ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Data Science and Marketing Analytics

Cryptocurrency and Purchase Intention based on Tweet properties

A study on the influence of Tweets on the Purchase Intention of First-time Millennial Investors

Name: Rohan Ghosh

Student ID: 497188

Supervisor: Marina Khismatullina

Second assessor: Vardan Avagyan

Abstract

For companies and brands in the cryptocurrency and altcoin space, the importance of Twitter and Tweets are of utmost importance in convincing users to invest in their product(s). Many altcoins are being marketed and promoted on Twitter and thus the focus on XRP creates relevance for the study on altcoins. Understanding how best to manipulate tweets from a company standpoint and what changes, small or big, to make with respect to how the tweets are presented to these users can significantly influence whether a user will invest in their altcoin. Hence, this study focuses on the largest demographic of cryptocurrency investors: millennials and examines how first-time millennial cryptocurrency investors are impacted by the different tweet property (likes, retweets, etc.) values that are seen at first glance. Thus, a choice based conjoint framework was constructed using Conjointly and distributed on Qualtrics in a survey format that collected 282 respondents of the target cohort. The choice-based analyses, supervised machine learning model of logistic regression was used to analyse the data. The study found that while likes, retweets and blue tick verification were significant and positively influenced the users, it was not the case for quote tweets which had a lesser but also significant negative impact on the purchase intention probability of a tweet influencing a user to want to purchase XRP. The study concludes with recommendations for both companies and marketeers alike as well as limitations to be understood regarding this study. These recommendations might aid them in more effectively promoting their altcoins on Twitter for the target cohort.

Table of Contents

Introduction

1.1 Problem Statement

Since the origin of cryptocurrency with Bitcoin in 2009, Twitter has assumed a crucial role in providing information about cryptocurrency and its developments due to the reluctance and lack of interest from traditional media sources to do so themselves (Nakamoto, 2018; Kraaijeveld & de Smedt, 2020). Understanding how to properly utilize Twitter and its supply of information to the general public and cryptocurrency investors is critical due to its aforementioned integral role as a platform that enables the provision of information to the public. However, there is a lack of research on how to do so. The majority of research conducted on Twitter and its capability to influence cryptocurrency pertains solely to the content of the tweet and its sentiment, thus the focus of this study will be on how the properties of tweets (e.g. likes) influence investors (Kraaijeveld & de Smedt, 2020; Naeem et al., 2020; Edrogan & Canayaz, 2018). While the area of tweet sentiment analysis is extensively researched, investigating what other aspects of a tweet drive the influence on purchase intention for first-time investors with no cryptocurrency investing history has not been focused on or explored in this same way before. The importance in taking this approach comes with unique insights, ranging from how to attract new investors to investing in new kinds of cryptocurrency through the manipulation of tweet properties which are further discussed in Section 2.4. Moreover, rather than solely Bitcoin (the cryptocurrency with the largest market cap), the use of a relatively smaller coin would make for a more interesting research design where newer, emerging cryptocurrencies are focused upon, which is why this study implements and has a specific focus on a relatively smaller cryptocurrency: Ripple (XRP).

Ciaian et al. (2018) denote that there is a high interdependency between altcoins (meaning alternative coins other than bitcoin) and bitcoin prices in both the long and short term, however, Corbet et al. (2020) argued that the extent of interdependency varies based on the altcoin's unique features. There has been frequent discussion on the interdependence between bitcoin and altcoin prices, however, there continues to be debate as to how interdependent they are (Aysan et al., 2021; Demir et al., 2021; Meynkhard, 2020; Hu et al., 2019). Moreover, while there has been studies on what factors influence investor purchase intention on cryptocurrency, there has been little study on investor attitudes with a focus on purchase intention that highlights the difference between bitcoin and altcoins thus there is no conclusive information regarding this. Instead, there are only studies that focus exclusively on Bitcoin's attitudes and purchase intention with no focus or comparison with Altcoins (Kaplan, C. et al, 2018). Irrespective of the lack of conclusive information on the relationship between altcoins and bitcoin, this study will use a single altcoin and will hold it constant throughout (XRP) which will be done to ensure a more realistic depiction of how altcoins are interacted with, by potential investors, rather than focusing on bitcoin. This will be done in which the only exposure that survey recipients will have during the data collection stage will be solely to the altcoin XRP to ensure that the effect of this study is focused only on altcoins.

The proposed research examines the properties of a tweet while holding sentiment/content, age, investment experience and price (keeping it constant) to focus on the effect on properties of tweets such as likes and retweets shown in tweets that users would be exposed to when interacting with a tweet in real life. Users on social media and Twitter are known to have shorter attention spans as put forth by various studies and consequently, the first few stimuli they see when scrolling through tweets will be the tweet property values (Ventola et al., 2014; von Muhlen et al., 2012; Gabielkov et al., 2016). Thus, the properties were chosen on the basis that this is everything a user sees when they click on any given tweet to allow for a realistic simulation of a user interacting with a tweet. These properties of tweets (independent variables) are: the likes of the tweet, amount of retweets, quote tweets, whether the user is verified on Twitter (blue tick) and how these factors would influence the purchase intention of someone interacting with the tweet while also determining what tweet properties most motivate them to invest in Ripple (XRP) for first-time millennial investors (Daga et al., 2020; Elendner et al., 2018). These properties are further discussed in Section 2.4. Additionally, it is important to note that the tweet property "Replies" was excluded for two main reasons. Firstly, unlike the other properties mentioned, the number of replies is not immediately visible when a user clicks on a tweet. Secondly, the content of the replies are immediately visible below the tweet itself. This makes its effect on purchase intention fundamentally different from the other tweet properties.

1.2 Research Question

This paper, therefore, investigates the influence of tweets on first-time millennial investors, thereby providing a benchmark for firms using Twitter to reach goals and targets related to increasing their investor base by allocating resources to tailoring their tweets. As such, the research question for this paper is:

“How do the properties and social media engagement of a Twitter ‘Tweet’ influence the purchase intention of first-time millennial investors for the Ripple (XRP) coin?”

Literature Review

2.1 Social Media and Twitter’s Importance with Cryptocurrency

Firstly, throughout history, there has been a multitude of communication channels that have stood in the foreground, influencing society as well as daily life. In the 21st Century, social media has taken this mantle, a platform that provides individuals to express their opinions, beliefs, and ideas in a revolutionary way that can reach a vast audience in the blink of an eye. The rapid growth of companies that leverage this industry and continue to innovate like Twitter, Meta and TikTok has ushered in the era of social media. In essence, social media is any type of online content or media that stimulates participation, evokes open discussion and establishes a sense of community (M. Saravanakumar & T. Sugantha Lakshmi, 2012).

Cryptocurrencies are vastly different and are a digital asset class that leverage blockchain technology, a decentralized and innovative technology that enables the digitalization of transparency and trust. Blockchain technology obsoletes financial institutions and governments in their roles as producers of currency and as intermediary entities that verify transactions. The cryptocurrency space was created on the 3rd of January in 2009 with the launch of Bitcoin (BTC). It was only throughout 2017 and early 2018 that a large surge in interest and innovation in the space took place. This was a consequence of the trend in media coverage of the unprecedented returns that created a subsequent pseudo-gold rush (Nakamoto, 2018). As a result of the cryptocurrency being so young and new, traditional media outlets did not always report events promptly, which created a gap for which social media outlets became primary sources of information for cryptocurrency investors. Smaller news outlets and influencers in the cryptocurrency space filled the need for information in the industry. The platform that enables this transfer and reporting of information is Twitter, in which Twitter users 'tweet' (write anything in text and send it out on the platform for anyone to read), which denotes its importance in its ability to attract new investors for cryptocurrency as well and is why it will be the focus of this paper (Kraaijeveld & de Smedt, 2020). In this paper, the primary focus on Twitter will be the way that tweets are presented to potential first-time investors and how the information in a tweet influences the purchase intention of the latter on cryptocurrency.

2.2 The importance and relevance of the Ripple (XRP) coin

Moreover, the Ripple (XRP) coin is of notable interest in the cryptocurrency space with its use case. It leverages blockchain technology, evolving completely independently from BTC, with its main purpose to act as both a payment system and digital currency. Ripple's main function pertains to sending international remittance payments and how its unique application with blockchain ledger technology provides it disruptive advantages to its current main competitor, SWIFT (Armknecht et al., 2015). While SWIFT are the current market leader in the provision of the remittance service, Ripple is quickly growing in reaching its potential to drastically improve. The benefits Ripple provides include low transaction costs, real-time delivery and 2x7 service availability (Rosner & Kang, 2015; Qiu et al., 2019). Thus, as a trending cryptocurrency coin, XRP will be used as a focal point for this study to provide insight as to how larger, emerging coins such as XRP interact with potential investors. Lastly, there are extremely few studies that focus on XRP's interaction with Twitter's platform that does not pertain primarily to the sentiment analysis of tweets themselves as aforementioned, which indicates that a study of the XRP coin offers novelty in its research.

2.3 Millennial Investors and Their Relevance

While the majority of investors in cryptocurrency are of the Generation Z demographic (1997-2012 born), the average spending of a Gen Z investor is the lowest when compared to Generation X (1965-1980 born) and Millennials (1981-1996 born) investors, as a result of them not having reached their earnings potential (being older means a greater likelihood of a higher income and a thus greater propensity for investment per individual). Hence, in the short-term, there is a great incentive for Owners of cryptocurrencies and their affiliates to encourage investment in attracting Millennial and Gen X investors. Gen X buyers purchased an average of $9,611 crypto last year while Millennials bought $8,596 and Gen Z investors bought $6,120 (Gogol, 2021).

Due to the lack of availability of data for owners of Ripple (XRP), looking at the current cryptocurrency market instead, the most common investor in the market are millennial investors, consisting of 76.45% of the market (Gogol, 2021). Various studies argue that the reason for millennials being the demographic most invested in cryptocurrency is that their age indicates maturity and financial literacy that enables them to be more receptive to learning about and investing in a new modern digital financial class such as cryptocurrency (Bhilawadikar & Garg, 2020; Patil, 2019; Sanders, 2022). On top of this, millennials have also claimed that there is social exposure on the social media platforms that they use that motivates their purchasing behaviour of cryptocurrency (Gogol, 2021). Consequently, understanding how users leverage Twitter's platform to influence the purchase intention of Millennials specifically will provide deeper insight into how to target this specific demographic. The studies aforementioned also suggest they are of higher likelihood to invest in new technologies and are thus an ideal target market for companies investing in blockchain and cryptocurrency.

2.4 Tweet Properties (Independent Variables)

When focusing on the Tweet properties that will be included in the survey that will be sent out, the idea is to determine how important they are in a Tweet. Below are the reasons that the specified tweet properties are important to users as well as why they are relevant in influencing users that interact with tweets with their purchase intention.

Trust has played a central role in previous research regarding looking at the influence of social media engagement on online purchase decisions. Weisberg, Te'eni and Arman (2011) studied the value of trust and its influence on consumer past-purchase behaviour in the context of e-commerce and this is applicable in the context of tweets and how they can influence purchase intention. Rousseau et al. (1998) define trust as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions of the behaviour of another". This trust can be built on Twitter with the presence of

specific properties of tweets. These properties can increase the trustworthiness or conversely decrease said trust depending on their characteristics and context.
More than solely credibility and trust being driving factors for tweet properties, in the modern era with huge levels of exposure to mass stimuli on social media, the levels of attention for each tweet are scarce in the context of Twitter. Attention can be seen as a finite currency spread across all the tweets they interact with (Ventola et al., 2014; von Muhlen et al., 2012; Gabielkov et al., 2016). This attention enables users to retain the information they gain on a priority basis based on the more attentive they are to the tweet. Additionally, the levels of interest that a user has in the tweet they interact with have a similar effect on their purchase intention as a consumer of the product being discussed in the tweet. As a result, understanding how tweet properties can make users more attentive or more interested in the information they are receiving in a tweet based on the tweet properties is argued to be a driving factor in their purchase intention (Ko et al., 2015 Richard & Guppy, 2014).
Below are the three properties of the tweets that will be focused on. These were selected as the features that users see without taking any action when they come into contact with a tweet:

2.4.1 Tweet Likes

Building on the aforementioned explanation of Social media, in a marketing context, social media acts as a means of communication and promotion between customers and firms in different ways. The first way is through the nature of social media applications and their functionality. The connectedness is enabled by networking sites such as social networking sites (e.g, Facebook, Snapchat, etc.), microblogging sites (e.g, Reddit, Tumblr, Twitter, etc.), and content platforms (Youtube, Instagram, Vimeo, etc.). It is from the likes of these social media outlets that social connectedness is built (Kaplan & Haenlein, 2010).
'Liking' is a feature on Twitter that enables a user to engage with a tweet. This feature/action can be seen across all forms of social media and has been studied quite thoroughly on how consumer engagement, measured in likes, influences purchase intention of products and the existing research dictates that likes positively influence purchase intention and thus as there are more likes on a post/tweet on social media platforms, consumers are more likely to purchase the products being marketed due to the increased level of interest associated with the tweets made possible by the number of likes (Mas'od et al., 2019; Hutter et al., 2013; Kırcova et al., 2021). The justification for likes increasing purchase intention is the same as

mentioned throughout the studies, either partially or explicitly explained in which the likes' take the role of a metric of quality assurance for the social media being posted, in which the more likes means the tweet is more credible, in line with the central theme of the study put forth by Weisberg et al. (2011).

