

RISCO DE CRÉDITO

Apresentado por: Clevia Bento de Oliveira

Introdução

Uma organização deseja prever quem são os possíveis inadimplentes no produto de empréstimos ao consumidor.

Eles têm dados sobre o comportamento histórico do cliente com base no que observaram. Portanto, quando adquirem novos clientes, eles desejam prever quem é mais arriscado e quem não é.



Objetivo

Identificar qual perfil de clientes possui mais risco para a organização de serem inadimplentes.

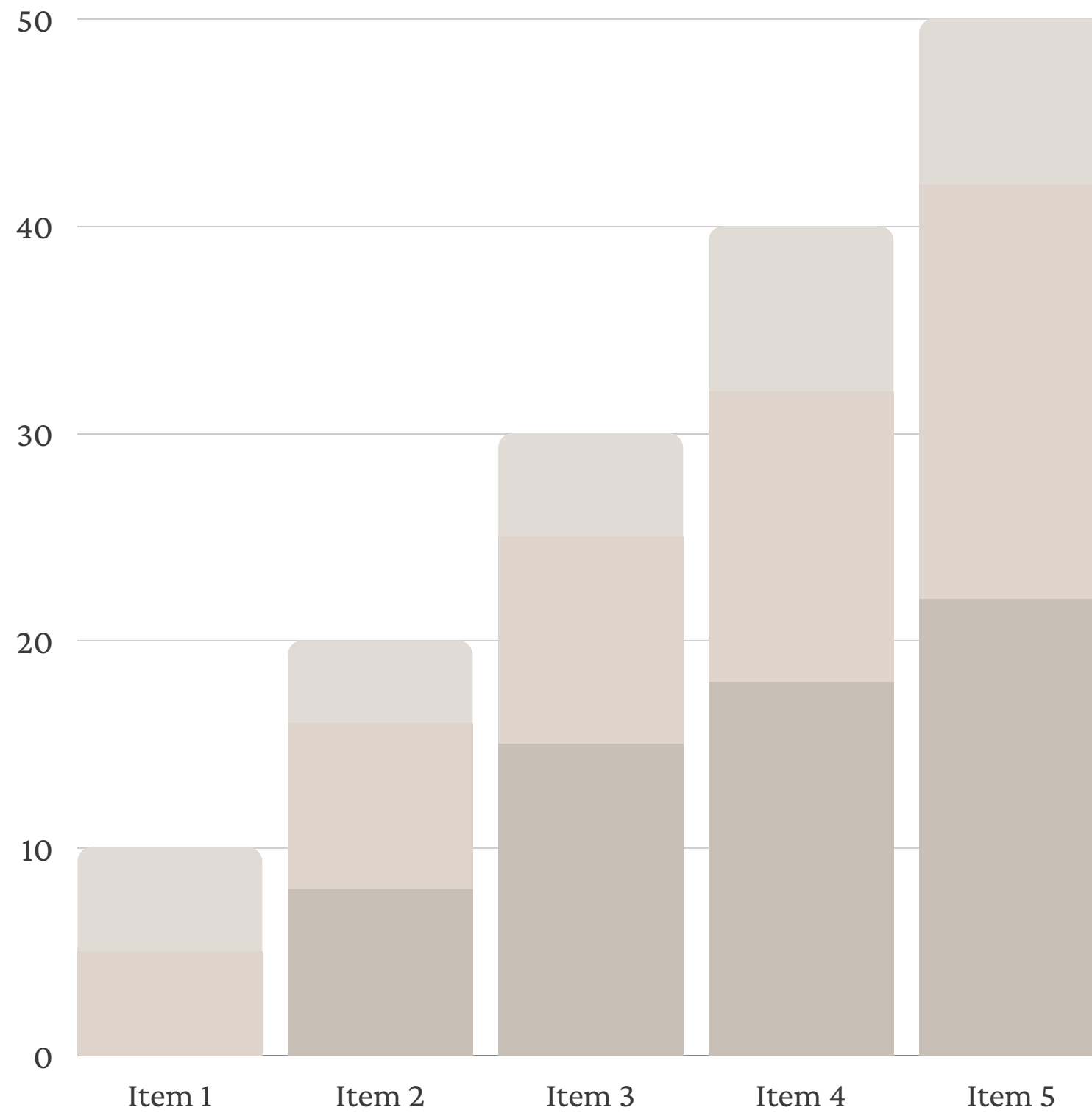


Metodologia

Será aplicado um teste de hipóteses Qui-quadrado para identificar associações entre a variável resposta e variáveis explicativas.

O conjunto de dados de treinamento servirá para identificar padrões de clientes inadimplentes.

