**Assignment 1**

Name: Vansh Lal Tolani

Roll no.:22BDS061

**Detailed Report on the Structural Properties of the Network for Engineering/Architectural Firms in World Cities**

**Dataset Overview**

The dataset "Engineering/Architectural Firms in World Cities: Nodal Size and Network Connectivity" is derived from the GaWC (Globalization and World Cities) Research Network. It provides insights into the distribution and network connectivity of engineering and architectural firms across various global cities.

The dataset is structured to capture the nodal size (importance or connectivity) of cities in a global network based on the presence and operations of these firms. It helps to understand how different cities are interconnected through the activities of engineering and architectural firms, highlighting the significance of certain cities in the global urban hierarchy.

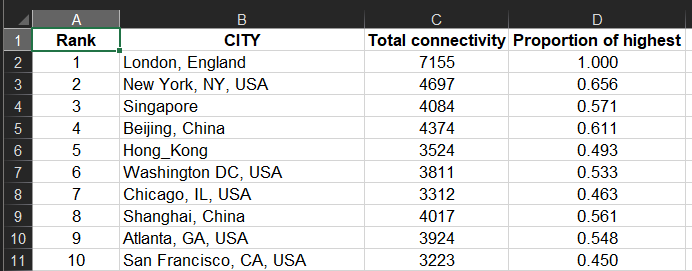
**Nodes** = Cities where engineering/architectural firms have offices. (856)

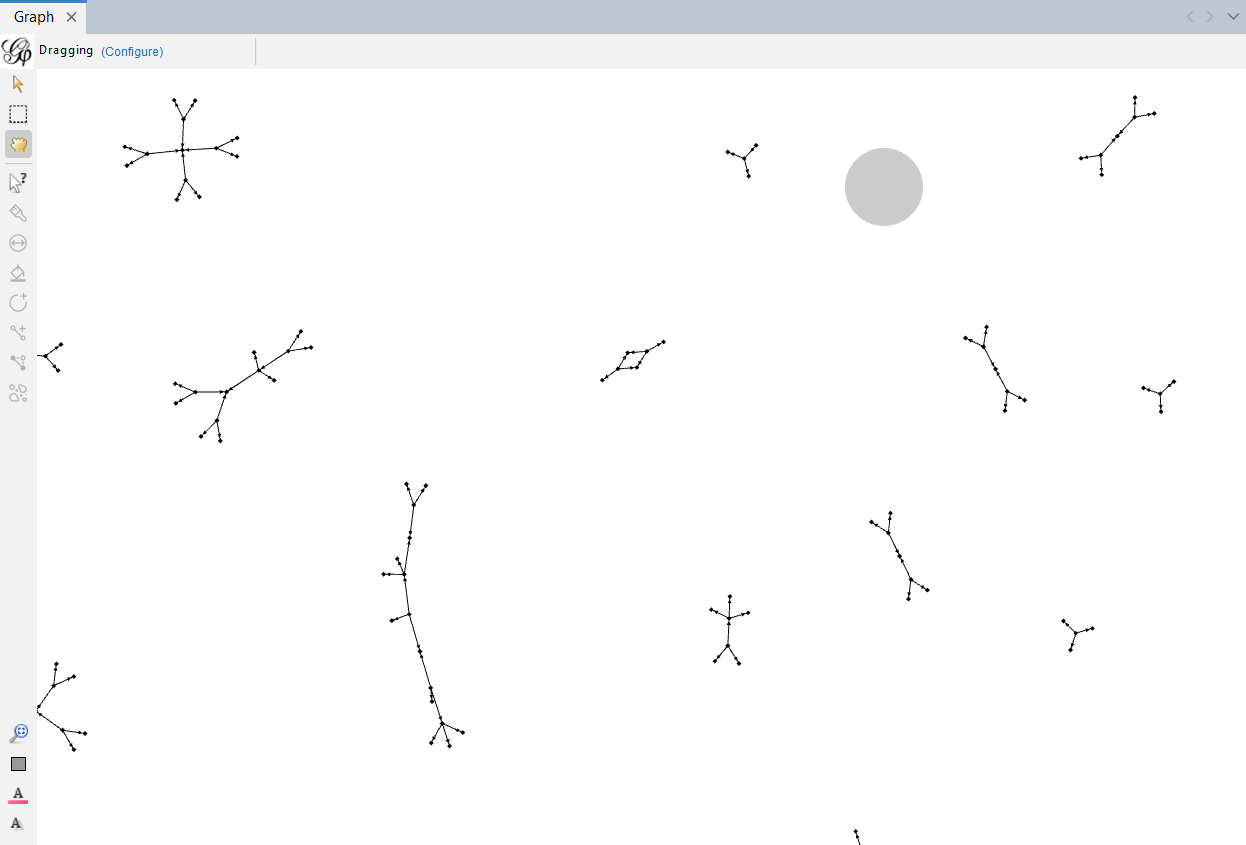
**Edges** = Connections between these cities, established by the presence and operations of firms in multiple locations. (705)

