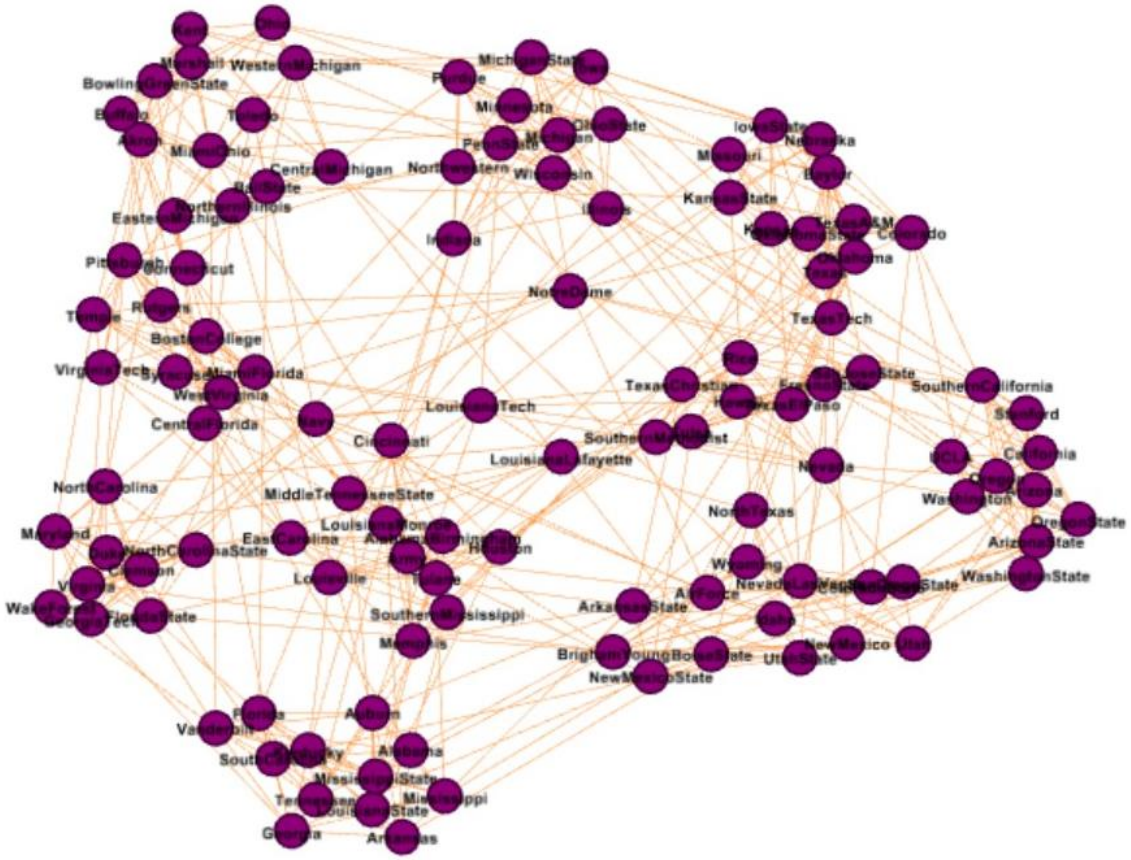


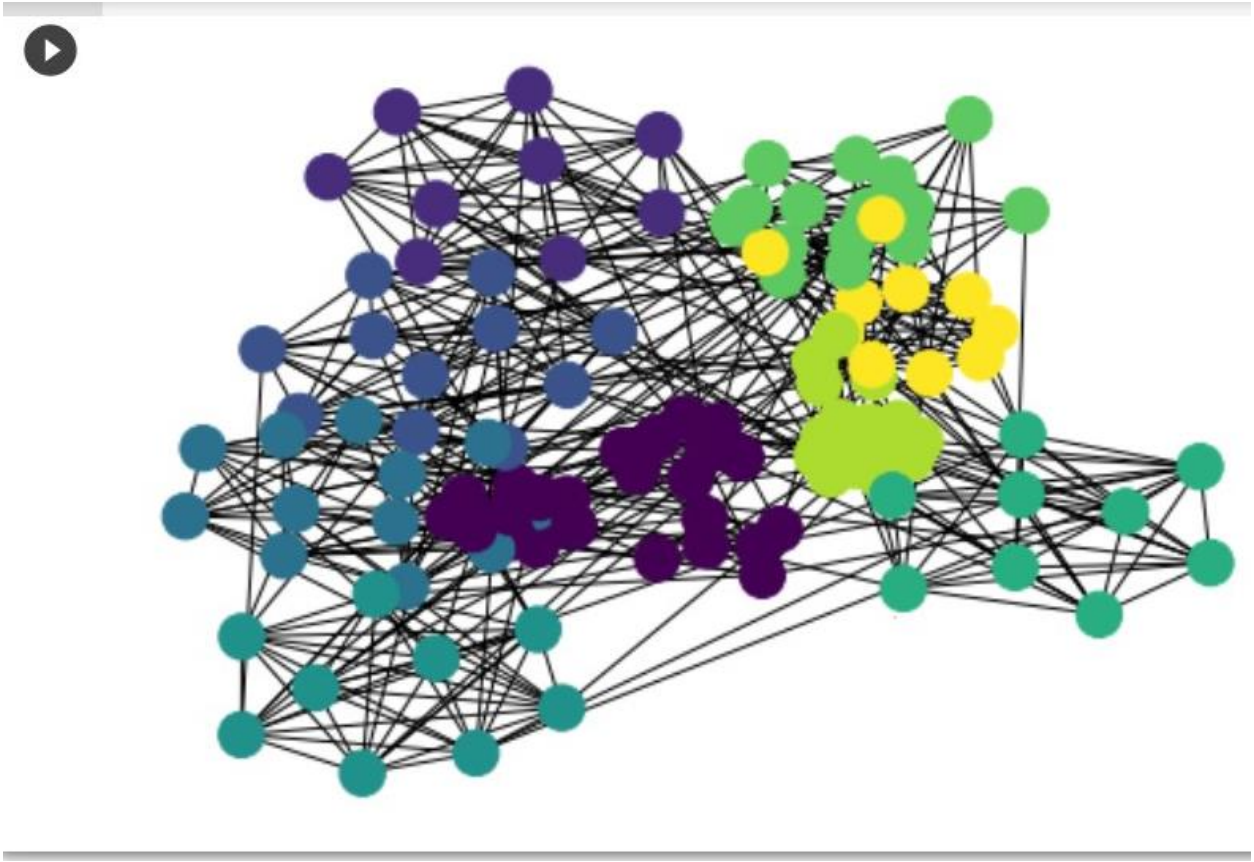
This is the network of American football games between Division IA colleges during regular season Fall 2000.

