

# **Data Mining e Graph Mining**

## **Inteligência Artificial**

Professor André Luiz Esperidião

# LABORATÓRIO

Nesta apresentação, exploraremos a aplicação da metodologia CRISP-DM para identificar e mitigar a perda de clientes em uma empresa de telecomunicações. Nosso objetivo é transformar dados brutos em insights açãoáveis.



# Nossa abordagem: CRISP-DM em 6 Fases

## METODOLOGIA CRISP-DM

Seguiremos a estrutura do CRISP-DM, analisando cada fase sob três perspectivas cruciais: o que o negócio almeja, o papel do cientista de dados e o resultado concreto esperado.

1

### **Business Understanding**

Definição clara dos objetivos de negócio.

2

### **Data Understanding**

Exploração e compreensão dos dados disponíveis.

3

### **Data Preparation**

Limpeza e transformação dos dados para modelagem.

4

### **Modeling**

Construção e seleção de modelos preditivos.

5

### **Evaluation**

Avaliação do desempenho dos modelos.

6

### **Deployment**

Implementação da solução e acompanhamento.

# CRISP-DM

FASE 02

## Data Understanding: Desvendando Nossos Dados

A fase de compreensão dos dados nos permite uma visão 360° do cliente, integrando perfil, contrato, uso do serviço, experiência e comportamento financeiro.



### Histórico de Uso

Serviços utilizados e padrões de consumo.



### Reclamações e Atendimento

Registros de interações com o suporte técnico.



### Perfil Demográfico

Gênero, idade (idoso), parceiro, dependentes.



### Dados de Faturamento

Valores mensais e totais, métodos de pagamento.

# CRISP-DM

## Estrutura do Dataset: Visão Geral

Nosso dataset possui **6000 linhas e 27 colunas**, oferecendo uma riqueza de informações para análise.

Identificação	customer_id	Identificador único do cliente
Demográfico	gender, senior_citizen, partner, dependents	Informações de perfil do cliente
Relacionamento	tenure_months	Tempo de relacionamento em meses
Serviços	phone_service, internet_service, etc.	Detalhes dos serviços contratados
Contrato/Cobrança	contract, monthly_charges, payment_method	Condições de contrato e faturamento
Qualidade	complaints_last_3m, outages_last_3m	Registros de problemas recentes
Financeiro	late_payments_last_6m	Histórico de pagamentos em atraso
Variável Alvo	churn	Indica se o cliente cancelou (1) ou não (0)

# CRISP-DM

## Análise Descritiva dos Dados

- **Count:** Quantidade de valores não nulos.
- **Mean:** Média aritmética, sensível a outliers.
- **Std:** Desvio padrão, mede a dispersão dos dados.
- **Min/Max:** Valores mínimo e máximo, identificando limites.
- **Quartis (25%, 50%, 75%):** Indicadores de distribuição central.

# CRISP-DM

## Análise da Taxa de Churn (Base Rate)

A proporção de clientes que cancelaram o serviço é um indicador crítico da saúde do negócio.

**Clientes Ativos**

**85.95%**

Dos 6000 clientes, **5157** permanecem utilizando os serviços.

**Clientes em Churn**

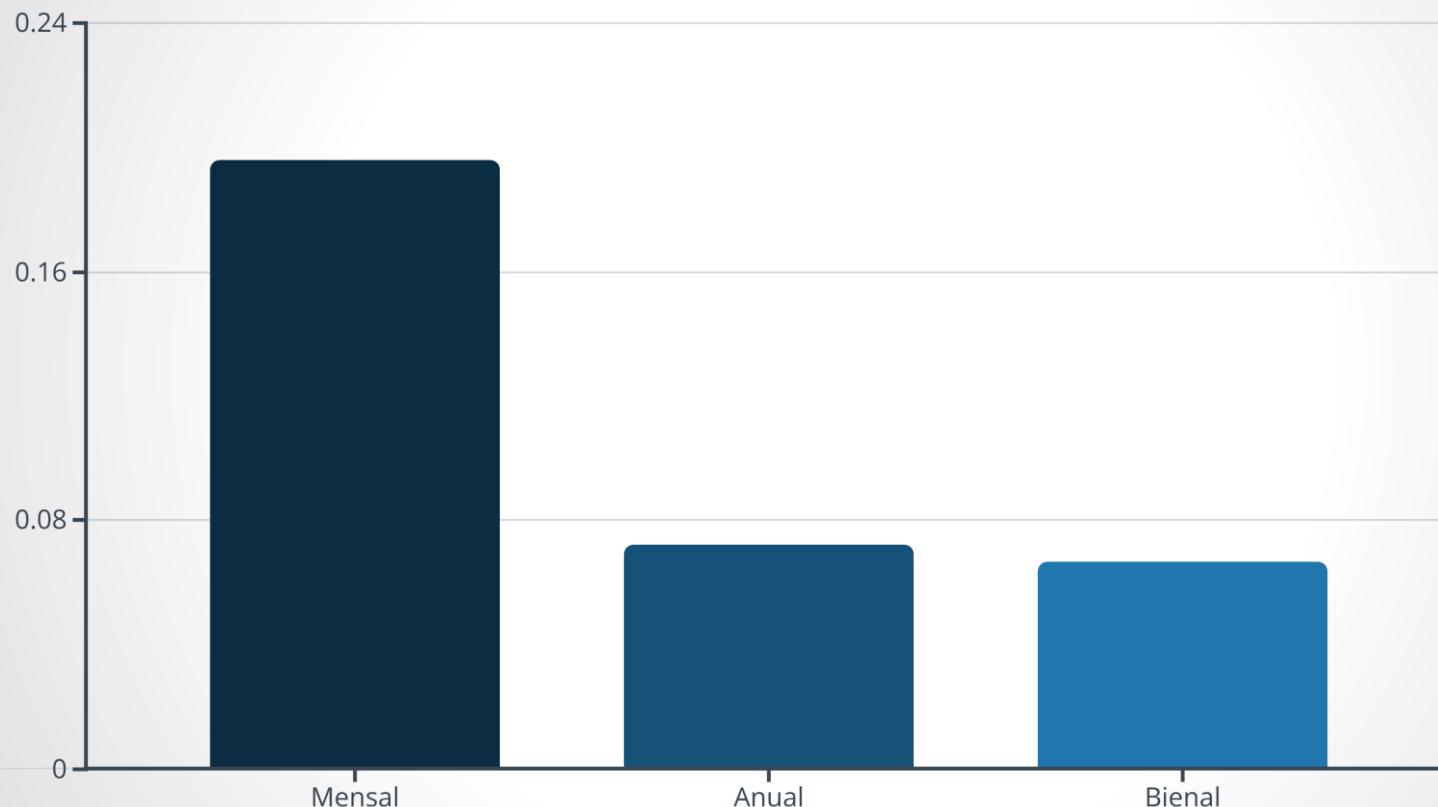
**14.05%**

Uma parcela de **843** clientes optou por cancelar seus serviços, representando nossa variável alvo.

**Observação didática:** Este é um cenário real de desbalanceamento de classes, comum em problemas de churn, e que exige atenção especial na modelagem.

# CRISP-DM

## Churn por Tipo de Contrato: Um Fator Chave



O contrato **mensal** (Month-to-month) apresenta a maior taxa de churn, quase **três vezes maior** que os contratos anuais e bienais.

## Insight de Negócio: Foco nos Contratos Mensais

Os dados revelam uma oportunidade clara para a empresa reduzir o churn através de estratégias de retenção.

### Alto Risco no Mensal

Contratos mensais são o principal fator de risco de churn, exigindo atenção prioritária.

### Migração Estratégica

Migrar clientes de contratos mensais para anuais pode reduzir o churn em mais de 60%.

### Incentivos Direcionados

Ofertas personalizadas para clientes mensais com sinais de risco podem impulsionar a conversão para contratos de maior prazo.

### Ganho Marginal do Bienal

O contrato bienal oferece uma redução marginal em relação ao anual, questionando a necessidade de incentivos agressivos.

## Conexão com Data Mining: A Força da Feature 'Contract'

A variável 'contract' é uma "feature" (característica) **fortemente preditiva** e uma excelente candidata para diversas aplicações no modelo.



Tratá-la como uma **feature ordinal** (Month-to-month < One year < Two year) permitirá que o modelo capture a hierarquia e o risco associado a cada tipo de contrato, otimizando a precisão das previsões de churn.

# CRISP-DM

Próximo passo natural:

- Cruzar contrato × tenure
- Ver contrato × reclamações



# **Data Mining e Graph Mining**

## **Inteligência Artificial**

Professor André Luiz Esperidião