



# ASSIGNMENT 1

## MILESTONE 1

BIG DATA ANALYTICS  
(7230 ICT)

Shehryar Mallick  
S5328488

**Contents**

1.1) ABOUT ARTIST:..... 2

1.2) DATA SELECTION & EXPLORATION: ..... 3

    STRATEGY:..... 3

    COLLECTION HISTORY: ..... 4

    SELECTION OF “TheWeeknd” HASHTAG:..... 5

1.3) TOP 5 INFLUENTIAL USERS: ..... 6

    TOP 5 INFLUENTIAL USERS:..... 6

    SIMILARITIES AMONG THE TOP 5 INFLUENTIAL USERS:..... 11

1.4) TOP 10 IMPORTANT TERMS: ..... 13

1.5) UNIQUE USERS IN DATA SET:..... 19

REFERENCE:..... 20

### 1.1) ABOUT ARTIST:

The artist that I have chosen to represent as the artist's manager is **Abel Makkonen Tesfaye** or better known by his stage name "**The Weeknd**". Abel is a Canadian singer and song writer who started his journey in 2009 when he collaborated with renowned music producer Jeremy Ross which resulted in three songs. But Abel really made his mark in the music world when he was noticed by none other than a fellow Canadian singer, rapper, and song writer Drake, who identified his potential, gave him support by promoting Able on his personal blog, and giving him access to recording studio to pursue his dream. With Drake as his guardian angel Abel released nine song Mixtape by the name House of Ballon which was shortlisted for Canada's Polaris Music Prize. And from that point on he never looked back and came to be known to the world as The Weeknd.



**The Weeknd has been active since 2009** and during this period have collaborated with several other big names in the industry some of which are Drake, Kendrick Lamar, Kanye West, Ariana Grande, Beyonce, Daft punk and many others. The complete list of **The Weeknd's songs and albums** is as follows:

Category	Frequency
Studio Albums	5
Live Albums	1
Compilation Albums	3
EPs	3
Singles	67
Soundtrack Albums	1
Mixtapes	3
Promotional Singles	6

Some notable characteristics of The Weeknd's journey are that he had the **honor of performing at the 2021 Super Bowl as the first Canadian solo singer**, which was attended by around 25000 individuals.

Another recent characteristic/highlight of The Weeknd's career is (ELLINGTON, 2023) that: **"The Weeknd has tied Michael Jackson's longstanding Billboard chart record of being the only artist to have multiple No. 1 hits from numerous albums."**

And that is not all another record which The Weeknd made is (Will, 2023) that **"the Canadian superstar (real name Abel Tesfaye) has broken the all-time record for most monthly global listeners among all artists in Spotify history."**

Last notable character that I would like to mention is that **The Weeknd has now become the most popular artist** and was officially accredited by the Guinness World Records (Atwal, 2023) on this achievement so it would not be wrong to declare him as a living legend.

## **1.2) DATA SELECTION & EXPLORATION:**

### **STRATEGY:**

- A) Create a shortlist of hashtags
- B) Write the code
- C) Run the code to accumulate as much data as possible
- D) Select the hashtag that has the potential to retrieve most data
- E) Use that data for analysis

In-order to collect the data from the artist I represent I shortlisted a total of 10 keywords which are as follows:

- 1) #weeknd
- 2) #TheWeeknd
- 3) #Starboy
- 4) #Kissland
- 5) #Afterhours
- 6) #Theweekndxo
- 7) #Blindinglights
- 8) #Abeltesfaye
- 9) #XOTWOD
- 10) #FaveTourStyle

After the shortlisting of various hashtags, I ran the R code to accumulate the data associated with the said hashtags and stored the retrieved results in a list called `twitter_data_list`.

**(Note:** The code has comments to describe what each code block does)

```

17 # QUESTION NO 1.2
18 '''
19 DATA COLLECTION OF THE ARTIST: The Weeknd
20 '''
21 #creating a list of hashtags to iterate and collect data on
22 my_hashtags = list("#Weeknd", "#TheWeeknd", "#StarBoy", "#Kissland", "#AfterHours",
23                   "TheWeekndXO", "BlindingLights", "#AbelTesfaye", "#XOTWOD", "#FaveTourStyle")
24
25 #creating an empty list that will be populated with the retrieved data
26 twitter_data_list = list()
27
28 #this for loop will run from 1 to 10 and will loop through the my_hashtags list and retrieve
29 # the appropriate key word's data
30 for (i in 1:10) {
31   print(my_hashtags[i])
32   twitter_data_list[i] <- Authenticate("twitter",
33                                       appName = my_app_name,
34                                       apiKey = my_api_key,
35                                       apiSecret = my_api_secret,
36                                       accessToken = my_access_token,
37                                       accessTokenSecret = my_access_token_secret) %>%
38   Collect(searchTerm = my_hashtags[[i]],
39           searchType = "mixed",
40           numTweets = 1500,
41           lang = "en",
42           includeRetweets = TRUE,
43           writeToFile = TRUE,
44           verbose = TRUE) # use 'verbose' to show download progress
45
46   print(i)
47 }

```

### COLLECTION HISTORY:

Date	Time	Total Data Collected	Search Type
25/March/2023	10:38	4952	Mixed

Collecting tweets for search query... Search term: #weeknd Requested 1500 tweets of 17000 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  <table><tr><th>tweet</th><th>status_id</th><th>created</th></tr><tr><td colspan="3">-----</td></tr><tr><td>Latest Obs</td><td>  1639591937069178880</td><td>  2023-03-25 11:35:54</td></tr><tr><td>Earliest Obs</td><td>  1636346578104631297</td><td>  2023-03-16 12:40:00</td></tr><tr><td colspan="3">Collected 81 tweets.</td></tr><tr><td colspan="3">RDS file written: 2023-03-25_123848-TwitterData.rds</td></tr><tr><td colspan="3">Done.</td></tr></table>	tweet	status_id	created	-----			Latest Obs	1639591937069178880	2023-03-25 11:35:54	Earliest Obs	1636346578104631297	2023-03-16 12:40:00	Collected 81 tweets.			RDS file written: 2023-03-25_123848-TwitterData.rds			Done.			Collecting tweets for search query... Search term: #Theweeknd Requested 1500 tweets of 16800 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  <table><tr><th>tweet</th><th>status_id</th><th>created</th></tr><tr><td colspan="3">-----</td></tr><tr><td>Latest Obs</td><td>  1639040603437826048</td><td>  2023-03-23 23:05:06</td></tr><tr><td>Earliest Obs</td><td>  1637676831024246784</td><td>  2023-03-20 04:45:57</td></tr><tr><td colspan="3">Collected 1500 tweets.</td></tr><tr><td colspan="3">RDS file written: 2023-03-25_123945-TwitterData.rds</td></tr><tr><td colspan="3">Done.</td></tr></table>	tweet	status_id	created	-----			Latest Obs	1639040603437826048	2023-03-23 23:05:06	Earliest Obs	1637676831024246784	2023-03-20 04:45:57	Collected 1500 tweets.			RDS file written: 2023-03-25_123945-TwitterData.rds			Done.		
tweet	status_id	created																																									
-----																																											
Latest Obs	1639591937069178880	2023-03-25 11:35:54																																									
Earliest Obs	1636346578104631297	2023-03-16 12:40:00																																									
Collected 81 tweets.																																											
RDS file written: 2023-03-25_123848-TwitterData.rds																																											
Done.																																											
tweet	status_id	created																																									
-----																																											
Latest Obs	1639040603437826048	2023-03-23 23:05:06																																									
Earliest Obs	1637676831024246784	2023-03-20 04:45:57																																									
Collected 1500 tweets.																																											
RDS file written: 2023-03-25_123945-TwitterData.rds																																											
Done.																																											

