

Alumna: Paula Brück

Padron: 107533

TEORIA DE ALGORITMOS II

(75.30)

Trabajo Final



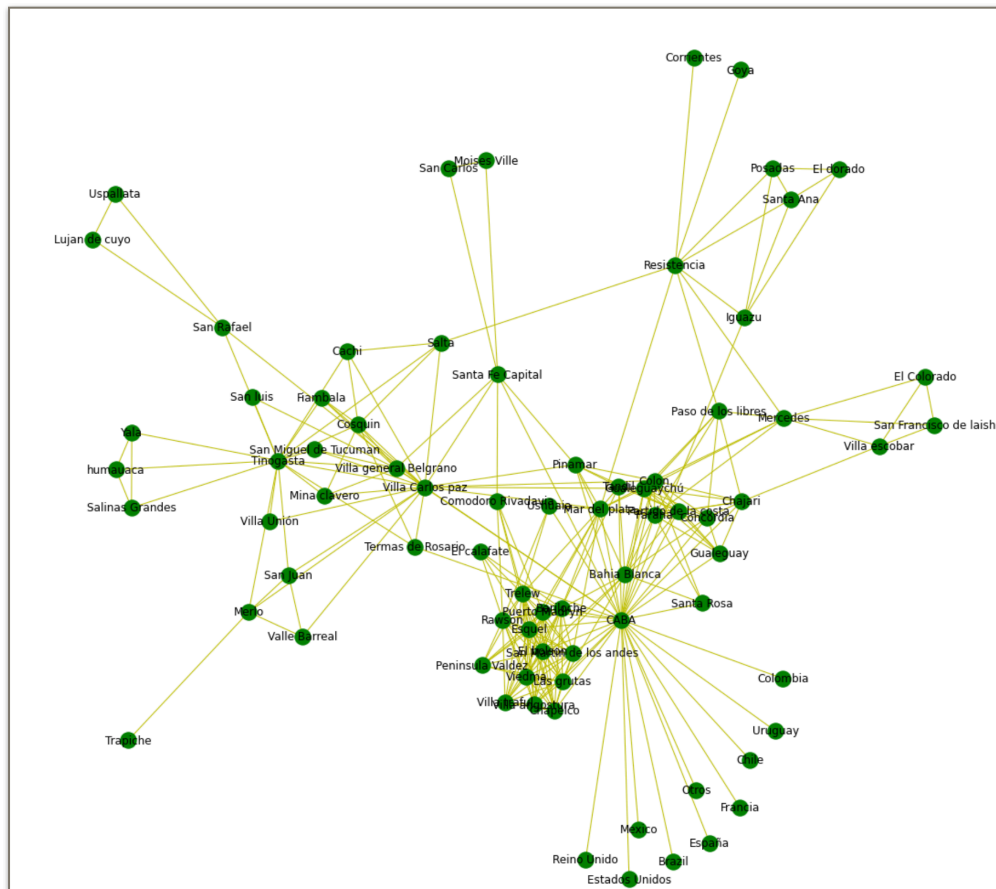
Cómo el “PREVIAJE” modifico la forma de viajar de los Argentinos.

Previjaje es un programa de preventa turística que fue implementado por el gobierno en el que se reintegra el 50% del valor del viaje comprado en crédito, para viajar y disfrutar de todos los destinos de Argentina. El mismo tiene el objetivo de impulsar el turismo nacional y revalorizar los lugares turísticos en Argentina. La iniciativa del Ministerio de Turismo y Deportes (Minturdep) tuvo su primera edición en 2020 la cual logro cumplir exitosamente con las expectativas de la propuesta y a su vez su resultado permitió llegar a conclusiones muy interesantes en cuanto a la forma de vacacionar de la población, pero por sobretodo: Transformó la forma de viajar de los Argentinos .

