

Atividades práticas de apoio – Semana 5: Análise de agrupamento

Use o algoritmo **complete-linkage** para analisar um conjunto de dados socioeconômicos de países. Para fazer a análise, selecione os atributos **child_mort** e **life_expec** usando 4 grupos, e, como a medida de similaridade, use a distância euclidiana. Gere o dendograma e o gráfico de dispersão dos dados.

Responda também: quais são os países que têm as piores taxas de mortalidade e expectativa de vida? Os que têm as melhores taxas? Sugestão: filtre o dataframe usando os valores apontados no gráfico de dispersão ou inclua os nomes dos países no gráfico.

Descrição dos dados:

- **country**: nome do país
- **child_mort**: mortalidade infantil de crianças até 5 anos a cada mil nascidos
- **exports**: percentual de participação das exportações no PIB
- **health**: percentual de gastos com saúde em relação ao PIB
- **imports**: percentual de participação das importações no PIB
- **income**: renda anual per capita
- **inflation**: taxa de crescimento anual do PIB
- **life_expec**: expectativa de vida de acordo com a taxa de mortalidade
- **total_fer**: número de crianças que poderiam nascer por mulher de acordo com as taxas de fertilidade/idade.
- **gdpp**: PIB per capita

Os dados originais foram retirados do Kaggle (Country Socioeconomic Data: <https://www.kaggle.com/datasets/ashydv/country-socioeconomic-data>) e estão disponíveis neste link: <https://raw.githubusercontent.com/higoramario/univesp-com360-mineracao-dados/main/country-data.csv>.