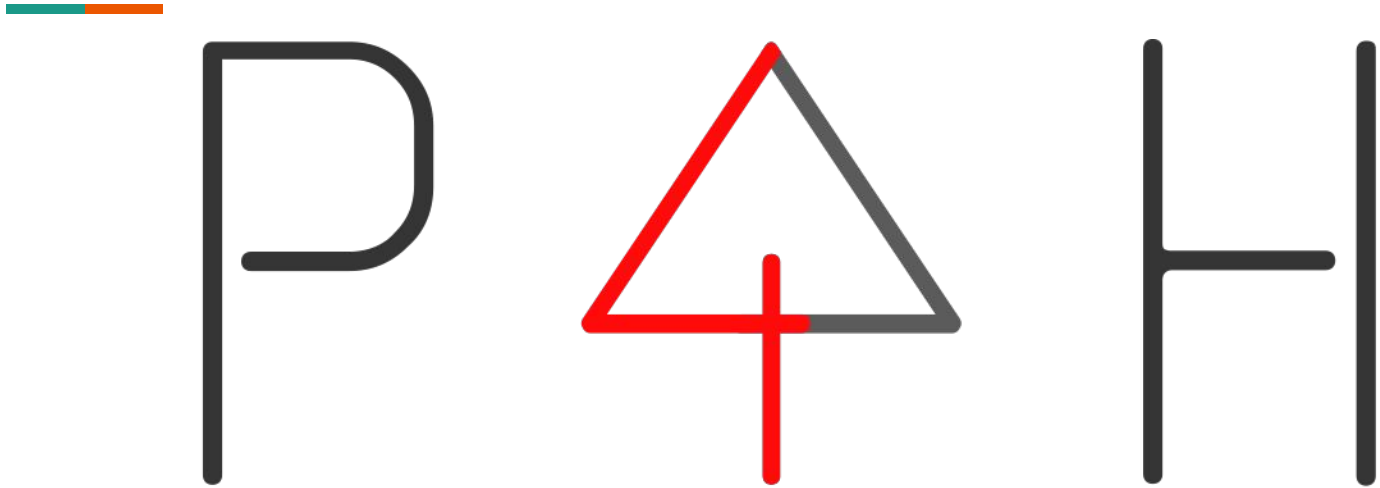
Workshop R: Manipulação de Dados

PROGRAMAÇÃO PARA HUMANIDADES



PROGRAMAÇÃO PARA HUMANIDADES

Acompanhe a aula pelo nosso site!

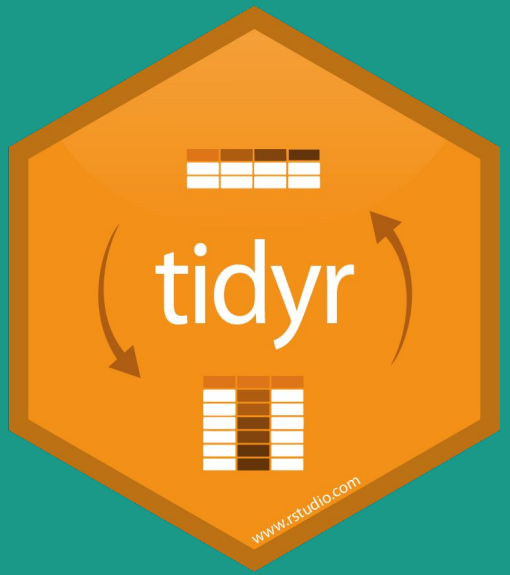
github

bit.ly/p4h_workshop

P4H

PROGRAMAÇÃO PARA HUMANIDADES

Tidy data e tidyr



P4H
PROGRAMAÇÃO PARA HUMANIDADES

Tidy data

Por que precisamos nos importar com isso?

1. Linhas: observações
2. Colunas: variáveis
3. Células: valores

P4H
PROGRAMAÇÃO PARA HUMANIDADES

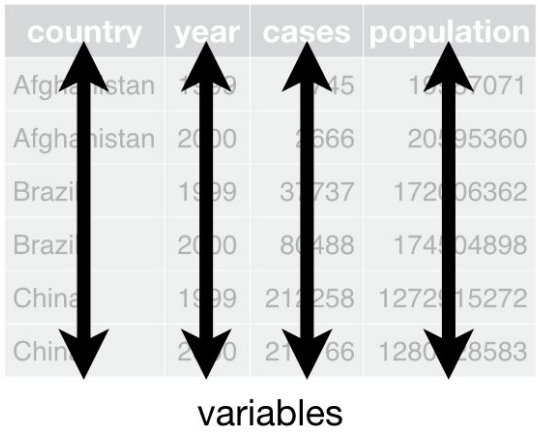
Exemplo de um banco não *tidy*

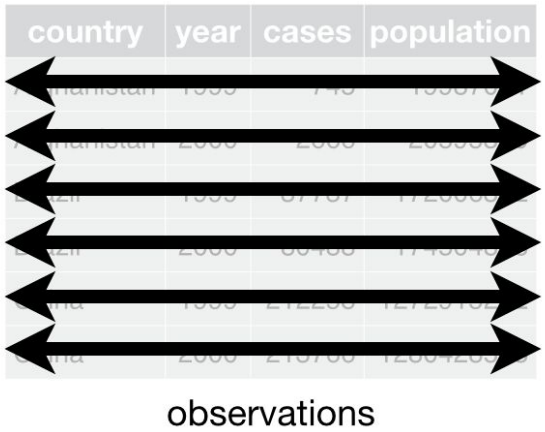
País	2016	2017
Brasil	9.2	12.2
China	7.1	9.3

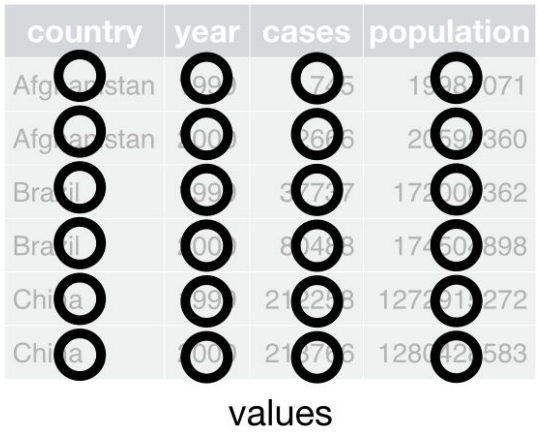
Exemplo de um banco *tidy*

País	Ano	Desemprego
Brasil	2016	9.2
Brasil	2017	12.2
China	2016	7.1
China	2017	9.2

Modelo *tidy*







PROGRAMAÇÃO PARA HUMANIDADES

Os verbos do pacote dplyr

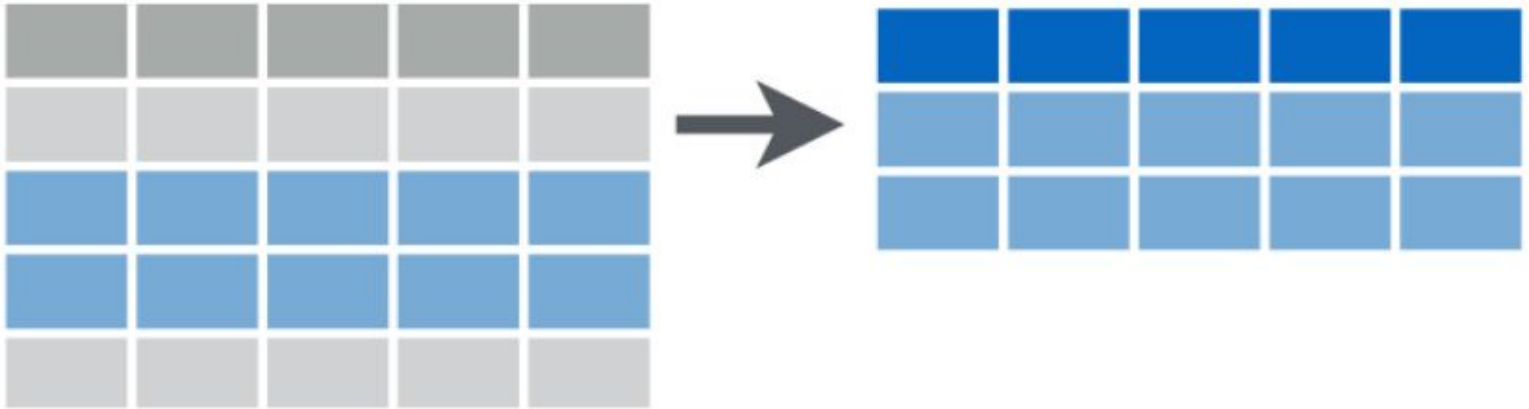
P4H
PROGRAMAÇÃO PARA HUMANIDADES

dplyr para manipular bancos de dados

1. filter()
2. mutate()
3. rename()

P4H
PROGRAMAÇÃO PARA HUMANIDADES

filter()



mutate()

dplyr para resumir bancos de dados

1. count()
2. group_by() e summarise()

PROGRAMAÇÃO PARA HUMANIDADES

group_by() e summarise()

Próximo módulo: visualização de dados

1ª turma: **09/10**

2ª turma: **11/10**

P4H
PROGRAMAÇÃO PARA HUMANIDADES