university of groningen

faculty of economics and business

Explaining the value of Non-Fungible Tokens through the analysis of social media sentiment

University of Groningen

Faculty of Economics and Business

Department of Marketing

MSc Marketing Analytics & Data Science

Alejandro Villanueva Lledo

a.j.villanueva.lledo@student.rug.nl

S4632613

First Supervisor: A.E. Tatar
Second Supervisor: F. Schaaf
Third examiner: T. Bijmolt

Abstract

Non-fungible tokens are unique digital assets that use blockchain technology to convey and guarantee ownership and distinctiveness, adding a layer of security to the digital art world. Being a fairly new but booming market, this thesis aims to study and understand the extent to which sentiment analysis on electronic word of mouth can affect the value of NFTs. With this goal, a sentiment analysis is performed in a series of tweets related to different NFT assets, linking OpenSea and Twitter parameters with tweet valence. An ordered logistic regression is carried out to assess the impact of not only tweet sentiment, but different Twitter and OpenSea metrics on the value of NFTs. The results of the regression reveal that tweet sentiment and Twitter metrics do not have a significant impact on NFT value, meanwhile, the previous number of sales of an asset have a positive impact on NFT value. These results can give a better understanding of the concept of NFTs as an investment opportunity and be a starting point for further research on the link between social media sentiment analysis and NFTs.

Table of Contents

1. Introduction

In 2008, Satoshi Nakamoto was credited with the creation of revolutionary blockchain technology. This technology will set the foundation for the birth of bitcoin, a now famous cryptocoin and the first mainstream application of blockchain technology. Massimo Di Pierro (2017) explains how blockchain technology created a distributed storage of time-stamped documents that cannot be altered without detection. In the blockchain a transaction includes the transaction plus a timestamp and every computer involved in the transaction keeps a copy of the history of the transaction. If a new record or transaction is inserted into a blockchain this is broadcasted to every interested party.
Following the growing popularity of cryptocurrencies and its use of blockchain technology came a wide range of applications, such as smart contracts and non-fungible tokens. Non-fungible tokens (NFTs) are defined as *"cryptographic assets on a blockchain with unique identification codes and metadata that distinguish them from each other"* (Investopedia, 2022). The non-fungibility characteristic of these tokens means that each token is unique and non-interchangeable with one another. For example, in fiduciary currency a 50€ bill has the same value and can be equally exchanged for two 20€ bills plus a 10€ bill or a 1€ coin can be equally exchange for a different 1€ coin and value will remain, on the other hand, NFTs don't have an agreed-upon value, but are instead valued by different unique properties, and one NFT cannot be easily exchanged for another. One of the unique traits comes from NFT's use of blockchain technology to keep a record of ownership, allowing confirmation of origin, custody and history of transfer of different digital, intangible or tangible assets.

In 2017 two Canadian artists created an experimental project named CryptoPunks, an NFT collection on the booming Ethereum blockchain. The project consisted of 10.000 unique characters, or tokens, inspired by London's punk scene. Not two tokens would be the same and it would be a limited release. Back then all CryptoPunks were free of purchase, with interested parties needing only to pay for the computational power required to create the tokens (gas prices). Since then NFTs trading volume has skyrocketed, going from $100 million in 2020 to $23 billion in 2021, a staggering increase of over 22000%[1] and it's a pace that does not seem to slow down. The NFT platform OpenSea has already reached a trading volume of $5 billion as of January 2022, breaking its previous record of $3.4 billion in August of 2021[2]. Beyond this, companies such as Starbucks, Pepsi or Pringles have

[1] https://www.financialexpress.com/market/nfts-generated-over-23-billion-in-trading-volume-in-2021-amid-craze-for-digital-assets/2390144/
[2] https://hypebeast.com/2022/2/opensea-new-record-nft-sales-january-2022

announced their plans to drop their own NFTs, and artists like Shawn Mendes, Mila Kunis and Justin Bieber have shown interest in the NFT market, buying and creating projects, highlighting the rise of NFTs in pop-culture. More recently, NFTs have been used to help fund the Ukrainian government on its war against Russia, with anonymous donors gifting valuable NFTs to Ukraine's digital wallet, as simultaneously, Ukraine creates and sells its own NFTs based on a timeline of its conflict with Russia. This shows how different industries, countries and policy makers are increasingly adopting and making use of digital art and blockchain technology.

Giving these assets a layer of ownership and authenticity throughout the use of blockchain technology was necessary with the increasingly digitalisation of art. The value of digital art is often subjective and generally all factors taken into consideration when evaluating physical art can be applied to digital art and NFTs, encompassing factors such as, the creator's personality or reputation, the narrative and the marketing associated with it (Irina Watkins, n.d.; Arnav Kapoor et al., 2022). For example, the value of one of Ukraine's NFTs may come from its authorship, the socio-political context in which it was created and/or its predicted historical relevance.

1.1 Problem Statement

NFTs can still be considered a new research area and with its increasing popularity and the financial pull it exhibits, assessing what adds to the value of NFTs can be an interesting and profitable area for research. An earlier study by Arnav Kapoor et al. (2022) takes a more marketing related approach into NFTs, looking at Twitter as a marketing tool that builds public perception and attracts buyers. Considering Twitter as a tool for NFT creators to promote their work, they make use of Twitter and OpenSea's quantitative data to predict NFT value, including number of likes and retweets of the promoting tweet and/or number of followers on the creator's Twitter account. While research has been done on the possible effect of Twitter and the price of NFTs on the same collection (Arnav Kapoor et al, 2022) or visual features on value and sales prediction (Yasin Uygun et al, 2021), to the best of my knowledge there are no studies that examine the possible effect of tweets sentiment analysis on NFT value. Research also shows that people trust seemingly disinterested opinions from people outside their immediate network (Cantallops et al., 2014) and Twitter puts this information out for everyone almost immediately. This leads to eWOM having a great influence on consumers when choosing products and/or services and, consequently, purchase intention.

This thesis aims to better understand the drivers behind the value of NFTs and will further investigate whether Twitter's quantitative data can significantly affect NFT value. Additionally, it will focus on including text sentiment analysis and broadening the scope of the study from promotional tweets to include all relevant tweets and their content. Consequently, the research question central to this paper is:

Research Question: *To what extent can Twitter sentiment analysis help predict NFT value?*

1.2 Contribution to literature

This study builds on previous literature by including non-promotional tweets and by investigating the effect of sentiment in electronic word-of-mouth (eWOM) in the value of NFTs. Previous studies have established the effect of sentiment analysis on other blockchain applications, such as cryptocurrencies (Naeem, Mbarki, & Shahzad, 2021; Abraham, Higdon, Nelson, & Ibarra, 2018) but these effects are yet to be studied for the NFT market.

Consequently, this paper will study if the value of NFTs can be predicted by studying eWOM spread through the microblogging site, Twitter. It is an exploratory research that will use both quantitative and qualitative data. Results could help artists/creators understand the effect of eWOM better and use it as a marketing tool, and investors/collectors can use it to assess potential NFT investments. At the very least this study will help better understand a field that is still scarcely researched. The data for this research will be mined from Twitter and OpenSea and will include randomly selected NFT projects sold in the first four months of 2022.