However, there is little information on how likes influence the purchase intention of assets like stocks or cryptocurrency. Hence, this hypothesis is formed through extrapolating the findings from how existing literature denotes how likes impact consumer purchase intention for goods and services and consumer purchase intention/intent to own cryptocurrency.

H1: The amount of likes associated with a Tweet will positively influence the purchase intention of XRP.

2.4.2 Tweet Retweets

A 'retweet' is a value that is shown denoting the number of re-shares of the tweet. A re-share is a user that interacts with the tweet and shares that same tweet with their network. The amount of retweets for any given tweet is shown as one of the values when a user interacts with a tweet.

It is important to once again note that Twitter Replies were not included in the study as they provided the same information as the retweet count at a glance, however, the replies only provided more information upon reading them due to the heavy importance placed on the sentiment of the reply to extract information (Kim J. et al., 2012). However, the study is focused solely on the stimuli and properties of a tweet that are made available upon initial interaction to see what users might experience within a short period. This is the case due to Twitter users' shorter attention span on the app, as such trying to deduce the important aspects of a tweet from quick and initial interactions gauging attention and interest to closely mimic reality (Ventola C. L. et al., 2014; von Muhlen et al, 2012).

Once again, there is little information regarding how retweets might influence purchase intention for cryptocurrencies in contrast to goods and services as were studied in the papers mentioned, however, in essence, the meaning of the retweets is argued to be universal by Metaxas, P., et al., (2021): "retweeting indicates not only interest in a message, but also trust in the message and the originator, and agreement with the message contents". Thus, given that the tweet has a positive sentiment about XRP (as will be controlled for), the second hypothesis dictates:

H2: The amount of retweets associated with a Tweet will positively influence the purchase intention of XRP.

2.4.3 Quote Tweets

A ‘quote tweet’ much like a retweet is a value that is shown denoting the number of re-shares of the tweet. However, the key difference is that Quote Tweets allow users to quote the tweet they share and add a short comment (Garimella et al., 2016). This difference, although subtle, is key as it opens up a range of new use cases that drastically changes the connotation of the tweet property as will be explained in the next paragraph. The amount of Quote Tweets for any given tweet is shown as one of the values when a user interacts with a tweet.

Limited research exists about Quote Tweets because the feature was introduced by Twitter as recently as 2015. Research about Quote Tweets in the crypto space is even more scarce. This being said, an illuminating study by Matalon et al. (2021) found that quoted tweets tend to follow the Opinion Inversion (O.I.) phenomenon, meaning that people tend to disagree with the original content of the source tweet they were quoting. In other words, Matalon et al. (2021) seem to find that retweets work as a tentative “dislike” button. However, it is important to keep in mind these findings were based on research in political communication and not cryptocurrency. Nevertheless, the basic notion can be cautiously extrapolated to the context at hand implying that a higher number of quoted tweets, given that the source tweet promotes the purchase of XRP, will dissuade consumers from buying the cryptocurrency as it indicates that many people disagree with the content of the tweet. Thus, given that the tweet has a positive sentiment about XRP (as will be controlled for), the third hypothesis dictates:

H3: The amount of Quote Tweets associated with a Tweet will negatively influence the purchase intention of XRP.

2.4.4 Blue Tick Verification

In 2009, Twitter responded to claims of impersonations on their platform by implementing the blue tick verification status, providing a solution to a context of illegitimacy on the platform (Kanalley, 2013). In principle, the blue tick verification is afforded to public figures with substantial followings, however, over the years the process for verification has remained arbitrary and has had controversy surrounding it regarding whether Twitter only verifies users whom they want on their platform or if there is a standardized verification process.

Consequently, there are quite a few users with large followings that condemn the verification and see it as an insult due to the seemingly unknown process for verification (Bishop 2017; Fairchild P., 2018). “The Blue Tick Brigade” is a pejorative term referring to users that are blue tick verified, referring to them as the “elite” of media professionals that are hand-picked by Twitter as if selected to act as ambassadors for Twitter’s brand. There are even some verified users who see this verification bestowed on them as a means of conformity to Twitter’s political agenda (Maragkou, E., 2019).
Thus, while the aim of the blue tick verification for a user on Twitter was to increase the credibility and reliability of said users, thus suggesting further weight being given to users’ tweets and consequent higher purchase intention for XRP, the references mentioned have indicated otherwise. The increase in credibility would enhance the influence of the tweet, however, it seems that Twitter’s inconsistency and lack of transparency with the verification process have created a negative connotation for the verification process in the eyes of Twitter users, both verified and unverified alike (Weisberg, Te’eni and Arman, 2011). Thus, it is expected that:

H4: The user of a tweet being blue-ticked verified negatively influences the purchase intention of XRP.

2.5 Academic Relevance

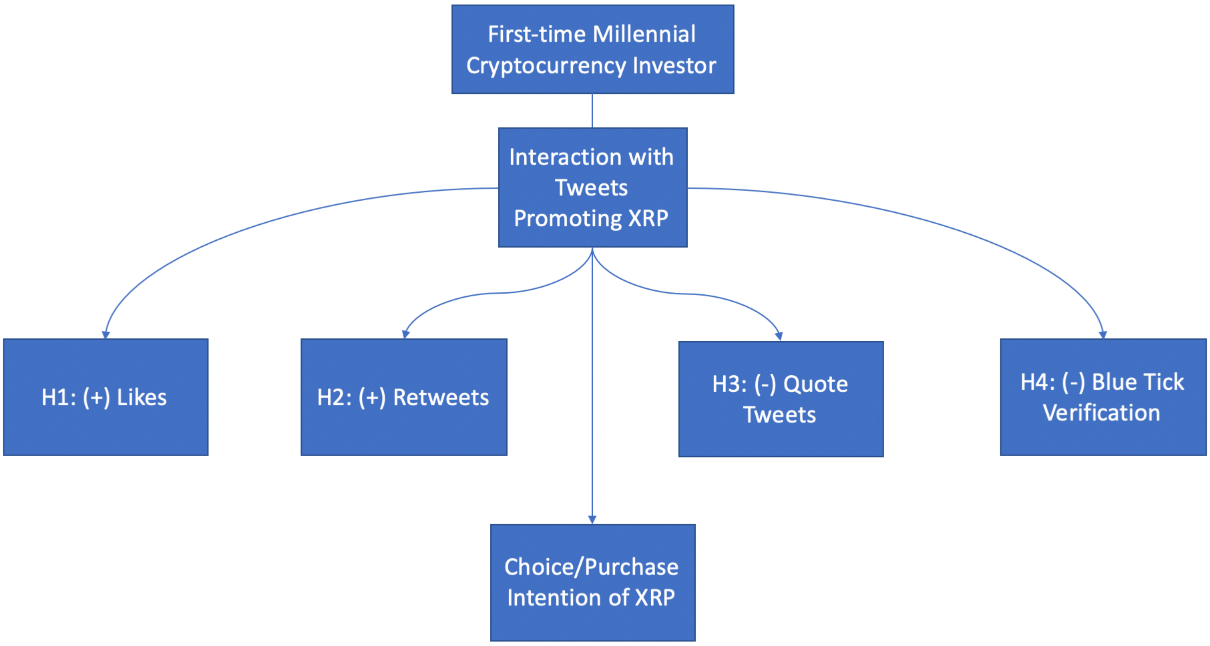
As aforementioned, Twitter’s role in transferring and reporting cryptocurrency information is crucial and thus its focus in this paper as a platform to influence potential investors is significant. There are a plethora of papers that focus on Twitter but fixate solely on sentiment analysis and how it influences investor sentiment and the price of financial assets, however, there seems to be a large gap in research that corresponds to how specific engagement with tweets influences purchase intention, as most research with Twitter and Cryptocurrency is saturated with sentiment analysis studies (Zimbra et al., 2018; Devi, G. D., & Kamalakkannan, S., 2020; Kraaijeveld & de Smedt, 2020; Naeem et al., 2020; Erdogan & Canayaz, 2018). More specifically, there is no research on how the properties (e.g, likes, retweets, etc.) of tweets influence how tweets impact millennial, first-time investor Twitter users’ purchase intention. There exists a gap in understanding how this specific demographic is influenced by the previously mentioned twitter properties and how they might affect one’s purchase intention for an altcoin such as XRP. The academic space is saturated with studies on sentiment analysis and price. Thus, an approach focusing on the way in-built metrics on the platform (Twitter properties) work with users, offers novelty. To reiterate, there is a severe lack of research about how Twitter metrics such as the ones focused on in this paper (likes, retweets and user blue tick verification) might impact consumer attitudes and thus purchase intention for cryptocurrency. On top of this, there is no focus for tweets focused on altcoins that incorporate the relevance of these Twitter properties, and hence this paper offers the chance to explore a gap in the research around how these properties influence a specific demographic of users around the purchase intention of an altcoin – specifically XRP.

Understanding how to use Twitter properties (having more/fewer likes for example) can be a significant advantage in the market for entities trying to leverage the platform to appeal to potential investors for their new or existing altcoin (Kraaijeveld & de Smedt, 2020). Furthermore, using XRP as the central focus of this study is to take a perspective focused on how altcoins operate in the space. This is important to mention as the results and conclusions from this study will be most relevant in application for companies working with altcoins and emerging coins and the focus on XRP thus provides novelty in the research. Particularly, the utilization of Choice-Based Conjoint analysis is also quite scarcely applied to social media. Lastly, the focus on millennials develops the novelty and specificity of this study, motivated by their high average investment ($8,596) in cryptocurrency. They make up the majority of the crypto investor market share (76.45%) (Gogol, 2021). These stats motivate the relevance of millennials and suggest that existing crypto firms should be targeting this demographic as they seem to be.

2.6 Hypotheses Formation

Figure 1

Conceptual Research Model



H_1: The amount of likes associated with a Tweet will positively influence the purchase intention of XRP.

H_2: The amount of retweets associated with a Tweet will positively influence the purchase intention of XRP.

H_3: The amount of Quote Tweets associated with a Tweet will negatively influence the purchase intention of XRP.

H_4:The user of a tweet being blue-ticked verified negatively influences the purchase intention of XRP.

Research Methodology and Data

3.1 Research Method

In accordance with the paper put forth by Malhotra (2006), a survey is defined as "a formalized set of questions for obtaining information from respondents and it enables quantitative data to be collected in a standardized way so that the data is internally consistent and coherent for analysis". Additionally, in contrast to observational data, descriptive data collected using a survey facilitates a greater understanding of respondent perspectives. Within budget, size and time constraints, surveys aid in collecting data that is easy to analyze, code and interpret, as well as maintaining high reliability as opposed to longitudinal studies and experiments (Couper et al., 2004). This is due to standardization in the survey experience in which identical questions, layout and user interface are all utilized for all respondents. Fox et al. (1988) in their study suggest that sponsorship from educational institutions increase response rates in surveys. Therefore, the online-web survey tool Qualtrics, as mentioned earlier, will be utilized with provision from the Erasmus University Rotterdam to create, monitor and export the survey.

The data obtained from the survey will attempt to quantify how much potential first-time millennial investors value the four different properties of a tweet when a tweet with positive sentiment is being made about XRP and how different properties contribute to the purchase intention of XRP as a result. The choice sets for the conjoint analysis were created using Conjointly, which is explained in detail in Section 3.4.1 and the choice sets were implemented in a survey that was created and distributed using Qualtrics. The complete questionnaire can be found in Appendix B.

3.2 Variables

3.2.1 Dependent Variable - Choice/Purchase intention

The dependent variable will be a binary variable called Purchase Intention that is calculated based on the Choice Based Conjoint model that is further discussed in Section 3.4. The other questions will be regarding the tweets that will be shown and their respective properties as well as the respondent's characteristics to gauge what properties of tweets are most important for a first-time millennial investor when they are investing in XRP.

3.2.2 Demographic Variables

The demographic variables aid in providing context for the sample of the survey being created and sent out. They are not included in the logistic regression model as the model fixates on the attributes/tweet properties and their respective levels rather than the individual demographic factors of the respondents. Thus, while gender and income are not included in the model, they are used to describe the sample used for analysis.

3.2.2.1 Gender of Respondent

The gender of the respondent has been shown to influence the purchase intention for cryptocurrency, in which more men have a higher propensity towards purchasing and using cryptocurrency. Hence, including the gender variable was done to account for differences in preference based on gender (Nandal & Jora, 2020). It is argued for various reasons that men may be more likely to invest in cryptocurrency and altcoins. For example, men may have a greater risk appetite as they invest more often than women do and thus are more likely to invest in altcoins and cryptocurrency in general, and they also are more financially literate (Xi et al, 2020). The gender of the respondent was taken as either male or female as there were no third-gender respondents or respondents who preferred not to say their gender. The options can be referred to in Appendix A.

3.2.2.2 Income of Respondent

The income of a respondent can also influence one's purchase intention for cryptocurrency. Specifically, having a higher (disposable) income means that you often can afford to have a greater risk appetite and thus newer asset classes such as cryptocurrency and altcoins within the space such as XRP become more appealing to invest in (Xi et al, 2020).

The income options for the respondent were broken down into five categories:

<€20,000
€20,000-€39,999
€40,000-€59,999
€60,000-€79,999
€80,000+

3.2.3 Constant Variables

Separate from Control variables, constant variables do not change throughout the entirety of an experiment. However, control variables can vary but are deliberately kept constant to isolate interrelations between an independent variable and the dependent variable.

3.2.3.1 The age of the respondent

To focus on determining the preferences of millennial first-time investors, in line with the research goal of this study, the survey will only take responses from respondents that are millennials, excluding anyone who is not 26-41 years old.

3.2.3.2 First-time investors

In accordance with the goals of the research, to understand what preferences are most important for first-time cryptocurrency investors, only respondents with no experience in investing in cryptocurrency will be taken into consideration in the survey.

3.2.3.3 The content/sentiment of the tweets

As the focus of the study is not to assess the importance with which the sentiment and content of a tweet influences user purchase intention, to ensure that these variables do not influence measuring preferences for twitter properties, they will be held constant (Devi et al., 2020). Hence, every tweet will have the content: “Now is a great time to invest in Ripple (XRP)!”. Refer to Figure 2 for an example of a tweet.

3.2.3.4 The user profile posting the tweets

The user posting a tweet can influence the engagement and its nature between a user and content. The username and the public image of users can greatly vary and consequently impact a user’s trust in their content (Jaakonmäki et al., 2017). Hence, keeping this variable constant by maintaining the same user for every tweet in the survey will ensure that this variable does not interfere with the accuracy of identifying the preferences for different tweet properties of the respondents of the survey. Thus, the username and user ID are held constant with the user ID being “@Crypto_123” and the username being “CryptoAdvisor”. Refer to Figure 2 for an example of a tweet.

3.2.3.5 The cryptocurrency being promoted (XRP)

A large part of this study is maintaining the focal point on an altcoin – specifically XRP. Thus, in every tweet, XRP is mentioned and is being promoted. Refer to Figure 2 for an example of a tweet.

3.2.3.6 The price of the coin

The price of the coin influences the purchase intention of investors for any financial asset (Gogol, 2021; Harcourt & Kenyon, 1976; Jora & Nandal, 2020; Nandal & Jora, 2020; Zhao & Zhang, 2021). Thus, to ensure that this variable does not interfere with gauging the preferences of the respondents it is held constant.

The incorporation of the coin price is to simulate real conditions for a potential investor who would have instant access to that information, which is motivated to play a large role in the investment decision, however as the price is not the focal point of the study it will be held constant along with the sentiment/content of the tweet (Harcourt & Kenyon, 1976). The main justification for price not being selected as an independent variable is that for that business application, firms trying to promote cryptocurrencies on Twitter will have no control over the price and thus the price is held constant as a control variable.

Each question has the text “Please use the price of Ripple and Tweets below as reference: Price of XRP at time of Tweet: $0.398085. This can be seen in Appendix B.

3.2.4 Independent Variable/Attribute Levels

The tweets will have varying levels for the independent variables that will provide sufficient variability and realism (e.g, Tweets will not have a level showing 1m likes as it is unheard of). Tweepy is an open-sources python library for accessing the Twitter API that was used to access the compiled data on 1000 of the most recent tweets that used the hashtag XRP. From the resulting dataset, the average number of likes, retweets and quoted tweets per tweet was calculated and taken as the base value. The base value was used to construct 2 more values (or levels) for each tweet property, one 50% higher and one 50% lower to ensure variability in the values across tweets. Using averages based on a sizable sample of recent and relevant tweets increases the external validity of the findings as the artificially constructed tweets used in the survey would mimic reality as closely as possible while maintaining just enough variance between each tweet to conduct the required analysis. The calculations can be found in Appendix C. Below in table 1 is the calculated values that will be used as the various levels for each attribute/independent variable.

Table 1

Attributes and Levels

Attributes	Levels
Likes	402
	803
	1205
Retweets	65
	130
	195
Quote Tweets	7
	13
	20
Blue Tick Verification	Yes
	No

3.3 Sampling

Kabadayi et al. (2013), in their paper about the impact of the properties of a car on consumer preferences and behaviours, found correlations between demographic information that was included in the initial section of the survey and the analysed factors and choices. It, therefore, provides a better understanding of the justification behind the choices made by respondents. These connections are explored in this study by asking respondents to provide the following details: gender, income, and age (only 26 - 41 years old).

When it comes to sampling size, the numbers in prior studies vary greatly. Studies that use convenience sampling rather than commercially-used or institutionalized funded techniques usually range from sample sizes between 200 and 300 (McCullough, 2002). Despite receiving a sample size of n = 500 through Surveyswap: an online paid service that provides survey responses and charges a premium for specific demographic categories, n = 292 were utilized for the data analysis. Only 292 responses were analysed after eliminating incomplete responses that did not fall into the age range of millennials and respondents that had previous investment experience. A sample size above the recommended amount also helped avoid under-representation of the desired audience. Additionally, a simplified random sampling technique was utilized in which Surveyswap distributed the survey and collected responses randomly across the millennial population (Majid, 2018).

To provide more context, Surveyswap enabled the targeting of millennial respondents for the survey at a financial cost which they were delivered quickly. However, there was no guarantee that these millennial respondents were first-time investors as the platform did not have functionality for segregating this way. Hence, 208 responses were unusable.

3.4 Designing a Conjoint Experiment

A Choice-based Conjoint (CBC) Analysis will be employed to determine what properties of a Tweet are important to users and to what extent those properties influence their purchase intention/propensity for the altcoin XRP, mentioned in the tweet. This method determines how much value and importance investors place on different attributes which helps identify the features that provide them with the highest utility and therefore have a significant effect on their purchase intentions and decisions. This technique is implemented by presenting profiles/choices containing different combinations of attributes that the respondents are required to either rank (ranking-based conjoint analysis), rate (rating-based conjoint analysis), or pick one of the choice profiles that are most appealing to them (choice-based conjoint analysis) (Hauber, et al., 2016). Through gathering data of preferences based on choices made by respondents, preference scores/utilities are generated and in this case are automatically passed as input for a logistic regression model, which is the main data science model of this paper. More about this method and its application is discussed in Section 3.6.

The use of CBC is highly relevant as it has been commonly utilized to help understand how and/or why consumers engage with or react to social media (Eggers, F. et al, 2022). The main idea behind the CBC analysis is that tweets with varying tweet property (independent variable) values will be shown to respondents who will in turn select which tweet motivates them the most to invest in XRP. This analysis and data collection framework suit the research needs well to determine how important different tweet properties are for millennial first-time investors.

When designing conjoint experiments, several critical steps are commonly followed or applied in previous literature, as mentioned below (Green and Srinivasan 1978; Green & Srinivasan 1990; Hauser & Rao 2004).

1. Decomposition of product on product attributes (In this case a Tweet on Tweet properties/attributes)
2. Stimuli representation
3. Stimuli configuration

4. Data collection
5. Estimation

The first step in designing a CBC takes into account defining the product attributes and the levels that each one will be tested at. It also defines the function of each attribute and its levels to create a model of preferences for the respondents' to choose from. These functions are preferences that can be vector monotone (approximately linear), ideal point (convex or concave) and part-worth (discrete or categorical). In this case, the attributes are categorical variables that

The second step describes the way that the stimuli in the survey will be presented to the target audience. In this case, artificially created digital tweets will be displayed, along with an introduction and description of instructions. To increase external validity, the stimuli representation was created to look as realistic and identical to a tweet as possible. The third step in this section relates to the existence of an extensive amount of product profiles that are not feasible to be shown to the respondents. To neutralize potential bias, product profiles are reduced using an orthogonal design approach. Its use and application will be explained in more detail in the coming sections of this chapter. The fourth and fifth steps go hand-in-hand since the way the data is collected needs to be exported and fit the model estimation for its format and design. For this study, the choice-based conjoint analysis will be used to collect the data and this data will be further analysed using logistic regression. The following sections will go into depth regarding its relevance, application and interpretation.

Within the context of this research, choice-based conjoint analysis (CBC) is performed and the analysis is done utilizing Conjointly, Qualtrics and Sawtooth software. CBC is chosen because this type of conjoint analysis, as compared to rating or ranking-based analysis, is known to be more "representative of real-life decision-making" because respondents are asked to make trade-offs and make a choice between options, which is the most realistic way of gauging purchase intention and preference (Louviere & Islam, 2008). It is therefore constructed to add realism to the survey design, thus making the results more externally valid. (Toubia et al., 2004).

3.4.1 Survey Design

The survey includes the following sections:

1. A cover page: Includes a brief description of the goal of the study (to evaluate the preferences of millennial first-time investors on different tweet properties on their likelihood and intention to invest in the cryptocurrency Ripple (XRP))
2. Instructions: An explanation of how the survey will be structured and what is expected of the respondents'
3. Main section (All questions are marked as required)
    a. A few demographic questions in multiple-choice format
    b. A Choice-based Conjoint (CBC) exercise with 12 trade-off tasks, where each task consists of three artificial profiles/tweets with different combinations of attribute levels from which the respondent has to pick one that motivates them to purchase XRP the most

Rather than a full factorial survey design, an orthogonal array design was used to help lower the number of choices or tweets shown to the respondents (Huber et al., 2003). It is not necessary to use data collected from all possible combinations since it results in lower robustness, and reliability (if the survey was filled out again, results would not be consistent) while increasing respondents' burden. Using an orthogonal array thereby helps reduce respondents' fatigue due to the less-time consuming nature of the survey while allowing the main effects of the attributes to be measured. In this design, the extreme combinations of the tweets' features are examined. In other words, "the levels of the features are chosen such that, for each pair of features, say a and b, the high level appears equally often in profiles that have a high-level b as in profiles that have a low level of b, and vice versa" (Hauser 2007). According to Rao (2013), the orthogonal design also results in a prime prediction of the preferred attributes despite some unrealistic situations being presented.

This technique has been investigated and implemented by numerous studies. They have a common viewpoint of it being a "robust design in the use of metric analysis" (Dahan & Hauser, 2002; Evgeniou et al., 2005). Egeniou et al. (2005), following the steps of Green & Srinivasan (1978) and Nickerson et al. (1990), also state that an individual should not be presented with more than 30 choice tasks in a single survey to avoid information overload by using an orthogonal factorial design to ensure accurate value estimation. In their paper, an extreme set of 1536 (4*3*27) product profiles was reduced to only 16 by applying an orthogonal design to be able to estimate the main effects. However, such simplifications and reductions do come with a trade-off. In this case, in accordance with the preference independence assumption, estimation of the interactions among the different attributes is disallowed (Keeney, 1971).

Therefore, to summarize the design of the choice experiment using a questionnaire, the function on Conjointly, a choice-based conjoint modelling software, takes the input values of the number of attributes and their respective levels and runs a full factorial design. In accordance with the orthogonal design, it decreases the number of alternatives and generates the optimal, balanced fractional factorial design. A full factorial design for this study would yield 54 combinations (3*3*3*2) of attributes and after the orthogonal array tool is applied, then creates 12 choice tasks with a random combination of the 12 conjoint profiles created through the orthogonal design, presented to the respondent: each containing three profiles/tweets with varying attributes. As seen below, Table 2 displays the result generated by Conjointly after using an orthogonal design and the number is reduced to 12 combinations for each respondent.