**Source of Dataset**: https://www.lboro.ac.uk/microsites/geography/gawc/datasets/da14.html.

**Key Attributes of the Dataset**

* Rank: This attribute ranks the cities based on their connectivity or importance within the network. The ranking is likely based on the "Total Connectivity" metric, where a higher rank indicates a higher level of connectivity.
* City: This attribute lists the names of the cities that are part of the network. Each city is considered a node in the network graph.
* Total Connectivity: This metric represents the overall connectivity of a city within the network. It could be a measure of the number of firms operating in the city, the strength of the connections with other cities, or the city's role as a hub in the network.
* Proportion of Highest: This attribute shows how the city's connectivity compares with the most connected city in the network, expressed as a proportion or percentage. It helps to visualize the relative importance of each city.



****

**The Actual Graph where the Components are farly placed indicating clustering coefficient 0.**

**2. Structural Properties of the Network**

**Network Diameter:**

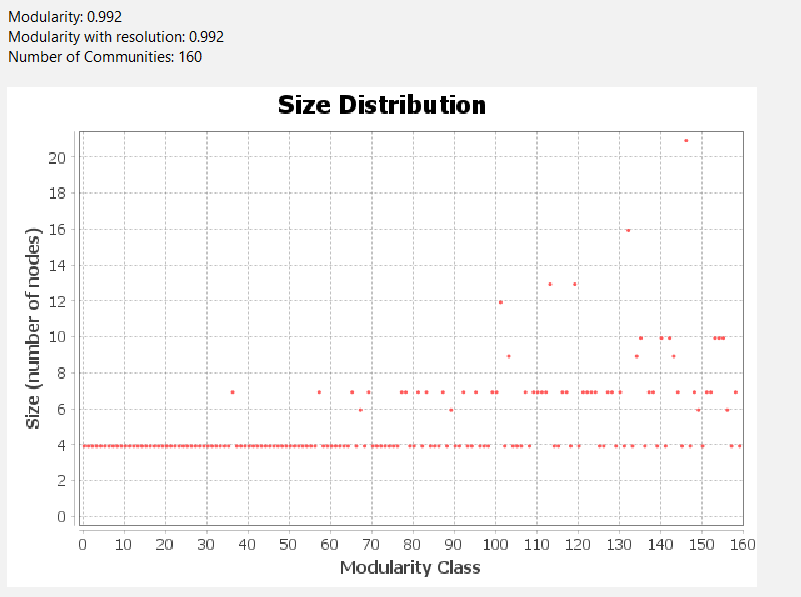
* **Value: 3**
* I**nterpretation:** The network diameter of 3 indicates that the longest shortest path between any two nodes in the network is just 3 steps. This small diameter suggests a highly interconnected network where cities are closely linked through the presence of engineering and architectural firms.

**Graph Density:**

* **Value: 0.001**
* **Interpretation:** The extremely low graph density of 0.001 signifies that only a tiny fraction (0.1%) of all possible connections between cities and firms are present. This sparse connectivity indicates that while cities may be connected, the number of direct links between them is minimal, reflecting the selective and strategic placement of offices by firms in key global cities.

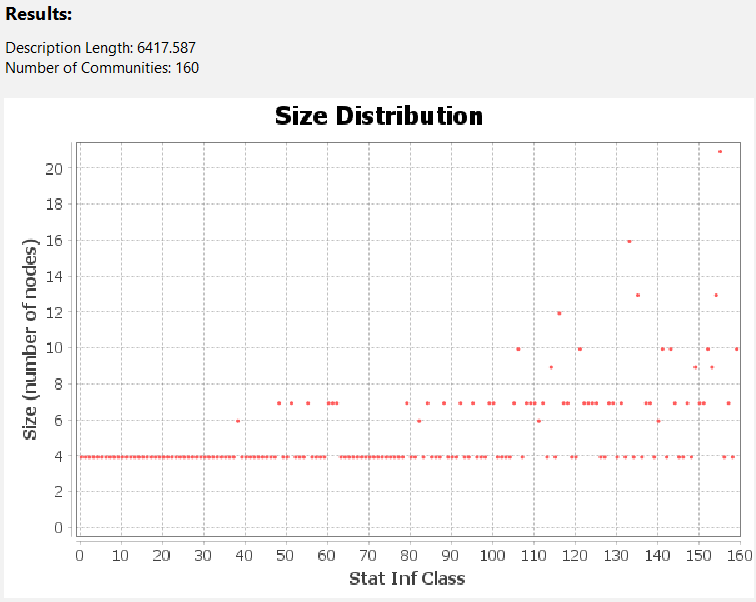
**Modularity:**

* **Value: 0.992**
* **Interpretation:** The very high modularity score of 0.992 suggests that the network is composed of highly distinct communities or clusters. These communities represent groups of cities that are more densely connected to each other than to the rest of the network.

