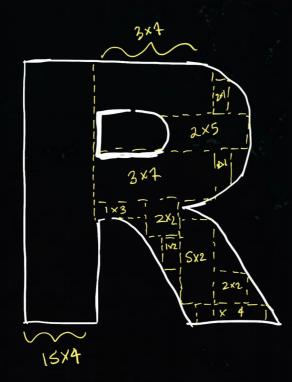
#### Datos, objetos, funciones

Pepi Amodeo Instituto Argentino de Oceanografía CONICET-UNS CCT Bahía Blanca







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

• Numérico (num, double)

2.23 2.0 0.00002 1.45e-2

• Entero (int)

2 10 25

• Texto (char, string)

"control"
"hembra"
"campaña dic\_17"
"pehuen"
"Nitrógeno 0.5mg"
"area\_51\$#%xfiles"
"bill\_boton"

Hay más libertad... pero iOJO con el encoding! Recomendación: mantenerlo simple respetando las reglas básicas Priorizar facilidad para tipear y minimizar posibles errores

• Lógico (logical)

TRUE
FALSE



# Objetos



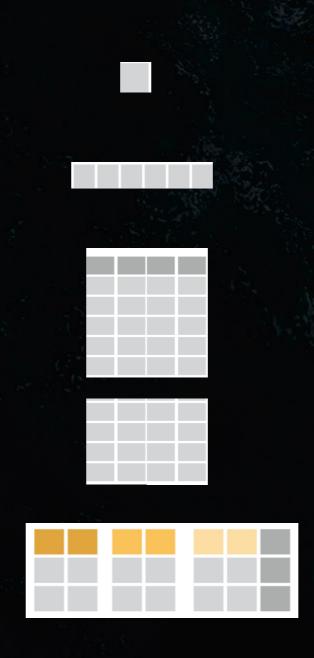
Valor

Vector

• Data frame

Matriz

• Lista



Valor

2.23 0.00002 "bill" "cuatreros topsecret3" TRUE

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

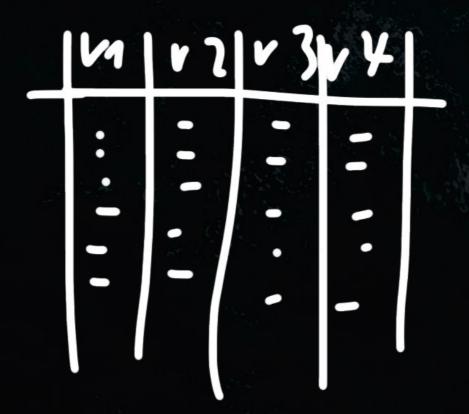
Vector lógico de 4

Vector

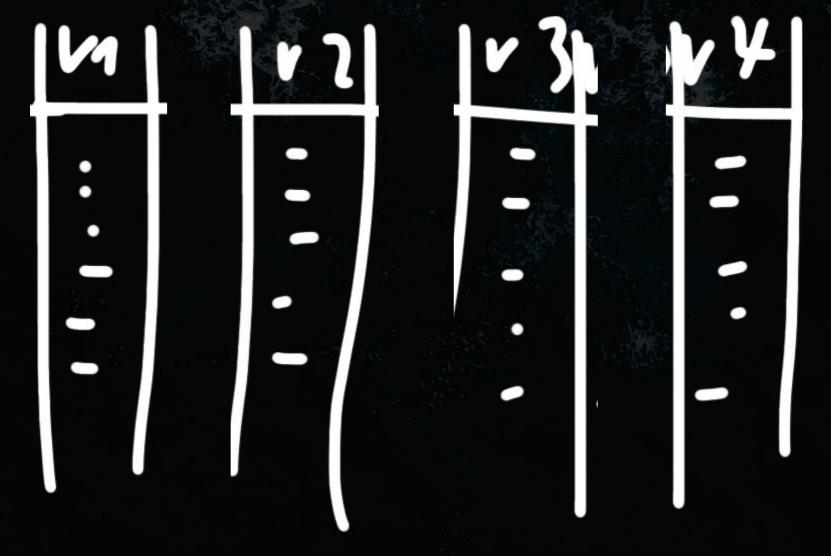
**c**()

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

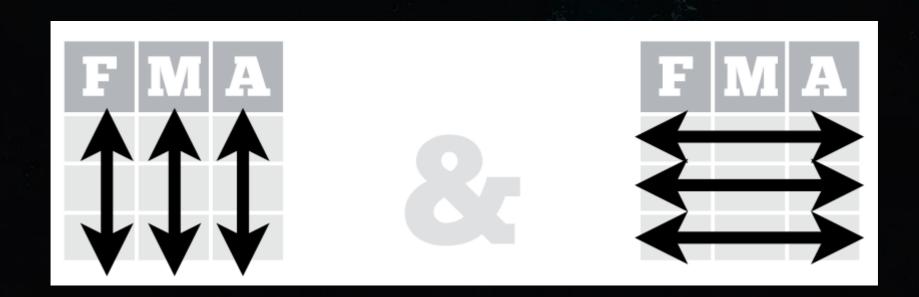
• Data frame



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)

1.20.258.32.23

Vector numérico

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

### Tipos de variables

• Variable de conteo (int)

1082352210

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

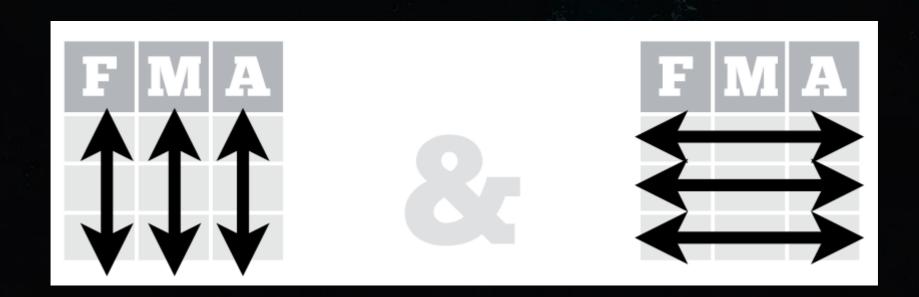
#### Factor

• Variable categórica (Factor)

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

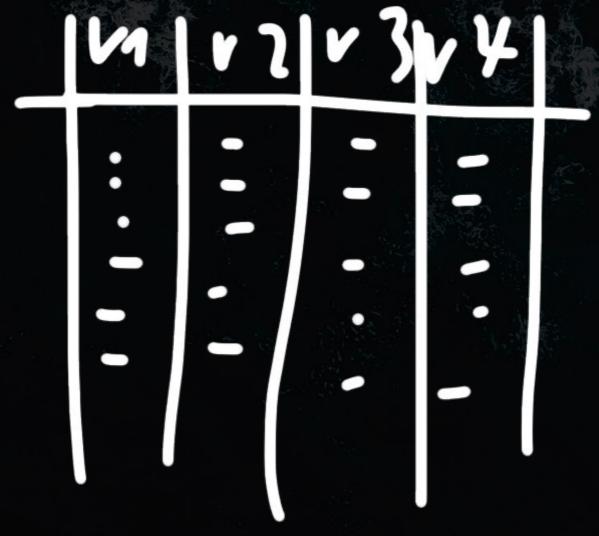
Tres niveles: "cuatreros", "ing white", "rosales"

as.factor()



Variables

Observaciones



Planilla interpretable

Simple, plana, completa

- Respetar Reglas Básicas para Nombres
- Unificar rótulos
   ("cuatr", "cuatre," "cuatreros", "cuatrero", "Cuatreros")
- 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

March	Englis Process  COL	201 000 148 000 000 477 000 000 000 477 000 000 000 000 000 000 000 000 000 0	200.51 Par himital 1 100 h 100
6 Miras 10 (10 Gar) 24 19524 19524 1131 27845 CM CM CM	CCC CCC CCC CCC CCC CCC CCC CCC CCC CC	CC2   CC5   148   CC5   CC5	0.10 0.10 79.45 5.1 0.11 17.81 2.5 0.21 17.81 2.5
6 Miras 10 (10 Gar) 24 19524 19524 1131 27845 CM CM CM	CCC CCC CCC CCC CCC CCC CCC CCC CCC CC	CC2   CC5   148   CC5   CC5	C75 5118 27
	0.12 0.12 0.12 0.12 0.12 0.12 0.12 0.12	0.15 0.01 12.45 0.05 0.00 4.59 1.371 0.00 97.46 0.01 0.00 2.05	0.76 68.18 8.2
	0.12 0.13 0.64 0.66 0.05 0.05 0.05 0.05 1.75 1.75	2372 8488 9746 002 000 207 002 000 206	0.76 68.18 8.2
1	005 005 148 148	002 000 206	
1	005 005 148 148	002 000 206	
	1.18 1.18 0.17 0.17		
Name			POC-ST SPM INTO EL-12
March-14   Di-Scham   Sept.			136 113.06 10
1		0.081 0.007 6.88	POC (WA 6 MC SPM (NO. 1) 0.18 1055 1.5
		0.125 0.010 10.45	
	2.72 (6 0.19 0.19 0.07 0.07	0.091 0.008 7.61	0.15 12.68 1.3
10   10   10   10   10   10   10   10	3.59	0.091 0.009 7.61 0.015 0.004 7.97 0.757 0.064 67.70	
	0.18 0.18	0.156 0.018 12.96 0.050 0.005 5.08	0.25 2094 0.8
Part	131 66	0.620 0.052 51.64	
Part			POC-51 PM hones-1
April   Sec.   April   Sec.   April			045 MAC 12036 1993
No.   175 From increases   116   125   1	056 056	0.255 0.022 22.13	POC (W/L 6 MC SPM (NO.1)
Note   10   10   10   10   10   10   10   1	035 035 3.93 66	0.164 0.014 18.60 1.861 0.155 155.05	555 5557 75
3.8 http://doi.org/10.000/04.00000000000000000000000000000	0,00 00,00	0,00	067 5556 79
\$000 (60-200um cumma/outposes) 92 196-00 196-00 196-00 2604-00 2004-00 2004-00 2004-00-00-00-00-00-00-00-00-00-00-00-00-	047 047 251 66	0174 0014 1446 1185 0099 9874	
2	0.18 0.18	0.181 0.011 10.90	D56 55.DE 8.4
100 21859 21859 1494 CD 14148 34 145 14149	122 66	0.000 6430	1
30			
46			
36			
98			
32			
37			
Manuscript Committee Change of the Change of	0.30		9 dicenter
5 0 10 5 0 10 10 10 10 10 10 10 10 10 10 10 10 1	236		
50 000 0.16 5-27 0.25 29-28 5-83 44-56 23-07 5-97 5-97 5-97 5-97 5-97 5-97 5-97 5-9	297 270		
500 Sept. C14 5.65 0.00 0.00 1935 107 1107 1107 1107 1107 1107 1107 1107	156		
\$2 tax 1 72/2000(51)			
87 No. 10	Parformation 4.61		
86	4.61		N domber
200 CCC 114 CZ 1781 Z1 18C 21	0.00		
58   Dennic 0.07   Z65   0.07   0.07   14.72	154		
77	0.00 154 2.66 154		

# Planilla NO interpretable

Celdas vacías, caracteres especiales

	STATE OF THE STATE	A difficult		(8)	147	
	Α	В	С	D	E	F
1	fecha	n.obs	vel.max	v el.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						
0.7						

# Planilla interpretable

Simple, plana, completa, con NA