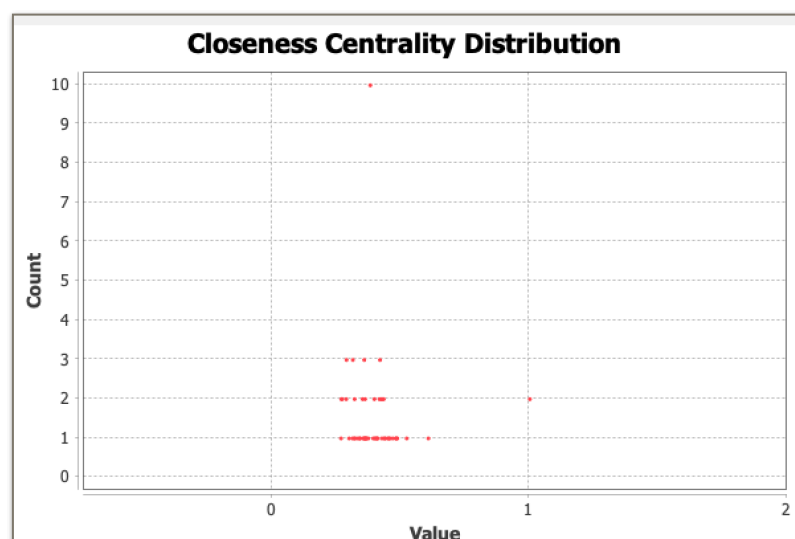
En este trabajo la idea es analizar cómo fue la evolución de viajes por el país mediante esta medida conocido como “PREVIAJE” . Para esto a continuación se analizaran 2 redes formadas con la base de datos proporcionada por una agencia turística donde una de ellas tiene los viajes que fueron vendidos previos a la implementación del Previjaje , es decir en el año 2019 , y otra del 2020 , año en el que tuvo lugar la primera edición de este beneficio.

Padron: 107533

Red de Viajes 1 —————> Previo al PREVIAJE/2019



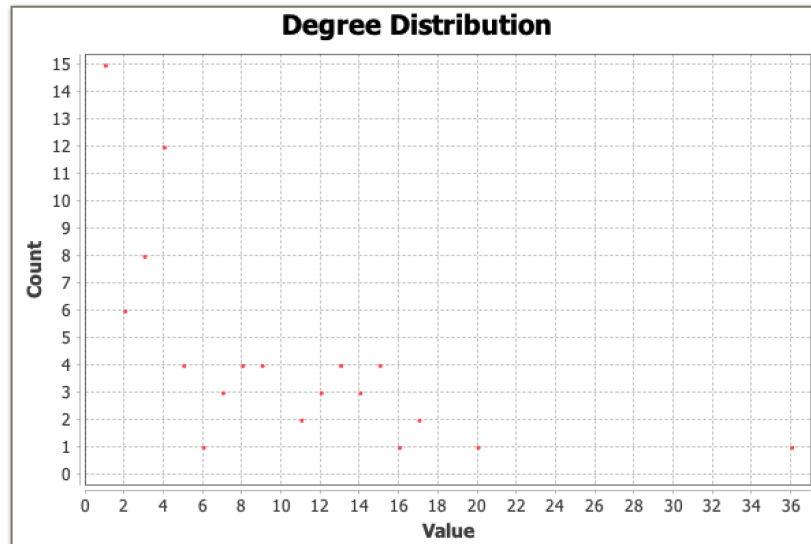
- * Según la Red de datos en el 2019 se realizaron 13887 ventas de viajes a 76 destinos . 2559 de estas ventas fueron viajes internacionales .
- * El diámetro de la red es de 5. Este valor representa la máxima distancia existente entre dos nodos en toda la red siendo la distancia media 2,5.



