

## Analysis of Users' Perception of Competing Video Gaming Consoles from Twitter Data

**SONY**

PlayStation®5

vs



Microsoft



XBOX

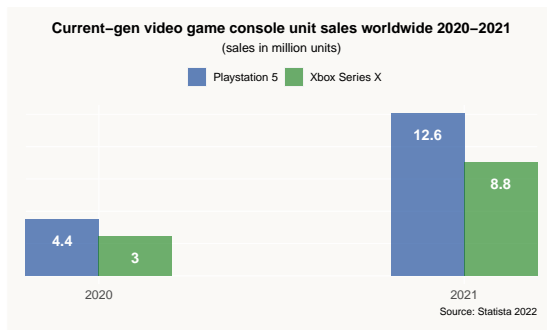
Group 5 - 10778940|10790672|10787306|10799756

07 December, 2022

# Introduction

## Why 'PlayStation 5' vs 'Xbox Series X'?

- ▶ Microsoft and SONY are two of the biggest players in the non hand-held console gaming market with their feature products 'Xbox' & 'PlayStation' being 2 of the most purchased gaming consoles globally in the last 2 decades.(Statista (2022)).
- ▶ In November 2020, Microsoft released the "Xbox Series X"(Xbox) to rival SONY's "Playstation 5"(Ps5) released in same month.(The Economist (2020)).



## Who needs to understand User Perception of these Products?

**International Retailers** (Amazon, Walmart etc.), **Video Game Developers** (Activision, Gameloft etc.), **Microsoft** and **SONY**.

# Data Gathering

- ▶ Data sourced using the 'rtweet' package in R. (Michael W. Kearney and Lluís Revilla Sancho and Hadley Wickham and Andrew Heiss and Francois Briatte and Jonathan Sidi (2022)).
- ▶ **16470** tweets downloaded for Xbox Series X
- ▶ **34567** tweets downloaded for PlayStation 5
- ▶ Data considered tweets within **6-9 days** leading up to **November 24th 2022**. (Michael W. Kearney and Lluís Revilla Sancho and Hadley Wickham and Andrew Heiss and Francois Briatte and Jonathan Sidi (2022)).
- ▶ English speaking markets were the focus of analysis
- ▶ We did a Preliminary Study to understand how users mention these products in their tweets and used this to structure our keyword search which included multiple keywords.

## *Xbox Series X* search keywords:

"xbox series x", "xboxseriesx" or #xboxseriesx

## *Playstation 5* search keywords:

"ps5", "playstation 5", "playstation5", #playstation5 or #ps5

# Data Cleaning and Transformation

- The Location related data was analysed to measure the extent of reliability.

Table 1: Analysis of location related variables in downloaded tweet data

	Tweet Count	% of Total Tweets
Tweets with 'country' data	561	1.10%
Tweets with 'latitude' data	0	0.00%
Tweets with 'longitude' data	0	0.00%
Tweets with 'location' data	30348	59.00%

*Note:*

This table considers a total of 51,037 tweets downloaded for both products in the analysis

Table 2: Tweets by unique locations showing inconsistent formatting and strange location values

Location	Tweet Count
NA	20689
Pittsburgh, PA	1964
United States	1860
United Kingdom	352
Worldwide	336

- The readable Location values were mapped to the associated Countries.
- The mapped Locations values account for approximately 30% of the available Location data.

# Product Popularity Analysis

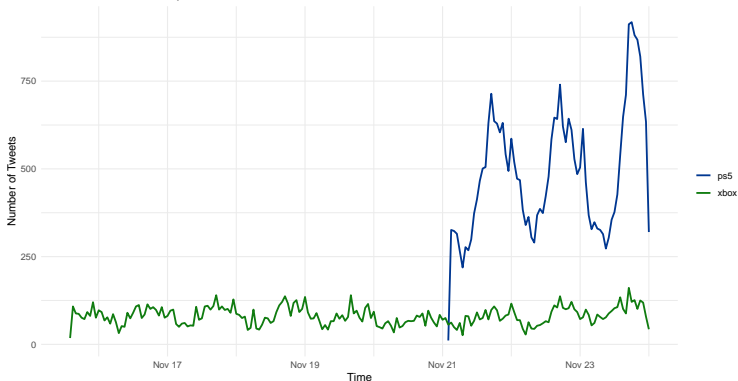
*Is the 'PlayStation 5' more tweeted about than the 'Xbox Series X' from our data?*

Product popularity on Twitter can be quantified by how much a Product is talked about. On Twitter we can measure this in Tweet Frequency.

## Frequency of Tweets related to 'Xbox Series X' and 'PlayStation 5'

Tweet counts aggregated using hourly intervals

34000 tweets related to PlayStation 5 downloaded in much less time than 16000 tweets on Xbox Series X



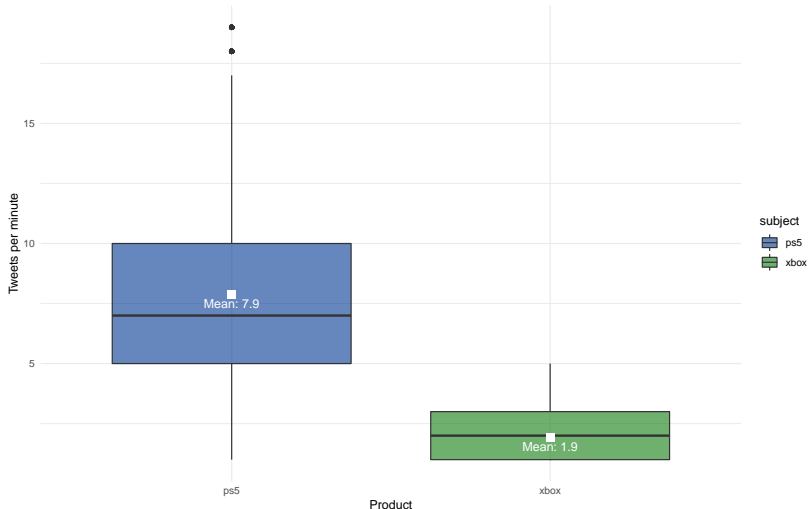
Source: Data collected from Twitter's REST API via rtweet

# Product Popularity Analysis

*Is the 'PlayStation 5' more tweeted about than the 'Xbox Series X' from our data?*

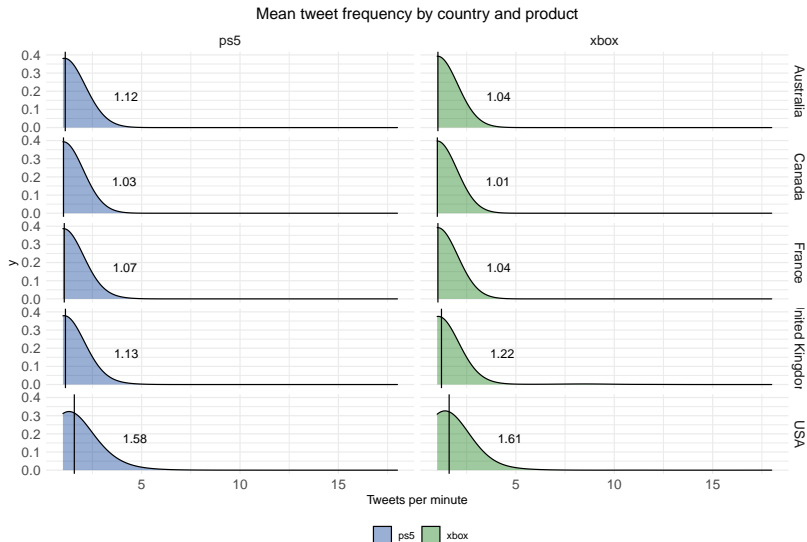
Boxplot of Tweet frequency in Tweets per minute

Tweet frequency for both products shows normal distribution. Mean and median are similar