Collecting tweets for search query... Search term: #StarBoy Requested 1500 tweets of 15300 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  <table><tr><th>tweet</th><th>status_id</th><th>created</th></tr><tr><td colspan="3">-----</td></tr><tr><td>Latest Obs</td><td>  1639604165159579649</td><td>  2023-03-25 12:24:30</td></tr><tr><td>Earliest Obs</td><td>  1636157129529622531</td><td>  2023-03-16 00:07:12</td></tr><tr><td colspan="3">Collected 324 tweets.</td></tr><tr><td colspan="3">RDS file written: 2023-03-25_124002-TwitterData.rds</td></tr><tr><td colspan="3">Done.</td></tr></table>	tweet	status_id	created	-----			Latest Obs	1639604165159579649	2023-03-25 12:24:30	Earliest Obs	1636157129529622531	2023-03-16 00:07:12	Collected 324 tweets.			RDS file written: 2023-03-25_124002-TwitterData.rds			Done.			Collecting tweets for search query... Search term: #Kissland Requested 1500 tweets of 14800 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  <table><tr><th>tweet</th><th>status_id</th><th>created</th></tr><tr><td colspan="3">-----</td></tr><tr><td>Latest Obs</td><td>  1639397106057195521</td><td>  2023-03-24 22:41:43</td></tr><tr><td>Earliest Obs</td><td>  1636603074172293121</td><td>  2023-03-17 05:39:14</td></tr><tr><td colspan="3">Collected 9 tweets.</td></tr><tr><td colspan="3">RDS file written: 2023-03-25_124004-TwitterData.rds</td></tr><tr><td colspan="3">Done.</td></tr></table>	tweet	status_id	created	-----			Latest Obs	1639397106057195521	2023-03-24 22:41:43	Earliest Obs	1636603074172293121	2023-03-17 05:39:14	Collected 9 tweets.			RDS file written: 2023-03-25_124004-TwitterData.rds			Done.		
tweet	status_id	created																																									
-----																																											
Latest Obs	1639604165159579649	2023-03-25 12:24:30																																									
Earliest Obs	1636157129529622531	2023-03-16 00:07:12																																									
Collected 324 tweets.																																											
RDS file written: 2023-03-25_124002-TwitterData.rds																																											
Done.																																											
tweet	status_id	created																																									
-----																																											
Latest Obs	1639397106057195521	2023-03-24 22:41:43																																									
Earliest Obs	1636603074172293121	2023-03-17 05:39:14																																									
Collected 9 tweets.																																											
RDS file written: 2023-03-25_124004-TwitterData.rds																																											
Done.																																											

Collecting tweets for search query... Search term: #AfterHours Requested 1500 tweets of 14600 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  tweet   status_id   created ----- Latest Obs   1639429987202899969   2023-03-25 00:52:22 Earliest Obs   1636149788776316931   2023-03-15 23:38:02 Collected 1047 tweets. RDS file written: 2023-03-25_124032-TwitterData.rds Done.	Collecting tweets for search query... Search term: TheWeekndXO Requested 1500 tweets of 13400 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  tweet   status_id   created ----- Latest Obs   1639603367143149568   2023-03-25 12:21:19 Earliest Obs   1636204475357573120   2023-03-16 03:15:20 Collected 42 tweets. RDS file written: 2023-03-25_124034-TwitterData.rds Done.
Collecting tweets for search query... Search term: BlindingLights Requested 1500 tweets of 13200 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  tweet   status_id   created ----- Latest Obs   1639585359393615873   2023-03-25 11:09:46 Earliest Obs   1636399297150304258   2023-03-16 16:09:29 Collected 50 tweets. RDS file written: 2023-03-25_124038-TwitterData.rds Done.	Collecting tweets for search query... Search term: #AbelTesfaye Requested 1500 tweets of 13000 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  tweet   status_id   created ----- Latest Obs   1638988013383344129   2023-03-23 19:36:07 Earliest Obs   1636292829998731264   2023-03-16 09:06:26 Collected 17 tweets. RDS file written: 2023-03-25_124039-TwitterData.rds Done.
Collecting tweets for search query... Search term: #XOTWOD Requested 1500 tweets of 12800 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  tweet   status_id   created ----- Latest Obs   1639547551232983044   2023-03-25 08:39:32 Earliest Obs   1636224225521582080   2023-03-16 04:33:49 Collected 54 tweets. RDS file written: 2023-03-25_124043-TwitterData.rds Done.	Collecting tweets for search query... Search term: #FaveTourStyle Requested 1500 tweets of 12600 in this search rate limit. Rate limit reset: 2023-03-25 12:44:36  tweet   status_id   created ----- Latest Obs   1637967260609310726   2023-03-21 00:00:01 Earliest Obs   1637538200183791616   2023-03-19 19:35:05 Collected 1495 tweets. RDS file written: 2023-03-25_124132-TwitterData.rds Done.

## SELECTION OF “TheWeeknd” HASHTAG:

Now from the above results it is evident that the hashtag: “TheWeeknd” has the potential to have the most tweets associated with it hence we will select this as our main hashtag and set a limit of 15000.

And here we have it a total of 2280 tweets associated with the hashtag collected:

Collecting tweets for search query... Search term: #TheWeeknd Requested 15000 tweets of 18000 in this search rate limit. Rate limit reset: 2023-03-26 02:46:50  tweet   status_id   created ----- Latest Obs   1639378582794371075   2023-03-24 21:28:06 Earliest Obs   1636467067636559873   2023-03-16 20:38:47 Collected 2280 tweets. RDS file written: 2023-03-26_023306-TwitterData.rds Done.
---

status_id	is_reply	is_quote	is_retweet	created_at	text	full
1	1639378582794371075	FALSE	FALSE	2023-03-24 21:28:06	🇬🇧 UKGae 🇬🇧 FLOWERS #MileyCyrus 🇬🇧 MIRACLE #CALVI...	
2	1639040603437826048	FALSE	FALSE	2023-03-23 23:05:06	W 🇬🇧 RLDs 🇬🇧 ALBUMS 🇬🇧 Ready To Be #TWICE 🇬🇧 One Thi...	
3	1639588050925289473	FALSE	FALSE	2023-03-25 11:23:38	🇬🇧 🇬🇧 🇬🇧 🇬🇧 Flowers #MileyCyrus 🇬🇧 Kill Bill #SZA 🇬🇧 An...	
4	1639809635002748928	FALSE	FALSE	2023-03-26 02:00:57	2023-03-26T02:00:57.4658027Z ▶ 50% OFF! Get yours NO...	
5	1639808037757104128	FALSE	TRUE	2023-03-26 01:54:37	RT @mamapiinoo: #FridayFeeling #goldenhour Ready for #...	
6	1639807307285512193	FALSE	TRUE	2023-03-26 01:51:42	RT @ERNJUAQU: After Hours Be Like: 🇬🇧 🇬🇧 #Ernjuaqu #Aft...	
7	1639805931964727304	FALSE	TRUE	2023-03-26 01:46:14	RT @ERNJUAQU: Me Watching Both Elemental & Ruby...	
8	1639805760736755712	FALSE	TRUE	2023-03-26 01:45:34	RT @XOTillWereOver: do y'all remember when kanye poste...	
9	1639805404543700994	FALSE	FALSE	2023-03-26 01:44:09	do y'all remember when kanye posted this before he dropp...	
10	1639804400578248706	FALSE	FALSE	2023-03-26 01:40:09	#OnMUNow &gt; Blinding Lights - #TheWeeknd &gt; Listen=...	

