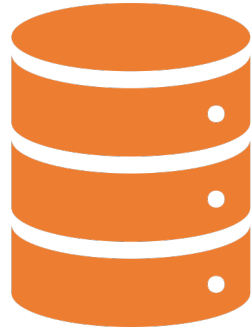


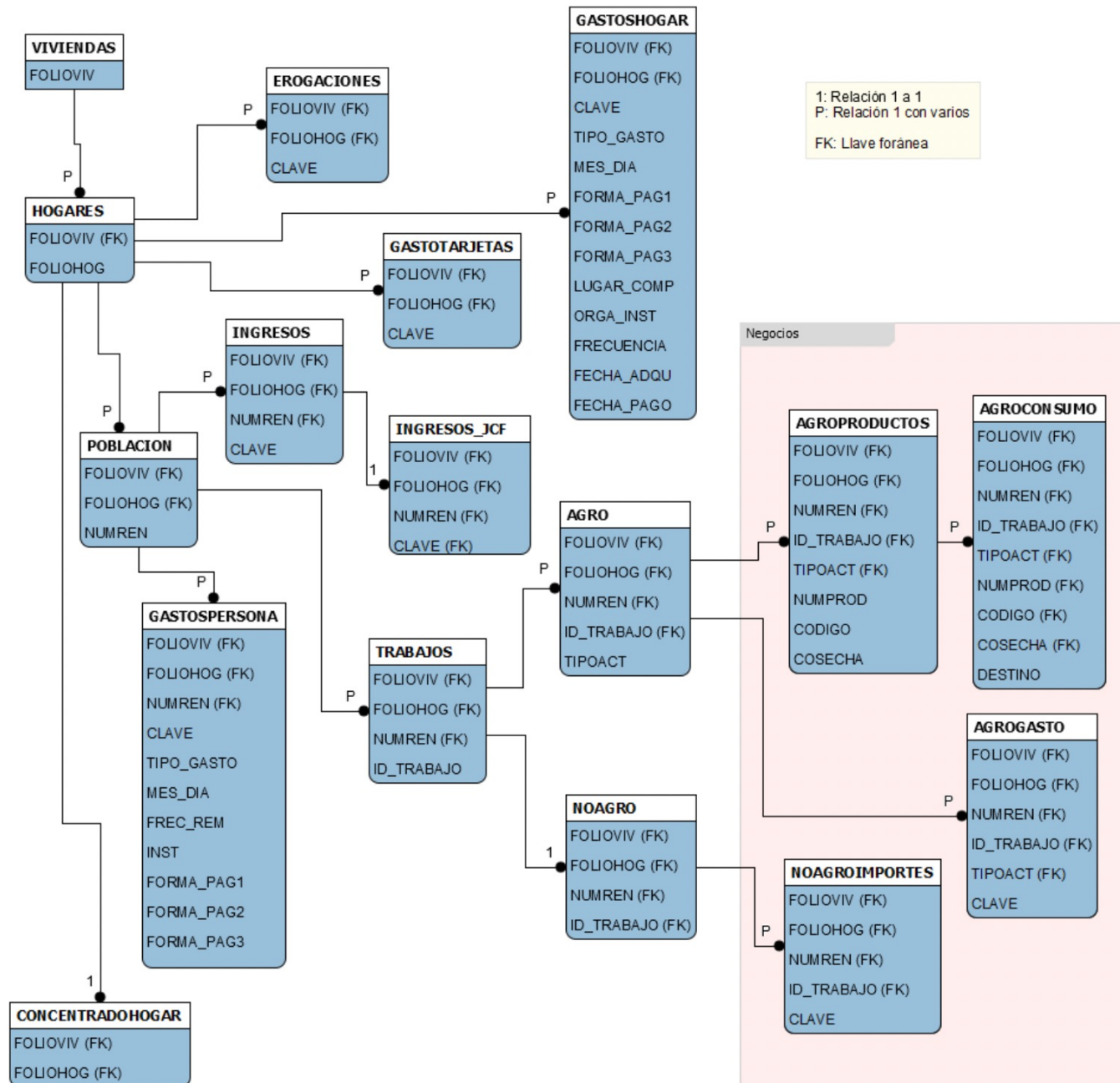
# ¡Descargamos la base de datos????



Una **base de datos** (del inglés: database) se encarga no solo de almacenar datos, sino también de conectarlos entre sí en una unidad lógica



El **dataframe o conjuntos de datos** normalmente es una tabla o matriz de datos



# Cardinalidad

## Correspondencia de cardinalidades [\[ editar \]](#)

Dado un conjunto de relaciones en el que participan dos o más conjuntos de entidades, la cardinalidad de la correspondencia indica el número de entidades con las que puede estar relacionada una entidad dada.

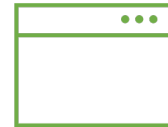
Dado un conjunto de relaciones binarias y los conjuntos de entidades A y B, las cardinalidades pueden ser:

- **Uno a Uno:** (1:1) Un registro de una entidad A se relaciona con solo un registro en una entidad B. (ejemplo dos entidades, profesor y departamento, con llaves primarias, código\_profesor y jefe\_depto respectivamente, un profesor solo puede ser jefe de un departamento y un departamento solo puede tener un jefe).
- **Uno a Varios:** (1:N) Un registro en una entidad en A se relaciona con uno o muchos registros en una entidad B. Pero los registros de B solamente se relacionan con un registro en A. (ejemplo: dos entidades, vendedor y ventas, con llaves primarias, código\_vendedor y venta, respectivamente, un vendedor puede tener muchas ventas pero una venta solo puede tener un vendedor).
- **Varios a Uno:** (N:1) Una entidad en A se relaciona exclusivamente con una entidad en B. Pero una entidad en B se puede relacionar con 1 o muchas entidades en A (ejemplo empleado-centro de trabajo).
- **Varios a Varios:** (N:M) Una entidad en A se puede relacionar con 1 o con muchas entidades en B y viceversa (ejemplo asociaciones-ciudadanos, donde muchos ciudadanos pueden pertenecer a una misma asociación, y cada ciudadano puede pertenecer a muchas asociaciones distintas).

