

Python Insights - Analisando Dados com Python

Case - Cancelamento de Clientes

Você foi contratado por uma empresa com mais de 800 mil clientes para um projeto de Dados. Recentemente a empresa percebeu que da sua base total de clientes, a maioria são clientes inativos, ou seja, que já cancelaram o serviço.

Precisando melhorar seus resultados ela quer conseguir entender os principais motivos desses cancelamentos e quais as ações mais eficientes para reduzir esse número.

Base de dados e arquivos: https://drive.google.com/drive/folders/1uDesZePdkhiraJmiyeZ-w5tfc8XsNYFZ?usp=drive_link

```
In [1]: # Pseudocodigo
# Passo 1: Importar a base dados OK
# Passo 2: Visualizar a base de dados OK
# Passo 3: Tratamento de erros (Corrigir os erros da base de dados) OK
# Tratar valores vazios OK
# Retirar informacoes inuteis (informacao que nao te ajuda, te atrapalha) OK
# Passo 4: Analise inicial da base de dados (entender como estao os cancelamentos)
# Passo 5: Analise profunda da base de dados (encontrando a causa dos cancelamentos)
```

```
In [2]: pip install pandas numpy openpyxl plotly
```

```
Requirement already satisfied: pandas in c:\users\nicho\anaconda3\lib\site-packages (1.4.4)
Requirement already satisfied: numpy in c:\users\nicho\anaconda3\lib\site-packages (1.21.5)
Requirement already satisfied: openpyxl in c:\users\nicho\anaconda3\lib\site-packages (3.0.10)
Requirement already satisfied: plotly in c:\users\nicho\anaconda3\lib\site-packages (5.9.0)
Requirement already satisfied: pytz>=2020.1 in c:\users\nicho\anaconda3\lib\site-packages (from pandas) (2022.1)
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\nicho\anaconda3\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: et_xmlfile in c:\users\nicho\anaconda3\lib\site-packages (from openpyxl) (1.1.0)
Requirement already satisfied: tenacity>=6.2.0 in c:\users\nicho\anaconda3\lib\site-packages (from plotly) (8.0.1)
Requirement already satisfied: six>=1.5 in c:\users\nicho\anaconda3\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

```
In [3]: # Passo 1 e 2
import pandas
tabela_cancel = pandas.read_csv('cancelamentos.csv')
#Passo 3
# linha --> axis = 0
# coluna --> axis = 1
tabela_cancel = tabela_cancel.drop('CustomerID', axis=1)
display(tabela_cancel)
```

	idade	sexo	tempo_como_cliente	frequencia_uso	ligacoes_callcenter	dias_atraso	assinatura	duracao
0	30.0	Female	39.0	14.0	5.0	18.0	Standard	
1	65.0	Female	49.0	1.0	10.0	8.0	Basic	
2	55.0	Female	14.0	4.0	6.0	18.0	Basic	
3	58.0	Male	38.0	21.0	7.0	7.0	Standard	
4	23.0	Male	32.0	20.0	5.0	8.0	Basic	
...
881661	42.0	Male	54.0	15.0	1.0	3.0	Premium	
881662	25.0	Female	8.0	13.0	1.0	20.0	Premium	
881663	26.0	Male	35.0	27.0	1.0	5.0	Standard	
881664	28.0	Male	55.0	14.0	2.0	0.0	Standard	
881665	31.0	Male	48.0	20.0	1.0	14.0	Premium	

881666 rows × 11 columns

```
In [4]: # Passo 3 continue
display(tabela_cancel.info())
tabela_cancel = tabela_cancel.dropna()
display(tabela_cancel.info())

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 881666 entries, 0 to 881665
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   idade                                881664 non-null  float64
1   sexo                                881664 non-null  object
2   tempo_como_cliente                  881663 non-null  float64
3   frequencia_uso                      881663 non-null  float64
4   ligacoes_callcenter                881664 non-null  float64
5   dias_atraso                        881664 non-null  float64
6   assinatura                          881661 non-null  object
7   duracao_contrato                    881663 non-null  object
8   total_gasto                         881664 non-null  float64
9   meses_ultima_interacao             881664 non-null  float64
10  cancelou                            881664 non-null  float64
dtypes: float64(8), object(3)
memory usage: 74.0+ MB
None

