GRAPH ANALYTICS

AND

ALGORITHMS – 18MAT333

CASE STUDY

FOOD WEB ANALYSIS

ROLL NUMBER	NAME
CB.SC.I5DAS19015	DHANUSHRI M
CB.SC.I5DAS19018	KEERTHANA PRIYA D

ABSTRACT:

- A food web is the natural interconnection of food chains and a graphical representation of what-eats-what in an ecological community.
- Food web is a complex network which is analyzed to understand the biodiversity, species interactions (relationship between the predator and prey), and ecosystem structure.
- Human action has induced strong pressures over time periods which makes it too short for the ecosystems to adapt which causes a series of changes in the food web.
- Network analysis on food web helps biologists to understand the key species and the species that are endangered and in the verge of extinction.

DATASET DESCRIPTION:

- Food web network contains:
  - NODE - Species (predators as source and preys as target)
  - Each species is represented using label which are unique numbers, along with their names and to which group they belong to.
  - EDGE - Relationship between the prey and the predator.
- It is an unweighted directed graph as it shows the relationship between prey and predator.

	label	id	name	group
0	0	0	2um Spherical Phytoplankt	unknown
1	1	1	Synedococcus	unknown
2	2	2	Oscillatoria	unknown
3	3	3	Small Diatoms (<20um)	unknown
4	4	4	Big Diatoms (>20um)	unknown
5	5	5	Dinoflagellates	unknown
6	6	6	Other Phytoplankton	unknown
7	7	7	Benthic Phytoplankton	Demersal Producer
8	8	8	Thalassia	Seagrass Producer
9	9	9	Halodule	Seagrass Producer

NUMBER OF NODES:

```
print("Number of nodes: ",vertices.shape[0])
```

Number of nodes: 128

source	target
71	105
57	122
123	40
61	116
1	124
35	88
99	104
70	115
64	127
13	64

NUMBER OF EDGES:

```
print("Number of edges: ",edges.shape[0])
```

Number of edges: 2106
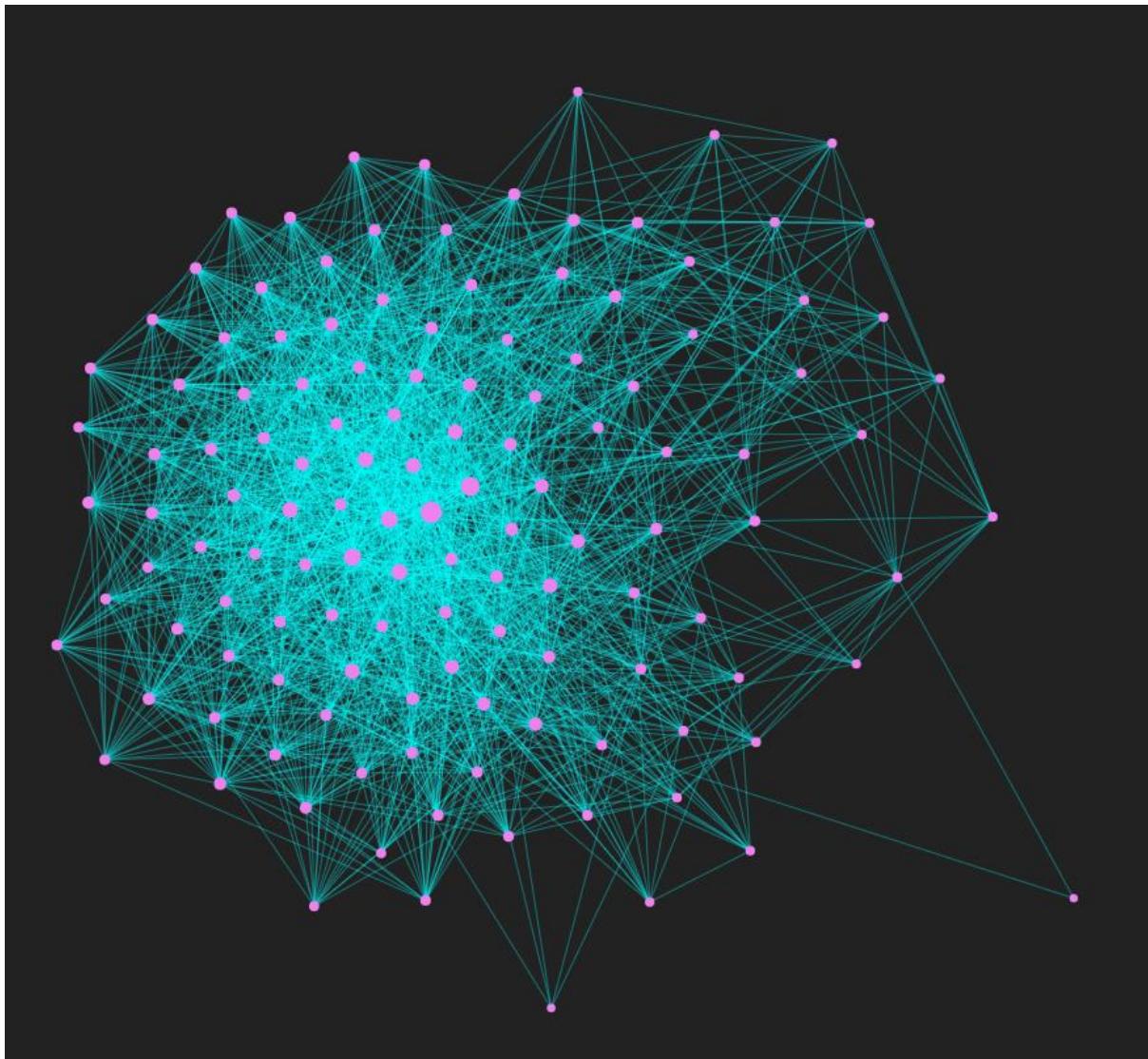
|	source	target
0	125	0
1	125	1
2	125	2
3	125	3
4	125	4
5	125	5
6	125	6
7	125	7
8	125	8
9	125	9

FOOD WEB NETWORK VISUALIZATION USING PYVIZ:
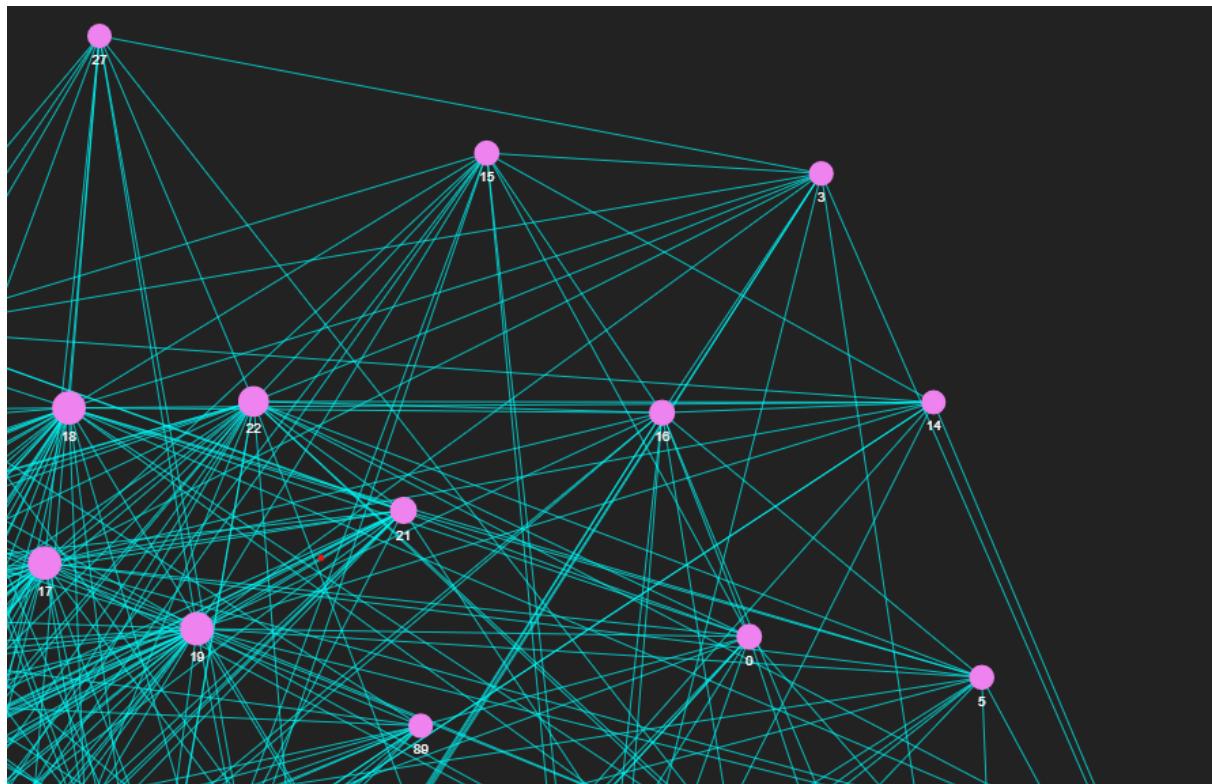
Pyviz is a Python library that enables visualizing and interactively manipulating network graphs.

Nodes – Violet

Edges – Cyan

Each of these nodes are labelled with unique number (label attribute) according to the species they belong to.

NEIGHBOUR NODES:

Consider the node labelled as "3" which is "Small Diatoms" belonging to an Unknown species group.

```
vertices.loc[vertices['label']=='3']
```

	label	id	name	group
3	3	3	Small Diatoms (<20um)	unknown

"Node 3" has 13 neighbors.

13 neighbors of Node 3 are:

|	source	target	source_name	source_group	target_name	target_group
3	125	3	Input	Detritus	Small Diatoms (<20um)	unknown
17	3	15	Small Diatoms (<20um)	unknown	Water Flagellates	Microbial Microfauna
22	3	16	Small Diatoms (<20um)	unknown	Water Cilitaes	Microbial Microfauna
32	3	17	Small Diatoms (<20um)	unknown	Acartia Tonsa	Zooplankton Microfauna
43	3	18	Small Diatoms (<20um)	unknown	Oithona nana	Zooplankton Microfauna
54	3	19	Small Diatoms (<20um)	unknown	Paracalanus	Zooplankton Microfauna
64	3	20	Small Diatoms (<20um)	unknown	Other Copepoda	Zooplankton Microfauna
79	3	22	Small Diatoms (<20um)	unknown	Other Zooplankton	Zooplankton Microfauna
98	3	26	Small Diatoms (<20um)	unknown	Sponges	Macroinvertebrates
106	3	27	Small Diatoms (<20um)	unknown	Coral	unknown
153	3	30	Small Diatoms (<20um)	unknown	Bivalves	Macroinvertebrates
194	3	36	Small Diatoms (<20um)	unknown	Suspension Feeding Polych	Macroinvertebrates
1946	3	124	Small Diatoms (<20um)	unknown	DOC	unknown
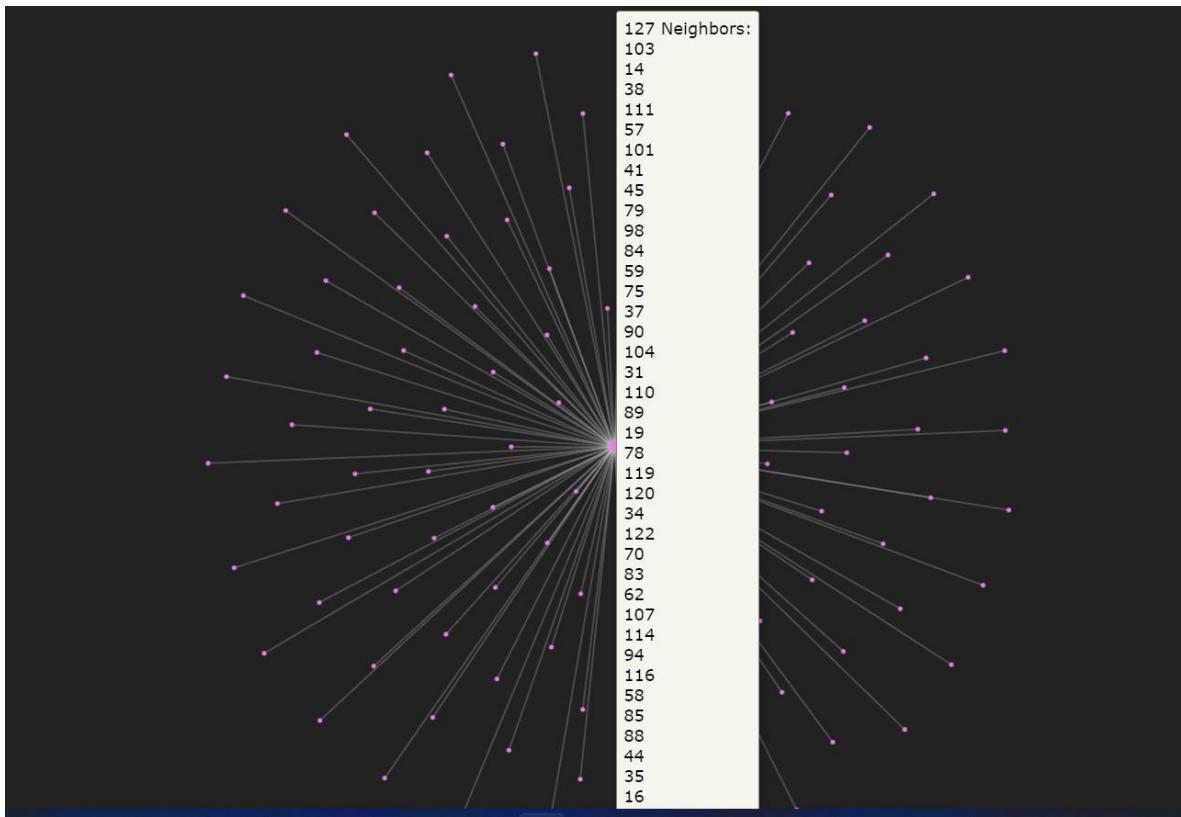
KEYSTONE SPECIES:

An organism without which the ecosystem would be dramatically different or cease to exist altogether is keystone species.
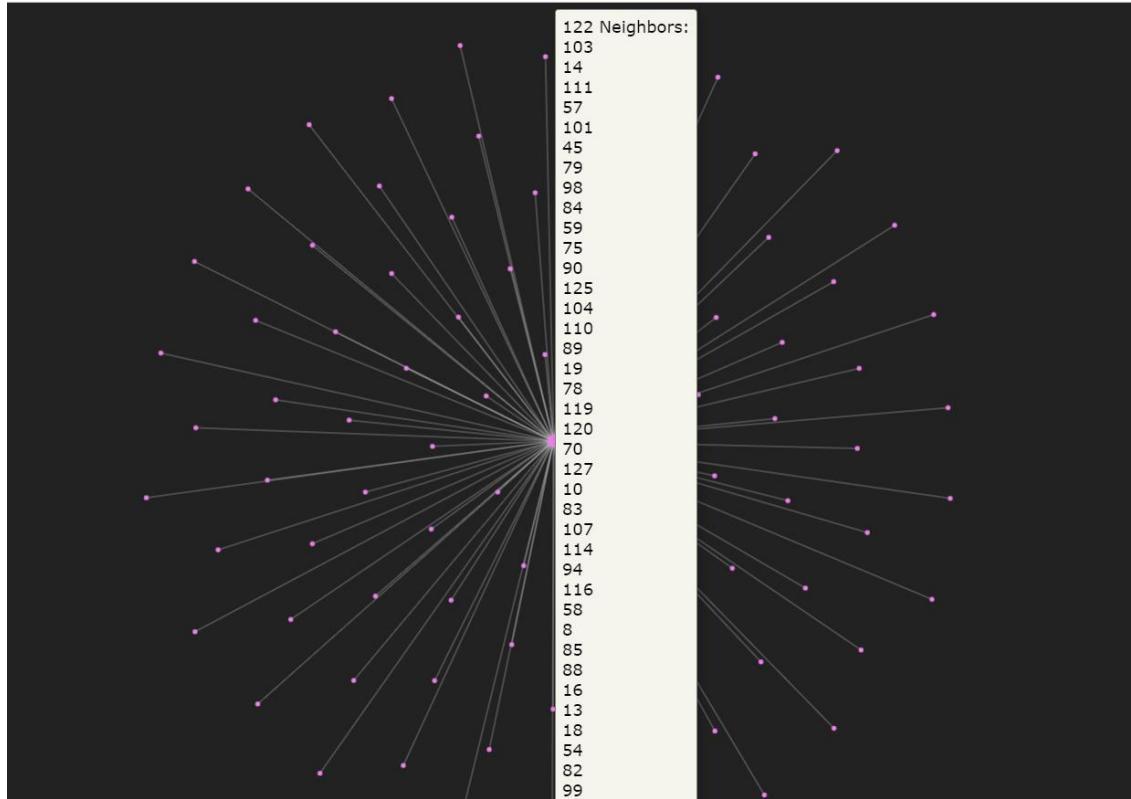
This can be analyzed by degree. The node that has the most degree can be said as the keystone species.

TOP 3 KEYSTONE SPECIES:

- NODE 127:

127 Neighbors:
103
14
38
111
57
101
41
45
79
98
84
59
75
37
90
104
31
110
89
19
78
119
120
34
122
70
83
62
107
114
94
116
58
85
88
44
35
16

- NODE 122:

122 Neighbors:
103
14
111
57
101
45
79
98
84
59
75
90
125
104
110
89
19
78
119
120
70
127
10
83
107
114
94
116
58
8
85
88
16
13
18
54
82
99
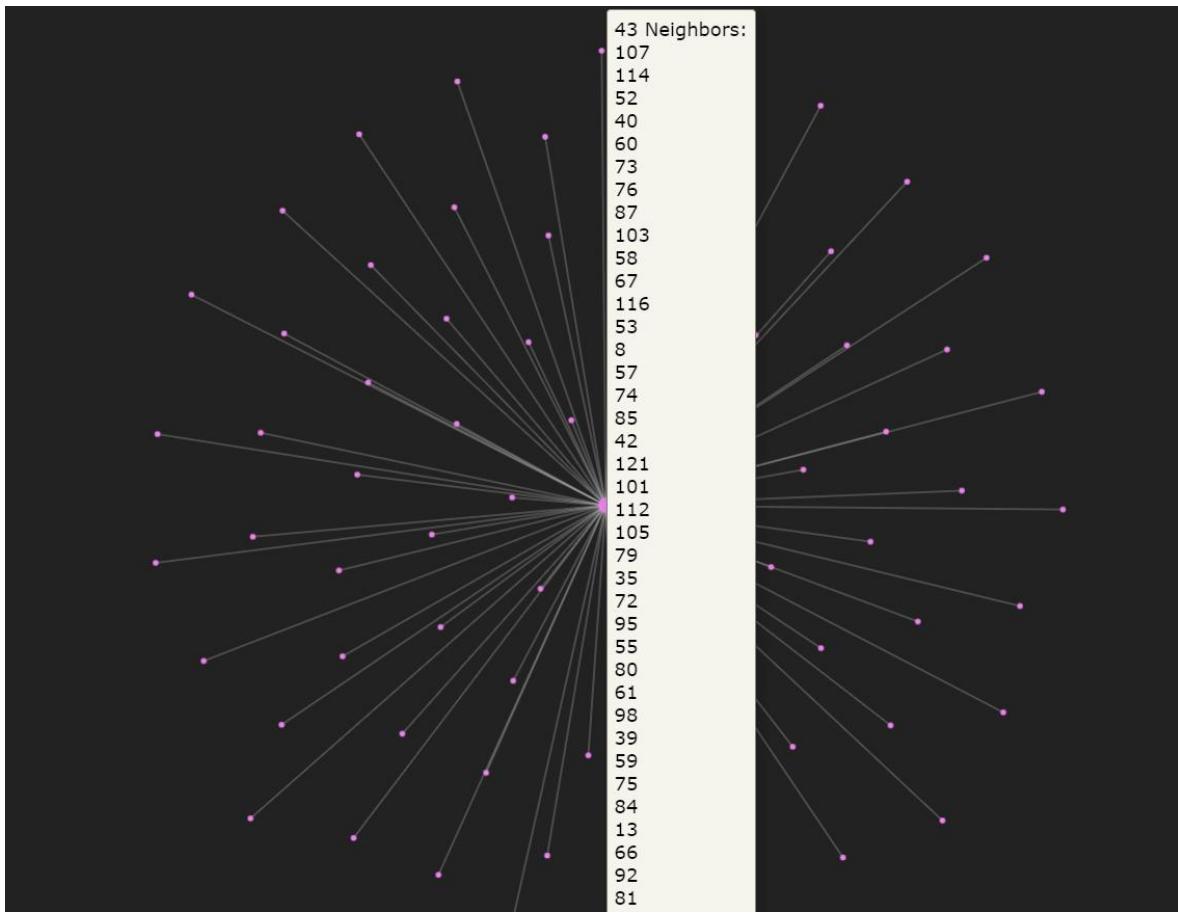
- **NODE 43:**

CENTRALITIES:

- Centrality is a very important concept in identifying important nodes in a graph.
- Computing centrality cut through noisy data, revealing parts of the network that needs attention. But they all work differently. Each measure has its own definition of 'importance'.

