



**ESCUELA POLITÉCNICA NACIONAL**  
**FACULTAD DE INGENIERÍA DE SISTEMAS**  
**INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN**

---

**Laboratorio de:** Inteligencia de negocios

**Práctica No.:** 6 de minería de datos

**Tema:** Análisis de redes sociales con Ghepi

**Nombre:** Danny Díaz

**Fecha:** 02/02/2020

**1. Objetivos:**

**1.1. Objetivo General**

*Explorar la herramienta Ghepi realizando un grafo*

**1.2. Objetivos Específicos**

*Graficar dos medidas de centralidad*

*Definir los conceptos de centralidad*

**2. Marco teórico:**

**Betweenness centrality**

Mide el número de conversaciones (o caminos más cortos entre nodos aleatorios) que potencialmente pueden pasar por un nodo de la red. Indica en cierta manera si ese nodo es un punto de distribución entre comunidades.

**Closeness centrality**

La centralidad de proximidad (o cercanía) de un nodo es una medida de centralidad en una red, calculada como el recíproco de la suma de la longitud de las rutas más cortas entre el nodo y todos los demás nodos en el gráfico. Por lo tanto, cuanto más central es un nodo, más cerca está de todos los demás nodos. Cuando un grafo no está fuertemente conectado, una idea generalizada es la de usar la suma de recíprocos de distancias, en lugar del recíproco de la suma de distancias, con la convención  $1 / \infty = 0$

$$H(x) = \sum_{y \neq x} \frac{1}{d(y, x)}.$$

## Gephi

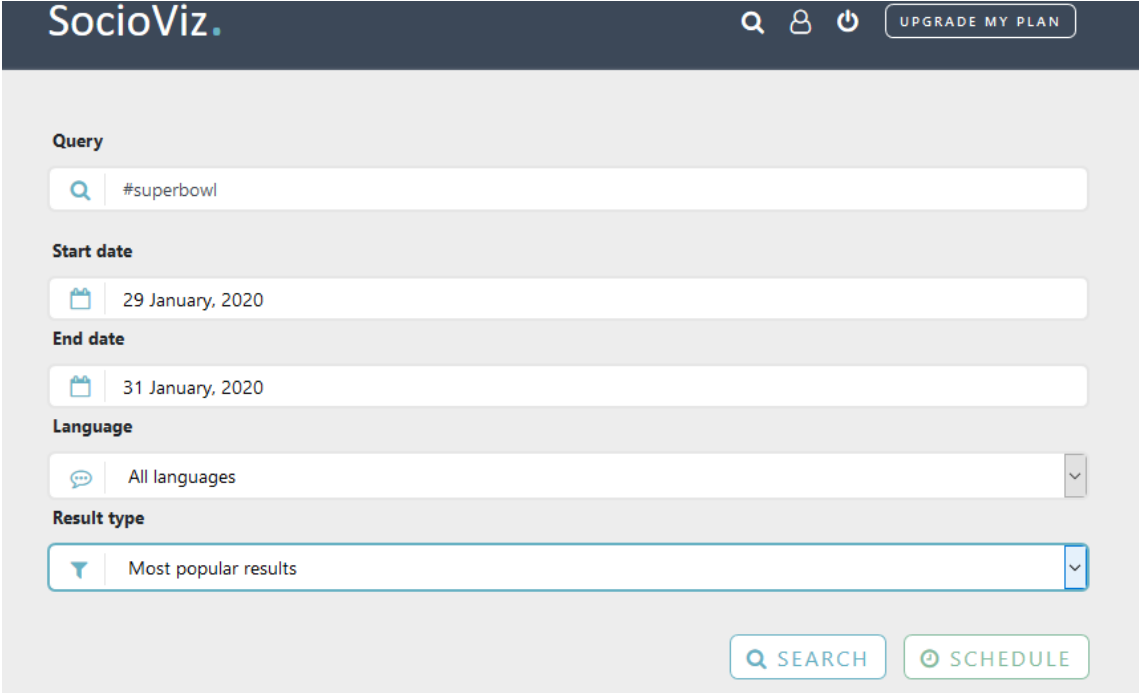
Es una herramienta open-source desarrollada en Java para visualizar y analizar grandes gráficos de red.

Usa un motor de renderizado 3D para mostrar gráficos en tiempo real y permite explorar, analizar, filtrar, clusterizar, manipular y exportar diversos tipos de gráficos.

### 3. Desarrollo de la práctica:

#### 3.1. Obtener datos para Gephi a partir de Twitter

El primer paso es registrarse en la API de SocioViz para realizar la consulta. En este caso se ingresa el hashtag de #superbowl con el rango de fechas máximo que nos permite la aplicación



The screenshot displays the SocioViz API interface. At the top, the header includes the 'SocioViz.' logo, navigation icons (search, user, power), and an 'UPGRADE MY PLAN' button. The main form area contains several sections: 'Query' with a search bar containing '#superbowl'; 'Start date' with a calendar icon and the date '29 January, 2020'; 'End date' with a calendar icon and the date '31 January, 2020'; 'Language' with a dropdown menu set to 'All languages'; and 'Result type' with a dropdown menu set to 'Most popular results'. At the bottom right, there are two buttons: 'SEARCH' and 'SCHEDULE'.