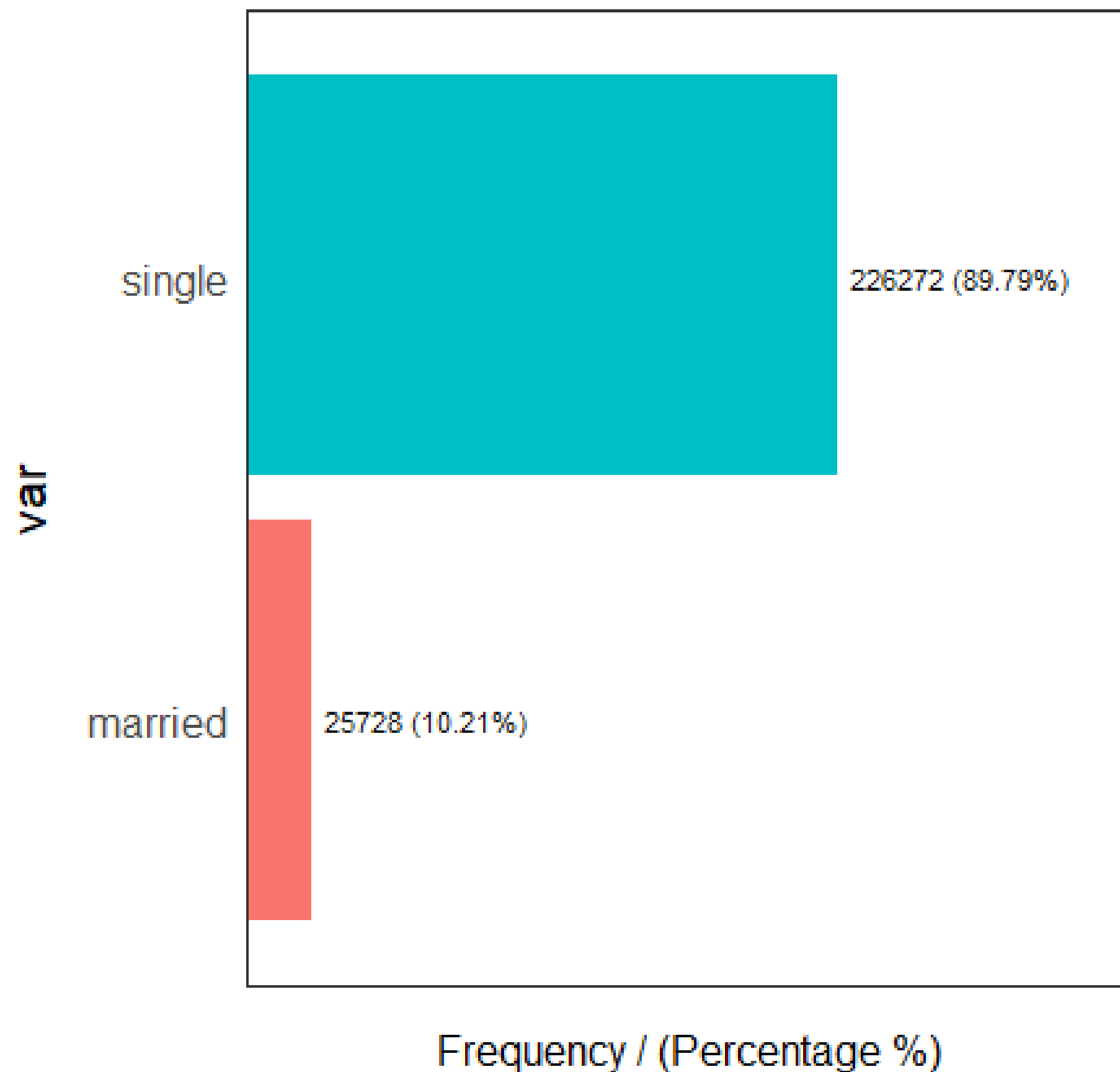
Para esta análise o software utilizado será o RStudio.



