SPONSORED BY: MICROSOFT MINECRAF

Introduction

Goal One:

First, we were asked to expand upon the social media listener project that started the previous year.

These listeners are valuable tools for our partners at Microsoft because they allow us to gather and analyze large amounts of Minecraft-related social media data. This meant:

- -Improving upon our already established Reddit and Twitter listeners.
- -Developing entirely new listeners for YouTube and Instagram.

Goal Two:

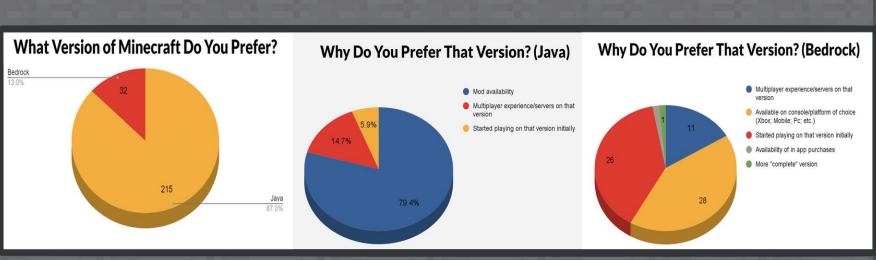
Our second goal was to delve into the success of Minecraft. We set out to understand the reasons behind its success. We approached this objective in a variety of ways,

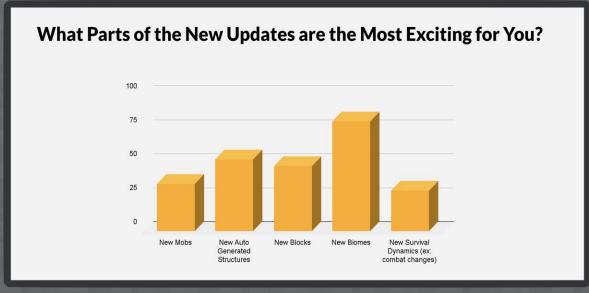
- including:
 - -Personally conducted surveys.
 - -Analyzing social media data.
 - -Using Twitter data for topic modeling.

Survey Results

The survey we designed was sent out to various Discord and Reddit communities. Approximately 250 people completed the survey.

The motivation for conducting a survey was to go directly to the player base and ask them directly for their opinions about Minecraft.





New Data Sources



TikTok



We wanted to collect data from TikTok, but TikTok has a unique way of posting content and can be challenging to analyze. Ultimately we had to move to Youtube and Instagram because:

- TikTok's userbase is too broad.
- Tiktok's tagging system.
- Users value views over community.
- -TikTok's terms of service.



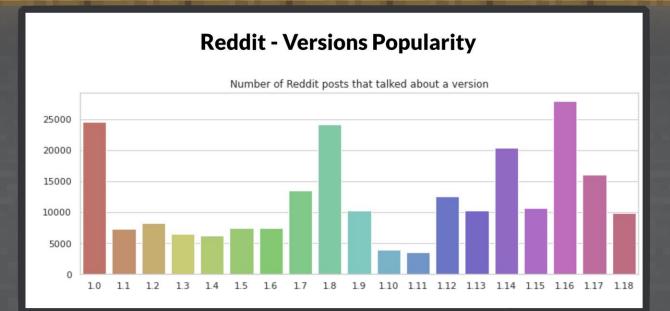
Instagram

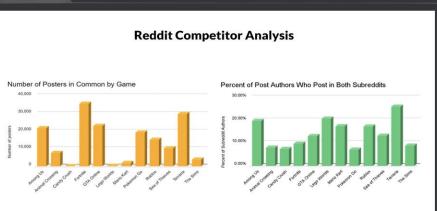
Reddit Data

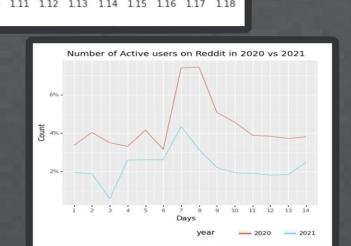
Building on the data we collected from Reddit last year, we looked at frequently discussed topics.

