

DMTM Gacha Games Twitter Analysis

Code ▾

Hide

```
## load rtweet package
library("prettydoc")
library("rtweet")
library("tidyverse")
library("httpuv")
library("ggplot2")
library("dplyr")
library("tm")
library("wordcloud2")
library("tidyr")
library("tidytext")
```

Hide

```
auth_setup_default()
```

```
Using default authentication available. Reading auth from 'C:\Users\Sydney\AppData\Roaming\R/config/R/rtweet/default.rds'
```

Hide

```
GI <- search_tweets("#GenshinImpact", n = 5000, lang = "en", include_rts = FALSE)
```

[illegible]

Hide

View(GI)

Hide

```
Revenue <- search_tweets("Genshin Revenue", n = 5000, lang = "en", include_rts = FALSE)
View(Revenue)
```

Hide

```
Gacha <- search_tweets("Gacha Games", n = 5000, lang = "en", include_rts = FALSE)
```

```
Downloading multiple pages =>-----
Downloading multiple pages ==>-----
Downloading multiple pages ===>-----
Downloading multiple pages ====>-----
Downloading multiple pages =====>-----
```

Hide

```
View(Gacha)
```

Hide

```
head(GI$text)
```

```
[1] "\"Haha, it's interesting to see you scrambling around to find something.\\n\\n\"Looking for t
his? Here, you're welcome.\\n\\n#Wanderer #GenshinImpact https://t.co/RF8X0bsV06\"
[2] \"Traveler, look! The roof! Isn't he...!\\n\\n#GenshinImpact #Wanderer https://t.co/s6uIgdU60t\"
[3] \"The web event for Genshin Impact's new character: Wanderer is now available. Take part in t
he event to obtain Primogems!\\n\\nSumeru's environment is different from that of Inazuma. It is g
reen as far as the eye can see.\\n\\n>Click to Join Event<\\nhttps://t.co/IOXmGCNMrn\\n\\n#Gens
hinImpact https://t.co/GKlnVpKaLz\"
[4] \"GOROU #PS5Share, #GenshinImpact https://t.co/ly4HMG5JOj\"
[5] \"This luck hello??? Thanks so much!!! #wanderer #GenshinImpact https://t.co/3qZvXWv9e6\"
[6] \"I ALMOST GAVE UP HOPE BUT MY BOY CAME HOOOOMMMME~! 💜💜💜\\n\\n#GenshinImpact #Tighnari #S
caramouche #Wanderer https://t.co/iFwy0P1RP4\"
```

Hide

```
head(Revenue$text)
```

[1] "@sdsadsd28169914 @ShadyEvan4830 @DandroPrz_19 @ShirooNekoooo @xaviersonline yeah because it's soooo fair to compare an entire franchise's revenue to one game. genshin still outperforms every individual sonic game <https://t.co/b6fGWAYtT6>"

[2] "@sdsadsd28169914 @ShadyEvan4830 @DandroPrz_19 @ShirooNekoooo @xaviersonline no individual sonic game, free or not, has made more revenue than genshin. hell I'm being generous and combining the revenue of all mobile sonic games. guess what? 6+ years and it still made less money than genshin in one year!"

[3] "@sdsadsd28169914 @kkkhhhhh234 @Genshinmem You're comparing individual revenue to historical revenue. That's disingenuous, and an actually ridiculous, stupid comparison. Genshin Impact made 2 billion dollars in its first three months. Sonic Frontiers hasn't made anywhere near that amount of money in a single month."

[4] "@sdsadsd28169914 @kkkhhhhh234 @Genshinmem Genshin surpasses any individual Sonic game by a mile in terms of yearly revenue. Can you point me to a single Sonic game that has made 2 billion or more dollars in the first three months, let alone yearly?"

[5] "@sdsadsd28169914 @DisposedZero @Genshinmem I've posted links proving that genshin makes \$5 billion a year, while Sonic's total historical revenue is \$6 billion."