Table 2

Orthogonal Design Results

Conjoint Profiles	Likes	Retweets	Quote Tweets	Blue Tick Verification
1	803	65	13	No
2	1205	195	13	Yes
3	402	130	20	Yes
4	1205	195	7	No
5	402	65	13	Yes
6	803	195	20	Yes
7	402	65	7	No
8	1205	195	20	No
9	803	130	13	Yes
10	402	130	7	No
11	1205	195	20	No
12	803	65	7	Yes

Figure 2 below is an example of an artificially created tweet that will be displayed next to two others (three tweets in total) for each task/choice set that the respondent has to pick from. The profiles from Table 2 above will be randomly combined into 12 distinct tasks. It is important to note that when 1205 is inputted as the number of likes, it automatically changes the displayed number to "1.2K" which mimics the way Twitter, in reality, would show the likes that cross the 1000 mark. The full survey can be found in Appendix B.

Figure 2

Example Tweet of CBC Profile 1

3.5 Testing for Bias

Roberts & Priest (2006) defined validity as the "closeness of what is being measured to what the research intends or aims to measure". It is important to clearly relate the theoretical framework and its learnings to the survey constructs while making sure that each independent variable is expressed accurately in the survey. The chosen attributes of tweets and their respective levels that will be used in the survey are therefore backed up by previous literature that has deployed similar methods.

Bias can be defined as "any tendency which prevents unprejudiced consideration of a question" (Pannucci & Wilkins, 2010). It helps eliminate any systematic error that exists in a sample or a testing method that hinders the results and consequently interpretation of scientific research. Some level of bias is said to be nearly always present in a study so it is imperative to look into how it might influence the research outcome and how it can be neutralized beforehand (Roberts & Priest, 2006). Multiple biases need to be highlighted and considered. First and foremost, omitted-variable bias is introduced when there is a discrepancy between an estimator's expected value and the true value that is a result of leaving out one or more relevant variables from the analysis (Jargowsky, 2005). To prevent or eliminate this bias, this research will add relevant control variables so no external varying factor reduces the validity of the data as mentioned in the previous section. Secondly, selection bias is when the sample used is not an accurate representation of the target respondent group. Since this survey will only be answered by millennials and first-time investors, the age will be asked in the initial demographic section of the survey and the ones that do not conform to the target will be eliminated from the sample.

In addition, using a choice-based conjoint model increases the reliability of the data by providing a holistic view of the tweets being shown by analysing the different attributes altogether in one tweet in a simulated format rather than questioning the consumers about the product attributes separately.

3.6 Logistic Regression

To gain further insight into how the tweet properties influence their choice of which tweet motivates them most to invest/purchase XRP, a logit model will be implemented on the results of the CBC data. The independent variables - the tweet properties to be focused on in this paper - will be evaluated in the logit model to understand which are the main contributors to their choice.

The logit model has been utilized for over three decades as the standard for the analysis of CBC data due to its ability to estimate choices accurately (Sawtooth Software, 2022). The logit model allows for the calculation of utilities and translates this into choice probabilities for different values/levels of tweet properties in tweets and can identify how the tweet properties would influence purchase intention for XRP of millennial first-time investors. It is a powerful and useful diagnostic data science tool that aids in assessing the quality of the experimental design as well as estimating preferences for the sample (Sawtooth Software, 2022).

Logistic regression is built on the foundation of the estimation of maximum likelihood in predicting the probabilities of the two classes of a binary dependent variable *Y*. The dependent variable is nominal, signifying that there is no intrinsic ordering in the different classes. In this paper, purchase intention is the dependent variable which takes a value of 1 indicating a respondent will invest in XRP and 0 otherwise (Warner, P, 2008). The preferences for different levels per attribute are estimated by calculating the utilities of each attribute level. The utility of an attribute is the sum of all its levels. The logit model assumes that the decision maker chooses the alternative with the highest utility. The probability that the purchase intention takes value 1 is given by the following formula:

$$\Pr(\text{Purchase Intention} = 1 \mid x) = \frac{exp(B_0+B_1x_1+\ldots+B_px_p)}{1 + exp(B_0+B_1x_1+\ldots+B_px_p)}$$

To illustrate how Logistic Regression works, it is imperative to examine the odds formula for the model below. Establishing notation, let *Y* be the outcome of the binary dependent. The interpretation of the weights in logistic regression differs from linear regression as the outcome for logistic regression is a probability between 0 and 1. As the weights do not influence the probability linearly the formula above is reformulated by taking the logarithm.

$$\log\left(\frac{Pr(Purchase\ Intention = 1 \mid x)}{Pr(Purchase\ Intention = 0 \mid x)}\right) = B_0 + B_1x_1 + \ldots + B_px_p$$

The formula within the log term is known as the odds ratio which becomes the log odds ratio when wrapped in the logarithm function. This formula shows that logistic regression is a linear model for the log odds value. Using this, we can compare what happens when a feature value increases by one unit. This is done by looking at the ratio of two log odds predictions:

$$\frac{odds_{x_j+1}}{odds} = \frac{B_0+B_1x_1+.B_j(x_j+1) + \ldots+B_px_p}{B_0+B_1x_1+B_jx_j + \ldots+B_px_p}$$

Simplifying and removing terms results in the following output:

$$\frac{odds_{x_j+1}}{odds} = exp(B_j(x_j + 1) - B_jx_j) = exp(B_j)$$

A change in a feature x_j by one unit changes the odds ratio by a factor $exp(B_j)$. This can also be interpreted as a one-unit change in x_j increases the log-odds ratio by B_j. The parameter estimates for B are obtained using maximum likelihood estimation. The logarithm of the maximum likelihood estimation is taken which is then maximized.

There are several advantages that Logit models possess over other models which makes it preferable in many situations. Compared to popular machine learning models, it not only provides values for feature importance but also indicates the direction of the effect (positive or negative). Additionally, it provides a good accuracy score in cases where the dataset is linearly separable. A dataset is said to be linearly separable if different clusters can be separated by a linear hyperplane.

One of the disadvantages of the model, however, is that it assumes there is no multicollinearity between the independent variables. Multicollinearity is a common occurrence which poses problems in multiple regression models. It is a situation in which an independent variable is highly correlated to another. Multicollinearity reduces the performance of the model as it reduces the statistical significance of the independent variables.

To deal with Multicollinearity, Variance Inflation Factor (VIF) values will be incorporated into the model to identify the extent of multicollinearity present among the variables. Mathematically, VIF for a regression model variable equates to the ratio of the model variance to the model variance with the inclusion of solely a single independent variable. This ratio value is computed for each of the independent variables. In other words, this is calculated by taking an independent variable and regressing said variable against the other independent variable predictors. This provides R^2 values that can be plugged into the VIF formula. ‘i’ is the independent variable being looked at (e.g x_1 or x_2) the formula for this calculation can be seen below (Glenn, 2020):

$$VIF = \frac{1}{1 - R_i^2}$$

The specific threshold for how large a VIF value can be before the multicollinearity causes significant issues is a subject of debate, however, there is a general rule of thumb for interpretation (Dodge, 2008; Everitt 2010):

- 1 = Uncorrelated
- Values between 1 and 5 = Moderate Correlation
- Greater than 5 = Strong Correlation

Although, what is known is that the higher VIF, the less reliable the output of a regression is going to be. For this study, any VIF value greater than 5 that denotes a strong correlation between predictor variables will be a sufficient cause for concern and will be removed from the model. Along with the application of the logistic regression model, using Sawtooth Software, the incorporation of VIF values was also computed.

3.6.1 Logit Application

While the CBC profiles were constructed on Conjointly, created in a survey format and distributed on Qualtrics, it was analysed with a logit model utilizing Sawtooth Software. The logistic regression model that is run provides both the log odds ratio and odds ratio to determine which independent variables/attributes of the tweets were most influential in motivating the respondents’ choice of which tweet made them want to invest in XRP the most.

Using CBC in a survey format to gather data and then applying a logit model to extract insights, the logit model will be able to predict whether a potential first-time investor is more likely to invest in XRP after interacting with a tweet and to what extent each property of a tweet influences investment decision (Zhang, J., et al, 2021; Harrell, F. E., 2015). The insights from this study aim to provide entities with the information to optimize their Twitter campaigns (e.g. use and allocate resources towards partnerships accounts that are blue tick verified when promoting crypto on Twitter) and use of tweets to influence the purchase intention of large and/or emerging cryptocurrencies like XRP, based on the importance of various factors.

To estimate the preferences for the varying levels of each attribute, the respective utility values for all the levels are computed. The utility of an attribute is fundamentally the summation of all of its associated levels. Put forth in the study by Walker and Ben-Akiva, 2002: "utilities are latent variables and are assumed to be a function of a set of explanatory variables X".

As the decision maker/respondent will select the choice alternative with the highest utility and greatest likelihood to purchase:

$$\text{Pr(Purchase Intention = 1 | Likes, Retweets, Quote\_Tweets, Blue\_Tick\_Verification)}$$
$$= F(B_0 + B_1 Likes + B_2 Retweets + B_3 Quote\_Tweets + B_4 Blue\_Tick\_Verification)$$

Thus, The formula for the logistic regression in this study - denoting the probability of selecting choice alternative i - is:

$$\text{Pr(Purchase Intention = 1} \mid Likes, Retweets, Quote\_Tweets, Blue\_Tick\_Verification) = \frac{exp(B_0+B_1Likes+B_2Retweets+B_3Quote\_Tweets+B_4Blue\_Tick\_Verification)}{1 + exp(B_0+B_1Likes+B_2Retweets+B_3Quote\_Tweets+B_4Blue\_Tick\_Verification)}$$

3.6.3 Prediction and Precision

While the main purpose of the model is for explanation, further application of the model can be relevant in terms of its predictive capability and model validity, in which it may be able to predict, to a certain degree, whether a respondent will select a specific tweet.

To test the choice-model-based logit output validity, a prediction-realization (hit-rate) table will be constructed with values computed that reflect, for each observed choice in the data, whether the predicted choice from the logit model was correct. The proportion of the predicted choices that are correct, known as hits (same as the respective observed choice), of the incorrectly predicted choices and the correctly predicted choices is the calculation for the hit-rate or precision metric (Orme, 2002; Chrzan & Sawtooth Software, Inc., 2015).

There will be In-sample predictions and Out-of-sample predictions. In-sample prediction refers to having a hit-rate value for observed choices within the training sample utilized to run the model in which a hit would be if the model predicted the correct tweet being chosen for a respondent. While an out-of-sample hit-rate table will also be calculated based on a training set that is set aside for this specific case. A random subset of 70% of the data (n=205) was utilized for the Logit model and the other 30% (n=87) of the total dataset (n=292) was purposely left out as a test set. The out-of-sample set is calculated based on the model's estimation of the test set.

This will be used to create a confusion matrix that will assess the accuracy of the model. How this works is that the confusion matrix for each In-sample and Out-of-Sample hit-rate table will provide precision values for each respective sample. This will be done by calculating the precision of the data provided by the confusion matrix. The confusion matrix will provide only True positive (TP) and False Positive (FP) proportions based on the fitting of the model to each sample and from there the precision/hit rate can be calculated for each sample (Lee et al, 2018; Chrzan & Sawtooth Software, Inc., 2015):

$$Precision = (TP)/(TP + FP)$$

It is important to denote that true and false negatives, which are usually included in a confusion matrix will not be utilized here as the way that the hit rate is calculated is based on whether the model predicts that a respondent has chosen the tweet that is in-line with the prediction based on the conjoint task. The way this is calculated is that the model only selects the tweet with the highest probability of being chosen based on the likelihood of the tweet being chosen, calculated by the marginal probabilities of each attribute measured (e.g. likes, retweets, etc.) hence the values in the hit rate table will only either be true positives that are correctly predicted or false positives where the tweet to be chosen is incorrectly predicted based on the conjoint task at hand. This is due to the model inherently assuming that a respondent will always choose to purchase XRP from the tweets being presented.

Precision establishes how close the measurements are to one another across the in-sample and out-of-sample predictions as each measurement in a series holds a portion of random error. This causes measurements to vary to a degree even while measuring the same item. Thus, precision relates to reproducibility and how repeatable the data is when measuring choice over different samples. Higher precision means measurements being closer together and hence greater reproducibility.

Results

In total there were 500 responses collected for the survey. However, out of the 500 respondents, only 292 were usable for the analysis, hence excluding 208 respondents based on them not being first-time investors. Utilizing the service Surveyswap: a platform where you submit surveys that are distributed to a tailored audience, the survey was able to target millennial investors with a fee charged by Surveyswap, of which only the 292 respondents claimed they were first-time investors in cryptocurrency. The platform did not have the functionality to provide millennials that are first-time investors without a sizeable premium. Out of the 292 respondents, 156 were female and 136 were male with proportions similar across both genders. In terms of income, (rounded to the nearest hundredth) 14.63% of respondents made an income below €20,000, 22.93% had an income between €20,000-€39,999, 26.83% had an income between €40,000-€59,999, 17.56% had an income Between €60,000-€79,999, and 18.05% had an income greater than €80,000.