<class 'pandas.core.frame.DataFrame'>
Int64Index: 881659 entries, 0 to 881665
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   idade                                881659 non-null  float64
1   sexo                                881659 non-null  object
2   tempo_como_cliente                  881659 non-null  float64
3   frequencia_uso                      881659 non-null  float64
4   ligacoes_callcenter                881659 non-null  float64
5   dias_atraso                        881659 non-null  float64
6   assinatura                          881659 non-null  object
7   duracao_contrato                    881659 non-null  object
8   total_gasto                         881659 non-null  float64
9   meses_ultima_interacao             881659 non-null  float64
10  cancelou                            881659 non-null  float64
dtypes: float64(8), object(3)
memory usage: 80.7+ MB
```

None

```
In [5]: # Passo 4
display(tabela_cancel['cancelou'].value_counts())
display(tabela_cancel['cancelou'].value_counts(normalize=True))

1.0    499993
0.0    381666
Name: cancelou, dtype: int64
1.0    0.567105
0.0    0.432895
Name: cancelou, dtype: float64
```

```
In [6]: # Passo 4 continue

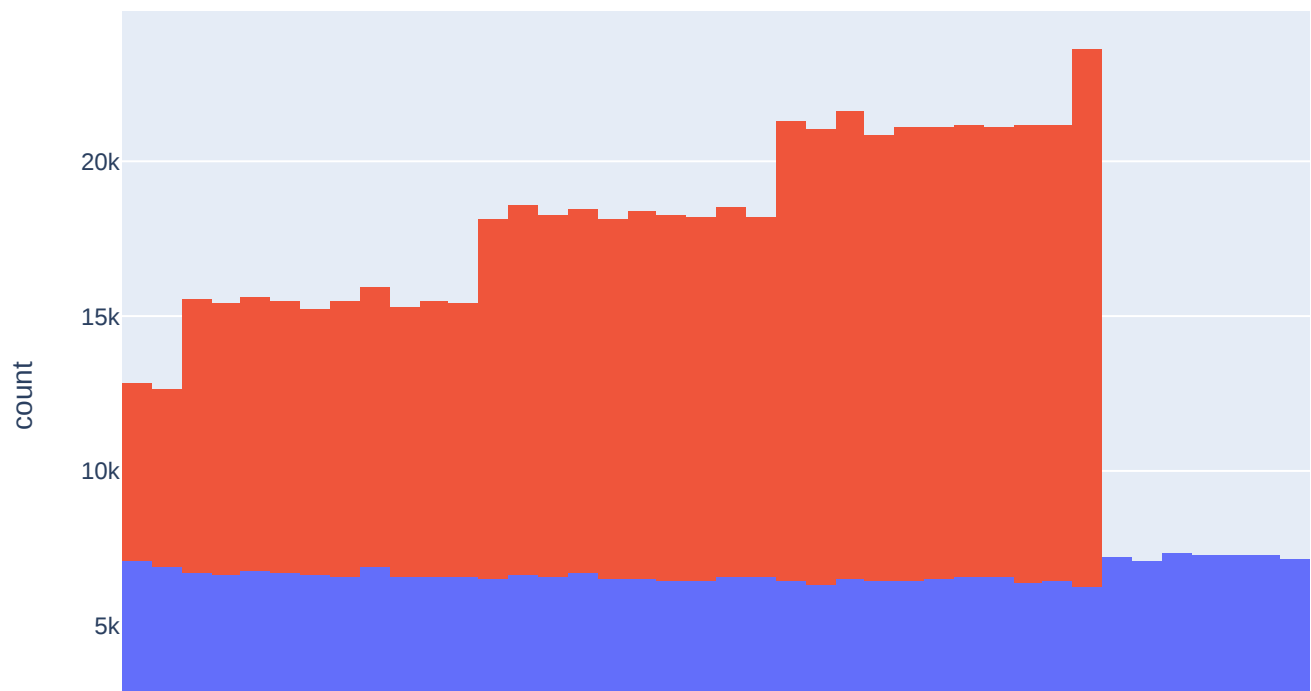
display(tabela_cancel.groupby("duracao_contrato").mean())
# Clientes que tem a assinatura Monthly tem 100% de cancelamento
```

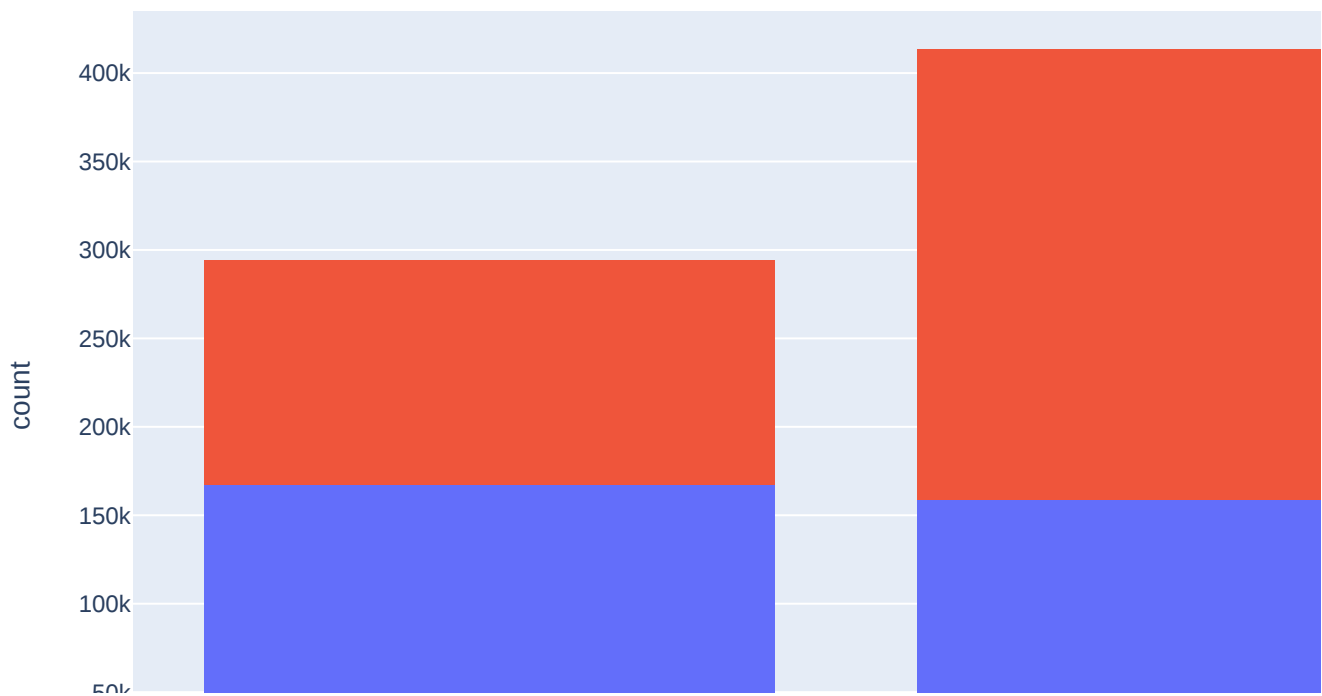
	idade	tempo_como_cliente	frequencia_uso	ligacoes_callcenter	dias_atraso	total_gasto
duracao_contrato						
Annual	38.842165	31.446186	15.880213	3.263401	12.465156	651.697738
Monthly	41.552407	30.538555	15.499274	4.985649	15.007267	550.616435
Quarterly	38.830938	31.419916	15.886662	3.265245	12.460863	651.427783

```
In [7]: tabela_cancel = tabela_cancel[tabela_cancel['duracao_contrato']!='Monthly']
display(tabela_cancel['cancelou'].value_counts())
display(tabela_cancel['cancelou'].value_counts(normalize=True))

