

Relatório Final: Análise de Evasão de Clientes (Churn) - Telecom X

1. Introdução

A Telecom X enfrenta um desafio significativo relacionado à alta taxa de evasão de clientes (Churn). A perda contínua de clientes impacta negativamente a receita e o crescimento da empresa. Para combater este problema, foi iniciado o projeto "Churn de Clientes", com o objetivo de compreender os fatores que levam os clientes a cancelar seus serviços.

Este relatório documenta a análise de dados exploratória realizada sobre os dados de clientes da Telecom X. O objetivo principal desta análise foi identificar padrões e características associadas aos clientes que cancelaram seus serviços, fornecendo insights valiosos que possam subsidiar a criação de modelos preditivos e o desenvolvimento de estratégias eficazes para a retenção de clientes.

A análise abrange desde a coleta e tratamento dos dados brutos até a exploração visual e estatística das variáveis mais relevantes, culminando em conclusões e recomendações práticas para a empresa.

2. Coleta, Limpeza e Tratamento de Dados

A análise partiu de dados brutos em formato JSON, disponibilizados via API (link do GitHub). As seguintes etapas foram cruciais para preparar os dados para a análise exploratória:

- Coleta e Carregamento:** Os dados foram carregados diretamente da fonte JSON utilizando Python e a biblioteca Pandas.
- Normalização:** A estrutura JSON continha dados aninhados (informações de cliente, telefone, internet e conta). Estes foram

normalizados e expandidos em colunas individuais no DataFrame principal para facilitar a manipulação.

3. Verificação de Qualidade:

- **Tipos de Dados:** A coluna `Charges.Total` (`Cobranca_Total`) foi identificada como tipo `object` devido a valores não numéricos (espaços em branco) e convertida para tipo numérico (`float64`).
- **Valores Ausentes:**
 - Foram encontrados 224 valores ausentes na coluna `Churn`. Como esta é a variável alvo, as linhas correspondentes foram removidas do dataset.
 - Foram encontrados 11 valores ausentes na coluna `Charges.Total`. Verificou-se que todos estes casos correspondiam a clientes com `tenure` (`Meses_Contrato`) igual a 0. Estes valores foram imputados com 0, assumindo que clientes novos ainda não possuem cobrança total registrada.
- **Duplicados:** Nenhuma linha duplicada foi encontrada.
- **Inconsistências:** Não foram identificadas inconsistências significativas nos valores das demais colunas categóricas ou numéricas.

4. **Engenharia de Features:** Foi criada a coluna `Contas_Diarias` (`Cobranca_Diaria`), calculada dividindo a `Cobranca_Mensal` por 30, para oferecer uma perspectiva adicional sobre o custo percebido pelo cliente.

5. Padronização e Transformação:

- **Conversão Binária:** Colunas com respostas "Yes"/"No" (`Churn`, `Partner`, `Dependents`, `PhoneService`, `PaperlessBilling`) foram convertidas para formato numérico (1 para "Yes", 0 para "No").
- **Renomeação:** Todas as colunas foram renomeadas para português (ex: `tenure` -> `Meses_Contrato`, `Charges.Monthly` -> `Cobranca_Mensal`) para maior clareza e acessibilidade.

Após estas etapas, o dataset final (`TelecomX_Data_Standardized.csv`), contendo 7043 linhas e 22 colunas, estava pronto e limpo para a análise exploratória.

3. Análise Exploratória de Dados (EDA)

Com os dados limpos e padronizados, realizamos uma análise exploratória para entender a distribuição das variáveis e identificar relações com a variável alvo, Churn.

3.1 Distribuição Geral do Churn

Primeiramente, analisamos a proporção geral de clientes que cancelaram versus os que permaneceram. Observou-se um desbalanceamento, com a maioria dos clientes (aproximadamente 73.5%) permanecendo na base, enquanto 26.5% cancelaram.

