

✓ Desafio I Zetta Lab 2025

Para esse desafio, foram baixados dados de PRODES [INPE/TerraBrasilis](#). As variáveis socioeconômicas foram extraídas da página do [IBGE](#). Há também o link do [Github](#), em que se encontra os códigos abaixo.

PRODES

O PRODES (Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite), conduzido pelo Instituto Nacional de Pesquisas Espaciais (INPE) desde 1988, é o principal sistema brasileiro de monitoramento do desmatamento por corte raso na Amazônia Legal. Utilizando imagens de satélite, o PRODES realiza anualmente o mapeamento das áreas desmatadas, fornecendo dados oficiais que são fundamentais para a formulação e avaliação de políticas públicas ambientais no Brasil. Esses dados são amplamente utilizados por órgãos governamentais, pesquisadores e organizações não governamentais para compreender a dinâmica do desmatamento e implementar ações de conservação e desenvolvimento sustentável na região.

✓ Primeiro vamos visualizar o desmatamento no estado do Pará desde 2020.

Para esse primeiro momento, vamos utilizar esses pacotes:

```
1 Sys.setenv(MAKEFLAGS= paste0("-j",parallel::detectCores()))
2
3 print(parallel::detectCores())
4
5 if (!requireNamespace("geobr", quietly = TRUE)) {
6   install.packages("geobr", verbose = TRUE, quiet = FALSE, dependencies = TRUE, INSTALL_opts = "--no-multiarch")
7 }
8
9 if (!requireNamespace("tidyverse", quietly = TRUE)) {
10  install.packages("tidyverse", verbose = TRUE, quiet = FALSE, dependencies = TRUE, INSTALL_opts = "--no-multiarch")
11 }
12
13 if (!requireNamespace("colorspace", quietly = TRUE)) {
14  install.packages("colorspace", verbose = TRUE, quiet = FALSE, dependencies = TRUE, INSTALL_opts = "--no-multiarch")
15 }
16
17 if (!requireNamespace("tidyr", quietly = TRUE)) {
18  install.packages("tidyr", verbose = TRUE, quiet = FALSE, dependencies = TRUE, INSTALL_opts = "--no-multiarch")
19 }
20
21 if (!requireNamespace("ggplot2", quietly = TRUE)) {
22  install.packages("ggplot2", verbose = TRUE, quiet = FALSE, dependencies = TRUE, INSTALL_opts = "--no-multiarch")
23 }
24
25 library(geobr)
26 library(tidyverse)
27 library(colorspace)
28 library(tidyr)
29 library(ggplot2)
```

```
🔄 Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

system (cmd0): /usr/lib/R/bin/R CMD INSTALL --no-multiarch

also installing the dependencies ‘proxy’, ‘e1071’, ‘wk’, ‘lazyeval’, ‘classInt’, ‘s2’, ‘units’, ‘assertthat’, ‘checkmate’, ‘duckdb’, ‘

foundpkgs: proxy, e1071, wk, lazyeval, classInt, s2, units, assertthat, checkmate, duckdb, rex, sf, arrow, censobr, covr, geobr, /tmp/

files: /tmp/Rtmp01Lj9W/downloaded_packages/proxy_0.4-27.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/e1071_1.7-16.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/wk_0.9.4.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/lazyeval_0.2.2.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/classInt_0.4-11.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/s2_1.1.7.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/units_0.8-7.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/assertthat_0.2.1.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/checkmate_2.3.2.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/duckdb_1.2.2.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/rex_1.2.1.tar.gz,
       /tmp/Rtmp01Lj9W/downloaded_packages/sf_1.0-20.tar.gz,
```

```

/tmp/Rtmp01Lj9W/downloaded_packages/arrow_20.0.0.tar.gz,
/tmp/Rtmp01Lj9W/downloaded_packages/censobr_0.4.1.tar.gz,
/tmp/Rtmp01Lj9W/downloaded_packages/covr_3.6.4.tar.gz,
/tmp/Rtmp01Lj9W/downloaded_packages/geobr_1.9.1.tar.gz

```

```

1): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/pro
2): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/wk_
3): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/laz
4): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/uni
5): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/ass
6): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/che
7): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/duc
8): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/e10
9): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/s2_
10): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/re
11): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/ar
12): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/cl
13): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/ce
14): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/co
15): succeeded '/usr/lib/R/bin/R CMD INSTALL --no-multiarch -l '/usr/local/lib/R/site-library' '/tmp/Rtmp01Lj9W/downloaded_packages/sf