As every respondent had 12 conjoint tasks in which the respondent needed to choose between three conjoint profiles per task, the dataset contained 292*36=8064 conjoint profile choice observations. It is important to note that the train set is composed of 205 responses, of which 106 are male and 99 are female and the test set is composed of 87 responses, of which 30 are male and 57 are female.

4.1 Models and Coefficients

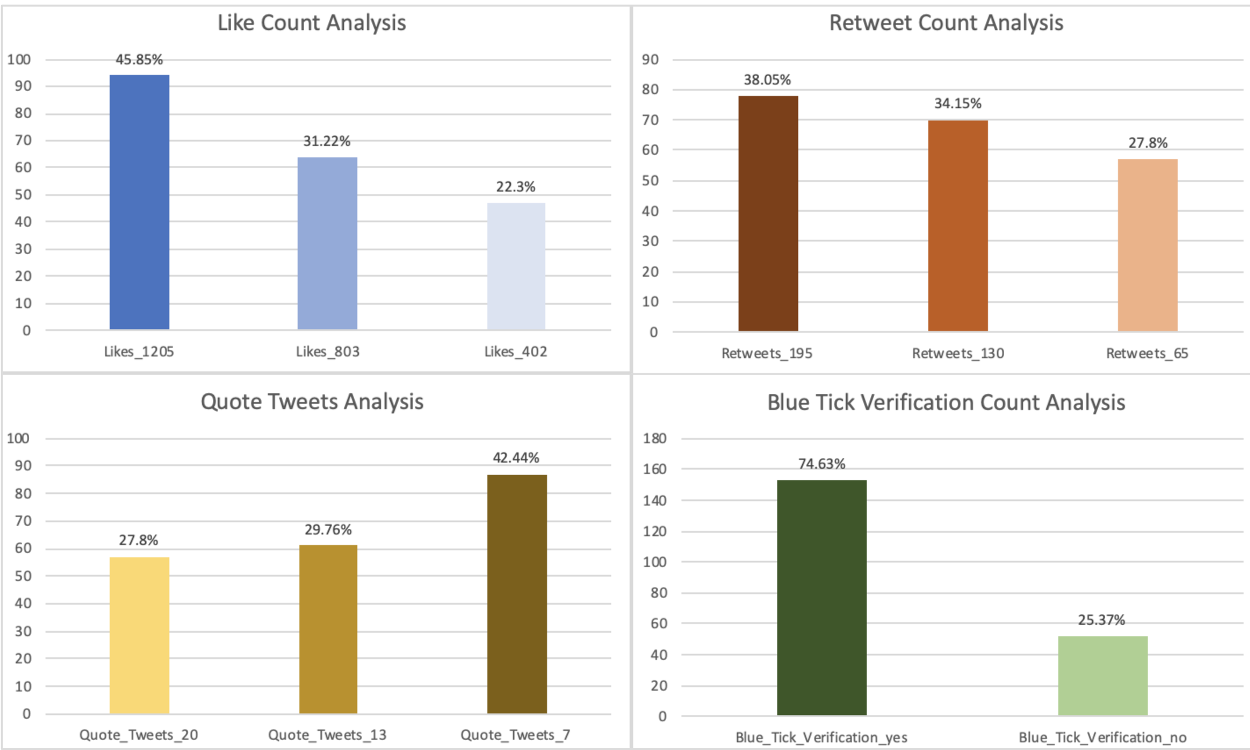
There are four independent variables in the output of the conjoint analysis results. Three variables were taken as categorical (Likes, Retweets, Quote Tweets) with the attribute levels for each resembling a high, medium, or low breakdown for their values, which is explained to be a suitable means for analysis instead of incorporating these variables into continuous data (Roh et al, 2007). It is argued that by keeping them categorical, if done correctly, the indifference area between levels (e.g., respondents may have no marginal utility gain from interacting with a tweet with 50 retweets versus 55 retweets) will be accounted for to more effectively measure sensitivity to the three variables for purchase intention of XRP. The last independent variable: Blue Tick verification is a binary categorical variable.

4.2 Count Analysis

As specified by Orme, B. (2002), count analysis is an interesting method to quickly and efficiently explore the data. In this specific analysis, the amount of times that particular attributes were chosen by the respondent, in their choice selection during the conjoint analysis, is counted. The count analysis is shown In Figure 2 below.

Figure 3

Count Analysis Graphs



The count analysis provides interesting insights. The Likes_1205 within the Likes attribute was shown to be the significantly most popular Likes value in the choice experiment. This seemed to be the case for a good majority of the conjoint profiles/tweets shown making up 45.85% of all selected tweets, almost half.

For the attribute Retweets, there seems to be a propensity for respondents to select the highest amount of retweets (Retweets_195) among the three levels for the attributes. However, the difference between the proportion of respondents who selected Retweets_195 and Retweets_130 is only 3.9% - a small marginal difference. However, it can be seen that most respondents would be more motivated to purchase XRP if the retweet values were at least 130.

For the attribute Quote_Tweets, similar to Likes, there is one level that was selected the most, Quote_Tweets_7. It is interesting to note that this is the only attribute that reflects a negative correlation between the proportion of respondents who choose Tweets with the respective values and the size of the values, as the lowest value, 7, was the most commonly selected.

Lastly, for the attribute Blue_Tick_Verification, the largest difference in the proportion of respondents can be seen between Blue_Tick_Verification_yes and Blue_Tick_Verification_no. Approximately, 3 in 4 respondents selected a Tweet with the user being blue tick verified.

4.3 Logistic Regression Output

The reference category levels for the logistic regression will take the values of the middle category of values for Likes, Retweets and Quote Tweets (the values being Likes_803, Retweets_130, and Quote_Tweets_13 respectively) as a means of gauging sensitivity to the variables from the middle values (Johfre et al., 2021). Furthermore, the Blue Tick Verification variable reference category value was taken as ‘no’.

Below, Table 3 provides information about the direction of the relationship between the variables and the choice/purchase intention as well as its significance. As can be seen in Table 3, all variables incorporated in the logit model above are significant at least to a 5% significance level and thus all variables have an explanatory influence on choice. Therefore, all variables are interpreted. It is important to note that the Akaike Information Criterion (AIC) is utilized to compare the goodness of fit of a model between different models, as a relative measure. When the AIC is lower, the model has a better fit. However, as there is no model comparison, this measure is informative and irrelevant.

Table 3

Logistic Regression Output

Variable	Estimate	Std. Error	p-value	VIF
Intercept	-9.465	0.3143	<2*10^-16 ***	
Likes_1205	0.07428	0.02579	2*10^-16 ***	1.98
Likes_402	-0.09093	0.01670	8.46*10^-11 ***	2.42
Retweets_195	0.04560	0.03596	4.13*10^11 ***	1.45
Retweets_65	-0.09179	0.02611	7.08*10^-7 ***	3.21
Quote_Tweets_20	-0.00968	0.02498	4.13*10^11 ***	1.64
Quote_Tweets_7	0.03046	0.03617	7.08*10^-7 ***	1.96
Blue_Tick_Verification_Yes	0.07175	0.02191	2*10^-16 ***	3.91
Observations		205		
Log Likelihood		-1312.155		
Akaike Information Criterion		761.480		

p 0.1, * p 0.05, ** p 0.01, *** p 0.001

Exploring the Odds Ratio is important to provide a clearer way to assess the attribute importance of a hypothetical tweet being interacted with to gauge purchase intention. To calculate the Odds Ratio from the Log-Odds in the logistic regression, the estimated coefficients were used as such:

$$e^{\beta j} = Odds\ Ratio\ Value\ for\ Variable$$

The values for the Odds Ratio provide insight into which an attribute level change can make the likelihood of a user purchasing XRP x times more likely to do so. It is important to note that an odds ratio >1 shows a positive association between purchase intention and the attribute and the same applies to vice versa.

To calculate the effect of Odds based on the values below:
(-1 + Odds Ratio Value)*100=Δ% in the odds of purchase intention of XRP after interacting with a tweet with a specific attribute.

Table 4

Odds Ratio

Variable	Odds Ratio
Likes_1205	1.077
Likes_402	0.9131
Retweets_195	1.047
Retweets_65	0.9123
Quote_Tweets_20	1.031
Quote_Tweets_7	0.9904
Blue_Tick_Verification_Yes	1.074

4.4 Assessing Hypotheses

This section deals with the assessment of all four of the hypotheses as aforementioned in Section 4. Each hypothesis has been assessed separately. Referring to Table 3, showing the results of the logistic regression model on the CBC analysis results, in line with the formula and explanation for the model mentioned in Section 3.6, the coefficients of the logit model will be interpreted with a 10% significance level along with the Odds Ratio Table 4, that enables a deeper interpretation. The variables are all the properties/attributes of a Tweet and how they influence the choice of which Tweet made them most likely to purchase XRP. Additionally, it is important to note that the VIF values in the output in Table 3 above represent the extent to which each independent variable attribute is correlated to all the others. Across all variables, these respective values for each independent variable satisfies that there is a lack of multicollinearity in the data as all VIF values < 5 and ensures that the logit model will not succumb to its sensitivity to multicollinearity.

H_1: The amount of likes associated with a Tweet will positively influence the purchase intention of XRP.

Here, the variables of interest are the Like_1205 and Like_402 variables that denote the influence of these different categorical variables on purchase intention. With the reference category being selected for the model as Likes_803, the variables dictate how purchase intention for XRP is different when the like values vary from Likes_803 for tweets. Referring to the Odds Ratio table 4, it is clear that if Likes decrease to Likes_402 there is an estimated

8.7% decrease in the odds of purchase intention for XRP after interacting with a tweet. Lastly, an increase for the attribute Likes, from Likes_803 (reference category) to Likes_1205 leads to a 7.7% increase in the odds of purchase intention for XRP after interacting with a tweet. Thus, the hypothesis is confirmed as the likes of a tweet have a positive influence on the purchase intention for XRP.

H_2: The amount of retweets associated with a Tweet will positively influence the purchase intention of XRP.

Here, the variables of interest are the Retweets_195 and Retweets_65 variables that denote the influence of these different categorical variables on purchase intention. With the reference category being selected for the model as Retweets_130, the variables dictate how purchase intention for XRP is different when the like values vary from Retweets_130 for tweets. Referring to the Odds Ratio table 4, it is clear that if Retweets decrease to Retweets_65 there is an approximate 8.8% decrease in the odds of purchase intention for XRP after interacting with a tweet. Lastly, an increase from Retweets_130 (reference category) to Retweets_195 leads to a 4.7% increase in the odds of purchase intention for XRP after interacting with a tweet. Thus, the hypothesis is confirmed as the number of Retweets has a positive influence on the purchase intention for XRP.

H_3: The amount of Quote Tweets associated with a Tweet will negatively influence the purchase intention of XRP.

Here, the variables of interest are the Quote_Tweets_7 and Quote_Tweets_20 variables that denote the influence of these different categorical variables on purchase intention. With the reference category being selected for the model as Quote_Tweets_13, the variables dictate how purchase intention for XRP is different when the like values vary from Quote_Tweets_13 for tweets. Referring to the Odds Ratio table 4, it is clear that if quote tweets decrease to Quote_Tweets_7 there is a 3.1% increase in the odds of purchase intention for XRP after interacting with a tweet. Lastly, an increase from Quote_Tweets_13 (reference category) to Quote_Tweets_20 leads to a 0.96% decrease in the odds of purchase intention for XRP after interacting with a tweet. Thus, the hypothesis is confirmed, although reflecting very small increments, as the number of quote tweets has a negative influence on the purchase intention for XRP.

H_4: The user of a tweet being blue-ticked verified negatively influences the purchase intention of XRP.

Here, the binary variable of interest is the Blue_Tick_Verification_Yes which denotes the influence of a verified profile on the purchase intention of XRP. With the reference category being selected for the model as Blue_Tick_Verification_No, the variable in question dictates how purchase intention for XRP differs when there is a presence and absence of a verified blue tick on the user profile. Referring to the Odds Ratio table 4, it is clear that if the user is verified, there is a 7.4% increase in the odds of purchase intention for XRP after interacting with a tweet. Thus, the hypothesis cannot be confirmed as the user of a tweet being blue-ticked verified has a positive influence on the purchase intention for XRP.