1.3 Report Structure

The remainder of this paper will proceed as follows: In section 2 existing literature is discussed and reviewed and hypotheses of relevance to the study are specified. The methodology is presented in section 3, specifying the data collection and data analysis methods. Section 4 will include the results of the different models and the results of the different hypotheses. The discussion and conclusion will follow in sections 5 and 6, respectively.

2. Literature Review

2.1 Non-fungible tokens

Due to the novelty of the subject, literature and research are limited. Much of the existing literature related to NFTs functions as an introduction to the concept of NFT and the link between NFTs, cryptocurrency and the blockchain. Similarly these articles discuss future opportunities and challenges that can be found in the NFT market. Often these opportunities relate to possible applications of NFT technology in the digital world such as the gaming industry or the increasingly hyped, metaverse (Qin Wang, 2021).

Iryna Watkins (n.d.) explains how through the use of blockchain technology the origin, history of transfer and custody of a digital creation can be easily accessed, giving digital art and creations the property of ownership, authenticity and uniqueness. Simultaneously, the author argues that the same factors taken into consideration when evaluating physical art can be extrapolated to the digital world. This idea is broadly supported by existing literature. Qin Wang et al (2021) agree on the benefits attached to creators being able prove the existence and ownership of digital assets as one of the most relevant characteristics of NFTs.

Recently, Raeesah Chohan et al (2021) dived into the future of NFT and how this developing technology creates an opportunity for marketers, giving a framework on how NFTs can be marketed based on the same principles of ownership and uniqueness explained before. The author argues that NFTs may significantly transform marketing functions. Additionally, this creates a comparison between decentralised applications (NFT) and more traditional products and the related marketing implications. When buying or acquiring a product or service, customers consider trust and the risk linked to a purchase (Kim, Ferrin, & Rao, 2008), echoing this Gefen and Pavlou (2012) argue that guarantees by the vendor can influence the perception of the risk taken in the purchase decision. The security inherent in blockchain technology can help reduce the perception of risk due to the upkeep of its transaction and ownership records. The author also uses the concept of scarcity to explain how the digital scarcity of NFTs is an asset to be exploited by marketers under the idea of consumers seeking to gain competitive advantage over others by identifying scarce resources that convey the feelings of distinctiveness and/or uniqueness. Moreover, the main notion of non-fungibility and not being limited by physical distribution are some of the aspects the author considers of relevance when marketing NFTs.

This thesis focuses on NFTs as a new form of art that can still be affected by some of the same principles that affect non-digital forms of art.

2.2 Twitter

The increasing use of social media marketing is obvious (Lan Jiang & Mehmet Erdem, 2017; Abdullah Alhidari et al., 2015) and a platform that has growingly and consistently been used as a marketing tool is Twitter. In contradiction with some authors that denied the plausibility of Twitter as a marketing tool (Martin Giles, 2010; Dorbian, 2010), Reema Aswani et al (2018) explain how platforms like twitter are often used by organisations for marketing purposes. Already in 2008 Cooke and Buckley (2008) predicted social media becoming a marketing tactic, and as soon as 2009, 53% of marketers claimed to plan increasing their social media investment further (Davidson, 2009). Authors such as Lan Jiang et al (2017) and Marius Bulearca (2010) support this view of twitter as more than just a microblogging platform, but a new *"path for businesses to attract and retain customers".*

$\boldsymbol{H_{1:}}$ Twitter quantitative data will have a positive effect on NFT value

Existing literature has studied the influence of Twitter on the value of NFTs. Arnav Kapoor et al (2022) analyse Twitter data on users promoting NFT assets through the social media platform. This data consisted of number of likes, retweets and time stamp of the promoting tweet (tweet with a link to an OpenSea project) and number of followers, following, bio and date of account creation of the account linked to the tweet. Similar to Seung-A, Annie Jin and Joe Phua (2014), who studied the effect of number of followers on brand-related outcomes and argue that the larger the number of followers on Twitter the greater the social influence, Arnav Kapoor et al. (2022) found a positive correlation between number of followers and NFT value, albeit a week one.

$\boldsymbol{H_{1a:}}$ Number of followers will have a positive effect on NFT value

In addition, retweets can be a powerful message reinforcing tool (Cha et al., 2010; Hung and Li, 2007), helping the original tweet reach an even broader audience facilitating the flow of information (Sun et al. 2006). Retweeting can be a medium to validate and engage with others, strengthening the original message and its valence. Already Kapoor et al (2022) found that the number of retweets and other metrics of engagement such as number of likes have an influence in their model output on NFT value.

H_{1b}: Engagement variables will have a positive effect on NFT value

Social media is now considered a significant part of the marketing mix (Jiang & Erdem, 2017; Withiam, 2010; Palmer & Koenig-Lewis, 2009) and it has opened an opportunity for consumers to offer consumption-related advice to a multitude of people and institutions by means of electronic word-of- mouth (eWOM) (Hennig-Thurau et al., 2015).
Considering Twitter as a communication tool that facilitates eWOM (Eunice Kim et al. 2014; Jansen et al., 2009; Zhao and Rosson, 2009), which has been acknowledged as one of the most effective forms of marketing (Jansen et al., 2009), can help shine some light into the relation between the value of NFTs and Twitter as a marketing tool and the development of eWOM.

Furthermore, Thorsten Hennig-Thurau et al. (2015) argue that Twitter eWOM can affect early product adoption and it can be especially relevant on products that depend on a hyped release, which can be the case for NFTs.

2.2 eWOM

Antoni Serra Cantallops et al (2014) categorise online reviews, recommendations and opinions within the scope of eWOM, similarly, Litvin et al (2008) define eWOM as "*all informal communications directed at consumers through Internet-based technology related to the usage or characteristics of particular goods and services, or their sellers*".
Traditional word-of-mouth (WOM) is an already important factor in consumer's decision-making process and product evaluation (Antoni Serra Cantallops et al., 2014; Sun-Jae Doh et al., 2009; Litvin et al, 2008; Engel et al. 1969; Gilly et al. 1998), a factor that only gains relevance when considering the increase in time spend on social media by consumers (Abdullah Alhidari et al., 2015) and the broader impact-reach and speed-of-interaction of eWOM (Antoni Serra Cantallops et al., 2014; Sun et al., 2006).

There is no previous literature on the effect of WOM or eWOM and its effect on art consumption, an existing study by Andrea Hausmann (2012) validates the effect of WOM and eWOM and museum visits, but nothing beyond this. However, the relationship between eWOM and purchase intention has already been established by existing research (King et al., 2014; Park & Kim, 2009). See-To and Ho (2014) agree that eWOM influences online consumer's purchase intention and it impacts other consumers' buying behaviour by creating brand value, moreover, eWOM is more important in persuading consumers to try new products (Alreck and Settle 1995), such as NFTs. This study will follow on the steps of previously established eWOM conceptual background investigating the effect of the direction of a tweet in NFT value. Christin Seifert et al (2019) carried sentiment analysis of eWOM in social networking sites (SNS) defining SNS users narrative on a brand as positive, negative or neutral, in accordance with literature that established direction of the tweet as whether the information within the tweet is positive, negative or neutral (Lee and Youn, 2009). The valence of eWOM has been proven to have a significant impact on product evaluation and purchase decision (Antoni Serra Cantallops et al, 2014). More specifically in Twitter 20% of tweets mention a specific brand and of these tweets 1 in 5 express negative or positive feelings towards the aforementioned brand (Jansen et al., 2009).