```

Utilizando o pacote geobr, vamos extrair os polígonos dos municípios para depois utilizar na plotagem

```
1 mun_pa <- read_municipality(year = 2020, code_muni = "PA")
```

→ Using year/date 2020

Adicionando nossa base de dados

```
1 dados <- read_delim("data.csv", delim = ";", quote = "\"", escape_double = TRUE,
2 locale = locale(encoding = "UTF-8"), na = c("", "NA"))
```

→ Rows: 144 Columns: 11
 — Column specification —
 Delimiter: ";"
 chr (3): Município, Escolarizacao, IDHM
 dbl (8): pop_censo, Mortalidade_infantil, PIB, 2020, 2021, 2022, 2023, 2024
 Use `spec()` to retrieve the full column specification for this data.
 Specify the column types or set `show_col_types = FALSE` to quiet this message.

Adicionando a coluna geom nos nossos dados

```
1 dados$Geom <- mun_pa$geom
```

✓ Agora vamos plotar o gráfico dos valores de desmatamento por ano

Primeiro vamos organizar a nossa tabela

```

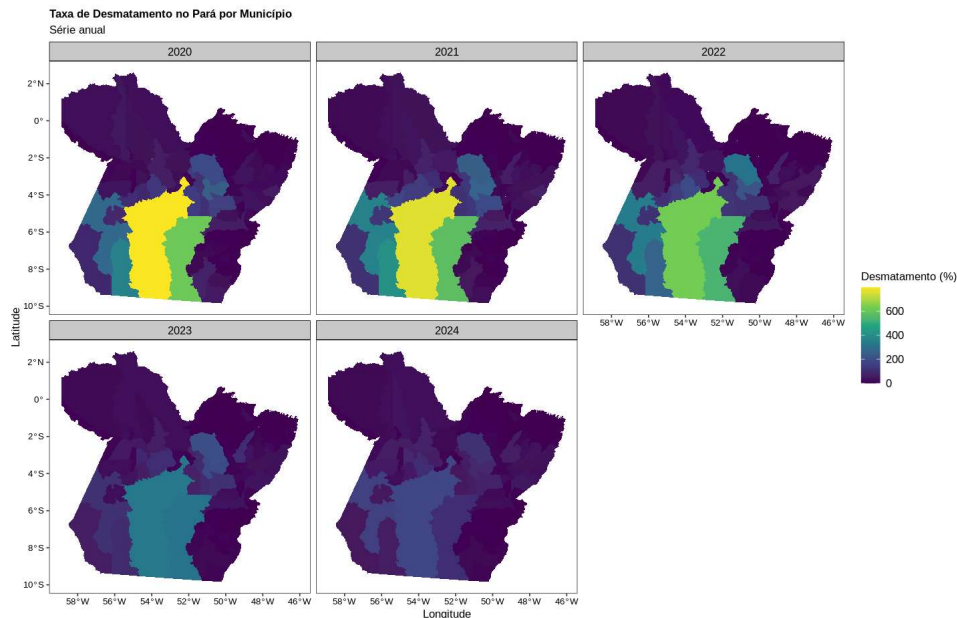
1 comp_thin <- dados %>%
2   gather(., key=year, value=Desmatamento, 7:11)
3
4 comp_thin$Desmatamento <- as.numeric(comp_thin$Desmatamento)
5 comp_thin$year <- as.factor(comp_thin$year)
6 levels(comp_thin$year) <- c("2020", "2021", "2022", "2023", "2024")

```

✓ Agora vamos usar o pacote ggplot2 para gerar gráficos.

Nota-se que estamos utilizando cores viridis para que pessoas com daltonismo não tenham dificuldade na visualização.

```
1 grafico_1 <- ggplot(comp_thin) +
2   geom_sf(
3     data = comp_thin$Geom,
4     aes(fill = comp_thin$Desmatamento),
5     color = NA)+ facet_wrap(~ comp_thin$year) +
6   scale_fill_viridis_c(
7     option = "viridis",
8     name = "Área Desmatada",
9     na.value = "transparent"
10  ) +
11  labs(
12    x = "Longitude",
13    y = "Latitude",
14    title = "Taxa de Desmatamento no Pará por Município",
15    subtitle = "Série anual"
16  ) +
17  theme_bw(base_size = 10) +
18  theme(
19    axis.text = element_text(color = "black"),
20    axis.title = element_text(color = "black"),
21    plot.title = element_text(color = "black", size = 10, face = "bold"),
22    plot.subtitle = element_text(color = "black", size = 10),
23    legend.text = element_text(color = "black", size = 10),
24    legend.title = element_text(color = "black", size = 10),
25    strip.text = element_text(color = "black", size = 10),
26    strip.background = element_rect(fill = "#CCCCCC", colour = "black"),
27    panel.grid.major = element_blank(),
28    panel.grid.minor = element_blank()
29  )
30 options(repr.plot.width = 12, repr.plot.height = 10)
31 grafico_1
32
```



