

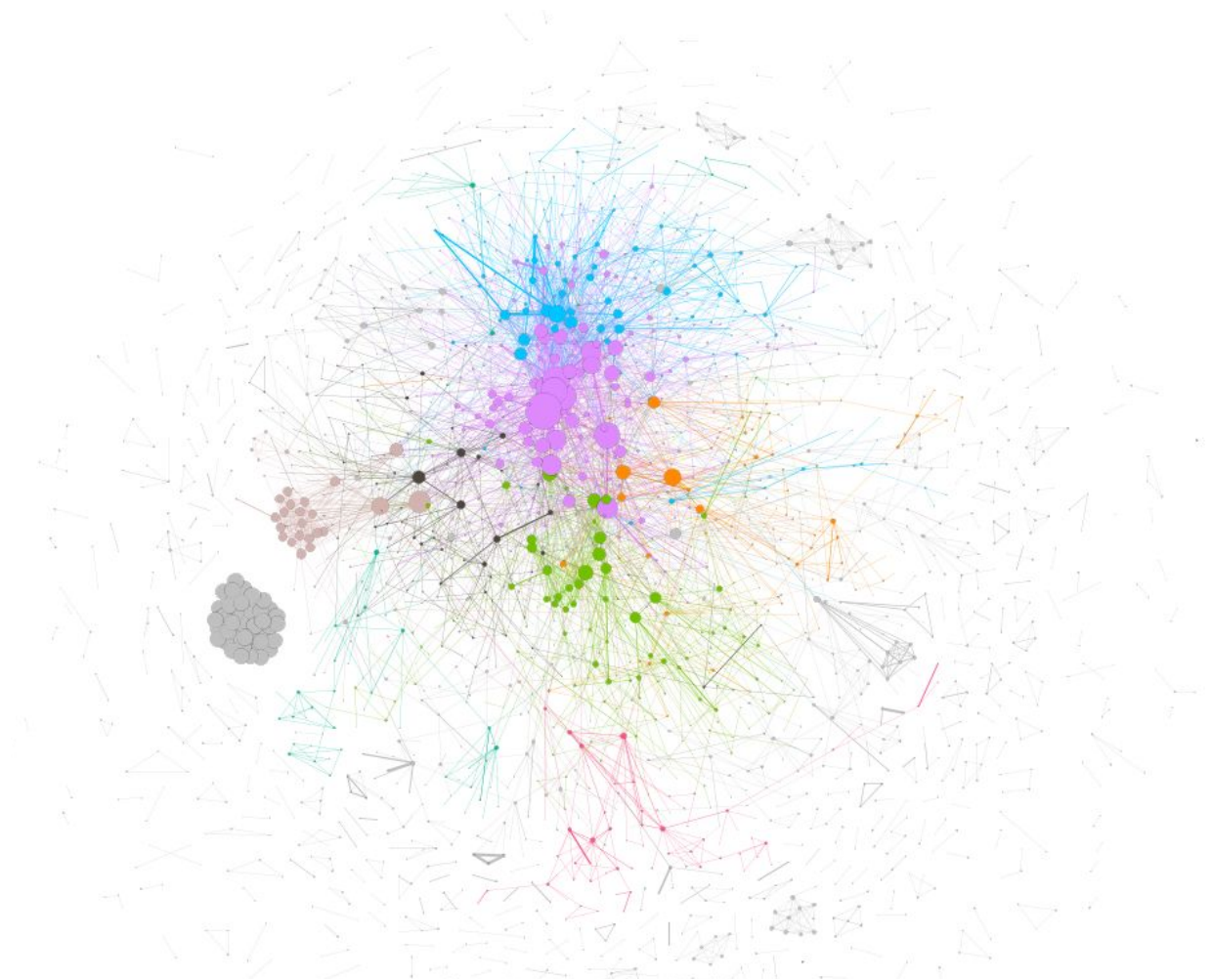
Proyecto final de Análisis de Redes Sociales

-

Estudio de las colaboraciones en la industria musical

Alberto García Doménech & Eric Heresi Medina

Enero 2021



Índice

1. Introducción	2
2. Selección de datos	3
3. Manipulación de datos y creación de las redes	3
4. Estudio de los datos procedentes de Billboard	6
4.1 Estudio global de los datos de Billboard	6
4.2 Estudio temporal de los datos de Billboard	10
4.2.1. Red con datos entre [1958 - 1979]	10
4.2.2. Red con datos entre [1980 - 1999]	11
4.2.3. Red con datos entre [2000 - 2019]	13
5. Estudio de los datos procedentes de Spotify	17
5.1 Estudio global de los datos de Spotify	17
5.2 Estudio regional de los datos de Spotify España	21
5.3 Comparaciones	25
6. Conclusiones	26

1. Introducción

En este proyecto vamos a estudiar la evolución y características de las colaboraciones en la industria musical.

A lo largo de los años hemos estado notando una cierta tendencia al alza en el número de colaboraciones que llenan las listas de éxitos de las canciones más populares. Existen numerosos artículos que hablan de esta tendencia. La gran mayoría se basan en un análisis de datos cuantitativo, es un estudio del aumento del número de estas colaboraciones en las listas de éxitos. De entre todas estas listas destaca la de Billboard Hot 100 debido a la gran cantidad de datos que contiene desde 1958 hasta la actualidad. En este top 100 de canciones que sacan semanalmente se ha podido ver cómo desde 2010 la lista ha estado dominada por colaboraciones entre artistas.

Al ver toda esta evolución nosotros también pensamos en estudiar este fenómeno desde otro punto de vista, en forma de red. Queremos estudiar de todo este aumento de colaboraciones cuáles han sido los artistas que más han influenciado en esta evolución, qué dúo o grupo de artistas han tenido más éxito trabajando juntos, si existen grupos diferenciados dentro de toda esta red de artistas, cómo los factores de popularidad o género influyen en estas colaboraciones y cómo se diferencia lo más escuchado en nuestro país actualmente en comparación a lo global.

2. Selección de datos

Los datos que se han utilizado en este proyecto han sido extraídos de dos principales fuentes, el *Hot100* de Billboard y los *Regional Charts* de Spotify.

Billboard es una revista estadounidense que desde 1958 publica semanalmente una recopilación de las 100 canciones más populares utilizando datos de la venta de discos, reproducciones en radio y (desde 2013) reproducciones en plataformas online como Spotify o Youtube. Además de haber seleccionado esta fuente por el gran número de canciones recogidas (100 canciones * 52 semanas * 62 años = 322400, aunque el número de canciones diferentes será menor porque no todas las canciones del top cambian semanalmente) se ha seleccionado Billboard porque está considerado como un estándar en la industria musical.

Estos datos los hemos extraído del dataset obtenido de *data.world* que recopila todo los *Hot100* desde el 2 de agosto de 1958 hasta el 28 de diciembre de 2019, que cuenta con los siguientes atributos importantes : songid, performer, song, spotify_track_id.

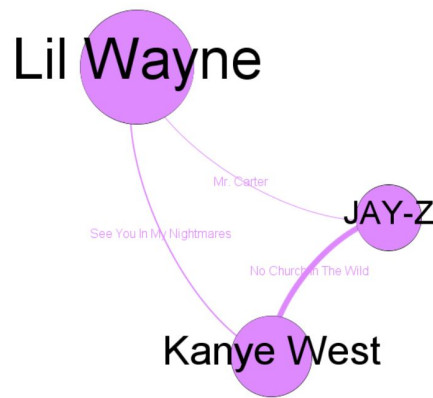
Debido a que Billboard es una revista estadounidense su top de canciones solo se basa en los números que hacen las canciones en ese mercado por lo que hemos seleccionado datos a su vez de las listas de éxitos de Spotify (con 200 canciones semanales), una de las plataformas digitales más usadas para la reproducción de música. Spotify solo lleva desde 2017 haciendo estas listas, pero las tiene divididas por territorios por lo que no será muy útiles para sacar datos de España; además tienen una lista global que usaremos para comparar con la española.

Estos datos los hemos extraído usando la librería de Python *spotifycharts* 2020.5.1 que nos permite descargar directamente desde *spotifycharts.com* a archivos .csv.

3. Manipulación de datos y creación de las redes

La red que queremos implementar consiste en utilizar los nodos como los artistas con alguna colaboración en los charts y las aristas como dicha colaboración. El peso por defecto de una arista es de 1, aunque cuantas más veces colaboren los artistas o más tiempo tengan en los tops sus canciones más peso tendrá la arista que los une. Aunque esto eliminará muchas canciones de los charts, esas canciones no nos sirven para estudiar la evolución de las colaboraciones, por lo que no pasa nada por eliminarlas.

Ejemplo de tres nodos, los cantantes Jay-Z, Lil Wayne y Kanye West, y las aristas que los unen porque los tres tienen canciones en común (por ejemplo "Mr. Carter" en el caso de Lil Wayne y Jay-Z). En la imagen podemos ver cómo la arista que une a Jay-Z con Kanye es más ancha ya que tiene peso 8 mientras que las otras dos tienen peso 1 o peso 2.



Los datos del dataset de Billboard vienen en un archivo .xlsx que utilizando la librería pandas hemos extraído para formar un dataframe con los atributos del archivo excel. Inicialmente para la creación de la red de Billboard se iba a utilizar solamente datos de este dataframe pero debido a que en canciones con muchos artistas estos no estaban propiamente acreditados (como en la canción *Almost Like Praying*) y esto supondría pérdidas de datos, decidimos utilizar la API de Spotify para la obtención de los artistas.

Canción “Almost Like Praying” acreditado a Lin-Manuel Miranda y artistas de Puerto Rico:

	SongID	Performer	Song	spotify_genre	spotify_track_id	spotify_track_preview_url
1490	Almost Like PrayingLin-Manuel Miranda Featuring Artists For Puert...	Lin-Manuel Miranda Featuring Artists For Puert...	Almost Like Praying	['hollywood', 'show tunes']	3dskm5fAe8OkBtM9E0VWwX	https://p.scdn.co/mp3-preview/758d2c0cdd4520bf...

Canción “Almost Like Praying” acreditada en Spotify a Lin-Manuel Miranda, Alex Lacamoire, Ana Villafañe, Anthony Ramos, Camila Cabello,...



Utilizando la API de Spotify, con una simple query utilizando el atributo *spotify_track_id* del dataframe conseguimos obtener una lista con información de dicha canción(nombre, artistas, álbum, fecha,...)