H_2**:** *Valence of a tweet impacts NFT value*

Additionally, Abdullah Alhidari et al. (2015) hypothesise that due to eWOM generally helping with better informed purchase decisions *"eWOM on SNS will be positively associated with purchase intentions on SNS."*

H_{2a}**:** *Tweets with a neutral sentiment will have a positive impact on NFT value*

This phenomenon was equally explained by Olson and Mitchell (2000) under the term "brand attitude". In line with studies by Wu and Wang (2011) and Park et al. (2008), they explain that brand attitude is an important predictor of consumer behaviour and positive brand attitude results in continuous preference from the customer toward that brand and has a positive effect of purchase intention.

H_{2b}**:** *Positive sentiment towards an NFT project will lead to a higher NFT value*

On the contrary, negative brand attitude has been associated with a negative effect on recommendations to friends (Lee and Youn, 2009), hotel choice (Vermeulen and Seegers, 2009) and lower movie sales (Rui et al., 2013).

H_{2c}**:** *Negative sentiment towards an NFT project will lead to a lower NFT value*

Additionally, Skowronski and Carlston (1989) find that negative information has a larger impact on customers than positive information. Negative information and WOM is more "attention grabbing" and has a larger influence on brand's perception and purchase intention than positive WOM (Arndt, 1967; Mizerski, 1982; Richins, 1984; Wright, 1974; Brown and Reingen, 1987; Weinberger et al., 1981 ; Homer and Yoon, 1992; Baumeister et al. 2001; Rozin and Royzman, 2001).

H_{2d}: *Negative eWOM will have a larger impact on NFT value than positive eWOM*

2.4 OpenSea

OpenSea is the largest NFT platform and being a relatively new idea driving large sums of money, it is fair to question whether this value holds. Traditional art has long been collected for aesthetic pleasure, nonetheless, pieces of artwork have algo gained value over time (Jianping Mei, 2005) turning art into investments with returns (Jianping Mei, 2005; Rachel Campbell, 2008) that outperform some fixed-income securities. Recent studies have also shown a similar trend in digital art, specifically NFTs. Mieszko Mazur (2021) analyses the return of investment of NFTs and the study shows that most NFTs change hands on the first day of listing and provide long-term returns, however due to the volatility of the crypto market and NFTs own market, the level of risk is relatively high. Simultaneously, this study shows an impressive return of investment of NFTs trading on the secondary market. The secondary market for NFTs, albeit a small one, is rapidly growing and secondary sale prices are usually higher than that of primary sales mostly due to a raise in the floor price. For most NFTs and artist this raise in price is just a little over the price in the primary market, but for some the value of resale has been over $774000[3].

H_3: *Number of sales has a positive effect on NFT value*

As argued in section 2.1, studies use the idea of scarcity in the NFT market as an advantage point for assets to be exploited by marketers and investors who rely on a feeling of uniqueness to gain competitive advantage, and as a consequence affect price and/or return of investment. Interestingly not all NFTs are completely unique. OpenSea allows for semi-fungible NFTs, letting artists create (mint) NFTs with multiple copies, hence, selling

[3]https://decrypt.co/63678/nfts-are-selling-millions-reselling

more than one or buying more than one of the same NFT is possible as long as it is a semi-fungible token. Evidently this goes against the idea of uniqueness possibly affecting the price of these semi-fungible tokens.

***H**$_4$: Semi-fungible tokens will have a negative effect on the value of NFT*

2.5 Conceptual Model

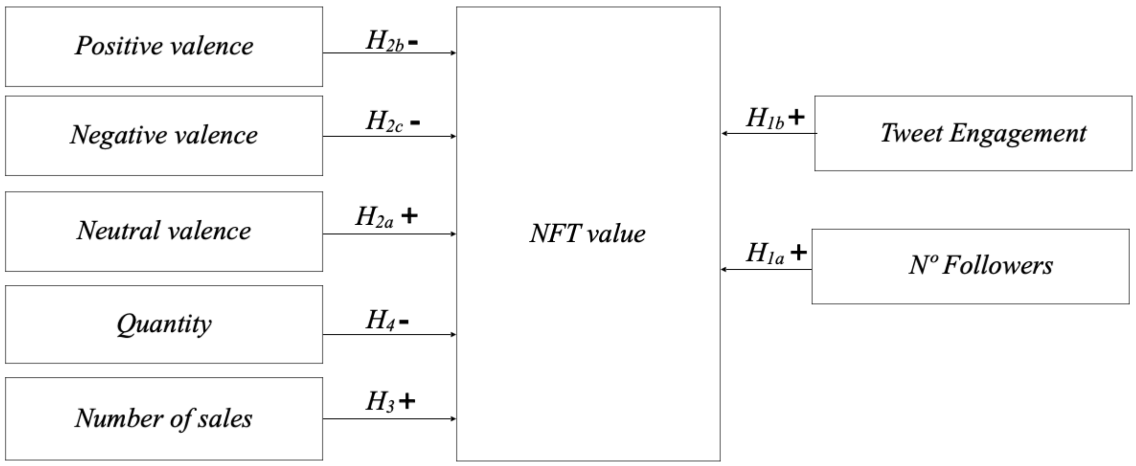

Figure 1 - Conceptual Model

3. Methodology

This research proposes to expand on the aforementioned study by Arnav Kapoor et al. (2022) by analysing the effect of eWOM on Twitter on the value of NFTs through the implementation of sentiment analysis, additionally, I aim to analyse twitter data not only as a promotional tool by those creators publicising their NFTs, but as a measure of reach and popularity in microblogging.

This paper aims to elaborate an explanatory model that will follow a quantitative approach.

An ordered logistic regression model will be used to classify assets (NFTs) in different price classes. In this case the classes are divided based on the average price of the NFT and represent NFT value. The maximum selling price was located at \$89.371, therefore classes were divided over powers of 10. Thus, class 1 includes NFTs selling below \$100, class 2 between \$100 and \$1000, class 3 between \$1000 and \$10000 and class 4 from \$10000 onwards.

Three different models will be used in the regression, a first full model for interpretation, an OpenSea model including only the OpenSea variables and a third model including only Twitter and valence variables for comparison on goodness of fit and explanatory power. The estimations and goodness of fit of these different models will be compared to determine whether sentiment analysis has any effect on the value of NFTs.

3.1 Data collection

The data consists of two datasets retrieved from two different sources that are merged into one big dataset.
As previously stated this research makes use of both qualitative and quantitative data. Data is scrapped from OpenSea, the largest NFT platform, with over 120 million visits in January 2022 only, and Twitter the top social media network directing traffic to OpenSea with over 64% of Opensea's social network traffic coming from Twitter alone[4].

