

# Análise exploratória dos anos de estudo no Brasil com pessoas de 15 anos ou mais (2023)

João Pedro

16 de Julho de 2025

## 1 Introdução

Esse relatório tem como objetivo demonstrar a média dos anos de estudo no Brasil entre pessoas de 15 anos de idade ou mais. A análise é feita com base nos dados da PNAD Contínua Anual de 2023, disponibilizada pelo IBGE.

Os dados foram obtidos a partir da plataforma SIDRA do IBGE: acesse o link, e foi importado utilizando a API do SIDRA com o auxílio do pacote (sidrar). Isso permitiu acessar diretamente as informações necessárias e automatizar o processo de coleta de dados para análise e visualização.

A seguir, mapas e gráficos são apresentados para ilustrar as diferenças regionais e raciais na média de anos de estudo.

## 2 Importando a base de dados e tratando os dados

```
dados <- get_sidra(api = "/t/7127/n3/all/v/3593/p/2023/c86/allxt/c58/2795/d/v3593%201")
```

```
dados_selecionados <- dados %>%  
  select(  
    "anos_estudo" = Valor,  
    "uf" = 'Unidade da Federação (Código)',  
    "raca_cor" = 'Cor ou raça')
```

## 3 Criando o mapa

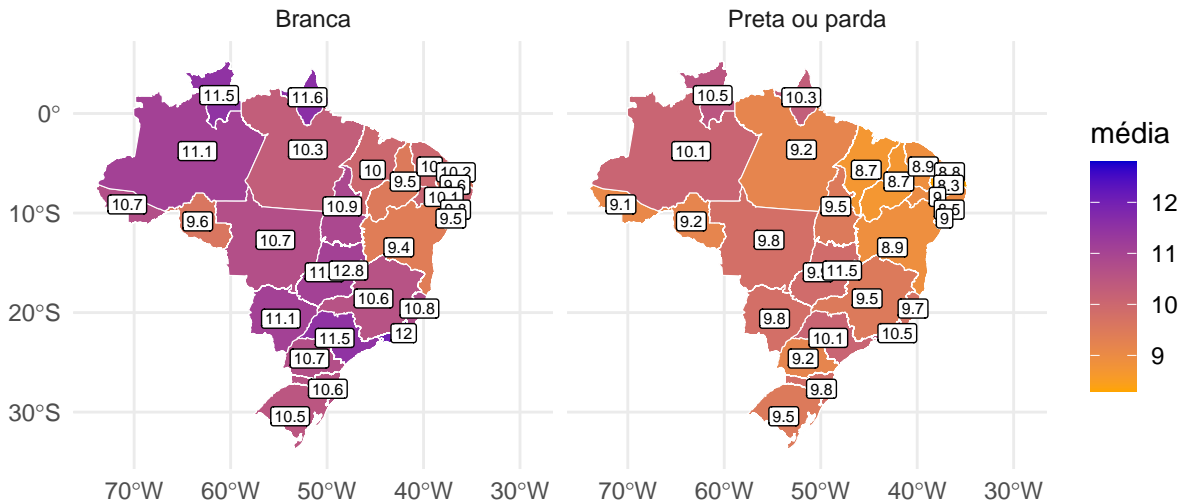
```
br_map_estado <- read_state(code_state = "all", year = 2020)  
br_map_estado$code_state = as.character(br_map_estado$code_state)
```

```
dados_juntos <- left_join(br_map_estado, dados_selecionados, by = c("code_state" = "uf"))
```

```
ggplot() +  
  geom_sf(data = dados_juntos, aes(fill = anos_estudo), color = "white") +  
  scale_fill_gradient(low = "orange", high = "blue3", name = "média") +  
  geom_sf_label(data = dados_juntos, aes(label = anos_estudo),  
    label.padding = unit(0.5, "mm"), size = 2) +  
  facet_wrap(~raca_cor) +  
  theme_minimal() +  
  labs (title = "Média de anos de estudo entre pessoas de 15 anos ou mais  
por estado e raça/cor (2023)",  
    caption = "Dados: PNAD-Continua Anual - 2023 | Elaboração: João Pedro",
```

```
x = "",
y=""')+
theme(plot.title = element_text(face = "bold", size = 14),
plot.subtitle = element_text(face = "italic"))
```