1 a 1



Aquí hay  
unas variables

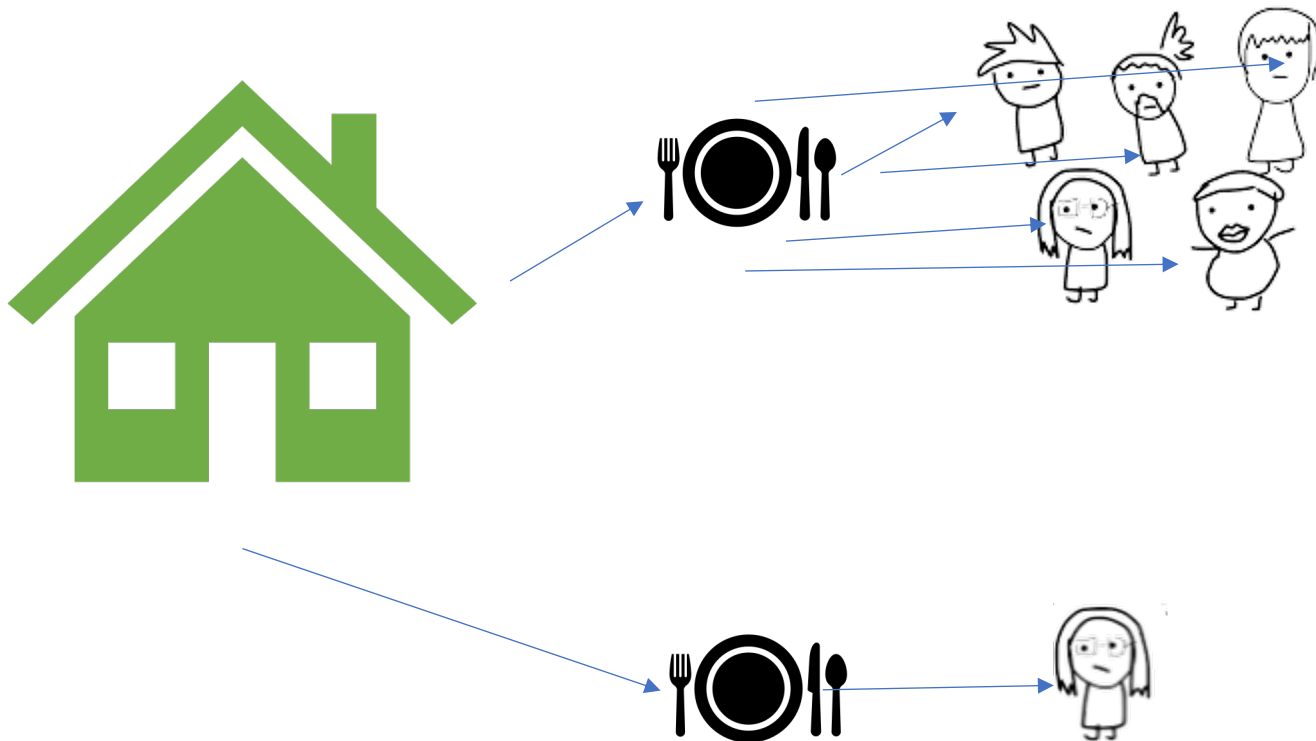


Aquí hay  
otras variables

merge  
join  
fusionado



Aquí tendremos  
todas las variables



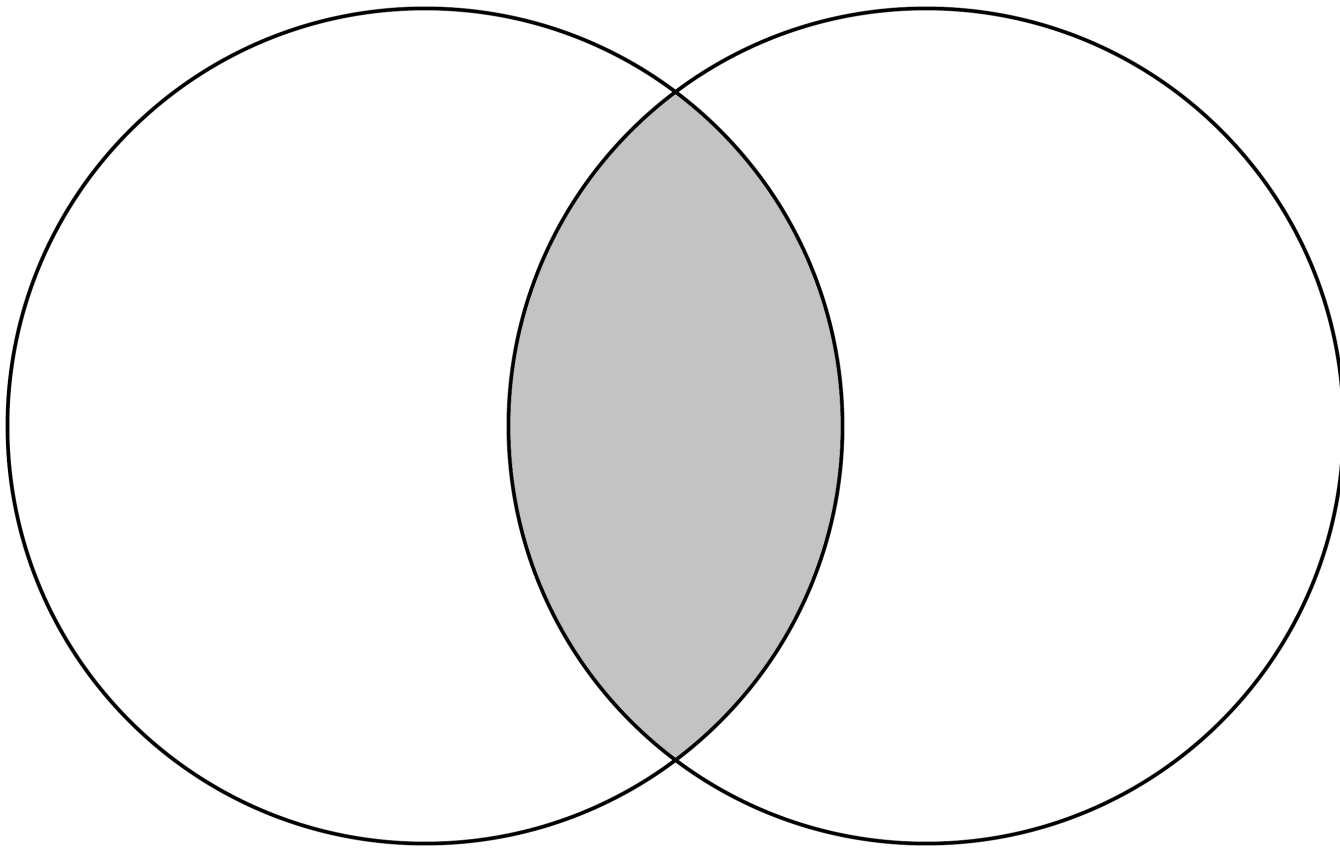
Hay dirección.

El sentido contrario

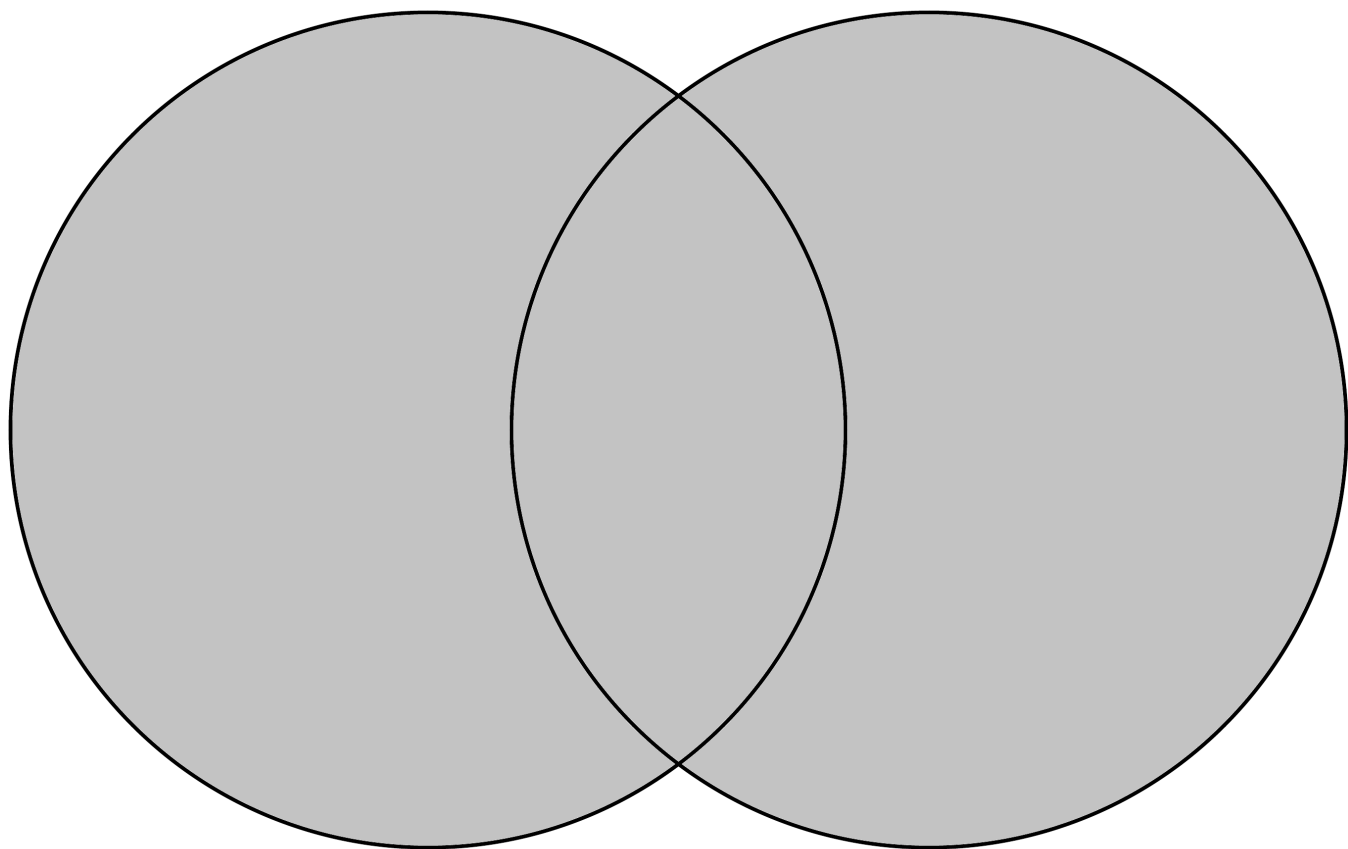
**NO SE PUEDE SIN OPERACIONES  
DE AGREGACIÓN**