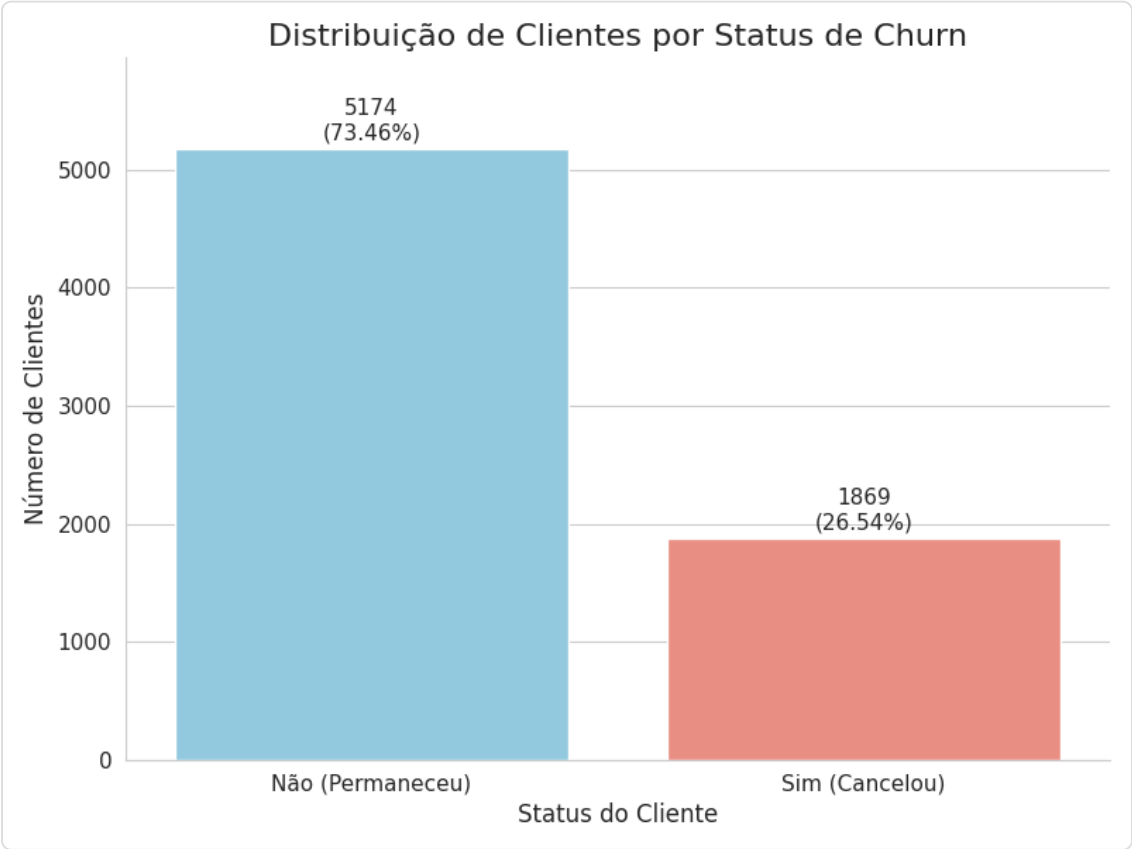


Figura 1: Proporção de clientes que permaneceram (Não) e cancelaram (Sim).

3.2 Análise de Churn por Variáveis Categóricas

Investigamos como a taxa de churn varia entre diferentes categorias de clientes. Alguns padrões notáveis emergiram:

- **Tipo de Contrato:** A taxa de churn é drasticamente maior para contratos mensais (>42%) em comparação com contratos de longo prazo (Anual: ~11%, Dois anos: <3%).

