

# Modelo preditivo para aumento de conversões do setor de telemarketing do Banco Santander

## Grupo:

- Fábio Duarte
- Lorena Bastos
- Mateus Falcão
- Philipe Vasconcelos





## Problema de negócio

O Banco Santander está com um novo produto em sua carteira de investimentos.

A fim de obter um ROI (*return on investment*) significativo nesta campanha, o diretor de vendas do Banco Santander, Alex Lima, pediu para a equipe de cientistas de dados do banco desenvolver um modelo preditivo que identifique entre seus potenciais clientes aqueles que possuem a maior probabilidade de fazer uma aplicação neste novo produto.

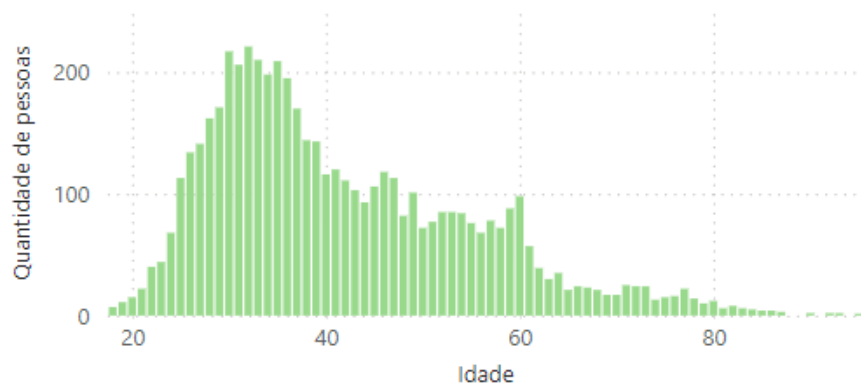
Desta forma, sua equipe de telemarketing poderá obter resultados muito eficientes e tirar sua corda do pescoço!

