

# Análisis de Redes Sociales

Curso 2018 – 2019

## Análisis de una red con Gephi

Práctica 1

Grupo 1

Javier Sesé García

Leila Ruiz Casanova

Víctor del Pino Castilla

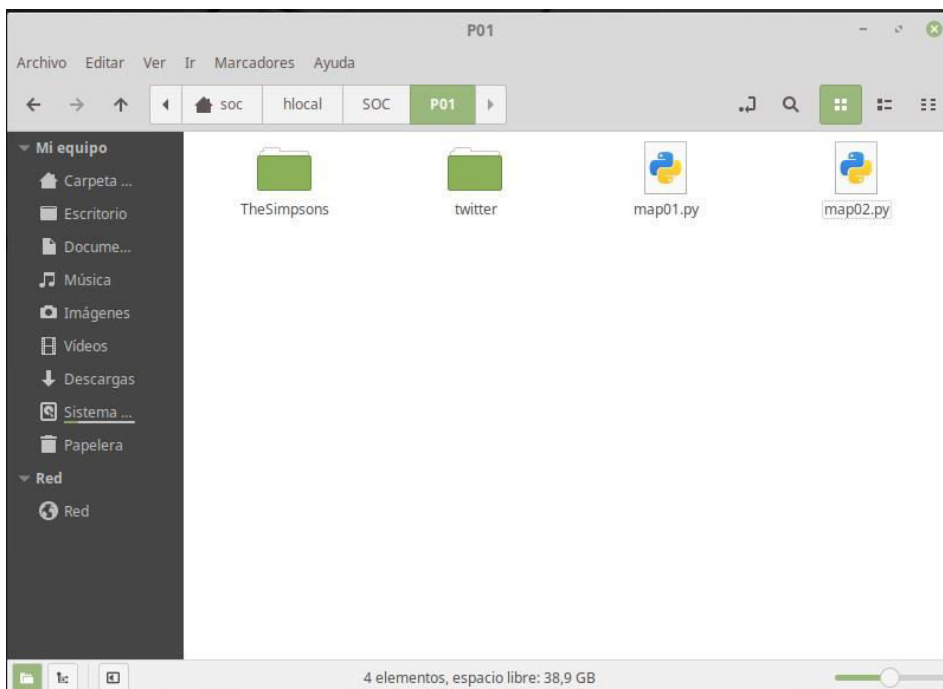
## 1. Introducción

Hemos elegido analizar las redes de Twitter en Reino Unido y España, pero una vez que hicimos los análisis de ambos países se nos ocurrió investigar las relaciones entre los diferentes países centrándonos en cómo es la interacción entre las diferentes redes de los países y los nodos que relacionan países.

## 2. Generación de la red

Para obtener las redes hemos creado dos scripts en Python `map01.py` y `map02.py`, en el primero generamos los ficheros con la información de los nodos de las redes, en formato csv, todos los ficheros tienen la nomenclatura `usuariosXXX.csv` y en el segundo script generamos la información de los enlaces entre nodos, en formato csv, todos los ficheros tienen la nomenclatura `seguidoresXXX.csv`.

Para funcionar ambos scripts deben situarse al mismo nivel que la carpeta *twitter* que debe contener los ficheros con los datos en bruto originales, para su ejecución hay que ejecutar ambos scripts en orden consecutivo



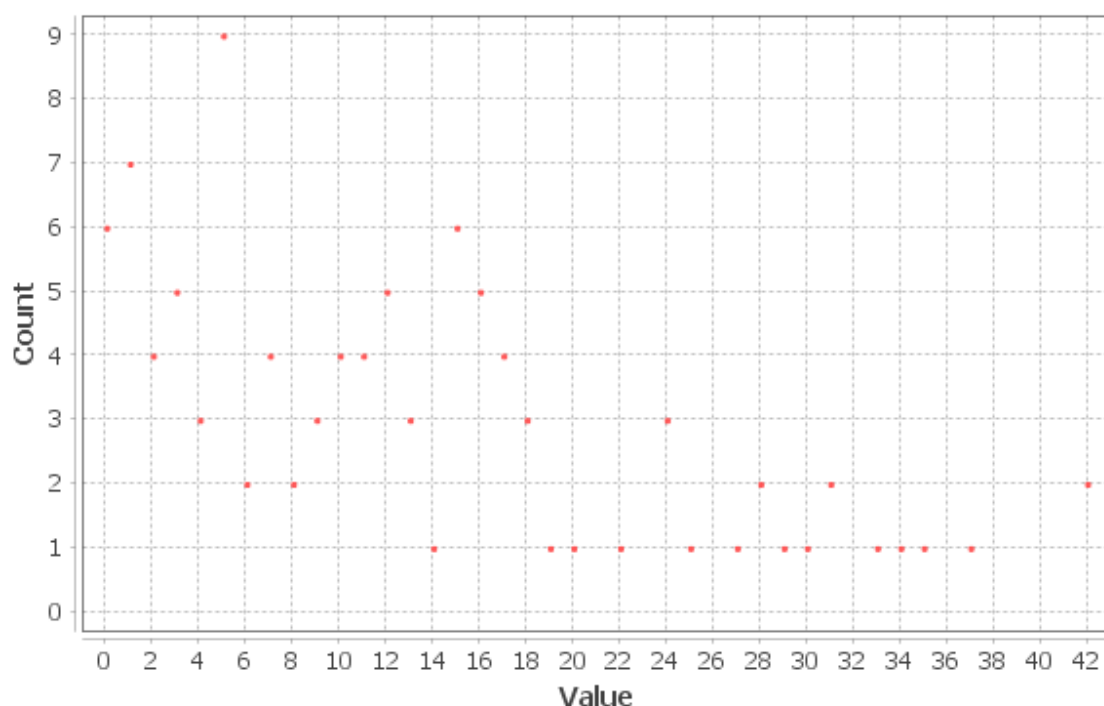
```
soc@soc-virtual-machine: ~/hlocal/SOC/P01
Archivo  Editar  Ver  Buscar  Terminal  Ayuda
soc@soc-virtual-machine:~/hlocal/SOC/P01$ ./map01.py
Ok CSV spain
Ok CSV uk
Ok read partial France
Ok read partial Germany
Ok read partial Global
Ok read partial Italy
Ok read partial UK
Ok read partial US
Ok read partial spain
Ok write global
Ok all
soc@soc-virtual-machine:~/hlocal/SOC/P01$ ./map02.py
Ok seguidoresSpain.csv
Ok seguidoresUK.csv
Ok France
Ok Germany
Ok Global
Ok Italy
Ok United_states
Ok Spain
Ok United_kingdom
Ok write global countries
Ok All
soc@soc-virtual-machine:~/hlocal/SOC/P01$
```

