

ANÁLISE PREDITIVA DE CANCELAMENTO DE CARTÕES DE CRÉDITO COM DATA MINING: PERFIS DE UTILIZAÇÃO E ESTRATÉGIAS DE RETENÇÃO

Náthalie Patzer - 102817

102817@aluno.uricer.edu.br

INTRODUÇÃO

O cenário atual do mercado financeiro é marcado por uma intensa competição e pela necessidade de oferecer serviços personalizados que atendam às expectativas dos clientes. O cartão de crédito, uma ferramenta essencial para a gestão financeira pessoal, tem sido um dos produtos mais impactados por essa dinâmica. No entanto, o cancelamento de cartões de crédito por parte dos clientes representa um desafio significativo para as instituições financeiras, que buscam não apenas reter clientes, mas também entender os fatores que levam a essa decisão.

A mineração de dados (data mining) surge como uma abordagem poderosa para explorar grandes volumes de informações, permitindo a extração de padrões e insights valiosos. Neste trabalho, serão utilizadas técnicas de data mining em uma base de dados contendo informações de clientes de um banco, com o objetivo de identificar os principais motivos que levam ao cancelamento de cartões de crédito. A análise incluirá variáveis estatísticas, comportamentais e transacionais dos clientes, permitindo uma compreensão mais aprofundada do perfil daqueles que optaram pelo cancelamento.

A relevância deste estudo se manifesta na capacidade de fornecer recomendações estratégicas para a melhoria dos serviços oferecidos, além de contribuir para a implementação de políticas de retenção de clientes mais eficazes. Ao final, espera-se que os resultados desta análise possam auxiliar o banco em sua jornada de aprimoramento contínuo, promovendo a satisfação do cliente e a sustentabilidade dos seus negócios.

FERRAMENTAS E TECNOLOGIAS UTILIZADAS

Kaggle: O Kaggle é uma plataforma de ciência de dados onde usuários podem acessar e compartilhar conjuntos de dados, competir em competições de machine learning e colaborar em projetos de data science. Para este trabalho, o Kaggle é útil como fonte de conjuntos de dados para análise e como ambiente colaborativo, onde podem ser realizados experimentos com dados e técnicas de mineração de dados.

Link da base de dados utilizada: <https://www.kaggle.com/sakshigoyal7/credit-card-customers>

Google Colab: O Google Colab é uma ferramenta de notebook online gratuita que permite escrever e executar código Python diretamente no navegador. Integrado ao Google Drive e suportando GPUs, o Colab facilita a execução de projetos de ciência de dados sem a necessidade de configurar um ambiente local complexo. No contexto deste trabalho, o Colab é utilizado para desenvolver, testar e visualizar o código de mineração de dados em Python, bem como os gráficos gerados com a análise.

Google Drive: O Google Drive é uma plataforma de armazenamento em nuvem que permite salvar, compartilhar e sincronizar arquivos. Em conjunto com o Google Colab, o Google Drive facilita o acesso e o armazenamento dos conjuntos de dados utilizados no projeto, garantindo que eles estejam disponíveis para análise a qualquer momento e em qualquer dispositivo.

Python: Python é uma linguagem de programação amplamente utilizada em ciência de dados devido à sua simplicidade e à vasta quantidade de bibliotecas disponíveis, como Pandas, NumPy, Matplotlib e Scikit-Learn. Essas bibliotecas oferecem funcionalidades para manipulação, análise e visualização de dados, bem como para o desenvolvimento de modelos de machine learning. Neste projeto, Python é a linguagem escolhida para implementar as técnicas de data mining e geração dos gráficos utilizados na análise.

METODOLOGIA E ETAPAS DA ANÁLISE DE DADOS

Para realizar a análise dos motivos de cancelamento de cartões de crédito, será utilizada uma metodologia estruturada em cinco etapas, detalhadas a seguir:

Passo 1: Importar a Base de Dados

A primeira etapa consiste em importar a base de dados que contém informações sobre os clientes e seus respectivos cartões de crédito. Utilizaremos a biblioteca *Pandas* para carregar o arquivo e preparar o ambiente de análise.

Passo 2: Visualizar e Tratar a Base de Dados

Após a importação, a base de dados é visualizada para verificar sua estrutura inicial. Realizaremos o tratamento de dados, que inclui a limpeza de valores nulos, a remoção de dados duplicados e a padronização de informações, para garantir que a base esteja consistente e pronta para a análise.

Passo 3: Visualização Geral da Base de Dados

Em seguida, faremos uma visualização geral das colunas e dos tipos de informações contidos na tabela, utilizando métodos como *.info()* e *.describe()* do *Pandas*. Essa etapa permite que identifiquemos as variáveis disponíveis e compreendamos melhor os dados, ajudando a direcionar as próximas etapas da análise.

Passo 4: Construir uma Análise com Base nos Dados Disponibilizados

Nesta etapa, aplicaremos técnicas de análise descritiva e exploratória. Utilizamos gráficos e estatísticas para investigar padrões e tendências relevantes, como idade média dos clientes, faixas de renda, uso do cartão e outros fatores que podem influenciar o cancelamento. A biblioteca *Matplotlib* será utilizada para gerar os gráficos que auxiliarão na análise dos dados.

Passo 5: Identificar os Principais Motivos dos Clientes Estarem Cancelando o Cartão de Crédito

Por fim, com base na análise exploratória e nas informações observadas, identificaremos os principais fatores que contribuem para o cancelamento dos cartões. Fatores como taxas elevadas, baixa utilização do cartão e estatísticas de uso no período de um ano são examinados para encontrar correlações e padrões que expliquem os cancelamentos.

ANÁLISE DOS DADOS

Antes de começarmos com a análise da nossa base de dados, precisamos importá-la via Google Drive para o Google Colab, através do seguinte comando:

```
[ ] from google.colab import drive
    drive.mount('/content/drive')
```

O primeiro passo após importar a base de dados, é realizar a importação da biblioteca *Pandas*, que será utilizada na análise, visualização e manipulação dos dados da base de dados. A informações serão armazenadas em uma variável, para que possamos manipulá-las e visualizá-las. Utilizamos o parâmetro *encoding* para tratar os caracteres especiais, e em seguida, o comando *display*, passando como argumento a variável na qual armazenamos os dados, para exibi-los. Pode ser utilizado também o comando *.head()* para a visualização dos dados.

```
import pandas as pd

tabela = pd.read_csv("/content/drive/MyDrive/Tópicos Especiais/ClientesBanco.csv", encoding='latin1')

display(tabela)
```

	CLIENTNUM	Categoria	Idade	Sexo	Dependentes	Educação	Estado Civil	Faixa Salarial Anual	Categoria Cartão	Meses como Cliente	...	Inatividade 12m	Contatos 12m	Limite	Limite Consumido	Limite Disponível	Mudanças Transacoes_Q4_Q1	Valor Transacoes 12m	Qtde Transacoes 12m	Mudança Qtde Transações_Q4_Q1	Taxa de Utilização Cartão
0	768805383	Cliente	45	M	3	Ensino Médio	Casado	\$60K - \$80K	Blue	39	...	1	3	12691.0	777	11914.0	1335.000	1144	42	1625.000	0.061
1	818770008	Cliente	49	F	5	Ensino Superior	Solteiro	Less than \$40K	Blue	44	...	1	2	8256.0	864	7392.0	1541.000	1291	33	3714.000	0.105
2	713982108	Cliente	51	M	3	Ensino Superior	Casado	\$80K - \$120K	Blue	36	...	1	0	3418.0	0	3418.0	2594.000	1887	20	2333.000	0.000
3	769911858	Cliente	40	F	4	Ensino Médio	Não informado	Less than \$40K	Blue	34	...	4	1	3313.0	2517	796.0	1405.000	1171	20	2333.000	0.760
4	709106358	Cliente	40	M	3	Sem ensino formal	Casado	\$60K - \$80K	Blue	21	...	1	0	4716.0	0	4716.0	2175.000	816	28	2.500	0.000
...
10122	772366833	Cliente	50	M	2	Ensino Superior	Solteiro	\$40K - \$60K	Blue	40	...	2	3	4003.0	1851	2152.0	0.703	15476	117	0.857	0.462
10123	710638233	Cancelado	41	M	2	Não informado	Divorciado	\$40K - \$60K	Blue	25	...	2	3	4277.0	2186	2091.0	0.804	8764	69	0.683	0.511
10124	716506083	Cancelado	44	F	1	Ensino Médio	Casado	Less than \$40K	Blue	36	...	3	4	5409.0	0	5409.0	0.819	10291	60	0.818	0.000
10125	717406983	Cancelado	30	M	2	Ensino Superior	Não informado	\$40K - \$60K	Blue	36	...	3	3	5281.0	0	5281.0	0.535	8395	62	0.722	0.000
10126	714337233	Cancelado	43	F	2	Ensino Superior	Casado	Less than \$40K	Silver	25	...	2	4	10388.0	1961	8427.0	0.703	10294	61	0.649	0.189

10127 rows x 21 columns

10127 rows x 21 columns

Após realizar uma visualização inicial da base de dados, identificamos as colunas que não possuem informações relevantes para a análise que faremos, e então, as removemos, para deixar a base de dados mais “limpa” e adequada à análise, tornando a manipulação dos dados mais fácil e eficiente. No caso da base de dados que utilizamos, iremos remover a coluna “CLIENTNUM”, que é apenas um identificador numérico de cada cliente na base de dados e não é relevante à análise do motivo pelo qual os clientes vêm cancelando o cartão de crédito, a qual é nosso objetivo. Para removê-la, utilizamos o comando `.drop()`, utilizando o parâmetro `axis=1`, para identificar que o objeto que estamos removendo é uma coluna (se fosse linha, utilizaríamos `axis=0`). Em seguida, utilizamos o comando `display` novamente para verificar se as alterações foram feitas corretamente.

```
[ ] tabela = tabela.drop("CLIENTNUM", axis=1)

[ ] display(tabela)
```

	Categoria	Idade	Sexo	Dependentes	Educação	Estado Civil	Faixa Salarial Anual	Categoria Cartão	Meses como Cliente	Produtos Contratados	Inatividade 12m	Contatos 12m	Limite	Limite Consumido	Limite Disponível	Mudanças Transacoes_Q4_Q1	Valor Transacoes 12m	Qtde Transacoes 12m	Mudança Qtde Transações_Q4_Q1	Taxa de Utilização Cartão
0	Cliente	45	M	3	Ensino Médio	Casado	\$60K - \$80K	Blue	39	5	1	3	12691.0	777	11914.0	1335.000	1144	42	1625.000	0.061
1	Cliente	49	F	5	Ensino Superior	Solteiro	Less than \$40K	Blue	44	6	1	2	8256.0	864	7392.0	1541.000	1291	33	3714.000	0.105
2	Cliente	51	M	3	Ensino Superior	Casado	\$80K - \$120K	Blue	36	4	1	0	3418.0	0	3418.0	2594.000	1887	20	2333.000	0.000
3	Cliente	40	F	4	Ensino Médio	Não informado	Less than \$40K	Blue	34	3	4	1	3313.0	2517	796.0	1405.000	1171	20	2333.000	0.760
4	Cliente	40	M	3	Sem ensino formal	Casado	\$60K - \$80K	Blue	21	5	1	0	4716.0	0	4716.0	2175.000	816	28	2.500	0.000
...
10122	Cliente	50	M	2	Ensino Superior	Solteiro	\$40K - \$60K	Blue	40	3	2	3	4003.0	1851	2152.0	0.703	15476	117	0.857	0.462
10123	Cancelado	41	M	2	Não informado	Divorciado	\$40K - \$60K	Blue	25	4	2	3	4277.0	2186	2091.0	0.804	8764	69	0.683	0.511
10124	Cancelado	44	F	1	Ensino Médio	Casado	Less than \$40K	Blue	36	5	3	4	5409.0	0	5409.0	0.819	10291	60	0.818	0.000
10125	Cancelado	30	M	2	Ensino Superior	Não informado	\$40K - \$60K	Blue	36	4	3	3	5281.0	0	5281.0	0.535	8395	62	0.722	0.000
10126	Cancelado	43	F	2	Ensino Superior	Casado	Less than \$40K	Silver	25	6	2	4	10388.0	1961	8427.0	0.703	10294	61	0.649	0.189

10127 rows x 20 columns

Em seguida, analisamos quais informações relevantes cada coluna da base de dados pode nos trazer. No nosso exemplo:

- Categoria: Identifica se a pessoa ainda é cliente ou já cancelou o cartão.
- Idade: Informa a idade do cliente.
- Sexo: Informa o sexo do cliente.
- Dependentes: Informa a quantidade de dependentes do cliente.
- Educação: Informa o nível de formação da pessoa.
- Faixa salarial anual: Categoria de salário na qual o usuário se encaixa.

- Categoria do cartão: Indica o tipo de cartão.
- Meses como cliente: Quantos meses a pessoa foi cliente.
- Produtos contratados: Quantos produtos do banco o cliente utiliza.
- Inatividade nos últimos 12 meses: Por quanto tempo o usuário ficou sem utilizar o cartão na faixa de um ano.
- Contatos nos últimos 12 meses: Quantas vezes o cliente precisou de alguma atividade de suporte do banco na faixa de um ano.
- Limites de crédito: Indica os limites de crédito do cliente, o limite que ele já utilizou, e o limite disponível.
- Valor transações: Quanto o cliente gastou no cartão de crédito nos últimos 12 meses.
- Quantidade transações: Qual foi a quantidade de transações que o cliente realizou em um período de 12 meses.
- Taxa de utilização do cartão: Indica um índice de usabilidade do cartão pelo cliente.

