



Dipartimento
di Marketing Analytics & Metrics

Cattedra di Customer Intelligence & Big Data

Console War: PS5 vs Xbox Series X

A sentiment analysis

Prof. Luigi Laura

RELATORE

Prof. Marco Querini

CORRELATORE

Roberto D'Amore
Matricola 722111

CANDIDATO

ANNO ACCADEMICO 2020/2021

Alla mia famiglia.

A Miei genitori che non mi hanno mai fatto mancare nulla.

A Nonna Giovanna, da sempre collante della famiglia.

Al caro Luigi, non sai quanto mi piacerebbe che tutto tornasse come prima.

Agli amici, quelli che ci sono sempre stati e ci saranno sempre.

Al prof. Laura per i suoi preziosi consigli e il suo supporto nella stesura.

Grazie.

INDEX

INTRODUCTION	8
CHAPTER 1: TEXT MINING TECHNIQUES.....	10
1.1. Unstructured data	10
1.2. Text mining.....	11
1.2.1. Text mining techniques.....	14
1.3. Natural Language Processing	15
1.4. Sentiment Analysis.....	16
1.4.1. Lexicon Based Techniques	19
1.5. Machine Learning	20
1.5.1. Traditional approaches	20
1.5.2. Deep Learning approaches	23
2. CHAPTER 2: CONSOLE WAR.....	26
2.1. Console War.....	26
2.1.1. SEGA vs Nintendo.....	27
2.1.2. Sony vs Microsoft.....	29
2.2. PlayStation 5 vs XBOX Series X/S.....	33
2.2.1. PS5 SWOT Analysis	34
2.2.2. Xbox Series X/S SWOT Analysis	38
2.2.3. Technical performance at comparison.....	41
2.2.4. Marketing strategies at comparison.....	42
2.3. Consoles shortage	43

3. CHAPTER 3: SENTIMENT ANALYSIS RUN	45
3.1. Twitter API.....	45
3.2. Sentiment Analysis with R.....	46
 4. CHAPTER 4: FINDINGS.....	 53
4.1. Findings about PS5.....	53
4.2. Findings about Xbox Series X.....	57
4.3. Discussion about findings	60
 CONCLUSION.....	 62
 APPENDIX.....	 63
R script.....	64
 BIBLIOGRAPHY	 81

FIGURE LIST

Figure 1: Predicted growth of unstructured data (Source: forbes.com).....	11
Figure 2: Venn Diagram of Text Mining (Turegun, 2019).....	12
Figure 3: Text mining process steps (Source: Research Gate)	14
Figure 4: Sentiment Analysis Methodology Use (Amado et al, 2017).....	17
Figure 5: Taxonomy of sentiment analysis techniques (Bhavitha, et. Al, 2020).....	18
Figure 6: Differences between two classification approaches of sentiment polarity, machine learning and deep learning (Dang, Garcia, De La Prieta, 2020).....	23
Figure 7: Deep Neural Network (Dang, Garcia, De La Prieta, 2020)	24
Figure 8: Convolutional Neural Network (Dang, Garcia, De La Prieta, 2020)	25
Figure 9: A recurrent neural network (RNN)	25
Figure 10: Magnavox Odyssey, first ever home console released (1972).....	26
Figure 11: Mario (Nintendo) and Sonic (Sega): the two most famous games in 90s	29
Figure 12: PlayStation 1 (1994) and the iconic PlayStation logo	30
Figure 13: Playstation 2 vs Xbox (year 2001)	30
Figure 14: Estimated Lifetime Sales of Each Console (Niu, 2021).....	32
Figure 15: SWOT Analysis.....	33
Figure 16: Playstation®5 in standard and digital edition	34
Figure 17: New PS5 fancy controller (Source: Sony)	35
Figure 18: Xbox Serie S (in white) and Xbox series X (in black).....	38
Figure 19: APIs functioning (Source: tibco.com).....	45
Figure 20: Twitter credentials.....	46
Figure 21: Bar Plot showing the count of words in the text, associated with each emotion (Source: RStudio).....	48
Figure 22: Sentiment Score Definition	49
Figure 23: Barplot of sentiment score (Source: RStudio).....	50
Figure 24: R output of sentiment score frequency (Source: RStudio).....	50
Figure 25: Wordcloud of tweets related to PlayStation 5 (Source: RStudio)	51
Figure 26: Wordcloud of tweets related to Xbox Series (Source: RStudio).....	52

TABLE LIST

Table 1: Worldwide PS5 Sales in millions, data retrieved in January 2022 (Source: SIE)	37
Table 2: Comparison table between Microsoft Xbox Series X and Series S.....	38
Table 3: Worldwide PS5 Sales in millions, data retrieved in January 2022	41
Table 4: Emotion classification PS5	53
Table 5: Sentiment Score Table (PlayStation 5)	54
Table 6: Word Frequency of some relevant words (PlayStation 5).....	55
Table 7: Relevant topic according to Word Frequency (PS5)	56
Table 8: Emotion classification Xbox Series X S	57
Table 9: Sentiment Score Table (Xbox Series X S).....	58
Table 10: Word Frequency of some relevant words (Xbox Series X S).....	59
Table 11: Relevant topic according to Word Frequency (Xbox Series X S)	59

INTRODUCTION

Today, due to the advancement of information technology, social media are even more popular and create a forum where people can freely exchange and express ideas and opinions, connect with other people around the world to seek advice and support.

Especially social networks like Twitter or Facebook have also a huge impact on customers' decisions, leading brands and organizations to incorporate information found on these platforms in their marketing strategies.

In addition, businesses across the world generate a huge amount of data, merely through their online presence. These data come from multiple sources and can be stored in big data warehouses or on cloud platforms. Traditional tools and methods are often not sufficient to analyze such amount of data that grows exponentially minutes by minutes, representing a major challenge for worldwide companies.

Text mining is probably the best tool to identify hidden patterns, uncovers relationships, and makes inference relying on patterns discovered through a deep analysis within different layers of textual big data. One of the main text mining applications is Sentiment analysis (or opinion mining), a natural language processing (NLP) technique used to determine whether data are positive, negative or neutral. Sentiment analysis is generally performed on textual data to help businesses monitor product and brand sentiment, looking at customer feedbacks, and investigate their needs and desires.

The challenge of my thesis is to perform a sentiment analysis to analyze the gaming console market, focusing mainly on PlayStation 5 and Xbox Series X both released in November 2020. Both were worldwide sold-out because of the shortage of units produced, generating considerable controversy. So, it could be really interesting exploring social media reactions, especially on Twitter: the most relevant social network for opinion mining.

The work will be divided into four chapters:

- 1) The first one, articulated in four paragraphs, is introductory. It will address the phenomenon of *text mining* and one of its main applications: *Sentiment Analysis*, a natural language processing (*NLP*) technique performed on textual data to determine whether data is positive, negative or neutral.

- 2) In the second chapter, after arguing about the famous Console War in the first paragraph; in the next sections PlayStation 5 and Xbox Series X, the two consoles object of the sentiment analysis, are presented through a SWOT analysis and focusing on some insights about technical features, marketing strategies and trend purchases. Further, there's an analysis about reasons why PS5 and Xbox Series X are worldwide sold-out and hence so difficult to buy.
- 3) In the third chapter I'll introduce how I obtained a Twitter API and then I'll explain how I performed the sentiment analysis using RStudio through sentiment scores, emotion classification and word clouds
- 4) In the final chapter I'll display the findings and then I'll discuss the results of analysis.

CHAPTER 1: TEXT MINING TECHNIQUES

1.1. Unstructured data

“Big Data is currently globally spread and widely accepted, representing also a synonym of vanguard in terms of information management, although this does not come without controversy” (Fan and Bifet, 2016). Nowadays Big Data is everywhere, both in the form of structured data, such as organizations traditional databases (CRM) or unstructured data, driven by new communication technologies and user editing platforms (text, audio, images and videos). Enormous amounts of data are collected and stored by organizations, with the hope of being useful in the future. However, the vast majority of new data being generated today is unstructured, prompting the emergence of new platforms and tools capable of managing and then analyzing all those data. Due to those tools organizations can effectively take advantage of unstructured data for business intelligence (BI) and analytics applications.

Unstructured data simply refers to the fact that data are not organized in a typical structured database format. They can be textual or non-textual; human-generated or machine-generated. Can be thought such as information that doesn't follow conventional data models, making it complex to store and manage in a mainstream relational database. Unstructured data has an internal structure but does not contain a predetermined data model or schema. Unstructured data stored often contain a lot of valuable but still hidden insights and there are many datasets that aren't considered useful until their importance is realized with an appropriate analysis. However, there's no doubt that the most common types of unstructured data are text which can have multiple forms: social media posts, Word documents, emails, PowerPoint or Canva presentations, survey responses, etc. Text information is a source of information that is ideal to communicate ideas and thoughts, express emotion in an accurate shape. Hence, textual data has an undoubted quality and utility.

Unstructured data makes up a whopping 80% or more of all enterprise data, and the percentage keeps growing. They will power analytics, machine learning, and business intelligence. Recent prediction said that unstructured data should increase by 175 billion zettabytes (1 zettabytes is equal to 10 billions of terabytes) by 2025.

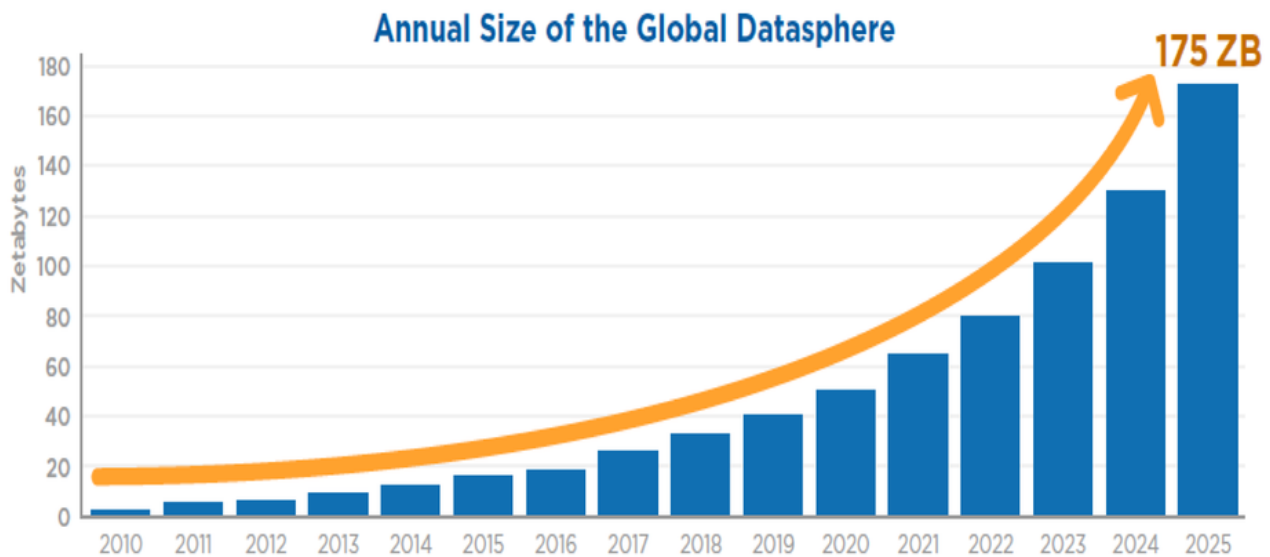


Figure 1: Predicted growth of unstructured data (Source: forbes.com)

The continuous growth of unstructured data is the result of the incorporation of intelligent agents that exploits machine learning or other forms of artificial intelligence to analyze the growing amount of data coming from the digital devices that we use every day. These data often come from virtual assistants or IoT devices inside our houses, factories and cities, which can create high video resolution content or even augmented reality. This is made possible by building networks and data centers capable to analyze, communicate and finally store a similar big amount of data. The creation of all this digital technology is nowadays called “*digital transformation.*”

1.2. Text mining

A big challenge for marketers is to manage such loads of data and extract appropriate and useful knowledge supporting decisions. A good text mining work provides valuable insights to help marketeres in understanding what is the most suitable product for a specific market and what is the best marketing strategy to advertise a product. Data online can be generated by every user, becoming a considerable non-filtered channel of knowledge for marketers’ analysis.

Text mining can be defined as the “*process of extracting interesting and non-trivial patterns from huge amount of text documents*” (Talib, et al., 2016) ¹. Text mining aims at processing text data whose principal characteristics are sparsity and high dimensionality. The major interest is linked to the fact that sometimes text allows to draw conclusions that otherwise wouldn’t be obtained from any other means of communication. From this perspective, text mining can be thought as gaining new information from a

¹ Text Mining: Techniques, Application and Issues. (2016)

certain database. Anyway, text mining can't be considered as a standing-alone discipline but rather has a strong connection with other disciplines like data mining, statistics, machine learning and computational linguistics. Below the Venn diagram is presented. It helps to contextualize text mining inside the broad data and computer science area.

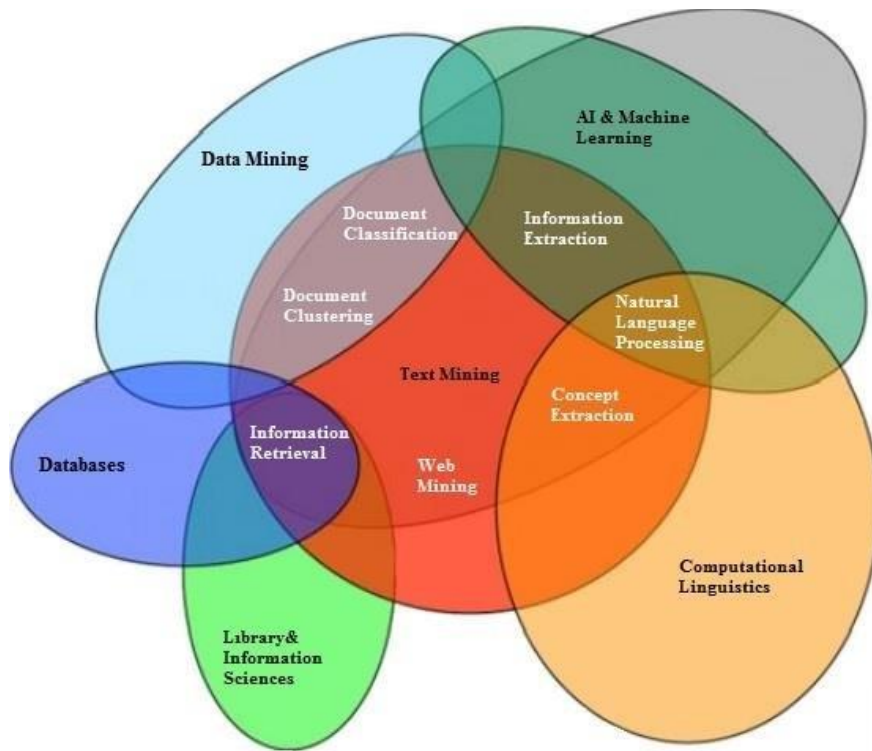


Figure 2: Venn Diagram of Text Mining (Turegun, 2019)

Text mining involves a series of five activities to be performed in order to efficiently “mine” the information:

1. **Gathering:** Depending by the different goal of the research, it's the collection of the useful data from different sources, like website, customer reviews, social media posts, books, document. With a proper application, the collection can also be fully automated.
2. **Pre-processing**, such as content identification / extraction of representative features. Once data have been collected, they need to be pre-processed so unnecessary information are removed, allowing the programming tools and algorithm to give us better qualitative output. Specific technique can be applied:
 - **Text clean-up** (also known as **stop word-removal**): it involves removing non useful or unwanted information such as ads from pages or words like articles, prepositions, pronouns, html links from the text. *Stop words* are words (is, in, to etc.) which do not hold any kind

of opinion. When we encounter some items with no value for further processing, it's better to get rid of them. Even words that occur very frequently can be deleted when necessary.

- **Stemming:** Stemming is a task that involves reducing words into their base form, or root. *Stem* refers to the set of words that have equal or similar meaning. For instance, past forms of verbs or plural nouns may be translated in their root form thanks to the stemming process. An example could be “waiting”, “waits” and “waited” that are all converted into “wait”, same for "games" converted into "game".
 - **Tokenization:** a computer only reads text as a string of characters, without for instances being able to identify various paragraphs, sentences or words inside a text. Tokenization splits the text in meaningful entities (words, sentences, etc.) considering white spaces and punctuations. Anyway, splitting a sentence into multiple unit parts can be translated in a change of original phrase semantic meaning. Thus, its adoption is not always useful for the goal of the research.
 - **Lemmatization:** it consists in returning words into their “lemma”, or dictionary form. It's a clearly refined process that groups together synonyms into a single word and at the same time convert verb in the infinitive. For example, “bought” are converted in “buy” and “worst” becomes “bad”.
 - **Feature extraction** (also called attribute selection): it involves the characterization of the text in order to obtain a set a quantitative measurement. For example, type of words, frequency of words in a text, syntactic information. These features can then be used for concrete further processing.
3. **Index:** introducing an index of accurate terms, to have an easy and quick access and to structuring the data under exam. Indexing can be defined as “*act of processing a text in order to extract statistics considered important for representing the information available and/or to allow fast search on its content*”.² Text indexing can be performed not only on natural language texts, but virtually on any type of textual information like for example stored in traditional database systems.
 4. **Mining:** at this point, text has been properly pre-processed and can now be ‘mined’. For that, is possible to apply different data exploration techniques to reveal new knowledge. These can, for example, include identifying mention of specific terms, linking of these terms with a dictionary and identifying relationships between different terms. This step involves usage of AI and machine learning tools or algorithms.

² Definition retrieved from Springer Encyclopedia of Database Systems

5. **Analysis:** is the final step, dataset is now ready to be analyzed but we still have raw results. Human intelligence is necessary to draw meaningful results from the analysis. However, these data need to be organized, evaluated, visualized, and finally interpreted looking at the initial purpose the text-miner wants to study in deep.

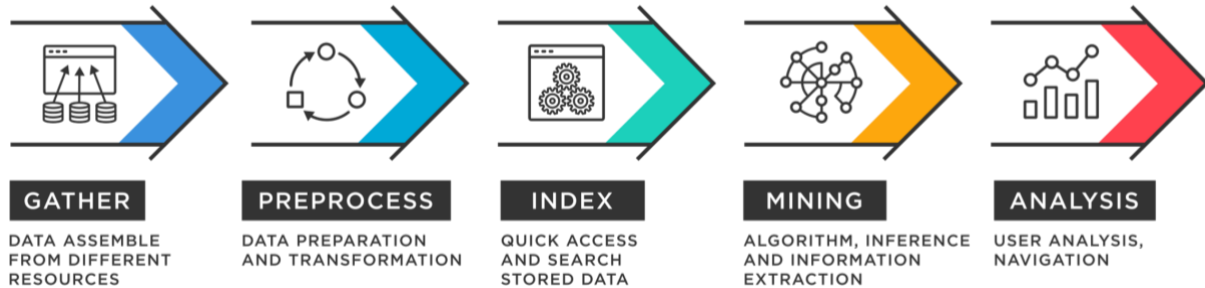


Figure 3: Text mining process steps (Source: Research Gate)

1.2.1. Text mining techniques

Different text mining techniques are available for analyzing the text patterns and their mining process. Most used are:

- Information extraction (IE)** that attempts to extract valuable information from a massive amount of textual data. It's the initial step to decipher unstructured text, discovering and underlining key phrases and recurrent relationships within text. Hence, once information has been extracted, is then stored in a database for future access and retrieval.
- Information retrieval (IR)** which involves the process of extracting relevant and associated patterns from the given set of words or phrases. IR systems exploits different algorithms to track and monitor user behaviors and discover relevant data accordingly.
- Clustering** in machine learning can be defined as *“the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. It is basically a collection of objects on the basis of similarity and dissimilarity between them”*.³ (Priy, 2021). In text mining, clustering seeks to identify intrinsic structures in textual information and organizes them into relevant subgroups or ‘clusters’ for further analysis. The most challenging task in clustering process is to build meaningful clusters from textual data without having any prior information on them. It should keep track of topics for each document, measuring at the same time weightage of how

³ Clustering in Machine Learning, (2021)

documents fit well inside each cluster. Cluster analysis often serves as a pre-processing step for other text mining algorithms running.

- d. **Summarization** that automatically generates a compressed version of a text which, however, summarizes the most important information for the end-user. The purpose is to browse multiple sources of text to create summaries of text which holds a substantial amount of information in a concise format, while keeping the general meaning and intent of the original documents essentially unaltered.
- e. **Categorization** which consists in gathering text documents, processing and analyzing them to uncover the right topics or indexes for each document. Depending on texts content, predefined classes are assigned to each text documents. Some useful analytical classification models, used to categorize text, are naive Bayesian classifier, decision trees, nearest neighbor classifier, and support vector machines (SVM).

1.3. Natural Language Processing

Natural language processing (NLP) is the branch of computer science and specifically of AI, which consist in giving computers, through some specific tools or algorithms, the capability to understand written text and spoken words as a human being can. Yse ⁴ defined NLP such as “*a field of Artificial Intelligence that gives the machines the ability to read, understand and derive meaning from human languages*”

Human language is of often full of ambiguities, thus is quite difficult to ideate a software that can determine, with a good degree of accuracy, intended meaning of text or spoken words. Let's think to metaphors, homonyms, grammar exceptions, sarcasm, sentence structure variations, which are only a small portion of all the irregularities of our language that we need years to learn well. We can only imagine how programmers' task is difficult since they have to develop natural language-driven applications, and those applications in order to be useful must be capable to recognize and understand faithfully every information from the beginning.

There are some NLP tasks that break down text and voice data, helping computers to make sense of what it's ingesting:

- **Speech recognition**: it involves converting voice data into text data. It is required for those applications which follow voice commands literally and can also answer to spoken questions (e.g. Amazon Alexa, Apple Siri, Google Home). Often people talk quickly, varying emphasis and intonation, sometimes even with an incorrect grammar: this is what mainly makes speech recognition a strong challenging task.

