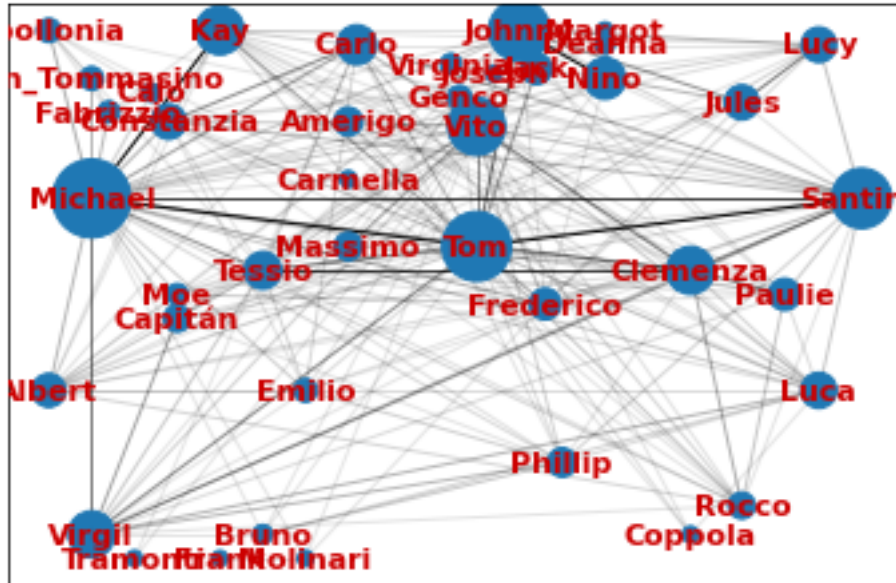


Analisis red

March 17, 2022



1 Cantidades

1.1 Tamaño

El tamaño de una red está caracterizada por el número de nodos y aristas en ella.

Cantidad de nodos: 41

Cantidad de aristas: 234

1.2 Densidad

La densidad de una red es un número $\rho \in [0,1]$ que nos dice la proporción de todas las posibles aristas de las que realmente suceden en la red. Es decir, para una red G consistente de n nodos y m aristas, la densidad $\rho(G)$ está dada por

$$\rho(G) = \frac{m}{\frac{n(n-1)}{2}} = \frac{2m}{n(n-1)}$$

	Vito	Michael	Tom	Santino	Clemenza	Tessio	Johnny	Kay	Virgil	Constanza	Emilio	Phillip	Frederico	Carlo	Lucy	Nino	Jack	Capitán	Carmella	Jules	Albert	Luca	Amerigo	Rocco	Genco	Bruno	Paulie	Massimo	Moe	Virginia	Apollonia	Don Tommasino	Calo	Fabrizio	Margot	Coppola	Deanna	Joseph	Tramonti	Frank	Molinari		
Vito	1																																										
Michael		1																																									
Tom			1																																								
Santino				1																																							
Clemenza					1																																						
Tessio						1																																					
Johnny							1																																				
Kay								1																																			
Virgil									2																																		
Constanza										2																																	
Emilio											1																																
Phillip												2																															
Frederico													2																														
Carlo														2																													
Lucy															1																												
Nino																2																											
Jack																	2																										
Capitán																		2																									
Carmella																			2																								
Jules																				2																							
Albert																					2																						
Luca																						2																					
Amerigo																							2																				
Rocco																								2																			
Genco																									2																		
Bruno																										2																	
Paulie																											2																
Massimo																												2															
Moe																																											

Densidad = 0.28536585365853656

2 Camino

Un camino es una sucesión de aristas continuos que conectan dos nodos.

2.1 Tamaño del camino más corto

De las sucesiones de caminos entre dos nodos, se elige el camino más corto (ver tabla ??).