3.1.1 NFTs Price Dataset

The individual assets and their respective information are scraped from the OpenSea API in Python. A timeframe of four months was set from January 2022 to April 2022 and only assets that were actually sold were considered for this study. OpenSea's API has strong limitations for data mining, hence only 2341 unique assets were retrieved. The dataset is aggregated at the asset level, thus price and number of sales are averaged per asset, while quantity (for semi-fungible tokens) was kept at its original value. For ease of further analysis the name of the assets were cleaned, eliminating spaces and punctuation symbols. The price in the dataset is given in the cryptocurrency ethereum (ETH), which is simultaneously the blockchain that holds NFTs sold in OpenSea. This price comes together with the exchange rate from ETH to USD (US dollar) at the time of the purchase, which is used to turn all ETH prices to USD prices for better analysis and interpretation.

3.1.2 Twitter Sentiment Dataset

The open software snscrape for Python was used to scrape Tweets made in the above-stated time period. All tweets mentioning or using a hashtag with the name of one of our assets in that time frame were collected which lead to a dataframe of 80961 different tweets. This

[4] https://www.similarweb.com/website/opensea.io/#overview

dataset was again aggregated at the asset level and included the text attached to the tweet, the username of the twitter account that wrote the tweet, the followers the account had, and the likes, retweets, replies and favourites (engagement variables) that each tweet amassed.
Once more, for ease of further analysis the text part of the tweets are cleaned from external links and emojis used.
Methodologically, this study relates to the existing literature on unstructured text analysis and text mining methods and will follow a supervised, pre-trained, neural network model based on RoBERTa (Liu et al. 2019) developed by Hartmann et al. (2021) that classifies tweet sentiment into positive, neutral and negative and analyses the intensity of the emotion giving the different emotion a score between 0 and 1. This deep learning model takes *"context dependencies of individual text elements into account"* (Hartmann et al., 2021) being able to differentiate similar words through context resulting in high accuracy levels. Once a score was given to each tweet all scores were averaged to keep the dataset aggregated at the asset level, similarly, all other variables were averaged to get e.g. the mean of the number of followers or likes per each unique asset.
The final dataset has a total of 11 columns and 998 rows.

3.2 Data analysis

Some basic descriptive statistics are shown in Table 1, here we see that there are big differences between the mean and maximum values in some of the variables, namely, number of sales (mean = 121.6, max = 14810), USD Price (mean = 582.86, max = 89371.5), retweets (mean = 6.6639, max = 220.25), replies (mean = 4.4607, max = 233.25), likes (mean = 14.979, max = 259.25), followers (mean = 4216, max = 226265) and favourites (mean = 6821, max = 170906). This is logically explained by the nature of NFTs, where only a few become very popular and in demand while the majority of them have a more humble path. As an example, the asset that sold for the maximum amount is a CryptoPunk NFT, mentioned in section 1, one of the first minted NFTs and one of the collections that has garnered the most attention. Normally, these assets showing a significant pull on the data would be considered outliers, however, the volatility that comes with the popularity of different assets is an innate characteristic of this market and consequently considered of interest for this research.

	Quant.	Sales	USD Price	Negative	Neutral	Posit.	Retw.	Replies	Likes	Follow.	Fav.
Min.	1.00	1.00	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.000	8	0
1st Qu.	1.00	1.00	68.43	0.0005209	0.6010	0.1172	0.9723	0.8489	5.267	1046	3215
Mean	1.003	121.6	582.86	0.0452284	0.7025	0.2473	6.6639	4.4607	14.979	4216	6821
3rd Qu.	1.00	3.00	378.59	0.0557546	0.8420	0.3376	5.0282	3.2712	17.023	4394	8174
Max.	2.00	14810	89371.50	0.9322788	0.9993	0.9995	220.25	233.25	259.25	226265	170906

Table 1 - Descriptive Statistics

The plot below (Figure 2) seems to indicate strong correlation between retweets, replies and likes and positive and neutral valence. These variables will be further studied to detect the existence of multicollinearity and the need to have some of them left out of the model. Collinearity is an issue that often appears in regression models, these models require predictors to be orthogonal, which allows to determine the significance of the different estimated coefficients independently of which of the other coefficients are significant (Trevor A. Craney, 2002). However, if the independent variables are not orthogonal, collinearity will exist between at least two of the independent variables and the p-values associated to the different parameters become meaningless measures of explanatory power. To identify the existence of multicollinearity a Variance Inflation Factor (VIF) is used. This tool reports how much of a regressor's variability is explained by other regressors in the model due to correlation among those regressors (Trevor A. Craney, 2002).

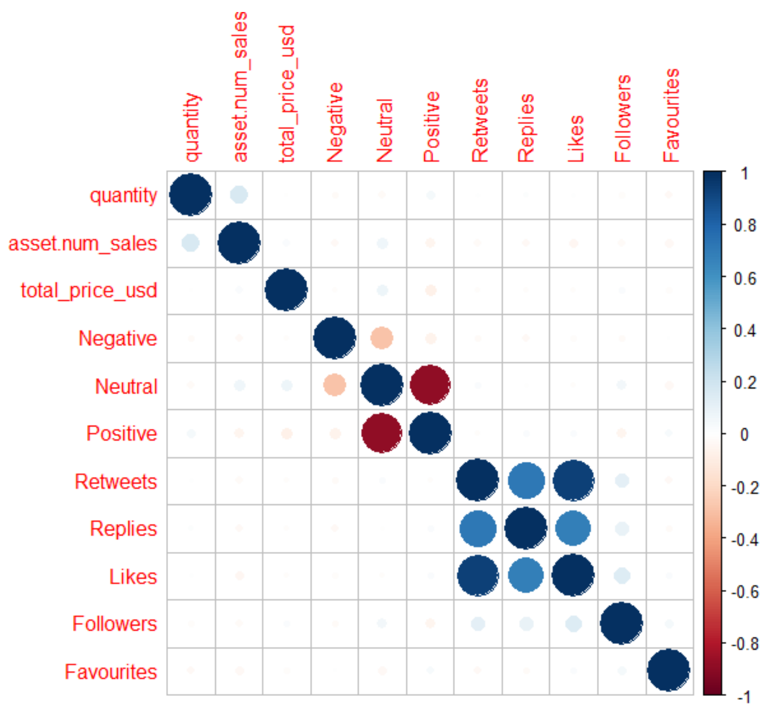

Figure 2 - Correlation Matrix

As a rule of thumb when using VIF, values greater than 5 indicate strong correlation or multicollinearity. Our model sees values of over 5 in the variables: Positive (8.98), Neutral (9.73), Retweets (8.78) and Likes (8.09).

3.2.1 Ordered Logistic Model

Ordered logit models, much like multinomial logit models, help classify observations. Ordered logit models focus on ordered dependent variables, meaning this variable is not continuous but takes discrete values. This differs from regular multinomial models by the fact that assets now face a ranked variable. As stated in section 3, assets have been assigned to three different categories according to their average price.