Total Nodes = 115

Total Edges = 613

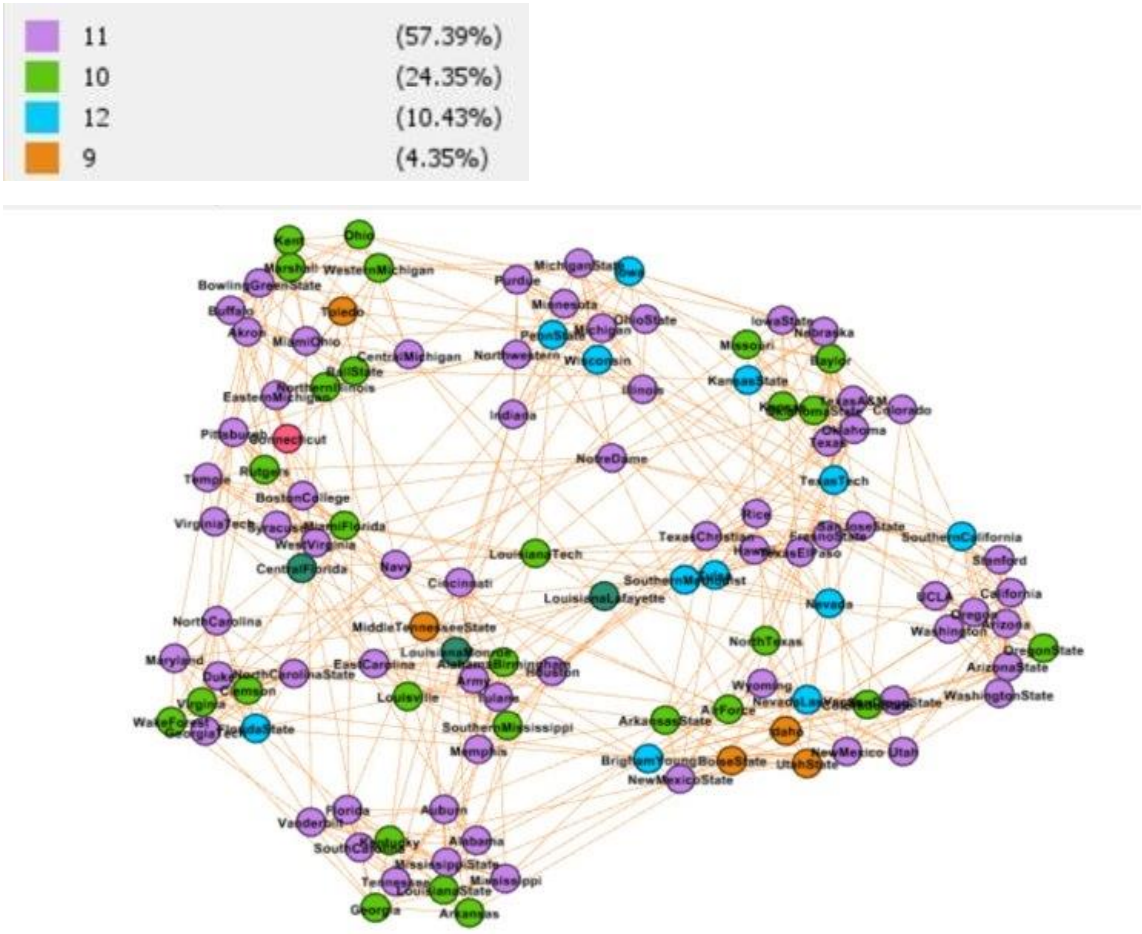
Nodes here denote colleges from different parts of US which participated in American football.

Edges here denote football matches between the colleges.



This is the network dataset based on community partition. Analyzed on python with networkx. As we can see the communities, so, it simply describes the regions which contains several teams and when won within them goes to inter regional matches.