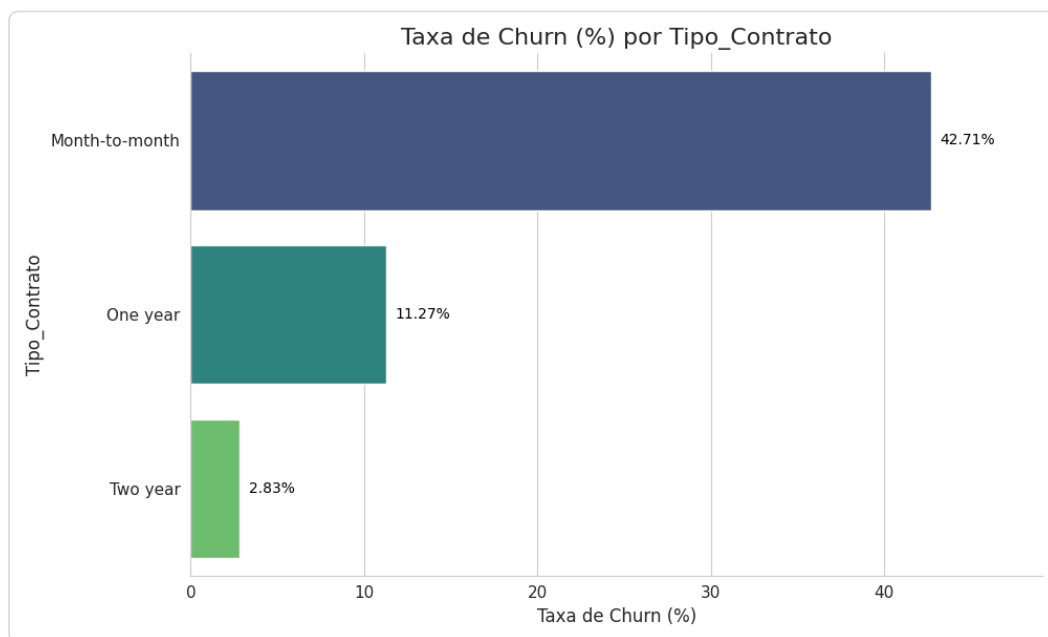


Figura 2: Churn por Tipo de Contrato.

- **Serviço de Internet:** Clientes com Fibra Óptica apresentam churn significativamente maior (>41%) que clientes com DSL (<20%). Clientes sem serviço de internet têm churn muito baixo (<8%).

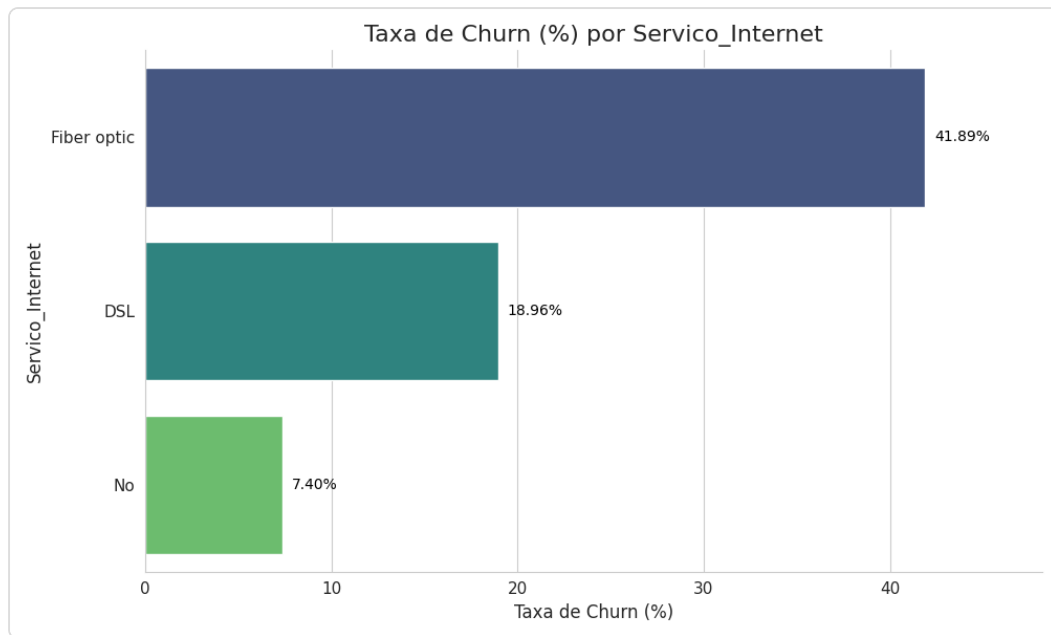


Figura 3: Churn por Serviço de Internet.

- **Método de Pagamento:** O pagamento via Débito Eletrônico (Electronic check) está associado à maior taxa de churn (>45%).