Showing 1 to 10 of 2,280 entries, 54 total columns

### 1.3) TOP 5 INFLUENTIAL USERS:

```

70 # QUESTION NO 1.3 #
71 "Top 5 most influential people"
72
73 #creating actor network graph
74 the_weeknd_twitter_actor_network <- the_weeknd_twitter_data %>% Create("actor")
75 the_weeknd_twitter_actor_graph <- the_weeknd_twitter_actor_network %>% Graph()
76 V(the_weeknd_twitter_actor_graph)$name <- V(the_weeknd_twitter_actor_graph)$screen_name
77
78 write.graph(the_weeknd_twitter_actor_graph, file = "TheWeekndTwitterActor.graphml",
79             format = "graphml")
80
81 #running the ranking algo to find the top 5 most influential people
82 rank_twitter_actor_the_weeknd <- sort(page_rank(the_weeknd_twitter_actor_graph)$vector,
83                                       decreasing = TRUE)
84
85 #making a barplot of top 5 people and their influence
86 top_5_inf_people = data.frame(head(rank_twitter_actor_the_weeknd, n = 5))
87 colnames(top_5_inf_people) <- c('Influence')
88 view(top_5_inf_people)
89 H <- top_5_inf_people[['Influence']]
90 M <- rownames(top_5_inf_people)
91 barplot(H,cex.names=0.8,names.arg=M,,xlab="User Name",ylab="Influence",
92         col="lightblue",main="Top 5 Influential People")

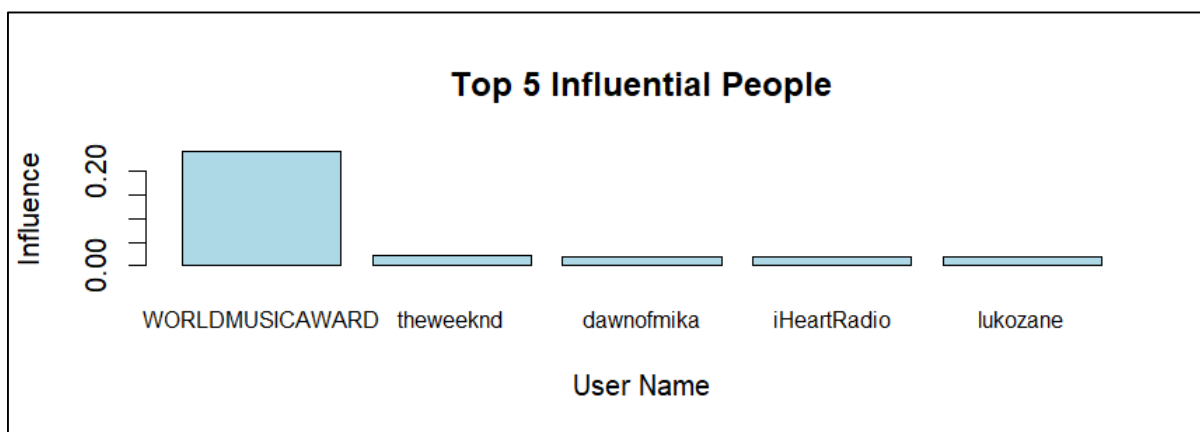
```

### TOP 5 INFLUENTIAL USERS:

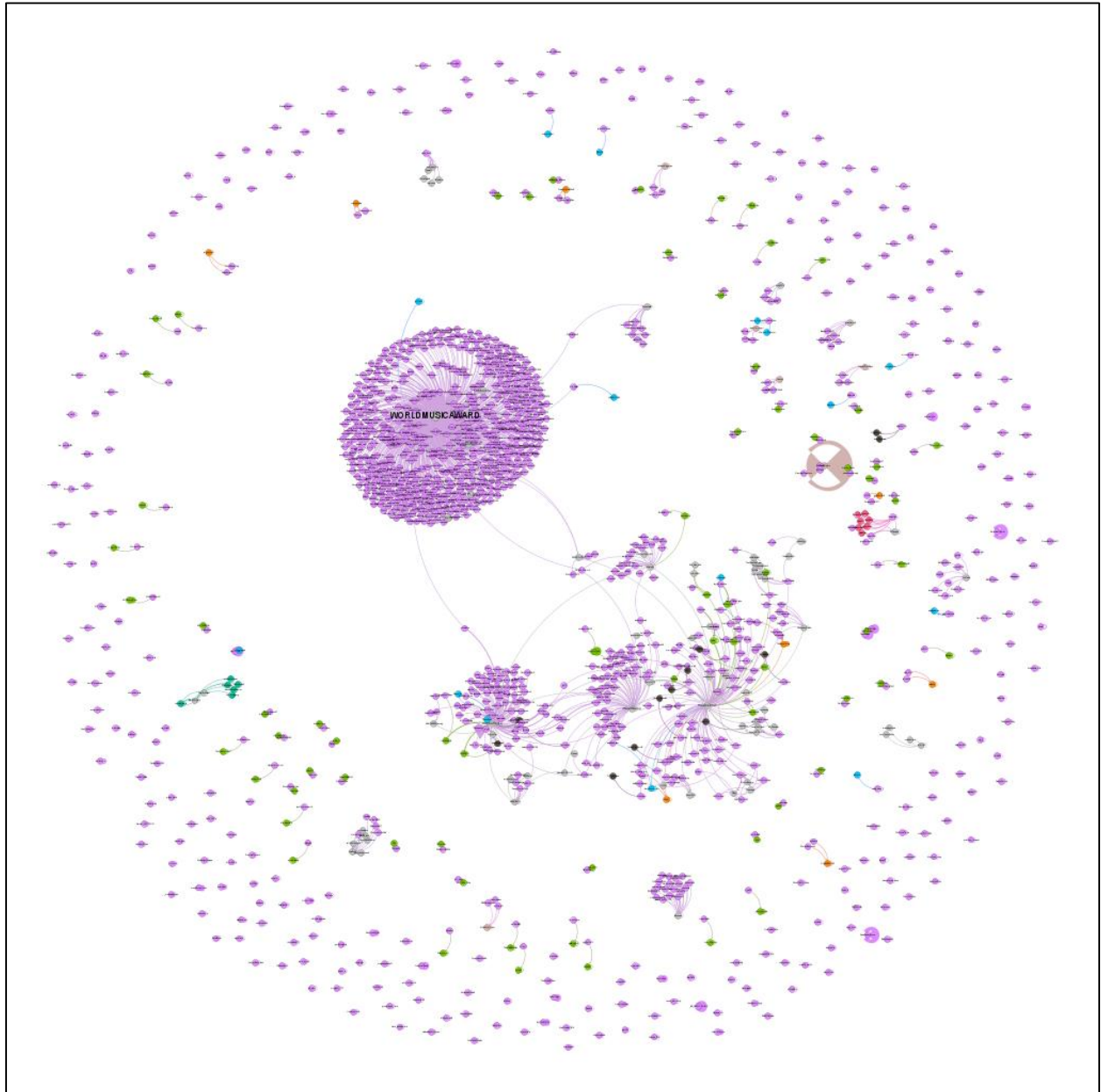
The following table depicts the twitter handle of the top 5 influential users of the hashtag under study and their influence on the rest of the users in the dataset. The influence is calculated by the PageRank algorithm of R language.

	Influence
<b>WORLDMUSICAWARD</b>	0.24267023
<b>theweeknd</b>	0.02250216
<b>dawnofmika</b>	0.01933454
<b>iHeartRadio</b>	0.01838609
<b>lukozane</b>	0.01763661

The bar diagram helps users to perceive how influential the five users are comparatively to each other.