### 3. Reino Unido

#### Medidas Globales:

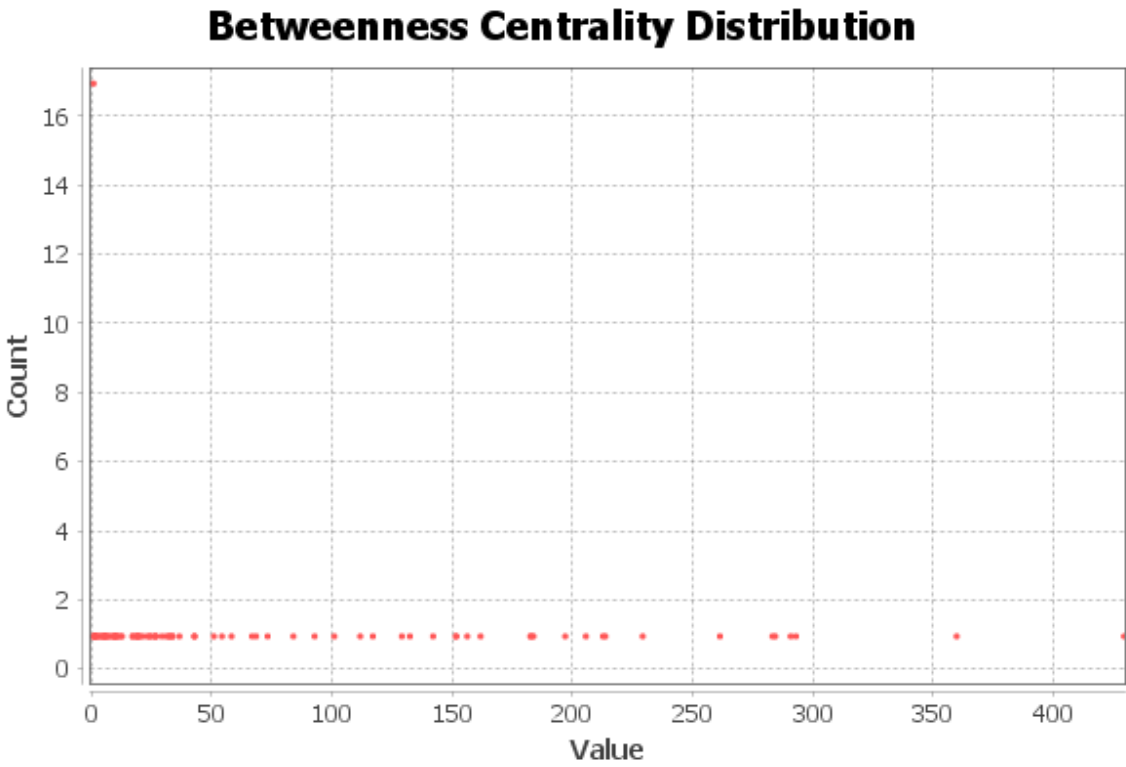
Nº de nodos: 100  
Nº Aristas: 622  
Densidad de la red: 0,126  
Grado medio: 12,44  
Diámetro: 7  
Distancia media: 2,469  
Coeficiente medio de clustering: 0,496

