UNIVERSITY OF LJUBLJANA
SCHOOL OF ECONOMICS AND BUSINESS

MASTER'S THESIS

THE LINK BETWEEN SOCIAL MEDIA SENTIMENT AND STOCK PRICES OF SELECTED FIRMS IN THE GAMING INDUSTRY

Ljubljana, July 2025

NEJC BRANCELJ

AUTHORSHIP STATEMENT

The undersigned Nejc Brancelj, a student at the University of Ljubljana, School of Economics and Business, (hereafter: UL SEB), author of this written final work of studies with the title “The link between social media sentiment and stock prices of select firms in the video game industry”, prepared under supervision of Jurij Jaklič, PhD

D E C L A R E

1. this written final work of studies to be based on the results of my own research;
2. the printed form of this written final work of studies to be identical to its electronic form;
3. the text of this written final work of studies to be language-edited and technically in adherence with the UL SEB Technical Guidelines for Written Works, which means that I cited and / or quoted works and opinions of other authors in this written final work of studies in accordance with the UL SEB Technical Guidelines for Written Works;
4. to be aware of the fact that plagiarism (in written or graphical form) is a criminal offence and can be prosecuted in accordance with the Criminal Code of the Republic of Slovenia;
5. to be aware of the consequences a proven plagiarism charge based on the this written final work could have for my status at the UL SEB in accordance with the relevant UL SEB Rules;
6. to have obtained all the necessary permits to use the data and works of other authors which are (in written or graphical form) referred to in this written final work of studies and to have clearly marked them;
7. to have acted in accordance with ethical principles during the preparation of this written final work of studies and to have, where necessary, obtained permission of the Ethics Committee;
8. my consent to use the electronic form of this written final work of studies for the detection of content similarity with other written works, using similarity detection software that is connected with the UL SEB Study Information System;
9. to transfer to the University of Ljubljana free of charge, non-exclusively, geographically and time-wise unlimited the right of saving this written final work of studies in the electronic form, the right of its reproduction, as well as the right of making this written final work of studies available to the public on the World Wide Web via the Repository of the University of Ljubljana;
10. my consent to publication of my personal data that are included in this written final work of studies and in this declaration, when this written final work of studies is published;
11. that I have verified the authenticity of the information derived from the records using artificial intelligence tools.

Ljubljana, July 5th, 2025 Author’s signature:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

sl. – Slovene
ADF – Augmented Dickey-Fuller
AMC – AMC Entertainment Holdings Inc
API – (sl. Vmesniki aplikacijskih programov); Application Programming Interface
ATVI – Activision Blizzard
BERT – Bidirectional Encoder Representations from Transformers
BF 2042 – Battlefield 2042
BL 3 – Borderlands 3
CDR – CD Projekt Red
CoD – Call of Duty
CP 2077 – Cyberpunk 2077
CRISP-DM – Cross-Industry Standard Process for Data Mining
CSV – (sl. Vrednosti, ločene z vejico); Comma-separated values
DJIA – Dow Jones Industrial Average
DLC – (sl. Prenosljiva vsebina); Downloadable content
EA – Electronic Arts
GaaS – (sl. Igra kot storitev); Game as a Service
GME – GameStop
GTA V – Grand Theft Auto V
HTML – (sl. Jezik za označevanje nadbesedila); Hypertext Mark-up language
LLM – (sl. Veliki jezikovni model); Large Language Model
ML – (sl. Strojno učenje); Machine Learning
MMO – (sl. Masovna spletna igra); Massively Multiplayer Online
NASDAQ – National Association of Securities Dealers Automated Quotations
NLP – (sl. Obdelava naravnega jezika); Natural Language Processing
NYSE – New York Stock Exchange
OW – Overwatch
PC – (sl. Osebni računalnik); Personal Computer
RDR2 – Red Dead Redemption 2
RoBERTa – (sl. Robustno optimiziran BERT model); Robustly Optimized BERT Approach
S&P – Standard and Poor's

SENSEX – Stock Exchange Sensitive Index
TTWO – Take Two Interactive
VADER – **(**sl.Valenčno občutljiv slovar in analizator sentimenta**);** Valence Aware Dictionary and sentiment Reasoner
WSB – Wall Street Bets
ZST – (sl. Zstandard stisnjene datoteke); Zstandard compressed files

1 INTRODUCTION

Financial markets have long been influenced by a plethora of factors, ranging from earnings reports to macroeconomic policies (Cutler et al., 1998). However, with the onslaught of the digital age, new opportunities to influence have arisen. With the widespread access to information and new platforms to express one's opinions, social media has emerged as a new and increasingly influential driver of stock price movements. Sites like Twitter (now X) and Reddit provide a real-time feed of mass sentiment. Because of this, research on the relationship between stock prices and social media sentiment has gained significant traction in recent years. Aside from Twitter, financial blogs and news are the main sources of sentiment (Garg & Tiwari, 2021). However, the findings remain inconsistent. Some researchers suggest there is a strong correlation (Ho et al., 2017; Pagolu et al., 2016; Ranco et al., 2015), while some find no significance (Behrendt & Schmidt, 2018; Piñeiro-Chousa et al., 2017; Sprenger et al., 2014), others observe that the impact is short-lived and context-dependent (Chahine & Malhotra, 2018; Li et al., 2017). Unlike Twitter, Reddit only came to the forefront of the research after the GameStop saga in 2021. This caused a surge in publications, such as Corbet et al. (2022), Hu et al. (2021), Long et al. (2023), Mancini et al. (2022), and Wang & Luo (2021). Meanwhile, the overall area of research still lacks volume and diversity, inviting further efforts. Reddit differs from platforms such as Twitter in its very rigid structure, enforcing a strict separation between communities, making it an ideal environment for analysing specific online communities.

Within such communities varying by engagement and importance in their respective industries, there is one that stands out. That is the gaming industry, where consumer engagement and opinions are of utmost importance for the success of a game (Ahmad et al., 2017; Yee, 2016). Gaming companies operate in a highly competitive market, fighting for the consumers' attention, which is crucial due to the monetisation techniques used in video games (Ivanov et al., 2021; Vaudour & Heinze, 2020). This creates a unique relationship between consumers, products, and companies. Hence, instead of focusing on the financial discussion as was done in previous research, we will explore whether the customers' opinion in any way reflects the stock prices of developers. This notion is based on three pillars. Firstly, cultivating a strong online community is crucial for the success of video games. Secondly, the gaming community has a strong presence on Reddit. Lastly, players' satisfaction is reflected in the game's revenue via monetisation techniques, affecting the developer's stock price. Therefore, the main research question is: What, if any, is the relationship between the sentiment on Reddit and stock prices of video game companies?

The purpose of this master's thesis is to deepen the understanding of how customer satisfaction on social media influences stock prices, addressing a critical research gap in both customer sentiment analysis and its connection to the financial markets, particularly within the video game industry. The findings of this study are relevant to companies that want to understand the value of monitoring social platforms, the importance of user opinions and their connection to stock prices; to investors looking for new indicators for potential investments; to consumers interested in the value of their opinion; to further research in the areas of social media, the Reddit platform, the video game industry, sentiment analysis and its relation to stock movements.

Consequently, the goals of this research are:

1. To obtain relevant data and prepare it for sentiment analysis within the framework of an appropriate data processing model.
2. To make the necessary augmentations to capture the financial, sentiment, and emotion data.
3. To perform sentiment analysis, considering both sentiment and emotion classification scores within the context of the studied companies.
4. To identify any significant local events that might carry a stronger connection between sentiment and stock datasets.
5. To calculate the correlation between the datasets to discover possible connections for both global and local periods.
6. To perform a Granger causality test for the stock–sentiment and stock–emotion interactions for global and local periods.
7. To discover the nature of the connection between the datasets, whenever it exists on a long-term level, or only for individual events.

To achieve set goals, we use the well-established Cross-Industry Standard Process for Data Mining (CRISP-DM) framework devised by Wirth & Hipp (2000) to develop our methods for data collection, processing and analysis. This allows us to apply an iterative approach, which yields optimal techniques for the respective steps of our research. We begin by defining our scope of research, i.e., which companies and subreddits to study. We do so by devising a process that considers the companies' size, revenue structure, released games, monetisation types, size of online communities, etc. Then we move to processing and analysis, for which we use Python, a widely used programming environment that offers a variety of libraries for the tasks at hand (with all the used scripts available at https://github.com/nejcb-00/masters).

We split our data collection and preparation process into two parts, as the financial and sentiment datasets are completely different. The former we acquired from the National Association of Securities Dealers Automated Quotations (NASDAQ) stock exchange and the financial portal Investing.com. The data is in a very simple format and, therefore, easy to prepare for the analysis. The opposite could be said for our sentiment data, which we gathered from the Pushshift dataset published by Baumgartner et al. (2020) on the Academic Torrents repository. This data contains comments and submissions for the selected subreddits. As noted by Baldwin et al. (2013), social media text is full of quality issues and overall hard to analyse. Hence, we apply some data-processing techniques by Chai (2023), but skip most of the steps, since our analysis methods, based on Valence Aware Dictionary and sentiment Reasoner (VADER) and Robustly Optimized BERT Approach (RoBERTa) do not require rigorous data preprocessing. With the sentiment data ready, we then conduct the first part of our analysis in which we calculate the sentiment scores and emotion classification scores. The sentiment is calculated using VADER, a classifier developed by Hutto & Gilbert (2014), specifically for social media texts, and is, therefore, the foremost tool used amongst researchers (Huynh et al., 2021; Long et al., 2023; Mancini et al., 2022; Reichenbach & Walther, 2023; Wang & Luo, 2021). The emotion classification scores are calculated using a RoBERTa-based model made by Liu et al. (2019), which is an ever more popular approach (Rahman et al., 2023; Semenova et al., 2021; Sousa et al., 2019; Wang & Luo, 2021). The model uses a Large Language Model (LLM) developed by Devlin et al. (2018) called Bidirectional Encoder Representations from Transformers (BERT) to create a classification system based on Twitter data. This creates a synergy with using VADER, as both approaches require similar data preparation and are based on the same principle, i.e., using existing social media data to calculate sentiment metrics.

With both financial and sentiment data prepared, we first test the stationarity of our datasets by applying the Augmented Dickey-Fuller (ADF) test, which ensures that our time series are suitable for further analyses. With stationarity ensured, we then identify significant local events, e.g., game releases, which we will further analyse. For the analysis, we follow Mudinas et al. (2019), Ranco et al. (2015), and Sun et al. (2017) and use a combination of correlation and Granger's causality to determine the nature of the connection. First, we calculate Pearson correlation coefficients for our time series and conduct a cross-correlation analysis for the overall sentiment. Then, we calculate Granger causality between the datasets, which tells us if sentiment or emotions can be considered as a predictor for the behaviour of the stock prices (Granger, 1969). We do so for both the overall observational period and the identified significant local events. Finally, we study these connections on the level of the whole company as well as on the level of individual communities. Hence, we will be able to identify both local and global connections between sentiment and stock prices, which we can then interpret within the context of the industry.

To do that, we contextualise our research proposal by presenting the field of sentiment, social media and the stock market, where we review the related works and highlight the lack of research about Reddit as a source of sentiment data within the context of the video game industry. We then explain the history and the current state of the social media landscape, while pointing out the uniqueness of Reddit and why it is best suited for use in our analysis. Next, we draw our focus to the video game industry, explaining the unique customer-game-developer relationship within it and how the monetisation of its products in combination with the strong online presence of its customers makes for a unique case to study. Now, we move to the methodological part of the thesis, where we define the scope and the overall approach. Then, we describe the data collection, preparation and analysis methods which lead us to the results. Thereafter, we present our findings from the sentiment analysis and their link to the stock market via the correlation and causality tests. We then contextualise our observations with the discussion of the results and finally conclude by presenting the key takeaways and their implications.

2 SENTIMENT, SOCIAL MEDIA, AND THE STOCK MARKET

2.1 Social Media and Stock Prices

The connection between social media sentiment and stock prices is a relatively new area of interest for research. Garg & Tiwari (2021) identify that the trend started around 2010, with the number of publications growing by the year. This is supported by Nyakurukwa & Seetharam (2023) in a similar literature overview study. Furthermore, both studies find that the sentiment data is almost exclusively obtained from Twitter. This matches the findings by Janková (2023), who explores the corpora behind sentiment analysis for stock market predictions, identifying financial news and Twitter as the main corpora behind sentiment analysis. Whereas the amount of research in this field is plentiful, it is by no means consistent. This can be partially attributed to the nature of stock price movements since they are influenced by numerous factors, which makes it difficult to determine if the changes are caused solely by the predictors we identify (Cutler et al., 1998). And with the growing number of information channels available today, the complexity of influence over stock prices increases.

Another factor contributing to the inconsistency is the nature of social media sentiment, as well as how it is linked to the stock market. Naturally, this depends on the goal of the research. Pagolu et al. (2016) use basic correlation to study the link, while Mudinas et al. (2019), Long et al. (2023), and Ranco et al. (2015) take it a step further by using Granger causality to see if sentiment could forecast stock prices. Some papers (Chahine & Malhotra, 2018; Corbet et al., 2022; Duz Tan & Tas, 2021; Ho et al., 2017; Huynh et al., 2021; Piñeiro-Chousa et al., 2017; Reichenbach & Walther, 2023; Sprenger et al., 2014; Sul et al., 2017) use various regression models to see if any more meaningful links exist. Newer methods, such as machine learning, are also regularly present in research (Chen & Ma, 2024; Derakhshan & Beigy, 2019; Teoh et al., 2019). Furthermore, many papers use machine learning as a basis for their custom models (Awan et al., 2021; Biswas et al., 2020; Broadhurst et al., 2022; Li et al., 2017; Mudinas et al., 2019; Nguyen et al., 2015; Nguyen & Shirai, 2015; Semenova et al., 2021). As we can see, the methods used are very diverse, which also translates to the diversity in findings.

Ho et al. (2017), Pagolu et al. (2016), and Ranco et al. (2015) all find that there is a strong correlation between online opinions and stock prices. Additionally, Chen et al. (2014) find that these opinions strongly predict future stock returns, even after accounting for the influence of other sources such as financial analysis and news. Duz Tan & Tas (2021) present similar findings, while also implying that the frequency of posts plays a role as a predictor.

On the contrary, Behrendt & Schmidt (2018) consider online sentiment as not particularly useful for investors. This is supported by Piñeiro-Chousa et al. (2017) and Sprenger et al. (2014), who do not find any meaningful link to the stock prices. Chahine & Malhotra (2018) and Li et al. (2017) observe that sentiment can be a strong predictor of specific events, with the opinions having a short diffusion time of only a few days. This is contradicted by Sul et al. (2017), stating that the opinions normally diffuse over a longer period of up to 20 days. The lack of consistency can also be observed in publications focused on different sentiment-based prediction models. Models from Derakhshan & Beigy (2019), Nguyen et al. (2015), and Nguyen & Shirai (2015) all predict around 50–60 % accuracy. And even though Awan et al. (2021) and Li et al. (2017) managed to achieve 70–95% accuracy with some of the models, these were stock-specific results and therefore cannot be translated to a more general setting.

While Twitter, financial blogs and the news have been at the forefront as the source of sentiment data for financial analysis, this is not the case for Reddit. According to Proferes et al. (2021), publications about Reddit followed a similar trend to the rest of social media sites, with 2015 being the turning point, after which research has substantially increased. However, the papers covering business and economics have been scarce (only six publications fall into that category). Moreover, if we peruse the scientific database Google Scholar for the papers published before 2021, only Biswas et al. (2020), Mudinas et al. (2019), Oncharoen & Vateekul (2018), and Teoh et al. (2019) used Reddit as a part of the sentiment data when constructing prediction models. The lull of research ended swiftly in early 2021 after a financial community on Reddit orchestrated one of the most unique events in recent stock market history. As explained by Duterme (2023) and Malz (2021), the GameStop saga came about after a hedge fund, Melvin Capital, put out a short position on the GameStop stock. This incited the members of the subreddit r/wsb to drive up the stock price to force a short squeeze (when a stock price moves up to the point when the short holders are forced to cover their positions or face margin calls). The Reddit community emerged victorious, and Melvin Capital was forced to close down due to the losses incurred. This event brought on a new wave of research surrounding the GameStop stock and the r/wsb community. However, the authors do not reach a consensus on the effect on stock movements. Hu et al. (2021) consider sentiment as a significant predictor, while Mancini et al. (2022) and Wang & Luo (2021) doubt that sentiment had any concrete effect. With most research (Corbet et al., 2022; Huynh et al., 2021; Long et al., 2023; Machavarapu, 2022) we observe a short-term correlation, which is unlikely to carry overall significance. But regardless of the surge in publications, the overall lack of volume and diversity in the field remains. The post-2021 research focuses on financial communities and popular stocks and stock indices (Broadhurst et al., 2022; Reichenbach & Walther, 2023; Semenova et al., 2021). Meanwhile, Chen & Ma (2024) and Lindskog & Serur (2020) do not limit themselves to only financial communities and gather sentiment data from all of Reddit. Nevertheless, the overall narrow focus of the research remains. To better understand the lack of diversity, Table 1 displays the sources of sentiment and financial data used by publications.

Table 1: Overview of published studies regarding social media and stock prices

	Sentiment data		Financial data
	Source	Type	
Sprenger et al. (2014)	Twitter	financial	S&P 100 index
Chen et al. (2014)	Seeking Alpha	financial	DJIA index
Ranco et al. (2015)	Twitter	financial	DJIA index
Nguyen et al. (2015)	Yahoo Finance	financial	Industry leaders' stocks
Nguyen & Shirai (2015)	Yahoo Finance	financial	Industry leaders' stocks
Pagolu et al. (2016)	Twitter	general	NASDAQ-100 index
Sul et al. (2017)	Twitter	financial	S&P 500 index
Piñeiro-Chousa et al. (2017)	StockTwits.com	financial	S&P 500 index
Li et al. (2017)	Twitter	financial	30 stocks on NASDAQ
Ho et al. (2017)	Yahoo Finance	financial	DJIA index
Oncharoen & Vateekul (2018)	News	general	S&P 500 and DJIA indices
Chahine & Malhotra (2018)	Twitter	general	312 stocks on S&P 500
Behrendt & Schmidt (2018)	Twitter	general	DJIA index
Teoh et al. (2019)	Reddit (r/news)	general	stocks on NASDAQ
Mudinas et al. (2019)	Financial times, Reddit (r/rwnc), Twitter	financial, general	DJIA index
Biswas et al. (2020)	Twitter, Yahoo Finance	financial, general	SENSEX index
Wang & Luo (2021)	Reddit (r/wsb)	financial	$GME
Huynh et al. (2021)	Reddit (r/wsb), Yahoo Finance	financial	$GME
Hu et al. (2021)	Reddit	general	Industry leaders
Duz Tan & Tas (2021)	Twitter	financial	S&P 500, S&P 350, S&P EMC
Awan et al. (2021)	News, Twitter, Yahoo Finance, various reviews	financial, general	Industry leaders' stocks
Semenova et al. (2021)	Reddit (r/wsb)	financial	Industry leaders' stocks
Mancini et al. (2022)	Reddit (r/wsb)	financial	$GME and stock indices
Machavarapu (2022)	Reddit (r/wsb)	financial	$GME, $AMC
Corbet et al. (2022)	Reddit (r/wsb)	financial	$GME and industry leaders

To be continued

Table 1: Overview of published studies regarding social media and stock prices (cont.)

Broadhurst et al. (2022)	Reddit	general	stocks on NASDAQ
Reichenbach & Walther (2023)	Reddit (r/wsb, r/investing)	financial	stocks on NYSE and NASDAQ
Long et al. (2023)	Reddit (r/wsb)	financial	$GME, Russel2000 index

Source: Own work.

There are two main observations we can make regarding the overview in Table 1. Firstly, publications use financial-themed sentiment or non-filtered general sentiment. Secondly, financial data mostly consists of stock indices, industry leaders' stocks, and the GameStop (GME) stock. This raises the question, what if instead, we were to look at the link between customer satisfaction and the stock prices in an industry that has a unique way of tying them together? That is the video game industry. Studies in this field have been few and far between. Zhu et al. (2010) laid out the groundwork for all further research by studying the impact of online reviews on video game sales. Xiong & Bharadwaj (2014) expanded the scope to include the stock market while studying the effect of the pre-release discussions online. However, further studies on this topic have been scarce. With exceptions like Suh & Lee (2011) who studied the events in online games and stock prices, and more recently Piñeiro-Chousa et al. (2023) who explored the influence of the streaming platform Twitch on pre- and post-COVID-19 stock returns. Besides published papers, there are also some relevant theses on this topic. Surminski (2023) studies the role of game quality in the financial markets. Canto (2019) and Hermansen & Bratli (2016) explore the effect of online game reviews on the firm value and stock prices, respectively. Mertová (2023) is the only one focusing on the connection between social media and the stocks of video game companies. While their broader research concept matches ours, the data used differs (general and financial sentiment from Twitter combined with news headlines). Hence, the usage of gaming-specific communities on Reddit in such a case is yet to be explored, whereas the overall field of analysing social media sentiment has no scarcity of research.

2.2 Social Media Sentiment

The collection and analysis of online sentiment has been an established field of research since the start of the mainstream popularity of the first web-based news outlets, forums, blogs, etc. As observed by Medhat et al. (2014), research in the early 2010s was based on data from news articles, reviews, forums and personal blogs. Whereas Rodríguez-Ibánez et al. (2023) find that the papers focused on social media came to the forefront after 2015, the number of publications growing steadily in the following years. Building on this shift in focus towards social media, let us see how we can harness online sentiment.

The journey from raw data gathered online to a meaningful analysis is quite complex. Therefore, data science frameworks have been established to structure such endeavours.

Saltz & Hotz (2020) identify CRISP-DM, Scrum, and Kanban as the most used among industry professionals. While these frameworks are not designed for social media sentiment analysis, their universality allows us to apply them to any problem, since they revolve around collecting, processing, modelling, and analysing data. The most common approach for gathering mass sentiment data online is an API (Application Programming Interface). API offers a simple way of connecting to and integrating with any piece of software (Biehl, 2014, p. 15). In our case, these are websites. Hence, they are widely used in research to obtain data from online platforms (Corbet et al., 2022; Duz Tan & Tas, 2021; Huynh et al., 2021; Lindskog & Serur, 2020; Pagolu et al., 2016; Piñeiro-Chousa et al., 2017, 2023; Ranco et al., 2015; Sprenger et al., 2014)**.** The other main source of data is databases (Awan et al., 2021; Chahine & Malhotra, 2018; Chen & Ma, 2024; Mudinas et al., 2019; Wang & Luo, 2021). These are usually filled with data acquired from APIs over time and have become increasingly popular since platforms like Twitter and Reddit started to charge for access to their data (Reddit, 2024a; X, 2024). Such databases can be found on Academic Torrents, a community-made repository that houses many datasets intended for research (Cohen & Lo, 2014). Amongst these, the Pushshift Reddit Dataset is one of the most updated and used on the website (Baumgartner et al., 2020). Therefore, gathering the data is very straightforward. However, the same cannot be said for processing the data.

Social media text is especially difficult to process. As observed by Baldwin et al. (2013), Han et al. (2013), and Salloum et al. (2017), this is due to the non-structured nature of the text, which also contains a lot of slang and jargon. Moreover, there is no shortage of images, links, emoticons, and other non-textual elements embedded in the text. Hence, the stage of processing the data is the most meticulous and time-consuming. Chai (2023), Krouska et al. (2016), and Vijayarani et al. (2015) identify five steps that are commonly used in text processing. Firstly, the formatting of the text must be removed. Secondly, the text is tokenised, i.e., split into words or phrases. Thirdly, the text is normalised by removing punctuation, converting it to lowercase, etc. Fourthly, stop-words, i.e., frequent words with no sentiment value, are removed from the text. And lastly, stemming removes the suffix of the words and only keeps their base, e.g., “hates” and “hateful” both get assigned to “hate”. The alternative approach is using pre-made models, packages or libraries that come with universal tools for data processing. Such as Python-based NLTK and TextCL (Petukhova & Fachada, 2022; M. Wang & Hu, 2021). The actual combination of techniques depends on a case-by-case basis, as it relies on the type of data and the type of analysis.

Drus & Khalid (2019) note that sentiment analysis is done via two main methods. The first is the lexical approach, which uses a predetermined sentiment lexicon to score the polarity of each word and hence the overall text. VADER is among the more popular lexicons constructed by Hutto & Gilbert (2014) specifically for social media sentiment and is used in a plethora of research (Huynh et al., 2021; Long et al., 2023; Reichenbach & Walther, 2023; Wang & Luo, 2021). The second common approach is machine learning (ML). Authors tend to use out-of-the-box algorithms (Broadhurst et al., 2022; Chen & Ma, 2024; Derakhshan &

Beigy, 2019; Duz Tan & Tas, 2021; Li et al., 2017; Mudinas et al., 2019; Pagolu et al., 2016; Ranco et al., 2015; Sprenger et al., 2014), or use ML as a basis for custom-made models (Awan et al., 2021; Nguyen et al., 2015; Nguyen & Shirai, 2015; Piñeiro-Chousa et al., 2023). The next evolution of approaches is based on Natural Language Processing (NLP), where a program uses linguistic models and statistics to train itself to understand a given language or text. Biswas et al. (2020), Lindskog & Serur (2020), and Piñeiro-Chousa et al. (2017) develop models that use this technique. There are also out-of-the-box models like BERT, which is developed by Google and can be used to classify social media sentiment (Semenova et al., 2021; Sousa et al., 2019; Wang & Luo, 2021).

3 REDDIT AND SOCIAL MEDIA

3.1 Social Media

Since the start of the digital era, social media has played a key role in transforming the flow of information and communication. The rise of platforms such as Facebook, YouTube and Twitter enabled widespread public discourse on a variety of topics and created a global community consisting of millions of online users. This, in hand, incentivised other entities, amongst them companies, to establish an online presence to capitalise on the ever-growing amount of information available on social platforms. Therefore, today's digital landscape is filled with companies in a race for potential customers' attention. But before we dive into the current state of affairs, let us explore what led to social media as we know it today.

Today, the term is mostly associated with social media platforms, e.g. Twitter, Facebook, Instagram, etc. However, according to Carr & Hayes (2015), older definitions of social media refer to it as a group of digital technologies focused on user-generated content or interaction. They also find that there is a lack of consistency throughout the academic community when it comes to defining what social media is. Therefore, they redefine the term as internet-based channels for mass communication that facilitate interactions between users and thus provide a user-generated value. Their observation about inconsistency seems to be true when we look at other authors' definitions. Sajithra & Rajindra (2013) define it as an extension and explosion of traditional networks for communication with the help of technology. Kaplan & Haenlein (2010) acknowledge that the term is too broad and split it into six categories based on social presence and self-presentation. These are Blogs, Social networking sites (e.g., Facebook), Virtual social worlds (e.g., Second Life), Collaborative projects (e.g., Wikipedia), Content communities (e.g., YouTube) and Virtual game worlds (e.g., World of Warcraft (WoW)). Even though not all of these are relevant today, they show us the broad spectrum of what can be regarded as social media. This is to be expected since the ever-evolving landscape of the internet constantly changes our perception of what social media is. For example, a few decades ago, people would associate the term with blog sites and various forums, while today we think of platforms such as TikTok and Instagram.

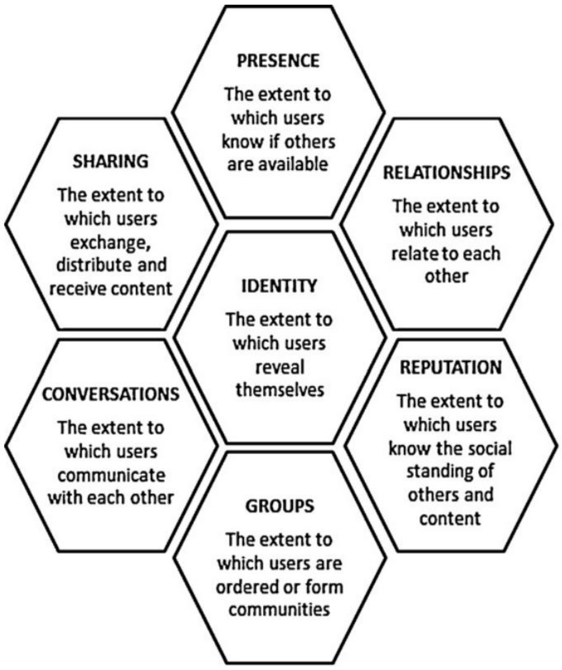
Therefore, to understand what social media is, we must first look at how it evolved in the past few decades. Kaplan & Haenlein (2010) place the beginning of its evolution in the late 1960s when the first online diaries, now known as blogs, began to form on the internet. A few years later, in 1971, Email was born, and with it, the ability to communicate online became an essential part of all future social media platforms (Sajithra & Rajindra, 2013). The rapid expansion of internet infrastructure and access in the next decades gave birth to the first personal websites in the 1990s and the first social networking sites such as MySpace and Facebook in the early 2000s (Kaplan & Haenlein, 2010; Sajithra & Rajindra, 2013). A new era dawned in 2005 with Web 2.0, which introduced web applications that display interactive information, allowing users to freely collaborate and interact as creators and consumers of user-generated content in a social media setting (Sajithra & Rajindra, 2013).

With new technologies, better infrastructure and wider available access, the early social networking sites grew into massive social media platforms that we know today. Carr & Hayes (2015) identify three principal areas of change that allowed for the growth of these platforms. Firstly, infrastructure changes that were brought on by the rapid diffusion of mobile devices improved the accessibility and convenience of consuming digital media, which resulted in a substantial increase in users. This infrastructure improvement altered the way information is stored, processed, and analysed, which enabled the use of data-driven tools that alter user experiences using a huge amount of data and computing power. Secondly, constant interactions on a massive scale allowed individuals and other entities such as governments, organisations, and companies, to reach millions of users from diverse backgrounds with relative ease. In such cases, the illusion of intimacy was achieved by clever use of wording devised by the social media teams and helped by the algorithms designed to reach the target audiences. This forever changed the landscape of advertising and introduced various ethical and legal issues about the users' privacy on such platforms. Thirdly, users could connect over common interests with ease, which led to the formation of communities within the platforms. Over time, these communities grew and became more organised to the point where one can find a strong online community for virtually any interest. These factors contributed to a dramatic increase in social media usage. Perrin et al. (2015) find that in the decade following the inception of Web 2.0 percentage of internet users who use social media rose from 10 % (in 2005) to 76 % (in 2015). Moreover, the authors found that the increase was even across all demographic groups. To put this into perspective, in 2017, social media was used by 2.73 billion people worldwide, and the number of users has been steadily increasing each year and reached 5.17 billion in May of 2024 (Jo Dixon, 2024a). This rapid evolution of social media has brought us to the current landscape occupied by behemoths such as Facebook, YouTube, Instagram, TikTok, Snapchat and Twitter, which are synonymous with the term social media today.

3.2 Social Media Platforms

As established, social media platforms are an integral part of today's society, with billions of daily users, and represent the mainstream way people consume digital content. As of 2024, the biggest amongst them, Facebook, engaged over 3 billion users per month, followed by YouTube with 2.5 billion and Instagram with 2 billion monthly users respectively (Jo Dixon, 2024b). Meanwhile, even if not the most well-known, Reddit still attracted 1.2 billion monthly users, which was almost double the user base of its closest rival, Twitter (Ceci, 2024; Jo Dixon, 2024b). Even though these platforms are designed for different purposes and appeal to different audiences, the philosophy behind their structure is the same. Kietzmann et al. (2011) identify seven core functional blocks in what they call the honeycomb of social media, as seen in Figure 1.

Figure 1: The honeycomb of social media



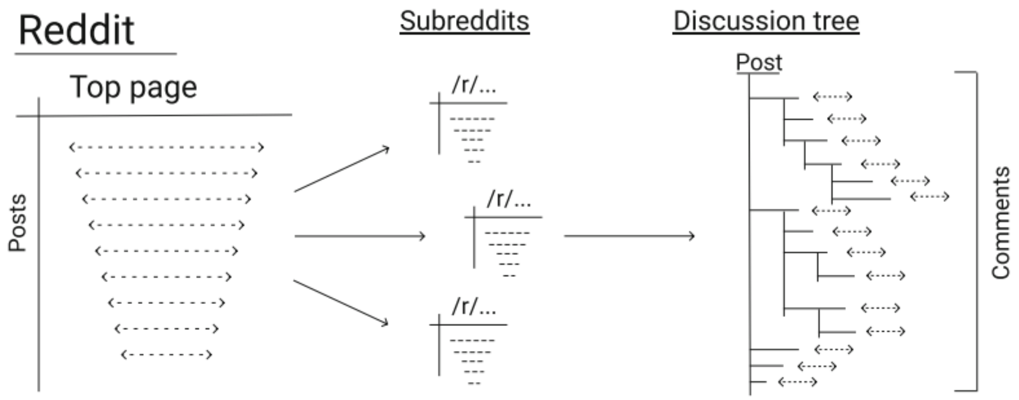
Source: Kietzmann et al. (2011, p. 243).

These building blocks revolve around how users exist, behave, interact and connect on social media platforms. But for this interaction-based framework to function, platforms require a scoring system to rank users' presence and reputation amongst peers. Facebook, Instagram and Twitter allow users to like and repost content to show positive sentiment, while platforms like YouTube and Reddit also allow dislikes to express negative sentiment (Castillo et al., 2014; Khan, 2017). Although the mechanisms to achieve this vary from platform to platform, we can generally divide them into direct and indirect ones. The former are bound to users' posts and comments and usually come in the form of likes, dislikes, reposts, etc., while the latter are bound to users' profiles and reflect their overall standing in the community, which is usually expressed by the number of followers. Now that we have explained the guiding principles behind social media platforms, let us take a closer look at why Reddit is a unique fit for our analysis.

3.3 Why Reddit

As alluded to in the introduction, Reddit differs from other social media platforms. It was founded in 2005 as a simple website that allowed users to post submissions, with the ability to comment being added a year later and the ability to form communities in 2008 (Anderson, 2015). The latter is the factor that differentiates it from other social media sites. Whereas other platforms like Twitter use a single feed system where a user can see posts regardless of their content or category, Reddit enforces a more rigid structure. As explained by Anderson (2015) the platform is divided into interest groups, which are called subreddits and bear the "r/" prefix. These form around users' common interests and span across diverse topics, which allows users to join any community that aligns with their interests. The largest communities boast millions of users. As of 2024, the biggest subreddits r/funny, r/AskReddit, and r/gaming attracted 61 million, 47 million and 42 million users respectively (Reddit, 2024b). Within these subreddits, users can then post and comment within the context of the respective community. To better understand this hierarchy, we can look to Medvedev et al. (2019) who visualised Reddit's structure as we can see in Figure 2.

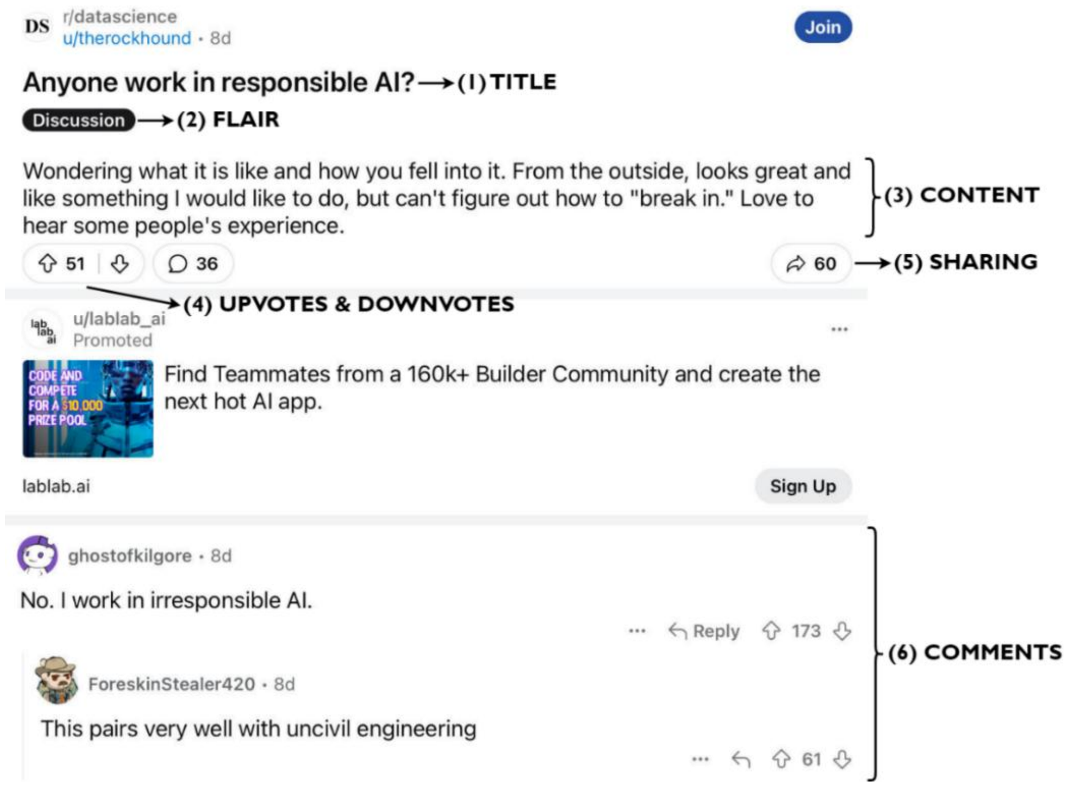
Figure 2: Structure of the Reddit platform



Source: Medvedev et al. (2019, p.185).

The entry point into the platform is a feed of the most popular posts. This varies for a registered and an unregistered user. While an unregistered user only sees posts that the algorithm deems popular across the platform, a register user may switch between that view and only seeing posts from the subreddits they subscribe to. Users are then able to navigate to specific subreddits and see posts that are exclusive to that respective community. These communities are usually well organised, with sets of rules for posting enforced by the moderators. Those are users vested with certain powers to keep the posts in compliance with the respective subreddit's rules. The posts themselves have a structure to them as well. As highlighted in Figure 3, there are three main parts: title (1), flair (2), which represents categories of discussion within subreddits, and content (3), which may contain text, images, videos and hyperlinks.

