DESCRIPTION:

The network was generated using email data from a large European research institution. We have anonymized information about all incoming and outgoing email between members of the research institution. There is an edge (u, v) in the network if person u sent person v at least one email. The e-mails only represent communication between institution members (the core), and the dataset does not contain incoming messages from or outgoing messages to the rest of the world.

The dataset also contains "ground-truth" community memberships of the nodes. Everyone belongs to exactly one of 42 departments at the research institute. This network represents the "core" of the email-Eu-All network, which also contains links between members of the institution and people outside of the institution (although the node IDs are not the same).

OBJECTIVES:

Tasks to be performed:

1. Title of app should be \Full Name Social Network Analysis"
2. User should be able to select and load the two relevant files for this assignment using the app: email-Eu-core-department-labels.txt, email-Eu-core.txt.
3. Using any package or function, display any n connections from the file, “email-Eu-core.txt"; where n is some number input by the user.
4. Programmatically compute the number of emails sent by each person; display this information in a tabular format.
5. Programmatically compute the number of emails received by each person; display this information in a tabular format.
6. Display up to 2-hop neighbors of the top 10 from (4) and (5).
7. Assume that each email sent or received is a connection. Compute the degree centrality of each person. Display/visualize up to 2-hop neighbors of 10 people with the highest centrality. The degree centrality of a node(person) *i*, can be defined as the total number of nodes connected to node n_i. Also, color code nodes according to the department to which they belong.
8. Assume that each email sent or received is a connection. Compute the betweenness centrality of each person. Display/visualize up to 2-hop neighbors of 10 people with the highest betweenness.
9. Display/visualize 2-hop neighbors of nodes with the top 10 indegree centrality. Color code nodes according to the department.
10. Aggregate the emails sent per person, to the department level. After aggregation, you should have a new table that indicates the number of emails sent and received between each department. The table should have three columns. Column A, indicates the department from which emails are originating, Column B, indicates the department to which the emails are being sent, and Column C indicates the total number of emails sent from A to B. Display the table, and visualize the directed connections.
11. In a few sentences describe your observations when comparing the visualizations from 7, 8 and 9