

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.

```
In [ ]: # !pip install pandas plotly numpy openpyxl nbformat ipkernel
# install library with ! to use the jupyter celule
import pandas as pd

table = pd.read_csv('cancelamentos_sample.csv') # read csv file
table = table.drop(columns=['CustomerID']) # remove a unnecessary column
display(table)
```

	idade	sexo	tempo_como_cliente	frequencia_uso	ligacoes_callcenter	dias_atras
0	23.0	Male	13.0	22.0	2.0	1.
1	49.0	Male	55.0	16.0	3.0	6.
2	30.0	Male	7.0	1.0	0.0	8.
3	26.0	Male	40.0	5.0	3.0	8.
4	27.0	Female	17.0	30.0	5.0	6.
...
49995	62.0	Female	35.0	7.0	2.0	8.
49996	36.0	Male	43.0	21.0	2.0	30.
49997	55.0	Male	42.0	8.0	1.0	12.
49998	40.0	Female	14.0	19.0	1.0	17.
49999	64.0	Male	41.0	29.0	5.0	9.

50000 rows × 11 columns

```
In [ ]: table = table.dropna() # remove rows with missing values

# display(table.info()) # show the data type of each column

display(table["cancelou"].value_counts(normalize=True)) # show the count of each
```

```
cancelou
1.0    0.567905
0.0    0.432095
Name: proportion, dtype: float64
```

```
In [ ]: import plotly.express as px # import the library to plot the graph

color_map = {0.0: '#1f77b4', 1.0: '#d62728'} # create a color map to use in the

for coluna in table.columns:
    if coluna == 'cancelou':
        continue
    graph = px.histogram(table, x=coluna, color='cancelou', color_discrete_map=c
    graph.show()
```

```
In [ ]: # clientes com contrato mensal cancelam mais
        # oferecer desconto em outros tipos de contratos

table = table[table["duracao_contrato"] != 'Monthly']
display(table["cancelou"].value_counts(normalize=True))

# clientes com mais ligações no callcenter cancelam mais
# melhorar o atendimento
# alertar quando tem mais 3 ligações
table = table[table["ligacoes_callcenter"] <= 3]
display(table["cancelou"].value_counts(normalize=True))

# clientes com atraso de mais de 20 dias cancelam mais
# acompanhamento para evitar atrasos
table = table[table["dias_atraso"] <= 20]
display(table["cancelou"].value_counts(normalize=True))
```

```
cancelou
0.0    0.538567
1.0    0.461433
Name: proportion, dtype: float64

cancelou
0.0    0.766424
1.0    0.233576
Name: proportion, dtype: float64

cancelou
0.0    0.841236
1.0    0.158764
Name: proportion, dtype: float64
```