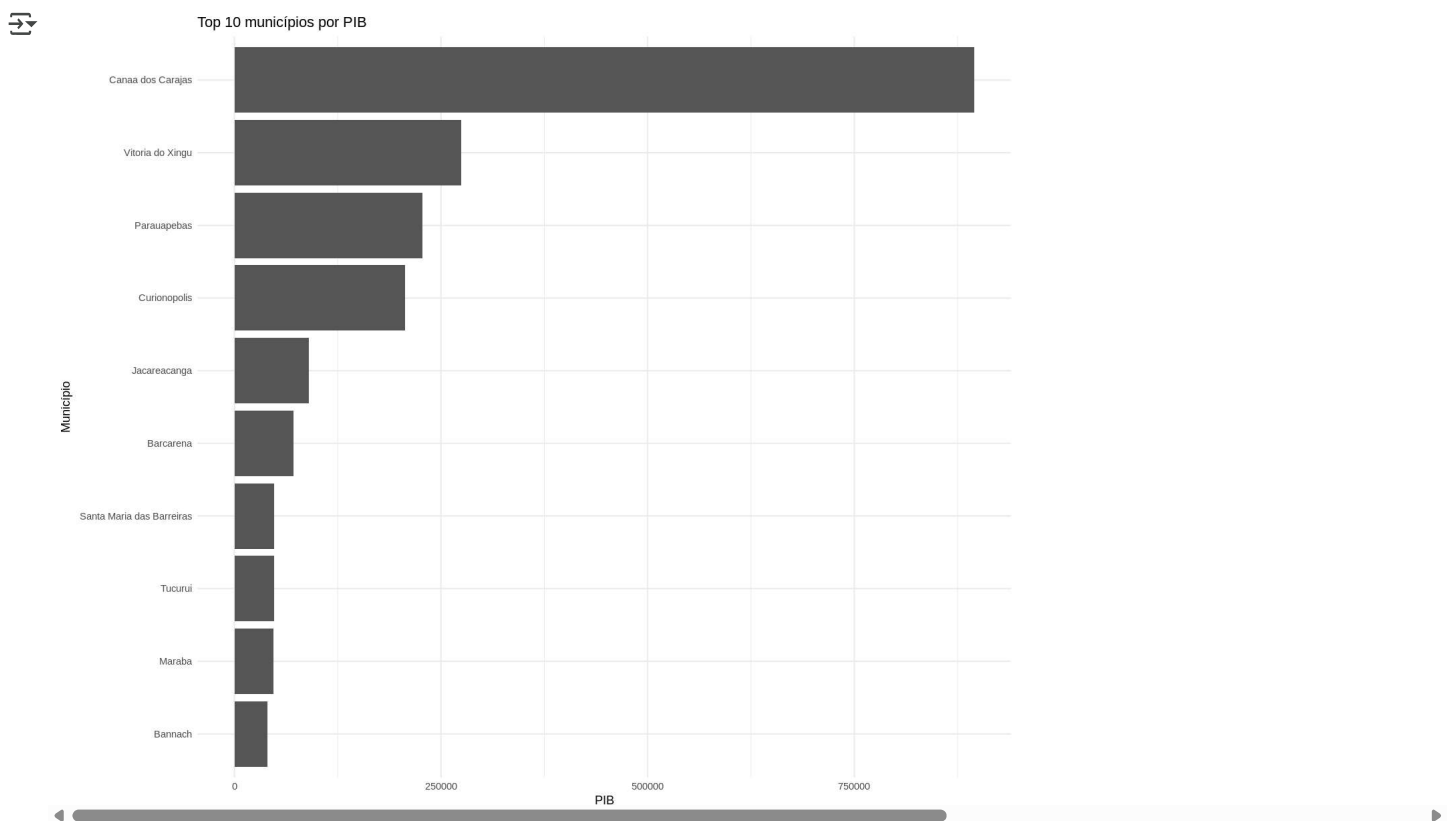
A partir dos gráficos gerados, percebe-se que a taxa de desmatamento foi mais intensa no período de 2020 a 2022, concentrados no sul do estado, localizado no arco do desmatamento. No último ano, os municípios com a maior taxa de desmatamento foram Altamira, Itaituba, Portel, São Felix do Xingu e Uruará. Entre 2023 e 2024, o Pará teve uma redução de 28,4% no desmatamento, resultado de uma maior fiscalização, maior consciencia ambiental da população e execução da Política Estadual sobre Mudanças Climáticas do Pará.