Habiendo resuelto ese problema, procedemos a crear la red haciendo uso de un script que recorriendo el dataframe, dado un id de una canción, busque ese id en la API de Spotify e introduzca en la red una arista que represente la colaboración entre los artistas que participan(networkx crea los nodos de los artistas si no existían previamente), aumentando el peso en 1 si la arista ya existía(por dos posibles motivos: la canción lleva más de una semana en el top o los artistas ya habían colaborado antes).

```
def formarGrafo(dataframe, _id,name):
    G = nx.Graph()
    errores = 0
    canciones_totales = 0
    colabos = 0
    for index, row in dataframe.iterrows():
        res = str(row[_id])
        if len(res) > 3:
            canciones_totales += 1
            try:
                result = sp.track(res)
            except:
                errores +=1
            else:
                artists = result["artists"]
                index = 0
                if(len(artists) > 1):
                    colabos += 1
                    while index < len(artists) -1:
                        indexb = index+1
                        while indexb < len(artists):
                            #G.add_edge(artists[index]["name"], artists[indexb]["name"], Song = str(row[name]))
                            if G.has_edge(artists[index]["name"], artists[indexb]["name"]):
                                # we added this one before, just increase the weight by one
                                G[artists[index]["name"]][artists[indexb]["name"]]['weight'] += 1
                            else:
                                # new edge. add with weight=1
                                G.add_edge(artists[index]["name"], artists[indexb]["name"], weight=1, Song = str(row[name]))
                                nx.set_node_attributes(G, values = {artists[index]["name"]:artists[index]["id"], artists[indexb]["name"]:artists[indexb]["id"]}, name='artist_id')
                                indexb+=1
                        index+=1
                    return G,errores, canciones_totales, colabos
```

Podemos ver que estamos contando el número de canciones totales y de colaboraciones para ver el porcentaje de canciones que están entre las 100 más populares y son colaboraciones. Después de ejecutarlo observamos que de 23.743 canciones 2868 son colaboraciones, es decir más del 12% de las canciones en el *Hot100*.

Además de una red global de las colaboraciones que aparecen en el *Hot100* a lo largo del tiempo,modificando un poco el código usado anteriormente para conseguir la fecha de salida de la canción, creamos tres redes distintas: una con todas las colaboraciones entre 1958(la creación del *Hot100*) hasta el año 1980(sin incluir), otra con las colaboraciones entre 1980 y el año 2000, y otra desde el año 2000 hasta 2019(el último dato obtenido). Haremos uso de estas redes para ver la evolución tanto en el número de las colaboraciones como la variación de los géneros a través del tiempo.

Para la creación de las redes de Spotify primero creamos dos dataframes con los datos de los archivos *GLOBAL_top200_weekly.csv* y *SPAIN_top_200.csv* que contienen la información de las listas semanales creadas por Spotify con datos globales y españoles respectivamente, obtenidas gracias a la biblioteca *spotifycharts*. Con los dataframes definidos podemos reutilizar el código usado anteriormente para la creación de la red de Billboard y obtener así dos redes, una con la información global y otra con la información española.

4. Estudio de los datos procedentes de Billboard

4.1 Estudio global de los datos de Billboard

Habiendo procesado el dataframe de Billboard y creado la red conseguimos un total de 2338 nodos lo que significan 2338 artistas distintos con alguna colaboración en el *Hot100* y 5102 aristas(lo que no significa que haya 5102 colaboraciones ya que cuando colaboren dos artistas más de una vez solo aumentará el peso de la arista única que comparten). En este dataframe no aparecen canciones repetidas, por lo que solo aumentará el peso de las aristas cuando colaboren en nuevas canciones los artistas implicados.

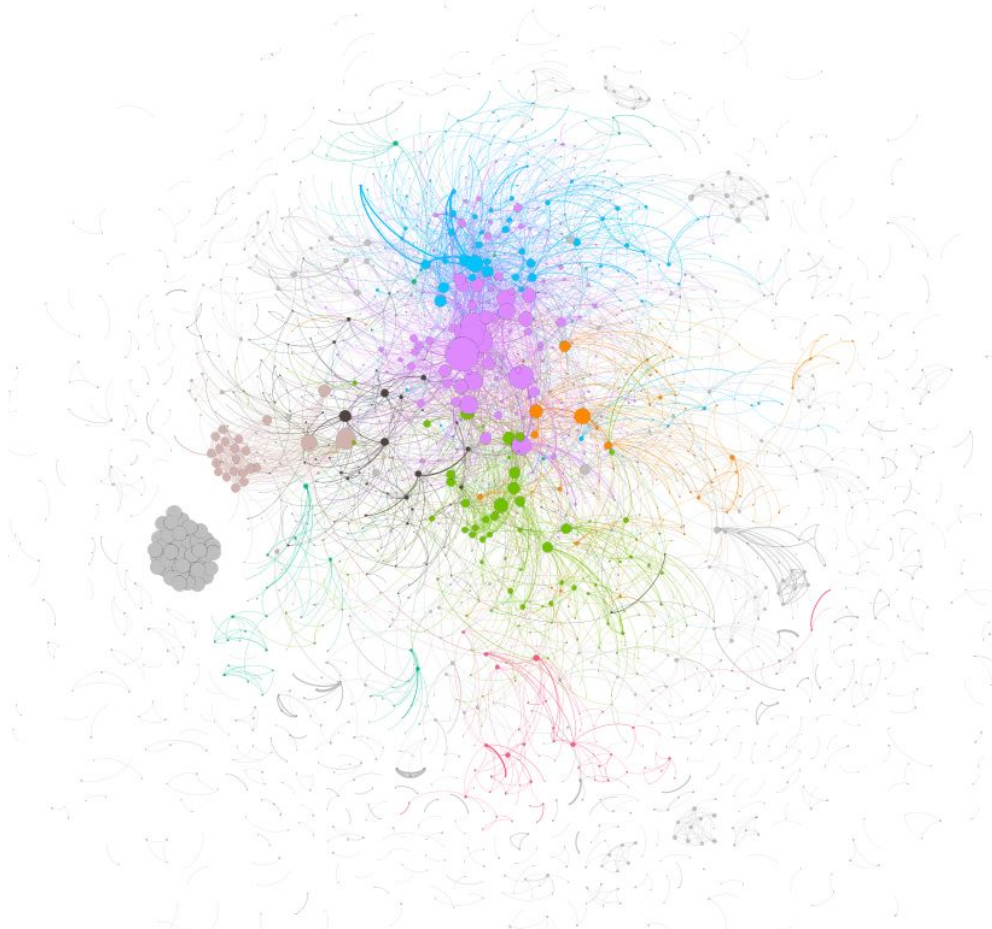
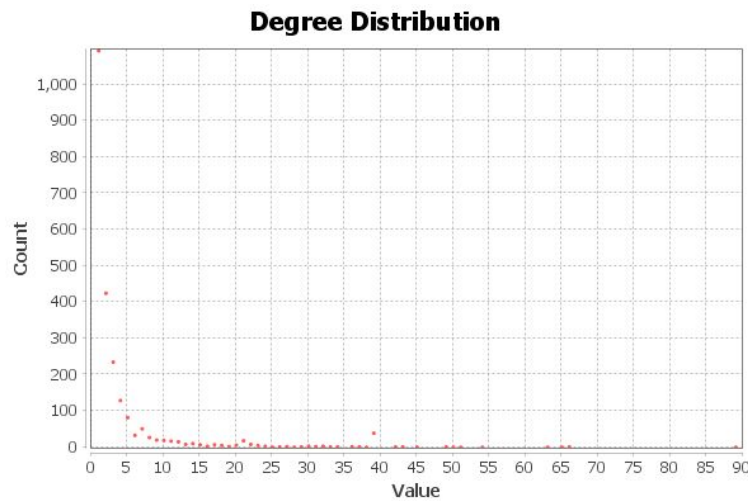


Imagen de la red con el layout Force Atlas 2

Como podemos observar en la imagen la red tiene muchas componentes conexas, exactamente 376, esto tiene sentido porque muchos de los artistas que aparecen solo han colaborado con otro artista y no se relacionan con los demás.

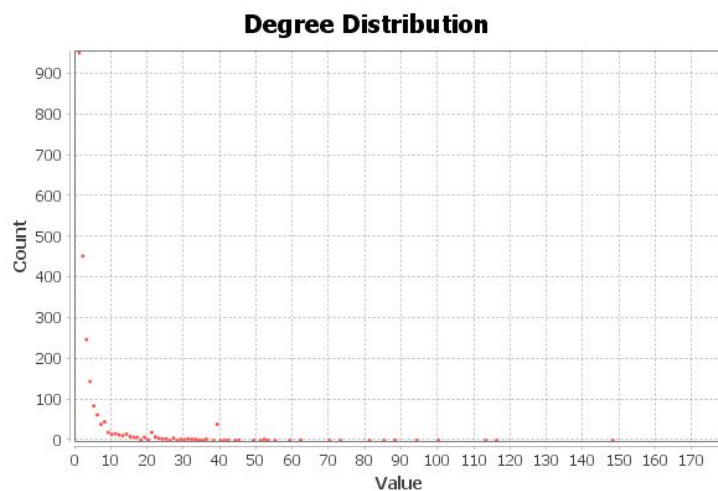
Aunque muchos de los artistas solo han colaborado con otro artista(es decir su grado es 1), el grado medio de los nodos es de 4.364, es decir que de media un artista colabora con 4 artistas. Los 5 artistas que más han colaborado con otros artistas son Lil Wayne con grado 89, Nicki Minaj con grado 66, Drake con 66, Chris Brown con 65 y Kanye West con grado

63. Todos estos artistas pertenecen a la época del 2000 en adelante y hacen música en torno al género Hip-Hop.



Distribución de grado medio

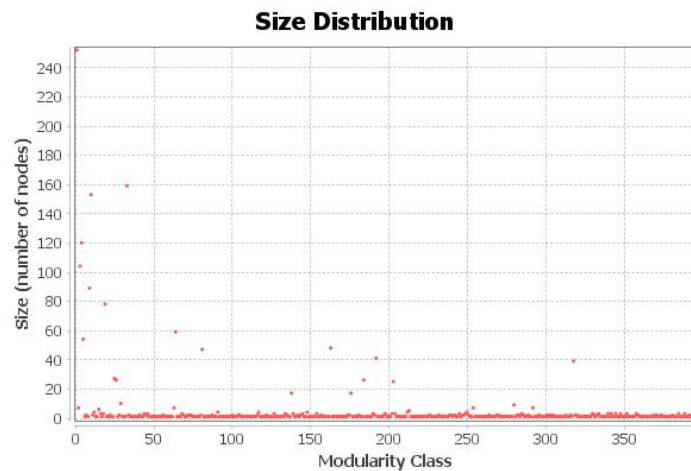
El grado medio sin tener en cuenta los pesos nos indica el número de artistas con los que han colaborado los nodos, pero si tenemos en cuenta los pesos de las aristas además de tener en cuenta los diferentes artistas contamos las diferentes colaboraciones entre los mismos artistas. Por eso, el grado medio con pesos (5.328) es ligeramente superior al grado medio, resultando que cada artistas hace de media 5 colaboraciones.



