Social network analysis project

	Full Name	Department/Section	E-mail
Member 1 Group leader	ضحى عاطف حسن	عربي/احصاء	Doha.atef2018@feps.edu.eg
Member 2	فيرينا صموئيل عيسى	عربي/اقتصاد	Verena.samuel2018@feps.edu.eg
Member 3	سارة محمود	عربي/احصاء	Sarah.mohamed2018@feps.edu.eg

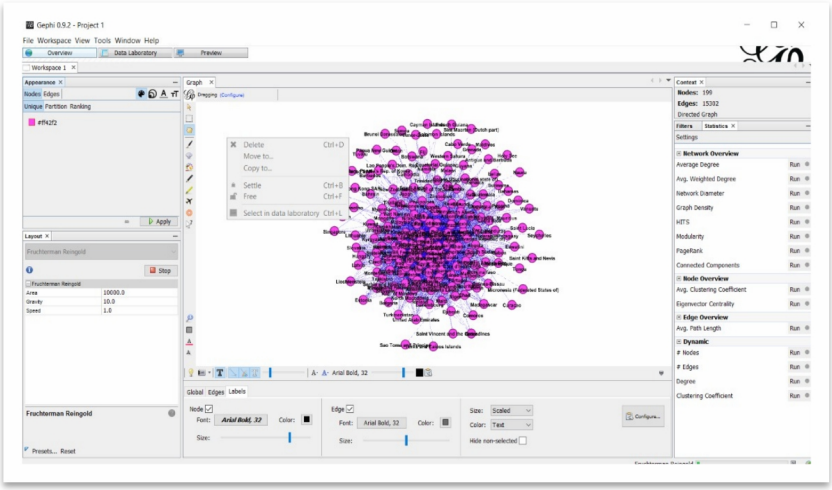
Our data is about the number of applications submitted from the country of origin to the country of asylum and we have separated our data into two files:

-nodes file: which includes all the countries whether it's a country of origin or asylum, and each country has its own unique id.

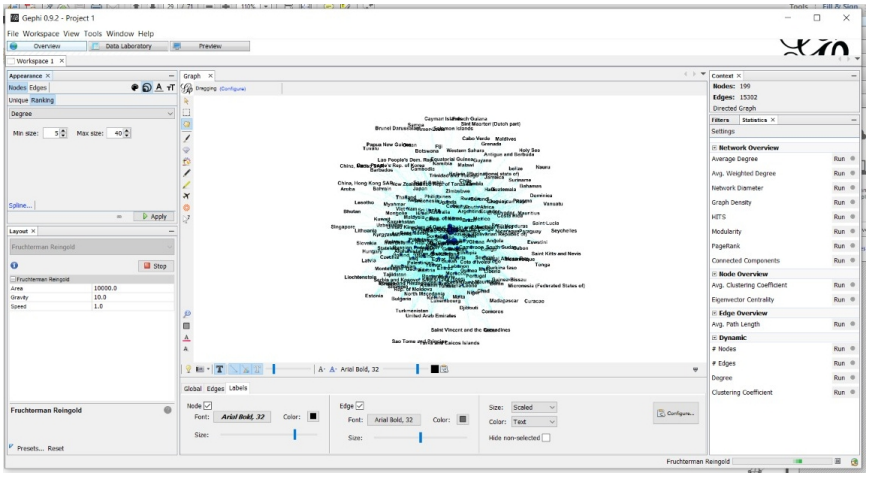
-edges file: which includes the country of origin as a source of refugee's applications, and the country of asylum as the target. And also, the weight for each link between these countries. And we used Gephi to visualize the network and to collect the data

And before we go through the analysis of network properties, we noticed from the data that:

- This network is one mode network because all its nodes are all on the same aggregate level (countries).
- The network is considered to be a directed network as the applications are going only in one direction: from the country of origin to the country of asylum.
- Nodes in this network represent all the countries whether it's a country of origin or asylum.
- As for the edges or links they represent the weighted direct relationship between the refugee's applications from the country of origin to the country of asylum.