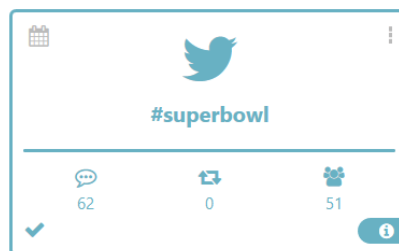


**ESCUELA POLITÉCNICA NACIONAL**  
**FACULTAD DE INGENIERÍA DE SISTEMAS**  
**INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN**

---

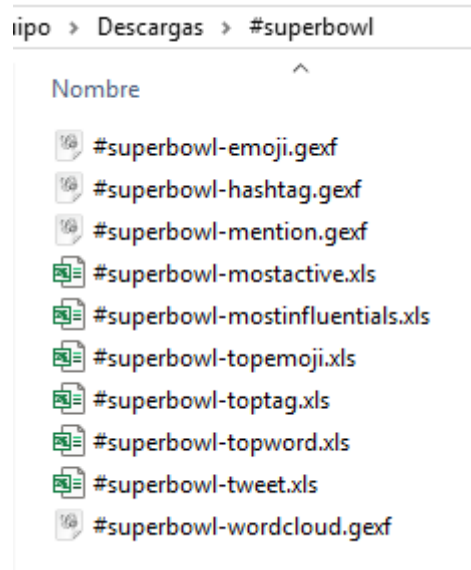
Después de un tiempo nos marcará que recolectó una cierta cantidad de Tweets

SocioViz.



Descargamos el conjunto de archivos generado pulsando en el menú de la parte superior derecha y luego descargar

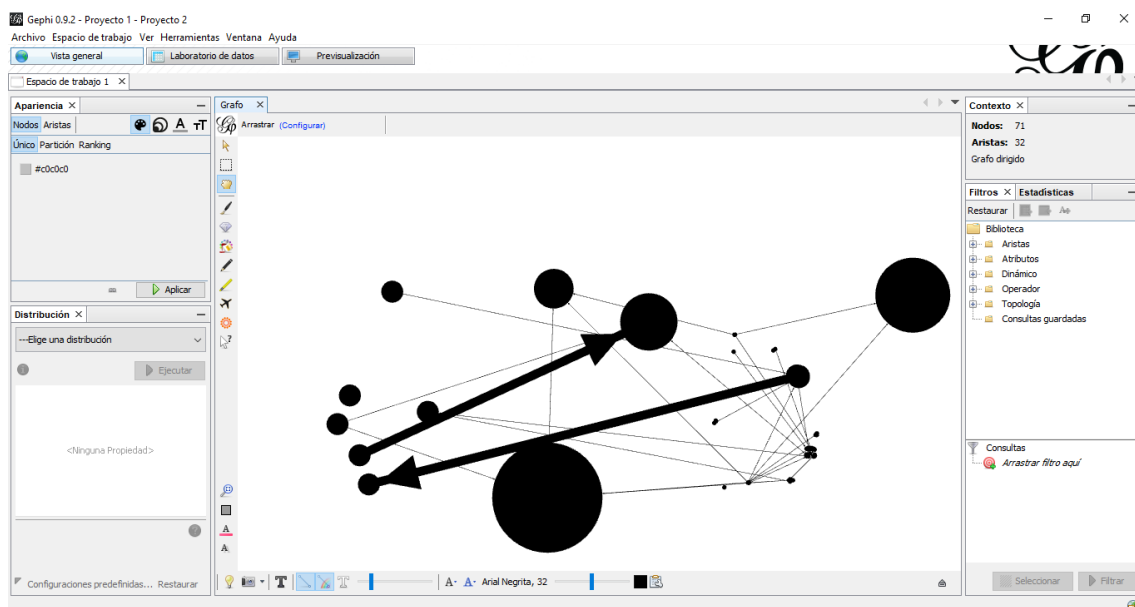
### 3.2. Cargar al archivo descargado en el punto anterior a Gephi.