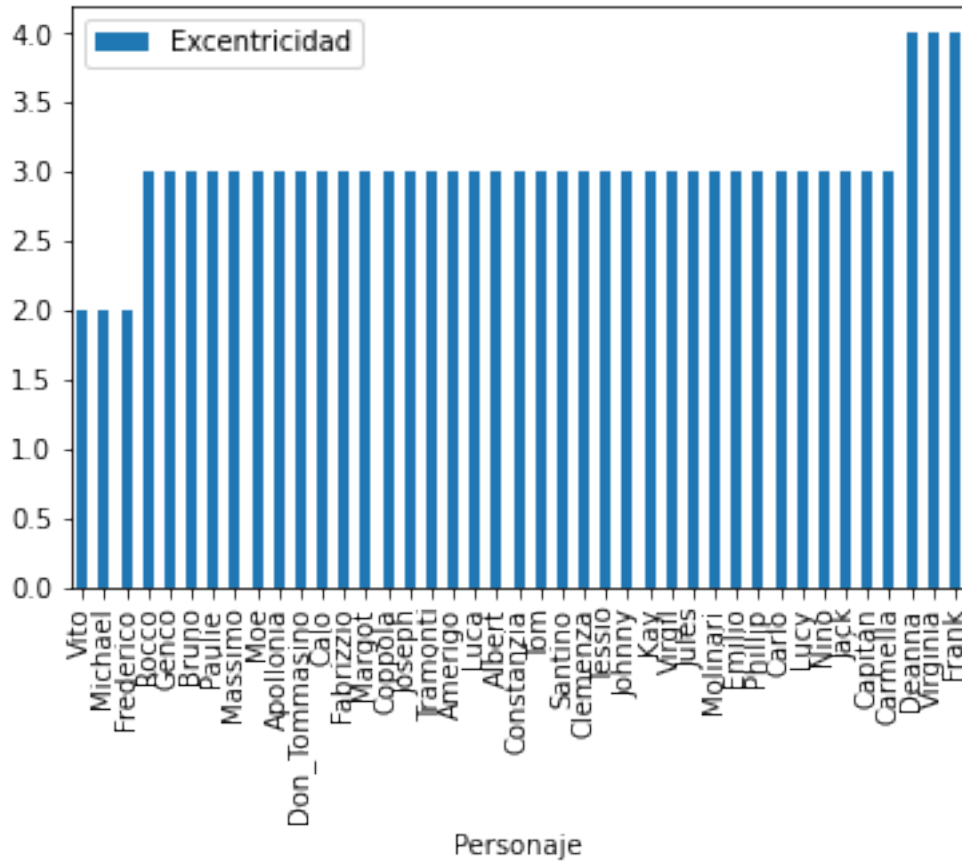
2.2 Tamaño de camino característico

Es el valor esperado de la longitud de caminos más cortos entre todo par de nodos.

Tamaño del camino característico: 1.8426829268292684

2.3 Excentricidad

La excentricidad de un nodo v en una red, es la distancia máxima del nodo v a cualquier otro nodo.



2.4 Diametro

La longitud más grande entre las longitudes de los caminos más cortos entre los nodos. Es decir, la máxima excentricidad.

Diametro: 4

2.5 Radio

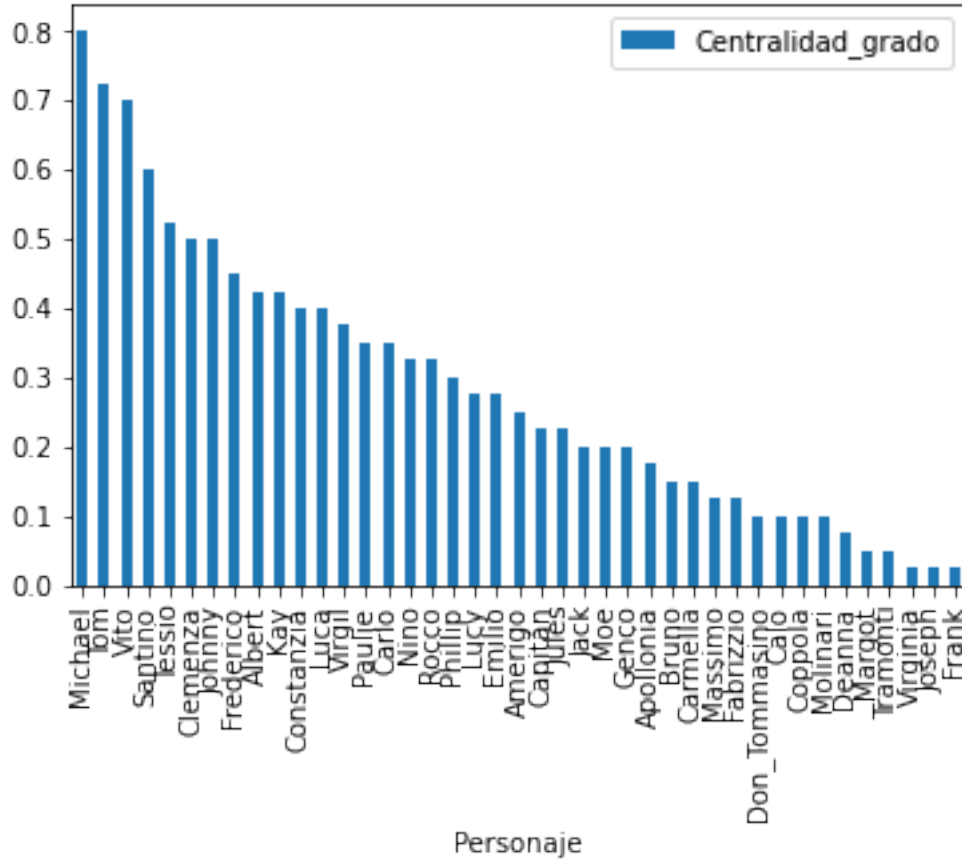
Es el mínimo de las excentricidades.