Distribución de grado medio con pesos

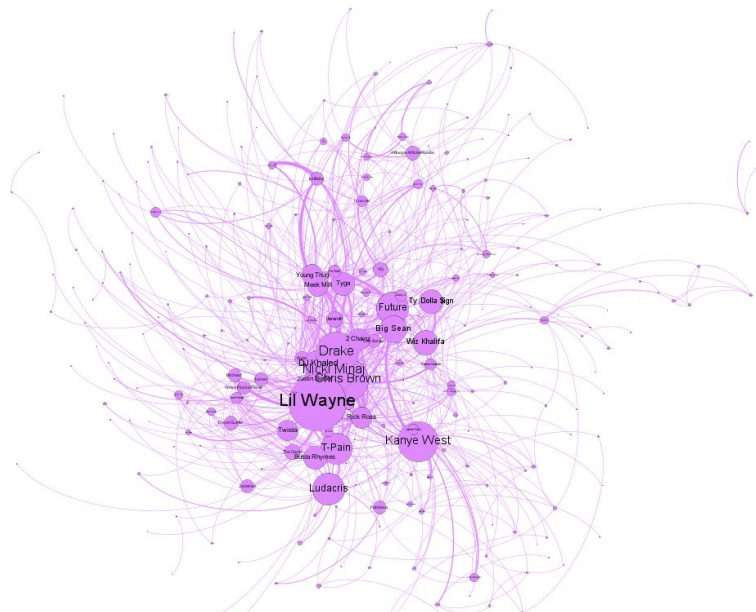
Como podemos observar en las gráficas, aunque las dos siguen una distribución exponencial, la distribución de grado medio con pesos se mueve un poco a la derecha resultando en un grado medio superior. Los nodos con grado medio con pesos superior vuelven a ser los mismos que antes, pero el orden cambia un poco: Lil Wayne sigue siendo el primero con 178 colaboraciones pero Drake ahora es segundo con 148 colaboraciones, una de las razones por las que Drake tiene tanto peso es por la arista que comparte con Lil Wayne, con 19 de peso, es decir, 19 canciones que han hecho juntos que han entrado en el *Hot100*.

Respecto a la distribución en comunidades, Gephi devuelve 397 comunidades, aunque la gran mayoría de ellas tienen tamaño muy pequeño (menor que 5). Esto se debe a la aparición de muchos artistas en el *Hot100* con una única colaboración



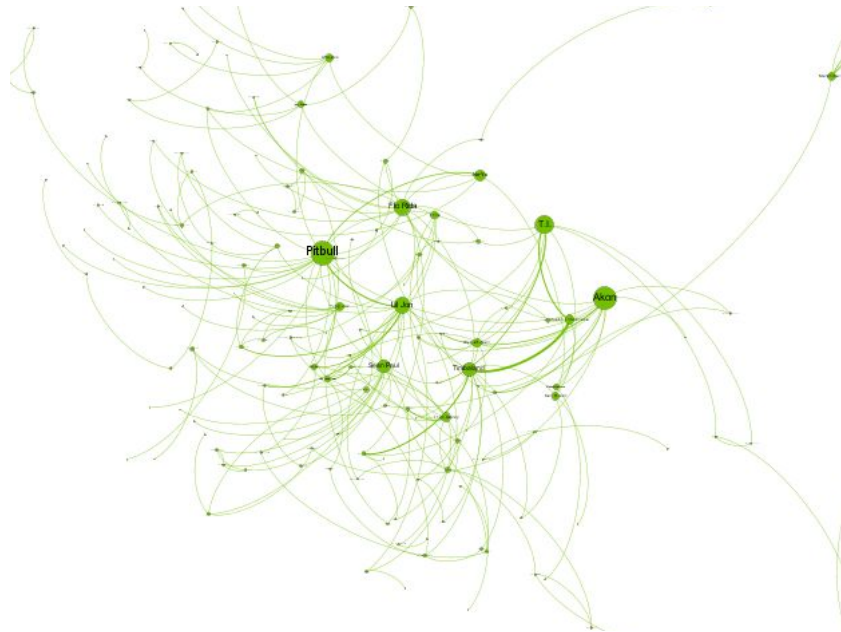
Distribución en comunidades mediante atributo modularity

Aunque la gran mayoría sean comunidades pequeñas, existen grandes comunidades en la red con tamaños como 253 nodos o 160. Vamos a ver las 3 comunidades más grandes a ver qué artistas pertenecen a cada una y si podemos encontrar un género al que pertenezcan.



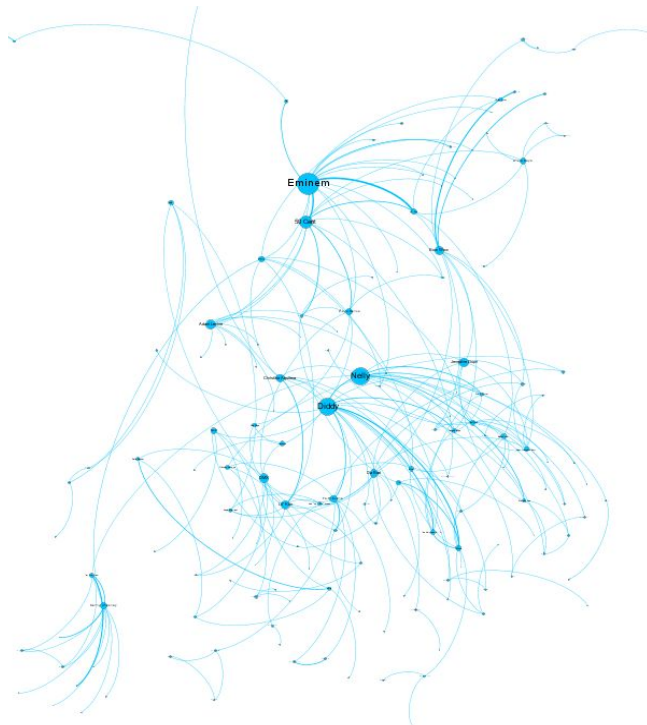
Comunidad más grande de la red con 253 nodos

Esta comunidad está formada por artistas como Lil Wayne, Kanye West, Future, Drake, T-Pain, Busta Rhymes... por lo que podemos concluir que es la comunidad perteneciente al género Hip-Hop. Esta comunidad además de ser muy grande en tamaño podemos ver que los nodos que la conforman se relacionan mucho entre sí, teniendo un grado medio (5.905 sin pesos, 8.877 con pesos) muy alto comparado al resto de la red.



Segunda comunidad más grande de la red con 160 nodos

Esta comunidad está formada por artistas como Pitbull, Flo Rida, Lil Jon... la gran mayoría de los artistas pertenecientes a esta comunidad hacen hip-hop también, pero pese a ello no colaboran con artistas de la primera comunidad porque los ritmos de los de la segunda suelen ser mucho más fiesteros y bailables, con toques latinos. Esta comunidad pese a ser muy grande en tamaño podemos ver que los nodos que la conforman se relacionan poco entre sí, teniendo un grado medio (2.75 sin pesos, 3.05 con pesos) inferior al de la media de la red. Esto se debe a que la gran mayoría de ellos solo aparecen una vez en el *Hot100* colaborando con dos artistas. Pitbull es el artista que más ha colaborado en esta comunidad con un total de 21 artistas colaborados.



Tercera comunidad más grande de la red con 154 nodos

Esta comunidad está formada por artistas como Eminem, Diddy(Puff Daddy),Nelly ... la gran mayoría de los artistas son raperos, que hacen hip-hop también, aunque no colaboran con los de la primera ni de la segunda comunidad. Esta comunidad es muy parecida a la segunda, ya que pese a ser muy grande en tamaño los nodos que la conforman se relacionan poco entre sí, ya que el estilo de rap que hacen es más individual, teniendo un grado medio(2.935 sin pesos, 3.377 con pesos) inferior al de la media de la red.

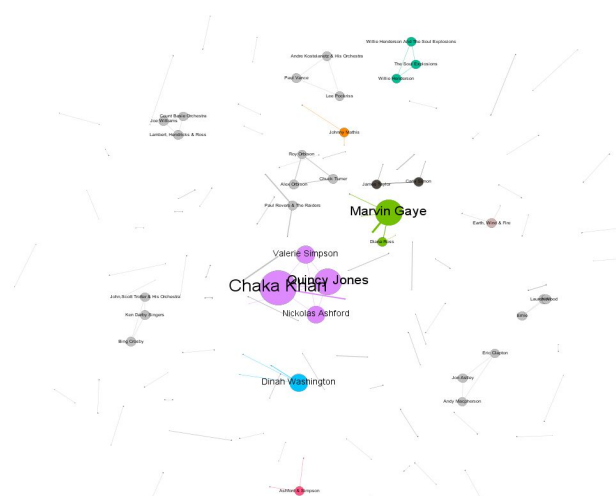
Podemos observar que las tres grandes comunidades pertenecen al mismo género, Hip-Hop, un género que aparte de ser uno de los más escuchados en los últimos tiempos (por lo que aparecerá mucho en las listas de canciones más populares) se caracteriza porque normalmente hay más de un artista en la canción, por lo que sus artistas más populares tendrán muchas colaboraciones.

4.2 Estudio temporal de los datos de Billboard

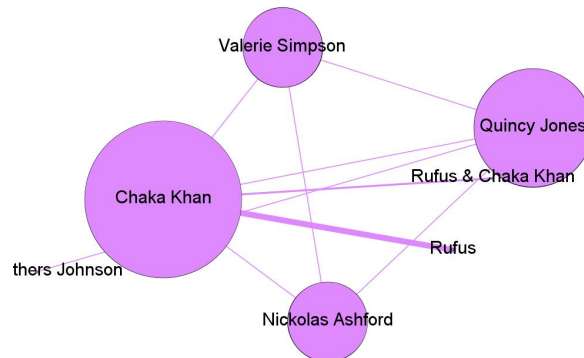
Para ver la evolución tanto en números como en géneros de las colaboraciones hemos creado tres redes pertenecientes cada una a un período de tiempo diferente: la primera desde agosto de 1958 hasta diciembre de 1979, la segunda desde enero de 1980 hasta diciembre de 1999 y la última desde el 2000 hasta diciembre de 2019.

Observando el tamaño de cada red podemos ver que efectivamente, la premisa de que las colaboraciones en la industria musical han aumentado considerablemente con el paso del tiempo es verdadera. La primera red consta de 163 nodos, la segunda de 412 y la tercera de 2293 nodos. Para verlo más claro: entre 1958 y 1980 las colaboraciones que entraban en el *Hot100* constituían el 3,24% de las canciones totales; entre 1980 y el 2000 formaban el 6,15% (casi el doble) y en las dos últimas décadas, desde el 2000 hasta 2020 el porcentaje de colaboraciones que formaban parte del *Hot100* se ha disparado hasta un 19,05%. Para explicar este cambio vamos a analizar cada una de estas redes.