Para analisar os impactos socioeconômicos do desmatamento, iremos utilizar dados da cobertura florestal com as variáveis:

- Índice de Desenvolvimento Humano Municipal(IDHM);
- Produto Interno Bruto por pessoa (PIB per capita);
- Mortalidade infantil;
- Escolarização de crianças entre 6 a 14 anos do estado.

Agora vamos analisar os municípios com os maiores valores de IDHM, PIB per capita, Mortalidade infantil e escolarização.

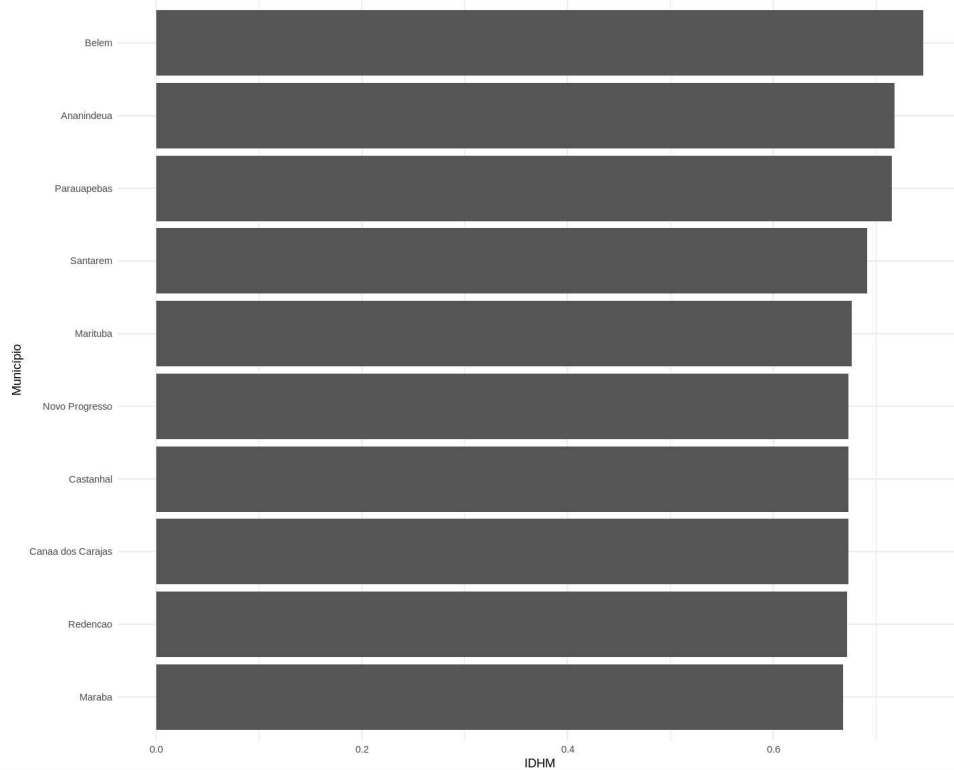
```
1 pib_10 <- dados %>%
2   arrange(desc(PIB)) %>%
3   slice(1:10)
4
5 top10_pib<-ggplot(pib_10, aes(x = reorder(Municipio, PIB), y = PIB)) +
6   geom_col() +
7   coord_flip() +
8   labs(x = "Município", y = "PIB", title = "Top 10 municípios por PIB") +
9   theme_minimal()
10 top10_pib
```



```
1 dados$IDHM<-as.numeric(dados$IDHM)
2 idhm_10 <- dados %>%
3   arrange(desc(IDHM)) %>%
4   slice(1:10)
5
6 top10_IDHM<-ggplot(idhm_10, aes(x = reorder(Município, IDHM), y = IDHM)) +
7   geom_col() +
8   coord_flip() +
9   labs(x = "Município", y = "IDHM", title = "Top 10 municípios por IDHM") +
10  theme_minimal()
11 top10_IDHM
```