Top 10 Nodes with Highest Degrees in our network



1) Brigham Young: 12

2) Florida State: 12

3) Iowa: 12

4) Kansas State: 12

5) New Mexico: 12

6) Texas Tech: 12

7) Penn State: 12

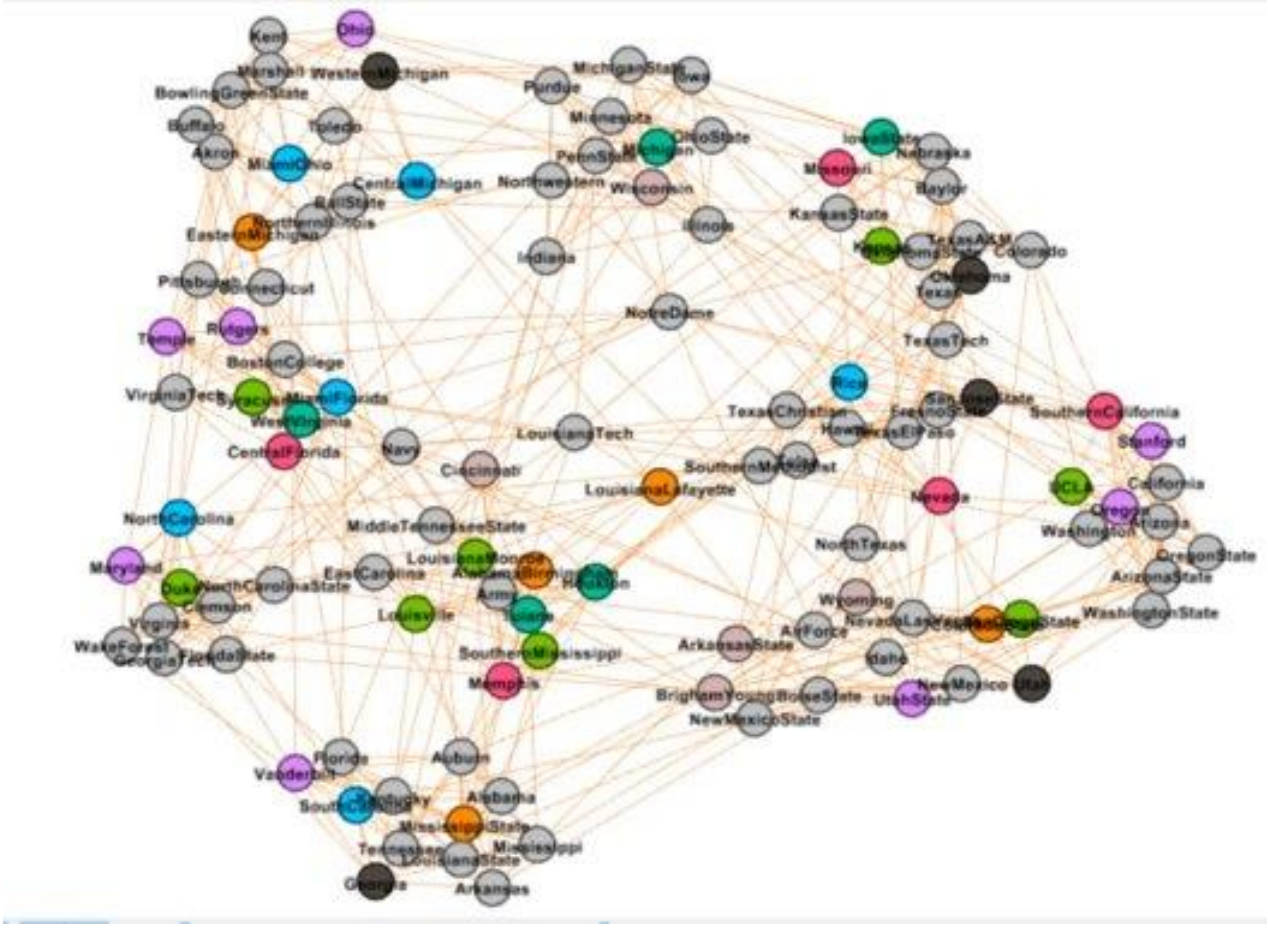
8) Southern California: 12

9) Arizona State: 11

10) San Diego State: 11

Generally speaking, “Degree” of a node represents the number of edges it is connected to. In our network, degree of a node represents the number of matches the college has played in the term.