4.2.1. Red con datos entre [1958 - 1979]



Esta red está formada por tan solo 194 nodos y 119 aristas, y como podemos observar en la imagen tiene un gran porcentaje de componentes conexas(85). La red está caracterizada por un grado medio muy bajo(1.227 sin pesos y 1.918 con pesos) y por la existencia de numerosas comunidades(ya que la gran mayoría no se relaciona más que con otro artista). Prueba de ello es que la comunidad más grande está formada por solo 7 nodos.



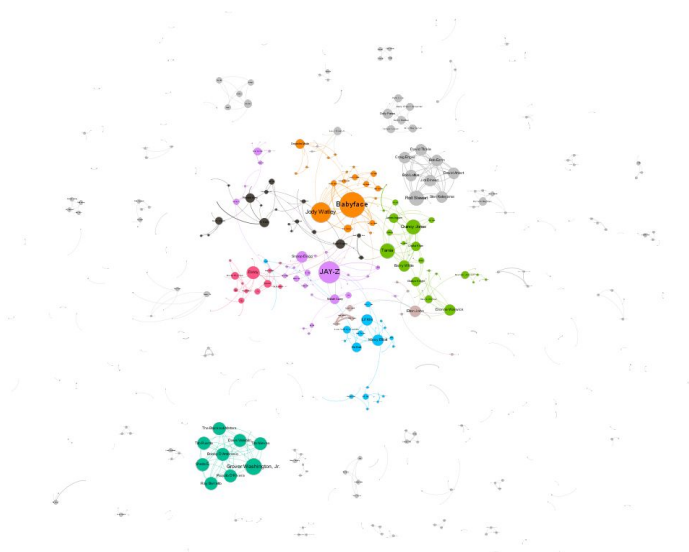
Los miembros de esta comunidad además son los artistas con más colaboraciones en esta época. Utilizando la API de Spotify vamos a ver con qué géneros están relacionados estos artistas(los artistas que aparecen con una lista vacía [] se debe a que Spotify no tiene géneros asociados a ellos, pero como se relacionan con los otros mediante colaboraciones podemos asumir que también comparten géneros):

```
# Los artistas con mas colaboraciones y con que generos estan relacionados
get_generos(G60,10)
```

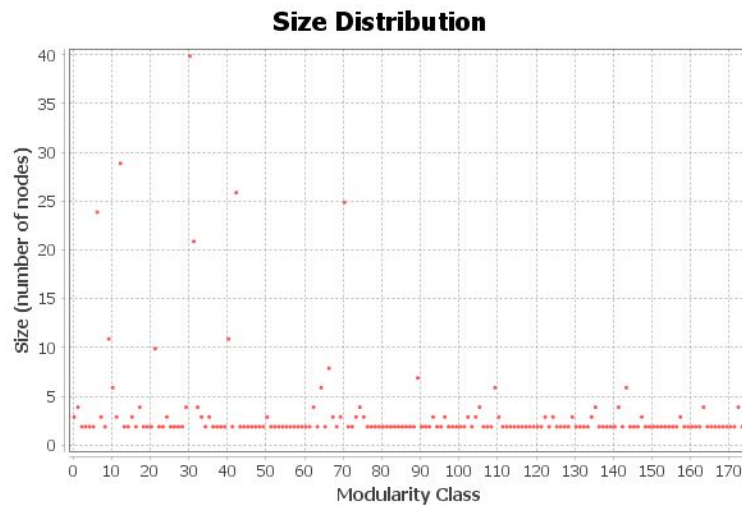
```
Artista: Chaka Khan generos: ['dance pop', 'disco', 'funk', 'motown', 'neo soul', 'post-disco', 'quiet storm', 'soul', 'urban contemporary']
Artista: Marvin Gaye generos: ['classic soul', 'funk', 'motown', 'northern soul', 'quiet storm', 'soul']
Artista: Quincy Jones generos: ['adult standards', 'disco', 'funk', 'jazz funk', 'motown', 'quiet storm', 'soul']
Artista: Dinah Washington generos: ['adult standards', 'jazz blues', 'lounge', 'soul', 'swing', 'vocal jazz']
Artista: Nickolas Ashford generos: []
Artista: Valerie Simpson generos: []
```

Como podemos ver la mayoría están relacionados con los géneros musicales que triunfaban en la época(disco,pop,soul, funk,...)

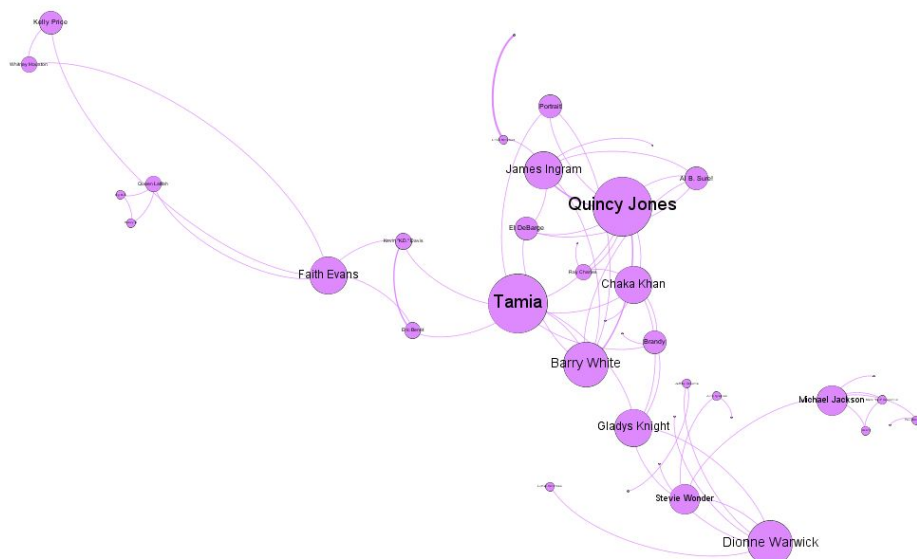
4.2.2. Red con datos entre [1980 - 1999]



Esta red está formada por 601 nodos y 610 aristas. Podemos observar que hay menos cantidad relativa de componentes conexas y que las comunidades que se crean tienen mayor tamaño y están muy relacionadas entre sí. El número medio de artistas con los que colaboran aumenta ligeramente respecto a la red anterior(2.03) y las colaboraciones aumentan también como es lógico ligeramente (2.576).

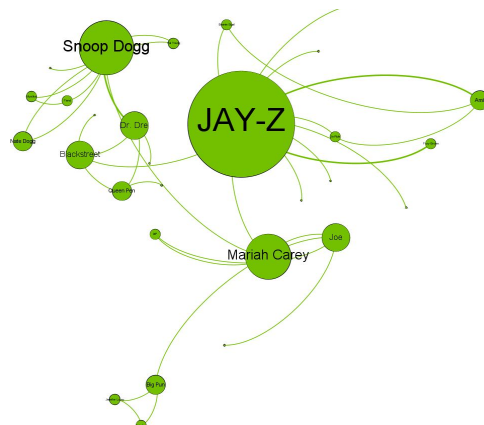


Gephi distribuye la red en 175 comunidades y como podemos ver en la distribución de estas la gran mayoría tienen tamaño 2 (ya que solo dos artistas colaboran entre sí) aunque también hay comunidades más grandes llegando hasta ser 40 artistas en la comunidad más grande. Si le echamos un vistazo a las 3 comunidades más grandes podremos ver en qué géneros triunfan más las colaboraciones y si hay artistas que habiendo aparecido en la anterior red siguen apareciendo aquí.

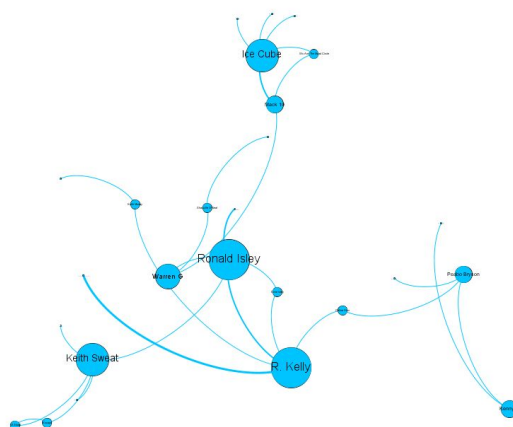


Como podemos ver, la comunidad más grande consta de 40 artistas, siendo los más reconocidos Tamia, Quincy Jones o Michael Jackson. Algunos artistas como Quincy Jones o Chaka Khan siguen apareciendo como en la red anterior pero la mayoría son artistas

nuevos. El género varía en torno al pop y al soul, aunque podemos ver que el grado medio de estos artistas ha aumentado en su comunidad respecto a la red de la anterior época.



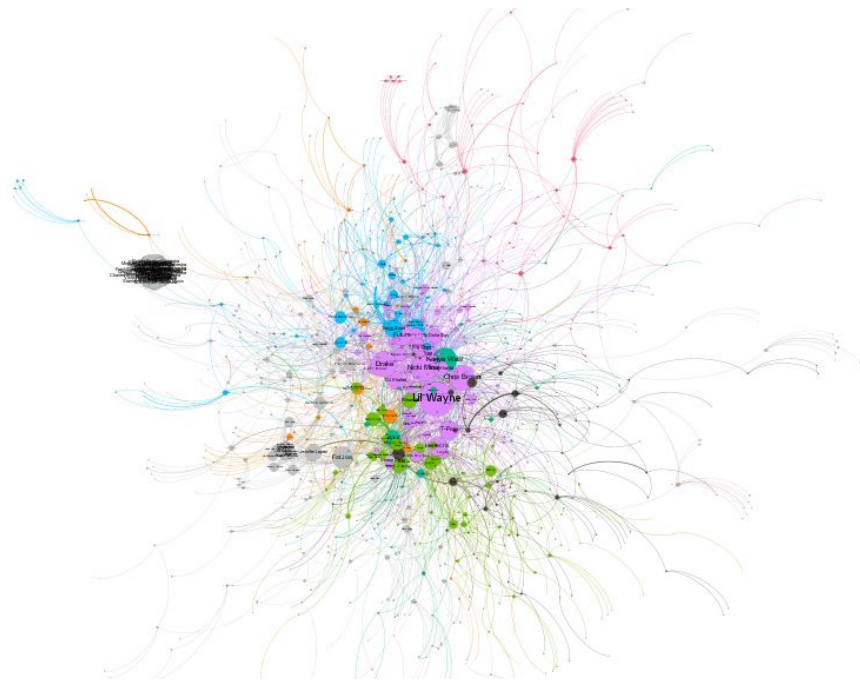
Esta segunda comunidad está formada por 29 artistas y podemos observar que giran en torno al hip-hop(género que surgió a finales de la década de los 70 en las block parties del Bronx y que se popularizó en la siguiente década). En esta comunidad vemos a artistas como Jay-Z, Snoop Dogg, Mariah Carey o Dr.Dre, muy populares en los últimos años del milenio. Aunque pueda parecer que este género tendría más grado medio al participar normalmente más artistas en las canciones, la comunidad tiene un grado medio de 2.483 artistas colaborados y 2.69 colaboraciones por artista.