[6] "@sdsadsd28169914 @kkkhhhhh234 @Genshinmem Just because Sonic's free games all surpass Genshin's total downloads it doesn't necessarily follow that Sonic has more revenue or that Sonic has a larger, more active community. Besides, how do we know that most of the people who download Sonic Dash are actually Sonic fans?"

Hide

```
head(Gacha$text)
```

[1] "Attention~!! You like gacha games?? Ready for another rabbit hole? 🐇🌟\n@emishadow_rpg just launched globally! Check it out at <https://t.co/ZRi9gaswYE> and retweet this tweet and follow their page for a chance to win 10,000 gems! Count those zero's 000\n\nRules: <https://t.co/SvnAEmlz3D> <https://t.co/HXvkXYXLaT>"

[2] "huge 'thank you' to the ramen shop patrons for not stealing the backpack of an aloof tourist who abandoned it to go play gacha games for hours"

[3] "\"Gacha games should be excluded from game of the year awards.\" <https://t.co/BdZsh6YZYU>"

[4] "\"// Gacha games should always be excluded from game of the year awards. <https://t.co/eMuqkiX0Cf>"

[5] "Gacha games should always be excluded from game of the year awards. <https://t.co/2Lx0zIVPUS>"

[6] "@Axiuno_WasTaken @whisperwoIf Amazes me that the only thing people bring to the table with Genshin being trash is that \"Gacha system bad\" mentality. Like give other reasons why it shouldn't win. Gacha games aren't automatically bad or worse than others, if they're done right that is."

Hide

```
# remove http elements manually
GI$stripped_text <- gsub("http.*","", GI$text)
GI$stripped_text <- gsub("https.*","", GI$stripped_text)

Revenue$stripped_text <- gsub("http.*","", Revenue$text)
Revenue$stripped_text <- gsub("https.*","", Revenue$stripped_text)

Gacha$stripped_text <- gsub("http.*","", Gacha$text)
Gacha$stripped_text <- gsub("https.*","", Gacha$stripped_text)
```

Hide

```
# note the words that are recognized as unique by R
a_list_of_words <- c("Genshin", "genshin", "genshin impact", "Genshin Impact", "gi", "games", "game", ",", ",")
unique(a_list_of_words)
```

```
[1] "Genshin"      "genshin"      "genshin impact" "Genshin Impact"
[5] "gi"          "games"        "game"         ",",
```

Hide

```
# remove punctuation, convert to lowercase, add id for each tweet!
GI_clean <- GI %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(word, stripped_text)
```

Hide

```
# remove punctuation, convert to lowercase, add id for each tweet!
Revenue_clean <- Revenue %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(word, stripped_text)
```

Hide

```
# remove punctuation, convert to lowercase, add id for each tweet!
Gacha_clean <- Gacha %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(word, stripped_text)
```

Hide

```
data("stop_words")
head(stop_words)
```

word	lexicon
<chr>	<chr>
a	SMART
a's	SMART
able	SMART
about	SMART
above	SMART
according	SMART
6 rows	

Hide

```
nrow(GI_clean)
```

```
[1] 78861
```

Hide

```
nrow(Revenue_clean)
```

```
[1] 1440
```

Hide

```
nrow(Gacha_clean)
```

```
[1] 18218
```

Hide

```
# remove stop words from your list of words
GI_stopwords <- GI_clean %>%
  anti_join(stop_words)
```

```
Joining, by = "word"
```

Hide

```
# there should be fewer words now
nrow(GI_stopwords)
```

```
[1] 47510
```

Hide

```
Revenue_stopwords <- Revenue_clean %>%  
  anti_join(stop_words)
```

```
Joining, by = "word"
```

Hide

```
# there should be fewer words now  
nrow(Revenue_stopwords)
```

```
[1] 725
```

Hide

```
Gacha_stopwords <- Gacha_clean %>%  
  anti_join(stop_words)
```

```
Joining, by = "word"
```