0.0    381666
1.0    325788
Name: cancelou, dtype: int64
0.0    0.539492
1.0    0.460508
Name: cancelou, dtype: float64
```

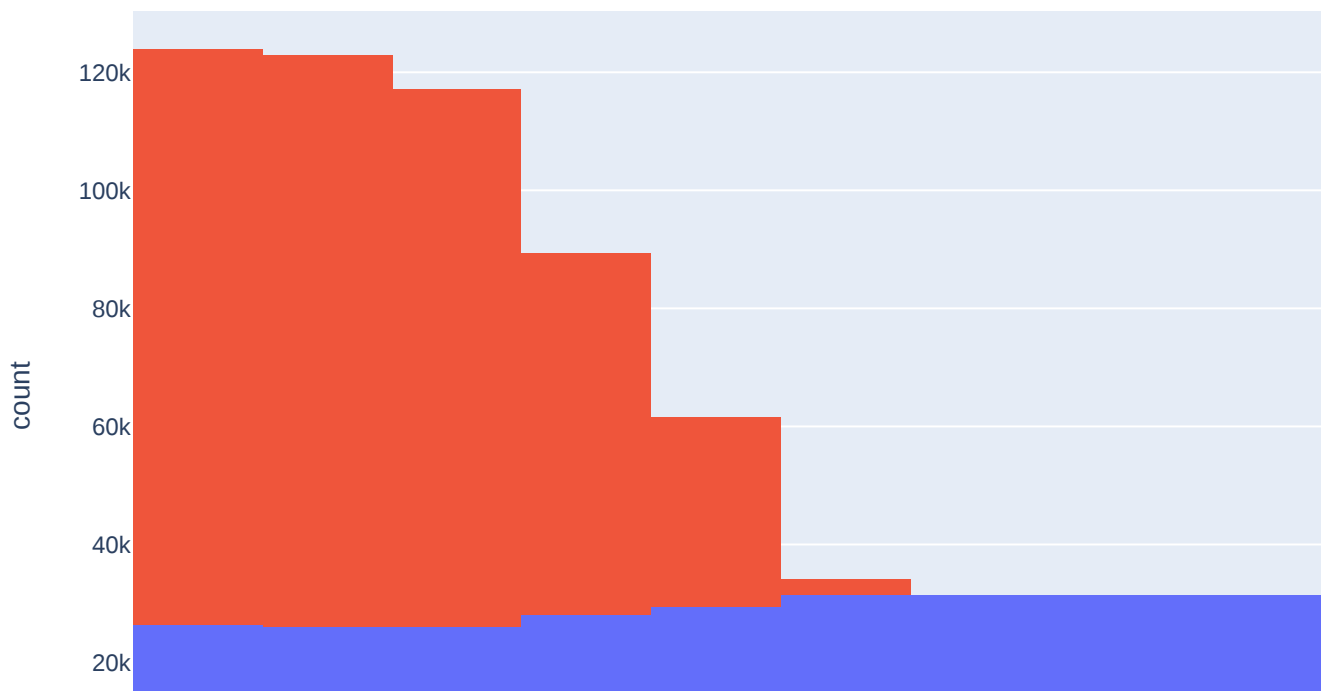
```
In [8]: # Passo 5
import plotly.express as px
for coluna in tabela_cancel.columns:
    grafico = px.histogram(tabela_cancel, x=coluna, color='cancelou')
    grafico.show()
```

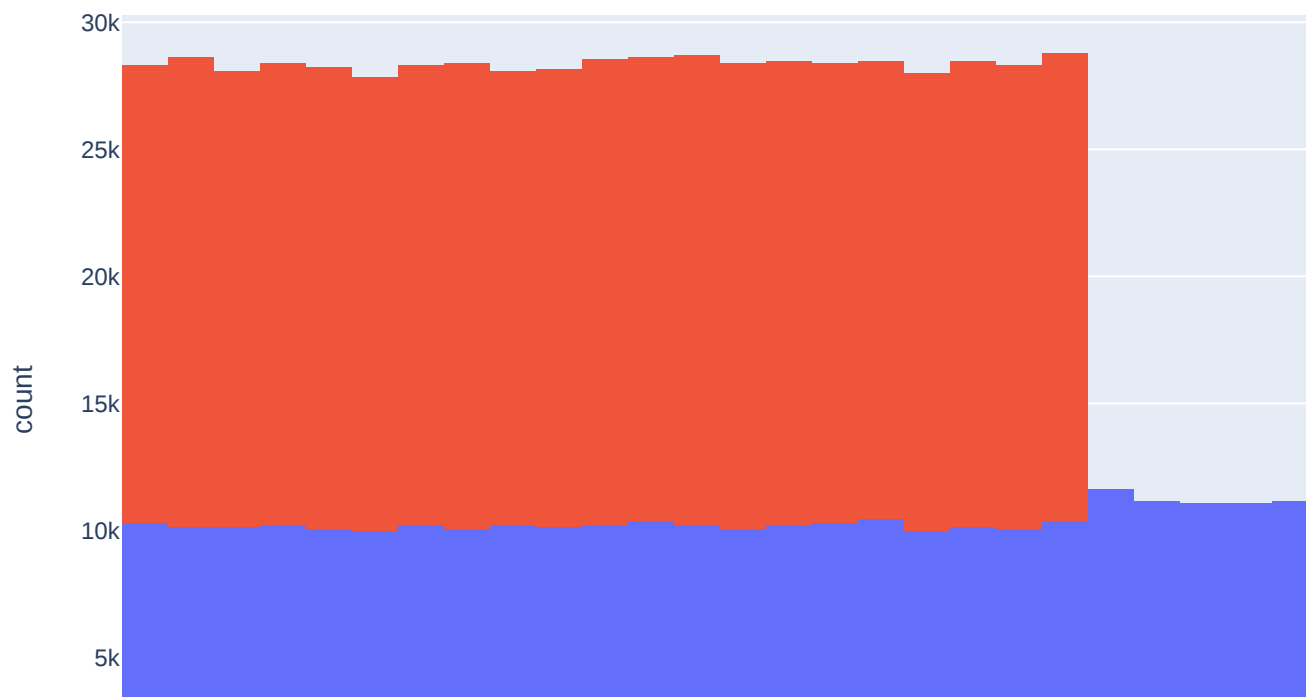




