

Insights

Case - Investigation regarding client service cancelations

```
In [ ]: import pandas as pd

        table = ""

        table = pd.read_csv("cancelamentos.csv")
        table = table.drop(columns="CustomerID")

        display(table)
```

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

881666 rows × 11 columns



```
In [ ]: display(table.info())
        table = table.dropna()
        display(table.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'>
Index: 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 [ ]: display(table["cancelou"].value_counts())
display(table["cancelou"].value_counts(normalize=True).map("{:.1%}".format))
```

```
cancelou
1.0    499993
0.0    381666
Name: count, dtype: int64
cancelou
1.0    56.7%
0.0    43.3%
Name: proportion, dtype: object
```

```
In [ ]: display(table["duracao_contrato"].value_counts(normalize=True))
display(table["duracao_contrato"].value_counts())
```

```
duracao_contrato
Annual    0.401964
Quarterly 0.400448
Monthly   0.197588
Name: proportion, dtype: float64
```

```
duracao_contrato
Annual      354395
Quarterly   353059
Monthly     174205
Name: count, dtype: int64
```

```
In [ ]: display(table.groupby("duracao_contrato").mean(numeric_only=True))
```

	idade	tempo_como_cliente	frequencia_uso	ligacoes_callcenter	dia
duracao_contrato					
Annual	38.842165	31.446186	15.880213	3.263401	12.507054
Monthly	41.552407	30.538555	15.499274	4.985649	12.433427
Quarterly	38.830938	31.419916	15.886662	3.265245	12.450690

```
In [ ]: table = table[table["duracao_contrato"]!="Monthly"]
# display(table)
#display(table["cancelou"].value_counts())
display(table["cancelou"].value_counts(normalize=True).map("{:.1%}".format))
```

```
cancelou
0.0    53.9%
1.0    46.1%
Name: proportion, dtype: object
```

```
In [ ]: # display(table)
display(table["assinatura"].value_counts(normalize=True))
display(table.groupby("assinatura").mean(numeric_only=True))
```

```
assinatura
Standard    0.339648
Premium     0.338138
Basic       0.322215
Name: proportion, dtype: float64
```

	idade	tempo_como_cliente	frequencia_uso	ligacoes_callcenter	dias_atraso
assinatura					
Basic	38.904813	32.316031	15.876921	3.310021	12.507054
Premium	38.817814	30.977869	15.889673	3.235886	12.433427
Standard	38.790478	31.048621	15.883393	3.249275	12.450690

```
In [ ]: import plotly.express as px

for column in table.columns:
    chart = px.histogram(table, x=column, color="cancelou")
    chart.show()
```

```
In [ ]: display(table["cancelou"].value_counts(normalize=True).map("{:.1%}".format))

table = table[table["ligacoes_callcenter"]<5]
table = table[table["dias_atraso"]<=20]
```

```
display(table["cancelou"].value_counts(normalize=True).map("{:.1%}".format))
```

```
cancelou
0.0    53.9%
1.0    46.1%
Name: proportion, dtype: object

cancelou
0.0    81.6%
1.0    18.4%
Name: proportion, dtype: object
```

More Relevant Perceptions:

1. All "Monthly" contracts have been canceled. Incentivise users to move to another contract model or revise the "Monthly" so it gets more attractive.
2. The user's issues need to be resolved within 4 user calls to the call center.
3. Need to improve the follow-up of past dues since 100% of clients canceled after 20 days past due.