Radio: 2

3 Centralidades

3.1 De grado

Se define para cada Nodo como el porcentaje de nodos que conectan con él por medio de una arista.

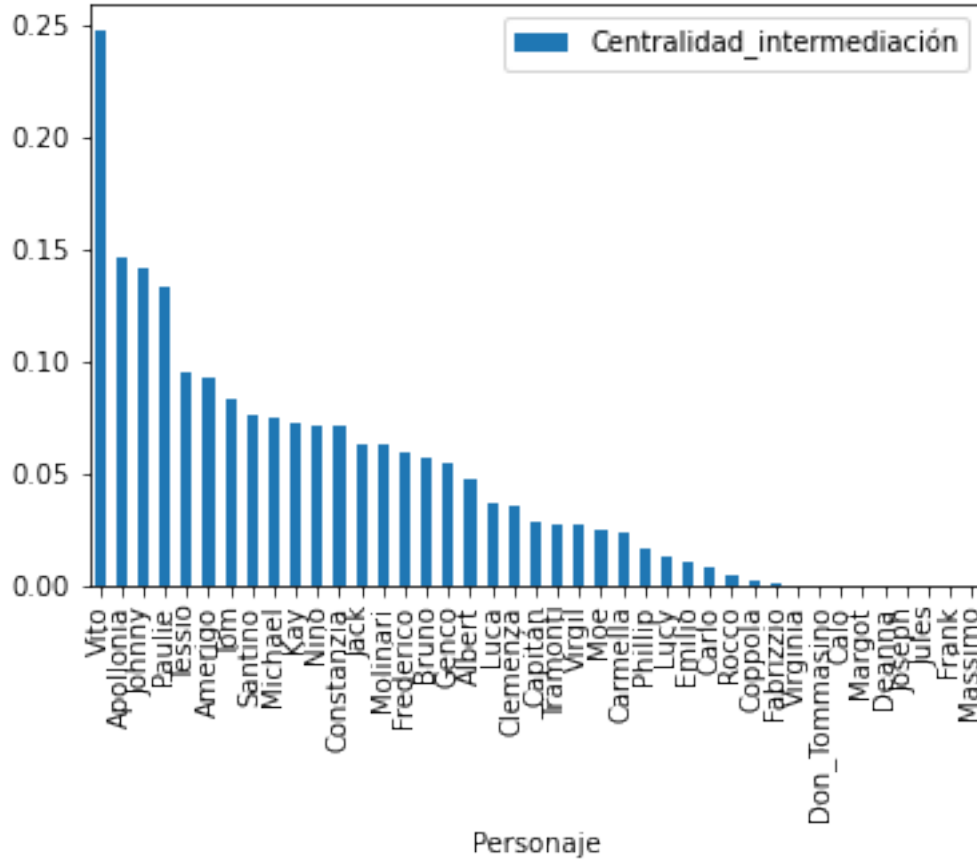


3.2 De intermediación

La centralidad de intermediación de un nodo v es la suma de la fracción de todos los caminos más cortos a otro nodo que pasan por v .