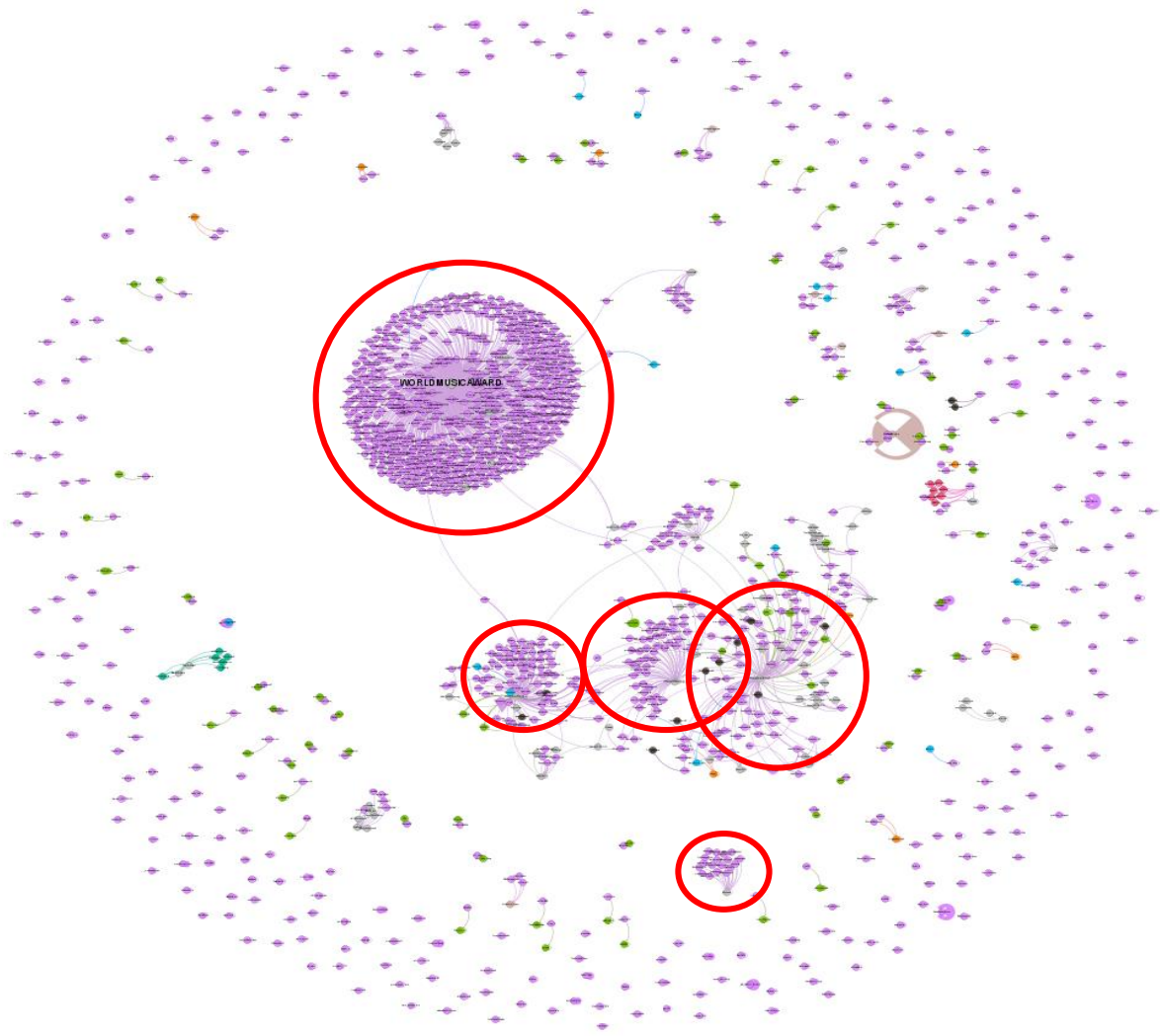


The following is the network graph of all the actors (users) that have participated in the #TheWeeknd trend.





If you observe closely, you can see the clusters of the most influential users.



Now we will filter out the top 5 users that we extracted from the R page rank algorithm to see:

- 1) How does their clusters look like?
- 2) What is the number of nodes connecting to them with depth=1,

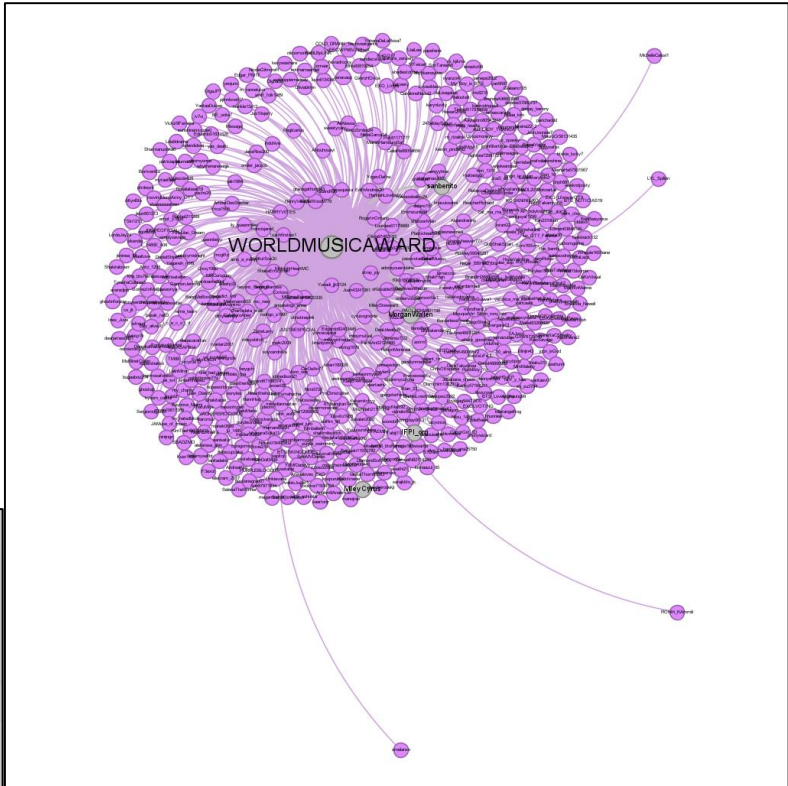
WORLDMUSICAWARD:

Context ×

**Nodes:** 560 (38.97% visible)

**Edges:** 560 (36.11% visible)

Directed Graph



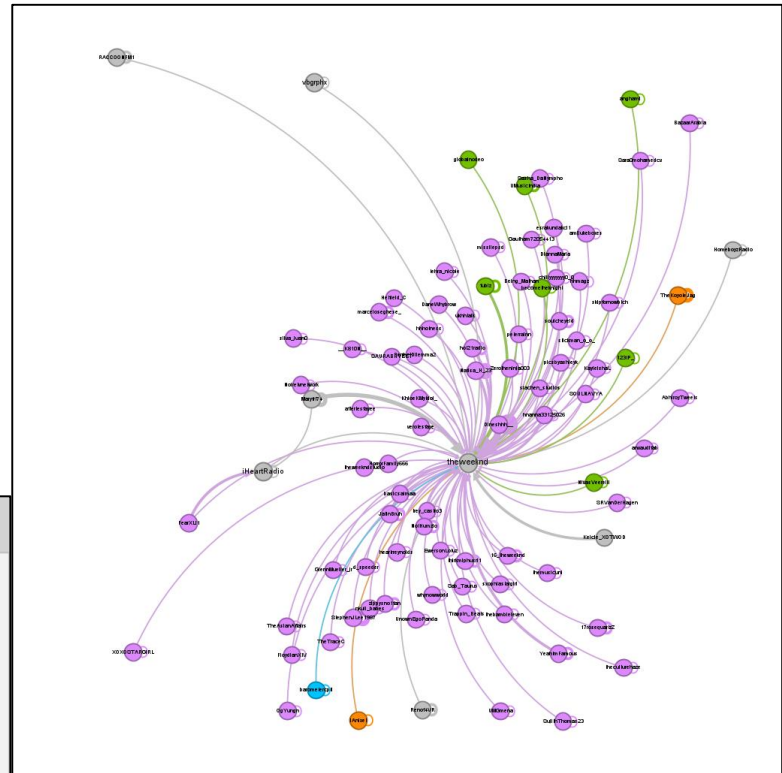
Theweeknd:

Context ×

**Nodes:** 92 (6.4% visible)

**Edges:** 174 (11.22% visible)

Directed Graph



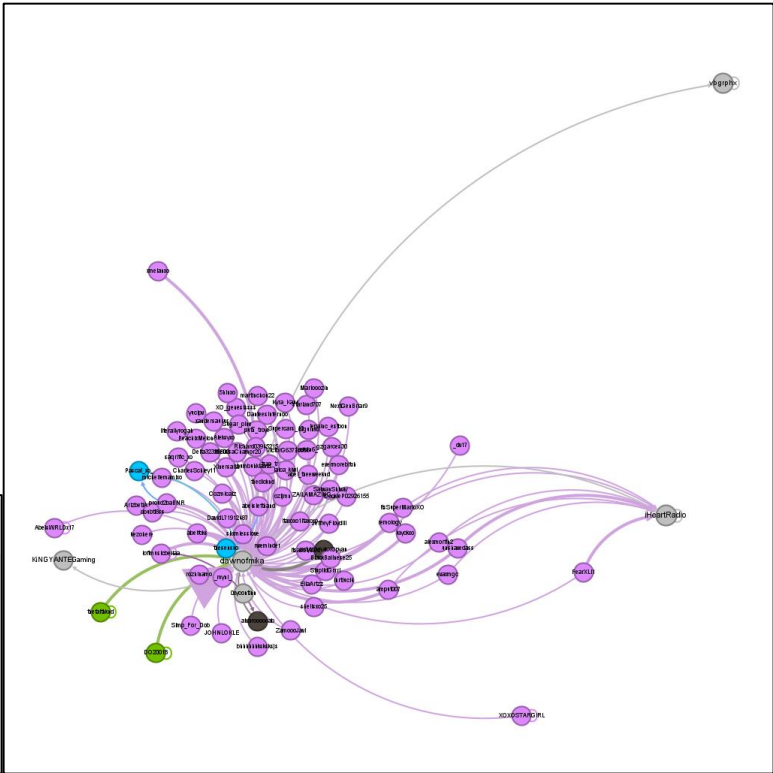
Dawnofmika:

Context ×

**Nodes:** 86 (5.98% visible)

**Edges:** 104 (6.71% visible)

Directed Graph



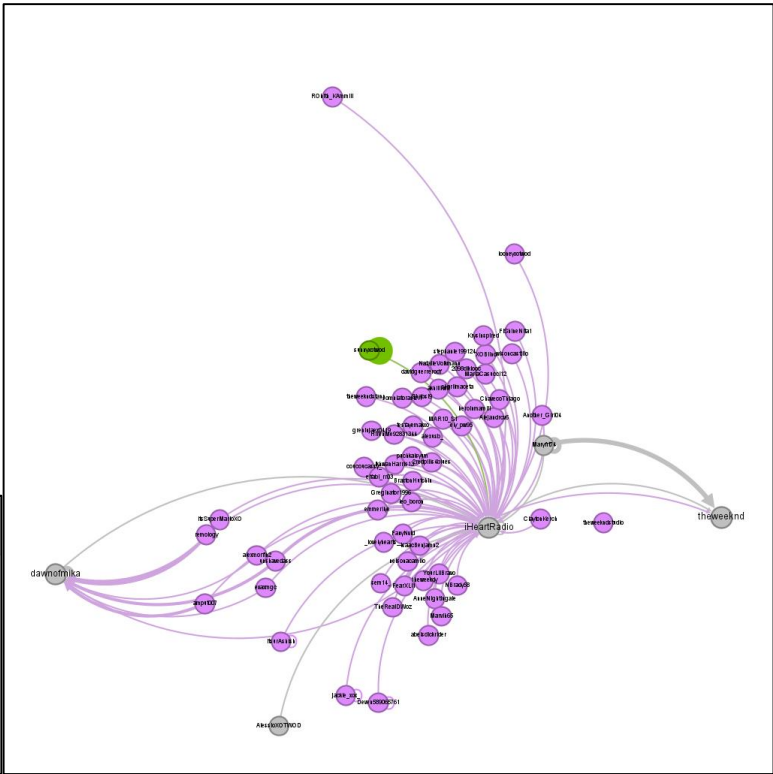
IHeartRadio:

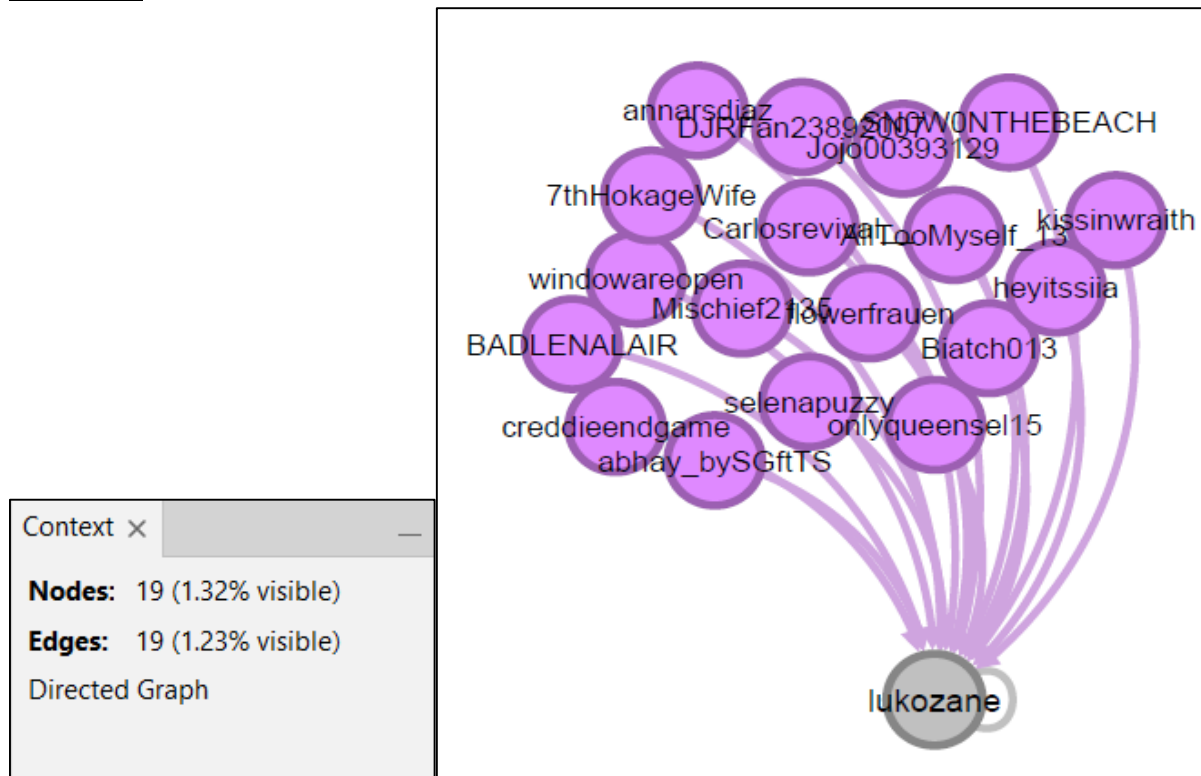
Context ×

**Nodes:** 65 (4.52% visible)

**Edges:** 81 (5.22% visible)

Directed Graph



**Lukozone:****SIMILARITIES AMONG THE TOP 5 INFLUENTIAL USERS:**

To find the similarities among the top 5 users, I went onto their twitter profiles and went through their following.

Following on twitter lists all the accounts that the users follow, the prime reason to follow someone is because your interests align with them, or you are inspired by them. After the analysis the **interesting finding regarding the similarities was**

All of the users are heavily invested in music and follow various artists, bands, and pages to keep up with the latest news in the music world.

