528.076	0.001	0.168
std	17556.437	0.038	0.050
min	3399.472	-0.465	-0.113
25%	9143.582	-0.015	0.135
50%	17128.725	0.001	0.161
75%	38705.980	0.018	0.197
max	67566.828	0.172	0.436

Ethereum:

	Price	Log Return	Sentiment
mean	1313.076	0.001	0.184
std	1280.902	0.049	0.061
min	104.535	-0.551	-0.153
25%	204.056	-0.020	0.148
50%	737.803	0.001	0.178
75%	2160.768	0.027	0.217
max	4812.087	0.231	0.485

Cardano:

	Price	Log Return	Sentiment
mean	0.571	0.001	0.268
std	0.679	0.056	0.094
min	0.024	-0.504	-0.015
25%	0.058	-0.028	0.200
50%	0.181	0.001	0.261
75%	1.053	0.027	0.338
max	2.968	0.279	0.519

Appendix Table 2: Descriptive statistics for the considered coins.

Bitcoin:

	AIC	BIC	FPE	HQIC
0	13.2	13.21	540200	13.2
1	7.6	7.622	1999	7.608
2	7.583	7.620*	1965	7.597
3	7.579	7.631	1957	7.598
4	7.569	7.635	1938	7.594
5	7.567	7.648	1934	7.597
6	7.557	7.652	1915	7.593*
7	7.558	7.668	1917	7.599
8	7.552	7.676	1904	7.598
9	7.547	7.686	1894	7.599
10	7.543*	7.697	1888.*	7.601

Appendix Table 3: VAR Order Selection based on AIC (* highlights the minimums) for Bitcoin.

Ethereum:

	AIC	BIC	FPE	HQIC
0	8.629	8.637	5594	8.632
1	2.953	2.975	19.17	2.962
2	2.891	2.927	18	2.904
3	2.863	2.914*	17.51	2.882
4	2.849	2.915	17.27	2.874
5	2.843	2.924	17.17	2.873
6	2.837	2.932	17.06	2.872
7	2.83	2.94	16.95	2.871*
8	2.83	2.955	16.95	2.877
9	2.834	2.973	17.02	2.886
10	2.836	2.99	17.05	2.893
11	2.836	3.004	17.05	2.899
12	2.832	3.015	16.99	2.901
13	2.833	3.031	17	2.907
14	2.834	3.047	17.02	2.914
15	2.826	3.053	16.88	2.911
16	2.828	3.07	16.92	2.918
17	2.826	3.082	16.88	2.922
18	2.821*	3.091	16.79*	2.922

Appendix Table 4: VAR Order Selection based on AIC (* highlights the minimums) for Ethereum.

Cardano:

	AIC	BIC	FPE	HQIC
0	-5.713	-5.706	0.003302	-5.711
1	-11.87	-11.85	0.000007	-11.86
2	-11.93	-11.89	0.000006579	-11.92
3	-11.95	-11.9	0.000006435	-11.93
4	-11.99	-11.92	0.000006217	-11.96
5	-12.01	-11.92*	0.000006112	-11.98
6	-12.02	-11.92	0.00000603	-11.98*
7	-12.02	-11.91	0.000006036	-11.98
8	-12.02	-11.89	0.000006051	-11.97
9	-12.01	-11.87	0.000006062	-11.96
10	-12.03	-11.88	0.000005946	-11.97
11	-12.03*	-11.86	5.943e-06*	-11.97

Appendix Table 5: VAR Order Selection based on AIC (* highlights the minimums) for Cardano.

Bitcoin:

Test statistic	Critical value	p-value	df
5.22	9.488	0.265	4

Ethereum:

Test statistic	Critical value	p-value	df
13.62	15.51	0.092	8

Cardano:

Test statistic	Critical value	p-value	df
5.527	9.488	0.237	4

Appendix Table 6: Portmanteau Autocorrelation test results.

Bitcoin:

	Returns	Compound
Maximum eigenvalue statistic	120.072	17.922
Critical values (95%)	14.264	3.842

Ethereum:

	Returns	Compound
Maximum eigenvalue statistic	67.250	13.154
Critical values (95%)	14.264	3.842

Cardano:

	Returns	Compound
Maximum eigenvalue statistic	95.162	12.153
Critical values (95%)	14.264	3.842

Appendix Table 7: Johansen Trace and Maximum Eigenvalue test results.

8. References

[1] Ashford, K. (2020, November 20). *What Is Cryptocurrency?* Forbes Advisor. https://www.forbes.com/advisor/investing/cryptocurrency/what-is-cryptocurrency/

[2] Beg, M. (2022, March 3). *Is Bitcoin replacing gold, oil to become "Store of Value" of the digital age?* Business Today. https://www.businesstoday.in/crypto/story/is-bitcoin-replacing-gold-oil-to-become-store-of-value-of-the-digital-age-324657-2022-03-03

[3] Begatta, M. (2021, October 4). *What is Crypto Mining? How Cryptocurrency Mining Works*. Blockgeeks. https://blockgeeks.com/guides/what-is-crypto-mining/

[4] Bevans, R. (2020, March 26). *Akaike Information Criterion | When & How to Use It*. Scribbr. https://www.scribbr.com/statistics/akaike-information-criterion/

[5] BOTS. (2022, July). *Crypto Coin Vs. Token: Understanding the Difference | BOTS*. Www.bots.io. https://www.bots.io/news/crypto-coin-vs-token-difference#:~:text=Key%20Takeaways%3A

[6] Caferra, R. (2020). Good vibes only: The crypto-optimistic behavior. *Journal of Behavioral and Experimental Finance*, *28*(28), 100407. https://doi.org/10.1016/j.jbef.2020.100407

[7] Chen, C. Y.-H., Guo, L., & Renault, T. (2019). What Makes Cryptocurrencies Special? Investor Sentiment and Return Predictability. *IRTG 1792 Discussion Paper*, *2019-016*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3398423

[8] Cramer, D. (1997). Fundamental statistics for social research. Routledge.