$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

Donde V es el conjunto de nodos, $\sigma(s,t)$ es la cantidad de caminos más cortos entre s y t y $\sigma(s,t|v)$ es la cantidad de estos caminos que además pasan por v , con v distinto de s y t . Si $s = t$, $\sigma(s,t) = 1$

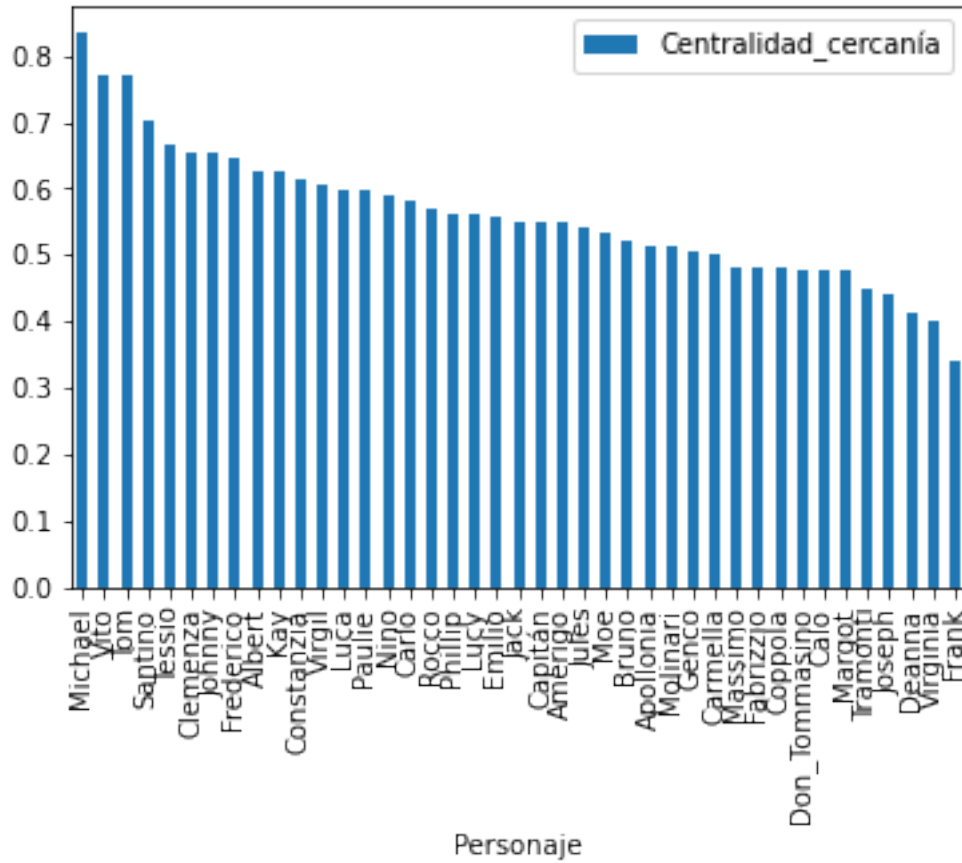


3.3 De cercanía

La centralidad de cercanía de un nodo v es el recíproco de la distancia promedio del camino más corto a v sobre todos los $n - 1$ nodos alcanzables.

$$C(v) = \frac{n - 1}{\sum_{u=1}^{n-1} d(v, u)}$$

donde $d(v, u)$ es la distancia del camino más corto entre v y u , y n es el número de nodos que pueden alcanzar v .