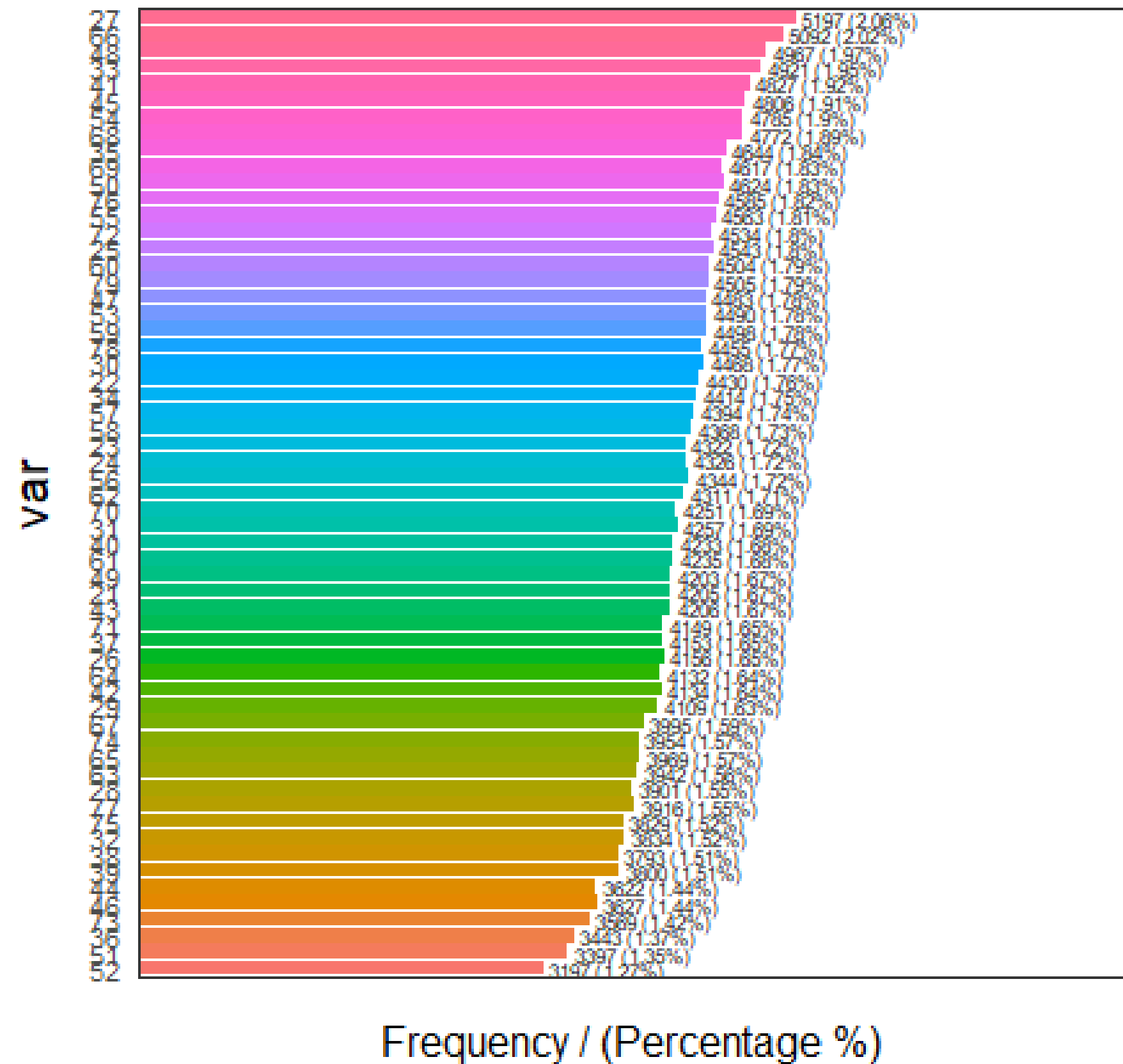
Estudo Descritivo das Variáveis

Frequência da variável Estado Civil

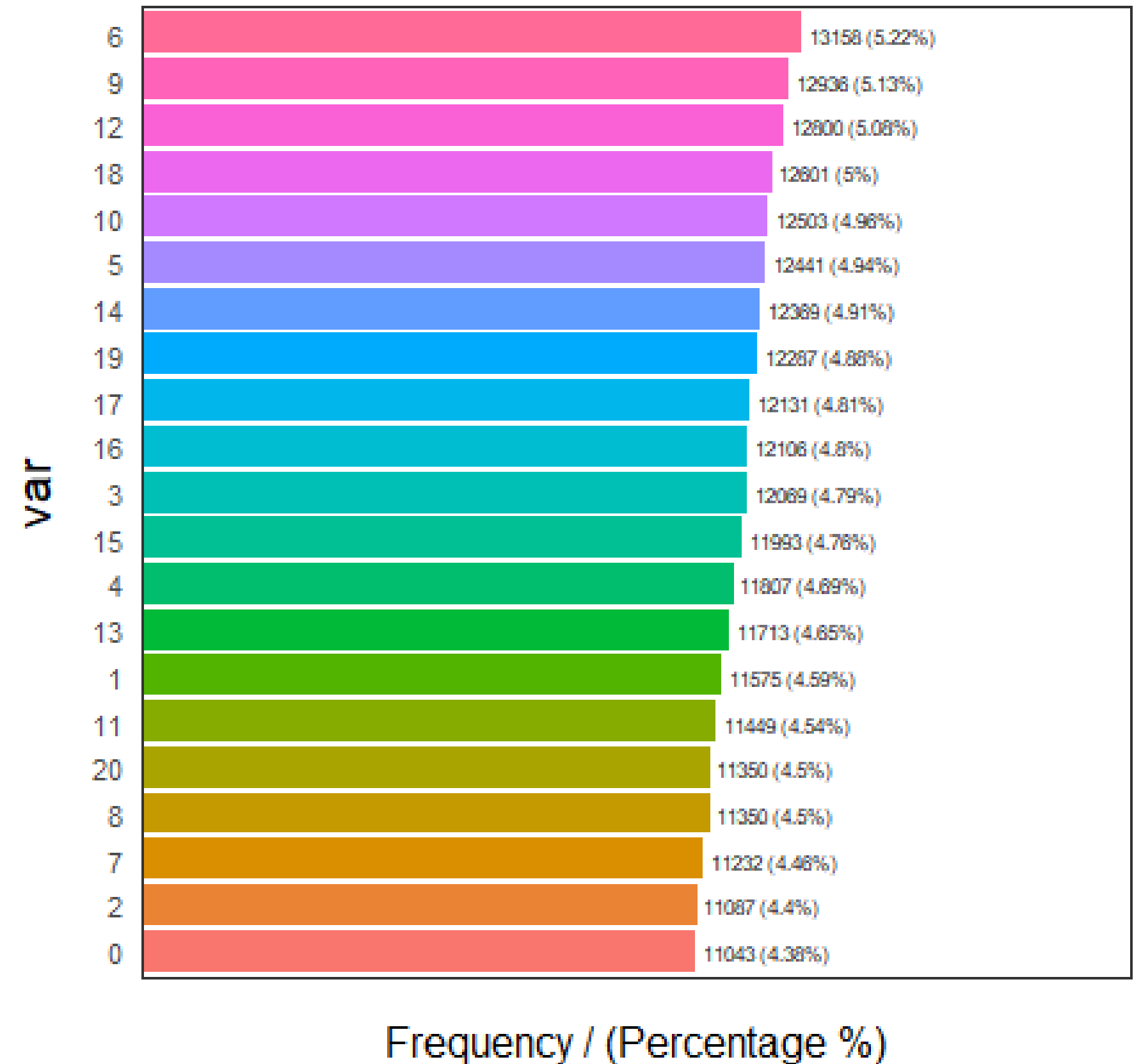
Podemos ver que 89,79%
dos clientes são solteiros e
19,21% são casados.