****

**Statistical Inference:**

* **Value: 6417.587**
* **Interpretation:** This value likely represents an aggregate measure derived from the network's structural properties.



**Average Clustering Coefficient:**

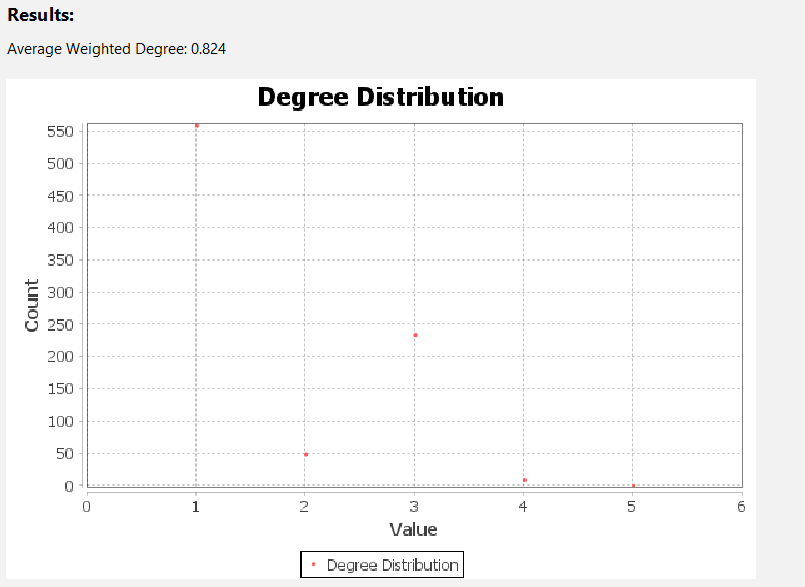
* **Value: 0**
* **Interpretation:** The average clustering coefficient of 0 indicates that there is no clustering within the network, meaning that the network lacks tightly connected triads. In practical terms, if a city A is connected to city B, and city B is connected to city C, it is unlikely that city A is also directly connected to city C. This absence of clustering reflects the sparsity of the network and may suggest that cities operate in isolated clusters with minimal overlap.

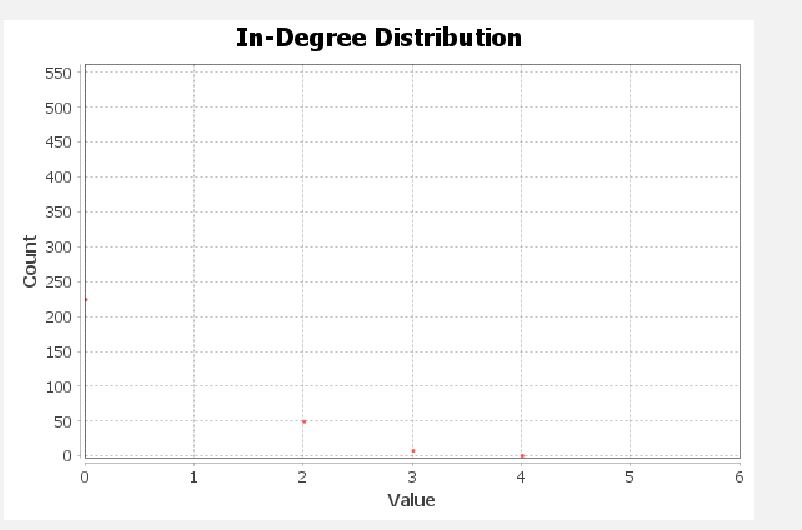
**Average Path Length:**

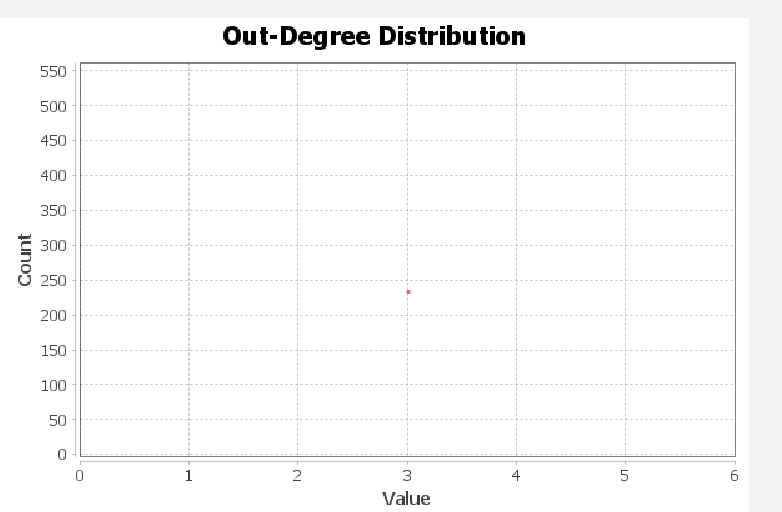
* **Value: 1.049**
* **Interpretation:** The low average path length of 1.049 shows that most cities are separated by only one intermediary city, making the network highly efficient in terms of connectivity. Despite the low density, the short path length indicates that cities are generally very close to each other in network terms, facilitating quick information flow and collaboration across the network.