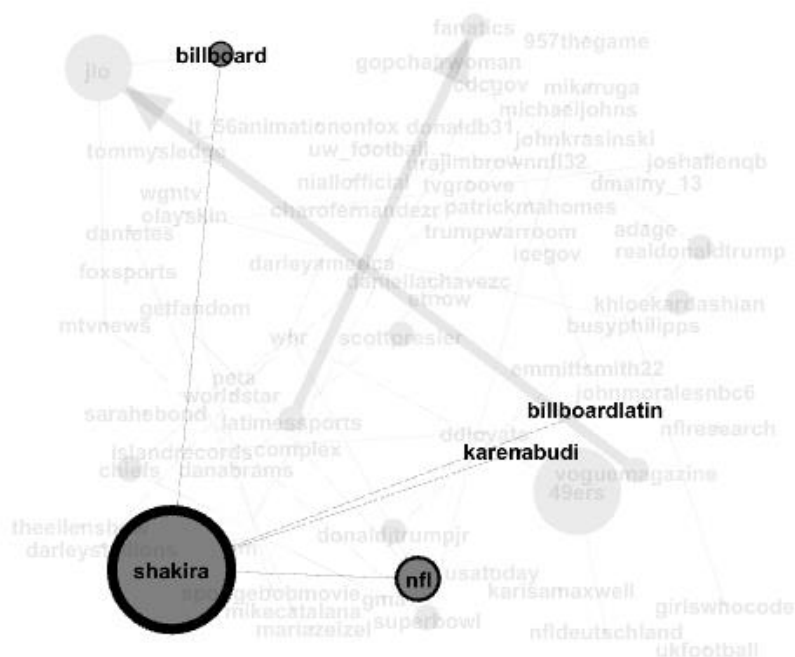
De todos los archivos generados se elige #superbowl-mention.gexf para cargar a Gephi

### 3.3. Visualizar y reportar peso de nodos y conectividad entre nodos.

Existen 71 nodos y 32 aristas. El peso de los nodos se representa gráficamente.



Si establecemos etiquetas además de los pesos se mostrarán los nombres de los referenciados.





## ESCUELA POLITÉCNICA NACIONAL FACULTAD DE INGENIERÍA DE SISTEMAS INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Se puede observar también la tabla de datos origen

Vista general | Laboratorio de datos | Previsualización

Espacio de trabajo 1

Tabla de datos

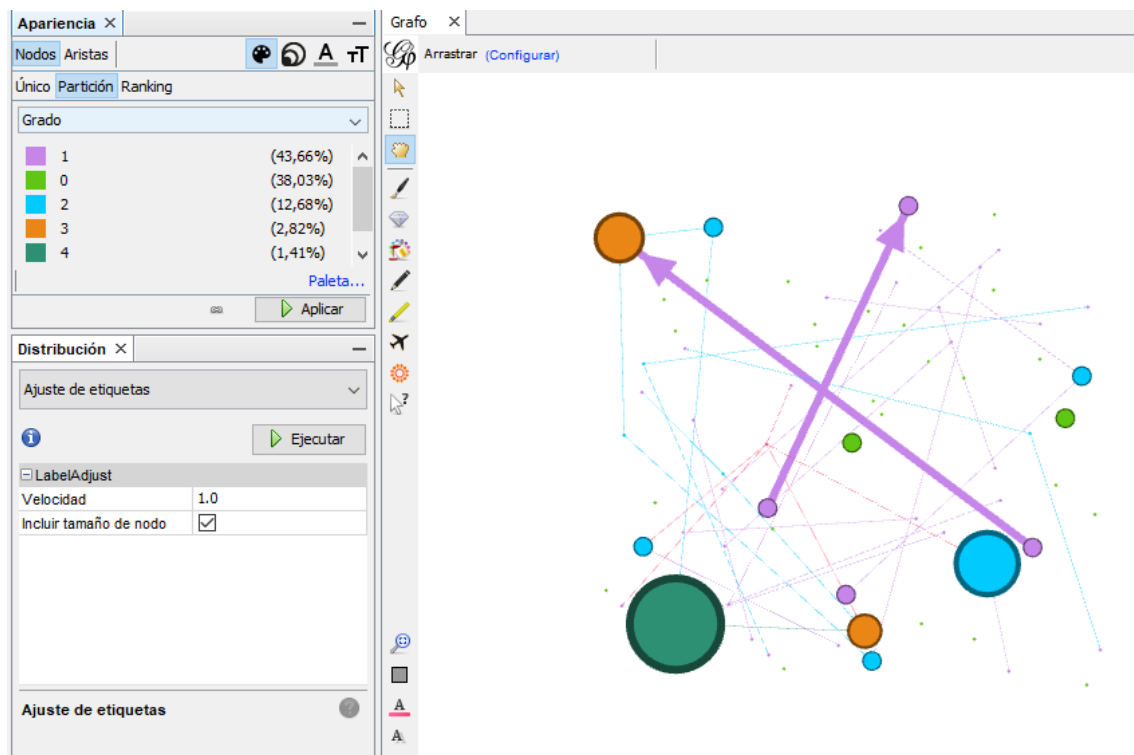
Nodos | Aristas | Configuración | Añadir nodo | Añadir arista | Buscar/Reemplazar | Importar hoja de cálculo | Exportar tabla | Más acciones

