

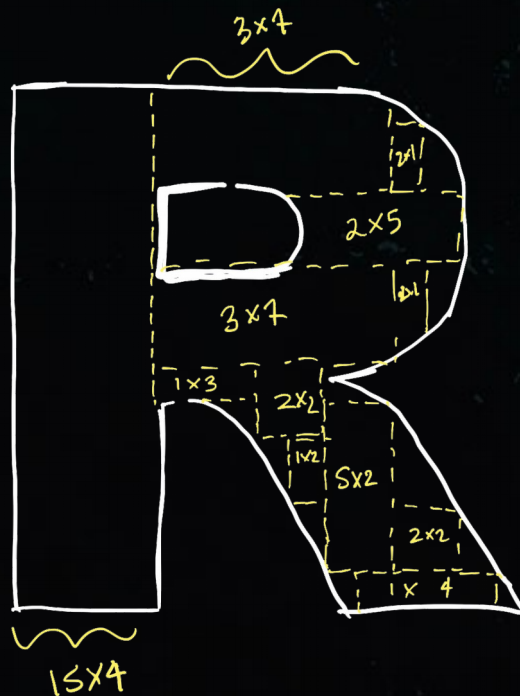
Datos, objetos, funciones

Martín Amodeo

DBBF, UNS

IADO CONICET-UNS

CCT Bahía Blanca



Tipos de datos

- Numérico (*num, double, float*)
- Entero (*int*)
- Texto (*character, string*)
- Lógico (TRUE/FALSE)
- Complejo

Tipos de datos

- Numérico (num, double)

2.23

2.0

0.00002

1.45e-2

Tipos de datos

- Entero (int)

2
10
25

Tipos de datos

- Texto (char, string)

“control”
“hembra”
“campaña dic_17”
“pehuen”
“Nitrógeno 0.5mg”
“area_51\$#%xfiles”
“bill_boton”

Hay más libertad... pero ¡OJO con el encoding!

Recomendación: mantenerlo simple respetando las reglas básicas

Priorizar facilidad para tipear y minimizar posibles errores

Tipos de datos

- Lógico (logical)

TRUE
FALSE



Objetos



Tipos de objetos

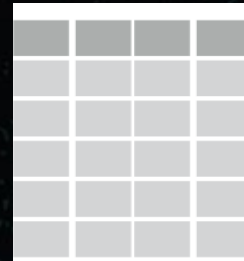
- Valor



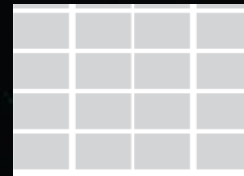
- Vector



- Data frame



- Matriz



- Lista



Tipos de objetos

- Valor

2.23
0.00002
"bill"
"cuatreros topsecret3"
TRUE

Tipos de objetos

- Vector

1.2 0.25 8.3 2.23
"bill" "wysiwyg" "boton"
"cuatreros" "ing white"
TRUE FALSE TRUE TRUE

Vector numérico de 4

Vector de texto de 3

Vector de texto de 2

Vector lógico de 4

Serie de valores

Tipos de objetos

- Vector

c()

Serie de valores CONCATENADOS

Tipos de objetos

- Vector
- `c(1.2, 0.25, 8.3, 2.23)`
- `c("bill", "wysiwyg", "boton")`
- `c("cuatreros", "ing white")`
- `c(TRUE, FALSE, TRUE, TRUE)`

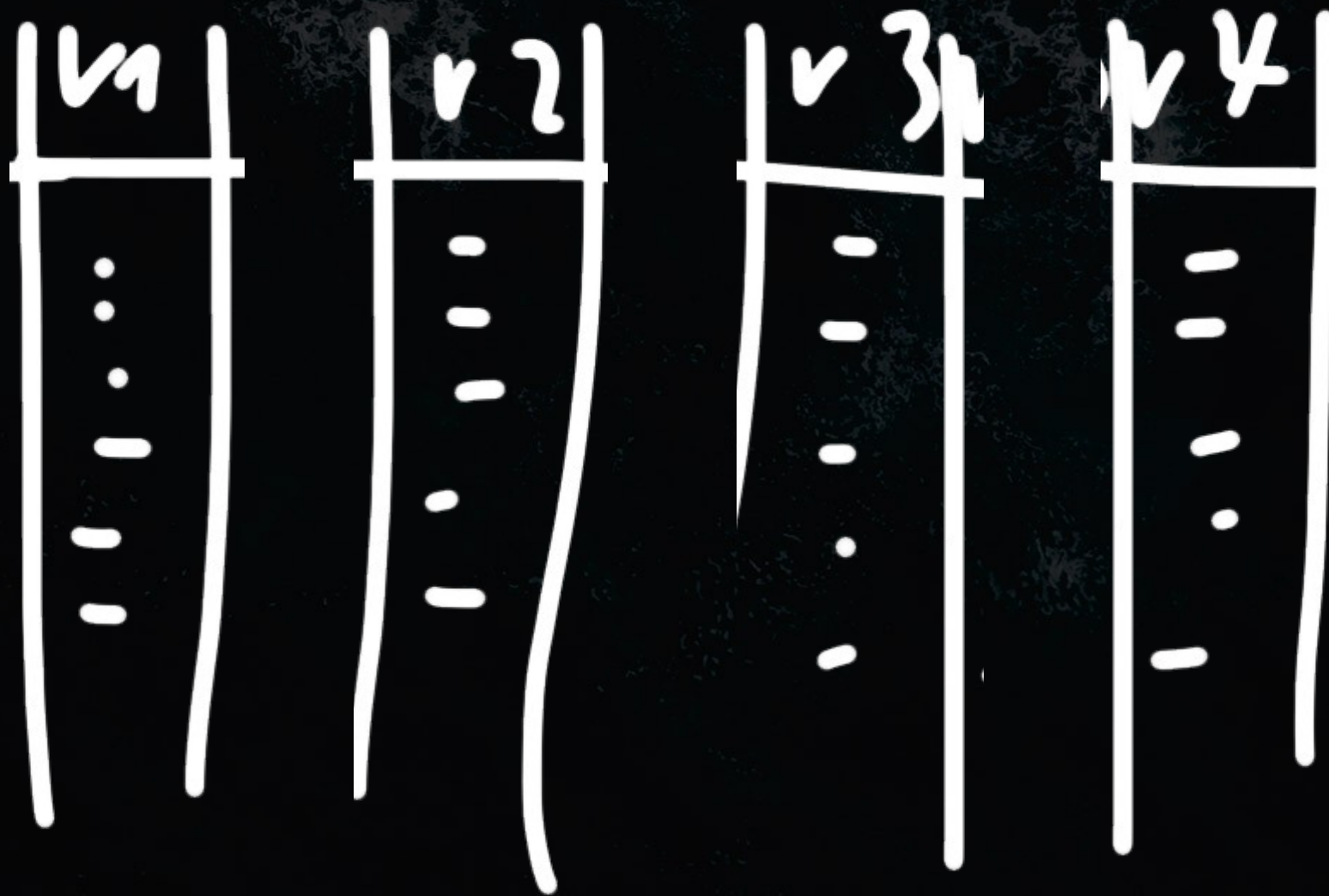
`c()` y separados por comas

Tipos de objetos

- Data frame

v_1	v_2	v_3	v_4
:	-	-	-
:	-	-	-
:	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-

Varios vectores del mismo largo alineados



Vectores del mismo largo
alineados



Variables

Observaciones

\$

datos\$variable1
datos\$variable2
datos\$variable3

Tipos de variables

- Variable numérica (num)

1.2 0.25 8.3 2.23

Vector numérico

temperatura, clorofila, ph, altura, concentración, salinidad, velocidad

Tipos de variables

- Variable de conteo (int)

1 0 8 2 3 5 22 10

Vector de enteros

Cantidad de cangrejos por cueva,
Cantidad de conos por árbol,
Número de frutos
Cantidad de fibras plásticas por muestra
Número de aves vistas por recorrido

Tipos de variables

- Variable categórica (Factor)

"bill" "wysiwyg" "boton"

"cuatreros" "ing white"

Sitios

Temporada

Estación

Campaña

Operador/a

Tratamiento

Vector de texto convertido en factor

Factor

Factor

- Variable categórica (Factor)

"cuatreros"	1
"ing white"	2
"rosales"	3
"cuatreros"	1
"cuatreros"	1
"rosales"	3
"ing white"	2
"ing white"	2

Tres niveles: "cuatreros", "ing white", "rosales"

`as.factor()`



Variables

Observaciones

[illegible]

Planilla interpretable

Simple, plana, completa

- Respetar Reglas Básicas para Nombres
- Unificar rótulos
("cuatr", "cuatr.", "cuatrerros", "cuatrero", "Cuatrerros")
- Chequear errores de tipeo y mayúsculas/minúsculas
- Sin columnas ni filas vacías
- Sin medidas de resumen, tablitas, grafiquitos....
- SOLO DATOS CRUDOS
- Definir los datos faltantes como NA

Planilla interpretable

NA

Dato faltante

- Se perdió la muestra,
- Se perdió el número
- No se hizo ese punto
- El equipo tiró error

Cero

- Hay dato y el valor es CERO

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
MAY-14	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
	MAY-14																0.00	0.00	0.00
MAY-15	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
	MAY-15																0.00	0.00	0.00
MAY-16	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
	MAY-16																0.00	0.00	0.00
MAY-17	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
	MAY-17																0.00	0.00	0.00
MAY-18	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00
	MAY-18																0.00	0.00	0.00

Planilla NO interpretable

Celdas vacías, caracteres especiales

	A	B	C	D	E	F
1	fecha	n.obs	vel.max	vel.media	dir	
2	2013-06-30 10:50:00	612	45.6	27.7	NW	
3	2013-07-07 10:50:00	608	33.3	19.1	NW	
4	2013-08-08 10:50:00	514	NA	30.5	NW	
5	2013-08-17 10:50:00	508	NA	33.7	NW	
6	2013-09-25 10:50:00	567	50.6	3.3	NW	
7	2013-10-20 10:50:00	539	66.2	21.9	NW	
8	2013-11-21 10:50:00	430	93.1	38.4	NA	
9	2013-12-07 10:50:00	584	52.4	24.7	NA	
10	2013-12-23 10:50:00	588	53.5	18.1	NW	
11	2014-04-21 10:50:00	539	26.8	14.4	NW	
12	2014-06-17 10:50:00	454	NA	NA	NA	
13	2014-06-24 10:50:00	394	NA	NA	NA	
14	2014-07-10 10:50:00	654	NA	NA	NA	
15	2014-07-26 10:50:00	568	NA	NA	NA	
16	2014-08-04 10:50:00	641	42.4	32.9	NW	
17	2014-08-20 10:50:00	125	54.4	16.6	NW	
18	2014-08-27 10:50:00	109	16.5	7.4	NW	
19	2014-09-12 10:50:00	72	49.5	17.5	NW	
20	2014-11-15 10:50:00	554	41.3	7.4	NW	
21	2014-11-24 10:50:00	206	45.4	NA	NA	
22	2014-12-01 10:50:00	622	40.5	16.6	NW	
23						
24						
25						
26						
27						

Planilla interpretable

Simple, plana, completa, con NA