

# Bank Marketing e Estatística

João Palma Neto (237550) e Rodrigo Gomes (186777)

## Introdução

“A Propaganda é a alma do negócio” você já deve ter se deparado com essa frase antes, acontece que desde as grandes até as pequenas empresas utilizam desse conhecimento popular para competir pela atenção e a preferência do consumidor. E a estatística claramente não ficaria de fora deste jogo comercial, seu papel na competição é de auxiliar os conhecimentos do comportamento do consumidor, a uma otimização efetiva, reconhecendo padrões, analisando tendências e concebendo previsões. Este trabalho traz análise dos dados de uma campanha de telemarketing ((??)) efetuada por um banco comercial, cujo objetivo era convencer os clientes a abrir uma conta bancária na instituição – *term deposit*, ou depósito com retorno periódico de juros. Dentre as variáveis, as que serão abordadas referem-se ao perfil econômico (emprego, estado\_civil, escolaridade, idade), diretamente ligadas às decisões de consumo no setor de crédito; e à duração das ligações efetuadas, afim de entender o comportamento geral de um caso de sucesso.

## Banco de Dados

A base é composta originalmente por 21 variáveis, listadas na tabela abaixo conforme a documentação, das quais 5 foram utilizadas no projeto, para fins de foco no objetivo supracitado. Dentre as variáveis ignoradas, 6 referem-se a dados de apoio e contexto, como taxas de preços para o consumidor e nível de emprego da economia nacional em determinado período. Fora excluída da análise também a variável default, booleana e referente à situação de inadimplência dos participantes do experimento, por haver apenas 3 entradas positivas em um total de 41188 observações, sendo portanto irrelevante para fins de comparação.

## Variáveis

### Atributos do cliente

- idade (categórico)
- emprego (categórico)
  - Área de trabalho, podendo ser:
    - admin (administrativo)
    - blue-collar (trabalhador de colarinho-azul, atividade braçal)
    - entrepreneur (empreendedor(a))
    - housemaid (trabalho doméstico)
    - management (gerência)
    - retired (aposentado(a))
    - self-employed (trabalhador(a) individual)
    - services (servidor(a))
    - student (estudante)
    - technician (técnico(a))
    - unemployed (desempregado(a))

- unknown (sem informação)
- marital (categórico)
 

Estado civil, podendo ser:

  - divorced (divorciado(a) ou viuvo(a))
  - married (casado(a))
  - single (solteiro(a))
  - unknown (sem informação)
- education (categórico):
 

Nível de educação:

  - basic.4y (ensino básico até o 4º ano)
  - basic.6y (ensino básico até o 6º ano)
  - basic.9y (ensino básico até o 9º ano)
  - high.school (ensino médio)
  - illiterate (analfabeto)
  - professional.course (curso profissionalizante)
  - university.degree (formação universitária)
  - unknown (sem informação)

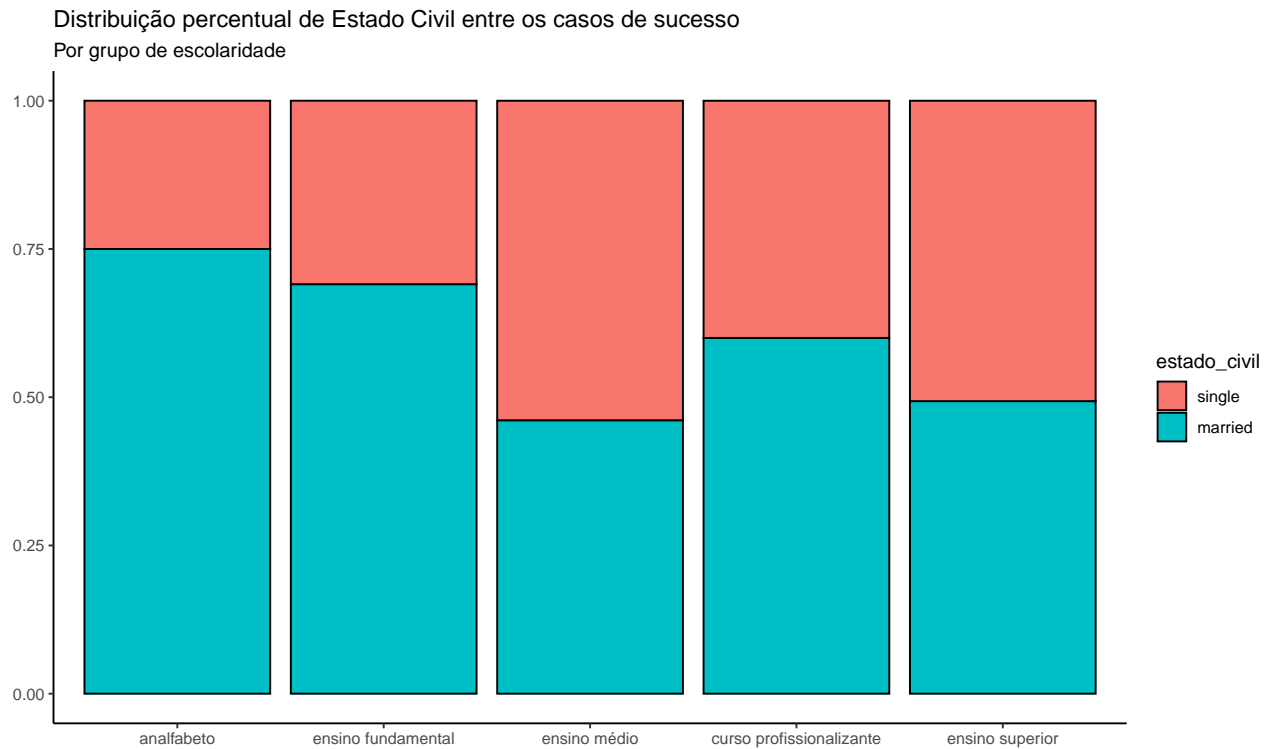
*Nota:* Para tornar mais claras as análises comparativas por faixa de educação, optou-se por agrupar as três categorias de ensino básico sob a categoria “ensino fundamental”.

#### **Atributos relacionados ao último contato da campanha atual**

- Duration (numérico): Duração em segundos do último contato.
- y (categórico): Booleano e referente à obtenção de sucesso na campanha atual.

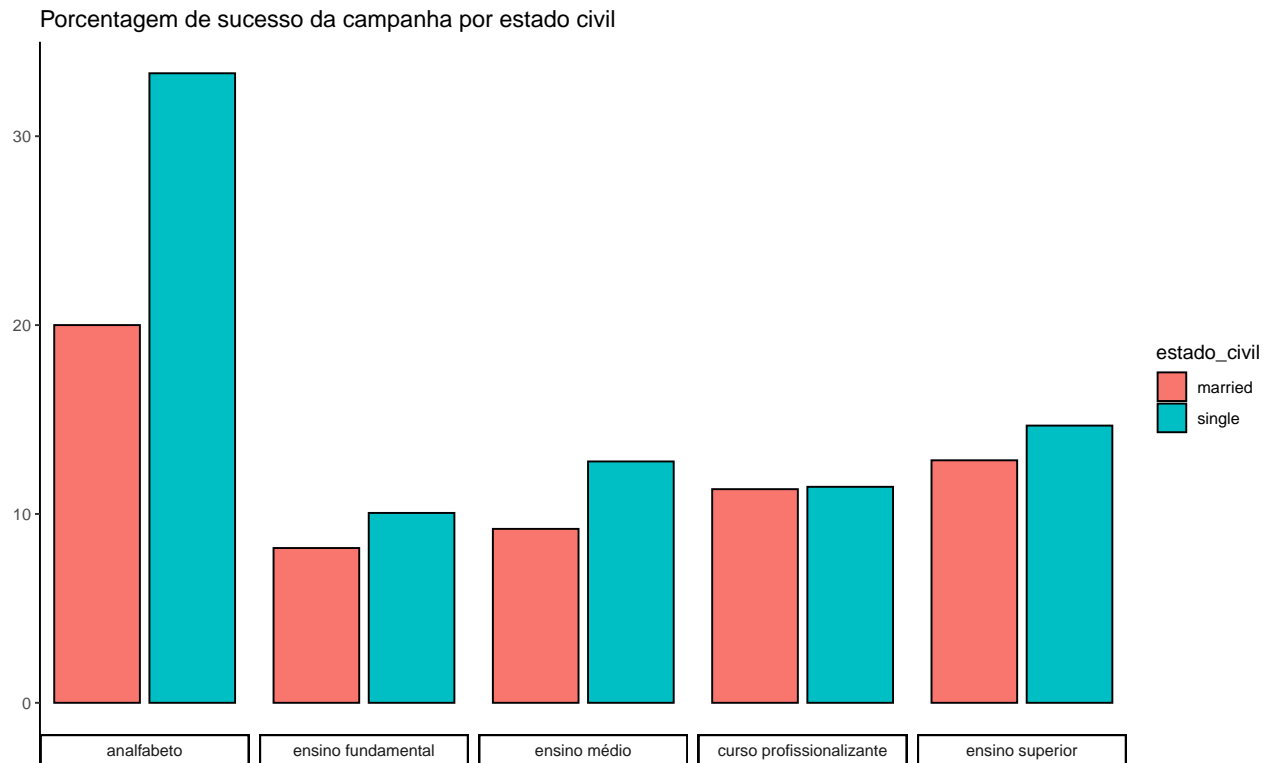
# Analise exploratoria

## Por estado civil



A variável escolaridade é utilizada como um dos parâmetros da análise de maneira intencional, utilizando das vantagens extraíveis pelo distinto comportamento entre as suas categorias, influenciadas por uma formação mais ou menos instrutiva, essa diferença busca evita distorções estatísticas para o estudo como esta sendo desenvolvido. Da mesma maneira foi parametrizado o estado civil, onde a existência de uma união matrimonial pode influenciar na prosperidade ou nos interesses econômicos do cliente.

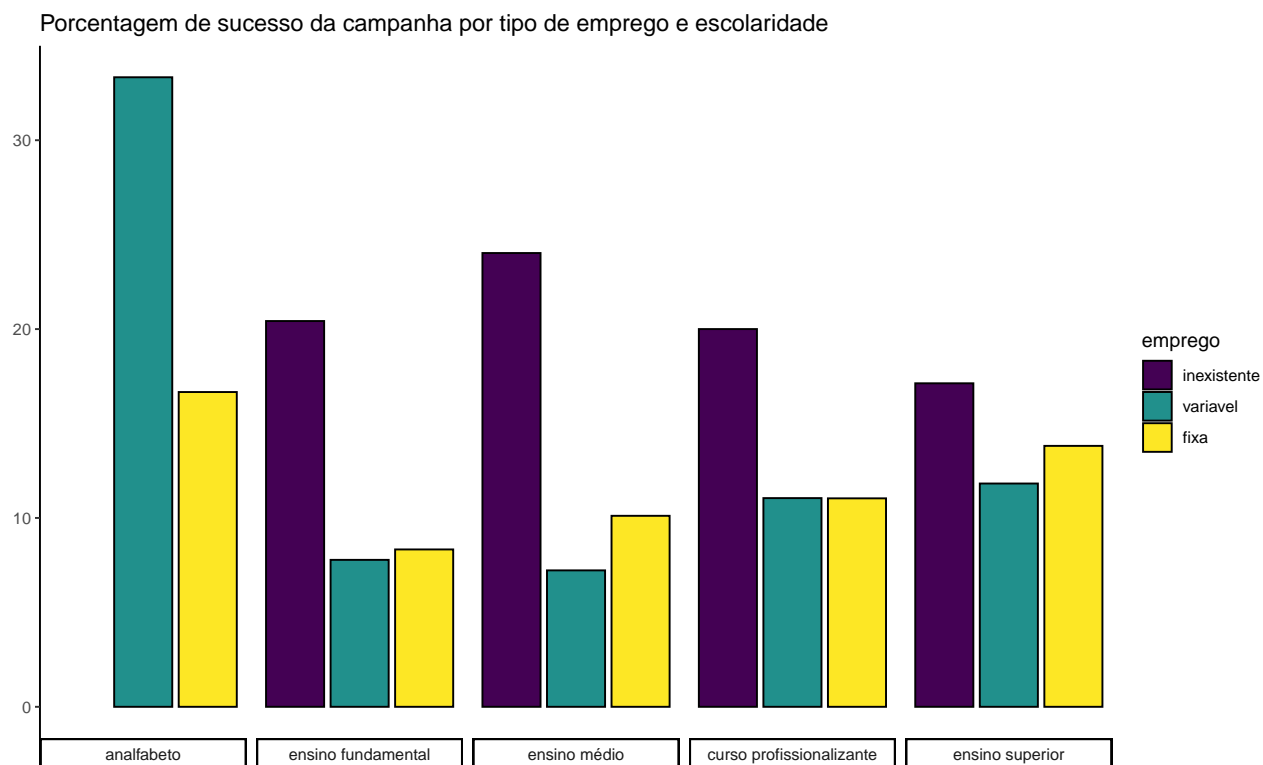
É entendível que apesar da educação, a população casada esta mais apta a abertura de conta do que a população solteira.



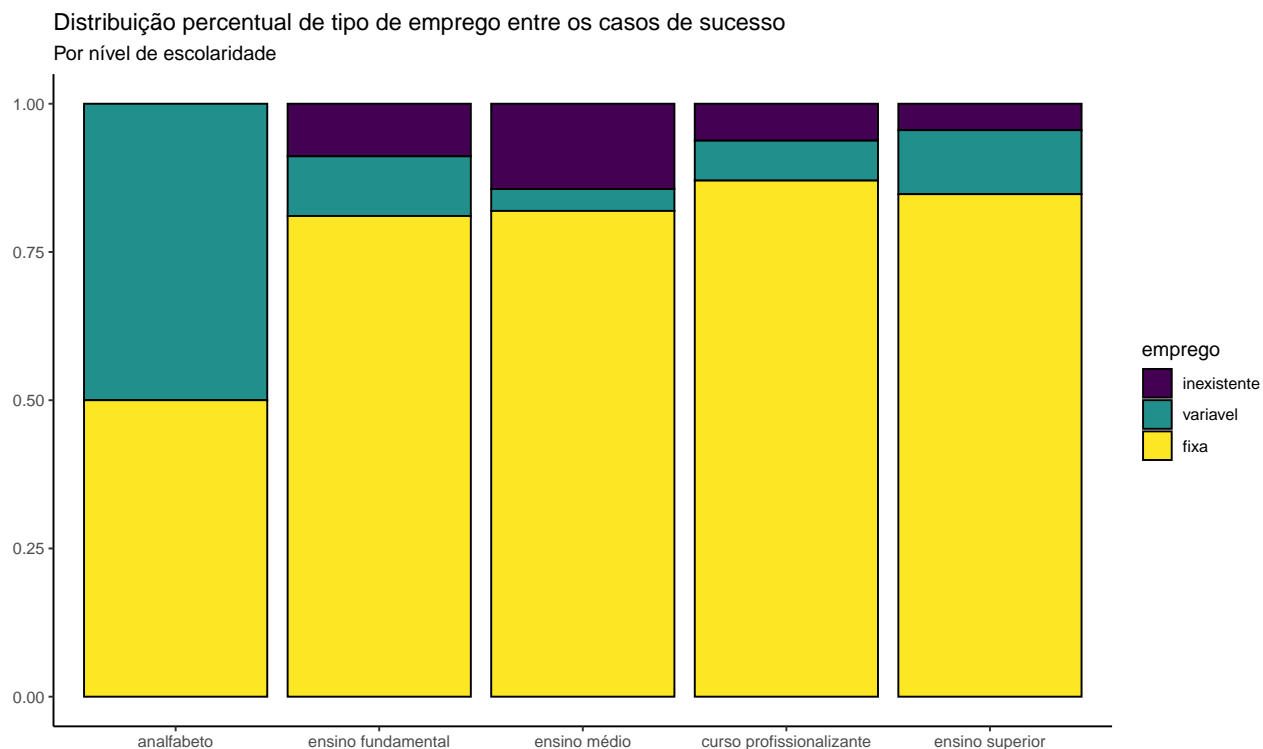
Aqui a população total é separada em grupos conforme estado civil, escolaridade e sucesso da campanha, então apenas os que aceitaram a campanha são selecionados (a porcentagem se refere ao total de pessoas no grupo). Tendo em vista o par ordenado analfabetos solteiros, formados pela união das duas categorias, temos maiores taxas de sucesso.

## Por emprego

Além do estado civil, outra dado econômico de grande relevância é o emprego atual do cliente, aqui classifica em tipos de emprego de renda fixa, variavel ou inexistente (desemprego), e avaliados novamente levando em conta o nível de escolaridade. É possível notar que os grupos proporcionalmente mais atingidos pela campanha são aqueles em situação econômica mais fragiizada ou instável, o que faz sentido tendo em vista a natureza financeira do produto, com expectativas de retorno no longo prazo.



Há, por outro lado, um volume muito maior de casos de sucesso com clientes de renda fixa em todos os grupos de escolaridade (salvo o grupo de analfabetos), dado que esta também é a categoria de emprego majoritária na base de dados.



## Por duração das chamadas

Por fim, para a análise quantitativa das durações de chamada, a tabela a seguir apresenta estatísticas sumárias da variável.

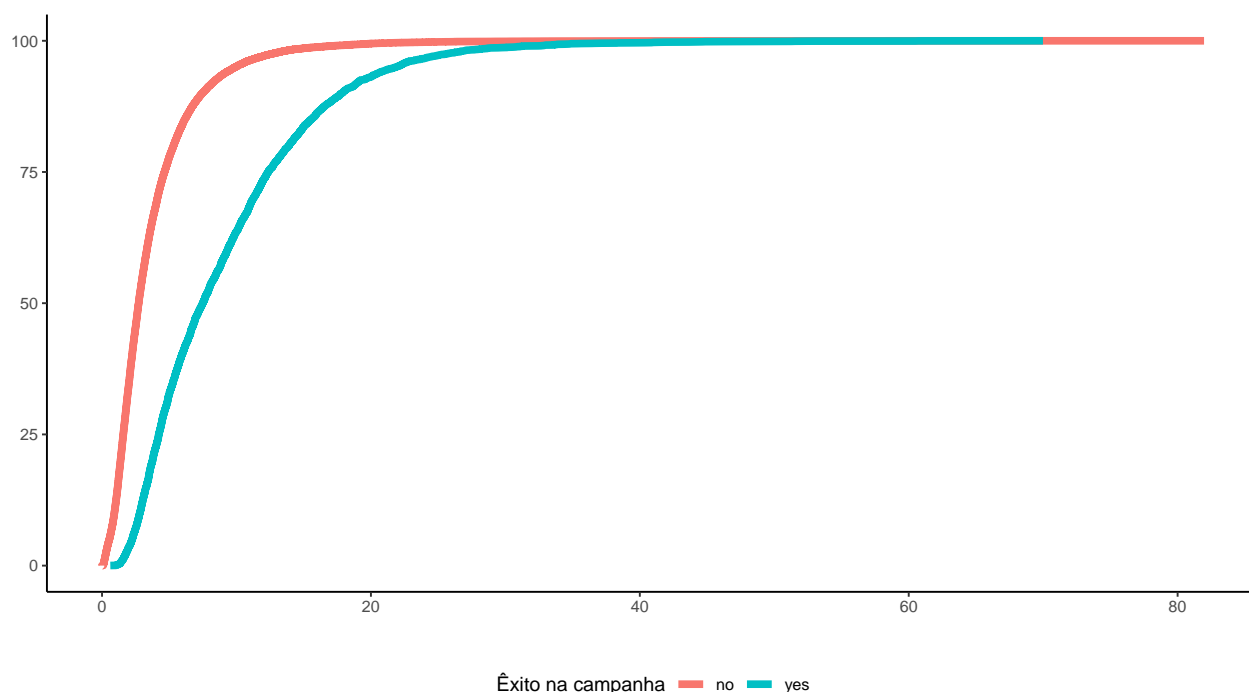
Table 1: Estatísticas Descritivas do tempo das chamadas(min) para casos de fracasso e sucesso

Estatísticas	Fracasso	Sucesso
Média	3.68	9.22
Desvio	3.45	6.69
Mediana	2.73	7.48
MIN	0.00	0.62
MAX	81.97	69.98

Avaliando as estatísticas observadas na tabela @ref(tab:tabX) é notável que apesar da média das chamadas no caso de sucesso serem superiores ao caso de fracasso, a chamada que teve maior duração (81.97 min) pertence a um caso de fracasso, onde 75% não ultrapassa os 4.65 minutos, se assim um outlier muito singular. Pensando na efetividade das chamadas, 25% das chamadas bem sucedidas precisaram de menos de 5 minutos para vender o produto.

O gráfico a seguir apresenta a frequência relativa e cumulativa de sucesso, e permite visualizar os limites máximo e mínimo de duração da chamada que abarcam a maior densidade de casos de sucesso e fracasso. Nesse caso, ligações bem sucedidas tem variam conforme um intervalo mais amplo de tempo, começando ao redor de 2 minutos e terminando em cerca de 20.

Porcentagem cumulativa do resultado da campanha por duração da chamada



## Por idade

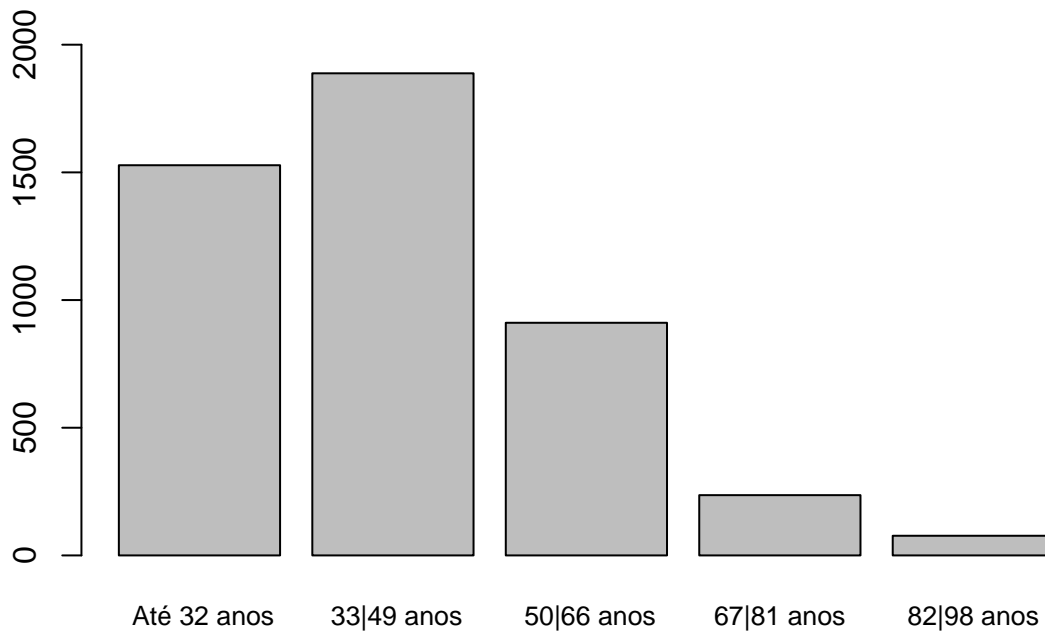
A faixa etária é uma variável que não só estimula o comportamento e a forma de agir financeiramente, como também mantém relações com a variável estado civil. Tendo em vista o intervalo entre a menor idade da

amostra (17 anos) e a maior (89 anos), foi criado hiatos interessados em uma homogeneidade entre cada faixa, tornando assim possível extrair informações isoladas pertinente ao grupo etário e também tornando visualmente atraente.

## Conclusão

É notável que, em geral, condições mais vulneráveis financeiramente, restrições ao acesso de informação e instrução, também associadas a condições de união matrimonial e construção familiar, gera um cenário mais favorável ao interesse do cliente no produto ofertado pelo banco. Ter ciência a respeito de outros tipos de investimento ou produtos do mesmo banco e entendimento a respeito da rentabilidade desses, pode influenciar com que grupos em situação mais segura financeiramente, recusem a proposta.

### Publicidade com sucesso por faixa etária



## Referências