A quantitative study of the reception of Xbox One versus PlayStation 4 among its users using Twitter data.

Among the users of the two gaming consoles, which group is more satisfied?

Rohit Rokde  
 Luddy School of Informatics, Computing, and Engineering  
 Indiana University, Bloomington, United States of America  
 [rrokde@iu.com](mailto:rrokde@iu.com)

Introduction

Twitter is a social networking service used specifically for voicing one’s opinions. Users of Twitter post their content through text which is called as a “tweet”. Twitter can be accessed using SMS, their official website or mobile applications on various mobile platforms. As a USP (Unique Selling Proposition) of this service, users are only allowed to send a tweet consisting of 140 characters in most languages and 280 in some languages.

“The 140-character limit was originally established to reflect the length of SMS messages, which was how tweets were distributed prior to the development of mobile apps. SMS messages are limited to 160 characters; Twitter reserved the remaining 20 for the username. As often happens in creative mediums, the constraint spurred creativity, and Twitter became a fast-moving, newsy, jokey, weirdo playground.” (sic) [1]

Twitter has had a huge impact on society, especially since few years. It has been used to disseminate news, inform people about the opinions from their friends or public figures. Public figures like politicians, singers, actors, actresses, economists, academics, etc. have been using Twitter to connect to their bases with few clicks on their computing devices. Like any other social networking service, Twitter also has several accounts created by the company representatives or other citizens who seem to represent a company. These other accounts are also called as “Fan clubs”.

Background

Since the 1980s, video games have been popularized in the media and have gained prominence since then. It’s estimated that the console segment’s market share amounts to $45.3 billion, which is 30% of the gaming market share in the gaming industry, and is expected to grow at around 7.3% year on year [2]. Although a major reason for a user to buy these consoles is to play video games, these consoles are equivalents of a full-fledged computer and provide several other services like video streaming services, internet browsers, movie rental services, music services, and many other services found on other computing devices.

As of today, the two major vendors of the most popular TV gaming consoles, Microsoft and Sony, have sold more number of units than any other company since their release in 2013. Apart from that, these two gaming consoles are competing products because of their similarities (on a similar price point), and there has been a history of such competition between the two companies selling these products.

The purpose of this research is to get a sense of people’s opinion on how they view these products and possibly deduce their feelings about these products using data obtained from Twitter.

Methods and Data

An open source python library called as GetOldTweets3 was used to collect the Twitter data for this research. Two search strings were finalized to collect data belonging to the two groups, viz “Xbox One” and “PlayStation 4”. Only the tweets containing these search-strings were thus included in our dataset. Additionally, the data collection is done in such a way so as 1000 tweets were collected every by setting the setUntil() function call. Here we have chosen 15th of every month starting from December 2013 to 2019 (discarding data from other dates).

Following data fields were recorded for the purpose of this research:

1. Date - String: The date field includes year, month, day and time to nearest second when the tweet was recorded.
2. Username - String: This is the username used for identifying the account holder.
3. To - String: The username of the account holder to whom the tweet was sent.
4. Replies - Integer: The number of tweets sent as replies that tweet.
5. Retweets - Integer: The number of times the tweet was shared by random account holders.
6. Favorites - Integer: The number of times Favorites button was clicked on that tweet.
7. Text - String: The actual content of the tweet.
8. Geo - Float: The latitude and longitude of the location from where the tweet was sent.
9. Mentions - String: The username(s) of other account holder(s) mentioned inside the tweet.
10. Hashtags - String: Usage of text preceded with a hash character (#)
11. Id - Integer: Unique integer to keep track of a tweet. Every tweet ever generated will have a unique number associated with it.
12. Permalink - String: A URL to access the tweet directly.

Let’s take a look at the sample distribution of our unit of measurement by categories:

|  |  |  |
| --- | --- | --- |
| Total tweets | | |
| Xbox One | | PlayStation 4 |
| 73000 | 57071 | |

As can be seen from the above data, we have slightly skewed data amongst the two categories.

Let us look at engagement metrics viz. number of replies (per thousand), number of retweets (per thousand) and number of favorites (per thousand) on Fig. 1. All the graphs are placed next to each other so that it is easier to see the contrast.