**Average Degree:**

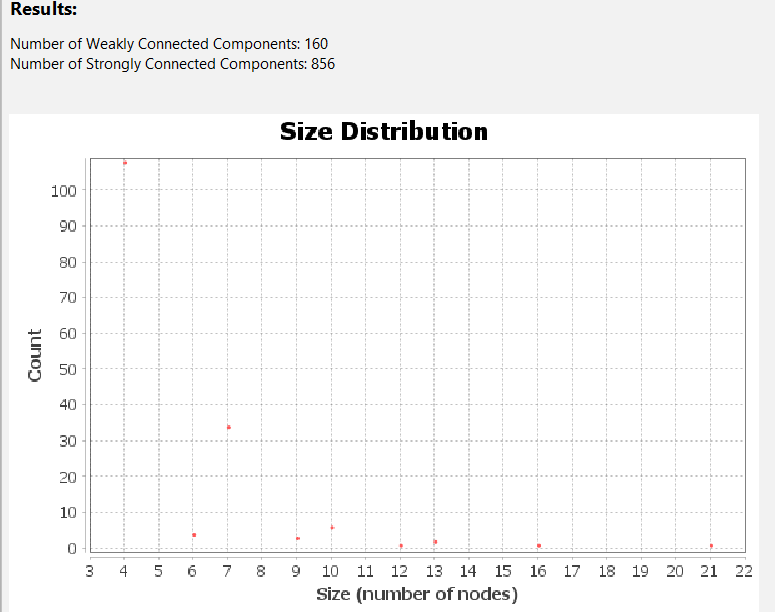
* **Value: 0.824**
* **Interpretation:** The average degree of 0.824 indicates that, on average, each city is connected to less than one other city. This reflects a sparse network where most cities have few connections, suggesting limited direct interaction between cities.

****

**In Degree:** The network exhibits a hierarchical structure, with 66.36% of cities having a minimal weighted in-degree of 1 and 26.40% entirely isolated. Only a few cities act as influential hubs, with higher weighted in-degrees of 2, 3, or 4, indicating their central role in the global network of engineering and architectural firms. Most cities remain on the periphery, with limited influence or connectivity.

Out Degree: The network shows a clear division between a passive majority, with 72.54% of cities having a weighted out-degree of 0, and an active minority, with 27.45% of cities making connections with three others. This minority drives the network's interaction and connectivity, highlighting a concentration of influence in a few key cities, while most remain passive or disconnected.

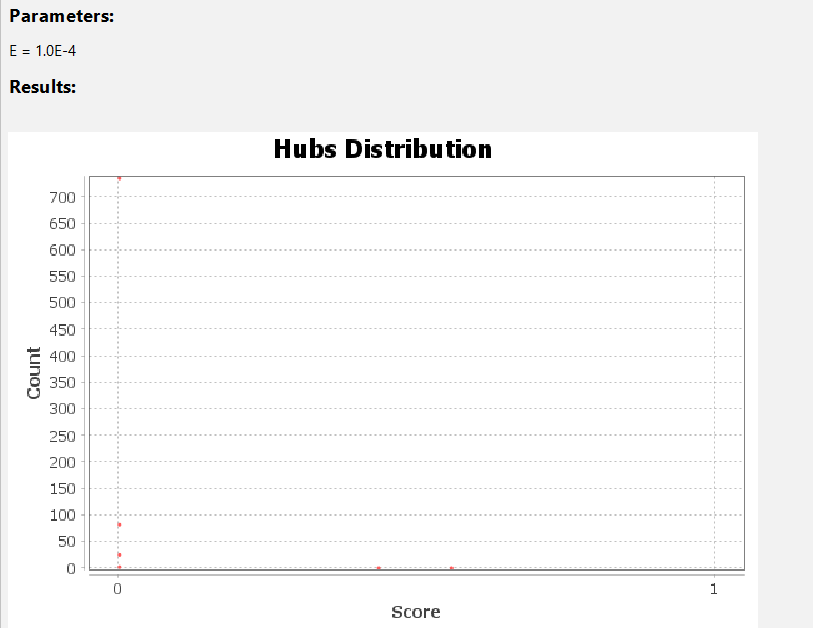
**Connected Components:**   
The presence of 160 connected components in the network indicates a high level of fragmentation, with many isolated clusters of cities that are not connected to each other. This suggests that the network is composed of numerous small, independent groups of cities with limited or no interaction between them, further highlighting the sparse connectivity and lack of cohesive integration across the global network of engineering and architectural firms.