Top 10 nodes with highest Closeness centrality



1) Louisiana Tech: 0.436782

2) Navy: 0.435115

3) Tulsa: 0.430189

4) Indiana: 0.426966

5) Penn State: 0.425373

6) Brigham Young: 0.423792

7) Cincinnati: 0.423792

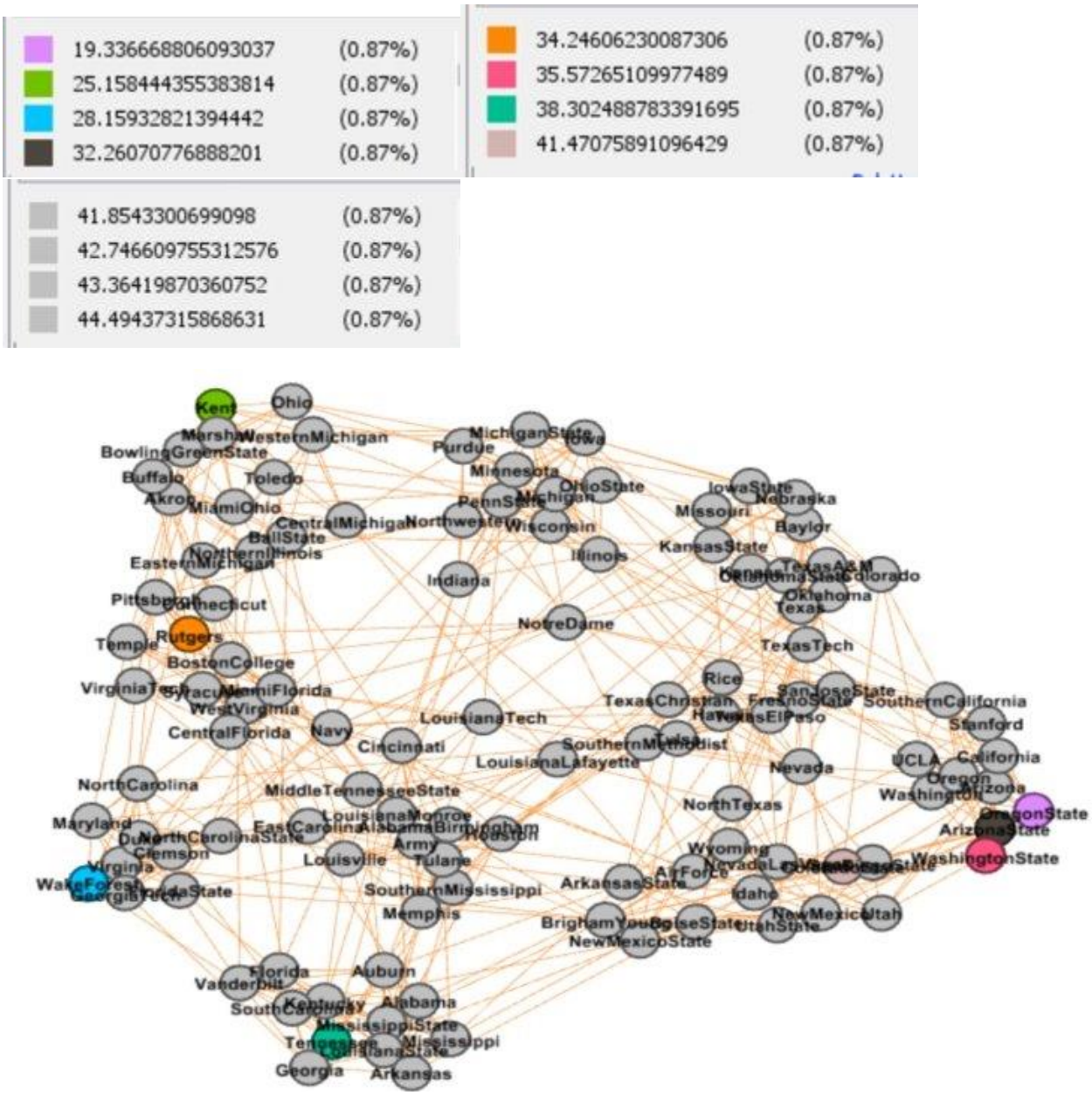
8) Wyoming: 0.423792

9) Wisconsin: 0.423792

10) Arkansas State: 0.423792

The closeness centrality of a node measures its average farness to all other nodes to all other nodes. So, in terms of this network, we can see that the teams are named after colleges from different parts of the US. So, the node with highest closeness centrality defines that there were so many matches of that team with teams from nearest places from that specific place.