The reason why we have written “per thousand” is because the numbers are cumulative across the y coordinates. The x coordinate data is the same for all the graphs in which each x coordinate represents all the top 1000 tweets until the 15th of a specific month of the given year. Each year has 12 months of recorded data except for 2013, in which only the data from the month of December is recorded. This was done because both the consoles were released on November 2013. So, it was a prudent measure to start collecting tweets from December 2013 till December 2019.

There is a clear trend that can be seen by comparing the engagement metrics in Fig.1. More on this, in the discussion section.

Next, we are going to look at the word counts from the textual data collected from both the categories. We have tokenized the textual data (tweets), converted all the words into lower case, removed common english stop words from it, and removed commonly occuring words in the tweets (e.g. http, www, games, ps4 etc.)

The word count is given on Fig 2 and Fig 3.

This part is left blank intentionally to accommodate the figure on a single page.

Figure 1

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

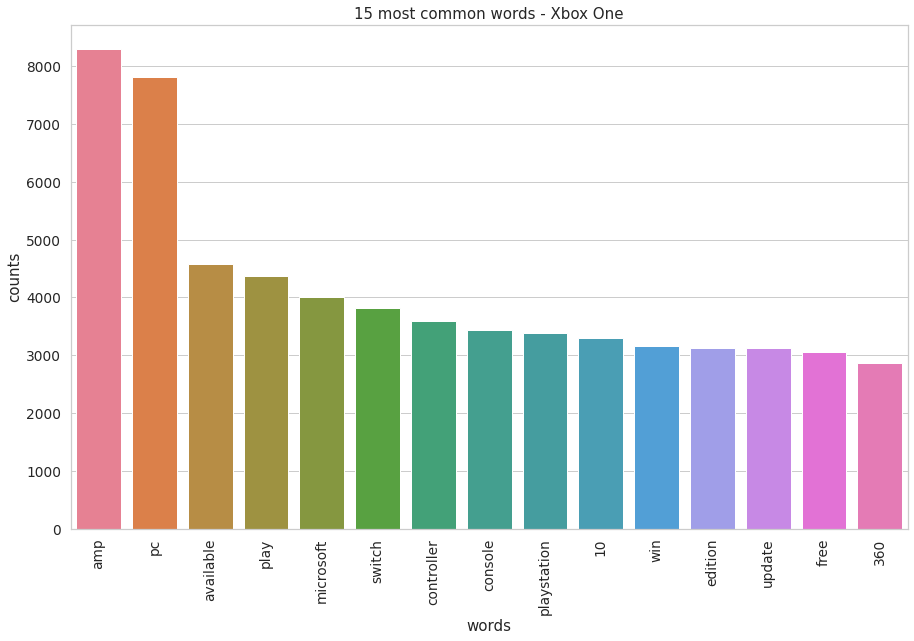
Description automatically generated

A screenshot of a cell phone

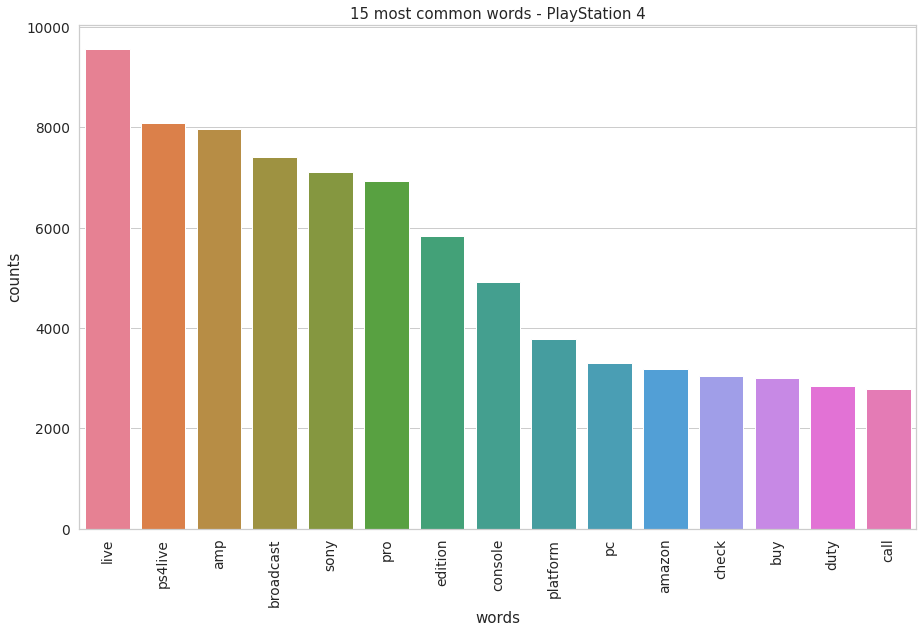
Description automatically generated

A screenshot of a cell phone

Description automatically generated

****

**Figure 2**

****

**Figure 3**

**Discussion**

Although the engagement metrics can be deceiving in the sense that the conclusions formed out of it can be sometimes challenged fundamentally by looking at other research related to basic Psychology, we can still deduce some conclusions with some degree of certainty. Even though the number of tweets related to PlayStation 4 is about 2/3rd the number of tweets related to Xbox One, the engagement metrics show patterns that are sufficiently different from each other.

Few highlights observed from the data:

1. At the end of 2019, the maximum number of replies related to Xbox One is around 3 times that of PlayStation 4. (Xbox - ~78K/thousand, PS4 - ~18K/thousand)
2. Overall, there is huge variation in the number of tweets related to PlayStation 4 whereas the number tweets related to Xbox One have been quite stable.
3. The engagement metrics have been on the rise for both types of tweets since 2013. This finding correlates to the trend in the number of units sold since 2013 [4].
