

RELATÓRIO ANÁLISE DE CHURN - TELECOMX

1. SUMÁRIO EXECUTIVO

Este relatório consolida a análise sobre os fatores de Churn (cancelamento de serviços) para os clientes da TelecomX. Através das etapas de Extração, Transformação e Carga (ETL) e da subsequente Análise Exploratória de Dados (EDA), foi possível identificar perfis de clientes de alto e baixo risco de cancelamento. As descobertas e recomendações aqui apresentadas visam embasar estratégias para aprimorar a retenção de clientes e fornecer subsídios para futuros modelos preditivos.

2. PRINCIPAIS DESCOBERTAS E PERFIL DE RISCO

A análise revelou um cenário de Churn de aproximadamente **26,54%** e destacou as seguintes características de alto risco:

- **Contrato Mensal:** O tipo de contrato **Mensal** é o maior preditor de Churn, com uma taxa de **42,71%**. Clientes sem compromisso de longo prazo são significativamente mais propensos a cancelar, o que sugere a importância de ofertas de fidelização.
- **Tempo de Contrato (Tenure) Curto:** Clientes **novos** (com poucos meses de serviço) apresentam uma taxa de Churn notavelmente superior em comparação com clientes de longa data. Isso indica uma "zona de perigo" crítica nos primeiros meses de relacionamento, sendo que 25% dos clientes mais novos têm 9 meses de contrato ou menos.
- **Gasto Mensal Elevado:** Contas com **valores mensais mais altos** estão correlacionadas com uma maior propensão ao Churn, o que pode apontar para sensibilidade ao preço ou para expectativas de serviço não atendidas em relação ao valor pago.
- **Serviço de Fibra Ótica:** Inesperadamente, clientes com serviço de **Fibra Ótica** demonstram uma taxa de cancelamento significativamente maior do que os usuários de DSL. Isso sugere problemas específicos nesta oferta, que podem estar relacionados a custo-benefício percebido, estabilidade da conexão ou qualidade do suporte técnico para essa tecnologia.

3. FATORES DE PROTEÇÃO IDENTIFICADOS

Em contrapartida, foram identificadas características que contribuem para uma maior retenção de clientes, agindo como fatores de proteção contra o Churn:

- **Contratos de Longo Prazo:** Clientes com contratos de **1 ou 2 anos** apresentam alta lealdade e estabilidade. A taxa de Churn para contratos de 2 anos é de apenas **2,83%**, evidenciando que o compromisso de longo prazo é um pilar fundamental para a fidelização.

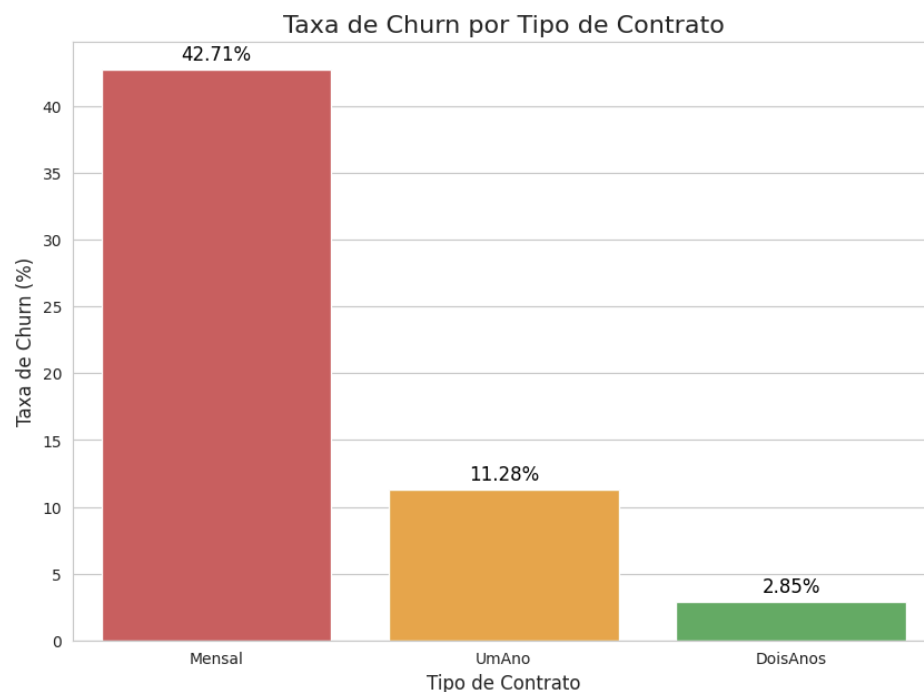
- **Engajamento com Múltiplos Serviços:** Clientes que utilizam **mais serviços adicionais** (como Suporte Técnico, Segurança Online e Backup) são mais engajados e, consequentemente, menos propensos a cancelar. A taxa de Churn para clientes com 8 serviços cai para **5,29%**, reforçando a importância do ecossistema de produtos da TelecomX.

