

Bases de datos: librería dplyr

Librería dplyr: verbos para la administración de bases de datos

- `select ()` → Selecciona columnas
- `filter ()` → Selecciona filas
- `arrange ()` → Ordena con respecto a una variable
- `mutate ()` → Crea nuevas variables
- `summarise ()` → Resumen de estadísticas

Librería dplyr: select

select()



Selecciona un
subconjunto de variables
o columnas

```
> select(BasedeDatos, variable1, variable3, variable4,...)
```

Librería dplyr: mutate

mutate()

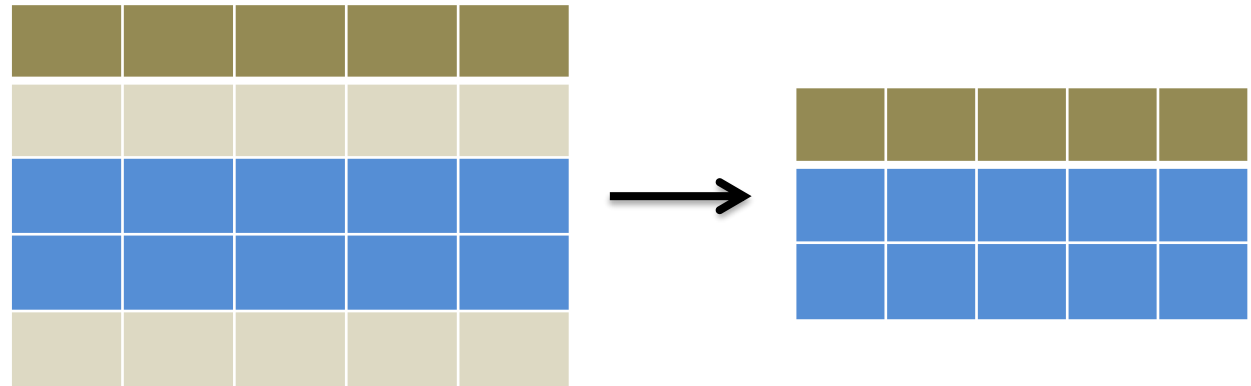


Crea nuevas
variables a partir
de las existentes

```
> mutate(BasedeDatos, Nueva_Variable=Expresión)
```

Librería dplyr: filter

filter()



olive	olive	olive	olive	olive
tan	tan	tan	tan	tan
blue	blue	blue	blue	blue
blue	blue	blue	blue	blue
tan	tan	tan	tan	tan



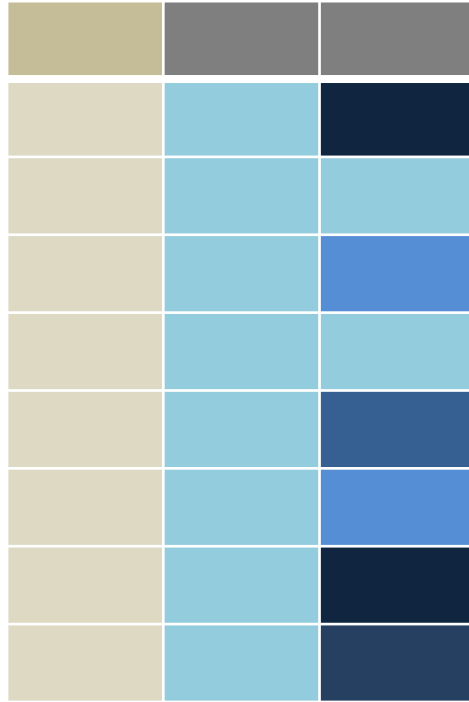
olive	olive	olive	olive	olive
blue	blue	blue	blue	blue
blue	blue	blue	blue	blue

