

Programación con R

Marzo 2019

Edgar Ruiz



@theotheredgar



linkedin.com/in/edgararuiz



github.com/edgararuiz

El
“tidyverse”



¿Que es el “tidyverse”?



Una colección de paquetes de R que son diseñados para la Ciencia de Datos.

Todos los paquetes utilizan la misma filosofía de diseño, gramática y estructuras de datos.

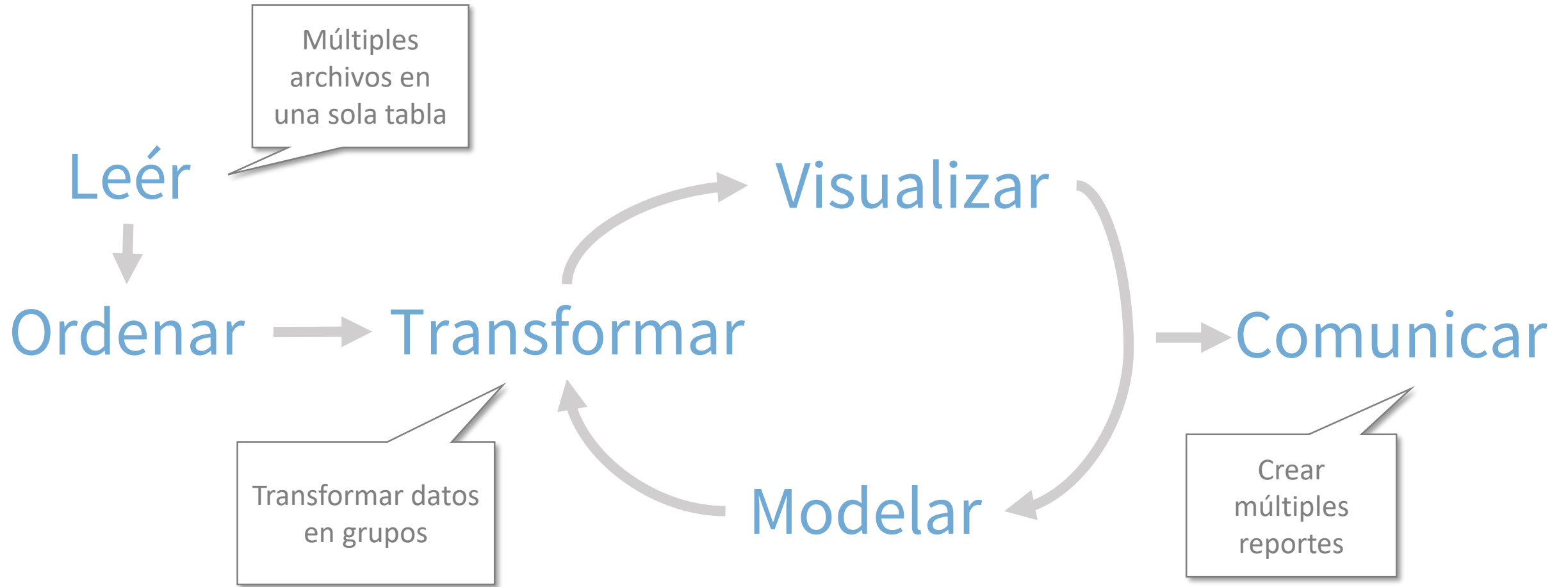
La meta del “tidyverse”



Fóto de [Kelly Sikkema](#) en [Unsplash](#)

Resolver problemas complicados mediante la combinación de diferentes piezas que son consistentes unas con otras

Ciencia de Datos



1. Funciones vectoriales

`x <-`

a	b	c
---	---	---

Vectores con uno o más valores

paste {base}

R Documentation

Concatenate Strings

Description

Concatenate vectors after converting to character.

Usage

```
paste (... , sep = " ", collapse = NULL)
paste0(... , collapse = NULL)
```

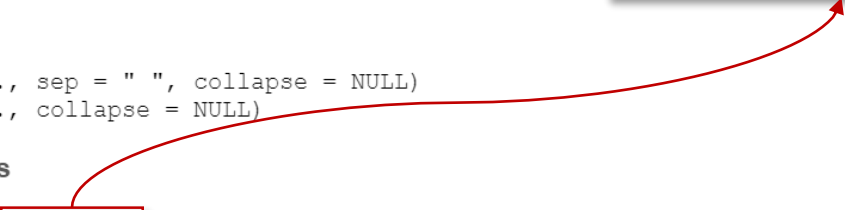
Arguments

... one or more R objects, to be converted to character vectors.

sep a character string to separate the terms. Not [NA_character_](#).

collapse an optional character string to separate the results. Not [NA_character_](#).

uno o más

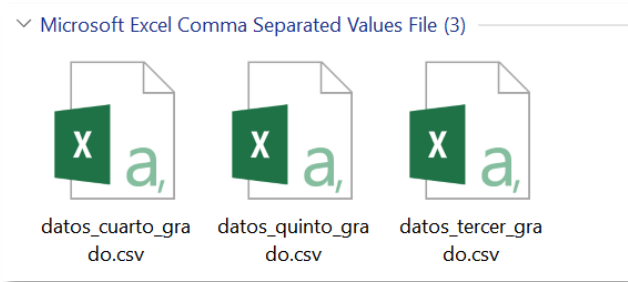


```
x <- c("a", "b", "c")
paste0("letra_", x)
## [1] "letra_a" "letra_b" "letra_c"
```

2. purrr



Leer múltiples archivos



`read_csv(`  `)`
datos_tercer_grado.csv

`read_csv(`  `)`
datos_cuarto_grado.csv

`read_csv(`  `)`
datos_quinto_grado.csv

dia	mes	año	apellido	primer	matricula	origen	grado	materia	puntos
1	1	2015	Hernandez	Rodrigo	100	datos_cuarto_grado.csv	cuarto	ingles	60
1	1	2015	Perez	Roberto	105	datos_cuarto_grado.csv	cuarto	ingles	60
1	1	2015	Ramirez	Alberto	99	datos_cuarto_grado.csv	cuarto	ingles	60
1	3	2015	Perez	Roberto	105	datos_cuarto_grado.csv	cuarto	ingles	62
1	5	2015	Perez	Roberto	105	datos_cuarto_grado.csv	cuarto	ingles	63
1	4	2015	Ramirez	Alberto	99	datos_cuarto_grado.csv	cuarto	ingles	63
1	6	2015	Perez	Roberto	105	datos_cuarto_grado.csv	cuarto	ingles	64
1	3	2015	Ramirez	Alberto	99	datos_cuarto_grado.csv	cuarto	ingles	64
1	2	2015	Ramirez	Alberto	99	datos_cuarto_grado.csv	cuarto	ingles	65
1	2	2015	Perez	Roberto	105	datos_cuarto_grado.csv	cuarto	ingles	66
1	5	2015	Ramirez	Alberto	99	datos_cuarto_grado.csv	cuarto	ingles	66
1	2	2015	Hernandez	Rodrigo	100	datos_cuarto_grado.csv	cuarto	ingles	70
1	4	2015	Perez	Roberto	105	datos_cuarto_grado.csv	cuarto	ingles	70

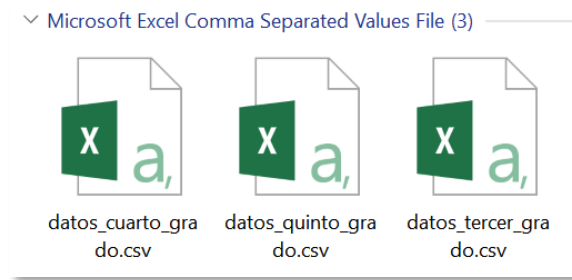
Archivos

Leer cada uno

Combinarlos en una
tabla

Usando purrr

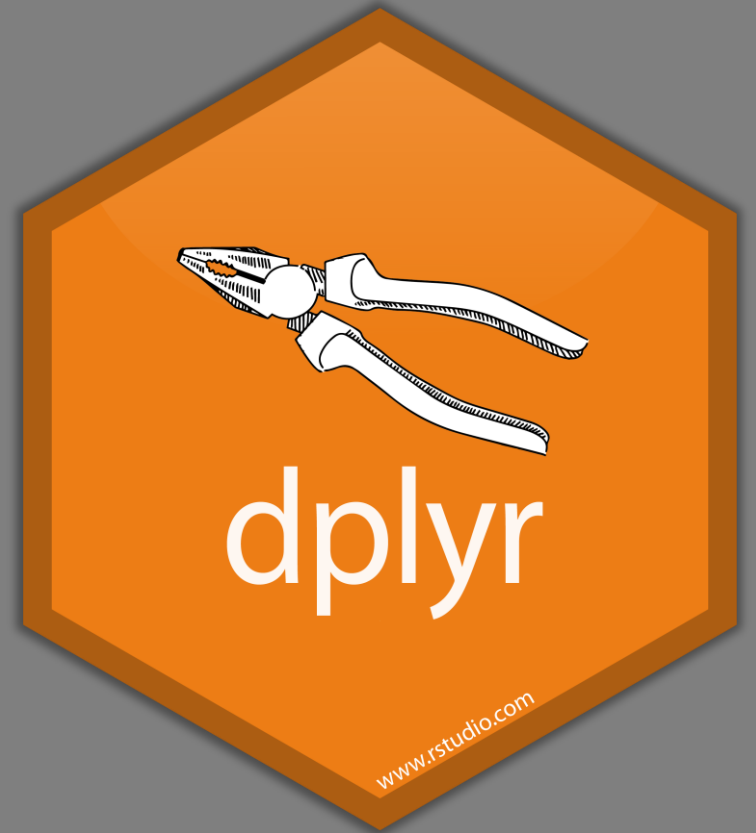
map(



, read_csv)

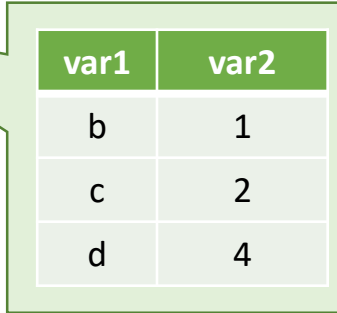
Requiere una conjunto y una función. Siempre devuelve un objeto tipo lista.

3. group_nest()



Más allá que tablas, tibbles

letra	datos
a	<tibble>
b	<tibble>
c	<tibble>



var1	var2
b	1
c	2
d	4

tibble permite crear variables de
tipo tibble y otros tipos

Agrupar datos dentro tibbles

`group_nest(nombre, apellido)`

nombre	apellido	mes	valor
Juan	Uno	Enero	10
Sofía	Dos	Febrero	20
Juan	Uno	Febrero	10
Sofía	Dos	Enero	5



nombre	apellido	data
Juan	Uno	<tibble>
Sofía	Dos	<tibble>

Sumariza la table por cada valor único del grupo,
y pone los detalles dentro tibbles individuales

Operaciones por grupo

```
mutate(m = map(lm(x~y, data = .x)))
```

nombre	apellido	data
Juan	Uno	<tibble>
Sofía	Dos	<tibble>



nombre	apellido	data	m
Juan	Uno	<tibble>	<lm>
Sofía	Dos	<tibble>	<lm>

Después de agrupar, se pueden correr operaciones por cada grupo usando `mutate()` y `map()`

¡Gracias por
su atención!

Foto de [Gary Bendig](#) en [Unsplash](#)

