



# Analisando dados de empregabilidade com o R

Dados da PNAD Contínua e PNAD covid-19

Mônica Rocabado e Renata Leal

Youth Voices Brasil - Open Data Day 2021

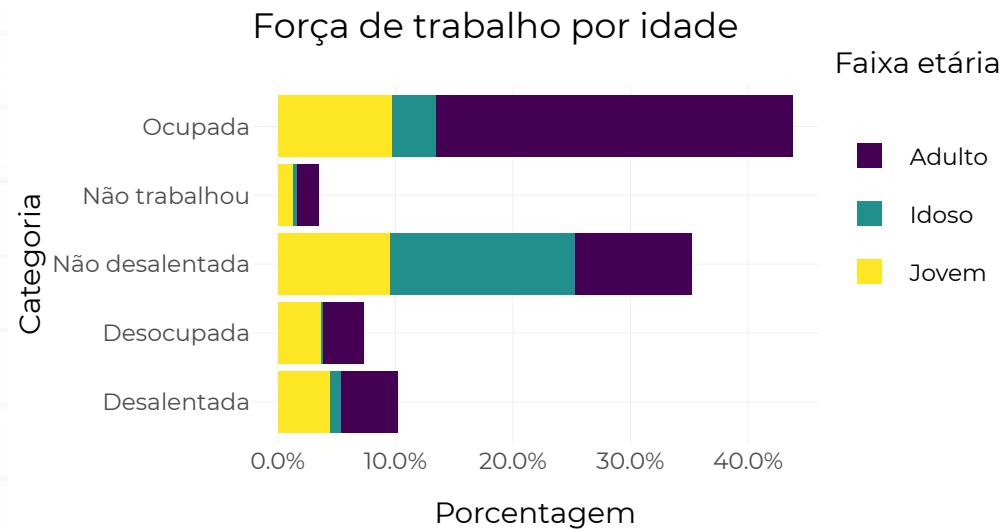
(updated: 2021-03-04)

## Agenda e expectativas

EVERYONE  
CAN COOP

## O que vamos ver hoje

- O que é a linguagem R?
- R 101 - conceitos rápidos
- Compreendendo os conceitos na prática com a PNAD Contínua
- Visualizando os dados
- Compreendo os dados e a importância do tema



## O que *não* vamos ver hoje

- Não é um curso aprofundado da linguagem R
- Não trataremos de conceitos estatísticos e econométricos
- Não é curso sobre Big Data



## Avisos

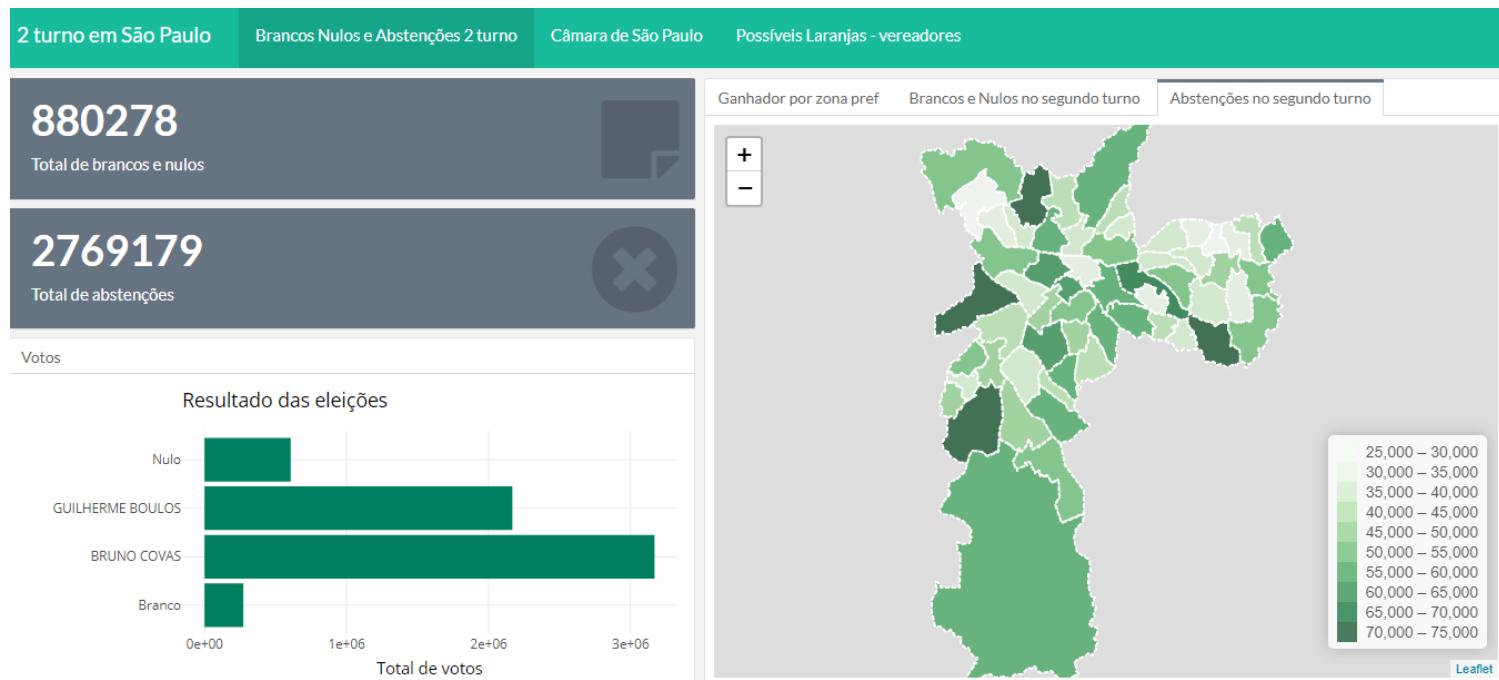
- Dúvidas pelo link
- Não precisa copiar nada
- Microfone no mudo
- Na parte mão na massa vou pedindo interação pelo chat

## A linguagem R

# O que é a linguagem R? Qual sua importância

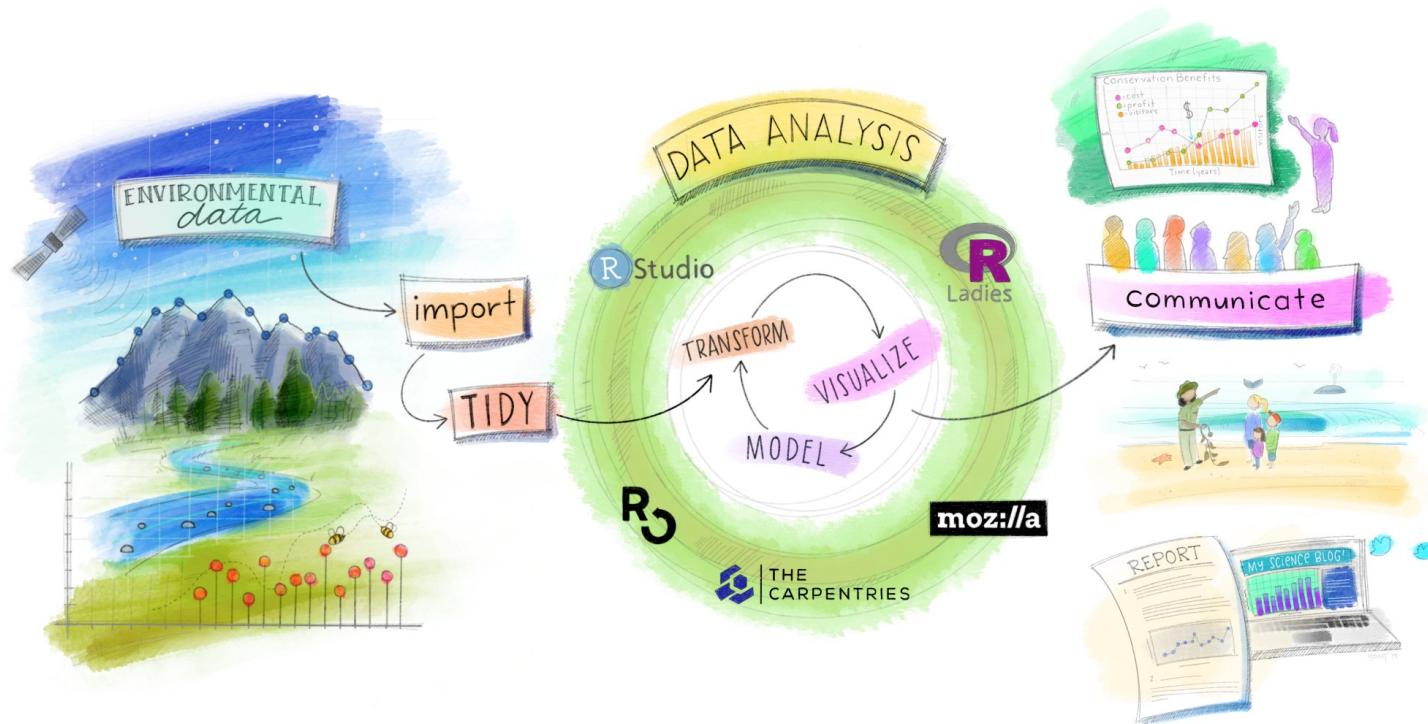
- Lógica de programação e compreensão do mundo
- Linguagem de programação gratuita e open-source (funções pré-prontas)
- Ambiente de desenvolvimento: análise de dados, Modelos, Dashboards

## Mais exemplos do possibilidades no R



Fonte: [https://monimazz.github.io/sp\\_segundoturno\\_2020/#brancos-nulos-e-abstenc%C3%A7%C3%B5es-2-turno](https://monimazz.github.io/sp_segundoturno_2020/#brancos-nulos-e-abstenc%C3%A7%C3%B5es-2-turno)

# Autonomia de pesquisa

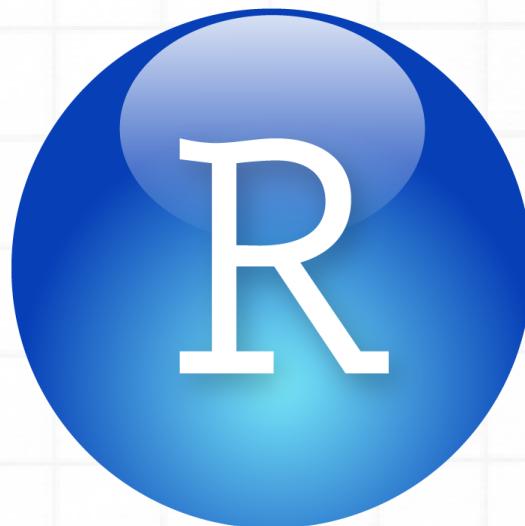


# A interface do R

R: Engine



RStudio: Dashboard



Studio®

## A interface do R: Rstudio (IDE)

R 101



## A lógica do R

- Computador precisa de contexto

## E se um computador fosse fazer um sanduíche?

- Declarar e informar os ingredientes
- Operacionar as instruções do preparo

```
#Declarar e atribuir valor  
oi <- "Olá mundo"  
  
# Instrução de imprimir valor  
print(oi)
```

## A lógica do R

- R base: instruções pré-prontas

```
numeros <- c(100 , 50 , 20)
print(numeros)
```

```
## [1] 100 50 20
```

## A lógica do R

- R base: instruções pré-prontas

```
mean(numeros)
```

```
## [1] 56.66667
```

Vamos para o projeto!

<https://rstudio.cloud/project/2247259>



## Exercício #1

- Atribua a um objeto a frase "olá mundo" no R Console
- Chame o seu objeto

## Formatos no R

- pdf
- excel
- csv
- rds
- etc

## Tipos de dados no R

- data
- integer: números inteiros
- numeric: números decimais
- complex: números complexos
- logical: FALSE, TRUE, NA
- factor: categórica ex: (“ótimo”, “bom”, “médio”, “ruim”)
- character: texto ex: (“eu amo bolo de chocolate”)

## Pacotes



## Pacotes ou nossa caixa de ferramentas

| Packages are the fundamental units of reproducible R code.

- Pacotes = Caixa de ferramentas
- Open source e a infinidade de pacotes
- Coleção de funções, dados, documentação, etc

```
# install.packages("kittyR")
library(kittyR)

kittyR::kittyR(meow = FALSE, only_kitten = TRUE)
```



# Explicando na prática! Dados da PNAD



# Dados da PNAD Contínua

- ◦ Foco em algumas variáveis
    - Identificação e controle
    - Características dos moradores
    - Características de trabalho das pessoas de 14 anos ou mais de idade
    - Pessoas ocupadas
- Microdados 3 trimestre de 2020
- UF
    - idade
    - cor
    - sexo
    - Na semana de ... a ...  
(semana de referência), ... trabalhou ou estagiou,  
durante pelo menos 1 hora,  
em alguma atividade  
remunerada em dinheiro?
    - Na semana de ... a ....  
(semana de referência), por  
que motivo ... estava  
afastado desse trabalho?
    - No período de ... a ...  
(período de referência de 30  
dias), ... Tomou alguma  
providência para conseguir

# Filtrando os dados

# O pacote dplyr

- Pacote que permite manipulação de dados
- filtragem de linhas com condicionais
- selecionar colunas
- mudar valores das colunas
- Resumir valores
- Agrupar valores

```
library(dplyr)
```

# dplyr::filter()

KEEP ROWS THAT  
satisfy  
*your CONDITIONS*

keep rows from... this data... ONLY IF... type is "otter" AND site is "bay"  
filter(df, type == "otter" & site == "bay")