Filtro: | Id

Id	Label	Timestamp	language	urlUser/Image	postGeoLat	postGeoLon	location	wordType	Grado de entrada	Grado de salida	Grado
shakira	shakira			https://pbs.twimg.com/...			Barranquilla	USER	3	1	4
49ers	49ers			https://pbs.twimg.com/...				USER	2	0	2
jlo	jlo			https://pbs.twimg.com/...				USER	3	0	3
nfl	nfl			https://pbs.twimg.com/...				USER	3	0	3
khloekardashian	khloekardashian			https://pbs.twimg.com/...				USER	0	0	0
realdonaldtrump	realdonaldtrump			https://pbs.twimg.com/...				USER	2	0	2
scottpresler	scottpresler			https://pbs.twimg.com/...			Invite me to your d...	USER	0	0	0
donalddtrumpjr	donalddtrumpjr			https://pbs.twimg.com/...				USER	0	1	1
superbowl	superbowl			https://pbs.twimg.com/...				USER	2	0	2
latimesports	latimesports			https://pbs.twimg.com/...			Los Angeles, CA	USER	0	1	1
fanatics	fanatics			https://pbs.twimg.com/...				USER	1	0	1
vogue magazine	vogue magazine			https://pbs.twimg.com/...			New York, NY	USER	0	1	1
chiefs	chiefs			https://pbs.twimg.com/...				USER	2	0	2
ballboard	ballboard			https://pbs.twimg.com/...				USER	0	2	2
complex	complex			https://pbs.twimg.com/...			New York	USER	0	0	0
nra	nra			https://pbs.twimg.com/...			Fairfax, VA	USER	0	0	0
gopchairwoman	gopchairwoman			https://pbs.twimg.com/...				USER	0	1	1
karisamaxwell	karisamaxwell			https://pbs.twimg.com/...			Charlotte, NC	USER	0	0	0
nialofficial	nialofficial			https://pbs.twimg.com/...			Mulling, Westmea...	USER	0	0	0
theellenshow	theellenshow			https://pbs.twimg.com/...			California	USER	0	0	0
worldstar	worldstar			https://pbs.twimg.com/...			USA	USER	0	1	1
mikeruga	mikeruga			https://pbs.twimg.com/...				USER	1	0	1
animationonfox	animationonfox			https://pbs.twimg.com/...				USER	0	0	0
danfetes	danfetes			https://pbs.twimg.com/...			Rochester, NY	USER	0	2	2

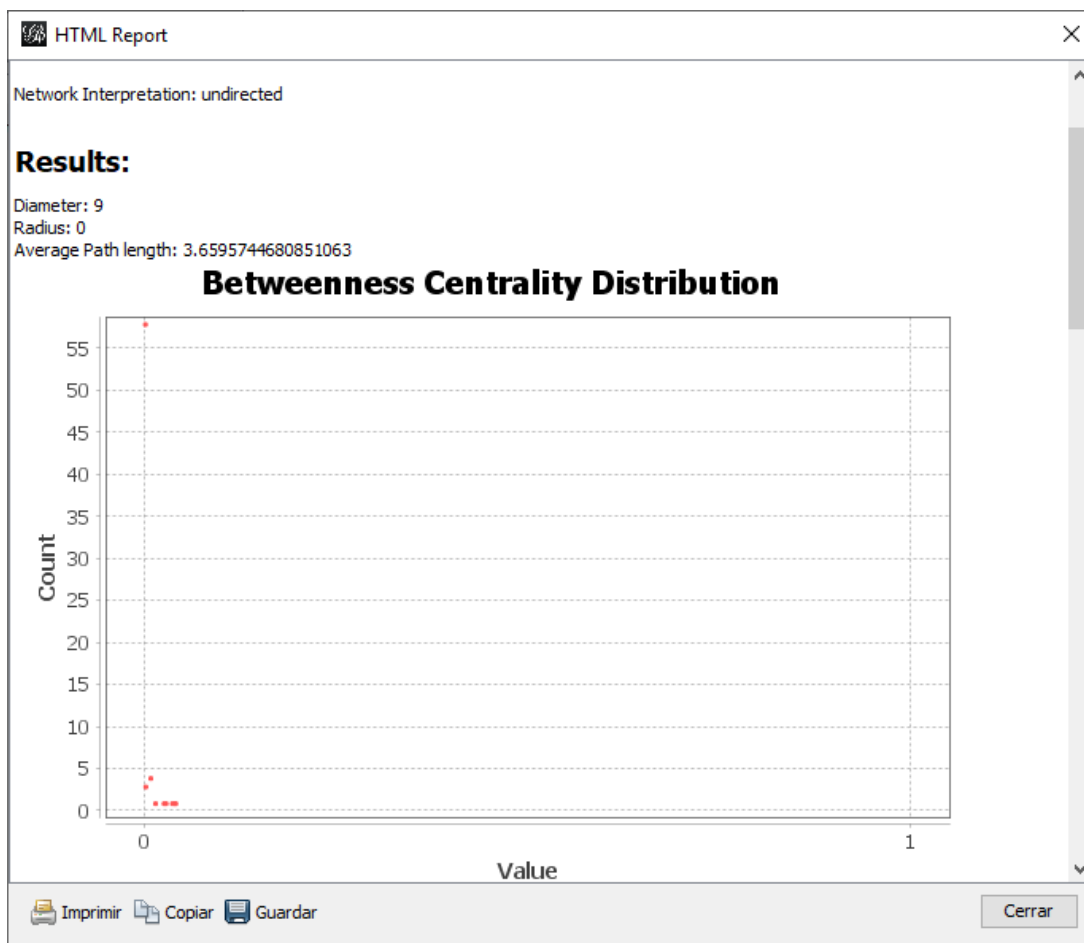
Añadir nueva columna | Mezclar columnas | Borrar columna | Borrar datos de columna | Copiar datos a otra columna | Rellenar columna con un valor | Duplicar columna | Crear columna booleana a partir de expresión regular | Crear columna con lista de grupos que se ajustan a una expresión regular | Negar columna booleana | Convertir columna a dinámica

