COURSEWORK: ANALYSIS OF A NETWORK

You are a junior data scientist who has been given the task to analyze a network of a specific BlueSky account.

Your role is to fetch and clean the data, import them to Gephi or Python or R or any other Network analysis tool and perform a network analysis approach to understand which users play a critical role in the network.

RULES:

- Analysis and visualization to be performed through use of the Gephi Software Tool (https://gephi.org) or Python or R with relevant Network Analysis packages (iGraph, NetworkX).

- Choose a network you wish to analyze, explain why you chose to analyze it and describe what you want to achieve through this analysis. Feel free to experiment with various datasets before you make your choice.

EXPECTATIONS:
- Provide a succinct analysis of the basic properties of the network. This should include:
    - Graphical representation of the network
    - Basic topological properties, such as numbers of nodes and edges, network diameter, and average path length.
    - Component measures, such as number of connected components, existence of a giant component and component size distribution.
    - Degree measures, such as maximum and average node degrees, as well as degree distribution.
    - Centrality measures (degree, betweenness, closeness, eigenvector)
    - Clustering effects in the network: average clustering coefficient, number of triangles, clustering coefficient distribution, existence of the triadic closure phenomenon in the friendship neighborhood.
    - Bridges and local bridges.
    - Gender and homophily.
    - Graph density
    - Community structure (modularity)
    - PageRank

Hint: It might be a good idea to produce ranked or partitioned versions of the network graph based on the criteria by which you analyze it (use the Partition and Ranking tabs in the top left window in Gephi).

Good Luck!!!

Evaluation – Assessment Methods

- **Theoretical Foundation:** 40%

  Explain the theoretical understanding of the objectives covered throughout the course

- **Empirical Validation:** 50%

  Demonstrate the ability to use a software tool to analyze and visualize a complex problem, through network analysis

- **Assignment Layout:** 10%

  A professional or academic layout should be taken care of