A cartoon illustration featuring a yellow smiley face character pointing towards a map of a coastal area labeled 'BAY'. To the right is a table of data with four rows. Below the table is a small caption '@allison\_horst'.

type	food	site
otter	urchin	bay
Shark	seal	channel
otter	abalone	bay
otter	crab	wharf

@allison\_horst

# As condicionais no R

- Maior ou igual a: `>=`
- Igual a: `==`
- Menor ou igual a: `<=`
- Diferente de: `!=`

# imagem mutate

# imagem case\_when

dplyr::case\_when()

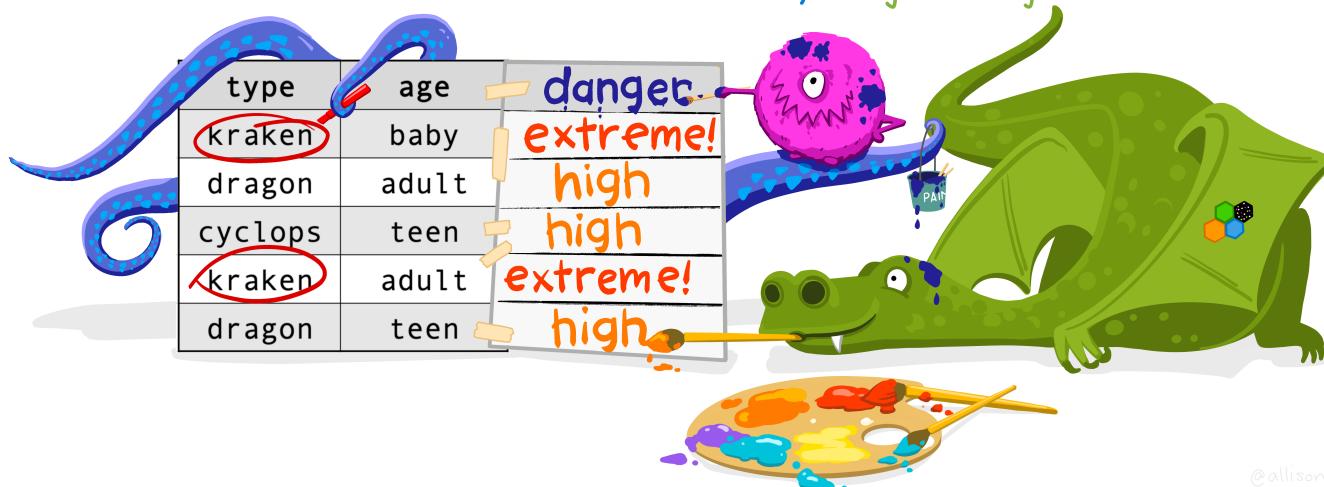
IF ELSE...  
(but you love it?)

df %>% ADD COLUMN  
'danger'

mutate(danger = case\_when(type == "kraken" ~ "extreme!",  
TRUE ~ "high"))

OTHERWISE, danger is high.

danger is  
extreme!



@allison\_horst

# Intervalo - 5 minutos

background-size: cover background-image: ("Apresentacao\_files/break.jpg")

# visualizando os dados (imagem)



# ggplot e camadas de bolo

# entendendo os dados

imagem dos 3 gráficos e análise

# Como baixar os dados da PNAD?

```
#subir o pacote
library(PNADcIBGE)

#determinar as variáveis
variaveis <- c("UF",
              "V1027",
              "V1028",
              "V2007",
              "V2009",
              "V2010",
              "V4001",
              "V4005",
              "V4071",
              "V4073",
              "V4074A",
              "V4077")

#Download dos dados da PNAD Contínua
dadosPNADc_2020_3tri <- get_pnadc(year = 2020, quarter = 3,
                                      vars = variaveis, labels = TRUE)

#Colocando os dados em um objeto - atribuindo valor a um objeto
pnad_2020_3tri <- dadosPNADc_2020_3tri$variables
```

# bem vindos a comunidade



# Referencias e próximos passos

- r4ds
- [http://rstudio-pubs-static.s3.amazonaws.com/528973\\_5c974ffebf02480f80b091daa4fd2514.html#47\\_regs](http://rstudio-pubs-static.s3.amazonaws.com/528973_5c974ffebf02480f80b091daa4fd2514.html#47_regs)
- documentos da PNAD CONTINUA
- [https://biblioteca.ibge.gov.br/visualizacao/periodicos/2421/pnact\\_2020\\_3tri.pdf](https://biblioteca.ibge.gov.br/visualizacao/periodicos/2421/pnact_2020_3tri.pdf)

# dicas de conteúdo