[9] Dablander, F. (2021, February 7). *Causal effect of Elon Musk tweets on Dogecoin price*. Fabian Dablander. https://fabiandablander.com/r/Causal-Doge.html

[10] Engle, R. F., & C. W. J. Granger. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

[11] Fama, E. F. (1970). Efficient Capital Markets: a Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383–417. https://doi.org/10.2307/2325486

[12] Frankenfield, J. (2023, February 4). *Cryptocurrency Explained With Pros and Cons for Investment*. Investopedia. https://www.investopedia.com/terms/c/cryptocurrency.asp

[13] Genç, E. (2022, June 24). *Bitcoin ATMs: What You Need to Know Before Buying Bitcoin at an ATM*. Www.coindesk.com. https://www.coindesk.com/learn/what-you-need-to-know-before-buying-bitcoin-at-an-atm/

[14] Gianti, S. (2022, May 4). *Bitcoin vs Risk: Understanding Volatility*. Swissquote. https://medium.com/swissquote-education/bitcoin-vs-risk-understanding-volatility-472efe96e439

[15] Hassan, M. K., Hudaefi, F. A., & Caraka, R. E. (2021). Mining netizen's opinion on cryptocurrency: sentiment analysis of Twitter data. *Studies in Economics and Finance*, *39*(3). https://doi.org/10.1108/sef-06-2021-0237

[16] Hatemi-J, A. (2004). Multivariate tests for autocorrelation in the stable and unstable VAR models. *Economic Modelling*, *21*(4), 661–683. https://doi.org/10.1016/j.econmod.2003.09.005

[17] Hjalmarsson, E., & Österholm, P. (2007). Testing for Cointegration Using the Johansen Methodology when Variables are Near-Integrated; Erik Hjalmarsson and Pär Österholm; IMF Working Paper 07/141, June 1, 2007. https://www.imf.org/external/pubs/ft/wp/2007/wp07141.pdf

[18] Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

[19] Iansiti, M., & Lakhani, K. (2018, March 6). *The Truth About Blockchain*. Harvard Business Review. https://hbr.org/2017/01/the-truth-about-blockchain

[20] John, D. L., & Stantic, B. (2022). Forecasting Cryptocurrency Price Fluctuations with Granger Causality Analysis. *Recent Challenges in Intelligent Information and Database Systems*, *1716*, 201–213. https://doi.org/10.1007/978-981-19-8234-7_16

[21] Karalevicius, V., Degrande, N., & De Weerdt, J. (2018). Using sentiment analysis to predict interday Bitcoin price movements. *The Journal of Risk Finance*, *19*(1), 56–75. https://doi.org/10.1108/jrf-06-2017-0092

[22] Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, *65*(65), 101188. https://doi.org/10.1016/j.intfin.2020.101188

[23] Mark, J., & de Vynck, G. (2022, December 18). "Crypto winter" has come. And it's looking more like an ice age.. *Washington Post*. https://www.washingtonpost.com/business/2022/12/18/crypto-winter-ftx-collapse-bitcoin-prices/

[24] Pano, T., & Kashef, R. (2020). A Complete VADER-Based Sentiment Analysis of Bitcoin (BTC) Tweets during the Era of COVID-19. *Big Data and Cognitive Computing*, *4*(4), 33. https://doi.org/10.3390/bdcc4040033

[25] Reutzel, B. (2021, September 22). *What is cryptocurrency? Here's what you need to know about blockchain, coins and more*. CNBC. https://www.cnbc.com/select/what-is-cryptocurrency/

[26] Rouhani, S., & Abedin, E. (2019). Crypto-currencies narrated on tweets: a sentiment analysis approach. *International Journal of Ethics and Systems*, *vol. 36*(1), pages 58-72. https://doi.org/10.1108/ijoes-12-2018-0185

[27] R. Parekh et al., "DL-GuesS: Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction," in IEEE Access, vol. 10, pp. 35398-35409, 202

[28] *SEC.gov | Investor Bulletin: Initial Coin Offerings*. (2017, July 25). Sec.gov. https://www.sec.gov/oiea/investor-alerts-and-bulletins/ib_coinofferings

[29] *Sentiment Analysis: Nearly Everything You Need to Know | MonkeyLearn*. (2018, June 20). MonkeyLearn. https://monkeylearn.com/sentiment-analysis/

[30] Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict Bitcoin? *Economics Letters*, *174*(174), 118–122. https://doi.org/10.1016/j.econlet.2018.11.007

[31] Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, *66*(1-2), 225–250. https://doi.org/10.1016/0304-4076(94)01616-8

[32] Westra, D. (2018, March 20). *Smart contracts uitleg – Wat zijn smart contracts? | AllesOverCrypto*. Allesovercrypto.nl. https://allesovercrypto.nl/blog/smart-contracts-uitleg-zijn-smart-contracts

[33] Yellin, T., Aratari, D., & Pagliery, J. (2018). *What is bitcoin? - CNNMoney*. CNNMoney. https://money.cnn.com/infographic/technology/what-is-bitcoin/index.html

[34] Yilmaz, B. (2023, January 23). *Stock Market Sentiment Analysis: How it works & 7 data sources*. Research.aimultiple.com. https://research.aimultiple.com/sentiment-analysis-stock-market/