Rebeca de Jesus Carvalho

FGV CEPESP rebeca.jesus.carvalho@gmail.com

- Introdução
- Bases relacionais com a gramática do dplyr
- R e Power Bl
- Laboratório

### Conteúdo da aula de hoje

- Veremos como combinar *data frames* de diferentes origens que se relacionam por meio de uma ou mais variáveis chave.
- Para isso, utilizaremos novos verbos do *dplyr*, de sufixo \_join para trabalhar com bases de dados relacionais.
- Por fim, aprenderemos como integrar R e Power Bl.

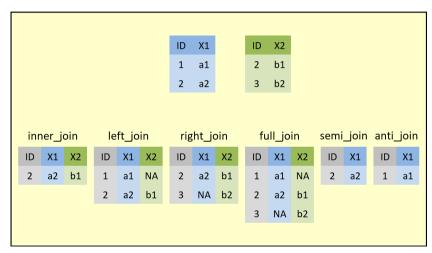
#### Combinando data frames

- Combinar data frames é necessário quando as informações que serão utilizadas na análise estão presentes em mais de uma fonte de dados.
- As variáveis usadas para conectar um par de data frames são chamadas de chaves. Uma chave é uma variável (ou conjunto de variáveis) que identifica exclusivamente uma observação. Em casos simples, uma única variável é suficiente para identificar uma observação.

#### Tipos de *join*

- Há 6 tipos de "joins" em R:
  - Interseção dos dois conjuntos, x e y: inner\_join();
  - Mantém todas as observações em x, ou seja, à esquerda: left\_join();
  - Mantém todas as observações em y, ou seja, à direita: right\_join();
  - Mantém todas as observações em x e y: full\_join();
  - Mantém todas as observações em x que correspondem a y: semi\_join(); e
  - Elimina todas as observações em x que correspondem a y: anti\_join():

#### Visão geral das funções de join do dplyr



algoritmos implementados na linguagem.

# • É possível importar e modelar dados dentro do Power BI com scripts de R, tais quais os que trabalhamos no curso, além de produzir gráficos ou rodar

R e Power BI

- Entre as vantagens dessa integração, destacam-se, sobretudo:
  - A possibilidade de criar visualizações personalizadas, que não estão disponíveis nativamente no Power BI;
  - A vasta biblioteca de pacotes estatísticos do R que podem ser incorporados ao Power BI; e
  - Importação e análise de dados de diferentes fontes diretamente no Power BI.

#### Laboratório

Agora é o momento de partir para a ação! Temos dois tutoriais agendados para hoje, prontos para serem explorados. Eles estão disponibilizados no GitHub, e você pode acessá-los clicando neste **link**.

Se deixou algum tutorial inacabado dos encontros anteriores, comece por eles. Caso contrário, prossiga.

## Dúvidas?