# Análise dos dados demográficos

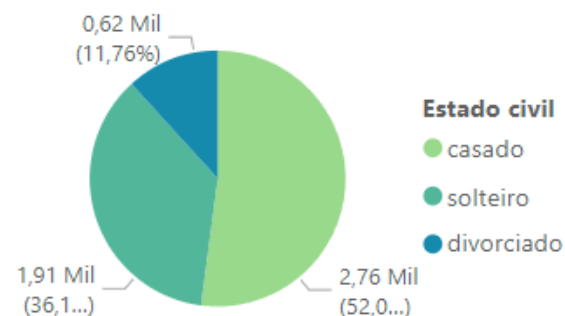
## Conversão

- ☐ Não realizada
- ☒ Realizada

## Idade da população

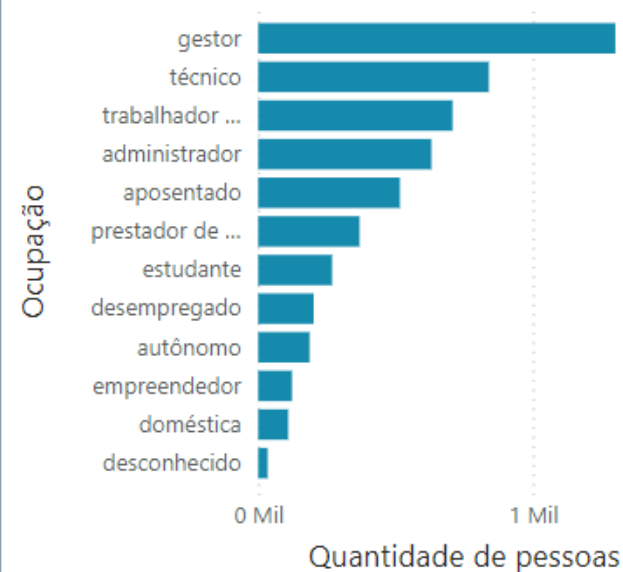


## Estado civil da população

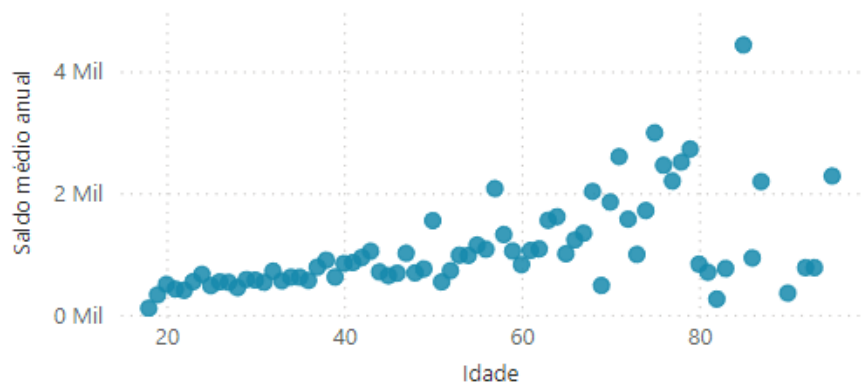


## Ocupação da população

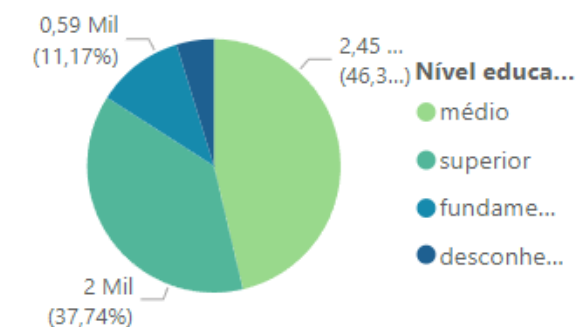
Conversão ● Realizada



## Saldo médio anual por idade



## Nível educacional da população



## Empréstimos contratados



# Análise dos dados da campanha

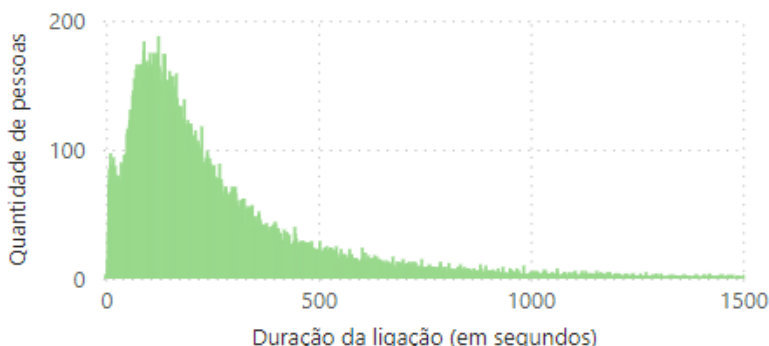
## Conversão

- ☐ Não realizada
- ☐ Realizada

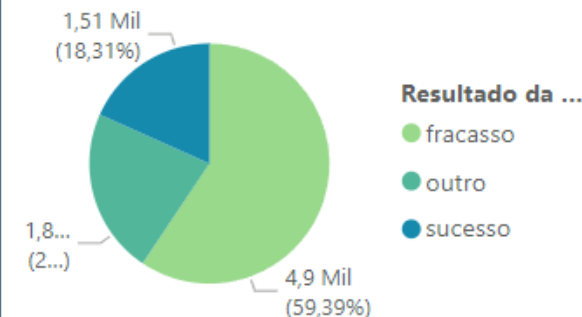
### Quantidade de pessoas por dia do mês



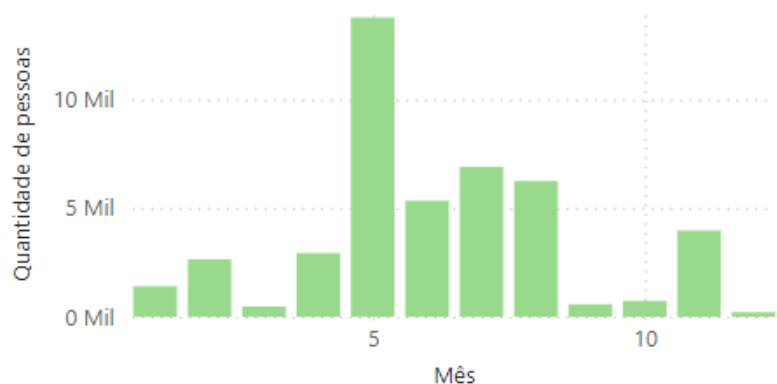
### Quantidade de pessoas pela duração da ligação



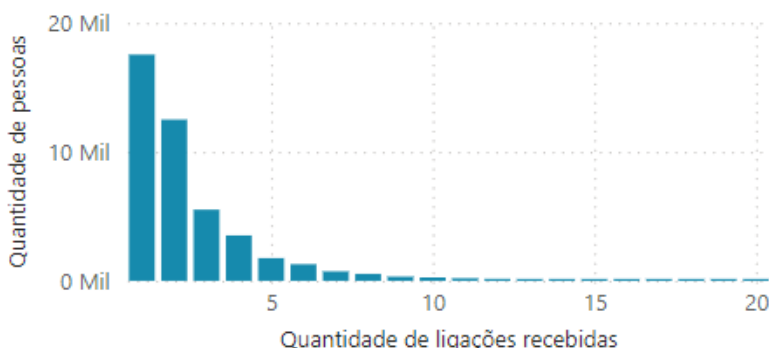
### Resultado da última campanha



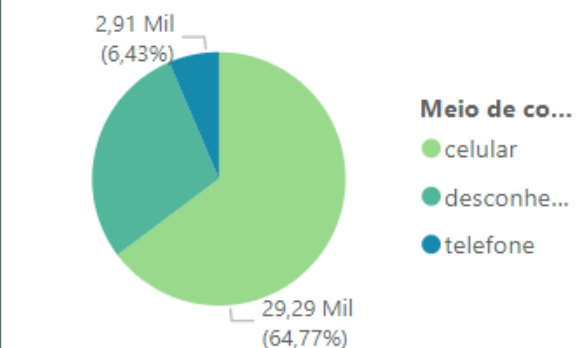
### Quantidade de pessoas por mês



### Quantidade de pessoas pela quantidade de ligações recebidas



### Meio de contato



# Insights para aumento da conversão

## Principais influenciadores Principais segmentos

O que influencia Conversão a ser Realizada

Realizada

?