4. On an average, the engagement metrics across the three categories (number of replies, retweets and favorites) show that there is a disproportionate response to the tweets related to Xbox One as compared to the tweets related to the PlayStation 4. It should be noted that the number of PlayStation 4 units sold (~109 million) is more than twice that of Xbox One units (~47 million) [5].
5. The fact that overwhelming amount of data was generated on Xbox One as compared to the PlayStation 4 can be attributed to the fact, perhaps users of Xbox are happier or excited about using the Xbox One. Because the number of consoles sold in North America in case of both types of consoles is almost the same (30 million units) [6] [7], our conclusion might hold water.

We can also deduce certain information from textual data. For example, the use of word “amp” which is short hand for Amplifier could indicate that the user is enjoying his/her console. This should be a reasonable assumption because only that person would choose to invest in subsidiary technologies (like an amplifier) who is happy with the product and chooses to use it. We can clearly see that the word “amp” is used more by the users of Xbox One.

**Conclusion**

Since there was virtually no difference in the data collection strategy for both categories viz. “Xbox One” and “PlayStation 4”, the over whelming number of tweets obtained belonged to the category of Xbox One. Add to the fact that the overwhelming number of replies, retweets and favorites belonged to the category of Xbox One. Just by looking at these two observations we can say that people are talking about Xbox One more than PlayStation 4. We can say with certain degree of certainty that, at least in North America, people are probably more satisfied with Xbox One as compared to PlayStation 4. Although a sentiment analysis could have shed some more light and could perhaps strengthened/weakened the conclusion.

**REFERENCES**

[1] <https://www.theverge.com/2017/9/26/16363912/twitter-character-limit-increase-280-test>

[2] <https://newzoo.com/key-numbers>

[3] <https://pypi.org/project/GetOldTweets3>

[4] <https://www.statista.com/statistics/651576/global-ps4-console-unit-sales/>

[5] <https://en.wikipedia.org/wiki/List_of_best-selling_game_consoles>

[6] <https://www.statista.com/statistics/1005418/xbox-one-console-unit-sales-country/>

[7] <https://en.wikipedia.org/wiki/PlayStation_4>

Conference Name:ACM Woodstock conference

Conference Short Name:WOODSTOCK’18

Conference Location:El Paso, Texas USA

ISBN:978-1-4503-0000-0/18/06

Year:2018

Date:June

Copyright Year:2018

Copyright Statement:rightsretained

DOI:10.1145/1234567890

RRH: F. Surname et al.

Price:$15.00