

Introduction to Data Science

HW1

2023 June Unemployment Rate.

Unemployment rate by country

About the dataset:

When someone who is able and ready to work cannot find adequate job, they are said to be unemployed. It is an important economic metric that measures the strength of a labor market as well as the economy's overall health. The complex issue of unemployment, which is influenced by several circumstances, can have important social and economic repercussions.

Features in the dataset:

Series ID of Country:

This feature most likely serves as a special code or identify for each nation. It can be utilized to refer to and connect every nation to the appropriate region.

Region name of country:

The name or label of the region to which each country belongs is provided by this feature. It is necessary for understanding the geographical classification of countries.

Region code of country:

This feature is most likely a numerical or alphanumeric code connected with each location. It can be used to handle and categorize data efficiently.

Possible problems to investigate:

Geographical Analysis: Use the region data to examine how countries are distributed throughout different regions. You can calculate and view the number of countries in each region.

Mapping Countries to Regions: Using this dataset, create a mapping or lookup table that associates each country with its associated region. This is especially handy when working with other datasets that contain references to areas.

Data Validation: Ensure data consistency and accuracy by cross-referencing region names and codes with other sources or databases. Determine any discrepancies or missing information.

Regional Comparisons: Determine whether certain regions share common characteristics or trends in economic, social, or political aspects. This can include comparing regions in terms of development, governance, or other pertinent characteristics.

Cluster Analysis: Using the region data, use cluster analysis to identify groups of countries that share similar features. This can be valuable for comparative research or policy analyses.

Data Integration: Combine this region information with other datasets containing country-level data to improve your study and provide a more comprehensive knowledge of the regions' impact on numerous elements.