Dependiendo el grado del nodo es posible agregar el color que los identifique



- 3.4. Seguir el video: [https://www.youtube.com/watch?v=L0C\\_D68E1Q0](https://www.youtube.com/watch?v=L0C_D68E1Q0) para revisar características de Ghepi, e identificar dos medidas de centralidad del grafo.

Se tiene un diámetro de 9 y un promedio de 3.66. Con las medidas de: Betweenness Centrality Distribution

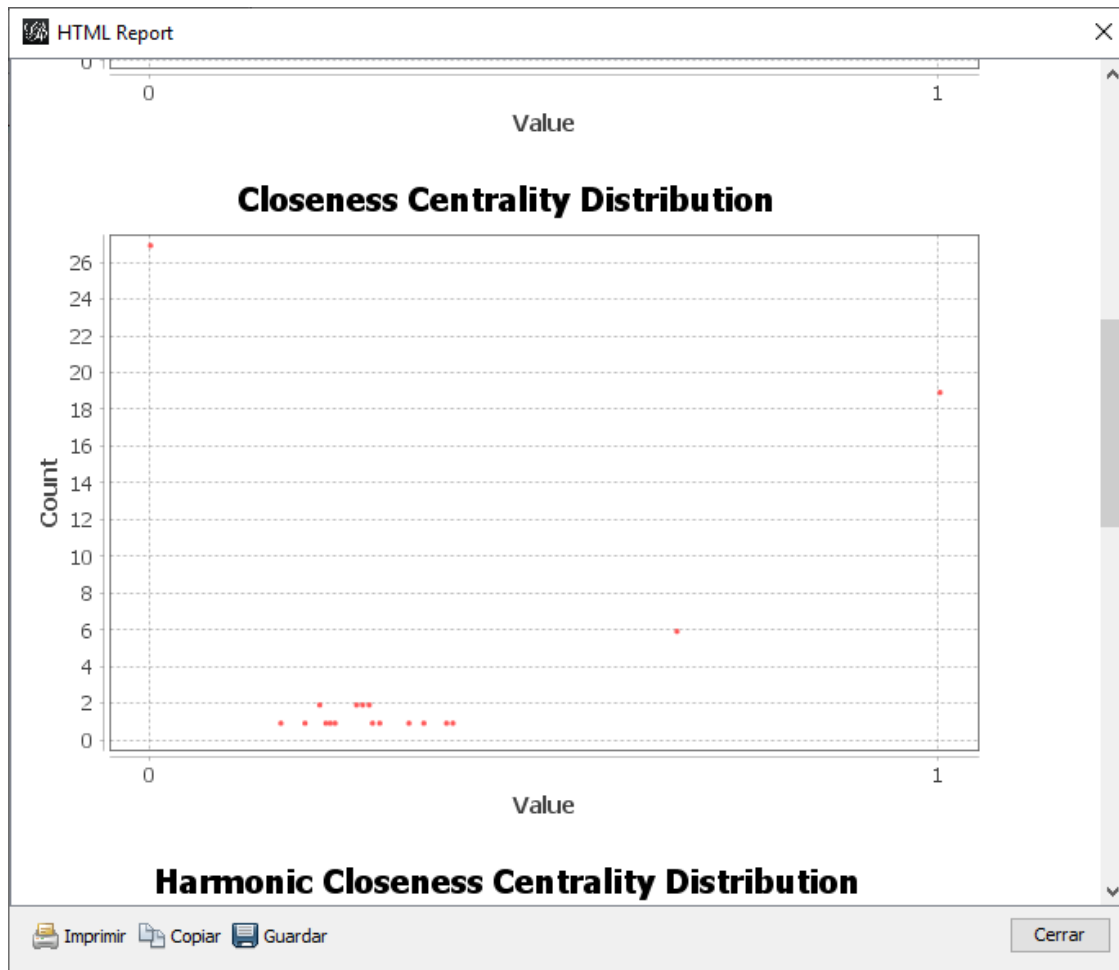




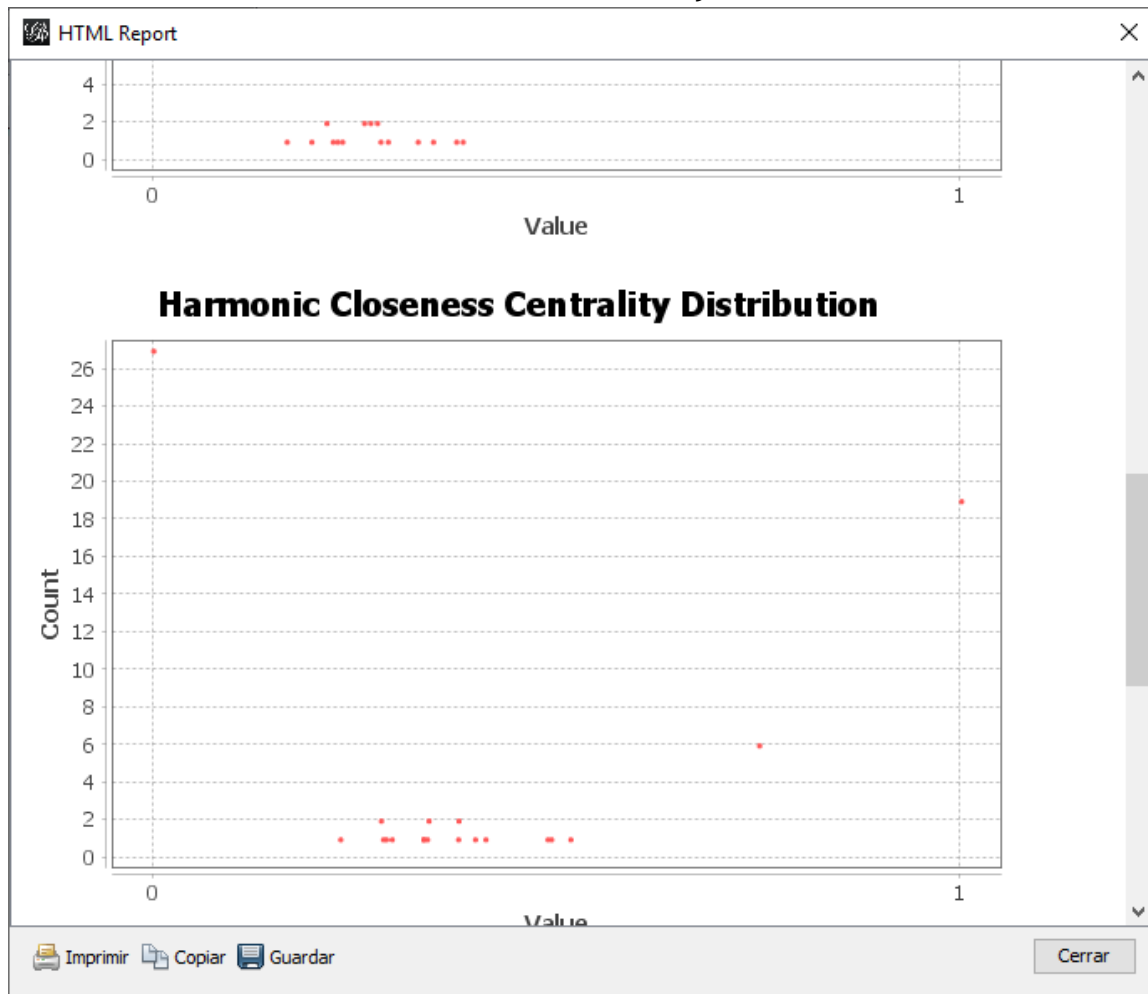
**ESCUELA POLITÉCNICA NACIONAL**  
**FACULTAD DE INGENIERÍA DE SISTEMAS**  
**INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN**

