



Centro Internacional de Agricultura Tropical  
Desde 1967 *Ciencia para cultivar el cambio*

# Alineación de bases de datos de clima, suelo y manejo agronómico con eventos de cosecha.

Julio 2018

## Autores

Hugo Andrés Dorado B.

Juan Camilo Rivera

[h.a.dorado@cgiar.org](mailto:h.a.dorado@cgiar.org) , [j.c.rivera@cgiar.org](mailto:j.c.rivera@cgiar.org)



El CIAT es un Centro de Investigación de CGIAR

# Obtención de una base de datos estructurada

	A	B	C	D	E
1	ID	Sowing_Date	Harvest_Date	Variety	Yield
2	RC61_2008_989	2008-03-07	2008-07-05	ACARIGUA	6700
3	RC62_2010_207	2010-07-22	2010-11-25	ACD 2526	9125
4	RC62_2011_275	2011-03-11	2011-07-15	ACD 2526	6375
5	RC62_2012_361	2011-09-08	2012-01-12	ACD 2526	6875
6	RC62_2011_303	2011-04-25	2011-08-29	ACD 2528	7500
7	RC62_2011_213	2010-08-30	2011-01-03	ACD 2540	6563
8	RC62_2011_274	2011-03-09	2011-07-13	caracoli	6250
9	RC62_2010_76	2009-12-19	2010-04-24	CHICALA	5600
10	RC62_2011_336	2011-08-06	2011-12-10	CHICALA	4625
11	RC62_2011_345	2011-08-22	2011-12-26	CHICALA	4687
12	RC62_2011_348	2011-08-23	2011-12-27	CHICALA	5163
13	RC62_2012_372	2011-09-14	2012-01-18	CHICALA	6875
14	ENA_2007a_106386	2007-02-21	2007-07-01	CIMARRON BARINAS	6937.5
15	ENA_2007a_100234	2007-03-21	2007-07-25	CIMARRON BARINAS	7500
16	ENA_2007a_102633	2007-04-14	2007-09-25	CIMARRON BARINAS	8187.5
17	ENA_2007a_101504	2007-05-14	2007-10-09	CIMARRON BARINAS	8000
18	ENA_2007a_100400	2007-05-26	2007-10-06	CIMARRON BARINAS	5187.5
19	ENA_2007a_100150	2007-05-26	2007-10-13	CIMARRON BARINAS	7812.5
20	ENA_2008a_101504	2008-03-01	2008-07-02	CIMARRON BARINAS	6562.5
21	ENA_2008a_100234	2008-04-28	2008-08-08	CIMARRON BARINAS	7000

## Esfuerzos

- Alinear en un único formato bases de datos aisladas que pueden estar incluso con otras estructuras.
- Resumir bases de datos, con observaciones repetidas (manejo agronómico).
- Unir bases de datos utilizando un identificador
- Procesar datos climáticos diarios en indicadores de cultivo.

# Contenido

- Funciones para re-estructurar bases de datos del paquete tidyr.
- Función **ddply** para sumarización de base de datos.
- Función **merge** para unir base de datos.
- Funciones script para generar indicadores climáticos.

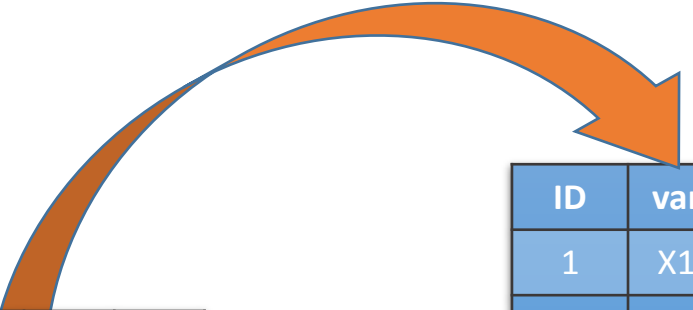
# Funcion gather

```
ID <- c(1,2,3)
x1 <- c(3,2,3)
x2 <- c(5,3,1)
x3 <- c(7,5,3)
```

```
data <- data.frame(id,x1,x2,x3)
```

```
data_gather <- gather(data,var,value,-id)
```

```
data_gather
```