# Lógica de la fusión de datos

1 a 1

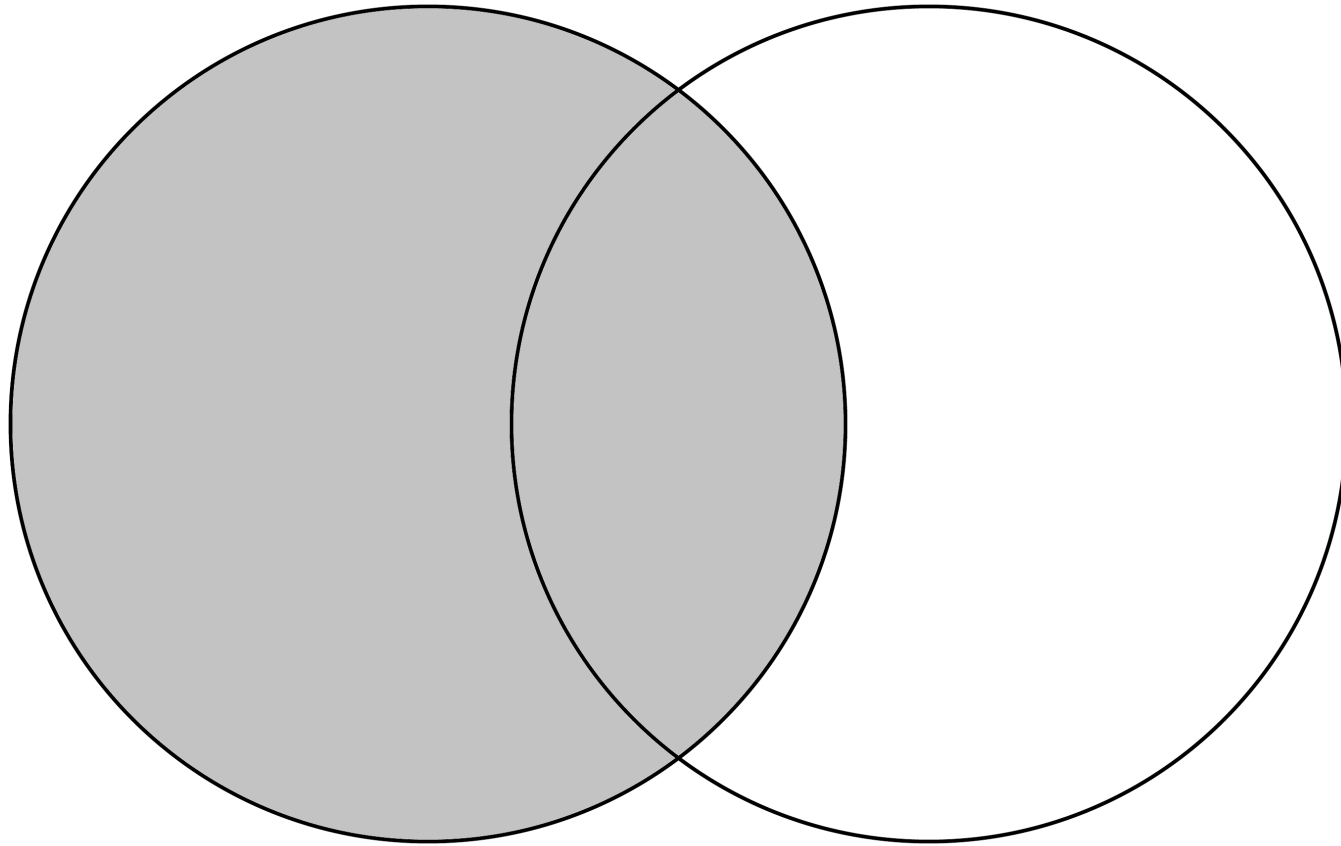


```
merge()  
all=FALSE  
dplyr::inner_join()
```

