Project 4

This project focuses on network analysis and blockmodeling by using the **LastFM Asia Social 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 connected. If not, calculate the fraction of vertices belonging to the largest component. If it is connected, find its connectivity.
- Diameter and radius of G.
- Clustering coefficient of G.

2) Degree Distribution Analysis: Analyze the degree distribution of G.

3) Calculate the cosine similarity matrix, Pearson correlation matrix, and Euclidean distance matrix for the network.

4) Apply two of the following methods to perform a blockmodeling analysis:

- Multidimensional Scaling (MDS)
- Clustering-based method
- Stochastic Block Model (SBM)

For the first two methods, you can directly use the Euclidean distance matrix as the dissimilarity matrix. Alternatively, you can transform the cosine similarity matrix or Pearson correlation matrix into a distance matrix. Ensure you follow these steps for each method:

- Obtain the membership IDs from the clustering process.
- Permute the adjacency matrix based on the membership IDs and plot the heatmap of the permuted matrix.
- Compute the image matrix.
- Plot the blockmodel.
- Evaluate the goodness of fit.

5) Interpret the blockmodel by analyzing the patterns of within-block and between-block connections.

6) Bonus (5 points): Try different approaches to classify the nodes as accurately as possible and compare your classification results with the labels provided in lastfm_asia_target.csv.