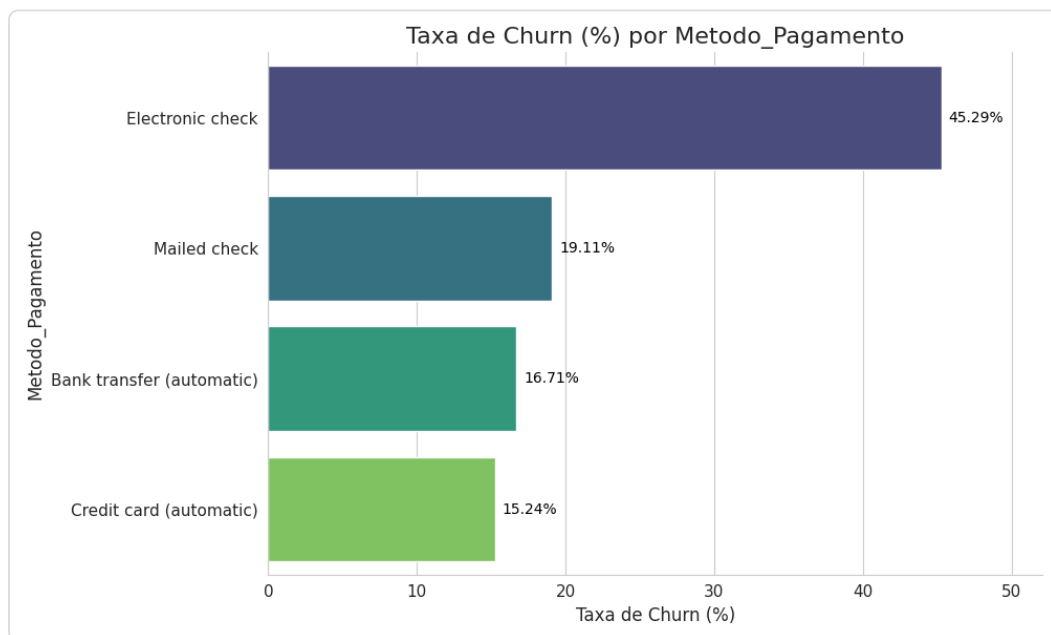


Figura 4: Churn por Método de Pagamento.

- **Serviços Adicionais:** A ausência de serviços como Segurança Online e Suporte Técnico está fortemente correlacionada com taxas de churn mais altas (>41%).

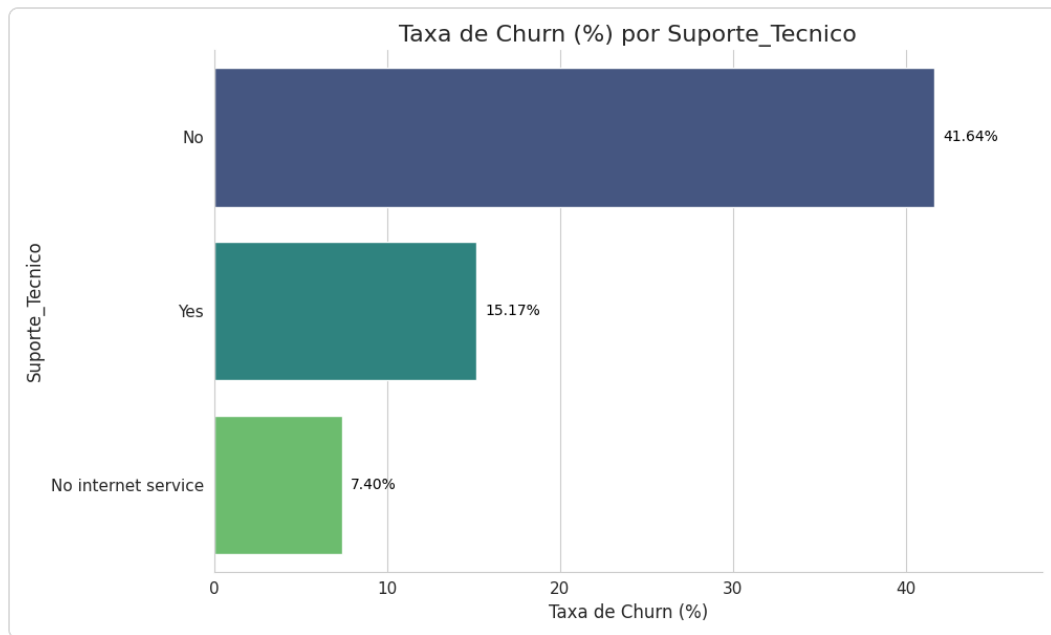


Figura 5: Churn por Suporte Técnico.