Filtra un subconjunto
de observaciones o
filas

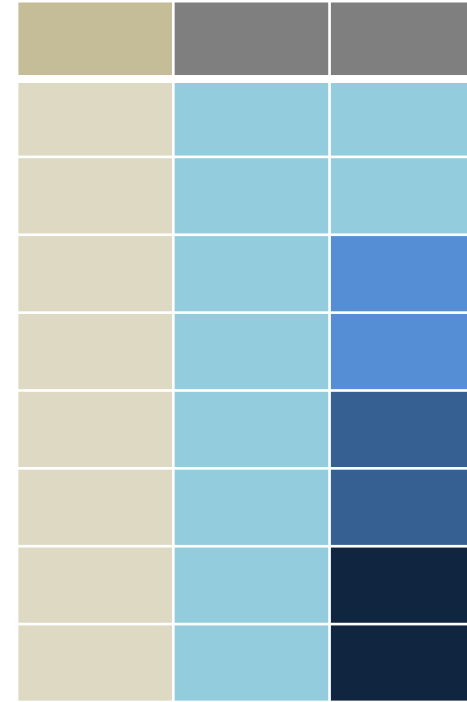
```
> filter(BaseDatos, condición_logica1, condición_logica2,...)
```

Librería dplyr: arrange

arrange()



Light Brown	Grey	Grey
Light Brown	Light Blue	Dark Blue
Light Brown	Light Blue	Light Blue
Light Brown	Light Blue	Blue
Light Brown	Light Blue	Light Blue
Light Brown	Light Blue	Dark Blue
Light Brown	Light Blue	Blue
Light Brown	Light Blue	Dark Blue
Light Brown	Light Blue	Dark Blue



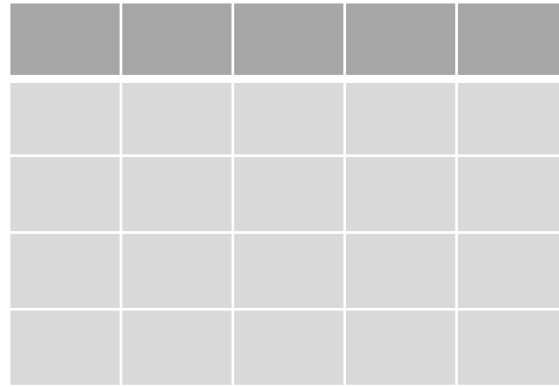
Light Brown	Grey	Grey
Light Brown	Light Blue	Light Blue
Light Brown	Light Blue	Light Blue
Light Brown	Light Blue	Blue
Light Brown	Light Blue	Blue
Light Brown	Light Blue	Dark Blue
Light Brown	Light Blue	Dark Blue
Light Brown	Light Blue	Dark Blue
Light Brown	Light Blue	Dark Blue

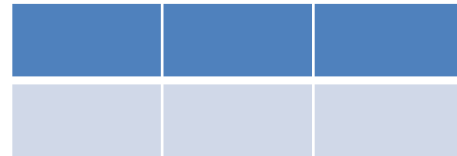
Ordena la base de
datos en función
de una variable

```
> arrange(BaseDatos, Variable)
```

Librería dplyr: summarise

summarise()



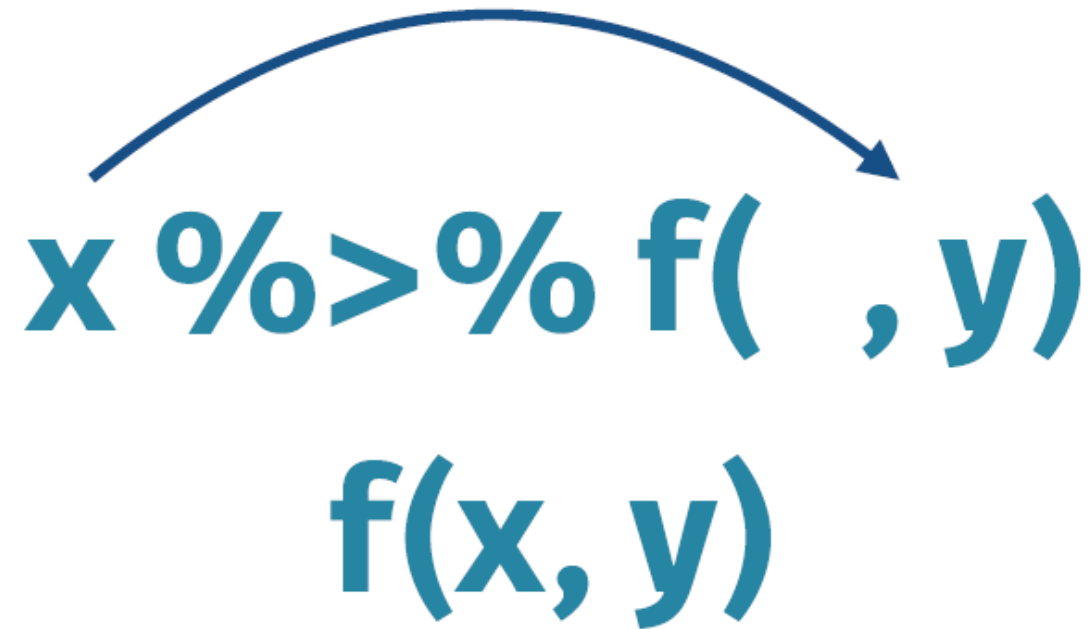


Obtiene las estadísticas en
una sola fila de una variable

```
> summarise(BasedeDatos,media=mean(variable))
```

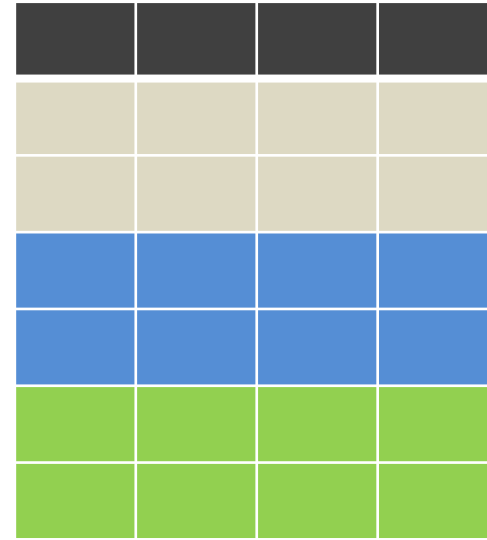
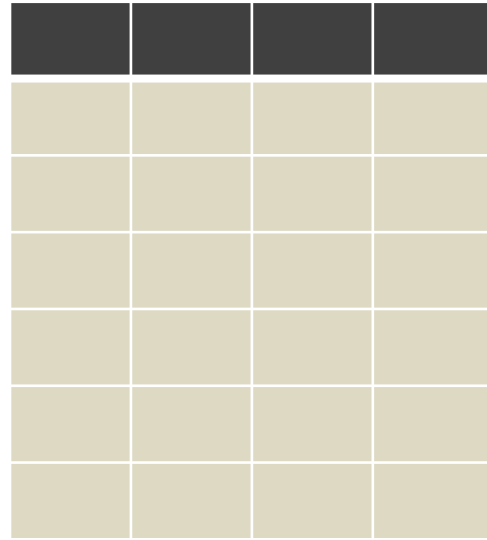
El operador pipe

`%>%`



Librería dplyr: group

`group_by()`



`group_by() %>%
summarise()`

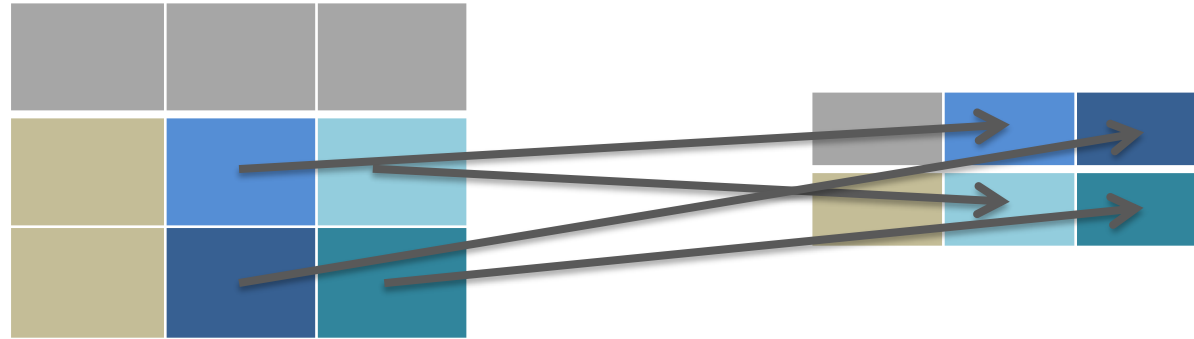
Agrupar variables
para después
obtener sus
estadísticas

Bases de datos: librería tidy

- `pivot_longer ()` → Distribuye columnas en filas
- `pivot_wider ()` → Distribuye filas en columnas
- `separate ()` → Separa una columna en múltiples
- `unite ()` → Une varias columnas en una sola

Librería tidyr: spread

`pivot_wider()`



Las filas se
convierten en
columnas

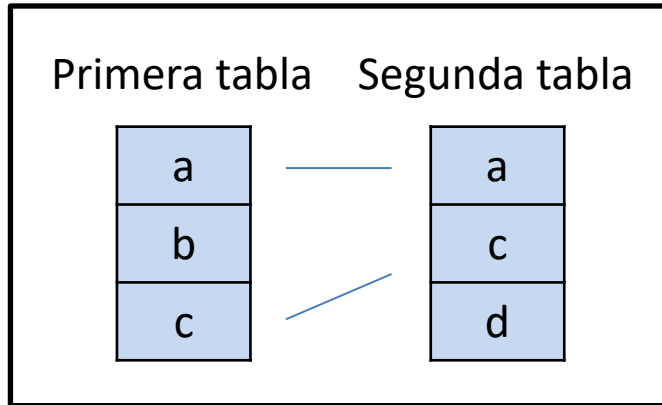
```
> spread(BaseDatos, NombreColumna con Valores Llave, NombreColumna con Valores)
```

Bases de datos: merge

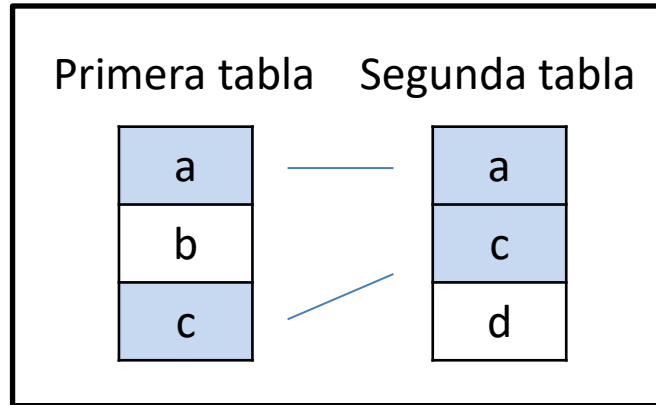
Librería dplyr: verbos de fusión de bases de datos

- `full_join()` \longrightarrow Fusión completa
- `inner_join()` \longrightarrow Fusión parcial
- `left_join()` \longrightarrow Fusión con respecto a la primera base
- `right_join()` \longrightarrow Fusión con respecto a la segunda base
- `anti_join ()` \longrightarrow Elementos que no hacen fusión

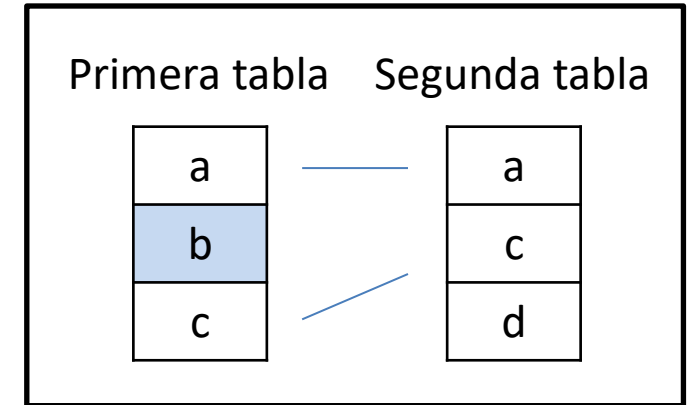
Full join



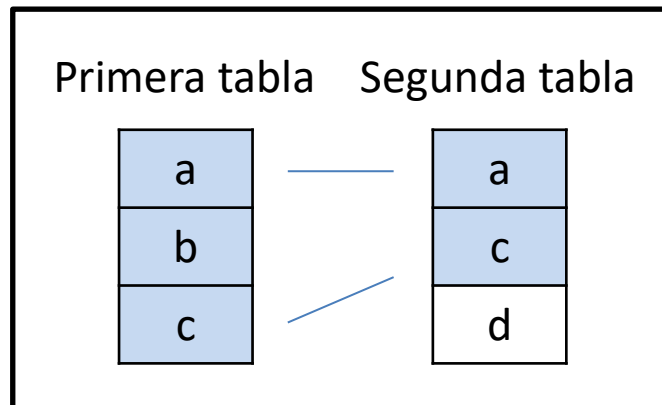
Inner join



Anti join



Left join



Right join