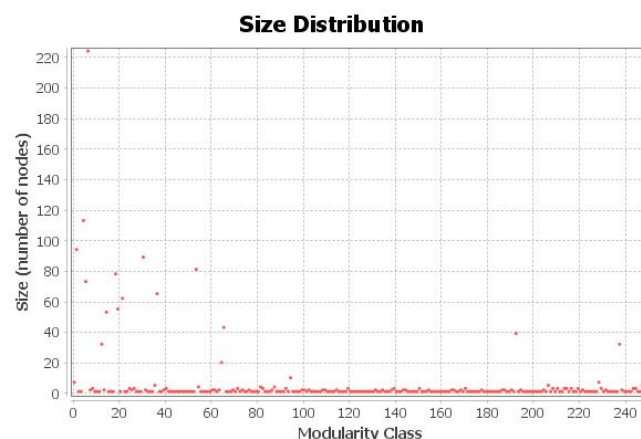
La última comunidad está formada por artistas como Ronald Isley, R. Kelly,Keith Sweat... Es una comunidad formada por 26 nodos con un grado inferior a las otras dos(2.154) y menos colaboraciones por artistas(2.53) y que gira en torno al género soul y al R&B.

4.2.3. Red con datos entre [2000 - 2019]

Esta última red es la más grande de las tres, formada por 1759 nodos y 4424 aristas. Además es la red con menos componentes conexas relativas, teniendo un total de 234 componentes conexas. En esta red cada artista colabora de media con otros 5.03 artistas y produce un total de 5.99 colaboraciones de media, como podemos observar un grado medios muy superior al de las anteriores épocas.

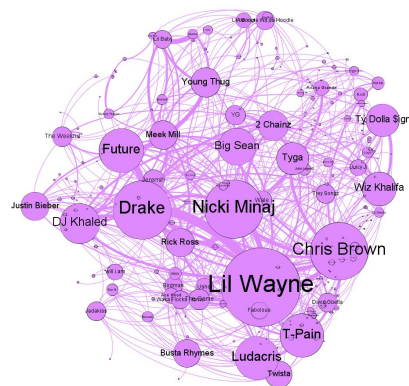


Como podemos observar en la imagen los nodos tienen un grado alto, y las comunidades que se forman están muy relacionadas entre sí. Viendo la distribución por comunidades podemos ver que hay 254 comunidades diferentes y que aunque la mayoría (como en el resto de las redes) tienen un tamaño menor que 5, aunque hay comunidades muy grandes con tamaños que llegan a 228 artistas diferentes.



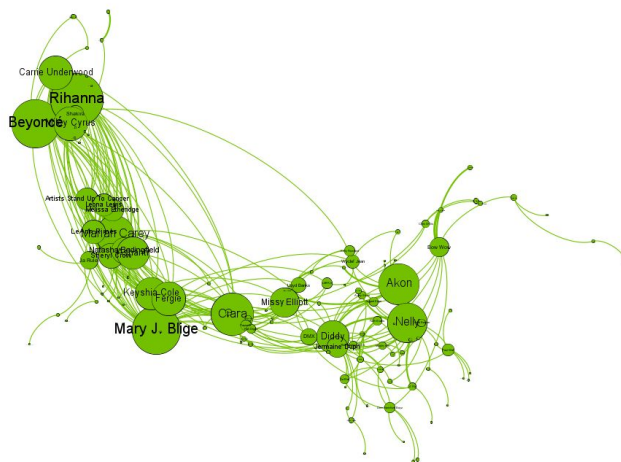
Las 4 comunidades más grandes tienen cada una de ellas más de 100 artistas, a continuación vamos a ver qué artistas son, si se relacionan mucho entre sí y si el género al que pertenecen está relacionado.

La comunidad más grande está formada por 228 artistas, siendo los más sonados entre ellos Lil Wayne, Drake, Future, Nicki Minaj, Ludacris,... Estos artistas además de ser los que han colaborado con un mayor número de artistas en esta época también son los que más grado tienen en la red global vista antes.



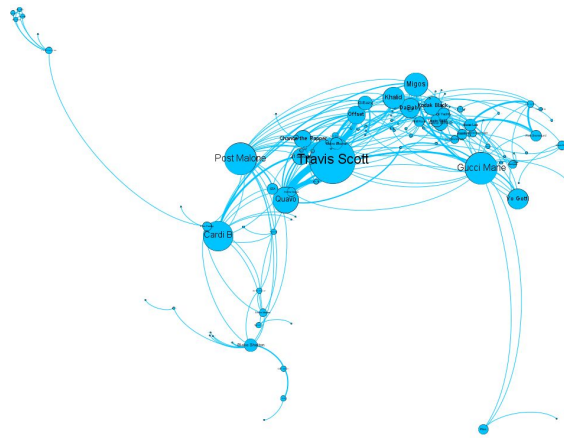
Además de ser la comunidad con más nodos, esta comunidad está relacionada de manera más fuerte que la media, con un 5.76 de artistas colaborados de media y 8.76 colaboraciones por cada artista. Como hemos dicho anteriormente, estos artistas hacen música perteneciente al género Hip-Hop.

La segunda comunidad más grande tiene 121 artistas que han colaborado entre ellos y tiene artistas como Rihanna, Beyoncé, Mariah Carey, Miley Cyrus,...



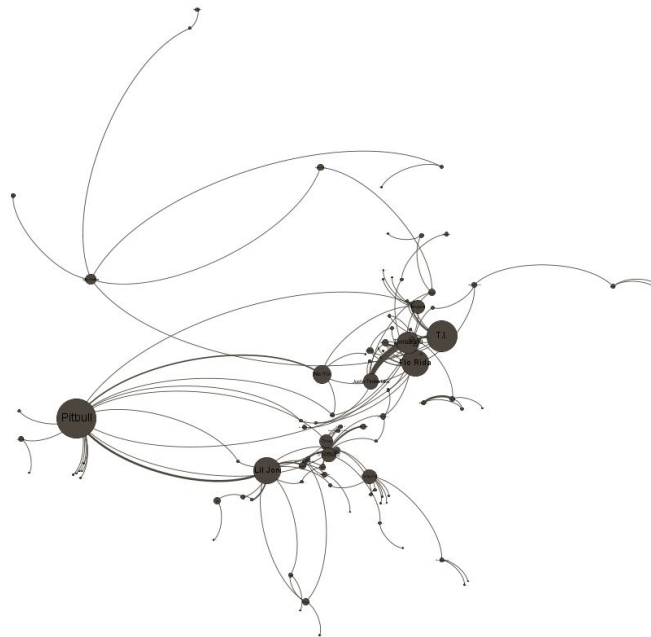
La comunidad tiene ligeramente menor grado medio que la red entera (4.89 colaboradores de media por artista y menor número de colaboraciones (5.19 de media). Estas artistas pertenecen mayoritariamente al género del pop o al EDM.

La tercera comunidad tiene 107 artistas, entre ellos Travis Scott, Post Malone, Migos, Cardi B, Khalid ... Aunque esta comunidad tiene un grado bastante menor al global de la red (3.82 colaboradores de media) el número de colaboraciones por artista no disminuye tanto llegando a 5.533, lo que nos podría indicar que colaboran más veces entre los mismos artistas.



Al igual que la primera comunidad, estos artistas hacen música relacionada con el Hip-Hop, aunque la mayoría se podría clasificar dentro del género trap, un subgénero del hip-hop.

La cuarta comunidad está formada por 104 artistas como Pitbull, Flo Rida, Lil Jon, T.I... Esta comunidad ya nos aparecía en la red global del principio y no ha cambiado demasiado al respecto.



La red sigue teniendo menor número de artistas colaboradores(2.654 de media) y solo 3 colaboraciones por artista.

Si analizamos los grados de los artistas con más peso en la red(Lil Wayne, Drake, Kanye West, Chris Brown) podemos ver que la gran mayoría de las colaboraciones en las que han participado pertenecen a esta época(2000-2019): de las 189 colaboraciones que Lil Wayne tiene en *Hot100* a lo largo de su carrera 141 forman parte de este grafo, de las 148 totales que tiene Drake 108 son de ahora o la totalidad de las de Kanye West (100 colaboraciones) que solo ha llegado al *Hot100* en este momento.

5. Estudio de los datos procedentes de Spotify

5.1 Estudio global de los datos de Spotify

Habiendo procesado el dataframe de Spotify global y creado la red conseguimos un total de 1081 nodos lo que significan 1081 artistas distintos con alguna colaboración en los charts de Spotify y 2721 aristas. En este dataframe hay canciones repetidas (es decir que llevan más de una semana en la lista de canciones más populares) así que el peso de las aristas que unen los nodos se verá altamente incrementado (así como el grado medio con pesos de cada nodo).

