

17-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 churrumais3, 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), estamos tomando valor absoluto de ventas y devoluciones porque hay valores negativos. Suponemos que tiene que ver con promociones en las que la gente llega al OXXO y canjea un cupón de “papitas gratis”. No nos importa distinguir entre papas consumidas y pagadas y papas consumidas y gratis porque al final es demanda de papas. Quizá tenga más sentido omitir los valores negativos, pero a mí me late que no.

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

Sabritas normales

- 1: CUSTKEY_CUSTNUMBER - 830I000000002
- 2: COD_BARRAS, - N/A todos
- 3: PRO.ID_PROD_ODS, - 0436
- 4: CALENDARDATE, - 2012-03-16
- 5: CURRENCY, - PESOS
- 6: MEDIDA, - EACH
- 7: VENTA_CAPT_PES, -> SALESAMOUNT - 1720.0000
- 8: VENTA_CAPT_UNI, -> SALESQUANTITY - 50.0000
- 9: DEVOL_PES, -> UNSALEABLEAMOUNT - 0.0000
- 10: DEVOL_UNI -> UNSALEABLEQUANTITY - 106
- 11: RUT.ID_RUTA

```
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("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)
write.table(datos,"papas3.dat",sep="|",row.names = FALSE,col.names =FALSE)
#datos<-as.data.frame(read.csv("papas3.dat",sep="|"))
#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])
```

1	10
3376	2
11	2
1	3
20	23
5	1
26	34
3	8

4	ABARROTES	AUTO LATA
1	191669	2
BALNEARIO	BAR	BODEGAS
260	472	16617
CASETA TELEFONICA	CENTRO DEPORTIV	CENTRO RECREATI
224	261	542
CERVECENTRO	CERVEFRIO	CIBERCAFE
20	8	1979
CINE / TEATRO	CLUB DE PRECIOS	CORONA
349	41	21
CREMERIA	DEP. DE CERVEZA ULTRAMARINOS	DEPOSITO DE CERVEZA
947	2	12511
DTS	ESCUELA	ESTADIO
2	5212	3416
EVENTUAL	FABRICA	FARMACIA
113	132	4905
FRUTERIA Y VERD	GASOLINERIA	HIPERMERCADO
2562	178	788
HOTEL	KINDER DE GOBIERNO	KINDER PARTICULAR
335	110	68
LA CANTERA	LICORERIA	LONCHERIA / FOND
543	1264	513
MAQUILADORAS	MINISUPER	MISCELANEA
695	25145	70365
MODELORAMA	NO VALIDO	OASIS
1216	14	6
OM	OTROS	PALETERIA
12	43226	1794
PAPELERIA	PREP A UNIV GOBIERNO	PREP A UNIV PARTICULAR
8800	82	88
PREP Y UNIV GOBIERNO	PREP Y UNIV PARTICULAR	PREPARATORIA DE GOBIERNO
37	24	969
PREPARATORIA PARTICULAR	PRIM A PREP GOBIERNO	PRIM A PREP PARTICULAR
403	24	22
PRIM A SEC GOBIERNO	PRIM A SEC PARTICULAR	PRIMARIA A SEC PARTICULAR
1003	146	111
PRIMARIA A SECUNDARIA GOBIERNO	PRIMARIA DE GOBIERNO	PRIMARIA PARTICULAR
81	3250	345
PUESTO	PUESTO DE REVIS	PUESTO FIJO
15684	347	76
PUNTO MODELO	REFRESQUERIA	RESTAURANT
11	1036	2836
ROSTICERIA	SEC Y PREP GOBIENO	SEC Y PREP PARTICULAR
34506	36	14
SECUNDARIA A PREP GOBIERNO	SECUNDARIA A PREP PARTICULAR	SECUNDARIA DE GOBIERNO
68	75	1252
SECUNDARIA PARTICULAR	SUPER C	SUPER MERCADO
681	6	598
TECATE SIX	TIENDA DE CONVE	TIENDA DE GOBIE
2227	4528	1515
TIENDA DEPARTAMENTAL	TODOS LOS TIPOS DE CLIENTE	TORTILLERIAS
89	8	1727
TOSTICENTRO	UNIVERSIDAD DE GOBIERNO	UNIVERSIDAD PARTICULAR
4549	413	232
VARIOS DE GOBIERNO	VARIOS GOBIERNO	VARIOS PARTICULAR
61	42	192
VENTANA CON VENTA	VIDEOCLUB	

