**Analysis of Economic Indicators Using Clustering and Trend Analysis**

Prepared by**:** Shirisha Jakkula

Student Id: 2309638

https://github.com/jakshirisha/Clustering-and-Fitting

**Introduction**: This report focuses on analysing economic indicators to understand patterns among countries and GDP growth trends over time. The methods used include K-Means clustering and linear regression trend analysis. The results are visualized and interpreted to derive insights into economic behaviours.

The analysis begins by loading a dataset of economic indicators. The dataset includes variables such as GDP (current USD), GDP growth (annual %), and unemployment rates (% of total labour force). The dataset was read into Python using the following code:

**1. Data Cleaning:** Cleaning the data is essential to ensure accurate results. Key indicators such as GDP, GDP growth, and unemployment were retained, and missing values were dropped.

**2. K-Means Clustering**

A group of graphs showing different colored dots

Description automatically generatedK-Means clustering was applied to group countries based on GDP (current USD), GDP growth (annual %), and unemployment rate (%). The steps included normalizing the data and running the K-Means algorithm.

**3. Visualization:**

Clustering Results:

**Cluster 0 (Purple):** - Composed of countries with relatively low GDP values but moderate to low unemployment rates. These could represent developing economies with stable labour markets but limited overall economic size.

**Cluster 1 (Teal):** - Countries with moderate GDP growth and unemployment rates. Likely includes emerging economies experiencing steady growth but with structural challenges in employment.

- Example countries: Emerging markets in Southeast Asia or Latin America.

**Cluster 2 (Yellow):**- Represents countries with high unemployment rates and moderate GDP growth. May include nations facing economic instability or recovering from crises.

-These could be countries in regions like Sub-Saharan Africa or post-conflict zones.

**Interpretation:** The clustering highlights distinct economic patterns and provides insights into how countries group together based on their economic conditions. These clusters could guide policymakers in understanding their nation's economic standing relative to others.

**4: Linear Regression for GDP Growth Trend :** Linear regression was used to analyse GDP growth trends for individual countries. The example focused on India.

A graph with blue dots

Description automatically generated**Visualization:**

The graph displayed India’s GDP growth rates as blue dots and a red regression line indicating the trend. The trend line was relatively flat, suggesting consistent GDP growth over time.

**5. Final Insights**

**Key Observations:**

1. **Clustering Analysis**: -

- Cluster 0: Developing countries with low GDP but moderate stability in employment.

- Cluster 1: Emerging economies balancing moderate GDP growth and unemployment.

- Cluster 2: Economies with structural challenges reflected in high unemployment rates despite moderate GDP growth.

**2. India's GDP Growth Trend:** -

- The fitted line shows stable GDP growth over time, averaging around 6%.

- Despite external shocks (e.g., COVID-19), the economy demonstrates resilience.

**Real-World Implications of Clustering:** -

- Cluster 0 (Purple): Developing nations need targeted investment in infrastructure and economic diversification.

- Cluster 1 (Teal): Emerging economies should focus on improving employment opportunities while sustaining growth.

- Cluster 2 (Yellow): Countries in this cluster may benefit from addressing structural issues such as labour market inefficiencies and political stability to reduce unemployment and boost growth.

**Conclusion: -** This analysis provides insights into global economic patterns and individual country trends. Clustering helps identify groups of similar economies, while linear regression reveals historical trends for specific countries. These methods combined can guide policymakers and economists in understanding and comparing economic performance.