

** Script em R utilizado para a análise do banco de migrações**

- São retirados os parlamentares com *status Primeiro mandato* para considerar apenas os **Migrantes e Não Migrantes** possibilitando avaliar, efetivamente, aqueles que mudaram de partido;
- O segundo filtro é aplicado em razão da identificação de partidos de **Esquerda** tem fluxo de parlamentares entrando ou saindo muito baixo em comparação com direita e centro, de modo que sem estes o ajuste do modelo (sua capacidade preditiva) aumenta, sem prejuízo para as análises;
- Aqui estão descritas apenas duas etapas da análise dos dados: 1. a classificação dos dados; e 2. os testes de hipótese;

Contém os comandos básicos para organização e limpeza dos dados (*tidy*)

```
library(tidyverse)
```

Contém os comandos para limpeza de texto

```
library(stringr)
```

Contém os comandos para o modelo de regressão logística

```
library(stats)
```

1) Organização dos dados

O banco de dados utilizado como base foi coletado por Sainz, N. & Codato, A. (2025), considerando os deputados titulares eleitos entre os anos de 1988 e 2022 a partir de dados do Tribunal Superior Eleitoral.

O projeto ocorre no âmbito do INCT Redem (Representação e Legitimidade Democrática).

```
banco_de_migrações <-  
read_xlsx("C:/Users/felip/Downloads/BD_Sainz&Codato_ClassePoliticadoBrasil_INCT_ReDem  
_2025.xlsx")
```

É criada uma coluna contendo a ideologia dos partidos, baseado em Coppedge (1997), Wiesehomeier e Benoit, (2007) Power e Zucco (2011), Tarouco e Madeira (2015), e Bolognesi, Ribeiro e Codato (2022), **de maneira agregada**.

```
banco_de_migrações_cap <- banco_de_migrações_cap |>  
mutate(  
  Ideologia = case_when(  
    Sigla_Partido == "PT DO B" ~ "Direita",  
    Sigla_Partido == "PTdoB" ~ "Direita",  
    Sigla_Partido == "AVANTE" ~ "Direita",  
    Sigla_Partido == "PPS" ~ "Centro",
```

```
Sigla_Partido == "CIDADANIA" ~ "Centro",
Sigla_Partido == "PSDC" ~ "Direita",
Sigla_Partido == "DC" ~ "Direita",
Sigla_Partido == "PFL" ~ "Direita",
Sigla_Partido == "DEM" ~ "Direita",
Sigla_Partido == "PMDB" ~ "Centro",
Sigla_Partido == "MDB" ~ "Centro",
Sigla_Partido == "NOVO" ~ "Direita",
Sigla_Partido == "PAN" ~ "Direita",
Sigla_Partido == "PEN" ~ "Direita",
Sigla_Partido == "PATRIOTA" ~ "Direita",
Sigla_Partido == "PC DO B" ~ "Esquerda",
Sigla_Partido == "PCdoB" ~ "Esquerda",
Sigla_Partido == "PDT" ~ "Esquerda",
Sigla_Partido == "PHS" ~ "Direita",
Sigla_Partido == "PR" ~ "Direita",
Sigla_Partido == "PL" ~ "Direita",
Sigla_Partido == "PMN" ~ "Direita",
Sigla_Partido == "PTN" ~ "Direita",
Sigla_Partido == "PODE" ~ "Direita",
Sigla_Partido == "PPB" ~ "Direita",
Sigla_Partido == "PP" ~ "Direita",
Sigla_Partido == "PP**" ~ "Direita",
Sigla_Partido == "PL*" ~ "Direita",
Sigla_Partido == "PPL" ~ "Direita",
Sigla_Partido == "PRB" ~ "Direita",
Sigla_Partido == "PMR" ~ "Direita",
Sigla_Partido == "REPUBLICANOS" ~ "Direita",
Sigla_Partido == "PRONA" ~ "Direita",
Sigla_Partido == "PROS" ~ "Direita",
Sigla_Partido == "PRP" ~ "Direita",
Sigla_Partido == "PRTB" ~ "Direita",
Sigla_Partido == "PSB" ~ "Centro",
Sigla_Partido == "PSC" ~ "Direita",
Sigla_Partido == "PSD" ~ "Direita",
Sigla_Partido == "PSDB" ~ "Centro",
Sigla_Partido == "PSL" ~ "Direita",
Sigla_Partido == "PSOL" ~ "Esquerda",
Sigla_Partido == "PST" ~ "Direita",
Sigla_Partido == "PT" ~ "Esquerda",
Sigla_Partido == "PTB" ~ "Direita",
Sigla_Partido == "PTC" ~ "Direita",
Sigla_Partido == "PTN" ~ "Direita",
Sigla_Partido == "PV" ~ "Centro",
Sigla_Partido == "REDE" ~ "Centro",
Sigla_Partido == "SD" ~ "Direita",
Sigla_Partido == "SDD" ~ "Direita",
Sigla_Partido == "SOLIDARIEDADE" ~ "Direita",
Sigla_Partido == "PRD" ~ "Direita",
Sigla_Partido == "PMB" ~ "Direita",
Sigla_Partido == "PATRI" ~ "Direita",
Sigla_Partido == "UNIÃO" ~ "Direita",
```

```
TRUE ~ "Indefinido"
```

```
) ,
```