Frequência da variável Idade

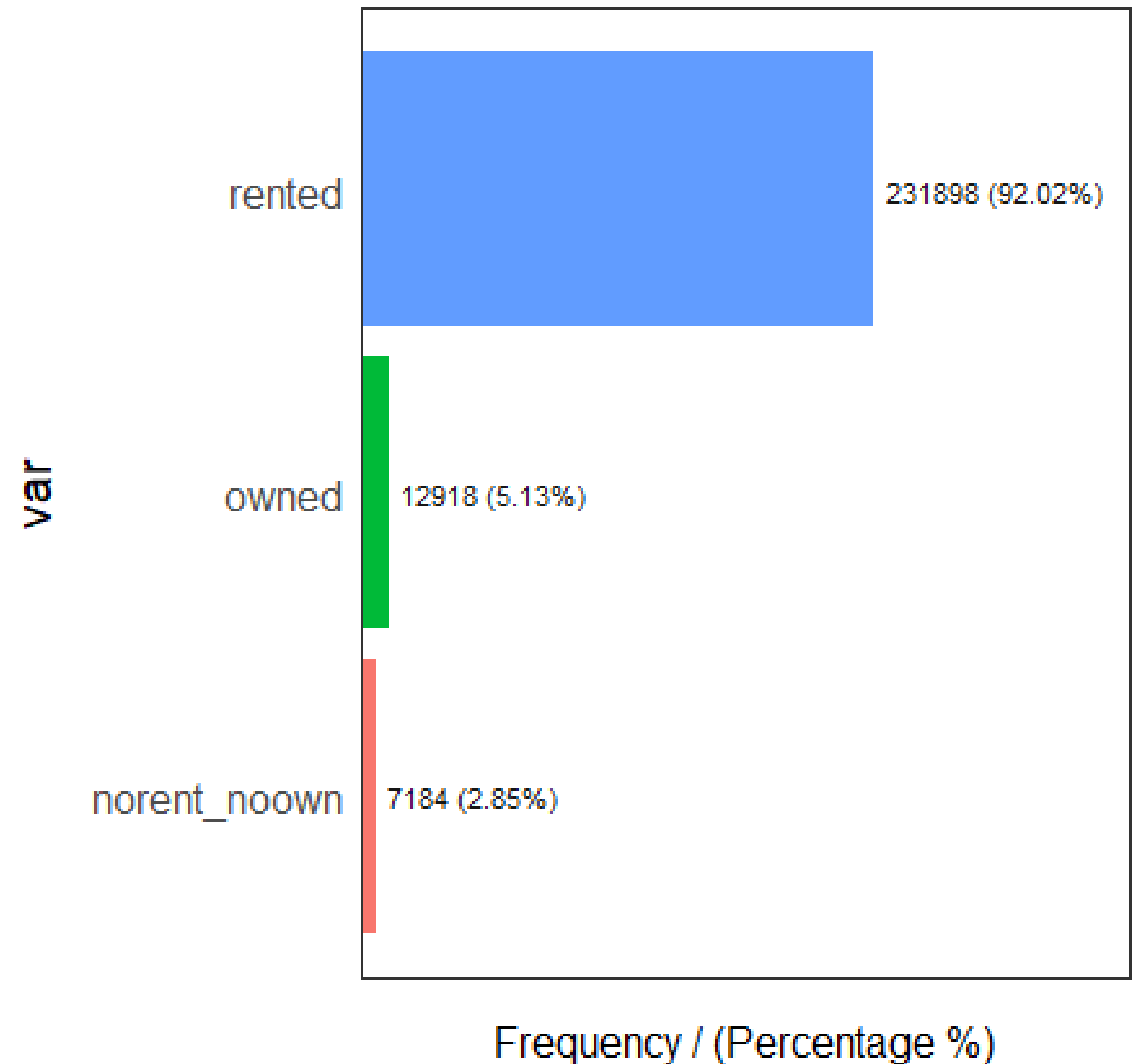


Frequência do
tempo (em anos)
em que cada
cliente pertence à