Figure 3: Structure of a Reddit post



Source: Adapted from Reddit (2024c).

Posts also allow users to interact with them. As highlighted in Figure 3, the main ways to do so come in the form of upvotes and downvotes (4), sharing (5) and comments (6). The latter are organised in discussion trees, i.e., a user can post a comment on a comment, which can create multi-level discussion trees on posts. In a similar fashion to posts, users may also upvote and downvote specific comments to show their sentiment. The combination of all upvotes and downvotes that a user receives on their posts and comments then counts towards their "karma", which represents the overall popularity or unpopularity of the user's opinions on the platform. This is the first part of the answer to our original question. Why Reddit? Because by making use of the platform's scoring mechanism, not only can we extract users' sentiment from the textual data, but we can also weigh the significance of the expressed sentiment using upvotes and downvotes.

This makes Reddit an excellent source of sentiment data for a variety of fields, including financial markets, upon which we will expand later while talking about the link between social media and stock markets. Furthermore, Reddit's structure allows us to filter out the noise by targeting specific communities via subreddits. As pertains to this research, we will be focusing on video game communities, the strong presence of which makes up the second part of our answer. With the social aspect representing one of the key factors for success of a video game, the link between this industry and social media has been a constant since the start of the internet era in the 1980s (Bankov, 2019). This strengthened substantially with the release of the first massively multiplayer online (MMO) games, such as WoW in the early 2000s, which incited millions of players to form online communities (Carr & Hayes, 2015). The growth of online communities was therefore linked to the growth of video game users, the number of which reached 2.58 billion in 2024 (Clemet, 2024). Of course, Reddit is not the only platform used by these communities, as noted by Bergstrom & Poor (2021), who give Usenet, AOL and Twitter as examples of more notable platforms used during the past decade. This is expanded upon by Fiesler & Dym (2020) and Puente & Tosca (2013) who find that while communities constantly migrate between platforms, they also tend to maintain a presence on multiple platforms at the same time. This is supported by Clement (2022) who surveys gamers (self-proclaimed term for video game players) in the U.S. and finds that both Twitter and Reddit attract over 20 % of respondents respectively, while platforms like YouTube, Facebook and Instagram attract over 40 % of respondents respectively which indicates an overlap in user bases. While Reddit may not be the biggest platform that gamers use, it is by no means an irrelevant one. The overarching subreddit r/Gaming represents the third-largest community on the platform, boasting 43 million members, noting that more than twenty game-specific subreddits have over a million members (Reddit, 2024d). Moreover, cultivating strong online communities is also in the interest of the video game developers themselves. As noted by Ruggles et al. (2005), video game developers believe that a thriving online community contributes to a successful video game, as it allows the developers to stay in contact with consumers and thus obtain a constant feedback channel. Furthermore, it plays a significant role in marketing as it allows companies to raise brand awareness, increase customer loyalty and reach new customers (Wawrowski & Otola, 2020). In the case of Reddit, the main way of communication is via the moderators of subreddits. These are either notable community figures or employees of the video game companies. Game developers can therefore survey the users' opinions and organise community events such as Ask Me Anything to connect with the community. Let us now put this relationship into the broader context of the video game industry.

4 VIDEO GAME INDUSTRY

4.1 About the video game industry

Before we dive into the intricacy of the player-game-developer relationship, we should get broadly familiar with the video game industry. Wolf (2012) divides the early history of the industry into pre-crash and post-crash, which occurred in 1983 and 1984 and saw the at the time industry giants experience heavy losses (Atari) or exit the market altogether (Mattel). In the pre-crash period, video games were played on arcade machines, with the golden age in the 1970s, when releases such as PONG, Breakout and Space Invaders were dominating the markets. Arcades peaked in the early 1980s, reaching 400.000 street locations with 1.5

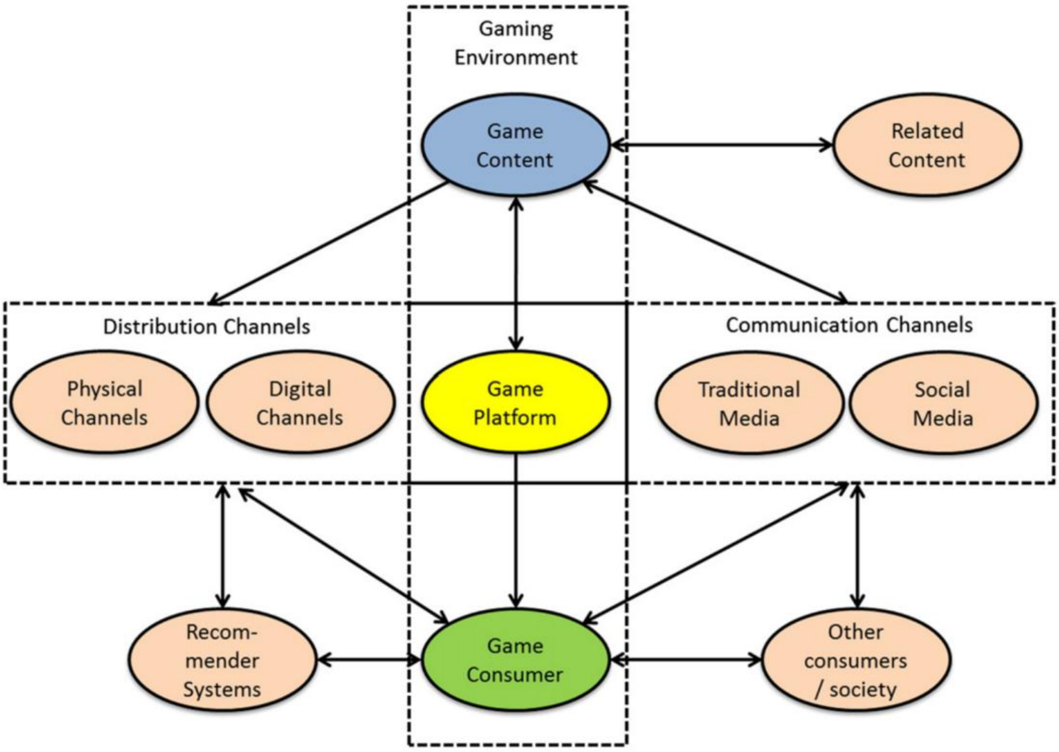
million video game copies in circulation. With rapidly evolving technology, devices became increasingly smaller, which led to the rise of home game consoles that completely overthrew the arcades by the mid-1980s. The post-crash landscape was defined by Nintendo's NES, which paved the way for the release of the consoles we know today, e.g., Sony's PlayStation and Microsoft's Xbox. In the following decade, the introduction of 3D graphics and the evermore-used World Wide Web demanded better and better hardware to run the games. This steered the industry towards personal computer (PC) video games and led to a consolidation of the home console market, which by the 2000s was shared only between Nintendo, Microsoft, and Sony (the three players that dominate it to this day). The introduction of the PC to the video game industry also had a second and much more influential effect. As explained by Zackariasson & Wilson (2010), this separated the hardware and software aspects from the development point of view, which allowed independent developers to make games that all PCs could run. This shift gave birth to game studios such as Blizzard, Activision, Electronic Arts, Ubisoft, etc., which are the titans of the industry today. Blizzard, being the most influential amongst them, changed the name of the game in 2004 with the release of WoW, a MMO game that shifted the distribution of games to online and championed the subscription-based payment as opposed to a one-time purchase. Wallach (2020) describes this era as the online boom, as most industry-defining releases heavily incorporated the online aspect. The most notable are League of Legends (2009), Minecraft (2010), Grand Theft Auto V (GTA V) (2013) and Fortnite (2017) as it pertains to PC and console games. Meanwhile, the release of Angry Birds (2009) and Candy Crush (2012) caused another shift, as the market rapidly expanded into the mobile territory, which by 2020 was bigger than the PC and console markets combined.

In a 2023 market analysis by Statista (2023), the video game market reached a revenue of $406 billion with hardware sales ($161 billion), in-game advertising ($94 billion) and mobile games ($89 billion) being the main sources of revenue, eclipsing the sales of games online ($26 billion). The leading companies by revenue in Q3 of 2023 are Tencent ($30 billion), Sony ($18 billion), Microsoft ($15 billion) and Apple ($11 billion). Whereas this paints a picture of the state of the industry today, only a subgroup of these companies is suitable for our research. This is due to different revenue streams, which in turn affect the stock price. Since we are interested in the connection between the sentiment expressed by communities and the stock price, the primary income stream of these companies must be based on their games. In the Q1 2023 market capitalisation analysis by Clement (2024d), the leaders of this segment were Activision-Blizzard ($67 billion), Electronic Arts ($35 billion), Roblox ($24 billion) Take-Two Interactive ($21 billion). Now that we have become familiar with the industry, we can further explore what makes it unique from the viewpoint of research.

4.2 The Customer-Game-Developer Relationship

The uniqueness of the video game industry stems from the relationship between customers, games and developers. The customer-developer relationship is in part facilitated by social media platforms, which allow for a constant loop of feedback to be funnelled directly to the game developers. This connection is a part of the value creation chain in the video game industry, as visualised by Marchand & Hennig-Thurau (2013) in Figure 4.

Figure 4: Value creation in the video game industry



Source: Marchand & Hennig-Thurau (2013, p. 142).

Hence, the opinions of users shape the state of the games over time as customer feedback is an integral part of a successful video game (Ahmad et al., 2017). Moreover, these customers can be considered genuine experts in the field as they have years of experience playing a variety of video games and are an important source of information for the developers (Burger-Helmchen & Cohendet, 2011). The importance of engaging these players is expanded upon by Yee (2006), who lists socialising, relationships and teamwork as the main components of player motivation in games. This is reaffirmed by Yee (2016) when identifying social collaboration and competition as one of the main clusters of gaming motivation. Therefore, it is in the best interest of the developers to cultivate a strong community of players. Ruggles et al. (2005) find that companies benefit from supporting user-made communities and should encourage effective interaction with the players as well as involve them in the design and development decisions, ensuring that players are satisfied with the overall quality of the game. This, as noted by Yee (2016), is directly connected to satisfying the need for action, mastery, achievement, creativity and immersion, which represent the main clusters of motivation for playing video games. Therefore, the customer-developer relationship is inherently connected to the customer-game relationship. If the state of the game is lacklustre, it has an immediate negative effect on how players will perceive the developers and vice versa. Furthermore, Johnson et al. (2016) find that overall satisfaction is a strong predictor of time spent playing video games. In addition to this, Souza & Freitas (2017) conclude that the intention to play has a high degree of influence on the intention to pay. The long-term success of a video game is therefore dependent not only on the immediate sales but also on the income brought on by an active user base that engages with the product over a prolonged period. This is due to the monetisation techniques used by the industry, many of which rely on the customers' continuous engagement.

4.3 Monetisation techniques

The monetisation techniques have vastly evolved during the past few decades; long gone are the days of the fixed pricing models where games were considered as a one-time purchase. According to Ivanov et al. (2021), the original sin that brought about the micro-transaction and subscription-riddled utopia of today were the so-called "expansion" packs, which could be bought to unlock additional content and increase the product lifecycle as well as generate extra revenue. The next evolutionary step was that these packs came in the form of downloadable content (DLC) packs of either a cosmetic or gameplay nature, and could be purchased digitally. At the same time subscription-based monetisation was gaining popularity amongst MMO games, which were popularised by the release of WoW in 2004. These types of games laid the foundation for monetisation via in-game currencies and microtransactions, which are the biggest income streams for video games today. Microtransactions brought on the latest revolution of monetisation, shifting the industry towards releasing free-to-play games to grow the user base and then introducing in-game stores with cosmetics, currencies, loot boxes, etc., as a main revenue source. This technique is especially effective for mobile games, which tend to attract younger audiences. Hence, a large part of the video game industry has replaced traditional monetisation with the Games as a Service (GaaS) model, which relies on small but constant purchases from a larger base of consumers (Vaudour & Heinze, 2020). According to Ivanov et al. (2021) and Vaudour & Heinze (2020) the main monetisation types in this new landscape are:

- Standard product price is the base amount a customer pays for the base game. Most triple-A titles (main releases of major video game companies) were priced at 60 €/$60. On the other hand, the base game can also be free, which is then offset via other revenue streams.
- DLC packs are digital additions to the base game. These are divided into gameplay packs, which add new playable content to the base game, and non-gameplay packs, which usually include cosmetics. Their main intent is to prolong the game's lifecycle and attract new players.
- Subscription services are used in online games and represent a monthly payment (around 20 €/$20) which allows the user to play the game for that period. They are mainly used for MMO games that rely on a consistently large player base.
- Season passes are a combination of DLC packs and subscriptions. A customer makes a single payment per season (usually 3-4 months) and gets access to all DLC content

released within that period. They are mainly offered in free-to-play online games and mostly include items and cosmetics.

- In-game currencies are virtual currencies that can be purchased and spent in-game. They enable players to access premium content, e.g., loot boxes, items, cosmetics, etc., inside the basic game.
- Loot boxes contain randomised standard and premium content. The customer does not know what the content will be until after the purchase. This is considered to be the most predatory monetisation tactic, since a plethora of research (King & Delfabbro, 2018; Zendle & Cairns, 2019) connects it to developing a gambling addiction, leading some countries, such as Japan, Belgium, and the Netherlands, to impose partial bans on their sales (Xiao, 2022).
- Microtransactions is an umbrella term for all manner of small purchases which a game offers. The three main types are in-game currencies, random chance purchases (loot boxes) and in-game items. They are present in almost every newly released game and make up the bulk of the revenue stream for video game developers.
- In-game advertising comes in many forms, the most used methods being pop-up ads, loading screen ads, ads to earn in-game currencies, etc. Mobile games usually offer an option for a one-time purchase to hide all ads. This technique is mostly present in the mobile market but can also be seen in some free-to-play PC video games.

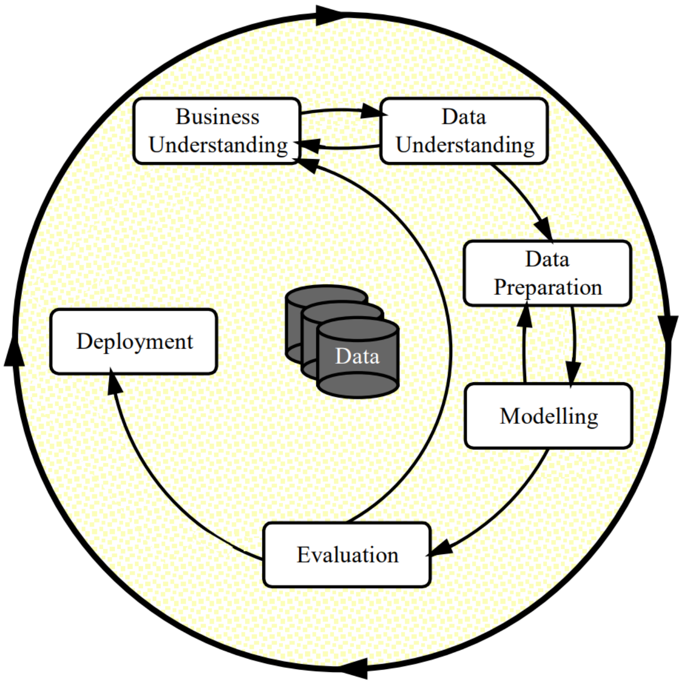
The implementation of these techniques varies from game to game, but their deployment is most prevalent in three ways. First, there are games with a fixed price for the standard product, with the addition of DLC packs. Second, online subscription-based games with microtransactions. And third, free-to-play or "freemium" games with season passes and microtransactions. Vaudour & Heinze (2020) find that all of these rely on good customer relationship management to ensure the loyalty and engagement of players. The importance of this is underlined by the conclusions of Ivanov et al. (2021) that link time spent playing the games with the likelihood of spending money via microtransactions. Therefore, the revenue stream is dependent on the long-term satisfaction of players. And since most of these companies are publicly traded, this begs the question of whether we can link the satisfaction of their customers with their success on the stock market.

5 METHODOLOGY

5.1 CRISP-DM

To ensure a clear alignment between our research objectives and the data analysis process, we use the CRISP-DM framework established by Wirth & Hipp (2000). Its flexibility and iterative nature allow us to refine and improve our approaches, which contributes to the overall quality of our analysis. The framework consists of six key phases through which we will form our methodology. While the phases iteratively interact with each other, the sequence of their execution is not strict. The purpose of the model is to emphasise the importance of dependencies across different stages, with the implementation being left to a case-by-case basis.

Figure 5: Phases of the CRISP-DM framework



Source: Wirth & Hipp (2000, p. 5).

In the following, we outline each of the phases seen in Figure 5:

- Business Understanding: We establish the objectives and what we want to accomplish, convert the problem into a data mining context, and form a preliminary plan.
- Data Understanding: First we collect the initial data, which we then explore to better understand and discover any potential quality issues or other shortcomings.
- Data Preparation: This stage is the most important and time-consuming since we prepare the whole dataset for the final analysis. We determine which data will be used, clean it, and transform it.
- Modelling: In this phase, we select, test, assess, and apply various modelling techniques which suit our use case.
- Evaluation: Before we proceed to the final deployment of our models, we must first assess them. This entails evaluating the results to see if we meet the criteria defined at the start, while also reviewing the process to spot if we overlooked any crucial steps.
- Deployment: Depending on the requirements, this phase can be both simple and complex. It generally consists of deploying our model, monitoring and maintaining it throughout the deployment, then producing the final report and reviewing our work.

5.2 Scope of Research

5.2.1 Selection process

Before we delve into the analysis, we must first define the scale of research. Our scope of research depends on three main factors. The observation time frame, which companies we choose to analyse, and the Reddit communities we link to the chosen companies. As for the time frame, we will focus on the 2020–2023 period. This is to ensure the relevancy of our research, as it deals with recent events. Moreover, we must acknowledge the constraints we have regarding the processing power, which does not allow us to analyse a decade's worth of data. On the other hand, the choice of companies requires more attention. Considering the scarcity of research in this field, there is no universally adopted procedure or framework for the choice of companies. This is done on a case-by-case basis, with most research focusing on specific companies, e.g., papers about GME (as shown in Table 1). Meanwhile, publications covering the broader industry (Piñeiro-Chousa et al., 2023; Xiong & Bharadwaj, 2014), tend to use the top-rated stocks or industry leaders.

This is inadequate for our analysis. Therefore, we must define an approach of our own. The goal is to select up to five companies that capture the variety of the video game industry, while also garnering a large social media community. First, we must select potential candidates. For a company to qualify, it must meet two conditions. Firstly, they must be publicly traded on a stock exchange. And secondly, their revenue must be based on video games. This is in place to ensure that no other revenue streams are affecting the stock prices. Therefore, industry giants such as Microsoft, Tencent, Sony, Nintendo, etc., are not suitable for our analysis, since they all earn a large share of revenue from other sources. The first draft of candidates is based on a list of video game stocks curated by Yahoo Finance (2024a), with the addition of Activision Blizzard (no longer publicly traded due to acquisition by Microsoft in October of 2023). For the complete list of candidates, see Appendix 2.

Now we evaluate these prospects. The main criteria are how well we can link the respective video games to the company revenue and the size of the Reddit communities associated with these games. This is not an easy task, since developers release many games each year with different monetisation models, which makes it hard to link specific games to the revenue generated. Therefore, we use earnings reports and third-party industry overviews to approximate which games generate the most revenue. Besides this, we also account for the type of video games, the relevancy of the video games, and the monetisation techniques used.

As a result, we are left with four companies for our analysis. Three of them are industry giants with different revenue structures (Activision Blizzard, Electronic Arts, and Take-Two Interactive), and one smaller but unique developer who released some of the most successful games in the industry's history (CD Projekt Red). On the other hand, we could not consider NetEase a mobile game giant, since the size of its online communities was too small. The same applies to NEXON, which develops mostly free-to-play online games. Ubisoft, a well-established name in the industry, is excluded because its revenue structure, types of games, and monetisation techniques are very similar to the three industry leaders we did include. Therefore, it does not provide any additional value to our analysis. For an in-depth overview of the candidates, see Appendix 3.

5.2.2 Activision Blizzard

Blizzard has been a staple of the video game industry since the release of WoW in 2004. The status was reinforced by the merger with Activision in 2008, making it the largest video game developer in the world (Activision Blizzard, 2008). The company has since been known for its long-running franchises, such as Diablo, Overwatch (OW) and Call of Duty (CoD). With the somewhat recent addition of Candy Crush Saga, the company added to its portfolio after acquiring King, a leading developer in the mobile market (Activision Blizzard, 2016). The company having been acquired by Microsoft in October of 2023 is irrelevant to our analysis, since it has operated through most of our observational timeframe (Microsoft, 2022). The company generates most of its revenue via live services, i.e., subscriptions, microtransactions, etc., with its games spread across console, PC and mobile platforms Clement (2024a), the latter being the largest by the share of revenue, with PC and consoles also contributing a sizeable share (Clement, 2024a). Hence, this company has diverse revenue streams which are linked to various types of games. The list of chosen games and their communities can be seen in Table 2. One big omission from the list is the Candy Crush Saga, which generates most of the mobile platform revenue, but was excluded due to a small online community.

Table 2: Activision Blizzard games and subreddits

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
World of Warcraft	PC	Subscription, microtransactions	r/wow	2.800.000
Overwatch/ Overwatch 2	PC, Console	Fixed price, microtransactions	r/Overwatch	5.800.000
CoD: MWII	PC, Console	Fixed price, microtransactions	r/ModernWarfareII	4.800.000
CoD: Warzone	PC. Console	Free-to-play, subscription, microtransactions	r/CODWarzone	1.400.000
CoD: Mobile	Mobile	Free-to-play, subscription, microtransactions	r/CallOfDutyMobile	338.000

Source: Adapted from Activision Blizzard (2024); Clement (2024a); Reddit (2024b).

5.2.3 Electronic Arts

Electronic Arts is another giant in the industry. Its portfolio consists of long-running franchises, such as FIFA, Battlefield, The Sims, Madden NFL, and Need for Speed. Like Activision Blizzard, it generates most of its revenue from live services, but differs when it comes to revenue by platform, as consoles are bringing in more revenue than PC and mobile combined. Hence, it makes an interesting company to analyse. While it is similar to Activision Blizzard in the way it makes money, it does so from a completely different segment of consumers. This is reflected in the games we analyse, which are listed in Table 3.

Table 3: Electronic Arts games and subreddits

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
Apex Legends	PC, Console	Free-to-play, Subscription, Microtransactions	r/apexlegends	2.900.000
Battlefield 2042	PC, Console	Fixed price, DLC, Microtransactions	r/battlefield2042	238.000
Madden franchise	PC, Console	Fixed price, Microtransactions	r/Madden	247.000
Star Wars: Galaxy of Heroes	Mobile	Free-to-play, Microtransactions	r/SWGalaxyOfHeroes	140.000

Source: Adapted from Clement (2024e); Electronic Arts (2024); Reddit (2024b).

5.2.4 Take-Two Interactive

Take-Two Interactive is the third industry leader chosen for our analysis. It is mostly known for the Grand Theft Auto and Red Dead Redemption franchises, developed by its subsidiary Rockstar Games. Civilization and NBA 2K franchises are also prominent within their respective genres, developed by the second subsidiary studio 2K. The company's revenue structure by platform is not dissimilar to Electronic Arts, with mobile making a surge in the share of revenue in 2023 as Take-Two Interactive acquired a mobile industry giant in 2022, Zynga (Clement, 2024b; Take-Two-Interactive, 2022). This also boosted the share of revenue from live services, which was almost four times the revenue of product sales in 2023 (Clement, 2024b). This is a contrast to previous years, where the product sales represented about a third of the overall revenue (Clement, 2024b). However, as already established, mobile games do not have a large community presence on social media. Hence, we cannot adequately analyse them. The list of games and their communities can be seen in Table 4.

Table 4: Take-Two Interactive games and subreddits

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
GTA V	PC, Console	Fixed price, Microtransactions	r/gtaonline r/GTAV	1.600.000 476.000
Red Dead Redemption 2	PC, Console	Fixed price, Microtransactions	r/reddeadredemption r/RedDeadOnline	1.900.000 426.000
NBA 2K franchise	PC, Console	Fixed price, Microtransactions	r/NBA2k	569.000
Borderlands 3	PC	Fixed price, DLC, Subscription, Microtransactions	r/borderlands3	413.000

Source: Adapted from Clement (2024b); Reddit (2024b); Take-Two Interactive (2024).

5.2.5 CD Projekt Red

While not an industry giant like Electronic Arts and Activision Blizzard. This developer from Poland released two of the biggest games in the history of the industry, with over 80 million copies sold (CD Projekt Red, 2024). Witcher 3 and Cyberpunk 2077 (CP 2077), both of which are single-player, story-driven games that are sold at a fixed price without any microtransactions or other live service monetisation techniques (CD Projekt Red, 2024). This makes for a unique case for our analysis. Furthermore, these two games alone represent almost all the revenue of this developer (CD Projekt Red, 2024). This contrasts with the bigger names in the industry, whose revenue is based on many different games. In addition, the release of CP 2077 was met with a lot of controversy that was kept alive through much of the game's history, which makes it especially interesting for our case (Frank, 2020).

Table 5: CD Projekt Red games and subreddits

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
Cyberpunk 2077	PC, Console	Fixed-price, DLC	r/cyberpunkgame	1.800.000

Source: Adapted from CD Projekt Red (2024); Reddit (2024b).

5.3 Data Collection

5.3.1 Financial data

We collected both financial and sentiment data for the 2020–2023 period. The financial data consists of the stock prices for the companies included in our scope of research. We acquired the data from the NASDAQ stock exchange and a financial portal, Investing.com (for the Activision Blizzard stock since it is no longer publicly traded, and the CD Projekt Red stock that is traded on the Warsaw Stock Exchange) (Investing.com, 2024a, 2024b; NASDAQ, 2024).

5.3.2 Sentiment data

The sentiment data is collected from Academic Torrents, an open-source repository of datasets for research (Cohen & Lo, 2014). From here, we acquire the comment and submission data for the chosen subreddits (see Table 1, Table 2, Table 3, Table 4, and Table 5), which are a part of the *Subreddit comments/submissions 2005-06 to 2023-12* dataset (Watchful1, 2024). This dataset is part of a broader project of the Pushshift Reddit Dataset by Baumgartner et al. (2020), which contains a variety of regularly updated Reddit-based datasets. This now serves as a successor to a widely used Pushshift API, which served as a main data source for Reddit-related research (Corbet et al., 2022; Huynh et al., 2021; Mancini et al., 2022; Reichenbach & Walther, 2023; Semenova et al., 2021; Wang & Luo, 2021). The end of the Pushshift API use came when the pricing change enforced by Reddit in 2023 ended its viability for public use (Reddit, 2024a).

5.4 Data Preparation

5.4.1 Financial Data

To compensate for the different formats and to clean the data, we develop a Python script with which we standardise the data from all sources into a universal format. Apart from the formatting issues, there is another problem we must address. Because the financial markets do not operate on weekends and holidays, we must estimate those prices. French (1980) implies that the prices over the weekend tend to fall as the information released tends to be unfavourable. However, this cannot be taken for granted in all instances. Since the stock prices usually follow a concave function, we follow a basic approach by Mittal & Goel (2012). We fill in the missing values (p_t) with the average price of the closing price of the first available preceding day ($p_{CLOSE\ (t-n)}$) and the opening price of the first available following day ($p_{OPEN\ (t+m)}$).

$$p_t = \frac{p_{CLOSE\ (t-n)} + p_{OPEN\ (t+m)}}{2} \tag{1}$$

As our methods of analysis (correlation and Granger's causality) require stationary time series, i.e., the series cannot follow any trends, we must augment our data, since stock prices tend to follow long-term trends. In line with common practice in financial applications (see Behrendt & Schmidt (2018), Corbet et al. (2022), Ranco et al. (2015), and Sprenger et al. (2014)), we do so by calculating stock price returns:

$$r_t = \frac{p_t}{p_{t-1}} - 1 \tag{2}$$

with r_t denoting the return of the day and p_t the closing price of an asset at a time t.

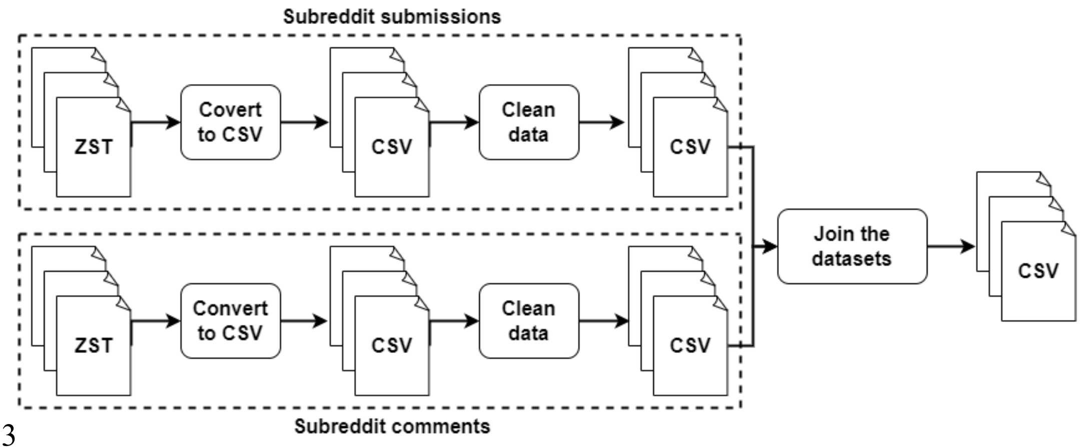
5.4.2 Sentiment Data

Our approach to data preparation depends on the type and quality of data, while also considering the type of analysis we will conduct. Since we will be using the VADER package for the sentiment analysis and the RoBERTa model for emotion classification, we skip most of the conventional steps noted by Chai (2023), such as normalisation, tokenisation, stop-word removal, and stemming. These packages evaluate sentiment based on the whole text, which includes elements such as punctuation, capitalised words, etc. Hence, we keep these so as not to lose any sentimental value. This is in fashion with related research (Huynh et al., 2021; Long et al., 2023; Reichenbach & Walther, 2023; Wang & Luo, 2021). Our approach to cleaning the data also depends on its characteristics. Therefore, we must first extract it from its raw form. For this, we modified an existing Python script by Baumgartner (2024) to convert all our data into comma-separated values (CSV) files, keeping only the relevant attributes listed in Tables 9 and 10. As we peruse the data, we can identify the following elements that have to be dealt with:

- Embedded links, images and videos come in the form of various types of hyperlinks, e.g., "[https://www.reddit.com]".
- Bad encoding and decoding of some symbols, e.g., apostrophe presenting as "â€™".
- Hypertext Mark-up Language (HTML) elements in the text, e.g., "&".
- Line breaks and white spaces.
- Deleted or removed text elements, which are marked as "[deleted]", "[removed]", and "[removed by user]".

To tackle these issues, we develop Python scripts to clean and process the data as outlined in Figure 6. We do this separately for the submissions and the comments dataset. For both datasets, the core logic of the text cleaning is the same. First, we convert the UNIX timestamp into a conventional date format and cut off all the data not within our time frame. Then we remove hyperlinks, embedded images and videos, HTML tags, line breaks, whitespaces, numbers, and stray non-alphabetical characters. Now that the text is clean, we must fix the encoding issues that stem from different formats of source data. And lastly, we remove all rows where there is no textual data, or the text values are insignificant, e.g. only non-alphabetical text. We then join both datasets, so we have one CSV file per subreddit.

Figure 6: Data cleaning process



3

Source: Own work.

5.5 Data Analysis

5.5.1 Sentiment analysis

In line with Huynh et al. (2021), Long et al. (2023), Mancini et al. (2022), Reichenbach & Walther (2023), and Wang & Luo (2021), we use VADER as our tool for sentiment analysis. Developed by Hutto & Gilbert (2014) it is a sentiment package that is purpose-made for analysing social media text. The advantage of using this approach is threefold. Firstly, it is easily implemented with a corresponding library available in Python. Secondly, it does not require substantial text preprocessing, as it is designed to consider the structure of text. And thirdly, it does not require any training data, since it uses a human-validated sentiment lexicon which includes all the rules related to grammar and syntax. In addition to the stock lexicon, we include some custom words and phrases from the slang used in the gaming community, e.g., buggy ‘containing many bugs’, which relates to performance issues. As explained by Hutto & Gilbert (2014) VADER evaluates the words with valence scores ranging from [-4] as the most negative, [0] as neutral, and [+4] as the most positive. Based on these, it then provides four rankings for the text: negative, positive, neutral and composite. The latter is calculated from the former three and represents the overall sentiment of the text, with [-1] as the most negative, [0] as neutral, and [+1] as the most positive. When the compound score is greater than 0.05, it denotes a positive sentiment, a score between 0.05 and -0.05 denotes a neutral sentiment, and a score below -0.05 denotes a negative sentiment. Hence, it provides us with a clear metric of the sentiment in our data. We then apply this to our data to calculate the daily sentiment for each of the subreddits. We do this by taking the composite sentiment score of respective comments or posts (S_n) and calculating the average for each day (S_sub).

$$S\_sub = \frac{S_1 + \cdots + S_n}{n} \tag{3}$$

Furthermore, we also calculate the daily sentiment by company (S_comp), which is defined as an average of the sentiment scores of posts/comments from all the subreddits associated with the company ($S_{k,i}$)

$$S\_comp = \frac{\sum_{k=1}^{x} \sum_{i=1}^{n_k} S_{k,i}}{n} \tag{4}$$

where x is the number of different subreddit series and n_k is the length of the k-th series. As with the financial dataset, we must ensure the stationarity of the sentiment data. Hence, we take a first difference, i.e. we calculate the change between one observation and the next. We apply this to both the daily sentiment by subreddit as well as the daily sentiment by company.

$$S\_sub\_diff = S\_sub_t - S\_sub_{t-1} \tag{5}$$

$$S\_comp\_diff = S\_comp_t - S\_comp_{t-1} \tag{6}$$

5.5.2 Emotion classification

In addition to conventional sentiment analysis, we also perform emotion classification. This approach allows us to identify the basic emotions in the text, such as anger, joy, disappointment, etc. We can achieve this by using emotion classifiers, which are libraries or models that gauge the text and tell us how strong the presence of each emotion is. In our case, we use a model based on BERT, a LLM developed by Devlin et al. (2018), which is commonly used for both sentiment analysis and emotion classification (Rahman et al., 2023; Semenova et al., 2021; Sousa et al., 2019; Wang & Luo, 2021). The model we use was developed by Antypas et al. (2023) and was trained based on 154 million tweets using RoBERTa, which is a BERT-based pretraining approach by Liu et al. (2019). Its advantages can be likened to those of VADER. It classifies emotions from the whole text (preprocessing of text can therefore be largely skipped), does not require any training data, and is relatively easy to use. The model classifies eight emotions: fear, joy, love, optimism, pessimism, sadness, surprise and trust. For all of this, it provides a score ranging from [-1] as the most negative to [+1] as the most positive. To implement this, we use a Python script based on Antypas et al. (2023). Consistent with the logic applied to sentiment analysis scores, we calculate the daily scores for each emotion (Es_sub) as an average of the respective emotion scores of posts/comments (E_n) on a given day.

$$E\_sub = \frac{E_1 + \cdots + E_n}{n} \tag{7}$$

Additionally, the daily score for a respective emotion by a company (E_comp) is calculated as an average of the emotion scores of posts/comments from all the subreddits associated with the company ($E_{k,i}$)

$$E\_comp = \frac{\sum_{k=1}^{x}\sum_{i=1}^{n_k} E_{k,i}}{n} \tag{8}$$

where x is the number of different subreddit series and n_k is the length of the k-th series. Like with the sentiment data, we also calculate the first difference for both the emotion by subreddit and the emotion by company series.

$$E\_sub\_diff = E\_sub_t - E\_sub_{t-1} \tag{9}$$

$$E\_comp\_diff = E\_comp_t - E\_comp_{t-1} \tag{10}$$

5.5.3 Stationarity test

Before we can proceed to our analysis, we must establish the stationarity of our time series. As explained by Shrestha & Bhatta (2018), the stationarity of time series is crucial for practically all analysis methods. Therefore, we apply the ADF test, which is defined as:

$$\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \sum_{i=1}^{p}(\delta_i \Delta y_{t-i}) + e_t \tag{11}$$

where α is a constant, β the coefficient of the time trend, p the lag order of the autoregressive process, and e the residual (Holmes et al., 2021). The null hypothesis is that the data are non-stationary. We test this by the unit root test, in which the null hypothesis $H_0: \gamma = 0$ is tested against the alternative hypothesis $H_1: \gamma < 0$. If we do not reject the null hypothesis (p-value < 0.05), the series is stationary, whereas rejection (p-value > 0.05) means that the series is non-stationary.

5.5.4 Local Events Identification

Before we delve into more advanced methods, we first identify local events in our datasets where the volume of submissions is abnormal. The importance of this is twofold. We get a better grasp of our data. And more importantly, we can spot any local surges of activity in our sentiment dataset. The latter is an integral part of our overall approach. Since findings about the stock-sentiment connection are inconsistent (Behrendt & Schmidt, 2018; Chen et al., 2014; Ranco et al., 2015), we will take a closer look at significant local events in our data, in line with Chahine & Malhotra (2018) and Li et al. (2017).

The process is split into two steps. Firstly, we identify events on the level of a subreddit (game) by visualising and analysing the dataset containing the volume of daily submissions. Additionally, the identified events are then cross-referenced with the events in the industry (see Appendices 5–8). And secondly, we determine which of these events are significant on the level of the company. This will be assessed on a case-by-case basis to best represent different types of local events as well as ensure that they are relevant within the context of the respective company. The correlation and Granger's causality test will then be performed for the entire period, as well as the smaller timeframes surrounding the selected local surges (minimum of 15 days before and after an event, to ensure an adequate sample size). Thus, we will be able to determine the nature of the connection for both local events (e.g., game releases, expansion announcements, game updates) as well as the broader timeline.