#### Degree Distribution

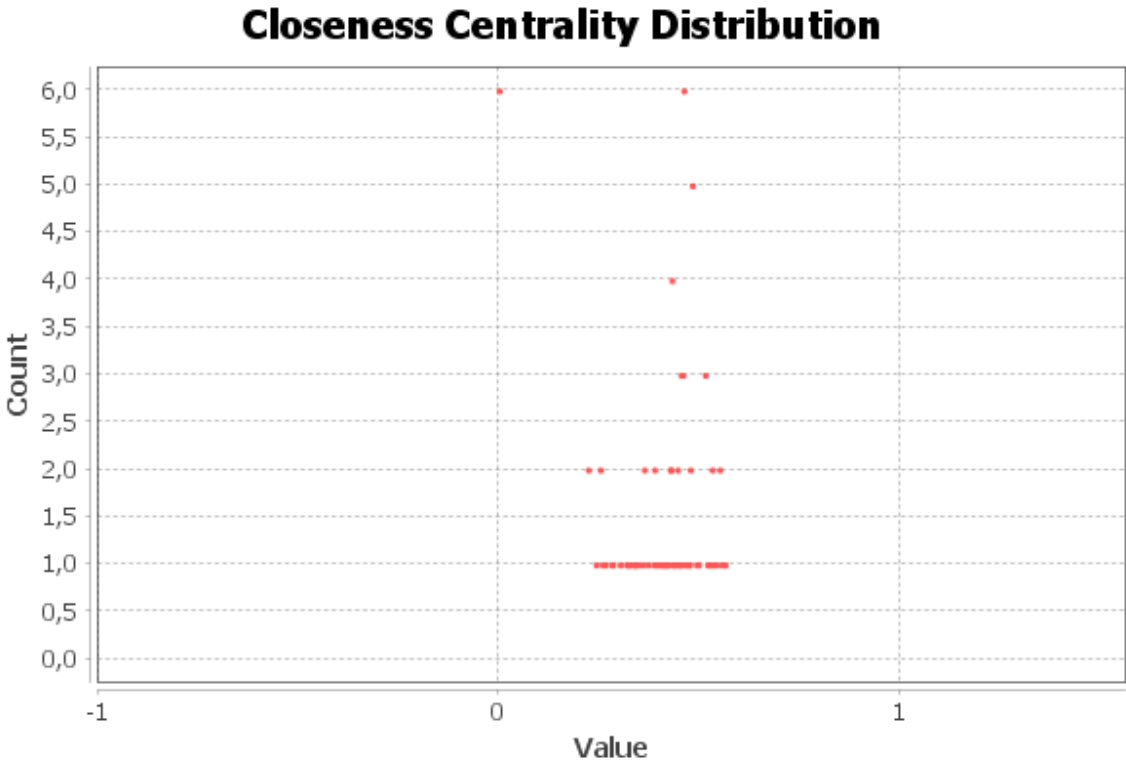


Medidas de centralidad:

Intermediación:

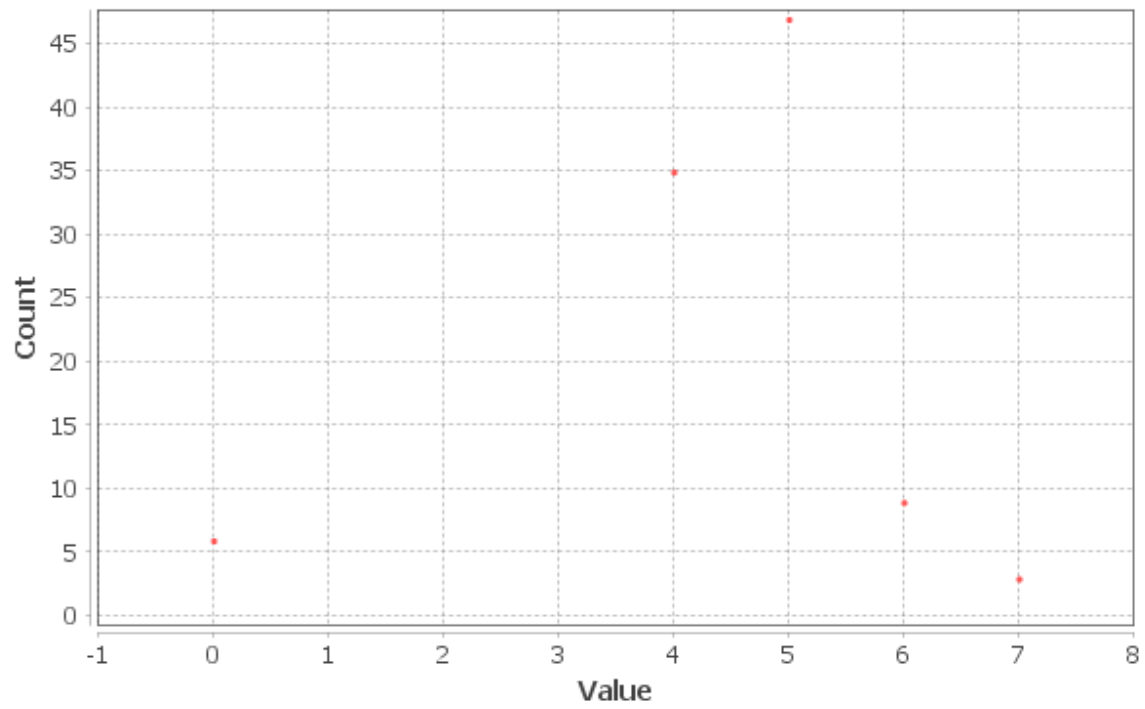


Cercanía:



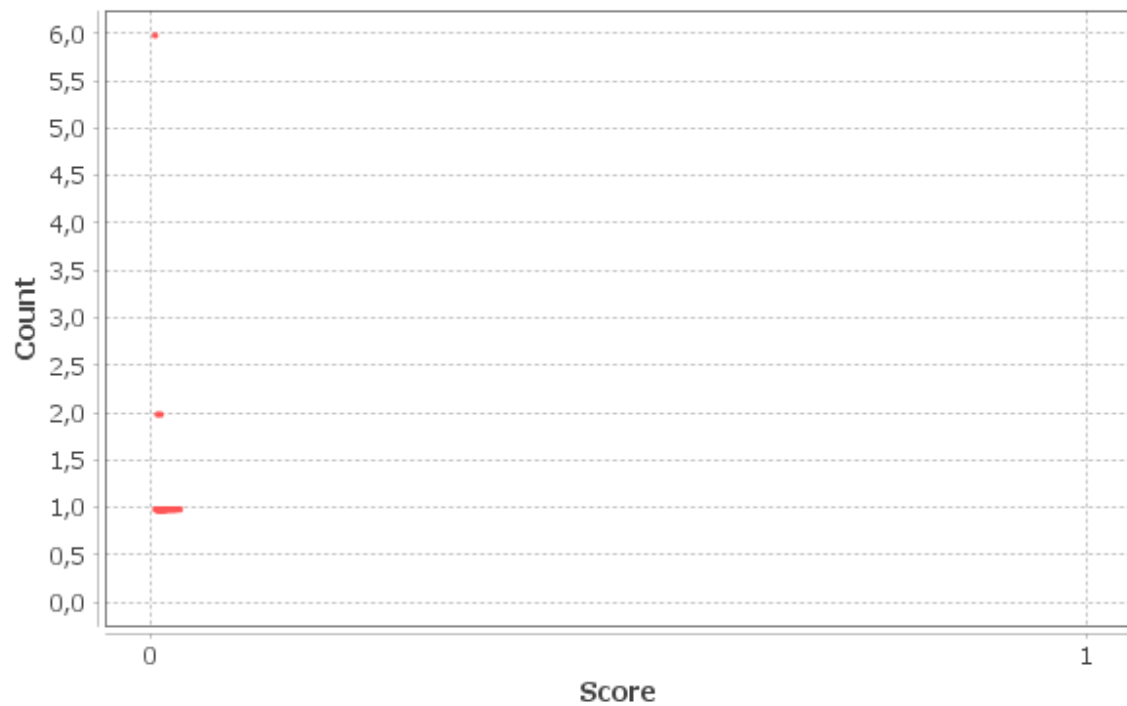
Excentricidad:

### Eccentricity Distribution



PageRank:

### PageRank Distribution

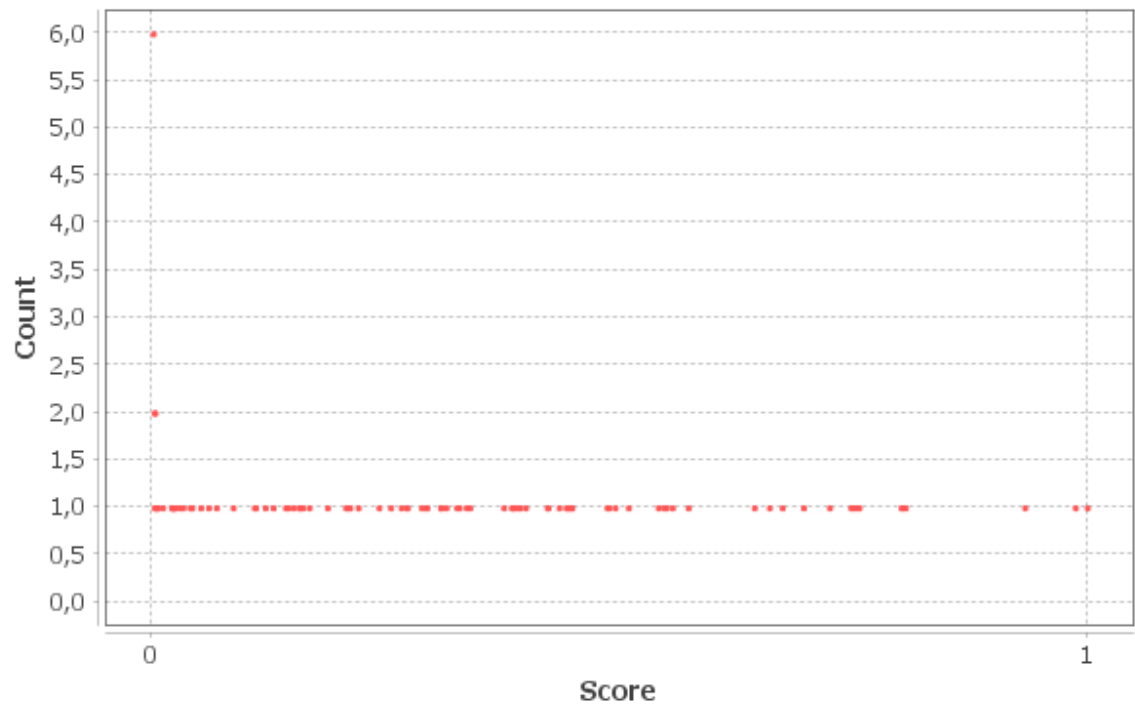


Epsilon = 0.001

Probability = 0.85

Centralidad de vector propio:

### Eigenvector Centrality Distribution



Nº iteraciones: 100

Sum change: 0.0011643116512642666

Conectividad:

Nº de componentes conexas: 7

Modularidad: 0,273

Nº de comunidades: 13

### Size Distribution



Resolution = 1.0

Randomize: ON

Edge weights: ON

Como podemos comprobar existen ocho comunidades de las cuales cuatro constan de cinco o menos miembros (Naranja, Verde Oscuro, Rojo y Marrón), dedicadas a el sector noticias, cultura y a los Beatles.

Destacan seis usuarios que no están conectados con ningún otro nodo, consultando los perfiles de estos usuarios en la red social se comprueba que o no siguen ningún usuario o solo siguen a miembros de su ámbito que no están en los datos de origen.

Atendiendo a las medidas del grafo se pueden identificar dos Hubs, los nodos @JKCorden y @wossy, lo que nos da un diámetro de la red de 7 pasos por lo que es una red medianamente conectada entre comunidades.