É criada uma coluna contendo o tamanho dos partidos ao longo do tempo, controlado pela **bancada no início da legislatura**

```
Tamanho = case_when(
```

==== LEGISLATURA 57 (2022-2026) ====

```
# Grandes (50+)
```

```
Sigla_Partido %in% c("PL", "PT", "UNIÃO") & idLegislatura == 57 ~ "Grande",
```

```
# Médios (21-49)
```

```
Sigla_Partido %in% c("REPUBLICANOS", "PSD", "MDB", "PP") & idLegislatura == 57  
~ "Médio",
```

```
# Pequenos (5-20)
```

```
Sigla_Partido %in% c("CIDADANIA", "PRD", "AGIR", "PSC", "PSOL", "AVANTE",  
"PODE", "PSB", "PDT", "PSDB") & idLegislatura == 57 ~ "Pequeno",
```

```
# Nanicos (até 4)
```

```
Sigla_Partido %in% c("PTB", "PC DO B", "PV", "REDE", "PMB", "UP", "DC",  
"PATRIOTA", "PROS", "SOLIDARIEDADE", "NOVO") & idLegislatura == 57 ~ "Nanico",
```

==== LEGISLATURA 56 (2019-2023) ====

```
# Grandes (50+)
```

```
Sigla_Partido %in% c("PT", "PSL", "PL") & idLegislatura == 56 ~ "Grande",
```

```
# Médios (21-49)
```

```
Sigla_Partido %in% c("PP", "PSD", "MDB", "PR", "PSB", "PRB", "DEM", "PSDB",  
"PDT", "UNIÃO") & idLegislatura == 56 ~ "Médio",
```

```
# Pequenos (5-20)
```

```
Sigla_Partido %in% c("SOLIDARIEDADE", "PODE", "PSOL", "PTB", "PCdoB", "NOVO",  
"PPS", "PROS", "PSC", "AVANTE", "PHS", "PC DO B") & idLegislatura == 56 ~ "Pequeno",
```

```
# Nanicos (até 4)
```

```
Sigla_Partido %in% c("PATRI", "PRP", "PV", "PMN", "PTC", "DC", "PPL", "REDE") &  
idLegislatura == 56 ~ "Nanico",
```

==== LEGISLATURA 55 (2015-2019) ====

```
# Grandes (50+)
```

```
Sigla_Partido %in% c("PT", "PMDB", "PSDB", "MDB") & idLegislatura == 55 ~  
"Grande",
```

```
# Médios (21-49)
```

```
Sigla_Partido %in% c("PP**", "PSD", "PSB", "PR", "PTB", "PRB", "DEM", "PDT",
```

```
"UNIÃO") & idLegislatura == 55 ~ "Médio",
# Pequenos (5-20)
Sigla_Partido %in% c("SD", "PSC", "PROS", "PPS", "PCdoB", "PV", "PSOL", "PHS",
"SOLIDARIEDADE", "PMB", "REDE", "PC DO B") & idLegislatura == 55 ~ "Pequeno",
# Nanicos (até 4)
Sigla_Partido %in% c("PTN", "PRP", "PMN", "PEN", "PSDC", "PTC", "PTdoB", "PSL",
"PRTB", "AVANTE", "PODE", "PATRI", "PPL", "PATRIOTA", "PT DO B") & idLegislatura ==
55 ~ "Nanico",
```