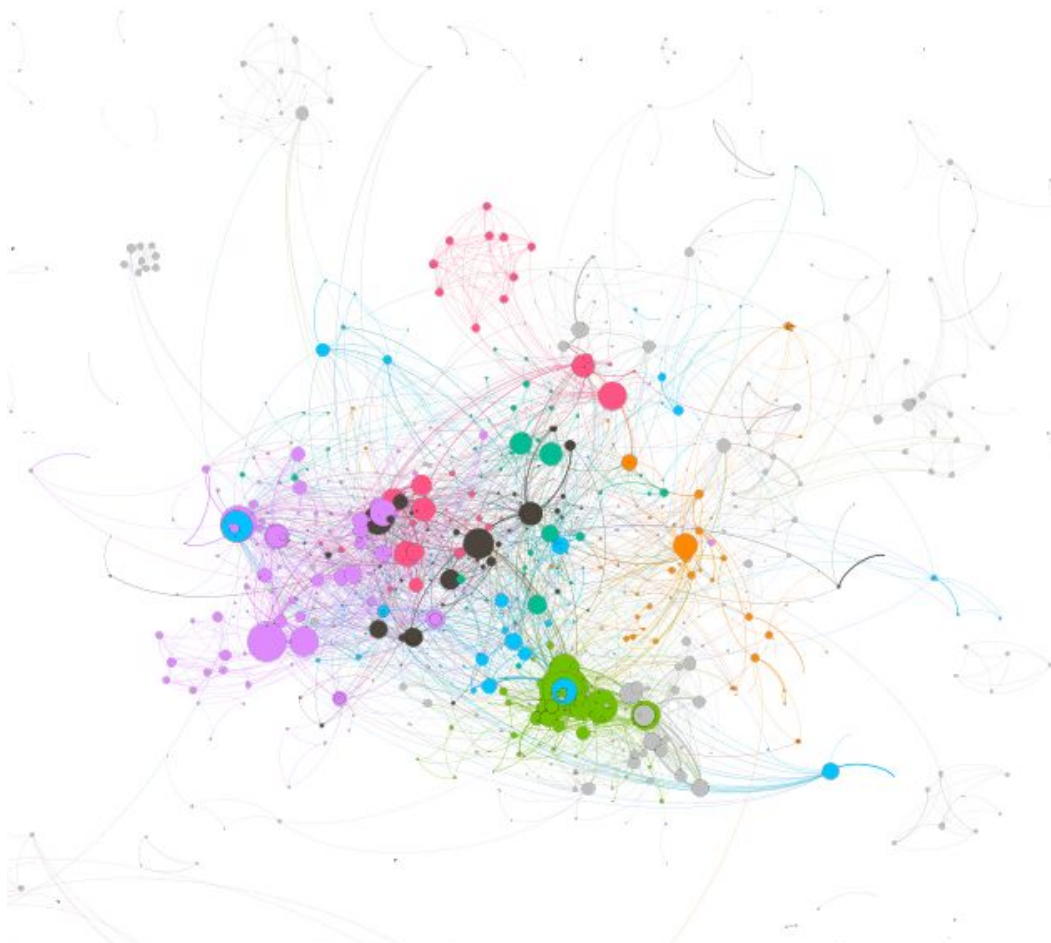
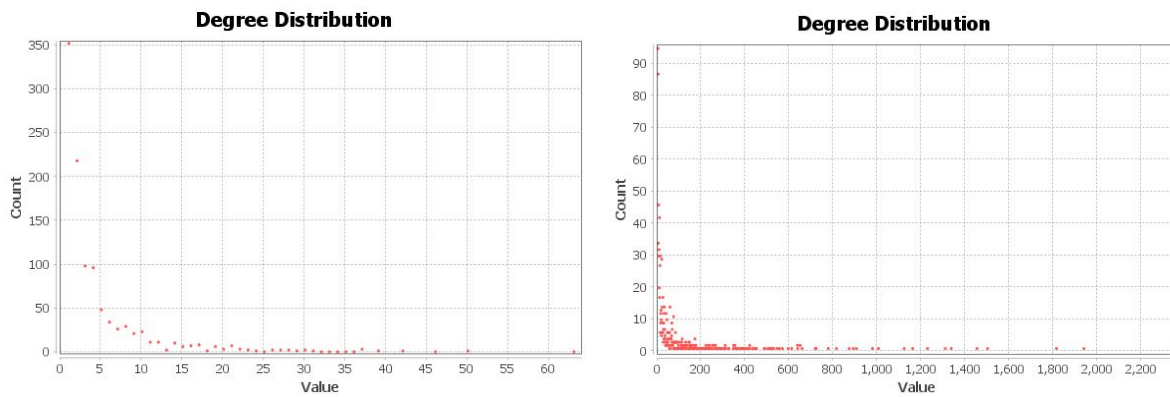


Imagen de la red con el layout OpenOrd

Los artistas tienen una media de colaboradores de 5.034 pero una media de colaboraciones muy alta, de 81.52 por artista. Esto se debe a lo que hemos comentado anteriormente sobre la repetición de canciones en las listas.

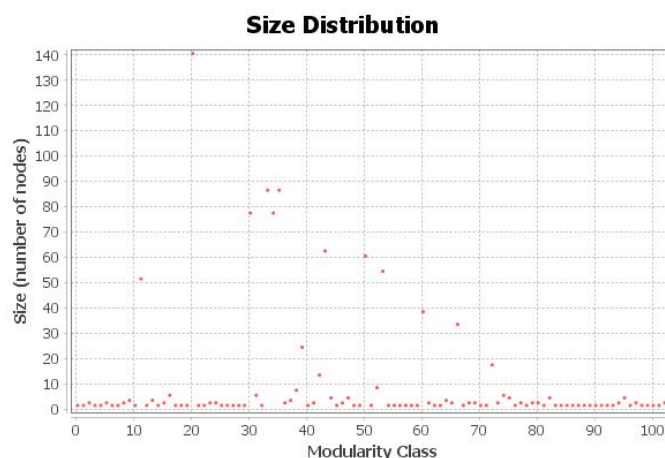


Podemos ver que la distribución de grado sin pesos(izq) se asemeja a la de con pesos(der) solo que esta última desplazada a la izquierda por los grandes valores de los nodos. Las dos siguen una distribución exponencial.

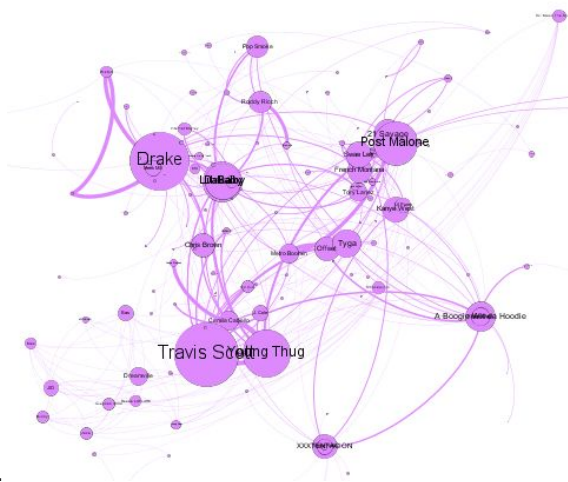
El artista con más grado es J Balvin habiendo colaborado con 63 artistas y el segundo Bad Bunny habiendo colaborado con 50. Cuando tenemos en cuenta los pesos de las aristas J Balvin sigue siendo el primero con un peso de 2360 debido ,entre otras cosas, a la arista que comparte con Bad Bunny con 288 de peso ya que los dos son artistas que suelen colaborar y cuando colaboran las canciones que hacen suelen estar entre las más escuchadas de spotify duramente semanas.

Como podemos observar, estos artistas pertenecen al género de música latina o reggaeton, posteriormente vamos a dividir la red en distintas comunidades para ver si los distintos géneros musicales varían en tamaño o en grado.

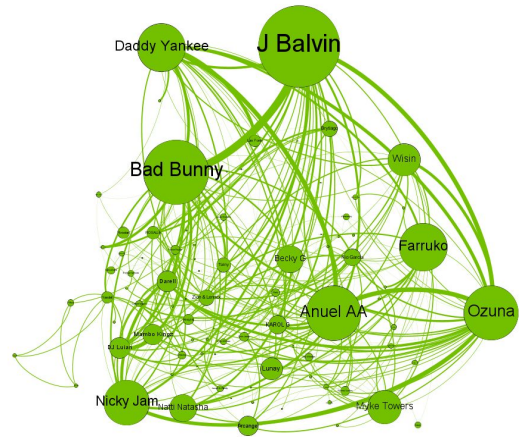
Gephi nos divide la red en 104 comunidades, pero como podemos ver en la distribución la gran mayoría de ellas no superan los 10 nodos de tamaño, solo superando ese umbral 14 comunidades.



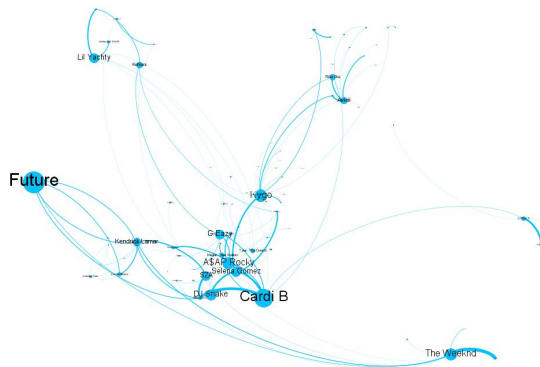
Las 5 comunidades más grandes superan los 75 nodos de tamaño, por lo que vamos a estudiarlas para ver si pertenecen a algún género específico



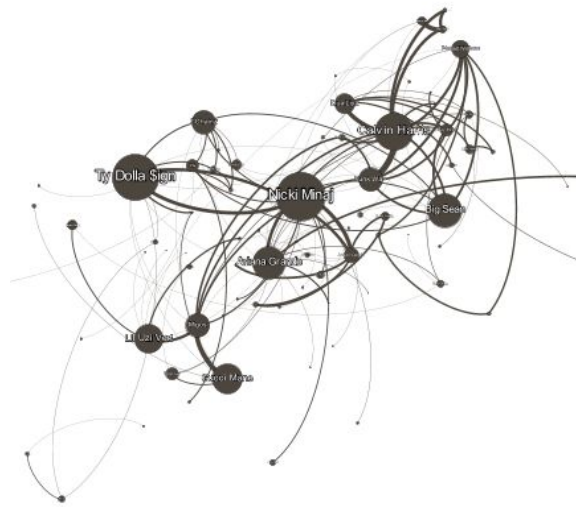
1.



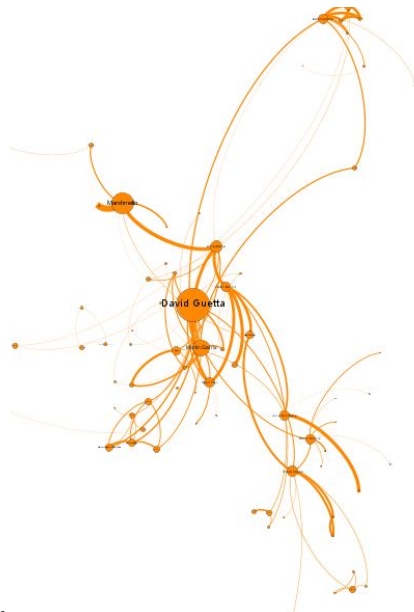
2.



3.



4.

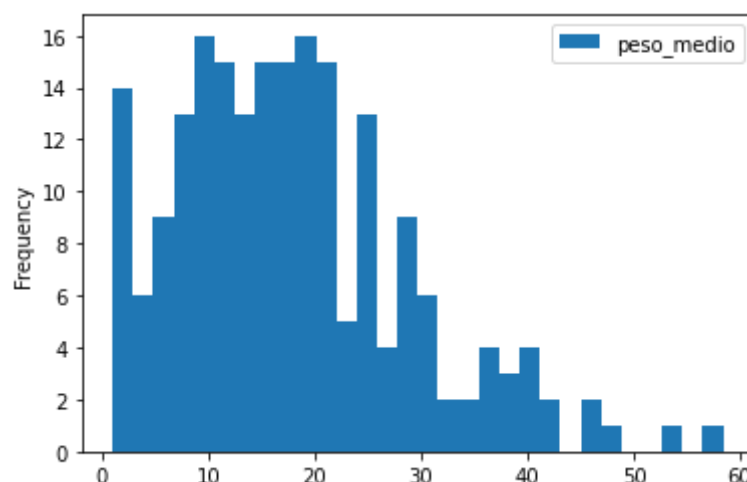


5.

