Como vota Deputado? Uma análise votação que barrou o financiamento de campanhas por Pessoas Jurídicas na PEC 182/2007

Francisco Antonio Sousa de Araújo

11 jul, 2022 (atualizado: 17 jul, 2022)

# Introdução

Além desta introdução, este capítulo é composto por mais seis. O próximo trata da Teoria Espacial do Voto, o que permite a construção dos mapas de votação. Depois é feito o levantamento bibliográfico com os trabalhos correlatos ao tema. A seção seguinte descreve a metodologia utilizada. O capítulo 5 descreve a base de dados e traça um panorama geral da legislatura estudada. Na seção 6 são discutidos os resultados da pesquisa, seguido das considerações finais.

# Revisão Bibliográfica

# Metodologia

## Modelo

Na tentativa de encontrar quais os fatores mais correlacionados com o voto dos Deputados foram elencadas algumas variáveis relacionadas às características pessoais, aos aspectos políticos e de financiamento de campanha dos parlamentares. Estimou-se o seguinte modelo linear.

Onde é um vetor representando o voto de cada dos ’s Deputados, , e , são vetores dos coeficientes a serem estimados, é o vetor dos distúrbios não observados e as variáeveis entre parênteses estão descritas no Quadro 3.1.

| **Variável** | **Descrição** |
| --- | --- |
| EMA22 | Variável com valor igual a 1 caso o Deputado tenha votado SIM e 0 caso tenha votado NÃO. |
| **Pessoais** |  |
| Superior | *Dummy* para Deputado com grau de escolaridade de nível superior - valor 1 para candidatos com nível superior completo e valor 0 caso contrário. |
| Casado | *Dummy* para o parlamentar casado - valor 1 para casado e 0 para os demais. |
| Político | *Dummy* para Deputado cuja ocupação declarada ao TSE no pleito de 2014 - valor 1 para era deputado, senador ou vereador e valor 0 caso contrário. |
| **Políticas** |  |
| Ideologia | Variável construída no capítulo anterior pelo método W-NOMINATE. |
| Orientação Contra | *Dummy* para o parlamentar cujo partido orientou a votar NÃO na EMA 22. |
| Oposição | *Dummy* para os parlamentares que não faziam parte da coligação vencedora nas eleições de 2014 - valor 1 caso pertencesse e 0 caso contrário. |
| Região | *Dummy* para a região a qual a UF cujo Deputado foi eleito pertence. |
| **Financiamento** |  |
| Rec.PJ | Percentual da Receita Total de campanha advinda de doações de pessoas jurídicas. |
| Rec.PF | Percentual da Receita Total de campanha advinda de doações de pessoas físicas. |
| Rec.Próprios | Percentual da Receita Total de campanha advinda de recursos próprios do Deputado. |
| Rec.Partido | Percentual da Receita Total de campanha advinda do partido político do parlamentar. |

## Método

Devido a existência de uma grande quantidade de variáveis disponíveis e nenhum arcabouço teórico consolidado sobre o tema optou-se por utilizar um método de *Machine Learning* (ML) conhecido como *Elastic Net*.

O conjunto de explicativas utilizado para o ML foi composto pela ideologia do parlamentar, a orientação partidária em relação a votação em questão, se o Deputado mudou de partido após a eleição até o momento da votação, se o Deputado era um suplente, o total de votos obtidos na eleição, e se o congressista fazia parte da coligação do Governo. Também foram incluídos o patrimônio declarado à Justiça Eleitoral, os valores absolutos das doações feitas por PF, PJ, outros candidatos e comitês, pelo partido político do próprio Deputado, os recursos próprios utilizados como receita de campanha e também o valor total desta. Além de valores absolutos montou-se o percentual advindo de cada fonte. Como carcaterísticas pessoais mantiveram-se a idade, o sexo, o grau de instrução, o estado civil, a raça e a região a qual a UF cujo Deputado representava pertencia.

O ML consiste basicamente em três etapas. Inicialmente, divide-se a amostra em duas partes, aleatoriamente, uma chamada de teste e outra de treino. Utilizou-se a proporção de 75% para teste e 25% para treino, mantendo-se a proporção na variável resposta. Ou seja, na amostra como um todo cerca de 55% dos Deputados votou a favor da emenda e o restante contra. Essa proporcionalidade foi mantida na base de teste e de treino.