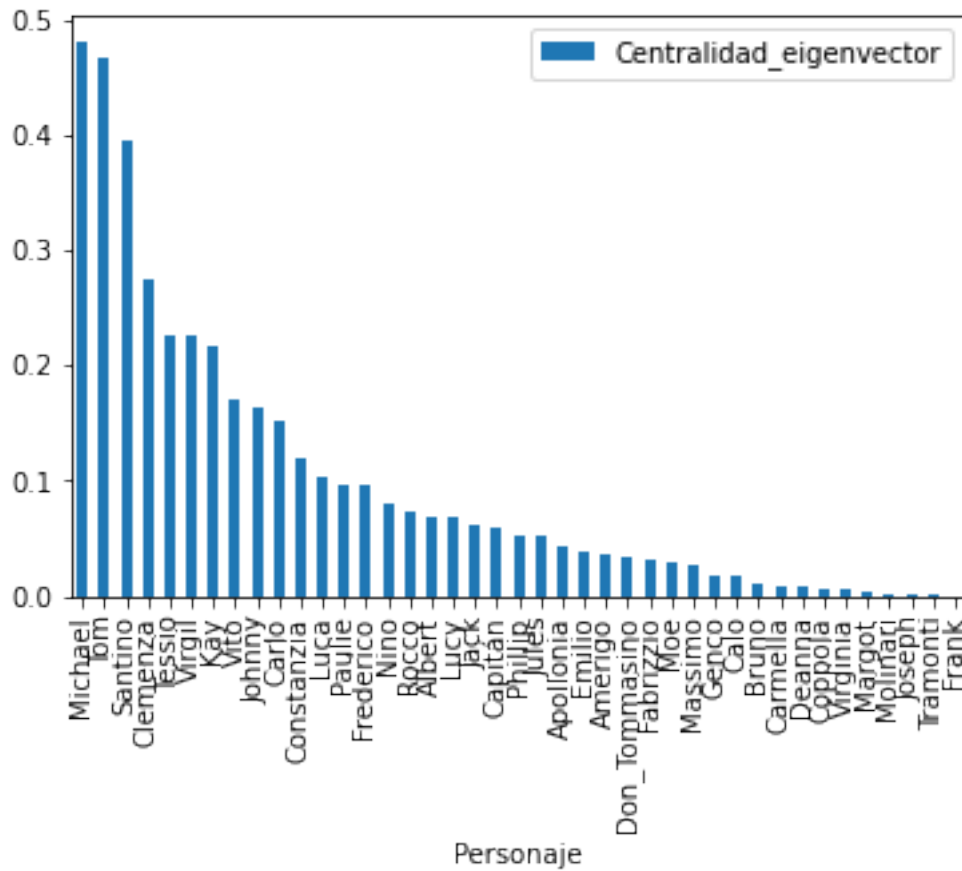


3.4 De eigenvector

La centralidad de eigenvector (vector propio) calcula la centralidad de un nodo en función de la centralidad de sus vecinos. La centralidad del vector propio para el nodo i es el i -ésimo elemento del vector x definido por la ecuación