- **Perfil Demográfico:** Clientes idosos e aqueles sem parceiro(a) ou dependentes também mostraram taxas de churn mais elevadas.

(Para uma análise detalhada de todas as categorias, consulte o arquivo `visual_analysis_summary.md` gerado anteriormente).

3.3 Análise de Churn por Variáveis Numéricas

Comparamos a distribuição de variáveis numéricas chave entre os grupos que cancelaram e os que permaneceram:

- **Meses de Contrato:** Clientes que cancelam têm um tempo de contrato significativamente menor (mediana de 10 meses vs. 38 meses).

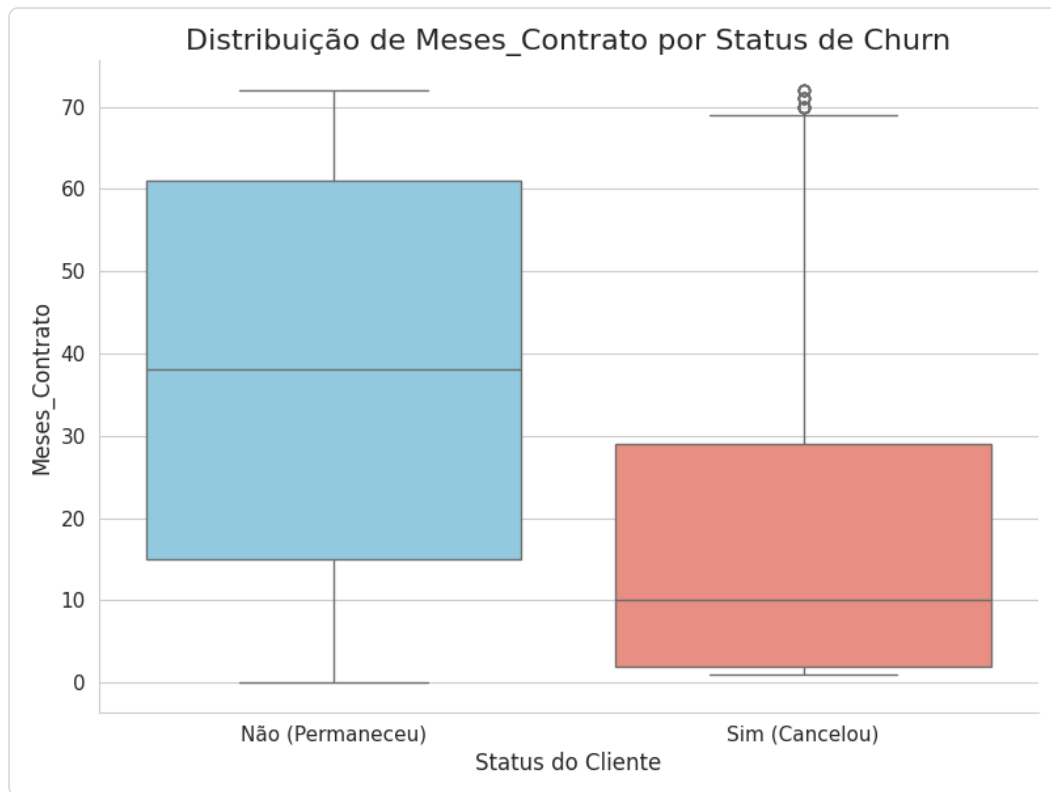


Figura 6: Boxplot de Meses de Contrato por Status de Churn.

- **Cobrança Mensal:** Clientes que cancelam tendem a ter cobranças mensais mais altas (mediana de R\$ 79.65 vs. R\$ 64.43).