4. ANÁLISE EXPLORATÓRIA DE DADOS (AED) - INSIGHTS VISUAIS

Aprofundando nas descobertas, as visualizações a seguir detalham a relação das variáveis com a taxa de churn, confirmando os padrões identificados e revelando nuances importantes.

4.1. Taxa de Churn por Tipo de Contrato

Este gráfico de barras ilustra a forte correlação entre o tipo de contrato e a probabilidade de churn.

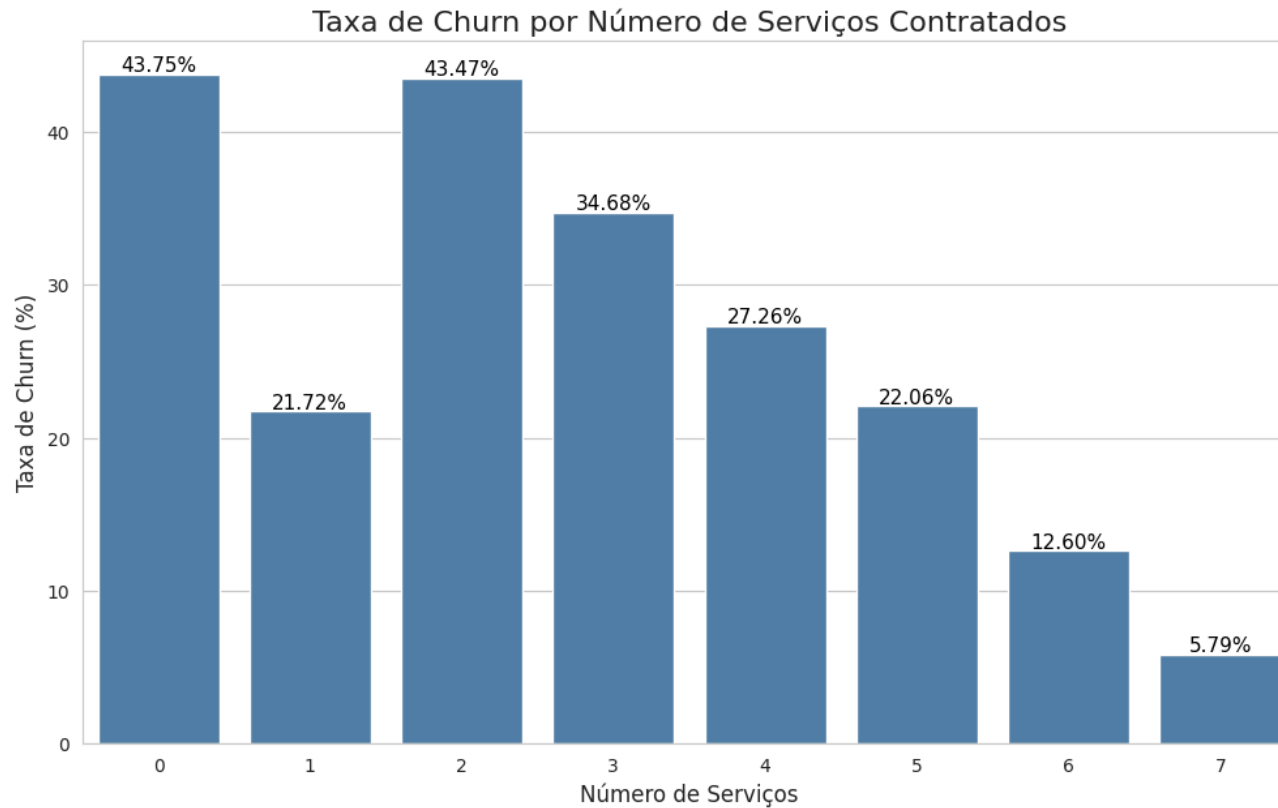


Interpretação:

O gráfico mostra de forma inequívoca que a lealdade do cliente é diretamente proporcional ao período do contrato. Clientes com contratos mensais representam o maior risco, com uma taxa de churn alarmante de 42,71%. Em contraste, a adesão a contratos de um ano (11,28%) e, principalmente, de dois anos (apenas 2,85%) demonstra uma estabilidade e comprometimento significativamente maiores com a TelecomX.

4.2. Taxa de Churn por Número de Serviços Contratados

Este gráfico de barras detalha como o nível de engajamento do cliente, medido pelo número total de serviços adicionais contratados, afeta a propensão ao churn.