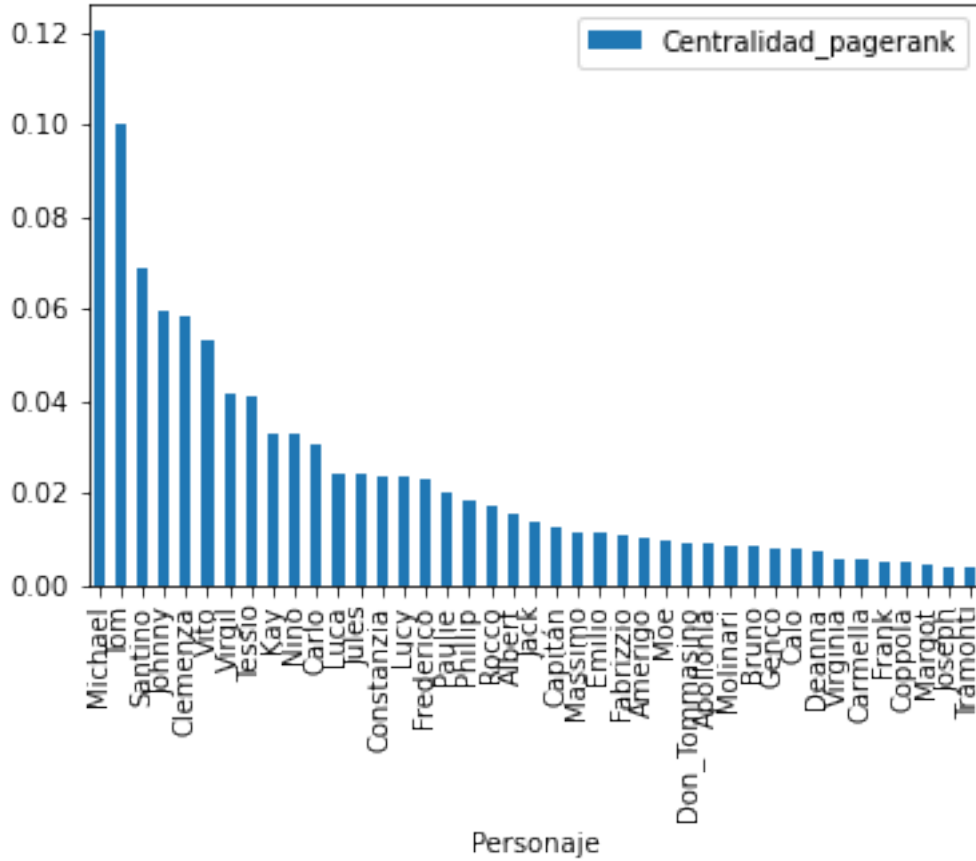
$$Ax = \lambda x$$

donde A es la matriz de adyacencia del grafo G con eigenvalor λ .



4 Page-rank

PageRank calcula una clasificación de los nodos en el gráfico G en función de la estructura de los enlaces entrantes. Originalmente fue diseñado como un algoritmo para clasificar páginas web.



5 Clustering

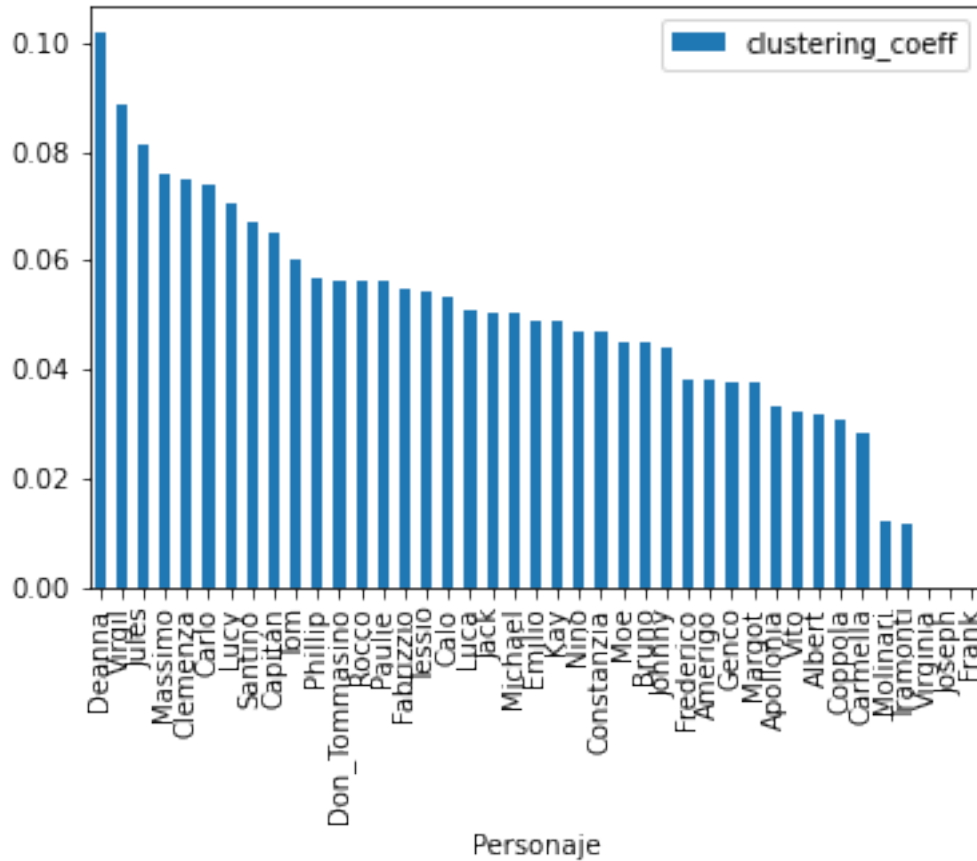
5.1 Coeficiente de clustering

Un triplete son tres nodos que están conectados por dos o tres aristas no dirigidas. Por lo tanto, un gráfico triangular (triángulo) incluye tres tripletes cerrados (de 3 aristas), uno centrado en cada uno de los nodos (nota: esto significa que los tres tripletes en un triángulo provienen de selecciones superpuestas de nodos).