1.DEGREE CENTRALITY:

- Degree centrality measures how many direct connections each node has to other nodes in the network.
- Degree Centrality is used for finding very connected species, popular species, species who are likely to be extinct or likely to be keystone species.

$D(i) = \sum_j m(i,j)$ where m(i,j) = 1 if there is a link from node "*i*" to node "*j*"

Since it is a directed graph, both indegree and outdegree is computed.

INDEGREE: Number of edges directed into a vertex in a directed graph.
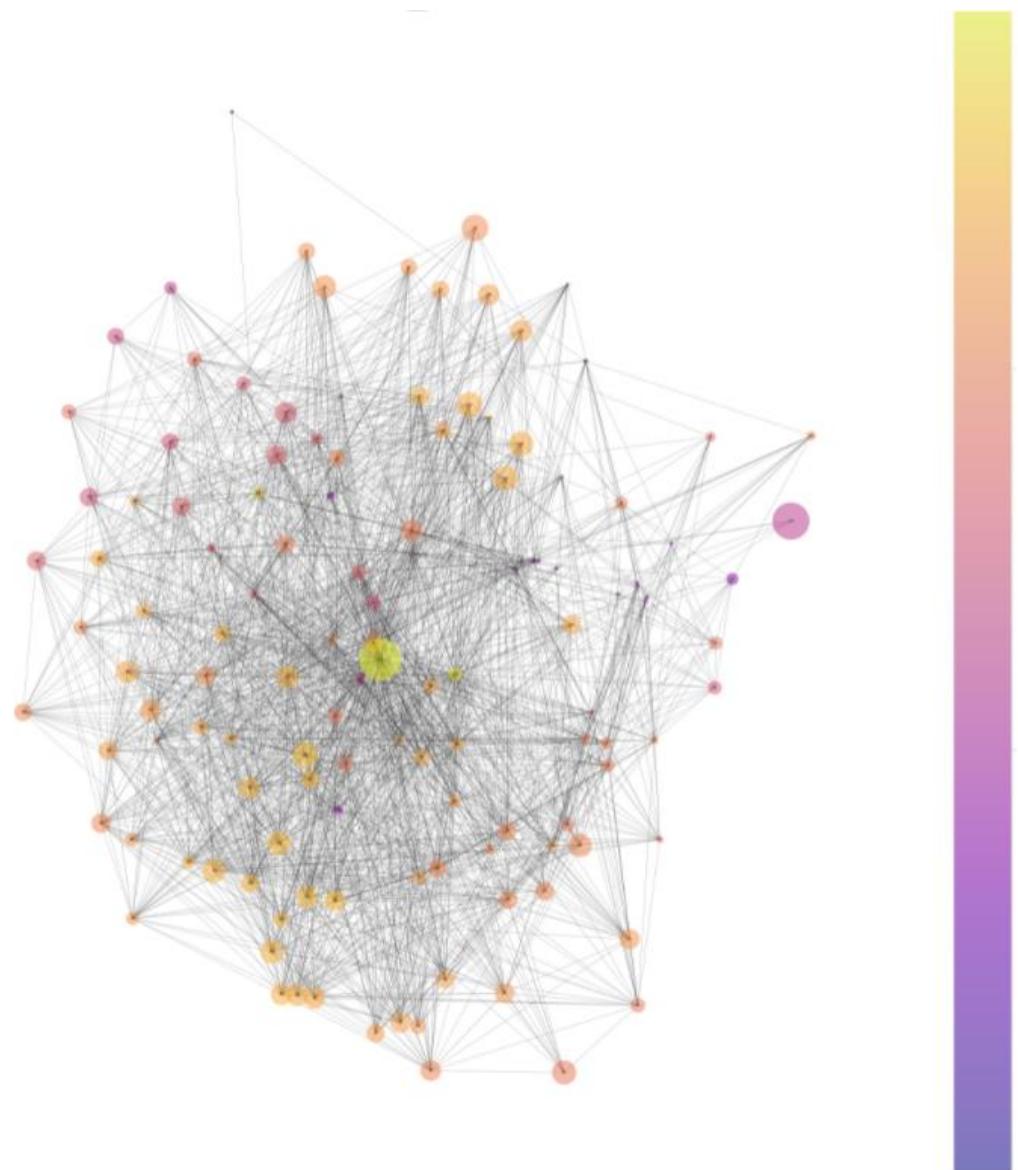
OUTDEGREE: Number of edges directed outside a vertex in a directed graph.

VISUALIZATION OF INDEGREE CENTRALITY:

Lighter the colour and larger the size of the node, more is the indegree centrality. Darker the colour and smaller the size of the node, lesser is the indegree centrality.

- Pale yellow nodes are the nodes with higher connections or indegree centrality, which means those species are more influential preys in the food web (keystone species).
- Dark blue nodes are the nodes with lesser connections or indegree centrality, which means those species are least important in the food web as a

prey. So, they are very less likely to be preys to the predators.

INFERENCE:

- "Node 127" has the highest indegree centrality which means the species labelled as 127 is the top keystone species in the food web as prey. If this prey doesn't exist then the predators depending upon this prey would cease to exist.

- "Node 122" is the second species which has the highest degree centrality which means this species is the second keystone species in the food web as a prey.
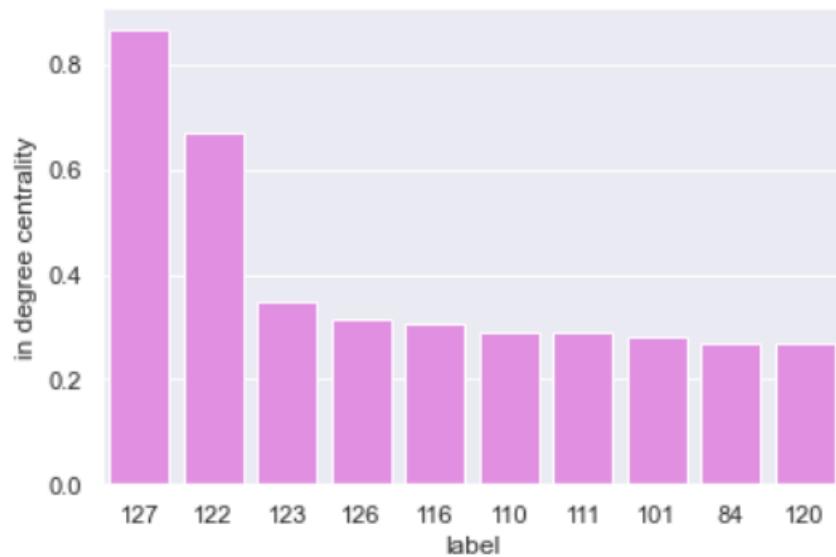
INDEGREE CENTRALITIES OF EACH NODE:
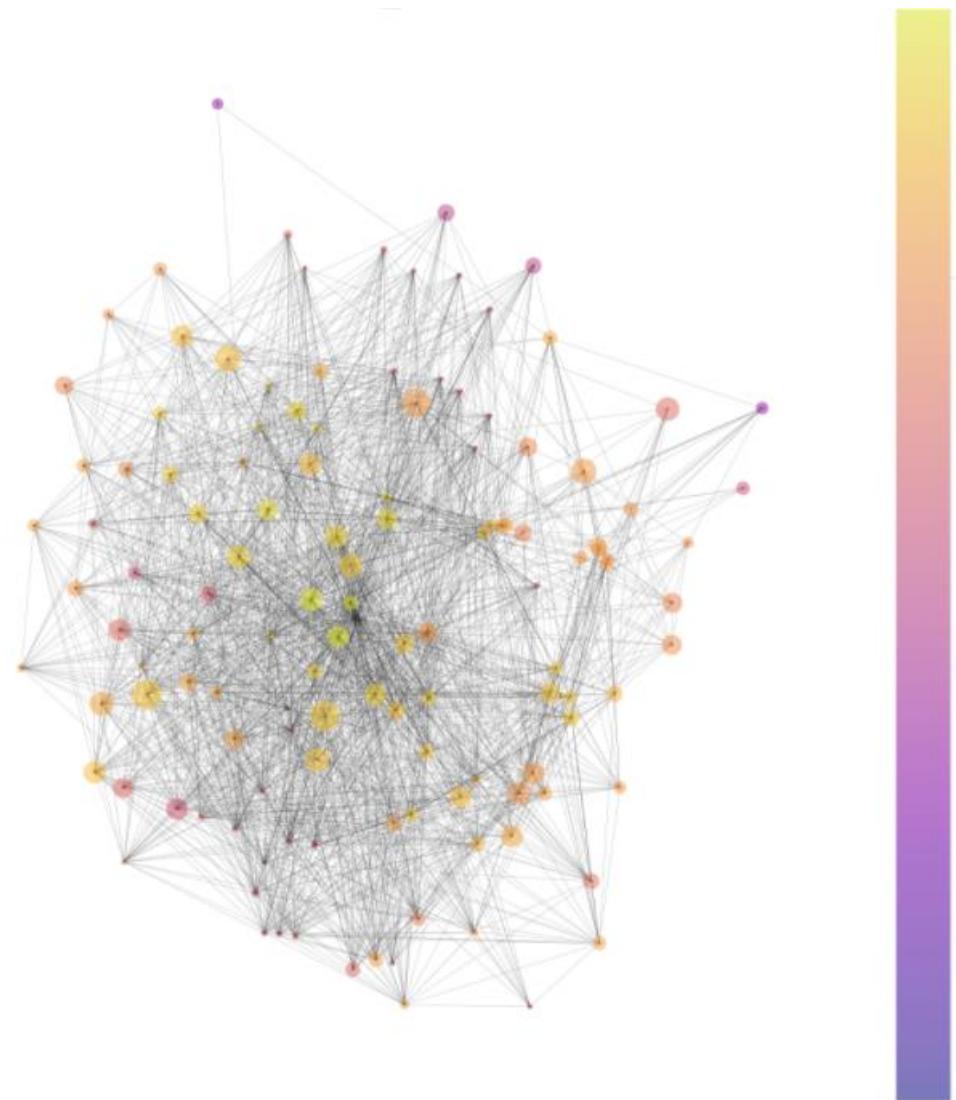
```
vertices
```

	label	id	name	group	counts	in degree centrality
0	0	0	2um Spherical Phytoplankt	unknown	16	0.0079
1	1	1	Synedococcus	unknown	24	0.0079
2	2	2	Oscillatoria	unknown	11	0.0079
3	3	3	Small Diatoms (<20um)	unknown	13	0.0079
4	4	4	Big Diatoms (>20um)	unknown	15	0.0079
...	...	...	...	...	...	...
123	123	123	Benthic POC	Detritus	64	0.3465
124	124	124	DOC	unknown	14	0.1024
125	125	125	Input	Detritus	18	0.0000
126	126	126	Output	unknown	40	0.3150
127	127	127	Respiration	unknown	110	0.8661

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_deg_cent_in(vertices)

VISUALIZATION OF OUTDEGREE CENTRALITY:

INFERENCE:
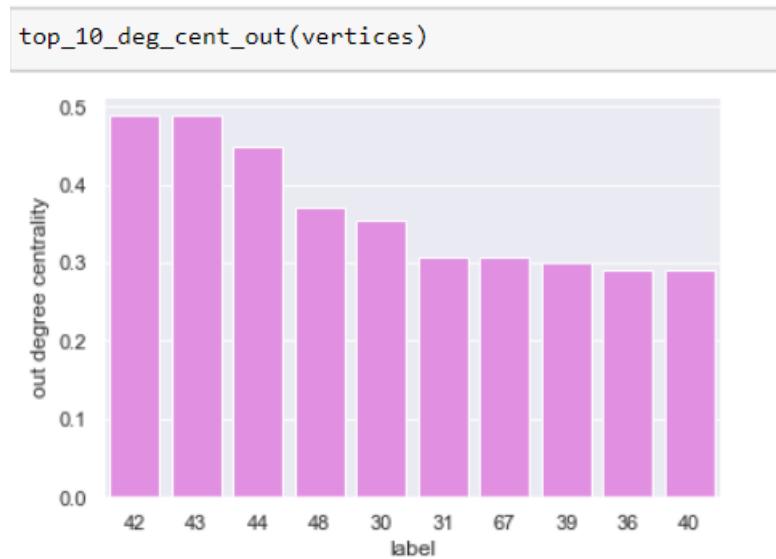
- "Node 42, 43" has the highest
  outdegree centrality which means
  these species are the top keystone
  species in the food web as predators. If
  these predators doesn't exist then the

other predators depending upon these
species would cease to exist.

OUTDEGREE CENTRALITIES OF EACH NODE:

vertices

	label	id	name	group	counts	in degree centrality	out degree centrality
0	0	0	2um Spherical Phytoplankt	unknown	16	0.0079	0.1181
1	1	1	Synedococcus	unknown	24	0.0079	0.1811
2	2	2	Oscillatoria	unknown	11	0.0079	0.0787
3	3	3	Small Diatoms (<20um)	unknown	13	0.0079	0.0945
4	4	4	Big Diatoms (>20um)	unknown	15	0.0079	0.1102
...	...	...	...	...	...	...	...
123	123	123	Benthic POC	Detritus	64	0.3465	0.1575
124	124	124	DOC	unknown	14	0.1024	0.0079
125	125	125	Input	Detritus	18	0.0000	0.1417
126	126	126	Output	unknown	40	0.3150	0.0000
127	127	127	Respiration	unknown	110	0.8661	0.0000

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_deg_cent_out(vertices)
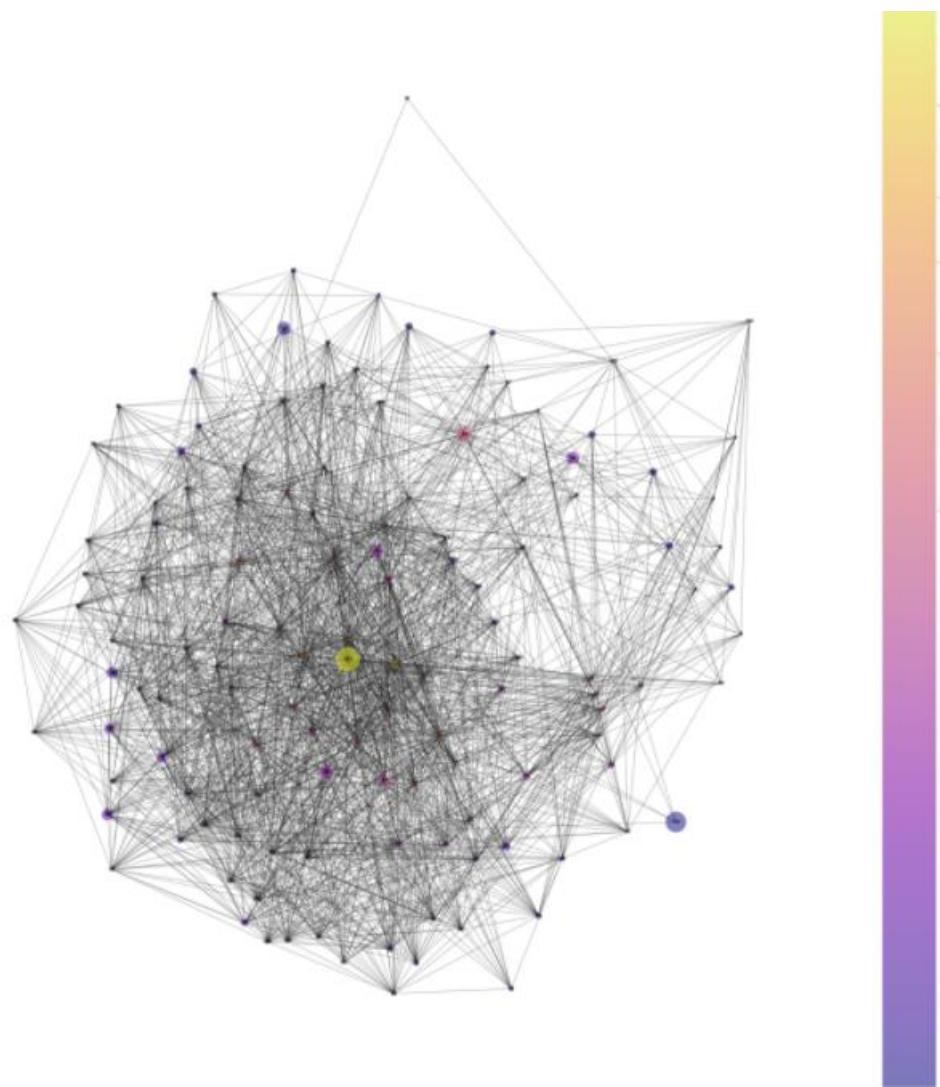
2. BETWEENESS CENTRALITY:

- Betweenness centrality measures the number of times a node lies on the shortest path between all pairs of nodes in a graph.
- This measure shows which nodes are 'bridges' between nodes in a network.
- It is used to find the individuals who influence the flow around a system.

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Where σ_{st} is total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of paths that pass-through v.

VISUALIZATION OF BETWEENESS CENTRALITY:

INFERENCE:

- "Node 127" has the highest betweenness centrality among other nodes which means the "Respiration" species belonging to unknown species

group is the bridge to many other pair of species.

- So, it means that this organism is an important species in the food web either as prey or as predator and it is the key determinant for the energy flow in the food web.
- This species with high betweenness centrality is important because it mediates many indirect interactions between species.
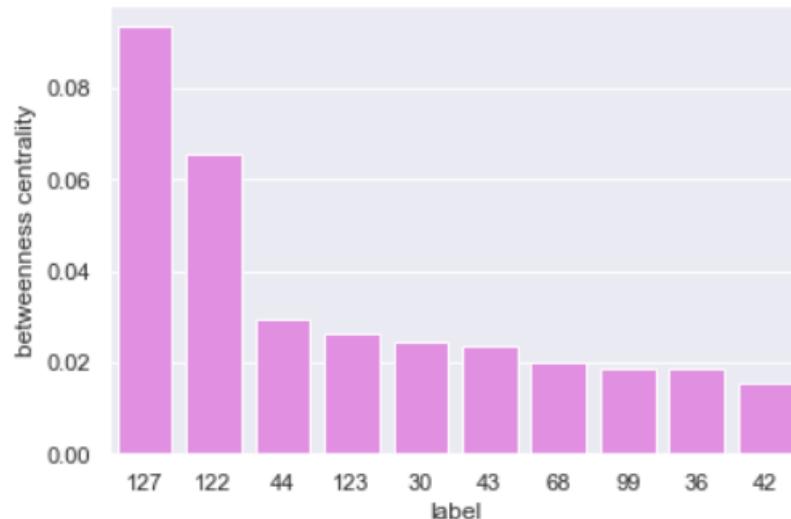
BETWEENESS CENTRALITIES OF EACH NODE:

```
vertices
```

	label	id	name	group	counts	degree centrality	betweenness centrality
0	0	0	2um Spherical Phytoplankt	unknown	16	0.1260	0.0016
1	1	1	Synedococcus	unknown	24	0.1890	0.0058
2	2	2	Oscillatoria	unknown	11	0.0866	0.0011
3	3	3	Small Diatoms (<20um)	unknown	13	0.1024	0.0012
4	4	4	Big Diatoms (>20um)	unknown	15	0.1181	0.0017
...	...	...	...	...	...	...	...
123	123	123	Benthic POC	Detritus	64	0.3622	0.0263
124	124	124	DOC	unknown	14	0.1102	0.0026
125	125	125	Input	Detritus	18	0.1417	0.0079
126	126	126	Output	unknown	40	0.3150	0.0104
127	127	127	Respiration	unknown	110	0.8661	0.0935

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

```
top_10_bet_cent(vertices)
```
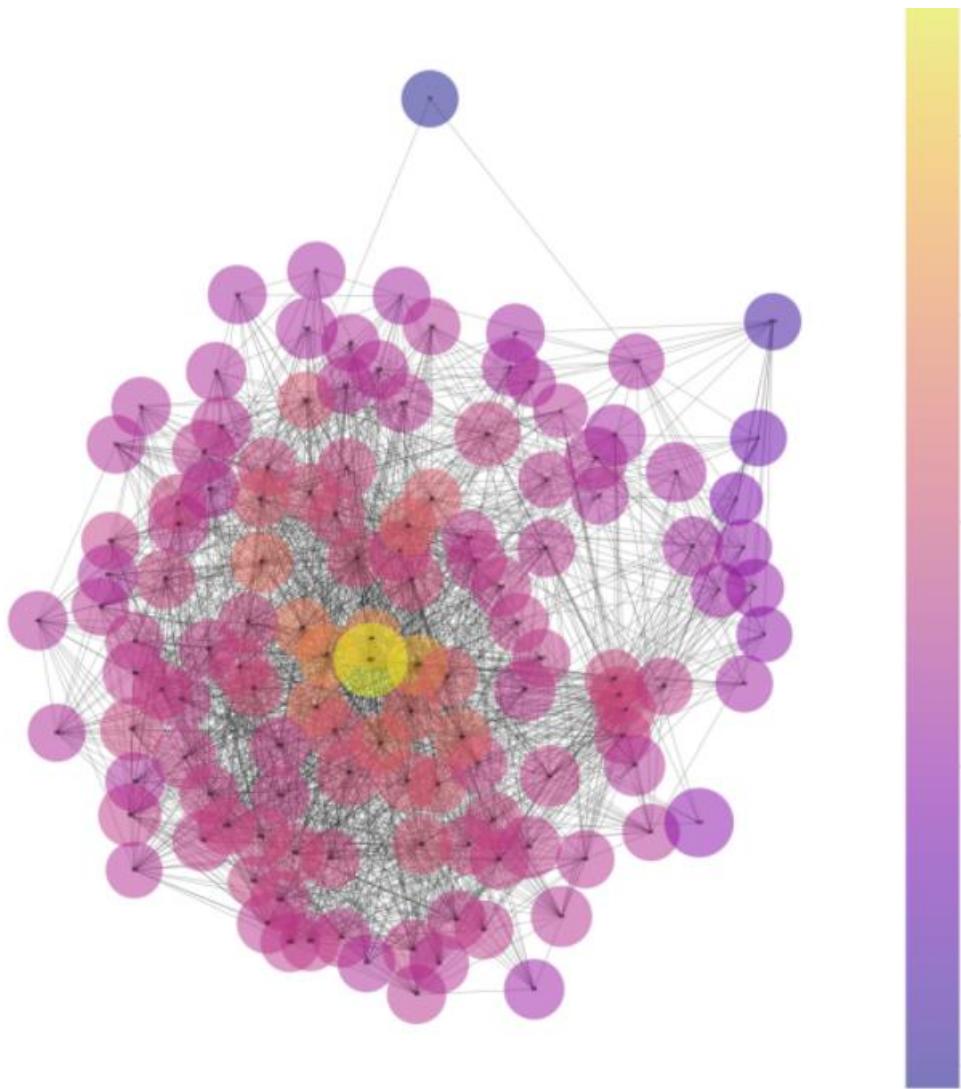
3.CLOSENESS CENTRALITY:

- Closeness centrality measures each node based on their 'closeness' to all other nodes in the network.
- A vertex with a high closeness centrality would mean it has close relationships with many vertices.
- If a node has strong closeness centrality, it is in a position, with its relationships, to spread information quickly.
- These species can be important influencers of the network.

$$C(x) = \frac{1}{\sum_y d(y,x)}$$

where d (x, y) is the distance between vertices x and y.