## 4. España

### Medidas Globales:

Nº de nodos: 100

Nº Aristas: 1562

Densidad de la red: 0,158

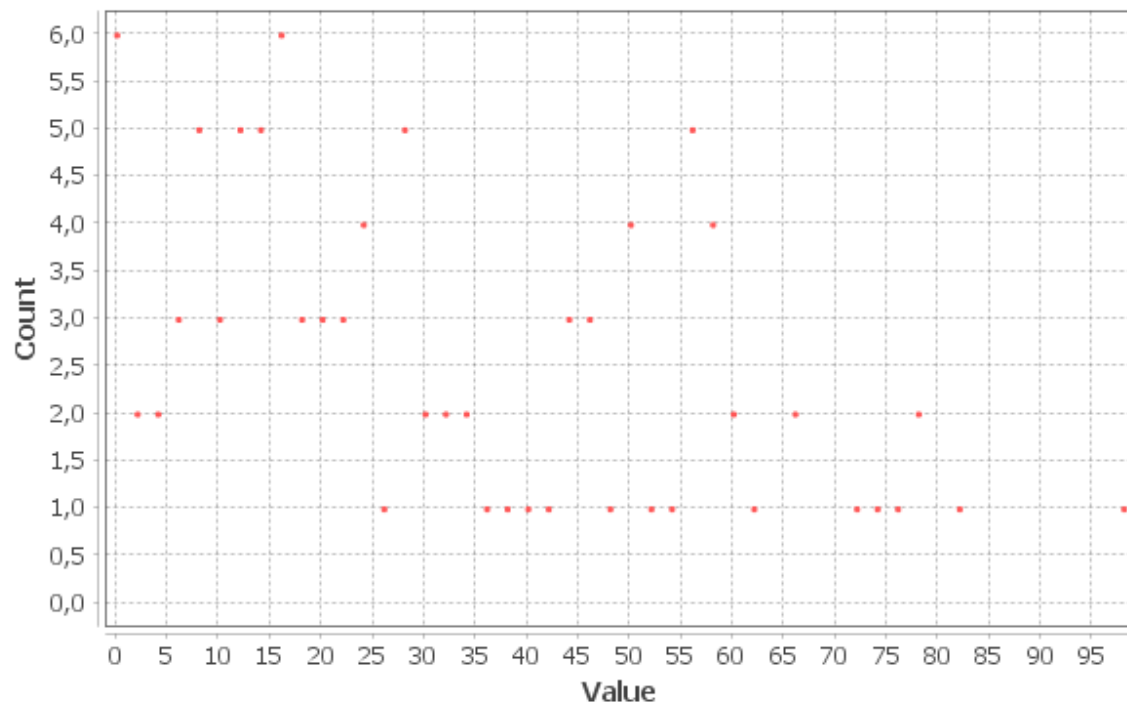
Grado medio: 15,62

Diámetro: 4

Distancia media: 2,104

Coefficiente medio de clustering: 0,546

