Project 3

This project focuses on community detection and network analysis by using the **email-Eu-core network** data set. It involves the following steps:

1) Network Structure Analysis: Characterize the overall structure of the network G by calculating the following metrics:

   - $n(G)$.

   - $e(G)$.

   - Density of G

   - Connectivity: Determine if G is strongly connected. If not, calculate the fraction of vertices belonging to the largest strongly connected component.

   - Diameter and radius of G.

   - Clustering coefficient of G (using measures appropriate for directed graphs. The DirectedClustering package in R is required; see Chapter 4 notes for details).

2) Degree Distribution Analysis: Analyze the degree distribution of G. Generate and plot histograms for in-degree, out-degree, and total degree.

3) Triad Census Analysis: Perform a triad census analysis of G (see Chapter 4 notes) and compare the results with the communities you detected in Step 5 (for example, if nodes that are contained in the same triad with three edges (types 030T, 120U, etc.) are frequently contained in the same community).

4) Modularity/Assortativity Analysis: Calculate the modularity value of G based on enumerative characteristics (see Chapter 4 notes).

5) Community Detection: Apply at least two distinct community detection algorithms discussed in Chapter 5 to identify communities within G.

6) Evaluation: Compare the detected communities with the provided ground-truth communities. Quantify the accuracy of the community detection results using appropriate metrics.