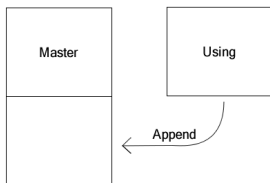


Econometría I

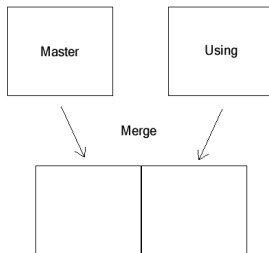
Bases de datos

Hernando Hernandez & Julián Rojas

Bogotá, 2 de mayo de 2020



(a) Concatenación.



(b) Fusión.

Figura 1: Diferencias al manipular bases de datos.

Dependiendo de qué tantos datos se quiera conservar, se puede:

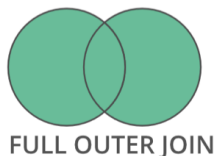
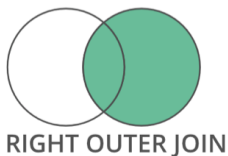
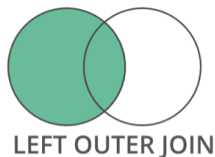
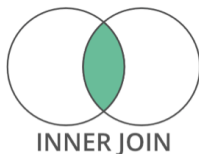


Figura 2: Distintas formas de fusionar.

Al usar la opción (1b), hay varias alternativas para proceder. A saber:

- **Uno-a-uno:** La más sencilla de todas.

	Empleado	Profesión		Empleado	Contratación
0	Mateo	Contador público	0	Valeria	2023
1	Sara	Economista	1	Andrés	2021
2	Andrés	Ingeniero industrial	2	Sara	2020
3	Valeria	Administradora de empresas	3	Mateo	2022

Cuadro 1: Problema enfrentado, *master* vs. *using*.

	Empleado	Profesión	Contratación
0	Mateo	Contador público	2022
1	Sara	Economista	2020
2	Andrés	Ingeniero industrial	2021
3	Valeria	Administradora de empresas	2023

Cuadro 2: Solución planteada.

- ▶ **Muchos-a-uno:** La clave debe ser única para los datos en disco.
- ▶ **Uno-a-muchos:** La clave debe ser única para los datos en memoria.
- ▶ **Muchos-a-muchos:** Suele ser mejor evitarlo.

Aclaración

La información ha sido extraída de Stata. Por tanto, datos en disco hace referencia a las bases de datos que no se encuentren cargadas en el programa econométrico. Otras fuentes de información se encuentran en GitHub y RPubS. Se recomienda su exploración.