Quando...

... a probabilidade de Conversão ser Realizada aumenta em

Resultado da última campanha é sucesso

6.62x

Mês é mais de 11

4.55x

Idade é mais de 60

4.30x

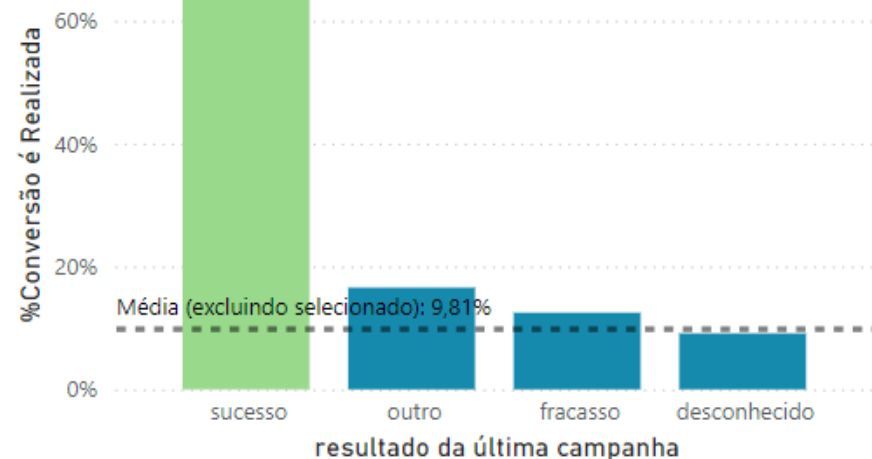
Duração da ligação aumenta para 262.72

4.25x

Mês é 8 - 10

4.03x

← É mais provável que Conversão seja Realizada quando Resultado da última campanha é sucesso do que o contrário (normalmente).