5.5.5 Correlation

Now that we have calculated the sentiment and emotion classification scores, we link them to the stock price movements. For an initial investigation, we use basic Pearson's correlation as well as cross-correlation to get the first hint if there is any connection between the sentiment and stock data. The former calculates the Pearson coefficient (r_{xy}) by substituting estimates of variances and covariances based on the following formula

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{12}$$

where n is the sample size, x_i and $y_i$are individual samples indexed with i, and $\bar{x}$, $\bar{y}$ are the means of the respective samples.

Since sentiment may not be instantaneously reflected in the stock price movements. We also look at how the two series correlate at different time lags. With no consensus among researchers (Chahine & Malhotra, 2018; Li et al., 2017; Sul et al., 2017) for how long it takes the sentiment to diffuse, we use cross-correlation. We follow suit of Bu & Pi (2014), Guo et al. (2017), Sun et al. (2017), and Xu et al. (2021), who use this technique to identify connections between stocks and sentiment. This approach measures the similarity of two series at different time lags, which tells us if one series (e.g., sentiment) could impact the other series (e.g., stock price returns). The correlation z of two time series x and y is defined as

$$z[k] = (x * y)(k - N + 1) = \sum_{l=0}^{||x||-1} x_l y_{l-k+N-1}^{*} \tag{13}$$

for $k = 0, 1, \ldots, ||x|| + ||y|| - 2$ where $||x||$ is the length of x, $N = \max(||x||, ||y||)$, and y_m is 0 when m is outside the range of y.

5.5.6 Granger causality

As the final step of our analysis, we calculate the Granger causality between our datasets, a common approach when analysing effects on stock prices (Long et al., 2023; Mudinas et al., 2019; Ranco et al., 2015; Sun et al., 2017). Conceptualised by Granger (1969) Granger causality allows us to determine whether one time series can predict another. Contrary to its name, it cannot determine true causality, i.e., in the sense of regression; nevertheless, it can tell us if one series predicts the other. The key idea is that if a time series Y Granger causes time series X, then the past values of Y provide information that helps predict the future value of X beyond what could be predicted from past values of X alone. In layman's terms, Y has predictive power over X. The assumption is that both series remain stationary, i.e., their variance is constant over time. To ensure all requirements are met, we follow the theoretical steps of Ranco et al. (2015) and implement them in Python in line with Borja et al. (2020).

With the stationarity of our series already tested, we build and fit a Vector Autoregressive (VAR) model. These models are used to capture relationships between variables and their lagged values. They differ in complexity and applications, but for our purposes, we use the VAR(p) model for order p defined as

$$y_t = a_0 + \sum_{j-1}^{p} A_j y_{t-j} + e_t \tag{14}$$

where y_t is a $N \times 1$ vector of N endogenous variables, a_0 is a $N \times 1$ vector of constants, $A_1 \dots A_P$ are the p $N \times N$ matrices of autoregressive coefficients, and e_t is a $N \times 1$ vector of residuals. We first split the data into training and test sets and select the VAR order p by computing multivariate information criteria (AIC, BIC, HQIC, and FPE), according to which we choose the value of p to fit the model with. Then, we can finally perform the Granger causality test

$$X_t = \alpha + \sum_{i=1}^{p} \beta_i X_{t-i} + \sum_{i=1}^{p} \gamma_j Y_{t-j} + \epsilon_t \tag{15}$$

where X is an independent variable, α the intercept, $\sum_{i=1}^{p} \beta_i X_{t-i}$ the sum of lagged values of X, capturing its past influence, $\sum_{i=1}^{p} \gamma_j Y_{t-j}$ the sum of lagged values of Y, capturing its potential influence on X and ϵ_t is the error. This test checks whether knowing past values of Y helps predict X better than just using other variable information. Specifically, if the variance of prediction errors for X is smaller when including Y compared to when excluding it, we say that Y "Granger-causes" X, denoted as $Y_t \Rightarrow X_t$.

6 RESULTS

6.1 Data description

6.1.1 Financial data

The financial data is comprised of the CSV files for the respective stocks. Due to the different data sources, i.e., Investing.com and NASDAQ, the format of the data differs as displayed in Tables 6 and 7.

Table 6: Description of the NASDAQ dataset

Field	Data type	Description
Date	Date	Date in the MM/DD/YYYY format.
Close/Last	String	Close price of the day, e.g., $141.5.
Volume	Integer	Volume of traded stocks of the day, e.g., 10000.
Open	String	Open price of the day, e.g., $140.6.
High	String	High price of the day, e.g., $141.1.
Low	String	Low price of the day, e.g., $139.6.

Source: Adapted from NASDAQ (2024).

Table 7: Description of the Investing.com dataset

Field	Data type	Description
Date	Date	Date in the MM/DD/YYYY format.
Price	Float	Close price of the day, e.g., 141.5.
Open	Float	Open price of the day, e.g., 140.6.
High	Float	High price of the day, e.g., 141.1.
Low	Float	Low price of the day, e.g., 139.6.
Vol.	String	Volume of traded stocks of the day, e.g., 164.25K.
Change%	String	Percent change in price to the previous day, e.g., 0.5%.

Source: Adapted from Investing.com (2024a) and Investing.com (2024b).

To standardise the data, we then applied our processing steps to arrive at the standard format shown in Table 8.

Table 8: Description of the processed financial dataset

Field	Data type	Description
date	Date	Date in the YYYY-MM-DD format.

To be continued

Table 8: Description of the processed financial dataset (cont.)

open	Float	Open price of the day, e.g., 141.5.
close	Float	Close price of the day, e.g., 149.5.
return	Float	Calculated stock price return.

Source: Own work.

6.1.2 Sentiment data

The collected data is in the form of Zstandard compressed files (ZST), which contain compressed submissions and comments of the respective subreddits. For each subreddit, there are two files, one for comments and one for submissions, e.g., files for the subreddit r/wow are wow_submissions.zst and wow_comments.zst. Each contains the data for all submissions and comments made up to the end of 2023. As documented by Baumgartner et al. (2020), the structure of the two datasets differs (for the full data description, see Appendix 4). Nevertheless, for our research, this is not that impactful, since we are only interested in sentiment-related data. Namely, the title and content of submissions and the content of comments. Hence, we outline the relevant fields in Table 9 for the Submissions dataset and in Table 10 for the Comments dataset. The notable omission from both is a score metric. While accounting for the number of likes and dislikes for each submission and comment could provide additional value, it is not included in any related research (see Table 1). Therefore, we will not weigh the sentiment as there is no apparent indication that it improves how well it is captured.

Table 9: Description of the Pushshift Submissions dataset

Field	Data type	Description
id	String	The submission's identifier, e.g., "5lcgjh".
created_utc	Integer	UNIX timestamp that refers to the time of the submission's creation, e.g., 1483228803.
subreddit	String	Name of the subreddit the submission is posted on. Note that it excludes the prefix /r/. E.g., 'AskReddit'.
selftext	String	The text of the submission.
title	String	The title of the submission.

Source: Adapted from Baumgartner et al. (2020, p.833).

Table 10: Description of the Pushshift Comments dataset

Field	Data type	Description
id	String	The comment's identifier, e.g., "dbumnq8".

To be continued

Table 10: Description of the Pushshift Comments dataset (cont.)

created_utc	Integer	UNIX timestamp referring to the time of the submission's creation, e.g., 1483228803
subreddit	String	The name of the subreddit of the submission is posted. Note that it excludes the prefix /r/. E.g., 'AskReddit'
body	String	The comment's text, e.g., "This is an example comment".

Source: Adapted from Baumgartner et al. (2020, p.834).

As we apply our processing steps to both datasets, we get a standardised CSV file for each of the subreddits within the scope of our research. The structure of the file can be seen in Table 11.

Table 11: Description of the processed sentiment dataset

Field	Data type	Description
id	String	The identifier of a submission/comment.
date	Date	The date a submission/comment was posted in a YYYY-MM-DD format.
text	String	For a comment, this is the same as the body field. For a submission, the title and the selftext are combined in a single field.

Source: Own work.

6.2 Data analysis

To fully grasp the connection between the stock returns and sentiment data, we analysed it from two viewpoints. The first one tackles the length of observations, i.e., whether we study the overall timeline or short, local periods. Meanwhile, with the second one, we distinguish between the connections at the level of the whole company and those at the level of a single game or subreddit. Thus, as these two viewpoints interact, it is important to clarify some key terms we use to describe our findings.

- Company level – pertains to correlations and causalities for the sentiment and emotion scores calculated for the entire company, i.e., the combination of all the subreddit/game scores.
- Subreddit/game level – relates to correlations and causalities for the sentiment and emotion scores calculated for specific subreddit/game communities, i.e., the combination of scores for those respective subreddits.
- Long-term/overall – references the entire period of our observation.

- Short-term/local - references one or more individual local events we defined (see Tables 12, 14, 16, 18).

6.2.1 Stationarity test

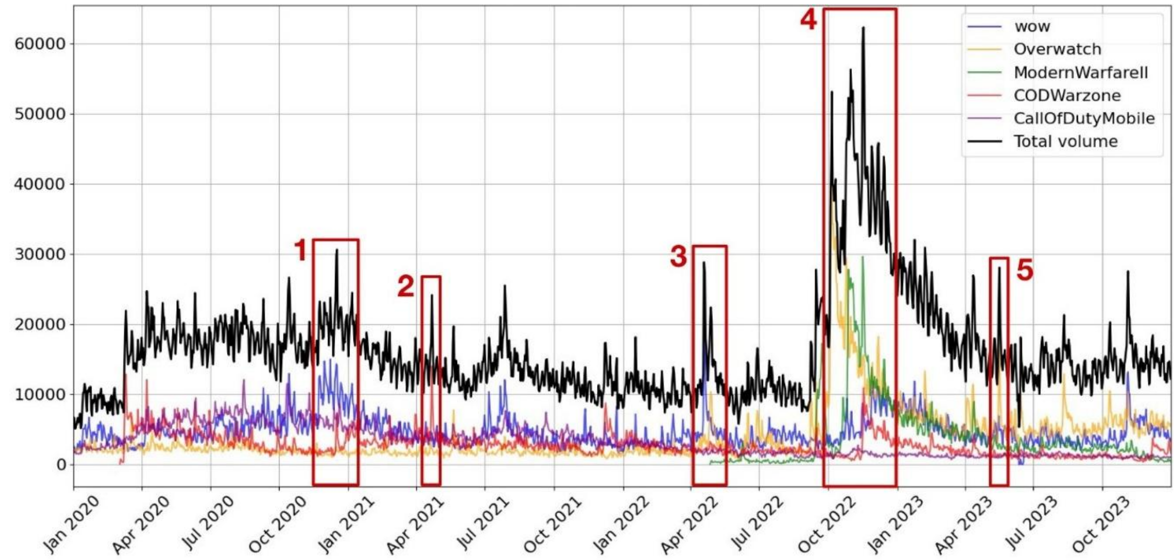
The stationarity tests, detailed in Appendix 9, reveal varying degrees of stationarity across financial, sentiment and emotion datasets. Regarding the financial data, the open and close variables are mostly non-stationary since the stock prices tend to follow long-term trends. Therefore, we use the stock return data in the next steps of our analysis, a common practice in the econometric literature. On the other hand, the sentiment and the emotion datasets have varying stationarity across variables (see Appendix 9). The raw sentiment data exhibits mixed stationarity, with negative and compound being stationary for most series and positive being inconsistent.

However, after differencing the sentiment data, all series achieve robust stationarity ($p<0.01$). This also holds for the emotion data. While certain emotions like anger, joy and disgust are stationary for most series, others, such as love and anticipation, fail to achieve stationarity. After differencing (see Appendix 9), all the variables are stationary ($p<0.01$). Hence, the transformation of the datasets via differentiation ensures that we have time series that exhibit consistent stationarity across all variables, which is crucial for the following analyses.

6.2.2 Activision Blizzard

The case of Activision Blizzard is unique since the publicly traded company was made up of two game studios, Activision and Blizzard. The former champions the CoD franchise, and the latter is making its mark with OW and WoW. The chosen events displayed in Table 12 represent the most impactful releases of these two game studios. With greater detail provided in Appendix 5, the omission of the two mobile games is due to the low volume of submissions, as can be seen in Figure 7. Some high-volume peaks are also omitted, since the events surrounding them are not related to the games but rather events surrounding the companies.

Figure 7: Significant events for Activision Blizzard



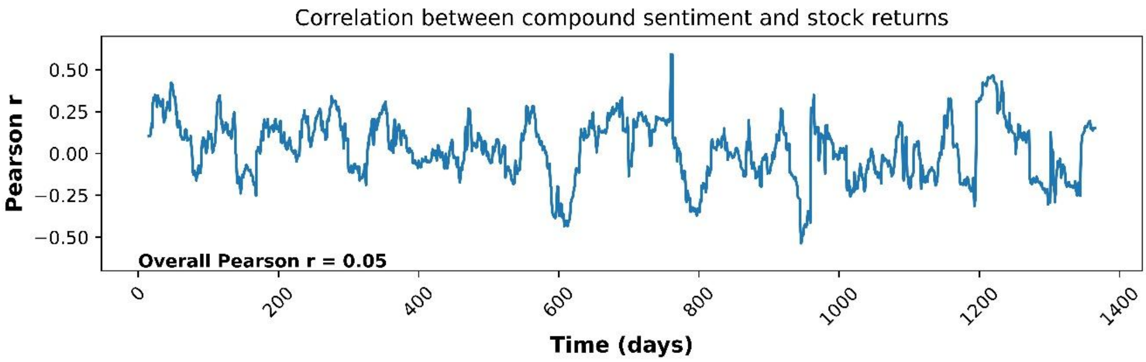
Source: Own work.

Table 12: Significant events for Activision Blizzard

	Period	Events
(1)	18/11/2020–14/01/2021	WoW: Shadowlands release
(2)	07/04/2021–07/05/2021	Call Of Duty: Warzone Verdansk 84 release
(3)	10/04/2022–10/05/2022	WoW: Dragonflight revealed Overwatch 2 PvP Beta release
(4)	01/10/2022–01/01/2023	Overwatch 2 release and patches WoW: Dragonflight release Call Of Duty MW2 game release Call Of Duty MW2: Season 1 release Call Of Duty MW2 Season 1 Reloaded release Call Of Duty: Warzone Mobile announced
(5)	03/05/2023–03/06/2023	Overwatch 2 patch

Source: Adapted from Blizzard Entertainment (2020b), Blizzard Entertainment (2022a), Blizzard Entertainment (2022b), Blizzard Entertainment (2022d), Blizzard Entertainment (2022e), and Blizzard Entertainment (2023b); Call Of Duty (2022); Cotten (2022); Goslin (2022); Heaney (2022b) and Heaney (2022c); Hodgson (2021); Noel (2020e) and Noel (2022).

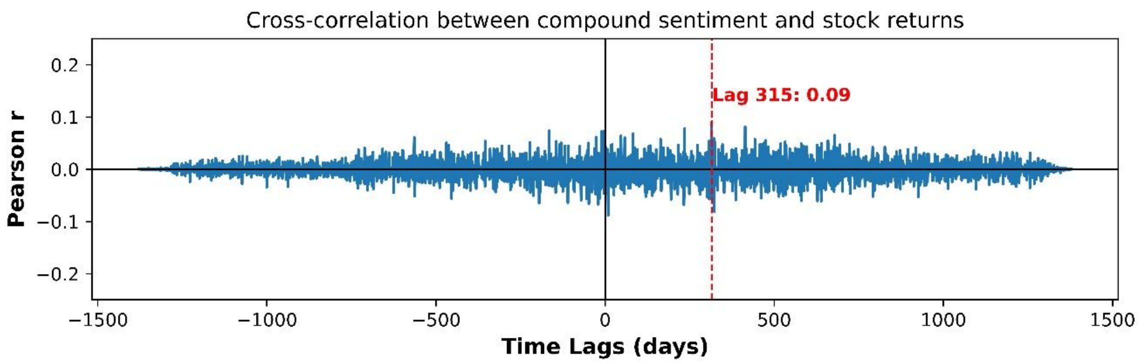
Figure 8: Rolling correlation (30 days) for Activision Blizzard



Source: Own work.

The long-term correlation trends between compound sentiment and stock returns, as illustrated in Figures 8 and 9, are highly inconsistent. The overall correlation is negligible (r = 0.05), and the cross-correlation similarly demonstrates very weak associations across all time lags, peaking at r = 0.09 at lag 315. This pattern extends to all variable interactions and series, as detailed in Appendix 10. Nevertheless, there are some local peaks of medium correlation present, which can indicate a more significant connection for some local events.

Figure 9: Cross-correlation for Activision Blizzard



Source: Own work.

As shown in Table 13 and Appendix 10, there is almost no correlation between stock returns and sentiment or emotion variables on a company level. This trend persists at the subreddit/game level, where even the sentiment and emotion series for major releases such as WoW and OW show negligible correlations with stock returns. The same is true for all CoD games, with the sole exception of the optimism variable for CoD: MW2, which exhibits a weak negative correlation (r = -0.116).

The lack of correlation is evident in local observations as well. On the company level, weak associations are observed for negative sentiment, compound sentiment, and optimism around the release of WoW: Shadowlands (1). Similarly, anticipation and sadness show weak connections around the release of Call of Duty: Warzone Verdansk 84 (2) and the reveal of WoW: Dragonflight (3). The most notable activity occurs around the release of an OW 2 patch (5), where most variables display weak, albeit detectable, correlations. The same can be said for the connection between specific subreddits and stock returns around local events. There are sparse correlations for most interactions, with two exceptions. The second studied event (2) for CoD: Warzone exhibits weak correlations across the variables. Moreover, there is a consistency amongst significant variables, as positivity is correlated with the returns. The same can be said for the fifth event (5) for OW. Nevertheless, most of these interactions still do not produce a clear-cut picture of the situation.

Table 13: Correlation and Granger causality results for Activision Blizzard

		Correlation			Granger
		weak	medium	high	
ATVI	all				
	(1)	X			X
	(2)	X			X
	(3)	X			
	(4)	X			
	(5)	X			
Wow	all				X
	(1)	X			
	(3)	X			
	(4)	X			
OW	all				
	(3)	X			X
	(4)	X			
	(5)	X			
CoD: MW2	all	X			
	(4)	X			
CoD: Warzone	all				X
	(2)	X			
	(4)	X			X
CoD: Mobile	all				

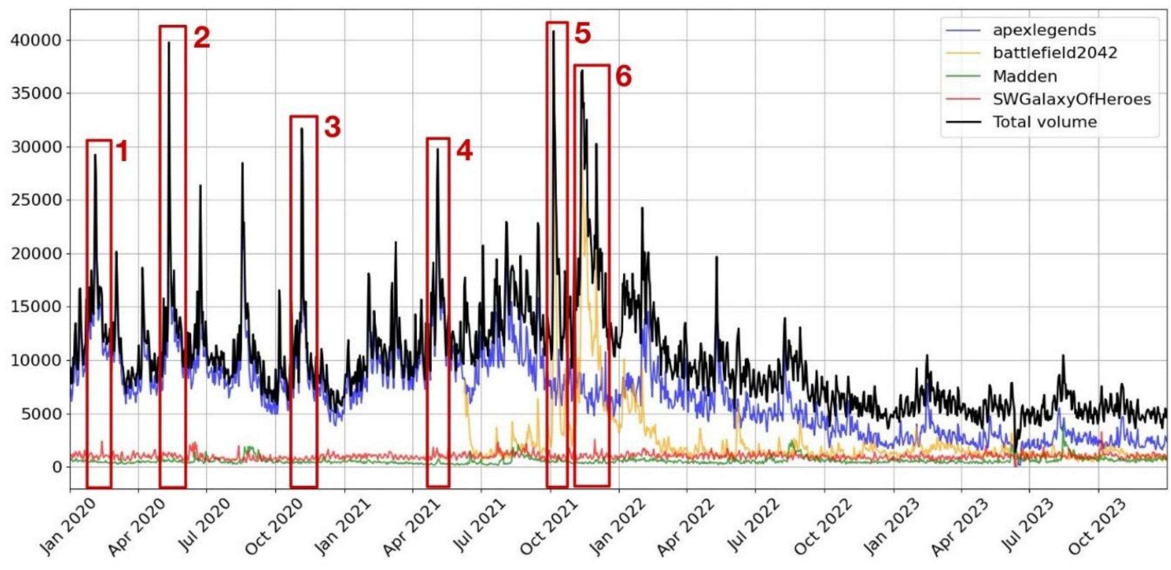
Source: Own work.

Consistent with the correlation findings, Granger causality analysis (see Table 13 and Appendix 10) reveals sparse causal relationships across the dataset. No causality is observed at the company level. Similarly, causality is absent for most games, except WoW and CoD: Warzone. For the former, the anger is significant for causality. And for the latter, compound sentiment, as well as emotions such as anger, anticipation, trust, optimism, surprise, and joy, Granger-cause stock returns. On the local level, the first observed period (1) presents causality for joy. Meanwhile, the second period (2) shows multiple significant variables with negative connotations, including anger, disgust, pessimism, and sadness, alongside joy and optimism. If we dive deeper, the causality on the subreddit level tells a different tale. The third period (3) shows us causality solely for the positive variables for both WoW and OW 2 subreddits, apart from the compound variable for the WoW series. Still, generally, the causality findings are inconsistent, with both positive and negative variables exhibiting sporadic causal relationships. Furthermore, there is somewhat of a disconnect between the company level and the subreddits.

6.2.3 Electronic Arts

Like Activision Blizzard, Electronic Arts has its flagship games that make up most of the Reddit activity (Apex Legends and Battlefield). Looking at Figure 10, we can see the periodical spikes, which are caused by in-game releases of Apex Legends, shown in Table 14, as well as the activity surrounding the release of Battlefield 2042 (BF 2042). The other titles could not be seriously considered due to the low volume of submissions.

Figure 10: Significant events for Electronic Arts



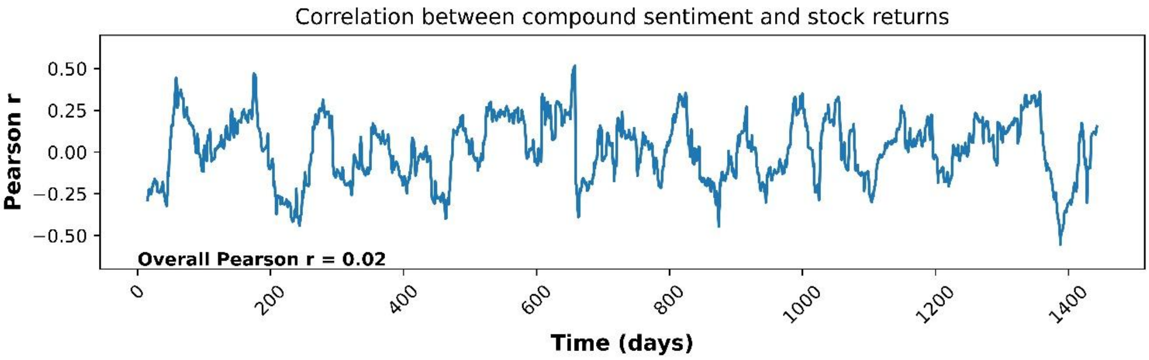
Source: Own work.

Table 14: Significant events for Electronic Arts

	Period	Events
(1)	20/01/2020–20/02/2020	Apex Legends: Season 4 release
(2)	27/04/2020–27/05/2020	Apex Legends: Season 5 release
(3)	20/10/2020–20/11/2020	Apex Legends: Season 7 release
(4)	20/04/2021–20/05/2021	Apex Legends: Season 9 release
(5)	21/09/2021–21/10/2021	Battlefield 2042 Open Beta
(6)	01/11/2021–15/12/2021	Battlefield 2042 Pre-Release Battlefield 2042 General Release Battlefield 2042 Update

Source: Adapted from Battlefield (2021); Battlefield Wiki (2021); Bull (2023); Nelson (2021).

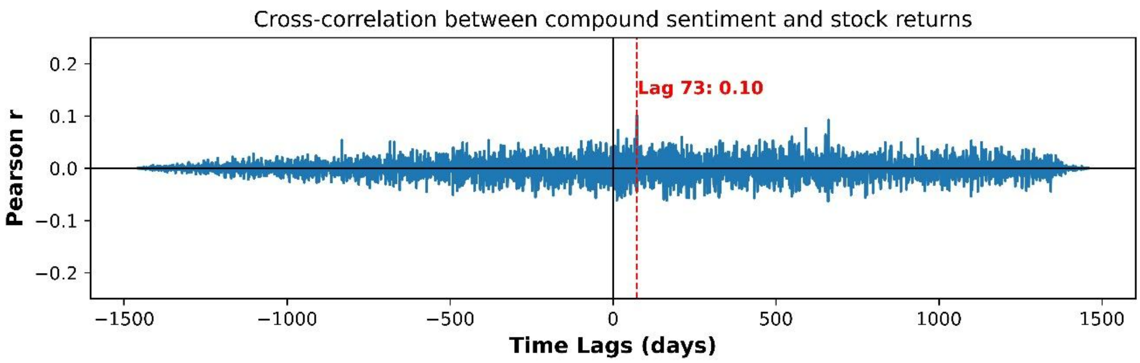
Figure 11: Rolling correlation (30 days) for Electronic Arts



Source: Own work.

Like with the previous company, there is a very weak overall correlation between compound sentiment and stock returns, as shown in Figure 11 (r = 0.02). This pattern persists with the cross-correlation results in Figure 12, where the correlation coefficient does not exceed r = 0.1. Moreover, this extends to all variables on the company level, as well as the subreddits (see Table 15 and Appendix 11). Nonetheless, there are some local surges of correlation.

Figure 12: Cross-correlation for Electronic Arts



Source: Own work.

In contrast to Activision Blizzard, the local events of Electronic Arts are defined only by two big games. Moreover, as displayed in Table 15, all these events correspond to one game. And not multiple, as is the case with the previously studied company. The local events related to Apex Legends reveal weak to moderate correlations for certain variables. For instance, the releases of Season 4 (1) and Season 5 (2) show weak correlations with negative, as well as surprise, and anticipation, respectively. In contrast, the release of Season 7 (3) paints a clearer picture. The negative variables have a positive correlation, while the positive ones have a negative correlation, with sadness and trust being the most significant. This implies the connection between the negative sentiment and stock returns for this period. The opposite is true for the Season 9 release (4), as the positive variables positive correlation, meanwhile the negative ones have a negative correlation. Hence, the positive attitudes seem to be reflected in the stock returns. Furthermore, the beta release of BF 2042 (5) also has a noticeable trend with correlation. Once again, positive variables have positive correlations and negative variables have negative correlations. Demonstrating another connection between positive sentiment and stock returns. For our last observed event (6), the premise shifts again, the correlation favouring the negative sentiment, but with less intensity than in previous cases. If we look at the results for a connection between the isolated subreddits and these events, we get consistent findings. The first two periods (1) and (2) are a mixed bag, with no clear results, whereas the periods (3)–(6) have correlations consistent with findings on the company level. With the connections being slightly stronger, we eliminate the "noise" by only focusing on one subreddit at a time for the respective events.

Table 15: Correlation and Granger causality results for Electronic Arts

		Correlation			Granger
		weak	medium	high	
EA	all				
	(1)	X			X
	(2)	X			
	(3)	X	X		X
	(4)	X	X		X
	(5)	X	X		
	(6)	X			
Apex Legends	all				
	(1)	X			X
	(2)	X			
	(3)	X	X		X
	(4)	X	X		X
BF 2042	all				
	(5)	X	X		X
	(6)	X	X		

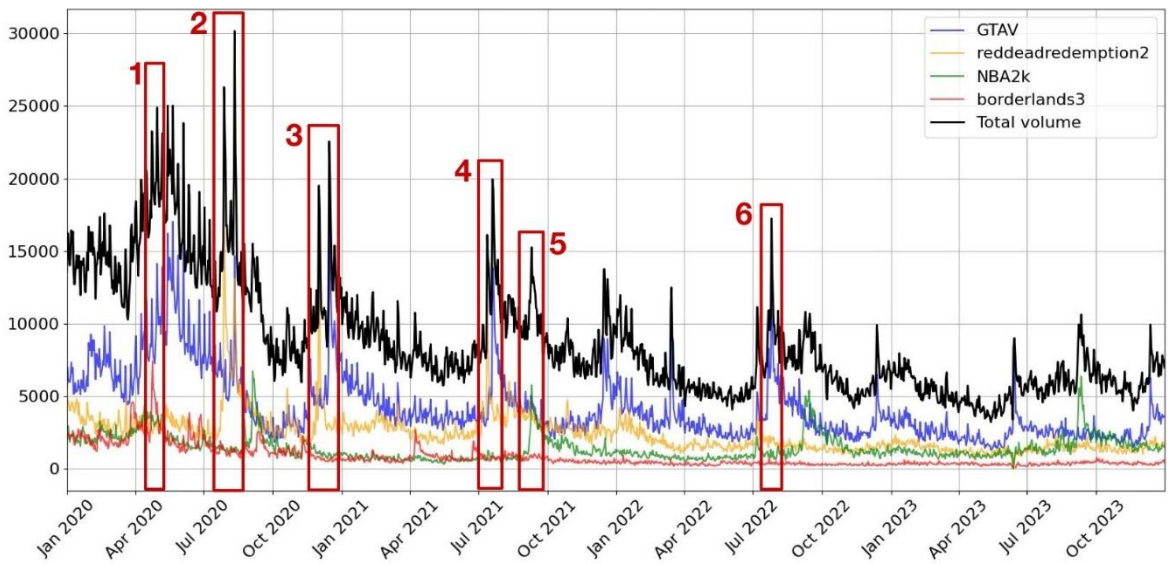
Source: Own work.

Similarly, Granger causality reveals no long-term causal relationships for none of the series (see Table 15 and Appendix 11). However, causality is detected on the company level for certain Apex Legends releases. The releases of Season 4 (1) and Season 9 (4) show very mixed results, with causality being detected for variables across the board. Conversely, the Season 7 (3) release indicates causality only for the negative emotions such as anger, pessimism, and sadness. Additionally, on the subreddit level, the findings are mostly the same. Events (1), (4), and (5) show significance for both positive and negative variables. On the other hand, the third observed event (3) exhibits causality solely for the negative variables, i.e., anger, pessimism, and sadness. This makes this event the only one consistently associated with negative markers across local correlation as well as Granger causality. Which makes it an exception since the overall causality is predominantly mixed.

6.2.4 Take-Two Interactive

The trend of big games eclipsing the others continues with Take-Two Interactive. The Grand Theft Auto and Red Dead Redemption franchises are by far the biggest ones by the volume of submissions. Hence, the events we will focus on are predominantly from those franchises but will also include other releases, as seen in Figure 13 and Table 16.

Figure 13: Significant events for Take-Two Interactive



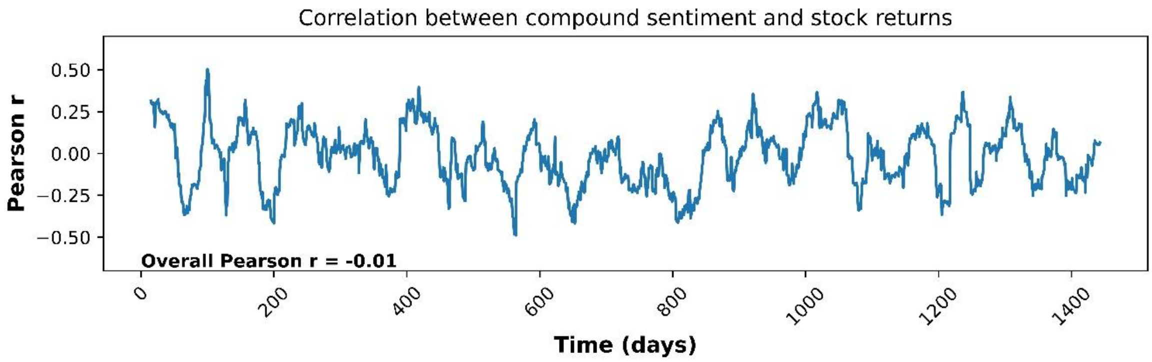
Source: Own work.

Table 16: Significant events for Take-Two Interactive

	Period	Events
(1)	08/04/2020–08/05/2020	Borderlands 3 update
(2)	18/07/2020–21/08/2020	GTA Online: Los Santos Summer Special Red Dead Online update
(3)	20/11/2020–25/12/2020	Red Dead Online update GTA Online: The Cayo Perico Heist
(4)	03/07/2021–03/08/2021	Red Dead Online update GTA Online: Los Santos Tuners
(5)	26/08/2021–26/09/2021	NBA® 2K22 Release
(6)	11/07/2022–11/08/2022	GTA Online: The Criminal Enterprises

Source: Adapted from 2K (2021) and 2K (2024); Marshall (2020); Rockstar Games (2020b), Rockstar Games (2020c), Rockstar Games (2020d), Rockstar Games (2021a), Rockstar Games (2021b), and Rockstar Games (2022b).

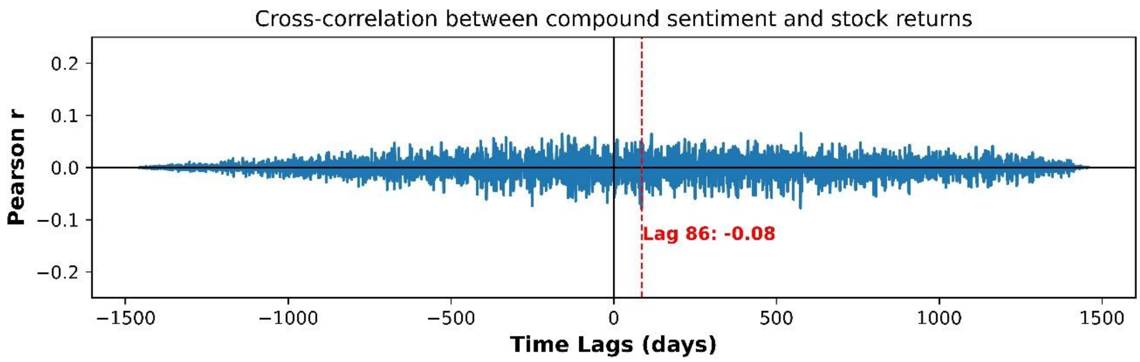
Figure 14: Rolling correlation (30 days) for Take-Two Interactive



Source: Own work.

Consistent with trends observed in previous companies, there is virtually no correlation at the company level for Take-Two Interactive (r = -0.01) at the company level, as shown in Figure 14. This pattern is mirrored in the cross-correlation results in Figure 15, where the correlation peaks at r = -0.08. These findings extend to all variables for the company and its game series for the long term, as shown in Table 17 and detailed in Appendix 12.

Figure 15: Cross-correlation for Take-Two Interactive



Source: Own work.

On the contrary, local-level analyses on the company level reveal more meaningful correlations. The timeframe surrounding the release of the Borderlands 3 (BL 3) update (1) demonstrates weak to moderate correlations associated with positive variables, and vice versa for the negative variables. This indicates a connection with the positive attitude and stock returns. The same can be observed for the third period (3) surrounding the releases of GTA V and Red Dead Redemption 2 (RDR2). Conversely, all the other observed events, i.e., (2), (4), (5), and (6), show a mix of weak and medium correlation for both sides of the sentiment. The same can be said for the observations on the subreddit level, as most events provide no clear results. The BL 3 update (1) is not consistent with the correlations on the company level, as it shows connections for both positive and negative variables. The same applies to events (3), (5), and (6). Still, there are consistent correlations for the positive variables around the second timeframe (2) for the RDR2 subreddit. Additionally, the results for the GTA V series around the fourth period (4) indicate a connection for the positive variables, apart from love, which correlates negatively. Besides that, both on the company level and the subreddit/game level, there is no clear-cut association between sentiment, emotions and stock returns.

Table 17: Correlation and Granger causality results for Take-Two Interactive

		Correlation			Granger
		weak	medium	high	
TTWO	all				X
	(1)	X	X		X
	(2)	X			X
	(3)	X	X		
	(4)	X			
	(5)	X	X		X
	(6)	X	X		
GTAV	all				X
	(2)	X	X		
	(3)	X	X		X
	(4)	X	X		X
	(6)	X	X		X
RDR2	all				
	(2)	X			X
	(3)	X			X
	(4)	X	X		X
NBA2K	all				
	(5)	X			X
BL3	all				X
	(1)	X	X		

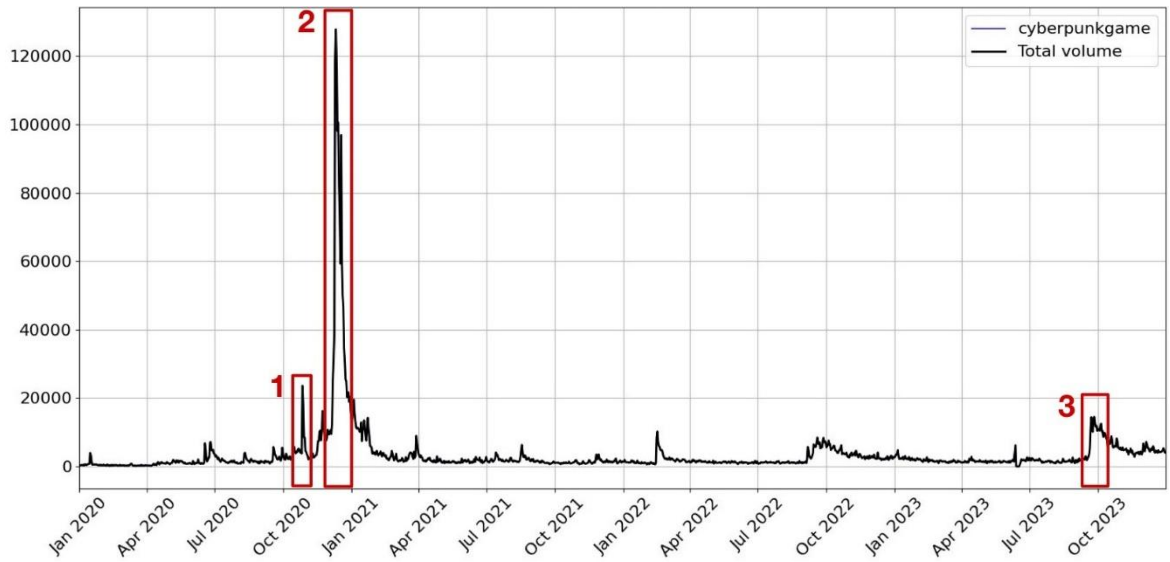
Source: Own work.

