

Nota técnica

Método de cálculo utilizado para estimar o número de crianças e adolescentes em situação de não frequência escolar utilizando dados do módulo de educação da Pesquisa Nacional de Amostra por Domicílios Contínua Anual

A Pesquisa Nacional de Amostra por Domicílios Contínua (PNADc) é uma pesquisa realizada anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE), cujo objetivo principal é acompanhar características socioeconômicas da população brasileira ao longo do tempo.

A *frequência escolar*, ou seja, a presença regular dos estudante nas atividades escolares, é um dos temas tratados no módulo de educação da PNADc. Neste sentido, a pergunta “Frequenta a escola?” (variável V3002), aplicada para todos os moradores dos domicílios da amostra¹, busca compreender se os indivíduos estavam ou não participando regularmente de atividades escolares no momento da entrevista.

Isto posto, as seguintes etapas foram realizadas para estimar o número de crianças e adolescentes que não frequentam a escola, a partir dos dados da PNADc:

1. Seleção dos indivíduos em idade escolar (0 a 17 anos)

Utilizou-se a variável contínua **V2009** (idade do morador na data de referência) com o filtro para indivíduos com menos de 17 anos de idade, selecionando apenas aqueles em idade escolar.

2. Seleção dos indivíduos com a educação básica incompleta

Utilizou-se a variável categórica **VD3004** (Nível de instrução mais elevado alcançado) com filtro para excluir as categorias “Médio completo ou equivalente”, “Superior incompleto ou equivalente”, “Superior completo”, mantendo apenas aqueles que não concluíram a educação básica.

3. Seleção dos indivíduos que não frequentam a escola

Utilizou-se a variável categórica **V3002** com o filtro para os indivíduos que responderam “não” para essa pergunta, selecionando apenas aqueles que não estavam frequentando a escola.

4. Desagregação dos dados por estratos da população

Foram selecionadas as seguintes variáveis para desagregar os dados de não frequência escolar para diversos estratos sociais de interesse:

- **V1022:** Localização
- **V2007:** Sexo
- **V2010:** Cor ou raça
Foram agrupadas as categorias 1 e 3 (Branca e amarela) e 2, 4 e 5 (Preta, parda e indígena)
- **VDI5011:** Rendimento domiciliar per capita
Foram definidos os percentis 20% e 80% para categorizar os indivíduos entre os grupos “20% mais ricos” e “20% mais pobres”.
 - Para o grupo “20% mais ricos”, foi considerado rendimento domiciliar per capita acima de R\$ 2.381,00.
 - Para o grupo “20% mais pobres”, foi considerado rendimento domiciliar per capita abaixo de R\$ 533,00.

¹ Nota técnica do IBGE sobre a PNADc, disponível em:
https://biblioteca.ibge.gov.br/visualizacao/livros/liv102172_notas_tecnicas.pdf

A extração dos dados foi feita com o **Software R**, utilizando os seguintes pacotes:

- **PNADcIBGE**
- **Survey**

A sistematização dos dados foi feita utilizando o **Software Excel**.

Script utilizado para a coleta dos dados:

```
install.packages("PNADcIBGE")
install.packages("survey")
library(PNADcIBGE)
library(survey)
variaveis_selecionadas <-
c("UF", "V1022", "V2001", "V2007", "V2009", "V2010", "V3002", "VD3004", "VDI5011", "VD4020", "V3009A")
dadosPNADc <- get_pnadc(year=ANO, topic=2, labels=TRUE, deflator=TRUE, design=TRUE,
vars=variaveis_selecionadas)
dadosPNADc$variables <- transform(dadosPNADc$variables, faixa=ifelse((V2009>=0 & V2009<4), "0 a
3", ifelse((V2009>=4 & V2009<6), "4 e 5", ifelse((V2009>=6 & V2009<11), "6 e 10", ifelse((V2009>=11 &
V2009<15), "11 e 14", ifelse((V2009>=15 & V2009<18), "15 e 17", "18 ou mais")))))
dadosPNADc$variables <- transform(dadosPNADc$variables, fora_escola=ifelse((V2009<18 &
V3002=="Não"), 1, 0))
dadosPNADc$variables <- transform(dadosPNADc$variables, concluintes=ifelse((VD3004=="Médio
completo ou equivalente" | VD3004=="Superior incompleto ou equivalente" | VD3004=="Superior
completo"), 1, 0))
dadosPNADc$variables <- transform(dadosPNADc$variables, fora_b=ifelse((fora_escola=="1" &
concluintes=="0"), 1, 0))
dadosPNADc$variables <- transform(dadosPNADc$variables, faixa_4=ifelse((V2009>=0 &
V2009<=4), 1, 0))
dadosPNADc$variables <- transform(dadosPNADc$variables, faixa_pre_f=ifelse((faixa_4==1 &
V3002=="Não"), 1, 0))
dadosPNADc$variables <- transform(dadosPNADc$variables, fora_d=ifelse((fora_b=="1" |
faixa_pre_f==1), 1, 0))
dadosPNADc$variables <- transform(dadosPNADc$variables, raca=ifelse((V2010=="Branca" |
V2010=="Amarela"), "1) Branca e amarela", ifelse((V2010=="Preta" | V2010=="Parda" |
V2010=="Indígena"), "2) Preta, parda e indígena", "3) Outro")))
svytable(~V2007, design=dadosPNADc)
quantisrenda <- svyquantile(x=~VDI5011, design=dadosPNADc, quantiles=c(0.2, 0.8), ci=FALSE,
na.rm=TRUE)
quantisrenda
dadosPNADc$variables <- transform(dadosPNADc$variables, quantil=ifelse((VDI5011<=533), "20%
mais pobres", ifelse((VDI5011>=2381), "20% mais ricos", "60% com renda intermediária")))
Planilha0 <- svytotal(x=~interaction(UF, V1022, faixa), design=subset(dadosPNADc, V2009<18,
weights=v1032), na.rm=TRUE)
write.csv2(Planilha0, "C:/Users/...")
Planilha1 <- svytotal(x=~interaction(UF, V1022, faixa, fora_d), design=subset(dadosPNADc, V2009<18,
weights=v1032), na.rm=TRUE)
write.csv2(Planilha1, "C:/Users/...")
Planilha2 <- svytotal(x=~interaction(UF, V1022, faixa, fora_d, V2007),
design=subset(dadosPNADc, V2009<18 & V3002=="Não", weights=v1032), na.rm=TRUE)
write.csv2(Planilha2, "C:/Users/...")
Planilha3 <- svytotal(x=~interaction(UF, V1022, faixa, fora_d, raca),
design=subset(dadosPNADc, V2009<18 & V3002=="Não", weights=v1032), na.rm=TRUE)
write.csv2(Planilha3, "C:/Users/...")
Planilha4 <- svytotal(x=~interaction(UF, V1022, faixa, quantil, fora_d),
design=subset(dadosPNADc, V2009<18 & V3002=="Não", weights=v1032), na.rm=TRUE)
write.csv2(Planilha4, "C:/Users/...")
Planilha5 <- svytotal(x=~interaction(UF, V1022, fora_d, V3009A),
design=subset(dadosPNADc, V2009<18 & V3002=="Não", weights=v1032), na.rm=TRUE)
write.csv2(Planilha5, "C:/Users/...")
```