Em seguida, podemos utilizar o comando *info*, pertencente também à biblioteca *Pandas*, para exibir um resumo informativo sobre os dados da nossa base de dados. Isso é útil para que possamos ter uma ideia geral dos tipos de dados de cada coluna, a quantidade de registros em cada coluna, e também para identificar se há campos nulos.

```
[ ] display(tabela.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10127 entries, 0 to 10126
Data columns (total 20 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Categoria                            10127 non-null  object
1   Idade                                10127 non-null  int64
2   Sexo                                 10127 non-null  object
3   Dependentes                          10127 non-null  int64
4   Educação                             10127 non-null  object
5   Estado Civil                         10127 non-null  object
6   Faixa Salarial Anual                 10127 non-null  object
7   Categoria Cartão                     10126 non-null  object
8   Meses como Cliente                  10127 non-null  int64
9   Produtos Contratados                 10127 non-null  int64
10  Inatividade 12m                      10127 non-null  int64
11  Contatos 12m                         10127 non-null  int64
12  Limite                               10127 non-null  float64
13  Limite Consumido                     10127 non-null  int64
14  Limite Disponível                    10127 non-null  float64
15  Mudanças Transacoes_Q4_Q1           10127 non-null  float64
16  Valor Transacoes 12m                 10127 non-null  int64
17  Qtde Transacoes 12m                  10127 non-null  int64
18  Mudança Qtde Transações_Q4_Q1       10127 non-null  float64
19  Taxa de Utilização Cartão            10127 non-null  float64
dtypes: float64(5), int64(9), object(6)
memory usage: 1.5+ MB
None
```

Analisando as informações que foram apresentadas, podemos notar que a coluna “Categoria de Cartão” possui um registro a menos do que as demais, o que significa que provavelmente ela possui uma linha com registro nulo. Para que isso não atrapalhe a análise, a melhor maneira de contornar a situação é apenas excluindo essa entrada incompleta da base de dados. Para tal, utilizaremos o comando *dropna*, que removerá campos com informações ausentes, ou faltantes.

```
[ ] tabela = tabela.dropna()
display(tabela.info())
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 10126 entries, 0 to 10126
Data columns (total 20 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   Categoria                                10126 non-null  object
1   Idade                                    10126 non-null  int64
2   Sexo                                    10126 non-null  object
3   Dependentes                             10126 non-null  int64
4   Educação                                10126 non-null  object
5   Estado Civil                             10126 non-null  object
6   Faixa Salarial Anual                     10126 non-null  object
7   Categoria Cartão                         10126 non-null  object
8   Meses como Cliente                      10126 non-null  int64
9   Produtos Contratados                    10126 non-null  int64
10  Inatividade 12m                         10126 non-null  int64
11  Contatos 12m                            10126 non-null  int64
12  Limite                                  10126 non-null  float64
13  Limite Consumido                        10126 non-null  int64
14  Limite Disponível                       10126 non-null  float64
15  Mudanças Transacoes_Q4_Q1              10126 non-null  float64
16  Valor Transacoes 12m                   10126 non-null  int64
17  Qtde Transacoes 12m                    10126 non-null  int64
18  Mudança Qtde Transações_Q4_Q1          10126 non-null  float64
19  Taxa de Utilização Cartão              10126 non-null  float64
dtypes: float64(5), int64(9), object(6)
memory usage: 1.6+ MB
None
```

Após feito isso, utilizamos novamente o comando *info*, para verificar se os dados foram tratados corretamente, e se eles se mantêm íntegros. Agora podemos verificar que a base de dados está uniforme, e com a mesma quantidade de registros em todas as colunas, o que a deixa pronta para as análises necessárias. Podemos usar também o método *describe*, que fornece um resumo estatístico de colunas numéricas ou categóricas, nos ajudando a entender a distribuição dos dados.

```
[ ] display(tabela.describe())
```

	Idade	Dependentes	Meses como Cliente	Produtos Contratados	Inatividade 12m	Contatos 12m	Limite	Limite Consumido	Limite Disponível	Mudanças Transacoes_Q4_Q1	Valor Transacoes 12m	Qtde Transacoes 12m	Mudança Qtde Transações_Q4_Q1	Taxa de Utilização Cartão
count	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000	10126.000000
mean	46.327079	2.346139	35.928995	3.812463	2.341300	2.455264	8631.962088	1162.764369	7469.197719	108.713167	4404.391764	64.861841	66.579995	0.274901
std	8.016420	1.298956	7.986593	1.554440	1.010584	1.106267	9089.225428	815.012238	9091.132356	353.912795	3397.157925	23.471594	284.001363	0.275704
min	26.000000	0.000000	13.000000	1.000000	0.000000	0.000000	1438.300000	0.000000	3.000000	0.000000	510.000000	10.000000	0.000000	0.000000
25%	41.000000	1.000000	31.000000	3.000000	2.000000	2.000000	2555.000000	358.000000	1324.250000	0.631000	2156.000000	45.000000	0.582000	0.023000
50%	46.000000	2.000000	36.000000	4.000000	2.000000	2.000000	4549.000000	1276.000000	3473.000000	0.736000	3899.500000	67.000000	0.702000	0.175500
75%	52.000000	3.000000	40.000000	5.000000	3.000000	3.000000	11067.750000	1784.000000	9860.000000	0.859000	4741.000000	81.000000	0.818000	0.503000
max	73.000000	5.000000	56.000000	6.000000	6.000000	6.000000	34516.000000	2517.000000	34516.000000	3397.000000	18484.000000	139.000000	3714.000000	0.999000

Analisando a descrição dos campos da tabela, conseguimos extrair várias informações úteis, como:

- Count: quantos itens há na tabela
- Mean: média
- Std: desvio padrão
- Min: valor mínimo
- Max: valor máximo
- 25% na coluna Idade, por exemplo, isso significa que 25% dos clientes estão abaixo dos 41 anos, 50% estão abaixo dos 46 anos, e 75% estão abaixo dos 52 anos.
- Coluna de inatividade: em média, nos últimos 12 meses, os clientes ficaram 2.3 meses inativos. Metade dos clientes ficaram 2 meses ou menos inativos. No máximo, um cliente ficou 6 meses inativo.

Para uma melhor visualização dos dados, podemos utilizar o método `round()`, para realizar o ajuste da precisão das casas decimais das informações.

```
[ ] display(tabela.describe().round(1))
```

	Idade	Dependentes	Meses como Cliente	Produtos Contratados	Inatividade 12m	Contatos 12m	Limite	Limite Consumido	Limite Disponível	Mudanças Transacoes_Q4_Q1	Valor Transacoes 12m	Qtde Transacoes 12m	Mudança Qtde Transações_Q4_Q1	Taxa de Utilização Cartão
count	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0	10126.0
mean	46.3	2.3	35.9	3.8	2.3	2.5	8632.0	1162.8	7469.2	108.7	4404.4	64.9	66.6	0.3
std	8.0	1.3	8.0	1.6	1.0	1.1	9089.2	815.0	9091.1	353.9	3397.2	23.5	284.0	0.3
min	26.0	0.0	13.0	1.0	0.0	0.0	1438.3	0.0	3.0	0.0	510.0	10.0	0.0	0.0
25%	41.0	1.0	31.0	3.0	2.0	2.0	2555.0	358.0	1324.2	0.6	2156.0	45.0	0.6	0.0
50%	46.0	2.0	36.0	4.0	2.0	2.0	4549.0	1276.0	3473.0	0.7	3899.5	67.0	0.7	0.2
75%	52.0	3.0	40.0	5.0	3.0	3.0	11067.8	1784.0	9860.0	0.9	4741.0	81.0	0.8	0.5
max	73.0	5.0	56.0	6.0	6.0	6.0	34516.0	2517.0	34516.0	3397.0	18484.0	139.0	3714.0	1.0

A primeira coluna que é interessante analisarmos, é a coluna de “Categoria”, para identificarmos a quantidade (absoluta e em percentual) de clientes que não cancelaram o cartão, e de clientes que cancelaram. Para isso, podemos utilizar o método `.value_counts()`, que vai retornar a contagem de valores únicos na coluna desejada. Utilizando o mesmo método, com o parâmetro `normalize`, obtemos esses dados em valor percentual. Podemos concluir que de todos os usuários analisados, aproximadamente 84% se manteve como cliente do banco, e 16% cancelou o cartão de crédito.

```
[ ] qtde_categoria = tabela['Categoria'].value_counts()
display(qtde_categoria)
```

	count
Categoria	
Cliente	8499
Cancelado	1627

dtype: int64

```
[ ] qtde_categoria_perc = tabela['Categoria'].value_counts(normalize=True)
display(qtde_categoria_perc)
```

	proportion
Categoria	
Cliente	0.839325
Cancelado	0.160675

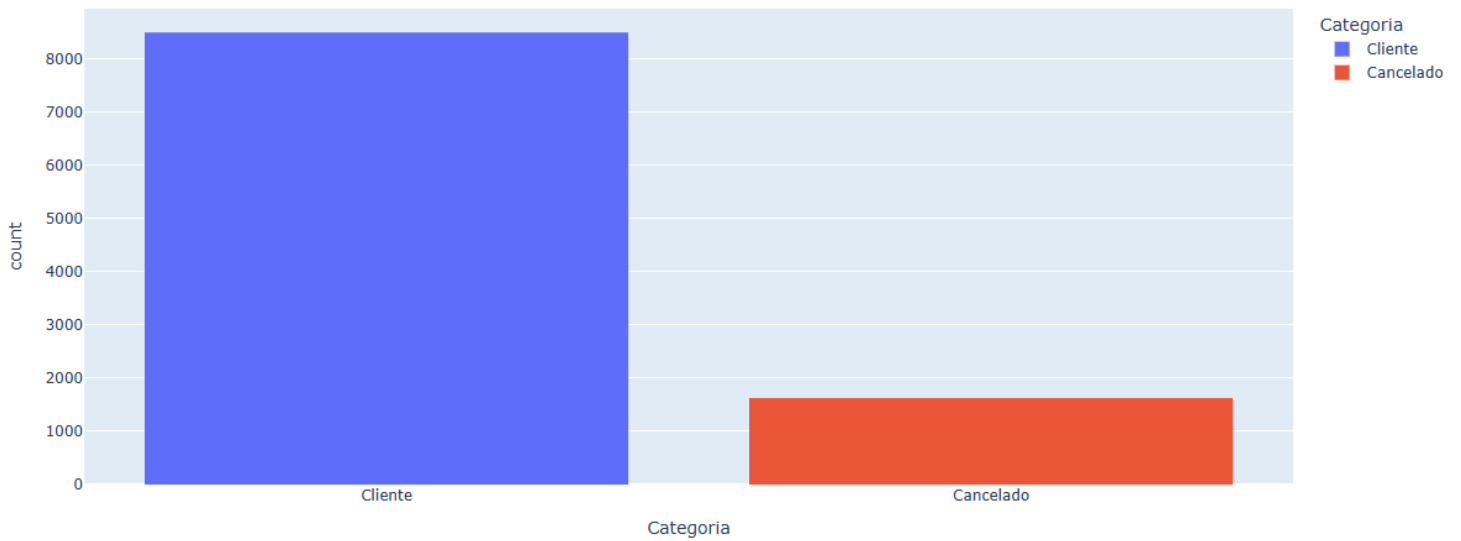
dtype: float64

Após essa análise comparativa mais minuciosa entre os percentuais de cada cenário, podemos analisar a comparação entre “Clientes” e “Cancelados” em cada uma das colunas da base de dados, gerando gráficos de Histograma e Dispersão. Para tal, podemos utilizar um looping *for*, usando como variável cada coluna da base de dados, e definindo o parâmetro *color* do gráfico como “Categoria”. E em seguida, utilizar o método *show* para visualizar os resultados obtidos, e continuar com a análise.

```
[ ] for coluna in tabela:
    grafico = px.histogram(tabela, x=coluna, color="Categoria")
    grafico.show()
```

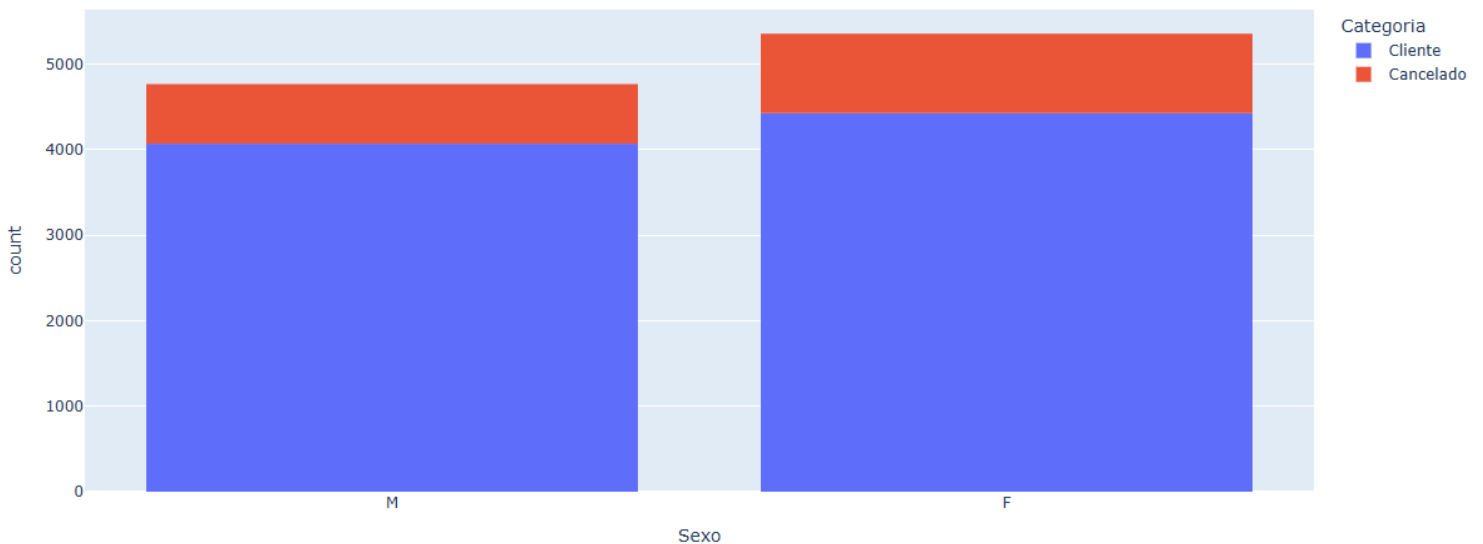
```
[ ] for coluna in tabela:
    grafico = px.scatter(tabela, x=coluna, color="Categoria")
    grafico.show()
```

Categoria



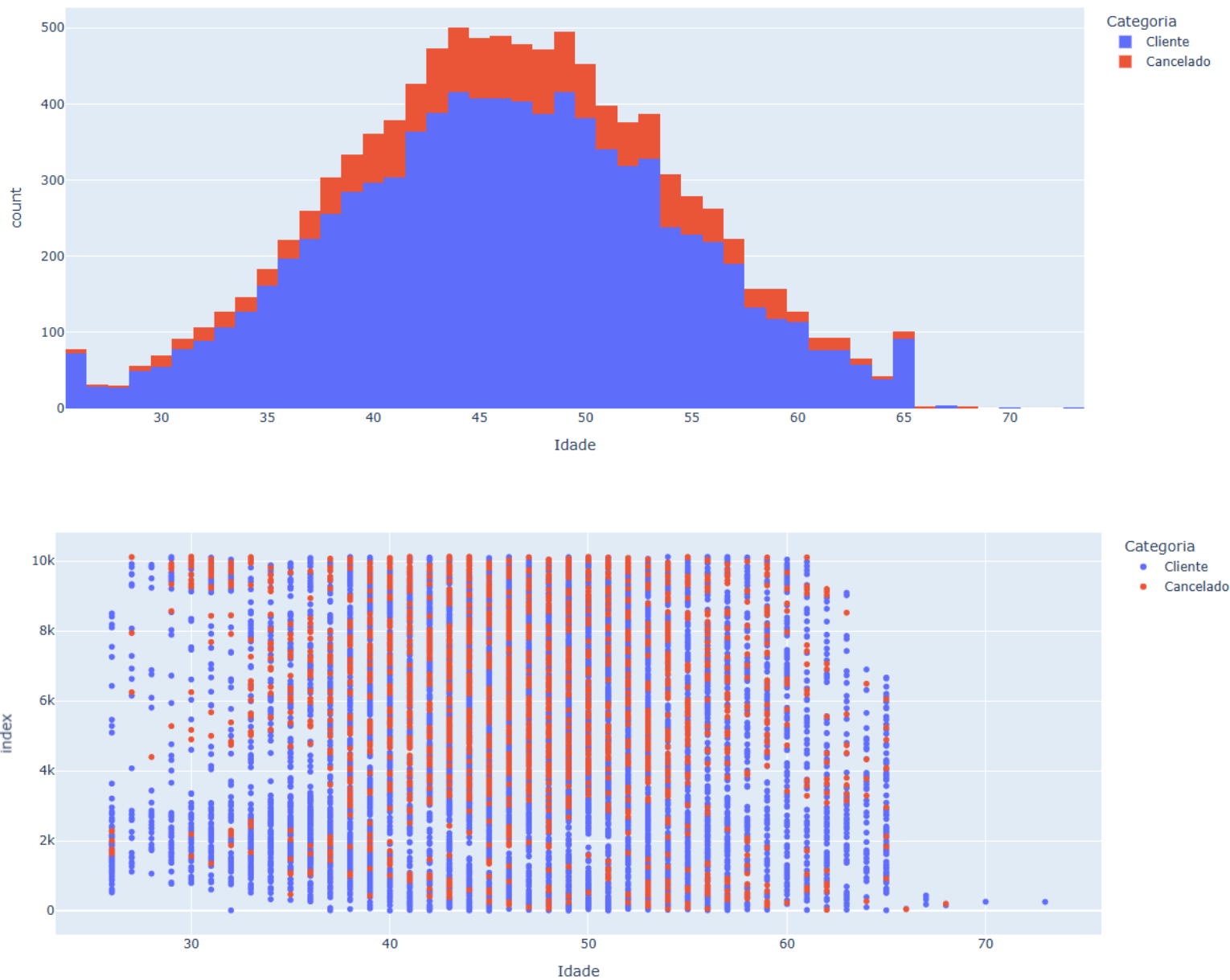
Nesse gráfico é possível visualizar a comparação entre usuários que ainda são clientes, e clientes que cancelaram o cartão de crédito. Podemos perceber que o percentual calculado anteriormente está correto, aproximadamente 16% cancelou o cartão, enquanto 84% ainda é cliente.