The network after fruchterman reingold

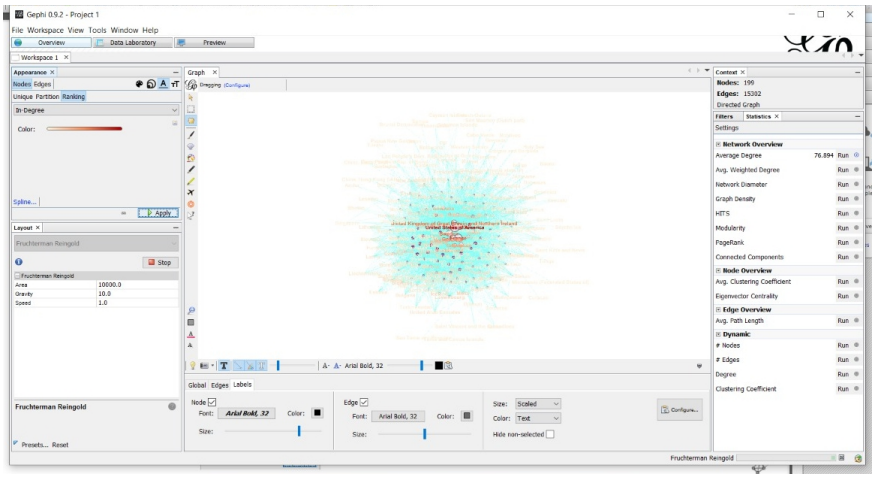


Adjust the range of nodes size

1. **Degree centrality**

It indicates the number of connections (links) each country has in the network. On average, the number of refugees' application flow between countries is 67.89.

- The United States of America is reported to be the most country that receives refugees' applications, with the highest degree centrality all over the network (=1260).
- USA (1186), France (999), and Sweden (751) has the highest in-degree centrality, which means that the number of received refugees' applications (asylum countries).
- While Syria (345), Afghanistan (304), and Pakistan (293) have the highest out-degree centrality which means that the number of applied refugees' applications to other countries (origin countries).

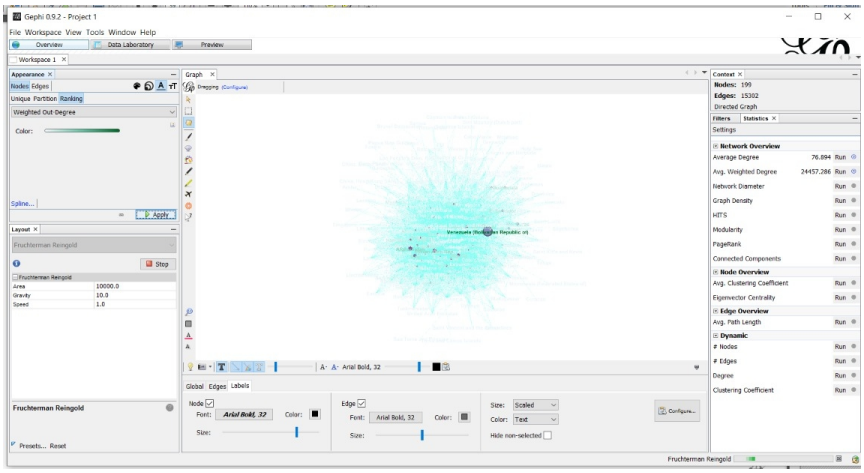


Degree centrality

Weighted Degree Centrality:

The degree here depends on the weights of the links (edges) between countries, and not just the number of the links between them, as it measures how strong the centrality of each country is. The weights here are represented by the number of refugees applied for and to each country in the network.

- The average applied refugees' application flow from the country of origin to the country of asylum is 24457 applications.
- The highest weighted degree all over the network is Venezuela (868564).
- Venezuela (866846), Afghanistan (288828), and Syria (257353) have the highest number of out-going applications (out-degree).
- While USA (719522), Germany (556708), and France (433023) have the highest incoming applications in the network (in-degree).