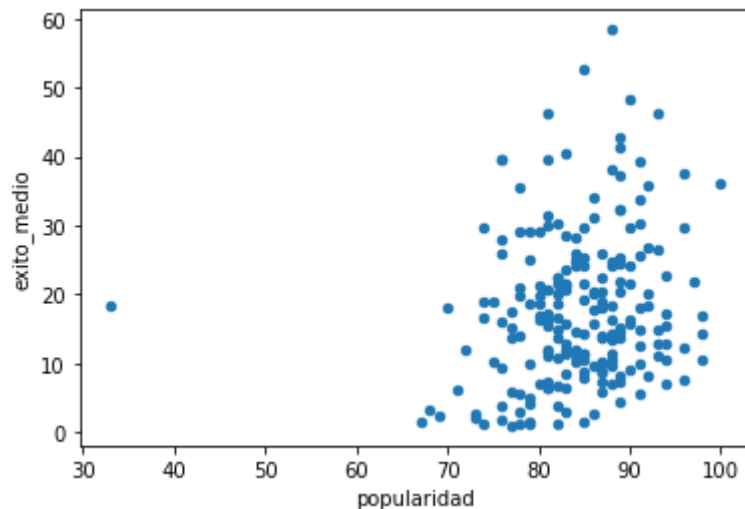
Podemos ver que la primera, tercera y cuarta comunidad más grande, que representan un 28.31% de todos los datos, pertenecen al mismo género, el hip-hop. La segunda comunidad, que representa el 8.05% de la red está formada por los nodos con mayor peso, J Balvin, Bad Bunny, Ozuna, etc. Esta comunidad representa al género reggaeton y podemos ver que aunque no colaboren con muchos artistas (colaboran de media con 6.23 artistas) sí que repiten muchas colaboraciones o las colaboraciones que realizan pasan mucho tiempo en las listas de éxitos (su peso medio es de 207.54). La última comunidad pertenece al género de música electrónica con artistas como David Guetta, Martin Garrix o Marshmello. Esta comunidad tiene un grado medio muy bajo, de 2.94 artistas colaboradores de media por artista y un peso medio de 54.97, que aunque no llega al nivel de la comunidad de reggaeton podemos ver que también en vez de colaborar con muchos artistas repiten colaboraciones.

Lo siguiente que hemos podido sacar es la media de los pesos de las aristas que tenga un nodo para poder ver cómo de exitosas son las colaboraciones de un artista. Para ver mejor esta media hemos cribado por los nodos que tengan más de tres colaboraciones diferentes con otros tres artistas diferentes.

A continuación tenemos la distribución de este éxito medio en forma de histograma donde tenemos un grueso en el rango hasta 30 y podemos ver unos pocos que destacan en la franja entre 50 y 60. Estos son Lenny Tavárez, cantante puertorriqueño, y Dalex, cantante estadounidense que también ha colaborado bastante con Lenny Tavárez, este dúo además se encuentra el segundo en la lista de pesos de aristas con un valor de 220.



Además, tenemos de cada artista su valor de popularidad, que es un valor asignado por Spotify que indica entre 0 y 100 cómo de popular es un artista y que se calcula a través de la popularidad de todas sus canciones. Hemos relacionado este dato con el que acabamos de ver de éxito medio



Aquí podemos ver cómo estos datos aparecen en la parte superior de la franja de popularidad y cabe destacar el punto que representa a Offset, el único que aparece con un índice de popularidad inferior a 65 con un valor de 33. En este caso Offset es más conocido como miembro de su grupo de trap Migos, por lo que se puede entender que haya entrado en este ranking a pesar de su bajo índice de popularidad.

5.2 Estudio regional de los datos de Spotify España

Habiendo procesado el dataframe de Spotify España y creado la red conseguimos un total de 986 nodos lo que significan 986 artistas distintos con alguna colaboración en los charts de Spotify y 3470 aristas, lo que ya nos indica que aunque haya menos artistas que colaboren respecto a las listas globales las colaboraciones en España triunfan más, ya que hay más colaboraciones en las listas de éxitos. Al igual que en el anterior dataframe hay canciones repetidas (es decir que llevan más de una semana en la lista de canciones más populares).

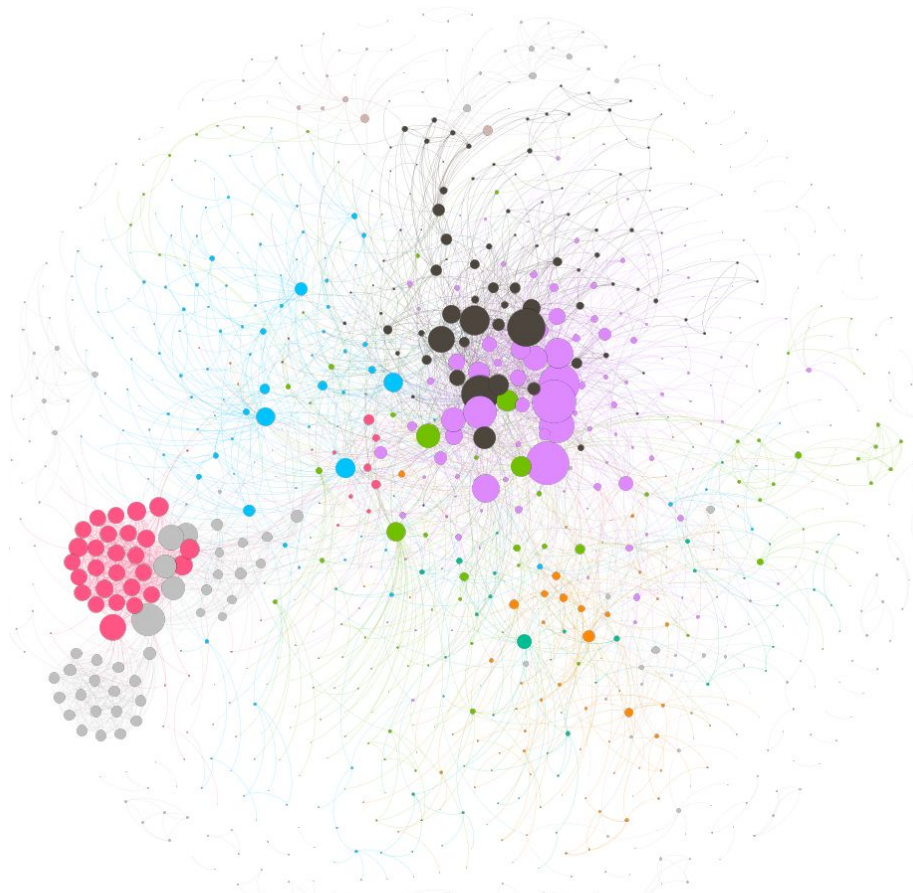
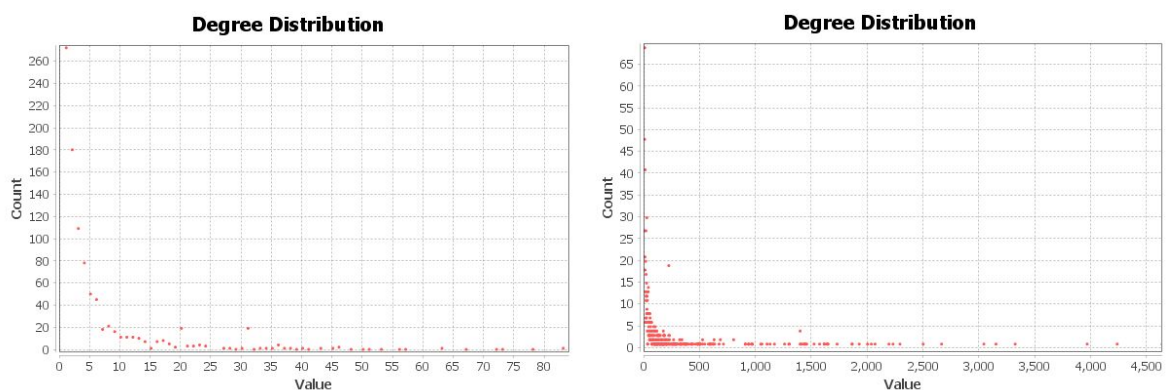


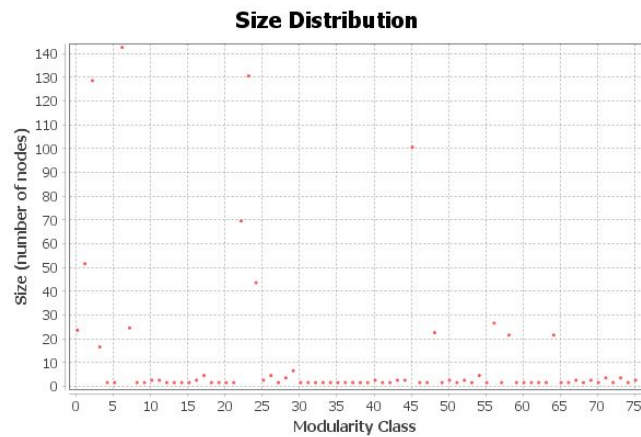
Imagen de la red con el layout Fruchterman Reingold

De media en esta red los artistas colaboran con otros 7.039 artistas y tienen de media 165.19 colaboraciones(peso). Como podemos ver en las gráficas de la distribución, siguen también una distribución exponencial. El artista con más grado medio vuelve a ser J Balvin, esta vez empatado con Bad Bunny en 83 artistas colaborados. En cambio, el artista con más peso en Spotify España es Ozuna con 4631 de peso. Este cantante comparte aristas con los otros con 238 de peso en el caso de Bad Bunny y 310 con J Balvin.