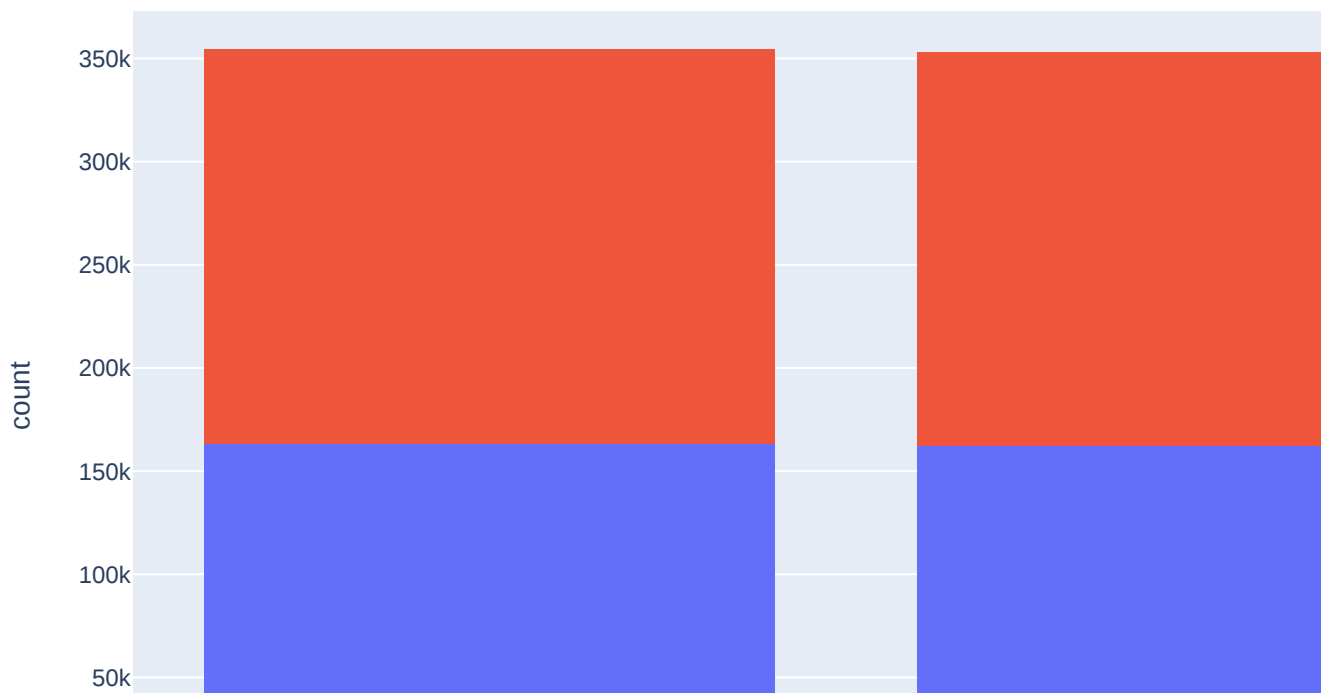


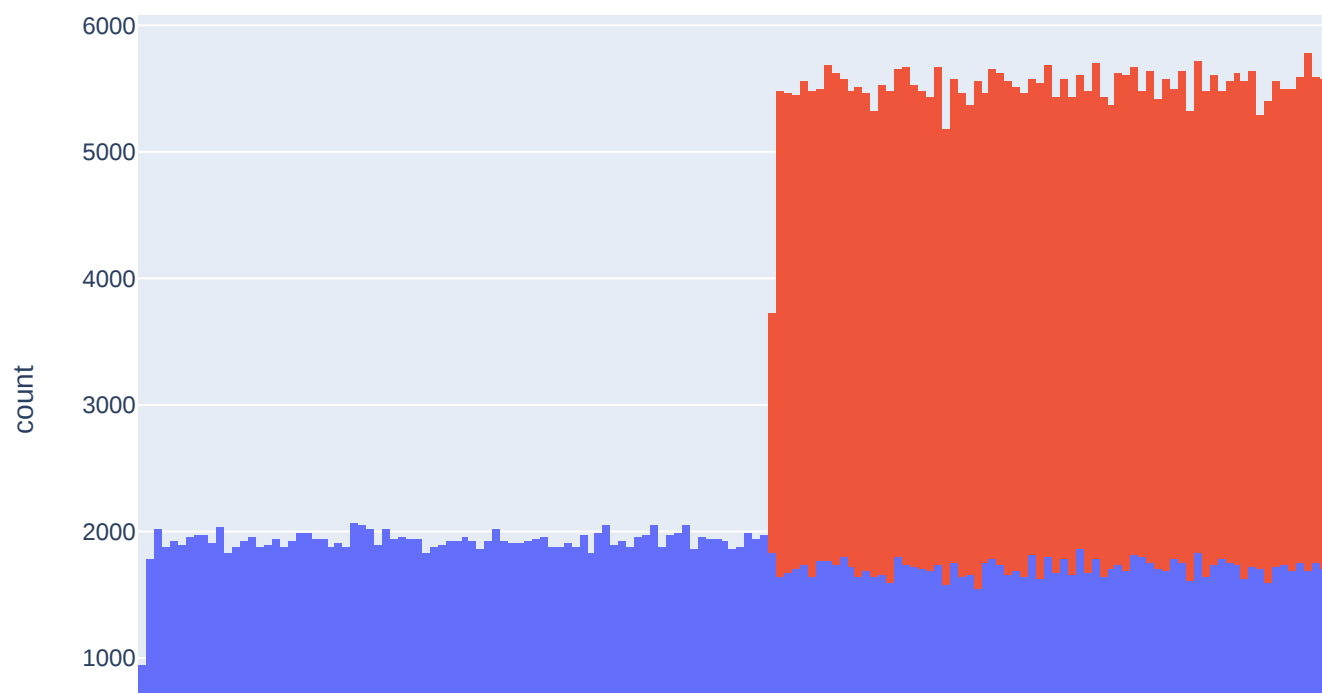




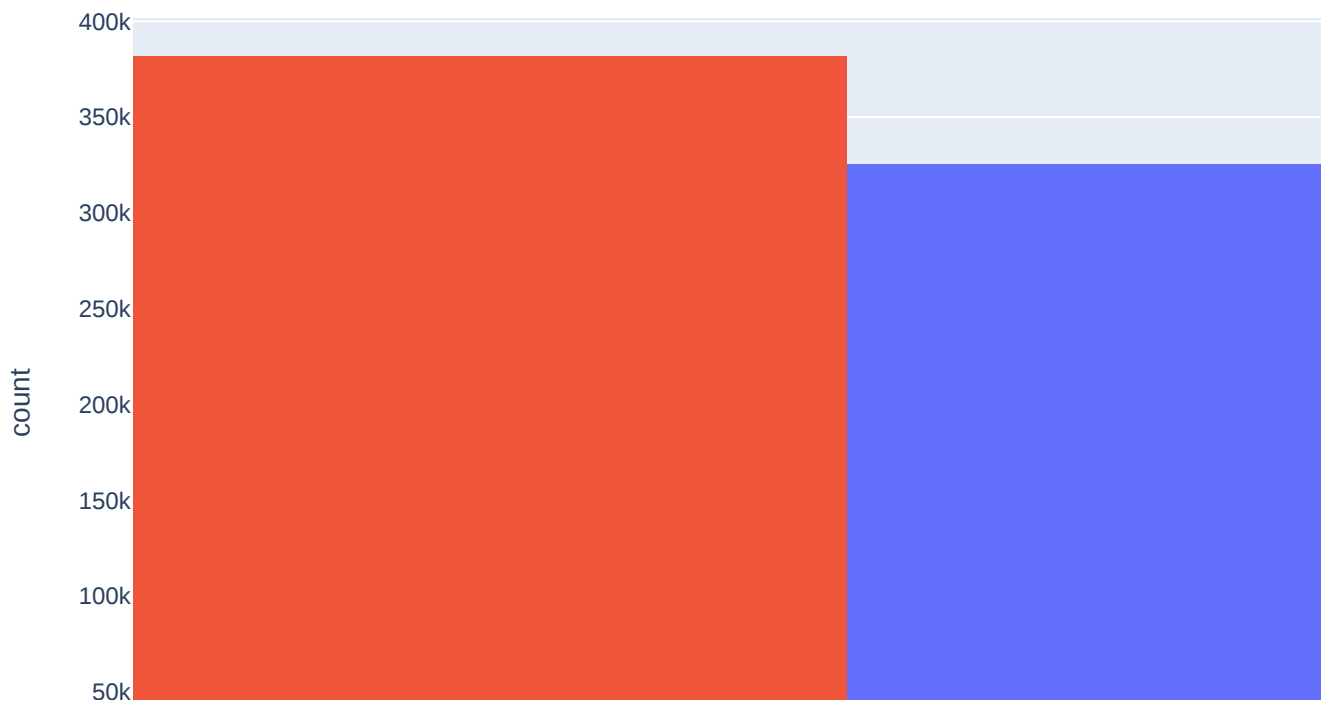