Top 10 nodes with highest Betweenness Centrality



1) Notre Dame: 215.985774

2) Brigham Young: 209.267763

3) Navy: 187.826344

4) Louisiana tech: 185.647806

5) Central Michigan: 162.228191

6) New Mexico State: 155.482003

7) Cincinnati: 153.529962

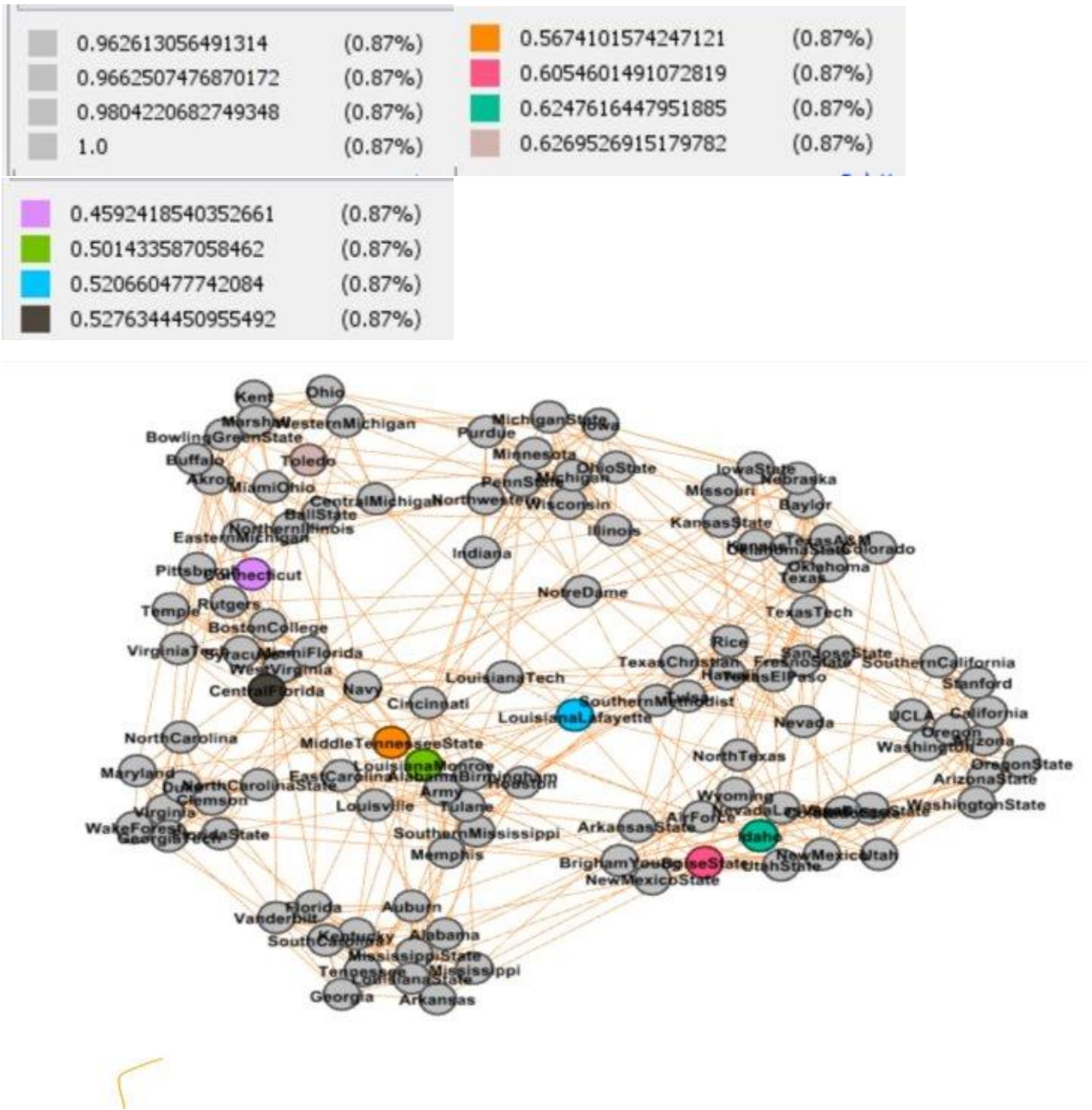
8) Kansas State: 148.594507

9) Alabama: 148.441358

10) Wyoming: 143.076815

Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. In our network, as we can see there are colleges in certain regions and whose centrality is low, it means there might be some matches between the colleges within the same region and the specific college which has low betweenness centrality hasn't played many inter-regional matches because it might have lost itself in the inter-colleges of the same region. There are some regions which has very less colleges. In our network, Notre Dame is from that certain region. It has to play matches from other regions because maybe it's the only team from that region.

Top 10 nodes with highest Eigenvector centrality



1) Nevada: 1

2) Southern Methodist: 0.980422

3) Tulsa: 0.966251

4) Iowa: 0.962613

5) Wisconsin: 0.95925

6) Southern California: 0.954954

7) Penn State: 0.940001

8) Nevada Las Vegas: 0.933449

9) San Jose State: 0.933405

10) Fresno State: 0.930233

Eigenvector centrality is a measure of the influence of a node in a network. A high score means that a node is connected to other nodes that have high scores. In this network's context, it simply means that the team who has won many matches with the strongest teams, that is, who is possibly the winner of the term will have highest eigenvector value because the team which played matches with other teams also has high eigenvector centrality.

Analysis of Centralities

The difference in degrees between the nodes is not much as most of the teams played 7 to 12 at most matches. Many teams have 12 degrees but there is only one with 7 degree which is Connecticut.

The closeness centrality lies in the range between 0.354 and 0.436. Louisiana Tech has the highest and Oregon State has the lowest. There is a very tiny difference between their closeness centrality which is 0.001 to 0.01.

The betweenness centrality lies in the range between 19.33 and 215.985. Notre Dame has the highest betweenness centrality while Oregon State being the lowest. It has a difference of almost +2 in every team.

The eigenvector centrality lies in the range between 0.459 and 1. Nevada being the highest and Connecticut being the lowest. The difference between the teams was almost 0.01.

Connecticut has the lowest degree and the lowest eigenvector centrality, it implies that this team is the last on the leaderboard of football season.

Brigham Young has the highest degree, it might be the winner of the football season.

PageRank

The PageRank measures the importance of each node within the graph, based on the number incoming relationships and the importance of the corresponding source nodes.

```
dict = nx.pagerank(G)
top = {}
temp_val = 0
temp_name = ''
for x in range(5):
  for i in dict:
    if i in top.keys():
      continue
    else:
      if dict[i] > temp_val:
        temp_val = dict[i]
        temp_name = i
  top[temp_name] = temp_val
  temp_val = 0

for i in top:
  print(i, ' ', top[i])
```

```
TexasTech   0.009676913422795617
FloridaState   0.009641122343366627
KansasState   0.009617056478393123
BrighamYoung   0.009587665333872 54
PennState   0.009548177032061969
```