Para gráficas no-pesadas, el coeficiente de clustering de un nodo u es la fracción de posibles triángulos que existen a través de ese nodo,

$$c_u = \frac{2T(u)}{\deg(u)(\deg(u) - 1)}$$

donde $T(u)$ es el número de triángulos a través del nodo y $\deg(u)$ es el grado del nodo u . Si $\deg(u) < 2$ entonces $c_u = 0$.



5.2 Transitividad

Es la fracción de todos los triángulos posibles presentes en G.

Los posibles triángulos se identifican por el número de “tríadas” (dos aristas con un vértice compartido).

La transitividad está dada por

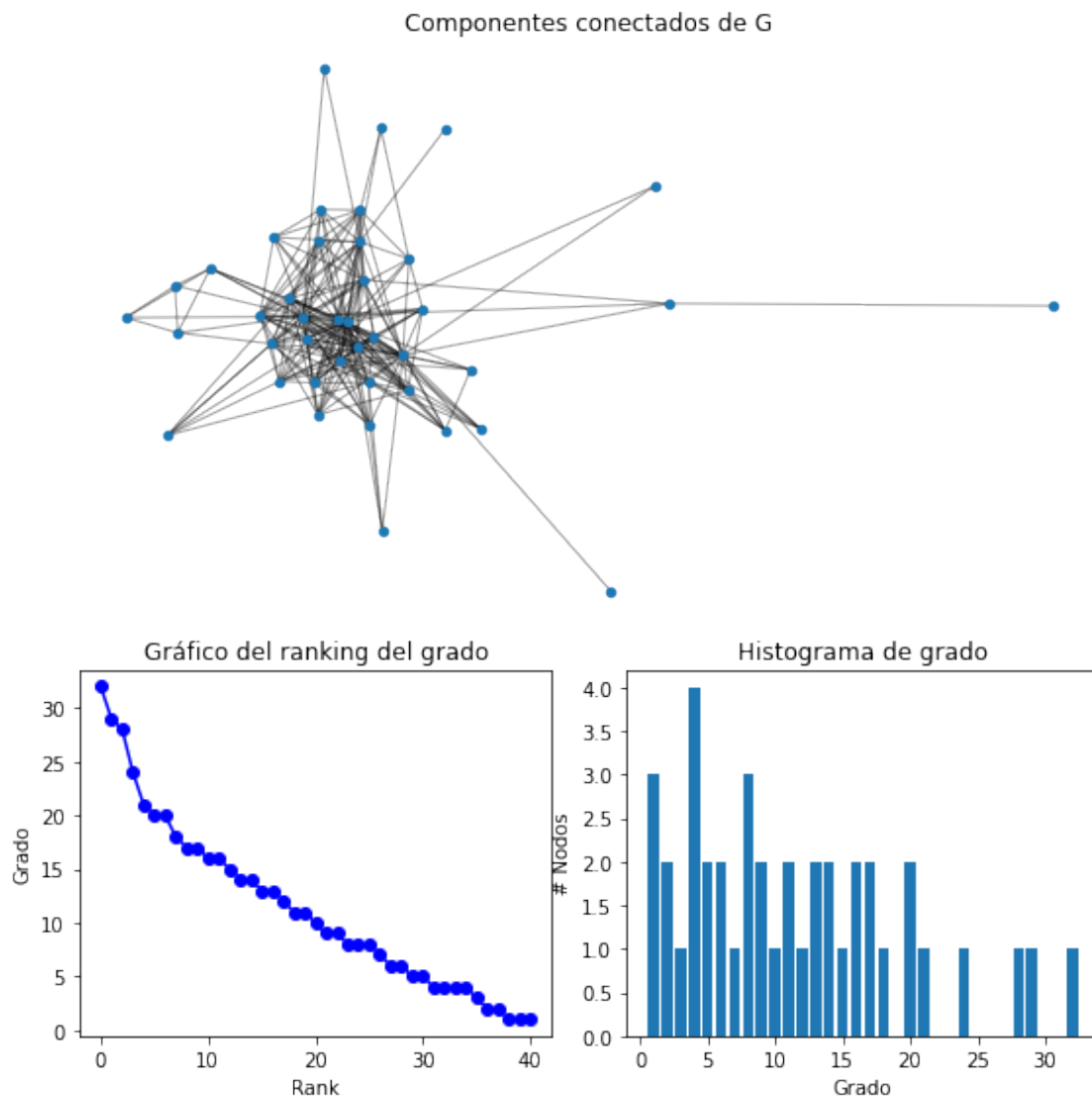
$$T = 3 \frac{\#triangulos}{\#triadas}$$

