

29-oct-2013

Cálculo de series de tiempo para clientes agrupados por tipo para la zona norte de pepsi y 3 productos

Makefile: pandoc -s -V geometry:margin=0.7in -V lang=spanish 2_seriesdetiempo.md -o Series.pdf

Después de filtrar en el exploratorio los sku's colados, en este archivo ahora se trata de producir un archivo con las ventas y devoluciones promedio de cada tipo de cliente y de cada fecha. Después, con estos promedios, calcularemos inventarios.

Para poder dividir a los clientes por tipo, tenemos en la tabla de cliente info de eso: misceláneas, hospitales, escuelas, etc. La llave para unir el catálogo de clientes con la info de ventas es un Id de clientes que consiste en ID_SUC_CORP y ID_CLI_CORP concatenados.

De este archivo se generan fritos3, papas3 y churruvais3, que son tablas con los valores de ventas y devoluciones en unidades y pesos para cada tipo de cliente y cada día.

NOTA IMPORTANTE: Para calcular los inventarios de los clientes (tienditas) y sus ventas a los consumidores (gente que compra papas en las tienditas), estábamos tomando valor absoluto de ventas y devoluciones porque hay valores negativos. Pero volvimos a hacer esto porque resulta que ventas negativas = devoluciones que no llegaron hasta el centro de distribución, sino que se recolocaron en otra tienda. No nos importa distinguir entre papas devueltas hasta el centro de distribución y las que sólo son vendidas en otra tienda porque al final son devoluciones.

En el archivo de inventario ya se generan grafiquitas con las series de tiempo de cada tipo de cliente para cada producto.

Churruvais

```
setwd("/Users/PandoraMac/Documents/David/pepsi2/Datos Norte 3 papitas/")
clientes<-as.data.frame(read.table("SF_ExtCustomerData.txt", sep="|", quote="", header=FALSE,fill=TRUE,strip=FALSE))
head(clientes)
```

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10						
1	10I000000002	10I000000002	EVENTUAL AJUSTES	0	N/A	N/A	0	MEXICO								
2	26I000000002	26I000000002	EVENTUAL AJUSTES	0	N/A	N/A	0	MEXICO								
3	800I000000002	800I000000002	EVENTUAL AJUSTES	0	N/A	N/A	0	MEXICO								
4	804I000000001	804I000000001	CLIENTE EVENTUAL	EVENTUAL	N/A	N/A	S/CP	MEXICO	0.0000	0.0000						
5	804I000000002	804I000000002	EVENTUAL AJUSTE	EVENTUAL AJUSTE	N/A	N/A	S/CP	MEXICO	0.0000	0.0000						
6	804I000000003	804I000000003	UNICORNIO	LUCIO BLANCO	N/A	N/A	S/CP	MEXICO	22.1499	-98.1772						
	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26
1	2	DTS	21	OTROS	0	N/A	0	N/A				3	NORTE	55N360	DIV COAHUILA	10
2	2	DTS	21	OTROS	0	N/A	0	N/A				3	NORTE	55N830	DIV AS NORTE	26
3	2	DTS	21	OTROS	0	N/A	0	N/A				3	NORTE	55M800	DIV MAYOREO NORTE	800
4	2	DTS	37	EVENTUAL	0	N/A	0	N/A				3	NORTE	55N320	DIV TAMPICO	804
5	2	DTS	37	EVENTUAL	0	N/A	0	N/A				3	NORTE	55N320	DIV TAMPICO	804
6	2	DTS	19	MISCELANEA	0	N/A	0	N/A				3	NORTE	55N320	DIV TAMPICO	804
	V27	V28	V29	V30	V31	V32										
1		SALTILLO	EDI		N/A	N/A	N/A									
2		SALTILLO	C.A. EDI		N/A	N/A	N/A									
3			MAYOREO NORTE		N/A	N/A	N/A									
4			TAMPICO		N/A	N/A	N/A									
5			TAMPICO		N/A	N/A	N/A									
6			TAMPICO		N/A	N/A	N/A									

#v14 es la buena.

```

clientes_sum<-clientes[,c(1,14)]
rm(clientes)
names(clientes_sum)<-c("id_cliente","tipo_cli")

#Habíamos guardado en papas2.dat las papas con los sku's buenos

nombres<-c("id_cliente","codbarr","id_prod","fecha","moneda","medida","vta_din","vta_uni","dev_din",
"dev_uni","id_ruta")
s<-as.data.frame(read.table("churrumais2.dat", sep="|",header=FALSE,
colClasses=c(rep("character",6),rep("numeric",4),"integer")))
names(s)<-nombres