==== LEGISLATURA 54 (2011-2015) ===

```
# Grandes (50+)
Sigla_Partido %in% c("PT", "PMDB", "PSDB", "PSD") & idLegislatura == 54 ~
"Grande",
# Médios (21-49)
Sigla_Partido %in% c("PP**", "DEM", "PR", "PSB", "PDT", "PTB") & idLegislatura ==
54 ~ "Médio",
# Pequenos (5-20)
Sigla_Partido %in% c("PSC", "PCdoB", "PV", "PPS", "PRB", "SD", "SDD", "PROS",
"PC DO B") & idLegislatura == 54 ~ "Pequeno",
# Nanicos (até 4)
Sigla_Partido %in% c("PMN", "PTdoB", "PSOL", "PHS", "PRTB", "PRP", "PTC",
"PSL", "PEN", "PSDC", "PT DO B") & idLegislatura == 54 ~ "Nanico",
```

==== LEGISLATURA 53 (2007-2011) ===

```
# Grandes (50+)
Sigla_Partido %in% c("PMDB", "PT", "PSDB", "PFL", "DEM") & idLegislatura == 53
~ "Grande",
# Médios (21-49)
Sigla_Partido %in% c("PP**", "PR", "PSB", "PDT", "PTB") & idLegislatura == 53 ~
"Médio",
# Pequenos (5-20)
Sigla_Partido %in% c("PPS", "PV", "PCdoB", "PSC", "PC DO B") & idLegislatura ==
53 ~ "Pequeno",
# Nanicos (até 4)
Sigla_Partido %in% c("PAN", "PSOL", "PMN", "PTC", "PHS", "PTdoB", "PRB",
"PRTB", "PRONA", "PT DO B") & idLegislatura == 53 ~ "Nanico",
```

==== LEGISLATURA 52 (2003-2007) ===

```
# Grandes (50+)
Sigla_Partido %in% c("PT", "PFL", "PMDB", "PSDB") & idLegislatura == 52 ~
```

```

"Grande",

# Médios (21-49)
Sigla_Partido %in% c("PPB", "PTB", "PL*", "PSB", "PPS", "PP**", "PR") &
idLegislatura == 52 ~ "Médio",

# Pequenos (5-20)
Sigla_Partido %in% c("PDT", "PCdoB", "PRONA", "PV", "PC DO B") & idLegislatura
== 52 ~ "Pequeno",

# Nanicos (até 4)
Sigla_Partido %in% c("PMN", "PSC", "PSL", "PHS", "PSOL", "PST", "PTC", "PMR",
"PRB", "PRP", "PSD", "PSDC") & idLegislatura == 52 ~ "Nanico",

```

==== LEGISLATURA 51 (1999-2003) ====

```

# Grandes (50+)
Sigla_Partido %in% c("PFL", "PSDB", "PMDB", "PPB", "PT") & idLegislatura == 51
~ "Grande",

# Médios (21-49)
Sigla_Partido %in% c("PTB", "PDT") & idLegislatura == 51 ~ "Médio",

# Pequenos (5-20)
Sigla_Partido %in% c("PSB", "PL", "PCdoB", "PC DO B") & idLegislatura == 51 ~
"Pequeno",

# Nanicos (até 4)
Sigla_Partido %in% c("PPS", "PSD", "PMN", "PSC", "PRONA", "PSL", "PST", "PV") &
idLegislatura == 51 ~ "Nanico",

# Padrão para partidos não listados ou legislaturas futuras
TRUE ~ "Indefinido"
)
)

```

Por fim, é criada a coluna para saber se o deputado foi reeleito ou não:

São retirados as acentuações e pontuações de nomes de deputados para facilitar sua integração

```

banco_de_migrações_cap$Nome_Deputado <-
stri_trans_general(banco_de_migrações_cap$Nome_Deputado, "LATIN-ASCII")

```

Os dados são ordenados por ordem alfabética e temporalmente

```

banco_de_migrações_cap <- banco_de_migrações_cap |>
arrange(Nome_Deputado, idLegislatura)