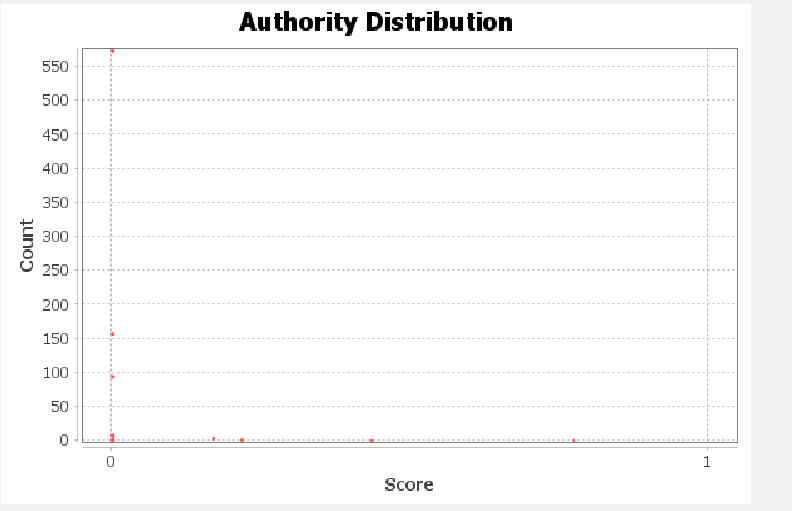
****

**3. Centrality Measures**

**Hub and Authority Scores:**

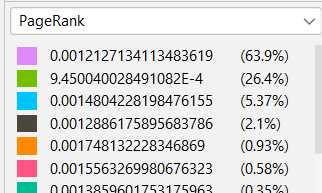
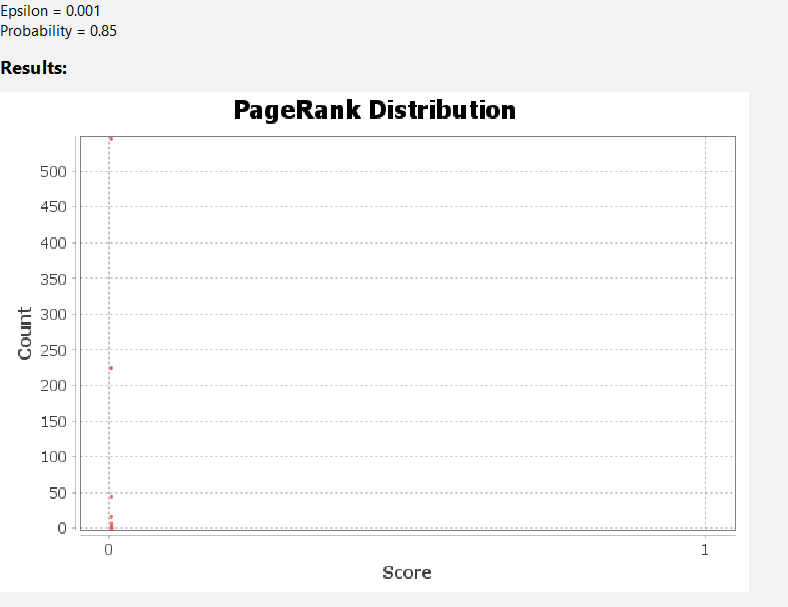
* **Hub Scores:**
  + The distribution of hub scores is heavily skewed towards zero, indicating that only a few cities serve as major connectors or hubs, linking to many authoritative firms. These hubs are likely critical nodes within the network, driving much of the global connectivity.

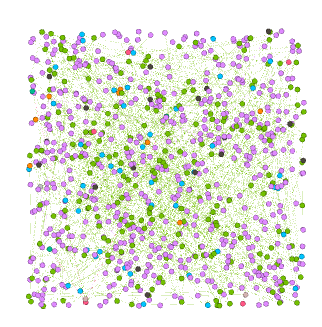


* **Authority Scores:**
  + The distribution of authority scores is similarly concentrated at the lower end, with only a few cities being recognized as major authorities within the network. These cities might host key headquarters or primary offices of major firms, establishing their significance in the global engineering and architectural landscape.
  + 