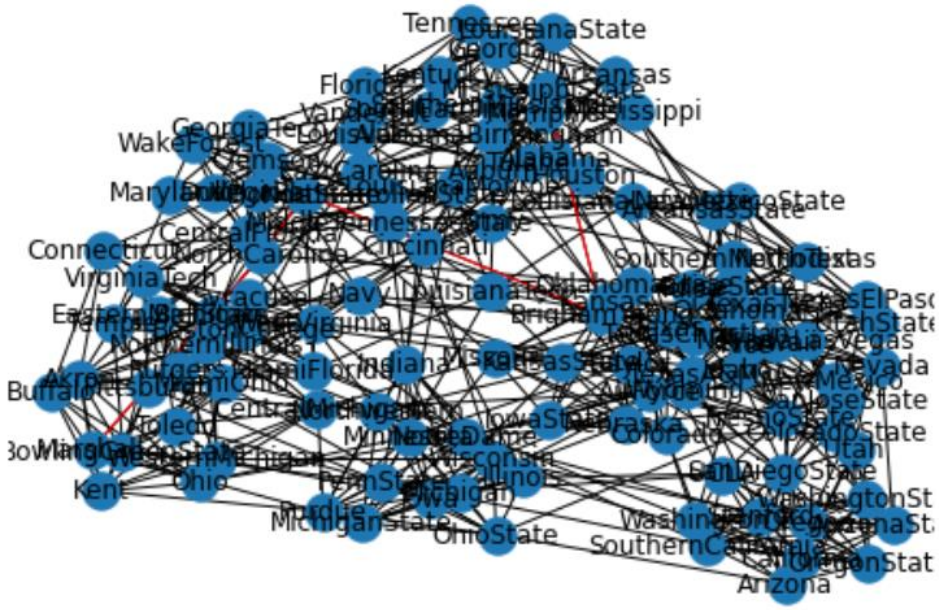
Here, Teams with high PageRank means that they had matches with the strongest teams in the season. These top 5 teams were the strongest

teams and matches with the other strongest teams. Because PageRank is all about importance of the node and quality. So, here it refers to the link between the strongest teams. OfCourse, no weak team will be played against strongest one in the finals. So, these teams should be the list of teams in Finals.

The diameter of graph is the maximum distance between the pair of vertices. It can also be defined as the maximal distance between the pair of vertices. The diameter of this graph is 4.

```
l_s_path_edges = list(zip(l_s_path, l_s_path[1:]))
pos = nx.spring_layout(G1)
nx.draw(G1, pos=pos, with_labels=True)
nx.draw_networkx_edges(G1, edge_color='r', edgelist=l_s_path_edges, pos=pos
print(nx.diameter(G1))
```

```
4
```



The diameter here, indicates that every source team is related to destination team via at most 2 other teams.

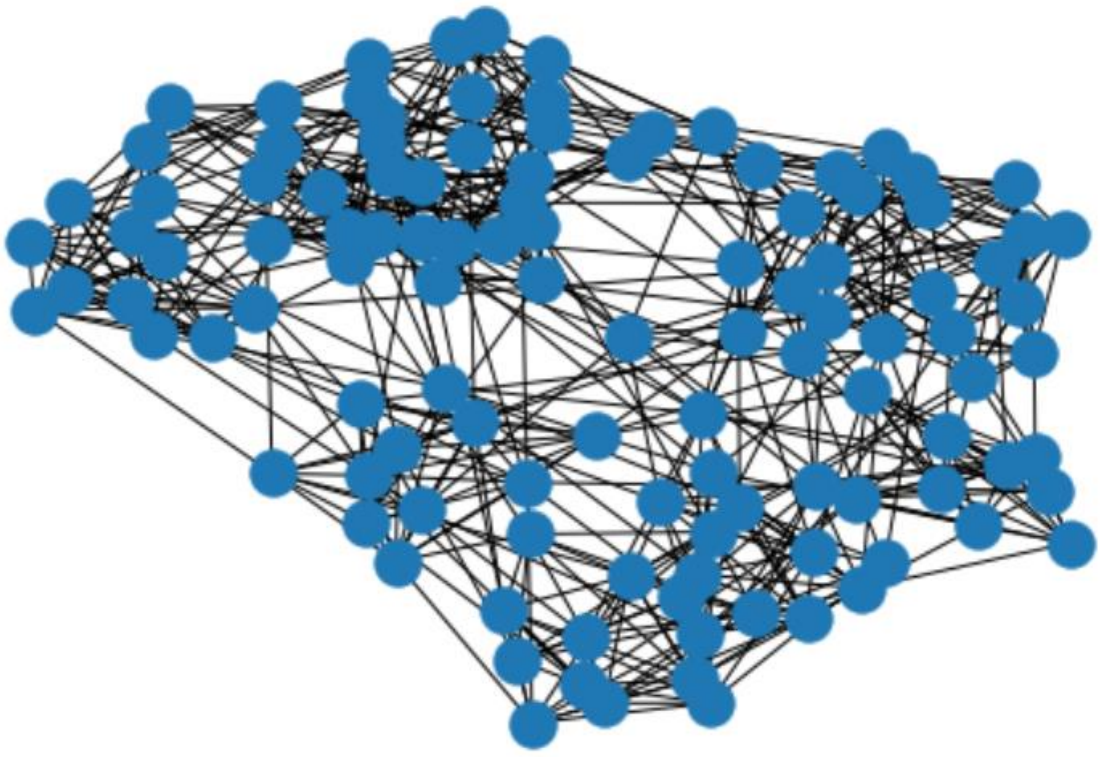
Few Questions to be discussed

1) What if we remove teams with highest degrees?