VISUALIZATION OF CLOSENESS CENTRALITY:

INFERENCE:

- "Node 127" is a Respiration species belonging to the "Unknown" species
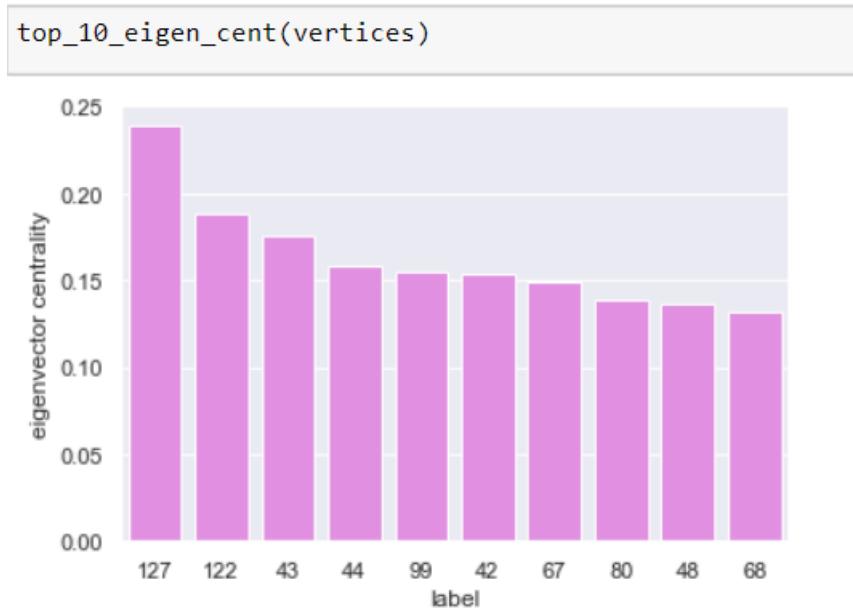
group. It is seen that this species has the highest closeness centrality.

- It means this organism is close to many other species in the food web unlike other species. So, this species is the most influential one.

- Extinction of this species will have a drastic impact on the other species close to it.

CLOSENESS CENTRALITIES OF EACH NODE:

vertices

	label	id	name	group	counts	degree centrality	betweenness centrality	closeness centrality
0	0	0	2um Spherical Phytoplankt	unknown	16	0.1260	0.0016	0.5060
1	1	1	Synedococcus	unknown	24	0.1890	0.0058	0.5451
2	2	2	Oscillatoria	unknown	11	0.0866	0.0011	0.4811
3	3	3	Small Diatoms (<20um)	unknown	13	0.1024	0.0012	0.4601
4	4	4	Big Diatoms (>20um)	unknown	15	0.1181	0.0017	0.5184
...	...	...	...	...	...	...	...	...
123	123	123	Benthic POC	Detritus	64	0.3622	0.0263	0.6106
124	124	124	DOC	unknown	14	0.1102	0.0026	0.4150
125	125	125	Input	Detritus	18	0.1417	0.0079	0.5270
126	126	126	Output	unknown	40	0.3150	0.0104	0.5721
127	127	127	Respiration	unknown	110	0.8661	0.0935	0.8819

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_eigen_cent(vertices)

4. EIGENVECTOR CENTRALITY:
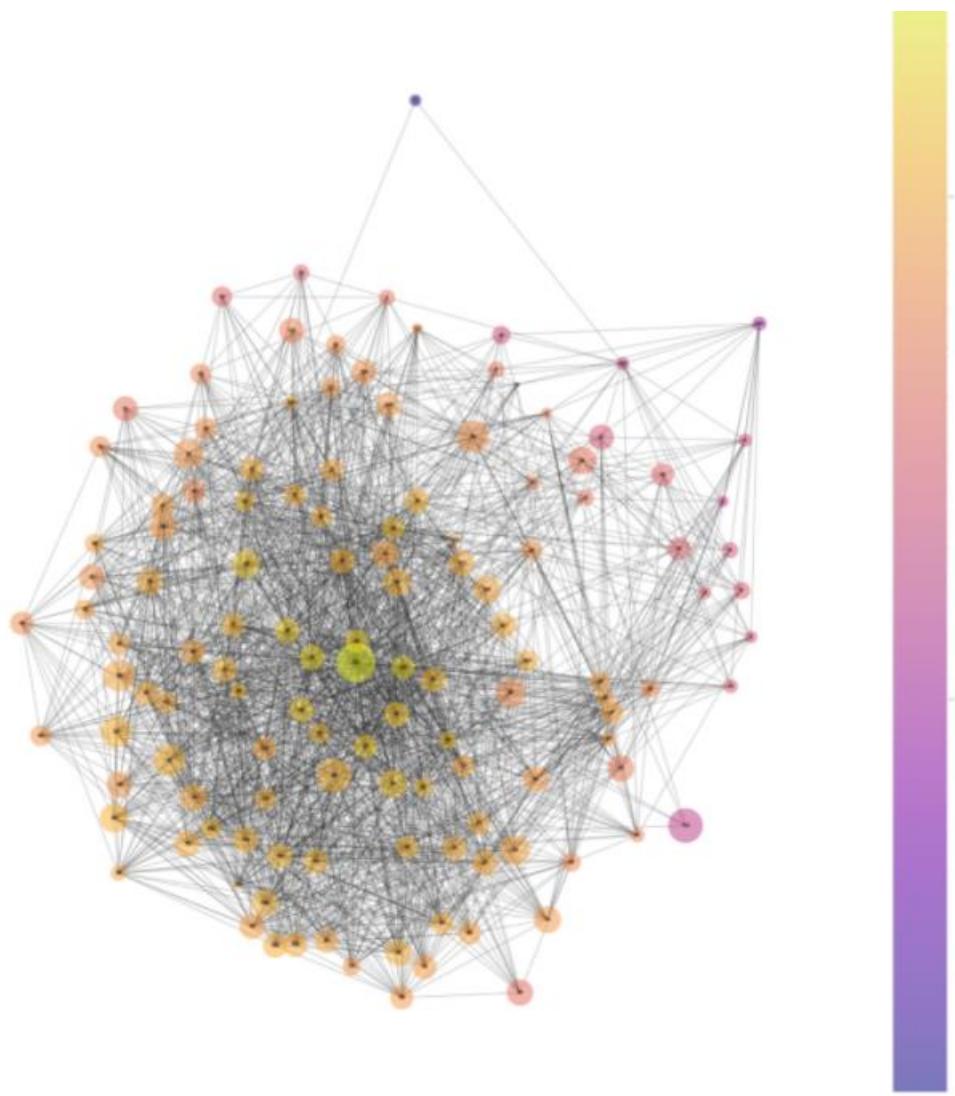
- This metric measures the importance of a node in a graph as a function of the importance of its neighbours.
- If a node is connected to highly important nodes, it will have a higher Eigen Vector Centrality score as

compared to a node which is connected to lesser important nodes.

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{N} A_{i,j} x_j$$

where λ is constant and A_{ij} is the ij^{th} element of the adjacency matrix.

VISUALIZATION OF EIGENVECTOR CENTRALITY:

INFERENCE:

- "Node 127" has the highest eigenvector centrality compared to the others which means it is the most influential species in the food web.
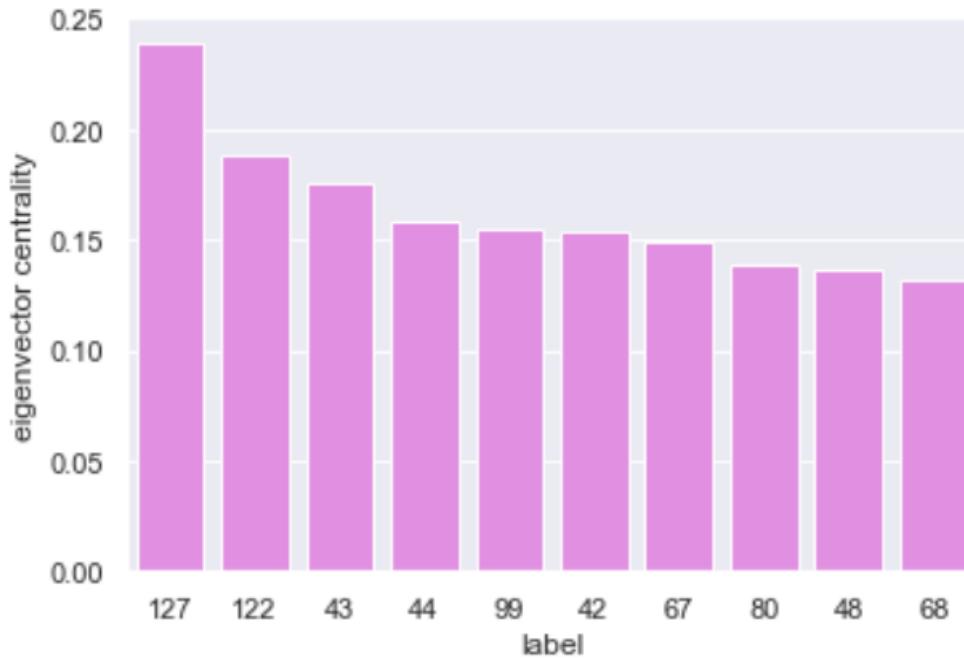
- As this respiratory species is the prey or predator of many other species which in turn has connections with other influential species, "Node 127" turns out to be the most influential species in the network.

EIGENVECTOR CENTRALITIES OF EACH NODE:

	label	id	name	group	counts	degree centrality	betweenness centrality	closeness centrality	eigenvector centrality
0	0	0	2um Spherical Phytoplankt	unknown	16	0.1260	0.0016	0.5060	0.0266
1	1	1	Synedococcus	unknown	24	0.1890	0.0058	0.5451	0.0454
2	2	2	Oscillatoria	unknown	11	0.0866	0.0011	0.4811	0.0170
3	3	3	Small Diatoms (<20um)	unknown	13	0.1024	0.0012	0.4601	0.0193
4	4	4	Big Diatoms (>20um)	unknown	15	0.1181	0.0017	0.5184	0.0268
...	...	...	...	...	...	...	...	...	...
123	123	123	Benthic POC	Detritus	64	0.3622	0.0263	0.6106	0.1002
124	124	124	DOC	unknown	14	0.1102	0.0026	0.4150	0.0106
125	125	125	Input	Detritus	18	0.1417	0.0079	0.5270	0.0178
126	126	126	Output	unknown	40	0.3150	0.0104	0.5721	0.0821
127	127	127	Respiration	unknown	110	0.8661	0.0935	0.8819	0.2386

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

```
top_10_eigen_cent(vertices)
```


5. CLUSTERING COEFFICIENT CENTRALITY:
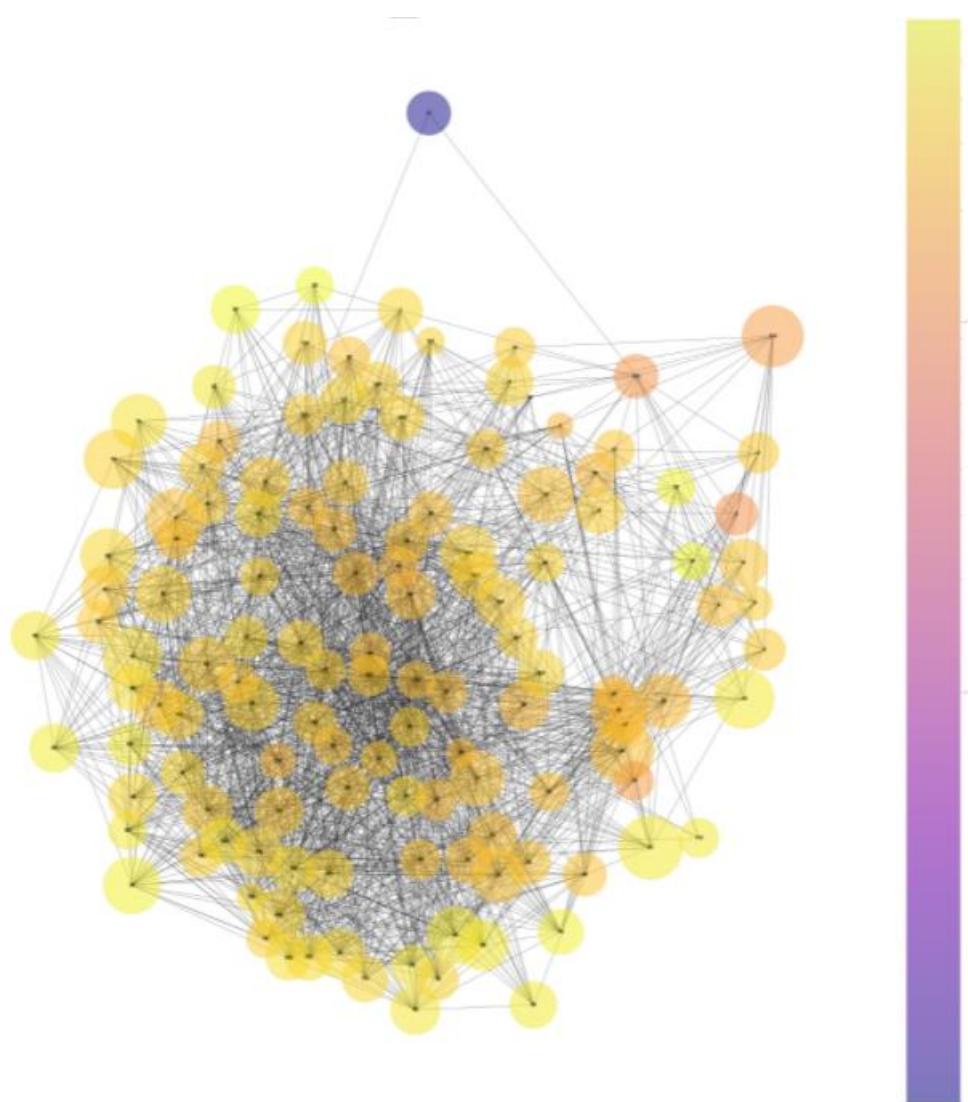
- Clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together.

$$C_i = \frac{\text{\# of links between 'n' neighbors}}{n(n-1)/2}$$

CLUSTERING COEFFICIENT OF THE WHOLE NETWORK:

$$C = \frac{1}{N} \sum C_i$$

VISUALIZATION OF CLUSTERING COEFFICIENT CENTRALITY:

INFERENCE:
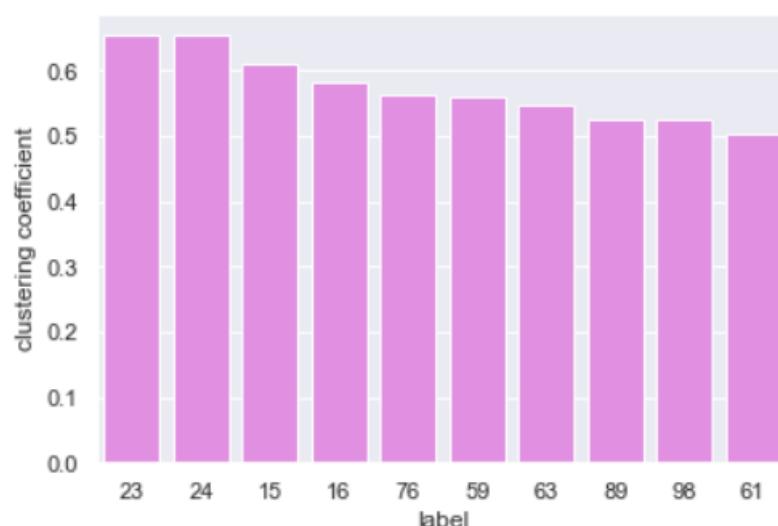
- "Node 23" and "Node 24" has more clustering coefficient than the other species which means these nodes tend to form clusters.

CLUSTERING COEFFICIENT CENTRALITIES OF EACH NODE:

vertices

	label	id	name	group	counts	degree centrality	betweenness centrality	closeness centrality	eigenvector centrality	clustering coefficient
0	0	0	2um Spherical Phytoplankt	unknown	16	0.1260	0.0016	0.5060	0.0266	0.3167
1	1	1	Synedococcus	unknown	24	0.1890	0.0058	0.5451	0.0454	0.1957
2	2	2	Oscillatoria	unknown	11	0.0866	0.0011	0.4811	0.0170	0.1091
3	3	3	Small Diatoms (<20um)	unknown	13	0.1024	0.0012	0.4601	0.0193	0.2949
4	4	4	Big Diatoms (>20um)	unknown	15	0.1181	0.0017	0.5184	0.0268	0.2667
...	...	...	...	...	...	...	...	...	...	...
123	123	123	Benthic POC	Detritus	64	0.3622	0.0263	0.6106	0.1002	0.3130
124	124	124	DOC	unknown	14	0.1102	0.0026	0.4150	0.0106	0.1319
125	125	125	Input	Detritus	18	0.1417	0.0079	0.5270	0.0178	0.1242
126	126	126	Output	unknown	40	0.3150	0.0104	0.5721	0.0821	0.2077
127	127	127	Respiration	unknown	110	0.8661	0.0935	0.8819	0.2386	0.2859

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:
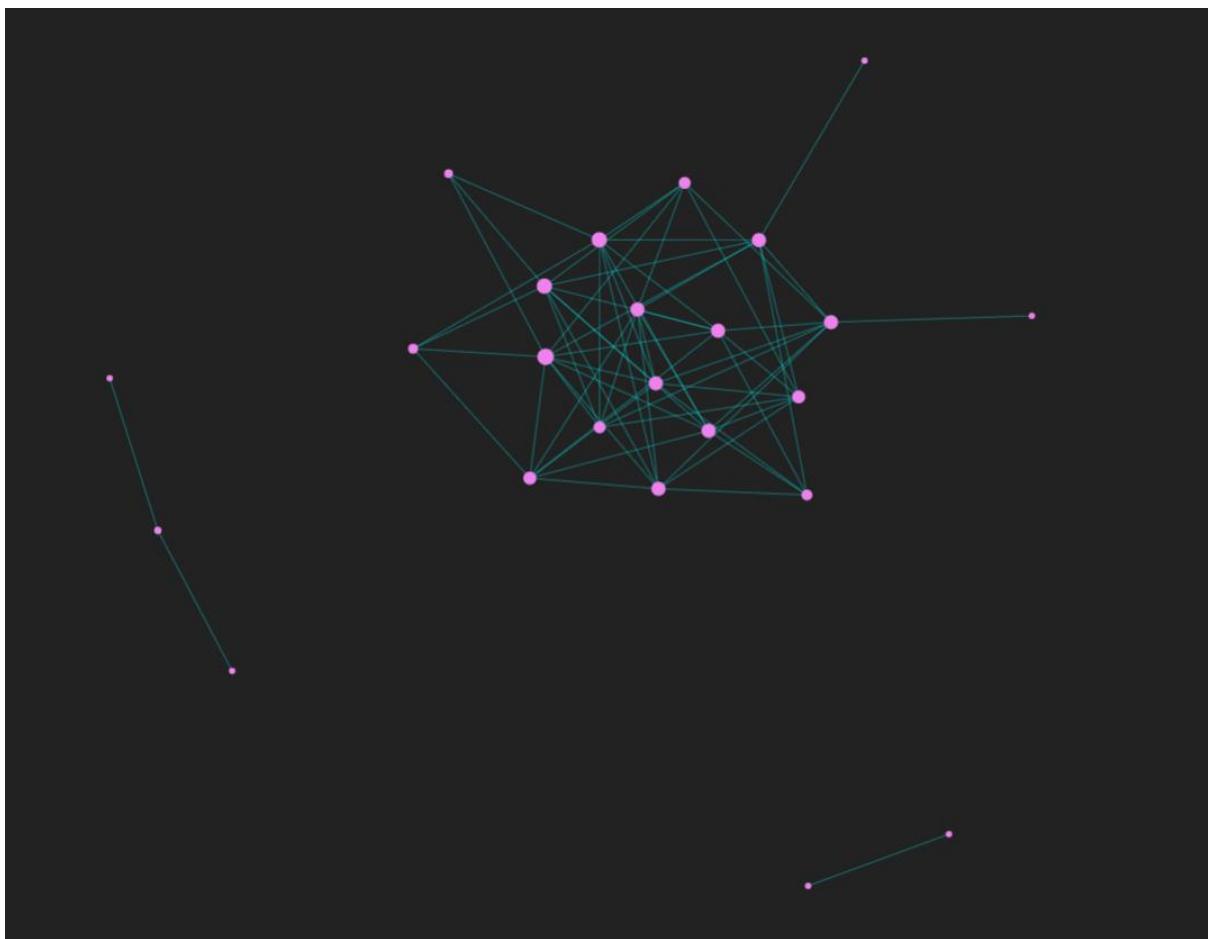
CENTRALITIES FOR SPECIES GROUP:

1. SPECIES GROUP: UNKNOWN

NODES IN THE NETWORK THAT BELONGS TO "UNKNOWN" SPECIES GROUP:
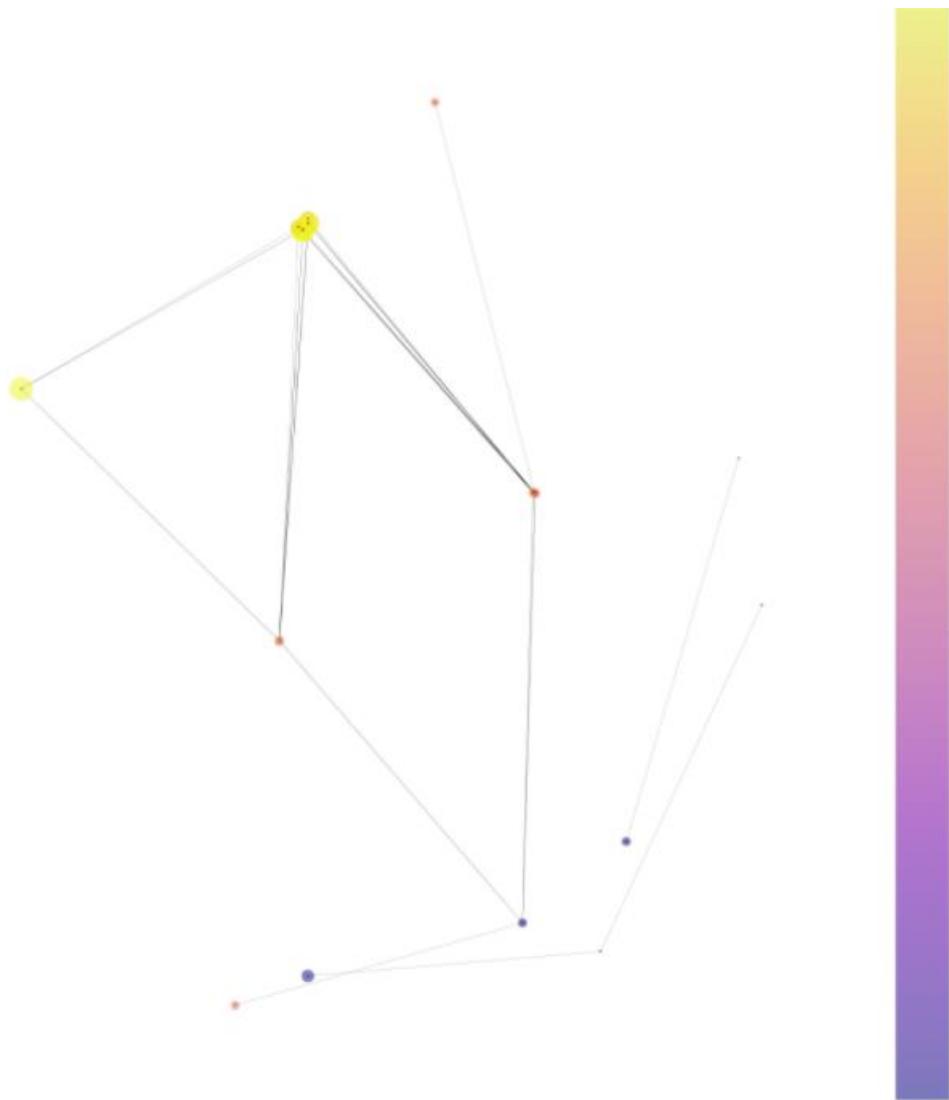
	vertices
	label
0	0
1	125
2	15
3	16
4	17
5	18
6	19
7	20
8	22
9	26
10	1
11	2
12	3
13	27
14	4
15	5
16	21
17	6
18	11
19	28
20	45
21	66
22	14
23	124

UNKNOWN SPECIES GROUP VISUALIZATION USING PYVIZ:

DEGREE CENTRALITY:

VISUALIZATION OF INDEGREE CENTRALITY:

INFERENCE:
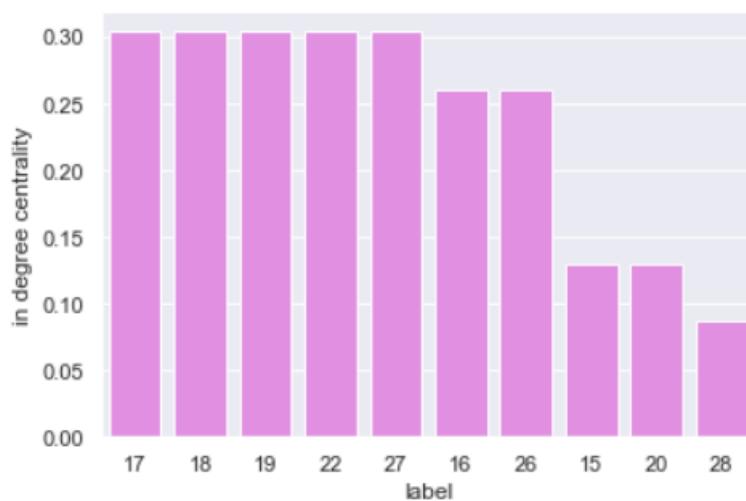
- "Node 17, 18, 19, 22, 27" has the maximum indegree centrality which means these species are more likely to be prey to many predators while

"Unknown" species group is taken into account.
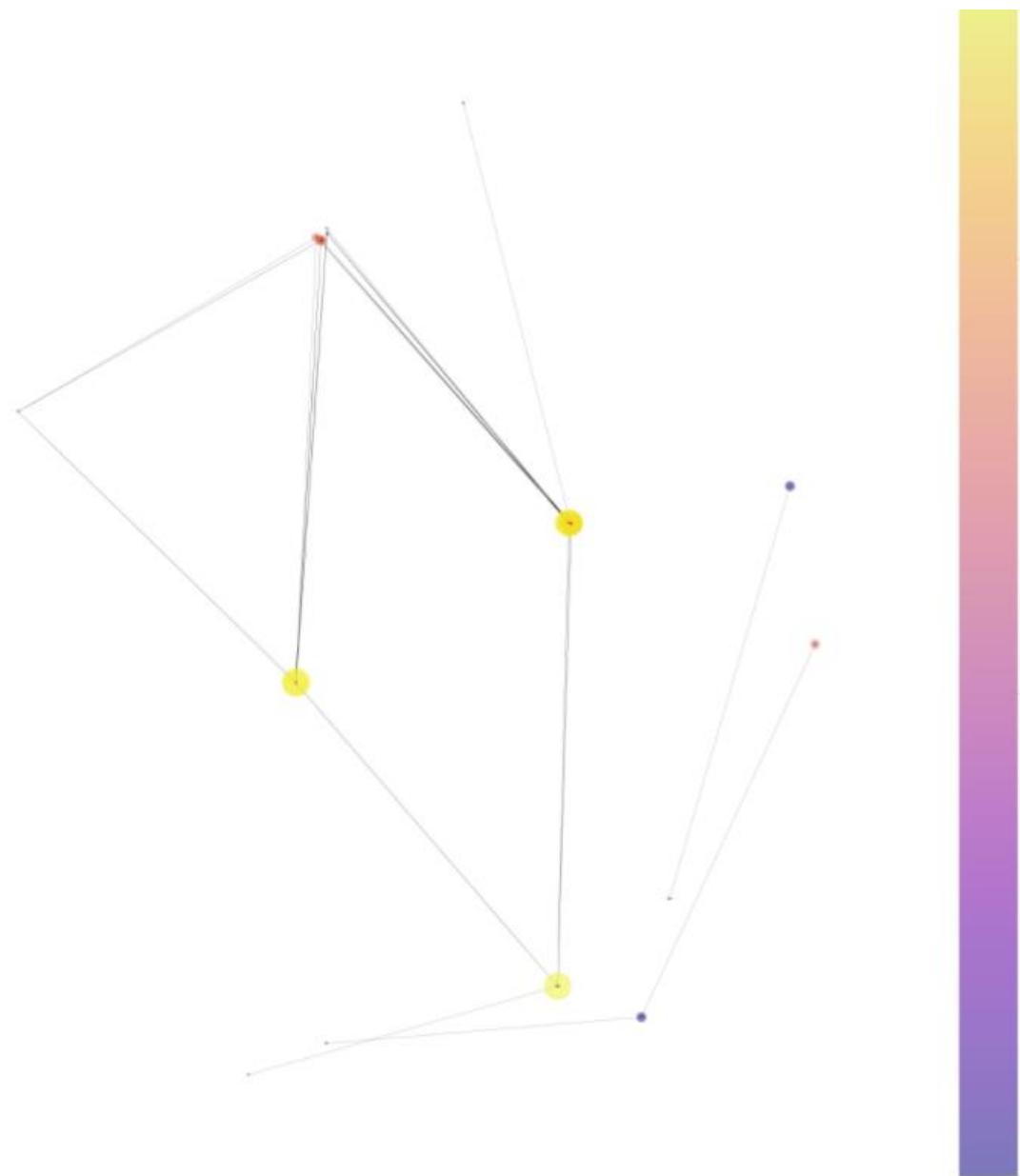
- So, these preys are more influential in unknown species group. Without these preys the predators depending on them would cease to exist.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:
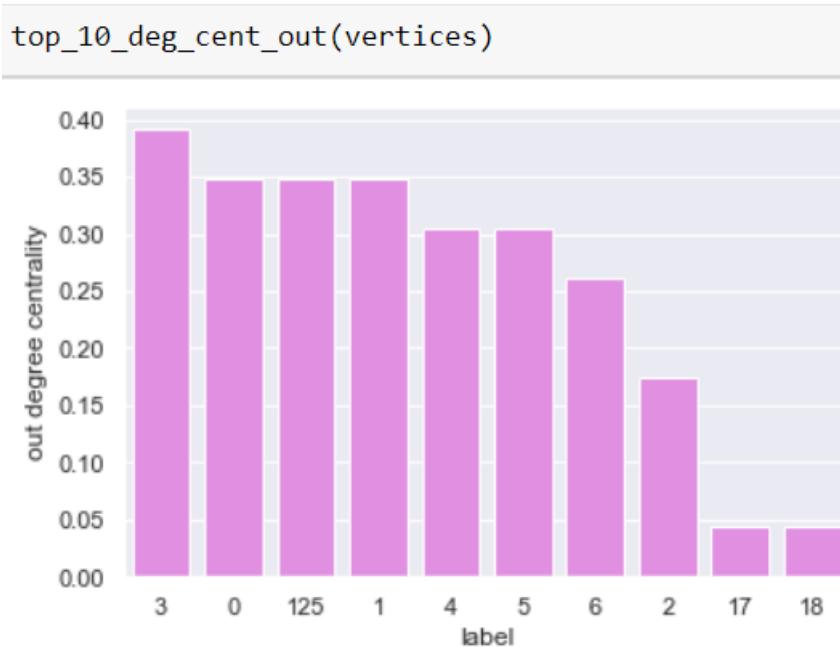
VISUALIZATION OF OUTDEGREE CENTRALITY:

INFERENCE:
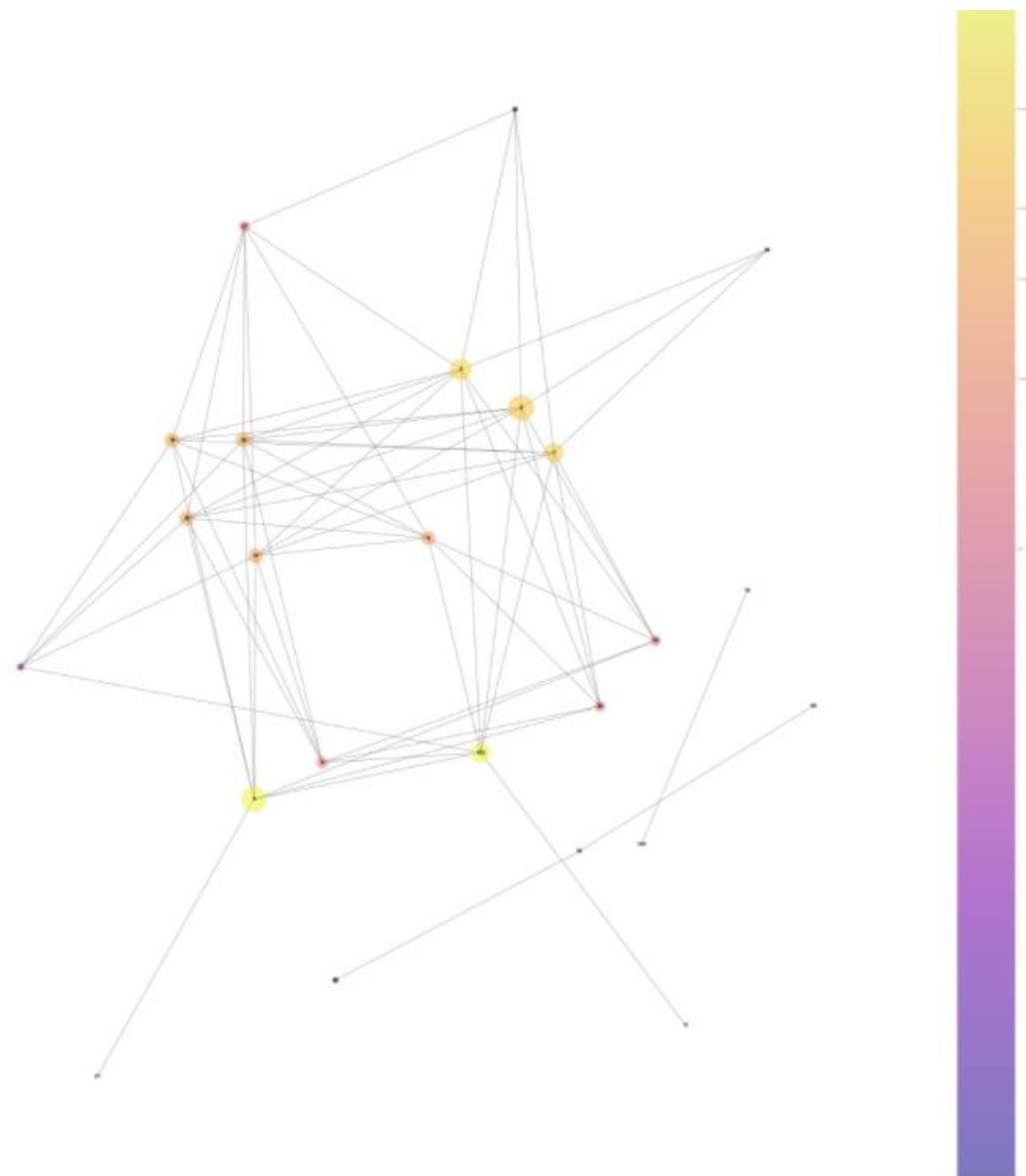
- "Node 3" has the maximum outdegree centrality which means this organism is likely to be a predator while "Unknown" species group is taken into account.
- So, this predator is more influential in unknown species group. Without this species the other predators depending on this organism would cease to exist.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

BETWEENESS CENTRALITY:
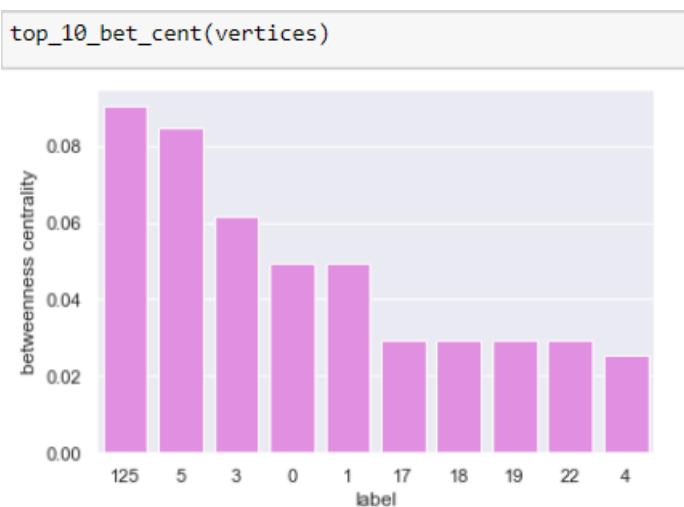
VISUALIZATION OF BETWEENESS CENTRALITY:

INFERENCE:
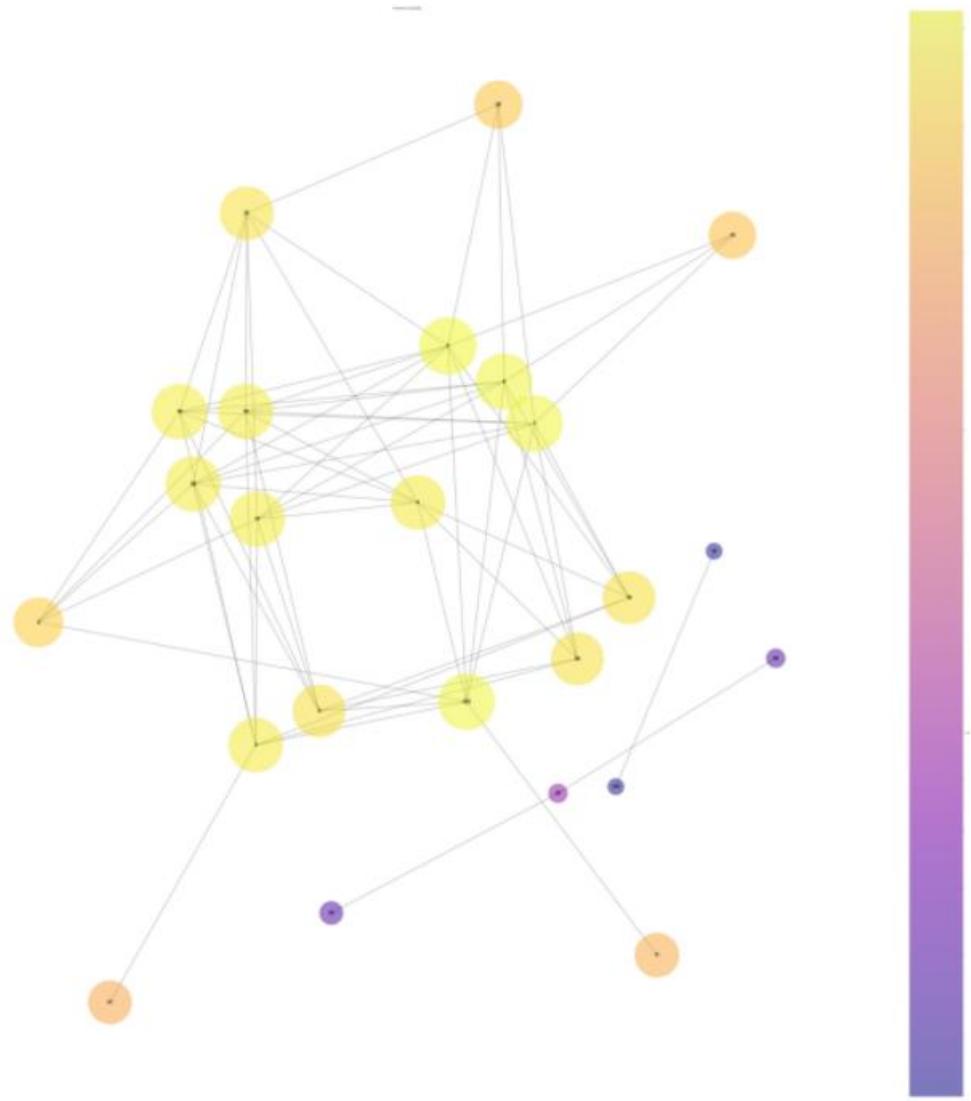
- "Node 125" has the highest betweenness centrality while "Unknown" species group is taken into account as either prey or predator.
- It means this species is well connected and it acts as a bridge between many pair of species.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

