

# Resumen - Conclusiones

# Tipo de datos

- **Categoricos** (Sexo [Femenino, masculino], Regiones [Este, Oeste, Norte,..])
  - Ordinales por ej.:
    - Número de cajas de cigarros que una persona fuma en el dia:  
0,1,2,etc...
    - Variable picante: fuerte medio y bajo
  - No ordinales
- **Numéricos** (Población, índice de asesinatos, alturas,...)
  - Discretos or ej.:
    - Cantidad de cigarros que una persona fuma en el dia:  
0,1,2,3,...,36
  - Continuos por ej.:
    - Medida de altura de personas

# Librerías y paquetes utilizados

- dslabs: Data Science Labs - Datasets de ejemplo utilizados
- tidyverse:
  - [dplyr](#): Funciones para manipular datos
    - Mutate - Select - Filter - Summarise - arrange
  - [ggplot2](#): Librería para gráficos
  - [tidyr](#): Librería para crear **tidy data**
  - [readr](#): Librería para cargar, parsear. formatear datos csv, tsv, fwf

# Funciones importantes : dplyr :: [cheat sheet](#)

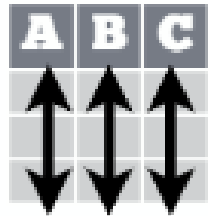
- mutate() Agrega nuevas variables a un data set
- select() Selecciona variables por su nombre
- filter() filtra objetos de acuerdo a su valor
- summarise() crea un resumen de valores importantes
- arrange() reordena filas
- Todas se pueden combinar con group\_by()
- Otras: Split , cut , reduce, identical, unique

# Importar datos: tidyr :: [cheat sheet](#)

- Tidy data:

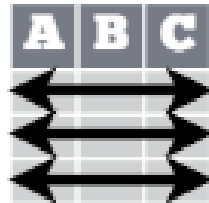
**Tidy data** is a way to organize tabular data. It provides a consistent data structure across packages.

A table is tidy if:



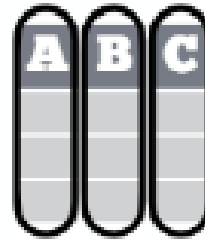
Each **variable** is in its own **column**

&

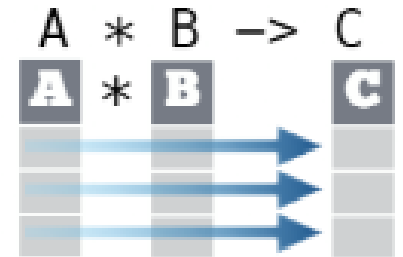


Each **observation**, or **case**, is in its own **row**

Tidy data:



Makes variables easy to access as vectors



Preserves cases during vectorized operations

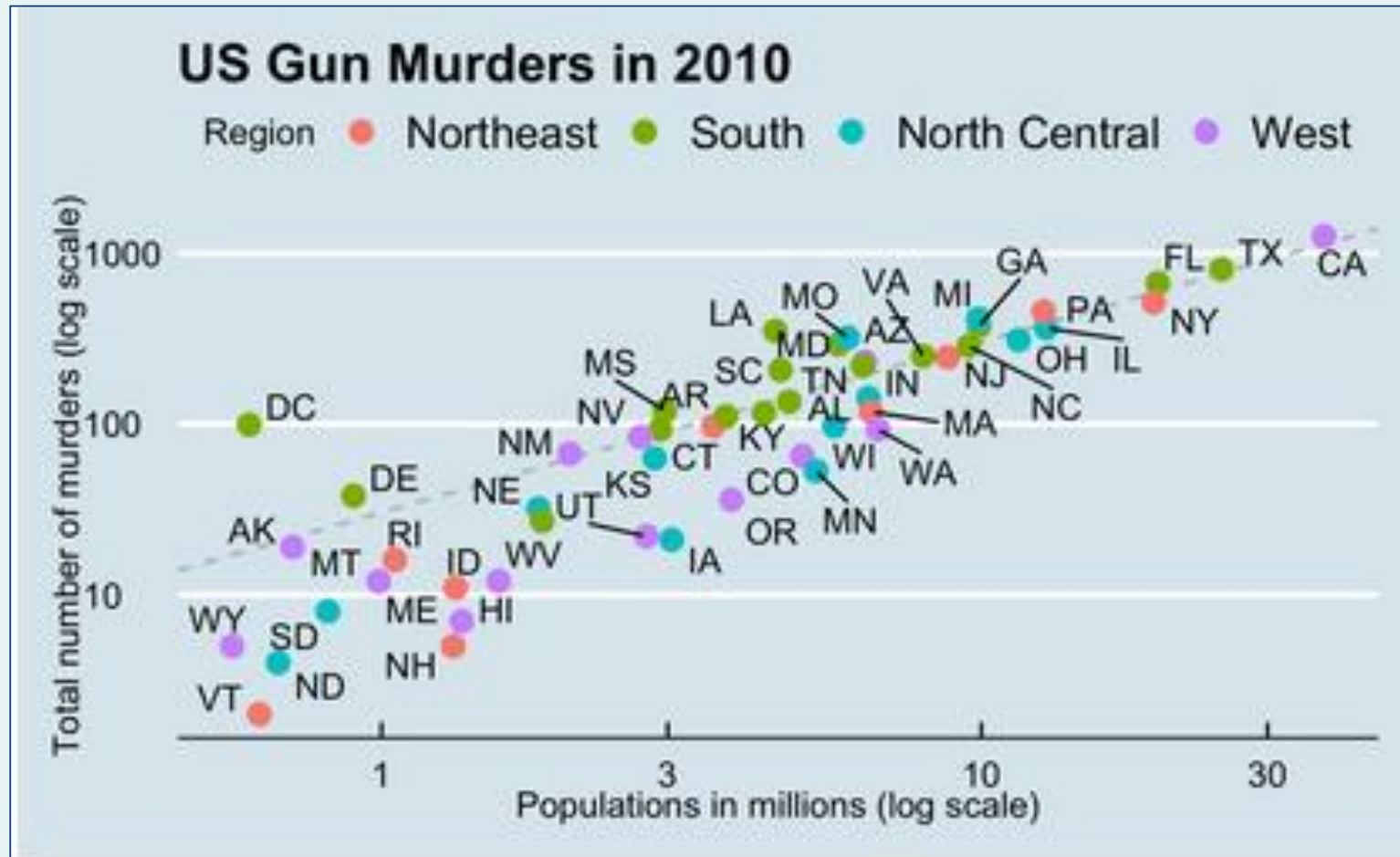
- Cada variable se representa en una columna
- Cada observación o muestra se representa en una fila

¿Qué conclusión puedes sacar de esta tabla de datos?

```
> library(dslabs)
> data(murders)
> head(murders)
```

	state	abb	region	population	total
1	Alabama	AL	South	4779736	135
2	Alaska	AK	West	710231	19
3	Arizona	AZ	West	6392017	232
4	Arkansas	AR	South	2915918	93
5	California	CA	West	37253956	1257
6	Colorado	CO	West	5029196	65

# Una imagen vale más que mil palabras...



# Casos de sarampión en USA