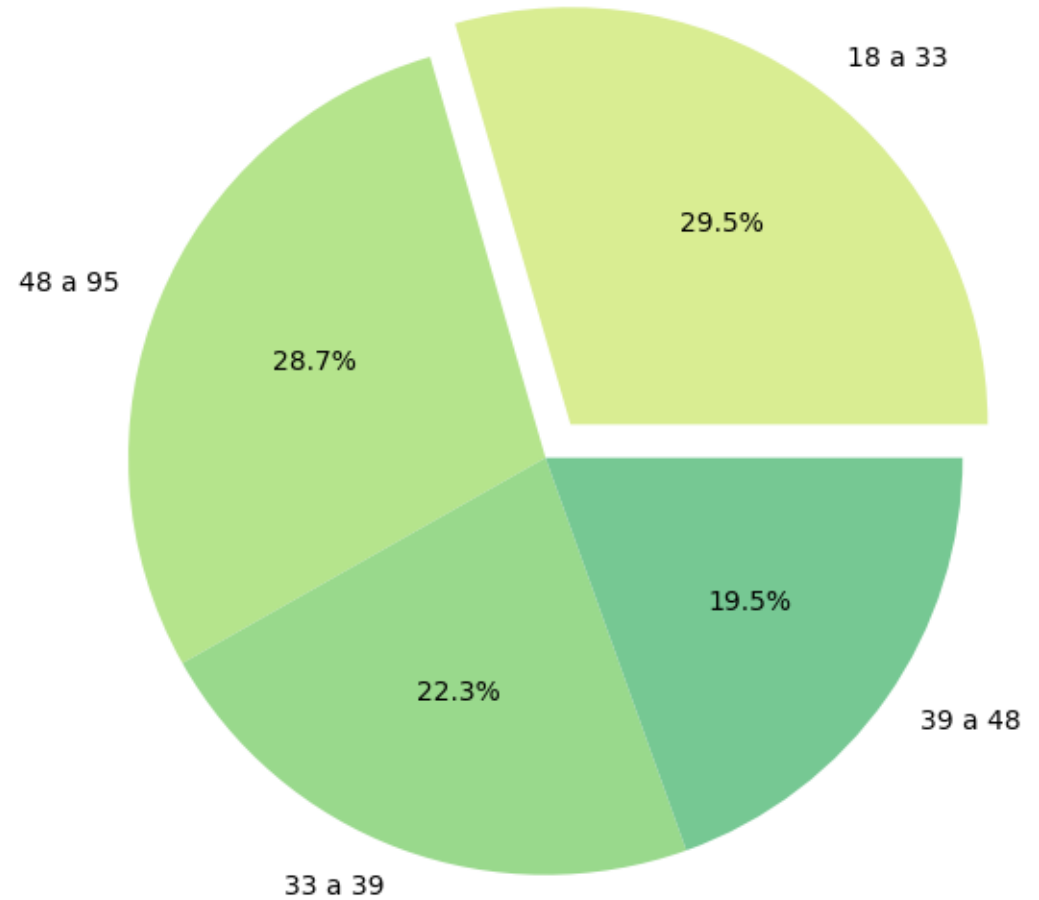


☐ Mostrar apenas os valores que são influenciadores

# FEATURE: idade

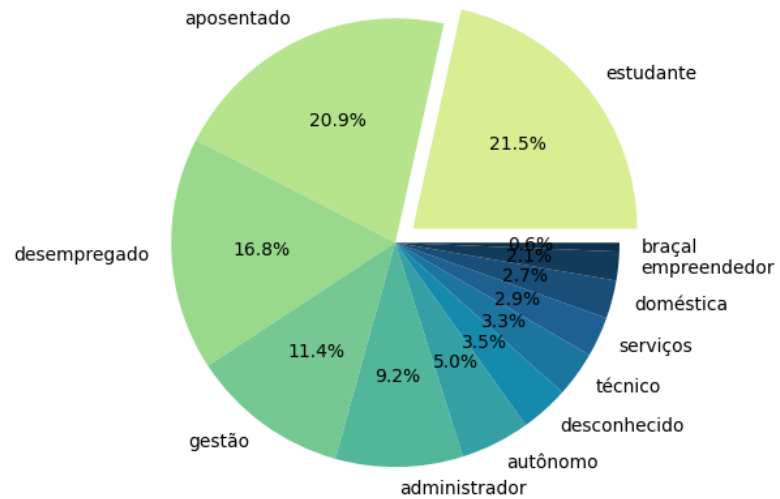


**Conversão por faixa de idade**

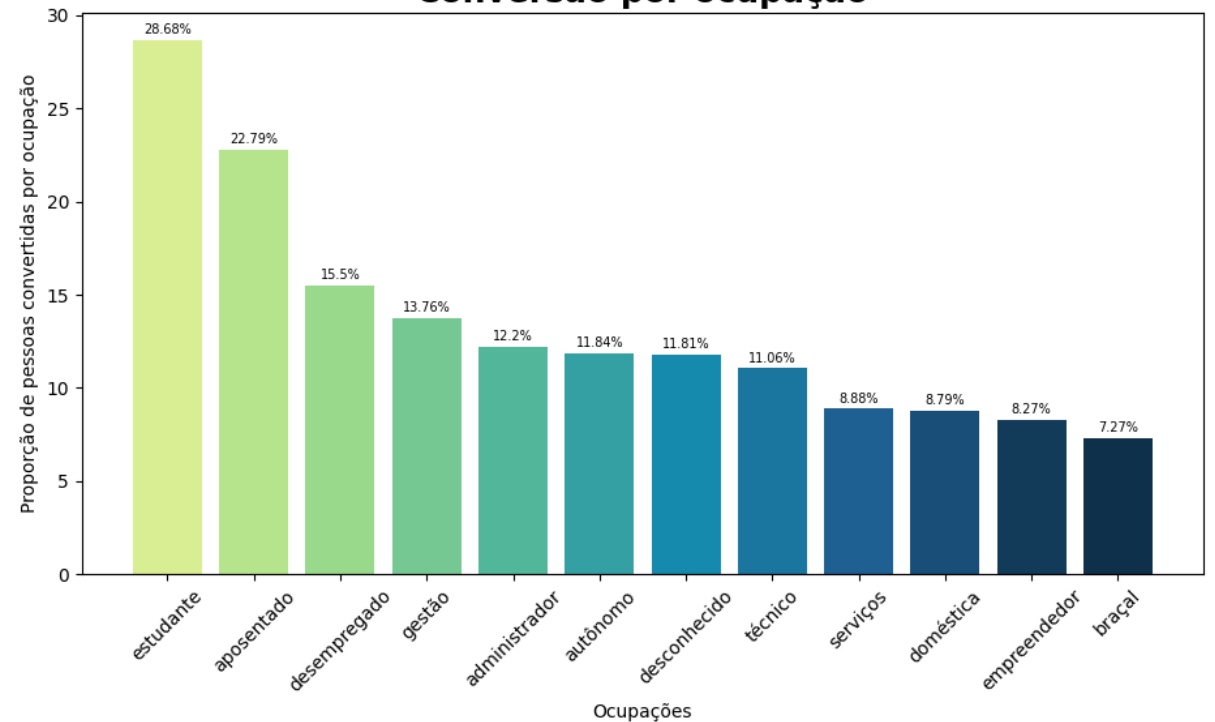