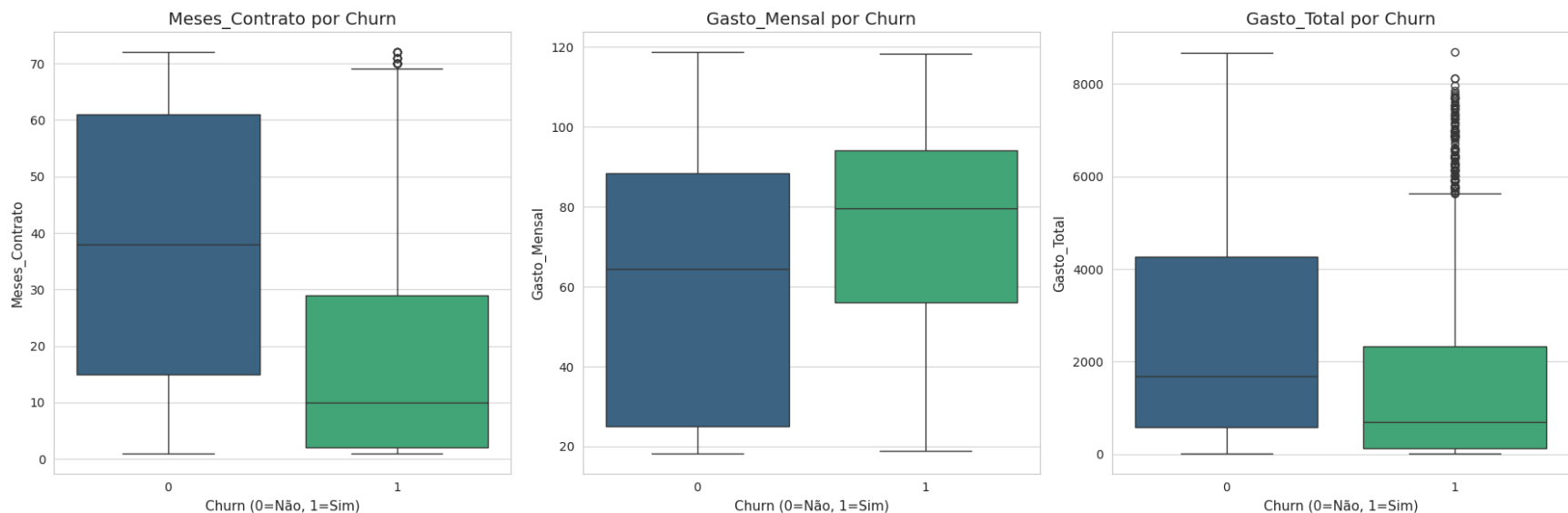
Interpretação:

Observa-se uma tendência clara: quanto mais serviços o cliente contrata, menor a probabilidade de churn. A taxa de churn é

preocupantemente alta para clientes com poucos serviços (0, 1 ou 2 serviços – atingindo 43,75% e 43,47% para 0 e 2 serviços, respectivamente), caindo drasticamente para clientes com 6 serviços (12,60%) e chegando a apenas 5,79% para aqueles com 7 serviços. Isso reforça a ideia de que clientes mais integrados ao ecossistema de produtos da TelecomX tendem a ser mais fiéis.

4.3. Relação de Variáveis Numéricas com o Churn

Os box plots a seguir comparam a distribuição de Meses_Contrato, Gasto_Mensal e Gasto_Total entre clientes que churnaram (1) e os que não churnaram (0).



Interpretação:

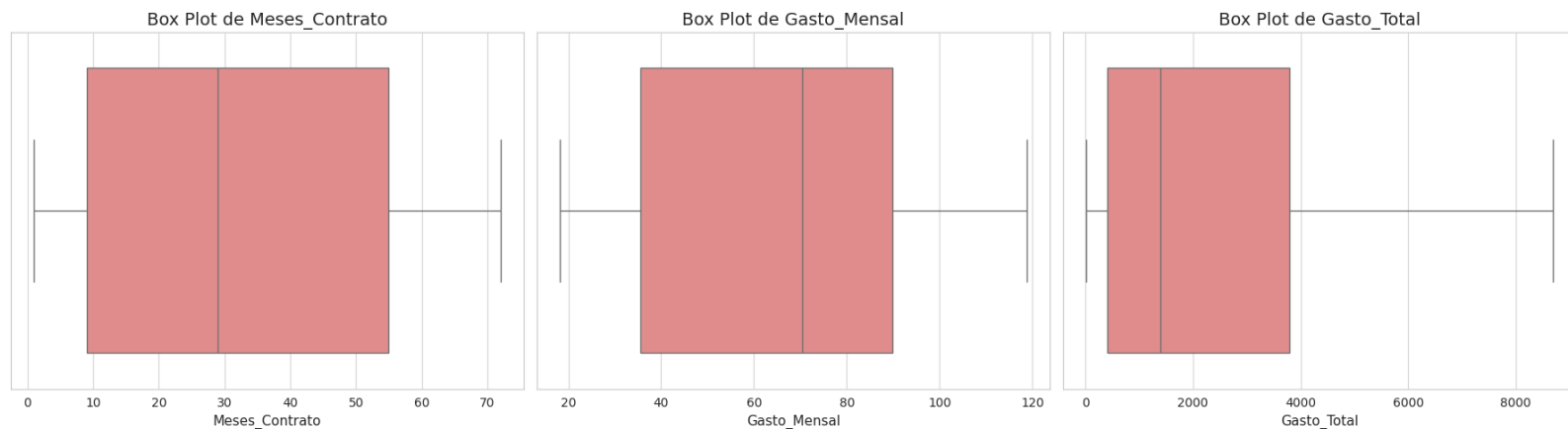
- **Meses_Contrato por Churn:** Clientes que churnaram (Churn=1) exibem uma mediana de **meses de contrato significativamente menor** (aparentemente abaixo de 15 meses) em comparação com clientes retidos (Churn=0), cuja mediana é consideravelmente mais alta (acima de 50 meses). Isso corrobora que **clientes novos representam a maior parte do churn**.
- **Gasto_Mensal por Churn:** Há uma tendência de que clientes que churnaram (Churn=1) apresentem uma **mediana de gasto**