⁴ Your Guide to Natural Language Processing (NLP), (2020)

- **Part of speech (PoS) tagging**, also called **grammatical tagging**: it's about categorizing (as verbs, nouns, determiners, etc.) every word in a text in correspondence with a particular part of speech depending on the definition of the word and its context. For example, POS identifies 'work' as a verb in 'I work tomorrow,' and as a noun in 'I'm going work'
- **Natural language generation**: that is almost like the opposite of speech recognition; it involves converting structured information into human language.
- **Co-reference resolution**: it involves identifying situations in which two words refer to the same entity. A common example is determining the person or object to which a certain pronoun refers (e.g., 'he = 'Roberto), but it may also be about identifying a metaphor in the text.
- **Named entity recognition (NEM)**: which categorizes words or sentences as useful entities. For example, Named Entity Recognition identifies 'Rome1 as a location or 'Alessia' as a women's name.
- **Word sense disambiguation**: it involves identifying which sense of a word is used in a sentence. It happens through a process of semantic analysis that determine the word that makes the most sense in the given context. For example, in the two sentences: "The *bank* will not be accepting cash on Saturdays." and "The river overflowed the *bank*." The word *bank* in the first sentence refers to the commercial (finance) banks, whereas in second sentence it refers to the riverbank.
- **Sentiment analysis**: which has the main purpose to extract subjective qualities (such as amotions, feelings, attitudes, sarcasm, confusion) from text. It to determine positive or negative sentiment from a given data sources and to trace changes in customer behaviour over a specific tie period. The next section will study in deep sentiment analysis.

1.4. Sentiment Analysis

Sentiment analysis, commonly known as opinion mining or contextual mining, is a research and developing stream of Natural Language Processing (NLP) methods of the AI, which helps to identify, systematically extract and quantify, the subjective information. Sentiment analysis algorithms should be capable to identify a precise sentiment attached to a word by linking it to a pre-set dictionary.

Sentiment analysis might be applied on a lot of sources, especially on the web: social media platforms, magazine, influencer pages, blogs, and in general on any websites of interest. Especially social media are a relevant information source for a lot of decision-making processes and for the development of modern and creative marketing strategies It is a fast-growing trend in text mining since usage of text data of social media platforms has drastically increased in the last years.

Analysis of sentiment on social networks is a powerful means to learn about the users' opinions and has a wide range of applications; lot of interesting patterns about people can be mine from analysis and can

represent a mean to deeply analyze customer response to an event, a product, a news, etc. measuring like and dislike sentiments of the people.

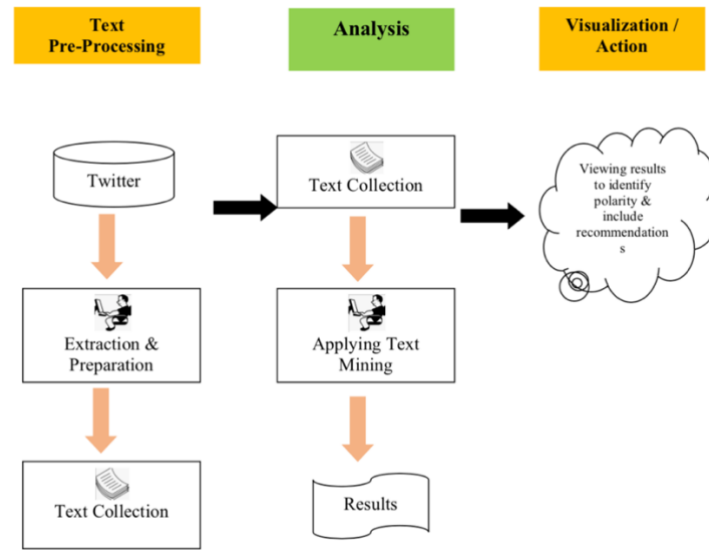


Figure 4: Sentiment Analysis Methodology Use (Amado et al, 2017)

According to Sharma, et al. ⁵ “Sentiment analysis uses three terms to define sentiment. These are, (1) object about which opinion is given, (2) features of that object, (3) opinion holder who give his opinion about the object. Sentiment analysis handles various challenges such as identification of the object, feature extraction and finds the orientation of opinion.”

Sentiment Analysis performs the classification task in 3 steps:

- *Document level*
- *Sentence level*
- *Feature level or Aspect level*

Document level of classification is performed when the main goal is to investigate the overall polarity of a topic irrespective of opinion holder. It assumes opinion about the single entity is expressed inside the text. This happens, for example, when we face with a movie review or a product review etc. where a document merely expresses the opinion about a single movie or product.

Instead, **sentence level** classification relies on the assumption that each sentence holds a single opinion. The sentence is a shorter form of document as collection of sentence makes a document.

To conclude, the **feature level** is performed through an analysis of various features of an object. For instances, let's think to a customer who bought a new iPhone, then he observed that the camera quality is

⁵ Sentiment Analysis Techniques for Social Media Data: A Review, (2019)

pretty good, whereas sound quality of the phone is not. When the purpose is to deep analyze different aspect of a topic, then an **aspect level** analysis is performed.

A sentiment may be represented by either a subtle or by a complex way in a text. The mixture of objective and subjective information on a given topic can generate noise (stop words, emojis, emoticons, ironies, etc.), which are commonly found in most of the available data sets, making it necessary to clean or modify them with specific techniques (de Oliveira Júnior, de Sousa Jr., de Oliveira Albuquerque et al., 2021).

Sentiment Analysis includes Data Pre-processing, Feature Selection, and Classification then find the polarity of data. Data pre-processing includes all the steps introduced in the previous paragraph: text clean-up, stemming, tokenization, lemmatization etc.

The figure above illustrates a taxonomy of methods for sentiment analysis.

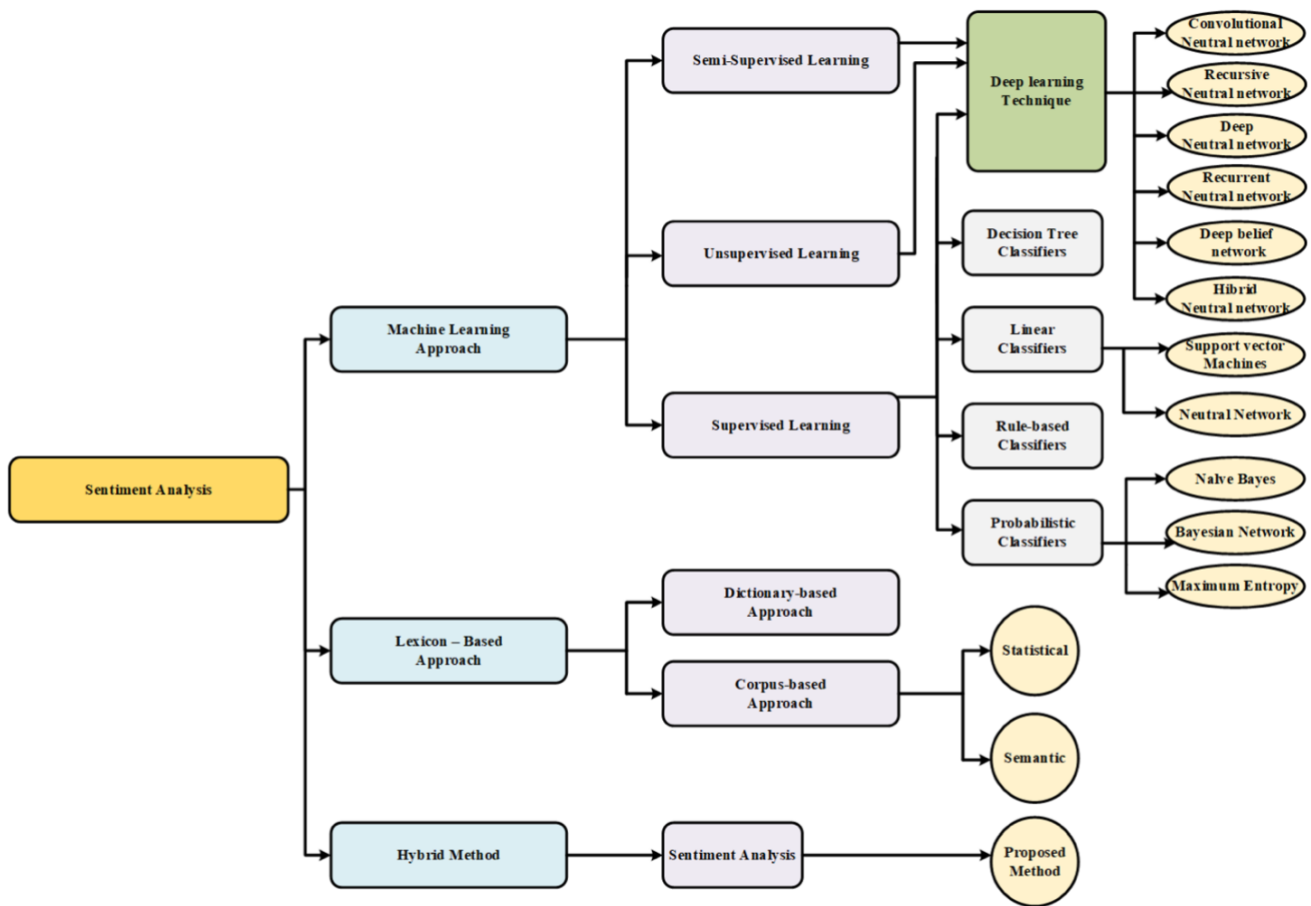


Figure 5: Taxonomy of sentiment analysis techniques (Bhavitha, et. Al, 2020)

According to Dang et al. ⁶, there are three major approaches to address the problem of sentiment analysis: (1) Lexicon-based technique, (2) Machine learning based technique and (3) Hybrid approach which combines the two previous approaches. Sentiment lexicons commonly play a key role in latter strategies.

1.4.1. Lexicon Based Techniques

Lexicon-based technique was the first to be used with sentiment analysis purpose. It is further subdivided into two different approaches: **dictionary-based** and **corpus-based**.

- On one hand, with dictionary-based sentiment classification is simply performed by using a dictionary of terms, distinguishing between terms with positive and negative valence. In addition, according to the topic some *seed word*, namely opinion words that are unique and important in a corpus, can be added to positive or negative list of terms.
- On the other hand, corpus-based sentiment analysis does not rely on a predefined dictionary but rather on statistical analysis of the contents of a collection of documents, using techniques based on k-nearest neighbors (k-NN), conditional random field (CRF), and hidden Markov models (HMM) among others.

Lexicon-based method is faster and widely implemented, although it has a downside since it doesn't consider a potential target for the sentiment expressed. A most common case of lexicon limitation is given by negations: for instance, the word "good" taken alone has a positive meaning, but if followed by a negation such as "not", its meaning could become negative. Dictionaries can easily be found surfing on the web; Some of the most used dictionaries in sentiment analysis studies are:

- (1) **SenticNet**: this dictionary provides a set of semantics, sentics, and polarity associated with 200,000 natural language concepts. Semantics define information associated with words, sentics define the connotative information associated with natural language concepts (i.e., emotion categorization values expressed in terms of four affective dimensions) and polarity is a floating number between -1 (negative) and +1 (positive).
- (2) **SentiWordNet**: this dictionary can assign different scores to the same word depending on its different meanings. For example, the word "attractive" can have at least two meanings: pleasing to the eye; having the properties of a magnet; SentiWordNet has multiple sets of scores, one for each meaning
- (3) **SentiWords**: this dictionary covered more than 150.000 English. Each of them has a polarity score that ranges from -1 and 1
- (4) **VADER**: Vader is a lexicon and a rule-based sentiment analysis tool for social media text. The lexicon has been built manually, by aggregating ratings coming from 10 human annotators. it contains just over

⁶ Sentiment Analysis Based on Deep Learning: A Comparative Study, (2020)

7000 words. Nevertheless, its precision should be higher than the resources created automatically. Moreover, being specifically tuned for social media, it also covers emojis and abbreviations (such as “lmao”, “lol”) that other dictionaries normally don’t. (Elia, 2020) ⁷.

- (5) **WordStat**: this dictionary is a little bit more complicated. A negative score is assigned not only to negative words but also to the precedent three words if there are no negations, or to positive words when preceded by negations within three words. The mechanism is the same also viewed from positive perspective.
- (6) **AFINN**: introduced in 2011, words scores range from minus five (negative) to plus five (positive). It consists of 2,477 coded words.

1.5. Machine Learning

Approaches Machine-learning-based techniques can be further divided into two groups: Traditional models that refer to classical machine learning techniques, such as the naïve Bayes classifier, maximum entropy classifier, or support vector machines (SVM) and Deep learning models that can even provide better results than traditional models. Such approaches address classification problems at the document level, sentence level, or aspect level.

1.5.1. Traditional approaches

Keyword-Based

It’s the simplest method, it uses a positive and negative keyword list and for each review there is a count about the number of negative, neutral and positive reviews. The polarity of the highest count is returned by classifier.

Naïve Bayes (NB) classifier

It’s the most used classifier. Based on allocation of the words in document, this categorization model can compute the subsequent probability of a class. This model works with the BOWs feature extraction ignoring the position of words in the document. Bayes Theorem is used to calculate the probability that specified feature set is part of particular tag (Wahyudi, Kristiyanti, 2016). This is shown by the equation below:

⁷ Sentiment Analysis Dictionaries (2021)

$$P(features|tag) = \frac{P(tag) * P(features - tag)}{P(features)}$$

Where:

- P(tag): It's preceding probability of tag or the possibility that arbitrary feature sets tag.
- P(features): It's preceding probability that specified feature set has occurred.
- P(features|tag): It's preceding probability that specified feature set is characterized as tag.

Considering Naive assumption which says that all the features are independent, the equation can also be written as:

$$P(features - tag) = \frac{PP(tag) * P(f1|tag) * ... P(fn|tag)}{P(features)}$$

Improved NB classifier was introduced by Kang and Yoo that solved the problem of trend of positive classification accuracy to be 10% more than negative classification accuracy. Therefore, when the accuracies of two classes are shown as an average value, the problem of reducing the average accuracy is formed. Naive Bayes is an easy model working well on text categorization problems (32) (33). Multinomial Naive Bayes model is used by us. In which Class c is assigned to review r (20), where value of c is as shown in this equation:

$$c^* = \arg \max_c P_{NB}(c|r)$$

$$P_{NB}(c|r) := \frac{P(c) \sum_{i=1}^m P(f_i|c)^{n_i(r)}}{P(r)}$$

In the last equation feature is f and the count of feature f_i found in review r is $n_i(r)$. m is total number of features.

Maximum Entropy (ME) Classifiers

This classifier is also known as a conditional exponential classifier. It translates labeled feature sets to vectors through encoding. This encoded vector after that is utilized to estimate weights of each feature that

are united to choose the most liable label of the feature set. ME classifier has parameters as a set of *weights* (W), which combines joint features generated from feature-set by *encoding* (E). The encoding maps each C{(featureset, label)} pair to a vector (Kang, Yoo, Han, 2012). To compute probability of every label we have used the following equation:

$$P(f|s) = \frac{\text{dotprod}(\text{weights}, \text{encode}(fs, \text{label}))}{\text{sum}(\text{dotprod}(\text{weights}, \text{encode}(fs, l)) \text{ for } l \text{ in labels})}$$

By using small amounts of training data Kaufmann (39) used it for detecting parallel sentences between any language pairs. Machine Entropy (ME) classifiers can generate constructive results for almost all language pairs. So that parallel corpora for numerous new languages can be created.

Maximum Entropy model is used with idea that we must choose the most uniform models that suit the given restraint (40) (41). They are feature-based models. Like Naives Bayes, ME formulate no autonomous assumptions for its features.

Maximum Entropy can handle enhanced feature overlap therefore it performs better than Naive Bayes theoretically. But, in practice, Naive Bayes (NB) classifier performs good on a range of problems

Support Vector Machines (SVM) Classifiers

SVM lies on the principle of determining linear separators in the search space that can fines divide the diverse classes. For SVM classification test data is well suited due to raw nature of text, in which some aspects are related to each other and are arranged into linearly divisible categories.

One of the important applications of SVM is to classify reviews according to their quality. Considering product reviews as a categorization problem, Chen and Tseng projected method for calculating information quality/value in product reviews. To find information-oriented feature set, on iPhone reviews, they also adopted an information quality (IQ) framework. They categorized reviews in provision of their quality.

Further, Li and Li used SVM as sentiment polarity classifier. They classified the opinions using SVM after recognizing and extracting the topics on Twitter in opinions related with queries of the users. They proved that for aggregating micro-blog opinions the consideration of opinion subjectivity and user credibility is necessary.

1.5.2. Deep Learning approaches

“Deep learning adapts a multilayer approach to the hidden layers of the neural network. In traditional machine learning approaches, features are defined and extracted either manually or by making use of feature selection methods. However, in deep learning models, features are learned and extracted automatically, achieving better accuracy and performance. In general, the hyper parameters of classifier models are also measured automatically” (Dang, Garcia, De La Prieta, 2020)

The picture above shows the differences in sentiment polarity classification between traditional machine learning approaches (Support Vector Machine, Bayesian networks, or decision trees) and deep learning approaches.

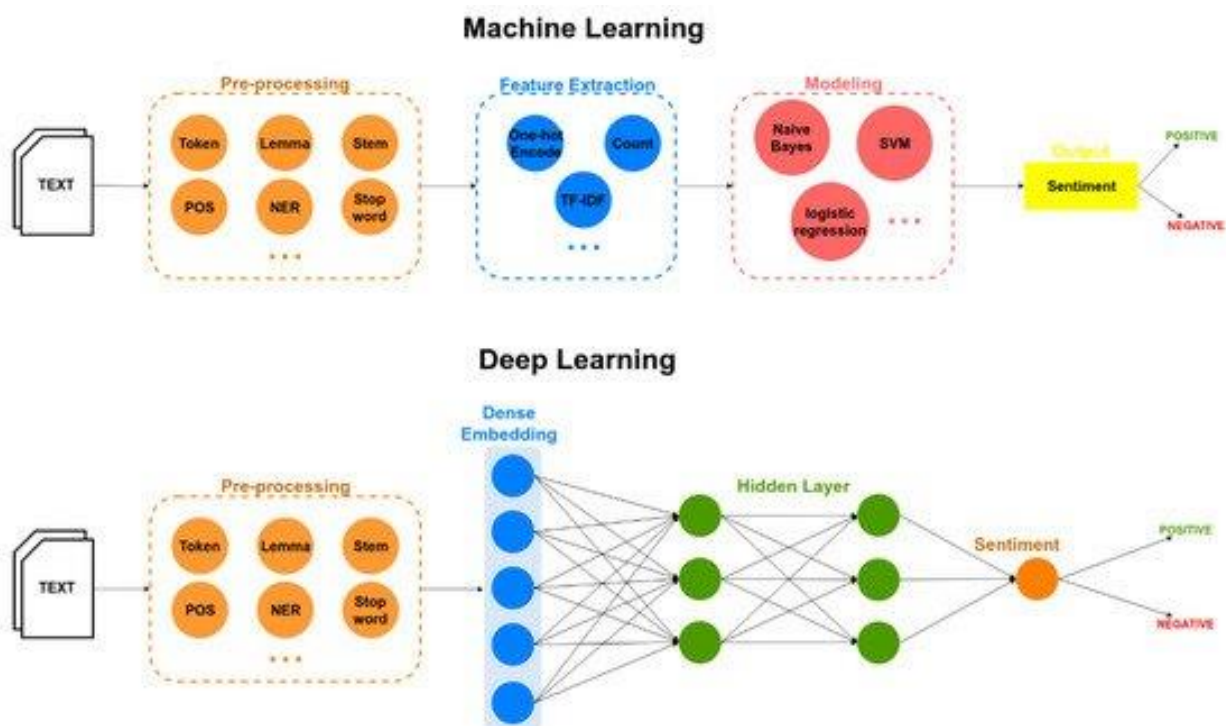


Figure 6: Differences between two classification approaches of sentiment polarity, machine learning and deep learning (Dang, Garcia, De La Prieta, 2020)

Neural Networks

Deep learning utilizes artificial neural networks to analyze data. “Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another. Artificial neural networks (ANNs) are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold. If the

output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.”⁸

Deep Neural Networks

A deep neural network (DNN) consists of a neural network with more than two layers, some of which are hidden. With DNN definitely complex mathematical modeling are involved in order to be able to process data in many different ways. A neural network is an adjustable model of outputs as functions of inputs, that is made by different layers: an input layer that includes input data; an hidden layers that includes processing nodes called neurons; and an output layer that includes one or several neurons, in which outputs are the network outputs.

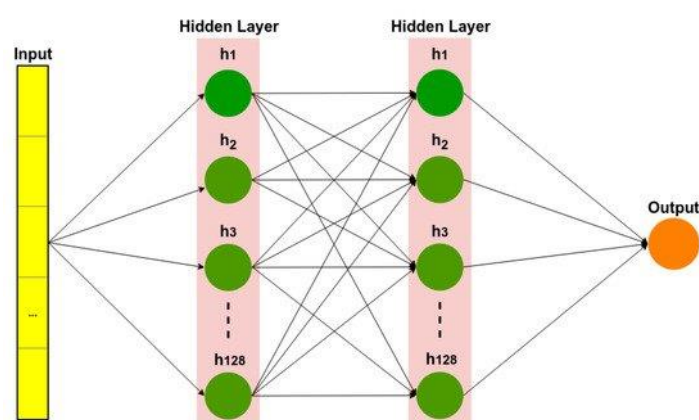


Figure 7: Deep Neural Network (Dang, Garcia, De La Prieta, 2020)

Convolutional Neural Networks

“A convolutional neural network (CNN) is a special type of feed-forward neural network originally employed in areas such as computer vision, recommender systems, and natural language processing. It is a deep neural network architecture, typically composed of convolutional and pooling or subsampling layers to provide inputs to a fully-connected classification layer. Convolution layers filter their inputs to extract features; the outputs of multiple filters can be combined. Pooling or subsampling layers reduce the resolution of features, which can increase the CNN’s robustness to noise and distortion. Fully connected layers perform classification tasks.”