```
In [9]: # Problema 1: Clientes assima de 51 anos de idade tem 100% de chance de cancelar
tabela_cancel = tabela_cancel[tabela_cancel['ligacoes_callcenter']<=51]
# Problema 2: Clientes que ligam mais de 5 vezes ao callcenter tem 100% de chance de cancelar
tabela_cancel = tabela_cancel[tabela_cancel['ligacoes_callcenter']<5]
# Problema 3: Clientes que atrasam o pagamento a cima de 21 dias tem 100% de chance de cancelar
tabela_cancel = tabela_cancel[tabela_cancel['dias_atraso']<=20]
# Problema 4: Clientes que gastam menos de R$500 tem 100% de chance de cancelar
tabela_cancel = tabela_cancel[tabela_cancel['total_gasto']>500]
# Problema 5: Clientes que adquirem a assinatura Monthly tem 100% de chance de cancelar
tabela_cancel = tabela_cancel[tabela_cancel['duracao_contrato']!='Monthly']
display(tabela_cancel['cancelou'].value_counts())
display(tabela_cancel['cancelou'].value_counts(normalize=True))

0.0    379020
1.0     37474
Name: cancelou, dtype: int64
0.0    0.910025
1.0    0.089975
Name: cancelou, dtype: float64
```

Possivel resolucao para os problemas encontrados

Problema 1: Clientes assima de 51 anos de idade tem 100% de chance de cancelar

Facilidade de uso e acessibilidade: Adaptar a interface dos serviços, aplicativos ou site da empresa para torná-los mais amigáveis e acessíveis aos usuários mais velhos. Considerar a ampliação de fontes, uso de alto contraste e uma navegação intuitiva. Ofertas e descontos especiais: Oferecer ofertas,

descontos ou programas de fidelidade exclusivos para clientes acima de 51 anos pode incentivá-los a permanecer com a empresa por mais tempo. Programas de educação e suporte: Oferecer programas de educação e suporte para ajudar os clientes mais velhos a utilizar melhor os serviços e recursos disponíveis. Isso pode ser feito através de tutoriais, webinars, workshops ou até mesmo suporte técnico especializado. Feedback contínuo: Estabelecer um canal de comunicação aberto para receber feedback contínuo dos clientes acima de 51 anos, permitindo que eles expressem suas opiniões e sugestões para melhorias. Campanhas de marketing direcionadas: Desenvolver campanhas de marketing direcionadas especificamente para clientes mais velhos, destacando os benefícios do serviço para essa faixa etária. Parcerias com organizações e associações: Estabelecer parcerias com organizações ou associações que atendam a pessoas acima de 51 anos, para aumentar a conscientização sobre os serviços e oferecer benefícios exclusivos. Monitoramento de métricas: Monitorar regularmente as métricas de satisfação e retenção dos clientes acima de 51 anos para acompanhar o progresso e identificar áreas que precisam de melhorias contínuas.

Problema 2: Clientes que ligam mais de 5 vezes ao callcenter tem 100% de chance de cancelar

Melhorar o atendimento inicial: Garantir que o atendimento inicial seja rápido, eficiente e resolva o problema do cliente logo na primeira ligação. Treinar os atendentes para entenderem as necessidades do cliente desde o início e fornecer soluções eficazes. Atendimento ao cliente de qualidade: Investir em treinamento contínuo para o pessoal do call center, capacitando-os a lidar com situações desafiadoras de forma empática, cortês e profissional. Registro de histórico de chamadas: Manter um registro detalhado do histórico de chamadas e interações anteriores com cada cliente. Isso permitirá que os atendentes tenham acesso rápido às informações relevantes quando o cliente ligar novamente, evitando a repetição de perguntas e problemas.

Problema 3: Clientes que atrasam o pagamento a cima de 21 dias tem 100% de chance de cancelar

Opções de pagamento flexíveis: Oferecer opções de pagamento flexíveis, como parcelamentos ou datas de vencimento personalizadas, para acomodar as diferentes situações financeiras dos clientes. Descontos por pagamento antecipado: Oferecer descontos ou benefícios especiais para clientes que pagam suas faturas antes da data de vencimento, incentivando o pagamento pontual. Política de tolerância: Implementar uma política de tolerância para atrasos menores, permitindo que os clientes evitem taxas de atraso ou penalidades caso paguem a fatura dentro de um prazo razoável após a data de vencimento. Programas de fidelidade: Criar programas de fidelidade ou recompensas para clientes que mantêm um histórico de pagamentos pontuais, oferecendo incentivos para que eles continuem a cumprir com os prazos de pagamento.

Problema 4: Clientes que gastam menos de reais 500 tem 100% de chance de cancelar

Parcerias estratégicas: Estabelecer parcerias com outras empresas ou estabelecimentos para oferecer benefícios e descontos exclusivos aos clientes que gastam menos, incentivando o uso contínuo dos serviços. Ofertas e promoções personalizadas: Analisar o histórico de gastos de cada cliente e oferecer ofertas e promoções personalizadas que atendam às suas necessidades e interesses. Isso pode incentivar os clientes a permanecer com o serviço e aproveitar as vantagens exclusivas. Pacotes e planos flexíveis:

Oferecer pacotes e planos flexíveis que sejam mais adequados às necessidades e ao orçamento dos

clientes que gastam menos. Isso pode incluir opções mais econômicas ou pacotes com serviços essenciais para essa faixa de consumidores.

Problema 5: Clientes que adquirem a assinatura Monthly tem 100% de chance de cancelar

Oferecer benefícios adicionais: Incluir benefícios ou recursos exclusivos para os assinantes mensais, tornando a assinatura mais atraente e valiosa em relação às opções avulsas ou concorrentes. Ofertas de upgrade: Oferecer oportunidades de upgrade para planos com mais benefícios ou serviços adicionais, incentivando os assinantes a permanecerem na empresa. Personalização: Personalizar a experiência dos assinantes mensais com recomendações de conteúdo, serviços ou produtos com base em suas preferências e histórico de uso.

A implementação das orientações propostas tem o potencial de contribuir significativamente para a redução do problema de cancelamento dos clientes, resultando em uma diminuição substancial na taxa de cancelamento, possivelmente alcançando níveis entre 18% e 9%. No entanto, é importante reconhecer que a eficácia dessas ações pode ser influenciada por uma série de fatores adicionais, incluindo a resposta dos clientes às estratégias implementadas, as condições do mercado e a capacidade da empresa de se adaptar e melhorar continuamente suas práticas de atendimento e retenção de clientes.