## Média de anos de estudo entre pessoas de 15 anos ou mais por estado e raça/cor (2023)



Dados: PNAD–Continua Anual – 2023 | Elaboração: João Pedro

### 4 Criando o gráfico - Os 3 Estados com maiores e menores médias de anos de estudo por cor/raça

```
library(DT)
library(dplyr)
library(ggplot2)

# Preparando os dados para o gráfico
maiores_branca <- dados_juntos %>%
  filter(raca_cor == "Branca") %>%
  arrange(desc(anos_estudo)) %>%
  slice_head(n = 3) %>%
  select(name_state, anos_estudo) %>%
  mutate(grupo = "Maiores - Branca")

maiores_preta_parda <- dados_juntos %>%
  filter(raca_cor == "Preta ou parda") %>%
  arrange(desc(anos_estudo)) %>%
  slice_head(n = 3) %>%
  select(name_state, anos_estudo) %>%
  mutate(grupo = "Maiores - Preta/Parda")
```

```

menores_branca <- dados_juntos %>%
  filter(raca_cor == "Branca") %>%
  arrange(anos_estudo) %>%
  slice_head(n = 3) %>%
  select(name_state, anos_estudo) %>%
  mutate(grupo = "Menores - Branca")

menores_preta_parda <- dados_juntos %>%
  filter(raca_cor == "Preta ou parda") %>%
  arrange(anos_estudo) %>%
  slice_head(n = 3) %>%
  select(name_state, anos_estudo) %>%
  mutate(grupo = "Menores - Preta/Parda")

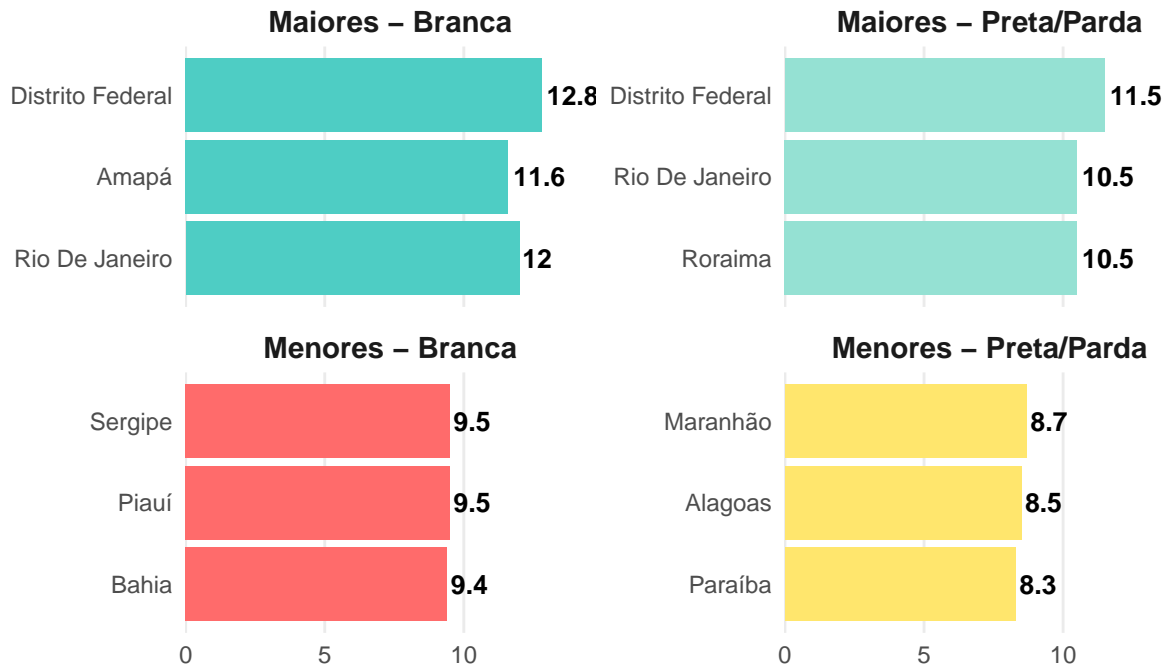
# Criando dataset combinado
dados_grafico <- bind_rows(
  maiores_branca,
  maiores_preta_parda,
  menores_branca,
  menores_preta_parda
)

# Criando gráfico horizontal
ggplot(dados_grafico, aes(x = reorder(name_state, anos_estudo), y = anos_estudo, fill = grupo)) +
  geom_col() +
  geom_text(aes(label = round(anos_estudo, 1)),
    hjust = -0.1,
    size = 3.5,
    fontface = "bold") +
  coord_flip() +
  facet_wrap(~grupo, scales = "free_y") +
  labs(
    title = "Os 3 Estados com Menores e Maiores Médias
de Anos de Estudo por Raça/Cor",
    caption = "Dados: PNAD-Continua Anual - 2023 | Elaboração: João Pedro",
    x = "",
    y = "",
    fill = "Categoria"
  ) +
  theme_minimal() +
  theme(
    legend.position = "none",
    plot.title = element_text(hjust = 0.5, size = 14, face = "bold"),
    plot.subtitle = element_text(hjust = 0.5, size = 12),
    strip.text = element_text(size = 11, face = "bold"),
    panel.grid.major.y = element_blank(),
    panel.grid.minor = element_blank()
  ) +
  scale_fill_manual(values = c(
    "Menores - Branca" = "#FF6B6B",
    "Maiores - Branca" = "#4ECDC4",
    "Menores - Preta/Parda" = "#FFE66D",
    "Maiores - Preta/Parda" = "#95E1D3"
  ))