Ordered multinomial models, similar to the binary logit models, rely on an unobserved variable Y_i^*. This variable is related to an intercept and a vector of asset-specific variables x_i, that is (Leflang et al, 2017):

$$Y_i^* = \alpha + x'_i \beta + \varepsilon_i, \text{ with } E[\varepsilon_i] = 0 \qquad \text{(Eq.1)}$$

This latent variable has a direct relation with the categories that form the independent variable, Y_i. Mapping Y_i^* to have more than two categories can be written:

$$\begin{aligned} Y_i &= 1 \text{ if } \alpha_0 < Y_i^* \leq \alpha_1 \\ Y_i &= j \text{ if } \alpha_{j-1} < Y_i^* \leq \alpha_j \text{ for } j = 2,..., J - 1 \\ Y_i &= J \text{ if } \alpha_{j-1} < Y_i^* \leq \alpha_J \end{aligned} \qquad \text{(Eq. 2)}$$

Where α_0 to α_J are thresholds on the latent scale. The α are set to be $[-\infty,\infty]$, such that every value Y_i^* corresponds to one of the categories in Y_i. In summary, a latent scale is define and the thresholds α_J are used to divide this scale into J different regions. The probability of an asset belonging to Class J is equal to the probability that the latent variable is in between the thresholds α_{j-1} and α_j. The parameters in this model are estimated using Maximum Likelihood methods.

One of the assumptions of this model is that those thresholds are far away from each other, if the thresholds get close to each other it may become difficult to correctly classify categories (classes) that might be in the middle of two colliding thresholds.

Another important assumption in this model is that all IVs have the same coefficients in all equations, this is called parallel lines assumption, for assuming that regression lines are parallel. To check whether this equality holds a test is proposed by Brant (1990).

If the probabilities on the Brant test are above 0.05, then we can assume that the parallel lines assumption holds.

- **Validation and Fit**

In ordered models the R^2 cannot be directly computed, therefore, unlike GLM, R^2 cannot be used to assess the fit of a choice model, instead, pseudo R^2 statistics that have a similar interpretation have been developed for logistic models. Pseudo R2 are based on comparing the log-likelihood (LL) of a null model (model with only an intercept) with the log-likelihood of a model with K explanatory variables and take values between 0 and 1. If the pseudo R-squares are 0 then the LL of the model with K parameter is equal to the LL of the null model (Leeflang et al, 2015), concluding that the covariates have no explanatory power.
For comparison between models that are estimated with maximum likelihood it is common to use information criteria such as AIC and/or BIC, these criteria are full sample criteria (not data splitting required) and seek to locate the best approximation to reality. These information criteria also measure parsimony in parameterization of a model, meaning they punish model complexion. Information criteria is only interpreted as a way to compare two models, considering the model with the smallest or most negative value the better fit model. Another test that is commonly used is the LR test, used to investigate two models that are nested. This can help determine the goodness of fit and significance of each explanatory variable or set of variables in the model (using nested models). In the LR test the H_0 = The full model and the nested model fit the data equally well, hence, if the p-value is below 0.05, H_0 would be rejected.

4. Results

4.2 Ordered Logistic Regression

The classes in our DV are ordered in terms of average price with the lower class having a lower average price and the highest class including observations with the highest average price.

4.2.1 Check of Assumptions

Section 3.2.1 highlights some of the assumptions necessary for this regression model. The checks and tests discussed in this section were performed in each of the estimated models (OpenSea, Twitter and Full). Firstly, in equation 2, the importance of the unobserved thresholds not colliding is explained, so that all classes in the DV should be significantly different from each other. Therefore, we check that α_0 to α_J do not overlap. Through the use of the MASS package in R, the intercepts and standard errors for the different classes are estimated and used to check that the maximum possible value for one class does not overlap with the minimum possible value of the next class. This check proves that an ordered model can be applied.

Intercepts	**Value**	**Std. Error**	**α**
1|2	-1.2425	0.1077	[-1.4579, -1.0271]
2|3	1.6639	0.1500	[1.3639, 1.9639]
3|4	4.4988	0.4011	[3.6966, 5.301]

Table 2 - Thresholds of the latent scale (Full Model)

Second, in this model regression lines are assumed to be parallel, meaning that all IVs have the same effect on each level (Class) of the DV, a Brant (1990) test is applied to this model, showing that all probabilities are above 0.05 (Table 3) not rejecting the null hypothesis of parallel regression assumption. These assumptions make the ordered logit a more restrictive model than multinomial, but also more efficient.

Brant Test	**Omnibus**	**N. of Sales**	**Quantity**	**Likes**	**Replies**	**Negative**	**Neutral**	**Followers**	**Favourites**
Probability	0.75	0.77	1	0.73	0.32	0.84	0.09	0.76	0.86

Table 3 - Brant Test (Full Model)

4.2.2 Model Estimation

The estimation from the full model can be seen in Table 4. The parameters in the ordered model are interpreted in terms of the latent scale, as mentioned in section 3.2.1. The *polr* function from the MASS library in R does not estimate the significance of the variables, as before, in p-values, instead use of the LR test is recommended. The parameter estimates and sign can also be an indication of the effect of the different parameters on Class classification.

	Coefficients	Std. Error	Odds Ratio
N. of Sales	0.0002423	0.0001064	1.0002424
Quantity	-0.8876667	0.1210413	0.4116151
Likes	0.0055865	0.0045375	1.0056021
Replies	-0.0043895	0.0070047	0.9956202
Negative	0.1986853	0.0513730	1.2197981
Neutral	0.3209085	0.2968698	1.3783795
Followers	0.0011995	0.0006847	1.0012002
Favourites	-0.0004584	0.0008138	0.9995417

Table 4 - Ordered model parameter estimates

Oddly, replies and favourites are estimated to have a negative effect on class when all other variables are held constant. However, as the odds column shows, the values are very close to 1 and therefore it is likely they don't have a significant effect on price class. Likewise, e.g. likes and number of sales have a positive effect on price class, but a 1 unit increase in one of those variables, given the other variables are held constant, translate into the odds of Class 4 being just above 1.005 and 1.0002 times greater than the rest of the classes combined. This can lead to assuming that these variables' effect on price class could be not significant (odds ratio = 1). Quantity shows a larger effect on the DV than the previous variables, with a 1 unit increase in the available quantity of an NFT (semi-fungible tokens) the odds of an NFT being Class 4 versus the combined other classes are 0.411 times lower, when all other parameters are held constant. Likewise, the odds of an NFT being classes 3 and 4 combined vs classes 1 and 2 combined are 0.411 times lower when quantity increases by 1 unit. To better understand these effects and their significance additional nested and not nested models are estimated.

	LL	AIC	BIC
OpenSea	-940.36*	1890.727	1915.255
Twitter	-943.04	1904.080	1948.232
Full	-936.94*	1895.879	1949.843

Table 5 - In-sample predictive and goodness of fit metrics (Ordered model)

4.2.4 Model Discussion

Seeing the small effects of some of the variables in Table 4, we argue that the full model might have free parameters that have none to little effect on the DV, so additional models are estimated. Table 5 shows different fit metrics for the different models, the LR test is performed to determine if the models are significantly better than the null model. Here, the LR test indicates that a model with only the Twitter and eWOM sentiment parameters is not significantly different from a random selection (null model), indicating that these predictor variables do not add any explanatory power or offer a significant improvement in fit to the model, adding up to the indications seen in Table 4. Only the OpenSea model and the Full model fit the data significantly better than a random selection, with the Full model having the largest difference from the null mode. Nonetheless, the OpenSea model has a better AIC and BIC. Here, as before, we argue that the full model might have too many free parameters which is why the BIC and AIC are punishing the model and benefiting the ,more parsimonious, OpenSea model. Additionally, since the Twitter model seems to have no significant association with the DV, it is possible that the significance in the full model comes only from the OpenSea variables hence, through means of a trial and error test, we find that the main variable making our model better than a null model and exhibits significant explanatory power (price class) is the number of sales. Finally, for a more complete comparison Nagelkerke and McFadden's pseudo R^2 are calculated resulting in values really close to 0 for both models.