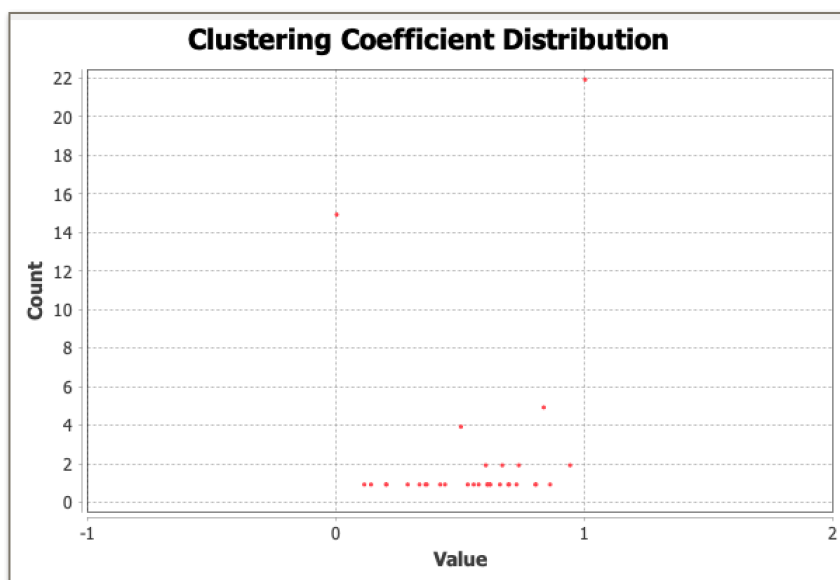
Alumna: Paula Brück

Padron: 107533

* En esta Red el grado Promedio es 6.973684210526316 lo cual quiere decir que cada nodo de la red está conectado con otros 6 en media y las gráficas de densidades de los grados son las siguientes:



* En el caso de esta Red al representar los viajes entre lugares turísticos y por lo tanto al ser no dirigida el promedio del coeficiente de Clustering, es decir el promedio de los coeficientes individuales para cada vértice, es de 0.6080231666338478.



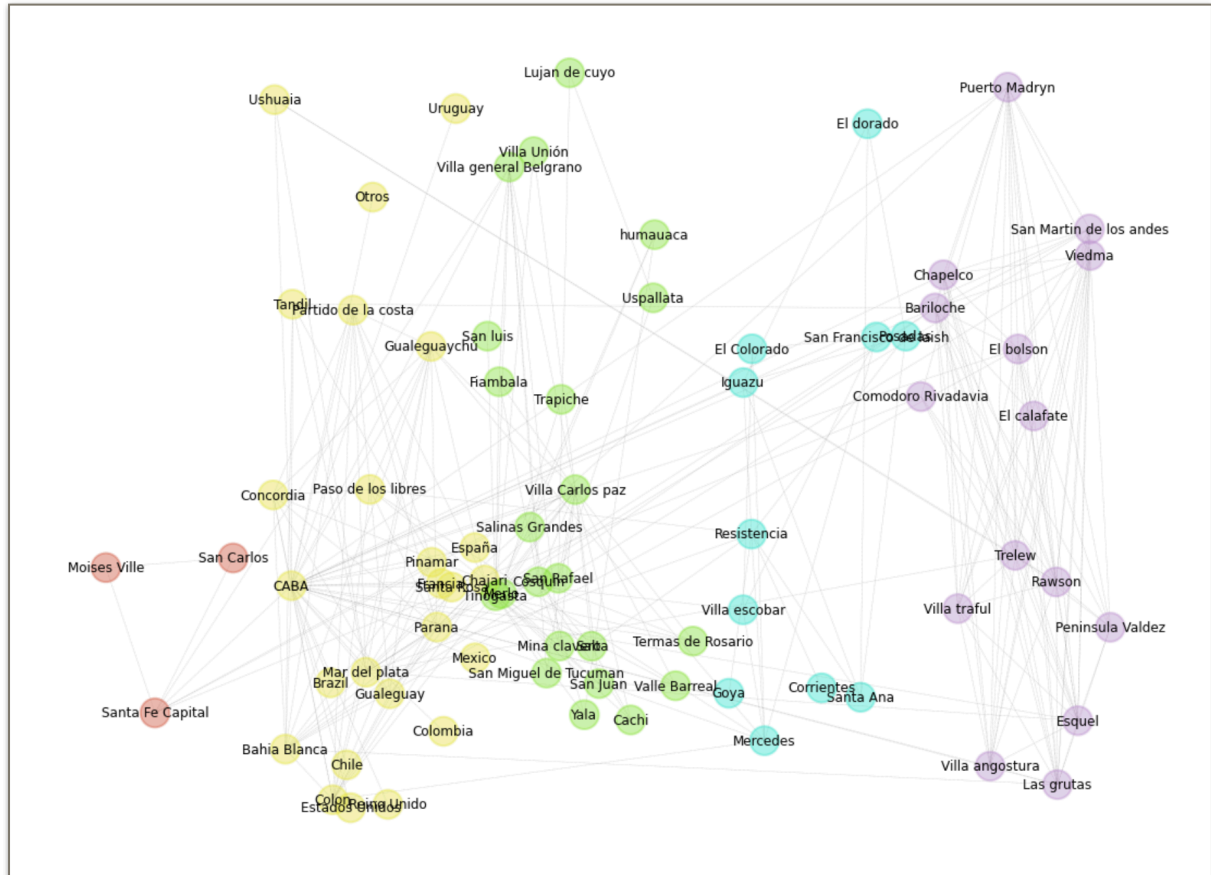
* En el caso de esta red de vuelos podemos analizar la homofilia por zonas, es decir el espacio geográfico al que pertenecen los lugares turísticos que se conectan mediante los viajes que realiza la población.