Sexo



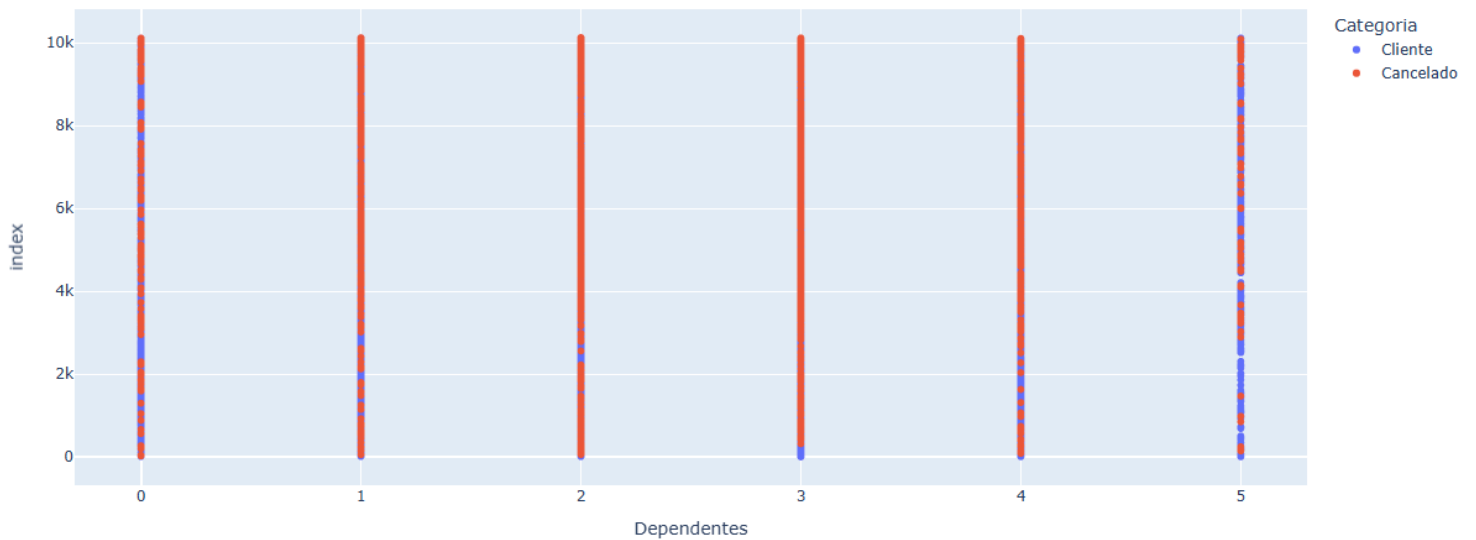
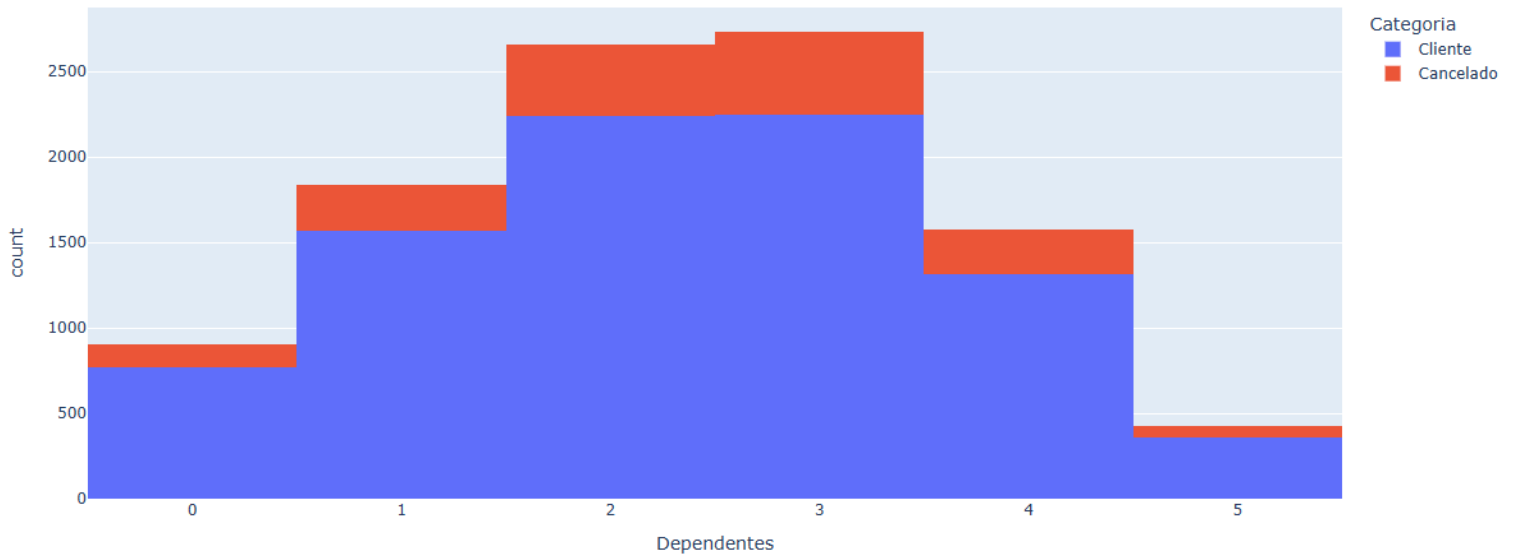
No gráfico que compara a taxa de cancelamento por sexo, podemos perceber que mais mulheres cancelaram o cartão de crédito. Visto que uma maior quantidade de clientes é formada por mulheres, podemos concluir que a taxa de cancelamento do cartão se mantém aproximadamente igual entre homens e mulheres. O sexo não parece aumentar a taxa de cancelamento do cartão.

Idade



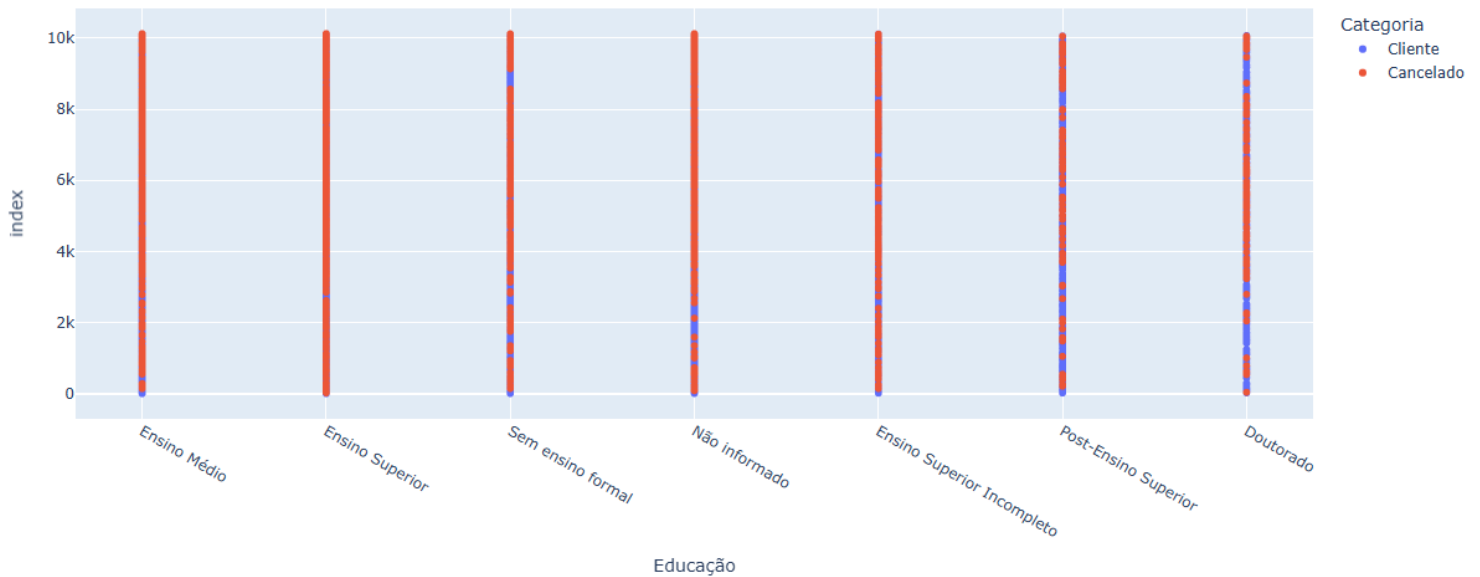
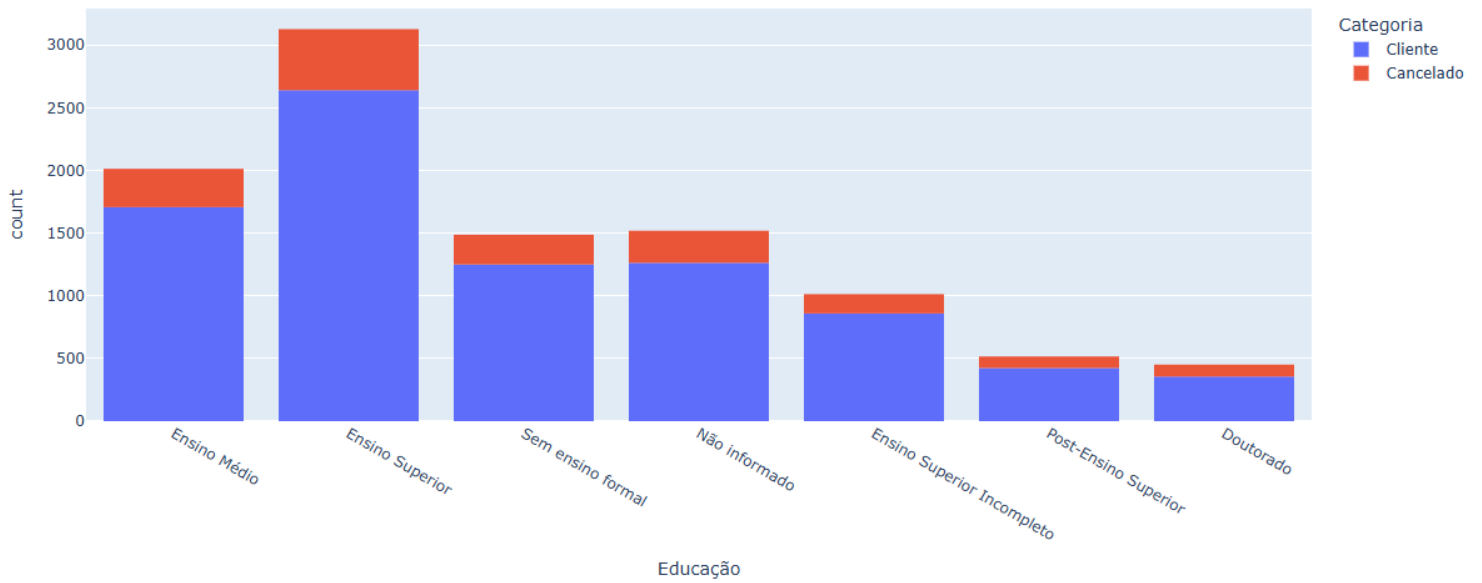
Analisando os gráficos de histograma e dispersão entre Clientes e Cancelados com base na idade, podemos concluir que a faixa etária na qual ocorrem mais cancelamentos é dos 40 aos 50 anos de idade.

Dependentes



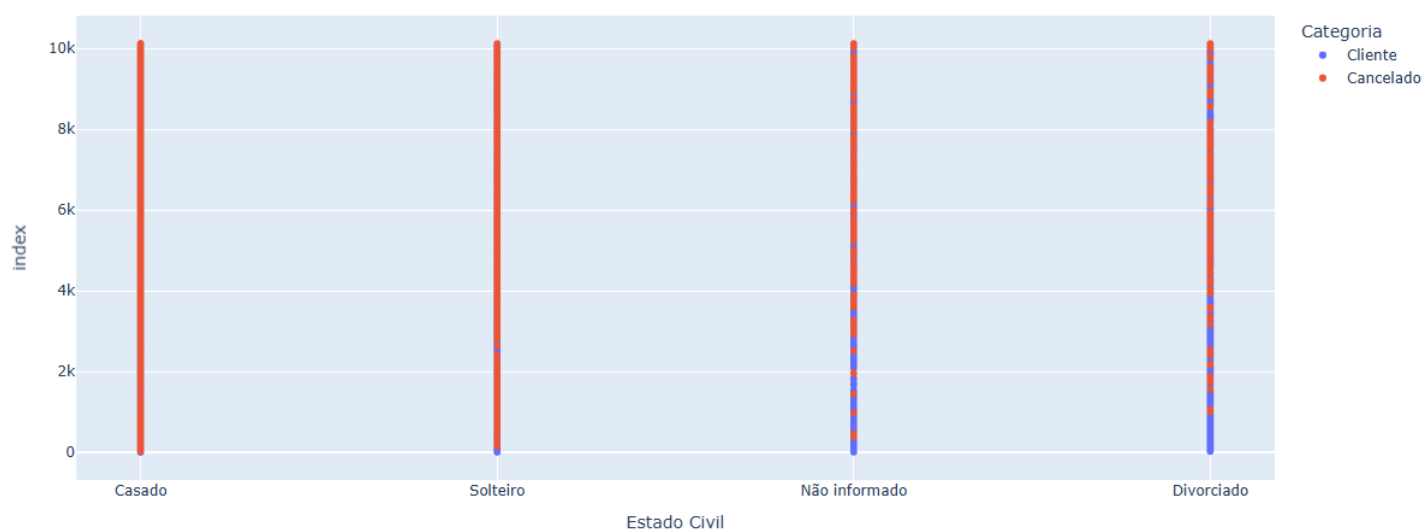
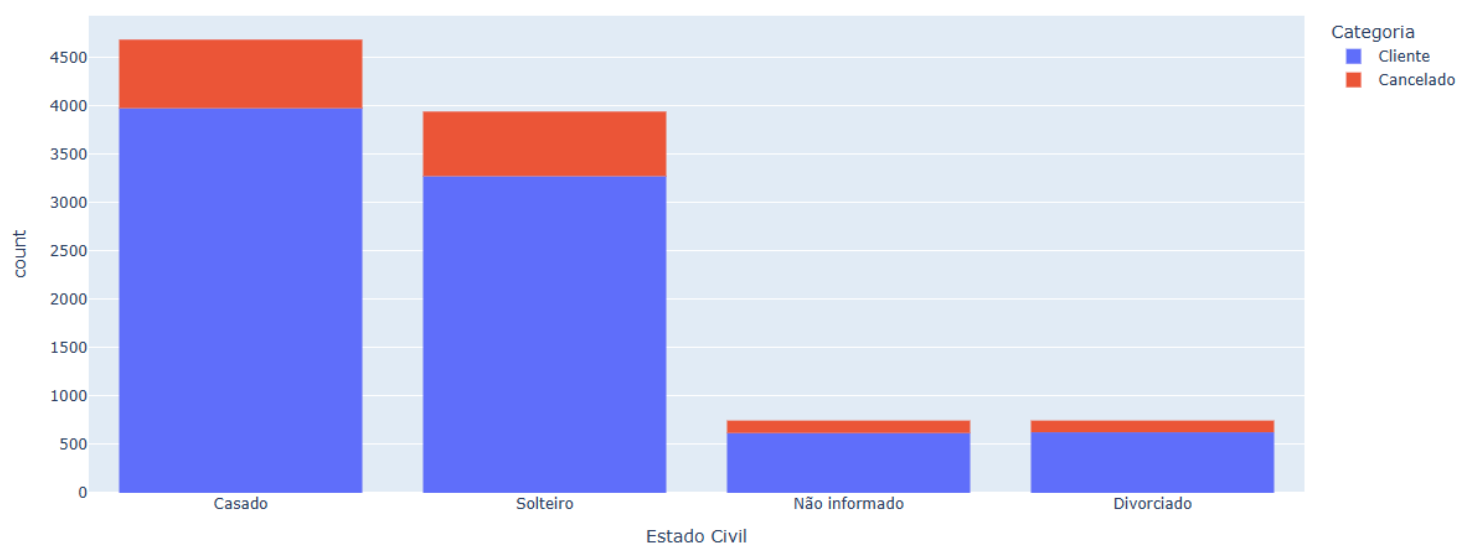
Analisando os gráficos de Clientes e Cancelados por Dependentes, concluímos que os clientes que possuem 2 ou 3 dependentes tiveram mais taxa de cancelamento do cartão de crédito. Logo depois, estão os clientes com 1 ou 4 dependentes, e por último, os clientes com 0 ou 5 dependentes.