```

Para definição dos parlamentares que brigam são aplicados alguns filtros lógicos

```

banco_de_migrações_cap <- banco_de_migrações_cap %>%
  group_by(Nome_Deputado) %>%
  mutate(
    primeira_aparicao = row_number() == 1, # É criada uma condição para a primeira vez que cada nome aparece no banco

    mudança_legitima = case_when(
      primeira_aparicao ~ FALSE, # Verifica se não é a primeira mudança
      lag(Sigla_Partido) == Sigla_Partido ~ FALSE, # Verifica se o partido associado ao deputado é diferente do partido anterior
  )

```

- Aqui é criada uma condição complementar, em que se verifica os partidos que foram aglutinados, fundiram-se, ou mesmo apenas mudaram de nome, verificando as siglas numa posição anterior com relação à posição atual, controlado pelo ano:

== MUDANÇAS DE NOME - usar período temporal correto ==

PSN -> PHS (mudança em 2000, aparece na eleição 2002)

```
(lag(Sigla_Partido) == "PSN" & Sigla_Partido == "PHS" &
  Ano_Eleição == 2002) ~ FALSE,
```

PRN -> PTC (mudança em 2001, aparece na eleição 2002)

```
(lag(Sigla_Partido) == "PRN" & Sigla_Partido == "PTC" &
  Ano_Eleição == 2002) ~ FALSE,
```

PPB -> PP (mudança em 2003, aparece na eleição 2006)

```
(lag(Sigla_Partido) == "PPB" & Sigla_Partido == "PP" &
  Ano_Eleição == 2006) ~ FALSE,
```

PFL -> DEM (mudança em 2007, aparece na eleição 2010)

```
(lag(Sigla_Partido) == "PFL" & Sigla_Partido == "DEM" &
  Ano_Eleição == 2010) ~ FALSE,
```

PMR -> PRB (mudança em 2009, aparece na eleição 2010)

```
(lag(Sigla_Partido) == "PMR" & Sigla_Partido == "PRB" &
  Ano_Eleição == 2010) ~ FALSE,
```

PMDB -> MDB (mudança em 2018, aparece na eleição 2018)

```
(lag(Sigla_Partido) == "PMDB" & Sigla_Partido == "MDB" &
  Ano_Eleição == 2018) ~ FALSE,
```

PT DO B -> AVANTE (mudança em 2017, aparece na eleição 2018)

```
(lag(Sigla_Partido) == "PT DO B" & Sigla_Partido == "AVANTE" &
  Ano_Eleição == 2018) ~ FALSE,
```

PTN -> PODE (mudança em 2017, aparece na eleição 2018)

```
(lag(Sigla_Partido) == "PTN" & Sigla_Partido == "PODE" &
  Ano_Eleição == 2018) ~ FALSE,
```

PRB -> REPUBLICANOS (mudança em 2019, aparece na eleição 2022)

```
(lag(Sigla_Partido) == "PRB" & Sigla_Partido == "REPUBLICANOS" &
  Ano_Eleição == 2022) ~ FALSE,
```

PPS -> CIDADANIA (mudança em 2019, aparece na eleição 2022)

```
(lag(Sigla_Partido) == "PPS" & Sigla_Partido == "CIDADANIA" &
  Ano_Eleição == 2022) ~ FALSE,
```

PR -> PL (mudança em 2019, aparece na eleição 2022)

```
(lag(Sigla_Partido) == "PR" & Sigla_Partido == "PL" &
  Ano_Eleição == 2022) ~ FALSE,
```

SD -> SOLIDARIEDADE (mudança em 2018, aparece na eleição 2018)

```
(lag(Sigla_Partido) == "SD" & Sigla_Partido == "SOLIDARIEDADE" &
  Ano_Eleição >= 2018) ~ FALSE,
```

== FUSÕES ==

PL/PRONA -> PR (fusão em 2006, aparece na eleição 2006)

```
(lag(Sigla_Partido) %in% c("PL", "PRONA") & Sigla_Partido == "PR" &
  Ano_Eleição >= 2006) ~ FALSE,
```

DEM/PSL -> UNIÃO (fusão em 2022, aparece na eleição 2022)

```
(lag(Sigla_Partido) %in% c("DEM", "PSL") & Sigla_Partido == "UNIÃO" &
  Ano_Eleição >= 2022) ~ FALSE,
```

Outras mudanças são migrações reais

```
TRUE ~ TRUE
```

),

Feitos os filtros, agora aplicamos as condições aos deputados para criar as três possibilidades da variável:

1. “Primeiro mandato”; 2. “Migrante”; e 3. “Não migrante”:

Classificar o deputado

```
Migrantes = case_when(
  primeira_aparicao ~ "Primeiro mandato",
  mudança_legitima ~ "Migrante",
  !mudança_legitima ~ "Não migrante",
  TRUE ~ "Erro" #' Para debug
)
) %>%
  #' Remover colunas auxiliares utilizadas apenas como filtro
select(-primeira_aparicao, -mudança_legitima) %>%
ungroup()
```

2) Testes de hipótese

- Em todos os casos foram retirados da análise aqueles que aparecem em **Primeiro mandato** ou que estão em partidos à **Esquerda**. Como verificado na literatura e nos dados exploratórios deste capítulo, a esquerda tem chances consideravelmente menores de migrar, de modo que sua consideração pode gerar ruídos no modelo.

H1) Deputadas mulheres tem maiores chances de migrar para partidos grandes (Wylie, 2018)

```
modelo_gênero <- banco_arrumado %>%
  filter(Migrantes != "Primeiro mandato" & Ideologia != "Esquerda")
modelo_gênero$Migrantes <- as.factor(modelo_gênero$Migrantes)
```

Primeiro é testado se há **diferença estatisticamente significativa** de mulheres migrarem considerando **gênero** e **tamanho** apenas com efeitos somados:

```
regressão_gênero_tam <- glm(Migrantes ~ Gênero + Tamanho,
                               data = modelo_gênero, family = "binomial")
summary(regressão_gênero_tam)
```

```
##
## Call:
## glm(formula = Migrantes ~ Gênero + Tamanho, family = "binomial",
##      data = modelo_gênero)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.8092   -1.2693    0.7411    0.8140    1.0902
##
```

```

## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) 1.4201    0.2805   5.062 4.15e-07 ***
## GêneroMasculino -0.2680    0.2771  -0.967  0.3334
## TamanhoMédio    -0.2174    0.1343  -1.619  0.1055
## TamanhoNanico   -0.9435    0.4845  -1.947  0.0515 .
## TamanhoPequeno  -0.9386    0.2171  -4.324 1.53e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1523.1 on 1291 degrees of freedom
## Residual deviance: 1501.4 on 1287 degrees of freedom
## AIC: 1511.4
##
## Number of Fisher Scoring iterations: 4

```

É testado como **gênero** interage com **tamanho** na avaliação da significância estatística:

```

regressão_gênero_e_tam <- glm(Migrantes ~ Gênero * Tamanho,
                               data = modelo_gênero, family = "binomial")
summary(regressão_gênero_e_tam)

```

```

##
## Call:
## glm(formula = Migrantes ~ Gênero * Tamanho, family = "binomial",
##      data = modelo_gênero)
##
## Deviance Residuals:
##      Min       10      Median       30       Max
## -1.8802  -1.2435   0.7267   0.8217   1.1278
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) 0.7376    0.3666   2.012  0.0442 *
## GêneroMasculino 0.4592    0.3804   1.207  0.2274
## TamanhoMédio 0.8429    0.5538   1.522  0.1280
## TamanhoNanico 13.8285  882.7434  0.016  0.9875
## TamanhoPequeno 13.8285  441.3718  0.031  0.9750
## GêneroMasculino:TamanhoMédio -1.1272    0.5709  -1.974  0.0483 *
## GêneroMasculino:TamanhoNanico -14.9075  882.7436 -0.017  0.9865
## GêneroMasculino:TamanhoPequeno -14.8711  441.3719 -0.034  0.9731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1523.1 on 1291 degrees of freedom
## Residual deviance: 1492.2 on 1284 degrees of freedom
## AIC: 1508.2