An example of a CNN architecture can be seen in the figure above.

⁸ Definition retrieved from IBM Cloud Education (2020)

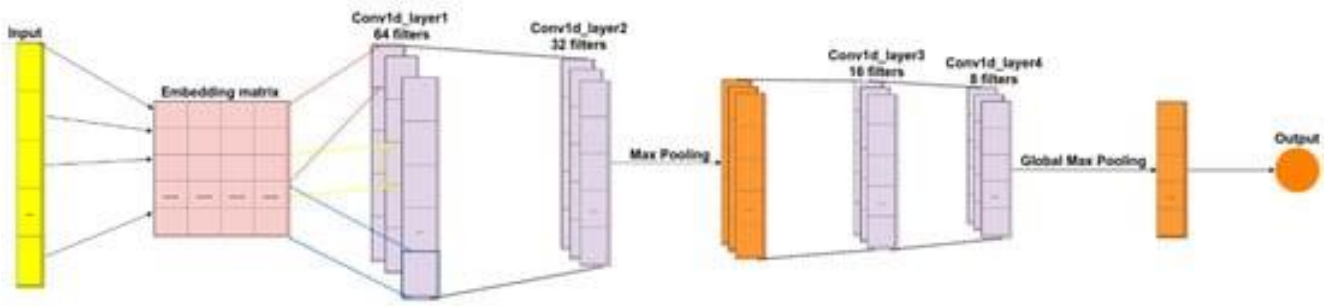


Figure 8: Convolutional Neural Network (Dang, Garcia, De La Prieta, 2020)

“The input data was preprocessed to reshape it for the embedding matrix. The figure shows an input embedding matrix processed by four convolution layers and two max pooling layers. The first two convolution layers have 64 and 32 filters, which are used to train different features; these are followed by a max pooling layer, which is used to reduce the complexity of the output and to prevent the overfitting of the data. The third and fourth convolution layers have 16 and 8 filters, respectively, which are also followed by a max pooling layer. The final layer is a fully connected layer that will reduce the vector of height 8 to an output vector of one, given that there are two classes to be predicted (Positive, Negative).”

(Dang, Garcia, De La Prieta, 2020)

Recurrent Neural Networks (RNN)

“Recurrent neural networks are a class of neural networks whose connections between neurons form a directed cycle, which creates feedback loops within the RNN”. The most important function of RNN is about the processing of sequential information based on the internal memory captured by the directed cycles. Unlike traditional neural networks, recurrent neural networks are capable to remember previous computation of information and reuse it by applying it to the next element in the sequence of inputs.

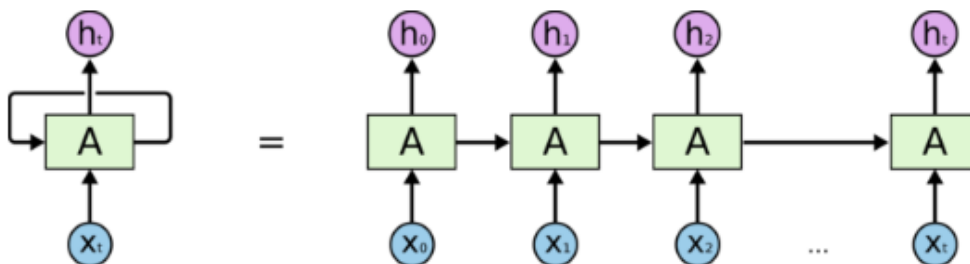


Figure 9: A recurrent neural network (RNN)

CHAPTER 2: CONSOLE WAR

2.1. Console War

In the video game industry, **console war** refers to “*the competition between two or more video game console manufacturers in trying to achieve better consumer sales through more advanced console technology, an improved selection of video games, and general marketing around their consoles.*”⁹

In these console wars videogame producers focused mainly on direct tactics to compare their offerings directly against their competitors with the attempt to disparage the competition in contrast to their own, instead of trying to outperform the direct competitor in sales. These wars saw different actors on different continents, since traditionally the four main markets (Japan, Australia, Europe and USA) have been treated as separate entities, with machines and games released at different times or even completely different games being released.

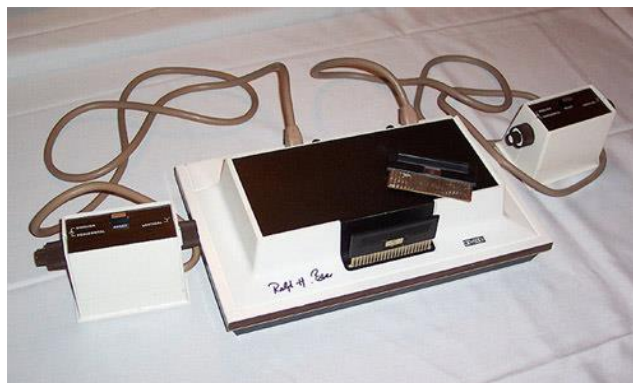


Figure 10: Magnavox Odyssey, first ever home console released (1972)

The video game console market started in 1972 with the release of the first home console, the *Odyssey* produced by Magnavox. As more manufacturers entered the home videogame console market and technology rapidly improved, market began to merge around the concept of console generations, groupings consoles with similar technical features competing inside the same consumer marketplace.

Media labelled their competition as an endless wrestling match, using really often terms such as “battle” or “war”. Initially, a strong focus was on the size of central processor unit, emphasizing that games had better capabilities with 16-bit processors rather than 8-bit ones.

⁹ Definition retrieved on Wikipedia (Console Wars)

2.1.1. SEGA vs Nintendo

Before Sony and Microsoft entered in the industry, the most famous rivalry was between SEGA and Nintendo during late 1980s and early 1990s.

Some years later the birth of home game console, videogame industry, especially in USA, suffered a severe market crash in first 80s due to the increasing popularity of computers as a videogame platform combined with a strong market recession and a lost in publishing control due the proliferation of unauthorized but compatible games. There was also a consistent drop in programmers team ability, thus it was difficult to ideate and sell good quality games.

In Japan, Nintendo released its Famicom console in those years, one of the first consoles of the third generation ("8-bit" generation). Nintendo looked to release Famicom in the United States but recognized that the market was still struggling from the crash. Thus, they had enough time to redesign the Nintendo, adding a special "10NES" lockout system that worked as a lock-and-key system to further prevent unauthorized games from being published for the system and avoid the loss of publishing control that had caused the crash of that period. The NES was able to revitalize the U.S. video game industry and established Nintendo as the undisputed dominator in videogame consoles over Atari. In his lifetime sales, the NES had sold about 62 million units worldwide.

At the same time, SEGA was looking to join the videogame business as well. Having been a successful arcade game manufacturer, after the downturn in arcade game industry, they attempted to exploit those experience in other useful ways. They released the SG-1000 console in Japan the same day of Famicom in 1983 but selling only 160k units in the first year. Sega redesigned the SG-1000 twice in order to challenge Nintendo's dominance, but unfortunately they failed to gain further sales.

In 1990 Nintendo released Super Mario Bros that further drove sales away from Sega.

Even if the superiority of Nintendo was undisputed, SEGA didn't give up in this console war.

SEGA tried to develop an iconic mascot character and build a game around, to challenge Nintendo's Mario mascot. They ideated Sonic the Hedgehog, a fast anthropomorphic character with an "attitude" that would appeal to teenagers and incorporating the blue color of Sega's logo. At this point, SEGA developed a new strategy for SEGA push with four key decisions, which included a strong price-cut of their console (Genesis) from \$189 to \$149 and started with aggressive marketing advertising to convince people that SEGA Genesis look "cool" over the Nintendo NES. Further, they pushed hard for American developers like EA Sports to create games on the Genesis that would better fit American preferences, particularly sports simulation games for which the console gained an excellent reputation. Finally, they insisted on making Sonic the Hedgehog the bundled game.

SEGA also renewed the advertising approach, with a young adult audience as a target, while Nintendo was still seen as a child-friendly console for many. Advertising focused mainly on Sonic, the edgier games in the Genesis library, and its larger collection of sports games. Their epic television ads for the Genesis and

its games always ended with the "Sega Scream"¹⁰, one character shouting the name "Sega" to the camera in the final shot to catch on attention quickly.

Another important factor was that Genesis had a larger library of games with over 150 titles compared to Nintendo SNES which had only eight games. A combination of those strategic moves led Sega to gain over Nintendo in the market.

In 1991 Nintendo understood and recognized that gradually they were losing ground. In response to SEGA, Nintendo tried to focus and develop advanced features and specific for SNES console that instead lacked in the Genesis (most famous was Mode 7 which introduced a simulated 3D perspective effects). A stock of 1 million of SNES was bought quickly and a total of 3.4 million of SNES were sold by the end of the year: a record for a new console launch. Despite all, SEGA Genesis still maintained strong sales against the SNES, convincing several Nintendo's third-party developers (e.g. Konami, Acclaim, etc.) to break their exclusive development agreements with Nintendo and seek out licenses to work also for Genesis. In 1992, Nintendo publicly acknowledged that they were no longer in the dominant position. They were forced to lower SNES's price to \$149 to match the Genesis, but in response Sega reduced in few time the Genesis price to \$129. In 1993 Genesis reached almost 55% of market share, an astonishing result, especially thinking that in 1990 Nintendo had 90 % of control over the market. Though Sega outperformed Nintendo in 1992/93, it still had corporate debt. In 1994, to continue to fight Nintendo, Sega's next console was the Sega Saturn, released in 1994, who incorporated technology of Sega's arcade games that used 3d polygonal graphics. They choose to put more effort and attention onto the Saturn line, neglecting Genesis despite its sales still being good. Suddenly, a new competitor emerged, Sony who launched the PlayStation in December 1994. Sega, aware of Sony's potential competition, had enough Saturn units ready to be sold, with the final goal of overwhelming Sony's offering.

Sega believed they had the stronger position over Sony, as gaming publications, comparing the Saturn to the PlayStation, rated the Saturn as the better system. Sega announced Saturn's various features and its selling price of \$399. When Sony announced that the launch price of PlayStation was \$299, this surprise price cut caught Sega off-guard (since they were sure \$399), and, in addition to several stores pulling Sega from their lineup due to being shunned from early Saturn sales, the higher price point made it more difficult for them to sell Saturn console. When the PlayStation officially launched in the USA in September 1995, its sales over the first two days exceeded what the Saturn had sold over the prior five months. Because Sega had invested heavily on Saturn into the future, Sony's competition drastically hurt the company's finances. In the case of Nintendo, their next offering was the Nintendo 64, which had powerful capabilities such as 3D graphics better than Saturn and PlayStation, but a not sufficient memory to store information for each

¹⁰ "Sega scream compilation" video at this link: <https://www.youtube.com/watch?v=Z6ibh4qWeWY&t=38s>

game. The first PlayStation game in the series, Final Fantasy VII, drove sales of the PlayStation, further weakening Nintendo's position and driving Sega further out of the market

By this point, the console war between Nintendo and Sega ended. In few years Sony disrupted the whole videogame market. Sega left the home console hardware business in 2002 to focus on software development and licensing. Nintendo remains a key player in the home console business, but more recently has taken a "blue ocean strategy" approach to avoid competing directly with Sony or Microsoft on a feature-for-feature basis with consoles like the Wii, Nintendo DS, and Nintendo Switch.

Anyway, SEGA vs Nintendo, in the collective imaginary, it's often seen as the first and original example of console war since it established the use of aggressive advertising and marketing tactics by the two companies with the specific goal of gaining control of the marketplace. Their rivalry led even to the spread of video game magazines. This war owes much of its fame to the iconic clash between Sonic and Mario. On one side the speediest hedgehog of the world; on the other side the plumber with the Mushroom Kingdom's most famous moustache.



Figure 11: Mario (Nintendo) and Sonic (Sega): the two most famous games in 90s

2.1.2. Sony vs Microsoft

Gamers, investors, and analysts use the terms *Console wars* to refer to the never-ending popularity contest between Sony, Microsoft, and (to a lesser extent) Nintendo. Countless articles have been written about PlayStation and Xbox's unending feud. Gamers debate it for weeks in online forums, and there's even an upcoming CBS documentary based on a book about the console wars. (Paez, 2021) ¹¹

This led to a generation of gamers who frequently adopted the expression "console war" to describe the constant marketing shots gaming console manufacturers take at each other.

¹¹ How a Retro Rivalry Ignited the Most Heated Feud in Video Games, (2021)

Sony made its entrance in the gaming console market with the leading PlayStation in 1994 in Japan and 1995 in Europe.



Figure 12: PlayStation 1 (1994) and the iconic PlayStation logo

Microsoft entrance in the console market

Sony progressed on with the next console, namely the PlayStation 2 in 2001. That same year, we saw the Microsoft entrance inside the video game console industry with its well-received Xbox, which featured also online gaming service, the Xbox Live. In 2001 Microsoft made his entrance in the home console market with the Xbox console. On one side PlayStation 2 was developed from mostly custom components, whereas on the other side Microsoft approached the Xbox as a highly refined personal computer based on Microsoft Window. Anyway, first Xbox did not achieve good sales result compared to PlayStation 2, selling only 24 million units worldwide against 155 million units of PlayStation 2 sold (that still today make PlayStation 2 the best-selling console ever). According to public opinion even if Microsoft did not reach good profit on the console hardware, they were anyhow satisfied with their performance thus they decided to continue to compete in the marketplace.



Figure 13: Playstation 2 vs Xbox (year 2001)

PlayStation 3 vs Xbox 360

Microsoft was able to take lessons learned from the first Xbox to its second model: Xbox 360 released in 2005 with an improved design, anticipating Sony's release of the PlayStation 3 in 2006. Microsoft was also able to ensure more first-party developers in its Microsoft Game Studios. On the other side, PlayStation 3 had less exclusives at launch and the higher price at launch represented a significant obstacle who gave Xbox 360 a consistent competitive advantage in first years after the release. However, unfortunately, Xbox 360 suffered from the "Red Ring of Death", a hardware fault on a significant number of his models, costing Microsoft over \$ 1.1 billion in repairs during the console's lifetime.

Further, in the attempt to include multimedia feature into high-definition movie playback, Microsoft made a huge mistake preferring HD-DVD standard for movie playback over the Blu-ray chosen by Sony, since shortly after the release, movies industry standardized on Blu-ray technology.

In addition, a fearful competitor was the Nintendo's Wii which presented an innovative technology with the WiiMote motion-sensing device. To react, both Microsoft and Sony released their own motion-sensing systems, respectively Kinect (Microsoft) and PlayStation Move, respectively, for their consoles. Both companies also released console refreshes mid-generation. Microsoft introduce two new versions: the low-cost Xbox 360 S, shipped with less internal storage space, and the Xbox 360 E, shipped with more storage space and the Kinect sensor; Sony release two different slim models of the PS 3 that reduced the system size with a subsequential retail price which helped improve sales.

Surprisingly, the Nintendo Wii overperformed the other two consoles with 101 million of units sold. Instead, the war between Microsoft and Sony ended with a substantial breakeven: Xbox 360 had 84 million units of estimated sales, while PlayStation 3 sold 87 million unit.

PlayStation 4 vs Xbox One

In 2013 a new chapter of the console war began: Sony released the new PlayStation 4 whereas Microsoft introduced the Xbox One.

Microsoft initially had the ambitious purpose to establish Xbox One as a replacement for a cable box in the living room, as a single source for entertainment with features aimed around television viewing in addition to gaming. In fact, Xbox One was coupled with Kinect and an always-on Internet connection as to enable numerous features, such as the ability to share games with other family members. But probably this project was even too ambitious: when these features were first advertised, there were strong criticisms from public opinion and consumers, which considered all these features as unnecessary and not privacy-complaint. Microsoft was forced to remove many of these features from the Xbox One before the launch, in particular they removed the always-connected requirement and the need to always use Kinect. Sony cleverly took the opportunity in their PlayStation 4 marketing strategy to play off Microsoft's missteps; simplicity of game sharing by simply passing along the physical media to another person and its lower price point were

successful keys. Microsoft was able to course-correct the Xbox One after launch, but in the meanwhile Sony PlayStation 4 gained enough ground also thank to a strong library of console-exclusive titles. For these reasons, PlayStation 4 largely overperformed the Xbox One, 106 million of units sold against 52 million units.

The graph below is a recap of the lifetime sales of most successful consoles.

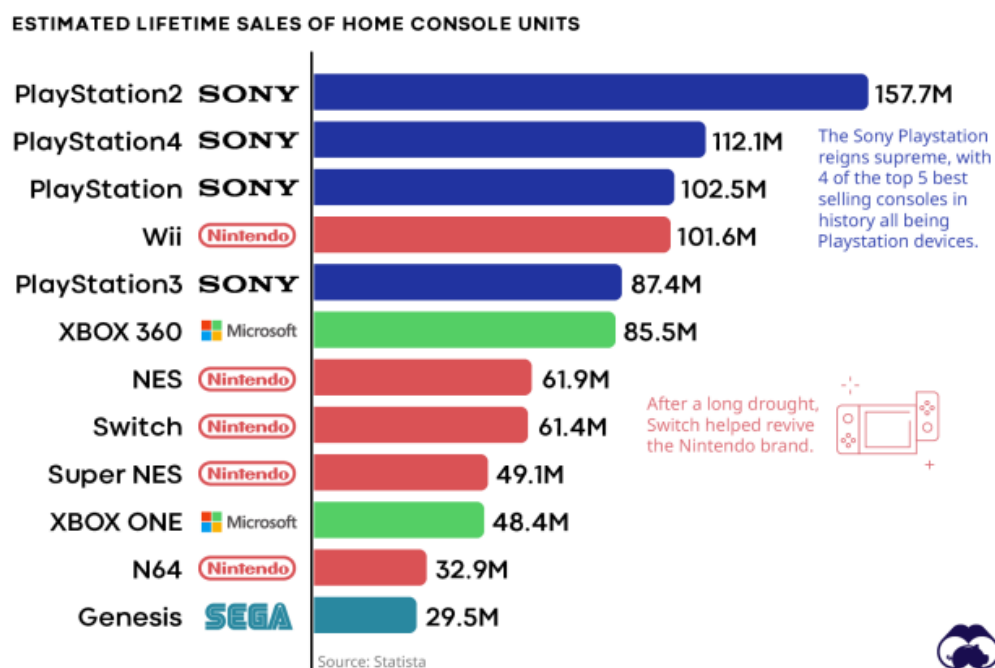


Figure 14: Estimated Lifetime Sales of Each Console (Niu, 2021)

Sony pricing game strategies over the years

A list of curiosity about Sony and competitors prices strategies over the years, extrapolated from an article by Perez on Forbes ¹², is here reported:

- The 1995 PlayStation 1 price tag, adjusted for inflation, would be about \$503.48 in US dollars today, meaning the PS5 may be priced in keeping with its predecessors after all. Console prices raised concurrently with the complexity of systems have evolved.
- Sony has always exploited price as a strategic mean to compete against its toughest competitors into the console market. Since the beginning, when Sony debuted with the first-ever PlayStation, they priced \$299, announcing it at the E3 video game convention, whereas at the same time Sega for its Saturn console chose the higher price of \$399.

¹² PlayStation 5 Rumored Price Tag Is \$500 As Console Pricing Wars Continue 2020

- For PlayStation 2, Sony was able to leverage its leading market share against Nintendo, Sega and Microsoft. They had the genial idea to guarantee backwards compatibility with PlayStation 1 games and controllers, pricing PlayStation 2 at \$299 in 2000 (\$443.29 in today's dollars) a year after Sega launched his discussed Sega Dreamcast at \$199.
- When in 2001 Nintendo released the GameCube at the competitive \$199 (\$289.38 today) to compete against Microsoft first Xbox at \$399 (\$577.29 today), Sony immediately reacted by discounting the one-year-old PS2 by a third in 2002, setting it on its way to becoming the best-selling console of all time with 155 million units sold.
- PlayStation 3 was the most expensive ever due the Blu-ray drive, costing up to \$599 (\$766.87 today), and launched later than Microsoft's Xbox 360, causing slow sales out the gate. Both consoles were decisively defeated by the innovative Nintendo Wii available at the competitive price of \$249 (\$318.78 today) in 2006.
- Sony's PlayStation 4 pricing of \$ 399 in 2013 against the \$ 499 for the Xbox One is maybe the main reason for the PlayStation 4 supremacy. In addition, Microsoft didn't help itself with several mistakes and bad decisions (which are mentioned earlier) in the lead up to launch.

2.2. PlayStation 5 vs XBOX Series X/S

Both companies released their next and last consoles in November 2020: PlayStation 5 and the Xbox Series X and Series S. Important technology improvements were present, including high-resolution and high framerates, high-speed internal storage, and backward compatibility with earlier systems.

To make a better comparison between the two console I'll perform a SWOT analysis to identify internal strengths and weaknesses and to investigate both internal and external factors that may affect their business.

How this analysis model work is shown in the figure below.



Figure 15: SWOT Analysis

2.2.1. PS5 SWOT Analysis

Sony has designed two versions of its next-generation console. For \$499.99 is possible to buy a standard PS5 with a 4K Blu-ray drive. But for \$100 less, at \$399.99, is also available the PS5 Digital Edition (which also looks noticeably thinner than its more expensive sibling). Unlike Microsoft, the only thing that differentiates those two versions is which console has a disc drive and how much each one costs.