Educação



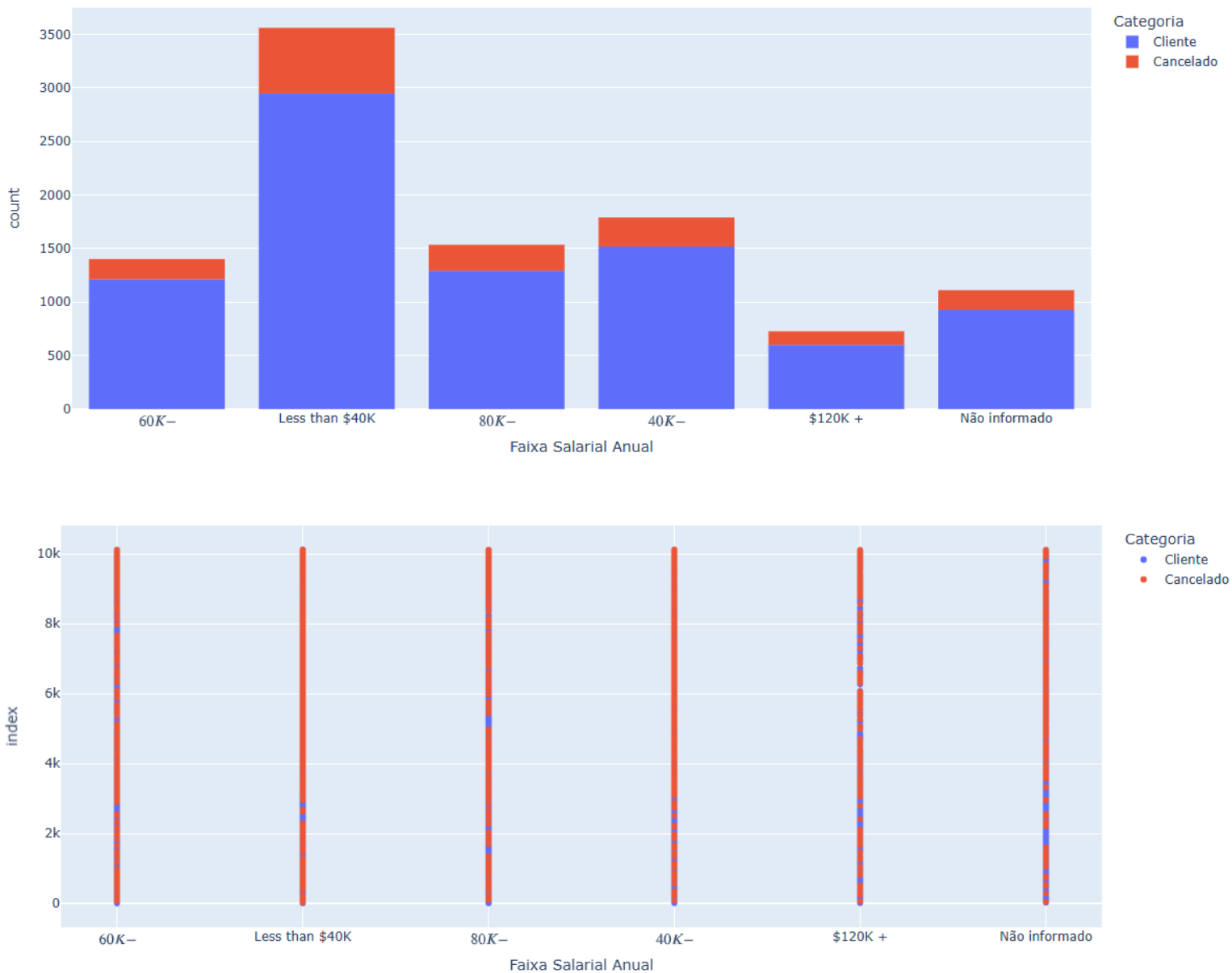
Analisando os gráficos de Clientes e Cancelados por Experiência Educacional, podemos concluir que usuários que possuem Ensino Superior formam a maior quantidade de clientes que cancelaram o cartão de crédito. Visto que a quantidade de clientes que possuem Ensino Superior é também maior do que as outras categorias, a média de desistentes levando em conta a quantidade de clientes em cada categoria não parece representar um valor significativo, não parecendo a experiência educacional interferir de forma relevante nesse quesito.

Estado Civil



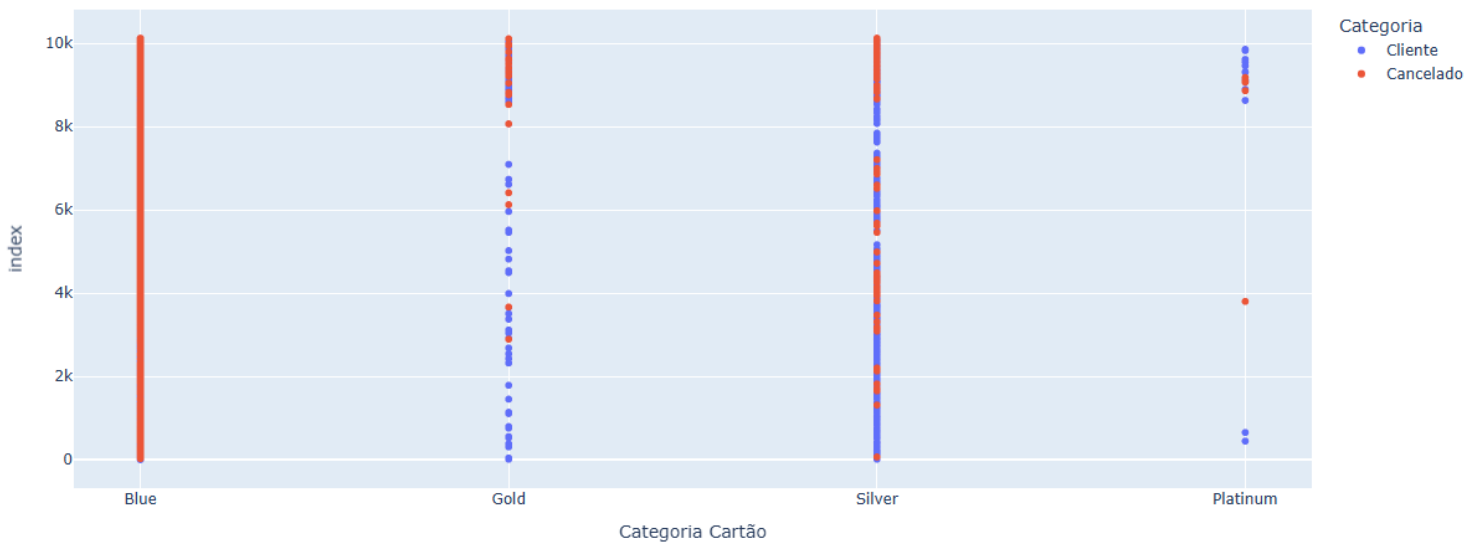
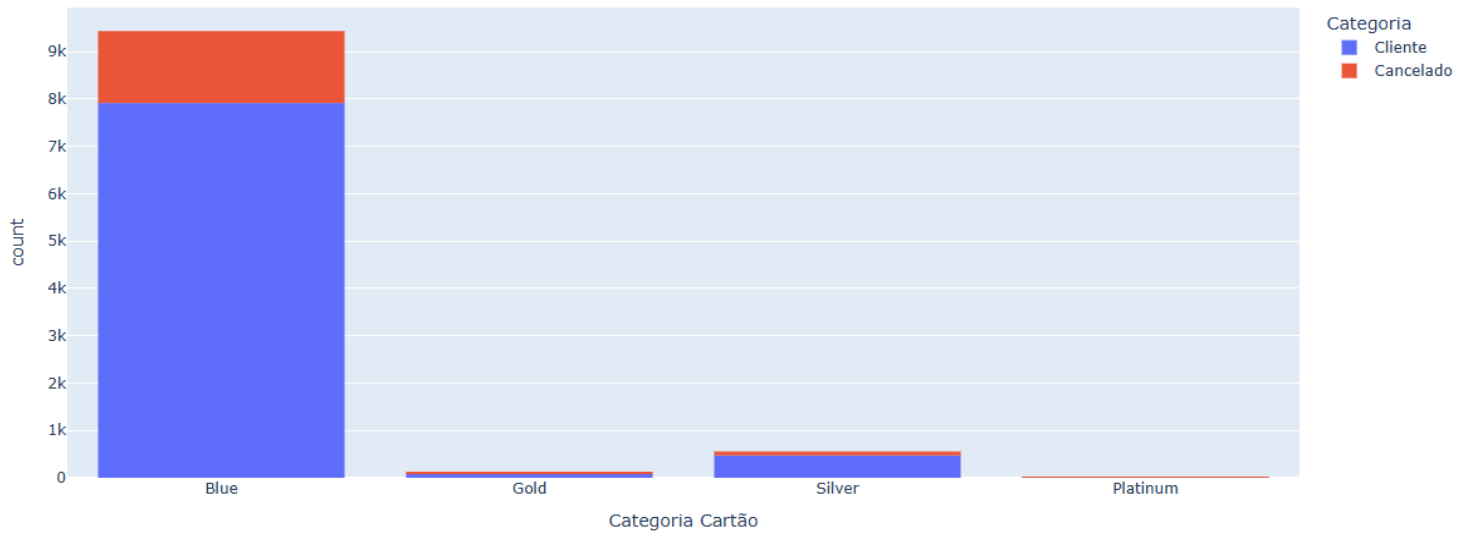
Analisando os gráficos por Estado Civil, concluímos que usuários casados ou solteiros representam a grande maioria dos cancelamentos de cartão de crédito. Visto que as outras duas categorias disponíveis não possuem muitos usuários cadastrados, essa informação também não parece ser significativa na avaliação do cenário de desistência dos cartões.

Faixa Salarial Anual



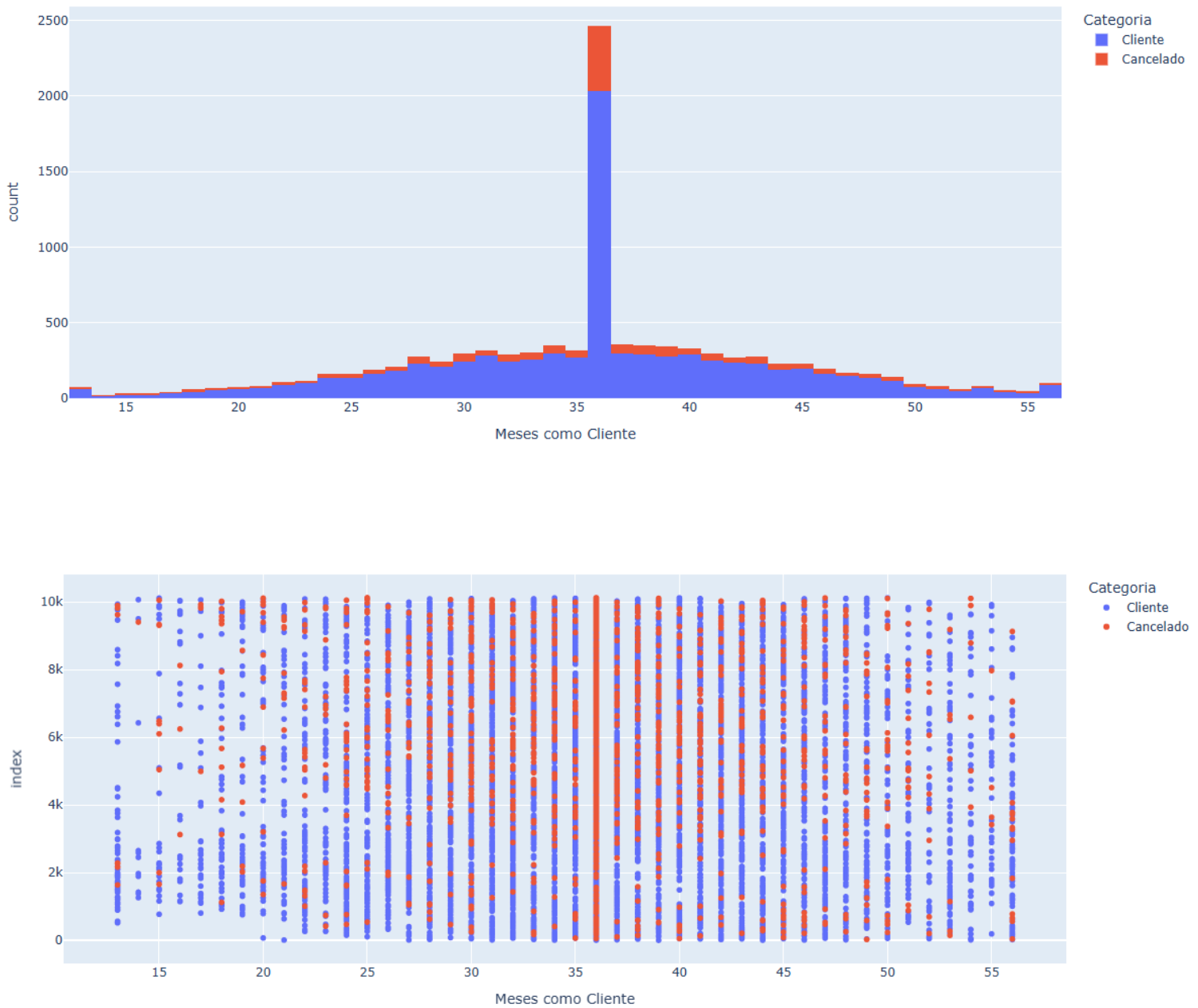
Analisando os gráficos de Clientes e Cancelados por Faixa Salarial Anual, podemos observar que clientes nos quais a faixa salarial anual é menor que 40 mil representam a maior taxa de cancelamentos. Isso pode ser relevante na avaliação dos possíveis motivos que levam os usuários a cancelarem o cartão, como por exemplo, o impacto potencial de taxas anuais e outras tarifas sobre seu orçamento. Para esses clientes, os benefícios do cartão (como milhas ou programas de recompensas) podem não compensar os custos, especialmente se o limite de crédito não for alto o suficiente para justificar o gasto com anuidade. Além disso, muitos desses clientes podem preferir evitar dívidas e focar no uso de cartões de débito ou dinheiro para ter maior controle financeiro e evitar juros elevados em casos de atraso no pagamento.