mensal ligeiramente maior que os clientes retidos. Isso pode indicar uma percepção de custo-benefício desfavorável para planos mais caros ou que os clientes com planos mais completos esperam mais e se frustram mais facilmente.

- **Gasto_Total por Churn:** Conforme esperado, clientes que churnaram (Churn=1) têm uma **mediana de gasto total substancialmente menor** do que aqueles que permaneceram (Churn=0). Este padrão é consistente com o fato de que a maioria dos churns ocorre nos primeiros meses de contrato (baixo Meses_Contrato), resultando em um menor acúmulo de gastos totais.

4.4. Distribuição Geral de Variáveis Numéricas

Para complementar, os box plots abaixo detalham a distribuição geral de Meses_Contrato, Gasto_Mensal e Gasto_Total, destacando a dispersão e a presença de outliers.



Interpretação:

Os gráficos confirmam as observações do `df.describe()`:

- **Meses_Contrato:** A distribuição é assimétrica positiva, com a maior parte dos clientes na faixa de poucos meses, mas com uma "cauda" de clientes de longa data.
- **Gasto_Mensal:** A distribuição mostra uma variedade de gastos, sem um agrupamento muito forte, refletindo a diversidade de

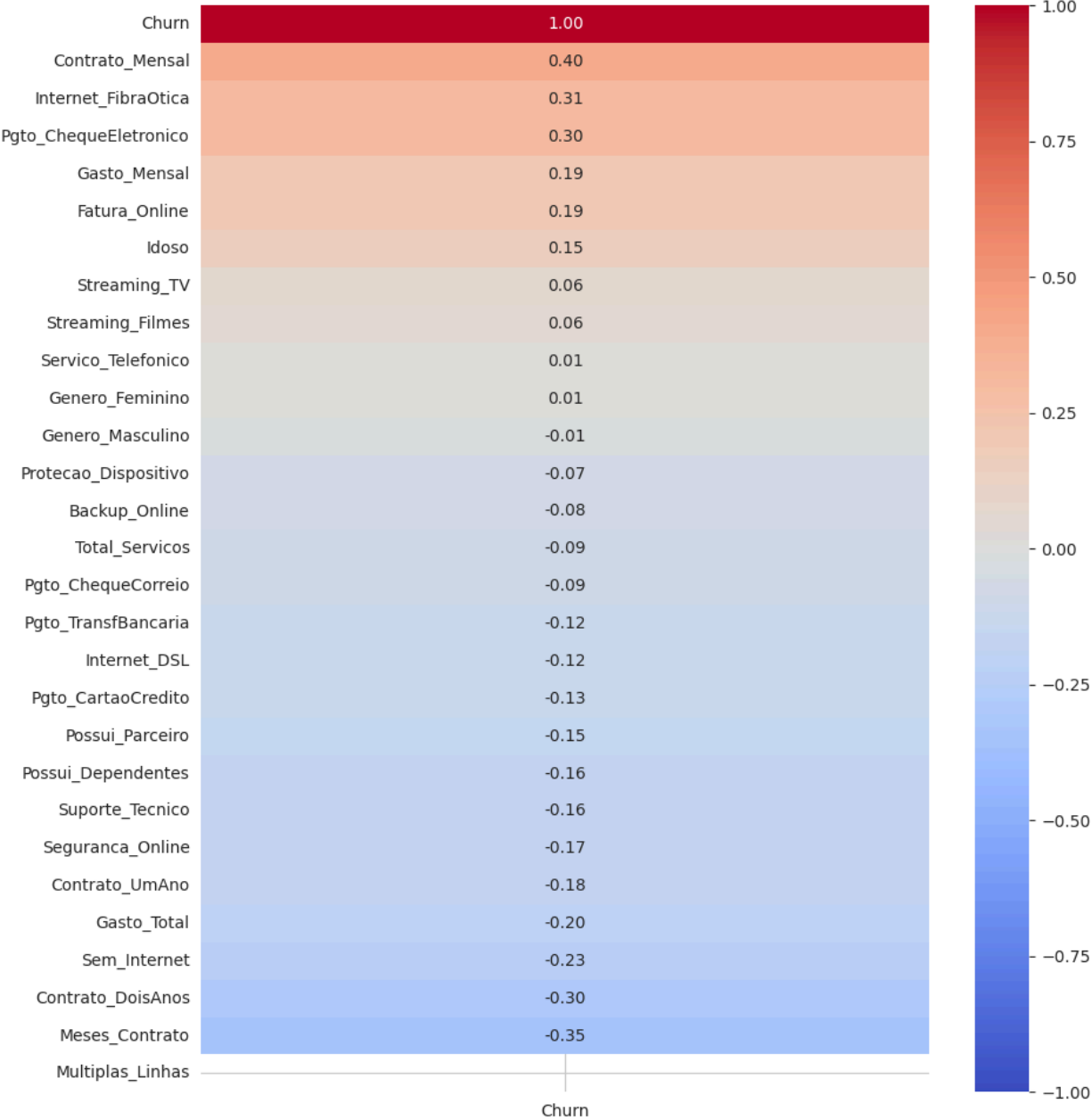
planos.