- Well, in this case there will be less teams which means less competition and less matches. Other centralities will also be affected by it but it totally depends on the team if it played a match with that team directly or not attributes in the network will change. For example, NevadaLasVegas's eigenvector centrality was 0.933 but after removal of all the 12-degree nodes and its adjacent edges, its centrality decreased but SanJoseState centrality increased from 0.933 to 0.95286.

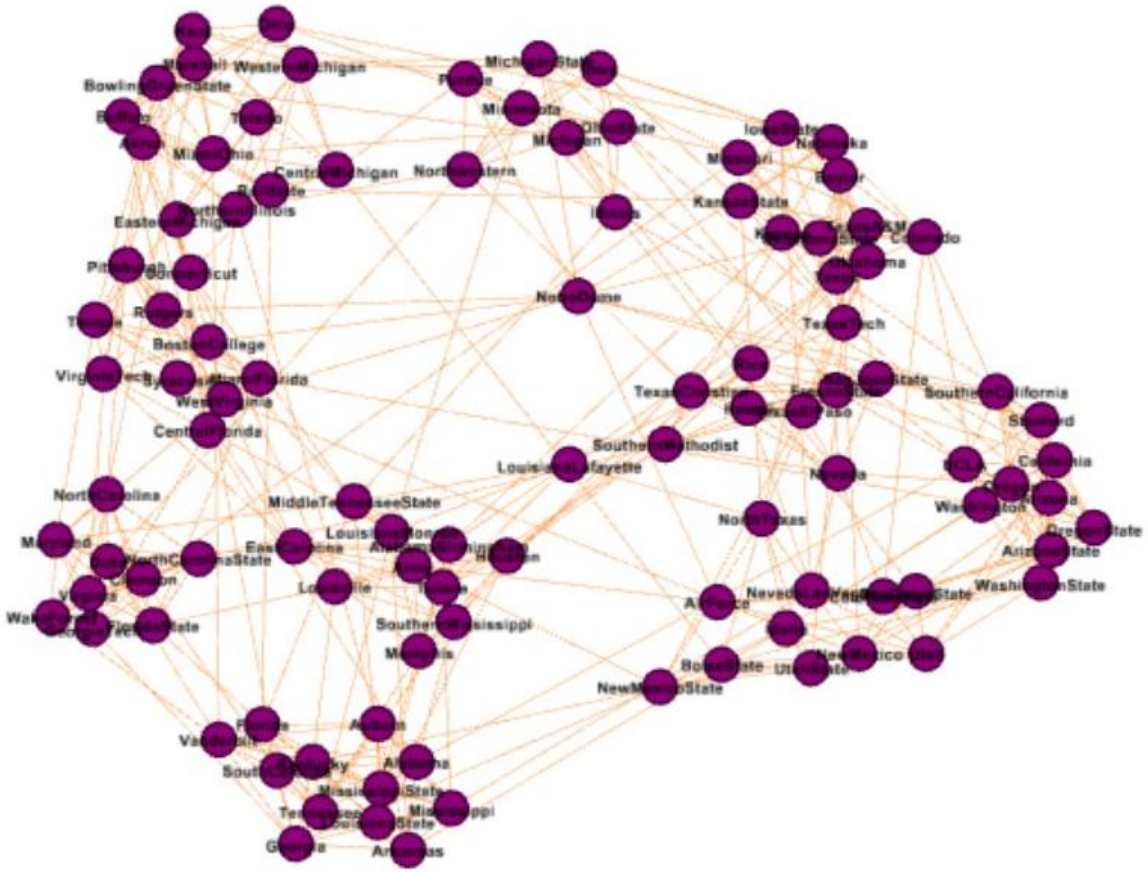
This is how our network will look after removing nodes whose degree is 12.

```
for node in G1.nodes():
  if G1.degree(G1.nodes[node]) == 12:
    G1.remove_node(node)
nx.draw(G1)
```