**World Music Awards**  
@WORLDMUSICAWARD

Followers

Following

**Clout News** ✓  
@CloutNewsMedia  
Clout News is your entertainment news website.

Follow

**Audacy** ✓  
@Audacy  
Your radio, music and podcasts. Your favorite listening, one free app 🎧  
🎧 Support: @AudacyHelp

Follow

**Most Requested Live** ✓  
@MostRequestLive  
#MostRequestedLive w/ @OnAirRomeo, the most interactive show on the radio! 🎧 Coast to coast on 170+ radio stations every Saturday night! 🌍

Follow

**Spin or Bin Music** ✓  
@spinorbinmusic  
Music musings in bite-sized pieces. World Bloggers and Social Media Awards Winner. 3-time Singapore Blog Awards Winner | Contact: spinorbinmusic@gmail.com

Follow

**JUNGKOOK AMERICA** 🇺🇸  
@AMERICAJUNGKOOK  
America jungkook is a fanbase/ fan account. dedicated solely for the support of all Jungkook projects, from Buying music, charting, voting, trending, etc.

Follow

**iHeartRadio**  
@iHeartRadio

Followers you know

Followers

Following

**American Music Awards** 🏆  
@AMAs  
THE WORLD'S LARGEST 🌟 FAN-VOTED 🌟 AWARDS SHOW! Congrats to the winners of the 2022 #AMAs 🎵

Follow

**Most Requested Live** ✓  
@MostRequestLive  
#MostRequestedLive w/ @OnAirRomeo, the most interactive show on the radio! 🎧 Coast to coast on 170+ radio stations every Saturday night! 🌍

Follow

**Columbia Records** ✓  
@ColumbiaRecords

Follow

**People's Choice** ✓  
@peopleschoice  
The people have spoken 🗣️ Catch up on everything you missed 📺

Follow

**96.5 TDY** ✓  
@965TDY  
Today's Hits! 🎵 🎧 Ask ya smart speaker to 🗣️ "Play 96.5 TDY" 🎵  
Always live on the free @Audacy app! 📱

Follow

**Z100 New York** ✓  
@Z100NewYork  
@ElvisDuranShow 6a-10a @RyanSeacrest 10a-2p @MaxwellsHouse & @LaCrystalRosas 2p-6p @YoJoshMartinez 6p-10p @Shelley\_Rome 10p-

Follow

**@lukozane**

Followers

Following

**jonny** ✓  
@smgiscoming  
| award winning fan account for #EMMY and #GRAMMY nominated actress/singer selena gomez | joshua shooter 📸

Follow

**Kylie Jenner** ✓  
@KylieJenner

Follow

**Rihanna**  
@meganfox592  
SZA, Selena, Rihanna, Miley fentybeauty.com - rarebeauty.com

Follow

**Selenagomez**  
@SLnrgmzz

Follow

**Pop Tingz** ✓  
@ThePopTingz  
Pop Culture. Updates & Opinions. DM for business inquiries. Turn notifications on!

Follow

**mika x0 • my dear melancholy day!** ❤️🧡  
@dawnofmika

Followers

Following

**The Weeknd News**  
@NewsWeeknd  
Your best source for charts, new music, pictures, videos, daily news and more! | fan account

Follow

**MEMENTO MORI XO** 🕯️  
@MementoMoriXO  
We are Team MMXO, your best source for @theweeknd's radio show on @applemusic + much more! | Fan Account 🔗: apple.co/mementomori

Follow

**mat** ✓  
@stxbboy2  
x0 | fan account 📍

Follow

**Grimes** ✓  
@Grimezs  
our lady of perpetual chaos. Apprentice Planetologist at @makesoil 🧙‍♀️  
Lore for Modern Magic 🧙‍♀️

Follow



### 1.4) TOP 10 IMPORTANT TERMS:

#QUESTION NO 1.4 #

"Finding out the top 10 important terms that appear in tweets"

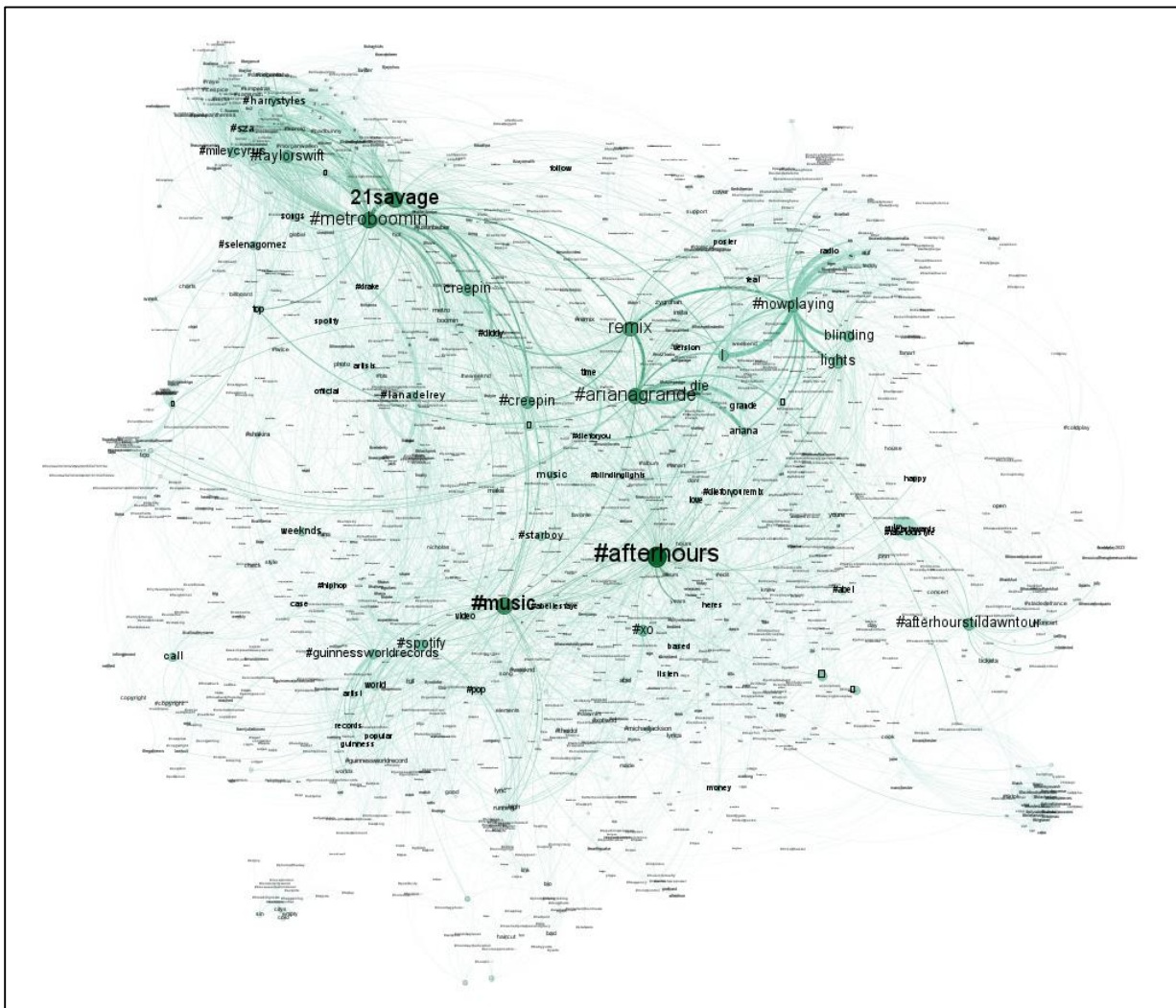
```
twitter_semantic_the_weeknd <- the_weeknd_twitter_data %>%
  Create("semantic", termFreq = 25, hashtagFreq = 75,
        removeTermsOrHashtags = c("#TheWeeknd", "weeknd", "weeknd"))
```

```
twitter_semantic_graph_the_weeknd <- twitter_semantic_the_weeknd %>% Graph()
write.graph(twitter_semantic_graph_the_weeknd, file = "TwitterSemanticTheWeeknd.graphml",
            format = "graphml")
```

```
# Run Page Rank algorithm to find important terms/hashtags
rank_twitter_semantic_the_weeknd <- sort(page_rank(twitter_semantic_graph_the_weeknd)$vector,
                                         decreasing = TRUE)
head(rank_twitter_semantic_the_weeknd, n = 10)
```

```
> head(rank_twitter_semantic_the_weeknd, n = 10)
```