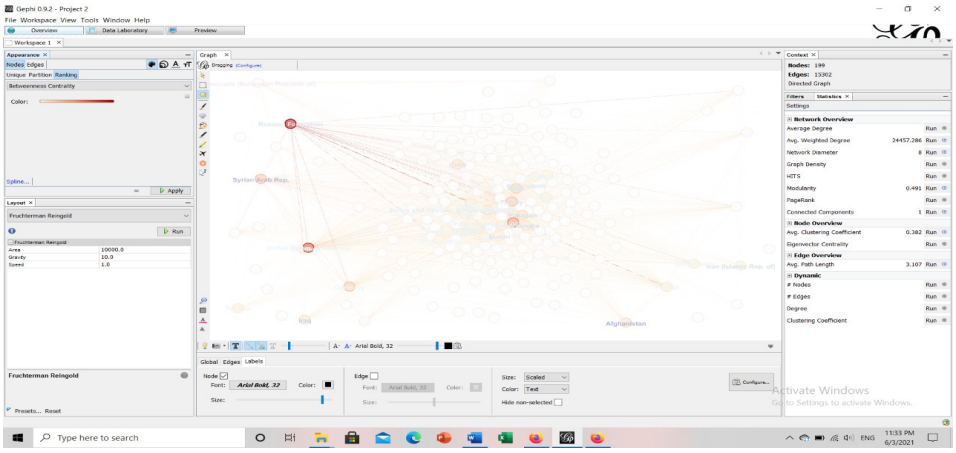


Weighted degree centrality

2. **Betweenness Centrality**

   Measures how often a country appears on the shortest paths in the network.

- It appears that Russia Federation (11414), USA (8802.07), and Egypt (5539.37) has the shortest paths among the network and are considered the mediator actors in the refugees' flow all over the network.

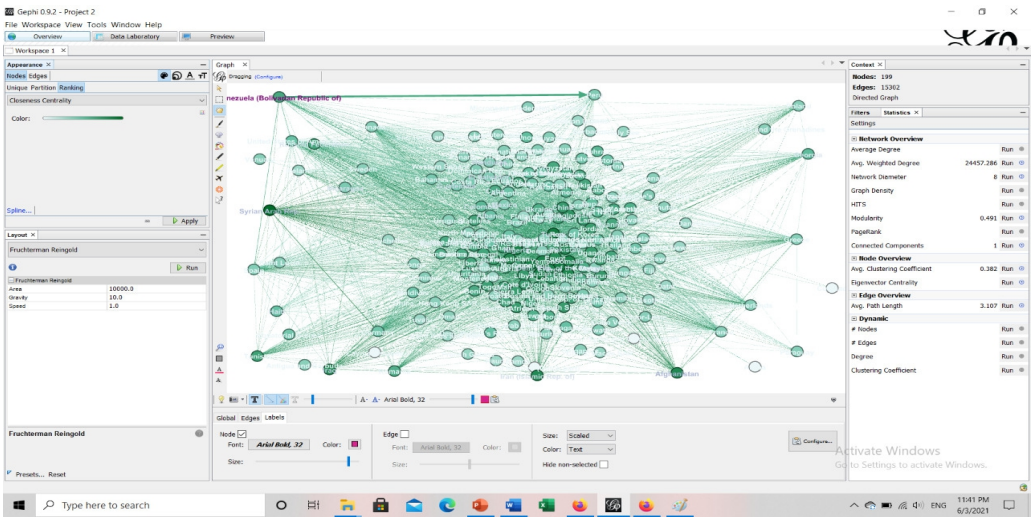


Betweenness centrality

Closeness Centrality

It measures the average distance from a given starting country to other countries in the network.

- The country that has High closeness centrality in this network is Syria (0.71) which also has a high weighted out-degree, which means that it has the most effective influence among the network.
- Whereas a country like Sudan (0.636) has a relatively high closeness centrality, but relatively low weighted out-degree (74915), which means its influence is not that impactful over the network.