# FEATURE: ocupação

Distribuição de ocupações



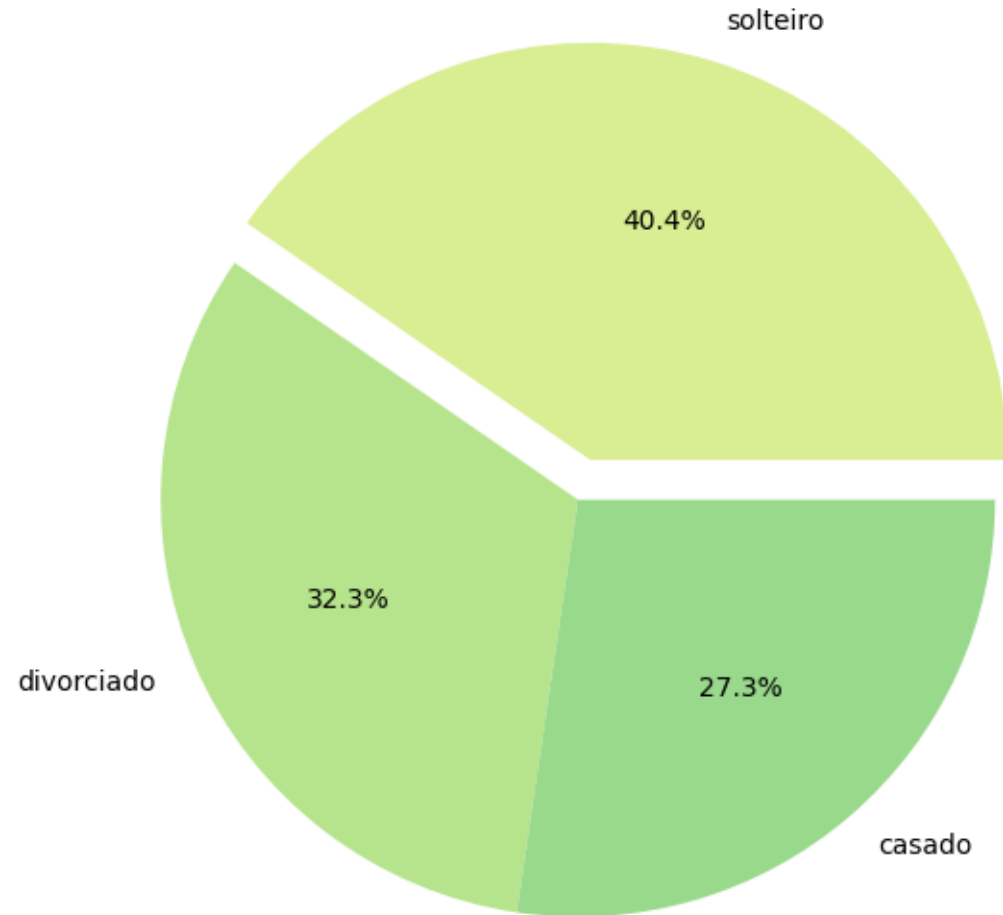
Conversão por ocupação



# FEATURE: estado civil



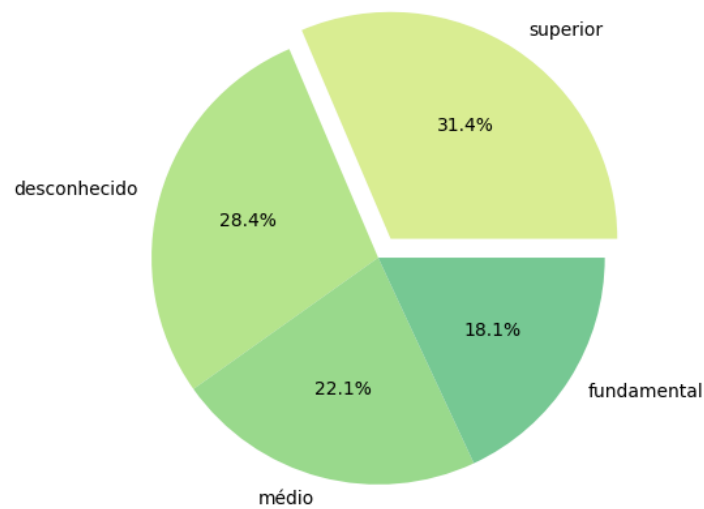
**Conversão por estado civil**



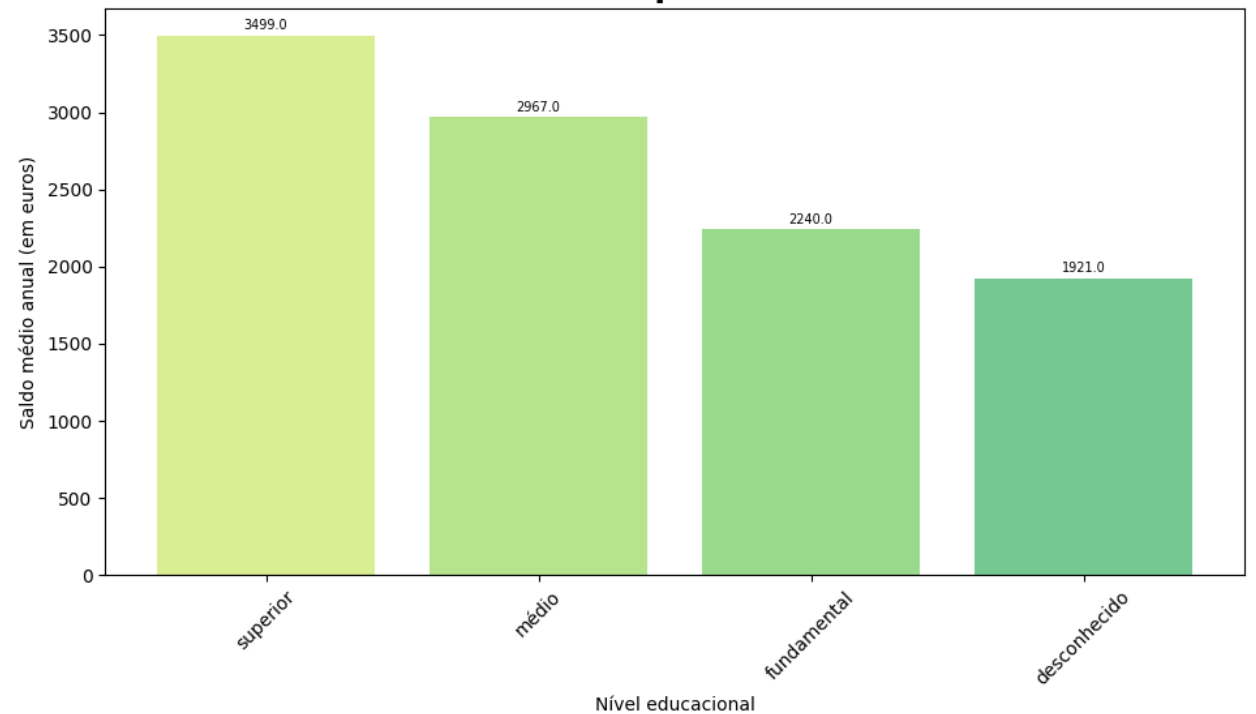


# FEATURE: nível educacional

Conversão por nível educacional



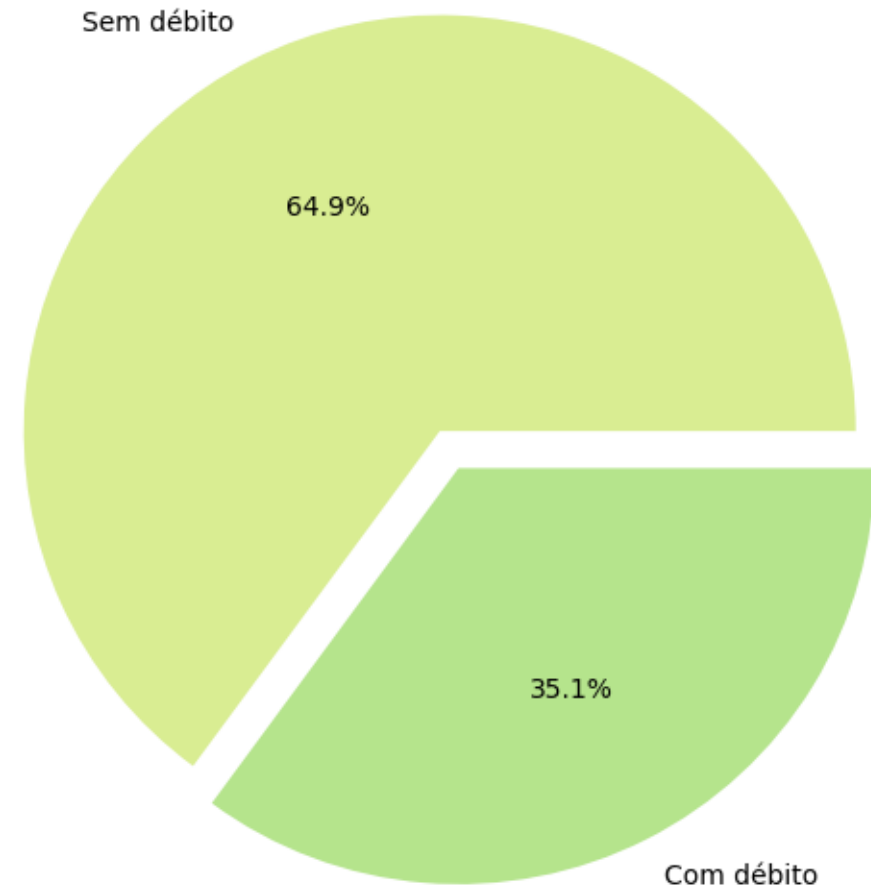
Saldo médio anual por nível educacional



# FEATURE: inadimplência

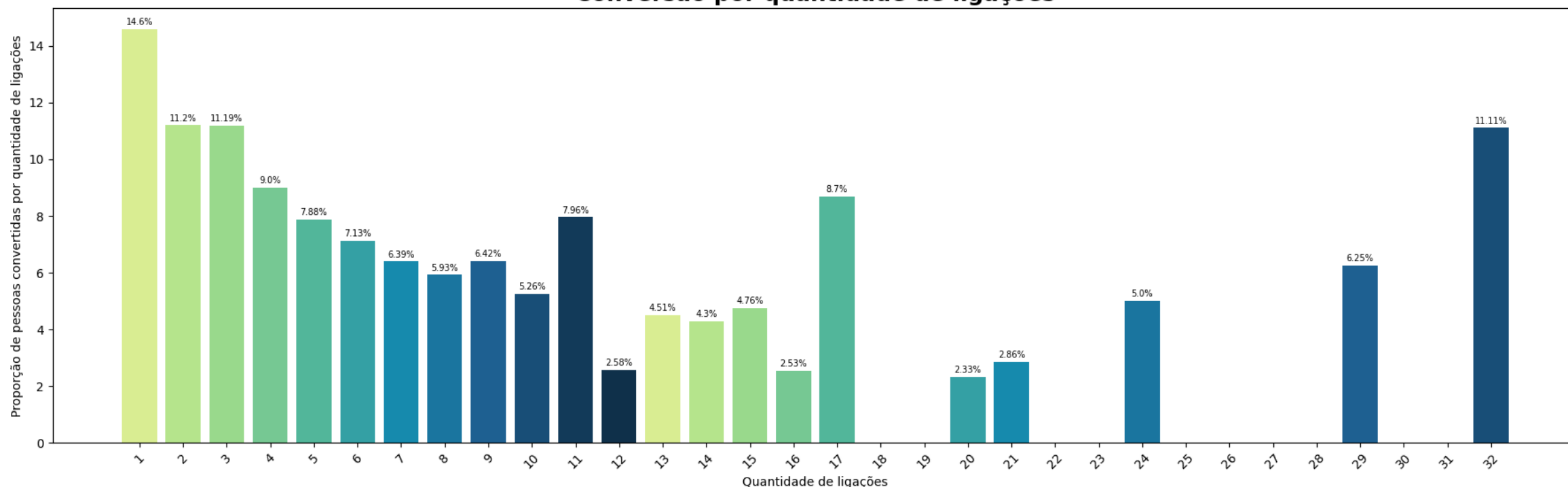