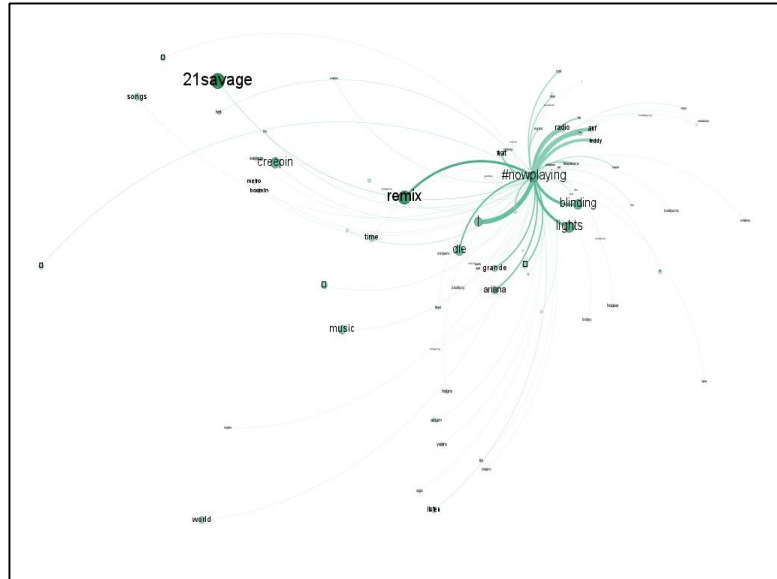
#nowplaying	#metroboomin	#afterhours	#arianagrade
0.012726266	0.010566705	0.010336954	0.009977044
21savage	remix	#music	#taylorswift
0.008600212	0.006930872	0.006876972	0.006443432
#creepin	#guinnessworldrecords		
0.005665335	0.005545086		



Now we will see the top 10 important terms and the words associated with our artist and different words that connect to them. The words selected are those that were rated as top 10 by the page rank algorithm of R language.

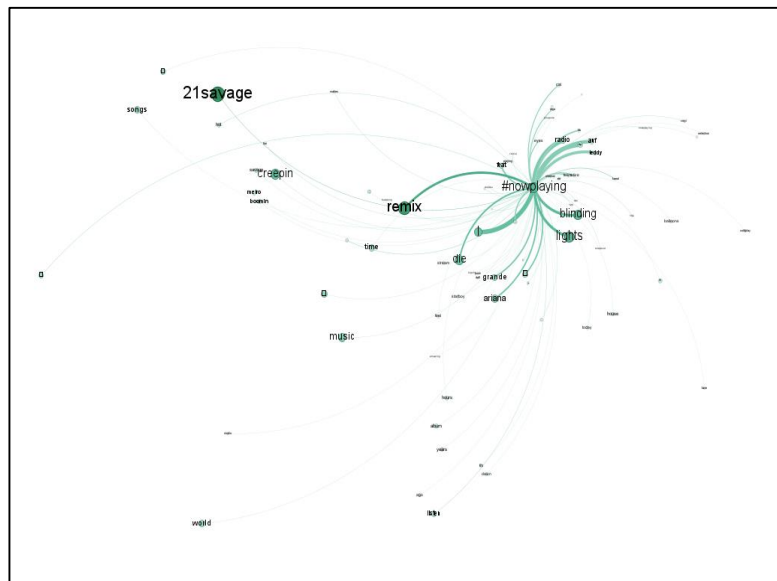
**#Nowplaying:**

This word was associated with the live album that The Weeknd released on 3<sup>rd</sup> March 2023.



## #metroboomin:

Metro Boomin is a music producer and DJ who has recently collaborated with The Weeknd in a song called **Creepin'** which also featured another singer named **21Savage**.



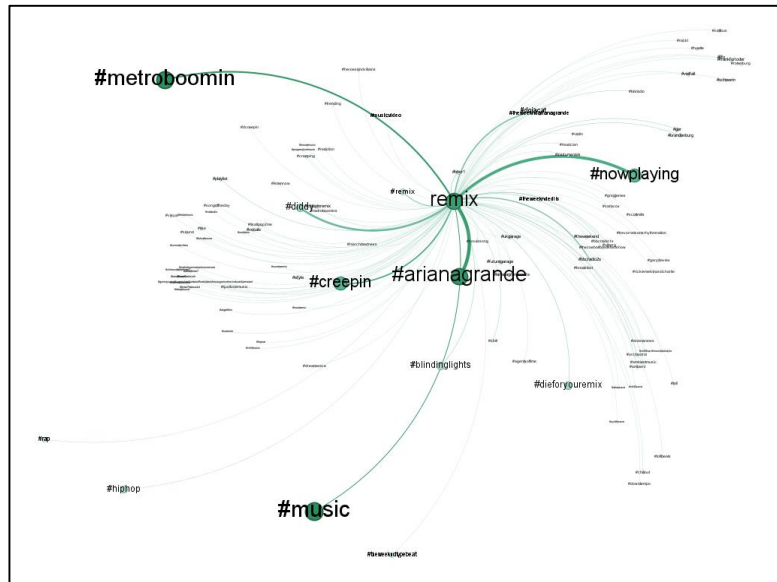






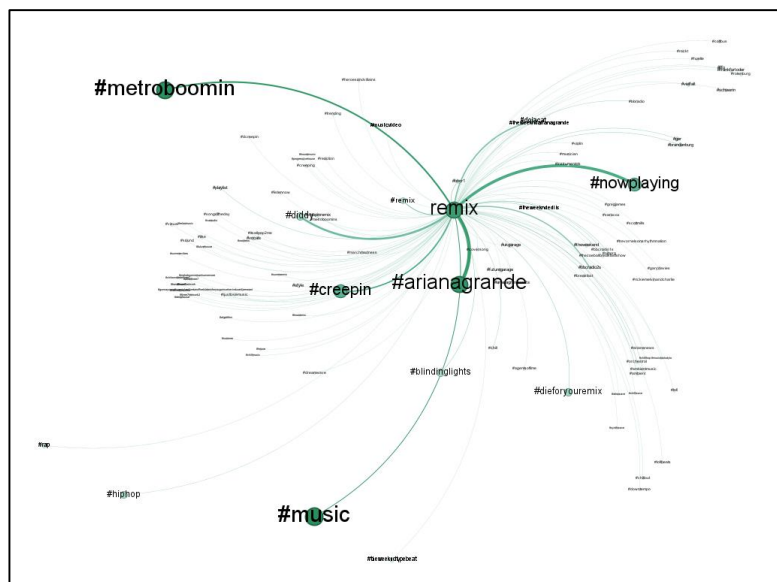
### #Music:

This hashtag is associated with the new live album that The Weeknd has released in 2023.



