**Documentação Técnica – Tratamento de Dados**

1. Introdução

Este documento descreve as etapas de tratamento de dados realizadas no notebook 'dados\_nao\_tratados.ipynb'. O objetivo principal é preparar a base de dados bruta para análises, garantindo consistência, padronização e melhor qualidade dos dados. O notebook foi desenvolvido em Python, utilizando a biblioteca Pandas para a manipulação de dados. A base de dados e o carregamento do arquivo ‘dados\_nao\_tratados.csv’.

1. Etapas de treinamento
   1. Importação de Bibliotecas

Foi importada a biblioteca Pandas para a manipulação e transformar dados.



* 1. Carregamento do Dataset

O dataset foi carregado a parti do arquivo ( .CSV )

Texto

O conteúdo gerado por IA pode estar incorreto.

* 1. Tratamento da coluna “ativos”

A coluna continha valores inconsistentes, "e", "yes", "sim", "true", "false", "não", "nao". Esses valores foram transformados para valores booleanos (True/False).

Texto

O conteúdo gerado por IA pode estar incorreto.

* 1. Convensão da coluna “idade”

A coluna ‘idade’ foi convertida para valores numéricos, eliminando os possíveis inconsistências ou valores invalídos.



* 1. Conversão da Coluna “data\_incritos”

A coluna ‘data\_incritos’ foi convertida para o formato datetime, considerando o padrão de datas brasileiro dia/mês/ano.



1. Conclusão

Após esses tratamentos, o dataset está padronizado, pronto para ser utilizados em analise estatísticas. As principais tratamentos desse dataset foram padronização de variáveis categóricas, conversão de tipos de dados e correção de inconsistências