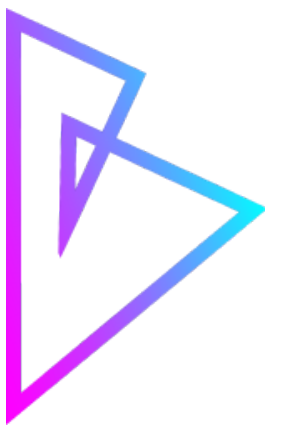


# Aula 22

Joining dados no R



# Relational data



# O que é **relational** data?

- **O que se espera ao trabalhar com dados?**
  1. Muitos dataframes;
  2. Banco de dados armazenando dados (SQL);
  3. Um ID para uni-los todos (felizmente).

# O que é **relational** data?

- **O que se espera ao trabalhar com dados?**

1. Muitos dataframes;
2. Banco de dados armazenando dados (SQL);
3. Um ID para uni-los todos (felizmente).

- **Então, como trabalhar com isso?**

Merging/Joining the data (as tabelas)!

Ou seja, colocamos juntas informações de diferentes tabelas.

# O que é **relational** data?

- **O que se espera ao trabalhar com dados?**
  1. Muitos dataframes;
  2. Banco de dados armazenando dados (SQL);
  3. Um ID para uni-los todos (felizmente).

student_id	name	age
001	Natalia	25
002	Arthur	30

student_id	course
001	EMP
002	MPP

joined

student_id	name	age	course
001	Natalia	25	EMP
002	Arthur	30	MPP

- **Então, como trabalhar com isso?**

Merging/Joining the data (as tabelas)!

Ou seja, colocamos juntas informações de diferentes tabelas.

# Tipos de **join** no R

- `inner_join`
  - `left_join`
  - `right_join`
  - `full_join`
- } **mutating joins**
- `semi_join`
  - `anti_join`
- } **filtering joins**

# Mutating joins

# inner\_join: Descarta observações não encontradas

**table1**

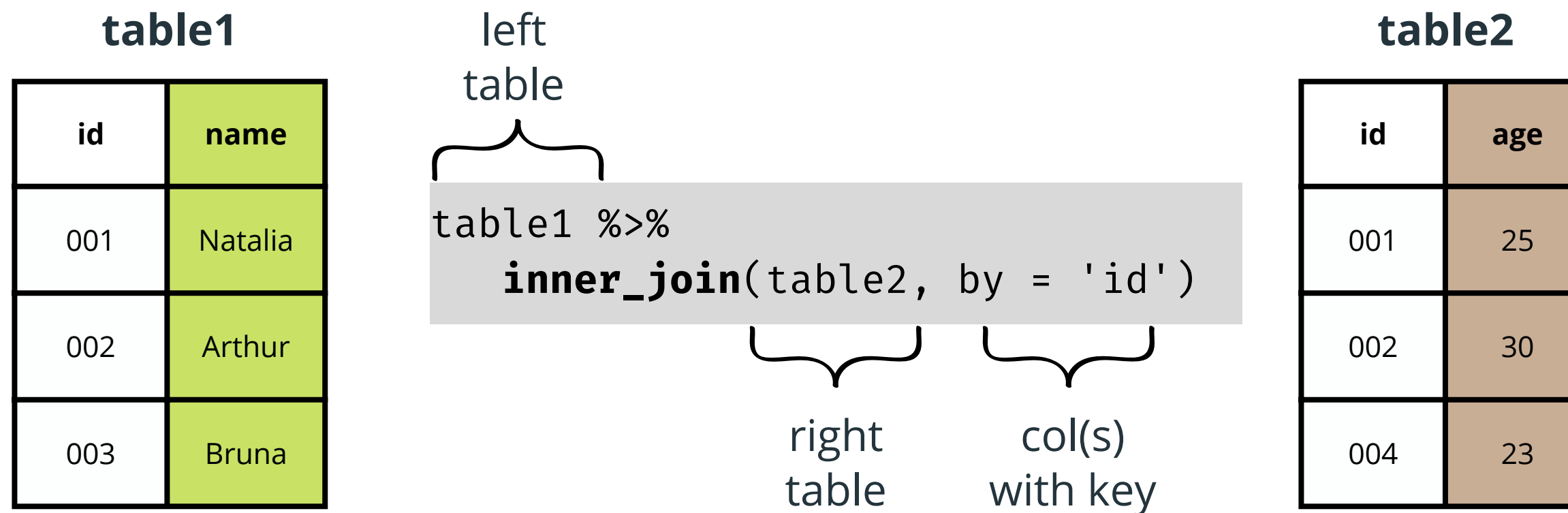
id	name
001	Natalia
002	Arthur
003	Bruna

**table2**

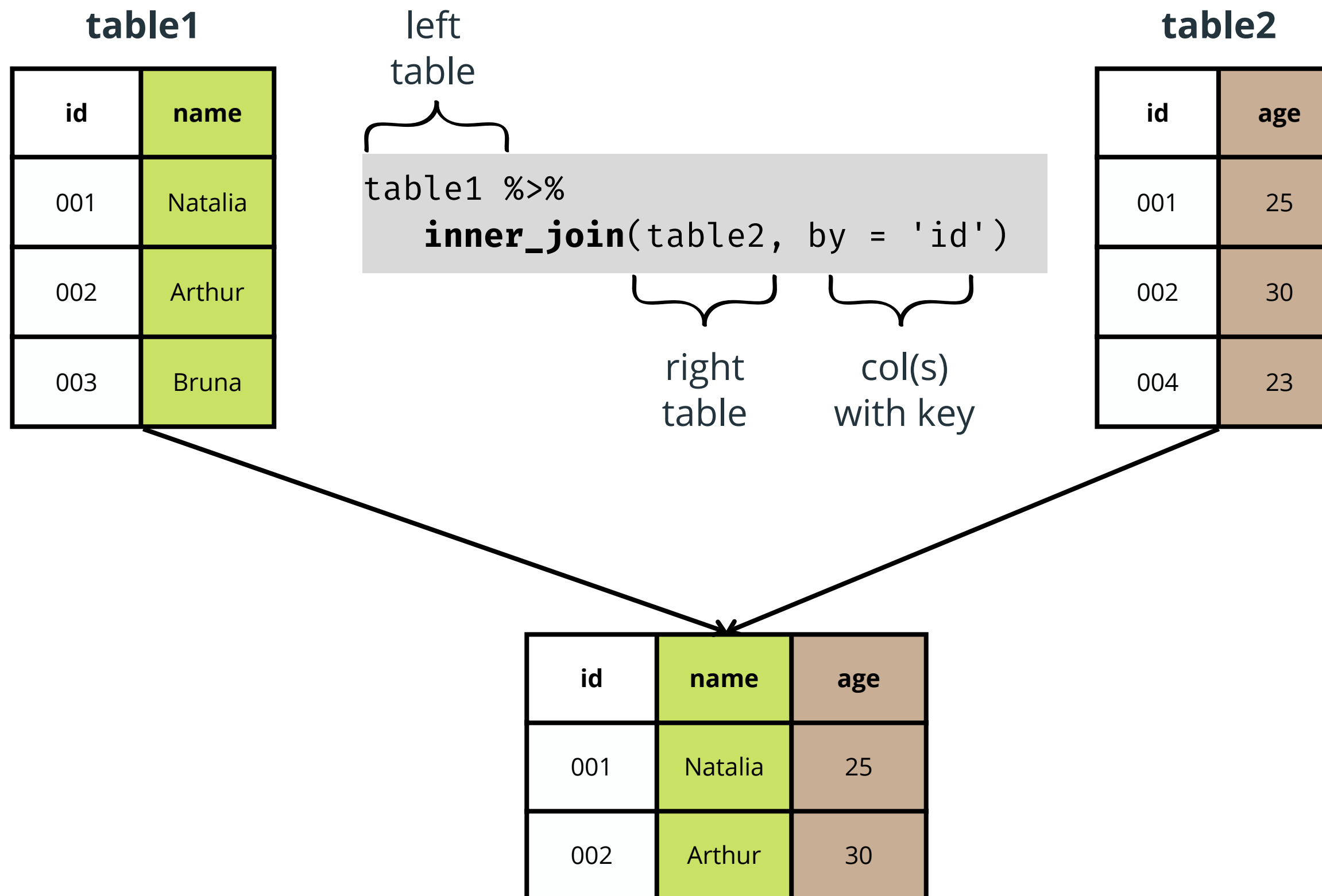
id	age
001	25
002	30
004	23



# inner\_join: Descarta observações não encontradas



# inner\_join: Descarta observações não encontradas



# left\_join: mantém as observações da table1

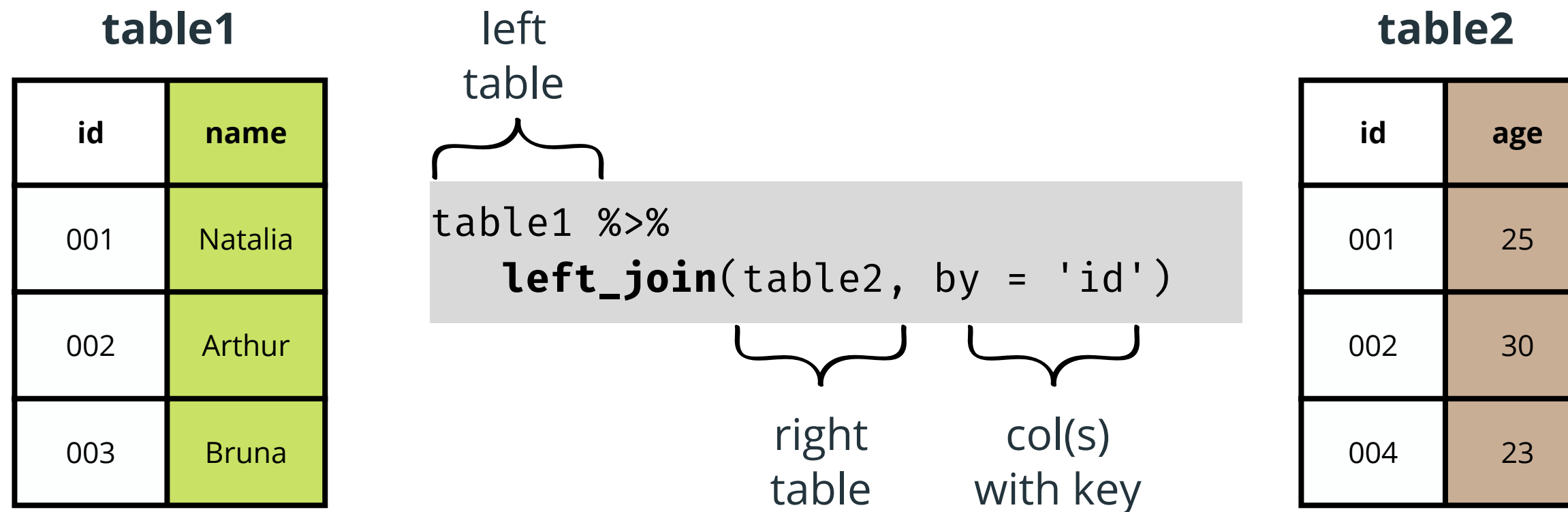
**table1**

id	name
001	Natalia
002	Arthur
003	Bruna

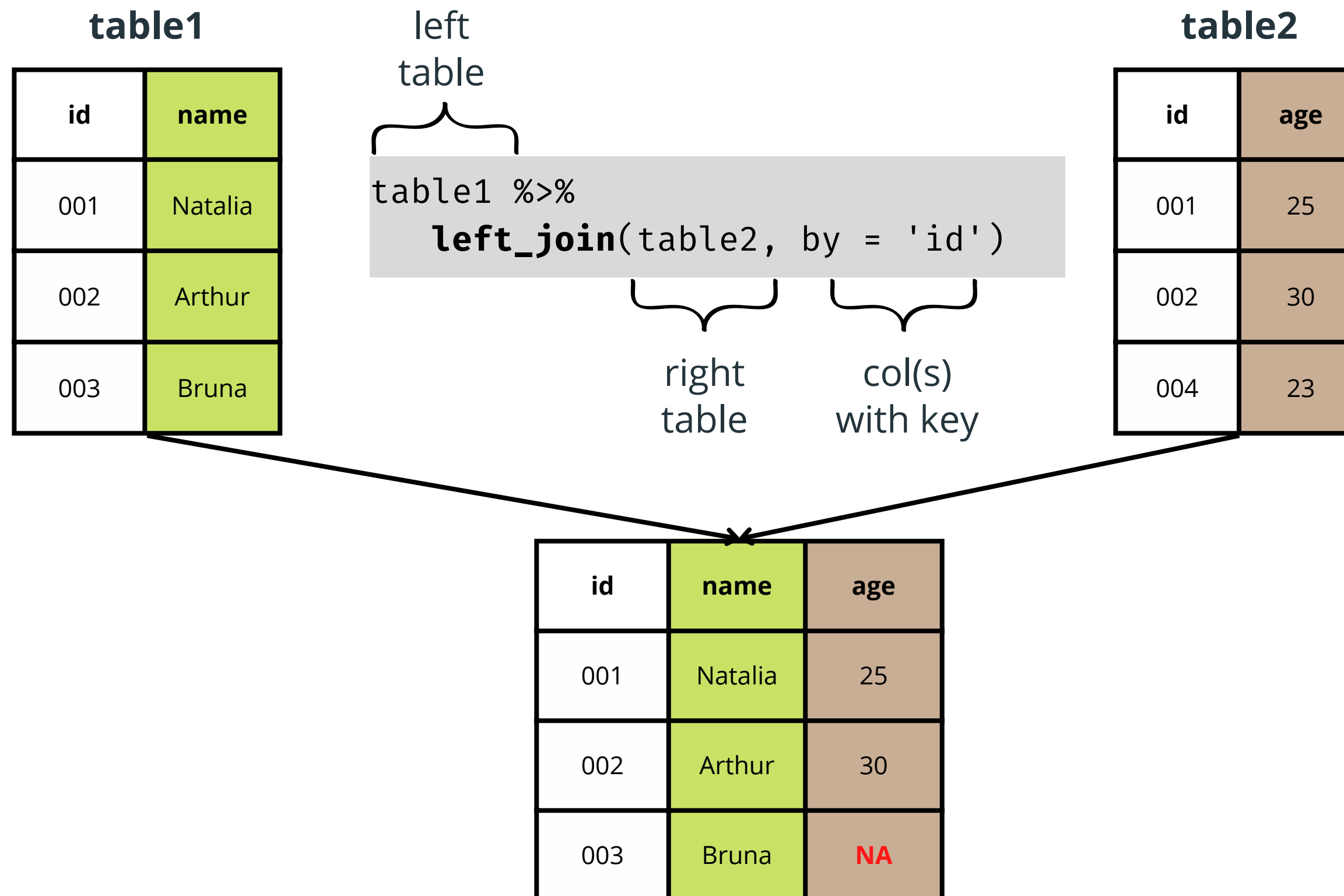
**table2**

id	age
001	25
002	30
004	23

# left\_join: mantém as observações da table1



# left\_join: mantém as observações da table1



# **right\_join:** mantém as observações da table2

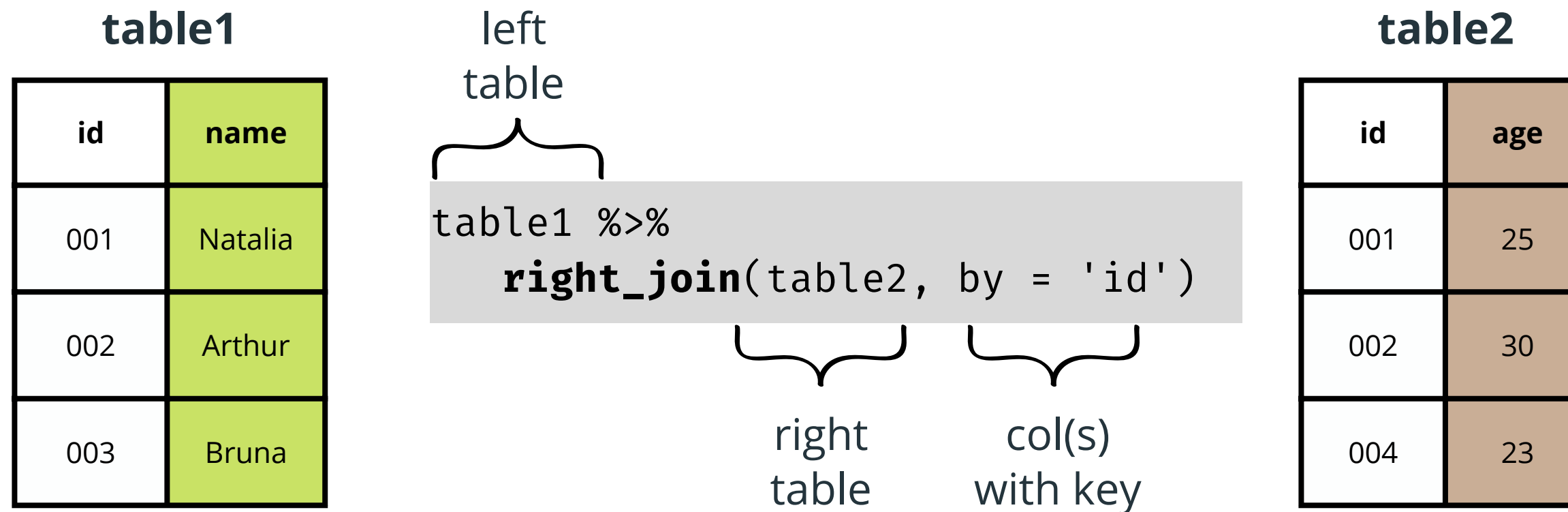
**table1**

id	name
001	Natalia
002	Arthur
003	Bruna

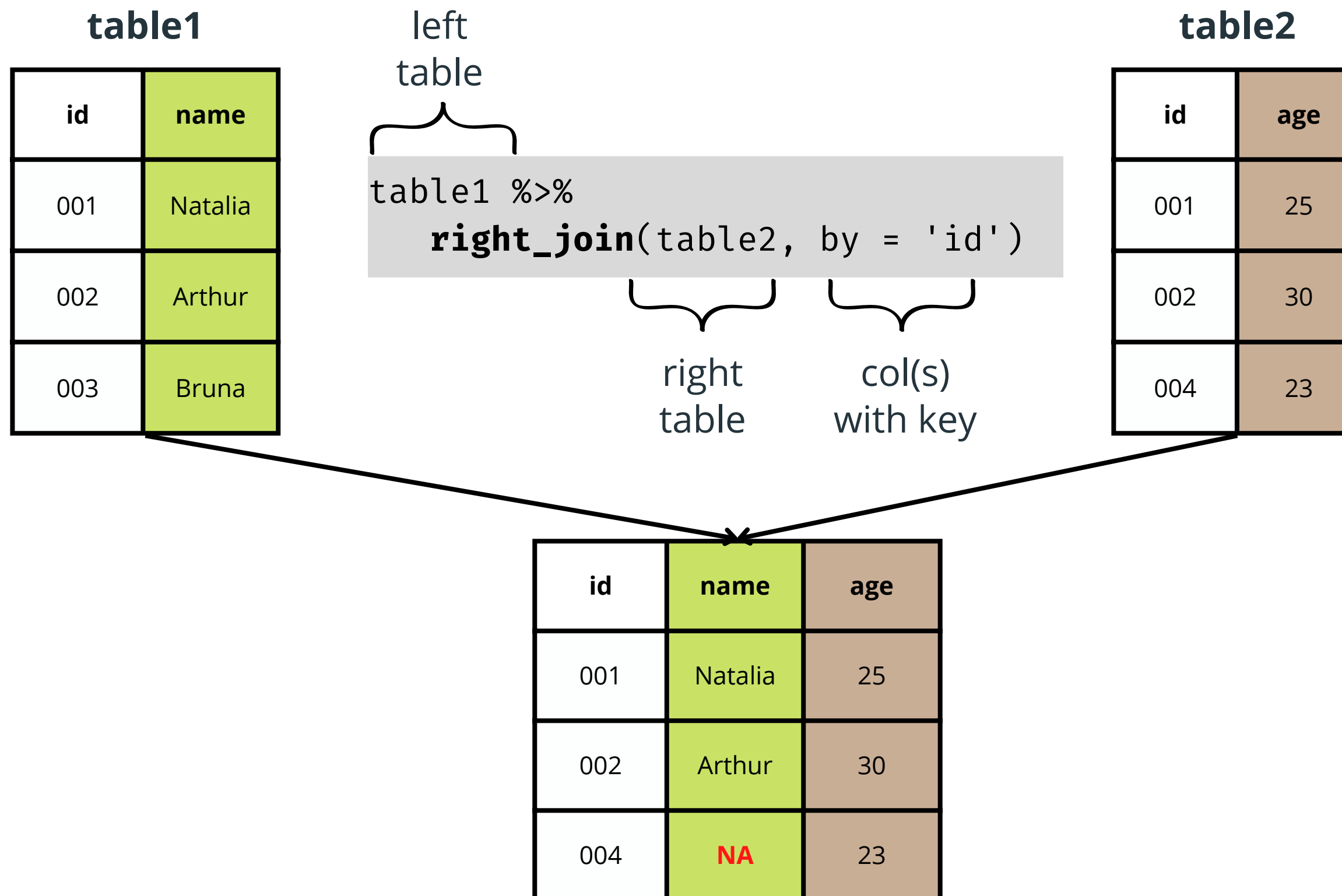
**table2**

id	age
001	25
002	30
004	23

# **right\_join:** mantém as observações da table2



# right\_join: mantém as observações da table2





# **full\_join:** mantém as informações de ambas tables

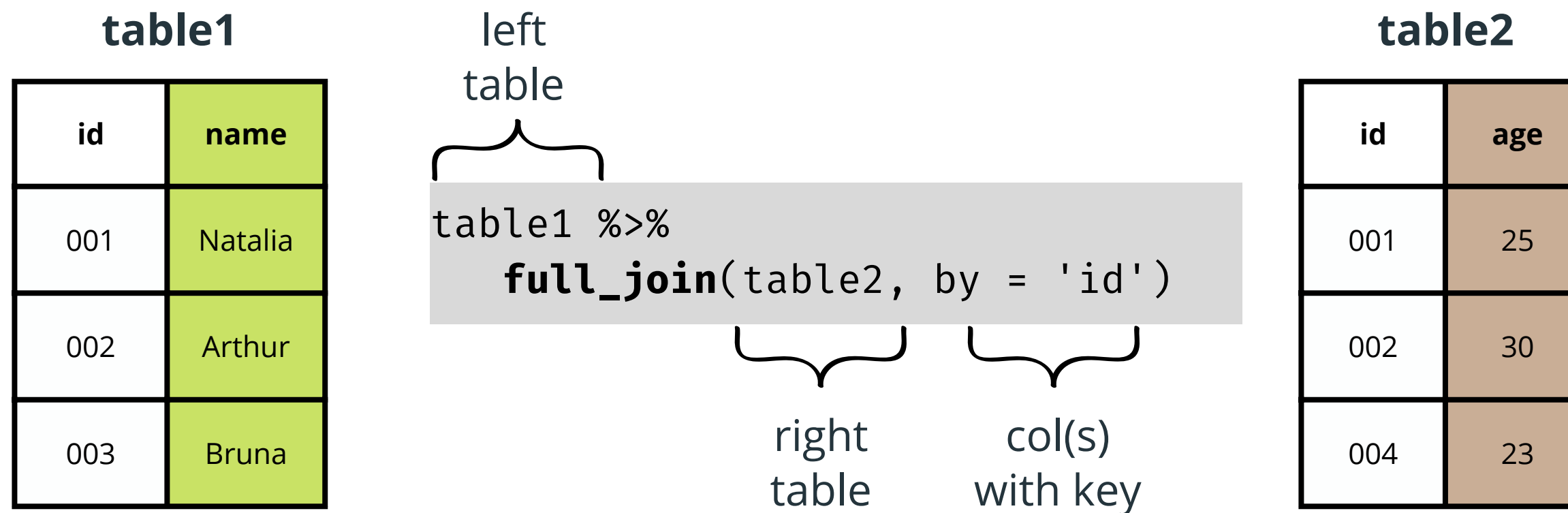
**table1**

id	name
001	Natalia
002	Arthur
003	Bruna

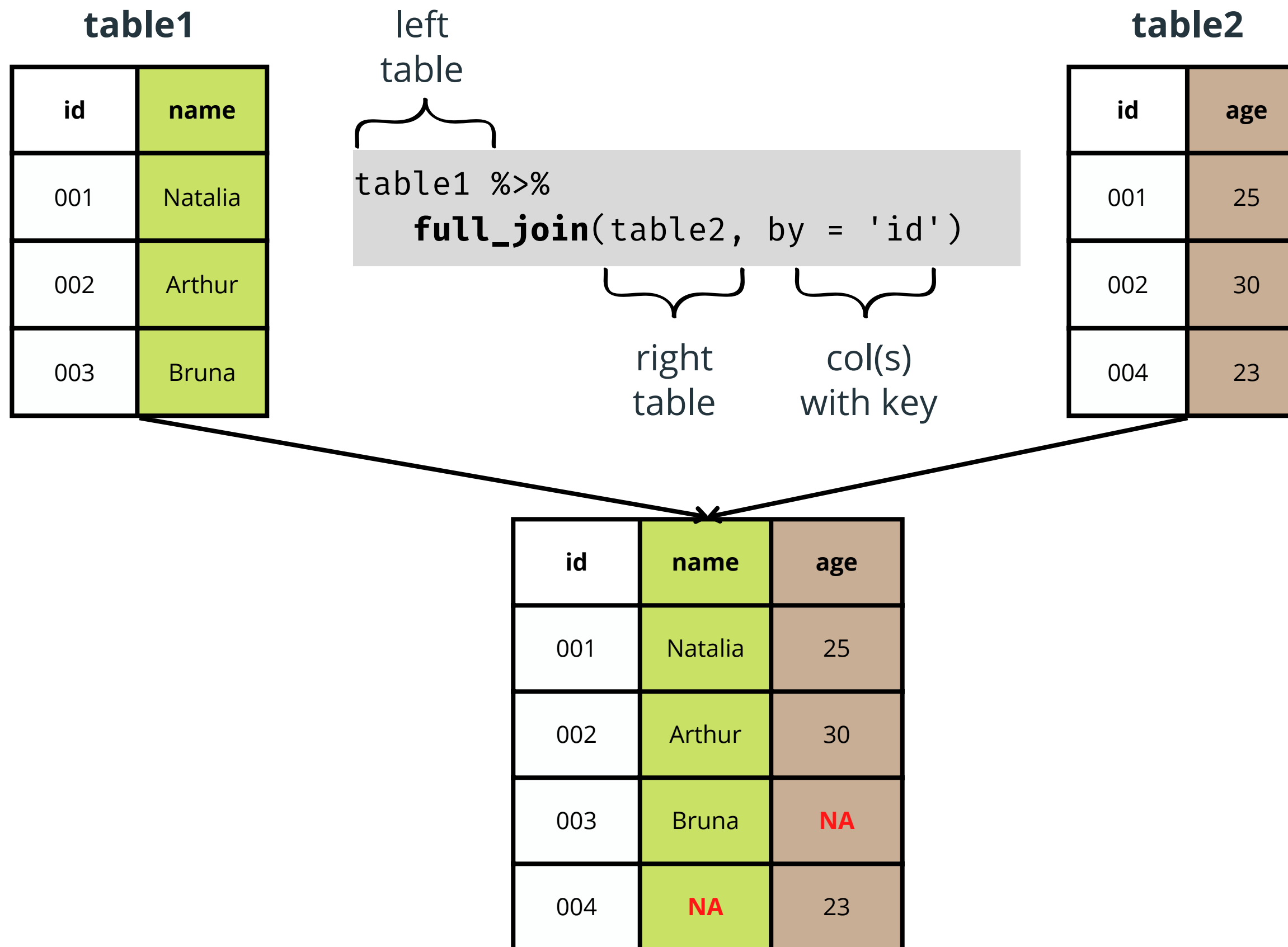
**table2**

id	age
001	25
002	30
004	23

# **full\_join:** mantém as informações de ambas tables



# full\_join: mantém as informações de ambas tables



# Filtering joins

**semi\_join:** mantém, na table1, as informações matched

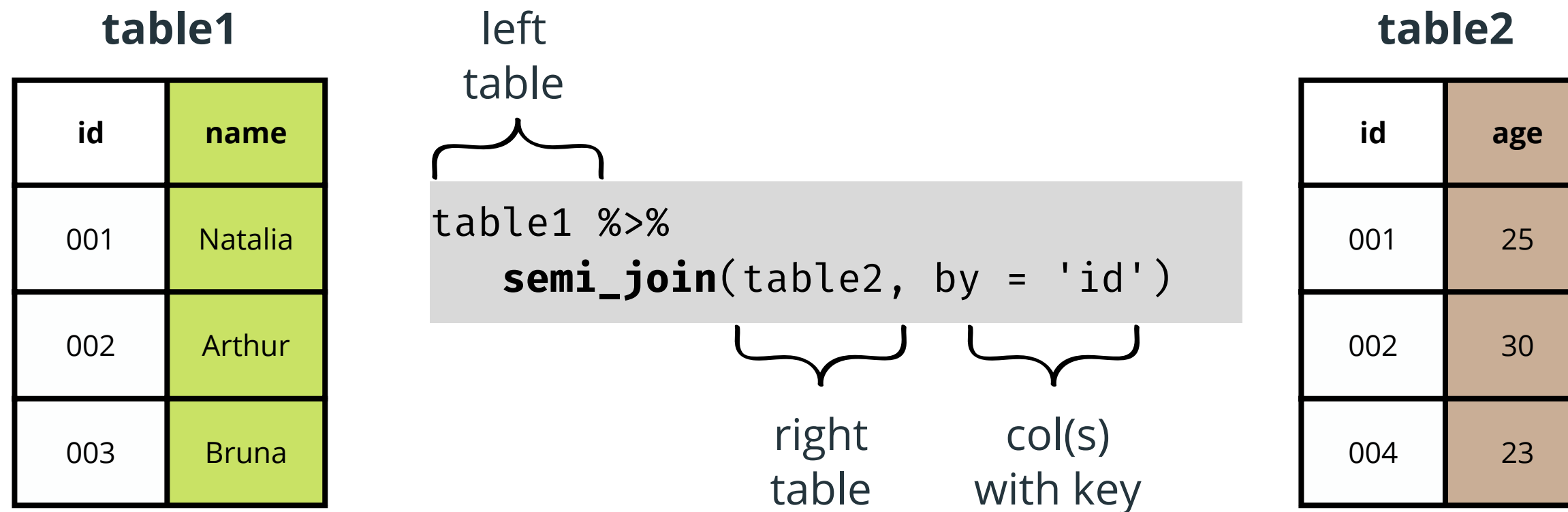
**table1**

id	name
001	Natalia
002	Arthur
003	Bruna

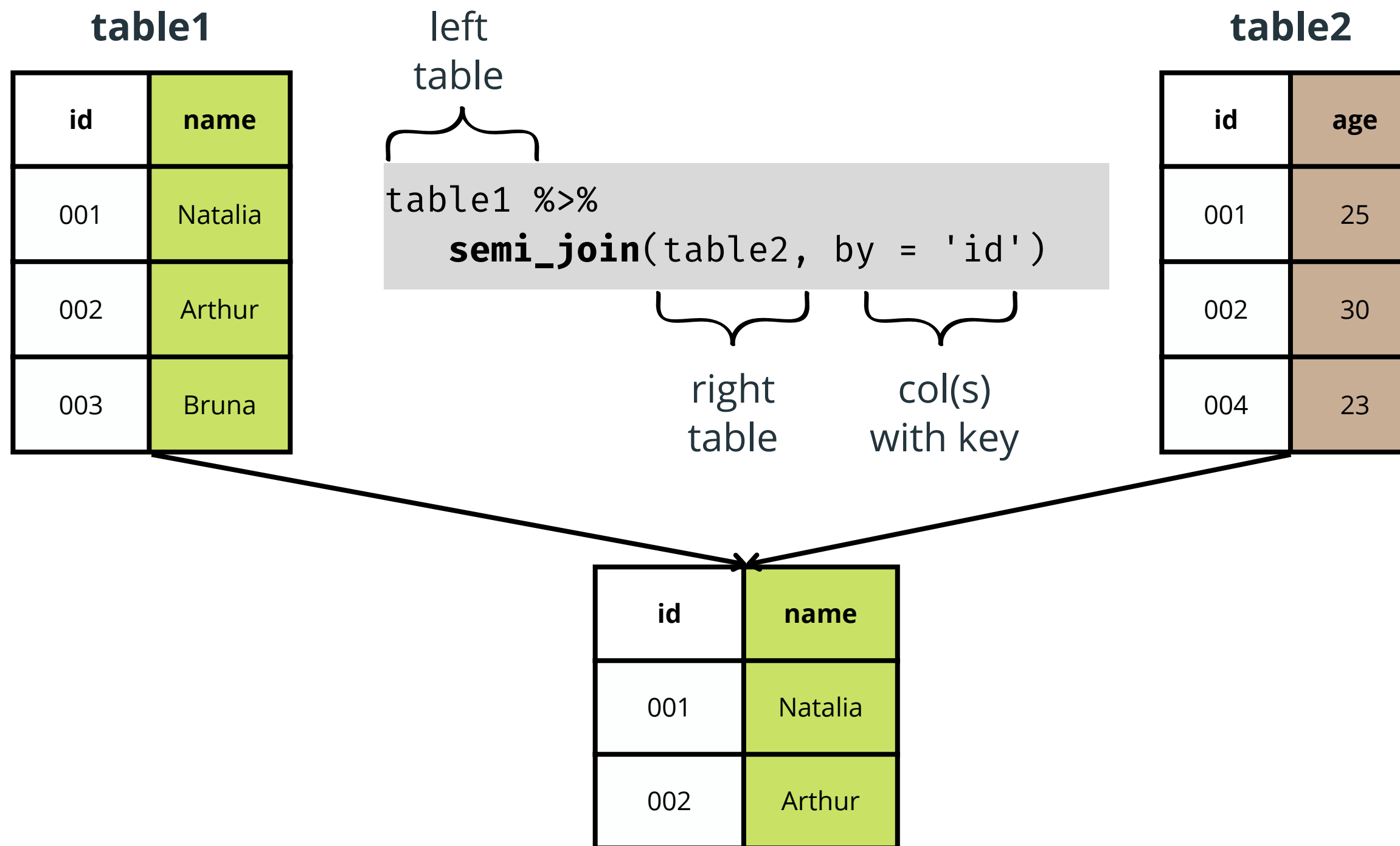
**table2**

id	age
001	25
002	30
004	23

**semi\_join:** mantém, na table1, as informações matched



**semi\_join:** mantém, na table1, as informações matched



# anti\_join: mantém as non- matched observations

**table1**

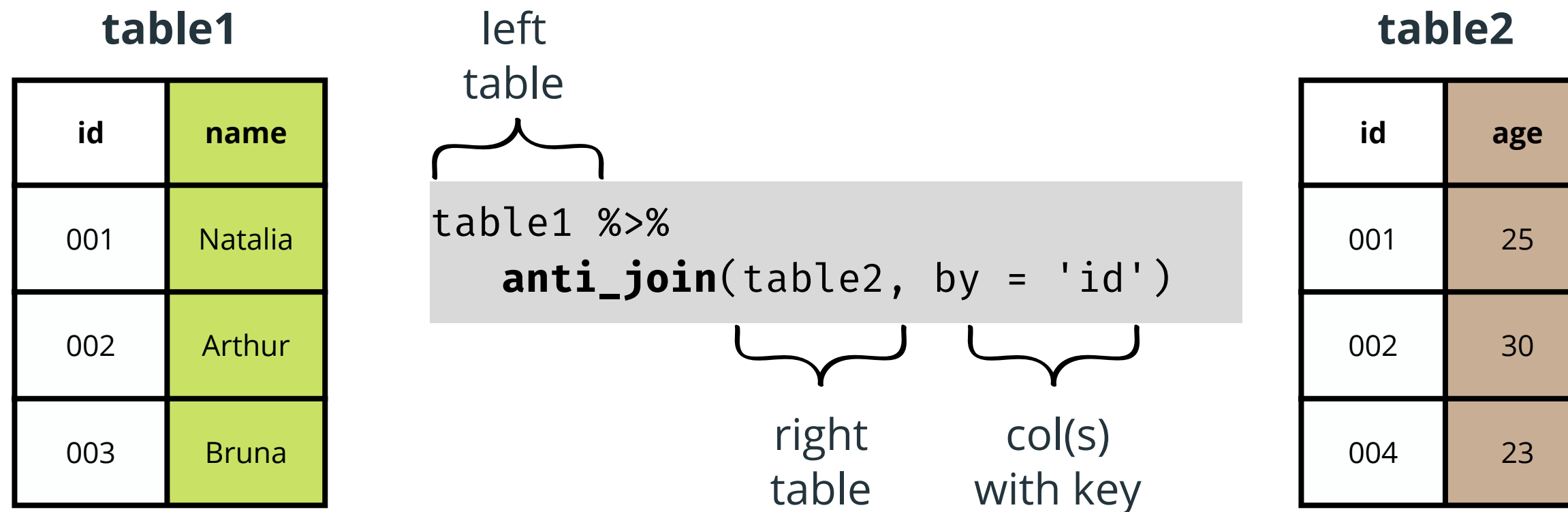
id	name
001	Natalia
002	Arthur
003	Bruna

**table2**

id	age
001	25
002	30
004	23



# anti\_join: mantém as non- matched observations



# anti\_join: mantém as non- matched observations