4.5 Hit-Rate Tables

Below are the hit-rate tables for both in-sample and out-of-sample predictions, obtained by crossing the number of predictions for choice estimations against the observed/actual choices. As estimates take values of probabilities, the probability value that is greatest for each choice task is predicted as the tweet that is chosen for a respondent.

For example, there are three tweets for a conjoint task that a respondent in the test set has chosen. The model takes into account the preferences and estimated probability derived from running the model on the training set and calculates a probability for the respondent selecting each of the three tweets. If the tweet with the highest probability of being chosen based on the logit model's predicted parameters is the same as the tweet that was chosen by the respondent in the test set, then it counts as a true positive as it was correctly predicted to be a choice.

4.5.1 In-Sample

Table 5

In-Sample Prediction Hit-Rate Table

Predicted	Actual	Result
1	1	0.748
1	0	0.252

Precision=0.748

The in-sample precision value is 0.748 and denotes how often the model is correct in predicting the correct choice of a respondent. Hence, approximately 3 in 4 choices are correctly predicted, which reflects a moderately high precision.

4.5.2 Out-of-Sample

To reiterate Section 3.6.2, the Out-of-Sample predictions will be based on the trained model, which was trained on 70% of the data (n=205) and its fitting on the other 30% of the data that was kept aside as a test set (n=87).

Table 6

Out-of-Sample Prediction Hit-Rate Table

Predicted	Actual	Result
1	1	0.554
1	0	0.446

Precision = 0.554

There seems to be overfitting of the model on the training data as the precision value for the in-sample hit-rate calculation is approximately 0.194 greater than that of the out-of-sample hit-rate precision value of 0.554. The out-of-sample precision reflects a moderate capability of the model to correctly predict when a respondent chooses a specific tweet with the model predicting approximately 55% of all choices correct, and thus reflects moderate reliability in the model correctly classifying a choice.

4. Conclusion and Discussion

5.1 Discussion

The central research question of this thesis is: “How do the properties and social media engagement of a Twitter ‘Tweet’ influence the purchase intention of first-time millennial investors of the Ripple (XRP) coin?”. The findings of this research dictate that the amount of likes (H_1) and retweets (H_2) have a positive influence on the purchase intention of first-time millennial investors of the Ripple (XRP) coin. These results were in line with the literature used in Chapter 2 of this study. Also in line with the literature, a negative influence was found between quote tweets (H_3) and purchase intention. Finally, a positive influence on purchase intention was also found when the profile supporting the XRP had a blue tick verification (H_4), which was not in line with the literature used to develop this hypothesis.

5.2 Implications

This paper makes significant contributions in the areas of social media marketing, influencer marketing and consumer decision-making behaviour on Twitter. The relevance of this platform stems from the rapid rise of active Twitter usage around the world - it is estimated to have almost a total of 80 million users. This can be used by brand managers to set benchmarks and goals within budget for marketing their altcoins. For example, by understanding that having a lower amount of retweets has the greatest dissuasive power for a respondent to invest in XRP after seeing a Tweet (based on the lowest Odds Ratio value of 0.9123 in Table 4), brand managers can ensure that they allocate sufficient resources to increasing retweets for a tweet they are focused on promoting. Alternatively, with the information from this study, the brands can work to prevent quote tweets to increase the likelihood of an investor purchasing/investing in their altcoin cryptocurrency. A lot of Twitter brands and users employ Twitter bots to manipulate their Twitter metrics in which they can set specific values for their tweets thus understanding the ideal thresholds and values for tweets is insightful as these brand managers can essentially decide on the number of retweets, likes, etc. utilizing Twitter bots (fake accounts) (Sayegh, 2022). They can thereby set thresholds for response rates, and deploy better data analytics and growth marketing strategies via Twitter that can cause higher results and better performance.

Since this study also targets first-time inventors, it helps understand how to ease them into the realm of cryptocurrency and consequently grow their millennial customer base. As the findings of this study pertain solely to this cohort, it is ideal for brands to understand exactly what preferences users and potential investors have and how to manipulate tweet properties to tailor to them.

Additionally, the use of the altcoin XRP is a good example as a benchmark for altcoins as the way the study was conducted ensured minimal explanation of XRP. This makes for a realistic depiction of the tweet for how a user would interact with it, irrespective of whether the user would know about the altcoin or not. This reflects the implications that the results of this study would apply to other brands that would leverage Twitter to promote their altcoins.

Furthermore, the research contributes to the emerging stream of literature that investigates the effect of social media properties and popularity/following metrics that can influence the advertising and persuasion power of users for fin-tech, also a growing and innovative field, based information. Since there is a lack of research that amalgamates Twitter and cryptocurrency, this paper helps bring to light the importance of this social media platform and how its properties have shown quantifiable effects. A choice-based conjoint has not been applied in this context before and therefore provides a novel perspective into the factors that influence the incentives to invest.

5.3 Limitations and Recommendations for Future Research

There are several limitations of this research study that are important to account for and rectify in future research. Firstly, to make the choice-based design easier and straightforward for the respondents, the alternatives could be lowered to two choices per set instead of three. This could take up less decision-making time and decrease fatigue. This also ties in with adding an option to the experimental design and following code. This makes the task more realistic and helps provide more accurate or truthful data since some of the respondents might, in reality, not be incentivized or want to invest in any of the options presented against each other. This creates an unrealistic assumption that each respondent’s preferences can be expressed or illustrated by the same combination of choice sets (Shan et al, 2017).

The issue with replication of real-life purchasing environments or in other words, the hypothetical nature of the alternatives deployed in the choice sets is a general concern brought up by numerous practitioners when using conjoint methodologies and argue that respondents are not motivated while answering the survey due to the discrepancy (Toubia et al., 2012; Ding et al., 2005). This is thereby a limitation since external validity is higher in an experimental context that better imitates reality. For future studies, this can be rectified by adding incentives within the survey to reveal their true preferences. According to Carson and Groves (2007), respondents have a higher motivation about being conscientious about their answers in hypothetical settings if they think that their answers that if their answers "have an impact on decisions made by businesses or governments for outcomes that the agents care about, they will respond in such a way as to maximize their payoffs and welfare". Therefore, having an incentive-based survey design that evokes a sense of importance in the respondents, makes them believe that their answers have an impact or that they will be rewarded, and external reliability can be increased.

Another limitation pertains to the lack of variation in the attribute levels. While, the data scraped from tweets mentioning XRP provided us average values for the attributes: Likes, Retweets and Quote Tweets, these values were quite small. Thus, with the proper investment from firms promoting altcoins, there could be a significantly greater (or lesser) influence on first-time millennial investors and their purchase intention for the respective altcoin being mentioned after interacting with a tweet if these attribute values were significantly higher.

Furthermore, the execution of conjoint analysis with the obtained survey results was difficult and time-consuming due to the lack of access and financial resources leading to limited access to software that can perform conjoint analysis behind paywalls since these are mainly catered to companies (Qualtrics Conjoint Capabilities, Sawtooth Software, etc.). The crux of this limitation stems mainly from the limited access to Sawtooth Software, a paid software provided specifically to companies. Unfortunately, the Erasmus University Rotterdam license did not encompass Qualtrics, hence I has to utilize Qualtrics to build and distribute the survey, Conjointly to build the conjoint profiles and Sawtooth Software to analyse the data. The license for Sawtooth Software was provided with limited access which posed issues in terms of adding specific nuances to the research (such as the "None" Option).

In future research, additional variables could also be added to test for the influence of attributes such as follower count, the addition of hashtags, emojis, etc. Moreover, demographic data such as time spent on Twitter or a measure for a history of gambling could also be included in future studies to test for whether more avid users of Twitter have different perspectives on what attributes influence them to invest. The same applies to income level - as a first-time investor, a higher income level could indicate less hesitation to invest and could thereby have an effect on the results. In that case, it could also be added as a controlled variable to see the comparison.

A good way to improve the model in terms of its predictive capabilities and its application to other samples would be to increase the training sample size to account for greater variability and would thus make the model more robust to different samples as a result. Furthermore, with a larger dataset, the utilization of cross-validation would be beneficial, providing a better understanding of how the model would fit on different subsets of data. However, in this case, due to the aforementioned smaller sample size, this was unfeasible.

Lastly, incorporating demographic variables into the analysis as a controlled variable would also be extremely insightful, however, when referring to other papers, it was the case where demographic variables were either left out or held constant rather than controlled and included in the model as the CBC logit model does not incorporate these factors (Soutar et al, 2002; Larsen et al, 2021; Jelena et al, 2019; G.B. Dimitrov, 2017; Booij, 2021). A solution for further research could be to implement clustering using a Latent Class Clustering analysis for the respondents and then post-doc assigning demographic values to see how different demographic categories may influence preferences within first-time millennial cryptocurrency investors.

References

Ante, L. (2021). The non-fungible token (NFT) market and its relationship with Bitcoin and Ethereum. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3861106

Armknecht, F., Karame, G. O., Mandal, A., Youssef, F., & Zenner, E. (2015). Ripple: Overview and Outlook. Trust and Trustworthy Computing, 163–180. https://doi.org/10.1007/978-3-319-22846-4_10

Aysan, A. F., Khan, A. U. I., & Topuz, H. (2021). Bitcoin and Altcoins Price Dependency: Resilience and Portfolio Allocation in COVID-19 Outbreak. Risks, 9(4), 74. https://doi.org/10.3390/risks9040074

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, 114, 24-31.

Bermel, N. and Knittl, L. (2012). Corpus frequency and acceptability judgments: A study of morphosyntactic variants in Czech. Corpus Linguistics and Linguistic Theory 8 (2): 241– 275. https://doi.org/10.1515/cllt-2012-0010

Bernard, S., Heutte, L., & Adam, S. (2009, June). Influence of hyperparameters on random forest accuracy. In International workshop on multiple classifier systems (pp. 171-180). Springer, Berlin, Heidelberg.

Bhilawadikar, V. S., & Garg, E. (2020). Investment attitude of millennials towards cryptocurrencies. Social Business. https://doi.org/10.1362/204440820x15929907056652

Bishop. (2017, March 30). “Verified” is now a derogatory term on Twitter. The Outline. https://theoutline.com/post/1323/verified-blue-checkmark-derogatory-insult-twitter?zd=1&zi=5p22jutz

Biau, G., & Scornet, E. (2016). A random forest guided tour. Test, 25(2), 197-227.

Booij, R.M. (2021, August 26). Brand Loyalty and Online Decision-Making. Business Economics. Retrieved from http://hdl.handle.net/2105/59195

Budiman, S. (2012). Analysis of consumer attitudes to purchase intentions of counterfeiting bag product in Indonesia. International Journal of Management, Economics and Social Sciences, 1(1), 1-12.

Carson, Richard T, and Theodore Groves. 2007. "Incentive and Informational Properties of Preference Questions." Environmental and Resource Economics 37 (1). Springer: 181–210.

Chrzan, K. & Sawtooth Software, Inc. (2015). How Many Holdout Tasks for Model Validation? Sawtooth Software Research Paper Series. https://sawtoothsoftware.com/resources/technical-papers/how-many-holdout-tasks-for-model-validation

Chu, S. C., Chen, H. T., & Sung, Y. (2016). Following brands on Twitter: An extension of theory of planned behavior. International Journal of Advertising, 35(3), 421-437.

Ciaian, P., Rajcaniova, M., & Kancs, D. (2018). Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. Journal of International Financial Markets, Institutions and Money, 52, 173–195. https://doi.org/10.1016/j.intfin.2017.11.001

Collins, T. U. T. (2021, April 8). Have social platforms reached their peak? Pew Research survey shows little user growth since last year. USA TODAY. https://eu.usatoday.com/story/tech/2021/04/07/pew-social-media-survey-facebook-youtube-gen-z-instagram-snapchat-tiktok/7107392002/

Corbet, S., Katsiampa, P., & Lau, C. K. M. (2020). Measuring quantile dependence and testing directional predictability between Bitcoin, altcoins and traditional financial assets. International Review of Financial Analysis, 71, 101571. https://doi.org/10.1016/j.irfa.2020.101571

Couper, M. P., Tourangeau, R., & Kenyon, K. (2004). Picture this! Exploring visual effects in web surveys. Public Opinion Quarterly, 68(2), 255-266.

Couronné, R., Probst, P., & Boulesteix, A. L. (2018). Random forest versus logistic regression: a large-scale benchmark experiment. BMC bioinformatics, 19(1), 1-14.