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

```
datos$vta_din2 <- abs(datos$vta_din)
datos$vta_uni2 <- abs(datos$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)
tipo_cliente<-subset(clientes_sum,id_cliente%in%agregados_fecha$id_cliente)
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 804I000000001 2011-12-31  2377.7    403         0         0 EVENTUAL
2 804I000000001 2012-10-23   590.0    100         0         0 EVENTUAL
3 804I000000001 2013-01-29  1563.5    265         0         0 EVENTUAL
4 804I000000001 2013-04-13  1073.8    182         0         0 EVENTUAL
5 804I000000001 2013-03-27   29.5      5         0         0 EVENTUAL
6 804I000000001 2012-02-11  1687.4    286         0         0 EVENTUAL

x <- as.POSIXlt(agregados_fecha_tipo$fecha) #tarda sigllooooo.... ni pex...
agregados_fecha_tipo$semana<-strftime(x,format="%W")
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli semana
1 804I000000001 2011-12-31  2377.7    403         0         0 EVENTUAL    52
2 804I000000001 2012-10-23   590.0    100         0         0 EVENTUAL    43
3 804I000000001 2013-01-29  1563.5    265         0         0 EVENTUAL    04
4 804I000000001 2013-04-13  1073.8    182         0         0 EVENTUAL    14
5 804I000000001 2013-03-27   29.5      5         0         0 EVENTUAL    12
6 804I000000001 2012-02-11  1687.4    286         0         0 EVENTUAL    06
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")

agregados_semana_tipo<-aggregate(agregados_fecha_tipo[,3:6],list(FactorA=agregados_fecha_tipo$tipo_cli,
                                                                    FactorB=agregados_fecha_tipo$semana),mean)
head(agregados_semana_tipo)
  FactorA FactorB vta_din2 vta_uni2 dev_din2 dev_uni2
1      00  78.18643 13.251938 0.00000000 0.00000000
2      21      00  5.90000  1.000000 0.00000000 0.00000000
3 ABARROTES      00 46.22394  7.834565 0.01118545 0.001888523
4 BALNEARIO      00 38.35000  6.500000 0.00000000 0.00000000
5      BAR      00 57.52500  9.750000 0.00000000 0.00000000
6  BODEGAS      00 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","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 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("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 804I000000001 N/A 7379 2012-11-05 PESOS EACH 45.54 9 0 0 7339 EVENTUAL
2 804I000000001 N/A 1963 2013-08-26 PESOS EACH 10.12 2 0 0 7337 EVENTUAL
3 804I000000001 N/A 6191 2012-03-01 PESOS EACH 20.24 4 0 0 7339 EVENTUAL
4 804I000000001 N/A 7378 2012-08-23 PESOS EACH 10.12 2 0 0 7337 EVENTUAL
5 804I000000001 N/A 7378 2012-09-12 PESOS EACH 80.96 16 0 0 27302 EVENTUAL
6 804I000000001 N/A 7379 2012-10-13 PESOS EACH 25.30 5 0 0 8143 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:

```
datos$vta_din2 <- abs(datos$vta_din)
datos$vta_uni2 <- abs(datos$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)
tipo_cliente<-subset(clientes_sum,id_cliente%in%agregados_fecha$id_cliente)
head(agregados_fecha)
  FactorA      FactorB 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
names(agregados_fecha)<-c("id_cliente","fecha","vta_din2","vta_uni2","dev_din2","dev_uni2")
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$fecha) #tarda siglooooooooos.... ni pex...
agregados_fecha_tipo$semana<-strftime(x,format="%W")
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli semana
1 804I0000000001 2011-12-31   222.64      44         0         0 EVENTUAL    52
2 804I0000000001 2012-06-29   166.98      33         0         0 EVENTUAL    26
3 804I0000000001 2012-10-24   278.30      55         0         0 EVENTUAL    43
4 804I0000000001 2012-03-13   323.84      64         0         0 EVENTUAL    11
5 804I0000000001 2012-06-15   242.88      48         0         0 EVENTUAL    24
6 804I0000000001 2012-08-22   465.52      92         0         0 EVENTUAL    34

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")

agregados_semana_tipo<-aggregate(agregados_fecha_tipo[,3:6],list(FactorA=agregados_fecha_tipo$tipo_cli,
  FactorB=agregados_fecha_tipo$semana),mean)
```

```
head(agregados_semana_tipo)
  FactorA FactorB vta_din2 vta_uni2 dev_din2 dev_uni2
1          00 34.27742 6.774194 0.0816129 0.01612903
2          1  20.24000 4.000000 0.0000000 0.00000000
3 ABARROTES  00 26.87393 5.311054 0.1205700 0.02389776
4 BALNEARIO  00 32.04667 6.333333 0.0000000 0.00000000
5          BAR  00 20.80222 4.111111 3.3733333 0.66666667
6  BODEGAS    00 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","vta_din2","vta_uni2","dev_din2","dev_uni2")
```

Churrumais

```
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:

```

datos$vta_din2 <- abs(datos$vta_din)
datos$vta_uni2 <- abs(datos$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")
head(agregados_fecha_tipo)
  id_cliente      fecha vta_din2 vta_uni2 dev_din2 dev_uni2 tipo_cli semana
1 804I0000000001 2011-12-31   101.10     30        0        0 EVENTUAL    52
2 804I0000000001 2013-03-14   219.44     52        0        0 EVENTUAL    10
3 804I0000000001 2012-05-18   111.21     33        0        0 EVENTUAL    20
4 804I0000000001 2013-09-18    42.20     10        0        0 EVENTUAL    37
5 804I0000000001 2013-02-11   295.40     70        0        0 EVENTUAL    06
6 804I0000000001 2012-06-13   161.76     48        0        0 EVENTUAL    24

write.table(agregados_fecha_tipo,"churruvais_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")

agregados_semana_tipo<-aggregate(agregados_fecha_tipo[,3:6],list(FactorA=agregados_fecha_tipo$tipo_cli,
                                                                    FactorB=agregados_fecha_tipo$semana),mean)
head(agregados_semana_tipo)
  FactorA FactorB vta_din2 vta_uni2 dev_din2 dev_uni2
1              00 33.72964  7.992806 0.00000000 0.03597122
2             19    0 16.88000  4.000000 0.00000000 0.00000000
3 ABARROTES    00 27.83766  6.601336 0.04677287 0.01109969
4 BALNEARIO    00 27.43000  6.500000 0.00000000 0.00000000
5          BAR    00 31.64400  8.700000 0.00000000 0.00000000
6   BODEGAS    00 26.17081  6.201613 0.00000000 0.00000000

write.table(agregados_semana_tipo,"churruvais3.dat",sep="|",row.names = FALSE,col.names =FALSE)
#names(agregados_semana_tipo)<-c("tipo","semana","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, stringsAsFactors = FALSE, colClasses = rep("character",
  32)))

# v14 es la buena.

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

# Hab<U+00ED>amos guardado en fritos2.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)
```

```
write.table(datos,"fritos3.dat",sep="|",row.names = FALSE,col.names =FALSE)
```