Figure 16: Playstation®5 in standard and digital edition

Strengths

- + **Leader in console industry:** Sony's PlayStation is the favourite and most popular console according to a huge percentage of gamers. PlayStation has been the undisputed industry leader in gaming market for long time. PlayStation over the years became so iconic and still has a lot of strengths that attract customers.
- + **Exclusive PS title:** PlayStation 5 has a huge number of exclusive games thus gamers who want to play those games are forced to buy a PlayStation platform. Exclusive titles include blockbusters like Ratchet and Clank: Rift Apart, Final Fantasy 7, Marvel's Spider-Man and Demon's Soul. Sony, thanks to these exclusive agreements, has been able to attract many gamers to come and buy their device to enjoy these games.
- + **Exciting and revolutionary gameplay:** In addition to the exclusive range of games available only for PlayStation, gamers can experience a totally different, innovative, and often improved gameplay when playing those games on PlayStation 5 rather than on other consoles. This is due by two main factors:
4K resolution and a wonderful technology of their controllers. 1) **Controller:** PlayStation has dual shock controllers, which can give gamers more enjoyment and a significant improved game experience, also compared with Xbox controllers. Considering for example Ratchet and Clank, PlayStation

controller gives the real feeling of shooting with a gun. Depending on how hard the trigger is held in, a player will either utilize the gun's primary or secondary function. This helps each weapon feel unique. In addition to an eccentric design released in quirky colors like Starlight Blue, Nova Pink and Galactic Purple (pictures below), Sony PS5 new dual sense controller offers haptic feedback and dynamic adaptive triggers which ensure a gaming experience improved in a very large degree. 2) **Graphics and Resolution:** An amazing gaming experience shouldn't leave out the graphic. Since 2017 PlayStation was the first console that touched 4k technology, reaching right away compatibility with 4k televisions and displays. Even if other consoles like new Xbox are catching up with the 4k resolution, Sony's technology will be more mature and stable for long time. PS5's GPU (Graphical Processing Unit) is around eight times more powerful than that of PS4. The combinations of ultra-HD resolution and super graphics gives customers the ultimate gaming experience.



Figure 17: New PS5 fancy controller (Source: Sony)

- + **Price:** Sony's PlayStation represent the immediate second-best alternative for a gaming PC. However, we have to take in consideration that a good gaming PC which support high refresh rate and high resolution would cost over thousand dollars, which equals to at least two PlayStation 5 (which can be bought for 499.99 €/€). Therefore, the cost-effectiveness will always be a point to attract customers.
- + **Backward compatibility and customer service:** Last but not least, PlayStation's customer service is efficient and extremely user-friendly. To be user-friendly, Sony each month on PlayStation Plus (a subscription service from PlayStation, which allows online gaming) provides gamers free access to 3 to 5 games per month and to around 20 games from the PS4 catalogue. In addition, PS5 supports streaming services like Netflix, Amazon Prime, Apple TV, etc. so customers could enjoy with movies, watch sports and TV shows. Finally, Sony perfectly acknowledges that technology is developing surprisingly fast and there will be higher generation console come out, thus they are able to ensure gamers backward compatibility for their device (PS4 games are compatible with PS5), which translates into the fact that each PlayStations has a relative longer life cycle.

Weakness

- **Games prices:** Although price of the PlayStation 5 is quite lower than PCs, players should also take in consideration the prices of games. Basically, the games on the PlayStation platform have higher prices. Even the same game price can be different depending on the store.
- **Limited Storage:** The size (in term of memory) of the games is quite big, taking up till 100GB of system storage. Thus, there may be some issues with the current storage system given that it only offers 825 GB of storage (of which only 667 GB are usable). Hence console memory is likely to fill up quickly.
- **PCs competition:** Anyway, if a gamer does not matter about money since he wants the best possible game experience, PC is ahead about technology. A high-end PC is for sure much higher in resolution and gaming technology than the latest PlayStation. In addition, PlayStation is not very suitable for those gamers who do not prefer using controllers to play games.

Opportunities

- **Covid-19 pandemic:** The background today is the pandemic, with the majority people that are stuck at home and need something to kill the time. According to this favorable scenery, videogaming industry gets good opportunity to develop. Sony caught that opportunity, releasing PlayStation 5. They were able to offer *Play At Home* activity that give some free games to players and encourage people to spend more time at home to protect themselves from SARS-CoV-2 virus.
- **Virtual reality:** VR is nowadays a hot topic among a lot of industries, including cinema and gaming. Platform like Steam has already released latest at-home VR devices. Sony should be able to avoid wasting time in developing VR technology application which will lead to the launch of its own virtual reality console, attracting to their brand also consumers that wanna face with new virtual reality experience.
- **Multitasking:** Apart gaming, PS 5 can be used as well as device to watch movies and TV shows. It has established strong cooperative relations with allies like Netflix and Disney. Next step should be the building up of its own platform, which can provide PS5 owners a series of Sony prime exclusive services (movies, series, sports). This will offer additional possibilities of entertainment and give a push to the expansion of target customers. Excluding gamers, whole family would exploit these PlayStation new functionalities.

Threats

- **Drop in sales:** Other platforms (especially Xbox) and their sales amount represent a threat. PlayStation is now dropping sales: from the release of PS4 in the far 2010, PlayStation console's sale became less and less. Concerning about PlayStation 5, this problem is still more serious since only a limited number

of customers can get their own console because of a restricted number of units produced. Limited availability, as often happens, leads to off the market prices in secondary market platform such as eBay or StockX (a popular reseller site who claims more than 130k of console sold). As a consequence, gamers are forced to choose PC and other console like Xbox S or Nintendo Switch.

- **Competitors:** As mentioned above, PC has both a technology advantage and lower prices of games, while mobile is more convenient. Further, since the release date of the PlayStation 5 is relatively new, the number of exclusive games on PlayStation is still limited when compared to Nintendo Switch, especially considering PS5. Sony should be worried about this and try to develop more exclusive games for their console in the next future.

Sales

Below, a table retrieved on the Sony Interactive Entertainment official website ¹³, shows the number of PS5 sold since his release on 12th November 2020 to the end of 2021: 25 million.

There has been a sharp decline in the past three quarters, and the main cause is the console shortage.

In next paragraph this aspect will be better analyzed.

	Q1 (April - June)	Q2 (July - September)	Q3 (October – December)	Q4 (January - March)	FY (April - March)
FY2020	-	-	4.5	3.3	7.8
FY2021	2.3	3.3	3.9		

Table 1: Worldwide PS5 Sales in millions, data retrieved in January 2022 (Source: SIE)

¹³ Available at <https://www.sie.com/en/corporate/data.html>

2.2.2. Xbox Series X/S SWOT Analysis

Xbox Series X and Xbox Series S are home video game consoles released on 10th November 2020 as the fourth generation of the Xbox console family.



Figure 18: Xbox Serie S (in white) and Xbox series X (in black)

The table below ¹⁴ illustrates the main difference between the two different series lineups. Series X can be considered as the real PS5 competitor, whereas Series S represents a valid alternative for customer who worry about price or are tired of waiting a bust in PS5 or Series X market availability.

Parameters of Comparison	Xbox Series X	Xbox Series S
Price	499\$	299\$
Performance	12 Teraflops of processing power.	4 Teraflops of processing power.
Storage	Up to 1TB.	Up to 512GB.
Accessories	Box includes 1 Ultra High-Speed HDMI cable.	Box includes 1 High-Speed HDMI cable.
Display Resolution	True 4K.	1440p.

Table 2: Comparison table between Microsoft Xbox Series X and Series S

¹⁴ Data retrieved by the New York Times article “Choosing the Right Xbox: Series X or Series S” (Gies, 2021)

Strengths

- + **Big player and loyal community:** Xbox has the second highest market share in the videogame industry and it is right behind Sony's PlayStation. In addition, Xbox's owners literally love the console, and many swear that Xbox is better than PlayStation.
- + **Appreciated exclusive games:** Some exclusive games of Xbox such as Halo Infinite and Gears of War series are considered among lot gamers as masterpiece games. They are so popular that many people buy Xbox just to play these exclusive games.
- + **Larger controller:** lot of gamers feel that the controller of an Xbox is its selling point mainly because the controller is large and fits the hand perfectly.
- + **Impressive backward compatibility:** Xbox Series X and Series S support all existing games playable on Xbox One, including Xbox 360 and original Xbox games currently supported through backward compatibility on the Xbox One, thus allowing the new consoles to support four generations of games. This ensures each Xbox console a significant long-life cycle.
- + **Deep pockets:** As Xbox comes from the house of Microsoft, it is safe to say that Xbox has deep pockets and can survive a price war for a long time.
- + **Competitive pricing:** Xbox has consistently sold its product at competitive pricing. It is priced below the PlayStation but above the Nintendo. New gamers might prefer an Xbox Serie S (the cheaper alternative) over a PlayStation 5 for the cheaper price, saving 200 €.

Weaknesses

- **Always right behind PlayStation in sales:** Xbox is always behind PlayStation in its overall sales and has rarely taken over the number 1 spot. However, although it is consistently at number 2 spot, a company like Microsoft always want to be the market leader in a technology segment.
- **Design:** Xbox Series X is generally considered as not having an attractive design. It is, for all intents and purposes, little more than a big black box. Microsoft may have packed a lot of power inside of it, but it doesn't seem like they put anywhere near as much effort in to designing the exterior
- **Games price:** \$499 is a lot of money to a lot of people. That's a common problem with PlayStation 5 since the price of games is very high on console if compared to desktop or online. As a result, many customers buy lesser games, or they buy it from the resale market thereby saving money. Although Xbox offers a cheaper alternative (Series X9), is reasonable to believe that the majority of gamers are going to want the most powerful version.

Opportunities

- **Virtual Reality:** Since VR is one of the hottest topic, it is also one of the most favorable opportunity. Virtual reality is with no doubt the future of gaming and social interactions and Microsoft needs to catch up soon via the Xbox. Microsoft still commercialize virtual reality headset and accessories but, at least for the moment, these latter are not compatible with Xbox Series X|S.
- **Lowering prices of games:** this is a strategy that could help Xbox in its console sales is to lower the selling price of games especially in developing markets where it is exporting the games. Doing so, Microsoft could gain advantage over its competitors where the cost of 10 games is sometimes equal to the cost of the whole console.
- **Exclusive games:** A reason why PlayStation is ahead of Xbox is the huge number of exclusive games launched on PlayStation such as God of war, Uncharted and others. Microsoft has Halo and Gears, but there are not enough games which are a platform seller for the Xbox.
- **Covid-19 pandemic:** as just said for PS5, pandemic situation is an ideal scenery for console industry with a lot of growth opportunities. New Xbox Series consoles give people an enjoying alternative to spend more time at home, protecting themselves from the pandemic.
- **Online Gaming:** Online gaming is catching up in a big way. Xbox should ideate its own arcade type online games to make his community happy to play.

Threats

- **Dropping console sales** – The worst threat is that Xbox is the dropping sale of consoles. The whole console market is dropping in the number of sales. In addition, only a limited number of Xbox Series X have been produced and put on the market compared to the demand which was much higher. Many consumers, tired of all this wait, may definitively give up on buying the Xbox Series.
- **Sony and Nintendo competition** – Xbox has been in constant competition with PlayStation. Microsoft is not able to do the further step to be number 1. In fact, it's quite far behind PlayStation and Nintendo could soon catch up. Thus, the market is strongly dynamic in nature and Xbox needs to reinvent itself from time to time.

Sales

On 10th January 2022, the CEO of Microsoft Gaming Phil Spencer said that “At this point, we’ve sold more of Xbox Series X|S than we had any previous version of Xbox.”

Even if there's no official sales numbers, several sources suggest that Xbox Series X/S selling are in the range of 11/12 million consoles since their launch.

2.2.3. Technical performance at comparison

The table below shows technical features of PS5 (standard and digital edition) and Xbox Series XS.

<i>Categories</i>	<i>PS5</i>	<i>PS5 (digital)</i>	<i>Xbox Series X</i>	<i>Xbox Series S</i>
CPU	Eight Zen 2 Cores @ 3.5GHz with SMT (variable frequency)	Eight Zen 2 Cores @ 3.5GHz with SMT (variable frequency)	Eight-core AMD Zen 2 CPU @ 3.8GHz (3.6GHz with SMT enabled)	Eight-core AMD Zen 2 CPU @ 3.6GHz (3.4GHz with SMT enabled)
GPU	AMD RDNA 2 GPU 36 CUs @ 2.23GHz (variable frequency)	AMD RDNA 2 GPU 36 CUs @ 2.23GHz (variable frequency)	AMD RDNA 2 GPU 52 CUs @ 1.825GHz	AMD RDNA 2 GPU 20 CUs @ 1.565GHz
GPU Power	10.28 TFLOPs	10.28 TFLOPs	12.15 TFLOPS	4 TFLOPS
RAM	16GB GDDR6 RAM	16GB GDDR6 RAM	16GB GDDR6 RAM	10GB GDDR6 RAM
Performance Target	Target TBD. Up to 8K. Up to 120fps	Target TBD. Up to 8K. Up to 120fps	Target 4K @ 60fps. Up to 120fps	Target 1440p @ 60fps. Up to 120fps
Storage	825GB PCIe Gen 4 NVMe SSD (5.5GB/s uncompressed, typical 8-9GB/s compressed). Usable storage is 667.2GB	825GB PCIe Gen 4 NVMe SSD (5.5GB/s uncompressed, typical 8-9GB/s compressed). Usable storage is 667.2GB	1TB PCIe Gen 4 NVMe SSD (2.4GB/s uncompressed, 4.8GB/s compressed). Usable storage is 802GB	512GB PCIe Gen 4 NVMe SSD (2.4GB/s uncompressed, 4.8GB/s compressed). Usable storage is 364GB
Expandable Storage	NVMe SSD slot	NVMe SSD slot	1TB expansion card	1TB expansion card
Backward Compatibility	The "overwhelming majority" of the more than 4,000 PS4 games. Some PS3 and PS2 titles playable via PlayStation Now.	The "overwhelming majority" of the more than 4,000 PS4 games. Some PS3 and PS2 titles playable via PlayStation Now.	"Thousands" of Xbox One, Xbox 360, original Xbox games. Xbox One accessories.	"Thousands" of Xbox One, Xbox 360, original Xbox games. Xbox One accessories.
Disc Drive	4K UHD Blu-ray	None	4K UHD Blu-ray	None
Display Out	HDMI 2.1	HDMI 2.1	HDMI 2.1	HDMI 2.1
Price	\$499 - €499	\$399 - €399	\$499 - €499	\$299 - €299

Table 3: Worldwide PS5 Sales in millions, data retrieved in January 2022

PS5 Digital Edition differs from Standard PS5 only for the absence of the disc drive. All consoles guarantee a good degree of backward compatibility. Anyway, that's not an aspect to ignore as the owners of Digital Edition will be permanently forced to buy games (often at a higher price) on the PlayStation Online Store against an initial saving of 100€. A factor that stands out for PS5 is the performance target that goes up to 8K, ensuring players an HD excellent graphics, devised for the latest generation of TVs.

Xbox Series S has more powerful CPU (Central Processing Unit) and GPU (Graphics Processing Unit) with a broader usable storage space (802 GB vs 667 GB). All consoles have a good degree of backward compatibility.

However, looking at the is evident how Xbox Series S cannot be considered at the same level of PS5 and her sister Series X, given the very limited storage, the disc drive absence, the reduced power of CPU, GPU, RAM and the lack of 4K/8K technology.

2.2.4. Marketing strategies at comparison

Xbox Series X were released on 10th November, while PlayStation 5 were released on 12th November 2020. Basically, Sony and Microsoft have been competing since day zero from market entry of their next generation console.

According to the data researcher MediaRadar, during the first two weeks of November, the videogame industry spent more than \$45 million on print, TV and digital ads. Especially Sony spent more than \$15 million to promote the PlayStation 5 in the month leading up to launch, more than triple what Microsoft spent on ads for the Xbox Series X|S.

According to Maheshwari, in the period just before launch Sony was better than Microsoft to create the right hype on social media, revealing just enough to keep their consumers talking. Sony kept it pricing strategy a mystery till the PS5 launch, opting for the price of \$499.99 (\$399.99 for the digital edition) after a long market analysis. Whereas Microsoft's price leaked to be at \$299 for the Xbox Series S (which, however, is certainly lower in terms of performance, as also highlighted in the technical comparison), and \$499 for the real competitor of PS5: Xbox Series X. Although, Microsoft pricing strategies of offering an alternative console on the market for a cheaper price was not wrong at all, they have never been able to outsell PlayStation 5. Overall, Sony has always tried to convince consumers that PlayStation is the “queen” of the consoles, and they have the best contents.

But, Sony and Microsoft haven't been able to keep up with demand because the coronavirus pandemic disrupted the supply chain for electronic components, leading to widespread shortage. In this favorable scenery, also a common competitor like Nintendo increased its social media presence trying to capture a consistent part of consoles demand.

However, both Sony and Microsoft came up with fresh marketing strategies that involved the online. For example, platforms such as Twitch and YouTube, where gamers streamed their live play, suddenly became very popular and attractive to marketers. Covid-19, which forced people at home, acted such as a mean of reaching a wider range of audience on these streaming platforms. Sony to build hype around PS5 and some of their best exclusive games, such the new Spider-Man game (which also featured the first black Spider-Man: Miles Morales), has entered into sponsorship agreements with multiple streamers.

2.3. Consoles shortage

It's been more than a year since PlayStation 5 and Series X were released, but it is still hard today to find one. Even people that book the console in a store, have to wait months and months before receiving it. The main reason of the lack of supply, relies on the continued semiconductor shortage. There has been a global shortage (due to Covid-19) affecting all industries, raising interest in electronic entertainment, ranging from video games to medical equipment. Microchip today goes in everything from electric cars to phones, including next generation games consoles like the Xbox Series X and PlayStation 5.

According to Robert Cialdini: *"Opportunities seem more valuable to us when their availability is limited"*. Cialdini suggested that the "scarcity", or limited availability of a product, produces a greater desire to try to obtain the thing itself. Often people establish a mental equation: rare = of value; the rarer something is, the higher its value. In addition, people like to own what others can't have. Given the situation, it's reasonable think that Sony and Microsoft are saving money and the remaining stocks, elevating at the same time PS5 and Xbox Series X in their image, importance, and prestige.

Further, the few times PS5 and Xbox Series X are available, console is often coupled with a bundle. Doing so, they successfully managed to increase sales games and accessories that otherwise most people wouldn't have bought. An example is GameStop which periodically sells a limited stock of consoles on its website. Bypassing the fact that finish in few minutes, the only option to have a PS5 is to pay 750€, buying a bundle with an additional controller, headsets and other games instead of 500€ as it should be.

For Sony a further indicator that states the lack of PS5 is instead a good deal for the company, comes from the stock values over the last two years. Before PS5 release, the average value of a Sony Group Corporation stock was about 70\$ per share; now it's above 100\$ per share.

Anyway, as everything has its positives it also has negatives. Scalpers represent a major issue, although some stores take drastic measures to counter next-gen scalpers, many simply let bots purchase consoles in a second. Over the last two years, lot of scalpers sold PS5 and Xbox Series X for insane prices forcing

people to pay so much, because of the lack of availability. In addition, consumers had to join several and queues on many websites just to have an opportunity to buy Series X or PS5. This led into an increasing dissatisfaction among users due to frequent server crashes and a lot of time wasted.

To conclude, this situation is probably benefiting Sony and Microsoft from an economic point of view, but on the other hand it's weakening their image and relationship with customers.

Given these premises, I decided to focus my analysis precisely on the next generation consoles (the PlayStation 5 and the Xbox Series X/S) because I strongly believe that is a real and interesting challenge to investigate the popular opinion and sentiment towards these consoles that despite being on the market for over 1 year are still extremely difficult to buy, generated enormous controversy in the last months.

Could it be true that waiting increases desire? Or are gamers losing patience? Let's see it in the next chapters.

CHAPTER 3: SENTIMENT ANALYSIS RUN

3.1. Twitter API

An API (Application Program Interface) is a software intermediary that allows two unrelated applications to talk to each other. They connect everything together and make software systems work in harmony.

APIs are mostly invisible to the business user but open up a wide range of possibilities for software programs. They work by opening a small portion of the software's features and data in a controlled manner. This allows developers to access that program, piece of hardware, data, or app without having to access code for the entire system.

The goal is to achieve a higher-level abstraction, usually between the hardware and the developer, simplifying the programming work. The API, in fact, allows developers to avoid rewriting every time from scratch (low level) all the functions necessary for a given program, thus falling within the broader concept of code reuse.

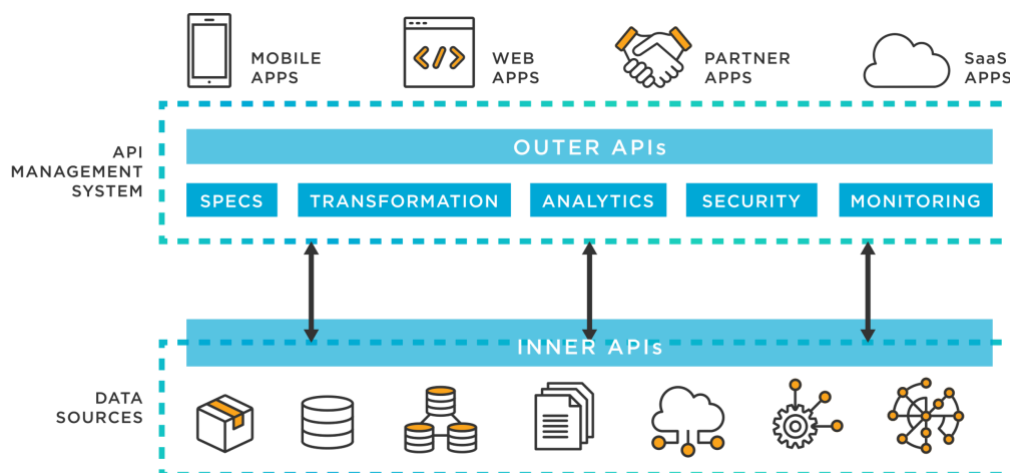


Figure 19: APIs functioning (Source: tibco.com)

Twitter API is an interface through which a website or an app interact directly with Twitter. It allows access to the main features of the social platform, such as posting tweets, retweeting, and also finding tweets that contain a particular word.

To obtain a Twitter API, the first step is to connect on Twitter application and to use the API service in order to get four secret keys (figure x.x).

Application Settings

Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

Consumer Key (API Key)

Consumer Secret (API Secret)

Your Access Token

This access token can be used to make API requests on your own account's behalf. Do not share your access token secret with anyone.

Access Token

Access Token Secret

Figure 20: Twitter credentials

Once obtained access credentials, the program can be implemented, and tweets can be processed. A tweet is a message with a maximum length of 208 characters.