Granger causality, however, does not align with the correlation findings. First of all, the causality is detected long-term for both the company level, with anticipation, as well as the subreddits. GTA V has multiple significant variables across the spectrum, trust demonstrating causality for RDR2, and fear for BL 3. In line with correlations, causality appears across various variables for the first (1) and second (2) events. The last event to show causality is the release of NBA2K (5), where all but one (pessimism) of the significant variables are positive. Apart from that, the general findings are not definitive. This translates to the subreddit level as well. There is mixed causality for the RDR2 series in the second period (2). The same applies for the GTA V series for the third period (3) and the NBA2K series around its release (5). On the other hand, there is causality for the positive variables for the RDR2 series around the third event (3). And around the fourth event (4), the same series demonstrates causality with the surprise variable. In the same timeframe, the GTA V subreddit shows causality for the negative emotion. Additionally, both pessimism and surprise indicate causality around the last observed event (6). Overall, the causality is far and few between, with most of it showing no clear direction, apart from a few exceptions. Moreover, the disconnect with correlation seems to be a constant across all the studied companies.

6.2.5 CD Projekt Red

CD Projekt Red is a unique case in our analysis as it only has one big game release receiving regular updates in the studied period. Therefore, we focus on three major events surrounding this game as seen in Figure 16 and Table 18. The controversy surrounding the release of CP 2077 makes it a prime candidate for the context of our study.

Figure 16: Significant events for CD Projekt Red



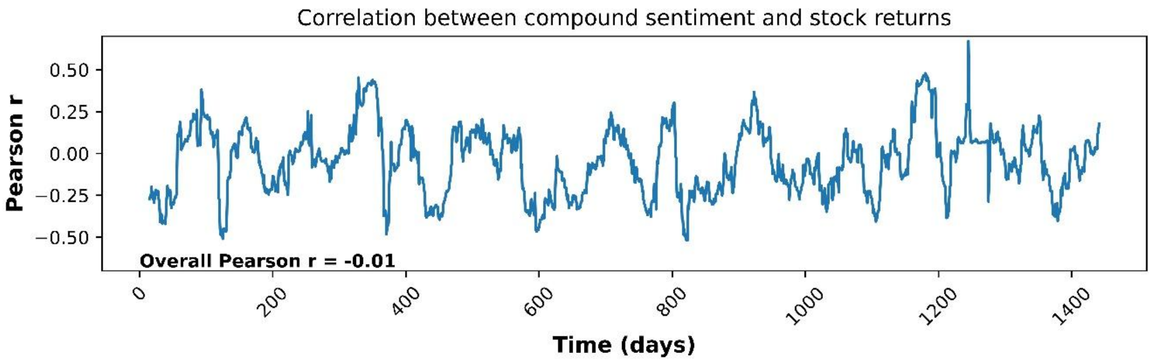
Source: Own work.

Table 18: Significant events for CD Projekt Red

	Period	Events
(1)	12/10/2020–12/11/2020	Cyberpunk 2077 delay announcement
(2)	01/12/2020–01/01/2021	Cyberpunk 2077 game release
(3)	11/09/2023–11/10/2023	Cyberpunk 2077: Phantom Liberty release

Source: Adapted from Bankhurst (2020b); Kim (2020); Sirani (2023).

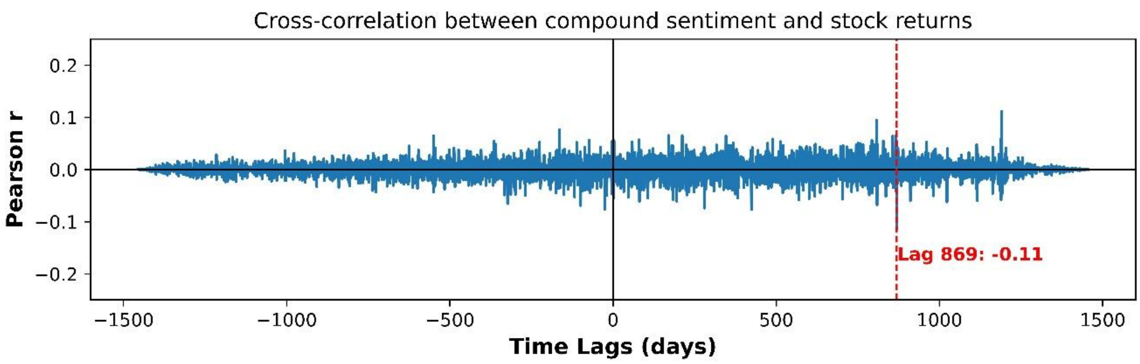
Figure 17: Rolling correlation (30 days) for CD Projekt Red



Source: Own work.

The overall correlation trend between compound sentiment and stock returns, as shown in Figure 17, is highly inconsistent. The correlation is negligible (r = -0.01), and the cross-correlation (see Figure 18) similarly reveals very weak associations across all time lags, peaking at r = -0.11. This pattern persists throughout all variables at the company level. For CD Projekt Red (CDR), CP 2077 represents the sole game in the series analysed. Consistent with trends observed in other companies, there is almost no correlation between compound sentiment and stock returns over the entire timeline. This applies to all emotion and sentiment variables, as demonstrated in Appendix 13.

Figure 18: Cross-correlation for CD Projekt Red



Source: Own work.

This lack of correlation extends to the local level as well. As summarised in Table 19, neither the pre-release delay announcement (1) nor the Phantom Liberty expansion (2) shows any substantial connection to stock returns. However, significant activity is observed surrounding the controversial release of CP 2077. Here, moderate correlations are detected for all sentiment variables, as well as for emotions such as anger, disgust, pessimism, sadness, anticipation, joy, and love. Once again, the data presents mixed signals, though the average Pearson's r is higher for negative variables compared to positive ones. Despite this, no clear conclusion can be drawn in favour of either side of the emotional spectrum.

Table 19: Correlation and Granger causality results for CD Projekt Red

		Correlation			Granger
		weak	medium	high	
CDR	all				X
	(1)				X
	(2)	X	X		
	(3)	X			
CP 2077	all				X
	(1)				X
	(2)	X	X		
	(3)	X			

Source: Own work.

In contrast, Granger causality paints a markedly different picture (see Appendix 13). Unlike correlation, nearly all variables except fear, love, pessimism, and sadness exhibit causality for the CDR series at the global level. At the local level, however, the period with moderate correlation, i.e., the CP 2077 release (2), shows no causality. Instead, causality is observed during the month preceding the release, when the game's delay was announced (1). During this period, all sentiment variables are significant, while among emotions, a clear negative bias emerges as anger, disgust, fear, pessimism, and sadness are significant, whereas none of the positive emotions show causality. Nevertheless, the significance of positive sentiment muddles the waters, preventing a clear-cut interpretation.

7 DISCUSSION

With a surge of publications over the last few years, we aim to differentiate ourselves by offering a fresh point of view from which to tackle the question of social media and its usefulness in understanding the movement of stock prices better. In contrast to papers like Li et al. (2017), Pagolu et al. (2016), Ranco et al. (2015), and others listed in Table 1, we investigate whether the consumer sentiment in specific online communities (subreddits) can in any way be connected to the movement of stock prices in a specific industry. To take on this challenge, we first had to devise an ironclad methodology to enable our analysis. Most of the steps taken here do not merit a lengthy discussion. Nevertheless, due to the specific nature of our research, some parts provide useful insight and could advance the understanding of analysis in this field. Thus, we highlight two steps of our analysis, which we believe provide merit for future work.

First, we need to tackle our selection of the studied companies, as the lack of previous research left us to devise our own process. The goal was to strike a balance between being concise and capturing the most relevant parts of the video game industry. To achieve this, we devised a set of criteria to find the best-fitting companies. In the first part, based on the previous works of Piñeiro-Chousa et al. (2023) and Xiong & Bharadwaj (2014), we chose the top players in the industry to create a pool of candidates. But this alone was not sufficient, considering the size of the online communities was also an important factor. Hence, we selected only the companies with substantial online community presence. Therefore, many companies in the mobile and online sector had to be eliminated (see Appendix 2). Another main criterion was that only companies which are solely video game developers are to be considered, eliminating giants such as Amazon, Microsoft, Sony, and Nintendo. Here we deviated both from the general approach used by the likes of Xiong & Bharadwaj (2014) and others listed in Table 1, as well as from the method used by Mertová (2023), who also focused on the video game industry. Moreover, we accounted for the different game types, monetisation techniques and the dynamics this creates for each of the companies. Hence, we constructed a framework which could provide some basis for future works regarding the video game industry.

The second element we need to address is the identification of local events. As alluded to in the first chapter, previous research has been inconsistent at best when it comes to finding the link between social media sentiment and stock prices. Consequently, we decided to study local events, which could give us a better understanding of the correlation and causality around high-volume periods in our data. To do so, we first had to identify them. Since there were no sources to lean upon as a foundation, we devised our approach. Firstly, we looked at the volume of posts and submissions in our datasets and noted any surges in activity. Secondly, we researched industry events for each of the games and their respective developers. Thirdly, we cross-referenced the peaks in volume with the events. And lastly, we chose periods for observation which best reflected various types of important events for each company (see Appendices 5-8). With this, we identified twenty events across the four studied companies. By far the biggest challenge was to decide which events are relevant enough to consider. We could not formulate a universal criterion across all companies, and at the same time, we could not just arbitrarily decide which events are relevant. Therefore, we decided on a case-by-case basis, which is by no means perfect, but necessary in this case. Additionally, we must acknowledge the bias towards bigger subreddits, as the volume of submissions overwhelms the smaller subreddits, which prevents us from spotting all local events. However, the size of the community reflects the popularity of the game and the size of its player base. Therefore, it is very unlikely that we missed any important events. Regardless of those concerns, it cannot be said that the chosen events do not reflect the peaks of activity in our data, which was the primary goal. Overall, our methodology somewhat resembles the work of Ranco et al. (2015), who relied purely on volume peaks to decide which events to study. However, we refined this approach via cross-analysis of industry events, which allowed us to both contextualise the peaks as well as eliminate the events which are not related to the games themselves (e.g., Activision Blizzard sexual harassment lawsuit, which generated significant volume, as seen in Appendix 5). Thus, our methodology separates the product (game) related events from the general discourse, which can provide a greater insight into the customer-product-company relationship in the video game industry.

The important methodological issues having been addressed, we now turn to answer the most important question of this research. What is the nature of the connection between sentiment, emotions and stock returns, and on what level is it present? To better understand our findings, we have synthesised them in Table 20.

Table 20: Key findings

F1	Long-term correlation at the company, as well as the subreddit level, is negligible.
F2	Long-term Granger causality on the company, as well as the subreddit level, is inconsistent, but much less sporadic than correlation.
F3	Short-term correlations around local events on the company, as well as the subreddit level, vary across the companies, with some instances showing clear associations, while others are unclear.
F4	Short-term correlations around the same local events on the subreddit level do not always match with correlations at the company level, but those that do tend to exhibit stronger correlations.
F5	Short-term Granger causality around local events on the company, as well as the subreddit level, varies across the companies, with some instances showing clear direction, while others are unclear.
F6	Short-term Granger causality does not translate from the company to the game level as well as correlation, but there are still instances of consistency.
F7	Correlation and causality are more definitive on a subreddit level than on a company level.
F8	Emotions are stronger predictors than sentiment for both correlation and Granger causality.
F9	There is some disconnect between correlation and Granger causality.

Source: Own work.

To expand upon our first finding (**F1**), the correlation seems to be negligible across all the datasets (see Figures 8, 11, 14, 17). Furthermore, when we calculated overall cross-correlation between the stock return and compound variables, we were met by similar results showing that even the lagged connections are very weak (see Figures 9, 12, 15, 18). This is further reflected in all the variables, as shown in Appendices 10–13. This also translates to the Granger causality test, which is our second finding (**F2**). While the causality is inconsistent, presenting itself either for a single variable or multiple variables from both the positive and negative fields, it is still almost universally present across the datasets. Therefore, even if we cannot draw a clear connection between a specific set of variables, the Granger causality still tells us that sentiment and emotions have some predictive value for long-term observations. Moreover, if we consider the frequency of significant variables and when they are observed, we can identify which are the key sentiments and emotions to observe in the future. Additionally, significant predictors are much more often found in emotions rather than in sentiment. This is true for both the more complex relationships on the level of the company, where all the sentiment and emotional scores of all the games are combined, as well as the more direct game-level analysis. To underline the first set of our findings, we find ourselves in conjunction with Ranco et al. (2015), observing a very low overall connection for our variables in the long-term setting. Nevertheless, there are clear hints of higher local correlations, which could be connected to the nature of the industry since updates and game releases are when the new information about the game is introduced, and customers express the majority of their opinions. Furthermore, we found that emotions Granger-cause the stock returns in some instances, which makes them worth noting when looking at trends. Still, with the complete picture being far from clear, it seems that one of the culprits for such results could be the length of the observation period itself, i.e., 2020–2024. While it is important to note that Chen et al. (2014) found a strong, long-term connection, and Mudinas et al. (2019) discovered the same for some longer periods; most research is still done for shorter timeframes. Behrendt & Schmidt (2018), Ho et al. (2017), Pagolu et al. (2016), and Ranco et al. (2015) elected that there should be observation periods of up to two years. However, this does not detract from the validity of our research, as we anticipated this due to the mixed findings of the previous research (detailed in chapter 2.1) and therefore focused on the shorter, more defined timeframes. This decision was further reinforced by the nature of the studied industry (game release cycles) and the trace findings when studying overall correlation.

To preface our findings around local events, it is important to note two key points. First, our local periods are based on the significant events that pertain to the games, and not company-related news such as earnings reports. Second, the collected data is not from financial communities, but rather from the communities discussing the state of the studied games, i.e., subreddits. Hence, while we will contextualise the results with the related research, the nature of the connection differs from most of the published papers (see Table 1 for an overview of related research). But we digress. The short-term findings paint a very different picture from their long-term counterparts. As we summarise in our third finding (**F3**), the correlations around local events are somewhat of a mixed bag. Unlike Corbet et al. (2022), Huynh et al. (2021), Long et al. (2023), and Machavarapu (2022), we do not study a singular local event, but rather twenty across different games and companies. Therefore, the lack of a universal connection does not detract from the significance of clear connections observed for some of the events. The strength and the nature of these connections vary. Activision Blizzard has both the lowest number and intensity of correlations, with only two events showing a weak correlation that connects positive attitude to the increase in stock returns (see Appendix 10). One could find this surprising, since we are dealing with the biggest communities, and hence the biggest volume of submissions. Nevertheless, this also invites more complexity. On the other hand, all other companies produced more enticing results. For Electronic Arts, we observed consistent weak to medium correlations for half of the observed local events (see Appendix 11). Moreover, the connections remained consistent when we moved from the company to the subreddit level, with intensity increasing. The same can be said for some local events surrounding GTA V and RDR2 of Take-Two

Interactive (see Appendix 12) with clear signs of a connection between the behaviour of stock returns and either the positive or the negative attitudes. However, somewhat surprisingly, when it comes to the release of CD Projekt Red's CP 2077, which was surrounded by controversy, there was no exclusive correlation with negativity and the stock returns. Rather, there were medium correlations amongst all variables, with the negative ones having higher coefficients (see Appendix 13). In summary, we have identified clear-cut examples of correlations for some of the local events. Even though most of these connections were of a medium strength, our findings align with Chahine & Malhotra (2018) and Li et al. (2017), who acknowledge the role of sentiment in short-term connections. Moreover, as we highlighted in our fourth finding (**F4**), when we look at the links on the micro (game) level, the detected connections from the macro (company) level grow stronger. This demonstrates that following the subreddits of respective games can bear fruit when looking for clues about stock performance. This further reinforces the usefulness of sentiment and emotion analysis when studying significant events within the industry. Hence, it is beneficial for companies to track the response of the customers to their releases, to avoid future mistakes and, in turn, limit the possibility of financial turmoil.

Now that we have established the usefulness of the association, we can ask the same about the causality. More specifically, Granger causality has provided us with some insight on the long-term level. While correlation can give us an idea if two series share a positive or negative connection, causality can indicate if one series predicts the other. In a similar fashion to correlation, our findings (**F5**) about causality are not ubiquitous. The Activision Blizzard data is, once again, the least revealing, with only three events demonstrating causality. One of them shows that positive emotions have some predictive power over stock returns (see Appendix 10), with similar findings when we move to other companies. Apex Legends demonstrates causality for negative variables for one event on both company and subreddit level (see Appendix 11), RDR2 has significant positive and GTA V significant negative variables for respective observations (see Appendix 12), and CP 2077 demonstrates causality for almost all negative variables within the month leading up to its release (see Appendix 13). Overall, we are dealing with three types of results: there are either no significant variables, the significant variables are mixed, or we get a clear indication of causality amongst positive or negative variables. While, like Ranco et al. (2015) and Sun et al. (2017), we cannot claim that we have found strong overall causality, we also cannot join Behrendt & Schmidt (2018) in discarding its significance altogether. Thus, in conjunction with Long et al. (2023), Mittal & Goel (2012), and Mudinas et al. (2019), we find some, albeit sparse, definitive causality. Moreover, there is a strong presence of mixed causality for numerous events, with only a quarter of the studied events having no causal relationships. Conversely, it is important to note another key finding (**F6**), that, unlike correlation, Granger causality does not translate that well from the company to the subreddit level. This is most evident with our usual culprit, Activision Blizzard, as it is observed for completely different events on the company and the game level. Meanwhile, for the other companies, we observed instances of matching mixed causalities, as well as some definitive connections.

Nevertheless, we should not reject the potential of Granger causality on these grounds, as it still has predictive powers. Granted, some contradictions appear when we switch between the company and the game level, but the latter takes precedence. While the findings for the company as a whole are important, the real value lies in how specific communities react to individual game performance. This is due to the nature of the industry, as video game developers stick to release cycles for their games, which means that for any given local event in our data, only one game is significant to study. This is evident if we look at the subreddit activity around such events, where we see considerable spikes in volume (see Appendices 5-8). This is corroborated by our finding (**F7**) that both correlation and causality are more present on the subreddit level. Hence, the use of sentiment and emotion analysis to explain the stock returns is at its best when it is for a short period around a significant event for a specific community. Moreover, to further refine the approach to analysis, we should consider another important discovery (**F8**), that emotional variables are more frequent and are stronger indicators of correlation and causality. We already alluded to this in our findings for the long-term trends, and it seems to translate to our entire scope. Li et al. (2017), Mudinas et al. (2019), Sprenger et al. (2014), and Wang & Luo (2021) note that including emotional analysis improves the results when studying the connection with the stock movements. Moreover, it allows us to better understand that link, since while sentiment captures the overall polarity, emotions tell us what its roots are. Hence, the more frequent significance of emotional variables is no surprise, as they are the building blocks of sentiment. This does not take away from the importance of sentiment analysis but rather contextualises it.

While the clarity increases if we observe the data on the game level or pay closer attention to emotional variables, overall, we still find a disconnect between correlation and Granger causality (**F9**). The most glaring example is the CD Projekt Red data (see Appendix 13). There, for the long term, we observe no correlation whatsoever; meanwhile, ten variables, both negative and positive, are significant for causality. Additionally, in the short term, there is an abundance of correlation for the release of CP 2077 and the subsequent update, with no causality detected. But when looking at the month leading up to the release, there are Granger-causing variables across the board, with no correlation detected. While the data for the other companies does not show this level of contradiction, in most instances, the causality still fails to follow through where the correlation is present. Moreover, even when present for the same local events, the variables that correlate do not always match those that are significant for causality. However, this does not invalidate either set of results, as correlation and Granger causality are two independent methods, with the former focusing on the connection between series at a given time, and the latter testing if the past values of emotions and sentiment can predict the stock returns. Moreover, it is to be expected that correlation would be more frequent than causality, as it is inherently easier to detect. In general, we should consider these as two tools at our disposal, with each telling an important story on its own, as well as together.

Ultimately, all the differences we encountered show the complex and layered nature of the relationship between social media sentiment and stock market behaviour. Hence, rather than seeking a universal explanation for the interactions between sentiment, emotions, and stock returns, it is better to understand the importance of different approaches which allow us to decode this connection. Be it for the long term or the short term, on the company or the game level, the combination of correlation and causality helps us to better understand the complex dynamic we study. The former captures the broader patterns in the data, while the latter identifies the causal relationships. When we use these techniques in combination with emotion classification and sentiment analysis, we allow for a deeper understanding of otherwise unseeming connections. As we highlighted, these techniques shine when used on specific events for specific communities, where we found multiple, definitive connections between the sentiment and stock data. On the other hand, they also provide important insight for more complex cases, i.e., long-term trends and whole companies. Granted, the results were not as definitive as in the short term, but still, they reveal important information, such as overall happiness with the games, most frequent Granger-causing variables, most frequent high-correlating variables, etc. Thus, creating significant analytical value for companies who want a better understanding of how their customers perceive their products, and how that ties back to their financial success.

8 CONCLUSION

This thesis aimed to investigate how customer satisfaction on social media influences stock prices of firms in the video game industry. More specifically, our goal was to select the appropriate data, study it with sentiment analysis and emotion classification, explore the links with the stock returns, and establish an understanding of how and where the connections are present. With this, we strived to enrich the lacking research surrounding social media, especially Reddit, the video game industry and the stock market. Moreover, we wanted to demonstrate that companies within this industry can use social media to great effect, deepening the understanding of their customers, thereby improving their products and ensuring better financial performance.

We achieved set goals through our analysis in which we set out to study four industry leaders, twelve of their games and associated communities on Reddit from 2020 to 2024. We successfully processed the data, calculated the sentiment and emotion scores, and finally conducted the correlation and causality tests. This has left us with some key findings. Firstly, the overall correlation and causality for the long term is negligible, but it still provides us with an insight into the most common sentiment and emotions, which were significant predictors for the stock movements. Secondly, when looking at significant industry events, we can identify some definitive correlative and causal links between sentiment, emotions and stock returns. Thirdly, studying the connections on the game level rather than on the company level is more fruitful, as we observe more definitive results for both correlation and causality, which is consistent with the mechanics of the video game industry. Fourthly, emotional variables are stronger indicators of correlation and causality, since they act as the building blocks of sentiment. And lastly, there is some disconnect between correlation and causality results, which does not detract from their validity, as we should consider them as two independent tools that offer many synergies. To synthesise, correlation and causality can provide us with limited insight as it pertains to long-term observations, while having the capacity to show clear-cut connections when focusing on specific events and communities in a shorter timeframe. Therefore, providing a deeper understanding of how social media sentiment can impact the stock movements of companies within the video game industry.

The implications of our research are twofold. Firstly, we contribute to the better understanding of social media sentiment and its connection with the stock market, especially within the context of Reddit and the video game industry, which are sparsely researched. This is partially visible in our findings, but there is another side to our contribution, which we achieve by both differentiating from and building upon existing methodology. First and foremost, we offer a more elaborate process for the selection of local events. Building upon the work of Ranco et al. (2015), we additionally consider industry events, which provide a basis for a more accurate identification. Furthermore, we define a framework for selecting suitable companies for social media analysis within the video game industry, thus moving away from the arbitrary process used by the previous works (Mertová, 2023). Moreover, our overall approach differs in the data used. Unlike Pagolu et al. (2016), Ranco et al. (2015), Sprenger et al. (2014), and others listed in Table 1, our sentiment data is comprised of customers' opinions rather than the commonly used financial discourse. This has implications for our results as well. We are in alignment with Chahine & Malhotra (2018), Li et al. (2017), and Ranco et al. (2015) when it comes to acknowledging the significance of local events. And while we do not find any strong long-term links like Chen et al. (2014), Ho et al. (2017), and Pagolu et al. (2016), we are in parallel with Li et al. (2017), Mudinas et al. (2019), Sprenger et al. (2014), and Wang & Luo (2021) in recognising the role of emotions in the long-term. Nevertheless, as our results stem from non-financial discourse, they provide a fresh insight into these relationships. Hence, we advance the understanding of the link between the customer, product and the stock value. Secondly, our findings demonstrate that companies can extract significant value from such analysis. This particularly applies to the firms in the video game industry, as they can identify the most significant sentiments and emotions around their games, which are reflected in their stock prices. Thereby, allowing them to understand the overall satisfaction of their customers, and which of their products need changing. Furthermore, by studying local events, they can discover how their past releases were received amongst the player base, how they affected their stock prices, and use this to avoid mistakes in their future releases. Additionally, they can continuously improve their products by monitoring sentiment on social media channels and harnessing the constant feedback their customers provide. Hence, taking advantage of social media data is paramount to any company, especially within the video game industry.

However, we must note some limitations of our study. First, the choice of games and subreddits does not fully reflect the entirety of the discourse about respective companies. The second point of contention is that social media texts are considered one of the hardest to evaluate in terms of sentiment analysis. Moreover, the video game community uses very specific language which conventional sentiment and emotion analysis tools are not designed to evaluate. To offset this, we added custom words into the lexicons of these classifiers, but still, this does not solve the broader issue. Another limitation pertains to the content of our data, as the discussion on subreddits is not strictly limited to the quality of the game, which can decrease the accuracy of the sentiment and emotion scores. Additionally, these scores are not weighted. When it comes to correlation and causality analysis, it is worth noting that the observational period for local events is 30 days, or more specifically, 15 days before and after the date of the event. This was done to prevent the crossover between the events and to capture only the peaks in the volume of submission, which might detract from the accuracy of correlation and causality results.

These limitations invite further research in this field, especially when it comes to better classification and understanding of the language in the video game community. Research leading to a proper lexicon of terms would be an essential step to allow for better sentiment analysis in this field. More generally, we invite a greater focus on Reddit as a source of data for financial and stock market analysis. And crucially, we hope that researchers and companies capitalise on the value of customer opinions as opposed to purely focusing on financial discourse when studying the links between social media and the stock market.

REFERENCE LIST

1. 2K. (2020, September 4). *Everything is Game: NBA® 2K21 Now Available Worldwide*. https://newsroom.2k.com/news/everything-is-game-nbar-2k21-now-available-worldwide
2. 2K. (2021, July 14). *Anyone, Anywhere: NBA® 2K22 Features Luka Dončić and NBA Scoring Legends – Kareem Abdul-Jabbar, Dirk Nowitzki, and Kevin Durant – as Cover Athletes*. https://newsroom.2k.com/news/anyone-anywhere-nbar-2k22-features-luka-doncic-and-nba-scoring-legends-kareem-abdul-jabbar-dirk-nowitzki-and-kevin-durant-as-cover-athletes
3. 2K. (2024). *Update Notes*. https://borderlands.2k.com/borderlands-3/update-notes/
4. Activision Blizzard. (2008, July 10). *Vivendi and Activision Complete Transaction to Create Activision Blizzard*. Activision Blizzard. https://investor.activision.com/news-releases/news-release-details/vivendi-and-activision-complete-transaction-create-activision
5. Activision Blizzard. (2016, February 23). *Activision Blizzard Completes King Acquisition Becomes the Largest Game Network in the World with over 500 Million Users*. Activision Blizzard. https://investor.activision.com/news-releases/news-release-details/activision-blizzard-completes-king-acquisition-becomes-largest

6. Activision Blizzard. (2024). *Battle.net*. https://us.shop.battle.net/en-us
7. Ahmad, N. B., Barakji, S. A. R., Shahada, T. M. A., & Anabtawi, Z. A. (2017). How to launch a successful video game: A framework. *Entertainment Computing*, *23*, 1–11. https://doi.org/10.1016/j.entcom.2017.08.001
8. Anderson, K. E. (2015). Ask me anything: What is reddit? *Library Hi Tech News*, *32*(5), 8–11. https://doi.org/10.1108/LHTN-03-2015-0018
9. Antypas, D., Ushio, A., Barbieri, F., Neves, L., Rezaee, K., Espinosa-Anke, L., Pei, J., & Camacho-Collados, J. (2023). *SUPERTWEETEVAL: A Challenging, Unified and Heterogeneous Benchmark for Social Media NLP Research*. https://doi.org/https://doi.org/10.48550/arXiv.2310.14757
10. Apex Legends. (2021, May 9). *Apex Legends Launches on Nintendo Switch on March 9th, 2021*. Apex Legends. https://www.ea.com/en-gb/games/apex-legends/news/switch-launch-march-2021
11. Apex Legends Wiki. (2020a, April 7). *April 07, 2020 Patch*. Fandom. https://apexlegends.fandom.com/wiki/April_07,_2020_Patch
12. Apex Legends Wiki. (2020b, June 23). *June 23, 2020 Patch*. Fandom. https://apexlegends.fandom.com/wiki/June_23,_2020_Patch
13. Apex Legends Wiki. (2023, February 14). *February 14, 2023 Patch*. Fandom. https://apexlegends.fandom.com/wiki/February_14,_2023_Patch
14. Awan, M. J., Rahim, M. S. M., Nobanee, H., Munawar, A., Yasin, A., & Zain, A. M. (2021). Social Media and Stock Market Prediction: A Big Data Approach. *Computers, Materials and Continua*, *67*(2), 2569–2583. https://doi.org/10.32604/cmc.2021.014253
15. Baldwin, T., Cook, P., Lui, M., MacKinlay, A., & Wang, L. (2013). How Noisy Social Media Text, How Diffrnt Social Media Sources? In R. Mitkov (Ed.), *Proceedings of the Sixth International Joint Conference on Natural Language Processing* (pp. 356–364). Asian Federation of Natural Language Processing. http://rankings.big-boards.com
16. Bankhurst, A. (2020a, November 23). *Cyberpunk 2077 Physical Copies Appear to Be in the Wild, So Beware of Spoilers [Update: Gameplay Leaks]*. IGN. https://www.ign.com/articles/cyberpunk-2077-physical-copies-appear-to-be-in-the-wild-so-beware-of-spoilers
17. Bankhurst, A. (2020b, December 5). *Cyberpunk 2077 Global Release Times Revealed*. IGN. https://www.ign.com/articles/cyberpunk-2077-global-release-times-revealed
18. Bankhurst, A. (2023, November 4). *BlizzCon 2023: Everything Announced and Revealed*. IGN. https://www.ign.com/articles/blizzcon-2023-everything-announced-and-revealed
19. Bankov, B. (2019). *The Impact of Social Media on Video Game Communities and the Gaming Industry*. https://www.researchgate.net/publication/337144821
20. Battlefield. (2021, December 2). *BATTLEFIELD 2042 UPDATE #3*. https://www.ea.com/en-gb/games/battlefield/battlefield-2042/news/battlefield-2042-update-notes-3

21. Battlefield Wiki. (2021). *Battlefield 2042*. Fandom. https://battlefield.fandom.com/wiki/Battlefield_2042
22. Battlefield Wiki. (2022). *Battlefield 2042: Zero Hour*. Fandom. https://battlefield.fandom.com/wiki/Battlefield_2042:_Zero_Hour
23. Baumgartner, J. (2024). *zreader*. GitHub. https://github.com/pushshift/zreader
24. Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., Blackburn, J., & Io, P. (2020). The Pushshift Reddit Dataset. *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media*, *2020*. www.aaai.org
25. Behrendt, S., & Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. *Journal of Banking and Finance*, *96*, 355–367. https://doi.org/10.1016/j.jbankfin.2018.09.016
26. Bergstrom, K., & Poor, N. (2021). Reddit Gaming Communities During Times of Transition. *Social Media and Society*, *7*(2). https://doi.org/10.1177/20563051211010167
27. Biehl, M. (2014). *API architecture* (Vol. 2). API-University Press.
28. Biswas, S., Bose, R., Roy, S., Sarkar, I., & Das, P. (2020). *Examining the Effects of Pandemics on Stock Market Trends through Sentiment Analysis*. https://www.researchgate.net/publication/342083661
29. Blizzard Entertainment. (2020a, October 14). *The Shadowlands Pre-Expansion Patch is Now Live*. https://worldofwarcraft.blizzard.com/en-us/news/23529212/the-shadowlands-pre-expansion-patch-is-now-live
30. Blizzard Entertainment. (2020b, October 29). *Shadowlands Arrives November 23*. https://worldofwarcraft.blizzard.com/en-us/news/23558953/shadowlands-arrives-november-23
31. Blizzard Entertainment. (2022a, April 25). *Welcome to the first Overwatch 2 PvP Beta*. https://overwatch.blizzard.com/en-gb/news/23798985/welcome-to-the-first-overwatch-2-pvp-beta/
32. Blizzard Entertainment. (2022b, June 12). *Overwatch 2 Launches October 4 as a Free-to-Play Live Experience*. https://overwatch.blizzard.com/en-us/news/23814216/overwatch-2-launches-october-4-as-a-free-to-play-live-experience/
33. Blizzard Entertainment. (2022c, September 15). *Overwatch 2 Explained: Battle Pass, Shop, Hero Unlocks, and more*. https://overwatch.blizzard.com/en-us/news/23824003/overwatch-2-explained-battle-pass-shop-hero-unlocks-and-more/
34. Blizzard Entertainment. (2022d, November 28). *World of Warcraft: Dragonflight Now Live!* https://worldofwarcraft.blizzard.com/en-us/news/23873409/world-of-warcraft-dragonflight-now-live
35. Blizzard Entertainment. (2022e, December 6). *Overwatch 2 Retail Patch Notes - December 6, 2022*. https://us.forums.blizzard.com/en/overwatch/t/overwatch-2-retail-patch-notes-december-6-2022/758323/1
36. Blizzard Entertainment. (2023a, January 25). *Dragonflight: 10.0.5 Content Update Notes*. https://worldofwarcraft.blizzard.com/en-us/news/23892227/updated-january-24-dragonflight-1005-content-update-notes

37. Blizzard Entertainment. (2023b, May 12). *Overwatch 2 Retail Patch Notes - May 17, 2023*. https://overwatch.blizzard.com/en-us/news/patch-notes/live/2023/05/
38. Borja, B. E., Chua, G. M. G., Corpuz, F. J., Dajac, C. V., Ibañez, S. C., Javier, P. J. E., Liponhay, M. P., & Ventura, M. E. M. (2020). *Time Series Handbook*. Asian Institute of Management.
39. Broadhurst, C. N., Abuomar, O. Ls., & Szczurek, P. (2022). Data Analytics On Nasdaq Stock Prices: Reddit Social Media Case Study. *IEEE International Conference on Electro Information Technology*, *2022-May*, 53–60. https://doi.org/10.1109/eIT53891.2022.9813913
40. Bu, H., & Pi, L. (2014). Does investor sentiment predict stock returns? The evidence from Chinese stock market. *Journal of Systems Science and Complexity*, *27*(1), 130–143. https://doi.org/10.1007/s11424-013-3291-y
41. Bull, T. (2023, November 28). *Every Season of Apex Legends: Dates, Content, and More*. Esports.Gg. https://esports.gg/news/apex-legends/every-season-of-apex-legends-dates-content-and-more/
42. Burger-Helmchen, T., & Cohendet, P. (2011). User communities and social software in the video game industry. *Long Range Planning*, *44*(5–6), 317–343. https://doi.org/10.1016/j.lrp.2011.09.003
43. Call Of Duty. (2022, August 9). *Call of Duty®: Modern Warfare® II Editions, Benefits Detailed [Updated]*.
44. Call Of Duty Wiki. (2021). *Caldera*. Fandom. https://callofduty.fandom.com/wiki/Caldera
45. Call Of Duty Wiki. (2022). *Season Three (Modern Warfare)*. Fandom. https://callofduty.fandom.com/wiki/Season_Three_(Modern_Warfare)
46. Call Of Duty Wiki. (2023). *Season Two (Modern Warfare II)*. Fandom. https://callofduty.fandom.com/wiki/Season_Two_(Modern_Warfare_II)
47. Canto, A. A. C. (2019). *The Effect of Online Reviews on the Shares of Video Game Publishing Companies*. Dublin Business School.
48. Carr, C. T., & Hayes, R. A. (2015). Social Media: Defining, Developing, and Divining. *Atlantic Journal of Communication*, *23*(1), 46–65. https://doi.org/10.1080/15456870.2015.972282
49. Castillo, C., El-Haddad, M., Pfeffer, J., & Stempeck, M. (2014). Characterizing the life cycle of online news stories using social media reactions. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 211–223. https://doi.org/10.1145/2531602.2531623
50. CD Projekt Red. (2021, March 29). *Patch 1.2*. https://www.cyberpunk.net/en/news/37843/patch-1-2
51. CD Projekt Red. (2022, February 15). *Patch 1.5 & Next-Generation Update — list of changes*. https://www.cyberpunk.net/en/news/41435/patch-1-5-next-generation-update-list-of-changes
52. CD Projekt Red. (2024). *Games*. CD Projekt Red. https://www.cdprojektred.com/en/our-games

53. Ceci, L. (2024, February 27). *Number of monthly active Reddit users worldwide from 2018 to 2026*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/forecasts/1309791/reddit-mau-worldwide
54. Chahine, S., & Malhotra, N. K. (2018). Impact of social media strategies on stock price: the case of Twitter. *European Journal of Marketing*, *52*(7–8), 1526–1549. https://doi.org/10.1108/EJM-10-2017-0718
55. Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, *29*(3), 509–553. https://doi.org/10.1017/S1351324922000213
56. Chen, H., De, P., Hu, Y., Hwang, B.-H., Faccio, M., Kim, S., Liberti, J., Lou, D., Loughran, T., Polak, I., Xu, J., & participants, seminar. (2014). Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media Get access Arrow. *The Review of Financial Studies*, *27*(5), 1367–1403. http://ssrn.com/abstract=1807265Electroniccopyavailableat:https://ssrn.com/abstract=1807265Electroniccopyavailableat:http://ssrn.com/abstract=1807265
57. Chen, Y., & Ma, H. (2024). *Predicting Stock Market Using Reddit Comments*. https://hal.science/hal-04667405
58. Clement, J. (2022, December 9). *Most popular social media platforms for gamers in the United States in regards to gaming as of August 2022*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/statistics/1351103/us-most-used-social-related-to-gaming/
59. Clement, J. (2024a, February 29). *Activision Blizzard - Statistics & Facts*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/topics/8033/activision-blizzard/#topicOverview
60. Clement, J. (2024b, February 29). *Take-Two Interactive - Statistics & Facts*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/topics/8891/take-two-interactive/#topicOverview
61. Clement, J. (2024c, February 29). *Ubisoft - Statistics & Facts*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/topics/11717/ubisoft/#topicOverview
62. Clement, J. (2024d, May 22). *Market capitalization of the largest gaming companies worldwide as of April 2024*. Statista. https://www.statista.com/statistics/1197213/market-value-of-the-largest-gaming-companies-worldwide/
63. Clement, J. (2024e, July 18). *Electronic Arts (EA) - Statistics & Facts*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/topics/7545/electronic-arts-ea/#topicOverview
64. Clemet, J. (2024, August 22). *Number of video game users worldwide from 2019 to 2029*. Statista. https://www.statista.com/statistics/1344682/global-video-gamers-by-segment/
65. Cohen, J. P., & Lo, H. Z. (2014). Academic Torrents. *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment*, 1–2. https://doi.org/10.1145/2616498.2616528