CLOSENESS CENTRALITY:
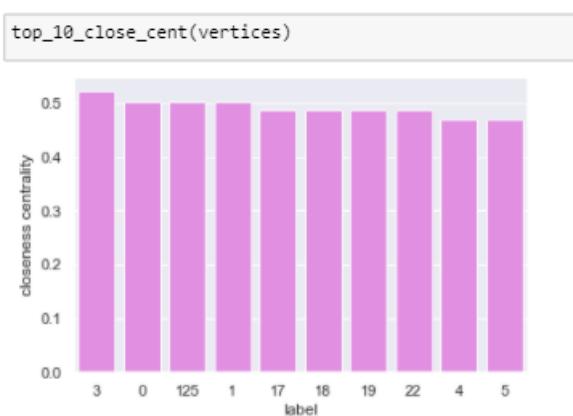
VISUALIZATION OF CLOSENESS CENTRALITY:

INFERENCE:
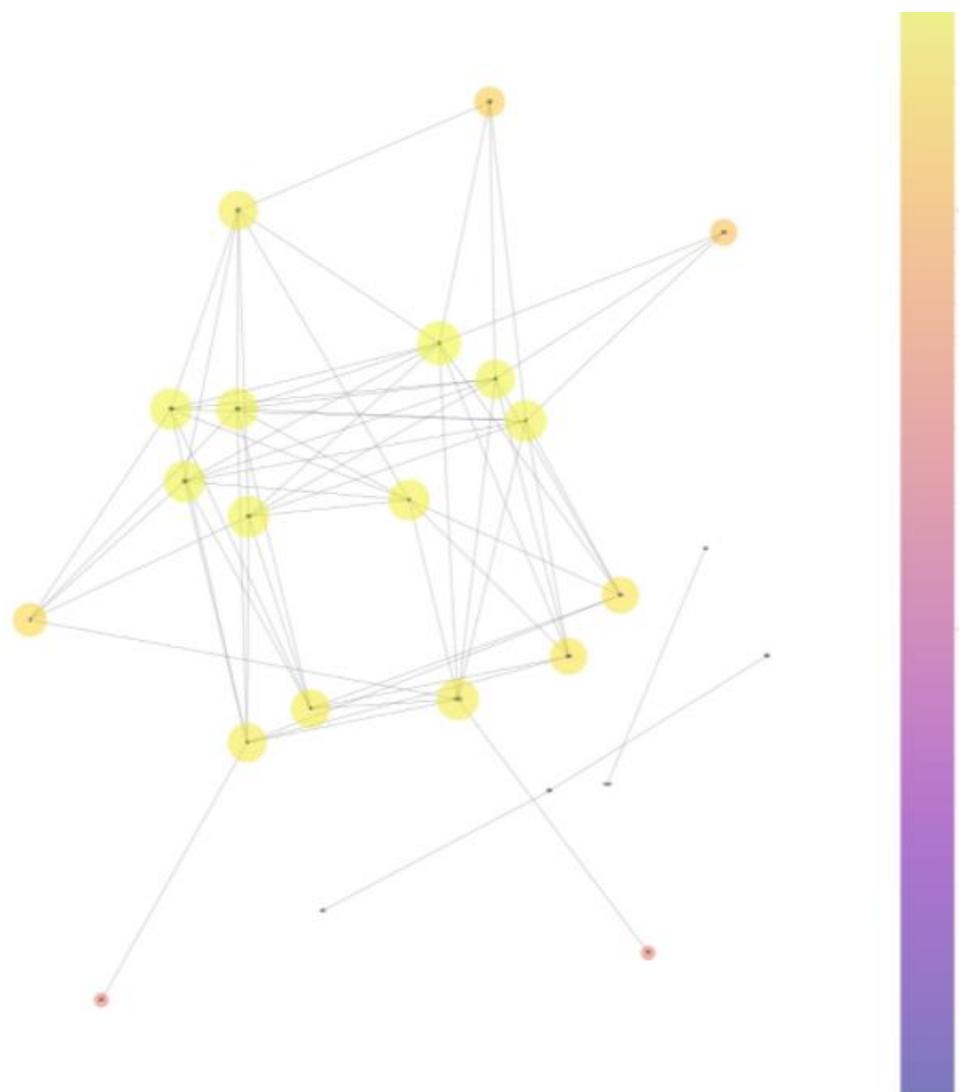
- "Node 3" has the highest closeness centrality while "Unknown" species group is taken into account as either prey or predator.
- It means this organism is closely connected to many species. So, the other species around Node 3 can either be prey or predator to this organism.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

EIGENVECTOR CENTRALITY:
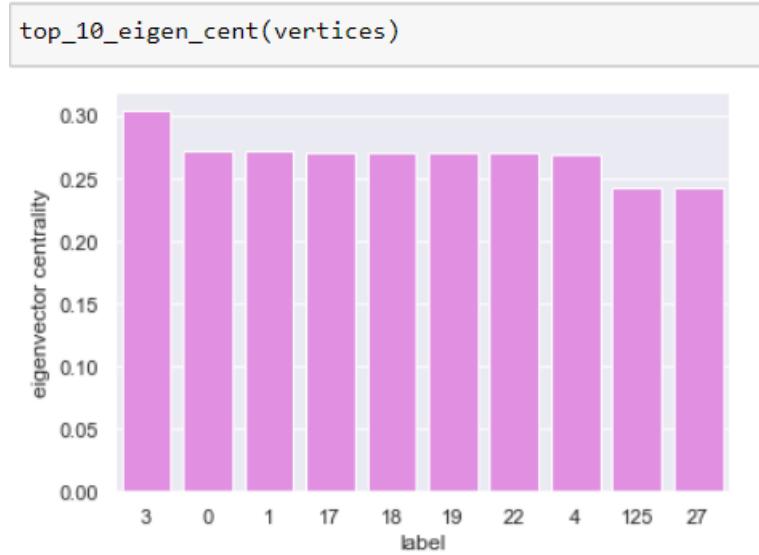
VISUALIZATION OF EIGENVECTOR CENTRALITY:

INFERENCE:
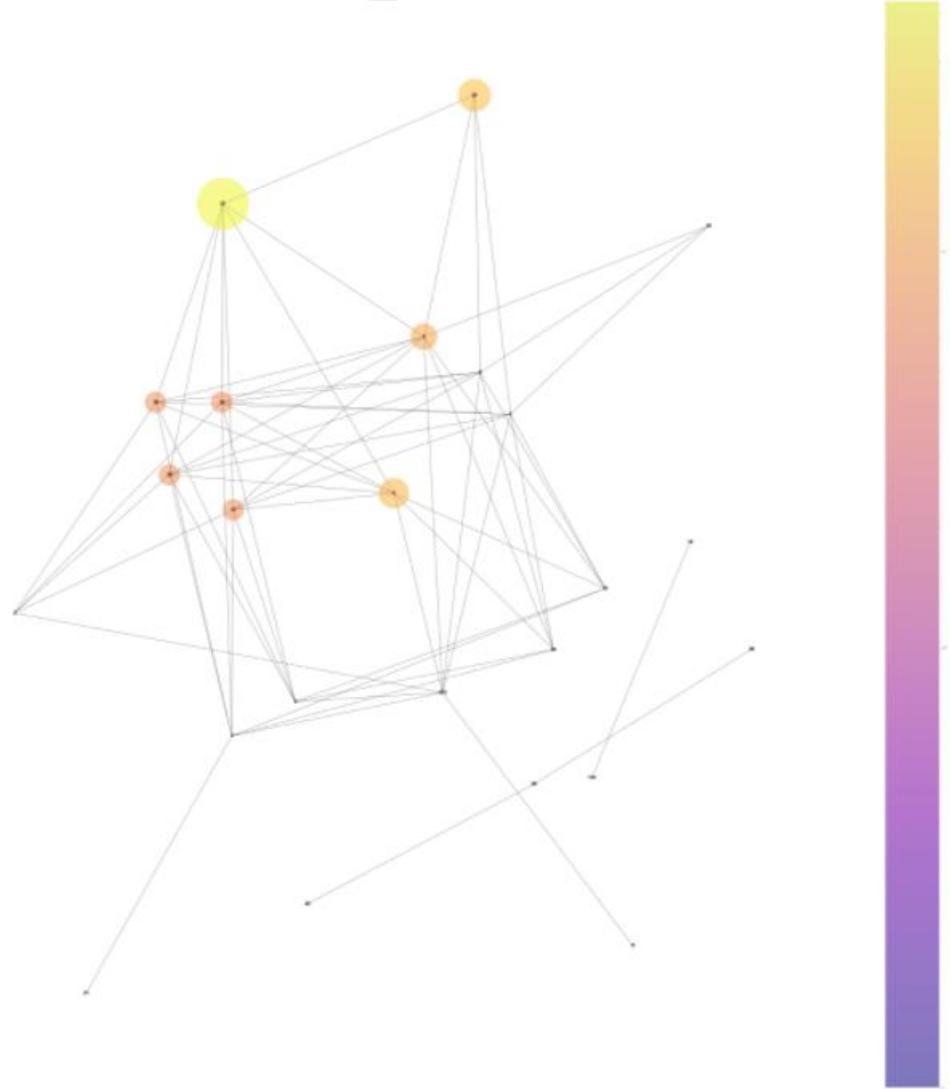
- "Node 3" belonging to the unknown species has the highest eigenvector centrality compared to the others which means it is the most influential species in the food web.
- As this species is the prey or predator of many other species which in turn has connections with other influential species, "Node 3" turns out to be the most influential species in the network.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_eigen_cent(vertices)

CLUSTERING COEFFICIENT:
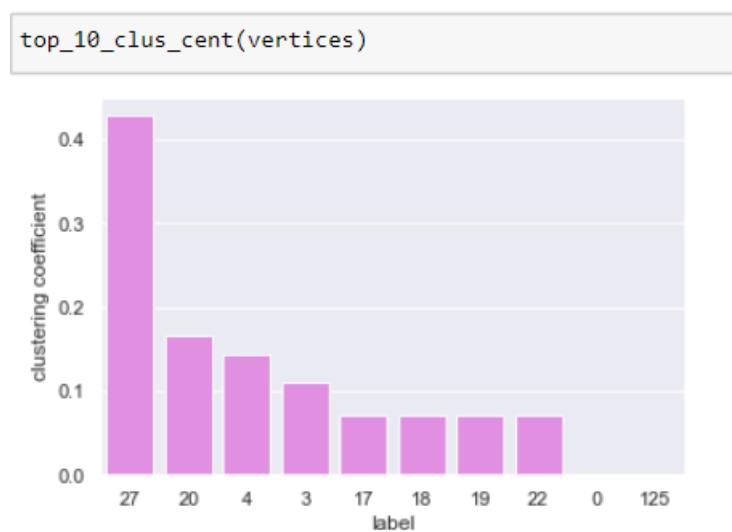
VISUALIZATION OF CLUSTERING CENTRALITY:

INFERENCE:

- "Node 27" has the highest clustering coefficient which means this species belonging to the unknown species group tend to form bigger cluster than the other species.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_clus_cent(vertices)

CENTRALITIES OF EACH NODE IN THE SPECIES GROUP "UNKNOWN":

vertices

	label	counts	degree centrality	betweenness centrality	closeness centrality	eigenvector centrality	clustering coefficient
0	0	9	0.3913	0.0493	0.5031	0.2716	0.0000
1	125	8	0.3478	0.0902	0.5031	0.2430	0.0000
2	15	3	0.1304	0.0013	0.3522	0.1141	0.0000
3	16	6	0.2609	0.0145	0.4402	0.2146	0.0000
4	17	8	0.3478	0.0292	0.4858	0.2713	0.0714
5	18	8	0.3478	0.0292	0.4858	0.2713	0.0714
6	19	8	0.3478	0.0292	0.4858	0.2713	0.0714
7	20	4	0.1739	0.0029	0.3707	0.1468	0.1667
8	22	8	0.3478	0.0292	0.4858	0.2713	0.0714
9	26	6	0.2609	0.0145	0.4402	0.2146	0.0000
10	1	9	0.3913	0.0493	0.5031	0.2716	0.0000
11	2	5	0.2174	0.0075	0.3913	0.1788	0.0000
12	3	11	0.4348	0.0614	0.5217	0.3043	0.1111
13	27	9	0.3043	0.0143	0.4544	0.2430	0.4286
14	4	9	0.3478	0.0253	0.4696	0.2692	0.1429
15	5	8	0.3478	0.0846	0.4696	0.2409	0.0000
16	21	1	0.0435	0.0000	0.2997	0.0324	0.0000
17	6	7	0.3043	0.0174	0.4402	0.2365	0.0000
18	11	1	0.0435	0.0000	0.3130	0.0327	0.0000
19	28	2	0.0870	0.0040	0.0870	0.0000	0.0000
20	45	1	0.0435	0.0000	0.0580	0.0000	0.0000
21	66	1	0.0435	0.0000	0.0580	0.0000	0.0000
22	14	1	0.0435	0.0000	0.0435	0.0000	0.0000
23	124	1	0.0435	0.0000	0.0435	0.0000	0.0000
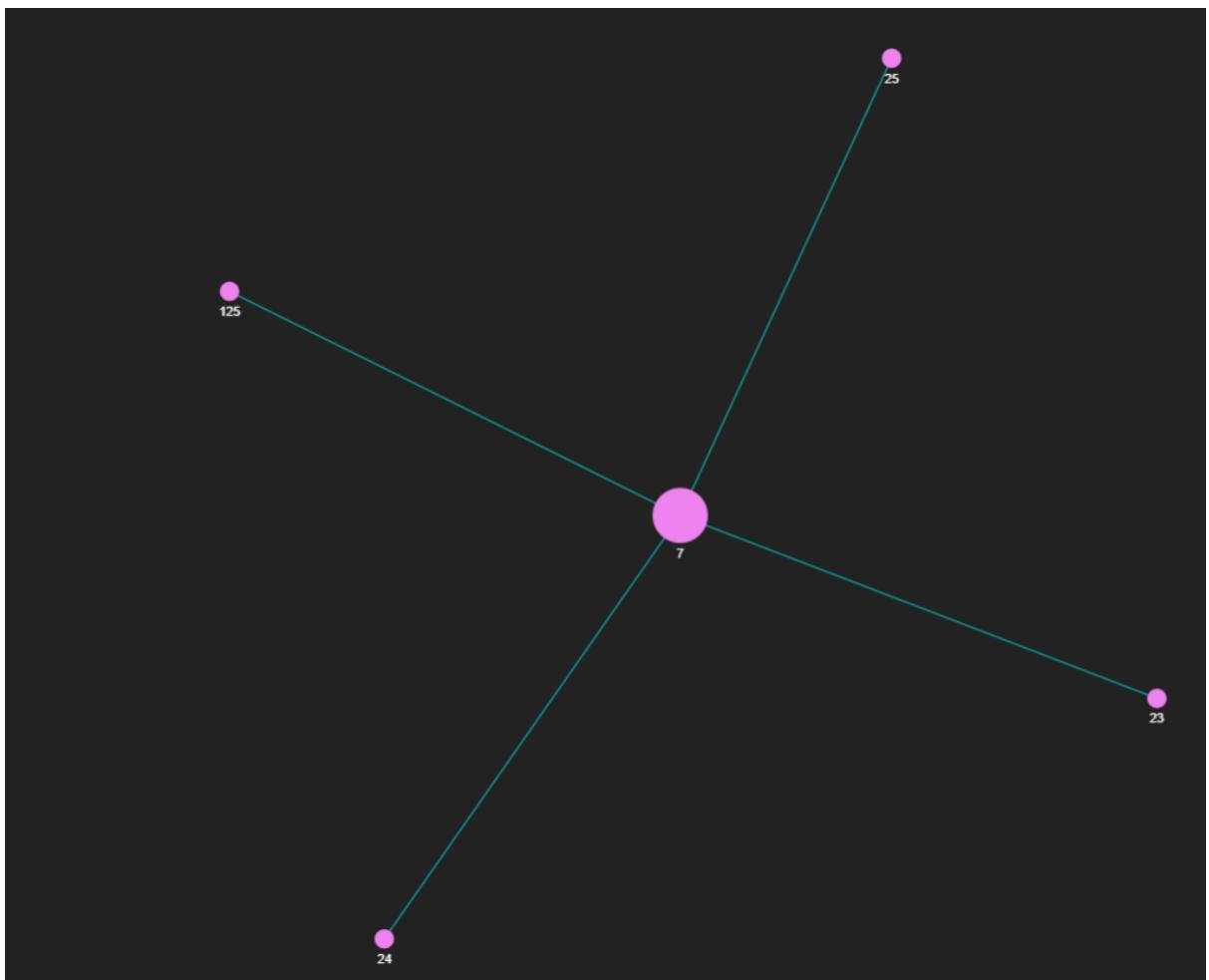
2. SPECIES GROUP: DEMERSAL PRODUCER

NODES IN THE NETWORK THAT BELONGS TO "DEMERSAL PRODUCER" SPECIES GROUP:

vertices
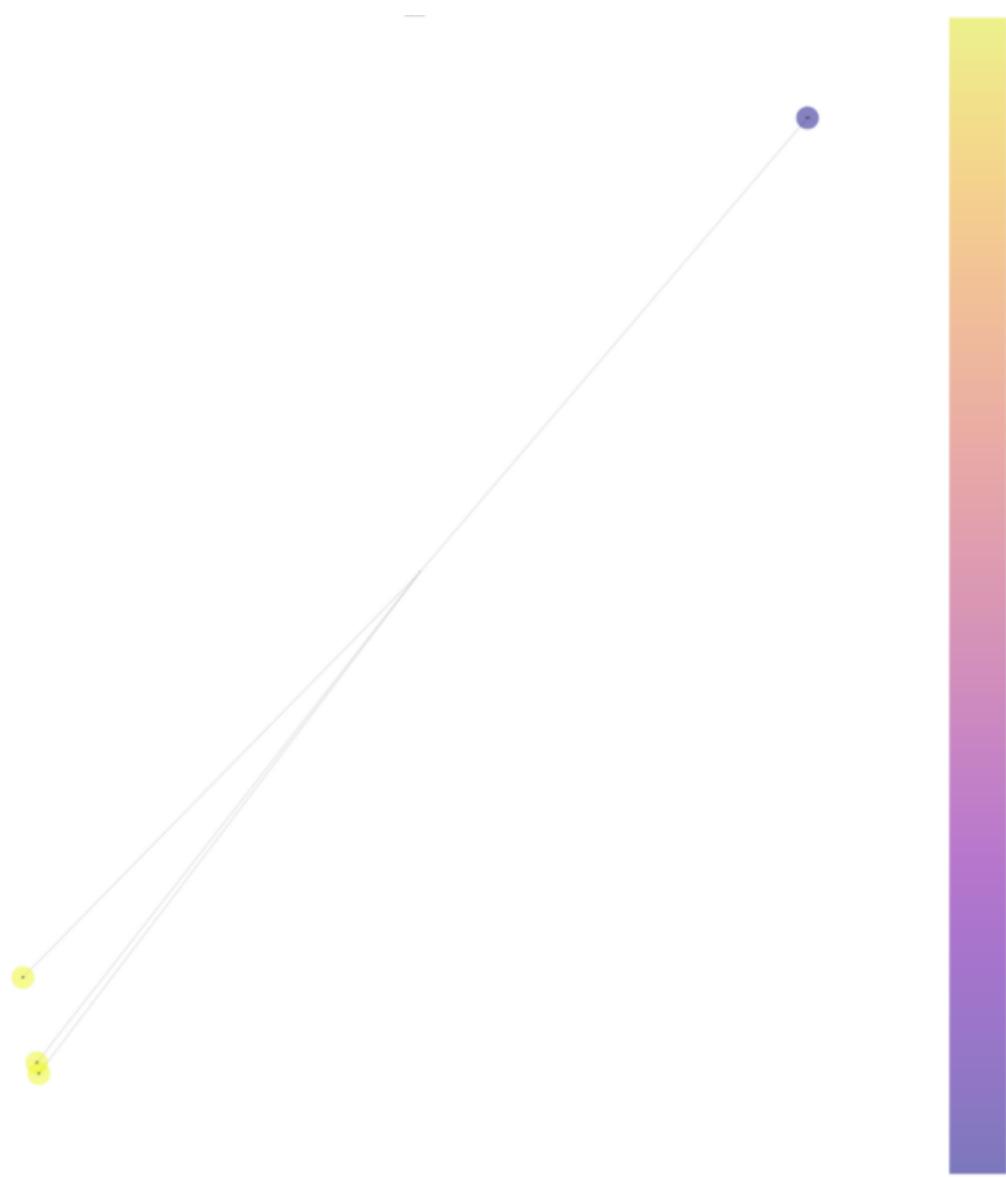
	label
0	7
1	125
2	23
3	24
4	25

DEMERSAL PRODUCER SPECIES GROUP VISUALIZATION USING PYVIZ:

DEGREE CENTRALITY:

VISUALIZATION OF INDEGREE
CENTRALITY:
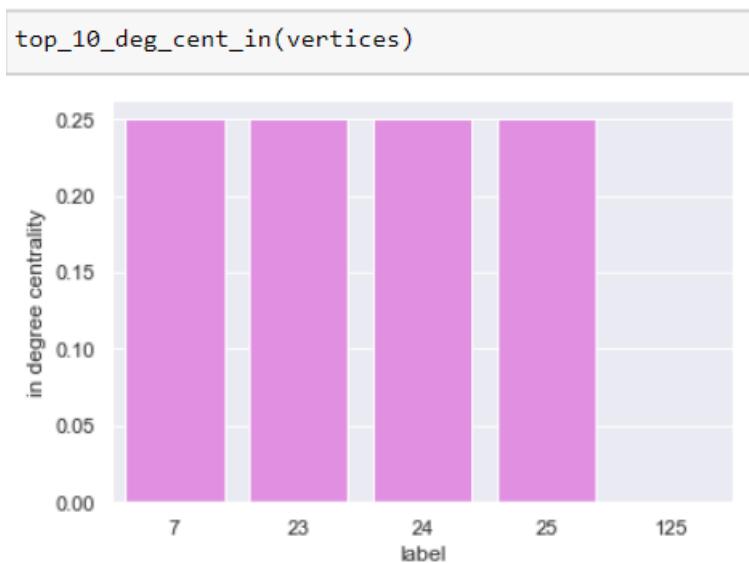
INFERENCE:

- "Node 7, 23, 24, 25" has the maximum indegree centrality while "Demersal producer" species group is taken into account, which means these species are the top keystone species in the food
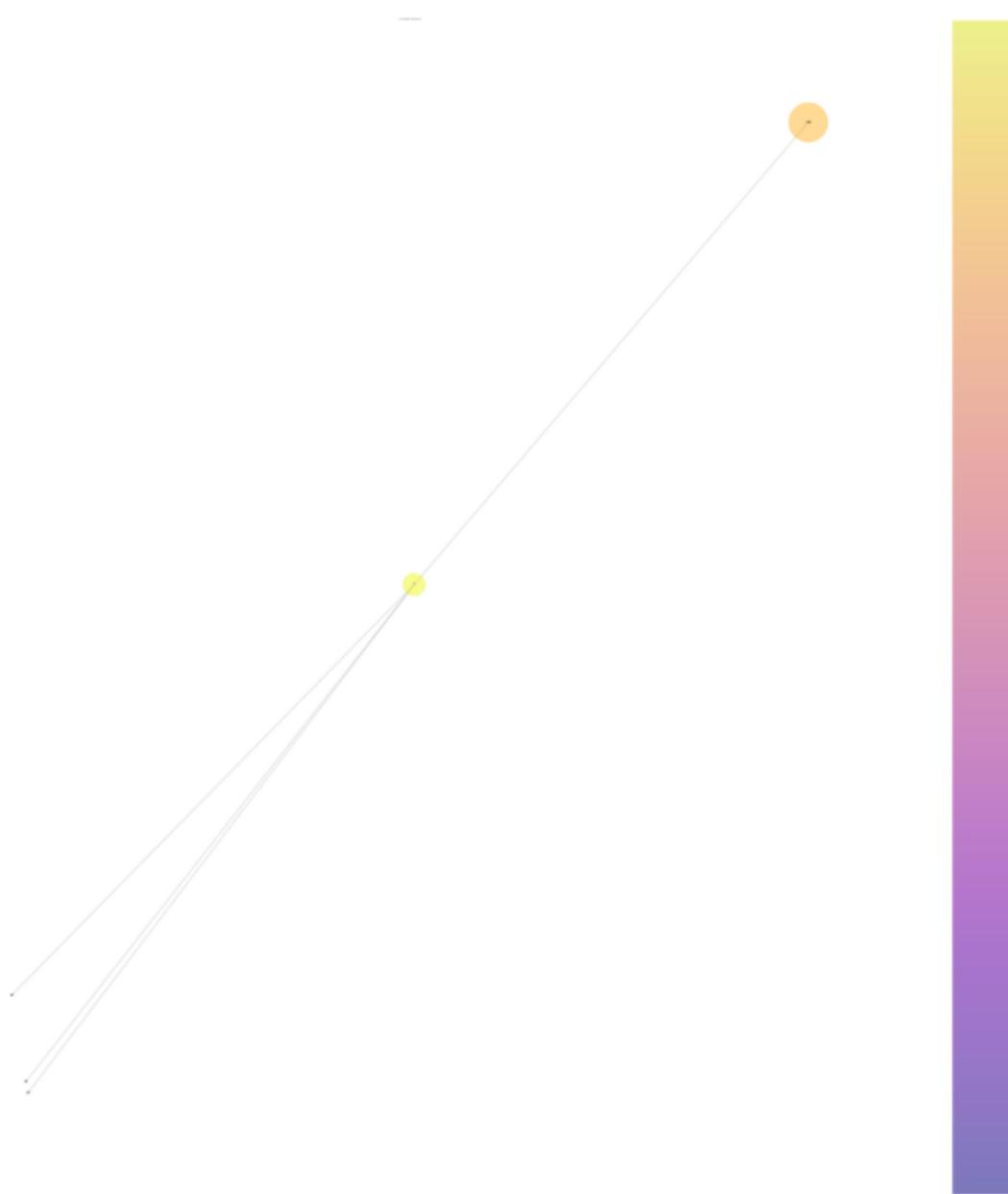
web as prey. If this prey doesn't exist then the predators depending upon this prey would cease to exist.

- It means that this species is the most influential organism in this particular species group.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

VISUALIZATION OF OUT

DEGREE CENTRALITY:

INFERENCE:
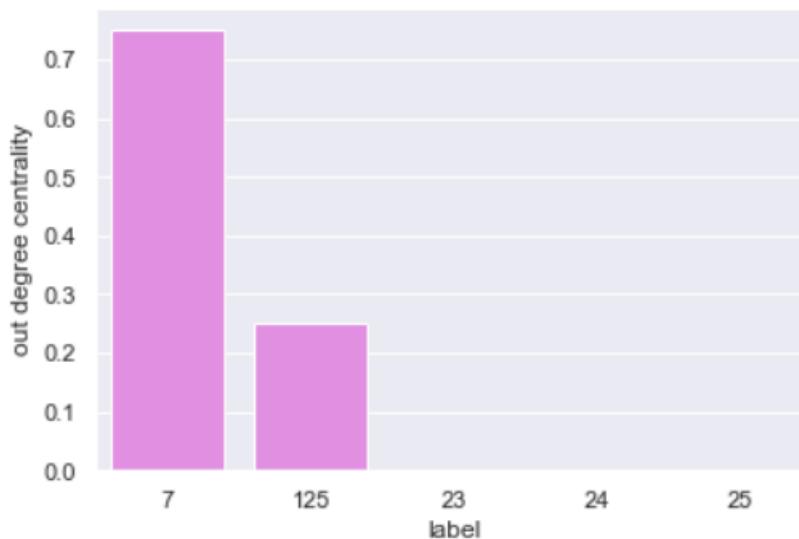
- "Node 7, 125" has the highest outdegree centrality while "Demersal producer" species group is taken into

account, which means these species are the top keystone species in the food web as predator. If this predator doesn't exist then the other predators depending upon this species would cease to exist.

- "Node 23, 24, 25" has no outdegree centrality which means these species have no preys.
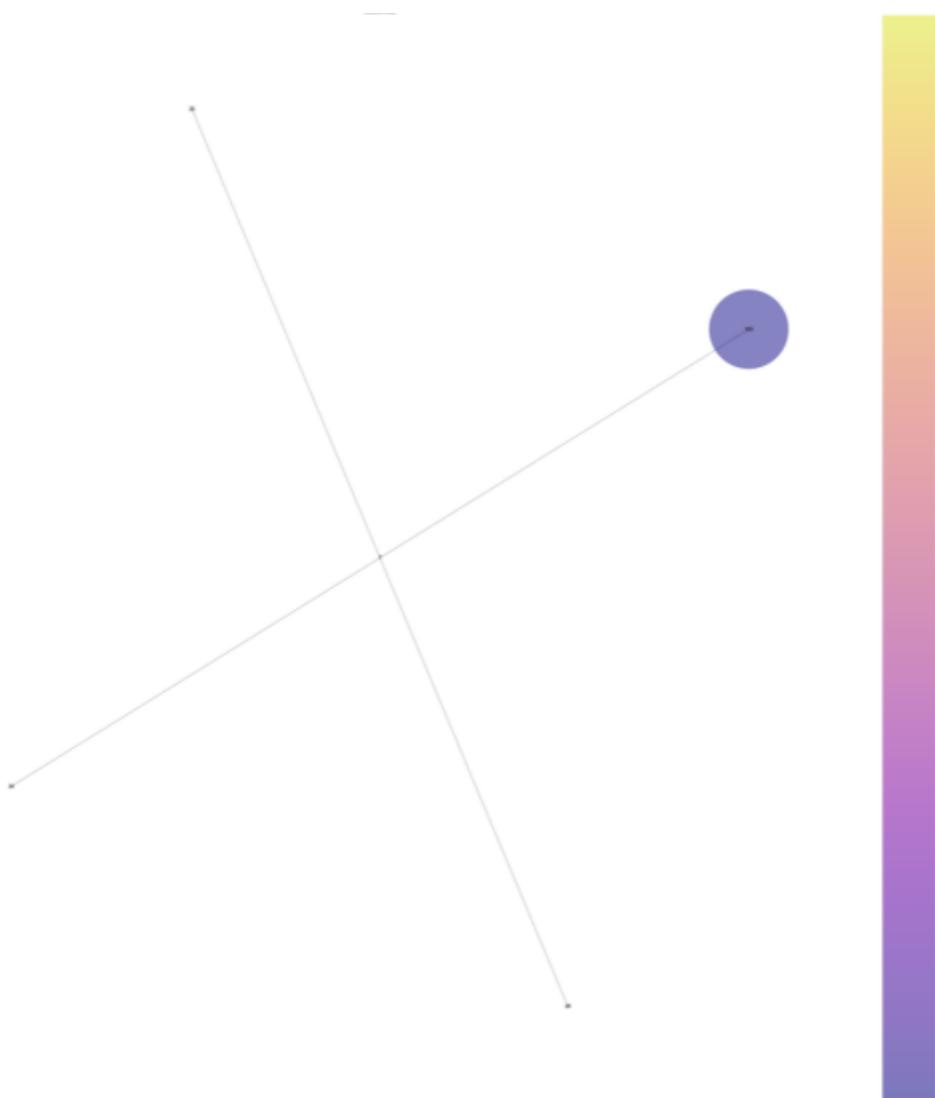
DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

BETWEENESS CENTRALITY:

VISUALIZATION OF BETWEENESS CENTRALITY:
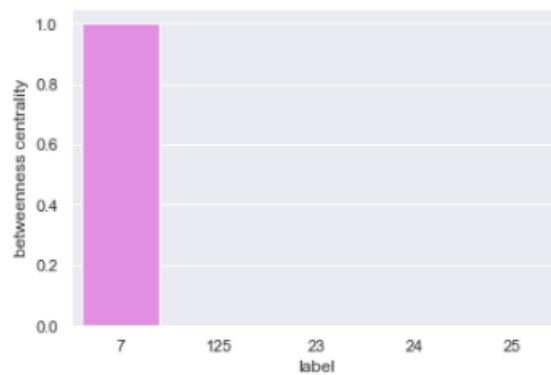
INFERENCE:

- "Node 7" has the highest betweenness centrality when "Demersal producer" species group is taken into account as either prey or predator.
- It means this species is well connected and it acts as a bridge between many pair of species.
- While Node 125, 23, 24, 25 has no betweenness centrality as the other pairs of species do not pass through these nodes.
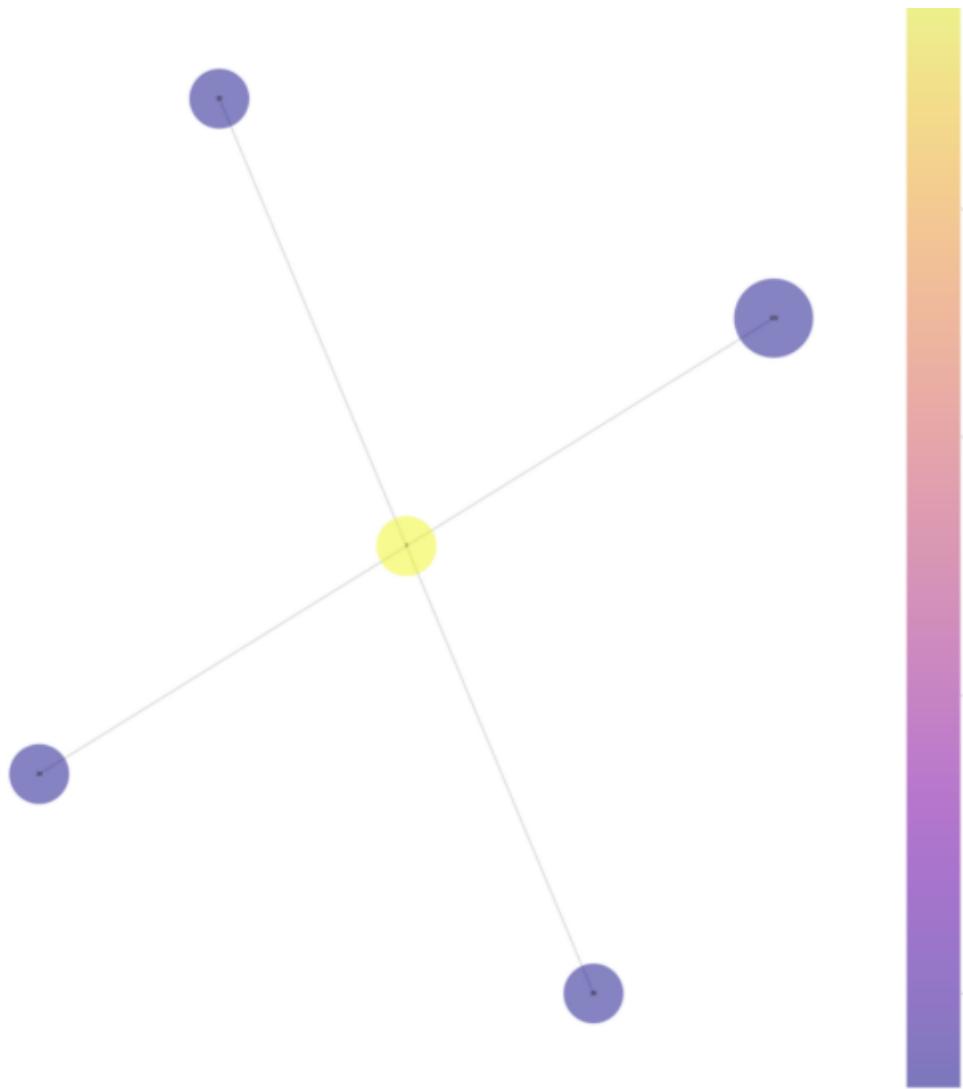
DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_bet_cent(vertices)

CLOSENESS CENTRALITY:

VISUALIZATION OF CLOSENESS CENTRALITY:
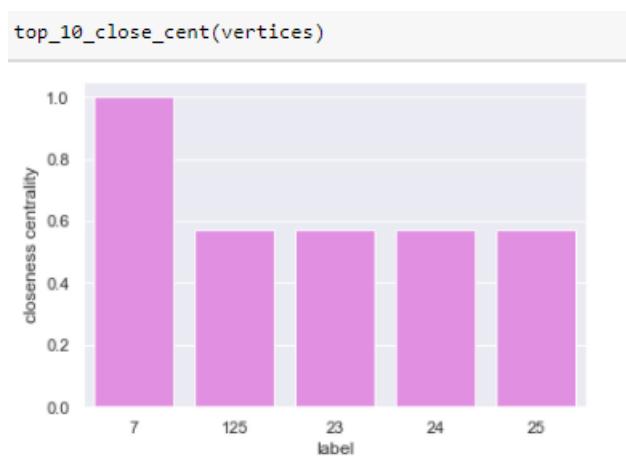
INFERENCE:

- "Node 7" has the highest closeness centrality while "Demersal producer" species group is taken into account as either prey or predator.
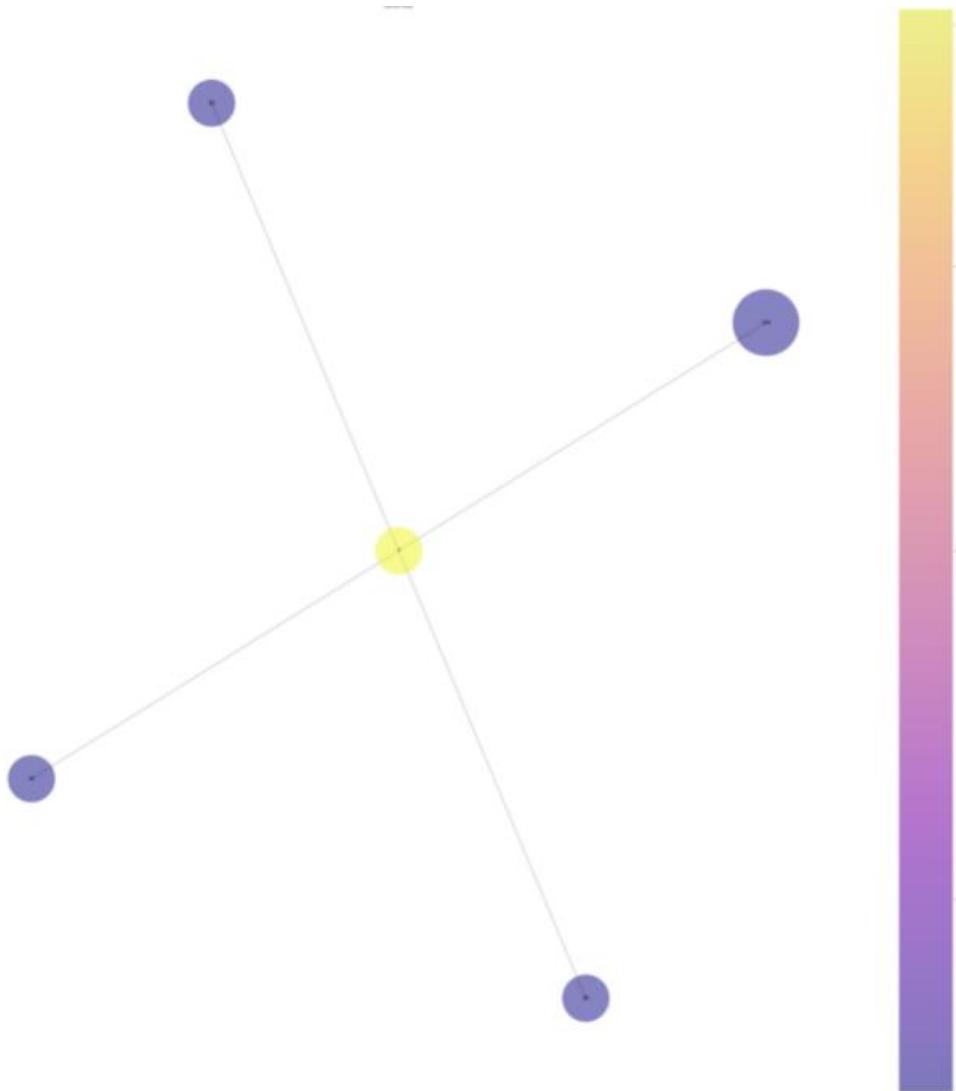
- It means this organism is closely connected to many species. So, the other species around Node 7 can either be prey or predator to this organism.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

EIGENVECTOR CENTRALITY:

VISUALIZATION OF EIGENVECTOR CENTRALITY:
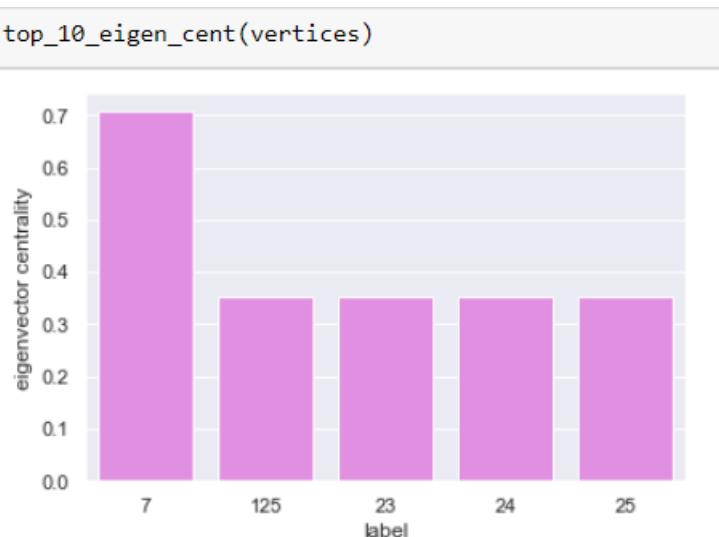
INFERENCE:

- "Node 7" belonging to the "Demersal producer" species has the highest eigenvector centrality compared to the others which means it is the most influential species in the food web.
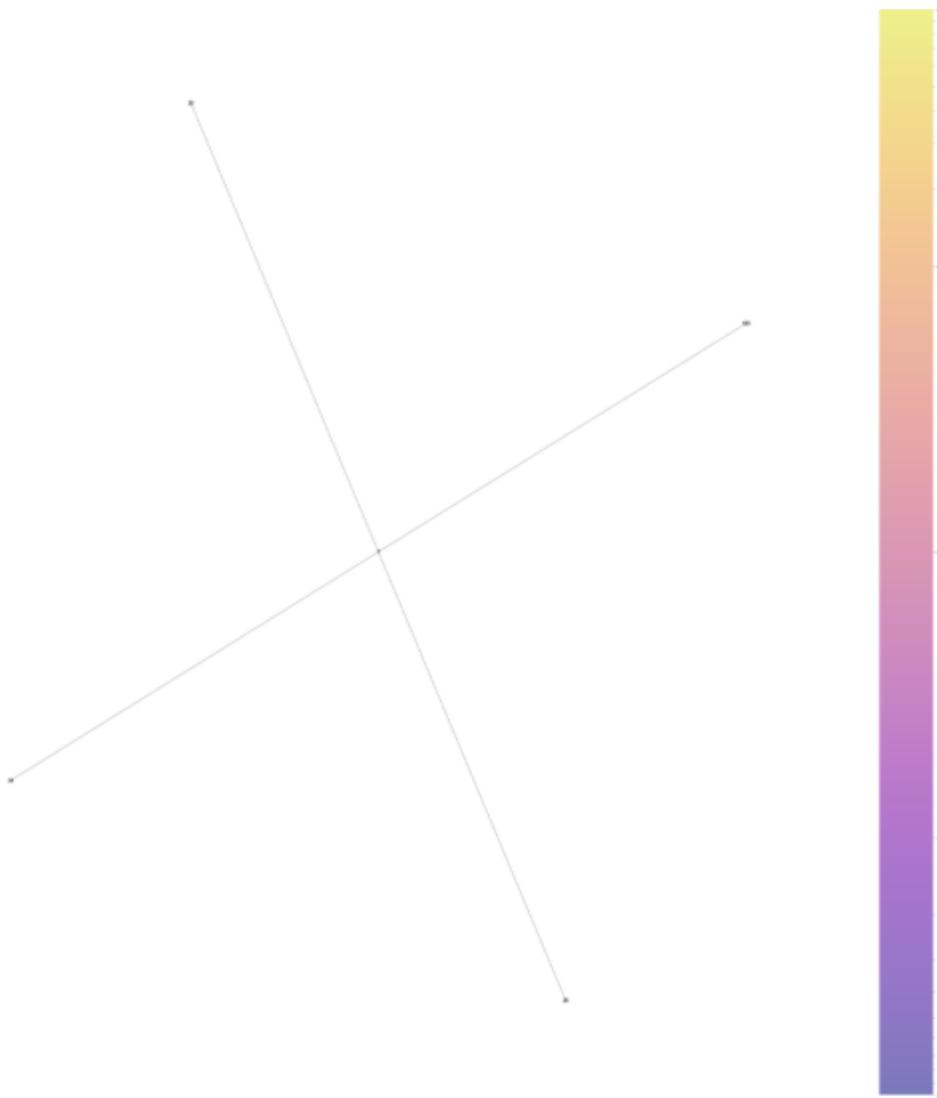
- As this species is the prey or predator of many other species which in turn has connections with other influential species, "Node 7" turns out to be the most influential species in the network.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

CLUSTERING COEFFICIENT:

VISUALIZATION OF CLUSTERING CENTRALITY:
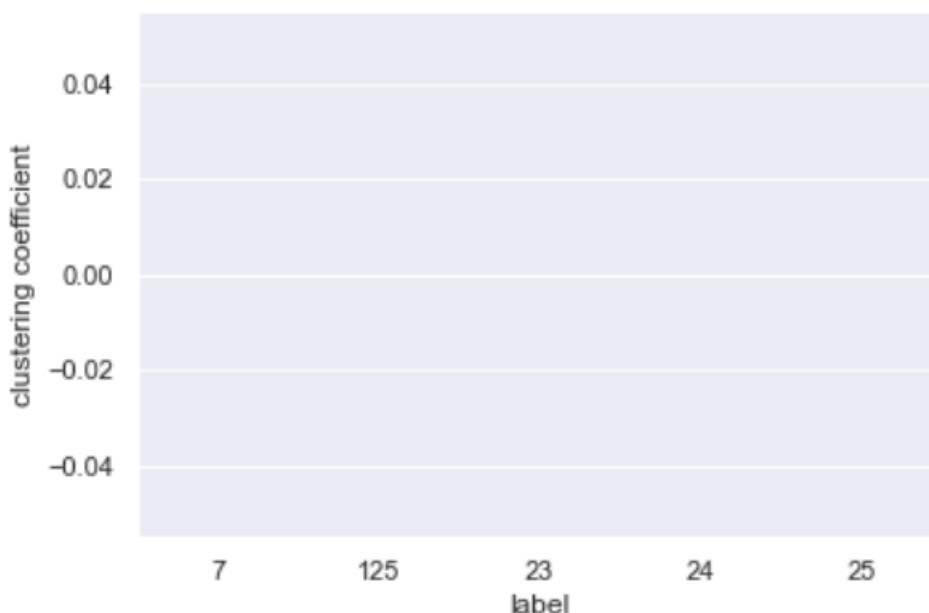
INFERENCE:

- No node belonging to the Demersal producer species group has the clustering coefficient which means this species do not form cluster.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

```
top_10_clus_cent(vertices)
```

CENTRALITIES OF EACH NODE IN SPECIES GROUP "DEMERSAL PRODUCER":

vertices

	label	counts	degree centrality	betweenness centrality	closeness centrality	eigenvector centrality	clustering coefficient
0	7	4	1.00	1.0	1.0000	0.7071	0.0
1	125	1	0.25	0.0	0.5714	0.3536	0.0
2	23	1	0.25	0.0	0.5714	0.3536	0.0
3	24	1	0.25	0.0	0.5714	0.3536	0.0
4	25	1	0.25	0.0	0.5714	0.3536	0.0
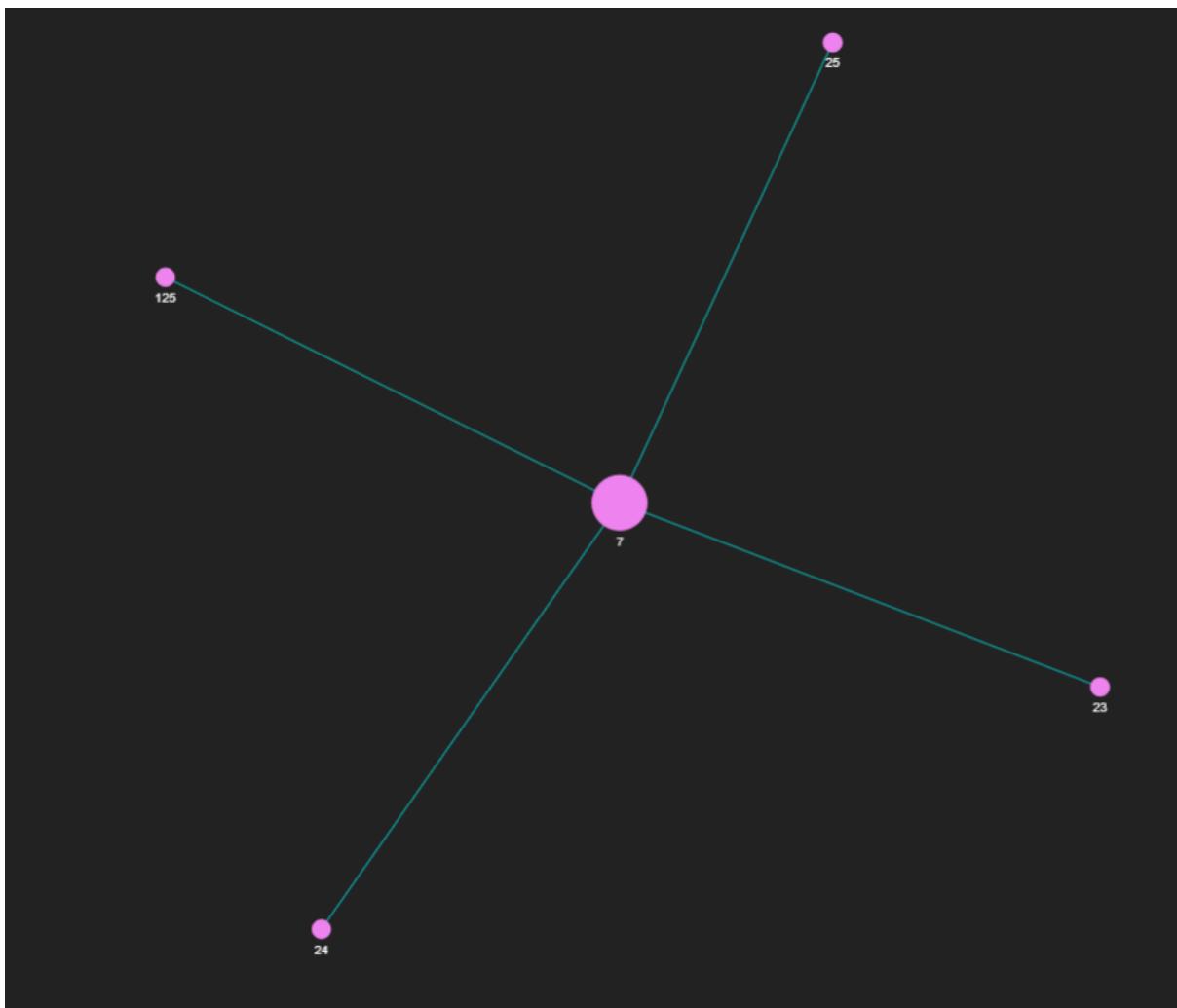
3. SPECIES GROUP: SEAGRASS PRODUCER

NODES IN THE NETWORK THAT BELONGS TO "SEAGRASS PRODUCER" SPECIES GROUP:

vertices
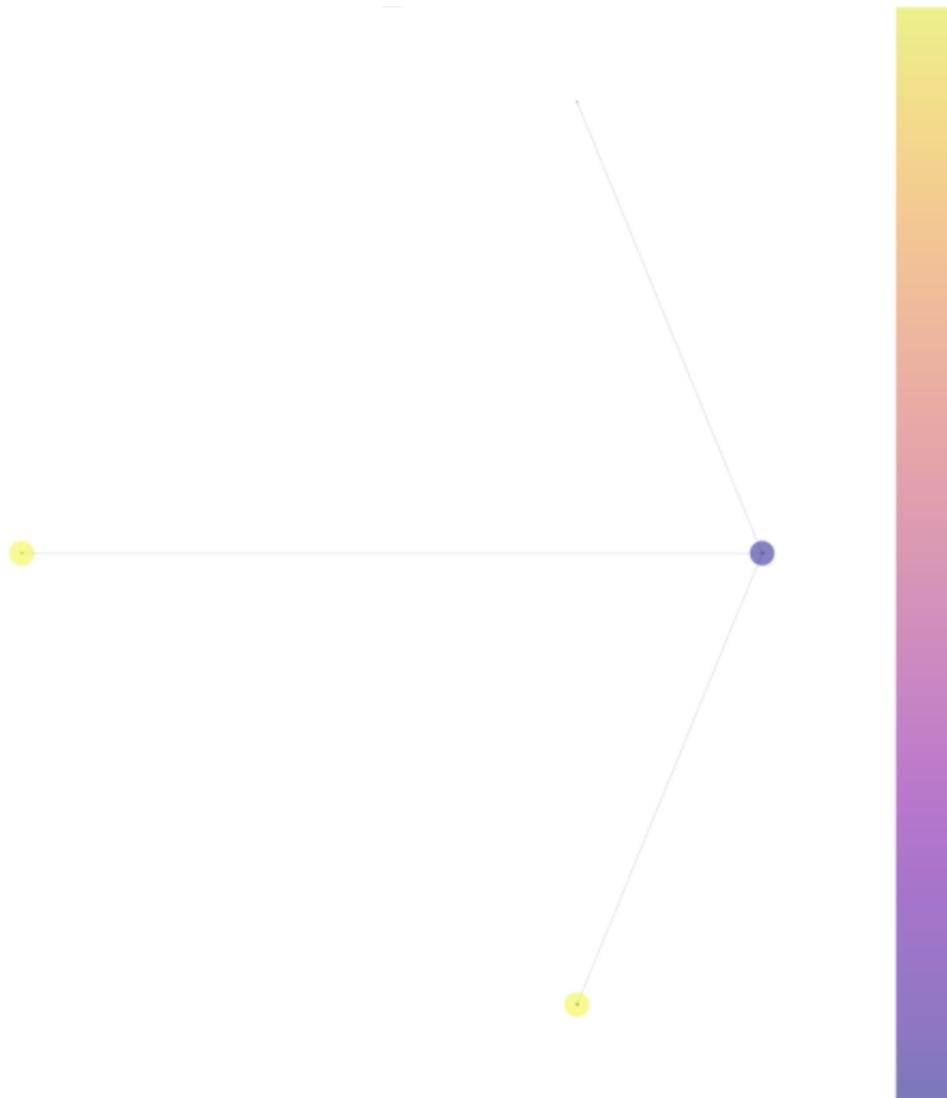
	label
0	8
1	125
2	9
3	10

SEAGRASS PRODUCER SPECIES GROUP VISUALIZATION USING PYVIZ:

DEGREE CENTRALITY:

VISUALIZATION OF INDEGREE CENTRALITY:
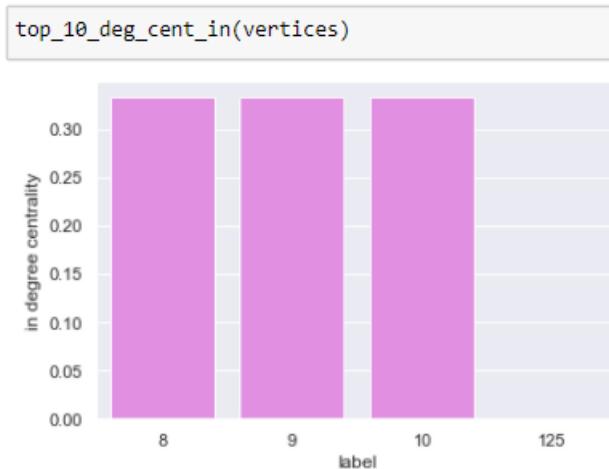
INFERENCE:

- "Node 8, 9, 10" has the maximum indegree centrality when "Seagrass producer" species group is taken into account, which means these species are
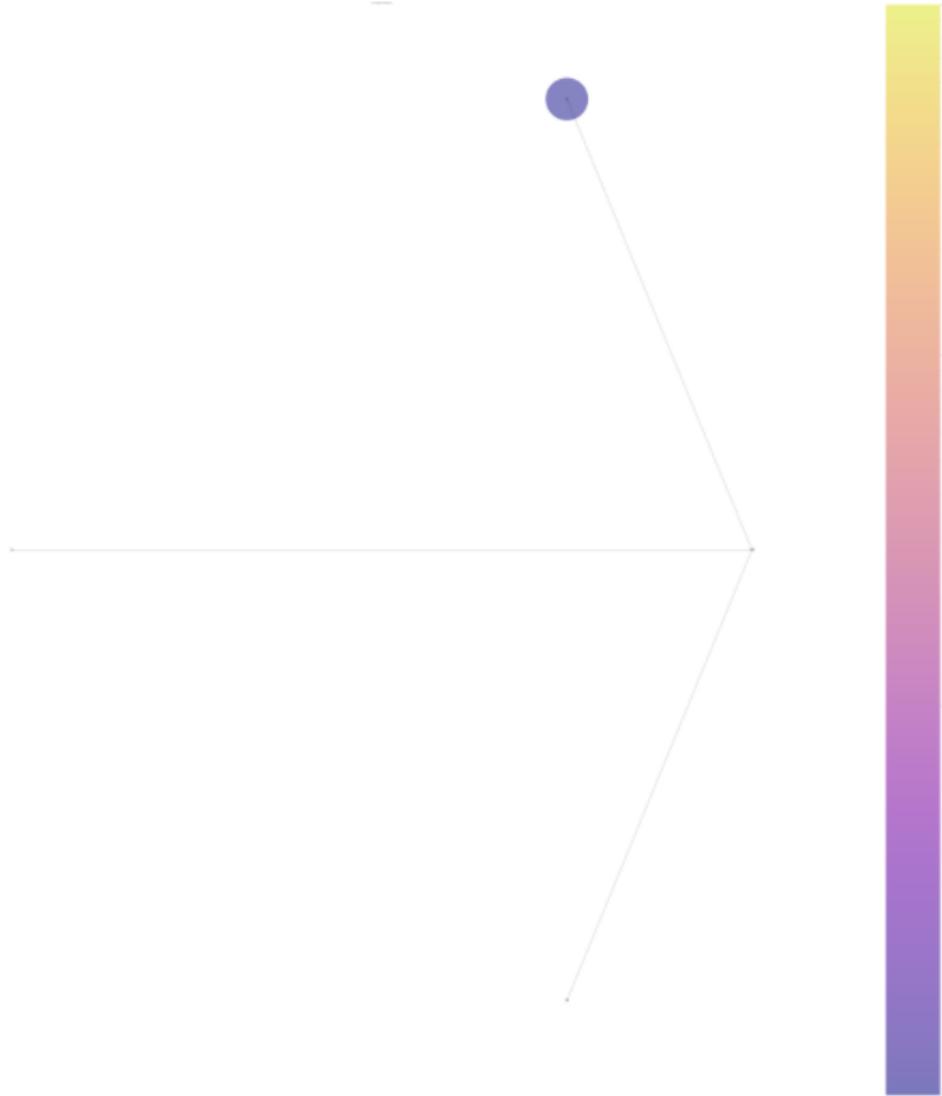
the top keystone species in the food web as prey. If these preys don't exist then the predators depending upon these preys would cease to exist.

- It means that these species are the most influential organism in this particular species group.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

VISUALIZATION OF OUTDEGREE CENTRALITY:
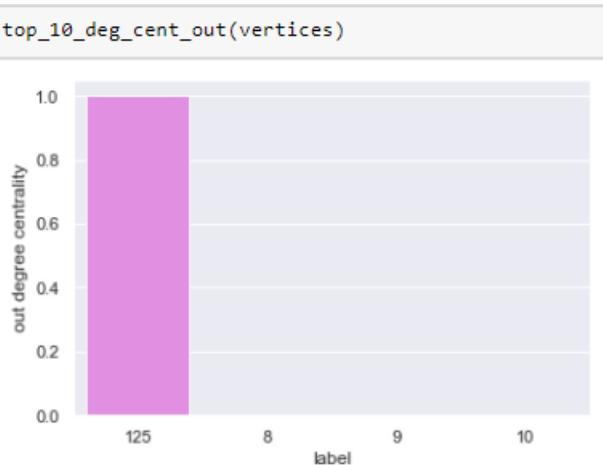
INFERENCE:

- "Node 125" has the highest outdegree centrality while "Seagrass producer" species group is taken into account, which means the species labelled as 125 is the top keystone species in the food

web as predator. If this predator doesn't exist then the other predators depending upon this species would cease to exist.
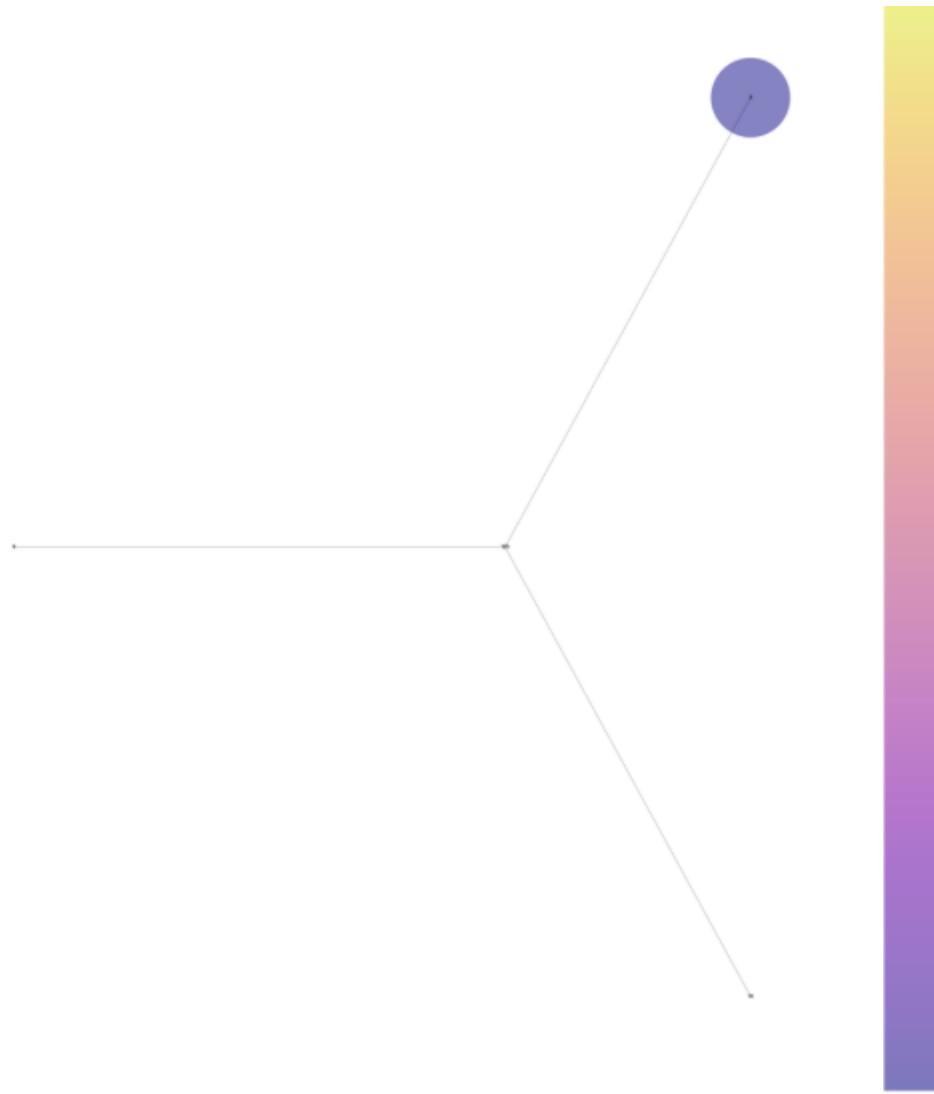
- "Node 8, 9, 10" do not have any outdegree centrality which means these species do not have any prey.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

# BETWEENESS CENTRALITY:

VISUALIZATION OF BETWEENESS CENTRALITY:

INFERENCE:
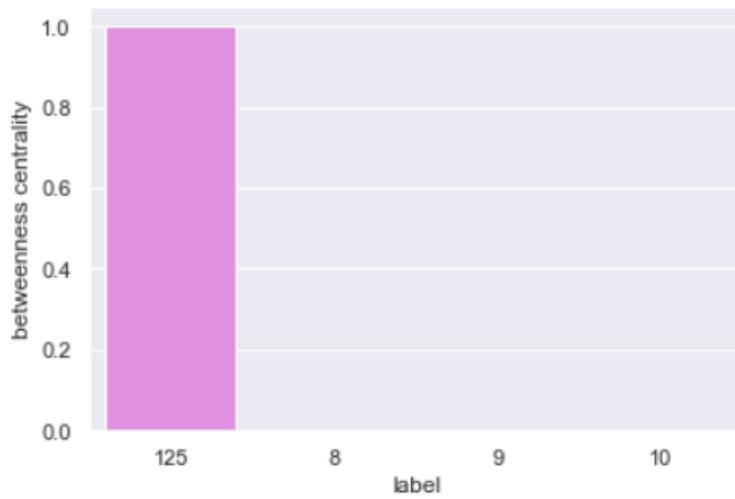
- "Node 125" has the highest betweenness centrality when "Seagrass

producer" species group is taken into account as either prey or predator.