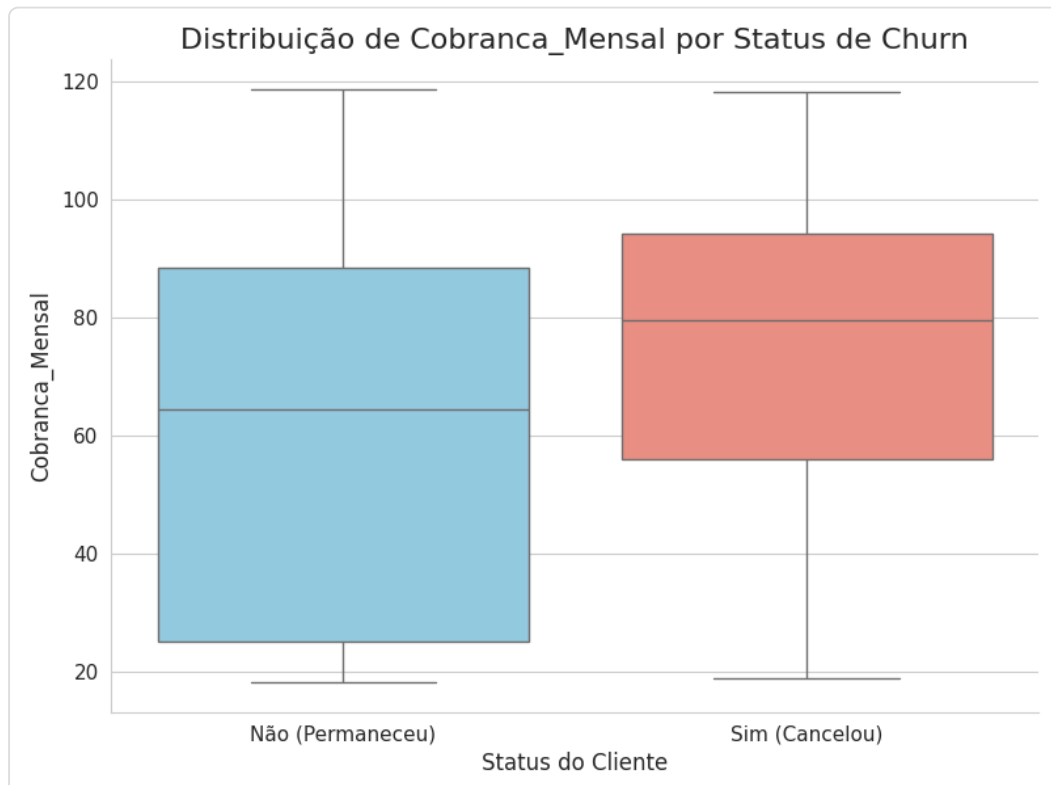


Figura 7: Boxplot de Cobrança Mensal por Status de Churn.

(Para estatísticas descritivas detalhadas por grupo, consulte o arquivo `numerical_desc_by_churn.md` gerado anteriormente).

4. Conclusões e Insights Principais

A análise exploratória dos dados da Telecom X revelou insights importantes sobre os fatores associados à evasão de clientes:

- **Perfil de Risco:** Existe um perfil de cliente com maior probabilidade de churn, caracterizado por:
 - **Contrato Mensal:** A falta de um compromisso de longo prazo é o indicador mais forte de churn.
 - **Clientes Recentes:** O churn é muito mais comum nos primeiros meses de contrato.
 - **Cobranças Mensais Elevadas:** Clientes que pagam mais mensalmente tendem a cancelar mais.
 - **Serviço de Fibra Óptica:** Embora seja um serviço premium, está associado a um churn maior, talvez por custo, instabilidade percebida ou expectativas não atendidas.
 - **Método de Pagamento Específico:** O uso de Débito Eletrônico (Electronic check) se destaca com alta taxa de churn.
 - **Ausência de Serviços de Valor Agregado:** Clientes sem serviços como Segurança Online e Suporte Técnico são mais propensos a sair.
 - **Demografia:** Clientes idosos e aqueles sem parceiros ou dependentes também apresentam maior risco.
- **Fatores de Retenção:** Por outro lado, fatores associados à permanência do cliente incluem:
 - **Contratos de Longo Prazo (1 ou 2 anos):** Clientes com esses contratos são significativamente mais leais.
 - **Maior Tempo de Contrato (Tenure):** Clientes mais antigos tendem a permanecer.