66. Corbet, S., Hou, Y., Hu, Y., & Oxley, L. (2022). We Reddit in a Forum: The Influence of Message Boards on Firm Stability. *Review of Corporate Finance*, *2*, 151–190. https://doi.org/10.1561/114.00000014_app
67. Cotten, S. (2022, October 26). *Overwatch 2 Patch Notes Oct 25*. https://www.ign.com/wikis/overwatch-2/Overwatch_2_Patch_Notes_Oct_25
68. Cotten, S. (2023, February 7). *Overwatch 2 Season 3 Release Date and Details*. https://www.ign.com/wikis/overwatch-2/Overwatch_2_Season_3_Release_Date_and_Details
69. Cutler, D. M., Poterba, J. M., & Summers, L. H. (1998). What Moves Stock Prices? (Spring 1989). In *Streetwise* (pp. 56–64). Princeton University Press. https://doi.org/10.1515/9781400829408-008
70. Derakhshan, A., & Beigy, H. (2019). Sentiment analysis on stock social media for stock price movement prediction. *Engineering Applications of Artificial Intelligence*, *85*, 569–578. https://doi.org/10.1016/j.engappai.2019.07.002
71. Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. https://arxiv.org/abs/1810.04805
72. Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia Computer Science*, *161*, 707–714. https://doi.org/10.1016/j.procs.2019.11.174
73. Duz Tan, S., & Tas, O. (2021). Social Media Sentiment in International Stock Returns and Trading Activity. *Journal of Behavioral Finance*, *22*(2), 221–234. https://doi.org/10.1080/15427560.2020.1772261
74. Electronic Arts. (2024). *Games*. Electronic Arts. https://www.ea.com/en-gb
75. Fiesler, C., & Dym, B. (2020). Moving Across Lands: Online Platform Migration in Fandom Communities. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1). https://doi.org/10.1145/3392847
76. Frank, A. (2020, December 18). *How one of the biggest games of 2020 became one of the most controversial*. Vox. https://www.vox.com/culture/22187377/cyberpunk-2077-criticism-ps4-xbox-one-bugs-glitches-refunds
77. Freeman, G. (2023, August 11). *Overwatch 2: Invasion is Here With New Adventures, and More*. Tech Power Up. https://www.techpowerup.com/312266/overwatch-2-invasion-is-here-with-new-adventures-and-more
78. French, K. R. (1980). STOCK RETURNS AND THE WEEKEND EFFECT. In *Journal of Fmancml Economms* (Vol. 8). North-Holland Pubhshmg Company.
79. Garg, D., & Tiwari, P. (2021). Impact of social media sentiments in stock market predictions: A bibliometric analysis. *Business Information Review*, *38*(4), 170–182. https://doi.org/10.1177/02663821211058666
80. Glover, S. (2020, June 11). *Captain Price leads the charge in a packed new Season of Call of Duty®: Modern Warfare® including Warzone*. https://blog.activision.com/call-of-duty/2020-06/Captain-Price-leads-the-charge-in-a-packed-new-Season-of-Call-of-Duty-Modern-Warfare-including-Warzone

81. Goslin, A. (2022, April 19). *World of Warcraft's next expansion, Dragonflight, revealed*. Polygon. https://www.polygon.com/23031271/world-of-warcraft-expansion-dragonflight-wow-new-class-cinematic-trailer-release
82. Goslin. Austen. (2020, May 4). *Apex Legends players are finally getting a bigger inventory*. Polygon. https://www.polygon.com/battle-royale/2020/3/4/21164528/apex-legends-system-override-patch-notes-inventory-upgrade
83. Granger, C. W. J. (1969). *Investigating Causal Relations by Econometric Models and Cross-spectral Methods* (Vol. 37, Issue 3).
84. Guo, K., Sun, Y., & Qian, X. (2017). Can investor sentiment be used to predict the stock price? Dynamic analysis based on China stock market. *Physica A: Statistical Mechanics and Its Applications*, *469*, 390–396. https://doi.org/10.1016/j.physa.2016.11.114
85. Han, B., Cook, P., & Baldwin, T. (2013). Lexical normalization for social media text. *ACM Transactions on Intelligent Systems and Technology*, *4*(1). https://doi.org/10.1145/2414425.2414430
86. Heaney, S. (2022a, October 17). *Beta Guide: All Available Maps, Game Modes, and Rewards*. IGN. https://www.ign.com/wikis/call-of-duty-mw2/Beta_Guide:_All_Available_Maps,_Game_Modes,_and_Rewards
87. Heaney, S. (2022b, November 17). *MW2 Season 1 Update: 1.10 Patch Notes*. IGN. https://www.ign.com/wikis/call-of-duty-mw2/MW2_Season_1_Update:_1.10_Patch_Notes
88. Heaney, S. (2022c, December 14). *MW2 Season 1 Reloaded Release Time and Details*. IGN. https://www.ign.com/wikis/call-of-duty-mw2/MW2_Season_1_Reloaded_Release_Time_and_Details
89. Hermansen, A. K., & Bratli, R. J. (2016). *How third party reviews affect firm value An event study in the video game industry*.
90. Ho, C. S., Damien, P., Gu, B., & Konana, P. (2017). The time-varying nature of social media sentiments in modeling stock returns. *Decision Support Systems*, *101*, 69–81. https://doi.org/10.1016/j.dss.2017.06.001
91. Hodgson, D. (2021, April 22). *Welcome to the New Warzone: An Overview of Verdansk '84*. https://www.callofduty.com/blog/2021/04/Warzone-Season-Three-Overview-Verdansk-84
92. Holmes, E. E., Scheuerell, M. D., & Ward, E. J. (2021). Dickey-Fuller and Augmented Dickey-Fuller tests. In *Applied Time Series Analysis for Fisheries and Environmental Sciences*.
93. Hu, D., Jones, C. M., Zhang, V., & Zhang, X. (2021). The Rise of Reddit: How Social Media Affects Retail Investors and Short-sellers' Roles in Price Discovery. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3807655
94. Hume, M. (2021, April 20). *Project Magma: The untold origin of Verdansk, the Gulag and 'Call of Duty: Warzone'*. Washington Post. https://www.washingtonpost.com/video-games/2021/04/20/warzone-call-of-duty-history/

95. Hutto, C. J., & Gilbert, E. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. http://sentic.net/
96. Huynh, D., Audet, G., Alabi, N., & Tian, Y. (2021). Stock Price Prediction Leveraging Reddit: The Role of Trust Filter and Sliding Window. *Proceedings - 2021 IEEE International Conference on Big Data, Big Data 2021*, 1054–1060. https://doi.org/10.1109/BigData52589.2021.9671412
97. Investing.com. (2024a). *Activision Blizzard Inc (ATVI)*. Investing.Com. https://www.investing.com/equities/activision-inc-historical-data
98. Investing.com. (2024b). *CD PROJEKT SA (CDR)*. Investing.Com. https://www.investing.com/equities/cdproject
99. Ivanov, M., Wittenzellner, H., & Wardaszko, M. (2021). Video Game Monetization Mechanisms in Triple A (AAA) Video Games. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11988 LNCS*, 389–404. https://doi.org/10.1007/978-3-030-72132-9_33
100. Janková, Z. (2023). CRITICAL REVIEW OF TEXT MINING AND SENTIMENT ANALYSIS FOR STOCK MARKET PREDICTION. *Journal of Business Economics and Management*, *24*(1), 177–198. https://doi.org/10.3846/jbem.2023.18805
101. Jo Dixon, S. (2024a, May 17). *Number of social media users worldwide from 2017 to 2028*. Statista. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/
102. Jo Dixon, S. (2024b, May 22). *Most popular social networks worldwide as of April 2024, ranked by number of monthly active users*. Statista. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/
103. Johnson, D., Gardner, J., & Sweetser, P. (2016). Motivations for videogame play: Predictors of time spent playing. *Computers in Human Behavior*, *63*, 805–812. https://doi.org/10.1016/j.chb.2016.06.028
104. 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
105. Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, *66*, 236–247. https://doi.org/10.1016/j.chb.2016.09.024
106. Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241–251. https://doi.org/10.1016/j.bushor.2011.01.005
107. Kim, M. (2020, October 27). *Cyberpunk 2077 Twitter Account Reckoning With Its Old Tweets After Delay*. IGN. https://www.ign.com/articles/cyberpunk-2077-delay-old-tweets

108. King, D. L., & Delfabbro, P. H. (2018). Predatory monetization schemes in video games (e.g. 'loot boxes') and internet gaming disorder. *Addiction*, *113*(11), 1967–1969. https://doi.org/10.1111/add.14286
109. Krouska, A., Troussas, C., & Virvou, M. (2016). *The effect of preprocessing techniques on Twitter Sentiment Analysis*. http://tweenator.com/index.php?page_id=13
110. Lawler, R. (2021, July 4). *Apex Legends hackers interrupt games with messages about Titanfall hacking*. The Verge. https://www.theverge.com/2021/7/4/22563257/save-titanfall-hackers-apex-legends-ea-respawn
111. Li, B., Chan, K. C. C., Ou, C., & Ruifeng, S. (2017). Discovering public sentiment in social media for predicting stock movement of publicly listed companies. *Information Systems*, *69*, 81–92. https://doi.org/10.1016/j.is.2016.10.001
112. Lindskog, S., & Serur, J. A. (2020). Reddit Sentiment Analysis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3887779
113. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. http://arxiv.org/abs/1907.11692
114. Long, S., Lucey, B., Xie, Y., & Yarovaya, L. (2023). "I just like the stock": The role of Reddit sentiment in the GameStop share rally. *Financial Review*, *58*(1), 19–37. https://doi.org/10.1111/fire.12328
115. Machavarapu, A. (2022). Reddit Sentiments Effects on Stock Market Prices. *Proceedings of Fifth International Conference on Smart Computing and Informatics (SCI 2021)*, 75–84. https://doi.org/10.1007/978-981-16-9669-5_7
116. Malz, A. M. (2021). The GameStop Episode: What Happened and What Does It Mean? *Cato Journal*, *41*(3), 529–550. https://doi.org/10.1111/jacf.12481
117. Mancini, A., Desiderio, A., Di Clemente, R., & Cimini, G. (2022). Self-induced consensus of Reddit users to characterise the GameStop short squeeze. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-17925-2
118. Marchand, A., & Hennig-Thurau, T. (2013). Value creation in the video game industry: Industry economics, consumer benefits, and research opportunities. *Journal of Interactive Marketing*, *27*(3), 141–157. https://doi.org/10.1016/j.intmar.2013.05.001
119. Marshall, C. (2020, July 28). *Red Dead Online's long-awaited update brings a new role: the naturalist*. Polygon. https://www.polygon.com/2020/7/28/21337182/red-dead-online-naturalist-update-outlaw-pass-gameplay-wardrobe-rewards
120. Marshall, C. (2021, July 22). *World of Warcraft players are hosting sit-in protests after Blizzard allegations*. Polygon. https://www.polygon.com/22589046/world-of-warcraft-players-protest-blizzard-lawsuit-california-discrimination-harassment-allegations
121. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

122. Medvedev, A. N., Lambiotte, R., & Delvenne, J.-C. (2019). *The Anatomy of Reddit: An Overview of Academic Research* (pp. 183–204). https://doi.org/10.1007/978-3-030-14683-2_9
123. Mertová, V. (2023). *The Impact of News on Videogame Stock Market Prices and Volatility* (Master's Thesis). Charles University Faculty of Social Sciences Institute of Economic Studies.
124. Microsoft. (2022, January 18). *Microsoft to acquire Activision Blizzard to bring the joy and community of gaming to everyone, across every device*. Microsoft News Center. https://news.microsoft.com/2022/01/18/microsoft-to-acquire-activision-blizzard-to-bring-the-joy-and-community-of-gaming-to-everyone-across-every-device/
125. Mittal, A., & Goel, A. (2012). Stock Prediction Using Twitter Sentiment Analysis. In *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*. IEEE. https://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf
126. Mudinas, A., Zhang, D., & Levene, M. (2019). *Market Trend Prediction using Sentiment Analysis: Lessons Learned and Paths Forward*.
127. NASDAQ. (2024). *Historical Data*. NASDAQ. https://www.nasdaq.com/market-activity/quotes/historical
128. Nelson, M. (2021, October 6). *Xbox Game Pass Ultimate Members with EA Play Can Join the Battlefield 2042 Open Beta Starting Today*. Xbox Wire. https://news.xbox.com/en-us/2021/10/06/battlefield-2042-open-beta-starts-today/
129. NetEase Games. (2024). *Game Portfolio*. https://www.neteasegames.com/game/hot/
130. NEXON. (2024). *All Games*. NEXON. https://www.nexon.com/main/en
131. Nguyen, T. H., & Shirai, K. (2015). Topic Modeling based Sentiment Analysis on Social Media for Stock Market Prediction. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 1354–1364. https://doi.org/10.3115/v1/P15-1131
132. Nguyen, T. H., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, *42*(24), 9603–9611. https://doi.org/10.1016/j.eswa.2015.07.052
133. Noel, D. (2020a, June 11). *Radioactive Agent, the new Season of Call of Duty®: Mobile, is now live!* Https://Blog.Activision.Com/Call-of-Duty/2020-06/Radioactive-Agent-the-New-Season-of-Call-of-Duty-Mobile-Is-Now-Live.
134. Noel, D. (2020b, August 14). *Introducing Conquest, the new Season of Call of Duty®: Mobile.*
135. Noel, D. (2020c, October 14). *Call of Duty®: Mobile turns 1! Introducing Anniversary, the Latest Season and a Celebration of this Milestone Event*. https://www.callofduty.com/blog/2020/10/Call-of-Duty-Mobile-Introducing-Anniversary
136. Noel, D. (2020d, November 10). *Night Descends on Call of Duty®: Mobile in Going Dark, the Latest Season Launching November 11.*

https://www.callofduty.com/blog/2020/11/Night-Descends-on-Call-of-Duty-Mobile-in-Going-Dark
137. Noel, D. (2020e, December 18). *Blast into the Holidays on Call of Duty®: Mobile Winter War, the Latest Season Launching December 21*. https://www.callofduty.com/blog/2020/12/call-of-duty-mobile-season-13-winter-war-announcement
138. Noel, D. (2021a, July 26). *Announcement: Turn Up The Heat in Season Six of Call of Duty®: Mobile*. https://www.callofduty.com/blog/2021/07/Turn-Up-the-Heat-in-Season-Six-of-Call-of-Duty-Mobile
139. Noel, D. (2021b, September 22). *Call of Duty®: Mobile Celebrates its second anniversary with a Major Update in Season 8*. https://www.callofduty.com/blog/2021/09/call-of-duty-mobile-season-8-second-anniversary-blackout
140. Noel, D. (2022, November 17). *Call of Duty®: Warzone™ Mobile Pre-Orders Now Available for iOS on the App Store*. https://www.callofduty.com/blog/2022/11/call-of-duty-warzone-mobile-pre-order-now-available-app-store
141. Nyakurukwa, K., & Seetharam, Y. (2023). The evolution of studies on social media sentiment in the stock market: Insights from bibliometric analysis. In *Scientific African* (Vol. 20). Elsevier B.V. https://doi.org/10.1016/j.sciaf.2023.e01596
142. Oncharoen, P., & Vateekul, P. (2018). Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators. *2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, 19–24. https://doi.org/10.1109/ICAICTA.2018.8541310
143. Pagolu, V. S., Reddy, K. N., Panda, G., & Majhi, B. (2016). Sentiment analysis of Twitter data for predicting stock market movements. *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, 1345–1350. https://doi.org/10.1109/SCOPES.2016.7955659
144. Perrin, A., Duggan, M., Rainie, L., Smith, A., Greenwood, S., Porteus, M., & Page, D. (2015). Social Media Usage: 2005-2015. In *Pew Research Center*. www.pewresearch.org/internet
145. Petukhova, A., & Fachada, N. (2022). TextCL: A Python package for NLP preprocessing tasks. *SoftwareX*, *19*. https://doi.org/10.1016/j.softx.2022.101122
146. Piñeiro-Chousa, J., López-Cabarcos, M. Á., Pérez-Pico, A. M., & Caby, J. (2023). The influence of Twitch and sustainability on the stock returns of video game companies: Before and after COVID-19. *Journal of Business Research*, *157*. https://doi.org/10.1016/j.jbusres.2022.113620
147. Piñeiro-Chousa, J., Vizcaíno-González, M., & Pérez-Pico, A. M. (2017). Influence of Social Media over the Stock Market. *Psychology and Marketing*, *34*(1), 101–108. https://doi.org/10.1002/mar.20976
148. Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social Media and Society*, *7*(2). https://doi.org/10.1177/20563051211019004

149. Puente, H., & Tosca, S. (2013). *Proceedings of DiGRA 2013: DeFragging Game Studies. The Social Dimension of Collective Storytelling in Skyrim*.
150. Rahman, A. M. M., Uddin, A., & Wang, G. G. (2023). HODL: The Hold of Reddit Over the Stock Market. *Proceedings - 2023 IEEE 39th International Conference on Data Engineering Workshops, ICDEW 2023*, 36–39. https://doi.org/10.1109/ICDEW58674.2023.00010
151. Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The Effects of Twitter Sentiment on Stock Price Returns. *PLOS ONE*, *10*(9), e0138441. https://doi.org/10.1371/journal.pone.0138441
152. Reddit. (2024a). *Developer Platform & Accessing Reddit Data*. Reddit. https://support.reddithelp.com/hc/en-us/articles/14945211791892-Developer-Platform-Accessing-Reddit-Data
153. Reddit. (2024b, June 19). *Top Communities*. Reddit. https://www.reddit.com/best/communities/1/
154. Reddit. (2024c, August 22). *r/datascience*. Reddit. https://www.reddit.com/r/datascience/
155. Reddit. (2024d, August 26). *r/gaming list-sorted-by-subscribers*. Reddit. https://www.reddit.com/r/gaming/wiki/list-sorted-by-subscribers/
156. Reichenbach, F., & Walther, M. (2023). Financial recommendations on Reddit, stock returns and cumulative prospect theory. *Digital Finance*, *5*(2), 421–448. https://doi.org/10.1007/s42521-023-00084-y
157. Rockstar Games. (2020a, May 14). *GTAV: Premium Edition Available Free on the Epic Games Store Until May 21st*. https://www.rockstargames.com/newswire/article/o349k55255a257/GTAV-Premium-Edition-Available-Free-on-the-Epic-Games-Store-Until-May-
158. Rockstar Games. (2020b, August 11). *Los Santos Summer Special in GTA Online*. https://www.rockstargames.com/newswire/article/1748koo9oo3557/Los-Santos-Summer-Special-in-GTA-Online
159. Rockstar Games. (2020c, December 1). *Red Dead Redemption 2 Title Update 1.24 Notes (PS4 / Xbox One / PC / Stadia)*. https://support.rockstargames.com/articles/360060216733/Red-Dead-Redemption-2-Title-Update-1-24-Notes-PS4-Xbox-One-PC-Stadia
160. Rockstar Games. (2020d, December 15). *Out now: The Cayo Perico Heist*. https://www.rockstargames.com/newswire/article/1748koo9oo583k/out-now-the-cayo-perico-heist
161. Rockstar Games. (2021a, July 13). *Red Dead Redemption 2 Title Update 1.25 Notes (PS4 / Xbox One / PC / Stadia)*. https://support.rockstargames.com/articles/4403603010579/Red-Dead-Redemption-2-Title-Update-1-25-Notes-PS4-Xbox-One-PC-Stadia
162. Rockstar Games. (2021b, July 15). *Los Santos Tuners Coming July 20*. https://www.rockstargames.com/newswire/article/25o1829o5o4993/los-santos-tuners-coming-july-20

163. Rockstar Games. (2021c, December 15). *GTA Online: The Contract - Out Now*. https://www.rockstargames.com/newswire/article/ak759k91592728/gta-online-the-contract-out-now
164. Rockstar Games. (2022a, March 15). *Grand Theft Auto V and GTA Online Out Now on PlayStation 5 and Xbox Series X|S*. https://www.rockstargames.com/newswire/article/4ka18ao8a4o227/grand-theft-auto-v-and-gta-online-out-now-on-playstation-5-and-xbox-se
165. Rockstar Games. (2022b, July 21). *The Criminal Enterprises, Coming July 26 to GTA Online*. https://www.rockstargames.com/newswire/article/25k44k95aa452a/the-criminal-enterprises-coming-july-26-to-gta-online
166. Rockstar Games. (2022c, December 13). *GTA Online: Los Santos Drug Wars Now Available*. https://www.rockstargames.com/newswire/article/75258k488okk33/gta-online-los-santos-drug-wars-now-available
167. Rockstar Games. (2023a, June 13). *GTA Online: San Andreas Mercenaries Now Available*. https://www.rockstargames.com/newswire/article/175k8294o31ooo/gta-online-san-andreas-mercenaries-now-available
168. Rockstar Games. (2023b, December 12). *GTA Online: The Chop Shop Now Available*. https://www.rockstargames.com/newswire/article/75k3149ko747ka/gta-online-the-chop-shop-now-available
169. Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P.-M. (2023). A review on sentiment analysis from social media platforms. *Expert Systems with Applications*, *223*, 119862. https://doi.org/10.1016/j.eswa.2023.119862
170. Ruggles, C., Wadley, G., & Gibbs, M. R. (2005). Online Community Building Techniques Used by Video Game Developers. *LNCS*, *3711*, 114–125.
171. Sajithra, K., & Rajindra, P. (2013). Social Media – History and Components. *IOSR Journal of Business and Management (IOSR-JBM)* , *7*(1), 69–74.
172. Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2017). A Survey of Text Mining in Social Media: Facebook and Twitter Perspectives. *Advances in Science, Technology and Engineering Systems Journal*, *2*(1), 127–133. https://doi.org/10.25046/aj020115
173. Saltz, J. S., & Hotz, N. (2020). Identifying the most Common Frameworks Data Science Teams Use to Structure and Coordinate their Projects. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, 2038–2042. https://doi.org/10.1109/BigData50022.2020.9377813
174. SEC. (2023, July 19). *ACTIVISION BLIZZARD ANNOUNCES SECOND QUARTER 2023 FINANCIAL RESULTS*. https://www.sec.gov/Archives/edgar/data/718877/000162828023025102/atvi63023ex991prtables.htm
175. Semenova, V., Winkler, J., Van Der Ploeg, R., Farmer, J. D., Dong, X., Bond, S., Ditraglia, F., Sheppard, K., Scholl, M., Pougué-Biyong, J., Strauss, I., Yang, J., Heinrich, T., Lafond, F., Winkler, M., Moran, J., Ravasan, F., Bordalo, P., Hommes,

C., … Lambiotte, R. (2021). *Reddit's self-organised bull runs: Social contagion and asset prices*.
176. Shrestha, M. B., & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *Journal of Finance and Data Science*, *4*(2), 71–89. https://doi.org/10.1016/j.jfds.2017.11.001
177. Sirani, J. (2023, September 21). *Cyberpunk 2077: Phantom Liberty – Release Date, Gameplay Updates, Story Details, and Latest News*. IGN. https://www.ign.com/articles/cyberpunk-2077-phantom-liberty-release-date-gameplay-changes-story-news
178. Solon, O. (2021, July 28). *Activision Blizzard workers walk out after sexual harassment lawsuit*. NBC News. https://www.nbcnews.com/tech/tech-news/activision-blizzard-workers-walk-sexual-harassment-lawsuit-rcna1525
179. Sousa, M. G., Sakiyama, K., Rodrigues, L. D. S., Moraes, P. H., Fernandes, E. R., & Matsubara, E. T. (2019). BERT for stock market sentiment analysis. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, *2019-November*, 1597–1601. https://doi.org/10.1109/ICTAI.2019.00231
180. Souza, L. L. F. de, & Freitas, A. A. F. de. (2017). Consumer behavior of electronic games' players: a study on the intentions to play and to pay. *Revista de Administração*, *52*(4), 419–430. https://doi.org/10.1016/j.rausp.2017.08.004
181. Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. *European Financial Management*, *20*(5), 926–957. https://doi.org/10.1111/j.1468-036X.2013.12007.x
182. Statista. (2023, July). *Video game industry worldwide*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/study/136728/video-game-industry-worldwide/
183. Suh, C., & Lee, B. (2011). *An Analysis of Events in Online Game Industry and Stock Price Reactions*. http://aisel.aisnet.org/pacis2011/185
184. Sul, H. K., Dennis, A. R., & Yuan, L. I. (2017). Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns. *Decision Sciences*, *48*(3), 454–488. https://doi.org/10.1111/deci.12229
185. Sun, T., Wang, J., Zhang, P., Cao, Y., Liu, B., & Wang, D. (2017). Predicting Stock Price Returns Using Microblog Sentiment for Chinese Stock Market. *Proceedings - 2017 3rd International Conference on Big Data Computing and Communications, BigCom 2017*, 87–96. https://doi.org/10.1109/BIGCOM.2017.59
186. Surminski, N. (2023). *The role of video game quality in financial markets* (Master's Thesis). Uppsala University.
187. Take Two Interactive. (2022, September 9). *Answer The Call: NBA® 2K23 Now Available Worldwide*. https://www.take2games.com/ir/news/answer-call-nbar-2k23-now-available-worldwide
188. Take Two Interactive. (2023, September 8). *See You on the Court: NBA® 2K24 Now Available Worldwide*. https://www.take2games.com/ir/news/see-you-court-nbar-2k24-now-available-worldwide

189. Take-Two Interactive. (2024). *Games*. Take-Two Interactive. https://www.take2games.com/
190. Take-Two-Interactive. (2022, May 23). *Take-Two Interactive Software, Inc. Completes Combination with Zynga Inc.* Take-Two-Interactive. https://www.take2games.com/ir/news/take-two-interactive-software-inc-completes-combination-zynga
191. Teoh, T.-T., Lim, W. T., Koh, K. W., Soh, J. J., Tan, T., Liu, S. Y., & Nguwi, Y.-Y. (2019). From Technical Analysis to Text Analytics: Stock and Index Prediction with GRU. *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, 496–500. https://doi.org/10.1109/CIS-RAM47153.2019.9095772
192. Thomala, L. L. (2024, May 21). *Revenue of NetEase Inc. from 2017 to 2023, by segment*. Statista. https://www-statista-com.nukweb.nuk.uni-lj.si/statistics/1118186/netease-revenue-by-segment/
193. Trading View. (2024a). *NEXON CO LTD revenue breakdown overview*. Trading View. https://www.tradingview.com/symbols/TSE-3659/financials-revenue/
194. Trading View. (2024b, September 12). *World Stocks Overview*. Trading View. https://www.tradingview.com/markets/world-stocks/
195. Ubisoft. (2024). *Games*. Ubisoft. https://www.ubisoft.com/en-us/games
196. Vaudour, F., & Heinze, A. (2020). Software as a service: Lessons from the video game industry. *Global Business and Organizational Excellence*, *39*(2), 31–40. https://doi.org/10.1002/joe.21982
197. Vijayarani, S., Ilamathi, J., & Nithya, M. (2015). Preprocessing Techniques for Text Mining-An Overview. *International Journal of Computer Science & Communication Networks*, *5*, 7–16. https://www.researchgate.net/publication/339529230
198. Wallach, O. (2020, November 27). *The history of the gaming industry in one chart*. World Economic Forum. https://www.weforum.org/agenda/2020/11/gaming-games-consels-xbox-play-station-fun/
199. Wang, C., & Luo, B. (2021). Predicting $GME Stock Price Movement Using Sentiment from Reddit r/wallstreetbets. *Proceedings of the Third Workshop on Financial Technology and Natural Language Processing*, 22–30. https://rapidapi.com/apidojo/api/yahoo-finance1
200. Wang, M., & Hu, F. (2021). The application of nltk library for python natural language processing in corpus research. *Theory and Practice in Language Studies*, *11*(9), 1041–1049. https://doi.org/10.17507/tpls.1109.09
201. Watchful1. (2024, February 5). *Subreddit comments/submissions 2005-06 to 2023-12*. Academic Torrents. https://academictorrents.com/details/56aa49f9653ba545f48df2e33679f014d2829c10
202. Wawrowski, B., & Otola, I. (2020). Social media marketing in creative industries: How to use social media marketing to promote computer games? *Information (Switzerland)*, *11*(5). https://doi.org/10.3390/INFO11050242

203. Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a Standard Process Model for Data Mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*.
204. Wolf, M. J. P. (2012). *Before the crash : early video game history*.
205. X. (2024). *X API*. X. https://developer.x.com/en/products/x-api
206. Xiao, L. Y. (2022). Which Implementations of Loot Boxes Constitute Gambling? A UK Legal Perspective on the Potential Harms of Random Reward Mechanisms. In *International Journal of Mental Health and Addiction* (Vol. 20, Issue 1, pp. 437–454). Springer. https://doi.org/10.1007/s11469-020-00372-3
207. Xiong, G., & Bharadwaj, S. (2014). Prerelease Buzz Evolution Patterns and New Product Performance. *Marketing Science*, *33*(3), 401–421. https://doi.org/10.1287/mksc.2013.0828
208. Xu, K., Pang, Y., & Han, J. (2021). Dynamic Cross-Correlation between Online Sentiment and Stock Market Performance: A Global View. *Discrete Dynamics in Nature and Society*, *2021*. https://doi.org/10.1155/2021/6674379
209. Yahoo Finance. (2024a, September 10). *Video Game Stocks*. Yahoo Finance. https://finance.yahoo.com/u/yahoo-finance/watchlists/video-game-stocks/
210. Yahoo Finance. (2024b, September 12). *Yahoo Finance Lookup*. Yahoo Finance. https://finance.yahoo.com/lookup
211. Yee, N. (2006). Motivations for Play in Online Games. *CYBERPSYCHOLOGY & BEHAVIOR*, *9*(6). www.mmogchart.com/.
212. Yee, N. (2016). The Gamer Motivation Profile. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, 2–2. https://doi.org/10.1145/2967934.2967937
213. Yu, J. (2023, April 11). *When does Modern Warfare 2 and Warzone 2.0 season 3 start?* https://www.polygon.com/guides/2023/4/11/23678705/modern-warfare-warzone-2-season-3-release-time
214. Zackariasson, P., & Wilson, T. L. (2010). Paradigm shifts in the video game industry. *Competitiveness Review*, *20*(2), 139–151. https://doi.org/10.1108/10595421011029857
215. Zendle, D., & Cairns, P. (2019). Correction: Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PLOS ONE*, *14*(3), e0214167. https://doi.org/10.1371/journal.pone.0214167
216. Zhu, F., Xiaoquan, &, & Zhang, M. (2010). Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing*, *74*, 133–148.

APPENDICES

Appendix 1: Povzetek (Summary in Slovene language)

Finančni trgi so že od nekdaj pod vplivom različnih dejavnikov, ki segajo vse od letnih poročil do makroekonomskih politik (Cutler in drugi, 1998). Vzpon digitalne dobe je v ospredje prinesel nov in vpliven dejavnik: družbena omrežja. Posledično je povezava med cenami delnic in razpoloženjem na družbenih omrežjih postala predmet povečanega zanimanja med raziskovalci. Pri tem lahko opazimo dva trenda, ki odpirata vrata novim raziskavam. Prvič, dosedanje raziskave kot glavni vir za razpoloženjske podatke uporabljajo Twitter (zdaj X), finančne bloge, ali novice (Garg in Tiwari, 2021). In drugič, avtorji prihajajo do mešanih ugotovitev. Nekateri, kot Ranco in drugi (2015), poročajo o močni korelaciji, medtem ko drugi, kot Sprenger in drugi (2014), te povezave ne potrjujejo. Za razliko od ostalih platform ima Reddit rigidno strukturo, kar ga naredi primernega za raziskovanje specifičnih spletnih skupnosti. Ena takšnih je skupnost igralcev videoiger, kjer sta zadovoljstvo in vključenost potrošnikov ključna faktorja za uspešnost izdelka (Ahmad in drugi, 2017; Yee, 2016). Naša raziskava se tako oddaljuje od finančnih forumov in se osredotoča na vprašanje, ali se razpoloženje potrošnikov odraža v cenah delnic razvijalcev videoiger.

Namen magistrskega dela je poglobiti razumevanje razmerja med zadovoljstvom kupca na družbenih omrežjih in cenami delnic podjetij ter s tem prispevati k zapolnitvi raziskovalne vrzeli tako na področju panoge videoiger kot tudi na področju pomena razpoloženja kupcev in njegove povezave s finančnim stanjem podjetij. Cilji raziskave so tako: (1) pridobiti relevantne podatke in jih pripraviti za analizo razpoloženja v okviru primernega modela za obdelavo podatkov; (2) izvesti potrebne transformacije za zajem finančnih, razpoloženjskih in čustvenih podatkov; (3) izvesti analizo razpoloženja, pri čemer je treba upoštevati tako ocene razpoloženja kot ocene čustev v kontekstu preučevanih podjetij; (4) identificirati pomembne lokalne dogodke, ki bi lahko razkrili močnejše povezave med podatki o razpoloženju in finančnimi podatki; (5) izračunati korelacijo med podatkovnimi nizi in odkriti povezave na lokalnem ter globalnem nivoju; (6) izvesti Grangerjev test vzročnosti za interakcije med finančnimi in razpoloženjskimi podatki ter finančnimi in čustvenimi podatki, za globalna in lokalna obdobja; (7) ugotoviti naravo povezave med podatkovnimi nizi, kadar ta obstaja na dolgoročni ravni ali zgolj v okviru posameznih dogodkov.

Metodološki pristop sledi CRISP-DM modelu, ki sta ga razvila Wirth in Hipp (2000). Najprej smo izbrali podjetja in skupnosti na platformi Reddit, katerih podatke smo analizirali s programskim jezikom Python. Finančne podatke smo pridobili iz Nasdaqa in spletnega portala Investing.com, medtem ko smo podatke iz platforme Reddit pridobili iz podatkovne baze Pushshift (Baumgartner in drugi, 2020). Glede na znane težave, ki jih izpostavijo Baldwin in drugi (2013), smo sledili nekaterim korakom za čiščenje besedila po Chai (2023). Za analizo razpoloženja smo uporabili leksikon VADER, za klasifikacijo čustev pa model RoBERTa (Hutto in Gilbert, 2014; Liu in drugi, 2019). Statičnost podatkov smo potrdili z uporabo testa ADF. Nato smo po vzoru Ranco in drugi (2015) uporabili korelacijsko analizo in Grangerjev test vzročnosti, in sicer tako za splošna časovna obdobja kot tudi za specifične dogodke – tako na ravni celotnih podjetij kot posameznih spletnih skupnosti.

Rezultati kažejo, da korelacija in vzročnost ponujata le omejen pogled v dolgoročne trende, vendar lahko izpostavita jasne povezave pri analizi specifičnih dogodkov na ravni posameznih skupnosti v krajšem časovnem okvirju. Implikacije raziskave lahko gledamo iz dveh vidikov. Prvič, prispevamo k boljšemu razumevanju vpliva razpoloženja na gibanje cen delnic znotraj panoge videoiger. Opredelimo postopek za izbiro ustreznih podjetij v panoge videoiger za analizo razpoloženja na socialnih omrežjih, s čimer izboljšamo pristope v preteklih delih (Mertová, 2023). Poleg tega, z upoštevanjem dogajanja v panogi, nadgradimo delo Ranco in drugi (2015) za izbiro pomembnih lokalnih dogodkov. Za povrh naši rezultati, ki izhajajo iz nefinančnih diskurzov, ponujajo svež vpogled v obravnavano problematiko. Drugič, naše ugotovitve kažejo, da lahko podjetja iz tovrstnih analiz pridobijo vpogled v splošno zadovoljstvo svojih strank in prepoznajo, pri katerih izdelkih so potrebne spremembe. Hkrati analiza lokalnih dogodkov omogoča vpogled v odzive igralcev na pretekle igre ter njihov vpliv na gibanje cen delnic, kar lahko služi kot osnova za izogibanje napakam v prihodnosti.

Kljub temu pa ima raziskava nekatere omejitve. Razprave na platformi Reddit ne zajamejo nujno celotnega spektra javnega mnenja, poleg tega pa je jezik, poln žargona o videoigrah, izziv za konvencionalna orodja za analizo sentimenta. Naš 30-dnevni časovni okvir za analizo dogodkov (15 dni pred in po vsakem dogodku) prav tako morda ne zajame vseh premikov v razpoloženju.

Appendix 2: List of potential candidates

Candidate*	C1**	C2***	Market cap (in Billion USD) ****
Activision Blizzard	X	X	64.43
NetEase	X	X	60.35
Nintendo	X		60.01
Electronic Arts	X	X	37.36
Roblox		X	35.10
Take-Two Interactive	X	X	22.31
NEXON	X	X	19.92
BiliBili	X		15.39
Capcom	X		6.35
Ubisoft	X	X	5.32
CD Projekt Red	X	X	4.52
Playtika		X	3.99
Skillz	X		2.25
Gravity	X	X	0.53
DoubleDown Interactive			0.46
Sohu.com	X		0.45
PLAYSTUDIOS			0.45
Golden Matrix Group	X		0.28
GD Culture Group	X		0.03
Motorsport Games		X	0.02
Gaxos.ai			0.003

Source: Adapted from Trading View (2024b); Yahoo Finance (2024b).

**Chosen candidates are highlighted*

***C1: The company has been publicly traded on the stock market for all the years of our observation (source: Yahoo Finance (2024b)).*

****C2: The company's revenue is based on video games (source: Trading View (2024b)).*

*****Calculated based on the average end-of-year market capitalisation for the 2020-2023 period, as shown in the formula below (formula was adjusted for companies not publicly traded in all years):*

$$\mu_{20-23} = \frac{P_{20} \times Q_{20} + P_{21} \times Q_{21} + P_{22} \times Q_{22} + P_{22} \times Q_{23} + P_{23} \times Q_{23}}{4}$$

μ - Average market capitalisation.

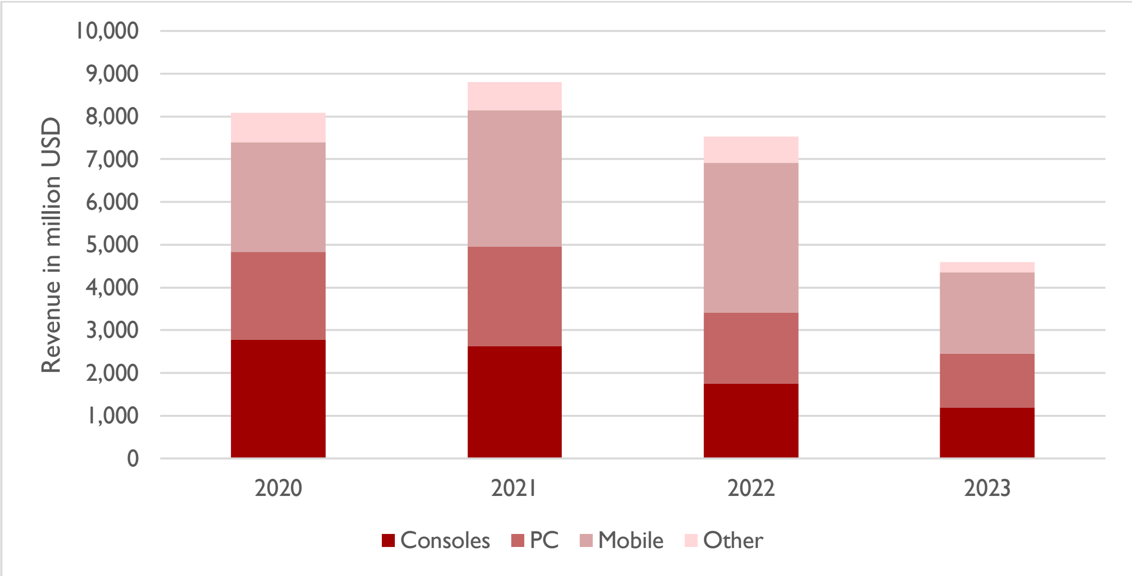
P – End-of-year share price.

Q – End-of-year share quantity.

Appendix 3: In-depth overview of the candidates

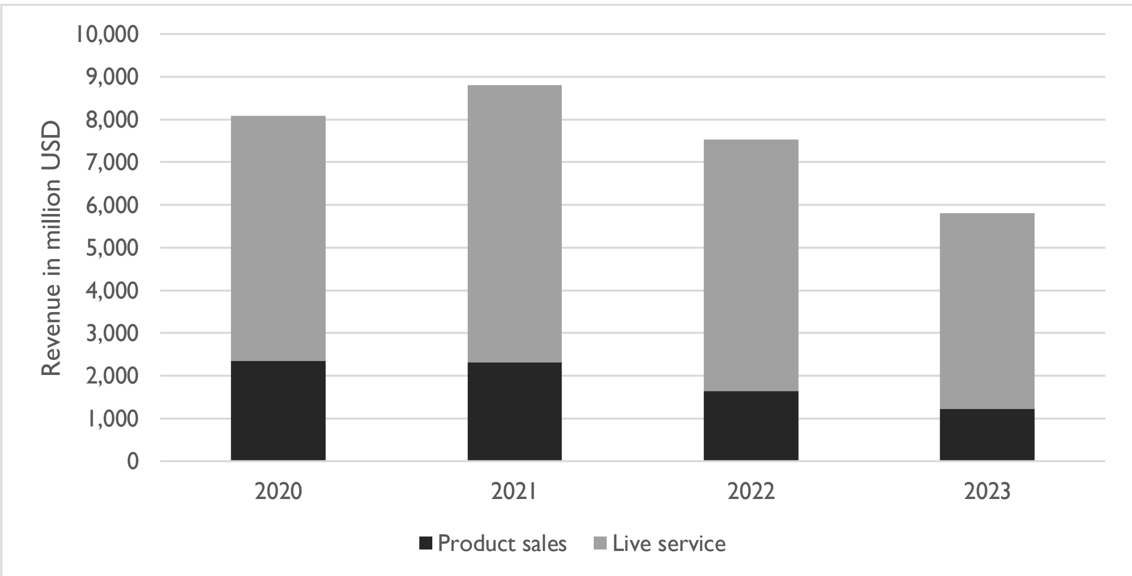
Activision Blizzard

*Activision Blizzard net revenue 2020-2023, by platform**



Source: Adapted from Clement (2024a); SEC (2023).

*Activision Blizzard net revenue 2020-2023, by composition**



Source: Adapted from Clement (2024a); SEC (2023).

**2023 data only consist of revenue in Q1 and Q2 SEC filings due to acquisition by Microsoft in Q3*

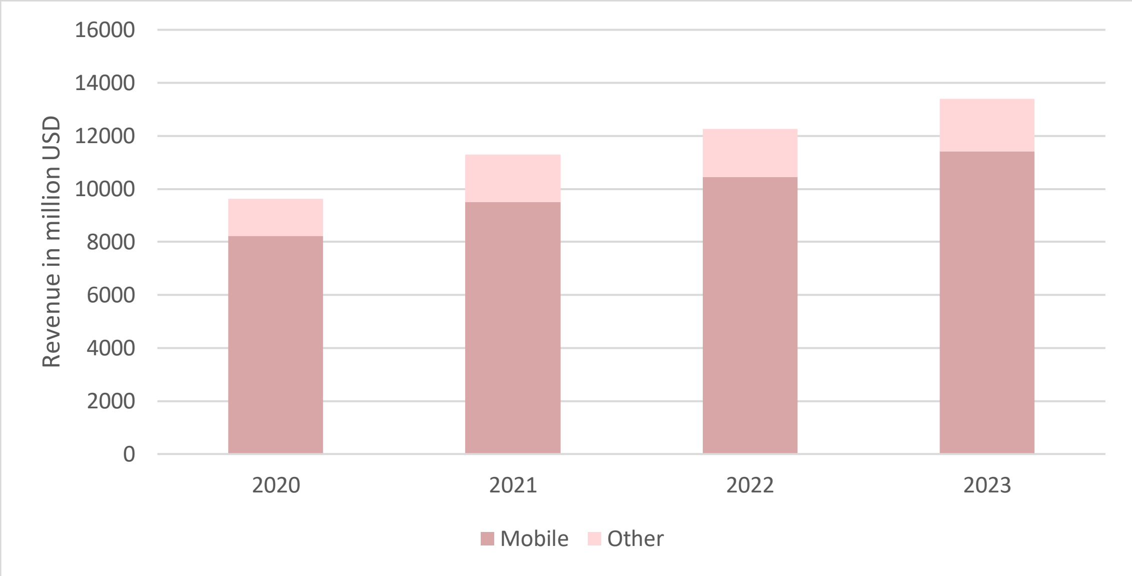
Notable releases by Activision Blizzard

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
World of Warcraft	PC	Subscription, Microtransactions	r/wow	2.800.000
Overwatch Overwatch 2	PC, Console	Fixed price, Microtransactions	r/Overwatch r/overwatch2	5.800.000 148.000
CoD:MWII	PC, Console	Fixed price, Microtransactions	r/ModernWarfareII	4.800.000
CoD:Warzone	PC. Console	Free-to-play, Subscription, Microtransactions	r/CODWarzone	1.400.000
CoD: Warzone Mobile	Mobile	Free-to-play, subscription, microtransactions	r/WarzoneMobile	28.000
CoD: Mobile	Mobile	Free-to-play, Subscription, Microtransactions	r/CallOfDutyMobile	338.000
Candy Crush Saga	Mobile	Free-to-play, Microtransactions	r/candycrush	14.000

Source: Adapted from Activision Blizzard (2024); Clement (2024a); Reddit (2024b).

NetEase

NetEase net revenue 2020-2023, by platform



Source: Adapted from Thomala (2024).

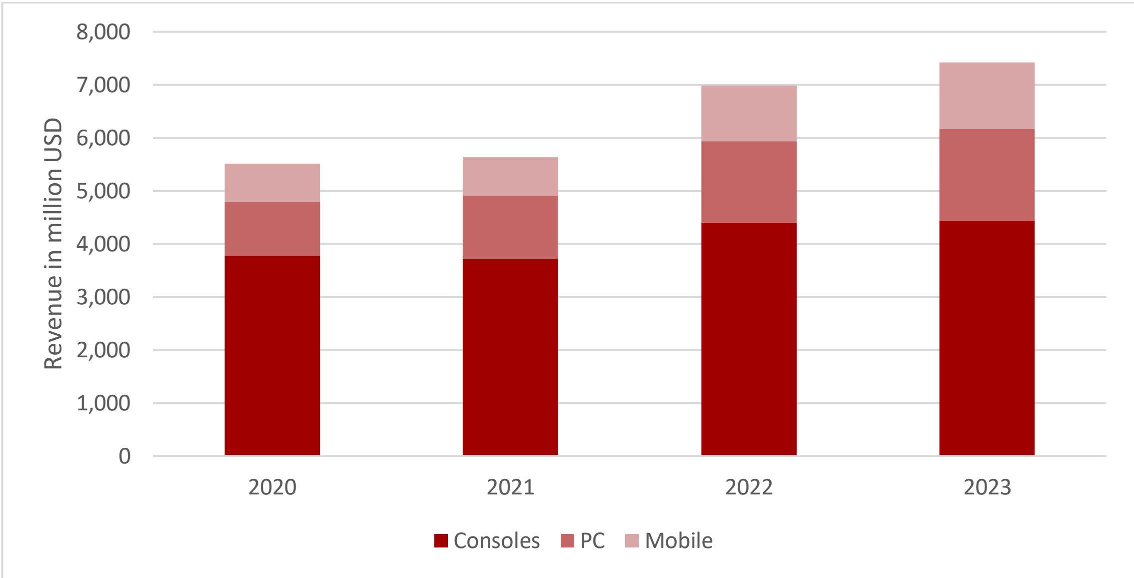
Notable releases NetEase

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
Blood Strike	Mobile	Free-to-play, Microtransactions	r/ProjectBloodstrikeBR	<10.000
Identity V	Mobile	Free-to-play, Microtransactions	r/IdentityV	67.000
Blood Strike MENA	Mobile	Free-to-play, Microtransactions	/	<10.000
Knives Out	Mobile	Free-to-play, Microtransactions	r/KnivesOutGame	<10.000
LifeAfter	Mobile	Free-to-play, Microtransactions	r/lifeafter	13.000
Onmoyoji Arena	Mobile	Free-to-play, Microtransactions	r/OnmyojiArena	12.000

Source: Adapted from NetEase Games (2024); Reddit (2024b); Thomala (2024).

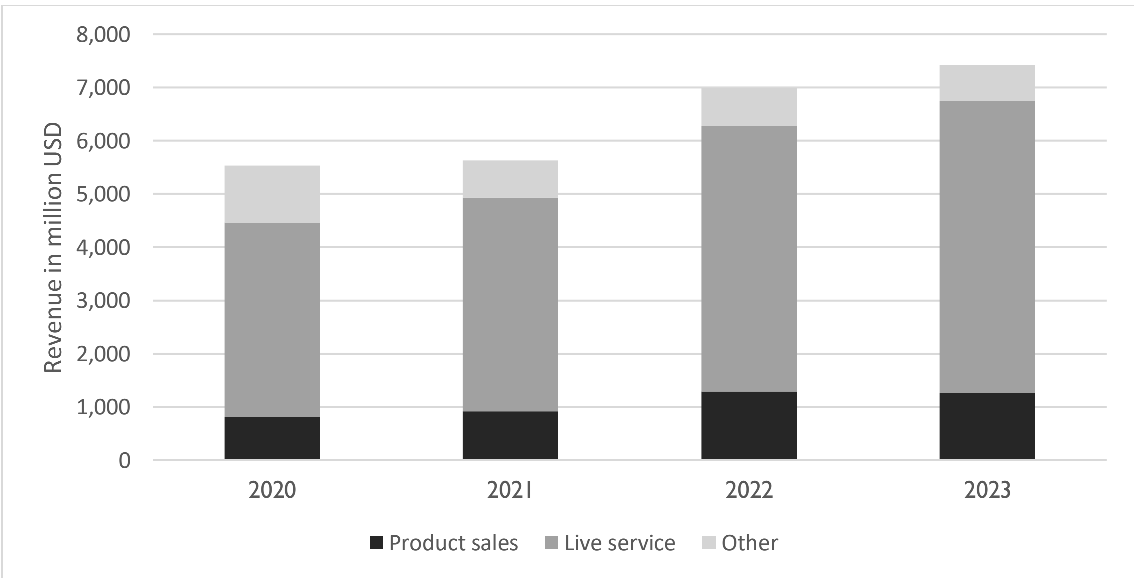
Electronic Arts

Electronic Arts net revenue 2020-2023, by platform



Source: Adapted from Clement (2024e).

Electronic Arts net revenue 2020-2023, by composition



Source: Adapted from Clement (2024e).

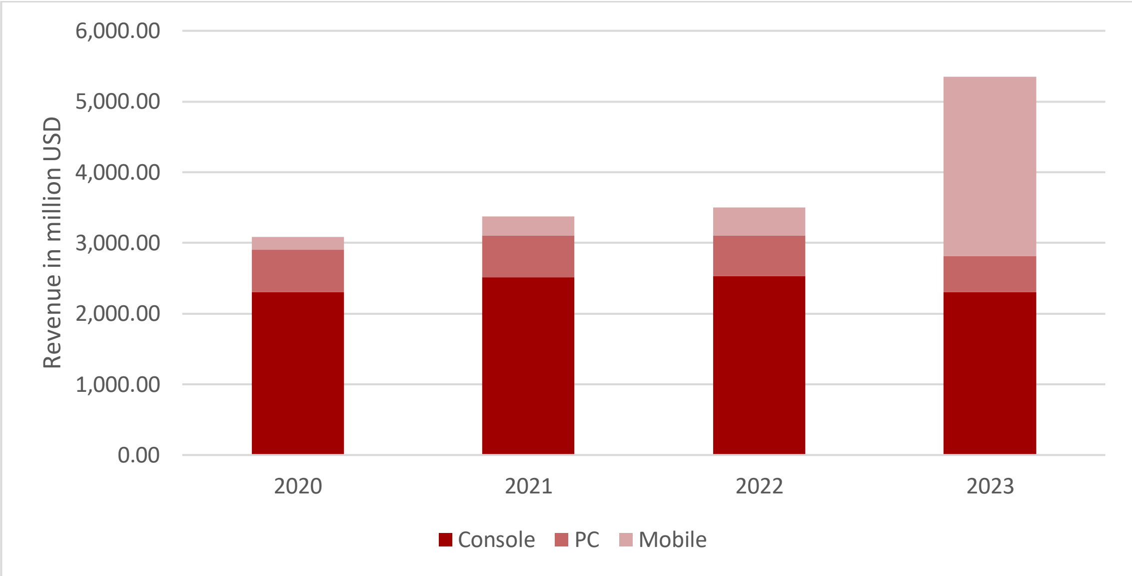
Notable releases by Electronic Arts

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
FIFA franchise	PC, Console	Fixed price, Microtransactions	r/EASportsFC	1.100.000
Madden franchise	PC, Console	Fixed price, Microtransactions	r/Madden	247.000
Battlefield 2042	PC, Console	Fixed price, DLC, Microtransactions	r/battlefield2042	238.000
Battlefield V	PC, Console	Fixed price, DLC, Microtransactions	r/BattlefieldV	252.000
Sims 4	PC	Fixed price, DLC Microtransactions	r/Sims4	1.400.00
Apex Legends	PC, Console	Free-to-play, Subscription, Microtransactions	r/apexlegends	2.900.000
Star Wars Battlefront	PC, Console	Fixed price, Microtransactions	r/StarWarsBattlefront	476.000
Star Wars: Galaxy of heroes	Mobile	Free-to-play, Microtransactions	r/SWGalaxyOfHeroes	140.000
FC Mobile	Mobile	Free-to-play, Microtransactions	r/FUTMobile	98.000

Source: Adapted from Clement (2024e); Electronic Arts (2024); Reddit (2024b).

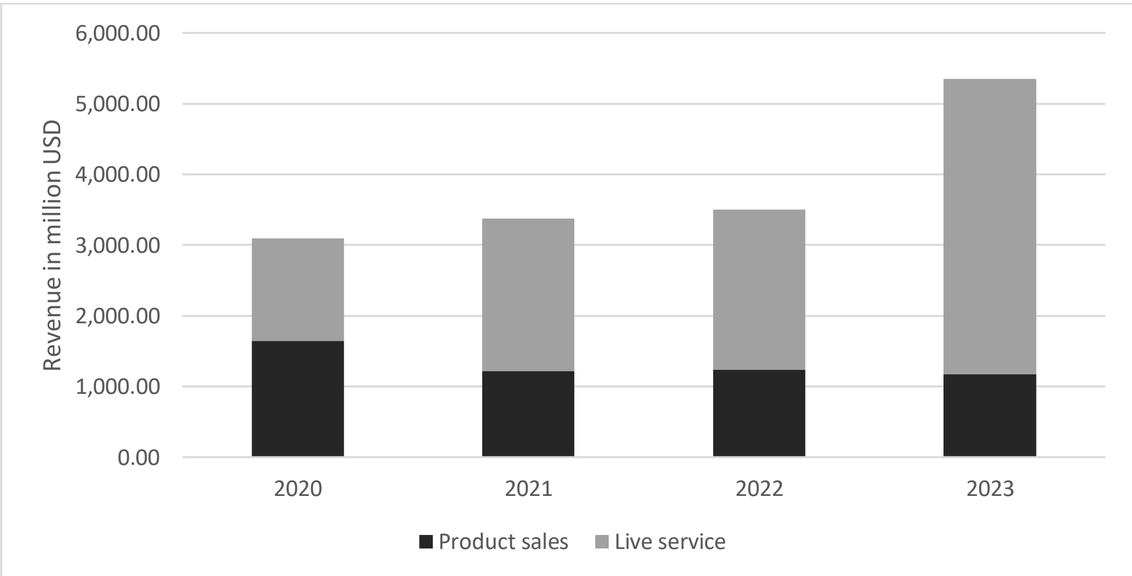
Take-Two Interactive

Take-Two Interactive net revenue 2020-2023, by platform



Source: Adapted from Clement (2024b).

Take-Two Interactive net revenue 2020-2023, by composition



Source: Adapted from Clement (2024b).

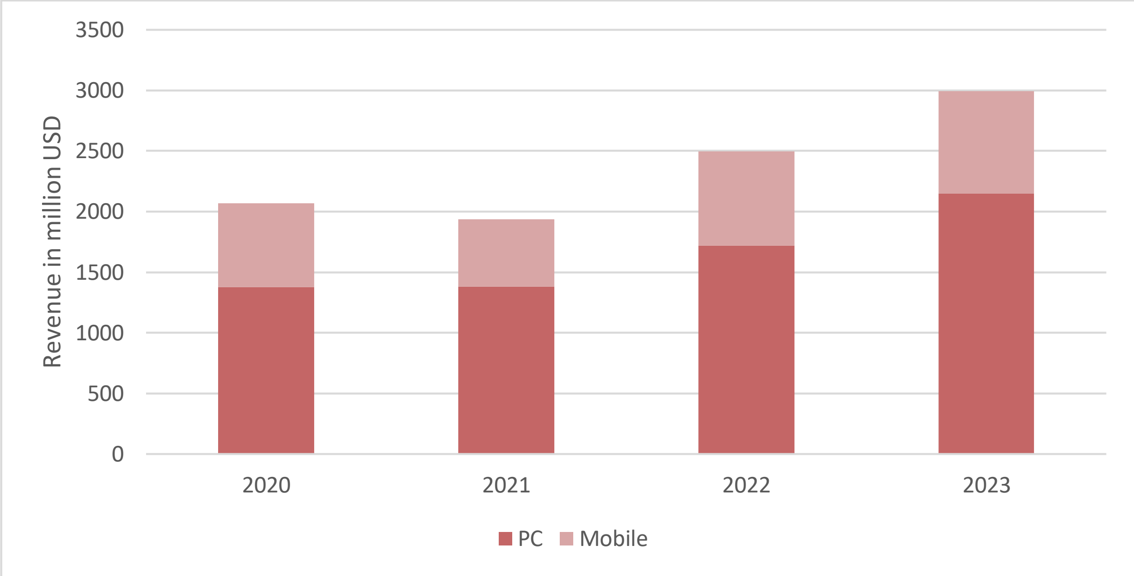
Notable releases by Take-Two Interactive

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
GTA V	PC, Console	Fixed price, Microtransactions	r/gtaonline r/GTAV	1.600.000 476.000
Read Dead Redemption 2	PC, Console	Fixed price, Microtransactions	r/reddeadredemption r/RedDeadOnline	1.900.000 426.000
NBA 2K franchise	PC, Console	Fixed price, Microtransactions	r/NBA2k	569.000
Borderlands 3	PC	Fixed price, DLC, Subscription, Microtransactions	r/borderlands3	413.000
Empires & Puzzles: Match-3 RPG	Mobile	Free-to-play, Microtransactions	r/EmpiresAndPuzzles	29.000
Merge Dragons!	Mobile	Free-to-play, Microtransactions	r/MergeDragons	46.000
Zynga Poker	Mobile	Free-to-play, Microtransactions	r/zyngapoker	<10.000
CSR 2 Drag Racing Car Games	Mobile	Free-to-play, Microtransactions	r/CSRRacing2	53.000

Source: Adapted from Clement (2024b); Reddit (2024b); Take-Two Interactive (2024).

NEXON

NEXON net revenue 2020-2023, by platform



Source: Adapted from Trading View (2024a).

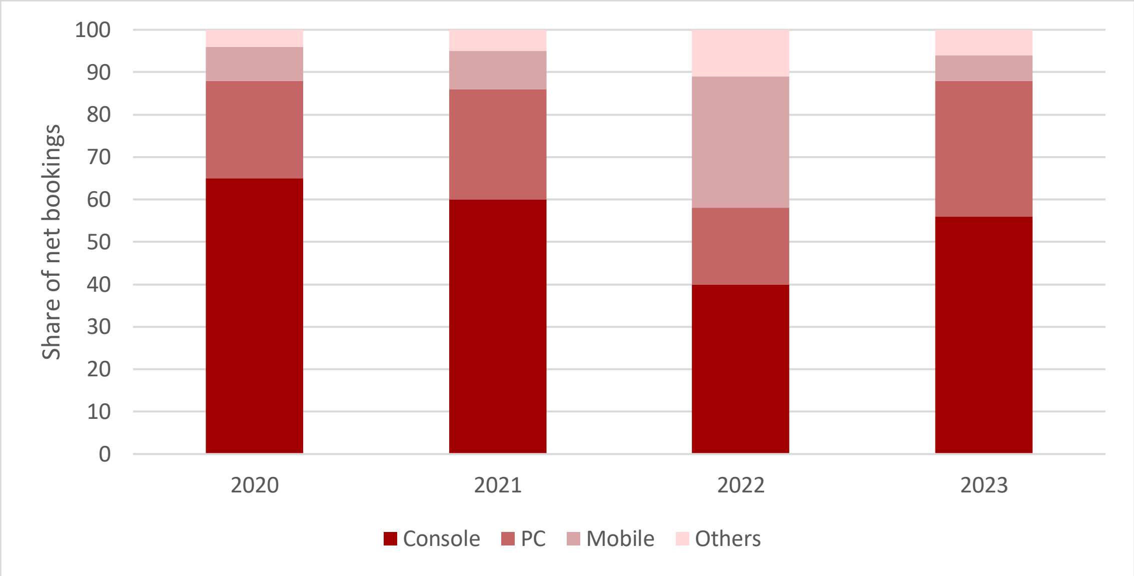
Notable releases by NEXON

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
MapleStory	PC	Free-to-play, Microtransactions	r/Maplestory	126.000
Mabingoi V	PC	Free-to-play, Microtransactions	r/Mabinogi	18.000
Vindictus	PC	Free-to-play, Microtransactions	r/Vindictus	16.000
THE FINALS	PC, Console	Free-to-play, Microtransactions	r/thefinals	134.000
MapleStory M	Mobile	Free-to-play, Microtransactions	r/MapleStoryM	30.000
Blue Arhive	Mobile	Free-to-play, Microtransactions	r/BlueArchive	121.000

Source: Adapted from NEXON (2024); Reddit (2024b).

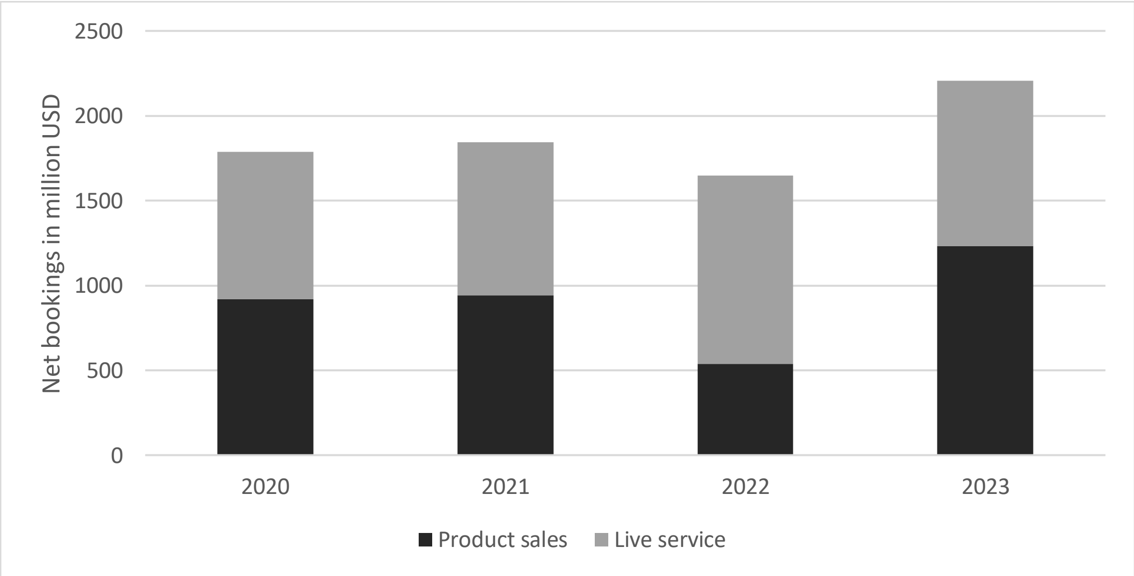
Ubisoft

Ubisoft share of net bookings by platform 2020-2023



Source: Adapted from Clement (2024c).

Ubisoft net bookings by composition 2020-2023



Source: Adapted from Clement (2024c).

Notable releases by Ubisoft

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
Assasin’s Creed franchise	PC, Console	Fixed price, DLC, Subscription, Microtransactions	r/assassinscreed	731.000
Far Cry franchise V	PC, Console	Fixed price, DLC, Subscription, Microtransactions	r/farcry	241.000
Tom Clancy's franchise	PC, Console	Fixed price, DLC, Subscription, Microtransactions	r/Rainbow6 r/GhostRecon r/thedivision	2.000.000 141.000 380.000
Brawlhalla	PC, Console, Mobile	Free-to-play, Microtransactions	r/Brawlhalla	175.000
Watch Dogs	PC, Console	Fixed price, DLC, Subscription, Microtransactions	r/watch_dogs	120.000
For Honor	PC, Console	Free-to-play, Microtransactions	r/forhonor	424.000

Source: Adapted from Clement (2024c); Reddit (2024b); Ubisoft (2024).

CD Projekt Red

Notable releases by CD Projekt Red

Game	Platform	Monetisation	Reddit	
			Subreddit	Members
Cyberpunk 2077	PC, Console	Fixed-price, DLC	r/cyberpunkgame	1.800.000
Witcher 3	PC, Console	Fixed-price, DLC	r/Witcher3	477.000

Source: Adapted from CD Projekt Red (2024); Reddit (2024b).

Appendix 4: Pushshifts Reddit Dataset Submissions and Comments dataset description

Submissions data description

Field	Description
id	The submission’s identifier, e.g., “5lcgjh” (String).
url	The URL that the submission is posting. This is the same with the permalink in cases where the submission is a self post. E.g., “https://www.reddit.com/r/AskReddit/”.
permalink	Relative URL of the permanent link that points to this specific submission, e.g., “/r/AskReddit/comments/5lcgj9/what did you think of the ending of rogue one/” (String).
author	The account name of the poster, e.g., “example username” (String).
created_utc	UNIX timestamp referring to the time of the submission’s creation, e.g., 1483228803 (Integer).
subreddit	Name of the subreddit that the submission is posted. Excludes the prefix /r/. E.g., ’AskReddit’ (String).
subreddit_id	The identifier of the subreddit, e.g., “t5_2qh1i” (String).
selftext	The text that is associated with the submission (String).
title	The title that is associated with the submission, e.g., “What did you think of the ending of Rogue One?” (String).
num_comments	The number of comments associated with this submission, e.g., 7 (Integer).
score	The score that the submission has accumulated. The score is the number of upvotes minus the number of downvotes. E.g., 5 (Integer). NB: Reddit fuzzes real score to prevent spam bots.
is_self	Flag that indicates whether the submission is a self post, e.g., true (Boolean).

To be continued

Submissions data description (cont.)

over_18	Flag that indicates whether the submission is Not-Safe-For-Work, e.g., false (Boolean).
distinguished	Flag to determine whether the submission is distinguished by moderators. “null” means not distinguished (String).
edited	Indicates whether the submission has been edited. Either a UNIX timestamp that indicates the submission was edited at, or “false” otherwise.
domain	The domain of the submission, e.g., self.AskReddit (String).
stickied	Flag indicating whether the submission is set as sticky in the subreddit, e.g., false (Boolean).
locked	Flag indicating whether the submission is currently closed to new comments, e.g., false (Boolean).
quarantine	Flag indicating whether the community is quarantined, e.g., false (Boolean).
hidden_score	Flag indicating if the submission’s score is hidden, e.g., false (Boolean).
retrieved_on	UNIX timestamp referring to the time we crawled the submission, e.g., 1483228803 (Integer).
author_flair_css_class	The CSS class of the author’s flair. This field is specific to the subreddit (String).
author_flair_text	The text of the author’s flair. This field is specific to the subreddit (String).

Source: Adapted from Baumgartner et al. (2020, p.833).

Comments data description

Field	Description
id	The comment's identifier, e.g., "dbumnq8" (String).
author	The account name of the poster, e.g., "example username" (String).
link_id	Identifier of the submission that this comment is in, e.g., "t3_5l954r" (String).
parent_id	Identifier of the parent of this comment, might be the identifier of the submission if it is a top-level comment or another comment, e.g., "t1_dbu5bpp" (String).
created_utc	UNIX timestamp referring to the time of the comment's creation, e.g., 1483228803 (Integer).
subreddit	Name of the subreddit where the comment is posted. Excludes the prefix /r/. E.g., 'AskReddit' (String).
subreddit_id	Identifier of the subreddit where the comment is posted, e.g., "t5_2qh1i" (String).
body	The comment's text, e.g., "This is an example comment" (String).
score	The score of the comment. The score is the number of upvotes minus the number of downvotes. Note that Reddit fuzzes the real score to prevent spam bots. E.g., 5 (Integer).
distinguished	Flag to determine whether the comment is made by moderators or admins. "null" means not distinguished (String).
edited	Flag indicating if the comment has been edited. Either the UNIX timestamp of the edit or "false" if unedited.
stickied	Flag indicating whether the comment is set as sticky in the subreddit, e.g., false (Boolean).
retrieved_on	UNIX timestamp referring to the time we crawled the comment, e.g., 1483228803 (Integer).

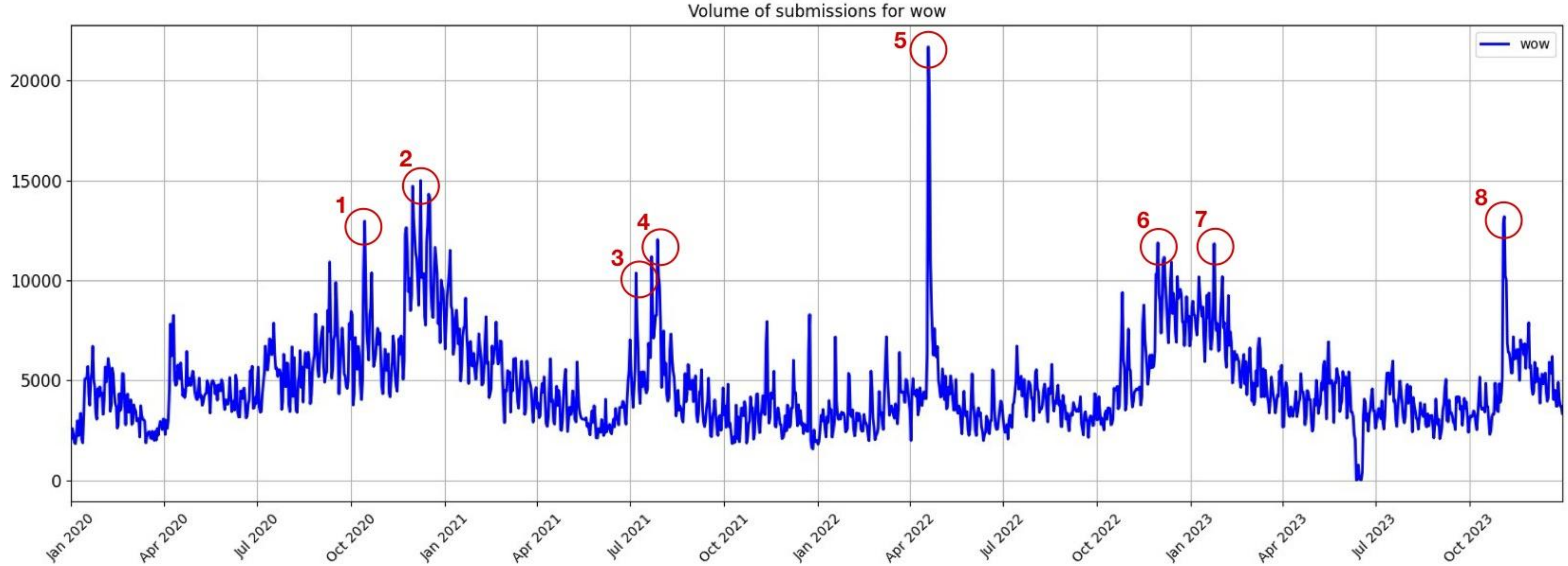
To be continued

Comments data description (cont.)

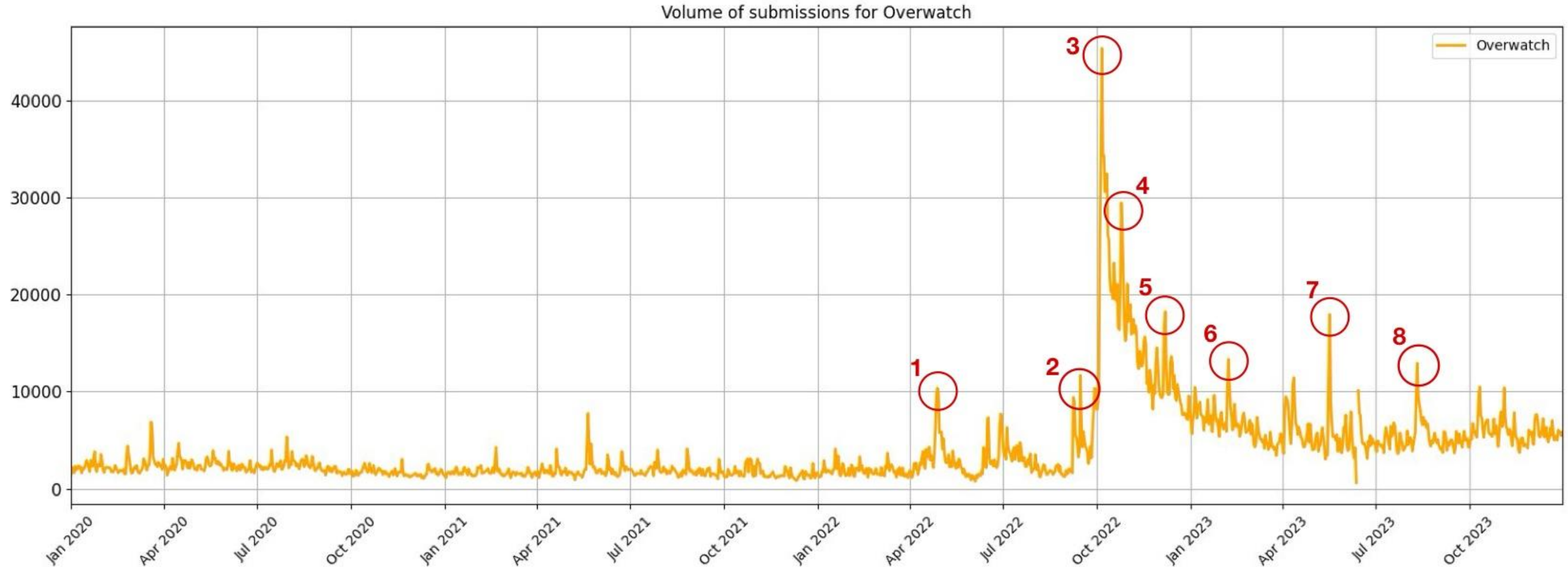
gilded	The number of times this comment received Reddit gold, e.g., 0 (Integer).
controversiality	Number that indicates whether the comment is controversial, e.g., 0 (Integer).
author_flair_css_class	The CSS class of the author's flair. This field is specific to the subreddit (String).
author_flair_text	The text of the author's flair. This field is specific to the subreddit (String).