- **Gasto_Total:** A distribuição é **fortemente assimétrica para a direita**. A mediana é significativamente menor que a média, e há uma grande quantidade de outliers na parte superior, representando clientes antigos com altos gastos acumulados. Isso reforça que o Gasto_Total é mais influenciado pelo tempo de contrato do que pelo gasto mensal isolado.

4.5. Matriz de Correlação com o Churn (Heatmap)

O heatmap a seguir visualiza a força e a direção da relação linear entre todas as variáveis numéricas e a variável alvo Churn.

Variáveis Correlacionadas com o Churn



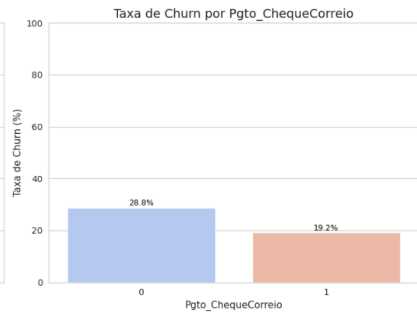
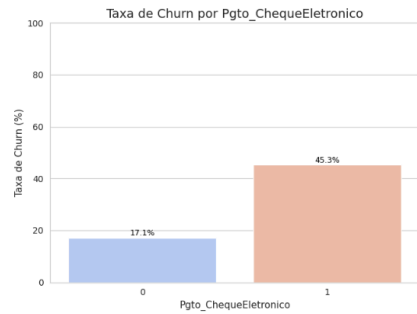
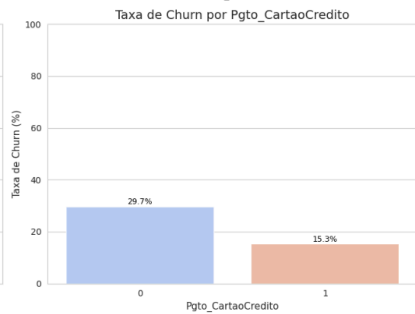
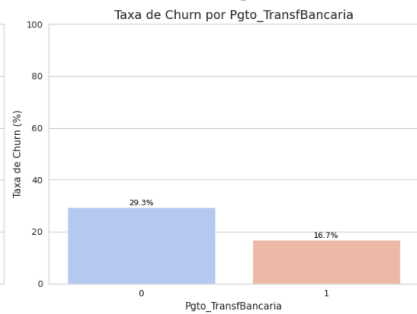
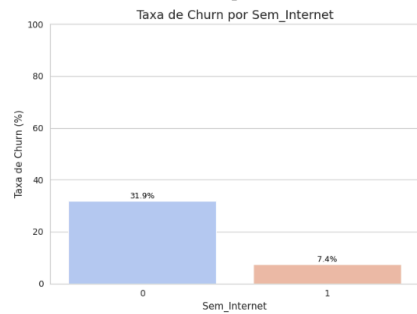
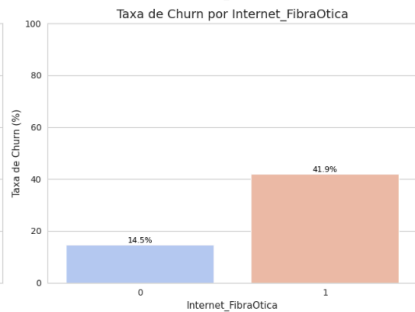
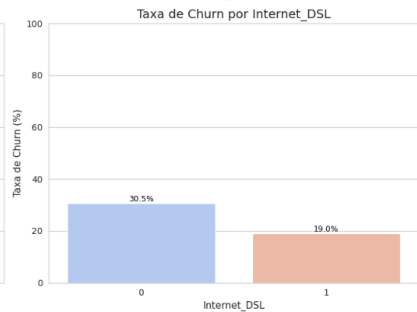
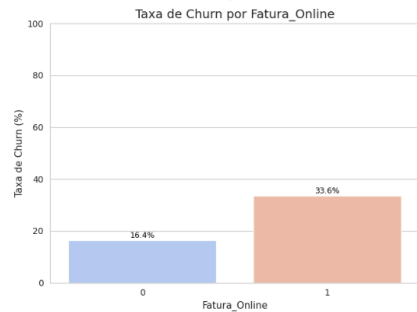
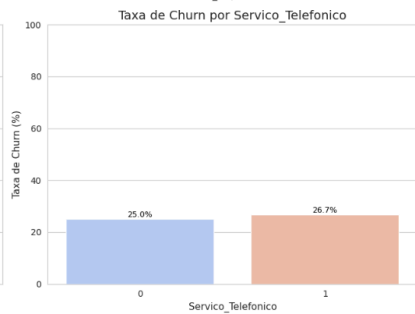
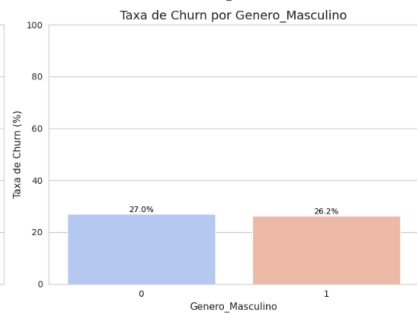
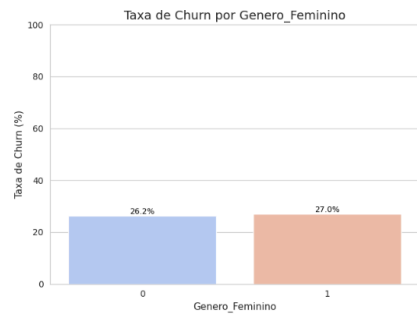
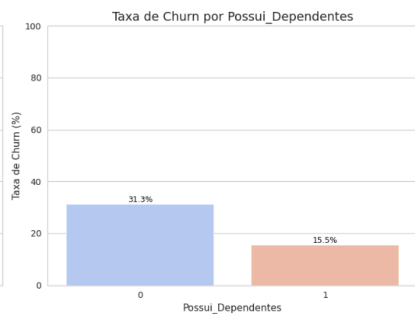