organização



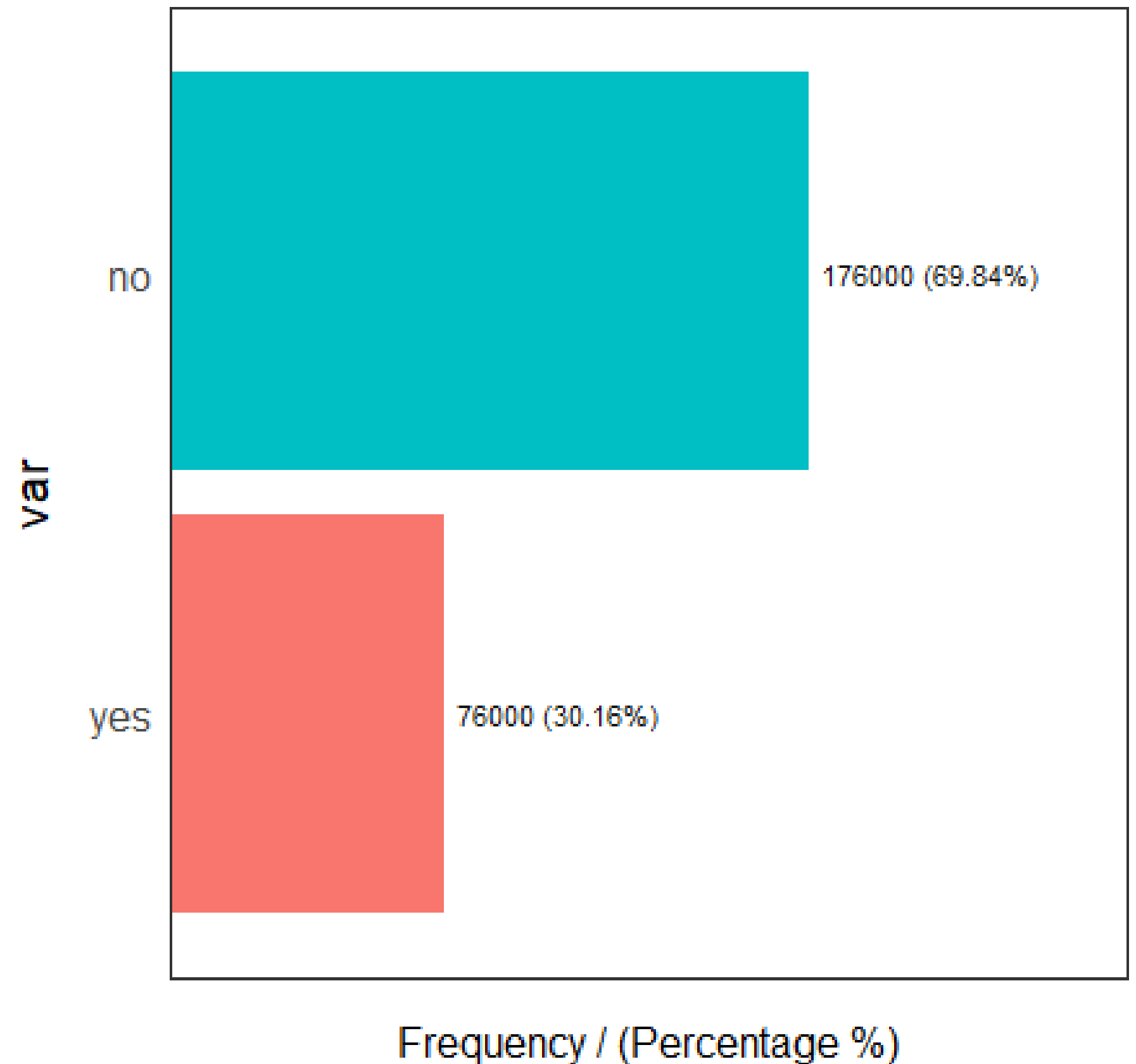
Frequência dos clientes que possuem residência ou moram de aluguel

Podemos ver que aproximadamente
92% dos clientes moram de aluguel



Frequência de clientes que possuem carro ou não

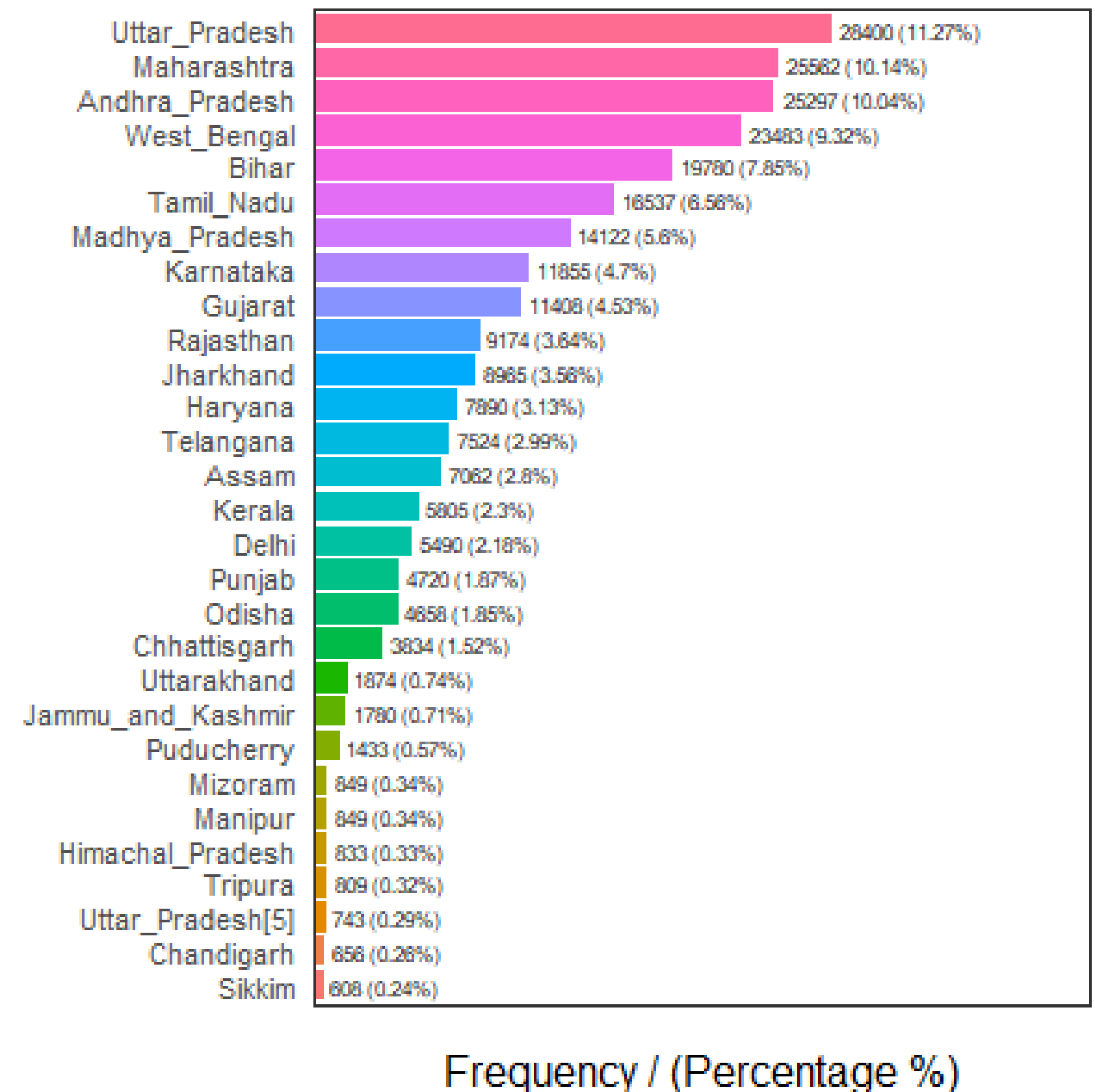
Vemos que
aproximadamente 70% dos
clientes não possuem carro



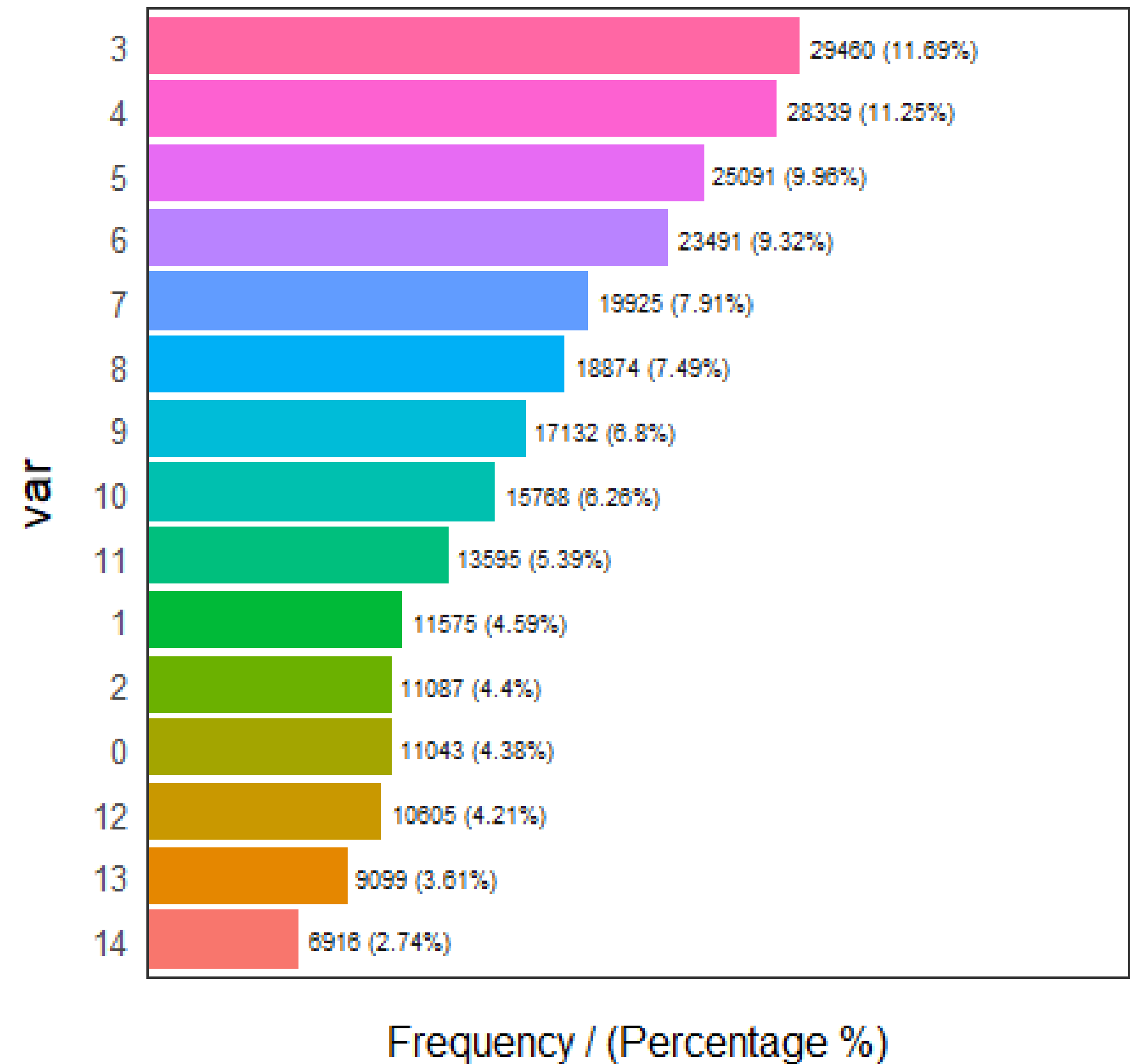
Frequência de clientes por Estado.

Podemos ver que há mais clientes nos estados de Utta, Maharashtra e Andhra.

var

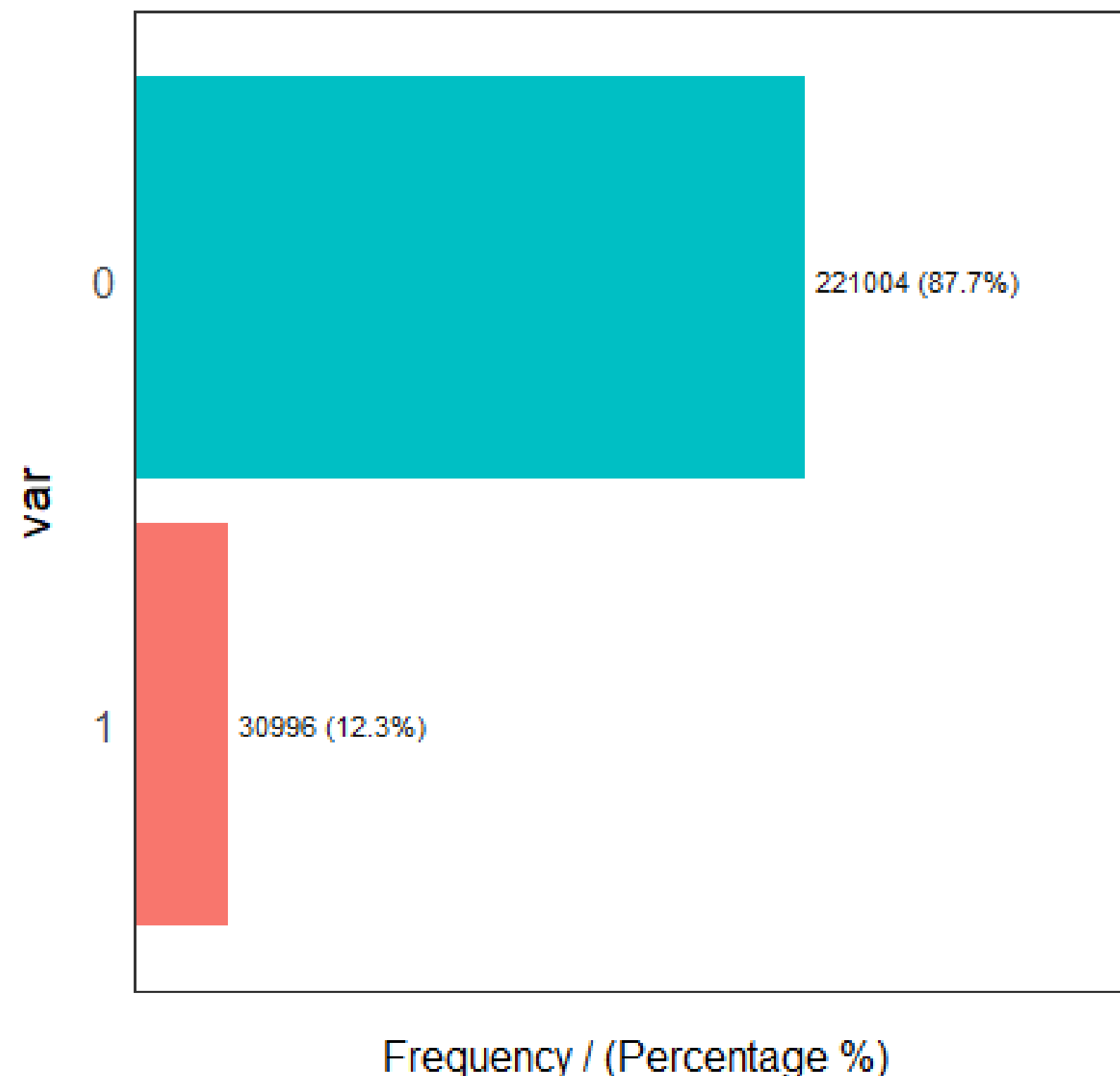


Frequência dos anos de trabalho de cada cliente



Frequência dos clientes que são inadimplentes ou não

zero se refere aos clientes adimplentes, vemos que 87% aproximadamente são bons pagadores



TESTE QUI- QUADRADO

Nossa variável resposta é risk_flag (bandeira de risco)

Assumindo as Hipóteses como:



H0: As variáveis são independentes.



H1: As variáveis são dependentes.

Para estado civil o resultado foi:

$X\text{-squared} = 111.89, df = 1, p\text{-value} < 0.00000000000000000022$

Assim, rejeitamos H_0 e concluimos que o estado civil influencia ao risco de crédito

Para idade o resultado do teste foi:

$X\text{-squared} = 776.37, df = 58, p\text{-value} < 0.00000000000000000022$

Podemos concluir que a idade influencia no risco de crédito.

Para a variável possuir casa ou aluguel temos:

$X\text{-squared} = 182.99, df = 2, p\text{-value} < 0.00000000000000000022$

Concluimos que morar de aluguel ou possuir moradia influencia no risco de crédito.

Para a variável estado o resultado foi:

$X\text{-squared} = 733.4, df = 28, p\text{-value} < 0.00000000000000000022$

Podemos concluir que o estado em que o cliente mora influencia no risco de crédito.

Para a variável anos de trabalho temos:

$X\text{-squared} = 733.4, df = 28, p\text{-value} < 0.00000000000000000022$

Concluimos que os anos de trabalho possuem influência no risco de crédito.

Para a variável anos de casa temos:

$X\text{-squared} = 6.115, df = 4, p\text{-value} = 0.1907$

Podemos concluir que esta variável não influencia no risco de crédito.

CONCLUSÃO

Concluimos que pelo teste Qui-Quadrado as variáveis, estado civil, idade, possuir casa, estado e anos de trabalho influenciam no risco de crédito.