

# 1. Introdução

A HealthCare Solutions está implementando uma iniciativa de transformação digital para melhorar a qualidade do atendimento ao paciente. O projeto visa analisar dados provenientes de EHRs, dispositivos de monitoramento, pesquisas de satisfação e sistemas administrativos para **gerar insights que reduzam readmissões e aumentem a satisfação dos pacientes**.

---

## 2. Metodologia

### 2.1 Coleta de Dados

- Dataset simulado com 500 pacientes
  - Variáveis: `patient_id`, `age`, `gender`, `length_of_stay`, `satisfaction_score`, `num_procedures`, `chronic_condition`, `readmitted`
  - Dados salvos em CSV: `healthcare_data_500.csv`
- 

### 2.2 Limpeza e Pré-processamento

- Conversão de `gender` para numérico
  - Conferência de duplicados e valores ausentes
  - Preparação para modelagem
- 

### 2.3 Análise Exploratória

- Histogramas: Idade, Tempo de Internação
- Heatmap de correlação
- Scatter plot: Satisfação x Tempo de Internação

## 2.4 Modelagem Preditiva

- **Modelo:** Árvore de Decisão
- **Variáveis:** age, gender, length\_of\_stay, satisfaction\_score, num\_procedures, chronic\_condition
- **Acurácia:** 55%
- **Matriz de Confusão:**

### Insights:

- Condições crônicas aumentam risco de readmissão
  - Internações longas aumentam risco
  - Pacientes mais satisfeitos tendem a não ser readmitidos
- 

## 3. Conclusão

- É possível identificar pacientes com maior risco de readmissão
- Visualizações e métricas auxiliam na tomada de decisões
- Próximos passos: integrar dados reais, testar outros modelos e criar dashboards interativos

## 4. Referências

- LGPD: [http://www.planalto.gov.br/ccivil\\_03/leis/L13709.htm](http://www.planalto.gov.br/ccivil_03/leis/L13709.htm)
- SMITH, João. *Melhores Práticas em Ciência de Dados na Saúde*. Revista de Ciência de Dados em Saúde, 2023
- Kaggle / GitHub – dados simulados