# Product Popularity Analysis

*How popular is the 'Playstation 5' vs 'Xbox Series X' in major Markets?*



Source: Data collected from Twitter's REST API via rtweet

# Word Analysis and Text Mining

### Wordcloud for Xbox related Tweets



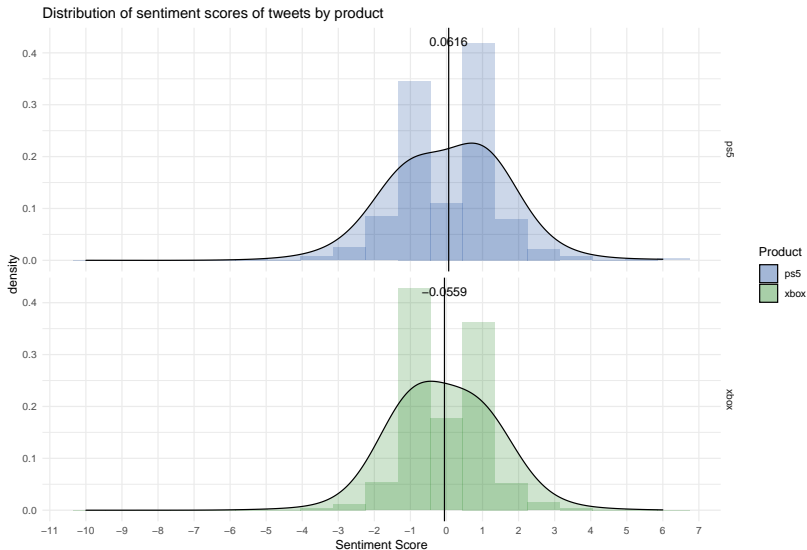
Wordcloud for Ps5 related Tweets





# Sentiment Analysis

*Is the Twitter User Sentiment better toward the 'PlayStation 5' than the 'Xbox Series X' from our data set?*



Source: Data collected from Twitter's REST API via rtweet

# Sentiment Analysis

*Is the Twitter User Sentiment better toward the 'PlayStation 5' than the 'Xbox Series X' for the wider population of twitter users?*

- ▶ 10,000 Randomly sampled Tweets for each Product
- ▶ Unequal Variances of Sentiment Scores for Tweets of both Products
- ▶ Normally distributed Sentiment Score Values for Tweets on both Products

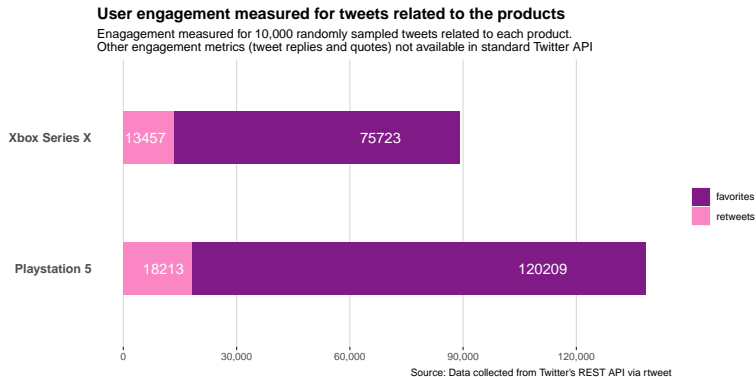
**The Welch Test will help answer our question**

Table 3: Results of statistical test of the sentiment means of both products (Null Hypothesis rejected)

Test Description	H0 (Null Hyp)	H1 (Alt. Hyp)	p-value
Two-tailed Welch's test	mPs5 = mXbox	mPs5 != mXbox	<b>1.57352e-05</b>
One-tailed Welch's test	mPs5 < mXbox	mPs5 > mXbox	<b>7.867598e-06</b>

# User engagement and Influencer Analysis

*Are Twitter Users engaging more with 'Xbox Series X' or 'Playstation 5 Tweets'?*



# User engagement and Influencer Analysis

## Who is driving the engagements with these products on Twitter?

- We developed an algorithm to rank the top twitter accounts (*Influencers*) driving user engagement around these products which could be of marketing interest.

Table 4: Influencer Rank for tweets related to Xbox Series X

User	Rank	Followers	Tweet count	Mean Retweets per tweet	Mean favorite per tweet	Mean Sentiment Score
Atlas_West	<b>2907</b>	421736	3	281	2136	0.33
Wario64	<b>2903</b>	1155165	3	29	319	0.33
JamieMoranUK	<b>2897</b>	23558	26	6	94	0.54
IGNDeals	<b>2893</b>	287215	3	17	178	1.00
gundamevolution	<b>2891</b>	35672	3	253	1257	0.67

*Note:*

This table considers 10,000 randomly sampled tweets

Table 5: Influencer Rank for tweets related to PlayStation 5

User	Rank	Followers	Tweet count	Mean Retweets per tweet	Mean favorite per tweet	Mean Sentiment Score
NextGenPlayer	<b>9639</b>	66869	6	114	899	0.17
GermanStrands	<b>9630</b>	36592	10	81	956	0.70
videogamedeals	<b>9599</b>	475742	10	7	42	0.30
LimitedRunGames	<b>9568</b>	237760	6	6	31	0.17
TCMF2	<b>9550</b>	11917	4	65	558	0.50

*Note:*

This table considers 10,000 randomly sampled tweets

# Limitations of the analysis

## Twitter API Limitations

- ▶ Variables such as reply\_count and quote\_count were not available to us to improve our analysis of consumer engagement and reach (Twitter Developer Platform (2020)).
- ▶ Twitter location data is not reliable.

## Limitations of the Statistical tests

- ▶ Social media data is not representative of the consumer population.
- ▶ Sentiment score can be skewed by singular events/incidents e.g a likely hacking incident related to Xbox Series X in our case.

## Limitations of Lexicon-based approach

- ▶ Studies show that the lexicon based approach not as accurate as supervised-learning approach. (Kharde and S.Sonawane (2016)).
- ▶ lexicon-based methods are restricted by their lexicons, and more particularly, by the use of static prior sentiment values of terms regardless of their contexts. (Kharde and S.Sonawane (2016)).
- ▶ In the context of twitter where non-traditional language is largely used with different colloquialisms generated each day, it's difficult for lexicon based approaches to be manually updated to account for new words. (Kharde and S.Sonawane (2016)).
- ▶ Algorithm does not cover for context of speech, sarcasm and other nuances of speech based communication.

# Key Insights & Strategic Suggestions

- ▶ Considering the PlayStation 5 has more traction on twitter than the Xbox Series X, **International Retailers** could strongly consider twitter as a major marketing channel for PlayStation 5.
- ▶ **International Retailers** can leverage on influencers for each product as part of their marketing strategy to improve the engagement and sentiment of users around that product.
- ▶ **Video Game Developers** can focus their resources on developing more games for the PlayStation 5 versus the Xbox Series X.
- ▶ **Microsoft** could look into a possible hacking related incident revealed from the Word Analysis for Xbox Series X related tweets.
- ▶ Also from the Word Analysis, **SONY** could look into PlayStation 4(Ps4) related tweets. Perhaps there is a preference in terms of features in the Ps4 console.

## Other Reference Notes

- ▶ Packages: Tidytext (Gabriela De Queiroz and Colin Fay and Emil Hvitfeldt and Os Keyes and Kanishka Misra and Tim Mastny and Jeff Erickson and David Robinson and Julia Silge (2022)); kableExtra (Zhu (2020)).
- ▶ Script Guide: R markdown guide I (Xie, Dervieux, and Riederer (2022)); R markdown guide II (Xie, Allaire, and Grolemund (2022)).

# References

- Gabriela De Queiroz and Colin Fay and Emil Hvitfeldt and Os Keyes and Kanishka Misra and Tim Mastny and Jeff Erickson and David Robinson and Julia Silge. 2022. *Package “Tidytext”*. Vienna, Austria: The R Foundation.
- Kharde, Vishal A., and S. S.Sonawane. 2016. “Sentiment Analysis of Twitter Data:a Survey of Techniques.” *International Journal of Computer Applications* 139 (11): 1–15.
- Michael W. Kearney and Lluís Revilla Sancho and Hadley Wickham and Andrew Heiss and Francois Briatte and Jonathan Sidi. 2022. *Package “Rtweet”*. Vienna, Austria: The R Foundation.
- Statista. 2022. “Current Generation Video Game Console Unit Sales Worldwide from 2017 to 2021.” <https://www.statista.com/statistics/276768/global-unit-sales-of-video-game-consoles/>.
- The Economist. 2020. “PlayStation 5 v Xbox Series x.” <https://www.economist.com/business/2020/11/07/playstation-5-v-xbox-series-x>.
- Twitter Developer Platform. 2020. “Twitter API V2 Data Dictionary.” <https://developer.twitter.com/en/docs/twitter-api/data-dictionary/introduction>.
- Xie, Yihui, J. J. Allaire, and Garrett Grolemund. 2022. *R Markdown: The Definitive Guide*. Boca Raton, Florida: Chapman & Hall/CRC.
- Xie, Yihui, Christophe Dervieux, and Emily Riederer. 2022. *R Markdown Cookbook*. Boca Raton, Florida: Chapman & Hall/CRC.
- Zhu, Hao. 2020. *Create Awesome LaTeX Table with Knitr::kable and kableExtra*. Vienna, Austria: R Foundation.