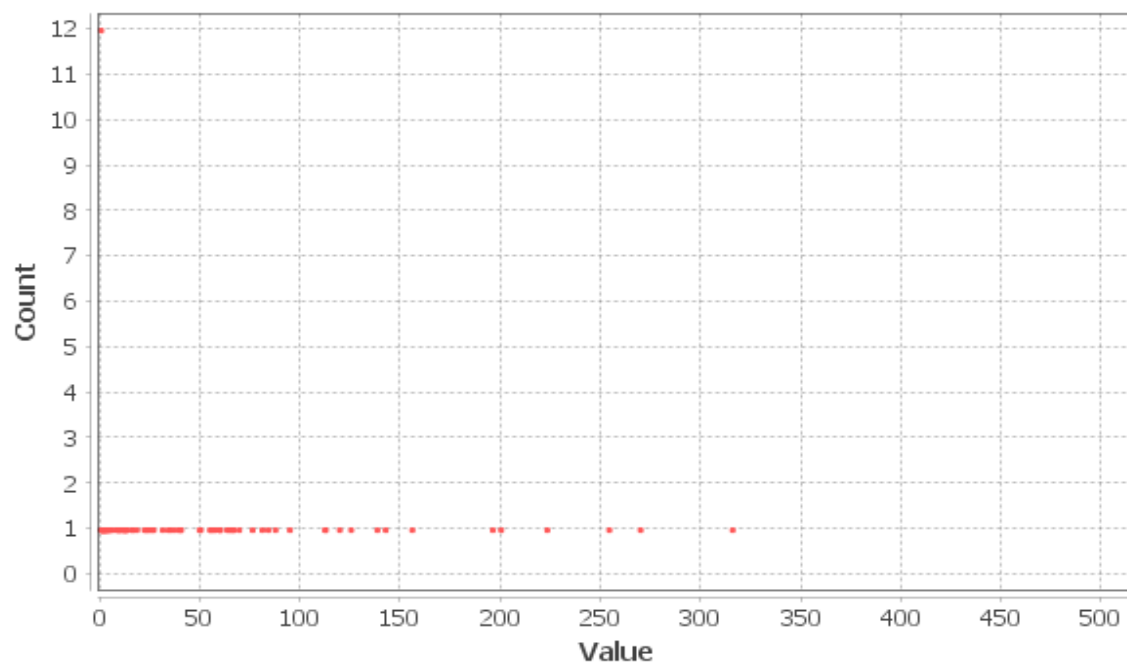
### Degree Distribution



### Medidas de centralidad:

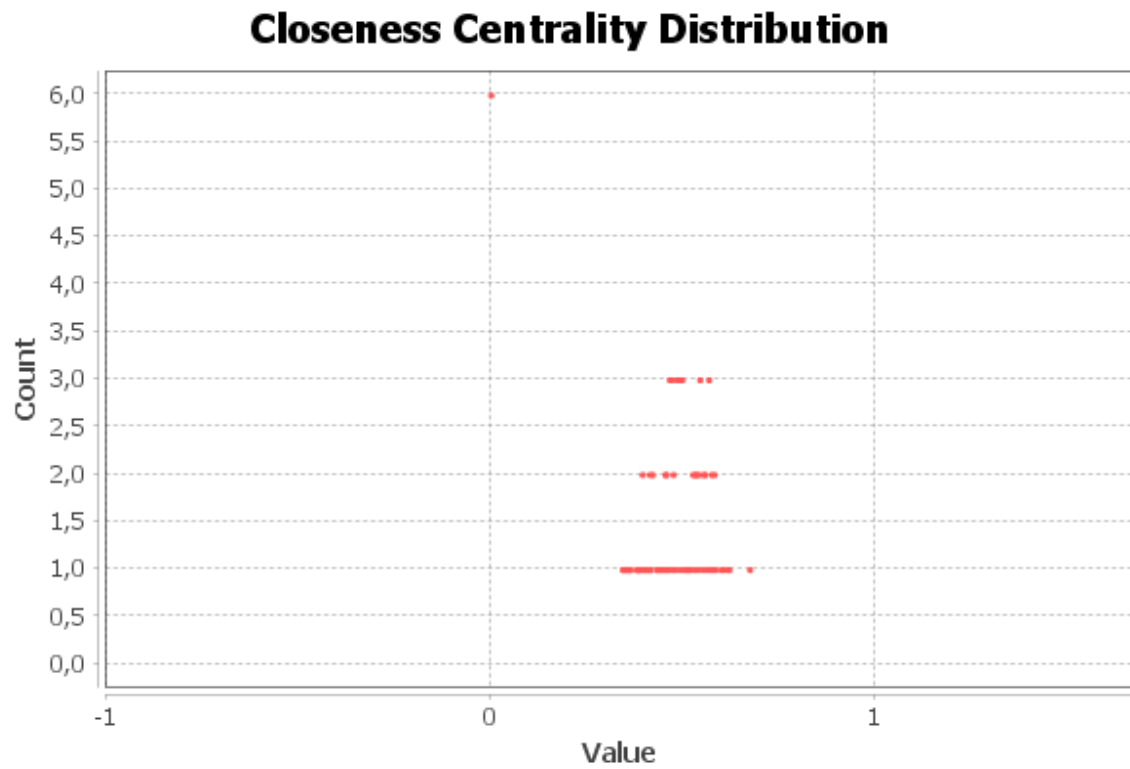
Intermediación:

### Betweenness Centrality Distribution

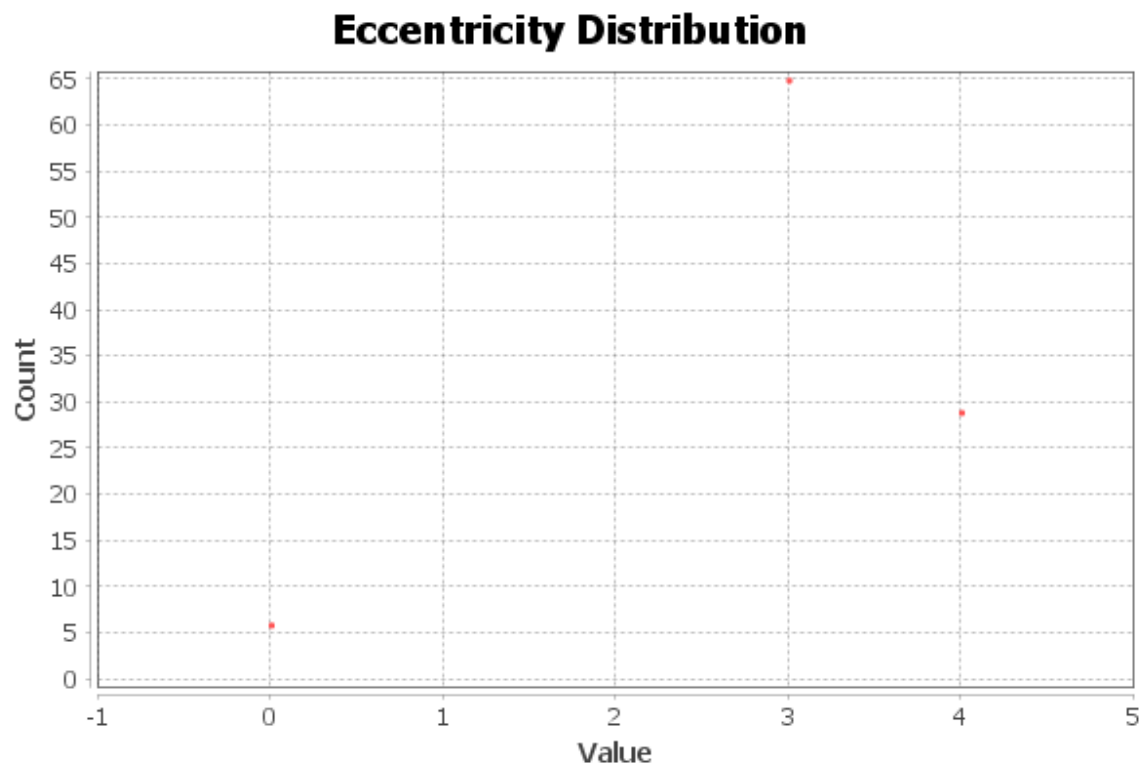




Cercanía:



Excentricidad:



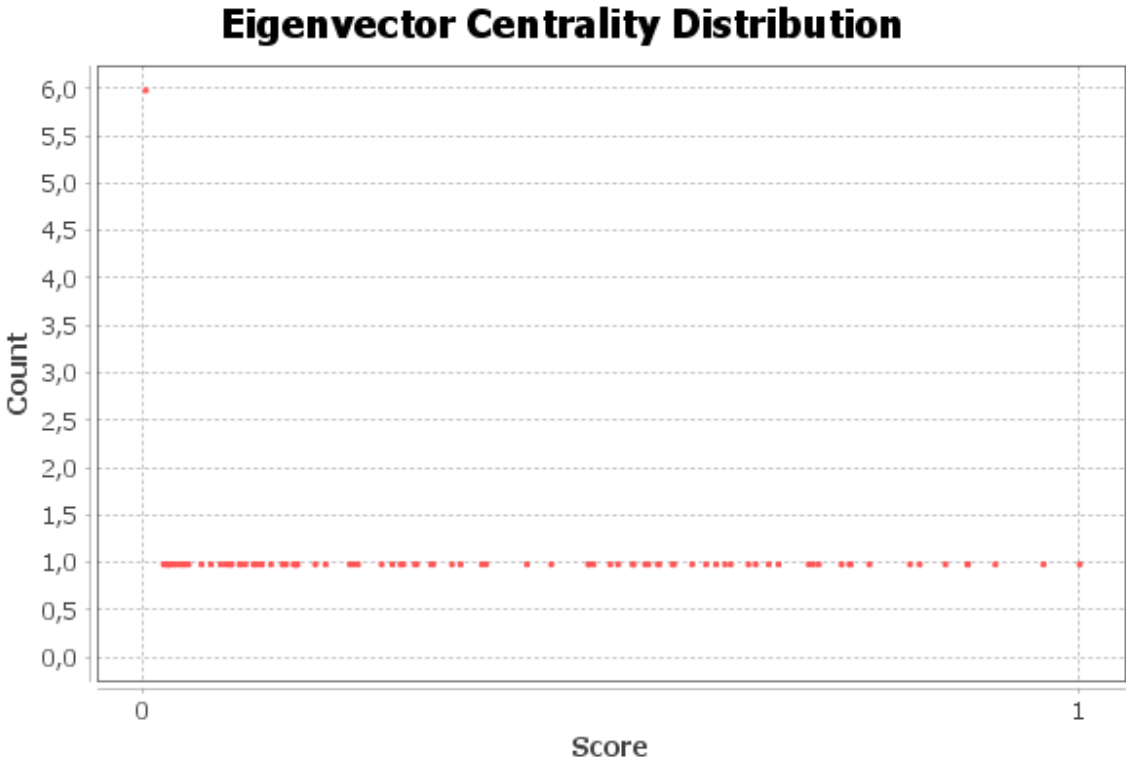
PageRank:



Epsilon = 0.001

Probability = 0.85

Centralidad de vector propio:



Nº Iteraciones = 100

Sum change: 0.003950658456527267

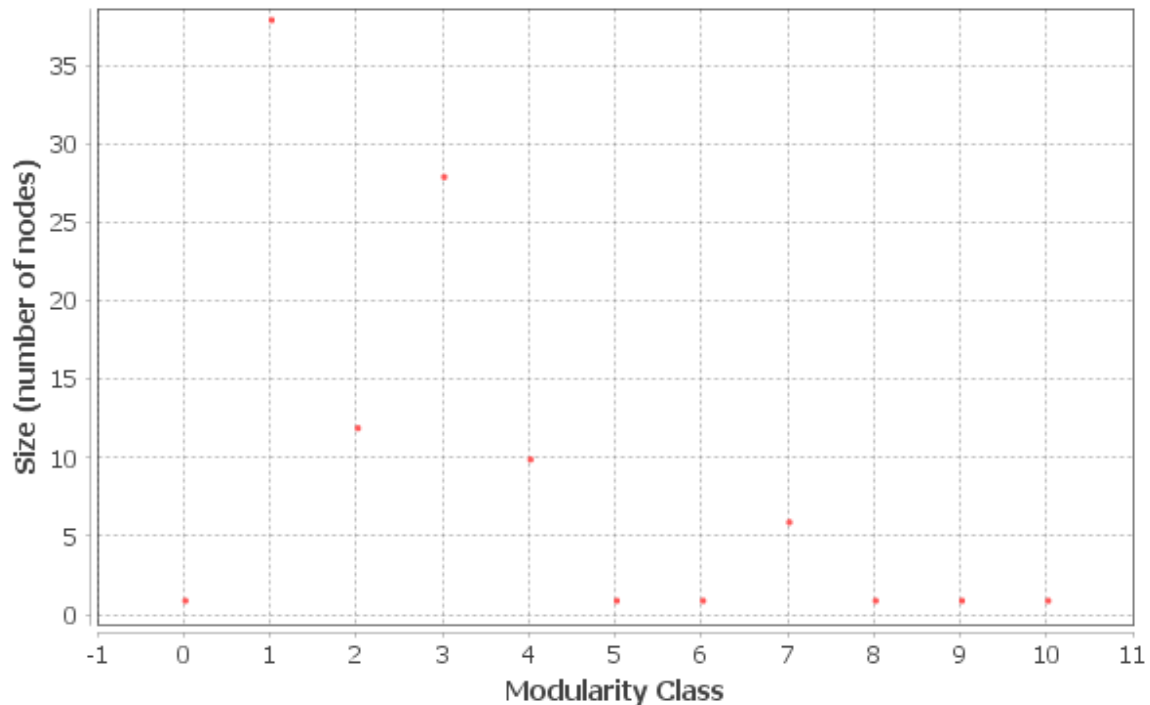
## Conectividad:

Nº de componentes conexas: 7

Modularidad: 0,292

Nº de comunidades: 11

## Size Distribution

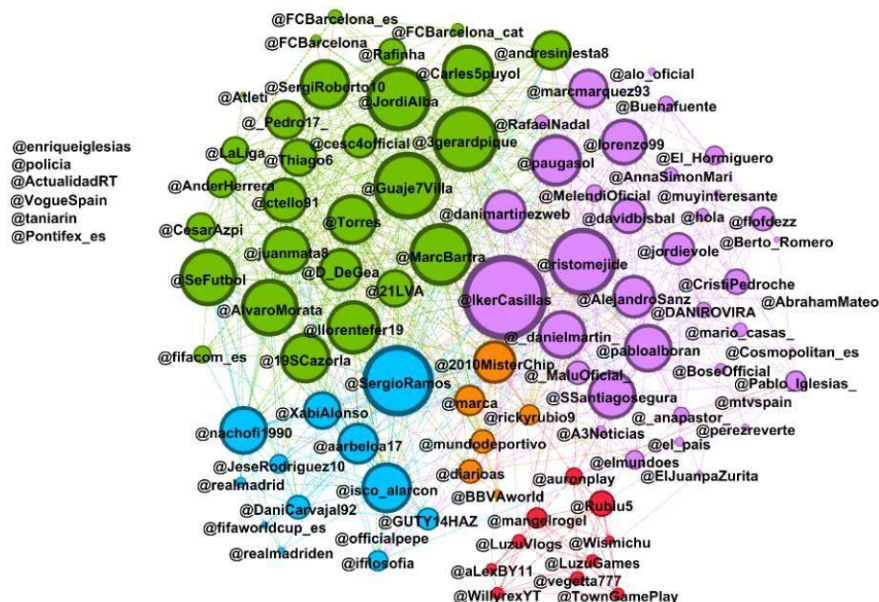


Resolution = 1.0

Randomize: ON

Edge weight: ON

## Discusión:



Hemos pintado con colores las diferentes comunidades encontradas y el tamaño según el número de seguidores de cada uno de los nodos.

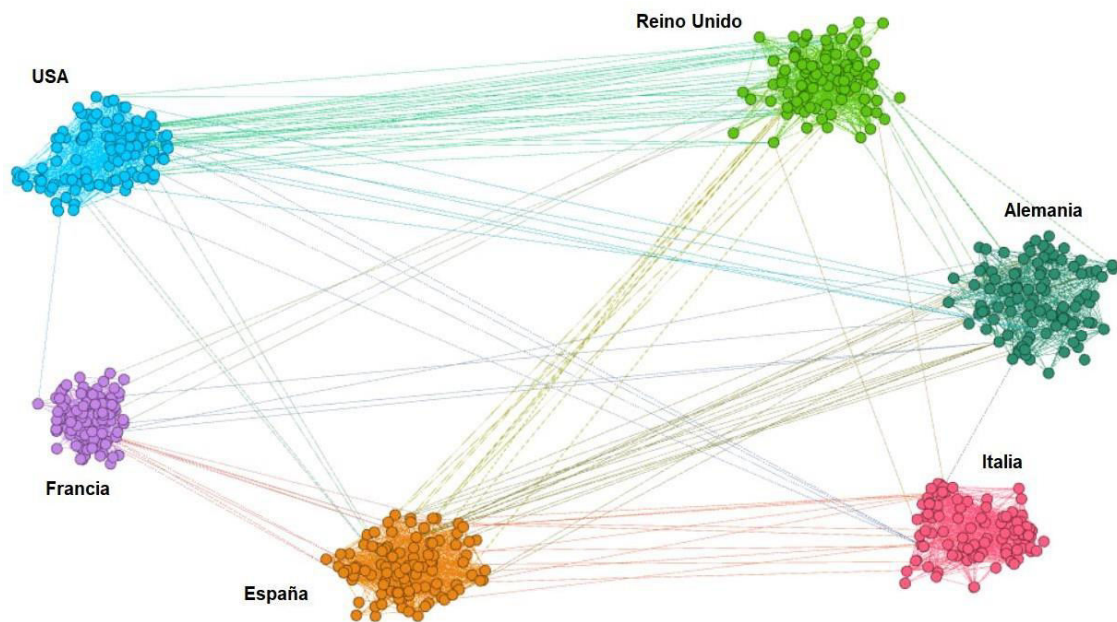
Como podemos comprobar existen tres comunidades predominantes, jugadores del Fútbol Club Barcelona (Verde), jugadores del Real Madrid (Azul) y la tercera comunidad que podemos englobarla como mundo del espectáculo (Morado), de estas tres comunidades destaca el hecho que el usuario @IkerCasillas se le engloba en el mundo del espectáculo aun siendo jugador del Real Madrid.

Adicionalmente tenemos dos comunidades más pequeñas una de prensa deportiva (Naranja) y Youtubers (Rojo).

Destacan seis usuarios que no están conectados con ningún otro nodo, consultando los perfiles de estos usuarios en la red social se comprueba que la cuenta @policia no sigue a ningún usuario en la red y el resto solo siguen a usuarios de comunidades no representadas en los datos de origen.

Atendiendo a las medidas del grafo se pueden identificar tres hubs, los nodos @IkerCasillas, @ristomejide y @SergioRamos lo que nos da un diametro de la red de 4 pasos por lo que es una red bastante conectada entre comunidades.

## 5. Global



Cuando analizamos los datos de todos los nodos de todos los países podemos comprobar como hay varios flujos de comunicación entre países de lenguas parecidas, España tiene bastantes enlaces en común con Italia y Alemania sobre todo por jugadores de futbol, pero en contraste Alemania no tiene casi enlaces con Italia. También destaca la comunicación entre los dos países de habla inglesa y sobre todo destaca los pocos enlaces que tiene Francia con el resto.

Para visualizar mejor estos datos hemos generado otra vista en la que establecemos el grosor de la arista en función de los enlaces entre los países que hemos agrupado en un solo nodo.

Con lo que confirmamos nuestras apreciaciones en los grafos anteriores