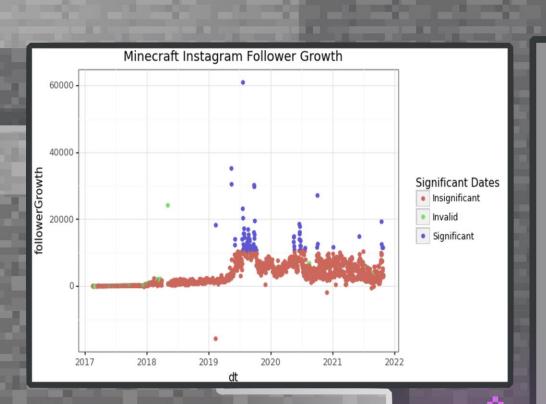
First, We asked ourselves what version of the game people preferred. This can help us gather insight on what kinds of content people are most receptive to.

We also looked at the active users in 2020 vs. 2021 to understand how outside factors and time impact how players talk about the game.









YouTube

For Instagram, we collected data from Minecraft and their competitors. The data includes: -images.

- likes.
- -comments.
- -and captions.

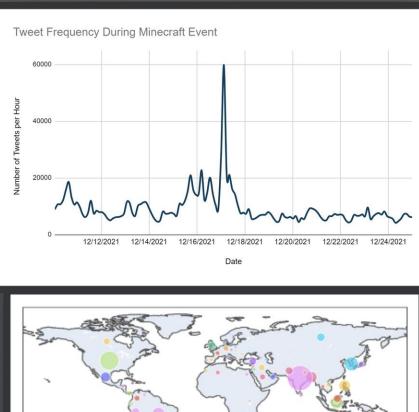
We noticed higher engagement during: -Minecraft Live.

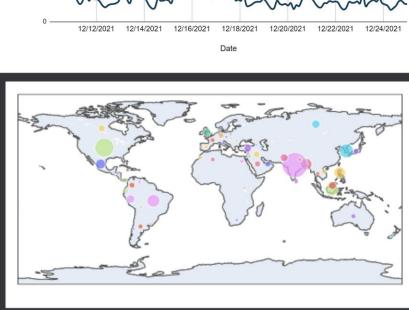
- -game updates. -Minecraft Earth.
- -announcements.

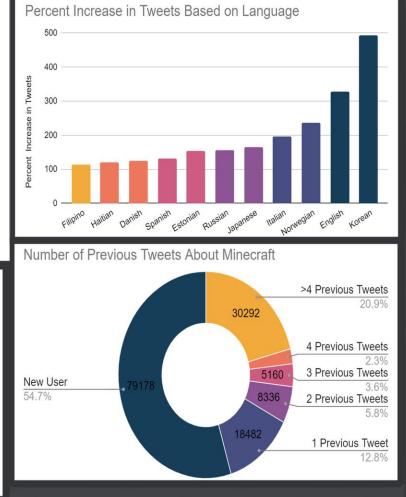
YouTube Events

Minecraft is huge for content creators and hosts many different events. The following graphs focused on tweets and reactions to Minecraft's Trillion Views event and the BTS concert.

Overall we concluded that these events spark responses from different countries and locations and that events on one platform can affect others.







Our YouTube Listener:

The script runs daily to monitor trending, influential, and official content related to Minecraft and its competitors

We collect data regarding channels and videos. This data can range from metadata (e.g., title, description, tags) to stats (e.g., number of views, likes, comments) to more niche information (e.g. how YouTube classifies a video)

This listener was built with YouTube's algorithms in mind to help reduce noise and increase our reach.

Conclusion

Topic Analysis

For topic modeling, we wanted to make sure we could run code fast, so we wrote a custom implementation of the GSDMM clustering algorithm. This helps us go more quickly by using C libraries, and it is 31x faster than any other implementation.

We use it semantically to cluster Tweets, YouTube videos, Twitch stream titles, etc. because: -It provides a richer understanding of the social media corpus. -Facilitates advanced analysis of community discussion.

So why is Minecraft Successful?

Minecraft's ability to stay relevant through multiple social media platforms is a significant factor in its success.

The buzz it creates keeps it relevant, and this is caused by:

Frequent Updates. **Modding Capabilities** Creativity and Replayability.

Our data analysis clarifies that different people have different reasons for playing Minecraft, but the game supports all of these players.

BRIGHU SHREE, MATTHEW SKIRVIN, AND THE DATAMINE STAFF

References:

Purdue University. "Submitting an IRB Protocol Application." 3 April 2022. irb.Purdue.edu. Document.

University, Jianhua Yin Tsinghua, et al. "A Dirichlet Multinomial Mixture Model-Based Approach for Short Text Clustering: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining." ACM Conferences, 1 Aug. 2014, https://dl.acm.org/doi/10.1145/2623330.2623715.