Top 10 municípios por IDHM



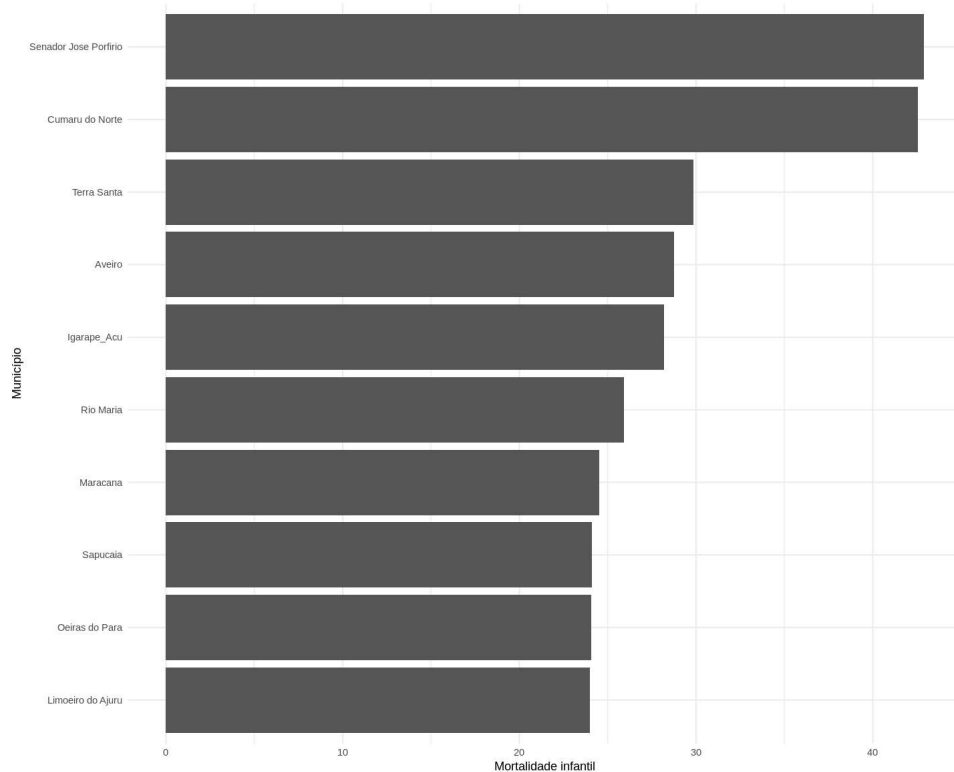
```

1 mortalidade_10 <- dados %>%
2   arrange(desc(Mortalidade_infantil)) %>%
3   slice(1:10)
4
5 top10_mortalidade<-ggplot(mortalidade_10, aes(x = reorder(Município, Mortalidade_infantil), y = Mortalidade_infantil)) +
6   geom_col() +
7   coord_flip() +
8   labs(x = "Município", y = "Mortalidade infantil", title = "10 municípios com maiores taxas de mortalidade infantil") +
9   theme_minimal()
10 top10_mortalidade