	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])
```

	1	10
3376	19	2
11	19	2
1	3	3
20	21	23
5	6	1
26	33	34
3	2	8
4	ABARROTES	AUTO LATA
1	191669	2
BALNEARIO	BAR	BODEGAS
260	472	16617
CASETA TELEFONICA	CENTRO DEPORTIV	CENTRO RECREATI
224	261	542
CERVECENTRO	CERVEFRIO	CIBERCAFE
20	8	1979
CINE / TEATRO	CLUB DE PRECIOS	CORONA
349	41	21
CREMERIA	DEP. DE CERVEZA ULTRAMARINOS	DEPOSITO DE CERVEZA
947	2	12511
DTS	ESCUELA	ESTADIO
2	5212	3416
EVENTUAL	FABRICA	FARMACIA
113	132	4905
FRUTERIA Y VERD	GASOLINERIA	HIPERMERCADO
2562	178	788
HOTEL	KINDER DE GOBIERNO	KINDER PARTICULAR
335	110	68
LA CANTERA	LICORERIA	LONCHERIA / FOND
543	1264	513
MAQUILADORAS	MINISUPER	MISCELANEA
695	25145	70365
MODELORAMA	NO VALIDO	OASIS
1216	14	6
OM	OTROS	PALETERIA
12	43226	1794
PAPELERIA	PREP A UNIV GOBIERNO	PREP A UNIV PARTICULAR
8800	82	88
PREP Y UNIV GOBIERNO	PREP Y UNIV PARTICULAR	PREPARATORIA DE GOBIERNO
37	24	969
PREPARATORIA PARTICULAR	PRIM A PREP GOBIERNO	PRIM A PREP PARTICULAR
403	24	22
PRIM A SEC GOBIERNO	PRIM A SEC PARTICULAR	PRIMARIA A SEC PARTICULAR
1003	146	111
PRIMARIA A SECUNDARIA GOBIERNO	PRIMARIA DE GOBIERNO	PRIMARIA PARTICULAR

81	3250	345
PUESTO	PUESTO DE REVIS	PUESTO FIJO
15684	347	76
PUNTO MODELO	REFRESQUERIA	RESTAURANT
11	1036	2836
ROSTICERIA	SEC Y PREP GOBIENO	SEC Y PREP PARTICULAR
34506	36	14
SECUNDARIA A PREP GOBIERNO	SECUNDARIA A PREP PARTICULAR	SECUNDARIA DE GOBIERNO
68	75	1252
SECUNDARIA PARTICULAR	SUPER C	SUPER MERCADO
681	6	598
TECATE SIX	TIENDA DE CONVE	TIENDA DE GOBIE
2227	4528	1515
TIENDA DEPARTAMENTAL	TODOS LOS TIPOS DE CLIENTE	TORTILLERIAS
89	8	1727
TOSTICENTRO	UNIVERSIDAD DE GOBIERNO	UNIVERSIDAD PARTICULAR
4549	413	232
VARIOS DE GOBIERNO	VARIOS GOBIERNO	VARIOS PARTICULAR
61	42	192
VENTANA CON VENTA	VIDEOCLUB	
3159	757	

Quedamos en tomar las ventas como positivas; tanto en unidades como en dineros:

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

Para el análisis agregado que queremos, hay que promediar por fecha y por tipo_cli ventas y devoluciones en unidades y dinero.

```
rm(s)
rm(clientes_sum)
rm(tipo_cliente)
#como 10 minutos tardó
agregados_tipo<-aggregate(datos[,13:16],list(FactorA=datos$tipo_cli,FactorB=datos$fecha),mean)
head(agregados_tipo)
  FactorA      FactorB vta_din2 vta_uni2  dev_din2  dev_uni2
1      2011-12-31 26.10960 5.160000 0.00000000 0.00000000
2      1 2011-12-31 20.24000 4.000000 0.00000000 0.00000000
3  ABARROTES 2011-12-31 27.74547 5.517098 0.04185492 0.008290155
4   BODEGAS 2011-12-31 27.32400 5.400000 0.00000000 0.00000000
5 CASETA TELEFONICA 2011-12-31 15.18000 3.000000 0.00000000 0.00000000
6  CENTRO DEPORTIV 2011-12-31 20.24000 4.000000 0.00000000 0.00000000
#para clientes seguro tardará más...
#agregados_cliente<-aggregate(datos[,13:16],list(FactorA=datos$id_cliente,FactorB=datos$fecha),mean)
#head(agregados_cliente)

write.table(agregados_tipo,"fritos_agregados_tipo.csv",sep="," ,row.names = FALSE)
#write.csv(agregados_cli,"fritos_agregados_cli.csv",sep="," ,row.names = FALSE)
```

Churrumais

```
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, stringsAsFactors = FALSE, colClasses = rep("character",
32)))

# v14 es la buena.

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

# Hemos guardado en fritos2.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("churruvais2.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)
write.table(datos,"churruvais3.dat",sep="|",row.names = FALSE,col.names =FALSE)

```

```

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])

```

1	10
3376	2
11	2
1	3
20	23
5	1
26	34
3	8
4	AUTO LATA
1	2
BALNEARIO	BAR
260	16617
CASETA TELEFONICA	CENTRO DEPORTIV
224	542
CERVECENRO	CIBERCAFE
20	1979
CINE / TEATRO	CLUB DE PRECIOS
349	21
CREMERIA	DEPOSITO DE CERVEZA
947	12511

DTS	ESCUELA	ESTADIO
2	5212	3416
EVENTUAL	FABRICA	FARMACIA
113	132	4905
FRUTERIA Y VERD	GASOLINERIA	HIPERMERCADO
2562	178	788
HOTEL	KINDER DE GOBIERNO	KINDER PARTICULAR
335	110	68
LA CANTERA	LICORERIA	LONCHERIA / FOND
543	1264	513
MAQUILADORAS	MINISUPER	MISCELANEA
695	25145	70365
MODELORAMA	NO VALIDO	OASIS
1216	14	6
OM	OTROS	PALETERIA
12	43226	1794
PAPELERIA	PREP A UNIV GOBIERNO	PREP A UNIV PARTICULAR
8800	82	88
PREP Y UNIV GOBIERNO	PREP Y UNIV PARTICULAR	PREPARATORIA DE GOBIERNO
37	24	969
PREPARATORIA PARTICULAR	PRIM A PREP GOBIERNO	PRIM A PREP PARTICULAR
403	24	22
PRIM A SEC GOBIERNO	PRIM A SEC PARTICULAR	PRIMARIA A SEC PARTICULAR
1003	146	111
PRIMARIA A SECUNDARIA GOBIERNO	PRIMARIA DE GOBIERNO	PRIMARIA PARTICULAR
81	3250	345
PUESTO	PUESTO DE REVIS	PUESTO FIJO
15684	347	76
PUNTO MODELO	REFRESQUERIA	RESTAURANT
11	1036	2836
ROSTICERIA	SEC Y PREP GOBIENO	SEC Y PREP PARTICULAR
34506	36	14
SECUNDARIA A PREP GOBIERNO	SECUNDARIA A PREP PARTICULAR	SECUNDARIA DE GOBIERNO
68	75	1252
SECUNDARIA PARTICULAR	SUPER C	SUPER MERCADO
681	6	598
TECATE SIX	TIENDA DE CONVE	TIENDA DE GOBIE
2227	4528	1515
TIENDA DEPARTAMENTAL	TODOS LOS TIPOS DE CLIENTE	TORTILLERIAS
89	8	1727
TOSTICENTRO	UNIVERSIDAD DE GOBIERNO	UNIVERSIDAD PARTICULAR
4549	413	232
VARIOS DE GOBIERNO	VARIOS GOBIERNO	VARIOS PARTICULAR
61	42	192
VENTANA CON VENTA	VIDEOCLUB	
3159	757	

Quedamos en tomar las ventas como positivas; tanto en unidades como en dineros:

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

Para el análisis agregado que queremos, hay que promediar por fecha y por tipo_cli ventas y devoluciones en unidades y dinero.

```

rm(s)
rm(clientes_sum)
rm(tipo_cliente)
#como 10 minutos tardó
agregados_tipo<-aggregate(datos[,13:16],list(FactorA=datos$tipo_cli,FactorB=datos$fecha),mean)
head(agregados_tipo)

```

	FactorA	FactorB	vta_din2	vta_uni2	dev_din2	dev_uni2
1		2011-12-31	26.22281	7.781250	0.00000000	0.00000000
2	1	2011-12-31	13.48000	4.000000	0.00000000	0.00000000
3	ABARROTES	2011-12-31	22.48717	6.675319	0.04026782	0.01194891
4	BAR	2011-12-31	25.27500	7.500000	0.00000000	0.00000000
5	BODEGAS	2011-12-31	23.25300	6.900000	0.00000000	0.00000000
6	CASETA TELEFONICA	2011-12-31	6.74000	2.000000	0.00000000	0.00000000

```

#para clientes seguro tardará más...
agregados_cliente<-aggregate(datos[,13:16],list(FactorA=datos$id_cliente,FactorB=datos$fecha),mean)
head(agregados_cliente)

write.table(agregados_tipo,"churrumais_agregados_tipo.csv",sep="," ,row.names = FALSE)
write.table(agregados_cli,"churrumais_agregados_cli.csv",sep="," ,row.names = FALSE)

```