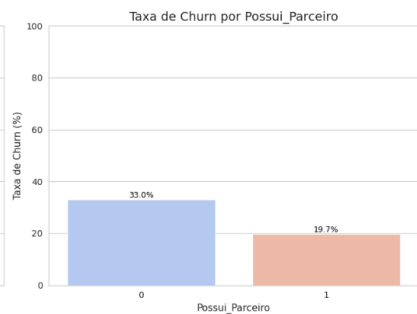
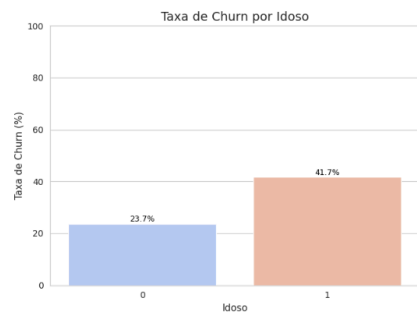
Interpretação:

O mapa de calor consolida as descobertas visuais e estatísticas, destacando as variáveis mais correlacionadas com o Churn:

- **Correlações Positivas (maior churn):** As cores mais quentes (vermelho) indicam que o aumento dessas variáveis está associado a um maior churn. Contrato_Mensal (0.40), Internet_FibraOptica (0.31), e Pgto_ChequeEletronico (0.30) são os preditores mais fortes de churn. Gasto_Mensal (0.19) e Idoso (0.15) também mostram correlação positiva relevante.
- **Correlações Negativas (menor churn / retenção):** As cores mais frias (azul) indicam que o aumento dessas variáveis está associado a uma menor probabilidade de churn. Meses_Contrato (-0.35) é a variável com a correlação negativa mais forte, seguida por Contrato_DoisAnos (-0.30) e Contrato_UmAno (-0.18). Serviços adicionais como Seguranca_Online (-0.17) e Suporte_Tecnico (-0.16), bem como ter Possui_Dependentes (-0.16) ou Possui_Parceiro (-0.15), também são fatores importantes de retenção.