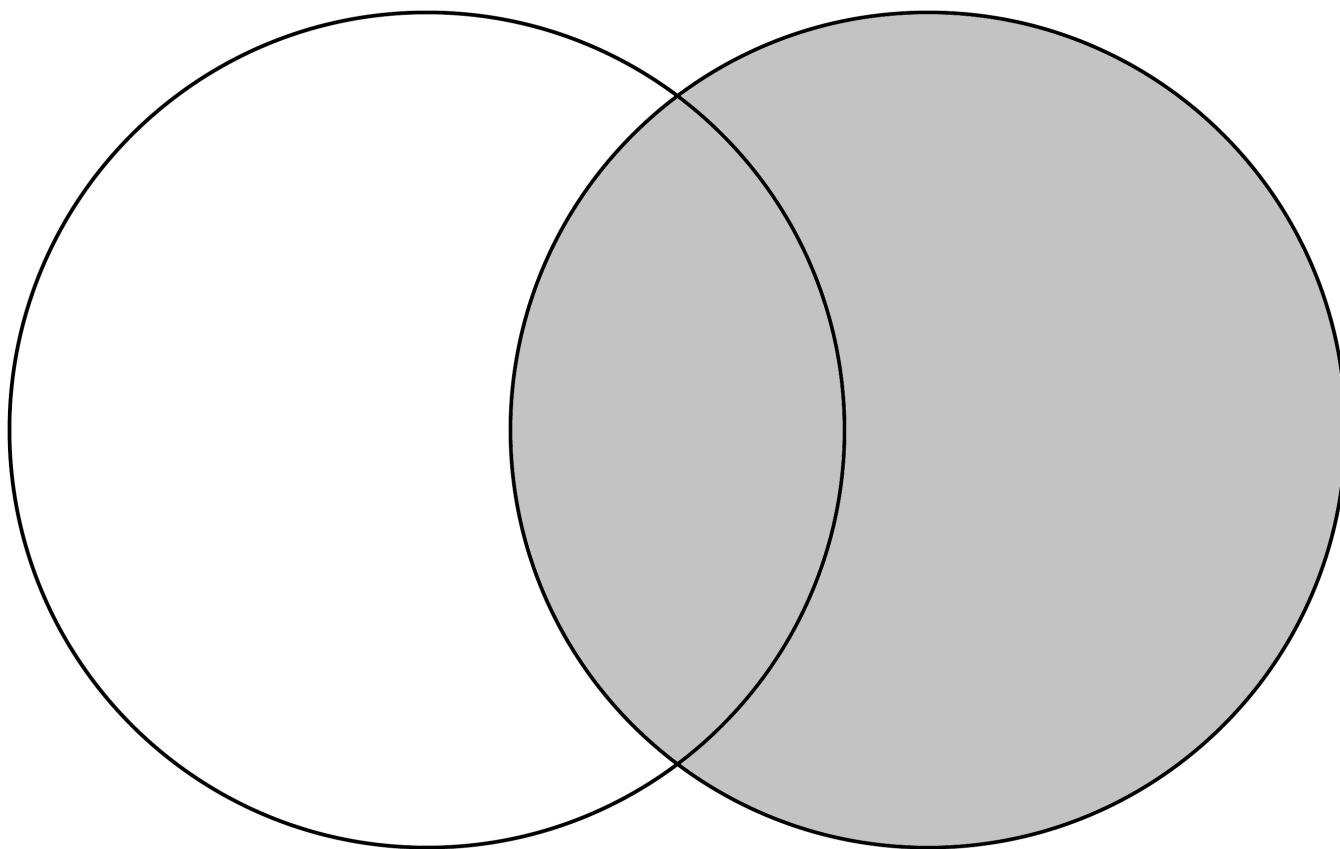


```
merge()  
all=TRUE  
dplyr::full_join()
```





```
merge()  
all.x=TRUE  
dplyr::left_join()
```



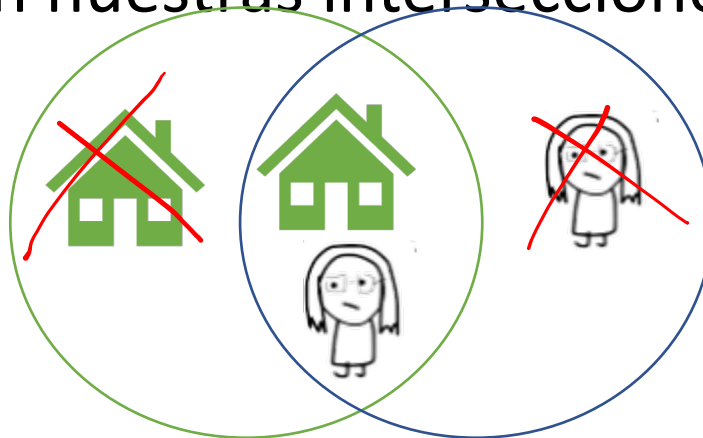
```
merge()  
all.y=TRUE  
dplyr::right_join()
```

# Lógica de la fusión de datos

1 a muchos

# La lógica de la fusión cambia un poco

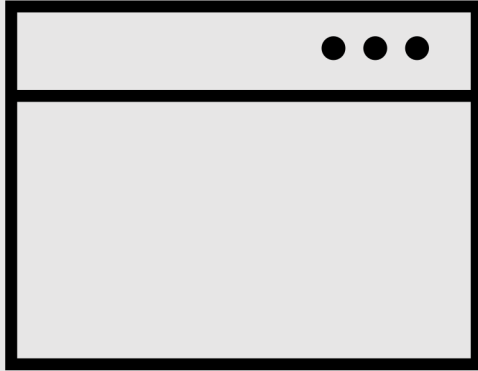
- Lo ideal es empezar desde la unidad “más grande” a las más pequeñas.
- Pensemos cómo se ven nuestras intersecciones ¿cuál sería la opción ideal?



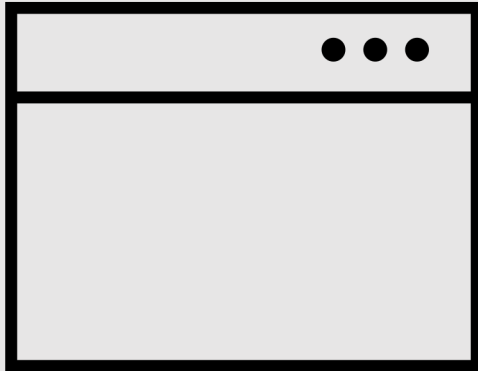
# Agregar datos

append

¿Cómo juntar los 32 estados?



Los nombres deben ser iguales  
Si no los entenderá como otra variable



- `rbind`
- `dplyr::bind_rows`