* Nuestra red de viajes cuenta con 76 Nodos (lugares turísticos) y con 265 aristas (viajes).

Alumna: Paula Brück

Padron: 107533

* Utilizando el algoritmo de Louvain para detectar comunidades podemos considerar en esta red 5 comunidades distintas :



* Comunidad -----> 0 : 3 nodos

['Santa Fe Capital', 'Moises Ville', 'San Carlos']

* Comunidad -----> 1 : 25 nodos

['Mar del plata', 'Tandil', 'Pinamar', 'Partido de la costa', 'CABA', 'Parana', 'Colon', 'Guaquaychú', 'Ushuaia', 'Bahia Blanca', 'Brazil', 'Uruguay', 'Chile', 'Estados Unidos', 'Mexico', 'Colombia', 'España', 'Francia', 'Reino Unido', 'Otros', 'Concordia', 'Chajari', 'Guaquay', 'Santa Rosa', 'Paso de los libres']

* Comunidad -----> 2 : 22 nodos

['Villa Carlos paz', 'Villa general Belgrano', 'Termas de Rosario', 'Tinogasta', 'Fiambala', 'Cosquin', 'Mina claver', 'humauaca', 'Yala', 'Salinas Grandes', 'Villa Unión', 'San Rafael', 'Cachi', 'Salta', 'San Juan', 'Merlo', 'San luis', 'San Miguel de Tucuman', 'Valle Barreal', 'Uspallata', 'Lujan de cuyo', 'Trapiche']

Alumna: Paula Brück

Padron: 107533

* Comunidad -----> 3 : 11 nodos

['Resistencia', 'Iguazu', 'Corrientes', 'Goya', 'Mercedes', 'Santa Ana', 'Posadas', 'Villa escobar', 'San Francisco de laish', 'El Colorado', 'El dorado']

* Comunidad -----> 4 : 15 nodos

['Trelew', 'Puerto Madryn', 'Esquel', 'San Martin de los andes', 'Bariloche', 'Las grutas', 'El bolson', 'Viedma', 'Villa angostura', 'Chapelco', 'Rawson', 'Comodoro Rivadavia', 'Peninsula Valdez', 'Villa traful', 'El calafate']

* Analizando las comunidades resultantes del algoritmo de Louvain podemos decir que habría homofilia por ubicación geográfica. La comunidad 0 incluiría los lugares turísticos de la provincia de Santa Fe , la comunidad 1 los de provincia de Buenos Aires , Entre ríos , La Pampa, Tierra del Fuego y el Extranjero, la comunidad 2 las provincias de Cordoba , Jujuy , Salta, San Miguel de Tucumán, la zona de cuyo (San Luis, San Juan , Mendoza , La Rioja), la comunidad 3 las provincias de Misiones , Formosa , Chaco y Corrientes y la comunidad 4 la Patagonia (Chubut, Neuquen, Rio Negro, Santa Cruz)(excepto Tierra del Fuego). De esta manera nos quedaría Argentina dividida por zonas agrupándose las provincias en base a su cercanía entre ellas. Sin embargo podemos encontrar excepciones como es el caso de los países fuera de Argentina que se encuentran en la comunidad 1 , lo cual podemos pensar que se debe a que la mayoría de las personas que viajan al extranjero lo hacen partiendo desde la capital del país buenos aires . Otro caso también presente en la comunidad 1 es el de Ushuaia , dicha ciudad debería estar incluido en la comunidad 4 de la Patagonia , podemos asociar esta diferencia a las grandes cantidades de personas que viajan desde lugares en provincia de Buenos Aires a Ushuaia y viceversa.