3.2. Sentiment Analysis with R

R is a free software environment for statistical computing and graphics. It provides a wide variety of statistical and graphical techniques and is highly extensible.

R has a rich set of packages for Natural Language Processing (NLP) and to generate plots. The main steps involve loading the text file into an R Corpus, then cleaning and stemming the data before performing analysis. I will perform my analysis including Word Frequency, Word Clouds, Sentiment Scores and Emotion Classification using several plots and charts.

I collected data from December 1st to January 31st. Every day I analyzed with R 10.000 tweets related to the PlayStation 5 and 10.000 tweets related to the Xbox Series. I tried to be as punctual as possible, making the analysis each day in a time slot between 18.00 and 20.00 when in USA, the country where the majority of tweets comes from, is morning. I was forced not to cross the threshold of 10 thousand tweets per analysis because R cannot go beyond that limit. In this section I'll explain which techniques I used for the sentiment analysis, whereas findings will be discussed only in the next chapter.

In writing the R script I took a cue from GitHub Pages with the guide "Twitter Sentiment Analysis in R"¹⁵ of Shivani Modi and Sriram Dommeti as well as from the article "Text Mining and Sentiment Analysis: Analysis with R"¹⁶ of Sanil Mhatre.

¹⁵ Available at <https://jtr13.github.io/cc21/twitter-sentiment-analysis-in-r.html#twitter-sentiment-analysis-in-r> (2021)

¹⁶ Available at <https://www.red-gate.com/simple-talk/databases/sql-server/bi-sql-server/text-mining-and-sentiment-analysis-with-r/>

In R, the fundamental unit of shareable code is the package. A package bundles together code, data, documentation, and tests, and it is easy to share with others. The following R packages were used for my analysis:

- ~ **twitteR** which provides access to the Twitter API
- ~ **igraph** for simple graphs and network analysis
- ~ **tidytext** which makes many text mining tasks easier, more effective, and consistent with tools already in wide use
- ~ **tm** for text mining operations like removing numbers, special characters, punctuations and stop words (removing words that have no value for NLP and should be filtered out).
- ~ **wordcloud** for generating the word cloud plot
- ~ **networkD3** for network, tree, dendrogram, and Sankey graph
- ~ **rtweet** for interacting with Twitter's APIs
- ~ **snowballc** for stemming, which is the process of reducing words to their base or root form
- ~ **ggplot2**, **ggeasy** and **plotly** for plotting graphs
- ~ **magrittr** that make the code more readable
- ~ **widyr** that wraps the pattern of un-tidying data into a wide matrix, performing some processing, then turning it back into a tidy form
- ~ **readxl** which makes it easy to get data out of Excel and into R
- ~ **syuzhet** for sentiment scores and emotion classification
- ~ **lubridate** to work with date-times and timespans: fast and user-friendly parsing of date-time data, extraction and updating of components of a date-time (years, months, days, hours, minutes, and seconds).

As a first step, I needed to receive authorized credentials from Twitter to use the API for extracting the tweets. As I said in the previous chapter, I created a Twitter developer account and a custom app in order to obtain my personal credentials directly from Twitter.

At this point, though the proper script R extracted 10.000 tweets about PS5 (or Xbox RSeries).

Then, I started with the cleaning of the text data. First, I removed special characters from tweets, using the `tm_map()` function to replace special characters like `/`, `@` and `|` with a space.

Next step was about removing unnecessary whitespace and converting text to lower case.

Then I removed the stopwords: most commonly occurring words in a language which have insignificant little value in terms of gaining useful information. They should be removed before performing further

analysis. Examples of stopwords in English are “the, is, at, on”. I also removed frequent words with no value for my analysis like “dont, will, can, tweet, retweet, rt”.

Final last step was stemming. It is the process of reducing the word to its root form. The stemming process simplifies the word to its common origin. For example, the stemming process reduces the words “fishing”, “fished” and “fisher” to its stem “fish”.

At this point, tweets were ready to be processed.

1) EMOTION CLASSIFICATION

Emotion classification was built on the NRC Word-Emotion Association Lexicon.

The definition of “NRC Emotion Lexicon” is the following: “*The NRC Emotion Lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive).*”¹⁷

Negative emotions: Anger, Fear, Sadness, Disgust; Positive emotions: Anticipation, Joy, Surprise, Trust.

With the proper script, R shows a plot that displays the total number of times words into the tweets are associated with each of these eight emotions.

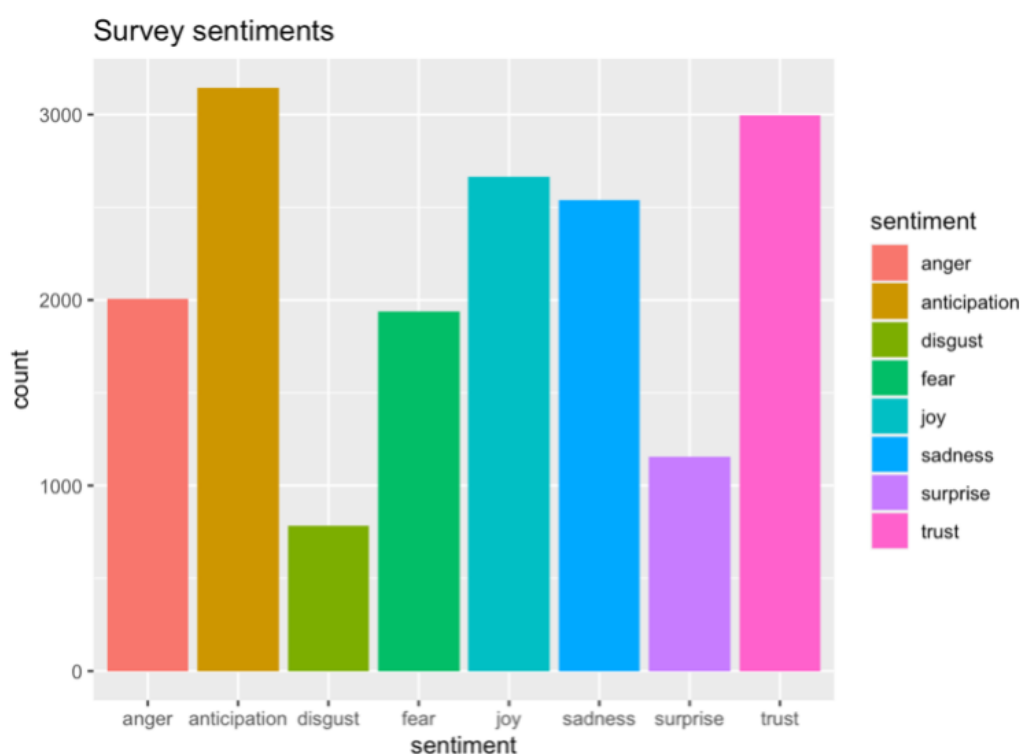


Figure 21: Bar Plot showing the count of words in the text, associated with each emotion (Source: RStudio)

¹⁷ Retrieved from <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

Each day for 2 months, I took notes how into the tweets under examination were present these eight different emotions.

2) ESTIMATING SENTIMENT SCORE

For estimating sentiment score I made use of a very simple algorithm which assigns sentiment score of the text by simply counting the number of occurrences of “positive” and “negative” words in a tweet.

Sentiment Score	
Neutral	Score is 0 if both words either positive or negative does not exist Score is 0 if both words (positive or negative) are equal
Positive Polarity)	Score is positive if total positive words is > total negative words
Negative Polarity	Score is negative if total positive words is < total negative words

Figure 22: Sentiment Score Definition

Hu & Liu published an “Opinion Lexicon” that categorizes approximately 6,800 words as positive or negative, which can be freely downloaded online.

Hence, after charging their positive and negative dictionary in my script, I also added some specific words to those dictionaries, with the following instruction:

```
pos.words = c(positive,'upgrade','congrats','prizes','prize','thnx',
  'great', 'love', 'leader', 'fun', 'like', 'best', 'amazing',
  'awesome', 'bargain', 'beautiful', 'benefit', 'benefits', 'bless', 'incredible',
  'funny', 'champ', 'competitive', 'defeat', 'defeats', 'enjoy',
  'enjoyable', 'entertain', 'excellent', 'exciting', 'excited', 'fortune',
  'lucky', 'luckily', 'gain', 'genial', 'gold', 'god', 'good', 'fine', 'great',
  'experience', 'happy', 'modern', 'pleasant', 'pleasure',
  'improve', 'improved', 'better', 'master', 'outperform', 'passion', 'popular',
  'recommend', 'recommended', 'revolutionary', 'reliable', 'rich', 'satisfy',
  'satisfied', 'satisfying', 'satisfies', 'smart', 'sensational', 'spectacular',
  'speedy', 'strong', 'sublime', 'success', 'superb', 'supreme', 'upgraded',)
neg.words = c(negative,'wtf','wait','waiting','epicfail','no', 'not',
  'anger', 'bad', 'confusion', 'cons', 'disappoint', 'disappointed', 'disaster', 'dislike', 'dissatisfy',
  'dissatisfaction', 'fail', 'fuck', 'fucking', 'frustrated',
  'hate', 'hated', 'hater', 'haters', 'horrified', 'idiot', 'idiots', 'idiocy',
  'impatient', 'overpriced', 'overrated', 'ridiculous', 'sh*t', 'shame', 'shameless',
```

*'stupid', 'terrible', 'threat', 'unacceptable', 'unaccessible',
'unaffordable', 'unavailable', 'unavoidably', 'out of stock',
'unsuccessful', 'waste', 'worst', 'worse', 'wrong')*

Then, after a series of instruction, R computed the sentiment score and printed a barplot like the one shown below.

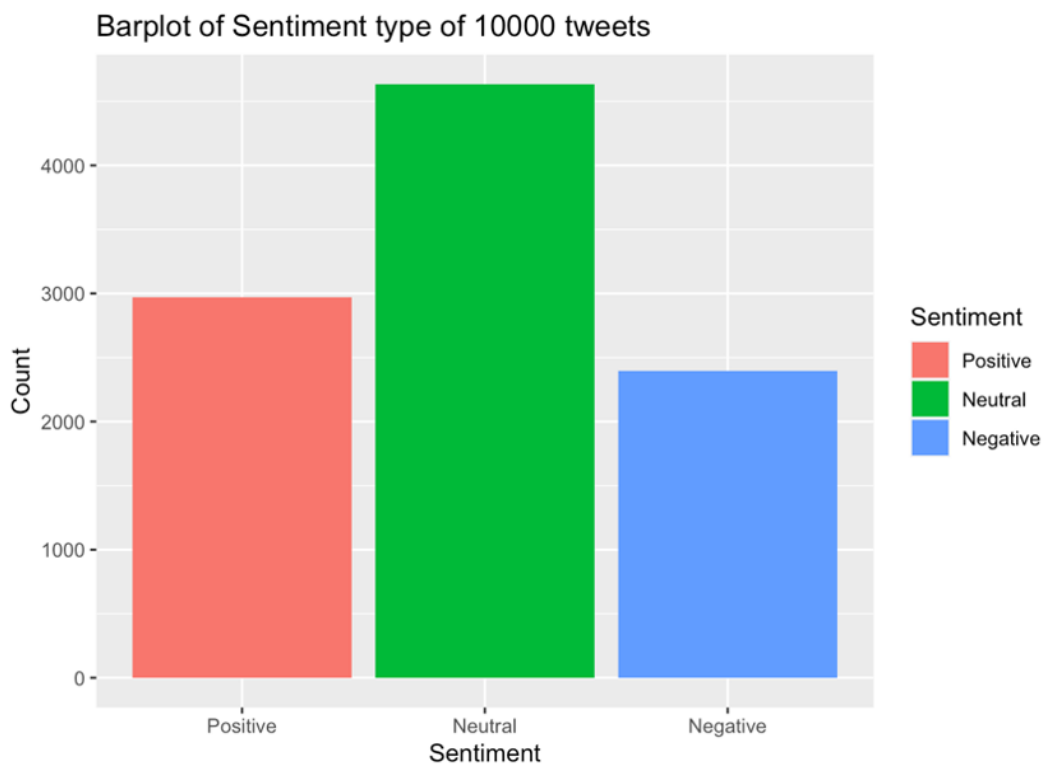


Figure 23: Barplot of sentiment score (Source: RStudio)

As still done for emotion classification, I daily reported sentiment score for both PS5 and Xbox Series. However, I did not limit to report only the frequency of negative or positive tweets, but I also reported the positive/negative valence, obtaining an output, like the one below.

```
> # sentiment score frequency table
> table(analysis$score)
```

Score	Frequency
-4	9
-3	35
-2	224
-1	1250
0	5353
1	2165
2	664
3	295
4	5

Figure 24: R output of sentiment score frequency (Source: RStudio)

3) WORD FREQUENCY

Further, I also took notes daily of the 20 top-used words inside the daily tweets' under-exam.

R built a term-document matrix and sorted words by a decreasing value of frequency. Then, through the instruction *head*, R printed the most frequent word.

This work was interesting in understand whose the main daily topic about PS5 and Xbox Serie are (like a new game/controller or people who complain because it's almost impossible find a next generation console both online or in-store), and at the same time to investigate how many times competitor were mentioned (for example taking Xbox related tweets, how many times “*PlayStation*” or “*Nintendo Switch*” were mentioned?). Into the count I didn't consider the console itself since it's the keyword of each tweet.

4) WORDCLOUD

A wordcloud is one of the most popular ways to visualize and analyze qualitative data. It's an image composed of keywords found within a body of text, where the size of each word indicates its frequency in that body of text.

Below two examples: one related with Xbox Series X and the other one with PlayStation 5. Of course, the name of the console itself have been excluded since it's the keyword and it's present in all the tweets.

For this reason, “PlayStation” cannot be found in the tweets related to PS5.



Figure 25: Wordcloud of tweets related to PlayStation 5 (Source: RStudio)

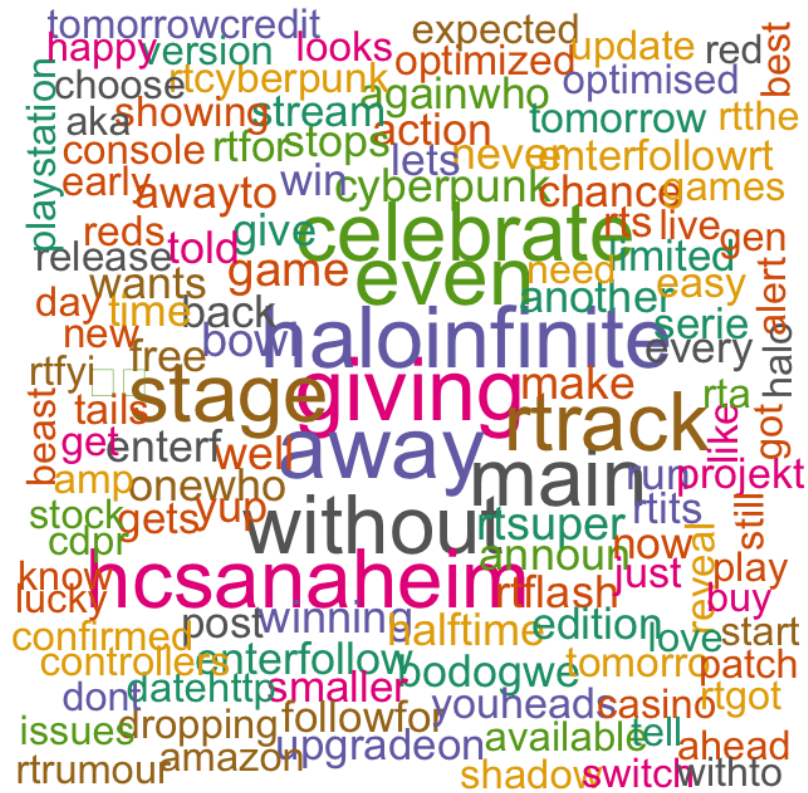


Figure 26: Wordcloud of tweets related to Xbox Series (Source: RStudio)

For the sake of convenience, I didn't report the wordcloud resulting from the output of R for 60 days in a row, since I preferred to focus for practical issues on word frequency.

CHAPTER 4: FINDINGS

In this section, to sum up I displayed the result of my SA¹⁸ which highlights the most valuable insights. All the tables shown in this chapter have been made with Microsoft Excel. Results will be fully report in the Appendix.

A total of 610.000 tweets were examined (10.000 per day) with the keyword “*PlayStation 5*” and “*Xbox Series*”, from the beginning of December 2021 to the end of January 2022. I took notes everyday about the SA results on an Excel sheet.

4.1. Findings about PS5

EMOTION CLASSIFICATION

EMOTION	TIMES TOP EMOTION	DECEMBER	JANUARY	TOTAL
ANGER	5	1641,6	1551,2	1596,4
ANTICIPATION	20	2453,2	3053,1	2753,2
DISGUST	2	1311,5	1221,8	1266,6
FEAR	2	1635,0	1794,2	1714,6
JOY	11	2590,9	2370,2	2480,5
SADNESS	0	1070,1	1616,3	1343,2
SURPRISE	4	2189,2	1787,8	1988,5
TRUST	17	2779,1	3185,4	2982,2

Table 4: Emotion classification PS5

According to the SA, the most frequent emotion in tweets in **Trust** (on average 2982 tweets out of 10000 were classified as containing trust). It is about believe that something is reliable, good, honest, effective, etc.

¹⁸ Sentiment Analysis

Anticipation, however, was the top emotion most of the time (20 out of 61 times). Anticipation is the feeling of excitement about something that is going to happen.

In addition, **Joy** reported good result in term of average presence on tweets (2480) and as the top daily detected emotion (11).

Only 9 times the predominant emotion was a negative one (sadness, fear, anger, disgust) according to R. On the one side, **Anger** was the top emotion 5 times, while on the other one **Sadness** was never the top emotion even if it increased exponentially in absolute value between December and January. On average the less detected emotion was **Disgust**.

SENTIMENT SCORE

	POSITIVE	NEUTRAL	NEGATIVE	POS > NEG	NEG > POS
1-10 DEC	3262	3867	2871	7	3
11-20 DEC	3390	4196	2414	7	3
21-31 DEC	3546	4055	2399	9	1
DECEMBER	3399	4039	2562	23	7
1-10 JAN	3734	4381	1885	9	1
11-20 JAN	2963	4688	2349	8	2
21-31 JAN	3238	4347	2415	10	1
JANUARY	3312	4472	2216	27	4
<i>TOTAL</i>	<i>3355</i>	<i>4256</i>	<i>2389</i>	<i>50</i>	<i>11</i>

Table 5: Sentiment Score Table (PlayStation 5)

On average just under half of the tweets (4256) were classified as neutral, giving the list of positive and negative words (published by Hu & Liu) used for this SA.

The absolute difference between Positive and Negative tweets is 966 (837 in December and 1096 in January). Negative tweets gradually decreased (2872 in the first 10 days against a mean of 2389, and also comparing December (2562) to January (2216).

50 times (out of 61) positive tweets outperformed negative ones (11 times but 6 between the period 1st December – 20th December). In the period between 5-7 December negative tweets were more than positives for three days in a row.

WORD FREQUENCY

Each day I took notes of the top 20 most frequent words in tweets related to PS5. In addition, I added a column which displays how many times the word was in top 3 and thus can be classified as a hot trend.

		TOP 20	TOP 3
Competitor	Nintendo	5	0
	Switch	5	0
	Xbox	45	23
	Series	18	4
	One	26	7
Store (reseller)	Amazon	14	1
	GameStop	7	0
	Walmart	6	0
Important words in relation to the shortage	Available	22	8
	Stock/restock	12	0
	Still	18	0
	Wait/waiting	11	0
	Buy	23	1
	Soldout	4	0
	Got	22	6
	Promo	7	0
Other platform	Twitch	3	1
	Youtube	2	0
-	Consolewinner	6	0

Table 6: Word Frequency of some relevant words (PlayStation 5)

Of course, the main competitor of PlayStation 5 is Xbox (45 times in the top 20 words). Anyway, surprisingly One¹⁹ were mentioned more times than Series. However, “one” can be used also to refer to the number 1, therefore this data must be taken with care. Nintendo and Switch were mentioned only 5 times. During the period in which the analysis was running, I also noticed that One was paired with the verb “got”: “I got one” to refer to the fact that they buy a PS5.

¹⁹ Xbox One is the Microsoft console prior to Series X|S

Looking at the reseller, Amazon is on top (14 times in top 20), GameStop is in the trend 7 times, whereas Walmart 6 times.

Platform like Twitch and YouTube are in the list respectively 3 and 2 times.

The word “consolewinner”, used as hashtag (#) was in the top 20 list 6 times.

most freq neg words	crazy (8), insane (4), clowns (8), wtf (5), nightmare (3 times 5-7 dec), bad (3), damn (1)
most freq pos words	enjoy (2), like (15, 2 times top word), best (9), luck(y), fun (3), happy (11, but most of all during holidays), well (7), good (3)
PS5 related (games or service)	Sony (17), digital (13, 8 times in top 3) edition (12), PS5 games (Spiderman, Fifa (10), Final Fantasy, Ratchet & Clank, Gran Turismo, Uncharted, Fortnite (2)), psplus (7 only in first 20 days), controller (18), psn., gameplay, bundle (18), console (19), rpg
Really used word with no particular value for SA	new, now, covid, game, play, update, store, Christmas, year, month, son, kids, home, gift, man, holiday
VIP	Lebron James, Kyrie Irving, Shaq, Drake, Lewandowski, Cristiano Ronaldo, Lukaku, Mount

Table 7: Relevant topic according to Word Frequency (PS5)

Adjectives with negative meanings are more frequent in PS5 related tweets in comparison with Xbox Series X tweets. We will see it better later. A strong word such as “clowns” to describe Sony was often used, probably by unsatisfied people who are frustrated because they cannot buy a PS5. Even “nightmare” is three times in top 20.