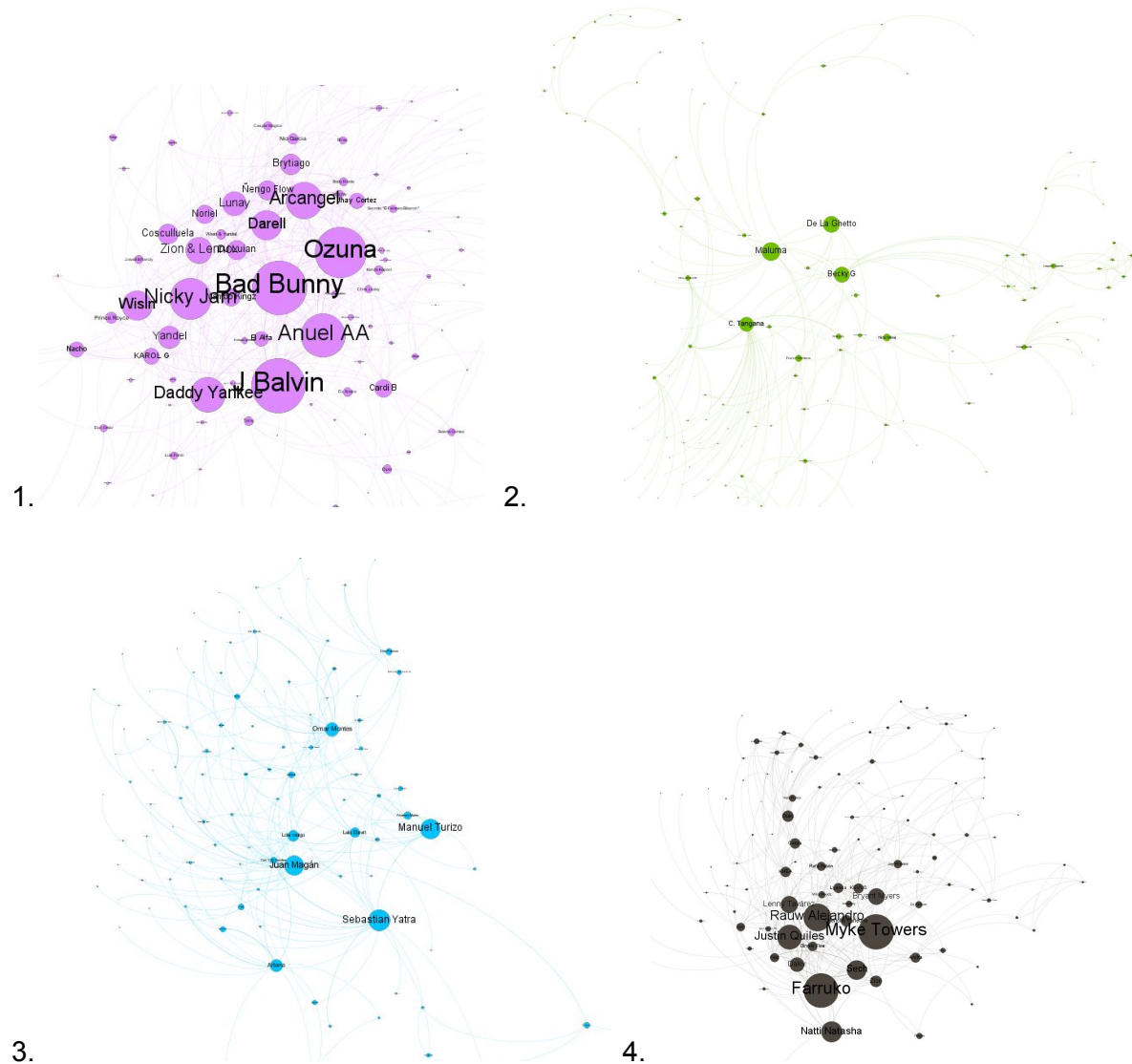


La red consta de 60 componentes conexas y Gephi la divide en 76 comunidades distintas. Al igual que en el resto de las redes la gran mayoría de las comunidades tienen menos de 10 nodos, concentrándose más del 75% de la red en las 10 comunidades más grandes. Las

4 comunidades más grandes llegan a más de 100 nodos: la primera más de 140 nodos, la segunda y tercera alrededor de 130 nodos y la cuarta un centenar de nodos.



Vamos a ver las 4 comunidades más grandes para ver si coinciden en alguna manera con las comunidades más grandes formadas en la red global de Spotify:

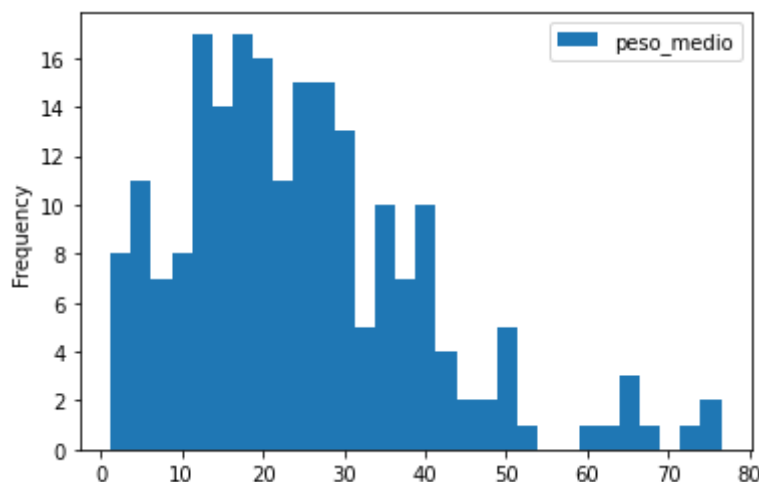


La única que se parece a las comunidades que nos encontrábamos en la red global es la primera comunidad, formada por cantantes de reggaeton como Bad Bunny, Daddy Yankee, J Balvin, etc. Esta comunidad formada por 143 nodos tiene ligeramente menor grado medio que la red completa, bajando a que cada artista de esta comunidad colabora de media con 6.79 artistas; sin embargo, el peso medio de los artistas aumenta hasta 278.

Las demás comunidades también están formadas en su mayoría por cantantes de reggaeton (Maluma, C. Tangana, Juan Magan, Myke Towers, Farruko,...) pero estos no aparecen (o si lo hacen no forman comunidades claramente separables) en la red global de Spotify.

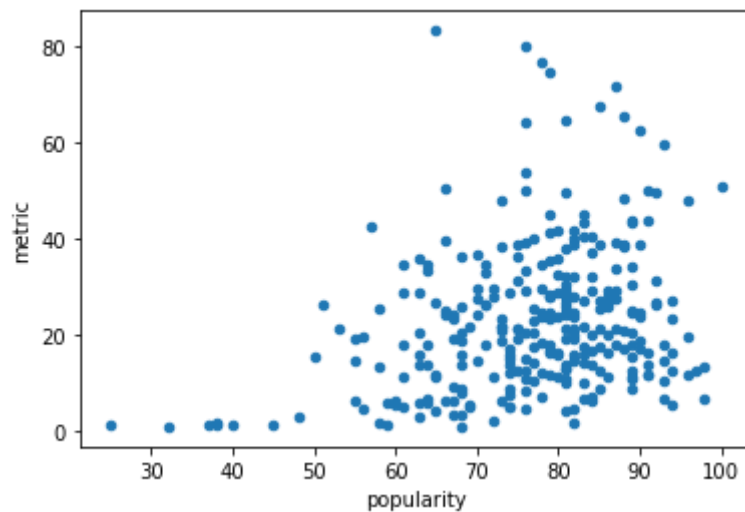
Quizás también cabe destacar la aparición de dos componentes bastante densas dentro del grafo debido a dos canciones que aglutinan a un gran número de artistas. Por un lado tenemos al pop español juntándose con Alejandro Sanz en una versión de “Y Si Fuera Ella?” donde colaboran 17 artistas diferentes con motivo de su 20º aniversario y con fines benéficos, por otro lado vemos cómo el indie-pop español hacía algo parecido en 2020 cuando más de 20 artistas trabajaron juntos en “Los Abrazos Prohibidos” y destinaban todos los beneficios a la investigación sobre el coronavirus SARS-CoV-2.

Aquí también podemos calcular la media de los pesos de las aristas para los nodos.



En este caso tenemos a tres posiciones por encima de 70 que son Juanes, Pablo Alborán y Shakira, que son artistas que funcionan muy bien en nuestro país.

Volviendo a ver este valor contrastado con la popularidad de un artista nos sale la siguiente distribución



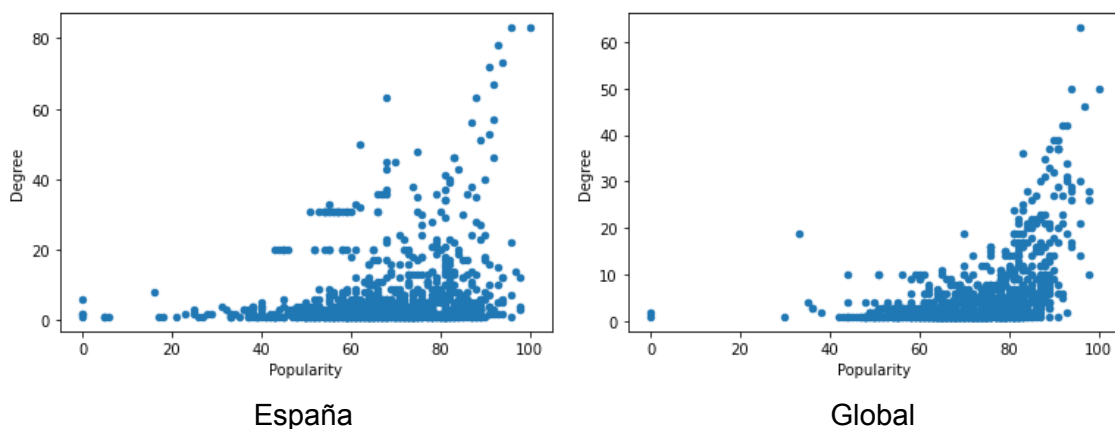
De aquí cabe destacar unos cuantos puntos por debajo de la franja de los 50, aquí tenemos a concursantes del talent show Operación Triunfo, cuyos concursantes interpretaban en grupo canciones en el programa cada semana y en alguna ocasión esta interpretación llegaba a las listas de Spotify España.

5.3 Comparaciones

En la red española el peso medio de los artistas es más del doble que el peso medio de los artistas en la red global. Esto puede ser por dos motivos: las colaboraciones triunfan más en el mercado español y pasan más tiempo en las listas de éxitos o hay más colaboraciones en el mercado español.

Estos dos motivos pueden deberse al tipo de música que triunfa entre los oyentes españoles, mayoritariamente géneros en torno al reggaeton. En este género es normal que haya al menos dos artistas por canción por lo que explicaría el gran número de colaboraciones y al ser el género que más triunfa en España explicaría que pasasen tanto tiempo en las listas de éxitos.

También hemos calculado en ambas listas la distribución de la popularidad de un artista con su grado, el número de artistas con los que ha colaborado.



Aquí podemos ver cómo en ambas imágenes aparece que los artistas que colaboran con más artistas diferentes son los más populares, pero también se ve que en la lista de España hay un mayor espacio para que artistas no tan populares tengan colaboraciones exitosas.

6.Conclusiones

Para concluir, hemos podido comprobar el gran aumento que ha habido en el número de colaboraciones en la industria musical desde 1958 hasta la actualidad.

Este aumento ha sido provocado entre otras razones por el cambio de popularidad de los géneros musicales, además de la aparición de dos de ellos principalmente: el hip-hop y el reggaeton.

Tras la popularización del hip-hop a finales de siglo se puede ver como los artistas pertenecientes a este género son los que más colaboraciones tienen, siendo el claro ejemplo Lil Wayne con 141 colaboraciones en las listas de éxitos en los últimos 20 años. Cada vez vemos con mayor frecuencia estas colaboraciones en la música más popular y hemos comprobado cómo los artistas que más colaboran entre ellos acaban siendo los más populares del panorama musical y esto hace que haya ciertos artistas que siempre aparezcan en las listas de canciones más escuchadas ya sea porque una colaboración con uno o con otro ha conseguido el éxito.