**PageRank:**

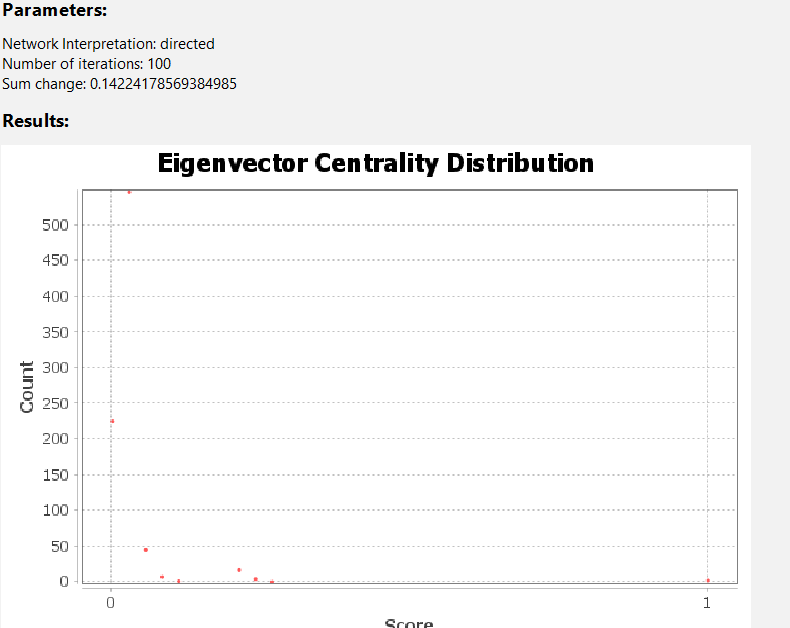
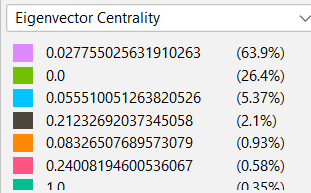
* **Distribution:**
  + The PageRank distribution is expected to mirror that of the hub and authority scores, with a few cities having high PageRank values, signifying their importance in terms of connectivity and influence within the network. A skewed distribution would further underscore the hierarchical nature of the network, where a few cities dominate in terms of global influence.
  + The PageRank distribution shows that the majority of cities (547 elements) have a very low PageRank of 0.0012, and another significant group (226 elements) has an even lower PageRank of 9.5x10⁻⁴. This suggests that most cities in the network have minimal importance and influence, with only a few connections to more influential nodes. The network appears to be dominated by a small number of central cities, while the majority of cities have limited roles in the global network of engineering and architectural firms.

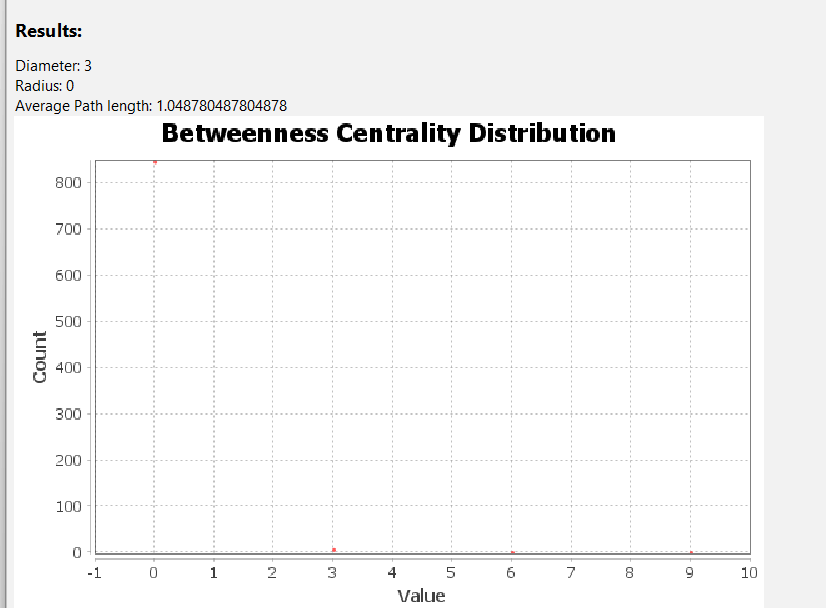
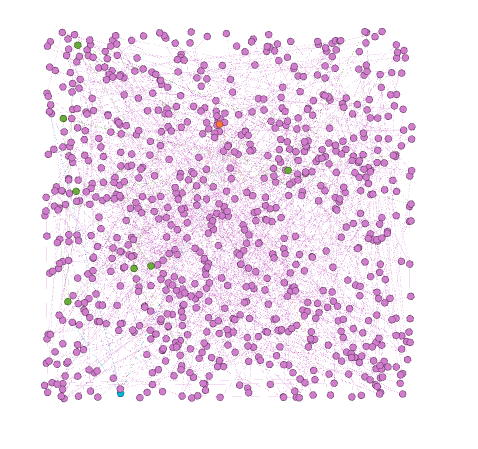