In January, the announcement of Sony about the release of new controllers in fancy colors such as galactic purple, starlight blue and nova pink did a lot of talks. “Controller” is in the top 20 list for 18 times. Sony was into the trend 17 times. In general, PS5 community discusses a lot about games, console itself and platform to play online: PlayStation Plus (7). FIFA was a trend topic after the publication of the team of the year, whereas “Spiderman” (an exclusive PS5 title) was a hot topic especially during the first days after (around 17th December) the SpiderMan movie was in cinemas. Also other successful games like Ratchet & Clank or Uncharted are in the list multiple times.

I will not further dwell on words not valuable for the SA.

Finally, some VIPs were mentioned in the tweets: Some football players like Cristiano Ronaldo, Mount, Lewandowski or Lukaku since if they scored in a match that day, there would be the chance to win a PS5 for a lucky person. Shaq (Shaquille O’Neal, a famous) because he donated to the children of a poor school

1000 PlayStation, madding headlines. Also, the famous rapper Drake and the famous NBA basketball stars Lebron James and Kyrie Irving were in the top-20 tweets one day.

4.2. Findings about Xbox Series X

EMOTION CLASSIFICATION

EMOTION	DAYS TOP EMOTION	DECEMBER	JANUARY	TOTAL
ANGER	2	1059,7	1606,2	1332,9
ANTICIPATION	20	2775,6	2831,1	2803,3
DISGUST	0	746,9	941,9	844,4
FEAR	1	1295,9	1554,0	1424,9
JOY	10	2441,3	2030,5	2235,9
SADNESS	0	1194,8	1603,3	1399,1
SURPRISE	3	1589,1	1541,9	1565,5
TRUST	25	3208,4	3199,5	3203,9

Table 8: Emotion classification Xbox Series X|S

For Xbox Series X|S positive emotions completely outperformed negative ones. 58 times (out of 61) a positive emotion was the top emotion. The three emotions that monopolize tweets are trust (25 times top emotions), anticipation (20) and joy (10).

Trust is longer the most detected emotion: on average about 1/3 of tweets were classified as containing element of trust. Only 3 times (out of 61), trust had a value less than 2000.

Disgust was the least detected emotion, on average 844 times in 10000 tweets, and together with Sadness have never been the most the top emotion. Disgust never exceeded 2000 in 61 days.

Anger (+ 547) and Sadness (+ 408) had a strong increase comparing December to January, whereas Joy (- 411) strongly decreased.

These data, however, will be further discussed in the next section to make a comparison between PlayStation and Xbox findings.

SENTIMENT SCORE

	POSITIVE	NEUTRAL	NEGATIVE	POS > NEG	NEG > POS
1-10 DEC	3469	3999	2532	7	3
11-20 DEC	3246	4212	2542	8	2
21-31 DEC	4166	4492	1342	10	0
DECEMBER	3627	4234	2139	25	5
1-10 JAN	4181	3915	1904	9	1
11-20 JAN	3590	4150	2191	8	2
21-31 JAN	3960	4151	1889	10	1
JANUARY	3910	4072	1995	27	4
<i>TOTAL</i>	<i>3768</i>	<i>4153</i>	<i>2067</i>	<i>52</i>	<i>9</i>

Table 9: Sentiment Score Table (Xbox Series X|S)

For Xbox the number of tweets classified as neutral was on average 4153, considered the list of positive and negative words used for the SA.

The absolute difference between Positive (3710) and Negative (2068) tweets for Xbox is broader: 1701 (1488 in December and 1915 in January). Especially, between 21st December and 10th January, perhaps in conjunction with Christmas holidays, tweets detected as positive have largely exceeded tweets classified as negative.

52 times (out of 61) positive tweets were in greater number than negatives (9 times).

WORD FREQUENCY

		IN LIST	TOP 3
competitor	Nintendo	20	0
	Switch	19	0
	Playstation	40	9
store (reseller)	Amazon	10	0
	Gamestop	6	0
	Walmart	15	1
Words related to the shortage	Available	38	4
	Stock/restock	17	1
	Promo	3	0
	Still	4	0
	Wait/waiting	8	0

	Got	17	5
	Drop	5	0
	Buy	19	0
	Bundle	18	1
Other platform	Twitch	8	0
	Youtube	2	0
	Netflix	1	0
-	One	25	4
-	Consolewinner	22	2

Table 10: Word Frequency of some relevant words (Xbox Series X|S)

PlayStation, as might be expected is the main competitor of Xbox Series X according to the results. However, also Nintendo Switch received a huge number of mentions in Xbox related tweets: 20. For PS5 Nintendo was in top 20 only 6 times.

Looking at the reseller, Walmart here is on top (15 times into the trend), followed by Amazon (10) and GameStop (6).

The word “available”, probably to refer to the fact that somewhere the console was available, is in the top 20 for 38 times (against the 22 time of PS5).

Twitch, a popular streaming platform, is in top 20 for 8 times (whereas for PlayStation just 3).

To conclude the word “consolewinner”, used as hashtag (#) for Xbox Series was in the top 20 list 22 times (and twice in the top 3).

most frequent neg words	crazy (5), insane (4), wtf (2), nightmare (1), bad (1)
most frequent pos words	super (2), well (8 with 4 in top 3), excited 6), happy 8 (but has relative value since it's in the New Year's Eve days), like (20 – 5 times in top 3), lucky (5)
Xbox Series related (games or service)	Microsoft (17), Vrr (5), Kinect (5), Live Gold, games (Gta (Grand Theft Auto), Fortnite (8), Halo (11), Cod (Call of Duty), Fifa, Nba, Wwe, Gears, Hitman), controller (4), console (4)
Really used word with no particular value for SA	new, gen, now, week, month, year, time, online, game, Christmas, win, online, purchase, people, son
VIP	Cristiano Ronaldo, Mason Mount, Shaq, Pringles*

Table 11: Relevant topic according to Word Frequency (Xbox Series X|S)

Positive words far exceeded negative words.

Considering Xbox related words, VRR is in the top for 5 times. VRR (variable refresh rate) is an Xbox Series X technology to get a smooth, artefact-free picture when gaming, ensuring a clean image. Also Kinect (sensor add-on for the Xbox 360 gaming console) is 5 times into the trend. Lot of discussions were done also for the exclusive Xbox title Halo (10) and the famous battle royal videogame Fortnite (11). Microsoft was in the trend for 17 times.

As done for PS5, I will not further discuss about words with no additional value for SA.

Looking At the VIPs, only Shaq (O’Neal), Cristiano Ronaldo and Mason mount were mentioned for the same reason described for PS5. Curiously, Pringles was twice in the top-20 due to a contest in which people could win an Xbox, buying a package of Pringles crisps.

4.3 Discussion about findings

On average negative emotions like Anger (PS5: 1596²⁰ vs Xbox Series X|S: 1332), Fear (1714 vs 1424) and Disgust (1266 vs 844) are widely more present in PS5 related tweets.

Concurrently, the absolute difference between tweets classified as Positive and ones classified as Negative for Xbox is broader: 1701 vs 966 of PS5, definitely a huge difference.

In 61 days negative tweets exceeded positives 11 times for PS5 and 9 times for Xbox Series.

These data are coherent with the usage of strong words like “clowns” or “nightmare” in PS5 tweets. Furthermore, words as “bad”, “wtf”, “crazy” are more used in comparison to Xbox tweets.

These insights put together highlight that the sense of expectation/impatience for the console shortage is more widespread to the PS5 compared to the Xbox Series X. Probably the public opinion is angrier with Sony than with Microsoft, as the number of people who would like to have but cannot buy a PS5 is greater than those who cannot find the Series X.

Even in tweets related to next gen console, worldwide spread hot topics like Covid-19 pandemic, Christmas and New Year greetings (during Christmas holidays) or Spiderman blockbuster movies were really talked. What is certain is that the lack of electronic components which provoked the next generation console shortage has generated discontent. But at the same time, in many people this long waiting is contributing to increase the hype around the consoles. A very frequent topic in tweets both for Series X and PS5 is about (often potential) availability in some resellers store such as Amazon, GameStop, or Walmart. Online users often advise each other how and where to buy a PS5 or an Xbox and are in solidarity with each other

²⁰ Out of 10000

Another interesting insight is given by Nintendo Switch (Nintendo's console) which is often (20 times out of 61) a trend topic for Xbox tweets (only 6 times for PlayStation). It can be considered a direct competitor only for Xbox and especially for Series S model (the economic but less performant alternative to Series X). The hashtag #consolewinner (22 times in the trend topic and 2 times in the top 3) suggests a major support for Xbox community towards the Series X|S console; whereas #consolewinner for PS5 is in top-20 only 6 times. Accordingly, more Xbox tweets (on average 3204 out of 10000) were classified as containing Trust element (trust refers to something that is reliable, good, honest, effective) against a score of 2982 for PS5. Despite all, Sony does a good job in inducing people to discuss about PS5, his games, his accessories, and his service. Lot of talks online were done after PS5 introduced new fancy controllers, FIFA (football simulation game) become a hot topic among gamers after the publication of the team of the year with the best football players, whereas PlayStation Plus was in the top-20 during a period in which the annual subscription was on sale.

Even Xbox players on Twitter speak about games: Halo Infinite (the Xbox flagship game) was several times in the top trend, such as the VRR technology which contributes to the spectacular and clean Series X graphic. Further, contests contribute to create hype towards the console. For example, buying Pringles people had the chance to win a Series X, or still if famous football players like Ronaldo or Lukaku scored, someone had the possibility to win a PS5.

Anyway, Sony with its marketing communication, its announcement, and its exclusive games, does a great job in generating interest and hype towards PS5, probably better than one did by Microsoft for Series X. Sales number (25 million for PS5 vs 12 million for Xbox Series X|S) confirms this claim.

Coherently, also surprise on average were significantly more detected in PS5 (on average on 1988 tweets out of 10000) compared to Xbox (1565).

CONCLUSION

In the first chapter, I addressed the importance of text data and introduced concepts of Text Meaning, Natural Language Processing (NLP) and sentiment analysis (SA).

Text mining is the best way to identify hidden patterns, to uncover relationships, and to make inference. Instead, sentiment analysis is extremely useful in social media monitoring as it allows us to gain an overview of the wider public opinion behind certain topics.

In Chapter 2 I really enjoyed my time to retrace from the beginning the console war developments. I started with the original console war between Nintendo and Sega of the 80s and 90s years, up to the present day with the battle Sony (PlayStation) vs Microsoft (Xbox). Then I compared the last generation of console.

There are so many reasons for what a gamer could prefer one console over the other: an exclusive game, a more convincing marketing communication, a better graphic resolution, a broader memory storage, a most attractive design, a new ultra-modern technology, etc.

But there is a common problem: the shortage due to the lack of microchip provoked by the Covid-19 pandemic that boosted the electronic industry. Nowadays, is difficult for Sony and Microsoft to find such a number of electrical components to be able to satisfy all the console demand. In this situation a 3rd player like Nintendo with his Nintendo Switch Console well managed to take advantage. Is evident that Series S Xbox cannot be put at the same level of PS5 ad Series X model and the only reason for which someone would buy a Series S is the lower price.

Given that premises the question that spontaneously arises is: “Are the console shortage increases desire? Or are consumers losing patience?” In my opinion, although many consumers are a little tired of this situation, their desire to own a next gen console is increasing day by day. I strongly believe that the equation *rare = of value* is true and that humans have the tendency to desire what they cannot easily have.

In the 3rd chapter I introduced the basis on which my sentiment analysis relies on. In the Chapter 4 is highlighted that people are angrier to the shortage of PS5 instead of Xbox Series X. However, the hashtag #consolewinner confirms that Xbox’s owners literally love the console, sustaining that Xbox is better than PlayStation. Despite of this, Sony probably did a better job of Microsoft in generating hype and arguments to talk. Twitter has also proven to be a platform where people fraternize and unite towards a common goal: buy a PS5/Series X. News about a (“literary”) temporary presumed availability, new restocks or tactics to buy them have always been hot topics.

Anyhow, PlayStation 5 vs Xbox Series S|X it’s just a stage of the infinite console war. Technology will make more and more progress; new generations of consoles will be released.

Maybe one day (perhaps not even that far) we ourselves will be teleported inside video games through a sophisticate three-dimensional helmet.

APPENDIX

PS5 Sentiment Analysis Full Result (1st Dec 2021 – 31st Jan 2022)

EMOTION CLASSIFICATION

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
01-dic	2898	2743	1727	1860	1762	1359	2512	2066
02-dic	1357	2077	891	1474	3079	915	3489	2564
03-dic	989	2805	703	1490	3654	822	3036	3599
04-dic	1484	2511	1127	1984	3992	1092	1947	3207
05-dic	1356	2779	1138	2018	2532	706	1306	3692
06-dic	1496	2440	1302	1367	2519	1172	1993	3412
07-dic	1899	3285	1889	1920	4078	801	1749	2789
08-dic	2024	3571	1916	2193	3019	1033	2225	2487
09-dic	2388	1771	3124	2642	2601	1379	2181	2669
10-dic	2617	2025	2280	2655	1678	505	1994	1819
11-dic	1628	1826	706	1136	2811	789	2438	3294
12-dic	2021	2997	890	1348	2388	671	3789	3010
13-dic	1035	2440	1020	906	2004	1013	3251	2929
14-dic	2913	1794	2088	1328	1991	958	1996	2775
15-dic	1269	3063	1923	1888	1759	1042	2520	4541
16-dic	1704	1875	2968	2347	2024	1295	2656	2773
17-dic	728	3782	1178	2003	1623	1806	2309	2886
18-dic	1003	2069	1024	1908	2771	1784	2869	2442
19-dic	886	2476	906	1509	2983	991	2000	1955
20-dic	1096	2507	652	1459	1803	903	2408	3251
21-dic	1512	1864	800	926	2550	1119	1438	2891
22-dic	1437	1352	1294	990	3159	853	1171	2790
23-dic	1984	2780	1001	927	2482	1103	823	2925
24-dic	1300	3669	788	1211	3470	1182	1307	3379
25-dic	1903	1429	589	689	2860	716	1871	1194
26-dic	2866	1272	2351	2647	1683	1610	2790	1504
27-dic	1302	3181	906	2455	2122	1284	1347	2837
28-dic	2317	2009	892	1020	1996	1306	979	2084
29-dic	994	3342	667	1546	2023	979	1684	3582
31-dic	842	1862	604	1203	4310	915	3599	2026

SENTIMENT SCORE

	POSITIVE	NEUTRAL	NEGATIVE			-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
01-dic	3983	3212	2805	10000			4	17	67	304	2413	3212	2514	1309	127	26	7	
02-dic	4149	3611	2240	10000			4	11	43	280	1902	3611	2824	1141	171	8	4	1
03-dic	3114	5046	1840	10000		1	2	12	59	257	1509	5046	2241	676	181	11	5	
04-dic	3721	3410	2869	10000			1	19	102	341	2406	3410	2775	813	120	13		
05-dic	2954	3563	3483	10000		3	6	22	139	504	2809	3563	1996	863	83	9	3	
06-dic	2512	3612	3876	10000			3	31	98	706	3038	3612	1984	472	46	10		
07-dic	2430	4046	3524	10000			5	29	117	612	2761	4046	1812	519	78	17	3	1
08-dic	3171	4411	2418	10000			4	18	96	486	1814	4411	2470	601	55	44	1	
09-dic	3434	3913	2653	10000			1	18	68	552	2014	3913	2114	1137	129	51	3	
10-dic	3147	3848	3005	10000				36	129	402	2438	3848	2289	712	104	42		
11-dic	3414	4096	2490	10000			3	14	99	390	1984	4096	2573	745	49	45	2	
12-dic	2869	4487	2644	10000		1	8	48	181	644	1762	4487	1939	793	94	37	5	1
13-dic	3221	5448	1331	10000				21	73	275	962	5448	2096	931	185	9		
14-dic	3163	4923	1914	10000			2	16	48	325	1523	4923	1889	1133	101	39	1	
15-dic	3838	3994	2168	10000		2	6	37	89	456	1578	3994	2886	762	159	28	3	
16-dic	3075	3712	3213	10000		1		19	92	627	2474	3712	2072	876	106	21		
17-dic	3234	3509	3257	10000			1	44	109	562	2541	3509	2224	905	86	16	3	
18-dic	4062	4173	1765	10000		1	2	8	75	483	1196	4173	2838	1097	71	52	4	
19-dic	2506	4319	3175	10000			5	37	112	568	2453	4319	1723	614	116	47	6	
20-dic	4514	3300	2186	10000			3	18	54	291	1820	3300	3364	953	164	23	9	1

	POSITIVE	NEUTRAL	NEGATIVE			-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
21-dic	4034	3719	2247	10000		2	8	28	71	401	1737	3719	3083	712	194	36	8	1
22-dic	3318	4156	2526	10000			5	11	43	276	2191	4156	2412	872	15	19		
23-dic	2770	4584	2646	10000		2	5	13	51	361	2214	4584	2269	415	72	13	1	
24-dic	3469	3712	2819	10000		3	7	69	82	271	2387	3712	2502	891	65	7	3	1
25-dic	4712	2938	2350	10000		1	2	5	46	303	1993	2938	3882	687	101	33	9	
26-dic	2169	4554	3277	10000			1	9	56	366	2845	4554	1807	305	39	11	7	
27-dic	3207	3879	2914	10000		2		6	50	308	2548	3879	2598	494	104	9	2	
28-dic	3098	4523	2379	10000		1	1	13	65	377	1922	4523	2591	396	94	9	6	2
29-dic	3241	4958	1801	10000		1	1	11	39	280	1469	4958	2729	440	57	15		
30-dic	5092	3796	1112	10000				3	32	156	921	3796	4224	738	63	66	1	
31-dic	5442	3528	1030	10000				4	43	188	795	3528	3525	1644	227	31	12	3
01-gen	5654	3406	940	10000				4	18	130	788	3406	4341	1031	143	133	6	
02-gen	4441	3764	1795	10000			2	6	38	239	1510	3764	3545	753	81	61	1	
03-gen	4438	4543	1019	10000				3	20	115	881	4543	2439	1908	69	17	4	1
04-gen	2967	5386	1647	10000			1	5	58	228	1355	5386	2603	305	36	21	1	1
05-gen	1912	5774	2314	10000				7	78	427	1802	5774	1632	220	47	13		
06-gen	2548	5070	2382	10000			2	3	40	866	1471	5070	1455	963	92	38		
07-gen	4415	3538	2047	10000			4	16	105	390	1532	3538	2954	1332	90	35	4	
08-gen	3976	4445	1579	10000			3	15	58	274	1229	4445	2237	1571	152	12	4	
09-gen	3672	3659	2669	10000		1	2	9	50	441	2166	3659	2373	997	264	31	7	
10-gen	3316	4226	2458	10000			3	23	71	377	1984	4226	1893	665	712	45	1	

	POSITIVE	NEUTRAL	NEGATIVE			-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
11-gen	3304	4568	2128	10000			1	13	80	395	1639	4568	2469	669	158	7	1	
12-gen	3060	4763	2177	10000			2	4	74	261	1836	4763	1491	1471	91	5	2	
13-gen	3946	4339	1715	10000				6	82	235	1392	4339	2608	1259	70	9		
14-gen	2720	4308	2972	10000			2	8	70	492	2400	4308	1912	495	300	12	1	
15-gen	3371	3918	2711	10000		4	5	4	58	360	2280	3918	2504	627	232	8		
16-gen	2214	5432	2354	10000				12	41	506	1795	5432	1687	418	92	17		
17-gen	3384	4424	2192	10000				25	38	241	1888	4424	2074	1101	197	11	1	
18-gen	2420	5146	2434	10000			2	9	57	334	2032	5146	1873	406	131	7	3	
19-gen	2591	4973	2436	10000			2	7	56	379	1992	4973	2097	432	53	9		
20-gen	2622	5005	2373	10000				10	49	344	1970	5005	2047	483	67	20	5	
21-gen	3149	4082	2769	10000			2	8	42	580	2137	4082	2602	466	67	12	1	1
22-gen	3401	3759	2840	10000			1	14	24	492	2309	3759	3011	308	59	19	4	
23-gen	2995	3448	3557	10000		1		11	59	668	2818	3448	1912	934	115	26	6	2
24-gen	3258	3912	2830	10000				7	26	391	2406	3912	2390	813	46	7	2	
25-gen	2991	4864	2145	10000			3	3	40	477	1622	4864	2412	509	56	9	3	2
26-gen	3462	4547	1991	10000				8	63	307	1613	4547	2033	1152	271	5	1	
27-gen	2613	5904	1483	10000				3	34	175	1271	5904	1888	394	173	158		
28-gen	3556	4724	1720	10000				3	15	212	1490	4724	1937	1290	251	74	4	
29-gen	3889	4088	2023	10000			1	2	9	449	1562	4088	2691	973	203	19	3	
30-gen	2669	4101	3230	10000			2	7	33	593	2595	4101	2002	484	148	33	1	1
31-gen	3631	4390	1979	10000			1	6	28	700	1244	4390	1996	1514	92	27	2	

WORD FREQUENCY

[illegible]

Xbox Series X|S Sentiment Analysis Full Result (1st dec 2021 – 31st Jan 2022)

EMOTION CLASSIFICATION

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
01-dic	883	2475	807	928	3122	1029	967	3149
02-dic	1205	2860	1029	1451	3641	1232	1363	2008
03-dic	912	3033	916	844	3221	754	1176	2156
04-dic	607	3394	1246	1288	2778	1517	2018	1799
05-dic	618	4196	781	1792	2809	807	1336	1652
06-dic	1011	2583	552	2623	2426	1475	2021	2278
07-dic	1259	3569	689	1834	2507	1370	1832	2743
08-dic	1599	2611	506	1996	2399	1141	1922	3395
09-dic	2306	2517	930	1257	2727	822	2155	2831
10-dic	1872	2842	1129	1148	3129	1420	2647	2692
11-dic	909	2883	1291	780	2641	1397	3015	2334
12-dic	1214	3353	872	1576	2283	934	2288	3020
13-dic	780	2960	1187	1219	2301	926	1960	3441
14-dic	622	3091	805	886	1856	1184	2026	2839
15-dic	951	4087	654	972	1677	741	1739	3226
16-dic	1334	2760	721	1476	2781	607	1344	2651
17-dic	817	2310	398	1361	2049	1033	1624	3246
18-dic	1322	1646	621	1036	1675	990	811	5079
19-dic	1019	2159	417	821	1941	891	1552	3904
20-dic	1156	1991	349	1427	2379	1223	1420	3153
20-dic	702	1612	490	1164	3013	1509	1580	2457
21-dic	873	1996	512	752	2196	1912	1593	2651
22-dic	967	2456	736	1094	1943	1641	1666	2806
23-dic	1391	3890	691	1100	1777	1708	1790	3043
24-dic	477	1729	305	694	872	1170	645	6881
25-dic	803	2045	717	812	3082	729	1002	3406
26-dic	951	2502	660	1259	3782	1254	1424	3379
27-dic	804	2695	497	1391	2546	1494	1320	4055
28-dic	796	2871	464	1679	2346	1353	962	4338
29-dic	959	3041	488	1868	2190	1280	875	5858
31-dic	1829	3101	1297	1775	1528	1525	1021	2936

		Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
01-gen		940	1654	660	1188	1295	900	863	5976
02-gen		3195	2570	339	2697	993	2650	505	2875
03-gen		2701	2158	638	2171	1274	2472	891	4286
04-gen		1606	2697	843	1262	1828	1719	2282	3613
05-gen		1496	3051	784	1156	2084	1829	1299	2904
06-gen		1405	3671	582	955	1983	1428	1550	3369
07-gen		1391	3891	691	1100	1778	1708	1791	3502
08-gen		1801	2997	743	1264	2103	1847	1320	4043
09-gen		1634	3614	990	1177	1575	1709	1765	4367
10-gen		1161	3716	928	1428	2132	1570	1435	3433
11-gen		1002	2865	773	1119	2720	991	1616	4819
12-gen		1840	2849	1853	2038	1771	1868	785	2471
13-gen		3073	2650	633	1119	2022	1789	1952	1894
14-gen		1792	3212	1299	2307	2184	2287	1312	3171
15-gen		1740	3079	1396	2893	3711	1942	1348	3034
16-gen		1282	3201	1098	2003	2070	1307	1027	2390
17-gen		1204	4660	1287	1602	2635	1417	2157	2166
18-gen		1111	2925	793	1192	1956	1460	854	2176
19-gen		1569	2658	823	1714	3022	1221	1397	2652
20-gen		955	3075	658	1246	2149	1352	1827	3340
21-gen		897	1450	887	629	1635	940	3753	1959
22-gen		1023	1239	932	841	1512	799	2712	2349
23-gen		1350	2227	679	1098	1448	1034	1971	3170
24-gen		1643	3493	1157	1311	1348	1431	811	2215
25-gen		1237	2690	647	1149	2749	1122	957	2364
26-gen		1724	3019	787	1168	2024	957	1245	3299
27-gen		1377	2048	912	1912	2837	1660	1589	4194
28-gen		3022	1450	1967	2911	3746	3179	2903	3644
29-gen		1813	3437	1041	2049	1961	1554	1463	2723
30-gen		1625	3121	1002	1564	1002	1564	1097	2834
31-gen		1182	2398	1377	1942	1398	1996	1321	3952

SENTIMENT SCORE

	POSITIVE	NEUTRAL	NEGATIVE			-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
01-dic	3858	4601	1541	10000			2	8	87	275	1169	4601	2720	998	131	9		
02-dic	3934	4407	1659	10000				6	94	307	1252	4407	3106	785	43			
03-dic	3439	4261	2300	10000			1	4	102	404	1789	4261	2193	1199	36	7	3	1
04-dic	4102	3891	2007	10000				2	37	312	1656	3891	3140	784	177	1		
05-dic	3603	4045	2352	10000			3	14	52	611	1672	4045	2097	1287	209	8	2	
06-dic	4265	3219	2516	10000				7	59	534	1916	3219	3119	970	162	13	1	
07-dic	3643	3859	2498	10000				1	6	91	407	3859	3026	539	72	6		
08-dic	2826	3782	3392	10000		1		3	9	126	707	3782	2306	452	64	3	1	
09-dic	2141	4026	3833	10000			2	11	143	954	2723	4026	1682	391	53	8	5	2
10-dic	2882	3896	3222	10000			5	19	138	562	2498	3896	2358	468	50	6		
11-dic	3151	4551	2298	10000		1		11	90	455	1741	4551	2619	437	89	5	1	
12-dic	3606	3940	2454	10000			2	9	58	709	1676	3940	2420	1054	111	17	3	1
13-dic	2843	4029	3128	10000			4	3	122	511	2488	4029	2126	621	78	14	4	
14-dic	3647	3728	2625	10000				11	85	583	1946	3728	2576	924	138	9		
15-dic	2953	3655	3392	10000			2	19	153	789	2429	3655	1911	878	151	10	3	
16-dic	3348	4042	2610	10000				4	37	352	2217	4042	1822	1393	122	9	2	
17-dic	3650	3952	2398	10000				8	144	419	1827	3952	3026	517	96	11		
18-dic	3268	4364	2368	10000				7	49	513	1799	4364	2473	692	87	10	4	2
19-dic	2510	5251	2239	10000			1	11	87	401	1739	5251	1796	658	52	3	1	
20-dic	3484	4610	1906	10000		1	3	8	102	536	1256	4610	2364	1001	107	10	2	

	POSITIVE	NEUTRAL	NEGATIVE			-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
11-gen	3869	4612	1519	10000		1	2	9	112	268	1127	4612	2107	1636	120	5	1	
12-gen	5238	3586	1176	10000				3	42	163	968	3586	2335	2778	121	4		
13-gen	4807	3176	2017	10000				4	13	149	1851	3176	3046	1721	39		1	
14-gen	3331	3632	3037	10000				2	4	210	274	3632	2096	950	276	8	1	
15-gen	3799	3779	2422	10000				2	3	220	310	3779	2056	887	845	10	1	
16-gen	3364	4705	1931	10000			2	6	128	317	1478	4705	2323	637	395	9		
17-gen	3083	5260	1657	10000				7	54	320	1276	5260	2033	960	72	16	1	1
18-gen	875	4994	4131	10000				3	83	928	3117	4994	762	97	12	2	2	
19-gen	4255	3634	2111	10000			1	5	52	320	1733	3634	3632	492	124	5	2	
20-gen	3975	4121	1904	10000				7	66	402	1429	4121	3179	638	147	8	3	
21-gen	5337	3252	1411	10000			1	5	59	302	1044	3252	3958	1271	103	5		
22-gen	4901	3870	1229	10000				2	31	375	821	3870	4091	737	60	11	2	
23-gen	3530	3901	2569	10000			2	11	68	439	2049	3901	2813	592	119	4	2	
24-gen	3738	3954	2308	10000			3	12	56	413	1824	3954	2987	657	88	4	1	1
25-gen	3376	4096	2528	10000				10	39	325	2154	4096	2406	669	296	5		
26-gen	3937	4764	1299	10000				2	7	51	227	4764	3080	702	146	8	1	
27-gen	4587	4591	822	10000				4	29	136	653	4591	3467	849	264	3	3	1
28-gen	3109	4929	1962	10000		1	1	11	37	370	1542	4929	2456	532	114	6	1	
29-gen	2883	4090	3027	10000		1	3	11	92	681	2239	4090	2176	509	186	8	4	
30-gen	3984	4314	1702	10000				2	9	196	1495	4314	2980	870	95	15	22	2
31-gen	4177	3896	1927	10000			3	7	49	489	1379	3896	3093	623	409	36	16	

	POSITIVE	NEUTRAL	NEGATIVE			-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
11-gen	3869	4612	1519	10000		1	2	9	112	268	1127	4612	2107	1636	120	5	1	
12-gen	5238	3586	1176	10000				3	42	163	968	3586	2335	2778	121	4		
13-gen	4807	3176	2017	10000				4	13	149	1851	3176	3046	1721	39		1	
14-gen	3331	3632	3037	10000				2	4	210	274	3632	2096	950	276	8	1	
15-gen	3799	3779	2422	10000				2	3	220	310	3779	2056	887	845	10	1	
16-gen	3364	4705	1931	10000			2	6	128	317	1478	4705	2323	637	395	9		
17-gen	3083	5260	1657	10000				7	54	320	1276	5260	2033	960	72	16	1	1
18-gen	875	4994	4131	10000				3	83	928	3117	4994	762	97	12	2	2	
19-gen	4255	3634	2111	10000			1	5	52	320	1733	3634	3632	492	124	5	2	
20-gen	3975	4121	1904	10000				7	66	402	1429	4121	3179	638	147	8	3	
21-gen	5337	3252	1411	10000			1	5	59	302	1044	3252	3958	1271	103	5		
22-gen	4901	3870	1229	10000				2	31	375	821	3870	4091	737	60	11	2	
23-gen	3530	3901	2569	10000			2	11	68	439	2049	3901	2813	592	119	4	2	
24-gen	3738	3954	2308	10000			3	12	56	413	1824	3954	2987	657	88	4	1	1
25-gen	3376	4096	2528	10000				10	39	325	2154	4096	2406	669	296	5		
26-gen	3937	4764	1299	10000				2	7	51	227	4764	3080	702	146	8	1	
27-gen	4587	4591	822	10000				4	29	136	653	4591	3467	849	264	3	3	1
28-gen	3109	4929	1962	10000		1	1	11	37	370	1542	4929	2456	532	114	6	1	
29-gen	2883	4090	3027	10000		1	3	11	92	681	2239	4090	2176	509	186	8	4	
30-gen	3984	4314	1702	10000				2	9	196	1495	4314	2980	870	95	15	22	2
31-gen	4177	3896	1927	10000			3	7	49	489	1379	3896	3093	623	409	36	16	

WORD FREQUENCY

<div><div>01-dec</div><div><div>TOP 20 WORDS</div><div>console 1095 game 719 great 625 buy 587 lucky 573 players 571 too 554 new 540 online 525 deceember 525 console 483 bundle 452 month 449 home 438 wtf 435 fortnite 432 hustlefield 430 doubt 417 itmo 417 con 416</div></div></div> <div><div>02-dec</div><div><div>TOP 20 WORDS</div><div>new 1579 game 1479 lucky 728 win 703 buy 587 available 653 win 628 pringles 579 shane 578 chips 573 contest 561 great 486 live 460 gold 442 buy 449 off 418 console 407 release 395 game 385 deceember 370 bundle 370</div></div></div> <div><div>03-dec</div><div><div>TOP 20 WORDS</div><div>play 577 available 958 game 716 win 704 mech 683 game 571 game 552 nintendo 536 hulu 484 well 429 console 421 compatibility 389 win 389 community 389 win 389 hulu 438 rule 341 mms 339 pringles 331</div></div></div> <div><div>04-dec</div><div><div>TOP 20 WORDS</div><div>playstation 1111 well 812 mech 730 like 684 switch 638 game 632 nintendo 622 available 599 midnight 573 advancement 521 win 519 grr 483 digital 462 edition 462 game 436 bust 408 high 406 human 382 cot 375 excited 371 grr 358 switch 425</div></div></div> <div><div>05-dec</div><div><div>TOP 20 WORDS</div><div>new 911 console 916 roms 886 roms 882 cristiano 882 chance 771 win 760 available 700 best 635 win 608 Hund 661 Hope 442 rule 441 game 436 bust 408 high 406 human 382 cot 375 excited 371 grr 358</div></div></div> <div><div>06-dec</div><div><div>TOP 20 WORDS</div><div>update 947 playstation 835 new 774 new 758 available 693 win 613 halo 601 live 580 game 573 switch 536 last 532 win 516 nintendo 487 bundle 464 need 462 gray 458 like 449 launch 446 nintendo 440</div></div></div> <div><div>07-dec</div><div><div>TOP 20 WORDS</div><div>new 1568 need 927 hulu 886 console 762 new 573 infense 527 grr 510 start 509 switch 457 drop 441 playstation 433 edge 433 live 421 Incredible 430 week 408 people 401 game 387</div></div></div> <div><div>08-dec</div><div><div>TOP 20 WORDS</div><div>new 1518 release 977 available 728 people 653 out 618 tired 566 waiting 543 console 527 wtf 486 out 477 world 476 live 451 edge 447 gold 422 game 422 war 418 con 409 card 407 way 406 off 403</div></div></div> <div><div>09-dec</div><div><div>TOP 20 WORDS</div><div>new 1571 gray 908 issue 884 game 745 week 719 release 637 out 636 playstation 611 wtf 569 out 526 announcement 524 game 503 available 487 day 440 partner 439 sports 439 production 438 youtube 420 multiplayer 416 store 414</div></div></div> <div><div>10-dec</div><div><div>TOP 20 WORDS</div><div>game 1542 free 1499 now 1456 game 1354 week 824 gray 721 win 700 launch 571 console 550 promo 518 drop 516 gta 474 time 469 boy 424 release 422 multiplayer 386 playstation 383 money 381 play 370 streaming 360</div></div></div>
<div><div>11-dec</div><div><div>TOP 20 WORDS</div><div>console 1215 new 1093 game 860 nintendo 774 switch 763 players 609 available 513 waiting 449 online 422 market 410 drop 407 live 396 release 357 hulu 348 year 331 like 317 alpha 312 announcement 309 original 295</div></div></div> <div><div>12-dec</div><div><div>TOP 20 WORDS</div><div>new 2034 game 1583 console 1578 game 820 available 743 gameplay 606 win 511 original 510 better 514 trailer 497 drop 407 live 396 release 357 hulu 348 year 331 like 317 alpha 312 announcement 309 original 295</div></div></div> <div><div>13-dec</div><div><div>TOP 20 WORDS</div><div>need 1537 new 1509 game 1583 game 820 available 743 gameplay 606 win 511 original 510 better 514 trailer 497 drop 407 live 396 release 357 hulu 348 year 331 like 317 alpha 312 announcement 309 original 295</div></div></div> <div><div>14-dec</div><div><div>TOP 20 WORDS</div><div>playstation 1006 new 1009 game 826 hulu 838 console 615 better 574 vrr 458 buy 414 like 381 shane 377 hulu 438 human 372 now 370 support 369 game 332 smaller 328 controller 326 best 315 time 294 ago 294 tonight 290</div></div></div> <div><div>15-dec</div><div><div>TOP 20 WORDS</div><div>new 911 new 707 game 682 nintendo 569 game 556 brand 531 buy 480 nintendo 530 super 522 gta 509 new 496 celebrate 460 buy 377 controller 368 mega 332 Incredible 357 launch 355 share 348 best 344 Jordan 391 legend 336 restock 331</div></div></div> <div><div>16-dec</div><div><div>TOP 20 WORDS</div><div>playstation 983 console 855 better 811 game 712 switch 660 available 641 buy 604 nintendo 604 one 586 new 581 game 556 amazon 513 hulu 480 like 480 restock 479 buy 433 online 407 Jordan 391 live 388 Jon 386</div></div></div> <div><div>17-dec</div><div><div>TOP 20 WORDS</div><div>new 1012 like 911 money 772 win 751 mech 648 play 610 win 588 record 559 now 535 win 548 live 480 spiderman 452 michelle 448 brave 430 gta 430 gray 426 like 392 many 390 sightly 365 con 358</div></div></div> <div><div>18-dec</div><div><div>TOP 20 WORDS</div><div>like 1280 new 909 mech 878 spiderman 878 online 662 money 619 win 588 cray 553 one 549 win 548 console 544 game 521 buy 520 walmart 509 console 495 controller 493 game 487 fortnite 485 switch 480 launch 472</div></div></div> <div><div>19-dec</div><div><div>TOP 20 WORDS</div><div>game 1086 playstation 1022 launch 814 game 775 vrr 616 like 605 grr 579 buy 579 one 561 game 517 microsoft 543 today 537 spiderman 534 available 496 nintendo 473 Nathan 426 live 411 gold 406 beat 399 record 386</div></div></div> <div><div>20-gen</div><div><div>TOP 20 WORDS</div><div>well 791 got 784 playstation 707 game 686 today 512 happy 497 game 485 win 462 one 458 jack 450 win 437 purchase 430 illigible 401 console 399 partner 397 vrr 380 gameplay 380 lie 372 microsoft 369 switch 366</div></div></div>
<div><div>21-dec</div><div><div>TOP 20 WORDS</div><div>ps 802 game 750 day 688 playstation 643 wait 623 new 607 hard 598 win 586 new 541 console 527 sold 515 back 449 got 437 spiderman 401 one 378 like 368 christmas 347 present 333 nintendo 332</div></div></div> <div><div>22-dec</div><div><div>TOP 20 WORDS</div><div>day 1006 got 999 game 880 win 680 one 671 playstation 628 update 534 need 501 console 477 edition 429 game 428 promo 400 console 384 new 377 spiderman 334 hulu 328 like 326 christmas 347 present 333 nintendo 332</div></div></div> <div><div>23-dec</div><div><div>TOP 20 WORDS</div><div>chance 729 game 681 playstation 609 win 646 update 606 edition 608 crazy 602 nintendo 593 new 488 walmart 468 microsoft 449 gameplay 436 fortnite 434 infinite 426 available 423 nintendo 409 bundle 406</div></div></div> <div><div>24-dec</div><div><div>TOP 20 WORDS</div><div>christmas 2656 santa 864 game 548 win 540 Ultimate 537 new 530 update 508 island 515 santa 512 new 506 game 428 ps 374 game 395 community 395 thank 391 drawing 392 like 390 live 370 play 358 like 353 edition 319</div></div></div> <div><div>25-dec</div><div><div>TOP 20 WORDS</div><div>christmas 2656 santa 864 game 548 win 540 Ultimate 537 new 530 update 508 island 515 santa 512 new 506 game 428 ps 374 game 395 community 395 thank 391 drawing 392 like 390 live 370 play 358 like 353 edition 319</div></div></div> <div><div>26-dec</div><div><div>TOP 20 WORDS</div><div>christmas 1315 new 830 walmart 772 checker 766 best 712 day 629 restock 620 got 611 games 549 great 539 reason 525 hulu 503 hulu 480 console 477 new 451 thank 391 fun 379 like 362 dad 359 dream 357 high 356</div></div></div> <div><div>27-dec</div><div><div>TOP 20 WORDS</div><div>mind 1024 christmas 779 best 600 console 577 giftcard 547 ps 523 day 507 win 504 hulu 503 holiday 492 source 478 reason 472 game 452 like 429 ultimate 408 playstation 406 one 403</div></div></div> <div><div>28-dec</div><div><div>TOP 20 WORDS</div><div>one 918 christmas 779 well 718 playstation 661 best 651 users 628 dian 625 easily 620 issue 620 christmas 554 hulu 551 ale 526 brand 462 switch 461 nintendo 460 edge 447 like 429 holiday 432 game 426</div></div></div> <div><div>29-dec</div><div><div>TOP 20 WORDS</div><div>game 1418 mount 1371 tonight 1369 sore 1351 mason 1347 win 1107 dante 1081 stock 614 switch 602 nintendo 599 new 488 year 404 brand 402 console 381 live 378 vrr 376 like 375 best 365 crazy 352</div></div></div> <div><div>31-gen</div><div><div>TOP 20 WORDS</div><div>new 1518 year 1659 win 909 win 790 like 676 happy 661 new 6125 great 593 like 563 people 540 jack 539 photos 533 rest 457 original 426 celebrate 419 purchase 393 vrr 391 time 380 play 377 best 373</div></div></div>
<div><div>01-gen</div><div><div>TOP 20 WORDS</div><div>new 4405 year 3794 one 882 got 771 happy 712 year 542 hug 463 great 415 il 404 jack 382 seen 380 photos 375 itmo 364 game 362 stock 349 goodbye 339 playstation 327 good 326 amazon 324 microsoft 317</div></div></div> <div><div>02-gen</div><div><div>TOP 20 WORDS</div><div>new 2336 year 2123 happy 1183 game 810 console 673 fortnite 612 people 593 available 592 edition 588 playstation 569 microsoft 489 away 467 out 459 game 457 stock 422 amazon 420 bundle 388 time 351 itmo 348 buy 335</div></div></div> <div><div>03-gen</div><div><div>TOP 20 WORDS</div><div>new 1827 year 1687 game 991 game 960 just 885 win 741 new 712 switch 679 pass 677 playstation 598 like 509 out 459 game 457 stock 422 amazon 420 bundle 388 time 351 itmo 348 buy 335</div></div></div> <div><div>04-gen</div><div><div>TOP 20 WORDS</div><div>new 1278 playstation 1002 game 901 like 850 console 703 kids 609 game 588 like 591 deal 558 game 549 dic 538 pass 526 market 515 indemend 514 tracking 505 game 376 update 413 nintendo 409 trailer 429</div></div></div> <div><div>05-gen</div><div><div>TOP 20 WORDS</div><div>new 967 game 857 game 830 console 560 play 527 game 510 community 507 just 466 available 464 drop 419 play 411 buy 404 pass 396 bundle 387 first 374 online 373 stock 364</div></div></div> <div><div>06-gen</div><div><div>TOP 20 WORDS</div><div>new 1220 game 1204 new 1045 like 752 time 734 console 702 january 647 available 594 new 552 robby 548 world 543 got 512 slobod 505 steam 504 launching 497 enting 455 like 411 microsoft 411 cod 400</div></div></div> <div><div>07-gen</div><div><div>TOP 20 WORDS</div><div>chance 759 game 681 win 699 console 613 win 513 time 494 new 472 chance 465 digital 404 like 379 money 357 production 355 live 339 play 331 best 320 think 307 able 303 buy 299 away 288</div></div></div> <div><div>08-gen</div><div><div>TOP 20 WORDS</div><div>game 682 win 699 console 613 win 513 time 494 new 472 chance 465 digital 404 like 379 money 357 production 355 live 339 play 331 best 320 think 307 able 303 buy 299 away 288</div></div></div> <div><div>09-gen</div><div><div>TOP 20 WORDS</div><div>game 682 win 699 console 613 win 513 time 494 new 472 chance 465 digital 404 like 379 money 357 production 355 live 339 play 331 best 320 think 307 able 303 buy 299 away 288</div></div></div> <div><div>10-gen</div><div><div>TOP 20 WORDS</div><div>win 2895 new 2218 like 1613 year 993 now 674 game 567 stock 442 console 371 amazon 402 ever 358 wait 343 fortnite 339 switch 363 playstation 336 available 320 enjoy 319 time 302 game 299 rba 292 hulu 287 money 281</div></div></div>
<div><div>11-gen</div><div><div>TOP 20 WORDS</div><div>win 1114 console 801 new 782 giving 780 away 777 game 764 celebrate 647 best 502 point 449 console 455 previous 433 time 393 new 374 million 373 like 364 buy 356 new 341 selling 329 gamestop 321 bundle 318</div></div></div> <div><div>12-gen</div><div><div>TOP 20 WORDS</div><div>new 1287 win 1278 console 690 game 582 away 777 game 764 celebrate 647 best 502 point 449 console 455 previous 433 time 393 new 374 million 373 like 364 buy 356 new 341 selling 329 gamestop 321 bundle 318</div></div></div> <div><div>13-gen</div><div><div>TOP 20 WORDS</div><div>new 1542 win 1448 microsoft 1107 game 968 console 799 officially 1034 nintendo 990 like 771 switch 682 nintendo 629 digital 519 owners 402 fortnite 387 game 382 walmart 332 restock 305 game 303 playstation 295 game 281 gamestop 354 well 279</div></div></div> <div><div>14-gen</div><div><div>TOP 20 WORDS</div><div>microsoft 1565 game 929 production 892 new 888 available 866 game 856 game 793 like 771 switch 640 nintendo 567 win 558 owners 402 fortnite 387 game 382 walmart 332 restock 305 game 303 playstation 295 game 281 gamestop 354 well 279</div></div></div> <div><div>15-gen</div><div><div>TOP 20 WORDS</div><div>one 1169 thru 719 game 751 time 694 like 613 coming 576 buy 561 out 554 win 538 game 412 activision 459 breaking 454 lot 440 new 419 offically 417 theshywhatsaga 402 regatarwars 401 rainbow 394</div></div></div> <div><div>16-gen</div><div><div>TOP 20 WORDS</div><div>like 1254 thru 719 game 681 time 654 like 646 halo 613 edition 616 microsoft 608 like 582 discoun 517 beat 388 nintendo 374 model 357 playstation 352 production 309 got 307 money 299</div></div></div> <div><div>17-gen</div><div><div>TOP 20 WORDS</div><div>game 952 new 638 playstation 583 switch 523 got 515 microsoft 371 game 358 dame 340 nintendo 338 nintendo 338 game 321 play 308 purchase 300 game 300 available 292 game 287 people 286</div></div></div> <div><div>18-gen</div><div><div>TOP 20 WORDS</div><div>got 1105 mind 1024 best 600 win 574 giftcard 547 ps 523 day 507 win 504 hulu 503 checker 492 source 478 reason 472 win 467 game 452 like 429 ultimate 408 ready 391 gamestop 388 switch 382</div></div></div> <div><div>19-gen</div><div><div>TOP 20 WORDS</div><div>win 3021 fortnite 880 like 887 win 870 buy 581 new 565 game 548 gamestop 483 fortnite 476 microsoft 526 gamestop 483 playstation 487 hulu 419 help 415 ready 412 game 410 volley 396 amazon 394 nightmare 394 lucky 388</div></div></div> <div><div>20-gen</div><div><div>TOP 20 WORDS</div><div>game 964 console 701 playstation 595 available 508 switch 480 game 476 week 454 nintendo 424 console 358 purchase 363 game 334 game 334 readers 329 lower 329 reasons 296 earn 284</div></div></div> <div><div>21-gen</div><div><div>TOP 20 WORDS</div><div>new 3790 hulu 3820 year 2799 new 760 like 553 hulu 542 refraction 402 bottli 379 pig 342 peppa 341 hulu 339 upgrade 338 flea 310 microsoft 307 playstation 302 restock 300 planned 294 read 294</div></div></div> <div><div>22-gen</div><div><div>TOP 20 WORDS</div><div>new 1890 year 1219 hulu 1220 legay 907 straight 828 available 517 nintendo 452 win 412 amazon 392 game 392 jio 381 collection 367 beavers 304 edition 350 building 345 game 345 star 317 avenger 304 luna 303</div></div></div> <div><div>23-gen</div><div><div>TOP 20 WORDS</div><div>new 1542 win 1448 microsoft 1107 game 968 console 799 officially 1034 nintendo 990 like 771 switch 682 nintendo 629 digital 519 owners 402 fortnite 387 game 382 walmart 332 restock 305 game 303 playstation 295 game 281 gamestop 354 well 279</div></div></div> <div><div>24-gen</div><div><div>TOP 20 WORDS</div><div>game 1227 available 1113 hulu 1105 playstation 877 nintendo 877 coming 576 buy 561 out 554 win 538 game 412 activision 459 breaking 454 lot 440 new 419 offically 417 theshywhatsaga 402 regatarwars 401 rainbow 394</div></div></div> <div><div>25-gen</div><div><div>TOP 20 WORDS</div><div>one 1169 thru 719 game 751 time 694 like 613 coming 576 buy 561 out 554 win 538 game 412 activision 459 breaking 454 lot 440 new 419 offically 417 theshywhatsaga 402 regatarwars 401 rainbow 394</div></div></div> <div><div>26-gen</div><div><div>TOP 20 WORDS</div><div>one 1851 thru 719 game 681 time 654 like 646 halo 613 edition 616 microsoft 608 like 582 discoun 517 beat 388 nintendo 374 model 357 playstation 352 production 309 got 307 money 299</div></div></div> <div><div>27-gen</div><div><div>TOP 20 WORDS</div><div>new 1212 win 993 available 821 platformer 796 game 789 remember 786 win 784 restock 559 anyone 528 league 527 game 526 celebrate 514 console 453 game 400 game 356 grr 357 nintendo 327 avenger 323 luna 323</div></div></div> <div><div>28-gen</div><div><div>TOP 20 WORDS</div><div>gta 1105 mind 1024 best 600 win 574 giftcard 547 ps 523 day 507 win 504 hulu 503 checker 492 source 478 reason 472 win 467 game 452 like 429 ultimate 408 ready 391 gamestop 388 switch 382</div></div></div> <div><div>29-gen</div><div><div>TOP 20 WORDS</div><div>win 3021 fortnite 880 like 887 win 870 buy 581 new 565 game 548 gamestop 483 fortnite 476 microsoft 526 gamestop 483 playstation 487 hulu 419 help 415 ready 412 game 410 volley 396 amazon 394 nightmare 394 lucky 388</div></div></div> <div><div>30-gen</div><div><div>TOP 20 WORDS</div><div>game 964 console 701 playstation 595 available 508 switch 480 game 476 week 454 nintendo 424 console 358 purchase 363 game 334 game 334 readers 329 lower 329 reasons 296 earn 284</div></div></div> <div><div>31-gen</div><div><div>TOP 20 WORDS</div><div>new 1289 win 1215 chance 1026 giftcard 996 new 835 nintendo 710 one 685 infect 671 game 659 start 542 available 513 one 461 game 452 undrafted 409 playstation 397 chose 385 love 381 collected 384</div></div></div>