- It means this species is well connected and it acts as a bridge between many pair of species while Node 8, 9, 10 belonging to this species group does not have any betweenness centrality which means they do not act as a bridge for any other species in the network.
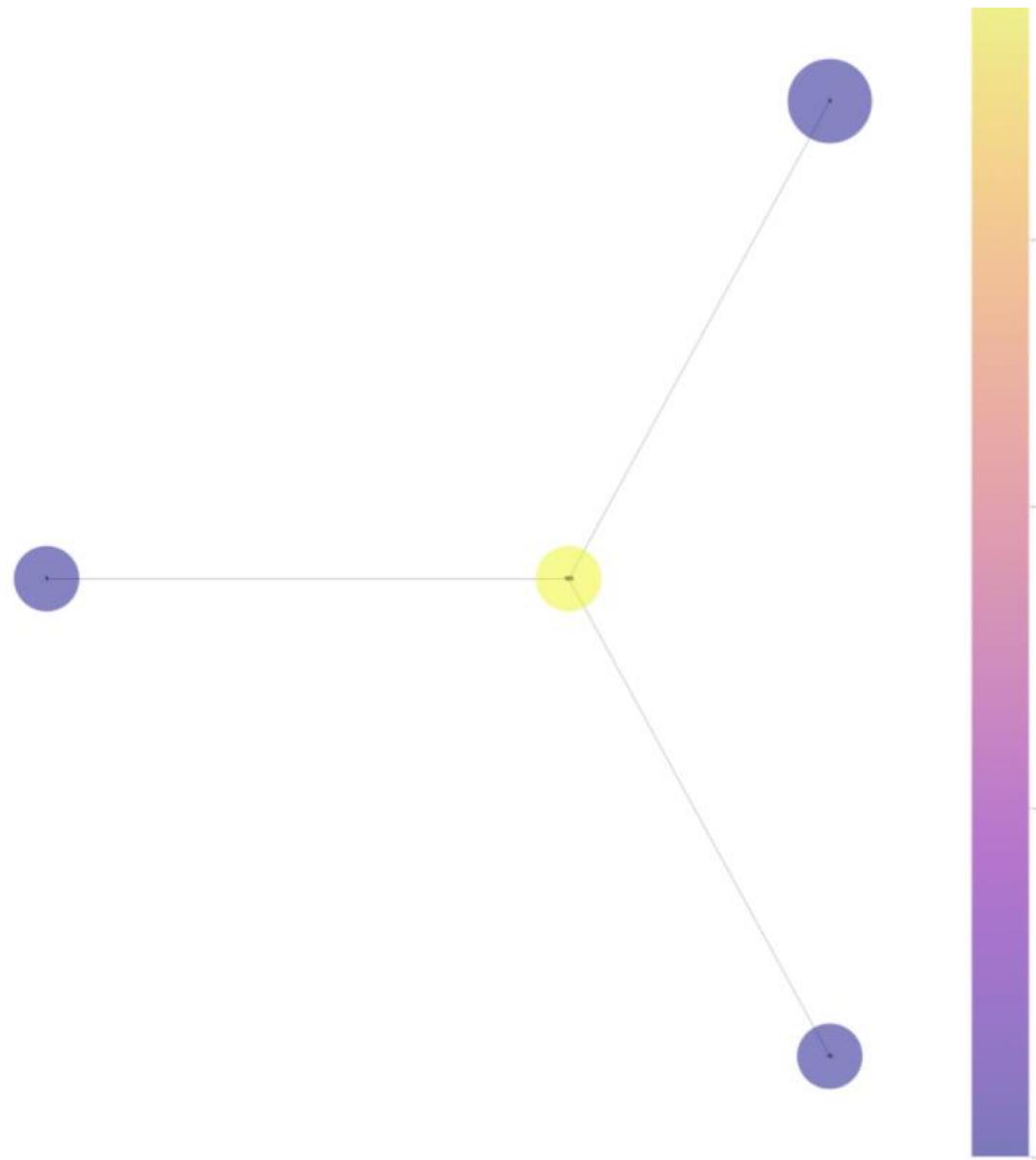
DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

top_10_bet_cent(vertices)

CLOSENESS CENTRALITY:

## VISUALIZATION OF CLOSENESS CENTRALITY:
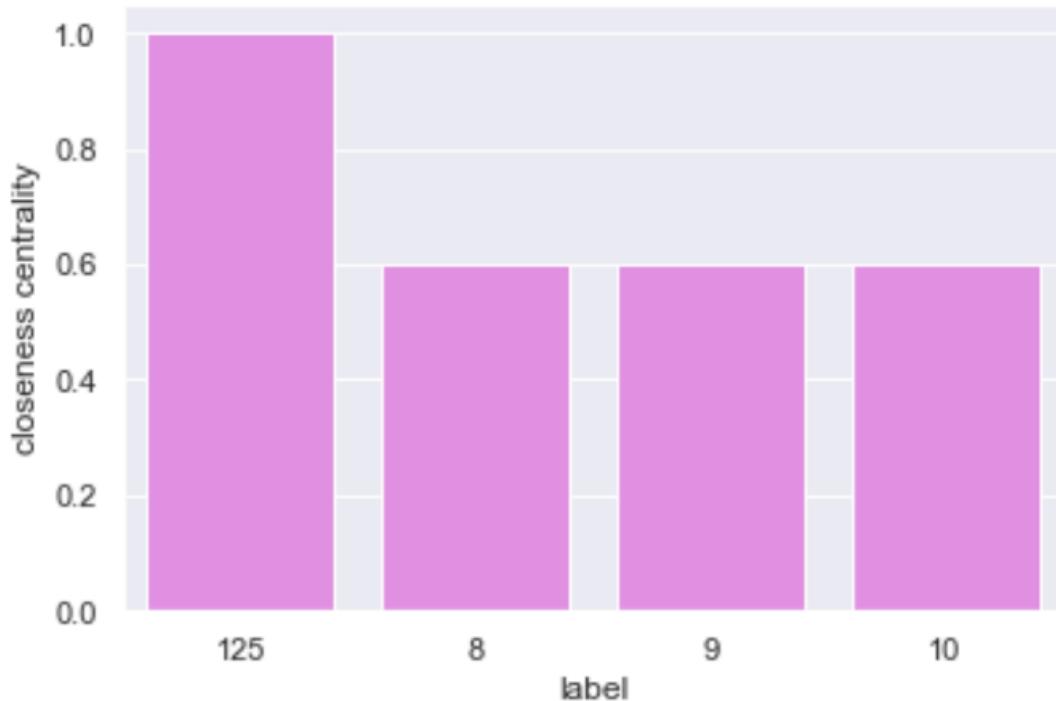
INFERENCE:

- "Node 125" has the highest closeness centrality while "Seagrass producer" species group is taken into account as either prey or predator.

- It means this organism is closely connected to many species. So, the other species around Node 125 can either be prey or predator to this organism.
- While Node 8, 9, 10 has the same closeness centrality as they are equally and closely connected to other species in the network.
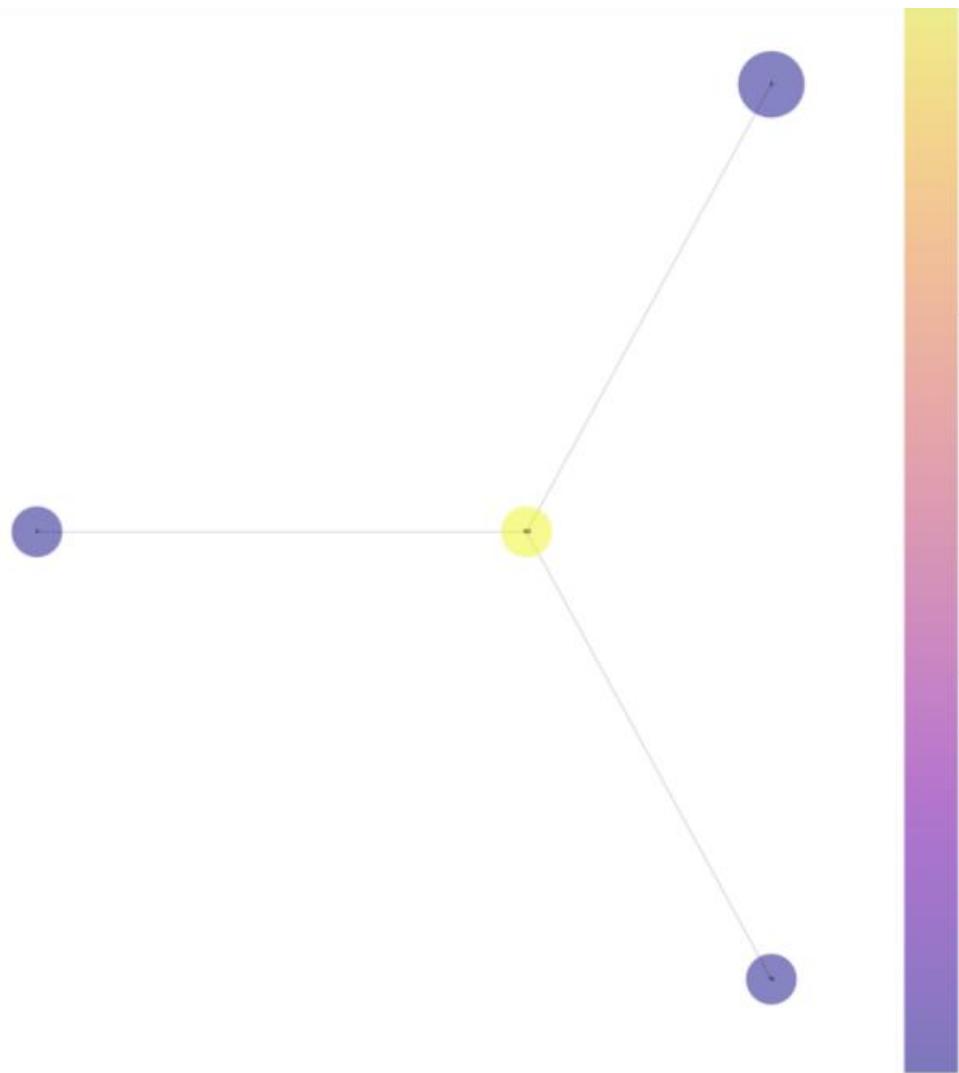
DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

```
top_10_close_cent(vertices)
```


EIGENVECTOR CENTRALITY:

# VISUALIZATION OF EIGENVECTOR CENTRALITY:

INFERENCE:
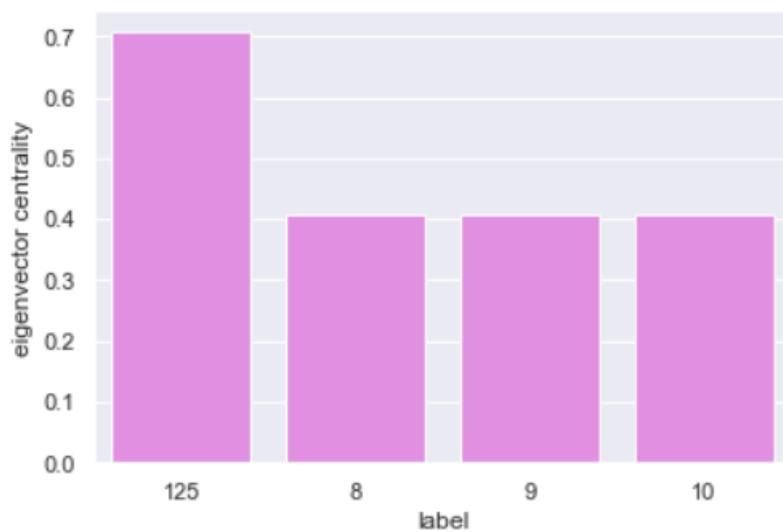
- "Node 125" belonging to the "Seagrass producer" species has the highest eigenvector centrality compared to the

others which means it is the most influential species in the food web.

- As this species is the prey or predator of many other species which in turn has connections with other influential species, "Node 125" turns out to be the most influential species in the network.
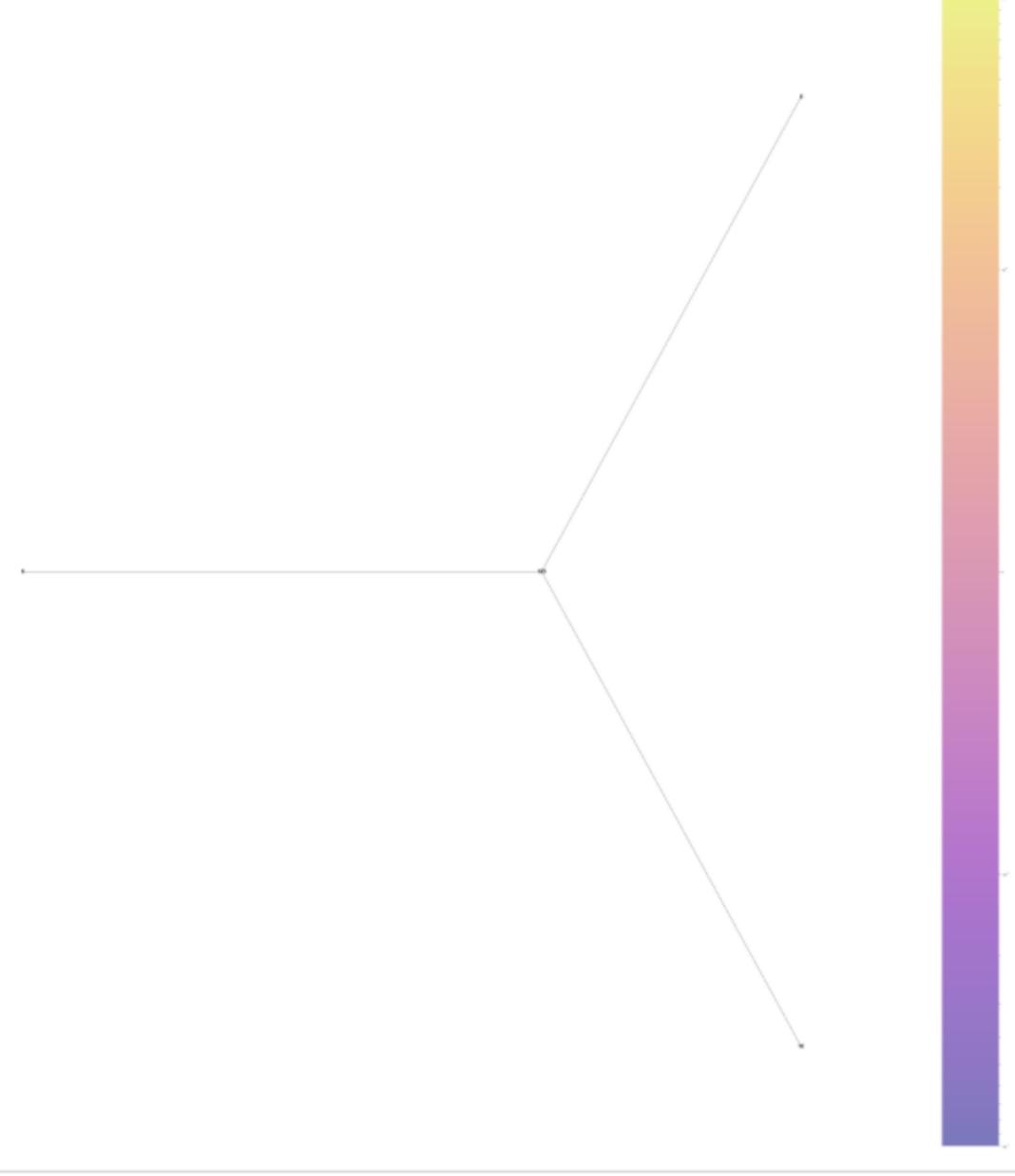
DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

CLUSTERING COEFFICIENT:

VISUALIZATION OF CLUSTERING CENTRALITY:
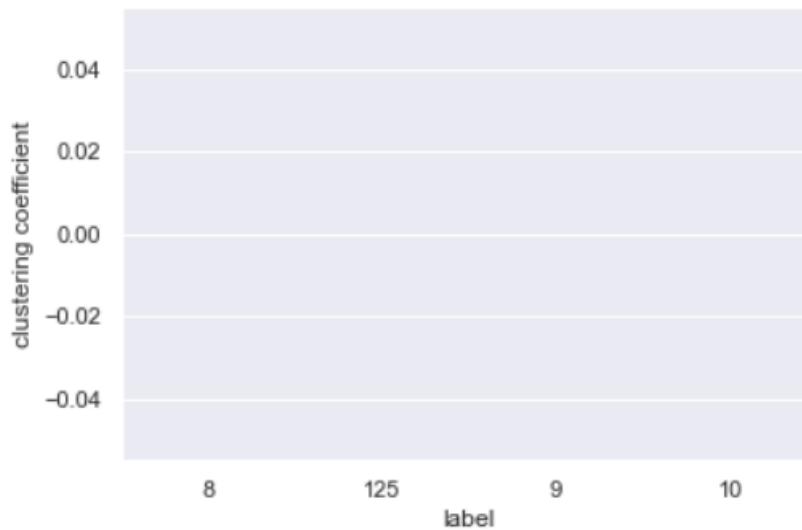
INFERENCE:

- No node belonging to the Seagrass producer species group has the clustering coefficient which means this species do not form cluster.

DISTRIBUTION OF TOP 10 NODES WITH MAXIMUM CENTRALITY:

CENTRALITIES OF EACH NODE IN THE SPECIES GROUP "SEAGRASS PRODUCER":

```
vertices
```

	label	counts	degree centrality	betweenness centrality	closeness centrality	eigenvector centrality	clustering coefficient
0	8	1	0.3333	0.0	0.6	0.4082	0.0
1	125	3	1.0000	1.0	1.0	0.7071	0.0
2	9	1	0.3333	0.0	0.6	0.4082	0.0
3	10	1	0.3333	0.0	0.6	0.4082	0.0

CONCLUSION:

Nodes with maximum centralities:

Node 127, 122, 43, 44, 123

- The analysis of the food web network through different kinds of graph centralities have shown that "Node 127, 122, 43, 44" are the most influential species and are connected to other influential species in the food web.
- While "Node 123" is the bridge that connects other pair of species that are totally unrelated.

- "Node 23 and Node 24" form bigger clusters in the food web.

Hence, it is evident that graph centralities can be used to draw inferences about the network's nodes.

TOOLS USED:

- Jupyter notebook to implement the graph centralities.
- Python libraries like Networkx and Pyviz are used for graph visualization.
- HTML is used to navigate through the interactive graph visualizations.

REFERENCES:

- https://towardsdatascience.com/graph-analytics-introduction-and-concepts-of-centrality-8f5543b55de3

- https://www.geeksforgeeks.org/eigenvector-centrality-centrality-measure/
- https://cambridge-intelligence.com/keylines-faqs-social-network-analysis/