Source: Adapted from Baumgartner et al. (2020, p. 834).

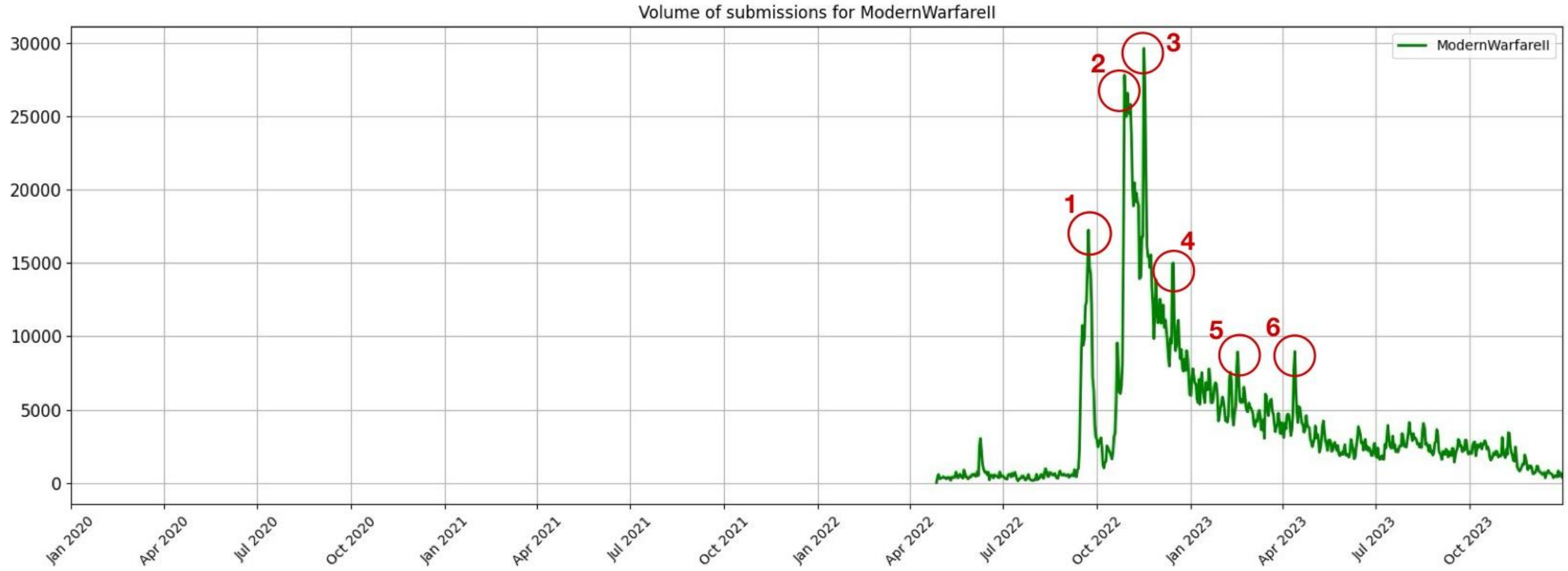
Appendix 5: Activision Blizzard volume of submissions



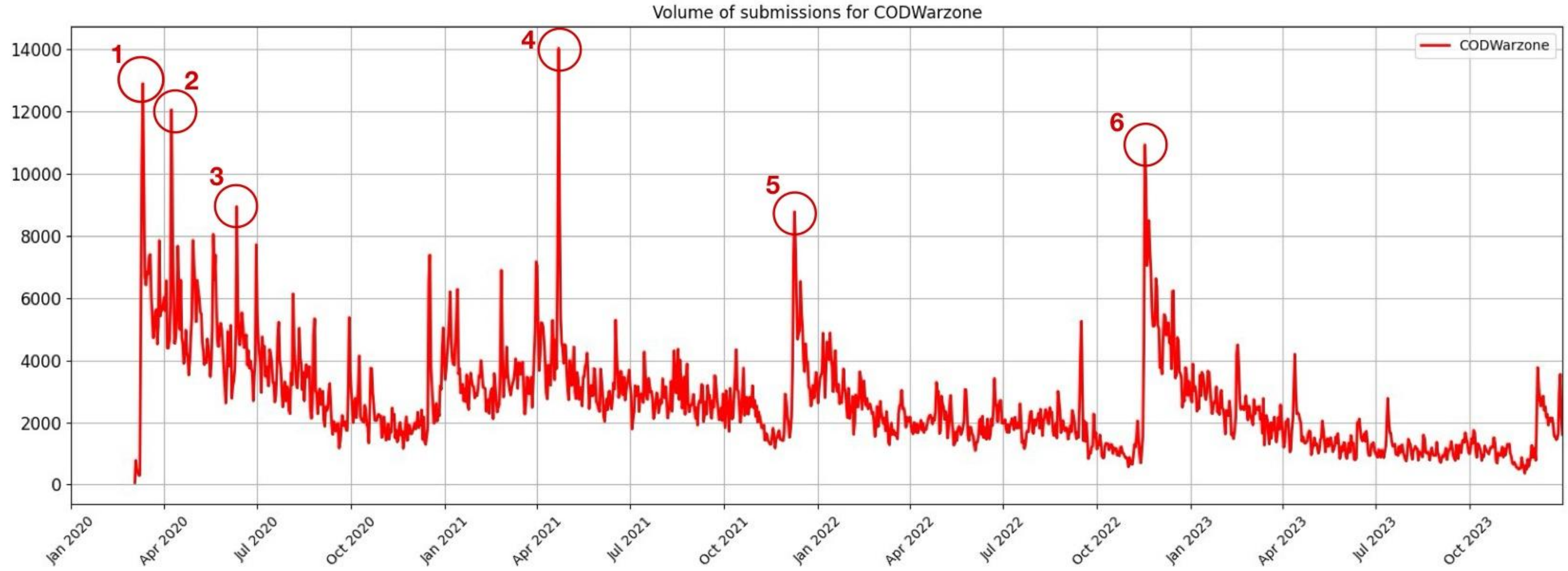
	Date	Event
(1)	13/10/2020	WoW: Shadowlands Pre-Expansion Patch (Blizzard Entertainment, 2020a).
(2)	23/11/2020	WoW: Shadowlands release (Blizzard Entertainment, 2020b).
(3)	22/07/2021	World of Warcraft players are hosting sit-in protests after Blizzard allegations (Marshall, 2021).
(4)	28/07/2021	Activision Blizzard workers walk out after sexual harassment lawsuit (Solon, 2021).
(5)	19/04/2022	WoW: Dragonflight revealed (Goslin, 2022).
(6)	28/11/2022	WoW: Dragonflight release (Blizzard Entertainment, 2022d).
(7)	24/01/2023	WoW: Dragonflight 10.0.5 Content Update (Blizzard Entertainment, 2023a).
(8)	03/11/2023	BlizzCon 2023 (Bankhurst, 2023).



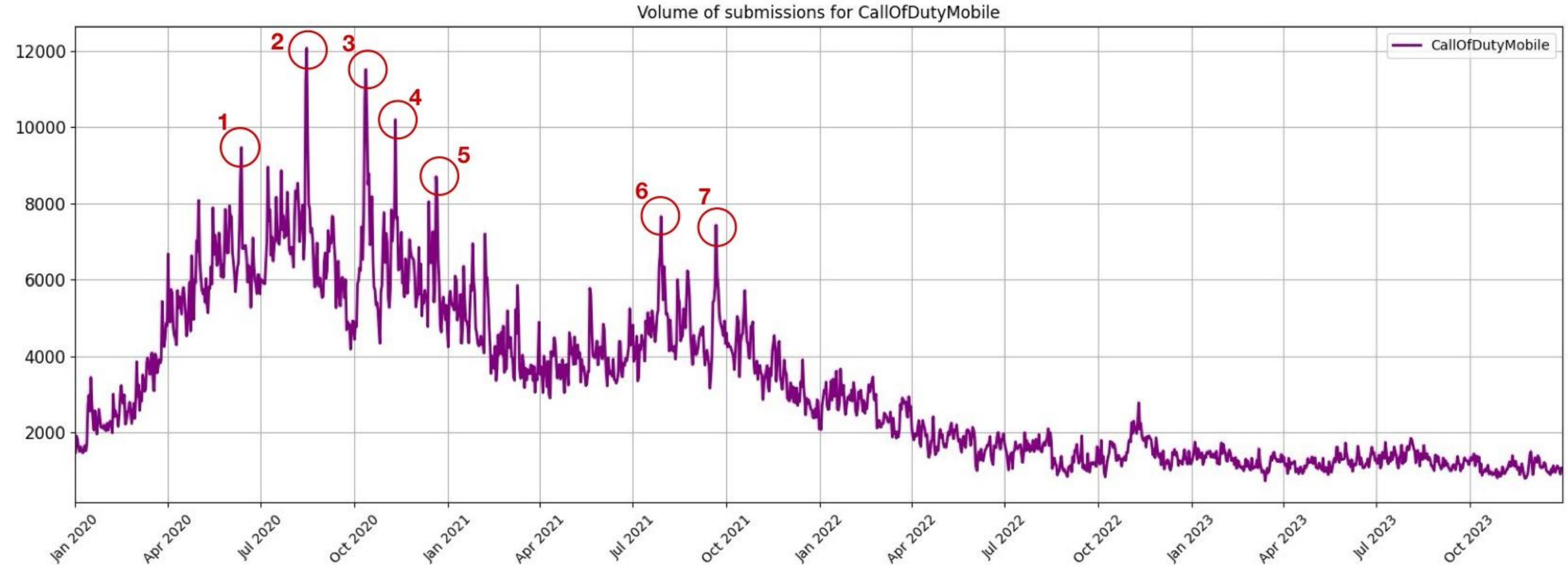
	Date	Event
(1)	25/04/2022	Overwatch 2 PvP Beta release (Blizzard Entertainment, 2022a).
(2)	15/09/2022	Overwatch 2 announcement (Blizzard Entertainment, 2022c).
(3)	04/10/2022	Overwatch 2 release (Blizzard Entertainment, 2022b).
(4)	25/10/2022	Overwatch 2 patch (Cotten, 2022).
(5)	06/12/2022	Overwatch 2 patch (Blizzard Entertainment, 2022e).
(6)	07/02/2023	Overwatch 2 Season 3 (Cotten, 2023).
(7)	17/05/2023	Overwatch 2 patch (Blizzard Entertainment, 2023b).
(8)	11/08/2023	Overwatch 2: Invasion release (Freeman, 2023).



	Date	Event
(1)	22/09/2022	Call Of Duty MW2 beta release (Heaney, 2022a).
(2)	28/10/2022	Call Of Duty MW2 game release (Call Of Duty, 2022).
(3)	16/11/2022	Call Of Duty MW2: Season 1 release (Heaney, 2022b).
(4)	14/12/2022	Call Of Duty MW2 Season 1 Reloaded release (Heaney, 2022c).
(5)	15/02/2023	Call Of Duty MW2 Season 2 release (Call Of Duty Wiki, 2023).
(6)	12/04/2023	Call Of Duty MW2 Season 2 release (Yu, 2023).

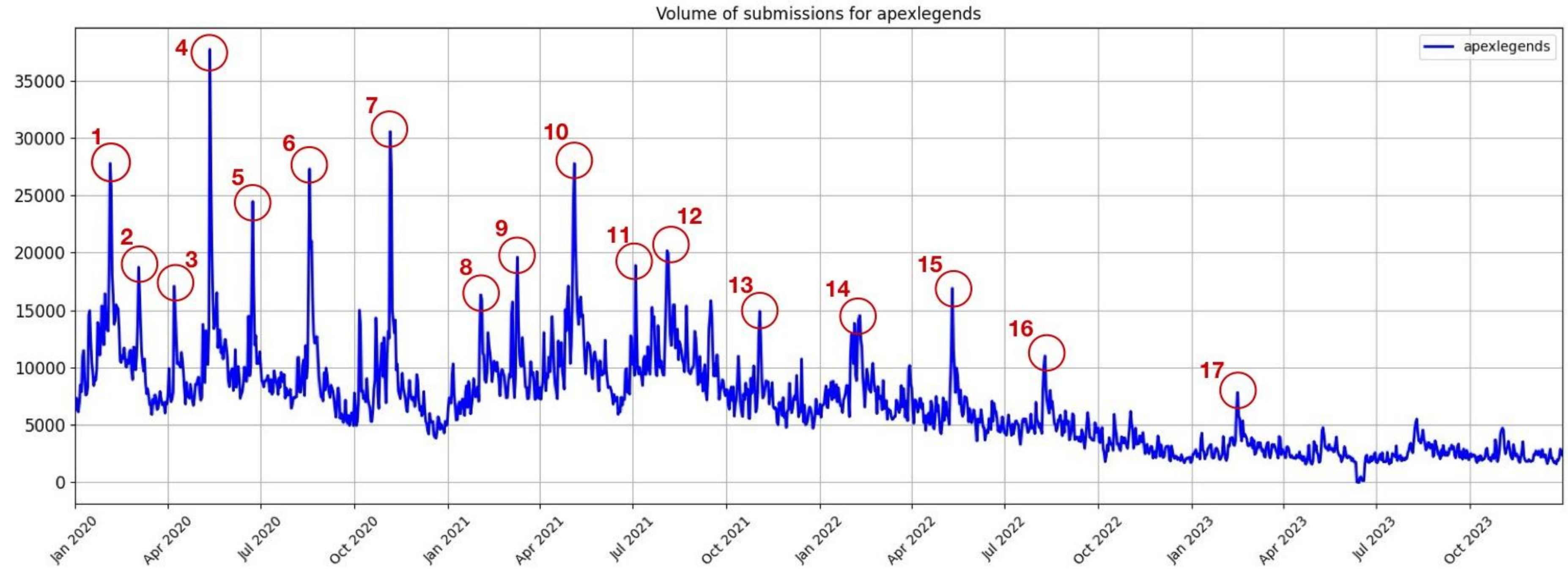


	Date	Event
(1)	10/03/2020	Call Of Duty: Warzone release (Hume, 2021).
(2)	08/04/2020	Call Of Duty: Warzone Season 3 release (Call Of Duty Wiki, 2022).
(3)	11/06/2020	Call Of Duty: Warzone Season 4 release (Glover, 2020).
(4)	22/04/2021	Call Of Duty: Warzone Verdansk 84 release (Hodgson, 2021).
(5)	08/12/2021	Call Of Duty: Warzone Caldera map release (Call Of Duty Wiki, 2021).
(6)	17/11/2022	Call Of Duty: Warzone Mobile announced (Noel, 2022).



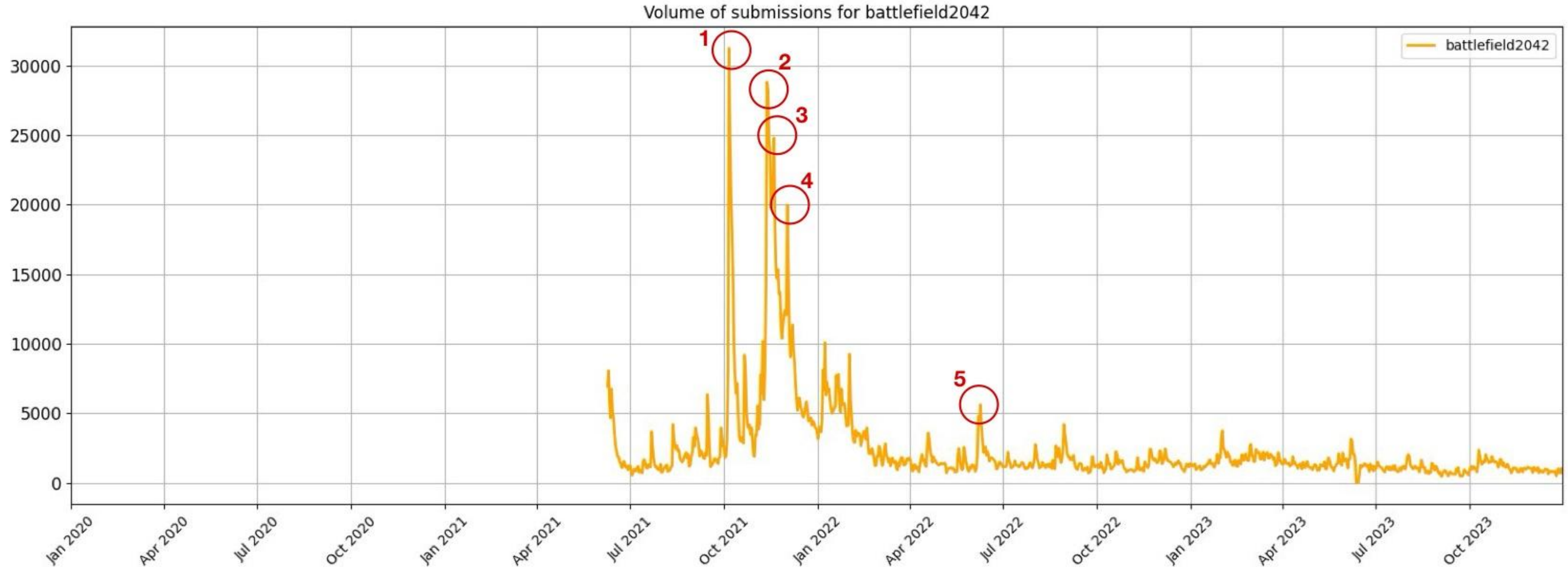
	Date	Event
(1)	11/06/2020	Call Of Duty: Mobile Season 7 release (Noel, 2020a).
(2)	14/08/2020	Call Of Duty: Mobile Season 9 release (Noel, 2020b).
(3)	14/10/2020	Call Of Duty: Mobile anniversary (Noel, 2020c).
(4)	11/11/2020	Call Of Duty: Mobile Season 12 release (Noel, 2020d).
(5)	21/12/2020	Call Of Duty: Mobile Season 13 release (Noel, 2020e).
(6)	29/07/2021	Call Of Duty: Mobile Season 6 release (Noel, 2021a).
(7)	22/09/2021	Call Of Duty: Mobile Season 8 release (Noel, 2021b).

Appendix 6: Electronic arts volume of submissions



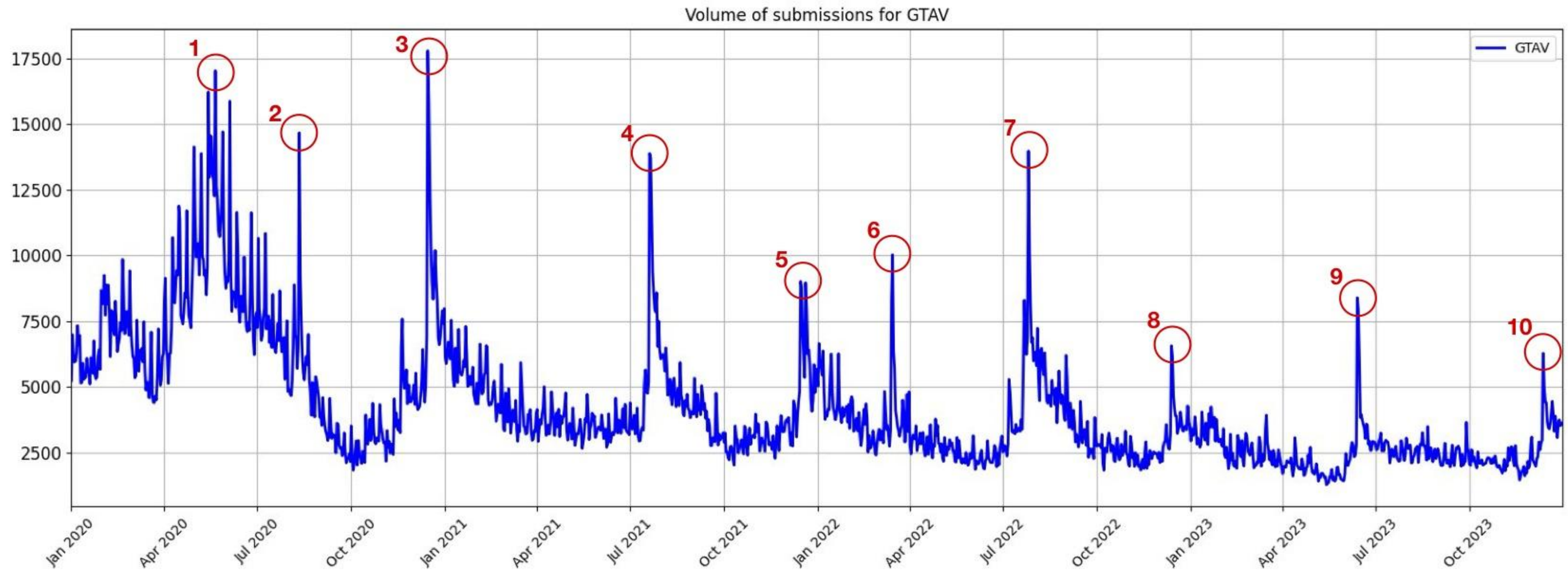
	Date	Event
(1)	04/02/2020	Assimilation - Season 4 (Bull, 2023).
(2)	04/03/2020	Game update (Goslin. Austen, 2020).
(3)	07/04/2020	Game update (Apex Legends Wiki, 2020a).
(4)	12/05/2020	Fortune's Favor - Season 5 (Bull, 2023).
(5)	23/06/2020	Game update (Apex Legends Wiki, 2020b).
(6)	18/08/2020	Boosted - Season 6 (Bull, 2023).
(7)	04/11/2020	Ascension - Season 7 (Bull, 2023).
(8)	02/02/2021	Mayhem - Season 8 (Bull, 2023).

(9)	09/03/2021	Apex Legends launches on Nintendo Switch (Apex Legends, 2021).
(10)	04/05/2021	Legacy - Season 9 (Bull, 2023).
(11)	04/07/2021	Hacking attack (Lawler, 2021).
(12)	03/08/2021	Emergence - Season 10 (Bull, 2023).
(13)	02/11/2021	Escape - Season 11 (Bull, 2023).
(14)	08/02/2022	Defiance - Season 12 (Bull, 2023).
(15)	10/05/2022	Saviors - Season 13 (Bull, 2023).
(16)	09/08/2022	Hunted - Season 14 (Bull, 2023).
(17)	14/02/2023	Game update (Apex Legends Wiki, 2023).



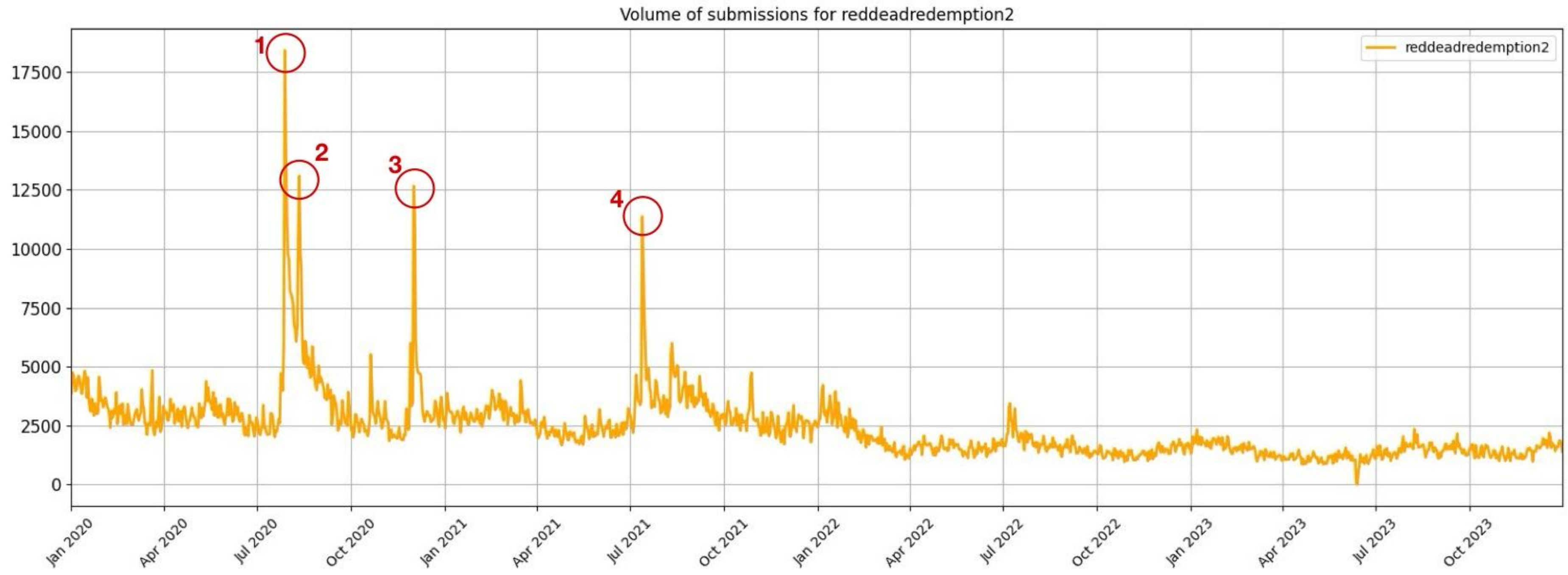
	Date	**Event**
(1)	06/10/2021	Battlefield 2042 Open Beta (Nelson, 2021).
(2)	12/11/2021	Battlefield 2042 Pre-Release (Battlefield Wiki, 2021).
(3)	19/11/2021	Battlefield 2042 General Release (Battlefield Wiki, 2021).
(4)	02/12/2021	Battlefield 2042 Update #3 (Battlefield, 2021).
(5)	07/06/2022	Battlefield 2042: Zero Hour (Battlefield Wiki, 2022).

Appendix 7: Take-Two Interactive volume of submissions

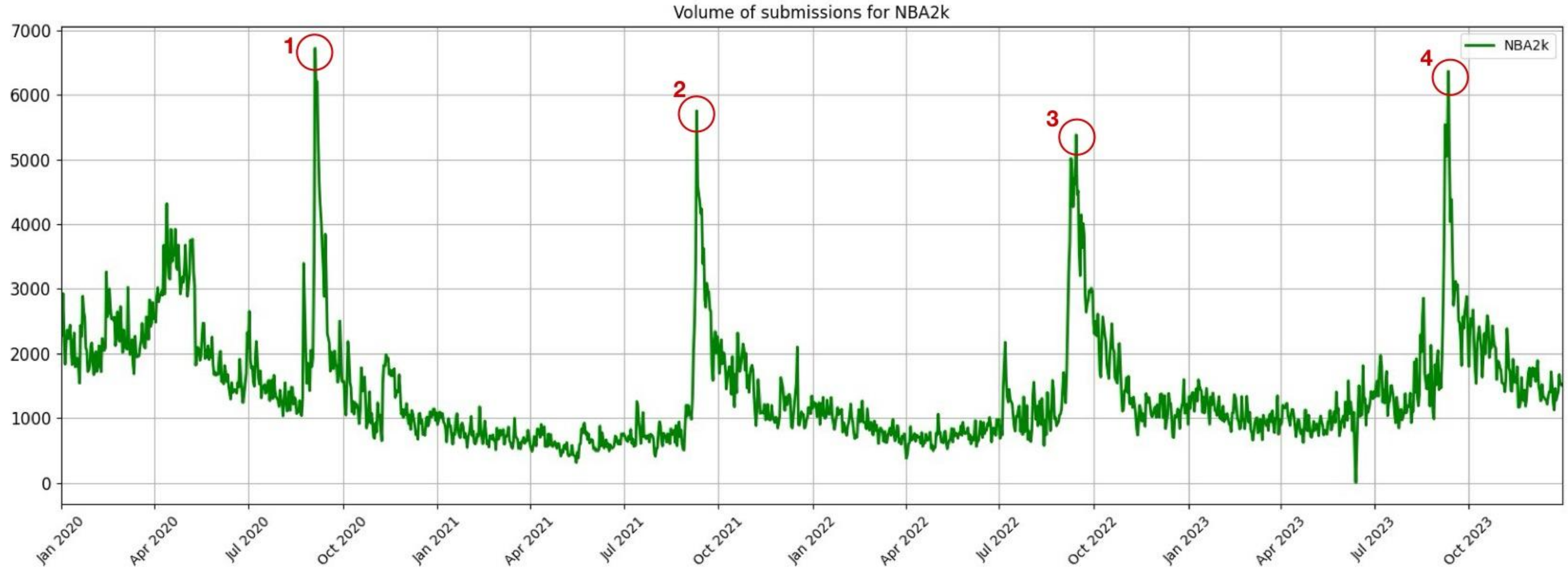


	Date	Event
(1)	14/05/2020	GTAV: Premium Edition Available Free on the Epic Games Store (Rockstar Games, 2020a).
(2)	11/08/2020	GTA Online: Los Santos Summer Special (Rockstar Games, 2020b).
(3)	15/12/2020	GTA Online: The Cayo Perico Heist (Rockstar Games, 2020d).
(4)	20/07/2021	GTA Online: Los Santos Tuners (Rockstar Games, 2021b).
(5)	15/12/2021	GTA Online: The Contract (Rockstar Games, 2021c).
(6)	15/03/2022	Grand Theft Auto V and GTA Online Out Now on PlayStation 5 and Xbox Series X|S (Rockstar Games, 2022a).
(7)	26/07/2022	GTA Online: The Criminal Enterprises (Rockstar Games, 2022b).
(8)	12/12/2022	GTA Online: Los Santos Drug Wars (Rockstar Games, 2022c).

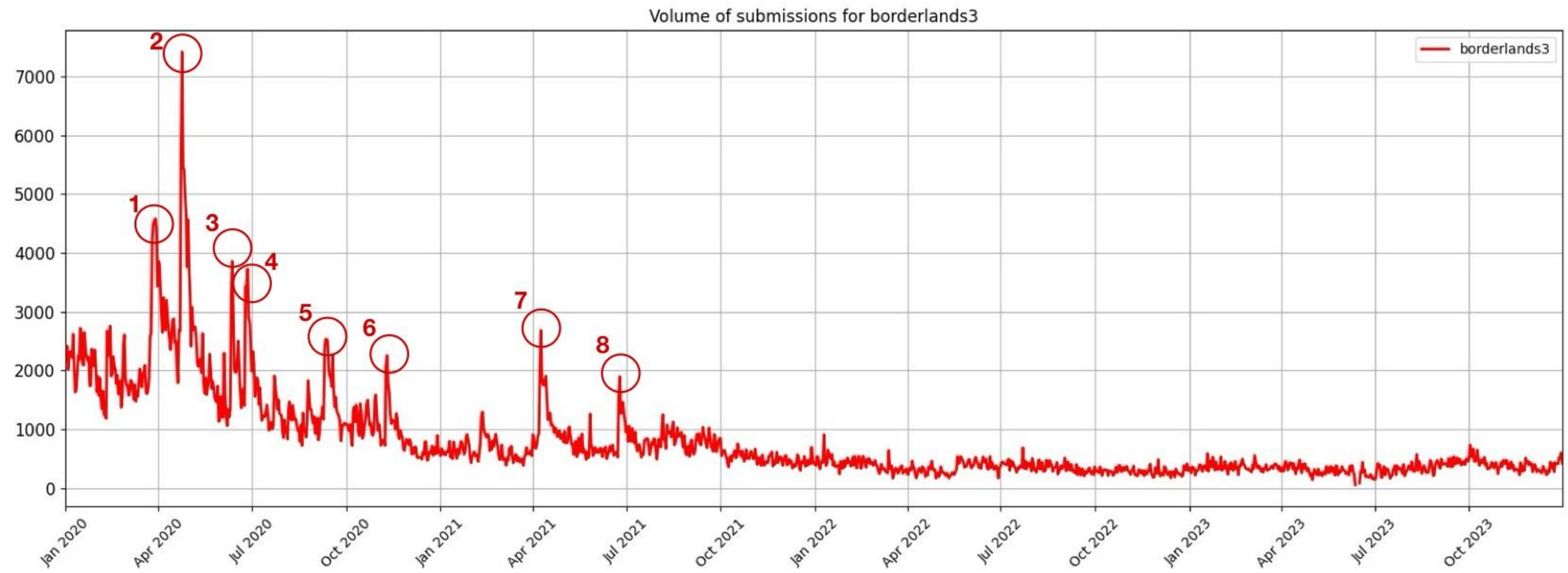
(9)	13/06/2023	GTA Online: San Andreas (Rockstar Games, 2023a).
(10)	12/12/2023	GTA Online: The Chop Shop (Rockstar Games, 2023b).



	Date	Event
(1), (2)	28/07/2020	Red Dead Online's long-awaited update brings a new role: the naturalist (Marshall, 2020).
(3)	01/12/2020	Red Dead Redemption 2 Title Update 1.24 (Rockstar Games, 2020c).
(4)	13/07/2021	Red Dead Redemption 2 Title Update 1.25 (Rockstar Games, 2021a).

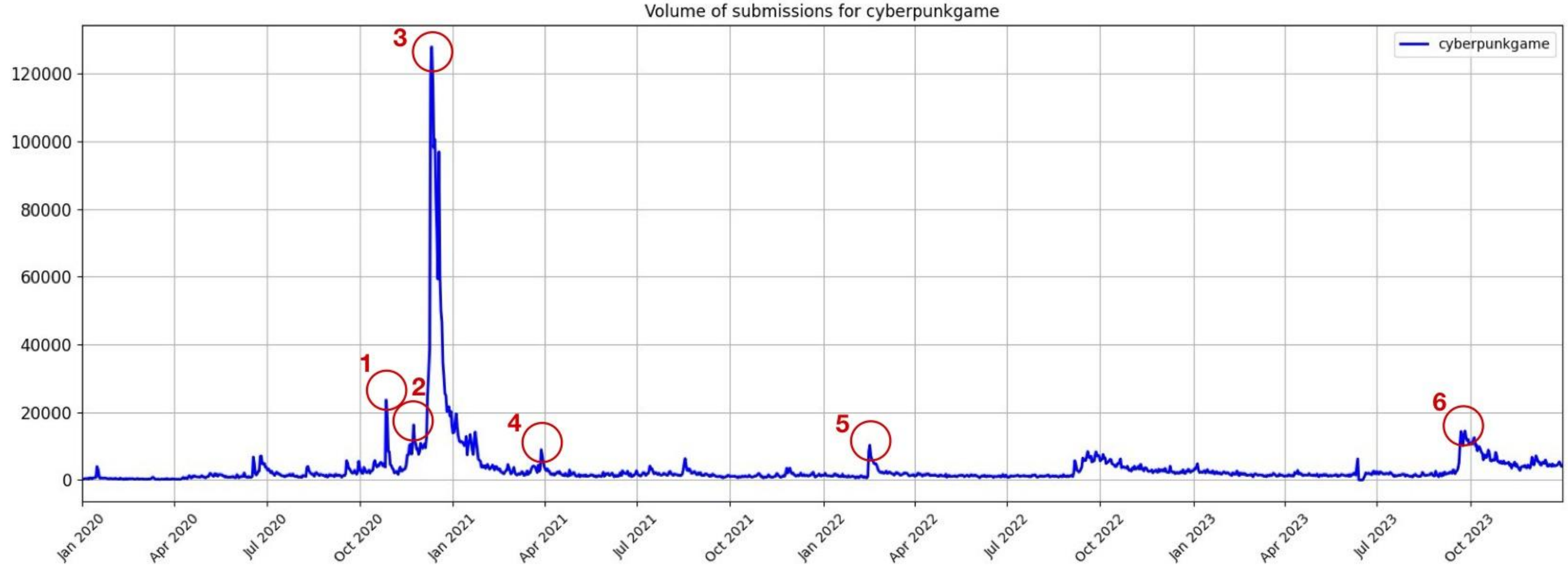


	Date	Event
(1)	04/09/2020	NBA® 2K21 Release (2K, 2020).
(2)	10/09/2021	NBA® 2K22 Release (2K, 2021).
(3)	09/09/2022	NBA® 2K23 Release (Take Two Interactive, 2022).
(4)	08/09/2023	NBA® 2K24 Release (Take Two Interactive, 2023).



	Date	Event
(1)	26/03/2020	Borderlands 3 Update and Hotfixes (2K, 2024).
(2)	23/04/2020	Borderlands 3 Update and Hotfixes (2K, 2024).
(3)	11/06/2020	Borderlands 3 Update and Hotfixes (2K, 2024).
(4)	25/06/2020	Borderlands 3 Update and Hotfixes (2K, 2024).
(5)	10/09/2020	Borderlands 3 Update and Hotfixes (2K, 2024).
(6)	09/11/2020	Borderlands 3 Update and Hotfixes (2K, 2024).
(7)	08/04/2021	Borderlands 3 Update and Hotfixes (2K, 2024).
(8)	24/06/2021	Borderlands 3 Update and Hotfixes (2K, 2024).

Appendix 8: CD Projekt Red volume of submissions



	Date	Event
(1)	27/10/2020	Cyberpunk 2077 delay announcement (Kim, 2020).
(2)	22/11/2020	Cyberpunk 2077 physical copies leak (Bankhurst, 2020a).
(3)	10/12/2020	Cyberpunk 2077 game release (Bankhurst, 2020b).
(4)	29/03/2021	Cyberpunk 2077 patch 1.2 released (CD Projekt Red, 2021).
(5)	15/02/2022	Cyberpunk 2077 patch 1.5 released (CD Projekt Red, 2022).
(6)	26/09/2023	Cyberpunk 2077: Phantom Liberty release (Sirani, 2023).