## Conversão por status de inadimplência



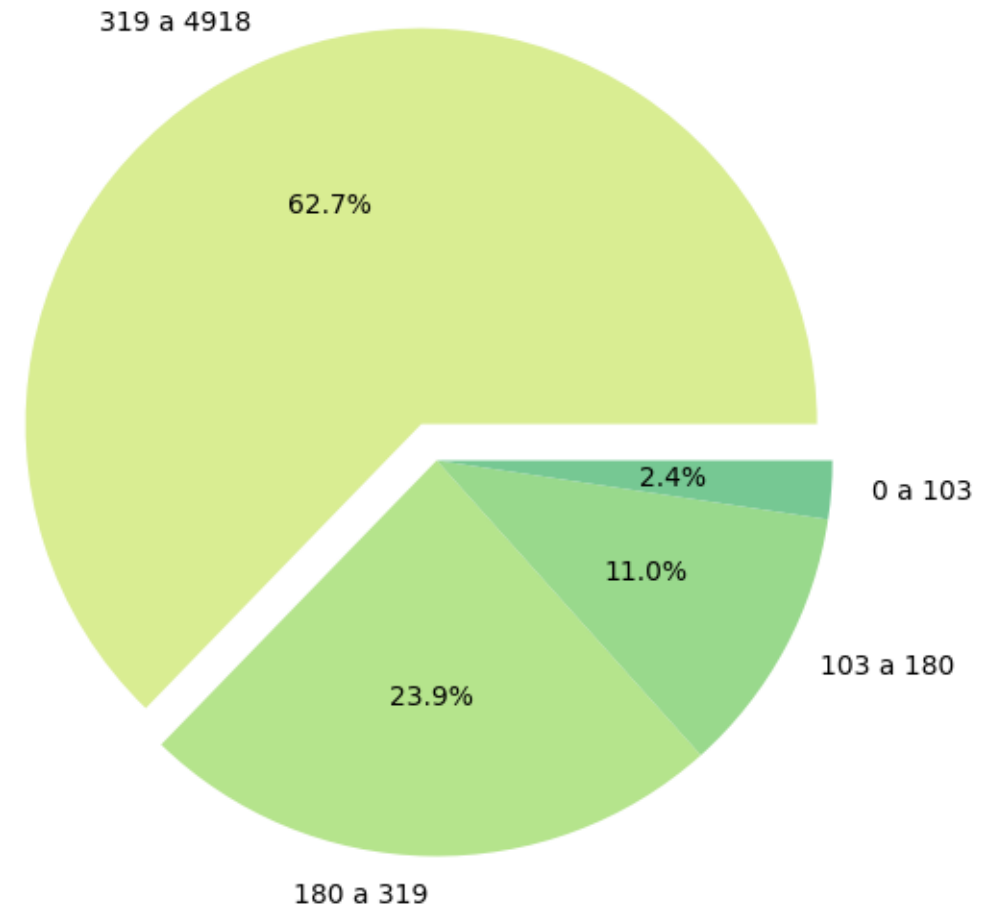
# FEATURE: quantidade de ligações

Conversão por quantidade de ligações



# FEATURE: duração da ligação

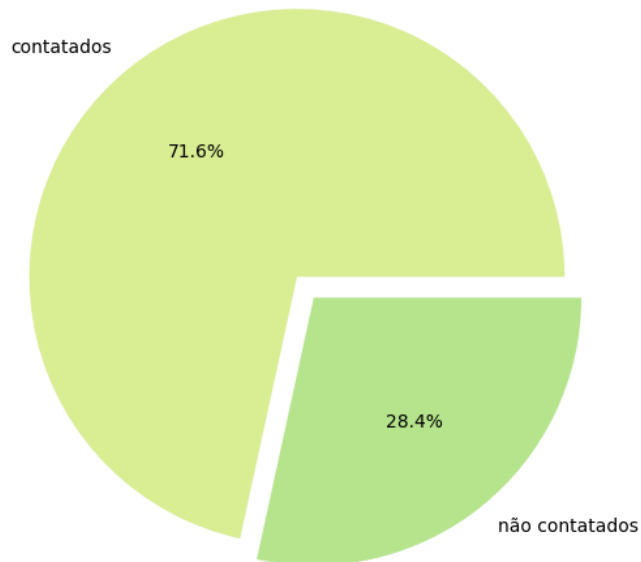
**Conversão por duração do contato  
(em segundos)**



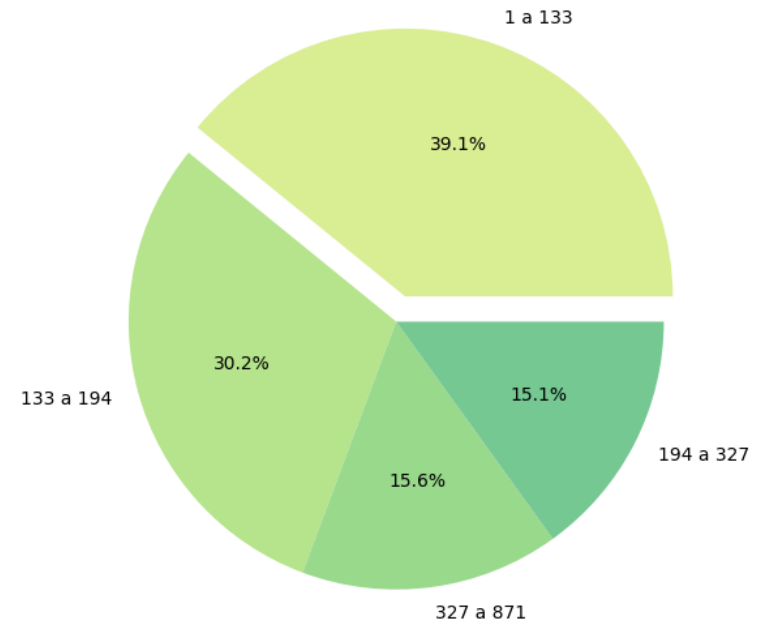
# FEATURE: resultado de campanhas anteriores



**Conversão por contatados  
anteriormente vs não-contatados**



**Conversão por quantidade de dias  
passados desde último contato**





## Sobre o modelo escolhido

Dado que o objetivo do banco é aumentar a quantidade de conversões no setor de telemarketing, escolhemos o modelo de classificação que obteve as maiores pontuações para a revocação e a acurácia.

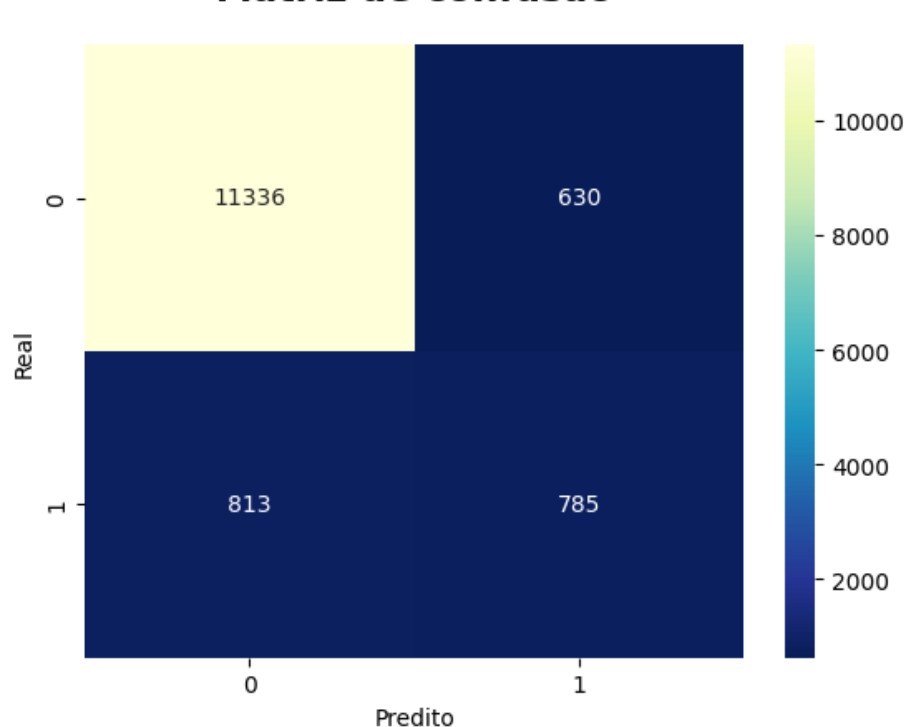
Levando em consideração que o banco aloca um montante financeiro significativo em campanhas de conversão, acreditamos que o melhor modelo é aquele que apresenta o maior percentual de classificações corretas enquanto busca a menor quantidade possível de erros do tipo II (ou falsos negativos), visto que os vendedores estariam perdendo oportunidades de negócio.

# MODELO: Árvore de decisão

## AVALIAÇÃO DO MODELO - MÉTRICAS DE CLASSIFICAÇÃO

- Acurácia: 0.8936154526688292
- Precisão: 0.5547703180212014
- Revocação: 0.49123904881101377
- F1-Score: 0.5210753401924991
- AUC: 0.7192949380775777

**Matriz de confusão**



**Curva ROC-AUC**

