**Project Written Submission**

The project leverages a data set containing global education statistics from the United Nations (UN) for the year 2020. The main goal is to analyze the data using graph clustering and partitioning methods to gain insights into relationships between different countries or regions based on their education statistics. The project involves creating a network where each node represents a country or region, and the connections (edges) between the nodes are weighted based on similarity in education metrics. The aim is to identify clusters within the network and evaluate how well the chosen clusters represent different aspects of the global education landscape.

The code provided achieves this through the following key steps:

1. Data Loading and Preprocessing:
   * The program begins by loading data from a CSV file containing the UN global education statistics for 2020.
   * It processes the data to handle missing values and prepares it for graph construction.
2. Graph Construction:
   * Once the data is loaded, the code constructs an undirected graph (petgraph::Graph) where each node represents a country or region, and the edges between nodes are weighted based on some metric (e.g., similarity in education statistics).
   * A mapping between node indices and country names is created to facilitate the interpretation of clusters later.
3. Graph Clustering and Partitioning:
   * The program uses the Louvain method for community detection to cluster the graph nodes.
   * It produces clusters of nodes that can represent groups of countries or regions with similar education statistics.
4. k-Means Clustering:
   * In addition to graph clustering, the program uses the smartcore library to perform k-means clustering on the graph's adjacency matrix.
   * This additional clustering method provides another layer of analysis on the data.
5. Counting Nodes and Edges:
   * The program includes a function to count the number of nodes and edges in the graph, providing insights into the size and complexity of the network.
6. Testing:
   * Tests are written to verify the functionality of each key function in the program.
   * These tests validate data loading, graph construction, clustering, and other functions to ensure the code behaves as expected.

### **Conclusion:**

This project aims to provide a better understanding of global education statistics through graph clustering and partitioning. By finding meaningful clusters and representatives, you can gain insights into educational trends and potential areas for policy improvement.