Appendix 9: ADF test results

ADF test results for financial data

open	close	return	
0.1501	0.5265	3.4515e-18**	ATVI
0.0160*	0.0212*	8.6258e-18**	EA
0.3246	0.2995	< 0.01**	TTWO
0.7238	0.6474	< 0.01**	CDR
	p-value		

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

ADF test result for sentiment and the first difference in sentiment

negative	positive	compound	negative_diff	positive_diff	compound_diff	
0.0019**	0.1560	0.0002**	6.33e-19**	1.31e-23**	1.80e-28**	ATVI
3.22e-21**	1.06e-18**	3.18e-13**	5.96e-26**	1.17e-24**	4.69e-24**	WoW
0.0006**	0.3706	0.0174*	1.94e-26**	8.12e-28**	1.06e-22**	OW
0.3065	0.0004**	0.0369*	1.00e-14**	4.66e-26**	3.91e-25**	CoD: MW2
8.01e-05**	4.07e-06**	1.98e-13**	2.73e-25**	4.82e-24**	6.53e-27**	CoD: Warzone
0.0131*	0.0276*	0.0205*	8.95e-28**	1.15e-25**	8.71e-26**	CoD: Mobile
0.0318*	0.3825	0.0148*	4.70e-26**	5.57e-25**	4.59e-22**	EA
2.10e-10**	0.3733	3.08e-06**	1.27e-26**	1.91e-23**	5.20e-23**	Apex Legends
0.0014**	3.05e-08**	0.0379*	9.38e-19**	6.22e-19**	1.01e-18**	BF 2042
2.93e-08**	3.34e-07**	5.80e-05**	8.30e-29**	5.63e-24**	1.49e-27**	Madden
2.39e-09**	9.27e-06**	4.39e-16**	2.61e-23**	3.32e-26**	8.02e-25**	SW: GOH
0.0002**	0.2093	6.84e-07**	5.97e-28**	8.40e-23**	1.48e-24**	TTWO
0.0475*	0.2819	0.0077**	9.33e-29**	3.37e-23**	1.78e-28**	GTAV
1.78e-08**	4.69e-09**	0.0002**	6.63e-24**	5.32e-22**	1.91e-27**	RDR2
0.1127	0.0466*	0.1230	3.12e-26**	9.09e-24**	8.02e-26**	NBA2K
0.0005**	2.57e-08**	0.0100*	2.48e-26**	5.07e-24**	1.06e-24**	BL 3
3.54e-07**	0.0041**	1.06e-05**	2.09e-21**	4.07e-24**	1.49e-26**	CDR
3.54e-07**	0.0041**	1.06e-05**	2.09e-21**	4.07e-24**	1.49e-26**	CB 2077
p-value						

Source: Own work.

* Significant at p < 0.05; ** Significant at p < 0.01.

ADF test results for emotion data

anger	anticipation	disgust	fear	joy	love	optimism	pessimism	sadness	surprise	trust	
0.0236*	0.0001**	0.0255*	0.0247*	0.0001**	0.0003**	0.0068**	0.0004**	0.0008**	6.17e-05**	0.0002**	ATVI
0.0036**	0.0003**	0.0043**	3.42e-16**	5.52e-07**	8.29e-12**	5.09e-08**	0.0130*	0.0061**	8.74e-08**	7.05e-06**	WoW
0.1965	6.87e-12**	0.1841	3.04e-07**	0.0211*	0.0886	0.0697	0.0003**	0.0019**	0.0518	0.0235*	OW
1.70e-06**	0.2232	2.57e-06**	1.02e-16**	3.33e-07**	2.98e-10**	0.0060**	3.00e-12**	0.0777	7.81e-06**	0.0378*	CoD: MW2
3.24e-07**	2.54e-12**	5.61e-07**	0.0437*	0.0011**	0.0107*	4.76e-09**	0.0235*	1.18e-08**	6.28e-06**	6.48e-11**	CoD:Warzone
0.0040**	0.0005**	0.0061**	7.19e-13**	0.0013**	6.76e-09**	0.1188	5.35e-13**	2.29e-17**	2.41e-14**	0.0221*	CoD: Mobile
0.0004**	7.45e-05**	0.0007**	0.4127	2.67e-05**	0.1626	1.51e-07**	0.0011**	0.0003**	7.05e-06**	2.04e-13**	EA
0.0125*	9.04e-11**	0.0159*	0.1011	0.0195*	0.0896	0.0003**	0.0266*	0.0004**	0.0032**	0.0013**	ApexLegends
0.0088**	0.0404*	0.0105*	0.0170*	0.0188*	0.0317*	0.0385*	0.1310	0.0607	0.0009**	0.0433*	BF 2042
0.0020**	0.0634	0.0035**	2.00e-06**	0.0004**	3.72e-11**	0.0013**	0.0053**	0.0061**	4.77e-15**	0.0002**	Madden
9.49e-14**	1.68e-05**	5.28e-09**	3.42e-22**	2.06e-17**	1.05e-09**	2.02e-15**	2.13e-12**	9.30e-12**	2.90e-06**	2.44e-16**	SW: GOH
1.31e-13**	8.62e-07**	1.11e-13**	3.28e-06**	8.48e-06**	9.97e-08**	0.0127*	5.45e-10**	1.76e-11**	0.0026**	3.04e-05**	TTWO
2.59e-07**	0.0001**	2.26e-06**	3.56e-06**	0.0009**	6.17e-05**	0.0015**	4.19e-05**	0.0006**	3.22e-05**	4.88e-05**	GTAV
5.84e-07**	1.22e-08**	4.47e-07**	5.12e-07**	5.78e-10**	8.14e-06**	9.59e-08**	1.62e-12**	4.56e-15**	3.14e-13**	1.49e-16**	RDR2
0.1953	0.1089	0.1909	1.75e-07**	1.03e-05**	0.0013**	0.0043**	0.0294**	0.0025**	1.83e-06**	0.0605	NBA2K
9.24e-06**	8.20e-07**	6.67e-06**	7.01e-16**	0.0029**	0.0003**	5.72e-27**	0.0084**	4.68e-07**	1.01e-20**	1.87e-23**	BL 3
0.0298*	0.0686	0.0405*	0.0226*	1.29e-06**	0.0001**	0.0023**	0.0002**	0.0113*	8.81e-05**	0.0283*	CDR
0.0298*	0.0686	0.0405*	0.0226*	1.29e-06**	0.0001**	0.0023**	0.0002**	0.0113*	8.81e-05**	0.0283*	CB 2077
p-value											

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

ADF test results for the first difference of emotion data

anger	anticipation	disgust	fear	joy	love	optimism	pessimism	sadness	surprise	trust	
3.00e-21**	7.90e-19**	1.59e-21**	4.02e-20**	1.60e-22**	5.21e-28**	1.35e-25**	3.78e-22**	4.04e-23**	1.77e-24**	2.07e-23**	ATVI
2.28e-27**	6.48e-25**	3.83e-28**	6.68e-27**	6.35e-26**	5.93e-26**	2.92e-28**	1.63e-26**	3.82e-24**	3.86e-30**	5.37e-28**	WoW
2.74e-26**	2.48e-25**	2.72e-26**	3.06e-26**	2.16e-26**	1.44e-26**	1.23e-23**	1.57e-28**	2.99e-26**	1.31e-29**	1.48e-25**	OW
8.86e-16**	1.32e-10**	1.06e-15**	1.41e-19**	7.66e-24**	2.08e-23**	1.28e-19**	1.28e-17**	3.52e-16**	2.04e-13**	1.04e-23**	CoD: MW2
2.56e-30**	2.37e-28**	1.68e-26**	3.11e-22**	8.75e-26**	2.21e-25**	1.40e-27**	1.76e-25**	1.45e-24**	6.86e-24**	2.24e-27**	CoD:Warzone
2.66e-27**	5.94e-27**	5.87e-28**	5.85e-25**	6.74e-22**	1.12e-23**	1.58e-23**	6.22e-23**	6.77e-24**	1.81e-24**	8.31e-23**	CoD: Mobile
1.10e-28**	2.54e-24**	1.82e-29**	1.74e-24**	6.47e-25**	2.12e-26**	8.91e-26**	5.14e-25**	7.33e-22**	9.53e-25**	3.30e-22**	EA
6.11e-30**	2.33e-23**	5.58e-30**	1.82e-26**	8.65e-25**	7.20e-24**	5.30e-24**	6.04e-22**	5.72e-23**	2.38e-26**	6.44e-23**	ApexLegends
7.32e-25**	4.44e-23**	4.48e-25**	4.29e-21**	8.40e-21**	1.25e-21**	1.88e-19**	4.19e-25**	7.89e-24**	2.55e-24**	5.28e-20**	BF 2042
1.63e-29**	1.21e-28**	8.02e-29**	3.56e-26**	9.51e-30**	2.99e-22**	6.48e-29**	2.37e-29**	9.18e-29**	4.81e-27**	4.03e-26**	Madden
1.11e-22**	6.90e-27**	8.91e-23**	1.94e-26**	4.72e-24**	2.33e-25**	4.45e-25**	1.25e-22**	1.23e-22**	2.51e-26**	1.54e-25**	SW: GOH
<0.01**	4.39e-30**	1.18e-29**	<0.01**	4.63e-29**	2.91e-30**	5.74e-29**	1.12e-29**	2.05e-30**	2.09e-30**	4.20e-29**	TTWO
1.39e-21**	3.52e-17**	8.54e-22**	5.60e-25**	1.40e-19**	1.05e-19**	4.29e-17**	1.43e-23**	8.79e-15**	1.34e-17**	1.44e-18**	GTAV
6.85e-25**	3.36e-24**	1.03e-24**	1.42e-24**	6.48e-24**	3.27e-23**	1.63e-23**	2.45e-23**	4.29e-23**	4.82e-23**	1.87e-22**	RDR2
3.52e-24**	8.87e-22**	5.86e-24**	7.63e-30**	1.95e-24**	2.69e-24**	1.94e-26**	2.25e-24**	3.66e-23**	4.25e-24**	1.09e-24**	NBA2K
3.24e-23**	3.00e-24**	1.29e-23**	3.30e-24**	7.62e-25**	4.48e-24**	8.48e-25**	2.91e-24**	1.69e-23**	5.00e-24**	1.18e-24**	BL 3
1.84e-25**	9.30e-26**	2.64e-25**	6.19e-28**	2.24e-23**	2.38e-25**	5.69e-27**	1.47e-23**	3.77e-23**	7.76e-23**	6.86e-25**	CDR
1.84e-25**	9.30e-26**	2.64e-25**	6.19e-28**	2.24e-23**	2.38e-25**	5.69e-27**	1.47e-23**	3.77e-23**	7.76e-23**	6.86e-25**	CB 2077
					p-value						

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

Appendix 10: Correlation and Granger causality results for ATVI

Correlation results

	ATVI	WoW	OW	CoD: MW2	CoD: Warzone	CoD: Mobile	(1)	(2)	(3)	(4)	(5)	
negative	-0.049	-0.036	-0.041	-0.013	0.028	0.012	**-0.236***	0.081	-0.042	**0.121***	**-0.271***	Pearson r
positive	0.066	0.051	0.074	0.022	-0.012	-0.003	-0.03	0.004	**-0.121***	0.091	**0.239***	
compound	0.051	0.037	0.049	-0.045	-0.023	-0.017	**0.104***	-0.054	0.008	-0.088	**0.287***	
anger	-0.023	-0.027	-0.044	0.031	0.068	0.024	0.009	-0.051	-0.061	0.089	**-0.241***	
anticipation	-0.009	0.012	-0.021	-0.085	-0.069	0.001	0.01	**0.229***	**0.178***	**-0.223***	0.091	
disgust	-0.027	-0.03	-0.047	0.038	0.063	0.026	0.003	-0.065	-0.029	0.074	-0.24	
fear	-0.03	-0.023	-0.047	0.028	-0.01	-0.007	-0.003	-0.071	-0.033	0.042	-0.015	
joy	0.039	0.035	0.047	-0.053	-0.036	-0.019	0.028	0.052	0.018	-0.049	**0.165***	
love	0.04	0.025	0.046	-0.022	-0.001	-0.009	0.043	-0.098	-0.04	0.001	**0.111***	
optimism	0.044	0.05	0.044	**-0.116***	-0.043	-0.01	**0.169***	0.051	0.024	**-0.135***	0.084	
pessimism	-0.038	-0.036	-0.056	-0.007	0.001	0.047	0.085	-0.056	0.004	-0.056	**-0.202***	
sadness	-0.054	-0.047	-0.057	0.041	-0.006	0.005	0.023	**-0.118***	**-0.161***	0.05	**-0.147***	
surprise	-0.021	-0.037	0.009	0.018	-0.046	-0.008	**-0.225***	0.086	**-0.004**	**-0.165***	**0.113***	
trust	0.023	0.038	0.027	-0.1	-0.05	-0.015	0.049	-0.023	**0.176***	**-0.138***	0.04	
	return											

Source: Own work.

* Weak correlation at 0.1 < r < 0.3; ** Medium correlation at 0.3 < r < 0.7.

(1) – (5) correlations between return and ATVI for local events in Table 13.

Correlation results for local events

	(1)	(2)	(3)		(4)				(5)	
	WoW	CoD: Warzone	WoW	OW	WoW	OW	CoD: MW2	CoD: Wazone	OW	
negative	-0.092	**-0.254***	**0.152***	-0.016	**0.103***	0.08	-0.01	0.035	**-0.287***	Pearson r
positive	-0.072	**0.125***	**-0.148***	**0.125***	-0.006	**0.121***	0.03	-0.058	**0.227***	
compound	-0.037	**0.234***	-0.051	0.003	-0.044	-0.003	-0.017	**-0.187***	**0.265***	
anger	0.012	**-0.109***	**0.144***	**-0.129***	0.026	0.085	**0.128***	**0.163***	-0.18	
anticipation	-0.027	0.095	0.058	-0.012	-0.071	**-0.203***	-0.17	-0.08	0.052	
disgust	0.006	**-0.129***	**0.168***	-0.102	0.028	0.069	**0.123***	**0.161***	**-0.185***	
fear	0.087	0.006	**0.109***	-0.061	0.069	-0.022	-0.029	-0.026	**-0.122***	
joy	0.023	**0.143***	-0.074	0.059	0.015	0.018	-0.084	**-0.219***	**0.212***	
love	-0.01	**0.125***	**-0.1***	**0.104***	0.019	0.064	-0.076	**-0.143***	**0.204***	
optimism	**0.121***	0.076	-0.088	-0.052	-0.028	-0.035	-0.099	**-0.206***	**0.232***	
pessimism	0.005	**-0.107***	**0.145***	-0.068	0.026	-0.071	0.088	0.036	**-0.168***	
sadness	0.011	-0.064	**0.112***	**-0.199***	0.072	-0.034	**0.198***	0.082	**-0.163***	
surprise	**-0.143***	0.028	-0.055	**0.1***	-0.02	**-0.14***	**-0.136***	**-0.109***	0.009	
trust	0.031	0.085	**0.145***	-0.027	-0.035	-0.058	-0.16	**-0.189***	**0.181***	
	Return									

Source: Own work.

* Weak correlation at 0.1 < r < 0.3; ** Medium correlation at 0.3 < r < 0.7.

(1) – (5) local events in Table 13.

Granger causality results

	ATVI	WoW	OW	CoD: MW2	CoD: Warzone	CoD: Mobile	(1)	(2)	(3)	(4)	(5)
negative	0.1015	0.1102	0.2289	0.8466	0.1576	0.7337	0.5369	**0.0172***	0.7930	0.4454	0.1807
positive	0.4427	0.5545	0.0823	0.1499	0.3789	0.6769	0.0948	0.4339	0.2221	0.9768	0.4883
compound	0.1271	0.3305	0.2182	0.3590	**0.0068****	0.3803	0.1395	0.1479	0.8445	0.7915	0.2861
anger	0.4133	0.6751	0.4355	0.4691	**0.0257***	0.4137	0.1554	**0.0038****	0.4988	0.3779	0.7150
anticipation	0.7926	0.6906	0.4892	0.3529	**0.0093****	0.1978	0.2467	0.0501	0.5357	0.3756	0.5875
disgust	0.4825	0.6796	0.5497	0.4073	**0.0250***	0.5280	0.1380	**0.0019****	0.4930	0.4122	0.7121
fear	0.1833	0.0343	0.1399	0.4412	0.5925	0.3788	0.2292	0.1366	0.3473	0.5648	0.1695
joy	0.5910	0.5544	0.2180	0.9254	**0.0033****	0.2002	**0.0491***	**0.0285***	0.2101	0.2214	0.6804
love	0.1649	0.5447	0.5343	0.5973	0.3846	0.3487	0.0693	0.2631	0.0976	0.1914	0.8674
optimism	0.6166	0.3968	0.4161	0.3107	**0.0016****	0.4323	0.2497	**0.0363***	0.6557	0.0706	0.2665
pessimism	0.3439	0.8805	0.8972	0.4800	0.2921	0.1973	0.0842	**0.0091****	0.4772	0.9619	0.6519
sadness	0.2959	0.4958	0.8421	0.1783	0.5783	0.3320	0.2575	**0.0052****	0.5179	0.5052	0.4783
surprise	0.1273	**0.0252***	0.1596	0.5212	**0.0274***	0.4233	0.3137	0.8248	0.9268	0.1518	0.5933
trust	0.3707	0.3847	0.5720	0.3365	**0.0092****	0.4477	0.6991	0.1174	0.2666	0.1181	0.5715

return

p-value

Source: Own work.

* Significant at p < 0.05; ** Significant at p < 0.01.

(1) – (5) causality between return and ATVI for local events in Table 13.

Granger causality results for local events

	(1)	(2)	(3)		(4)				(5)	
	WoW	CoD: Warzone	WoW	OW	WoW	OW	CoD: MW2	CoD: Wazone	OW	
negative	0.1219	0.7498	0.1390	0.4133	0.7897	0.1714	0.7069	0.2642	0.2755	
positive	0.7056	0.1642	0.0807	0.3353	0.6735	0.6136	0.4803	0.4684	0.4216	
compound	0.1613	0.3954	**0.0060****	0.5296	0.8367	0.6012	0.7728	0.3592	0.2631	
anger	0.4011	0.1459	0.1090	0.2139	0.7305	0.3240	0.5874	0.7592	0.3172	
anticipation	0.1967	0.1320	0.2753	0.2537	0.2833	0.3087	0.8090	0.4535	0.4511	
disgust	0.5222	0.1307	0.1153	0.2488	0.6442	0.3435	0.6750	0.7031	0.4211	
fear	0.0352	0.5836	0.2504	0.0598	0.3296	0.8256	0.5960	**0.0311***	0.3415	p-value
joy	0.5382	0.2005	**0.0010****	0.0908	0.3552	0.1529	0.1427	0.8663	0.6895	
love	0.7079	0.2164	0.0583	**0.0342***	0.7133	0.2032	0.7043	0.7107	0.6933	
optimism	0.5561	0.5918	**<0.01****	0.1534	0.5485	0.1194	0.5322	0.5637	0.3484	
pessimism	0.3143	0.2414	0.5886	0.7235	0.7296	0.9590	0.2341	0.9481	0.8279	
sadness	0.0550	0.5826	0.2852	0.7527	0.6872	0.8701	0.8810	0.3386	0.5551	
surprise	0.0846	0.3239	0.3396	0.4372	0.6636	0.0505	0.2134	0.2987	0.8508	
trust	0.4444	0.7220	**0.0285***	**0.0271***	0.8267	0.2129	0.6050	0.6689	0.1852	
	return									

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

(1) – (5) local events in Table 13.

Appendix 11: Correlation and Granger causality results for EA

Correlation results for Electronic Arts

	EA	Apex Legends	BF 2042	(1)	(2)	(3)	(4)	(5)	(6)	
negative	-0.059	-0.026	-0.057	**-0.161***	**-0.142***	**0.121***	-0.076	**-0.139***	**0.135***	Pearson r
positive	0.014	0.033	0.008	-0.072	-0.019	-0.012	**0.108***	0.28	**-0.15***	
compound	0.018	0.016	0.049	-0.093	0.071	**-0.104***	**0.121***	**0.247***	**-0.116***	
anger	-0.026	-0.018	-0.065	0.017	-0.041	0.097	**-0.177***	**-0.293***	**0.172***	
anticipation	0.019	0.008	0.033	0.019	**0.111***	**-0.263***	**0.304****	**0.2***	-0.037	
disgust	-0.03	-0.02	-0.062	0.014	-0.046	0.098	**-0.173***	**-0.333****	**0.159***	
fear	0.007	0.005	0.019	-0.048	-0.093	0.035	**0.13***	-0.099	0.007	
joy	0.016	0.026	0.011	0.056	0.041	**-0.128***	**0.291***	**0.344****	**-0.154***	
love	0.001	0.014	0.002	-0.088	-0.031	**-0.108***	**0.265***	**0.422****	-0.098	
optimism	0.033	0.031	0.016	0.087	-0.021	**-0.224***	**0.483****	**0.425****	**-0.135***	
pessimism	-0.049	-0.023	-0.054	-0.047	0.016	**0.223***	**-0.47****	**-0.399****	0.093	
sadness	-0.055	-0.037	-0.045	-0.064	-0.031	**0.335****	**-0.614****	**-0.494****	0.097	
surprise	-0.035	-0.018	0.045	**-0.143***	0.008	**-0.169***	**0.214***	**-0.161***	**0.236***	
trust	0.017	0.016	0.03	-0.055	-0.003	**-0.311****	**0.481****	**0.311****	-0.06	
	return									

Source: Own work.

* Weak correlation at 0.1 < r < 0.3; ** Medium correlation at 0.3 < r < 0.7.

(1) – (6) correlations between return and EA for local events in Table 15.

Correlation results for local events

	(1)	(2)	(3)	(4)	(5)	(6)	
	Apex Legends	Apex Legends	Apex Legends	Apex Legends	BF 2042	BF 2042	
negative	**-0.128***	**-0.216***	0.086	-0.02	**-0.39****	**0.218***	Pearson r
positive	-0.085	0.065	-0.005	**0.132***	**0.39****	**-0.189***	
compound	**-0.114***	**0.154***	-0.086	0.095	**0.311****	**-0.165***	
anger	0.048	**-0.128***	0.09	**-0.186***	**-0.376****	**0.172***	
anticipation	-0.014	**0.146***	**-0.242***	**0.267***	**0.347****	**-0.116***	
disgust	0.047	**-0.136***	0.091	**-0.177***	**-0.413****	**0.179***	
fear	-0.044	-0.085	0.018	**0.176***	0.026	0.029	
joy	0.043	**0.105***	**-0.121***	**0.328****	**0.254***	**-0.22***	
love	**-0.125***	0.011	**-0.101***	**0.313****	**0.108***	**-0.178***	
optimism	0.062	0.039	**-0.214***	**0.5****	**0.383****	**-0.243***	
pessimism	-0.012	-0.057	**0.24****	**-0.479****	**-0.436****	**0.114***	
sadness	-0.02	-0.083	**0.338****	**-0.622****	**-0.517****	**0.153***	
surprise	**-0.157***	0.031	**-0.16***	**0.284***	0.019	**0.151***	
trust	**-0.111***	0.036	**-0.312****	**0.487****	**0.222***	**-0.199***	
	return						

Source: Own work.

* Weak correlation at 0.1 < r < 0.3; ** Medium correlation at 0.3 < r < 0.7.

(1) – (6) local events in Table 15.

Granger causality results

	EA	Apex Legends	BF 2042	(1)	(2)	(3)	(4)	(5)	(6)	
negative	0.5577	0.5658	0.4607	0.2704	0.6675	0.4548	0.6679	0.4674	0.5317	
positive	0.5330	0.7249	0.8946	**0.0006****	0.5257	0.2543	**<0.01****	0.5468	0.7999	
compound	0.8120	0.7912	0.9391	**0.0456***	0.8502	0.5836	**0.0410***	0.8220	0.5676	
anger	0.0867	0.0976	0.7291	**0.0260***	0.5904	**0.0180***	**<0.01****	0.5993	0.5303	
anticipation	0.3174	0.0673	0.3594	**0.0232***	0.2048	0.0940	0.2801	0.2037	0.2149	
disgust	0.0900	0.1041	0.7771	**0.0193***	0.6242	**0.0169***	**<0.01****	0.5008	0.5405	
fear	0.0719	0.0920	0.8012	0.5699	0.1389	0.1637	0.5193	0.4016	0.1441	
joy	0.2546	0.4423	0.4490	**0.0022****	0.3892	0.2147	**<0.01****	0.8748	0.1870	p-value
love	0.3665	0.4721	0.6506	0.7161	0.2361	0.3888	**0.0273***	0.4880	0.1370	
optimism	0.3963	0.4354	0.8567	**0.0005****	0.4271	0.2409	**0.0219***	0.6374	0.1096	
pessimism	0.1055	0.2135	0.9128	**0.0075****	0.8558	**0.0077****	0.1837	0.4849	0.3008	
sadness	0.2827	0.3933	0.8986	**0.0034****	0.4933	**0.0048****	0.3683	0.7567	0.1195	
surprise	0.2944	0.1208	0.8170	**0.0049****	0.2053	0.1309	**0.0087****	0.7569	0.8443	
trust	0.9671	0.6291	0.6214	0.1128	0.6160	0.2859	0.3377	0.9035	0.0527	
					return					

Source: Own work.

* Significant at p < 0.05; ** Significant at p < 0.01.

(1) – (6) causality between return and EA for local events in Table 15.

Granger causality results for local events

	(1)	(2)	(3)	(4)	(5)	(6)	
	Apex Legends	Apex Legends	Apex Legends	Apex Legends	BF 2042	BF 2042	
negative	0.1293	0.4875	0.6530	0.7491	0.9676	0.4327	p-value
positive	**0.0010****	0.3625	0.3521	**<0.01****	0.9355	0.7697	
compound	0.0730	0.4737	0.6952	**0.0356***	0.4364	0.5049	
anger	**0.0067****	0.2890	**0.0277***	**0.0002****	0.4839	0.7429	
anticipation	**0.0006****	0.2544	0.1663	0.3437	0.7363	0.5504	
disgust	**0.0037****	0.2772	0.0273	**0.0002****	0.4156	0.6889	
fear	0.2039	0.1489	0.1645	0.4710	**0.0065****	0.1942	
joy	**0.0001****	0.3041	0.2603	**<0.01****	0.5798	0.5312	
love	0.8422	0.0447	0.4722	**0.0139***	0.2796	0.4275	
optimism	**<0.01****	0.2507	0.3078	**0.0111***	0.1958	0.5730	
pessimism	**0.0017****	0.6485	**0.0193***	0.1833	0.5493	0.3394	
sadness	**<0.01****	0.6145	**0.0141***	0.4505	0.1573	0.3748	
surprise	**0.0003****	0.3666	0.1661	**0.0081****	0.6558	0.6411	
trust	0.1325	0.3633	0.3668	0.2786	**0.0188***	0.4828	
	return						

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

(1) – (6) local events in Table 15.

Appendix 12: Correlation and Granger causality results for TTWO

Correlation results

	TTWO	GTAV	RDR2	NBA2K	BL 3	(1)	(2)	(3)	(4)	(5)	(6)
negative	0.031	0.008	0.03	-0.008	0.009	**-0.251***	**-0.217***	**-0.188***	0.068	0.09	-0.001
positive	-0.0	-0.066	-0.011	0.041	0.002	-0.063	**-0.232***	**0.151***	0.065	**0.157***	-0.046
compound	-0.01	-0.033	-0.017	0.037	0.017	0.035	0.01	**0.164***	**-0.166***	**-0.175***	0.032
anger	0.005	0.044	-0.027	-0.031	-0.005	**-0.321****	0.064	**-0.115***	0.086	-0.019	**-0.133***
anticipation	-0.032	0.01	-0.016	0.015	-0.027	**0.333****	0.049	**0.263***	**-0.297***	**-0.277***	0.036
disgust	0.001	0.045	-0.028	-0.031	-0.01	**-0.315****	0.023	**-0.123***	**0.104***	-0.032	**-0.172***
fear	-0.019	-0.03	-0.017	-0.007	-0.03	**-0.183***	**0.112***	0.032	-0.075	**0.447****	**-0.125***
joy	0.008	-0.04	0.022	0.032	0.0	**0.223***	**0.101***	**0.19***	**-0.144***	0.075	-0.016
love	-0.001	-0.026	0.004	0.008	0.003	-0.038	**0.115***	-0.016	0.056	0.092	**-0.142***
optimism	-0.013	-0.028	-0.008	0.013	0.001	**0.331****	**0.247***	**0.309****	**-0.17***	-0.057	**-0.14***
pessimism	-0.01	-0.023	-0.058	0.009	-0.011	**-0.16***	**-0.284***	-0.051	**0.111***	**-0.425****	**-0.335****
sadness	0.014	-0.02	-0.028	0.006	0.002	**-0.229***	**-0.234***	-0.031	**0.173***	**-0.349****	**-0.292***
surprise	0.001	-0.007	0.018	0.013	-0.03	**0.053**	**-0.255***	0.008	**-0.137***	**0.154***	-0.041
trust	-0.018	-0.02	-0.008	-0.003	-0.02	**0.258***	**0.238***	**0.288***	**-0.2***	-0.034	**-0.2***

return

Pearson r

Source: Own work.

* Weak correlation at 0.1 < r < 0.3; ** Medium correlation at 0.3 < r < 0.7.

(1) – (6) correlations between return and TTWO for local events in Table 17.

Correlation results for local events

	(1)	(2)		(3)		(4)		(5)	(6)	
	BL 3	GTAV	RDR2	GTAV	RDR2	GTAV	RDR2	NBA2K	GTAV	
negative	0.036	-0.002	**-0.348****	**-0.167***	**-0.139***	**-0.287***	**0.166***	-0.089	**-0.247***	Pearson r
positive	**-0.461****	**-0.223***	-0.014	**-0.178***	**0.159***	-0.08	**0.302****	**0.297***	-0.029	
compound	**-0.133***	**-0.34****	**0.257***	**-0.263***	0.058	**0.193***	-0.057	0.093	**0.208***	
anger	**0.128***	**-0.168***	**-0.214***	-0.094	-0.045	**-0.123***	**-0.272***	-0.057	**-0.176***	
anticipation	0.096	**0.231***	**0.331****	**-0.305****	**0.118***	**0.38****	**-0.332****	-0.049	0.053	
disgust	**0.155***	**-0.118***	**-0.235***	-0.06	-0.046	**-0.139***	**-0.24***	-0.056	**-0.181***	
fear	0.038	**0.128***	-0.088	**-0.134***	-0.066	**0.1***	-0.007	**0.172***	**0.463****	
joy	**-0.205****	0.065	**0.319****	-0.051	-0.067	0.059	**0.152***	**0.14***	0.001	
love	**-0.185***	0.011	**0.216***	**0.194***	**-0.216***	**-0.291***	**0.497****	0.044	**0.117***	
optimism	-0.087	0.052	**0.373****	**-0.26***	-0.072	**0.341****	**-0.12***	**0.234***	**-0.311****	
pessimism	**0.373****	**0.237***	**-0.247***	-0.13	-0.042	**-0.181***	**-0.172***	**-0.121***	**-0.164***	
sadness	**0.222***	**0.245***	**-0.143***	-0.15	-0.044	**-0.372****	-0.044	**-0.175***	**-0.38****	
surprise	**-0.366****	**0.169***	**0.236***	-0.014	**0.215***	**0.217***	-0.06	**0.18***	0.012	
trust	**-0.125***	**0.232***	**0.327****	**-0.111***	**-0.108***	**0.201***	-0.068	**0.193***	**-0.302****	
	return									

Source: Own work.

* Weak correlation at $0.1 < r < 0.3$; ** Medium correlation at $0.3 < r < 0.7$.

(1) – (6) local events in Table 17.

Granger causality results

	TTWO	GTAV	RDR2	NBA2K	BL 3	(1)	(2)	(3)	(4)	(5)	(6)	
negative	0.2112	0.0544	0.2281	0.8401	0.2302	**0.0117***	0.1423	0.4435	0.3337	0.0644	0.3643	
positive	0.0664	0.0748	0.2526	0.4529	0.9358	0.0838	**0.0333***	0.7201	0.1951	**<0.01****	0.0568	
compound	0.1384	**0.0084****	0.2892	0.8861	0.6157	**0.0084****	**0.0246***	0.3372	0.4777	**0.0464***	0.1929	
anger	0.0904	**0.0134***	0.8578	0.6187	0.4574	**0.0244***	**0.0430***	0.8668	0.7032	0.5043	0.1490	
anticipation	**0.0346***	0.0841	0.1310	0.2889	0.1480	**0.0028****	0.7397	0.4779	0.1784	**0.0418***	0.3258	
disgust	0.1159	**0.0077****	0.9730	0.7194	0.5041	**0.0126***	0.0651	0.8833	0.6248	0.3506	0.2400	
fear	0.2938	0.1064	0.6895	0.2145	**0.0335***	0.0790	0.6056	0.6569	0.6666	0.0890	0.1820	p-value
joy	0.5249	**0.0011****	0.3995	0.8295	0.3073	**0.0016****	0.1116	0.3565	0.6683	**0.0170***	0.3097	
love	0.2781	**0.0178***	0.0923	0.7861	0.3148	0.2037	**0.0004****	0.1323	0.2795	**0.0155***	0.6035	
optimism	0.1659	0.1329	0.0506	0.7279	0.6375	**0.0007****	0.4163	0.5755	0.4545	**0.0291***	0.3058	
pessimism	0.3870	0.2882	0.5214	0.7765	0.9003	**0.0054****	**0.0114***	0.7188	0.8820	**0.0494***	0.6291	
sadness	0.2714	0.2265	0.1846	0.7429	0.8398	**0.0001****	**0.0009****	0.4242	0.6434	0.1702	0.4645	
surprise	0.3324	0.1368	0.6720	0.5204	0.1493	0.3369	0.4024	0.6394	0.1452	0.2403	0.3192	
trust	0.1485	0.6979	**0.0498***	0.7106	0.7645	**0.0081****	0.4880	0.2363	0.3181	**0.0160***	0.3378	
						return						

Source: Own work.

* Significant at p < 0.05; ** Significant at p < 0.01.

(1) – (6) causality between return and TTWO for local events in Table 17.

Granger causality results for local events

	(1)	(2)		(3)		(4)		(5)	(6)	
	BL 3	GTAV	RDR2	GTAV	RDR2	GTAV	RDR2	NBA2K	GTAV	
negative	0.0297	0.1090	**0.0373***	0.2971	0.3619	**0.0262***	0.1229	0.2884	0.3422	P-value
positive	0.0938	0.4272	**0.0073****	**0.0279***	0.4353	0.3363	0.7626	0.1572	0.2119	
compound	0.1836	0.4015	**0.0087****	0.0782	0.6424	0.1911	0.3307	**0.0303***	0.4638	
anger	0.8897	0.4628	**0.0265***	**0.0030****	0.7113	0.6785	0.0829	**0.0481***	0.1509	
anticipation	0.1126	0.4517	0.1816	0.7675	0.0558	0.0560	0.0631	0.0864	0.0617	
disgust	0.8452	0.6698	**0.0235***	**0.0028****	0.6941	0.5629	0.1403	**0.0115***	0.2008	
fear	0.1747	0.2992	0.4060	0.8851	0.1446	0.2514	0.2273	0.1071	0.2291	
joy	0.2827	0.1283	**0.0165***	**0.0086****	0.1338	0.5338	0.2653	0.0780	0.1516	
love	0.9465	0.1341	**0.0041****	**0.0039****	**0.0309***	0.8655	0.0228	**0.0127***	0.7616	
optimism	0.0694	0.0484	0.1571	0.3419	**0.0318***	0.1271	0.0663	0.3321	0.3647	
pessimism	0.3765	0.7925	**0.0012****	**0.0326***	0.7880	0.3999	0.3322	0.0954	**0.0112***	
sadness	0.6494	0.9824	**0.0002****	**0.0148***	0.4674	0.8323	0.3001	0.1985	0.4956	
surprise	0.0562	0.2265	0.7971	0.6783	0.8833	0.7762	**0.0172***	0.1224	**0.0086****	
trust	0.2061	0.2165	0.1412	0.8611	**0.0138***	0.0672	0.2479	0.4227	0.2108	
	return									

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

(1) – (6) local events in Table 17.

Appendix 13: Correlation and Granger causality results for CDR

Correlation results

	CDR	CP 2077	(1)	(2)	(3)	
negative	0.016	0.016	0.088	**-0.421****	0.006	Pearson r
positive	-0.05	-0.05	-0.012	**0.379****	**-0.164***	
compound	-0.012	-0.012	-0.006	**0.43****	-0.099	
anger	0.029	0.029	0.006	**-0.527****	0.08	
anticipation	0.011	0.011	0.024	**0.303****	0.066	
disgust	0.023	0.023	-0.007	**-0.529****	0.068	
fear	0.002	0.002	-0.027	0.066	**-0.119***	
joy	-0.058	-0.058	-0.019	**0.372****	0.046	
love	-0.027	-0.027	-0.08	**0.356****	**-0.196***	
optimism	-0.055	-0.055	-0.003	**0.221***	**0.127***	
pessimism	-0.046	-0.046	-0.076	**-0.485****	0.023	
sadness	-0.029	-0.029	-0.016	**-0.406****	-0.038	
surprise	0.006	0.006	-0.078	**0.291***	**-0.161***	
trust	-0.046	-0.046	0.004	**0.125***	**-0.15***	
			return			

Source: Own work.

* Weak correlation at 0.1 < r < 0.3; ** Medium correlation at 0.3 < r < 0.7.

(1) – (3) correlations between return and CDR for local events in Table 19.

Granger causality results

	CDR	CP 2077	(1)	(2)	(3)	
negative	**0.0002****	**0.0002****	**0.0075****	0.7216	0.4240	p-value
positive	**0.0325***	**0.0325***	**0.0339***	0.5269	0.7734	
compound	**<0.01****	**<0.01****	**0.0255***	0.6775	0.6438	
anger	**0.0009****	**0.0009****	**0.0242***	0.7940	0.5415	
anticipation	**0.0015****	**0.0015****	**0.0135***	0.3073	0.4851	
disgust	**0.0010****	**0.0010****	**0.0274***	0.7654	0.6714	
fear	0.3073	0.3073	**0.0308***	0.0085	0.4032	
joy	**0.0032****	**0.0032****	0.0628	0.6363	0.5641	
love	0.1066	0.1066	0.1690	0.6986	0.5347	
optimism	**0.0072****	**0.0072****	0.1970	0.3283	0.2793	
pessimism	0.0851	0.0851	**0.0100***	0.5799	0.9868	
sadness	0.0643	0.0643	**0.0023****	0.2582	0.8316	
surprise	**0.0308***	**0.0308***	0.1880	0.1230	0.0882	
trust	**0.0045****	**0.0045****	0.0890	0.2981	0.7344	
			return			

Source: Own work.

* Significant at $p < 0.05$; ** Significant at $p < 0.01$.

(1) – (3) causality between return and CDR for local events in Table 19.