

# Analizando a PNADC com o pacote **survey**

## Exercício

Guilherme Jacob

16/08/2021

O conjunto de dados que vamos utilizar neste exercício está guardado no arquivo **pnadc-recorte.Rds** na pasta **dados** e pode ser lido usando a função **readRDS**. Para ler estes dados e guardá-los no objeto **pnadc.df**, use o comando abaixo (verifique o endereço do arquivo no seu computador!):

```
pnadc.df <- readRDS( "dados/pnadc-recorte.Rds" )
```

Este arquivo contém um recorte dos microdados da 1ª visita da PNAD Contínua. Com ele, realize as tarefas a seguir:

1. Crie um objeto de plano amostral;
2. Estime as seguintes características:
  - a) Média do rendimento domiciliar per capita (**def.rdpc**), usando a função **svymean**;
  - b) Média do rendimento domiciliar per capita da sua região usando a função **subset**;
  - c) Média do tamanho dos domicílios (**vd2003**) por região, usando a função **svyby** e a variável **regiao**.
  - d) A soma dos rendimentos de todos os trabalhos (**def.vd4019**) por trimestre na sua região, usando **svyby** combinado com **svytotal** para o objeto de plano da sua região.
3. Escolha uma região e crie um objeto de plano amostral para ela usando a função **subset**. Com ele:
  - a) Estime a média do rendimento domiciliar per capita (**def.rdpc**) por situação do domicílio, usando a função **svyby** e a variável **v1022**. Calcule também o coeficiente de variação e o intervalo de confiança (95%) das estimativas usando a opção **vartype** e o efeito do plano amostral com a opção **deff = TRUE**;
  - b) Usando a função **update**, crie a variável **p.trab = rtrab / rdom**, indicando a proporção do trabalho no rendimento do domicílio;
    - Dica: Isso vai criar outro objeto. Guarde-o! Vamos utilizá-lo daqui pra frente.
  - c) Agora, usando as técnicas utilizadas da questão 3, calcule a média desta proporção entre os domicílios;
    - Dica: use a função **subset( objeto , vd2002 == 1 )**. Na definição do IBGE, domicílios só tem um responsável!
  - d) Usando a função **svyplot**, faça um diagrama de dispersão comparando **log(rtrab)** vs. **p.trab** entre os domicílios com renda do trabalho positiva.
    - Dica: filtre domicílios onde **rtrab > 0**!