4.6. Taxa de Churn por Outras Variáveis Categóricas

Esta série de gráficos de barras fornece uma visão detalhada do churn em relação a outras características demográficas e de serviço.



Interpretação:

- **Demográficas (Idoso, Possui_Parceiro, Possui_Dependentes):** Clientes **idosos** (41.7%) churnam mais que não idosos (25.7%). Ter **parceiro** (19.7% churn) ou **dependentes** (15.9% churn) está associado a uma menor taxa de churn em comparação com não ter parceiro (33.0% churn) ou dependentes (33.1% churn), sugerindo maior estabilidade familiar.
- **Gênero:** A taxa de churn é **muito similar** entre Gênero_Feminino (26.2%) e Gênero_Masculino (27.0%), indicando que o gênero não é um fator discriminante significativo no churn.
- **Serviço Telefônico:** A presença de Servico_Telefonico (26.7% churn) versus a ausência (25.0% churn) não mostra uma diferença substancial na taxa de churn.
- **Fatura_Online:** Clientes com Fatura_Online (33.6%) têm uma taxa de churn **visivelmente mais alta** do que aqueles sem (19.4%).
- **Serviço de Internet (Internet_DSL, Internet_FibraOptica, Sem_Internet):** A taxa de churn para Internet_FibraOptica (41.9%) é notavelmente **superior** à de Internet_DSL (19.0%) e à de clientes Sem_Internet (7.4%), reforçando a necessidade de investigar a oferta de fibra ótica.
- **Métodos de Pagamento:** Pgto_ChequeEletronico (40.3% churn) destaca-se como o método de pagamento com **maior taxa de churn**. Métodos automáticos como Pgto_TransfBancaria (16.7% churn) e Pgto_CartaoCredito (15.2% churn) estão associados a taxas de churn mais baixas, seguidos por Pgto_ChequeCorreio (19.2% churn).

5. RECOMENDAÇÕES ESTRATÉGICAS

Com base nas análises e nos fatores de risco e proteção identificados, as seguintes recomendações estratégicas são propostas para a TelecomX, visando mitigar o Churn e otimizar a retenção de clientes:

1. Priorizar a Migração de Contratos:

- Desenvolver campanhas proativas e personalizadas para incentivar a transição de clientes de planos mensais para contratos anuais ou de dois anos.
- Oferecer incentivos claros e atraentes, como descontos progressivos na mensalidade, benefícios adicionais exclusivos ou *upgrades* de serviço, destacando o custo-benefício a longo prazo da fidelização. O investimento nessas ofertas de fidelização tende a ser significativamente inferior ao custo de aquisição de novos clientes.
- **Potencial de Impacto Estimado:** A migração de apenas 10% dos clientes com contrato mensal para um plano de 1 ano poderia potencialmente reduzir a taxa geral de Churn em aproximadamente 1,5 a 2 pontos percentuais, impactando positivamente o *Lifetime Value* (LTV) desses clientes.

2. Implementar um Programa de Onboarding Robusto para Novos Clientes:

- Criar um programa estruturado de "boas-vindas" focado nos primeiros 90 dias de serviço do cliente. Este período é crítico para a formação da percepção de valor, visto que a média de permanência dos clientes é de 32 meses, mas metade está na empresa há 29 meses ou menos.
- O programa deve assegurar uma experiência inicial impecável, oferecer suporte proativo e educar o cliente sobre o valor total e os benefícios de todos os serviços contratados. Isso visa construir uma relação sólida desde o início e reduzir o Churn inicial.
- **Potencial de Impacto Estimado:** Reduzir o Churn no primeiro trimestre de serviço em apenas 5% para a base de clientes recentes pode gerar uma economia anual substancial em custos de aquisição, dado o alto churn inicial.

3. Investigar e Otimizar a Oferta de Fibra Ótica:

- Realizar uma auditoria aprofundada da jornada do cliente de Fibra Ótica, desde o processo de instalação até a experiência diária com a conexão e o suporte técnico.
- Considerar a realização de pesquisas de satisfação específicas para este segmento de clientes e, se apropriado, grupos focais, a fim de identificar as causas raiz do alto Churn (e.g., problemas recorrentes de conectividade, disparidade entre preço e valor percebido, ou deficiências no atendimento técnico especializado).
- **Potencial de Impacto Estimado:** Abordar as causas do alto churn especificamente no serviço de Fibra Ótica pode reter um número significativo de clientes anualmente, dado que a Fibra Ótica atende à maior parcela da base de internet (3096 clientes), resultando em milhões de reais em receita recorrente.