Depois, o modelo proposto é estimado utilizando a amostra de treino. Após a escolha da melhor especificação com base em métricas definidas pelo pesquisador, verifica-se o poder preditivo do modelo na base de teste. Esse processo é feito até que o melhor resultado seja obtido.

Algumas etapas auxiliares foram implementadas. Primeiramente, fez-se um pré-processamento dos dados na base de teste. Todas as variáveis categóricas foram transformadas em binárias. Depois, as variáveis monetárias, em valores absolutos, foram logaritmizadas. Aquelas cujo valor era zero foram transformadas em um, antes de aplicar o operador. Após, foram imputados valores para os dados ausentes, de acordo com o algoritmo de K Vizinhos mais próximos (KNN). Por fim, foram eliminadas as variáveis com variância próxima de zero.

Na etapa em que o modelo foi treinado, usou-se a validação cruzada conhecida como *CV k-Fold* que consiste em dividir a amostra *k* partes iguais para que sejam escolhidos os hiperparâmetros do modelo. Cada método de estimação tem um conjunto de hiperparâmetros a serem definidos pelo pesquisador. Nesta pesquisa, o *k* escolhido foi igual a 4. Dessa forma, separou-se a primeira parte () e juntaram-se as demais ( + + ) onde o modelo foi estimado. Com base nos parâmetros obtidos, testou-se a previsão do modelo em . Isso é feito até que todas as subamostras sejam utilizadas como teste. Depois, associado a cada combinação de hiperparâmetros, calcula-se uma medida de qualidade média para as *k* subamostras.

O algoritmo de ML utilizado nesta pesquisa é uma generalização dos mínimos quadrados ordinários, onde é aplicada uma penalização ao número de coeficientes angulares estimados. Portanto, é feita a seguinte minimização:

Onde é o hiperparâmetro de penalização e é o hiperparâmetro de mistura entre os métodos *Ridge* () ou *Lasso* (). Como destacado, seus valores são escolhidos pelo pesquisador. Além disso, representa as variáveis explicativas, a variável resposta e . Por fim, e .

Voltando ao método de validação cruzada, buscou-se um conjunto de combinações (*grid*) entre e que alcançasse a melhor média, entre as *k* subamostras, da métrica de avaliação do modelo. Como o intuito desta pesquisa não é previsão, definiu-se como métrica o erro quadrático médio (EQM), outros modelos de previsão e classificação focam em acurácia, precisão e curva ROC (*Receiver Operating Characteristic*). Uma vez escolhidos os melhores hiperparâmetros, estimou-se o modelo no conjunto de teste, onde foram extraídas as variáveis mais importantes, a partir do valor dos parâmetros estimados e padronizados.

Além do método de ML, empregou-se uma regressão logarítmica usual. Como será visto, o valor de encontrado foi próximo de zero, permitindo que fosse realizado uma regressão clássica. Nesse caso, mantiveram-se as variáveis explicativas que foram consideradas importantes para explicar o voto do parlamentar. Assim, características pessoais como Sexo e Raça foram excluídas. Dentre as variáveis de financiamento, optou-se pelo uso dos percentuais em relação à receita total de campanha. O patrimônio foi desconsiderado por ter muitos dados ausentes e ter uma disparidade muito grande entre os Deputados. Quanto às políticas, aquela que indicava se o mesmo era suplente não foi mantida, porque o número de votantes nessa condição era pequeno, e também o total de votos obtidos.

Em suma, mantiveram-se as variáveis apresentadas no Quadro 3.1. A ideologia foi construída no capítulo 2. Optou-se pelo uso daquela construída a partir do W-NOMINATE por considerar o método mais consolidado na literatura sobre o tema. É importante ressaltar que a variável foi construída para toda a legislatura, portanto é fixa e só muda conforme a migração partidária. Espera-se que parlamentares mais à direita sejam favoráveis à EMA 22.

Quanto à orientação partidária contra a emenda, a hipótese é que esta seja capaz de influenciar a decisão do congressista. Para o coeficiente da variável binária que indica se o Deputado era da Oposição não há um sinal esperado a priori, uma vez que o Governo não se posicionou sobre o tema, liberando os Deputados para votarem conforme queiram.

Para as variáveis de financiamento, espera-se que aqueles cuja maior fonte de receita seja PJ votem pela aprovação da proposta. Para aqueles cuja maior fonte de financiamento advinda do próprio partido ou de recursos próprios espera-se um sinal negativo. A hipótese é de que ao barrar o financiamento PJ possa ser implementada uma barreira à entrada de novos concorrentes. Pelo mesmo motivo, de diminuir a concorrência, espera-se um parâmetro negativo para *dummy* que indica se o Deputado já era político.

Quanto às regiões, utilizou-se a Centro-Oeste como categoria de referência. Assim como as demais variáveis de características pessoais, não são esperados sinais a priori para os coeficientes estimados.

# Dados

Todas as informações obtidas provêm do portal de Dados Abertos da Câmara dos Deputados e do portal de Dados Abertos do Tribunal Superior Eleitoral (TSE). As variáveis referentes às características pessoais e de financiamento de campanha advém do TSE e foram montadas com base no pleito de 2014.

Os votos para a EMA 22 foram obtidos conforme a data e horário em que a votação foi feita. Como já apresentado, outras EMA’s trouxeram propostas de financiamento de campanha eleitoral para a discussão da PEC 182/2007. A variável dependente considerada nesta pesquisa trata apenas dos votos contra ou a favor da EMA 22.

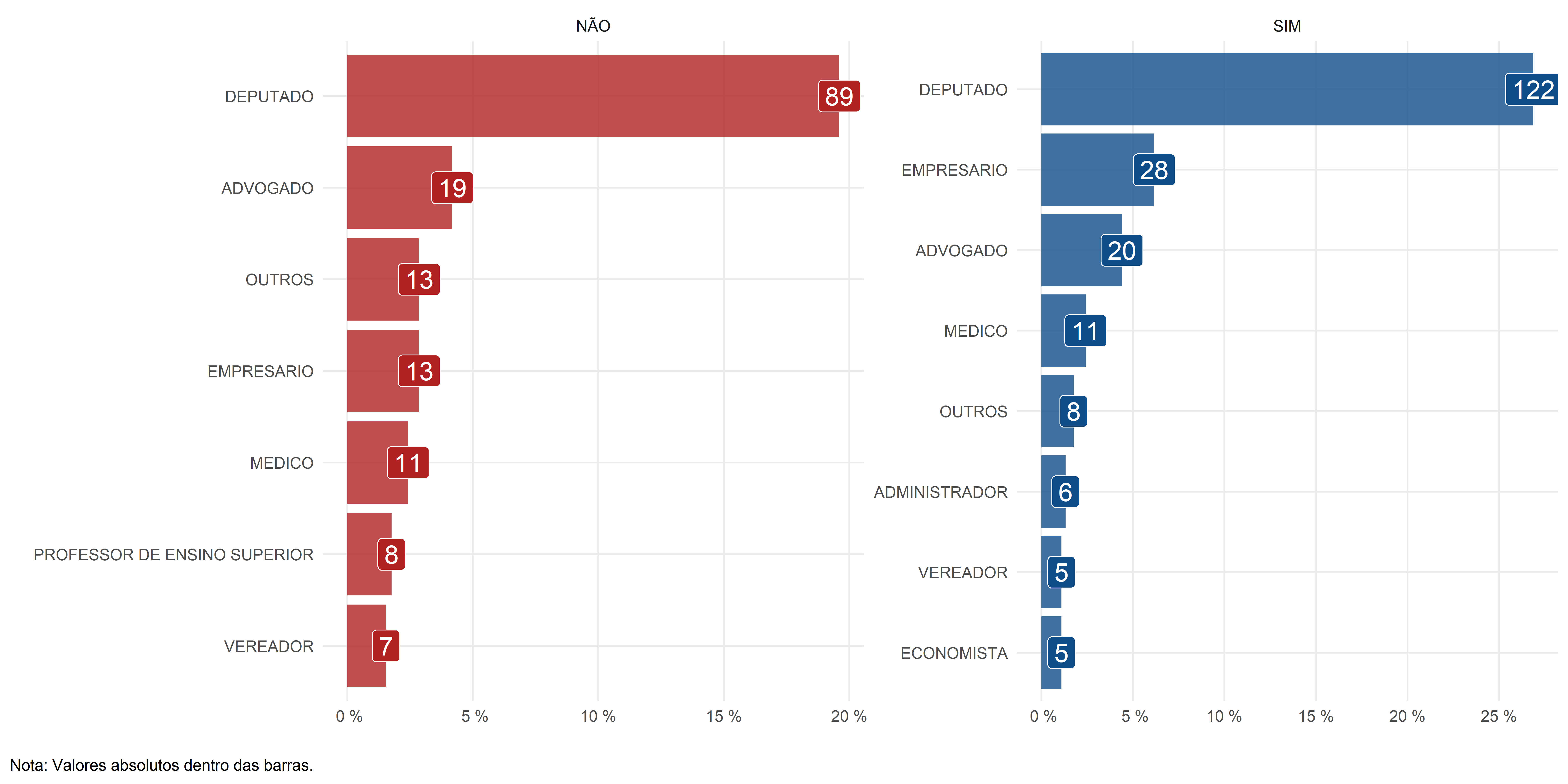
No total, foram 459 votos. Destes, 251 foram a favor, 203 contra, e 4 abstenções, além do presidente da Câmara que não votou. A votação ocorreu no dia 26 de maio de 2015, a partir das 23h33m, sendo a última do dia. Foram considerados nas estimações apenas votos Sim e Não.

É importante ressaltar que 30 Deputados não participaram desta votação mas no dia seguinte votaram a proposta sobre a EMA 10 e EMA 32, que também tratavam do tema finaciamento de campanha.

## Características Pessoais

Buscou-se traçar um perfil dos Deputados votantes, diferenciando-os de acordo com o voto. A Figura @ref(fig.id:ocupacao) traz as dez ocupações mais frequentes dos Deputados que votaram contra e a favor a emenda. Os percentuais indicados são sempre em relação aos 454 Deputados.

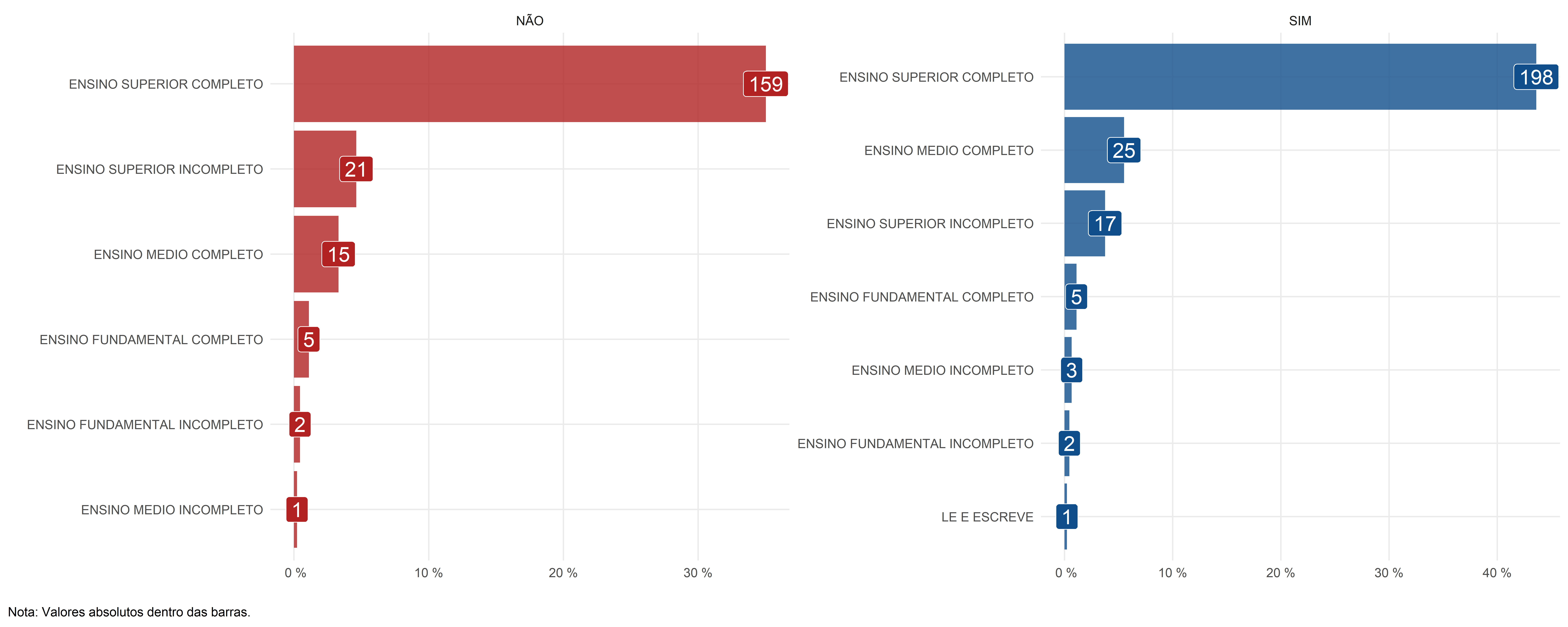
**Figura** : Dez principais ocupações dos votantes.



Mais de 25% de quem votou a favor já definiu sua ocupação, no momento da candidatura, como Deputado. Quase 20% de quem votou contra também já era congressista. Além disso, parece haver uma variabilidade maior nas ocupações do grupo que votou Sim. Por fim, é possível que estes percentuais estejam subestimados, uma vez que um Deputado pode considerar sua ocupação principal como advogado e declarar isso à justiça eleitoral.

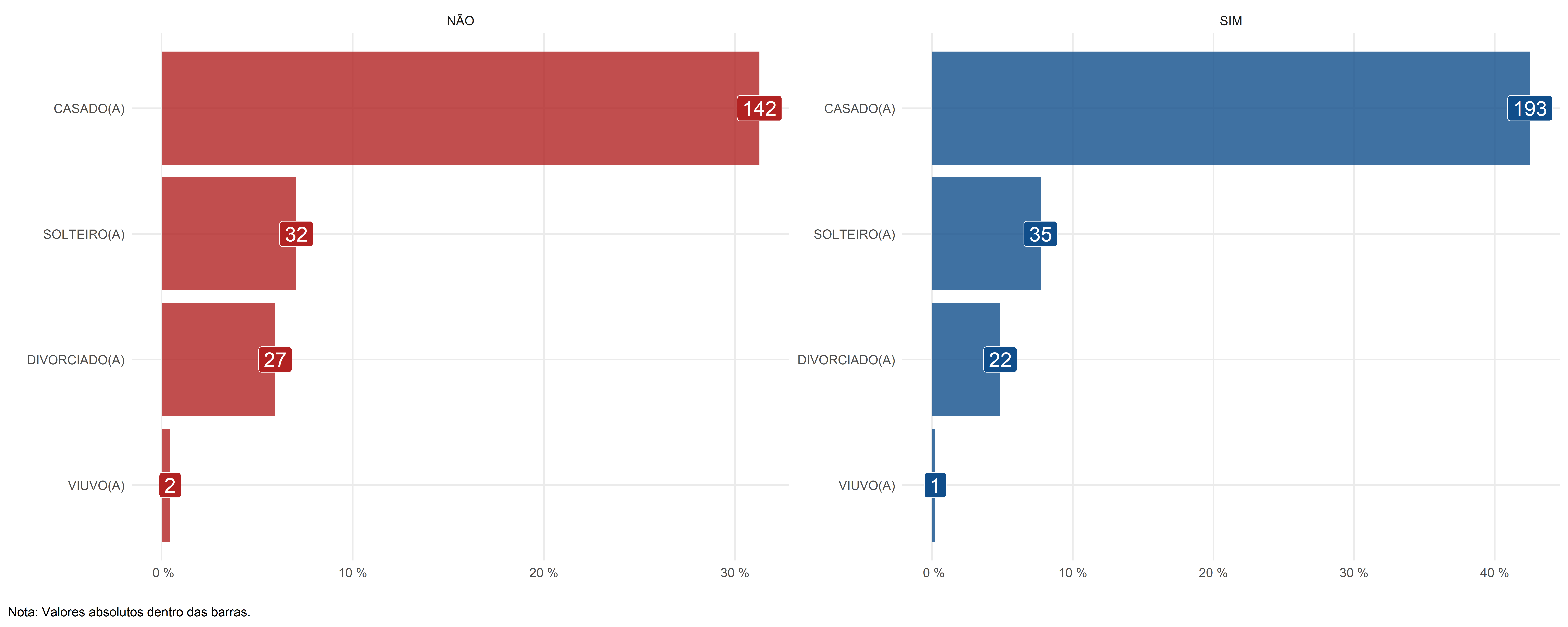
A Figura exibe a escolaridade dos Deputados. Embora a maioria, nos dois grupos, tenha ensino superior completo, chama atenção que um apenas lê e escreve.

**Figura** : Grau de instrução dos votantes.

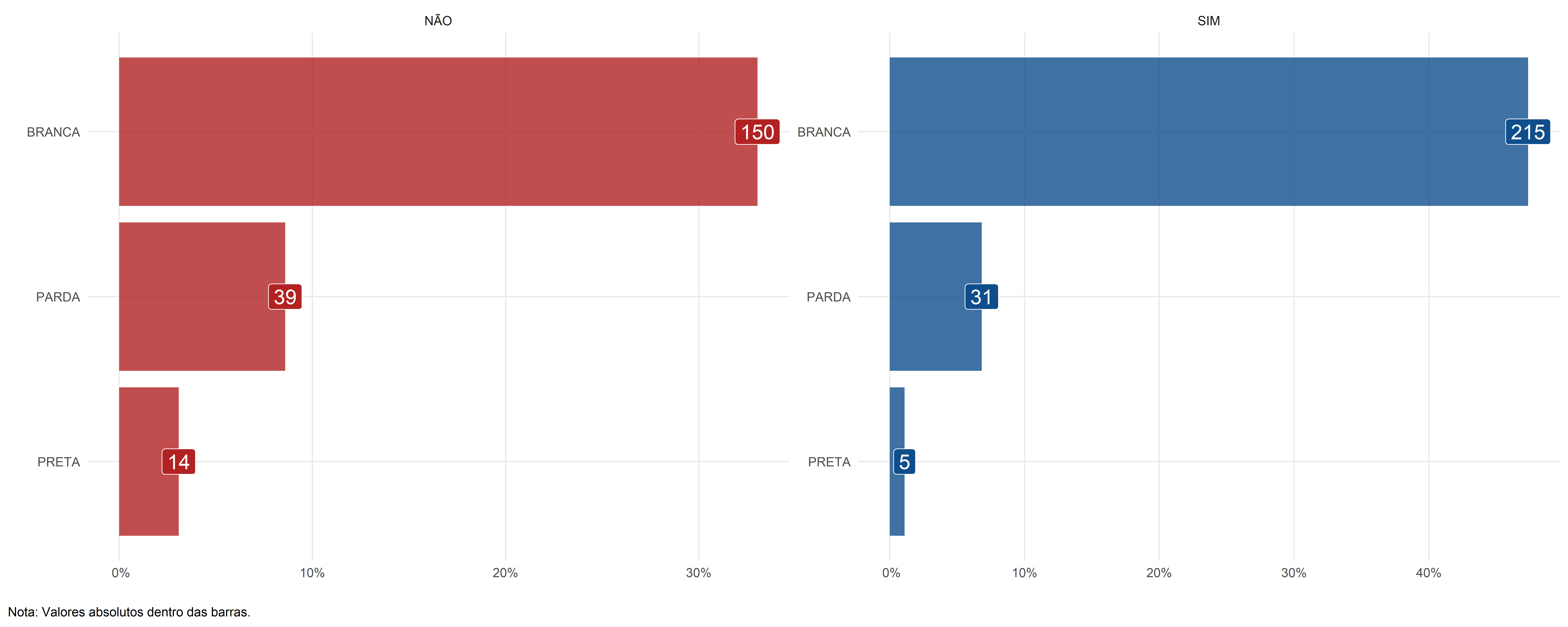


Quanto ao estado civil, Figura , mais de 70% são casados. A composição em ambos os grupos é similar. Além disso, observa-se pouca representatividade racial. Conforme a Figura , 80% são brancos. Não há nenhum indígena ou amarelo. Segundo a Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2019, aproximadamente 42% dos brasileiros se declararam como brancos, 46% como pardos, cerca de 9% como pretos, e 1% como amarelos ou indígenas.

**Figura** : Estado Civil dos votantes.

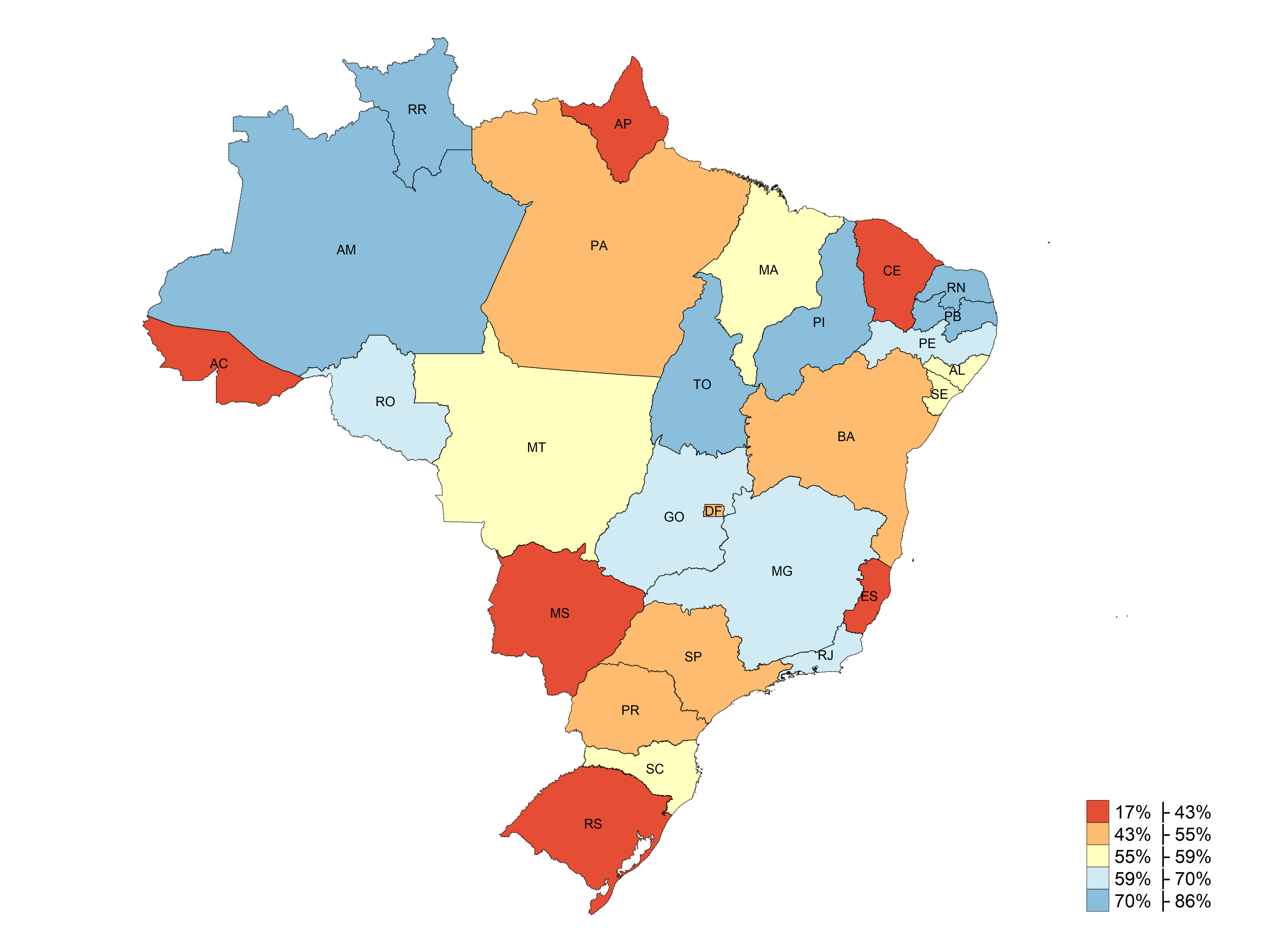


**Figura** : Raça dos votantes.



O mapa seguinte exibe o percentual de Deputados que votaram Sim para a EMA 22 dentro de cada UF. Os intervalos das classes são fechados à esquerda e foram construídos conforme os quintis. Dessa forma, as UF ’s em vermelho são aquelas cujos deputados mais votaram contra a emenda. No MS, por exemplo, apenas 1 Deputado, dos 6, votou a favor. Logo, cerca de 83% dos Deputados Federais do estado foram contra a proposta. No Ceará, 30% votaram a favor. Por outro lado, no TO, 86% (6 de 7) votaram Sim.

**Figura** : Percentual dos Deputados que votaram à favor, por UF.



Por fim, a Tabela sintetiza o grupo de variáveis através das suas estatísticas descritivas e compara os dois grupos. A variável patrimônio, em milhares de reais, é autodeclarada. Além de alguns casos de dados ausentes existem alguns casos “curiosos”. Alguns Deputados declararam ter apenas R$1.800,00. Por outro lado, o valor máximo foi de mais de R$ 1.000.000,00. O teste de médias apontou que os que votaram Sim têm um patrimônio maior do que os que votaram Não.

Quanto a ser político, fato que ocorre quando o parlamentar declara ser Deputado, Vereador ou Senador no momento da eleição, os grupos têm média estatisticamente igual.

A idade média também foi semelhante para os dois agrupamentos. O mínimo atingido foi de 23 anos (votou contra) e o máximo 84 anos (votou a favor). A média de pessoas do Sexo Masculino foi maior para os Deputados que votaram Sim, assim como a média de pessoas autodeclaradas da raça branca. Quanto a ter ensino superior e ser casado, a média de ambos os grupos não é estatisticamente diferente.

**Tabela** : Estatísticas Descritivas e Teste de Média, características pessoais.

| Estatísticas | Não, N = 203 | Sim, N = 251 | Valor p |
| --- | --- | --- | --- |
| Patrimônio |  |  | 0,012 |
| Média | 1.688,39 | 3.266,57 |  |
| Desvio Padrão | 3.554,67 | 8.862,84 |  |
| Mínimo | 0,18 | 10,75 |  |
| Máximo | 24.522,00 | 108.581,71 |  |
| Político |  |  | 0,3 |
| Média | 0,47 | 0,52 |  |
| Desvio Padrão | 0,50 | 0,50 |  |
| Mínimo | 0,00 | 0,00 |  |
| Máximo | 1,00 | 1,00 |  |
| Idade |  |  | 0,7 |
| Média | 50,83 | 51,24 |  |
| Desvio Padrão | 11,16 | 11,94 |  |
| Mínimo | 23,00 | 25,00 |  |
| Máximo | 75,00 | 84,00 |  |
| Sexo Masculino |  |  | 0,002 |
| Média | 0,86 | 0,94 |  |
| Desvio Padrão | 0,35 | 0,23 |  |
| Mínimo | 0,00 | 0,00 |  |
| Máximo | 1,00 | 1,00 |  |
| Ensino Superior |  |  | 0,9 |
| Média | 0,78 | 0,79 |  |
| Desvio Padrão | 0,41 | 0,41 |  |
| Mínimo | 0,00 | 0,00 |  |
| Máximo | 1,00 | 1,00 |  |
| Casado |  |  | 0,10 |
| Média | 0,70 | 0,77 |  |
| Desvio Padrão | 0,46 | 0,42 |  |
| Mínimo | 0,00 | 0,00 |  |
| Máximo | 1,00 | 1,00 |  |
| Branco |  |  | 0,002 |
| Média | 0,74 | 0,86 |  |
| Desvio Padrão | 0,44 | 0,35 |  |
| Mínimo | 0,00 | 0,00 |  |
| Máximo | 1,00 | 1,00 |  |

## Características Políticas

## Financiamento de Camoanha

E

# Resultados

Nesta seção serão apresentados os pontos ideais, estimados pelos métodos descritos e suas respectivas métricas que permitem inferir o quão adequados são. Buscar-se-á determinar o número de dimensões importantes para explicar a totalidade das votações e uma interpretação para estas. Por fim, uma vez definidos o mapa de votação espacial para cada um dos métodos, será proposto um número ótimo, a partir de medidas de similaridade, de partidos políticos na legislatura em questão.

## *Elastic Net*

## *Logit*

# Considerações Finais

# Referências

# Apêndice