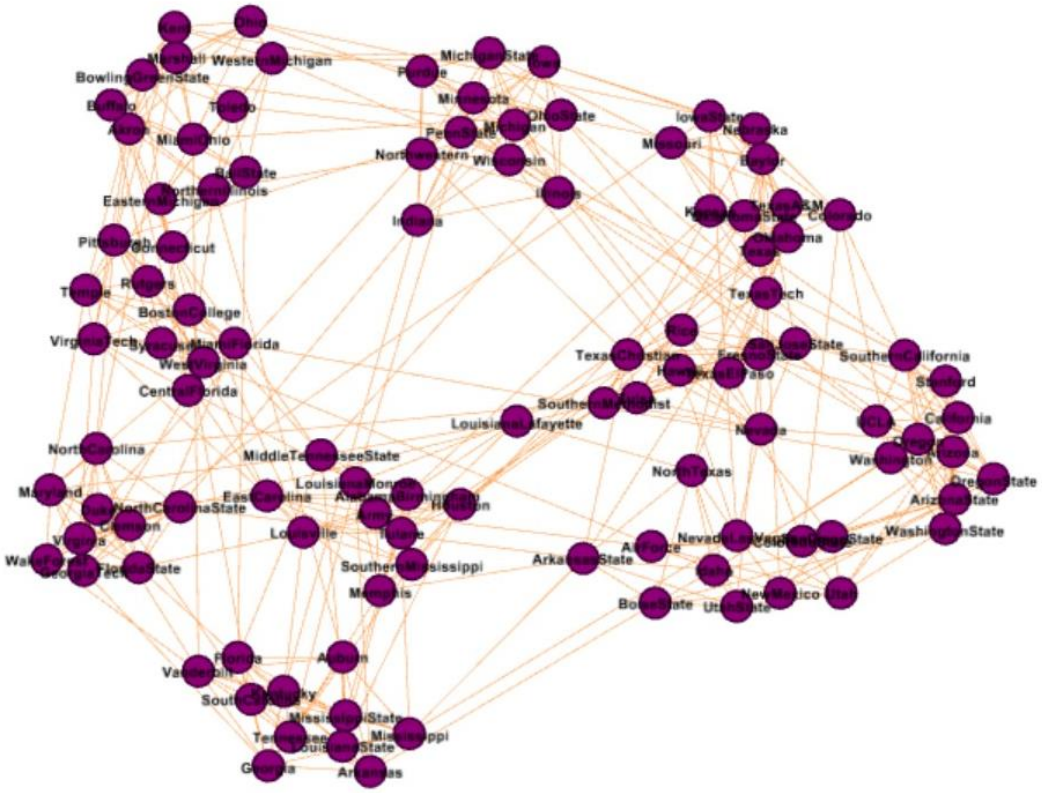
Graph on python with networkx

2) What if we remove teams with highest eigenvector centrality? - Then, we are not going to have the teams which were the strongest. If we see it, on the other hand, audience won't be excited about it too because people are excited about their favourite player/favourite team which is obviously a stronger team. So, less likely teams which will have then highest eigenvector centrality will be most likely to win.



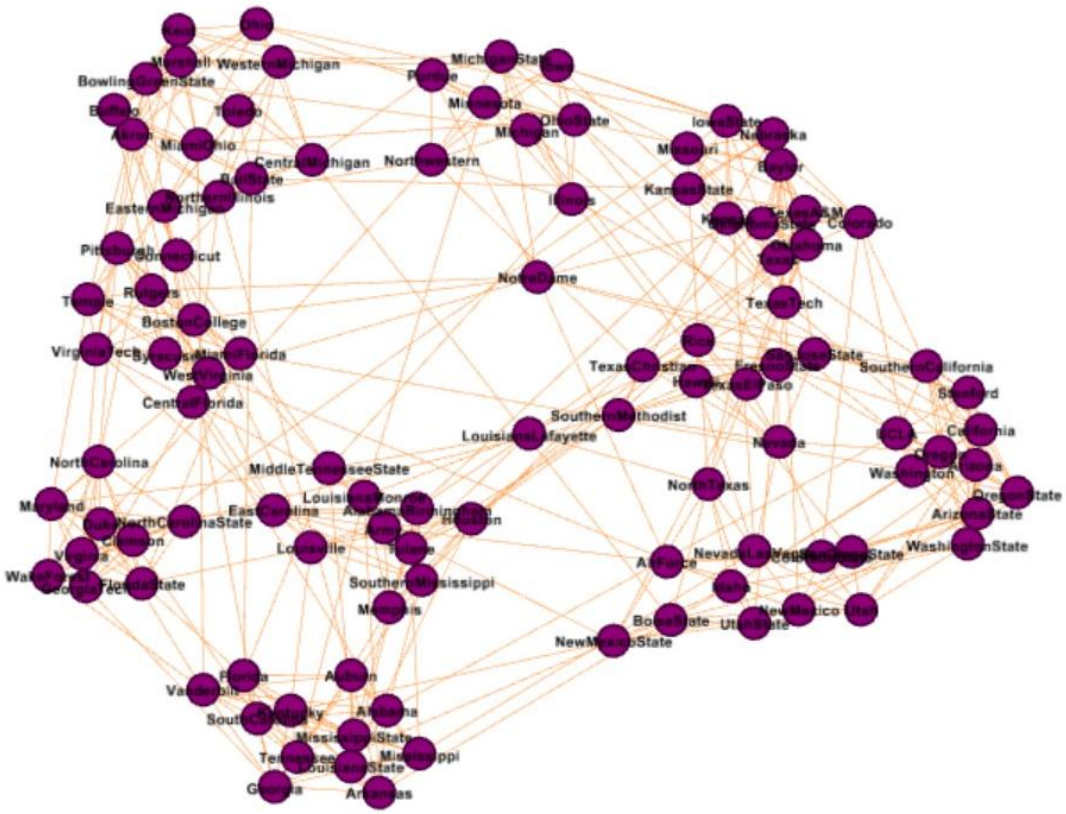
3) What if we remove teams with highest betweenness centrality?

- If we remove teams with highest betweenness centrality, then we will see more inter-regional matches than now. As there is not bridge between them, they will have to play it directly.



4) What if we remove teams with highest closeness centrality? -

If we remove teams with high closeness centrality, then we would simply not have teams which played with far away distanced teams.



Effect of Centralities on our network

Our network will still be stable even if removed all highest of the centralities because our network is about a football games where nodes are teams and edges are match played between teams. So, there just needs to be at least 2 teams in every region to make our network stable. These centralities are going to change and will depend on our data. So, it can be changed and that is not going to affect us in any way because there can be any number of teams in the season which can participate.