Hide

```
# there should be fewer words now  
nrow(Gacha_stopwords)
```

```
[1] 7960
```

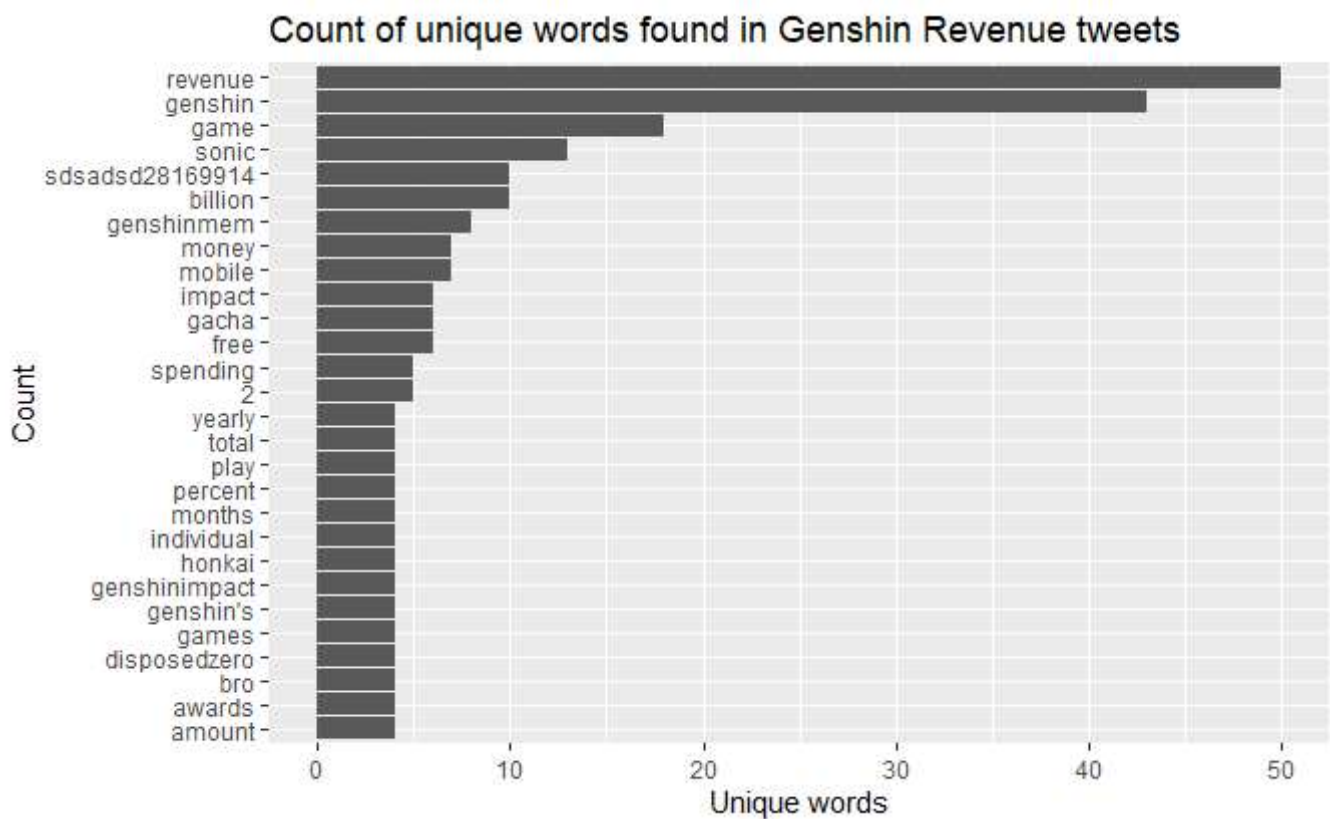
Hide

```
# plot the top 15 words  
GI_stopwords %>%  
  count(word, sort = TRUE) %>%  
  top_n(15) %>%  
  mutate(word = reorder(word, n)) %>%  
  ggplot(aes(x = word, y = n)) +  
  geom_col() +  
  xlab(NULL) +  
  coord_flip() +  
  labs(x = "Count",  
       y = "Unique words",  
       title = "Count of unique words found in #GenshinImpact tweets")
```