---

Closeness Centrality Distribution



## Harmonic Closeness Centrality Distribution

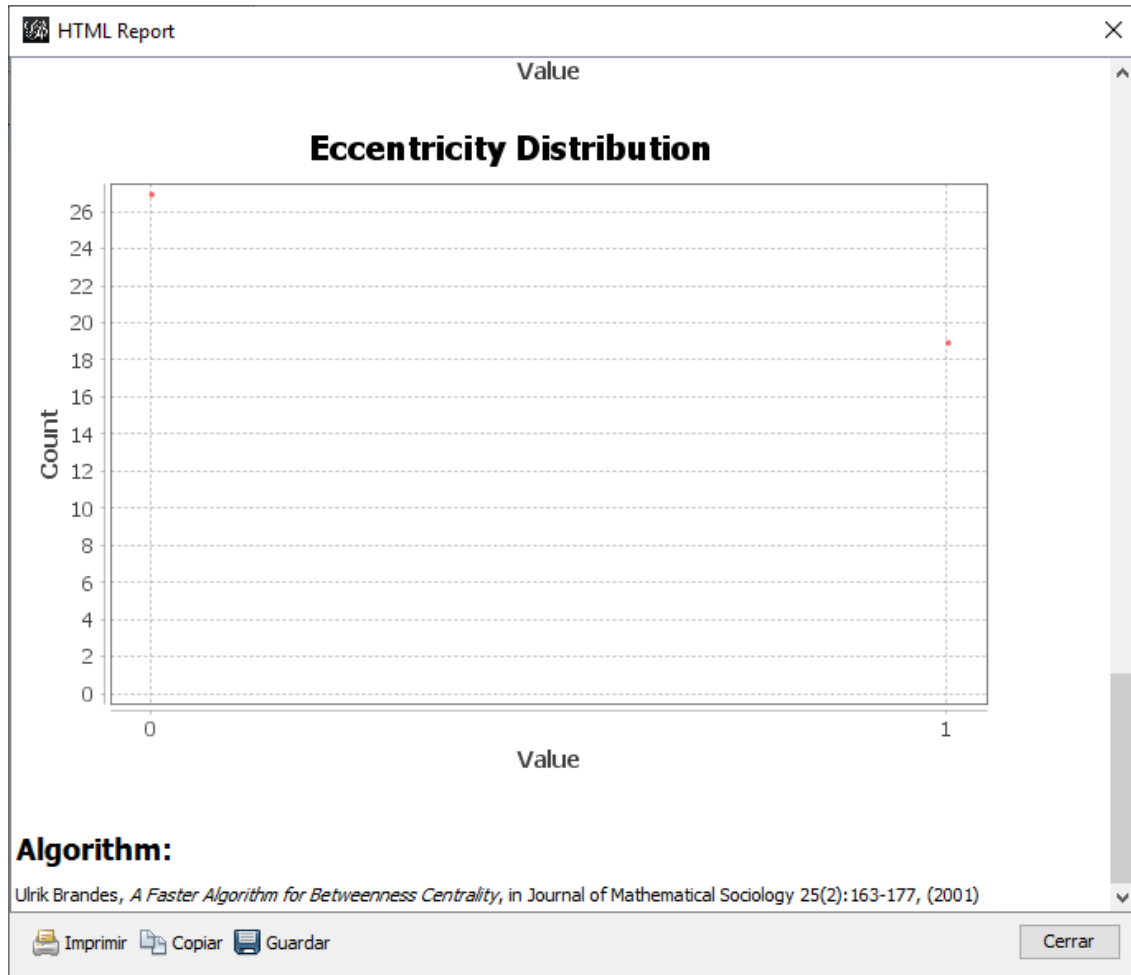




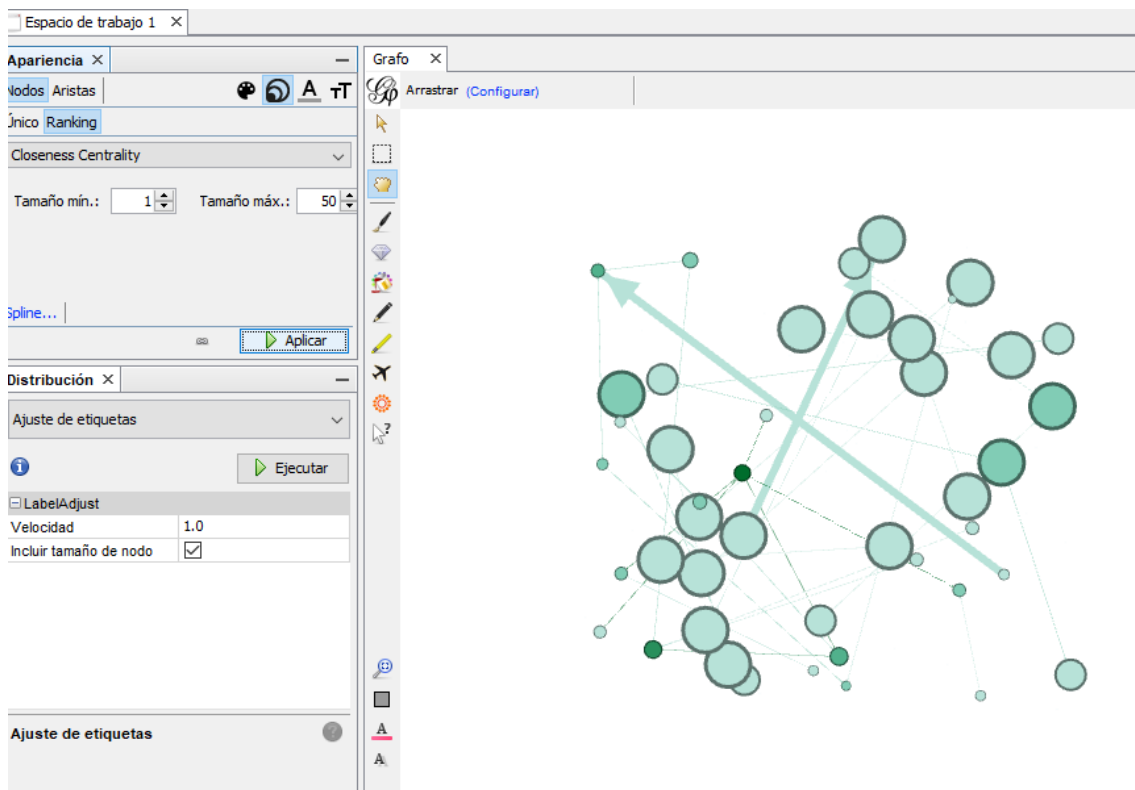
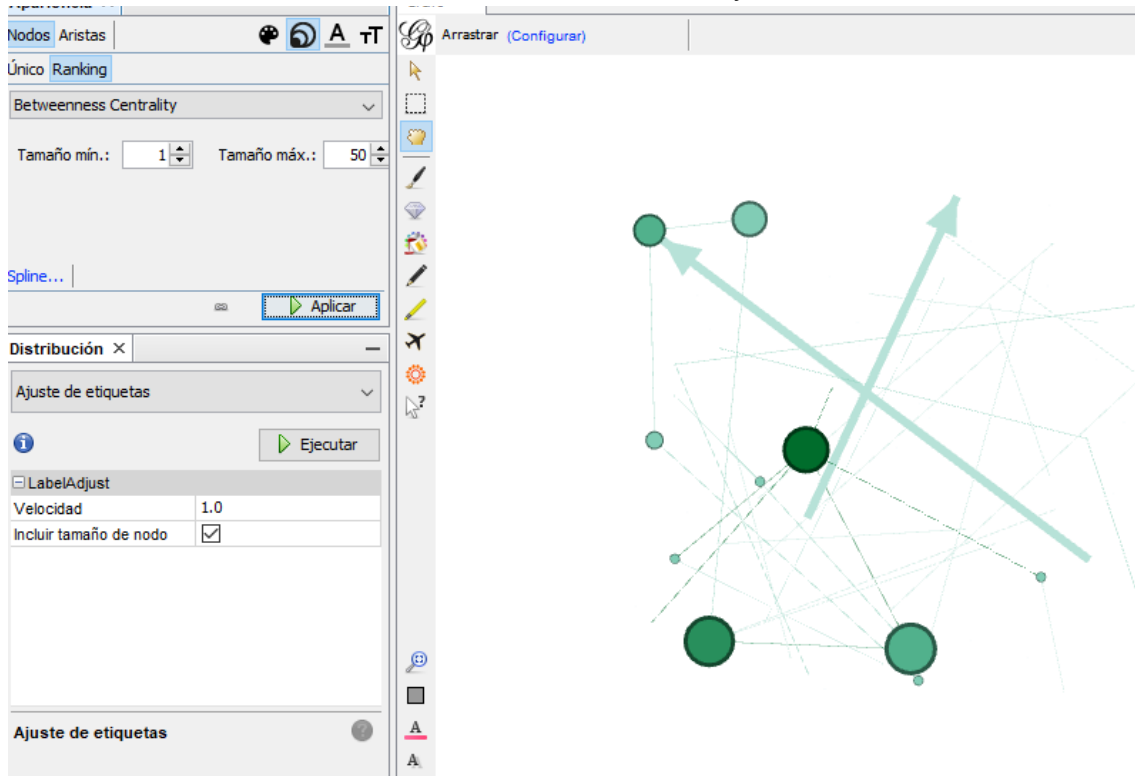


ESCUELA POLITÉCNICA NACIONAL  
FACULTAD DE INGENIERÍA DE SISTEMAS  
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Excentricity Distribution



De forma gráfica es posible establecer que tan profundo son las primeras dos medidas tomadas de Betweenness y Closeness





**ESCUELA POLITÉCNICA NACIONAL**  
**FACULTAD DE INGENIERÍA DE SISTEMAS**  
**INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN**

---

#### **4. Análisis de resultados:**

Aunque los tweets fueron pocos se notó una enorme tendencia a referenciar a Shakira.

La cantidad máxima de grado era 4, lo cual puede ser muy poco tomando en cuenta lo famoso de algunas personas.

#### **5. Conclusiones y recomendaciones:**

- Se logró entender la herramienta Gephi
  - Se graficó 4 medidas de centralidad
  - Se identificó los grados de cada nodo
  - Se definió los conceptos de centralidad
- 
- Es recomendable utilizar filtros para identificar los colores
  - La herramienta solo provee pocos tweets es importante elegir un tema muy sonado

*Explorar la herramienta Gephi realizando un grafo*

##### **5.1. Objetivos Específicos**

*Graficar dos medidas de centralidad*

*Definir los conceptos de centralidad*

#### **6. Bibliografía:**

- <http://www.joserodriguez.info/bloc/ayuda-para-entender-grafos-que-es-betweenness-intermediacion/>
- [https://en.wikipedia.org/wiki/Closeness\\_centrality](https://en.wikipedia.org/wiki/Closeness_centrality)
- <https://unpocodejava.com/2014/09/11/que-es-gephi/>