- **Serviços Adicionais:** A adesão a serviços como Suporte Técnico e Segurança Online parece aumentar a retenção.
- **Clientes sem Serviço de Internet:** Este grupo tem uma taxa de churn muito baixa, indicando que o churn está mais concentrado entre usuários de internet.
- **Importância do Valor Percebido:** A combinação de cobranças mensais altas e a ausência de serviços de suporte/segurança sugere que o valor percebido pelo cliente pode ser um fator crucial. Clientes com fibra óptica (geralmente mais cara) e sem suporte/segurança podem sentir que não estão recebendo um bom custo-benefício.

Esses insights fornecem uma base sólida para entender *quem* está saindo e *por quê*, permitindo que a Telecom X desenvolva ações mais direcionadas para reduzir a evasão.

5. Recomendações

Com base nos insights obtidos na análise exploratória, apresentamos as seguintes recomendações para a Telecom X, visando reduzir a taxa de evasão de clientes:

1. Incentivar Contratos de Longo Prazo:

- Desenvolver campanhas ativas para migrar clientes de contratos mensais para planos anuais ou bienais, oferecendo descontos progressivos, bônus de serviço ou outros benefícios claros.
- Implementar programas de fidelidade que recompensem clientes com maior tempo de contrato (tenure).

2. Revisar Estratégia de Preços e Valor (Especialmente Fibra Óptica):

- Analisar a estrutura de preços da Fibra Óptica, comparando com o mercado e avaliando a percepção de valor dos clientes, especialmente aqueles com churn elevado.
- Considerar o *bundling* (pacotes) de serviços de valor agregado, como Segurança Online e Suporte Técnico, especialmente para planos de Fibra Óptica, para aumentar o valor percebido e a dependência do serviço.

- Oferecer promoções ou descontos direcionados para clientes com alto custo mensal e baixo tempo de contrato.

3. Otimizar Métodos de Pagamento:

- Investigar as razões por trás da alta taxa de churn associada ao Débito Eletrônico (Electronic check). Pode haver problemas no processo, falhas, ou ser um indicativo de menor estabilidade financeira do cliente.
- Incentivar a adoção de métodos de pagamento com menor churn, como débito automático em conta bancária ou cartão de crédito, talvez oferecendo pequenos descontos ou benefícios.

4. Fortalecer o Suporte e Serviços Adicionais:

- Promover ativamente os serviços de Segurança Online e Suporte Técnico, destacando seus benefícios, especialmente para clientes de Fibra Óptica e Idosos.
- Considerar incluir um nível básico desses serviços em planos mais caros ou oferecer um período de teste gratuito para demonstrar valor.

5. Melhorar a Experiência do Cliente Recente (Baixo Tenure):

- Reforçar o processo de *onboarding* para novos clientes, garantindo que compreendam os serviços, saibam como obter suporte e percebam valor desde o início.
- Monitorar de perto a satisfação dos clientes nos primeiros meses e intervir proativamente se houver sinais de insatisfação ou baixo uso.

6. Desenvolver Campanhas de Retenção Segmentadas:

- Utilizar os perfis de risco identificados (ex: Idosos, clientes sem parceiro/dependentes, usuários de Fibra sem suporte) para criar ofertas e comunicações de retenção personalizadas.

7. Aprofundar a Análise:

- Continuar a análise com técnicas mais avançadas, como a construção de modelos preditivos de churn, para identificar com maior precisão os clientes em risco e os fatores de maior impacto.
- Coletar dados adicionais, se possível, sobre a satisfação do cliente, motivos de contato com o suporte e uso específico dos serviços para enriquecer a análise.

Implementando estas recomendações, a Telecom X pode atuar de forma mais estratégica para reter seus clientes, reduzir a taxa de churn e fortalecer sua posição no mercado.