```



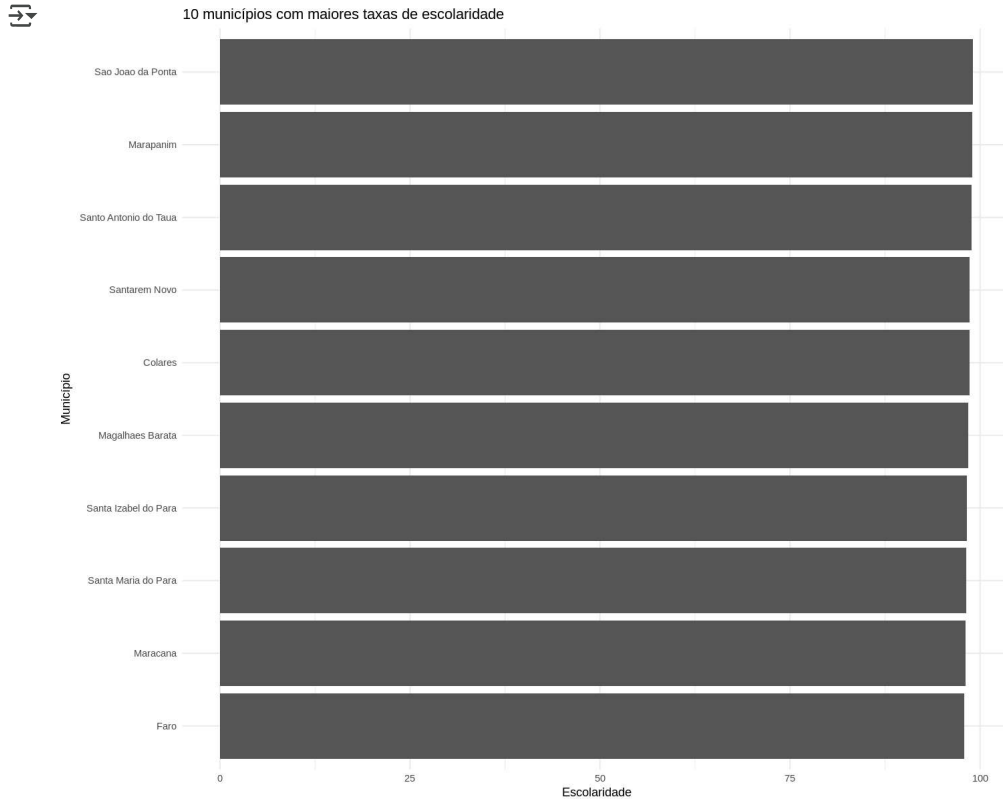
10 municípios com maiores taxas de mortalidade infantil



```

1 dados$Escolarizacao<-as.numeric(dados$Escolarizacao)
2 escolaridade_10 <- dados %>%
3   arrange(desc(Escolarizacao)) %>%
4   slice(1:10)
5
6 top10_escolaridade<-ggplot(escolaridade_10, aes(x = reorder(Municipio, Escolarizacao), y = Escolarizacao)) +
7   geom_col() +
8   coord_flip() +
9   labs(x = "Município", y = "Escolaridade", title = "10 municípios com maiores taxas de escolaridade") +
10  theme_minimal()
11 top10_escolaridade

```



Para analisar a influencia das variaveis seleccionadas no desmatamento, iremos fazer uma regressão linear simples.

```

1 library(readxl)
2 library(dplyr)
3 library(readr)
4
5 df <- read_excel("dados2.xlsx") %>%
6   rename(
7     idhm          = IDHM,
8     mort_inf      = Mortalidade_infantil,
9     escolarizacao = Escolarizacao
10  ) %>%
11  mutate(
12    idhm          = parse_number(as.character(idhm), na = c("-", "")),
13    mort_inf      = parse_number(as.character(mort_inf), na = c("-", "")),
14    escolarizacao = parse_number(as.character(escolarizacao), na = c("-", "")),
15    pib           = as.numeric(PIB),
16    desmat_2024   = as.numeric(`2024`)
17  ) %>%
18  filter(
19    !is.na(desmat_2024),
20    !is.na(idhm),
21    !is.na(pib),
22    !is.na(mort_inf),
23    !is.na(escolarizacao)
24  )
25
26 model <- lm(
27   desmat_2024 ~ idhm + pib + mort_inf + escolarizacao,
28   data = df
29 )

```

```
30
31 summary(model)
```



```
Call:
lm(formula = desmat_2024 ~ idhm + pib + mort_inf + escolarizacao,
    data = df)
```

Residuals:

Min	1Q	Median	3Q	Max
-42.364	-12.987	-7.647	1.464	140.228

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.032e+02	5.423e+01	3.746	0.000263 ***
idhm	1.041e+02	4.871e+01	2.137	0.034384 *
pib	-1.678e-05	2.911e-05	-0.576	0.565316
mort_inf	4.835e-01	3.141e-01	1.539	0.125974
escolarizacao	-2.708e+00	6.605e-01	-4.099	7.04e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 27.07 on 138 degrees of freedom
Multiple R-squared: 0.1308, Adjusted R-squared: 0.1056
F-statistic: 5.191 on 4 and 138 DF, p-value: 0.0006319

A variável IDHM e a Escolarização foram significativas no modelo, em que a escolarização foi a mais significativa. O IDHM significa que um aumento de 0,01 ponto no IDHM está associado a +1,04 na taxa de desmatamento em 2024 ($p=0.034$). Para a escolarização, cada ponto percentual a mais na taxa de escolarização reduz, em média, 2,71 de desmatamento ($p<0.001$). PIB e mortalidade infantil não foram estatisticamente significativos ($p>0.1$), indicando pouca relação linear direta com a taxa de desmate em 2024. O modelo resultou em um R ajustado baixo, em que poderia ser incluída mais variáveis explicativas no modelo.