```

Y ahora, veamos de qué tipo son los clientes que tenemos. Primero, filtrar todos los clientes del catálogo con únicamente los que están en la tabla de papitas. Luego ya el merge. Y lo escribimos en un archivo, por si crashea R o lo que sea. Hacer el merge tardó como 5 minutos pero de todos modos.

```

tipo_cliente<-subset(clientes_sum,id_cliente%in%s$id_cliente)
datos<-merge(x = s, y = tipo_cliente, by = "id_cliente", all.x=TRUE)
nombres<-c("id_cliente","codbarr","id_prod","fecha","moneda","medida","vta_din","vta_uni","dev_din",
"dev_uni","id_ruta","tipo_cli")
names(datos)<-nombres
head(datos)

```

	id_cliente	codbarr	id_prod	fecha	moneda	medida	vta_din	vta_uni	dev_din	dev_uni	id_ruta	tipo_cli
1	804I0000000001	N/A	5545	2012-06-08	PESOS	EACH	13.48	4	0	0	17276	EVENTUAL
2	804I0000000001	N/A	5545	2012-06-16	PESOS	EACH	26.96	8	0	0	8136	EVENTUAL
3	804I0000000001	N/A	7390	2012-12-27	PESOS	EACH	12.66	3	0	0	15383	EVENTUAL
4	804I0000000001	N/A	7389	2012-11-19	PESOS	EACH	63.30	15	0	0	8311	EVENTUAL
5	804I0000000001	N/A	5545	2012-04-02	PESOS	EACH	3.37	1	0	0	19777	EVENTUAL
6	804I0000000001	N/A	1679	2013-07-27	PESOS	EACH	21.10	5	0	0	8139	EVENTUAL

¿Y cuántos clientes hay de cada tipo?

```
# table(clientes_sum[,2])
```

Quedamos en tomar las ventas como positivas; tanto en unidades como en dineros y hay que agregar una columna con la semana:

```

ind <- 1:nrow(datos)
ind_vta_din <- ind[datos$vta_din < 0]
datos$dev_din[ind_vta_din] <- datos$dev_din[ind_vta_din] + abs(datos$vta_din[ind_vta_din])
datos$vta_din[ind_vta_din] <- 0
ind_vta_uni <- ind[datos$vta_uni < 0]
datos$dev_uni[ind_vta_uni] <- datos$dev_uni[ind_vta_uni] + abs(datos$vta_uni[ind_vta_uni])
datos$vta_uni[ind_vta_uni] <- 0

datos$vta_din2 <- datos$vta_din
datos$vta_uni2 <- datos$vta_uni
datos$dev_din2 <- abs(datos$dev_din)
datos$dev_uni2 <- abs(datos$dev_uni)

```

Para el análisis agregado que queremos, hay que sumar por semana y por tipo_cli las ventas y devoluciones en unidades y dinero.

```

rm(s)
rm(clientes_sum)