Categoria do Cartão

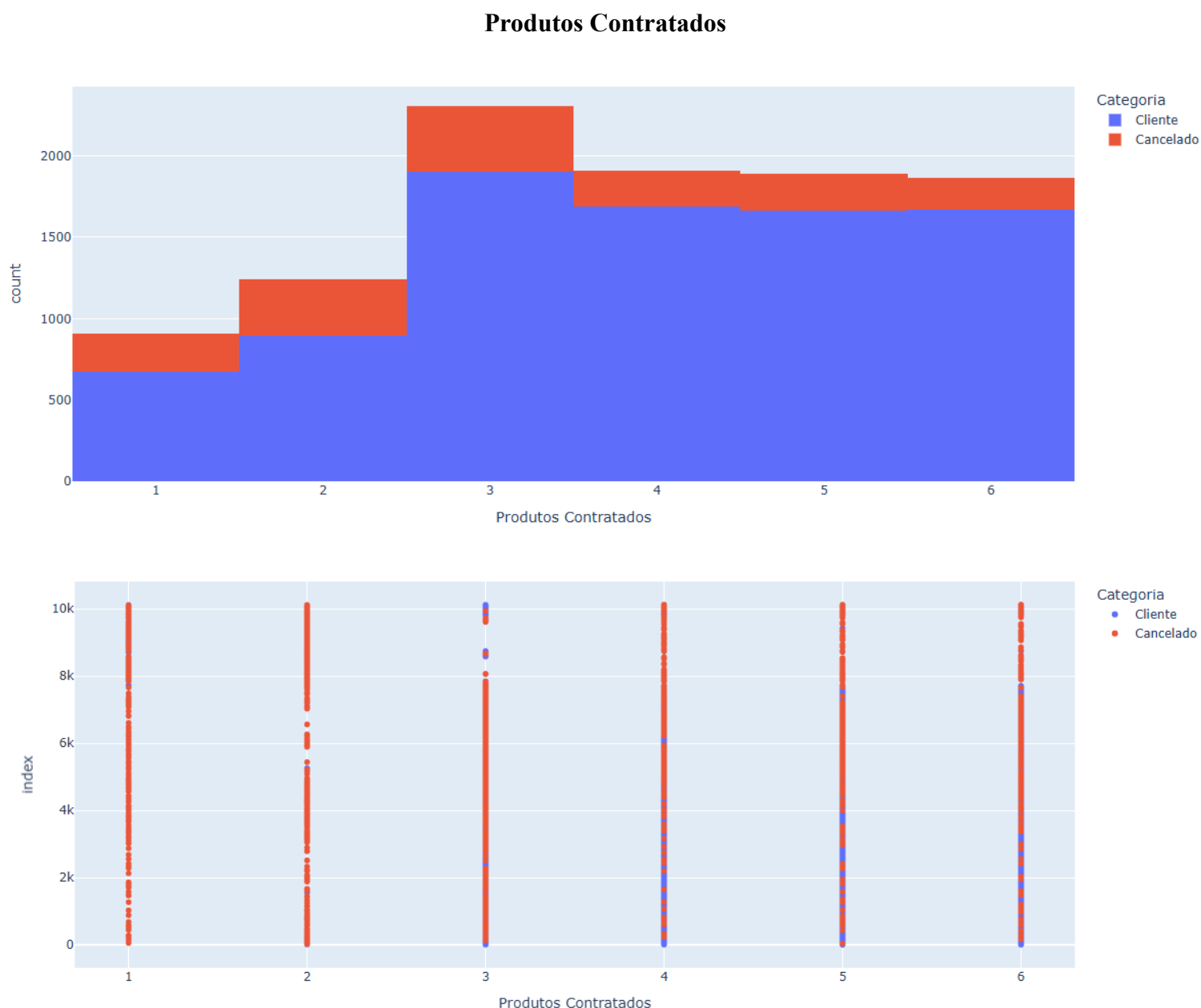


Analisando os gráficos com embasamento na categoria do cartão de crédito, concluímos que a maioria dos cancelamentos provém de clientes que possuíam o cartão Blue. Isso pode ser muito relevante na análise do motivo desses cancelamentos, sendo recomendado investigar mais especificamente sobre as condições que envolvem essa categoria de cartão em específico, taxas e juros, benefícios, etc.

Meses como Cliente

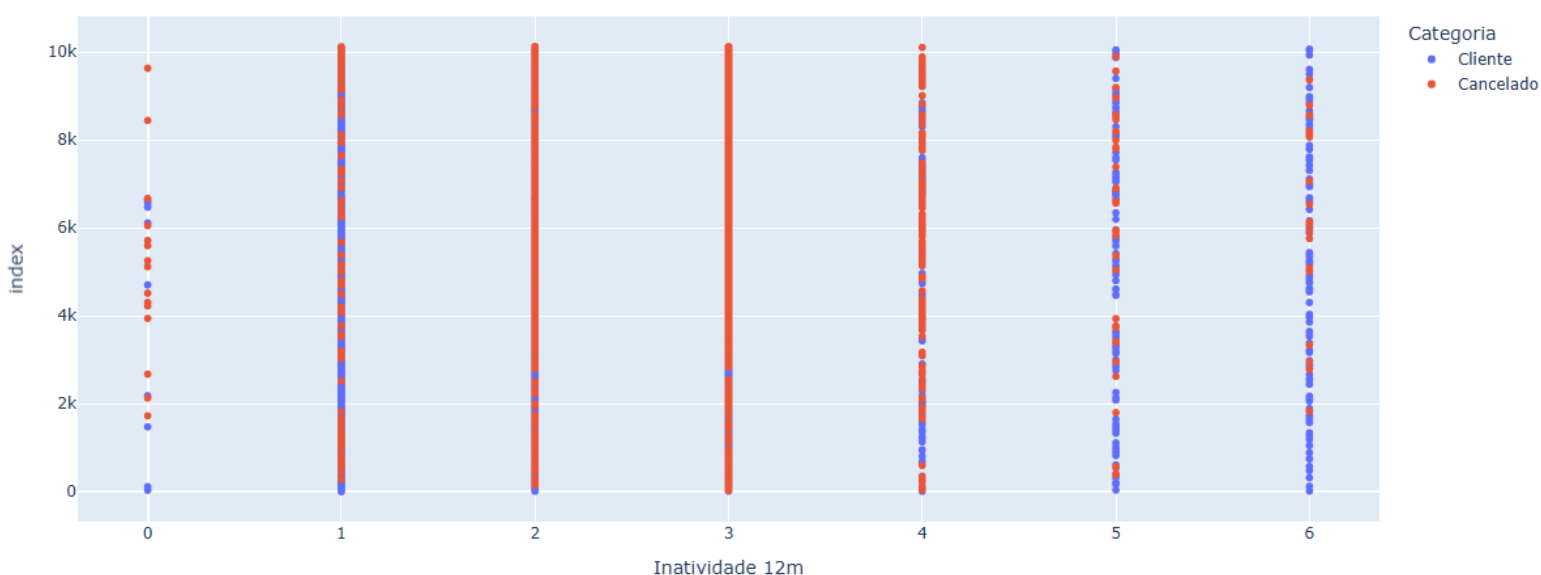
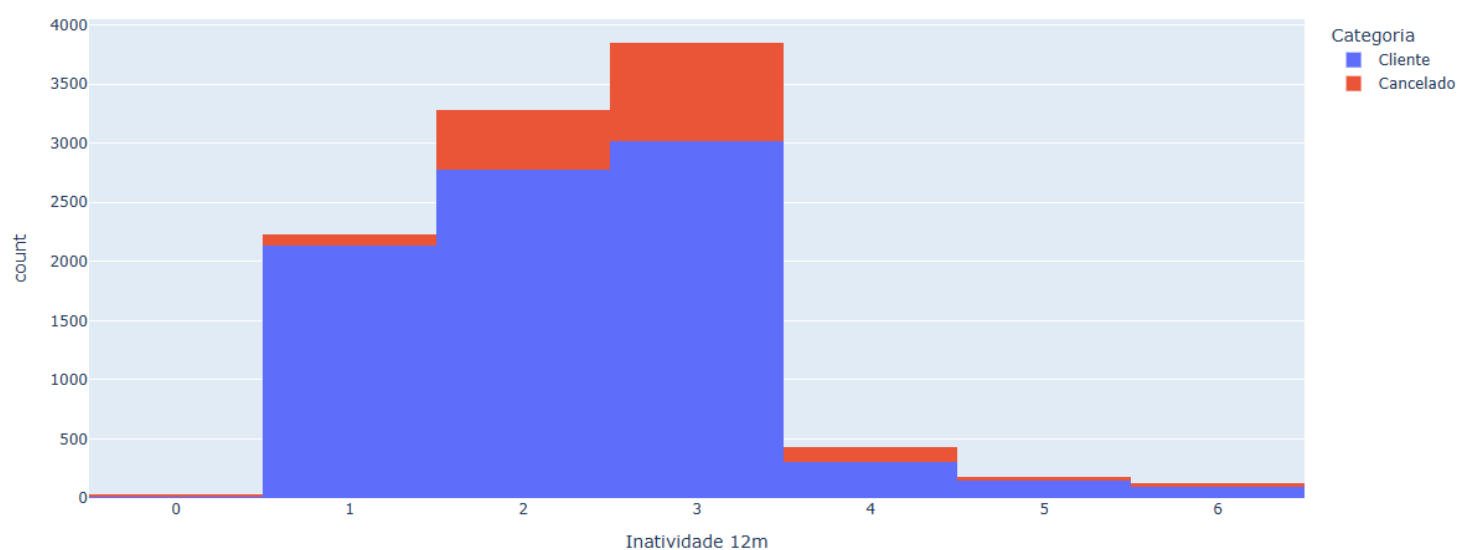


Analisando os gráficos baseando-se na quantidade de meses em que um usuário foi cliente, podemos concluir que a maioria foi cliente por aproximadamente 35 meses até cancelar o serviço, mas isso não parece ser relevante na análise, visto que a maioria dos clientes que não cancelaram o serviço também estão nesse mesmo patamar de tempo.



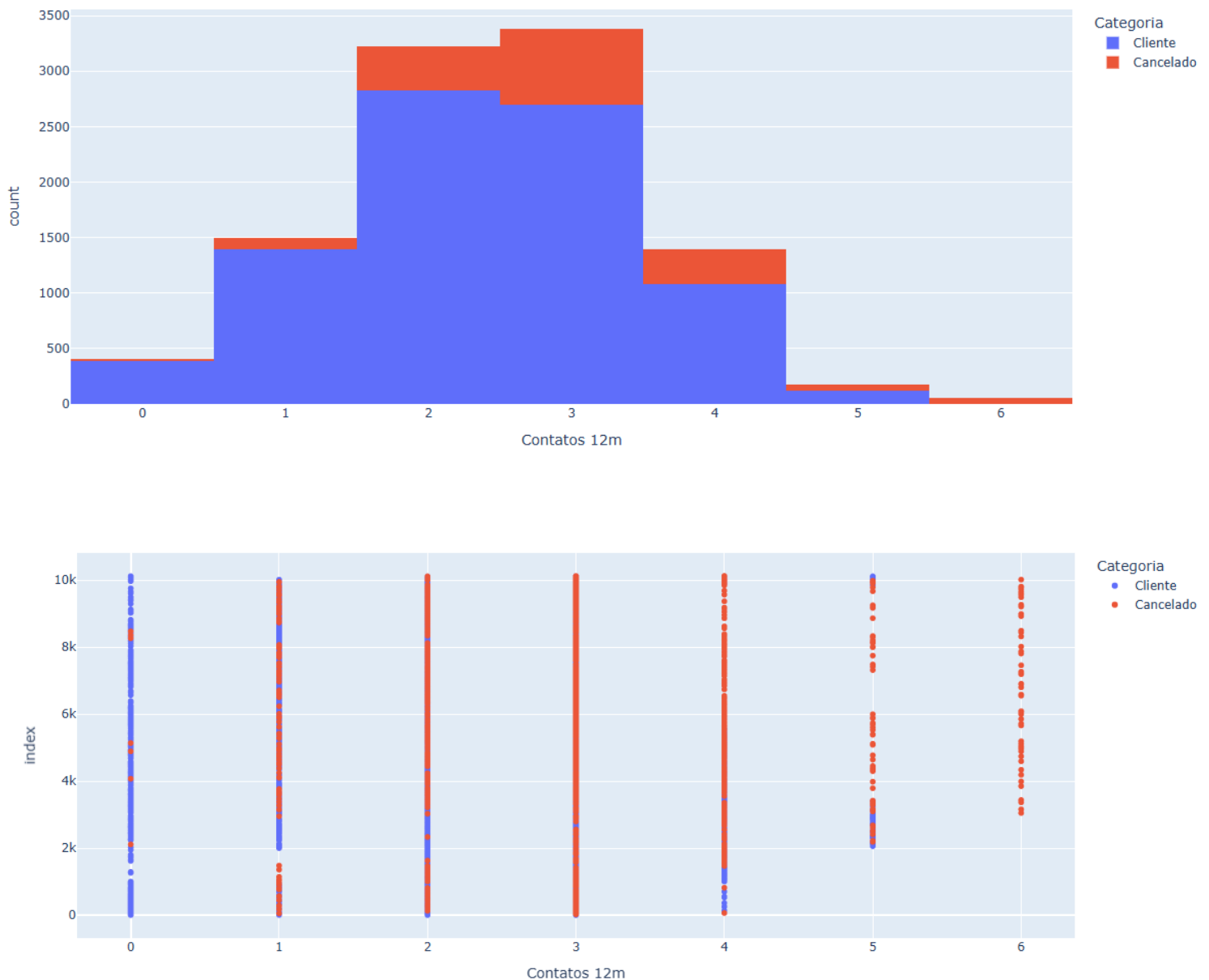
Analisando os gráficos com base na quantidade de produtos contratados pelos clientes, temos uma conclusão muito relevante e importante para a análise. Clientes que contratam mais produtos representam a menor taxa de cancelamentos de cartão de crédito. Isso pode ser devido a fatores como: clientes que possuem múltiplos serviços no banco, como conta corrente, investimentos, empréstimos ou seguros, tendem a cancelar menos seus cartões de crédito porque geralmente obtêm benefícios adicionais e tarifas reduzidas por manterem um "pacote" de serviços. Além disso, esses clientes têm uma relação mais integrada com o banco, o que facilita o gerenciamento financeiro em uma única instituição e cria um senso de fidelidade e conveniência. O cartão de crédito, nesse contexto, se torna mais vantajoso, com possíveis isenções de anuidade e recompensas que complementam os demais produtos bancários que eles já utilizam.

Inatividade em 12 meses



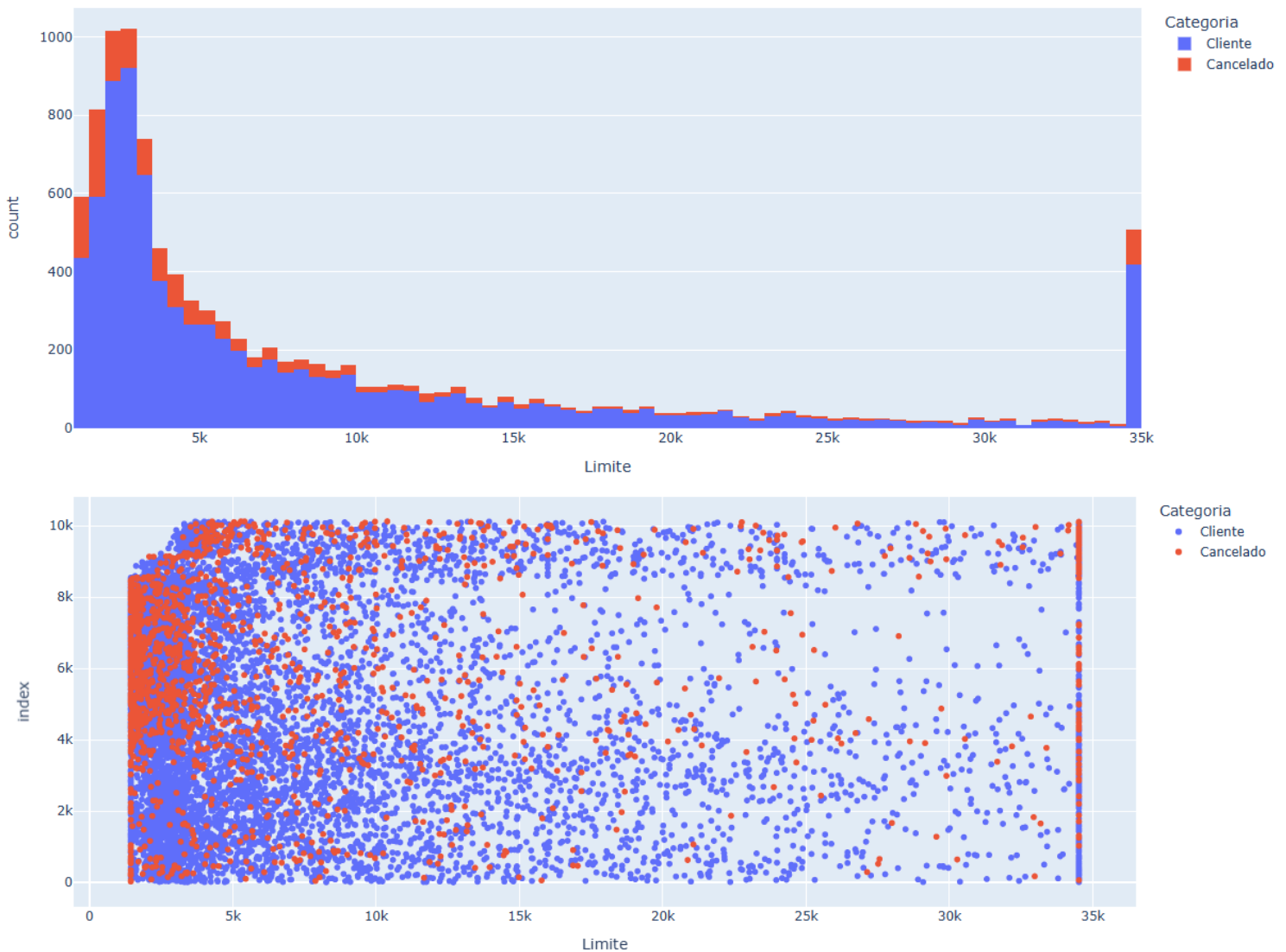
Cientes com inatividade no cartão por aproximadamente 2 a 3 meses representam a maior quantidade de desistentes do serviço. Isso pode se dar ao fato de que maioria de usuários que ainda continuam como clientes também pertencem à essa faixa temporal, mas também que em períodos mais curtos, a ausência de uso pode indicar que o cliente está reavaliando seu relacionamento com o cartão ou mesmo com a instituição financeira, especialmente em situações de mudanças recentes, como ajustes no orçamento ou novas escolhas de produtos financeiros. Períodos de inatividade mais longa, por sua vez, pode sugerir que o cliente ainda vê algum valor no cartão, mesmo que não o use regularmente. Ele pode estar aguardando um momento mais adequado para retomar o uso, como períodos de maior gasto (festas de fim de ano ou férias). Além disso, esses clientes podem estar mantendo o cartão como uma "reserva" para futuras necessidades, o que reduz a probabilidade de cancelamento imediato.

Contatos em 12 meses



Analisando os gráficos pela quantidade de contatos de clientes com o banco no período de um ano, outra conclusão relevante para a análise é gerada. Clientes que necessitaram mais vezes do suporte, que nos leva a concluir que tiveram mais problemas ou dúvidas, formam a maior quantidade de desistentes. Isso pode significar que algum serviço oferecido pelo cartão pode estar gerando muitos problemas, ou que usuários que tiveram dúvidas e necessitaram de ajuda/suporte não foram bem atendidos ou não tiveram seus problemas solucionados com eficiência por parte do banco.

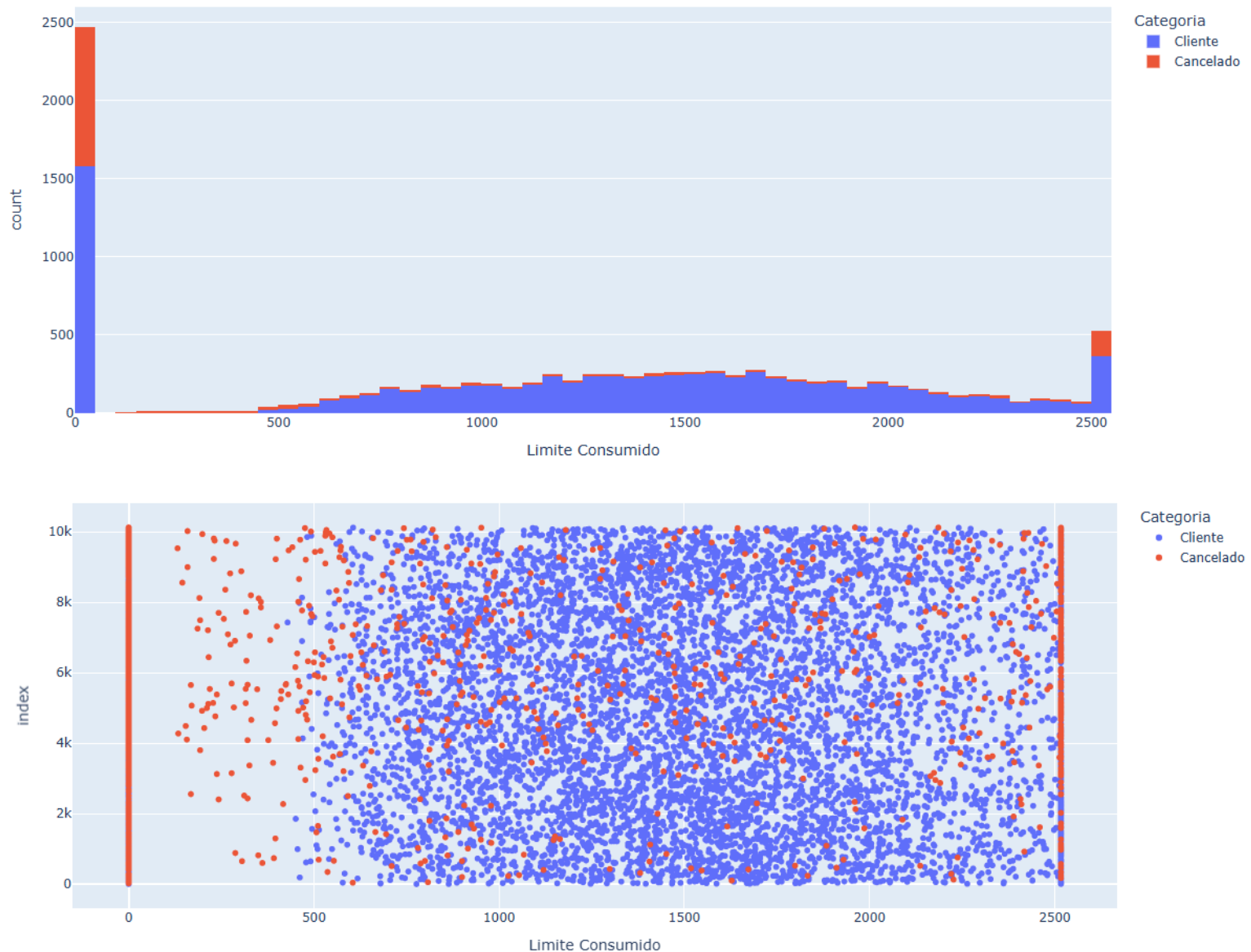
Limite



Ao analisar os gráficos baseados no cancelamento de cartão relacionados ao limite de crédito, podemos concluir que clientes que possuem cartões de crédito com valores baixos ou altos tendem a cancelar mais o cartão do que usuários com o limite mais próximo de um valor médio, o que é um fator muito relevante. Isso pode ocorrer devido a usuários com limite baixo estarem insatisfeitos porque o limite muitas vezes não atende a suas necessidades, tornando o cartão menos útil em situações em que eles precisam de crédito adicional. Eles podem sentir que o cartão não agrega valor e procurar outras opções com limites mais altos ou melhores condições. Além disso, eles podem considerar que taxas, como anuidades, não valem a pena em um cartão de uso limitado. Usuários com limites mais altos, em contrapartida, podem cancelar o cartão por questões de custo-benefício. Esse grupo pode ter múltiplos cartões e, se percebem que o custo (anuidades, taxas) não compensa os benefícios oferecidos, podem optar por cancelar. Além disso, eles podem considerar que ter um limite muito alto desnecessariamente pode impactar o perfil de crédito ou aumentar o risco de endividamento. Os usuários de limite médio, por outro lado, podem encontrar um equilíbrio entre o valor do limite e o custo do cartão. Esse grupo geralmente vê o limite como suficiente para suas necessidades, evitando a insatisfação que pode levar ao cancelamento. Além disso,

muitos cartões com limite médio oferecem benefícios atrativos e anuidades mais acessíveis, o que contribui para que eles mantenham o cartão. Essas tendências mostram como o valor do limite pode influenciar a percepção de custo-benefício do usuário e, consequentemente, as taxas de cancelamento em cada faixa.

Limite Consumido



Quando falamos de limite consumido (ou seja, o quanto o usuário efetivamente utiliza do seu limite), as razões para cancelar o cartão também variam com base em como esse limite é utilizado em cada faixa. Usuários que consomem quase todo o limite frequentemente podem cancelar o cartão por algumas razões:

- Juros e Endividamento: Com o limite quase todo utilizado, eles podem acabar pagando mais em juros rotativos, especialmente se não conseguem quitar a fatura

integralmente. Isso pode gerar uma insatisfação com o cartão e um desejo de cortar despesas.

- **Estresse Financeiro:** Estar próximo do limite pode ser financeiramente desgastante, aumentando o risco de inadimplência e trazendo impacto negativo no crédito. Eles podem preferir cancelar o cartão para evitar essa pressão.
- **Busca por Alternativas:** Esses usuários podem optar por um cartão com condições mais favoráveis, como juros menores ou programas de recompensas que justifiquem os gastos elevados.

Usuários que não consomem boa parte do limite podem optar pelo cancelamento devido a:

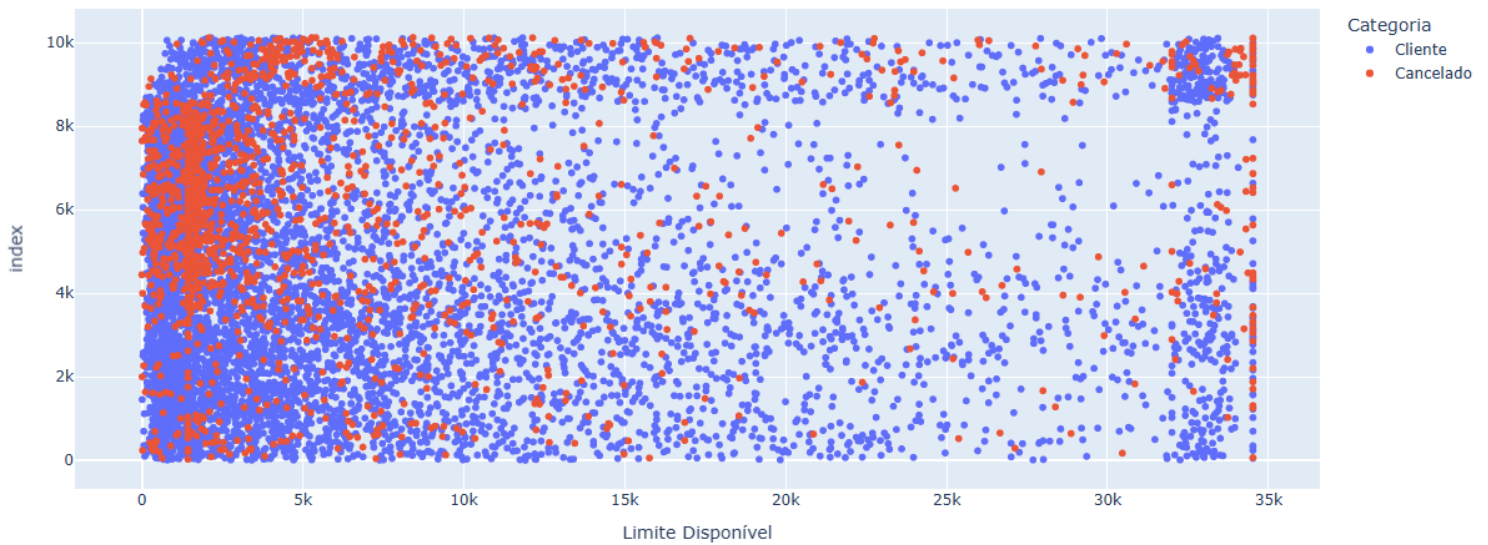
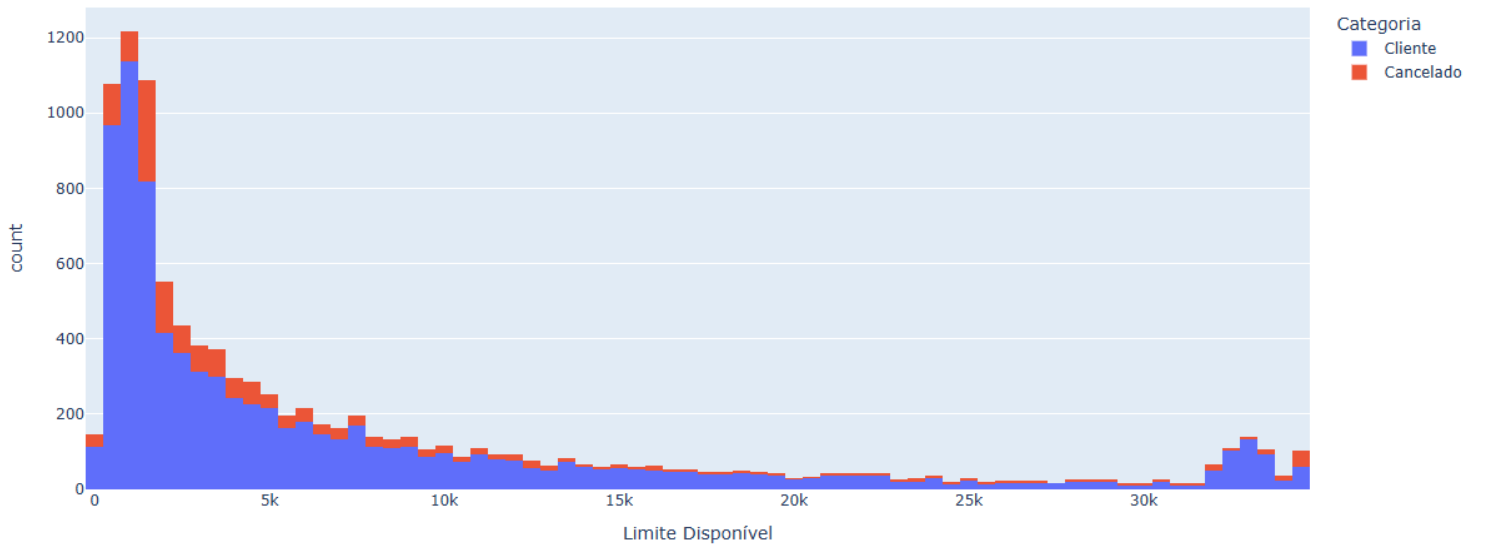
- **Baixa Utilidade:** Se o cartão é pouco usado, eles podem sentir que ele não oferece benefícios significativos. Esses usuários podem preferir cancelar o cartão para evitar anuidades e taxas, especialmente se têm outros cartões que suprem suas necessidades.
- **Custo-Benefício Desfavorável:** Muitos cartões cobram anuidades ou oferecem benefícios que são percebidos como úteis apenas se o cartão for utilizado com frequência. Se o limite não é utilizado, esses usuários podem considerar o custo desnecessário.

Para aqueles que consomem uma parte moderada do limite, o cartão geralmente encontra um equilíbrio entre utilidade e custo-benefício:

- **Satisfação com Benefícios e Custo:** Esse grupo vê valor no cartão, pois usa o limite de forma equilibrada, sem pressionar o crédito disponível. O uso moderado muitas vezes mantém o custo das taxas de juros e anuidades justificável.
- **Menor Risco de Endividamento:** Como o limite não é consumido em excesso, esses usuários evitam a pressão de juros rotativos, o que reduz o risco de inadimplência e aumenta a satisfação com o cartão.

Assim, o limite consumido influencia a percepção do valor do cartão e a probabilidade de cancelamento: usuários que estão nas extremidades de uso (muito alto ou muito baixo) tendem a cancelar mais, enquanto os de uso moderado frequentemente encontram um equilíbrio mais sustentável.

Limite Disponível



Quando se trata do limite disponível (o quanto resta do limite total), as razões para cancelar o cartão também podem variar conforme a faixa desse limite. Para quem tem pouco limite disponível, o cancelamento pode ocorrer devido a:

- **Dificuldade de Uso do Cartão:** Com o limite já bastante comprometido, esses usuários podem encontrar restrições para realizar novas compras. Isso pode tornar o cartão menos útil e, ao invés de buscar um aumento de limite, alguns optam pelo cancelamento.

- Custos com Juros: Quando o limite disponível é baixo, é mais provável que estejam pagando juros rotativos, especialmente se não conseguem liquidar o saldo integral. Esses custos aumentam o descontentamento e podem levar ao cancelamento.
- Alternativas com Limites Mais Altos: Em alguns casos, esses usuários preferem buscar cartões com limites maiores que atendam melhor suas necessidades, especialmente se o custo-benefício do atual não é satisfatório.

Quem tem um limite disponível alto, ou seja, usa o cartão muito pouco em relação ao que poderia gastar, também pode optar pelo cancelamento por motivos como:

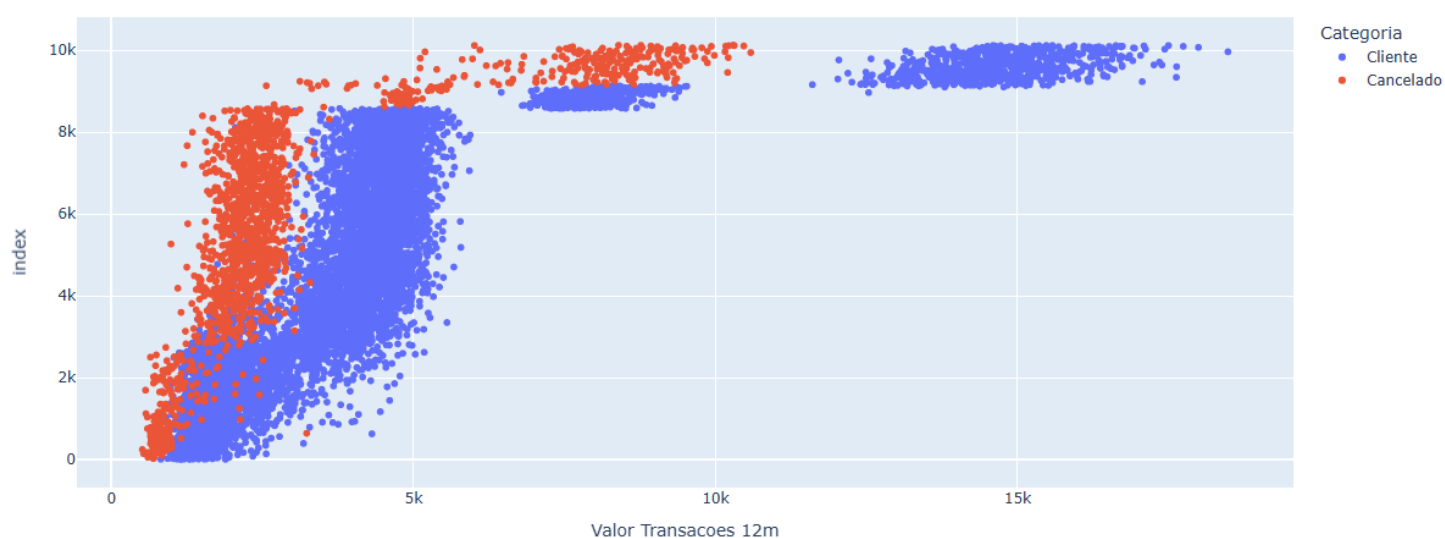
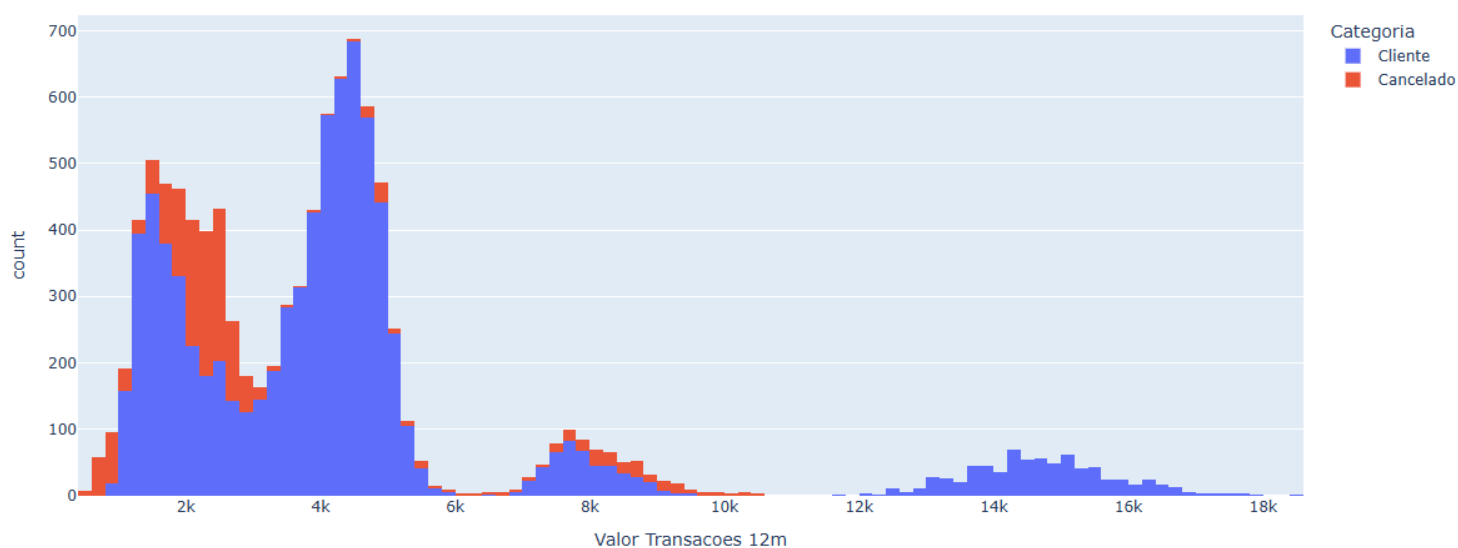
- Percepção de Baixa Necessidade: Se o usuário utiliza o cartão com pouca frequência, ele pode sentir que o limite alto é desnecessário. Esse cartão pode ser considerado um custo sem utilidade, especialmente se ele precisa pagar anuidade ou outras taxas.
- Busca por Benefícios Melhores: Usuários com muito limite disponível às vezes percebem que poderiam ter um cartão com benefícios mais alinhados ao uso esporádico, como isenção de anuidade ou cashback em vez de manter um cartão com limite elevado.
- Redução de Risco Financeiro: Ter um limite alto disponível pode ser percebido como um risco potencial de endividamento. Esses usuários podem preferir cancelar o cartão para reduzir o crédito disponível e evitar gastos impulsivos.

Para quem mantém um limite disponível moderado, o cancelamento é menos provável, porque:

- Equilíbrio entre Necessidade e Custo: Esse grupo geralmente usa o cartão de maneira que aproveita o crédito disponível sem ultrapassar os limites de controle financeiro. Com isso, os custos de manutenção do cartão são justificados, e o limite restante é suficiente para novas compras sem gerar estresse financeiro.
- Satisfação com Flexibilidade: Manter um limite disponível razoável oferece flexibilidade para emergências e compras, criando uma percepção positiva do cartão. Esse equilíbrio contribui para uma experiência satisfatória e reduz a necessidade de buscar outros cartões.

Assim, o limite disponível impacta a probabilidade de cancelamento com usuários que têm muito ou pouco limite disponível sendo mais propensos a cancelar, enquanto aqueles que mantêm um equilíbrio de uso tendem a ver o cartão como mais útil e, portanto, menos provável de cancelar.

Valor das transações em 12 meses



Quando analisamos o valor das transações feitas com o cartão de crédito, os padrões de cancelamento entre usuários de diferentes faixas de gasto refletem suas necessidades e expectativas de benefícios. Usuários que fazem transações de valores baixos podem ser mais propensos a cancelar o cartão por alguns motivos:

- **Percepção de Baixo Benefício:** Muitos benefícios de cartões de crédito (milhas, cashback, programas de pontos) são mais atrativos para usuários com maior volume de gastos. Aqueles que fazem transações menores podem sentir que os benefícios oferecidos não justificam taxas ou anuidades, especialmente se o retorno em pontos ou cashback for baixo.

- Alternativas de Pagamento: Para transações pequenas, esses usuários podem preferir outras formas de pagamento, como cartões de débito ou contas digitais, que não envolvem custos adicionais. Isso pode fazer com que o cartão de crédito pareça dispensável, levando ao cancelamento.
- Sensibilidade a Taxas: Mesmo que esses clientes usem o cartão com frequência, os valores baixos das transações não trazem grandes retornos. Portanto, o custo de manutenção, como a anuidade, pode ser percebido como um fardo desproporcional ao valor que eles extraem do cartão.

Usuários que fazem transações de valores médios também podem apresentar uma tendência maior ao cancelamento:

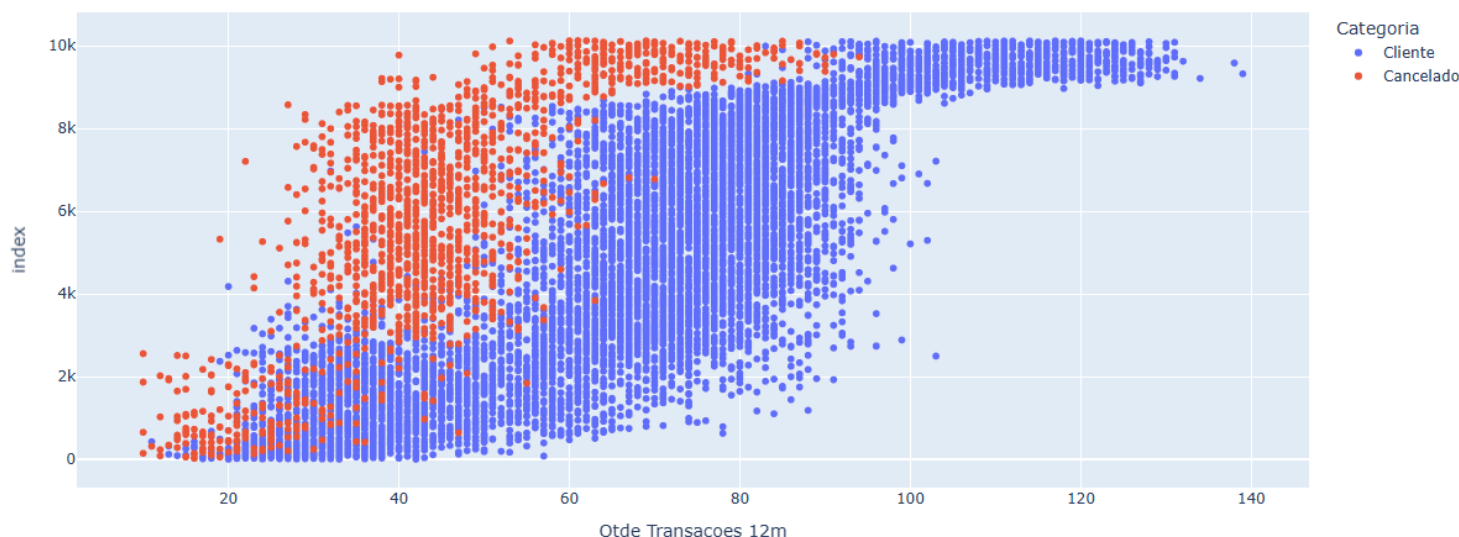
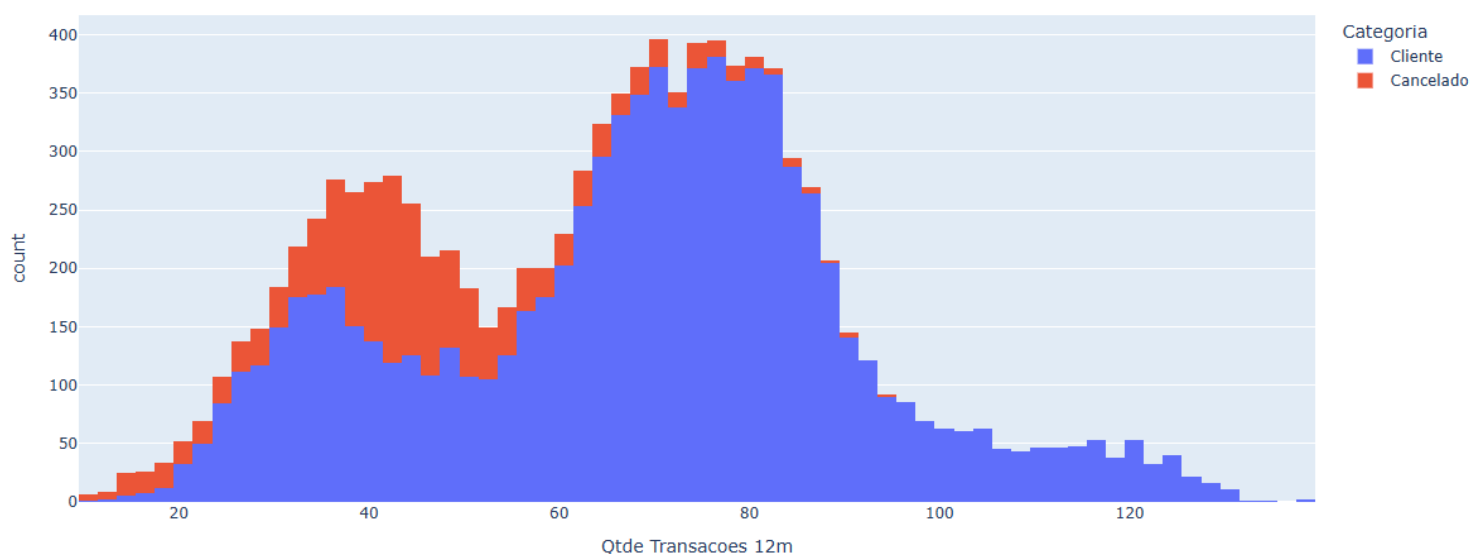
- Insatisfação com Benefícios Limitados: Os usuários de transações de valor médio estão em uma posição intermediária em que seus gastos podem não ser suficientes para aproveitar ao máximo os programas de recompensas, mas ainda assim eles pagam taxas e anuidades. Esse grupo pode acabar cancelando o cartão se perceber que outros cartões oferecem mais benefícios por menos custos.
- Busca por Alternativas com Melhor Custo-Benefício: Clientes com transações de valores médios podem avaliar que conseguem mais valor com cartões de débito ou outros produtos financeiros que não impliquem em taxas anuais ou juros rotativos. Alternativas mais baratas e com benefícios voltados para o perfil de uso médio podem parecer mais vantajosas.