R script

```
# 21 PlayStation 5 R script 22
```

```
library('twitteR')
```

```
library('ROAuth')
```

```
library('tidytext')
```

```
library('tm')
```

```
library('wordcloud')
```

```
library('igraph')
```

```
library('glue')
```

```
library('networkD3')
```

```
library('rtweet')
```

```
library('plyr')
```

```
library('stringr')
```

```
library('ggplot2')
```

```
library('ggeasy')
```

```
library('plotly')
```

```
library('dplyr')
```

```
library('hms')
```

```
library('lubridate')
```

```
library('magrittr')
```

```
library('widyr')
```

```
library("readxl")
```

```
library("syuzhet")
```

```
library("lubridate")
```

²¹ # indicates a comment

²² I report only the script for the PS5 as it is the same as the one for the Xbox Series (excepted the tweet search keyword)

```

library("scales")

api_key <- "XXXXXXXXXXXXXXXXXXXXXXXXXXXX" 23
api_secret <- "XXXXXXXXXXXXXXXXXXXXXXXXXXXX"
access_token <- "XXXXXXXXXXXXXXXXXXXXXXXXXXXX"
access_token_secret <- "XXXXXXXXXXXXXXXXXXXXXXXXXXXX"

setup_twitter_oauth(api_key,api_secret,access_token,access_token_secret)

# extracting 10000 tweets related to PS5
tweets <- searchTwitter("playstation 5", n=10000, lang="en") 24
n.tweet <- length(tweets)
print(n.tweet)

# convert tweets to a data frame
tweets.df <- twListToDF(tweets)

tweets.txt <- sapply(tweets, function(t)t$getText())

# Ignore graphical Parameters to avoid input errors
tweets.txt <- str_replace_all(tweets.txt,"[^[:graph:]]", " ")

#####

## pre-processing text:
clean.text = function(x)
{

```

²³ For obvious reason, real key and tokens were omitted

²⁴ For Xbox Analysis, the query research was “xbox series”


```

# convert to lower case

x = tolower(x)

# remove rt

x = gsub("rt", "", x)

# remove at

x = gsub("@\\w+", "", x)

# remove punctuation

x = gsub("[[:punct:]]", "", x)

# remove links http

x = gsub("http\\w+", "", x)

# remove tabs

x = gsub("[ \\t]{2,}", "", x)

# remove blank spaces at the beginning

x = gsub("^ ", "", x)

# remove blank spaces at the end

x = gsub(" $", "", x)

# some other cleaning text

x = gsub('https://', '', x)

x = gsub('http://', '', x)

x = gsub('[^[:graph:]]', ' ', x)

x = gsub('[[:punct:]]', ' ', x)

x = gsub('[[:cntrl:]]', ' ', x)

x = gsub("\\d+", ' ', x)

x = gsub('tweet', ' ', x)

x = gsub('retweet', ' ', x)

x = gsub('enterretweet', ' ', x)

```

```

x = gsub('will', '', x)

x = gsub('dont', '', x)

x = gsub('can', '', x)

x = str_replace_all(x,"[:graph:]", " ")

return(x)

}

#####

cleanText <- clean.text(tweets.txt)

# remove empty results (if any)

idx <- which(cleanText == " ")

cleanText <- cleanText[cleanText != " "]

print(cleanText)

#####

# run nrc sentiment analysis to return data frame with each row classified as one of the following
# emotions, rather than a score:

# anger, anticipation, disgust, fear, joy, sadness, surprise, trust

# It also counts the number of positive and negative emotions found in each row

d<-get_nrc_sentiment(cleanText)

# head(d,10) - to see top 10 lines of the get_nrc_sentiment dataframe

head (d,10)

#transpose

td<-data.frame(t(d))

#The function rowSums computes column sums across rows for each level of a grouping variable.

```

```

td_new <- data.frame(rowSums(td[1:10000]))

#Transformation and cleaning

names(td_new)[1] <- "count"

td_new <- cbind("sentiment" = rownames(td_new), td_new)

rownames(td_new) <- NULL

td_new2<-td_new[1:8,]

#Plot One - count of words associated with each sentiment

quickplot(sentiment, data=td_new2, weight=count, geom="bar", fill=sentiment,
ylab="count")+ggtitle("Survey sentiments")

print(td_new2)


# charge positive and negative list of words

positive = scan("/Users/positions-in-my-pc/positive-words.txt", what = 'character', comment.char = ';')

negative = scan("/Users/ positions-in-my-pc/negative-words.txt", what = 'character', comment.char = ';')

# add your list of words below as you wish if missing in above read lists

pos.words = c(positive,'upgrade','congrats','prizes','prize','thnx',

               'great','love','leader', 'fun', 'like', 'best', 'amazing',

               'awesome', 'bargain','beautiful', 'benefit', 'benefits', 'bless', 'incredible',

               'funny', 'champ', 'competitive', 'defeat', 'defeats', 'enjoy',

               'enjoyable', 'entertain', 'excellent', 'exciting', 'excited', 'fortune',

               'lucky', 'luckily', 'gain', 'genial', , 'god', 'good', 'fine', 'great',

               'experience', 'happy', 'modern', 'pleasant', 'pleasure',

               'improve', 'improved', 'better', 'master', 'outperform', 'passion', 'popular',

               'recommend', 'recommended', 'revolutionary', 'reliable', 'rich', 'satisfy',

               'satisfied', 'satisfying', 'satisfies', 'smart', 'sensational', 'spectacular',

               'speedy', 'strong', 'sublime', 'success', 'superb', 'supreme', 'upgraded')

```

```
neg.words = c(negative,'wtf','wait','waiting','epicfail','no','not',
              'anger', 'bad', 'confusion', 'cons', 'disappoint', 'disappointed','disaster', 'dislike','dissatisfy',
              'dissatisfaction', 'fail', 'fuck', 'fucking','frustrated',
              'hate', 'hated','hater', 'haters', 'horrified', 'idiot', 'idiots', 'idiocy',
              'impatient', 'overpriced', 'overrated', 'ridiculous', 'sh*t', 'shame', 'shameless',
              'stupid', 'terrible', 'threat', 'unacceptable', 'unaccessible',
              'unaffordable','unavailable', 'unavoidably','out of stock',
              'unsuccessful', 'waste', 'worst', 'worse', 'wrong')
```

```
#score sentiment
```

```
score.sentiment = function(sentences, pos.words, neg.words, .progress='none')
```

```
{
```

```
  require(plyr)
```

```
  require(stringr)
```

```
  # we are giving vector of sentences as input.
```

```
  # plyr will handle a list or a vector as an "l" for us
```

```
  # we want a simple array of scores back, so we use "l" + "a" + "ply" = laply:
```

```
  scores = laply(sentences, function(sentence, pos.words, neg.words) {
```

```
    # clean up sentences with R's regex-driven global substitute, gsub() function:
```

```
    sentence = gsub('https://',"",sentence)
```

```
    sentence = gsub('http://',"",sentence)
```

```
    sentence = gsub('[^[:graph:]]', ' ',sentence)
```

```
    sentence = gsub('[[:punct:]]', " ", sentence)
```

```
    sentence = gsub('[[:cntrl:]]', " ", sentence)
```

```

sentence = gsub("\\d+', ", sentence)

sentence = str_replace_all(sentence,"[^[:graph:]]", " ")

# and convert to lower case:

sentence = tolower(sentence)


# split into words. str_split is in the stringr package

word.list = str_split(sentence, "\\s+")

# sometimes a list() is one level of hierarchy too much

words = unlist(word.list)


# compare our words to the dictionaries of positive & negative terms

pos.matches = match(words, pos.words)

neg.matches = match(words, neg.words)


# match() returns the position of the matched term or NA

# we just want a TRUE/FALSE:

pos.matches = !is.na(pos.matches)

neg.matches = !is.na(neg.matches)


# TRUE/FALSE will be treated as 1/0 by sum():

score = sum(pos.matches) - sum(neg.matches)


return(score)

}, pos.words, neg.words, .progress=.progress )

scores.df = data.frame(score=scores, text=sentences)

```

```

return(scores.df)

}

#Calculating the sentiment score

analysis <- score.sentiment(cleanText, pos.words, neg.words)

# sentiment score frequency table

table(analysis$score)

#histogram of sentiment scores

analysis %>%

  ggplot(aes(x=score)) +

  geom_histogram(binwidth = 1, fill = "lightblue")+

  ylab("Frequency") +

  xlab("sentiment score") +

  ggtitle("Distribution of Sentiment scores of the tweets") +

  ggeasy::easy_center_title()

#Barplot of sentiment type

neutral <- length(which(analysis$score == 0))

positive <- length(which(analysis$score > 0))

negative <- length(which(analysis$score < 0))

Sentiment <- c("Positive","Neutral","Negative")

Count <- c(positive,neutral,negative)

print(Count)

output <- data.frame(Sentiment,Count)

output$Sentiment<-factor(output$Sentiment,levels=Sentiment)

```

```

ggplot(output, aes(x=Sentiment,y=Count))+

  geom_bar(stat = "identity", aes(fill = Sentiment))+

  ggtitle("Barplot of Sentiment type of 10000 tweets")


#WordCloud

text_corpus <- Corpus(VectorSource(cleanText))

text_corpus <- tm_map(text_corpus, content_transformer(tolower))

text_corpus <- tm_map(text_corpus, function(x)removeWords(x,stopwords("english")))

text_corpus <- tm_map(text_corpus, removeWords, c("ps5","playstation"))

tdm <- TermDocumentMatrix(text_corpus)

tdm <- as.matrix(tdm)

tdm <- sort(rowSums(tdm), decreasing = TRUE)

tdm <- data.frame(word = names(tdm), freq = tdm)

set.seed(123)

wordcloud(text_corpus, min.freq = 1, max.words = 100, scale = c(2.2,1),

  colors=brewer.pal(8, "Dark2"), random.color = T, random.order = F)


# Build a term-document matrix

TextDoc_dtm <- TermDocumentMatrix(text_corpus)

dtm_m <- as.matrix(TextDoc_dtm)

# Sort by decreasing value of frequency

dtm_v <- sort(rowSums(dtm_m),decreasing=TRUE)

dtm_d <- data.frame(word = names(dtm_v),freq=dtm_v)

# Display the top 20 most frequent words

head(dtm_d, 30)

```

```
#WordFrequencyPlot  
ggplot(tdm[1:50,], aes(x=reorder(word, freq), y=freq)) +  
  geom_bar(stat="identity") +  
  xlab("Terms") +  
  ylab("Count") +  
  coord_flip() +  
  theme(axis.text=element_text(size=7)) +  
  ggtitle('Most common word frequency plot') +  
  ggeasy::easy_center_title()
```


BIBLIOGRAPHY

Amado A., et al. (2017), “*Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis*”

Bhavitha B.; Rodrigues A.P.; Chiplunkar, N.N (2017). “*Comparative study of machine learning techniques in sentimental analysis.* “

Bifet Albert, Fan Wei (2016) “*Mining Big Data: Current Status, and Forecast to the Future*”. SIGKDD Explorations

Coughlin Tom (2018) “*175 Zettabytes By 2025*”. Available at:
<https://www.forbes.com/sites/tomcoughlin/2018/11/27/175-zettabytes-by-2025/?sh=6f25c0154597>

CRAN, <https://CRAN.R-project.org/>

De Moura Edleno (2009), “*Text Indexing Techniques*”. Springer Link

Elia Francesco (2020), “*Sentiment Analysis Dictionaries*”. Available at
<https://www.baeldung.com/cs/sentiment-analysis-dictionaries>

Gies Arthur (2021), “*Choosing the Right Xbox: Series X or Series S*”. New York Times, Available at
<https://www.nytimes.com/wirecutter/reviews/choosing-best-xbox-series-x-vs-series-s/>

Gruet Sam (2021) “*Console shortages: Why can't I buy the Xbox Series X or PlayStation 5?*” Available at
<https://www.bbc.com/news/newsbeat-59476611>

Kabir Amhed, et. al (2018), “*The Power of Social Media Analytics: Text Analytics Based on Sentiment Analysis and Word Clouds on R*”. Informatica Economică vol. 22, no. 1/2018

Maata R., Pineda A., Epoc F., Cordova R. (2021) “*Application of Text Mining to Analyze Customer Opinions on Social Media*”. Global Business and Management Research: An International Journal Vol. 13, No. 3

Maheshwari P. (2021), “*PlayStation 5 v/s Xbox Series X/S - Gaming Marketing Strategies*” Available at
<https://www.mktsoc.club/post/playstation-5-v-s-xbox-series-x-s-gaming-marketing-strategies>

- Mathre Sanil (2020), “Text Mining and Sentiment Analysis: Analysis with R” Available at <https://www.red-gate.com/simple-talk/databases/sql-server/bi-sql-server/text-mining-and-sentiment-analysis-with-r/>
- Migliaccio Valerio (2021), “*PlayStation 5: Lack of Supply or Marketing Strategy?*”
- Modi S., Dommeti S. (2021), “Twitter sentiment analysis in R” Available at <https://jtr13.github.io/cc21/twitter-sentiment-analysis-in-r.html#twitter-sentiment-analysis-in-r>
- Perez Matt (2020), “*PlayStation 5 Rumored Price Tag Is \$500 As Console Pricing Wars Continue*”. Available at <https://www.forbes.com/sites/mattperez/2020/02/14/playstation-5-rumored-price-tag-is-500-as-console-pricing-wars-continue/?sh=47a4f978482c>
- Poh Michael (2011), “*Evolution of Home Video Game Consoles (1967 – 2011)*” Available at: <https://www.hongkiat.com/blog/evolution-of-home-video-game-consoles-1967-2011/>
- Qingjie L., et al. (2020), “*Development and application of massive unstructured big data retrieval technology based on cloud computing platform*”.
- R – <https://www.r-project.org/>
- Rathor A., Agarwal A., Dimri P. (2018), “*Comparative Study of Machine Learning Approaches for Amazo Reviews*”. Procedia Computer Science 132 (2018) 1552–1561
- Saif M. (2021), “NRC Word-Emotion Association Lexicon”. Available at <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- Sony Interactive Entertainment Inc. (2022), “*SIE Business Development | Sony Interactive Entertainment*” Retrieved in January ’22, available at <https://www.sie.com/en/corporate/data.html>
- Turegun Nida (2019), “*Text Mining in Financial Information*”
- Vinay Kumar (2019), “*Sentiment Analysis Techniques for Social Media Data: A Review*”.
- Williams Robert (2020), “PS5, Xbox rollouts drive 80% jump in ad spending for video games” Available at <https://www.marketingdive.com/news/ps5-xbox-rollouts-drive-80-jump-in-ad-spending-for-video-games/591698/>

Yann Niu (2021). “*SWOT Analysis on Sony’s PlayStation Under COVID-19*”. Advances in Social Science, Education and Humanities Research, volume 573

Yse D. L., (2020), “*Your Guide to Natural Language Processing (NLP)*”. Available at:
<https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1>

Zhang L., Wang S., Liu B. (2018) “*Deep learning for sentiment analysis: A survey*”. WIREs Data Min. Knowl. Discov. 2018, 8, e1253.