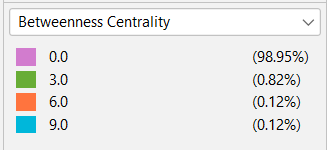
****

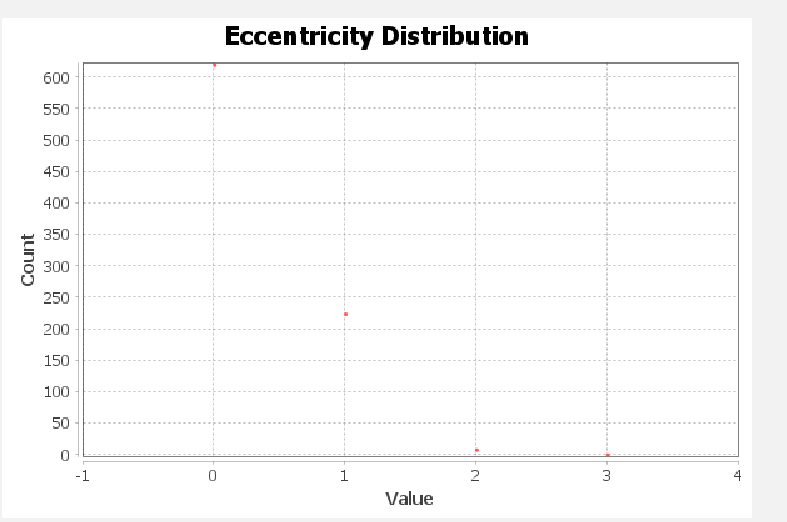
**Eigenvector Centrality:**

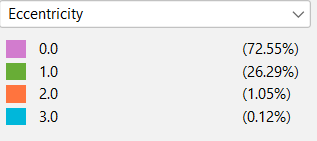
* **Change: Sum change: 0.14224178569384985**
* **Interpretation:** The eigenvector centrality indicates the influence of a city based not just on its direct connections but also on the importance of the nodes it is connected to. The stable sum change suggests that the central cities in the network remain consistently influential across different iterations, likely representing the major global hubs of engineering and architectural firms**.**
* The low eigenvector centrality for most cities indicates that influence is concentrated in a few central cities, while the majority are peripheral and less connected. This suggests a core-periphery network structure, where a small number of key cities dominate the global network, and many cities remain less influential and more isolated.

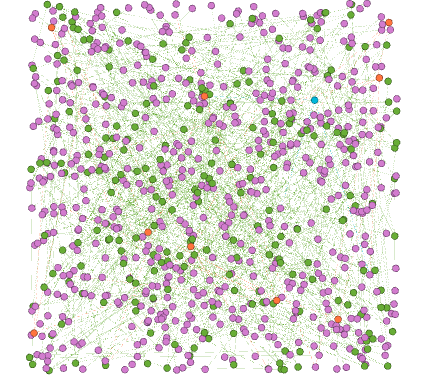
**** ****

**Betweenness Centrality:**The betweenness centrality distribution in this network reveals that the vast majority of elements (847 out of 856) have a centrality of 0, indicating that they do not play a significant role in facilitating connections between other nodes. Only a small number of elements (7) have a betweenness centrality of 3, and just 2 elements have higher values of 6 and 9, respectively. This suggests a highly centralized structure where only a few key cities or firms act as critical intermediaries in the global network of advanced producer service firms.

****

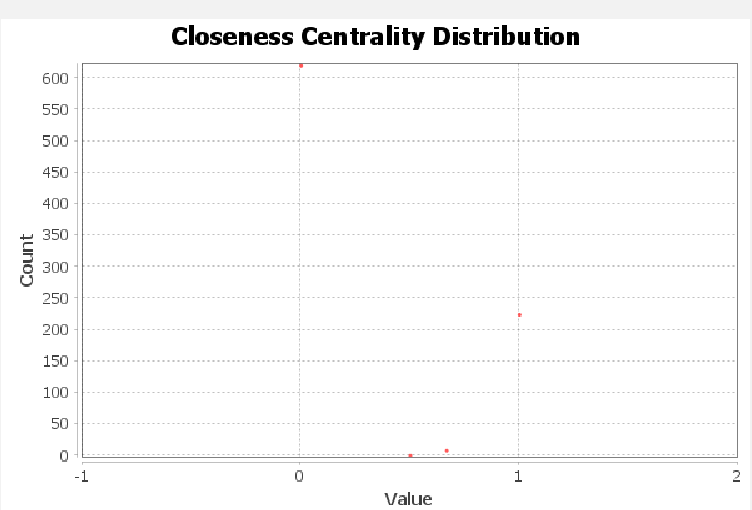
**Eccentricity:** The network is highly fragmented, with 72.57% of nodes either isolated or in small, disconnected components. A smaller, more connected portion (26.29%) is central to the network, likely serving as key hubs. The few nodes with higher eccentricity act as connectors or bridges. This reflects a hierarchical structure, where central cities dominate, while most others remain on the periphery with limited influence.