Considering the results presented above the OpenSea model seems to be a better fitting model. However, finding the best fitting model is out of the scope of this thesis, redirecting our focus to assessing the effect of the different parameters in the value of NFT and the different hypotheses.

4.3 Hypothesis Testing

Hypothesis 1 aimed to check the effect of Twitter quantitative data on the value of NFT. To evaluate this, two different sub-hypotheses regarding the influence of followers and different engagement variables were studied. Our multinomial ordered model exhibits no significant effect of Twitter variables on the price class of NFTs, with coefficients and pseudo R^2 close to 0 and odds-ratio close to 1. These findings contradict those of Arnak Kapoor et al (2022) on weak correlation between number of followers, retweets and likes and NFT value, rejecting H_{1a}, H_{1b} and consequently fully rejecting H_1.

Valence of a tweet was the Twitter variable included that measures eWOM sentiment. Hypotheses 2 and sub-hypotheses 2_a to 2_d were concerned with the effect of these variables on the price of NFTs. As observed in table 4, neutrality of a tweet displays a positive effect on price class with the coefficient being positive and odds-ratio over 1, however, as indicated by the LR tests done to the nested models this impact is not significant. Furthermore, negative valence of a tweet indicates a positive effect on the price class of NFTs, contradicting face-validity and the findings by Lee and Youn (2009), Vermeulen and Seegers (2009) and Rui et al. (2013), simultaneously. What's more and as previously discussed, by means of an LR test these effects prove to be not significant leading to the rejection of H_2 to H_{2d}.

The parameter number of sales shows interesting results, with a positive but close to 0 coefficient and an odds-ratio of just above 1. Number of sales shows a positive effect on the dependent variable, moreover, this effect is tested by means of trial and error and the LR test and these shows that the number of sales has a significant explanatory power, therefore accepting H_3, in accordance with Mieszko Mazur's ideas on NFTs return of investment.

Finally, quantity in semi-fungible tokens shows, as expected, a negative effect on the value of NFT with a negative coefficient and an odds-ratio below 1, this could be an indication of the importance of the concept of scarcity as explained by Raeesah Chonan et al (2021). Nevertheless, when dropping the quantity variable from the full and OpenSea model, no significant differences are detected in the LR test, leading to the assumption that much like the Twitter variables, quantity's explanatory power is not significant, rejecting H_4.

Hypothesis	Result
$H_{1:}$ *Twitter quantitative data will have a positive effect on NFT value*	Rejected
$H_{1a:}$ *Number of followers will have a positive effect on NFT value*	Rejected
$H_{1b:}$ *Number of retweets will have a positive effect on NFT value*	Rejected
H_{2}: *Valence of a tweet impacts NFT value*	Rejected
H_{2a}: *Tweets with a neutral sentiment will have a positive impact on NFT value*	Rejected
H_{2b}: *Positive sentiment towards an NFT project will lead to a higher NFT value*	Rejected
H_{2c}: *Negative sentiment towards an NFT project will lead to a lower NFT value*	Rejected
$H_{2d:}$ *Negative eWOM will have a larger impact on NFT value than positive eWOM*	Rejected
$H_{3:}$ *Number of sales has a positive effect on NFT value*	Accepted
$H_{4:}$ *Semi-fungible tokens will have a negative effect on the value of NFT*	Rejected

Table 6 - Hypothesis Overview

5. Discussion

This thesis intended to answer the following research question: *'To what extent can Twitter sentiment analysis help explain NFT value?'* and through existing academic work different hypotheses and variables were drawn.

The analysis reveals that the value of NFT is not being influenced by eWOM on Twitter. In fact, all parameters representing Twitter data, whether it was quantitative or qualitative, show no evidence of Twitter or the valence of tweets having any significant effects on the selling price of the studied NFTs. This is evidenced by the poor explanatory power of each variable when added to the full model and the only Twitter model itself.

Being Twitter the main social media network directing traffic to OpenSea it is odd to see Twitter variables having no effect on the value of the NFTs. One possible explanation for this is that the ideas on the effect of negative and positive eWOM on brand and purchase intention argued by Wu and Wang (2011) and Park et al (2008), simply do not translate equally to the NFT market. The same can occur with Alhidari et al (2015) ideas' on neutral sentiment conveying better informed purchase decisions.

The data shows most tweets come across as neutral (neutral valence) and very few have a negative valence, concurring with findings by Jansen et al (2009). However, tweets with a neutral valence still do not show a significant effect on the value of NFTs.
A possible explanation for this can be the very early stage of the NFT market. Still a largely unknown phenomena[5] creators and other parties are sharing their new creations, sales or price drops on Twitter, tweets that are simply informative and lack sentiment (Appendix - Figure 3). Additionally, NFTs can be considered investment pieces of relatively high risk (Mieszko Mazur, 2021) and studies have found that investment advice in social media can be influential but have little predictive value (Kadous et al, 2017), therefore Twitter influence might not translate into our model.
The data also shows large differences amongst the assets in price and number of sales, with 70% of the assets in our dataset selling for less than \$300 and only 0.7% of the NFTs selling for over \$10000. This large variability can be hard to capture only by the parameters used in this study and additional control variables can reveal different results. Another possible explanation comes from the volatility of the NFT market itself, as mentioned in section 2.4. NFTs are a young, volatile market with big disparities that sits on top of another volatile market (cryptocurrency) making it difficult to infer or predict value.
Overall, these results prove that there's still a long way to go with NFTs, opening a new, broad and interesting field of research that is still fairly untouched.

5.1 Limitations

The data used for this thesis was primarily collected from the OpenSea API. OpenSea sets strong restrictions on the information that can be collected, the amount of information that can be collected and who can collect this data. Through the use of Python 'GET' requests are limited by OpenSea to 4/sec with a default limit of 20 assets and capped at maximum 50 assets. As stated in the previous section, extra added control variables from OpenSea can improve model accuracy and estimation.
The sentiment analysis was done with a supervised, pre-trained, neural network model based on RoBERTa, this model is not specific to NFTs, leaving possible NFT related jargon (e.g. Gas) to be misinterpreted by the sentiment analysis. This could also relate to the unbalanced dataset, with the large majority of the tweets being neutral, interpretation and model estimation get compromised.

[5] https://finbold.com/study-90-of-adults-in-japan-dont-know-what-is-an-nft/
https://www.entrepreneur.com/article/423653

Another limitation comes from the MASS package in R not retrieving p-values, this makes the significance of the variables less clear and can lead to wrong inferences.

6. Conclusion

Through the use of an ordered logistic regression this thesis aimed to estimate the effect of different Twitter engagement variables and tweet sentiment on the value of NFTs.

This research paper establishes that these variables exhibit no significant observable explanatory power on the value of NFTs. An explanation for this behaviour could be the lack of polarity (positive vs negative) of sentiment present in the data with over 80% of the assets in the dataset being associated with neutral sentiments.

A significant effect on the value of NFT was found in the total number of sales an asset has, adding to the idea of NFTs as potential and interesting forms of investment. This thesis shows that, much like traditional art, the more an NFT is sold the larger the probability of its value increasing, proving that NFTs that are popular and highly requested can show interesting returns on investment.

Finally, this study is just a small stepping stone in the study of such a new market. Further research should be conducted with different control variables and different tools for sentiment analysis to study, to a greater extent, the possible effect of eWOM and sentiment analysis on the NFT market. Additionally, another potential area of research can be found in the study of the volatility of the market and the understanding of which factors affect the greater success of some assets over others as the data has shown big differences in price and number of sales between different assets.

References

Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). *Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis*. *1*(3), 22.

Alhidari, A., Iyer, P., & Paswan, A. (2015). Personal level antecedents of eWOM and purchase intention, on social networking sites. *Journal of Customer Behaviour*, *14*(2), 107–125. https://doi.org/10.1362/147539215X14373846805707

Alreck, P. L., & Settle, R. B. (1995). The importance of word-of-mouth communications to service buyers. In *Proceedings of American Marketing Association* (Vol. 6, pp. 188-193). Chicago, IL: American Marketing Association.

Anderson, T. R., & Slotkin, T. A. (1975). Maturation of the adrenal medulla—IV. Effects of morphine. *Biochemical Pharmacology*, *24*(16), 1469–1474. https://doi.org/10.1016/0006-2952(75)90020-9

Arndt, J. (1967). Role of Product-Related Conversations in the Diffusion of a New Product. *Journal of Marketing Research*, *4*(3), 291–295. https://doi.org/10.1177/002224376700400308

Aswani, R., Kar, A. K., Ilavarasan, P. V., & Dwivedi, Y. K. (2018). Search engine marketing is not all gold: Insights from Twitter and SEOClerks. *International Journal of Information Management*, *38*(1), 107–116. https://doi.org/10.1016/j.ijinfomgt.2017.07.005

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is Stronger than Good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Blattberg, R. C., Kim, B.-D., & Neslin, S. A. (2008). Statistical Issues in Predictive Modeling. In R. C. Blattberg, B.-D. Kim, & S. A. Neslin, *Database Marketing* (Vol. 18, pp. 291–321). Springer New York. https://doi.org/10.1007/978-0-387-72579-6_11

Bose, K. S., & Sarma, R. H. (1975). Delineation of the intimate details of the backbone conformation of pyridine nucleotide coenzymes in aqueous solution. *Biochemical and Biophysical Research Communications*, *66*(4), 1173–1179. https://doi.org/10.1016/0006-291x(75)90482-9

Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics*, *46*(4), 1171. https://doi.org/10.2307/2532457

Brown, J. J., & Reingen, P. H. (1987). Social Ties and Word-of-Mouth Referral Behavior. *Journal of Consumer Research*, *14*(3), 350. https://doi.org/10.1086/209118

Bulearca, M., & Bulearca, S. (n.d.). *Twitter: A Viable Marketing Tool for SMEs? 2*(4), 15.

Campbell, Rachel. (2008). Art as a Financial Investment. *The Journal of Alternative Investments*, *10*(4), 64–81. https://doi.org/10.3905/jai.2008.705533

Cha, M. (n.d.). *Measuring User Influence in Twitter: The Million Follower Fallacy*. 8.

Chakraborty, I. S. K. (n.d.). *Three Essays on the Role of Unstructured Data in Marketing Research*. 147.

Chohan, R., & Paschen, J. (2021). What marketers need to know about non-fungible tokens (NFTs). *Business Horizons*, S0007681321002202. https://doi.org/10.1016/j.bushor.2021.12.004

Cooke, M., & Buckley, N. (2008). Web 2.0, Social Networks and the Future of Market Research. *International Journal of Market Research*, *50*(2), 267–292. https://doi.org/10.1177/147078530805000208

Craney, T. A., & Surles, J. G. (2002). Model-Dependent Variance Inflation Factor Cutoff Values. *Quality Engineering*, *14*(3), 391–403. https://doi.org/10.1081/QEN-120001878

Daniels, L. K. (1976). Rapid in-office and in-vivo desensitization of an injection phobia utilizing hypnosis. *The American Journal of Clinical Hypnosis*, *18*(3), 200–203. https://doi.org/10.1080/00029157.1976.10403798

Davidson, D. (2009). How to… put a price on your social media strategy. *Revolution Magazine–The Insider's Guide to Digital Marketing*, *2009*, 29-33.

Di Pierro, M. (2017). What Is the Blockchain? *Computing in Science & Engineering*, *19*(5), 92–95. https://doi.org/10.1109/MCSE.2017.3421554

Doh, S.-J., & Hwang, J.-S. (2009). How Consumers Evaluate eWOM (Electronic Word-of-Mouth) Messages. *CyberPsychology & Behavior*, *12*(2), 193–197. https://doi.org/10.1089/cpb.2008.0109

Dorbian, I. (2010). Social media taps its way into B2B marketing plans. *Min's B2B*, *13*(5), 5-6.

Engel, J. F. (n.d.). *How information is used to adopt and innovation*.

Franses, P. H., & Paap, R. (2001). *Quantitative models in marketing research*. Cambridge University Press.

Gefen, D., & Pavlou, P. A. (2012). The Boundaries of Trust and Risk: The Quadratic Moderating Role of Institutional Structures. *Information Systems Research*, *23*(3-part-2), 940–959. https://doi.org/10.1287/isre.1110.0395

Giles, M. (2010). A world of connections. *The Economist*, *394*(8667), 3.

Gilly, M. C., Graham, J. L., Wolfinbarger, M. F., & Yale, L. J. (1998). A Dyadic Study of Interpersonal Information Search. *Journal of the Academy of Marketing Science*, *26*(2), 83–100. https://doi.org/10.1177/0092070398262001

Gu, T. (2020). *Analyzing Unstructured Data for Marketing Insights* (Doctoral dissertation, The University of Arizona).

Hartmann, J., Heitmann, M., Schamp, C., & Netzer, O. (2021). The Power of Brand Selfies. *Journal of Marketing Research*, *58*(6), 1159–1177. https://doi.org/10.1177/00222437211037258

Hausmann, A. (2012). Creating 'buzz': Opportunities and limitations of social media for arts institutions and their viral marketing: Creating 'buzz'. *International Journal of Nonprofit and Voluntary Sector Marketing*, *17*(3), 173–182. https://doi.org/10.1002/nvsm.1420

Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2015). Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science*, *43*(3), 375–394. https://doi.org/10.1007/s11747-014-0388-3

Hilbe, J. M. (2009). *Logistic Regression Models* (0 ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781420075779

Homer, P. M., & Yoon, S.-G. (1992). Message Framing and the Interrelationships among Ad-Based Feelings, Affect, and Cognition. *Journal of Advertising*, *21*(1), 19–33. https://doi.org/10.1080/00913367.1992.10673357

Hung, K. H., & Li, S. Y. (2007). The Influence of eWOM on Virtual Consumer Communities: Social Capital, Consumer Learning, and Behavioral Outcomes. *Journal of Advertising Research*, *47*(4), 485–495. https://doi.org/10.2501/S002184990707050X

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, *60*(11), 2169–2188. https://doi.org/10.1002/asi.21149

Jiang, L., & Erdem, M. (2017). Twitter-marketing in multi-unit restaurants: Is it a viable marketing tool? *Journal of Foodservice Business Research*, *20*(5), 568–578. https://doi.org/10.1080/15378020.2016.1222746

Jin, S.-A. A., & Phua, J. (2014). Following Celebrities' Tweets About Brands: The Impact of Twitter-Based Electronic Word-of-Mouth on Consumers' Source Credibility

Perception, Buying Intention, and Social Identification With Celebrities. *Journal of Advertising*, *43*(2), 181–195. https://doi.org/10.1080/00913367.2013.827606

Kadous, K., Mercer, M., & Zhou, Y. (2017). Undue influence? The effect of social media advice on investment decisions.

Kapoor, A., Guhathakurta, D., Mathur, M., Yadav, R., Gupta, M., & Kumaraguru, P. (2022). TweetBoost: Influence of Social Media on NFT Valuation. *ArXiv:2201.08373 [Cs]*. http://arxiv.org/abs/2201.08373

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, *44*(2), 544–564. https://doi.org/10.1016/j.dss.2007.07.001

Kim, E., Sung, Y., & Kang, H. (2014). Brand followers' retweeting behavior on Twitter: How brand relationships influence brand electronic word-of-mouth. *Computers in Human Behavior*, *37*, 18–25. https://doi.org/10.1016/j.chb.2014.04.020

King, R. A., Racherla, P., & Bush, V. D. (2014). What We Know and Don't Know About Online Word-of-Mouth: A Review and Synthesis of the Literature. *Journal of Interactive Marketing*, *28*(3), 167–183. https://doi.org/10.1016/j.intmar.2014.02.001

Lee, M., & Youn, S. (2009). Electronic word of mouth (eWOM): How eWOM platforms influence consumer product judgement. *International Journal of Advertising*, *28*(3), 473–499. https://doi.org/10.2501/S0265048709200709

Leeflang, P. S. H., Wieringa, J. E., Bijmolt, T. H. A., & Pauwels, K. H. (Eds.). (2017). *Advanced Methods for Modeling Markets*. Springer International Publishing. https://doi.org/10.1007/978-3-319-53469-5

Li, X., & Shi, M. (2019). Video mining: Measuring visual information using automatic methods. *International Journal of Research in Marketing*, 16.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458–468. https://doi.org/10.1016/j.tourman.2007.05.011

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. https://doi.org/10.48550/ARXIV.1907.11692

Mazur, M. (2021). Non-Fungible Tokens (NFT). The Analysis of Risk and Return. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3953535

Mei, J., & Moses, M. (2005). Beautiful asset: Art as investment. *Journal of Investment Consulting*, *7*(2), 45-51.

Mitchell, A. A., & Olson, J. C. (1981). Are Product Attribute Beliefs the Only Mediator of Advertising Effects on Brand Attitude? *Journal of Marketing Research*, *18*(3), 318–332. https://doi.org/10.1177/002224378101800306

Mizerski, R. W. (1982). An Attribution Explanation of the Disproportionate Influence of Unfavorable Information. *Journal of Consumer Research*, *9*(3), 301. https://doi.org/10.1086/208925

Naeem, M. A., Mbarki, I., & Shahzad, S. J. H. (2021). Predictive role of online investor sentiment for cryptocurrency market: Evidence from happiness and fears. *International Review of Economics & Finance*, *73*, 496–514. https://doi.org/10.1016/j.iref.2021.01.008

Palmer, A., & Koenig‐Lewis, N. (2009). An experiential, social network‐based approach to direct marketing. *Direct Marketing: An International Journal*, *3*(3), 162–176. https://doi.org/10.1108/17505930910985116

Park, D.-H., & Kim, S. (2008). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, *7*(4), 399–410. https://doi.org/10.1016/j.elerap.2007.12.001

Richins, M. L. (1984). Word of mouth communication as negative information. *ACR North American Advances*.

Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2

Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems*, *55*(4), 863–870. https://doi.org/10.1016/j.dss.2012.12.022

See-To, E. W. K., & Ho, K. K. W. (2014). Value co-creation and purchase intention in social network sites: The role of electronic Word-of-Mouth and trust – A theoretical analysis. *Computers in Human Behavior*, *31*, 182–189. https://doi.org/10.1016/j.chb.2013.10.013

Seifert, C., & Kwon, W.-S. (2019). SNS eWOM sentiment: Impacts on brand value co-creation and trust. *Marketing Intelligence & Planning*, *38*(1), 89–102. https://doi.org/10.1108/MIP-11-2018-0533

Serra Cantallops, A., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, *36*, 41–51. https://doi.org/10.1016/j.ijhm.2013.08.007

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, *52*(4), 689–699. https://doi.org/10.1037/0022-3514.52.4.689

Sun, T., Youn, S., Wu, G., & Kuntaraporn, M. (2006). Online Word-of-Mouth (or Mouse): An Exploration of Its Antecedents and Consequences. *Journal of Computer-Mediated Communication*, *11*(4), 1104–1127. https://doi.org/10.1111/j.1083-6101.2006.00310.x

Uygun, Y., Oruc, M. F., & Sefer, E. NFT Sales Price Prediction Using Visual Feature Extraction.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, *30*(1), 123–127. https://doi.org/10.1016/j.tourman.2008.04.008

Wang, Q., Li, R., Wang, Q., & Chen, S. (2021). Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges. *ArXiv:2105.07447 [Cs]*. http://arxiv.org/abs/2105.07447

Watkins, I. (n.d.). *NON-FUNGIBLE TOKENS (NFT) AND THE EVOLUTION OF ART*. 3.

Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology*, *59*(5), 555–561. https://doi.org/10.1037/h0037186

Weinberger, M. G., Allen, C. T., & Dillon, W. R. (1981). Negative information: Perspectives and research directions. *ACR North American Advances*.

Withiam, G. (2011). Social media and the hospitality industry: Holding the tiger by the tail.

Wu, P. C. S., & Wang, Y. (2011). The influences of electronic word‐of‐mouth message appeal and message source credibility on brand attitude. *Asia Pacific Journal of Marketing and Logistics*, *23*(4), 448–472. https://doi.org/10.1108/13555851111165020

Zhao, D., & Rosson, M. B. (2009). How and why people Twitter: The role that micro-blogging plays in informal communication at work. *Proceedinfs of the ACM 2009 International Conference on Supporting Group Work - GROUP '09*, 243. https://doi.org/10.1145/1531674.1531710

Appendix

Figure 3 - Informative Tweet