Transitividad = 0.5684803001876173

6 Distribución de grado

La distribución de grado en una red se representa habitualmente como $P(k)$ y es definida como la fracción de nodos en la red con un cierto grado k . Es como decir que si la red posee una cantidad de n nodos en total y n_k de ellos repartidos en cada grado k , de esta forma se tiene que $P(k) =$

nk/n . La misma información se presenta en forma de una distribución acumulativa de grado que viene a indicar la fracción de nodos con un grado mayor o igual que k .



7 Comunidad

7.1 Estructuras de Comunidad

Se dice que una red tiene estructura de comunidad si los nodos de la red se pueden agrupar fácilmente en conjuntos de nodos (potencialmente superpuestos) de modo que cada conjunto de nodos esté densamente conectado internamente.

```
[['Albert', 'Amerigo', 'Apollonia', 'Bruno', 'Calo', 'Capitán', 'Carlo',
'Carmella', 'Clemenza', 'Constanzia', 'Coppola', 'Deanna', 'Don_Tommasino',
'Emilio', 'Fabrizzio', 'Frank', 'Frederico', 'Genco', 'Jack', 'Johnny', 'Jules',
```

'Kay', 'Luca', 'Lucy', 'Margot', 'Massimo', 'Michael', 'Moe', 'Molinari',
 'Nino', 'Paulie', 'Phillip', 'Rocco', 'Santino', 'Tessio', 'Tom', 'Tramonti',
 'Virgil', 'Vito'], ['Joseph'], ['Virginia']]

7.2 Modularidad

La modularidad es la fracción de los vértices que caen dentro de los grupos dados menos la fracción esperada si los vértices se distribuyeran al azar. El valor de la modularidad para gráficos no ponderados y no dirigidos se encuentra en el rango $[-1/2, 1]$.

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta(c_i, c_j)$$

Donde m es el número de vértices, A es la matriz de adyacencia de G . k_i es el grado de i , γ el parámetro de resolución y $\delta(c_i, c_j) = 1$ si i y j están en la misma comunidad y $\delta(c_i, c_j) = 0$ en otro caso

[58]: -1.2174804562347645e-05

8 Conclusiones

El resultado que más llama mi atención es el **Clustering**. Se puede notar que los personajes con clustering más alto no son los principales, sin embargo son personajes que solo aparecen en las subtramas de la historia.

Deanna por ejemplo, es amiga de Johnny Fontain y se relacionaba con Nino Valenti e interacciona con los personajes que están en el medio.

Virgil Solozzo está dentro de otra subtrama con los Corleone y las demás familias. Siendo el principal antagonista en la parte relacionada al tráfico de drogas.

Jules Seagal es un doctor que podría considerarse especial de la familia Corleone, atendiendo casos concretos de acciones turbias.

Finalmente, Massimo Fanucci es el antiguo Don local de Nueva York. Representa el principal antagonista en la juventud de Vito Corleone.

Otro aspecto a destacar es la centralidad. A pesar de que Michael es el personaje con mayor centralidad de grado y de cercanía, dándonos a notar que él será el nuevo Don, Vito Corleone es, por mucho, el personaje con mayor centralidad de intermediación. Y es que, en la novela, Vito Corleone ya no es el Don que era, se nos presenta sus últimos años como Don, vemos el cambio de generación y es sabiendo que esto ocurrirá que él actúa como intermediario del nuevo don (Michael) con las conexiones que él tenía anteriormente. Podemos pensar que su papel es introducir al nuevo Don en ese mundo.