Clientes com transações de valores elevados tendem a cancelar menos, por algumas razões principais:

- Maximização de Benefícios: Para esses usuários, os programas de recompensas, cashback e acúmulo de pontos são geralmente mais vantajosos, pois o volume elevado de transações gera retornos significativos. Isso torna o custo do cartão mais fácil de justificar, criando uma relação custo-benefício favorável.
- Benefícios Adicionais Exclusivos: Muitos cartões de crédito com benefícios premium (como seguros de viagem, e atendimento personalizado) são projetados para quem faz transações altas. Esses usuários veem valor agregado nas vantagens exclusivas que justificam a anuidade e outros custos.
- Satisfação e Conveniência: Como as vantagens e facilidades atendem bem a suas necessidades, esses clientes geralmente estão satisfeitos com o cartão. O volume de transações permite a eles utilizar o cartão de forma conveniente e contínua, reduzindo a necessidade de buscar outras opções.

Em resumo, clientes que fazem transações de valores baixos e médios podem perceber que os benefícios não compensam os custos, o que os leva a cancelar o cartão com mais frequência. Já os usuários que fazem transações elevadas tendem a aproveitar melhor os programas de benefícios, sentindo maior valor na manutenção do cartão.

Quantidade de transações em 12 meses



Ao analisar os gráficos baseados na quantidade de transações dos usuários no período de um ano, algumas conclusões significativas podem ser elaboradas. A quantidade de transações com cartão de crédito influencia diretamente a percepção de valor e o custo-benefício, o que pode impactar a taxa de cancelamento. Em geral, usuários com poucas ou médias transações podem ser mais propensos ao cancelamento do cartão do que aqueles que fazem muitas transações. Usuários que realizam poucas transações no cartão podem cancelar o cartão com mais frequência devido a:

- **Percepção de Utilidade Limitada:** Com poucas transações, o cartão acaba sendo pouco utilizado, levando esses usuários a considerarem o cartão desnecessário. A falta de frequência no uso pode gerar uma percepção de que o cartão é dispensável.
- **Custo Desproporcional:** Cartões de crédito com anuidades ou taxas mensais são menos atraentes para quem usa o cartão esporadicamente, já que o custo fixo pode

parecer desproporcional ao uso real. Esses usuários tendem a preferir alternativas sem anuidades ou cartões de débito.

- **Foco em Outros Métodos de Pagamento:** Clientes com poucas transações podem preferir pagar com outros métodos, como dinheiro, débito ou carteiras digitais, especialmente se esses métodos forem mais convenientes ou se não há incentivo forte para usar o cartão de crédito.

Clientes com um volume médio de transações também podem cancelar o cartão com mais frequência por motivos como:

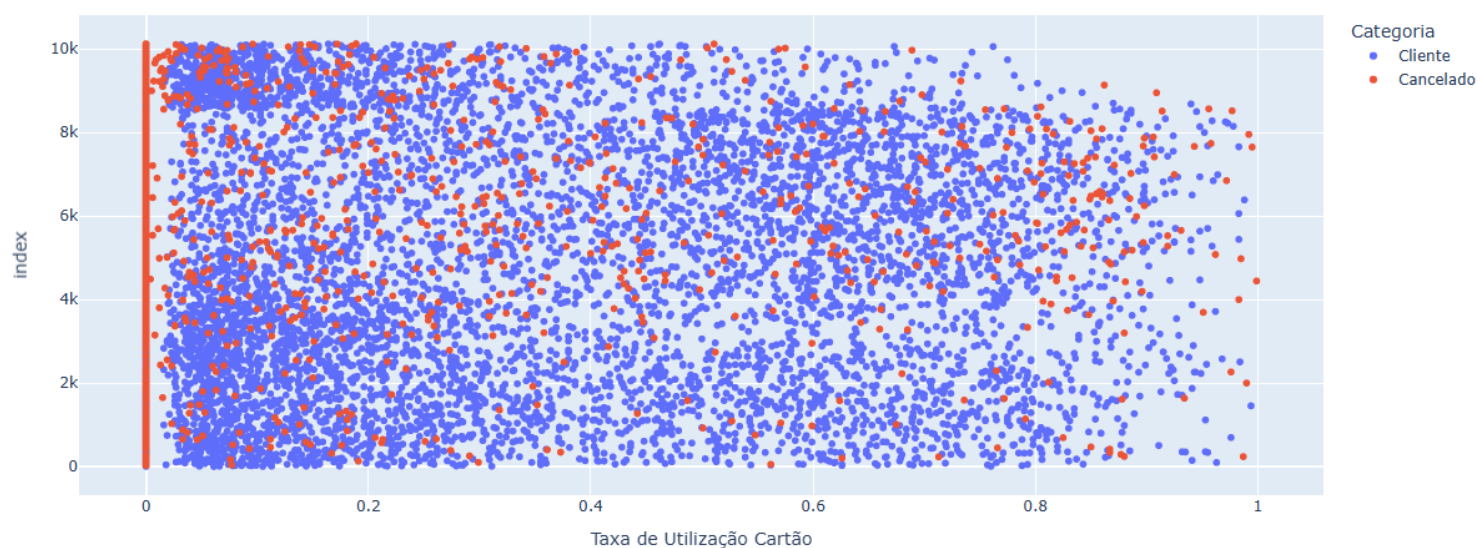
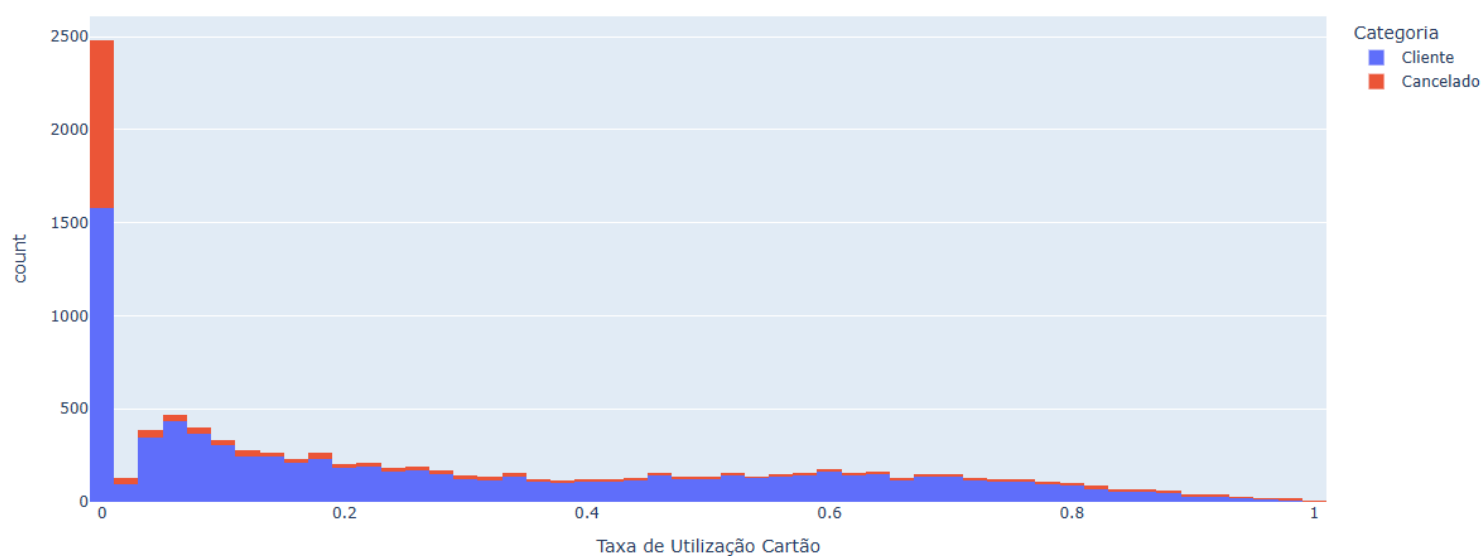
- **Expectativas Não Atendidas de Benefícios:** Quem usa o cartão regularmente, mas com quantidade moderada de transações, pode perceber que o retorno em benefícios (como pontos e cashback) é inferior ao esperado. O uso moderado muitas vezes não é suficiente para compensar a anuidade, reduzindo a satisfação com o cartão.
- **Busca por Alternativas:** Esses clientes podem se sentir mais inclinados a buscar outros cartões com melhores condições ou mais alinhados ao seu perfil de consumo, como cartões sem anuidade ou com programas de recompensas mais vantajosos para transações de menor frequência.
- **Custos Menores com Outros Produtos Financeiros:** Muitos usuários com transações moderadas percebem que a manutenção de um cartão de crédito pode ser desnecessária, e preferem canais financeiros que evitem dívidas e taxas fixas.

Clientes com alta quantidade de transações são geralmente menos propensos a cancelar o cartão, pois:

- **Aproveitamento Máximo de Benefícios:** Com uma alta frequência de uso, esses usuários conseguem maximizar os benefícios oferecidos, como programas de pontos, milhas ou cashback. A quantidade de transações resulta em recompensas significativas, justificando o custo do cartão.
- **Facilidade e Conforto no Uso:** Usuários que fazem transações frequentes tendem a preferir a praticidade do crédito para simplificar compras recorrentes e concentrar os gastos em um único meio de pagamento, o que facilita o controle financeiro e aumenta a satisfação.
- **Benefícios Extras Atraentes:** Muitas instituições oferecem benefícios adicionais para quem utiliza o cartão de forma frequente, como isenção de anuidade ou bônus em programas de recompensas, o que faz com que esses clientes vejam valor em manter o cartão.

Portanto, a quantidade de transações impacta diretamente a percepção de valor do cartão. Usuários com poucas ou moderadas transações frequentemente acham que os benefícios não compensam os custos, levando-os a cancelar o cartão. Em contraste, aqueles com alta quantidade de transações geralmente obtêm retornos vantajosos e conveniências que justificam a manutenção do cartão.

Taxa de Utilização do Cartão



Os clientes com menor taxa de utilização foram os que mais cancelam o cartão, isso sugere que eles percebem o cartão como tendo pouco valor em relação ao custo de mantê-lo, ou que suas necessidades de crédito são menores, tornando o cartão dispensável. Clientes com uma baixa taxa de utilização provavelmente não veem o cartão como uma ferramenta essencial para o dia a dia. Como eles utilizam pouco do limite disponível, o cartão pode parecer supérfluo, especialmente se ele acarreta custos de manutenção, como anuidades ou taxas de serviços. Esses clientes podem preferir outras formas de pagamento, como cartões de débito ou carteiras digitais, que não trazem custos adicionais e são mais alinhados ao seu perfil de uso. Para quem usa pouco o limite, os benefícios do cartão (como cashback ou milhas) podem ser insuficientes para justificar o custo de manter o cartão. Como eles têm uma

utilização reduzida, o retorno em recompensas ou benefícios é limitado, e o custo de anuidade ou outras taxas pode parecer desproporcional ao benefício obtido. Esses clientes, então, optam por cancelar o cartão em busca de alternativas mais econômicas. Clientes com baixa taxa de utilização também podem sentir que os incentivos oferecidos pelo cartão não são adequados ao seu perfil. Se o cartão não oferece recompensas atrativas para transações esporádicas ou de baixo valor, esses clientes podem preferir produtos financeiros com melhores incentivos, como cartões de débito com cashback ou contas digitais gratuitas. Os clientes com baixa taxa de utilização podem ser mais conservadores em relação ao crédito, evitando o uso do limite para evitar o risco de endividamento. Para eles, o crédito disponível em um cartão de crédito não é tão relevante e pode até representar uma tentação para gastos indesejados. Esse perfil mais conservador tende a ver o cartão como menos necessário, optando pelo cancelamento. A alta taxa de cancelamento entre clientes com menor taxa de utilização sugere que eles não veem um valor agregado no cartão, dado o uso reduzido e o custo de mantê-lo. Eles estão em busca de produtos financeiros mais alinhados ao seu perfil de gastos e que não representem custos desnecessários. Para reter esses clientes, as instituições poderiam oferecer alternativas com taxas reduzidas ou benefícios mais flexíveis, adequados a usuários com menor frequência de uso.

CONCLUSÃO

A análise dos padrões de uso, cancelamento e percepção de valor entre os clientes do cartão de crédito revela algumas tendências importantes. Clientes com alta utilização e alto volume de transações tendem a ver mais valor no cartão, aproveitando benefícios que justificam os custos, enquanto usuários com baixa taxa de utilização ou transações esporádicas encontram menor custo-benefício, levando a um maior índice de cancelamento. Já clientes com uma taxa de utilização moderada encontram um equilíbrio que os mantém satisfeitos e menos inclinados a cancelar.

Algumas sugestões para evitar ou reduzir os cancelamentos, seriam:

1. **Oferecer Cartões com Anuidades Variáveis:** Criar opções de cartões com anuidades reduzidas ou mesmo sem anuidade para clientes de uso esporádico ou com menores volumes de transação. Isso tornaria o cartão mais atraente para quem faz menos transações e valoriza o baixo custo.
2. **Ajustar Benefícios ao Perfil de Uso:** Proporcionar recompensas diferenciadas, como cashback ou pontos extras para quem realiza transações menores ou usa o cartão com menor frequência. Alternativamente, permitir que clientes escolham benefícios adequados ao seu perfil (milhas, descontos, isenções).
3. **Oferecer Incentivos para Utilização Moderada:** Implementar programas de fidelidade específicos para usuários de uso moderado, como descontos em anuidade para quem mantém uma taxa de utilização média e consistente. Isso pode aumentar o valor percebido para clientes com utilização intermediária e incentivar o engajamento.
4. **Simplificar o Processo de Solicitação de Aumento de Limite:** Para clientes que demonstram alta utilização do limite, oferecer um processo mais fácil e rápido para solicitação de aumento de limite. Esse grupo provavelmente verá o cartão como uma ferramenta valiosa e estará disposto a mantê-lo.
5. **Educação e Orientação Financeira:** Oferecer conteúdo educativo para clientes com baixa utilização, mostrando as vantagens de um uso equilibrado e controlado do cartão de crédito. Isso pode gerar maior compreensão do valor do cartão e incentivar um uso mais ativo.

Essas ações permitem que o banco atenda de maneira mais eficaz às necessidades variadas dos seus clientes, melhorando a experiência e reduzindo a taxa de cancelamento ao alinhar os produtos oferecidos aos perfis e comportamentos dos usuários.