Datathon

2025

*Khrystyna Budzak*

*Rodrigo Boto*

*Tiago*

*Leandro*

**Introdução**

No presente projeto tentamos identificar padrões de consumo sustentável e identificar estratégias baseadas nos dados para melhorar a sustentabilidade sem comprometer o lucro.

Objetivos:

* Vendas: Análise das tendências de consumo sustentável
* Impacto ambiental: análise do impacto ambiental dos produtos
* Segmentação de clientes: identificação de segmentos de consumidores com base nas suas preferências e comportamentos de compra:
  + Correlacionar a frequência de compra com país
  + Correlacionar idade com compra de produtos mais ou menos sustentáveis
* Feedback de clientes: análise das perceções e feedback em relação aos produtos sustentáveis
  + Dos produtos com índice mais sustentável qual foi o feedback e quais foram as reclamações mais frequentes
* Estratégias: desenvolvimento de estratégias para aumentar a venda de produtos sustentáveis

**Análise e Tratamento de Dados**

O nosso dataset tem 10674 dados de vendas e 3000 dados de clientes.

Para definir as categorias e produtos mais e menos sustentáveis criamos uma tabela nova com normalização das variáveis para uma escala de 0 e 1 e calculamos um índice ponderado da seguinte forma:

Índice de sustentabilidade = (×

– 20% peso atribuído ao consumo de água

– 30% peso atribuído ao consumo de energia

– 50% peso atribuído a emissão de CO2

Classificação utilizada:

* > 0.75 sustentável
* 0.5 -0.75 médio
* < 0.50 baixa sustentabilidade

Na sheet clientes verificamos 10 ID de clientes duplicados e eliminamos (ao todo 20 linhas eliminadas).

**Resultados**

Com base neste índice de sustentabilidade verificamos o top 3 de produtos mais e menos sustentáveis por cada categoria: