# Introdução e Motivação

## Introdução

A avaliação de dados educacionais é crucial para a criação de políticas públicas, o planejamento de investimentos e a promoção de uma sociedade mais justa. Em um cenário europeu, onde as desigualdades entre regiões e grupos sociais persistem, entender os níveis de educação pode fornecer dados valiosos para tomadas de decisões estratégicas.

## Motivação para o Projeto

A motivação para este projeto vem da necessidade de investigar as diferenças no nível educacional entre países, idades e gêneros na Europa. Com essa avaliação, busca-se reconhecer padrões, agrupamentos e potenciais disparidades que possam orientar futuras pesquisas ou iniciativas de aprimoramento por parte de instituições de ensino e governos.

## Descrição do Problema

O conjunto de dados empregado inclui dados sobre o grau de instrução da população europeia, divididos por idade, gênero e localização geográfica. Education attainment is given by International Standard Classification of Education (ISCED11).

ISCED11 education levels are the following:

**X** - No schooling

**0** - Early childhood education

**1** - Primary education

**2** - Lower secondary education

**3** - Upper secondary education

**4** - Post-secondary non-tertiary education

**5** - Short-cycle tertiary education

**6** - Bachelor’s or equivalent level

**7** - Master’s or equivalent level

**8** - Doctoral or equivalent level

**9** - Not elsewhere classified

Devido à vasta quantidade e variedade dos dados, é um desafio identificar padrões pertinentes sem o uso de métodos adequados de análise e agrupamento. Portanto, o projeto visa utilizar algoritmos de aprendizado de máquina (especialmente de clusterização) para identificar agrupamentos naturais nos dados que possuam características semelhantes.

## Objetivos do Projeto

Os objetivos centrais deste projeto são:

* Executar uma análise exploratório na base de dados de educação europeia;
* Utilizar métodos de agrupamento para agrupar observações com atributos parecidos;
* Analisar o rendimento de variados algoritmos de agrupamento;
* Analisar os agrupamentos identificados para obter percepções pertinentes sobre o perfil educacional dos europeus.