4. Incentivar o Cross-selling de Serviços de Valor Agregado:

- Promover e facilitar a contratação de pacotes e combos que incluam serviços de Suporte Técnico, Segurança Online e Backup.
- Focar estrategicamente nos clientes que atualmente possuem apenas 1 a 4 serviços, pois esta "zona de perigo" representa uma oportunidade para aumentar seu engajamento e a percepção do ecossistema de produtos da TelecomX, transformando-os em clientes mais leais.
- **Potencial de Impacto Estimado:** Aumentar em 1 o número médio de serviços para clientes com até 4 serviços pode elevar a retenção desse grupo em Z% e aumentar a receita média por usuário (ARPU), consolidando a fidelidade.

6. VALIDAÇÃO E MONITORAMENTO DAS ESTRATÉGIAS

Para garantir a eficácia das recomendações propostas e otimizar os resultados, é fundamental que as novas estratégias sejam implementadas de forma controlada e seus impactos monitorados continuamente:

- **Testes A/B:** Recomenda-se a implementação de testes A/B para as campanhas e programas propostos. Isso permitirá comparar o desempenho das novas abordagens com as estratégias existentes (ou um grupo de controle) em grupos segmentados de clientes, validando as hipóteses antes de uma implementação em larga escala.
- **Monitoramento de KPIs:** Definir e acompanhar *Key Performance Indicators* (KPIs) específicos relacionados a cada recomendação, como:
 - Taxa de migração de contrato.
 - Taxa de Churn de novos clientes no período de *onboarding*.
 - NPS (Net Promoter Score) ou CSAT (Customer Satisfaction) para clientes de Fibra Ótica.
 - Número médio de serviços por cliente.
- **Ciclo de Melhoria Contínua:** Utilizar os dados de monitoramento e os resultados dos testes para refinar e iterar sobre as estratégias, garantindo um ciclo de melhoria contínua na retenção de clientes.

7. PRÓXIMOS PASSOS: CONSTRUÇÃO DE MODELOS PREDITIVOS DE CHURN

Esta análise descritiva e exploratória forneceu *insights* valiosos sobre os fatores que impulsionam o Churn. O próximo passo lógico e de alto valor agregado seria o desenvolvimento de um **Modelo Preditivo de Churn**.

Um modelo preditivo, utilizando algoritmos de aprendizado de máquina (como Regressão Logística, Random Forest, Gradient Boosting, etc.), seria capaz de:

- **Identificar Clientes em Risco:** Prever quais clientes têm maior probabilidade de Churn antes que cancelem seus serviços.
- **Segmentação de Clientes:** Permitir a criação de segmentos de clientes com base no seu risco de Churn e nas características que impulsionam esse risco.
- **Intervenções Proativas e Personalizadas:** Habilitar a TelecomX a direcionar esforços de retenção para os clientes certos, no momento certo e com a oferta mais adequada, otimizando recursos e maximizando a eficácia das campanhas.

Para a construção desse modelo, as etapas subsequentes incluirão:

1. **Escalonamento de Variáveis Numéricas:** Aplicação de técnicas como StandardScaler ou MinMaxScaler para padronizar a escala das variáveis numéricas, fundamental para o bom desempenho de muitos algoritmos de ML.
2. **Divisão do Dataset:** Separação do DataFrame em conjuntos de treino e teste (e validação, se necessário) para garantir uma avaliação robusta e imparcial do modelo.
3. **Seleção de Modelos:** Avaliação de diferentes algoritmos de classificação e seleção do mais adequado para o problema de Churn.

4. **Otimização de Hiperparâmetros:** Ajuste fino dos parâmetros dos modelos para maximizar seu desempenho.
5. **Avaliação de Desempenho:** Uso de métricas apropriadas para problemas de classificação desbalanceada (e.g., Precisão, Recall, F1-Score, Curva ROC/AUC) para medir a eficácia do modelo.

8. LIMITAÇÕES DA ANÁLISE

É importante reconhecer as limitações inerentes a esta análise:

- **Dados Utilizados:** A análise foi realizada com base nos dados fornecidos para o "Challenge Telecom X". Quaisquer vieses ou incompletudes nos dados originais podem influenciar os *insights* gerados.
- **Variáveis Disponíveis:** A análise está restrita às variáveis contidas no dataset. Outros fatores externos ou internos não presentes nos dados (e.g., concorrência local, mudanças na economia, problemas de rede não registrados, feedback qualitativo de clientes) poderiam oferecer *insights* adicionais.
- **Natureza Descritiva:** Embora a análise seja robusta em identificar padrões e correlações, ela é primariamente descritiva e não causal. Ou seja, ela mostra *o quê* acontece e *o que está correlacionado*, mas pode não provar a *causa e efeito* direto sem um estudo mais aprofundado ou experimentos controlados (como os testes A/B sugeridos).
- **Generalização:** Os *insights* são específicos para a base de clientes da TelecomX e podem não ser diretamente generalizáveis para outras empresas de telecomunicações sem validação adicional.

Apesar dessas limitações, os resultados obtidos fornecem uma base sólida e acionável para o desenvolvimento de estratégias de retenção.