Cryptocurrency Marketing Agency. (2021, May 14). Single Grain. https://www.singlegrain.com/cryptocurrency-marketing-agency/#:%7E:text=What%20is%20cryptocurrency%20marketing%3F,coins%20need%20to%20stand%20out

Dąbrowska, E. (2010). Naive vs. expert intuitions: An empirical study of acceptability judgments. Te Linguistic Review 27: 1–23. https://doi.org/10.1515/tlir.2010.001

Daga, I., Gupta, A., Vardhan, R., & Mukherjee, P. (2020). Prediction of Likes and Retweets Using Text Information Retrieval. Procedia Computer Science, 168, 123–128. https://doi.org/10.1016/j.procs.2020.02.273

DAM, T. C. (2020). Influence of brand trust, perceived value on brand preference and purchase intention. The Journal of Asian Finance, Economics and Business, 7(10), 939-947.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. International journal of market research, 50(1), 61-104.

Dawes, J.G. (2002) Five point vs eleven point scales: does it make a difference to data characteristics? Australasian Journal of Market Research , 10, 1, pp. 39–47.

Demir, E., Simonyan, S., García-Gómez, C. D., & Lau, C. K. M. (2021). The asymmetric effect of bitcoin on altcoins: evidence from the nonlinear autoregressive distributed lag (NARDL) model. Finance Research Letters, 40, 101754. https://doi.org/10.1016/j.frl.2020.101754

Devgan, S. (2021, June 29). 100 Social Media Statistics For 2021 [+Infographic] |

Statusbrew. Statusbrew Blog. https://statusbrew.com/insights/social-media-statistics/

Ding, Min, Rajdeep Grewal, and John Liechty. 2005. “Incentive-Aligned Conjoint Analysis.” Journal of Marketing Research 42 (1). American Marketing Association: 67–82.

Devi, G. D., & Kamalakkannan, S. (2020). Literature Review on Sentiment Analysis in Social Media: Open Challenges toward Applications. Test Eng. Manag, 83(7), 2466-2474.

Dodge, Y. (2008). The Concise Encyclopedia of Statistics. Springer.

Eggers, F., Sattler, H., Teichert, T., & Völckner, F. (2022). Choice-Based conjoint analysis. In Handbook of market research (pp. 781-819). Springer, Cham.

ERDOGAN, M. C., & CANAYAZ, M. (2018). Crypto-Currency Sentiment Analyse on Social Media. 2018 International Conference on Artificial Intelligence and Data Processing (IDAP). https://doi.org/10.1109/idap.2018.8620724

Elendner, H., Trimborn, S., Ong, B., & Lee, T. M. (2018). The Cross-Section of Crypto-Currencies as Financial Assets 1 1Financial support from the Deutsche Forschungsgemeinschaft via CRC 649 “Economic Risk” and IRTG 1792 “High Dimensional Non Stationary Time Series,” Humboldt-Universität zu Berlin, is gratefully acknowledged. Handbook of Blockchain, Digital

Everitt, B. S.; Skrondal, A. (2010), The Cambridge Dictionary of Statistics, Cambridge University Press.

Dahan, E., & Hauser, J. R. (2002). The virtual customer. The Journal of Product Innovation Management , 19, 332-353.

GREEN, P. E., '" SRlNIVASAN, V. (1978). Conjoint analysis in consumer research: Issues and outlook. Journal of Consumer Research,S, 103-123.

Evgeniou, T., Boussios, C., & Zacharia, G. (2005). Generalized Robust Conjoint Estimation. Marketing Science , 24 (3), 415–429.

Endresen, A., & Janda, L. A. (2017). Five statistical models for Likert-type experimental data on acceptability judgments. Journal of Research Design and Statistics in Linguistics and Communication Science, 3(2), 217–250. https://doi.org/10.1558/jrds.30822

Fairchild, P. (2018, October 2). The Toxicity Of Blue Tick Culture - Phaylen Fairchild. Medium. https://phaylen.medium.com/the-toxicity-of-blue-tick-culture-b93be3cfb827

Finance, and Inclusion, Volume 1, 145–173. https://doi.org/10.1016/b978-0-12-810441-5.00007-5

Fox, R. J., Crask, M. R., & Kim, J. (1988). Mail survey response rate: A meta-analysis of selected techniques for inducing response. Public opinion quarterly, 52(4), 467-491.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016, June). Social clicks: What and who gets read on Twitter?. In Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science (pp. 179-192).

Gareth James, D. W. (2013). An introduction to statistical learning : with applications in R. Newyork: Springer.

Garimella, K., Weber, I., & de Choudhury, M. (2016). Quote RTs on Twitter. Proceedings of the 8th ACM Conference on Web Science. https://doi.org/10.1145/2908131.2908170

G.B. Dimitrov. (2017, August 29). Bayesian Truth Serum Fused Conjoint. Business Economics. Retrieved from http://hdl.handle.net/2105/39589

Glenn, S. (2020, December 16). Variance Inflation Factor. Statistics How To. https://www.statisticshowto.com/variance-inflation-factor/

Gogol, F. (2021, April 23). Study: 94% of Crypto Buyers are Gen Z/Millennial, but Gen X is Outspending Them. Stilt Blog. https://www.stilt.com/blog/2021/03/vast-majority-crypto-buyers-millennials-gen-z/

Grilli, L. and Rampichini, C. (2012). Multilevel models for ordinal data. In R. S. Kenett and S. Salini (Eds) Modern Analysis of Customer Surveys: with Applications using R, 391–408. Chichester: John Wiley and Sons.

Riquelme, F., & González-Cantergiani, P. (2016). Measuring user influence on Twitter: A survey. Information processing & management, 52(5), 949-975.

Green, Paul E, and Venkat Srinivasan. 1990. "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice." The Journal of Marketing. JSTOR, 3–19.

Guzman, A. (2018). The Ripple Effect of Cryptocurrencies. [online] Forbes.com. Available at: https://www.forbes.com/sites/forbesproductgroup/2018/01/11/the-rippleeffect-of-cryptocurrencies/#66286cc86080

Harcourt, G. C., & Kenyon, P. (1976). PRICING AND THE INVESTMENT DECISION. Kyklos, 29(3), 449–477. https://doi.org/10.1111/j.1467-6435.1976.tb01984.

Hastie, T. T. (2009). The elements of statistical learning: data mining, inference, and prediction. New York: Springer.

Harrell, F. E. (2015). Ordinal logistic regression. In Regression modeling strategies (pp. 311-325). Springer, Cham.

Hauser, John R, and Vithala R Rao. 2004. "Conjoint Analysis, Related Modeling, and Applications." In Marketing Research and Modeling: Progress and Prospects, 141–68. Springer.

Homer, P. M., & Kahle, L. R. (1990). Source expertise, time of source identification, and involvement in persuasion: An elaborative processing perspective. Journal of advertising, 19(1), 30-39.

Hu, A. S., Parlour, C. A., & Rajan, U. (2019). Cryptocurrencies: Stylized facts on a new investible instrument. Financial Management, 48(4), 1049–1068. https://doi.org/10.1111/fima.12300

Hutter, K., Hautz, J., Dennhardt, S., & Füller, J. (2013). The impact of user interactions in social media on brand awareness and purchase intention: the case of MINI on Facebook. Journal of Product & Brand Management.

Janitza, S., Tutz, G., & Boulesteix, A. L. (2016). Random forest for ordinal responses: Prediction and variable selection. Computational Statistics & Data Analysis, 96, 57–73. https://doi.org/10.1016/j.csda.2015.10.005

Jaakonmäki, R., Müller, O., & Vom Brocke, J. (2017). The impact of content, context, and creator on user engagement in social media marketing.

Jelena, A. L., & Ana, N. (2019). Designing e-learning environment based on student preferences: conjoint analysis approach. International Journal of Cognitive Research in Science, Engineering and Education, 7(3), 37-47.

Johfre, S. S., & Freese, J. (2021). Reconsidering the reference category. Sociological Methodology, 51(2), 253-269.

Jora, M. N., & Nandal, N. (2020). Investors Attitude towards Cryptocurrency-based on Gender. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 11(2), 622-630.

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. British journal of applied science & technology, 7(4), 396.

Keeney, R. L. (1971). Utility independence and preferences for multi attributed consequences. Operations Research, 19(4), 875-893.

Larsen, A., Tele, A., & Kumar, M. (2021). Mental health service preferences of patients and providers: a scoping review of conjoint analysis and discrete choice experiments from global public health literature over the last 20 years (1999–2019). BMC health services research, 21(1), 1-13.

Nickerson, C. A., McCLELLAND, G. H., & Petersen, D. M. (1990). Solutions to some problems in the implementation of conjoint analysis. Behavior Research Methods, Instruments, & Computers, 22(4), 360-374.

Kanalley, C. (2013, May 12). Why Twitter Verifies Users: The History Behind the Blue Checkmark. HuffPost. https://www.huffpost.com/entry/twitter-verified-accounts_b_2863282?guccounter=1

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

Huber, F., Herrmann, A., & Gustafsson, A. (2003). On the Influence of the Evaluation Methods in Conjoint Design — Some Empirical Results. In Conjoint Measurements: Methods and Applications (pp. 209-234). Springer Berlin Heidelberg.

Kaplan, C., Aslan, C., & Bulbul, A. (2018). Cryptocurrency Word-of-Mouth Analysis via Twitter. ResearchGate.

Kaushal, S. K., & Kumar, R. (2016). Influence of Attitude Towards Advertisement on Purchase Intention: Exploring the Mediating Role of Attitude Towards Brand Using SEM Approach. IUP Journal of Marketing Management, 15(4).

Kirasich, Kaitlin; Smith, Trace; and Sadler, Bivin (2018) "Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets," SMU Data Science Review: Vol. 1: No. 3, Article 9.

Kırcova, İ. (2021). Instagram, Facebook or Twitter: which engages best? A comparative study of consumer brand engagement and social commerce purchase intention.

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees?. Nature biotechnology, 26(9), 1011-1013.

Kinange, U. M., & Joshi, V. An Empirical Study on the Effect of User Generated Content on Consumer Purchase Intention. JOURNAL OF MANAGEMENT &

ENTREPRENEURSHIP, 87.

Kotsiantis, S. B. (2013). Decision trees: a recent overview. Artificial Intelligence Review, 39(4), 261-283.

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, 101188. https://doi.org/10.1016/j.intfin.2020.101188

Lavrakas, P. J. (2008). Encyclopedia of Survey Research Methods. Tousand Oaks, CA: SAGE Publications. https://doi.org/10.4135/9781412963947

Lee, D., Derrible, S., & Pereira, F. C. (2018). Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling. Transportation Research Record, 2672(49), 101-112.

Lee, K. H., & Shin, D. (2010). Consumers' responses to CSR activities: The linkage between increased awareness and purchase intention. Public Relations Review, 36(2), 193-195.

Lim, Y. S., & Lee-Won, R. J. (2017). When retweets persuade: The persuasive effects of dialogic retweeting and the role of social presence in organizations' Twitter-based communication. Telematics and informatics, 34(5), 422-433.

Louviere, J. J., & Islam, T. (2008). A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best–worst scaling. Journal of Business Research, 61(9), 903-911.

Majid, U. (2018). Research fundamentals: Study design, population, and sample size. Undergraduate research in natural and clinical science and technology journal, 2, 1-7.

Maragkou, E. (2019) "Blue Tick Brigade": A critique of Twitter's verification system as a shaper of online self-presentation.

Mas'od, A., Idris, U. N., Sulaiman, Z., & Chin, T. A. (2019, December). The Influence of Facebook Features and Activities on Consumers' Purchase Intention. In 2019 6th

International Conference on Research and Innovation in Information Systems (ICRIIS) (pp. 1-6). IEEE.

McCullough, D. (2002). A user's guide to conjoint analysis. Marketing Research, 14(2).

Metaxas, P., Mustafaraj, E., Wong, K., Zeng, L., O'Keefe, M., & Finn, S. (2021). What Do Retweets Indicate? Results from User Survey and Meta-Review of Research. Proceedings of the International AAAI Conference on Web and Social Media, 9(1), 658-661. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14661

Meynkhard, A. (2020). Effect of Bitcoin Volatility on Altcoins Pricing. Software Engineering Perspectives in Intelligent Systems, 652–664. https://doi.org/10.1007/978-3-030-63322-6_55

Menon, R. V., & Sigurdsson, V. (2016). Conjoint analysis for social media marketing experimentation: Choice, utility estimates and preference ranking. Managerial and Decision Economics, 37(4-5), 345-359.

Mildon. (2021). What Is Instagram and Why Should You Be Using It? Lifewire. https://www.lifewire.com/what-is-instagram-3486316

M. Saravanakumar, M. S., & T. Sugantha Lakshmi, T. S. L. (2012). Social Media Marketing. Life Science Journal, 9(4), 4444–4451. http://www.lifesciencesite.com/lsj/life0904/670_13061life0904_4444_4451.pdf

Naeem, M. A., Mbarki, I., Suleman, M. T., Vo, X. V., & Shahzad, S. J. H. (2020). Does Twitter Happiness Sentiment predict cryptocurrency? International Review of Finance, 21(4), 1529–1538.

Netzer, Oded, Olivier Toubia, Eric T Bradlow, Ely Dahan, Theodoros Evgeniou, Fred M Feinberg, Eleanor M Feit, et al. 2008. "Beyond Conjoint Analysis: Advances in Preference Measurement." Marketing Letters 19 (3-4). Springer: 337–54.

Nakamoto, S., 2008. Bitcoin: A Peer-to-Peer Electronic Cash System.

Nandal, N., & Jora, M. N. (2020a). Impact of Household Income on Investors Attitude towards Crypto Currency. Annals of the Romanian Society for Cell Biology, 449-458.

Nitze, I., Schulthess, U., & Asche, H. (2012). Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. Proceedings of the 4th GEOBIA, Rio de Janeiro, Brazil, 79, 3540.

Orme, B. (2002). Interpreting conjoint analysis data. Sawtooth Software Research Paper Series.

Patil, P. S. (2019). Attitude of the millennial generation towards cryptocurrency. DBS eSource. https://esource.dbs.ie/handle/10788/3817

Peterson, R.A. & Wilson, W. (1992) Measuring customer satisfaction: fact and artifact. Journal of the Academy of Marketing Science , 20, 1, pp. 61–71.

Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: data mining and knowledge discovery, 9(3), e1301.

Qiu, T., Zhang, R., & Gao, Y. (2019). Ripple vs. SWIFT: Transforming Cross Border Remittance Using Blockchain Technology. Procedia Computer Science, 147, 428–434. https://doi.org/10.1016/j.procs.2019.01.260

Quinlan, J. R. (1987). Simplifying decision trees. International journal of man-machine studies, 27(3), 221-234.

Raileanu, L. E., & Stoffel, K. (2004). Theoretical comparison between the gini index and information gain criteria. Annals of Mathematics and Artificial Intelligence, 41(1), 77-93.

Rawale, S. (2018, June 14). Understanding Decision Tree, Algorithm, Drawbacks and Advantages. Medium.

Richard, J., & Guppy, S. (2014). Facebook: Investigating the influence on consumer purchase intention.

Roh, E. K., & Kim, S. H. (2007). Conjoint analysis of high-performance fabrics for mountaineering jacket. Fibers and Polymers, 8(6), 649-653.

Rokach, L., & Maimon, O. (2005). Decision trees. In Data mining and knowledge discovery handbook (pp. 165-192). Springer, Boston, MA.

Rosner, M. T., & Kang, A. (2015). Understanding and Regulating Twenty-First Century Payment Systems: The Ripple Case Study Notes. Michigan Law Review 114. Mich. L. Rev. https://heinonlineorg.eur.idm.oclc.org/HOL/Page?collection=journals&handle=hein.journals/mlr114&id=687&men_tab=srchresults

Rousseau, D. M., Sitkin, S. B., Burt, R. S. & Camerer, C. (1998), "Not so different after all: a cross-discipline view of trust". Academy of Management review, vol. 23, no. 3, pp. 393-404.

Sanders, S. (2022). Invest Like a Millennial. Sabinet African Journals. https://doi.org/10.10520/ejc-monpf_v2021_n485_a4

Sayegh, E. (2022, August 4). Twitter's Bots Were The Business All Along. Forbes. Retrieved September 5, 2022, from https://www.forbes.com/sites/emilsayegh/2022/08/04/twitters-bots-were-the-business-all-along/

Shan, L. C., De Brún, A., Henchion, M., Li, C., Murrin, C., Wall, P. G., & Monahan, F. J. (2017). Consumer evaluations of processed meat products reformulated to be healthier–A conjoint analysis study. Meat Science, 131, 82-89.

Shi, T., & Horvath, S. (2006). Unsupervised learning with random forest predictors. Journal of Computational and Graphical Statistics, 15(1), 118-138.

Soutar, G. N., & Turner, J. P. (2002). Students' preferences for university: A conjoint analysis. International journal of educational management.

Statista. (2022, January 28). Twitter: distribution of global audiences 2021, by age group.

https://www.statista.com/statistics/283119/age-distribution-of-global-twitter-users/#:%7E:text=As%20of%20April%202021%2C%20Twitter,share%20of%20almost%2021%20percent

Strobl, C., Malley, J., and Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classifcation and regression trees, bagging, and random forests. Psychological Methods 14 (4): 323–348. https://doi.org/10.1037/a0016973

Toubia, O., Hauser, J. R., & Simester, D. I. (2004). Polyhedral methods for adaptive choice-based conjoint analysis. Journal of Marketing Research, 41(1), 116-131.

TripleA crypto payment. (2021, October 14). Global Cryptocurrency Ownership Data 2021. TripleA. https://triple-a.io/crypto-ownership/

Ventola C. L. (2014). Social media and health care professionals: benefits, risks, and best practices. P & T : a peer-reviewed journal for formulary management, 39(7), 491–520.

von Muhlen, M., & Ohno-Machado, L. (2012). Reviewing social media use by clinicians: Table 1. Journal of the American Medical Informatics Association, 19(5), 777–781. https://doi.org/10.1136/amiajnl-2012-000990

Warner, P. (2008). Ordinal logistic regression. Journal of Family Planning and Reproductive Health Care, 34(3), 169.

Walker, Joan, and Moshe Ben-Akiva. 2002. “Generalized Random Utility Model.” Mathematical Social Sciences 43 (3). Elsevier: 303–43.

Weisberg, J., Te'eni, D. & Arman, L. (2011), "Past purchase and intention to purchase in ecommerce. The mediation of social presence and trust". Internet Research, vol. 21, no. 1, pp. 82-96.

Weismueller, J., Harrigan, P., Wang, S., & Soutar, G. N. (2020). Influencer endorsements: How advertising disclosure and source credibility affect consumer purchase intention on social media. Australasian marketing journal, 28(4), 160-170.

Xi, D., O'Brien, T. I., & Irannezhad, E. (2020). Investigating the investment behaviours in cryptocurrency. The Journal of Alternative Investments, 23(2), 141-160.

Xu, H., Phoa, F. K., & Wong, W. K. (2009). Recent developments in nonregular fractional factorial designs. Statistics Surveys, 3, 18-46.

Zhao, H. and Zhang, L. (2021), "Financial literacy or investment experience: which is more influential in cryptocurrency investment?", International Journal of Bank Marketing, Vol. 39 No. 7, pp. 1208-1226. https://doi-org.eur.idm.oclc.org/10.1108/IJBM-11-2020-0552

Zimbra, D., Abbasi, A., Zeng, D., & Chen, H. (2018). The State-of-the-Art in Twitter Sentiment Analysis. ACM Transactions on Management Information Systems, 9(2), 1–29. https://doi.org/10.1145/3185045

Hauber, A. B., Gonzáles, J. M., Prior, C. G., Marshall, D. A., Cunningham, C., IJzerman, M. J., & Bridges, J. F. (2016). Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR conjoint analysis good research practices task force. Value in Health, 19(4), 300- 315.

Zhang, J., Wei, X., Fukuda, H., Zhang, L., & Ji, X. (2021). A Choice-based conjoint analysis of social media picture posting and souvenir purchasing preference: A case study of social analytics on tourism. Information Processing & Management, 58(6), 102716.

Appendices

Appendix A

Cover Page & Demographic Questions

Cover Page:
Hi! I am Rohan Ghosh, a Data Science and Marketing Analytics student who is majoring at Erasmus University Rotterdam. In this survey I am conducting for my Master's thesis, I explore how Tweet properties on Twitter influence consumer purchase/investment intention of Millennial first-time investors of the cryptocurrency Ripple (XRP). Please only answer this survey if you are between the ages of 26-41 and are a first-time investor in Cryptocurrency.

There will be 12 Choice Tasks with three tweets that will be displayed. Please choose the tweet that you feel motivates you the most to purchase XRP.

It is important to note that the price of the cryptocurrency being mentioned in the tweets (XRP), the content of the Tweets and the username/user of the Tweets are held constant throughout all Tweet options in the survey.

The survey should take no longer than 6 minutes. If you have any remarks, questions or concerns please feel free to contact me at my university email at 497188rg@eur.nl. Thank you for your time!

Demographic Questions:

What gender are you?

- ◯ Male
- ◯ Female
- ◯ Non-binary / third gender
- ◯ Prefer not to say

How old are you?

- ◯ Younger than 18
- ◯ 18-25
- ◯ 26-41
- ◯ 42+

Have you invested in cryptocurrency before?

- ◯ Yes
- ◯ No

What is your household income?

- ◯ <€20,000
- ◯ €20,000-€39,999
- ◯ €40,000-€59,999
- ◯ €60,000-€79,999
- ◯ €80,000+

7.2 Appendix B

Orthogonal Design:

Choice Set	Choice ID	Likes	Retweets	Quote Twe	Tick
1	1	2	1	1	no
1	2	3	3	1	yes
1	3	1	2	3	yes
2	1	3	3	1	no
2	2	1	1	2	yes
2	3	2	3	3	yes
3	1	1	1	1	no
3	2	3	3	3	no
3	3	2	2	2	yes
4	1	1	2	1	no
4	2	3	3	3	no
4	3	2	1	1	yes
5	1	2	3	1	no
5	2	1	2	2	yes
5	3	3	2	3	yes
6	1	1	1	1	no
6	2	1	1	2	no
6	3	2	3	3	yes
7	1	3	1	3	no
7	2	1	2	1	yes
7	3	2	2	2	yes
8	1	3	3	3	no
8	2	2	1	1	yes
8	3	1	2	2	yes
9	1	3	2	2	no
9	2	2	1	1	yes
9	3	2	3	3	yes
10	1	3	2	1	no
10	2	2	3	2	no
10	3	1	1	3	yes
11	1	2	2	1	no
11	2	3	1	2	no
11	3	2	3	3	yes
12	1	1	1	2	no
12	2	2	3	3	no
12	3	3	2	1	yes

Main Section (Survey Screenshots):


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 7 Quote Tweets 402 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 13 Quote Tweets 803 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 7 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 13 Quote Tweets 402 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 803 Likes

Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085

Select the Tweet that motivates you most to purchase XRP.

CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 7 Quote Tweets 402 Likes

CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 13 Quote Tweets 1.2K Likes

CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 20 Quote Tweets 402 Likes

Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085

Select the Tweet that motivates you most to purchase XRP.

CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 7 Quote Tweets 402 Likes

CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 13 Quote Tweets 402 Likes

CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 803 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 7 Quote Tweets 803 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 13 Quote Tweets 402 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 20 Quote Tweets 1.2K Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 7 Quote Tweets 803 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 13 Quote Tweets 402 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 20 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 7 Quote Tweets 402 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 13 Quote Tweets 803 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 13 Quote Tweets 803 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 7 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 20 Quote Tweets 402 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 13 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 7 Quote Tweets 803 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 803 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 7 Quote Tweets 402 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 7 Quote Tweets 803 Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 13 Quote Tweets 402 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
195 Retweets 20 Quote Tweets 803 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 7 Quote Tweets 1.2K Likes


Please use the price of Ripple and the Tweets below as reference:
Price of XRP at time of Tweet: $0.398085
Select the Tweet that motivates you most to purchase XRP.
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
130 Retweets 7 Quote Tweets 803 Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
65 Retweets 13 Quote Tweets 1.2K Likes
CryptoAdvisor
@CryptoAdvisor
Now is a great time to invest in Ripple (XRP)!
19:00 · 07/17/2022 · Twitter Web App
190 Retweets 20 Quote Tweets 803 Likes

7.3 Appendix C

When the aforementioned method was applied to the tweet property "likes" the following values were found:

Average from dataset (base value/value 2): 803

1. Value/Level 1 (50% lower) = (803/100)*50 = 402 (rounded)
2. Value/Level 3 (50% higher) = (803/100)*1.5 = 1205 (rounded)

When applied to Retweets:

Average from dataset (base value/value 2): 130

1. Value/Level 1 (50% lower) = (803/100)*50 = 65(rounded)
2. Value/Level 3 (50% higher) = (803/100)*1.5 = 195 (rounded)

When applied to Quote Tweets:

Average from dataset (base value/value 2): 13

1. Value/Level 1 (50% lower) = (803/100)*50 = 7 (rounded)
2. Value/Level 3 (50% higher) = (803/100)*1.5 = 20 (rounded)

Blue Tick Verification is a binary variable that only has two values:

1. Yes
2. No