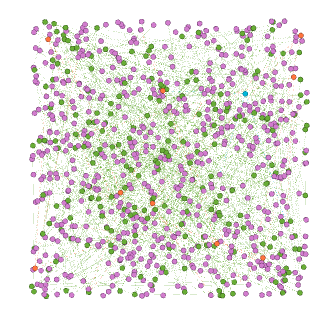
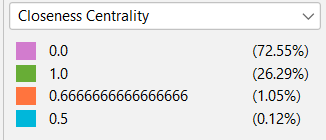


****

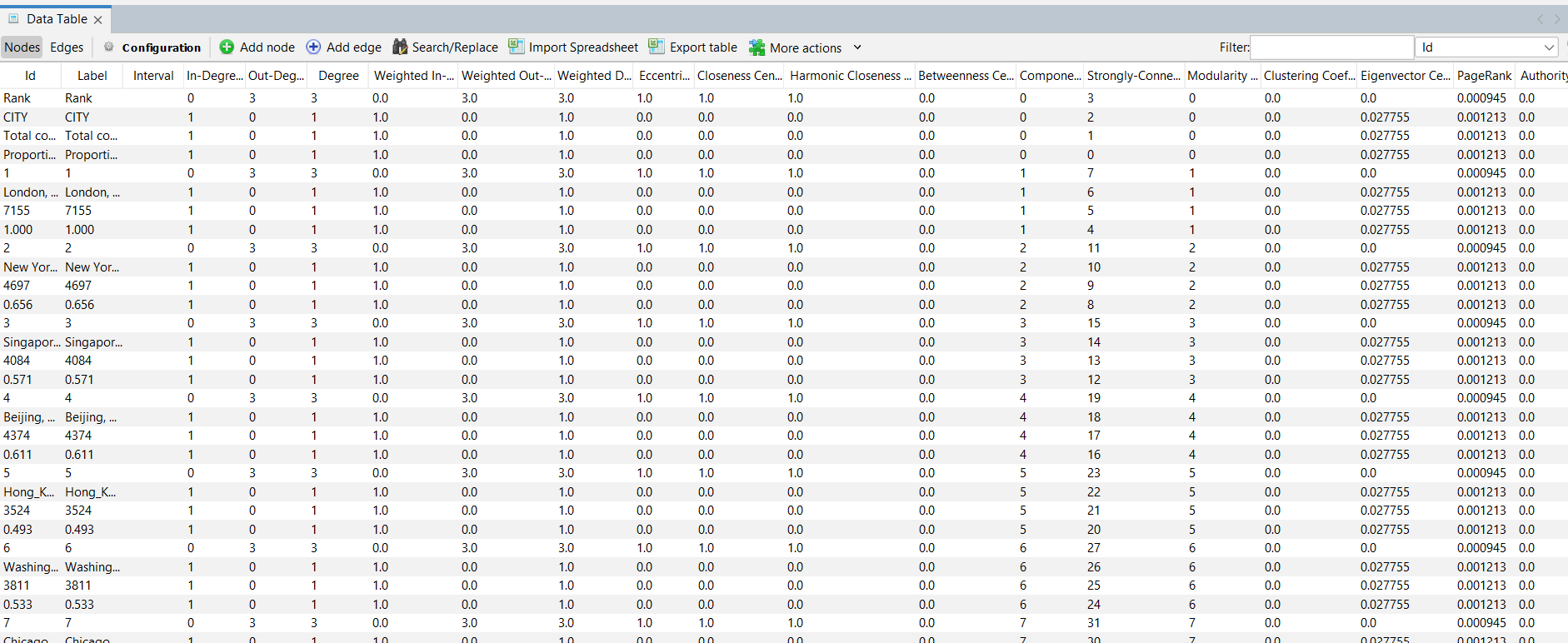
**Closeness Centrality:**

The distribution shows that the majority of cities (621 elements) have a closeness centrality of 0, indicating they are disconnected or poorly connected within the network. A smaller group (225 elements) has a closeness centrality of 1, representing a well-connected core. The presence of just a few cities with intermediate values (0.6 and 0.5) suggests that there are very few bridging cities between the core and the periphery, emphasizing a highly polarized network structure.



**Data Laboratory**

****

**4.Summary and Implications**

* The "Engineering/Architectural Firms in World Cities" network is characterized by sparse but highly efficient connectivity, with a small-world property evidenced by its short diameter and low average path length. The network's high modularity reveals the presence of distinct clusters, likely representing regional or sectoral groupings of cities that dominate in specific areas of engineering and architecture.
* Centrality measures indicate a hierarchical network structure, with a few key cities acting as global hubs and authorities, driving the connectivity and influence within the network. These cities are critical for maintaining the network's overall structure, making the network potentially vulnerable to disruptions if these key nodes are targeted.
* The absence of clustering and the network's sparse nature suggest that cities within the same cluster may operate independently, with little overlap or interaction with cities outside their immediate community. This structure highlights the importance of strategic connections between clusters to ensure the resilience and global reach of the engineering and architectural firms within this network.
* Overall, the network reflects the strategic placement and global operation of engineering and architectural firms, with key cities serving as the main drivers of connectivity and influence within this highly modular and efficient global network.