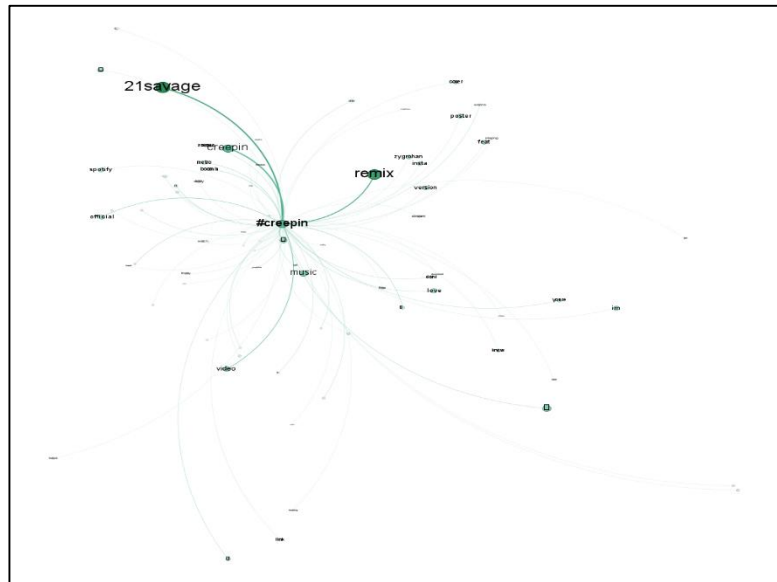
### #TaylorSwift:

Taylor Swift is renowned American singer and songwriter; she is also part of the list of popular singers who have collaborated with The Weeknd. Their collaboration happened back in **June 2021** for the song **Can't feel my face**.



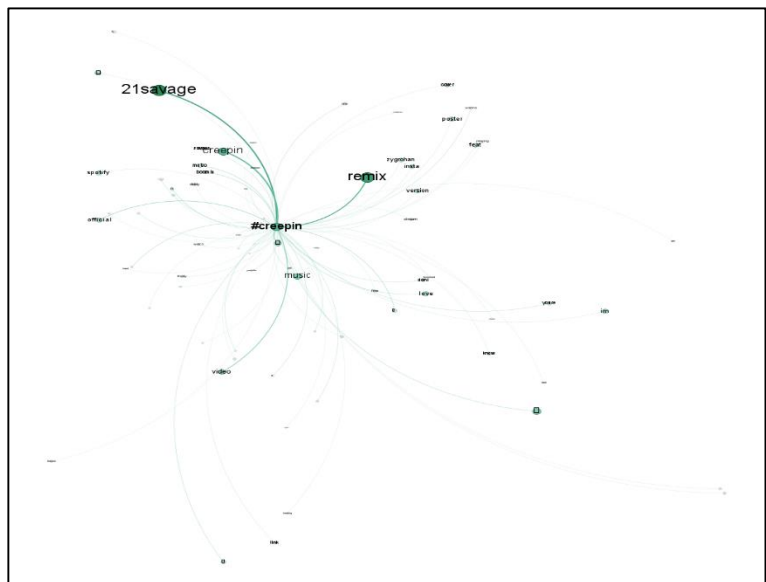
### #Creepin:

Creepin' is associated with the song that The Weeknd released with collaborators **21 Savage** and **Metro Boomin**.



### #GuinnessWorldRecords:

The final word is associated with the most recent achievement of The Weeknd as he has been officially regarded as the **Most popular artist of all time** by none other than Guinness World Records.



## 1.5) UNIQUE USERS IN DATA SET:

```
# QUESTION NO 1.5 #
unique_users_in_data_set <- length(unique(the_weeknd_twitter_data$tweets$user_id))
print(unique_users_in_data_set)
```

```
> print(unique_users_in_data_set)
[1] 1320
```

In the above code we first extracted the user\_id from the data that we had collected. The reason for extracting user\_id and not user\_name or screen\_name is that more than one user can have the same screen name, but no two users can have the same user\_id.

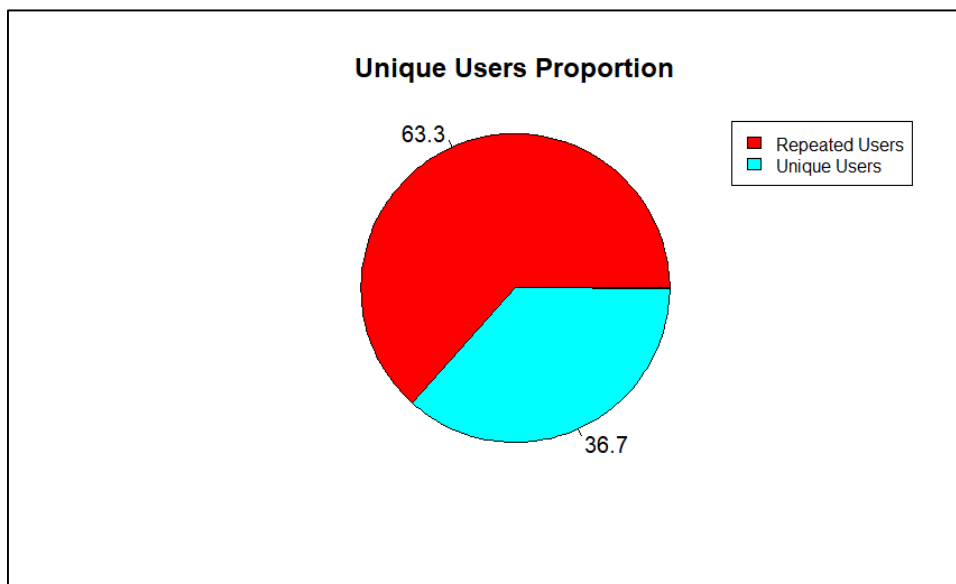
Once I had the user\_id column I used the R language's unique function to separate out all the unique ids that were in the dataset.

Then I applied the length function of R which calculates the number of occurrences.

After that I stored the length in a variable called: "unique\_users\_in\_data\_set"

Finally, I printed the number of unique users which was equal to 1320.

```
#plotting the unique users
x <- c(total_users_in_data_set, unique_users_in_data_set)
label <- c("Repeated Users", "Unique Users")
piepercent <- round(100*x/sum(x),1)
pie(x, labels = piepercent, main = "Unique Users Proportion", radius=1,
    col = rainbow(length(x)))
legend("topright", label, cex = 0.8, fill = rainbow(length(x)))
```



The above results indicate that over the past seven days around 1320 users actively engaged in tweets that were related to artist The Weeknd with the specific hashtag “#TheWeeknd”. The sudden surge of engagement is a distinctive factor of the artist’s popularity and also solidifies the notion that he is currently trending on twitter.

## **REFERENCE:**

Bauer, P. (2023, February 12). *The Weeknd*. Encyclopedia Britannica. Retrieved from <https://www.britannica.com/biography/the-Weeknd>

M. (2023, February 13). *The Weeknd's biggest collaborations: From Ariana Grande to Drake*. Planet Radio. Retrieved from <https://planetradio.co.uk/kiss/entertainment/music/the-weeknd-collaborations/>

VANDERHOOF, E. (2021, February 7). *The Weeknd's Super Bowl 2021 Halftime Show Broke Longstanding Traditions*. Vanity Fair. Retrieved from <https://www.vanityfair.com/style/2021/02/the-weeknd-super-bowl-2021-halftime-show>

ELLINGTON, A. (2023, March 8). *THE WEEKND EQUALS MICHAEL JACKSON'S BILLBOARD CHART RECORD*. HipHopDX. <https://hiphopdx.com/news/the-weeknd-equals-michael-jacksons-billboard-chart-record>

Will, I. (2023, January 14). *THE WEEKND BREAKS SPOTIFY'S ALL-TIME MONTHLY LISTENERS RECORD*. HipHopDX. <https://hiphopdx.com/news/weeknd-spotify-most-monthly-listeners-all-time-record>

Atwal, S. (2023, March 20). *The Weeknd is officially the world's most popular artist*. <https://www.guinnessworldrecords.com/news/2023/3/the-weeknd-is-officially-the-worlds-most-popular-artist-741774>