```

```

rm(tipo_cliente)
#como 10 minutos tardó
agregados_fecha<-aggregate(datos[,13:16],list(FactorA=datos$id_cliente,FactorB=datos$fecha),sum)

names(agregados_fecha)<-c("id_cliente","fecha","vta_din2","vta_uni2","dev_din2","dev_uni2")
tipo_cliente<-subset(clientes_sum,id_cliente%in%agregados_fecha$id_cliente)
head(agregados_fecha)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2
1 804I0000000001 2011-12-31   101.10      30         0         0
2 804I0000000135 2011-12-31    16.85       5         0         0
3 804I0000000151 2011-12-31    33.70      10         0         0
4 804I0000000154 2011-12-31    33.70      10         0         0
5 804I0000000393 2011-12-31    20.22       6         0         0
6 804I0000000534 2011-12-31    13.48       4         0         0

agregados_fecha_tipo<-merge(x = agregados_fecha, y = tipo_cliente, by = "id_cliente", all.x=TRUE)
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli
1 804I0000000001 2011-12-31   101.10      30         0         0 EVENTUAL
2 804I0000000001 2013-03-14   219.44      52         0         0 EVENTUAL
3 804I0000000001 2012-05-18   111.21      33         0         0 EVENTUAL
4 804I0000000001 2013-09-18    42.20      10         0         0 EVENTUAL
5 804I0000000001 2013-02-11   295.40      70         0         0 EVENTUAL
6 804I0000000001 2012-06-13   161.76      48         0         0 EVENTUAL

x <- as.POSIXlt(agregados_fecha_tipo$fecha) #tarda sigllooooo... ni pex...
agregados_fecha_tipo$semana<-strftime(x,format="%W")
agregados_fecha_tipo$anio<-strftime(x,format="%Y")
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli semana anio
1 804I0000000001 2011-12-31   101.10      30         0         0 EVENTUAL    52 2011
2 804I0000000001 2013-03-14   219.44      52         0         0 EVENTUAL    10 2013
3 804I0000000001 2012-05-18   111.21      33         0         0 EVENTUAL    20 2012
4 804I0000000001 2013-09-18    42.20      10         0         0 EVENTUAL    37 2013
5 804I0000000001 2013-02-11   295.40      70         0         0 EVENTUAL     6 2013
6 804I0000000001 2012-06-13   161.76      48         0         0 EVENTUAL    24 2012

write.table(agregados_fecha_tipo,"churrumais_sum_fecha_cli.csv",sep=" ",row.names = FALSE)
#names(agregados_fecha_tipo)<-c("id_cliente","fecha","vta_din2","vta_uni2", "dev_din2", "dev_uni2","tipo_cli")

sem_anio<-paste(agregados_fecha_tipo$semana,agregados_fecha_tipo$anio,sep="-")
agregados_semana_tipo<-aggregate(agregados_fecha_tipo[,3:6],list(FactorA=agregados_fecha_tipo$tipo_cli,
                                                                    FactorB=sem_anio),mean)
head(agregados_semana_tipo)
  FactorA FactorB vta_din2 vta_uni2 dev_din2 dev_uni2
1      00-2013 33.72964  7.992806 0.00000000 0.03597122
2      19 00-2013 16.88000  4.000000 0.00000000 0.00000000
3 ABARROTES 00-2013 27.82882  6.599075 0.05561973 0.01336074
4 BALNEARIO 00-2013 27.43000  6.500000 0.00000000 0.00000000
5      BAR 00-2013 31.64400  8.700000 0.00000000 0.00000000
6  BODEGAS 00-2013 26.17081  6.201613 0.00000000 0.00000000

write.table(agregados_semana_tipo,"churrumais3.dat",sep="|",row.names = FALSE,col.names =FALSE)
#names(agregados_semana_tipo)<-c("tipo","semana_anio","vta_din2","vta_uni2","dev_din2","dev_uni2")

```

Fritos

```
setwd("/Users/PandoraMac/Documents/David/pepsi2/Datos Norte 3 papitas/")
clientes<-as.data.frame(read.table("SF_ExtCustomerData.txt", sep="|", quote="", header=FALSE, fill=TRUE, strip.white=TRUE))
head(clientes)
```

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
1	10I0000000002	10I0000000002	EVENTUAL AJUSTES	0	N/A	N/A	0	MEXICO		
2	26I0000000002	26I0000000002	EVENTUAL AJUSTES	0	N/A	N/A	0	MEXICO		
3	800I0000000002	800I0000000002	EVENTUAL AJUSTES	0	N/A	N/A	0	MEXICO		
4	804I0000000001	804I0000000001	CLIENTE EVENTUAL	EVENTUAL	N/A	N/A	S/CP	MEXICO	0.0000	0.0000
5	804I0000000002	804I0000000002	EVENTUAL AJUSTE	EVENTUAL AJUSTE	N/A	N/A	S/CP	MEXICO	0.0000	0.0000
6	804I0000000003	804I0000000003	UNICORNIO	LUCIO BLANCO	N/A	N/A	S/CP	MEXICO	22.1499	-98.1772

	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26
1	2	DTS	21	OTROS	0	N/A	0	N/A				3	NORTE	55N360	DIV COAHUILA	10
2	2	DTS	21	OTROS	0	N/A	0	N/A				3	NORTE	55N830	DIV AS NORTE	26
3	2	DTS	21	OTROS	0	N/A	0	N/A				3	NORTE	55M800	DIV MAYOREO NORTE	800
4	2	DTS	37	EVENTUAL	0	N/A	0	N/A				3	NORTE	55N320	DIV TAMPICO	804
5	2	DTS	37	EVENTUAL	0	N/A	0	N/A				3	NORTE	55N320	DIV TAMPICO	804
6	2	DTS	19	MISCELANEA	0	N/A	0	N/A				3	NORTE	55N320	DIV TAMPICO	804