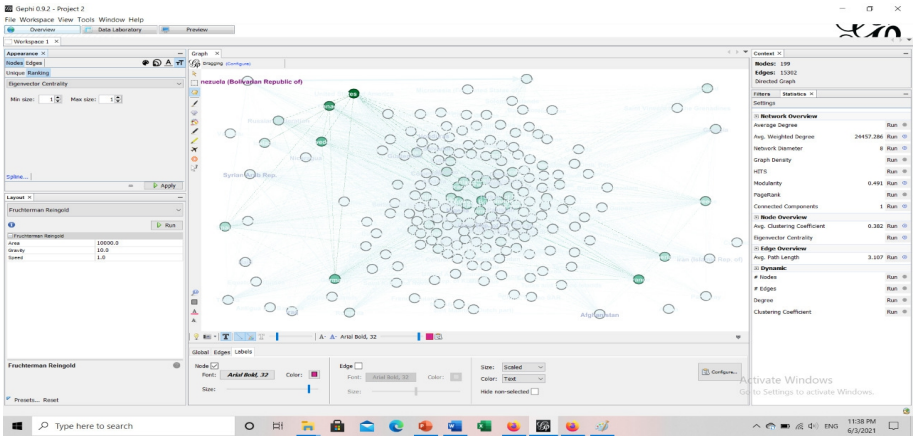
Closeness centrality

Eigenvector Centrality

It is a measure of the influence that a country has on a network. If a country has a flow of applications with the most central country in the network, then this country has many connections all over the network.

- The country that has the highest eigenvector score is the USA (=1) which indicates that it has a leading role in the flow refugees' applications all over the network. In

addition to that, it has a high in-degree centrality therefore it's an important asylum country in the network.



Eigenvector centrality

3. Eccentricity and Network Diameter

It indicates the maximum distance from one country to any other country in the network. While the network diameter measures the maximum distance between all countries in the network.

- Highest eccentricity is the network diameter, which means that the furthest flow of refugees among the network equals 8 links and the highest eccentricity country in the network is Australia and China, Macau (=8) "so far away", while the lowest eccentricity country is Cayman Islands "quite close".

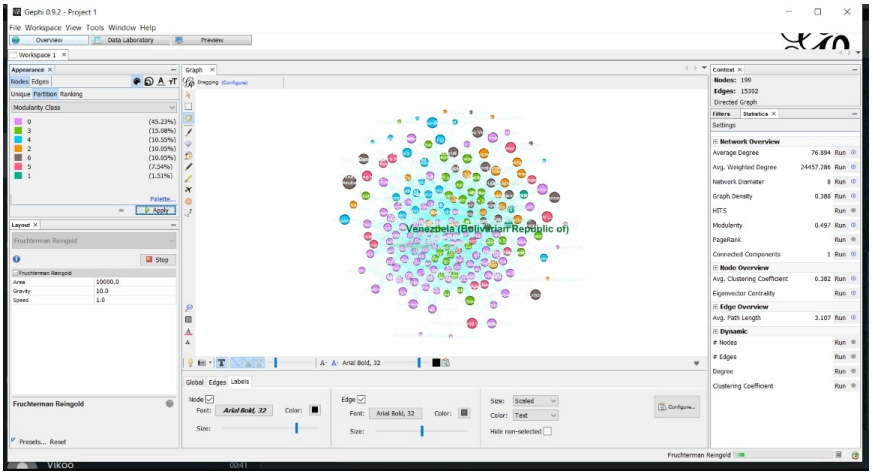
Graph density

This network has 38.8% density, which means that all possible refugees' flow among the network were represented, hence network is considered to be a sparse network and the network is not completely connected.

4. Modularity

It is a measure of the structure of the network, and the strength by which the network can be divided with into modules, while in each modules the countries are strongly connected, but there is a sparse connection between countries from different modules.

- this network has the ability to divide into 6 small groups by 49%

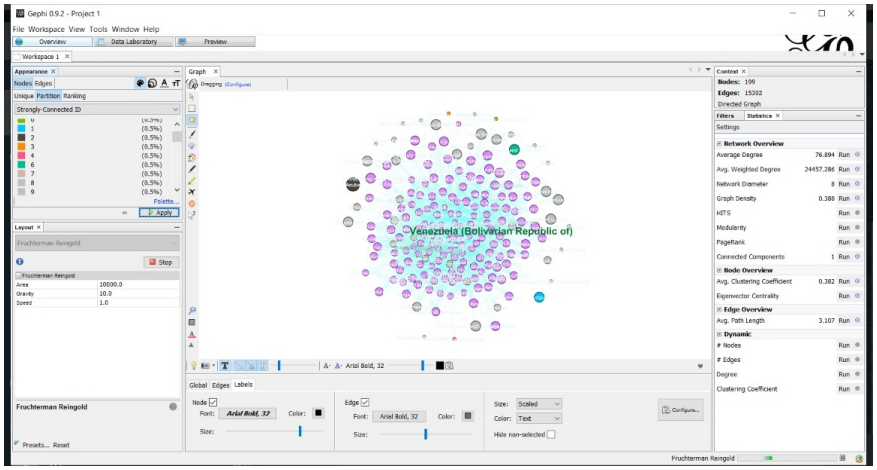


Modularity

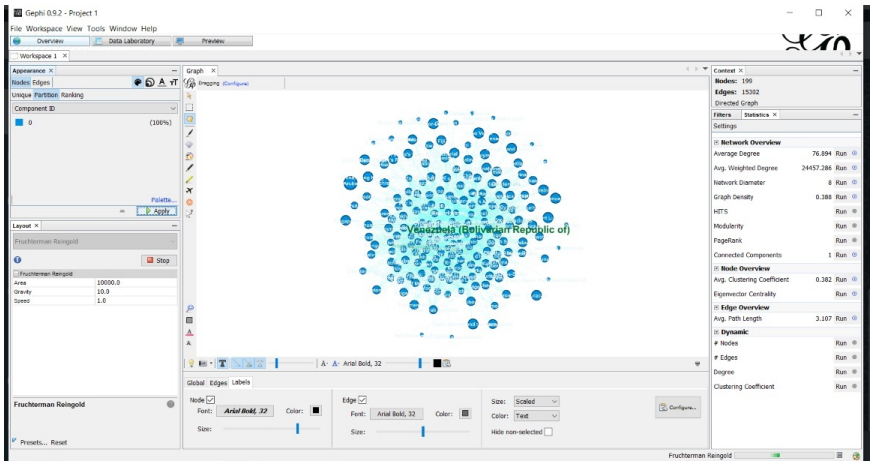
Connected Components

Components connected by in and out edges are strongly connected, while components connected by in edges only or out edges only are weakly connected.

- The strongly connected countries: Western Sahara (42), Vanuatu (41), Uzbekistan (40) and Tuvalu (39) receive asylum applications, and send the applications received to other countries in the network to host asylum seekers which help regulate the flow of refugees' applications through the network.



Strongly component ID

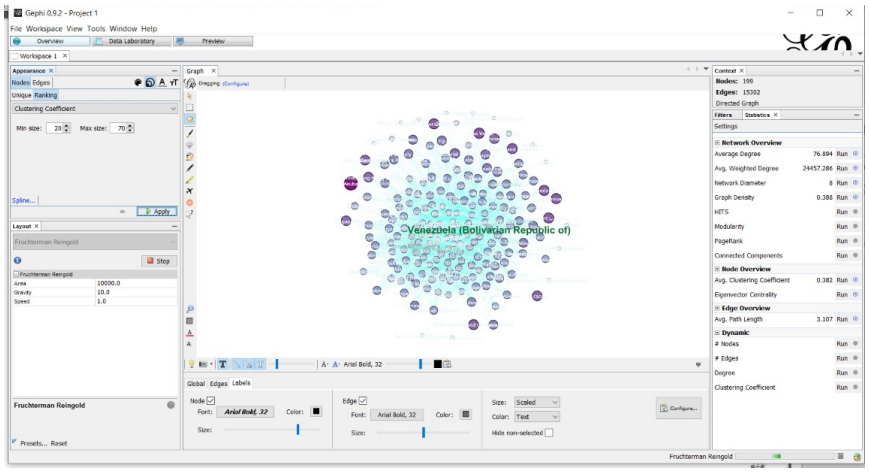


Component ID

5. Clustering Coefficient and the Average Path Length

The clustering coefficient measures how well a country is connected to its neighbors, while the average path length calculates the average distance between all pairs of countries.

- This network has a relatively low clustering coefficient which equals 0.38, While having a relatively high average path length which equals 3.107
- This interpretation indicates that **the small world phenomena** does not comply on this network, as the path length between countries is quite large and the ability to create a cluster between countries is quite small.



Clustering coefficient