```

```
)) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.15)))
```

## Os 3 Estados com Menores e Maiores Médias de Anos de Estudo por Raça/Cor



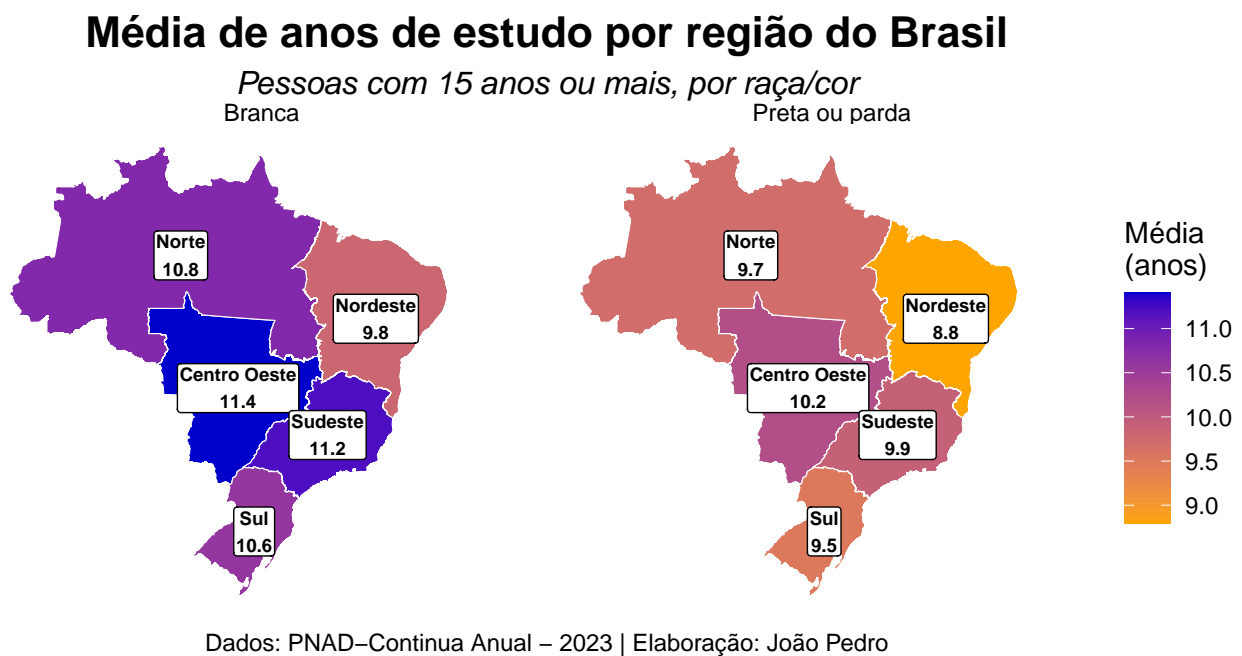
Dados: PNAD–Continua Anual – 2023 | Elaboração: João Pedro

### 5 Mapeando as regiões

```
dados_regioes <- dados_juntos %>%
  group_by(name_region, raca_cor) %>%
  summarise(
    anos_estudo_media = round(mean(anos_estudo, na.rm = TRUE), 1),
    .groups = 'drop'
  )

ggplot() +
  geom_sf(data = dados_regioes,
    aes(fill = anos_estudo_media,
      color = "white",
      size = 0.8) +
    scale_fill_gradient(low = "orange",
      high = "blue3",
      name = "Média\n(anos)") +
    geom_sf_label(data = dados_regioes,
      aes(label = paste0(name_region, "\n", anos_estudo_media)),
      label.padding = unit(0.3, "mm"),
      size = 2.5,
      fontface = "bold") +
    facet_wrap(~raca_cor, ncol = 2) +
```

```
theme_void() +
labs(title = "Média de anos de estudo por região do Brasil",
      subtitle = "Pessoas com 15 anos ou mais, por raça/cor",
      caption = "Dados: PNAD-Contínua Anual - 2023 | Elaboração: João Pedro") +
theme(
  plot.title = element_text(face = "bold", size = 16, hjust = 0.5),
  plot.subtitle = element_text(face = "italic", size = 12, hjust = 0.5),
  plot.caption = element_text(size = 9, hjust = 0.5),
  legend.position = "right"
)
```



O mapa revela importantes desigualdades regionais nos anos médios de estudo no Brasil. A região Centro-Oeste apresenta as maiores médias tanto para pessoas brancas (11,4 anos) quanto para pretas ou pardas (10,2 anos), seguida da região Sudeste. Já o Nordeste registra as menores médias em ambos os grupos raciais: 9,8 anos para brancos e apenas 8,8 anos para pretos ou pardos.

## 6 Conclusão

A análise dos dados da PNAD Contínua 2023 evidenciou disparidades significativas nos anos médios de estudo entre diferentes estados, regiões e grupos raciais do Brasil. Estados como o Distrito Federal e Rio de Janeiro apresentaram os maiores índices médios, enquanto estados do Nordeste como Maranhão e Alagoas ficaram entre os menores.

Além disso, observa-se uma diferença clara entre os grupos de cor/raça: pessoas brancas possuem, em média, mais anos de estudo do que pessoas pretas ou pardas em praticamente todas as regiões do país. Esses achados demonstram como a desigualdade social e racial ainda se reflete fortemente no acesso à educação no Brasil. A análise reforça a importância de investimentos em políticas educacionais inclusivas e regionalmente equilibradas para promover maior equidade no país.