* Analizando la homofilia:

* Primero procedemos a calcular la cantidad de aristas que cruzan de un grupo a otro:

* Para esto evaluamos por cada comunidad que países tienen como vecino a uno que no pertenece al mismo grupo

- Comunidad 0: 5
- Comunidad 1: 37
- Comunidad 2: 10
- Comunidad 3: 9
- Comunidad 4: 21

* Por ende la cantidad total es : $82/2$ (al cruzar de un grupo a interactúan dos nodos se suma en ambos grupos por eso se suman $*2$) = 41 aristas que cruzan desde un grupo a otro.

Alumna: Paula Brück

Padron: 107533

* Luego calculamos el valor ideal/ teórico de sin homofilia: $P = p_0 (1 - p_0) + p_1 (1 - p_1) + p_2 (1 - p_2) + p_3 (1 - p_3) + p_4 (1 - p_4)$

* El calculo de p_i es el resultado de la cantidad de nodos de el grupo i / la cantidad de nodos totales de la red.

$$p_0 = n_0 / n_T, p_1 = n_1 / n_T, p_2 = n_2 / n_T, p_3 = n_3 / n_T, p_4 = n_4 / n_T$$

$$p_0 = 3/76 = 0,039, p_1 = 25/76 = 0,32, p_2 = 22/76 = 0,28, p_3 = 11/76 = 0,14,$$

$$p_4 = 15/76 = 0,19$$

* Reemplazando p_0, p_1, p_2, p_3 y p_4 en P :

$$P = 3/76(1-3/76) + 25/76(1-25/76) + 22/76(1-22/76) + 11/76(1-11/76) + 15/76(1-15/76)$$

$$* P = 0,037 + 0,21 + 0,20 + 0,12 + 0,15$$

* **$P = 0,71$** ———> Ideal sin homofilia (threshold teórico)

* Para calcular el valor real de sin homofilia presente en nuestra red tenemos que hacer:
Aristas totales que cruzan / Aristas Totales

$$* \text{Esto es igual a : } 41/265 = 0,16$$

* Entonces teniendo el valor ideal/ teórico y el valor real de sin homofilia calculamos el porcentaje de sin homofilia:

$$0,71 \text{ ———} > 100\%$$

$$0,16 \text{ ———} > X = \mathbf{22,53 \%} \text{ sin homofilia}$$

* Por lo tanto dado que hay un 22,53% sin homofilia podemos decir que hay un 77,46% de homofilia por ubicación geográfica. Por lo tanto podemos decir como conclusión que se cumple que la mayoría de los viajes se realizan entre ciudades cercanas que se encuentran en la misma zona y que por esta característica hay homofilia .

* En cuanto a Puentes Locales podemos decir que hay 13:

[('CABA', 'Brazil'),

('CABA', 'Uruguay'),

('CABA', 'Chile'),

('CABA', 'Estados Unidos'),

('CABA', 'Mexico'),

('CABA', 'Colombia'),

('CABA', 'España'),

('CABA', 'Francia'),

('CABA', 'Reino Unido'),

('CABA', 'Otros'),

Alumna: Paula Brück

Padron: 107533

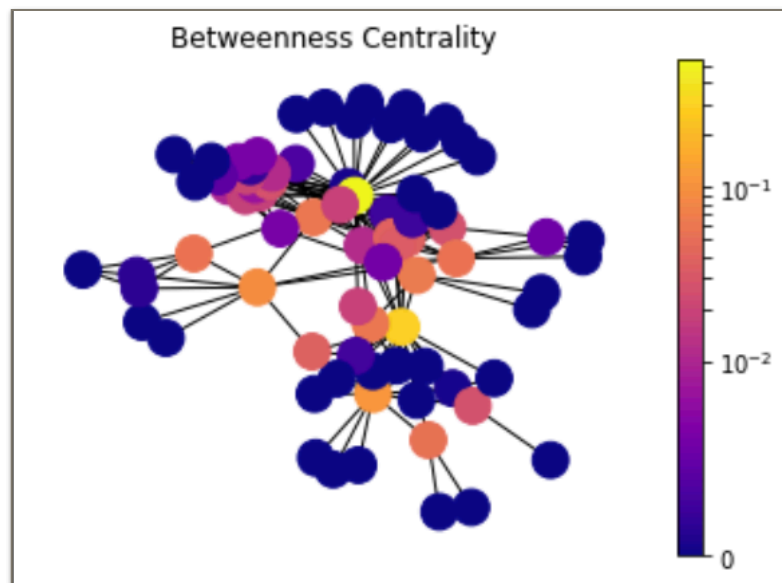
('Resistencia', 'Corrientes'),

('Resistencia', 'Goya'),

('Merlo', 'Trapiche')]