ID	x1	X2	X3
1	3	5	7
2	2	3	5
3	3	1	3

Formato ancho

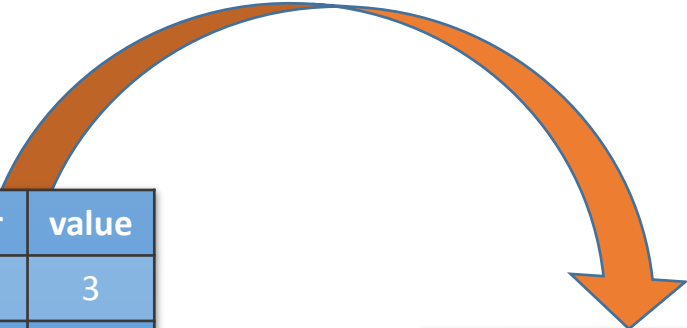
ID	var	value
1	X1	3
1	X1	2
1	X1	3
2	X2	5
2	X2	3
2	X2	1
3	X3	7
3	X3	5
3	X3	3

Formato Largo

Source: <https://diarium.usal.es/jaortega/2016/02/06/introduccion-a-tidyr-datos-ordenados-en-r/>

# Función spread

`spread(data_gather, var, value )`



ID	var	value
1	X1	3
1	X1	2
1	X1	3
2	X2	5
2	X2	3
2	X2	1
3	X3	7
3	X3	5
3	X3	3

Formato Largo

ID	x1	X2	X3
1	3	5	7
2	2	3	5
3	3	1	3

Formato ancho



# Función separate

```
> data_gather$pais_ciudad <- paste('Honduras',c(rep('copan',4),rep('tegucicalpa',5)),sep='_')
> data_gather
  id var value      pais_ciudad
1  1  x1     3 Honduras_copan
2  2  x1     2 Honduras_copan
3  3  x1     3 Honduras_copan
4  1  x2     5 Honduras_copan
5  2  x2     3 Honduras_tegucicalpa
6  3  x2     1 Honduras_tegucicalpa
7  1  x3     7 Honduras_tegucicalpa
8  2  x3     5 Honduras_tegucicalpa
9  3  x3     3 Honduras_tegucicalpa
> separate(data_gather,pais_ciudad,c('pais','ciudad'),sep = '_')
  id var value      pais      ciudad
1  1  x1     3 Honduras      copan
2  2  x1     2 Honduras      copan
3  3  x1     3 Honduras      copan
4  1  x2     5 Honduras      copan
5  2  x2     3 Honduras tegucicalpa
6  3  x2     1 Honduras tegucicalpa
7  1  x3     7 Honduras tegucicalpa
8  2  x3     5 Honduras tegucicalpa
9  3  x3     3 Honduras tegucicalpa
```

# Funciones de summarización ddply

```
library(plyr)

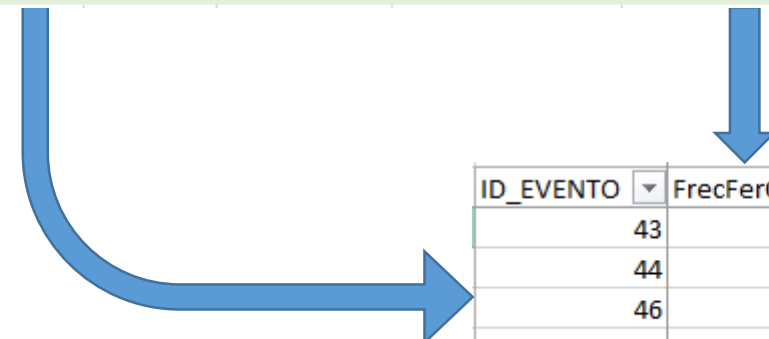
id_evento      <- c(43,43,44,44,44,46,46,46)
cantidad_prod  <- c(300,225,300,250,100,300,234,550)

data <- data.frame(id_evento,cantidad_prod)

data

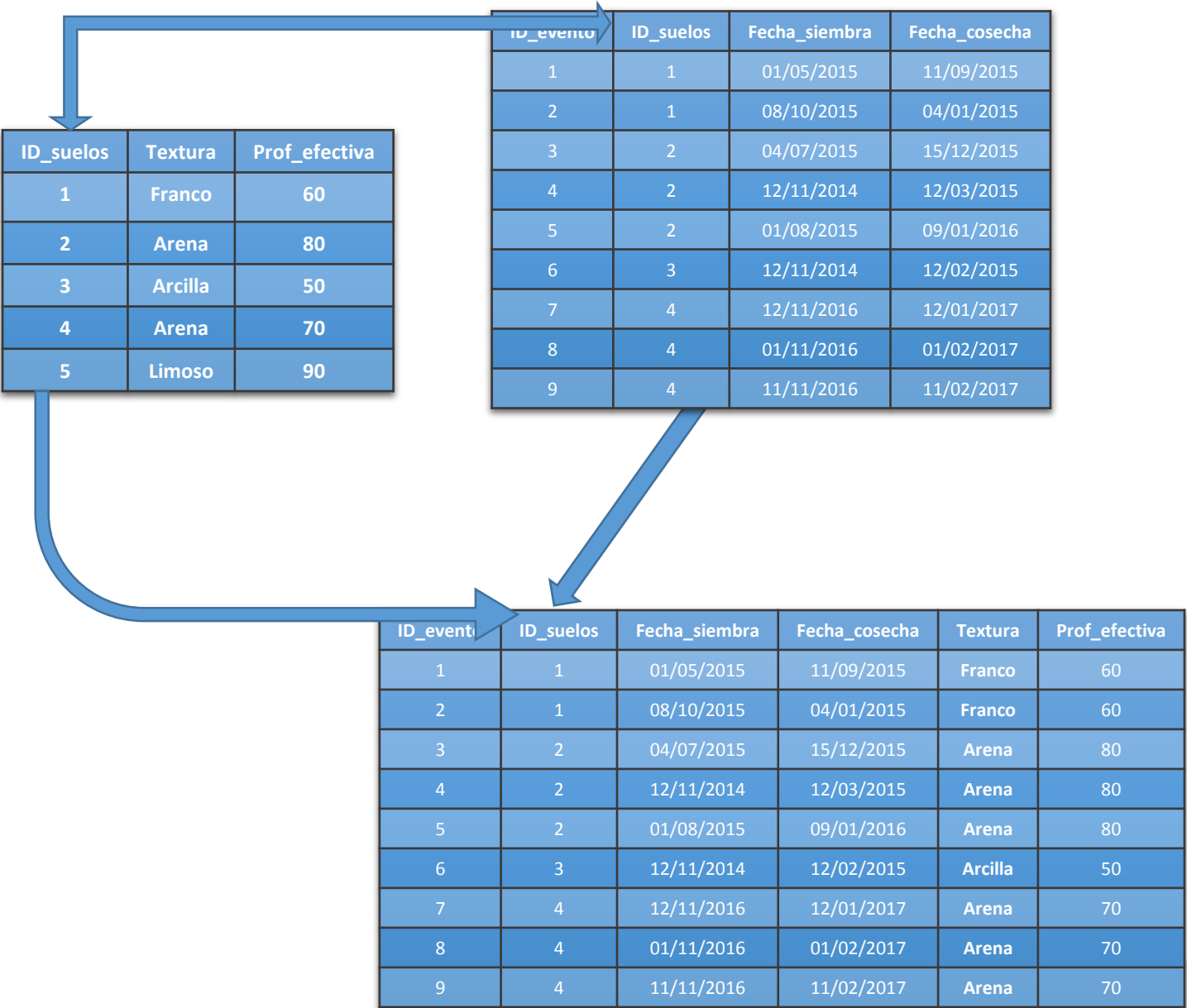
ddply(data,
      ~id_evento,
      summarise,
      acmProd = sum(cantidad_prod),
      freqProd=length(cantidad_prod),
      avg = length(cantidad_prod))
```

ID_EVENTO	ID_PROD	FECHA_FERTI	TIPO_PROD_FERTI	CANTIDAD_PROD_FERTI
43	52	4/13/2013	Quimica	300
43	52	5/15/2013	Quimica	225
44	54	4/25/2013	Quimica	300
44	54	5/25/2013	Quimica	250
44	54	5/25/2013	Quimica	100
46	55	3/27/2013	Quimica	300
46	55	4/26/2013	Quimica	234
46	55	4/26/2013	Quimica	550



ID_EVENTO	FrecFerQu	TotFerQuir
43	2	525
44	3	650
46	3	1084
53	1	100

# Merge





# Merge

```
id_evento <- 1:9
id_suelos <- c(1,1,2,2,2,3,4,4,4)

fecha_siembra <- c("1/5/2015", "8/10/2015", "4/7/2015", "12/11/2014", "1/8/2015",
                  "12/11/2014", "12/11/2016", "1/11/2016", "11/11/2016")

fecha_cosecha <-
c("11/9/2015", "4/1/2015", "15/12/2015", "12/3/2015", "9/1/2016",
  "12/2/2015", "12/1/2017", "1/2/2017", "11/2/2017")

enventos_cosecha <-
data.frame(id_evento, id_suelos, fecha_siembra, fecha_cosecha)

id_suelos <- 1:5
textura <- c("Franco", "Arena", "Arcilla", "Arena", "Limoso")
prof_efectiva <- c(60, 80, 50, 70, 90)

suelo <- data.frame(id_suelos, textura, prof_efectiva)

merge(enventos_cosecha, suelo, by.x='id_suelos', by.y='id_suelos', all.x=T, all.y=
F, sort=T)
```

# Indicadores climáticos

Información diaria

	A	B	C	D	E	F
1	DATE	ESOL	RAIN	RHUM	TMAX	TMIM
.557	4/5/2009	412.8747	0	70.99139	36	24.3016
.558	4/6/2009	513.9043	0	75.20833	34.8	24.9
.559	4/7/2009	396.5338	0	73.85714	34.1	25.6
.560	4/8/2009	397.8491	0	74.09524	33.9	25.4
.561	4/9/2009	448.4498	0	76.82609	34.6	24.9
.562	4/10/2009	481.8188	0	66.20671	39	24.8
.563	4/11/2009	448.1053	0	73.66386	35.9	25.4

Indicadores climáticos (Acumulados, promedios, frecuencia, máximo, mínimo)

ID	FECHA_SIEMBRA	FECHA_COSECHA	ANO_COS	RENDIMIENTO_HA	TMAXavg	TMINavg	TEMPavg	GDaccu11	RANGO_Diurno_avg	Eneraccu
RC38_2009_5	4/5/2009	8/3/2009	2009	5600	33.11977441	23.67722572	28.39850006	1957.651791	9.442548692	43981.57
RC38_2009_6	4/5/2009	8/3/2009	2009	5775	33.11977441	23.67722572	28.39850006	1957.651791	9.442548692	43981.57
RC38_2009_7	4/5/2009	8/3/2009	2009	4200	33.11977441	23.67722572	28.39850006	1957.651791	9.442548692	43981.57
RC27_2013_3037	10/22/2012	2/19/2013	2013	5262	34.06942149	24.20578512	29.13760331	1880.2564	9.863636364	43883.31
RC38_2013_129	10/22/2012	2/19/2013	2013	5265	34.06942149	24.20578512	29.13760331	1880.2564	9.863636364	43883.31
RC38_2013_130	10/24/2012	2/21/2013	2013	5284	34.16363636	24.2107438	29.18719008	1873.7553	9.952892562	43962.81
RC38_2013_134	11/2/2012	3/2/2013	2013	6720	34.30661157	24.30743802	29.30702479	1862.2728	9.999173554	44100.78

Información por evento de cultivo

<https://github.com/hdorado/Indicadores-climaticos>

# ¡Gracias!



NOS ENORGULLECE  
HABER CELEBRADO 50 AÑOS  
DE INVESTIGACIÓN AGRÍCOLA  
PARA EL DESARROLLO

## Centro Internacional de Agricultura Tropical - CIAT

Sede Principal y Oficina Regional  
para Suramérica y el Caribe

+57 2 445 0000

Km 17 Recta Cali-Palmira  
A.A. 6713, Cali, Colombia

✉ [ciat@cgiar.org](mailto:ciat@cgiar.org)

🌐 [ciat.cgiar.org](http://ciat.cgiar.org)



El CIAT es un Centro de Investigación de CGIAR