```

```
##  
## Number of Fisher Scoring iterations: 13
```

H1.1) O efeito do gênero varia conforme o tamanho e a ideologia do partido:

```
regressão_gênero_tam_ideo <- glm(Migrantes ~ Gênero * Tamanho * Ideologia,  
                                   data = modelo_gênero, family = "binomial")  
summary(regressão_gênero_tam_ideo)
```

```
##  
## Call:  
## glm(formula = Migrantes ~ Gênero * Tamanho * Ideologia, family = "binomial",  
##       data = modelo_gênero)  
##  
## Deviance Residuals:  
##      Min        1Q     Median        3Q       Max  
## -1.9214   -0.9665    0.6335    0.8556    1.4042  
##  
## Coefficients: (1 not defined because of singularities)  
##                                     Estimate Std. Error z value  
Pr(>|z|)  
## (Intercept)                 1.2993    0.6513  1.995  
0.0461 *  
## GêneroMasculino             0.3558    0.6811  0.522  
0.6015  
## TamanhoMédio                0.3747    0.9056  0.414  
0.6790  
## TamanhoNanico               13.2668   882.7436  0.015  
0.9880  
## TamanhoPequeno              13.2668   624.1942  0.021  
0.9830  
## IdeologiaDireita            -0.8938    0.7953 -1.124  
0.2611  
## GêneroMasculino:TamanhoMédio -0.5257    0.9676 -0.543  
0.5869  
## GêneroMasculino:TamanhoNanico -29.4879  1019.3045 -0.029  
0.9769  
## GêneroMasculino:TamanhoPequeno -13.2796   624.1944 -0.021  
0.9830  
## GêneroMasculino:IdeologiaDireita 0.2351    0.8286  0.284  
0.7766  
## TamanhoMédio:IdeologiaDireita  0.7239    1.1550  0.627  
0.5308  
## TamanhoNanico:IdeologiaDireita 15.8126   509.6525  0.031  
0.9752  
## TamanhoPequeno:IdeologiaDireita 0.8938    882.7437  0.001  
0.9992  
## GêneroMasculino:TamanhoMédio:IdeologiaDireita -0.7529   1.2143 -0.620  
0.5353
```

```

## GêneroMasculino:TamanhoNanico:IdeologiaDireita      NA      NA      NA
NA
## GêneroMasculino:TamanhoPequeno:IdeologiaDireita   -2.3961   882.7439  -0.003
0.9978
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1523.1 on 1291 degrees of freedom
## Residual deviance: 1448.7 on 1277 degrees of freedom
## AIC: 1478.7
##
## Number of Fisher Scoring iterations: 13

```

H2) Deputados pretos e pardos apresentam maiores chances de migração em comparação a brancos (Wylie, 2018)

- São retiradas da análise as auto declarações **Amarela** e **Indígena** pois representam apenas 5 e 6 casos, respectivamente;
- São desconsideradas as **eleições anteriores a 2014** pois não havia auto declaração racial nesse momento;
- A raça é reclassificada;

```

modelo_raça <- banco_arrumado |>
  filter(Migrantes != "Primeiro mandato" & Ideologia != "Esquerda" & Ano_Eleição >
2010 &
  Cor_Raça_TSE != "Amarela" & Cor_Raça_TSE != "Indígena")
modelo_raça$Migrantes <- as.factor(modelo_raça$Migrantes)
modelo_raça <- modelo_raça |>
  mutate(raça_reclassificada = case_when(
    Cor_Raça_TSE == "Preta" ~ "Não branca",
    Cor_Raça_TSE == "Parda" ~ "Não branca",
    TRUE ~ "Branca"
  ))

```

É testado se há **diferença estatisticamente significativa** de pessoas pretas ou pardas (Não brancas) migrarem considerando **raça** e **tamanho** apenas com efeitos somados:

```

regressão_raça <- glm(Migrantes ~ raça_reclassificada + Tamanho,
                      data = modelo_raça, family = "binomial")
summary(regressão_raça)

```

```

##
## Call:
## glm(formula = Migrantes ~ raça_reclassificada + Tamanho, family = "binomial",
##       data = modelo_raça)
##
## Deviance Residuals:

```

```

##      Min      1Q   Median      3Q      Max
## -1.6498 -1.3365  0.7698  0.8730  1.2739
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 0.7683    0.1482  5.185 2.17e-07 ***
## raça_reclassificadaNão branca -0.4017    0.2084 -1.928  0.0539 .
## TamanhoMédio                  0.2964    0.1868  1.587  0.1126
## TamanhoNanico                -0.3293    0.6043 -0.545  0.5858
## TamanhoPequeno                -0.5906    0.2785 -2.121  0.0339 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 808.06 on 644 degrees of freedom
## Residual deviance: 793.14 on 640 degrees of freedom
## AIC: 803.14
##
## Number of Fisher Scoring iterations: 4

```

É testado como **raça** interage com **tamanho** na avaliação da significância estatística:

```

regressão_raça_e_tam <- glm(Migrantes ~ raça_reclassificada * Tamanho,
                             data = modelo_raça, family = "binomial")
summary(regressão_raça_e_tam)

##
## Call:
## glm(formula = Migrantes ~ raça_reclassificada * Tamanho, family = "binomial",
##      data = modelo_raça)
##
## Deviance Residuals:
##      Min      1Q   Median      3Q      Max
## -1.6345 -1.2858  0.7812  0.8605  1.4823
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 0.8030    0.1608  4.993 5.93e-07 ***
## raça_reclassificadaNão branca       -0.5516    0.3324 -1.659  0.0970
## TamanhoMédio                   0.2277    0.2121  1.074  0.2830
## TamanhoNanico                  -0.1098    0.7252 -0.151  0.8796
## TamanhoPequeno                  -0.6331    0.3068 -2.063  0.0391
## raça_reclassificadaNão branca:TamanhoMédio      0.3095    0.4498  0.688  0.4915
## raça_reclassificadaNão branca:TamanhoNanico     -0.8346    1.4528 -0.575  0.5656
## raça_reclassificadaNão branca:TamanhoPequeno     0.1994    0.7386  0.270  0.7871
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 808.06 on 644 degrees of freedom
## Residual deviance: 792.14 on 637 degrees of freedom
## AIC: 808.14
##
## Number of Fisher Scoring iterations: 4
```

H2.1) O efeito de raça varia conforme o tamanho e a ideologia do partido:

```
regressão_raça <- glm(Migrantes ~ raça_reclassificada * Tamanho * Ideologia,
                      data = modelo_raça, family = "binomial")
summary(regressão_raça)
```

```
##
## Call:
## glm(formula = Migrantes ~ raça_reclassificada * Tamanho * Ideologia,
##      family = "binomial", data = modelo_raça)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -2.3459 -1.2202  0.6681  0.9005  1.6651
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.6856
## 0.5967  4.501 6.77e-06 ***
## raça_reclassificadaNão branca -1.5870
## 1.0113 -1.569 0.116594
## TamanhoMédio -0.8938
## 0.7084 -1.262 0.207070
## Tamanhonoico -3.7842
## 1.2998 -2.911 0.003598 **
## Tamanhopequeno -0.2007
## 1.1997 -0.167 0.867165
## IdeologiaDireita -2.2614
## 0.6223 -3.634 0.000279 ***
## raça_reclassificadaNão branca:TamanhoMédio 1.1815
## 1.3392  0.882 0.377659
## raça_reclassificadaNão branca:Tamanhonoico -15.9351
## 650.8748 -0.024 0.980468
## raça_reclassificadaNão branca:Tamanhopequeno 14.6681
## 1029.1225  0.014 0.988628
## raça_reclassificadaNão branca:IdeologiaDireita 1.2629
## 1.0743  1.176 0.239790
## TamanhoMédio:IdeologiaDireita 1.3438
## 0.7454  1.803 0.071415 .
```

```

## TamanhoNanico:IdeologiaDireita          18.9261
650.8749  0.029  0.976802
## TamanhoPequeno:IdeologiaDireita        -0.4859
1.2486  -0.389  0.697196
## raça_reclassificadaNão branca:TamanhoMédio:IdeologiaDireita      -1.0384
1.4251  -0.729  0.466199
## raça_reclassificadaNão branca:TamanhoNanico:IdeologiaDireita       NA
NA      NA      NA
## raça_reclassificadaNão branca:TamanhoPequeno:IdeologiaDireita     -14.7748
1029.1228 -0.014  0.988545
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 808.06 on 644 degrees of freedom
## Residual deviance: 740.41 on 630 degrees of freedom
## AIC: 770.41
##
## Number of Fisher Scoring iterations: 14

```

H3) Deputadas mulheres e negras tem maiores chances de migrar para partidos maiores

```

modelo_misto <- banco_arrumado |>
  filter(Migrantes != "Primeiro mandato" & Ideologia != "Esquerda")
modelo_misto$Migrantes <- as.factor(modelo_misto$Migrantes)
modelo_misto <- modelo_misto |>
  mutate(raça_reclassificada = case_when(
    Cor_Raça_TSE == "Preta" ~ "Não branca",
    Cor_Raça_TSE == "Parda" ~ "Não branca",
    TRUE ~ "Branca"
  ))

```

Primeiro é testado se há **diferença estatisticamente significativa** de deputados migrarem de acordo com o **gênero e raça**, somados ao **tamanho** do partido:

```

regressão_mista <- glm(Migrantes ~ raça_reclassificada * Gênero + Tamanho,
                        data = modelo_misto, family = "binomial")
summary(regressão_mista)

```

```

##
## Call:
## glm(formula = Migrantes ~ raça_reclassificada * Gênero + Tamanho,
##       family = "binomial", data = modelo_misto)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.8263  -1.2958   0.7234   0.7912   1.3122
##
## Coefficients:

```

```

##                                     Estimate Std. Error z value
Pr(>|z|)
## (Intercept)                   1.4587    0.3029   4.816 1.47e-
06 ***
## raça_reclassificadaNão branca -0.3076    0.7397  -0.416
0.6775
## GêneroMasculino                -0.2515    0.3017  -0.834
0.4045
## TamanhoMédio                  -0.2063    0.1350  -1.528
0.1266
## TamanhoNanico                 -0.8990    0.4877  -1.843
0.0653 .
## TamanhoPequeno                 -0.9330    0.2179  -4.282 1.85e-
05 ***
## raça_reclassificadaNão branca:GêneroMasculino -0.2780    0.7672  -0.362
0.7171
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1523.1  on 1291  degrees of freedom
## Residual deviance: 1493.3  on 1285  degrees of freedom
## AIC: 1507.3
##
## Number of Fisher Scoring iterations: 4

```

É testado como as variáveis todas interagem entre si

```

regressão_mista_geral <- glm(Migrantes ~ raça_reclassificada * Gênero + Tamanho,
                               data = modelo_misto, family = "binomial")
summary(regressão_mista_geral)

```

```

## 
## Call:
## glm(formula = Migrantes ~ raça_reclassificada * Gênero + Tamanho,
##      family = "binomial", data = modelo_misto)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q       Max
## -1.8263  -1.2958   0.7234   0.7912   1.3122
##
## Coefficients:
##                                     Estimate Std. Error z value
Pr(>|z|)
## (Intercept)                   1.4587    0.3029   4.816 1.47e-
06 ***
## raça_reclassificadaNão branca -0.3076    0.7397  -0.416
0.6775
## GêneroMasculino                -0.2515    0.3017  -0.834
0.4045

```

```

## TamanhoMédio           -0.2063    0.1350   -1.528
0.1266
## TamanhoNanico          -0.8990    0.4877   -1.843
0.0653 .
## TamanhoPequeno         -0.9330    0.2179   -4.282  1.85e-
05 ***
## raça_reclassificadaNão branca:GêneroMasculino -0.2780    0.7672   -0.362
0.7171
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1523.1  on 1291  degrees of freedom
## Residual deviance: 1493.3  on 1285  degrees of freedom
## AIC: 1507.3
##
## Number of Fisher Scoring iterations: 4

```

H4) Deputados tem maiores chances de migrar para partidos grandes e médios

```

modelo_tam <- banco_arrumado |>
  filter(Migrantes != "Primeiro mandato" & Ideologia != "Esquerda")
modelo_tam$Migrantes <- as.factor(modelo_tam$Migrantes)

```

É testado se há **diferença estatisticamente significativa** de deputados migrarem de acordo com o **tamanho** do partido:

```

regressão_tam <- glm(Migrantes ~ Tamanho,
                      data = modelo_tam, family = "binomial")
summary(regressão_tam)

```

```

##
## Call:
## glm(formula = Migrantes ~ Tamanho, family = "binomial", data = modelo_tam)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q       Max
## -1.6958   -1.2735    0.7363    0.8079    1.0842
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.16679   0.09784 11.926 < 2e-16 ***
## TamanhoMédio -0.21457   0.13425 -1.598   0.1100
## TamanhoNanico -0.94365   0.48433 -1.948   0.0514 .
## TamanhoPequeno -0.94365   0.21696 -4.349  1.37e-05 ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)

```

```
##  
## Null deviance: 1523.1 on 1291 degrees of freedom  
## Residual deviance: 1502.4 on 1288 degrees of freedom  
## AIC: 1510.4  
##  
## Number of Fisher Scoring iterations: 4
```