* Para esta red ya que representa los viajes entre destinos turísticos me parece más útil calcular la centralidad de Intermediación(betweenness) para poder determinar influencia. Si buscamos cual lugar turístico es mas importante , influyente podremos decir que será cual tenga mayor conexiones con otros ya que esto significara que una mayor cantidad de lugares se encuentran conectados a través de personas que viajan entre ellos . Cuanto más conocidos, importantes, mas popularizados más razones hay para que personas de diferentes lugares quieran viajar hacia ellos.

* Grafico de Centralidad de Intermediacion:



* Analizando el gráfico podemos ver claramente cómo CABA , Villa Carlos paz son algunos de los mas influyentes ,populares . Al ser el gráfico de la centralidad de intermediación se esta teniendo en cuenta todas las veces que un lugar aparece en la ruta de caminos mas cortos de otro lugar. Es decir que en este gráfico el ranking de influencia va a estar dado por aquellos lugares que tienen mayor conexión con otros .

* Del mas popular al menos :

CABA 0.5347428241600841

Villa Carlos paz 0.28869344329450447

Tinogasta 0.1077335517335517

Resistencia 0.09194325540745443

Alumna: Paula Brück

Padron: 107533

Santa Fe Capital 0.06535952546824686
Villa general Belgrano 0.06050211634302065
Mar del plata 0.059671352607430866
Iguazu 0.0577951302951303
Mercedes 0.055676771117947596
San Rafael 0.052612612612612616
Colon 0.04103299717889837
Parana 0.03815444406842676
Salta 0.03731992311722042
Gualectuaychú 0.0342790686930514
Bariloche 0.02871482794304967
Merlo 0.02723723723723724
Chajari 0.02528096744235764
El bolson 0.02094669157141412
Bahia Blanca 0.019039985744720742
Termas de Rosario 0.018765739274863524
Puerto Madryn 0.01783611562111574
Esquel 0.01783611562111574
Pinamar 0.011457982567843533
Las grutas 0.011359683984406531
Trelew 0.009958280726296025
San Martin de los andes 0.00861350875016521
Villa angostura 0.006348120662795698
Chapelco 0.006348120662795698
Comodoro Rivadavia 0.006122803786363899
Paso de los libres 0.005881595881595884
Villa escobar 0.0053012363012363015
Rawson 0.004386256386256387
Ushuaia 0.003327331982114875
Cosquin 0.002822822822822823
Partido de la costa 0.002576132302138494
Concordia 0.0025710119909161154

Alumna: Paula Brück

Padron: 107533

Viedma 0.0017731241615070545
Santa Ana 0.0015592515592515591
Posadas 0.0015592515592515591
San Juan 0.0005705705705705706
Tandil 0.0005224510224510224
Brazil 0.0
Uruguay 0.0
Chile 0.0
Estados Unidos 0.0
Mexico 0.0
Colombia 0.0
España 0.0
Francia 0.0
Reino Unido 0.0
Otros 0.0
Guaileguay 0.0
Santa Rosa 0.0
Fiambala 0.0
Mina claverio 0.0
humauaca 0.0
Yala 0.0
Salinas Grandes 0.0
Villa Unión 0.0
Cachi 0.0
San luis 0.0
San Miguel de Tucuman 0.0
Corrientes 0.0
Goya 0.0
Peninsula Valdez 0.0
Villa traful 0.0
El calafate 0.0
Valle Barreal 0.0