Hide

```
# plot the top 15 words
Revenue_stopwords %>%
  count(word, sort = TRUE) %>%
  top_n(15) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col() +
  xlab(NULL) +
  coord_flip() +
  labs(x = "Count",
       y = "Unique words",
       title = "Count of unique words found in Genshin Revenue tweets")
```

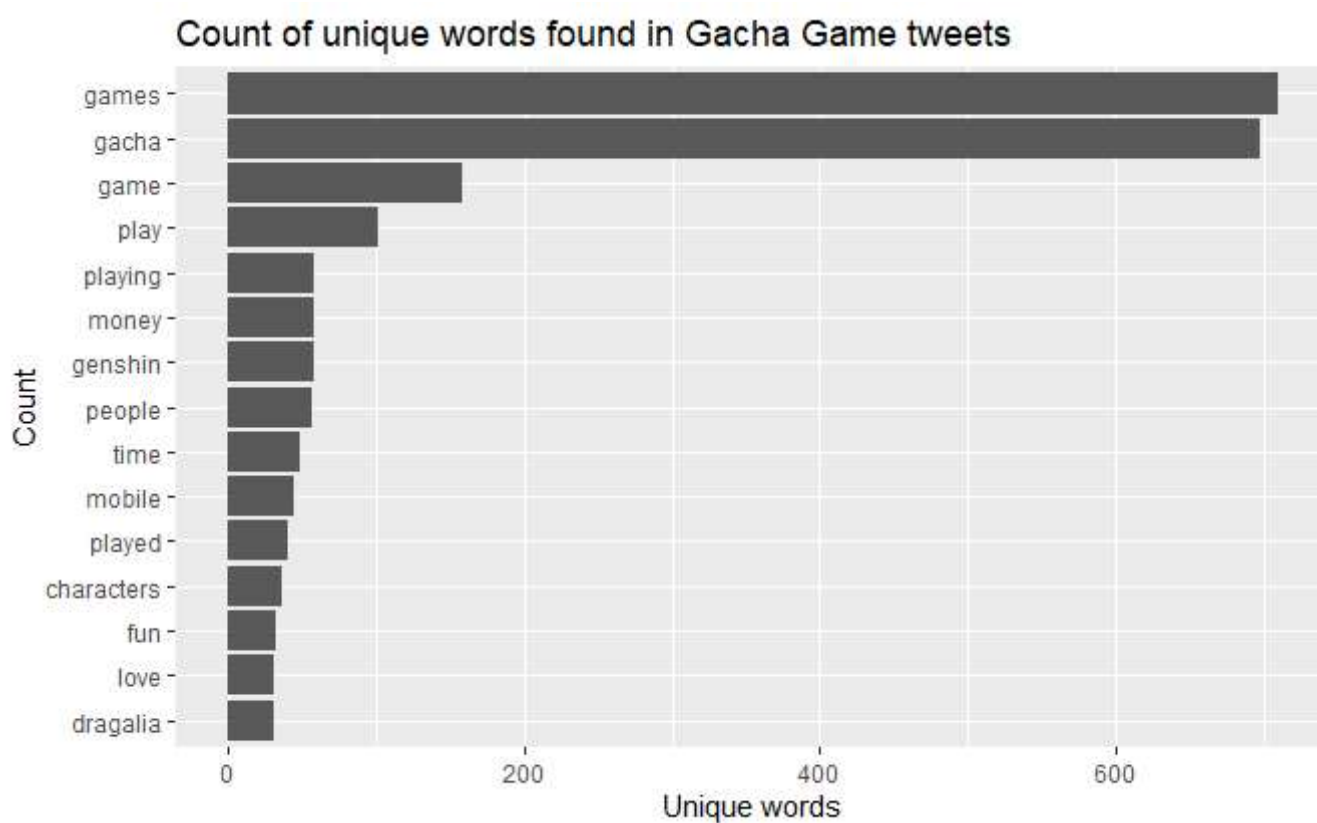
Selecting by n



Hide


```
# plot the top 15 words
Gacha_stopwords %>%
  count(word, sort = TRUE) %>%
  top_n(15) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col() +
  xlab(NULL) +
  coord_flip() +
  labs(x = "Count",
       y = "Unique words",
       title = "Count of unique words found in Gacha Game tweets")
```

Selecting by n



Words that occur together in tweets

Hide

```
#library(devtools)
#install_github("dgrtwo/widyr")
#library(widyr)
# Genshin Impact
# remove punctuation, convert to lowercase, add id for each tweet!
GI_paired_words <- GI %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(paired_words, stripped_text, token = "ngrams", n = 2)

GI_paired_words %>%
  count(paired_words, sort = TRUE)
```

paired_words	n
<chr>	<int>
原神	696
genshin impact	387
genshinimpact 原	267
happy birthday	234
genshinimpact genshin	228
mommy sorry	212
scaramouche wanderer	199
genshin genshinimpact	168
神 genshinimpact	163
genshinimpact scaramouche	160

1-10 of 43,321 rows

Previous123456...100Next

Hide

```
#Revenue
Revenue_paired_words <- Revenue %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(paired_words, stripped_text, token = "ngrams", n = 2)

Revenue_paired_words %>%
  count(paired_words, sort = TRUE)
```

paired_words	n
<chr>	<int>
the revenue	8
genshin impact	6
in the	5

paired_words	n
<chr>	<int>
of money	5
disposedzero genshinmem	4
sdsadsd28169914 disposedzero	4
sonic game	4
the first	4
2 billion	3
amount of	3
1-10 of 1,240 rows	Previous 1 2 3 4 5 6 ... 100 Next

Hide

```
#Gacha
Gacha_paired_words <- Gacha %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(paired_words, stripped_text, token = "ngrams", n = 2)

Gacha_paired_words %>%
  count(paired_words, sort = TRUE)
```

paired_words	n
<chr>	<int>
gacha games	526
of the	76
one of	56
gacha game	38
games are	36
in the	35
to play	34
play gacha	32
a gacha	31
other gacha	31
1-10 of 11,920 rows	Previous 1 2 3 4 5 6 ... 100 Next

Hide

```
library(tidyr)
GI_separated_words <- GI_paired_words %>%
  separate(paired_words, c("word1", "word2"), sep = " ")

GI_filtered <- GI_separated_words %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word)

# new bigram counts:
GI_words_counts <- GI_filtered %>%
  count(word1, word2, sort = TRUE)
```

[Hide](#)

```
#revenue
Revenue_separated_words <- Revenue_paired_words %>%
  separate(paired_words, c("word1", "word2"), sep = " ")

Revenue_filtered <- Revenue_separated_words %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word)

# new bigram counts:
Revenue_words_counts <- Revenue_filtered %>%
  count(word1, word2, sort = TRUE)
```

[Hide](#)

```
#Gacha
Gacha_separated_words <- Gacha_paired_words %>%
  separate(paired_words, c("word1", "word2"), sep = " ")

Gacha_filtered <- Gacha_separated_words %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word)

# new bigram counts:
Gacha_words_counts <- Gacha_filtered %>%
  count(word1, word2, sort = TRUE)
```

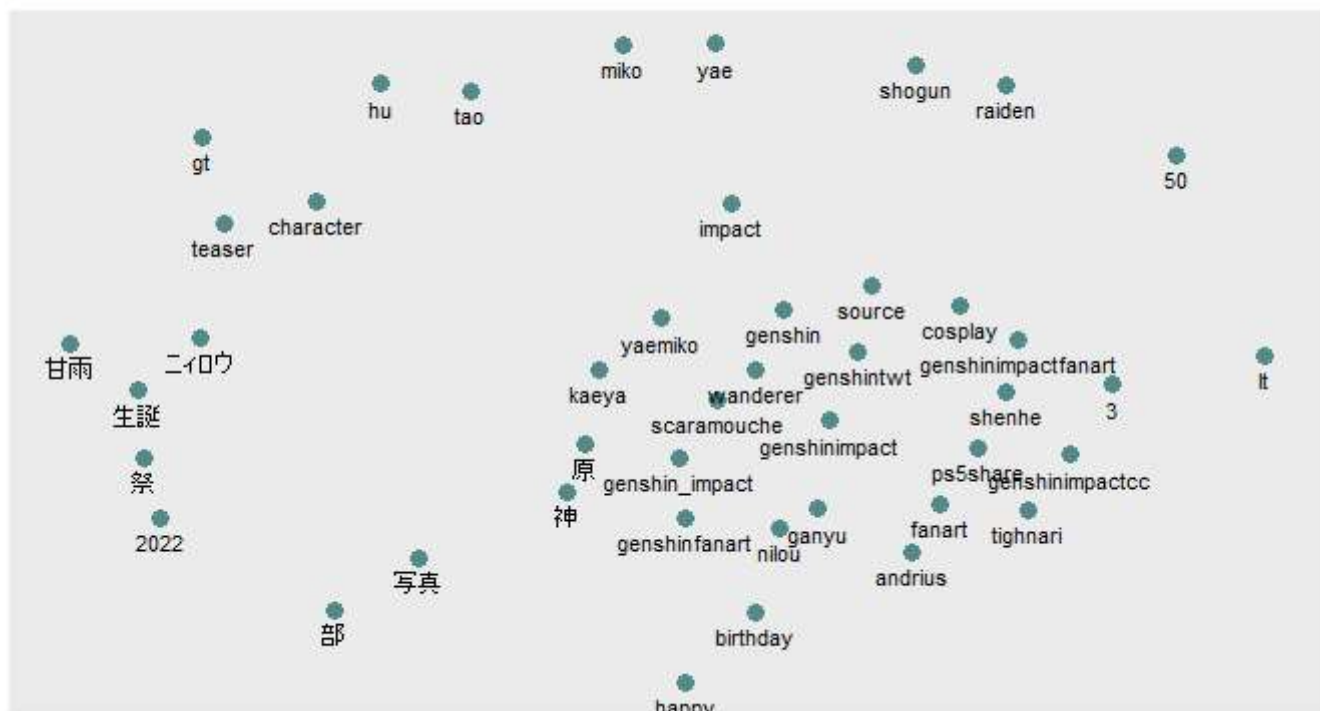
[Hide](#)

```
# (plotting graph edges is currently broken)
GI_words_counts %>%
  filter(n >= 24) %>%
  graph_from_data_frame() %>%
  ggraph(layout = "fr") +

  geom_node_point(color = "darkslategray4", size = 3) +
  geom_node_text(aes(label = name), vjust = 1.8, size = 3) +
  labs(title = "Word Network: Tweets using the hashtag - #GenshinImpact",
        subtitle = "Text mining twitter data ",
        x = "", y = "")
```

Word Network: Tweets using the hashtag - #GenshinImpact

Text mining twitter data



Hide

```
library(wordcloud)
library(RColorBrewer)
```

Hide

```
wordcloud(GI_stopwords, max.words=75,
          random.order=FALSE, colors=brewer.pal(8, "Dark2"))
```



Hide

```
wordcloud(Gacha_stopwords, max.words=75,  
          random.order=FALSE, rot.per=0.35, colors=brewer.pal(8, "Dark2"))
```