	V27	V28	V29	V30	V31	V32
1	SALTILLO	EDI		N/A	N/A	N/A
2	SALTILLO	C.A. EDI		N/A	N/A	N/A
3	MAYOREO	NORTE		N/A	N/A	N/A
4	TAMPICO			N/A	N/A	N/A
5	TAMPICO			N/A	N/A	N/A
6	TAMPICO			N/A	N/A	N/A

#v14 es la buena.

```
clientes_sum<-clientes[,c(1,14)]
rm(clientes)
names(clientes_sum)<-c("id_cliente", "tipo_cli")
```

#Habíamos guardado en papas2.dat las papas con los sku's buenos

```
nombres<-c("id_cliente", "codbarr", "id_prod", "fecha", "moneda", "medida", "vta_din", "vta_uni", "dev_din",
"dev_uni", "id_ruta")
s<-as.data.frame(read.table("fritos2.dat", sep="|", header=FALSE,
colClasses=c(rep("character", 6), rep("numeric", 4), "integer")))
names(s)<-nombres
```

Y ahora, veamos de qué tipo son los clientes que tenemos. Primero, filtrar todos los clientes del catálogo con únicamente los que están en la tabla de papitas. Luego ya el merge. Y lo escribimos en un archivo, por si crashea R o lo que sea. Hacer el merge tardó como 5 minutos pero de todos modos.

```
tipo_cliente<-subset(clientes_sum, id_cliente%in%s$id_cliente)
datos<-merge(x = s, y = tipo_cliente, by = "id_cliente", all.x=TRUE)
nombres<-c("id_cliente", "codbarr", "id_prod", "fecha", "moneda", "medida", "vta_din", "vta_uni", "dev_din",
"dev_uni", "id_ruta", "tipo_cli")
names(datos)<-nombres
head(datos)
```

	id_cliente	codbarr	id_prod	fecha	moneda	medida	vta_din	vta_uni	dev_din	dev_uni	id_ruta	tipo_cli
1	804I0000000001	N/A	1963	2013-03-13	PESOS	EACH	50.60	10	0	0	8142	EVENTUAL
2	804I0000000001	N/A	7379	2012-11-05	PESOS	EACH	45.54	9	0	0	7339	EVENTUAL
3	804I0000000001	N/A	1963	2013-08-26	PESOS	EACH	10.12	2	0	0	7337	EVENTUAL
4	804I0000000001	N/A	6191	2012-03-01	PESOS	EACH	20.24	4	0	0	7339	EVENTUAL
5	804I0000000001	N/A	7378	2012-08-23	PESOS	EACH	10.12	2	0	0	7337	EVENTUAL
6	804I0000000001	N/A	7378	2012-09-12	PESOS	EACH	80.96	16	0	0	27302	EVENTUAL

¿Y cuántos clientes hay de cada tipo?

```
# table(clientes_sum[,2])
```

Quedamos en tomar las ventas como positivas; tanto en unidades como en dineros y hay que agregar una columna con la semana:

```
ind <- 1:nrow(datos)
ind_vta_din <- ind[datos$vta_din < 0]
datos$dev_din[ind_vta_din] <- datos$dev_din[ind_vta_din] + abs(datos$vta_din[ind_vta_din])
datos$vta_din[ind_vta_din] <- 0
ind_vta_uni <- ind[datos$vta_uni < 0]
datos$dev_uni[ind_vta_uni] <- datos$dev_uni[ind_vta_uni] + abs(datos$vta_uni[ind_vta_uni])
datos$vta_uni[ind_vta_uni] <- 0

datos$vta_din2 <- datos$vta_din
datos$vta_uni2 <- datos$vta_uni
datos$dev_din2 <- abs(datos$dev_din)
datos$dev_uni2 <- abs(datos$dev_uni)
```

Para el análisis agregado que queremos, hay que sumar por semana y por tipo_cli las ventas y devoluciones en unidades y dinero.

```
rm(s)
rm(clientes_sum)
rm(tipo_cliente)
#como 10 minutos tardó
agregados_fecha<-aggregate(datos[,13:16],list(FactorA=datos$id_cliente,FactorB=datos$fecha),sum)

names(agregados_fecha)<-c("id_cliente","fecha","vta_din2","vta_uni2","dev_din2","dev_uni2")
tipo_cliente<-subset(clientes_sum,id_cliente%in%agregados_fecha$id_cliente)
head(agregados_fecha)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2
1 804I0000000001 2011-12-31   222.64      44         0         0
2 804I0000000091 2011-12-31    20.24       4         0         0
3 804I0000000093 2011-12-31    40.48       8         0         0
4 804I0000000114 2011-12-31    15.18       3         0         0
5 804I0000000151 2011-12-31    50.60      10         0         0
6 804I0000000283 2011-12-31    15.18       3         0         0

agregados_fecha_tipo<-merge(x = agregados_fecha, y = tipo_cliente, by = "id_cliente", all.x=TRUE)
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli
1 804I0000000001 2011-12-31   222.64      44         0         0 EVENTUAL
2 804I0000000001 2012-06-29   166.98      33         0         0 EVENTUAL
3 804I0000000001 2012-10-24   278.30      55         0         0 EVENTUAL
4 804I0000000001 2012-03-13   323.84      64         0         0 EVENTUAL
5 804I0000000001 2012-06-15   242.88      48         0         0 EVENTUAL
6 804I0000000001 2012-08-22   465.52      92         0         0 EVENTUAL

x <- as.POSIXlt(agregados_fecha_tipo$fecha) #tarda sigllooooo.... ni pex...
agregados_fecha_tipo$semana<-strftime(x,format="%W")
agregados_fecha_tipo$anio<-strftime(x,format="%Y")
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli semana anio
1 804I0000000001 2011-12-31   222.64      44         0         0 EVENTUAL    52 2011
```

```

2 804I0000000001 2012-06-29 166.98 33 0 0 EVENTUAL 26 2012
3 804I0000000001 2012-10-24 278.30 55 0 0 EVENTUAL 43 2012
4 804I0000000001 2012-03-13 323.84 64 0 0 EVENTUAL 11 2012
5 804I0000000001 2012-06-15 242.88 48 0 0 EVENTUAL 24 2012
6 804I0000000001 2012-08-22 465.52 92 0 0 EVENTUAL 34 2012

```

```

write.table(agregados_fecha_tipo,"fritos_sum_fecha_cli.csv",sep=",",row.names = FALSE)
#names(agregados_fecha_tipo)<-c("id_cliente","fecha","vta_din2","vta_uni2", "dev_din2", "dev_uni2","tipo_cli")
sem_anio<-paste(agregados_fecha_tipo$semana,agregados_fecha_tipo$anio,sep="-")
agregados_semana_tipo<-aggregate(agregados_fecha_tipo[,3:6],list(FactorA=agregados_fecha_tipo$tipo_cli,
FactorB=sem_anio),mean)

```

```

head(agregados_semana_tipo)
  FactorA FactorB vta_din2 vta_uni2 dev_din2 dev_uni2
1          00-2013 34.27742 6.774194 0.0816129 0.01612903
2          1 00-2013 20.24000 4.000000 0.0000000 0.00000000
3 ABARROTES 00-2013 26.84807 5.305942 0.1464358 0.02900958
4 BALNEARIO 00-2013 32.04667 6.333333 0.0000000 0.00000000
5          BAR 00-2013 20.80222 4.111111 3.3733333 0.66666667
6  BODEGAS 00-2013 28.56658 5.645570 0.0000000 0.00000000

```

```

write.table(agregados_semana_tipo,"fritos3.dat",sep="|",row.names = FALSE,col.names =FALSE)
#names(agregados_semana_tipo)<-c("tipo","semana_anio","vta_din2","vta_uni2","dev_din2","dev_uni2")

```

Papas Sabritas

```

setwd("/Users/PandoraMac/Documents/David/pepsi2/Datos Norte 3 papitas/")
clientes<-as.data.frame(read.table("SF_ExtCustomerData.txt", sep="|", quote="", header=FALSE,fill=TRUE,strip=FALSE))
head(clientes)
  V1          V2          V3          V4 V5 V6 V7 V8 V9 V10
1 10I0000000002 10I0000000002 EVENTUAL AJUSTES          0 N/A N/A 0 MEXICO
2 26I0000000002 26I0000000002 EVENTUAL AJUSTES          0 N/A N/A 0 MEXICO
3 800I0000000002 800I0000000002 EVENTUAL AJUSTES          0 N/A N/A 0 MEXICO
4 804I0000000001 804I0000000001 CLIENTE EVENTUAL          EVENTUAL N/A N/A S/CP MEXICO 0.0000 0.0000
5 804I0000000002 804I0000000002 EVENTUAL AJUSTE EVENTUAL AJUSTE N/A N/A S/CP MEXICO 0.0000 0.0000
6 804I0000000003 804I0000000003 UNICORNIO LUCIO BLANCO N/A N/A S/CP MEXICO 22.1499 -98.1772
  V11 V12 V13          V14 V15 V16 V17 V18 V19 V20 V21 V22 V23 V24 V25 V26
1  2 DTS 21 OTROS 0 N/A 0 N/A          3 NORTE 55N360          DIV COAHUILA 10
2  2 DTS 21 OTROS 0 N/A 0 N/A          3 NORTE 55N830          DIV AS NORTE 26
3  2 DTS 21 OTROS 0 N/A 0 N/A          3 NORTE 55M800 DIV MAYOREO NORTE 800
4  2 DTS 37 EVENTUAL 0 N/A 0 N/A          3 NORTE 55N320          DIV TAMPICO 804
5  2 DTS 37 EVENTUAL 0 N/A 0 N/A          3 NORTE 55N320          DIV TAMPICO 804
6  2 DTS 19 MISCELANEA 0 N/A 0 N/A          3 NORTE 55N320          DIV TAMPICO 804
  V27 V28 V29 V30 V31 V32
1  SALTILLO EDI          N/A N/A N/A
2 SALTILLO C.A. EDI          N/A N/A N/A
3  MAYOREO NORTE          N/A N/A N/A
4  TAMPICO          N/A N/A N/A
5  TAMPICO          N/A N/A N/A
6  TAMPICO          N/A N/A N/A

```

#v14 es la buena.

```

clientes_sum<-clientes[,c(1,14)]
rm(clientes)
names(clientes_sum)<-c("id_cliente","tipo_cli")

```

#Habíamos guardado en papas2.dat las papas con los sku's buenos

```
nombres<-c("id_cliente","codbarr","id_prod","fecha","moneda","medida","vta_din","vta_uni","dev_din",
           "dev_uni","id_ruta")
s<-as.data.frame(read.table("papas2.dat", sep="|",header=FALSE,
                           colClasses=c(rep("character",6),rep("numeric",4),"integer")))
names(s)<-nombres
```

Y ahora, veamos de qué tipo son los clientes que tenemos. Primero, filtrar todos los clientes del catálogo con únicamente los que están en la tabla de papitas. Luego ya el merge. Y lo escribimos en un archivo, por si crashea R o lo que sea. Hacer el merge tardó como 5 minutos pero de todos modos.

```
tipo_cliente<-subset(clientes_sum,id_cliente%in%s$id_cliente)
datos<-merge(x = s, y = tipo_cliente, by = "id_cliente", all.x=TRUE)
nombres<-c("id_cliente","codbarr","id_prod","fecha","moneda","medida","vta_din","vta_uni","dev_din",
           "dev_uni","id_ruta","tipo_cli")
names(datos)<-nombres
head(datos)
```

	id_cliente	codbarr	id_prod	fecha	moneda	medida	vta_din	vta_uni	dev_din	dev_uni	id_ruta	tipo_cli
1	804I000000001	N/A	7398	2013-01-15	PESOS	EACH	70.80	12	0	0	398	EVENTUAL
2	804I000000001	N/A	517800	2013-08-19	PESOS	EACH	60.66	9	0	0	1272	EVENTUAL
3	804I000000001	N/A	7398	2012-11-13	PESOS	EACH	141.60	24	0	0	17276	EVENTUAL
4	804I000000001	N/A	7398	2013-04-22	PESOS	EACH	35.40	6	0	0	8140	EVENTUAL
5	804I000000001	N/A	1446	2012-05-03	PESOS	EACH	94.40	16	0	0	17217	EVENTUAL
6	804I000000001	N/A	1892	2012-04-02	PESOS	EACH	259.60	44	0	0	1272	EVENTUAL

¿Y cuántos clientes hay de cada tipo?

```
# table(clientes_sum[,2])
```

Quedamos en tomar las ventas como positivas; tanto en unidades como en dineros y hay que agregar una solumna con la semana:

```
ind <- 1:nrow(datos)
ind_vta_din <- ind[datos$vta_din < 0]
datos$dev_din[ind_vta_din] <- datos$dev_din[ind_vta_din] + abs(datos$vta_din[ind_vta_din])
datos$vta_din[ind_vta_din] <- 0
ind_vta_uni <- ind[datos$vta_uni < 0]
datos$dev_uni[ind_vta_uni] <- datos$dev_uni[ind_vta_uni] + abs(datos$vta_uni[ind_vta_uni])
datos$vta_uni[ind_vta_uni] <- 0

datos$vta_din2 <- datos$vta_din
datos$vta_uni2 <- datos$vta_uni
datos$dev_din2 <- abs(datos$dev_din)
datos$dev_uni2 <- abs(datos$dev_uni)
```

Para el análisis agregado que queremos, hay que sumar por semana y por tipo_cli las ventas y devoluciones en unidades y dinero.

```
rm(s)
rm(clientes_sum)
rm(tipo_cliente)
#como 10 minutos tardó
agregados_fecha<-aggregate(datos[,13:16],list(FactorA=datos$id_cliente,FactorB=datos$fecha),sum)
names(agregados_fecha)<-c("id_cliente","fecha","vta_din2","vta_uni2","dev_din2","dev_uni2")
tipo_cliente<-subset(clientes_sum,id_cliente%in%agregados_fecha$id_cliente)
```



```

head(agregados_fecha)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2
1 804I0000000001 2011-12-31  2377.7    403      0      0
2 804I0000000007 2011-12-31   29.5      5      0      0
3 804I0000000034 2011-12-31   47.2      8      0      0
4 804I0000000074 2011-12-31   94.4     16      0      0
5 804I0000000088 2011-12-31   29.5      5      0      0
6 804I0000000093 2011-12-31   94.4     16      0      0

agregados_fecha_tipo<-merge(x = agregados_fecha, y = tipo_cliente, by = "id_cliente", all.x=TRUE)
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli
1 804I0000000001 2011-12-31  2377.7    403      0      0 EVENTUAL
2 804I0000000001 2012-10-23   590.0    100      0      0 EVENTUAL
3 804I0000000001 2013-01-29  1563.5    265      0      0 EVENTUAL
4 804I0000000001 2013-04-13  1073.8    182      0      0 EVENTUAL
5 804I0000000001 2013-03-27   29.5      5      0      0 EVENTUAL
6 804I0000000001 2012-02-11  1687.4    286      0      0 EVENTUAL

x <- as.POSIXlt(agregados_fecha_tipo$fecha) #tarda siglooooo... ni pex...
agregados_fecha_tipo$semana<-strftime(x,format="%W")
agregados_fecha_tipo$anio<-strftime(x,format="%Y")
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli semana anio
1 804I0000000001 2011-12-31  2377.7    403      0      0 EVENTUAL    52 2011
2 804I0000000001 2012-10-23   590.0    100      0      0 EVENTUAL    43 2012
3 804I0000000001 2013-01-29  1563.5    265      0      0 EVENTUAL     4 2013
4 804I0000000001 2013-04-13  1073.8    182      0      0 EVENTUAL    14 2013
5 804I0000000001 2013-03-27   29.5      5      0      0 EVENTUAL    12 2013
6 804I0000000001 2012-02-11  1687.4    286      0      0 EVENTUAL     6 2012

write.table(agregados_fecha_tipo,"papas_sum_fecha_cli.csv",sep="," ,row.names = FALSE)
#names(agregados_fecha_tipo)<-c("id_cliente","fecha","vta_din2","vta_uni2", "dev_din2", "dev_uni2","tipo_cli")
sem_anio<-paste(agregados_fecha_tipo$semana,agregados_fecha_tipo$anio,sep="-")
agregados_semana_tipo<-aggregate(agregados_fecha_tipo[,3:6],list(FactorA=agregados_fecha_tipo$tipo_cli,
                                                                FactorB=sem_anio),mean)

head(agregados_semana_tipo)
  FactorA FactorB vta_din2 vta_uni2 dev_din2 dev_uni2
1      00-2013 78.18643 13.251938 0.00000000 0.00000000
2      21 00-2013  5.90000  1.000000 0.00000000 0.00000000
3 ABARROTES 00-2013 46.18446  7.827875 0.05066098 0.008579291
4 BALNEARIO 00-2013 38.35000  6.500000 0.00000000 0.00000000
5      BAR 00-2013 57.52500  9.750000 0.00000000 0.00000000
6  BODEGAS 00-2013 45.12159  7.647727 0.00000000 0.00000000

write.table(agregados_semana_tipo,"papas3.dat",sep="|",row.names = FALSE,col.names =FALSE)
#names(agregados_semana_tipo)<-c("tipo","semana_anio","vta_din2","vta_uni2","dev_din2","dev_uni2")

```

Top de ventas y devoluciones en baro:

```

agregados_fecha_tipo <- read.table("churrumais_sum_fecha_cli.csv", header = TRUE,
  sep = ",")
cosa <- aggregate(agregados_fecha_tipo[, 3], list(FactorA = agregados_fecha_tipo$tipo_cli),
  sum)
write.table(cosa, "churrus.csv", sep = ",", row.names = FALSE)

agregados_fecha_tipo <- read.table("fritos_sum_fecha_cli.csv", header = TRUE,

```



```

    sep = ",")
cosa <- aggregate(agregados_fecha_tipo[, 3], list(FactorA = agregados_fecha_tipo$tipo_cli),
    sum)
write.table(cosa, "frits.csv", sep = ",", row.names = FALSE)

agregados_fecha_tipo <- read.table("papas_sum_fecha_cli.csv", header = TRUE,
    sep = ",")
cosa <- aggregate(agregados_fecha_tipo[, 3], list(FactorA = agregados_fecha_tipo$tipo_cli),
    sum)
write.table(cosa, "paps.csv", sep = ",", row.names = FALSE)

#
agregados_fecha_tipo <- read.table("churruvais_sum_fecha_cli.csv", header = TRUE,
    sep = ",")
cosa <- aggregate(agregados_fecha_tipo[, 5], list(FactorA = agregados_fecha_tipo$tipo_cli),
    sum)
write.table(cosa, "churru2.csv", sep = ",", row.names = FALSE)

agregados_fecha_tipo <- read.table("fritos_sum_fecha_cli.csv", header = TRUE,
    sep = ",")
cosa <- aggregate(agregados_fecha_tipo[, 5], list(FactorA = agregados_fecha_tipo$tipo_cli),
    sum)
write.table(cosa, "frits2.csv", sep = ",", row.names = FALSE)

agregados_fecha_tipo <- read.table("papas_sum_fecha_cli.csv", header = TRUE,
    sep = ",")
cosa <- aggregate(agregados_fecha_tipo[, 5], list(FactorA = agregados_fecha_tipo$tipo_cli),
    sum)
write.table(cosa, "paps2.csv", sep = ",", row.names = FALSE)

```

Y otro para abrrotes

```

cosa<-read.table("churruvais_sum_fecha_cli.csv",header=TRUE,sep=",")
datos<-subset(cosa,tipo_cli=="ABARROTES")

> length(table(datos$id_cliente))
[1] 124,764
>
> summary(datos$vta_din)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.00  16.85   20.22   24.36  26.96 37970.00
>
> summary(datos$vta_uni)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.000  4.000   5.000   6.452  8.000 11270.000
>
> summary(datos$dev_din)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000 0.0000  0.0000  0.0364  0.0000 240.0000
>
> summary(datos$dev_uni)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.00000 0.00955 0.00000 60.00000

```

```
cosa<-read.table("fritos_sum_fecha_cli.csv",header=TRUE,sep=",")
datos<-subset(cosa,tipos_cli=="ABARROTES")
```

```
> length(table(datos$id_cliente))
[1] 120,153
>
> summary(datos$vtas_din)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00  15.18   25.30   26.10   25.30 13410.00
>
> summary(datos$vtas_uni)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000   3.000   5.000   5.163   5.000 2650.000
>
> summary(datos$dev_din)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.00000 0.09362 0.00000 190.00000
>
> summary(datos$dev_uni)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.00000 0.01856 0.00000 38.00000
```

```
cosa<-read.table("papas_sum_fecha_cli.csv",header=TRUE,sep=",")
datos<-subset(cosa,tipos_cli=="ABARROTES")
```

```
> length(table(datos$id_cliente))
[1] 133,223
>
> summary(datos$vtas_din)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00  23.60   29.50   43.57   47.20 49560.00
>
> summary(datos$vtas_uni)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000   4.000   5.000   7.151   8.000 8400.000
>
> summary(datos$dev_din)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.00000 0.0139 0.00000 350.00000
>
> summary(datos$dev_uni)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.00000 0.00224 0.00000 56.00000
```

```
length(table(datos$id_cliente))
```

```
summary(datos$vtas_din)
```

```
summary(datos$vtas_uni)
```

```
summary(datos$dev_din)
```

```
summary(datos$dev_uni)
```