Alumna: Paula Brück

Padron: 107533

San Francisco de laish 0.0

El Colorado 0.0

Uspallata 0.0

Lujan de cuyo 0.0

El dorado 0.0

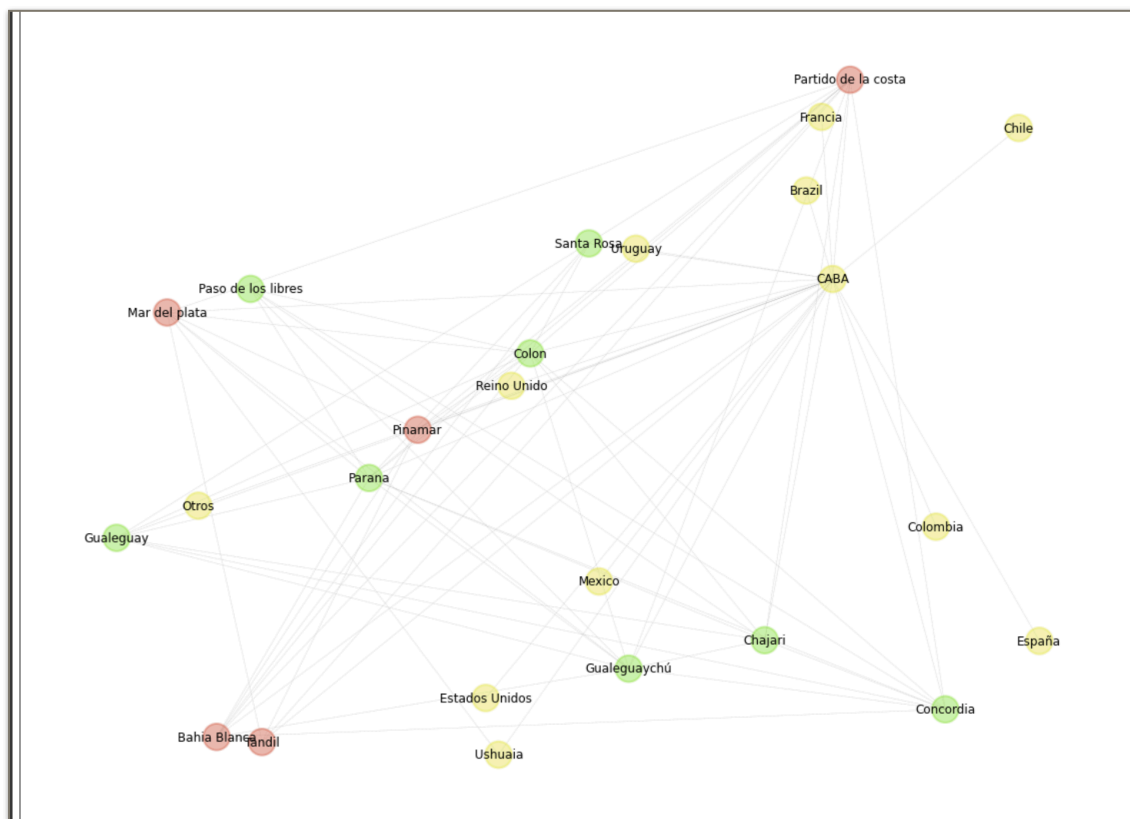
Trapiche 0.0

Moises Ville 0.0

San Carlos 0.0

* Visualización de sub-comunidades:

* Contamos con 5 comunidades por ende 5 subredes. Si analizamos las sub-comunidades para comunidad 1 :



Sub-

Comunidad -----> 0 : 5 nodos (lugares de Provincia de Buenos Aires)

['Pinamar', 'Bahia Blanca', 'Tandil', 'Mar del plata', 'Partido de la costa']

Sub-Comunidad -----> 1 : 12 nodos (Países extranjeros / lugares que se conectan solo con CABA(Ushuaia que esta solo conectada con CABA y MDQ))

Alumna: Paula Brück

Padron: 107533

['Ushuaia', 'Mexico', 'Reino Unido', 'CABA', 'Estados Unidos', 'Uruguay', 'España', 'Chile', 'Francia', 'Brazil', 'Colombia', 'Otros']

Sub-Comunidad -----> 2 : 8 nodos (lugares de Provincia de Entre Rios, Santa Rosa y Corrientes)

['Guaeguay', 'Chajari', 'Parana', 'Guaeguaychú', 'Concordia', 'Paso de los libres', 'Santa Rosa', 'Colon']

Red de Viajes 2 —————> Implementación del PREVIAJE/2020

ANALISIS:

- ***Diametro***
- ***Grado promedio***
- ***Puentes***
- ***Betwness***

Alumna: Paula Brück

Padron: 107533

- *Clustering*
- *Homofilia / Comunidades*
- *Subcomunidades*
- *Roles*
- *Cascadas*

CONCLUSIONES: